

Magnus Grønset

# Wisdom of crowds i en digital byggebransje

Masteroppgave i Bygg- og miljøteknikk

Veileder: Ole Jonny Klakegg

Juni 2023



Magnus Grønset

# Wisdom of crowds i en digital byggebransje

Masteroppgave i Bygg- og miljøteknikk  
Veileder: Ole Jonny Klakegg  
Juni 2023

Norges teknisk-naturvitenskapelige universitet  
Fakultet for ingeniørvitenskap  
Institutt for bygg- og miljøteknikk



Kunnskap for en bedre verden





# Sammendrag

Gode estimat er noe byggebransjen er avhengig av å gi, ettersom estimatene ofte er styrende for byggeprosjekt. For å gi estimat benyttes det i dag ulike fremsynsmetoder. Wisdom of crowds har vist seg å være en metode som gir svar med oppsiktsvekkende stor grad av presisjon for en rekke problem. Likevel er det lite litteratur om hvordan fenomenet anvendes i dagens byggebransje. Masteroppgavens hensikt er å besvare problemstillingen: «Hvordan kan wisdom of crowds brukes til å minimere kostnadsrisiko?».

Problemstillingen tar for seg temaene wisdom of crowds og kostnadsrisiko i byggebransjen. For å skaffe en oversikt over eksisterende litteratur om temaene har det blitt gjennomført en litteraturstudie. Et eksperiment har blitt utført for å skaffe empirisk data. Eksperimentet gikk ut på at én supereksperter og 33 ingeniørstudenter skulle gi kostnadsestimat for en totalentreprise for et kombinert kontor- og undervisningsbygg. De 33 studentene var fordelt på fire mindre flokker. Kontor- og undervisningsbygget er et ekte prosjekt, og inngår som en del av Norsk havteknologisenter. Estimaterne til supereksperter og studentene ble hentet inn ved hjelp av en Mentimeter-presentasjon. Presentasjonen inneholdt kun utvalgte deler av tilbudsunderlaget for totalentreprisen. Supereksperter og de 33 studentene ga ett estimat hver i to runder, i sum to estimat hver. Mellom den første og den andre runden fikk supereksperter og studentene opplysninger om pris for tidligere prosjekt.

I resultatdelen ble det for begge rundene regnet ut et aritmetisk gjennomsnitt og medianen av estimatene studentene ga. Dette ble gjort for hver av de fire flokkene av studenter og for en flokk bestående av estimatene til alle studentene.

Resultatene til supereksperter og de 33 studentene har blitt analysert. Analysen har sett på kostnadsestimatene til supereksperter og de 33 studentene, samt estimat for totalentreprisen gitt av byggherren (Statsbygg) og tilbudspriser fra tre ulike entreprenører. Ettersom totalentreprisen ikke er ferdig, vil det være umulig å si hva den faktiske prisen for byggeprosjektet blir. Derfor er entreprenørenes priser det nærmeste den aktuelle totalentreprisen kommer en pris. Dermed ble entreprenørenes priser valgt som sammenligningsgrunnlag for estimatet til Statsbygg, estimatene til supereksperter og estimatene til studentene. Medianen for flokken bestående av alle studentene var i den andre runden nær entreprenørenes priser. Statsbyggs estimat var også nær entreprenørenes priser. Supereksperter har gitt det minst presise estimatet. Likevel kan det ikke utelukkes at supereksperter har gitt pris for hele prosjektet, og ikke bare for totalentreprisen.

Presise estimat ble også observert for to av de fire mindre flokkene av studenter. Imidlertid har det blitt konkludert med at usikkerheten forbundet med gjennomsnittet av estimatene er høyere for mindre flokkstørrelser enn for større flokkstørrelser.

Supereksperter og studentene fikk i eksperimentet kun overordnet informasjon om totalentreprisen. I den første runden hadde supereksperter og studentene en top-down-tilnærming. I den andre runden ble top-down-tilnærmingen kombinert med et utsideperspektiv, ettersom priser for andre prosjekt ble gjort kjent. Det ble konkludert med at kombinasjonen av en top-down-tilnærming og utsideperspektivet har vært en medvirkende årsak for at medianen til studentene har vært nær entreprenørenes priser. På bakgrunn av dette anbefales det å bruke wisdom of crowds for estimeringsoppgaver hvor det foreligger lite tilgjengelig informasjon.

Diskusjonen har pekt på at ved å bruke wisdom of crowds har bias som ofte inntreffer i estimeringsprosesser i byggebransjen blitt eliminert. Det ble konkludert med at

flokkmentalitet, underestimering, overoptimisme og the planning fallacy er bias som motvirkes ved å bruke wisdom of crowds. Dog har diskusjonen av resultatene avslørt at det er sannsynlig at supereksperter og studentene i dette eksperimentet har blitt påvirket av forankring. Forankringen knyttes til at det er sannsynlig at deltakerne har sett medieoppslag for kostnadene til hele Norsk havteknologisenter.

Det ble i analysen konstatert at usikkerheten forbundet med gjennomsnittet av estimatene var høyere for mindre flokkstørrelser enn for større flokkstørrelser. Derfor vil et tiltak for å redusere risiko være å inkludere flere personer i flokken når wisdom of crowds brukes. Flokkstørrelsen bør justeres etter risikoen forbundet med hva som skal estimeres. Det ble konkludert med at digitale verktøy har gjort det lettere enn noen gang å inkludere personer i flokken. Basert på drøftingen er det kommet frem at uerfarne personer er godt egnet for å justere flokkstørrelsen.

# Abstract

Good estimates are something the construction industry depends on giving, as estimates often guide construction projects. To provide estimates, various foresight methods are currently used. Wisdom of crowds has proven to be a method that provides answers with startling degrees of precision for a variety of problems. Nevertheless, there is little literature on how the phenomenon is applied in today's construction industry. The purpose of the master's thesis is to answer the question: "How can the wisdom of crowds be used to minimize cost risk in today's construction industry?".

The issue addresses the topics of wisdom of crowds and cost risk in the construction industry. In order to obtain an overview of existing literature on the topics, a literature review has been conducted. An experiment has been conducted to obtain empirical data. The experiment involved one super expert and 33 engineering students providing cost estimates for a turnkey contract for a combined office and teaching building. The 33 students were divided into four smaller crowds. The office and teaching building is a real project, and is part of the Norwegian Centre for Ocean Technology. The estimates of the super expert and the students were collected using a Mentimeter presentation. The presentation contained only selected parts of the tender basis for the turnkey contract. The super expert and the 33 students gave one estimate each in two rounds, in total two estimates each. Between the first and second round, the super expert and the students received information about the prize for previous projects.

In the results section, an arithmetic mean and the median of the estimates given by the students were calculated for both rounds. This was done for each of the four smaller crowds of students and for a crowd consisting of the estimates of all the students.

The results of the super expert and the 33 students have been analyzed. The analysis has looked at the cost estimates of the super expert and the 33 students, as well as estimates for the turnkey contract given by the developer (Statsbygg) and bid prices from three different contractors. Both Statsbygg and the contractors had access to the entire tender documentation when they prepared estimates and tenders respectively. As the turnkey contract is not finished, it will be impossible to say what the actual price of the construction project will be. Therefore, the contractors' prices are the closest the relevant turnkey contract comes to a price. Thus, the contractors' prices were chosen as a basis for comparison for Statsbygg's estimate, the estimates of the super expert and the estimates of the students. The median for the crowd of all the students in the second round was close to the contractors' prices. Statsbygg's estimate was also close to the contractors' prices. The super expert has given the least precise estimate. Nevertheless, it cannot be ruled out that the super expert has given a prize for the entire project, and not just for the turnkey contract.

Precise estimates were also observed for two of the four smaller crowds of students. However, it has been concluded that the uncertainty associated with the mean of the estimates is higher for smaller crowd sizes than for larger crowd sizes.

In the experiment, the super expert and the students received only general information about the turnkey contract. In the first round, the super expert and students had a top-down approach. In the second round, the top-down approach was combined with an outside perspective, as prices for other projects were announced. It was concluded that the combination of a top-down approach and the outside perspective has been a contributing factor to the median of the students being close to contractors' prices. Considering this, it is recommended to use wisdom of crowds for estimation tasks where little information is available.

The discussion has pointed out that by using the wisdom of crowds, bias that often occurs in estimation processes in the construction industry has been eliminated. It was concluded that herd mentality, underestimation, over-optimism, and the planning fallacy are biases countered by using the wisdom of crowds. However, the discussion of the results has revealed that it is likely that the super expert and the students in this experiment have been influenced by anchoring. The anchor is linked to the fact that it is likely that the participants have seen media reports about the costs of the entire Norwegian Ocean Technology Centre.

The analysis found that the uncertainty associated with the mean of the estimates was higher for smaller crowd sizes than for larger crowd sizes. Therefore, one measure to reduce risk would be to include more individuals in the crowd when using the wisdom of crowds. Crowd size should be adjusted according to the risk associated with what is to be estimated. It was concluded that digital tools have made it easier than ever to include people in the crowd. Based on the discussion, it has emerged that inexperienced people are well suited for adjusting the crowd size.

# Forord

Denne masteroppgaven er skrevet våren 2023 som en del av emnet TBM4900 Bygg- og miljøteknikk, masteroppgave. Oppgaven er skrevet hos Norges teknisk-naturvitenskapelige universitet ved Institutt for bygg- og miljøteknikk, og utgjør 30 studiepoeng. Oppgaven markerer slutten på min femårige integrerte mastergrad på bygg- og miljøteknikk, med studieretning bygg- og anlegg og hovedprofil byggeprosess.

Inspirasjonskilden for denne masteroppgaven var en forelesning i emnet TBA4130 Produksjonsledelse i BA-prosjekt på NTNU. Her ble jeg for første gang introdusert for fenomenet wisdom of crowds. Masteroppgaven har gitt meg muligheten til å fordype meg i dette fascinerende fenomenet. Dagens byggebransjen er helt avhengig av gode estimat, og forhåpentligvis kan denne masteroppgaven bidra til å belyse en alternativ fremsynsmetode.

En stor takk rettes til min veileder ved NTNU, Ole Jonny Klakegg, for gode råd, kontakter og faglige innspill. Takk rettes også til deltakerne i eksperimentet som har latt meg gjennomføre eksperimentet for dem, og benytte resultatene deres i min forskning. Uten klassene, Statsbygg og den eksterne eksperten hadde ikke forsøket latt seg gjennomføre.

Til slutt vil jeg takke mine venner og studievenner for å ha hjulpet meg med å komme gjennom sivilingeniørutdanningen.

Trondheim, 11. juni 2023

A handwritten signature in black ink, reading "Magnus Grønset", is written over a horizontal line.

Magnus Grønset



# Innhold

Figurer .....	xiv
Tabeller .....	xv
Forkortelser/symboler .....	xvi
Begrepsforklaring .....	xvi
1 Innledning .....	20
1.1 Bakgrunn .....	20
1.2 Formål og problemstilling .....	21
1.3 Avgrensninger .....	22
1.4 Leserveiledning .....	23
2 Metode .....	24
2.1 Metodelære .....	24
2.1.1 Triangulering .....	24
2.1.2 Valg av forskningsmetoder .....	24
2.1.3 Reliabilitet og validitet .....	25
2.2 Litteraturstudie .....	25
2.2.1 Fremgangsmåte .....	25
2.2.2 Styrker og svakheter med litteraturstudiets reliabilitet og validitet .....	28
2.3 Eksperiment .....	29
2.3.1 En forklaring av eksperimentet .....	29
2.3.2 Styrker og svakheter med eksperimentets reliabilitet og validitet .....	31
2.3.3 Beregninger gjort med eksperimentets dataverdier .....	32
3 Teoretisk rammeverk .....	34
3.1 Wisdom of crowds .....	34
3.1.1 Oppgavetyper .....	34
3.1.2 Flokkens sammensetning .....	35
3.1.3 Bias som kan påvirke flokkens deltakere .....	37
3.1.4 Arbeidsmetodikker for å gi estimat om fremtiden .....	40
3.2 Dagens byggebransje .....	45
3.2.1 Megaprojekt .....	45
3.2.2 Tilgjengelig informasjon og det digitale miljøets rolle for å redusere usikkerheten i byggeprosjekt .....	47
4 Resultat .....	51
4.1 Estimaten til byggherren .....	51
4.2 Entreprenørens estimat .....	52
4.3 Studentenes estimat .....	53

4.3.1	Flokk 1 av studenter .....	53
4.3.2	Flokk 2 av studenter .....	55
4.3.3	Flokk 3 av studenter .....	56
4.3.4	Flokk 4 av studenter .....	57
4.3.5	Flokken av alle studentene.....	58
4.4	Sammenligning av standardavvikene.....	59
4.5	Estimatene til supereksperten.....	59
4.6	Sammenligning av estimatverdiene .....	60
4.7	Fordeling av estimatene til Flokken av alle studentene.....	62
4.8	Antall studenters påvirkning på standardavviket til gjennomsnittet.....	65
4.9	Standardavvikets utvikling for ulike flokkstørrelser av studenter .....	67
4.10	Ekspertens punkter over nøkkelinformasjon .....	68
4.11	Individuelle estimat kontra flokkens samlede estimat .....	69
5	Analyse .....	75
5.1	Analyse av litteraturstudien .....	75
5.2	Tolkning av eksperimentets resultatet.....	75
5.2.1	Entreprenørens priser som sammenligningsgrunnlag .....	75
5.2.2	Median eller aritmetisk gjennomsnitt.....	76
5.2.3	Konsekvensene av for høye og for lave estimat .....	77
5.2.4	Størrelsens påvirkning på usikkerhet.....	78
6	Diskusjon.....	79
6.1	Hvordan kan digitale verktøy brukes for å hente inn kunnskap fra en flokk? ...	79
6.2	Hvordan bør sammensetningen til flokken være? .....	82
6.3	Hvilken informasjon bør flokken ha tilgjengelig?.....	85
6.4	Bias som påvirker personene som gir estimat .....	87
7	Konklusjon .....	92
7.1	Hvordan kan digitale verktøy brukes for å hente inn data fra en flokk? .....	92
7.2	Hvordan bør sammensetningen til flokken være? .....	92
7.3	Hvilken informasjon bør flokken ha tilgjengelig?.....	93
7.4	Hvilke bias påvirker flokkens deltakere?.....	93
8	Teoretiske implikasjoner .....	95
9	Praktiske implikasjoner.....	96
10	Et kritisk blikk på masteroppgaven .....	97
11	Videre arbeid .....	98
	Referanser.....	99
	Bilag .....	103





# Figurer

Figur 1: Viser fordelingen av svar for deltakere som fikk se anker, samt de som ikke fikk se et anker. Figuren viser spredningen av svarene (Görzen, 2019). .....	40
Figur 2: Viser et datasett med en symmetrisk fordeling av verdiene. Det aritmetiske gjennomsnittet og medianen er tegnet inn med stiplede linjer. Figuren er egenprodusert. ....	44
Figur 3: Viser et datasett med en skjevfordelt fordeling av verdiene. Det aritmetiske gjennomsnittet og medianen er tegnet inn med stiplede linjer. Figuren er egenprodusert. ....	44
Figur 4: Viser hvordan usikkerhet i prosjekt avhenger av størrelse, kompleksitet og innovasjon (Samset, 2017). .....	45
Figur 5: En illustrasjon av statens prosjektmodell (Finansdepartementet, 2019). .....	46
Figur 6: Illustrerer hvordan usikkerheten for et prosjekt endres over tid. Figuren er hentet fra kapittel 1 i Sunnevåg (2007). .....	47
Figur 7: Figuren viser sammenhengen mellom nytte og kostnad av informasjon for en økende informasjonsmengde (Samset, 2021). .....	48
Figur 8: En illustrasjon av forskjellen mellom bottom-up og top-down. Figuren viser hvordan kostnadene brytes ned, og følgerne det har for hva som har blitt utelatt. Figuren er hentet fra (Torp, 2021). .....	50
Figur 9: En grafisk fremstilling av hvordan estimatene til alle studentene er fordelt. Figuren viser hvordan estimatene fordelte seg i begge rundene. De samlede estimatene til eksperimentets deltakere er tegnet med stiplede linjer. ....	62
Figur 10: En grafisk fremstilling av hvordan estimatene til de mindre flokkene av studenter er fordelt. Figuren viser hvordan estimatene fordelte seg i begge rundene. De samlede estimatene til eksperimentets deltakere er tegnet med stiplede linjer. ....	63
Figur 11: Viser hvordan standardavviket til gjennomsnittet blir påvirket av antall deltakere. Figuren viser dette for både runde 1 og runde 2. ....	65
Figur 12: Viser standardavviket til gjennomsnittet av gjennomsnittet av 1 000 000 ulike kombinasjoner for hvert antall deltakere. Figuren viser dette for både runde 1 og runde 2. ....	67
Figur 13: Individuelle estimat og gjennomsnittsestimatet til Flokk 1, Flokk 2, Flokk 3, Flokk 4 og Flokk av alle studentene i runde 1, målt mot entreprenørenes gjennomsnittspris. ....	70
Figur 14: Individuelle estimat og gjennomsnittsestimatet til Flokk 1, Flokk 2, Flokk 3, Flokk 4 og Flokk av alle studentene i runde 1, målt mot Entreprenør 3. ....	71
Figur 15: Individuelle estimat og medianen til Flokk 1, Flokk 2, Flokk 3, Flokk 4 og Flokk av alle studentene i runde 1, målt mot entreprenørenes gjennomsnittspris. ....	72
Figur 16: Individuelle estimat og medianen til Flokk 1, Flokk 2, Flokk 3, Flokk 4 og Flokk av alle studentene i runde 1, målt mot Entreprenør 3. ....	73

# Tabeller

Tabell 1: Viser en oversikt over søkeordene brukt i hver av databasene, en beskrivelse av hvorfor søkeordene ble valgt og antall treff.....	27
Tabell 2: Viser forklaringer for underestimering i byggeprosjekt, samt betydningen av forklaringene. Tabellen er egenprodusert.....	38
Tabell 3: Entreprenørens tilbud.....	52
Tabell 4: Entreprenørens samlede estimat. ....	52
Tabell 5: Standardavviket til det aritmetiske gjennomsnittet til entreprenørens estimat. ....	52
Tabell 6: Kvartilene til entreprenørens estimat. ....	52
Tabell 7: Aritmetisk gjennomsnitt og median for Flokk 1 for hver av rundene.....	53
Tabell 8: Standardavviket til det aritmetiske gjennomsnittet for Flokk 1. ....	53
Tabell 9: Kvartilene til estimatene for Flokk 1. ....	53
Tabell 10: Aritmetisk gjennomsnitt og median for Flokk 2 for hver av rundene. ....	55
Tabell 11: Standardavviket til det aritmetiske gjennomsnittet for Flokk 2. ....	55
Tabell 12: Kvartilene til estimatene for Flokk 2.....	55
Tabell 13: Aritmetisk gjennomsnitt og median for Flokk 3 for hver av rundene. ....	56
Tabell 14: Standardavviket til det aritmetiske gjennomsnittet for Flokk 3. ....	56
Tabell 15: Kvartilene til estimatene for Flokk 3.....	56
Tabell 16: Aritmetisk gjennomsnitt og median for Flokk 4 for hver av rundene. ....	57
Tabell 17: Standardavviket til det aritmetiske gjennomsnittet for Flokk 4. ....	57
Tabell 18: Kvartilene til estimatene for Flokk 4.....	57
Tabell 19: Aritmetisk gjennomsnitt og median for Flokken av alle studentene for hver av rundene. ....	58
Tabell 20: Standardavviket til det aritmetiske gjennomsnittet for Flokken av alle studentene. ....	58
Tabell 21: Kvartilene til datasettet for Flokken av alle studentene.....	58
Tabell 22: En sortering av hvilke flokker av studenter som hadde lavest og høyest standardavvik for sine respektive gjennomsnitt. Øverst i tabellen er flokken med lavest standardavvik, og nederst er flokken med høyest standardavvik. Tabellen viser rekkefølgen for runde 1 og runde 2.....	59
Tabell 23: Estimaten til den eksterne eksperten.....	59
Tabell 24: Prosentvis sammenligning av studentenes og superekspertens estimat målt mot entreprenørene. + indikerer over og - indikerer under. ....	60
Tabell 25: Statsbyggs estimat sammenlignet med entreprenørens gjennomsnitt og Entreprenør 3. ....	60
Tabell 26: Viser en oversikt over antall treff for ulike databaser og søkeord fra litteraturstudien. En beskrivelse av søkeordene er også gitt. ....	104
Tabell 31: Samling av alle estimatene gitt av den første flokken.....	130
Tabell 32: Samling av alle estimatene gitt av den andre flokken.....	130
Tabell 33: Samling av alle estimatene gitt av den tredje flokken. ....	131
Tabell 34: Samling av alle estimatene gitt av den fjerde flokken.....	131

# Forkortelser/symboler

NTNU	Norges teknisk-naturvitenskapelige universitet
PDF	Portable Document Format
S&P500	Standard & Poor's 500
mva	Merverdiavgift
NOK	Norske kroner
kr	Kroner
MNOK	Millioner norske kroner
ekskl.	Eksklusiv
BTA	Bruttoareal
m <sup>2</sup>	Kvadratmeter

## Begrepsforklaring

### Aritmetisk gjennomsnitt

Aritmetisk gjennomsnitt er det som på folkemunne bare kalles «gjennomsnitt». Alle tall summeres, og summen deles på antallet tall. Formelen er gitt i Formel 1. I masteroppgaven brukes begrepet «gjennomsnitt» utelukkende for å omtale det aritmetiske gjennomsnittet.

Formel 1: Aritmetisk gjennomsnitt.

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$$

### Bias

Innenfor tematikken om wisdom of crowds og estimering er bias knyttet til personers atferd sentral. Flyvbjerg (2021) sier at atferdsbias består av både kognitive bias og politiske bias. Innenfor prosjektledelse finnes det en rekke ulike atferdsbias, hvor hver av biasene har sin egen definisjon.

### Byggherre

De som står for «utviklingen og gjennomføringen av prosjektet» (Torp, 2020).

### Collective intelligence

Collective intelligence er ifølge Lévy (1997, sitert i Blesik et al., 2022) «en prosess for å dele, modifisere og vurdere informasjon gitt av andre gruppemedlemmer, for å forstå verden bedre.» (min oversettelse). Definisjonen av begrepet viser at collective intelligence legger vekt på forståelse av verden.

### Crowdsourcing

For crowdsourcing finnes det flere ulike definisjoner. Crowdsourcing kjennetegnes av at en stor, ekstern, folkemengde er involvert, informasjonsinnhenting skjer via internett, terskelen for å delta er lav, deltakerne kjenner ikke hverandre og at deltakerne har liten kunnskap til produksjonsprosessene de er involvert i (Dvergsdal, 2023; Estellés-Arolas & González-Ladrón-de-Guevara, 2012). Dvergsdal (2023) sier at «betegnelsen er avledet av ordet outsourcing og handler om å legge til rette for at en anonym mengde arbeidere skal

utføre oppgaver som leder til et ønsket produkt.» og at det er en måte å anskaffe varer og tjenester.

Ettersom innhentingene ofte er ekstern og via internett er det få begrensninger for antall deltakere. Begrensningene vil være relatert til hvor mange som ønsker å være med og/eller et forhåndsdefinert antall satt av organisatorene. Flokker bestående av over 1000 deltakere er funnet for crowdsourcing (Estellés-Arolas & González-Ladrón-de-Guevara, 2012).

Dvergsdal (2023) sier at den mest vellykkede bruken av crowdsourcing er knyttet til: «markeds plasser for småsalg av programvare og digitale åndsverk, kunnskapskilder og databaser».

### **Data**

Kjell J. Sunnevåg sier i Del 4 kapittel 13 i Sunnevåg (2007) at data bare er en grov form for informasjon.

### **Eksoterisk problem**

Et eksoterisk problem er det motsatte av et esoterisk problem (Nilstun, 2018). Begrepet brukes om det som er allmennfattelig.

### **Ekspert**

En ekspert er et begrep som er definert ulikt fra flere hold. Smeby (2021) har definert en ekspert til å være en person som er spesielt dyktig eller sakkyndig. Budescu & Chen (2015) har sagt at det er vanskelig å definere ekspert. Ofte brukes begrepet om en person med historisk gode beslutninger eller estimat, lang erfaring, høy utdanning, høytstående stilling og/eller inngående kompetanse om typen problem som skal løses.

### **Esoteriske problem**

Et esoterisk problem er ifølge (Nilstun, 2018) et problem som er vanskelig å forstå for de fleste mennesker, og det krever vanligvis spesialisert kunnskap eller erfaring for å kunne bli fullt ut forstått.

### **Flokk**

I masteroppgaven har ordet flokk vært brukt som substitutt for det som i engelsk litteratur heter «crowd».

### **Gruppe**

Gruppe kan brukes overordnet for å omtale ulike sammensetninger av mennesker. For eksempel kan ordet gruppe brukes for å omtale gruppen av personer i en flokk, mens begrepet «flokk» kan ikke brukes for å omtale andre grupper.

### **Informasjon**

Kjell J. Sunnevåg sier i Del 4 kapittel 13 i Sunnevåg (2007) at data bare er en grov form for informasjon. Data blir informasjon når det blir relevant for beslutningsproblemet. I masteroppgaven har jeg valgt å bruke ordet informasjon for å omtale dataen jeg valgte ut fra tilbudsunderlaget, og ga til deltakerne i eksperimentet. Til tross for at dataen som jeg presenterte til deltakerne var nytt for dem, har jeg med rammene til eksperimentet ilagt innholdet en relevans.

## **Kollektiv visdom**

En samlebetegnelse for når flere mennesker bidrar med sin kunnskap og ekspertise.

## **Kommunikasjonskanal**

Ifølge Sander (2021) defineres kommunikasjonskanal som den fysiske og tekniske måten å spre et budskap på.

## **Kunnskap**

Kunnskap er ifølge Kjell J. Sunnevåg i Del 4 kapittel 13 i Sunnevåg (2007) «en form for bevissthet omkring, eller forståelse av fakta, etablerte sannheter eller informasjon ervervet gjennom erfaring eller læring (a posteriori), eller selvinnsikt (a priori)».

## **Median**

Medianen er verdien som deler et sortert tallmateriale i to. Den ene delen har verdier som er lavere enn medianen, mens den andre delen har verdier som er høyere enn medianen.

## **Mengde**

Brukes for å beskrive fysiske størrelser. Eksempelvis en mengde sand og en mengde data. Fra min side presiseres det at begrepet i oppaven ikke er brukt for å omtale en sammensetning av mennesker.

## **Mentimeter**

Mentimeter er et responssystem på nett. Deltakerne får tilgang til en Mentimeter-presentasjonen ved å taste inn en kode gitt av administratoren for presentasjonen. En presentasjon med Mentimeter innebærer at lysbildene i presentasjonen er synlig på deltakernes personlige enhet. For en Mentimeter-presentasjon er det mulig å benytte interaktive lysbilder, noe som innebærer at de som har presentasjonen åpen på en personlig enhet kan skrive inn tekst og/eller tall i en boks på den personlige enheten. Svarene lagres i presentasjonen.

## **Mikroflokker**

Gecer Ulu et al. (2019) har omtalt flokker med 5-14 personer som «mikroflokker» (microcrowds). Begrepet mikroflokker ble brukt om flokker hvor deltakerne hadde blitt trukket tilfeldig ut av en større flokk, og satt i flokker på 5-14 personer. Utvelgelsen skjedde fra en flokk som var satt sammen av tilfeldige ingeniørstudenter.

I eksperimentet i masteroppgaven ble studentene hentet fra én undervisningsklasse og tre forskjellige lesesaler. Samlingen av personene i undervisningsklassen og lesesalene ble kalt for Flokk 1 (9 personer), Flokk 2 (6 personer), Flokk 3 (14 personer) og Flokk 4 (4 personer). Alle personene i Flokk 1, Flokk 2, Flokk 3 og Flokk 4 var studenter. For eksperimentet ble det antatt at det var lite som skilte studentenes kompetanse. Derfor ble estimatene til studentene i de fire flokkene lagt sammen til et datasett. Datasettet som inneholdt estimatene til alle studentene, ble omtalt som Flokk av alle studentene (33 personer). I masteroppgaven er mikroflokker brukt for å omtale Flokk 1, Flokk 2, Flokk 3 og Flokk 4.

## **Personlig enhet**

Personlig enhet er en samlebetegnelse på håndholdte elektroniske enheter som kan sende signal over internett. Eksempler er telefon, nettbrett og bærbare datamaskiner.

## Prosjekteier

Den juridiske personen som «har eieransvaret for og eierrettighetene til prosjektet, og som i utgangspunktet bærer risikoen for prosjektets kostnader og bruksverdi» (Torp, 2020).

## Ressursbruk

En ressurs er av Nilstun (2023) definert som elementer eller menneskelige egenskaper som kan utnyttes økonomisk, eller som er til nytte på annen måte.

## Standardavvik for gjennomsnittet

Usikkerheten til et gjennomsnittsestimat uttrykkes ved hjelp av standardavviket til gjennomsnittet (Jelle, 2022). Formelen er gitt i Formel 2.

Formel 2: Standardavviket av gjennomsnittet.

$$\Delta x = \sigma_{\bar{x}} = \frac{\sqrt{\frac{\sum(x_i - \bar{x})^2}{N - 1}}}{\sqrt{N}}$$

## Støy

Støy er forhold som forstyrrer budskapet fra sender til mottaker (Olseng & Sundbye, 2020).

## Totalentreprise

Ifølge Welde et al. (2018) er en totalentreprise definert til å være en kontrakt der «Byggherren overlater både prosjekteringen og utførelsen til entreprenøren. Detaljeringsgraden i grunnlaget for entreprenørens prosjektering kan variere mellom alt fra overordnede funksjonsbeskrivelser, til mer detaljerte tegninger og beskrivelser. Det sentrale er imidlertid at ansvaret og risikoen knyttet til prosjekteringen overføres fra byggherre til entreprenør ved kontraktsinngåelse. Merk at selv om prosjekteringsrisikoen i prinsippet overføres til entreprenør, vil byggherre fortsatt sitte med noe risiko.»

## Uerfaren

I masteroppgaven har jeg valgt å bruke uerfaren for å betegne en person som har basiskunnskaper på et område, men ikke er en ekspert. Eksempler er studenter og nyutdannede.

## Wisdom of crowds

Begrepe wisdom of crowds, wisdom of the crowd og the wisdom of the crowd brukes om hverandre. Begrepe henviser alle til det samme fenomenet om at en flokk er i stand til å fatte en beslutning med høyere presisjon enn et enkelt individ. I mangelen på en god norsk oversettelse bruker oppgaven begrepet wisdom of crowds for å beskrive fenomenet wisdom of crowds.

# 1 Innledning

## 1.1 Bakgrunn

I min tidligere prosjektoppgave er det kommet frem at wisdom of crowds handler om at en flokk mennesker er i stand til å gi et samlet estimat med høyere presisjon enn et enkelt individ. Det første avsnittet i innledningen er hentet fra prosjektoppgaven. Fenomenets første vitenskapelige observasjoner kan spores tilbake til en artikkel skrevet av Francis Galton i 1907 (Ugander et al., 2015; Wallis, 2014; Yu et al., 2017). Artikkelen tar for seg en konkurranse hvor deltakere på en landbruksmesse betalte for å gjette vekten til en okse, hvor forslagene nærmest den faktiske vekten ble premiært (Wallis, 2014). Totalt 787 gyldige billetter ble vurdert, og Galton (1907) observerte at medianen av forslagene bare var 4 kg over den faktiske vekten av oksene. Etter publikasjonen har flere studier tatt for seg resultatene på nytt, og funnet at gjennomsnittet av målingene var 0,45 kg over oksens vekt. Flokkens sammenlagte forslag var bedre enn alle ekspertenes individuelle anslag, og lå bemerkelsesverdig nær den faktiske verdien. Fenomenet har siden blitt gjenskapt i flere studier på tvers av ulike disipliner, og resultatene peker i den retning at en gruppe mangfoldige mennesker vanligvis slår individuelle eksperter (Goldstein et al., 2014; Mannes et al., 2014; Surowiecki, 2005; Wallis, 2014; Yu et al., 2017). Fenomenet skiller seg fra vanlig gruppearbeid i den forstand at løsningene, eller forslagene, skapes individuelt, før de til slutt legges sammen. Som oftest legges estimatene sammen med å ta gjennomsnittet eller medianen av de individuelle estimatene.

Mitt første møte med fenomenet wisdom of crowds var i en forelesning i emnet «TBA4130 Produksjonsledelse i BA-prosjekt» på NTNU i Trondheim. Forelesningen ble gitt av en tidligere konsernsjef i en av Norges største bygg- og anleggsentreprenører. I forelesningen ble det fortalt hvordan entreprenøren brukte fenomenet wisdom of crowds i sin virksomhet. Det ble fortalt at bruken var relatert til å identifisere og tallfeste risiko og muligheter for byggeprosjekt. Dette ble gjort ved at en flokk mennesker fra firmaet ble samlet i et rom for å vurdere usikre kostnadsposter ved byggeprosjekt. Et eksempel på en slik post er grunnforhold. Deltakerne skulle gi sine vurderinger for hvor sannsynlig de anså det at forholdene beskrevet i postene var som forventet, verre enn forventet og bedre enn forventet, såkalte tripplestimat. For eksempel at grunnen inneholdt masser som er dyrere å håndtere enn forventet, som forventet eller rimeligere enn forventet. Ved å slå sammen medlemmenes estimat ble fenomenet wisdom of crowds utnyttet for å skaffe et samlet estimat for posten. Gjennom å identifisere risiko og muligheter har det vært mulig å håndtere usikkerheten i estimatene entreprenøren opererer med. For entreprenøren handlet det om å enten ikke påta seg risikoen eller å prise den inn i sine tilbud. Ved å bruke fenomenet har entreprenøren levert tilbud hvor kostnadsestimatene for prosjektene har stemt godt overens med de faktiske kostnadene.

Gode estimat er noe byggebransjen er avhengig av å gi, ettersom estimatene ofte er styrende for byggeprosjekt. Eksempler på dette er fremdriftsplaner, budsjett og beskrevet kvalitet. Ettersom et prosjekt måles mot disse parameterne er det en fordel at estimatene samsvarer i størst mulig grad med virkeligheten. Fascinasjonen min for at wisdom of crowds hadde bidratt til å oppnå gode estimat har ført til ønsket om å se videre på hvordan fenomenet kan brukes i dagens byggebransje.



## 1.2 Formål og problemstilling

Introduksjonen av digitale systemer som bygningsinformasjonsmodeller (BIM) og digitale kommunikasjonskanaler har ført til en betydelig økning i mengden av tilgjengelige data. Kjell J. Sunnevåg sier i Del 4 kapittel 13 i Sunnevåg (2007) at data bare er en grov form for informasjon. Data blir informasjon når det blir relevant for beslutningsproblemet. Knut Samset sier i Del 1 kapittel 1 i Sunnevåg (2007) at en vanlig definisjon som kopler usikkerhet til en beslutningssituasjon, er at usikkerhet er et uttrykk for mangel på informasjon for å kunne ta den rette beslutningen.

Austeng et al. (2005) sier at det er mye å hente på å forbedre usikkerhetsanalysene for prosjekt. To av punktene som fremheves er å forbedre kvaliteten på inngangsdataene og gjette bedre om fremtiden. I eksempelet brukt i innledningen er det vist at wisdom of crowds har blitt brukt til å forbedre inngangsdataene. Likevel er det ikke sett på hvordan fenomenet kan brukes i et nyere og digitalt miljø hvor det er større tilgang på data. Derfor ble problemstillingen til denne oppgaven valgt til å være:

*«Hvordan kan wisdom of crowds brukes til å minimere kostnadsrisiko i dagens byggebransje?».*

Eksempelet med veiingen av oksen viser flere av kjennetegnene for wisdom of crowds. Kjennetegnene er blant annet at det ble benyttet et stort antall lekfolk (787 personer), at oppgaven har latt seg besvare uten bakgrunnskunnskap og at flokkens samlede estimat var bedre enn ekspertene. Imidlertid krever flere av oppgavene innenfor byggebransjen en form for kunnskap for å kunne løses. Kunnskap er ifølge Kjell J. Sunnevåg i Del 4 kapittel 13 i Sunnevåg (2007) «en form for bevissthet omkring, eller forståelse av fakta, etablerte sannheter eller informasjon ervervet gjennom erfaring eller læring (a posteriori), eller selvinnsikt (a priori)». Et eksempel på en oppgave som krever kunnskap er å estimere prisen for et byggeprosjekt.

Så vidt jeg vet finnes det ikke litteratur som har sett på om et stort antall lekfolk gi mer presise estimat enn eksperter for prisen til et byggeprosjekt. Derfor har masteroppgaven gjennomført et eksperiment for å skaffe empiriske data. Eksperimentet har gått ut på at en flokk personer med lite erfaring fra byggebransjen, (studenter) og en supereksperter med lang erfaring fra å estimere byggekostnader, har gitt estimat for det samme byggeprosjektet.

Resultatene fra eksperimentet har deretter blitt kommentert og analysert. Analysen og resultatene fra eksperimentet var det som dannet grunnlaget for diskusjonen, og dermed vært en del av grunnlaget for å svare på oppgavens problemstilling. For å hjelpe til med å besvare problemstillingen ble det i tillegg lagd fire forskningsspørsmål:

1. *Hvordan kan digitale verktøy brukes for å hente inn data fra en flokk?*
2. *Hvordan bør sammensetningen til flokken være?*
3. *Hvilken informasjon bør flokken ha tilgjengelig?*
4. *Hvilke bias påvirker flokkens deltakere?*

Det første forskningsspørsmålet så på hvordan digitale verktøy kunne brukes for å hente inn data, fordi digitale verktøy har gjort det lettere enn noen gang å komme i kontakt med ressurspersoner uten å møtes fysisk.

Det andre forskningsspørsmålet så på sammensetningen av flokken, fordi eksperimentet utforsket hvor presise estimat en flokk uerfarne personer og en supereksperter var i stand til å gi. Dersom flokken av de uerfarne personene skulle vise seg å gi presise estimat, ville

det være interessant å se på hvordan størrelsen og ekspertisen til flokken påvirket estimatet.

Det tredje forskningsspørsmålet så på hvilken informasjon flokken burde ha tilgjengelig, fordi i mengden av data som de digitale verktøyene har gitt er det viktig å plukke ut dataen som er relevant for beslutningsproblemet. Dette er viktig ettersom det kun er relevant data som kan bli til informasjon. I henhold til definisjonen gitt innledningsvis i dette delkapittelet kan økt tilgang på informasjon være med på å redusere usikkerhet. Ved å redusere usikkerhet er muligheten for å redusere kostnadsrisiko større.

Det fjerde forskningsspørsmålet så på hvordan bias påvirket flokkens deltakere, fordi hvordan informasjonen i eksperimentet blir vurdert vil være avhengig av forutsetningene til personene som analyserer informasjonen. Paralleller kan bli trukket til Francis Galtons konkurranse hvor deltakerne skulle gjette vekten av en okse. Til tross for at alle deltakerne fikk se nøyaktig samme okse, hadde de ulike forutsetninger for å gjette korrekt vekt. En bonde som har jobbet tett på dyr vil mest sannsynlig ha en annen formening om vekten til oxen enn hva en advokat vil ha. Likevel vil en advokat ha kjennskap til vekten av andre gjenstander, og fra dette komme med et estimat. Hver av deltakerne vil inneha egne bias. I dette eksempelet er tanken med wisdom of crowds at de for høye estimatene kansellerer de for lave estimatene, og sannheten ligger sannsynligvis en plass mellom dem. Dette avhenger at det ikke foreligger systematisk bias som flytter intervallet vekk fra sannheten.

### 1.3 Avgrensninger

Litteraturen har vist at det er flere fenomener som stammer fra tankegangen om at en flokk mennesker kan ta bedre beslutninger enn individuelle eksperter. Ettersom masteroppgaven har hatt en gitt tidsramme, var det begrenset hvor mange sider av fenomenet som kunne belyses. En konsekvens av dette var at det har vært nødvendig å avgrense litteratursøkene.

Det ble valgt å utelate begrepet «crowdsourcing» som søkeord. Dette kommer av at litteraturen har pekt på at flokkene det er snakk om for crowdsourcing er store. Flokker med 1000 personer har blitt funnet (Estellés-Arolas & González-Ladrón-de-Guevara, 2012). Det ville vært både utfordrende å skaffe et så stort utvalg, og samtidig vanskelig å håndtere en så stor flokk. Andre søkeord som førte til treff der hvor begrepet crowdsourcing likevel var inkludert har blitt lest.

I min tidligere prosjektoppgave ble det avdekket at det finnes mye litteratur om fenomenet wisdom of crowds, og at litteraturen om wisdom of crowds befinner seg innenfor flere fagdisipliner. Ettersom fenomenet har vist seg i flere fagdisipliner var det naturlig at treff fra andre bransjer enn byggebransjen dukket opp i søkene. For å kunne svare detaljert på masteroppgavens problemstilling og forskningsspørsmål var det dermed nødvendig at søkene ble avgrenset. Basert på funnene i prosjektoppgaven, og råd fra veileder, ble det besluttet å avgrense litteratur om wisdom of crowds til å befinne seg innenfor ingeniørdisiplinene, atferdsøkonomi, økonomi og psykologi.

Litteraturen viste at det finnes flere beregningsmetoder for å komme frem til et statistisk oppsummeringstall. Eksempler på statistikk oppsummeringstall er median, aritmetisk gjennomsnitt, geometrisk gjennomsnitt, gjennomsnittet av median og aritmetisk gjennomsnitt, trimmet aritmetisk gjennomsnitt, korrigert aritmetisk gjennomsnitt, korrigert median, bayesiansk nettverk og «maximum-likelihood aggregation». Median og aritmetisk gjennomsnitt er målene som har blitt mest brukt i litteraturen. Oppgavens

problemstilling og forskningsspørsmål hadde ikke til hensikt å gå i dybden på hvordan ulike beregningsmetoder kunne brukes. Derfor ble oppgaven avgrenset til å kun ta for seg median og aritmetisk gjennomsnitt.

Av praktiske årsaker ble det besluttet å bare søke etter litteratur utgitt på engelsk og norsk.

I eksperimentet skulle studentene og supereksperter gi estimat for prisen av en totalentreprise for et byggeprosjekt som befinner seg i Norge. Byggeprosjektet har en norsk byggherre, og det var kun norske entreprenører som leverte tilbud. Derfor kan det sies at det i denne masteroppgaven er sett på wisdom of crowds for norske byggeprosjekt.

Den aktuelle totalentreprisen, som var gjenstand for prisestimatene i eksperimentet, var et kombinert kontor- og undervisningsbygg. Prosjektet er dermed et byggeprosjekt, og ikke et anleggsprosjekt. Ergo er oppgaven avgrenset til å omtale kostnader for byggeprosjekt.

Bias finnes i flere ulike varianter, og settinger. For byggeprosjekt er forankringseffekt, flokkmentalitet, underestimering, optimisme, planleggingsfeilen, søkelys på det unike og basisratefeil dratt frem å være blant de viktigste biasene (Flyvbjerg, 2021). Oppgaven er avgrenset til å kun undersøke de nevnte biasene.

## 1.4 Leserveiledning

Disposisjonen til masteroppgaven følger et IMTRAD-format. Formatet er delt opp i seks hoveddeler: introduksjon, metode, teori, resultat, analyse og diskusjon. Introduksjonsdelen gir en oversikt over problemstillingene som oppgaven skal besvare. Her presenteres også formålet med oppgaven, samt en kort beskrivelse av bakgrunnen for problemstillingen. Metoden gir en beskrivelse av hvordan problemstillingene ble studert. Her redegjøres det for hvilke metoder som ble benyttet, samt hvordan dataene ble samlet inn. Teorien danner et rammeverk for diskusjonen av resultatene. Resultatdelen inneholder funnene fra undersøkelsene. Analysen tolker og forklarer resultatene. Diskusjonen tar for seg betydningen av funnene og deres implikasjoner. Konklusjoner trekkes basert på resultatene og drøftingen i diskusjonen. I tillegg til de seks hoveddelene, inneholder masteroppgaven kapitlene videre arbeid, teoretiske implikasjoner, praktiske implikasjoner og et kritisk blikk på masteroppgaven.

## 2 Metode

Metodekapittelet beskriver hva en metode er, gir begrunnelser for valg av metoder og beskriver fremgangsmåten for å velge metoder. Kapittelet inkluderer en drøfting av styrker og svakheter med de valgte metodene.

### 2.1 Metodelære

Flere forfattere definerer metodebegrepet ulikt. Aubert (sitert i Hellevik 1999) gir at en metode er «en fremgangsmåte, et middel til å løse problemer og komme frem til ny kunnskap. Et hvilket som helst middel som tjener dette formålet, hører med i arsenalet av metoder.». Hellevik (1999) sier at metodelæren «gir råd om hvordan forskeren kan samle inn, behandle og analysere data.». Østbye og Knapskog (1994) viser at en metode er «en planmessig måte for å nå fram til kunnskap på.». Til tross for at definisjonene er formulert ulikt, er forståelsen av hva en metode er relativt lik. Videre beskrives ulike forskningsmetoder, valg av forskningsmetoder, begrunnelse for valg av forskningsmetoder og validitet og reliabilitet.

#### 2.1.1 Triangulering

Ettersom det finnes flere ulike metoder er det viktig at valget av metode er tilpasset målet (Dalland, 2001). Valg av metode må dermed vurderes både med bakgrunn i hva som skal oppnås, og samtidig hensynta hva som er praktisk gjennomførbart. Ofte kombineres flere ulike metoder, og dette kalles triangulering. Triangulering bidrar til å redusere svakheter med de valgte forskningsmetodene. Både kvalitative og kvantitative forskningsmetoder har til hensikt å belyse det samme målet, og det er anbefalt å kombinere metodene.

#### 2.1.2 Valg av forskningsmetoder

Oppgavens formål og forskningsspørsmål la føringer for hvilke metoder det egnet seg best å bruke. Problemstillingen innebar å beskrive fenomenet wisdom of crowds, og samtidig redegjøre for hvordan fenomenet passer inn i byggebransjen. Hensikten til oppgaven var dermed å utdype hvordan et kjent fenomen opptrer i byggebransjen, og problemstillingen kan sies å ha vært eksplorerende (Jacobsen, 2022). Jacobsen (2022) gir at en eksplorerende problemstilling ofte krever en metode som får frem nyanserte data, går i dybden og er følsom for uventede forhold og dermed er åpen for kontekstuelle forhold. En følge av å studere nyanser er at forskningen ofte begrenses til å konsentrere seg om noen få enheter. Metoden beskrevet over egner seg til innsamling av kvalitative eller åpne data. At problemstillingen skulle beskrive et fenomen i en gitt kontekst, talte for at det burde brukes en metode som var i stand til å samle inn kvalitative data.

For å besvare forskningsspørsmålene ble et eksperiment gjennomført. Eksperimentet tok sikte på å teste hvordan fenomenet wisdom of crowds slår ut for en flokk studenter og en supereksperter for en oppgave innenfor byggebransjen. Så vidt jeg vet finnes det ikke litteratur som har sett på bruken av wisdom of crowds for en flokk studenter der en oppgave innenfor byggebransjen skal løses. For å skaffe empiriske data som ville være med på å besvare forskningsspørsmålene, ble det i masteroppgaven gjennomført et eksperiment. Fra min side ble det besluttet at oppgaven flokken studenter og supereksperter skulle løse var å estimere kostnaden for en totalentreprise. Derfor har det

i eksperimentet vært nødvendig å benytte en metode hvor det samles inn kvantitative data.

Samlet sett ga dette at for delen av oppgaven som omhandlet eksperimentet ble det benyttet en kvantitativ metode, og for delene hvor fenomenet skulle belyses ble det benyttet en kvalitativ metode.

Ved å se separat på wisdom of crowds og byggebransjen, finnes det omfattende litteratur om begge temaene. Den kvalitative metoden som skulle brukes måtte dermed være i stand til å undersøke eksisterende litteratur om wisdom of crowds og byggebransjen. Samtidig var det ønskelig at litteraturen som ble funnet kunne gjennomgås nøye. Ved å gjennomgå litteraturen nøye var tanken at elementer fra litteraturen om wisdom of crowds og byggebransjen kunne brukes for å gi føringer for eksperimentets utforming og utførelse. Samlet sett skulle litteraturen danne et teoretisk rammeverk. Det teoretiske rammeverket skulle fungere som et rammeverk for diskusjonen av resultatene til eksperimentet. Det fremstår følgelig at det var behov for en metode som var i stand til å utforske litteratur, og samtidig gå gjennom litteraturen nøye. En metode som passet til dette formålet var en litteraturstudie. Derfor ble en litteraturstudie valgt som kvalitativ metode.

Føringene for eksperimentets gjennomførelse bygget i stor grad på hva de kvalitative undersøkelsene viste. Eksperimentet var et resultat av å gå fra teori til empiri, og var styrt av tidligere teoretiske antakelser og empiriske funn. Følgelig kan det sies at forskningen har hatt en deduktiv tilnærming (Jacobsen, 2022).

### 2.1.3 Reliabilitet og validitet

Forskning som tar sikte på å samle inn empiri bør tilfredsstillende krav til validitet og reliabilitet (Jacobsen, 2022). Med validitet menes det at empirien må være gyldig og relevant, og med reliabilitet menes det at empirien må være pålitelig og troverdig. Gyldighet og relevans knyttes til om empirien som blir samlet inn gir svar på spørsmålene stilt til forskningen. Pålitelighet og troverdig er relatert til om forskningen er til å stole på. Dette innebærer at forskningen må gjennomføres på en troverdig og tillitsvekkende måte.

## 2.2 Litteraturstudie

### 2.2.1 Fremgangsmåte

I min tidligere prosjektoppgave ble store deler av litteraturen om wisdom of crowds kartlagt. Litteraturen strakk seg over flere fagdisipliner. Prosjektoppgaven brukte en scoping review for kartleggingen. Litteraturen ble ikke studert inngående. Masteroppgaven har gjennom en litteraturstudie gått i dybden på deler av stoffet funnet i prosjektoppgaven. Gjennom litteraturstudien ble det også utført nye litteratursøk. Søkene i litteraturstudien var mer spisset mot masteroppgavens problemstilling og forskningsspørsmål. Litteraturen funnet ved hjelp av litteraturstudien var med på å danne det teoretiske rammeverket i masteroppgaven.

Litteratur som ble funnet i prosjektoppgaven, og samtidig var relevant for masteroppgavens problemstilling og forskningsspørsmål, ble valgt ut etter en skjønnsmessig vurdering. Tekst i masteroppgaven, som baserer seg på referansene funnet som en del av prosjektoppgaven, omtales spesifikt der hvor det er aktuelt.

Gjennom litteraturstudien har det blitt søkt etter ny litteratur. Søkene har hatt til hensikt å finne litteratur om wisdom of crowds, bias, kostnadsrisiko og byggebransjen. Temaene

er godt dokumentert i vitenskapelige artikler. Derfor er det å forvente at flere databaser vil kunne inneholde treff om temaene. Dog har masteroppgaven hatt en tidsfrist, og dermed har det vært begrenset hvor mye litteratur det har vært mulig å rekke over. Derfor var det vært nødvendig å avgrense søkene til få databaser.

For å velge databaser ble det tatt utgangspunkt i avgrensningene i masteroppgaven. Søkene ble avgrenset til å stamme fra ingeniørdisiplinene, atferdsøkonomi, økonomi og psykologi. For å få treff innenfor alle disse fagfeltene måtte databasene være tverrfaglig. For at litteraturen skulle være relevant for masteroppgaven, måtte litteraturen i databasene være vitenskapelig. Fra problemstillingen var det gitt at dagens byggebransje skulle utforskes. Derfor var det nødvendig at databasene kontinuerlig innlemmet ny litteratur. Databasene måtte også være tilgjengelig uten en finansiell hindring for meg. Databasene Scopus og Oria passet begge til beskrivelsen, og ble dermed valgt.

For å få et skarpere bilde over hvordan det jobbes i dagens byggebransje ble det besluttet å finne litteratur innenfor Concept-programmet (*Forskningsprogrammet Concept - NTNU, u.å.*). Concept-programmet er et forskningsprogram finansiert av Finansdepartementet. Programmet har til hensikt å «utvikle kunnskap som sikrer bedre konseptvalg, ressursutnytting og effekt av store statlige investeringer.» (*Forskningsprogrammet Concept - NTNU, u.å.*). Concept-programmet har en rekke publikasjoner. For denne masteroppgaven ble det benyttet litteratur fra «Concept rapportserie», «Concept Temahefter» og «Arbeidsrapporter». Publikasjonene innenfor «Concept rapportserie» godkjent som vitenskapelig publiseringskanal på nivå 1, «Concept Temahefter» er populariserte kortversjoner av doktorgradsavhandlinger og andre større publikasjoner og «Arbeidsrapporter» er «rapporter som enten dekker spesialtema som er for smale til å bli utgitt som egen Concept-rapport, eller rapporter som presenterer underveisresultater som grunnlag for videre arbeid. Arbeidsrapporter vil normalt ikke ha vært underlagt de samme strenge rutinene for kvalitetssikring fra Concept-programmets side, som Concept-rapporter i den offisielle serien.» (*Forskningsprogrammet Concept - NTNU, u.å.*).

Søkene ble først gjort i Scopus. Kun engelske søkeord ble brukt, ettersom databasen ikke inkluderer litteratur skrevet på norsk. De første søkeordene favnet bredt innenfor de gitte avgrensningene. Mot slutten ble søkene mer spisset. Etter en skjønnsmessig vurdering for at aktuell litteratur var funnet i Scopus ble de samme søkeordene brukt i Oria. Deretter ble det i Oria søkt etter litteratur med norske søkeord. Et utdrag fra oversikten over benyttede databaser, søkeord, beskrivelse av søkeord og antall treff er gitt i Tabell 1. Den fullstendige oversikten finnes i Bilag 1. For å finne litteratur innenfor Concept-programmet ble det ikke benyttet søkeord. Litteraturen ble funnet ved å bla gjennom publikasjonene innenfor kategoriene «Concept rapportserie», «Concept Temahefter» og «Arbeidsrapporter» på nettsiden til Concept-programmet.

Tabell 1: Viser en oversikt over søkeordene brukt i hver av databasene, en beskrivelse av hvorfor søkeordene ble valgt og antall treff.

Database	Søkeord	Beskrivelse	Treff
Scopus	TITLE-ABS-KEY ( wisdom AND of AND the AND crowd AND engineering )	Søket inneholder kjernen i problemstillingen, wisdom of crowds. Søket er forsøkt avgrenset til byggebransjen for å sikre relevans til oppgavens problemstilling.	99
Scopus	TITLE-ABS-KEY ( wisdom AND of AND the AND crowd AND information AND available ) AND ( LIMIT-TO ( SUBJAREA , "BUSI" ) OR LIMIT-TO ( SUBJAREA , "DECI" ) OR LIMIT-TO ( SUBJAREA , "SOCI" ) OR LIMIT-TO ( SUBJAREA , "ENGI" ) OR LIMIT-TO ( SUBJAREA , "ECON" ) OR LIMIT-TO ( SUBJAREA , "ARTS" ) )	Søket er relevant for forskningsspørsmål nummer 3 inneholder avgrensninger relatert til fagfelt.	36
Oria	wisdom of the crowd information available  Filter: Fra fagfelleverderte tidsskrift Artikler Konferanseforedrag Bokkapitler Avhandlinger Engelsk Science & Technology Social Sciences Decision Making Wisdom Of Crowds Business & Economics Psychology Management Fra fagfelleverderte tidsskrift Artikler Konferanseforedrag Bokkapitler Avhandlinger Engelsk Science & Technology Social Sciences Decision Making Wisdom Of Crowds Business & Economics Psychology Management	Relatert til forskningsspørsmål 4. Relativt likt som det tidligere søket i Scopus.	175

Utvelgelseskriteriene var overskrift, sammendrag og innhold. Litteratur ble plukket ut fra søkene basert på en skjønsmessig vurdering av overskriften. Deretter ble sammendragene for den valgte litteraturen lest. Dersom litteraturen fortsatt var ansett som relevant, ble den lest gjennom nøye. Etter å ha lest den aktuelle litteraturen ble referanselisten lest. Å lese referanselisten ble gjort for å identifisere mer litteratur om temaet (Wohlin, 2014). Ved å lese referanselistene fikk jeg også en oversikt over hvilken litteratur som gikk igjen på området.

Som nevnt tidligere ble det først søkt i databasen Scopus, før det deretter ble søkt i Oria. En konsekvens av dette var at flere av treffene i Oria var duplikat av litteraturen funnet i Scopus. Utvelgelsen av litteratur fra Oria ble avsluttet da det for meg virket som at hvert resultat som dukket opp allerede hadde blitt funnet i Scopus. Gjennom litteraturstudiet for masteroppgaven har det dukket opp litteratur som også ble brukt i prosjektoppgaven. Litteraturen ble funnet uavhengig av prosjektoppgaven, og må dermed anses som en del av arbeidet med denne masteroppgaven.

I Excel ble det ført logg over når søk ble utført, hvilken database det ble søkt i, benyttede søkeord, antall treff hvor hvert søk, antall valgt litteratur, overskrift for valgt litteratur, personlige tanker og dato for når hele artikkelen var lest. For å analysere litteraturen har programmet NVivo blitt benyttet. NVivo er et program som støtter kvalitative data og blandede forskningsmetoder, og har derfor egnet seg som analyseverktøy for litteraturstudiet (*NVivo - Kunnskapsbasen - NTNU, u.å.*).

### 2.2.2 Styrker og svakheter med litteraturstudiets reliabilitet og validitet

Litteraturstudien i masteroppgaven hadde både styrker og svakheter. En styrke med litteraturstudien var at litteraturen ble funnet i anerkjente forskningsdatabaser. At søkene i både Scopus og Oria ga mange like treff om wisdom of crowds taler for at det meste av relevant litteratur om wisdom of crowds ble funnet. Det kan likevel ikke utelukkes at bruk av flere databaser kunne gitt flere unike treff om wisdom of crowds. Imidlertid har masteroppgaven hatt en tidsavgrensning, og det har dermed vært hensiktsmessig å begrense søkene til få databaser.

Det meste av litteraturen om dagens byggebransje ble funnet i Concept-programmets «Concept rapportserie», «Concept Temahefter» og «Arbeidsrapporter». En konsekvens av dette var at litteraturen som tok for seg dagens byggebransje i all hovedsak tok for seg dagens byggebransje i Norge. Det er sannsynlig at dagens byggebransje i Norge er annerledes enn dagens byggebransje i India. Problemstillingen «Hvordan kan wisdom of crowds brukes til å minimere kostnadsrisiko i dagens byggebransje» har ikke spesifisert hvilken «dagens byggebransje» det er snakk om. Derfor kan det sies at avgrensningen til å bruke norsk litteratur for å beskrive dagens byggebransje har vært med på å svekke metodens validitet.

Imidlertid har valget av å beskrive dagens byggebransje med hovedsakelig norsk litteratur bidratt til å gi et klarere bilde av én av verdens «dagens byggebransjer». Dersom alle «dagens byggebransjer» hadde blitt undersøkt ville litteraturen pekt i ulike retninger. Derfor har avgrensningen med å undersøke hovedsakelig norsk litteratur for å beskrive dagens byggebransje bidratt til å gi et mer konsentrert bilde av én «dagens byggebransje». Derfor har avgrensningen med å bruke norsk litteratur for å beskrive «dagens byggebransje», gitt god reliabilitet.

Selv om data i utgangspunktet er relevante, må de også være samlet inn på en slik måte at de er pålitelige (Dalland, 1993). Dette stilte krav til at informasjonen måtte hentes inn



med høy grad av nøyaktighet. Ettersom jeg har stått for utvelgelsen, har det følgelig vært en risiko for at data har blitt misforstått eller oversatt. Dette utgjorde en svakhet med metodens reliabilitet. At det er blitt ført en detaljert logg av søkene har imidlertid gjort at det har vært mulig å gå tilbake og sjekke hvorfor litteratur ble valgt eller valgt bort. Loggen har følgelig vært med på å øke metodens reliabilitet.

Søkene inneholdt avgrensninger til både språk og fagdisipliner. Selv om det er søkt bredt innenfor avgrensningene kan det ikke utelukkes at informasjon fra andre disipliner og på andre språk kunne vært relevant. En styrke med avgrensningene har vært at det har gjort det mulig å gå nøye gjennom litteraturen som ble funnet. Dersom litteratur på andre språk hadde blitt inkludert, måtte litteraturen ha blitt oversatt. Dette hadde det åpnet opp for en annen feilkilde i form av feiltolkning eller feil oversettelse.

En svakhet med litteraturstudien var at det ikke var avgrensninger knyttet til hvor gammel litteraturen kunne være. Dette førte til at søkene ga treff med gammel litteratur. I masteroppgaven skulle dagens byggebransje utforskes. I og med at byggebransjen gjennom de siste årene har utviklet seg, er det ikke sikkert at gammel forskning fortsatt er gjeldende for å beskrive dagens situasjon. Derfor hadde det vært en styrke om kun nyere forskning hadde blitt brukt for å beskrive byggebransjen. Derfor burde det vært en avgrensning knyttet til publiseringsdato for søkene som skulle finne litteratur om dagens byggebransje.

For å beskrive fenomenet wisdom of crowds var det likevel en styrke at gammel litteratur ble inkludert. Dokumenterte observasjoner av fenomenet har blitt sporet tilbake til 1907. Fra 1907 og frem til i dag er det en rekke dokumenterte observasjoner av fenomenet. Med en avgrensning for publiseringsdato kunne interessante observasjoner blitt utelatt.

## 2.3 Eksperiment

### 2.3.1 En forklaring av eksperimentet

Eksperimentet gikk ut på at én superekspert og 33 studenter skulle gi kostnadsestimat for en totalentreprise for et kombinert kontor- og undervisningsbygg. De 33 studentene var fordelt på fire mindre flokker med ulik størrelse. Studentene i de fire mindre flokkene befant seg i én undervisningsklasse og på tre forskjellige lesesaler på NTNU i Trondheim. Alle studentene som ga kostnadsestimat, studerte ved studieprogrammet bygg- og miljøteknikk. Samlingen av studentene i undervisningsklassen og de tre forskjellige lesesalene omtales i denne masteroppgaven som Flokk 1 (9 personer), Flokk 2 (6 personer), Flokk 3 (14 personer) og Flokk 4 (4 personer). Alle personene i Flokk 1, Flokk 2, Flokk 3 og Flokk 4 var studenter. For eksperimentet er det antatt at det er lite som skiller studentenes kompetanse. Dette ble begrunnet med at studentene studerer på det samme universitet og innehar lite erfaring fra arbeidslivet. Basert på denne antakelsen har kostnadsestimatene til studentene i de fire mindre flokkene blitt lagt sammen til et datasett. Datasettet som inneholdt kostnadsestimatene til alle studentene omtales i denne masteroppgaven som Flokk av alle studentene (33 personer). Årsaken til at studentene var fordelt i mindre flokker var fordi det ikke lyktes å finne én klasse hvor det befant seg mer enn 30 personer samtidig.

Supereksperten hadde 34 års erfaring med å estimere kostnaden for byggeprosjekt. Kostnaden for byggeprosjekt utgjør alle midlene som går med for å realisere byggeprosjektet. Kostnaden for en totalentreprise er bare en del av midlene som går med på å realisere hele byggeprosjektet. 34 års erfaring innenfor fagfeltet å estimere kostnaden

for byggeprosjekt er likevel lenger erfaring enn de aller fleste i byggebransjen. Derfor kan personen kalles en superekspert for å estimere kostnaden for byggeprosjekt. Fra et yrkeslivs perspektiv representerer supereksperten og studentene ytterpunktene av erfaring.

Studentene og supereksperten skulle gi prisestimat for er en totalentreprise for et kontor- og undervisningsbygg. Totalentreprisen for byggeprosjektet er ekte, og inngår som en del av Norsk havteknologisenter. Studentene og supereksperten skulle gi kostnadsestimat for denne totalentreprisen fordi prisen for totalentreprisen ikke hadde blitt gjort tilgjengelig for allmennheten.

Til tross for ulik erfaring hos supereksperten og studentene skulle de basert på samme mengde informasjon gi kostnadsestimat for totalentreprisen. Av meg ble informasjonen presentert for supereksperten og studentene gjennom en Mentimeter-presentasjon. Mentimeter-presentasjonen innebar at supereksperten og hver av studentene hadde lysbildene til presentasjonen synlig på personlige enheter. Mentimeter-presentasjonen hadde åtte lysbilder med informasjon, og to lysbilder som var interaktive. Interaktive lysbilder innebar at de som hadde presentasjonen åpen på en personlig enhet kunne skrive inn tekst og/eller tall i en boks på den personlige enheten. Gjennom de interaktive lysbildene supereksperten og studentene sendte inn estimat. Etter at estimatene ble sendt inn i de interaktive lysbildene ble de lagret i presentasjonen. Svarene var ikke synlig for andre enn meg. Lysbildene med informasjon inneholdt kun utvalgte deler av tilbudsunderlaget for totalentreprisen. Delene av tilbudsunderlaget som ble inkludert i presentasjonen er listet opp under. Før presentasjonen startet ble det for studentene og supereksperten opplyst at Statsbygg er byggherre for prosjektet.

- Skisser av bygget
- Situasjonsplan for bygget
- Aksonometri for bygget
- Bilder fra en BIM-modell av bygget
- Formål med bygget
- Bruttoareal for bygget
- Opplysning om at rivning av eksisterende bebyggelse er en del av kontrakten

Supereksperten og studentene skulle gi ett estimat hver i to runder, i sum to estimat hver. I den første runden hadde supereksperten og studentene kun data jeg hadde hentet fra tilbudsunderlaget. Ettersom jeg hadde ansett den utvalgte dataen fra tilbudsunderlaget som relevant for beslutningsproblemet, kan det sies at deltakerne har mottatt informasjon.

Den første runden ble avsluttet når hver av studentene i flokkene, og supereksperten, hadde gitt ett estimat. Den andre runden ble startet rett etter den første runden var avsluttet. Den andre runden startet med at det i Mentimeter-presentasjonen ble vist frem priser fra Norsk prisbok. Prisene fra Norsk prisbok var pris/BTA  $m^2$  (ekskl. mva.) for en tidligere videregående skole (undervisningsbygg) og priser for to kontorbygg med ulik størrelse. Utover prisene fra Norsk prisbok ble det ikke gitt mer informasjon enn hva som var gitt i den første runden. Den andre runden ble avsluttet når hver av studentene i flokkene, og supereksperten, hadde gitt et nytt estimat. Presentasjonen er gitt i Bilag 2.

Hver av estimatene til supereksperten og studentene ble gitt uavhengig av hverandre, og helt individuelt. Hverken eksperten eller studentene hadde lov til å diskutere oppgaven med andre. Det var kun lov å stille meg spørsmål dersom noe var uklart. Det var heller ikke lov å søke etter annen informasjon enn den som ble gitt i Mentimeter-presentasjonen.

Det var flere årsaker til at studentene og supereksperter kun fikk se utvalgte deler av tilbudsunderlaget i Mentimeter-presentasjonen. Fra min side var det ønskelig at studentene og supereksperter skulle vurdere den samme informasjonen. Dessuten var det fra min side ønskelig å se på hvor presise/upresise estimat det var mulig å gi for et byggeprosjekt, selv med lite tilgjengelig informasjon. Utover dette anser jeg det som sannsynlig at mange av studentene hadde takket nei til å være med i eksperimentet dersom det stilte krav til forberedelser.

### 2.3.2 Styrker og svakheter med eksperimentets reliabilitet og validitet

I delkapittel 1.2 ble det nevnt at så vidt jeg vet finnes det ikke litteratur som har sett på om et stort antall lekfolk gi mer presise estimat enn eksperter for prisen til et byggeprosjekt. Eksperimentet har vært med på å hente inn empirisk data for å se om det er mulig for et stort antall lekfolk å gi mer presise estimat enn eksperter for prisen til et byggeprosjekt. Dataen her har blitt lagt til grunn for å kunne besvare oppgavens problemstilling og forskningsspørsmål. Derfor har valget av å gjennomføre et eksperiment vært med på å gi oppgaven høy validitet.

Det første forskningsspørsmålet tok for seg hvordan digitale verktøy kunne bidra til å hente inn data fra en flokk. For å hente inn data fra en flokk benyttet eksperimentet det digitale verktøyet Mentimeter. Valget av å bruke Mentimeter var dermed med på å gi høy validitet for det første forskningsspørsmålet. Imidlertid var det en svakhet for det første forskningsspørsmålets validitet at Mentimeter var det eneste digitale verktøyet som ble brukt i eksperimentet.

Ved å bruke interaktive lysbilder i Mentimeter-presentasjon fikk studentene og eksperten selv taste inn estimat. Etter at estimatene ble sendt inn ble de lagret i presentasjonen. At estimatene ble gitt direkte fra studentene og eksperten, og samtidig lagret med en gang, har bidratt til at datainnsamlingen har vært presis. Dette har vært med på å styrke reliabiliteten til datainnsamlingen. Naturligvis var det en risiko for at personene kunne taste inn feil estimat, ettersom det ikke forelå noen begrensninger for hva de kunne taste inn. Imidlertid anser jeg det som sannsynlig at dersom dette var tilfellet så kunne den aktuelle personen sagt ifra til meg, og så kunne jeg rettet opp i feilen med en gang.

Det andre forskningsspørsmålet tok for seg sammensetningen til flokken. I eksperimentet var studentene fordelt på fire mindre flokker med ulik størrelse. At flokkene har hatt ulik størrelse har gjort det mulig å studere resultatene for ulike flokkstørrelser. Derfor har det å ha fire mindre flokker vært med på å gi økt validitet for det andre forskningsspørsmålet. At studentene var fordelt på mindre flokker har imidlertid åpnet opp for feilkilder for eksperimentet. Ettersom presentasjonen ble holdt av meg har det vært en risiko for at jeg har presentert ting på ulikt vis. Dette har vært med på å svekke eksperimentets validitet for det andre forskningsspørsmålet. Et manus burde vært fulgt for å minimere feilkilden.

Eksperimentet gikk ut på å hente estimat fra både én superekspert og flokker bestående av studenter. Studentene og supereksperter representerer ytterpunktene av erfaring. Samtidig representerer studentene og supereksperter forskjellen mellom at mange gir et estimat (studentene) og få som gir et estimat (supereksperter). Motsetningene ble brukt som en del av diskusjonen for det andre forskningsspørsmålet. Derfor har motsetningene vært med på å øke validiteten for det andre forskningsspørsmålet.

Det tredje forskningsspørsmålet tok for seg hvilken informasjon flokken burde ha tilgjengelig. I eksperimentet ga studentene og supereksperter estimat i to runder. Mellom den første og den andre runden fikk studentene og supereksperter ekstra informasjon.

Den ekstra informasjonen var pris for tidligere prosjekt. Ved å se på forskjellen mellom estimatene gitt i den første og andre runden har det vært mulig å se hvordan supereksperter og studentene har respondert på informasjonen om pris for tidligere prosjekt. Dette har vært med på å øke validiteten for det tredje forskningsspørsmålet. Imidlertid var ikke den ekstra informasjonen om pris på egenhånd avgjørende for hvilke estimat supereksperter og studentene ga. Informasjonen om pris ble gitt i tillegg til informasjon hentet fra tilbudsunderlaget. Derfor har det vært vanskelig å si hvor viktig informasjonen om pris på egenhånd har vært. Dette har vært med på å svekke reliabiliteten for det tredje forskningsspørsmålet.

Det fjerde forskningsspørsmålet tok for seg hvilke bias som kunne påvirke flokkens deltakere. Studentene og supereksperter fikk begge opplysninger om at kontor- og undervisningsbygget inngikk som en del Norsk havteknologisenter. Isolert sett var kontor- og undervisningsbygget ukjent for supereksperter og studentene. Imidlertid var konteksten, Norsk havteknologisenter, kjent for flere av dem. Norsk havteknologisenter har fått mye oppmerksomhet i mediene, med flere oppslag om totalkostnaden for hele prosjektet. Summene nevnt i artiklene har vært flere milliarder kroner høyere enn de kostnadene det er estimert at kontor- og undervisningsbygget i dette eksperimentet vil koste. Likevel har litteraturen pekt på at mennesker kan gi et nytt tallestimat basert på en gammel verdi. Tendensen med å basere nye tallestimat på gamle verdier omtales som forankringsbias. I eksperimentet har konteksten med Norsk havteknologisenter vært et mulig anker for studentene og supereksperter. Ved å være klar over at summene nevnt for Norsk havteknologisenter kunne fungere som et anker, har det vært mulig å se om studentene og supereksperter har latt seg påvirke av forankringsbias. Dette har vært med på å øke validiteten til det fjerde forskningsspørsmålet.

Imidlertid har eksperimentet ikke blitt holdt for personer som ikke visste om at kontor- og undervisningsbygget var en del av Norsk havteknologisenter. Dermed har det ikke vært mulig å sammenligne resultatet mot en kontrollgruppe uten forankringsbias. Dette har vært med på å svekke reliabiliteten til det fjerde forskningsspørsmålet.

Til tross for at jeg var klar over biaset på forhånd ble hverken studentene eller supereksperter spurt om de kjente til Norsk havteknologisenter. Dermed har det heller ikke vært mulig å se hvem som kunne vært påvirket av forankringsbiaset. Dette har vært med på å svekke reliabiliteten til det fjerde forskningsspørsmålet.

### 2.3.3 Beregninger gjort med eksperimentets dataverdier

En sentral del av wisdom of crowds er at forslagene til flokkens deltakere legges sammen til et samlet forslag. Et samlet forslag er i sammenheng med eksperimentet i denne masteroppgaven et statistisk oppsummeringstall for hver av flokkene av studenter. I masteroppgaven uttrykkes flokkenes statistiske oppsummeringstall som et aritmetisk gjennomsnitt og median.

Aritmetisk gjennomsnitt er det som på folkemunne bare kalles «gjennomsnitt». Alle tall summeres, og summen deles på antallet tall. Formelen er gitt i Formel 1. I masteroppgaven brukes begrepet «gjennomsnitt» utelukkende for å omtale det aritmetiske gjennomsnittet.

Usikkerheten til et gjennomsnittsestimat uttrykkes ved hjelp av standardavviket til gjennomsnittet (Jelle, 2022). Formelen er gitt i Formel 2.

Det har vært ønskelig å kunne si noe om hvordan antall studenter påvirker usikkerheten til gjennomsnittsestimatet. Til dette har det blitt laget et Python-skript. Python-skriptet beregner standardavviket for tilfeldig sammensatte flokker med ulik størrelse.

Datasettet til Flokken av alle studentene var delt i to. Den ene delen av datasettet inneholdt studentenes estimat gitt i runde 1 og den andre delen av datasettet inneholdt studentenes estimat gitt i runde 2. Ingen av studentene ga flere enn ett estimat i hver av rundene. Derfor var det én student bak hvert estimat i hver av de to delene av datasettet.

Python-skriptet trakk tilfeldig ut estimat fra den samme delen av datasettet og beregner så gjennomsnittet for de tilfeldige estimatene. Deretter beregnes standardavviket av gjennomsnittet til de tilfeldige estimatene. Denne prosessen utføres av Python-skriptet én gang for hver flokkstørrelse. Det presiseres at ved å trekke ut verdier fra «den samme delen av datasettet» så kan ikke estimat fra runde 1 kombineres med estimat fra runde 2. Denne beskrevne delen av python-skriptet utførte beregningene én gang.

For å finne ut om det var en trend for hvordan standardavviket utviklet seg for ulike flokkstørrelser ble python-skriptet utvidet til å utføre de nevnte beregningene 1 million ganger. På samme vis som tidligere trakk Python-skriptet ut estimat på tilfeldig vis fra datasettet med estimatene til studentene. Deretter ble det for estimatene som ble trukket ut beregnet et gjennomsnitt. Prosessen med å trekke ut estimat og beregne gjennomsnittet av estimatene ble utført 1 million ganger. Følgelig endte denne prosessen opp med å frembringe 1 million ulike gjennomsnitt for hver flokkstørrelse. Deretter tok Python-skriptet gjennomsnittet av de 1 million ulike gjennomsnittene. Til slutt ble det beregnet ett standardavvik for gjennomsnittet av de 1 million ulike gjennomsnittene, for hver flokkstørrelse.

Resultatet til alle beregningene er illustrert ved hjelp av Python-plots. Hele Python-koden er gitt i Bilag 3.

## 3 Teoretisk rammeverk

Teorikapittelet presenterer viktige aspekter ved fenomenet wisdom of crowds og byggeprosjekt. Teorien i dette kapittelet danner rammeverket for den senere diskusjonen. Rammeverket har også vært med å forme hvordan eksperimentet har blitt utført. Kapittelet er å anse som originalt, og omtrent hele rammeverket er et resultat av det nye litteraturstudiet. Underkapittel 3.1.2 er delvis basert på referanser funnet i prosjektoppgaven.

### 3.1 Wisdom of crowds

For å forstå hvordan fenomenet wisdom of crowds kan brukes i dagens byggebransje er det nødvendig å ha en forståelse av hva fenomenet faktisk er. Derfor vil dette delkapittelet beskrive fenomenet nærmere.

#### *Definisjon*

Wisdom of crowds har ingen entydig definisjon. Wisdom of crowds kan sies å være et begrep som viser til tanken om at en flokk menneskers samlede forslag kan være bedre enn enkeltindivids individuelle forslag (Da & Huang, 2020; Fiechter & Kornell, 2021; Gecer Ulu et al., 2019; Hill & Ready-Campbell, 2011; Keuschnigg & Ganser, 2017; Simmons et al., 2011; Yi et al., 2012). Et samlet forslag er i sammenheng med masteroppgaven et statistisk oppsummeringstall av flokkens estimat. I masteroppgaven uttrykkes flokkens statistiske oppsummeringstall som et aritmetisk gjennomsnitt og median.

#### 3.1.1 Oppgavetyper

Fenomenet wisdom of crowds har vist seg å være gjeldende for flere oppgaver hvor det er ønskelig å komme frem til et punktestimat (Yi et al., 2012). De fleste av oppgavene har vært oppgaver hvor det ikke har vært nødvendig å inneha noen nødvendig kompetanse for å gi et svar. Eksempler på dette er å gjette vekten av en okse, temperaturen i et klasserom, spå en fremtidig aksjeverdi og som nevnt i innledningen estimere prisen for poster i anbudsregning (Galton, 1907; Hill & Ready-Campbell, 2011; Simmons et al., 2011). Oppgavene kjennetegnes av at de har forhåndsdefinerte svar med ett klart svar. Slike oppgaver defineres som «velstrukturerte oppgaver» (well structured tasks) (Graf-Drasch et al., 2022). Dersom oppgaven ikke har et forhåndsdefinert og klart svar defineres den av Graf-Drasch et al. (2022) som en «ustrukturert oppgave» (unstructured tasks). Graf-Drasch et al. (2022) gjennomførte et litteratursøk hvor det ble identifisert 21 forskjellige studier som omhandlet bruken av kollektiv visdom. Videre ble spørsmålene stilt i de 21 studiene klassifisert etter definisjonen for strukturerte og ustrukturerte oppgaver. Ved å analysere resultatene til deltakerne innenfor de to oppgavetyperne ble det vist at fenomenet har hatt størst suksess for velstrukturerte oppgaver. Suksess innebærer her at de kollektive estimatene har hatt høy grad av presisjon og gjort det bedre enn de fleste enkeltstående estimatene. Det ble konkludert med at effekten av kollektiv wisdom bare har vært synlig for velstrukturerte oppgaver, og ikke for ustrukturerte oppgaver.

En av årsakene til at effekten av kollektiv visdom har vært synlig for kun velstrukturerte oppgaver er at svarene for ustrukturerte oppgaver ikke er diskret (Gecer Ulu et al., 2019). Svarene til en ustrukturert oppgave vil være åpen for en kvalitativ vurdering. Det er verdt

å merke seg at både velstrukturerte og ustrukturerte oppgaver kan ha ulike krav til kunnskap og ferdigheter hos deltakere. I velstrukturerte oppgaver vil det ofte være lite avhengighet av deltakernes ekspertise på et spesifikt felt, mens ustrukturerte oppgaver ofte vil kreve en viss grad av domenekunnskap. Dermed kan det være forskjellige faktorer som påvirker kvaliteten på wisdom of the crowd-resultatene i ulike typer oppgaver (Graf-Drasch et al., 2022).

Litteraturen har vist at effekten av wisdom of crowds kan ses i esoteriske problem så vel som i eksoteriske problem (Gecer Ulu et al., 2019). Et esoterisk problem er et problem som er vanskelig å forstå for de fleste mennesker, og det krever vanligvis spesialisert kunnskap eller erfaring for å kunne bli fullt ut forstått. Et eksoterisk problem er det motsatte (Nilstun, 2018). Å gjette vekten av en okse er et eksempel på et eksoterisk problem. Å gjette prisen på et byggeprosjekt er et eksempel på en esoterisk oppgave.

### 3.1.2 Flokkens sammensetning

Sammensetningen av flokken påvirker presisjonen til wisdom of crowds (Budescu & Chen, 2015; Galesic et al., 2018; Gecer Ulu et al., 2019; Goldstein et al., 2014; Hong et al., 2020; Keuschnigg & Ganser, 2017; Larrick, 2016; Mannes et al., 2014; Nguyen & Nguyen, 2017). Litteraturen har pekt på at flokkens størrelse og mangfold er faktorer som avgjør om flokken gir gode estimat. Sammensetning av flokken påvirker også ressursbruken forbundet med å få frem fenomenet. Flokken i wisdom of crowds har ofte kun inneholdt lekfolk. Likevel har ikke litteraturen sagt noe om at flokken ikke kan inneholde eksperter.

#### *Mangfold*

Begrepet mangfold har innenfor litteraturen om wisdom of crowds ingen klar definisjon. Litteraturen har vist at mangfold er et begrep som betegner variasjon i deltakernes kompetanse, meninger, innsikt og bakgrunn (Hong et al., 2020; Nguyen & Nguyen, 2017). Hong et al. (2020) har sett på hvordan mangfold av erfaringer påvirker presisjonen til flokken som brukes i wisdom of crowds. Studien tok for seg flokkens evne til å spå fremtidig aksjeverdier for S&P 500 selskaper. Omfanget av begrepet erfaring var i dette tilfellet hvor mange industrier deltakerne hadde gitt estimat for, og hvor mange estimat de hadde gitt innenfor hver industri. Et investeringsforum på internett kalt StockTwits ble brukt for å samle inn data. I denne studien ble mangfoldet av erfaringer brukt som en variabel for å undersøke hvordan det påvirker presisjonen til flokken. En tendens til at økt mangfold førte til bedre presisjon ble funnet.

Nguyen & Nguyen (2017) har sett på hvordan en endring av mangfold påvirker kvaliteten til kollektive estimat. Mangfold er i studien sagt å være ulikheten av deltakernes bakgrunn eller meninger om problemet som skal løses. I studien uttrykkes ulik bakgrunn og ulike meninger i form av forskjell mellom estimat. Forskningen skulle se på konsekvensene av at de som hadde de mest avvikende estimatene fra gjennomsnittsestimatet endret sine estimat. Prosessen med at deltakerne endret sine estimat ble gjentatt helt til avvikene var innenfor et predefinert område. Effekten av å fjerne de største avvikene ble også studert. Resultatet av eksperimentet viste at både det å endre de største avvikene og fjerne de største avvikene ikke ga bedre samlede estimat. Kvaliteten på estimatene gikk ned med et redusert mangfold.

Betydningen av mangfold har blitt målt opp mot kompetanse hos deltakerne (Keuschnigg & Ganser, 2017). Keuschnigg & Ganser (2017) har undersøkt om det å tilføre kompetanse eller mangfold er mer effektivt for å oppnå høyere presisjon i de samlede forslagene fra en flokk. Begrepet kompetanse er i studien brukt for å omhandle personer som gir svar som

er nærmest det korrekte svaret. Mangfold er i studien satt til å være personer som bidrar til en spredning av estimatene i datasettet. Oppgavene dreide seg om å gi tallestimat for ulike oppgaver, for eksempel å gjette antallet drops i en krukke. Resultatene fra studien viste at presisjonen til det samlede estimatet økte mer for en økning av mangfold blant deltakerne, enn det gjorde for en økning av kompetanse hos deltakerne. Studien viser betydningen av å ha et variert utvalg av perspektiver og erfaringer i en flokk når det er ønskelig å få frem presise estimat med wisdom of crowds.

### *Størrelse*

Litteraturen om wisdom of crowds har ikke gitt en definisjon for hvor mange bidragsyttere det må være for å kunne si at det er en flokk. Den første vitenskapelige observasjonen av fenomenet hadde 787 bidragsyttere, mens effekten også er funnet for flokker med bare 4 bidragsyttere (Galton, 1907; Keuschnigg & Ganser, 2017). Gecer Ulu et al. (2019) har omtalt flokker med 5-14 personer som «mikroflokker» (microcrowds). Litteraturen peker dermed i den retning at selv grupper med et lite antall bidragsyttere kan karakteriseres som flokker. Etter mitt syn ser det ut til at det er arbeidsmetodikken til gruppene som definerer bruken av ordet flokk.

Forskning har undersøkt hvilke resultat små flokker kan gi, og hvordan dette måler seg mot større flokker (Gecer Ulu et al., 2019; Mannes et al., 2014). Mannes et al. (2014) har sett på en strategi der det brukes selektive flokker. Selektive flokker består av utvalgte medlemmer av den originale flokken. Medlemmene som er valgt ut er de som anses å være de mest kvalifiserte i den opprinnelige flokken. Hvem som blir ansett som best er avhengig av ulike indikatorer. Studien har brukt historiske prestasjoner som indikator for å velge ut de beste medlemmene. Det fremholdes at andre indikatorer også kunne vært brukt, eksempelvis selvtillit, status og utdanning. Studien har vist at utvalgte flokker på fem personer har gitt gode resultat for velstrukturerte oppgaver hvor det søkes etter et kvantitativt svar. Resultatene til den selektive gruppen har for en signifikant del av oppgavene vært bedre enn både enkeltstående eksperter og hele flokken samlet.

Gecer Ulu et al. (2019) har sett på effekten av wisdom of crowds for flokker med ulik størrelse og kompetanse. Problemene som skulle løses var esoteriske ingeniørproblem. Studien brukte to forskjellige flokker. Den ene bestod av 100 tilfeldige mennesker fra et forum på internett, mens den andre bestod av 15 maskinstudenter. Amazon Mechanical Turk (AMT), CrowdFlower og Prolific Academic ble dratt fram som mulige forum å skaffe deltakere fra, men bare Amazon Mechanical Turk ble brukt i studien. Maskinstudentene representerer en gruppe mennesker med kjennskap og kunnskap til problemene som skal løses. Gruppens medlemmer er ikke nødvendigvis eksperter på området, men de kan sies å inneha basisferdigheter på området. En ekspert vil være en undergruppe av de som innehar basisferdigheter, og en med ekstra mye kompetanse på området. Flokken bestående av tilfeldige personer kan sies å representere den øvrige befolkningen. Deltakerne i studien skulle alle besvare fire spørsmål relatert til 3D-printing. Det første spørsmålet handlet om å bedømme hvor mye støttemateriale ulike objekt ville kreve, de to neste handlet om hvordan støttematerialet ville påvirke overflaten av ulike objekt og det siste handlet om hvor sterke objektene kom til å være. For de tre første spørsmålene skulle svarene gis på en skala fra 1-10 mens det siste spørsmålet hadde en skala fra 1-5. I det første spørsmålet tilsvarte 1 lite støttemateriale og 10 mye støttemateriale. I de to neste tilsa 1 at materialet ble svært påvirket og 10 nesten ikke påvirket. I det siste spørsmålet var 1 svak og 5 sterk. De tre første spørsmålene kan besvares ved hjelp av areal og volumbetraktninger, mens det siste spørsmålet krever en viss forståelse av



hvordan krefter virker på et objekt. Areal og volumberegninger er noe tilnærmet alle har vært ute for, mens det å vurdere styrken til et objekt ikke kan sies å være en del av hverdagen til folk flest. Følgelig kan det sies at spørsmålene hadde en økende vanskelighetsgrad.

Resultatene fra forsøket til Gecer Ulu et al. (2019) viste at de samlede forslagene for flokken bestående av den øvrige befolkningen ikke var i nærheten av de korrekte svarene. For flokken bestående av tilfeldige folk ble det heller ikke observert en effekt av wisdom of crowds. Flokken bestående av mennesker med basisferdigheter oppnådde derimot gode resultater. Det ble observert at antall deltakere som gjorde det bedre enn flokkens samlede estimat gikk ned da antallet deltakere i flokken gikk opp. Observasjonen viser en effekt av wisdom of crowds. Resultatet peker i den retning at for esoteriske problemer vil det være nødvendig at deltakerne innehar en viss kompetanse. For å studere effekten fenomenet hadde for mikroflokker ble en undergruppe med 500 tilfeldig genererte kombinasjoner av 5–14 individer skapt fra det opprinnelige settet med 15 deltakere. Selv for de minste flokkene ble effekten av wisdom of crowds funnet. Effekten ble styrket med flere deltakere.

### 3.1.3 Bias som kan påvirke flokkens deltakere

Bias kan være med på å svekke presisjonen til estimatene. Innenfor tematikken om wisdom of crowds og estimering er bias knyttet til personers atferd sentral. Flyvbjerg (2021) sier at atferdsbias består av både kognitive bias og politiske bias. Innenfor prosjektledelse finnes det en rekke ulike atferdsbias, hvor hver av biasene har sin egen definisjon. Flere av biasene kan påvirke estimatene til medlemmene av flokken. Derfor er det viktig å ha kjennskap til noen av biasene deltakerne kan inneha. Grunnet kapasitetsbegrensninger fra min side har det vært nødvendig å avgrense hvor mange bias som skal presenteres i denne masteroppgaven. I oppgaven er forankringseffekt, flokkmentalitet, underestimering, optimisme, planleggingsfeilen, søkelys på det unike og baseratefeil undersøkt nærmere (Belle et al., 2017; Flyvbjerg, 2021; Görzen, 2019; Jayles et al., 2020; Simmons et al., 2011; Stanovich & West, 2008). Etter mitt syn er disse de mest kjente biasene relatert til wisdom of crowds. Biasenes påvirkning på eksperter er også undersøkt.

#### *Forankring*

Begrepet forankringseffekt viser til den kognitive tendensen mennesker har til å basere nye tallestimat på en gammel verdi. (Belle et al., 2017; Görzen, 2019; Stanovich & West, 2008; Tversky & Kahneman, 1974). Den gamle verdien fungerer som et utgangspunkt for nye estimat. Utgangspunktet refereres til som et anker. Studier har vist at referanseverdien ikke behøver å være relevant for hva som skal estimeres. Tversky & Kahneman (1974) gjennomførte et eksperiment hvor deltakerne først skulle spinne et lykkehjul med tall mellom 1 og 100, for så å gjette hvor mange afrikanske land som var medlem av FN. Eksperimentet viste at deltakerne som fikk lave tall på lykkehjulet også ga lave gjett for antall afrikanske land i FN.

#### *Underestimering, optimisme og planleggingsfeilen*

Underestimering viser til den tendensen mennesker har for å underestimere omfanget eller størrelsen av store mengder eller tall (Jayles et al., 2020, 2021). Tendensen har vist seg for flere estimeringsoppgaver. Innenfor bygg- og anleggsbransjen har underestimering blitt dratt frem som en årsak til kostnadsoverskridelser (Flyvbjerg et al., 2002). Flyvbjerg et al. (2002) undersøkte hvordan kostnadsestimatene for offentlige samferdselsprosjekt holdt opp mot de faktiske kostnadene. En klar tendens til at de faktiske kostnadene ble

høyere enn de estimerte kostnadene ble funnet. Fire forklaringer for underestimering av kostnader ble holdt frem: teknisk, økonomisk, psykologisk og politisk. Flyvbjerg et al. (2002) diskuterte videre hvilke av forklaringene som samsvarte best med datamaterialet de hadde. En sammenfatning av forklaringene og betydningen av dem er gitt i Tabell 2.

Tabell 2: Viser forklaringer for underestimering i byggeprosjekt, samt betydningen av forklaringene. Tabellen er egenprodusert.

<b>Underestimeringsvariant</b>	<b>Årsak til underestimering og betydning</b>
Teknisk forklaring	At underestimeringen skyldes prognosefeil (Flyvbjerg et al., 2002). Dette innebærer blant annet at det er manglende informasjon, vanskelig å spå fremtiden, ærlige feil og manglende erfaring hos de som kommer med estimatene. Forklaringen samsvarte ikke med datamaterialet Flyvbjerg et al. (2002) hadde.00.00.0000 00:00:00
Økonomisk forklaring	To økonomiske forklaringer for underestimering finnes. Disse er økonomisk egeninteresse, og av hensyn til allmenne interesser (Flyvbjerg et al., 2002). Økonomisk egeninteresse er knyttet til at interessentene i prosjektet underestimerer kostnadene og overdriver nytten. Dette gjøres for å øke sannsynligheten for at prosjektet skal iverksettes. Den andre økonomiske forklaringen til underestimering er at de som lyser ut prosjektet bevisst underestimerer omfanget av kostnadene for prosjektet. Årsaken er at de mener at de dermed unngår entreprenører som sløser med skattebetalernes ressurser, ettersom de tvinges til å operere med lave priser. Begge forklaringene samsvarte med datamaterialet Flyvbjerg et al. (2002) hadde.
Psykologisk forklaring	Psykologiske forklaringer bruker bias i menneskers tankesett som en forklaring på underestimering (Flyvbjerg et al., 2002). Vurderingsoptimisme hos mennesker trekkes frem som den viktigste årsaken til at slike vurderinger forekommer. En psykologisk forklaring holdes frem som en mulig forklaring på underestimeringen. Flyvbjerg et al. (2002) mener forklaringen hadde passet best dersom det var uerfarne personer som estimerte prisen for prosjektene. Årsaken til dette er at disse personene ikke er kjent med flere av realitetene i byggeprosjekt. Datamaterialet strekker seg dog over en 70-årsperiode, og Flyvbjerg et al. (2002) anser det som usannsynlig de som estimerer kostnadene ikke ville ha lært av tidligere prosjekt.
Politisk forklaring	At politikere bevisst underestimerer kostnader slik at de får gjennomført prosjekt som sammenfaller med deres interesser og samtidig kan styrke deres maktposisjon. Flyvbjerg et al. (2002) mener at denne forklaringen samsvarer godt med store deler av tallmaterialet.

Flyvbjerg et al. (2002) konkluderer med at en blanding av politisk forklaring og økonomisk forklaring ligger bak de fleste tilfellene av underestimering i offentlige samferdselsprosjekt. En systematisk underestimering fører til to separate problem (Samset, 2017). Samset (2017) sier at for lave kostnadsestimat kan føre til problemer med å kontrollere kostnadsoverskridelser i gjennomføringen av prosjekt, samt at strategisk underestimering i tidligfasen påvirker selve valget av prosjektet. Strategisk underestimering i tidligfasen trekkes frem som det største problemet. Årsaken til at strategisk underestimering i tidligfasen er det største problemet er ifølge Samset (2017) at et bedre, det vil si mer korrekt, kostnadsestimat i tidligfasen trolig vil føre til at færre dårlige prosjekt velges. Nytten av investeringene vil dermed bli bedre.

I likhet med Flyvbjerg et al. (2002) mener Samset (2017) at det er en tendens til systematisk feilestimering av prosjektets nytte. Samset (2017) skriver at det som regel er en overestimering av nytten. Det påpekes dog at det ofte er vanskeligere å forutsi nytte enn kostnad av prosjekt. Dette kommer av at effekten av prosjektet kun kan måles en tid etter det er overlevert.

Samlet sett fører en overvurdering av nytte og en undervurdering kostnad til en for stor nytte-/kostnadsbrøk. En nytte-/kostnadsbrøk er et mye brukt mål for å måle om prosjekter er gode (Samset, 2017). I de tilfellene hvor det er mulig å bruke kronebeløp er et prosjekt ikke lønnsomt dersom brøken er mindre enn én.

Det finnes også litteratur som mener at det ikke er grunn til å tro at systematisk underestimering og vurderingsoptimisme er noen av årsakene til kostnadsoverskridelser i bygg- og anleggsbransjen. Love & Ahiaga-Dagbui (2018) går langt i å kritisere Flyvbjerg et al. (2002). Love & Ahiaga-Dagbui (2018) mener at datagrunnlaget til Flyvbjerg et al. (2002) har vært for lite til å kunne si at det har statistisk signifikants. Love & Ahiaga-Dagbui (2018) mener videre at Flyvbjerg et al. (2002) har tatt for lite hensyn til endring i prosjektomfang. Det konkluderes med at det ikke finnes bevis for at systematisk villfarelse av estimat er en av årsakene for kostnadsunderestimering (Love & Ahiaga-Dagbui, 2018).

En undergruppe av underestimering og optimisme er planleggingsfeilen (planning fallacy) (Flyvbjerg, 2021). Begrepet viser til de tilfellene hvor hendelser i for stor grad planlegges ut ifra at det best tenkelige scenarioet finner sted. I korte trekk vil det si at gjennomføringstid, arbeidsomfang, risiko og kostnader undervurderes, mens nytten og mulighetene for handlingene overvurderes.

#### *Ensidig søkelys på det unike (uniquess bias)*

Biasen sikter til at planleggere og prosjektledere ser på prosjekt som mer unike enn hva de faktisk er (Flyvbjerg, 2021). Flyvbjerg (2021) gjengir flere definisjoner av prosjekt, hvor fellestrekket til dem er at prosjekt blir ansett som en unik innsats for å realisere et mål. Derimot mener Flyvbjerg (2021) på sin side at en slik definisjon er uheldig, ettersom den bidrar til å rette søkelyset på det unike i prosjekt, fremfor å se likhetene mellom prosjektene. Uten å se likheten mellom prosjektene blir det også vanskelig å lære av tidligere feil og løsninger.

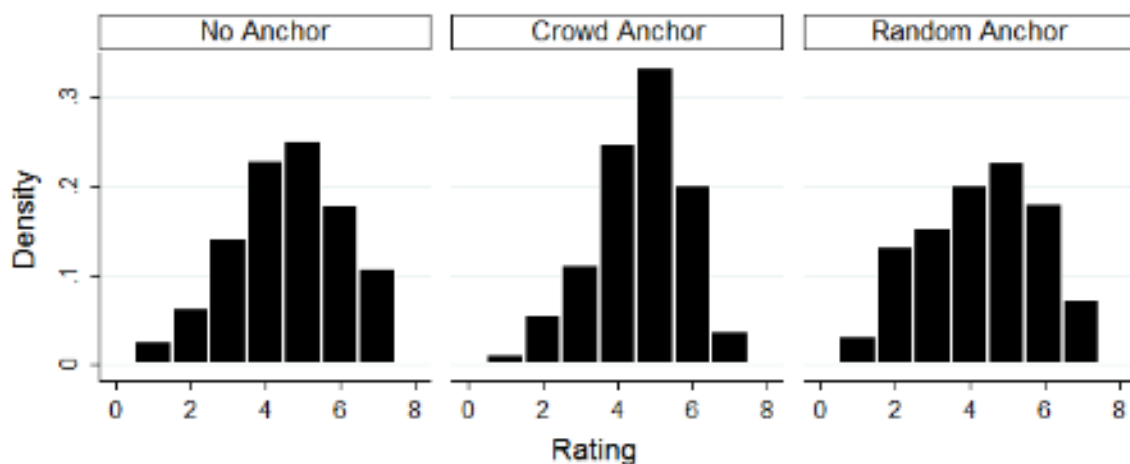
#### *Basisratefeil*

Basisratefeil viser til tendensen folk har til å henge seg opp i detaljer og personlig informasjon, på bekostning av å se de store linjene (Flyvbjerg, 2021).

### *Påvirkning av bias på eksperter*

Görtzen (2019) har vist at det finnes både forskning der det konkluderes med at eksperter blir påvirket av bias, og forskning der det konkluderes med at eksperter i liten grad påvirkes av bias. Görtzen (2019) har studert hvordan eksperter påvirkes av forankring. I studien ble det brukt deltakere fra det offentlig tilgjengelige forumet CrowdFlower. Deltakerne skulle bedømme ulike forretningsmodeller for parfymmer. Forretningsmodellene skulle av deltakerne få en karakter mellom 1 og 7 for ulike kategorier. Kategoriene som skulle bedømmes var kreativitet, nyskapning og nytte. For å studere forankring ble to forskjellige ankre introdusert. Et av ankrene var gjennomsnittet til de som tidligere hadde gitt en karakter, og det andre ankeret var en tilfeldig verdi mellom 1 og 7. Når deltakerne skulle gi svar hadde skalaen enten ett av ankrene vist over seg eller ingenting. Ekspertene ble identifisert ved at deltakerne selv svarte på spørsmål om deres kompetanse. Deltakerne ble stilt spørsmål om hvor lenge de hadde vært på CrowdFlower, hvor lenge de hadde jobbet med detaljhandel, spørsmål om parfymemerker og erfaring med forretningsmodeller. For å gi en fasit for hva som var den mest riktige bedømmelsen av de ulike forretningsmodellene ble to eksterne eksperter rekruttert. En av ekspertene hadde erfaring fra entreprenørseksjonen til et universitet, mens den andre jobbet som seniorkonsulent for et stort firma.

Resultatene viste at spredningen til estimatene hos deltakerne ble påvirket av ankrene (Görzen, 2019). Fordelingen av estimatene viste at de som hadde blitt gitt ankeret som viste gjennomsnittsverdien til de andre deltakerne hadde en høyere konsentrasjon rundt enkelte verdier, sammenlignet med de som ikke hadde blitt vist noe anker. Deltakerne som hadde blitt vist et tilfeldig anker kom med svar som var mer spredt sammenlignet med de som ikke fikk se et anker. En visualisering av resultatene kan ses i Figur 1.



Figur 1: Viser fordelingen av svar for deltakere som fikk se anker, samt de som ikke fikk se et anker. Figuren viser spredningen av svarene (Görzen, 2019).

Gjennom å studere resultatene fra eksperimentet ble det funnet at også ekspertene ble påvirket av ankrene (Görzen, 2019). Studien sier videre at ekspertise derfor ikke alene kan motvirke forankring.

#### 3.1.4 Arbeidsmetodikker for å gi estimat om fremtiden

Wisdom of crowds bygger på en annen arbeidsmetodikk enn tradisjonelt gruppearbeid. Mens tradisjonelt gruppearbeid involverer en liten gruppe mennesker som samarbeider for

å løse et problem, baserer wisdom of crowds seg på bidrag fra en flokk av individuelle medlemmer. Bidragene til flokken legges sammen ved hjelp av et statistisk oppsummeringstall. En arbeidsmetodikk som kan minne om den som brukes for wisdom of crowds er Delphimetoden. Delphimetoden vil bli presentert senere i dette delkapittelet. Alle de nevnte arbeidsmetodene har til hensikt å si noe om en usikker framtidssituasjon. Dette underkapittelet skal se nærmere på hvordan flokkens arbeidsmetodikk i wisdom of crowds skiller seg fra andre samarbeidsformer, samt hvilke teknikker det finnes for å legge sammen ulike bidrag.

### *Tradisjonelt gruppearbeid*

Tradisjonelt gruppearbeid kjennetegnes av at gruppemedlemmene samarbeider om å komme frem til en samlet løsning på et problem (Minson et al., 2018). Samarbeidet skjer gjerne ved at gruppemedlemmene presenterer sin løsning for resten av gruppen, før den deretter blir diskutert. En slik samarbeidsform er omdiskutert. Minson & Mueller (2012) utførte en studie der de undersøkte hvor mye informasjon som ble tatt opp av individer og grupper. Studien var utformet slik at deltakerne enten samarbeidet to og to eller jobbet alene for å finne en løsning. Deretter fikk deltakerne forslag fra den andre parten, enten som individ eller som gruppe. Resultatene viste at personene som samarbeidet var mer tilbøyelige for å ikke ta til seg ny informasjon, og at samarbeidet dermed førte til redusert informasjonsopptak. Resultatet av dette var at gruppens reviderte forslag var dårligere enn om de hadde valgt å ta til seg den nye informasjonen. Eksempelet viser til en situasjon hvor tradisjonelt gruppearbeid har prestert dårlig. Gruppens selvtillit til egne forslag ble holdt fram som en forklaring på hvorfor gruppen valgte å ikke ta til seg ny informasjon. Schultze et al. (2012) har på sin side vist at medlemmer av grupper kan øke presisjonen til sine forslag gjennom å diskutere løsningene i plenum. Likevel fremheves det av Schultze et al. (2012) at selv om gruppemedlemmene individuelt sett forbedret sine forslag, var gruppens endelige forslag ikke bedre enn gjennomsnittet av de individuelle estimatene. Eksempelet viser at tradisjonelt gruppearbeid kan oppnå gode resultat, men at en alternativ arbeidsmåte kan gi bedre resultat.

Et skille i den presenterte samarbeidsformen og wisdom of crowds er at flokken som inngår i wisdom of crowds kjennetegnes av at deltakerne kommer med individuelle forslag og at disse legges sammen (Da & Huang, 2020; Kao et al., 2018; Keuschnigg & Ganser, 2017; Lorenz et al., 2011; Minson et al., 2018). Flokken i wisdom of crowds inkluderer ofte kun lekfolk. I gruppearbeid er som regel de involverte ansett å være eksperter. Budescu & Chen (2015) sier at forskjellen mellom å bruke en gruppe av eksperter og en flokk lekfolk kan betegnes som forskjellen på å basere seg på kvalitet (ekspertise) og kvantitet (flokker). Likevel har ikke litteraturen sagt noe om at flokken ikke kan inneholde eksperter.

### *Store talls lov, uavhengighet og flokkmentalitet*

En teoretisk forklaring for hvorfor wisdom of the crowds kan være suksessfullt involverer store talls lov (Kao et al., 2018; Keuschnigg & Ganser, 2017). Store talls lov er et matematisk prinsipp som sier at når man tar gjennomsnittet av en stor mengde observasjoner, vil gjennomsnittet i en viss forstand nærme seg sin forventningsverdi når antallet observasjoner vokser (Frøslie, 2022). For kvantitative oppgaver er ideen at eventuelle overestimer og underestimer vil motvirke hverandre. Ved å kombinere et bredt spekter av estimer fra forskjellige individer, vil feilene for hvert estimat avbøtes og det faktiske svaret vil sannsynligvis ligge et sted imellom (Larrick & Soll, 2006). Flokkens medlemmer gir individuelle estimat uten å diskutere seg fram til en samlet løsning. Kao et

al. (2018) skriver at dersom flokkens medlemmer utveksler informasjon vil effekten av fenomenet reduseres. Feilen omtales som sosial påvirkning, og kan føre til flokkmentalitet. Begrepet flokkmentalitet beskriver tendensen til å kopiere hverandre (Jayles et al., 2021). Gjennom samtaler og diskusjoner vil svarene til flokkens medlemmer til en viss grad være korrelerte (Da & Huang, 2020; Kao et al., 2018). Dette kommer av at deltakere kan observere hva andre ytrer, og basere egne estimat på hva som allerede er sagt eller gjort (Da & Huang, 2020). Deltakerne påvirkes i så stor grad at de velger å se bort ifra personlig informasjon. Flokkmentalitet fører til at svarene blir mer sentrert rundt et punkt, og at distribusjonen av svarene blir snevrere.

Minson et al. (2018) har også studert hvordan diskusjon påvirker presisjonen til samlede estimat. Det er sett på hvordan diskusjon kan påvirke resultatet for grupper bestående av relativt få deltakere. Grupper med få deltakere er eksempelvis et ledelsesteam eller partnerskap. Studien har brukt grupper bestående av to personer i sine forsøk. Det er undersøkt om diskusjon kan motvirke grove feil i medlemmenes estimat. Det ble gitt estimat i to runder. I den første runden ble det gitt individuelle estimat og i den andre runden ble gruppemedlemmene enig om et samlet estimat. Diskusjonen fant sted etter at medlemmene hadde gitt individuelle estimater. To forskjellige typer spørsmål ble studert. I den ene situasjonen var det ingen begrensninger for hvilke svar deltakerne kunne gi, mens det i den andre forelå en naturlig begrensning. Spørsmålet «Hva er årslønnen til en daglig leder for et Fortune 500 selskap?» er et eksempel på et ikke begrensende spørsmål brukt i studien. Et spørsmål som inneholdt en begrensning, var «Hvor stor prosentandel av innbyggerne i Philadelphia har ikke helseforsikring?». Ettersom det i det første spørsmålet ikke foreligger en begrensning på hvilke svar en deltaker kan gi, finnes det heller ikke en begrensning på hvor langt unna det riktige svaret deltakeren kan komme. Dette åpner opp for at datagrunnlaget kan få ekstreme svar som kan påvirke det samlede gjennomsnittet. I det andre spørsmålet var svarene begrenset til å befinne seg en plass mellom 0 % og 100 %. Studien viste at diskusjon før runde to hindret deltakerne fra å gi ekstreme svar i de tilfellene hvor det ikke forelå begrensninger. I de tilfellene hvor grove feil ble eliminert ble gruppens samlede svar bedre enn gjennomsnittet av de individuelle svarene. I tilfellene hvor det var begrensninger var gruppens samlede estimat etter diskusjon dårligere enn gjennomsnittet av de individuelle svarene.

### *Delphimetoden*

Delphimetoden er en arbeidsmetodikk hvor samarbeid mellom eksperter kombineres med elementer av wisdom of crowds. Metoden ble utviklet av Norman Dalkey og Olaf Helmer på 1950-tallet (Dalkey & Helmer, 1963; Gürçay et al., 2015). Hensikten er ifølge Dalkey & Helmer (1963): «å hente inn den mest pålitelige konsensusen av meninger for en gruppe eksperter» (min oversettelse). Arbeidsmetodikken går ut på at eksperter gjennom flere runder blir spurt om å gi sine løsninger på et problem (Gürçay et al., 2015). Svarene blir hentet inn individuelt og anonymt av en moderator. Mellom rundene får ekspertene tilgang til ny informasjon (Dalkey & Helmer, 1963). Informasjonen det dreier seg om er primært data etterspurt av andre eksperter eller om hensyn ansett som relevant av de andre deltakerne. For å sikre anonymitet blir informasjonen gitt til deltakerne uten at de får vite hvem som kom med hvilke betraktninger. Mellom hver runde får deltakerne anledning til å endre sine svar. Prosessen blir gjentatt helt til en samlet konsensus er nådd. Etter den siste runden regnes et statistisk gjennomsnitt ut, som regel gjennomsnitt eller median.

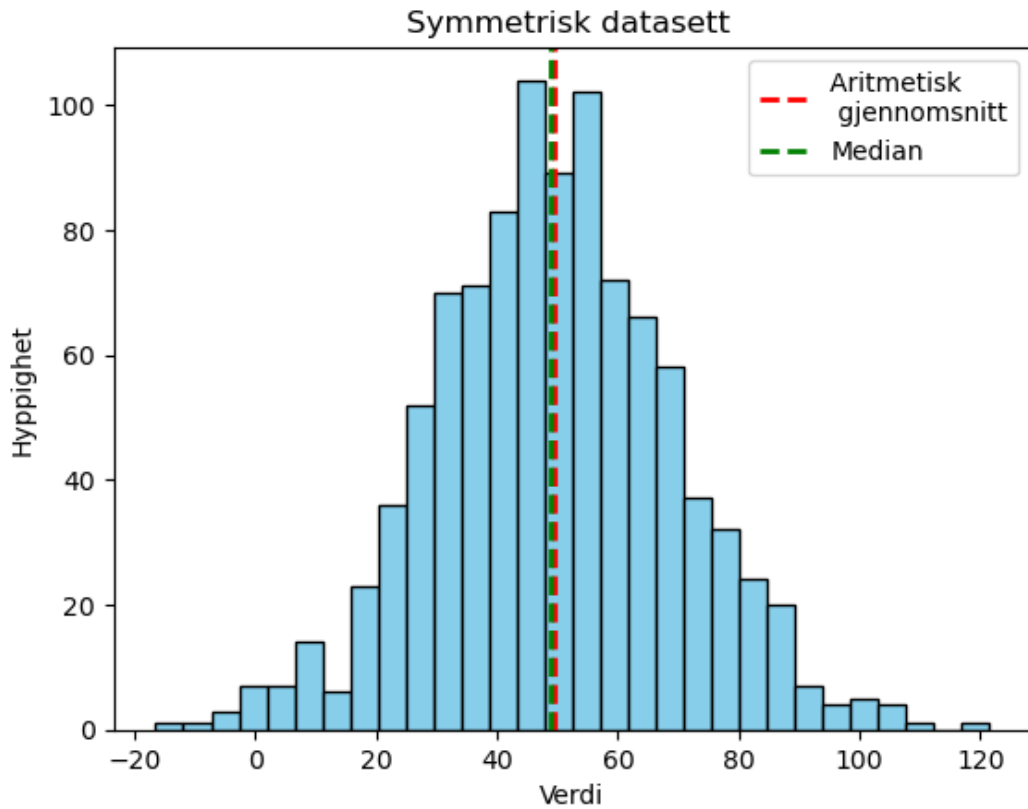
I likhet med wisdom of crowds har Delphimetoden vist seg å gi resultater med høyere presisjon enn tradisjonelt gruppearbeid for velstrukturerte oppgaver (Rowe & Wright, 1999). Rowe & Wright (1999) har sammenlignet en rekke studier hvor Delphimetoden har blitt benyttet. Gjennom å analysere studiene er det kommet frem at Delphimetoden har slått tradisjonelt gruppearbeid i fem studier, gjort det dårligere i én og like godt i to. Metoden har ført til forbedring av estimatene gitt i første runde i 12 studier, svekket resultatene i to og gjorde det like godt i to. Det påpekes likevel at de ulike studiene har benyttet ulike varianter av Delphimetoden. En følge av dette er at det har vært vanskelig å lage en generalisering av resultatene.

### *Utregning*

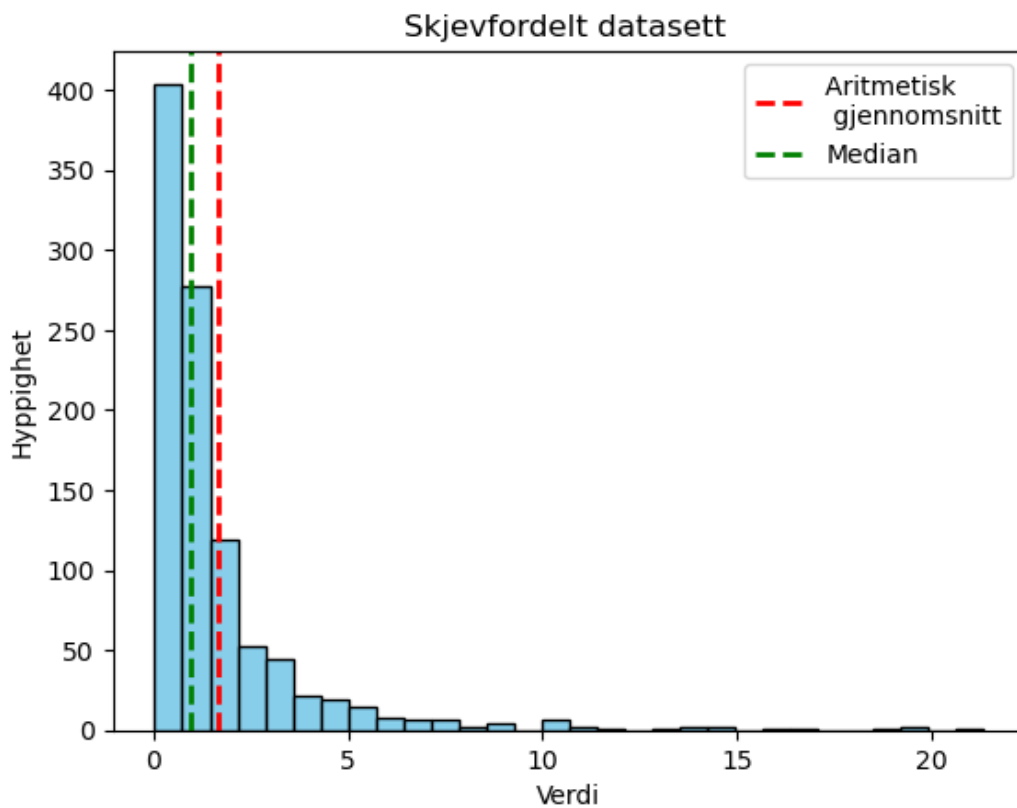
Flokkens samlede estimat uttrykkes ved hjelp av statistiske oppsummeringstall. Vanligvis brukes et aritmetisk gjennomsnitt eller median (Budescu & Chen, 2015; Gecer Ulu et al., 2019; Goldstein et al., 2014; Kao et al., 2018; Mannes et al., 2014; Wallis, 2014). Medianen er den midterste verdien i et tallmateriale (Frøslie, 2023). Medianen finnes ved å sortere alle tallene i stigende rekkefølge og deretter plukke ut tallet midt i rekken. Følgelig vil det være like mange tall over medianen som det er under den. Gjennomsnitt «er et tall som angir den «mest typiske» verdien for en mengde med tall» (Vatne, 2023). Aritmetisk gjennomsnitt, geometrisk gjennomsnitt og harmonisk gjennomsnitt er de tre mest vanlige målene for å beregne gjennomsnittet av en serie tall. Aritmetisk gjennomsnitt er det vanligste gjennomsnittet, og omtales gjerne som bare gjennomsnitt eller aritmetisk middel. For å regne ut et aritmetisk gjennomsnitt summeres alle tallene, og deles på summen av antall tall.

Vatne (2023) skriver at: «God bruk av statistikk handler om å beskrive data så riktig som mulig.», og at bruken av median og gjennomsnitt bør tilpasses datasettet. Der hvor datasettet er symmetrisk betyr valget av oppsummeringstall lite. Både median og gjennomsnitt beskriver de typiske verdiene. For slike sett foretrekkes likevel ofte gjennomsnitt. For skjevfordelte datasett har valget av oppsummeringstall større betydning. Dersom et datasett har en topp til en av sidene, og en lang hale til motsatt side, vil dette forskyve gjennomsnittet til den siden halen befinner seg. Gjennomsnittet vil dermed befinne seg lengre unna de mest typiske verdiene, og samtidig et stykke unna de få høye verdiene. Ettersom medianen har like mange observasjoner på hver side, vil målet befinne seg i senteret av observasjonene. Følgelig er det vanligst å bruke median for skjevfordelte data.

For å illustrere forklaringen gitt over har jeg laget to generelle figurer. Figur 2 viser et symmetrisk datasett og Figur 3 viser et skjevfordelt datasett. At Figur 3 har en topp til venstre, og en hale til høyre, vil si at fordelingen er høyreskjev. I begge figurene er det aritmetiske gjennomsnittet og medianen av datasettene tegnet inn med stiplede linje.



Figur 2: Viser et datasett med en symmetrisk fordeling av verdiene. Det aritmetiske gjennomsnittet og medianen er tegnet inn med stiplede linjer. Figuren er egenprodusert.



Figur 3: Viser et datasett med en skjevfordelt fordeling av verdiene. Det aritmetiske gjennomsnittet og medianen er tegnet inn med stiplede linjer. Figuren er egenprodusert.



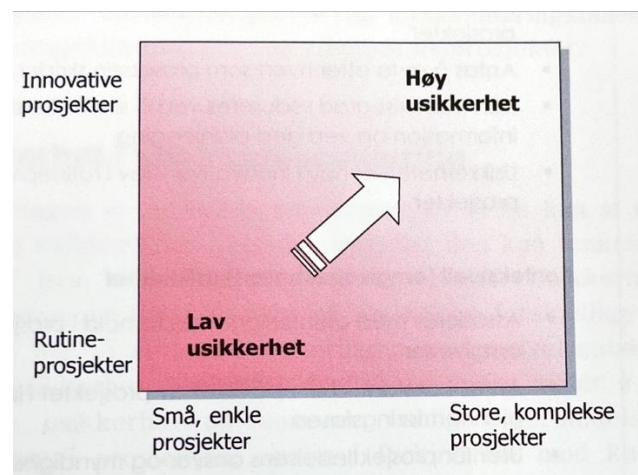
For å gi en mer komplett beskrivelse av datasettet, bør også tall som beskriver variasjonen i datamaterialet inkluderes (Vatne, 2023). Der hvor median benyttes for å beskrive det typiske, bør kvartiler oppgis i tillegg. Dette gjør det mulig å analysere fordelingen av data og forstå hvordan de er fordelt over hele spekteret. For gjennomsnittet bør standardavviket oppgis. Dette gir et mål på spredningen i dataene rundt gjennomsnittet, og usikkerheten knyttet til tallet (Jelle, 2022).

I det originale forsøket av Francis Galton i 1907 ble median benyttet (Galton, 1907). Ifølge Wallis (2014) var Galton selv en forkjemper for å bruke median. Begrunnelsen til Galton var gitt fra et demokratisk ståsted. Galton vektla at én stemme representerte én verdi. Videre mente Galton at det er den midterste verdien i datasettet som uttrykker verdien flertallet av stemmene er enig med. Alle andre verdier enn den midterste verdien vil majoritetene av stemmene anse å være enten for høy eller for lav.

## 3.2 Dagens byggebransje

### 3.2.1 Megaprojekt

Byggebransjen har lenge vært preget av «jernloven», at megaprojekt om og om igjen havner over tidsplanene, over budsjett og under nytte (Flyvbjerg, 2011). I Flyvbjerg (2011) er et megaprojekt definert til å være et prosjekt som koster 100 millioner dollar eller mer. Slike prosjekt kjennetegnes av høy grad av kompleksitet og usikkerhet. Kompleksiteten ligger i at det er store prosjekt med mange involverte som igjen fører til mange grensesnitt. Utover dette er prosjektene ofte innovative, med ikke standardisert teknologi og design. Samlet sett fører dette til at usikkerheten for prosjektene øker. Samset (2017) har illustrert dette med en figur. Figuren er gjengitt i Figur 4. Usikkerheten kan slå ut både positivt og negativt for prosjektets kostnader.



Figur 4: Viser hvordan usikkerhet i prosjekt avhenger av størrelse, kompleksitet og innovasjon (Samset, 2017).

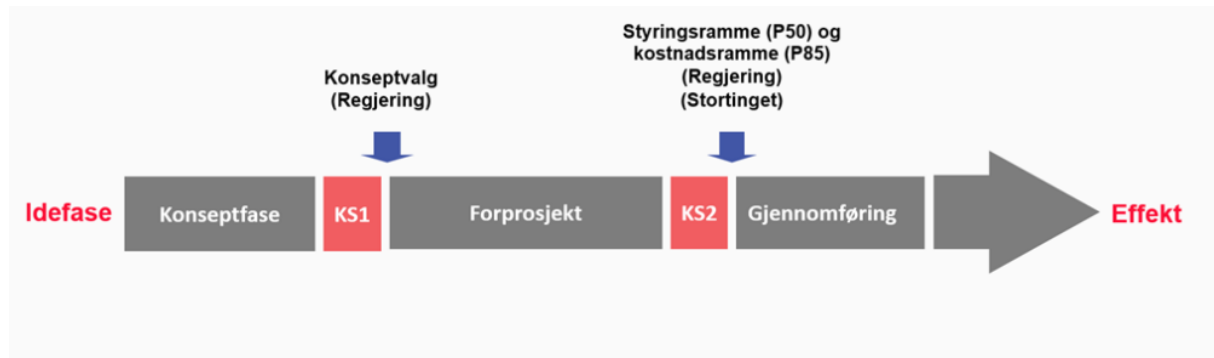
### *Risiko og muligheter*

De positive utfall av usikkerhet blir ansett som muligheter og de negative blir ansett som risiko (Samset, 2017). Ifølge to av Norges største entreprenører innenfor bygg og anlegg er det i dagens byggebransje viktigere enn noen gang å avdekke risiko, ettersom utfallet av å ikke gjøre det kan være dramatisk (Offergaard, 2022, 2023). Etter at pengene i

budsjettet er brukt opp finnes det nesten ingen øvre tapsgrense. Den store usikkerheten forbundet med dagens store og komplekse prosjekt medfører dermed at det er viktig å gi gode antakelser og estimat om fremtiden.

### *Statens prosjektmodell og KS-ordningen*

I skrivende stund stilles det krav til at statlige byggeprosjekt i Norge med en anslått kostnadsramme over 1 milliard kroner skal følge statens prosjektmodell (Finansdepartementet, 2019). Kort fortalt innebærer dette at investeringsprosjektene skal gjennomføres med en idéfase, en konseptfase, et forprosjekt og gjennomføring. Det gjennomføres eksterne kvalitetssikringer mellom konseptfasen og forprosjektet (KS1), samt mellom forprosjektet og gjennomføringen (KS2). Statensprosjekt modell er illustrert i Figur 5.



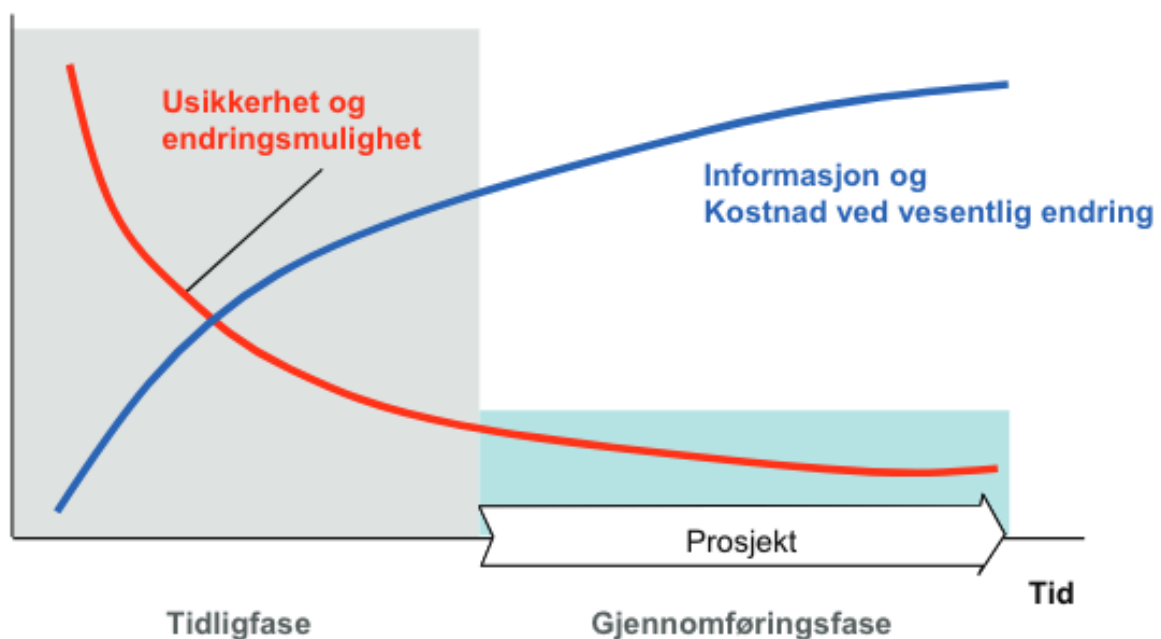
Figur 5: En illustrasjon av statens prosjektmodell (Finansdepartementet, 2019).

Ifølge Finansdepartementet (2019) skal konseptutredningen kvalitetssikres i KS1. Deretter fattet regjeringen en beslutning for hvilket konsept eller alternativ som eventuelt skal videreføres til forprosjektfasen. I KS2 kvalitetssikres det sentrale styringsdokumentet. I tillegg til dette foretas det en vurdering om planleggingen og kostnadene i forprosjektet for det valgte prosjektet er realistisk. Etter KS2 fremlegges saken for regjeringen, med en tilråding om investeringsbeslutning eller plan for videre oppfølging. Regjeringen tar deretter stilling til om saken skal fremmes for Stortinget. Stortinget fastsetter prosjektets kostnadsramme.

Welde (2016) har vist at kostnadsestimatene i KS1 i liten grad har stemt med kostnadsestimatene i KS2. I snitt har økningen fra KS1 til KS2 vært på om lag 40 % i faste kroner. Mangelfull estimering trekkes frem som en av årsakene til dette. Som nevnt tidligere kan dette føre til at feil prosjekt velges. Imidlertid er det i snitt et langt mindre avvik mellom KS2 og de faktiske kostnadene. Welde et al. (2019) har konkludert med at: «totalt sett så synes KS2-rammer og kostnadsestimatene å være realistiske og være rimelig godt kalibrerte». Likevel holdes det frem i samme rapport at utviklingen over tid er bekymringsfull, og det bør komme forbedringer med estimeringsarbeidet. Konkret nevnes det at estimatfordelingene bør anerkjenne at kostnadsusikkerheten typisk er større enn det som tidligere har kommet frem i estimeringsarbeidet. Utover dette bør estimeringen bli bedre på å skille mellom prosjekter med lav og med høy kostnadsusikkerhet.

### 3.2.2 Tilgjengelig informasjon og det digitale miljøets rolle for å redusere usikkerheten i byggeprosjekt

En definisjon for usikkerhet er mangelen på viten om fremtiden (Drevland, 2013; Samset, 2017; Torp et al., 2015). Manglende viten om omfang eller hvor mye en aktivitet kommer til å koste er eksempler på usikkerhet knyttet til kostnader i byggeprosjekt. Samset (2017) sier at usikkerhet kan ses på som «differansen mellom den nødvendige informasjonen for å ta en sikker beslutning og den tilgjengelige informasjonen på beslutningspunktet.». Fra Del 1, kapittel 1 i Sunnevåg (2007) skriver Knut Samset at det antas en sammenheng mellom usikkerhet og informasjonstilgang over tid. I tidligfasen er mulighetene for å redusere usikkerheten størst, og kostnaden av det lavest. For gjennomføringsfasen er situasjonen stort sett låst, endringer i denne fasen vil som regel medføre store kostnader. En illustrasjon av denne forklaringen kan ses i Figur 6.



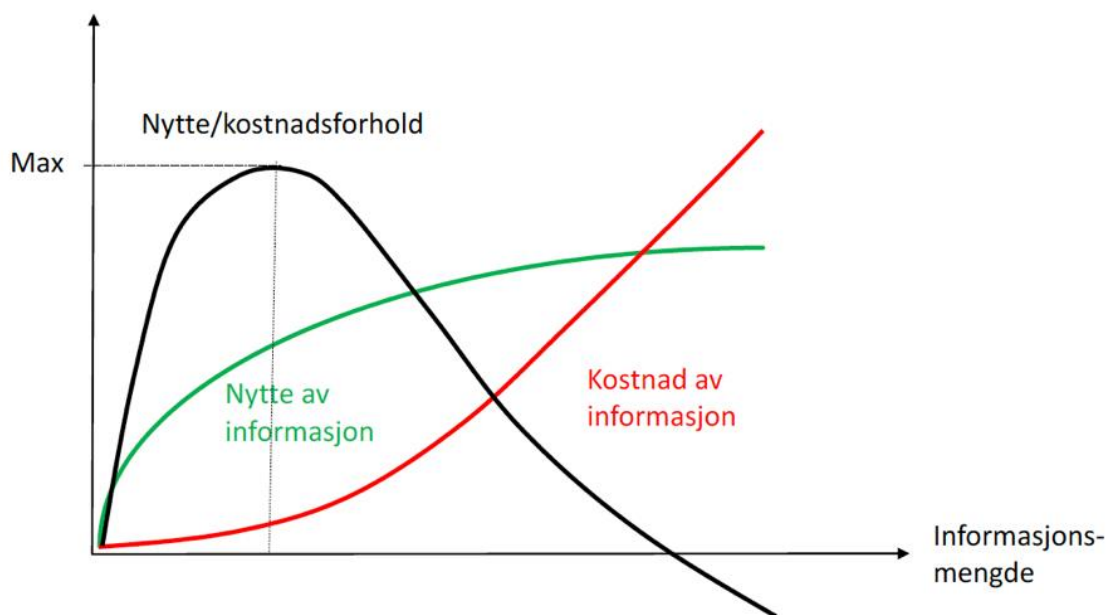
Figur 6: Illustrerer hvordan usikkerheten for et prosjekt endres over tid. Figuren er hentet fra kapittel 1 i Sunnevåg (2007).

For å redusere usikkerheten knyttet til beslutninger er det vanlig å søke etter relevant informasjon relatert til prosjektet. Fra delkapittel 1.2 kommer det frem at data blir til informasjon når det blir bearbeidet av ressurspersoner. Hvilken, og hvor mye, informasjon deltakerne i estimeringsprosess skal få er et omdiskutert tema. Torp et al. (2015) sier at det er særdeles viktig at deltakerne av en ressursgruppe ikke må få tilsendt for mye informasjon. Dette begrunnes med at dersom deltakerne får en stor informasjonsmengde, er det urealistisk å forvente at de vil sette seg inn i alt. Ressurspersonene er som regel travle mennesker, og har ikke mye tid til overs. Følgelig er det sannsynlig at de bare kommer til å se overfladisk på for mye informasjon. Konsekvensen av dette er at deltakerne stiller dårligere forberedt til når de skal gi estimat.

Samset (2021) har også sett på utfordringer relatert til informasjonsmengden. Her dras det frem at tilgangen til informasjon aldri har vært større enn hva den er i dag, og at den er voksende. Bidragsyttere til at informasjonsmengden øker er internett, digitalisering og papirløs informasjon. Hjelpemidlene har gjort at det er lett å produsere rapporter med mye informasjon. En av årsakene til at rapportene blir så informasjonsrike er at

detaljeringsnivået er høyt. Ofte er graden av detaljering høyere enn nødvendig. Torp et al. (2015) peker på at for mange detaljer kan føre til at ressurspersonene bruker like lang tid på å diskutere små kostnadsbærere som store kostnadsbærere. En konsekvens av dette er at uviktige ting får for stor oppmerksomhet i estimeringsprosessen. Videre kan dette føre til at ressurspersonene mister overblikket over totaliteten. Dette kan relateres til basisratefeilen. Basisratefeilen er tidligere omtalt i underkapittel 3.1.3 i denne masteroppgaven. Flere detaljer medfører samtidig flere sammenhenger og grensesnitt, som igjen fører til at det bør inkluderes enda flere usikkerhetsfaktorer. I tillegg til dette vil flere detaljer føre til at ressurspersonene bruker lengre tid på å gi estimat som ikke nødvendigvis er bedre enn dersom bare overordnet informasjonen ble behandlet. Følgelig har det blitt foreslått å gi ressurspersonene bearbeidet informasjon med få detaljer om det som skal estimeres.

Samset (2021) peker på at kostnaden ved å fremskaffe informasjon med høyere grad av presisjon også er høyere. I tidligfasen vil det være relativt rimelig å fremskaffe informasjon. Derimot vil kostnadene øke med større informasjonsmengde og økt presisjonsnivå. Samset (2021) sier videre at desto mer informasjon som er tilgjengelig, desto mindre vil nytten av detaljene være. Sammenhengen mellom nytte og kostnad av informasjon har Samset (2021) illustrert med en figur. Figuren er gjengitt i Figur 7.



Figur 7: Figuren viser sammenhengen mellom nytte og kostnad av informasjon for en økende informasjonsmengde (Samset, 2021).

Illustrasjonen viser at nytte/kostnadsforholdet er størst tidlig i prosessen. Samme illustrasjon viser at ut i prosessen blir kostnaden av å frembringe mer informasjon høyere enn nytten av informasjonen. Dette fører til at forholdet til nytte/kostnad kan bli negativt. Dersom dette er tilfellet betyr det at mindre presis informasjon kan gi et tilstrekkelig godt bilde av situasjonen. Følgelig vil det være bortkastet å bruke mer ressurser enn hva som er nødvendig.

Det er imidlertid utfordrende å fastslå hva som utgjør tilstrekkelig informasjon for å estimere kostnaden av et prosjekt. Litteraturen har vist at det ofte bør tas utgangspunkt i referanseprosjekt, dersom disse finnes (Flyvbjerg, 2021; Sunnevåg, 2007; Torp et al.,

2015). Som nevnt i underkapittel 3.1.3 sier Flyvbjerg (2021) at det er svært få prosjekt som er helt unike. Med et slikt syn vil det være mulig å lære fra andre prosjekt, for eksempel gjennom å bruke erfaringstall. Fra Del 3 kapittel 8 i Sunnevåg (2007) skriver Ole Jonny Klakegg at det ikke er helt uproblematisk å bruke erfaringstall. Det fremholdes at det kan være vanskelig å sammenligne prosjekt ettersom flere prosjekt stiller opp kostnadsoverslag på ulikt vis, blant annet gjennom å definere kostnadselementer ulikt. I tillegg til dette nevnes en rekke karakteristika som bør ses i sammenheng med erfaringstallene. De nevnte forholdene er:

- *Byggherreorganisasjon*
- *Omfang/dimensjon (størrelsesfaktor)*
- *Valg av kvalitet og løsninger, estetikk (materialer, overflater og former)*
- *Teknisk kompleksitet (integrasjon av systemer m.m.)*
- *Arealfordelinger på ulike funksjoner (noen funksjoner er dyrere enn andre)*
- *Grad av rehabilitering og provisorier, driftshensyn (konflikt mellom bygging og drift og flytting av funksjoner)*
- *Kontraktstrategi/entrepriseform (type entrepriser, størrelse på kontrakter).*

#### *Referanseklasseestimering*

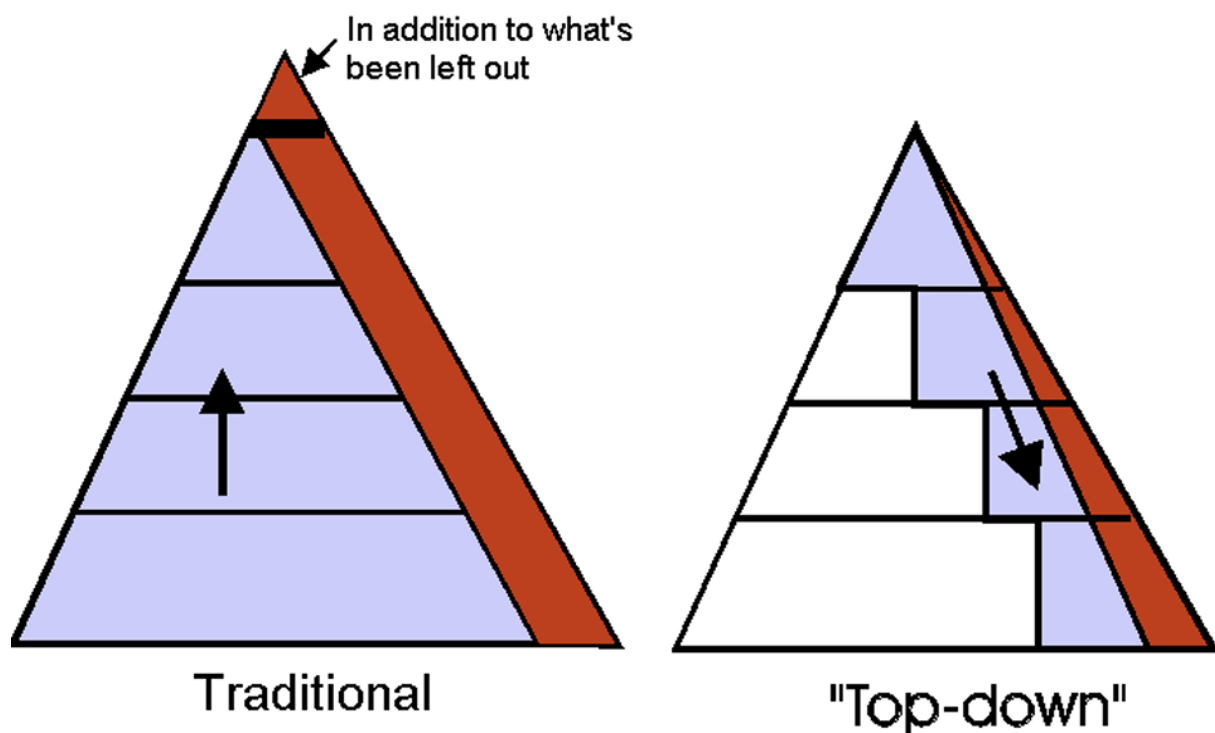
I Flyvbjerg (2006) er tankesettet om benytte erfaringstall videreutviklet til noe som kalles «referanseklasseestimering». Referanseklasseestimering baserer seg på teoriene om beslutninger under usikkerhet som Daniel Kahneman vant Nobelprisen i økonomi for i 2002. Teoriene handler om at det er for mange feilkilder til utarbeidelsen av tradisjonelle kostnadsestimat. For tradisjonelle kostnadsestimat er det vanlig å benytte et innsideperspektiv. Et innsideperspektiv vil si at prosjektet anses som noe unikt, og det bare vies oppmerksomhet til det aktuelle prosjektet sine mål. Det argumenteres for at det fra innsideperspektivet er lett å henge seg opp i detaljer, og at det dermed er lett å miste den totale oversikten. For å bøte på disse problemene har Flyvbjerg (2006) foreslått å benytte seg av et utsideperspektiv. Utsideperspektivet innebærer at man ser utover på andre prosjekt, i motsetning til å se innover på sitt eget prosjekt. Den metodiske tilnærmingen for å ta utsideperspektivet i evalueringen av planlagte hendelser er hva som kalles referanseklasseestimering. For et spesifikt prosjekt kreves det at det identifiseres sammenlignbare prosjekt som tidligere har blitt gjennomført. Her kreves det at de sammenlignbare prosjektene har møtt på de samme utfordringene som det aktuelle prosjektet, samtidig som de er relevante nok for å sammenlignes. De identifiserte prosjektene slås deretter sammen til en klasse. Basert på klassen av identifiserte prosjekt blir det dannet en sannsynlighetsfordeling. Til slutt sammenlignes det aktuelle prosjektet mot sannsynlighetsfordelingen av de sammenlignbare prosjektene.

Følgelig tar ikke referanseklasseestimering sikte på å spå usikkerheten til enkelthendelser. Det aksepteres at noe kommer til å skje uansett, men det gjøres ingen tiltak for å finne ut hva dette «noe» er, eller hvor sannsynlig det er for at det skal forekomme. Informasjonen estimatene baserer seg på er bare overordnet, og knyttet til prisene for sammenlignbare prosjekter.

#### *Top-down og bottom-up tilnærming*

Tilnærmingene top-down og bottom-up er for kostnadsestimering to forskjellige tilnærminger som er prinsipielt forskjellige (Torp, 2021). For å illustrere forskjellen har Torp (2021) laget en figur, figuren er gjengitt i Figur 8. Figuren viser at top-down-tilnærmingen innebærer en nedbrytning av kostnadene i prosjektet fra toppen av

hierarkiet. Først identifiseres kostnadene knyttet til de største operasjonene og oppgavene i prosjektet. Deretter blir arbeidsoppgavene ytterligere detaljert. På den annen side starter bottom-up-tilnærmingen med å ta for seg kostnadene til de minste delene av prosjektet, for deretter å arbeide seg oppover. Bottom-up-tilnærmingen tar sikte på å inkludere kostnadene knyttet til alle elementene i prosjektet, mens top-down-tilnærmingen ikke søker samme detaljeringsnivå. Top-down-tilnærmingen tar sikte på å fange opp kostnadene for de største delene av prosjektet. Imidlertid kan en bottom-up-tilnærming, på grunn av det høyere detaljeringsnivået, ha større risiko for at kostnader blir utelatt. I tillegg kan det detaljerte fokuset føre til tap av oversikt over de største delene av prosjektet. Dette medfører at flere kostnadsdrivere kan bli oversett i en bottom-up-tilnærming sammenlignet med en top-down-tilnærming.



Figur 8: En illustrasjon av forskjellen mellom bottom-up og top-down. Figuren viser hvordan kostnadene brytes ned, og følgerne det har for hva som har blitt utelatt. Figuren er hentet fra (Torp, 2021).

## 4 Resultat

I dette kapitlet vil resultatene fra eksperimentet beskrevet i underkapittel 2.3 presenteres. Som nevnt tidligere startet eksperimentet når jeg begynte å holde Mentimeter-presentasjonen, og ble avsluttet når hver av studentene i flokkene, og supereksperter, hadde gitt to estimat hver. Til tross for at det ikke ble benyttet en stoppeklokke for å dokumentere gjennomføringstiden, kan det med sikkerhet slås fast at supereksperter og studentene hver brukte under 30 minutter for å gi to individuelle estimat.

For å se hvor presise estimatene til studentene og supereksperter var, har det vært nødvendig å ha noe å sammenligne studentenes og supereksperterens estimat med. Byggherren for prosjektet, Statsbygg, hadde akkurat vurdert tre ulike tilbud for totalentreprisen for kontor- og undervisningsbygget. Tilbudene ble gitt av tre ulike entreprenører. Entreprenørene er i masteroppgaven omtalt som Entreprenør 1, Entreprenør 2 og Entreprenør 3. Entreprenørens tilbud ble vurdert 60 % på pris og 40 % på kvaliteten på tilbudt personell. Etersom totalentreprisen ikke er ferdig, vil det være umulig å si hva den faktiske prisen for byggeprosjektet blir. Derfor er entreprenørens priser det nærmeste den aktuelle totalentreprisen kommer en pris. Etersom tilbudene for totalentreprisen ble vurdert både på pris og kvaliteten på tilbudt personell, kunne i teorien entreprenøren med det høyeste tilbudet vunnet totalentreprisen. Derfor vil det etter mitt syn være mulig å både betrakte gjennomsnittsprisene til entreprenørene og tilbudet til entreprenøren som vant anbudskonkurransen (Entreprenør 3), som prisen for totalentreprisen. Jeg har valgt å benytte gjennomsnittsprisen til entreprenørene fordi jeg mener at medianen ikke viser et «samlet» forslag fra entreprenørene, men heller prisen til en av entreprenørene som ikke vant anbudskonkurransen. Gjennomsnittsprisen til entreprenørene og prisen til Entreprenør 3 er derfor benyttet som sammenligningsgrunnlag for supereksperterens estimat og studentenes estimat.

Statsbygg hadde i forkant av anbudskonkurransen gitt et eget estimat for totalentreprisen. Dette estimatet har også blitt sammenlignet gjennomsnittsprisen til entreprenørene og prisen til Entreprenør 3.

Til slutt vil jeg nevne at dersom erfaringsprisene fra Norsk prisbok multipliseres med arealet for dette kontor- og undervisningsbygget, vil det mest konservative estimatet være 368 518 700 kr.

### 4.1 Estimaterne til byggherren

Gjennom en liten dokumentgjennomgang og e-postkorrespondanse med Statsbygg kom det frem at byggherren estimerte prisen for prosjektet til 400 millioner norske kroner + mva (2022-tall).

## 4.2 Entreprenørenes estimat

For prosjektet var det totalt tre entreprenører som leverte ett tilbud hver. Tilbydernes priser kan ses i Tabell 3. Prisene inkluderer ikke mva. Vinneren av anbudskonkurransen var Entreprenør 3. Entreprenørenes samlede estimat ble også funnet. Aritmetisk gjennomsnitt og median ble brukt som statistisk oppsummeringstall. Resultatene kan ses i Tabell 4. Standardavviket til det aritmetiske gjennomsnittet av entreprenørenes estimat kan ses i Tabell 5. Kvartilene for datasettet kan ses i Tabell 6.

Tabell 3: Entreprenørenes tilbud.

Navn på entreprenør	Tilbudt pris (2022 NOK)
Entreprenør 1	392 437 852
Entreprenør 2	409 098 509
Entreprenør 3	382 740 705

Tabell 4: Entreprenørenes samlede estimat.

Statistisk oppsummeringstall	Samlet estimat (2022 NOK)
Aritmetisk gjennomsnitt for entreprenørene	394 759 022
Median til entreprenørene	392 437 852

Tabell 5: Standardavviket til det aritmetiske gjennomsnittet til entreprenørenes estimat.

Hva det skal tas standardavvik av	Standardavvik (2022 NOK)
Aritmetisk gjennomsnitt for entreprenørene	7 696 846

Tabell 6: Kvartilene til entreprenørenes estimat.

Kvartil av medianen	Resultat (2022 NOK)
Minimumsverdi	382 740 705
Første kvartil (25. persentil)	387 589 279
Middelverdi (50. persentil)	392 437 852
Tredje kvartil (75. persentil)	400 768 181
Maksimumsverdi	409 098 509

Tabell 5 og Tabell 6 viser variasjonen for om henholdsvis aritmetisk gjennomsnitt eller median benyttes for å beskrive det typiske i datasettet. Et lavt standardavvik er en indikator på at variasjonen har vært lav. Medianens kvartiler forteller hvor spredt estimatene i datasettet er.



## 4.3 Studentenes estimat

Flokken av alle studentene bestod av totalt fire mindre flokker, Flokk 1, Flokk 2, Flokk 3 og Flokk 4. Hver av personene i flokkene ga ett estimat i to runder. I dette delkapittelet er resultatene fra begge rundene for hver av de mindre flokkene og Flokken av alle studentene presentert. Et aritmetisk gjennomsnitt og medianen ble funnet for hver av flokkene. Standardavviket til gjennomsnittet er funnet for hver av flokkene. Kvartilene for hvert av datasettene er også regnet ut. Ingen av estimatene inkluderte mva. En fullstendig samling av alle estimatene gitt av studentene er gitt i Bilag 4.

### 4.3.1 Flokk 1 av studenter

Flokk 1 bestod av 9 studenter. Fra estimatene til disse 9 studentene ble det aritmetiske gjennomsnittet og medianen til Flokk 1 funnet for begge rundene. Resultatet kan ses i Tabell 7. Standardavviket til hver av de to aritmetiske gjennomsnittene til Flokk 1 kan ses i Tabell 8. Kvartilene for datasettet til Flokk 1 kan ses i Tabell 9.

Tabell 7: Aritmetisk gjennomsnitt og median for Flokk 1 for hver av rundene.

<b>Statistisk oppsummeringstall</b>	<b>Samlet estimat i runde 1 (2022 NOK)</b>	<b>Samlet estimat i runde 2 (2022 NOK)</b>
Aritmetisk gjennomsnitt for Flokk 1	411 988 889	409 833 333
Median til Flokk 1	400 000 000	380 000 000

Tabell 8: Standardavviket til det aritmetiske gjennomsnittet for Flokk 1.

<b>Hva det skal tas standardavvik av</b>	<b>Standardavvik (2022 NOK)</b>
Aritmetisk gjennomsnitt for Flokk 1 i runde 1	95 625 805
Aritmetisk gjennomsnitt for Flokk 1 i runde 2	30 021 289

Tabell 9: Kvartilene til estimatene for Flokk 1.

<b>Kvartilene til Flokk 1</b>	<b>Resultat fra runde 1 (2022 NOK)</b>	<b>Resultat fra runde 2 (2022 NOK)</b>
Minimumsverdi	12 000 000	330 000 000
Første kvartil (25. persentil)	230 000 000	334 000 000
Middelverdi (50. persentil)	400 000 000	380 000 000
Tredje kvartil (75. persentil)	500 000 000	450 000 000
Maksimumsverdi	1 000 000 000	600 000 000

Tabell 8 viser at standardavviket til Flokk 1 i runde 1 (95 625 805 NOK) var høyere enn entreprenørenes standardavvik (7 696 846 NOK). I runde 2 var standardavviket til Flokk 1 (30 021 289 NOK) også høyere enn entreprenørenes standardavvik. Mellom runde 1 og runde 2 gikk standardavviket ned.

Tabell 9 viser at estimatene til Flokk 1 i runde 1 var spredt fra en lavere verdi (12 000 000 NOK) enn entreprenørens laveste verdi (382 740 705 NOK), til en høyere verdi (1 000 000 000 NOK) enn entreprenørens høyeste verdi (409 098 509 NOK). For runde 2 var estimatene til Flokk 1 også spredt fra en lavere verdi (330 000 000 NOK) enn entreprenørens laveste verdi, til en høyere verdi (600 000 000 NOK) enn entreprenørens høyeste verdi. Mellom runde 1 og runde 2 ble fordelingen av estimat smalere.

### 4.3.2 Flokk 2 av studenter

Flokk 2 bestod av 6 studenter. Fra estimatene til disse 6 studentene ble det aritmetiske gjennomsnittet og medianen til Flokk 2 funnet for begge rundene. Resultatet kan ses i Tabell 10. Standardavviket til hver av de to aritmetiske gjennomsnittene til Flokk 2 kan ses i Tabell 11. Kvartilene for datasettet til Flokk 2 kan ses i Tabell 14.

Tabell 10: Aritmetisk gjennomsnitt og median for Flokk 2 for hver av rundene.

<b>Statistisk oppsummeringstall</b>	<b>Samlet estimat i runde 1 (2022 NOK)</b>	<b>Samlet estimat i runde 2 (2022 NOK)</b>
Aritmetisk gjennomsnitt for Flokk 2	395 525 000	382 416 667
Median til Flokk 2	315 000 000	370 000 000

Tabell 11: Standardavviket til det aritmetiske gjennomsnittet for Flokk 2.

<b>Hva det skal tas standardavvik av</b>	<b>Standardavvik (2022 NOK)</b>
Aritmetisk gjennomsnitt for Flokk 2 i runde 1	128 363 270
Aritmetisk gjennomsnitt for Flokk 2 i runde 2	17 483 524

Tabell 12: Kvartilene til estimatene for Flokk 2.

<b>Kvartilene til Flokk 2</b>	<b>Resultat fra runde 1 (2022 NOK)</b>	<b>Resultat fra runde 2 (2022 NOK)</b>
Minimumsverdi	48 000 000	345 000 000
Første kvartil (25. persentil)	171 250 000	348 125 000
Middelverdi (50. persentil)	315 000 000	370 000 000
Tredje kvartil (75. persentil)	657 612 250	406 500 000
Maksimumsverdi	800 000 000	450 000 000

Tabell 10 viser at standardavviket til Flokk 2 i runde 1 (128 363 270 NOK) var høyere enn entreprenørenes standardavvik (7 696 846 NOK). I runde 2 var standardavviket til Flokk 2 (17 483 524 NOK) også høyere enn entreprenørenes standardavvik. Mellom runde 1 og runde 2 gikk standardavviket ned. Mellom runde 1 og runde 2 gikk standardavviket ned.

Tabell 11 viser at estimatene til Flokk 2 i runde 1 var spredt fra en lavere verdi (48 000 000 NOK) enn entreprenørenes laveste verdi (382 740 705 NOK), til en høyere verdi (800 000 000 NOK) enn entreprenørenes høyeste verdi (409 098 509 NOK). For runde 2 var estimatene til Flokk 2 også spredt fra en lavere verdi (345 000 000 NOK) enn entreprenørenes laveste verdi, til en høyere verdi (450 000 000 NOK) enn entreprenørenes høyeste verdi. Mellom runde 1 og runde 2 ble fordelingen av estimat smalere. Mellom runde 1 og runde 2 ble fordelingen av estimat smalere.

### 4.3.3 Flokk 3 av studenter

Flokk 3 bestod av 14 studenter. Fra estimatene til disse 14 studentene ble det aritmetiske gjennomsnittet og medianen til Flokk 3 funnet for begge rundene. Resultatet kan ses i Tabell 13. Standardavviket til hver av de to aritmetiske gjennomsnittene til Flokk 3 kan ses i Tabell 17. Kvartilene for datasettet til Flokk 3 kan ses i Tabell 15.

Tabell 13: Aritmetisk gjennomsnitt og median for Flokk 3 for hver av rundene.

<b>Statistisk oppsummeringstall</b>	<b>Samlet estimat i runde 1 (2022 NOK)</b>	<b>Samlet estimat i runde 2 (2022 NOK)</b>
Aritmetisk gjennomsnitt for Flokk 3	585 428 571	481 458 576
Median til Flokk 3	515 000 000	442 500 000

Tabell 14: Standardavviket til det aritmetiske gjennomsnittet for Flokk 3.

<b>Hva det skal tas standardavvik av</b>	<b>Standardavvik (2022 NOK)</b>
Aritmetisk gjennomsnitt for Flokk 3 i runde 1	82 303 827
Aritmetisk gjennomsnitt for Flokk 3 i runde 2	43 504 752

Tabell 15: Kvartilene til estimatene for Flokk 3.

<b>Kvartilene til Flokk 3</b>	<b>Resultat fra runde 1 (2022 NOK)</b>	<b>Resultat fra runde 2 (2022 NOK)</b>
Minimumsverdi	250 000 000	341 000 000
Første kvartil (25. persentil)	332 000 000	392 500 000
Middelverdi (50. persentil)	515 000 000	442 500 000
Tredje kvartil (75. persentil)	753 750 000	489 855 017
Maksimumsverdi	1 300 000 000	950 000 000

Tabell 14 viser at standardavviket til Flokk 3 i runde 1 (82 303 827 NOK) var høyere enn entreprenørenes standardavvik (7 696 846 NOK). I runde 2 var standardavviket til Flokk 3 (43 504 752 NOK) også høyere enn entreprenørenes standardavvik. Mellom runde 1 og runde 2 gikk standardavviket ned.

Tabell 15 viser at estimatene til Flokk 3 i runde 1 var spredt fra en lavere verdi (250 000 000 NOK) enn entreprenørenes laveste verdi (382 740 705 NOK), til en høyere verdi (1 300 000 000 NOK) enn entreprenørenes høyeste verdi (409 098 509 NOK). For runde 2 var estimatene til Flokk 3 også spredt fra en lavere verdi (341 000 000 NOK) enn entreprenørenes laveste verdi, til en høyere verdi (950 000 000 NOK) enn entreprenørenes høyeste verdi. Mellom runde 1 og runde 2 ble fordelingen av estimat smalere.

#### 4.3.4 Flokk 4 av studenter

Flokk 4 bestod av 4 studenter. Fra estimatene til disse 4 studentene ble det aritmetiske gjennomsnittet og medianen til Flokk 4 funnet for begge rundene. Resultatet kan ses i Tabell 16. Standardavviket til hver av de to aritmetiske gjennomsnittene til Flokk 4 kan ses i Tabell 21. Kvartilene for datasettet til Flokk 4 kan ses i Tabell 18.

Tabell 16: Aritmetisk gjennomsnitt og median for Flokk 4 for hver av rundene.

<b>Statistisk oppsummeringstall</b>	<b>Samlet estimat i runde 1 (2022 NOK)</b>	<b>Samlet estimat i runde 2 (2022 NOK)</b>
Aritmetisk gjennomsnitt for Flokk 4	647 500 000	541 250 000
Median til den Flokk 4	545 000 000	492 500 000

Tabell 17: Standardavviket til det aritmetiske gjennomsnittet for Flokk 4.

<b>Hva det skal tas standardavvik av</b>	<b>Standardavvik (2022 NOK)</b>
Aritmetisk gjennomsnitt for Flokk 4 i runde 1	153 146 172
Aritmetisk gjennomsnitt for Flokk 4 i runde 2	98 569 455

Tabell 18: Kvartilene til estimatene for Flokk 4.

<b>Kvartilene til Flokk 4</b>	<b>Resultat fra runde 1 (2022 NOK)</b>	<b>Resultat fra runde 2 (2022 NOK)</b>
Minimumsverdi	300 000 000	330 000 000
Første kvartil (25. persentil)	440 000 000	335 000 000
Middelverdi (50. persentil)	647 500 000	541 250 000
Tredje kvartil (75. persentil)	650 000 000	650 000 000
Maksimumsverdi	1 200 000 000	850 000 000

Tabell 17 viser at standardavviket til Flokk 4 i runde 1 (153 146 172 NOK) var høyere enn entreprenørenes standardavvik (7 696 846 NOK). I runde 2 var standardavviket til Flokk 4 (98 569 455 NOK) også høyere enn entreprenørenes standardavvik. Mellom runde 1 og runde 2 gikk standardavviket ned.

Tabell 18 viser at estimatene til Flokk 4 i runde 1 var spredt fra en lavere verdi (300 000 000 NOK) enn entreprenørenes laveste verdi (382 740 705 NOK), til en høyere verdi (1 200 000 000 NOK) enn entreprenørenes høyeste verdi (409 098 509 NOK). For runde 2 var estimatene til Flokk 4 også spredt fra en lavere verdi (330 000 000 NOK) enn entreprenørenes laveste verdi, til en høyere verdi (850 000 000 NOK) enn entreprenørenes høyeste verdi. Mellom runde 1 og runde 2 ble fordelingen av estimat smalere.

#### 4.3.5 Flokken av alle studentene

Flokken av alle studentene bestod av totalt 33 studenter. Fra estimatene til disse 33 studentene ble det aritmetiske gjennomsnittet og medianen til Flokken av alle studentene funnet. Resultatene kan ses i Tabell 19. Standardavviket til hver av de to aritmetiske gjennomsnittene til Flokken av alle studentene kan ses i Tabell 20. Kvartilene for datasettet Flokken av alle studentene kan ses i Tabell 21.

Tabell 19: Aritmetisk gjennomsnitt og median for Flokken av alle studentene for hver av rundene.

<b>Statistisk oppsummeringstall</b>	<b>Samlet estimat i runde 1 (2022 NOK)</b>	<b>Samlet estimat i runde 2 (2022 NOK)</b>
Aritmetisk gjennomsnitt for Flokken av alle studentene	511 122 727	451 164 245
Median til Flokken av alle studentene	450 000 000	400 000 000

Tabell 20: Standardavviket til det aritmetiske gjennomsnittet for Flokken av alle studentene.

<b>Hva det skal tas standardavvik av</b>	<b>Standardavvik (2022 NOK)</b>
Aritmetisk gjennomsnitt for Flokken av alle studentene i runde 1	54 832 305
Aritmetisk gjennomsnitt for Flokken av alle studentene i runde 2	25 781 121

Tabell 21: Kvartilene til datasettet for Flokken av alle studentene.

<b>Kvartilene til Flokken av alle studentene</b>	<b>Resultat fra runde 1 (2022 NOK)</b>	<b>Resultat fra runde 2 (2022 NOK)</b>
Minimumsverdi	12 000 000	330 000 000
Første kvartil (25. persentil)	300 000 000	350 000 000
Middelverdi (50. persentil)	450 000 000	400 000 000
Tredje kvartil (75. persentil)	650 000 000	450 420 069
Maksimumsverdi	1 300 000 000	950 000 000

Tabell 19 viser at standardavviket til Flokk av alle studentene i runde 1 (54 832 305 NOK) var høyere enn entreprenørenes standardavvik (7 696 846 NOK). I runde 2 var standardavviket til Flokk av alle studentene (25 781 121 NOK) også høyere enn entreprenørenes standardavvik. Mellom runde 1 og runde 2 gikk standardavviket ned.

Tabell 20 viser er at estimatene til Flokk av alle studentene i runde 1 var spredt fra en lavere verdi (12 000 000 NOK) enn entreprenørenes laveste verdi (382 740 705 NOK), til en høyere verdi (1 300 000 000 NOK) enn entreprenørenes høyeste verdi (409 098 509 NOK). For runde 2 var estimatene til Flokk av alle studentene også spredt fra en lavere verdi (330 000 000 NOK) enn entreprenørenes laveste verdi, til en høyere verdi (950 000 000 NOK) enn entreprenørenes høyeste verdi. Mellom runde 1 og runde 2 ble fordelingen av estimat smalere.

## 4.4 Sammenligning av standardavvikene

Usikkerheten til et gjennomsnittsestimat uttrykkes ved standardavviket. For å sammenligne hvem av flokkene av studenter som har størst og lavest usikkerhet (standardavvik) for sitt gjennomsnittsestimat, har det blitt laget en tabell. Tabell 22 viser hvem av flokkene som i runde 1 og runde 2 hadde lavest og høyest usikkerhet for sitt gjennomsnitt. Tabellen har en rekkefølge der flokken som hadde lavest standardavvik er øverst. Fra toppen til bunnen av tabellen er det et stigende standardavvik for flokkene.

Tabell 22: En sortering av hvilke flokker av studenter som hadde lavest og høyest standardavvik for sine respektive gjennomsnitt. Øverst i tabellen er flokken med lavest standardavvik, og nederst er flokken med høyest standardavvik. Tabellen viser rekkefølgen for runde 1 og runde 2.

<b>Runde 1</b>	<b>Runde 2</b>
Flokk av alle studentene	Flokk 2
Flokk 3	Flokken av alle studentene
Flokk 1	Flokk 1
Flokk 2	Flokk 3
Flokk 4	Flokk 4

Flokk 1, Flokk 2, Flokk 3, Flokk 4 og Flokken av alle studentene hadde alle høyere standardavvik enn standardavviket til entreprenørens gjennomsnittspris.

## 4.5 Estimaterne til supereksperter

Supereksperterens estimat kan ses i Tabell 23. Supereksperterens estimat er gitt uten mva.

Tabell 23: Estimaterne til den eksterne eksperter.

<b>Estimatgiver</b>	<b>Estimat i den første runden (2022 NOK)</b>	<b>Estimat i den andre runden (2022 NOK)</b>
Supereksperter	686 850 000	686 850 000

Fra Tabell 23 kan det observeres at supereksperter valgte å gi det samme estimatet i begge rundene.

## 4.6 Sammenligning av estimatverdiene

I dette delkapittelet sammenlignes estimatverdiene gitt av de mindre flokkene studenter, Flokken av alle studentene, supereksperter og Statsbygg, prosentvis med entreprenørens gjennomsnittspris og prisen til Entreprenør 3. For studentene og supereksperter ble estimatene fra begge rundene sammenlignet med entreprenørens gjennomsnittspris og prisen til Entreprenør 3. En samling av de prosentvise sammenligningene for studentene og supereksperter er gitt i Tabell 24. Statsbygg hadde bare ett estimat. Dette estimat ble sammenlignet med gjennomsnittsprisen til entreprenørene og prisen til Entreprenør 3. Sammenligningen kan ses i Tabell 25.

Tabell 24: Prosentvis sammenligning av studentenes og supereksperterens estimat målt mot entreprenørene. + indikerer over og - indikerer under.

Gruppe	Sammenlignet med gjennomsnittet til entreprenørene i runde 1	Sammenlignet med Entreprenør 3 i runde 1	Sammenlignet med gjennomsnittet til entreprenørene i runde 2	Sammenlignet med Entreprenør 3 i runde 2
Flokk 1 median	+1,33 %	+4,51 %	-3,74 %	-0,72 %
Flokk 1 gjennomsnitt	+4,36 %	+7,64 %	+3,82 %	+7,08 %
Flokk 2 median	-20,20 %	-17,70 %	-6,27 %	-3,33 %
Flokk 2 gjennomsnitt	+0,19 %	+3,34 %	-3,13 %	-0,08 %
Flokk 3 median	+30,46 %	+34,56 %	+12,09 %	+15,61 %
Flokk 3 gjennomsnitt	+48,30 %	+52,96 %	+21,96 %	+25,79 %
Flokk 4 median	+38,06 %	+42,39 %	+24,76 %	+28,68 %
Flokk 4 gjennomsnitt	+64,02 %	+69,17 %	+37,11 %	+41,41 %
Flokk av alle studentene median	+13,99 %	+17,57 %	+1,33 %	+4,51 %
Flokk av alle studentene gjennomsnitt	+29,48 %	+33,54 %	+14,29 %	+17,88 %
Superekspert	+73,99 %	+79,46 %	+73,99 %	+79,46 %

Tabell 25: Statsbyggs estimat sammenlignet med entreprenørens gjennomsnitt og Entreprenør 3.

Gruppe	Sammenlignet med gjennomsnittet til entreprenørene	Sammenlignet med Entreprenør 3
Statsbygg	+1,33 %	+4,51 %

Fra Tabell 24 er det synlig at medianen til Flokk 1 var noe lengre unna gjennomsnittsprisene til entreprenørene i runde 2 enn i runde 1. Gjennomsnittet til Flokk 2 var også noe lengre unna gjennomsnittsprisene til entreprenørene i runde 2 enn i runde 1. Imidlertid var medianen til Flokk 1 og gjennomsnittet til Flokk 2 nærmere prisen til Entreprenør 3 i runde 2. For alle de andre flokkene var estimatene gitt i runde 2 nærmere gjennomsnittsprisen til entreprenørene og prisen til Entreprenør 3 enn estimatene i runde 1. Dette var tilfellet uavhengig av hvilket oppsummeringstall som ble



brukt. Samlet sett førte endringene av estimat til studentene i 18 av 20 tilfeller til mer presise estimat i runde 2 enn runde 1.

Tabell 24 viser at i både runde 1 og runde 2 har medianen for Flokk av alle studentene gitt et estimat som ligger nærmere entreprenørenes gjennomsnittspris, og prisen til Entreprenør 3, enn hva et aritmetisk gjennomsnitt gir. For Flokk 1, Flokk 3 og Flokk 4 ga medianen et estimat som var nærmere gjennomsnittsprisen til entreprenørene og prisen til Entreprenør 3, enn det aritmetiske gjennomsnittet i både runde 1 og runde 2. For Flokk 1 var gjennomsnittet det statistiske oppsummeringstallet som ga estimat nærmere gjennomsnittsprisen til entreprenørene og prisen til Entreprenør 3 i både runde 1 og runde 2.

Fra Tabell 24 og Tabell 25 er det mulig å se at det for runde 1 og runde 2 var supereksperten som ga estimatet som var lengst unna entreprenørenes gjennomsnittspris og prisen til Entreprenør 3. I runde 1 og runde 2 var gjennomsnittsestimatet til Flokk 2 nærmest entreprenørenes gjennomsnittspris og prisen til Entreprenør 3. Blant estimatene til de mindre flokkene var gjennomsnittsestimatet til Flokk 4 lengst unna entreprenørenes gjennomsnittspris og prisen til Entreprenør 3. I runde 1 var medianen til Flokk 2 lavere enn entreprenørenes gjennomsnittspris og prisen til Entreprenør 3. For runde 2 var medianen til Flokk 1, gjennomsnittsestimatet til Flokk 2 og medianen til Flokk 2 lavere enn entreprenørenes gjennomsnittspris og prisen til Entreprenør 3.

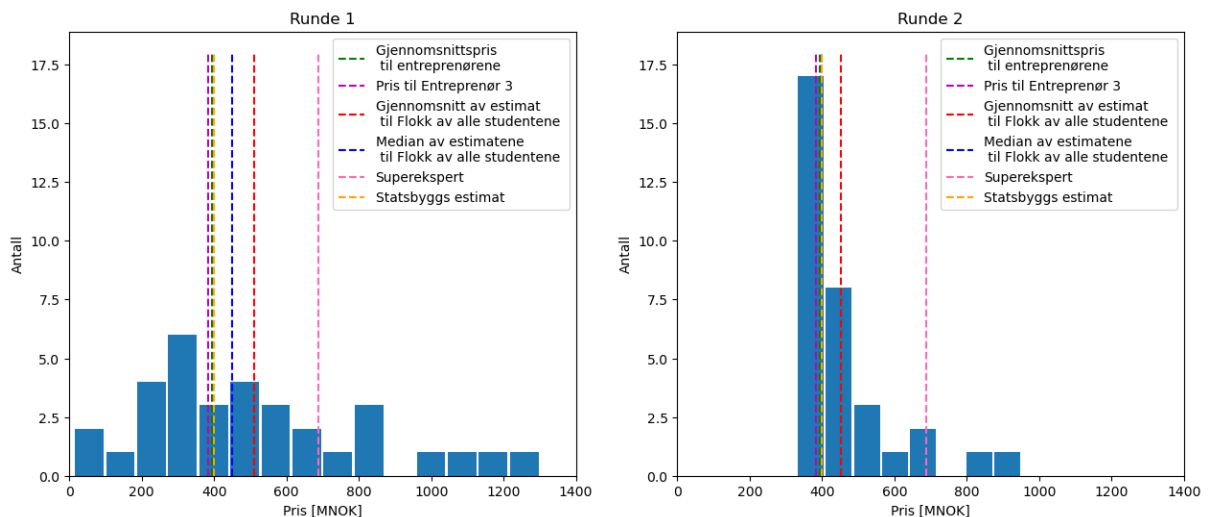
Av Tabell 24 og Tabell 25 er det mulig å se at i runde 1 var medianen til Flokk 1, gjennomsnittsestimatet til Flokk 1 og gjennomsnittsestimatet til Flokk 2 under 10 % unna entreprenørenes gjennomsnittspris og prisen til Entreprenør 3. For runde 2 var både medianen og gjennomsnittsestimatet til både Flokk 1 og Flokk 2, samt medianen til Flokken av alle studentene, under 10 % unna entreprenørenes gjennomsnittspris og prisen til Entreprenør 3. Statsbyggs estimat var også under 10 % unna entreprenørenes gjennomsnittspris og prisen til Entreprenør 3. At flere av de nevnte estimatene for studentene også er under 5 % unna enten entreprenørenes gjennomsnittspriser eller prisen til Entreprenør 3 er spesielt oppsiktsvekkende.

Innenfor rammene i eksperimentet har alle flokkene av studenter gitt mer presise estimat enn supereksperten. Det er oppsiktsvekkende at selv flokker bestående av få studenter med lite/ingen erfaring med å anslå priser har gitt estimat som er nærmere gjennomsnittsprisene til entreprenørene og Entreprenør 3 enn supereksperten. Det bemerkes også at Flokken av alle studentene har gitt det samme estimatet som Statsbygg.

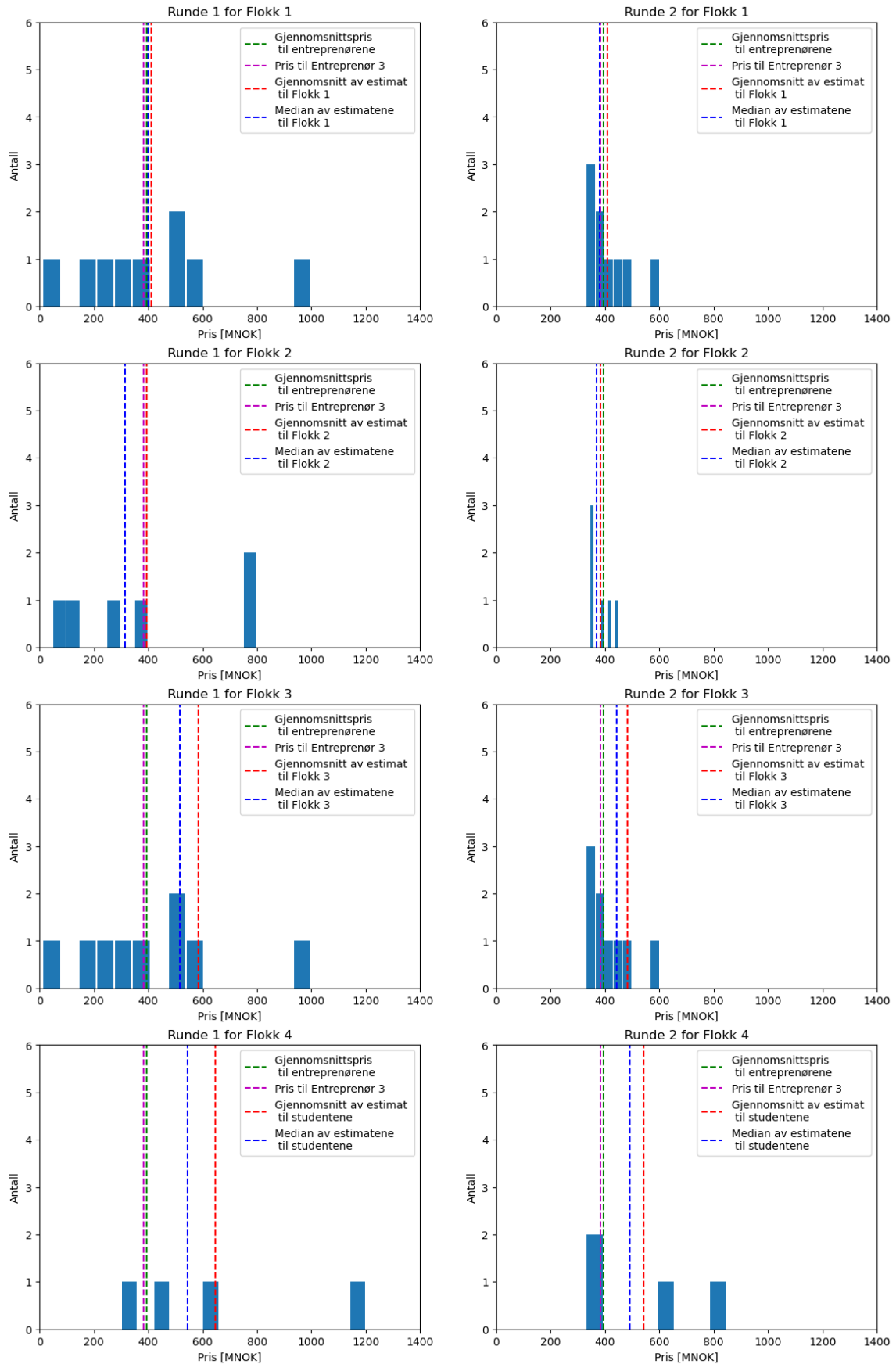
## 4.7 Fordeling av estimatene til Flokken av alle studentene

En grafisk fremstilling av hvordan estimatene til Flokk av alle studentene er fordelt kan ses i Figur 9. Figuren viser hvordan estimatene til Flokk av alle studentene fordelte seg i begge rundene. Gjennomsnittet til Flokk av alle studentene, medianen til Flokke av alle studentene, supereksperens estimat, Statsbyggs estimat, entreprenørens gjennomsnitt og vinneren av anbudskonkurransen (Entreprenør 3) er markert med stiplede linjer. En samling av fordelingen til Flokk 1, Flokk 2, Flokk 3 og Flokk 4 er gitt i Figur 10.

Ved å se hvordan søylene i Figur 9 er fordelt kan det for Flokken av alle studentene observeres at estimatene har en høyreskjev fordeling i runde 1 og runde 2. Annet enn å observere en høyreskjev fordeling fra søylene, kan det også observeres at den stiplede linjen som angir gjennomsnittet til flokken er et stykke til høyre for den stiplede linjen som angir medianen for flokken. Ved å se på forskjellene mellom disse to linjene kan det fra Figur 10 ses at Flokk 2, Flokk 3 og Flokk 4 har en høyreskjev fordeling av estimat i runde 1, mens Flokk 1 har en symmetrisk fordeling av estimat i runde 1. For runde 2 kan det ved bruk av de stiplede linjene for gjennomsnittet og medianen observeres at Flokk 3 og Flokk 4 har en høyreskjev fordeling, mens Flokk 1 og Flokk 2 har en forholdsvis symmetrisk fordeling.



Figur 9: En grafisk fremstilling av hvordan estimatene til alle studentene er fordelt. Figuren viser hvordan estimatene fordelte seg i begge rundene. De samlede estimatene til eksperimentets deltakere er tegnet med stiplede linjer.

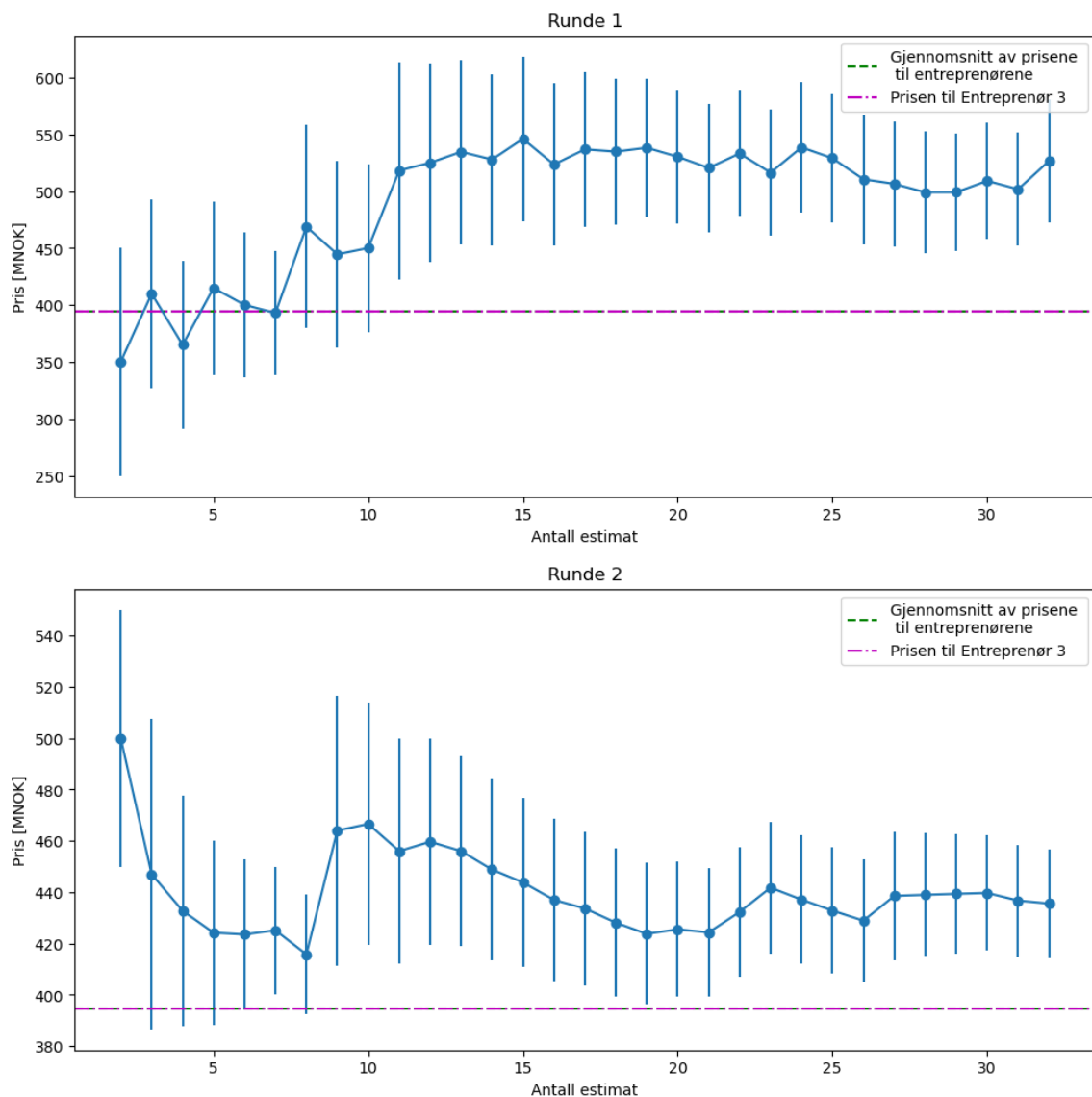


Figur 10: En grafisk fremstilling av hvordan estimatene til de mindre flokkene av studenter er fordelt. Figuren viser hvordan estimatene fordelte seg i begge rundene. De samlede estimatene til eksperimentets deltakere er tegnet med stiplede linjer.

Figur 9 og Figur 10 viser at for runde 2 er estimatene til studentene samlet rundt færre verdier enn runde 1. Av figurene ser det ut til at svarene til studentene blir mer sentrert rundt ett punkt, og at distribusjonen av svarene er snevrere enn i runde 1. Figurene viser at det for runde 2 er færre høye verdier i datasettet, sammenlignet med runde 1.

## 4.8 Antall studenters påvirkning på standardavviket til gjennomsnittet

Det er sett på hvordan standardavviket til gjennomsnittet av estimatene til studentene blir påvirket av antall deltakere. Figur 11 viser resultatene av å bruke den første delen av Python-koden beskrevet i underkapittel 2.3.3. I korte trekk ble det for ulike flokkstørrelser trukket ut tilfeldige estimat fra datasettet bestående av estimatene til studentene, tatt gjennomsnitt av estimatene som ble trukket ut, og til slutt beregnet et standardavvik for gjennomsnittet av estimatene. De blå sirklene i Figur 11 viser gjennomsnittet til estimatene som ble plukket ut av datasettet, og de blå strekene viser standardavviket til gjennomsnittet av estimatene som ble plukket ut. Gjennomsnittsprisen til entreprenørene og prisen til Entreprenør 3 er tegnet inn som stiplede linjer. De stiplede linjene er der for å gi en referanse for hvor langt unna gjennomsnittet til de utvalgte estimatene er.

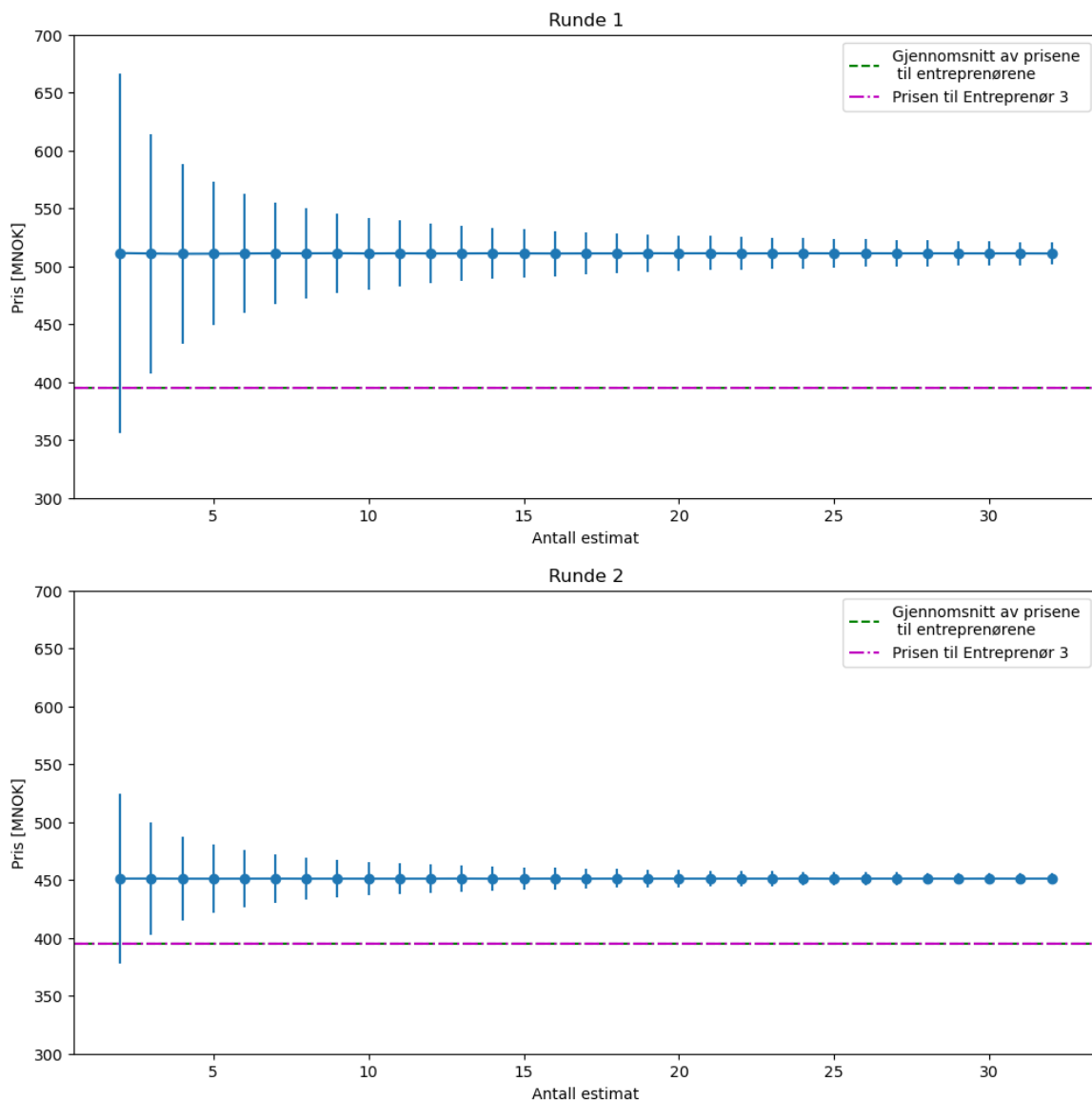


Figur 11: Viser hvordan standardavviket til gjennomsnittet blir påvirket av antall deltakere. Figuren viser dette for både runde 1 og runde 2.

Figur 11 viser at det er mulig å tilfeldig trekke ut et lavt antall estimat fra datasettet der gjennomsnittet av estimatene som trekkes ut ligger nærme entreprenørenes gjennomsnittspris og prisen til Entreprenør 3. Fra runde 1 i figuren kan det observeres at for tre estimat så ligger den blå sirkelen (gjennomsnittet av estimatene) nærme linjene som angir entreprenørenes gjennomsnittspris og prisen til Entreprenør 3. Fra samme figur er det dog synlig at de blå strekene er lange (høyt standardavvik) for gjennomsnittene der det er brukt et lavt antall estimat. Dermed er også usikkerheten til gjennomsnittene av et lavt antall estimat høyt. Figuren viser at det er mulig å oppnå ønsket resultat med få personer, men at det er avhengig av en del flaks i utvelgelsen av estimat fra datasettet til flokken. Runde 2 i Figur 11 viser hva som skjer dersom to av estimatene som plukkes ut fra datasettet ikke er gode. Da vil resultatet ligge lengre unna entreprenørenes samlede pris.

## 4.9 Standardavvikets utvikling for ulike flokkstørrelser av studenter

Det ble identifisert hvordan standardavviket utvikler seg for ulike flokkstørrelser. Figur 12 viser resultatene av å bruke den andre delen av Python-koden beskrevet i underkapittel 2.3.3. I korte trekk ble det for ulike flokkstørrelser trukket ut tilfeldige estimat fra datasettet bestående av estimatene til studentene og deretter tatt et gjennomsnitt av estimatene som ble trukket ut. Denne prosessen ble gjentatt 1 million ganger for hver flokkstørrelse. Deretter ble det tatt et gjennomsnittet av de 1 million gjennomsnittene, og så regnet ut ett standardavvik av gjennomsnittet av de 1 million gjennomsnittene. Dette ble gjort for hver flokkstørrelse. De blå sirklene i Figur 12 viser gjennomsnittet til verdiene som ble plukket ut av datasettet, og de blå strekene viser standardavviket til gjennomsnittet av verdiene som ble plukket ut. Gjennomsnittsprisen til entreprenørene og prisen til Entreprenør 3 er tegnet inn som stiplede linjer. De stiplede linjene er der for å gi en referanse for hvor langt unna gjennomsnittet til de utvalgte estimatene er.



Figur 12: Viser standardavviket til gjennomsnittet av gjennomsnittet av 1 000 000 ulike kombinasjoner for hvert antall deltakere. Figuren viser dette for både runde 1 og runde 2.

Figur 12 viser at for både runde 1 og runde 2 så blir de blå strekene kortere for hver eneste økning av antall estimat. Derfor kan man si at det er en trend for at standardavviket til gjennomsnittet går ned for et høyere antall estimat. Fra samme figur er det synlig at de blå strekene for runde 2 er kortere sammenlignet med et tilsvarende antall estimat i runde 1. Dette viser at det er mindre variasjon i datasettet for runde 2 enn for runde 1, og med det at gjennomsnittsestimatene har en lavere grad av usikkerhet i runde 2 enn i runde 1.

Figur 12 viser også at usikkerheten går kjappet ned for de laveste antallene av estimat. Usikkerheten reduseres mer av å legge til 2 personer i en flokk bestående av 3 personer, enn av å legge til 2 personer i en flokk bestående av 30 personer. Imidlertid vil den totale usikkerheten være lavere for en flokk bestående av 32 personer enn 5 personer.

Fra Figur 12 er det mulig å se at de blå sirklene (gjennomsnittet av estimatene) i runde 1 ligger på omtrent samme verdi, og at de blå sirklene i runde 2 ligger på omtrent samme verdi. I runde 1 ligger de blå sirklene omtrent på samme verdi som gjennomsnittet til datasettet bestående av alle studentene for runde 1. For runde 2 ligger de blå sirklene omtrent på samme verdi som gjennomsnittet til datasettet bestående av alle studentene for runde 2. Dette kommer av at når man ser på gjennomsnittet av gjennomsnittet til nesten alle kombinasjonene for hvert antall estimat, så vil verdien for hvert antall estimat gå mot gjennomsnittsverdien til datasettet. Dette viser at selv om usikkerheten for hver flokkstørrelse blir lavere, er flokken samlet sett fortsatt avhengig av å treffe i nærheten av det svaret som er ønskelig.

## 4.10 Ekspertens punkter over nøkkelinformasjon

Gjennom litteraturstudien er det kommet frem at det ikke finnes en samlet liste over den viktigste nøkkelinformasjonen for å kunne gi presise estimat for kostnaden av et byggeprosjekt. For å få en input til hvilken nøkkelinformasjon som burde være med i en liste over den viktigste nøkkelinformasjonen ble eksperten etter presentasjonen spurt om hva han anså å være de viktigste punktene. Eksperten listet opp følgende punkter:

- Grunnforhold
- Bygningsgeometri
- Bruttoareal
- Nettoareal
- Bruk av materialer
- Miljøambisjon
- Byggetid
- Hvilken fase prosjektet er inne i
- Finanskostnader
- Tilgang på arealer for rigg og drift
- Byggherreorganisasjon
- Prosjekterende

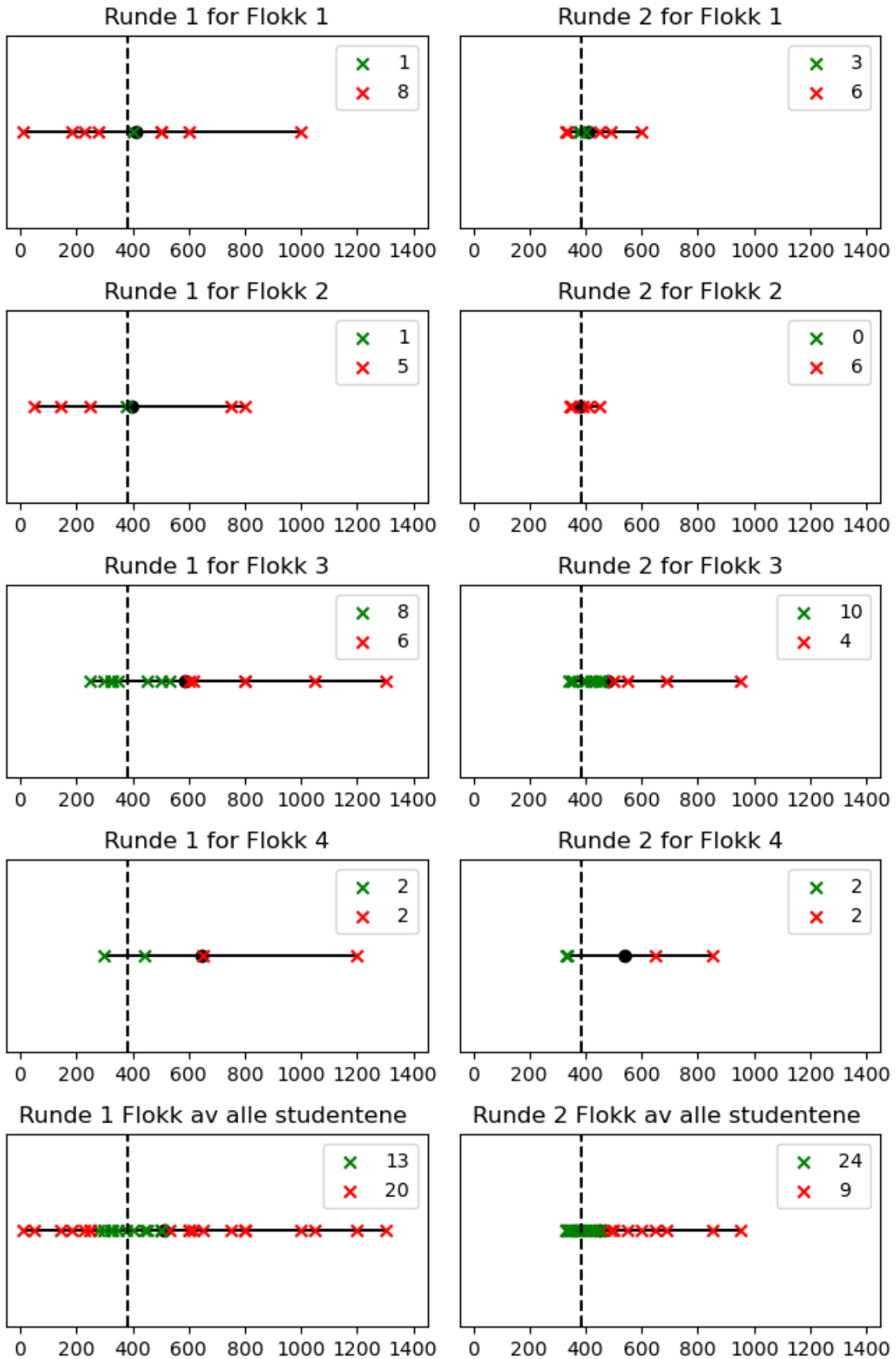


## 4.11 Individuelle estimat kontra flokkens samlede estimat

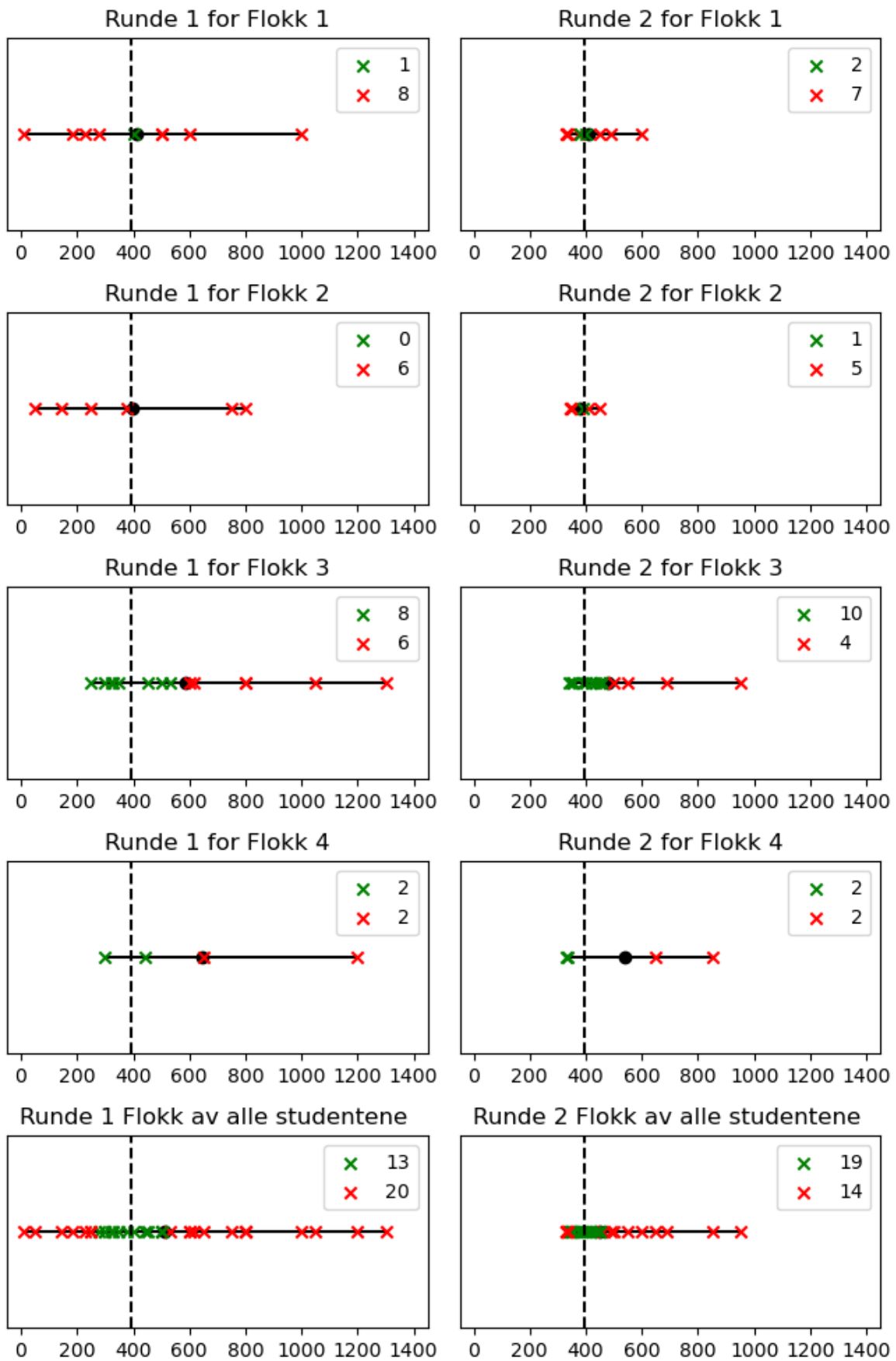
Et annet sentralt aspekt med wisdom of crowds er at det samlede estimatet til flokken ofte er mer presis enn de individuelle estimatene. For å finne ut om dette har vært tilfellet for flokkene av studenter i eksperimentet har de individuelle estimatene til studentene, og det samlede estimatet til flokken, begge blitt sammenlignet med gjennomsnittsprisen til entreprenørene og prisen til Entreprenør 3. Det samlede estimatet til flokkene av studenter ble uttrykt både som et aritmetisk gjennomsnitt og median. For å finne ut om det aritmetiske gjennomsnittet og/eller medianen hadde gitt mer presise estimat enn de individuelle estimatene var det nødvendig å undersøke følgende situasjoner for hver av flokkene:

- Individuelle estimat og gjennomsnittsestimatet til flokken i runde 1, målt mot entreprenørenes gjennomsnittspris
- Individuelle estimat og medianen til flokken i runde 1, målt mot entreprenørenes gjennomsnittspris
- Individuelle estimat og gjennomsnittsestimatet til flokken i runde 1, målt mot prisen til Entreprenør 3
- Individuelle estimat og medianen til flokken i runde 1, målt mot prisen til Entreprenør 3
- Individuelle estimat og gjennomsnittsestimatet til flokken i runde 2, målt mot entreprenørenes gjennomsnittspris
- Individuelle estimat og medianen til flokken i runde 2, målt mot entreprenørenes gjennomsnittspris
- Individuelle estimat og gjennomsnittsestimatet til flokken i runde 2, målt mot prisen til Entreprenør 3
- Individuelle estimat og medianen til flokken i runde 2, målt mot prisen til Entreprenør 3

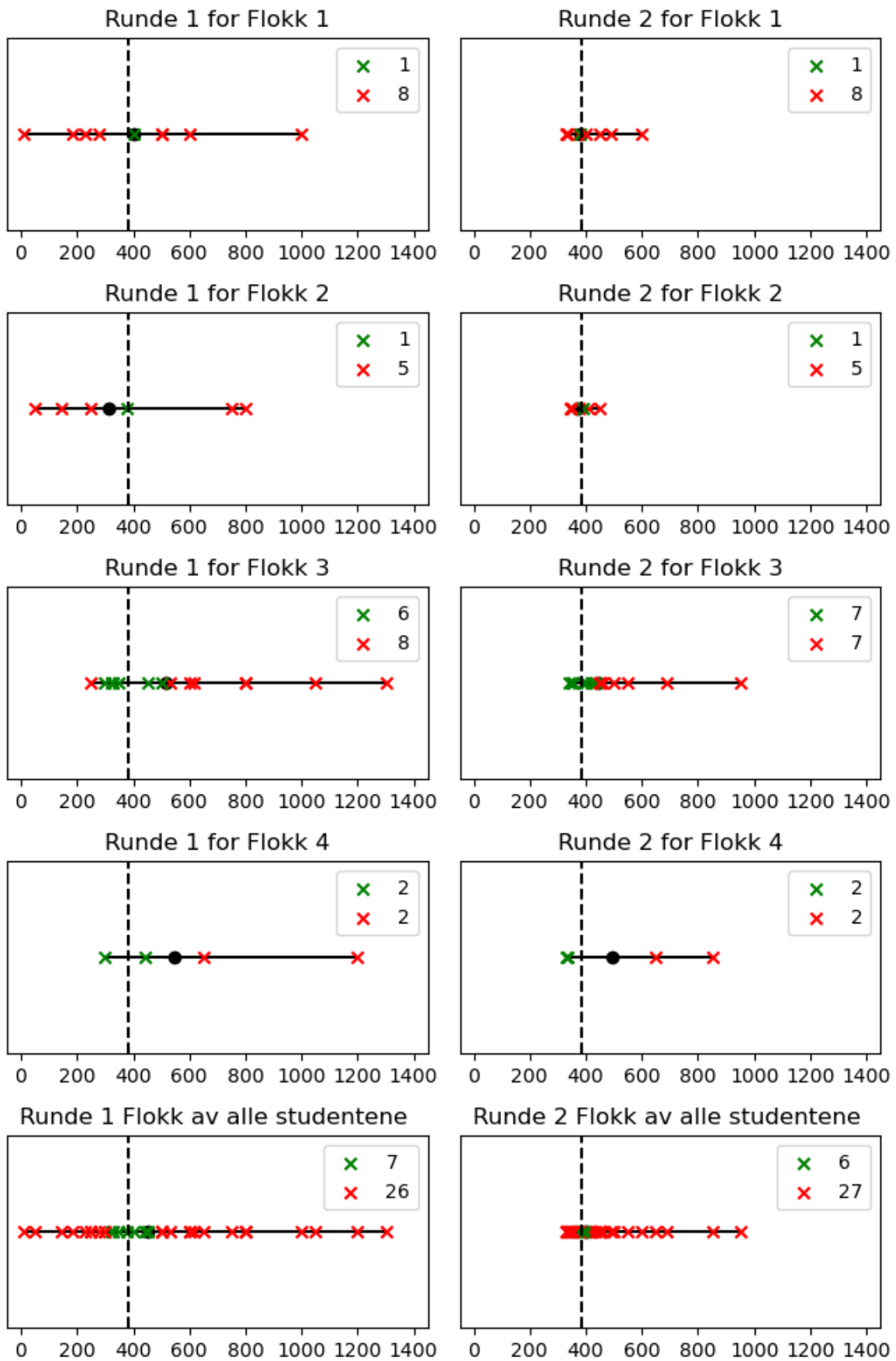
For hver av situasjonene listet opp ble det funnet ut hvor mange av de individuelle estimatene som var mindre presise enn det samlede estimatet. Dette ble gjort for hver av flokkene. Resultatene er gitt i Figur 13, Figur 14, Figur 15 og Figur 16. De røde kryssene representerer antallet individuelle estimat som var mindre presise enn det samlede estimatet til flokken, mens grønne kryss indikerer det motsatte. Det samlede estimatet til flokken er vist som en svart sirkel. Avhengig av hvilken situasjon det er sett på gir den vertikale stiplede linjen enten entreprenørenes gjennomsnittspris eller prisen til Entreprenør 3. Sirkelen og kryssene er fordelt på en svart horisontal linje som strekker seg fra den aktuelle flokkens laveste estimat til den aktuelle flokkens høyeste estimat. Alle situasjonene som er listet opp er undersøkt. Hver situasjon er oppgitt i figurtekstene.



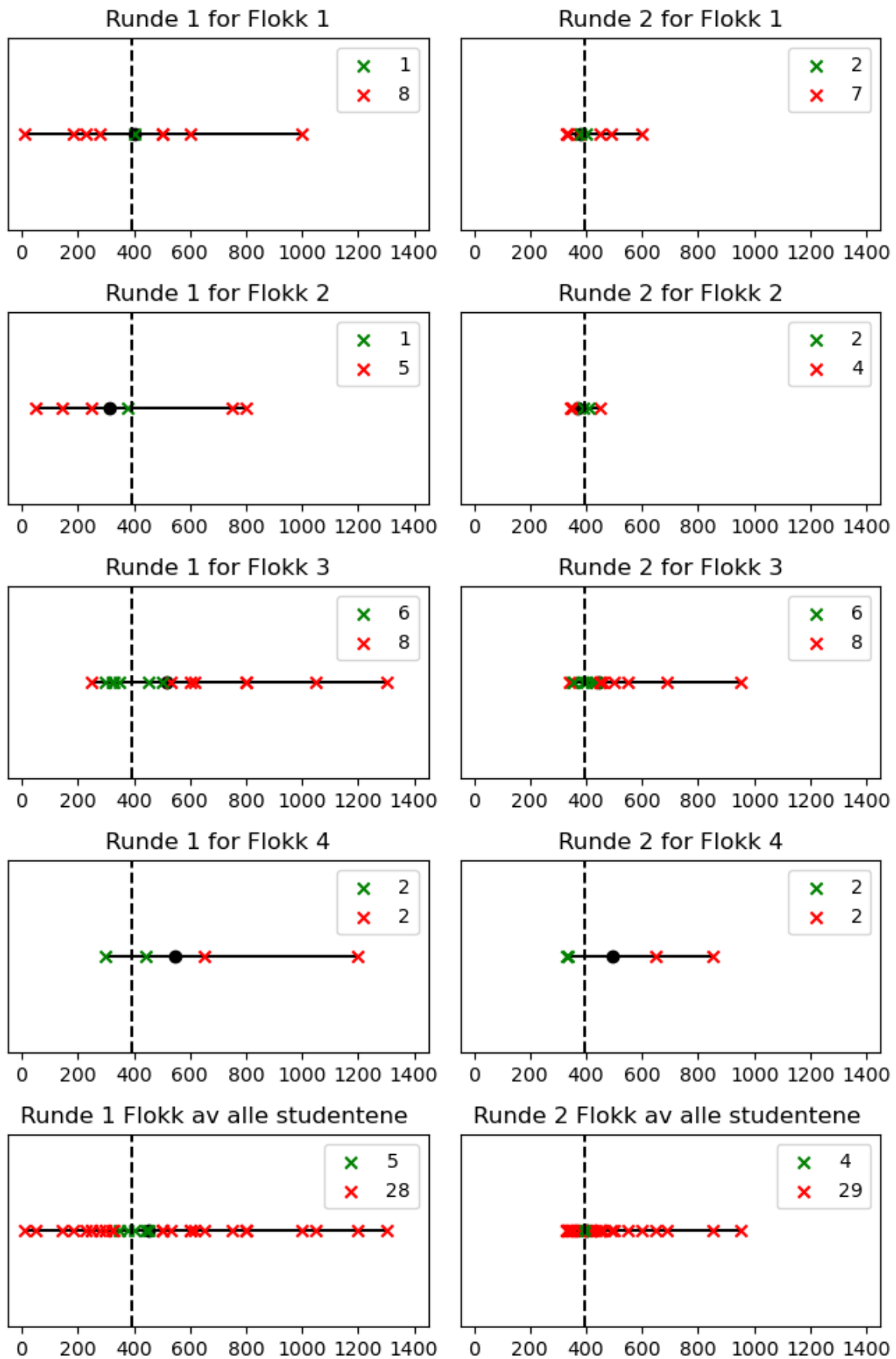
Figur 13: Individuelle estimat og gjennomsnittsestimatet til Flokk 1, Flokk 2, Flokk 3, Flokk 4 og Flokk av alle studentene i runde 1, målt mot entreprenørenes gjennomsnittspris.



Figur 14: Individuelle estimat og gjennomsnittsestimatet til Flokk 1, Flokk 2, Flokk 3, Flokk 4 og Flokk av alle studentene i runde 1, målt mot Entreprenør 3.



Figur 15: Individuelle estimat og medianen til Flokk 1, Flokk 2, Flokk 3, Flokk 4 og Flokk av alle studentene i runde 1, målt mot entreprenørenes gjennomsnittspris.



Figur 16: Individuelle estimat og medianen til Flokk 1, Flokk 2, Flokk 3, Flokk 4 og Flokk av alle studentene i runde 1, målt mot Entreprenør 3.

Fra Figur 13, Figur 14, Figur 15 og Figur 16 er det mulig å se at for de situasjonene der den svarte sirkelen er nær den stiplede linjen, så er antallet røde kryss som regel høyere enn antallet grønne kryss. Dette kan forstås som at for de situasjonene der det samlede estimatet til flokken er nærme entreprenørenes gjennomsnittspris, eller prisen til Entreprenør 3, så er antallet individuelle estimat som var mindre presise enn det samlede estimatet til flokken som regel høyere enn antallet individuelle estimat som var mer presise enn det samlede estimatet til flokken.

Den svarte horisontale linjen kan ses på som variasjonen til datasettet. Dersom variasjonen er lav, er den horisontale linjen kort. Er den horisontale linjen kort så har estimatene i datasettet høy grad av presisjon. Dersom den svarte sirkelen er i nærheten av den vertikale stiplede linjen, er det samlede estimatet til flokken nøyaktig. Følgelig har en flokk med en kort horisontal linje og sirkel nær den vertikale stiplede linjen høy grad av presisjon og nøyaktighet. Tilsvarende har en flokk med en lang horisontal linje og sirkel langt unna den vertikale stiplede linjen, lav grad av presisjon og nøyaktighet.

Figur 16 egner seg godt for å gi en visuell forståelse av fenomenet wisdom of crowds. For en situasjon der wisdom of crowds brukes er det ønskelig at den svarte sirkelen skal ligge i skjæringspunktet mellom den vertikale stiplede linjen og den horisontale linjen. For at dette skal være mulig er det først og fremst nødvendig at den vertikale stiplede linjen ligger oppå den horisontale linjen. Den horisontale linjen går som kjent fra den aktuelle flokkens laveste estimat til høyeste estimat. Dernest må sirkelen plasseres nærme den vertikale linjen. Plasseringen til sirkelen er avhengig av hvor kryssene settes på den horisontale linjen, og av hvilket statistisk oppsummeringstall som skal benyttes for å plassere sirkelen (det samlede estimatet til kryssene). I Figur 16 er median benyttet for å plassere sirkelen. Når median benyttes for å plassere sirkelen, vil sirkelen plasseres i midten av kryssene. For at sirkelen skal havne i skjæringspunktet mellom den vertikale linjen og den horisontale linjen er det nødvendig at det er like mange kryss på hver side av den vertikale linjen. På det viset kanselleres kryssene som er til venstre for den vertikale linjen mot kryssene som er til høyre for den vertikale linjen. Kryssene til høyre og venstre gir uttrykk for personlige bias.

Ekstra interessant for wisdom of crowds er det at effekten av at kryss kanselleres mot hverandre kan skje både for datasett der estimatene er presise og upresise. «Runde 2 for Flokk 2» og «Runde 2 Flokk av alle studentene» i Figur 16 gir et spesielt godt bilde på dette. Dette er også gjeldene for de gangene et gjennomsnitt benyttes som statistisk oppsummeringstall.

Forklaringen med at den vertikale stiplede linjen må ligge oppå den horisontale linjen, kan også være med på å forklare systematiske og personlige bias. Dersom den horisontale linjen av en eller annen grunn skyves vekk fra der den vertikale linjen er, kan årsaken til forskyvningen skyldes systematiske bias. En følge av dette er at alle kryssene flyttes vekk fra den vertikale linjen. Dette er naturligvis ikke ønskelig. Det som er ønskelig er at de personlige biasene på hver side av den vertikale linjen kan kanselleres mot hverandre.

## 5 Analyse

I dette kapittelet vil mine tolkninger av resultatene til eksperimentet presenteres. I kapittelet er det også gitt en forklaring på hvordan litteraturen i litteraturstudiet har blitt analysert.

### 5.1 Analyse av litteraturstudien

Litteraturen som har dannet det teoretiske rammeverket ble analysert ved hjelp av programmet NVivo. NVivo er et program som støtter kvalitative data og blandede forskningsmetoder (*NVivo - Kunnskapsbasen - NTNU*, u.å.). Den valgte litteraturen i litteraturstudien ble lastet ned som PDF-filer og så importert og lest i NVivo-programmets PDF-leser. I PDF-leseren til NVivo ble tekst markert og tildelt koder av meg. Strategien var å opprette koder til tema som var aktuelle for oppgavens problemstilling og forskningsspørsmål. Dersom ulike artikler tok for seg det samme tema, ble teksten som omhandlet temaet i hver av artiklene tildelt den samme koden. Kodene var ikke forhåndsbestemt, og ble dermed opprettet underveis i litteraturstudien. En samling av kodene finnes i en egen fane i NVivo. Under hver av kodene finnes all tekst som har blitt tildelt den samme koden. I dette vinduet vises også hvilket dokument den markerte teksten stammer fra. På det viset var det lett å finne tilbake til relevant informasjon om hvert av temaene. Årsaken til at jeg lagde kodene underveis var at det dermed var mulig å tilpasse hvilken informasjon som kunne markeres. Dersom det dukket opp interessant informasjon i litteraturen, som jeg ikke hadde tenkt på før søket, måtte disse blitt forkastet hvis kodene var låst på forhånd. At relevant tekst ble markert og kodet i NVivo har bidratt til å holde en god oversikt over alle funnene innenfor hvert av temaene. En slik samling har gjort det enkelt å finne tilbake til presis informasjon om spesifikke tema. Dermed har det ikke vært nødvendig å memorere, eller fysisk notere ned informasjonen underveis i litteraturstudiet. Dette har vært med på å gi en presis gjenfortelling av litteraturen.

### 5.2 Tolkning av eksperimentets resultatet

#### 5.2.1 Entreprenørenes priser som sammenligningsgrunnlag

I resultatene ble det slått fast at entreprenørenes priser er det nærmeste totalentreprisen for kontor- og undervisningsbygget kommer en pris. Basert på at tilbudene for totalentreprisen ble vurdert både på pris og kvaliteten på tilbudt personell, ble det besluttet å bruke gjennomsnittsprisen til entreprenørene og tilbudet til entreprenøren som vant anbudskonkurransen (Entreprenør 3), som prisen de andre estimatene skulle sammenlignes med.

For eksperimentet er det imidlertid viktig å påpeke at det aktuelle prosjektet ikke er bygget. Det er fortsatt en risiko for at jernloven kan gjøres gjeldende (Flyvbjerg, 2011). Til tross for at dette i utgangspunktet ikke er et megaprojekt, har historien vist at det finnes tilfeller av prosjekter som har utviklet seg til å bli det. Dermed vil det være feil å helt uten videre stemple de for høye estimatene til Flokk 3, Flokk 4 og supereksperter som dårlige. Spesielt gjelder dette for supereksperterens estimat. Supereksperter hadde 34 års erfaring med å estimere kostnaden for byggeprosjekt, og dette var ikke første gangen han hadde vært med å estimere prisen for et bygg.

Imidlertid er det en risiko for at supereksperter har estimert kostnadene for byggeprosjektet, og ikke bare prisen for totalentreprisen for kontor- og undervisningsbygget. Det vil si at det er en risiko for at han har gitt et estimat for alle midlene som går med for å realisere byggeprosjektet. Årsaken til dette er at det er byggekostnader supereksperter er ekspert på å estimere. I presentasjonen ble det imidlertid spesifisert at det var kun totalentreprisen for kontor- og undervisningsbygget det skulle gis et estimat for. Likevel kan tenke seg at supereksperter har vært så vant til å gi estimat for byggekostnadene at det er også byggekostnadene han har gitt et estimat for i dette eksperimentet. Likevel er estimatet for rammene gitt i dette eksperimentet det minst presise estimatet.

### 5.2.2 Median eller aritmetisk gjennomsnitt

Det ble i resultatkapittelet slått fast at medianen til estimatene til Flokk av alle studentene, Flokk 1, Flokk 3 og Flokk 4 har gitt et estimat som ligger nærmere entreprenørens samlede pris, og prisen til Entreprenør 3, enn hva et aritmetisk gjennomsnitt har gitt. Imidlertid ble det slått fast at det motsatte hadde vært tilfellet for Flokk 2. Resultatene har dermed vært divergerende for hvilket oppsummeringstall som gir resultatet med høyest presisjon. Hensikten med å bruke wisdom of crowds er å forutsi fremtidige hendelser. Følgelig vil det være behov for å vite hvilket oppsummeringstall som skal brukes før det er mulig å sammenligne med «fasiten», i dette tilfellet prisene til entreprenørene. Det var derfor interessant å se på om det er indikatorer eller teori som kan bidra til å ta det riktige valget før fasiten blir kjent.

I det teoretiske rammeverket er det gitt at Francis Galton, bakmannen til det første dokumenterte eksperimentet, foretrakk å bruke median (Wallis, 2014). Galton argumenterte fra et demokratisk synspunkt for at medianen var det målet som ga flest fornøyde deltakere. I forsøket med gjetting av vekten til oxen viste det seg dog at det aritmetiske gjennomsnittet ga et mer korrekt svar. På tidspunktet da estimatene ble gitt kan det hende at det var medianen som ga flest fornøyde deltakere, men hvilken rolle spiller dette dersom andre oppsummeringstall viser seg å være objektivt mer korrekt? Etter min mening ville de fleste akseptert å heller benytte seg av et aritmetisk gjennomsnitt dersom de visste at det ville føre til et mer riktig svar. Eksempelvis å gi en anbudspris som er nærmere de faktiske kostnadene. For meg fremstår det som sannsynlig at de involverte heller vil tjene noen ekstra kroner, mot det å være mer enig med estimatet som gis. Dermed bør det undersøkes om det finnes andre fundament enn demokrati som kan legges til grunn for å velge om et aritmetisk gjennomsnitt eller median skal benyttes.

Andre statistikere mener fordelingen til datamaterialet bør bestemme valget av statistisk oppsummeringstall (Frøslie, 2023). Frøslie (2023) sier at median bør benyttes for datasett hvor verdiene til datasettet er skjevfordelt, mens et aritmetisk gjennomsnitt egner seg for symmetriske datasett. Fra resultatene kom det frem at Flokk 2, Flokk 3, Flokk 4 og Flokken av alle studentene hadde en høyreskjev fordeling av estimat i runde 1, mens Flokk 1 hadde en symmetrisk fordeling av estimat i runde 1. For runde 2 hadde Flokk 3, Flokk 4 og Flokken av alle studentene en høyreskjev fordeling, mens Flokk 1 og Flokk 2 hadde en forholdsvis symmetrisk fordeling. Ved å anvende teorien om at medianen bør benyttes for skjevfordelte datasett og aritmetisk gjennomsnitt for symmetriske datasett, bør medianen benyttes for Flokk 2, Flokk 3, Flokk 4 og Flokken av alle studentene i runde 1, og det aritmetiske gjennomsnittet bør benyttes for Flokk 1 i runde 1. Med bruk av samme teori bør medianen benyttes for Flokk 3, Flokk 4 og Flokken av alle studentene i runde 2, og det aritmetiske gjennomsnittet bør benyttes for Flokk 1 og Flokk 2 i runde 2.



Dersom de teoretiske valgene av statistisk oppsummeringstall sammenlignes med hvilke oppsummeringstall som ga mest presise estimat i eksperimentet, kan det ses at de teoretiske valgene har stemt for nesten alle flokkene i begge rundene. Det er kun for estimatet til Flokk 1 i runde 1, og for Flokk 2 i begge rundene, at det teoretiske valget ikke har vært det mest presise valget av statistisk oppsummeringstall. For Flokk 1 var forskjellen mellom å velge det teoretiske oppsummeringstallet og oppsummeringstallet som viste seg å være mest presist 3,03 prosentpoeng og 3,13 prosentpoeng for henholdsvis entreprenørenes gjennomsnittspris og prisen til Entreprenør 3. For Flokk 2 var forskjellen mellom å velge det teoretiske oppsummeringstallet og oppsummeringstallet som viste seg å være mest presist i runde 1 20,01 prosentpoeng og 10,06 prosentpoeng for henholdsvis entreprenørenes gjennomsnittspris og prisen til Entreprenør 3. I runde 2 var forskjellen mellom å velge det teoretiske oppsummeringstallet og oppsummeringstallet som viste seg å være mest presist 3,14 prosentpoeng og 3,25 prosentpoeng, for henholdsvis entreprenørenes gjennomsnittspris og prisen til Entreprenør 3.

For dette eksperimentet kan ikke forskjeller på under 10 prosentpoeng sies å være nevneverdig feil. Derfor kan ikke valget av det teoretiske oppsummeringstallet sies å ha spilt en stor rolle for Flokk 1 i runde 1 og Flokk 2 i runde 2. Det er kun for Flokk 2 i runde 1 at valget mellom det teoretiske oppsummeringstallet og oppsummeringstallet som viste seg å være mest presist, har gitt et forholdsmessig «feil» svar.

Alt i alt virker det å velge medianen for skjevfordelte datasett og aritmetisk gjennomsnitt for symmetriske datasett å være en god taktikk for å få det mest presise estimatet fra et datasett. Taktikken har gitt det mest presise estimatet for nesten alle flokkene i begge rundene. Kanskje mest interessant er det at taktikken har gitt det mest presise estimatet for det største datasettet, Flokken av alle studentene. Resultatene fra eksperimentet harmonerer dermed godt med delen av teorien som sier at fordelingen til datamaterialet bør bestemme valget av statistisk oppsummeringstall.

### 5.2.3 Konsekvensene av for høye og for lave estimat

Til nå har det blitt sett på hvor nært ulike estimat prosentmessig har vært gjennomsnittsprisen til entreprenørene og Entreprenør 3. I byggebransjen er det imidlertid viktig å ta hensyn til den praktiske betydningen av å gi for høye, og ikke minst, for lave estimat.

Til tross for at Flokk 1 og Flokk 2 har gitt estimat som er nærme entreprenørenes priser er det ikke gitt at disse estimatene er gode i en praktisk sammenheng. Det samlede estimatet for Flokk 1 og 2 var under entreprenørenes gjennomsnittspris, og under prisen til Entreprenør 3. Fra en entreprenørs ståsted kan et for lavt estimat være verre enn et for høyt estimat. I byggebransjen blir det sagt at de dyreste prosjektene for entreprenørene er de prosjektene de vinner anbudskonkurransen for, men hvor de egentlig ikke har visst hva de har gitt tilbud på. En følge av dette er at tilbudet kan være for lavt priset fordi noe kostnadsdrivende er oversett. Kostnadene for prosjektet kan dermed øke for entreprenørens del, mens det ikke nødvendigvis er en økning i hva de får betalt. Som nevnt av sentrale aktører i byggebransjen kan dette føre til store tap, og i ytterste konsekvens konkurs (Offergaard, 2022, 2023). I en bransje preget av lave marginer er det dermed viktig å ikke gi for lave anbud. Imidlertid er det lite trolig at den negative forskjellen til Flokk 1 og 2 utgjør en risiko for konkurs. Likevel er det viktig å være klar over følgene av å gi estimat som er for lave.

Konsekvensene av å levere for høye anbud behøver ikke å være like alvorlige. Avhengig av tildelingskriteriene for prosjektet, kan en for høy pris føre til at entreprenøren ikke får jobben. Noe forenklet er tapet for denne situasjonen begrenset til hvor mye ressurser som ble benyttet for å levere anbudet. Dersom en bedrift gang på gang opplever å ikke få tildelt jobber vil dette naturligvis bli et problem. Likevel vil det være flere muligheter for å treffe med tilbudet sitt.

For høye og for lave estimat er også interessant fra en byggherres ståsted. Et for lavt estimat kan føre til at byggherren er nødt til å spørre prosjekteier om mer penger eller ikke kan gjennomføre prosjektet. Samset (2017) nevner at en overskridelse av kostnadene forbundet med prosjektet kan føre til at det aktuelle prosjektet etter overskridelsen er et dårligere prosjekt enn et annet prosjekt som heller kunne vært valgt. Et for høyt estimat kunne på den andre siden ført til at prosjektet i utgangspunktet ikke hadde fått finansiering, og dermed blitt forkastet. At Statsbygg hadde et høyere estimat enn gjennomsnittet til entreprenørene kunne i verste fall ført til at prosjektet ikke hadde fått finansiering. Imidlertid kan det fra resultatene ses at Statsbyggs estimat er under tilbudet til entreprenøren med det høyeste tilbudet. Det er lite trolig at overestimatet til Statsbygg kontra Entreprenør 3 ville satt en stopper for prosjektet. Ei heller kunne den ekstra summen med å bruke den dyreste entreprenøren medført at et annet prosjekt hadde vært et bedre alternativ.

#### 5.2.4 Størrelsens påvirkning på usikkerhet

Resultatene har vist at Flokk 1 (9 personer) og Flokk 2 (6 personer) har gitt samlede estimat i både runde 1 og runde 2 som er nærme entreprenørenes gjennomsnittspris og prisen til Entreprenør 3. Imidlertid har Flokk 3 (14 personer) og Flokk 4 (4 personer) gitt samlede estimat i både runde 1 og runde 2 som er et stykke unna entreprenørenes gjennomsnitt og Entreprenør 3. Dette peker i retning av at ved å bruke wisdom of crowds har det blitt gitt både presise og upresise estimat for mindre flokker. Resultatene sier ikke noe om hvor stor flokken bør være for en såkalt «mikroflokk». En «mikroflokk» ble i Gecer Ulu et al. (2019) omtalt som flokker med 5-14 personer. Flokk 2 har selv med relativt få deltakere gitt et presist estimat, mens Flokk 3 har med relativt mange deltakere gitt et upresist estimat. Dermed er det basert på dette eksperimentet vanskelig å komme med noe klart svar på hvorvidt fenomenet er gjeldende for mikroflokker.

En viktig antakelse brukt i resonnementet i eksperimentet, og det foregående avsnittet, har vært at alle studentene har hatt samme kompetanse. At Flokk 1 og Flokk 2 har gitt oppsiktsvekkende presise estimat i runde 1 og runde 2 kan skyldes at studentene i disse flokkene besitter en annen kunnskap om kostnadsestimering enn de andre studentene i eksperimentet. For å sjekke om dette kan ha vært tilfellet ble standardavviket til gjennomsnittet for Flokk 1 og Flokk 2 undersøkt. Standardavviket forteller oss hvor usikkert gjennomsnittsestimatet er. I delkapittel 4.4 ble standardavvikene til hver av flokkene sammenlignet. For runde 1 ble det funnet at hverken Flokk 1 eller Flokk 2 hadde lavere standardavvik enn flokken av alle studentene og Flokk 3. Dette viser at Flokk 1 og Flokk 2 ikke hadde sikrere estimat enn andre flokker i eksperimentet i runde 1. Dette taler for at Flokk 1 og Flokk 2 ikke har hatt en spesiell kompetanse som skiller dem fra de andre flokkene. Etter mitt syn fremstår det som at Flokk 1 og Flokk 2 enten har hatt flaks, eller at sammensetningen av flokkene har gitt ønskelig kanselleringsseffekt.

## 6 Diskusjon

Diskusjonen kommer til å bruke det teoretiske rammeverket for å diskutere resultatene fra eksperimentet, samt tolkningen av resultatene. Kapitlet er strukturert slik at forskningsspørsmålene belyses i kronologisk rekkefølge. Ved å svare på forskningsspørsmålene er hensikten å kunne svare på oppgavens problemstilling, «Hvordan kan wisdom of crowds brukes i dagens byggebransje».

1. *Hvordan kan digitale verktøy brukes for å hente inn data fra en flokk?*
2. *Hvordan bør sammensetningen til flokken være?*
3. *Hvilken informasjon bør flokken ha tilgjengelig?*
4. *Hvilke bias påvirker flokkens deltakere?*

### 6.1 Hvordan kan digitale verktøy brukes for å hente inn kunnskap fra en flokk?

Det har aldri vært større tilgang på data enn hva det er i dagens digitale hverdag. Ved hjelp av blant annet e-post, forum og skytjenester har det heller aldri vært lettere å dele informasjon, og komme i kontakt med ressurspersoner. Litteraturen har vist at det innenfor tematikken om wisdom of crowds har vokst frem flere digitale plattformer hvor det hentes inn svar på spørsmål fra en flokk mennesker (Gecer Ulu et al., 2019; Görzen, 2019; Hong et al., 2020). Plattformene kjennetegnes av at det ikke stilles krav til kompetanse hos deltakerne for å kunne delta. En følge av dette har vært at det har vært mulig å få tusenvis av svar fra ulike deltakere. Deltakerne på disse plattformene har stort sett blitt brukt til å besvare allmennfattelige spørsmål hvor det søkes etter et punkttestimat. Resultatene fra eksperimentet i denne masteroppgaven har vist at flere av flokkene av studenter har gitt svært presise estimat for en vanskelig oppgave, nemlig å anslå prisen for en totalentreprise.

Denne masteroppgaven har dog til hensikt å se på bruken av wisdom of crowds innenfor byggebransjen. Flere av oppgavene innenfor byggebransjen kjennetegnes av at de er esoteriske, det vil si det motsatte av allmennfattelige. Når det gjelder å løse esoteriske problem har det fra litteraturstudien kommet frem at det finnes lite forskning på hvor presise estimatene til deltakerne på de digitale plattformene er sammenlignet med folk med erfaring på området. Jeg har kun funnet én studie som omhandler dette temaet (Gecer Ulu et al., 2019). Konklusjonen fra Gecer Ulu et al. (2019) var at de uten kompetanse på området ikke var i nærheten av å gi presise estimat, mens de med litt kompetanse var i stand til å gi gode estimat.

Selv om det er lite forskning på om lekfolk er i stand til å gi nøyaktige estimater ved bruk av wisdom of crowds for en ingeniøroppgave, fremstår funnet til Gecer Ulu et al. (2019) etter min mening som logisk. Flere ingeniøroppgaver vil for allmennheten oppleves som svært vanskelig å besvare. Noen ganger kan det til og med være vanskelig å i det hele tatt forstå hva som skal besvares. En esoterisk oppgave kan for eksempel være å estimere den totale varmeoverføringen gjennom en vegg med 200 millimeter isolasjon, stående kledning av tre, vindsperre, dampspærre og gips. Det er usannsynlig at mannen i gata hadde klart å gi et fornuftig svar på spørsmålet. Følgelig fremstår det som at digitale plattformer som

ikke stiller krav til kompetanse hos deltakerne, ikke vil være særlig nyttig for å besvare spørsmål relatert til byggebransjen.

Hadde spørsmålet blitt stilt på et forum hvor kun bygningsfysikere hadde fått lov til å svare på spørsmålet, er det sannsynlig at deltakerne hadde kommet frem til det korrekte svaret. Etter hva jeg kjenner til finnes det ingen offentlig tilgjengelige plattformer hvor det er nødvendig for deltakerne å verifisere sin bakgrunn eller kompetanse. Et naturlig spørsmål vil dermed være om det er mulig å opprette tilsvarende digitale plattformer som er brukt i tidligere studier, men at deltakerne må gjennom en bakgrunnssjekk for å delta.

Det vil være flere utfordringer med å opprette en digital plattform som krever at deltakerne går gjennom en bakgrunnssjekk. For det første vil det kreve at deltakerne må sende inn dokumentasjon på kompetanse. For det andre må dokumentasjonen gås gjennom av en moderator, noe som vil være ressurskrevende. For det tredje må en nettside for plattformen opprettes, noe som også vil kreve ressurser. For det fjerde må de i målgruppen få vite om plattformen, noe som krever markedsføring. Problemene relatert til det rent praktiske med å sette opp og moderere plattformen kan relativt enkelt løses ved å sette av ressurser til dette. Det finnes millioner av nettsider i verden, og dermed er denne problematikken ansett å være relativt enkel å overkomme. Dog kan det være vanskeligere å få den ønskede målgruppen til å delta. En mulig løsning kan være å tilby en pengesum til de som registrerer seg og deltar. Virkemiddelet med å tilby en finansiell kompensasjon til deltakerne er et vanlig insentiv for at flere personer skal melde seg til å utføre oppgaver. Det er sannsynlig at det ville fungert her også. Likevel er det viktig å påpeke at den økte ressursbruken disse tiltakene medfører må være lavere enn inntjeningene som følge av estimatene deltakerne kommer med. Alt tatt til betraktning vil det være mulig å skaffe kompetente deltakere, samt å sette opp en nettside for plattformen.

Likevel vil en offentlig tilgjengelig plattform være problematisk i kraft av at den er offentlig. Byggebransjen er i stor grad preget av konkurranse. Dersom en bedrift setter opp en offentlig tilgjengelig plattform kan det føre til at konkurrenter får tilgang til informasjon bedriften ikke ønsker å dele. En mulig løsning på dette kan være å kreve at brukerne må logge inn for å bruke nettsiden, og at svarene kun er synlig for moderatorene. Problemet med en slik løsning vil imidlertid være at bedriften ikke har kontroll over hva brukerne gjør med informasjonen etter innlogging. At informasjon videresendes til konkurrenter, kan ikke utelukkes. Selv om deltakerne kun får se spørsmål, så kan spørsmålet i seg selv inneholde verdifull informasjon. Eksempelvis kan spørsmålet dreie seg om noe en konkurrent ikke har tenkt på. Dersom plattformen kun brukes til å stille forholdsvis enkle spørsmål, slik som spørsmålet stilt i denne masteroppgavens eksperiment: «Hvor mye koster dette prosjektet?», vil ikke spørsmålet gi noen informasjon til konkurrentene. Imidlertid vil det faktum at siden er tilgjengelig for folk utenfor bedriften medføre en risiko for at deltakere kommer med ondsinnede forslag.

I eksempelet som ble presentert i innledningen av masteroppgaven brukte entreprenøren bare folk fra egen virksomhet. Dette fører til at problemet med ondsinnede forslag minimeres. Utover dette vil en slik tilnærming innebære at deltakernes kompetanse allerede er vurdert gjennom en ansettelsesprosess og arbeid de tidligere har utført. Videre kan bedriften pålegge deltakelse som en del av jobben den ansatte skal gjøre. Dermed fremstår det som en god tilnærming å bruke personer allerede ansatt i firmaet.

Likevel ligger det et uforløst potensial knyttet til digitalisering i metoden entreprenøren beskrev. I eksempelet ble de ansatte fysisk samlet for å gi estimat. Fra det teoretiske rammeverket er det gitt at disse ressurspersonene ofte kan være vanskelig å fremskaffe

til samme tid (Torp et al., 2015). Dette kommer av at de er travle mennesker, ofte med midlertidige arbeidsplasser. Som et alternativ til at ressurspersonene møtes fysisk foreslås det fra min side at ressurspersonene gir estimat ved hjelp av digitale verktøy. I eksperimentet i denne masteroppgaven ble det holdt en presentasjon på nett ved hjelp av Mentimeter. Presentasjonen ble gjort tilgjengelig for deltakerne ved at de åpnet Mentimeter i nettleseren på en personlig enhet og skrev inn en unik kode. For en bedrift kan koden for eksempel gjøres tilgjengelig på intranettet til bedriften. Deretter kan en administrator starte presentasjonen og gå gjennom lysbilder med informasjon. Til slutt kan deltakerne gi estimat ved å taste inn svar på de personlige enhetene. Lysbildene vil til enhver tid være tilgjengelig på deltakernes enheter gjennom presentasjonen. En slik digital presentasjonsteknikk inneholder riktignok styrker og svakheter.

### *Styrker*

At det er mulig å besvare presentasjonen fra hvor som helst i verden hvor det er internett er en styrke. Dette fører til at deltakerne ikke er avhengig av å forflytte seg for å delta, og at eventuelle reisekostnader unngås. En annen styrke er at all informasjonen deltakerne trenger er gitt på deres egne skjermer. Dette fører til at deltakerne ikke er avhengig av å lete etter relevant data. Det er heller ikke en risiko for at støy i kommunikasjonskanalen fører til at informasjonen på lysbildene ikke når frem til deltakeren. At en administrator er til stede gjør det også mulig for deltakerne at eventuelle spørsmål kan besvares umiddelbart.

### *Svakheter*

Det faktum at en administrator må være til stede for å gjennomføre en presentasjon med Mentimeter kan også ses på som en svakhet. Svakheten ligger i at det enten er nødvendig å samle alle ressurspersonene til samme tid eller at administratoren må holde presentasjonen flere ganger. At administratoren måtte holde presentasjonen flere ganger var tilfellet for eksperimentet i denne masteroppgaven. Problemet med at en presentasjon holdes flere ganger er at det åpner opp for feilkilder i form av at deltakerne kan få ulike svar av administratoren, eller at noe blir presentert på ulikt vis. Til tross for at ressurspersonene ikke trenger å møtes fysisk, er det fortsatt en risiko for at timeplanen til deltakerne ikke samsvarer med hverandre.

En alternativ løsning til å holde en digital presentasjon med en administrator er å sende ut en spørreundersøkelse. Her må det imidlertid foretas en vurdering rundt viktigheten av at en administrator er til stede. Ettersom omtrent alle prosesser i byggebransjen er begrenset av tidsfrister, er det nødvendig at en spørreundersøkelse besvares innen en gitt frist. En administrator vil kunne sørge for at alle deltakerne på møtet gir et svar. For en spørreundersøkelse finnes det ingen slike sikkerhetsnett, og det er sannsynlig at spørreundersøkelsen ikke besvares. Likevel er det ikke gitt at det er en fiasko dersom for eksempel 29 av 33 personer gir et svar. Eksperimentet har vist at selv flokker med et mindre antall deltakere har gitt presise estimat. Dessuten vil en digital spørreundersøkelse gjøre at deltakerne selv bestemmer når de vil besvare den. Dermed er det mulig å besvare spørreundersøkelsen når hver av deltakerne har ledig tid. Dersom det oppstår ledig tid på kort varsel, for eksempel hvis et møte blir avlyst, kan den ledige tiden benyttes til å svare på spørreundersøkelsen. Flexibiliteten til når spørreundersøkelsen kan besvares øker sannsynligheten for at flere deltakere kan gi estimater.

Resultatet fra forsøket peker i den retning at studenter med basisferdigheter innenfor byggebransjen er i stand til å estimere prisen for en totalentreprise for et kontor- og

undervisningsbygg. Dermed bør muligheten for å inkludere estimat fra uerfarne personer utforskes. Ettersom en digital presentasjon ikke utløser et behov for at deltakerne må forflytte seg, er verktøyet godt egnet for å utvide størrelsen på flokken, og nå ut til flere personer.

## 6.2 Hvordan bør sammensetningen til flokken være?

Litteraturen innenfor wisdom of crowds har pekt på størrelse og mangfold som bidragsyttere for å få presise estimat (Budescu & Chen, 2015; Galesic et al., 2018; Gecer Ulu et al., 2019; Goldstein et al., 2014; Hong et al., 2020; Keuschnigg & Ganser, 2017; Larrick, 2016; Mannes et al., 2014; Nguyen & Nguyen, 2017). Resultatene til eksperimentet har vist at flokker av studenter med ulikt antall personer har gitt presise estimat. Fra resultatene og analysen er det kommet frem at usikkerheten for estimatene til studentene reduseres for en økende flokkstørrelse. Imidlertid er det vist at selv mikroflokker har gitt svært presise estimat. For problemstillingen vil det være interessant å videre diskutere hvordan sammensetningen av flokkene har bidratt til å gi presise estimat. Ettersom byggebransjen er en resultatorientert bransje er det interessant å diskutere resultatene opp mot ressursbruk.

### *Ressursbruk*

For situasjoner hvor flokker i wisdom of crowds benyttes er det knyttet ressursbruk til både hvem som deltar, og hvor mange som deltar. Som regel vil det være mer ressurskrevende å bruke eksperter enn uerfarne personer, for eksempel en nyutdannet person. Ved å anta at alle eksperter er like ressurskrevende vil en dobling av antall eksperter medføre en dobling av ressursbruken. Samme prinsipp gjelder for ferske på området. Imidlertid vil en dobling av ressursbruken for eksperter medføre en høyere total ressursbruk, enn hva en dobling av ressursforbruket for nyutdannede medfører. For eksempel, dersom én ekspert koster 2 kr, og en nyutdannet koster 1 kr, vil det være dobbelt så ressurskrevende å bruke én ekspert enn én nyutdannet. For prisen av én ekspert er det mulig å få to nyutdannede. Følgelig vil det for en beslutningstaker være interessant å se på kostnadene av å bruke flere eksperter kontra enda flere nyutdannede. En slik situasjon kan gjerne betegnes som en dragkamp mellom å legge til ekstra ekspertise kontra å legge til størrelse. Budescu & Chen (2015) beskrev som nevnt i underkapittel 3.1.2 denne forskjellen som forskjellen på å basere seg på kvalitet (ekspertise) eller kvantitet (flokker). I eksperimentet ble ytterpunktene av kvalitet og kvantitet utforsket. Superexperten representerte kvalitet og flokken av studenter representerte kvantitet. Resultatene fra eksperimentet pekte i retning av at ekspertise ikke er en garanti for suksess.

At studentene i eksperimentet har gitt estimat som er svært nærme entreprenørenes priser er interessant. Selv om spørsmålet stilt i eksperimentet i seg selv er enkelt, er det ikke enkelt å svare korrekt på spørsmålet. At studentene har truffet såpass godt med sine samlede estimat peker i den retning av at flokker av uerfarne kan brukes for å svare på estimeringsoppgaver. Tradisjonelt sett har kun eksperter blitt brukt for å svare på estimeringsoppgaver.

### *Størrelse*

Resultatene fra eksperimentet viste en klar tendens for at usikkerheten forbundet med estimatene gitt av flokker bestående av studenter er synkende for et økende antall studenter. I resultatene ble det trukket frem at usikkerheten gikk ned kjappast av å legge til ekstra personer i flokkene med få personer. Imidlertid vil den totale usikkerheten være

reduisert mer for flokker med flere personer. Basert på dette resultatet mener jeg at størrelsen for flokken bør justeres etter hvor presist det er ønskelig at estimatet skal være. For store prosjekter er det som regel større risiko enn for små prosjekter. Følgelig vil det være ønskelig å gi mer presise estimat for store prosjekter enn små prosjekter. Dersom wisdom of crowds benyttes som en fremsynsmetode bør flokkene være størst for de største prosjektene.

Som diskutert tidligere kan deltakerne enkelt involveres ved hjelp av en digital presentasjon eller et digitalt spørreskjema. Dermed kan disse digitale verktøyene være med på å justere flokkstørrelsen. For store prosjekter kan koden til presentasjonen, eller spørreskjemaet, sendes ut til flere personer enn for mindre prosjekter.

Fra resultatene kom det frem at mikroflokker har gitt presise estimat. Resultatene kan holdes frem som eksempler for at det med wisdom of crowds er mulig å oppnå presise estimat for prisen til byggeprosjekt, selv med få personer. Imidlertid ble det fra tolkningen av resultatene observert at det var knyttet større usikkerhet til estimatene gitt av mikroflokkene, enn for flokken av alle studentene. Tolkningen av resultatene pekte på flaks og en god kansellerings effekt som mulige årsaker til det presise estimatet.

Til tross for at de gode estimatene til mikroflokkene kan ha vært avhengig av flaks, mener jeg at man ikke bør se helt vekk fra resultatene av den grunn. Noen ganger er det flaks man behøver for å lykkes med byggeprosjekt. Hvis det finnes en mulighet for at mikroflokker kan bidra til å gi presise estimat, så bør muligheten utnyttes. I eksperimentet brukte deltakerne i mikroflokkene under 30 minutter på å gi estimat. For meg fremstår det som sannsynlig at dersom 3 - 15 tilfeldige ingeniører i løpet av 30 minutter kan være med på å øke presisjonen til et kostnadsestimat, så er det noe de fleste beslutningstakere bør være interessert i å gjøre. Om ikke estimatene til mikroflokkene brukes alene, så kan estimatene fra mikroflokkene brukes som et supplement til andre fremsynsmetoder.

### *Mangfold*

Tidligere i diskusjonen har det blitt vist at det ved en tilfeldig utvelgelse av estimat fra et datasett bestående av estimat gitt av studenter, er mulig å få samlede estimat med høy presisjon. Estimaten her har dog hatt høy grad av usikkerhet. En viktig antakelse brukt i resonnetet har vært at deltakerne i flokkene av studenter har hatt samme kompetanse. Antakelsen om at det er lite som skiller studentene kan begrunnes med at alle er studenter, de studerer på samme universitet og innehar begrenset erfaring fra arbeidslivet.

I den virkelige verden er det likevel langt flere ting som skiller personer fra hverandre. For eksempel kjønn, arbeidsoppgaver i sommerjobber og alder. Flere deler av litteraturen har sett på om det er mulig å identifisere eksperter innenfor flokken. Resultatene fra disse studiene har ikke gitt noe klart svar på om det er mulig eller ikke (Budescu & Chen, 2015; Mannes et al., 2014). Imidlertid er det ikke gitt at det samlede estimatet til en flokk bestående av utelukkende eksperter, er bedre enn en flokk av amatører. For eksempel, hvis et bygg koster 1 milliard kroner, og to eksperter begge anslår prisen til å være 1,3 milliarder kroner, mens to amatører anslår prisen til å være henholdsvis 500 millioner kroner og 1,5 milliarder kroner, så vil det samlede estimatet til amatørerne være bedre enn ekspertene. Å benytte seg av bare eksperters forslag for en estimeringsoppgave er noe som kjennetegner tradisjonelt gruppearbeid og Delphimetoden, ikke wisdom of crowds.

Derfor er det etter mitt syn ikke så interessant å se på hvordan det er mulig å identifisere eksperter innad i flokken når det er snakk om wisdom of crowds. Styrken til wisdom of crowds ligger i flokkens mangfold av meninger og estimat. Mangfoldet av estimatene kan uttrykkes med spredningen av estimat. I resultatene ble det illustrert hvordan de for høye estimatene kansellerer de for lave estimatene, og at det riktige svaret ofte ligger i midten av dem. Følgelig vil det etter mitt syn heller være nyttig å prøve å sette sammen en flokk som kan dra nytte av kanselleringseffekten.

Dessuten er det vist at det er den enkeltstående eksperten som har gitt estimat lengst unna entreprenørene. Dette peker i retning av at det kan være lurt å ikke stole helt og holdent på én person når det kommer til en estimeringsoppgave tilsvarende den i eksperimentet. Eksperten kunne muligens selv dratt nytte av en flokk. Statsbyggs estimat, som er kommet i stand ved bruk av flere eksperter, er for eksempel like nært entreprenørenes samlede estimat som Flokken av alle studentene.

#### *Oppgaver som løses best med ekspertise*

Til tross for at flere av flokkene av studenter har gitt presise estimat i eksperimentet, er det klart at flokker ikke kan benyttes for alle arbeidsoppgavene i byggebransjen. Dersom en akutt situasjon skulle oppstå er det nødvendig å gi et godt svar hurtig. For eksempel om forskalingen sprekker under en støp. For denne situasjonen vil det være lite hensiktsmessig å prøve å samle inn forslag fra mange personer. Det er en risiko for at betongen allerede er fast før alle personene har rukket å svare. Å kontakte en eller to eksperter, og handle etter hva de anser å være den beste løsningen, er etter mitt syn et mye bedre alternativ enn å bruke en flokk. Svaret til ekspertene trenger ikke å være perfekt. Konsekvensene av å ikke handle, eller handle for sent, kan være mye verre.

Litteraturen har påpekt at effekten av wisdom of crowds er størst for velstrukturerte oppgaver, slik som oppgaven i eksperimentet (Gecer Ulu et al., 2019; Graf-Drasch et al., 2022). Imidlertid vil ekspertise gjøre det bedre enn flokker for ustrukturerte oppgaver, slik som å svare på hvordan betongsøl bør håndteres.

#### *Følger av eksperimentets resultater for størrelse og mangfold i fremtidige wisdom of crowds forsøk*

Basert på eksperimentets resultater, og drøftingen av resultatene, kan det virke som at kvalitet (ekspertise) vinner dragkampen for ustrukturerte oppgaver, og kvantitet (flokker) vinner dragkampen for ustrukturerte oppgaver. Når det gjelder fremtidige eksperimenter med wisdom of crowds, mener jeg at det er nødvendig å undersøke hvilke sammensetninger av deltakere som bidrar til en best mulig kanselleringseffekt.



### 6.3 Hvilken informasjon bør flokken ha tilgjengelig?

I eksperimentet fikk studentene og supereksperter kun tildelt overordnet informasjon om prosjektet. Presentasjonen inneholdt bilder av modellen til bygget, lokasjon, formålet med bygget og bruttoareal. Presentasjonen kan ses i sin helhet i Bilag 2. Deler av litteraturen har argumentert for at estimeringsprosesser for byggeprosjekt ofte tar utgangspunkt i et innsideperspektiv, mens det heller burde vært benyttet et utsideperspektiv (Flyvbjerg, 2006). Utsideperspektivet innebærer at man ser utover på andre prosjekt, i motsetning til å se innover på sitt eget prosjekt. En konsekvens av utsideperspektivet er at de som skal gi estimat får informasjon med få detaljer. Et innsideperspektiv innebærer flere detaljer, og som regel mye informasjon. Samset (2021) har pekt på at kildene til informasjon, slik som rapporter, stadig blir mer detaljerte og lange. Flaskehalsen for informasjonsbehandlingen ligger i ressurspersonenes kapasitet. Dermed vil det være interessant å finne ut hvor lite informasjon ressurspersonene behøver for å kunne gi presise estimat.

Basert på informasjonen supereksperter og studentene fikk i den første runden kan det konkluderes med de har hatt en grov top-down-tilnærming i runde 1. I den andre runden hadde supereksperter og studentene en kombinasjon av top-down-tilnærming og utsideperspektiv, ettersom priser for andre prosjekt ble gjort kjent mellom rundene.

Supereksperter valgte å ikke endre på sitt estimat mellom runde 1 og runde 2. Dette vitner om at opplysninger om tidligere priser ikke har påvirket supereksperter. Supereksperter har gjennom sitt arbeid vært med på å utarbeide prisene som ble presentert. Informasjonen om pris var derfor godt kjent for supereksperter på forhånd. Derfor er det kanskje ikke så rart at han valgte å beholde estimatet sitt. At supereksperter på forhånd visste om priser for andre prosjekt, vitner også om at han gikk inn i eksperimentet med større grad av kunnskap enn studentene. Troen på egen kunnskap er muligens det som har gjort at supereksperter har valgt å se bort fra prisene som ble gitt. Tidligere litteratur har pekt på at nettopp eksperter ofte er med selvsikre (Mannes et al., 2014; Minson & Mueller, 2012). For supereksperter kan det hende at selvsikkerheten har ført til at han har sett bort ifra informasjon, selv om han egentlig burde vurdert den. Likevel må det igjen fremholdes at han kan ha gitt et estimat for hele byggeprosjektet.

Resultatene viste at informasjonen om pris for andre prosjekt var med på at studentene endret estimat fra runde 1 til runde 2. I 18 av 20 tilfeller førte endringen av estimat til at det samlede estimatet for de ulike flokkene ble mer presist i runde 2 enn runde 1. En av årsakene til at estimatet til de fleste flokkene ble mer presist, var at flere av de høyeste estimatene ble justert ned. Tidligere litteratur har sett på konsekvensene av nettopp det å endre eller fjerne ekstremalverdier fra et datasett (Nguyen & Nguyen, 2017). I Nguyen & Nguyen (2017) har det blitt konkludert med at å endre ekstremalverdier ikke gir mer presise samlede estimat. For dette eksperimentet har konklusjonen fra den tidligere studien kun vært gyldig for to tilfeller.

En mulig forklaring på hvorfor færre ekstremalverdier ga et bedre resultat i dette eksperimentet, var at det er mer rom for å gi estimat med ekstreme verdier oppover enn nedover. Gjetter noen 100 % over den faktiske prisen må noen gjette 0 kr for at de kansellerer hverandre. Det er ingen øvre grense, kun en nedre grense. Dette kan også holdes frem som en årsak til at estimatene har blitt skjevfordelt.

At informasjonen om pris endret estimatet til de fleste studentene kan vitne om at studentene i utgangspunktet ikke hadde nok informasjon, eller kunnskap, for å svare på

spørsmålet i eksperimentet. Poenget kan videre underbygges med at supereksperterens estimat er lengst unna entreprenørenes priser. Supereksperter har vært borti mange estimeringsoppgaver tidligere. At han er lengst unna kan vitne om at han ikke har fått nok informasjon til å gi et presist estimat. Fra tidligere litteratur har det ikke kommet noe entydig svar på hva som utgjør tilstrekkelig informasjon for estimeringsprosesser i byggebransjen. Derfor har det vært vanskelig å bedømme hvilken informasjon flokken i wisdom of crowds bør få for å kunne gi presise estimat.

Av informasjonen som ofte er aktuell i andre estimeringsprosesser enn i de hvor wisdom of crowds benyttes, trekkes erfaringstall frem. Fra underkapittel 3.2.1 kommer det imidlertid frem at Ole Jonny Klakegg i Del 3 kapittel 8 i Sunnevåg (2007) skriver at det bør gjøres tilpasninger for erfaringstallene, dersom erfaringstallene skal brukes til å gi estimat andre prosjekt enn for prosjektet erfaringstallene er hentet fra.

På bakgrunn av at det foreligger lite konkret litteratur om hva som utgjør tilstrekkelig informasjon i estimeringsprosesser hvor wisdom of crowds brukes, ble supereksperter i eksperimentet spurt om hva han anså å være den viktigste nøkkelinformasjonen. Listen er gitt i delkapittel 4.10. Ved å sammenligne listen til eksperter med informasjonen som ble gitt i presentasjonen er det kommet frem til at presentasjonen inneholdt kun 5 av 12 punkt. Disse var bygningsgeometri, bruttoareal, byggherreorganisasjon, nettoareal og tilgang på arealer for rigg og drift. Dette peker i retning av at informasjonen studentene og supereksperter fikk i den første runden ikke har vært tilstrekkelig.

På den annen side så hadde flere av flokkene av studenter presise estimat, selv i den første runden. Som diskutert tidligere inneholdt disse estimatene riktignok en høyere grad av usikkerhet enn entreprenørenes gjennomsnittspris. Likevel kan resultatene holdes frem som eksempel på at det er mulig å gi gode estimat selv med informasjonen gitt i runde 1. Utover dette har bruken av bilder fra BIM-modeller gjort det lettere enn noensinne å ha en top-down-tilnærming. Ved å se bilder fra en 3D-modell er det lettere å danne seg en formening om omfanget av det fremtidige arbeidet, sammenlignet med å måtte sette seg inn i tekniske 2D-tegninger.

Likevel er det viktig at personene som ser på tegningene, eller modellen, vet hva den viser. Kunnskap oppnås først når informasjonen blir forstått. Eksperter har vanligvis mye kunnskap fordi de har en dyp forståelse av informasjonen, mens det motsatte kan bli sagt for uerfarne. Følgelig kan det virke som at supereksperter vil dratt mer nytte av mer informasjon enn studentene ville gjort.

Flokken bestående av alle studentene hadde imidlertid sitt mest presise estimat etter de fikk informasjon om prisen til andre prosjekt. Likevel vil det være vanskelig å si noe om hvorvidt informasjon om pris alene er den viktigste. Dette skyldes blant annet at supereksperter og studentene i forkant av opplysninger om pris hadde mottatt annen informasjon. Dessuten ville informasjonen om pris og størrelse isolert sett gitt estimat som var mindre presise enn det den samlede flokken av studenter ga. Følgelig vil det være vanskelig å si om det er gjennomsnittsprisen fra andre prosjekt som er viktigst eller om det er en kombinasjon av priser og annen informasjon. Likevel kan det fra eksperimentet observeres at opplysningene om tidligere priser har hatt en innvirkning på estimatene til studentene. For et fremtidig eksperiment hadde det vært interessant å se på forskjellen mellom å gi all informasjonen på en gang, og det å splitte den opp.

Uavhengig om det er gitt tilstrekkelig med informasjon i dette eksperimentet, kan det fremheves at informasjonsmengden i stor grad minner om det informasjonsbildet som

foreligger tidlig i et prosjekt. For eksempel ved kvalitetssikring av konseptvalg (KS1). I denne fasen er det meste ved prosjektet ukjent, kun de overordnede linjene er kjent. Derfor er det heller ikke detaljer å ta hensyn til, ettersom de ikke finnes. I løpet av forprosjektet blir flere detaljer kjent, og det foreligger en mer komplett beskrivelse ved KS2. Welde (2016) har vist at kostnadsestimatene gitt i KS1 i liten grad har harmonert med kostnadsestimatene i KS2. Dette er problematisk ettersom Welde et al. (2019) har konkludert med at estimatene i KS2 stemmer godt med de faktiske kostnadene. Fra eksperimentet i denne masteroppgaven er det vist at det er mulig å gi presise estimat, til tross for lite informasjon. Derfor er wisdom of crowds etter min mening en fremsynsmetode som bør kunne benyttes for å gi bedre kostnadsestimat i KS1.

Fremsynsmetoden wisdom of crowds vil også være egnet for andre situasjoner hvor det foreligger lite informasjon, og det er ønskelig med et kjapt overslag. For eksempel kan en entreprenør kjapt finne ut om en jobb er innenfor den størrelsesorden av prosjekter som entreprenøren har lyst å levere tilbud for. Overslaget kan dermed være med på å øke sannsynligheten for at entreprenørene velger rett type prosjekt. Med å ha en sum å forholde seg til fra starten av, kan man unngå å levere tilbud for prosjekter man egentlig ikke har peiling på hva innebærer.

## 6.4 Bias som påvirker personene som gir estimat

Personer i estimeringsprosesser er som regel påvirket av ulike bias. Biasene kan være bevisst eller ubevisst. Som nevnt tidligere kjennetegnes wisdom of crowds av at de dårlige forslagene kansellerer hverandre. De for høye estimatene kansellerer de for lave estimatene, og sannheten ligger sannsynligvis en plass mellom dem. Dette delkapittelet vil se på hvilke bias som er observert i eksperimentet i masteroppgaven. Videre vil delkapittelet ta for seg hvilke bias tradisjonelle estimeringsprosesser støter på, og diskutere dem opp mot wisdom of crowds.

### *Generelt om bias for wisdom of crowds*

For wisdom of crowds er det ønskelig at deltakerne skal inneha personlige bias, men ikke systematiske bias (Larrick & Soll, 2006). En visuell forklaring av dette ble gitt i delkapittel 4.11. Resultatene viste at studentene i dette eksperimentet ikke hadde systematiske bias, men flere hadde personlige bias. Studentenes personlige bias kan knyttes til ulike kjente bias.

### *Forankringsbias*

At noen av studentene og den enkelte eksperten i eksperimentet har gitt relativt høye estimat kan til dels forklares med et forankringsbias. Prosjektet de skulle gi estimat for var isolert sett ukjent for dem. Konteksten til prosjektet, som er Norsk havteknologisenter, må imidlertid være å anse som kjent. Norsk havteknologisenter har fått mye oppmerksomhet i mediene, med flere oppslag om totalkostnaden for hele prosjektet. Summene nevnt i artiklene har vært flere milliarder kroner høyere enn de kostnadene det er estimert at det spesifikke prosjektet i dette eksperimentet vil koste. Til tross for at det er en åpenbar forskjell på tallene nevnt i media og for det aktuelle prosjektet, kan det ikke utelates at tallene i media har fungert som et anker for deltakerne. Tversky & Kahneman (1974) har vist at selv helt tilfeldige anker har hatt en påvirkning på deltakere i gjettekonkurranser. Dermed er det sannsynlig at studentene og supereksperter kan ha blitt påvirket av de tallene de har sett i media.

En interessant observasjon fra eksperimentet er at det er supereksperten som har gitt det høyeste estimatet, når det sammenlignes med flokkenes samlede estimat, Statsbyggs estimat og entreprenørens tilbud. Görtzen (2019) har vist at det finnes både forskning der det konkluderes med at eksperter blir påvirket av bias, og forskning der det konkluderes med at eksperter i liten grad påvirkes av bias. Görtzen (2019) har selv gjennom et eksperiment vist at ekspertise ikke alene er en garanti mot forankring. Et naturlig spørsmål som følge av ekspertens høye estimat vil være om ekspertens svar er mer påvirket av forankring enn studentenes svar. En ekspert er ofte mer bransjeorientert, og det er sannsynlig at han har lest mer om Norsk havteknologisenter enn studentene. Dette ble dog ikke undersøkt nærmere i dette eksperimentet. Problematikken kunne blitt undersøkt nærmere dersom deltakerne hadde svart på en undersøkelse om de hadde lest om prosjektet, og i hvor stor grad de hadde kjennskap til prosjektet på forhånd. Eventuelt kunne biasen blitt unngått ved å ikke fortelle deltakerne om at prosjektet er en del av Norsk havteknologisenter. Tiltaket ville imidlertid gjort at deltakerne hadde mistet informasjon om konteksten til prosjektet.

Fra datasettet ble det observert at flere av studentene ga estimat som både var mye høyere og mye lavere enn den supereksperten. Dette taler for at det er vanskelig å gi noe entydig svar om studentene eller supereksperten lar seg påvirke mest av forankring. Dessuten har ekspertene i Statsbygg gitt et estimat som ligger nærme entreprenørens samlede estimat. At flere av estimatene til individuelle studenter og supereksperten er et stykke unna entreprenørens gjennomsnittspris kan like så godt bare skyldes at estimatene i seg selv ikke er presise.

Likevel er det for de fleste studentene en tendens til at estimatene er høye. Som nevnt tidligere ville estimatene til studentene vært lavere dersom bare informasjonen om pris og størrelse hadde blitt benyttet. Dessuten var estimatene til studentene høyere i runde 1 enn i runde 2. Dette vitner om at studentene som ga høye estimat kan ha blitt påvirket av medieoppslagene.

Den samme argumentasjonen kan brukes for superekspertens høye estimat. Imidlertid er det en vesentlig forskjell på studentene brukt i dette eksperimentet og supereksperten. Flere av studentene har ikke holdepunkt for hva ting koster, mens eksperten har jobbet 34 år med kostnadsestimering. Dermed er det sannsynlig at studentene vil støtte seg på det lille de vet, i dette tilfellet prisene fra medieomtalen. Dette poenget kan videre underbygges med at flere av studentene nedjusterte sine estimat etter at informasjonen om pris på tidligere prosjekt ble synlig. Bidragene fra hver student ble ikke merket, så det har ikke vært mulig å se hvem som endret estimat. Likevel er det klart at flere av deltakeren nedjusterte sine estimat, ettersom datasettet i runde 2 inneholdt færre høye verdier enn datasettet i runde 1.

Supereksperten valgte på sin side å ikke endre sitt estimat etter at tidligere priser ble gjort kjent for han. Dette vitner videre om at eksperten har mer kontroll på hva ting koster. Samtidig vitner det om at eksperten innehar mer selvsikkerhet, noe som tidligere studier har sagt kjennetegner eksperter (Mannes et al., 2014; Minson & Mueller, 2012). Selvsikkerheten til supereksperten kan også brukes som et argument for at han ikke har latt seg påvirke av medienes oppslag. Ettersom han er klar over at de summene media presenterer ikke stemmer med det han fikk i oppgave å estimere, er det godt mulig at han i større grad enn studentene har vært i stand til å blokkere denne påvirkningen.

### *Tradisjonelt gruppearbeid og flokkmentalitet*

At flere av studentene endret estimat er interessant sett opp mot tradisjonelt gruppearbeid og flokkmentalitet. Resultatene fra Minson & Mueller (2012) viste at deltakere i tradisjonelt gruppearbeid er restriktive for å ta til seg ny informasjon. I dette eksperimentet endret imidlertid flere av studentene estimat etter at ny informasjon ble kjent. For eksperimentets del medførte endringen mer presise estimat. Eksperimentet skiller seg fra gruppearbeidet beskrevet i Minson & Mueller (2012) ved at studentene i masteroppgavens eksperiment ikke fikk se eller høre hva de andre studentene tenkte. Etter min mening kan det faktum at studentene valgte å ta til seg informasjonen forklares både med at estimatene ble gitt helt uavhengig av hverandre og fravær av flokkmentalitet.

Jayles et al. (2021), Da & Huang (2020) og Kao et al. (2018) skriver at flokkmentalitet fører til at folk ser bort fra personlig informasjon i sine estimat. Resultatene viste at for runde 2 var estimatene til studentene samlet rundt færre verdier enn runde 1. En umiddelbar tanke var at studentene hadde flokket seg til min opplysning om pris. Imidlertid hadde studentene flokket seg til en høyere verdi enn opplysningen om pris skulle tilsi. Dette viser at studentene fortsatt har vektlagt personlige meninger. Derfor mener jeg at samlingen av estimatene ikke skyldes flokkmentalitet.

En viktig årsak til at studentene fortsatt vektlegger personlig informasjon er at de har gitt estimat uavhengig av hverandre. I tradisjonelt gruppearbeid hender det støtt og stadig at personer slutter seg til det første forslaget. For enkelte kan det oppleves som ukomfortabelt å si imot en annen person som for eksempel innehar en høyere rang enn seg selv. Følgelig kan forslaget fra personen med høyest rang stå uimotsagt, og diskusjonen sentrerer seg rundt dette forslaget. Imidlertid har det fra resultatene og tidligere i diskusjonen kommet frem at ekspertise ikke er en garanti for suksess. Dersom forslaget fra en person med høy rang ikke er i nærheten av fasit, kan en diskusjon rundt det aktuelle forslaget føre til at fordelingen av svarene flyttes vekk fra riktig svar. Dermed kan diskusjoner i tradisjonelt gruppearbeid føre til dårligere samlede estimat enn for situasjoner for det benyttes wisdom of crowds. At en diskusjon fører til at estimatene flyttes vekk fra det riktige svaret er et eksempel på et systematisk bias. Å bruke wisdom of crowds der deltakerne har gitt uavhengige estimat, har vært med på å eliminere risikoen for at flokkmentalitet kunne finne sted.

I likhet med wisdom of crowds eliminerer også Delphimetoden flokkmentalitet. Deltakerne i en estimeringsoppgave hvor Delphimetoden brukes gir også estimat uavhengig av hverandre. Imidlertid skiller Delphimetoden seg fra wisdom of crowds på primært to områder.

- Delphimetoden benytter kun eksperter. Flokken i wisdom of crowds kjennetegnes av at den er mangfoldig og inkluderer lekfolk.
- Delphimetoden etterstreber å nå en samlet konsensus, og prosessen gjentas flere ganger. Wisdom of crowds kjennetegnes av at feil skal nulle hverandre ut, og at det ikke er et mål at alle deltakerne skal komme med like svar.

Til tross for at uavhengighet har bidratt til å eliminere flokkmentalitet i dette eksperimentet, så har Minson et al. (2018) pekt på at diskusjoner kan forhindre at medlemmer kommer med estimat som avviker mye fra det korrekte svaret. Studien viste at diskusjon forbedret de samlede forslagene til grupper for oppgaver hvor det ikke var begrensninger for hva svarene kunne være. Som nevnt tidligere er det ikke noen grense for hvor feil et kostnadsestimat for byggeprosjekter kan være. Imidlertid har Schultze et al. (2012) kommet frem til at selv om gruppemedlemmene individuelt sett forbedrer sine forslag med diskusjon, vil gruppens endelige forslag ikke være bedre enn gjennomsnittet av de individuelle estimatene. Derfor kan det se ut til at diskusjon ikke vil være til hjelp for wisdom of crowds. Det er heller uavhengigheten som bidrar til at fenomenet wisdom of crowds gir gode resultater.

### *Underestimering, optimisme og planleggingsfeilen*

En undergruppe av underestimering og optimisme er planleggingsfeilen (planning fallacy) (Flyvbjerg, 2021). Begrepet viser til de tilfellene hvor hendelser i for stor grad planlegges ut ifra at det best tenkelige scenarioet finner sted. I korte trekk vil det si at gjennomføringstid, arbeidsomfang, risiko og kostnader undervurderes, mens nytten og mulighetene for handlingene overvurderes.

Ettersom prosjektet ikke er fullført er det vanskelig å kunne slå fast om entreprenøren som vant anbudet har vært for optimistisk med sitt anslag. Følgelig er det usikkert om summen for totalentreprisen, som har blitt brukt som sammenligningsgrunnlag for studentenes, superekspertens og byggherrens estimat, er riktig. Om man likevel baserer seg på antakelsen om at summen er riktig, så er det ikke observert at hverken flokken av alle studentene eller supereksperten har drevet med underestimering og optimisme. Etter mitt syn kan dette forklares med at hverken studentene eller supereksperten har hatt noen direkte interesse av å gi lave estimat. Den eneste interessen de har hatt for dette prosjektet er å spå riktig pris. Estimaten er kun tuftet på et ønske om å gjøre det så godt som mulig. Etter mitt syn har dette vært en viktig bidragsyter for at studentenes estimat ikke har blitt for lave. Følgelig mener jeg at personer som ikke har interesse av utfallet til hva som skal estimeres, bør inkluderes i en estimeringsprosess. For bedrifter kan en mulig løsning på dette problemet være å bruke wisdom of crowds.

Tidligere i diskusjonen kom det frem at ved hjelp av digitale verktøy kan wisdom of crowds brukes med personer på tvers av deler i bedriften. En følge av dette er at personer i bedriften som ikke har en direkte interesse av utfallet til estimeringsprosessen, kan gi uavhengige estimat. For eksempel kan noen som sitter på et kontor i Oslo være med å estimere kostnaden for et prosjekt i Kristiansund. Personene i Oslo vil ikke ha den samme tilknytningen til prosjektet som personene i Kristiansund. Dersom kontoret i Kristiansund har hatt lav aktivitet over en lengre periode er det lett for at desperasjon sprer seg på kontoret. Ønsket om å vinne et anbud kan dermed prege anbudsregningen, gjennom en overoptimisme i tilbudet. Personene i Oslo vil være i bedre stand til å løsrive seg fra håpet om at alt kommer til å gå på skinner. I estimeringsprosessen vil personene fungere som en motvekt til optimismen. For å støtte denne tankerekken vises det til studentenes resultat i eksperimentet. Kanskje kan til og med personene i Oslo synes det er gøy å ta del i en utradisjonell estimeringsprosess med wisdom of crowd.

### *Ensidig søkelys på det unike og basisratefeil*

Etter mitt syn henger biasene ensidig søkelys på det unike og basisratefeil tett sammen for estimering av byggeprosjekt. Basisratefeil handler om at folk henger seg opp i detaljer og ikke ser de store linjene. Ofte er det disse detaljene som gjør prosjektene unikt i øynene for de som skal estimere prisen. Som nevnt tidligere fikk deltakerne i eksperimentet kun overordnet informasjon om prosjektet. En følge av dette er at deltakerne heller ikke har kunnet henge seg opp i detaljer og det unike med prosjektet. Dette har vært en av styrkene med å bare gi deltakerne overordnet informasjon.

## 7 Konklusjon

Masteroppgaven har hatt til hensikt å besvare problemstillingen «Hvordan kan wisdom of crowds benyttes for å redusere kostnadsrisiko i dagens byggebransje?». Problemstillingen besvares ved hjelp av å svare på forskningsspørsmålene. Konklusjonen for hvert av forskningsspørsmålene baserer seg på den foregående diskusjonen og analysen av resultatene til eksperimentet.

### 7.1 Hvordan kan digitale verktøy brukes for å hente inn data fra en flokk?

For å inkludere personer i flokken i wisdom of crowds konkluderes det med at det er mulig å holde digitale presentasjoner eller sende ut spørreskjema digitalt. Presentasjonen eller spørreskjemaet anbefales å kun være tilgjengelig for ansatte i virksomheten som benytter seg av flokkens estimat. Å gjennomføre innhenting av estimat digitalt gjør det mulig å få bidrag fra personer som ikke er fysisk til stede der hvor estimeringsprosessen ellers ville ha funnet sted. En følge av at ikke alle som deltar må være fysisk til stede på samme plass, er at flere vil ha anledning til å delta. Ettersom deltakerne ikke behøver å fysisk forflytte seg, vil ressursbruken forbundet med deres deltakelse være lavere enn om de hadde vært nødt til å forflytte seg. Dessuten innebærer en digital presentasjon, eller et spørreskjema, at det er forholdsvis enkelt å skalere opp og ned hvor mange deltakere som skal være med i flokken. Dette kan gjøres ved å sende ut koden til presentasjonen, eller spørreskjemaet, til et ønskelig antall personer.

### 7.2 Hvordan bør sammensetningen til flokken være?

Fra resultatene og diskusjonen har det kommet frem at usikkerheten til kostnadsestimatet til Flokken av alle studentene er lavere for et høyt antall personer i flokken, enn for et lavt antall personer. Følgelig vil det for wisdom of crowds være ønskelig å ha et stort antall personer i flokken når risikoen forbundet med det som skal estimeres er stor. Derfor bør det benyttes større flokker for store prosjekter, enn for små prosjekter.

Resultatene fra eksperiment har vist at gjennomsnittet og medianen til flere av flokkene studenter i eksperimentet har gitt presise estimat for kostnaden av en totalentreprise. Dermed fremstår det som at også uerfarne personer bør kunne benyttes for estimeringsoppgaver hvor wisdom of crowds brukes. Tradisjonelt sett har kun eksperter og ressurspersoner blitt brukt til estimeringsoppgaver i byggebransjen.

Uerfarne personer kjennetegnes av at de er mindre ressurskrevende å benytte, samt lettere å skaffe enn ressurspersoner. Fra diskusjonen trekkes det en konklusjon om at uerfarne personer bør brukes for å justere flokkstørrelsen i estimeringsoppgaver hvor wisdom of crowds brukes. Konklusjonen trekkes på bakgrunn av at uerfarne personer er ansett som relativt enkle å få tak i, samt at resultatet fra eksperimentet har vist at flokker med studenter er i stand til å gi presise estimat.

En mulig følge av å legge til uerfarne personer i flokken i wisdom of crowds, er at mangfoldet av estimat kan øke. At mangfoldet øker vil si at spredningen av estimat kan bli større. Dette gjør at sannsynligheten for at det korrekte svaret ligger innenfor spekteret av estimat øker.



### 7.3 Hvilken informasjon bør flokken ha tilgjengelig?

Det har ikke lyktes å komme til en klar konklusjon om eksakt hvilken informasjon deltakerne bør få i en estimeringsprosess tilsvarende den i forsøket. Flokken bestående av alle studentene har vist at det er mulig å gi presise estimat til tross for relativt lite informasjon. Imidlertid har den supereksperter ikke klart å gi et presist estimat basert på den samme informasjonen. Supereksperter har selv kommet med en liste over hvilken informasjon han mener er den viktigste for å estimere kostnadene for et byggeprosjekt. For studentene hadde prisen på tidligere byggeprosjekt en innvirkning på estimatene de ga. Imidlertid er det usikkert nøyaktig hvor viktig informasjonen om pris alene har vært, og om annen informasjon har spilt en større rolle.

Likevel kan resultatene fra eksperimentet holdes frem som et eksempel på at det er mulig å gi presise estimat, selv med lite informasjon. Lite informasjon er noe som ofte kjennetegner tidligfasen av prosjekter. For denne fasen er det vanlig at kostnadsestimatene avviker fra de senere kostnadsestimatene, for eksempel fra KS1 til KS2. Flokken av studenter har vært i stand til å gi presise kostnadsestimat, til tross for tilgang på relativt lite informasjon. Ettersom eksperimentet har vist at det er mulig å gi presise kostnadsestimat med liten mengde informasjon, bør bruken av wisdom of crowds vurderes for å gi kostnadsestimat tidlig i et byggeprosjekt.

### 7.4 Hvilke bias påvirker flokkens deltakere?

Totalentreprisen studentene og supereksperter skulle gi estimat for er en del av Norsk havteknologisenter. Isolert sett var totalentreprisen ukjent for deltakerne, mens Norsk havteknologisenter ikke var det. Personer i estimeringsprosesser er ofte påvirket av ulike bias, enten bevisst eller ubevisst. I eksperimentet er det sannsynlig at noen av studentene har blitt påvirket av et forankringsbias. Det er sannsynlig at studentene har blitt påvirket av medieomslag som har tatt for seg Norsk havteknologisenter. Konklusjonen dras på bakgrunn av at studentene har hatt en tendens til å gi høyere estimat enn hva informasjonen tilsier at de skulle gitt. Likevel er det vanskelig å konkludere med i hvor stor grad de ulike studentene har blitt påvirket. Under eksperimentet burde det blitt undersøkt hvem av studentene som hadde hørt om Norsk havteknologisenter på forhånd.

Det har ikke vært mulig å konkludere om supereksperter påvirkes mer eller mindre av forankringsbias enn studenter. Slutningen er begrunnet med at supereksperter har gitt et høyere estimat enn hva informasjonen om totalentreprisen skulle tilsi, men samtidig har han vært restriktiv med å ta til seg annen informasjon.

Det konkluderes med at hverken studentene eller supereksperter i eksperimentet har blitt utsatt for flokkmentalitet. Årsaken er at de ikke har kunnet kommunisere med andre. Dermed har de også gitt estimat uavhengig av andre. Dette skiller seg fra tradisjonelt gruppearbeid, hvor flokkmentalitet ofte oppstår. For studentenes samlede estimat har fraværet av flokkmentalitet vært en styrke. De individuelle estimatene har gitt uttrykk for hver enkelt students mening. Dette har muliggjort at flere av de for høye estimatene har blitt kansellert mot de for lave estimatene.

Innenfor byggebransjen holdes underestimering, overoptimisme og planleggingsfeilen frem som vanlige bias. For hverken flokken av alle studentene eller supereksperter er det observert underestimering og optimisme. Tvert imot har det vært en tendens til overestimering og pessimisme. En mulig forklaring for dette er at studentene og supereksperter ikke har hatt en direkte interesse av prosjektet. Estimaten til både

studentene og supereksperter er kun basert på et ønske om å gi presise estimat. Derfor vil det være klokt å involvere personer i prosessen som ikke har en direkte interesse av hva utfallet for hva som skal estimeres. Å benytte seg av digitale verktøy er ansett som en mulig løsning på dette problemet.

Det konkluderes med at tilgang til lite informasjon har bidratt til at de som ga estimat ikke har hengt seg opp i detaljer. Dette har muliggjort at studentene og supereksperter har evnet å bare se på bygget for hva det er.

## 8 Teoretiske implikasjoner

Ekspérimentets resultat går inn i rekken av observasjoner av fenomenet wisdom of crowds. Resultatene til eksperimentet sammenfaller med delene av teorien som har sagt at dersom en kombinerer et bredt spekter av estimater fra forskjellige individer, så vil feilene for hvert estimat avbøtes og det faktiske svaret vil sannsynligvis ligge et sted imellom (Larrick & Soll, 2006).

Frøslie (2023) sa at fordelingen til datasettet er hva som bør bestemme hvilket statistisk oppsummeringstall det er som skal brukes for å uttrykke det mest typiske. For eksperimentet stemte de teoretiske valgene av oppsummeringstall for nesten alle flokkene i begge rundene. Følgelig ble det konkludert med at det beregningsmetoden for å uttrykke det statistiske oppsummeringstallet for flokker i wisdom of crowds, bør gis basert på fordelingen av estimatene.

Fra resultatene har det kommet frem at usikkerheten til estimatene til Flokken av alle studentene avtar med størrelsen. At en økning i antall personer i flokken fører til en reduksjon av usikkerhet for det samlede estimatet er den samme tendensen Gecer Ulu et al. (2019) kom frem til.

Gecer Ulu et al. (2019) viste at ved å bruke wisdom of crowds så har selv en mikroflokk bestående av 15 personer gitt presise estimat for en ingeniøroppgave. For eksperimentet i denne masteroppgaver er det vist at selv flokker med 6 personer har vært i stand til å gi presise estimat for en ingeniøroppgave. Derfor samsvarer funnene om at selv små flokker kan gi presise estimat for estimeringsoppgaver innenfor ingeniørdisiplinen.

Selv om summene nevnt for Norsk havteknologisenter var irrelevant for totalentreprisen studentene skulle gi estimat for, er det mulig at studentene likevel har hengt seg opp i summene. Tversky & Kahneman (1974) viste at selv helt tilfeldige ankere kan påvirke deltakere i en gjettekonkurranse. At noen av studentene har latt seg påvirke av irrelevante summer peker på at teorien til Tversky & Kahneman (1974) har stemt for dette eksperimentet.

## 9 Praktiske implikasjoner

Fra masteroppgaven kan det trekkes ut at uavhengighet i estimeringsprosesser er viktig for å motvirke flokkmentalitet, underestimering, overoptimisme og planleggingsfeilen. Wisdom of crowds har vist seg som en fremsynsmetode hvor disse biasene i stor grad unngås. For å skaffe en flokk av personer som er uavhengig av utfallet til hva som skal estimeres, er det sett på om det er mulig å bruke digitale verktøy.

I eksperimentet ble en Mentimeter-presentasjon benyttet. En Mentimeter-presentasjon er en digital presentasjon hvor hver av deltakerne får opp presentasjonen på en personlig enhet. Alt deltakerne trenger for å delta er internett og en kode tildelt av administratoren for presentasjon. Dette har gjort at presentasjonen er forholdsvis lett å distribuere til personer man selv ønsker at skal delta. Å bruke en slik lukket presentasjon stiller seg ulikt fra crowdsourcing-plattformer, hvor det ikke er et filter for å kunne delta. Ved å holde en Mentimeter-presentasjon er det mulig å involvere så mange personer man selv ønsker at skal delta.

Resultatene fra eksperimentet har vist at for situasjoner hvor wisdom of crowds benyttes, reduseres usikkerheten ved estimatene som gis av en flokk, for enhver økning av antall personer i flokken. I eksperimentet er det vist at flokker bestående av uerfarne personer har gitt presise estimat. Derfor bør det se på om også uerfarne personer bør benyttes for estimeringsoppgaver, hvor wisdom of crowds benyttes.

Eksperimentet har vist at studenter, ved hjelp av fremsynsmetoden wisdom of crowds, har gitt et presist estimat for prisen til en totalentreprise for et kontor- og undervisningsbygg. Estimaten studentene ga var basert på lite informasjon. Derfor bør bruken av wisdom of crowds vurderes for å gi kostnadsestimat for prosjekter, eller konsepter, hvor det foreligger lite informasjon. Som regel vil det si tidlig i prosjektet. Som et eksempel på foreslått fremtidig bruk av wisdom of crowds, trekkes KS1 vurderinger frem.

## 10 Et kritisk blikk på masteroppgaven

Resultatene fra masteroppgaven kunne blitt gjort mer generalisert dersom det hadde blitt utført flere eksperimenter, og for flere typer prosjekt. Til tross for at funnet melder seg inn i rekken av tilfeller hvor en effekt av wisdom of crowds kan observeres, er det vanskelig å si om det bare er flaks, eller om dette vil gjenta seg for alle estimeringsoppgaver i byggebransjen.

Prosjektet skulle som diskutert tidligere ikke vurderes 100 % pris, også kvalifikasjon på tilbudt personell skulle vurderes. For dette eksperimentet har ikke dette vært å anse som et særlig stort problem, eller noe som har påvirket estimatene som ble gitt. Derimot er det en trend for at flere prosjekt som lyses ut på anbud har andre krav enn kun pris, for eksempel miljøkrav og krav til samspill. Dette er faktorer som i stor grad kan påvirke estimatene som gis, og kan knyttes til virksomhetens egen erfaring. Derfor kan det være vanskelig å si noe om hvorvidt wisdom of crowds kan brukes til å gi estimat for prosjekt med disse kravene. Derimot er det noen som må ende opp med å betale for prosjektet, så det vil alltid være interessant å ha en indikasjon på hva det kommer til å koste.

For å gi en mer konkret og spisset diskusjon rundt tradisjonelle samarbeidsformer, sett opp mot wisdom of crowds, burde Statsbyggs samarbeidsform blitt studert mer inngående. Derimot ville dette vært vanskelig for denne masteroppgaven. Statsbyggs estimat ble utarbeidet før arbeidet med masteroppgaven startet. Likevel kunne jeg spurt om en detaljert beskrivelse av hvordan Statsbygg endte opp med sitt estimat. Derimot ville dette åpnet opp for en feilkilde i form av at beskrivelsen til Statsbygg ikke ville samsvart med hva som faktisk ble gjort. Det hadde vært få måter å bekrefte hva på som ble sagt, og det er også mulig at beskrivelsen ville hatt en begrepsbruk som ikke hadde vært presis.

## 11 Videre arbeid

Masteroppgaven og eksperimentet har stort sett på bruken av wisdom of crowds som fremsynsmetode for kostnader. Resultatene fra eksperimentet har vist at det er mulig for uerfarne å gi presise estimat for en ingeniøroppgave. Derfor vil det være interessant å se videre på hvilke andre ingeniøroppgaver flokker av uerfarne kan løse. For byggebransjen er det etter mitt syn interessant å se på om en flokk er i stand til å gi et presist estimat for byggetiden til et prosjekt.

Den eksterne eksperter kom med en rekke punkter over hva han anser å være de viktigste punktene for å estimere kostnadene for et byggeprosjekt. Det vil være interessant å gjenta forsøket med denne informasjonen, for å se hvordan informasjonen kan påvirke estimatene til både eksperter og uerfarne.

Fra forsøket er det observert at usikkerheten til estimatene til Flokken av alle studentene avtar med størrelsen. Det er dog ikke konkludert med hva som er den beste flokkstørrelsen, ettersom dette blant annet vil være avhengig av ressursbruk og hva som skal estimeres. For fremtidig forskning vil det å se på sammenheng mellom flokkstørrelse og kompetanse for ulike arbeidsoppgaver være interessant.

I masteroppgaven er estimatene til studentene, supereksperter og byggherren sammenlignet med entreprenørens priser. Likevel er det ikke gitt at prisen for totalentreprisen kommer til å bli dette. Følgelig vil det være interessant å se på hva sluttsammen til totalentreprisen faktisk ender på.

For dette eksperimentet har flokkene vært en ren sammensetning av studenter. For et fremtidig eksperiment vil det være interessant på sammensetningen hvor flokkens deltakere består av både eksperter og uerfarne.

Det burde blitt undersøkt om de mindre flokkene som ga gode estimat faktisk var gode til å estimere, eller om de bare hadde flaks. Til dette formålet kunne de to mindre flokkene som ga gode estimat, fått estimere flere prosjekt. Avhengig av resultatene for estimeringen av disse prosjektene kunne det vært lettere å bedømme om disse mindre flokkene, samlet sett, inneholdt en unik evne til å gi presise estimat.

# Referanser

- Austeng, K., Torp, O., Midtbø, J. T., Helland, V., & Jordanger, I. (2005). *Usikkerhetsanalyse—Metoder* (Concept Rapport Nr. 12). Ex ante akademiske forlag. <https://ntnuopen.ntnu.no/ntnu-xmlui/handle/11250/228073>
- Belle, N., Cantarelli, P., & Belardinelli, P. (2017). Cognitive Biases in Performance Appraisal: Experimental Evidence on Anchoring and Halo Effects With Public Sector Managers and Employees. *Review of Public Personnel Administration*, 37(3), 275–294. <https://doi.org/10.1177/0734371X17704891>
- Blesik, T., Bick, M., & Kummer, T.-F. (2022). A Conceptualisation of Crowd Knowledge. *Information Systems Frontiers*, 24(5), 1647–1665. <https://doi.org/10.1007/s10796-021-10176-y>
- Budescu, D. V., & Chen, E. (2015). Identifying Expertise to Extract the Wisdom of Crowds. *Management Science*, 61(2), 267–280. <https://doi.org/10.1287/mnsc.2014.1909>
- Da, Z., & Huang, X. (2020). Harnessing the Wisdom of Crowds. *Management Science*, 66(5), 1847–1867. <https://doi.org/10.1287/mnsc.2019.3294>
- Dalkey, N., & Helmer, O. (1963). An Experimental Application of the DELPHI Method to the Use of Experts. *Management Science*, 9(3), 458–467. <https://doi.org/10.1287/mnsc.9.3.458>
- Dalland, O. (1993). *Metode og oppgaveskriving for studenter* (1. utgave). Universitetsforl.
- Dalland, O. (2001). *Metode og oppgaveskriving for studenter* (3. utgave, 2. opplag). Gyldendal.
- Drevland, F. (2013). *Kostnadsestimering under usikkerhet* (Temahefte Nr. 4; Concept temahefte, s. 28). NTNU. [https://www.ntnu.no/documents/1261860271/1262010610/CONCEPT\\_kostnadsestimering\\_til+WEB.pdf/7fe95f32-0477-4468-b0e5-54589687c16d](https://www.ntnu.no/documents/1261860271/1262010610/CONCEPT_kostnadsestimering_til+WEB.pdf/7fe95f32-0477-4468-b0e5-54589687c16d)
- Dvergsdal, H. (2023). Crowdsourcing. I *Store norske leksikon*. <https://snl.no/crowdsourcing>
- Estellés-Arolas, E., & González-Ladrón-de-Guevara, F. (2012). Towards an integrated crowdsourcing definition. *Journal of Information Science*, 38(2), 189–200. <https://doi.org/10.1177/0165551512437638>
- Fiechter, J. L., & Kornell, N. (2021). How the wisdom of crowds, and of the crowd within, are affected by expertise. *Cognitive Research: Principles and Implications*, 6(1), 5. <https://doi.org/10.1186/s41235-021-00273-6>
- Finansdepartementet. (2019, oktober 30). *Hva er statens prosjektmodell?* [Redaksjonellartikkel]. Regjeringen.no; regjeringen.no. <https://www.regjeringen.no/no/tema/okonomi-og-budsjett/statlig-okonomistyring/ekstern-kvalitetssikring2/hva-er-ks-ordningen/id2523897/>
- Flyvbjerg, B. (2006). From Nobel Prize to Project Management: Getting Risks Right. *Project Management Journal*, 37(3), 5–15. <https://doi.org/10.1177/875697280603700302>
- Flyvbjerg, B. (2011). *Over Budget, Over Time, Over and Over Again*. Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780199563142.003.0014>
- Flyvbjerg, B. (2021). Top Ten Behavioral Biases in Project Management: An Overview. *Project Management Journal*, 52(6), 531–546. <https://doi.org/10.1177/87569728211049046>
- Flyvbjerg, B., Holm, M. S., & Buhl, S. (2002). Underestimating Costs in Public Works Projects: *Error or Lie?* *Journal of the American Planning Association*, 68(3), 279–295. <https://doi.org/10.1080/01944360208976273>
- Forskningsprogrammet Concept—NTNU*. (u.å.). Hentet 8. mai 2023, fra <https://www.ntnu.no/concept>

- Frøslie, K. F. (2022). Store talls lov. I *Store norske leksikon*. [https://snl.no/store\\_talls\\_lov](https://snl.no/store_talls_lov)
- Frøslie, K. F. (2023). Median. I *Store norske leksikon*. <https://snl.no/median>
- Galesic, M., Barkoczi, D., & Katsikopoulos, K. (2018). Smaller crowds outperform larger crowds and individuals in realistic task conditions. *Decision*, 5(1), 1–15. <https://doi.org/10.1037/dec0000059>
- Galton, F. (1907). Vox Populi. *Nature*, 75(1949), Artikkel 1949. <https://doi.org/10.1038/075450a0>
- Gecer Ulu, N., Messersmith, M., Goucher-Lambert, K., Cagan, J., & Kara, L. B. (2019). Wisdom of Microcrowds in Evaluating Solutions to Esoteric Engineering Problems. *Journal of Mechanical Design*, 141(8), 081102. <https://doi.org/10.1115/1.4042615>
- Goldstein, D. G., McAfee, R. P., & Suri, S. (2014). The wisdom of smaller, smarter crowds. *Proceedings of the Fifteenth ACM Conference on Economics and Computation*, 471–488. <https://doi.org/10.1145/2600057.2602886>
- Graf-Drasch, V., Gimpel, H., Barlow, J. B., & Dennis, A. R. (2022). Task structure as a boundary condition for collective intelligence. *Personnel Psychology*, 75(3), 739–761. <https://doi.org/10.1111/peps.12489>
- Gürçay, B., Mellers, B. A., & Baron, J. (2015). The Power of Social Influence on Estimation Accuracy: The Power of Social Influence. *Journal of Behavioral Decision Making*, 28(3), 250–261. <https://doi.org/10.1002/bdm.1843>
- Görzen, T. (2019). Can Experience be Trusted? Investigating the Effect of Experience on Decision Biases in Crowdfunding Platforms. *Proceedings of the 52nd Hawaii International Conference on System Sciences*.
- Hellevik, O. (1999). *Forskningsmetode i sosiologi og statsvitenskap*. Universitetsforl. [https://www.nb.no/items/URN:NBN:no-nb\\_digibok\\_2008082604036?page=13](https://www.nb.no/items/URN:NBN:no-nb_digibok_2008082604036?page=13)
- Hill, S., & Ready-Campbell, N. (2011). Expert Stock Picker: The Wisdom of (Experts in) Crowds. *International Journal of Electronic Commerce*, 15(3), 73–102. <https://doi.org/10.2753/JEC1086-4415150304>
- Hong, H., Ye, Q., Du, Q., Wang, G. A., & Fan, W. (2020). Crowd characteristics and crowd wisdom: Evidence from an online investment community. *Journal of the Association for Information Science and Technology*, 71(4), 423–435. <https://doi.org/10.1002/asi.24255>
- Jacobsen, D. I. (2022). *Hvordan gjennomføre undersøkelser?: Innføring i samfunnsvitenskapelig metode* (4. utgave). Cappelen Damm Akademisk.
- Jayles, B., Escobedo, R., Cezero, S., Blanchet, A., Kameda, T., Sire, C., & Theraulaz, G. (2020). The impact of incorrect social information on collective wisdom in human groups. *Journal of The Royal Society Interface*, 17(170), 20200496. <https://doi.org/10.1098/rsif.2020.0496>
- Jayles, B., Sire, C., & Kurvers, R. H. J. M. (2021). Crowd control: Reducing individual estimation bias by sharing biased social information. *PLOS Computational Biology*, 17(11), e1009590. <https://doi.org/10.1371/journal.pcbi.1009590>
- Jelle, B. P. (2022, september 14). *Uncertainty Calculations* [TBM4500 Bygg- og miljøteknikk, fordypningsprosjekt (2022 HØST)]. [ntnu.blackboard.com](https://ntnu.blackboard.com)
- Kao, A. B., Berdahl, A. M., Hartnett, A. T., Lutz, M. J., Bak-Coleman, J. B., Ioannou, C. C., Giam, X., & Couzin, I. D. (2018). Counteracting estimation bias and social influence to improve the wisdom of crowds. *Journal of The Royal Society Interface*, 15(141), 20180130. <https://doi.org/10.1098/rsif.2018.0130>
- Keuschnigg, M., & Ganser, C. (2017). Crowd Wisdom Relies on Agents' Ability in Small Groups with a Voting Aggregation Rule. *Management Science*, 63(3), 818–828. <https://doi.org/10.1287/mnsc.2015.2364>
- Larrick, R. P. (2016). The Social Context of Decisions. *Annual Review of Organizational Psychology and Organizational Behavior*, 3(1), 441–467. <https://doi.org/10.1146/annurev-orgpsych-041015-062445>



- Larrick, R. P., & Soll, J. B. (2006). Intuitions About Combining Opinions: Misappreciation of the Averaging Principle. *Management Science*, 52(1), 111–127. <https://doi.org/10.1287/mnsc.1050.0459>
- Lorenz, J., Rauhut, H., Schweitzer, F., & Helbing, D. (2011). How social influence can undermine the wisdom of crowd effect. *Proceedings of the National Academy of Sciences*, 108(22), 9020–9025. <https://doi.org/10.1073/pnas.1008636108>
- Love, P. E. D., & Ahiaga-Dagbui, D. D. (2018). Debunking fake news in a post-truth era: The plausible untruths of cost underestimation in transport infrastructure projects. *Transportation Research Part A: Policy and Practice*, 113, 357–368. <https://doi.org/10.1016/j.tra.2018.04.019>
- Mannes, A. E., Soll, J. B., & Larrick, R. P. (2014). The wisdom of select crowds. *Journal of Personality and Social Psychology*, 107(2), 276–299. <https://doi.org/10.1037/a0036677>
- Minson, J. A., & Mueller, J. S. (2012). The Cost of Collaboration: Why Joint Decision Making Exacerbates Rejection of Outside Information. *Psychological Science*, 23(3), 219–224. <https://doi.org/10.1177/0956797611429132>
- Minson, J. A., Mueller, J. S., & Larrick, R. P. (2018). The Contingent Wisdom of Dyads: When Discussion Enhances vs. Undermines the Accuracy of Collaborative Judgments. *Management Science*, 64(9), 4177–4192. <https://doi.org/10.1287/mnsc.2017.2823>
- Nguyen, V. D., & Nguyen, N. T. (2017). The impact of diversity on the quality of collective prediction. *2017 IEEE International Conference on INnovations in Intelligent SysTems and Applications (INISTA)*, 149–154. <https://doi.org/10.1109/INISTA.2017.8001148>
- Nilstun, C. (2018). Eksoterisk. I *Store norske leksikon*. <https://snl.no/eksoterisk>
- Nilstun, C. (2023). Ressurs. I *Store norske leksikon*. <https://snl.no/ressurs>
- NVivo—Kunnskapsbasen—NTNU. (u.å.). Hentet 28. mars 2023, fra <https://i.ntnu.no/wiki/-/wiki/Norsk/nvivo>
- Offergaard, S. (2022, august 26). (+) *AF-sjefen: – Risikostyring viktigere enn noen gang*. <https://www.bygg.no/article/1505350/>
- Offergaard, S. (2023, mai 10). (+) *Fra ryddegutt til konsernsjef i Skanska*. <https://www.bygg.no/article/1526093/>
- Olseng, E. T., & Sundbye, L. M. T. (2020, desember 7). *Kommunikasjonsprosessen—Markedsføring og ledelse I—NDLA*. ndla.no. <https://ndla.no/nb/subject:1:47678c7b-bc09-4fc8-b2d9-a2e3d709e105/topic:1:aa643598-5490-42c9-8ce5-62a94b42bfe0/resource:556cac24-751c-4595-8b20-efaf2763d79c>
- Rowe, G., & Wright, G. (1999). The Delphi technique as a forecasting tool: Issues and analysis. *International Journal of Forecasting*, 15(4), 353–375. [https://doi.org/10.1016/S0169-2070\(99\)00018-7](https://doi.org/10.1016/S0169-2070(99)00018-7)
- Samset, K. (2017). *Prosjekt i tidligfasen—Valg av konsept* (2. utgave).
- Samset, K. (2021). *Utredninger og rapporter* (Concept Temahefte Nr. 15). Ex ante akademiske forlag. <https://www.ntnu.no/concept/concept-temahefter>
- Sander, K. (2021, november 22). *Valg av kommunikasjonskanaler—Innovasjon og markedsføring (SR-SRL vg2)—NDLA*. ndla.no. <https://ndla.no/nb/subject:1:a7c337ca-d3b6-492f-ace2-b05c45f54e93/topic:1:16b7d424-4674-4797-afcb-c68806c5d892/topic:1:c57feb7f-310a-43ee-89ff-bc79dd90ce6c/resource:1:162615>
- Schultze, T., Mojzisch, A., & Schulz-Hardt, S. (2012). Why groups perform better than individuals at quantitative judgment tasks: Group-to-individual transfer as an alternative to differential weighting. *Organizational Behavior and Human Decision Processes*, 118(1), 24–36. <https://doi.org/10.1016/j.obhdp.2011.12.006>
- Simmons, J. P., Nelson, L. D., Galak, J., & Frederick, S. (2011). Intuitive Biases in Choice versus Estimation:

- Implications for the Wisdom of Crowds. *Journal of Consumer Research*, 38(1), 1–15.  
<https://doi.org/10.1086/658070>
- Smeby, J.-C. (2021). Ekspert. I *Store norske leksikon*. <https://snl.no/ekspert>
- Stanovich, K. E., & West, R. F. (2008). On the relative independence of thinking biases and cognitive ability. *Journal of Personality and Social Psychology*, 94(4), 672–695. <https://doi.org/10.1037/0022-3514.94.4.672>
- Sunnevåg, K. J. (2007). *Beslutninger på svakt informasjonsgrunnlag* (Concept Rapport Nr. 17; Concept Rapport). NTNU. <https://www.ntnu.no/concept/concept-rapportserie>
- Surowiecki, J. (2005). *The wisdom of crowds* (Nachdr.). Anchor Books.
- Torp, O. (2020, august 25). *BYGGEPROSJEKTER (1) Byggverk, byggeprosesser og prosjektorganisasjoner* [TBA4135 Byggeprosess Grunnkurs (2020 HØST)]. [ntnu.blackboard.com](https://ntnu.blackboard.com)
- Torp, O. (2021, september 21). *Cost estimation under uncertainty* [TBA4155 Byggeprosess: Tidligfasen i prosjekter (2021 HØST)]. [ntnu.blackboard.com](https://ntnu.blackboard.com)
- Torp, O., Drevland, F., & Austeng, K. (2015). *Prosess for kostnadsestimering under usikkerhet* (Concept Temahefte Nr. 6). Ex ante akademiske forlag.
- Tversky, A., & Kahneman, D. (1974). Judgment under Uncertainty: Heuristics and Biases: Biases in judgments reveal some heuristics of thinking under uncertainty. *Science*, 185(4157), 1124–1131.  
<https://doi.org/10.1126/science.185.4157.1124>
- Ugander, J., Drapeau, R., & Guestrin, C. (2015). The Wisdom of Multiple Guesses. *Proceedings of the Sixteenth ACM Conference on Economics and Computation*, 643–660. <https://doi.org/10.1145/2764468.2764529>
- Vatne, J. E. (2023). Gjennomsnitt. I *Store norske leksikon*. <https://snl.no/gjennomsnitt>
- Wallis, K. F. (2014). Revisiting Francis Galton’s Forecasting Competition. *Statistical Science*, 29(3).  
<https://doi.org/10.1214/14-STS468>
- Welde, M. (2016). *Kostnadsutvikling i store statlige investerings- prosjekter fra KS1 til KS2* [Concept Arbeidsrapport]. <https://www.ntnu.no/concept/arbeidsrapporter>
- Welde, M., Dahl, R. E., Torp, O., & Aass, T. (2018). *Kostnadsstyring i entreprisekontrakter* (Concept Rapport Nr. 55). Ex ante akademiske forlag. <https://www.ntnu.no/concept/concept-rapportserie>
- Welde, M., Jørgensen, M., Larssen, P. F., & Halkjelsvik, T. (2019). *Estimering av kostnader i store statlige prosjekter: Hvor gode er estimatene og usikkerhetsanalysene i KS2-rapportene?* (Concept Rapport Nr. 59). Concept-programmet. <https://www.ntnu.no/concept/concept-rapportserie>
- Wohlin, C. (2014). Guidelines for snowballing in systematic literature studies and a replication in software engineering. *Proceedings of the 18th International Conference on Evaluation and Assessment in Software Engineering*, 1–10. <https://doi.org/10.1145/2601248.2601268>
- Yi, S. K. M., Steyvers, M., Lee, M. D., & Dry, M. J. (2012). The Wisdom of the Crowd in Combinatorial Problems. *Cognitive Science*, 36(3), 452–470. <https://doi.org/10.1111/j.1551-6709.2011.01223.x>
- Yu, C., Chai, Y., & Liu, Y. (2017). Collective Intelligence: From the Enlightenment to the Crowd Science. *Proceedings of the 2nd International Conference on Crowd Science and Engineering - ICCSE'17*, 111–115.  
<https://doi.org/10.1145/3126973.3126993>
- Østbye, H., & Knapskog, K. (1994). *Metodebok for mediestudenter: En innføring i kvantitativ metode* (3. foreløpige utg., (rev.)).

# Bilag

**Bilag 1:** Litteratursøk

**Bilag 2:** Presentasjon gitt til studentene og den enkelte eksperten

**Bilag 3:** Python-skript til beregningene og plottene

**Bilag 4:** Datasettet med estimat gitt av studentene og den enkelte eksperten

## Bilag 1: Litteratursøk

Tabell 26: Viser en oversikt over antall treff for ulike databaser og søkeord fra litteraturstudien. En beskrivelse av søkeordene er også gitt.

Database	Søkeord	Beskrivelse	Treff
Scopus	TITLE-ABS-KEY ( wisdom AND of AND the AND crowd AND engineering )	Søket inneholder kjernen i problemstillingen, wisdom of crowds. Søket er forsøkt avgrenset til byggebransjen for å sikre relevans til oppgavens problemstilling.	99
Scopus	TITLE-ABS-KEY ( wisdom AND of AND the AND crowd AND information AND available ) AND ( LIMIT-TO ( SUBJAREA , "BUSI" ) OR LIMIT-TO ( SUBJAREA , "DECI" ) OR LIMIT-TO ( SUBJAREA , "SOCI" ) OR LIMIT-TO ( SUBJAREA , "ENGI" ) OR LIMIT-TO ( SUBJAREA , "ECON" ) OR LIMIT-TO ( SUBJAREA , "ARTS" ) )	Søket er relevant for forsøkningsspørsmål nummer 3. Inneholder avgrensninger relatert til fagfelt.	36
Scopus	( TITLE-ABS-KEY ( the AND wisdom AND of AND the AND crowd ) ) AND ( ( construction ) ) AND ( engineering ) AND ( LIMIT-TO ( SUBJAREA , "ENGI" ) OR LIMIT-TO ( SUBJAREA , "SOCI" ) OR LIMIT-TO ( SUBJAREA , "BUSI" ) OR LIMIT-TO ( SUBJAREA , "DECI" ) OR LIMIT-TO ( SUBJAREA , "ECON" ) )	Ettersom problemstillingen ser på hvordan wisdom of crowds kan brukes i dagens byggebransje er det nyttig å skaffe seg kjennskap til hva tidligere litteratur har sett på.	29
Scopus	( TITLE-ABS-KEY ( "anchoring effect" ) ) AND ( ( ( bias ) ) AND ( judgement ) ) AND ( wisdom ) AND ( LIMIT-TO ( SUBJAREA , "ENGI" ) OR LIMIT-TO ( SUBJAREA , "SOCI" ) OR LIMIT-TO ( SUBJAREA , "BUSI" ) OR LIMIT-TO ( SUBJAREA , "ECON" ) )	Relateres til anchoring bias.	8
Scopus	( TITLE-ABS-KEY ( collective AND intelligence ) ) AND ( ( ( construction ) ) AND ( engineering ) ) AND ( estimates )	Ser på det samme som det første søket, men der hvor begrepet collective intelligence er benyttet.	27

Scopus	TITLE-ABS-KEY ( wisdom AND of AND the AND crowd AND estimation ) AND ( LIMIT-TO ( SUBJAREA , "ENGI" ) OR LIMIT-TO ( SUBJAREA , "PSYC" ) OR LIMIT-TO ( SUBJAREA , "BUSI" ) OR LIMIT-TO ( SUBJAREA , "DECI" ) )	Ser på bruken knyttet til å gi estimat.	48
Scopus	TITLE-ABS-KEY ( wisdom AND of AND crowds AND herding )	Ser på flokkmentalitet	23
Oria	wisdom of the crowd information available  Filter: Fra fagfelleverderte tidsskrift Artikler Konferanseforedrag Bokkapitler Avhandlinger Engelsk Science & Technology Social Sciences Decision Making Wisdom Of Crowds Business & Economics Psychology Management Fra fagfelleverderte tidsskrift Artikler Konferanseforedrag Bokkapitler Avhandlinger Engelsk Science & Technology Social Sciences Decision Making Wisdom Of Crowds Business & Economics Psychology Management	Relatert til forskningsspørsmål 3. Relativt likt som det tidligere søket i Scopus	175
Oria	wisdom of the crowd construction engineering  Filter: Fra fagfelleverderte tidsskrift Artikler Konferanseforedrag Bokkapitler Avhandlinger Engelsk Science & Technology Social Sciences	Søket er relatert til hvordan wisdom of crowds brukes innenfor byggebransjen og ingeniørdisiplinene . Relativt likt som det tidligere søket i Scopus.	29

	Decision Making Wisdom Of Crowds Business & Economics Psychology Management Fra fagfelleverderte tidsskrift Artikler Konferanseforedrag Bokkapitler Avhandlinger Engelsk Science & Technology Social Sciences Decision Making Wisdom Of Crowds Business & Economics Psychology Management		
Oria	Anchoring effect	Ingen avgrensninger ble brukt fordi det uansett hadde gitt svært mange treff.	1702 3
Oria	collective intelligence construction engineering estimates Filter: Engelsk	Bare avgrensningen til engelsk litteratur ble brukt ettersom søket ga få resultat. Det fantes ikke norsk litteratur.	16
	wisdom of the crowd estimation Filter: Fra fagfelleverderte tidsskrift Artikler Konferanseforedrag Science & Technology Social Sciences Technology Psychology Estimation Wisdom Of Crowds Decision Making Accuracy Estimates Behavioral Economics Behavioral Finance	Resultatene fra søket ser på bruken av wisdom of crowds til å gi estimat.	84
Oria	Visdom til flokk	Direkte oversettelse av	4

		wisdom of crowds til norsk	
Oria	Mengdens visdom	Alternativ oversettelse av wisdom of crowds	25

## Bilag 2: Presentasjon gitt til flokken studenter

Mentimeter



**Er dere flinkere enn ekspertene i byggebransjen?**

Kan en klasse på NTNU komme med bedre kostnadsestimater for et byggeprosjekt enn hva erfarne entreprenører er i stand til?

Ved hjelp av fenomenet the wisdom of the crowd kan dette være mulig, men det er avhengig av deres deltakelse og estimat.

Fenomenet dreier seg om at det samlede estimatet til en gruppe ofte har høyere grad av presisjon enn enkeltstående eksperter.

Eksperimentet går ut på at dere i løpet av kort tid skal anslå prisen for et prosjekt som for øyeblikket er ute på anbud. Etter dere har sendt inn deres første gjet vil dere få opplyst hva tilsvarende bygninger har kostet tidligere, og få en ny sjansse til å estimere prisen for det samme prosjektet.

Hele sekvensen vil ta under 15 min, og deltakelsen krever ingen forkunnskaper. Etter at byggherren har mottatt tilbud fra entreprenørene vil klassens priser sammenlignes med entreprenørene.



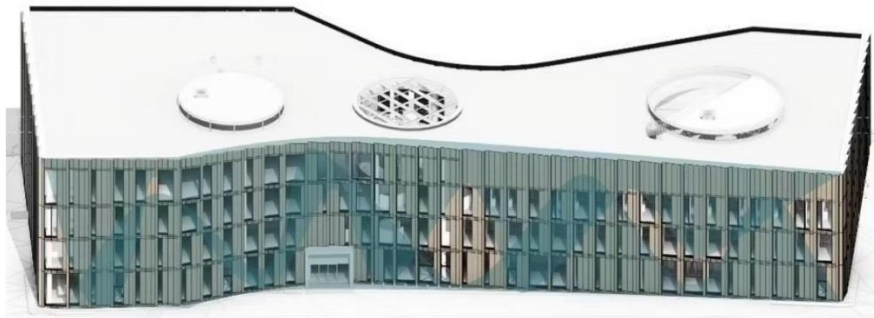
Mentimeter



Skisse

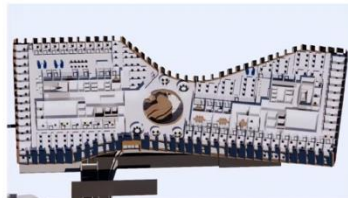








Aksjonmøtt Nybygg plan 1



Aksjonmøtt Nybygg plan 2



## Nøkkelinformasjon

- Blanding av kontor og undervisningsbygg
- Bruttoareal mellom 9 500 m<sup>2</sup> og 10 100 m<sup>2</sup>
- Riving av eksisterende bebyggelse er en del av kontrakten



Hva koster dette ekskl. mva.?



## Nøkkelinformasjon

- Blanding av kontor og undervisningsbygg
- Bruttoareal mellom 9 500 m<sup>2</sup> og 10 100 m<sup>2</sup>
- Riving av eksisterende bebyggelse er en del av kontrakten

Bygningstype	Pris/BTA m <sup>2</sup> (ekskl. mva.)
Kontorbygning 15 000 m <sup>2</sup>	32 365
Kontorbygning 5 000 m <sup>2</sup>	33 415
Videregående skole, studiespesialisering	36 487

Tallene er erfaringspriser hentet fra Norsk prisbok 2022



Hva koster dette ekskl. mva. (2. forsøk)?



### **Bilag 3: Python-skript til beregningene og plottene**

## Plots

```
import numpy as np
import matplotlib.pyplot as plt
import itertools
```

```
#Entreprenørens samlede pris
P = (392437852+409098509+382740705)/3
```

```
#Individuell ekspert
S = 686850000
```

```
#Statsbygg
SB = 400000000
```

```
#Entreprenør 3
W = 382740705
```

```
# Flokk 1
p1_try1 = [
    500,
    500,
    185.9,
    400,
    230,
    12,
    280,
    600,
    1000]
```

```
p1_try2 = [
    400,
    380,
    374,
    330,
    330.5,
    334,
    450,
    600,
    490,
]
```

```
# Flokk 2
p2_try1 = [
    250,
    380,
    48,
    750.15,
    800,
    145,
]
```

```
p2_try2 = [
    345,
    450,
    347.5,
    390,
    412,
    350,
]
```

```

# Flokk 3
p3_try1 = [
    530,
    1300,
    326,
    600,
    800,
    350,
    325,
    500,
    300,
    1050,
    615,
    450,
    250,
    800,
]

p3_try2 = [
    450,
    500,
    350,
    459.420069,
    690,
    390,
    450,
    355,
    435,
    400,
    550,
    420,
    341,
    950,
]

# Flokk 4
p4_try1 = [
    300,
    440,
    650,
    1200,
]

p4_try2 = [
    330,
    850,
    335,
    650,
]

p_try1 = np.array(p1_try1+p2_try1+p3_try1+p4_try1)
p_try2 = np.array(p1_try2+p2_try2+p3_try2+p4_try2)

np.random.shuffle(p_try1)
np.random.shuffle(p_try2)

fig1, (ax0, ax1) = plt.subplots(nrows=1,ncols=2, figsize = (15,6))
ax0.hist(np.sort(p_try1), label = "",rwidth=0.92, bins=15)
ax1.hist(np.sort(p_try2), label = "",rwidth=0.92,bins=8)
# Sett entreprenørenes samlede pris
ax0.vlines(P*1e-6, 0, 18, linestyle = "dashed", label =
    "Gjennomsnittspris \n til entreprenørene", color = "g")
ax1.vlines(P*1e-6, 0, 18, linestyle = "dashed", label =
    "Gjennomsnittspris \n til entreprenørene", color = "g")
# Sett Entreprenør 3
ax0.vlines(W*1e-6, 0, 18, linestyle = "dashed", label =
    "Pris til Entreprenør 3", color = "m")
ax1.vlines(W*1e-6, 0, 18, linestyle = "dashed", label =
    "Pris til Entreprenør 3", color = "m")
# Sett aritmetisk gjennomsnitt
ax0.vlines(np.mean(p_try1), 0, 18, linestyle = "dashed", label =
    "Gjennomsnitt av estimat \n til Flokk av alle studentene ", color = "r")
ax1.vlines(np.mean(p_try2), 0, 18, linestyle = "dashed", label =
    "Gjennomsnitt av estimat \n til Flokk av alle studentene ", color = "r")

```

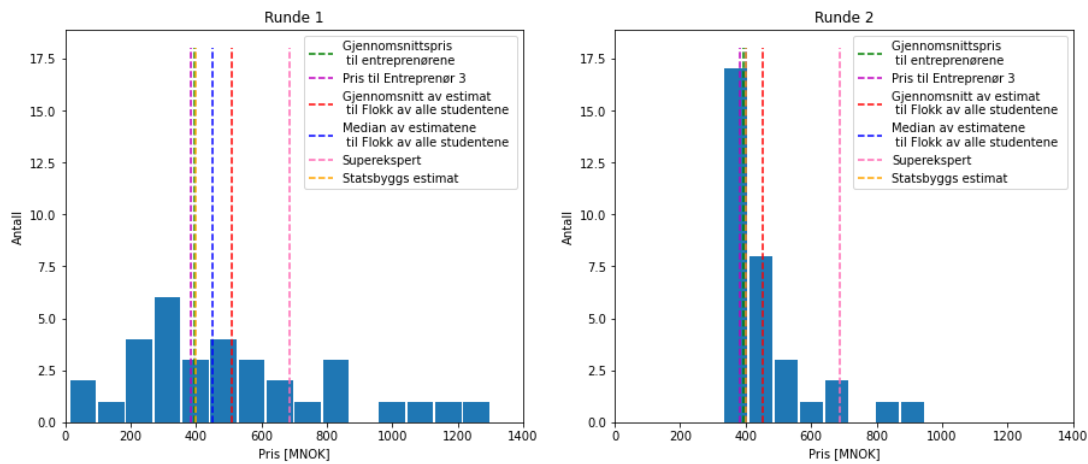
```

# Sett median pris
ax0.vlines(np.median(p_try1), 0, 18, linestyle="dashed", label =
"Median av estimatene \n til Flokk av alle studentene ", color = "b")
ax1.vlines(np.median(p_try2), 0, 18, linestyle="dashed", label =
"Median av estimatene \n til Flokk av alle studentene ", color = "b")
# Sett ekspert pris
ax0.vlines(S*1e-6, 0, 18, linestyle="dashed", label =
"Superekspert", color = "hotpink")
ax1.vlines(S*1e-6, 0, 18, linestyle="dashed", label =
"Superekspert", color = "hotpink")
# Sett Statsbygg
ax0.vlines(SB*1e-6, 0, 18, linestyle="dashed", label =
"Statsbyggs estimat", color = "orange")
ax1.vlines(SB*1e-6, 0, 18, linestyle="dashed", label =
"Statsbyggs estimat", color = "orange")

ax0.set_xlabel("Pris [MNOK]")
ax1.set_xlabel("Pris [MNOK]")
ax1.set_ylabel("Antall")
ax0.set_ylabel("Antall")
ax0.set_xlim(0,1400)
ax1.set_xlim(0,1400)
ax0.legend()
ax1.legend()
ax0.set_title('Runde 1')
ax1.set_title('Runde 2')

plt.show()

```



## Fordeling for hver av de mindre flokkene

```

fig,axs = plt.subplots(4,2, figsize = (14,22))
#Plott for flokk 1
p1_try1 = np.array(p1_try1)
p1_try2 = np.array(p1_try2)

np.random.shuffle(p1_try1)
np.random.shuffle(p1_try2)

axs[0,0].hist(np.sort(p1_try1), label = "",rwidth=0.92, bins=15)
axs[0,1].hist(np.sort(p1_try2), label = "",rwidth=0.92,bins=8)
# Sett entreprenørenes samlede pris
axs[0,0].vlines(P*1e-6, 0, 18, linestyle="dashed", label =
"Gjennomsnittspris \n til entreprenørene", color = "g")
axs[0,1].vlines(P*1e-6, 0, 18, linestyle="dashed", label =
"Gjennomsnittspris \n til entreprenørene", color = "g")
# Sett Statsbyggs pris

```



```

# Sett Entreprenør 3
axs[0,0].vlines(W*1e-6, 0, 18, linestyle = "dashed", label =
    "Pris til Entreprenør 3", color = "m")
axs[0,1].vlines(W*1e-6, 0, 18, linestyle = "dashed", label =
    "Pris til Entreprenør 3", color = "m")
# Sett aritmetisk gjennomsnitt
axs[0,0].vlines(np.mean(p1_try1), 0, 18, linestyle = "dashed", label =
    "Gjennomsnitt av estimat \n til Flokk 1", color = "r")
axs[0,1].vlines(np.mean(p1_try2), 0, 18, linestyle = "dashed", label =
    "Gjennomsnitt av estimat \n til Flokk 1", color = "r")
# Sett median pris
axs[0,0].vlines(np.median(p1_try1), 0, 18, linestyle = "dashed", label =
    "Median av estimatene \n til Flokk 1", color = "b")
axs[0,1].vlines(np.median(p1_try2), 0, 18, linestyle = "dashed", label =
    "Median av estimatene \n til Flokk 1", color = "b")

axs[0,0].set_xlabel("Pris [MNOK]")
axs[0,1].set_xlabel("Pris [MNOK]")
axs[0,1].set_ylabel("Antall")
axs[0,0].set_ylabel("Antall")
axs[0,0].set_xlim(0,1400)
axs[0,1].set_xlim(0,1400)
axs[0,0].set_ylim(0,6)
axs[0,1].set_ylim(0,6)
axs[0,0].legend()
axs[0,1].legend()
axs[0,0].set_title('Runde 1 for Flokk 1')
axs[0,1].set_title('Runde 2 for Flokk 1')

#Plott for flokk 2
p2_try1 = np.array(p2_try1)
p2_try2 = np.array(p2_try2)

np.random.shuffle(p2_try1)
np.random.shuffle(p2_try2)

axs[1,0].hist(np.sort(p2_try1), label = "",rwidth=0.92, bins=15)
axs[1,1].hist(np.sort(p2_try2), label = "",rwidth=0.92,bins=8)
# Sett entreprenørens samlede pris
axs[1,0].vlines(P*1e-6, 0, 18, linestyle = "dashed", label =
    "Gjennomsnittspris \n til entreprenørene", color = "g")
axs[1,1].vlines(P*1e-6, 0, 18, linestyle = "dashed", label =
    "Gjennomsnittspris \n til entreprenørene", color = "g")
# Sett Entreprenør 3
axs[1,0].vlines(W*1e-6, 0, 18, linestyle = "dashed", label =
    "Pris til Entreprenør 3", color = "m")
axs[1,1].vlines(W*1e-6, 0, 18, linestyle = "dashed", label =
    "Pris til Entreprenør 3", color = "m")
# Sett aritmetisk gjennomsnitt
axs[1,0].vlines(np.mean(p2_try1), 0, 18, linestyle = "dashed", label =
    "Gjennomsnitt av estimat \n til Flokk 2", color = "r")
axs[1,1].vlines(np.mean(p2_try2), 0, 18, linestyle = "dashed", label =
    "Gjennomsnitt av estimat \n til Flokk 2", color = "r")
# Sett median pris
axs[1,0].vlines(np.median(p2_try1), 0, 18, linestyle = "dashed", label =
    "Median av estimatene \n til Flokk 2", color = "b")
axs[1,1].vlines(np.median(p2_try2), 0, 18, linestyle = "dashed", label =
    "Median av estimatene \n til Flokk 2", color = "b")

axs[1,0].set_xlabel("Pris [MNOK]")
axs[1,1].set_xlabel("Pris [MNOK]")
axs[1,1].set_ylabel("Antall")
axs[1,0].set_ylabel("Antall")
axs[1,0].set_xlim(0,1400)
axs[1,1].set_xlim(0,1400)
axs[1,0].set_ylim(0,6)
axs[1,1].set_ylim(0,6)
axs[1,0].legend()
axs[1,1].legend()
axs[1,0].set_title('Runde 1 for Flokk 2')
axs[1,1].set_title('Runde 2 for Flokk 2')

```

```

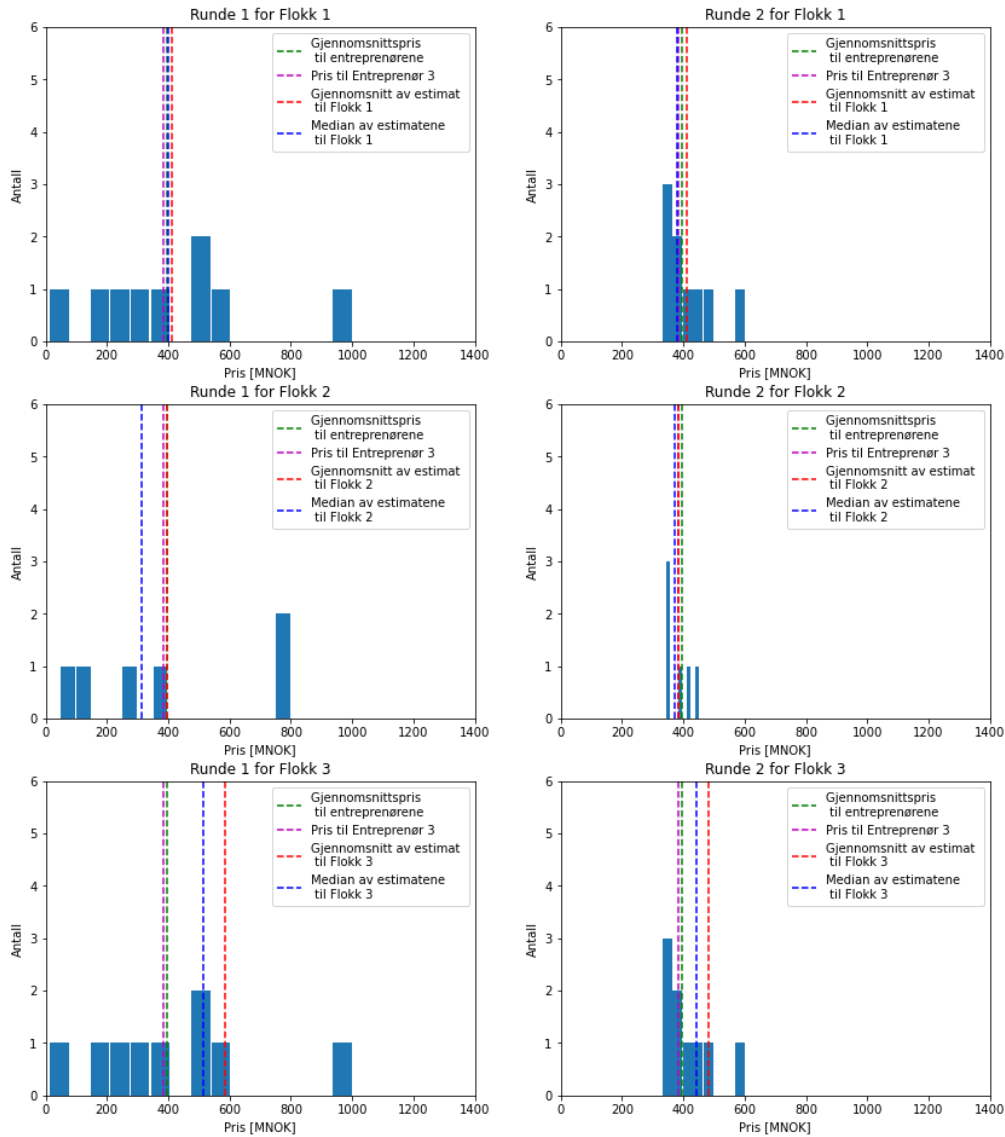
axs[3,0].set_xlabel("Pris [MNOK]")
axs[3,1].set_xlabel("Pris [MNOK]")
axs[3,1].set_ylabel("Antall")
axs[3,0].set_ylabel("Antall")
axs[3,0].set_xlim(0,1400)
axs[3,1].set_xlim(0,1400)
axs[3,0].set_ylim(0,6)
axs[3,1].set_ylim(0,6)
axs[3,0].legend()
axs[3,1].legend()
axs[3,0].set_title('Runde 1 for Flokk 4')
axs[3,1].set_title('Runde 2 for Flokk 4')

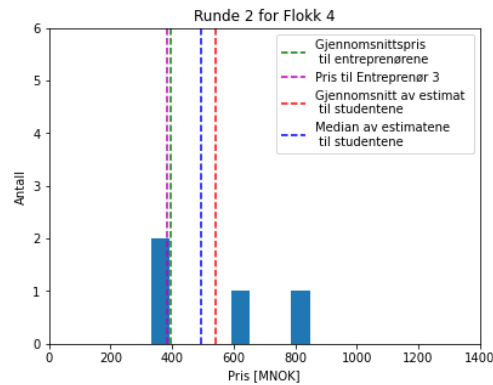
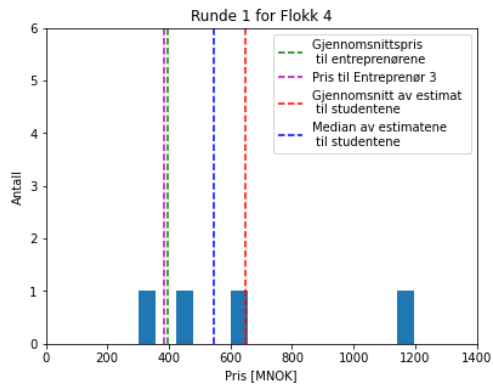
```

```

plt.show()
#plt.savefig("SamLet.pdf")

```





## Reduskjon av standardavvik. For tilfeldig valgte studenter til hver av flokkstørrelsene

```
# Corrected sample standard deviation:
def cstd(x):
    return np.sqrt(np.sum(abs(x - x.mean())**2)/(len(x)-1))

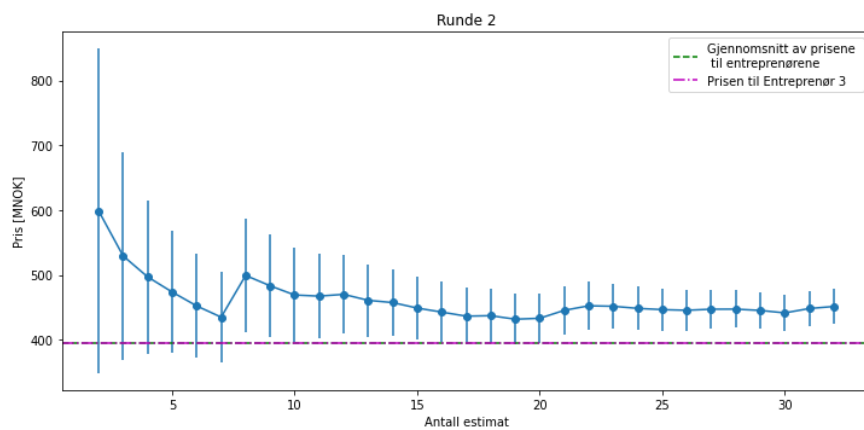
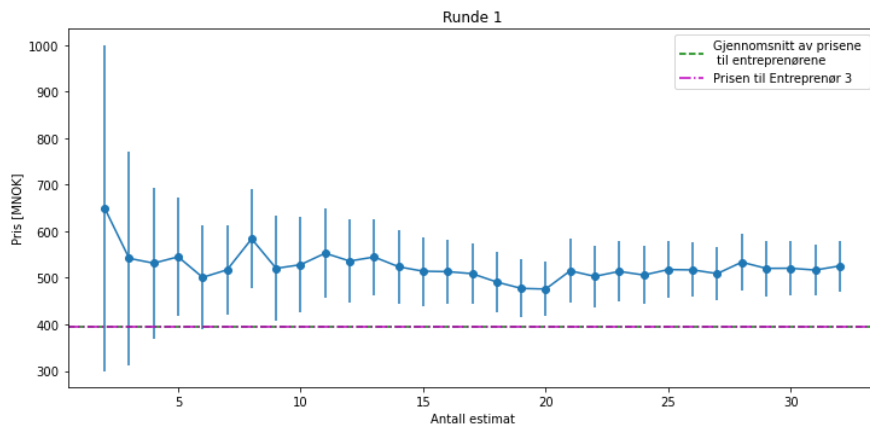
price_mean_try1 = np.zeros(len(p_try1))
price_std_try1 = np.zeros(len(p_try1))
price_mean_try2 = np.zeros(len(p_try1))
price_std_try2 = np.zeros(len(p_try1))

for i in range(2,len(p_try1)):
    price_mean_try1[i] = np.mean(p_try1[0:i])
    price_std_try1[i] = cstd(p_try1[0:i])/np.sqrt(i)
    price_mean_try2[i] = np.mean(p_try2[0:i])
    price_std_try2[i] = cstd(p_try2[0:i])/np.sqrt(i)

fig, (ax0, ax1) = plt.subplots(nrows=2, figsize = (12,12))
ax0.errorbar(np.arange(2,len(p_try1)),price_mean_try1[2:],
             yerr=price_std_try1[2:], fmt='-o')
ax0.set_title('Runde 1')

ax1.errorbar(np.arange(2, len(p_try2)),price_mean_try2[2:],
             yerr=price_std_try2[2:], fmt='-o')
ax1.set_title('Runde 2')
ax0.axhline(y = P*1e-6, color = 'g', linestyle = '--', label =
            "Gjennomsnitt av prisene \n til entreprenørene")
ax0.axhline(y = P*1e-6, color = 'm', linestyle = '-.', label =
            "Prisen til Entreprenør 3")
ax1.axhline(y = P*1e-6, color = 'g', linestyle = '--', label =
            "Gjennomsnitt av prisene \n til entreprenørene")
ax1.axhline(y = P*1e-6, color = 'm', linestyle = '-.', label =
            "Prisen til Entreprenør 3")
ax0.legend()
ax1.legend()
ax0.set_xlabel("Antall estimat")
ax1.set_xlabel("Antall estimat")
ax1.set_ylabel("Pris [MNOK]")
ax0.set_ylabel("Pris [MNOK]")

plt.show()
#Her får man nye plots for hver gang man kjører koden.
#Dette er årsaken til at plottene her ikke er nøyaktig de
#samme som i masteroppgaven
```



## Reduksjon av standardavvik for flokkstørrelser ved å bruke tilfeldig valgte delsett av dataene

```
# Start arrays
price_mean_try1 = np.zeros(len(p_try1))
price_std_try1 = np.zeros(len(p_try1))
price_mean_try2 = np.zeros(len(p_try1))
price_std_try2 = np.zeros(len(p_try1))

# Loop gjennom ulike flokkstørrelser
for i in range(2,len(p_try1)):
    # Velg N random indexes i flokk med størrelse i
    N = 1_000_000
    #idxs = [np.random.choice(33, i, replace=False) for j in range(N)]
    idxs = np.random.choice(33, (N, i), replace=True)
    comb1 = p_try1[idxs]
    price_mean_try1[i] = np.mean(np.mean(comb1,axis=1))
    price_std_try1[i] = np.std(np.mean(comb1,axis=1))/np.sqrt(i)
    comb2 = p_try2[idxs]
    price_mean_try2[i] = np.mean(np.mean(comb2,axis=1))
    price_std_try2[i] = np.std(np.mean(comb2,axis=1))/np.sqrt(i)
```

```
# Plott resultat
fig2, (ax0, ax1) = plt.subplots(nrows=2, figsize = (12,12))
ax0.errorbar(np.arange(2,len(p_try1)),price_mean_try1[2:],
             yerr=price_std_try1[2:], fmt='-o')
ax0.set_title('Runde 1')
```

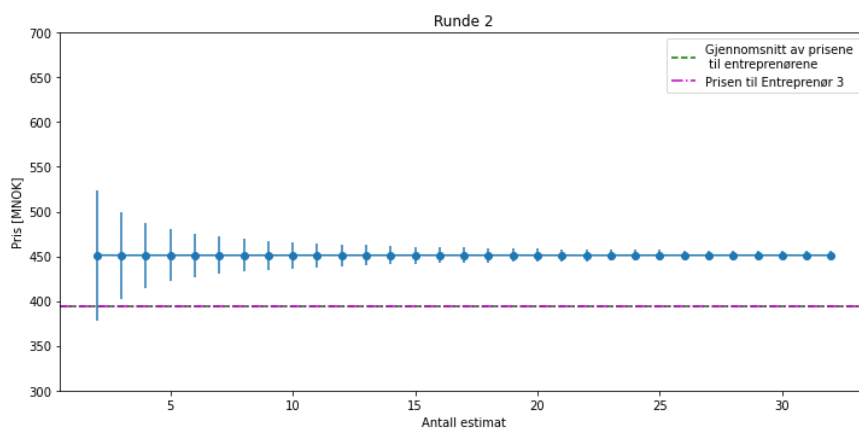
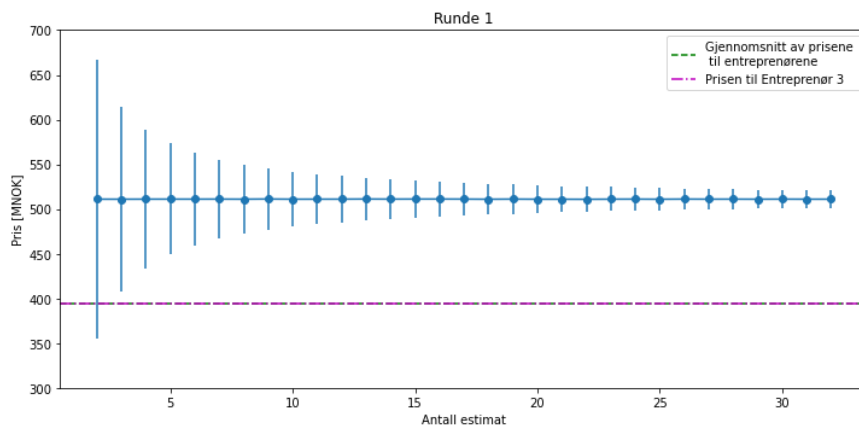
```

ax1.errorbar(np.arange(2, len(p_try2)), price_mean_try2[2:],
            yerr=price_std_try2[2:], fmt='-o')
ax1.set_title('Runde 2')

ax0.set_xlabel("Antall estimat")
ax1.set_xlabel("Antall estimat")
ax1.set_ylabel("Pris [MNOK]")
ax0.set_ylabel("Pris [MNOK]")
ax0.axhline(y = P*1e-6, color = 'g', linestyle = '--', label =
            "Gjennomsnitt av prisene \n til entreprenørene")
ax0.axhline(y = P*1e-6, color = 'm', linestyle = '-.', label =
            "Prisen til Entreprenør 3")
ax1.axhline(y = P*1e-6, color = 'g', linestyle = '--', label =
            "Gjennomsnitt av prisene \n til entreprenørene")
ax1.axhline(y = P*1e-6, color = 'm', linestyle = '-.', label =
            "Prisen til Entreprenør 3")
ax0.set_ylim(300,700)
ax1.set_ylim(300,700)
ax0.legend()
ax1.legend()

plt.show()

```



## Studentenes individuelle estimat og studentenes samlede estimat, sammenlignet med entreprenørenes sammenlagte pris, og Entreprenør 3

```
given_values = [382.740705, 394.759022]

#Først for hver av flokkene, i hver av rundene

subsets1 = np.array([p1_try1,p2_try1,p3_try1,p4_try1], dtype=object)
subsets2 = np.array([p1_try2,p2_try2,p3_try2,p4_try2], dtype=object)

subsets = np.vstack((subsets1,subsets2)).T

def make_plots_avg(given_value, row_titles):
    fig = plt.figure(figsize=(6.64, 8))
    ys = [-1, 1]
    i = 1
    for row_index, row in enumerate(subsets):
        for col_index, col in enumerate(row):
            col = np.array(col)
            ax = fig.add_subplot(4, 2, i)
            ax.set_yticks([])
            ax.set_xlim([-50, 1450])
            ax.set_ylim(ys)
            ax.set_xticks([0, 200, 400, 600, 800, 1000, 1200, 1400])
            ax.set_title(row_titles[row_index][col_index])
            i += 1
            #Lage svart linje
            ax.plot([col.min(), col.max()], [0, 0], "k")
            avg = col.mean()
            #Plot gjennomsnittsverdi av flokk
            ax.scatter(avg, 0, marker="o", color = "k", zorder=2, s=7)
            #Plot hver enkelt verdi
            numGreens = 0
            numReds = 0
            for val in col:
                color = "g"
                if abs(val - given_value) > abs(avg - given_value):
                    color = "r"
                    numReds +=1
                else:
                    numGreens += 1
            ax.scatter(val, 0, marker = "x", color = color, zorder=2, s=7)
            ax.scatter(val, 2, marker = "x", color = "g", label = numGreens)
            ax.scatter(val, 2, marker = "x", color = "r", label = numReds)
            #Plot gitt verdi
            ax.plot([given_value]*2, ys, "k--")
            ax.legend(loc="upper right")
    fig.tight_layout()
    return fig

def make_plots_med(given_value, row_titles):
    fig = plt.figure(figsize=(6.64, 8))
    ys = [-1, 1]
    i = 1
    for row_index, row in enumerate(subsets):
        for col_index, col in enumerate(row):
            col = np.array(col)
            ax = fig.add_subplot(4, 2, i)
            ax.set_yticks([])
            ax.set_xlim([-50, 1450])
            ax.set_ylim(ys)
            ax.set_xticks([0, 200, 400, 600, 800, 1000, 1200, 1400])
            ax.set_title(row_titles[row_index][col_index])
            i += 1
            #Lage svart linje
            ax.plot([col.min(), col.max()], [0, 0], "k")
            med = np.median(col)
            #Plot medianverdi til flokk
            ax.scatter(med, 0, marker="o", color = "k", zorder=2, s=7)
            #Plot hver enkelt verdi
            numGreens = 0
            numReds = 0
```

```

        for val in col:
            color = "g"
            if abs(val - given_value) > abs(med - given_value):
                color = "r"
                numReds +=1
            else:
                numGreens += 1
                ax.scatter(val, 0, marker = "x", color = color, zorder=2, s=7)
        ax.scatter(val, 2, marker = "x", color = "g", label = numGreens)
        ax.scatter(val, 2, marker = "x", color = "r", label = numReds)
        #Plot gitt verdi
        ax.plot([given_value]*2, ys, "k--")
        ax.legend(loc="upper right")
    fig.tight_layout()
    return fig
row_titles = [
    ["Runde 1 for Flokk 1", "Runde 2 for Flokk 1"],
    ["Runde 1 for Flokk 2", "Runde 2 for Flokk 2"],
    ["Runde 1 for Flokk 3", "Runde 1 for Flokk 3"],
    ["Runde 1 for Flokk 4", "Runde 2 for Flokk 4"]]

# Plot results

#For den samlede flokken

subsetsa1 = np.concatenate((p1_try1, p2_try1, p3_try1, p4_try1))
subsetsa2 = np.concatenate((p1_try2, p2_try2, p3_try2, p4_try2))

subsetsa = np.vstack((subsetsa1, subsetsa2)).T

def make_plotsa_avg(given_value, column_titles):
    fig = plt.figure(figsize=(6.64, 2))
    ys = [-1, 1]
    i = 1
    for col, title in zip(subsetsa.T, column_titles):
        # Transpose subsetsa for correct iteration
        col = np.array(col)
        ax = fig.add_subplot(1, 2, i)
        ax.set_yticks([])
        ax.set_xlim([-50, 1450])
        ax.set_ylim(ys)
        ax.set_xticks([0, 200, 400, 600, 800, 1000, 1200, 1400])
        ax.set_title(title)
        i += 1
        # Lage svart linje
        ax.plot([col.min(), col.max()], [0, 0], "k")
        avg = col.mean()
        # Plot gjennomsnittsverdi av flokk
        ax.scatter(avg, 0, marker="o", color="k", zorder=2, s=7)
        # Plot hver enkelt verdi
        numGreens = 0
        numReds = 0
        for val in col:
            color = "g"
            if abs(val - given_value) > abs(avg - given_value):
                color = "r"
                numReds += 1
            else:
                numGreens += 1
                ax.scatter(val, 0, marker="x", color=color, zorder=2, s=7)
        ax.scatter(val, 2, marker="x", color="g", label=numGreens)
        ax.scatter(val, 2, marker="x", color="r", label=numReds)
        # Plot gitt verdi
        ax.plot([given_value] * 2, ys, "k--")
        ax.legend(loc="upper right")
    fig.tight_layout()
    return fig

```

```

def make_plotsa_med(given_value, column_titles):
    fig = plt.figure(figsize=(6.64, 2))
    ys = [-1, 1]
    i = 1
    for col, title in zip(subsetsa.T, column_titles):
        # Transpose subsetsa for correct iteration
        col = np.array(col)
        ax = fig.add_subplot(1, 2, i)
        ax.set_yticks([])
        ax.set_xlim([-50, 1450])
        ax.set_ylim(ys)
        ax.set_xticks([0, 200, 400, 600, 800, 1000, 1200, 1400])
        ax.set_title(title)
        i += 1
        # Lage svart linje
        ax.plot([col.min(), col.max()], [0, 0], "k")
        med = np.median(col)
        # Plot medianverdi av flokk
        ax.scatter(med, 0, marker="o", color="k", zorder=2, s=7)
        # Plot hver enkelt verdi
        numGreens = 0
        numReds = 0
        for val in col:
            color = "g"
            if abs(val - given_value) > abs(med - given_value):
                color = "r"
                numReds += 1
            else:
                numGreens += 1
            ax.scatter(val, 0, marker="x", color=color, zorder=2, s=7)
        ax.scatter(val, 2, marker="x", color="g", label=numGreens)
        ax.scatter(val, 2, marker="x", color="r", label=(numReds))
        # Plot gitt verdi
        ax.plot([given_value] * 2, ys, "k--")
        ax.legend(loc="upper right")
    fig.tight_layout()
    return fig

column_titles1 = [("Runde 1, gjennomsnittet \n til Flokk av alle studentene "),
                  ("Runde 2, gjennomsnittet \n til Flokk av alle studentene ")]
column_titles2 = [("Runde 1, medianverdien \n til Flokk av alle studentene "),
                  ("Runde 2, medianverdien \n til Flokk av alle studentene ")]
#fig2 = make_plotsa_avg(given_values[0], column_titles1)
#make_plotsa_med(given_values[0], column_titles2)
#make_plotsa_avg(given_values[1], column_titles2)
#make_plotsa_med(given_values[1], column_titles2)

```

```

given_values = [382.740705, 394.759022]

#Først for hver av flokkene, i hver av rundene

subsets1 = np.array([p1_try1,p2_try1,p3_try1,p4_try1], dtype=object)
subsets2 = np.array([p1_try2,p2_try2,p3_try2,p4_try2], dtype=object)

subsets = np.vstack((subsets1,subsets2)).T

def add_plots_avg(fig, given_value, row_titles):
    ys = [-1, 1]
    i = 1
    for row_index, row in enumerate(subsets):
        for col_index, col in enumerate(row):
            col = np.array(col)
            ax = fig.add_subplot(5, 2, i)
            ax.set_yticks([])
            ax.set_xlim([-50, 1450])
            ax.set_ylim(ys)
            ax.set_xticks([0, 200, 400, 600, 800, 1000, 1200, 1400])
            ax.set_title(row_titles[row_index][col_index])
            i += 1
            #Lage svart linje
            ax.plot([col.min(), col.max()], [0, 0], "k")
            avg = col.mean()
            #Plot gjennomsnittsverdi av flokk
            ax.scatter(avg, 0, marker="o", color = "k", zorder=2, s=40)

```



```

#Plot hver enkelt verdi
numGreens = 0
numReds = 0
for val in col:
    color = "g"
    if abs(val - given_value) > abs(avg - given_value):
        color = "r"
        numReds +=1
    else:
        numGreens += 1
    ax.scatter(val, 0, marker = "x", color = color, zorder=2, s=40)
ax.scatter(val, 2, marker = "x", color = "g", label = numGreens)
ax.scatter(val, 2, marker = "x", color = "r", label = numReds)
#Plot gitt verdi
ax.plot([given_value]*2, ys, "k--")
ax.legend(loc="upper right")
fig.tight_layout()

def add_plots_med(fig, given_value, row_titles):
    ys = [-1, 1]
    i = 1
    for row_index, row in enumerate(subsets):
        for col_index, col in enumerate(row):
            col = np.array(col)
            ax = fig.add_subplot(5, 2, i)
            ax.set_yticks(ys)
            ax.set_xlim([-50, 1450])
            ax.set_ylim(ys)
            ax.set_xticks([0, 200, 400, 600, 800, 1000, 1200, 1400])
            ax.set_title(row_titles[row_index][col_index])
            i += 1
            #Lage svart linje
            ax.plot([col.min(), col.max()], [0, 0], "k")
            med = np.median(col)
            #Plot medianverdi til flokk
            ax.scatter(med, 0, marker= "o", color = "k", zorder=2, s=40)
            #Plot hver enkelt verdi
            numGreens = 0
            numReds = 0
            for val in col:
                color = "g"
                if abs(val - given_value) > abs(med - given_value):
                    color = "r"
                    numReds +=1
                else:
                    numGreens += 1
                ax.scatter(val, 0, marker = "x", color = color, zorder=2, s=40)
            ax.scatter(val, 2, marker = "x", color = "g", label = numGreens)
            ax.scatter(val, 2, marker = "x", color = "r", label = numReds)
            #Plot gitt verdi
            ax.plot([given_value]*2, ys, "k--")
            ax.legend(loc="upper right")
        fig.tight_layout()

row_titles = [["Runde 1 for Flokk 1", "Runde 2 for Flokk 1"],
              ["Runde 1 for Flokk 2", "Runde 2 for Flokk 2"],
              ["Runde 1 for Flokk 3", "Runde 2 for Flokk 3"],
              ["Runde 1 for Flokk 4", "Runde 2 for Flokk 4"]]

# Plot results

#For den samlede flokken

subsetsa1 = np.concatenate((p1_try1, p2_try1, p3_try1, p4_try1))
subsetsa2 = np.concatenate((p1_try2, p2_try2, p3_try2, p4_try2))

subsetsa = np.vstack((subsetsa1, subsetsa2)).T

```

```

def add_plotsa_avg(fig, given_value, column_titles):
    ys = [-1, 1]
    i = 1
    for col, title in zip(subsetsa.T, column_titles):
        # Transpose subsetsa for correct iteration
        col = np.array(col)
        ax = fig.add_subplot(5, 2, 8 + i)
        ax.set_yticks([])
        ax.set_xlim([-50, 1450])
        ax.set_ylim(ys)
        ax.set_xticks([0, 200, 400, 600, 800, 1000, 1200, 1400])
        ax.set_title(title)
        i += 1
        # Lage svart linje
        ax.plot([col.min(), col.max()], [0, 0], "k")
        avg = col.mean()
        # Plot gjennomsnittsverdi av flokk
        ax.scatter(avg, 0, marker="o", color="k", zorder=2, s=40)
        # Plot hver enkelt verdi
        numGreens = 0
        numReds = 0
        for val in col:
            color = "g"
            if abs(val - given_value) > abs(avg - given_value):
                color = "r"
                numReds += 1
            else:
                numGreens += 1
            ax.scatter(val, 0, marker="x", color=color, zorder=2, s=40)
        ax.scatter(val, 2, marker="x", color="g", label=numGreens)
        ax.scatter(val, 2, marker="x", color="r", label=numReds)
        # Plot gitt verdi
        ax.plot([given_value] * 2, ys, "k--")
        ax.legend(loc="upper right")
    fig.tight_layout()

def add_plotsa_med(fig, given_value, column_titles):
    ys = [-1, 1]
    i = 1
    for col, title in zip(subsetsa.T, column_titles):
        # Transpose subsetsa for correct iteration
        col = np.array(col)
        ax = fig.add_subplot(5, 2, 8+i)
        ax.set_yticks([])
        ax.set_xlim([-50, 1450])
        ax.set_ylim(ys)
        ax.set_xticks([0, 200, 400, 600, 800, 1000, 1200, 1400])
        ax.set_title(title)
        i += 1
        # Lage svart linje
        ax.plot([col.min(), col.max()], [0, 0], "k")
        med = np.median(col)
        # Plot medianverdi av flokk
        ax.scatter(med, 0, marker="o", color="k", zorder=2, s=40)
        # Plot hver enkelt verdi
        numGreens = 0
        numReds = 0
        for val in col:
            color = "g"
            if abs(val - given_value) > abs(med - given_value):
                color = "r"
                numReds += 1
            else:
                numGreens += 1
            ax.scatter(val, 0, marker="x", color=color, zorder=2, s=40)
        ax.scatter(val, 2, marker="x", color="g", label=numGreens)
        ax.scatter(val, 2, marker="x", color="r", label=(numReds))
        # Plot gitt verdi
        ax.plot([given_value] * 2, ys, "k--")
        ax.legend(loc="upper right")
    fig.tight_layout()

```

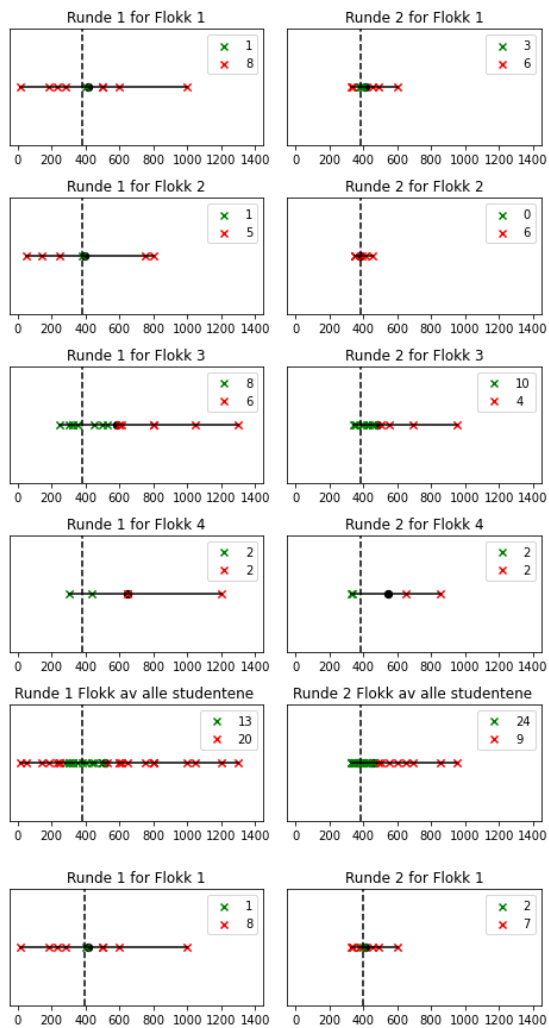
```

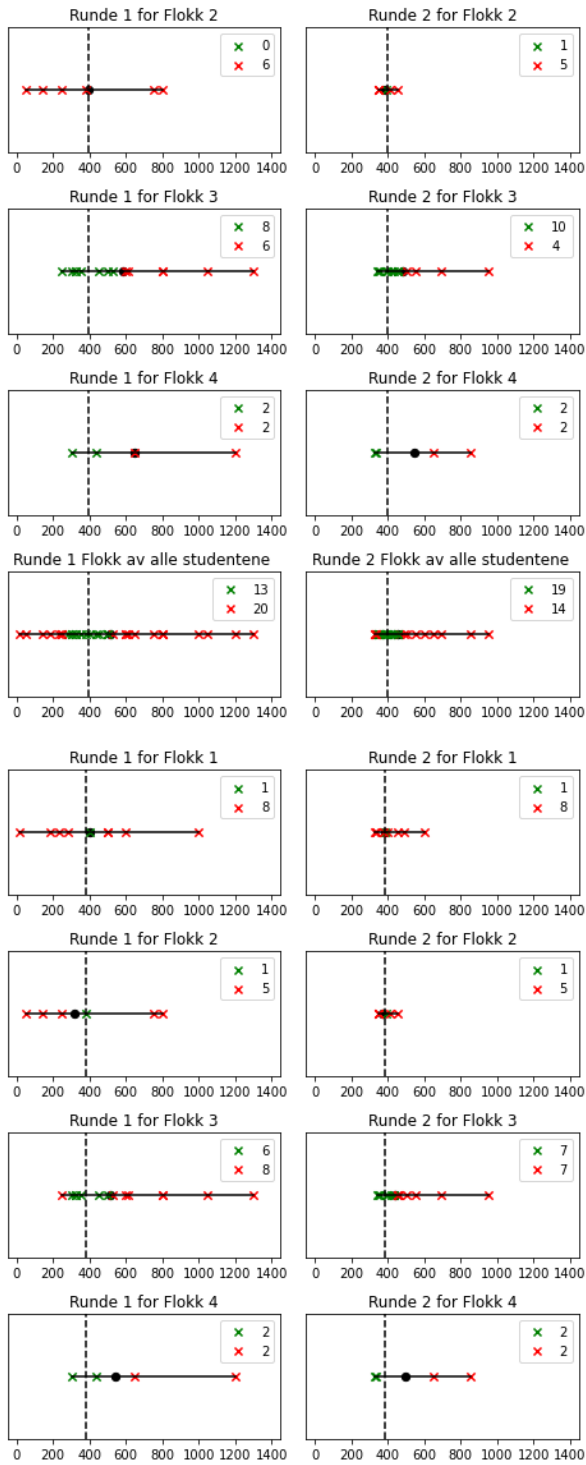
column_titles1 = [("Runde 1 Flokk av alle studentene "),
                 ("Runde 2 Flokk av alle studentene ")]
column_titles2 = [("Runde 1 Flokk av alle studentene "),
                 ("Runde 2 Flokk av alle studentene ")]

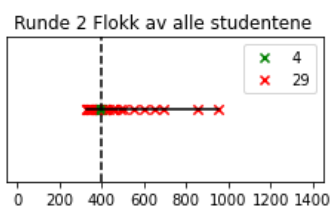
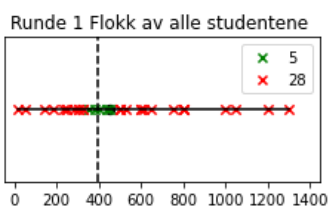
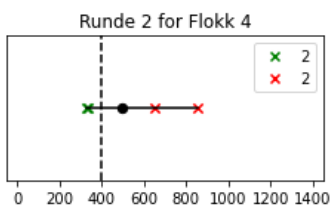
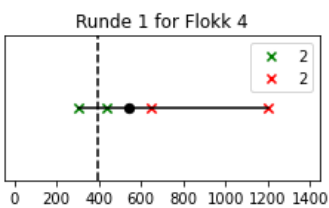
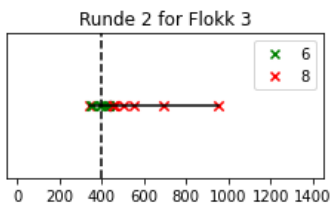
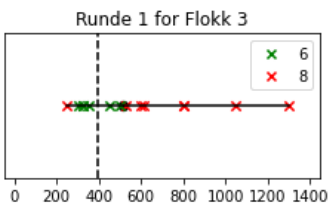
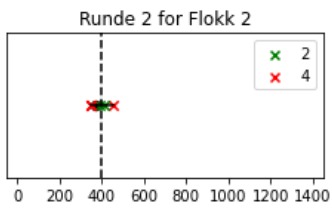
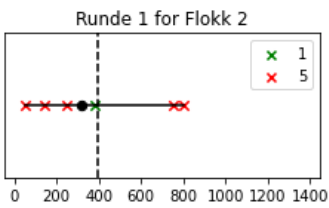
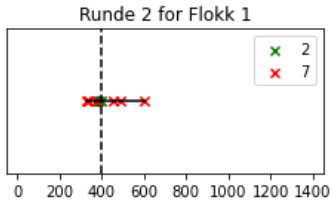
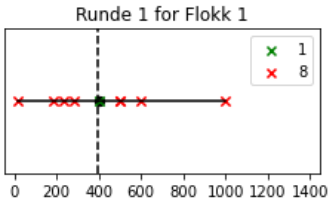
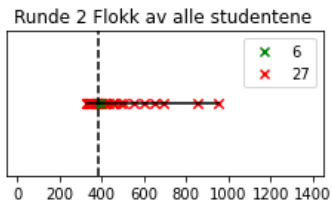
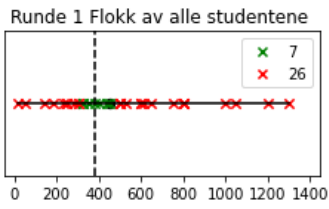
#Gjennomsnitt mot gjennomsnitt entreprenør
fig = plt.figure(figsize=(6.64, 10))
add_plots_avg(fig, given_values[0], row_titles)
add_plotsa_avg(fig, given_values[0], column_titles1)
#Gjennomsnitt mot Entreprenør 3
fig1 = plt.figure(figsize=(6.64, 10))
add_plots_avg(fig1, given_values[1], row_titles)
add_plotsa_avg(fig1, given_values[1], column_titles1)

#Median mot gjennomsnitt entreprenør
fig2 = plt.figure(figsize=(6.64, 10))
add_plots_med(fig2, given_values[0], row_titles)
add_plotsa_med(fig2, given_values[0], column_titles2)
#Median mot Entreprenør 3
fig3 = plt.figure(figsize=(6.64, 10))
add_plots_med(fig3, given_values[1], row_titles)
add_plotsa_med(fig3, given_values[1], column_titles2)

```







#### Bilag 4: Datasettene til studentene

En samling av alle estimatene gitt i hver av rundene for Flokk 1 er gitt i Tabell 27. Datasettets verdier er oppgitt i en tilfeldig rekkefølge i hver av kolonnene. Flokk 1 bestod av 9 studenter.

Tabell 27: Samling av alle estimatene gitt av den første flokken.

<b>Estimat gitt i første runde (2022 NOK)</b>	<b>Estimat gitt i andre runde (2022 NOK)</b>
500 000 000	400 000 000
500 000 000	380 000 000
185 900 000	374 000 000
400 000 000	330 000 000
230 000 000	330 500 000
12 000 000	334 000 000
280 000 000	450 000 000
600 000 000	600 000 000
1 000 000 000	490 000 000

En samling av alle estimatene gitt i hver av rundene for den Flokk 2 er gitt i Tabell 28. Datasettets verdier er oppgitt i en tilfeldig rekkefølge i hver av kolonnene. Flokk 2 bestod av 6 studenter.

Tabell 28: Samling av alle estimatene gitt av den andre flokken.

<b>Estimat gitt i første runde (2022 NOK)</b>	<b>Estimat gitt i andre runde (2022 NOK)</b>
250 000 000	345 000 000
380 000 000	450 000 000
48 000 000	347 500 000
750 150 000	390 000 000
800 000 000	412 000 000
145 000 000	350 000 000

En samling av alle estimatene gitt i hver av rundene for Flokk 3 er gitt i Tabell 29. Datasettets verdier er oppgitt i en tilfeldig rekkefølge i hver av kolonnene. Flokk 3 bestod av 14 studenter.

Tabell 29: Samling av alle estimatene gitt av den tredje flokken.

<b>Estimat gitt i første runde (2022 NOK)</b>	<b>Estimat gitt i andre runde (2022 NOK)</b>
530 000 000	450 000 000
1 300 000 000	500 000 000
326 000 000	350 000 000
600 000 000	459 420 069
800 000 000	690 000 000
350 000 000	390 000 000
325 000 000	450 000 000
500 000 000	355 000 000
300 000 000	435 000 000
1 050 000 000	400 000 000
615 000 000	550 000 000
450 000 000	420 000 000
250 000 000	341 000 000
800 000 000	950 000 000

En samling av alle estimatene gitt i hver av rundene for Flokk 4 er gitt i Tabell 30. Datasettets verdier er oppgitt i en tilfeldig rekkefølge i hver av kolonnene. Flokk 3 bestod av 14 studenter.

Tabell 30: Samling av alle estimatene gitt av den fjerde flokken.

<b>Estimat gitt i første runde (2022 NOK)</b>	<b>Estimat gitt i andre runde (2022 NOK)</b>
300 000 000	330 000 000
440 000 000	850 000 000
650 000 000	335 000 000
1 200 000 000	650 000 000

