Bacheloroppgave

NTNU
Norges teknisk-naturvitenskapelige universitet
Fakultet for økonomi
Institutt for samfunnsøkonomi

Kaja Sandvik Knudsen
Sara Rossi Stenevik

Fysisk aktivitet og studentprestasjoner

En analyse av effekten av fysisk aktivitet på studenters akademiske prestasjoner

Bacheloroppgave i Samfunnsøkonomi
Veileder: Per Tovmo
Mai 2021

NTNU
Kunnskap for en bedre verden

Kaja Sandvik Knudsen
Sara Rossi Stenevik

Fysisk aktivitet og studentprestasjoner

En analyse av effekten av fysisk aktivitet på studenters akademiske prestasjoner

Bacheloroppgave i Samfunnsøkonomi
Veileder: Per Tovmo
Mai 2021

Norges teknisk-naturvitenskapelige universitet
Fakultet for økonomi
Institutt for samfunnsøkonomi

NTNU
Kunnskap for en bedre verden

Innholdsfortegnelse

1. Introduksjon **1**

1.1 Problemstilling *1*

2. Empirisk grunnlag **2**

2.1. Fysisk aktivitet og helse *2*

2.2. Tidligere studier på fysisk helse og akademisk prestasjon *3*

3. Økonometriske modeller og teori **5**

3.1. Metode *5*

3.2. Modellens forklaringsgrad *7*

3.3. Korrelasjon *7*

3.4. Hypotesetester *8*

4. Deskriptiv analyse og presentasjon av data **9**

4.1. Om undersøkelsen *9*

4.2. Akademiske prestasjoner som måleenhet *9*

4.3. Presentasjon av variabler *9*

4.4. Deskriptiv statistikk *12*

4.5 Begrensninger *13*

5. Regresjonsanalyse **14**

5.1. Multippel regresjonsanalyse *14*

5.2 Korrelasjon *18*

5.3 Multivariat regresjonsanalyse *19*

5.4 Tester og robusthet *22*

6. Diskusjon og begrensninger **25**

7. Konklusjon **26**

Bibliografi **27**

Tabeller

Tabell 1: Variabler i modellen 10
Tabell 2: Deskriptiv statistikk N=578 (59% kvinner) 12
Tabell 3: Regresjonsanalyse 15
Tabell 4: Korrelasjonsmatrise 18
Tabell 5: Multivariat regresjonsanalyse 20
Tabell 6: Standardiserte betakoeffisienter i Modell 3 22
Tabell 7: Koeffisient og t-verdi med og uten robust verdi 24

Figurer

Figur 1: GPA vs. Steps 17

Formler

Formel 1: OLS 6
Formel 2: Justert forklaringsgrad 7
Formel 3: Korrelasjon 7
Formel 4: F-verdi 8
Formel 5: Regresjonsanalyse 14
Formel 6: Estimert regresjonsanalyse 16
Formel 7: Regresjonsmodell 1,2 og 3 19

1. Introduksjon

Høyere utdanning blir ofte sett på som nøkkelen til samfunnet, og veien mot å etablere seg som selvstendig voksen. Fokuset på høyere utdanning sender tydelige signaler om at det å prestere på skolen er viktig (Skaalvik & Skaalvik, 2013). Det er mange faktorer som kan være med på å påvirke studenters akademiske prestasjoner. Blant disse finner vi eksempelvis tidligere akademisk ytelse, foreldres inntektsnivå, bosituasjon, læringsevne, oppmøteprosent og studentenes egne vurderinger av hvor viktig det er å oppnå gode resultater på skolen. I tillegg nevnes fysisk aktivitet som en mulig medvirkende faktor til bedre karaktergjennomsnitt. I vår oppgave ønsker vi derfor å undersøke om økt fysisk aktivitet bedrer akademiske prestasjoner hos studenter.

1.1 Problemstilling

Oppgaven har som formål å besvare følgende problemstilling: **«Har fysisk aktivitet en effekt på studieprestasjoner til studenter på universitetsnivå?».**

Motivasjonen for denne problemstillingen er å se om det finnes et "gratis" prestasjonsfremmende middel som er tilgjengelig for alle. Vi er selv studenter og har kjent på mange av de positive effektene fysisk aktivitet gir. Dersom våre analyser viser at økt mengde fysisk aktivitet også bedrer akademisk prestasjon hos studenter ved universitetet, kan dette brukes i diskusjoner om fysisk aktivitet burde tilrettelegges for, og inkluderes i studieløpet hos studenter ved universitetet.

I denne oppgaven skal vi foreta en empirisk undersøkelse for å se om fysisk aktivitet har effekt på studenter sine akademiske prestasjoner. Ved å analysere studentenes daglige aktivitet ønsker vi å se om økte akademiske prestasjoner kan forklares ved hjelp av variablene i vår analyse. Formålet med denne oppgaven er å kunne bygge opp under, eller svekke, hypotesen om at fysisk aktivitet bedrer akademiske prestasjoner. Vi vil rette vårt fokus mot studenter på universitetsnivå.

Vi har funnet data med ulike variabler for fysisk aktivitet hos en gruppe studenter, og vil undersøke effektene disse variablene har på akademisk ytelse, målt ved et gjennomsnitt av semesterkarakterer (Grade point average heretter «GPA»). Vi antar, på bakgrunn av WHO sine anbefalinger, at flere minutter med aktivitet, opp til et visst punkt, i de forskjellige intensitetssonene (150-300 minutter i moderat intensitet og 75-150 minutter i høy intensitet) vil føre til bedre helse og prestasjon (World

Health Organization, 2020). Vi antar at vanlig mosjon av en viss varighet i løpet av uken har mest å si for aktivitetsnivået, og at det kanskje er den enkleste treningsformen for mange å kunne opprettholde over tid.

I kommende kapittel vil forholdet mellom fysisk aktivitet og skoleprestasjon utdypes på et teoretisk grunnlag. I kapittel 3 vil vi gå gjennom den metodiske tilnærmingen og forklare de statistiske analysene. Kapittel 4 presenterer den deskriptive statistikken som gir utgangspunkt for analysen som kommer i kapittel 5. Til slutt vil resultatene bli presentert, og funnene diskutert i henholdsvis kapittel 6 og 7.

2. Empirisk grunnlag

Amerikanere går i gjennomsnitt kun fem tusen skritt daglig. Dette viser at amerikanerne er mindre fysisk aktive enn mennesker i andre land (Johnson, 2016). Til sammenligning går en gjennomsnittlig australsk borger ti tusen daglige skritt. I tillegg til at den gjennomsnittlige amerikaneren er mindre fysisk aktiv har USA også sett en nedgang i akademiske prestasjoner sammenlignet med andre land. PISA-undersøkelsen (Programme for International Student Assessment) som måler leseevne, matematikkforståelse og realfagskompetanse hos 15-åringer rangerer amerikanske elever midt på tabellen. Av 71 land kom USA på 38. plass når det gjelder leseevne og matte- og realkompetanse hos 15-åringer (DeSilver, 2015).

Målt mot andre land regnes amerikanske borgere som mindre fysisk aktive, og under gjennomsnittet i akademiske prestasjoner (Johnson, 2016). At USA havner bak, både når det gjelder helsetrender og akademiske ytelser, gjør det interessant å undersøke om fysisk aktivitet og akademisk ytelse kan sees i sammenheng.

2.1. Fysisk aktivitet og helse

Per definisjon er fysisk aktivitet «all kroppslig bevegelse som er utført av skjelettmuskulatur, og som resulterer i en vesentlig økning i energiforbruket utover hvilenivå» (FHI, 2017). Som definisjonen impliserer, er fysisk aktivitet et svært vidt begrep. Begrepet gaper over flere former for aktivitet. Idrett, mosjon, kroppsøving, friluftsliv, turer og trim er former for fysisk aktivitet,

men også dagligdagse gjøremål som hagearbeid, å sykle til skolen og lek regnes som fysisk aktivitet. Den totale mengden fysisk aktivitet i løpet av en dag utgjør det fysiske aktivitetsnivået. Fysisk aktivitet kan gjennomføres med ulik intensitet, frekvens og varighet. Intensiteten forteller hvor hard aktiviteten er, og deles inn i lav, moderat og høy intensitet. Frekvens sier noe om hvor ofte aktiviteten gjentas. Varighet forteller hvor lenge aktiviteten pågår (Bahr, 2008).

En studie av Tudor-Locke, et al. (2011) anbefaler at en gjennomsnittlig voksen går 10 000 skritt per dag for å opprettholde god fysisk helse. Anbefalingen får støtte av Wattanapisit & Thanamee (2017) som gjennomførte en metaanalyse for å se effektene av å gå 10 000 skritt per dag. Studien fant at 10 000 skritt daglig kan være med på å forbedre kroppssammensetningen ved å redusere kroppsvekt, kroppsmasseindeks, midjeomkrets og hofteomkrets. I tillegg kunne man se en forbedring av blodtrykksverdier hos utvalget som var med i analysen. På bakgrunn av dette kan man si at antall daglige skritt kan være en god indikator på et individs fysiske helse (Booth, et al., 2014).

Verdens Helseorganisasjon, WHO, kom i 2020 med nye anbefalinger for hvor mye aktivitet som er optimalt for fysisk og psykisk helse (World Health Organization, 2020). Resultater fra et stort forskningsprosjekt med over 44 000 deltakere fra fire ulike land ligger til grunn for de nye anbefalingene. Personene ble utstyrt med aktivitetsmålere som overvåket aktivitetsnivået. Resultatene viste at personer som er fysisk inaktive i 10 eller flere timer daglig har en betydelig økt risiko for tidlig død. Studien viser også at moderat til hard fysisk aktivitet på 30-40 minutter hver dag vil være med å svekke risikoen betydelig. Anbefalingen fra WHO er en ukentlig innsats på 150-300 minutter med moderat fysisk aktivitet, eller 75-100 minutter hard fysisk aktivitet.

2.2. Tidligere studier på fysisk helse og akademisk prestasjon

En systematisk litteraturgjennomgang for å undersøke om helseprogrammer på skolene forbedret akademiske prestasjoner hos elevene, viste positive sammenhenger mellom skolebaserte fysiske aktiviteter og akademiske prestasjoner (Tomporowski, Lambourne, & Okumuraa, 2011). For barn og ungdom (alder 7-19 år) ble det rapportert sammenhenger mellom fysisk aktivitet, kardiovaskulær kondisjon og akademiske prestasjoner (Shin & So, 2012) (Booth, et al., 2014)

(Ardoy, et al., 2014). Andre studier fant ingen støtte for at fysisk aktivitet forbedret prestasjonen (LeBlanc, 2014). Når det gjelder variasjoner i forbindelse med kjønn er forskningsfunn sprikende i forholdet mellom akademisk prestasjon og fysisk aktivitet. I en studie av ungdommer fant man ingen sammenheng mellom fysisk aktivitet og akademisk ytelse hos gutter, mens man fant en positiv sammenheng hos jenter (Martínez-Gómez, et al., 2012). En annen studie, av koreanske ungdomsskolestudenter, fant at intensiv fysisk aktivitet var positivt korrelert med akademisk ytelse hos gutter, mens moderat fysisk aktivitet var positivt korrelert med akademiske prestasjoner hos begge kjønn (So, 2012).

En studie av Khan (2017) undersøkte om fysisk aktivitet og stressnivå var korrelert, og fant at fysisk aktivitet kan bidra til å redusere stress og forbedre GPA. GPA er et tall på hvor godt studentene scorer i løpet av studieløpet sitt. GPA beregnes ut fra et gjennomsnitt av alle karakterene studenten har oppnådd i løpet av semesteret, og skalaen går fra 0 til 4.

Fordi fysisk aktivitet kan se ut til å være knyttet til lavere stressnivå og høyere GPA, kan det antas at en som driver med regelmessig fysisk aktivitet vil ha høyere GPA enn en som er lite fysisk aktiv eller er stillesittende. Det finnes imidlertid flere forskere som ikke tror at GPA øker som resultat av å delta i fysisk aktivitet. Studier har funnet en positiv sammenheng mellom fysisk aktivitet og akademiske prestasjoner (målt i GPA) for studenter på et helse- og naturvitenskapsprogram ved et amerikansk universitet. Likevel mener forskere at for å kunne forbedre den enkeltes akademiske prestasjon, må han bruke mer tid på å fokusere på sitt fagfelt (Xu & Sansgiry, 2018), og at mer fokus på studiene vil medføre mindre tid til fysisk aktivitet (Gonzalez, Hernandez, Coltrane, & Mancera, 2014).

En studie, utført på medisinstudenter, fant en positiv sammenheng mellom fysisk aktivitet og studentenes GPA. Denne studien viste også en lineær økning i de deltakende medisinstudentenes generelle treningsnivå, som korrelerte med deres kumulative GPA (Stephens, Dong, & Durning, 2015). Studien konkluderte med at individer som begynner å delta i fysisk aktivitet, og fortsetter disse aktivitetene mens de går på universitetet, sannsynligvis vil oppleve forbedret GPA gjennom studieperioden.

En annen studie som har sett på sammenhengen mellom fysisk aktivitet og akademisk prestasjon hos studenter er «Excercise improves academic preformance» fra NHH (Cappelen, Charness, Ekström, Gneezy, & Tungodden, 2017). Dette var en intervensjonsstudie som ga universitetsstudenter insentiver til å trene, for å se effekten av dette på akademisk prestasjon. Studien hadde 778 deltakere, der 398 fikk gratis treningsabonnement og en premie på 1000 kr dersom de besøkte treningssenteret minst 2 ganger i uken i 7 av totalt 11 uker. Kontrollgruppen bestod av 380 studenter. Disse fikk ikke gratis treningsabonnement eller pengepremie. Funnene i denne studien var at det å gi intensiver til å trene på kort sikt, ga studentene vanedannende positiv adferd på lang sikt. Insentivene førte til økt treningsmengde og studien viste at det var en sterk årsaks-effekt av dette på akademisk prestasjon.

De fleste studiene viser altså en positiv sammenheng mellom fysisk aktivitet og akademiske prestasjoner.

3. Økonometriske modeller og teori

I dette kapittelet vil vi gjennomgå den økonometriske modellen, minste kvadraters metode, som er modellen vi skal bruke i våre analyser. I tillegg vil vi presentere hypotesetester, korrelasjon og forklaringsgrad og hvordan disse metodene blir anvendt til å besvare vår problemstilling.

3.1. Metode

I vår oppgave vil vi analysere en eventuell årsakssammenheng mellom fysisk aktivitet og akademisk prestasjon. Ettersom vi skal se på hvordan flere eksogene variabler vil påvirke den endogene variabelen må vi bruke multippel regresjonsanalyse. Minste kvadraters metode («Ordinary Least Squares» heretter OLS) gir oss den beste linjen som gjør at summen mellom den kvadrerte avstanden fra datapunktene og linjen er minst mulig (Thomas, 2005, ss. 266-267). Ved hjelp av OLS kan vi estimere effekten av de uavhengige parameterne X_i på den avhengige utfallsvariabelen Y. OLS gjør det mulig å undersøke hvorvidt det er en sammenheng mellom fysisk aktivitet og akademiske prestasjoner. Vi tar utgangspunkt i en lineær regresjonsmodell:

$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_i x_i + u_i$$

Formel 1: OLS

Den faktiske verdien på *Y* er ikke nødvendigvis lik forventet verdi på *Y*. Vi forutsetter lineær årsakssammenheng mellom utfallsvariabelen *Y*, og en eller flere av de uavhengige variablene.

Restleddet, *u*, er gitt som differansen mellom forventet verdi *E(Y)* i vår predikerte modell og faktisk verdi på *Y*. Størrelsen på restleddet, *u*, avhenger av de ikke-målbare faktorene som spiller inn på den forventede verdien på *Y*. Dette kan være faktorer som ikke er målbare, eller faktorer som vi ikke har data på. Jo større verdi dette leddet har, desto større andel av variasjonen i den avhengige variabelen vil forklares av variabler utenfor modellen.

Denne modellen er avhengig av visse forutsetninger for å kunne være effektiv, og gi et forventningsrett estimat (Thomas, 2005, ss. 356-359). Forutsetninger for restleddet er som følger:

- $E(u_i) = 0$: forventet verdi på restleddet settes lik null. Dette betyr at avviket mellom estimerte verdier og observerte verdier er lik 0.
- $V(u_i) = \sigma^2 < \infty$: Restleddet er homoskedastisk. Residualene har lik varians for alle X.
- $Cov(u_i, u_j) = 0$: Korrelasjonen mellom restledd er lik null. Dette betyr at restleddene er uavhengige av hverandre.
- $u_i \sim N(0, \sigma^2)$: Restleddet er normalfordelt, og har et gjennomsnitt på 0.

Forutsetninger til de uavhengige variablene:

- De uavhengige variablene kan ikke være stokastiske. Verdien til X er ikke tilfeldig, og avhenger ikke av andre variabler.
- Selv om de uavhengige variablene holder seg konstante over flere målinger betyr ikke dette at verdien på Y gjør det. Grunnen er at variasjonen i restleddet, u, ikke kontrolleres av størrelsen på verdiene på X.
- At utvalget (N) øker, betyr ikke nødvendigvis at variansen på de uavhengige variablene vil øke.

Fordelen ved å benytte kvantitativ metode er at vi kan se på kvantifiserte effekter. En slik metode egner seg godt til å sammenligne og analysere store mengder data.

3.2. Modellens forklaringsgrad

R^2 er modellens forklaringsgrad. Denne forteller oss hvor mye modellen forklarer av variasjonen i den avhengige variabelen (Thomas, 2005, ss. 273-277). Justert R^2 tar hensyn til at R^2 øker når flere variabler legges til modellen. Justert R^2 veier andel forklart varians mot antall variabler. Vi vil videre i oppgaven bruke justert R^2 ettersom denne tar hensyn til antall variabler, slik at ikke modellen får bedre forklaringsverdi uten grunnlag. Vi bruker R^2 for å se hvor godt regresjonslinjen beskriver datasettet (Thomas, 2005, s. 421).

$$R^2 = \frac{SSE}{SST}, \quad 0 < R^2 < 1$$

Formel 2: Justert forklaringsgrad

3.3. Korrelasjon

Korrelasjon måler sammenhengen mellom to eller flere variabler. Dette er et statistisk mål som enten kan være positivt eller negativt. Dersom to variabler har positiv korrelasjon, vil det bety at *Y* øker når *X* øker. Motsatt vil *Y* reduseres når *X* øker ved negativ korrelasjon. Vi kan bruke korrelasjonskoeffisienten *(p)* for å sammenligne korrelasjon. Denne har en verdi mellom -1 og 1. Ved en korrelasjonskoeffisient på 1, har vi perfekt korrelasjon. Dersom koeffisienten er tilnærmet lik 0, er størrelsene uavhengig og dette tilsier at det ikke er noen sammenheng mellom variablene. Korrelasjon indikerer ikke nødvendigvis kausalitet. Selv om variabler har sterk korrelasjon, betyr ikke dette at *X* forårsaker *Y*. Det kan være andre variabler som er årsaken til at det er korrelasjon mellom *X* og *Y* (Thomas, 2005, s. 258).

$$R = \frac{\Sigma(x-\bar{x})(y-\bar{y})}{\sqrt{\Sigma(x-\bar{x})^2}\sqrt{\Sigma(y-\bar{\bar{y}})^2}} \qquad -1 \leq R \leq 1$$

Formel 3: Korrelasjon

3.4. Hypotesetester

Hypotesetester anvendes for å avgjøre om OLS-estimatorene stemmer med virkeligheten under et gitt signifikansnivå. Metoden benytter en nullhypotese (H_0) og en alternativhypotese (H_A), hvor H_0 er konservativ, og man undersøker om man kan forkaste denne. Alternativhypotesen er en komplementær hypotese (Thomas, 2005, ss. 369-371) . Vi kan teste individuelle parametere eller utføre multiple tester med flere parametere. En t-test undersøker om effekten av enkeltparametere på avhengig variabel er signifikant når populasjonsvariansen er ukjent (Thomas, 2005, s. 428). For å kunne teste multiple hypoteser samtidig og se om de sammen har signifikant effekt på avhengig variabel kan vi benytte en F-test (Thomas, 2005, ss. 415-419). F-testen baseres på den relative endringen i forklaringskraft med, og uten restriksjoner (r) på modellen. F-verdien til betakoeffisientene er gitt ved formel 4.

$$F - stat = \frac{\frac{(R_{ur}^2 - R_r^2)}{q}}{\frac{1 - R_{ur}^2}{n - k - 1}}$$

Formel 4: F-verdi

Vi kan vurdere signifikansen til koeffisientene på to ulike måter; enten ved å vurdere koeffisientens t-verdi eller p-verdi (signifikanssannsynlighet). Dersom absoluttverdien til t-verdien er høyere enn kritisk verdi på et valgt signifikansnivå, er det grunn til å forkaste nullhypotesen og variabelen kan tolkes som signifikant. P-verdien beskrives som en standardisering av t-verdien. Dette er sannsynligheten for å observere at betakoeffisienten er lik observert verdi dersom nullhypotesen stemmer. Nullhypotesen forkastes dersom p-verdien er lavere enn signifikansnivået. I oppgaven vil vi bruke p-verdien for å tolke signifikansen til koeffisientene.

4. Deskriptiv analyse og presentasjon av data

I denne delen vil vi gi en grundig beskrivelse av undersøkelsen vi har valgt å se på. Først vil vi forklare de ulike variablene vi har valgt å fokusere på, og hvorfor vi har valgt disse variablene. Til slutt presenterer vi den deskriptive statistikken som gir utgangspunkt til analysen i kapittel 5.

4.1. Om undersøkelsen

Oppgaven tar utgangspunkt i undersøkelsen «Fitbits, field-tests, and grades» (Broaddus, et al., 2018). Dette er et datasett samlet inn gjennom høstsemesteret i 2017 fra 581 førsteårsstudenter på Oral Roberts University i Tulsa, Oklahoma i USA. Undersøkelsen gikk ut på at studentene brukte smartklokken Fitbit til å dokumentere antall skritt og aktive minutter i løpet av en dag. Studentenes gjennomsnittlige semesterkarakter ble lagt til på slutten av semesteret. Studentene gjennomførte også en *1-mile field test*, altså en løpetest på tid.

4.2. Akademiske prestasjoner som måleenhet

Akademiske prestasjoner kan måles på forskjellige måter; (1) Studentenes egen refleksjon over sine akademiske prestasjoner, (2) Studentenes subjektive vurdering av deres faglige ytelse sammenlignet med sine jevnaldrende, eller (3) En ekstern objektiv lærerevaluering av studentenes samlede akademiske prestasjon, som angitt av de endelige karakterene studentene har oppnådd i sine vurderinger. I våre analyser blir et gjennomsnitt av studentenes semesterkarakterer (GPA) brukt som mål på studentens akademiske prestasjon.

4.3. Presentasjon av variabler

Vi har valgt ut variabler som vi mener er relevante og tilgjengelige for vår analyse. Disse kan gi oss en god indikasjon på om fysisk aktivitet påvirker studenters prestasjoner. I datasettet var det totalt ti variabler. Vi har valgt åtte variabler vi vil fokusere på for å få bedre forståelse for om fysisk aktivitet har en positiv innvirkning på akademisk prestasjon. Vi har utelatt de variablene som ikke representerte det vi ønsket å fokusere på i denne oppgaven. Disse var *Age* og *Mode*.

Datasettet inneholder kun observasjoner fra førsteårsstudenter på universitetet, derfor er det lite variasjon i utvalget (20-23 år). På bakgrunn av dette har vi ekskludert variabelen *Age.* Vi har heller

ikke inkludert variabelen *Mode,* ettersom dette er en variabel som forklarer om studentene valgte å gå eller løpe 1-mile field-testen. Vi ekskluderte denne variabelen fra våre analyser, da vi antar at studentene som valgte å gå field-testen ikke er i god nok form til å løpe den. Andelen som valgte å gå var også svært lav (14%). Tabell 1 viser en oversikt over variablene som er inkludert i regresjonsmodellen.

Navn	**Beskrivelse**
GPA	Studentenes gjennomsnittlige semesterkarakter
Steps	Antall skritt daglig (hverdagsaktivitet)
Cardio	Antall minutter i Cardio sone (80-84% av makspuls)
Peak	Antall minutter i Peak sone (85% + av makspuls)
Fatburn	Antall minutter i Fatburn sone (50-69% av makspuls)
Minutes	Minutter brukt på 1-mile test
Female	Hvorvidt respondenten er kvinne
Life Score	Egenvurdert mål på livsstil

Tabell 1: Variabler i modellen

Først vil vi ta utgangspunkt i den avhengige variabelen GPA, som er studentenes gjennomsnittlige semesterkarakter. Dette er en variabel som er representativ for hvordan studenten presterer på skolen og vil være utgangspunktet for analysen vår. GPA er et gjennomsnitt av studentenes karakterer gjennom semesteret i mindre oppgaver, prosjekter, deltagelse og quizer. Verdien av variabelen er på kontinuerlig form.

Variabelen *Minutes* inneholder resultatet til studentene på en løpetest. Den forteller hvor lang tid studentene brukte på en engelsk mil, som tilsvarer ca. 1,6 kilometer. Variabelen gir oss informasjon om studentenes fysiske form. Dersom de er bedre trent, løper de testen på kortere tid.

Variabelen *Female* forteller oss hvorvidt studenten var kvinne. Vi antar at det kan være forskjeller i aktivitet, prestasjon på test og intensitetsnivå når det gjelder kjønn, og ønsket å ta høyde for dette. Variabelen *Female* er en dikotom dummy variabel på nominalnivå som angir at dersom studenten er en mann er *Female*=0 og dersom studenten er en kvinne er *Female*=1. Studien bestod av 344

kvinner (59%) og 237 menn (41%). Referansekategorien er med andre ord mann, når denne variabelen inkluderes i en modell.

Variabelen *Lifescore* er et mål på studentenes egenvurderte livsstil. Studentene gjennomgikk en undersøkelse der de evaluerte sin egen personlig helse, narkotika- og alkoholbruk, fysisk form, psykologisk tilstand, åndelig tilstand, personlig adferd og ernæring. Undersøkelsen baserte seg på 35 spørsmål innenfor disse emnene. Basert på evalueringene av studentene ble livsstil evaluert på en skala fra veldig sunn (<40) til veldig usunn (>130). Vi betrakter variabelen *Lifescore* som en kontinuerlig variabel. Denne variabelen er inkludert, da den kan bidra til å fange effekter fra andre livsstilsvariabler enn de aktivitetsspesifikke variablene (*Steps, Cardio, Fatburn* og *Peak*) vi har inkludert i modellene våre. For eksempel er det rimelig å anta at en student som opplyser om høy narkotikabruk og/eller alkoholkonsum i undersøkelsen, har en mindre aktiv livsstil, enn studenter som opplyser om ingen narkotikabruk og/eller alkoholkonsum. I tillegg kan det tenkes at en student med god psykologisk tilstand, personlig adferd, helse og ernæring, vil være mer aktive og verdsette gode karakterer høyere.

Dataene fra Fitbit var aggregert og hadde følgende variabler:
Variabelen *Steps* inneholder det estimerte gjennomsnittlige antall skritt tatt av en student per dag. Dette representerer hverdagsaktivitet som er en viktig faktor når det kommer til aktivitet. Vi tror denne faktoren påvirker aktiviteten og dermed resultatet til en student i størst grad.

Variabelen *Fatburn* gir det totale antall aktive minutter for hver student. Minutter i *Fatburn*-sonen korresponderer til en puls som er 50-69% av makspuls. Dette innebærer at *Fatburn*-sonen for studentene typisk ligger mellom 100-140 slag i minuttet.

Variabelen *Cardio* måler det totale antall minutter med en puls i *Cardio*-sone. *Cardio*-sonen er en middels høy intensitet pulssone som korresponderer til en puls som er 70-84% av makspuls. Dette vil si en puls på typisk 140-170 pulsslag i minuttet hos studentene.

Variabelen *Peak* måler det totale antall minutter studentene var i intensitetssonen *Peak*. Denne sonen korresponderer med en puls på mer enn 85% av makspuls og er typisk høyere enn 170 pulsslag i minuttet.

4.4. Deskriptiv statistikk

Vi vil nå presentere deskriptiv statistikk som gir en oversikt over variablene og deres egenskaper. Vi vektlegger mål på sentraltendenser som gjennomsnitt, standardavvik samt minimums- og maksimumsverdier, basert på observasjonene vi har tilgjengelig.

Variabler	Gjennomsnitt	Standardavvik	Min	Max
GPA	3,19	0,76	0	4
Steps	10 282,07	2985,02	986	20331
Peak	2,88	14,19	0	267,1
Cardio	9,05	16,50	0	180,65
Fatburn	149,50	96,50	1,37	729,67
Minutes	10,17	3,14	5,38	21,07
Life Score	71,20	18,62	35	130

Tabell 2: Deskriptiv statistikk N=578 (59% kvinner)

Fra de deskriptive dataene for vår avhengige variabel GPA kan vi se av Tabell 2 at elevene ligger på et gjennomsnitt på 79% av maks karakter. Vi skal se om det eksisterer noe grunnlag for å si at økt aktivitet fører til bedre prestasjon på studiet.

Studentene gikk i gjennomsnitt 10 282 skritt. Dette er i tråd med anbefalingen om å begrense stillesitting og være aktiv minst 60 minutter om dagen (Helsedirektoratet, 2019). Det er interessant å observere fra standardavviket at spredningen i antall skritt gjennomført daglig er relativt stor. Studentene som gjorde det best på løpetesten har betydelig bedre tid enn de som gjorde det dårligst, og her er det også ganske stor spredning i resultatene. *Lifescore* har gjennomsnittlig score på 71.2, noe som tilsvarer gjennomsnittlig sunn livsstil. Minimum for egenvurdering av studentenes livsstil var 35, som tilsvarer en veldig sunn livsstil. Maksimum var 130, noe som tilsvarer en usunn livsstil.

I observasjonene i datasettet så vi at det minste antall observerte daglige skritt var null. Dette indikerer at noen av studentene har gått null skritt i løpet av en dag, noe som er lite troverdig da bare en tur på toalettet vil gi utslag i antall daglige skritt. Også andre variabler fra Fitbit som *Fatburn*, *Cardio* og *Peak* viser minimumsverdi på null. På bakgrunn av dette er det rimelig å anta at noen fra utvalget ikke har hatt på seg Fitbit-klokken under datainnsamlingen. Dette kan gi skjevhet i resultatene. For å ta høyde for disse skjevhetene droppet vi alle observasjoner som viste daglig antall skritt på under 100. Totalt ble 3 observasjoner ekskludert fra våre analyser. Dette resulterte i at minimumsverdien var på 986 skritt som er betraktelig mer realistisk.

4.5 Begrensninger

Datasettet er samlet inn kun for elevene som tok faget «Introduction to whole person education». Et optimalt datasett hadde hatt et sannsynlighetsutvalg på nasjonalt nivå, om ikke internasjonalt nivå. Studentene som tok dette faget fikk vite at ved å gjennomføre 10 000 skritt om dagen, stod dette for 10% av karakteren i dette faget. For eksempel er det dermed mulig at noen av studentene produserte «falske» skritt ved å riste på hånden slik at aktivitetsklokken telte skritt. Dette kan potensielt føre til en inflasjon av skritt som inngår i variabelen *Steps*. De andre variablene har puls som forutsetning og er derfor vanskelig for studentene å fremstille kunstig. Dermed antar vi at disse variablene ikke har blitt manipulert. En annen begrensing kan være hvor nøyaktig Fitbit måler aktivitet[1]. Når det er sagt, kan vi ikke med sikkerhet si at slike sportsklokker måler variablene som inngår i modellen med en nøyaktighet som er nødvendig for å trekke en konklusjon. Variabelen *Lifescore* har ikke gjennomgått streng validering, så dette kan også begrense resultatene våre. Det er også verdt å nevne at denne variabelen er en egenvurdert undersøkelse, så vi har ingen garanti for at studentene svarer ærlig eller gir pålitelige svar.

---

[1] Flere studier har undersøkt Fitbit og dens nøyaktighet. En studie undersøkte antall skritt registrert i Fitbit og sammenlignet med manuell telling av skritt (Lynne, et al., 2018). Studien viste høy korrelasjon mellom trinn som ble registrert i Fitbit og trinn som ble registrert manuelt. I en annen studie av Thea, et al., (2015) var funnene at Fitbit Zip var den mest pålitelige enheten for måling av fysisk aktivitet sammenlignet med ni andre fysiske aktivitetssporingsenheter. Samlet har disse studiene vist at Fitbit er en tilstrekkelig nøyaktig og pålitelig enhet for å måle fysisk aktivitet.

Det kunne vært interessant å undersøke variabler som ikke inngår i dette datasettet, og om disse variablene har noen effekt på skoleprestasjonene. Dette kunne for eksempel vært variabler som søvn og kosthold som kan ha mye å si for prestasjon og aktivitet.

5. Regresjonsanalyse

Vi forventer å se at studentene med høyere aktivitetsnivå vil gjøre det bedre på skolen. I denne delen av oppgaven vil vi utføre forskjellige analyser for å finne ut om våre antagelser stemmer og forsøke å trekke en konklusjon til problemstillingen vår.

5.1. Multippel regresjonsanalyse

Vi har plukket ut variablene vi er interessert i å se på og vil nå utføre en lineær multippel regresjonsanalyse ved hjelp av minste kvadraters metode (OLS). Regresjonsanalysen gjennomføres ved hjelp statistikkverktøyet Stata. Den avhengige variabelen er gjennomsnittlig karakter i semesteret (GPA). Ettersom karakterene er målt i poeng er det fornuftig å bruke de faktiske verdiene. De estimerte koeffisientene i regresjonen forteller hvor mye karakterene endres dersom den respektive variabelen endres med en enhet, og de resterende variablene holdes uendret. Formel 5 viser ligningen for modellen.

$$\widehat{GPA} = \beta_0 + \beta_1 steps + \beta_2 female + \beta_3 minutes + \beta_4 lifescore + \beta_5 fatburn + \beta_6 peak + \beta_7 cardio + u$$

Formel 5: Regresjonsanalyse

Variabler	Koeffisienter og t-verdi
Steps	0.0001*** (9.06)
Female	0.3290*** (5.15)
Minutes	-0.0291** (-2.73)
Lifescore	-0.0028 (-1.80)
Fatburn	0.0002 (0.44)
Peak	-0.0019 (-0.87)
Cardio	0.0019 (0.90)
Konstant	2.4850*** (13.00)
N	578
adj. R^2	0.205

t-statistikk I parenteser
*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$

Tabell 3: Regresjonsanalyse

Variablenes signifikansnivå er vist med stjerner ved siden av tallene i tabellen for å vise om variablene er statistisk signifikant i utvalget. T-verdien under regresjonsanalysen ga oss denne informasjonen. Jo flere stjerner en variabel har, desto større er sikkerheten for at variabelen ikke er et resultat av tilfeldighet. Nederst i tabellen har vi lagt til modellens «adjusted R-squared». Denne forteller oss hvor mye av variansen i testresultatene som kan forklares av utvalget av eksogene variabler. Vi kan se at de tre variablene for intensitetssoner, samt variabelen for egenvurdert livsstil ikke er signifikante variabler for GPA resultatene. Vi kan dermed ikke si at

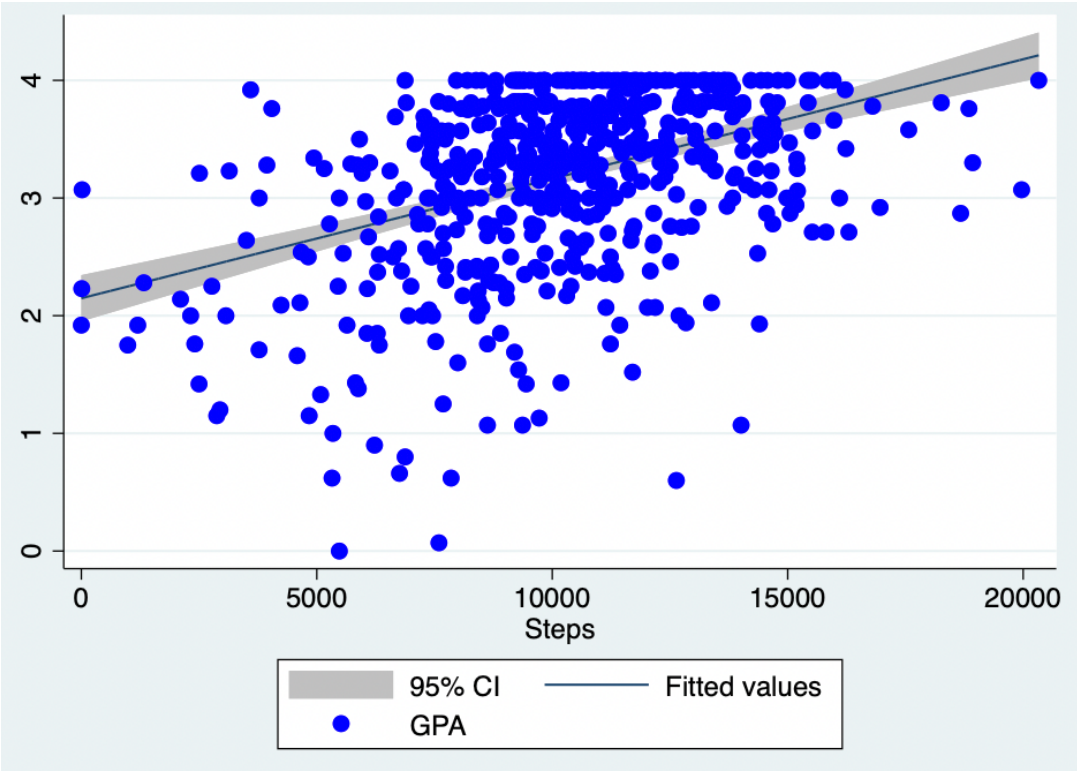
disse har noen effekt på studentenes prestasjoner og vi vil derfor videre fokusere på de andre variablene.

$$gpa_i = 2{,}4850 + 0{,}0001Steps + 0{,}3285Female - 0{,}0291Minutes - 0{,}0028Lifescore + 0{,}0002Fatburn - 0{,}0019Peak + 0{,}0019Cardio$$

Formel 6: Estimert regresjonsanalyse

Konstanten i regresjonsmodellen er 2,485. Dette vil si at GPA forventes å være lik 2,485 dersom alle andre variabler er satt til null. Ettersom *Female* er en dummy variabel, gir tolkningen av konstantleddet forventet verdi for menn på *Female*=0. Tolkningen av konstantleddet gir lite mening fordi det måler forventet verdi når alle andre variabler er lik null. Ettersom ingen av variablene i vår modell har null i snitt som referanseområde, er konstantleddet lite innsiktsfullt.

Ifølge vår modell kan en økning på 1000 gjennomsnittlig daglig skritt, *Steps,* øke den avhengige variabelen GPA med 0,1 for menn. Dette stemmer med vår antagelse om at mer hverdagsaktivitet fører til økt skoleprestasjon. Når variablene er omregnet til standardiserte betakoeffisienter i Modell 6 ser vi at det er variabelen *Steps* har størst innvirkning på GPA-resultatet til elevene, og vi kan derfor konkludere med at dette er den viktigste av variablene vi har inkludert i vår analyse. Det vil si at dersom studentene gjennomfører flere skritt, vil det ha positive effekter på karakterene. Figur 1 viser at studenter med høyere gjennomsnittlig daglig skritt tenderer mot høyere karakterer.



Figur 1: GPA vs. Steps

Om alt annet holdes likt, vil variablene ha følgende påvirkninger på semesterkarakterene: Ifølge regresjonsmodellen vil GPA øke med 0,0291 for menn dersom studentene løper løpetesten ett minutt raskere, altså reduserer tiden med ett minutt. Dette fanges opp i variabelen *Minutes.* Effekten i *Female* tar med forskjellen i effekten mellom kvinner og menn i alle de andre variablene. Ettersom de andre variablene i regresjonsmodellen, *Cardio, Fatburn, Peak* og *Lifescore,* ikke er signifikante kan vi ikke forkaste at disse har null effekt. Justert R^2 viser at 20,47% av variansen i GPA kan forklares av de uavhengige variablene i denne modellen. Våre analyser inkluderer kun data fra individnivå, og fanger ikke opp andre viktige faktorer, som sosioøkonomisk bakgrunn, karakterer fra videregående skole, arbeidsvaner eller andre faktorer. Derfor kan en forklaringsgrad på like over 20% sies å være svært høyt.

Modellen vår estimerer at økt antall skritt vil gi studentene bedre karakterer. På samme måte vil en reduksjon i minutter brukt på løpetesten gi studentene økt GPA. En reduksjon i *Lifescore*, som indikerer bedre helse, var ikke signifikant i vår modell. Regresjonsmodellen estimerer altså i hovedsak en positiv sammenheng mellom fysisk aktivitet og studentprestasjon.

5.2 Korrelasjon

Ettersom variablene består av mange av de samme komponentene er det her sannsynlig at det er sterk korrelasjon mellom variablene. Et potensielt problem er altså multikollinearitet. Multikollinearitet kan øke variansen til koeffisientestimatene og påvirke den statistiske signifikansen av de uavhengige variablene. Koeffisientestimatene kan da bli ustabile og følsomme for mindre endringer i modellen. Figuren nedenfor viser korrelasjonskoeffisientene for hvert par variabler. Korrelasjonene er i tabell 4 illustrert etter styrke med bakgrunnsfarge der ubetydelige sammenhenger er markert med svakest farge.

	GPA	**Cardio**	**Fatburn**	**Female**	**Lifescore**	**Minutes**	**Peak**	**Steps**
GPA	1,000							
Cardio	0,067	1,000						
Fatburn	0,141	0,442	1,000					
Female	0,146	0,034	0,227	1,000				
Lifescore	-0,117	-0,050	0,087	0,046	1,000			
Minutes	-0,146	0,012	0,205	0,413	0,194	1,000		
Peak	-0,012	0,321	0,006	-0,099	-0,083	-0,097	1,000	
Steps	0,407	0,054	0,231	-0,064	-0,101	-0,291	0,036	1,000

Tabell 4: Korrelasjonsmatrise

Variabelen *Minutes* inneholder tiden studentene brukte på Field-testen, en indikasjon på studentenes kardiovaskulære helse. Denne var negativt korrelert med GPA. Det er også en signifikant negativ korrelasjon mellom *Minutes* og *Steps*. Dette korresponderer med lavere (raskere) 1-mile test-tider for studenter som jevnlig tok flere skritt, også kontrollert for kjønn. Dette viser at skritt som en variabel, er en viktig faktor for vår analyse av fysisk aktivitet. Vi fant signifikant positiv korrelasjon (med korrelasjonskoeffisient >0,3) mellom variablene *Peak* og *Cardio*, *Fatburn* og *Cardio*, og *Minutes* og *Female*. Det var sterk korrelasjon mellom GPA og *Steps*. Dersom vi ser nærmere på statistikken for enkeltvariable lineære modeller finner vi at hypotesen vår stemmer. Det er en signifikant korrelasjon mellom sunne og fysisk disiplinerte

livsstilsvariabler *(Steps, Minutes, Fatburn, Lifescore*) og akademisk prestasjon. Korrelasjonen mellom *Minutes* og GPA er liten, men signifikant. Dette indikerer at studenter med lavere (raskere) test-tider på 1-mile testen, et indirekte mål på kardiovaskulær helse, tenderer mot høyere karakterer. Ved å inkludere kjønn i vår lineære regresjonsmodell kan vi observere at alle variablene som tidligere hadde signifikant korrelasjon med GPA, fortsatt var signifikante faktorer. Størrelsen på alle korrelasjonskoeffisientene var under 0,5, noe som tyder på at variablene i analysen har forskjellig informasjon.

5.3 Multivariat regresjonsanalyse

Med så mange faktorer med forskjellige informasjon, kan generelle multiple lineære modeller være nyttig for å få innsikt i de kumulative effektene av forskjellige variabler i kombinasjon med hverandre. Den første modellen vi ser på, er med variablene som har størst effekt på den avhengige variabelen GPA. I valg av modell to valgte vi å utelate variabelen *Steps* og se på en lignende modell der vi erstatter variabelen *Steps* med variabelen *Fatburn*. Den siste modellen vi ser på er en ubegrenset modell som inneholder alle variablene. Vi vil også i analysen se på hvor mye justert R^2 forandrer seg ved å legge til variabler. Dette gjør vi for å se om sammenhengen mellom variabler øker forklaringskraften til modellen. Slik kan vi se om valget av variabler er hensiktsmessige for vår problemstilling.

Vi estimerer følgende modeller:

(1) $gpa_i = \beta_0 + \beta_1 steps + \beta_2 female + \beta_3 minutes + \beta_4 lifescore + u_i$

(2) $gpa_i = \beta_0 + \beta_1 female + \beta_2 minutes + \beta_3 lifescore + \beta_4 fatburn + u_i$

(3) $gpa_i = \beta_0 + \beta_1 steps + \beta_2 female + \beta_3 minutes + \beta_4 lifescore + \beta_5 fatburn + \beta_6 peak + \beta_7 cardio + u_i$

Formel 7: Regresjonsmodell 1,2 og 3

	(1)	(2)	(3)
Variabel	GPA	GPA	GPA
Steps	0.0000966***		0.0000945***
	(9.80)		(9.06)
Female	0.339***	0.337***	0.329***
	(5.39)	(4.96)	(5.15)
Minutes	-0.0274**	-0.0603***	-0.0291**
	(-2.63)	(-5.60)	(-2.73)
Lifescore	-0.00273	-0.00376*	-0.00280
	(-1.77)	(-2.27)	(-1.80)
Fatburn		0.00119***	0.000159
		(3.68)	(0.44)
Peak			-0.00186
			(-0.87)
Cardio			0.00185
			(0.90)
Konstant	2.469***	3.693***	2.485***
	(13.14)	(25.80)	(13.00)
N	578	578	578
Adj. R2	0.206	0.094	0.205
$\Delta R2$		-0,112	+0,111

t-statistikk I parenteser
*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$

Tabell 5: Multivariat regresjonsanalyse

Modell 1 er en regresjonsanalyse mellom variablene GPA, *Steps, Female, Minutes og Lifescore.* Denne modellen inkluderer variablene som har størst effekt på den avhengige variabelen GPA. Modellen inneholder variabelen *Steps,* som er den mest signifikante faktoren for å predikere akademisk prestasjon. Modellens forklaringskraft er på 20,6%. Dette betyr at 20,6% av variasjonen i GPA kan forklares av livsstilsvariablene S*teps, Female, Minutes* og *Lifescore* i modell 1. Dette bekrefter at disse livsstilsvariablene forklarer mye av hva som påvirker variasjonen i akademisk prestasjon.

I modell 2 valgte vi å fjerne *Steps* som en variabel for å se på modeller både med og uten *Steps*. Dette gjorde vi for å se om gjennomsnittlig daglig skritt er et pålitelig mål på fysisk aktivitet og at det ikke er signifikant påvirket av studenter som kunstig produserte skritt mens de satt stille. Ettersom *Fatburn* er et annet mål på fysisk aktivitet, byttet vi ut variabelen *Steps* med *Fatburn* i denne modellen. Vi kan se at de negative koeffisientene for *Minutes* og *Lifescore* var større i denne modellen sammenlignet med modell 1, men det var også konstanten. Modell 2 sin forklaringskraft er på 0,094. Dette tilsier at modellen forklarer 9,4% av variasjonen i GPA. Dette tyder på at det er andre faktorer som har effekt på resultatet, men som ikke tas opp i modellen. Endringen i forklaringskraft fra modell 1 til 2 er negativ og forteller oss at ved å utelate *Steps* fra modellen, forklares variasjonen i prestasjon dårligere enn ved å inkludere denne variabelen. Dette gir oss økt grunnlag for å si at daglig skritt er et gyldig mål på fysisk aktivitet.

I modell 3 har vi tatt med variablene *Peak* og *Cardio,* og har dermed inkludert alle variabler vi har brukt i tidligere analyser. Disse variablene er ikke signifikante, og er derfor ikke hensiktsmessig å ta med som forklaring på endring i studentenes skoleprestasjoner. Vi kan altså anta at antall minutter aktivitet i intensitetssonene *Cardio* og *Peak* har liten effekt på hvor bra studentene presterer på skolen. Modellen som inkluderer alle variabler har en forklaringskraft på 20,5%. Dette vil si at variablene våre forklarer 20,5% av variasjonen i den avhengige variabelen GPA. Vi kan derfor se at økt aktivitetsnivå faktisk vil påvirke studentenes prestasjoner på skolen. Ved å inkludere flere variabler kan vi kontrollere for andre variabler, og med det fange opp renere effekter av interessevariabelen.

For å finne ut hvilken variabel i vår endelige modell som har størst effekt kan vi se på de standardiserte betakoeffisientene. De standardiserte betakoeffisientene er enkel måte å finne ut hvilke av koeffisientene som har størst innvirkning på *Y*. Tolkningen av standardiserte koeffisienter er "Når *Steps* øker med ett standardavvik, så øker GPA med 0,3801 standardavvik, kontrollert for de andre variablene i modellen." Dette gir lite mening, men det gjør at vi kan rangere koeffisientene etter absoluttverdi.

Variabler	Koeffisient	Standardisert betakoeffisient
Steps	0,0001 (***)	0,373
Female	0,3285 (***)	0,213
Minutes	-0,0291 (**)	-0,121
Life Score	-0,0028	-0,069
Cardio	0,0018	0,040
Peak	-0,0019	-0,035
Fatburn	0,0002	0,020
Konstant	2,4850	

Tabell 6: Standardiserte betakoeffisienter i Modell 3

5.4 Tester og robusthet

Vi vil nå benytte en F-test for å teste alle variablene i modellen samtidig. Med dette ønsker vi å teste om alle variablene vi har tilført i modellen utenom *Steps*, sammen har en signifikant effekt på GPA. Vi har da en restriktiv modell uten variablene og en ikke-restriktiv modell med variablene. Hypotesetesten blir slik:

$$H_0: \beta_2 = \beta_3 = \beta_4 = \beta_5 = \beta_6 = \beta_7 = 0$$
$$H_A: \beta_i \neq 0$$

Vi tester i Stata og får følgende F-statistikk:

$$F - stat = \frac{\frac{(R_{ur}^2 - R_r^2)}{q}}{\frac{1 - R_{ur}^2}{n - k - 1}} = 5{,}93$$

Kritisk verdi for et 5% signifikansnivå i en F-fordeling: $c_{6,578-7-1} = 2{,}0986 < 5{,}93$
Vi ser at testobservatoren overstiger kritisk verdi og kan dermed forkaste nullhypotesen om at variablene som er inkludert i modellen ikke har noen effekt på studentprestasjonene. P-verdien til testen er også 0,000 som betyr at sannsynligheten er null for å observere en F-statistikk på 5,93 dersom variablene i modellen vår er null. Variablene i modellen har altså samlet en effekt på studentprestasjonene.

Vi kan også benytte en F-test for å teste om de ikke-signifikante variablene *Cardio, Peak og Fatburn* samlet har noen effekt på GPA. Dette gjør vi for å ta hensyn til korrelasjonsproblemet som kan oppstå når variablene har sterk korrelasjon. Korrelasjon kan forklare lav signifikans når alle variablene er med. Denne testen kan forklare om grunnen til at disse variablene har lavt signifikansnivå er fordi de korrelerer med hverandre og dermed øker variansen til koeffisientestimatene.

$$H_0: \beta_5 = \beta_6 = \beta_7 = 0$$
$$H_A: \beta_i \neq 0$$

Vi tester i Stata og får at $F_{3,570} = 0{,}70$. Testens p-verdi er på 0,553 noe som tilsier at sannsynligheten for å observere en F-statistikk så ekstrem som 0,7 når nullhypotesen er sann er omtrent 55%. Det er altså svært sannsynlig å observere en slik statistikk når alle variablene er 0. Vi kan ikke forkaste nullhypotesen og konkluderer med at variablene for intensitetssonene ikke påvirker modellen vår, altså har koeffisienter lik 0. Det er med andre ord ikke korrelasjonen som gir variablene de lave signifikansnivåene.

OLS har forutsetninger som gir sterke restriksjoner på modellen. En forutsetning for en forventningsrett OLS-estimator er homoskedastisitet. For å teste kvaliteten på modellen kan vi teste for heteroskedastisitet. Her vil vi bruke en Breusch-Pagan test (Thomas, 2005, s. 505). En av

forutsetningene for OLS er som nevnt at det er konstant varians i restleddet for alle i. $Var\ (u_i) = \sigma^2$. Dersom modellen har heteroskedastiske støyledd vil estimatorens standardavvik bli feil (Thomas, 2005, s. 480). For å kunne si hva tusen skritt ekstra har å si for studentenes prestasjon må det være konstant varians på restleddene. Lik varians er også avgjørende for å kunne utføre hypotesetester. Testen går ut på å undersøke om variansen i støyleddene fra lineær regresjon er betinget av verdiene i de uavhengige variablene. Vi tester Modell 3 for heteroskedastisitet ved hjelp av Stata.

H_0: konstant varians, H_A: ikke konstant varians.

Vi tester en kji-kvadratfordelt nullhypotese om at variansen er konstant og vi har homoskedastiske restledd. Vi får en kji-kvadrat-verdi på 14,56 og en p-verdi på 0,0001. Vi må altså forkaste nullhypotesen som tilsier at modellen vår tester positivt for heteroskedasitet. Dette kan i utgangspunktet ugyldiggjøre resultatene vi har funnet ved hjelp av OLS. En måte å håndtere problemet med heteroskedastisitet på er å estimere regresjonene for robuste standardfeil. Vi gjør dette i Stata og finner at endringene i koeffisientene er uendret og endringen i t-verdiene er relativt små. Dette kan vi se i Tabell 7.

	T-verdi	T-verdi r.	Endring	Koeffisient	Koeffisient r.	endring
Steps	9.06	8.61	0,45	0.0001	0.0001	0
Peak	-0.87	-1.48	0,61	-0.0019	-0.0019	0
Cardio	0.90	0.93	0,03	0.0019	0.0019	0
Fatburn	0.44	0.45	0,01	0.0002	0.0002	0
Lifescore	-1.80	-1.60	0,20	-0.0028	-0.0028	0
Minutes	-2.73	-2.40	0,33	-0.0291	-0.0291	0
Female	5.15	5.23	0,08	0.3290	0.3290	0

Tabell 7: Koeffisient og t-verdi med og uten robust verdi

Modellene fra analysen kan altså beholdes ettersom de ikke har tilstrekkelig grad av heteroskedasitet til at de har alvorlige utslag på standardfeilene. På dette grunnlaget kan vi se bort fra små endringer og beholde modellen.

6. Diskusjon og begrensninger

Resultatene fra våre statistiske analyser er i samsvar med funn fra tidligere studier som har tatt for seg sammenhengene mellom fysisk aktivitet og akademisk prestasjon for barn i skolealder. (Tomporowski, Lambourne, & Okumuraa, 2011) (Booth, et al., 2014) (Ardoy, et al., 2014) (Shin & So, 2012). Vi ser, i likhet med tidligere studier (Stephens, Dong, & Durning, 2015), at fysisk aktivitet og bedre kardiovaskulær kondisjon har positiv korrelasjon med akademisk prestasjon for universitetsstudenter, basert på vårt datasett. Det var interessant å se at aktivitet og bevegelse målt i skritt og antall minutter i intensitetssonen *Fatburn* hadde større korrelasjon med forbedret GPA enn fysisk aktivitet i intensitetssonene *Cardio* og *Peak.* Våre statistiske undersøkelser viste at økt aktivitetsnivå har en signifikant positiv effekt på studentenes prestasjoner på skolen. På bakgrunn av dette kan fremtidige studier med fordel se nærmere på effektene av hvilken type trening studentene utfører, og hva dette betyr for deres akademiske prestasjon. Dette vil gi svar på om det finnes en spesifikk form for aktivitet som bør anbefales, eller om fysisk aktivitet kan anbefales på et generelt grunnlag.

Ytterligere forskning er hensiktsmessig for å bekrefte resultatene vi har funnet. Studiene bør inkludere kontrollgrupper, og i tillegg redegjøre for eventuelle forstyrrende faktorer som for eksempel ansvarsbevissthet. Det blir ikke tatt høyde for ansvarsbevissthet i våre analyser, men det er naturlig å anta at studentene som jobber hardt for å oppnå høy GPA vil være de samme som jobber hardt for å oppnå maksimale poeng relatert til antall daglige skritt og resultattid på løpstesten. Disse forbindelsene kan være interessante å undersøke videre.

Det er viktig å nevne at det er mange faktorer utenom fysisk aktivitet som påvirker studentenes gjennomsnittlige karakter og som ikke blir tatt i betraktning i våre statistiske analyser. For å kunne konkludere på effektene av variablene, er vi nødt til å være sikre på at forholdet vi ser på faktisk fanger opp effektene av aktivitet på studentprestasjonene og ikke forstyrres av årsaks-effekten fra andre variabler. Studentenes sosioøkonomiske bakgrunn, økonomi, tid, og kosthold kan spille inn på studentenes valg av fysisk aktivitet. Søvn, stress, mestring, og motivasjon er variabler som kan ha innvirkning på studenters akademiske prestasjoner. Dette er variabler som også bør inkluderes i videre studier. Dersom livsstilsvariablene som inkluderes i variabelen *Lifescore* hadde vært

atskilt, kunne det vært relevant å se på den isolerte effekten av disse på GPA. *Lifescore*-variabelen i vår modell kan likevel fange opp noen av disse effektene. I modellene våre vil andre livsstilsfaktorer som ikke inkluderes fanges opp i restleddet, som ved bruk av OLS vil estimeres til null. For å kontrollere virkningen av disse faktorene trenger man et mer omfattende datamateriale som vi i denne oppgaven ikke har tilgang på.

Det kan tenkes at skritt har avtakende marginalutbytte. En økning i aktivitet fra sedat (stillesittende) til litt mer aktiv livsstil kan tenkes å gi en større økning i prestasjon, enn en økning fra aktiv til veldig aktiv. Dermed kunne det vært interessant å dele opp hverdagsaktiviteten i bolker og se på ved hvilke nivåer økningen av prestasjonen har størst betydning for prestasjonsutbytte.

Det er vanskelig å argumentere for at våre funn gjelder for en større populasjon. Datasettet er gjort på en relativt liten utvalgsstørrelse (N=578). Ved undersøkelser med større utvalg kan feilmarginen reduseres. Undersøkelsen er utført på en gruppe førsteårsstudenter som tok ett fag ved universitetet. Utvalget var derfor svært begrenset, og resultatene kan derfor kun generaliseres for studenter som studerer første året i USA. Et sannsynlighetsutvalg på nasjonalt nivå ville gitt en mer pålitelig prediksjon.

7. Konklusjon

I denne analysen har vi undersøkt om og hvordan effektene fysisk aktivitet på akademisk prestasjon gjør seg gjeldende for studenter på universitetsnivå. Hensikten har vært å se om det faktisk eksisterer et prestasjonsfremmende middel som alle kan ta i bruk. Resultatene fra de estimerte modellene tilsier at økt aktivitetsnivå har en positiv effekt på studentprestasjoner. De fleste variablene talte for at hypotesen vår om at økt aktivitet bedret studentenes skoleprestasjon stemte. For en av variablene våre, antall minutter i *Peak*-sonen, var det motsatt. Denne effekten var dog ikke signifikant. Gjennomsnittlig daglig skritt er en signifikant faktor som spiller inn når man ser på prestasjon på studienivå. Dette støtter godt opp tesen om at vanlig mosjon av en viss varighet i løpet av uken er den beste treningsformen, og kanskje den enkleste for mange å kunne opprettholde over tid. På bakgrunn av våre analyser kan vi altså konkludere med at økt hverdagsaktivitet har en prestasjonsfremmende effekt. Vi har ikke grunnlag for å si at aktivitet med høyere intensitet er bedre for skoleprestasjon enn aktivitet på lavere intensitet.

Våre funn viser at akademiske prestasjoner for universitetsstudenter i første semester, målt ved semesterkarakterer (GPA), er betydelig korrelert med målinger av fysisk aktivitet, fysisk form og en sunn livsstil. Analysen vår utvider forståelsen av de positive effektene av å føre en sunn og fysisk aktiv livsstil ved å forklare akademisk prestasjon. Modellene viser at prestasjon påvirkes mest av hverdagslig aktivitet som skritt og at det er forskjeller mellom kjønn. Vi vil derfor konkludere med at økt aktivitet fører til økt akademisk prestasjon blant universitetsstudenter, der hverdagsaktivitet målt i antall daglig skritt har størst effekt.

Bibliografi

Ardoy, D. N., Fernández-Rodríguez, J. M., Jiménez-Pavón, D., Castillo, R., Ruiz, J. R., & Ortega, F. B. (2014). *A Physical Education trial improves adolescents' cognitive performance and academic achievement: the EDUFIT study.* Scand J Med Sci Sports.

Bahr, R. (2008). *Aktivitetshåndboken - Fysisk aktivitet i forebygging og behandling.* Helsedirektoratet.

Booth, J. N., Leary, S. D., Joinson, C., Ness, A. R., Tomporowski, P. D., Boyle, J. M., & Reilly, J. J. (2014). *Associations between objectively measured physical activity and academic attainment in adolescents from a UK cohort.* Br J Sports Med.

Broaddus, A., Jaquis, B., Jones, C., Jost, S., Lang, A., Li, A., . . . Spear, E. (2018, oktober 17). *Dataset: Fitbits, field-tests, and grades. The effects of a healthy and physically active lifestyle on the academic performance of first year college students.* Hentet fra https://doi.org/10.6084/m9.figshare.7218497.v1.

Cappelen, A. W., Charness, G., Ekström, M., Gneezy, U., & Tungodden, B. (2017). *Exercise Improves Academic Performance.* Bergen: Norges Handelshøyskole.

Carlson, S. A., Fulton, J. E., Lee, S. M., Maynard, L. M., Brown, D. R., Kohl III, H. W., & Dietz, W. H. (2008). *Physical Education and Academic Achievement in Elementary School: Data from the Early Childhood Longitudinal Study.* American Journal of Public Health.

DeSilver, D. (2015). *U.S. Students' Academic Achievement Still Lags that of their Peers in many other Countries.* Pew Research Center FactTank.

FHI. (2017). *Fysisk aktivitet i Noreg.* Folkehelserapporten.

Gonzalez, E. C., Hernandez, E. C., Coltrane, A. K., & Mancera, J. M. (2014). *The Correlation between Physical Activity and Grade Point Average for Health Science Graduate Students.* Occupation, Participation and Health,.

Helsedirektoratet. (2019, april 29). *Fysisk aktivitet for barn, unge, voksne, eldre og gravide .* Hentet fra Helsedirektoratet: https://www.helsedirektoratet.no/faglige-rad/fysisk-aktivitet-for-barn-unge-voksne-eldre-og-gravide

Johnson, T. G. (2016). *Communities of Practice: A Possible Professional-Development Model for Physical Activity Leadership.* Journal of Physical Education, Recreation and Dance.

Kari, J. T., Pehkonen, J., Hutri-Kähönen, N., Raitakari, O. T., & Tammelin, T. H. (2017). *Longitudinal Associations between Physical Activity and Educational Outcomes.* Medicine and Science in Sports and Exercise.

Khan, M. A. (2017). *Association of Internet Supported Physical Activity with Academic Performance of Medical Students.* Annals of King Edward Medical University.

LeBlanc, M. M. (2014). *Adiposity and physical activity are not related to academic achievement in school-aged children.* J Dev Behav Pediatr.

Lynne, M. F., Jasmina, G., Eric, C. S., Chance, P., Allison, M. E., Ju, Y. Y., . . . Linda, C. L. (2018). *Accuracy of Fitbit Devices: Systematic Review and Narrative Syntheses of Quantitative Data.* JMIR Mhealth Unhealth.

Martínez-Gómez, D., Veiga, O. L., Gómez-Martínez, S., Zapatera, B., Martínez-Hernández, D., Calle, M. E., & Marcos, A. (2012). *Gender-specific influence of health behaviors on academic performance in Spanish adolescents: the AFINOS study.* Nutr Hosp.

Reutz, L. T. (2018, september 18). *Forskning.no.* Hentet fra Helse: https://forskning.no/hoyskolen-kristiania-trening-partner/hadde-fysisk-aktivitet-vaert-en-pille-ville-alle-tatt-den/1239833

Shin, Y. S., & So, W. Y. (2012). *Association between Physical Inactivity and Academic Record in Korean Adolescents.* Iran J Public Health.

Skaalvik, E., & Skaalvik, S. (2013). *Skolen som læringsarena. Selvoppfatning, motivasjon og læring.* Oslo: Universitetsforlaget.

So, W. (2012). *Association between physical activity and academic performance in Korean adolescent students.* BMC Public Health.

Stephens, M. B., Dong, T., & Durning, S. J. (2015). *Physical Fitness and Academic Performance: A Pilot Investigation in USU Medical Students.* Military Medicine.

Thea, J. M., Manon, L. D., Siska, R. S., Wim, K., Cees, v. d., & Martijn, G. (2015). *Reliability and validity of ten consumer activity trackers.* BMC Sports Science, Medicine and Rehabilitation.

Thomas, R. (2005). *Using statistics in economics.*

Tomporowski, P., Lambourne, K., & Okumuraa, M. (2011). *Physical activity interventions and children's mental function: An introduction and overview.* Prev Med.

Trudeau, F., & Shephard, R. J. (2008). *Physical Education, School Physical Activity, School Sports and Academic Performance.* International Journal of Behavioral Nutrition and Physical Activity.

Tudor-Locke, C., Craig, C., Brown, W. J., Clemes, S., De Cocker, K., Giles-Corti, B., . . . Blair, S. (2011). How many steps/day are enough? For adults.

Wattanapisit, A., & Thanamee, S. (2017). *Evidence behind 10,000 steps walking.* Journal of Health and Research.

World Health Organization. (2020). *WHO guidelines on physical activity and sedentary behaviour.* Geneva: World Health Organization.

Xu, Q., & Sansgiry, S. S. (2018). Association between Physical Activity and Grade Point Average among a Cohort of Pharmacy Students in Didactic Years. *Currents in Pharmacy Teaching and Learning, 10*(3).

NTNU

Kunnskap for en bedre verden