

# Er det høyde som bestemmer inntekt?

Kine Maakestad

Susann Sivertsen

# Introduksjon

I denne oppgaven skal vi finne ut om høyde bestemmer inntekten til personer. Først skriver vi en liten kort litteraturgjennomgang, samt beskrivende statistikk. Deretter har vi analysert om høyde bestemmer inntekten. Til slutt konkluderer vi om dette stemmer.

## Kort litteraturgjennomgang

**The Effect of Physical Height on Workplace Success and Income: Preliminary Test of a Theoretical Model, av Judge and Cable (2004).**

Artikkelen ser på relasjonen mellom karriere suksess og fysisk høyde. Det baseres på en meta-analyse av litteraturen, hvor resultatene viser mot at høyde er relatert til hvordan andre ser deg, leder muligheter og opptreden. Artikkelen viser også til at høyde har en sterkere relasjon til karriere suksess for menn enn for kvinner, men denne forskjellen sees ikke på som signifikant. Ettersom det ikke finnes forskning på forholdet mellom høyde og inntekt, brukes det fire steg som tilsier at høyde er positivt for inntekt etter at det er tatt hensyn til kjønn, alder og vekt. Artikkelen tar for seg en omfattende analyse av relasjonen til høyde og karriere suksess, og resultatene tilsier at høye individer har et fortrinn i viktige aspekter av deres karrierer.

Artikkelen baseres på samfunnets inntrykk av at høye folk er mer suksessfulle. Dette går tilbake til dyreverden hvor høyde er en index for makt og styrke i kritiske situasjoner. Artikkelen viser til at høyde burde være relevant i arbeidsmiljø hvor overtalelse og makt er signifikant. Høyde påvirker ikke bare hvordan andre ser deg, men kan også påvirke hvordan du ser deg selv, og personers psykologiske tilpassning og selvtillit. Et resultat av de to faktorene vil igjen påvirke opptreden som videre påvirker karriere suksess (Judge and Cable 2004).

De fire studiene nevnt i artikkelen viser alle til det er en positiv korrelasjon mellom høyde og inntekt. Det konkluderes videre med at effekten til ens høyde ikke svekkes over tid, altså senere i ens karriere, men derimot øker litt. Videre i artikkelen diskuteres det igjen om effekten av høyde varierer basert på ulike typer yrker. Det generelle resultatet av dette var at sosiale yrker, med mye interaksjoner, avhenger mer av på utseende og dermed høyde enn mindre sosiale yrker (Judge and Cable 2004).

Denne artikkelen viser tydelig til at fysisk høyde påvirker karrierer og interaksjoner og at en videre forskning burde gjøres på dette. Ettersom det teoretisk er viktig for å skjønne hvorfor og hvordan det har effekt på suksess og inntekt. Og praktisk fordi at ledere lettere kan få en forståelse på om høyde påvirker suksess via hvordan andre ser deg, og ens egen selvtillit eller om det påvirker suksess helt uavhengig av andre faktorer (Judge and Cable 2004).

## Beskrivende statistikk

Beskrivende statistikk av datasettet “heights” er:

- **heights** = er høyde i tommer
- **income** = årlig inntekt. Top 2% gjennomsnitt av verdien og gjennomsnittet er brukt til å erstatte alle verdiene i topp inntekten.
- **weight** = det er vekten i pounds
- **age** = alder i år, mellom 47 og 56
- **martial** = sivilstatus
- **sex** = kjønn
- **education** = hvor mange år utdannelse
- **afgt** = persentilscore på Forsvarets kvalifikasjonsprøve

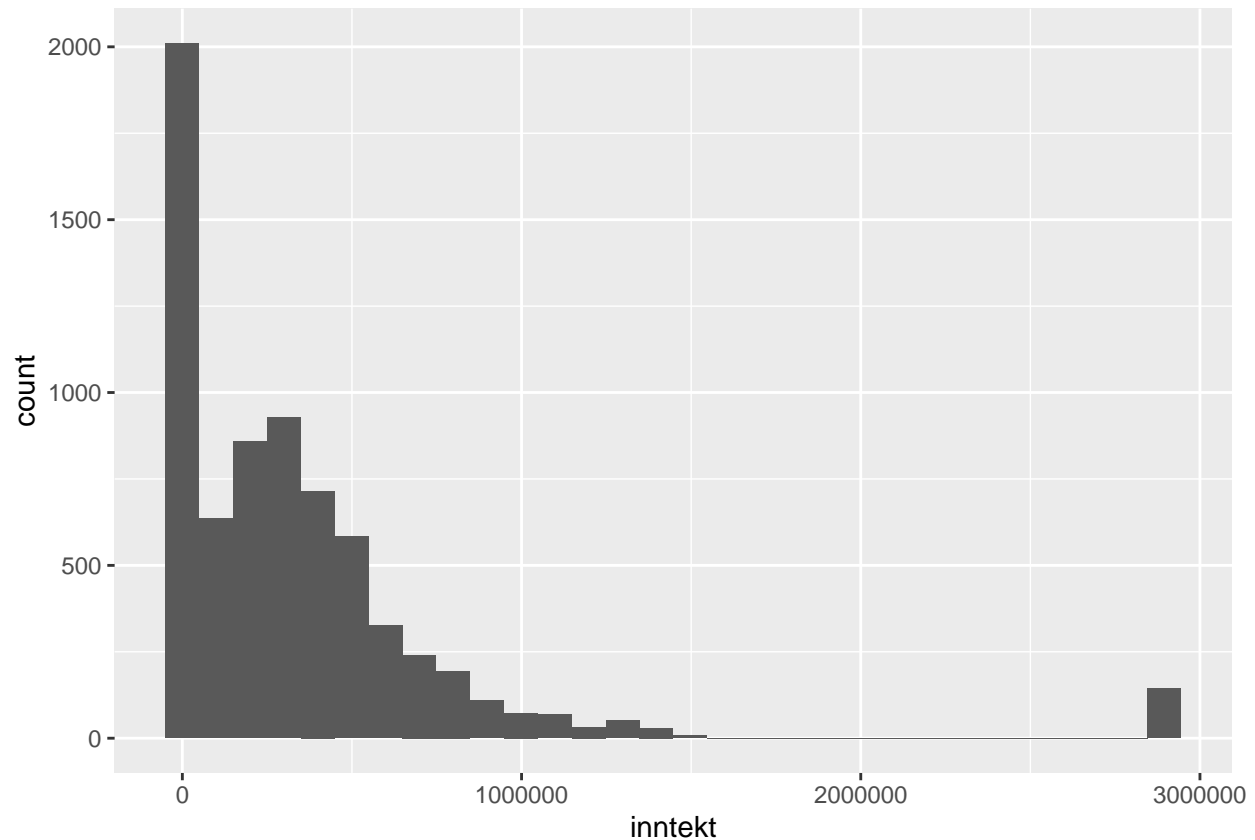
## Analyse:

For analyse delen lager vi først et histogram med variabelen inntekt. Vi har gjort om inntekt, høyde og vekt til metrisk standard. Som vil si at inntekten blir gjort om til norske kroner, høyde i cm og vekten i kg.

### EDA av datasett

Her har man gjort om variablene til metrisk standard. En har også lagt til tre nye variabler.

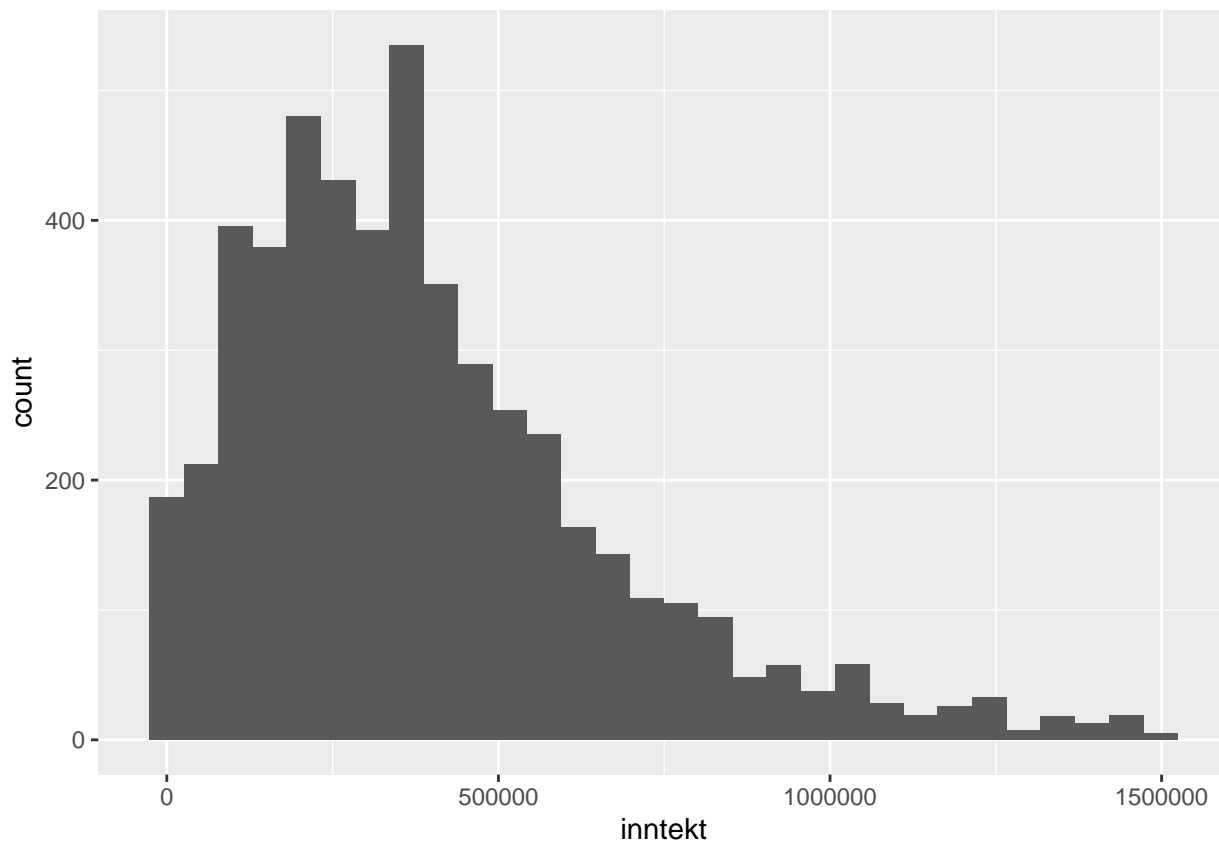
```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```



```
## geom_bar: na.rm = FALSE, orientation = NA
## stat_bin: binwidth = NULL, bins = 30, na.rm = FALSE, orientation = NA, pad = FALSE
## position_stack
```

I histogrammet ovenfor ser en at utliggerne ligger veldig langt til høyre. Grunnen for dette er at de har funnet gjennomsnittet av topp 2% inntekt.

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```



Som man kan se så er personer uten inntekt tatt med i dette datasettet.

```
## [1] 1740
```

Summen av personer som er tatt med i datasettet med uten inntekt er på 1740.

## Regresjonsanalyse

```
##
## Call:
## lm(formula = mod1, data = hoyde, subset = complete.cases(hoyde))
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -782810 -267359  -94513  123099 2699234
##
## Coefficients:
##              Estimate Std. Error t value      Pr(>|t|)
## (Intercept) -1361001.0    94430.0  -14.41 <0.0000000000000002 ***
## hoyde_cm      10047.9     552.8    18.18 <0.0000000000000002 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 467300 on 6643 degrees of freedom
## Multiple R-squared:  0.04737,    Adjusted R-squared:  0.04723
## F-statistic: 330.3 on 1 and 6643 DF,  p-value: < 0.00000000000000022
## [1] 377285.7
```

```
## [1] 256710.9
```

Man øker inntekten sin med 10047.9 kr per cm en øker i høyde.

```
##
## Call:
## lm(formula = mod2, data = hoyde, subset = complete.cases(hoyde))
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -843668 -263322  -92573   125798  2715000
##
## Coefficients:
##              Estimate Std. Error t value      Pr(>|t|)
## (Intercept) -1466873.6    96890.5  -15.139 < 0.0000000000000002 ***
## hoyde_cm      11430.3      624.3   18.308 < 0.0000000000000002 ***
## vekt_kg       -1518.4      320.5   -4.737    0.00000221 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 466600 on 6642 degrees of freedom
## Multiple R-squared:  0.05058,    Adjusted R-squared:  0.05029
## F-statistic: 176.9 on 2 and 6642 DF,  p-value: < 0.00000000000000022
## [1] 404280.3
## [1] 274708.7
```

Når høyden øker så går inntekten opp, mens når vekten økes går lønnen ned. Men en kombinasjon av disse gir økt inntekt.

```
##
## Call:
## lm(formula = mod3, data = hoyde, subset = complete.cases(hoyde))
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -886295 -261634  -93597   124905  2709981
##
## Coefficients:
##              Estimate Std. Error t value      Pr(>|t|)
## (Intercept) -2015890    447005  -4.510 0.0000066012 ***
## hoyde_cm      14669      2649   5.537 0.0000000319 ***
## vekt_kg       -4723      2567  -1.840    0.0658 .
## BMI           9224      7332   1.258    0.2084
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 466600 on 6641 degrees of freedom
## Multiple R-squared:  0.05081,    Adjusted R-squared:  0.05038
## F-statistic: 118.5 on 3 and 6641 DF,  p-value: < 0.00000000000000022
```

## Interaksjon med variabelen sex

```
##
## Call:
## lm(formula = mod4, data = hoyde)
```

	mod1	mod2	mod3
(Intercept)	-1361000.990 *** [-14.413]	-1466873.555 *** [-15.139]	-2015889.845 *** [-4.510]
hoyde_cm	10047.860 *** [18.175]	11430.259 *** [18.308]	14669.413 *** [5.537]
vekt_kg		-1518.381 *** [-4.737]	-4722.577 [-1.840]
BMI			9224.408 [1.258]
N	6645	6645	6645
R2	0.047	0.051	0.051
logLik	-96177.211	-96166.004	-96165.212
AIC	192360.423	192340.008	192340.424

Regresjonstabell 3: \*\*\* p < 0.001; \*\* p < 0.01; \* p < 0.05. T statistics in brackets.

```
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -849879 -244503  -90876   126516  2680060
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    -2310702.51   3080033.03  -0.750   0.453
## sexfemale       367086.85   3901078.89   0.094   0.925
## hoyde_cm       14261.90    17112.40   0.833   0.405
## vekt_kg         218.88     29727.93   0.007   0.994
## I(vekt_kg^2)    -29.66       71.43  -0.415   0.678
## BMI            17943.84    96811.11   0.185   0.853
## I(BMI^2)        -64.41       770.19  -0.084   0.933
## sexfemale:hoyde_cm -31.53    22594.40  -0.001   0.999
## sexfemale:vekt_kg -18366.14   40808.50  -0.450   0.653
## sexfemale:I(vekt_kg^2)  67.48     106.08   0.636   0.525
## sexfemale:BMI      25644.64   121508.62   0.211   0.833
## sexfemale:I(BMI^2)   -194.11     941.57  -0.206   0.837
##
## Residual standard error: 458400 on 6899 degrees of freedom
## (95 observations deleted due to missingness)
## Multiple R-squared:  0.06166,    Adjusted R-squared:  0.06016
## F-statistic: 41.21 on 11 and 6899 DF,  p-value: < 0.00000000000000022
```

Her har vi laget den endelige modellen uten 2% topp inntekt og 0 inntekt.

```
##
```

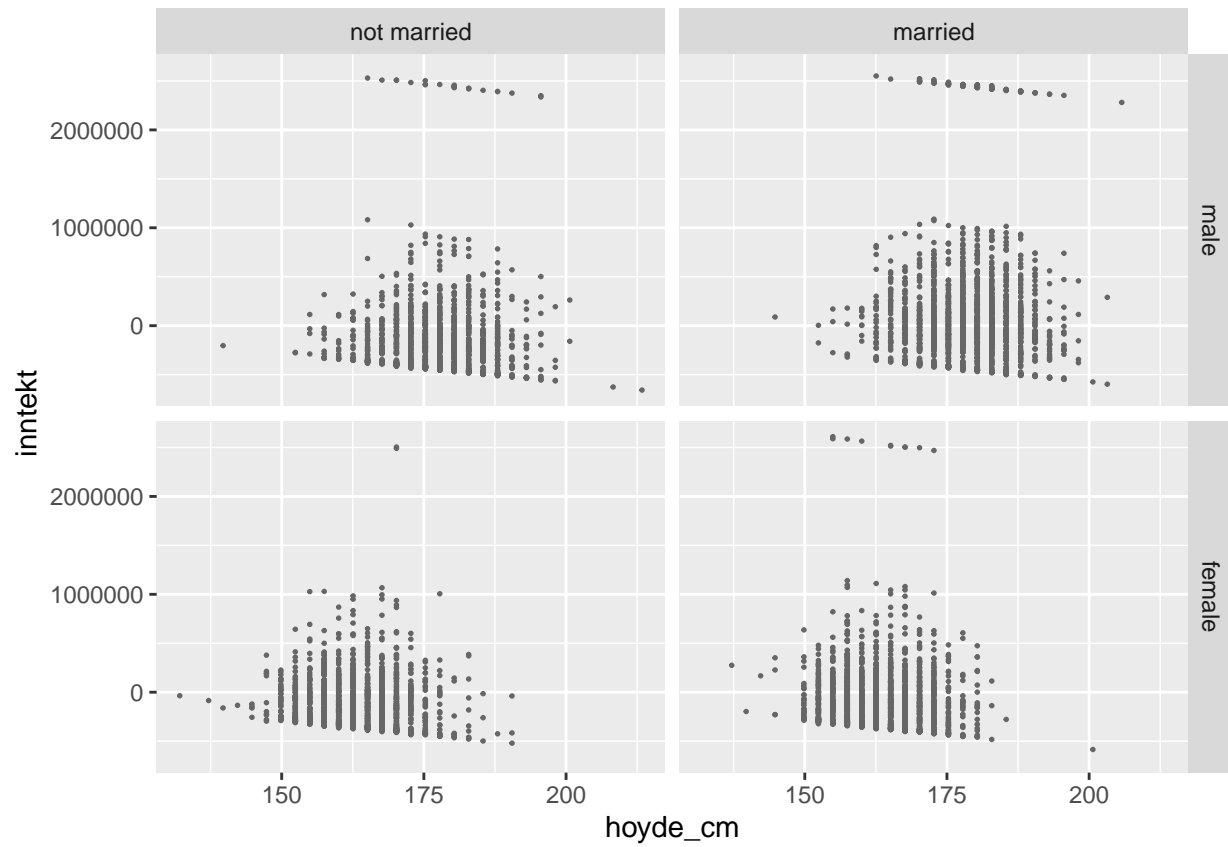
Res.Df	RSS	Df	Sum of Sq	F	Pr(>F)
6.90e+03	1.46e+15				
6.9e+03	1.45e+15	6	1.51e+13	12	1.81e-13

```
## Call:
## lm(formula = mod3, data = hoyde_begr, subset = complete.cases(hoyde))
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -537793 -189174  -56803  135653 1139990
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  -343501     328879  -1.044   0.2963
## hoyde_cm       4681       1940    2.413   0.0159 *
## vekt_kg        1393       1879    0.741   0.4586
## BMI          -6193       5430  -1.141   0.2541
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 274100 on 4800 degrees of freedom
## (1841 observations deleted due to missingness)
## Multiple R-squared:  0.05215,    Adjusted R-squared:  0.05156
## F-statistic: 88.03 on 3 and 4800 DF,  p-value: < 0.00000000000000022
```

## Residualer

ht	weight	age	marital	sex	education	afqt	inntekt	hoyde_cm	vekt_kg	BMI	married
60	155	53	married	female	13	6.84	1.6e+05	152	70.4	30.3	TRUE
70	156	51	married	female	10	49.4	2.95e+05	178	70.8	22.4	TRUE
65	195	52	married	male	16	99.4	8.84e+05	165	88.5	32.5	TRUE
63	197	54	married	female	14	44	3.37e+05	160	89.4	34.9	TRUE
66	190	49	married	male	14	59.7	6.32e+05	168	86.3	30.7	TRUE
68	200	49	divorced	female	18	98.8	8.59e+05	173	90.8	30.4	FALSE
74	225	48	married	male	16	82.3	0	188	102	28.9	TRUE
64	160	54	divorced	female	12	50.3	5.89e+05	163	72.6	27.5	FALSE
69	162	55	divorced	male	12	89.7	5.05e+05	175	73.5	23.9	FALSE
69	194	54	divorced	male	13	96	1.26e+06	175	88.1	28.7	FALSE

```
## Warning: Removed 95 rows containing missing values (geom_point).
```



## Konklusjon

Fra modellene kan vi konkludere med at høyde ikke har noe påvirkning på inntekt.



## Referanser

Judge, Timothy A., and Daniel M. Cable. 2004. "The Effect of Physical Height on Workplace Success and Income: Preliminary Test of a Theoretical Model." *Journal of Applied Psychology* 89 (3): 428–41. <https://doi.org/10.1037/0021-9010.89.3.428>.