

Er det høyde som bestemmer inntekt?

```
library(modelr)
library(ggplot2)
library(tinytex)
library(tidyverse)

## -- Attaching packages ----- tidyverse 1.3.1 --

## v tibble 3.1.3      v dplyr 1.0.7
## v tidyr 1.1.3      v stringr 1.4.0
## v readr 2.0.1     v forcats 0.5.1
## v purrr 0.3.4

## -- Conflicts ----- tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()     masks stats::lag()

library(ggpubr)
library(huxtable)

##
## Attaching package: 'huxtable'

## The following object is masked from 'package:ggpubr':
##
## font

## The following object is masked from 'package:dplyr':
##
## add_rownames

## The following object is masked from 'package:ggplot2':
##
## theme_grey

options(scipen = 999)
```

Introduksjon

Denne oppgaven skal avgjøre om høyde påvirker inntekt.

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

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.

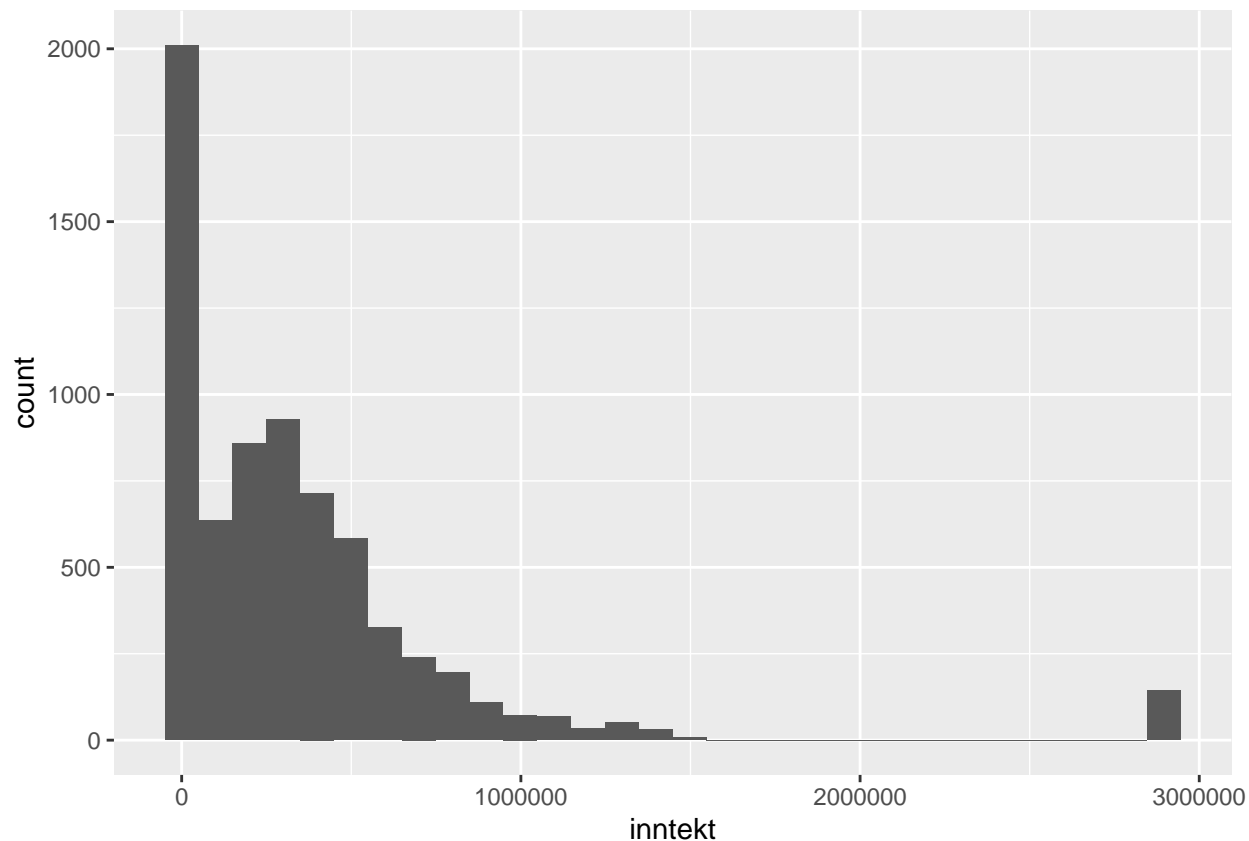
```
hoyde <- heights
```

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

```
hoyde <- hoyde %>%  
  mutate(inntekt = income * 8.42,  
         hoyde_cm = height * 2.54,  
         vekt_kg = weight * 0.454,  
         BMI = vekt_kg/(hoyde_cm/100)^2)
```

```
ggplot(data = hoyde, aes(x = inntekt)) +  
  geom_histogram()
```

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



```
geom_histogram(bins = 30)
```

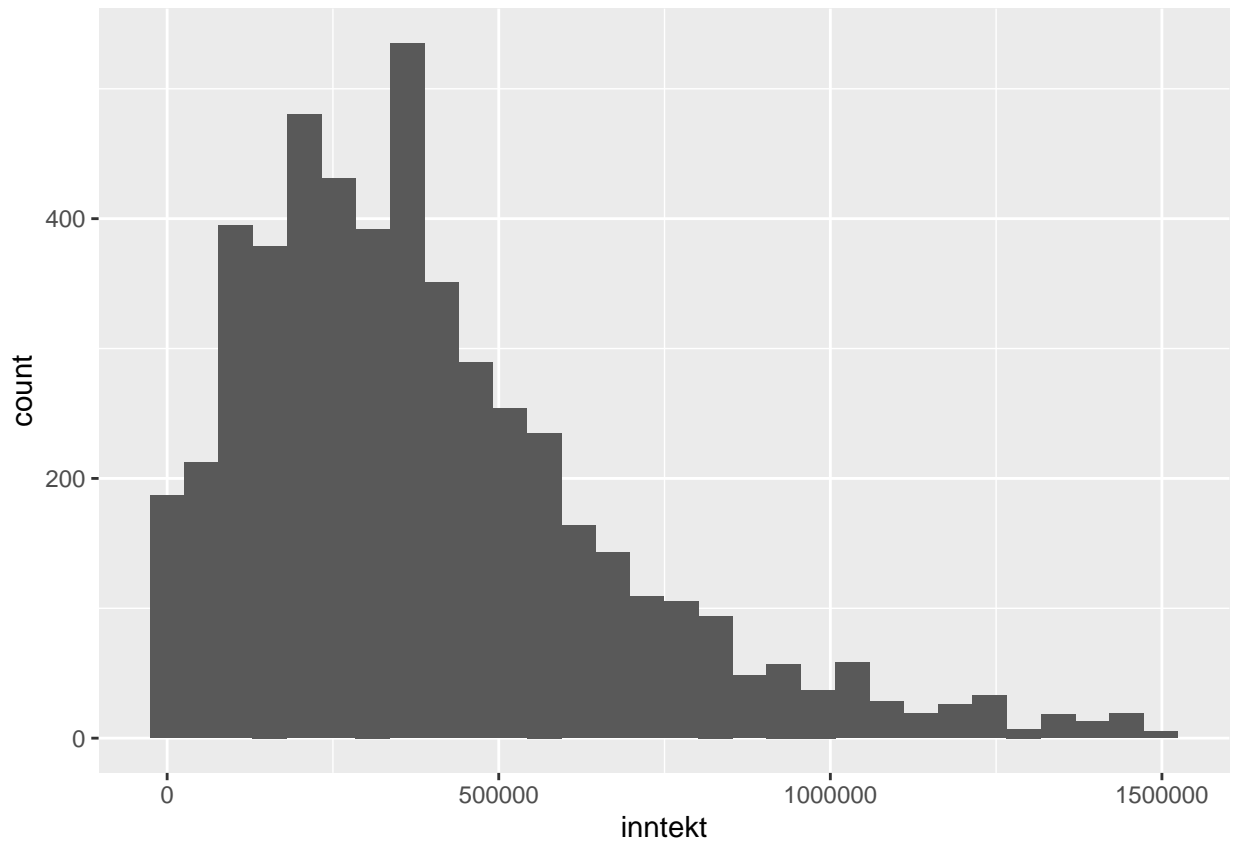
```
## geom_bar: na.rm = FALSE, orientation = NA
## stat_bin: binwidth = NULL, bins = 30, na.rm = FALSE, orientation = NA, pad = FALSE
## position_stack
```

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.

```
hoyde_begr <- hoyde %>%
  filter(inntekt < 1500000,
         inntekt > 1)
```

```
ggplot(data = hoyde_begr, aes(x = inntekt)) +
  geom_histogram()
```

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



Som man kan se så er personer uten inntekt tatt med i datasettet. Og summen er 1740 personer uten inntekt.

```
sum(hoyde$income == 0)
```

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

Regresjonsanalyse

```
mod1 <- "inntekt ~ hoyde_cm"
lm1 <- lm(mod1, data = hoyde, subset = complete.cases(hoyde))
```

```
summary(lm1)
```

```
##
## 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
```

```
-1361001.0 + (10047.9 * 173)
```

```
## [1] 377285.7
```

```
-1361001.0 + (10047.9 * 161)
```

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

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

```
mod2 <- "inntekt ~ hoyde_cm + vekt_kg"
lm2 <- lm(mod2, data = hoyde, subset = complete.cases(hoyde))
```

```
summary(lm2)
```

```
##
## 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
```

```
-1466873.6 + (11430.3 * 173) + (-1518.4 * 70)
```

```
## [1] 404280.3
```

```
-1466873.6 + (11430.3 * 161) + (-1518.4 * 65)
```

```
## [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.

```
mod3 <- "inntekt ~ hoyde_cm + vekt_kg + BMI"
lm3 <- lm(mod3, data = hoyde, subset = complete.cases(hoyde))
```

```
summary(lm3)
```

```
##
## 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
```

```
hoyde <- hoyde %>%
  mutate(
    married = factor(
      case_when(
        marital == 'married' ~ TRUE, TRUE ~ FALSE
      )
    )
  )
```

```
huxreg(
  list("mod1" = lm1, "mod2" = lm2, "mod3" = lm3),
  error_format = "[{statistic}]",
  note = "Regresjonstabell 3: {stars}. T statistics in brackets."
)
```

```
mod4 <- "inntekt ~ sex + hoyde_cm + vekt_kg + I(vekt_kg^2) + BMI + I(BMI^2)"
lm4 <- lm(mod4, data = hoyde)
summary(lm4)
```

```
##
## 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
## -702996 -248065 -89963  127248 2709094
##
## Coefficients:
##              Estimate Std. Error t value      Pr(>|t|)
## (Intercept) -225754.56  1214279.15  -0.186      0.853
## sexfemale   -111585.61   15869.36  -7.032 0.000000000000224 ***
## hoyde_cm     3702.50    7191.63   0.515      0.607
## vekt_kg       8228.69   12661.45   0.650      0.516
## I(vekt_kg^2)   -33.41     31.75  -1.052      0.293
## BMI          -20072.70   36148.90  -0.555      0.579
## I(BMI^2)        158.44    263.85   0.600      0.548
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 459000 on 6904 degrees of freedom
## (95 observations deleted due to missingness)
## Multiple R-squared:  0.0586, Adjusted R-squared:  0.05778
## F-statistic: 71.62 on 6 and 6904 DF,  p-value: < 0.00000000000000022
```

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>.