

Er det høyde som bestemmer inntekt?

Av Ann Elisabeth og Heidi Marie Rolfsnes

Innledning

I denne oppgaven vil vi se nærmere på om høyde har noe å si på inntekt. Utgangspunktet for denne oppgaven bygger på et datasett vi har fått utdelt. Det er hentet fra National Longitudinal study, som er gjennomført på vegne av U.S Bureau of Labor statistics.

Oppsummeringsstatistikker.

I denne tabellen har vi laget en statistikk der vi tar for oss: inntekt, høyde, vekt, alder, sivilstatus, kjønn, antall år med utdanning og scoren de fikk på AFQT. Statistikkene vi ser er min og max som vi kan sammenligne med minste inntekt og høyest inntekt, eller minst alder og høyest alder. Median er tallet vi finner i midten. Hvis vi setter opp alle inntektene i stigende rekkefølge er det tallet som er i midten vi kaller median (boka side 300). videre kan vi se på Mean. På norsk betyr det gjennomsnitt. Det vil si at vi plusser sammen alle aldrene til personene i utvalget og så deler så på antallet. Vi kommer frem til gjennomsnittsalderen for utvalget vårt.

Ut fra undersøkelsen kan vi se et utvalg på 7006 personer, der 3402 er menn og 3604 er kvinner. Aldersspennet i utvalget er fra 47 til 56 år, der medianen alder er 51 år og gjennomsnittsalderen er 51,33 år. Inntektene til de utvalgte går fra 0 USD til 343830 USD, her kan vi også se at medianen er 29589,50 USD mens gjennomsnittet er på 41203,90 USD. Om vi ser på høyden så er den laveste personen 52 inch og den høyeste 84 Inch.

Teori

hva er et boxplot?

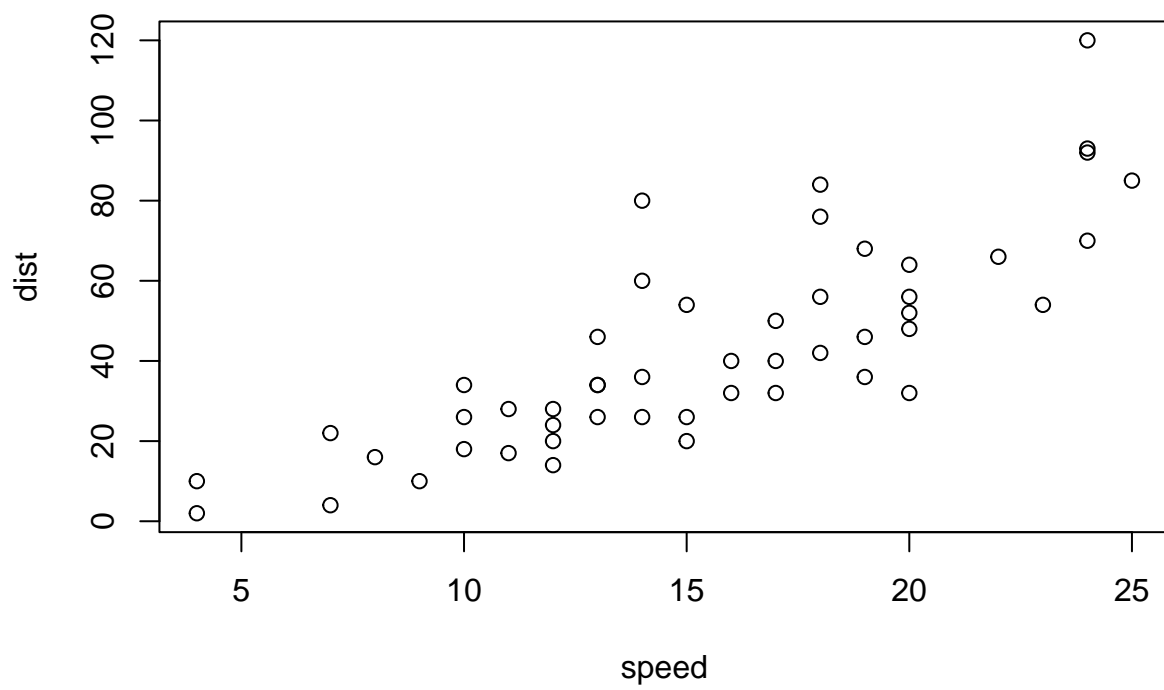
hva er punkt diagram?

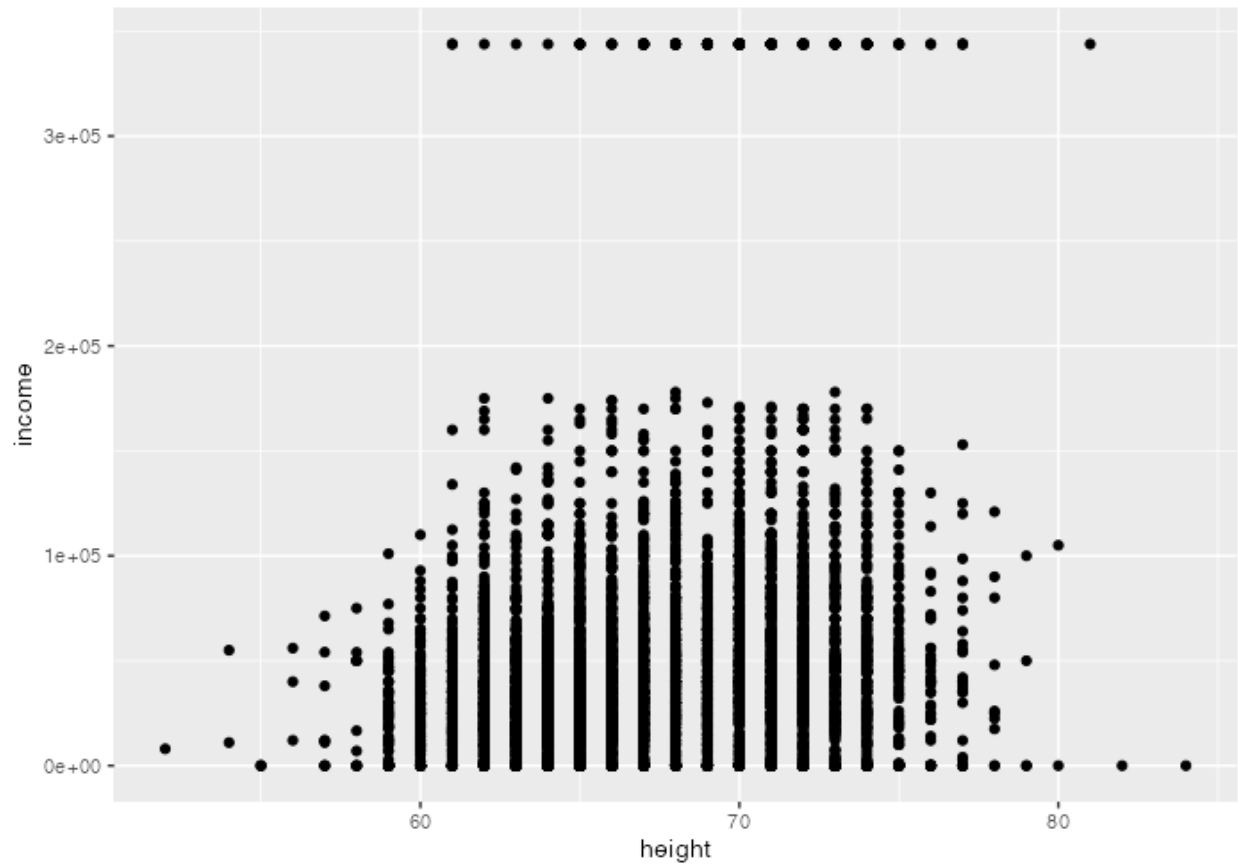
hva er regresjonsanalyse og hva blir den brukt til?

Statistikk

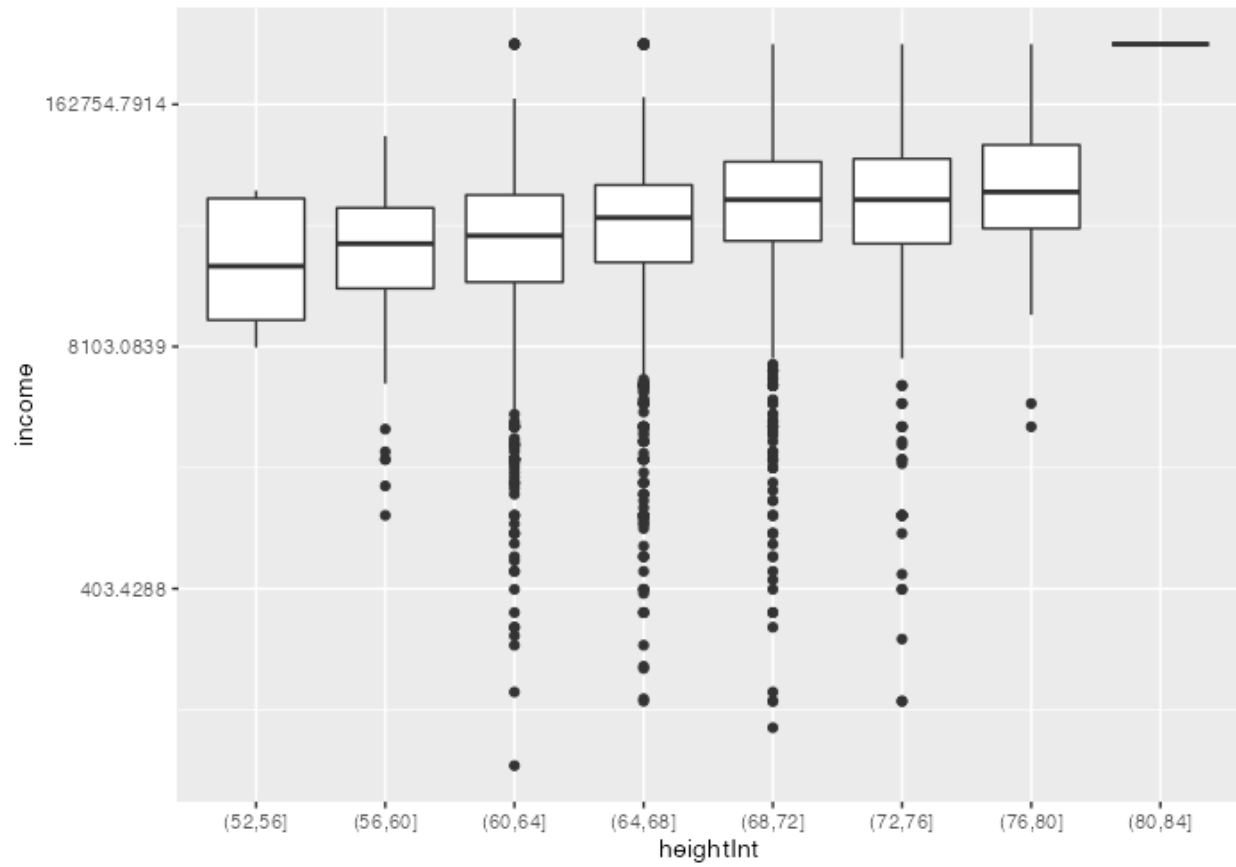
Innledende tekst til oppgave

```
plot(cars)
```

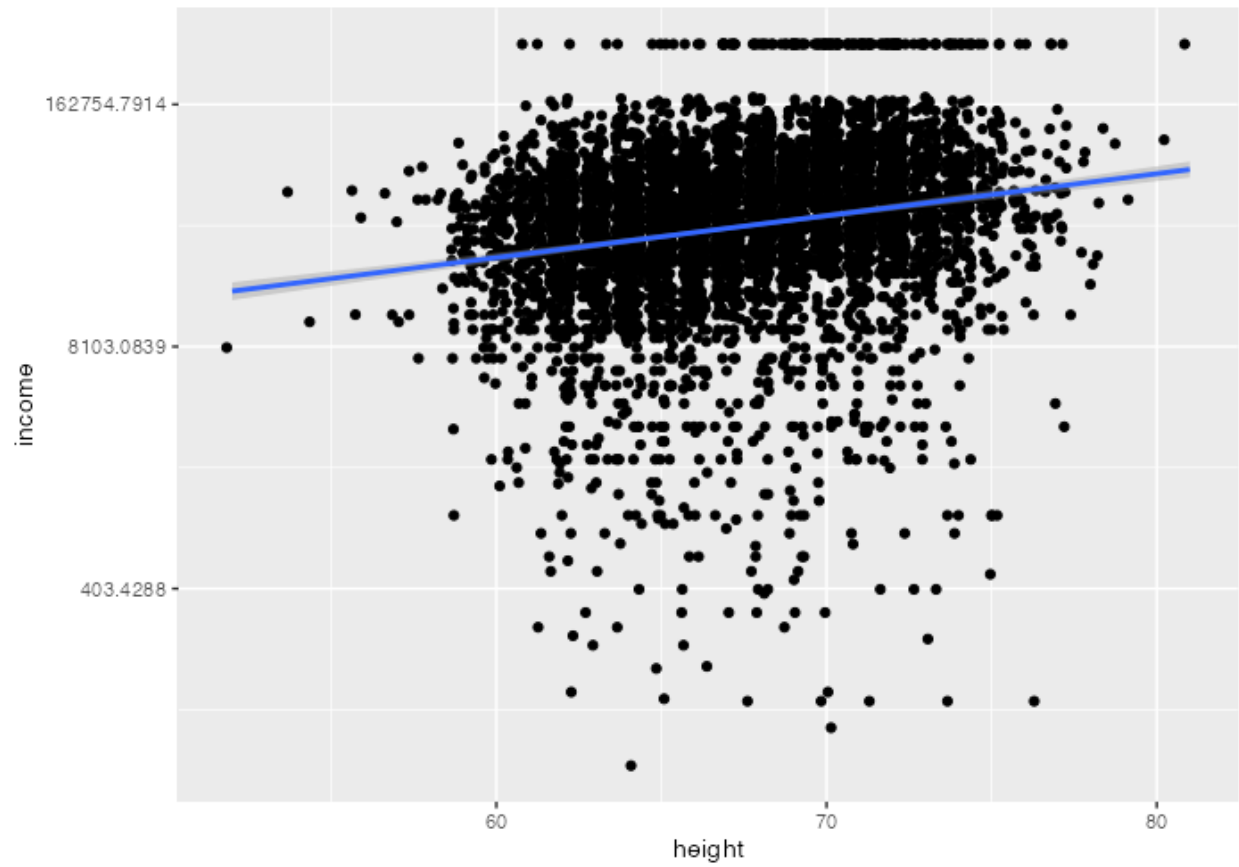




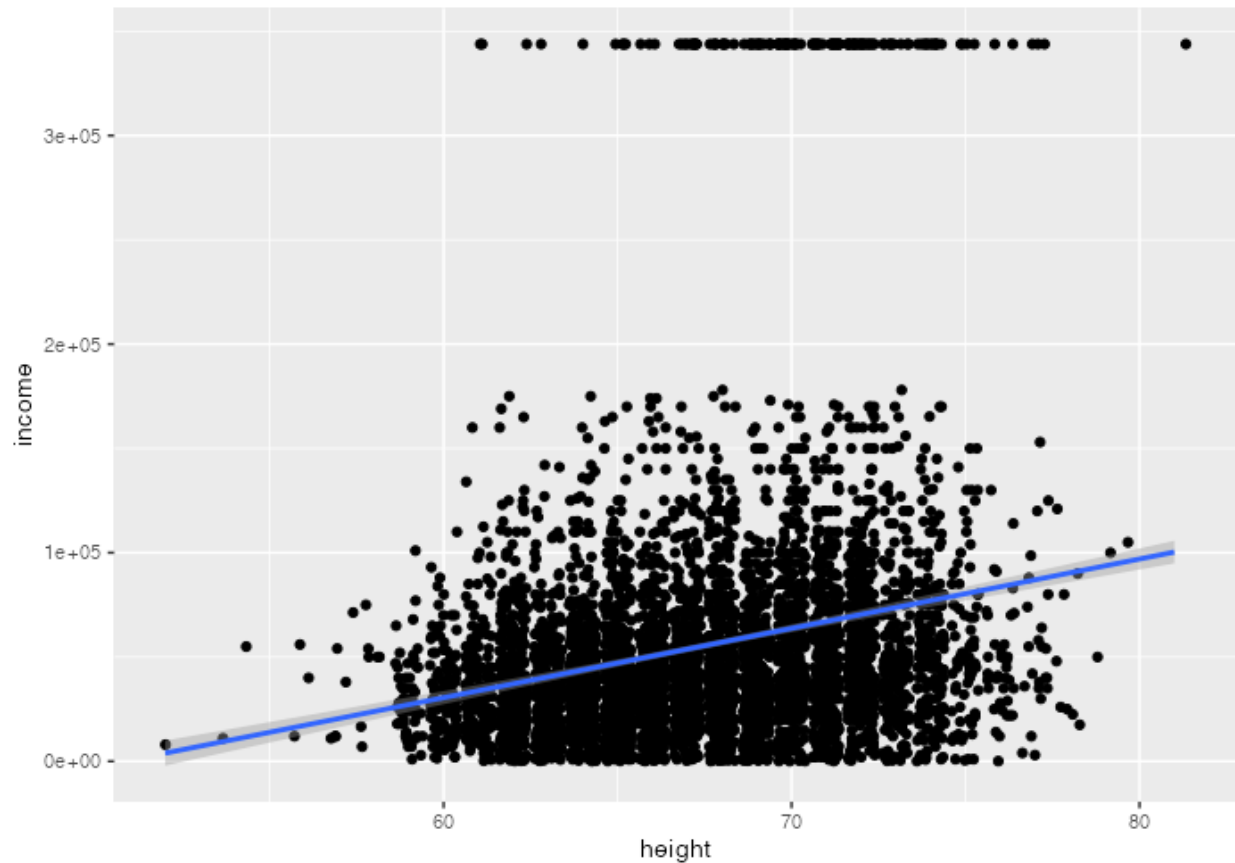
Diagrammet viser høyde i forhold til inntekt. På den vertikale linjen finner vi inntekten, og på den horisontale finner vi høyden. Her ser vi at en liten gruppe med folk i alle høyder som har en relativt høyere inntekt en resten, mellom denne gruppen og resten av utvalget er det et stort gap. noe vi kan kjenne oss igjen i når vi ser på den deskriptive statistikken.



I dette Boxplotet ser vi inntekt i forhold til HeightInt



I punkt diagrammet ser vi inntekt og høyde mot hverandre, hver prikk representerer en person i utvalget med personens høyde og inntekt.



Hvor mye tjener vi pr ekstra inch?

En ekstra inch (2,72cm) gir oss \$2982 ekstra i inntekt pr år.

Er det mer til historien?

Oppgave 2 - Hoyde

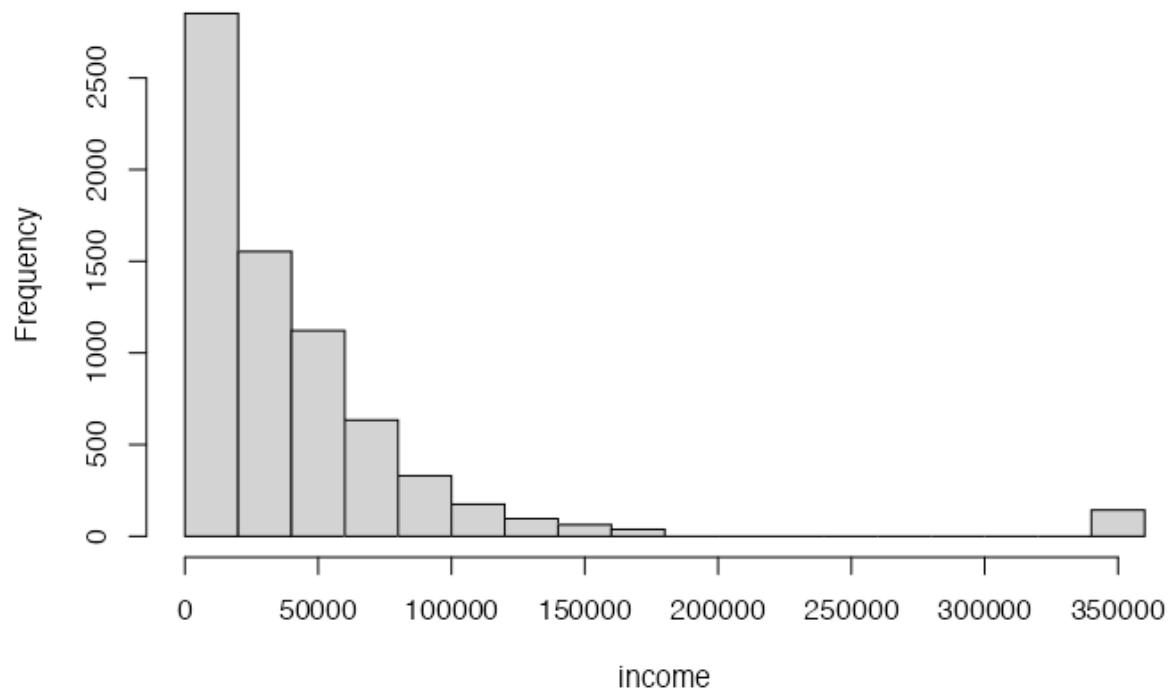
1. Beskrivende statistikk , dvs. kort beskrivelse av dataene

```
# A tibble: 6 × 8
```

	income	height	weight	age	marital	sex	education	afqt
	<int>	<dbl>	<int>	<int>	<fct>	<fct>	<int>	<dbl>
1	<u>19000</u>	60	155	53	married	female	13	6.84
2	<u>35000</u>	70	156	51	married	female	10	49.4
3	<u>105000</u>	65	195	52	married	male	16	99.4
4	<u>40000</u>	63	197	54	married	female	14	44.0
5	<u>75000</u>	66	190	49	married	male	14	59.7
6	<u>102000</u>	68	200	49	divorced	female	18	98.8

2. Lag et histogram av variabelen income

Income Histogram



2.1 Hva er forklaringen på utliggerne langt til høyre?

Grunnen til at det ligger en søyle helt til høyre er på grunn av at det er en andel personer i utvalget som har en vesentlig høyere inntekt enn det resterende utvalget. ut i fra histogrammet kan vi se at fleste tjener under 50000.

2.2 Har vi personer uten inntekt i datasettet?

Når vi ser i histogrammet kan vi se en stor andel med personer som har null i inntekt.

Regresjonsanalyse

Uten inntekt 0:

```
# A tibble: 5,266 × 8
  income height weight age marital sex education afqt
  <int>   <dbl>   <int> <int> <fct>   <fct>   <int> <dbl>
1  19000     60    155   53 married female     13  6.84
2  35000     70    156   51 married female     10 49.4
3 105000     65    195   52 married male      16 99.4
4  40000     63    197   54 married female     14 44.0
5  75000     66    190   49 married male      14 59.7
6 102000     68    200   49 divorced female     18 98.8
7  70000     64    160   54 divorced female     12 50.3
8  60000     69    162   55 divorced male      12 89.7
9 150000     69    194   54 divorced male      13 96.0
10 115000     64    145   53 married female     16 67.0
# ... with 5,256 more rows
```

Resultat

Konklusjon

Litteraturgjennomgang

Referanser