MSB106 - Assignment

Sindre H. Øveraas, Alen Colakovic, Mona Lisa Jones & Sebastian M. Fløysand

```
library(rgdal)
library(dplyr)
library(RSQLite)
library(sf)
library(tidyverse)
library(readr)
library(ggplot2)
library(kableExtra)
library(huxtable)
```

```
NOR_CBD <- read_csv("NOR_CBD.csv")
Dist_CBD_Dentist <- read_csv("Dist_CBD_Dentist.csv")
```

```
Warning: One or more parsing issues, see `problems()` for details
```

```
Zonal_CBD <- read_csv("Zonal_CBD.csv")
Dist_Mal_Dentist <- read.csv("Dist_Mal_Dentist.csv")
NOR_OSM_SHOP_MAL <- read.csv("NOR_OSM_SHOP_MAL.csv") |>
  rename("HubName" = fid)
Zonal_Den <- read.csv("Zonal_Den.csv") |>
  select(fid, X_sum)
Dist_Comp_Den <- read.csv("Dist_Comp_Den.csv") |>
  select(fid, distance)
NOR_KOMM <- read_csv("NOR_KOMM.csv")
NOR_KOMM_2 <- read_csv("NOR_KOMM_2.csv")
```

```
Commune_Data <- inner_join(NOR_KOMM, NOR_KOMM_2, by = "kommunenummer") |>
  select(kommunenummer, HubName, HubDist, Turnover_capita_retail_Omsetning) |>
  rename("knr" = kommunenummer, "DistMal" = HubDist, "Turnover_capita_retail" = Turnover_capita_reta
```

```
Commune_Data <- inner_join(Commune_Data, NOR_OSM_SHOP_MAL, by = "HubName") |>
  select(knr, DistMal, HubName, Turnover_capita_retail, shop, Vinmonopolet) |>
  rename("SizeMall" = shop, "Winemonopoly" = Vinmonopolet)
```

```
Commune_Data$SizeMall <- as.numeric(Commune_Data$SizeMall)
```

```
Warning: NAs introduced by coercion
```

```
Commune_Data$Winemonopoly <- as.integer(as.logical(Commune_Data$Winemonopoly))
```

```
Dentist_Data <- inner_join(Dist_Mal_Dentist, Dist_CBD_Dentist, by = "fid") |>
  select(fid, Juridisk.n, Antall.ans, Sum.Drifts, Sum.salgsi, Driftsresu.y, osm_id.y, latlong.y, Hubl
  rename("DistMal" = HubDist.x, "DistCBD" = HubDist.y, "HubName" = HubName.x, "HubNameCBD" = HubName
```

```
Dentist_Data <- inner_join(Dentist_Data, Zonal_Den, by = "fid")
```

```
Dentist_Data <- inner_join(Dentist_Data, Dist_Comp_Den, by = "fid") |>
  rename("DistComp" = distance)
```

```
Dentist_Data <- left_join(Dentist_Data, NOR_OSM_SHOP_MAL, by = "HubName") |>
  select(fid, Juridisk.n, Antall.ans, Driftsres, Sum.salgsi, latlong, DistCBD, DistMal, X_sum, DistC
  rename("SizeMall" = shop, "Winemonopoly" = Vinmonopolet)
```

```
Dentist_Data$SizeMall <- as.numeric(Dentist_Data$SizeMall)
```

```
Warning: NAs introduced by coercion
```

```
Dentist_Data$Winemonopoly <- as.integer(as.logical(Dentist_Data$Winemonopoly))
```

```
kable(summary(Commune_Data))
```

	knr	DistMal	HubName	Turnover_capita_retail	SizeMall	Winemonopoly
	Length:435	Min. : 0.01854	Min. : 6.0	Min. : 17366	Min. : 2.00	Min. :0.0000
	Class :character	1st Qu.: 9.41403	1st Qu.:220.0	1st Qu.: 59250	1st Qu.: 13.00	1st Qu.:0.0000
	Mode :character	Median :18.87373	Median :291.0	Median : 83203	Median : 25.00	Median :1.0000
	NA	Mean :23.82982	Mean :280.8	Mean : 84719	Mean : 37.22	Mean :0.5701
	NA	3rd Qu.:32.67127	3rd Qu.:363.0	3rd Qu.:106257	3rd Qu.: 50.00	3rd Qu.:1.0000
	NA	Max. :98.37355	Max. :424.0	Max. :218728	Max. :206.00	Max. :1.0000
	NA	NA	NA	NA	NA's :124	NA's :107

```
kable(summary(Dentist_Data[3:6]))
```

	Antall.ans	Driftsres	Sum.salgsi	latlong
	Min. : 0.00	Min. :-30319000	Min. : -28000	Length:5740
	1st Qu.: 1.00	1st Qu.: -5000	1st Qu.: 658250	Class :character
	Median : 1.00	Median : 330000	Median : 3176500	Mode :character
	Mean : 39.26	Mean : 4842620	Mean : 50123495	NA
	3rd Qu.: 3.00	3rd Qu.: 1201750	3rd Qu.: 6248250	NA
	Max. :6870.00	Max. :108100000	Max. :1058102000	NA
	NA	NA's :3838	NA's :3838	NA

```
kable(summary(Dentist_Data[7:10]))
```

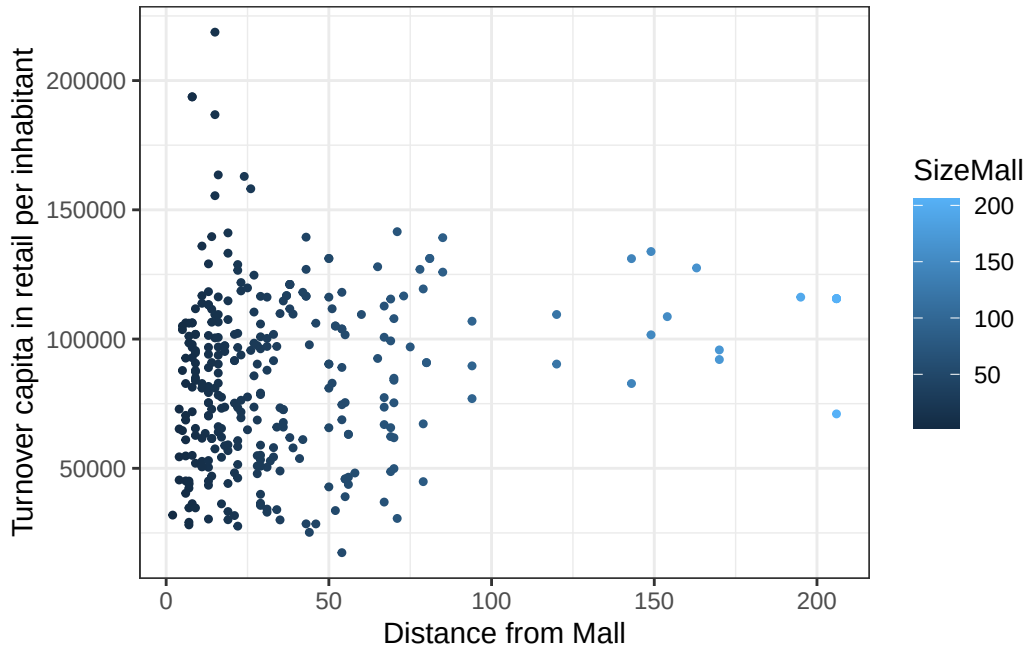
	DistCBD	DistMal	X_sum	DistComp
	Min. : 0.0046	Min. : 0.0008	Min. : 0	Min. : 0.0
	1st Qu.: 1.3912	1st Qu.: 0.3145	1st Qu.: 6129	1st Qu.: 0.0
	Median : 6.7067	Median : 0.9164	Median :14602	Median : 0.0
	Mean : 20.5000	Mean : 4.6331	Mean :22455	Mean : 614.4
	3rd Qu.: 19.8885	3rd Qu.: 2.5220	3rd Qu.:28904	3rd Qu.: 197.7
	Max. :470.1005	Max. :105.8867	Max. :92591	Max. :60577.8

```
kable(summary(Dentist_Data[11:12]))
```

	SizeMall	Winemonopoly
	Min. : 1.0	Min. :0.0000
	1st Qu.: 21.0	1st Qu.:0.0000
	Median : 40.0	Median :1.0000
	Mean : 51.1	Mean :0.6429
	3rd Qu.: 70.0	3rd Qu.:1.0000
	Max. :206.0	Max. :1.0000
	NA's :1388	NA's :1274

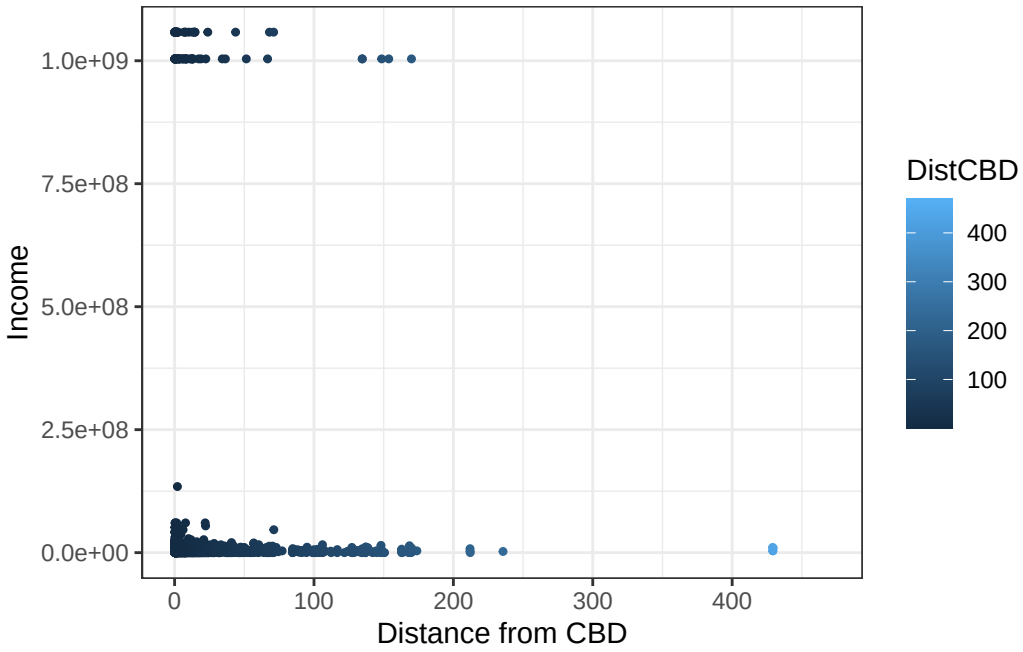
```
Commune_Data |>
  ggplot(aes(x = SizeMall, y = Turnover_capita_retail, colour = SizeMall)) +
  geom_point(lwd = .9) +
  labs(x = "Distance from Mall", y = "Turnover capita in retail per inhabitant") +
  theme_bw()
```

```
Warning: Removed 124 rows containing missing values (geom_point).
```



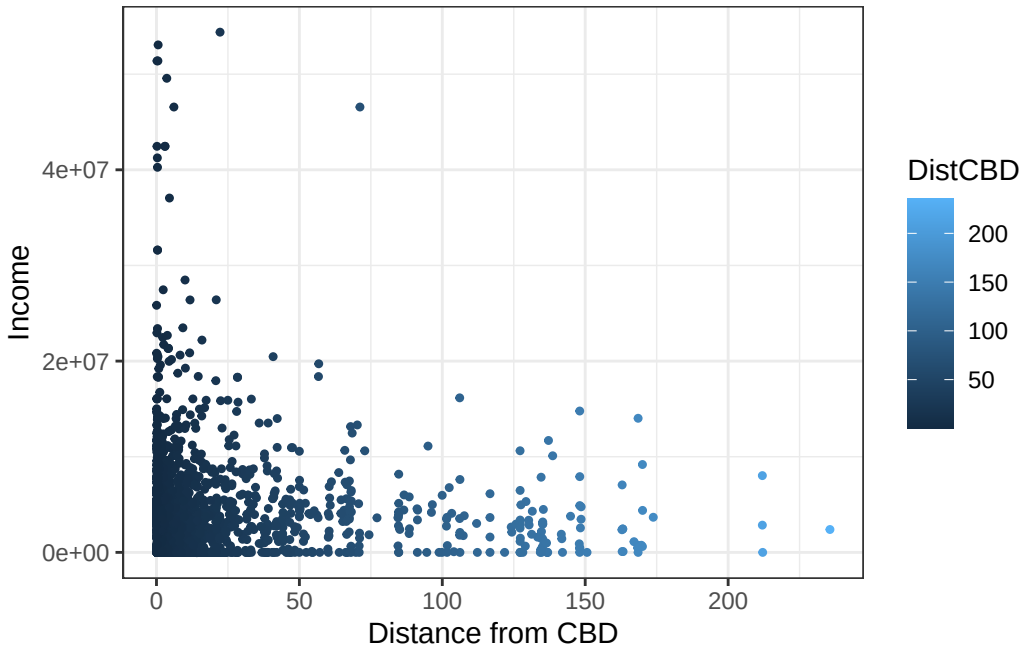

```
Dentist_Data |>
  ggplot(aes(x = DistCBD, y = Sum.salgsi, colour = DistCBD)) +
  geom_point(lwd = .9) +
  labs(x = "Distance from CBD", y = "Income") +
  theme_bw()
```

```
Warning: Removed 3838 rows containing missing values (geom_point).
```



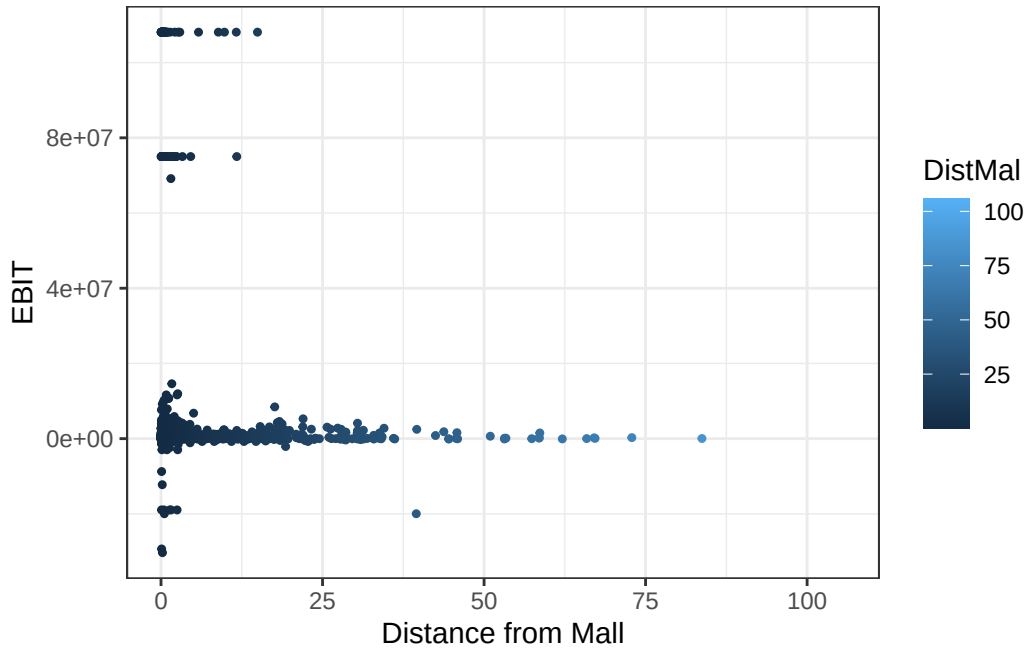

```
Dentist_Data_LIMIT <- Dentist_Data |>
  filter(Sum.salgsi < 100000000, DistCBD < 250,
         Driftsres < 20000000, Driftsres > -5000000, DistMal < 80)
```

```
Dentist_Data_LIMIT |>
  ggplot(aes(x = DistCBD, y = Sum.salgsi, colour = DistCBD)) +
  geom_point(lwd = .9) + labs(x = "Distance from CBD", y = "Income") +
  theme_bw()
```

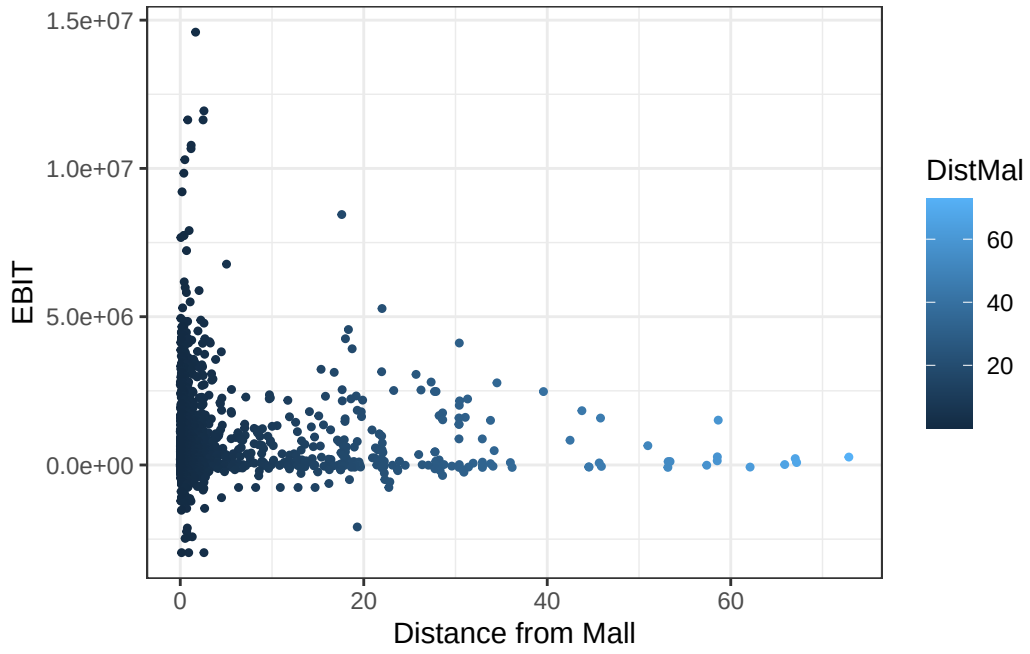



```
Dentist_Data |>
  ggplot(aes(x = DistMal, y = Driftsres, colour = DistMal)) +
  geom_point(lwd = .9) +
  labs(x = "Distance from Mall", y = "EBIT") +
  theme_bw()
```

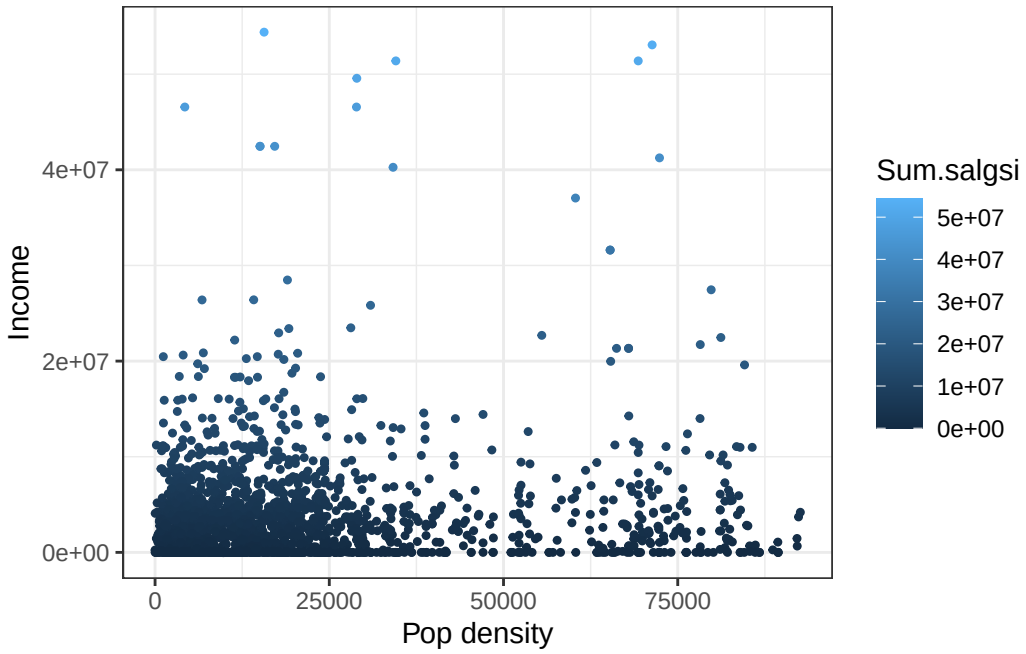
```
Warning: Removed 3838 rows containing missing values (geom_point).
```




```
Dentist_Data_LIMIT |>
  ggplot(aes(x = DistMal, y = Driftsres, colour = DistMal)) +
  geom_point(lwd = .9) +
  labs(x = "Distance from Mall", y = "EBIT") +
  theme_bw()
```




```
Dentist_Data_LIMIT |>
  ggplot(aes(x = X_sum, y = Sum.salgsi, colour = Sum.salgsi)) +
  geom_point(lwd = .9) +
  labs(x = "Pop density", y = "Income") +
  theme_bw()
```




```
lm1 <- lm(Sum.salgsi ~ DistComp + DistMal + SizeMall + (DistMal*SizeMall) + DistCBD + X_sum + (DistC
```

```
huxreg(list("Dentist Revenue" = lm1), statistics = c(N = "nobs", R2 = "r.squared"), const. = "(Inter
```

```
Den_W_Wine <- Dentist_Data_LIMIT |>
  filter(Winemonopoly == 1)
```

```
Den_WO_Wine <- Dentist_Data_LIMIT |>
  filter(Winemonopoly == 0)
```

```
lm2 <- lm(Turnover_capita_retail ~ DistMal + SizeMall + (DistMal*SizeMall) + Winemonopoly, data = Co
```

```
huxreg(list("Turnover per capita in retail" = lm2), statistics = c(N = "nobs", R2 = "r.squared"), co
```

Econometric approach

	Dentist Revenue
(Intercept)	3616068.187 ***
	(373568.663)
DistComp	-105.228
	(152.737)
DistMal	-13619.260
	(32749.952)
SizeMall	10063.209 *
	(4523.296)
DistCBD	7969.413
	(9756.588)
X_sum	13.257
	(7.925)
DistMal:SizeMall	-106.204
	(724.299)
DistCBD:X_sum	-1.105
	(1.065)
N	1364
R2	0.008

Note: *** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$ T statistics in brackets.

	Turnover per capita in retail
(Intercept)	90614.356 ***
	(4355.729)
DistMal	-393.142 **
	(139.667)
SizeMall	175.797 *
	(75.816)
Winemonopoly	-1319.795
	(3829.574)
DistMal:SizeMall	-3.390
	(3.293)
N	311
R2	0.116

Note: *** p < 0.001; ** p < 0.01; * p < 0.05 T statistics in brackets.