

# Task 4

Oscar Beck

6/5/2020

Our communication objective is “What characterises the areas where Liberal Alliance and Konservative are popular?”

I load in the data and merge the two files with the same code used earlier in the assignment

```
#Libraries used in the task
library(readxl)
library(dplyr)
library(tidyr)
library(stringr)
library(ggplot2)

#Loading the data and merging the two files with the same code as earlier in the assignment
FV19TOTA <- read_excel("FV19TOTA.xlsx")
FV19TOTA <- FV19TOTA %>%
  rename("Party" = "Folketingsvalget 2019 efter valgresultat, område og tid",
         "Area" = "...2",
         "Votes" = "...3")
FV19TOTA$Votes <- as.integer(FV19TOTA$Votes)
characteristics_data <- read_excel("PollingDistrictCharacteristics.xlsx",skip=4)

FV19TOTA <- fill(FV19TOTA, "Party", .direction = "down")
characteristics_data <- characteristics_data %>% rename("Area" = "Valgdistrikt - navn")
rough_merged <- full_join(FV19TOTA, characteristics_data, by = "Area")
vote_characteristics <- subset(rough_merged, !is.na(rough_merged$Storkreds))
vote_characteristics <- vote_characteristics %>% rename("Polling_district" = "Area","disponibel_income")

#Splitting the data into variables for "Det konservative folkeparti" and "Liberal Alliance"
konservative <- subset(vote_characteristics, Party == "C. Det Konservative Folkeparti")
konservative <- transform(konservative,disponibel_income = as.numeric(disponibel_income))
liberal <- subset(vote_characteristics, Party == "I. Liberal Alliance")
liberal <- transform(liberal,disponibel_income = as.numeric(disponibel_income))

#Merging the different polling districts in every "opstillingskreds" into one and in the process taking
liberal_merged <- liberal
liberal_merged <- liberal_merged %>% group_by(Party,Storkreds,Opstillingskreds) %>% summarise(Avg_Income

#Now the same for the Konservative
konservative_merged <- conservative
konservative_merged <- conservative_merged %>% group_by(Party,Storkreds,Opstillingskreds) %>% summarise
```

To answer our communication objective we will look at the income of the voters of both parties in the different areas.

```
#Liberal Alliance
votes_LA <- liberal_merged$Votes
income_LA <- liberal_merged$Avg_Income

#Linear Regression
lm_LA <-lm(income_LA~votes_LA)
lm_LA
```

```
##
## Call:
## lm(formula = income_LA ~ votes_LA)
##
## Coefficients:
## (Intercept)      votes_LA
##    195122.06         53.28
```

```
#Finding the R2 value
rsq <- function(x, y) cor(x, y) ^ 2
R2 <- rsq(votes_LA,income_LA)
R2
```

```
## [1] 0.4053055
```

```
#Plotting the scatterplot
png("LA.png", units="in", width=10, height=5, res=300)
ggplot(liberal_merged, aes(x=Votes,y=Avg_Income, colour = Storkreds)) + geom_point(position = position_
dev.off()
```

```
## pdf
##    2
```

```
#Det Konservative Folkeparti
votes_C <- konservative_merged$Votes
income_C <- konservative_merged$Avg_Income

#Linear Regression
lm_C <-lm(income_C~votes_C)
lm_C
```

```
##
## Call:
## lm(formula = income_C ~ votes_C)
##
## Coefficients:
## (Intercept)      votes_C
##    190029.19         20.67
```

```
#Finding the R2 value
```

```
rsq <- function (x, y) cor(x, y) ^ 2
```

```
R2 <- rsq(votes_C,income_C)
```

```
R2
```

```
## [1] 0.6398424
```

```
#Plotting the scatterplot
```

```
png("C.png", units="in", width=10, height=5, res=300)
```

```
ggplot(konservative_merged, aes(x=Votes,y=Avg_Income, colour = Storkreds)) + geom_point(position = position_jitter) +  
dev.off()
```

```
## pdf
```

```
## 2
```