

1. Packages -----

```
library("readxl")
library("openxlsx")
library("tidyverse")
library("ggrepel")
library("sjPlot")
library("stargazer")
library("ggpubr")
library("tidyr")
```

2. Regression -----

Importiere Daten und teile sie auf in Dispobestand bekannt und unbekannt

```
DataRegression <- read_excel("./Datensatz/DatensatzRegression.xlsx")
```

```
TrainData <- subset(DataRegression,!is.na(Dispobestand))
```

```
PredictionData <- subset(DataRegression,is.na(Dispobestand))
```

Orientierung: Diagramme TrainData (ohne unbekannte Dispobestände)

```
v1 <- ggplot(TrainData, aes(log(Marktkapitalisierung), Dispobestand))+geom_point()+theme_minimal()+
+ xlab("ln(Marktkapitalisierung)") + theme(text=element_text(family="Times New Roman"))
```

```
v2 <- ggplot(TrainData, aes(Stimmrechtsaktien, Dispobestand))+geom_point()+theme_minimal()+
+ theme(text=element_text(family="Times New Roman"))
```

```
v3 <- ggplot(TrainData, aes(Staatsbesitz, Dispobestand))+geom_point()+theme_minimal()+
+ theme(text=element_text(family="Times New Roman"))
```

```
v4 <- ggplot(TrainData, aes(Immobilien, Dispobestand))+geom_point()+theme_minimal()+
+ theme(text=element_text(family="Times New Roman"))
```

```
v5 <- ggplot(TrainData, aes(Investment, Dispobestand))+geom_point()+theme_minimal()+
+ theme(text=element_text(family="Times New Roman"))
```

```
v6 <- ggplot(TrainData, aes(StimmkraftBedeutendeAktionäre, Dispobestand))+geom_point()+theme_minimal()+
+ theme(text=element_text(family="Times New Roman"))
```

```
v7 <- ggplot(TrainData, aes(Dividendenrendite, Dispobestand))+geom_point()+theme_minimal()+
+ theme(text=element_text(family="Times New Roman"))
```

```
v8 <- ggplot(TrainData, aes(Streubesitz, Dispobestand))+geom_point()+theme_minimal()+
+ theme(text=element_text(family="Times New Roman"))
```

```
Zusammen <- ggarrange(v1, v2,v3,v4,v5,v6,v7,v8, ncol = 2, nrow = 4)
```

Zusammen

Modelle

Modell i) Komplexes Modell: Alle acht erklärenden Variablen

```
modell1 <- lm(Dispobestand ~ log(Marktkapitalisierung) + Stimmrechtsaktien + Staatsbesitz + Immobilien +
+ Investment + StimmkraftBedeutendeAktionäre + Dividendenrendite + Streubesitz, data = TrainData)
```

modell1

```
summary(modell1)
```

Modell2 ii) Vereinfachtes Modell: Nur statistisch signifikante Variablen (*** p≤0.001) des modell1:

Marktkapitalisierung und Streubesitz

```
modell2 <- lm(Dispobestand ~ log(Marktkapitalisierung) + Streubesitz, data = TrainData)
```

modell2

```
summary(modell2)
```

Vergleich der Modelle mit ANOVA: Modell2 ist das einfachere und Modell1 das komplexere Modell; es handelt sich um "nested Models".

```
VergleichAnova <- anova(modell2,model1)
```

VergleichAnova

Fazit: Das einfachere modell2 wird verwendet, denn der resultierende p-Wert aus der ANOVA-Analyse ist nicht signifikant (0.1476)

ANOVA-Tabelle für Word

Anmerkung: Die Tabelle wurde in Word nachträglich noch bearbeitet.

```
tab_df(VergleichAnova,file="TabelleAnova.doc")
```

Regressionstabellen für Word: Vergleich der beiden Modelle

Anmerkung: Die Tabelle wurde in Word nachträglich noch bearbeitet.

```
stargazer(modell1, modell2, type = "html", out = "Regressionstabelle.doc", title="Results", align=TRUE)
```

```

# Vergleich der Predictions von Modell1 und Modell2
PredictionsModell1 <- as.data.frame(predict(modell1,PredictionData,interval="confidence"))
PredictionsModell2 <- as.data.frame(predict(modell2,PredictionData,interval="confidence"))

Modell1 <- PredictionsModell1$fit
DataVergleichPredictions <- as.data.frame(Modell1)
DataVergleichPredictions$Modell2 <- PredictionsModell2$fit

VergleichPredictionsPlot <- ggplot(data=DataVergleichPredictions, aes(Model1, Model2),family ="Times New
Roman")+ geom_point()+theme_minimal()+
  geom_segment(aes(x = 0, y = 0, xend = 40, yend = 40), size=0.1)+ ggtitle("Vergleich der Schätzergebnisse",
subtitle="Komplexeres Modell (1) vs. einfacheres Modell (2)")+
  xlab("Modell 1")+ ylab("Modell 2")+ theme(text=element_text(family="Times New Roman"), plot.title =
element_text(face = "bold"))

VergleichPredictionsPlot

# Füge Predictions von modell2 (siehe Begründung für die Wahl weiter oben) in den ursprünglichen Datensatz
ein

PredictionDataNachRegression <- PredictionData
PredictionDataNachRegression$Dispobestand <- PredictionsModell2$fit

DataNachRegression <- rbind(TrainData,PredictionDataNachRegression)

# 3. 3-D Graphik des gewählten modell2 -----

library("scatterplot3d")

LogMarktkapitalisierung <- log(DataNachRegression$Marktkapitalisierung)
Daten3DModell <- as.data.frame(LogMarktkapitalisierung)
Daten3DModell$Streubesitz <- DataNachRegression$Streubesitz
Daten3DModell$Dispobestand <- DataNachRegression$Dispobestand

s3d <- scatterplot3d(Daten3DModell, type = "h", color = "black", pch = 16, angle= 55,main = "Graphische
Darstellung der Schätzung der unbekannten Dispobestände")

s3d$plane3d(modell2)

# Hebe die Schätzungen im Diagramm hervor: Reihen 149 bis 196 von Daten3DModell

Daten3DModellHervorhebung <- tail(Daten3DModell, 48)

s3d$points3d(Daten3DModellHervorhebung$LogMarktkapitalisierung,Daten3DModellHervorhebung$Streubesitz,
Daten3DModellHervorhebung$Dispobestand,
  col = "red", type = "h", pch = 16)

# 4. Speichere fertigen Datensatz -----

# Detaillierte Version als Excel
write.xlsx(DataNachRegression, 'DataNachRegressionAusführlich.xlsx')

# Reduzierte Version für Word

DataNachRegressionReduziert <- subset(DataNachRegression, select = -
c(Symbol,Valorennummer,Land,Handelswaehrung,Papierart))
tab_df(DataNachRegressionReduziert,file="DatensatzNachRegressionReduziert.doc")

# 5. Histogramm Dispobestände -----

HistogrammDaten <- DataNachRegression
HistogrammDaten$KategorieMarktkapitalisierung <-
ifelse(HistogrammDaten$Marktkapitalisierung<500,1,ifelse(HistogrammDaten$Marktkapitalisierung>=500 &
HistogrammDaten$Marktkapitalisierung < 5000,2,
ifelse(HistogrammDaten$Marktkapitalisierung>=5000,3,".")))

ggplot(HistogrammDaten, aes(x=Dispobestand,fill = HistogrammDaten$KategorieMarktkapitalisierung)) +
  geom_histogram(breaks=seq(0,max(HistogrammDaten$Dispobestand)+10,10), color="black")
+theme_minimal()+

```

```

xlab("Dispobestand einer Gesellschaft")+ ylab("Anzahl Gesellschaften")+
theme(text=element_text(family="Times New Roman"))+ scale_fill_grey()

```

```

HistogrammDatenNichtGeschätzt <- HistogrammDaten[-c(149:196), ]
ggplot(HistogrammDatenNichtGeschätzt, aes(x=Dispobestand, fill =
HistogrammDatenNichtGeschätzt$KategorieMarktkapitalisierung)) +
  geom_histogram(breaks=seq(0,max(HistogrammDatenNichtGeschätzt$Dispobestand)+10,10), color="black")
+theme_minimal()+
  xlab("Dispobestand einer Gesellschaft")+ ylab("Anzahl Gesellschaften")+
  theme(text=element_text(family="Times New Roman"))+ scale_fill_grey()+ scale_y_continuous(breaks=seq(0,
60, 20), limits=c(0, 60))

```

```

HistogrammDatenGeschätzt<- HistogrammDaten[-c(1:148), ]
ggplot(HistogrammDatenGeschätzt, aes(x=Dispobestand, fill =
HistogrammDatenGeschätzt$KategorieMarktkapitalisierung)) +
  geom_histogram(breaks=seq(0,max(HistogrammDatenGeschätzt$Dispobestand)+10,10), color="black")
+theme_minimal()+
  xlab("Dispobestand einer Gesellschaft")+ ylab("Anzahl Gesellschaften")+
  theme(text=element_text(family="Times New Roman"))+ scale_fill_grey()

```

6. Graphik Beherrschungsargument -----

```

# Schwellenwert: 50%, ohne Berücksichtigung der eigenen Aktien und fiktiver GV-Präsenzzahlen
# Importiere überarbeiteten Datensatz, ohne die 14 Fälle von Stimmrechtsaktien
DatenBeherrschungsargumentGraphik <- read_excel("./Datensatz/
DataNachRegressionAusführlichOhneStimmrechtsaktien.xlsx")
DatenBeherrschungsargumentGraphik$Geschätzt <- as.factor(DatenBeherrschungsargumentGraphik$Geschätzt
)

```

```

# Entferne einen Ausreisser aus Darstellungsgründen
DatenBeherrschungsargumentGraphik <- DatenBeherrschungsargumentGraphik[-c(77), ]

```

```

# Die Graphik wurde in Word noch weiterbearbeitet
Beherrschungsgraphik <- ggplot(DatenBeherrschungsargumentGraphik, aes(x=GrössterAktionär,
y=ZusätzlicheStimmkraftGrössterAktionär, fill = Geschätzt, size = log(Marktkapitalisierung))) +
  geom_point(shape=21) + geom_segment(x=50, y=0, xend=0, yend=50, size=0.2, color = "black") +
  geom_segment(x=50, y=0, xend=50, yend=100, size=0.2, color = "black")+ theme_minimal()+
  xlab("Stimmkraft des grössten Aktionärs (%)")+
  ylab("Zusätzliche Stimmkraft aufgrund des Dispobestandes (%)")+ theme(text=element_text(family="Times
New Roman"), plot.title = element_text(face = "bold"))+ scale_fill_grey()
(Beherrschungsgraphik)

```

7. Graphik beliebiger Schwellenwert -----

```

# (i) Ohne eigene Aktien, ohne Annahme über GV-Präsenzzahl

```

```

# Importiere überarbeiteten Datensatz, ohne die 14 Fälle von Stimmrechtsaktien
DatenBeliebigerSW <- read_excel("./Datensatz/DataNachRegressionAusführlichOhneStimmrechtsaktien.xlsx")

```

```

countfreqONE <- function(n,s){
  #Set k to 0
  k <- 0
  for(i in 1:n) {
    if (DatenBeliebigerSW$GrössterAktionär[i] +
DatenBeliebigerSW$ZusätzlicheStimmkraftGrössterAktionär[i] > s & DatenBeliebigerSW$GrössterAktionär[i]
< s)
      k <- k + 1
  }
  return(k)
}

```

```

# überprüfe Funktion: Für 182 (196 minus 14 Gesellschaften mit Stimmrechtsaktien) Gesellschaften und s = 50
sollten 15 Gesellschaften rauskommen
countfreqONE(182,50)

```

```

DataCountFall1 <- data.frame(Schwellenwerte=c(1:100))
VektorCountFall1 <- vector("numeric", 100L)
for(i in DataCountFall1$Schwellenwerte) VektorCountFall1[i] <- countfreqONE(182,i)

```

(VektorCountFall1)

```
DataCountFall1$AnzahlFälleBeherrschungsargumentErfüllt <- VektorCountFall1
```

(ii) Mit eigenen Aktien, ohne Annahme über GV-Präsenzzahl

Importiere überarbeiteten Datensatz, ohne die 14 Fälle von Stimmrechtsaktien; hier Eigene Aktien berücksichtigt

```
DatenBeliebigerSWInklEigeneAktien <- read_excel("./Datensatz/  
DataNachRegressionAusführlichOhneStimmrechtsaktienMitEigenenAktien.xlsx")
```

```
countfreqTWO <- function(n,s){  
  #Set k to 0  
  k <- 0  
  for(i in 1:n) {  
    if (DatenBeliebigerSWInklEigeneAktien$GrössterAktionär[i] +  
DatenBeliebigerSWInklEigeneAktien$ZusätzlicheStimmkraftGrössterAktionärInklEigeneAktien[i] > s &  
DatenBeliebigerSWInklEigeneAktien$GrössterAktionär[i] < s)  
      k <- k + 1  
    }  
  }  
  return(k)  
}
```

```
DataCountFall2 <- data.frame(Schwellenwerte=c(1:100))  
VektorCountFall2 <- vector("numeric", 100L)  
for(i in DataCountFall2$Schwellenwerte) VektorCountFall2[i] <- countfreqTWO(182,i)  
(VektorCountFall2)
```

```
DataCountFall2$AnzahlFälleBeherrschungsargumentErfüllt <- VektorCountFall2
```

(iii) Mit eigenen Aktien, und Annahme über GV-Präsenzzahl: Gem Zahlen der Studie von Ethos 67%: D.h. 33% nehmen nicht teil.

Importiere überarbeiteten Datensatz, ohne die 14 Fälle von Stimmrechtsaktien; hier Eigene Aktien berücksichtigt und Annahme über GV-Präsenzzahlen

```
DatenBeliebigerSWInklEigeneAktienAnnahmeGV <- read_excel("./Datensatz/  
DataNachRegressionAusführlichOhneStimmrechtsaktienMitEigenenAktienAnnahmeGV.xlsx")
```

```
countfreqTHREE <- function(n,s){  
  #Set k to 0  
  k <- 0  
  for(i in 1:n) {  
    if (DatenBeliebigerSWInklEigeneAktienAnnahmeGV$GrössterAktionär[i] +  
DatenBeliebigerSWInklEigeneAktienAnnahmeGV$ZusätzlicheStimmkraftGrössterAktionärInklEigeneAktienAnnahmeGV[i]  
> s & DatenBeliebigerSWInklEigeneAktienAnnahmeGV$GrössterAktionär[i] < s)  
      k <- k + 1  
    }  
  }  
  return(k)  
}
```

```
DataCountFall3 <- data.frame(Schwellenwerte=c(1:100))  
VektorCountFall3 <- vector("numeric", 100L)  
for(i in DataCountFall3$Schwellenwerte) VektorCountFall3[i] <- countfreqTHREE(182,i)  
(VektorCountFall3)
```

```
DataCountFall3$AnzahlFälleBeherrschungsargumentErfüllt <- VektorCountFall3
```

(iv) Daten zusammenfügen: Erstellung Diagramm

```
Schwellenwertgraphik <- data.frame(Schwellenwerte=c(1:100))  
Schwellenwertgraphik$Fall1 <- DataCountFall1$AnzahlFälleBeherrschungsargumentErfüllt  
Schwellenwertgraphik$Fall2 <- DataCountFall2$AnzahlFälleBeherrschungsargumentErfüllt  
Schwellenwertgraphik$Fall3 <- DataCountFall3$AnzahlFälleBeherrschungsargumentErfüllt
```

```
max(Schwellenwertgraphik$Fall1) # 35&36, 24  
max(Schwellenwertgraphik$Fall2) # 35, 25  
max(Schwellenwertgraphik$Fall3) # 75, 74
```

```
ggplot(Schwellenwertgraphik,aes(x = Schwellenwerte)) +
```

```

geom_line(aes(y=Fall1, color="black"), size=1.4)+
geom_line(aes(y=Fall2, color="grey"), size=1.4)+
geom_line(aes(y=Fall3, color="green"), size=1.4)+ theme_minimal()+scale_colour_grey()+
scale_x_continuous( breaks = seq(0, 100, 20))+ scale_y_continuous(breaks=seq(0, 100, 20), limits=c(0, 100))+
xlab("Notwendiger Stimmenanteil für die Beherrschung (%)")+
ylab("Anzahl Gesellschaften")+ theme(text=element_text(family="Times New Roman"), plot.title =
element_text(face = "bold"))+
geom_segment(x=50, y=0, xend=50, yend=100, size=0.2, color = "black", linetype=3)+
geom_segment(x=25, y=0, xend=25, yend=100, size=0.2, color = "black", linetype=3)+
geom_segment(x=33.3333, y=0, xend=33.3333, yend=100, size=0.2, color = "black", linetype=3)+
geom_point(aes(35,24),color="red", size=2.5)+
geom_point(aes(35,25),color="red", size=2.5)+
geom_point(aes(75,74),color="red", size=2.5)

```

8. Graphik FinfraG 135-----

```

# Importiere überarbeiteten Datensatz, ohne die 14 Fälle von Stimmrechtsaktien
FinfraGGraphik_wide <- read_excel("./Datensatz/
DataNachRegressionAusführlichOhneStimmrechtsaktien.xlsx")
colnames(FinfraGGraphik_wide)[14] <- "VRManagement"

```

```

FinfraGGraphik_wide <- subset( FinfraGGraphik_wide, select = -VRManagement ) #wird nicht berücksichtigt,
da ggf. doppelt
FinfraGGraphik_wide <-FinfraGGraphik_wide[-c(77), ] #Entferne Ausreisser

```

```

# Transformiere in langes Format
FinfraGGraphik_long <- gather(FinfraGGraphik_wide, Aktionärstyp, WertStimmkraft, Aktionär1:Aktionär9,
factor_key=TRUE)

```

```

FinfraGGraphik_long$EffektDispo <- (100*(FinfraGGraphik_long$WertStimmkraft / (100-
FinfraGGraphik_long$Dispobestand))) - FinfraGGraphik_long$WertStimmkraft
FinfraGGraphik_long$EffektDispEigeneAktien <- (100*(FinfraGGraphik_long$WertStimmkraft / (100-
FinfraGGraphik_long$Dispobestand-FinfraGGraphik_long$EigeneAktien))) -
FinfraGGraphik_long$WertStimmkraft
FinfraGGraphik_long$EffektDispEigeneAktienGVPräsenzzahl <- (100*(FinfraGGraphik_long$WertStimmkraft
/ (100-FinfraGGraphik_long$Dispobestand-FinfraGGraphik_long$EigeneAktien-33))) -
FinfraGGraphik_long$WertStimmkraft

```

```

FinfraGGraphik_long$Geschätzt <- as.factor(FinfraGGraphik_long$Geschätzt )

```

```

ggplot(FinfraGGraphik_long,aes(x=WertStimmkraft, shape=Geschätzt))+
geom_point(aes(y=EffektDispo), color="black", size=3)+
geom_point(aes(y=EffektDispEigeneAktien), color="grey80", size=3)+
geom_point(aes(y=EffektDispEigeneAktienGVPräsenzzahl), color="grey38", size=3)+theme_minimal()+
geom_segment(x=33.3333, y=0, xend=33.3333, yend=120, size=0.2, color = "black")+
geom_segment(x=33.3333, y=0, xend=0, yend=33.3333, size=0.2, color = "black")

```

Szenario 1

```

countfreqFinfraGEffektDispo <- function(n,s){
  #Set k to 0
  k <- 0
  for(i in 1:n) {
    if (FinfraGGraphik_long$WertStimmkraft[i] + FinfraGGraphik_long$EffektDispo[i] > s &
FinfraGGraphik_long$WertStimmkraft[i] < s)
      k <- k + 1
  }
  return(k)
}

```

```

countfreqFinfraGEffektDispo(1629,33.3333)

```

```

countfreqFinfraGEffektDispoGeschätzt <- function(n,s){
  #Set k to 0
  k <- 0
  for(i in 1:n) {
    if (FinfraGGraphik_long$WertStimmkraft[i] + FinfraGGraphik_long$EffektDispo[i] > s &
FinfraGGraphik_long$WertStimmkraft[i] < s & FinfraGGraphik_long$Geschätzt[i] == 1)
      k <- k + 1
  }
  return(k)
}

```

```

    }
    return(k)
}

countfreqFinfraGEffektDispoGeschätzt(1629,33.3333)

# Szenario 2

countfreqFinfraGEffektDispoEigeneAktien <- function(n,s){
  #Set k to 0
  k <- 0
  for(i in 1:n) {
    if (FinfraGGraphik_long$WertStimmkraft[i] + FinfraGGraphik_long$EffektDispEigeneAktien[i] > s &
FinfraGGraphik_long$WertStimmkraft[i] < s)
      k <- k + 1
  }
  return(k)
}

countfreqFinfraGEffektDispoEigeneAktien(1629,33.3333)

countfreqFinfraGEffektDispoEigeneAktienGeschätzt <- function(n,s){
  #Set k to 0
  k <- 0
  for(i in 1:n) {
    if (FinfraGGraphik_long$WertStimmkraft[i] + FinfraGGraphik_long$EffektDispEigeneAktien[i] > s &
FinfraGGraphik_long$WertStimmkraft[i] < s & FinfraGGraphik_long$Geschätzt[i] == 1)
      k <- k + 1
  }
  return(k)
}

countfreqFinfraGEffektDispoEigeneAktienGeschätzt(1629,33.3333)

# Szenario 3

countfreqFinfraGEffektDispoEigeneAktienGVPräsenzzahl <- function(n,s){
  #Set k to 0
  k <- 0
  for(i in 1:n) {
    if (FinfraGGraphik_long$WertStimmkraft[i] +
FinfraGGraphik_long$EffektDispEigeneAktienGVPräsenzzahl[i] > s &
FinfraGGraphik_long$WertStimmkraft[i] < s)
      k <- k + 1
  }
  return(k)
}

countfreqFinfraGEffektDispoEigeneAktienGVPräsenzzahl(1629,33.3333)

countfreqFinfraGEffektDispoEigeneAktienGVPräsenzzahlGeschätzt <- function(n,s){
  #Set k to 0
  k <- 0
  for(i in 1:n) {
    if (FinfraGGraphik_long$WertStimmkraft[i] +
FinfraGGraphik_long$EffektDispEigeneAktienGVPräsenzzahl[i] > s &
FinfraGGraphik_long$WertStimmkraft[i] < s & FinfraGGraphik_long$Geschätzt[i] == 1)
      k <- k + 1
  }
  return(k)
}

countfreqFinfraGEffektDispoEigeneAktienGVPräsenzzahlGeschätzt(1629,33.3333)

```