library(tidyverse)

library(readxl)

library(openxlsx)

options(scipen = 999)

data2021 <- read\_xlsx("./data/2021.xlsx", 1,

col\_types = c("numeric", "numeric", "text",

"text", "text", "text", "text", "text",

"text", "text", "text", "numeric",

"numeric", "numeric", "numeric",

"numeric", "text", "text", "text",

"text", "numeric", "numeric", "numeric",

"numeric", "text", "numeric", "numeric",

"numeric", "numeric", "numeric",

"text"))

data2020 <- read\_xlsx("./data/2020.xlsx", 1,

col\_types = c("numeric", "numeric", "text",

"text", "text", "text", "text", "text",

"text", "text", "text", "numeric",

"numeric", "numeric", "numeric",

"numeric", "text", "text", "text",

"text", "numeric", "numeric", "numeric",

"numeric", "text", "numeric", "numeric",

"numeric", "numeric", "numeric",

"text")) %>%

mutate(`Meno a priezvisko` = paste(`Meno bez medzier (pomocné)`, `Priezvisko bez medzier (pomocné)`))

data2019 <- read\_xlsx("./data/2019.xlsx", 1,

col\_types = c("numeric", "numeric", "text",

"text", "text", "text", "text", "text",

"text", "text", "text", "numeric",

"numeric", "numeric", "numeric",

"numeric", "text", "text", "text",

"text", "numeric", "numeric", "numeric",

"numeric", "text", "numeric", "numeric",

"numeric", "numeric", "numeric",

"text"))

data2018 <- read\_xlsx("./data/2018.xlsx", 1,

col\_types = c("numeric", "numeric", "text",

"text", "text", "text", "text", "text",

"text", "text", "text", "numeric",

"numeric", "numeric", "numeric",

"numeric", "text", "text", "text",

"text", "numeric", "numeric", "numeric",

"numeric", "text", "numeric", "numeric",

"numeric", "numeric", "numeric",

"text"))

data2017 <- read\_xlsx("./data/2017.xlsx", 1,

col\_types = c("numeric", "numeric", "text",

"text", "text", "text", "text", "text",

"text", "text", "text", "numeric",

"numeric", "numeric", "numeric",

"numeric", "text", "text", "text",

"text", "numeric", "numeric", "numeric",

"numeric", "text", "numeric", "numeric",

"numeric", "numeric", "numeric",

"text"))

data2016 <- read\_xlsx("./data/2016.xlsx", 1,

col\_types = c("numeric", "numeric", "text",

"text", "text", "text", "text", "text",

"text", "text", "text", "numeric",

"numeric", "numeric", "numeric",

"numeric", "text", "text", "text",

"text", "numeric", "numeric", "numeric",

"numeric", "text", "numeric", "numeric",

"numeric", "numeric", "numeric",

"text"))

data2015 <- read\_xlsx("./data/2015.xlsx", 1,

col\_types = c("numeric", "numeric", "text",

"text", "text", "text", "text", "text",

"text", "text", "text", "numeric",

"numeric", "numeric", "numeric",

"numeric", "text", "text", "text",

"text", "numeric", "numeric", "numeric",

"numeric", "text", "numeric", "numeric",

"numeric", "numeric", "numeric",

"text"))

data2014 <- read\_xlsx("./data/2014.xlsx", 1,

col\_types = c("numeric", "numeric", "text",

"text", "text", "text", "text", "text",

"text", "text", "text", "numeric",

"numeric", "numeric", "numeric",

"numeric", "text", "text", "text",

"text", "numeric", "numeric", "numeric",

"numeric", "text", "numeric", "numeric",

"numeric", "numeric", "numeric",

"text"))

data2013 <- read\_xlsx("./data/2013.xlsx", 1,

col\_types = c("numeric", "numeric", "text",

"text", "text", "text", "text", "text",

"text", "text", "text", "numeric",

"numeric", "numeric", "numeric",

"numeric", "text", "text", "text",

"text", "numeric", "numeric", "numeric",

"numeric", "text", "numeric", "numeric",

"numeric", "numeric", "numeric",

"text"))

data2012 <- read\_xlsx("./data/2012.xlsx", 1,

col\_types = c("numeric", "numeric", "text",

"text", "text", "text", "text", "text",

"text", "text", "text", "numeric",

"numeric", "numeric", "numeric",

"numeric", "text", "text", "text",

"text", "numeric", "numeric", "numeric",

"numeric", "text", "numeric", "numeric",

"numeric", "numeric", "numeric",

"text"))

data2011 <- read\_xlsx("./data/2011.xlsx", 1,

col\_types = c("numeric", "numeric", "text",

"text", "text", "text", "text", "text",

"text", "text", "text", "numeric",

"numeric", "numeric", "numeric",

"numeric", "text", "text", "text",

"text", "numeric", "numeric", "numeric",

"numeric", "text", "numeric", "numeric",

"numeric", "numeric", "numeric",

"text"))

dataall <- bind\_rows(data2021, data2020, data2019, data2018, data2017, data2016, data2015, data2014, data2013, data2012, data2011)

dataall <- dataall %>%

mutate(`Meno a priezvisko` = paste(`Meno bez medzier (pomocné)`, `Priezvisko bez medzier (pomocné)`)) %>%

filter(grepl("^Okresný", `Názov súdu`),

grepl("^T$|^Tk$|^Tv$|^Nt$|^Pp$|^Tp$|^Ntt$|^C$|^Cpr$|^Ca$|^Csp$|^Cr$|^Ccud$|^Cb$|^CbPv$|^CbR$|^CbBu$|^CbHs$|^Cbi$|^CbVO$|^S$|^P$|^PPOm$|^Pc$|^Ps$|^Po$|^Pu$|^D$|^K$|^R$|^NcKR$|^Nsre$|^Zm$|^CbZm$",

Agenda))

dataall <- left\_join(dataall, dataID,

by = "Meno a priezvisko") %>%

select(-Meno, -Priezvisko) %>%

mutate(ID.x = ID.y) %>%

select(-ID.y) %>%

rename(ID = ID.x)

sum(is.na(dataall$ID))

registre1 <- dataall %>%

filter(grepl("^O$", Typ)) %>%

group\_by(Agenda, Rok) %>%

summarize(Rozhodnutia\_vyssieho\_sudu = sum(`Počet1`, na.rm = T) + sum(`Počet2`, na.rm = T) + sum(`Počet3`, na.rm = T),

Potvrdene = sum(`Počet1`, na.rm = T),

Zmenene = sum(`Počet2`, na.rm = T),

Zrusene = sum(`Počet3`, na.rm = T),

Odmietnute = sum(`Počet4`, na.rm = T)) %>%

mutate(Test = if\_else(Rozhodnutia\_vyssieho\_sudu >= 100, 1, 0)) %>%

ungroup() %>%

group\_by(Agenda) %>%

mutate(Pocet = n()) %>%

filter(sum(Test) >= 0.5\*Pocet)

registre1 <- unique(registre1$Agenda)

registre2 <- dataall %>%

filter(grepl("^O$", Typ)) %>%

group\_by(Agenda, Rok) %>%

summarize(Rozhodnutia\_vyssieho\_sudu = sum(`Počet1`, na.rm = T) + sum(`Počet2`, na.rm = T) + sum(`Počet3`, na.rm = T),

Potvrdene = sum(`Počet1`, na.rm = T),

Zmenene = sum(`Počet2`, na.rm = T),

Zrusene = sum(`Počet3`, na.rm = T),

Odmietnute = sum(`Počet4`, na.rm = T),

Kvalita = Potvrdene/Rozhodnutia\_vyssieho\_sudu) %>%

mutate(Test = if\_else(Kvalita >= 0.8, 1, 0)) %>%

ungroup() %>%

group\_by(Agenda) %>%

mutate(Pocet = n()) %>%

filter(sum(Test) >= 0.5\*Pocet)

registre2 <- unique(registre2$Agenda)

registre <- unique(c(registre1, registre2))

registre

# Dopocitanie medianu a MAD

kvalita <- dataall %>%

filter(grepl("^O$", Typ),

grepl("^C$|^Cb$|^Cbi$|^CbZm$|^Cpr$|^Csp$|^D$|^Nt$|^P$|^Pc$|^T$|^Tp$|^Pp$|^PPOm$|^Pu$|^Tv$", Agenda)) %>%

group\_by(ID) %>%

summarize(Rozhodnutia\_vyssieho\_sudu = sum(`Počet1`, na.rm = T) + sum(`Počet2`, na.rm = T) + sum(`Počet3`, na.rm = T),

Potvrdene = sum(`Počet1`, na.rm = T),

Zmenene = sum(`Počet2`, na.rm = T),

Zrusene = sum(`Počet3`, na.rm = T),

Odmietnute = sum(`Počet4`, na.rm = T)) %>%

filter(Rozhodnutia\_vyssieho\_sudu > 20) %>%

ungroup() %>%

group\_by(ID) %>%

mutate(Kvalita = Potvrdene/(Potvrdene + Zmenene + Zrusene)) %>%

ungroup() %>%

mutate(Median = median(Kvalita),

MAD = mad(Kvalita),

DeviationScore = (Kvalita - Median) / MAD)

# Distribucia

ggplot(kvalita, aes(sample = Kvalita)) +

geom\_qq() + geom\_qq\_line()

ggplot(kvalita, aes(x = Kvalita)) +

geom\_histogram()

ggplot(kvalita, aes( y = Kvalita)) +

geom\_boxplot()

# Filter vzdialenych pozorovani + dopocitanie priemeru a SD

kvalitanooutliers <- kvalita %>%

mutate(Outlier = if\_else(DeviationScore <= 2 & DeviationScore >= -2, 0, 1)) %>%

filter(Outlier == 0)%>%

mutate(MeanCisty = mean(Kvalita),

SDCisty = sd(Kvalita))

# Distribucia bez vzdialenych pozorovani

ggplot(kvalitanooutliers, aes(sample = Kvalita)) +

geom\_qq() + geom\_qq\_line()

ggplot(kvalitanooutliers, aes(x = Kvalita)) +

geom\_histogram()

ggplot(kvalitanooutliers, aes(y = Kvalita)) +

geom\_boxplot()

# Vypocet z score a bodov

konstanta <- -1.96

konstanta2 <- 1.96

kvalitabody <- kvalita %>%

mutate(z\_score = (Kvalita - first(kvalitanooutliers$MeanCisty))/first(kvalitanooutliers$SDCisty),

Body = case\_when(z\_score < konstanta ~ 0,

z\_score >= konstanta & z\_score < konstanta + ((2\*konstanta2)/9) ~ 1,

z\_score >= konstanta + ((2\*konstanta2)/9) & z\_score < konstanta + 2\*((2\*konstanta2)/9) ~ 2,

z\_score >= konstanta + 2\*((2\*konstanta2)/9) & z\_score < konstanta + 3\*((2\*konstanta2)/9) ~ 3,

z\_score >= konstanta + 3\*((2\*konstanta2)/9) & z\_score < konstanta + 4\*((2\*konstanta2)/9) ~ 4,

z\_score >= konstanta + 4\*((2\*konstanta2)/9) & z\_score < konstanta + 5\*((2\*konstanta2)/9) ~ 5,

z\_score >= konstanta + 5\*((2\*konstanta2)/9) & z\_score < konstanta + 6\*((2\*konstanta2)/9) ~ 6,

z\_score >= konstanta + 6\*((2\*konstanta2)/9) & z\_score < konstanta + 7\*((2\*konstanta2)/9) ~ 7,

z\_score >= konstanta + 7\*((2\*konstanta2)/9) & z\_score < konstanta + 8\*((2\*konstanta2)/9) ~ 8,

z\_score >= konstanta + 8\*((2\*konstanta2)/9) & z\_score < konstanta + 9\*((2\*konstanta2)/9) ~ 9,

z\_score >= konstanta2 ~ 10))

kvalitabody <- left\_join(kvalitabody, dataID2,

by = "ID") %>%

relocate(`Meno a priezvisko`, Meno, Priezvisko, .after = ID)

ggplot(kvalitabody, aes(x = Body)) +

geom\_bar()