library(tidyverse)

library(readxl)

library(openxlsx)

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"))

data2021 <- left\_join(data2021, dataID,

by = "Meno a priezvisko") %>%

select(-Meno, -Priezvisko) %>%

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

select(-ID.y) %>%

rename(ID = ID.x)

# Filter a vypocet potrebnych veci

efektivita2021 <- data2021 %>%

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

grepl("^R$|^V$|^N$", Typ),

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)) %>%

group\_by(ID, Rok, Typ) %>%

summarize(Pocet = sum(Počet1, na.rm = T)) %>%

ungroup()

efektivita2021 <- pivot\_wider(efektivita2021,

id\_cols = c(ID, Rok),

names\_from = Typ,

values\_from = Pocet) %>%

rename("Rozhodnute" = "R",

"Nevybavene" = "N",

"Vybavene" = "V")

restancne2021 <- data2021 %>%

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

grepl("^N$", Typ),

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)) %>%

group\_by(ID, Rok) %>%

summarize(Restancne = sum(Počet2, na.rm = T)) %>%

ungroup() %>%

select(ID, Restancne)

napad2021 <- data2021 %>%

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

grepl("^P$", Typ),

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)) %>%

group\_by(ID, Rok) %>%

summarize(Pridelene = sum(Počet1, na.rm = T),

Prerozdelene = sum(Počet2, na.rm = T),

Napad = if\_else(Prerozdelene < 0, Pridelene, Pridelene + Prerozdelene),

Napad2 = Pridelene + Prerozdelene,

TestPrerozdelene = Pridelene - abs(Prerozdelene)) %>%

ungroup() %>%

select(ID, Pridelene, Prerozdelene, Napad, Napad2, TestPrerozdelene)

# Spojenie ciastkovych datasetov a filter tych sudcov, co rozhodli aspon 50 veci

efektivita2021 <- left\_join(efektivita2021, restancne2021,

by = "ID")

efektivita2021 <- left\_join(efektivita2021, napad2021,

by = "ID")

efektivita2021 <- efektivita2021 %>%

filter(Rozhodnute > 50)

# Dopocitanie medianu a MAD

efektivita\_12021 <- efektivita2021 %>%

mutate(Efektivita1 = Rozhodnute/Napad2) %>%

filter(Napad2 > 0)

efektivita\_22021 <- efektivita2021 %>%

mutate(Efektivita2 = (Nevybavene/Vybavene)\*365) %>%

filter(Nevybavene >= 0,

Napad2 > 0)

efektivita\_32021 <- efektivita2021 %>%

mutate(Efektivita3 = Restancne/Nevybavene) %>%

filter(Napad2 > 0,

!is.nan(Efektivita3),

!is.infinite(Efektivita3))

efektivita\_full2021 <- full\_join(efektivita\_12021, efektivita\_22021,

by = c("ID", "Rok", "Nevybavene", "Rozhodnute", "Vybavene", "Restancne", "Pridelene", "Prerozdelene", "Napad", "Napad2"))

efektivita\_full2021 <- full\_join(efektivita\_full2021, efektivita\_32021,

by = c("ID", "Rok", "Nevybavene", "Rozhodnute", "Vybavene", "Restancne", "Pridelene", "Prerozdelene", "Napad","Napad2"))

efektivita\_full2021 <- efektivita\_full2021 %>%

mutate(Median1 = median(Efektivita1, na.rm = T),

MAD1 = mad(Efektivita1, na.rm = T),

DeviationScore1 = (Efektivita1 - Median1) / MAD1,

Median2 = median(Efektivita2, na.rm = T),

MAD2 = mad(Efektivita2, na.rm = T),

DeviationScore2 = (Efektivita2 - Median2) / MAD2,

Median3 = median(Efektivita3, na.rm = T),

MAD3 = mad(Efektivita3, na.rm = T),

DeviationScore3 = (Efektivita3 - Median3) / MAD3)

# Distribucia 1

ggplot(efektivita\_full2021, aes(sample = Efektivita1)) +

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

ggplot(efektivita\_full2021, aes(x = Efektivita1)) +

geom\_histogram()

ggplot(efektivita\_full2021, aes( y = Efektivita1)) +

geom\_boxplot()

# Distribucia 2

ggplot(efektivita\_full2021 , aes(sample = Efektivita2)) +

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

ggplot(efektivita\_full2021, aes(x = Efektivita2)) +

geom\_histogram()

ggplot(efektivita\_full2021, aes( y = Efektivita2)) +

geom\_boxplot()

# Distribucia 3

ggplot(efektivita\_full2021, aes(sample = Efektivita3)) +

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

ggplot(efektivita\_full2021, aes(x = Efektivita3)) +

geom\_histogram()

ggplot(efektivita\_full2021, aes( y = Efektivita3)) +

geom\_boxplot()

# Efektivita bez vzdialenych pozorovani

efektivita12021nooutliers <- efektivita\_full2021 %>%

mutate(Outlier = if\_else(DeviationScore1 <= 2 & DeviationScore1 >= -2, 0, 1)) %>%

filter(Outlier == 0) %>%

mutate(MeanCisty1 = mean(Efektivita1),

SDCisty1 = sd(Efektivita1))

efektivita22021nooutliers <- efektivita\_full2021 %>%

mutate(Outlier = if\_else(DeviationScore2 <= 2 & DeviationScore2 >= -2, 0, 1)) %>%

filter(Outlier == 0) %>%

mutate(MeanCisty2 = mean(Efektivita2),

SDCisty2 = sd(Efektivita2))

efektivita32021nooutliers <- efektivita\_full2021 %>%

mutate(Outlier = if\_else(DeviationScore3 <= 2 & DeviationScore3 >= -2, 0, 1)) %>%

filter(Outlier == 0) %>%

mutate(MeanCisty3 = mean(Efektivita3, na.rm = T),

SDCisty3 = sd(Efektivita3, na.rm = T))

# Distribucia 1 bez vzdialenych pozorovani

ggplot(efektivita12021nooutliers, aes(sample = Efektivita1)) +

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

ggplot(efektivita12021nooutliers, aes(x = Efektivita1)) +

geom\_histogram()

ggplot(efektivita12021nooutliers, aes( y = Efektivita1)) +

geom\_boxplot()

# Distribucia 2 bez vzdialenych pozorovani

ggplot(efektivita22021nooutliers, aes(sample = Efektivita2)) +

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

ggplot(efektivita22021nooutliers, aes(x = Efektivita2)) +

geom\_histogram()

ggplot(efektivita22021nooutliers, aes( y = Efektivita2)) +

geom\_boxplot()

# Distribucia 3 bez vzdialenych pozorovani

ggplot(efektivita32021nooutliers, aes(sample = Efektivita3)) +

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

ggplot(efektivita32021nooutliers, aes(x = Efektivita3)) +

geom\_histogram()

ggplot(efektivita32021nooutliers, aes( y = Efektivita3)) +

geom\_boxplot()

# Vypocet z score a bodov

konstanta <- -1.96

konstanta2 <- 1.96

efektivita2021body <- efektivita\_full2021 %>%

mutate(z\_score1 = (Efektivita1 - first(efektivita12021nooutliers$MeanCisty1))/first(efektivita12021nooutliers$SDCisty1),

Body1 = case\_when(z\_score1 < konstanta ~ 0,

z\_score1 >= konstanta & z\_score1 < konstanta + ((2\*konstanta2)/9) ~ 1,

z\_score1 >= konstanta + ((2\*konstanta2)/9) & z\_score1 < konstanta + 2\*((2\*konstanta2)/9) ~ 2,

z\_score1 >= konstanta + 2\*((2\*konstanta2)/9) & z\_score1 < konstanta + 3\*((2\*konstanta2)/9) ~ 3,

z\_score1 >= konstanta + 3\*((2\*konstanta2)/9) & z\_score1 < konstanta + 4\*((2\*konstanta2)/9) ~ 4,

z\_score1 >= konstanta + 4\*((2\*konstanta2)/9) & z\_score1 < konstanta + 5\*((2\*konstanta2)/9) ~ 5,

z\_score1 >= konstanta + 5\*((2\*konstanta2)/9) & z\_score1 < konstanta + 6\*((2\*konstanta2)/9) ~ 6,

z\_score1 >= konstanta + 6\*((2\*konstanta2)/9) & z\_score1 < konstanta + 7\*((2\*konstanta2)/9) ~ 7,

z\_score1 >= konstanta + 7\*((2\*konstanta2)/9) & z\_score1 < konstanta + 8\*((2\*konstanta2)/9) ~ 8,

z\_score1 >= konstanta + 8\*((2\*konstanta2)/9) & z\_score1 < konstanta + 9\*((2\*konstanta2)/9) ~ 9,

z\_score1 >= konstanta2 ~ 10),

z\_score2 = (Efektivita2 - first(efektivita22021nooutliers$MeanCisty2))/first(efektivita22021nooutliers$SDCisty2),

Body2 = case\_when(z\_score2 < konstanta ~ 10,

z\_score2 >= konstanta & z\_score2 < konstanta + ((2\*konstanta2)/9) ~ 9,

z\_score2 >= konstanta + ((2\*konstanta2)/9) & z\_score2 < konstanta + 2\*((2\*konstanta2)/9) ~ 8,

z\_score2 >= konstanta + 2\*((2\*konstanta2)/9) & z\_score2 < konstanta + 3\*((2\*konstanta2)/9) ~ 7,

z\_score2 >= konstanta + 3\*((2\*konstanta2)/9) & z\_score2 < konstanta + 4\*((2\*konstanta2)/9) ~ 6,

z\_score2 >= konstanta + 4\*((2\*konstanta2)/9) & z\_score2 < konstanta + 5\*((2\*konstanta2)/9) ~ 5,

z\_score2 >= konstanta + 5\*((2\*konstanta2)/9) & z\_score2 < konstanta + 6\*((2\*konstanta2)/9) ~ 4,

z\_score2 >= konstanta + 6\*((2\*konstanta2)/9) & z\_score2 < konstanta + 7\*((2\*konstanta2)/9) ~ 3,

z\_score2 >= konstanta + 7\*((2\*konstanta2)/9) & z\_score2 < konstanta + 8\*((2\*konstanta2)/9) ~ 2,

z\_score2 >= konstanta + 8\*((2\*konstanta2)/9) & z\_score2 < konstanta + 9\*((2\*konstanta2)/9) ~ 1,

z\_score2 >= konstanta2 ~ 0),

z\_score3 = (Efektivita3 - first(efektivita32021nooutliers$MeanCisty3))/first(efektivita32021nooutliers$SDCisty3),

Body3 = case\_when(z\_score3 < konstanta ~ 10,

z\_score3 >= konstanta & z\_score3 < konstanta + ((2\*konstanta2)/9) ~ 9,

z\_score3 >= konstanta + ((2\*konstanta2)/9) & z\_score3 < konstanta + 2\*((2\*konstanta2)/9) ~ 8,

z\_score3 >= konstanta + 2\*((2\*konstanta2)/9) & z\_score3 < konstanta + 3\*((2\*konstanta2)/9) ~ 7,

z\_score3 >= konstanta + 3\*((2\*konstanta2)/9) & z\_score3 < konstanta + 4\*((2\*konstanta2)/9) ~ 6,

z\_score3 >= konstanta + 4\*((2\*konstanta2)/9) & z\_score3 < konstanta + 5\*((2\*konstanta2)/9) ~ 5,

z\_score3 >= konstanta + 5\*((2\*konstanta2)/9) & z\_score3 < konstanta + 6\*((2\*konstanta2)/9) ~ 4,

z\_score3 >= konstanta + 6\*((2\*konstanta2)/9) & z\_score3 < konstanta + 7\*((2\*konstanta2)/9) ~ 3,

z\_score3 >= konstanta + 7\*((2\*konstanta2)/9) & z\_score3 < konstanta + 8\*((2\*konstanta2)/9) ~ 2,

z\_score3 >= konstanta + 8\*((2\*konstanta2)/9) & z\_score3 < konstanta + 9\*((2\*konstanta2)/9) ~ 1,

z\_score3 >= konstanta2 ~ 0)) %>%

select(ID, Rok, Napad, Napad2, Rozhodnute, Vybavene, Nevybavene, Restancne, Pridelene, Prerozdelene, Efektivita1, Efektivita2, Efektivita3,

Median1, MAD1, DeviationScore1, Median2, MAD2, DeviationScore2, Median3, MAD3, DeviationScore3,

z\_score1, z\_score2, z\_score3, Body1, Body2, Body3) %>%

rowwise() %>%

mutate(EfektivitaSpolu = sum(Body1, Body2, Body3, na.rm = TRUE))

efektivita2021body <- left\_join(efektivita2021body, dataID2,

by = "ID") %>%

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

ggplot(efektivita2021body, aes(x = Body1)) +

geom\_bar()

ggplot(efektivita2021body, aes(x = Body2)) +

geom\_bar()

ggplot(efektivita2021body, aes(x = Body3)) +

geom\_bar()