## Dataskript pro program R

```
# načteme potřebné knihovny ####
library(forecast)
library(tseries)

# read data from file ####
data <- readRDS("./data/data.rds")
metadata <- read.csv2(file="./data/metadata.csv")

data <- data[c(1:32), ]

tsData <- ts(data, frequency = 4, start=2015)

# time-series analysis ####
idx_flights <- c(4:10)

for (i_f in idx_flights) {
  meta <- find_metadata(colnames(tsData)[i_f])
  ts <- tsData[, i_f]

  ## decomposition ####
  decomposedRes <- decompose(ts, type="additive") #"additive" for additive components
  plot(decomposedRes) # see plot below
  stlRes <- stl(ts, s.window = "periodic")

  ## autocorrelation ####
  acfRes <- acf(ts) # autocorrelation
  pacfRes <- pacf(ts)  # partial autocorrelation

  ## stationary test ####
  #Use Augmented Dickey-Fuller Test (adf test). A p-Value of less than 0.05 in adf.test()
indicates that it is stationary.
  adf_test <- adf.test(ts) # p-value < 0.05 indicates the TS is stationary

  ## arima ####
  autoarima <- auto.arima(ts, trace=F, stepwise = FALSE)
  plot(forecast(autoarima, h=8))
  write.csv2(forecast(autoarima, h=8), file=paste0("./results/ARIMA_forecast_", meta$Short,
".csv", sep=""))


## HDP a další

data <- readRDS("./data/data.rds")
metadata <- read.csv2(file="./data/metadata.csv")
```

```
tsData <- ts(data, frequency = 4, start=2015)

idx_flights <- c(11,13:14)

for (i_f in idx_flights) {
  meta <- find_metadata(colnames(tsData)[i_f])
  ts <- tsData[, i_f]

  ## dekomposition ####
  decomposedRes <- decompose(ts, type="additive") #mult,  use type = "additive" for
additive components
  plot(decomposedRes) # see plot below
  stlRes <- stl(ts, s.window = "periodic")

  ## autocorrelation ####
  acfRes <- acf(ts) # autocorrelation
  pacfRes <- pacf(ts)  # partial autocorrelation

  ## stationary test ####
  #Use Augmented Dickey-Fuller Test (adf test). A p-Value of less than 0.05 in adf.test()
indicates that it is stationary.
  adf_test <- adf.test(ts) # p-value < 0.05 indicates the TS is stationary

  ## arima ####
  autoarima <- auto.arima(ts, trace=F, stepwise = F)
  plot(forecast(autoarima, h=5))  # do konce roku 2024
  write.csv2(forecast(autoarima, h=5), file=paste0("./results/ARIMA_forecast_", meta$Short,
".csv", sep=""))



# GLM ####
idx_flights <- c(4:9)
idx_covariates <- c(1, 3, 11, 13, 14)

data <- readRDS("./data/data.rds")
metadata <- read.csv2(file="./data/metadata.csv")

data <- data[c(1:32), ]

tsData <- ts(data, frequency = 4, start=2015)

for (i_f in idx_flights) {
  meta <- find_metadata(colnames(tsData)[i_f])
  ts <- tsData[, c(i_f, idx_covariates)]
  glm1 <- glm(ts[,1] ~ ts[,2:6])
  plot(glm1)
```

```
# (S)ARIMA + GLM ####
data <- readRDS("./data/data.rds")
metadata <- read.csv2(file="./data/metadata.csv")

data_all <- data
data <- data[c(1:32), ]

tsData <- ts(data, frequency = 4, start=2015)

idx_flights <- c(4:9)
idx_covariates <- c(1, 3, 11, 13, 14)

for (i_f in idx_flights) {
  meta <- find_metadata(colnames(tsData)[i_f])
  ts_data <- tsData[, i_f]

  # Model ARIMA
  arima_model <- auto.arima(ts_data)

  # Odstranění trendu a sezónních efektů pomocí ARIMA
  residuals_arima <- residuals(arima_model)

  # Předpověď s GLM modelem
  autoarima <- auto.arima(ts_data, trace=F, stepwise = F)
  arima_forecast <- forecast(autoarima, h=8)
  autoarima_hdp <- auto.arima(data_all$HDP, trace=F, stepwise = F)
  forecast_hdp <- forecast(autoarima_hdp, h=8)
  autoarima_cpi <- auto.arima(data_all$CPI, trace=F, stepwise = F)
  forecast_cpi <- forecast(autoarima_cpi, h=8)
  autoarima_saldo <- auto.arima(data_all$Saldo, trace=F, stepwise = F)
  forecast_saldo <- forecast(autoarima_saldo, h=8)

  ts <- tsData[, c(i_f, idx_covariates)]
  ddata <- data.frame(
    Time=c(1:32),
    Quarter=data$Quarter,
    HDP=data$HDP,
    CPI=data$CPI,
    Saldo=data$Saldo
  )
  glm0 <- glm(data[, i_f] ~ Time + Quarter + HDP + CPI + Saldo, data=ddata)
  glm1 <- glm(residuals_arima ~ Time + Quarter + HDP + CPI + Saldo, data=ddata)
  print_verbatim(summary(glm1))

  newdata <- data.frame(
```

```
  Time=c(33:40),
  Quarter=rep(c(1:4), times=2),
  HDP=forecast_hdp$mean,
  CPI=forecast_cpi$mean,
  Saldo=forecast_saldo$mean)

 combined_forecast <- predict(glm0,
                      newdata=newdata,
                      type = "response", se.fit = TRUE)

 write.csv2(combined_forecast, file=paste0("./results/combined_forecast_", meta$Short,
".csv", sep=""))

 # Vykreslení výsledků
 plot(data[, i_f],
     col = "blue",
     xlim = range(c(1:40)),
     ylim = range(c(ts_data, combined_forecast)),
     main = "Kombinace ARIMA a GLM",
     t="l"
 )
 lines(x=c(33:40), as.vector(arima_forecast$mean), col = "red", lty = 2)
 lines(glm0$fit, col = "green", lty = 2)
 lines(x=c(33:40), y=combined_forecast$fit, col = "purple", lty = 1)
 legend("bottomleft", legend = c("Časová řada", "ARIMA", "GLM", "Kombinace"), col =
c("blue", "red", "green", "purple"), lty = c(1, 2, 2, 1))
}
```