

# 2016-10-08

*Paulius Kostickis ir Manvydas Sokolovas*

*10/8/2016*

Šią savaitę galutinai sutvarkėme duomenis, sudarėme “dataframe”. Pasirinkome naudoti duomenis nuo 2002 metų, nes tuomet litas buvo susietas su euru. Šiame kode yra nesutvarkytų duomenų ištraukimas ir sudėjimas į csv formato lentelę. Paskaitoje pateiksime sutvarkytų duomenų vizualizaciją.

Reikiamos bibliotekos:

```
library("quantmod")
library("forecast")
library("xts")
library("dplyr")
library("gridExtra")
```

Duomenų ištraukimo funkcija:

```
istraukimas <- function(salis,folderio_pav,failo_pav,indikatorius,versija){
  setwd(folderio_pav)
  data=read.csv(failo_pav)

  data = data[data$S_ADJ=="Unadjusted data (i.e. neither seasonally adjusted nor calendar adjusted data",]
  data = data[data$GEO == salis,]
  setwd('..')
  if(versija == 1){
    data = data[data$UNIT=="Index, 2010=100",]
    return(select(data,TIME,Value))
  }
  if(versija == 2){
    data = data[data$INDIC == indikatorius,]
    return(select(data,TIME,Value))
  }

  else{print("klaida del versiju pasirinkimo")}}
}
```

Nuskaitomi OMX Vilnius duomenys:

```
setwd("...")
setwd("duomenys")
omx=read.csv("omx.csv")
duom=select(omx,OMX.raw,OMXV.Value)
colnames(duom) <- c("Data", "Verte")

ts <- xts(duom$Verte, as.Date(duom$Data, "%m/%d/%Y"))
ts_m = apply.monthly(ts, last)
colnames(ts_m) <- c("OMX")
```

Gamintojų kainos:

```

setwd("..")
setwd("duomenys")
salis="Lithuania"
indikatorius = NA # kol kas indikatoriaus nebus
folderio_pav="Domestic producer prices - total industry (excluding construction)"
failo_pav="sts_inppd_m_1_Data.csv"
gamintoju_kainos=istraukimas(salis,folderio_pav,failo_pav,indikatorius,1)
colnames(gamintoju_kainos) <- c("Data", "gam_kain")

```

Industrinė produkcija:

```

setwd("..")
setwd("duomenys")
salis="Lithuania"
indikatorius = NA # kol kas indikatoriaus nebus
folderio_pav="Production in industry - monthly data"
failo_pav="sts_inpr_m_1_Data.csv"
industrine_produkcija=istraukimas(salis,folderio_pav,failo_pav,indikatorius,1)
colnames(industrine_produkcija) <- c("Data", "industr_prod")

```

Ilgalaikio vartojimo prekės:

```

setwd("..")
setwd("duomenys")
salis="Lithuania"
indikatorius = NA # kol kas indikatoriaus nebus
folderio_pav="Turnover in industry - consumer durables"
failo_pav="sts_intv_m_1_Data.csv"
ilgalaikio_vartojimo_prekes=istraukimas(salis,folderio_pav,failo_pav,indikatorius,1)
colnames(ilgalaikio_vartojimo_prekes) <- c("Data", "ilg_vart_prek")

```

Apyvarta:

```

setwd("..")
setwd("duomenys")
folderio_pav="Turnover and volume of sales in wholesale and retail trade - monthly data"
failo_pav="sts_trtu_m_1_Data.csv"
apyvarta=istraukimas(salis,folderio_pav,failo_pav,indikatorius,1) #didmeninės ir mažmeninės
colnames(apyvarta) <- c("Data", "Apyvarta")

```

Statybų pasitikėjimas:

```

setwd("..")
setwd("duomenys")
indikatorius="Construction confidence indicator"
folderio_pav="Sentiment indicators - monthly data"
failo_pav="ei_bssi_m_r2_1_Data.csv"
statybu_pasitikejimas=istraukimas(salis,folderio_pav,failo_pav,indikatorius,2)
colnames(statybu_pasitikejimas) <- c("Data", "statyb_pasitik")

```

Paslaugų pasitikėjimas:

```

setwd("..")
setwd("duomenys")
indikatorius="Services Confidence Indicator"
folderio_pav="Sentiment indicators - monthly data"
failo_pav="ei_bssi_m_r2_1_Data.csv"
paslaugu_pasitikejimas=istraukimas(salis,folderio_pav,failo_pav,indikatorius,2)
colnames(paslaugu_pasitikejimas) <- c("Data", "paslaugu_pasitik")

```

Vartotojų pasitikėjimas:

```

setwd("..")
setwd("duomenys")
indikatorius="Consumer confidence indicator"
folderio_pav="Sentiment indicators - monthly data"
failo_pav="ei_bssi_m_r2_1_Data.csv"
vartotoju_pasitikejimas=istraukimas(salis,folderio_pav,failo_pav,indikatorius,2)
colnames(vartotoju_pasitikejimas) <- c("Data", "vartot_pasitik")

```

Mažmeninis pasitikėjimas:

```

setwd("..")
setwd("duomenys")
indikatorius="Retail confidence indicator"
folderio_pav="Sentiment indicators - monthly data"
failo_pav="ei_bssi_m_r2_1_Data.csv"
mazmeninis_pasitikejimas=istraukimas(salis,folderio_pav,failo_pav,indikatorius,2) #mažmeninės prekybos
colnames(mazmeninis_pasitikejimas) <- c("Data", "mazm_pasitik")

```

Pramonės pasitikėjimas:

```

setwd("..")
setwd("duomenys")
indikatorius="Industrial confidence indicator"
folderio_pav="Sentiment indicators - monthly data"
failo_pav="ei_bssi_m_r2_1_Data.csv"
pramones_pasitikejimas=istraukimas(salis,folderio_pav,failo_pav,indikatorius,2)
colnames(pramones_pasitikejimas) <- c("Data", "pram_pasitik")

```

Verslo aktyvumas:

```

setwd("..")
setwd("duomenys")
indikatorius="Business activity (sales) development over the past 3 months"
folderio_pav="Mazmenine prekyba"
failo_pav="ei_bsrt_m_r2_1_Data.csv"
verslo_aktyvumas=istraukimas(salis,folderio_pav,failo_pav,indikatorius,2)
colnames(verslo_aktyvumas) <- c("Data", "versl_aktyv")

```

Turimos akcijos:

```

setwd("..")
setwd("duomenys")
indikatorius="Volume of stocks currently hold"
folderio_pav="Mazmenine prekyba"
failo_pav="ei_bsrt_m_r2_1_Data.csv"
turimos_akcijos=istraukimas(salis,folderio_pav,failo_pav,indikatorius,2)
colnames(turimos_akcijos) <- c("Data", "turimos_akc")

```

Užsakymų lūkesčiai:

```

setwd("..")
setwd("duomenys")
indikatorius="Expectations of the number of orders placed with suppliers over the next 3 months"
folderio_pav="Mazmenine prekyba"
failo_pav="ei_bsrt_m_r2_1_Data.csv"
uzsakymu_lukesciai=istraukimas(salis,folderio_pav,failo_pav,indikatorius,2)
colnames(uzsakymu_lukesciai) <- c("Data", "uzsak_luk")

```

Darbo Lygio lūkesčiai:

```

setwd("..")
setwd("duomenys")
indikatorius="Employment expectations over the next 3 months"
folderio_pav="Mazmenine prekyba"
failo_pav="ei_bsrt_m_r2_1_Data.csv"
darbo_lygio_lukesciai=istraukimas(salis,folderio_pav,failo_pav,indikatorius,2)
colnames(darbo_lygio_lukesciai) <- c("Data", "darb_lyg_luk")

```

Nedarbas:

```

setwd("..")
setwd("duomenys/nedarbas")
data = read.csv("ei_lmhu_m_1_Data.csv")
data = data[data$GEO == salis,]
nedarbas = select(data,TIME,Value)
colnames(nedarbas) <- c("Data", "nedarbas")
setwd("..")

```

Infliacija:

```

setwd("..")
setwd("duomenys/infliacija")
data = read.csv("prc_hicp_mv12r_1_Data.csv")
data = data[data$GEO == salis,]
infliacija = select(data,TIME,Value)
colnames(infliacija) <- c("Data", "Infliacija")
setwd("..")

```

S&P 500:

```

setwd("..")
setwd("duomenys/S&P 500")
data = read.csv("YAHOO-INDEX_GSPC.csv")
SP500=select(data,Date,Adjusted.Close)
colnames(SP500) <- c("Data", "SP500")
setwd("..")

```

S&P 350 Europe:

```

setwd("..")
setwd("duomenys/S&P 350 Europe")
data = read.csv("GOOG-NYSEARCA_IEV.csv")
SP350=select(data,Date,Close)
colnames(SP350) <- c("Data", "SP350")

```

Sujungiami duomenys į vieną lentelę:

```

data=full_join(SP350,SP500,by="Data")

```

```

## Warning in full_join_impl(x, y, by$x, by$y, suffix$x, suffix$y): joining
## factors with different levels, coercing to character vector

```

```

data = data %>% arrange(-row_number())

```

```

dataset=Reduce(function(x, y) merge(x, y, all=TRUE), list(apyvarta, darbo_lygio_lukescai,nedarbas, inf

```

```

Data=dataset[,1]
dataset=dataset[,-1]
n=length(dataset[,1])
datafin=cbind(data[1:n,c(2,3)], dataset)
OMX=data.frame(ts_m)
colnames(OMX)=c("OMX")
datafin=cbind( OMX,datafin)
rownames(datafin) <- NULL
rawdata=cbind( Data,datafin)
write.csv(rawdata, file = "rawdata.csv",row.names=FALSE)

```