<https://indexes.nasdaqomx.com/Home/Global>

R code:

#Load packages

library(RCurl)

library(plyr)

library(stringr)

Nasdaq <- getURL("[https://indexes.nasdaqomx.com//docs//AppendixNQGIIndexes.csv](https://indexes.nasdaqomx.com/docs/AppendixNQGIIndexes.csv)") # URL for all nasdaq indexes

Nasdaq <- as.data.frame(read.csv(textConnection(Nasdaq), header = TRUE, stringsAsFactors = FALSE)) # trasforming to dataframe

Nasdaq <- subset(Nasdaq, CURRENCY == "EUR") # concentrate only on indexes in EUR

Nasdaq <- subset(Nasdaq, select = c(Price.Return.Symbol,[INDEX.NAME](http://index.name/))) # leaving only needed columns

Nasdaq <- Nasdaq[grep("[0-9]{4}", Nasdaq$Price.Return.Symbol),] # leaving only rows, which contain 4-digit code, i.e specific sector

Nasdaq <- Nasdaq[!grepl("LMEUR", Nasdaq$Price.Return.Symbol),] # leaving only rows, which don't have "LM" in the name, i.e. no large, med small capitalization

Nasdaq <- Nasdaq[!grepl(c(" Cap ","BRIC", "Europe"), Nasdaq$[INDEX.NAME](http://index.name/)),] # leaving only rows, which don't have " Cap " in the name, i.e. no indication for capitalization (only relevant for US indexes). We need to exclude "BRIC" and "Europe" as well, as it would be difficult later

Country\_Code <- substr(Nasdaq$Price.Return.Symbol, 1, 4) # exctract first 4 symbols of the index code: this is used to define countries

Sector\_Code <- str\_extract(Nasdaq$Price.Return.Symbol, "[[:digit:]]{4}") # exctract a 4-digit code

Big\_Sector <- paste(substr(Sector\_Code, 1, 1), # for creation of separate column for the broad sector, we need to extract 1st digit from sector...

"000", Sep="") #... and add "000" to it

Big\_Sector <- gsub(" ", "", Big\_Sector) # somehow it creates it with space => we need to remove spaces

Big\_Sector <- gsub("0000", "0001", Big\_Sector) # Oil & Gas has code 0001

Nasdaq <- cbind(Nasdaq, Country\_Code, Big\_Sector, Sector\_Code) # now we could add 3 columns to the main database

Country\_Code <- c("NQJP", "NQGB", "NQCA", "NQAT", "NQBE", "NQFI", "NQFR", "NQDK", "NQDE", "NQGR", "NQIE", "NQIT", "NQNL", "NQNO", "NQES", "NQSE", "NQCH", "NQAU", "NQNZ", "NQHK", "NQSG", "NQIL", "NQBR", "NQCL", "NQCO", "NQMX", "NQPL", "NQRU", "NQTR", "NQCN", "NQIN", "NQID", "NQMA", "NQKR", "NQPH", "NQTW", "NQTH", "NQEG", "NQMY", "NQZA", "NQUS")

Country\_Name <- c("Japan", "United Kingdom", "Canada", "Austria", "Belgium", "Finland", "France", "Denmark", "Germany", "Greece", "Ireland", "Italy", "Netherlands", "Norway", "Spain", "Sweden", "Switzerland", "Australia", "New Zealand", "Hong Kong", "Singapore", "Israel", "Brazil", "Chile", "Colombia", "Mexico", "Poland", "Russia", "Turkey", "China", "India", "Indonesia", "Morocco", "Korea", "Philippines", "Taiwan", "Thailand", "Egypt", "Malaysia", "South Africa", "US")

Countries <- data.frame(Country\_Code, Country\_Name)

Sector\_Code <- c("0001", "0530", "0533", "0537", "0570", "0573", "0577", "0580", "0583", "0587", "1000", "1300", "1353", "1357", "1700", "1730", "1737", "1750", "1755", "1757", "1770", "1771", "1775", "1777", "2000", "2300", "2353", "2357", "2700", "2710", "2713", "2717", "2720", "2723", "2727", "2730", "2733", "2737", "2750", "2753", "2757", "2770", "2771", "2773", "2775", "2777", "2779", "2790", "2791", "2793", "2795", "2797", "2799", "3000", "3300", "3355", "3500", "3530", "3535", "3537", "3570", "3573", "3577", "3700", "3720", "3722", "3724", "3726", "3728", "3740", "3743", "3745", "3747", "3760", "3763", "3765", "3767", "3780", "4000", "4530", "4533", "4535", "4537", "4570", "4573", "4577", "5000", "5300", "5330", "5333", "5337", "5370", "5371", "5373", "5375", "5377", "5379", "5500", "5553", "5555", "5557", "5700", "5751", "5752", "5753", "5755", "5757", "5759", "6000", "6530", "6570", "7000", "7530", "7535", "7570", "7573", "7575", "7577", "8000", "8300", "8500", "8530", "8532", "8536", "8538", "8570", "8600", "8630", "8633", "8637", "8670", "8671", "8672", "8673", "8674", "8675", "8676", "8677", "8700", "8770", "8771", "8773", "8775", "8777", "8779", "9000", "9530", "9533", "9535", "9537", "9570", "9572", "9576", "9578")

Sector\_Name <- c("Oil & Gas", "Oil & Gas Producers", "Exploration & Production", "Integrated Oil & Gas", "Oil Equipment, Services & Distribution", "Oil Equipment & Services", "Pipelines", "Alternative Energy", "Renewable Energy Equipment", "Alternative Fuels", "Basic Materials", "Chemicals", "Commodity Chemicals", "Specialty Chemicals", "Basic Resources", "Forestry & Paper", "Paper", "Industrial Metals & Mining", "Nonferrous Metals", "Iron & Steel", "Mining", "Coal", "General Mining", "Gold Mining", "Industrials", "Construction & Materials", "Building Materials & Fixtures", "Heavy Construction", "Industrial Goods & Services", "Aerospace & Defense", "Aerospace", "Defense", "General Industrials", "Containers & Packaging", "Diversified Industrials", "Electronic & Electrical Equipment", "Electrical Components & Equipment", "Electronic Equipment", "Industrial Engineering", "Commercial Vehicles & Trucks", "Industrial Machinery", "Industrial Transportation", "Delivery Services", "Marine Transportation", "Railroads", "Transportation Services", "Trucking", "Support Services", "Business Support Services", "Business Training & Employment Agencies", "Financial Administration", "Industrial Suppliers", "Waste & Disposal Services", "Consumer Goods", "Automobiles & Parts", "Auto Parts", "Food & Beverage", "Beverages", "Distillers & Vintners", "Soft Drinks", "Food Producers", "Farming, Fishing & Plantations", "Food Products", "Personal & Household Goods", "Household Goods & Home Construction", "Durable Household Products", "Nondurable Household Products", "Furnishings", "Home Construction", "Leisure Goods", "Consumer Electronics", "Recreational Products", "Toys", "Personal Goods", "Clothing & Accessories", "Footwear", "Personal Products", "Tobacco", "Health Care", "Health Care Equipment & Services", "Health Care Providers", "Medical Equipment", "Medical Supplies", "Pharmaceuticals & Biotechnology", "Biotechnology", "Pharmaceuticals", "Consumer Services", "Retail", "Food & Drug Retailers", "Drug Retailers", "Food Retailers & Wholesalers", "General Retailers", "Apparel Retailers", "Broadline Retailers", "Home Improvement Retailers", "Specialized Consumer Services", "Specialty Retailers", "Media", "Broadcasting & Entertainment", "Media Agencies", "Publishing", "Travel & Leisure", "Airlines", "Gambling", "Hotels", "Recreational Services", "Restaurants & Bars", "Travel & Tourism", "Telecommunications", "Fixed Line Telecommunications", "Mobile Telecommunications", "Utilities", "Electricity", "Conventional Electricity", "Gas, Water & Multi-utilities", "Gas Distribution", "Multi-utilities", "Water", "Financials", "Banks", "Insurance", "Nonlife Insurance", "Full Line Insurance", "Property & Casualty Insurance", "Reinsurance", "Life Insurance", "Real Estate", "Real Estate Investment & Services", "Real Estate Holding & Development", "Real Estate Services", "Real Estate Investment Trusts", "Industrial & Office REITs", "Retail REITs", "Residential REITs", "Diversified REITs", "Specialty REITs", "Mortgage REITs", "Hotel & Lodging REITs", "Financial Services", "Financial Services", "Asset Managers", "Consumer Finance", "Specialty Finance", "Investment Services", "Mortgage Finance", "Technology", "Software & Computer Services", "Computer Services", "Internet", "Software", "Technology Hardware & Equipment", "Computer Hardware", "Semiconductors", "Telecommunications Equipment")

Sectors <- data.frame(Sector\_Code, Sector\_Name)

Nasdaq <- merge(Nasdaq, Countries, by = "Country\_Code")

Nasdaq <- merge(Nasdaq, Sectors, by = "Sector\_Code")

Big\_Sector <-Sectors[grep("000", Sectors$Sector\_Code),]

names(Big\_Sector) <- c("Big\_Sector","Broad\_Sector")

Nasdaq <- merge(Big\_Sector, Nasdaq, by = "Big\_Sector")

Nasdaq<-Nasdaq[,-c("Big\_Sector","Sector\_Code","Country\_Code")]

Nasdaq<-subset(Nasdaq, select = -c(Big\_Sector, Sector\_Code, Country\_Code)) # deleting not needed columns

head(Nasdaq)

---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

#Load packages

#Load packages

library(RCurl)

library(plyr)

library(stringr)

library (gdata)

#Indexes <- c("NQG1300EUR", "NQFR8000EUR", "NQEURO2000EUR")

Indexes <- c("NQG1300EUR")

End\_Date = Sys.Date()-3 # format should be “2015-01-19”

Start\_Date = End\_Date-5 #format should be “2015-01-19”

URL\_Nasdaq <- paste("<https://indexes.nasdaqomx.com/Index/ExportHistory/>",

Indexes,

"?startDate=",

Start\_Date,

"T00:00:00.000&endDate=",

End\_Date,

"T00:00:00.000&timeOfDay=EOD.xls") # URL for all nasdaq indexes

URL\_Nasdaq <- gsub(" ", "", URL\_Nasdaq) # somehow it creates it with space => we need to remove spaces

#Nasdaq <- read.xlsx(textConnection(URL\_Nasdaq)) # trasforming to dataframe

#Nasdaq <- read.xls (URL\_Nasdaq, sheet = 1, header = TRUE)

#installXLSXsupport()

Nasdaq <- read.xls(URL\_Nasdaq, sheet = 1, header = TRUE)

URL\_Nasdaq