Data From API

Max

September 12, 2018

Objective

Using API calls to gather data

In this example I'm using QUANDL package (www.quandl.com) to gather following economic data

- GDP
- Unemployment
- NASDAQ

Then I will merge then in a single dataset and remove any rows that has at least one NA value

Setup Connection

```
startDate=as.Date('2018/01/01')
allDates = data.table(Date=seq(from=startDate,to=Sys.Date(),by=1))

#setting up QUANDL to download load economic indicato
#https://www.quandl.com/searchr
Quandl.api_key('9d_xURs_x6eAjTwbszrb') #Max's API key

#set_config(use_proxy(url="http://proxy.micron.com",port=8080))  # setup proxy
httr::set_config( config( ssl_verifypeer = 0L ) ) #fixes issue with CURL and Micron Proxy
```

Getting GDP Data

```
### GDP Data ###

# Get GDP Data from Quandl
GDP = data.table(Quandl("FRED/GDP",start_date = startDate))

# Select Columns
GDP=GDP[,.(Date,GDP = Value)][order(Date)]

# create all missing Dates
GDP=merge(GDP,allDates,by='Date',all.y=T)[order(Date)]

# fill missing values
GDP[,GDP  := na.locf(GDP ,na.rm=F,fromLast=F,rule=1)] #rule=2 grabs the closest backwards or forward

head(GDP)
```

```
##          Date      GDP
## 1: 2018-01-01 20041.05
## 2: 2018-01-02 20041.05
```

```
## 3: 2018-01-03 20041.05
## 4: 2018-01-04 20041.05
## 5: 2018-01-05 20041.05
## 6: 2018-01-06 20041.05
```

Getting Unemployment Data

```
### Unemployment ###
# get data
Unemp = data.table(Quandl("FRED/UNRATE",start_date = startDate))
# select columns
Unemp=Unemp[,.(Date,Unemp = Value)]
# create all missing dates
Unemp=merge(Unemp,allDates,by='Date',all.y=T)[order(Date)]

# fill missing values
Unemp[,Unemp  := na.locf(Unemp ,na.rm=F,fromLast=F,rule=1)] #rule=2 grabs the closest backwards or forw

head(Unemp)
```

```
##          Date Unemp
## 1: 2018-01-01   4.1
## 2: 2018-01-02   4.1
## 3: 2018-01-03   4.1
## 4: 2018-01-04   4.1
## 5: 2018-01-05   4.1
## 6: 2018-01-06   4.1
```

Getting NASDAQ Data

```
### Unemployment ###
# get data
NASDAQ = data.table(Quandl("NASDAQOMX/NDX",start_date = startDate))
# select columns
NASDAQ=NASDAQ[,.(Date = `Trade Date`,NASDAQ = `Index Value`)]
# create all missing dates
NASDAQ=merge(NASDAQ,allDates,by='Date',all.y=T)[order(Date)]

# fill missing values
NASDAQ[,NASDAQ  := na.locf(NASDAQ ,na.rm=F,fromLast=F,rule=1)] #rule=2 grabs the closest backwards or f

head(NASDAQ)
```

```
##          Date  NASDAQ
## 1: 2018-01-01      NA
## 2: 2018-01-02 6511.34
## 3: 2018-01-03 6575.80
## 4: 2018-01-04 6584.58
## 5: 2018-01-05 6653.29
## 6: 2018-01-06 6653.29
```

Merge and Clean

```
dataset = merge(allDates,GDP,by='Date',all.x=T)
dataset = merge(dataset,Unemp,by='Date',all.x=T)
dataset = merge(dataset,NASDAQ,by='Date',all.x=T)

head(dataset)
```

```
##          Date      GDP Unemp  NASDAQ
## 1: 2018-01-01 20041.05   4.1      NA
## 2: 2018-01-02 20041.05   4.1 6511.34
## 3: 2018-01-03 20041.05   4.1 6575.80
## 4: 2018-01-04 20041.05   4.1 6584.58
## 5: 2018-01-05 20041.05   4.1 6653.29
## 6: 2018-01-06 20041.05   4.1 6653.29
```

```
# keep records don't have NAs
keep= complete.cases(dataset)
dataset=dataset[keep,]

head(dataset)
```

```
##          Date      GDP Unemp  NASDAQ
## 1: 2018-01-02 20041.05   4.1 6511.34
## 2: 2018-01-03 20041.05   4.1 6575.80
## 3: 2018-01-04 20041.05   4.1 6584.58
## 4: 2018-01-05 20041.05   4.1 6653.29
## 5: 2018-01-06 20041.05   4.1 6653.29
## 6: 2018-01-07 20041.05   4.1 6653.29
```