

# Харинаев 20.09.2021

## Описание датасета

Датасет содержит информацию об акциях компаний из индекса S&P 500 с 2013 по 2018г.

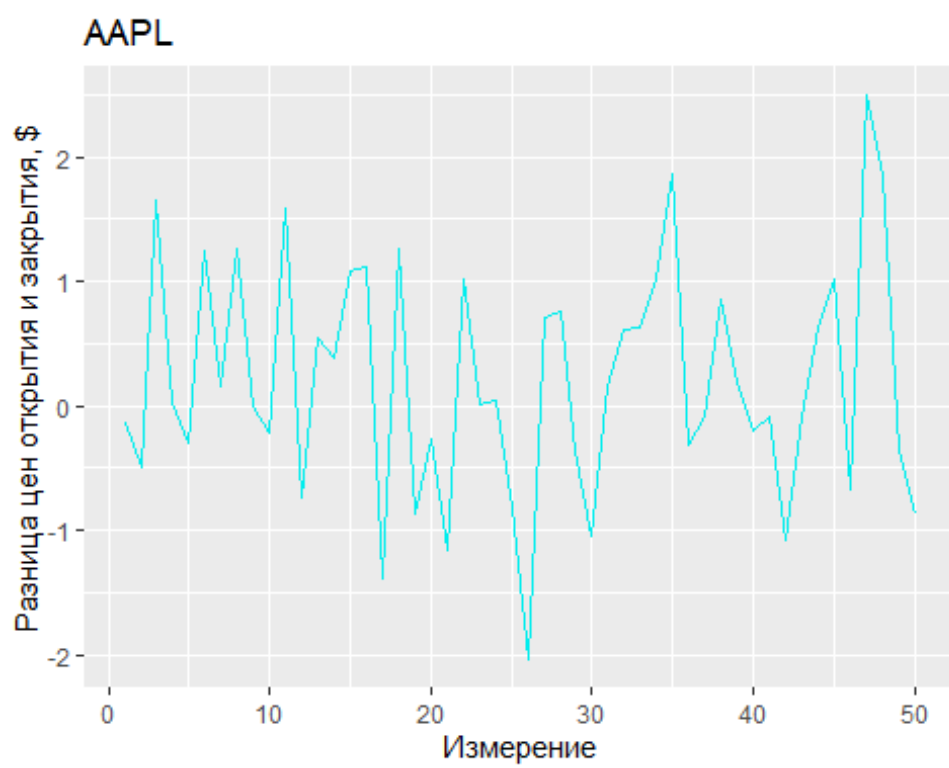
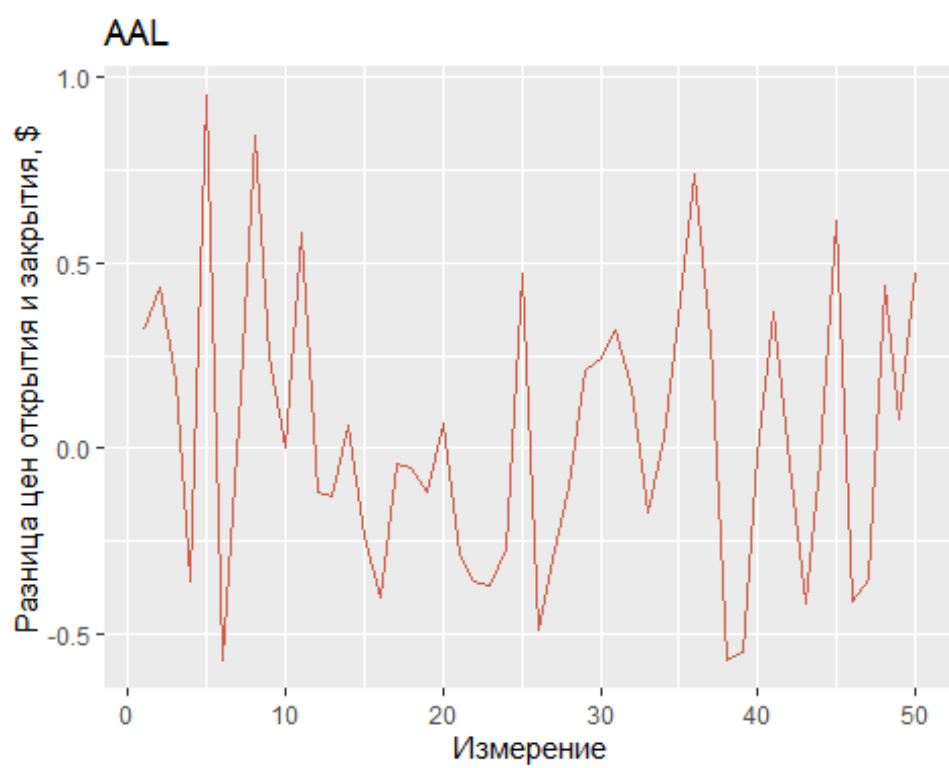
619 тыс. записей, 7 колонок:

1. дата
2. цена открытия
3. наивысшая цена в течении дня
4. наименьшая цена в течении дня
5. цена закрытия
6. объем продаж
7. тикер компании

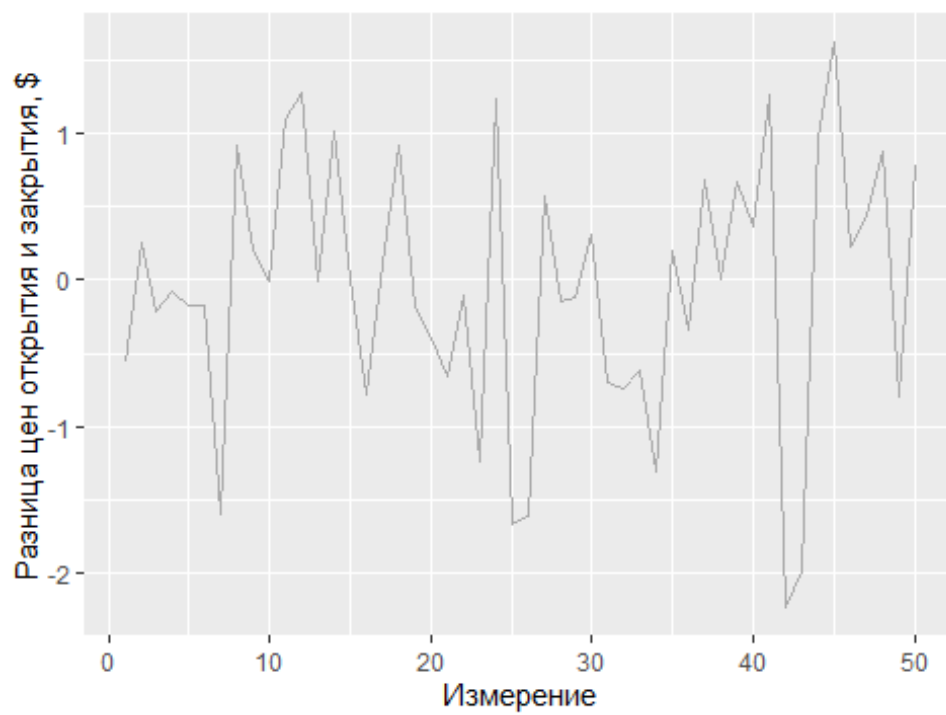
## 1. Функция

Динамика разности цены открытия и цены закрытия для n компаний за k измерений

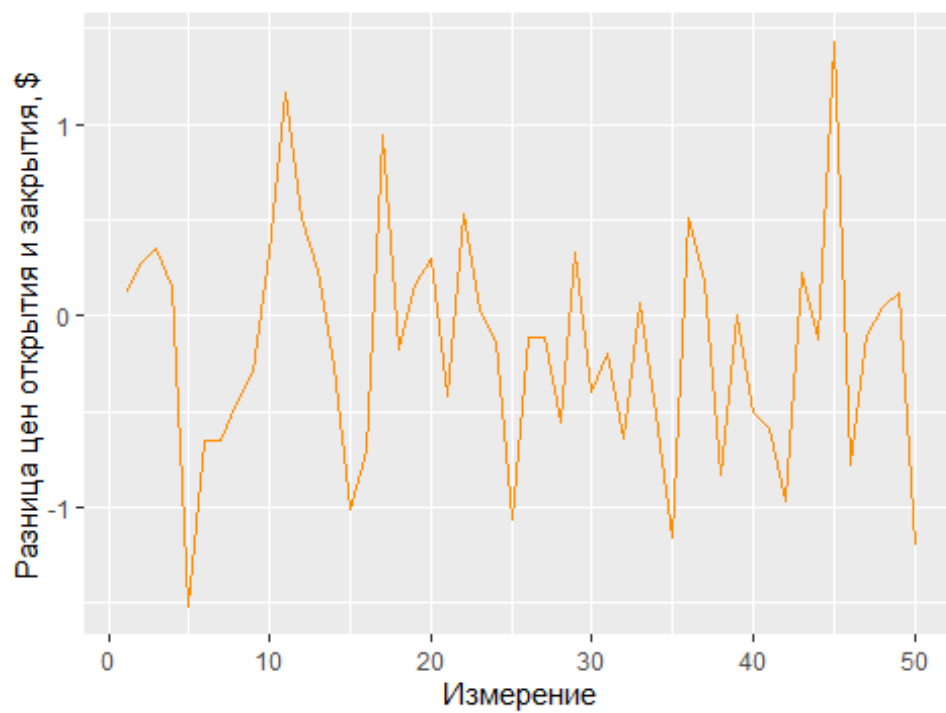
```
library(ggplot2)
mean_value <- function(n,k){
  data <- read.csv(file = '../dataset.csv')
  companies <- data['Name'][!duplicated(data$Name),][0:n]
  if (n >= 20){
    n = 20
  }
  if (k >= nrow(subset(data, data['Name'] == companies[1]))){
    k = nrow(subset(data, data['Name'] == companies[1]))
  }
  for (i in 1:n){
    company <- subset(data, data['Name'] == companies[i])
    value <- data.frame(difference = unlist(company['open'] - company['close']),
num = 1:nrow(company))
    plt <- ggplot(value[1:k,], aes(x=num)) + geom_line(aes(y=difference), color
= colors()[10*i+50]) + labs(title = companies[i], y = 'Разница цен открытия и
закрытия, $', x = 'Измерение')
    print(plt)
  }
}
mean_value(4,50)
```



AAP



ABBV

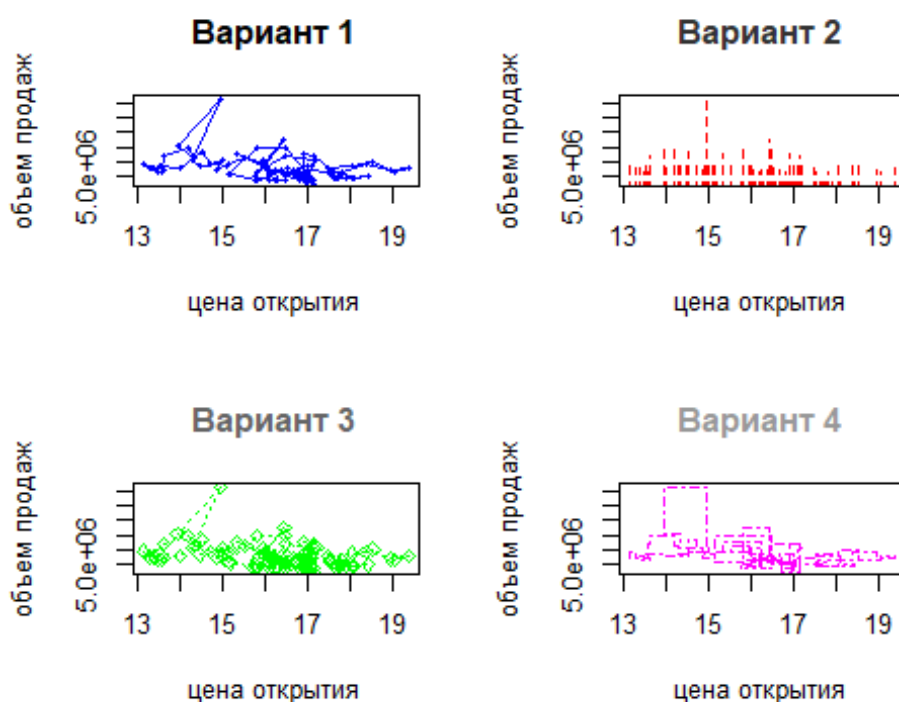


## 2. Способности визуализации функции plot

Рассмотрим зависимость цены открытия от объема сделок

```
data <- read.csv(file = '..\\dataset.csv')

par(mfrow = c(2,2))
my_color = c("blue", "red", "green", "magenta")
title = c('Вариант 1','Вариант 2','Вариант 3','Вариант 4')
title_color = c('grey1', 'grey20', 'grey40', 'grey60')
my_types = c('o','h','o','s')
my_pch = c(1,3,5,4)
my_cex = c(0.5, 1, 0.75, 2)
my_lty = c(1,2,3,4)
for (i in 1:4){
  plot(data$open[1:100], data$volume[1:100], xlab = 'цена открытия', ylab =
'объем продаж', col = my_color[i], main = title[i], col.main = title_color[i],
type=my_types[i], pch=my_pch[i], cex=my_cex[i], lty=my_lty[i])
}
```



## 3. Векторная обработка данных

```
my_data <- subset(data,
data['Name']=='AAPL')[c('open','high','low','close','volume')]
apply(my_data, 2, max)

##      open      high      low      close      volume
## 179.37 180.10 178.25 179.26 266833581.00

apply(my_data, 2, min)
```

```

##      open      high      low      close      volume
## 5.542420e+01 5.708570e+01 5.501420e+01 5.578990e+01 1.147592e+07

sapply(my_data['high'], min)

##      high
## 57.0857

sapply(my_data['low'], max)

##      low
## 178.25

lapply(my_data, median)

## $open
## [1] 108.97
##
## $high
## [1] 110.03
##
## $low
## [1] 108.05
##
## $close
## [1] 109.01
##
## $volume
## [1] 45668931

vapply(my_data['volume'], mean, numeric(1))

##      volume
## 54047900

mapply(function(num1, num2) max(c(num1, num2)), my_data['open'],
my_data['close'])

##      open
## 179.37

```