Tesla, Inc. Stocks Research

Task description

Nowadays it is quite popular though that any person can make money on stock exchange - all you need is starter money and some knowledge.
We decided to check if it is possible to get a big revenue from some big companies without much effort.

Data description

For our research we first wanted to use the **New York Stock Exchange dataset** that consists of the historical data on New York Stock Exchange from 2012 until 2016 parced from Yahoo Finance. But since tha data are quite outdated and we do not need the data on all the companies, we decided to get similar data.
For our final dataset we downloaded the historical data on **Tesla, Inc.** stocks from Yahoo Finanace (same resource the author of the previous dataset used) from January 5, 2015 until January 4, 2020.
You can find the .csv file created **here**.
We're going to load the data and see what we have.

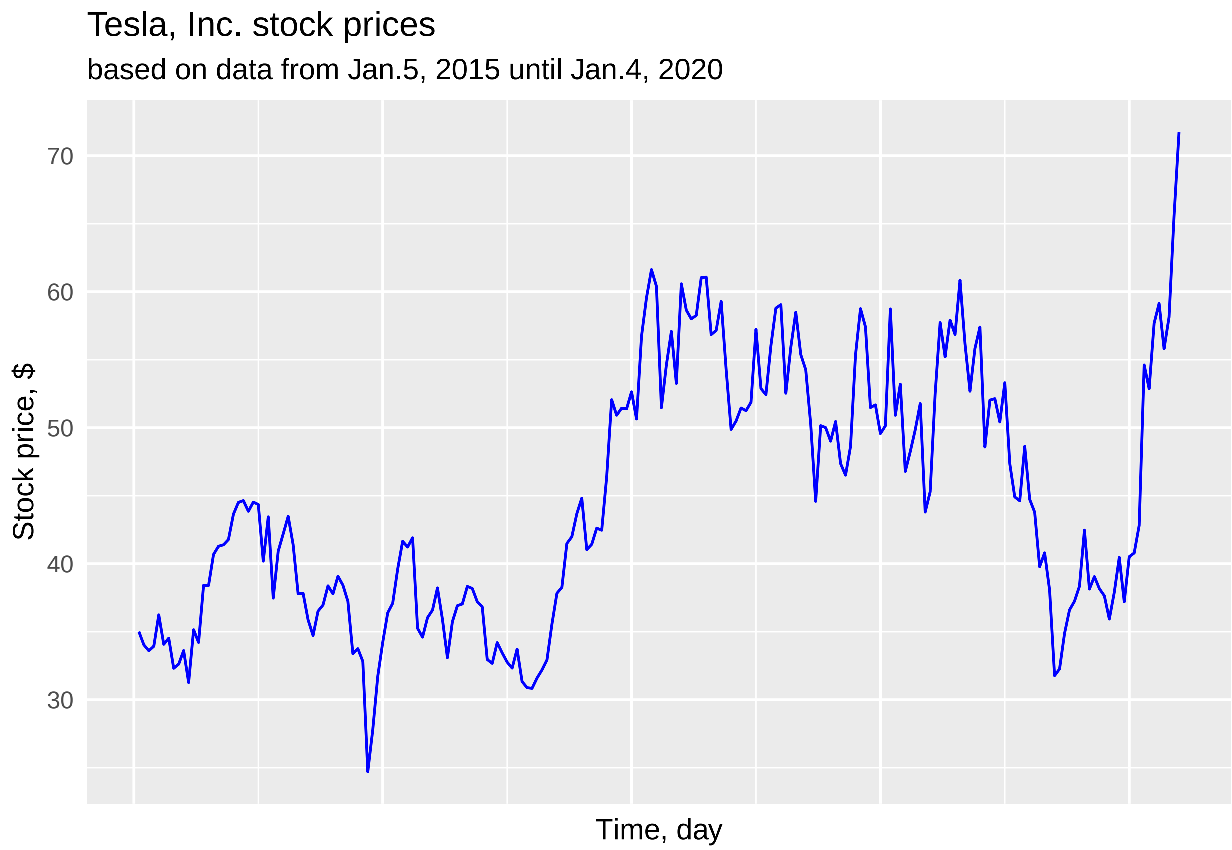
```
tsla.data <- read.csv("TSLA.csv")
head(tsla.data)
```

```
##         Date   Open   High    Low  Close Adj.Close  Volume
## 1 2015-01-05 214.55 216.50 207.16 210.09    210.09 5368500
## 2 2015-01-06 210.06 214.20 204.21 211.28    211.28 6261900
## 3 2015-01-07 213.35 214.78 209.78 210.95    210.95 2968400
## 4 2015-01-08 212.81 213.80 210.01 210.62    210.62 3442500
## 5 2015-01-09 208.92 209.98 204.96 206.66    206.66 4668300
## 6 2015-01-12 203.05 204.47 199.25 202.21    202.21 5950300
```

Here we have a data frame consisting of **7 different columns** and **1259** rows. What do these data mean? Each row in tha data frame represent a day, the stock was out on a market. The date of the day can be found in the **date** column, consisting of date objects.
All the follwoing columns, except for the last one, tell us about the **price** of the stock. The **open** and **close** close columns tells the price of the stock when the market opened and closed on that day. The **low** and **high** column consist of the lowest and the highest values the stock price reached during the day. **Adj.Close** column consists of the adjusted close price adjusted for both dividends and splits, althou it almost always is the same as the close one.
The last column - **volume** - consists of integer values refering to the numbers of shares that have been bought and sold for the day.

Data analysis

Let's first check how Tesla stock prices behaved themselves during the last five years.
For a better of the plot we're going to plot every 6^{th} day data, we'll show later why it doesn't change much.



As we can see here, the stock prices have significantly increased during last months. Let's then check if we can make some good money here :)

Introducing new metric

Since we wanted to check if we can make easy money on Tesla we'll test one of the easiest way - buying the shares one day and selling them just next day. Even easier way that does not envolve any tracking - buying and selling at the same time.
For our case let's take the buying and selling time to be just before the market is closing.
For this we'll need a **daily return** notion - the amount of stock price daily growth or descend. Daily return is calculated by a simple intuitive formula

$$R_t = \frac{C_t - C_{t-1}}{C_{t-1}} = \frac{C_t}{C_{t-1}} - 1$$

```
tsla.data$PrevClose <- tsla.data$Close
# since there is no previous day for the first day, we'll get NA value
# in order not to get NA in further calculations we'll get rid of the first day
tsla.data <- na.omit(transform(tsla.data, PrevClose = c(NA, PrevClose[-nrow(tsla.data)])))
tsla.data$DailyRet <- with(tsla.data, Close / PrevClose - 1)

head(tsla.data)
```

```
##         Date   Open   High    Low  Close Adj.Close  Volume PrevClose
## 2 2015-01-06 210.06 214.20 204.21 211.28    211.28 6261900    210.09
## 3 2015-01-07 213.35 214.78 209.78 210.95    210.95 2968400    211.28
## 4 2015-01-08 212.81 213.80 210.01 210.62    210.62 3442500    210.95
## 5 2015-01-09 208.92 209.98 204.96 206.66    206.66 4668300    210.62
## 6 2015-01-12 203.05 204.47 199.25 202.21    202.21 5950300    206.66
```

```
## 7 2015-01-13 203.32 207.61 200.91 204.25    204.25 4477300    202.21
##       DailyRet
## 2  0.005664254
## 3 -0.001561918
## 4 -0.001564361
## 5 -0.018801591
## 6 -0.021532938
## 7  0.010088487
```

Now we'll see what we can say on the daily return values we got.
We can start off with some simple numeric characterics.

```
cat("mean value: ", mean(tsla.data$DailyRet), "  ")
```

```
## mean value:  0.000991882
```

```
cat("max value: ", max(tsla.data$DailyRet), "  ")
```

```
## max value:  0.1766923
```

```
cat("min value: ", min(tsla.data$DailyRet))
```

```
## min value:  -0.1390154
```

Looking at these numbers we already can say that day-to-day changes in the stock prices were not that significant during the last five years. It drives us to a conculsion that the strategy of selling the shares just the next day is **not the best strategy for a good money income, although it's a safe one** - you will not loose much if something goes wrong.

Still, daily return can give even more information on some company stock prices - daily return can be used to derrive **annual return**.
Annual return can be calucated with a single daily return value by the following formula

$$Y_t = (R_t + 1)^{365} - 1$$

We're now going to add one more column with calucated annual return.

```
tsla.data$YearRet <- with(tsla.data, power(DailyRet + 1, 365) - 1)
head(tsla.data)
```

```
##         Date   Open   High    Low  Close Adj.Close  Volume PrevClose
## 2 2015-01-06 210.06 214.20 204.21 211.28    211.28 6261900    210.09
## 3 2015-01-07 213.35 214.78 209.78 210.95    210.95 2968400    211.28
## 4 2015-01-08 212.81 213.80 210.01 210.62    210.62 3442500    210.95
## 5 2015-01-09 208.92 209.98 204.96 206.66    206.66 4668300    210.62
## 6 2015-01-12 203.05 204.47 199.25 202.21    202.21 5950300    206.66
## 7 2015-01-13 203.32 207.61 200.91 204.25    204.25 4477300    202.21
##       DailyRet    YearRet
## 2  0.005664254  6.8586848
## 3 -0.001561918 -0.4347831
## 4 -0.001564361 -0.4352877
## 5 -0.018801591 -0.9990200
## 6 -0.021532938 -0.9996457
## 7  0.010088487 38.0111416
```

Now we can analyze what data we get.
Again we'll start with some simple characteristics.

```
cat("mean value: ", mean(tsla.data$YearRet), "  ")
```

```
## mean value:  6.782408e+22
```

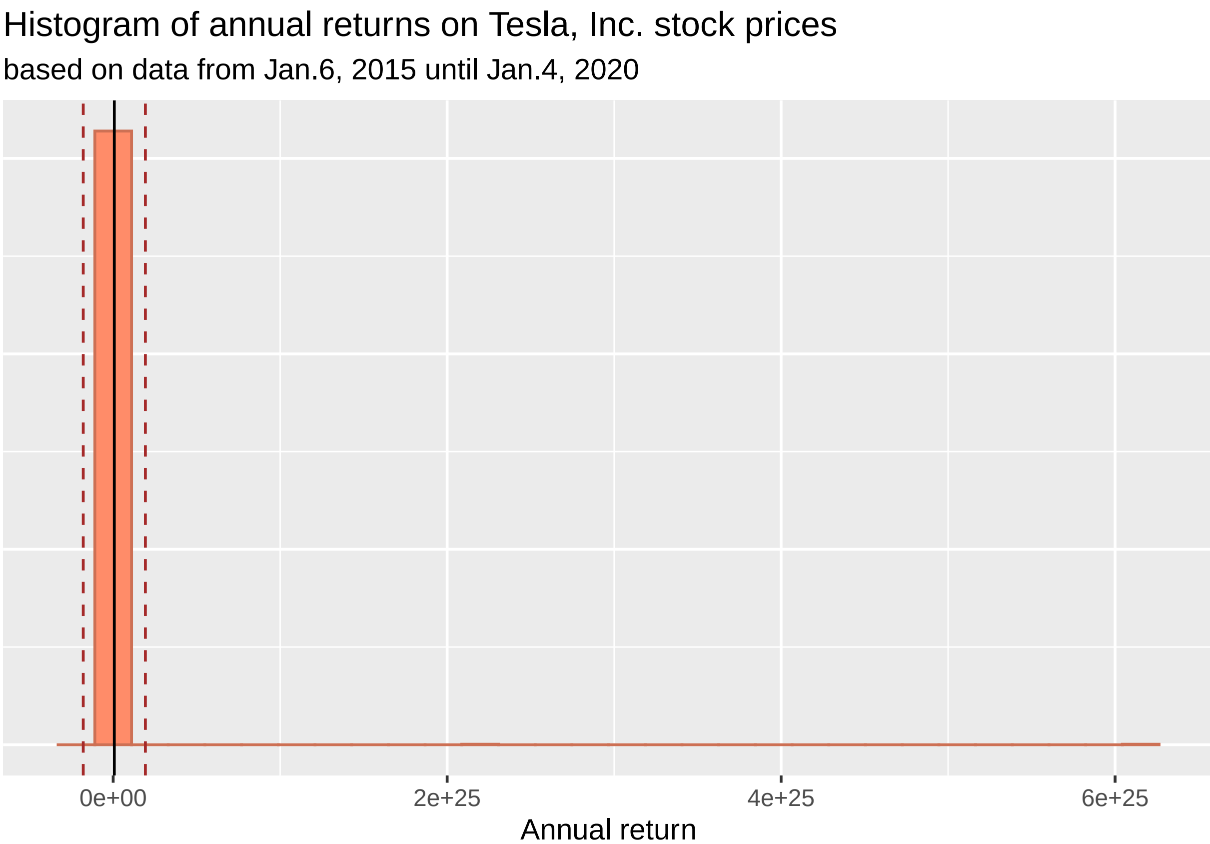
```r
cat("max value: ", max(tsla.data$YearRet), "  ")
```

```
## max value:  6.193926e+25
```

```r
cat("min value: ", min(tsla.data$YearRet))
```

```
## min value:  -1
```

The data we got seem quite interesting and not that self-explanatory. For further we'll first vizualise all the data we got on annual return.



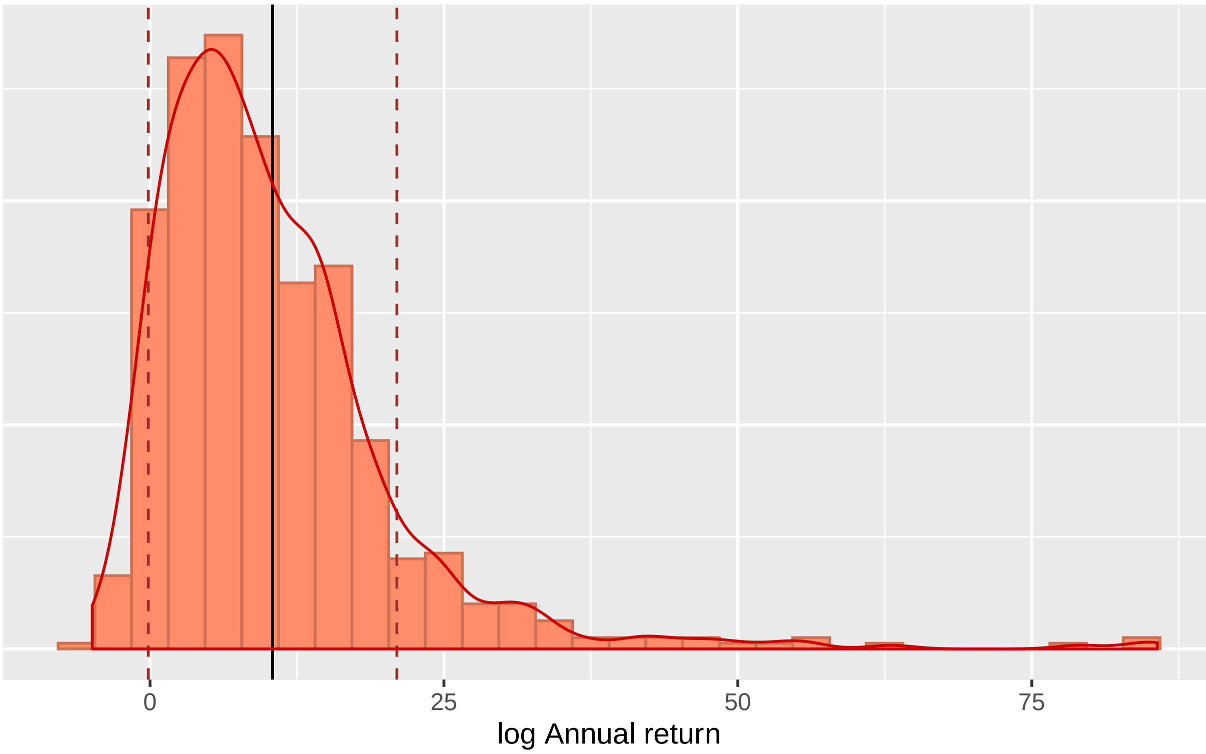
As we can see the data turned to have some quite extreme values and the plot does not help much. In order to fix this we're going to take the logarithm on the y axis, since we're working with ratios, and plot the data we got.

```r
log.ret <- with(tsla.data, log2(YearRet))
log.ret <- log.ret[!is.nan(log.ret) & !is.infinite(log.ret)]
head(log.ret)
```

```
## [1]  2.777932  5.248350  3.126778 12.578770 13.357224 13.583587
```

Histogram of annual returns on Tesla, Inc. stock prices

based on data from Jan.6, 2015 until Jan.4, 2020



It is natural to suggest that annual return follows a **lognormal distribution**, but as we can see the log of annual return is not that normal-lookig. For further analysis we're going to work with *log* **annual return**.

```
cat("skewness: ", skewness(log.ret), " ")
```

```
## skewness:  2.636748
```

```
cat("kurtosis: ", kurtosis(log.ret))
```

```
## kurtosis:  15.15702
```