# Statistics Assignment

Giampietro Ciancio 965991
Riccardo Valenti 979784

### Stock description, data collection and notation

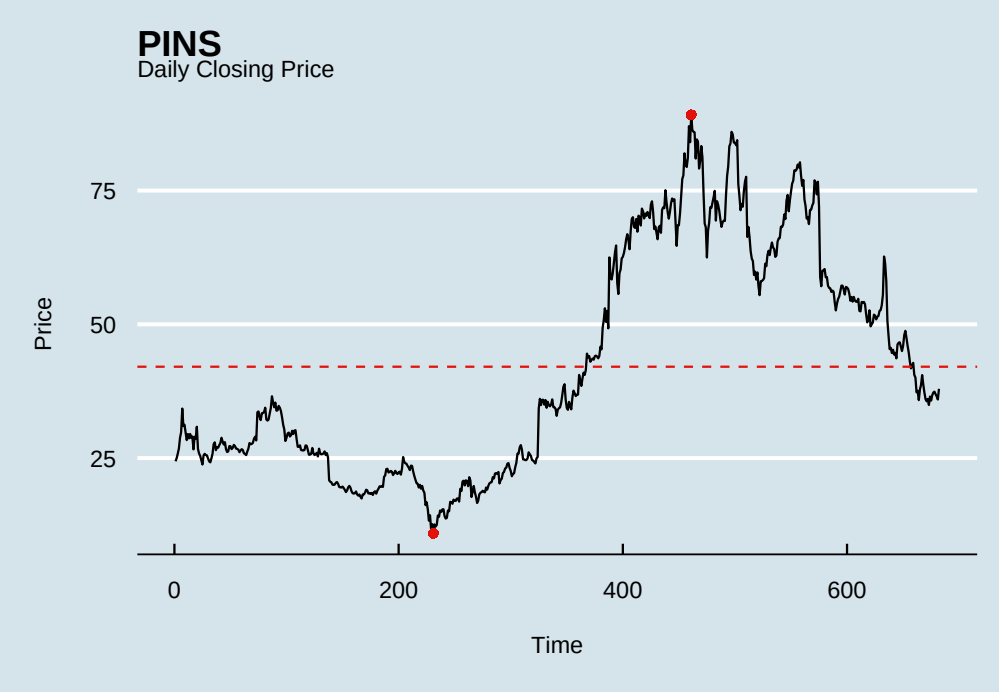
We start collectin data from the `tidyquant` stocks database. We'll be using Pinterest stock closing price from the IPO date (April 18, 2019) until the end of 2021 (December 31, 2021). Pinterest is an image sharing and social media service designed to enable saving and discovery of information on the internet using image. The statistical unit is the day, the variable is the price and the unit of measure of the price is dollar. We define the IPO date as $t = 1$ and the last day of 2021 as $T$. Hence, the stock closing price at $t = 1$ will be denoted as $P_1$, the last closing price as $P_T$ and the generic day closing price as $P_t$. We download the data and pipe it to extrapolate only the `close` column which we assign to vector `price`.

```
price <- tq_get("PINS", get = "stock.prices", from = "2019-04-18", to = "2021-12-31") %>%
  .$close
head(price)
```

```
## [1] 24.40 24.99 25.85 26.80 28.80 29.85
```

### Daily Closing Price

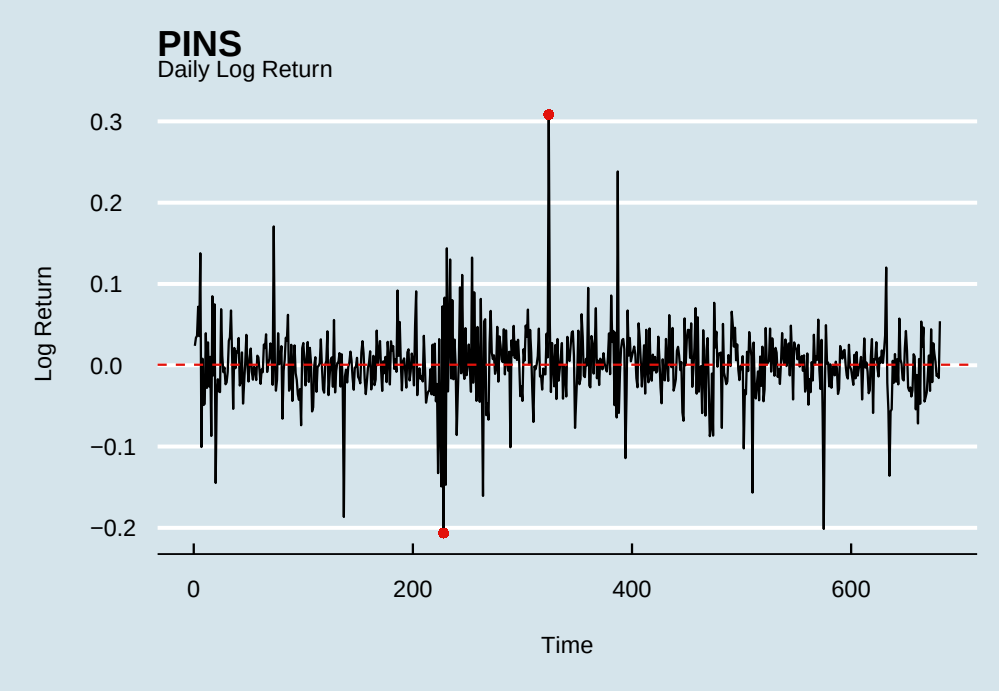
We plot the data as a line having $t$ on the x axis and $P_t$ on y axis.



From the plot we observe the evolution of the stock price in the above said interval. After the IPO, the stock prices moved in the interval between 25$ and 35$, before plunging from the 100th day circa. After day 200, the stock price showed a sign of recovery and then fell reaching its minimum shortly after. From that point on, the stock prices showed a sharp and constant growth until the 400th day, from which it slowed down for a while and then start growing even faster reaching the maximum price at day 450 circa. The stock then showed a very high volatility, showing both positive and negative spike in a overall decreasing trend. The average closing price of the stock is 40$ circa and we can approximately say that the stock price have been lower than mean in the first half of the period and greater in the second half.

### Daily Log Return

We plot the data as a line having $t$ on the x axis and $r_t$ on y axis.




### Mean

```
m.1
```

```
## [1] 0.0005699206
```

```
m.2
```

```
## [1] 0.0006144026
```

Comments on mean

### Standard deviation

```
s.1
```

```
## [1] 0.04283665
```

```
s.2
```

```
## [1] 0.04287659
```

comments on sd

### Correlation

```
corr.log.rt
```

```
## [1] -0.06192373
```

The correlation coefficient shows a very weak negative linear dependence. A negative linear dependence means that given a log return $r_k$ at day $k$, then $r_{k+1}$ will tend to move the stock price in the opposite direction with a more or less constant magnitude. However, the very low correlation coefficient tells us that the daily log-return of day $k+1$ cannot be expressed as linear function of $r_t$. Hence the linear dependence is negligible and we cannot draw any conclusion about the relationship of two consecutive days log-return.

### Absolute value

```
m.1.abs
```

```
## [1] 0.02883113
```

```
m.2.abs
```

```
## [1] 0.02887561
```

```
s.1.abs
```

```
## [1] 0.03166774
```

```
s.2.abs
```

```
## [1] 0.03168204
```

comments on means and sd

```
corr.log.rt.abs
```

```
## [1] 0.1568449
```

## Source

Giampietro Ciancio, Riccardo Valenti. 2022. "Stats Assignment." 2022. https://github.com/giampietrociancio/StatsAssignment.