# Business Intelligence for Financial Services Report

Leonardo Fraquelli
l.fraquelli@campus.unimib.it
820651

June 4, 2019

# Contents

# Part I

# Offline Project

# Chapter 1

# Introduction

In the following document I'll describe the work on the BISF 2018-2019 course project.
The student's required to choose a set of stocks and perform **analysis, visualization,prediction and portfolio management** .
A **Web Applet is also required**.

# Chapter 2

# Data Summary

The securities chosen for this project are:

**Google (GOOG)**

**Amazon (AMZN)**

**Electronic Arts (EA)**

**Activision-Blizzard (ATVI)**

**Philip Morris (PM)**

**Universal Tobacco (UVV)**

The groups represented by these securities reflect three different sectors: **Data**,**Entertainment**, **Tobacco**
The following paragraphs will explain the reasoning behind such choices.

## 2.1 Data

Google and Amazon securities are the representatives behind the **Data** tag.
They operate with a huge amount of users and have *some* similar services.
They were expected to have some close relationship but further analysis will reveal that they are, in fact, not related at all, suggesting a mistake on my part.

## 2.2 Entertainment

Electronic Arts and Activision-Blizzard securities are the representatives behind the **Entertainment** tag.
Their field of operation is the **Gaming Industry** and as such they were expected to have a similar trend.

## 2.3 Tobacco

Philip Morris and Universal securities are the representatives behind the **Tobacco** tag.
While both operate in the tobacco sector they are mostly unrelated, with PM producing tobacco-related products and UVV being a raw tobacco supplier. With both are in a different market I expected them to be somehow related based on the raw resource (tobacco) they both share.

## 2.4 Downloading the data

The securities in this project have been taken from the **NASDAQ** and **NYSE** markets, the former holds *Data and Entertainment* securities, while the latter holds *Tobacco*.

Listing 2.1: Downloading a security

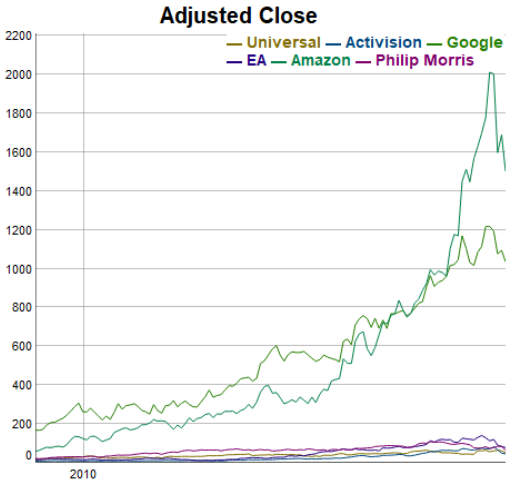
```
package(tseries)

PM_MK<- as.xts(
        na.omit(
                get.hist.quote(
                        instrument ="PM",
                        start =start_date,
                        end =end_date,
                        quote ="Low",
                        provider ="yahoo",
                        compression ="m",
                        retclass ="zoo"
                          )
                )
               )
```

The start and end date are, respectively, **2009-1-1** and **2019-1-1** for the whole program.

## 2.5 Analysis

After downloading the data we can get a general overview of the **close** of the securities in one whole using the *dygraph* package.
The graph doesn't allow for any analysis, due to the big differences in closing values, we therefore **scale** the dataset to allow for a better overview:



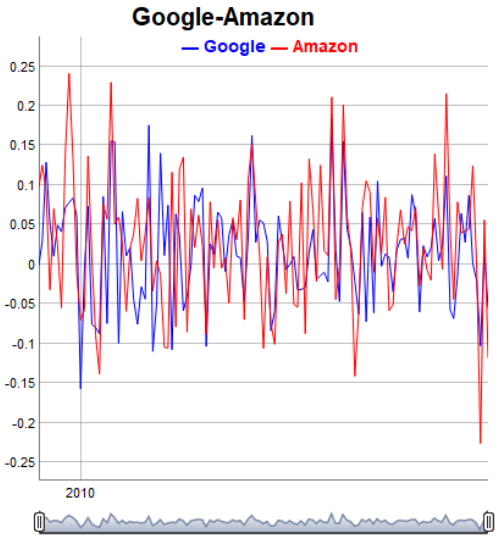


The graph shows an overall growth in all securities, with many highs and lows during the years. A good growth was expected from the markets in order to survive 2008's economical crysis.

## 2.6 Returns

Since an analysis of the whole returns of the whole dataset is messy and unreadable I've decided to analyse them in pairs.

### 2.6.1 Google-Amazon

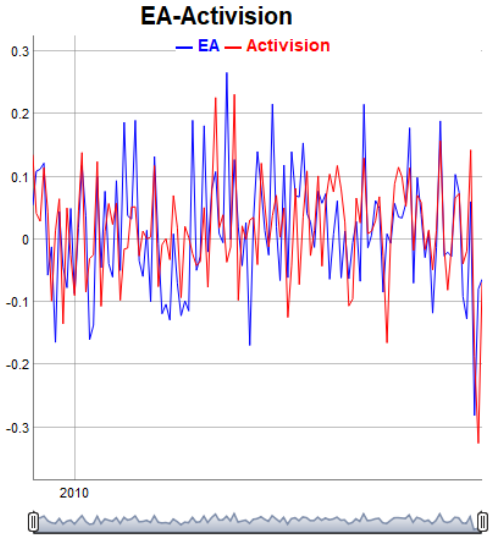



The graph shows what seems a huge volatility in the returns, with many different highs and lows.
Some of these are representatives of *things* happening inside the company,while others seem to have no particular headlines associated to them:

- Google's peak in **July 2015** is associated with both google's quarter's report [1] and with it's apointing of a new CFO, but at the same time Amazon has nothing associated.
- Amazon's first peak **2009** has no particular headlines associated to it
- Amazon's second peak in **2010, September** signals the $ 1tn threshold of the company
- Google's first, and worst, low **2010, January** has some news regarding it such as Google's refusal to keep censoring results in china
- **October 2018** is Amazon's worst month, associated with a heavy pressure applied from its competitors.

[1] https://www.fool.com/investing/general/2015/08/05/why-google-inc-stock-soared-22-in-july.aspx

### 2.6.2 ElectronicArts-Actvision

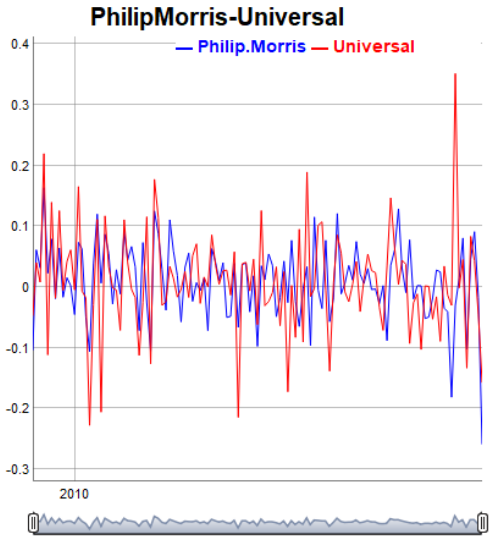



While looking as volatile as the previous graph, it appears at first glance that the two securities are more related, in comparison to the previous couple. Some interesting events:

- EA's first noticeable low in **August 2009** is interesting as EA's reports indicate it was **the #1 publisher in North America and Europe in the June quarter**
- EA's second peak **May 2011** seems normal: the company reports being top seller in iOS games and being # 1 publisher in europe.
- EA's biggest low, **October 2018** represents EA's failure in standing out against its competitors (such as ATVI too, which didn't fare much better)
- ATVI first noticeable peak is in **March 2010**, with fans having a *strong* demand for popular ATVI titles [2]
- ATVI's biggest low **November 2018** represents a huge decline in the player base in its player base.

[2]https://investor.activisionblizzard.com/news-releases/news-release-details/activision-blizzards-march-quarter-2010-financial-performance

### 2.6.3 PhilipMorris-Universal




The two companies while treating the same product have very different graphs, with UVV having big spikes in both highs and lows, and PM being less volatile in its swings

- Both UVV and PM report a sales increase in **May 2009**, reflected in both having a first peak after announcing their quarter's analysis
- UVV in **2010, May** reports a decrease in sales, corresponding with one of the first lows.
- UVV has an unexpected low in **August, 2013** after reporting a joint venture with *Avoca Inc.*
- UVV has a peak in **May 2018** after reporting huge successes in its quarter
- PM has its most relevant low in **December 2018** after being *downgrade* by Credit Suisse, in what was an overall rough year

## 2.7 Four panel diagnostics

For each financial histrument, 4 different graphics have been elaborated:

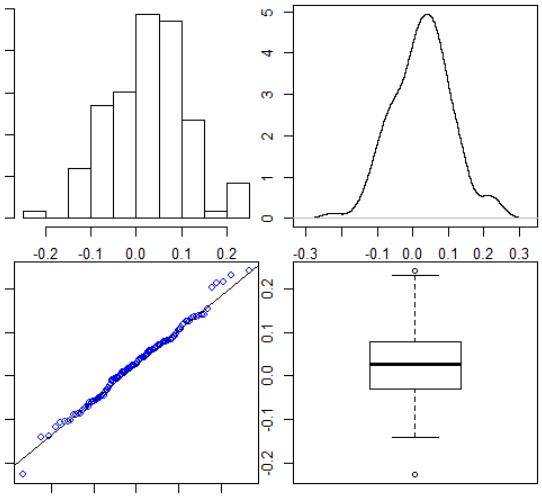
1. Histogram plot
2. QQ-Plot
3. Density plot

4. boxplot

Histogram and Density plot allows us to analyse the returns' density, QQ-plots the distribution and boxplots makes outliers easy to catch

### 2.7.1 AMZN

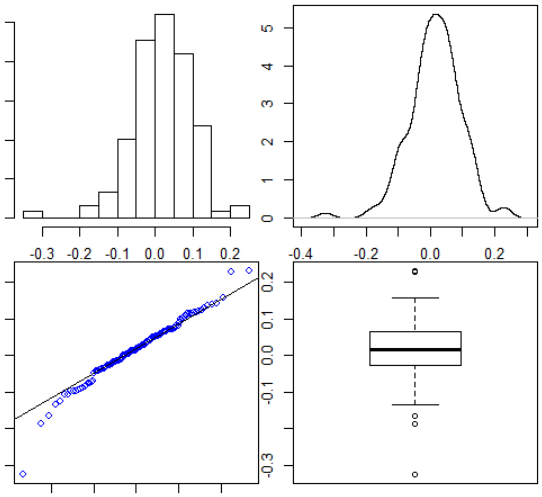
Amazon's panel indicates an almost normal shaped distribution, with almost




no outliers.

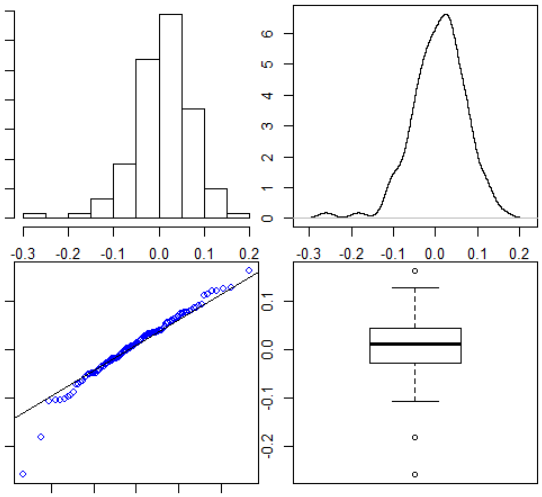
### 2.7.2 ATVI

Activision's panel distribution deviates from the previous, having visible outlier values, with most values being center on the 0.0-0.1 range



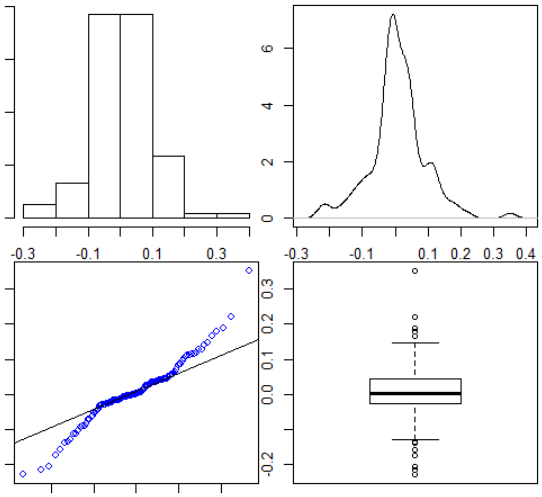
### 2.7.3 PM

PM's density curve peaks on the 0.0-0.1 range, with a few outliers on the -0.3 - -0.1 range




### 2.7.4 UVV

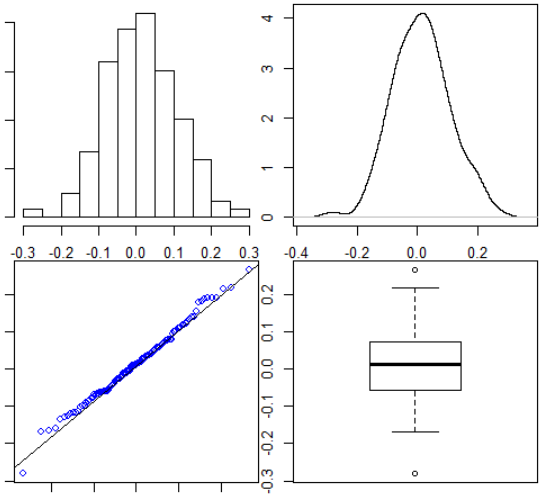
UVV's curve is the only one whose's density curve has a negative range of values, with the range going from -0.1 to 0.1 and outlying values being the biggers ones




### 2.7.5 EA

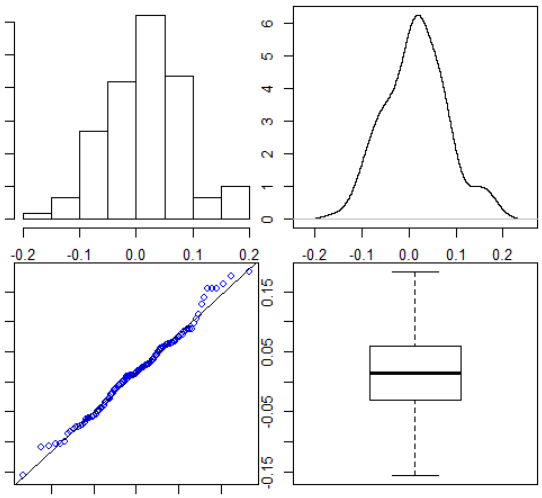
EA's graph as opposed to it's competitor, ATVI, follows a more normal distribtuion having a concentration of values on the -0.1 to 0.1 range

some outliers can still be found, these being extremes such as -0.3 and 0.3



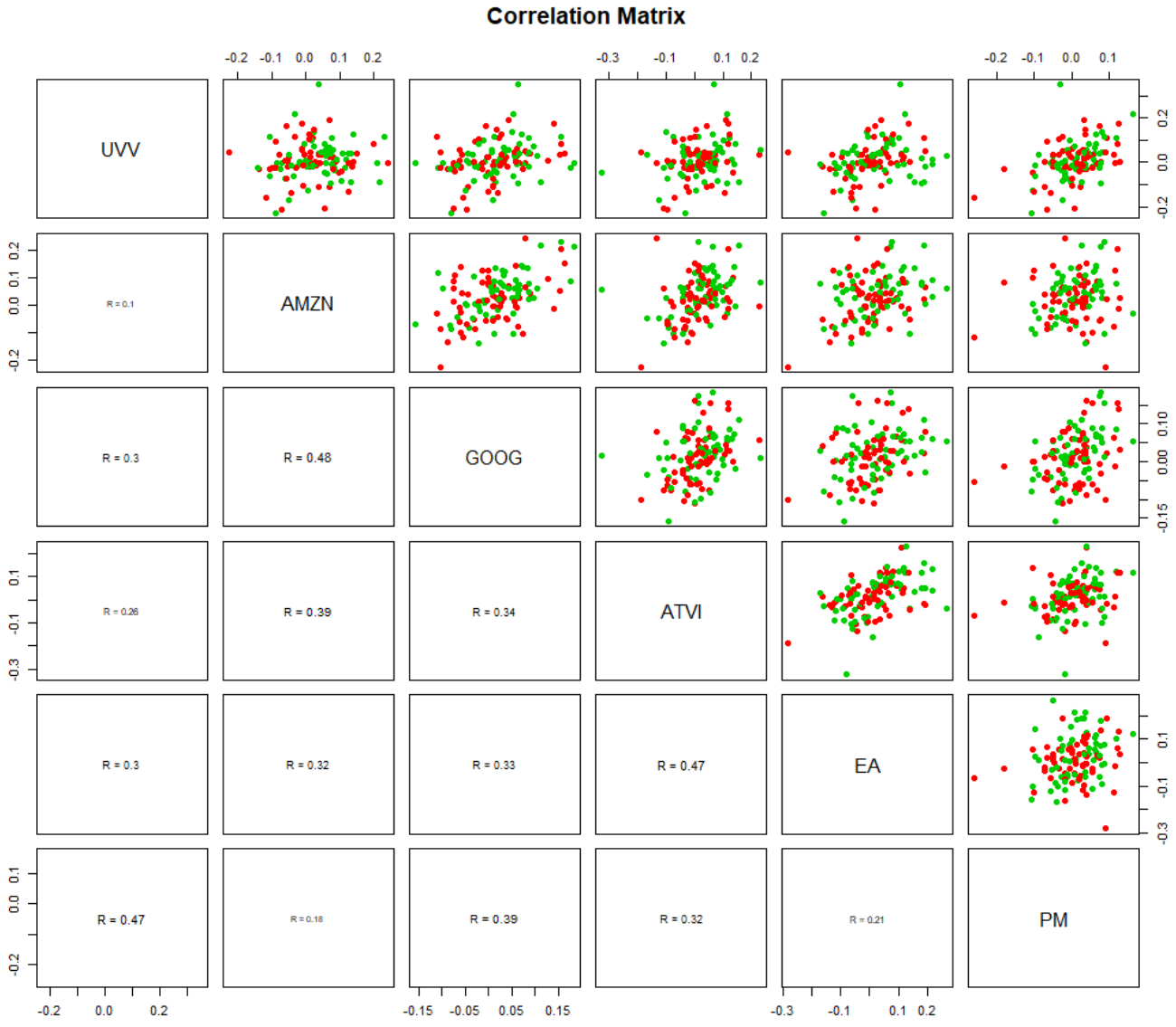

### 2.7.6 GOOG

Google's distribution shows no outliers in the boxplot and is almost normally shaped, with values at the extremes being few.
Google's distribution is particularly dense on the 0.0 range



### 2.7.7 Correlation Matrix

The pairs function has been "split" into two differnet panels.
The lower panel gives the Correlation between the given stocks, this gives us a written and more clear indication than the upper panel's scatterplot.
We can see that stocks in similar sector are, somehow, slightly related and therefore influence each other's price.



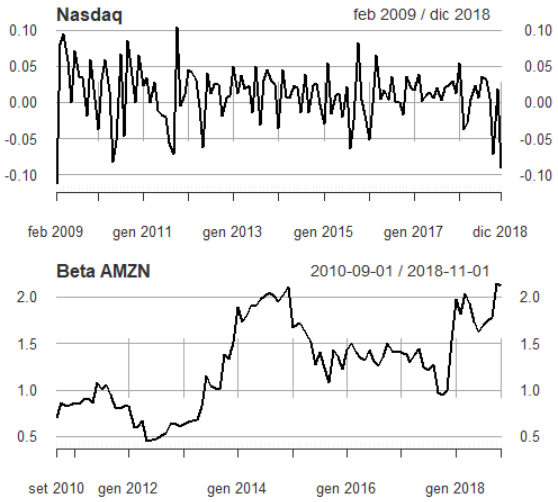
# Chapter 3

# Beta and Forecasting

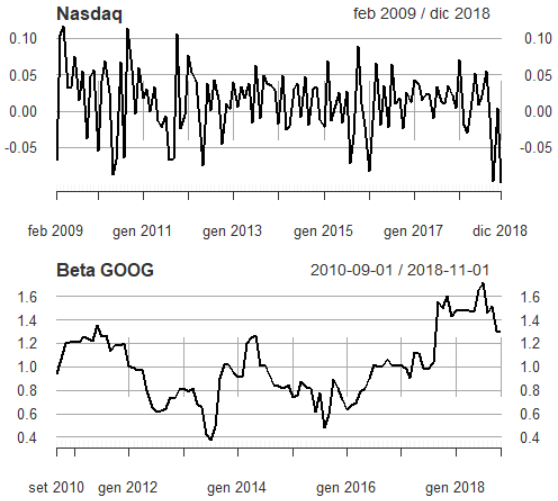
## 3.1 Beta

The Beta value is useful for analysing the risk of a financial security.
It is useful in portfolio management and forecasting. Higher Beta values means higher risk, thus higher reward, but higher volatility too.
All technological securities have been compared with the NASDAQ Composite index, while PM and UVV have been compared with the SPY Fund
NASDAQ Composite is heavily weighted towards information technology companies.
SPY tracks a market-cap-weighted index of US large- and midcap stocks selected by the S&P Committee.
In our portfolio we will choose the weights of our stocks based on their beta
In the following section we plot a comparison between the beta of the stocks and the selected market.

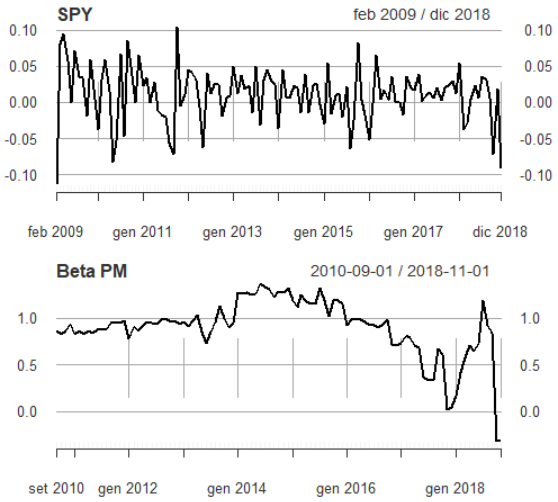
### 3.1.1 AMZN



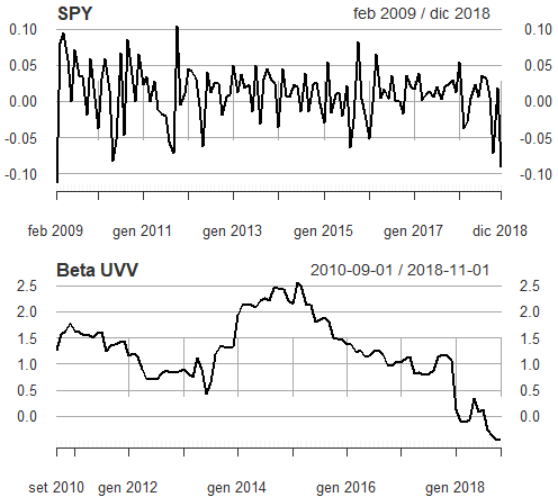

### 3.1.2 GOOG



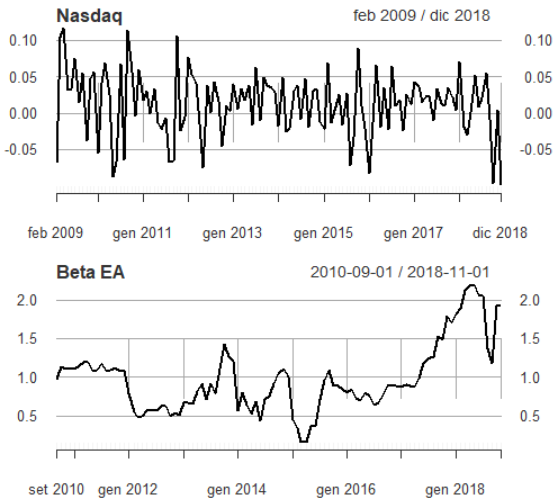
### 3.1.3 PM



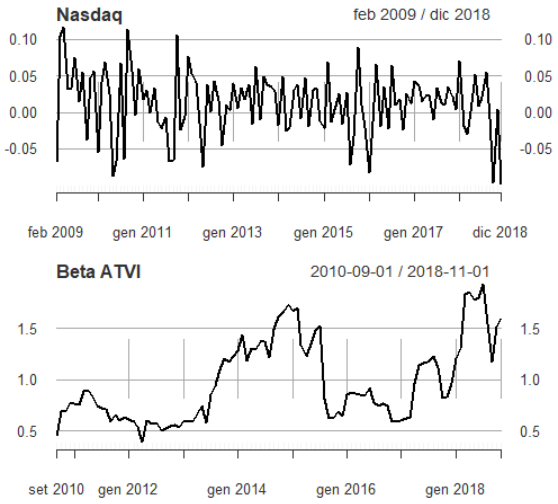

### 3.1.4 UVV



### 3.1.5 EA




### 3.1.6 ATVI




## 3.2 Predictive Analytics

Forecasting is another useful tool for a Portfolio.

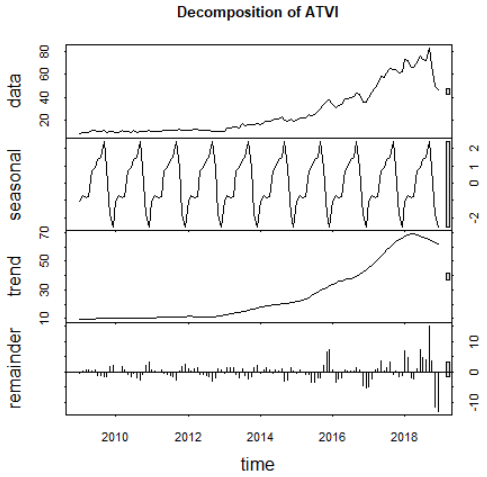
In order to make better investments we use the ARIMA model for forecasting.

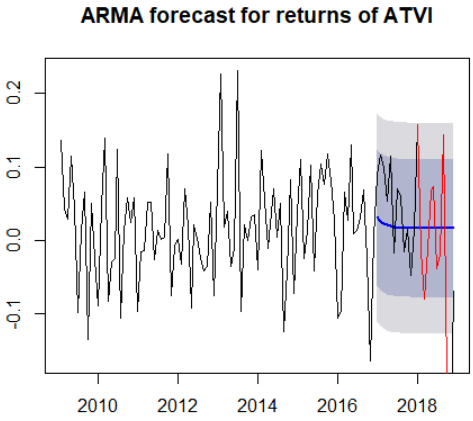
We first decompose our stocks to analyse Trends and Seasonal growth.

Due to some problems the auto.arima did not seem to be working and has not been used we instead loop through *some* possible values for the AR,MA parts of the function

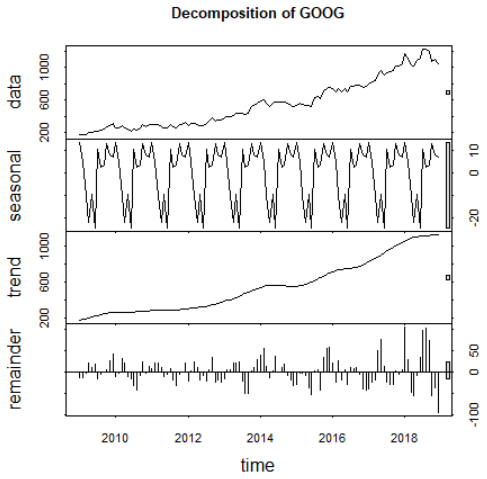
Follows the result of the analysis:
Blue line indicates the predicted values
Red line indicates the validation values
Training period: 80%
Testing period:80-90%
Validation: 90-100%

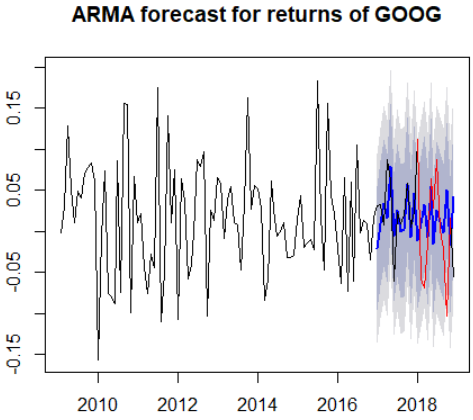
### 3.2.1 ATVI



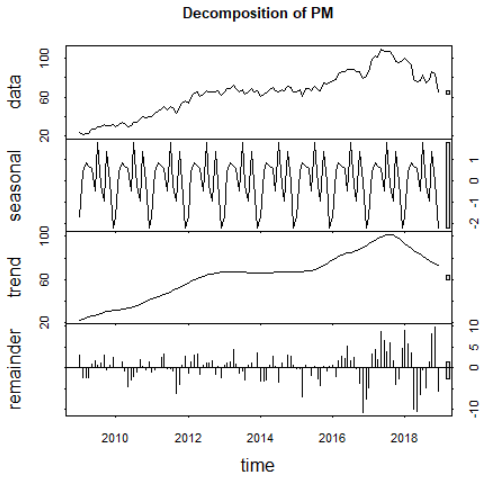



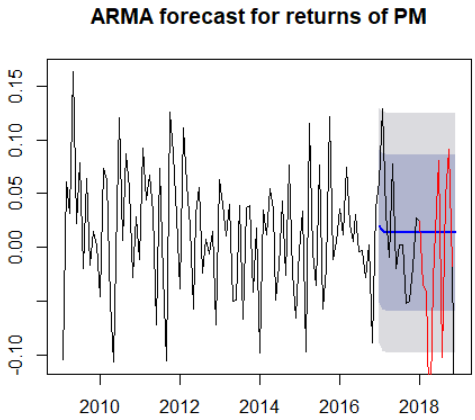

### 3.2.2 GOOG



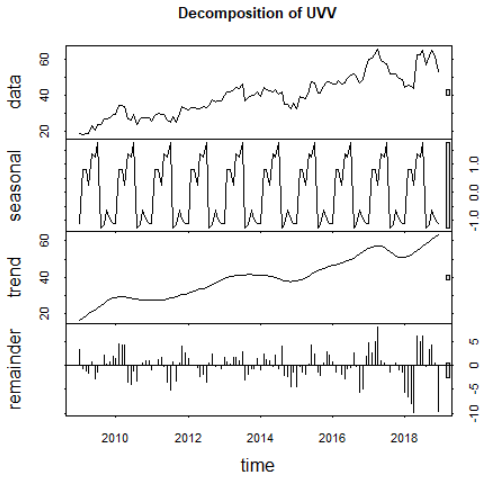



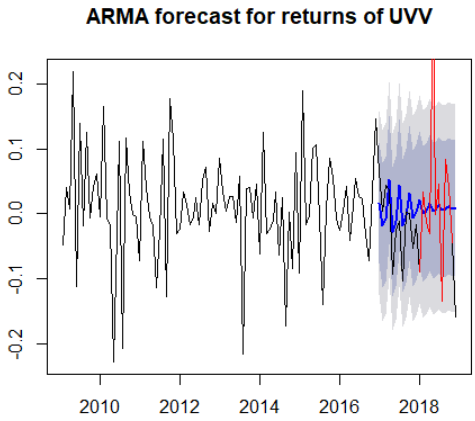
### 3.2.3 PM



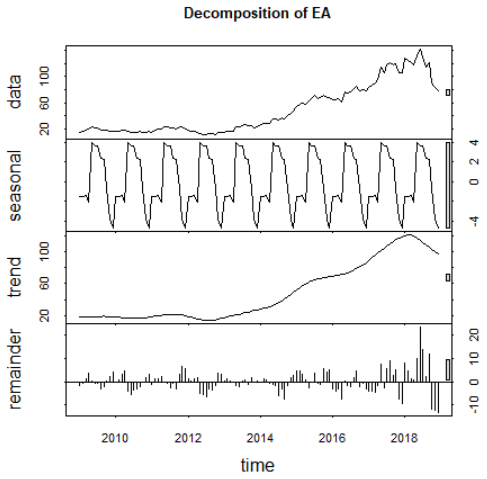



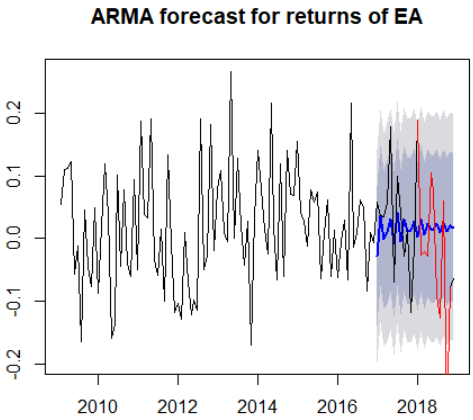

### 3.2.4 UVV



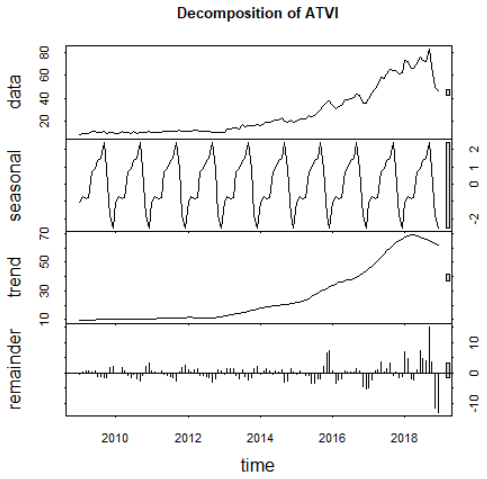



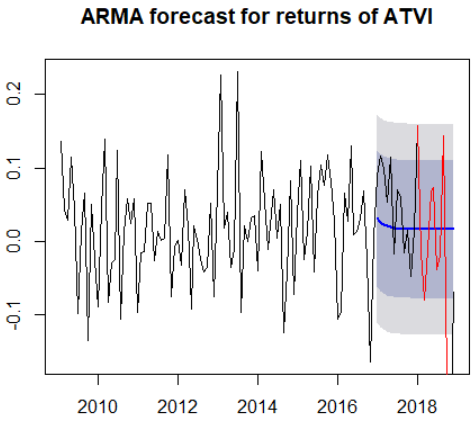
### 3.2.5 EA




### 3.2.6 ATVI



# Chapter 4

# Portfolio Management

" Portfolio management is the art and science of making decisions about investment mix and policy, matching investments to objectives, asset allocation for individuals and institutions, and balancing risk against performance."[1]

For the following part the following libraries have been used:

1. tidyquant
2. portfolioAnalysis
3. ROI

and a few minor more PortfolioAnalysis, with the ROI library allows for automatic portfolioOptimization.
The required portfolio is a MeanVariance Portfolio which maximizes returns while having a minimum variance

Listing 4.1: Portfolio and Constraints creation

```
p <- portfolio.spec(assets=fund.names)
#The default weights without using any constraints
print.default(p)

rp.seq <- generatesequence(min = 0, max = 1, by = 0.002)
p <- portfolio.spec(assets = fund.names,
                    weight_seq = rp.seq)
p <- add.constraint(portfolio = p, type = 'weight_sum',
                    min_sum = 0.99, max_sum = 1.01)
#p <- add.constraint(portfolio = p, type = 'box', min = 0, max = 1)
```

[1] https://www.investopedia.com/terms/p/portfoliomanagement.asp

```
p <- add.constraint(portfolio = p, type = 'box',
                    min = 0.05, max = 0.50, indexnum = 2)
p <- add.objective(portfolio = p, type = 'return', name = 'mean',
                   multiplier = 0)
p <- add.objective(portfolio = p, type = 'risk', name = 'StdDev')
```

We first optimize our portfolio via the *optimize.portfolio* function, rebalancing the portfolio quarterly.

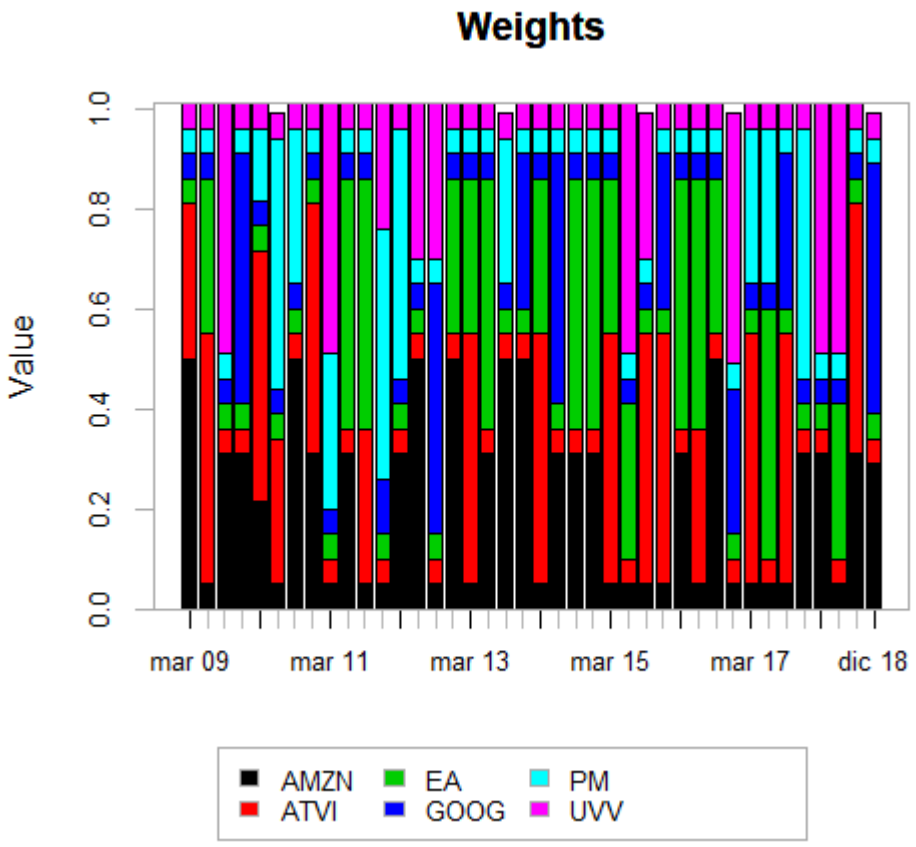



Figure 4.1: Chart of the weights' distribution over the rebalancing periods

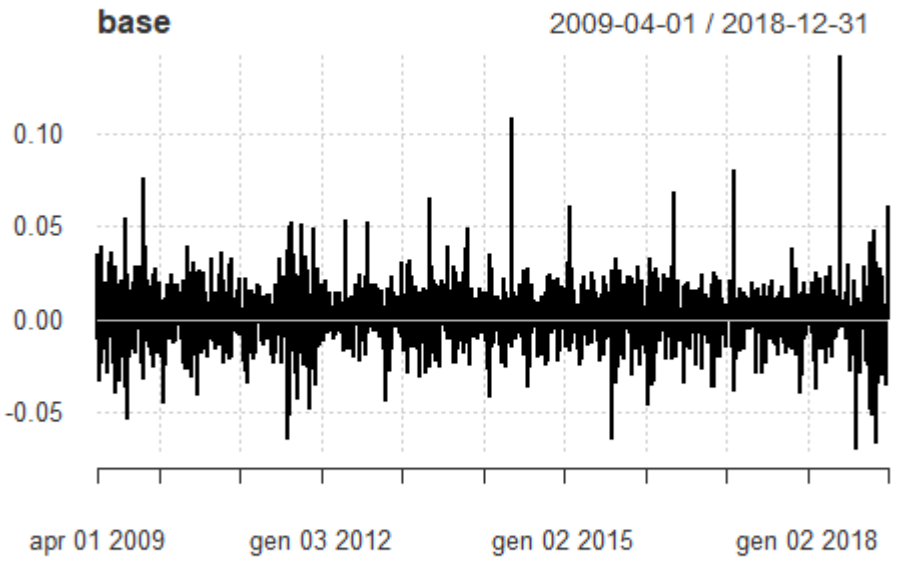



Figure 4.2: Plot of our portfolio returns

Finally we show how a small budget grows over the time period using this portfolio
For the rebalancing costs we considered a fee of 1% at every rebalancing period.

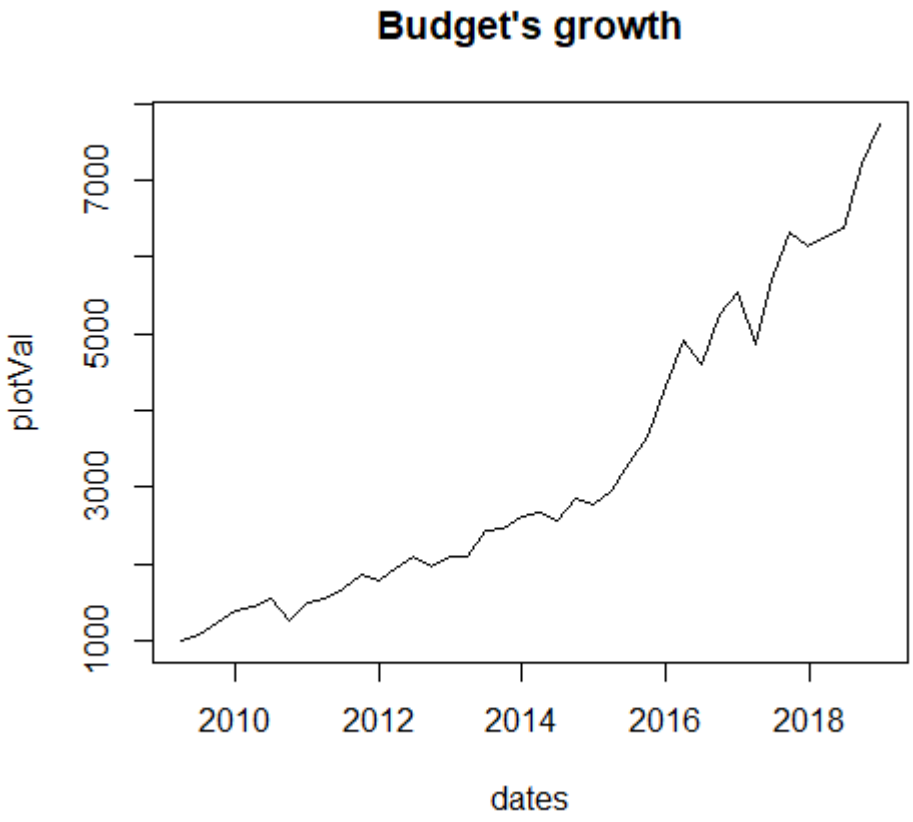



Figure 4.3: Plot of a starting budget of 1000$ showing its growth

# Part II

# Web App

# Chapter 5

# WebApp

The WebApp part of the project requires to create an RShiny webapp and to use its features to support one of the previous tasks.

In addition to the RShiny library I decided to use the shinydashboard library to add a nice-looking layout.

To populate the dashboard we need three main things:

1. The header
2. The sidebar
3. The main body

The header is used to indicate the title of the webapp
The sidebar is used to navigate between our features:

1. Descriptive analytics
2. Forecasting
3. Multivariate analysis

The body displays the features.

## 5.1 Descriptive analytics

The first section was a training ground for the webapp.
as in the following sections all the values have been changed to become *input reactive*, that is, the data will only update upon pushing the *update button*.
This section allows the user to choose **one** stock and to get basic data from yahoo.
The data can then be plotted with the graphs seen in the previous sections.
A summary of the data can be also be found

## 5.2 Forecasting

The forecasting section allows the user to choose a training period via the slider bar and to then plot both the forecasting graph and the decomposition graph.
The statistics box outputs the accuracy of the prediction (RMSE value)

## 5.3 Multivariate

Multivariate analysis allows the user to plot either dygraphs(which requires a different plotOutput in the shiny library) or scatterplots. both dygraph and scatterplots chart the returns of the desired stocks.
A limit of 4 stocks has been used to limit the size of the scatterplot.