RIGA TECHNICAL UNIVERSITY

Faculty of Engineering Economics and Management

Subject:

Business Analytics

Supervisor:

Ilze Birzniece

Project title:

“Analyzing stock market for decision-making and predictions”

Student: Marlen Dreyer

Student ID: [redacted]

Email address: [redacted]

Student: Séraphin Henry

Student ID: [redacted]

Email address: [redacted]

Student: Thomas Martin

Student ID: [redacted]

Email address: [redacted]

Student: Adil Mohammad

Student ID: [redacted]

Email address: [redacted]

Table of contents:

List of Figures

List of Tables

1. Introduction

1.1. Stock market and trading

Nowadays, a lot of successful companies make the choice to "go public". This means that they chose to put the shares of the company on a market, where people are ready to buy and then eventually sell them, under the name of "stocks", in order to make some profits.

The stock of a company is basically a share of the company, that constitutes its ownership.[1] Owning a share also means owning a fraction of a company and being entitled to a fraction of its earning. The market where people trade these shares is called "stock market".[2] All these transactions happen in stock exchanges, physical or virtual sites, such as the New York Stock Exchange and the Nasdaq. The stocks have a fluctuating price, which is defined by the basic law of the offer and the demand: the price of a stock might rise if the demand is high but will decrease if there are too many offers.

There are many things which can influence a trader to buy or sell a stock at a certain value: the company reputation at a certain moment, a trend regarding a sector, the increase or decrease of another stock value. Possibilities are numerous, and while a lot of them are unpredictable, some of them can be anticipated through analysis of statistics or pattern detection in the plots, which will give information about the characteristics and the current state of a stock. In the end, the goal for investors is to buy shares of company that are bound to be successful. Either they make money thanks to the benefits the company make, or by re-selling shares when their value is higher than at the time they bought them.

1.2.Problem Statement

Trading and investing in shares are a common form of investment for both companies and individuals, that can lead to making huge profits. One of the reasons that justify the enthusiasm of companies and individuals for this activity is that, in certain cases, the chances of making a profit are higher than for example with savings books.

To proceed to stock exchange, buyers and sellers must proceed in a defined period, depending on the exchange place of the shares they sell or buy. For example, the New York stock exchange (known as Wall Street or NYSE) opens at 9:30 am and closes at 4 pm.[3]

Consequently, the most important aspect is to **choose the right time for investment** in relation to the development of the company, so that no loss occurs.

In this context, the choice of the right company is also an important decision factor, although this is not considered in this paper. Fortunately, there are several techniques to take a decision about how money should be invested. But these techniques can reveal themselves more or less efficient according to the current health state of the company and its stocks, or to the field of activity of the targeted company, and it requires a detailed analysis to figure out which are the most relevant focus points that can help with decision making, knowing when it is the right time to buy or sell, or simply "hold", which means not doing anything.

[1] Investopedia (2020) (b).
[2] Investopedia (2020) (d).
[3] Nasdaq (2020).

2. Objective and purpose of study

The objective of this work will be to analyze various companies' history in stock market and find which are the most relevant metrics, attention points, methods and techniques to help with decision making, when it comes to buying or selling stocks. To do so, the focus will be put on a restricted selection of methods applied to a selection of companies coming from two different fields: the automotive and the social media industry. The first is the oldest industry and the other is an up-and-coming field that has risen in the past years. In this context data analysis can help finding a suitable time for investments, regarding any stock. A stock price evolution, or at least its probability of occurrence, can also be predicted from an analysis of the current state or the history of a stock. However, other kind of data can also be used, such as news articles, to analyze the link between the company's reputation and its stock value.

In addition, a distinction is made between two main purposes. On the one hand, it should be possible to help at decision making, buying, or selling at a particular moment. And on the other hand, it should be possible to give metrics and focus points in order to predict when to buy or sell in a near future, to detect a potential incoming increase or a decrease of a stock value.

Eventually, the direct benefit of this analysis is an increase of the potential profit from an investment made by companies, investors, day traders or swing traders, by focusing on the most relevant focus points of the targeted stock. Due to the variety of trading opportunities and strategies, it can have an enormous value contribution for many participants.

3. Proposed Framework for Metrics and KPI

3.1. Basic stock market values as metrics

Each stock has different characteristic values. For the sake of the study, it can focus on two basic ones: actual stock market value and the close market value.

Close market value is the value of a share at the end of the day. Stock exchanges close at a certain time. For example, Nasdaq usually closes at 4pm. After this time, trading shares is impossible and stock values are frozen to a value that they will keep overnight, until stock exchange opens again.

The first metric is related to the **Return On Investment,** or ROI[4]. It is a performance measure used to evaluate the efficiency of an investment. In this case, it will represent the gains, or the loss generated by a trader or company thanks to the study.

The second one is the **evaluated risk** associated to a stock, when it comes to investing. The risk is based on the stability of a stock, and the distribution of the return of investment on a defined period. The more unstable it is, the riskier it gets to invest in such a stock. So, just as the Return On Investment, it is a very important metric to look at in order to provide a relevant analysis which will eventually help decision making.

3.2.Key Performance Indicators

For this study, the main KPI is the **correlation value**. Throughout the analysis, correlations can be useful to figure out if the figures provided by a calculation or some findings are correlated to the stock price evolution. This way, it can target the most relevant focus points that can help with the analysis and, in the end, with decision making.

[4] Investopedia (2020) (c).

4. Dataset building and tools used

4.1. Origin of the Dataset and tools used

To build the dataset, **Finnhub** was used. Finnhub is a company based in New York that, according to their website, “offers a real-time (web) API for stocks, forex, and cryptocurrencies.” An API is an acronym for Application Programming Interface. It is a way of communicating with a particular computer program or Internet service.[5] Finnhub’s API gives access to real-time data from stock exchange. This ease of access and the relevance of their data were keys to building the dataset used for this study.

Data analysis was carried out using **Pandas**, an open-source data manipulation and analysis library for Python.[6] The graphs used in this study were built using **Matplotlib**, a comprehensive library for creating static visualizations in Python.[7]

4.2. Explanation of the Dataset

The dataset contains various stock prices from eight different companies, and is divided into different attributes, which will be briefly described in this chapter.

The attributes "Open", "High", "Low", and "Close" always refers to the market value of the day in stock trading. Accordingly, "Open" and "Close" are the prices at which a specific stock started and ended trading for a specific day. In addition, "High" and "Low" correspond to the highest and lowest prices for a specific day. Eventually, the column “Volume” corresponds to the number of shares traded during the day. The considered date is indicated in the present data record by the column "date" or “t”.

Even though all these columns are present in the initial dataset, this study will mostly focus on closing prices as it states the final value of a stock at the end of a trading day, as well as the date column. As already mentioned, the present data set contains history of stock prices of eight companies, on a period going from November 2019 to November 2020, divided into two sectors: four companies from the automotive industry and four companies from the social media industry. These two industries represent both the oldest and one of the newest industries. The companies from the different industries are the following:

[5] Cambridge Dictionary (2020).
[6] Pandas (2020).
[7] Matplotlib (2020).

From automotive industry		
Abbreviation	Designation	Description
PUGOY	Peugeot Société Anonyme	PSA is a French multinational automobile manufacturer. It regroups the Peugeot, Citroën, DS, Opel and Vauxhall brands.[8]
VLKAF	Volkswagen Group	Volkswagen is a German car manufacturer group. It comprises twelve car brands.[9]
BMWYY	Bayerische Motoren Werke AG	The BMW Group is an automobile and motorcycle manufacturer focused on premium products.[10]
GM	General Motors Company	GM is an American automobile manufacturer that owns four major car brands in the U.S.[11]
Social media companies		
FB	Facebook Inc.	Facebook is a US-American company and was founded as a social network. It is a kind of meeting place on the Internet where people communicate with friends and acquaintances.[12]
TWTR	Twitter, Inc.	Twitter is an American company developing and maintaining the eponymous social network. It is a platform allowing users to post short texts or media content.[13]
SNAP	Snap Inc.	Snap is an American camera company. They develop mobile apps, notably Snapchat, a social network allowing users to share pictures and videos to friends or globally.[14]
PINS	Pinterest Inc.	Pinterest is an American camera company and a social network that allows users to share pictures.[15]

Table 1: Description of companies

[8] Peugeot (2020).
[9] Volkswagen AG (2020).
[10] BMW AG (2020).
[11] General Motors Company (2020).
[12] Facebook (2020).
[13] Twitter Inc. (2020).
[14] Snap Inc (2020).
[15] Pinterest (2020).

5. Data Analysis, Pre-processing, Cleaning, and Transformation

To ensure that sufficient analyses can be carried out with the available data set, it was first necessary to clean the data set so that a high quality of the results can be achieved.

As said previously the dataset contains stock information of eight companies. Four of them are considered as social media companies and the other four are from the automotive sector. Concerning Facebook Inc., Twitter Inc., Snap Inc. and Pinterest Inc. it was not necessary to clean the data. As it stands, these data will be usable. However, for companies in the automotive sector a delta between daytime has been noticed (about 14 minutes delta between different companies). In order to be able to compare the values it had to be cleaned by matching the data daytime.

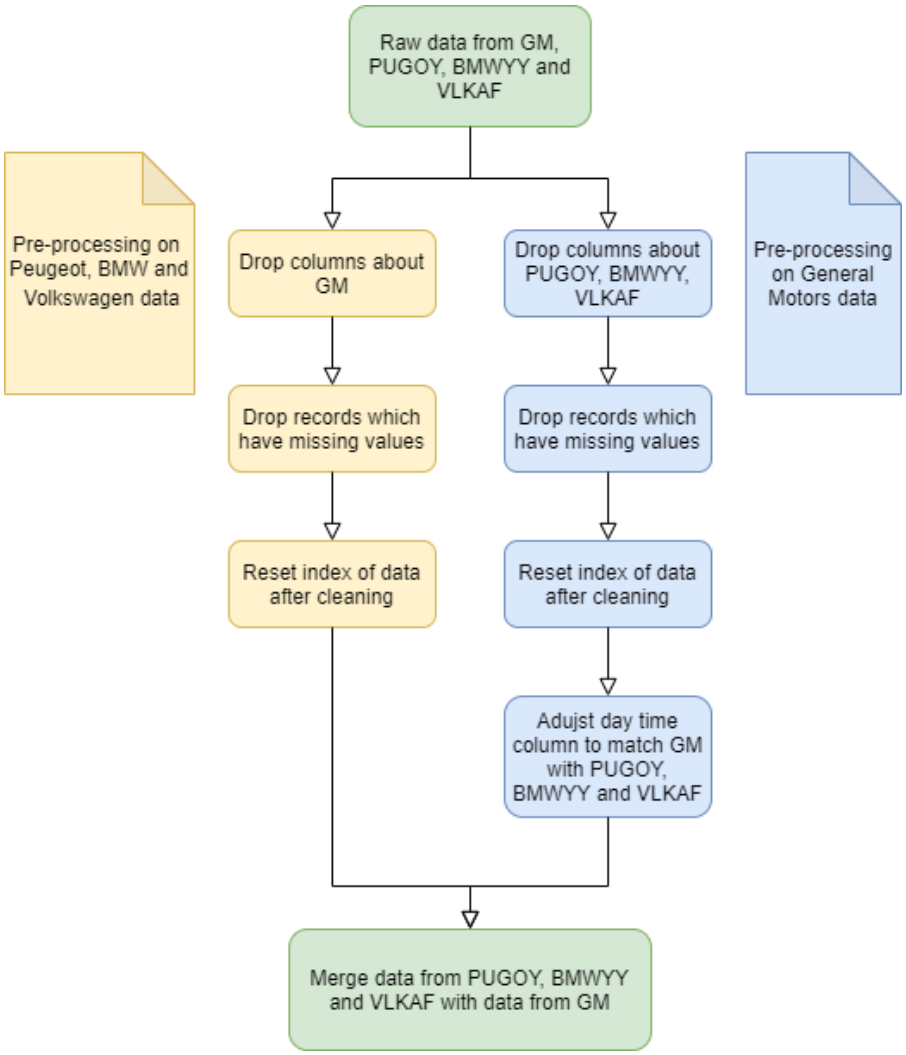

Figure 1: Flow diagram of pre-processing

6. Analysis and Findings

6.1. Data analysis to understand stocks behavior, risks, and trends

Before deep diving into analysis using specific statistical techniques and various stock market strategies, an introductory descriptive data analysis is necessary for this study, as it is important to get to know these eight companies stocks a bit better through different aspects.

Indeed, when it comes to investing, the recent evolution of the stock, the recent returns of investment it provided as well as the risk associated are relevant focus points. First, it is interesting to focus on the development of closing prices over one year for both industries.

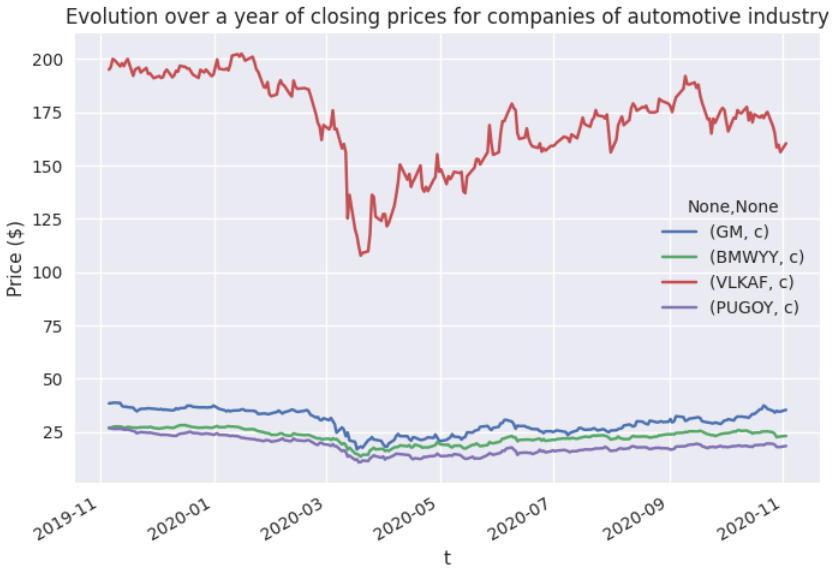

Figure 2: Evolution over a year of closing prices for companies of automotive industry

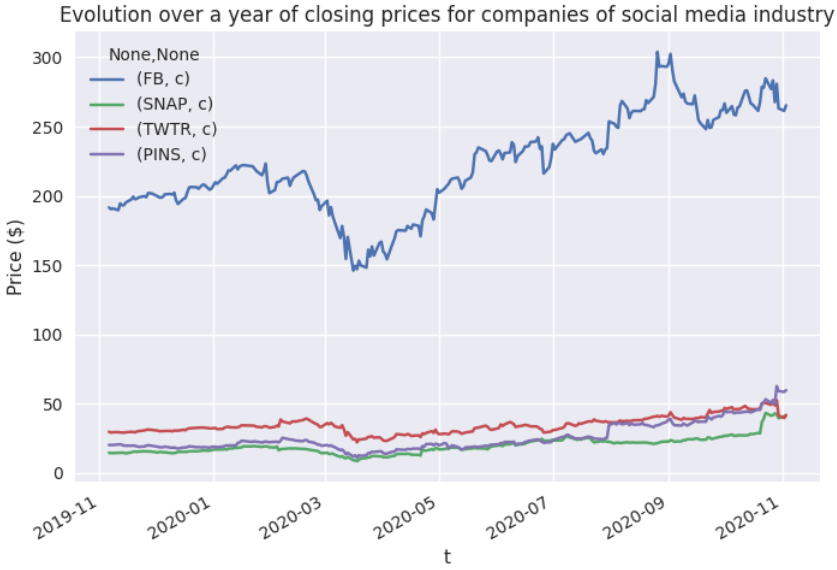

Figure 3: Evolution over a year of closing prices for companies of social media industry

These two graphs allow to compare the stock closing prices evolution of companies from two different industries: automotive and social media. Among the four chosen companies for each industry, it clearly appears that Volkswagen (VLKAF) and Facebook (FB) position themselves as the leaders of their respective field of activity.

However, Facebook stocks are more valuated than the Volkswagen stocks, as well as the other social media companies compared to the automotive companies. The first hypothesis is that investing in social media companies' stocks is a bit trendier these times, compared to automotive stocks. Indeed, this fad for social media industry, even allowed these companies to survive a bit better to the highest point of the coronavirus crisis (around March-April 2020). Facebook stocks for example, even had a huge increase since then.

Studying closing prices evolution for a stock is interesting in order to see the evolution of a stock value over a defined period, and also to help detecting which stocks are actually trendy and which are not.

However, the evolution of a stock value can also be visualized through its performance evolution over a period, as it is a more straight-forward manner to figure out if the stock value has been really improving or not.

To do so, every closing prices of the period has been divided by the very first closing price of the period, and then multiplied by one hundred, for each stock. This way, it can be visualized the performance evolution for all the stocks on both industries over a year period, from November 2019 to November 2020.

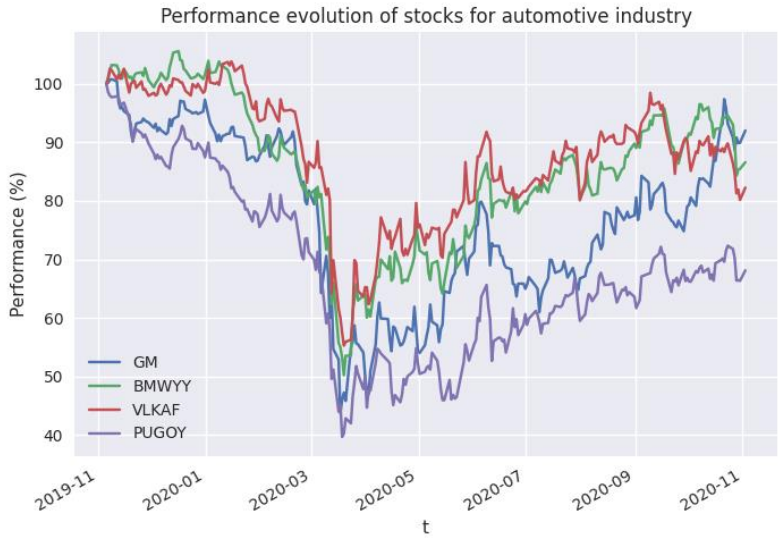

Figure 4: Performance evolution of stocks for the automotive industry

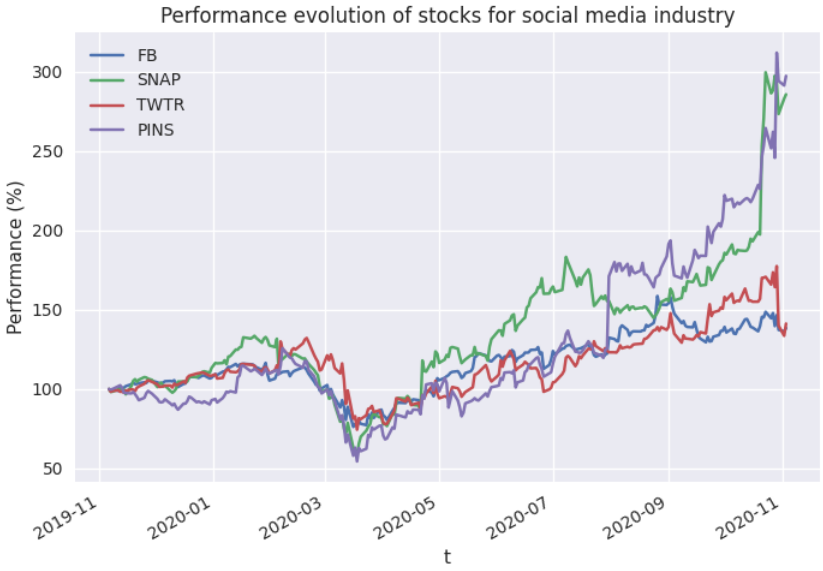

Figure 5: Performance evolution of stocks for the social media industry

The two graphs show the evolution in % of the closing prices for the two industries. By reading these graphs, it appears that the automotive industry suffered a lot of the March-April drop, as the stocks struggle to get back to their previous closing values. On the other hand, the social media stocks are going higher and higher, even going for a 300 % increase of the stock values in a year for Pinterest (PINS), going beyond the performances of Facebook, the leader of the social media stock market. All the picked social media companies saw their stock values increasing by at least 150 % over a year, where automotive companies saw their stock values diminishing by at least 10 %.

Performance of a stock can also be studied through the distribution of its return on investment values. Indeed, if return values are mainly distributed on the negative part of the scale, it means that the stock is not that worth for investment.

To visualize the performance of the stocks chosen for this study, the distribution of their daily returns was plotted over the past year, which is shown in the following figures.

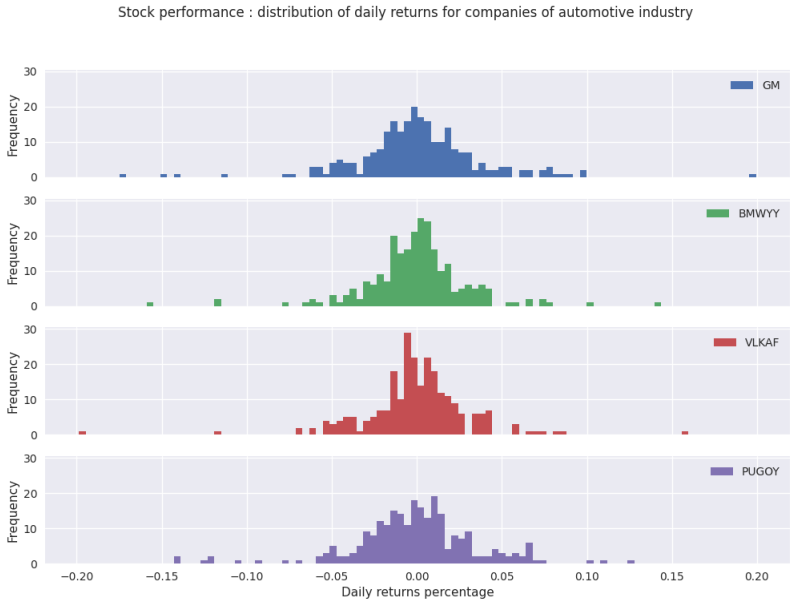

Figure 6: Stock performance: distribution of daily returns (automotive industry)

Figure 7: Stock performance: distribution of daily returns (social media industry)

The daily returns distribution also informs about the stock volatility, and it allows to give an idea about how likely investing in a stock will either bring profits or cause losses. Here, by comparing these two graphs, it appears that daily returns for automotive stocks are a bit more volatile compared to the

other industry, meaning that there are more risks regarding return on investment. Indeed, on the daily returns for social media stocks, the values are more regrouped on the performance scale.

Then the daily returns' means of stocks from both industries studied were plotted to get an idea of their performance over the past year. In addition, it is also useful to have an insight of the daily return's means of these stocks, in order to have a more general vision.

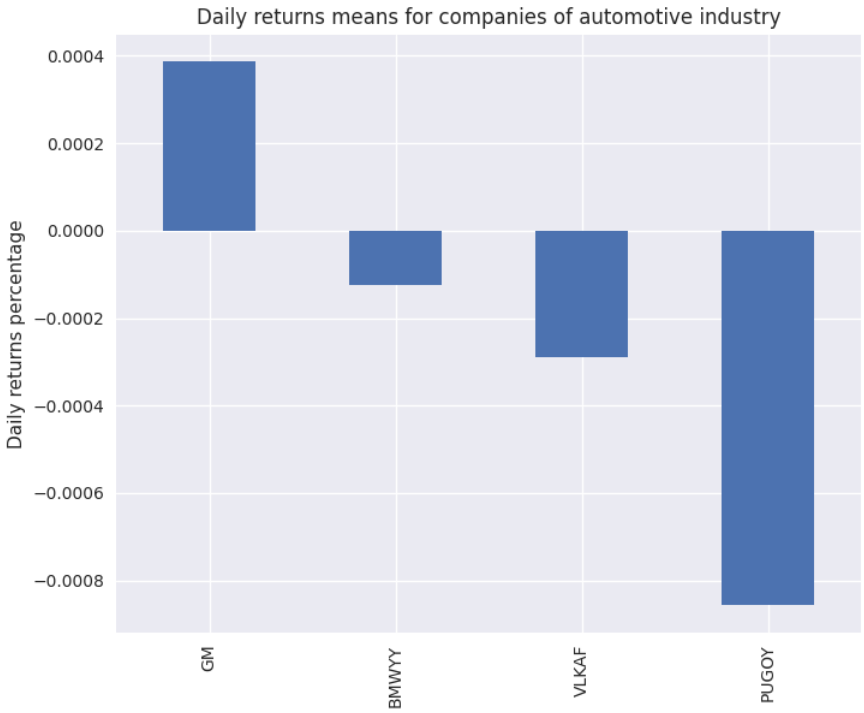

Figure 8: Daily returns means for companies of automotive industry

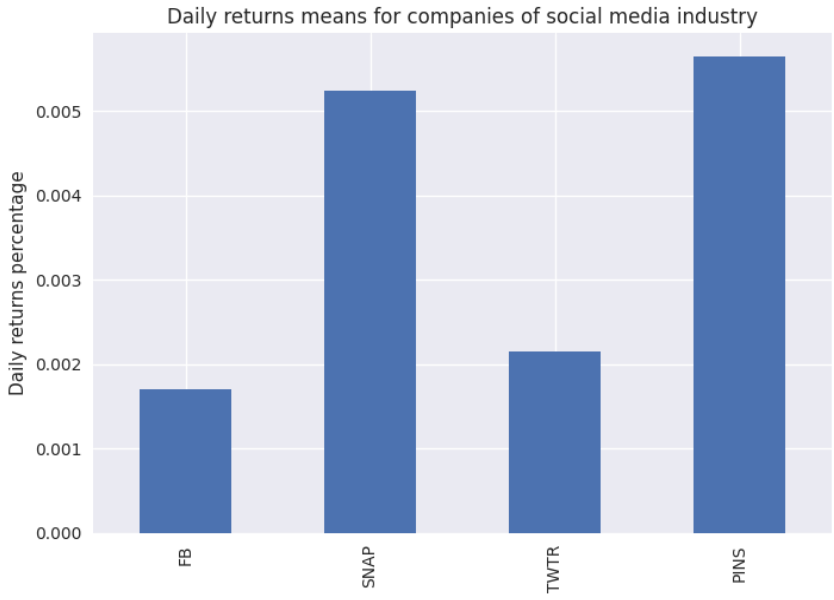

Figure 9: Daily returns means for companies of social media industry

To measure the volatility of each stock, the standard deviation of each stock was spotted. Finally, the standard deviation can give information about the volatility of the stock, which ultimately give insight of the stability of the stock.

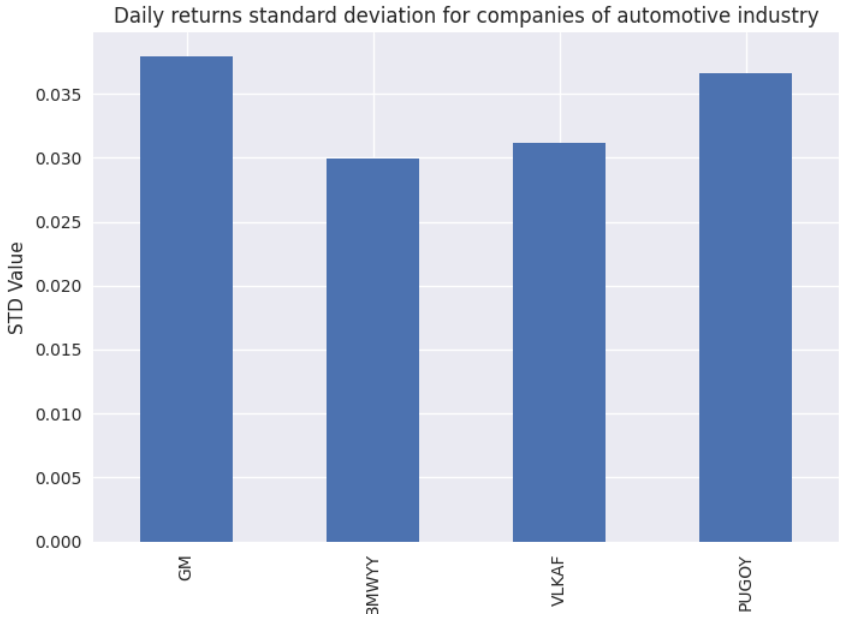

Figure 10: Standard deviation daily returns (automotive industry)

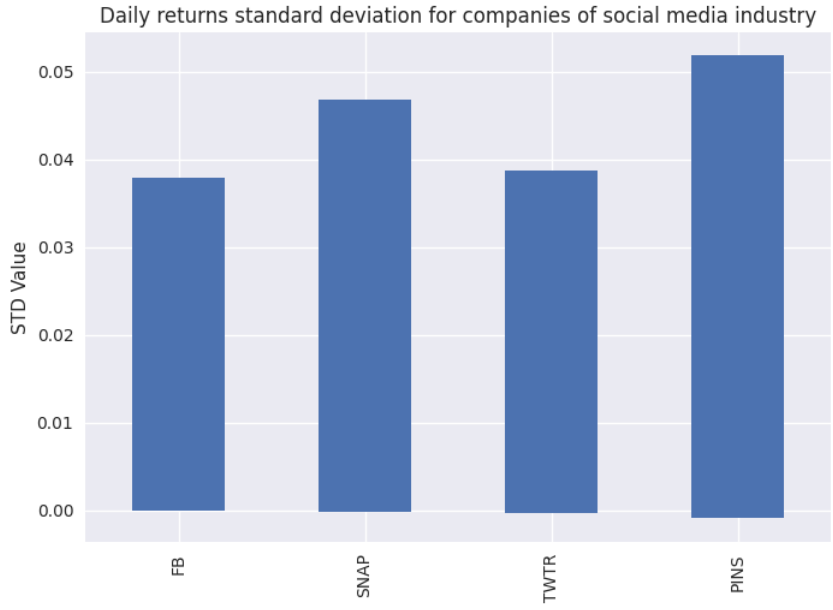

Figure 11: Standard deviation daily returns (social media industry)

These mean and standard deviation values give a clearer picture of what is represented in the daily returns' distribution graphs. Indeed, average daily returns is lower for all the picked companies of the automotive industry, and even if General Motors (GM) stocks are the only one above 0, the standard deviation tells that this stock is the most volatile of the four automotive companies.

In short, there is a great probability of loss or making profit by investing in this stock, however regarding the fact that the average daily return is very low (0.0004 % of profit expected on average), this might not be that worth it.

On the other hand, the social media stocks are also very volatile, even more than automotive stocks, and that is due to their recent increase. So, even if the average daily returns are good, it is still risky to invest in these companies, because it could drop at any time.

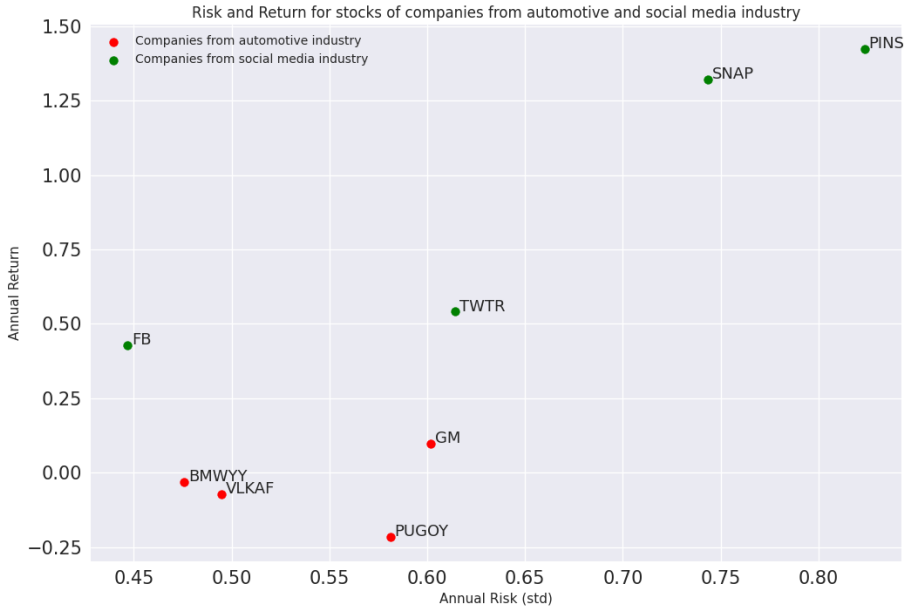

Figure 12: Risk and Return for stocks of the companies

This graph summarizes what it has been mentioned above about the intricacies of trading stocks from automotive and social media industries. The lower a stock is, the worse its annual return was. The further on the right it is, the highest the risk of losing money after investing on it. The stocks of the automotive industry present low annual risk, however only one among those studied had a positive annual return, GM. Although this stock seems profitable, it is the riskiest one from the automotive industry. Companies from the social media industry all have a positive annual return, ranging from 50 % for Facebook to 150 % for outliers such as Pinterest and Snapchat. Looking solely at these numbers would indicate that investing in Pinterest and Snapchat might be the best option because it has the potential of generating the most profits. It is important nonetheless to consider their annual risk, whose value is ranging from 0.75 to 0.8 meaning these two stocks are highly volatile. All in all, it appears Facebook is the best stock to invest in, having a positive annual return and the lowest risk.

In conclusion, this first introductive analysis gives interesting information to learn more about the current situation regarding both industries, on the stock market. Social media companies are interesting stocks to look at, due to their recent increase, and their overall daily returns. However, some of these companies are relatively new to stock market, and their volatility remains quite high, bringing a lot of risks when it comes to investing. On the other hand, it is also important to note that this recent increase might be due to the actual sanitary situation, where a lot of companies and people organize their life around online and remote events, making the value of social media and “internet solutions” increase.

Regarding automotive industry, the coronavirus outbreak really had a huge impact on stock values, as since March-April 2020, automotive stocks have not fully recovered their pre-crisis initial value. In the

end, this analysis shows that it is not the best time to invest in any of these four automotive companies, as their daily returns mean is too low. However, the stocks are still increasing, and seem like they will soon be back at their initial pre-crisis value, so it clearly is a stock to look at for future investments.

Furthermore, an understanding of the correlation between the individual companies within an industry is very interesting.

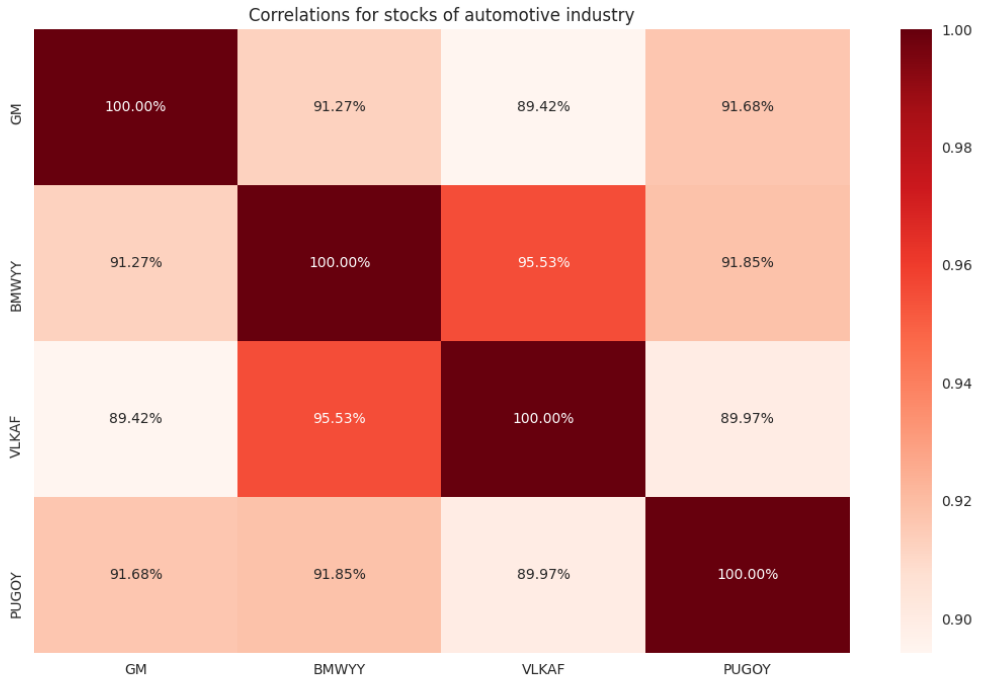

Figure 13: Correlation for stocks of automotive industry

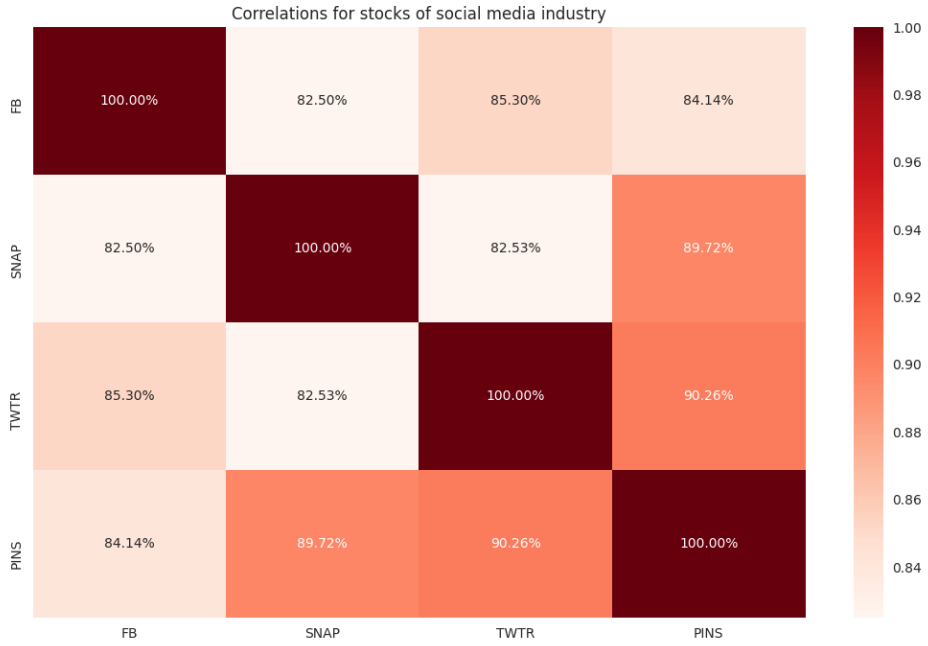

Figure 14: Correlation for stocks of social media industry

Figures show the correlation analysis between automotive stocks and social media stocks. Stocks can correlate positively if they move up or down in parallel. A correlation value of 1 (indicated by the color dark red in the figures) corresponds to a so-called perfect positive correlation. In concrete terms, this means that as soon as one considered stock of a company rises, the other considered stock of another company also rises. On the other hand, a perfect negative correlation means that one stock moves up and the other stock of a company moves down. For the two industries under consideration, it can be seen at first glance that automotive stocks correlate somewhat stronger than social media stocks. In the automotive industry, the correlation values generally differ only marginally within the four companies considered. Only the correlation between BMW and Volkswagen, at 95.53 %, is an almost perfect correlation. On the one hand, this could be because both operate in the same countries. On the other hand, the two car companies are characterized by a similar product range. Furthermore, the development directions of both are very closely related (e.g. regarding e-cars). If the car industry is doing well, both companies will rise and vice versa. An almost perfect correlation means for the automotive industry in general that an international pandemic such as Covid19 will have an equal impact on performance. This development was confirmed in the performance evolution (Figure 4), which was weak for all four companies between March and May 2020 in the wake of Covid19.

Although there is also a very strong correlation in the social media sector, the differences between the four companies considered are greater than in the automotive sector. This may be since although they are in the same industry and have similar target groups and thus a strong correlation, not all companies have the same focus in their activities. For example, Snapchat is geared towards sharing photos and videos, while Facebook as a complete social network offers more than these functionalities. Thus, the lowest correlation in this case of 82.50% can be justified. The same arguments can thus be used to justify almost the same correlation between Snapchat and Twitter. The highest correlation is between Pinterest and Twitter with a value of 90.26 %. These can be explained by the fact that for example both stocks have lower values and are very close together. An investor who decides to invest in this type of

stock will therefore probably try to increase the chances and spread the risk in the industry by investing in both companies.

In the context of the analysis, it is questionable to what extent the correlation can help decision-making. The correlation analysis can help to gain an overview that for example in a crisis like Covid19 all companies in the automotive industry are affected almost equally. For the social media industry, on the other hand, it is less meaningful because the correlation is lower. The results therefore produce relevant data, but it is rather doubtful at this point how valuable this information is, since only four companies from each industry were considered.

6.2. Data analysis for decision-making

6.2.1 Definitions of trade triggers and Bullish Engulfing Pattern

When deciding whether to make a trade or not, traders analyze the current situation of the market looking for trade triggers.

A trade trigger is any event that meets the criteria to initiate an automated securities transaction. It is usually a market condition, such as a rise or fall in the price of a stock.[16] There are various trade triggers, that translate to different trading strategies. One of them is the Bullish Engulfing Pattern.

Stock charts are often depicted as candlestick charts.

Figure 15: Example of a candlestick chart

Each candlestick represents one day of trading. This kind of chart help to quickly visualize the trend a stock value is following by providing key data in a concise manner.

[16] Investopedia (2020) (e).

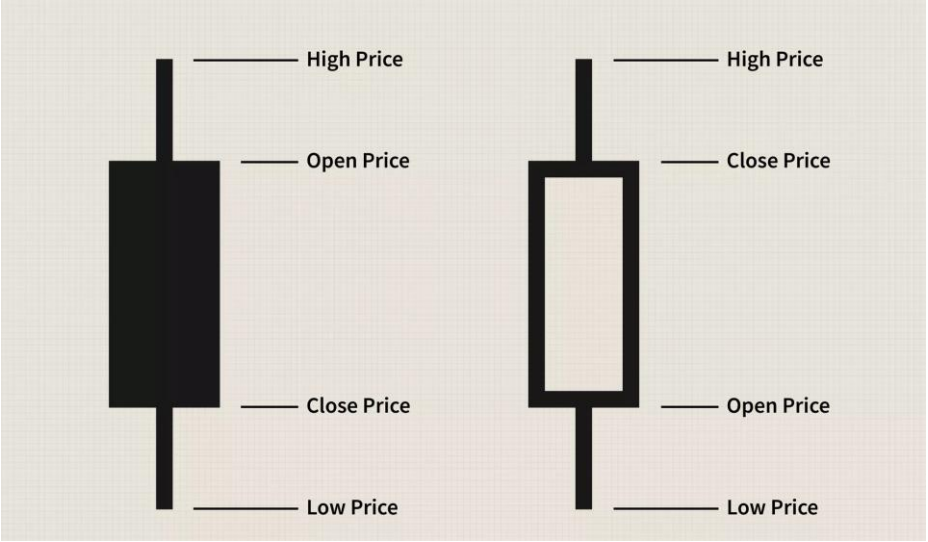

Figure 16: Meaning of the candlesticks

A candlestick is filled when the close price was lower than the open price. If the open price is higher, then the body is left empty.

The Bullish Engulfing Pattern is a candlestick chart pattern that takes shape as a small filled candlestick followed by a big shallow one the next day.

It basically is the moment when a stock value is about to go up after a negative trend. It is admitted that this pattern is more likely to occur if the shallow candlestick appears after several filled candlesticks.

Furthermore, for this pattern to appear, the stock must open at a lower price on Day 2 than it closed at on Day 1.[17]

Figure 17: Bullish Engulfing Pattern Example

In this example of Philipp Morris' stock, after an uptrend in 2011, the stock went down for a few days in the beginning of 2012, where a bullish engulfing pattern appeared as circled above. For traders, this pattern can mean several things: either buying as soon as possible is a certainty or waiting to make

[17] Investopedia (2020) (a).

sure it really means there is an upcoming positive trend. This trade trigger is common, and it might be considered as a decision-making tool.

6.2.2 Case examples of the Bullish Engulfing Pattern

As said previously, the Bullish Engulfing Pattern arises in a stock's value when certain conditions are met: the stock's candlestick chart displays one or preferably more black candlestick, meaning that each day the closing price is lower than the opening price, and a white candlestick appears.

To find instances of the Bullish Engulfing Pattern among the stocks studied here, an algorithm was used to identify span of days where it occurred.

The algorithm used to do so compared candles iteratively over the past year. Several of these instances fit the characteristics of the pattern, among them the case of Pinterest and BMW. It is important to note that these instances were only found for a few of the stocks studied here, and some of the cases of Bullish Engulfing Pattern the algorithm returned did not show any meaningful change of trends.

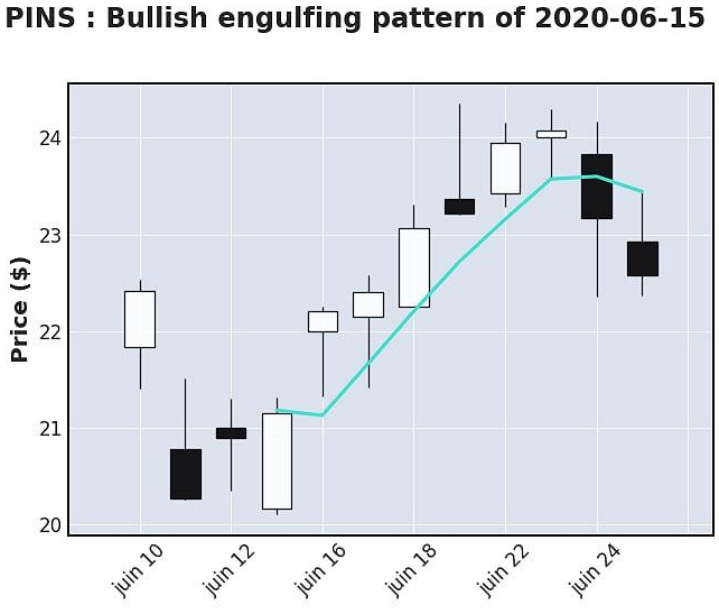

Figure 18: Bullish Engulfing Pattern in the case of Pinterest

As seen on this candlestick graph of Pinterest, a Bullish Engulfing Pattern appeared on June 15, after a negative trend lasted several days. It is shown by two black candlesticks, followed by a white one on June 15, and several others in the next days. The pattern heralded an upcoming positive trend.

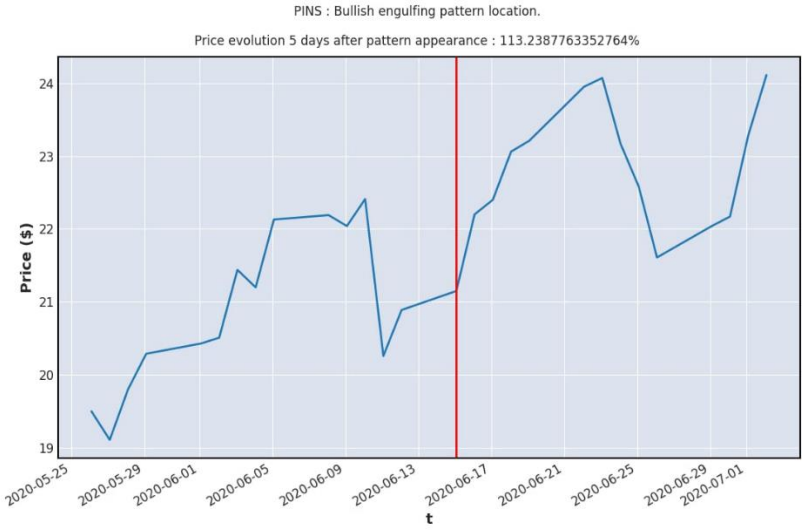

Figure 19: Pinterest's stock value evolution after the Bullish Engulfing Pattern

From June 15, the price of Pinterest's stock had a steep increase, gaining 13 % in the first five days. This stock being highly volatile as mentioned previously, the Bullish Engulfing entails a big reversal in trends.

The case of BMW is slightly different on that aspect.

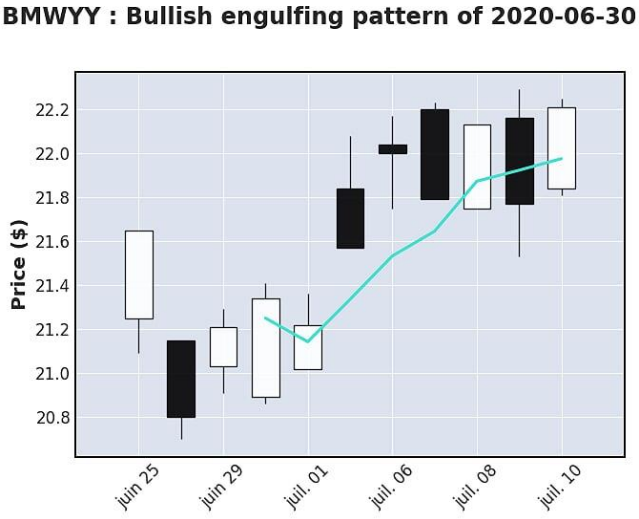

Figure 20: Bullish Engulfing Pattern in the case of BMW

BMW's stock showcased a Bullish Engulfing pattern on June 30. This case is different because the pattern occurred when a big white candlestick appeared after a small one on June 29. The trend reversal was also weaker from the one observed for Pinterest.

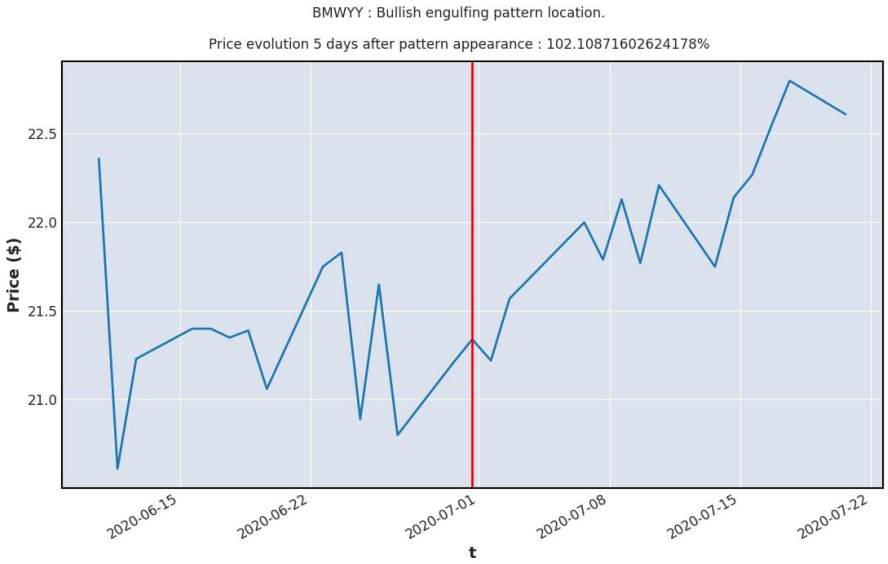

Figure 21: BMW's stock value evolution after the Bullish Engulfing Pattern

Five days after the Bullish Engulfing Pattern appeared, the stock only gained 2 %. BMW being a relatively stable stock, trend changes are not as pronounced as the ones in stocks from the social media industry. In this graph, it can be seen that the stock value of BMW hovers from $20.5 to $22.5.

Overall, being on the lookout for this pattern can be interesting. It is almost always synonym of a change of reversals, whether it is meaningful or not. Based on the stock, a Bullish Engulfing Pattern can foretell a significant gain in value.

The Bullish Engulfing Pattern is a great decision-making tool for buying stocks. Indeed, this pattern's appearance in the candlestick chart of a stock means the value of this stock is going to go up in the coming days, which means profits for an investor if they buy stocks when the pattern appears. It is even more effective on highly volatile stocks such as ones of the social media industry. The profits are potentially bigger than the ones an investor would get for automotive industry stocks. It is however important to point out that due to the nature of the Bullish Engulfing Pattern and social media industry stocks' high annual risk, this method is not fit for long-term investment. It is to be used as a tool to buy and sell stocks in a timespan shorter than a week.

6.2.3 Case examples of the Momentum trading

The present trading strategy observed is the Momentum Trading. This strategy of trading depends strongly on the good or bad shape of the stocks. Several tools can be used to determine the state of stocks such as Rolling statistics, Simple Moving Average (SMA), comparing plots with shorter and longer window.

Considering SMA, it is basically aggregating the mean over a defined period (window). Here, for graphs below, it is considered that one window is equal to one day. For the analysis, two windows are used. The first is at window = 10 (SMA 10), which captures the most localized and re-center trends. The second is at window = 50 (SMA 50), which captures a more general trend. With the momentum strategy, if the curve representing the SMA 10 is above the curve representing SMA 50, it might be good to consider buying.

Here only Facebook and Volkswagen are considered for clarity and because they are leaders of their respective groups.

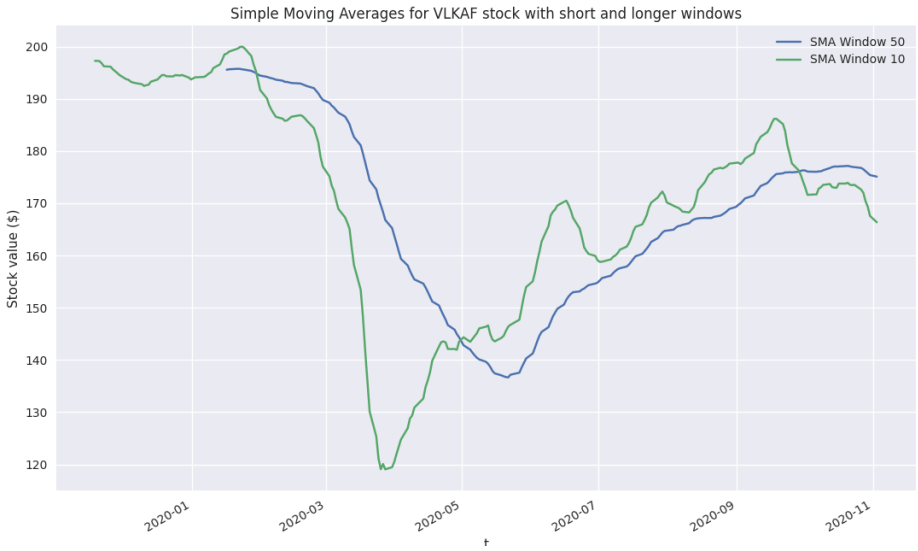

Figure 22: Simple Moving Averages for VLKAF stock

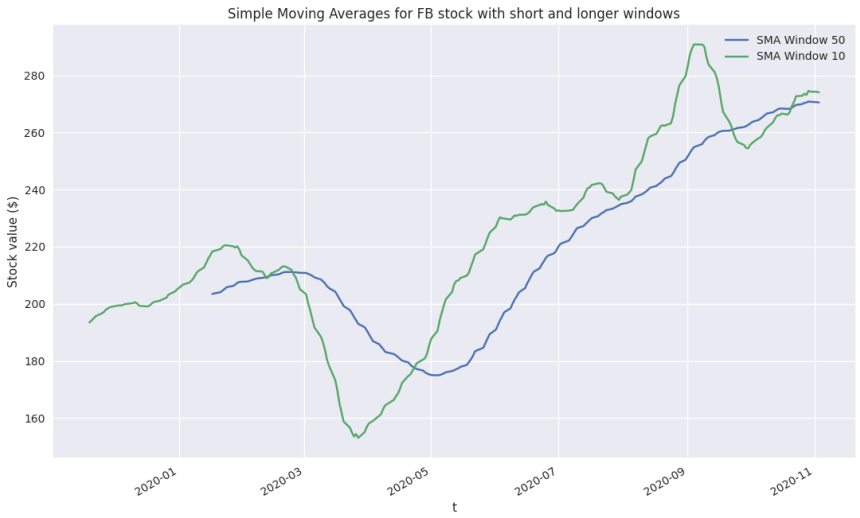

Figure 23: Simple Moving Averages for FB stock

These two graphs show the two waves of Covid19 have impacted the stocks. The SMA 10 curve is globally above the SMA 50 over the year, except during periods of high virus activity where the green curve goes below the blue one. Between late January/early February and late April for the first wave for both companies and for the second wave, the green curve goes back below the blue curve from mid-September for both firms too. A slight increase in Facebook is to be noted at the end of October because of its field of activity which is less impacted than the automotive industry. This is also readable thanks to the sharpness of green peaks during the waves, they are sharper for Volkswagen, where Facebook has smoother peaks.

All this shows it was very inadvisable to buy during these two waves of the virus. Even if the second wave had less impact on the stocks, it will be advised to act carefully at the end of this second wave and prepare for a possible third wave.

To help the decision making, the potential investment is another great tool. It can be calculated using the formula: $\frac{SMA\ 10}{SMA\ 50} - 1$.

It would allow the buyer to have another metric to check before taking any decision. So, if the investment rate is above 0, it means the SMA 10 is higher than the SMA 50 and that means buying is advised according to the strategy. If, on the contrary the investment rate is below 0, it means the SMA 10 is lower than the SMA 50 and buying is not advised.

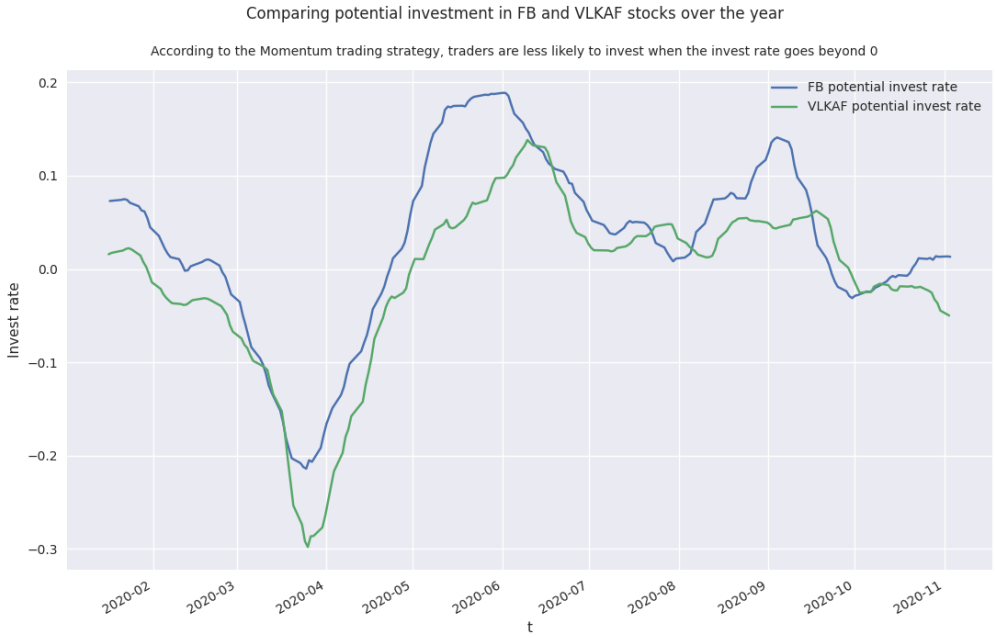

Figure 24: Comparing potential investment in FB and VLKAF stocks over the year

Finally, before taking this kind of decision, it is also great to have an overview on risks and returns with a short down, still using SMA with a window = 10.

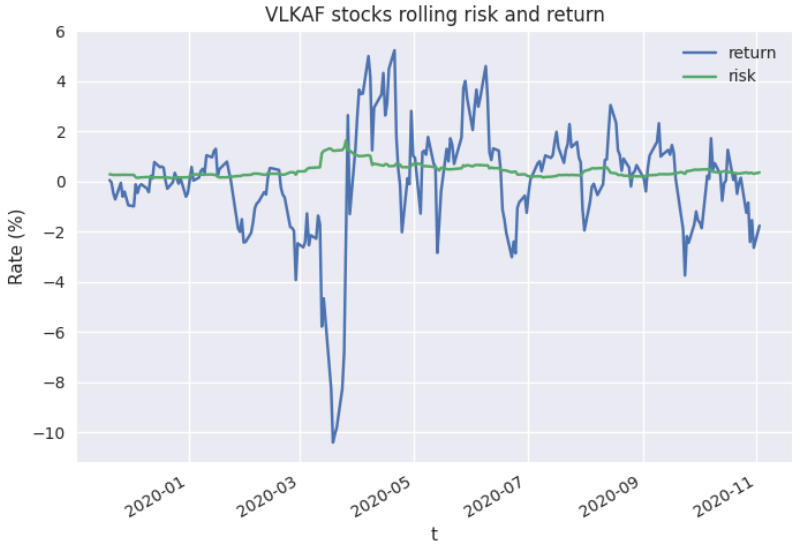

Figure 25: VLKAF stocks rolling risk and return

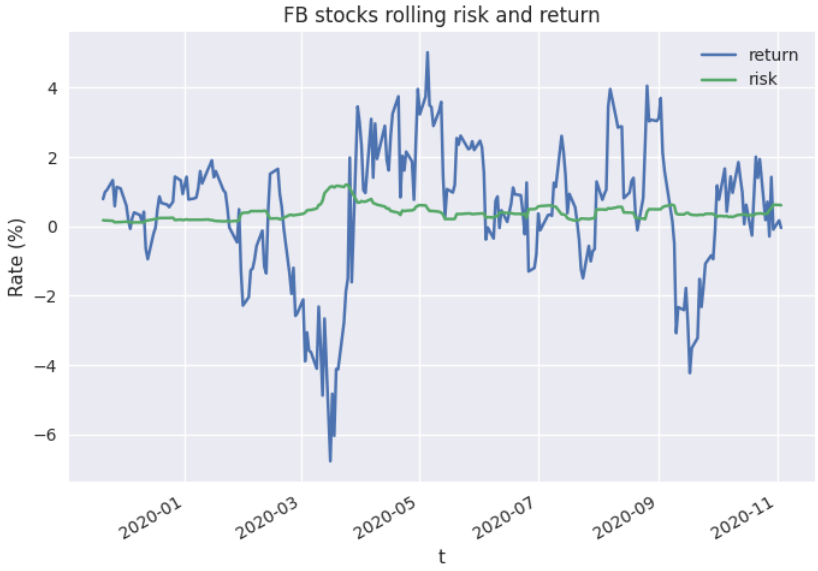

Figure 26: FB stocks rolling risk and return

These two graphs give an important insight regarding the risk to invest at given time. In fact, it can be a good metric to know if it is the right time to invest or not. Facebook for example has generally a return higher than zero. It means whenever someone invests in Facebook there is a high chance to have a positive return. However, these huge drops could be a real opportunity to invest because it increases the return if the curve goes back to a positive value.

For Facebook and Volkswagen, these numbers show once again the damages of these two waves of the virus. According to the Momentum Strategy, it would not be advised to invest but still, the opportunity to make a profit is palpable. Regarding the risk curve, it can be noticed that when the curve starts to increase, it often prevents a dive from the return curve. After a first wave around April, with the risk increased and the return dropped, it appears that around mid-September the second wave was more awaited since the risk did not increase as much as it did during the first wave.

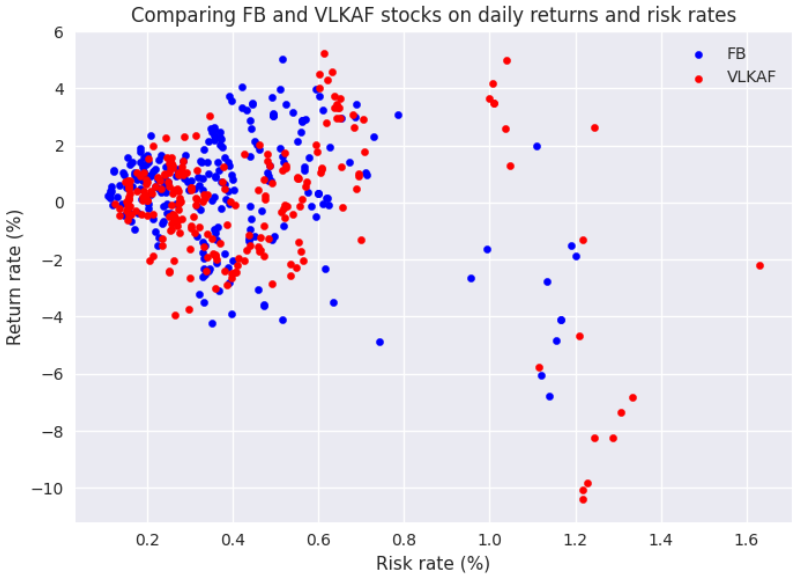

Figure 27: Comparing FB and VLKAF stocks on daily return and risk rates

Eventually, it is interesting to compare Facebook and Volkswagen to gather information whether to invest in Facebook or in Volkswagen. The graph gathers risk rate compared to the return rate. Each point represents a day of the health of the stock. After analysis, the graph gives a good image of Facebook and Volkswagen stocks. They seem to have their return rate around -4/+4 and their risk rate around 0.1/0.7. So, it is generally good to invest on these companies. However, outside these ranges, there are some dots, a major part being from Volkswagen. These dots are a direct consequence of the impact of the Covid19 where Volkswagen seems to have more troubles to recover of this situation than Facebook which seems to be in better shape.

In conclusion, there are several metrics to keep an eye on. The SMA, the return and the risk give huge information on the stock welfare. With these metrics, it can be safer to invest using the Momentum Trading strategy. Then the best moment to invest would be when the SMA 10 is higher than the SMA 50, the return has a positive value and the risk is under 0.8 %.

6.3. Data analysis for prediction

The objective of stock market predictions is generally to determine the future movements of the stock value of a financial exchange. To predict the market, the analysis of unstructured textual information such as from financial reports or news within the framework of fundamental analysis has proven to be efficient. This analysis is better known as text mining. Text mining strategies are used to extract important information for the analysis of market behavior. When predicting stock market values, online financial reports, or information other than social media are mainly used as source texts.[18]

A subfield of text mining is also the so-called sentiment analysis, or also known as opinion mining. This is an automatic evaluation of texts with the aim of identifying something written as positive, negative, or neutral. In this context, the sentiment score compares the number of positive reviews to the number of negative reviews. If the positive and negative reviews are the same, the sentiment score is 0.

[18] Alzazah, F. S. & Cheng, X. (2020).

The higher the sentiment score, the more a corresponding report has positively impressed a company. A negative sentiment score corresponds to a negative evaluation of the report for a company.[19]

The financial reports and news for all companies were analyzed with the help of text mining and the sentiment score was calculated by an algorithm. Since a presentation of all eight companies in this chapter would lead too far, the following section presents one company from each industry, which has the most concise results.

When comparing the development of the closing price of Peugeot with the development of the sentiment score, it can be seen that news articles have an effect on the development of the closing price. In August 2020 there was probably a bad article for the company (sentiment score of -0.50), which was also accompanied by a decline in the share price. Nevertheless, there were more positive articles in recent weeks, which were also linked to a price increase in November and December. However, there are still considerably more facts than news on the development of a company's share. For example, share prices fell towards the end of October, although at least three articles were rated positively for the company.

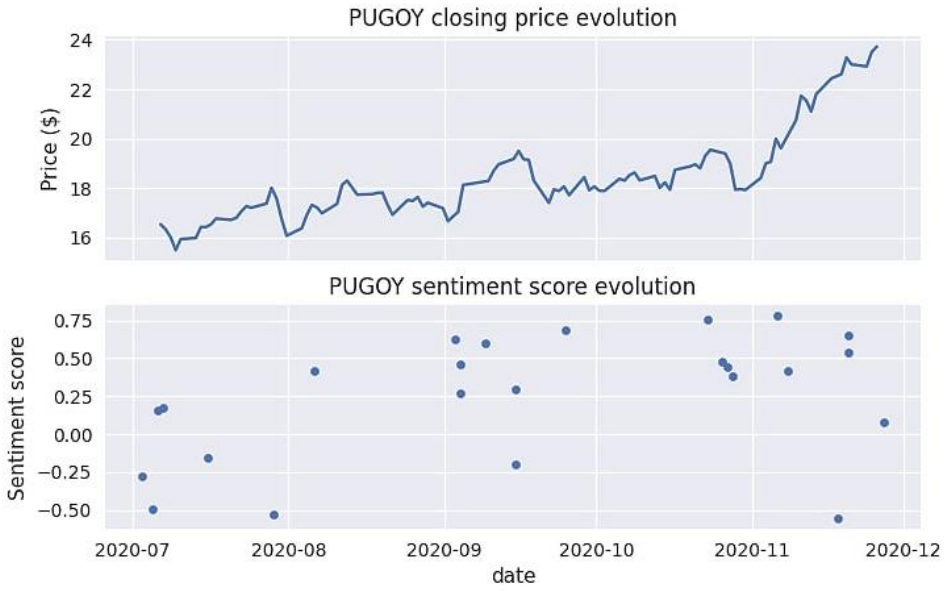

Figure 28: Comparison PUGOY closing price and sentiment score

In order to find out how much the closing price and the sentiment score correlate with each other and thus to determine the significance of these two variables, the correlation for Peugeot is shown in the following figure. As can be seen, there is only a very low correlation value of 0.36. This means that the stock value of Peugeot is not really influenced by the headlines of the news.

[19] Alzazah, F. S. & Cheng, X. (2020).

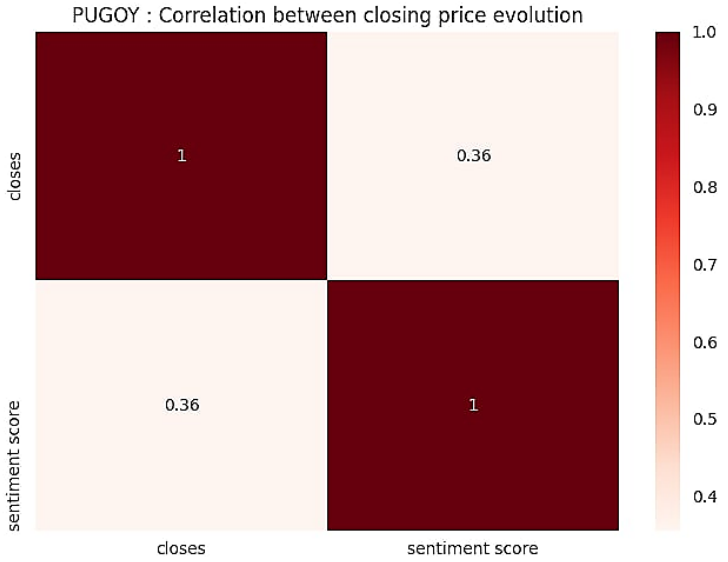

Figure 29: Correlation between closing price and sentiment score (PUGOY)

For Snapchat there were fewer articles in total than for Peugeot. Nevertheless, when comparing closing prices and sentiment score in the social media sector, it can be stated that news articles probably have a less strong impact on the share price development. For example, on November 9, 2020, there was a rapid decline in the share price, probably due to new global corona restrictions, although there was a very positive news contribution for Snapchat. Even so, at the beginning of November the share price development was not significantly influenced by worse news articles.

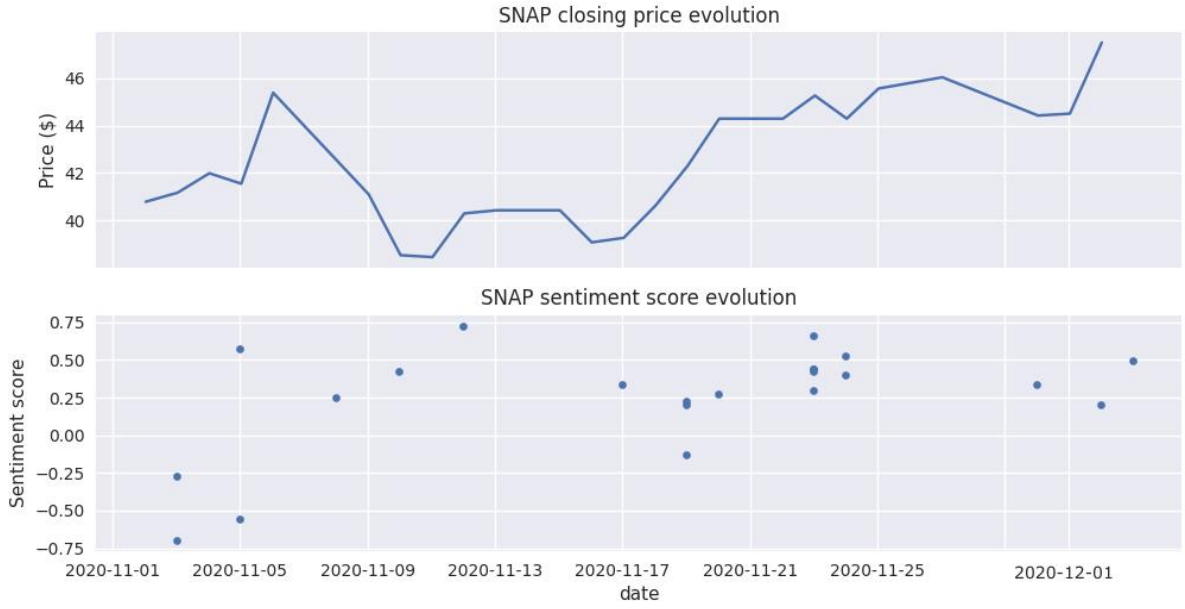

Figure 30: Comparison SNAP closing price and sentiment score

In the area of social media, it was also necessary to find out to what extent the closing price and the sentiment score correlate with each other in order to determine the significance. In contrast to Peugeot, the correlation for Snapchat is very high with a value of 0.7. Therefore, it can be said that the stock value of snapchat is influenced by the news.

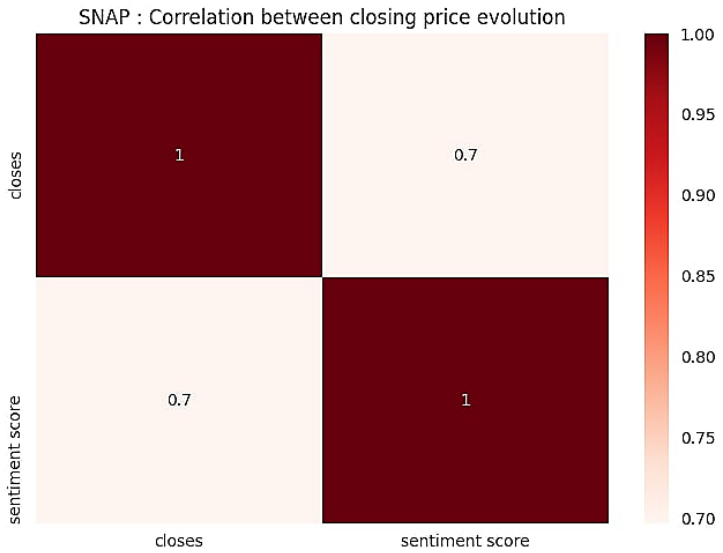

Figure 31: Correlation between closing price and sentiment score (SNAP)

In conclusion, for the social media industry there is a close connection between information from unstructured texts, such as social media and financial reports, and market movements on the stock exchange. The information thus helps to improve forecasting models. The more data is fed into the forecast model, the greater the forecast accuracy. In relation to this work, text mining based predictive models provide a practical solution to increase confidence in understanding market movements and when to make valuable investments. In the automotive industry, however, this statement cannot be confirmed due to the low correlation value.

7. Conclusion

Stock trading is one of the most common forms of investment for a wide range of parties, as it often offers higher chances of profit. Nevertheless, it remains a complex undertaking that requires the correct investment strategy, the right timing of the investment, and thus the ability to take advantage of such profit opportunities.

To be profitable when investing in the stock market, it is important to acknowledge the economic context of the stock considered. As seen in this study, every industry has its own characteristics and pace of evolution. Furthermore, investing in one stock or another may represent different levels of risk of losing money. Companies from the automotive industry seem to be rather stable compared to the ones from the social media industry. Depending on the kind of investment that is being considered and the trading strategy adopted, one case is more suitable than the other.

To make a maximum amount of profits, the timing of investment is a key factor. To know the right time to invest in a stock, several tools can be used in a more or less reliable way. The Bullish Engulfing Pattern, a phenomenon which occurrence indicates when to invest in a stock, has an efficiency that depends on the nature of the stock. In the case of a volatile one, the profits can potentially be higher than with a more stable one because it usually stagnates around the same value all year long, unless a big event occurs. Another technique to time investment is the Momentum Trading. Thanks to an analysis of a stock during the past couple of months, this principle defines rules that state when investing is advised. These two techniques work on short-term investments, ranging from five days to one month.

Text mining is also used in several prediction studies. In the case of companies from automotive industry and the social media companies, predictions are more or less conclusive. After a sentiment analysis on news articles on both domains, it appears the prediction are more likely usable for the social media companies than the automotive industry.

In conclusion, there are a lot of ways to help with decision-making. Some of them allow to give a predict insight on a near future, while others allow to capture the trend of the moment, the right time to take a decision. However, each technique or focus point has its relevancy and its efficiency depending on the nature of the stock, and its current state. No technique is totally efficient in all cases, and it is up to the analyst to use them in a way relevant and appropriate to their needs.

Bibliography

Code of our whole analysis, used to generate all the graphs of the report, publicly available at https://github.com/AppliNH/py-stockmarket-analysis.

Alzazah, F. S. & Cheng, X. (2020). Recent Advances in Stock Market Prediction Using Text Mining: A Survey. URL: https://www.intechopen.com/online-first/recent-advances-in-stock-market-prediction-using-text-mining-a-survey, access on 06.12.2020.

BMW AG (2020). About BMW.com. URL: https://www.bmw.com/en/footer/about-bmw-com.html, access on 04/11/2020.

Cambridge Dictionary (2020). API | Meaning in the Cambridge English Dictionary. URL: https://dictionary.cambridge.org/dictionary/english/api, access on 28/11/2020.

Facebook (2020). Our Culture. URL: https://about.fb.com/company-info/, access on 26/10/2020.

General Motors Company (2020). About GM. URL: https://www.gm.com/our-company/about-gm.html access on 04/11/2020.

Investopedia (2020) (a). Bullish Engulfing Pattern Definition. URL: https://www.investopedia.com/terms/b/bullishengulfingpattern.asp, access on 02/12/2020.

Investopedia (2020) (b). Investing, Stock Market Definition. URL: https://www.investopedia.com/terms/s/stockmarket.asp, access on 04/11/2020.

Investopedia (2020) (c). Return On Investment. URL: https://www.investopedia.com/terms/r/returnoninvestment.asp, access on 09/11/2020.

Investopedia (2020) (d). Shares. URL: https://www.investopedia.com/terms/s/shares.asp, access on 04/11/2020.

Investopedia (2020) (e). Trade Triggers. URL: https://www.investopedia.com/terms/t/trade-trigger.asp, access on 30/11/2020.

Matplotlib (2020). Python Plotting. URL: https://matplotlib.org, access on 28/11/2020.

Nasdaq (2020). Trading Hours for the Nasdaq Stock Markets. URL: https://www.nasdaq.com/stock-market-trading-hours-for-nasdaq, access on 04/11/2020.

Pandas (2020). Pandas | Python Data Analysis Library. URL: https://pandas.pydata.org, access on 28/11/2020.

Peugeot (2020). Peugeot in brief. URL: http://www.peugeot.com/en/brand/peugeot-in-brief/peugeot-in-brief#listen, access on 04/11/2020.

Pinterest (2020). About Pinterest. URL: https://about.pinterest.com/en, access on 04/11/2020.

Snap Inc. (2020). Snap Inc. URL: https://www.snap.com/en-US, access on 04/11/2020.

Twitter Inc. (2020). Our company. URL: https://about.twitter.com/en_us/company.html, access on 04/11/2020.

Volkswagen AG (2020). The Volkswagen Group. URL: https://www.volkswagenag.com/en/group/portrait-and-production-plants.html, access on 04/11/2020.