Contents

[Abstract 3](#_Toc90483334)

[Keywords 3](#_Toc90483335)

[I. Introduction 3](#_Toc90483336)

[II. Methods 4](#_Toc90483337)

[III. Results & Discussion 7](#_Toc90483338)

[IV. Problems and Shortcomings 11](#_Toc90483339)

[V. Conclusion 11](#_Toc90483340)

[References 12](#_Toc90483341)

Table of Figures

[Fig. 1.Code for importing the libraries and extracting the data 5](#_Toc90482012)

[Fig. 2 Code for constructing the dataframe 5](#_Toc90482013)

[Fig. 3 Code for creating the signals 6](#_Toc90482014)

[Fig. 4 Code for constructing the portfolio 7](#_Toc90482015)

[Fig. 5 Code for constructing the returns 7](#_Toc90482016)

[Fig. 6 Code for plotting the data 8](#_Toc90482017)

[Fig. 7 Line chart for the returns after searching "flowers" 8](#_Toc90482018)

[Fig. 8 Plot chart for the returns after searching "flowers" 9](#_Toc90482019)

[Fig. 9 Table depicting the results based on the "flowers" data 9](#_Toc90482020)

[Fig. 10 Line chart for the returns after searching "stock" 10](#_Toc90482021)

[Fig. 11 Plot chart for the returns after searching "stock" 11](#_Toc90482022)

[Fig. 12 Table depicting the results based on the "stock" data 11](#_Toc90482023)

[Fig. 13 Returns based on the "flowers" data 11](#_Toc90482024)

[Fig. 14 Returns based on the "stock" data 12](#_Toc90482025)

## Abstract

Search engines are contemporarily used to scrutinize user attention allocation, such insights being helpful in the interpretation of the functioning of the overall economic environment, but more specifically, the stock market. This paper examines the practicality of such engines, such as Google Trends, which can be used to construct an investing portfolio with higher returns than the benchmark, based on data reporting weekly concerns on some topics. With the use of specific programming languages, we can observe how a buy-and hold portfolio behaves, compared to one constructed based on search engines, in the 2016 - end of 2021 time frame.

## Keywords

stocks, Python, Google Trends, buy and hold, research, trading system, return, benchmark, investing

## I. Introduction

The stock market has become an ubiquitous subject in our lives, gaining an increasing popularity in the last decade and being considered a store of value in the ever-evolving economic environment. It dates back to the 18th century, when the market on Wall Street was opened after the Buttonwood Agreement signed by 24 brokers in that same place (Contributor 2020). Nonetheless, it has suffered several changes, both in its substance and in the trading style, correlated with the shift in trends and overall economic and societal environment behavior throughout the years.

The hypothesis that a regular buy-and-hold strategy over a stock portfolio is no longer satisfactory for an investor, considering all the refittings in the economic environment in the modern years, is what represented the catalyst for this paper. Thus, we find it important to correlate the investing behavior with the recent trends, having as a benchmark the returns of one of the most representative indexes of the stock market: Dow Jones.

Considering that the benchmark indicates a conservative return of the performance which is around 8%, the scope of our project is to create a trading system based on some key words inserted in the Google Trends engine which would overtake the initial benchmark, thus ‘beating the market’. This would sustain the idea that the economic and overall global trends must be accompanied by a specially-fitted trading model, which should be in-line with these societal movements. Of course, following a buy-and-hold strategy which offers returns of 8% annually is a satisfactory idea for any investor, due to the fact that his money outperforms the inflation, but **the value desired to bring** with this project is a better fit trading design with higher returns which are personalized with what the community undertakes.

A diversified portfolio is crucial to every investor, which needs to carefully build and modify it based on the moves of the market. This asset allocation must be fit with his personal and financial goals, and also with the level of risk he is willing to undertake on his wealth. Thus, in order to minimize the risks and maximize the profits, it is rather essential to stay connected to the market and the market trends, which is why we could raise the interest of every potential investor with this paper, which can observe one or more keywords on the Google Trends engine and try to better understand the economic environment and its reshaping in order to make his portfolio more profitable than the benchmark.

## II. Methods

The two main tools used in order to generate the results were Google Trends and Python, the former being employed for the extraction of data regarding the interest of internet users towards specific words which would be more or less relevant to the scope of the project: ‘flowers’ and ‘stock’. After the extraction, these results saved as Excel files, respectively CSV files, were introduced in the programming language Python, through the PyCharm interface.

Python served as an adjuvant for the processing of data and generation of results which would be further manipulated and compared in order to test the hypothesis.

Firstly, several libraries needed to be imported, as depicted by Fig. 1, in order to allow the import and display of data into different formats or plots. Next, the second part of the code takes the data from the CSV/ Excel files extracted from Google Trends and also downloads the data from the desired timeframe and splits it by weeks in order to easily observe the returns.



Fig. 1.Code for importing the libraries and extracting the data

Next, the stencil for the table of data is created by naming the Headers of the columns which will subsequently be displayed in this paper and constructing the data set.

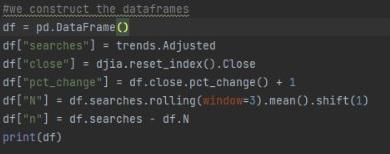


Fig. 2 Code for constructing the dataframe

The following step contains a for loop in order for the signals to be generated based on the previously imported data. Therefore, the value “1” was set for selling and “0” for buying. Also, the index was set to start after the first three weeks, in order to have some historical value before the first signal.

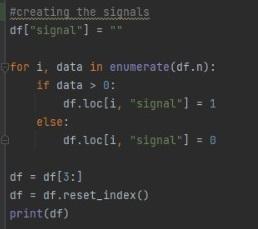


Fig. 3 Code for creating the signals

The next part of the code helps in processing the previously created signals and adjusting the buy and hold calculus in order to insert the results in the table which will be created.

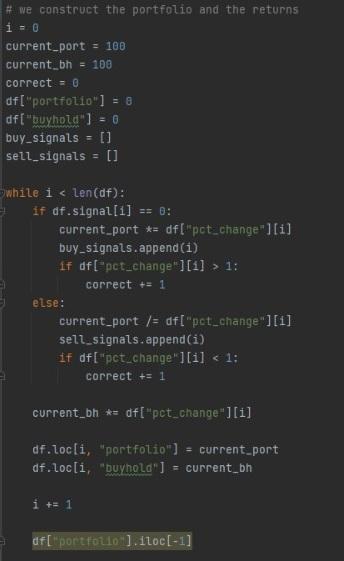


Fig. 4 Code for constructing the portfolio

Fig. 5. illustrates the subsequent step of computing and displaying the results for the “Annualised Buy-and-Hold Portfolio Return”, “Accuracy”, “Total Return” and “Annualised Google Portfolio Results”. These computations are standardised return formulas for investment portfolios.

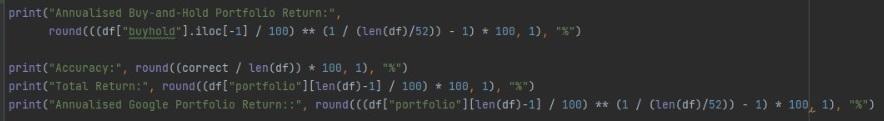


Fig. 5 Code for constructing the returns

The final step is necessary in order to obtain a visual representation of the results through plotting the outputs, thus creating two graphs, one being a line chart and one a scatter chart.

****

Fig. 6 Code for plotting the data

## III. Results & Discussion

Firstly, it would be worth mentioning that through trial and error, we noticed that it is very important to use relevant words, in order for the Google Trends portfolio to be properly assessed and to generate pertinent results.

As it can be observed in Fig. 7. the use of a random word, such as “flowers” does not offer satisfactory results, implying that the buy and hold strategy would be the best, depicting therefore parallel graph lines. On the vertical axis we have the portfolio return, wheres on the horizontal one we have the index values of the searches.

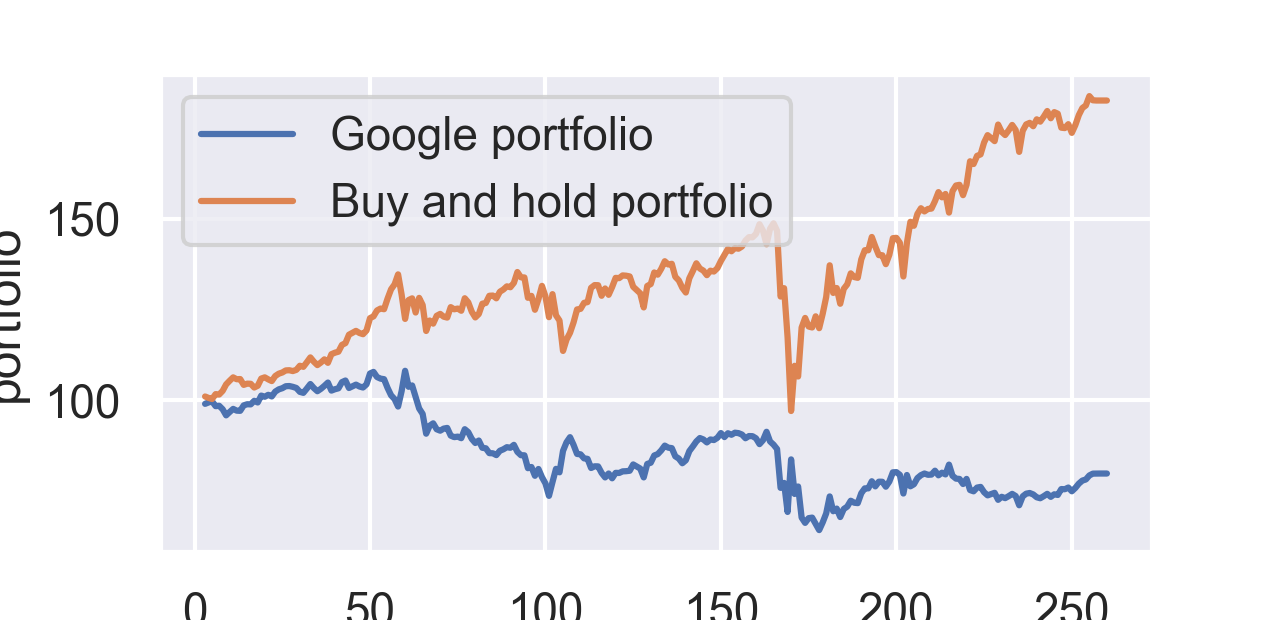
****

Fig. 7 Line chart for the returns after searching "flowers"

Moreover, a scatter chart regarding the decision of buying and selling was generated with the help of the imported data. On the vertical axes, we have the closing prices of the index at the end of the weeks, and on the horizontal axes, also the search index mentioned previously.

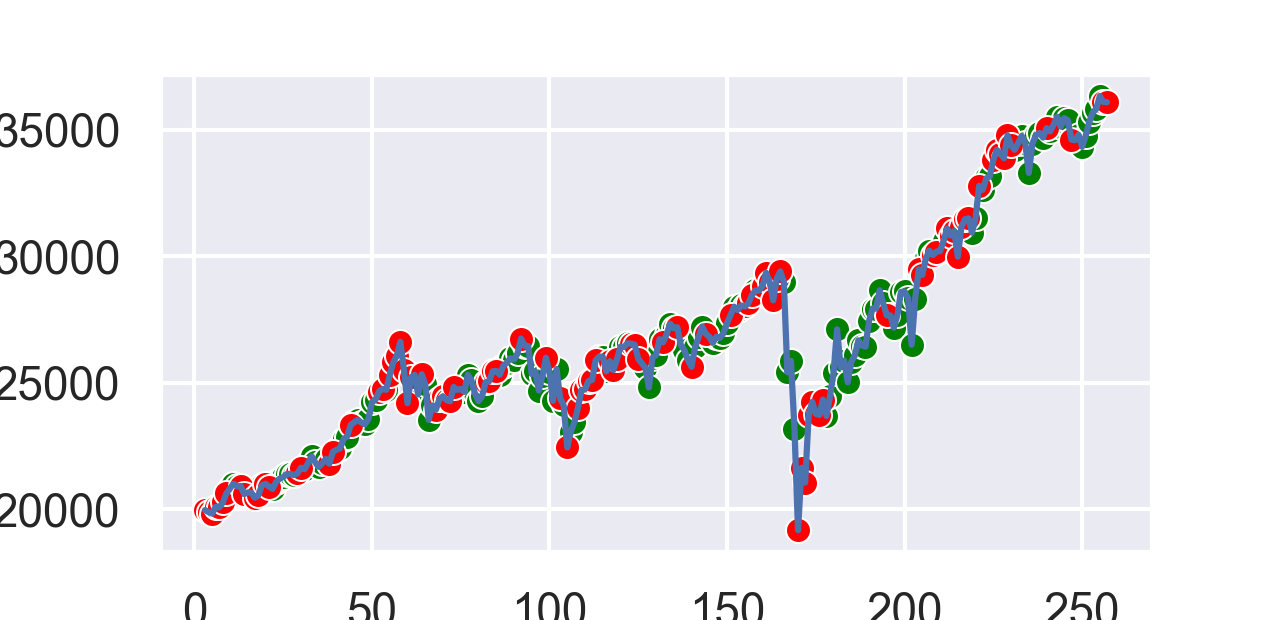
****

Fig. 8 Plot chart for the returns after searching "flowers"

The table below depicts the values generated for the ‘flowers’ search on Google trends and run into Python. As such, it can be seen that there are 258 entries of data, between 11-12-2016 and 5-12-2021. The index represents the number of the week, and its values start with 3 because we need at least 3 weeks for the first results to be generated. The ‘searches’ column illustrates the indexes used in order to show the interest upon the searched word, which range between 0-100, where 100 is the maximum level of interest of the people. The ‘close’ column indicates the closing price at the end of the week, while the following one shows the percentage change of the closing price from one week to another. The N column shows the mean of the prices from the last three weeks, while the n one represents the difference between the searches of that week and the current mean. Lastly, the signal column illustrates the decision which should be made, namely 1 for sell and 0 for buy.

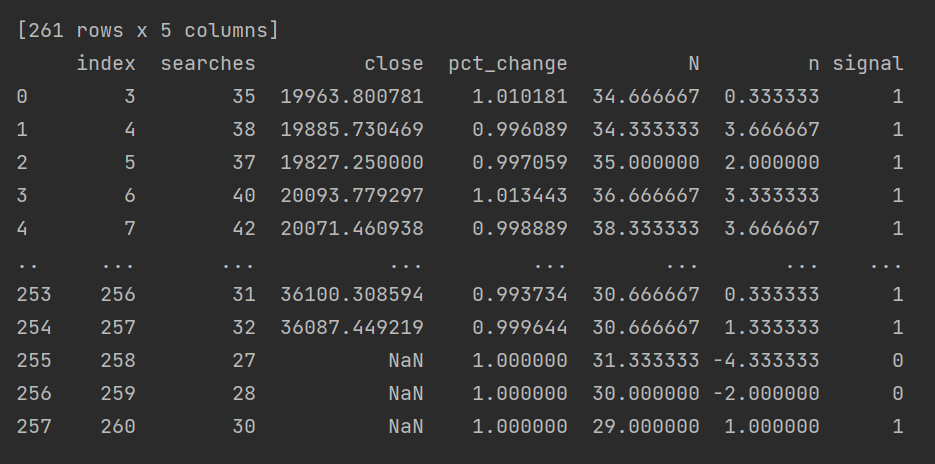
****

Fig. 9 Table depicting the results based on the "flowers" data

After the search of the random word ‘flowers’ and its irrelevant depictions, we went ahead with the search of one of the most important key words, namely ‘stock’, in order to truly discover the power of the search engine-based portfolio. Our underlying decision of choosing this word stood in the fact that whenever the interest of people in stocks grew, so did the profitability of this personalized portfolio.

From 2016 up until the end of 2019, the buy and hold strategy proved itself to be sufficient and profitable, but with the appearance of the Sars-Cov-2 virus, it is noticeable that the investing behavior has switched. Thus, the initial strategy is lacking, and the need of a personalized portfolio with the use of Google Trends obviously takes the spotlight with regards to the profitability over the benchmark of the market. Thus, the initial hypothesis that the economic and societal trends would imply a new investment behavior stands true. Of course, historical data would be rather important to keep in mind, but not sufficient anymore, so as to make decisions of buying and selling, having to replace it with a more relevant anchor: the index of searches for the word ‘stock’ in the search engine. Fig. 3.4. shows that the graph lines had a mirroring evolution throughout the period chosen, and thus it also depicts the switch previously mentioned, the takeaway being that it has modified the profitability of both strategies, turning the Google portfolio into the one which gives the best return.

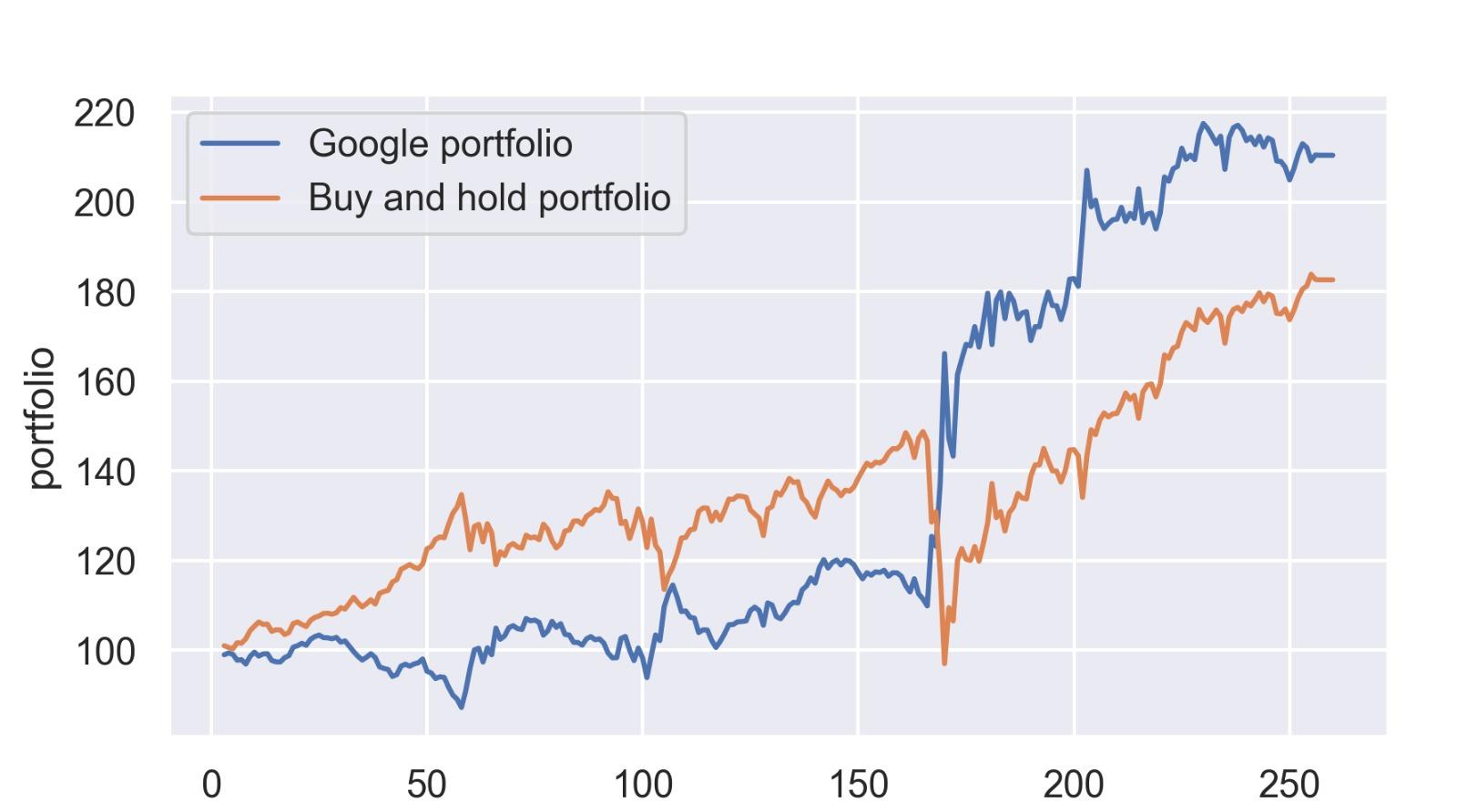
****

Fig. 10 Line chart for the returns after searching "stock"

Once again, the plot we created for this more relevant word depicts the actualized signals for buying and selling, in order to generate a higher return. There is an obvious improvement in the decision-making process, the signals being more appropriate.

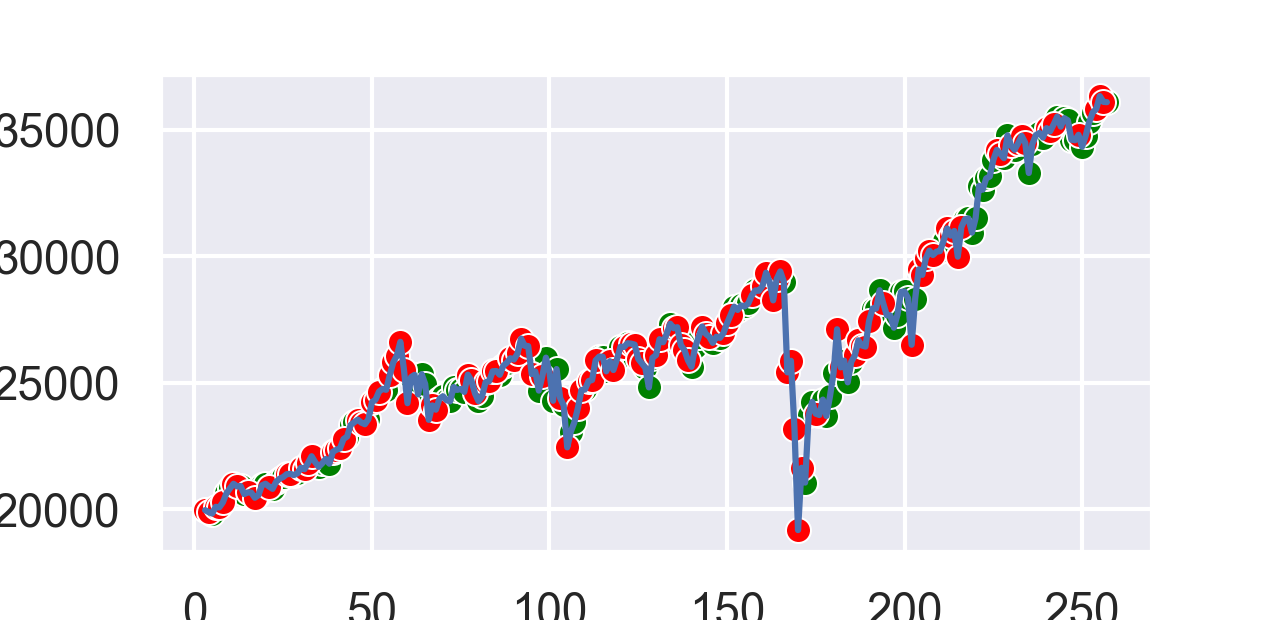
****

Fig. 11 Plot chart for the returns after searching "stock"

The table depicted by figure 3.6. has the same characteristics as Figure 3.3., the main difference arising from the index of the searches, which are significantly lower, but indicates an overall ascending trend in the interest over stocks and investing, from 15 in 2016 until 27 in the current year. This exact snapshot does not illustrate all the values, but as seen in our database, the searches for the word ‘stock’ grew significantly amid the rise of Covid-19, when most people started to look for ways of nurturing their wealth, reaching the maximum index of 100 January 2021.

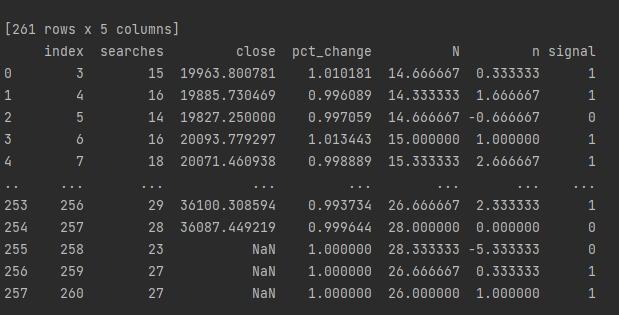
****

Fig. 12 Table depicting the results based on the "stock" data

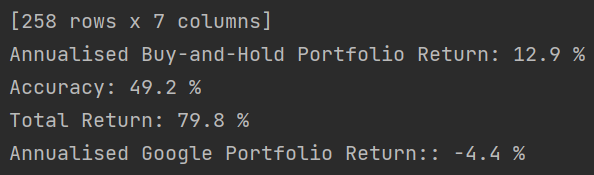
****

Fig. 13 Returns based on the "flowers" data

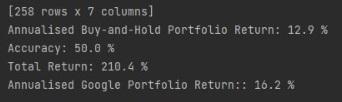
****

Fig. 14 Returns based on the "stock" data

Fig. 14 generates the reason for choosing a personalized portfolio based on Google Trends, which illustrates that the Annualised Google Portfolio Return exceeds the Buy-and-Hold one with almost 4%. The scope of any investor is to maximize their profits, thus one might consider this system to be more appropriate, as illustrated by the Total Return of 210.4%. The Accuracy, however, might be an issue, as 50% represents a high level of uncertainty.

## IV. Problems and Shortcomings

The main shortcoming encountered within this project would be the reliability of the system based on the searched word, mainly because one is unaware of the period during which that exact word will continue to be relevant, especially if it is taken into account that interest in stock investments may falter and return to the initial levels.

Another problem that should be highlighted is the fact that the system might be lacking in giving the appropriate signal due to the fact that it reacts to the prices rather than predicting them.

If the results depicted by Fig. 3.7. and Fig. 3.8. are considered, it is evident that an overall level of accuracy of around 50% is not necessarily desirable, as it does not offer enough certainty. Moreover, when a comparison is done, the levels of accuracy between the two words were not significantly different, as there was only a slight amelioration of 0.8% in the case of the relevant word, namely “stock”.

## V. Conclusion

In conclusion, as the market undergoes changes over time, catalysed by economic and societal trends and even moments of crisis, investors must continuously research new methods in order for their wealth to upscale and keep up with such shifts. Thus, a regular buy-and-hold portfolio is a good strategy to keep in mind in order to gain steady, but lower returns, whereas taking advantage of the search engines which become increasingly popular could result in the creation of new, more profitable strategies, without higher risks.

The initial hypothesis that the historical strategy could no longer be satisfactory and investors must develop a strategy in line with the global trends was confirmed, our results generated by using the programming language Python depicting higher returns in the case of employing Google Trends as a support for creating an investment strategy. Obviously, not every topic might be relevant for constructing a portfolio, as illustrated in the initial test for the term ‘flowers’, but rather one related with the investment activity.

Search engines are, therefore, powerful tools which allocate attention to users’ interests and can be used as adjuvants for the maximization of one’s investments based on their most relevant results. Combined with a dash of courage, every person or investor might seek exceeding the regular benchmark of important indexes (Dow Jones, S&P) by keeping track of the shifts and considering the signals which a Google portfolio might produce.

# References

Contributor, Chron. 2020. "How the Stock Market Was Started & by Whom." *CHRON.* 24 November. Accessed December 13, 2021. https://smallbusiness.chron.com/stock-market-started-whom-14745.html.

n.d. "Research model." *Colab Research Google.* Accessed December, 2021. https://colab.research.google.com/drive/1b3AzyEVZL2dEXWmq\_TjMRX2JDxwTdGAA.