**Capstone Project**

**Document Skeleton**

# Problem statement

* We are trying to take as much risk out of the financial markets as possible. Risk can never be taken out completely but the goal is to minimise it as much as possible.
* The goal is to take some of the stress and uncertainty out of the markets and help people plan better for their future.
* Too many people are losing their hard earned money and savings on investments because they either don’t know what to invest in.
* Hopefully our portfolio prediction will help people make money and prepare better for their future.
* This is not the first time predicting the stock market moves has been attempted. We are trying to make it more accessible and easier to understand for the average retail investor by taking out all the complicated formulas and industry jargon. We are going to show our stakeholders and clients what they really need to see and only that.

# Industry/ domain

* The industry we will be working in is the finance industry. This industry has been dominated by big banks and corporations for too long. We plan to help the individual make a living, or at least a side income, in the finance industry.
* The equity markets alone are worth billions of dollars. The biggest challenge facing individuals is where to start and how much to invest in what stock/stocks.
* At first glance it might seem straight forward to get started in the world of investing as it is such an established industry with what appear to be clear cut rules. This is not always the case, there are so many different options available to invest in that it can be very daunting for individuals. It’s not as easy as buy a share and sell it at a later day at hopefully a profit. Our goal is to take most of the hard work of investing out for the individual.
* The main challenges for anyone in the financial industry is what to invest in, how much to invest and how much risk are they prepared to take on.
* The project we have taken on here is purely designed to help investors with stock selection and what percentage of their capital to invest in each stock. This means this project is designed purely for the purpose of the financial markets.

# Stakeholders

* Our stakeholders are the CEO, COO and head of trading. This is done so that the portfolio ideas can be approved by senior management before it can be presented to the clients. Once the approval has been given then there will be a second set of stakeholders, this time be clients who this project is actually designed for.
* This will be an important project for senior management as it will help the clients decide what stocks/portfolios to invest in. Often clients sit on their cash and don’t invest because they don’t know what to invest in. When the clients invest it gives the company more opportunities to make money. For the clients this project is also important as it will guide them in the right direction to invest their money in a portfolio of stocks chosen by professionals. Just sitting on their cash and not investing it won’t help the clients secure a better future.
* The expectations from senior management and the clients is for this project to help make the clients make a profit. If they make a profit then they are more likely to stay with the company and invest which in turn means the company makes money too.

# Business question

* The most important question that needs to be answered to make this project viable is, can we accurately predict the future price of a share/portfolio.
* The value of answering this question is difficult to calculate. Billions of dollars’ worth of shares are traded every day. This means the only limit on what a client invests depends on the amount they are willing to risk. Please note, we will not allow our clients to risk more than they can afford to lose.
* In order for clients to trust us enough to invest with us again we want to make sure our forecast is very reliable. If their returns are below what we predict then it is likely they will stop investing with us. There is some room for error on the upside though. If we predict a certain outcome and they make greater returns than expected there is a very good chance they will invest with us again.

# Data question

* Stock data is easily available to anyone. There are many different information providers worldwide. The data is normally clean and not much EDA is required.

# Data

* We got our data from Yahoo Finance and Quandl, both being free for what we needed to do.
* We only used 10 years’ worth of data but it is possible to go back a lot more. We only used monthly data which is why we only had 123 observations and 6 features. It is also possible to use daily data for that same period and more.
* The data is very reliable as it is the end of day data which is time stamped and confirmed by the exchange the stock is traded on.
* The quality of the raw data is very high.
* As I just mentioned, the data is generated by the exchange the stock is traded on. Each day the closing price of each stock is posted by the exchange.
* The data is available on an ongoing basis. At the moment there are many sites which offer data for free and there are also some data providers who offer more in-depth data for a price.

# Data science process

## Data analysis

* As mentioned earlier, there is not much EDA required. Seeing as we used ARIMA and SARIMA, we had to make sure the data was in a time series format. We also had to make sure the data was stationary.
* The pipeline is reusable for future projects. We only used four stocks for our portfolio so if we wanted to create another portfolio with four stocks we would simply have to import the data for the new stocks. If we wanted to add more than four stocks we would use the same procedure as we did in the previous four stocks.

## Modelling

* We tested the ARIMA model first but noticed the SARIMA worked better in predicting the forecast.
* The model used is the ARIMA and SARIMA models.
* Training did not take long at all, close to instantaneous.
* We ran the entire model on our personal computer.
* What are the model performance metrics?
* The SARIMA model gave us the best predictions for future forecasting which is why we used SARIMA.

## Outcomes

* The main outcome for our project was that we were able to forecast the portfolio performance two years in advance. It is a difficult talk to try and predict the overall value of a portfolio 2 years in advance but we are happy to say that our conservative forecast was only 6.2% off that actual market for that portfolio of shares. The actual market performed slightly better.

## Implementation

* The only considerations for implementing the model are based on what risk a client is prepared to take on and what amount of capital they are able to put at risk. Our forecast was accurate but there is no such thing in the financial market as a sure thing and there will always be some risk attached with investing.

# Data answer

* We feel the question was answered in that the data helped us get the results we were after. The data was sufficient in answering our business question.
* The confidence level in the data is very high as we know where is comes from and that the original source of the data is regulated.

# Business answer

* The business question was answered satisfactorily as we were able to predict the future value of a stock portfolio with a high level of accuracy.
* We feel our stakeholder, both the senior management as well as our clients, will have a high level of confidence in the business answer.

# Response to stakeholders

* We recommend that the senior management stakeholders approve this portfolio recommendation as the forecast is accurate. We recommend to our clients that if they are in a financial situation where they are able to invest, that they invest in this portfolio as the forecast shows that we can predict the future value with a certain level of accuracy.

# References

* The data code used for this project can be found in Github, using the following link:

<https://github.com/BerndBE/IOD_DS/tree/main/CAP>STONE

The notebooks used are:

ARIMA-Monthly-MSFT.ipynb

ARIMA-Monthly-CSCO.ipynb

ARIMA-Monthly-IBM.ipynb

ARIMA-Monthly-AMZN.ipynb

Portfolio\_Allocation.ipynb

The presentation is:

Share Portfolio Prediction.pptx

Capstone Project Document:

Capstone Project Document.docx