



FACULTY OF SCIENCE & TECHNOLOGY  
Department of Computing & Informatics  
Forensic Computing & Security

---

Tools & Technologies of Data Science

3rd March 2020

Paul-David Jarvis  
[s5115232@bournemouth.ac.uk](mailto:s5115232@bournemouth.ac.uk)

---

# One Page Proposal - Predicting & Analysing Microsoft Corp (MSFT) Stocks

# 1.0 MSFT Stock

## 1.1 MSFT & Why It's a Big Data Problem

The NYSE produces roughly 2TB of data every day. MSFT is one of those stocks that it produces. The data that is MSFT grows in volume every day at the end of closing. This will cause challenges later on when it comes to storing the data at the current volume and velocity that it's growing. The data variety is structured which makes it easier to analyse. The data's veracity is trustworthy and sourced from Kaggle which is sourced from the New York Stock Exchange and the NASDAQ stock market. The data's value differs from the way it is used as it's historic data from several years ago but can be used to roughly predict the future stock.

## 1.2 Objectives

- Download the historic data and import it into Python.
- Analyse and visualise the data to discover trends using Python libraries like Scikit-Learn to produce Machine Learning Algorithms that will make a prediction for a given day price and to use Matplotlib to produce charts
- Discuss future work to discover more efficient possibilities such as Hadoop, AWS and Azure