Presentation of Project & Data Management Plan

Title: Toyota Stocks Market Prediction

Research Questions:

- How might Toyota's stock prices evolve according to past market data?
- What forecasting models are typically used to predict Toyota's stock price movements?
- How effective are ARIMA and LSTM models in predicting Toyota's stock performance?
- Which model is more suitable for short-term forecasts versus long-term projections for Toyota's stock?

Aims and Objectives:

- **Aim:** Forecasting future Toyota stock prices using time-series prediction models.
- **Objectives:** Objectives are as follows:
    - Gather and preprocess Toyota stock price data from a trusted financial source.
    - Implement ARIMA and LSTM models to perform time-series forecasting.
    - Compare the models' performance using key evaluation metrics.
    - Identify strengths and limitations of both models for stock price predictions.

Background and Summary:

Toyota is a global automotive leader, and its stock is of considerable interest to investors. Accurate predictions of stock price trends can offer valuable insights for financial decision-making. Previous studies have applied both traditional statistical methods and modern machine learning techniques for stock price forecasting.

ARIMA is a widely recognized technique for time-series forecasting, known for its ability to model linear patterns and seasonal fluctuations (Hyndman & Athanasopoulos, 2018). LSTM, introduced by Facebook, is adept at managing complex seasonal behaviors and is robust in dealing with missing values and irregularities in the data (Goswami, 2022). This study will utilize both models to forecast Toyota stock prices, focusing on historical data to evaluate their predictive performance for short-term and long-term trends.

List of References:

- Tsay, R.S. (2013) *An introduction to time series analysis*. Wiley Series in Probability and Statistics.
- Kourentzes, N. and Petropoulos, F. (2019) 'Forecasting with exponential smoothing: the state space approach', *European Journal of Operational Research*, 274(1), pp. 130-142.
- Box, G.E.P., Jenkins, G.M. and Reinsel, G.C. (2016) *Time series analysis: forecasting and control*. 5th edn. Wiley.

Project Task List and Timeline

Work Plan

Project Planning and Initial Setup:
- Establish the research objectives and formulate relevant questions.
- Gather Toyota stock price data from Yahoo Finance.
- Install essential software and libraries (e.g., Python, pandas, ARIMA).

Literature Review:
- Perform an in-depth examination of time-series forecasting techniques used for predicting stock prices.
- Summarize key insights from studies using ARIMA and LSTM models.

Data Preparation:
- Clean and organize the Toyota stock price dataset for analytical purposes.
- Handle any missing values, rectify outliers, and normalize the dataset as necessary.
- Perform Data Analysis to uncover trends and patterns in the dataset.

Model Implementation and Assessment:
- Apply LSTM, ARIMA and alternative forecasting models to project Toyota stock prices.
- Explain the rationale behind the chosen model parameters.
- Evaluate the performance of the models using metrics such as Root Mean Squared Error and Mean Absolute Error.
- Analyze and contrast the short-term and long-term efficacy of both models.

Results Evaluation:
- Examine the outcomes generated by both forecasting models.
- Discuss the advantages and disadvantages of each model in predicting stock prices.

Final Report Compilation and Submission:
- Assemble the findings, including sections on the introduction, literature review, methodology, and conclusions.
- Verify that all references adhere to academic standards (Harvard referencing).
- Submit the report and upload all relevant code to GitHub, ensuring professional documentation is included.

Timeline Chart:

Task List	Duration	Start Date	End Date
Planning of Project & Preliminary Setup	**1 week**	15 October 24	21 October 24
Literature Review	**1 week**	22 October 24	28 October 24
Data Preparation	**1 week**	29 October 24	04 November 24
Model Implementation	**2 weeks**	05 November 24	18 November 24
Result Analysis	**1 week**	19 November 24	25 November 24
Result Interpretation	**2 weeks**	26 November 24	16 December 24
Final Report Writing & Submission	**1 week**	17 December 24	23 December 24

Data Management Plan

Overview and Summary of Dataset

Dataset: Toyota Stock Prices
Source of Dataset: Yahoo Stocks Finance (Click [here](#))
Content: Daily historical stock prices for Toyota Motor Corporation, encompassing the date, opening price, highest and lowest prices, closing price, adjusted closing price, and trading volume.
Size: Several thousand records, with an approximate file size of approximately 1000 KB.
Access: The data is publicly available and sourced from Yahoo Finance, a trusted platform for historical financial data.

Document Management, Security, and Storage Solutions

File Organization and Version Management:
- **GitHub Repository:** Files, including datasets, model scripts, and documentation, will be hosted on GitHub for version control.
- **File Naming Conventions**: A consistent naming format will be applied, such as:
    - `toyota_stock_data.csv` for the dataset.
    - `data_cleaning_script.py` for data preprocessing.
    - `arima_model_forecasting.py` for the ARIMA forecasting model.
    - `lstm_model_forecasting.py` for the LSTM forecasting model.
    - `draft_report_v1.docx` for the initial report.
- **Backup Procedures:** Data and project files will be routinely saved to Google Drive and an external hard drive for backup.

Metadata
The dataset will be in CSV format with about 3500 records, totaling approximately 1000 KB. Metadata provides context about the data's structure.
- **Document Control:** The project will be hosted at [here](#), with bi-weekly code commits. A consistent naming convention will be used for files, ensuring easy access for markers.
- **ReadMe File:** A ReadMe file will summarize the project's objectives, input/output formats, and usage instructions to aid other coders.
- **Security and Storage:** Files will be backed up weekly and shared securely via GitHub and GoogleDrive. Data will be stored in both locations for version control and backup.

Ethical Requirements
- **GDPR Compliance:** The dataset contains only publicly available financial data, thus it does not fall under GDPR regulations.
- **Conformity to UH Ethical Policies:** The project aligns with University of Hertfordshire's ethical guidelines, as it involves no human participants or sensitive data.
- **Permission to Use Data:** The data is publicly accessible and adheres to Yahoo Finance's terms for non-commercial research.
- **Ethical Data Collection:** The data is sourced from Yahoo Finance, a reputable platform that ensures ethical collection practices.