Project Charter

Thales Stock Predictor

****

**DBAS3090 – Applied Data Analytics**

**James Laurence – w0211593**

**January 12th, 2024**

Table of Contents

[Section 1. Charter Introduction 3](#_Toc157032199)

[1.1 Document Change Control 3](#_Toc157032200)

[1.2 Executive Summary 4](#_Toc157032201)

[1.3 Authorization 5](#_Toc157032202)

[Section 2. Project Overview 6](#_Toc157032203)

[2.1 Project Summary 6](#_Toc157032204)

[2.1.1 Project Goals, Business Outcomes and Objectives 6](#_Toc157032205)

[2.1.2 Project Scope 6](#_Toc157032206)

[2.1.3 Scope Definition 6](#_Toc157032207)

[2.1.4 Boundaries 7](#_Toc157032208)

[2.2 Milestones 7](#_Toc157032209)

[2.3 Deliverables 8](#_Toc157032210)

[2.4 Project Risks, Assumptions, and Constraints 9](#_Toc157032211)

[2.4.1 Risks 9](#_Toc157032212)

[2.4.2 Assumptions 10](#_Toc157032213)

[2.4.3 Constraints 10](#_Toc157032214)

[Section 3. Project References 11](#_Toc157032215)

[Section 4. Document References 11](#_Toc157032216)

[Section 5. Appendix 12](#_Toc157032217)

[5.1 Project Schedule 12](#_Toc157032218)

[5.2 Descriptive Statistics 13](#_Toc157032219)

[5.3 Key Performance Indicators (KPIs) 13](#_Toc157032220)

# Charter Introduction

## Document Change Control

|  |  |  |  |
| --- | --- | --- | --- |
| Revision Number | Date of Issue | Author(s) | Brief Description of Change |
| 1.0 | January 12th, 2024 | James Laurence | Document Creation |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |

## Executive Summary

**Project Objective**

The Thales Stock Price Prediction project aims to provide valuable insights to both current Thales stockholders and prospective investors. This project will forecast Thales' stock prices by leveraging predictive modelling techniques, enabling informed investment decisions.

**Project Deliverables**

1. **Project Documentation**:
   * **Project Charter Document**:
     + The documentation contains project goals, objectives, and business outcomes.
   * **Project Risks**
     + Project risks and risk mitigations are outlined.
   * **Project Schedule**:
     + The project deliverables and specific milestones have their timeframe defined.
2. **Database Development**:
   * **Data Flow Diagram**:
     + Visualizes data movement within the system.
   * **Data Dictionary**:
     + Compiles data elements, definitions, and relationships.
   * **Entity Relationship Diagram**:
     + Models database entities and associations.
   * **SQL Script Creation:**
     + Execution of SQL statements for project database.
3. **Data Processing and Modeling**:
   * **ETL Process**:
     + Extract, transform and load the data using SQL Server Integration Services (SSIS).
   * **Build Predictive Model**:
     + Implementation and testing of machine learning algorithms for stock price predictions.
4. **Data Automation and Reporting**:
   * **SSIS Modelling Integration and Package Scheduling**:
     + Integrate predictive model and set schedule for SSIS automation.
   * **Final Report:**
     + **Provide stakeholders with analysis of dataset features, dataset KPIs, predictive analytics and visualizations using dynamic data modelling.**

## Authorization

This project charter formally authorizes the existence of the Thales Stock Predictor project. It gives the project manager the authority to apply organizational resources to the activities described herein. If there is a change in the project scope, the project charter will be updated and submitted for re-approval.

A close-up of a signature

Description automatically generated

January 24th, 2024

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

George Campanis Date

Project Sponsor

Faculty Advisor/Instructor, Nova Scotia Community College

Diagram

Description automatically generated with medium confidence

January 24th, 2024

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

James Laurence Date

Project Manager

Student, Nova Scotia Community College

# Project Overview

## Project Summary

The primary objective of this project is to guide current Thales Stockholders in determining whether they should hold, sell, or buy their Thales stock holdings. The secondary aim is to inform prospective investors whether Thales’s stock is a lucrative investment opportunity. The project deliverables include Project Documentation, Database Development, Data Extraction, Transformation and Loading, Predictive Modelling, Data Automation, Data Visualization and Report Generation.

### Project Goals, Business Outcomes and Objectives

The following table outlines the Project Goals, Measurable Objectives, and Business Outcomes.

|  | Goals | Objectives | Business Outcomes |
| --- | --- | --- | --- |
| 1 | Project Deliverables | * Project Charter * Entity Relationship Diagram * Data Dictionary * Data Flow Diagram | * Project Documentation |
| 2 | Backend Development | * DDL Script * Predictive Modelling | * Database Development * Predictive Modelling |
| 3 | Task Automation | * SSIS – Extract Data * SSIS – Transform Data * SSIS – Load Data * SSIS – Modelling Integration * SSIS – Package Scheduling | * Complete Project Automation |
| 4 | Report and Visualization | * Descriptive Statistics * Predictive Analytics * KPI’s | * Reports (multi-dimensional) * Data Visualization * Data Models |

### Project Scope

This project's scope is to utilize automation for hands-off predictive modelling for Thales’s stock prices.

### Scope Definition

Overall project deliverables: The objectives include creating a Project Charter, Entity Relationship Diagram, Data Dictionary, and Data Flow Diagram. The business outcome is comprehensive Project Documentation.

Backend Development: The objectives are to create a DDL Script and perform Predictive Modelling. The business outcomes are Database Development and Predictive Modelling.

Task Automation: The objectives include data extraction, transformation, and loading (ETL) using SSIS, modelling integration, and package scheduling. The business outcome is complete Data Automation.

Report and Visual Generation: The objective is to perform Predictive Analytics. Business outcomes include generating multi-dimensional reports, Data Visualization, Predictive Modelling, and KPIs.

### Boundaries

The following table contains the activities in and out of the project scope.

|  |  |  |  |
| --- | --- | --- | --- |
|  | Activities in Scope |  | Activities Out of Scope |
| 1 | Documentation | 1 | Front-end Development |
| 2 | Multiple Dataset Extractions | 2 | Ongoing Maintenance |
| 3 | Database Deployment | 3 | Data Security |
| 4 | Data Modelling | 4 | Data Governance |
| 5 | Data Analytics | 5 | Privacy Compliance |
| 6 | Report and Visualization |  |  |

For the activities Out of project scope:

Front-end Development:

* Ideally, the interactive dashboard should be first; then the dashboard should be included in the front-end web.

Ongoing Maintenance:

* Project maintenance would continue once completed, depending on available resources.

Data Security, Data Governance and Privacy Compliance:

* These features would have standard inclusion during backend development; however, they must be improved when pursuing front-end development.

## Milestones

The following table contains the project milestones, description and expected completion date.

|  | Project Milestone | Description | Expected Date |
| --- | --- | --- | --- |
| 1 | Project Charter Document | Completion of the project charter outlines project objectives, scope, stakeholders, and high-level requirements. | 2024-01-28 |
| 2 | Project Schedule | Finalization of the project timeline, including start and end dates for critical activities. | 2024-01-28 |
| 3 | Data Flow Diagram | Creation of a visual representation illustrating how data moves through the system. | 2024-02-11 |
| 4 | Data Dictionary | Compilation of a comprehensive list of data elements, their definitions, and relationships. | 2024-02-11 |
| 5 | Entity Relationship Diagram | Development of a graphical model depicting database entities and their associations. | 2024-02-11 |
| 6 | DDL Database Creation | Successful execution of SQL statements to create the project database. | 2024-02-18 |
| 7 | Build Predictive Model | Implementation of machine learning algorithms to generate predictive insights. | 2024-02-25 |

|  | Project Milestone | Description | Expected Date |
| --- | --- | --- | --- |
| 8 | SSIS Data Extraction | Extraction of data from Yahoo Finance for Thales Stock, France Market Index, Euro Market Index, and Standard & Poor Index using SQL Server Integration Services (SSIS) as CSV. | 2024-03-03 |
| 9 | SSIS Data Transformation | Transformation of extracted data into a usable format. | 2024-03-03 |
| 10 | SSIS Data Loading | Loading transformed data into the Thales Stock Prediction database. | 2024-03-03 |
| 11 | SSIS Modelling Integration | Automating the predictive model using dynamic data. | 2024-03-10 |
| 12 | SSIS Package Scheduling | Scheduling the SSIS package task to perform automatically | 2024-03-24 |
| 13 | Descriptive Statistics | Summarization and organization of the characteristics of the dataset | 2024-03-31 |
| 14 | Predictive Analytics/Modelling | Application of statistical techniques to predict future outcomes. | 2024-03-31 |
| 15 | KPI’s | Identification and measurement of critical project success metrics. | 2024-04-07 |
| 16 | Descriptive Statistics | Analysis of data to summarize and describe its main features. | 2024-04-07 |
| 17 | Visualization and Model | Creation of visual reports and presentation of predictive models. | 2024-04-07 |
| 18 | Presentation | Presentation of implemented Project to shareholders. | 2024-04-14 |

## Deliverables

The following table outlines the Key Project Deliverables with expected requirements.

|  |  |
| --- | --- |
| **Project Deliverable 1: Project Documentation** | |
| **Stakeholder:** | The project manager, Project sponsor, and key team members are required for project oversight. |
| **Description:** | Created proper project documentation (project charter, data dictionary, entity relationship diagram, data flow diagrams). |
| **Acceptance Criteria:** | Project Sponsor approval. |
| **Due Date:** | As per the project’s current timeline (See [Appendix 5.1 Project Schedule](#_Project_Schedule)). |
| **Project Deliverable 2: Database Development** | |
| **Stakeholder:** | Database administrators, data engineers, data architects, and data analysts. |
| **Description:** | Built, tested, and deployed the Thales Stock Predictor database (data security, data integrity). |
| **Acceptance Criteria:** | Successful database deployment. |
| **Due Date:** | As per the project’s current timeline (See [Appendix 5.1 Project Schedule](#_Project_Schedule)). |
| **Project Deliverable 3: Predictive Modelling** | |
| **Stakeholder:** | Data scientists, Data analysts, Machine learning experts, and business analysts. |
| **Description:** | We are building and fine-tuning machine-learning models for stock prediction. |
| **Acceptance Criteria:** | Model accuracy and successful predictions. |
| **Due Date:** | Continuous throughout the project timeline (See [Appendix 5.1 Project Schedule](#_Project_Schedule)). |
| **Project Deliverable 4: Complete Project Automation** | |
| **Stakeholder:** | System administrators, IT automation specialists, Database administrators, Data analysts |
| **Description:** | **Automation of data extraction, transformation, loading and prediction model integration.** |
| **Acceptance Criteria:** | Consistent and effective automated workflows. |
| **Due Date:** | As per the project’s current timeline (See [Appendix 5.1 Project Schedule](#_Project_Schedule)). |
| **Project Deliverable 5: Reports and Visualizations** | |
| **Stakeholder:** | Project sponsor, potential investors, current investors. |
| **Description:** | Presenting Descriptive Statistics, Project KPIs, and Visualization in an Interactive Dashboard and Report format |
| **Acceptance Criteria:** | Achievement of KPI targets |
| **Due Date:** | As per the project’s current timeline (See [Appendix 5.1 Project Schedule](#_Project_Schedule)). |

## Project Risks, Assumptions, and Constraints

### Risks

This risk assessment differs from a full risk assessment, which would have been completed during project proposal drafting. The following table contains identified risks and possible mitigation options.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **Risk Description** | **Probability**  **(H/M/L)** | **Impact**  **(H/M/L)** | **Risk Management Plan** | **Office of**  **Primary Interest** |
| 1 | – Market Volatility –  The market may have fluctuations due to economic, political, or environmental events. | High | High | For investors, having a diverse portfolio lowers the risk of investing in Thales Stock. | Project Management |
| 2 | – Data Quality –  Inaccurate data can cause flawed predictions. | Medium | Medium | Implement data validation, cleansing and quality assurance policies and audits continuously. | Project Management, Data Management |
| 3 | – Predictive Model Issues –  The model may perform well on training data but poorly on the test data. This could lead to inaccurate results. | Low | High | By including cross-validation by having a sub-set of the training data to test the model. Also, increase the features by feature engineering. | Data Management, Data Science |
| 4 | – Software Issues –  There could be coding errors, design flaws, integration issues and security concerns. | Medium | High | Have code reviews and implement secure industry standard coding practices. Implement continuous testing to identify any integration issues. | Project Management,  Software Development, Data Management |
| 5 | – Hardware Issues –  Hardware failures, power outages, and damage are significant concerns. | Medium | High | Have hot-swap hardware on server racks or use a cloud service that provides continuous hardware support. If not cloud-based, have a proper UPS to provide the needed power during any electrical issues. Control access to any hardware to prevent physical damage due to human error. | Project Management, IT Department,  Cloud Service Provider |

### Assumptions

The following table lists the items that cannot be proven or demonstrated when this project charter was prepared:

|  | Assumptions |
| --- | --- |
| 1 | Data Availability and Quality – Data is accurate, complete, frequently updated, and properly cleaned. |
| 2 | IT Infrastructure – Hardware and software required for project implementation are available and functional. |
| 3 | Predictive Modeling – Based on the data quality and algorithm development, the predictive model for Thales’s stock prices will generate accurate predictions. |
| 4 | Project Timeline – The timeline provides reasonable time intervals for each deliverable and expected milestones. |
| 5 | Security and Compliance – The project's security, compliance and privacy all adhere to laws and regulations as required. |
| 6 | Project Team – Project Department teams have the required quantity of members and are subject matter experts in their field and department. |
| 7 | Stakeholders – Current and potential investors find the stock predictions beneficial for picturing the future volatility of Thales stock. |

### Constraints

The constraints for the project are listed in the table below:

|  | Category | Constraints |
| --- | --- | --- |
| 1 | Time | Project documentation requires Approval by the milestone due date for work to continue. |
| 2 | Time | Project milestones and deliverables have set timelines which enable the project to reach completion by the project’s deadline. |
| 3 | Time | Team members are expected to reach milestones within expected due dates. |
| 4 | Technical | Database development requires project documentation approval. |
| 5 | Technical | Database testing and deployment require accurate and proper SQL code. |
| 6 | Technical | Predictive modelling requires the testing of multiple models to find the algorithm that provides the highest accuracy. |
| 7 | Data | Data extraction, transformation and loading require quality and frequently updated data. Data transformation requires secure coding techniques. |
| 8 | Data | The predictive model requires access to accurate, clean, and frequently updated data. |
| 9 | Resources | Having the right amount of team members with the skill sets required to perform project requirements |
| 10 | Budget | Meeting the milestones and deliverables will continue funding this project. |

# Project References

More information concerning this project's data sources can be found in the following resources:

| Document Title | Version # | Date | Author and Organization | Location (link or path) |
| --- | --- | --- | --- | --- |
| Thales Stock | 2024-1 | January 22, 2024 | Yahoo Finance | [Thales Stock (HO.PA)](https://ca.finance.yahoo.com/quote/HO.PA/?p=HO.PA) |
| S&P500 Index | 2024-1 | January 22, 2024 | Yahoo Finance | [S&P500 Index (SDI=F)](https://ca.finance.yahoo.com/quote/SDI%3DF/?p=SDI%3DF) |
| France Market Index | 2024-1 | January 22, 2024 | Yahoo Finance | [France Market Index (FCHI)](https://ca.finance.yahoo.com/quote/%5EFCHI/?p=%5EFCHI) |
| Euro Market Index | 2024-1 | January 22, 2024 | Yahoo Finance | [Euro Stoxx Market Index (GDAXI)](https://ca.finance.yahoo.com/quote/%5EGDAXI/?p=%5EGDAXI) |

# Document References

Brownlee, J. (2019, April 12). Overfitting and Underfitting with Machine Learning Algorithms. Retrieved from machinelearningmastery.com: https://machinelearningmastery.com/overfitting-and-underfitting-with-machine-learning-algorithms/

Elite Data Science. (2022, July 6). Overfitting in Machine Learning: What It Is and How to Prevent It. Retrieved from elitedatascience.com: https://elitedatascience.com/overfitting-in-machine-learning

EM360 Tech. (2022, July 6). Understanding Computer Hardware Risk Mitigation Methods. Retrieved from em360tech.com: https://em360tech.com/tech-article/understanding-computer-hardware-risk-mitigation-methods

Epoch Converter. (2024). Epoch & Unix Timestamp Conversion Tools. Retrieved January 12, 2024, from epochconverter.com: https://www.epochconverter.com/

Farrelly, J. (2023, March 28). The Ultimate Guide to IT Department Structure & Responsibilities. Retrieved from electric.ai: https://www.electric.ai/blog/guide-to-it-department

Institute of Data. (2023, November 21). Understanding Risk Exposure in Software Engineering. Retrieved from institutedata.com: https://www.institutedata.com/us/blog/risk-exposure-in-software-engineering/

Kassa, S. G. (2017, May 1). IT Asset Valuation, Risk Assessment and Control Implementation Model. Retrieved from isaca.org: https://www.isaca.org/resources/isaca-journal/issues/2017/volume-3/it-asset-valuation-risk-assessment-and-control-implementation-model

Kelchner, L. (2018, December 28). What Are the Duties of an IT Department? Retrieved from careertrend.com: https://careertrend.com/13374589/what-are-the-duties-of-an-it-department

Romeyko, O. (2023, October 25). Top 15 Software Project Risks and Mitigation Examples in 2023. Retrieved from agentestudio.com: https://agentestudio.com/blog/software-development-project-risks

Sridhar, R. (2023, October 30). Your ultimate guide to project assumptions. Retrieved from rocketlane.com: https://www.rocketlane.com/blogs/project-assumptions

Tully, M. (2023, August 30). How to Make Accurate Project Assumptions: A Complete Guide. Retrieved from getrodeo.io: https://www.getrodeo.io/blog/project-assumptions

# Appendix

## Project Schedule



## Descriptive Statistics

1. Descriptive Statistics: Dynamically summarize key dataset features once the dataset has been updated.

Target: Display Descriptive Statistics in an interactive dashboard and Report.

## Key Performance Indicators (KPIs)

1. Predictive Model Accuracy: Measure the accuracy of the predictive model against current market values using Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), or R-squared value.

Target: Achieve a prediction accuracy rate of at least 80% over the specified period.

1. Return on Investment (ROI): Based on the model's predictions relative to the money invested. It is usually expressed as a percentage and calculated as follows:

ROI= ((Current Value of Investment−Cost of Investment)​/ Cost of Investment) ×100%

Target: Achieve a positive ROI, indicating that the gains made on the stock, based on model predictions, outweigh the initial investment cost with a minimum ROI of 10%.

1. Compound Annual Growth Rate (CAGR): Provides a normalized view of the investment's growth, smoothing out the effects of volatility and providing a compounded annual rate of return.

Target: Conservative Growth – Achieve a specific CAGR of 5% minimum.

Target: Moderate Growth – Achieve a minimum CAGR of 8% to 10%.

# These KPIs are to be displayed as Visualization in the Report and used for Feature Engineering #

1. Moving Average Crossing:

Integration: Utilize moving averages as features in your predictive model.

Target Alignment: Track the frequency and accuracy of stock prices crossing the moving average. Higher accuracy indicates better prediction of price movements.

Potential Benefit: Moving averages can help the model identify trend reversals and potential entry/exit points for trades.

1. Bollinger Bands Width:

Integration: Include Bollinger Bands width as a feature in your predictive model.

Target Alignment: Assess the width of Bollinger Bands to identify periods of volatility. A wider width during market fluctuations can indicate increased volatility.

Potential Benefit: Bollinger Bands can provide insights into volatility, helping the model adjust predictions during periods of market uncertainty.

1. OBV Trend Confirmation:

Integration: Incorporate On-Balance Volume (OBV) as a feature in your predictive model.

Target Alignment: Confirm the trend direction using OBV and aim for a high correlation between OBV trends and subsequent stock price movements.

Potential Benefit: OBV can act as a confirmation tool, helping the model validate trends based on volume movements.

1. RSI Overbought/Oversold Accuracy:

Integration: Integrate RSI values as features in your predictive model.

Target Alignment: Measure the accuracy of RSI in identifying overbought or oversold conditions. A higher accuracy rate indicates better prediction of price reversals.

Potential Benefit: RSI can provide signals for potential trend reversals, assisting the model in making more informed predictions.

1. MACD Histogram Stability:

Integration: Use the MACD histogram as a feature in your predictive model.

Target Alignment: Assess the stability of the MACD histogram to understand the consistency of trends. A lower variability suggests more reliable trend signals.

Potential Benefit: Stable MACD histograms can help the model capture sustained trends and provide more accurate predictions.