

High Level Design

Stock Market Analysis

Document Version Control

Date Issued	Version	Description	Author

Contents

Document Version Control.		2
Abstract.		4
1	Introduction	5
1.1	Why this High-Level Design Document?.	5
1.2	Scope.	5
1.3	Definitions	5
2	General Description.	6
2.1	Product Perspective	6
2.2	Problem statement	6
2.3	PROPOSED SOLUTION	6
2.4	FURTHER IMPROVEMENTS	6
2.5	Technical Requirements.	6
2.6	Data Requirements	7
2.7	Tools used.	8
2.8	Constraints	8
2.9	Assumptions.	9
3	Design Details	9
3.1	Process Flow.	10
3.2	Event log	11
3.3	Error Handling	11
3.4	Performance.	12
3.5	Reusability.	12
3.6	Application Compatibility	12
3.7	Resource Utilization	12
3.8	Deployment.	12
4	Dashboards.	13
4.1	KPIs (Key Performance Indicators)	13
5	Conclusion	13

Abstract

Stock market prediction has always caught the attention of many analysts and researchers. Popular theories suggest that stock markets are essentially a random walk and it is a fool's game to try and predict them. Predicting stock prices is a challenging problem in itself because of the number of variables which are involved. In the short term, the market behaves like a voting machine but in the longer term, it acts like a weighing machine and hence there is scope for predicting the market movements for a longer timeframe. Application of machine learning techniques and other algorithms for stock price analysis and forecasting is an area that shows great promise. In this paper, we first provide a concise review of stock markets and taxonomy of stock market prediction methods. We then focus on some of the research achievements in stock analysis and prediction. We discuss technical, fundamental, short- and long-term approaches used for stock analysis. Finally, we present some challenges and research opportunities in this field.

1 Introduction

1.1 Why this High-Level Design Document?

The purpose of this High-Level Design (HLD) Document is to add the necessary detail to the current project description to represent a suitable model for coding. This document is also intended to help detect contradictions prior to coding, and can be used as a reference manual for how the modules interact at a high level.

The HLD will:

- Present all of the design aspects and define them in detail
- Describe the user interface being implemented
- Describe the hardware and software interfaces
- Describe the performance requirements
- Include design features and the architecture of the project
- List and describe the non-functional attributes
 - like: o Security
 - o Reliability
 - o Maintainability
 - o Portability
 - o Reusability
 - o Application compatibility
 - o Resource utilization
 - o Serviceability

1.2 Scope

The HLD documentation presents the structure of the system, such as the database architecture, application architecture (layers), application flow (Navigation), and technology architecture. The HLD uses non-technical to mildly-technical terms which should be understandable to the administrators of the system.

1.3 Definitions

<i>Term</i>	<i>Description</i>
-------------	--------------------

Database	Collection of all the information monitored by this system
Power BI	Data Analysis Tool

2. General Description

2.1 Product Perspective

Financial Crime Data Analysis using Power BI with some past data.

2.2 Problem statement

A stock market, equity market, or share market is a gathering of buyers and sellers of stocks (also known as shares), which represent ownership claims on businesses; these may include securities listed on a public stock exchange as well as stock that is only traded privately, such as shares of private companies sold to investors through equity crowdfunding platforms. Typically, investments are undertaken with an investment strategy in mind.

The National Stock Exchange of India Ltd.'s main index is the NIFTY 50. (NSE). The Index follows the performance of a basket of blue-chip businesses, India's largest and most liquid equities. It contains 50 of the nearly 1600 businesses listed on the NSE, accounts for approximately 65 percent of its float-adjusted market value, and is an accurate representation of the Indian stock market.

The NIFTY 50 index encompasses significant sectors of the Indian economy and provides investment managers with exposure to the Indian market through a single, efficient portfolio. The Index has been in operation since April 1996 and is ideal for benchmarking, index funds, and index-based derivatives.

India Index Services and Products Ltd owns and manages the NIFTY 50. (IISL). IISL is India's first specialized corporation with an index as its primary product.

2.3 PROPOSED SOLUTION

- Perform ETL (Extract, Transform and Load) tasks to prepare the data for analysis purposes.
- Perform various data visualization techniques, analyze the market trends and submit a detailed report such that the insights can directly help a person make better investing decisions and would maximize his/her returns in the long run.

2.4 FURTHER IMPROVEMENTS

Increase the pages in dashboard and use high volume of data.

2.5 Technical Requirements

Hard Disk Power BI Report Server requires a minimum of 1 GB of available hard-disk space.

Addition space will be required on the database server that is hosting the report server database.

Memory Minimum: 1 GB

Recommended: At least 4 GB

Processor speed Minimum: x64 Processor: 1.4 GHz

Recommended: 2.0 GHz or faster

Processor type x64 Processor: AMD Opteron, AMD Athlon 64, Intel Xeon with Intel EM64T support, Intel Pentium IV with EM64T support

Operating system Windows Server 2019 Datacenter

Windows Server 2019 Standard

Windows Server 2016 Datacenter

Windows Server 2016 Standard

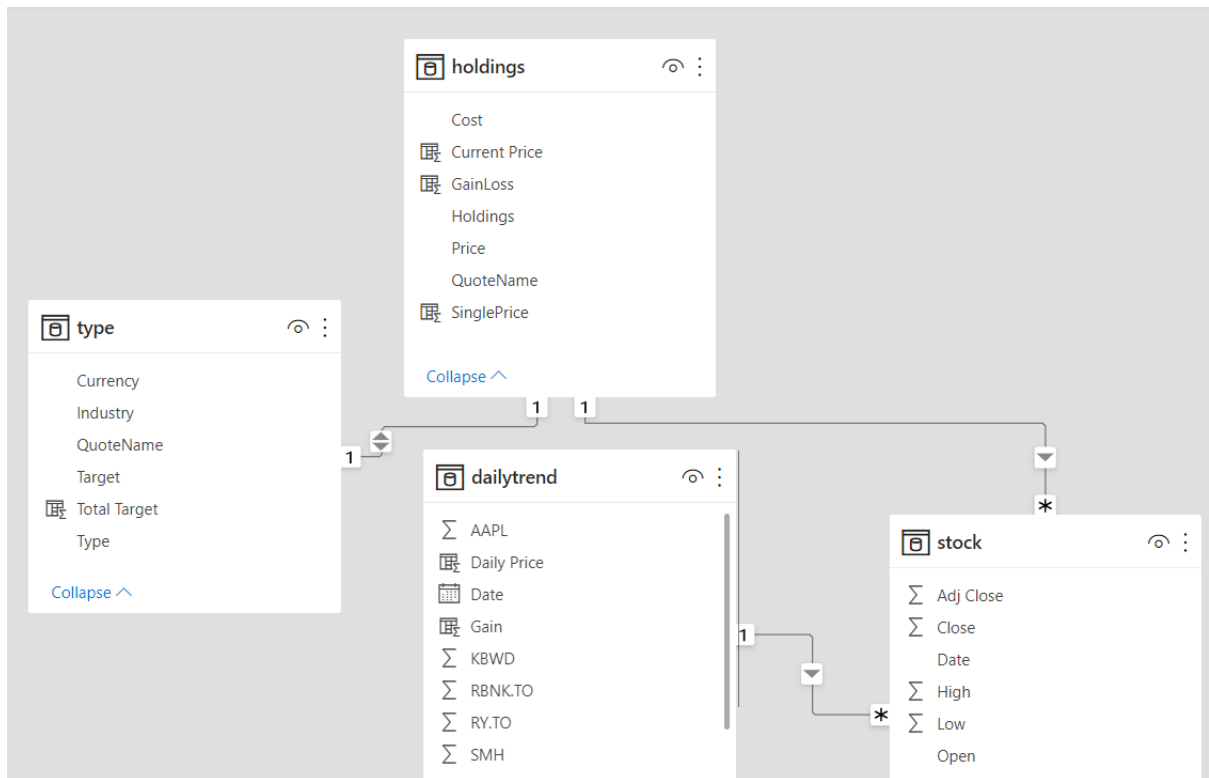
Windows 10 Home

Windows 10 Professional

Windows 10 Enterprise

Windows 11

2.6 Data Requirements



2.7 Tools used

Tools : Power BI, MS Excel

2.8 Constraints

Collecting meaningful data

Selecting the right tool

Consolidate data from multiple sources

2.9 Assumptions

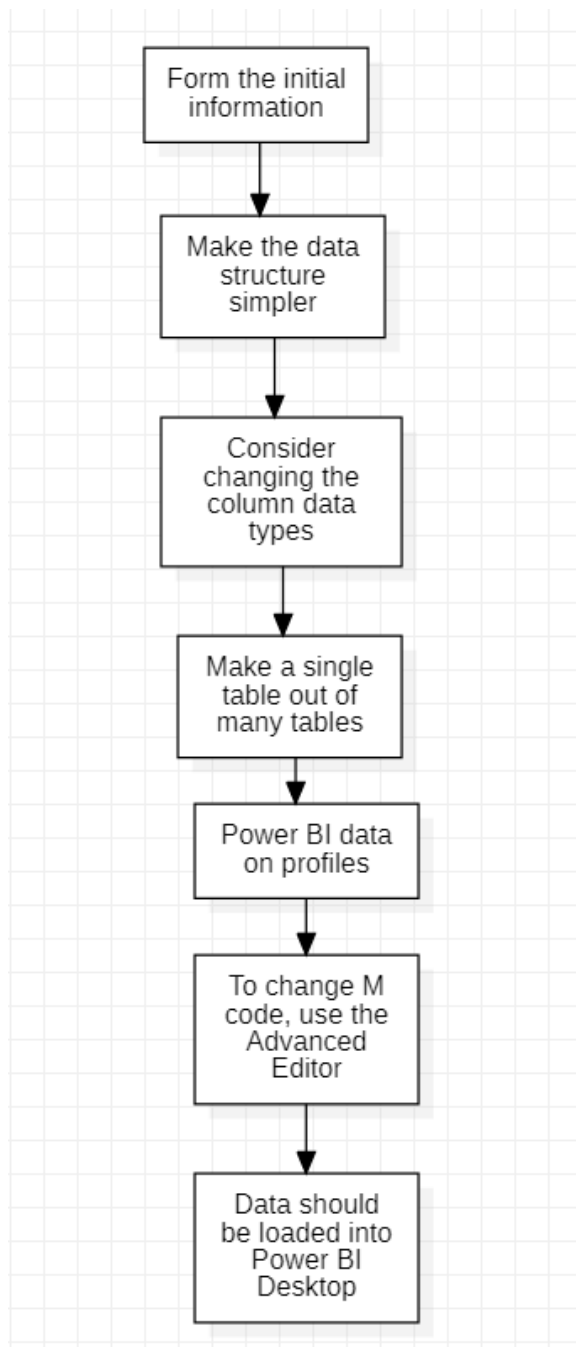
According to a study, the biggest obstacle in becoming a data-driven company lies in an organization's culture and not technologies. Only a 9.1% of executives have pointed out technology as a challenge in the path of data analysis. Many times, though top-level understand the importance of data analysis, they do not extend the desired support to their employees. Constant pressure and lack of support from the top and lower-level employees are among the most significant data analytics challenges.

3. Design Details

3.1 Process Flow

For identifying the different types of anomalies, we will use a deep learning base model. Below is the process flow diagram is as shown below.

Proposed methodology



Form the initial information

You can shape (alter) your imported data using Power Query Editor in Power BI Desktop. Renaming columns or tables, turning text to numbers, eliminating rows, setting the first row as headers, and much more are all possible. It's critical to structure your data so that it matches your requirements and can be used in reports.

Make the data structure simpler.

When you import data into Power BI Desktop from numerous sources, the data keeps its specified table and column names. You could wish to rename some of these items to make them more consistent, easier to work with, and meaningful to users. To make these name changes and simplify your data structure, use Power Query Editor in Power BI Desktop.

Consider changing the column data types.

Before loading data into a Power BI data model, it's a good idea to evaluate the column data types in Power Query Editor. You can modify a data type if you find it to be erroneous. You might also want to format the values in a column and modify the column's default summary.

Make a single table out of many tables.

Combining queries is useful since it allows you to append or merge data from several tables or searches. In the following situations, you can combine tables into a single table:

There are too many tables, making it difficult to manage a complex data model.

Several tables serve the same purpose.

Only one or two columns in a table can be moved to another table.

In a custom column, you wish to combine columns from multiple tables.

Power BI data on profiles

Profiling data entails finding abnormalities, inspecting and improving the underlying data structures, and querying data statistics like row counts, value distributions, minimum and maximum values, averages, and so on. This notion is significant because it helps you to structure and arrange data in such a way that dealing with it and determining its distribution is simple, making your job of working with data on the front end to create report elements nearly painless.

To change M code, use the Advanced Editor.

You generate a step in the Power Query process every time you shape data in Power Query. Where it makes sense, such steps can be reordered, eliminated, or adjusted. Although you probably generated each cleaning step using the graphical interface, Power Query uses the M

language behind the scenes. Using the Power Query Advanced Editor, you may read the combined steps. The M programming language is always open for direct reading and modification. It is not necessary to use M code in order to use Power Query. Although you will rarely need to create M code, it can be handy.

Data should be loaded into Power BI Desktop.

Select Open report and then Browse reports from the File tab.

3.2 Event log

The system should log every event so that the user will know what process is running internally.

Initial Step-By-Step Description:

1. The System identifies at what step logging required
2. The System should be able to log each and every system flow.
3. Developer can choose logging method. You can choose database logging/ File logging as well.
4. System should not hang even after using so many loggings. Logging just because we can easily debug issues so logging is mandatory to do.

3.3 Error Handling

Should errors be encountered, an explanation will be displayed as to what went wrong? An error will be defined as anything that falls outside the normal and intended usage.

4 Performance

4.1 Reusability

The code written and the components used should have the ability to be reused with no problems.

4.2 Application Compatibility

The different components for this project will be using Python as an interface between them. Each component will have its own task to perform, and it is the job of the Python to ensure proper transfer of information.

4.3 Resource Utilization

When any task is performed, it will likely use all the processing power available until that function is finished.

4.4 Deployment

Power BI

5 Dashboards

Dashboards will be implemented to display and indicate certain KPIs and relevant indicators for the unveiled problems that if not addressed in time could cause catastrophes of unimaginable impact.

5.1 KPIs (Key Performance Indicators)

- Gives better insights.
- Helps in easy flow.

6 Conclusion

The main purpose of our stock market data analysis is **to find meaning in data so that the derived knowledge can be used to make informed decisions.**