Stock Market Trends Analysis in Power BI

# Revision Number: 1.0

Last date of revision: 21/05/2022

Document Version Control

|  |  |  |  |
| --- | --- | --- | --- |
| Date Issued | Version | Description | Author |
| 21/05/2022 | 1 | Initial HLD — V1.0 | Prateek Jha |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |

**Contents**

[Document Version Control 2](#_TOC_250021)

[Abstract 4](#_TOC_250020)

1. [Introduction 5](#_TOC_250019)
   1. [Why this High-Level Design Document? 5](#_TOC_250018)
   2. [Scope 5](#_TOC_250017)
   3. [Definitions 5](#_TOC_250016)
2. [General Description 6](#_TOC_250015)
   1. [Product Perspective 6](#_TOC_250014)
   2. [Problem statement 6](#_TOC_250013)
   3. [PROPOSED SOLUTION 6](#_TOC_250012)
   4. [FURTHER IMPROVEMENTS 6](#_TOC_250011)
   5. [Technical Requirements 6](#_TOC_250010)
   6. [Data Requirements 7](#_TOC_250009)
   7. [Tools used 8](#_TOC_250008)
      1. Hardware Requirements 8
      2. ROS(Robotic Operating System) 9
   8. [Constraints 9](#_TOC_250007)
   9. [Assumptions 9](#_TOC_250006)
3. [Design Details 10](#_TOC_250005)
   1. [Process Flow 10](#_TOC_250004)
      1. [Model Training and Evaluation 10](#_TOC_250003)
      2. [Deployment Process 11](#_TOC_250002)
   2. [Event log 11](#_TOC_250001)
   3. [Error Handling 11](#_TOC_250000)
   4. Performance 12
   5. Reusability 12
   6. Application Compatibility 12
   7. Resource Utilization 12
   8. Deployment 12
4. Dashboards 13
   1. KPls (Key Performance lndicators) 13
5. Conclusion 14

# Abstract

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.

# Introduction

## **Why this High-Level Design Document:**

The main goal of this high level design (HLD) document is to detect all the inconsistencies in the design prior to code implementation and it can be used as a reference manual in case we need to understand the working of any manual for any of the modules.

This HLD will contain the following information needed:

* + - Present all of the design aspects and define them in detail
    - Describe the user interface being implemented
    - Description of user interfaces
    - Description of performance requirements
    - Project design as a whole unit
    - List and describe the non-functional attributes like:
      * + Reliability
        + Maintainability
        + Portability
        + Reusability
        + Application compatibility
        + Resource utilization
        + Serviceability

## Scope

This HLD documentation covers the overall architecture of the system, such as the database usage, application’s flow layer by layer both from UI and backend perspective and technology architecture. This HLD will contain non-technical to slightly technical terms related to the work but we will explain all of this here completely from scratch without any problem.

## Definitions

*Term Description*

|  |  |
| --- | --- |
| *PBI*  *ETL*  *CSV* | Microsoft Power BI |
| Used to denote the term Extract Transform and Load |
| Source files containing data to be put in PBI desktop directly |
|  |

# General Description

### **Problem statement**

We are trying to give the users a self service BI report that will give the users a brief idea as to help them focus and try to find the reasons and numbers that support the details about different stocks listed in NSE over the years and check out their performance based on the data provided for different stocks in equity share type and find the correct strategy for investment purposes.

It is very difficult to know about the overall idea just by looking at the raw file data in best possible manner but everyone understands data better in a visually appealing way. So we are designing a report in Power BI which would resolve all those issues in understanding data in raw form as it is generated.

### **PROPOSED SOLUTION**

We are proposing a solution where users with no or little background knowledge of self service report such as Power BI, one can simply come to this report that we’ve opened for our users, enter some filters on desired values based on the preferences from available list of options can get the most accurate results for available data with very less/no deviation.

The proposed solution here is a PBI report which is hosted in Power BI service which will capture all data changes from the source files generated for the same, perform an ETL using python and query editor steps in PBI itself to clean the data as much we can and using the logic appropriately for different visuals for representing different needs.

### **FURTHER IMPROVEMENTS**

Of course, this report will need to be refreshed with more data so a new file will be needed everytime with updated data which will be processed via python script to be dumped in MySQL DB and then finally to be refreshed in PBI service which is on cloud.

## **Technical Requirements**

* + - Data requirement completely depend on our problem statement.
    - We need data in a relational table format containing various columns/features where we will have one or multiple columns which will be used for accurate display of information.
    - We will need some columns as independent columns such as Date, Symbol, Series, open, high, low, last, close, VWAP, Volume, Turnover, Trades, Deliverable Volume, %Deliverable
    - The source of all this data will be a number of CSV files where each csv files contains data for one equity stock at a time and after collection of entire data, it will dumped directly into a Power query editor in Power BI desktop and use M code/ some UI to make data preprocessing changes. As far as technical knowledge is concerned, we need to be fundamentally cleared up on basics of some DAX functions in PBI for calculated columns and measures.

## **Data Requirements**

* + - Data requirement completely depend on our problem statement.
    - We data in a relational table format containing various columns/features where we will have one or multiple columns which will be used for prediction and there will be only one dependent feature/column which will be used for prediction.
    - The source of all these files will be a CSV file and after collection of entire data, it will dumped into a PBI desktop and from there, we will fetch it for cleaning and import in PBI accordingly needed per purpose.

## **Tools used**

Python programming language and frameworks such as NumPy, Pandas, Scikit-learn are used to build the entire backend part for machine learning modelling part.



* + - GitHub is used as version control system and contains the repository for this one.
    - Microsoft Power Bi for April 2022 release is currently used for the development of the report but can be used for changes in any upcoming versions of the release for this product.
* ~~- - - •~~ High Level Design 

## **Constraints**

The report created should be user friendly and should provide the user with correct details for which user is using the platform for with best accurate data to be shown in visual format.

**- 1**High Level Design (HLD) 

# Design Details

## **Process Flow**

For our created report to work properly, we will PBI desktop as ETL and reporting tool to fetch data from csv files and rest of steps and explained in detail below

Proposed methodology

