High Level Design (HLD)

Bank Marketing Campaign Analytics

Version: 1.0

Date: 01/10/2022

DOCUMENT CONTROL

CHANGE RECORD

Date Issue	Version	Description	Author
01/10/2022	1.0	Initial HLD – V 1.0	Aditya Gupta

APPROVAL STATUS

Version	Review Date	Review by	Approved by	Comments

CONTENTS

ABSTRACT

The data is related to direct marketing campaigns (phone calls) of a Portuguese banking institution. The classification goal is to get insights of the clients who subscribed to a term deposit so as to focus essential areas during bank marketing campaigns. The marketing campaigns were based on phone calls. Often, more than one contact to the same client was required, in order to access if the product (bank term deposit) would be subscribed or not.

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 THE DESGIN ASEPCTS AND DEFINE THEM IN DETAIL
- DESCRIBE THE USER INTERFACE BIENG IMPLEMENTED
- DESCRIBE THE HARDWARE AND SOFTWARE INTERFACE
- DESCRIBE THE PERFORMANCE REQUIREMENT
- INCLUDE DEFINE FEATURE AND ARCHITECHTURE OF THE PROJECT
- LIST AND DESCRIBE THE NON FUNCTIONAL ATTRIBUTES
  - SECURITY
  - RELIABLILTY
  - MAINTAINABILITY
  - PORTIBILITY
  - REUSEABILITY
  - APPLICATION COMPATIBILTY
  - RESOURCE UTILIZATION
  - 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 DEFINITION

TERM	DESCRIPTION
DATABASE	Collection of Information Monitored by the System
POWER BI	An interactive data visualization software focused on business intelligence
POWER QUERY	Power Query is a data transformation tool available in both Microsoft Excel and Power BI
ETL	Export Transform and Load

EDA	Exploratory Data Analysis
CSV	Comma Separated Values
ER	Entity Relation
BI	Business Intelligence

2. GENERAL DISCRIPTION

2.1 PROJECT PERSPECTIVE

Marketing Campaigns are important function in the banking industry. The Portuguese bank offers the service of Term deposits to its customers. In this project, Analytics of the Marketing campaign about the users who have either subscribed the term deposit or not are visualized.

2.2 PROBLEM STATEMENT

The objective of the project is to get insights through data visualization techniques to understand in detail about the customers.

2.3 PROPOSED SOLUTION

This project aims to use various visuals in Power BI to get insights and also get a visual understanding of the data.

2.4 FURTHER IMPROVEMENTS

More advancement can be done for better results.

2.5 TECHNICAL REQUIREMENTS

Power BI Desktop should be installed on the system. Further, to publish the report and to create dashboard, Power BI service account is required.
No hardware tool is required.

2.6 TOOLS USED

Power BI: Power BI is a collection of software services, apps, and connector that work together to turn unrelated sources of data into coherent, visually immersive and interactive insights. Data may be an Excel spreadsheet or a collection of cloud-based and on-premises hybrid data warehouses. Power BI lets you easily connect to data sources, visualize and discover what's important, and share that with anyone.

2.7 DATA REQUIREMENTS

The dataset bank-additional-full is in csv (comma separated values) format. Dataset is collected from https://archive.ics.uci.edu/ml/machine-learning-databases/00222/

2.8 CONSTRAINTS

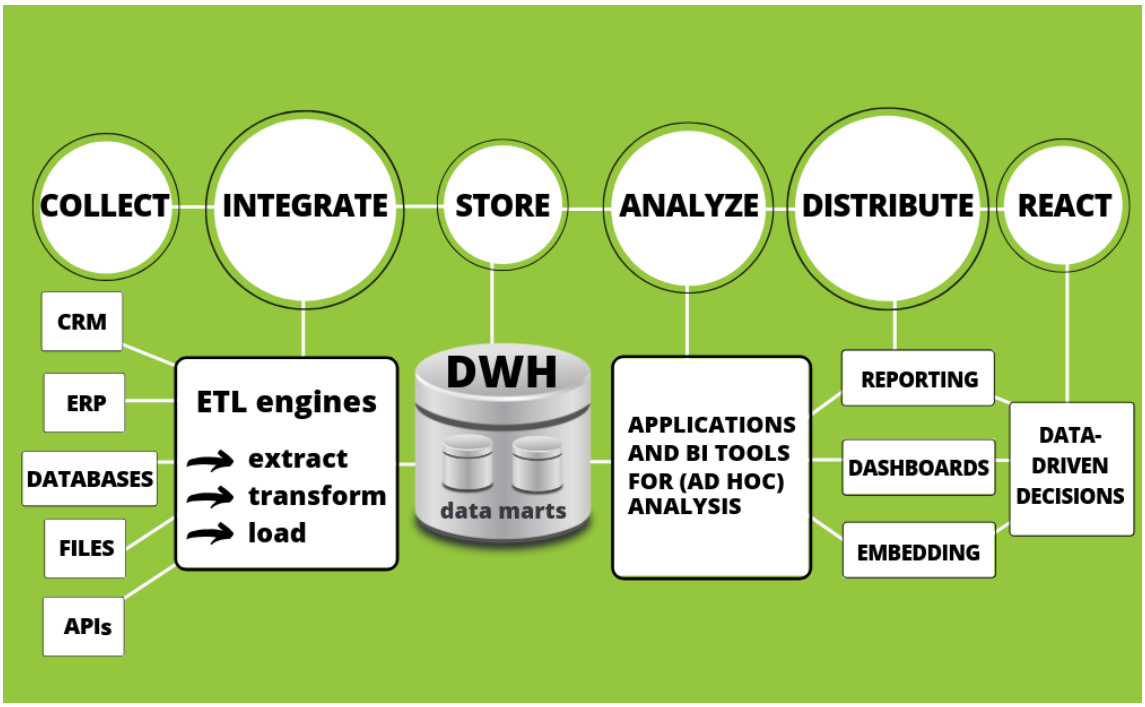
There are several missing values in some categorical attributes of dataset, all coded with the "unknown" label.

2.9 ASSUMPTIONS

These missing values in some categorical attributes of dataset are treated as a possible class label.

3. DESIGN DETAILS

3.1 PROCESS FLOW



A solid BI process flow consists of:

- Collection of data: The first step is related to the collection of relevant data from various external and internal sources which can be databases, ERP- or CRM systems, flat files, or APIs, just to name a few.
- Data integration: At this stage, the data collected is integrated into a centralized system, often with the help of ETL processes. Here the data is also cleaned and prepared for analysis.
- Storage of data: This is where data warehousing comes into the picture. A warehouse is a place in which structured data is stored. It makes it available for querying and analysis.

- Data analysis: After the information is processed, stored, and cleaned it is ready to be analyzed. With the help of the right tools, the data is visualized and used for strategic decision-making.
- Distribution of data: The data, now in the form of graphs and charts, is distributed in different formats. This can be online reporting, dashboarding, or embedding solutions.
- Reaction based on insights: The final stage of the architecture process is to extract actionable insights from the data and use them to make improved decisions to ensure company growth.

4. OPTIMIZATION

DATA STRATEGY DRIVES PERFORMANCE

- Minimize the number of fields
- Minimize the number of records
- Optimize extracts to speed up future queries by materializing calculations, removing columns and the use of accelerated views

REDUCE THE MARKS (DATA POINTS) IN YOUR VIEW

- Practice guided analytics. There's no need to fit everything planned to show in a single view. Compile related views and connect them with action filters to travel from overview to highly-granular views at the speed of thought.
- Remove unneeded dimensions from the detail shelf.
- Explore. Try displaying your data in different types of views.

LIMIT YOUR FILTERS BY NUMBER AND TYPE

- Reduce the number of filters in use. Excessive filters on a view will create a more complex query, which takes longer to return results. Double-check your filters and remove any that aren't necessary.
- Use an include filter. Exclude filters load the entire domain of a dimension, while include filters do not. An include filter runs much faster than an exclude filter, especially for dimensions with many members.
- Use a continuous date filter. Continuous date filters (relative and range-of-date filters) can take advantage of the indexing properties in your database and are faster than discrete date filters.
- Use Boolean or numeric filters. Computers process integers and Booleans (t/f) much faster than strings.
- Use parameters and action filters. These reduce the query load (and work across data sources).

OPTIMIZE AND MATERIALIZE YOUR CALCULATIONS

- Perform calculations in the database
- Reduce the number of nested calculations.
- Reduce the granularity of LOD or table calculations in the view. The more granular the calculation, the longer it takes.
- Where possible, use MIN or MAX instead of AVG. AVG requires more processing than MIN or MAX.
- Use Booleans or numeric calculations instead of string calculations. Computers can process integers and Booleans (t/f) much faster than strings.
  Boolean>Int>Float>Date>DateTime>String

5. KPIs

Dashboards will be implemented to display and indicate certain KPIs and relevant indicators for the business problem. As and when, the system starts to capture the data for a user, the dashboards will be included to display charts over time with progress on various indicators or factors.

5.1 KPIs (Key Performance Indicators)

Key indicators displaying a summary of the Bank marketing Campaign's results and subscriber's information based on various parameters:

1. Percentage conversion based on various parameters.
2. Number of subscriber's based on various parameters.

6. CONCLUSION

The Dashboard will be providing an interactive platform where insights of bank customers are presented in the form of visuals. The decision makers can identify the trend & predict to make future decisions based on the insights such as where to focus most in marketing campaigns to increase the subscribers in the least possible effort.