Low Level Design (LLD)

Budget Sales Analysis

Revision Number - 1.0

Last Date of Revision - 11/07/2022

Ayush Jaroli

Document Control

Date	Version	Description	Author
11/07/2022	1.0	Introduction, Problem Statement, Dataset Information, Architecture	Ayush Jaroli

Contents

1. Introduction

1.1 What is Low Level Design Document?

The goal of the Low-level design document (LLDD) is to give the internal logic design of the actual program code for the Heart Disease Diagnostic Analysis dashboard. LLDD describes the class diagrams with the methods and relations between classes and programs specs. It describes the modules so that the programmer can directly code the program from the document.

1.2 What is Scope?

Low-level design (LLD) is a component-level design process that follows a step-by-step refinement process. The process can be used for designing data structures, required software architecture, source code and ultimately, performance algorithms. Overall, the data organization may be defined during requirement analysis and then refined during data design work.

1.3 Project Introduction

Budget and Sales are by far most important attributes that defines a business's success and failure. Therefore, it is very important to keep a track on various features related to these attributes to keep on increasing the Sales and to allocate the Budget so that it can be utilized wisely and efficiently. So, it is very important for businesses to dig deep into the customer, sales, budget and product data to make better marketing strategy, to know the target customers, to make market friendly product upgrades and to keep a strong track on the budget efficiency. Good data driven systems can help achieve these goals and take the businesses forward towards success.

2. Problem Statement

Domain Sale process is structured to help potential buyers purchase the domain they want immediately without the hassle of contacting the seller directly. A seller lists a domain for sale at a specific price in our Marketplace. An interested buyer sees this domain for sale and decides to buy it. The dataset consists of information about Budget distribution of year 2016, previous customers, product details and sales details. As there is a huge amount of data with various features present with us, we have to find various key insights that can be useful in making business decisions and present the data in easy-to-understand visualizations.

3. Dataset Information

Various Excel files available in the dataset: -

- **Customer Data:** This file consists of the features related to the data about the customers i.e., 'CustomerKey', 'FirstName', 'LastName', 'FullName', 'BirthDate', 'MaritalStatus', 'Gender', 'YearlyIncome', 'TotalChildren', 'NumberChildrenAtHome', 'Education', 'Occupation', HouseOwnerFlag', 'NumberCarsOwned', AdressLine1', 'DateFirstPurchase', 'CommuteDistance'.

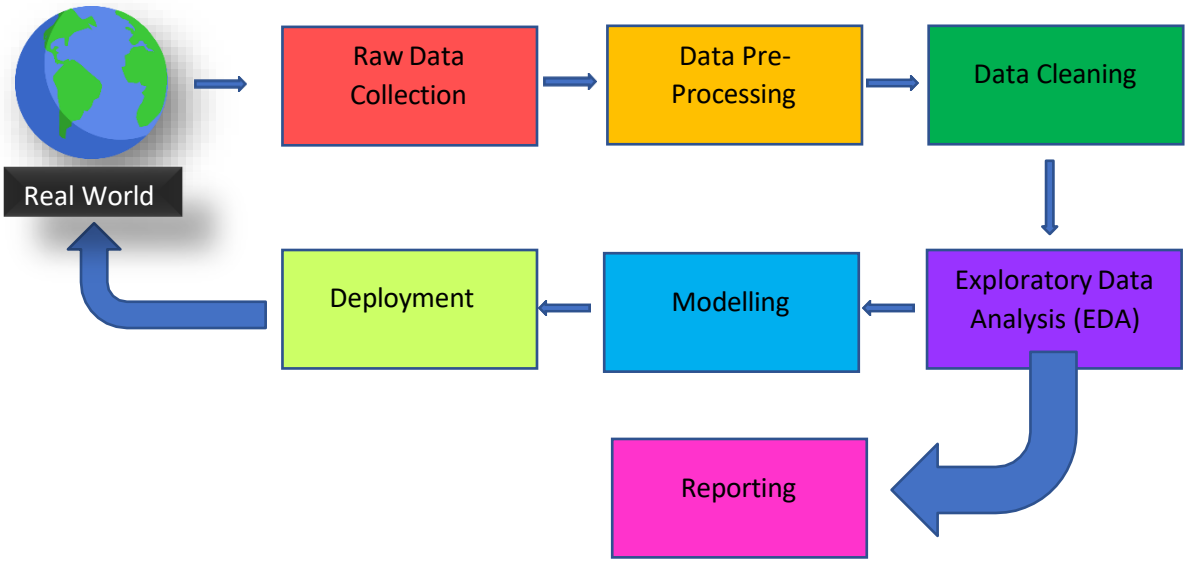
- **Product Data:** This file consists of the features related to the data about the product i.e., 'ProductKey', 'ProductName', 'Subcategory', 'Category', 'StandardCost', 'Color', 'List Price', 'DaysToManufacture', 'ProductLine', 'ModelName', 'Photo', 'ProductDescription', 'StartDate'.

- **Sales Data:** This file consists of the features related to the data about the Sales i.e., 'ProductKey', 'OrderDate', 'ShipDate', 'CustomerKey', 'PromotionKey', 'SalesTerritoryKey', 'SalesOrderNumber', 'SalesOrderLineNumber', 'OrderQuantity', 'UnitPrice', 'TotalProductCost', 'SalesAmount', 'TaxAmt'.

- **Territory Data:** This file consists of the features related to the data about the Territory i.e., 'SalesTerritoryKey', 'Region', 'Country', 'Group', 'RegionImage'.

- **Budget Data:** This file consists of the features related to the data about the Budget 2016 i.e., 'Category', 'Subcategory', 'ProductName', 'ProductKey', 'Jan, 2016', 'Feb, 2016', 'Mar, 2016', 'Apr, 2016', 'May, 2016', 'Jun, 2016', 'Jul, 2016', 'Aug, 2016', 'Sep, 2016', 'Oct, 2016', 'Nov, 2016', 'Dec, 2016', 'Grand Total'.

4. Architecture



4.1 Architecture Description

1. **Raw Data Collection**
   The Dataset was taken from iNeuron's Provided Project Description Document.
   https://drive.google.com/drive/folders/165Pjmfb9W9PGy0rZjHEA22LW0Lt3Y-Q8

2. **Data Pre-Processing**
   Before building any model, it is crucial to perform data pre-processing to feed the correct data to the model to learn and predict. Model performance depends on the quality of data feed to the model to train.
   This Process includes-
   a) Handling Null/Missing Values
   b) Handling Skewed Data
   c) Outliers Detection and Removal

3. **Data Cleaning**
   Data cleaning is the process of fixing or removing incorrect, corrupted, incorrectly formatted, duplicate, or incomplete data within a dataset.
   a) Remove duplicate or irrelevant observations
   b) Filter unwanted outliers
   c) Renaming required attributes

4. Exploratory Data Analysis (EDA)

Exploratory Data Analysis refers to the critical process of performing initial investigations on data to discover patterns, spot anomalies, test hypothesis and to check assumptions with the help of summary statistics and graphical representations.

5. Reporting

Reporting is a most important and underrated skill of a data analytics field. Because being a Data Analyst you should be good in easy and self-explanatory report because your model will be used by many stakeholders who are not from technical background.

a) High Level Design Document (HLD)
b) Low Level Design Document (LLD)
c) Architecture
d) Wireframe
e) Detailed Project Report
f) Power Point Presentation

6. Modelling

Data Modelling is the process of analyzing the data objects and their relationship to the other objects. It is used to analyze the data requirements that are required for the business processes. The data models are created for the data to be stored in a database. The Data Model's main focus is on what data is needed and how we have to organize data rather than what operations we have to perform.

7. Deployment

We created a Power BI Dashboard.