iNeuron

Low Level Design (LLD)

Budget Sales Analysis

Revision Number: 1.6
Last date of revision: 29/01/2023

Pranit Patil

Document Version Control

Date Issued	Version	Description	Author
14/12/2022	1.0	Abstract, Introduction	Pranit Patil
18/12/2022	1.1	Problem Statement	Pranit Patil
25/12/2022	1.2	Dataset Information	Pranit Patil
30/12/2022	1.3	Architecture	Pranit Patil
6/01/2023	1.4	Architecture Description	Pranit Patil
25/01/2023	1.5	Deployment	Pranit Patil
29/01/2023	1.6	Final Revision	Pranit Patil

Contents

1 Introduction

1.1 Why this Low-Level Design Document?

The purpose of this document is to present a detailed description of the heart disease diagnostic analysis. LLD 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. This document is intended for both the stakeholders and the developers of the system and will be proposed to the higher management for its approval.

The LLD will be focusing on the below objectives:

- Problem Understanding.

- Data Acquisition.

- Data Pre-Processing and Exploratory Analysis.

- Dashboard report for important activities.

1.2 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.

1.4 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, 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.

2 Technical specifications

2.1 Dataset

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', 'SalesTerritor yKey', 'SalesOrderNumber', 'SalesOrderLineNumber', 'OrderQuantity', 'UnitPrice', 'T otalProductCost', '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'.

3 Proposed Solution

3.1 Architecture
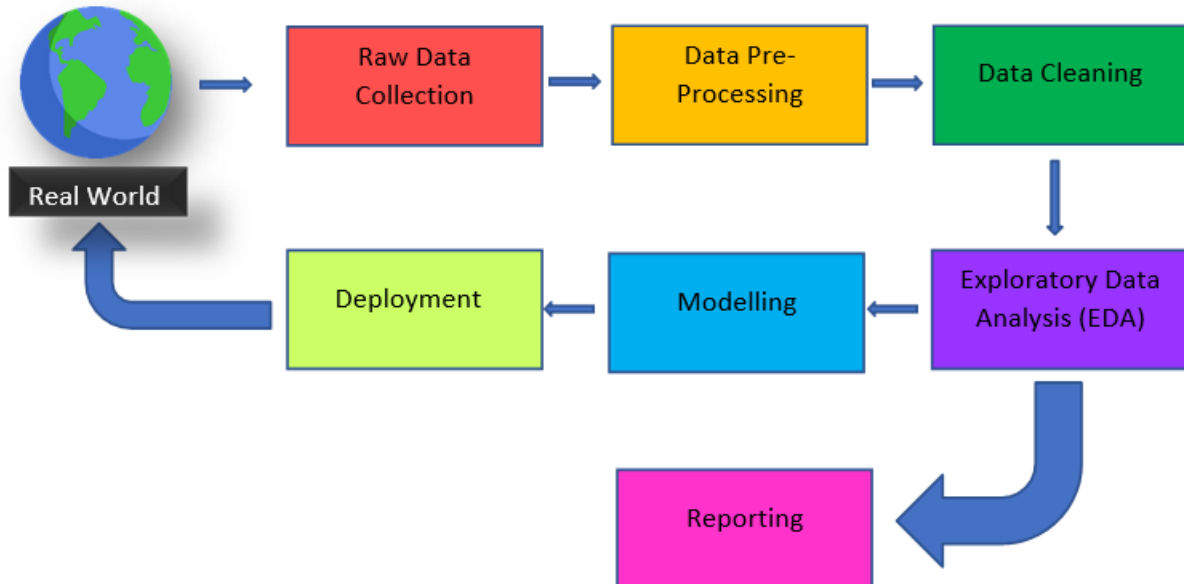
4 Architecture Description

4.1 Raw Data Collection:

The Dataset was taken from iNeuron's Provided Project Description Document.
Dataset link: Budget-Sales-Analysis/Datasets_at_main_·_Pp11112000/Budget-Sales-Analysis (github.com)

4.2 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 feeded to the model to train. This Process includes:

- Handling Null/Missing Values
- Handling Skewed Data
- Outliers Detection and Removal

4.2.1 Data Cleaning

Data cleaning is the process of fixing or removing incorrect, corrupted, incorrectly formatted, duplicate, or incomplete data within a dataset.

- Remove duplicate or irrelevant observations.
- Filter unwanted outliers.
- Renaming required attributes.

4.3 Exploratory Data Analysis

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.

4.4 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

4.5 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.

4.6 Deployment

We created a Power BI Dashboard and published it on Power BI Service.
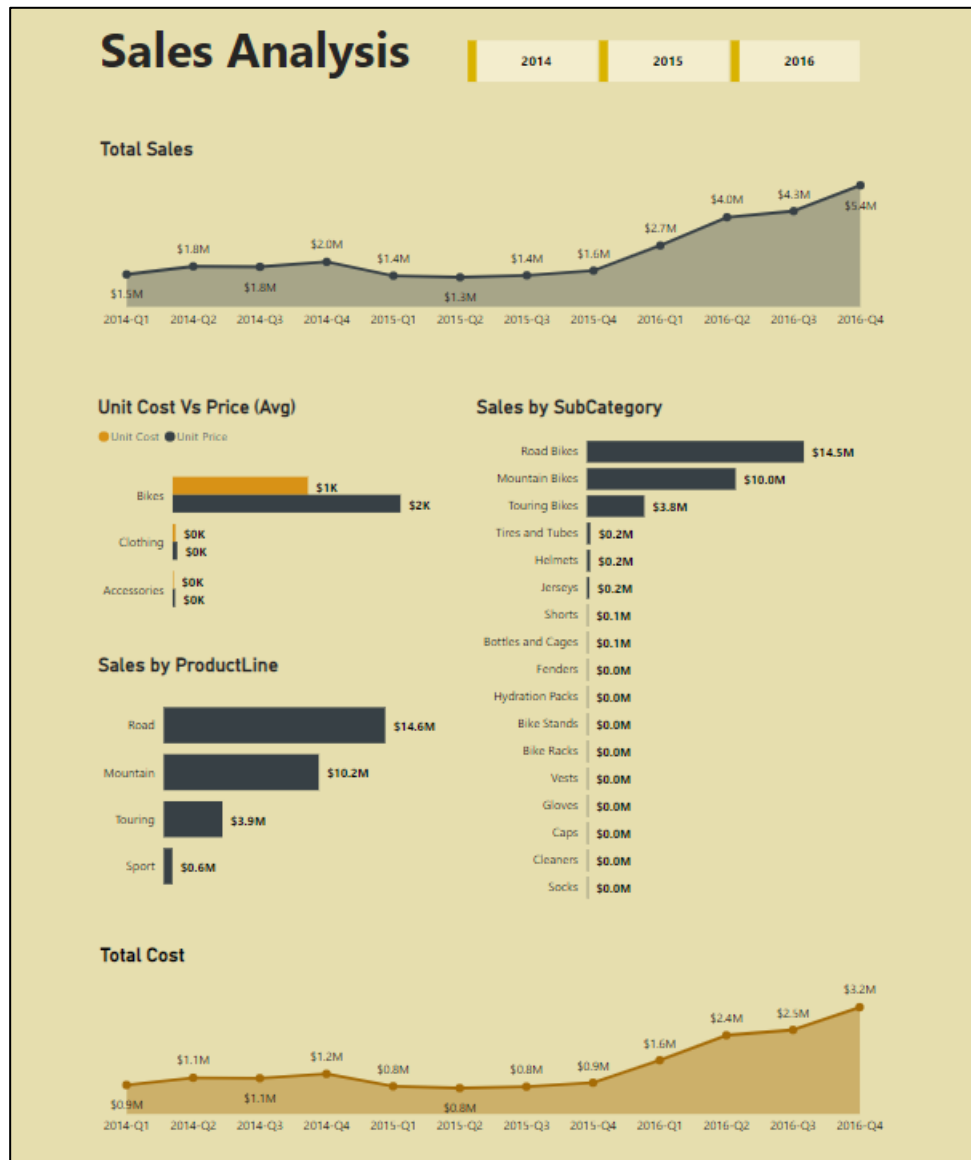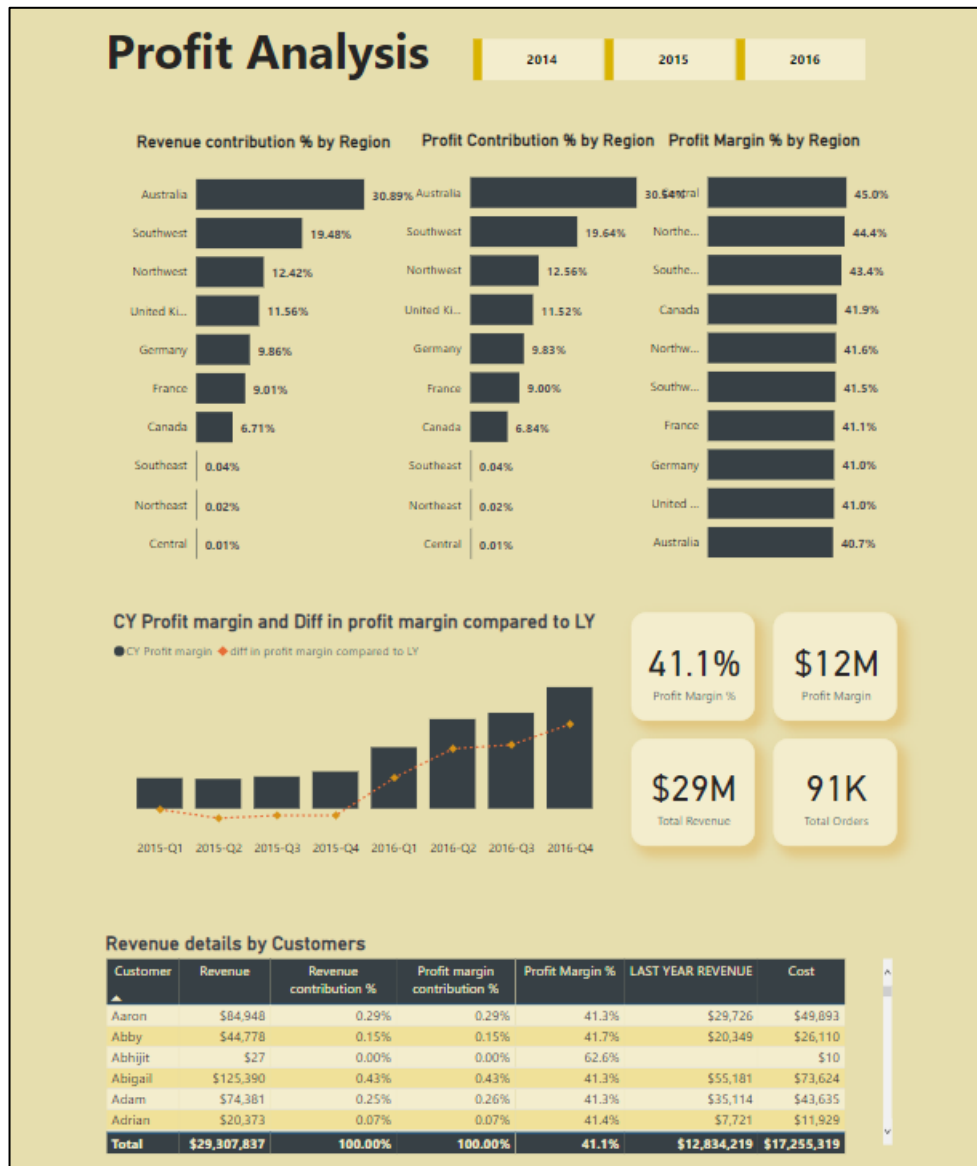
# Variance Analysis

Sales by Category & Variance to Target

Sales by Product & Variance to Target

Category	Sales	Target sales	Variance	Variance %
⊟ Clothing	$322,677	₹ 333,741	($11,064)	-3.3%
Vests	$33,465	₹ 34,565	($1,101)	
Socks	$4,882	₹ 5,449	($567)	
Shorts	$67,050	₹ 68,453	($1,403)	
Jerseys	$165,066	₹ 170,938	($5,872)	
Gloves	$33,380	₹ 33,965	($585)	
Caps	$18,834	₹ 20,371	($1,537)	
⊟ Bikes	$28,318,145	₹ 15,848,226	$12,469,919	78.7%
Touring Bikes	$3,844,801	₹ 4,055,965	($211,164)	
Road Bikes	$14,520,584	₹ 5,257,897	$9,262,687	
Mountain Bikes	$9,952,760	₹ 6,534,364	$3,418,396	
⊟ Accessories	$667,015	₹ 687,607	($20,592)	-3.0%
Tires and Tubes	$231,300	₹ 242,777	($11,477)	
Hydration Packs	$38,823	₹ 39,803	($980)	
Helmets	$215,923	₹ 221,905	($5,982)	
Fenders	$44,268	₹ 43,670	$598	
Cleaners	$6,869	₹ 6,465	$404	
Bottles and Cages	$55,031	₹ 57,369	($2,338)	
Bike Stands	$37,842	₹ 38,757	($915)	
Bike Racks	$36,960	₹ 36,861	$99	
Total	$29,307,837	₹ 16,869,574	$12,438,263	73.7%

Variance by Months
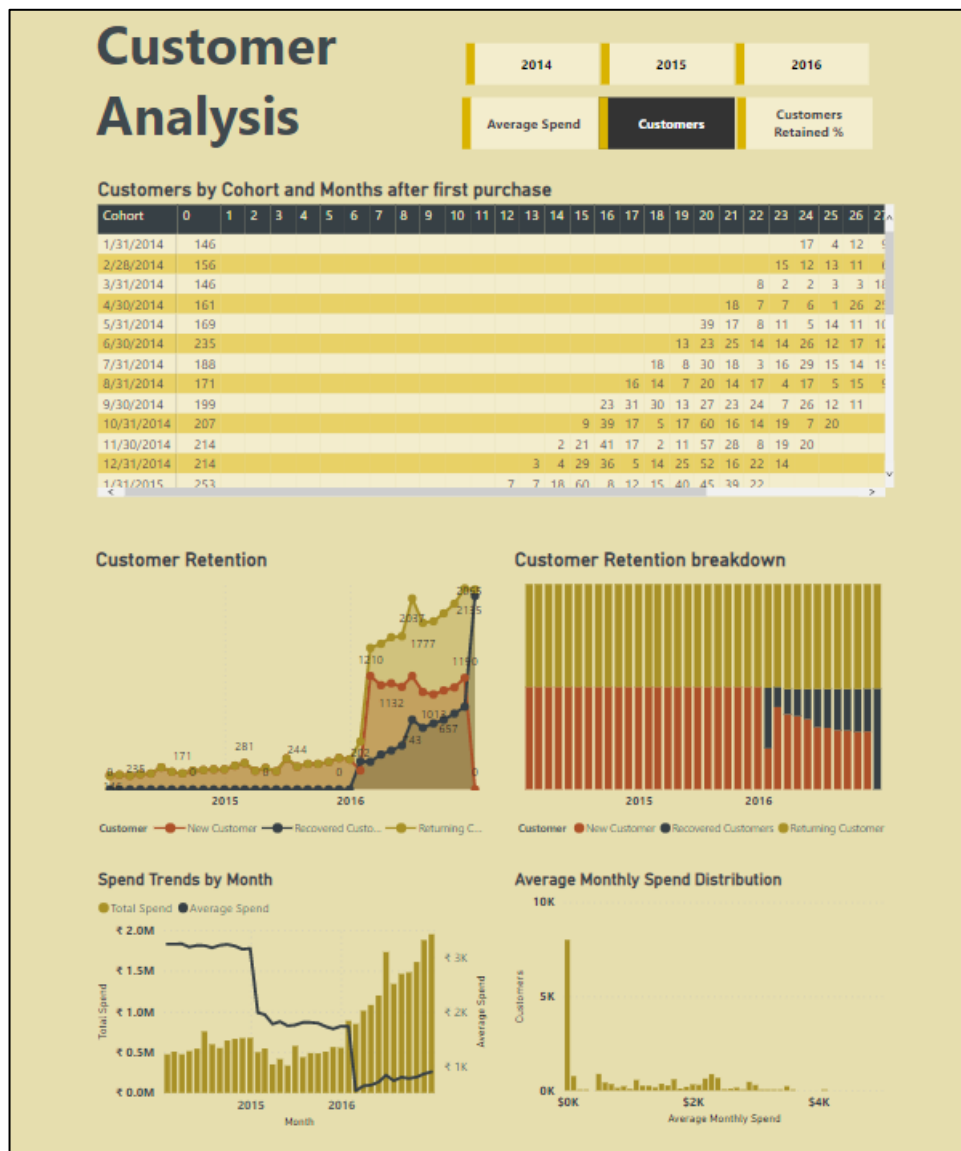
5 Key performance indicators (KPI)

Dashboards will be implemented to display and indicate certain KPIs and relevant indicators for the sales. When the system starts to capture the historical/periodic data for a user, the dashboards will be included to display charts over time with progress on various indicators or factors.

Key indicators displaying a summary of the sales and its relationship with different metrics.

1. Sales trend line
2. Cost trend line
3. Average unit cost and price
4. Revenue generated by Subcategory
5. Sales by Product Line
6. Revenue contribution by region
7. Profit contribution by region
8. Profit % by region
9. Current year profit margin vs difference in last year's profit margin
10. Total orders
11. Total revenue
12. Variance to target comparison by category
13. Variance by month line chart
14. Actual sales and target sales matrix
15. Cohort analysis table
16. Customer retention line chart
17. Monthly spending trend
18. Average monthly spend distribution