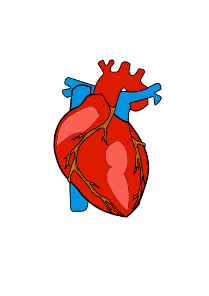
**Low Level Design (LLD)**

**Amazon Sales Data Analysis**



**Revision Number - 1.2**

**Last Date of Revision - 12/08/2022**

**Rajat Chauhan**

# Document Control

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Date** |  | **Version** | **Description** | **Author** |
| 21/07/2022 | 1.0 |  | Introduction,  Problem Statement | Rajat Chauhan |
| 02/08/2022 | 1.1 |  | Dataset Information,  Architecture  Description | Rajat Chauhan |
| 12/08/2022 | 1.2 |  | Final Revision | Rajat Chauhan |

**Contents**

Document Version Control…………………………………………………………………………2

1. Introduction……………………………………………………………………………………….4
   1. What is Low Level Design Document?.................................................................4
   2. Scope……………………………………………………………………………………4
   3. Project Introduction……………………………………………………………………4
2. Problem Statement.……………………………………………………………………………...5
3. Dataset Information……………………………………………………………………….……...5
4. Architecture…………………………………………………………………………………….…6
   1. Architecture Description …………………………………………….…....................6

# 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

Sales management has gained importance to meet increasing competition and the need

for improved methods of distribution to reduce cost and to increase profits. Sales

management today is the most important function in a commercial and business

enterprise.

# 2. Problem Statement

The goal of this project is to Do ETL : Extract-Transform-Load some Amazon dataset and to find

Sales-trend -> month wise , year wise , yearly\_month wise

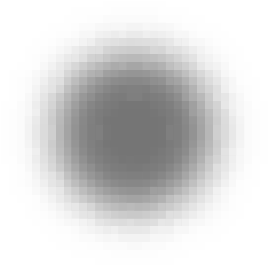
Find key metrics and factors and show the meaningful relationships between attributes.

To Do your own research and come up with your findings.

# 3. Dataset Information

* Region - Gives the different regions of the sales all over the world.
* Country - The country of the sale item.
* Item type - Tells about different item types that we have.
* Sales Channel - There are two sales channel modes - Online and Offline
* Order Priority - Four Letters C H L M used to represent order priority.
* Order Date - The date of placing the order in mixed format. (Non Uniform date formats)
* Ship Date - The shipping date of the product.
* Units Sold - Number of units of the particular item sold.
* Unit Price - Unit Price of the particular item at which it is sold.
* Unit Cost - Unit Cost i.e. per unit amount spend to make the product.
* Total Revenue - Total revenue generated for the item for the given region and country.
* Total Cost - Similarly Total cost
* Total Profit - and Total Profit

# 4. Architecture



Raw Data



Collection



Data



Pre



-



Processing



Data Cleaning



Exploratory Data



Analysis



EDA



)



(



Modelling



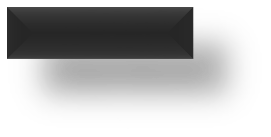
Deployment



R



eporting



Real World



## 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/1FkmFVL8wlJmQWP1z52TD8PlhOJhitTyI?usp=s>

### 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-

1. Handling Null/Missing Values
2. Handling Skewed Data
3. 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.

1. Remove duplicate or irrelevant observations
2. Filter unwanted outliers
3. 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.

1. High Level Design Document (HLD)
2. Low Level Design Document (LLD)
3. Architecture
4. Wireframe
5. Detailed Project Report
6. Power Point Presentation

### 6. Modelling

Data Modelling is the process of analysing the data objects and their relationship to the other objects. It is used to analyse 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

Final Tableau Dashboard

Chart

Description automatically generated