# Low Level Design (LLD)

# Analysing Google App Store Dataset

**Revision Number - 1.2**

**Last Date of Revision - 00/03/2023**

**Siddharth Bhat**

## Document Control

| Date | Version | Description | Author |
|---|---|---|---|
| 00/03/2023 | 1.0 | Introduction, Problem Statement | Siddharth Bhat |
| 00/03/2023 | 1.1 | Dataset Information, Architecture Description | Siddharth Bhat |
| 00/03/2023 | 1.2 | Final Revision | Siddharth Bhat |

# Contents

# 1. Introduction

## 1.1 What is Low-Level Design Document?

The goal of the Low-level design document (LLD) is to give the internal logic design of the actual program code for the Sales Analysis dashboard. 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.

## 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

This is a project about Google App Store Analysis. in this project, we aim on analysing Google play store that provides a particular app description and data such as reviews, ratings, price and number of downloads. The objective of this is to analyse the desire of the customer through the reviews provided in the feedback section and apps trend in the market to help the organization & developers. To this end, we provide an idea about app that managed to get maximum and minimum number of downloads and predicting the category of apps that is most likely to be downloaded in the coming years.

## 2. Problem Statement

Technology is the increasing need nowadays and used everywhere. One of the feature of Technology is android. Which we all use in our daily life. Android is a mobile operating system based on a modified version of the Linux kernel and other open source software, designed primarily for touchscreen mobile devices such as smartphones and tablets. Do ETL : Extract-Transform-Load the dataset and find for me some information from this large data. This is form of data mining.

## 3. Dataset Information

**Apps:** Different Types of Application

**Category:** Category Wise Application

**Rating:** Applications Performance Rating

**Reviews:** Comments of Application

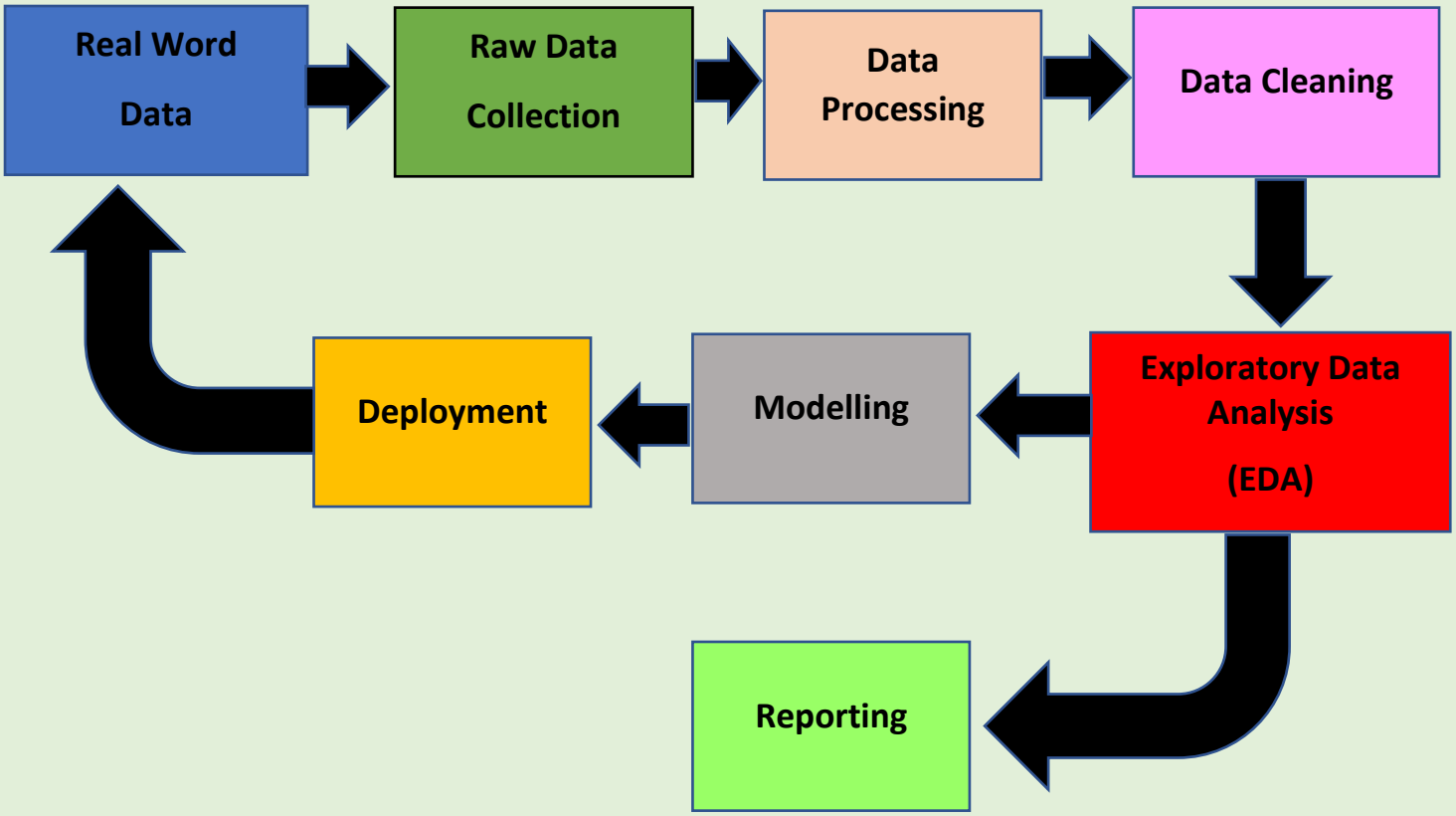
**Installs:** Number of Installation App

**Type:** Type of App (Free or Paid)

**Size:** Size of App

**Android Ver:** Android Version of App

## 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/file/d/1QpHsrW3XudFfarTD-e8jnMjGubdRPqvS/view

### 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 fed 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

Exploratory Data Analysis refers to the critical process of performing initial investigations on data to discover patterns, spot anomalies, test hypotheses and 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 the easy and self -explanatory report because your model will be used by many stakeholders who are not from a technical background.

a) High-Level Design Document (HLD)
b) Low-Level Design Document (LLD)
c) Architecture
d) Wireframe
e) Detailed Project Report
f) PowerPoint 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 to store the data 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