Low Level Design (LLD)
Google Play Store Analysis

Document Control

Date	Version	Description	Author
06/08/2022	1.0	Introduction, Problem Statement	Anil S Adiga
07/08/2022	1.1	Dataset Information, Architecture Description	Anil S Adiga
08/08/2022	1.2	Final Revision	Anil S Adiga

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 Google App Store 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.

2. Problem Statement

The are millions of apps on the play store for the users to use and each app has its own features and characteristics to know if its popular among the users. The aim of this project is to analyze the different features present in the dataset to gain insights about the popularity of the app in different categories and genres.

3. Dataset Information

app: The name of the apps in the play store

category: The app category to which it belongs to

rating: The ratings for each app by users

reviews: The count of reviews by users for a particular app

size: The size of the app in KB's and MB's

installs: The count of installs of a particular app

type: The type of the app, whether the app is a free or paid version.

price: The price of each app for the paid app type.

content category: The type of content category to which it belongs to.
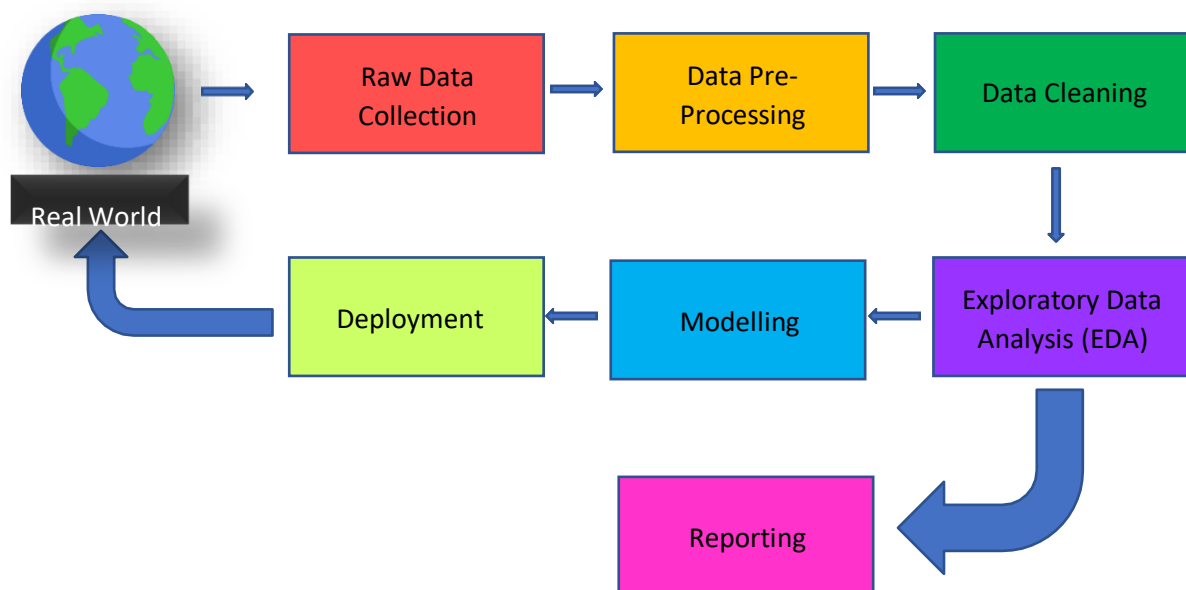
genre: The different genres available for the apps

last updated: The date at which the app was last updated in the play store

current version: The current version of the app after updating the app in the play store

android version: Android version of the app to check for compatibility.

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 feeded 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 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

I created a Power BI Dashboard