Low Level Design

Entertainer Data Analysis

Written By	Avs Jagannath
Document Version	0.3
Last Revised Date	

DOCUMENT CONTROL

Change Record:

VERSION	DATE	AUTHOR	COMMENTS
0.1	07-July-2024	Avs Jagannath	Introduction & Architecture define
0.2	07-July-2024	Avs Jagannath	Architecture & Architecture description appended & updated

Reviews:

VERSION	DATE	REVIWERS	COMMENTS
0.2	08-JULY-2024	Avs Jagannath	Unit test cases to be added

Approval Status:

VERSION	REVIEW DATE	REVIEWED BY		APPROVED BY	COMMENTS

Contents :

1. Introduction
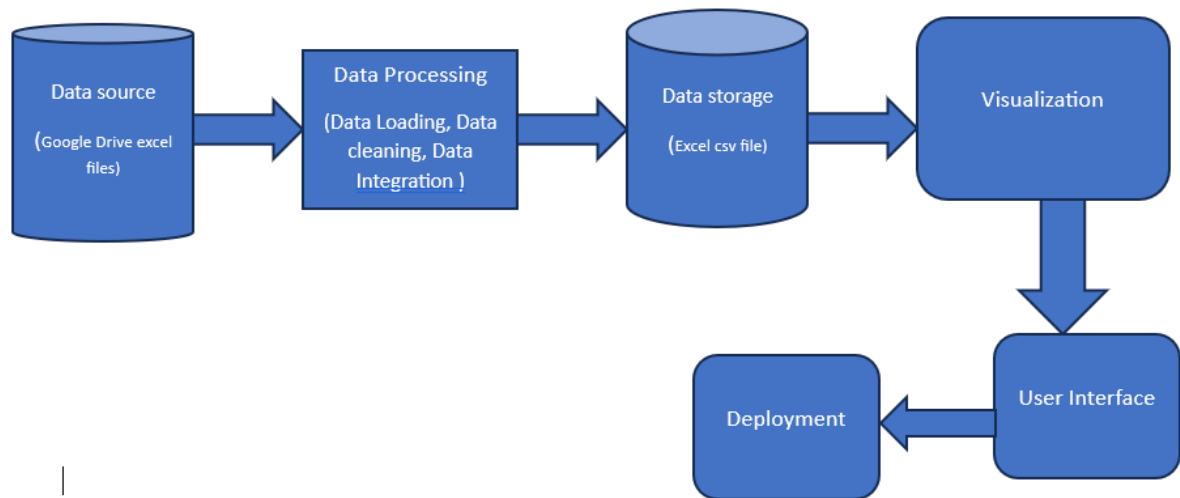
1.1 What is Low-Level Design Document?

This LLD document provides the internal logic design of the actual program code for the Entertainer Data Analysis dashboard. It describes the data flow, processing steps, and integration with visualization tools in detail.

1.2 Scope

This document covers the step-by-step process of data integration, transformation, analysis, and visualization for the Entertainer Data Analysis project.

2. Architecture

```
Data source          Data Processing        Data storage         Visualization
(Google Drive excel  (Data Loading, Data    (Excel csv file)
files)               cleaning, Data
                     Integration )
                                                                 User Interface
Deployment
```

3. Architecture Description

3.1 Data Description

The dataset contains information about entertainers including:

- Name

- Gender

- Birth Year

- Year of Breakthrough

- Breakthrough Name

- Year of First Major Award

- Age at Breakthrough

- Decade of Breakthrough

3.2 Data Integration

- Load data from multiple Excel files using Python's pandas library

- Combine data from different sources based on the 'Entertainer' column

3.3 Data Transformation

- Clean data (handle missing values, remove duplicates)

- Calculate additional features (Age, Decade of Breakthrough)

- Standardize data formats across all sources

3.4 Data Storage

- Store processed data in a CSV file for easy import into visualization tools

- Optionally, store data in a SQL database for more complex queries

3.5 Connection with Visualization Tool

- For Power BI:

  - Import data from CSV file or connect to SQL database

  - Set up data model in Power BI

  - Create DAX measures as needed

3.6 Export Data from Database

If using a database, export data as a CSV file for backup or alternative use

3.7 Deployment

- For Power BI:

  - Publish to Power BI Service

  - Set up scheduled refresh for the dataset

4. Unit Test Cases

Test Case Description	Expected Results
Decade of Breakthrough Slicer	Displays options for all decades from 1900s to 2020s
Count of Entertainer by Decade of Breakthrough	A bar chart should display the count of entertainers for each decade, matching the values shown in the image (e.g. 16 for 1960s, 9 for 1970s).
Count of Entertainer by Gender (traditional)	A pie chart should show 75% Male (48) and 25% Female (16), matching the data in the image.
Count of Entertainer by Year of Breakthrough	A line graph should display the count of entertainers by year, matching the peaks and valleys shown in the bottom chart of the image.
Average Age Difference calculation	The dashboard should calculate and display the average difference between Age at Breakthrough and Age at First Award across all selected entertainers.