LOW LEVEL DESGIN (LLD)

Entertainer Data Analysis

Last date of revision – 20/07/22

Naveen S

Document Version Control

Date Issued	Version	Description	Author
12/07/22	1.0	Introduction, Problem statement	Naveen S
14/07/22	1.1	Data information, Architecture	Naveen S
15/07/22	1.2	Deployment	Naveen S
20/07/22	1.3	Final revision	Naveen S

Table of Contents

1. Introduction

1.1 What is Low-Level design document?

The goal of the LDD or Low-level design document (LLDD) is to give the internal logic design of the actual program code for the House Price Prediction dashboard. LDD 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 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.

Project Introduction:

Entertainment Industry plays a major role in our day-to-day life, from how we talk to how we dress and how we behave in our life is largely influenced by movies and series. No matter what, we all have at least one close to heart movie or star because of some specific reason or without any reason also. People never missed to celebrate good movies and stars, maybe it was celebrated bit late, but obviously it will be celebrated if the work is genuine. As we all know, it's a very competitive career filed and it takes a great effort to sustain the in the industry as a star or technician or producer. This analysis is aimed to analysis the entertainer's filmography, awards and compare film over the years in different aspects and get insights out of it.

Problem Statement:

Normal life can be stressful, and people need to relax. Being entertained by others is a wonderful way to take some time out of life. It can reduce stress and make life's issues easier to face. The media and entertainment industry consists of film, television, radio and print. These segments include movies, TV shows, radio shows, news, music, newspapers, magazines, and books. Entertainment industry is a group of sub-industries devoted to entertainment. Entertainment industry is used to describe the mass media companies that control the distribution and manufacture of mass media entertainment.

Dataset Description:

Provided Dataset:

1. **Entertainer-Basic Info**: It consists of list of **70** Entertainers Name, Birth year and Gender
2. **Entertainer-Breakthrough Info:** It consists of details about the 70 entertainers like breakthrough year, first major award, breakthrough movie name
3. **Entertainer-Last major work Info:** It consists of the details about the 70 entertainers last major work and if died, Year of death details.

Populated Dataset:

As the provided dataset is very minimal and insufficient for any analysis, I have populated dataset for the given entertainers using IMDBpy/ Cinemagoer library.

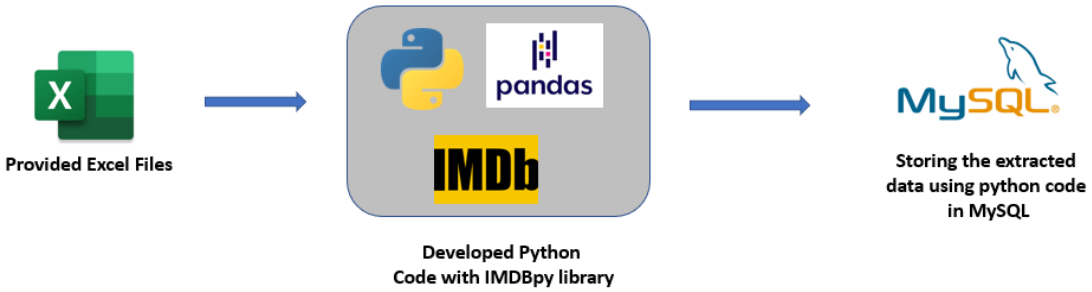
For populating the dataset, the code was developed. After populating the data using python code, extracted data were stored in MySQL local database. Database schema details as follow,

1. **Entertainers_basics_populated:**
    - Name: Entertainer's name
    - DOB: Date of birth of the Entertainer
    - Height: Height of the entertainer
    - Nicknames: Nicknames of the entertainers, that is used by fans or cinema industry
    - Quotes: Entertainers quote or statements they made in public
    - Mini-biography: Mini biography of the entertainer
    - Trademark: Trademark style or behaviour of the entertainer
    - Headshot: URL of the entertainers headshot
2. **Entertainers_film_list:**
    - Entertainer name: Name of the entertainer
    - Movie name: Movie name which the entertainer acted
    - Year: Release year of the movie
3. **Entertainers_awards:**
    - Name: Entertainer's name
    - Award: Award name

- Year: Year of the award
- Prize: Type of award
- Category: Under which category the award has been given
- Result: Whether entertainer won the award or just a Nominee (Winner/Nominee)
- Movie_name: For which movie the award was given
- Shared_with: with they shared the award with someone, there name

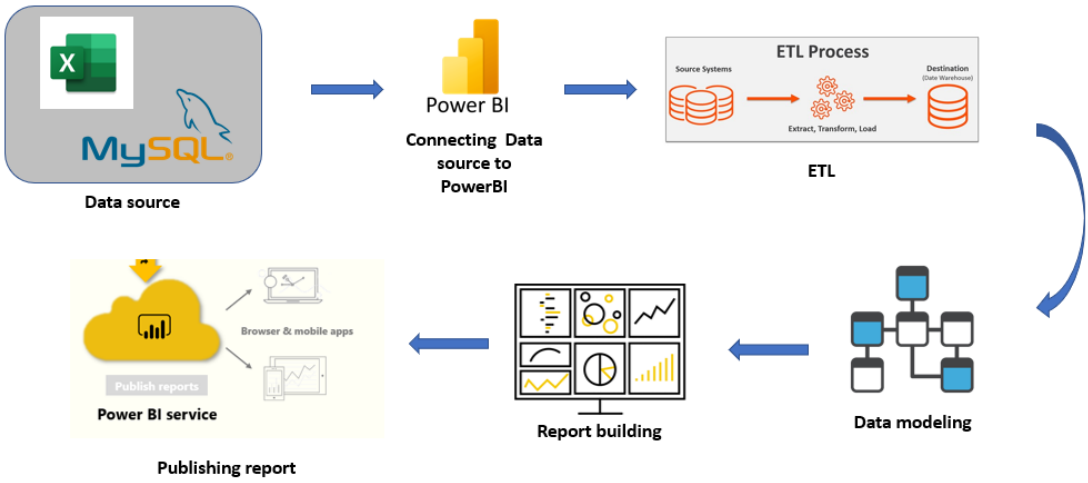
4. **Entertainer_salary:**
    - Name: Entertainer's name
    - Movie_name: Movie name of the salary they received
    - Year: Movie released year
    - Salary: Amount they received for that movie

Data populating flow



- Python code is developed by using the IMDBpy library in by following oops concepts.
- Using Pandas library the entertainers names from the excel/csv files is read and given as input to the code
- With the entertainers name as input, the code will generate all the needed data and store them in MySQL database in the mentioned table name

Architecture



Data sources:

- iNeuron provided dataset: 3 Excel files of the Entertainers details
- Populated data using python code which is stored in MySQL local database

Data connection:

Connecting all the data sources and importing the data into the PowerBi

ETL:

Extra, Load and Transform will be performed for the imported datasets in the PowerQuery editor.

- Data type change
- Duplicate Handling
- Date type
- Merge/ Append tables based on the need

Data 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.

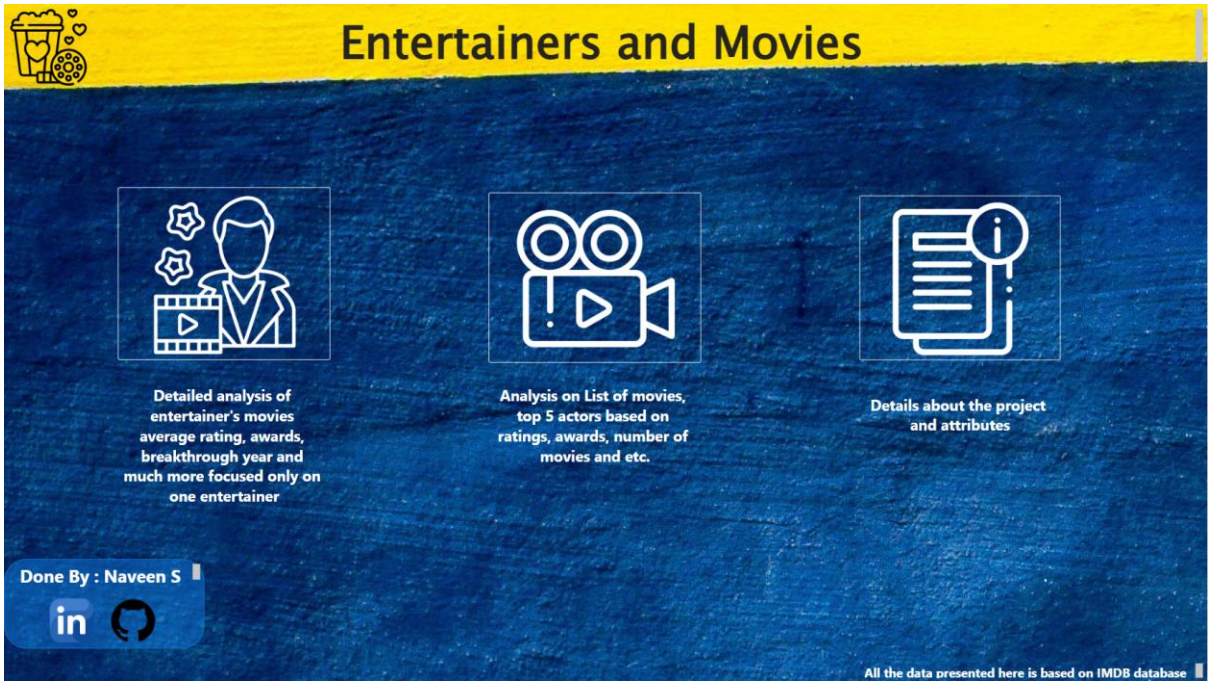
Report Building:

Report will be designed based on the mock ups which were created in the initial stage of the project. 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.

- High-Level Design Document (HLD)
- Low Level Design Document (LLD)
- Architecture
- Wireframe
- Detailed Project Report

Deployment

I have created a PowerBi Dashboard, which will be published in PowerBi service and can be shared with stakeholders and other users to consume.



LLD (Low Level Design)