HIGH LEVEL DESGIN (HLD)

Entertainer Data Analysis

Last Date of Revision – 4th Jul 2021

Naveen S

Document Version Control

Date Issued	Version	Description	Author
25th Jun 2021	1.0	Abstract, General Description	Naveen S
2nd Jul 2021	1.1	Design Details, KPI	Naveen S
3rd Jul 2021	1.2	Deployment	Naveen S
4th Jul 2021	1.3	Final revision	Naveen S

Abstract:

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.

1 Introduction

1.1 Why this High-Level Design Document?

The purpose of this High-Level Design (HLD) Document is to add the necessary detail to the current project description to represent a suitable model for coding. This document is also intended to help detect contradictions prior to coding, and can be used as a reference manual for how the modules interact at a high level.

The HLD will:

• Present all of the design aspects and define them in detail

• Describe the user interface being implemented

• Describe the hardware and software interfaces

• Describe the performance requirements

• Include design features and the architecture of the project

• List and describe the non-functional attributes like:

- o Security
- o Reliability
- o Maintainability
- o Portability
- o Reusability
- o Application compatibility
- o Resource utilization
- o Serviceability

1.2 Scope

The HLD documentation presents the structure of the system, such as the database architecture, application architecture (layers), application flow (Navigation), and technology architecture. The HLD uses non-technical to mildly-technical terms which should be understandable to the administrators of the system.

2 General Description

2.1 Product Perspective & Problem Statement

Housing prices are an important reflection of the economy, and housing price ranges are of great interest for both buyers and sellers. In this project, house prices will be predicted given explanatory variables that cover many aspects of residential houses.

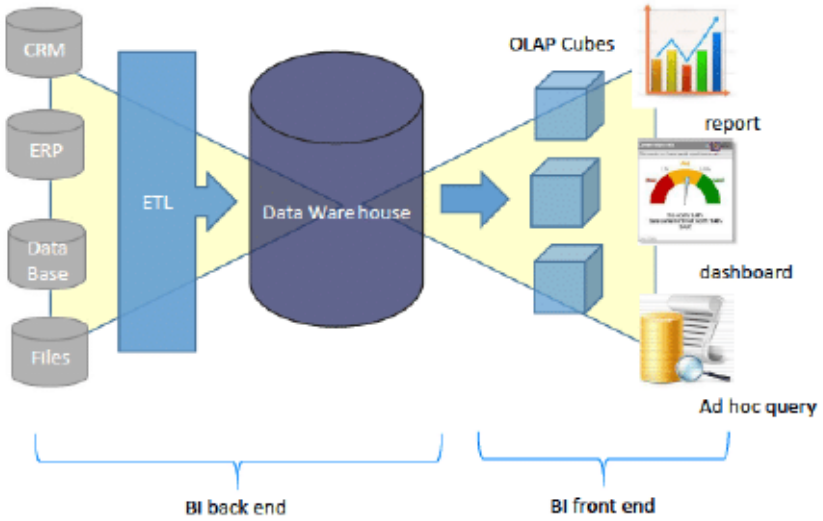
The objective of the project is to perform data visualization techniques to understand the insight of the data. This project aims apply various Business Intelligence tools such as Tableau or Power BI to get a visual understanding of the data.

2.2 Tools used

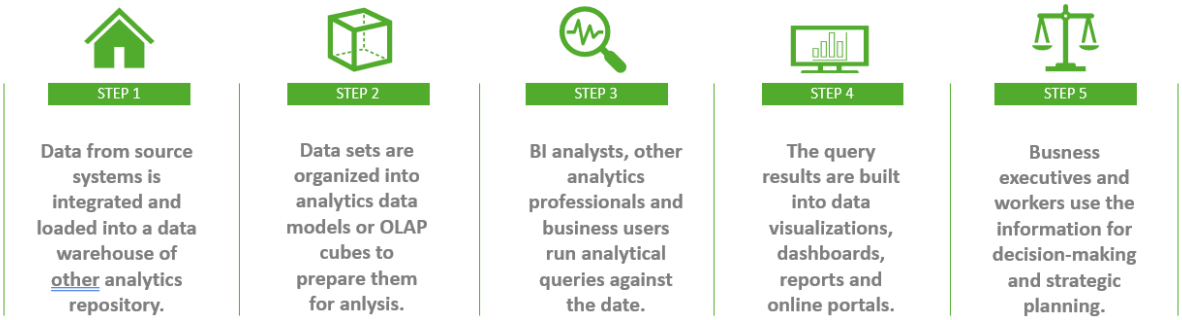
Business Intelligence tools and libraries works such as Python, MySQL, Pandas, Excel, Power BI are used to build the whole framework.

3 Design Details

3.1 Functional Architecture



How Bi works



3.2 Optimization

Your data strategy drives performance

• Minimize the number of fields

• Minimize the number of records

• Optimize extracts to speed up future queries by materializing calculations, removing columns and the use of accelerated views

Reduce the marks (data points) in your view

• Practice guided analytics. There's no need to fit everything you plan to show in a single view. Compile related views and connect them with action filters to travel from overview to highly-granular views at the speed of thought.

• Remove unneeded dimensions from the detail shelf.

• Explore. Try displaying your data in different types of views.

Limit your filters by number and type

• Reduce the number of filters in use. Excessive filters on a view will create a more complex query, which takes longer to return results. Double-check your filters and remove any that aren't necessary.

• Use an include filter. Exclude filters load the entire domain of a dimension, while include filters do not. An include filter runs much faster than an exclude filter, especially for dimensions with many members.

• Use a continuous date filter. Continuous date filters (relative and range-of-date filters) can take advantage of the indexing properties in your database and are faster than discrete date filters.

• Use Boolean or numeric filters. Computers process integers and Booleans (t/f) much faster than strings.

• Use parameters and action filters. These reduce the query load (and work across data sources).

Optimize and materialize your calculations

• Perform calculations in the database

• Reduce the number of nested calculations.

• Reduce the granularity of LOD or table calculations in the view. The more granular the calculation, the longer it takes.

o LODs - Look at the number of unique dimension members in the calculation.

o Table Calculations - the more marks in the view, the longer it will take to calculate.

• Where possible, use MIN or MAX instead of AVG. AVG requires more processing than MIN or MAX. Often rows will be duplicated and display the same result with MIN, MAX, or AVG.

• Make groups with calculations. Like include filters, calculated groups load only named members of the domain, whereas Tableau's group function loads the entire domain.

• Use Booleans or numeric calculations instead of string calculations. Computers can process integers and Booleans (t/f) much faster than strings.

Boolean>Int>Float>Date>DateTime>String

4 KPIs

Dashboards will be implemented to display and indicate certain KPIs and relevant indicators for the disease.

Power BI Desktop

As and when, the system starts to capture the historical/periodic data for a user, the dashboards will be included to display charts over time with progress on various indicators or factors

4.1 KPIs (Key Performance Indicators)

Key indicators displaying a summary of the Entertainer's filmography with different metrics

- Average Movie rating of the Entertainer
- Total number of movies acted by the Entertainer
- Total number of awards
- Breakthrough Movie year
- Break through Movie name
- First major award year
- Date of birth of the Entertainer
- Height of the Entertainer
- One of the famous Quote of the Entertainer
- Trademarks of the Entertainer
- Nicknames of the Entertainer

5 Deployment

Prioritizing data and analytics couldn't come at a better time. Your company, no matter what size, is already collecting data and most likely analyzing just a portion of it to solve business problems, gain competitive advantages, and drive enterprise transformation. With the explosive growth of enterprise data, database technologies, and the high demand for analytical skills, today's most effective IT organizations have shifted their focus to enabling self-service by deploying and operating PowerBi at scale, as well as organizing, orchestrating, and unifying disparate sources of data for business users and experts alike to author and consume content.

PowerBi prioritizes choice in flexibility to fit, rather than dictate, your enterprise architecture. PowerBi Desktop and PowerBi service leverage your existing technology investments and integrate into your IT infrastructure to provide a self-service, modern analytics platform for your users. With on-premises, cloud, and hosted options, there is a version of PowerBi to match your requirements.