High Level Design

BUDGET SALES ANALYSIS

MANISHA RAJ A.

Document Version Control

Data Issue	Version	Description	Author
	1.0	First version of Complete HLD	Manisha Raj A.

Table of Contents

Abstract

Analyzing data about entertainers helps us understand trends and patterns in the industry. It lets us look at what makes entertainers successful, their key moments of breakthrough, and how their careers develop over time. By examining this information, we can learn what factors contribute to an entertainer's success and how the industry changes.

This kind of analysis is also useful for discovering and nurturing new talent. By studying the traits and patterns of successful entertainers, industry professionals can spot emerging stars, offer guidance, and support their development. It also helps us understand what qualities lead to breakthrough achievements in an entertainer's career.

1. Introduction

1.1 Purpose of 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:

- Security
- Reliability
- Maintainability
- Portability
- Reusability
- Application compatibility
- Resource utilization
- 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 and Problem Statement.

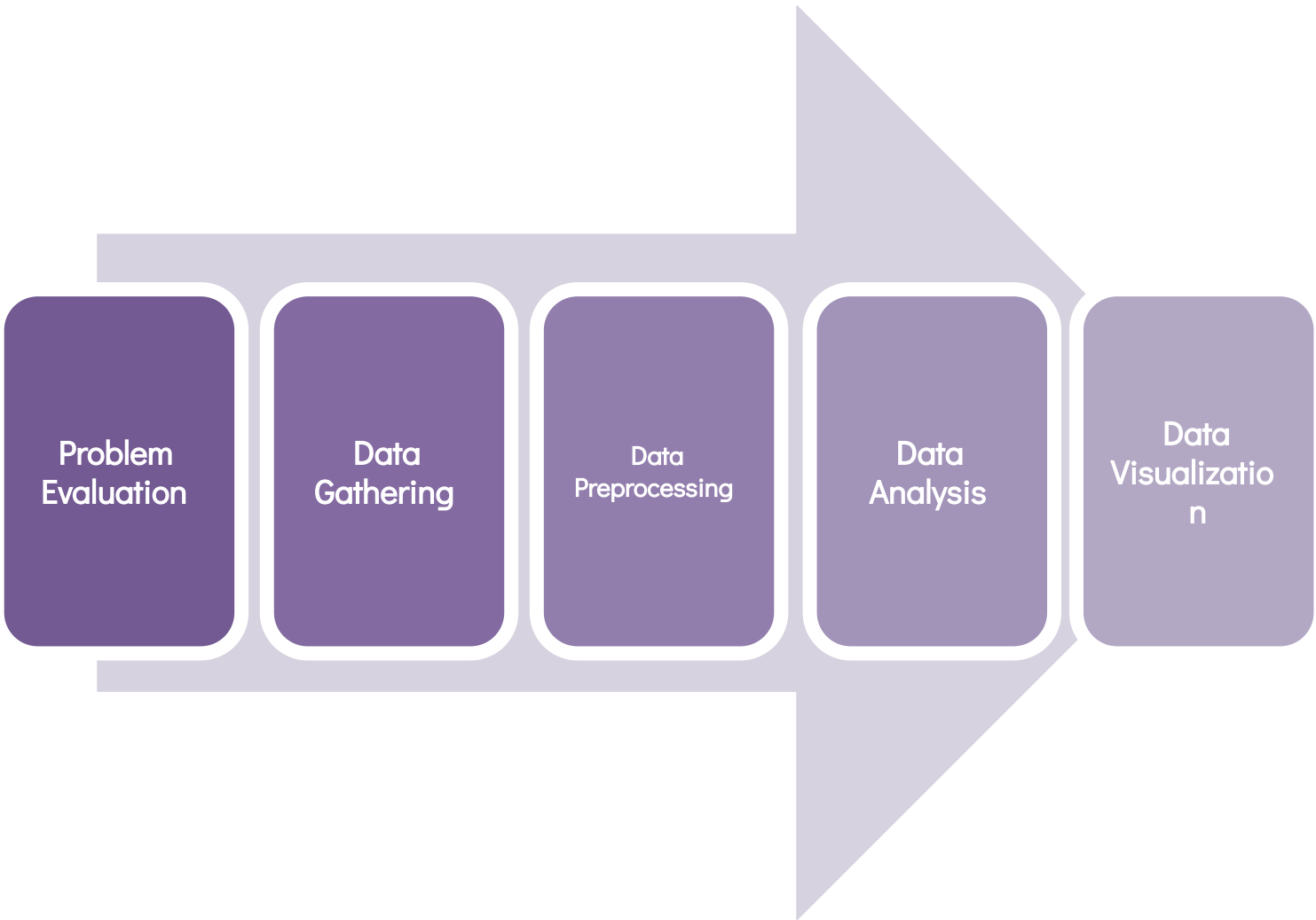
The dashboard is a central place where all important information about entertainers is collected. It combines details like names, genders, breakthrough years, awards, and professions into one convenient platform. This makes it easy for users to find and access all the relevant and current information about entertainers in an organized and straightforward way.

2.2 Tools Used

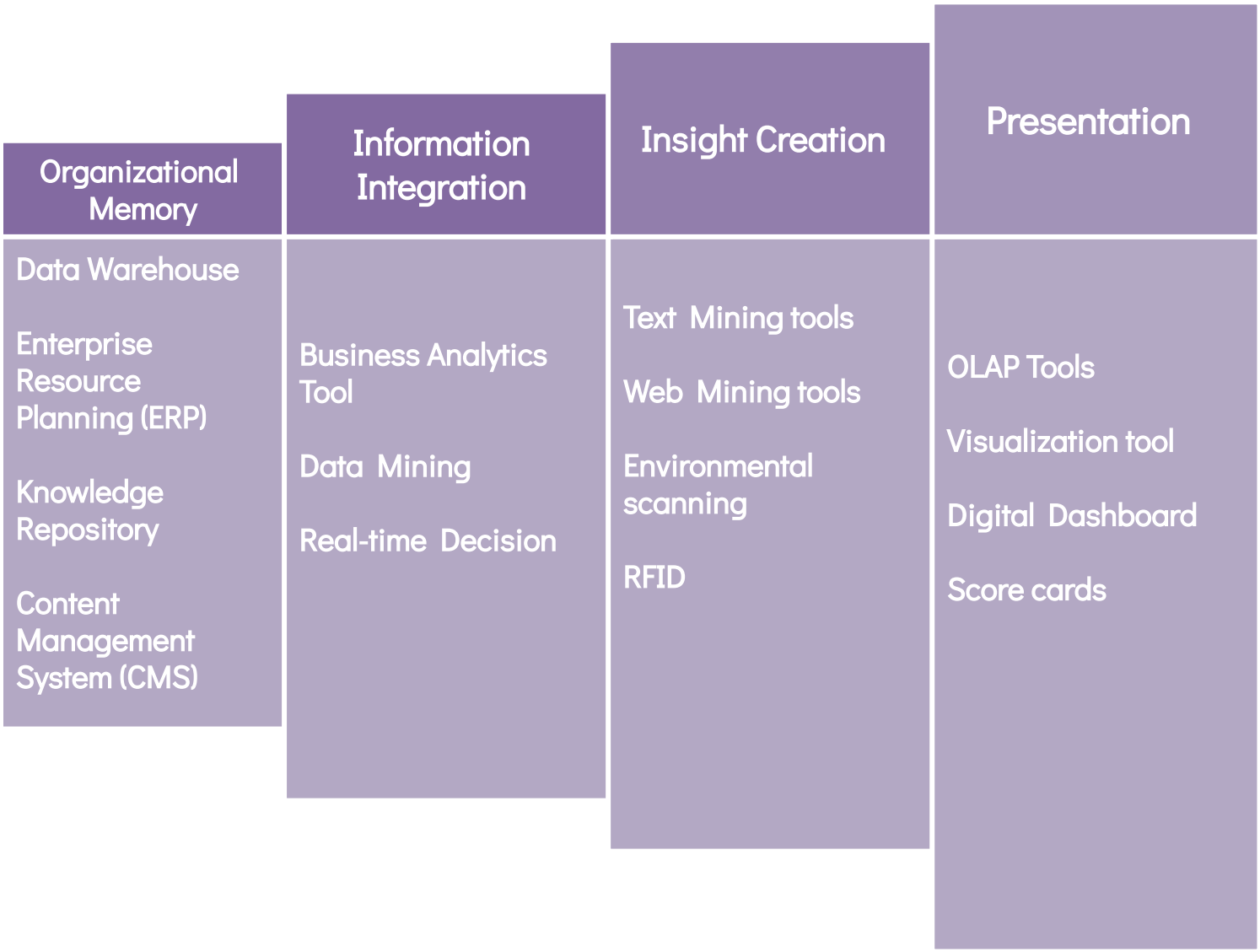
The framework is built using various Business Intelligence tools and libraries, including NumPy, Pandas, Seaborn, Matplotlib, MS Excel, MS Power BI, Jupyter Notebook, and the Python programming language.

3. Design Details.

3.1 Functional Architecture.



3.2 How BI Works



3.3 Optimization.

1. Data Processing Efficiency:

- Efficient Data Loading: Ensure that data is loaded efficiently to minimize delays. Consider using file formats like Parquet for faster read/write operations, especially with large datasets.

- Data Cleaning and Transformation: Reduce unnecessary data transformations by applying them only when required. For example, perform data type conversions and handle missing values efficiently at the loading stage.

- Data Structures: Use appropriate data structures for manipulation. In your code, leveraging Pandas DataFrames is appropriate, but ensure operations like `groupby` and `value_counts` are used efficiently to handle large datasets.

2. Query Performance:

- Indexing and Query Optimization: Although the code does not directly interact with a database, similar principles apply to data operations. For example, when filtering data, ensure that operations are vectorized and avoid iterative approaches.

- Query Caching: If you frequently run similar queries or calculations, consider caching intermediate results to avoid redundant computations.

3. Visualizations and Rendering:

- Rendering Speed: Optimize visualization rendering by reducing the complexity of plots where possible. For example, large histograms or scatter plots may benefit from downsampling or binning data to improve performance.

- Caching Mechanisms: Implement caching strategies for static visualizations to avoid re-computation. Libraries like Matplotlib offer ways to save plots as images to reduce rendering time for frequently accessed visualizations.

- Efficient Libraries: Use efficient visualization libraries or frameworks. While Matplotlib is widely used, consider libraries like Plotly or Seaborn for better performance and interactive features.

4. User Experience Optimization:

- Intuitive Layout and Design: Ensure that visualizations are clearly labeled and the layout is intuitive. Group related charts together and use consistent color schemes to enhance user understanding.

- Tooltips and Help Text: Provide tooltips or help text for interactive elements to guide users. This is especially useful in dashboards where users explore data dynamically.

- Logical Information Organization: Organize information logically to present insights clearly. For instance, group similar analyses (like award distributions and profession categories) together to facilitate user navigation and interpretation.

4. KPI.

4.1 KPIs (Key Performance Indicators:

1. User Engagement: Track how often users interact with the dashboard, including how many users are active, how long they spend on it, and how frequently they visit. This shows how interested and involved the audience is with the dashboard.

2. Data Completeness: Check how complete the entertainer information is in the dashboard. Look at the percentage of data available for things like achievements, awards, and nominations. This ensures the dashboard gives a full and accurate picture.

3. Data Accuracy: Monitor how accurate the data in the dashboard is by checking for inconsistencies, errors, or outdated info. Measure the percentage of accurate data to make sure the insights you get are reliable.

4. User Satisfaction: Find out how satisfied users are with the dashboard by using surveys, feedback ratings, or comments. This measures how users feel about the dashboard's usability and usefulness, helping to identify areas for improvement.

5. Impact on Decision-Making: Evaluate how the dashboard affects decision-making related to entertainers. Look at how often the dashboard's insights influence strategic decisions, industry trends, or research. This shows how valuable and effective the dashboard is for decision-making.

Key indicators displaying a summary of the sales generation and its relationship with different metrics

1. Total awards won
2. Total nominees
3. Oscar won
4. Emmy won
5. First Profession
6. First award won from breakthrough
7. First Entertainer
8. First Breakthrough Name
9. Sum of Total Award won by Gender (tradition)
10. Top 10 Entertainer
11. Sum of Oscar won bt Entertainer

5. Deployment

- Prioritizing data and analytics is more important than ever. Regardless of your company's size, you're already gathering data and likely only analyzing a small portion of it to address business issues, gain a competitive edge, and drive overall transformation.

- With the rapid increase in enterprise data and database technologies, and the high demand for analytical skills, today's most successful IT organizations are focusing on enabling self-service. They do this by using Power BI on a large scale and by organizing and integrating various data sources. This allows both business users and experts to easily create and access useful content.

- Power BI Desktop and Power BI Service make the most of your existing technology investments and integrate them into your IT infrastructure. They offer a modern, self-service analytics platform that can be tailored to your needs, with options available for on-premises, cloud, and hosted environments.