**High Level Design**

**AIR TRAVEL DATA ANALYSIS**

|  |  |
| --- | --- |
| **Written By** | **Nitish Pandey** |
| **Document Version** | **1.0** |
| **First Revised Date** | **01/07/2024** |
|  |  |
|  |  |

**DOCUMENT CONTROL**

**Change Record:**

|  |  |  |  |
| --- | --- | --- | --- |
| **VERSION** | **DATE** | **AUTHOR** | **COMMENTS** |
| 1.0 | 01 JULY 2024 | Nitish Pandey |  |
|  |  |  |  |

**Contents**

**Document Version Control......................................................................................................... 2**

**Abstract ...................................................................................................................................... 4**

**1 Introduction ............................................................................................................................ 5**

**1.1 Why this High-Level Design Document? ......................................................................... 5**

**1.2 Scope ................................................................................................................................ 6**

**2 General Description ................................................................................................................ 6**

**2.1 Product Perspective & Problem Statement ..................................................................... 6**

**2.2 Tools used ......................................................................................................................... 7**

**3 Design Details ......................................................................................................................... 8**

**3.1 Functional Architecture .................................................................................................... 8**

**3.2 Optimization ...................................................................................................................... 9**

**4 KPIs ......................................................................................................................................... 10**

**5 Deployment ........................................................................................................................... 10**

**Abstract**

This internship project focused on airline data analysis at PW (PhysicsWallah). The project encompassed rigorous data cleaning, transformation, and modeling processes, culminating in the generation of actionable insights.

Key activities included:

* **Data Cleaning**: Employed Python to meticulously clean and refine the dataset, ensuring accuracy by eliminating inconsistencies and errors.
* **Data Transformation**: Applied diverse transformation techniques to prepare the data for comprehensive analysis, including converting monthly data into yearly and quarterly formats.
* **Data Modeling**: Leveraged Power BI to create dynamic and interactive visualizations, facilitating deeper insights and informed decision-making.

This project has provided invaluable learning experiences in navigating complex datasets and deriving actionable insights in the aviation sector. These findings are pivotal for future applications in enhancing operational efficiency, strategic planning, and decision-making within the aviation industry.

**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**

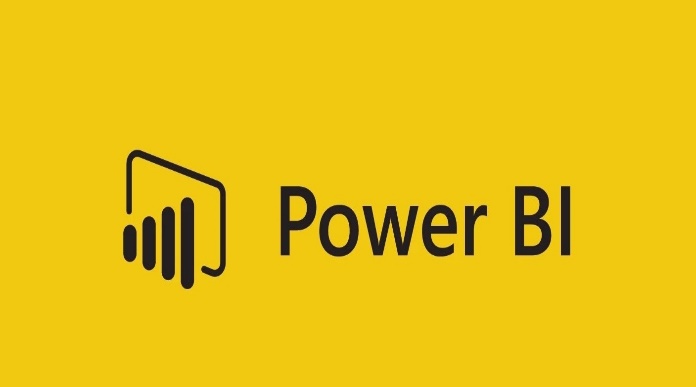
**2.1 Product Perspective & Problem Statement**

A banking domain is comprised of all the components needed to run a financial service end-to-end It covers the transaction and distribution process; the ways in which customers interact with the system, products, and services the organization offers; and the technology involved.

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 and helps in getting the clear insights from the data.

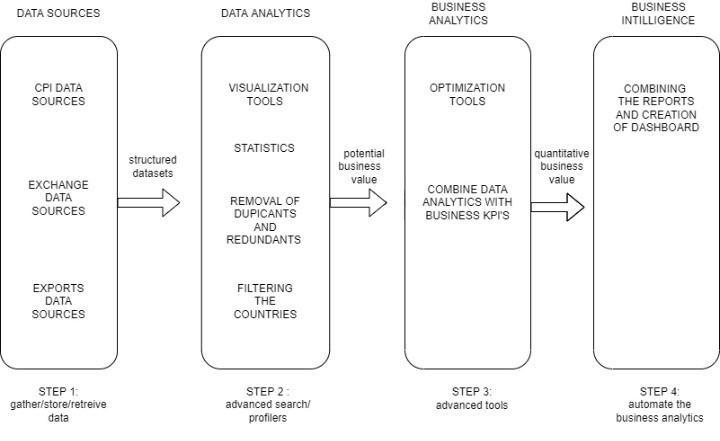
**2.2 Tools used**

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

**3 Design Details**

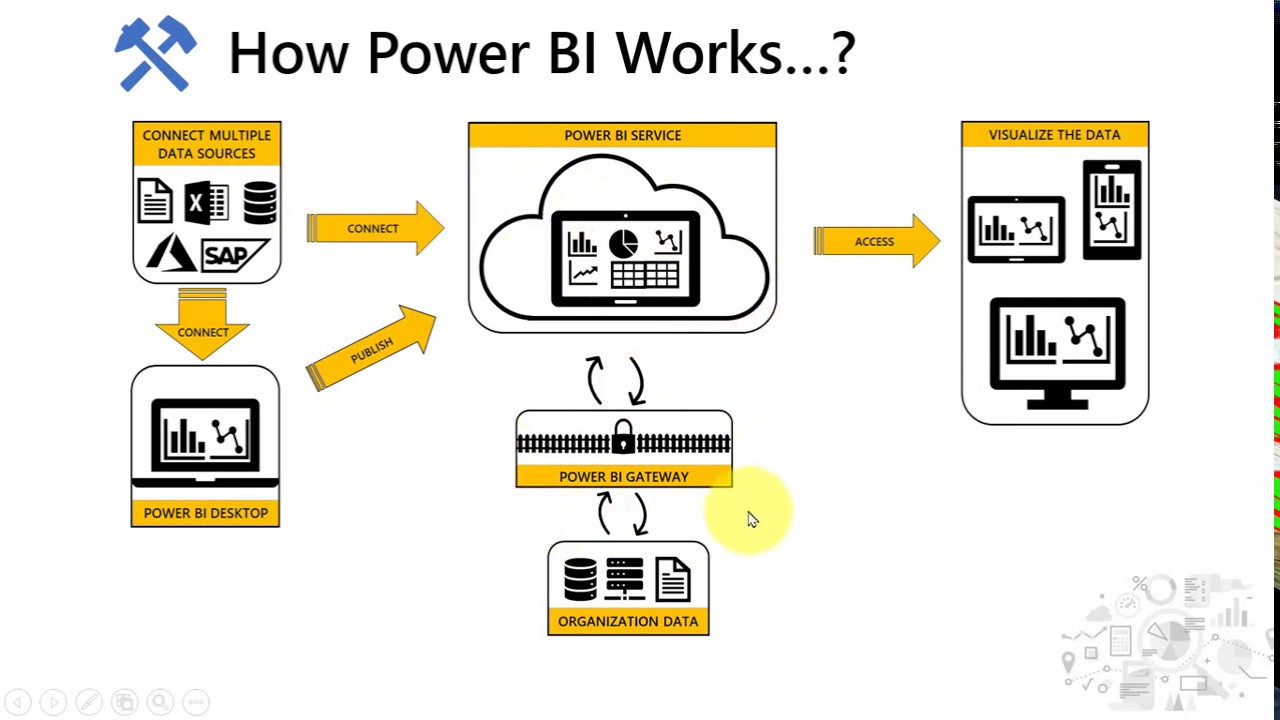
**3.1 Functional Architecture**



Airline

AIR TRAVEL

DATA SOURCE



**3.2 Optimization**

 **Removal of NaN Values**: Cleaned the dataset by identifying and removing any missing (NaN) values to enhance the accuracy of the analysis.

 **Data Conversion**: Transformed the data to different granularities, converting monthly data into yearly and quarterly data to facilitate various levels of analysis.

 **Country Filtering**: Applied filters to focus on specific countries, allowing for targeted analysis based on geographical regions.

 **Data Reshaping**: Reshaped the data from a wide format to a long format, which is more suitable for detailed analysis and visualization in Power BI.

**4 KPIs**

These KPIs can help in assessing various aspects of airline operations and marketing effectiveness, contributing to more informed decision-making in the aviation industry.

 **Busiest Route**:

* KPI: Number of flights between Los Angeles (LAX) and San Francisco (SFO).
* Metric: 2,220 departing flights from LAX and 2,338 arriving flights at SFO.

 **Top Marketing Airline**:

* KPI: Number of flights advertised/promoted by Air New Zealand.
* Metric: 4,500 flights.

 **Top Operating Airline**:

* KPI: Number of flights operated by United Airlines.
* Metric: 26,140 flights.

 **Longest Route**:

* KPI: Distance of the longest route.
* Metric: 13,583.42 kilometers between San Francisco and Singapore.

 **Traffic Volume**:

* KPI: Average number of departing flights on the busiest day.
* Metric: 7,900 flights on Sundays.

 **Punctuality**:

* KPI: On-time arrival rate of New Zealand Airline.
* Metric: 91.035% on-time arrival rate.

**5 Deployment**

 **Select Specific Routes**: Users can choose any route from the available data.

 **Filter by Date Range**: Users can filter the data by specific dates, ranging from the entire dataset timeframe.

 **Analyze Key Metrics**: Once a route and date range are selected, users can easily analyze various key

performance indicators such as the number of flights, on-time performance, and traffic volume.

 **Access Operational Insights**: Users can obtain detailed information on the operational performance of

different airlines, including the total number of flights operated, on-time arrival rates, and average delay

times.

 **Marketing Insights**: Users can explore marketing insights, such as the most advertised/promoted airlines

and their corresponding flight numbers.

 **Longest Routes and Traffic Volume**: Users can identify the longest routes and analyze traffic volume trends particularly on high-traffic days such as Sundays.

**6 Scope**

The scope of the airline data analysis project is to enhance the capabilities of the dataset by introducing new measure values. These enhancements will enable the following:

* **Predicting Airline Ranks**: Adding new measures will allow for the prediction and ranking of different airlines based on operational performance, punctuality, and customer satisfaction.
* **Investment Analysis**: Users will be able to assess investment levels in different airlines and routes based on their rankings, providing insights into profitable routes and high-performing airlines.
* **Operational Efficiency Analysis**: By incorporating measures that identify factors affecting operational efficiency, users can gain a deeper understanding of delays, cancellations, and other performance metrics, enabling them to take informed actions to improve efficiency.
* **Enhanced Data Visualization**: The addition of new measures will enable users to analyze the data in more visually appealing and comprehensive ways, improving the overall user experience and insight generation.
* **Passenger Traffic Trends**: Users can analyze trends in passenger traffic volume, including peak travel times and popular routes, to optimize scheduling and resource allocation.
* **Route Profitability**: Measures can be added to evaluate the profitability of different routes, helping airlines make strategic decisions about route planning and expansion.