Technical Documentation

Table of Contents

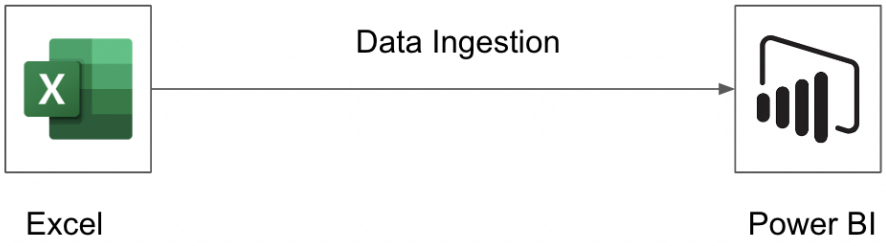
1. Architecture

High-Level Architecture

Due to the nature of this project, the architecture consists of the file excel file data source and the power bi desktop. There was no requirement for a live report on Power BI Service, hence the submission was as required in .pbix format. Consequently, there was no need for a data refresh via the Power BI Gateway.

Data Flow

The 23kb data provided is an Excel file based data stored on the local machine. Hence, the data flow consists of a manual data loading into Power BI using the excel file loading option.



2. Configuration

Data Sources

The data source consists of the Case Study Data Set.xlsx file loaded inot Power BI Desktop from the local storage.

Power BI Desktop Configuration

No additional configuration was needed for data ingestion and egestion since the whole data was provided in a file, and there was no requirement for an upload to the Power BI Service.

Data Transformation

This stage involves several transformation layers done to ensure the data is in the ready-mode for descriptive and predictive analysis as required in the specifications document. The processes included involve the following:

1. **Checks for Anomalies, Emptiness and Duplicates:** The dataset maintaines a perfect completeness and it's bereft of duplicates. Hence, no extra step was involved in filling any empty space, and no row of data was lost from data de-duplication exercise.

2. **Separation of Concerns:** The dataset consists of historicals of Budgets and project schedules. Hence, the data set was divided into Budget related tables (Budget Actuals and Budget Forecast), and the Scheduling related table (Case Study Data). Both Budget Actuals and Budget Forecast were referenced out of the 'Case Study Data' table, to ensure that future updates to the main table 'Case Study Data' is reflected in the two derived tables.

3. **Transposing of Tables:** The monthly reporting of the budget data requires a horizontally transposed table. Hence, both budget tables were transposed.

4. **Calendar Tables and Data Columns:** For proper attribution of budgets and project timelines, extra steps of Data variable columns were appended to the budget tables, and an extra Calendar table was introduced for the calculations of working days duration vs. non working days duration, and the project timelines.

5. **Column Renamings and Datatype Settings:** An additional step of renaming the columns after the table transpose, and setting of the appropriate data types, and using the currency metrics was also included to put the data in the best format for analysis.

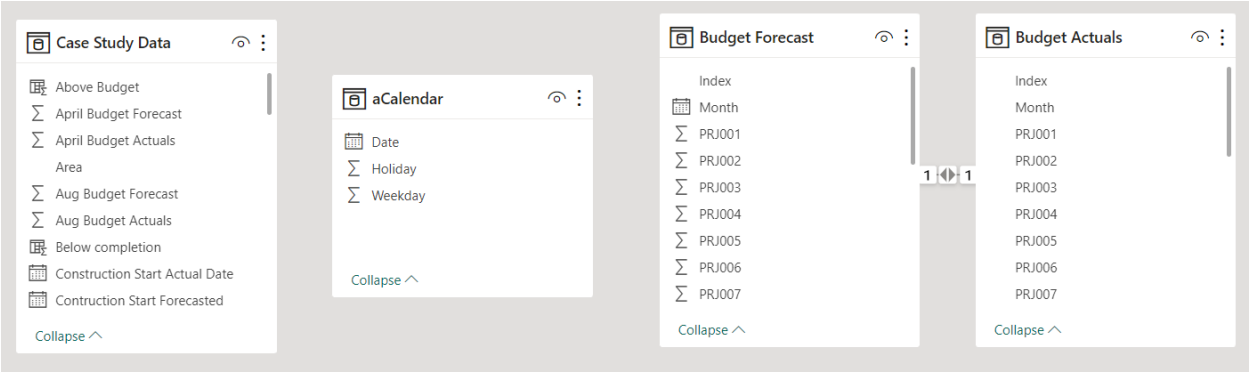
Advanced Analytics Configuration

For future budget predictions and to understand trends and seasonality, advanced analytics models such exponential smoothing and ETS-AAA Power BI forecasting algorithm was activated for future projections, understanding of trends and calculations of future budgets within Power BI to enable forecasting, trend analysis, and risk identification.

3. Setup

Power BI Data Model Design

The required table relationship was a one-to-one relationship between the Budget and the Actuals tables using the index as the primary key. No other table relationship was required.



Beside having the latest version of Power Bi installed on a windows machine, no other settings is required to open and interact with the dashboards.