

Power BI Project - Data Analysis and Visualization

The goal of this project is to assess your ability to clean, transform, analyze, and visualize data using **Power BI**. You will demonstrate your understanding of Power BI's key features, such as **Power Query Editor**, **DAX functions**, and **Dashboard creation**, to provide meaningful insights into the data. You are required to select one dataset from the list provided, clean and transform the data, apply DAX functions, and create a dynamic dashboard to visualize your findings.

Key Deliverables:

1. **Dataset Selection:** You need to choose one dataset from the list provided below. The datasets come from Kaggle and cover a variety of domains, such as sales, healthcare, sports, and more.

Available Datasets from Kaggle (Select One):

1. Sales Data – Online shop

- **Link:** <https://www.kaggle.com/datasets/gabrielramos87/an-online-shop-business>
- **Description:** This is a sales transaction data set of UK-based e-commerce (online retail) for one year. This London-based shop has been selling gifts and homewares for adults and children through the website since 2007.

2. Heart Disease Dataset

- **Link:** <https://www.kaggle.com/datasets/johnsmith88/heart-disease-dataset>
- **Description:** Data related to patients' medical information, which could be used to predict the likelihood of heart disease.

3. Global Health Data Analysis

- **Link:** <https://www.kaggle.com/datasets/kamaumunyori/global-health-data-analysis-1990-2019>
- **Description:** The "Global Health Data Analysis 1990-2019" dataset provides comprehensive health-related statistics across various countries, covering indicators such as life expectancy, disease prevalence, and healthcare expenditure over nearly three decades.

4. Car Price Prediction

- **Link:** <https://www.kaggle.com/datasets/zafarali27/car-price-prediction>
- **Description:** Contains car attributes such as make, model, year, mileage, etc., to predict car prices.

5. Sales Forecasting Data

- **Link:** <https://www.kaggle.com/datasets/tevecsystems/retail-sales-forecasting>
- **Description:** Sales forecasting dataset to predict future sales performance for retail businesses.

6. Employee Attrition Dataset

- **Link:** <https://www.kaggle.com/datasets/colearninglounge/employee-attrition>
- **Description:** Employee attrition dataset including employee details, job satisfaction, attrition status, and performance data.

7. Movie Ratings Dataset

- **Link:** <https://www.kaggle.com/datasets/asaniczka/tmdb-movies-dataset-2023-930k-movies>
- **Description:** A dataset containing movie ratings, genres, cast, crew, and other related data, suitable for sentiment analysis or recommendation systems.

8. COVID-19 Dataset

- **Link:** <https://www.kaggle.com/datasets/anandhuh/latest-covid19-india-statewise-data>
- **Description:** A dataset containing COVID-19 case details in India state wise.

9. Retail Store Transactions

- **Link:** <https://www.kaggle.com/datasets/marian447/retail-store-sales-transactions>
- **Description:** Dataset with retail store transactions data, including product purchases, customer details, and sales amounts.

10. World Happiness Report

- **Link:** <https://www.kaggle.com/datasets/utkarshsen/happiness-dataset>
- **Description:** Data on happiness rankings across countries, with factors such as GDP per capita, life expectancy, social support, and more.

2. Data Cleaning & Transformation:

- **Using Power Query Editor**, clean the selected dataset by:
 - Handling missing values.
 - Removing duplicates or irrelevant data.
 - Standardizing data types and formats.
 - Performing necessary transformations (e.g., splitting columns, merging data, etc.).
- Ensure the dataset is prepared and structured for analysis.

3. DAX Functions:

- Apply **at least 10 DAX functions** to the dataset to calculate various metrics, ratios, or aggregates. Examples of functions you might use:
 - SUM(), AVERAGE(), COUNT(), DISTINCTCOUNT(), IF(), CALCULATE(), DATEADD(), RELATED(), ALL(), and others as necessary.

4. Dashboard Creation:

- Create a **Power BI Dashboard** to visualize the results. Your dashboard should include:

- Multiple types of visualizations (e.g., bar charts, line graphs, pie charts, maps, etc.).
- Filters and slicers for interactivity.
- A clear and meaningful layout that presents your findings.

5. Insights & Analysis Report:

- Write a report summarizing the key insights and findings from your analysis. This should include:
 - A brief description of the dataset and the project's objective.
 - Key observations and trends discovered from the data.
 - Any significant patterns, anomalies, or relationships you identified.
 - Recommendations or actions based on your analysis.

6. PowerPoint Presentation:

- Create a PowerPoint presentation summarizing your work, including:
 - **Title Slide:** Project title, your name, and a brief introduction to your selected dataset.
 - **Dataset Overview:** Description of the dataset and the key attributes.
 - **Data Cleaning Process:** Brief summary of the cleaning and transformation steps you performed.
 - **DAX Functions:** Overview of the DAX functions applied and their impact on the analysis.
 - **Dashboard Overview:** Screenshot or overview of the dashboard with highlights of key insights.
 - **Key Insights & Recommendations:** Summary of the analysis and actionable insights.
 - **Conclusion:** Wrap-up with the overall findings and real-world applications.

Evaluation Rubric for Power BI project:

Criteria	Marks
Dataset Selection (Relevance and Appropriateness)	10
Data Cleaning & Transformation (Use of Power Query Editor)	20
DAX Functions (Application of at least 10 functions)	20
Dashboard Design (Variety, Aesthetic, Interactivity)	15
Insights & Analysis Report (Depth and Quality of Insights)	15
PowerPoint Presentation (Clarity and Structure)	20
Total	100