# Data Storage

# Data Dynamos

## (GitHub & Google Drive & Application)

For HR Data Analysis

- Project: Human Resources Dataset Analysis
- Team: Data-Dyanamos
- Github: github.com/0PeterAdel/Data-Dyanamos

## 1. Introduction

- **This document provides comprehensive documentation for the "Data-Dyanamos" project, which focuses on analyzing Human Resources datasets. It outlines the repository structure, explains the purpose of each file and folder, and describes each stage of the data analysis process—from data collection to final data mining and reporting.**

## 2. Repository Structure

```
Data-Dyanamos/
├── Project-Operations/          # Contains all steps of the data analysis project
│   ├── 01.Data-Collection/      # Scripts for gathering data from various sources
│   ├── 02.Data-Wrangling/       # Code for transforming and organizing data
│   ├── 03.Data-Cleaning/        # Procedures for cleaning and preparing data for analysis
│   ├── 04.Data-Exploration-&-Transformation/ # Exploratory data analysis and transformation techniques
│   ├── 05.Data-Modeling/        # Model development and training scripts
│   ├── 06.Data-Analysis/        # Analysis results and insights
│   ├── .....
│   └── 09.Data-Mining/          # Data mining techniques and dashboard reports
├── Instructions/                # Instructions for running the project and using the code
├── reports/                     # Final report and presentations
├── DataSet/                     # Original data
├── README.md                    # Project documentation
├── requirements.txt             # Dependencies
└── LICENSE                      # License information for the project
```

Repository link: github.com/0PeterAdel/Data-Dyanamos

## 3. Detailed Explanation of Folders and Files

A. Project-Operations/

This folder contains all the phases of the project, each corresponding to a step in the data analysis process:

1. 01.Data-Collection/
   - Purpose: *Gather data from various sources.*
   - Contents:

     EducationLevel.csv: *Contains education level data.*
     Employee.csv: *Contains employee data.*
     PerformanceRating.csv: *Contains performance rating data.*
     RatingLevel.csv: *Contains rating level data.*
     SatisfiedLevel.csv: *Contains satisfaction level data.*

2. 02.Data-Wrangling/
   - Purpose: Transform and organize the data to make it ready for further processing.
   - Contents:

     Data-Wrangling.pdf: Documentation that explains the steps and methods used in data transformation.

3. 03.Data-Cleaning/
   - Purpose: Clean and prepare the data for analysis.
   - Contents:

     Data-Cleaned/
     This folder contains the cleaned data, organized by category:
     Employee/
     1.png, 2.png: Screenshots or charts illustrating the cleaning process.
     Employee.sql: SQL scripts related to cleaning or extracting employee data.
     PerformanceRating/
     1.png, 2.png, 3.png: Screenshots or charts illustrating the cleaning process.
     PerformanceRating.sql: SQL scripts for cleaning or extracting performance rating data.

     Cleaned CSV files: (EducationLevel.csv, Employee.csv, PerformanceRating.csv, RatingLevel.csv, SatisfiedLevel.csv).
     Excel Files:
     HR-Data.xlsx: The main cleaned data file.
     EducationLevel.xlsx, Employee.xlsx, PerformanceRating.xlsx, RatingLevel.xlsx, SatisfiedLevel.xlsx: Individual Excel files for each data category.

4. 04.Data-Exploration-&-Transformation/
   - Purpose: Perform exploratory data analysis and apply data transformation techniques.
   - Contents:

     BQ-Team/

BQ-1.pdf, BQ-2.pdf: Documents containing business questions and analysis details.
Business-Questions.pdf: Document outlining the key business questions.
Business-Questions&KPI.pdf: Explains the business questions and key performance indicators (KPIs).
HR-Domain-Guide.pdf: A guide that explains the HR domain and key considerations during the analysis.

5. 05.Data-Modeling/
   - ❖ Purpose: Develop and train models based on the HR data.
   - ❖ Files:
     HR-Data.xlsx: The HR dataset used for modeling.
     Data-Modeling.png, Data-Modeling-2.png: Diagrams illustrating the data modeling process.

6. 06.Data-Analysis/
   - ❖ Purpose: Analyze the data to extract insights and results.
   - ❖ Contents:
     Analysis-Report.pdf: The final analysis report.
     Documentation/: Additional documentation related to the analysis.
     analysis-part1.ipynb and analysis-part2.ipynb: Jupyter Notebooks detailing the analysis process.
     Data.xlsx: Excel file containing the analysis results.
     readme.md: Explains the structure and contents of the Data Analysis folder.

7. 09.Data-Mining/
   - ❖ Purpose: Apply data mining techniques to extract patterns and insights.
   - ❖ Contents:
     Dashboard/
     Dashboard.pdf: Document describing the dashboard.
     Dashboard.pptx: PowerPoint presentation of the dashboard.
     Data-Mining.pdf: Document outlining the data mining process.
     Data-Mining.xlsx: Excel file with data mining results.

B. Other Folders and Files

1. Instructions/
   - ❖ Purpose: Contains detailed instructions for running the project and using the code.

2. reports/
   - ❖ Purpose: Stores the final project reports and presentation files.

3. DataSet/

- Purpose: Contains the original data files before any processing.

- Files:
  HR-Data.xlsx: The main dataset.
  original.zip: A compressed archive containing all the original data.

4. README.md

- Purpose: Provides a general overview of the project, usage instructions, and other relevant details.

5. requirements.txt

- Purpose: Lists all the libraries and dependencies required to run the project.

6. LICENSE

- Purpose: Contains the license information governing the use and distribution of the project.

## 4. Reference Links

- Main Repository Link: github.com/0PeterAdel/Data-Dyanamos
- Drive Link: ....

## 5. Conclusion

This document provides a comprehensive overview of the Data-Dyanamos project structure, detailing every stage of the data analysis process—from data collection to data mining—and explaining the role of each file and folder. It serves as an internal reference for team members and as documentation for external users. This document can be updated as the project evolves or when new files and folders are added.

**Data Dynamos**