**Project Introduction**

Welcome to the Data Analytics & Visualization Project! In this project, you will embark on an exciting journey to explore, clean, transform, model, and visualize data from various sources. By the end of this project, you will have gained hands-on experience in data manipulation, database management, and creating insightful visualizations using Power BI.

**Project Overview:**

This project aims to provide you with a comprehensive understanding of the end-to-end data workflow, from data acquisition to visualization. You will work with multiple datasets, including CSV, JSON, and Excel files, covering inventory, sales, customer information, product details, store information, and employee data.

**Key Steps in the Project:**

1. **Data Acquisition:** You will start by reading data from CSV, JSON, and Excel datasets using Python. This step involves loading the datasets into your environment for further analysis.
2. **Exploratory Data Analysis (EDA):** Dive into the datasets to identify key features and gain insights. Explore metadata information to understand the structure and characteristics of the data.
3. **Data Cleaning:** Address missing values, errors, and inconsistencies in the data. Convert data types as needed to ensure data quality and consistency.
4. **Data Transformation & Feature Engineering:** Reshape the data, create new columns, perform aggregations, replace values, and denormalize tables to prepare the data for analysis.
5. **Data Modeling & Documentation:** Design a data model, including entity-relationship diagrams (ERD) and business rules, to represent the relationships between different datasets and entities.
6. **Database Management:** Connect to a MySQL database using Python and create tables using Data Definition Language (DDL) statements. Insert data from the prepared datasets into the MySQL database tables.
7. **Verification:** Verify the tables and data in the MySQL Workbench to ensure the successful creation and population of tables.
8. **Connect to Power BI:** Connect the Power BI Desktop to the MySQL database and import the datasets. Recreate the data model within Power BI to reflect the relationships established in the database.
9. **Data Visualization:** Build insightful reports using Power BI's visualization capabilities. Utilize DAX measures to perform calculations and create meaningful insights.
10. **Deployment:** Load the report to Analysis Services for further analysis and sharing with stakeholders.

**Dataset Information:**

The project includes various datasets such as food inventory, sales by store, and dimensions like customers, products, stores, and employees. These datasets will provide rich insights into business operations and help drive informed decision-making.

**Best Practices:**

Throughout the project, adhere to best practices in data analytics & visualization, including:

* Documenting your process and decisions at each stage. (ERD, UML Diagrams)
* Maintaining data integrity and quality through rigorous cleaning and validation (Python Data Preparation).
* Optimizing database design for efficient querying and storage (MySQL DB).
* Creating clear and visually appealing reports for effective communication of insights (Power BI & PBI Analysis Services).
* Collaborating effectively with team members and stakeholders to gather requirements and iterate on solutions (GIT & Version Control).

This project will equip you with valuable data engineering, database management, and visualization skills, empowering you to tackle real-world data challenges effectively. Get ready to unleash the power of data and drive actionable insights!

Top of Form