**Use case: Sales Data Processing for Retail Optimization**

**Problem Statement**: XYZ Retail Inc., a prominent player in the retail industry, faces the challenge of efficiently managing and analysing its sales data. The company operates through multiple channels like online platforms, physical stores, and mobile apps. To enhance decision-making, XYZ Retail aims to create an end-to-end data pipeline that ingests, processes, and visualizes sales data.

**Objective:** Optimize retail operations by designing scalable data pipelines, integrating and cleaning diverse data sources, and enabling advanced analytics through robust storage, transformation, and governance frameworks.

**Tasks to perform:**

1. **Data Ingestion**:
   * Ingest sales data from a CSV file into Azure Data Lake Storage or Azure SQL Database. 
   * Ensure data quality and handle any data anomalies during ingestion.
2. **Data Transformation**:
   * Transform the raw sales data to create meaningful insights.
   * Calculate total sales, average order value, and other relevant metrics.
   * Enrich the dataset with additional fields (e.g., customer names, phone numbers).
3. **Power BI Dashboard**:
   * Create a Power BI dashboard to visualize sales trends, product performance, and customer insights.
   * Include interactive visuals, filters, and KPIs.

**Guidelines for Submission:**

1. **Design Document**:
   * Provide a high-level design document describing your proposed solution.
   * Explain the choice of data storage (ADLS or Azure SQL Database) and justify your decision.
   * Include an architectural diagram illustrating the data flow.
2. **Code Implementation**:
   * Write code/scripts to implement the data pipeline.
   * Document your code clearly with comments and explanations.
   * Handle PII data (e.g., credit card numbers) securely.
3. **Dataset Documentation**:
   * Describe the structure of the sales dataset (columns, data types).
   * Explain any data preprocessing steps performed.
4. **Power BI Dashboard**:
   * Create a Power BI report with relevant visuals.
   * Organize the dashboard logically (e.g., separate pages for different insights).
   * Add slicers and filters for user interactivity.

**Evaluation Criteria:** Your work will be evaluated based on the following criteria:

1. **Functionality**:
   * Does the solution meet the problem statement requirements?
   * Is the data pipeline functioning correctly?
2. **Scalability and Performance**:
   * Consider the scalability of your chosen storage solution.
   * Evaluate the efficiency of data transformation processes.
3. **Documentation Quality**:
   * Is the design document clear and comprehensive?
   * Are code comments and dataset documentation sufficient?
4. **Creativity and Attention to Detail**:
   * Did you explore additional features beyond the basic requirements?
   * Is the Power BI dashboard visually appealing and user-friendly?

Remember, this project is an opportunity to showcase your skills, creativity, and problem-solving abilities. Candidates are expected to showcase their analytical, logical thinking, and deduction skills throughout this challenge. Your approach to the problem, as well as your ability to present and defend your findings, will be central to the evaluation process.

Please ensure the work is your original work. Any kind of plagiarism or taking help from others will have adverse consequences.

**Best of Luck!**

**Team NeoStats.**