

Data Engineer Take-Home Case Study

Overview

Welcome to the Data Engineer take-home case study. In this exercise, you will demonstrate your skills in data cleaning, manipulation, joining, modeling, and visualization using multiple datasets. This case study consists of six data files related to a company's employee records, departments, salaries, attendance, projects, and project assignments. Your task is to uncover insights from these datasets and present them in a dashboard.

Instructions

Objective

- Clean and prepare the data.
- Perform data manipulation and joining.
- Analyze the data to uncover insights.
- Create a dashboard to visualize the insights.

Data Files

You will be provided with the following CSV files:

1. **Employees.csv**
2. **Departments.csv**
3. **Salaries.csv**
4. **Attendance.csv**
5. **Projects.csv**
6. **EmployeeProjects.csv**

Tasks

1. **Data Cleaning**
2. **Data Joining**
3. **Data Manipulation and Modeling**
4. **Data Visualization**

Deliverables

- Cleaned and merged dataset(s).
- A summary of the insights found.
- A dashboard visualizing the insights.

Submission Guidelines

- Please submit your cleaned data, analysis, and dashboard.
- You may use any tools or programming languages you prefer (e.g., Python, R, SQL) but not Excel. For the dashboard, you can use PowerBI or Tableau or a visualization tool of your choice.

Evaluation Criteria

- Accuracy of data cleaning and handling of errors.
- Effectiveness and correctness of data joins and manipulations.
- Quality and depth of the analysis.
- Clarity and informativeness of the visualizations.
- Overall presentation and documentation.

We look forward to seeing your work. Good luck!