**Project Prompt: Analyzing Loan Application Data**

**Objective:**

The objective of this project is to analyze loan application data to understand the process flow, activity frequency, and processing time. The project aims to transform raw data into an event log, generate various analyses and visualizations, and compile these into a final report.

**Data Source:**

* Excel workbook: credit\_cleaned.xlsx
* Sheet: "cleaned"

**Requirements:**

1. **Transform the Data into an Event Log**: Convert the raw data into a format suitable for process mining.
2. **Overview of Eventlog (Summary)**: Provide a summary of the event log, including the number of cases, activities, and other relevant statistics.
3. **Frequency of Activities**: Calculate the frequency of each activity in the event log.
4. **Generate the Process Map Graph**: Create a process map to visualize the flow of activities.
5. **Generate a Matrix with an Overview of Activity Tracker Frequency**: Create a matrix that shows the frequency of each activity for each case.
6. **Generate and Explore the Activity Trace**: Create a trace of activities for each case to understand the sequence of activities.
7. **Show Processing Time in Hours by Application Type**: Calculate and display the average processing time for each type of application.
8. **Show Processing Time in Hours per Loan (Credit) Purpose**: Calculate and display the average processing time for each loan purpose.

**Tools:**

* Python for data manipulation and analysis
* Libraries: pandas, matplotlib, seaborn, etc.
* Jupyter Notebook for code and documentation

**Steps:**

1. Load the data from the Excel file into a DataFrame.
2. Clean and preprocess the data as needed.
3. Transform the data into an event log format.
4. Generate an overview of the event log.
5. Calculate the frequency of activities and visualize it.
6. Create a process map graph.
7. Generate a matrix showing activity tracker frequency.
8. Generate and explore the activity trace.
9. Calculate the processing time in hours by application type.
10. Calculate the processing time in hours per loan purpose.
11. Compile all analyses and visualizations into a final report.

**Deliverables:**

* Jupyter Notebook containing all code, analyses, and visualizations
* Final report summarizing the findings