---

1. Data Preprocessing and Cleaning

1. **Data Filtering and KYC Verification**:
    - How can you use .loc[] to filter users from specific countries (e.g., GB, FR) in the **users.csv** dataset based on their **KYC status**?
2. **Data Cleanup and Index Reset**:
    - After filtering the dataset for failed transactions in **transactions.csv**, how can you apply reset_index() to maintain a clean DataFrame structure?
3. **String Manipulation**:
    - Can you apply lstrip() and rstrip() to clean up any string fields, such as **merchant categories** in **transactions.csv** or document properties in **doc_reports.csv**?
4. **Data Type Conversion**:
    - How would you use **.astype()** to convert the **iso_code** column in **currency.csv** to string for consistent formatting across the dataset?
5. **Data Extraction and Column Creation:**
    - How can you extract details like **gender, nationality, document_type, issuing_country**, and **date_of_expiry** from the nested properties field in **doc_reports.csv** and convert them into individual columns using **json_normalize** for better accessibility and analysis?

---

2. Data Analysis and Visualization

Facial Similarity Reports (facial_similarity_reports.csv)

6. **Facial Comparison and Filtering**:
    - How can you filter the dataset to identify users with **non-clear results** in **face_comparison_result** and analyze trends in failed attempts based on **created_at** timestamps?

Document Reports (doc_reports.csv)

7. **Document Verification Bar Chart**:

- How can you visualize the distribution of **visual_authenticity_result** and **image_integrity_result** using a bar chart to highlight potential discrepancies in document verification outcomes?

Bar Chart Visualization (transactions.csv)

8. **Transaction Status Analysis**:
    - How can you use go.Bar() to visualize the number of **completed** vs **declined** transactions in **transactions.csv**?

Pie Chart Visualization (users.csv)

9. **KYC Pass/Fail Distribution**:
    - Can you visualize the proportion of **KYC passed** vs **KYC failed** users in **users.csv** using a pie chart?

Table Representation (fx_rates.csv and transactions.csv)

10. **Currency Exchange Table**:
    - How can figure_factory.create_table() be used to display the top 5 **exchange rates** from **fx_rates.csv** for cryptocurrencies and fiat currencies?

11. **Transaction Summary Table**:

- Can you create a table summarizing the top 10 **countries** by failed transaction rates using **transactions.csv**?

Histogram Analysis (transactions.csv and users.csv)

12. **Transaction Amount Distribution**:

- How can you create a histogram to analyze the distribution of successful **transaction amounts** in **transactions.csv**?

13. **User Age Distribution**:

- Can you generate a histogram to analyze the **age distribution** of users based on the **BIRTH_YEAR** in **users.csv**?

Line Chart Visualization

14. **Transaction Volume Over Time**:

- Can you create a line chart to track **transaction volume** over time using **CREATED_DATE** from **transactions.csv**?

15. **Exchange Rate Fluctuations**:

- How would you visualize **exchange rate fluctuations** over time using data from **fx_rates.csv**, especially focusing on **cryptocurrencies**?

Scatter Plot with Randomized Data

16. **Transaction Analysis by Country**:

- How would you create a scatter plot to show the relationship between **transaction amounts** and **merchant country** in **transactions.csv**, and use different symbols for **completed** and **declined** transactions?

---

3. Advanced Data Exploration

17. **Fraud Detection**:

- How can you merge the **fraudsters.csv** with **users.csv** and **transactions.csv** to analyze fraudulent behavior patterns based on **transaction amounts**, **merchant categories**, and **countries**?

18. **Simulation of Fraudster Activity**:

- Can you generate a randomized line chart to simulate **fraudster activity trends** over time using data from **fraudsters.csv** and **users.csv**?

Conclusion

These questions will allow you to showcase your skills in:

- **Data Preprocessing** (filtering, string manipulation, type conversion).
- **Data Analysis** (merging datasets, identifying fraud).
- **Data Visualization** (bar charts, pie charts, tables, histograms, line charts, scatter plots).

This case study will provide a comprehensive analysis of the different datasets and demonstrate your ability to clean, explore, and visualize real-world data effectively.