PayPal Data Analyst End-to-End Project (Interview Ready)

Business Context / Problem Statement

PayPal operates a global digital payments platform handling millions of transactions daily. The company wants to analyze transaction performance, revenue trends, payment method reliability, user behavior, and chargeback risks. As a Data Analyst, your responsibility is to analyze PayPal-style transaction data and provide insights that help improve success rates, reduce failures, and maximize revenue.

Project Objective

1. Analyze 200,000+ payment transactions
2. Identify revenue trends and growth patterns
3. Evaluate transaction success and failure rates
4. Understand user behavior and high-value users
5. Detect potential risk and chargeback indicators

Tools & Skill Requirements

1. SQL (Intermediate): Joins, CTEs, Window Functions, Aggregations
2. Python (Basic): Pandas, Data Cleaning, GroupBy, Date Handling
3. Excel: Pivot Tables, LOOKUPs, Conditional Formatting
4. Power BI: Basic DAX, KPI Cards, Interactive Dashboard

Dataset Overview (200K+ Rows)

1. users.csv – 20,000 rows (user profile data)
2. transactions.csv – 200,000 rows (payment transactions)
3. payment_methods.csv – payment fee structure
4. chargebacks.csv – 5,000 rows (fraud & disputes)

Data Access Instructions

Generate the data using Python or Google Sheets as discussed. Upload CSV files to Google Sheets or import them into SQL / Power BI. You may host the files on Google Drive or GitHub for sharing. Recommended folder structure and generation logic are provided during project execution.

SQL Questions (Intermediate)

1. Calculate total and monthly revenue from successful transactions
2. Find top 10 users by total transaction value

3. Calculate transaction success and failure rate by payment method
4. Identify users with declining activity in the last 30 days
5. Calculate month-over-month revenue growth
6. Find revenue lost due to failed transactions
7. Compute chargeback rate across all transactions
8. Identify repeat users with more than 5 transactions
9. Compare business vs personal account performance
10. Calculate average gap between user transactions

Python Questions (Basic)

1. Clean invalid or negative transaction amounts
2. Calculate percentage of failed transactions
3. Find top 5 countries by revenue
4. Detect high-value outlier transactions
5. Calculate average transaction value per payment method
6. Identify inactive users (no transaction in last 60 days)
7. Calculate total PayPal fees collected
8. Analyze failed transactions by merchant category
9. Calculate revenue contribution percentage by payment method
10. Apply a simple fraud rule for large transactions

Final Deliverables

1. SQL scripts with comments
2. Python notebooks for cleaning and analysis
3. Excel file with business pivots
4. Power BI dashboard with KPIs
5. GitHub README with insights and screenshots