Pseudo Code:

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
from scipy.stats import describe, pearsonr

data = {
   "timestamp": generate_daily_timestamps(start_date="2024-01-01", total_days=100),
   "transaction_id": generate_sequence(1, 101),
   "user_id": generate_random_integers(low=1, high=20, count=100),
   "amount": generate_random_floats(low=-500, high=1000, count=100),
   "transaction_type": assign_random_values(["credit", "debit"], count=100)
}

df = pd.DataFrame(data)
print_shape(df)

print_missing_values(df)

duplicates = count_duplicates(df)
print(duplicates)

df = remove_duplicates(df)
print_shape(df)

print_summary_statistics(df)

print_transaction_type_counts(df)

plot_histogram(df["amount"], title="Distribution of Transaction Amounts", xlabel="Amount",
ylabel="Frequency")
plot_line(df["timestamp"], df["amount"], title="Transaction Amounts Over Time",
xlabel="Date", ylabel="Amount")
stats = compute_descriptive_statistics(df["amount"])
print(stats)

correlation, p_value = calculate_correlation(df["user_id"], df["amount"])
print(correlation, p_value)

top_users = find_top_users_by_amount(df, top_n=5)
print(top_users)

df["amount_category"] = categorize_transaction_amounts(df["amount"], bins=[-inf, 0, 500,
```

```
inf], labels=["Low", "Medium", "High"])

print_transaction_category_counts(df)

plot_bar_chart(df["amount_category"].value_counts(), title="Transaction Amount Categories", xlabel="Category", ylabel="Count")
```

Instruction:

You copy the script into a python environment with the libraries needed (pandas, numpy, matplotlib, scipy)

- Run the script to view analysis on the data
- Make changes to the data generation or analysis steps based on your specific needs

This script will cover:

- Checking for missing values or duplicates
- Displays summary statistics and counts transaction types
- Includes histograms and line plots for transaction amounts
- Provides descriptive statistics and examines correlations
- Highlights top users by transaction amounts and categorizes transactions

---

Analysis:

Plaid Transactions:
-The transaction amounts are widely spread, with a higher concentration of amounts near the lower and middle ranges. There is a noticeable tail in both directions, indicating outliers in both high expenses and high credits.

Open Exchange Rates:
-Exchange rates for USD to EUR, GBP, and JPY are stable with slight fluctuations. The USD to JPY exchange rate has the highest variance compared to USD to EUR and GBP.

Alpaca Stock Trades:
-Total Quantity: TSLA has the highest total quantity of trades, followed by GOOGLE, AMZN, and AAPL.
-Average Price: AAPL has the highest average trade price, while TSLA has the lowest, showing variability in stock pricing.

Firefly Budgets:
-Most budget categories are evenly distributed in the pie chart. However, specific categories like "Food" or "Rent" could dominate depending on allocation trends.
-There is a notable gap between allocated and spent amounts for some budgets, suggesting underspending or over-allocation in certain categories.

Visualization