Link for jupiter notebook-

```python
import pandas as pd
import numpy as np
import seaborn as sns
import matplotlib.pyplot as plt

%matplotlib inline


df = pd.read_csv(r"C:\Users\shree\Downloads\credit_card_fraud_50k.csv")


df.head()

df.info()
df.describe()

df.isnull().sum()

df.fillna(df.median(numeric_only=True), inplace=True)

df.drop_duplicates(inplace=True)
df.shape

numeric_cols = df.select_dtypes(include=np.number).columns

for col in numeric_cols:
    lower = np.percentile(df[col], 1)
    upper = np.percentile(df[col], 99)
    df[col] = np.clip(df[col], lower, upper)



sns.countplot(x="is_fraud", data=df)
plt.title("Fraud vs Non-Fraud Transactions")
plt.show()


sns.histplot(df["transaction_amount"], bins=50, kde=True)
plt.title("Transaction Amount Distribution")
plt.show()


plt.figure(figsize=(10,6))
sns.heatmap(df.corr(numeric_only=True), annot=True, cmap="coolwarm")
```

```python
plt.title("Feature Correlation Heatmap")
plt.show()

sns.boxplot(x="is_fraud", y="transaction_amount", data=df)
plt.title("Transaction Amount vs Fraud")
plt.show()

fraud_counts = df["is_fraud"].value_counts()

plt.figure(figsize=(6,6))
plt.pie(fraud_counts, labels=["Non-Fraud", "Fraud"], autopct="%1.1f%%", startangle=90)
plt.title("Fraud Transaction Percentage")
plt.show()

avg_amount = df.groupby("is_fraud")["transaction_amount"].mean()

avg_amount.plot(kind="bar")
plt.title("Average Transaction Amount by Fraud")
plt.xlabel("Fraud (0 = No, 1 = Yes)")
plt.ylabel("Average Amount")
plt.show()


df["transaction_amount"].head(200).plot(kind="line")
plt.title("Transaction Amount Trend")
plt.xlabel("Transaction Index")
plt.ylabel("Amount")
plt.show()


cols = ["transaction_amount", "customer_age", "transaction_hour", "is_fraud"]

sns.pairplot(df[cols], hue="is_fraud")
plt.show()

fraud_summary = df.groupby("is_fraud").mean(numeric_only=True)
fraud_summary
```