Analysis Deliverables

Submitted By:- Tejus Kumar Sharma

Contact No:- 9109594235

For Data Cleaning and Calculation of KPI Indicators

```
import pandas as pd
import numpy as np
# 1. Load
df = pd.read_excel("AmericanExpress_Data.xlsx")
# 2. Clean column names
df.columns = (
    df.columns
    .str.strip()
    .str.replace(" ", "_")
    .str.replace("($M)", "M", regex=False)
    .str.replace("%", "perc", regex=False)
    .str.replace("$", "", regex=False)  # remove $ sign safely
)
# 3. Explicit rename for consistency
df = df.rename(columns={
    "Active_Cards_(M)": "Active_Cards_M",
    "Avg_Spend_per_Card_(_)": "Avg_Spend_per_Card"  # optional cleanup
})

print("Final column names:", df.columns.tolist())
# 4. Ensure numeric columns
for col in ["Revenue_M", "Net_Income_M", "Active_Cards_M", "Delinquency_perc"]:
    if col in df.columns:
        df[col] = pd.to_numeric(df[col], errors="coerce")
```

 # 5. Core KPI calculations

```
df["Net_Income_Margin_perc"] = (df["Net_Income_M"] / df["Revenue_M"]) * 100

df["YoY_Revenue_Growth_perc"] = (
    df.groupby(["Region", "Segment", "Quarter"])["Revenue_M"]
    .pct_change(periods=1) * 100
)

df["Revenue_per_ActiveCard"] = df["Revenue_M"].div(df["Active_Cards_M"]).replace([np.inf, -np.inf], np.nan)

df["Delinq_to_Revenue_perc"] = (df["Delinquency_perc"] / df["Revenue_M"]) * 100
```

 # 6. Export

```
df.to_excel("Amex_Cleaned_Data2.xlsx", index=False)

print("Data cleaning & KPI calculations complete! File saved as Amex_Cleaned_Data2.xlsx")
```

DAX Variables

Net Income Margin % = DIVIDE(SUM(Amex_Cleaned_Data[Net_Income_M]), SUM(Amex_Cleaned_Data[Revenue_M]))

DateTable =

```
ADDCOLUMNS (
    CALENDAR ( DATE (2018, 1, 1), DATE (2025, 12, 31) ),
    "Year", YEAR ( [Date] ),
    "Quarter", "Q" & FORMAT ( [Date], "Q" )
)
```

Total Revenue = SUM (Sheet1[Revenue_M]) Previous Year Revenue =

```
CALCULATE (
    [Total Revenue],
    DATEADD ( 'DateTable'[Date], -1, YEAR )
) Previous Year Revenue =
CALCULATE (
```

[Total Revenue],

DATEADD ('DateTable'[Date], -1, YEAR)

)

Revenue per Active Card = DIVIDE(SUM(Amex_Cleaned_Data[Revenue_M]), SUM(Amex_Cleaned_Data[Active_Cards_M]))

Delinquency-to-Revenue Ratio = DIVIDE(SUM(Amex_Cleaned_Data[Delinquency_perc]), SUM(Amex_Cleaned_Data[Revenue_M]))