# payments\_accounting\_analysis.ipynb

# --- 1. Import Libraries ---

import pandas as pd

import numpy as np

# --- 2. Load Data ---

transactions = pd.read\_csv('transactions.csv', parse\_dates=['transaction\_date', 'settlement\_date'])

customers = pd.read\_csv('customers.csv')

merchants = pd.read\_csv('merchants.csv')

# --- 3. Basic Overview ---

print("Transactions Shape:", transactions.shape)

print("Date Range:", transactions['transaction\_date'].min(), "to", transactions['transaction\_date'].max())

print("Columns:", transactions.columns.tolist())

# --- 4. Data Cleaning ---

# Remove time component for date join

transactions['transaction\_date'] = transactions['transaction\_date'].dt.date

# Ensure numeric fields are correct

for col in ['amount', 'fee', 'tax', 'net\_amount', 'ledger\_amount', 'refund\_amount']:

transactions[col] = pd.to\_numeric(transactions[col], errors='coerce')

# --- 5. KPI Calculations ---

total\_amount = transactions['amount'].sum()

total\_fees = transactions['fee'].sum()

total\_tax = transactions['tax'].sum()

net\_revenue = total\_amount - total\_fees - total\_tax

chargeback\_count = transactions[transactions['status'] == 'chargeback'].shape[0]

chargeback\_rate = chargeback\_count / len(transactions)

unmatched\_count = transactions[transactions['reconciliation\_status'] == 'unmatched'].shape[0]

mismatch\_amount = (transactions['net\_amount'] - transactions['ledger\_amount']).abs().sum()

print(f"Total Amount: {total\_amount}")

print(f"Net Revenue: {net\_revenue}")

print(f"Chargeback Rate: {chargeback\_rate:.2%}")

print(f"Unmatched Count: {unmatched\_count}")

print(f"Reconciliation Mismatch: {mismatch\_amount}")

# --- 6. Grouped Insights ---

# Net Revenue by Region

net\_rev\_region = transactions.groupby('region')['net\_amount'].sum().reset\_index()

# Chargeback by Country

chargeback\_country = (transactions[transactions['status'] == 'chargeback']

.groupby('country')['transaction\_id']

.count().reset\_index()

.rename(columns={'transaction\_id': 'chargeback\_count'}))

# --- 7. Save Summary ---

summary = {

"Total Amount": total\_amount,

"Net Revenue": net\_revenue,

"Chargeback Rate": chargeback\_rate,

"Unmatched Transactions": unmatched\_count

}

summary\_df = pd.DataFrame([summary])

summary\_df.to\_csv("payments\_summary.csv", index=False)

print("Summary saved as payments\_summary.csv")