Data Analytics Project

Topic: B2B Courier Charges Accuracy Analysis
By: Marcel Basumatary (ET22BTHCS028) (6th Sem A) CSE

1.Company ABC - Pincode Zones
2.Company ABC - SKU Master
3.Company ABC- Order Report
4.Courier Company - Rates
5.courier_invoice

Code:-

```
import pandas as pd
import matplotlib.pyplot as plt

# Define file paths
base_path = r"C:\Users\basum\OneDrive\Documents\Downloads\Downloads\archive"
datasets = {
    "pincode_zones": f"{base_path}\\Company ABC - Pincode Zones.csv",
    "sku_master": f"{base_path}\\Company ABC - SKU Master.csv",
    "order_report": f"{base_path}\\Company ABC- Order Report.csv",
    "courier_rates": f"{base_path}\\Courier Company - Rates.csv"
}

# Load datasets
df_pincode = pd.read_csv(datasets["pincode_zones"])
df_sku = pd.read_csv(datasets["sku_master"])
df_order = pd.read_csv(datasets["order_report"])
df_rates = pd.read_csv(datasets["courier_rates"])

# Pincode Zone Analysis
zone_counts = df_pincode["Zone"].value_counts()
zone_counts.to_csv(f"{base_path}\\zone_analysis.csv", index=True)

# SKU Master Analysis
sku_counts = df_sku["SKU"].value_counts()
weight_stats = df_sku["Weight (g)"].describe()
sku_counts.to_csv(f"{base_path}\\sku_analysis.csv", index=True)

# Order Report Analysis
sku_order_qty = df_order.groupby("SKU")["Order Qty"].sum().sort_values(ascending=False)
sku_order_qty.to_csv(f"{base_path}\\order_analysis.csv", index=True)
```

```python
# Courier Rates Analysis
forward_fixed_avg = df_rates.iloc[0, :5].mean()
forward_additional_avg = df_rates.iloc[0, 5:10].mean()
return_fixed_avg = df_rates.iloc[0, 10:15].mean()
return_additional_avg = df_rates.iloc[0, 15:].mean()
df_rate_summary = pd.DataFrame({
    "Category": ["Forward Fixed", "Forward Additional", "Return Fixed", "Return Additional"],
    "Average Cost": [forward_fixed_avg, forward_additional_avg, return_fixed_avg, return_additional_avg]
})
df_rate_summary.to_csv(f"{base_path}\\rate_analysis.csv", index=False)

# Visualization - SKU Order Distribution
plt.figure(figsize=(10, 5))
plt.bar(sku_order_qty.index[:10], sku_order_qty.values[:10], color="skyblue")
plt.xlabel("SKU")
plt.ylabel("Total Order Quantity")
plt.title("Top 10 SKU Order Distribution")
plt.xticks(rotation=45, ha="right")
plt.grid(axis="y", linestyle="--", alpha=0.7)
plt.show()

print("Analysis complete! Files saved in the archive folder.")
```

Output:-

1.Pincode Zones Analysis

	A	B
1	Zone	count
2	b	79
3	d	38
4	e	7

2.SKU Master Analysis

	A	B
1	SKU	count
2	8904220000000	60
3	GIFTBOX202002	2
4	GIFTBOX202001	1
5	GIFTBOX202004	1
6	GIFTBOX202003	1
7	SACHETS001	1
8		

3.Order Report Analysis

	A	B
1	Category	Average Cost
2	Forward Fixed	30.9
3	Forward Additional	48.24
4	Return Fixed	23.58
5	Return Additional	46.24
6		

4.Rates Analysisya

	A	B	C	D	E	F	G	H	I	J
1	fwd_a_fixed	fwd_a_additional	fwd_b_fixed	fwd_b_additional	fwd_c_fixed	fwd_c_additional	fwd_d_fixed	fwd_d_additional	fwd_e_fixed	fwd_e_additional
2	29.5	23.6	33	28.3	40.1	38.9	45.4	44.8	56.6	55.5

K	L	M	N	O	P	Q	R	S	T
rto_a_fixed	rto_a_additional	rto_b_fixed	rto_b_additional	rto_c_fixed	rto_c_additional	rto_d_fixed	rto_d_additional	rto_e_fixed	rto_e_additional
13.6	23.6	20.5	28.3	31.9	38.9	41.3	44.8	50.7	55.5

5.invoice Analysis

	A	B
1	Zone	Billing Amount (Rs.)
2	d	11357.5
3	b	1969.9
4	e	320.8