# **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)
```

```
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

	А	В	
1	Zone	count	
2	b	79	
3	d	38	
4	е	7	

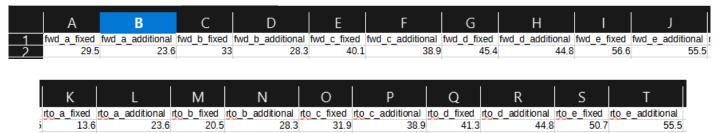
## 2.SKU Master Analysis

	Α	В	
1	SKU	count	
2	8904220000000	60	
3	GIFTBOX202002	2	
4	GIFTBOX202001	1	
5	GIFTBOX202004	1	
6	GIFTBOX202003	1	
7	SACHETS001	1	
 g			

## 3.<u>Order Report Analysis</u>

	Α	В	
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



55.5

### **5.invoice Analysis**

	Α	В
1	Zone	Billing Amount (Rs.)
2	d	11357.5
3	b	1969.9
4	е	320.8