Marwadi University Faculty of Technology Department of Information and Communication Technology		
Subject: Data Visualization and Dashboards (01CT0410)	**Aim:** Analysis of Superstore Dataset	
Case Study - 1	**Date:-** 16-03-2024	**Enrollment No:-** 92200133030

Aim: Analysis of Superstore Dataset

IDE: Microsoft Excel, Tableau , Spyder

Pre-Requisites:- Insert the Customer's Full Name, Gender, City, and Country From the Customer Sheet Using VLOOKUP Function In Order Dataset In Excel Itself on the Basis of Customer ID.

Now Import Necessary Libraries for Analysis:-

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

Dataset = pd.read_excel("./Superstore.xlsx",'Orders')
Dataset['Order_Date'] = pd.to_datetime(Dataset['Order_Date'])
Dataset["Year"] = Dataset["Order_Date"].dt.year
Dataset["Quater"] = Dataset["Order_Date"].dt.quarter
Dataset["Month"] = Dataset["Order_Date"].dt.month
```

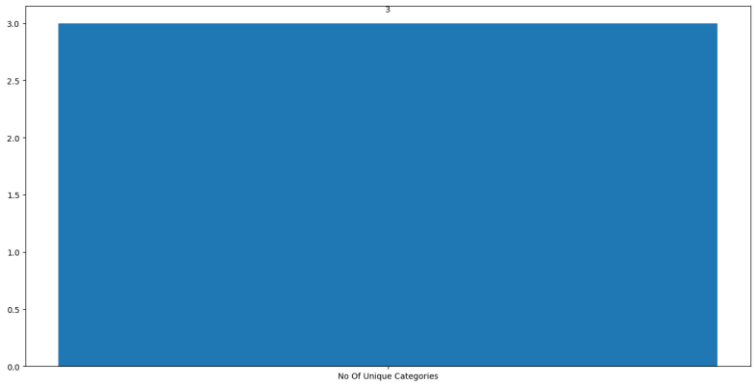
Questions:

1) **What are the product categories available for customers to shop?**

Code:-

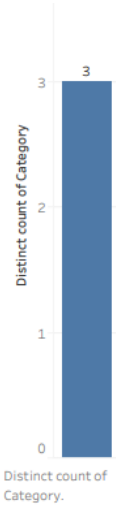
```
Unique_Categories = list(set(Dataset["Category"]))
plt.figure(figsize=(16, 8))
plt.bar(x=["No Of Unique Categories"], height=[len(Unique_Categories)], width=0.5)
for i, v in enumerate([len(Unique_Categories)]):
    plt.text(i, v + 0.1, str(v), ha="center")
plt.show()
```

Output :-



Marwadi University Marwadi Chandarana Group NAAC A+	**Marwadi University** **Faculty of Technology** **Department of Information and Communication Technology**	
Subject: Data Visualization and Dashboards (01CT0410)	**Aim:** Analysis of Superstore Dataset	
Case Study - 1	**Date:-** 16-03-2024	**Enrollment No:-** 92200133030

Tableau Workbook :-



2) How many products are there in each category?

Code:-

```
plt.figure(figsize=(20, 10))
ax = sns.countplot(x=Dataset["Category"], hue=Dataset["Category"])
for p in ax.patches:
   ax.annotate(
      f"{p.get_height()}",
      (p.get_x() + p.get_width() / 2.0, p.get_height()),
      ha="center",
      va="center",
      fontsize=12,
      color="black",
      xytext=(0, 5),
      textcoords="offset points",
   )

plt.show()
```

Marwadi University Marwadi Chandarana Group NAAC A+	**Marwadi University** **Faculty of Technology** **Department of Information and Communication Technology**	
Subject: Data Visualization and Dashboards (01CT0410)	**Aim:** Analysis of Superstore Dataset	
Case Study - 1	**Date:-** 16-03-2024	**Enrollment No:-** 92200133030

Output :-

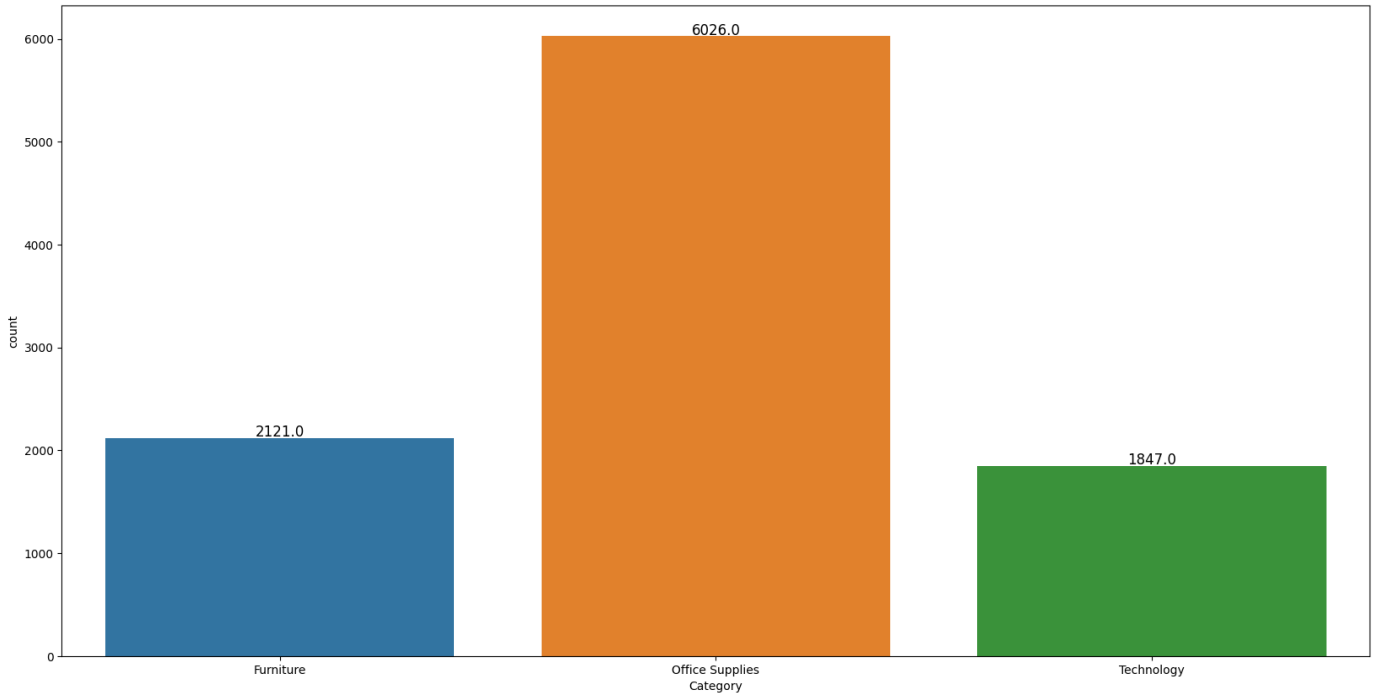
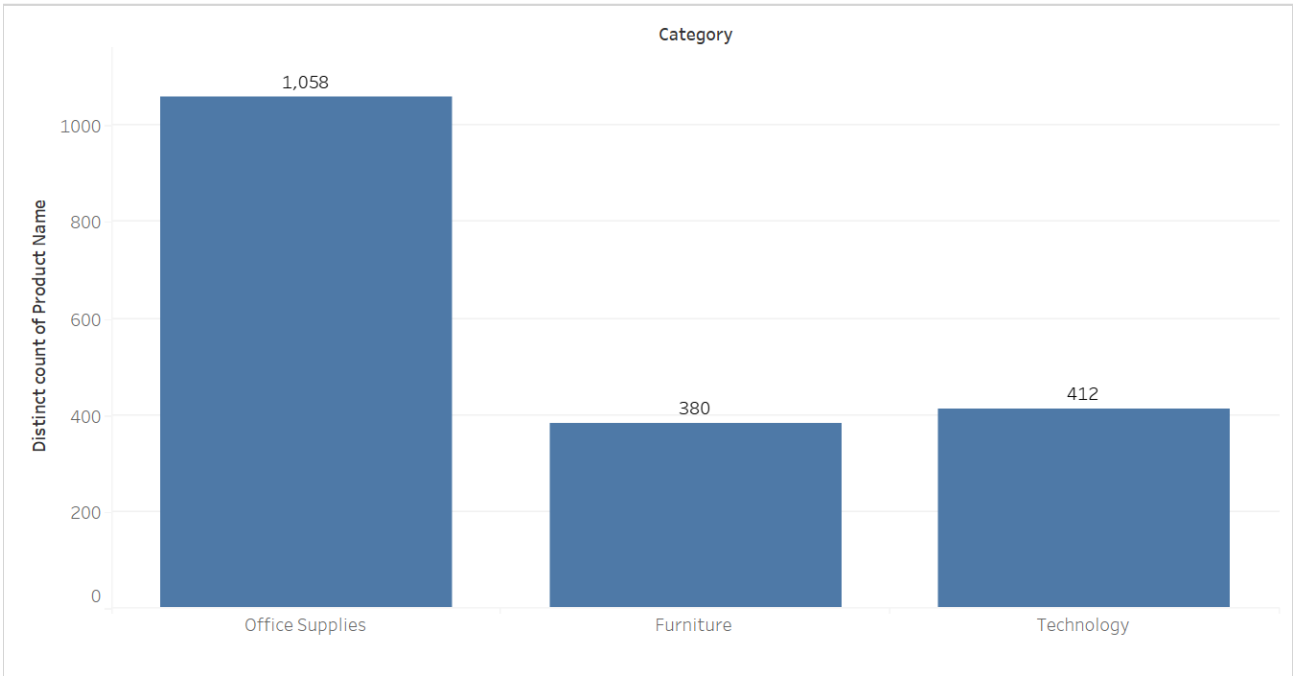


Tableau Workbook :-



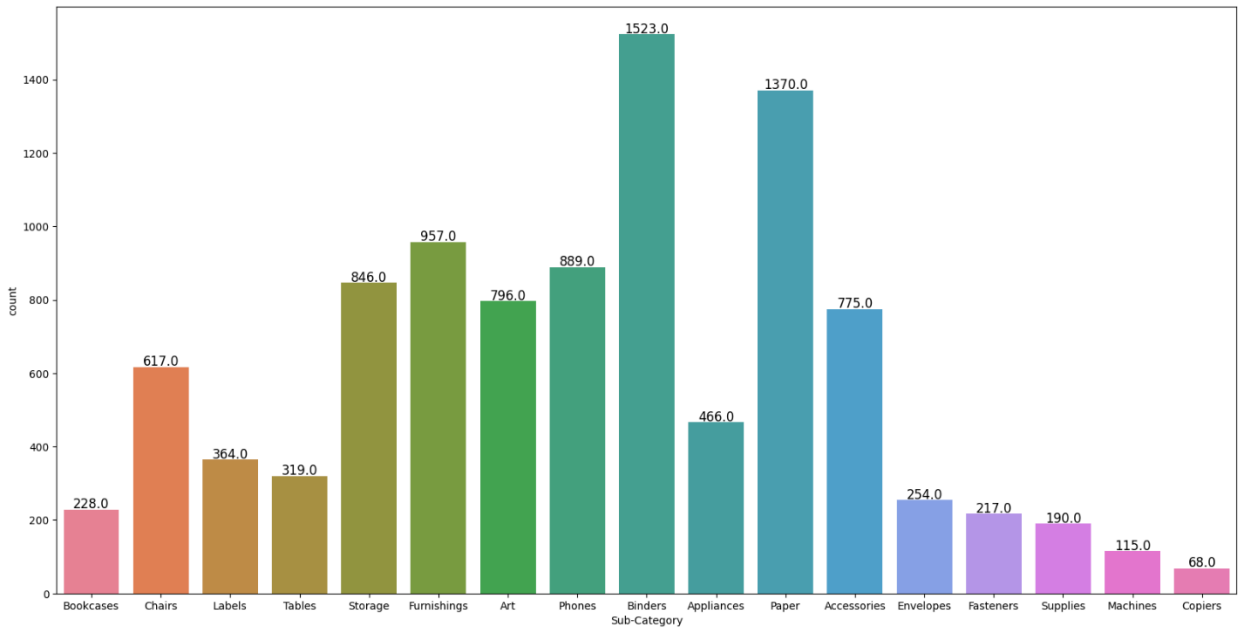
Marwadi University Marwadi Chandarana Group NAAC A+	**Marwadi University** **Faculty of Technology** **Department of Information and Communication Technology**	
Subject: Data Visualization and Dashboards (01CT0410)	**Aim:** Analysis of Superstore Dataset	
Case Study - 1	**Date:-** 16-03-2024	**Enrollment No:-** 92200133030

3) What is the count of products in each sub-category?

Code:-

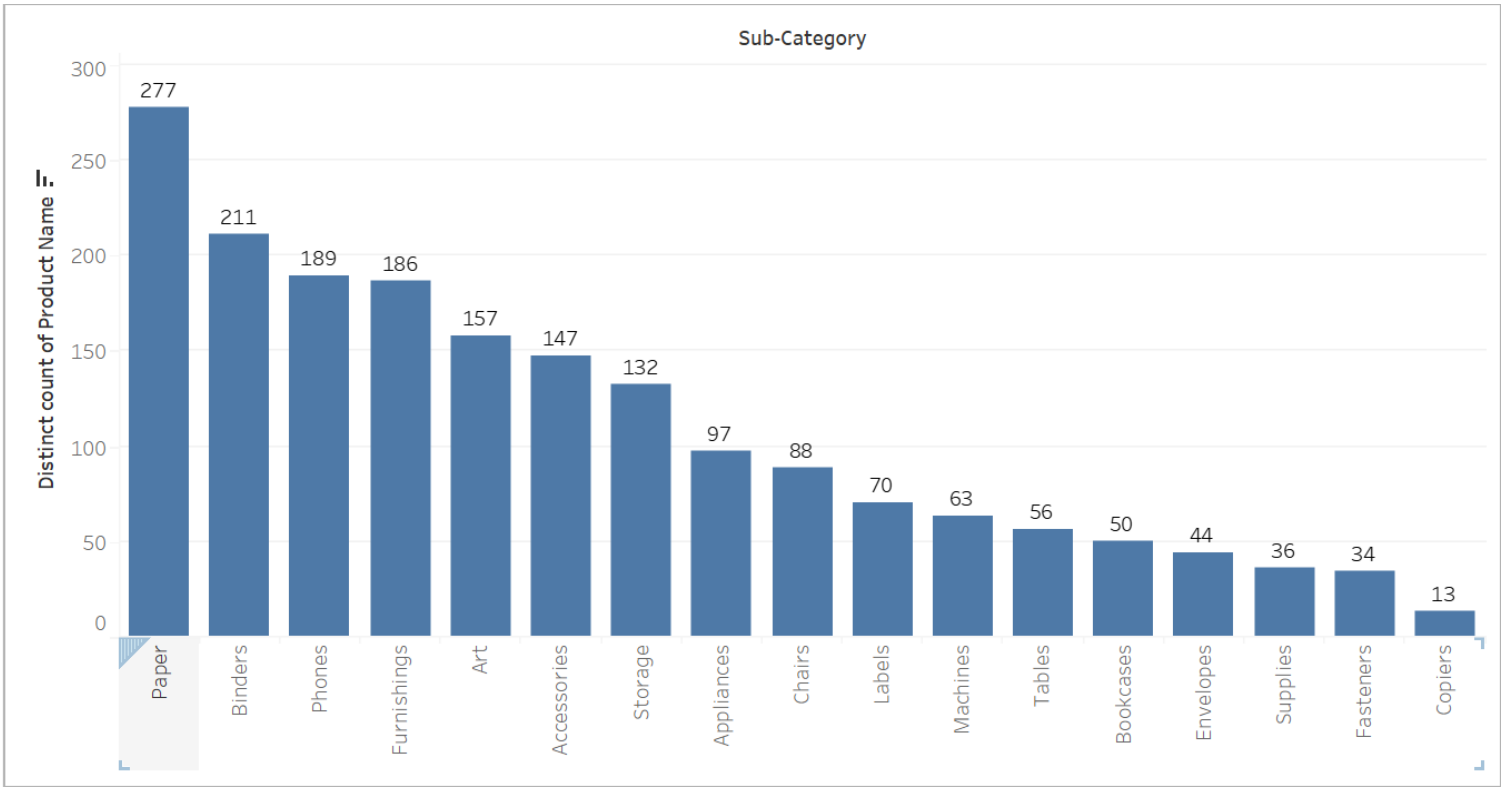
```
Sub_categgory_wise_product_distribuition = Dataset["Sub-Category"].value_counts()
print(Sub_categgory_wise_product_distribuition)
plt.figure(figsize=(20, 10))
ax = sns.countplot(x=Dataset["Sub-Category"], hue=Dataset["Sub-Category"])
for p in ax.patches:
    ax.annotate(
        f"{p.get_height()}",
        (p.get_x() + p.get_width() / 2.0, p.get_height()),
        ha="center",
        va="center",
        fontsize=12,
        color="black",
        xytext=(0, 5),
        textcoords="offset points",)
plt.show()
```

Output :-



Marwadi University Marwadi Chandarana Group NAAC A+	**Marwadi University** **Faculty of Technology** **Department of Information and Communication Technology**	
Subject: Data Visualization and Dashboards (01CT0410)	**Aim:** Analysis of Superstore Dataset	
Case Study - 1	**Date:-** 16-03-2024	**Enrollment No:-** 92200133030

Tableau Workbook :-



4) How are sub-categories distributed in relation to their respective categories?

Code:-

```
Distribution_of_sub_categgory_wrt_Category = Dataset.groupby('Category')['Sub-Category'].unique()
print(Distribution_of_sub_categgory_wrt_Category)
```

Output :-

```
Category
Furniture                    [Bookcases, Chairs, Tables, Furnishings]
Office Supplies    [Labels, Storage, Art, Binders, Appliances, Pa...
Technology                   [Phones, Accessories, Machines, Copiers]
Name: Sub-Category, dtype: object
```

Marwadi University Marwadi Chandarana Group NAAC A+	**Marwadi University** **Faculty of Technology** **Department of Information and Communication Technology**	
Subject: Data Visualization and Dashboards (01CT0410)	**Aim:** Analysis of Superstore Dataset	
Case Study - 1	**Date:-** 16-03-2024	**Enrollment No:-** 92200133030

Tableau Workbook :-

Category	Sub-Catego..	
Furniture	Bookcases	Abc
	Chairs	Abc
	Furnishings	Abc
	Tables	Abc
Office Supplies	Appliances	Abc
	Art	Abc
	Binders	Abc
	Envelopes	Abc
	Fasteners	Abc
	Labels	Abc
	Paper	Abc
	Storage	Abc
	Supplies	Abc
Technology	Accessories	Abc
	Copiers	Abc
	Machines	Abc
	Phones	Abc

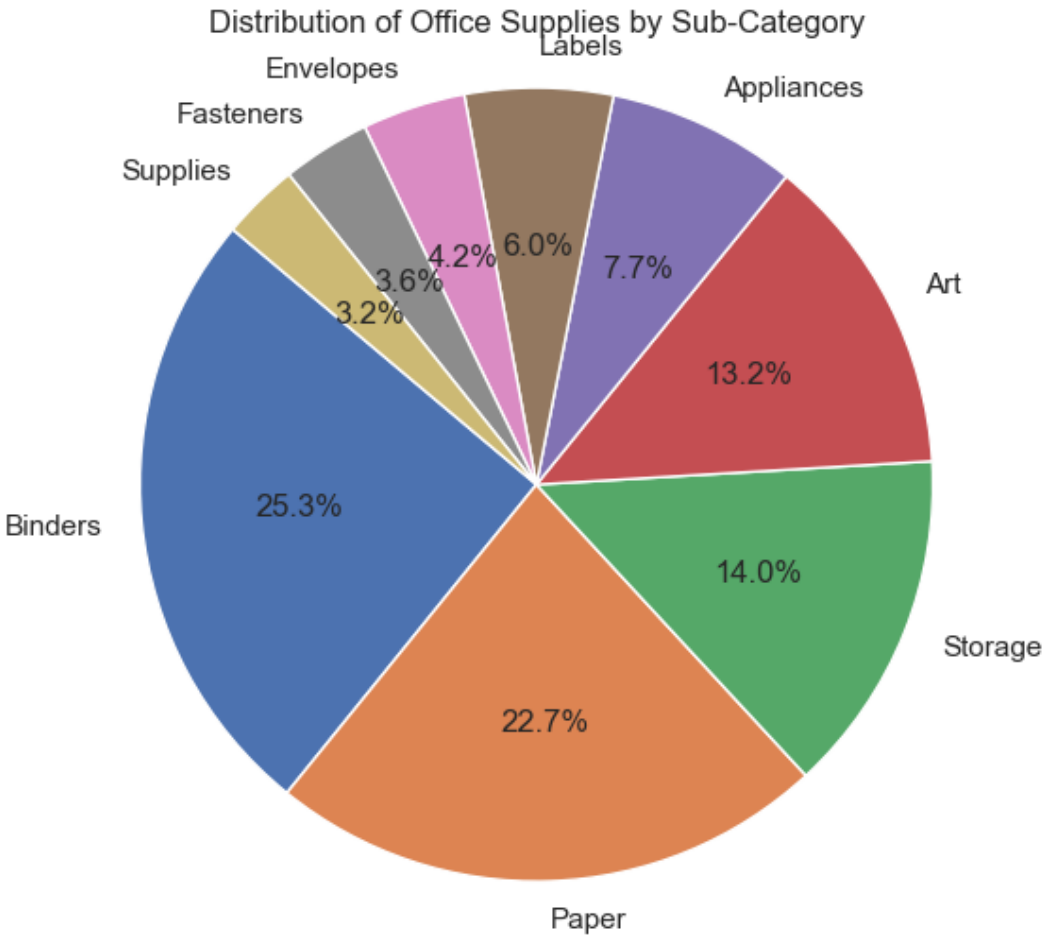
5) What is the percentage distribution of varieties of Office Supplies?

Code:-

```
Office_Supplies_Distribution = (
    Dataset[Dataset["Category"] == "Office Supplies"]["Sub-Category"]
    .value_counts()
    .rename("Counts")
)
total_office_supplies = Office_Supplies_Distribution.sum()
Office_Supplies_Distribution_with_percentage = (
    Office_Supplies_Distribution / total_office_supplies
) * 100
Office_Supplies_Distribution_with_percentage = (
    Office_Supplies_Distribution_with_percentage.rename("Percentage")
)
Distribution_of_office_Supplies = pd.concat(
    [Office_Supplies_Distribution, Office_Supplies_Distribution_with_percentage], axis=1
```

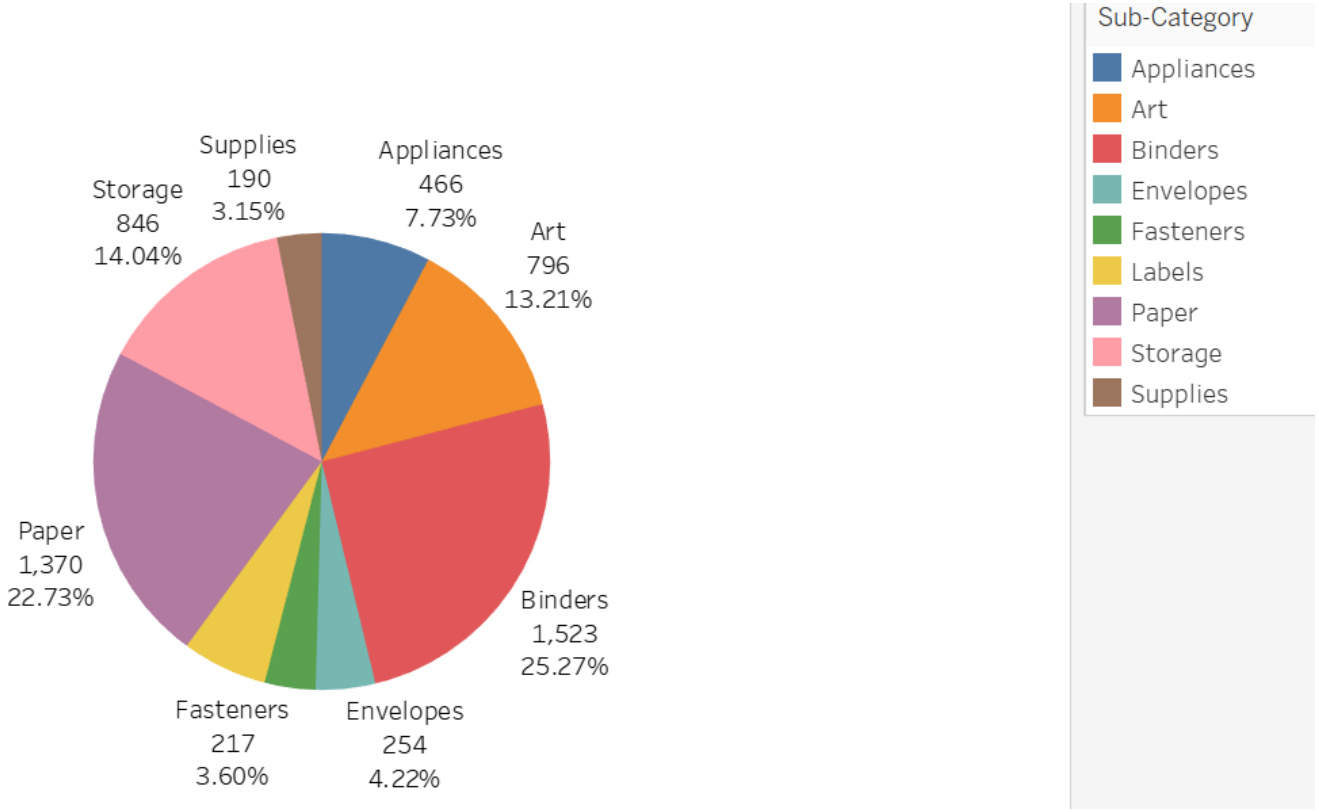
```
)

plt.figure(figsize=(10, 6))
plt.pie(
    Distribution_of_office_Supplies["Counts"],
    labels=Distribution_of_office_Supplies.index,
    autopct="%1.1f%%",
    startangle=140,
)
plt.title("Distribution of Office Supplies by Sub-Category")
plt.axis("equal")
plt.show()
```

Output :-



Marwadi University Marwadi Chandarana Group NAAC A+	**Marwadi University** **Faculty of Technology** **Department of Information and Communication Technology**	
Subject: Data Visualization and Dashboards (01CT0410)	**Aim:** Analysis of Superstore Dataset	
Case Study - 1	**Date:-** 16-03-2024	**Enrollment No:-** 92200133030

Tableau Workbook:-



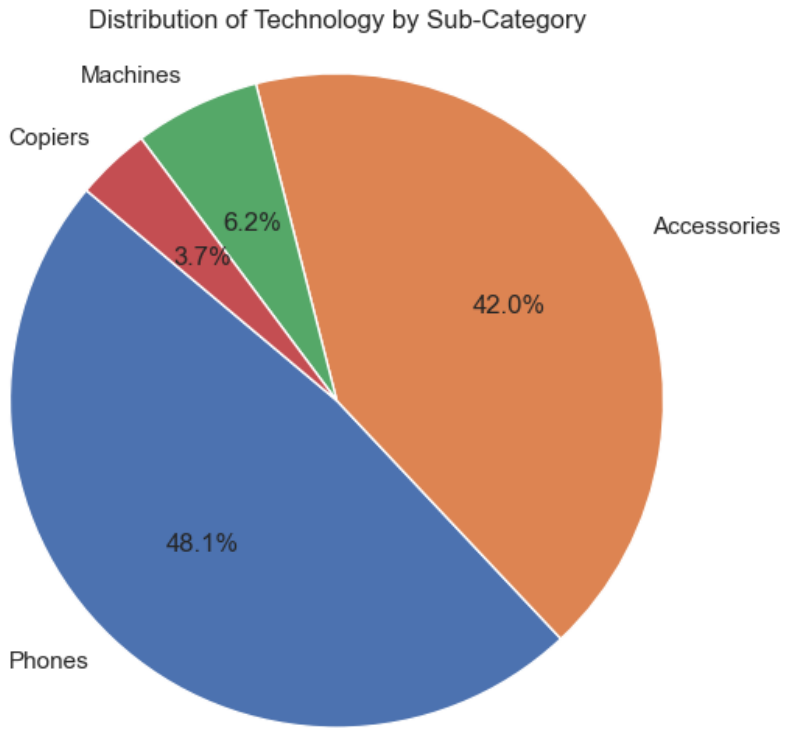
6) What is the percentage distribution of varieties of Technology?

Code:-

```
Technology_Distribution = (
    Dataset[Dataset["Category"] == "Technology"]["Sub-Category"]
    .value_counts()
    .rename("Counts")
)
total_technology = Technology_Distribution.sum()
Technology_Distribution_with_percentage = (
    Technology_Distribution / total_technology
) * 100
Technology_Distribution_with_percentage = (
    Technology_Distribution_with_percentage.rename("Percentage")
)
```

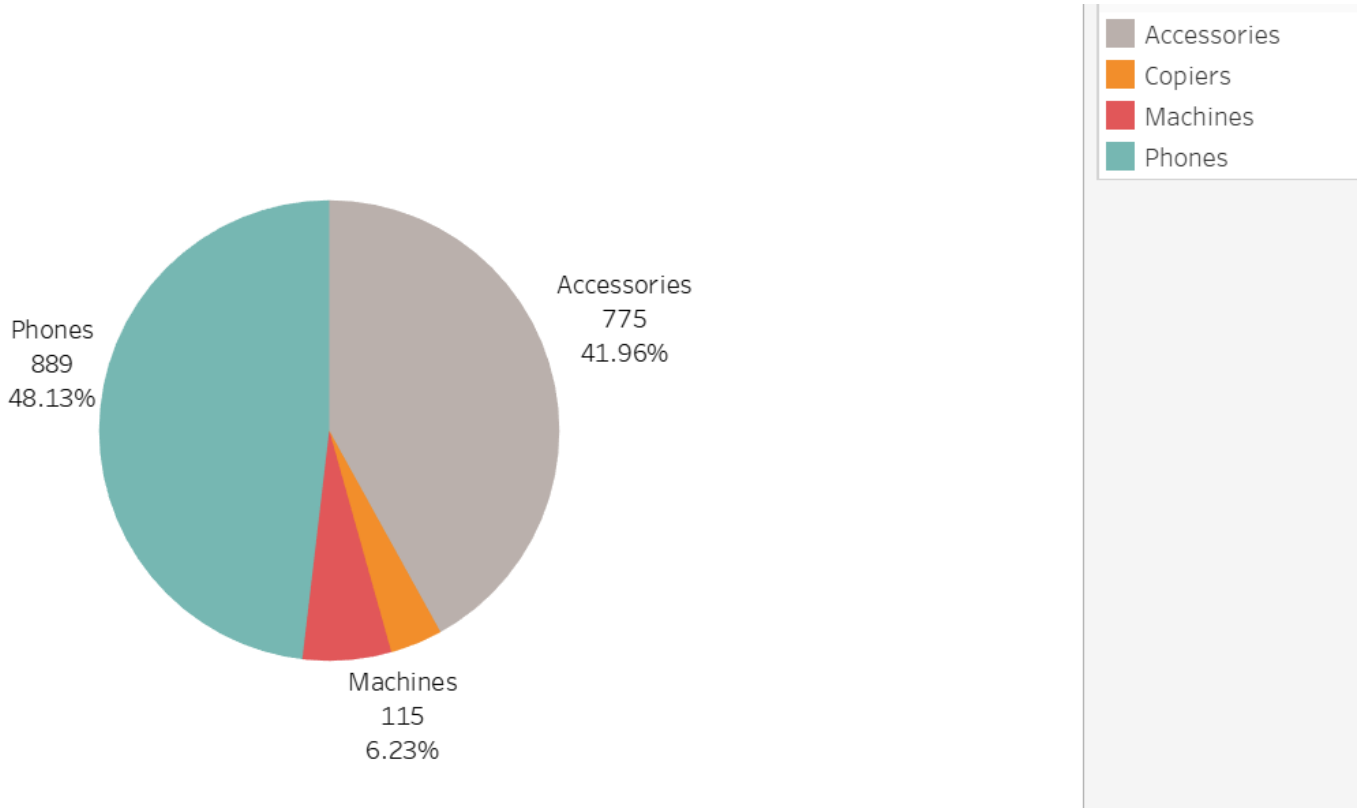
```
Distribution_of_Technology = pd.concat(
   [Technology_Distribution, Technology_Distribution_with_percentage], axis=1
)

print(Distribution_of_Technology)

plt.figure(figsize=(10, 6))
plt.pie(
   Distribution_of_Technology["Counts"],
   labels=Distribution_of_Technology.index,
   autopct="%1.1f%%",
   startangle=140,
)
plt.title("Distribution of Technology by Sub-Category")
plt.axis("equal")
plt.show()
```

Output:-



Marwadi University Marwadi Chandarana Group NAAC A+	**Marwadi University** **Faculty of Technology** **Department of Information and Communication Technology**	
Subject: Data Visualization and Dashboards (01CT0410)	**Aim:** Analysis of Superstore Dataset	
Case Study - 1	**Date:-** 16-03-2024	**Enrollment No:-** 92200133030

Tableau Workbook :-



7) What is the percentage distribution of varieties of Furniture?

Code:-

```
Furniture_Distribuition = (Dataset[Dataset["Category"] == "Furniture"]["Sub-
Category"].value_counts().rename("Counts"))
total_furniture = Furniture_Distribuition.sum()
Furniture_Distribution_with_percentage = (Furniture_Distribuition / total_furniture) * 100
Furniture_Distribution_with_percentage = Furniture_Distribution_with_percentage.rename("Percentage")
Distribution_of_Furniture = pd.concat([Furniture_Distribuition, Furniture_Distribution_with_percentage],
axis=1)
print(Distribution_of_Furniture)
plt.figure(figsize=(10, 8))
plt.pie(
    Distribution_of_Furniture["Counts"],
    labels=Distribution_of_Furniture.index,
    autopct="%1.1f%%",
    startangle=140,
```

```
)
plt.title("Distribution of Furniture Sub-Categories")
plt.axis("equal")
plt.show()
```

Output :-

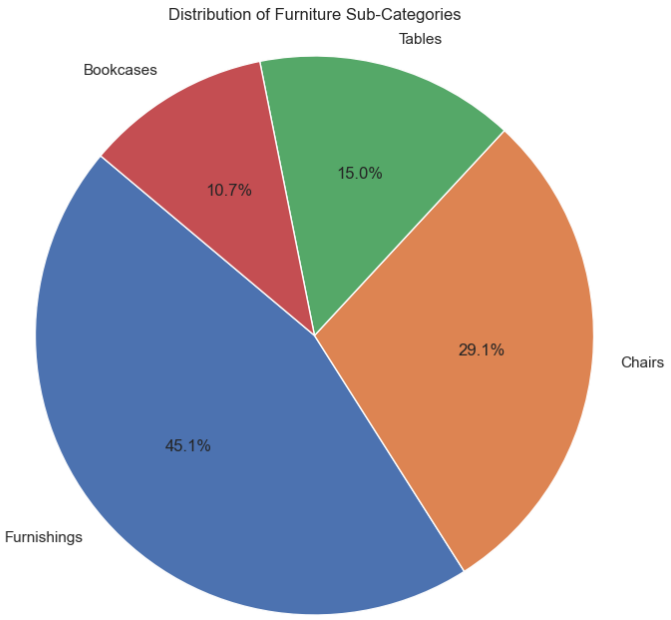
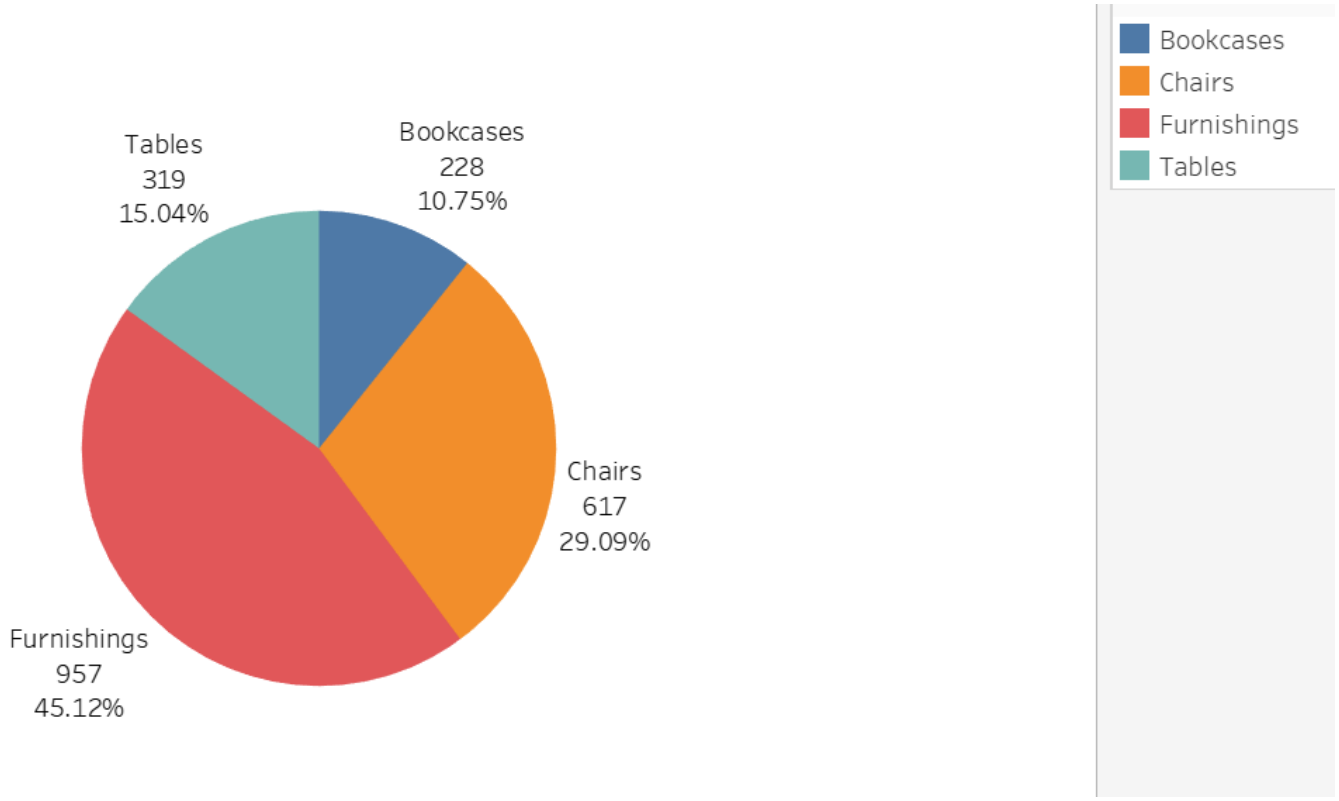


Tableau Workbook :-



8) What is the profit and sales data for each sub-category?

Code:-

```
Category_wise_Sales_and_profit = Dataset.groupby("Sub-Category")[["Profit", "Sales"]].sum()
print(Category_wise_Sales_and_profit)
ax = Category_wise_Sales_and_profit.plot(kind="line", figsize=(12, 6))
plt.title("Category-wise Sales and Profit")
plt.xlabel("Sub-Category")
plt.ylabel("Amount")
plt.grid(True)
plt.legend(loc="upper right")
for column in Category_wise_Sales_and_profit.columns:
    for index, value in enumerate(Category_wise_Sales_and_profit[column]):
        ax.text(index, value, str(round(value, 2)), ha="center", va="bottom")

plt.tight_layout()
plt.show()
```

Marwadi University Marwadi Chandarana Group NAAC A+	**Marwadi University** **Faculty of Technology** **Department of Information and Communication Technology**	
Subject: Data Visualization and Dashboards (01CT0410)	**Aim:** Analysis of Superstore Dataset	
Case Study - 1	**Date:-** 16-03-2024	**Enrollment No:-** 92200133030

Output :-

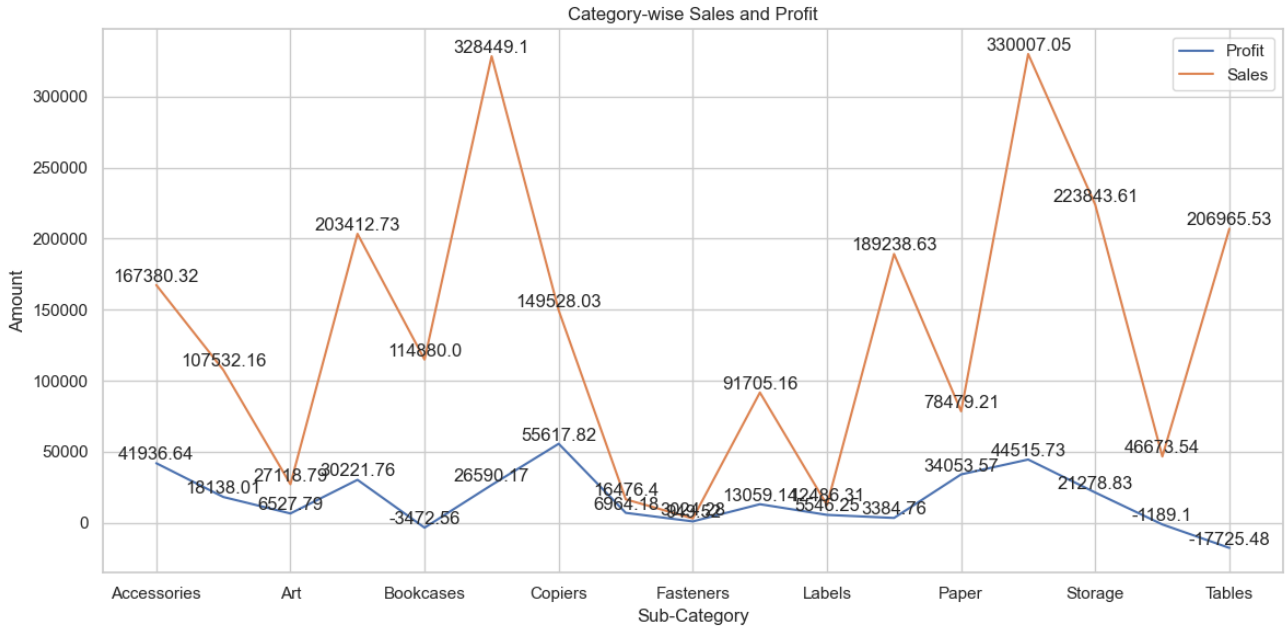
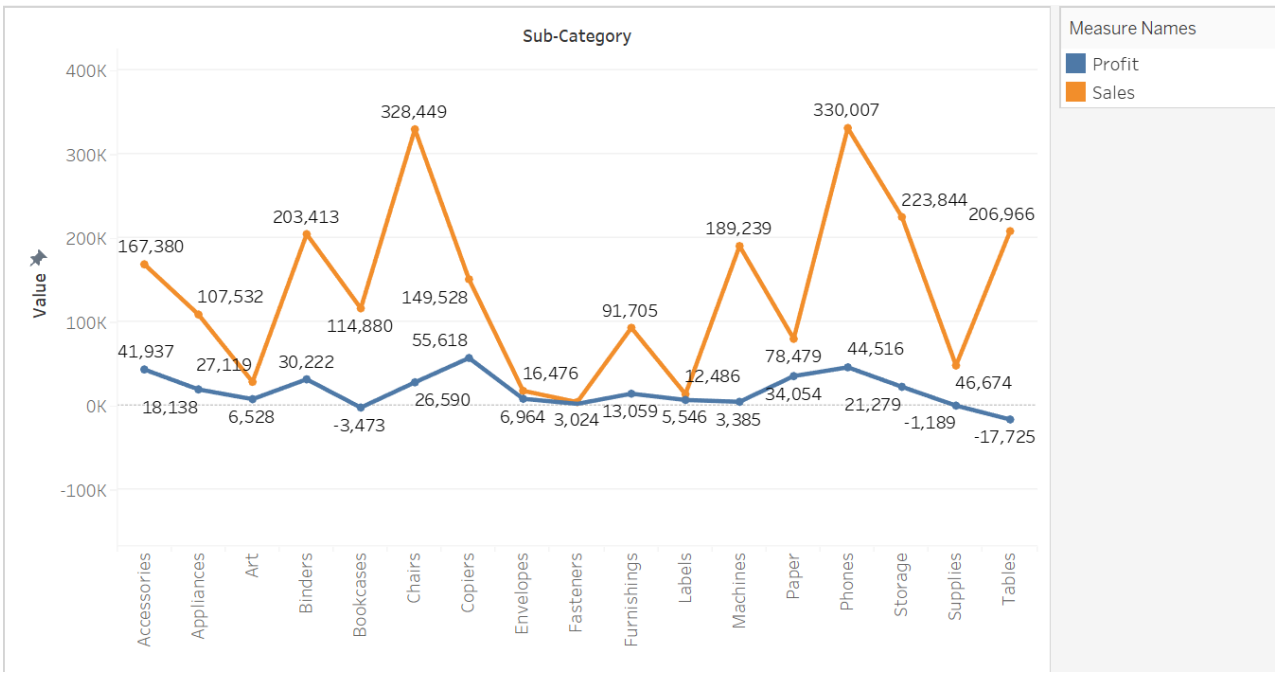


Tableau Workbook :-



9) How many unique products have been ordered in the store?

Code:-

```
Unique_Products = Dataset['Product Name'].unique()
No_of_Unique_Products = Dataset['Product Name'].nunique()
print(f"There are {No_of_Unique_Products} Unique Products :-\n{Unique_Products}")
plt.bar(
    x=["No Of Unique Products"], height=[len(Unique_Products)], width=0.5
)

for i, v in enumerate([len(Unique_Products)]):
    plt.text(i, v + 0.5, str(v), ha="center")

plt.show()
```

Output :-

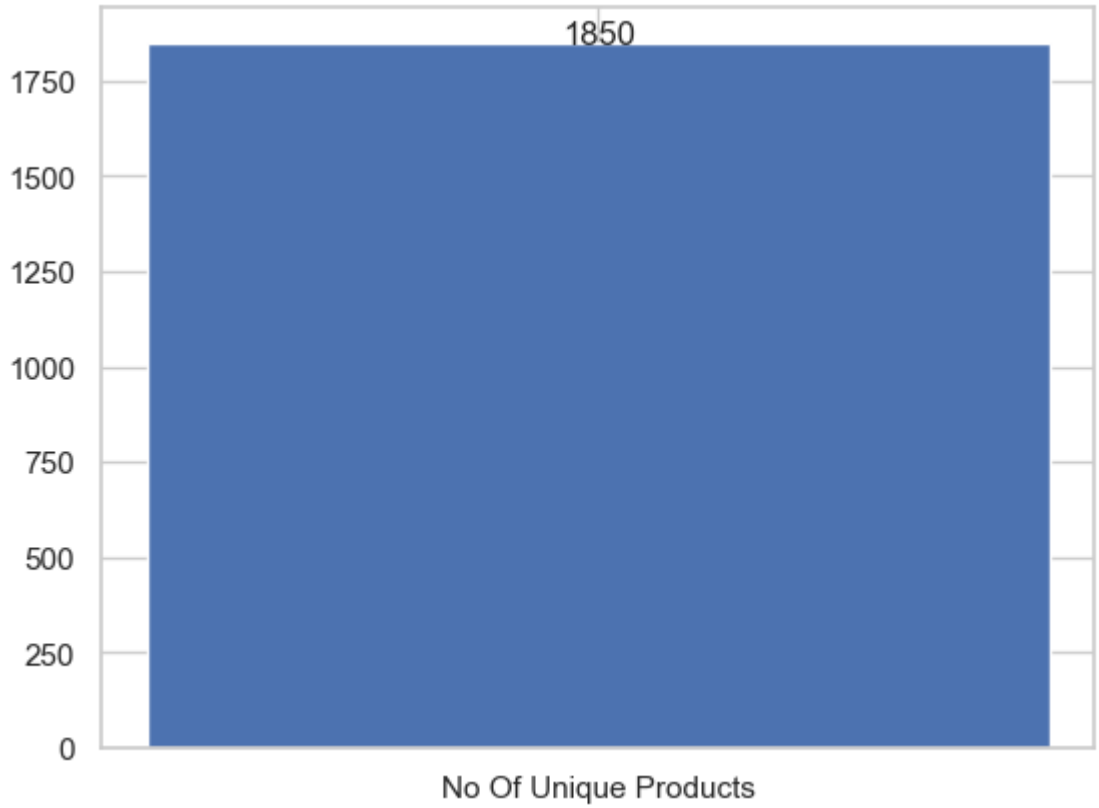
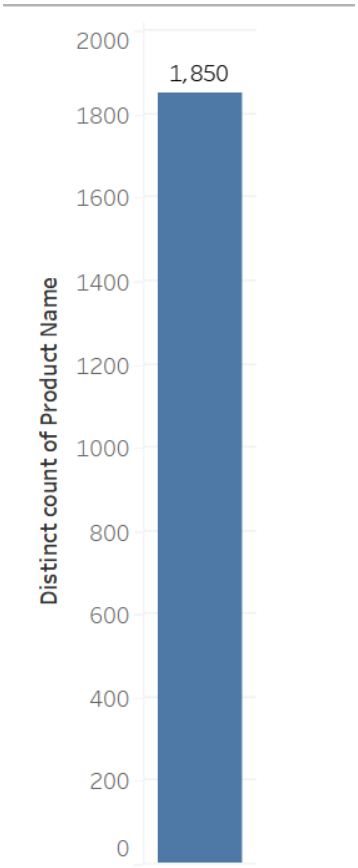


Tableau Workbook :-



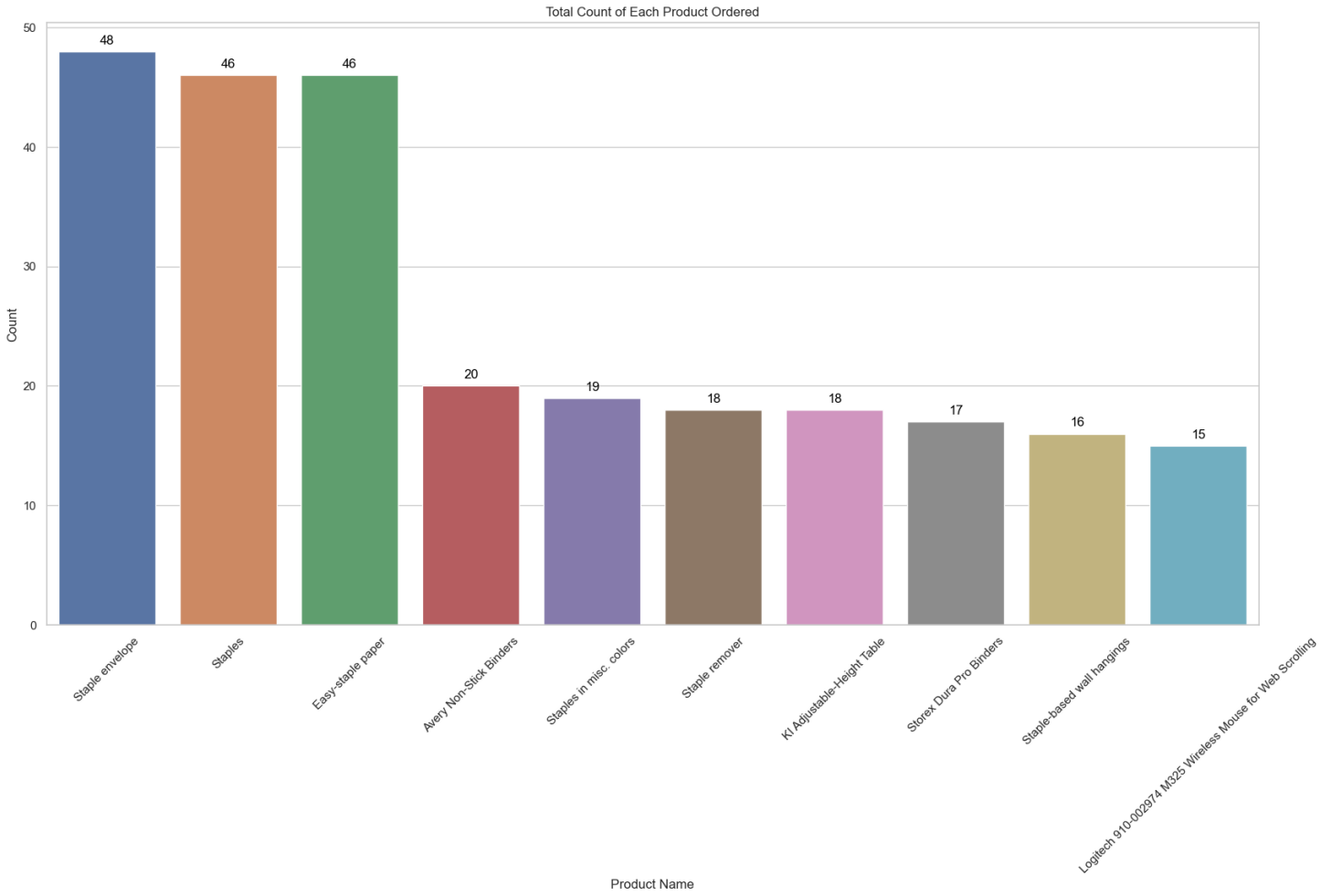
10) What is the total count of each product ordered in the store?

Code:-

```
Count_of_Each_Product = Dataset['Product Name'].value_counts()
Data = Count_of_Each_Product[:10]
Data_df = pd.DataFrame({"Product Name": Data.index, "Count": Data.values})
plt.figure(figsize=(20, 10))
ax = sns.barplot(data=Data_df, x="Product Name", y="Count", hue="Product Name")
for p in ax.patches:
    ax.annotate(
        format(p.get_height(), ".0f"),
        (p.get_x() + p.get_width() / 2.0, p.get_height()),
        ha="center",
```

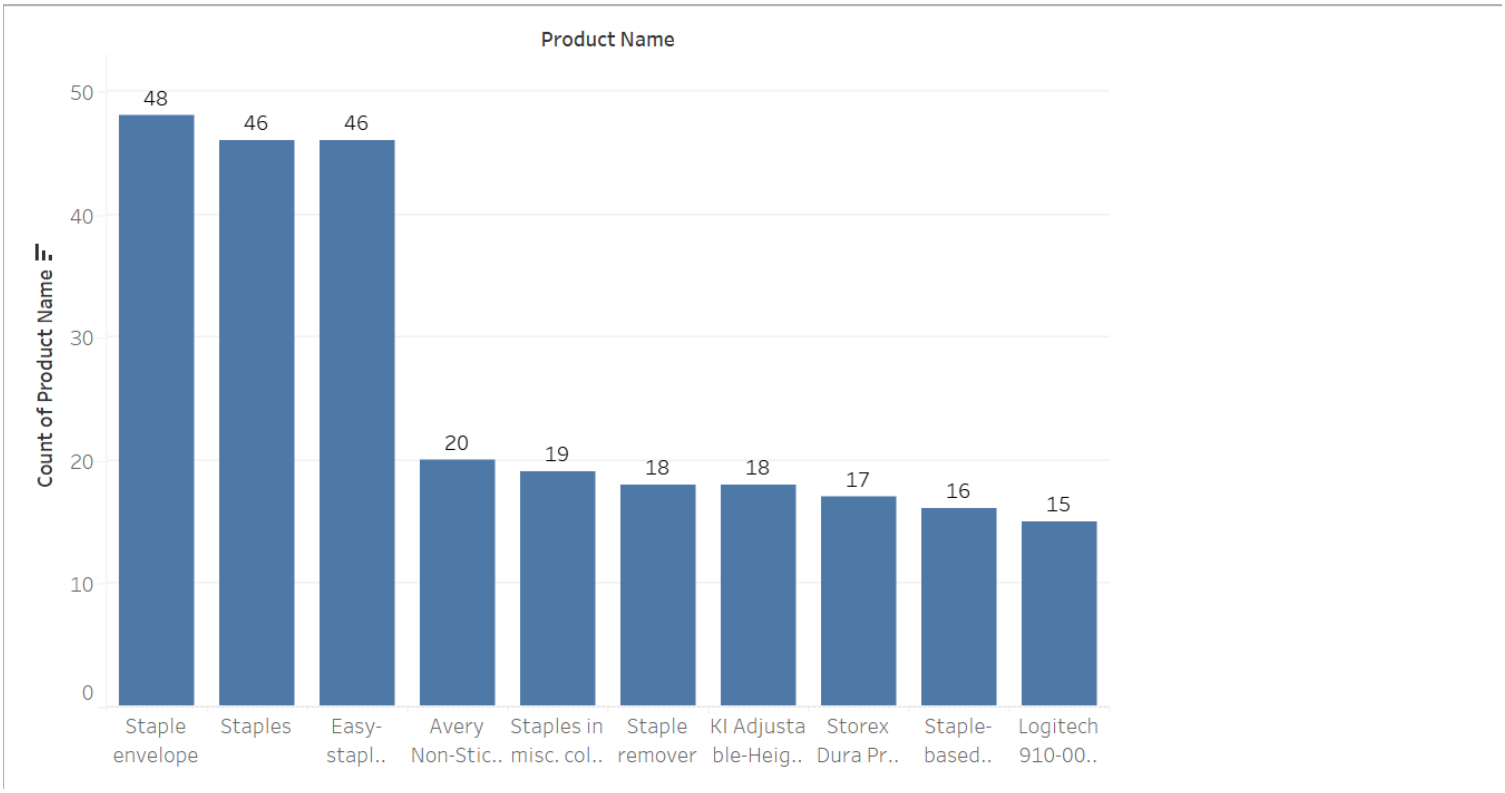
```
        va="center",
        xytext=(0, 10),
        textcoords="offset points",
        fontsize=12,
        color="black",
    )

plt.xticks(rotation=45)
plt.xlabel("Product Name")
plt.ylabel("Count")
plt.title("Total Count of Each Product Ordered")
plt.show()
```

Output :-



Marwadi University **Faculty of Technology** **Department of Information and Communication Technology**		
Subject: Data Visualization and Dashboards (01CT0410)	**Aim:** Analysis of Superstore Dataset	
Case Study - 1	**Date:-** 16-03-2024	**Enrollment No:-** 92200133030

Tableau Workbook :-



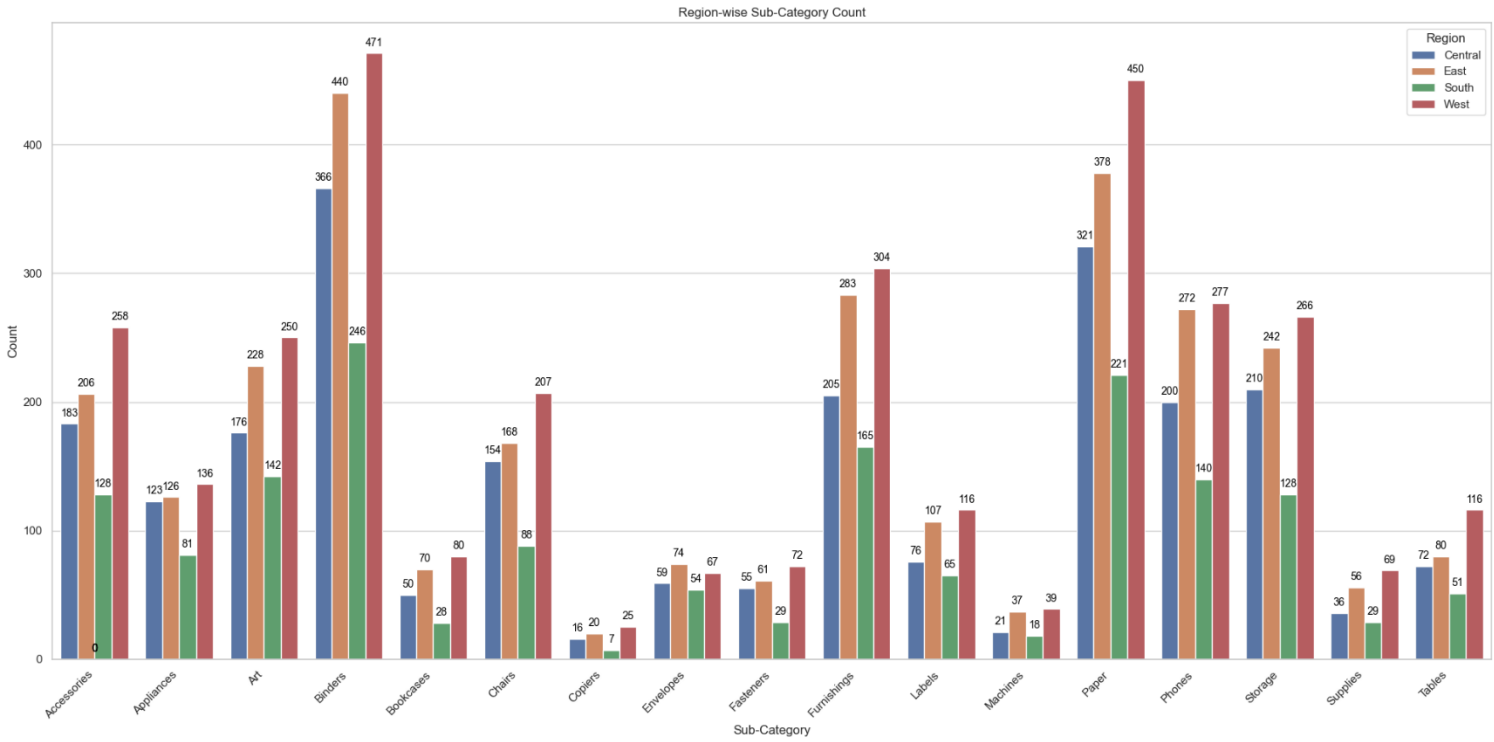
11) What is the total count of each sub-category region-wise?

Code:-

```
Region_wise_sub_category_count = Dataset.groupby(["Region", "Sub-
Category"]).size().reset_index(name="Count")
plt.figure(figsize=(20, 10))
ax = sns.barplot(
    data=Region_wise_sub_category_count, x="Sub-Category", y="Count", hue="Region"
)

for p in ax.patches:
    ax.annotate(
        format(p.get_height(), ".0f"),
        (p.get_x() + p.get_width() / 2.0, p.get_height()),
        ha="center",
```

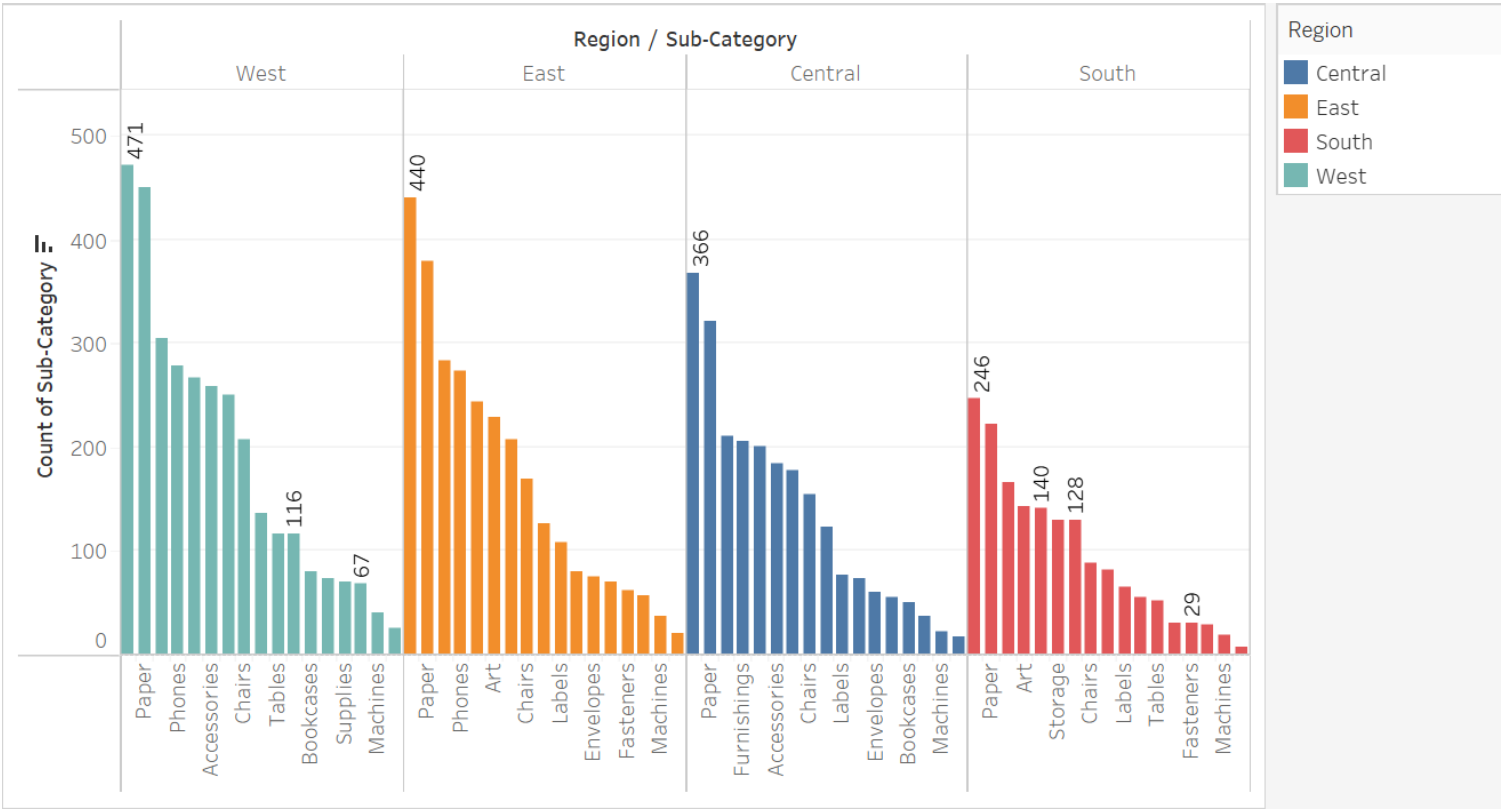
```
        va="center",
        xytext=(0, 10),
        textcoords="offset points",
        fontsize=10,
        color="black",)
plt.xticks(rotation=45, ha="right")
plt.xlabel("Sub-Category")
plt.ylabel("Count")
plt.title("Region-wise Sub-Category Count")
plt.legend(title="Region", loc="upper right")
plt.tight_layout()
plt.show()
```

Output :-



Marwadi University Marwadi Chandarana Group NAAC A+	**Marwadi University** **Faculty of Technology** **Department of Information and Communication Technology**	
Subject: Data Visualization and Dashboards (01CT0410)	**Aim:** Analysis of Superstore Dataset	
Case Study - 1	**Date:-** 16-03-2024	**Enrollment No:-** 92200133030

Tableau Workbook :-



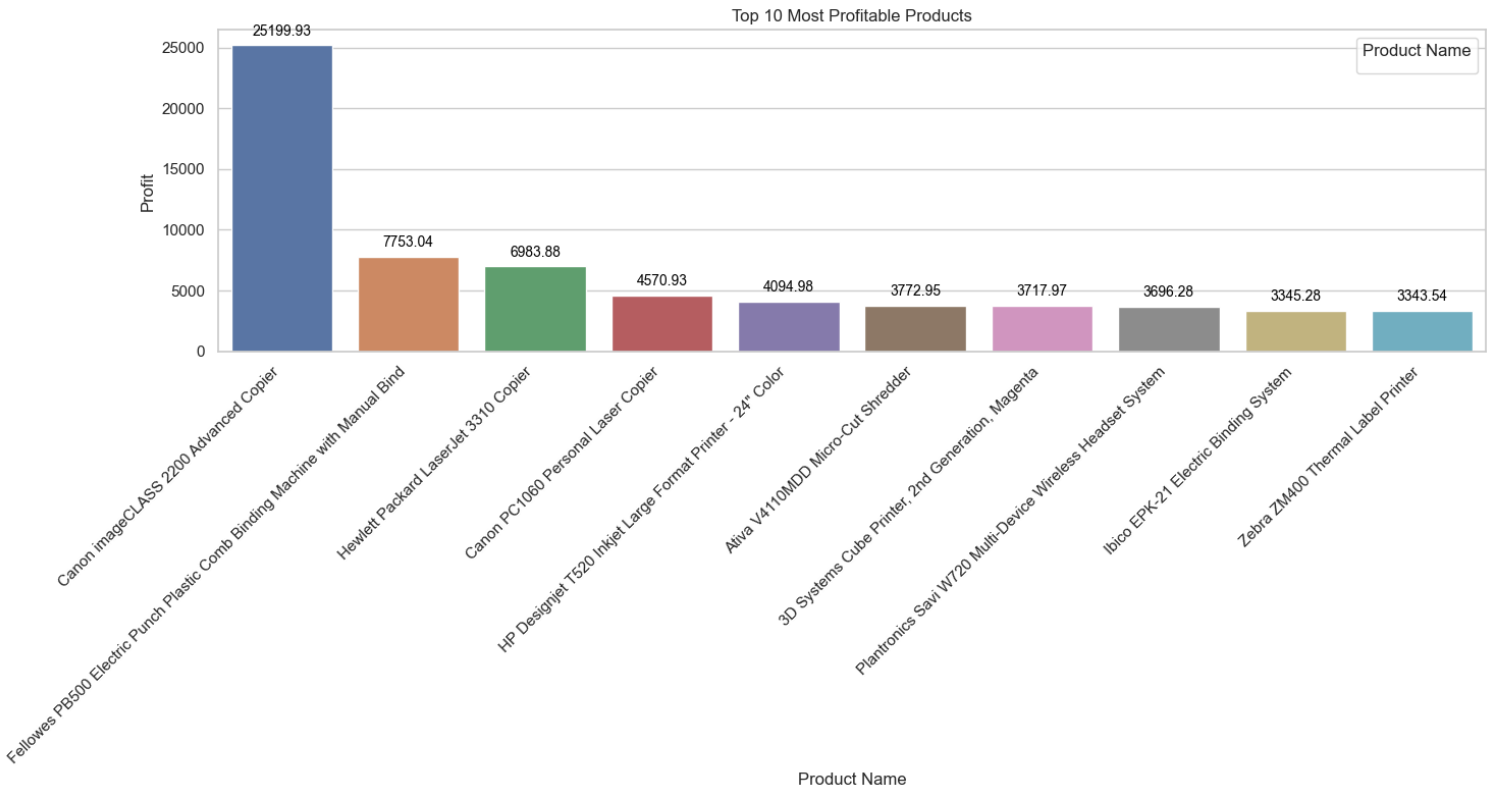
12) Which product has the highest profit?

Code:-

```
Max_Profitable_Product = (
   Dataset.groupby("Product Name")["Profit"].sum().sort_values(ascending=False)
)[:10]
Max_Profitable_Product_df = pd.DataFrame(
   {"Product Name": Max_Profitable_Product.index, "Profit": Max_Profitable_Product.values}
)
plt.figure(figsize=(15, 8))
ax = sns.barplot(
   data=Max_Profitable_Product_df, x="Product Name", y="Profit", hue="Product Name"
)
for p in ax.patches:
   ax.annotate(
      format(p.get_height(), ".2f"),
```

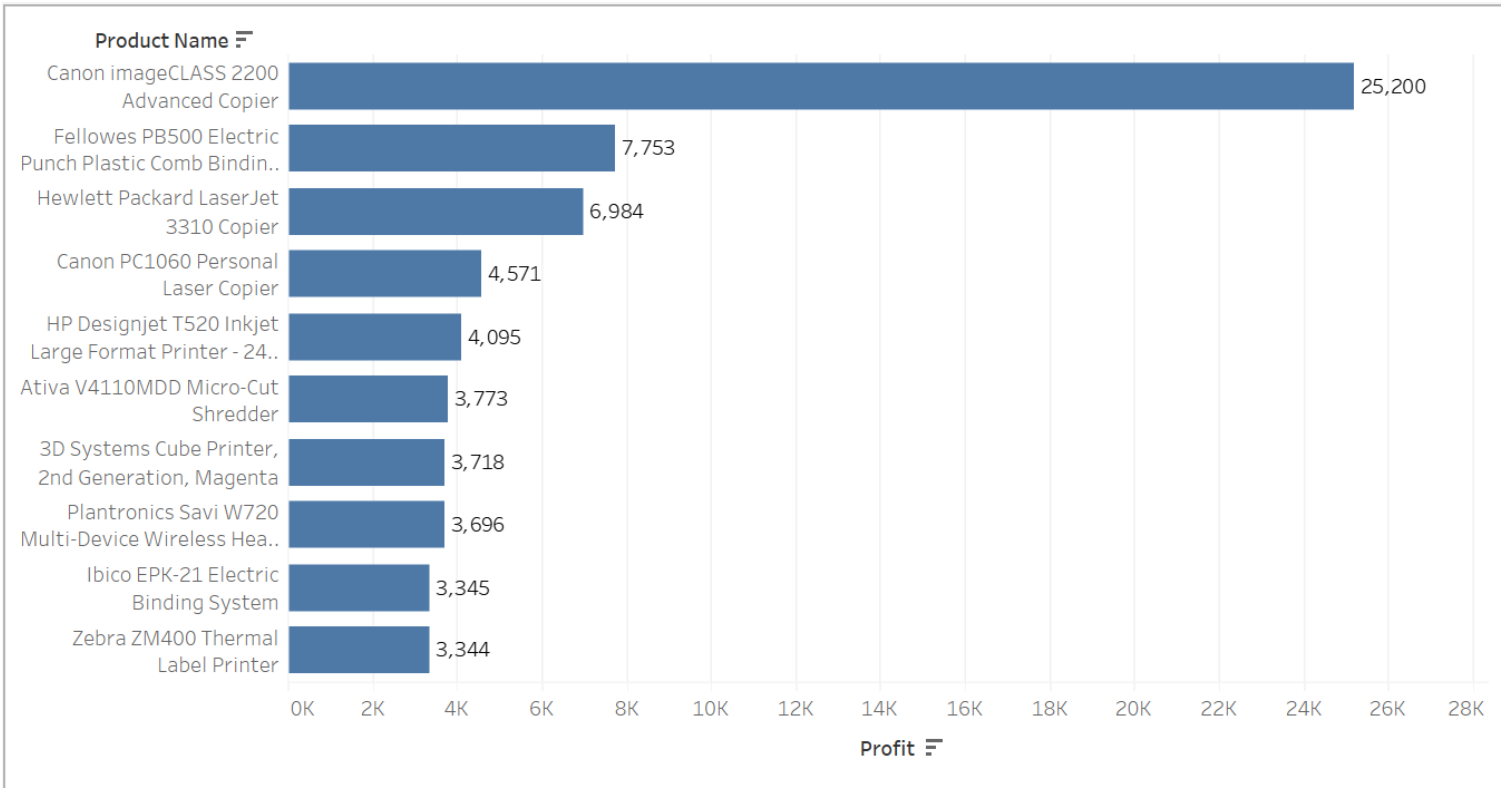
```
        (p.get_x() + p.get_width() / 2.0, p.get_height()),
        ha="center",
        va="center",
        xytext=(0, 10),
        textcoords="offset points",
        fontsize=10,
        color="black",
    )

plt.xticks(rotation=45, ha="right")
plt.xlabel("Product Name")
plt.ylabel("Profit")
plt.title("Top 10 Most Profitable Products")
plt.legend(title="Product Name", loc="upper right")
plt.tight_layout()
plt.show()
```

Output :-



Marwadi University Marwadi Chandarana Group NAAC A+	**Marwadi University** **Faculty of Technology** **Department of Information and Communication Technology**	
Subject: Data Visualization and Dashboards (01CT0410)	**Aim:** Analysis of Superstore Dataset	
Case Study - 1	**Date:-** 16-03-2024	**Enrollment No:-** 92200133030

Tableau Workbook :-



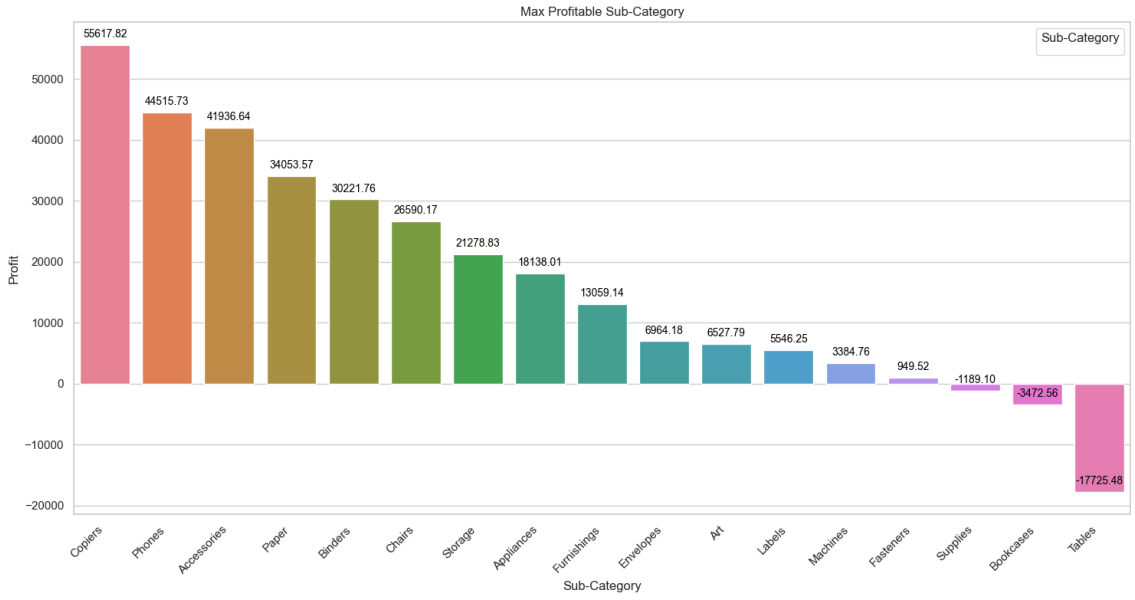
13) Which sub-category has the highest profit?

Code:-

```
Max_Profitable_Sub_Category = Dataset.groupby('Sub-
Category')['Profit'].sum().sort_values(ascending=False)
Max_Profitable_Sub_Category_df = pd.DataFrame(
    {
        "Sub-Category": Max_Profitable_Sub_Category.index,
        "Profit": Max_Profitable_Sub_Category.values,
    }
)

plt.figure(figsize=(15, 8))
ax = sns.barplot(
    data=Max_Profitable_Sub_Category_df,
    x="Sub-Category",
    y="Profit",
```

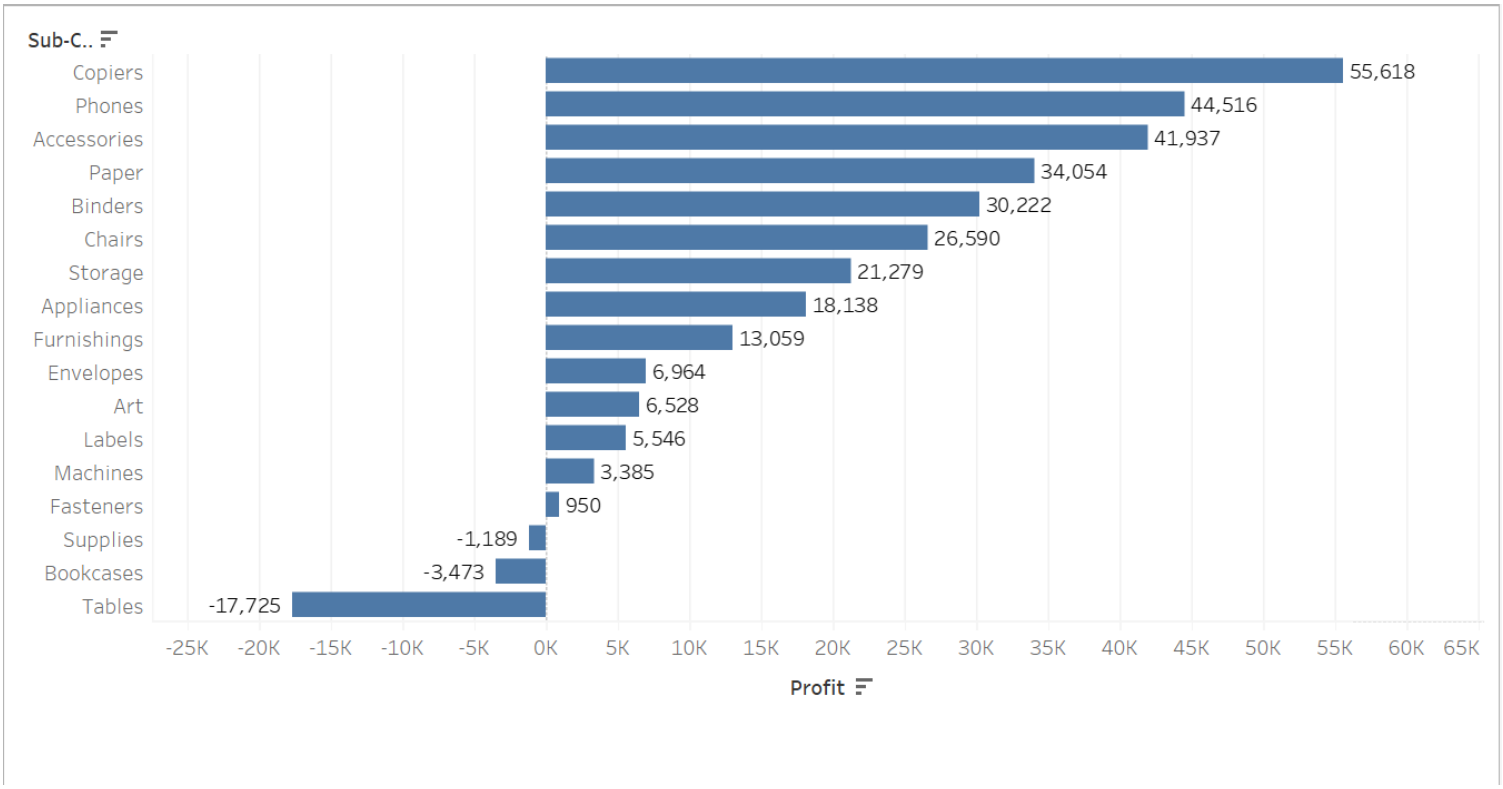
```
    hue="Sub-Category",
)

for p in ax.patches:
    ax.annotate(
        format(p.get_height(), ".2f"),
        (p.get_x() + p.get_width() / 2.0, p.get_height()),
        ha="center",
        va="center",
        xytext=(0, 10),
        textcoords="offset points",
        fontsize=10,
        color="black",
    )

plt.xticks(rotation=45, ha="right")
plt.xlabel("Sub-Category")
plt.ylabel("Profit")
plt.title("Max Profitable Sub-Category")
plt.legend(title="Sub-Category", loc="upper right")
plt.tight_layout()
plt.show()
```

Output :-



Marwadi University Marwadi Chandarana Group NAAC A+	**Marwadi University** **Faculty of Technology** **Department of Information and Communication Technology**	
Subject: Data Visualization and Dashboards (01CT0410)	**Aim:** Analysis of Superstore Dataset	
Case Study - 1	**Date:-** 16-03-2024	**Enrollment No:-** 92200133030

Tableau Workbook :-



14) Which product has the highest loss?

Code:-

```
Max_Loss_Product = Dataset.groupby("Product Name")["Profit"].sum().sort_values(ascending=True)[:10]
Max_Loss_Product_df = pd.DataFrame(
    {
        "Product Name": Max_Loss_Product.index,
        "Profit": Max_Loss_Product.values,
    }
)

plt.figure(figsize=(15, 8))
ax = sns.barplot(
    data=Max_Loss_Product_df, x="Product Name", y="Profit", hue="Product Name"
)

for p in ax.patches:
```

```
    ax.annotate(
        format(p.get_height(), ".2f"),
        (p.get_x() + p.get_width() / 2.0, p.get_height()),
        ha="center",
        va="center",
        xytext=(0, 10),
        textcoords="offset points",
        fontsize=10,
        color="black",
    )

plt.xticks(rotation=45, ha="right")
plt.xlabel("Product Name")
plt.ylabel("Profit")
plt.title("Top 10 Most Loss Making Products")
plt.legend(title="Product Name", loc="upper right")
plt.tight_layout()
plt.show()
```

Output :-

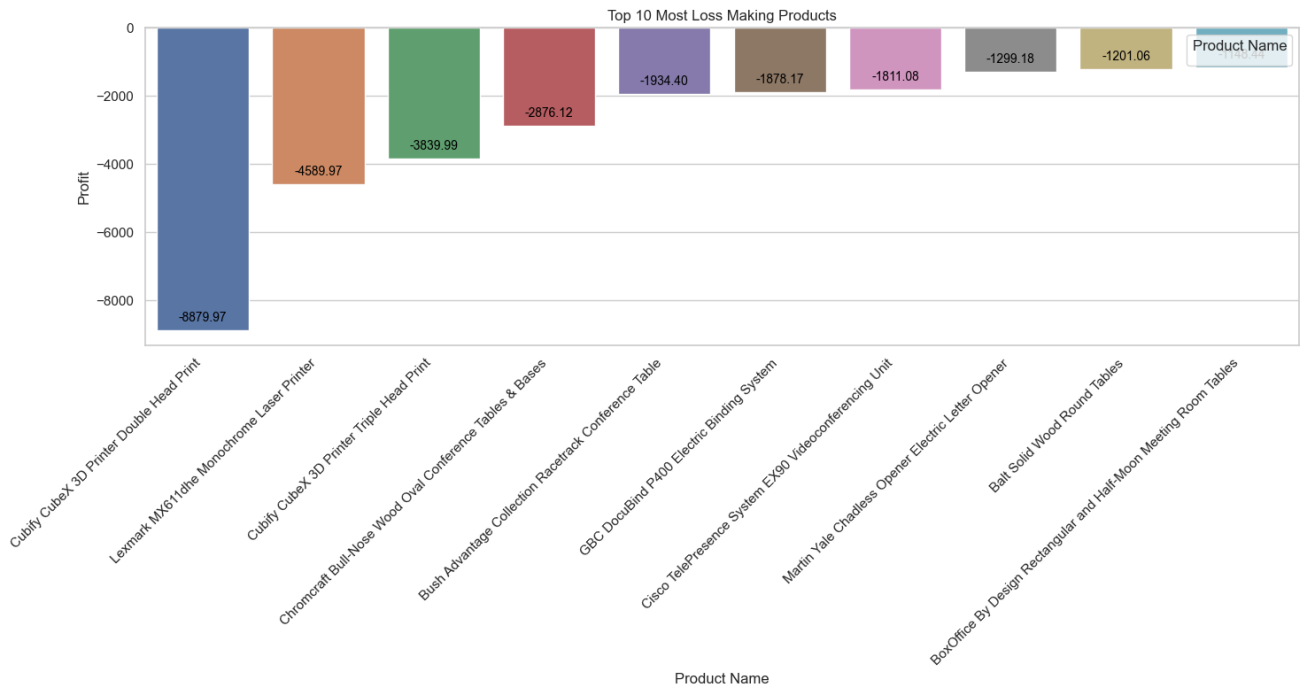
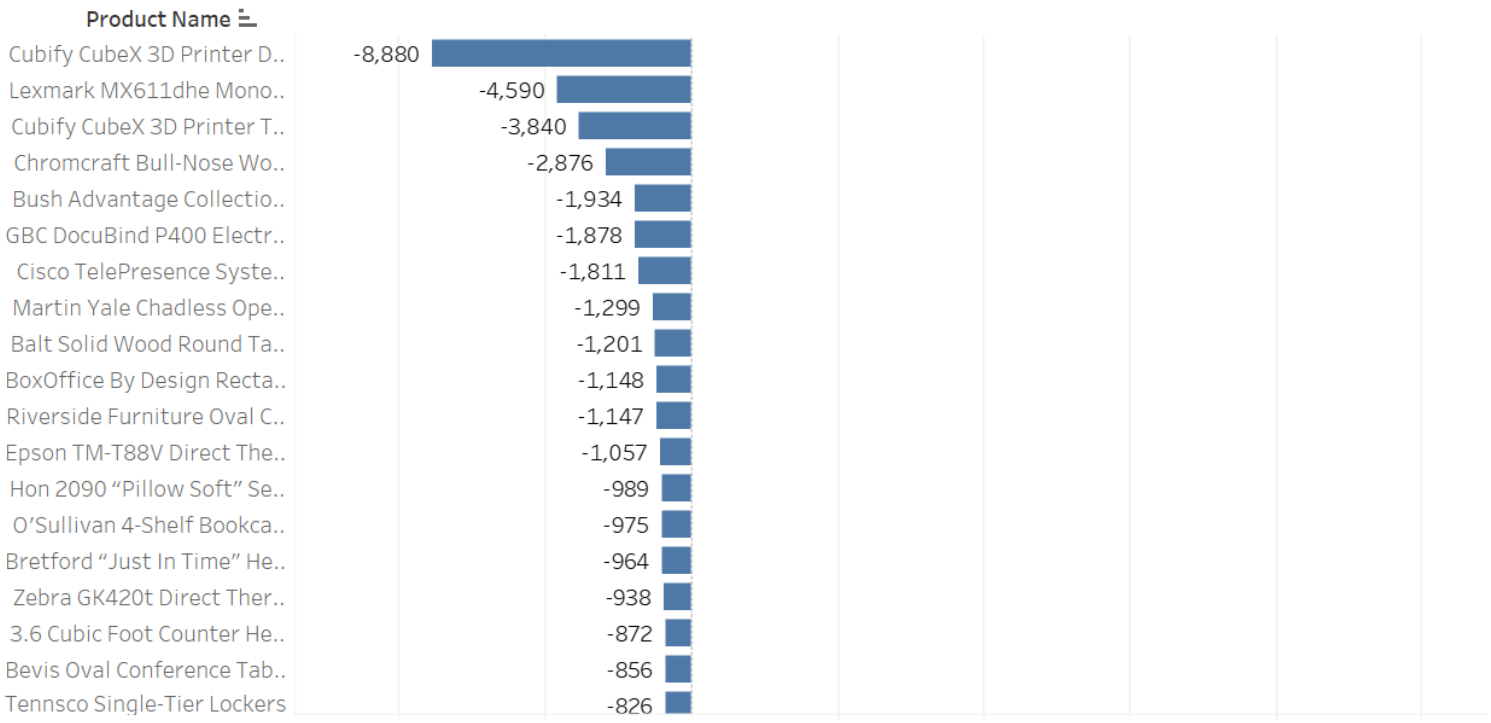


Tableau Workbook :-



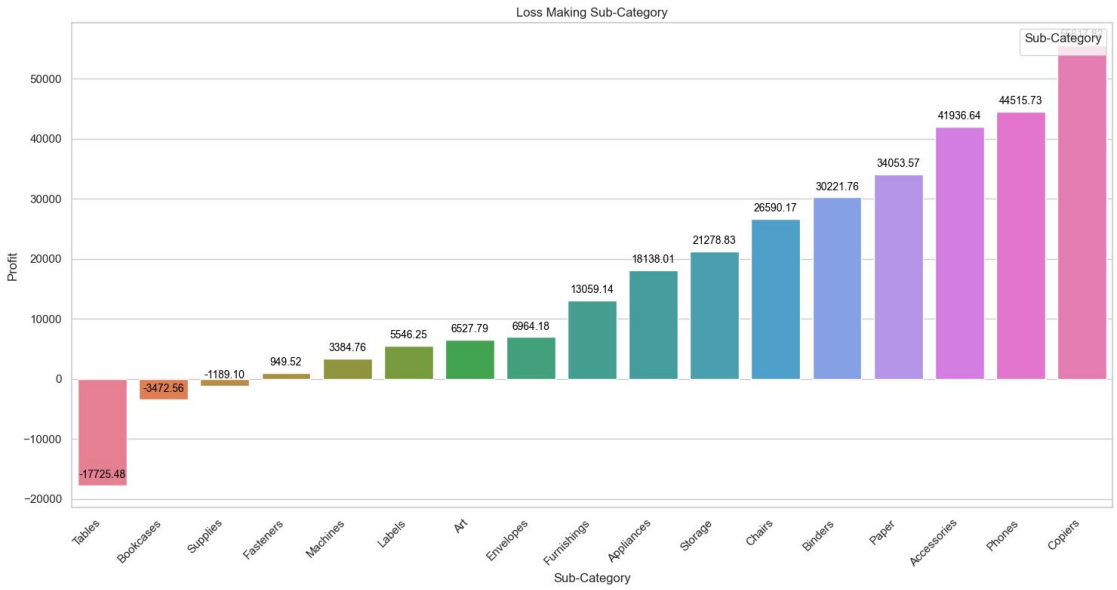
15) Which sub-category has the highest loss?

Code:-

```
Max_Loss_Sub_Category = Dataset.groupby("Sub-Category")["Profit"].sum().sort_values(ascending=True)
Max_Loss_Sub_Category_df = pd.DataFrame(
    {
        "Sub-Category": Max_Loss_Sub_Category.index,
        "Profit": Max_Loss_Sub_Category.values,
    }
)

plt.figure(figsize=(15, 8))
ax = sns.barplot(
    data=Max_Loss_Sub_Category_df,
    x="Sub-Category",
    y="Profit",
    hue="Sub-Category",
)
```

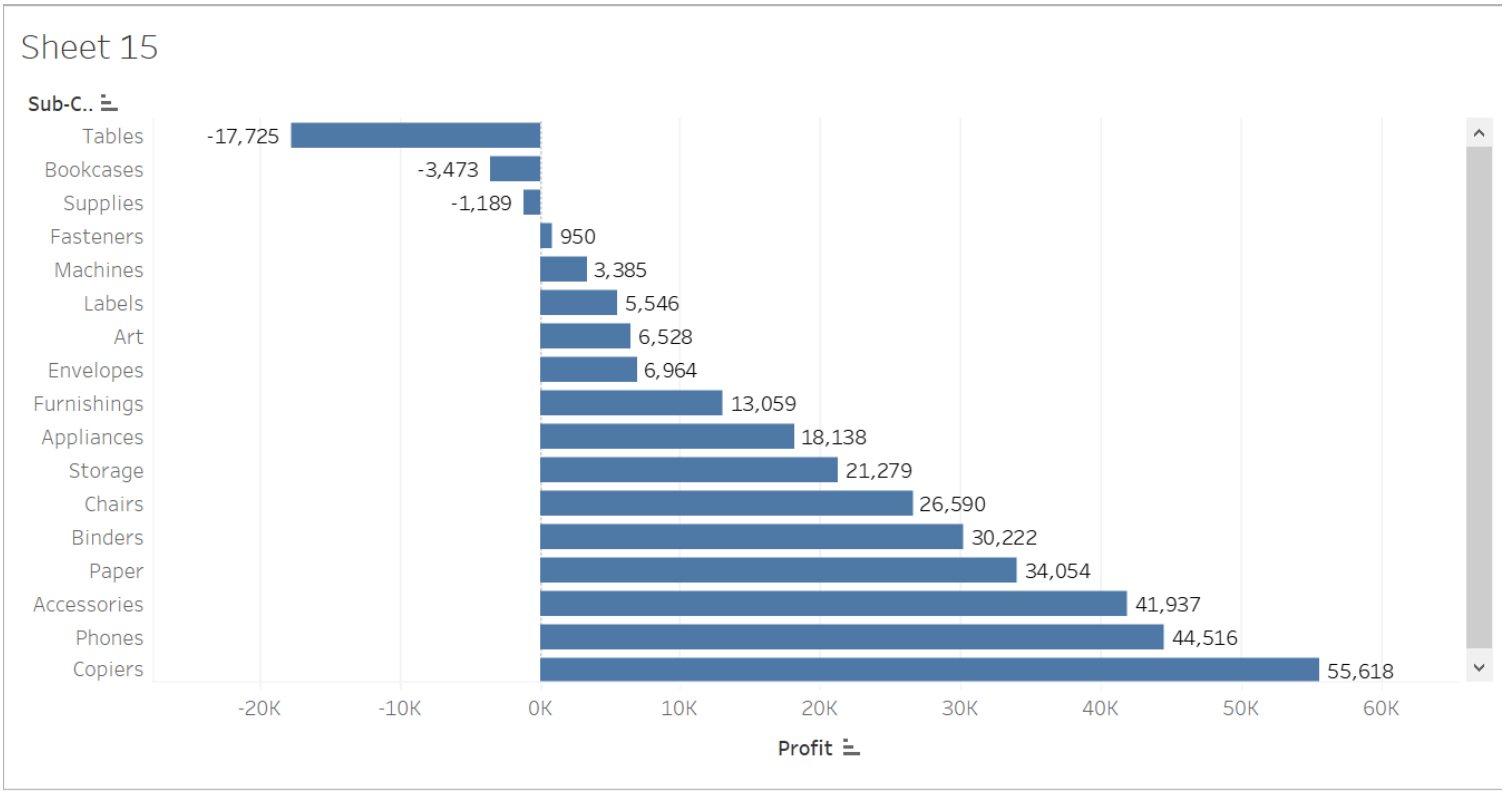
```
for p in ax.patches:
    ax.annotate(
        format(p.get_height(), ".2f"),
        (p.get_x() + p.get_width() / 2.0, p.get_height()),
        ha="center",
        va="center",
        xytext=(0, 10),
        textcoords="offset points",
        fontsize=10,
        color="black",
    )

plt.xticks(rotation=45, ha="right")
plt.xlabel("Sub-Category")
plt.ylabel("Profit")
plt.title("Loss Making Sub-Category")
plt.legend(title="Sub-Category", loc="upper right")
plt.tight_layout()
plt.show()
```

Output :-



Marwadi University Marwadi Chandarana Group NAAC A+	**Marwadi University** **Faculty of Technology** **Department of Information and Communication Technology**	
Subject: Data Visualization and Dashboards (01CT0410)	**Aim:** Analysis of Superstore Dataset	
Case Study - 1	**Date:-** 16-03-2024	**Enrollment No:-** 92200133030

Tableau Workbook :-



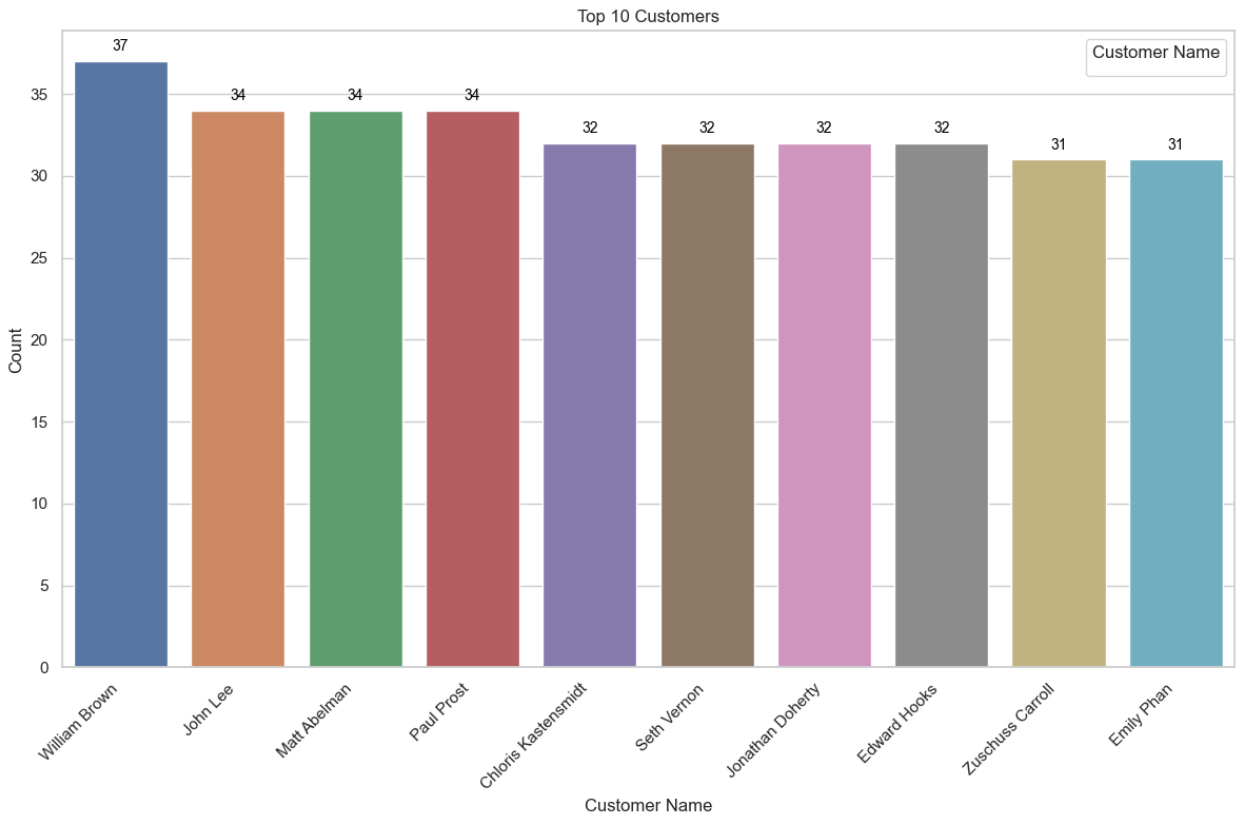
16) Who are the top 10 customers who order frequently?

Code:-

```
Top_10_Customers = Dataset['Customer Name'].value_counts()
Top_10_Customers_df = pd.DataFrame({
    "Customer Name": Top_10_Customers.index,
    "Count": Top_10_Customers.values
})[:10]
plt.figure(figsize=(12, 8))
ax = sns.barplot(
    data=Top_10_Customers_df, x="Customer Name", y="Count", hue="Customer Name"
)

for p in ax.patches:
    ax.annotate(
        format(p.get_height(), ".0f"),
```

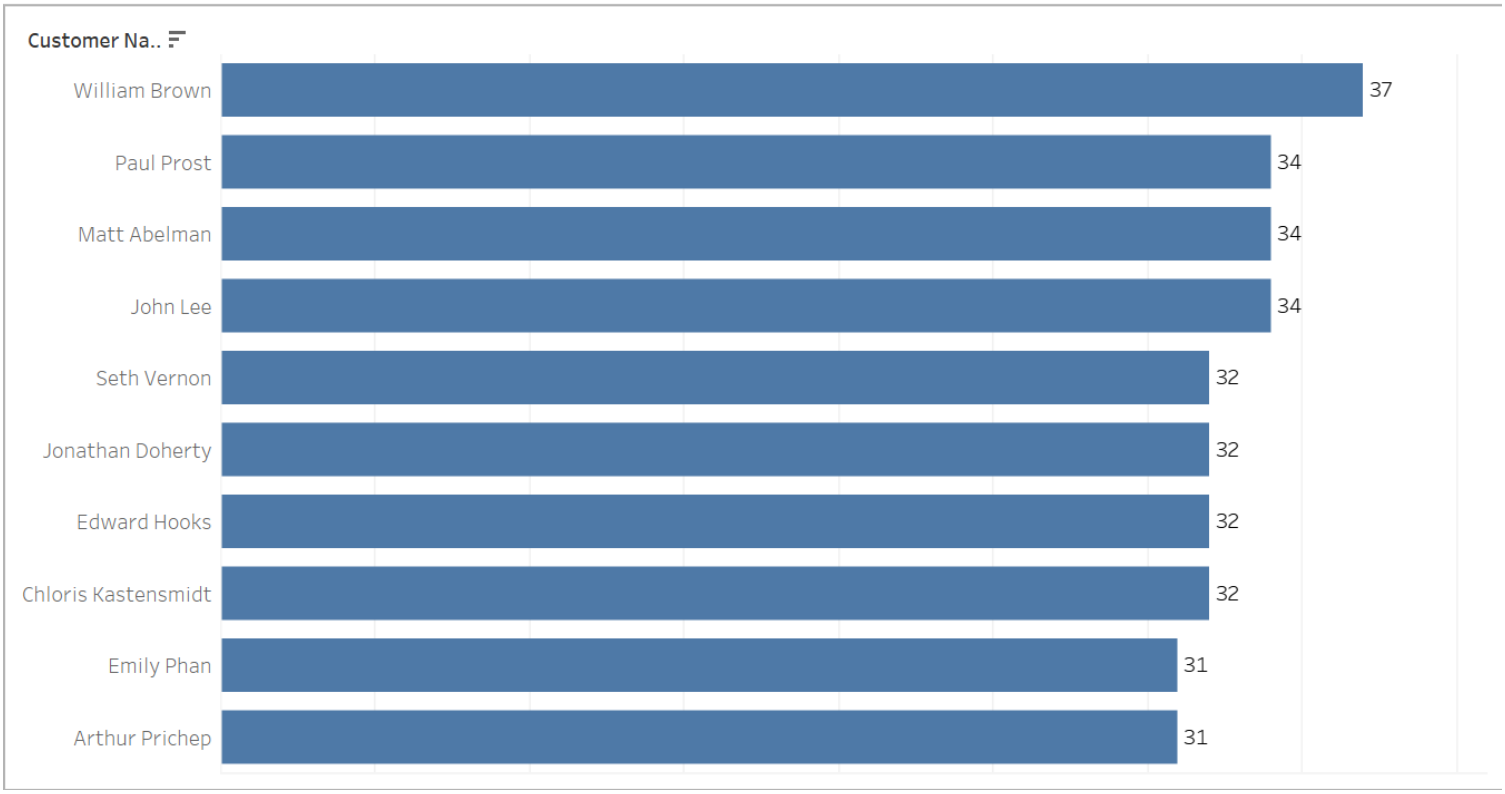
```
        (p.get_x() + p.get_width() / 2.0, p.get_height()),
        ha="center",
        va="center",
        xytext=(0, 10),
        textcoords="offset points",
        fontsize=10,
        color="black",
    )

plt.xticks(rotation=45, ha="right")
plt.xlabel("Customer Name")
plt.ylabel("Count")
plt.title("Top 10 Customers")
plt.legend(title="Customer Name", loc="upper right")
plt.tight_layout()
plt.show()
```

Output :-



Marwadi University Faculty of Technology Department of Information and Communication Technology		
Subject: Data Visualization and Dashboards (01CT0410)	**Aim:** Analysis of Superstore Dataset	
Case Study - 1	**Date:-** 16-03-2024	**Enrollment No:-** 92200133030

Tableau Workbook :-



17) How many unique customers are there in total?

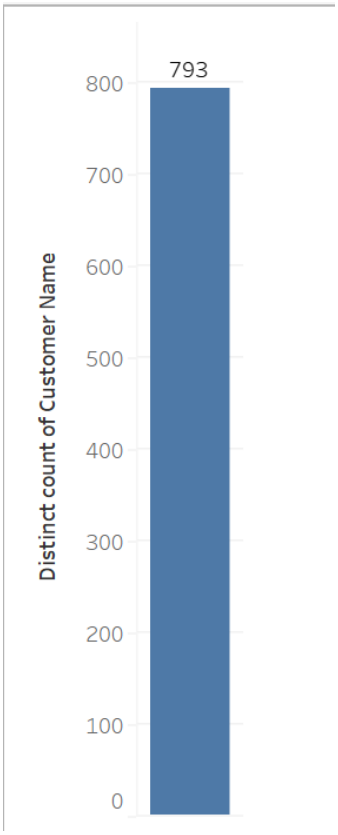
Code:-

```
unique_customers_count = Dataset["Customer Name"].nunique()
plt.figure(figsize=(8, 6))
bars = plt.bar(["Unique Customers"], [unique_customers_count], color="skyblue")
plt.xlabel("Category")
plt.ylabel("Count")
plt.title("Total Number of Unique Customers")

for bar in bars:
    yval = bar.get_height()
    plt.text(
        bar.get_x() + bar.get_width() / 2,
        yval,
        round(yval, 2),
        va="bottom",
```

```
    ha="center",
)

plt.xticks(rotation=45, ha="right")
plt.tight_layout()
plt.show()
```

Output :-

Tableau Workbook :-



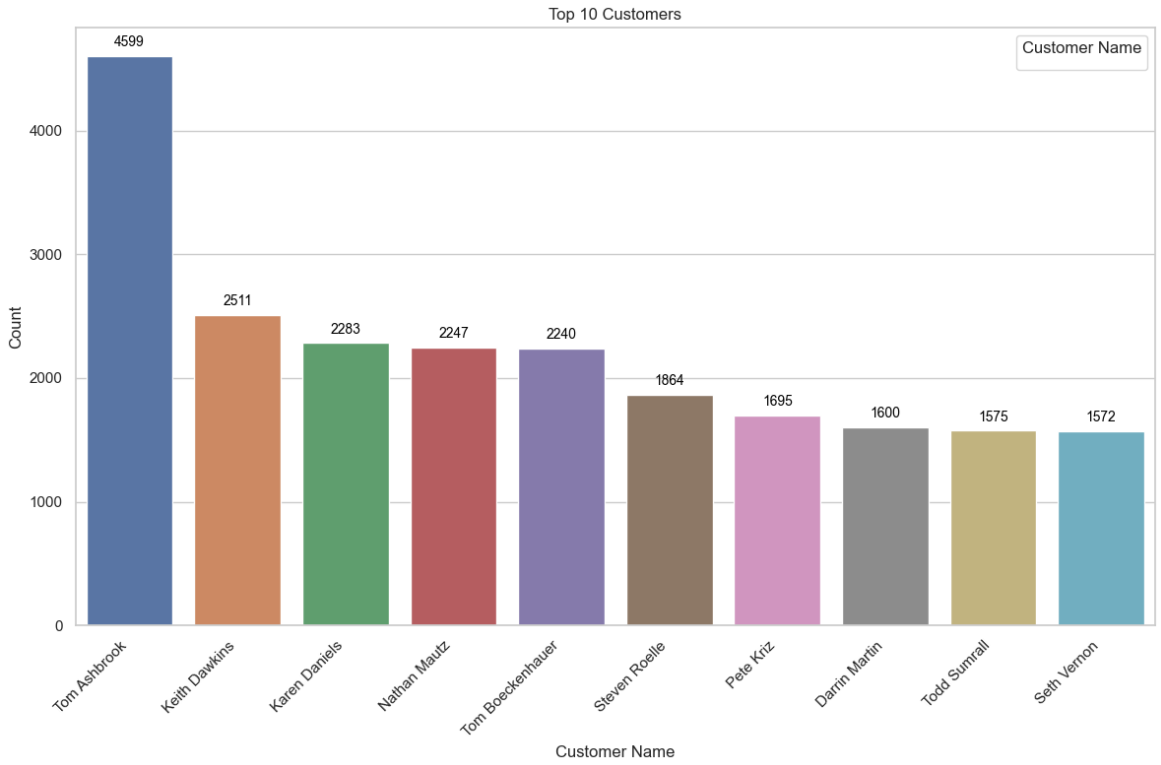
18) Who are the top 10 profitable customers in New York?

Code:-

```
New_York_Customers = Dataset[Dataset["State"] == "New York"]
Profitable_Customers = New_York_Customers.groupby("Customer Name")["Profit"].sum()
Top_10_Customers = pd.DataFrame(
    {"Customer Name": Profitable_Customers.index, "Profit": Profitable_Customers.values}
).sort_values(by="Profit", ascending=False)[:10]

plt.figure(figsize=(12, 8))
ax = sns.barplot(
    data=Top_10_Customers, x="Customer Name", y="Profit", hue="Customer Name"
)

for p in ax.patches:
    ax.annotate(
```

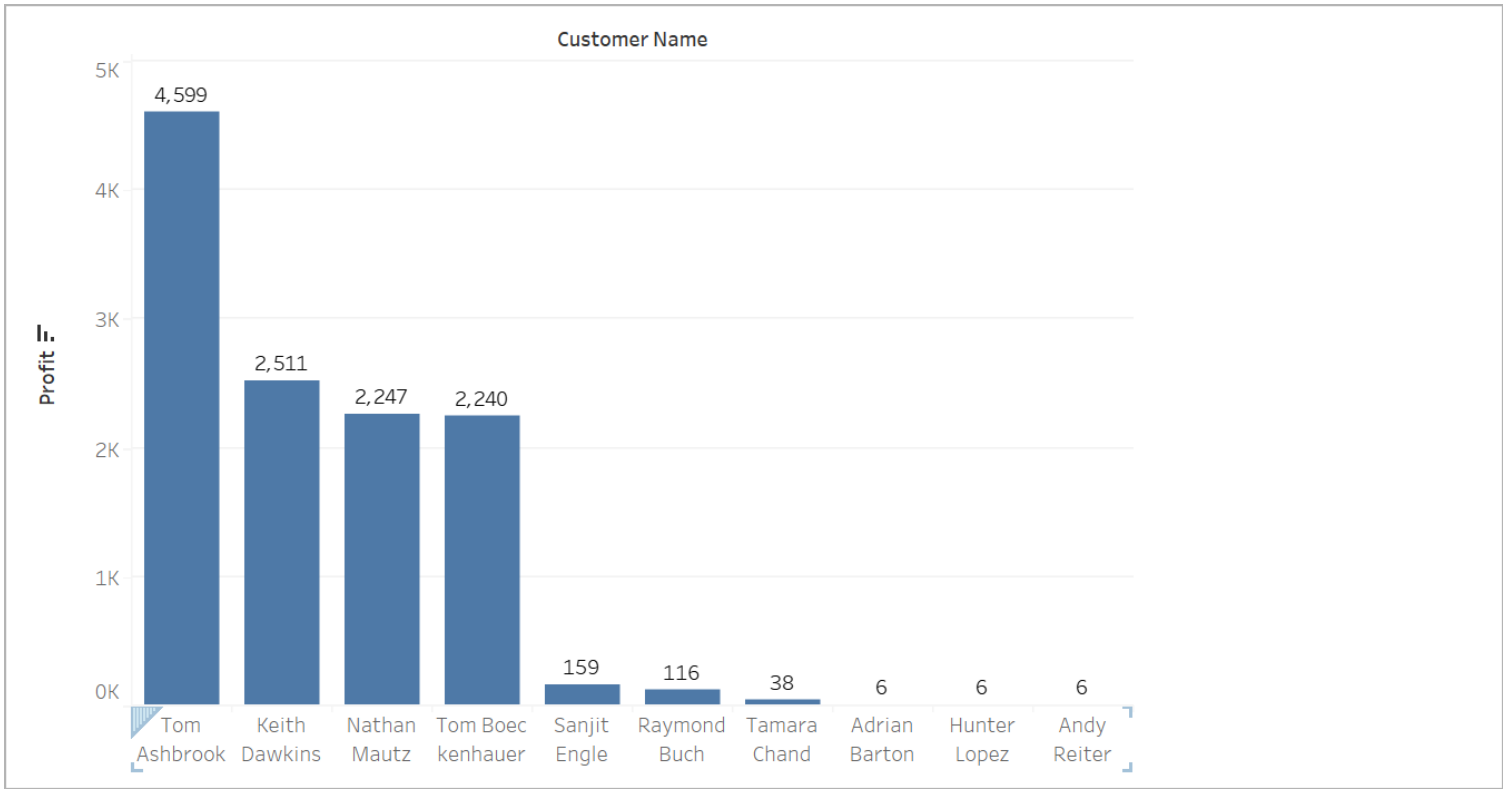
```
    format(p.get_height(), ".0f"),
    (p.get_x() + p.get_width() / 2.0, p.get_height()),
    ha="center",
    va="center",
    xytext=(0, 10),
    textcoords="offset points",
    fontsize=10,
    color="black",
  )

plt.xticks(rotation=45, ha="right")
plt.xlabel("Customer Name")
plt.ylabel("Count")
plt.title("Top 10 Customers")
plt.legend(title="Customer Name", loc="upper right")
plt.tight_layout()
plt.show()
```

Output :-



Marwadi University Marwadi Chandarana Group NAAC A+	**Marwadi University** **Faculty of Technology** **Department of Information and Communication Technology**	
Subject: Data Visualization and Dashboards (01CT0410)	**Aim:** Analysis of Superstore Dataset	
Case Study - 1	**Date:-** 16-03-2024	**Enrollment No:-** 92200133030

Tableau Workbook :-



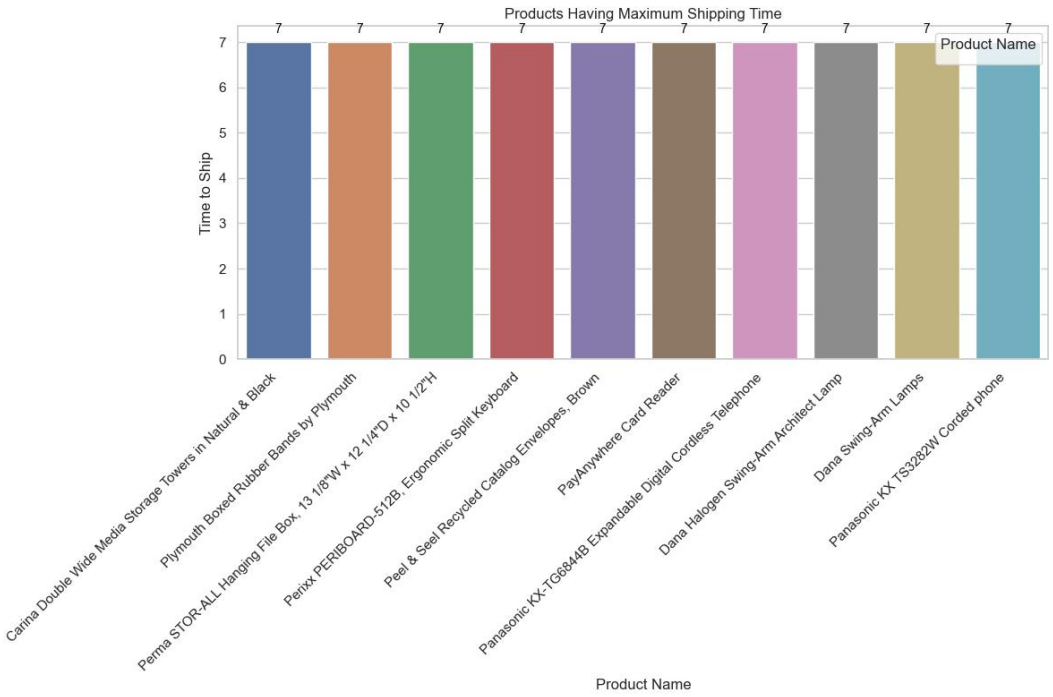
19) Which product has the maximum time taken to ship?

Code:-

```
Dataset['Time_to_Ship'] = (Dataset['Ship Date'] - Dataset['Order Date']).dt.days
Product_wies_Ship_Day = Dataset.groupby("Product Name")["Time_to_Ship"].max()
Product_wies_Ship_Day = pd.DataFrame(
    {
        "Product Name": Product_wies_Ship_Day.index,
        "Time to Ship": Product_wies_Ship_Day.values,
    }
).sort_values(by="Time to Ship", ascending=False)[:10]

plt.figure(figsize=(12, 8))
ax = sns.barplot(
    data=Product_wies_Ship_Day, x="Product Name", y="Time to Ship", hue="Product Name"
)
```

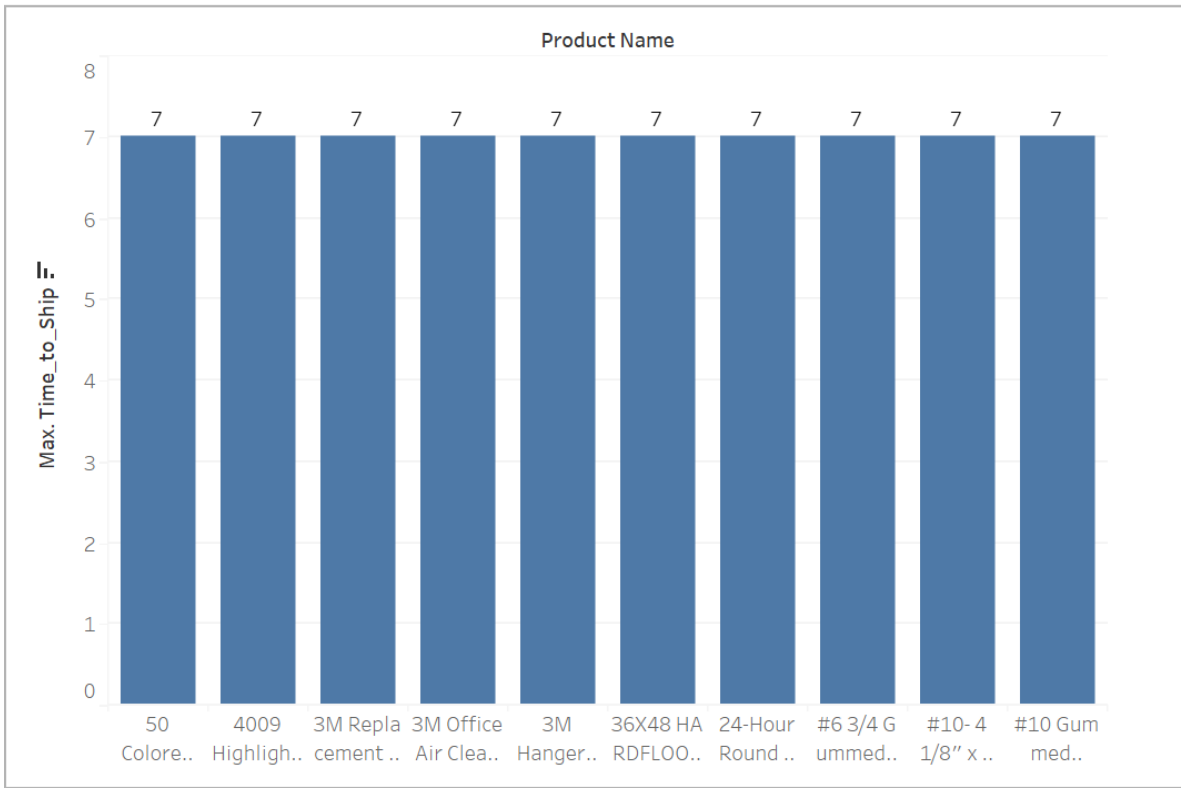
```
for p in ax.patches:
    ax.annotate(
        format(p.get_height(), ".0f"),
        (p.get_x() + p.get_width() / 2.0, p.get_height()),
        ha="center",
        va="center",
        xytext=(0, 10),
        textcoords="offset points",
        fontsize=10,
        color="black",
    )

plt.xticks(rotation=45, ha="right")
plt.xlabel("Product Name")
plt.ylabel("Time to Ship")
plt.title("Products Having Maximum Shipping Time")
plt.legend(title="Product Name", loc="upper right")
plt.tight_layout()
plt.show()
```

Output :-



Marwadi University Marwadi Chandarana Group NAAC A+	**Marwadi University** **Faculty of Technology** **Department of Information and Communication Technology**	
Subject: Data Visualization and Dashboards (01CT0410)	**Aim:** Analysis of Superstore Dataset	
Case Study - 1	**Date:-** 16-03-2024	**Enrollment No:-** 92200133030

Tableau Workbook :-



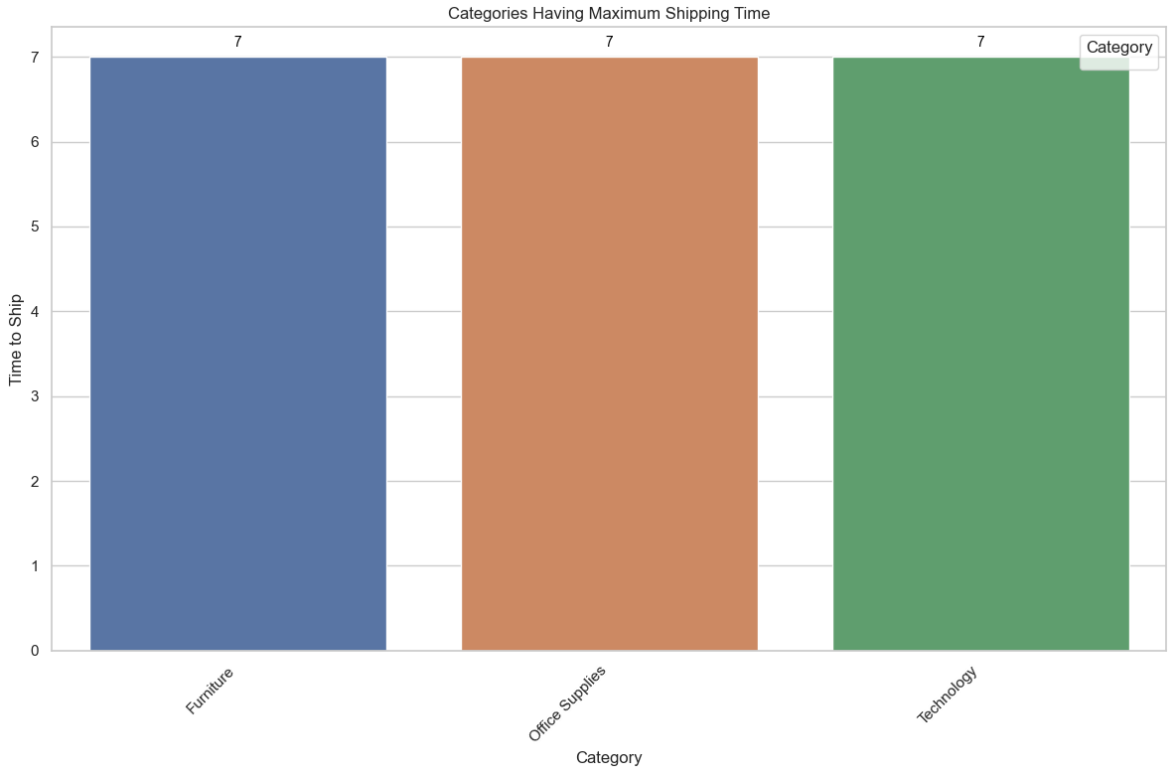
20) Which category has the maximum time taken to ship?

Code:-

```
Dataset['Time_to_Ship'] = (Dataset['Ship Date'] - Dataset['Order Date']).dt.days
Category_Ship_Wise_Groupping = Dataset.groupby('Category')['Time_to_Ship'].max()
Category_Ship_Wise_Groupping_df = pd.DataFrame(
    {
        "Category": Category_Ship_Wise_Groupping.index,
        "Time to Ship": Category_Ship_Wise_Groupping.values,
    }
).sort_values(by="Time to Ship", ascending=False)[:10]

plt.figure(figsize=(12, 8))
ax = sns.barplot(
    data=Category_Ship_Wise_Groupping_df, x="Category", y="Time to Ship", hue="Category"
)
```

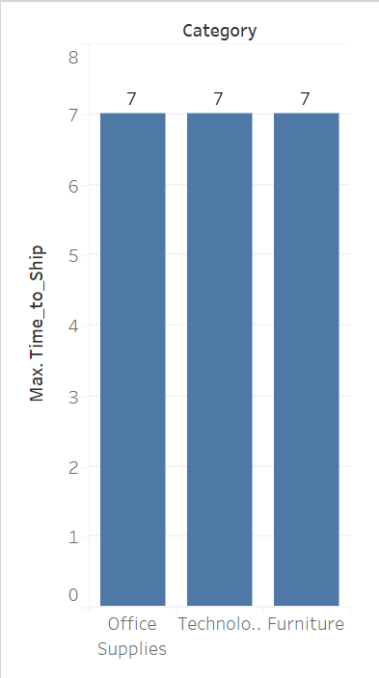
```
for p in ax.patches:
    ax.annotate(
        format(p.get_height(), ".0f"),
        (p.get_x() + p.get_width() / 2.0, p.get_height()),
        ha="center",
        va="center",
        xytext=(0, 10),
        textcoords="offset points",
        fontsize=10,
        color="black",
    )

plt.xticks(rotation=45, ha="right")
plt.xlabel("Category")
plt.ylabel("Time to Ship")
plt.title("Categories Having Maximum Shipping Time")
plt.legend(title="Category", loc="upper right")
plt.tight_layout()
plt.show()
```

Output :-



Marwadi University Marwadi Chandarana Group NAAC A+	**Marwadi University** **Faculty of Technology** **Department of Information and Communication Technology**	
Subject: Data Visualization and Dashboards (01CT0410)	**Aim:** Analysis of Superstore Dataset	
Case Study - 1	**Date:-** 16-03-2024	**Enrollment No:-** 92200133030

Tableau Workbook:-



21) What are the sales and profit figures for each year across all categories?

Code:-

```
Dataset['Year'] = Dataset['Order Date'].dt.year
Category_Ship_Wise_Groupping = Dataset.groupby('Category')['Time_to_Ship'].max()
Year_wise_profit_and_sales = Dataset.groupby(["Year", 'Category']).agg(
   {"Profit": "sum", "Sales": "sum"})
grouped_data = Dataset.groupby(['Year', 'Category'])[['Sales', 'Profit']].sum().reset_index()
pivot_data = grouped_data.pivot_table(index='Year', columns='Category', values=['Sales', 'Profit'])
pivot_data.plot(kind='bar', figsize=(12, 6))
plt.title('Sales and Profit for each year (for all categories)')
plt.xlabel('Year')
plt.ylabel('Amount')
plt.legend(title='Category')
plt.xticks(rotation=45)
plt.show()
```

Marwadi University Marwadi Chandarana Group NAAC A+	**Marwadi University** **Faculty of Technology** **Department of Information and Communication Technology**	
Subject: Data Visualization and Dashboards (01CT0410)	**Aim:** Analysis of Superstore Dataset	
Case Study - 1	**Date:-** 16-03-2024	**Enrollment No:-** 92200133030

Output :-

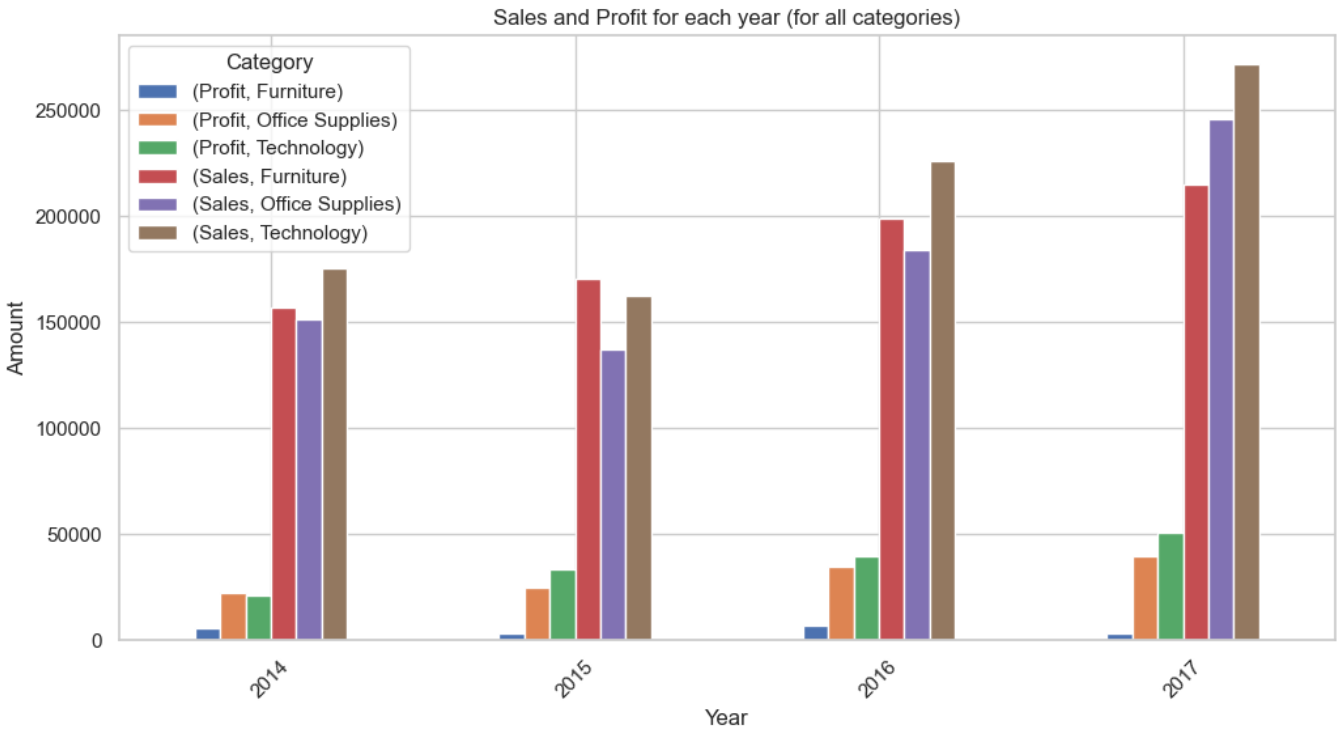
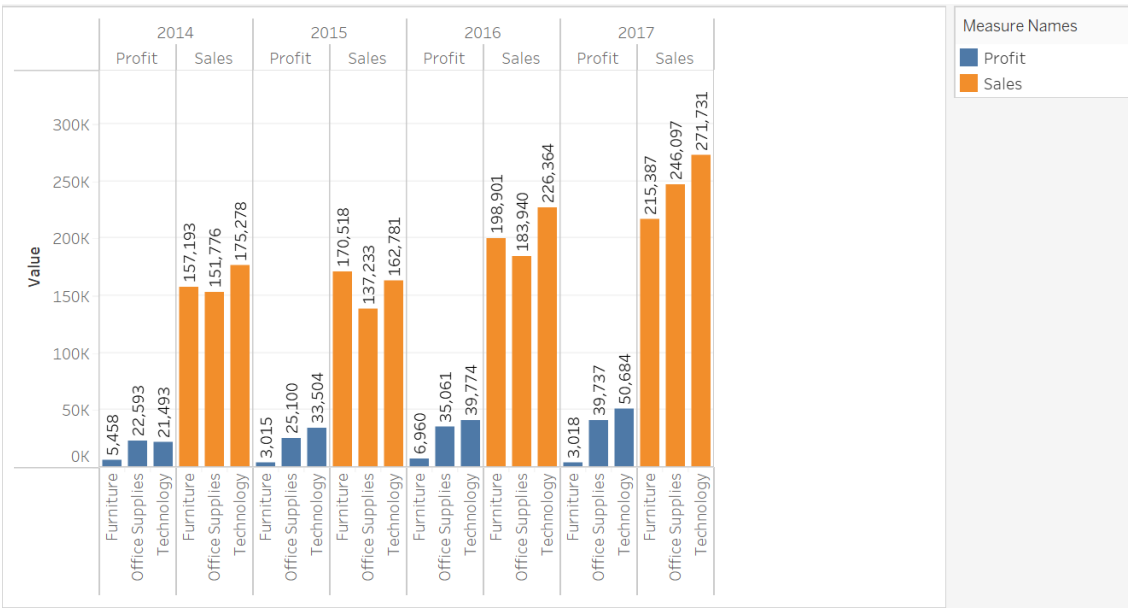


Tableau Workbook :-



22) Design the Dashboard of Questions 1, 16, 8, 9, and 21.

Code:-

```
fig, axs = plt.subplots(3, 2, figsize=(30, 15))
Unique_Categories = list(set(Dataset["Category"]))
axs[0, 0].bar(x=["No Of Unique Categories"], height=[len(Unique_Categories)], width=0.5)
axs[0, 0].set_title("Unique Categories")
Category_wise_Sales_and_profit = Dataset.groupby("Sub-Category")[
    ["Profit", "Sales"]
].sum()
Category_wise_Sales_and_profit.plot(kind="line", ax=axs[0, 1])
axs[0, 1].set_title("Category-wise Sales and Profit")
axs[0, 1].set_xlabel("Sub-Category")
axs[0, 1].set_ylabel("Amount")
axs[0, 1].grid(True)
axs[0, 1].legend(loc="upper right")
Unique_Products = Dataset["Product Name"].unique()
No_of_Unique_Products = Dataset["Product Name"].nunique()
axs[1, 0].bar(x=["No Of Unique Products"], height=[len(Unique_Products)], width=0.5)
axs[1, 0].set_title("Unique Products")
Top_10_Customers = Dataset["Customer Name"].value_counts()
Top_10_Customers_df = pd.DataFrame(
    {"Customer Name": Top_10_Customers.index, "Count": Top_10_Customers.values}
)[:10]
sns.barplot(
    data=Top_10_Customers_df,
    x="Customer Name",
    y="Count",
    hue="Customer Name",
    ax=axs[1, 1],
)
axs[1, 1].set_title("Top 10 Customers")
axs[1, 1].set_xlabel("Customer Name")
axs[1, 1].set_ylabel("Count")
axs[1, 1].legend(title="Customer Name", loc="upper right")
axs[1, 1].tick_params(axis="x", rotation=45)
pivot_data.plot(kind="bar", ax=axs[2, 0], figsize=(16, 6), legend=True)
axs[2, 0].set_title("Sales and Profit for each year (for all categories)")
axs[2, 0].set_xlabel("Year")
axs[2, 0].set_ylabel("Amount")
axs[2, 0].legend(title="Category")
```

```
axs[2, 0].tick_params(axis="x", rotation=45)
plt.show()
```

Output :-

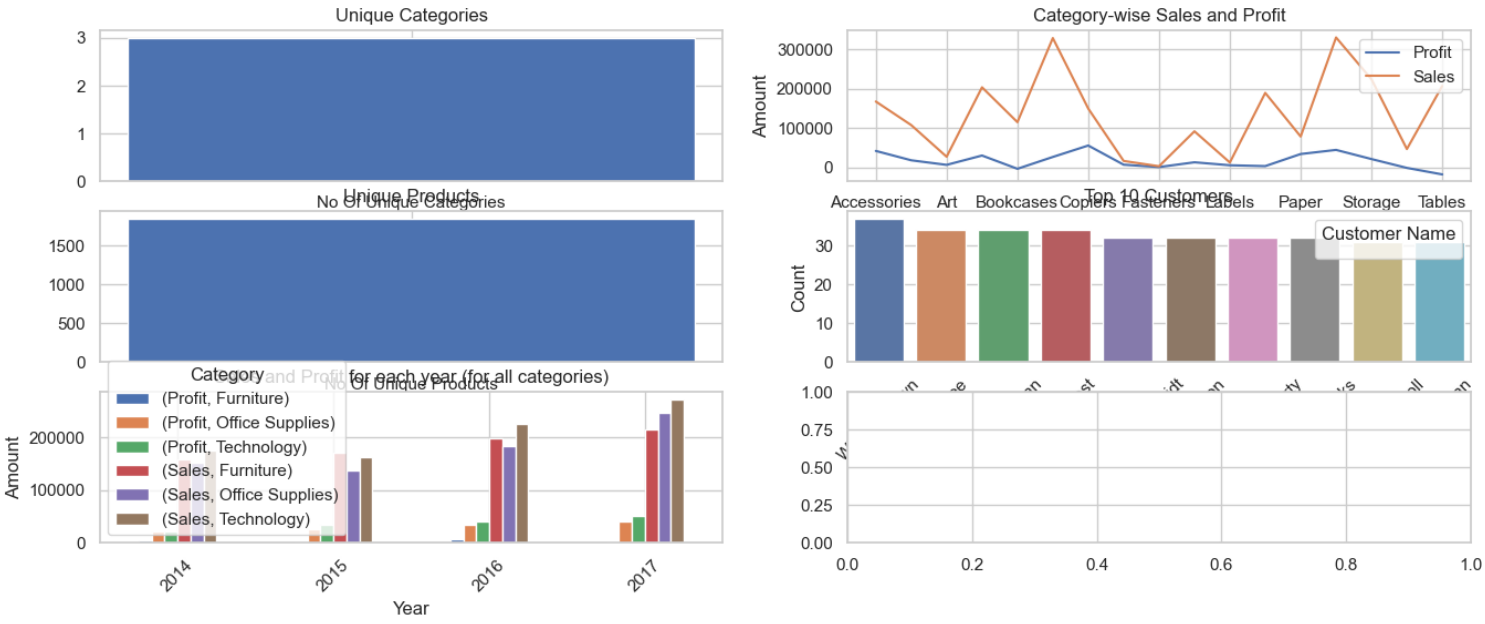


Tableau Workbook :-

Marwadi University **Faculty of Technology** **Department of Information and Communication Technology**		
Subject: Data Visualization and Dashboards (01CT0410)	**Aim:** Analysis of Superstore Dataset	
Case Study - 1	**Date:-** 16-03-2024	**Enrollment No:-** 92200133030