Objective of Case Study-

1. Monthly sales analysis of a Company.
2. Analysis of Sales by Category(Technology, Furniture, Office Supply)
3. Analysis of Monthly Profits
4. Profit Analysis by Category(Technology, Furniture, Office Supply)
5. Sales and Profit Analysis by customer segments

Importing Libraries-

In [1]:
```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import plotly.express as px
import plotly.graph_objects as go
import plotly.io as pio
import plotly.colors as colors

# Define the color palette
color_palette = colors.qualitative.Plotly

# Set default template for Plotly
pio.templates.default = "plotly_white"
```

Loading the dataset-

In [2]:
```
data = pd.read_csv("Business Case Study.csv", encoding="latin-1")
data.head()
```

Out[2]:

	Row ID	Order ID	Order Date	Ship Date	Ship Mode	Customer ID	Customer Name	Segment	Country	City	...	Postal Code	Region	Product ID	Category	S Categ
0	1	CA-2016-152156	11/8/2016	11/11/2016	Second Class	CG-12520	Claire Gute	Consumer	United States	Henderson	...	42420	South	FUR-BO-10001798	Furniture	Bookca
1	2	CA-2016-152156	11/8/2016	11/11/2016	Second Class	CG-12520	Claire Gute	Consumer	United States	Henderson	...	42420	South	FUR-CH-10000454	Furniture	Ch
2	3	CA-2016-138688	6/12/2016	6/16/2016	Second Class	DV-13045	Darrin Van Huff	Corporate	United States	Los Angeles	...	90036	West	OFF-LA-10000240	Office Supplies	La
3	4	US-2015-108966	10/11/2015	10/18/2015	Standard Class	SO-20335	Sean O'Donnell	Consumer	United States	Fort Lauderdale	...	33311	South	FUR-TA-10000577	Furniture	Ta
4	5	US-2015-108966	10/11/2015	10/18/2015	Standard Class	SO-20335	Sean O'Donnell	Consumer	United States	Fort Lauderdale	...	33311	South	OFF-ST-10000760	Office Supplies	Stor

5 rows × 21 columns

Checking Null Values-

In [3]:
```
data.isnull().sum()
```

Out[3]:
```
Row ID           0
Order ID         0
Order Date       0
Ship Date        0
Ship Mode        0
Customer ID      0
Customer Name    0
Segment          0
Country          0
City             0
State            0
Postal Code      0
Region           0
Product ID       0
Category         0
Sub-Category     0
Product Name     0
Sales            0
Quantity         0
Discount         0
Profit           0
dtype: int64
```

Checking the datatype-

In [4]:
```
data.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 9994 entries, 0 to 9993
Data columns (total 21 columns):
 #   Column         Non-Null Count  Dtype
---  ------         --------------  -----
 0   Row ID         9994 non-null   int64
 1   Order ID       9994 non-null   object
 2   Order Date     9994 non-null   object
 3   Ship Date      9994 non-null   object
 4   Ship Mode      9994 non-null   object
 5   Customer ID    9994 non-null   object
 6   Customer Name  9994 non-null   object
 7   Segment        9994 non-null   object
 8   Country        9994 non-null   object
 9   City           9994 non-null   object
 10  State          9994 non-null   object
 11  Postal Code    9994 non-null   int64
 12  Region         9994 non-null   object
 13  Product ID     9994 non-null   object
 14  Category       9994 non-null   object
 15  Sub-Category   9994 non-null   object
 16  Product Name   9994 non-null   object
 17  Sales          9994 non-null   float64
 18  Quantity       9994 non-null   int64
 19  Discount       9994 non-null   float64
 20  Profit         9994 non-null   float64
dtypes: float64(3), int64(3), object(15)
memory usage: 1.6+ MB
```

Descriptive Analysis-

In [5]:
```
data.describe()
```

Out[5]:

	Row ID	Postal Code	Sales	Quantity	Discount	Profit
count	9994.000000	9994.000000	9994.000000	9994.000000	9994.000000	9994.000000
mean	4997.500000	55190.379428	229.858001	3.789574	0.156203	28.656896
std	2885.163629	32063.693350	623.245101	2.225110	0.206452	234.260108
min	1.000000	1040.000000	0.444000	1.000000	0.000000	-6599.978000
25%	2499.250000	23223.000000	17.280000	2.000000	0.000000	1.728750
50%	4997.500000	56430.500000	54.490000	3.000000	0.200000	8.666500
75%	7495.750000	90008.000000	209.940000	5.000000	0.200000	29.364000
max	9994.000000	99301.000000	22638.480000	14.000000	0.800000	8399.976000

Converting Date into DateTime datatype-

In [6]:

```
data["Order Date"] = pd.to_datetime(data["Order Date"])
data["Ship Date"] = pd.to_datetime(data["Ship Date"])
data.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 9994 entries, 0 to 9993
Data columns (total 21 columns):
 #   Column         Non-Null Count  Dtype
---  ------         --------------  -----
 0   Row ID         9994 non-null   int64
 1   Order ID       9994 non-null   object
 2   Order Date     9994 non-null   datetime64[ns]
 3   Ship Date      9994 non-null   datetime64[ns]
 4   Ship Mode      9994 non-null   object
 5   Customer ID    9994 non-null   object
 6   Customer Name  9994 non-null   object
 7   Segment        9994 non-null   object
 8   Country        9994 non-null   object
 9   City           9994 non-null   object
 10  State          9994 non-null   object
 11  Postal Code    9994 non-null   int64
 12  Region         9994 non-null   object
 13  Product ID     9994 non-null   object
 14  Category       9994 non-null   object
 15  Sub-Category   9994 non-null   object
 16  Product Name   9994 non-null   object
 17  Sales          9994 non-null   float64
 18  Quantity       9994 non-null   int64
 19  Discount       9994 non-null   float64
 20  Profit         9994 non-null   float64
dtypes: datetime64[ns](2), float64(3), int64(3), object(13)
memory usage: 1.6+ MB
```

Extracting day, month, year from dates-

In [7]:

```
data["Order Month"]=data["Order Date"].dt.month
data["Order Year"]=data["Order Date"].dt.year
data["Order Day"]=data["Order Date"].dt.dayofweek
data.head()
```

Out[7]:

	Row ID	Order ID	Order Date	Ship Date	Ship Mode	Customer ID	Customer Name	Segment	Country	City	...	Category	Sub-Category	Product Name	Sales	Quantity
0	1	CA-2016-152156	2016-11-08	2016-11-11	Second Class	CG-12520	Claire Gute	Consumer	United States	Henderson	...	Furniture	Bookcases	Bush Somerset Collection Bookcase	261.9600	2
1	2	CA-2016-152156	2016-11-08	2016-11-11	Second Class	CG-12520	Claire Gute	Consumer	United States	Henderson	...	Furniture	Chairs	Hon Deluxe Fabric Upholstered Stacking Chairs,...	731.9400	3
2	3	CA-2016-138688	2016-06-12	2016-06-16	Second Class	DV-13045	Darrin Van Huff	Corporate	United States	Los Angeles	...	Office Supplies	Labels	Self-Adhesive Address Labels for Typewriters b...	14.6200	2
3	4	US-2015-108966	2015-10-11	2015-10-18	Standard Class	SO-20335	Sean O'Donnell	Consumer	United States	Fort Lauderdale	...	Furniture	Tables	Bretford CR4500 Series Slim Rectangular Table	957.5775	5
4	5	US-2015-108966	2015-10-11	2015-10-18	Standard Class	SO-20335	Sean O'Donnell	Consumer	United States	Fort Lauderdale	...	Office Supplies	Storage	Eldon Fold 'N Roll Cart System	22.3680	2

5 rows × 24 columns

Monthly sales analysis of a Company-

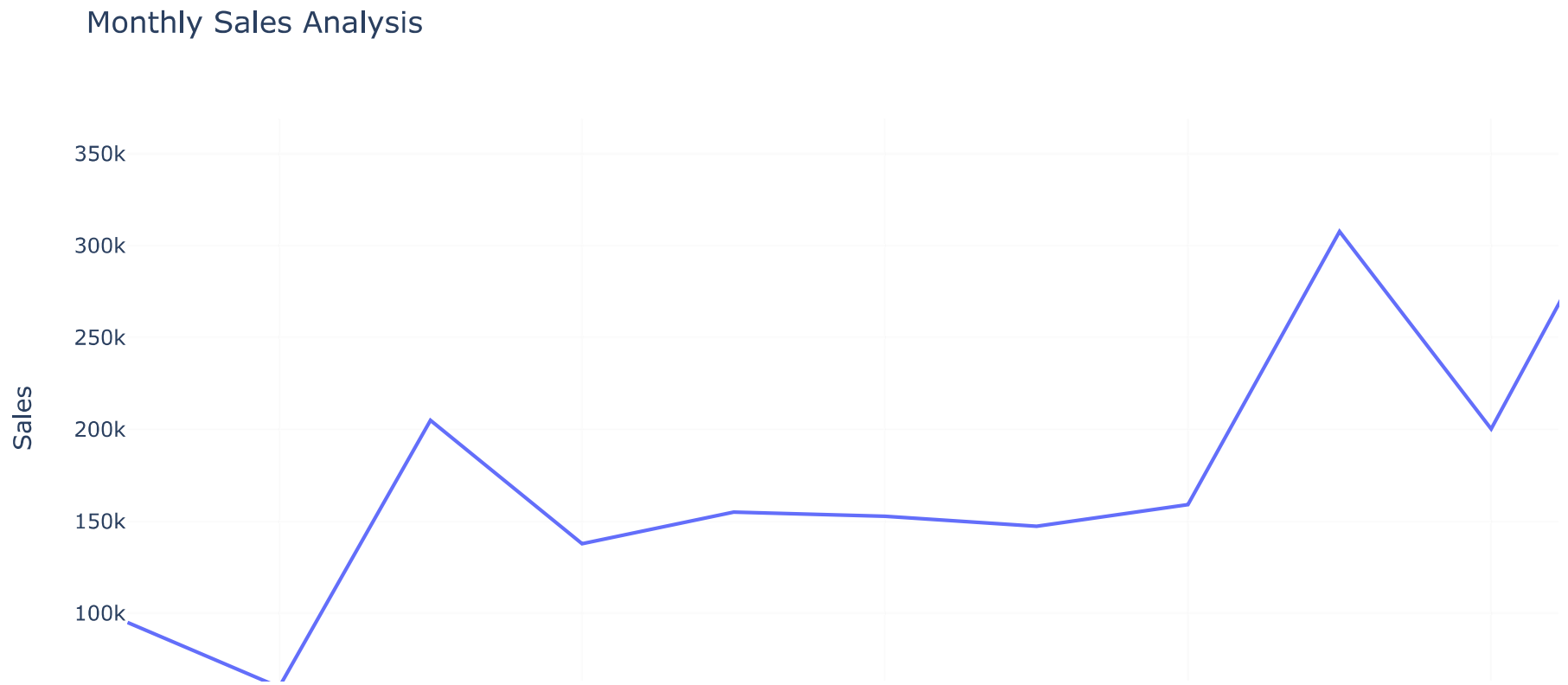
In [8]:
```
monthly_sales = data.groupby("Order Month")["Sales"].sum().reset_index()
monthly_sales
```

Out[8]:

	Order Month	Sales
0	1	94924.8356
1	2	59751.2514
2	3	205005.4888
3	4	137762.1286
4	5	155028.8117
5	6	152718.6793
6	7	147238.0970
7	8	159044.0630
8	9	307649.9457
9	10	200322.9847
10	11	352461.0710
11	12	325293.5035

In [9]:

```
fig=px.line(monthly_sales, x="Order Month", y="Sales", title="Monthly Sales Analysis")
fig.show()
```



Analysis of Sales by Category-

In [10]:
```
Category = data.groupby("Category")["Sales"].sum().reset_index()
Category
```

Out[10]:

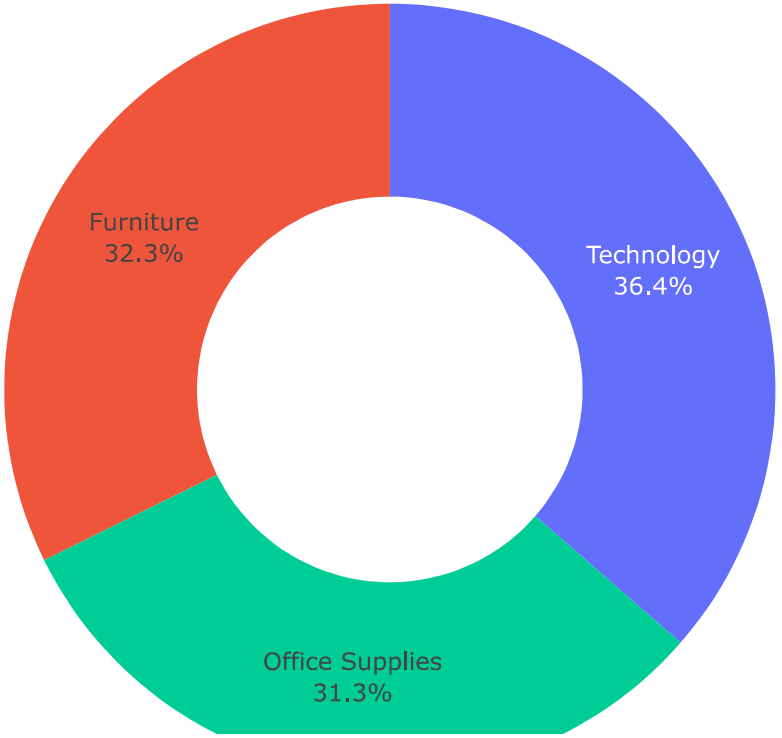
	Category	Sales
0	Furniture	741999.7953
1	Office Supplies	719047.0320
2	Technology	836154.0330

In [11]:

```
fig = px.pie(Category, names="Category", values="Sales", hole=0.5)

fig.update_layout(title_text="Analysis of Sales by Category")
fig.update_traces(textposition="inside", textinfo="percent+label")
fig.show()
```

Analysis of Sales by Category



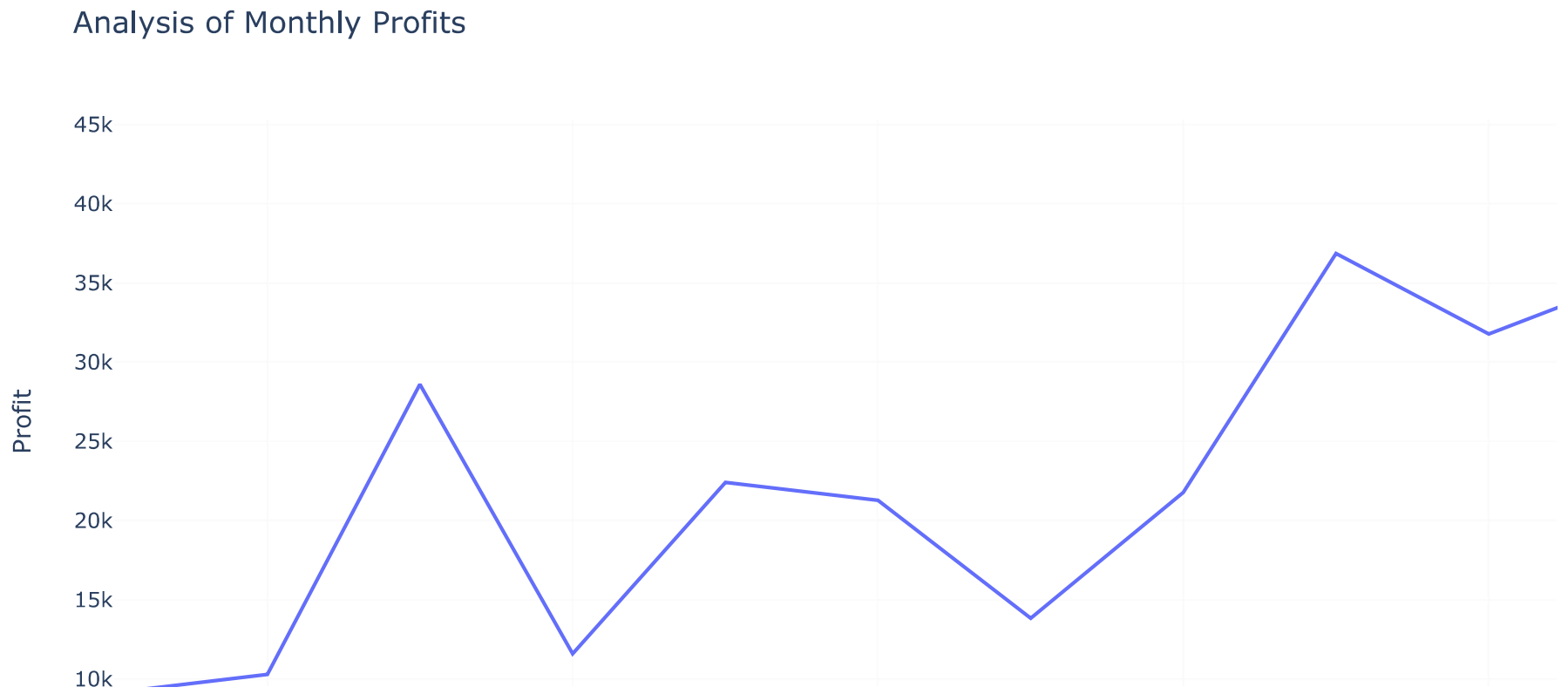
Analysis of Monthly Profits-

In [12]:
```
Monthly_Profit=data.groupby("Order Month")["Profit"].sum().reset_index()
Monthly_Profit
```

Out[12]:

	Order Month	Profit
0	1	9134.4461
1	2	10294.6107
2	3	28594.6872
3	4	11587.4363
4	5	22411.3078
5	6	21285.7954
6	7	13832.6648
7	8	21776.9384
8	9	36857.4753
9	10	31784.0413
10	11	35468.4265
11	12	43369.1919

In [13]:
```
fig = px.line(Monthly_Profit, x="Order Month", y="Profit", title="Analysis of Monthly Profits")
fig.show()
```



Profit Analysis by Category-

In [14]:
```
Profit_by_Category = data.groupby("Category")["Profit"].sum().reset_index()
Profit_by_Category
```

Out[14]:

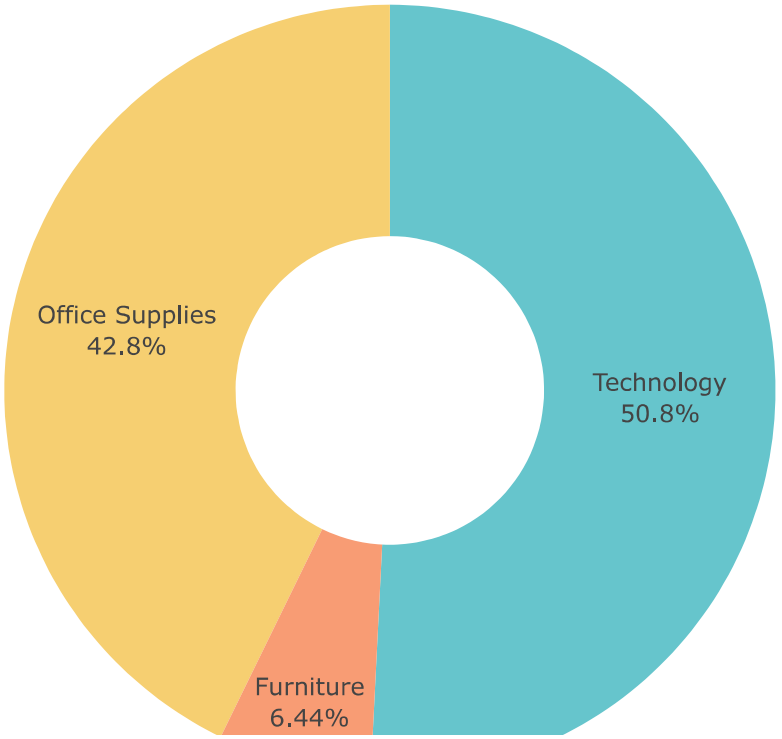
	Category	Profit
0	Furniture	18451.2728
1	Office Supplies	122490.8008
2	Technology	145454.9481

In [15]:

```
fig = px.pie(Profit_by_Category, names="Category", values="Profit", hole=0.4,
             color_discrete_sequence = px.colors.qualitative.Pastel)

fig.update_layout(title_text="Profit Analysis by Category", font_size=12)
fig.update_traces(textposition="inside", textinfo="percent+label")
fig.show()
```

Profit Analysis by Category



Sales and Profit Analysis by customer segments-

In [16]:
```
data.head(2)
```

Out[16]:

	Row ID	Order ID	Order Date	Ship Date	Ship Mode	Customer ID	Customer Name	Segment	Country	City	...	Category	Sub-Category	Product Name	Sales	Quantity	Di
0	1	CA-2016-152156	2016-11-08	2016-11-11	Second Class	CG-12520	Claire Gute	Consumer	United States	Henderson	...	Furniture	Bookcases	Bush Somerset Collection Bookcase	261.96	2	
1	2	CA-2016-152156	2016-11-08	2016-11-11	Second Class	CG-12520	Claire Gute	Consumer	United States	Henderson	...	Furniture	Chairs	Hon Deluxe Fabric Upholstered Stacking Chairs,...	731.94	3	

2 rows × 24 columns

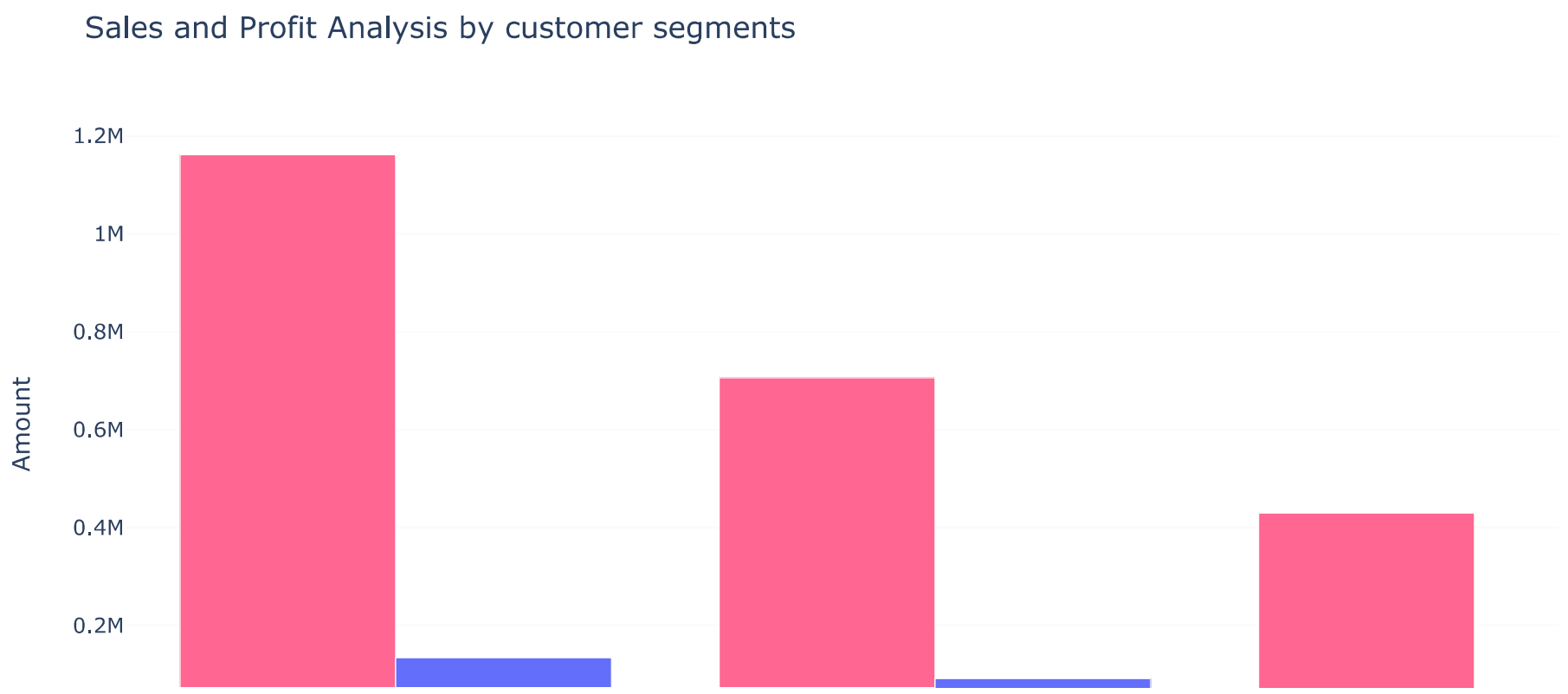
In [17]:
```
Sales_Profit_by_Segment = data.groupby("Segment").agg({"Sales": "sum", "Profit": "sum"}).reset_index()
Sales_Profit_by_Segment
```

Out[17]:

	Segment	Sales	Profit
0	Consumer	1.161401e+06	134119.2092
1	Corporate	7.061464e+05	91979.1340
2	Home Office	4.296531e+05	60298.6785

In [18]:
```
fig = go.Figure()

fig.add_trace(go.Bar(x=Sales_Profit_by_Segment["Segment"],
                     y=Sales_Profit_by_Segment["Sales"],
                     name="Sales",
                    marker_color=color_palette[6]))
fig.add_trace(go.Bar(x=Sales_Profit_by_Segment["Segment"],
                     y=Sales_Profit_by_Segment["Profit"],
                     name="Profit",
                    marker_color=color_palette[0]))
fig.update_layout(title="Sales and Profit Analysis by customer segments",
                  xaxis_title="Customer Segment",
                  yaxis_title="Amount")
fig.show()
```


Sales and Profit Analysis by customer segments
1.2M
1M
0.8M
0.6M
0.4M
0.2M
Amount