PYTHON PROJECT

E-COMMERCE STORE DATA ANALYSIS PROJECT

“NAVIGATING THE FUTURE OF ONLINE SHOPPING”

INTRODUCTION

In today's fast-paced digital economy, E-commerce has revolutionized the way businesses operate and customers shop. With the vast amount of data generated daily, analyzing sales data has become a cornerstone for making informed business decisions. This project, "E-commerce Store Sales Data Analysis," dives into the heart of data-driven decision-making by leveraging Python's robust analytical capabilities.

OBJECTIVE

The project aims to explore, analyze, and visualize sales data from an E-commerce store to uncover meaningful insights that can help optimize business strategies. By identifying trends, customer behavior patterns, and key performance metrics, this analysis empowers stakeholders to drive growth and enhance customer satisfaction.

MONTHLY SALES ANALYSIS

November has the Highest Sales and January has the Lowest Sales

Monthly Sales Analysis

```
sales_by_month = data.groupby('Order Month')['Sales'].sum().reset_index()
```

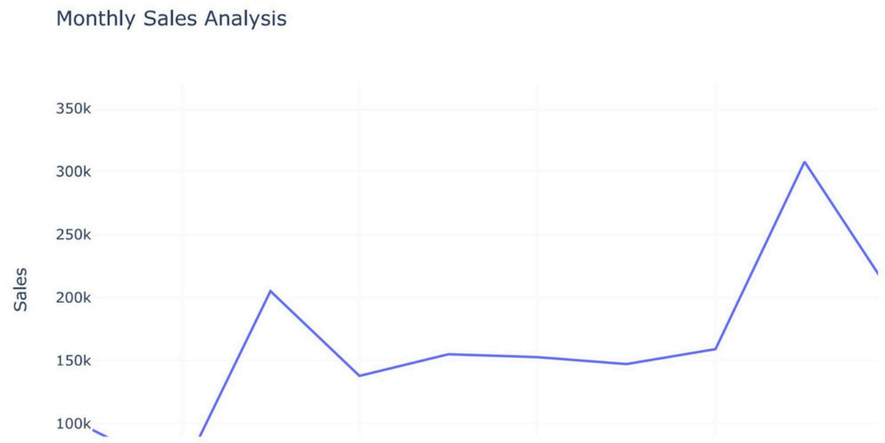
MONTHLY SALES ANALYSIS

November has the Highest Sales and January has the Lowest Sales

```
sales_by_month
```

	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

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



Conclusion- November had the Highest Sales and January had the Lowest Sales.

SALES BY CATEGORY

Sales By Category

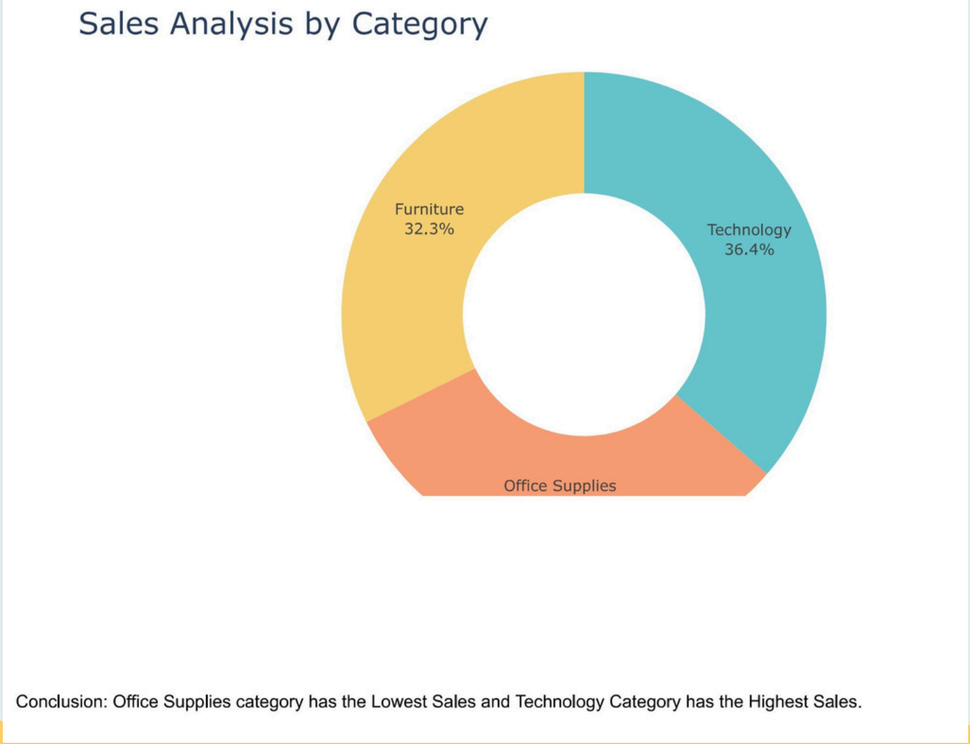
```
sales_by_category= data.groupby('Category')['Sales'].sum().reset_index()
```

```
sales_by_category
```

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

SALES BY CATEGORY

```
fig=px.pie(sales_by_category,
           values= 'Sales',
           names= 'Category',
           hole=0.5,
           color_discrete_sequence= px.colors.qualitative.Pastel)
fig.update_traces(textposition='inside', textinfo='percent+label')
fig.update_layout(title_text='Sales Analysis by Category', title_font=dict(size=24))
fig.show()
```



Conclusion: Office Supplies category has the Lowest Sales and Technology Category has the Highest Sales.

Office Suppliers has the Lowest Sales and Technology has the Highest Sales

SALES BY SUB-CATEGORY

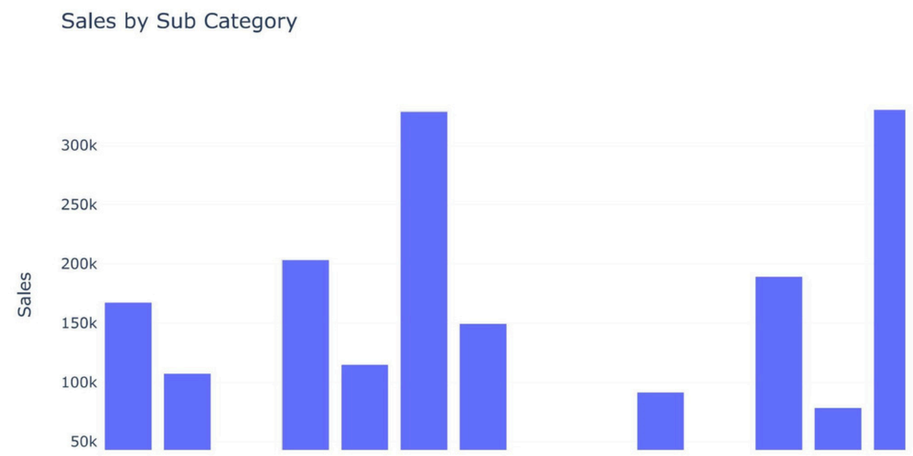
Phones has the First Highest Sub-Category and Chairs has the Second Highest Sub-Category

```
sales_by_subcategory= data.groupby('Sub-Category')['Sales'].sum().reset_index()
```

```
sales_by_subcategory
```

	Sub-Category	Sales
0	Accessories	167380.3180
1	Appliances	107532.1610
2	Art	27118.7920
3	Binders	203412.7330
4	Bookcases	114879.9963
5	Chairs	328449.1030
6	Copiers	149528.0300
7	Envelopes	16476.4020
8	Fasteners	3024.2800
9	Furnishings	91705.1640
10	Labels	12486.3120
11	Machines	189238.6310
12	Paper	78479.2060
13	Phones	330007.0540
14	Storage	223843.6080
15	Supplies	46673.5380
16	Tables	206965.5320

```
fig=px.bar(sales_by_subcategory, x='Sub-Category', y='Sales', title= "Sales by Sub Category")
fig.show()
```



Conclusion: Phones are the first highest selling Sub Category and then Chairs are the Second highest Selling Sub Category.

MONTHLY PROFIT

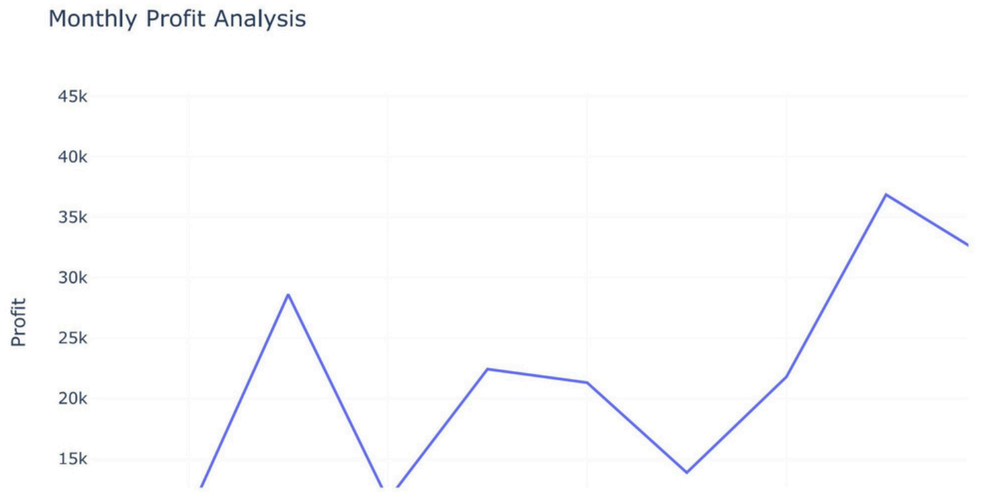
December has the Highest Profit and January has the Lowest Profit

```
profit_by_month= data.groupby('Order Month')['Profit'].sum().reset_index()
```

```
profit_by_month
```

	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

```
fig=px.line(profit_by_month, x= 'Order Month', y= 'Profit', title= "Monthly Profit Analysis")
fig.show()
```



Conclusion: December had the highest Profit and January had the lowest Profit.

PROFIT ANALYSIS BY CATEGORY

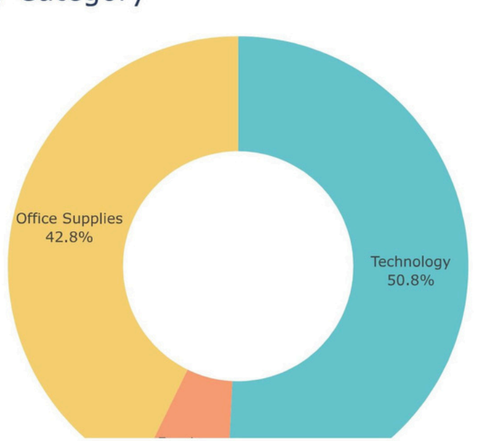
Technology has the Highest Profit

```
profit_by_category= data.groupby('Category')['Profit'].sum().reset_index()
```

```
profit_by_category
```

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

```
fig=px.pie(profit_by_category,
           values= 'Profit',
           names='Category',
           hole=0.5,
           color_discrete_sequence=px.colors.qualitative.Pastel)
fig.update_traces(textposition='inside', textinfo='percent+label')
fig.update_layout(title_text='Profit Analysis by Category', title_font=dict(size=24))
fig.show()
```



Conclusion: Technology has the highest Profit according to Category

PROFIT ANALYSIS BY SUB-CATEGORY

Copiers has the Highest Profit

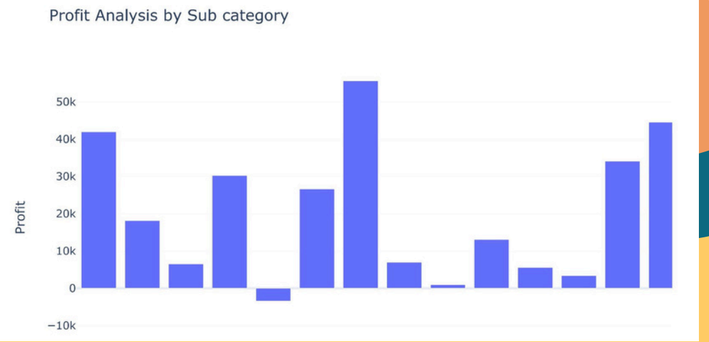
```
profit_by_subcategory= data.groupby('Sub-Category')['Profit'].sum().reset_index()
```

Conclusion: Copiers has the highest Profit accroding to Sub Category

```
profit_by_subcategory
```

	Sub-Category	Profit
0	Accessories	41936.6357
1	Appliances	18138.0054
2	Art	6527.7870
3	Binders	30221.7633
4	Bookcases	-3472.5560
5	Chairs	26590.1663
6	Copiers	55617.8249
7	Envelopes	6964.1767
8	Fasteners	949.5182
9	Furnishings	13059.1436
10	Labels	5546.2540
11	Machines	3384.7569
12	Paper	34053.5693
13	Phones	44515.7306
14	Storage	21278.8264
15	Supplies	-1189.0995
16	Tables	-17725.4811

```
fig=px.bar(profit_by_subcategory, x= 'Sub-Category', y='Profit', title="Profit Analysis by Sub ca
fig.show()
```



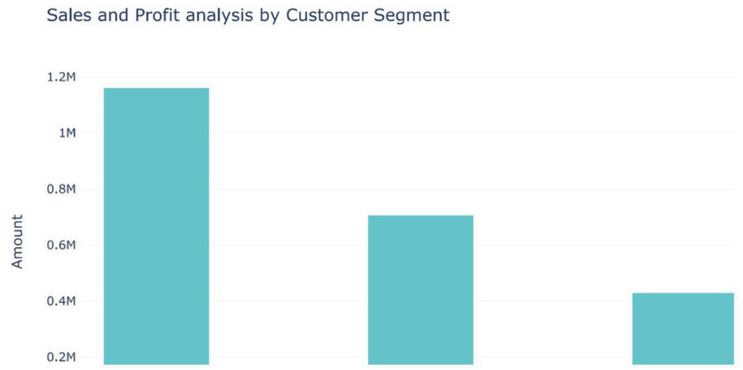
SALES AND PROFIT- CUSTOMER SEGMENT

```
sales_profit_by_segment=data.groupby('Segment').agg({'Sales':'sum', 'Profit':'sum'}).reset_index(
```

```
sales_profit_by_segment
```

	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

```
sales_profit_by_segment=data.groupby('Segment').agg({'Sales':'sum', 'Profit':'sum'}).reset_index(
color_palette= colors.qualitative.Pastel
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[0]))
fig.add_trace(go.Bar(x=sales_profit_by_segment['Segment'],
                     y= sales_profit_by_segment['Profit'],
                     name='Profit',
                     marker_color=color_palette[1]))
fig.update_layout(title='Sales and Profit analysis by Customer Segment',
                  xaxis_title='Customer Segment', yaxis_title='Amount')
fig.show()
```

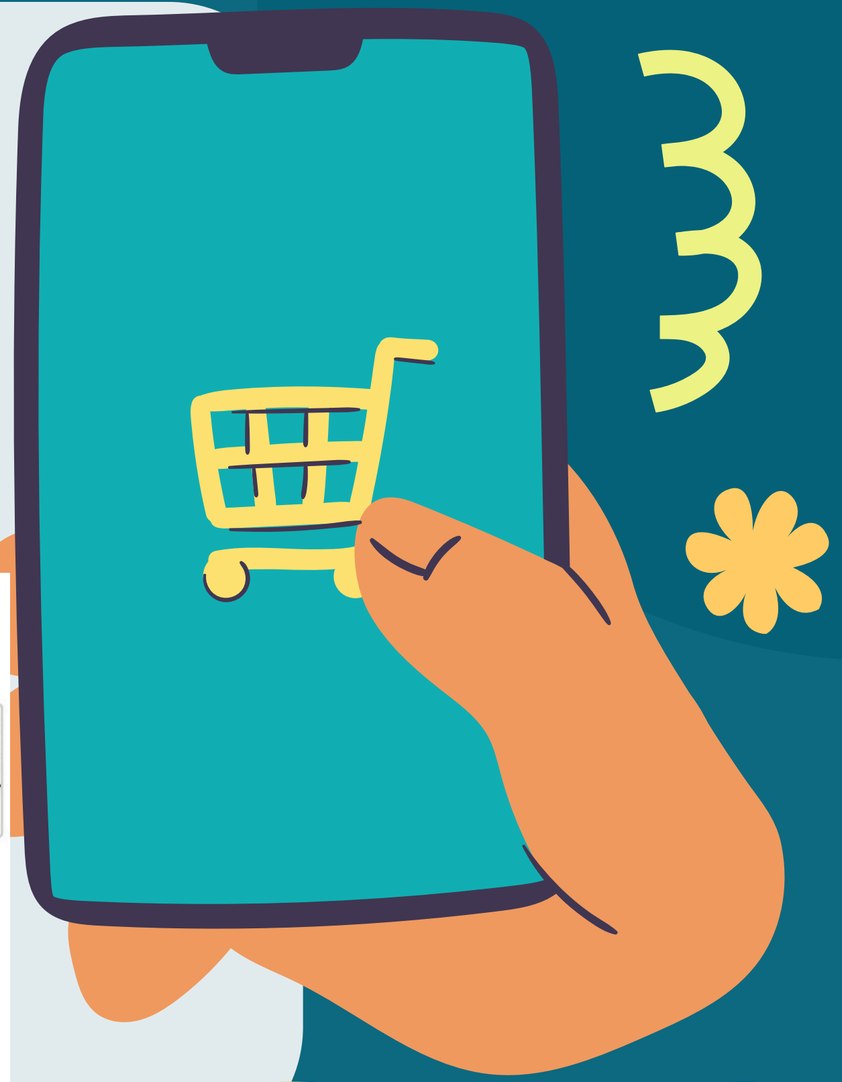


Conclusion: Accroding to Sales First Highest is Consumer, Second Highest is Corporate and Third Highest is Home Office. According to Profit First Highest is Consumer, Second Highest is Corporate and Third Highest is Home Office.

According to Sales First Highest is Consumer, Second Highest is Corporate and Third Highest is Home Office. According to Profit First Highest is Consumer, Second Highest is Corporate and Third Highest is Home Office.

SALES TO PROFIT RATIO

Sales to Profit Ratio for Consumer is 8.659471, Corporate is 7.677245 and Home Office is 7.125416



Sales to Profit Ratio

```
sales_profit_by_segment=data.groupby('Segment').agg({'Sales':'sum', 'Profit':'sum'}).reset_index(
sales_profit_by_segment['Sales_to_Profit_Ratio']=sales_profit_by_segment['Sales']/sales_profit_by
print(sales_profit_by_segment[['Segment','Sales_to_Profit_Ratio']])
```

```
       Segment  Sales_to_Profit_Ratio
0     Consumer               8.659471
1    Corporate               7.677245
2  Home Office               7.125416
```

Conclusion: Sales to Profit Ratio for Consumer is 8.659471, Corporate is 7.677245 and Home Office is 7.125416.

IMPACT AND APPLICATION

THIS PROJECT UNDERSCORES THE SIGNIFICANCE OF DATA ANALYTICS AS A CRITICAL BUSINESS TOOL IN THE E-COMMERCE LANDSCAPE. BY TRANSFORMING RAW DATA INTO MEANINGFUL INSIGHTS, BUSINESSES CAN ENHANCE OPERATIONAL EFFICIENCY, DRIVE CUSTOMER SATISFACTION, AND ACHIEVE SUSTAINABLE GROWTH.

FUTURE SCOPE

The foundation laid by this project opens doors for further enhancements, such as:

- Integrating advanced machine learning models for deeper predictive insights.
- Automating real-time data analysis for quicker decision-making.
- Exploring customer sentiment analysis through reviews and feedback to refine offerings further.

By embracing data-driven approaches, E-commerce businesses can confidently navigate the complexities of the digital marketplace and remain ahead of the competition.

WHY THIS PROJECT MATTERS

This project is a comprehensive example of how E-commerce platforms can harness data analytics to gain a competitive edge. By uncovering hidden patterns and trends in sales data, businesses can optimize their product offerings, enhance marketing campaigns, and improve operational efficiency.

THANK YOU