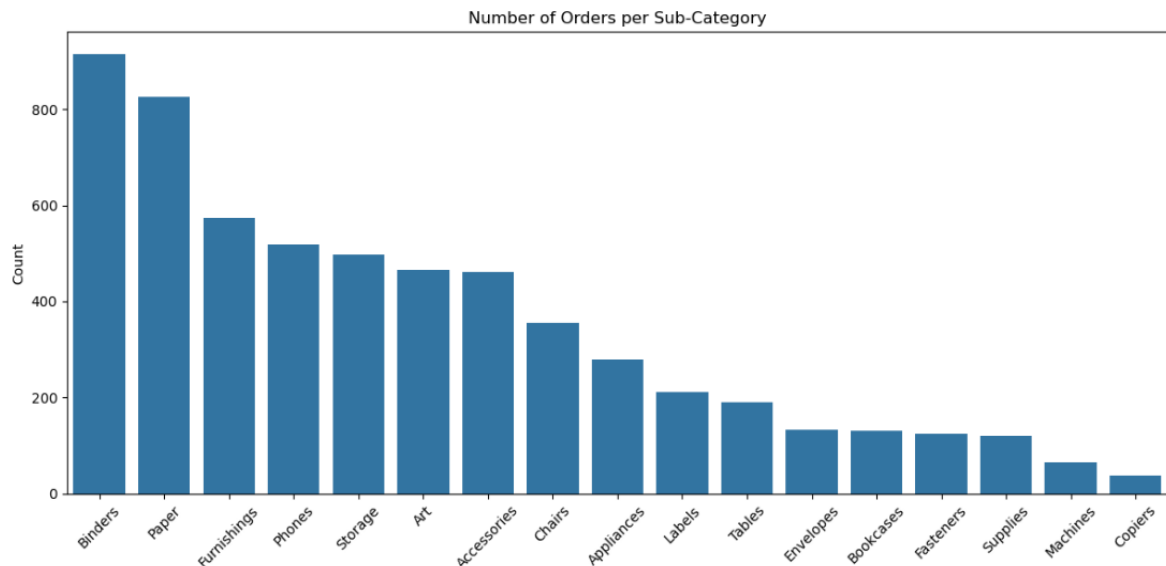


How to **Derive Actionable Data Insights** using data science and python
Here are snippets from my project “**Sales and Profitability Analysis on Amazon Store Data**”
take a look at how I derived valuable insights.

```
import seaborn as sns
import matplotlib.pyplot as plt

plt.figure(figsize=(12,6))
sns.countplot(data=df, x='sub-category', order=df['sub-category'].value_counts().index)
plt.title('Number of Orders per Sub-Category')
plt.xticks(rotation=45)
plt.xlabel('Sub-Category')
plt.ylabel('Count')
plt.tight_layout()
plt.show()
```



Key Insights:

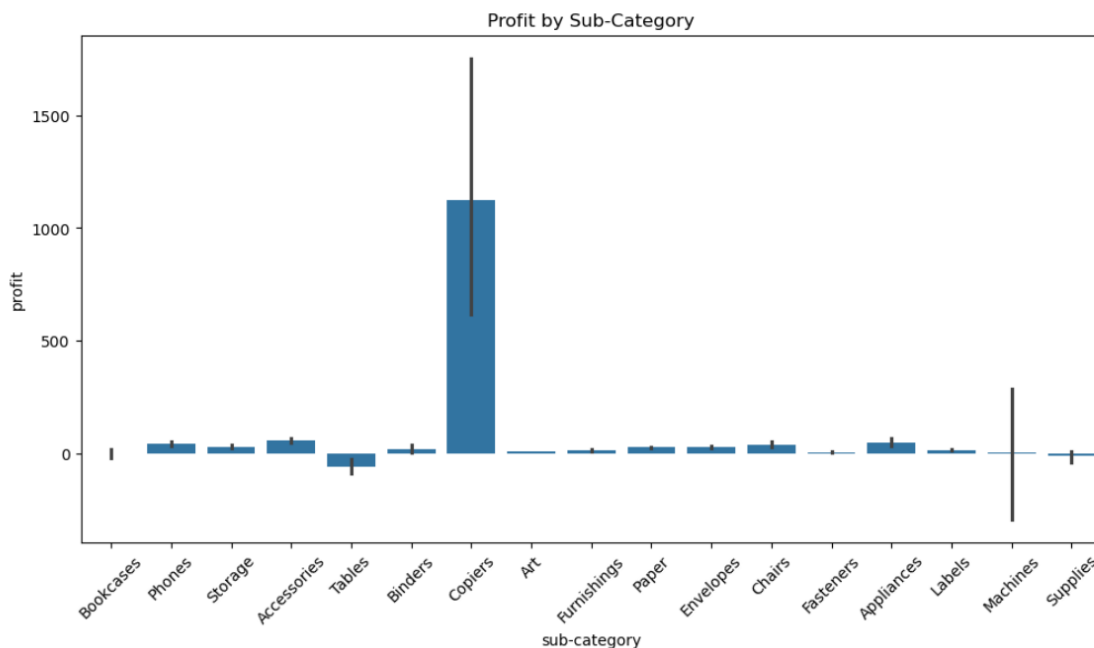
1. Binders and Paper are the top-selling sub-categories, with over 900 and 800 orders respectively.
2. Other popular categories include Furnishings, Phones, and Storage, all exceeding 500 orders.
3. Sub-categories such as Copiers, Machines, and Supplies have the lowest order volumes, indicating lower demand or possibly higher product cost limiting purchases.

Business Implications:

1. We should prioritize stock availability and marketing efforts for high-demand items like Binders, Paper, and Phones.
2. For low-demand items, consider reviewing pricing, customer feedback, or bundling strategies to boost movement.
3. This data can also be cross-analyzed with profit margins to identify which low-volume categories are still profitable.

How to **Derive Actionable Data Insights** using data science and python
Here are snippets from my project “**Sales and Profitability Analysis on Amazon Store Data**”
take a look at how I derived valuable insights.

```
3]: # Profit by Sub-Category
plt.figure(figsize=(12,6))
sns.barplot(data=df, x='sub-category', y='profit')
plt.title('Profit by Sub-Category')
plt.xticks(rotation=45)
plt.show()
```



Key Insights:

1. Copiers stand out with exceptionally high profits, despite having one of the lowest order volumes (from previous chart). This indicates a high profit margin per unit sold.
2. Other moderately profitable sub-categories include Accessories, Appliances, and Fasteners.
3. Interestingly, despite their high order volume, Binders and Paper have low profit contributions, suggesting these are high-volume, low-margin products.
4. Tables show a negative profit, which indicates losses — potentially due to high shipping costs, returns, or aggressive discounting.
5. Sub-categories like Machines, Supplies, and Labels have very low or negative profitability as well.

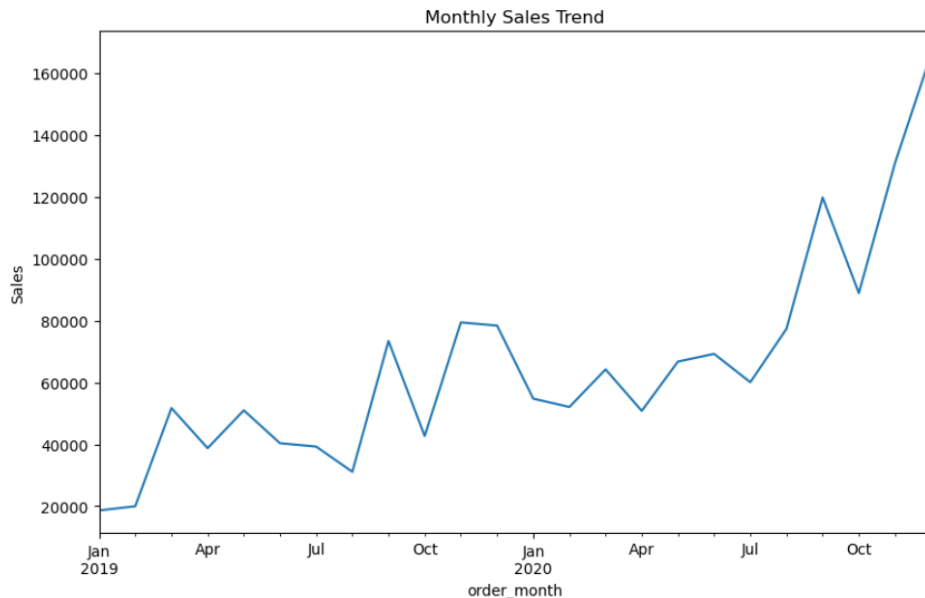
Business Implications:

1. We should promote and protect high-margin products like Copiers, possibly offering service add-ons or warranties to boost value.
2. For low-profit, high-volume items (Binders, Paper), we should:
 - a) Explore cost reduction strategies.
 - b) Consider bundling or cross-selling with profitable items.

How to **Derive Actionable Data Insights** using data science and python
Here are snippets from my project “**Sales and Profitability Analysis on Amazon Store Data**”
take a look at how I derived valuable insights.

```
df['order_month'] = df['order_date'].dt.to_period('M')
monthly_sales = df.groupby('order_month')['sales'].sum()

monthly_sales.plot(kind='line', figsize=(10,6), title='Monthly Sales Trend')
plt.ylabel('Sales')
plt.show()
```



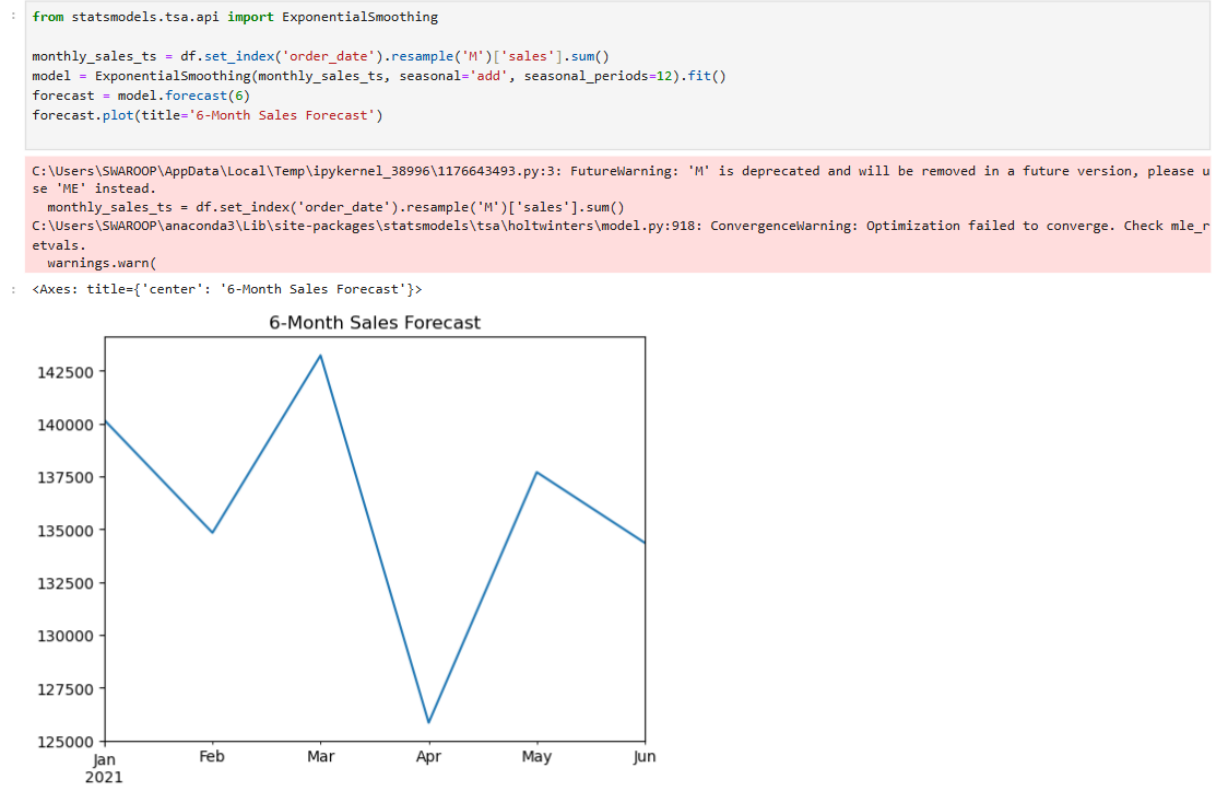
Key Insights:

1. There is a clear upward trend in sales from early 2019 to late 2020, indicating overall growth in business performance.
2. Notable sales spikes are observed during:
 - a) October 2019
 - b) November 2020
 - c) December 2020 — with December showing the highest sales peak, exceeding 160,000 in revenue.
3. Sales tend to rise significantly towards the end of each year, suggesting strong holiday season demand.

Business Implications:

1. We should capitalize on year-end sales spikes by:
 - a) Launching targeted marketing campaigns around Q4 (October to December).
 - b) Increasing inventory of high-demand items in advance.
 - c) Preparing logistics and customer support for higher order volumes.
2. The consistent upward trend suggests our strategies are working — but we should continue monitoring for anomalies or external impacts (like promotions or market disruptions).

How to **Derive Actionable Data Insights** using data science and python
Here are snippets from my project “**Sales and Profitability Analysis on Amazon Store Data**”
take a look at how I derived valuable insights.



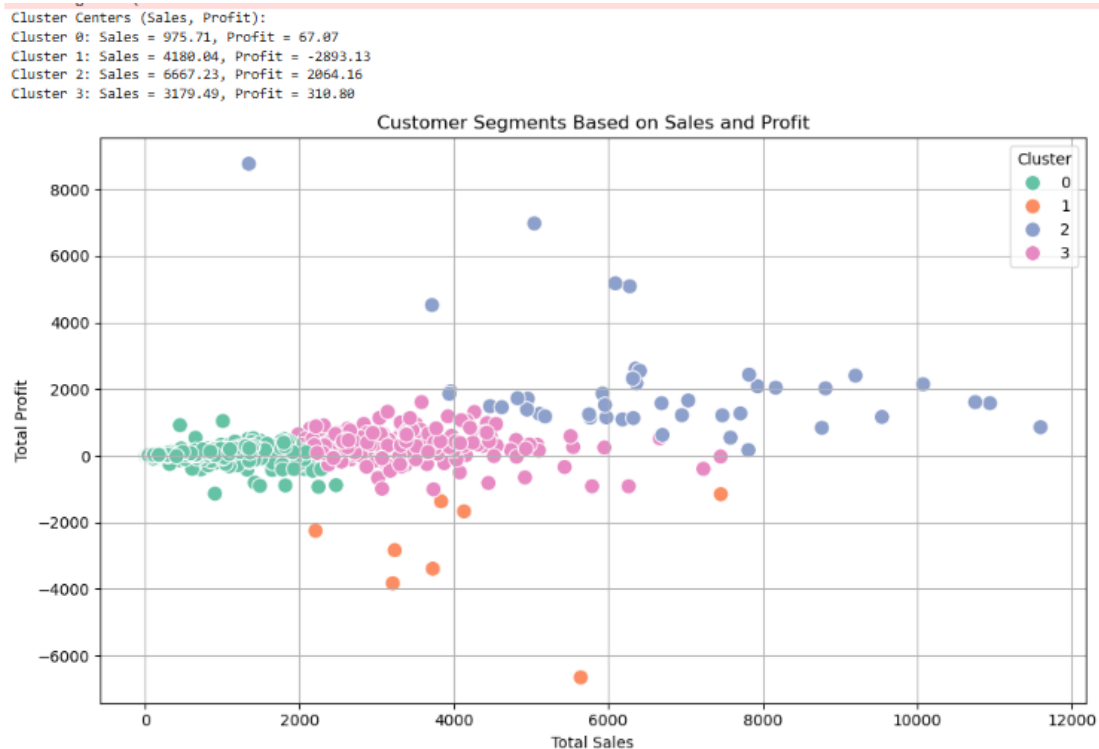
Key Insights:

1. Forecasted sales for this period range between ₹125,000 and ₹143,000 per month.
2. March is expected to experience the highest sales, while April shows a dip - suggesting possible seasonality or post-festival demand drop.
3. The sales trend is overall healthy and consistent, showing no signs of sudden decline or volatility.

Business Implications:

1. Forecasted growth supports confident planning for Q1 and Q2, such as scaling up promotions or increasing inventory of popular items.
2. Given the March peak, we may align product launches or major campaigns in that month to maximize ROI.
3. April's dip could be an opportunity to offer discounts or bundle deals to maintain customer engagement.

How to **Derive Actionable Data Insights** using data science and python
 Here are snippets from my project “**Sales and Profitability Analysis on Amazon Store Data**”
 take a look at how I derived valuable insights.



Objective:

To segment customers into meaningful groups based on their total sales contribution and associated profit, enabling more personalized marketing, retention strategies, and profit optimization.

Approach:

1. I used **K-Means Clustering** on customer data using two key features:
 - a) **Total Sales** per customer
 - b) **Total Profit** generated by each customer
2. The model identified **4 distinct clusters**, each representing a unique customer profile.

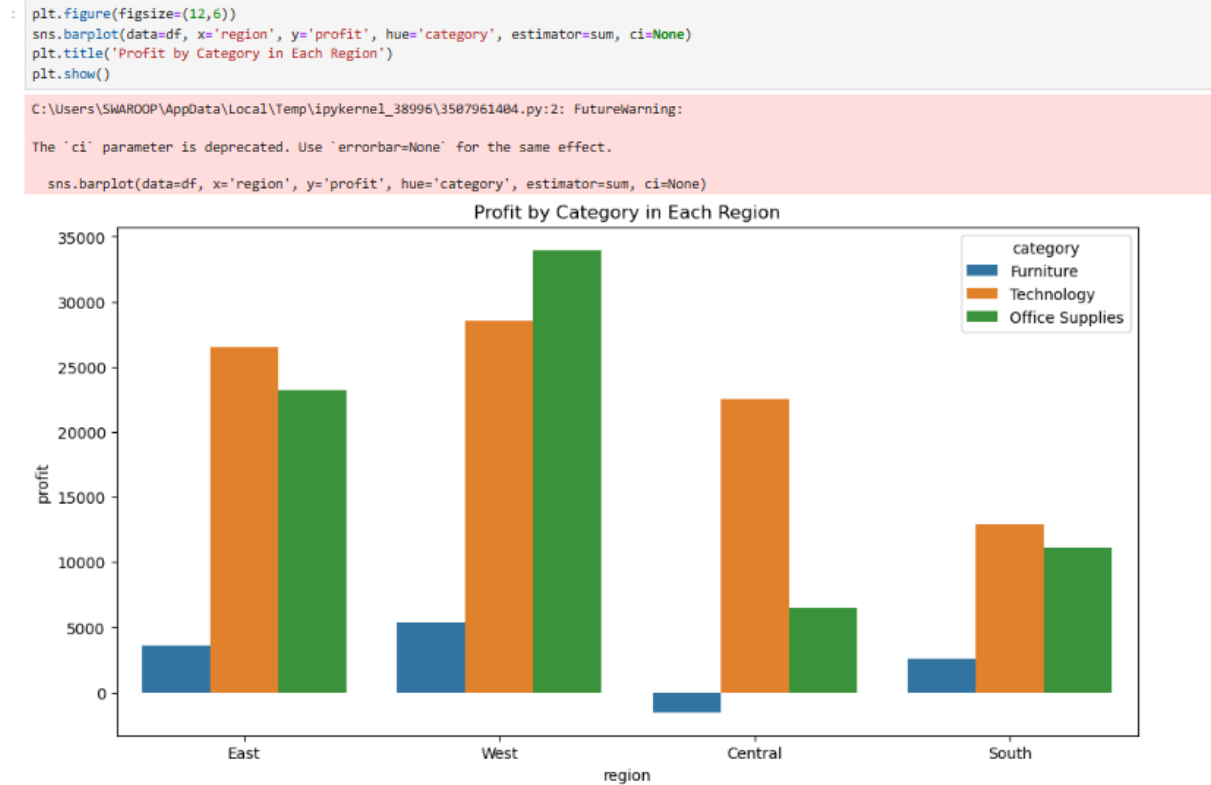
Cluster	Avg Sales	Avg Profit	Interpretation
0	₹976	₹67	Low sales, low profit — small but stable buyers
1	₹4180	-₹2893	High sales, very low profit — likely due to discounts, returns, or unprofitable deals
2	₹6687	₹2064	High sales and high profit — most valuable customers
3	₹3179	₹311	Moderate sales, low but positive profit — may be growth segment

How to **Derive Actionable Data Insights** using data science and python
Here are snippets from my project “**Sales and Profitability Analysis on Amazon Store Data**”
take a look at how I derived valuable insights.

Business Implications:

1. **Cluster 2 (blue):** Our most profitable customers — we should focus on retention, loyalty programs, and upselling to this group.
2. **Cluster 1 (orange):** Despite high sales, this group results in major losses. We need to:
 - a) Audit return/refund rates, discounting strategies, or supply chain costs.
 - b) Possibly restructure pricing or shift focus from these customers.
3. **Cluster 3 (pink):** Potential to become more profitable with targeted promotions or engagement.
4. **Cluster 0 (green):** Stable small buyers — ideal for automated email marketing or bundle offers to increase average order value.

How to **Derive Actionable Data Insights** using data science and python
 Here are snippets from my project “**Sales and Profitability Analysis on Amazon Store Data**”
 take a look at how I derived valuable insights.



Insights:

Region	Most Profitable Category	Observations
West	Office Supplies (₹34K+)	Highest total profit across all regions and categories. Technology also performs very well.
East	Technology (₹27K)	Strong profitability for Technology and Office Supplies; Furniture contributes marginally.
Central	Technology (₹23K)	Technology dominates, but Office Supplies and Furniture lag significantly.
South	Technology (₹12K)	Least profitable region overall, but Technology still leads among the three categories.

THANKYOU