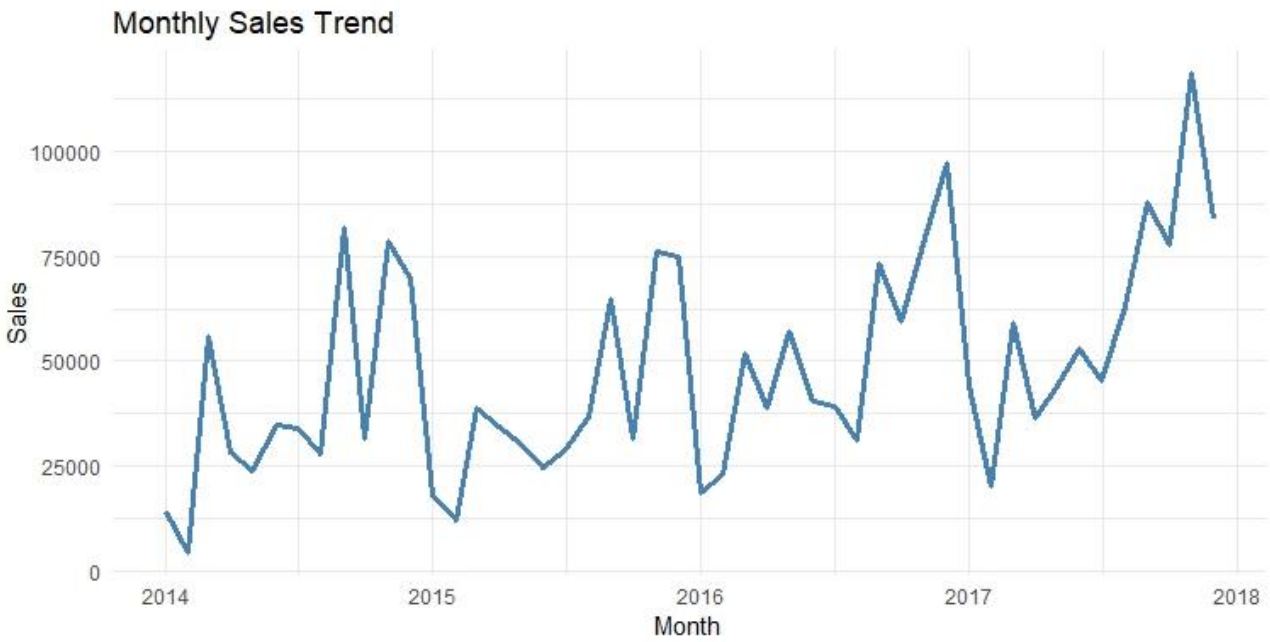
Task 2

We are using superstore dataset Which Contain Variable like

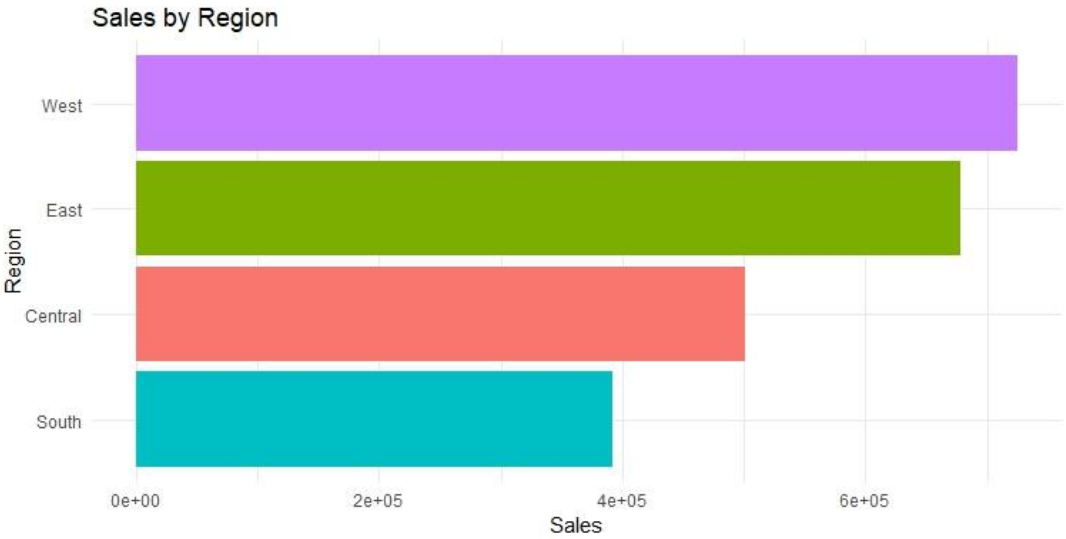
Row ID	Order ID	Order Date	Ship Date	Ship Mode	Customer	Customer	Segment	Country	City	State	Postal Cod	Region	Product ID	Category	Sub-Categ	Product Na	Sales	Quantity	Discount	Profit

Visualization Given As-

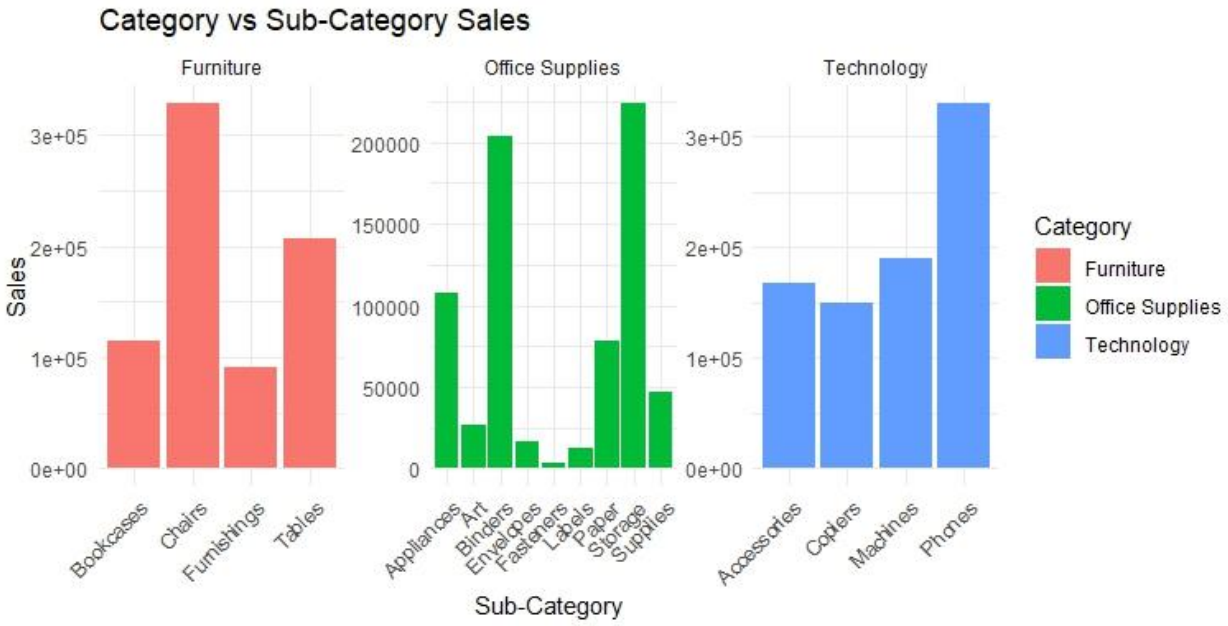
Creates a line plot showing total sales by month. This helps identify seasonal trends or sales growth over time.



Produces a horizontal bar chart showing total sales by each region. Helps identify top-performing regions.

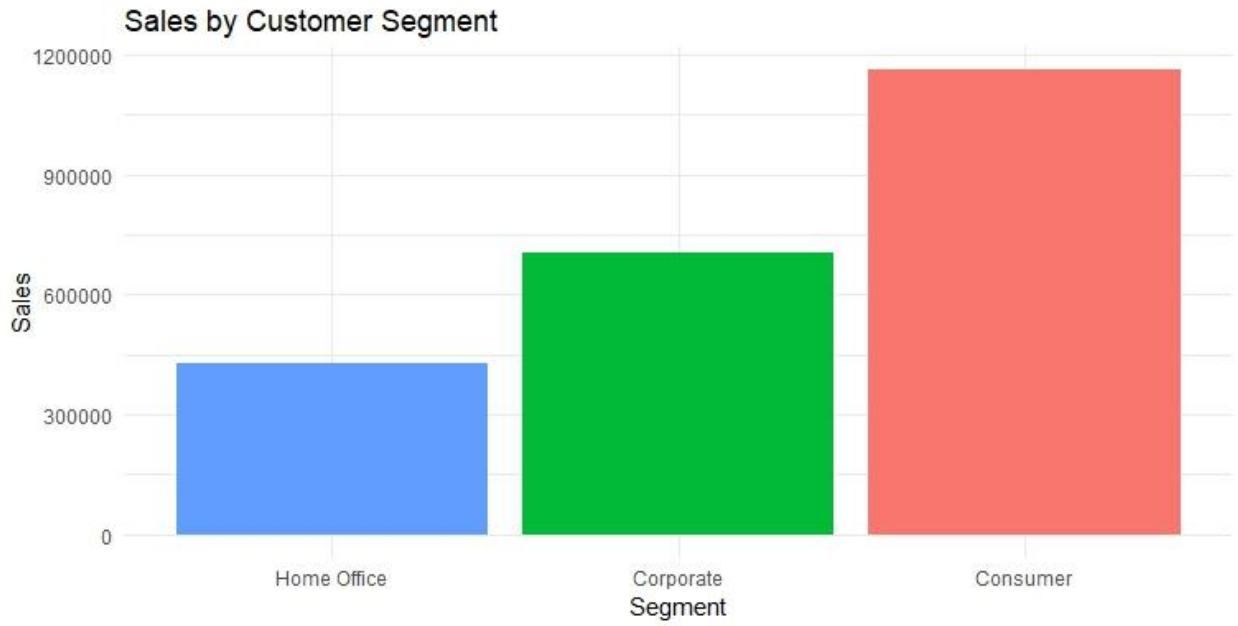


Creates facet bar plots showing how sub-categories perform within each main product category.



A scatter plot showing the relationship between discount and profit. A linear regression line is added to show the trend — helpful for identifying whether discounts reduce or increase profit.

A bar chart that compares sales across customer segments like Consumer, Corporate, and Home Office.



The importance of data visualization

Data visualization is crucial because it transforms raw data into a visual context like charts or graphs, making it easier to:

- Identify **trends, patterns, and outliers**.
- Communicate insights clearly to both technical and non-technical audiences.
- Support **data-driven decision-making**.
- Enhance storytelling by helping viewers **understand the "why" behind the numbers**.
- Save time — visuals are quicker to grasp than tables full of numbers.
- **use a pie chart vs bar chart?**

Pie Chart	To show **parts of a whole** (percentages) when there are **few categories** (ideally <6). Example: market share by company.
Bar Chart	To compare **absolute values** across categories, especially when you have **many items** or values aren't proportions. Example: sales by region.

make visualizations more engaging

- **Choose the right chart** for the data and message.
- Use **color meaningfully**, not just for decoration.
- Add **interactivity** (e.g., in tools like Power BI or Shiny apps).

- Include **clear titles, labels, and tooltips**.
- Use **annotations** to highlight key takeaways or surprising trends.
- Avoid clutter — **simplicity = clarity**.

Data storytelling is the art of combining **data, visuals, and narrative** to communicate a message or insight clearly and persuasively. It includes:

- A **beginning** (setting the context),
- A **middle** (highlighting the problem or discovery),
- And an **end** (drawing conclusions or recommending actions).

avoid misleading visualizations

- Use **consistent scales** (avoid distorted axes).
- **Label charts accurately** and don't exaggerate differences.
- Avoid **3D effects** that obscure actual values.
- Use **appropriate chart types** — e.g., don't use pie charts for time series.
- Be transparent about **data sources** and limitations.
- Never cherry-pick data — show the **complete picture**.

R (ggplot2, plotly, shiny) – for statistical and custom interactive visualizations.