Topic: Implementation of the Walmart Sales in Tableau

1. Introduction

This dataset provides valuable insights into the sales data of a Walmart Superstore from 2014 to 2017. The dataset contains 9,994 records, each representing a unique transaction. With 21 columns, the dataset offers a comprehensive view of various aspects of each transaction, including Order ID, Order Date, Ship Date, Ship Mode, Customer ID, Customer Name, Segment, Country, City, State, Postal Code, Region, Product ID, Category, Sub-Category, Product Name, Sales, Quantity, Discount, and Profit.

The dataset captures diverse data points such as the customer segments, product categories and sub-categories, and geographical information, making it a rich resource for understanding and analyzing the performance of the superstore across different dimensions. By utilizing this dataset, we can uncover patterns, trends, and relationships among the various factors influencing the superstore's sales and profit, enabling us to derive actionable insights to enhance its performance and decision-making.

2. Data Dictionary

Sr. No.	Field	Description	Format
1	Row ID	Unique identifier for each row	Integer
2	Order ID	Unique identifier for each order	String
3	Order Date	Date the order was placed	Date (YYYY-MM-DD)
4	Ship Date	Date the order was shipped	Date (YYYY-MM-DD)
5	Ship Mode	Shipping mode used for the order	String
6	Customer ID	Unique identifier for each customer	String
7	Customer Name	Name of the customer	String
8	Segment	Market segment of the customer	String
9	Country	Country of the order	String
10	City	City of the order	String
11	State	State of the order	String
12	Postal Code	Postal code of the order location	Integer
13	Region	Region of the order	String
14	Product ID	Unique identifier for each product	String
15	Category	Product category	String

16	Sub-Category	Product sub-category	String
17	Product Name	Name of the product	String
18	Sales	Sales revenue for the product	Float
19	Quantity	Number of items sold in the order	Integer
20	Discount	Discount rate applied to the order	Float (0.00 - 1.00)
21	Profit	Profit generated from the order	Float

Source of dataset: [Walmart Sales Analysis | Kaggle]

3. Persona & Research Questions:

The persona for this analysis is a **Store Operations Manager** at **Walmart**. This individual is responsible for overseeing the day-to-day activities in the store, optimizing sales, managing inventory, and ensuring customer satisfaction. Their main goal is to increase the efficiency of the store's operations while maximizing profits and minimizing costs.

As a Store Operations Manager, they need to make informed decisions based on data to drive better results for the business. They must analyse various factors, such as customer segments, product categories, geographical regions, and discount strategies, to identify patterns and trends that can impact the store's performance. This persona would greatly benefit from a dashboard that allows them to visualize the data and make more informed decisions for optimizing store performance.

Research Question (Q1): What are the most profitable categories and states, and how do their sales and profit trends change over time? Which period was the most profitable for the Walmart?

Requirement (R1): To answer the question, it requires an area line chart to display sales and profit trends by category and state over time, with tooltips for additional context.

Requirement (R2): It requires a stacked bar chart to represent sales and profit by region, further broken down by category and state. This chart should include filtering options to determine the most profitable time period for Walmart and tooltips for further context on the data points.

Research Question (Q2): How do sales and profit vary by the level of discounts offered, so discounts effect the sales and which product categories, and sub-categories are most sensitive to discounting?

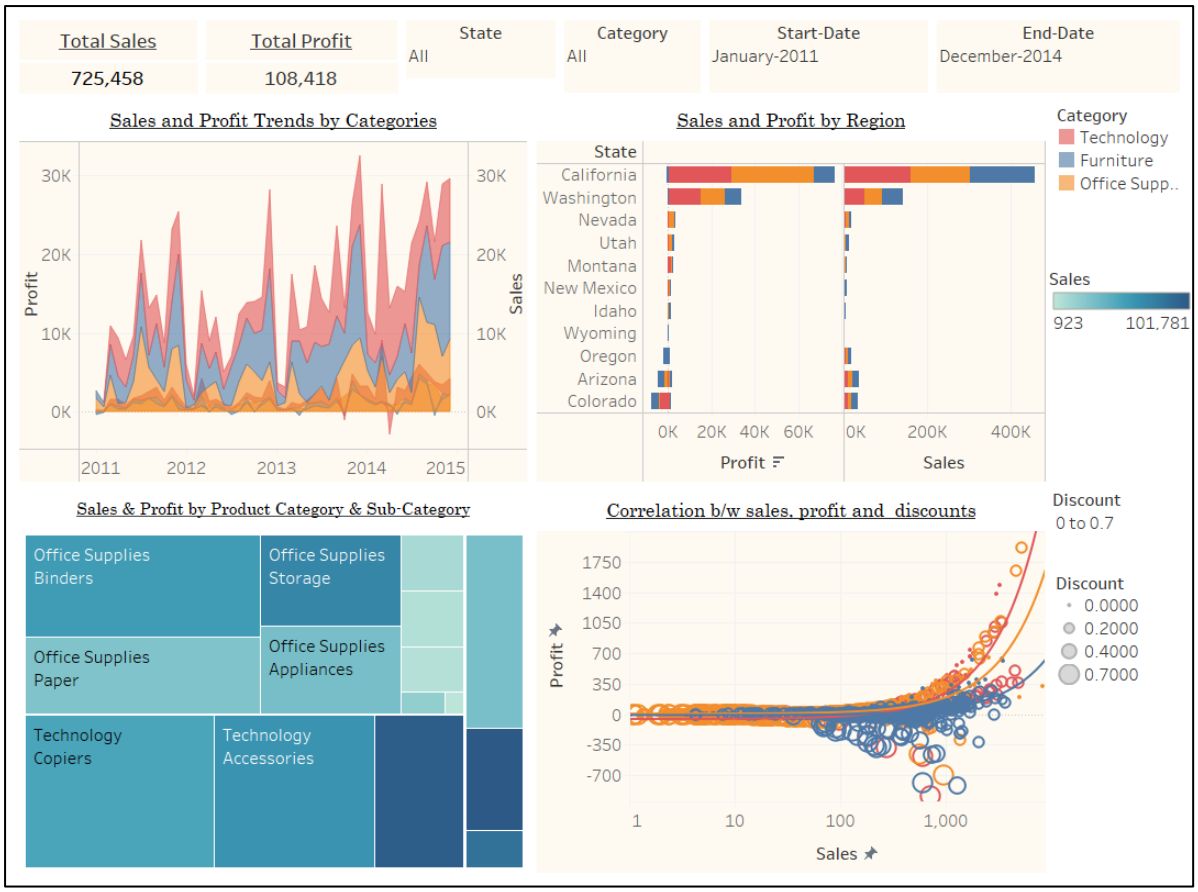
Requirement (R3): It requires a treemap to display sales and profit by product category and sub-category, allowing us to identify the categories most sensitive to discounting.

Requirement (R4): It requires a scatter plot to represent the correlation between sales, profit, and discounts, enabling us to analyze how discounts affect sales and profit across different categories and sub-categories.

4. Design

The given figure shows the prototype of the dashboard that I was planning to answer my research question before the feedback. However, after the feedback I got to change some parts of the dashboard and because of Tableau's features I can enhance the dashboard in terms of looks as well as feasible to understand the dashboard and answer the question easily through some interactions.

The evolved dashboard, we can see the image of the final dashboard here:

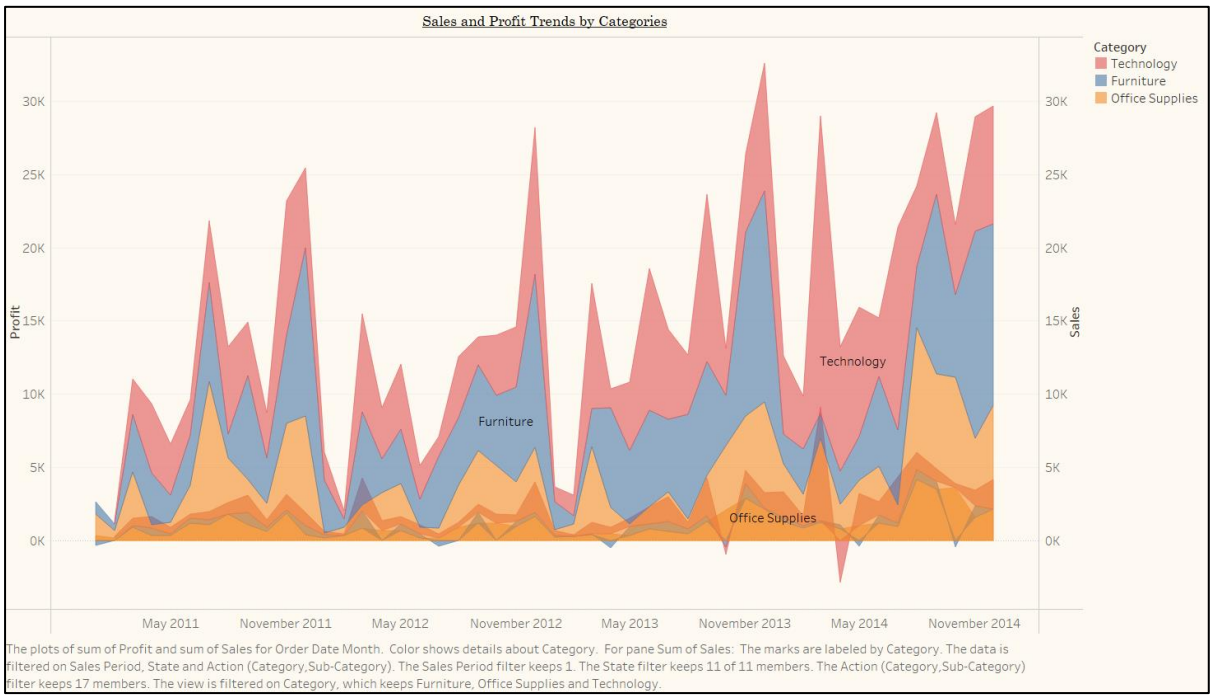


As it can clearly be seen that there are significant numbers of changes made in the final dashboard compared to the prototype. Because of the tableau's feature I was able to add more detailing to the graphs and can answer the questions more easily. The Line chart in the prototype evolved to area chart with more information loaded in compared to prototype. Moreover, I have converted the normal bar graph to stack bar chart to add more detail to answer the question in more detail. The treemap is added because to compare the subcategories with categories with broader picture. As the dataset had huge number of subcategories the user won't be able to understand it because it could get messed up. Therefore, implementation of treemap was the best visualization to answer our question.

5. Implementation

This section will explain the implementation of the graphs to the final dashboard.

This figure is the header section of the dashboard from which the user can get the information according to the need of the user. Moreover, for the convenience of our user they can see the statistics of the data at one glance whether they need total sales or profit in any state, category, or between any date.



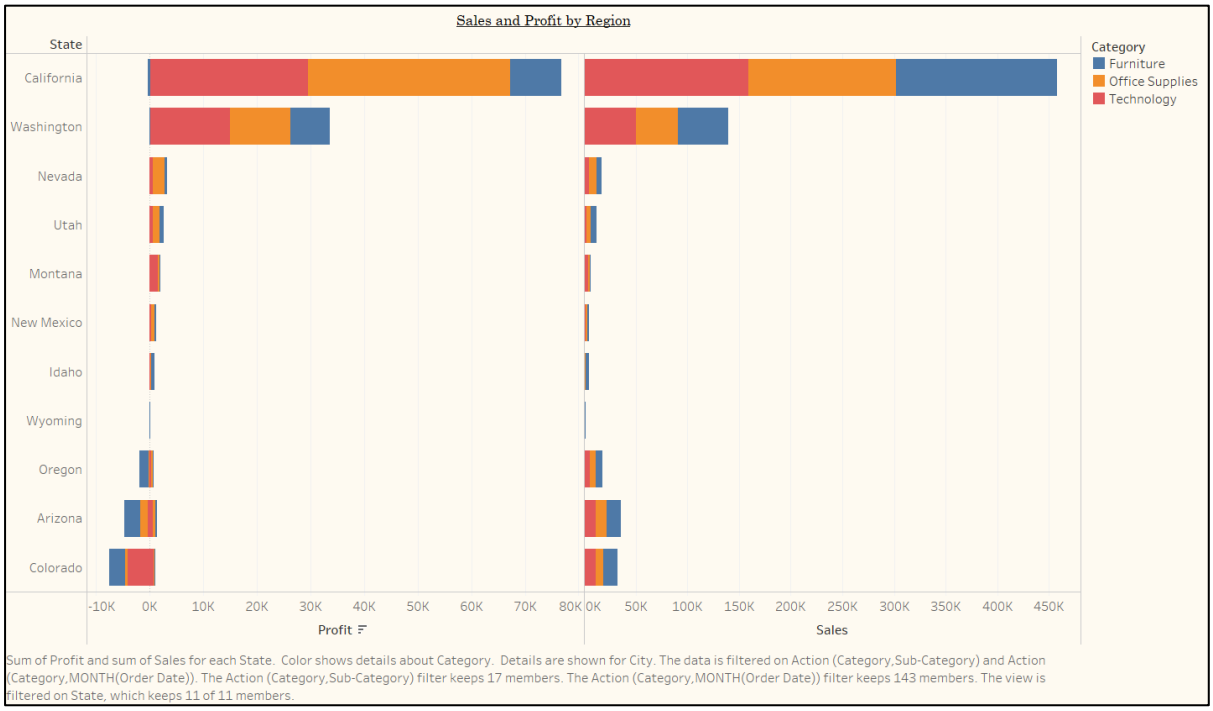
We designed an area chart to visualise sales and profit trends over time by category. The area chart efficiently conveys sales and profit changes for each category, allowing the Store Operations Manager to easily recognise trends and make informed decisions.To create the area chart, we followed these steps:

1. Created a new worksheet in Tableau and named it "Sales and Profit Trends by Category over Time."
2. Dragged the "Order Date" dimension to the Columns shelf and set it to display as a continuous variable (MONTH).
3. Dragged the "Sales" and "Profit" measures to the Rows shelf. This displays the sales and profit data on the same chart, with sales as the primary axis and profit as the secondary axis.

4. Dragged the "Category" dimension to the Color shelf to differentiate the categories in the area chart.
5. Adjusted the chart type from "Automatic" to "Area."
6. Included tooltips to provide additional context for the data points, such as specific sales, profit values, and category details.

The generated area chart enables the Store Operations Manager to quickly visualise how sales and profit trends vary over time for each category. The graphic also shows when sales and earnings were at their peak, allowing the management to identify the most profitable intervals and make informed decisions about resource allocation and marketing initiatives.

The shop Operations Manager can acquire useful insights into the performance of each category, detect seasonal trends, and establish strategies to optimise sales and profits for the shop by analysing the area chart.



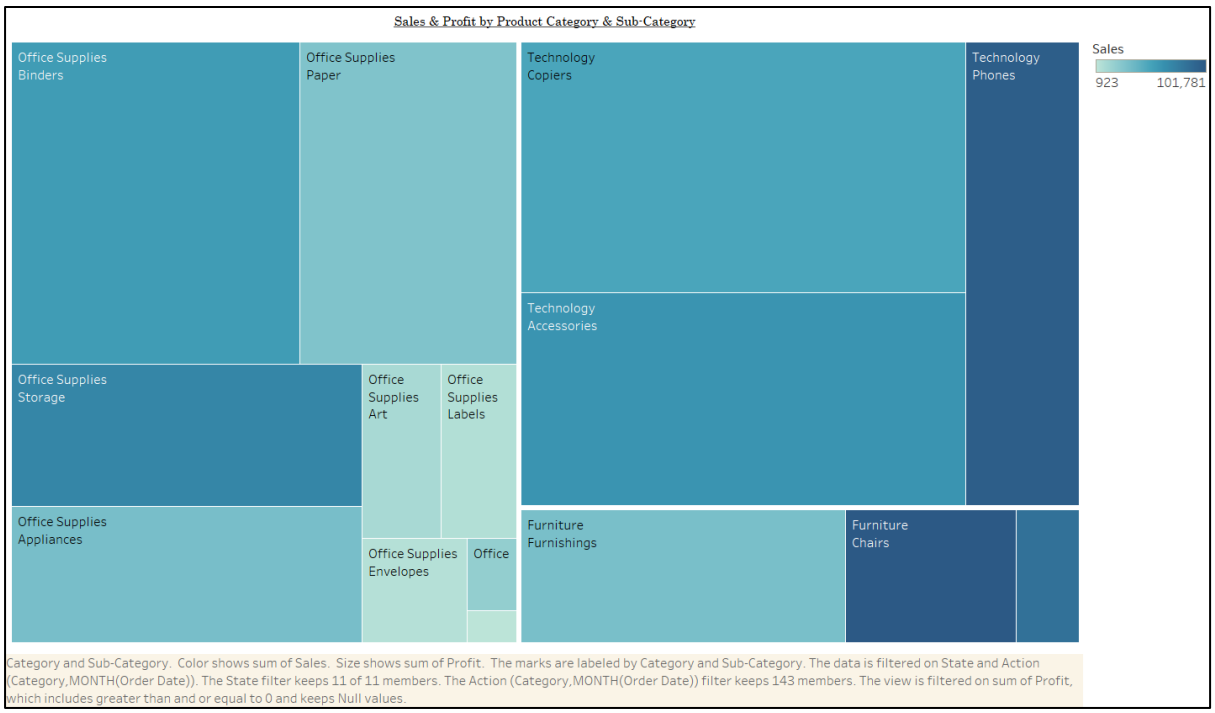
We developed a stacked bar chart to depict sales and profit by region, further subdivided by category. The stacked bar chart provides an easy-to-understand visual representation of the data, allowing the Store Operations Manager to compare the sales and profit growth of different regions and categories.

To create the stacked bar chart, we followed these steps:

1. Created a new worksheet in Tableau and named it "Sales and Profit by Region and Categories."
2. Dragged the "Region" dimension to the Columns shelf to display the data breakdown by region.
3. Dragged the "Sales" and "Profit" measures to the Rows shelf, showing the sales and profit data as bars for each region.

4. Dragged the "Category" dimension to the Color shelf to differentiate the categories within each bar. This created a stacked bar chart, with each category represented by a unique color within the bar.
5. Adjusted the chart type from "Automatic" to "Bar."
6. Included tooltips to provide additional context for the data points, such as specific sales, profit values, region, and category details.

The resulting stacked bar chart allows the Store Operations Manager to quickly compare the performance of each region in terms of sales and profits, as well as identify the top-performing categories within each region. This information can be used to make informed decisions about resource allocation, marketing efforts, and overall strategies for maximizing store performance.



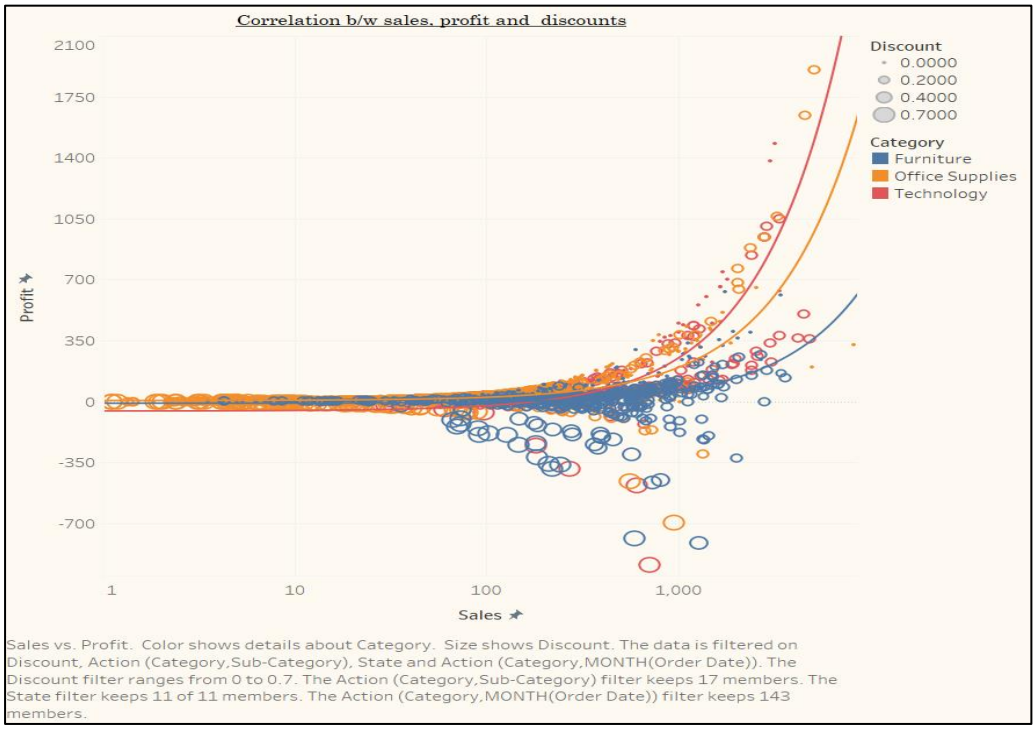
To visualise sales and profit by product category and sub-category, we designed a treemap. The treemap efficiently hierarchically presents the data, with the size of each box signifying profit and the colour gradient denoting sales. The Store Operations Manager may readily determine the most profitable and best-selling categories and sub-categories as a result of this.

To create the treemap, we followed these steps:

1. Created a new worksheet in Tableau and named it "Sales and Profit by Product Category & Sub-Category."
2. Dragged the "Category" dimension to the Rows shelf.
3. Dragged the "Sub-Category" dimension to the Rows shelf, placing it to the right of the "Category" pill. This creates a nested hierarchy for categories and sub-categories.

4. Dragged the "Profit" measure to the Size shelf, which determines the size of each box in the treemap based on profit.
5. Dragged the "Sales" measure to the Color shelf, creating a color gradient that reflects the sales values for each box.
6. Adjusted the chart type from "Automatic" to "Treemap."
7. Included tooltips to provide additional context for the data points, such as specific sales, profit values, and category details.

The resulting treemap allows the Store Operations Manager to quickly visualize the sales and profit performance of each category and sub-category. By analyzing the size and color gradient of each box, the manager can identify the most profitable and best-selling products, and make informed decisions about inventory management, marketing efforts, and overall strategies for maximizing store performance.



We created a scatter plot to visualize the relationship between sales and profit by category, with the size of each circle representing different discount bins. This visualization provides the Store Operations Manager with insights into the impact of discounts on sales and profits and helps them identify potential trends or outliers.

To create the scatter plot, we followed these steps:

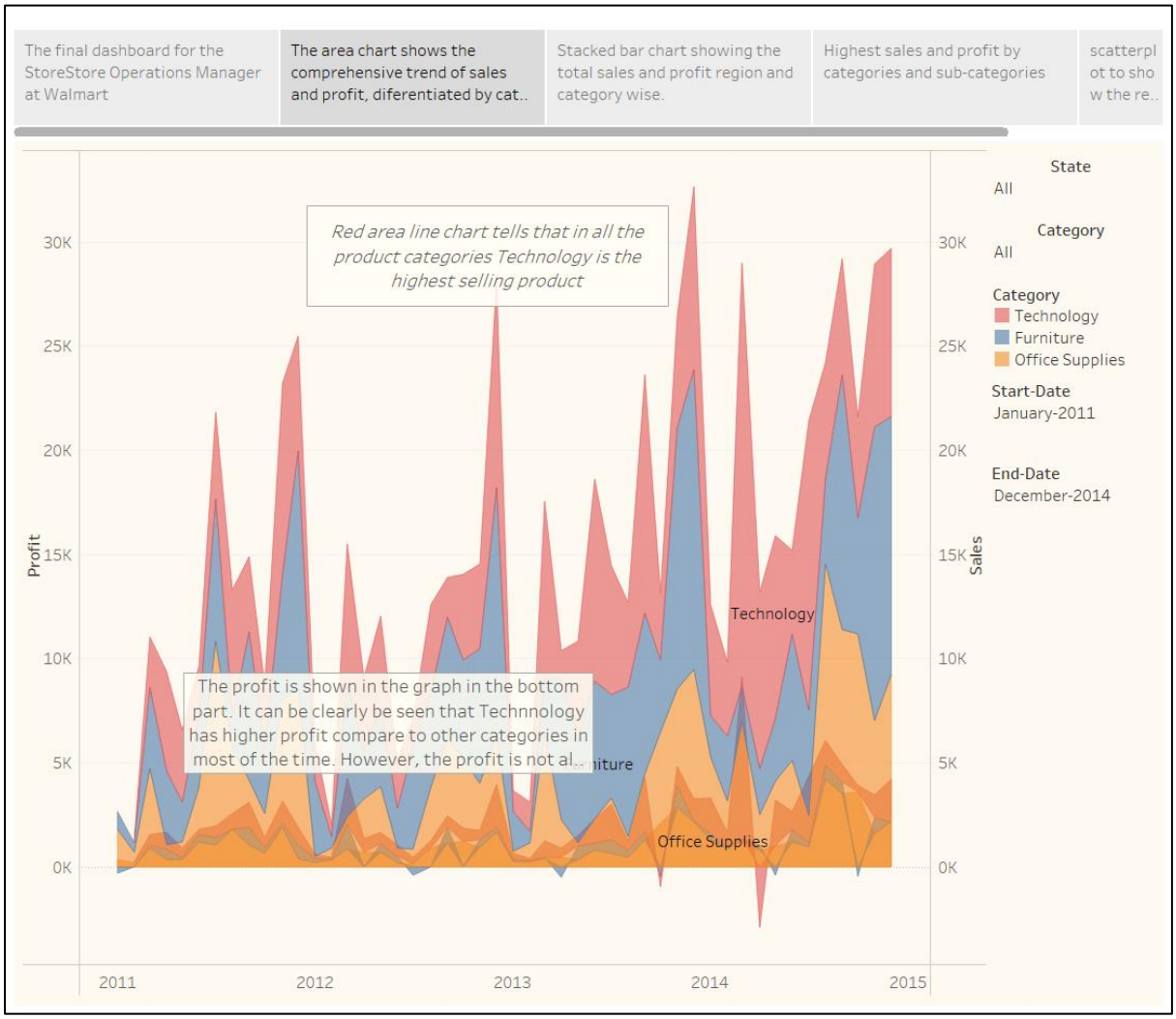
1. Created a new worksheet in Tableau and named it "Sales vs. Profit by Category and Discount Bins."
2. Dragged the "Sales" measure to the Columns shelf to display sales on the x-axis.
3. Dragged the "Profit" measure to the Rows shelf to display profit on the y-axis.

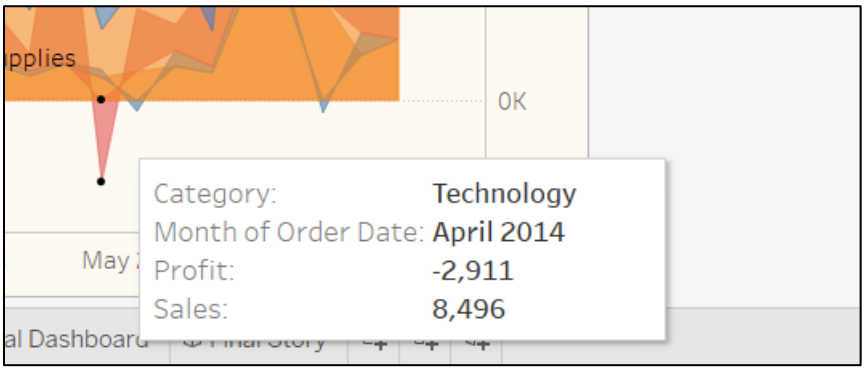
4. Dragged the "Category" dimension to the Color shelf to differentiate data points by category.
5. Created discount bins by right-clicking the "Discount" measure, selecting "Create" > "Bins," defining the bin size (e.g., 0.05 for 5% intervals), and clicking "OK."
6. Dragged the newly created "Discount (bin)" dimension to the Size shelf, determining the size of each circle based on the discount bins.
7. Adjusted the chart type from "Automatic" to "Scatter Plot."
8. Included tooltips to provide additional context for the data points, such as specific sales, profit values, category, and discount information.

The scatter plot that results allow the Store Operations Manager to easily visualise the link between sales and profit across different categories while accounting for discounts imposed. The manager can spot patterns or trends linked to discounts, sales, and profitability by evaluating the size and colour of each circle, and then make informed judgements about pricing tactics, promotions, and overall business performance.

6. Walkthrough

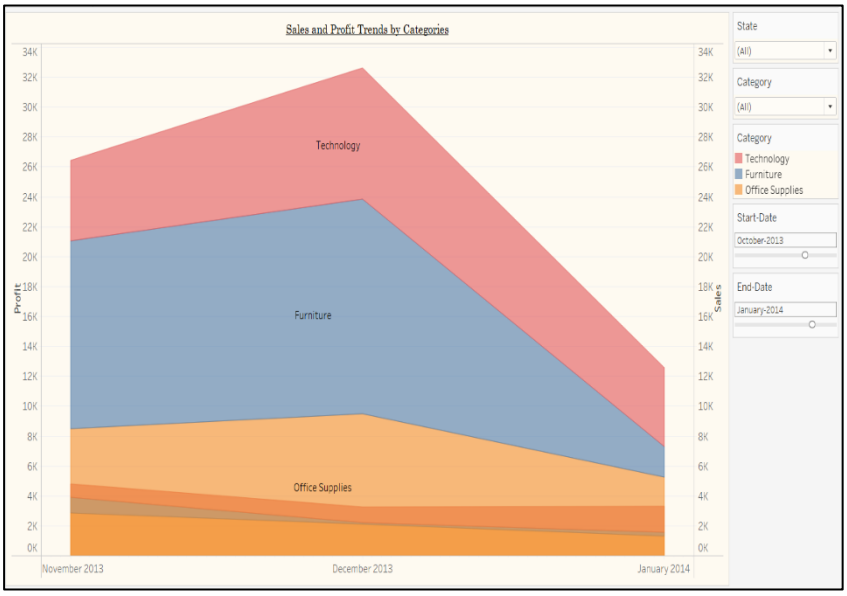
This section will walkthrough the dashboard with the focus of answering the research questions.

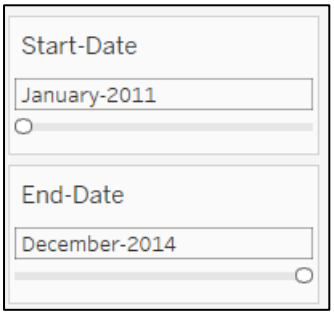


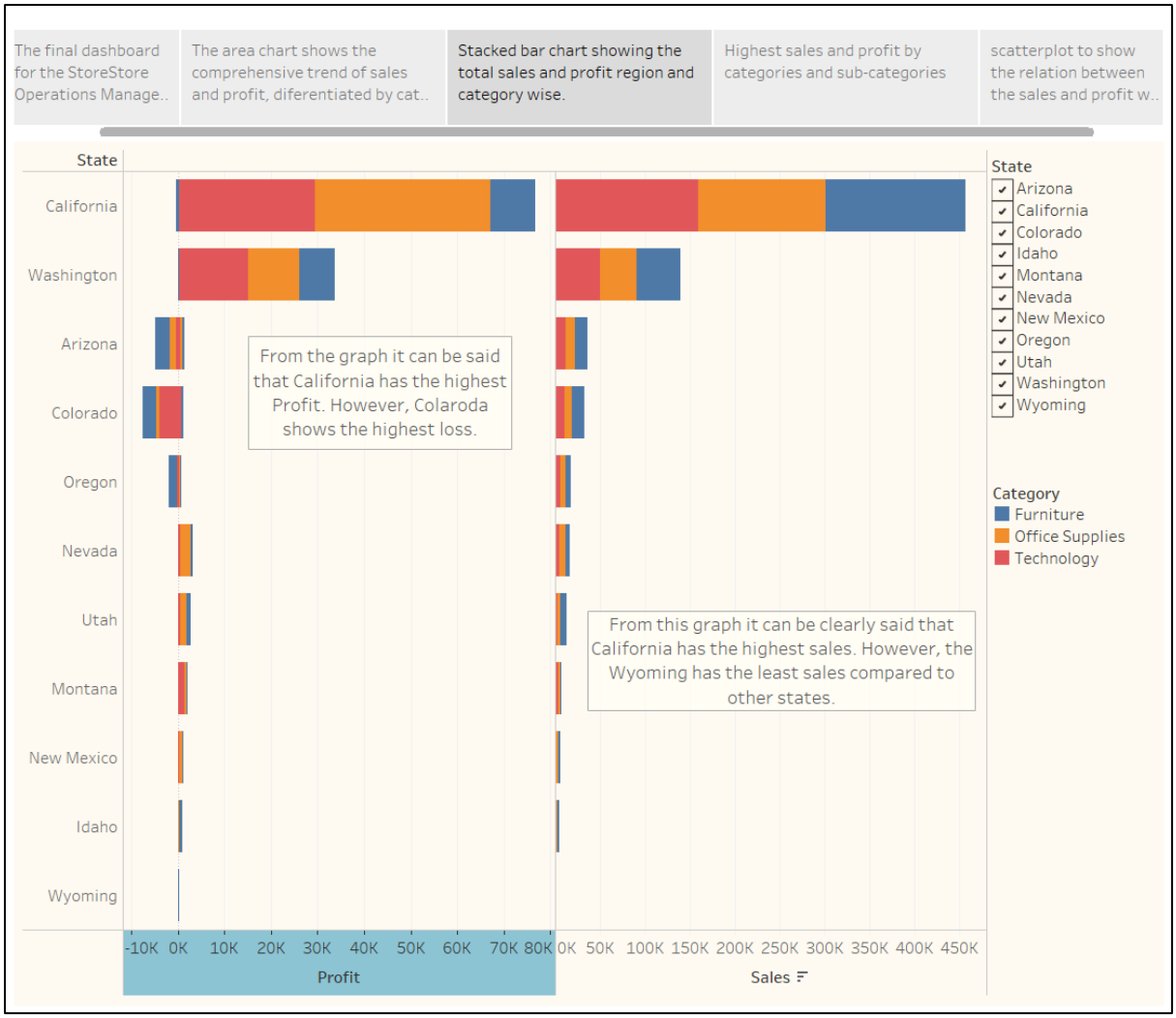


The Q1 asked about the most profitable categories and states and how does the profit and sales change over time. By observing the graph, we can answer half of the question. Over the time, the highest product sold from 2011 to 2015 is from the Technology category. However, the Office Supplies category is the least performed category almost each and every month and year. More details can be found when we hover on the graph, like here we hover on the date April, 2014 and we got to know that the sales were huge though the company had loss it may be possible due to heavy discounts going on in this period.

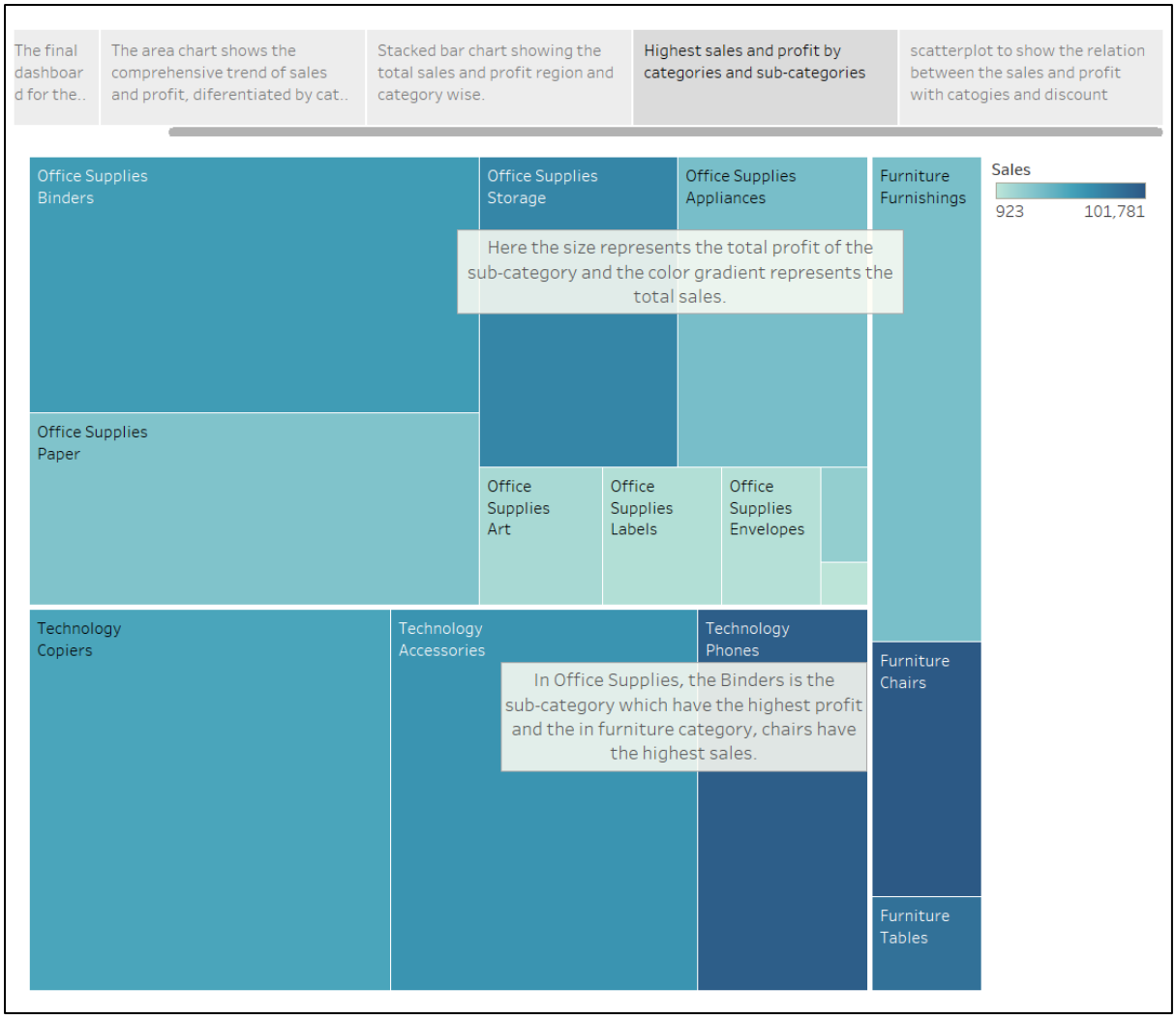
Similarly to see the most sales period then we can more the slider of Start-date and End-date given in the dashboard. By moving the slider, we got to know that period from November 2013- December 2013 has got the highest sales.



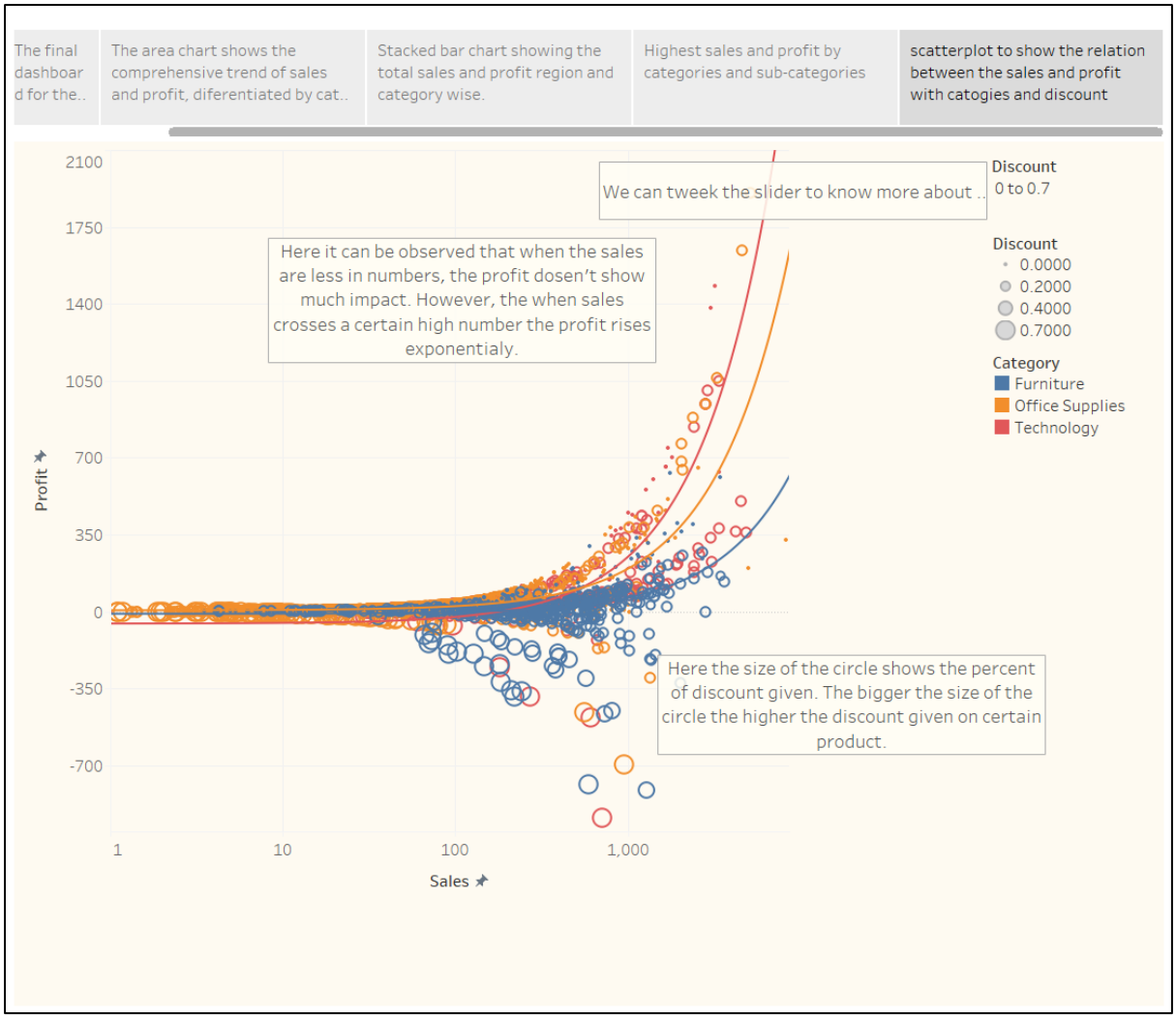




From the Story 2, we can answer second half of the question. We can observe that California has the highest profit. However, Colarido shows the highest loss. If we see the sales side, here also California tops the list in every Category but the Wyoming shows the least sales in comparison to other states.



From this story, it will help user to dig down dipper and will help to answer Q2. The sub-categories with categories which sold the higher products is Binders in Office Supplies. And the Phones sub-category in the Technology shows the highest profit among all the products we have. Now to check the correlation with discounts we need another graph that is scatter plot.



This story answers the remaining part of the Q2 question. Here, the size shows the per cent discount. It can be observed that when the sales are less in number, the profit doesn't have much impact. However, when sales cross a certain high number, the profit rises exponentially. Moreover, it can be clearly seen that as the size of the circle increases the profitability decreases.

7. Conclusion

The analysis conducted using the provided dataset and Tableau visualizations has allowed us to gain valuable insights into Walmart's sales and profit performance. We have answered the research questions and gained a better understanding of the relationship between product categories, states, and discounts.

From our analysis, we found that the Technology category consistently had the highest sales between 2011 and 2015, while the Office Supplies category performed the least. Notably, the period from November 2013 to December 2013 saw the highest sales. Heavy discounts during certain periods, such as April 2014, resulted in higher sales but also led to a loss in profit.

Our investigation into state-wise performance revealed that California consistently topped the list in terms of both sales and profit across all categories, while Wyoming showed the least sales. On the other hand, Colorado experienced the highest loss.

Further analysis of categories and sub-categories demonstrated that Binders in Office Supplies had the highest sales, while Phones in the Technology category had the highest profit. The scatter plot visualization provided insights into the impact of discounts on profitability, showing that profitability decreases as the discount percentage increases, particularly for high sales volumes.

The visualizations and analysis presented in this report provide the Store Operations Manager with a comprehensive understanding of the store's performance across various dimensions. These insights can be used to inform strategic decision-making, such as resource allocation, marketing efforts, and inventory management, ultimately driving better results for the store.