

# Analysis of Customer Sales Data Using Numerous Data Visualization Techniques

Dr T V Rajini Kanth<sup>1</sup>, Dr. D.V.Nagarjana Devi<sup>2</sup>, Ms. Kunduru Neha<sup>3</sup>, Mr T V S Aneesh<sup>4</sup>, Ms. Tanya<sup>5</sup>

<sup>1</sup>Professor & Head of CSE-AI & ML, Sreenidhi Institute of science and Technology(SNIST), Hyderabad,

<sup>2</sup>Ass. Prof. in IIIT Nuziveedu, AP,

<sup>3,5</sup>B Tech CSE Student from VIT,<sup>4</sup>B Tech CSE student, SNIST

<sup>1</sup>[rajinitv@gmail.com](mailto:rajinitv@gmail.com) <sup>2</sup>[devi.duvvuri@rgukt.in](mailto:devi.duvvuri@rgukt.in), <sup>3</sup>[kundurunehasivatmika@gmail.com](mailto:kundurunehasivatmika@gmail.com), <sup>4</sup>[aneeshthatiparti@gmail.com](mailto:aneeshthatiparti@gmail.com),  
<sup>5</sup>[tanya17gupta10.2000@gmail.com](mailto:tanya17gupta10.2000@gmail.com)

**ABSTRACT:** Analyzing customer sales data thrives businesses opportunities. It reveals hidden patterns and trends that drive effective decisions. Analyzing this data gives customer preferences, buying habits, and leads to targeted strategies. In modern business, in order to make effective decisions one must rely on various data visualization techniques for analysis of customer sales information. This includes 24 stunning visualizations techniques and requires inferences on each Graph. Some of these techniques are Doughnut Chart, Exploded Doughnut, line Chart, Area chart, Stacked bar chart, Bar Chart and triangle chart. Apart from that other visualization techniques Pyramid, Depleting, Mind Map, decision tree, Bubble chart, Gantt Chart, Nested Bubble Chart, Funnel chart, Radar Chart, Flow chart, Venn diagram, Concept Map, Bubble Tree map, Scatter plot, Periodic table, and Speedometer chart. In this paper, an attempt was made to analyze customer sales data using various data visualization techniques extensively to bring hidden insights like customer sales patterns etc. Sales data analysis fosters agility. Informed decision-making is the cornerstone of success, and data analysis lays the foundation through data visualization techniques. Data analysis acts as the foundation of success, and data visualization techniques facilitate informed decision-making.

**Keywords:** *Data Visualization Techniques, Customer Sales Data, Funnel chart, hidden insights, Sales patterns.*

## INTRODUCTION:

Dissecting the sales data helps to identify underperforming products or underexploited marketing channels. It's like having a GPS for the business that directs towards profitable paths. Predicting future trends becomes possible through data analysis. It is possible to anticipate seasonal highs and lows, optimize inventory, and prepare for upcoming customer demands, ensuring smooth sailing ahead. Data analysis fosters customer loyalty. By understanding customer needs and preferences, offerings and personalize experiences can be tailored, making them feel valued and boosting retention. Competitive insights emerge from sales data analysis. Data analysis

empowers effective resource allocation. Performance against rivals can be benchmarked, and their strengths identified in order to develop strategies that will allow outperforming them on the market. Efficient resource allocation is entrenched in data analysis. This makes to adapt quickly to market shifts, changing customer preferences, and unforeseen circumstances, ensuring the business remains adaptable and resilient. Every strategic move, from product development to pricing strategies, becomes data-driven and optimized for success. Ultimately, analyzing customer sales data is not just about numbers; it's about understanding your customers, anticipating their needs, and exceeding their expectations. It's the key to unlocking sustainable growth and building a business that thrives. One can identify sections where marketing products have the best return of investment, so that effectiveness is maximized and waste will be minimized.

#### **LITERATURE SURVEY:**

In their article named “Data Visualization in Marketing”, U.N.Umesh et al, [Ref1] assert that they tackled the issues of data explosion for marketing over past two decades. They proposed data visualization techniques as a solution to emphasize the significance of visualization to handle hierarchical data and provide dynamic visual analysis. The key objectives include - demonstrating the need for data visualization in marketing, analyzing hierarchical marketing data to provide managerial insights and comparing the insights derived from visualization with traditional tables. They promote the use of various linked visualizations in order to gain deep insight into multidimensional data in marketing. Robert F. Polichroniou et al. [Ref.2] stated that their paper focuses on designing effective dashboards for sales managers, leveraging data visualization techniques to track performance and identify areas for improvement. Zeel Patel et al. [Ref. 3] claimed in their article that “it offers a general perspective surrounding the role of big data analytics operations on sales marketing strategies; this concept highlights incorporation and usage of interrelated technologies focused toward effectiveness to enhance returns for business strategy, maximize revenues generating through increased organization operation proficiency.” Jeffrey Heer et al. [Ref.4] mentioned in their paper that the aim of this work is to fill out between computational methods for managing high-dimensional data and human capacity interpretation through visual exploration; they also suggest how on efficient analysis gets carried out by themselves such large set of datasets containing so many variables. Furthermore, Cole Nussbaumer Knafllic [Ref.5] says that there is a story in the data, but tools don't know what that

story is. It takes you—the analyst or communicator of the information—to bring that story visually and contextually to life. Ethan Munson [Ref.6] provides a practical and accessible guide for non-technical users to unlock the power of data analysis and visualization through Tableau, enabling them to become more informed and confident decision-makers. Mikael Lindqvist et al. [Ref. 7] stated that their paper explores the use of interactive visualizations to analyze customer journeys, revealing patterns and identifying friction points in the customer experience. Mikael Johansson et al. [Ref.8] stated that their paper demonstrates how self-organizing maps and radial visualizations can be used to segment customers based on their behavior and preferences, leading to more targeted marketing campaigns. Matthias Keller et al. [Ref.9] stated that their paper investigates how data visualization can be used to support predictive sales forecasting and optimize resource allocation in sales teams. Martijn ten Heuvelen [Ref.10] focuses on utilizing a specific data visualization technique, interactive heat maps, to gain insights into website user behavior. Essentially, it's like putting your website under a thermal scanner for user interaction. Jonathan Strayer et al. [Ref.11] demonstrated that how interactive sparklines can provide a user-friendly and effective way to analyze and visualize time series data with forecasting capabilities. This approach has the potential to enhance data exploration, model refinement, and decision-making across various domains. Stephen Few [Ref.12] focused on demonstrating how map visualizations can be a valuable asset for data analysis and exploration. By leveraging the power of spatial representation, maps can help us uncover hidden patterns, gain deeper insights, and communicate complex data effectively. Nathalie Henry et al. [Ref.13] focused on demonstrating how network visualization can be a valuable tool for unlocking the insights hidden within social media data. By understanding the underlying network structure, we can gain a deeper understanding of online communities, information flow, and user behavior. Bing Liu et al. [Ref.14] focused lies in showcasing how combining text analytics and data visualization can unlock valuable customer insights hidden within user reviews. By analyzing and visualizing sentiment, businesses can gain a deeper understanding of customer opinions, improve customer satisfaction, and make data-driven decisions for business growth. James R. Durbin et al. [Ref.15] focus lies in demonstrating how carefully chosen and designed visualizations can unlock the valuable insights hidden within time-series data. By leveraging these techniques, we can gain a deeper understanding of temporal trends, make informed decisions, and effectively communicate complex data insights. Daniel T. Rogers et al. [Ref.16] have focused on a case

study of a pharmaceutical company using data visualization for sales performance management. John A. Quelch et al. [Ref.17] focused on using data visualization to improve marketing campaign performance. Jeanne Liedtka et al. [Ref.18] focused on visualizing the customer experience through customer journey mapping. Peter Alderson [Ref.19] focuses on empowering marketers with practical skills and knowledge to utilize Tableau, is popular data visualization software, for driving successful marketing campaigns based on data insights. Thomas W.S. Lum et al. [Ref.20] focused to provide a clear understanding of its objectives and how it can benefit marketers seeking to leverage the power of R for data-driven marketing strategies. Tomasz Bednarzyk et al. [Ref.21] focused on exploring the use of augmented and virtual reality in data visualization. Zach Lipton [Ref.22] focused on explaining machine learning models and their outputs through data visualization. Catherine D. Manning et al. [Ref.23] focused on using natural language interfaces for exploring data interactively. Scott Murray [Ref.24] focused on making data visualization accessible and engaging for a wide audience through the power of interactive web-based charts and graphs. Ben Shneiderman [Ref.25] focused on future where data visualization is not just a static field but a dynamic and evolving ecosystem. It emphasizes the need to embrace new technologies and trends while addressing ethical and accessibility concerns to ensure data visualization empowers us to understand the world around us better. S. Aishwarya et al. [Ref.26] stated that their paper provides a broad overview of data visualization tools and techniques across various domains, including sales, marketing, and customer data analysis. Muhammad Fadzil Abdullah et al. [Ref.27] stated that their paper demonstrates the integration of data visualization within a customer relationship management system, using 360° virtual tours for enhanced customer engagement.

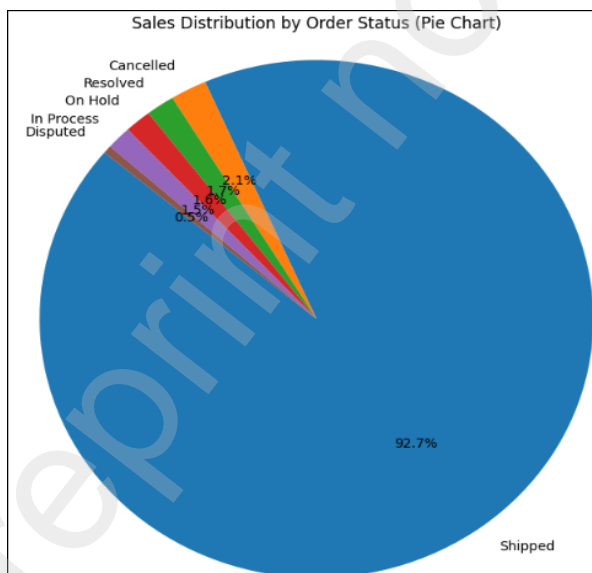
## **METHODOLOGY:**

The customer sales data was analyzed using 24 stunning visualizations techniques and their variants along with their inferences on each Graph. These techniques are Pie chart, Doughnut Chart, Exploded Doughnut, line Chart, Area chart, Stacked bar chart, Bar Chart and triangle chart. Apart from those other visualization techniques namely Pyramid, Depleting, Mind Map, decision tree, Bubble chart, Gantt Chart, Nested Bubble Chart, Funnel chart, Radar Chart, Flow chart, Venn diagram, Concept Map, Bubble Tree map, Scatter plot, Periodic table, and Speedometer chart etc. were used to analyze the customer data to track performance, identify

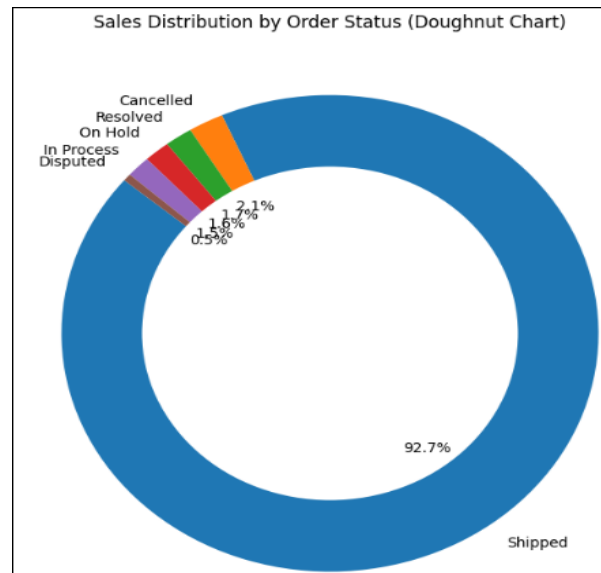
areas for improvement and analyze customer patterns. These visualization techniques revealed many insights of customer sales data.

## RESULTS AND ANALYSIS:

The Customer sales data consists of 2823 instances and 25 attributes. Some of the attributes are namely 'QUANTITYORDERED', 'PRICEEACH', 'SALES', 'ORDERDATE', 'STATUS', 'DEALSIZE', 'PRODUCTLINE', 'ADDRESSLINE1', 'CITY', 'STATE', 'COUNTRY' etc. The types of attributes are float64 (2), int64 (7) and object(16). This data set consists of Order Info, Sales, Customer, Shipping, etc. used for Segmentation, Customer Analytics, Clustering and more and is Inspired for retail analytics. But I found the data set could be useful for Sales Simulation training. In the Status attributes there are Shipped - 2617, Cancelled – 60, Resolved – 47, On Hold – 44, In Process – 41 and Disputed – 14 and is represented by the following pie diagram shown in Fig.1 and Doughnut chart with percentages shown in Fig.2. The Fig.1 pie graph is uses a circle to represent the whole, and slices of that circle, or “pie”, to represent the specific categories of status attribute that compose the whole. The Fig.2 used for visualizing proportions of different categories (like Shipped – 2617, Cancelled – 60, etc.) of the attribute status that make up a whole. Each slice ((like Shipped – 2617, Cancelled – 60, etc.) of the donut represents a category, and the size of the slice reflects its percentage of the total. Shipped category is highest percentage compared to others and disputed is only 14.

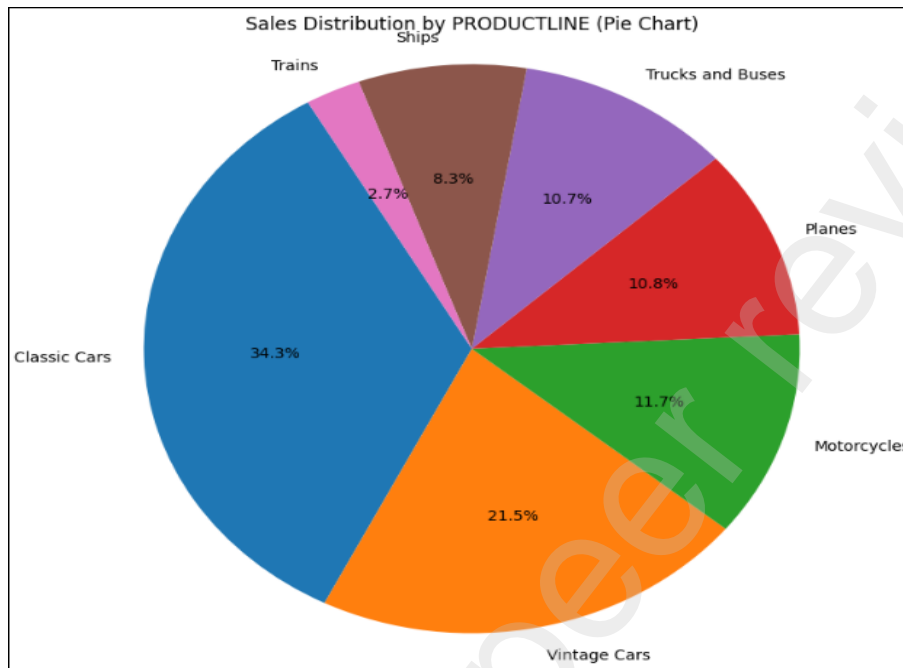


**Fig.1: Pie chart showing sales distribution by order status.**



**Fig.2: Doughnut Chart showing sales by order status**

The Fig.3 pie graph is uses a circle to represent the whole, and slices of that circle, or “pie”, to represent the specific categories (Ships - 8.3%, Planes – 10.8% etc.) of Product Line attribute that compose the whole. In this higher category is a classic car are 34.3%, followed by vintage cars 21.5% and lowest is Trains 2.7%.

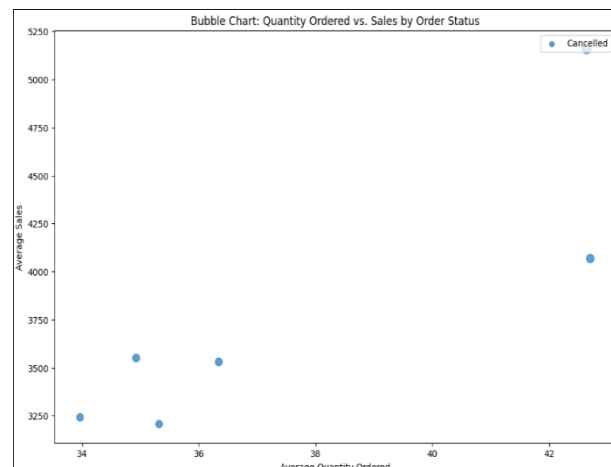


**Fig.3: Sales Distribution by Product Line (Pie Chart)**

**Fig.4** Shows word cloud in which it is observed that Shipped are more and disputed are less under the attribute Status.



**Fig.4: Word Cloud for Order Sales Distribution**



**Fig.5: Bubble Chart-Quantity Ordered vs. Sales by Order Status**

**Fig.5** shows Average Quantity Ordered vs Average Sales Bubble chart graph has more number at lower values. Bubble charts are a visual tool for representing relationships between three

numerical variables. While scatter plots can show the relationship between two variables, bubble charts add a third dimension by using the size of the bubbles to represent a third variable. This makes them useful for understanding complex data sets and identifying patterns that might not be apparent in other types of charts.

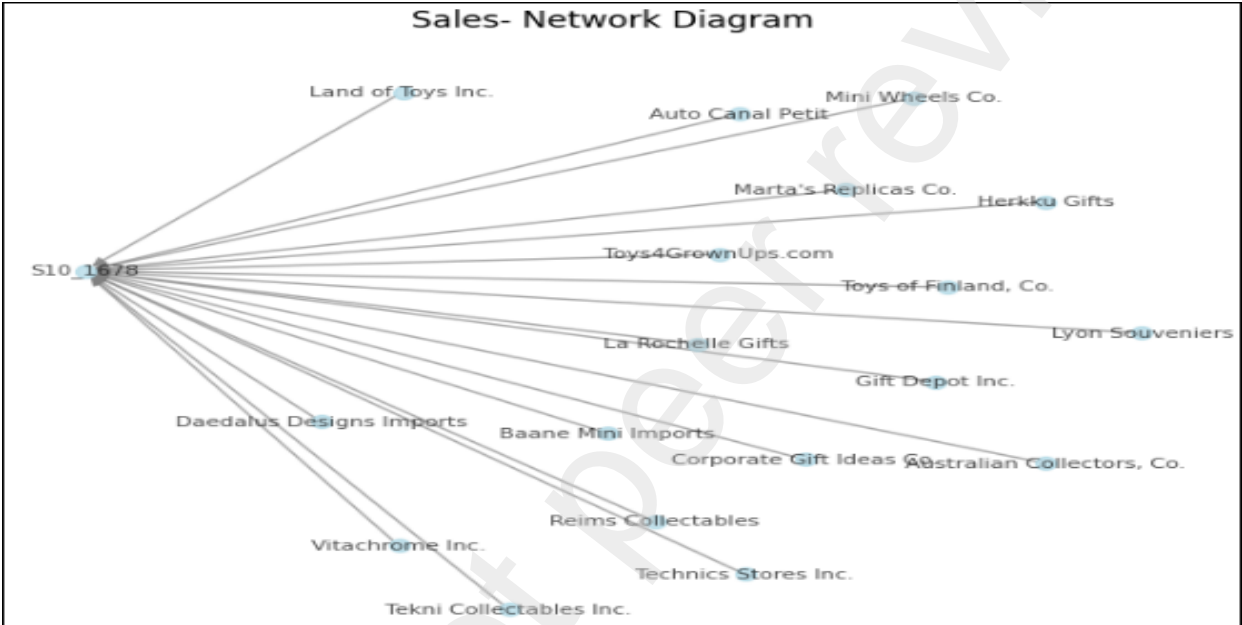


**Fig.6: Nested Bubble Chart: Quantity Ordered vs Sales by order Status and Product Line.**

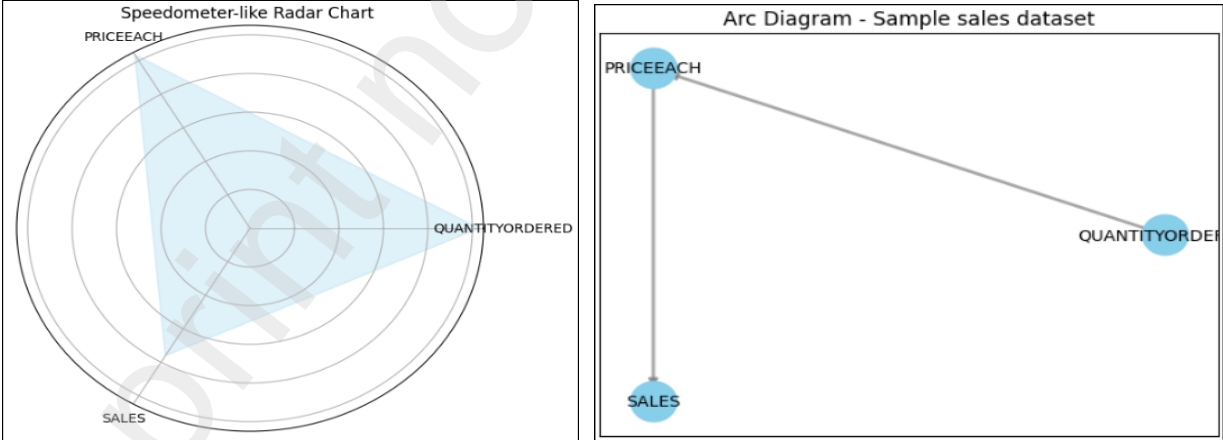
In **Fig.6** shows Nested Bubble Chart with Quantity Ordered vs Sales by order Status and Product Line in which average sales was taken along Y-axis an average Quantity ordered along X-axis. These Nested bubble charts, also known as hierarchical bubble charts or bubble tree maps, are a variation of bubble charts that visualize hierarchical data alongside numerical dimensions. They pack bubbles within bubbles to represent parent-child relationships while still using bubble size to encode a quantitative value.

**Fig.7:** Sales Network Diagrams are visual representations of the relationships and interactions between different entities within a sales organization or process. They can be used to map out various aspects of sales operations. In **Fig.8:** Speedometer – Radar Chart shows QUANTITY ORDERED, SALES AND PRICE EACH. A radar chart, also known as a spider chart, polar chart, or web chart, is a visual representation of multidimensional data that showcases the relative performance of different variables for multiple items or groups. Each variable is represented as an axis radiating from a central point, and the data points for each item are plotted along the corresponding axis at a distance proportional to their value. The resulting lines are then connected to form closed polygons, allowing for easy comparison of the overall performance

profiles of different items. This makes them effective for exploring complex datasets with multiple levels of organization. In **Fig.9: Arc Diagram – Sample sales Dataset** shows an Arc diagram is a special type of network graph that uses a one-dimensional layout of nodes and arcs to represent connections between them. Unlike traditional network diagrams that map connections in two dimensions, arc diagrams arrange nodes along a single line. They then use curved lines (arcs) to represent connections between nodes, with the thickness of the arc often indicating the strength of the relationship.



**Fig.7: Customer Name vs Product Code**



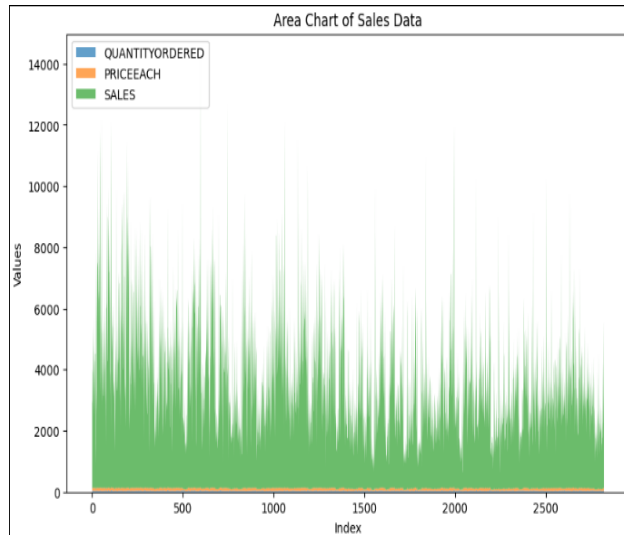
**Fig.8: Speedometer – Radar Chart**

**Fig.9: Arc Diagram – Sample sales Dataset**

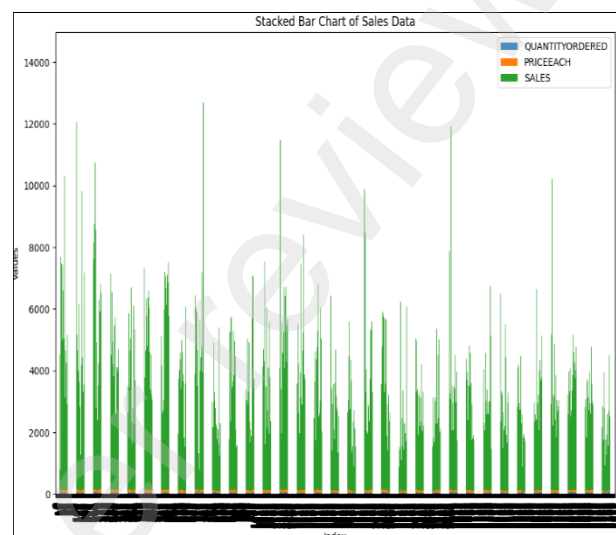
In Fig.10 shows Area chart with index along x-axis and values along y-axis. This Area charts are valuable tools for visualizing data that changes over time or across categories. They combine the features of line charts and bar charts to highlight both overall trends and the contribution of



individual components. In Fig.11 shows Stacked Bar Chart of Sales Data with index along x-axis and values along y-axis. This Stacked Bar Charts: Powerful Tool for Part-to-Whole Comparisons Stacked bar charts are a type of bar chart where multiple bars are stacked on top of each other, each representing a separate category within a main category.

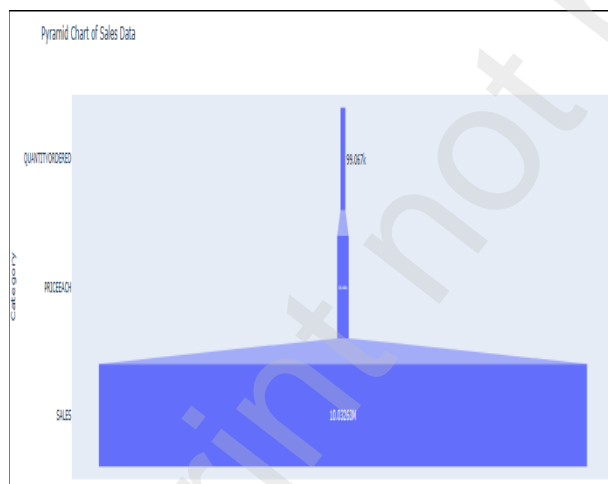


**Fig.10: Area Chart of Sales Data**

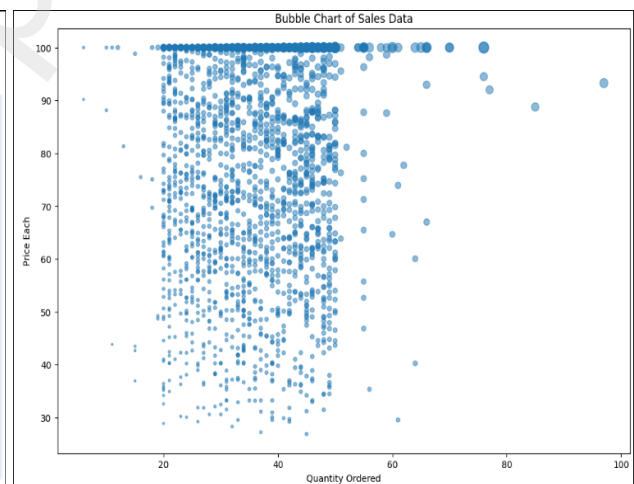


**Fig.11: Stacked Bar Chart of Sales Data**

They're particularly useful for visualizing the composition of a whole and how individual parts contribute to that whole.



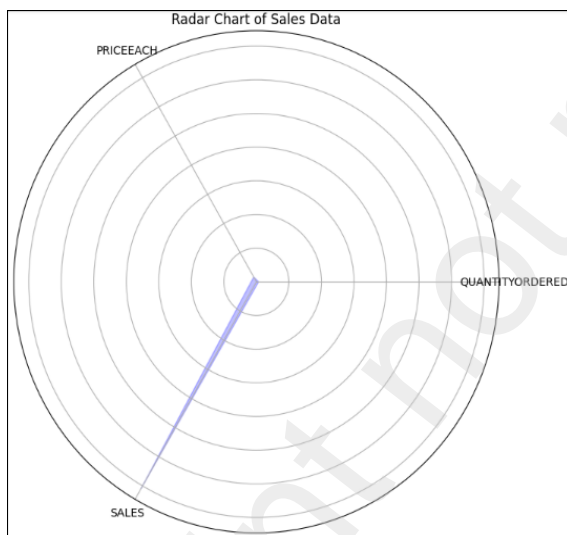
**Fig.12: Pyramid Chart of sales data**



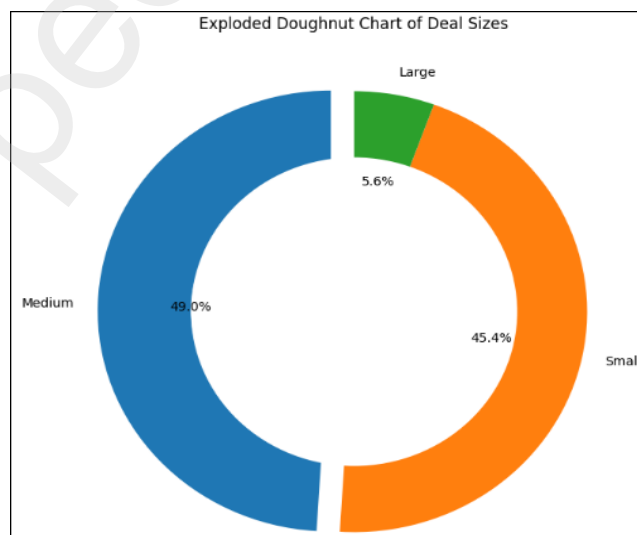
**Fig.13: Bubble Chart of Sales Data**

In Fig.12 shows Pyramid charts, also known as triangle charts, offer a visually compelling way to depict hierarchical structures, proportions, and narratives within your data. In this diagram Sales was taken along x-axis and category along y-axis. In Fig.13 shows Bubble Chart of sales data with QUANTITYORDERED along x-axis, PRICEEACH along y-axis and SALES represented by bubbles. A bubble chart is a type of chart that displays three dimensions of data.

Each data point is represented by a bubble, where the position of the bubble on the x and y axes represents two of the dimensions, and the size of the bubble represents the third dimension. Bubble charts can be used to show relationships between three variables, such as the population, area, and GDP of different countries. In between 20 to 50 more quantities are ordered and higher the price the quantities ordered are more with higher sales. These Bubble charts can be effective for displaying relationships between three variables. Bubble charts can be visually appealing and can help to draw attention to trends in data. Bubble charts can be used to show data for a large number of data points. In **Fig.14** Radar chart was with three attributes **Quantity ordered**, **Price each** and **Sales**. A radar chart, also known as a spider chart, polar chart, or star plot, is a graphical tool for displaying multivariate data – data with multiple quantitative variables. Imagine a web, where each spoke represents a variable, and the distance from the center along that spoke reflects the value of that variable. Data points are then plotted like points on the web, connected by lines to form closed shapes called polygons. Sales reduced as the higher values of Quantities ordered and Price each.



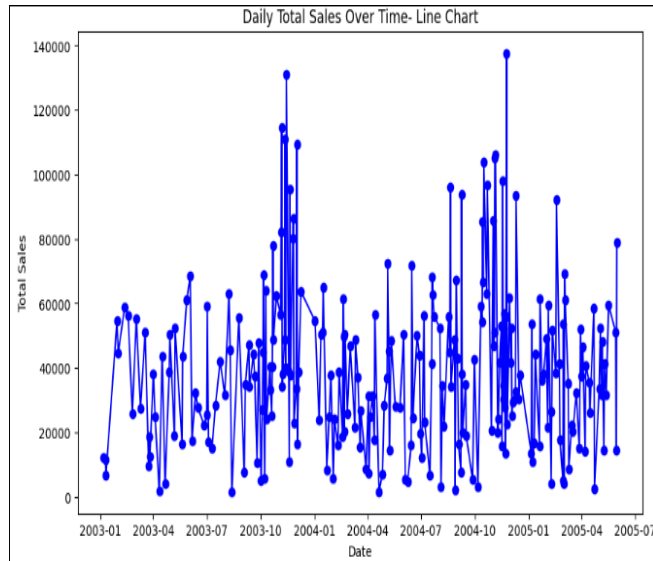
**Fig.14: Radar Chart of Sales data**



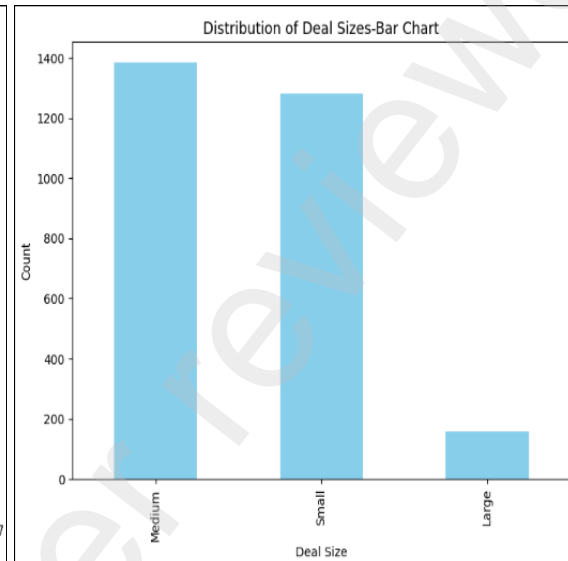
**Fig.15: Exploded Doughnut Chart of Deal Sizes**

In Fig.15 Exploded Doughnut Chart Deal size it shows that the medium Deal size is high followed by Small Deal size. An exploded doughnut chart is a variation of the classic doughnut chart, where one or more segments are "pulled out" from the center, visually emphasizing their significance. Imagine a regular doughnut, but some slices are proudly standing apart, demanding attention. When there are a few key categories you want to emphasize within a larger dataset. When dealing with numerous categories that might otherwise be too close together in a regular doughnut. When aiming for a visually engaging and attention-grabbing presentation of the data.

The Fig.16 shows Daily Total Sales over Time - Line Chart with Date along x-axis and Total Sales along y-axis. It reveals that November month sales are high irrespective of year and low sales in the month of April.



**Fig.16: Daily Total Sales over Time- Line Chart**



**Fig.17: Distribution of Deal sizes- Bar Chart**

The Fig.17 shows the Distribution of Deal size – Bar Chart with Deal size along x-axis and count along y-axis. This reveals that Medium sales are high followed by small sales and lowest is large sales. The Confusion Matrix is shown in below Table-1a), Table-1b) and Table-1c) with accuracy 1.00.

TABLE-1 a)

	QUANTITYORDERED	PRICEEACH	SALES
QUANTITYORDERED	41	0	0
PRICEEACH	0	272	0
SALES	0	0	252

Classification Report: TABLE-1b)

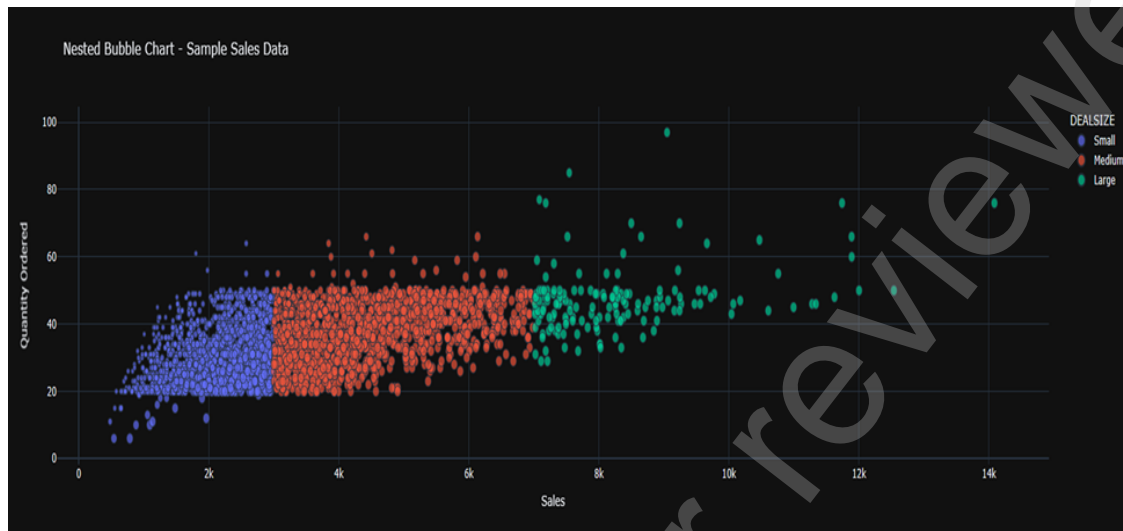
	precision	recall	f1-score	support
0	1.00	1.00	1.00	41
1	1.00	1.00	1.00	272
2	1.00	1.00	1.00	252

TABLE -1c)

accuracy			1.00	565
macro avg.	1.00	1.00	1.00	565
weighted avg.	1.00	1.00	1.00	565

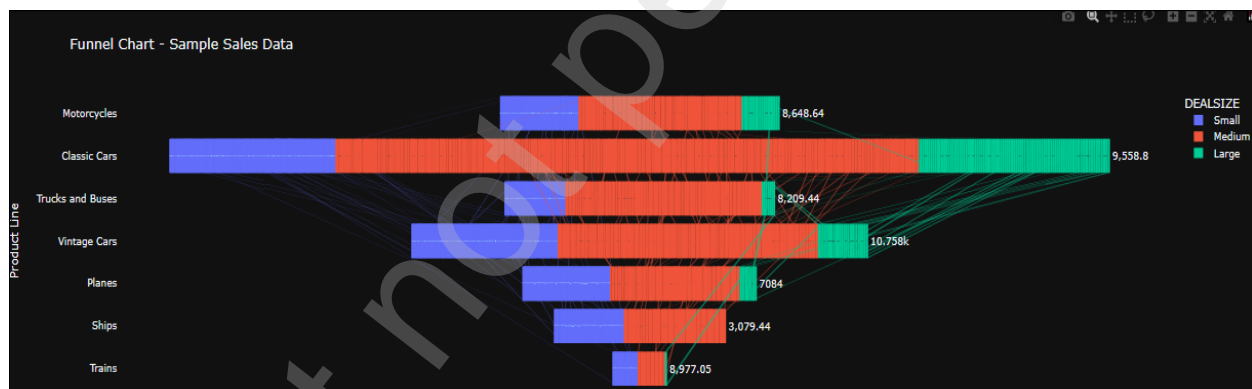
Micro averaging considers overall counts, potentially masking poor performance on smaller classes. A method for calculating overall performance metrics in multi-class classification problems. It treats all classes equally, regardless of their frequency in the dataset. It's calculated by first calculating the metric (e.g., precision, recall, F1-score) for each individual class, and then

taking the simple average of those scores. Macro avg. is just one way to evaluate model performance.



**Fig.18: Nested Bubble Chart – Deals**

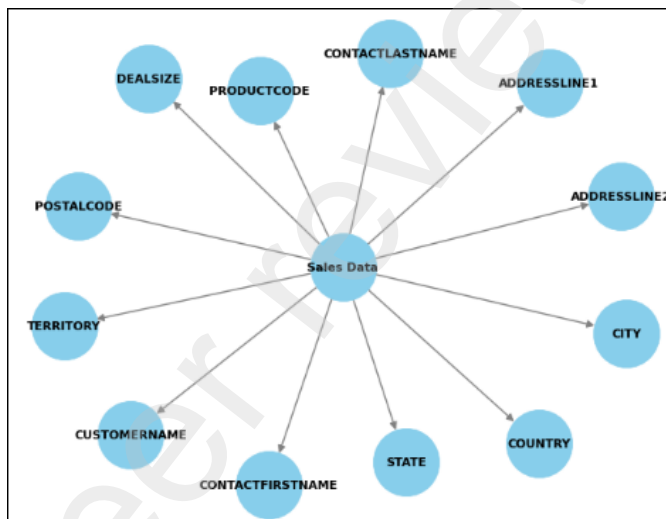
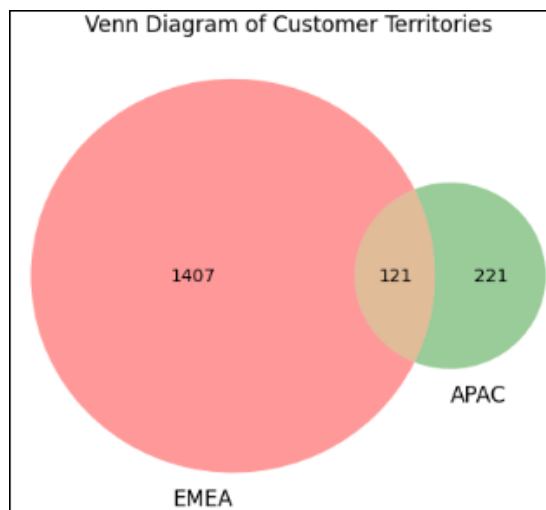
The Fig.18 shows nested bubble chart for Deal size and nested bubble charts were discussed in Fig. 6 and here sales was taken along x-axis and Quantity ordered along y-axis. Deal size small followed by medium followed by large as we move from left to right.



**Fig.19: Funnel Chart – Sample sales data**

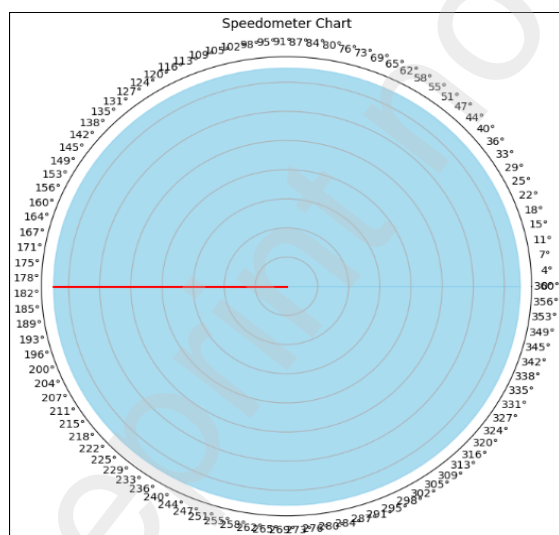
Fig.19 shows Funnel chart having sales with Deal sizes along x-axis and Product line along y-axis. Funnel charts are a valuable tool for understanding and optimizing processes, especially when dealing with sequential stages and conversion rates. By leveraging their visual clarity and intuitive design, you can effectively identify areas for improvement and drive better outcomes in your business or project. In this, highest deal size is classic Cars in which medium sales size is maximum followed by vintage Cars and lowest is Trains. The Fig.20 shows the Territories with EMEA – 1407, APAC – 221 and Japan – 121. In which Japan lies in both the Territories of EMEA (Europe, Middle East, and Africa) and APAC (Asia and Pacific). Venn diagrams are visual tools used to represent the relationships between sets of things. Think of them as colorful

circles on a canvas, where overlapping zones depict shared characteristics and separate areas showcase unique traits. These seemingly simple circles pack a powerful punch, helping us grasp complex relationships with ease. The graph Fig.21 shows 12 attributes network chart. Network charts, also known as graph visualizations, weave connections between data points, offering a unique perspective on relationships and patterns.

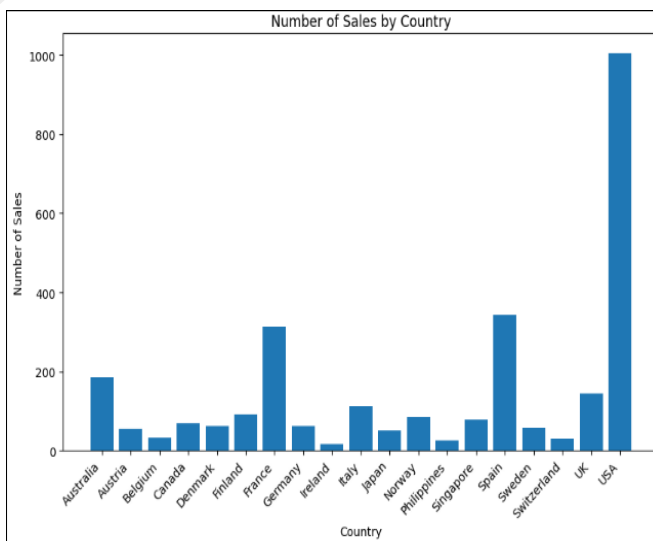


**Fig.20: Venn Diagram of customer Territories**    **Fig.21: Networks Graph**

Imagine a map where nodes (points) represent entities like people, objects, or concepts, and links (lines) symbolize connections between them. By analyzing these intricate webs, one can unveil hidden insights and glean invaluable knowledge.



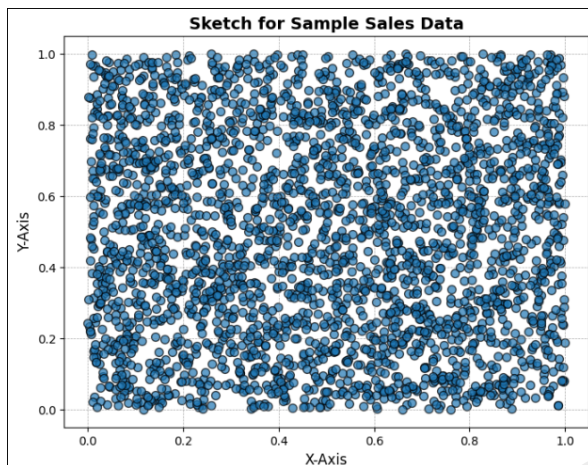
**Fig.22: Speedometer Chart**



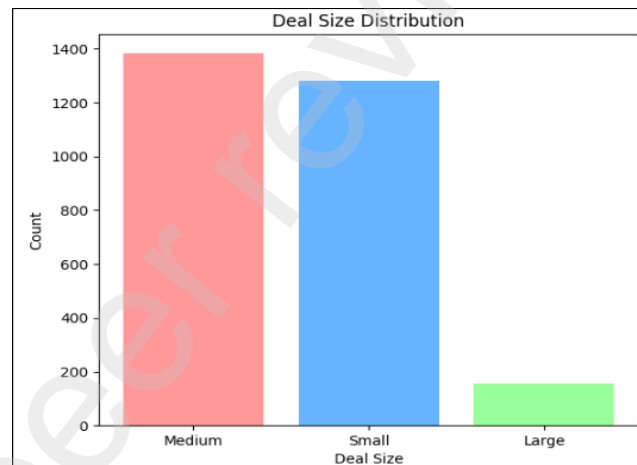
**Fig.23: Bar plot for Number of sales by country**

The Fig.23 shows Bar plot for Number of sales by country with country along x-axis and Number of Sales along y-axis. There is a highest sale in USA and lowest sales are in Ireland. Bar

plots are one of the most fundamental and versatile tools in data visualization. They use rectangular bars to represent categorical data and their corresponding numerical values. Imagine a set of colorful towers rising from a horizontal axis, each bar's height corresponding to the value it represents. This simple yet powerful visual language makes bar plots a popular choice for a variety of situations. The graph represented by **Fig.24** depicts sample sales data and the **Fig.25** shows Deal size distribution in which Deal size was taken along x-axis and count along y-axis. It indicates medium sales are high followed by small and lowest sales by large.

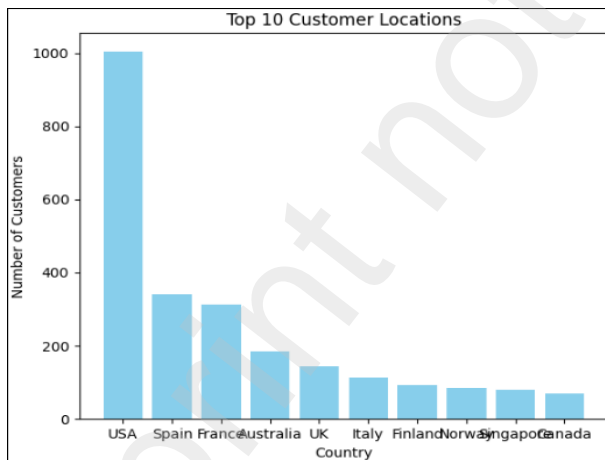


**Fig.24: Sketch for Sample Sales Data**

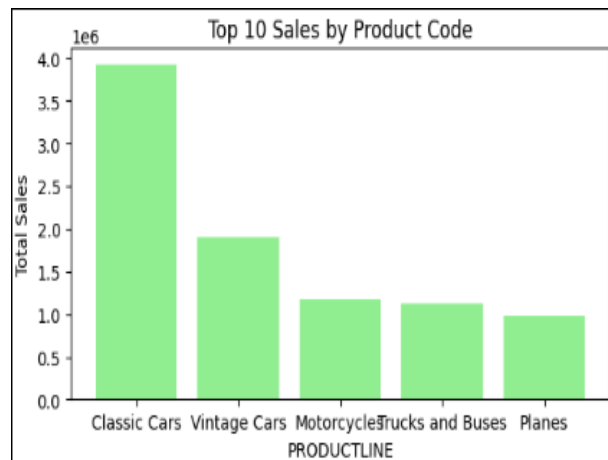


**Fig.25: Deal size Distribution**

The **Fig.26** reveals that USA is the first in top 10 customer locations followed by Spain, France and India occupies 10<sup>th</sup> position.



**Fig.26: Top 10 customer locations**

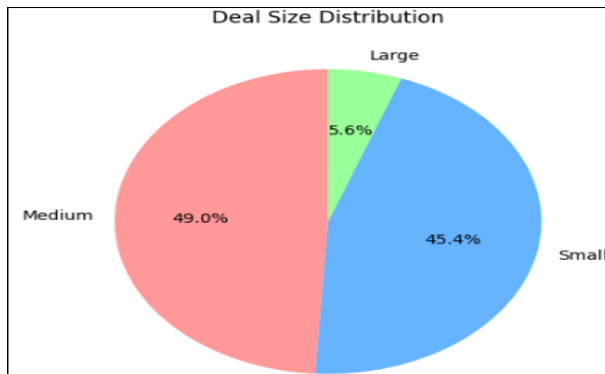


**Fig.27: Top 10 sales by Product code**

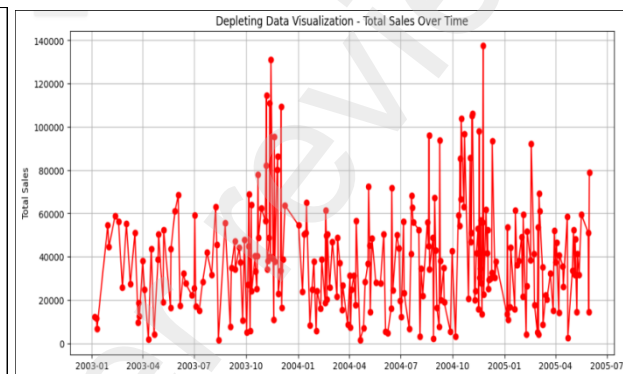
The **Fig.27** shows top 10 sales in which the highest sales product is Classic cars followed by vintage cars and lowest sales by planes sales. The **Fig.28** shows deal size with three categories large, medium and small in which medium is more followed by medium. The **Fig.29** shows depleting data visualization with Total sales over time period it indicates highest sales are in



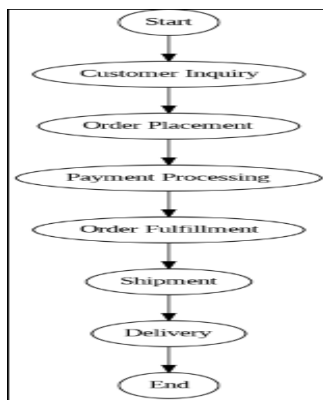
2003 December and 2004 December. The **Fig.30** shows sales flow chart with sequence of steps like start – customer inquiry – order placement – payment processing – order fulfillment – shipment – delivery – end. The **Fig.31** shows Triangle Chart with Sample Sales Data in which quantity ordered, sales and price each were considered. The **Fig.32** shows Periodic Table with Sample Sales Dataset in which order number, product line and customer name were considered. The **Fig. 33** shows digraph visualization of sample sales data.



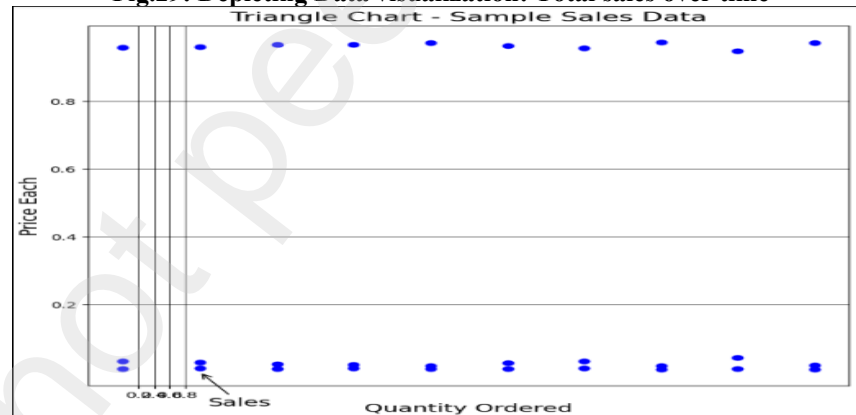
**Fig.28: Deal size distribution**



**Fig.29: Depleting Data visualization: Total sales over time**



**Fig.30: Sales Flow Chart**



**Fig.31: Triangle Chart – Sample Sales Data**

The **Fig.34** shows Iconographic Representation – sample sales data. The **Fig.35** shows Sales Timeline i.e. from Feb2003 to April 2005 and the interpretation from the graph is highest sales in November2003 and November2024 and lowest sales are in Febrauary2005. The **Fig. 36** shows Bubble Tree map Sample Sales data in which it indicates Classic Car sales are highest followed by vintage cars, and lowest sales in planes segment. The **Fig.37** shows Sankey Diagram Samples Sales data. The **Fig.38** shows Gant chart – Sample sales data.

Periodic Table with Sample Sales Dataset		
ORDERLINENUMBER	PRODUCTLINE	CUSTOMERNAME
0	2 Motorcycles	Land of Toys Inc.
1	5 Motorcycles	Reims Collectables
2	2 Motorcycles	Lyon Souveniers
3	6 Motorcycles	Toys4GrownUps.com
4	14 Motorcycles	Corporate Gift Ideas Co.
5	1 Motorcycles	Technics Stores Inc.
6	9 Motorcycles	Daedalus Designs Imports
7	1 Motorcycles	Herkku Gifts
8	2 Motorcycles	Mini Wheels Co.
9	14 Motorcycles	Auto Canal Petit

Fig.32: Periodic Table with Sample Sales Dataset

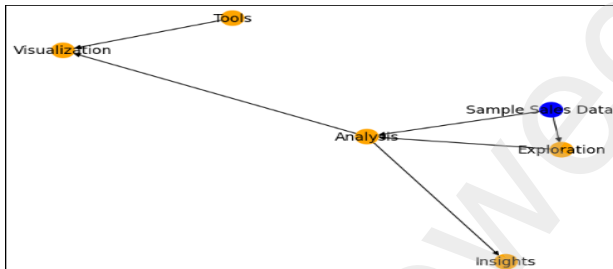


Fig.33: Digraph – visualization

Iconographic Representation - Sample Sales Data	
<=>	Deal Size
{?}	Order Status
?	Status and Performance
(+)	Customer Contacts
+	Customer Segmentation
*@*	Customer Analysis
/^	Product Trends
=?	Product Sales
^!^	Product Analysis
^@^	Global Distribution
^	Regional Sales
@	Geographic Analysis
/	Seasonality
/^	Trends
\$	Date Analysis
<!	Sales Overview

Fig.34: Iconographic Representation – sample sales data

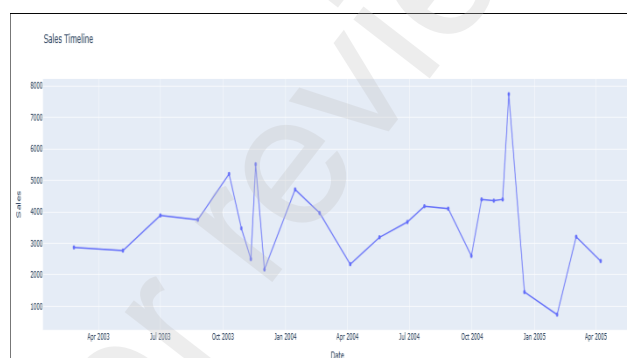


Fig.35: Sales Timeline



Fig.36: Bubble Tree map Sample Sales data

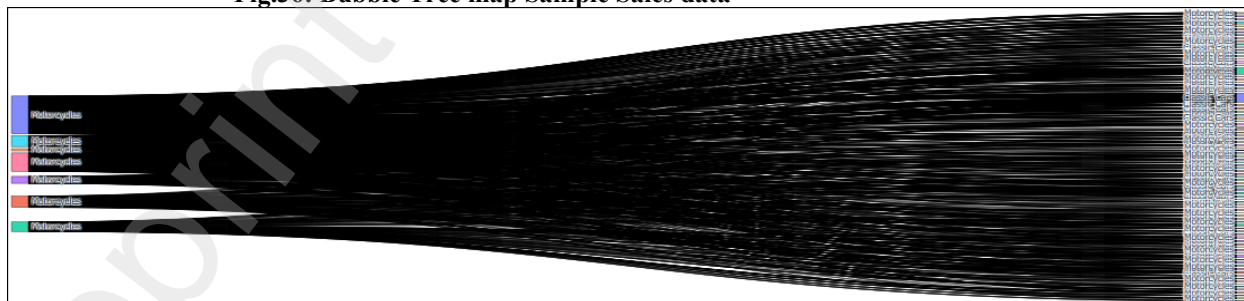
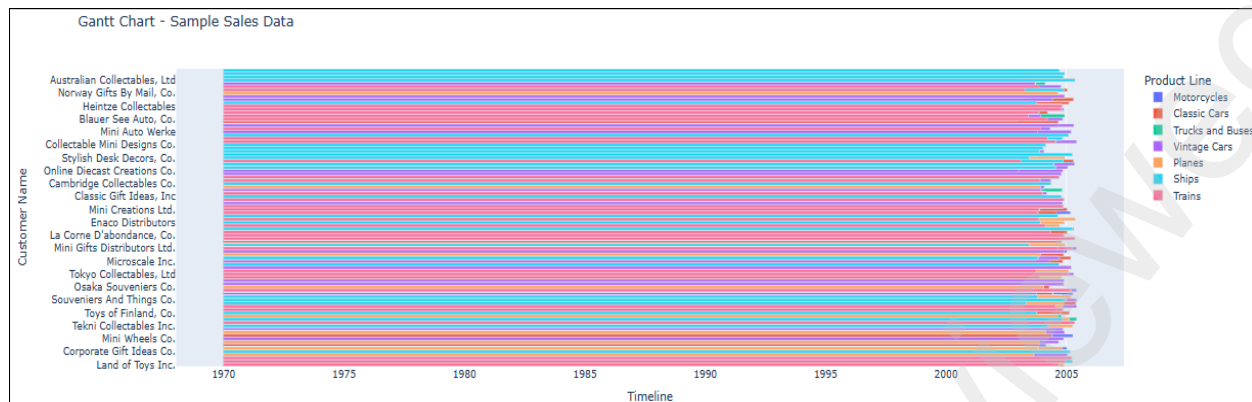


Fig.37: Sankey Diagram Samples Sales data





**Fig.38: Gant chart – Sample sales data**

## CONCLUSION:

The customer sales dataset is designed for segmentation, customer analytics, and clustering in retail, serving as a valuable resource for sales simulation training. Visualizations through pie and doughnut charts highlight that the shipped category holds the highest percentage, with only 14% disputed instances. In product line distribution, classic cars account for 34.3%, vintage cars for 21.5%, and trains for 2.7%. It is observed that Classic Car are the top-selling category across all categories. Word cloud analysis reveals more shipped instances and fewer disputed cases in the status attribute. Various visualizations, including a stacked bar chart, pyramid charts, and bubble charts, offer insights into changing data, relationships, and part-to-whole comparisons in sales data. Sales time lines shows peak total sales are in the month of December 2003 and December 2004. Deal sizes are categorized as large, medium, and small, in which medium-sized deals are dominating. In the top 10 customer sales locations, USA is the top leading country, followed by Spain, France, and India ranked 10<sup>th</sup>. In the Territory data EMEA with 1407, APAC with 221, and Japan with 121 were highlighted indicating Japan's presence is in both EMEA and APAC territories. The dataset is further explored through charts like radar charts, exploded doughnut charts, bar charts for deal sizes and sales, Venn diagrams, network charts, and a comprehensive sales flow chart outlining the sequential steps in the sales process.

**Acknowledgement:** Customer sales data was considered from Kaggle <https://www.kaggle.com/datasets/kyanyoga/sample-sales-data/>.

## Conflict of Interest Statement:

On behalf of all authors, the corresponding author states that there is no conflict of interest.

## REFERENCES:

1. U.N. Umesh, Martin Kagan, "Data Visualization in Marketing", DOI: 10.15640/jmm.v3n2a4.

2. Robert F. Polichroniou, Dimitris N. Chaliampoulakis, "Exploring Sales Performance Data: A Design Study of Dashboard Systems for Sales Managers",
3. Zeel Patel et al., "A Survey Paper on Big Data Analytics in Sales and Marketing" © 2022 IJCRT | Volume 10, Issue 11 November 2022 | ISSN: 2320-2882.
4. Jeffrey Heer, Michael Stone "Visual Analysis of Large, High-Dimensional Data Sets".
5. Cole Nussbaumer Knafllic, "Storytelling with Data: A Guide for Business Storytelling".
6. Ethan Munson, "Tableau for Business Intelligence: Data Analysis for Managers and Executives"
7. Mikael Lindqvist, Mikael Enquist, "Interactive Visualizations for Exploring Customer Journey Data".
8. Mikael Johansson, Magnus Andersson "Customer Segmentation Using Self-Organizing Maps and Radial Visualizations".
9. Matthias Keller, Markus Appelrath "Predictive Sales Force Management with Data Visualization".
10. Martijn ten Heuvelen, "Interactive Heatmaps for Visualizing Website Traffic".
11. Jonathan Strayer, Tamara Munzner, "Time Series Forecasting with Interactive Sparklines".
12. Stephen Few, "The Power of Maps: How Map Visualizations Can Reveal Patterns and Trends".
13. Nathalie Henry, Jean-Daniel Fekete, "Network Visualization for Analyzing Social Media Data".
14. Bing Liu, Mingqing Hu, "Text Analytics for Customer Insights: Extracting and Visualizing Sentiment from User Reviews".
15. James R. Durbin, Ronald A. Johnson, "Visualizing and Understanding Change: Techniques for Time-Series Data".
16. Daniel T. Rogers, Thomas W.S. Lum, "Data Visualization for Sales Performance Management: A Case Study of a Pharmaceutical Company".
17. John A. Quelch, Lisa N. Ferguson, "Using Data Visualization to Improve Marketing Campaign Performance".
18. Jeanne Liedtka, Tim Ogilvie, "Customer Journey Mapping: Visualizing the Customer Experience".
19. Peter Alderson, "Data-Driven Marketing with Tableau: A Practical Guide".
20. Thomas W.S. Lum, Garrett L. Curry, "R for Marketing Analytics: Techniques and Case Studies".
21. Tomasz Bednarzyk, Wojciech R. Smoktunowicz, "Augmented Reality and Virtual Reality for Data Visualization".
22. Zach Lipton, "Data Visualization for Machine Learning: Explainable AI and Model Interpretation".
23. Catherine D. Manning, Daniela Rus, "Natural Language Interfaces for Data Exploration".
24. Scott Murray, "Data Visualization for the Masses: Interactive Web-Based Charts and Graphs".
25. Ben Shneiderman, "The Future of Data Visualization: Trends and Challenges".
26. S. Aishwarya et al., "An Introduction to Data Visualization Tools and Techniques in Various Domains".
27. Muhammad Fadzil Abdullah et al., "EyEstate: Customer Relations Management System with Sales Data Visualization and 360° Virtual Tour Technology".