EDA / Descriptive Statistics

## Introduction:

Medical inventory optimization is essential for efficient healthcare management. Addressing challenges related to inventory management by analyzing transactional data to identify patterns and opportunities for improvement. With a focus on reducing bounce rates and minimizing inventory costs, our goal is to enhance patient satisfaction and operational efficiency in healthcare facilities. Good visualization is important to get clear insights from this vast amount of data. Through data analysis and visualization, we aim to provide actionable insights, dashboards and recommendations to optimize inventory management processes, ultimately improving patient care and resource utilization in the healthcare industry.

## Overall design strategy:

As the data contains over 14,000 records of transactional data covering a period of one year, from **January 1st, 2022, to December 31st, 2022**. These records encapsulate various aspects of medical inventory management, including sales, returns, final sales, final costs, and more. Through comprehensive analysis of this data, we aims to extract valuable insights into inventory utilization, operational efficiency, and financial performance within the healthcare facility.

Overall, we have used both the raw data and a custom SQL output to run the visualization. After the data is prepared, visualization color codes are identified to show the Sales and Returns metrics for each of the dashboard. Overall, we have use Green for Return and Red for Sales in the worksheet.

## Data Overview:

Data is extracted from Secondary source which contains medical inventory statistics and historical data. Sales and returns data will be collected to use in the visualization. Around **14,000 records** are extracted for medical inventory level data and later **Final Sales**, Return Quantity, **Final Cost**, Date of bill, Drug Name, Sub category, Specialization, Formulation, RtnMRP data is extracted.

The dataset comprises **14 columns and 14,219 rows**, capturing a wide array of information regarding pharmaceutical sales transactions. Among the columns, **seven are qualitative and discrete** in nature, including descriptors such as the type of sales, patient ID, specialization, formulation, drug name, and subcategories. Additionally, the **department column**, while **qualitative, is continuous**, providing further granularity in classifying transactions based on their departmental association.

On the quantitative front, the dataset encompasses three discrete variables—date of bill, quantity, and return quantity. Complementing these discrete metrics are **three continuous variables**—final cost, final sales, and RtnMRP—providing detailed numerical data on the financial aspects of each transaction.

## statistical insights:

Provide a summary table containing the key insights from the SQL analysis :

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Metric** | **Quantity** | **Return Quantity** | **Final Cost** | **Final Sales** | **Returned MRP** |
| Mean | 2.2317 | 0.2920 | 124.8239 | 234.0383 | 29.1326 |
| Median | 1.0000 | 0.0000 | 53.65 | 86.424 | 0.0000 |
| Mode | 1 | 0 | 49.352 | 0 | 0 |
| Variance | 26.3360 | 2.7003 | 216007.85 | 450560.41 | 33218.35 |
| Standard Deviation | 5.1319 | 1.6433 | 464.7664 | 671.238 | 182.259 |
| Min | 0 | 0 | 40 | 0 | 0 |
| Max | 150 | 50 | 33178 | 39490 | 8014 |
| Skewness | 11.3399 | 17.1705 | 34.5046 | 21.0045 | 15.7956 |
| Kurtosis | 183.0905 | 412.2706 | 2028.1537 | 951.1888 | 406.3615 |

## Business insights:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Metric** | **Quantity** | **Return Quantity** | **Final Cost** | **Final Sales** | **Returned MRP** |
| Average | 2.23 units | 0.29 units | $124.82 | $234.04 | $29.13 |
| Median | 1 unit | 0 units | $53.65 | $86.42 | $0 |
| Common Value | 1 unit | 0 units | $49.35 | $0 | $0 |
| Variation | Moderate | Low | High | High | High |
| Range | 0 - 150 units | 0 - 50 units | $40 - $33,178 | $0 - $39,490 | $0 - $8,014 |

* **Quantity**: Customers typically purchase around 2 units of medication per transaction, indicating a moderate demand.
* **Return Quantity**: Returns are infrequent, with the majority of transactions not involving any returned items, suggesting a low rate of dissatisfaction or product issues.
* **Final Cost**: The average cost of medication per transaction is $124.82, but costs vary significantly, highlighting the need for pricing strategies tailored to different customer segments.
* **Final Sales**: Sales figures are positively skewed, indicating that a few high-value transactions contribute significantly to total sales.
* **Returned MRP**: Most returned items have a retail price close to zero, indicating that returned products often have little resale value, which may impact profitability and inventory management decisions.

Data Preprocessing:

* Duplicates were identified and removed to ensure data integrity.
* Missing values were addressed using appropriate techniques:
* Imputing common values (mode) for 'Formulation' and 'DrugName' columns.
* Filling missing values in categorical columns with their respective modes.
* Replacing outliers in numerical columns with the median value (to handle potential skewness).

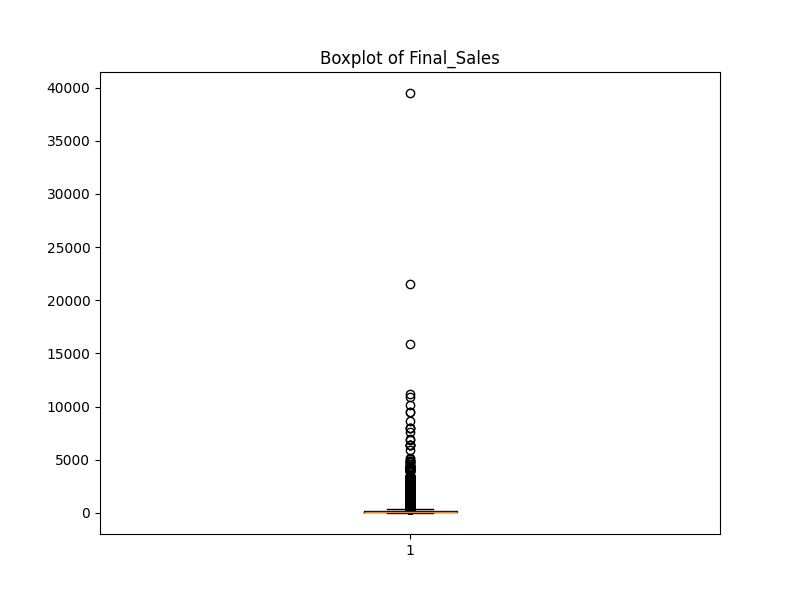


Fig-1 Sales Distribution

This graph shows the spread of final sales for medications in dataset. The center line indicates the typical sale, while the box highlights the range where most sales fall.

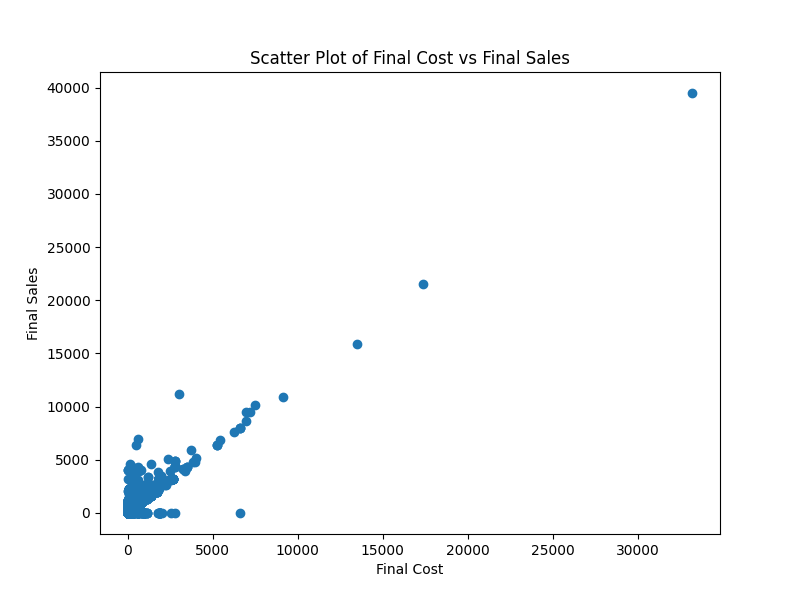


Fig-2 Cost vs Sales Distribution

This scatter plot depicts the relationship between final cost and final sales of a product. The horizontal axis (X-axis) represents the final cost. The vertical axis (Y-axis) represents the final sales price

Data Analysis:

The purpose of this Data Analysis is to provide a comprehensive understanding of the pharmaceutical sales dataset and derive actionable insights for business clients. This analysis aims to uncover patterns, trends, and relationships within the data to support strategic decision-making and business optimization.

**Sales and Return Trends:**

Line charts were created to visualize daily sales and return trends over time. This can help identify seasonal variations or promotional effects.

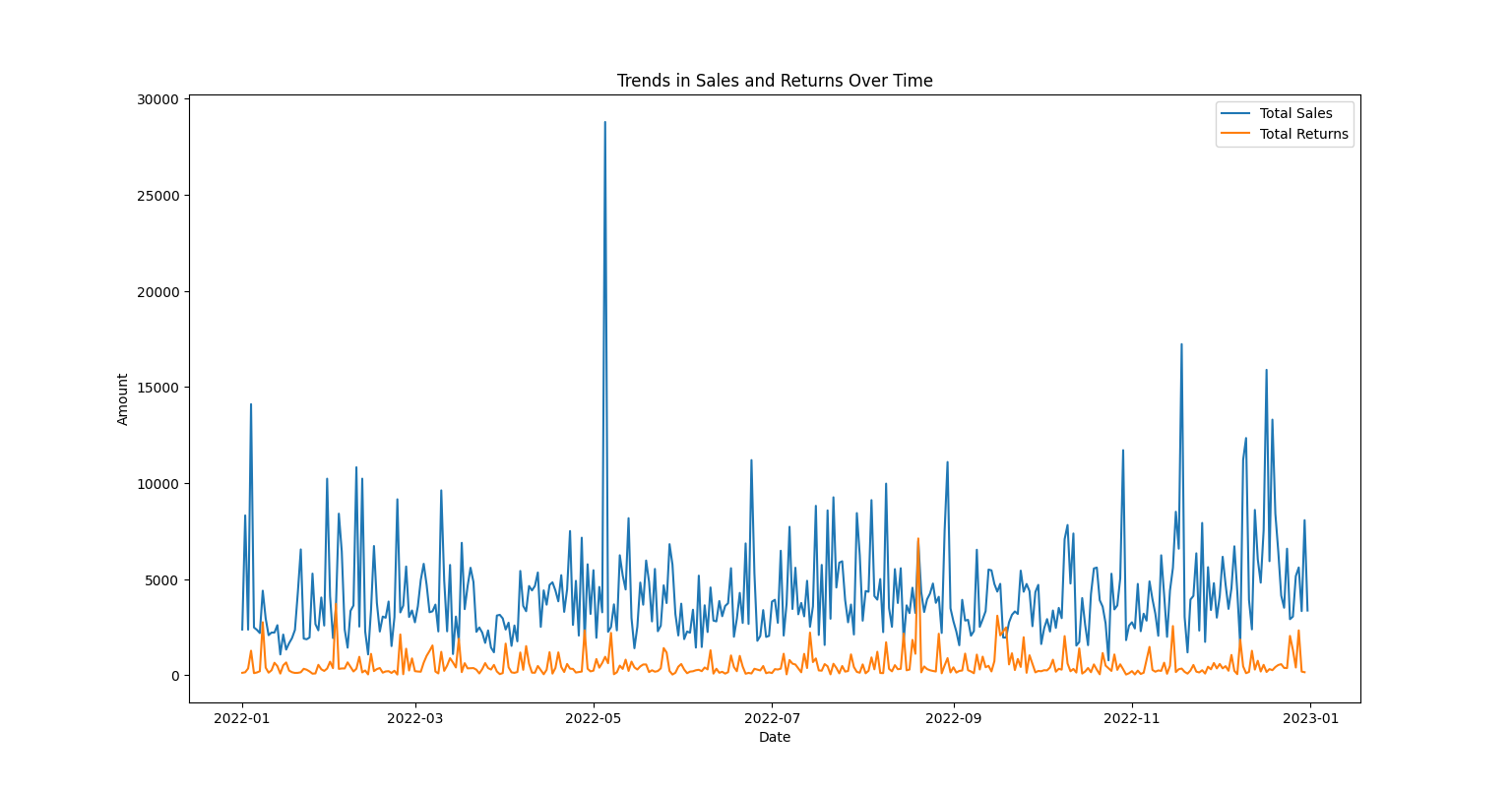


Fig-3 Trends in Sales and Returns Over Time

This line chart depicts trends in sales (blue line) and returns (red line) over time. The x-axis represents the date, and the y-axis represents the total amount.

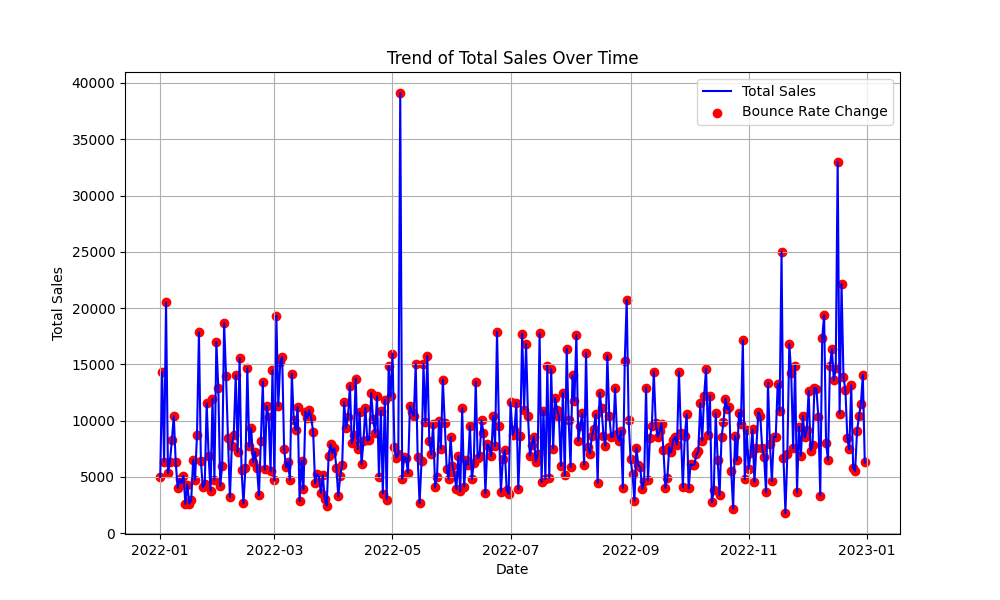


Fig-4 Trends in Sales and Bounce Rate change Over Time

The graph shows total sales increasing steadily over a year and tremendously increase in months May and November.

**Inventory Performance:**

The inventory turnover ratio was calculated to assess inventory efficiency. A higher ratio indicates better inventory management by selling through stock faster. The return rate was also calculated to understand the percentage of medication returned.

In our analysis, the inventory turnover ratio is calculated to be 14192.0, indicating that, on average, the inventory is turned over approximately **14192 times** within the analyzed period.

Our analysis reveals a return rate of 0.1314, indicating that approximately **13.14%** of the products sold are returned by customers.

**Product Analysis:**

The top 10 drugs and top 6 subcategories contributing to the highest sales were identified and displayed using bar charts. This helps prioritize focus on high-performing products.

Return rates were analyzed for individual drugs and subcategories. This can reveal insights into product quality or areas for improvement.

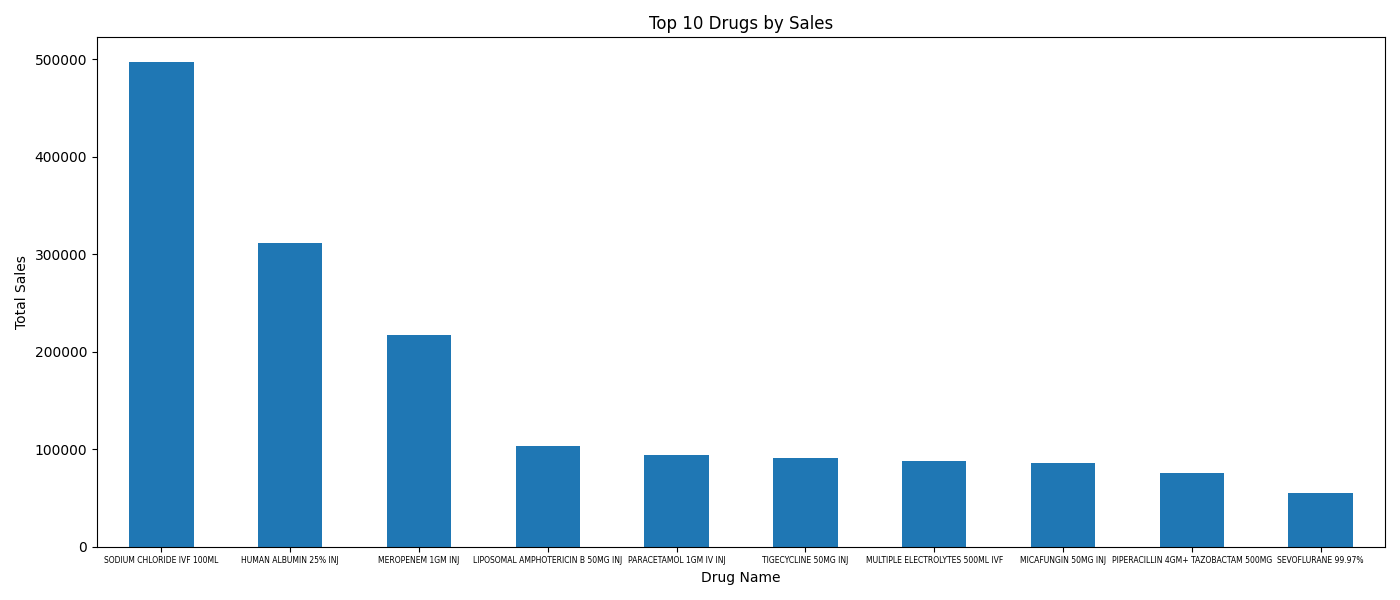


Fig-5 Top 10 Drugs by Sales

This bar chart highlights the ten medications contributing the most to overall sales in data set. The drug names are listed on the x-axis, and their corresponding total sales amount is represented on the y-axis.

From our analysis the most popular Drug by Sales is Sodium Chloride IVF 100ML with over 400000 sales and next to it is Human Albumin 25% INJ with 300000 sales. The third place is occupied by Meropenem 1GM INJ with 200000 over sales

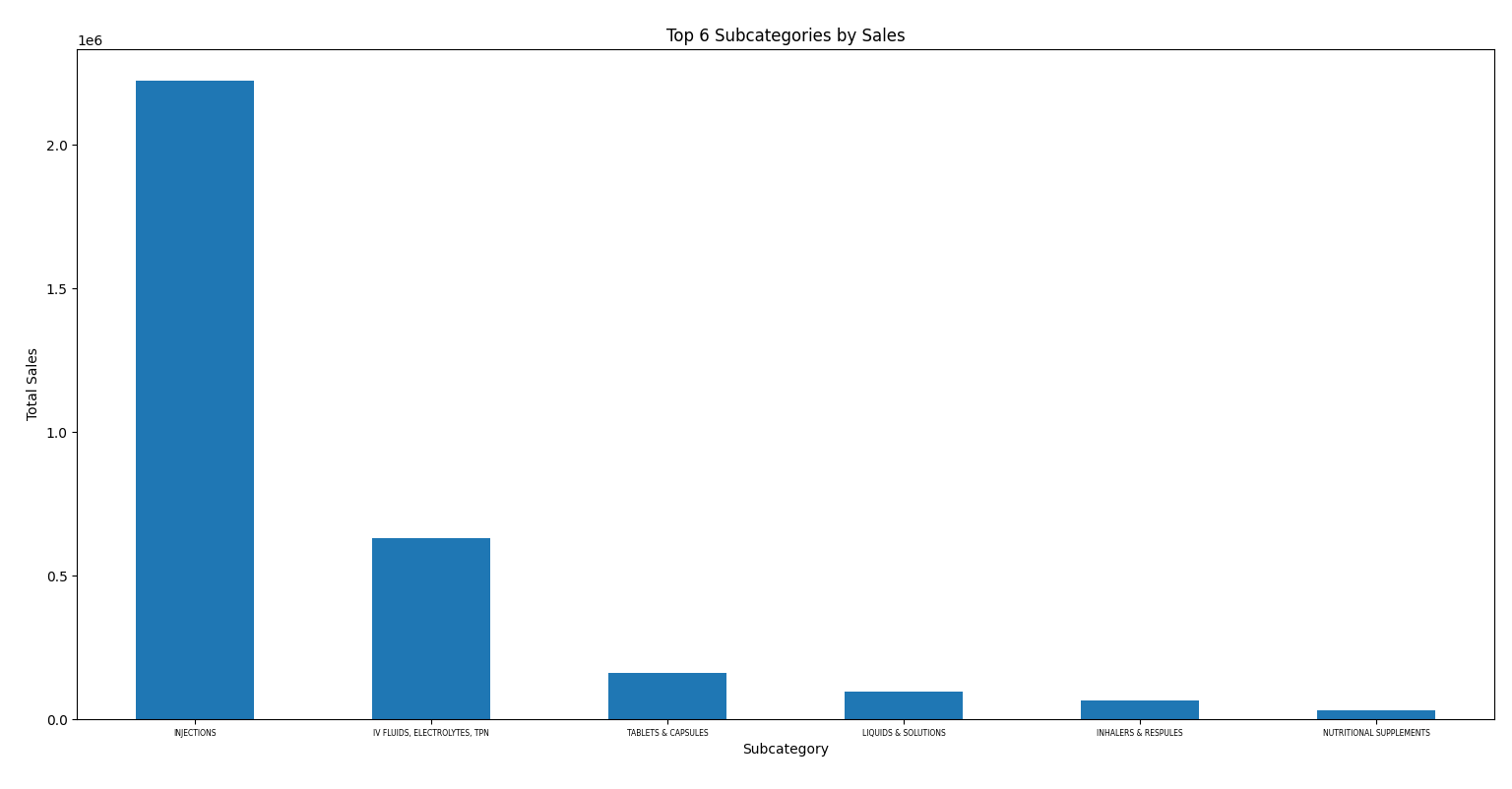


Fig-6 Top 6 Subcategories by Sales

This bar chart highlights the six subcategories of medications contributing the most to overall sales in data set. The subcategories are listed on the x-axis, and their corresponding total sales amount is represented on the y-axis.

**Sales Contribution by Department:**

The pie chart slices are labeled "Department1", "Department2", and "Department3".

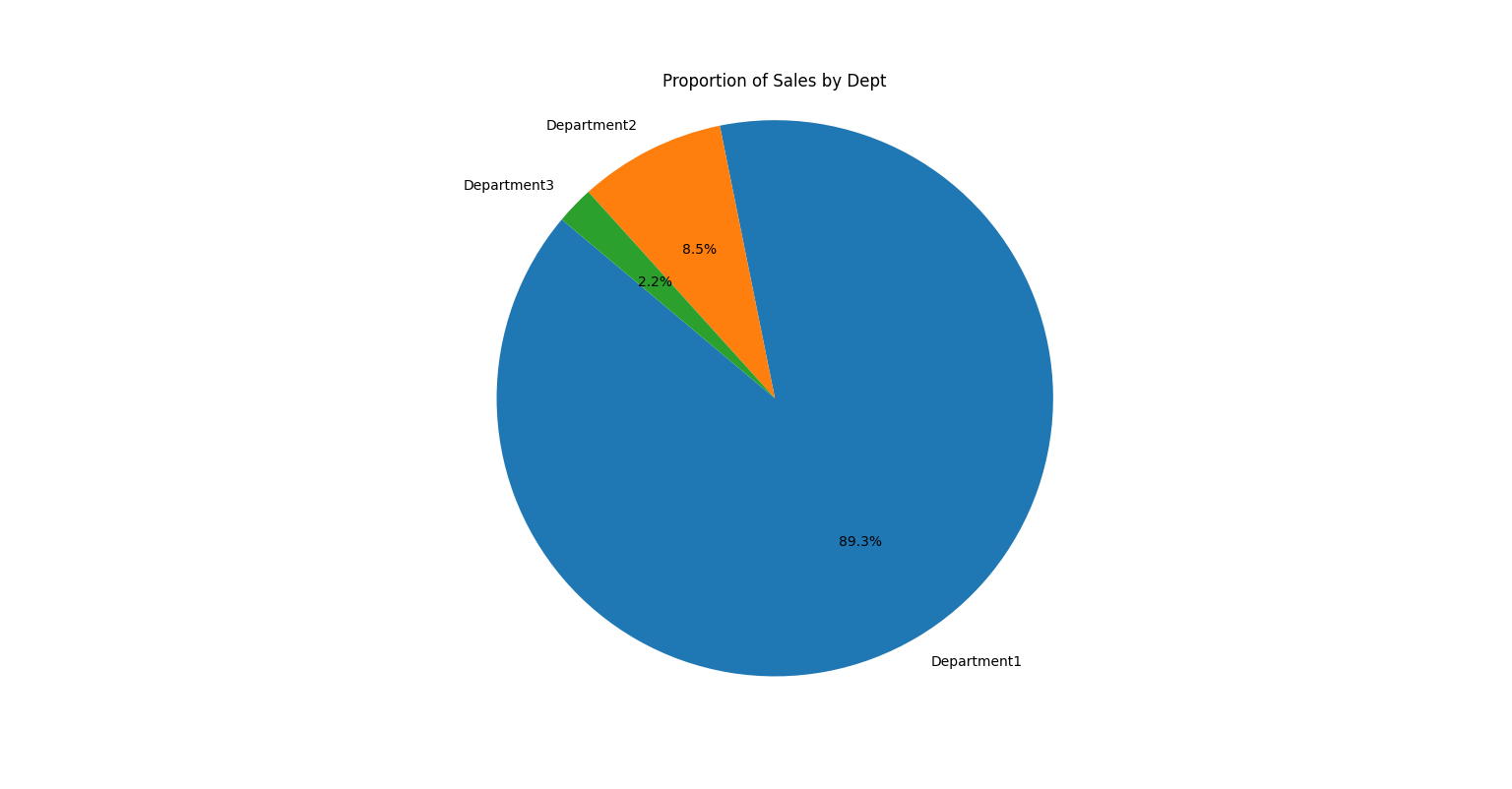


Fig-7 Sales Contribution by Department

Department3 makes up the vast majority of sales at 89.3%. Department 1 and Department2 contribute a much smaller portion of sales at 8.5% and 2.2% respectively.

## Recommendations:

Provide actionable recommendations based on the analysis

* Monitor inventory levels closely to optimize stock levels.
* Investigate the factors contributing to high return rates and take corrective actions.
* Implement pricing strategies to maximize profitability, considering the variance in final sales.

## Users:

Users of the visualization will be:

* Hospital Administrators : Hospital Administrators can use these visualizations to get insights into monitoring the bounce rate and inventory costs to optimize hospital operations and improve patient satisfaction.
* Pharmacists : Pharmacists can utilize the dashboard to manage medication inventory efficiently and ensure that essential drugs are always available for patient care.
* Supply Chain Managers : to optimize inventory levels, minimize stockouts, and reduce costs associated with excess inventory.

## Questions

Questions which will be answered by this visualization:

**Hospital Administrators:**

* What are the trends in bounce rates over time across different departments or specialties
* How does inventory cost vary across different departments

**Pharmacists:**

* What are the trends in bounce rates over time across different patient specializations
* How do final sales and costs trends inform inventory strategies

## Describe Visualization and how it answers the questions

**Hospital Administrators:**

* What are the trends in bounce rates over time across different departments or specialties
  + By visualizing bounce rates over time, Hospital Administrators can discern if certain departments consistently experience higher bounce rates compared to others.
  + This Visualization could provide insights into how bounce rates fluctuate over days, weeks, or months, enabling administrators to identify patterns and potential areas for improvement in patient flow and service delivery.
  + Through interactive charts and graphs, administrators can observe how bounce rates fluctuate over time, revealing any recurring patterns or seasonal variations.
* How does inventory cost vary across different departments
  + This pie chart gives the information that can help administrators comparing inventory costs across departments or specialties, allowing them to allocate resources effectively and identify opportunities for cost savings through better inventory management practices.

**Pharmacists:**

* What are the trends in bounce rates over time across different patient specializations
  + This Visualization can depict trends in bounce rates over time and analyze them by patient specializations or departments.
  + By correlating bounce rates with inventory levels and sales volumes, pharmacists can identify potential factors influencing bounce rates and develop targeted strategies to minimize them while optimizing inventory levels and sales performance.
* How do final sales and costs trends inform inventory strategies
  + A line graph could be used to illustrate the trends in final sales and final costs over time. Each line represents the respective metric, with time intervals (e.g., months) plotted on the x-axis and monetary values on the y-axis.
  + This visualization allows pharmacists to track how sales revenue and inventory costs fluctuate over different periods, identifying any seasonal patterns or irregularities.

## Conclusion

In conclusion, this EDA highlights key insights and recommendations derived from the analysis of pharmaceutical sales data. By leveraging these insights, businesses can make informed decisions to improve operational efficiency, enhance product performance, and maximize sales revenue. Analyzing the extensive medical inventory data, comprising over 14,000 records, even after aggregation, poses significant challenges in extracting comprehensive insights. Visualizing this data across various parameters can swiftly provide valuable insights unveiling additional insights into medication usage, inventory management, patient care practices and facilitate informed actions to enhance inventory management and patient satisfaction.