**Exploratory Data Analysis on E-commerce sales data**

**Data Source:** [**Kaggle**](https://www.kaggle.com/)

**Data:** [**E-commerce Sales Data**](https://drive.google.com/file/d/1W1V2YlF9eJvwHgCx0fJmaCt-ni-4kiij/view?usp=sharing)

Analysis Title: "E-commerce Sales and Profitability Analysis with EDA"

Objective: The objective of this analysis is to gain insights into e-commerce sales patterns and profitability drivers using the provided dataset.

Methodology: Data Cleaning and Preparation:

Data Cleaning and Preparation, clean the dataset to handle missing values, inconsistencies, and outliers. Convert date columns to datetime format for easier analysis. Explore the distribution of numerical variables and check for any anomalies.

**Exploratory Data Analysis (EDA):**

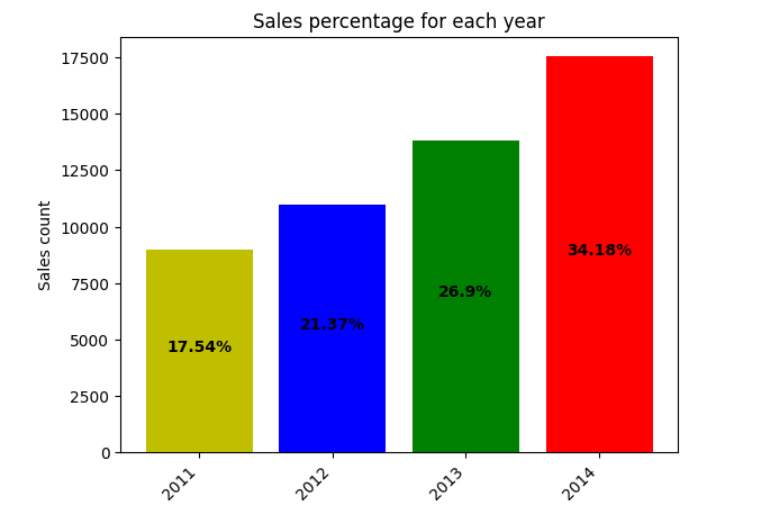
Below is the key Analysis in our data

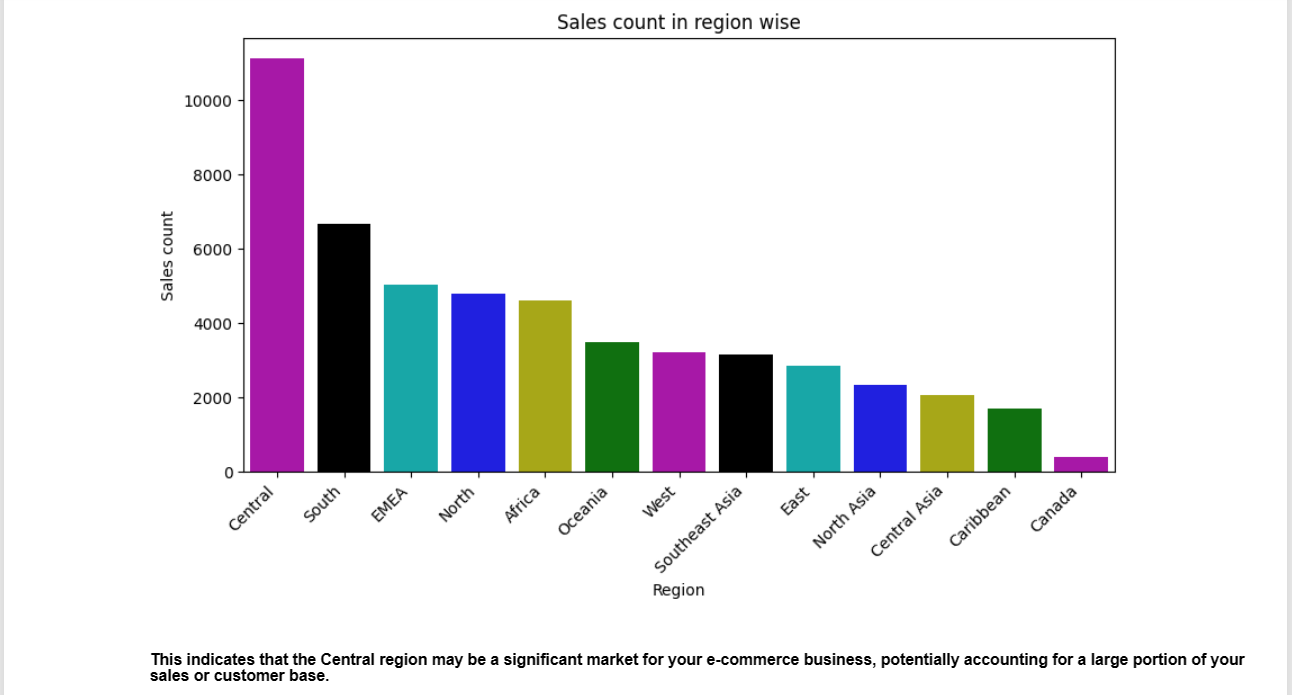
1. Explore the distribution of sales, profit, and other relevant numerical variables.
2. Analyze sales trends over time (e.g., monthly, quarterly).
3. Investigate the relationship between sales, profit, and other variables such as discount, quantity, and shipping cost.
4. Identify top-selling products, categories, and regions.
5. Assess the impact of discounts on sales and profitability.
6. Analyze shipping costs and their influence on overall profitability.
7. Examine the distribution of orders across different market segments, regions, and order priorities.

The task of exploring the distribution of sales, profit, and other relevant numerical variables involves analyzing the spread and characteristics of these variables within a dataset. Here's a summary of what this task typically entails:

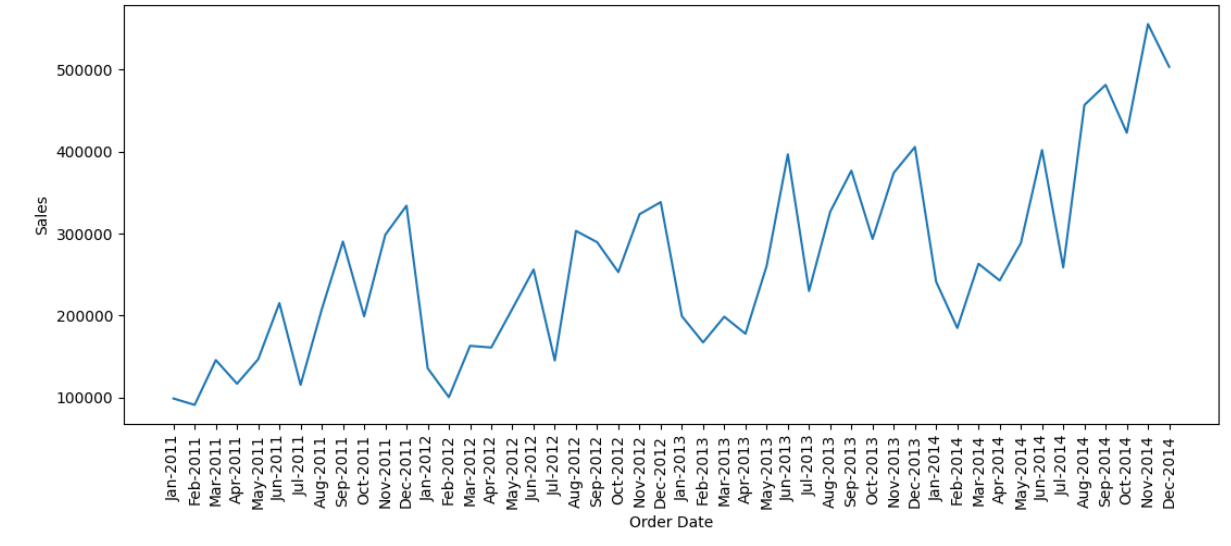
* Sales Distribution: Examining the distribution of sales values allows for understanding the range of sales amounts, the frequency of sales at different levels, and any patterns or anomalies in sales data. This analysis provides insights into the overall sales performance and variability within the dataset.
* Profit Distribution: Similarly, exploring the distribution of profit values enables the assessment of profitability levels, the presence of outliers or anomalies in profit data, and any trends or patterns in profit margins across different observations or categories.
* Other Numerical Variables: In addition to sales and profit, exploring the distribution of other relevant numerical variables, such as quantity sold, discount rates, shipping costs, or order quantities, provides a comprehensive understanding of various aspects of the dataset. Analyzing the distribution of these variables helps identify trends, outliers, and potential relationships with sales and profit metrics.

Overall, exploring the distribution of sales, profit, and other numerical variables involves visualizing the data through histograms, density plots, or box plots, assessing measures of central tendency and variability, and identifying any patterns, anomalies, or insights that can inform decision-making in areas such as pricing strategies, inventory management, or performance evaluation.





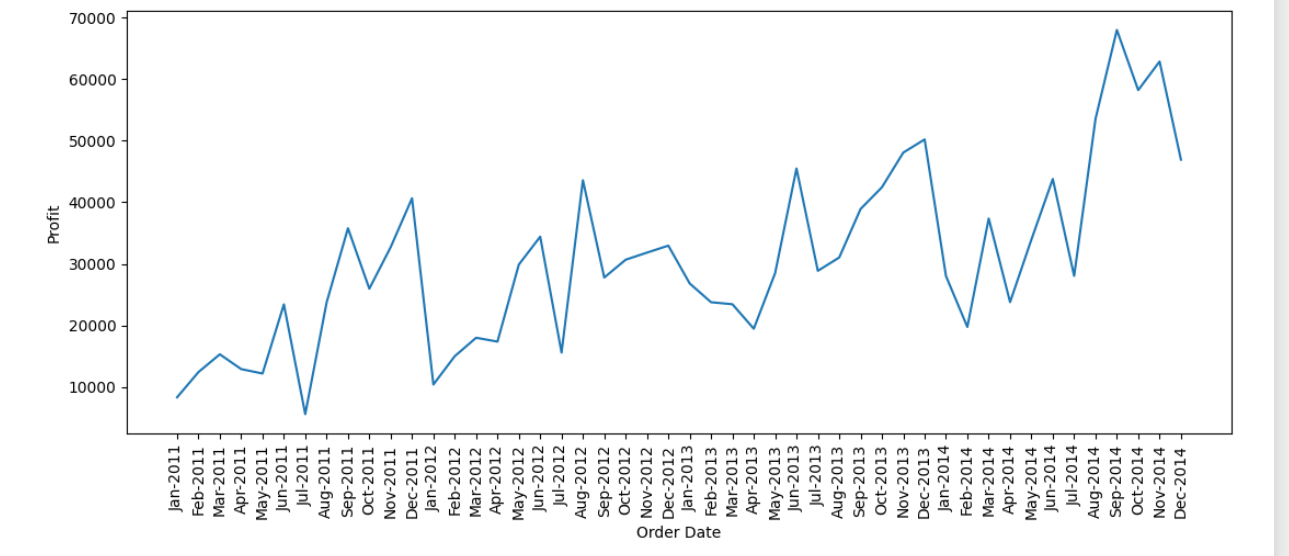
**Sales Trend**



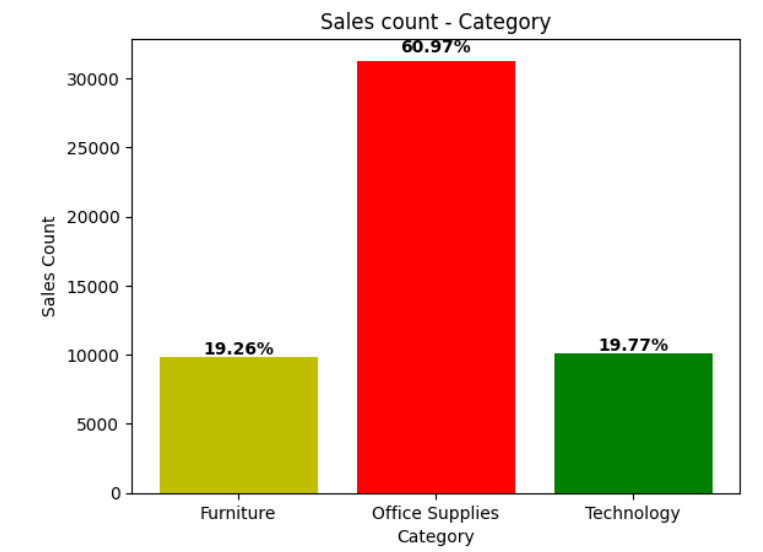
**Visualizing the percentage of sales for each year allows for a clear understanding of how sales distribution has evolved over time. Above findings indicate a noticeable increase in sales proportion from 2011 to 2014**

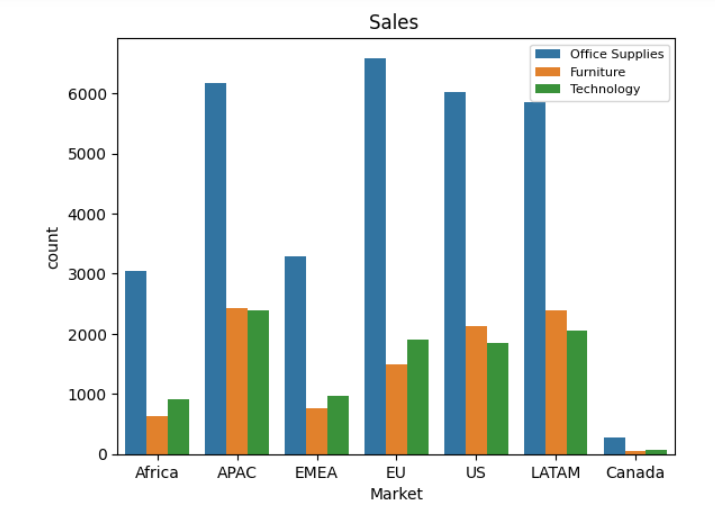
1. Yearly Sales Growth: The increase in the percentage of sales from 2011 to 2014 suggests a positive trend in your e-commerce business's performance over the years. This growth may be attributed to various factors such as expanding customer base, product offerings, or improved marketing strategies.
2. Acceleration of Growth: The acceleration in sales growth between 2013 and 2014, with a substantial jump from 26.9% to 34.18%, indicates a period of rapid expansion or heightened sales activity during that time.

**Profit Trend**



#### From the above plot we can see the trend of sales from jan-2011 till Dec-2014. From the pattern we can see increasing trend in sales, but in depth if you notice Feb and July of every year the sales have dropped which indicates poor sales strategy and need attention on those months





**Understanding the distribution of sales across different markets and product categories provides valuable insights into this e-commerce business's performance and customer preferences.**

**APAC Market Dominance**

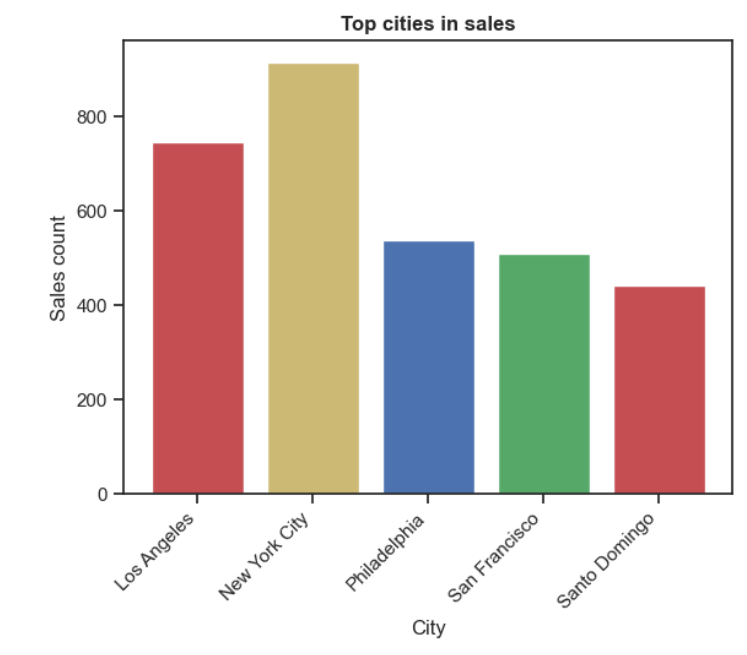
From the above observation that the APAC (Asia-Pacific) market has the highest count in overall suggests that this region is a significant contributor to our e-commerce sales. This insight highlights the importance of catering to the needs and preferences of customers in the APAC region and potentially allocating resources or marketing efforts accordingly to capitalize on this market opportunity.

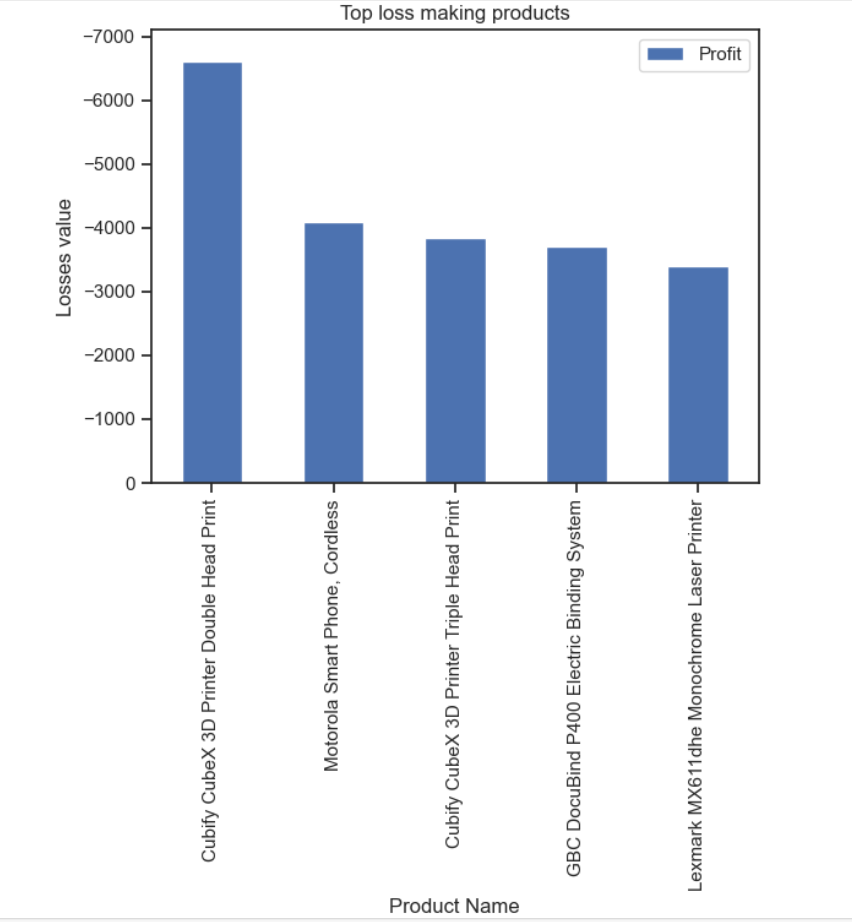
**Office Supplies Sales Leadership:**

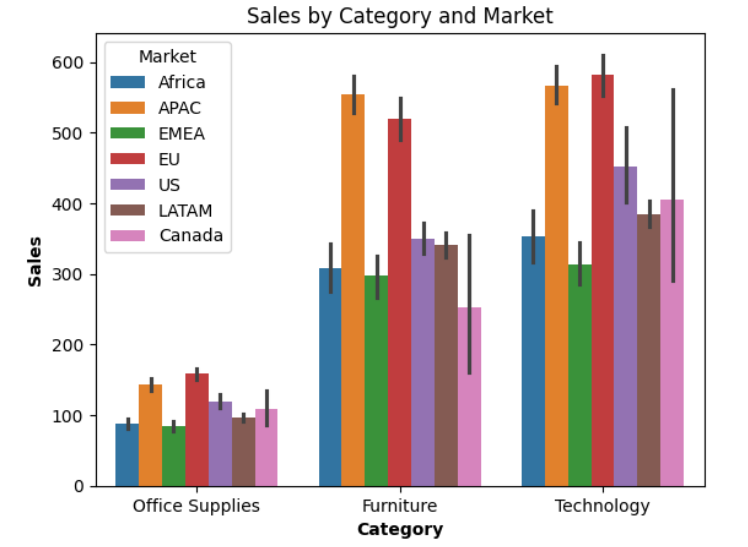
Across all markets, the "Office Supplies" category emerges as the top-selling product category. This finding indicates strong demand for office-related products and suggests that customers prioritize purchases related to workplace needs or office supplies. Understanding the popularity of office supplies can inform inventory management, marketing strategies, and product assortment decisions to meet customer demand effectively.

Identifying the Canada market as having the lowest count across all product categories provides valuable information for understanding market dynamics and potentially uncovering opportunities for growth or improvement. Here's how this observation can inform our analysis and strategic decision-making

1. Market Assessment: Analyzing the reasons behind this lower count, such as market saturation, competition, or economic factors, can help identify areas for improvement and potential strategies to revitalize sales in the Canadian market
2. Market Expansion Opportunities: Despite its current lower count, the Canadian market may present untapped opportunities for expansion and growth. Exploring targeted marketing efforts, promotional campaigns, or product offerings tailored to Canadian customers' preferences and needs can help stimulate demand and increase market share in this region
3. Customer Insights: Understanding customer behavior and preferences in the Canadian market can provide insights into unique challenges or opportunities specific to this region
4. Competitive Analysis: Assessing competitors' strategies and performance in the Canadian market can provide benchmarks for comparison and inform competitive positioning strategies. Identifying gaps or weaknesses in competitors' offerings can guide differentiation strategies
5. Resource Allocation: Based on the insights from our analysis, we can allocate resources, investments, and marketing budgets strategically to prioritize growth opportunities in the Canadian market.







**Analysis Out comes**

1. **Sales Trends:**
   * Overall, sales show an increasing trend over the years, with noticeable growth from 2011 to 2014.
   * 246.49 is the average sale price of a product from this data set
   * Sales distribution varies across different product categories and markets, with certain categories and regions showing higher sales volumes.
   * APAC and EU markets exhibit above-average sales performance, suggesting significant opportunities for business growth in these regions.
   * Furniture and Technology categories demonstrate similar sales levels across APAC and EU markets, indicating consistent consumer interest in these product categories.
2. **Profit Analysis:**
   * Profit margins vary across different product categories, with some categories showing higher profitability than others.
   * Outlier analysis reveals extreme profit ranges, indicating potential anomalies or exceptional profit events within the dataset.
   * Understanding profit distribution and trends is crucial for assessing the financial performance of the business and identifying areas for improvement or optimization.
   * Further analysis of profit in relation to sales, discount rates, and shipping costs can provide insights into factors influencing profitability and guide strategic decision-making.
3. **Customer Insights:**
   * Identification of top customers based on purchase frequency, such as Muhammed, highlights the importance of customer loyalty and engagement in driving sales and profitability.
   * Analyzing customer behavior and preferences can inform personalized marketing strategies, loyalty programs, and customer relationship management initiatives aimed at fostering long-term relationships and maximizing customer lifetime value.
4. **Market Segmentation:**
   * Market segmentation based on sales performance, product categories, and geographic regions allows for targeted marketing efforts and tailored strategies to address diverse customer needs and preferences.
   * Understanding differences in sales patterns and trends across markets and categories facilitates strategic resource allocation and market prioritization to optimize business growth and expansion opportunities.

**Conclusion:**

In conclusion, the analysis of sales and profit within the e-commerce dataset has provided valuable insights into the financial performance, market dynamics, and customer behavior. The examination of sales trends, profit margins, customer engagement, and market segmentation has illuminated key areas of strength and opportunities for improvement. The findings highlight the importance of strategic decision-making informed by data-driven insights to optimize sales growth, maximize profitability, and enhance customer satisfaction. Moving forward, leveraging these insights to refine marketing strategies, tailor product offerings, and prioritize market expansion efforts will be essential for achieving sustainable business success in the competitive e-commerce landscape. This project underscores the significance of data analytics in driving informed decision-making and guiding strategic initiatives to drive business growth and maximize value creation for stakeholders.