***Govindram Seksaria Institute of Management and Research INDORE, (M.P)***

******

Submitted in partial fulfilment of the requirements for the awardof the degree.

**Master of Business Administration (FT)**

**Year 2022-2024**

**“Major research report at A Strategic Analysis of Sales Data and Digital Consumer Behavior”**

**Submitted by: - Submitted to:-**

Prince Gupta C.A. Prem Dua

MBA (FT) III Semester

Enrolment no – DX2104234

**A close-up of a logo

Description automatically generated**

**STUDENT DECLARATION**

**Prince Gupta**, hereby declare that the presented Major research reportat “**Analysis of shopper sales data”** is uniquely prepared by me after the completion of, Madhya Pradesh.

I also confirm that the report is only prepared for my academic requirement not for any other purpose. It might not to be used with the interest of opposite party of the corporation.

**Sincerely yours**

Prince Gupta

MBA (FT) III SEM

Enrolment no – DX2104234

**A close-up of a logo

Description automatically generated**

**CERTIFICATE**

This is to certify that Prince Gupta the student of Master of Business Administration program of **Govindram Seksaria Institute of Management and research, Indore** has completed major research Project Report titled “**Analysis of shopper sales data**” under my guidance and supervision. As per my knowledge the work carried out by her/him is original and genuine. I wish further success to her/him.

Place- indore Signature of Guide

Date-

**TABLE OF CONTENT**

|  |  |  |
| --- | --- | --- |
| Sr.no. | content | Pg no. |
| 1 | Student declaration | 2 |
| 2 | Certification | 3 |
| 3 | Introduction | 5 |
| 4 | Objective | 5 |
| 5 | Problem statement | 5 |
| 6 | Data collection | 6 |
| 7 | Analysis of data | 7 |
| 8 | Data complexity | 9 |
| 9 | Profit and loss confusion | 9 |
| 10 | State wise Sales difficulty | 10 |
| 11 | Top buyer identification | 10 |
| 12 | Exploratory data analysis | 11 |
| 13 | Track and analysis by chart & Graphs | 14 |
| 14 | Analysis by Quarter | 21 |
| 15 | Regression analysis | 25 |
| 16 | Conclusion | 32 |

**Introduction-** Online sales analysis and track is the process of collecting, measuring, and interpreting data about online sales. This data can be used to understand how customers are interacting with your website, what products are selling well, and where you can improve your sales performance. Online sales analysis can help you increase sales, improve customer service, and make better decisions about your business. Through this, you can better understand how your online sales are performing and identify areas where you can improve. This information can then be used to make informed decisions about your marketing, product development, and customer service.

**Objective-** The main goal of tracking and analyzing online sales data is to understand and improve the performance of our digital store. By setting clear objectives and using key metrics, we aim to monitor sales in real-time, identify what products are resonating with customers, and The objective is to not only current performance but also predict future trends through forecasting models. This will help us make informed decisions on marketing strategies and product offerings. Through user-friendly dashboards, we want to present these insights in a way that is easily understandable, quick and effective decision-making. Ultimately, our focus is on continuous improvement, adapting to changing market conditions and customer behaviors to ensure sustained success in the online sales landscape.

**Problem statement-**

The organization is facing several challenges related to managing its product sales data:

1. Data complexity-Understanding the large amount of sales data is proving to be quite complicated.
2. Top buyer identification- It's hard to figure out who the top buyers are and understand their buying habits.
3. Profit and loss confusion-Making sense of whether the business is making a profit or facing losses is proving challenging.
4. State-wise sales difficulty-Determining which states are contributing the most to product sales is not straightforward.
5. Top product recognition-Identifying which products are selling the most is proving to be a difficult task.

**Data collection-** I takes this data from below link-

https://www.youtube.com/redirect?event=video\_description&redir\_token=QUFFLUhqbUZHclQwUy1sQ1ZaMGpDTVdWWjZFRGtpcktBUXxBQ3Jtc0trbFg1OFZSblVVSnphSEoteUgtWHBPUUYwdWpkc0dIRTJLOTQ5LUwxSEs0ZjJhUVB4akRmeTZqVWtLcGlabVBEdWFkVlFHUDhIY2tvZEpvSmpPNXJLSjBNOWVoeEM3eDBhOUdLem96TkJPMWd0T09UVQ&q=https%3A%2F%2Fbit.ly%2F3ZiHghb&v=6cV3OwFrOkk

**Analysis of data-**

1.Numerical-

|  |  |
| --- | --- |
|  | Amount |
| Mean | 291.8473 |
| Standard Error | 11.92684 |
| Median | 122 |
| Mode | 44 |
| Standard Deviation | 461.9246 |
| Sample Variance | 213374.4 |
| Kurtosis | 27.33664 |
| Skewness | 4.142391 |
| Range | 5725 |
| Minimum | 4 |
| Maximum | 5729 |
| Sum | 437771 |
| Count | 1500 |
| Largest(1) | 5729 |
| Smallest(1) | 4 |
| Confidence Level(95.0%) | 23.39507 |

1. Mean (Average):The typical amount is around 291.85.

2. Standard Error: The average amount can vary by about 11.93.

3. Median: Half of the amounts are below 122.

4. Mode: The most common amount is 44.

5. Standard Deviation: The amounts typically differ from the average by about 461.92.

6. Sample Variance: The amounts vary a lot, with a variance of 213374.4.

7. Kurtosis: The data has a high concentration of extreme values.

8. Skewness: The data is stretched more to the right.

9. Range: The difference between the highest and lowest amounts is 5725.

10. Minimum and Maximum: The smallest amount is 4, and the largest is 5729.

11. Sum :If you add up all the amounts, you get 437771.

12. Count: There are 1500 amounts in total.

13. Largest(1) and Smallest(1):The single largest amount is 5729, and the smallest is 4.

14. Confidence Level (95.0%): The average amount is expected to be within 23.40 of the calculated average, with 95% confidence.

2.Graphical-

A screenshot of a computer

Description automatically generated

**1.Data Complexity :**

* Challenge: Online sales generate vast amounts of data, including customer transactions, product details, and various metrics. Managing the volume, variety, and velocity of this data presents complexity.
* Approach: Utilize data modeling techniques to integrate and structure diverse data types. Leverage tools like filters and sorting options in your analysis platform to manage data complexity effectively.

**2.Top Buyer Identification:**

* Challenge: Determining who the top buyers are from a large customer base involves analyzing purchasing patterns, frequency, and overall contribution.
* Approach: Use data analytics tools to create customer profiles, focusing on metrics like total spending, purchase frequency, and loyalty. Implement segmentation and ranking to identify and visualize the top buyers.

**3.Profit and Loss Confusion:**

* Challenge: Analyzing profit and loss in online sales requires dissecting various cost components, revenue streams, and understanding overall financial performance.
* Approach: Develop financial dashboards using your analysis platform to visualize revenue, costs, and profit margins. Implement drill-down features for detailed insights, helping to clear up any confusion regarding profit and loss.

**4. State-Wise Sales Difficulty:**

* Challenge: Managing and analyzing sales data on a state-by-state basis may be complex due to regional variations, different regulations, and market dynamics.
* Approach**:** Leverage geographical mapping tools to create visualizations that highlight state-wise sales performance. Implement filters to focus on specific states and conduct trend analysis to understand difficulties and opportunities in different regions.

**5.Top Product Recognition:**

* Challenge:Identifying the top-performing products from a diverse catalog involves analyzing sales volume, revenue contribution, and customer preferences.
* Approach:Use data visualizations to display product-wise metrics such as sales, revenue, and customer ratings. Incorporate sorting and ranking features to easily recognize and showcase top-performing products.

**Exploratory data analysis:**

**Categories:**

1.Furniture: Furniture, a multifaceted domain in the realm of design and utility, comprises an extensive array of movable objects intended to facilitate various human activities across different environments. From the elegance of a well-crafted dining table to the comfort of a plush sofa, furniture not only fulfills practical needs but also contributes significantly to the aesthetics of living and working spaces.

The choice of materials is a pivotal aspect of furniture design, with options ranging from the timeless warmth of wood to the sleek modernity of metal and the versatility of plastic and fabric. Each material not only imparts distinct visual characteristics but also influences the durability and tactile experience of the furniture. Consequently, the careful selection of materials becomes an artful process, considering both form and function.

In the realm of interior design, furniture assumes a central role. Its arrangement and placement can transform a mere physical space into a harmonious and inviting environment. Whether creating cozy corners for relaxation or optimizing spaces for efficient work, the strategic use of furniture contributes to the overall atmosphere and functionality of a room.

2.Clothing: Turning our attention to clothing, this facet of daily life extends beyond mere practicality. Clothing is a canvas for personal expression, a medium through which individuals communicate their identity, cultural affiliations, and stylistic preferences. Beyond the functional aspects of protection and modesty, clothing becomes a form of wearable art that reflects the zeitgeist of a society.

The textile industry, a crucial component of the clothing ecosystem, produces a vast array of fabrics. Natural fibers like cotton and wool coexist with synthetic materials such as polyester, offering designers a diverse palette to create textures, patterns, and garments that cater to a myriad of tastes and preferences. The intricate processes of weaving, knitting, and dyeing contribute to the rich tapestry of textiles available for fashion and apparel.

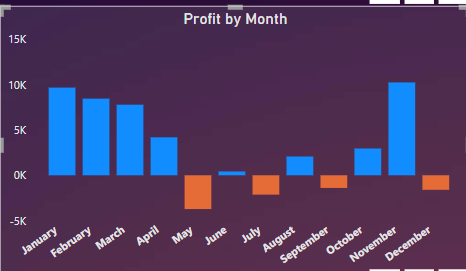
In the broader context of the fashion industry, clothing takes on a more dynamic and trend-driven dimension. Fashion becomes a reflection of societal changes, cultural influences, and artistic movements. Designers, influencers, and consumers collectively shape the ever-evolving landscape of styles, creating a dialogue between tradition and innovation.

In essence, furniture, electronics, and clothing form integral components of the human experience. Beyond their utilitarian functions, they encapsulate the essence of design, innovation, and self-expression, weaving a tapestry that reflects the diverse and ever-evolving nature of human culture and civilization.

**3.**Electronic: Moving on to electronics, this dynamic field of study revolves around the manipulation and control of electrical energy to enable a plethora of devices that have become integral to modern life. Smartphones, computers, televisions, and an array of gadgets are manifestations of electronic technologies that process information, transmit signals, and redefine communication.

Electronics, in its vastness, embraces both digital and analog realms. The digital domain, with its discrete binary code, governs the logic and precision of modern computing. Meanwhile, analog technology, dealing with continuous signals, finds its application in areas such as audio systems, offering a rich and nuanced experience. The relentless pace of innovation in electronics not only brings forth smaller and more powerful devices but also shapes the way societies connect, communicate, and navigate the digital landscape.

**Track and analysis by chart and graph:**

**1.Analysis profit and loss by month-**

**Financial Performance Summary:**

Profitable Months:

* Jan, Feb, Mar, Jun, Aug, Oct, Nov

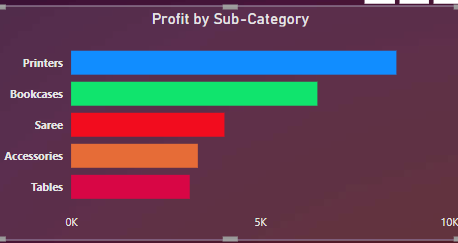
Loss Months:

* May, Jul, Sep, Dec

Highlight

* Highest Profit: November
* Highest Loss: May

**2.Profit by sub- category-**

****

**Profit Analysis by Sub-category:**

**Printers:**

* Highest selling and most profitable product.
* Key contributor to overall revenue.
* Consider exploring strategies to further capitalize on the success of this sub-category**.**

**Bookcases:**

* Second highest in sales and profitability.
* Demonstrates consistent market demand.
* Opportunities for optimizing marketing and sales efforts.

**Accessories:**

* Contributes positively to overall profit.
* Potential for further market penetration and product expansion.
* Monitor trends and customer preferences for potential growth areas**.**

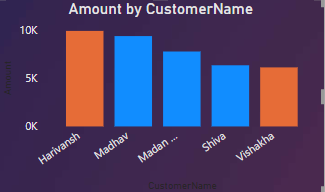
**Tables:**

* Notable impact on profit margins.
* Evaluate product positioning and marketing strategies for continued success.
* Identify opportunities for product enhancement or bundling.

**Saree:**

* Positioned within the product portfolio.
* Assess market demand and consider targeted marketing approaches.
* Explore potential synergies with other product categories**.**

**3.Analysis amount by customer name-**

****

**Harivansh:**

* Total Expenditure: 9.9k
* Consistent and substantial engagement, indicating high commitment.

**Madhav:**

* Total Expenditure: 9.4k
* Strong loyalty and recurrent interest in our product offerings.

**Madan:**

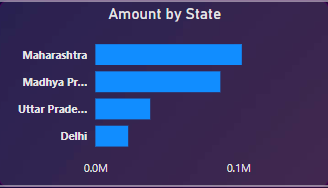
* Total Expenditure: 7.8k
* Demonstrates notable financial commitment and consistent patronage.

**Shiva:**

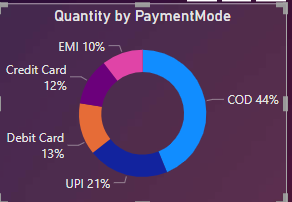
* Total Expenditure: 6.3k
* Significant contribution to revenue, reflecting continued interest.

**Vishaka:**

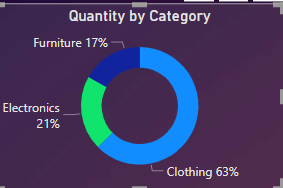
* Total Expenditure: 6.1k
* Sustained investment underlining commitment to our brand.

**4.Analysis amount by state-**

Collect sales data for Maharashtra, MP, UP, and Delhi, focusing on product amounts and revenue. Organize the data and calculate total sales for each city. Visualize the results using clear charts for easy comparison. Analyze seasonality and growth rates to identify trends. Compare demographic factors influencing sales differences. Gather customer feedback to refine marketing strategies. Tailor city-specific campaigns based on insights. Utilize statistical analyses to validate variations in sales amounts. Understand target audiences in each city for more effective marketing. Consider collaborations with local businesses for enhanced market presence. Implement predictive analytics for future sales forecasting. Develop contingency plans for unforeseen events affecting sales. Regularly monitor and update strategies to adapt to evolving market conditions.

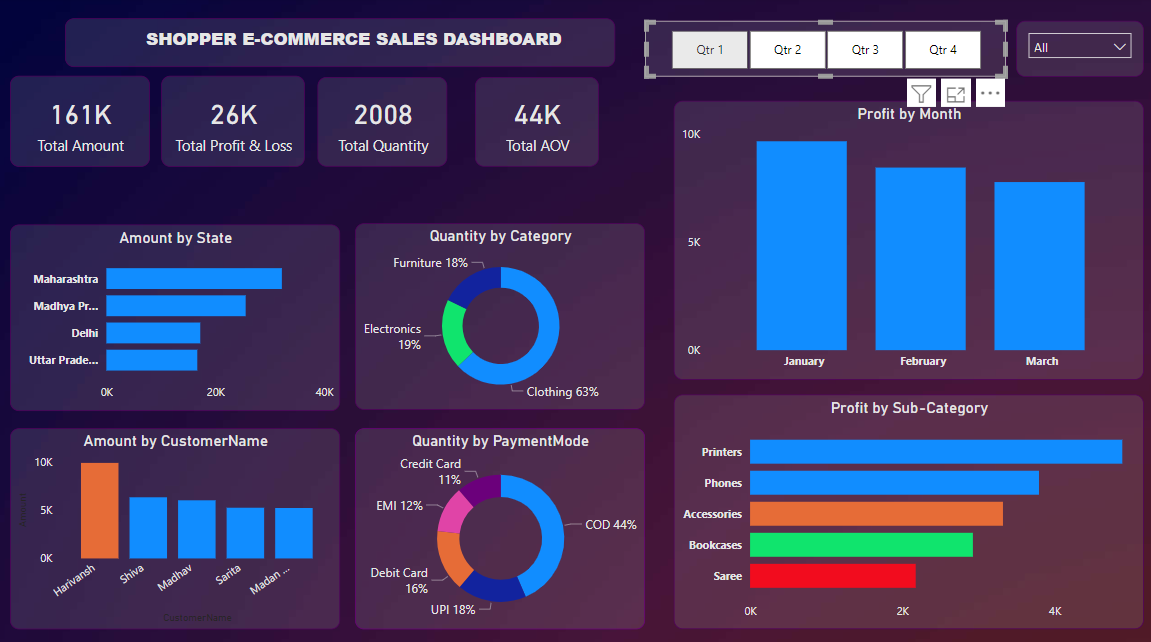
**5.Analysis Quantity by paymentmode-**

Analyze product quantity across payment modes: UPI (21%), credit card (12%), debit card (13%), EMI (10%), COD (44%). Explore correlations with product categories and demographics. Evaluate financial impact, considering cash flow and customer behaviors. Adapt marketing strategies, offer incentives, and monitor changes in payment options. Collaborate with financial institutions for efficiency and security. Address customer feedback, ensuring a seamless experience. Stay informed about industry trends and regulations. Communicate transparently with customers, adapting to technological advancements. Regularly review and update payment processes for optimization.

**6.Analysis Quantity by Category-**

Analyze product quantity distribution by category: furniture (17%), electronics (21%), and clothing (63%). Explore correlations with customer demographics and preferences within each category. Evaluate the impact of each category on overall sales quantity, considering seasonality or trends. Adapt marketing strategies to target specific categories, aligning promotions with customer preferences. Monitor changes in category preferences over time and adjust inventory or marketing efforts accordingly. Collaborate with suppliers or manufacturers to optimize product availability and pricing within each category. Address customer feedback related to specific categories to enhance the overall shopping experience. Stay informed about industry trends and competitor strategies within each category. Communicate transparently with customers about product offerings and any updates in category selections. Regularly review and update category-specific processes for continued optimization.

**Analysis by Quarter-**

**Qtr 1-**

In the first quarter, achieved sales of 161k with a net profit and loss of 26k. Sold a total of 2008 units, resulting in an average order value (AOV) of 44k. Monitor quarterly trends, assessing fluctuations in sales, profit, and quantity. Calculate profit margins to evaluate operational efficiency. Analyze the performance of individual products and customer segments. Evaluate operating expenses and identify potential cost-saving measures. Compare Q1 performance with previous quarters or industry benchmarks for insights. Consider season-specific factors influencing sales and adjust strategies accordingly. Continuously track and adapt marketing approaches based on customer behavior. Regularly communicate insights with key stakeholders for informed decision-making.

**Qtr 2-**

In the second quarter, recorded sales of 87k with a net profit and loss of 882. A total of 1181 units were sold, resulting in an average order value (AOV) of 24k. Analyze quarterly trends for insights into fluctuations in sales, profit, and quantity. Calculate profit margins to assess operational efficiency. Examine individual product or category performance for optimization. Segment sales and AOV based on customer behaviors. Evaluate operating expenses for cost-saving opportunities. Compare Q2 performance with previous quarters or industry benchmarks. Consider any season-specific factors influencing sales and adjust strategies accordingly. Continuously track and adapt marketing approaches based on customer behavior. Communicate insights regularly with key stakeholders for informed decision-making.

**Qtr 3-**

****

In the third quarter, the business recorded a total revenue of $72,000, despite facing challenges as evidenced by a net loss of -$1,469. The total quantity of goods or services sold reached 1,017 units. Notably, the average order value (AOV) for the quarter was $19,000, reflecting the average amount spent per transaction.Analyzing the financial snapshot, it's apparent that the business faced some profitability issues during this period. The negative profit and loss figure indicates that the company's expenses exceeded its revenue, contributing to the overall loss. To gain a deeper understanding, it would be helpful to further investigate the cost structure and identify areas where efficiencies could be improved.Despite the financial setback, the total quantity sold is a positive indicator, showcasing consistent sales activity. The AOV of $19,000 suggests that while the number of transactions may be relatively low, they contribute significantly to the overall revenue. Understanding customer behavior and preferences can provide insights into maximizing the AOV and, consequently, overall revenue.To improve the financial performance in the next quarter, the business should focus on cost management strategies, explore avenues for increasing transaction volume, and consider implementing targeted marketing campaigns to attract high-value customers. Regular monitoring of key performance indicators will be essential to track progress and make informed decisions for sustainable growth.

**Qtr 4-**

****

In the fourth quarter, the business experienced significant growth, with total revenue reaching $118,000, marking a notable improvement from the previous quarter. The company achieved a commendable net profit of $12,000, signaling a positive turnaround from the losses incurred in Q3. Sales volume saw a substantial increase, with a total of 1,409 units sold, reflecting heightened market demand or operational efficiency.A key highlight is the surge in the Average Order Value (AOV) to $33,000, indicating that customers, on average, spent more per transaction. This points to effective strategies that potentially upsell or capture higher-value transactions. The positive financial indicators underscore successful business initiatives implemented during the quarter.To sustain this momentum, the business should further analyze the specific factors contributing to this success, identify customer behavior patterns, and assess the impact of any implemented strategies. Continued vigilance over these key performance indicators will be crucial for adapting and optimizing strategies in the future. Regular analysis will help in maintaining growth and profitability.

**Regression analysis-**

**Electronics items-**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| SUMMARY OUTPUT | |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
| *Regression Statistics* | |  |  |  |  |  |  |  |
| Multiple R | 0.339388256 |  |  |  |  |  |  |  |
| R Square | 0.115184388 |  |  |  |  |  |  |  |
| Adjusted R Square | 0.103693276 |  |  |  |  |  |  |  |
| Standard Error | 428.3994314 |  |  |  |  |  |  |  |
| Observations | 79 |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
| ANOVA |  |  |  |  |  |  |  |  |
|  | *df* | *SS* | *MS* | *F* | *Significance F* |  |  |  |
| Regression | 1 | 1839625.155 | 1839625 | 10.02378 | 0.002215 |  |  |  |
| Residual | 77 | 14131507.6 | 183526.1 |  |  |  |  |  |
| Total | 78 | 15971132.76 |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
|  | *Coefficients* | *Standard Error* | *t Stat* | *P-value* | *Lower 95%* | *Upper 95%* | *Lower 95.0%* | *Upper 95.0%* |
| Intercept | 501.5150178 | 48.23252147 | 10.39786 | 2.53E-16 | 405.4718 | 597.5582 | 405.4718 | 597.5582 |
| Profit | 0.70137641 | 0.22153144 | 3.166036 | 0.002215 | 0.260251 | 1.142502 | 0.260251 | 1.142502 |
|  |  |  |  |  |  |  |  |  |

1. Multiple R (Multiple Correlation Coefficient):

Value: 0.3394

Explanation: This is a measure of the strength and direction of the linear relationship between the dependent variable (response) and the independent variables (predictors) in your regression model. It ranges from -1 to 1, where 0 indicates no linear relationship, 1 indicates a perfect positive relationship, and -1 indicates a perfect negative relationship.

2. R Square (Coefficient of Determination):

Value: 0.1152

Explanation:R Square represents the proportion of the variance in the dependent variable that is predictable from the independent variables. It ranges from 0 to 1, with higher values indicating a better fit. In this case, approximately 11.52% of the variance in the dependent variable is explained by the independent variables.

3. Adjusted R Square:

Value: 0.1037

Explanation: This is a modified version of R Square that adjusts for the number of predictors in the model. It penalizes the inclusion of irrelevant variables that do not improve the model. It is particularly useful when comparing models with different numbers of predictors.

4. Standard Error:

Value: 428.3994

Explanation: The standard error of the regression (or residual standard error) is a measure of the dispersion of the observed values around the regression line. It gives an indication of how well the model fits the data. Lower values indicate a better fit.

5. Observations:

Value: 79

Explanation: This represents the number of data points or cases used in the regression analysis. In your case, there are 79 observations.

ANOVA Table:

Regression: This section shows the analysis of the variance attributed to the regression model. It has one degree of freedom (df), a sum of squares (SS) of 1839625.155, mean square (MS) of 1839625, an F-statistic of 10.02378, and a significance F (p-value) of 0.002215.

Residual:Represents the unexplained variance or error in the model. It has 77 degrees of freedom, a sum of squares of 14131507.6, and a mean square of 183526.1.

Total: Sum of the regression and residual variances. It has 78 degrees of freedom and a total sum of squares of 15971132.76.

Regression Coefficients:

Intercept: The intercept (constant) of the regression model is 501.5150178. The standard error is 48.23252147, and the t-statistic is 10.39786. The p-value is very low (2.53E-16), indicating that the intercept is significantly different from zero. The 95% confidence interval for the intercept is between 405.4718 and 597.5582.

Profit: The coefficient for the predictor variable "Profit" is 0.70137641. The standard error is 0.22153144, and the t-statistic is 3.166036. The p-value is 0.002215, indicating that the Profit variable is statistically significant in predicting the response variable. The 95% confidence interval for the coefficient is between 0.260251 and 1.142502**.**

**Furniture-**



1. Multiple R (Multiple Correlation Coefficient):

Interpretation: The Multiple R value of 0.0914 indicates a weak linear relationship between the dependent variable and the set of independent variables in the regression model.

2. R Square (Coefficient of Determination):

Interpretation: The R Square value of 0.0084 implies that only approximately 0.84% of the variance in the dependent variable is explained by the linear combination of the independent variables. This suggests that the model has limited explanatory power.

3. Adjusted R Square:

Interpretation: The Adjusted R Square, being even lower at 0.0038, indicates that the inclusion of the predictors does not significantly improve the model's explanatory capability. It's important to be cautious about overinterpreting the model, considering the low adjusted R Square.

4. Standard Error:

Interpretation:The Standard Error of 507.5772 represents the typical dispersion of observed values around the regression line. A higher standard error indicates that the model may not be a precise fit to the data.

5. Observations:

Interpretation: With 220 observations, the dataset is relatively large, which can enhance the reliability of statistical analyses. However, the size alone doesn't compensate for the limited explanatory power of the model.

**Conclusion-** the analysis of shopper sales data has revealed valuable insights into consumer behavior and preferences. The top-selling products, bookcases, and printers, stand out as key contributors to overall sales. This information not only aids in inventory management but also provides a strategic focus for marketing efforts to enhance product visibility and customer engagement.

Geographically, Maharashtra and Madhya Pradesh emerge as the leading states in terms of top buys, indicating a significant market presence and potential for further expansion in these regions. Retailers can leverage this knowledge to tailor marketing strategies and optimize supply chain operations to meet the demands of these high-performing markets.

Furthermore, the identification of the best buyer, Harivansh, highlights the importance of building and nurturing relationships with key customers. Understanding the preferences and purchasing patterns of such individuals allows for personalized marketing approaches, fostering customer loyalty and repeat business.

The revelation that cash on delivery is the preferred payment mode among customers provides valuable information for payment processing and logistics planning. Retailers can streamline their operations to accommodate this payment preference, ensuring a seamless and satisfactory shopping experience for customers.

In essence, this research report serves as a foundational resource for retailers and businesses aiming to enhance their understanding of shopper behavior, refine marketing strategies, and optimize operational efficiency based on the identified trends and patterns in shopper sales data.