INCOME, GENDER, AND EDUCATION: A MICROSOFT EXCEL DASHBOARD APPROACH

Khushi Tiwari

School of Computer Science and Engineering

Lovely Professional University

Phagwara, Punjab, India (144411)

Abstract:

Supermarket Sales data can be harnessed to gain valuable insights and make informed decisions for business growth. In this study, we explore the process of analysing supermarket data and creating an interactive dashboard using Microsoft Excel, with a focus on the utilization of pivot tables. Our analysis begins with the collection and organization of various types of supermarket data, including sales, inventory, customer demographics, and product information. We then use Microsoft Excel's pivot table feature to perform data manipulation and aggregation. Pivot tables allow us to efficiently summarize and analyse data, facilitating the identification of trends, patterns, and anomalies within the dataset.

Keywords: EDA, machine learning, Microsoft excel, dashboard

Introduction

Supermarkets generate vast amounts of sales data on a daily basis, which, when harnessed and analysed effectively, can provide invaluable insights for optimizing operations, enhancing customer experiences, and boosting profitability. In this context, Microsoft Excel emerges as a powerful and accessible tool for supermarket sales data analysis. One of Excel's key features, the pivot chart, plays a pivotal role in transforming raw data into actionable insights by enabling dynamic and visual representation. This analysis focuses on the process of dissecting supermarket sales data using pivot charts within Microsoft Excel. We delve into the vast ocean of transaction records, product details, customer information, and more, and explore how pivot charts can be the power of Microsoft Excel to investigate key questions about income disparities in the United States.

Our research focuses on the following key questions:

The central research question guiding this study is: "What are the patterns of income distribution in the US, and how have they evolved in recent years?" To address this question, our research objectives include:

Data Collection and Compilation: Gathering relevant data from various sources, such as government reports, surveys, and census data, to create a comprehensive dataset for analysis.

Data Analysis Using Microsoft Excel: Employing Microsoft Excel to process and analyse the collected data, allowing for the identification of key trends, disparities, and changes in income distribution.

Dashboard Creation: Developing a user-friendly dashboard in Microsoft Excel to visually represent the research findings and make the information accessible and interpretable for a broad audience.

Gender Disparities: What is the percentage of males and females in the total dataset?

Age and Income: What is the average age of people earning more than $50,000 and those earning $50,000 or less annually?

Income Classification: What is the percentage of people earning $50,000 or less?

Working Hours: What are the average working hours for males and females, respectively?

Employment Type: What is the percentage of people employed in the private working class?

Low-Income Workers: What is the percentage of individuals in the private working class earning $50,000 or less?

Gender in the Private Working Class: What is the percentage of males and females in the private working class?

Literature Review

Retail and supermarket sales analysis is a critical area of study, given the competitive nature of the industry and the constant need for businesses to optimize operations, increase revenue, and improve customer satisfaction. Leveraging Microsoft Excel as a tool for sales analysis in the context of a superstore or retail setting has gained popularity due to its accessibility and versatility. In this literature review, we examine key themes and findings from various sources related to superstore sales analysis using MS Excel. Data Analysis Tools in Retail: Many studies highlight the importance of data analysis in the retail industry. Excel is often praised for its user-friendliness and ability to handle large datasets. Researchers emphasize the need for retail managers to harness the power of tools like pivot tables, pivot charts, and other Excel features to make data-driven decisions. Inventory Management: Effective inventory management is crucial for superstores. Literature explores how Excel can be used to track inventory turnover rates, optimize reorder points, and minimize carrying costs. The role of data analysis in preventing stockouts and overstock situations is a common theme. Sales Forecasting: Accurate sales forecasting is essential for superstores to ensure they have the right products in stock. Excel's functions and capabilities are often used to develop forecasting models based on historical sales data, seasonal trends, and other factors. These models can help superstores anticipate demand and adjust inventory levels accordingly. Customer Segmentation and Behaviour Analysis: Customer data is a valuable asset in the retail sector. Studies discuss how Excel can be used to segment customers based on their purchasing behaviour, demographics, and preferences. Understanding customer behaviour helps in targeted marketing and personalized shopping experiences. Visualizations and Dashboards: Excel's pivot charts and other visualization tools are highlighted for their role in creating interactive dashboards for superstore sales analysis. Researchers point out that such visualizations make complex data more understandable for non-technical stakeholders. Competitive Analysis: Some studies focus on how superstores can use Excel for competitive analysis. This includes comparing their sales data with competitors and identifying areas for improvement. Challenges and Limitations: It is important to acknowledge the limitations of using Excel for sales analysis. These may include scalability issues with large datasets, potential for errors in manual data entry, and the need for additional tools for more advanced statistical analysis. Integration with Other Tools: Research often suggests integrating Excel with other data analysis and business intelligence tools for more comprehensive insights. This might involve connecting Excel to databases, using Power BI for advanced visualizations, or utilizing programming languages like Python for more sophisticated analysis. In conclusion, the literature on superstore sales analysis using MS Excel highlights the significance of data analysis in the retail industry. Excel's ease of use, along with its powerful features, makes it a valuable tool for superstores aiming to understand their sales data, optimize operations, and enhance customer experiences. Researchers and practitioners continue to explore innovative ways to leverage Excel and other complementary tools to gain a competitive edge in the dynamic world of retails.

Significance of the Study

This research holds significant importance for both academic understanding and policy formulation. By leveraging Microsoft Excel for comprehensive data analysis and the creation of an interactive dashboard, we aim to provide valuable insights into income

disparities in the United States. The findings of this study can inform policymakers, sociologists, and economists about the impact of several factors on income, facilitating the design of more targeted and effective policies to reduce income inequality.

The structure of this paper will include a literature review, a methodology section detailing our data analysis process using Microsoft Excel, a results section presenting the findings of our research, a discussion section interpreting these results and their implications, and a conclusion summarizing the key insights and suggesting avenues for future research and policy recommendations.

Proposed Methodology:

Data Collection:

Data Source: The primary dataset for this research was obtained from Kaggle, a comprehensive database containing demographic and income information for a representative sample of individuals in the United States. This dataset was chosen for its breadth and depth of information, allowing for a comprehensive analysis of income-related factors.

Data Preprocessing: Prior to analysis, the dataset underwent rigorous preprocessing. Missing values were imputed using appropriate methods. Outliers were identified and addressed to ensure the integrity of the data. Additionally, any irrelevant or redundant variables were removed to streamline the analysis process.

Data Analysis:

Justification of Methods:

The chosen methods align with the specific research questions and the available dataset. Each method is tailored to the nature of the data and the research objectives. Descriptive statistics, frequency analysis, and binary classification were employed because they are suitable for summarizing and categorizing data, allowing for straightforward interpretations of the results. By utilizing these methods, we aimed to provide clear and concise answers to the research questions, ensuring the findings are accessible and easily interpretable for a broad audience.

Results and Discussions:

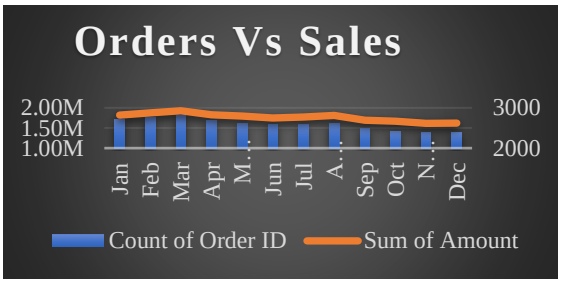

Fig (1) No. of more sales is having in the month of March compared to all Months.

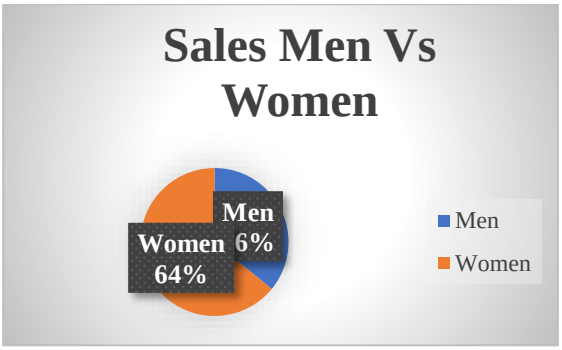

Fig (2) Around 64% of women do shopping compared to the Mens.

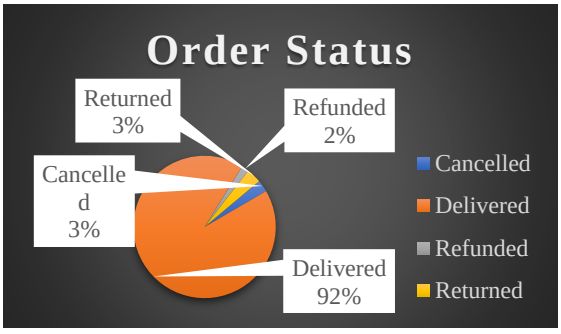

Fig (3) Here it is showing that no. of Orders is Delivered mostly .

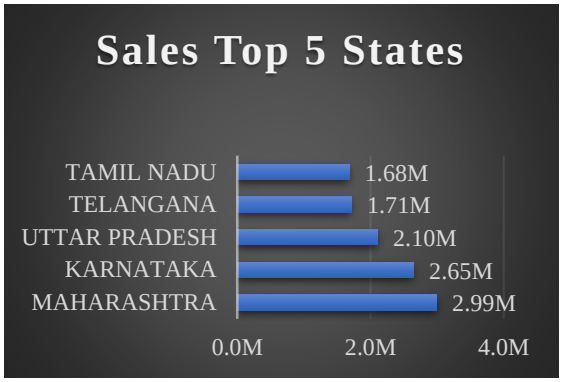

Fig (4) here is the analyse data that out of top 5 states Maharashtra have high no. of sales.

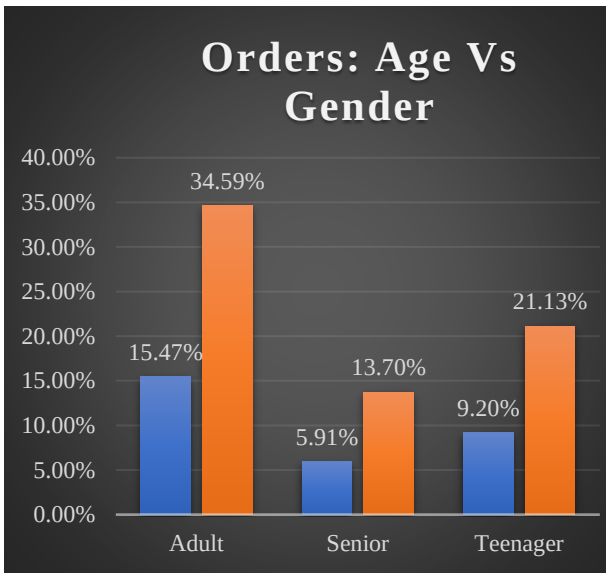

Fig (5) here is the data showing that out of 3 categories adult, senior and teenager whom no. of sales is greater.

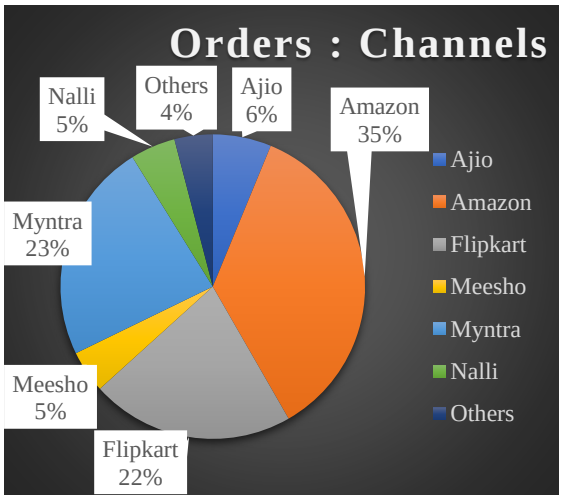

Fig (6) people purchased products from Amazon more compared to others.

Conclusion and Future Scope:

Supermarket sales analytics using pivot tables in MS Excel offers a powerful and accessible approach to gaining insights from complex sales data. This analytical process empowers supermarkets to make data-driven decisions, optimize operations, and enhance customer experiences. Here are some key takeaways from this approach: Data-Driven Decision Making: The use of pivot tables in MS Excel allows supermarkets to transform raw sales data into meaningful insights. These insights enable data-driven decision-making, which is critical for optimizing inventory, increasing sales, and improving overall business performance. Efficiency and Accessibility: Excel's pivot tables are user-friendly and require no specialized coding skills, making them accessible to a wide range of users within the organization. This efficiency in data analysis can lead to quicker response times and more informed decisions. Visualization and Interpretation: Pivot tables enable data to be summarized and presented in a visually comprehensible format. This not only aids data analysts but also helps non-technical stakeholders understand the data, fostering collaboration and alignment within the organization. Inventory Management: Supermarkets can use pivot tables to track inventory turnover rates, analyse stockouts and overstock situations, and make more accurate forecasts. This, in turn, can lead to reduced carrying costs and improved profitability. Customer Segmentation: Analysing customer data using pivot tables allows supermarkets to segment their customer base. Understanding customer behaviour, preferences, and demographics can lead to more targeted marketing campaigns and personalized shopping experiences

The future of supermarket sales analytics using pivot tables in MS Excel holds promising opportunities for further development and innovation: Integration with Advanced Analytics Tools: Supermarkets can explore integrating Excel with more advanced analytics tools or platforms like Python, R, or

dedicated business intelligence tools to enhance their analytical capabilities. Real-Time Data Analysis: Real-time data analytics can be an exciting avenue for supermarkets, allowing them to respond more rapidly to changing market conditions, customer preferences, and supply chain disruptions. Machine Learning and Predictive Analytics: The integration of machine learning models for predictive analytics can help supermarkets forecast sales trends with greater accuracy, identify anomalies, and make proactive decisions. Enhanced Customer Experience: Supermarkets can further refine their customer experience by analysing data in real-time to offer personalized discounts, recommend products, and streamline the shopping process. Sustainability and Eco-Friendly Practices: With growing environmental awareness, supermarkets can use data analytics to track the environmental impact of their operations and make decisions that align with sustainability goals. Market Basket Analysis: Utilizing pivot tables for market basket analysis can help supermarkets understand which products are often purchased together, enabling them to optimize product placement and cross-selling strategies. Supply Chain Optimization: Supermarkets can extend their use of pivot tables to analyse supply chain data for cost reduction and improved inventory management. In conclusion, the future of supermarket sales analytics using pivot tables in MS Excel is marked by ongoing innovation and integration with advanced technologies. The potential to gain deeper insights into sales data, enhance operations, and improve customer experiences is substantial, offering supermarkets a competitive edge in an ever-evolving retail landscape.

References:

1. Tariq, Dr.Tariq Hussain Sheikh and Aithal, P. S., Visualization and Explorative Data Analysis (March 26, 2023). International Journal of Enhanced Research in Science, Technology & Engineering, 12(3), 11-21. ISSN: 2319-7463, 2023, Available at SSRN: https://ssrn.com/abstract=4400256 or http://dx.doi.org/10.2139/ssrn.4400256
2. Potta Chakri, Saurabh Pratap, Lakshay, Sanjeeb Kumar Gouda,
3. An exploratory data analysis approach for analyzing financial accounting data using machine learning, Decision Analytics Journal, Volume 7, 2023, 100212, ISSN 2772-6622, https://doi.org/10.1016/j.dajour.2023.100212(https://www.sciencedirect.com/science/article/pii/S2772662223000528)
4. Nicodemo, C., & Satorra, A. (2022). Exploratory data analysis on large data sets: The example of salary variation in Spanish Social Security Data. BRQ Business Research Quarterly, 25(3), 283-294. https://doi.org/10.1177/2340944420957335
5. Vladimir Hlasny, Parametric representation of the top of income distributions: Options, historical evidence, and model selection, Journal of Economic Surveys, 10.1111/joes.12435, 35, 4, (1217-1256), (2021).
6. Perotti, R. Growth, income distribution, and democracy: What the data say. J Econ Growth 1, 149–187 (1996). https://doi.org/10.1007/BF00138861
7. Shibuya K, Hashimoto H, Yano E. Individual income, income distribution, and self rated health in Japan: cross sectional analysis of nationally representative sample BMJ 2002; 324 :16 doi:10.1136/bmj.324.7328.16
8. Anand Banerjee, Victor M. Yakovenko, T. Di Matteo, A study of the personal income distribution in Australia, Physica A: Statistical Mechanics and its Applications, Volume 370, Issue 1, 2006, Pages 54-59, ISSN 0378-4371, https://doi.org/10.1016/j.physa.2006.04.023. (https://www.sciencedirect.com/science/article/pii/S0378437106004353)