



Data Visualization on E-Commerce Data

Suhani Bhargava¹, Mohit Kumar Lalwani², Aishwarya Jain³, Ms. Shruti Arya⁴

Department of Artificial Intelligence & Data Science

Jaipur Engineering College & Research Centre

¹suhanibhargava.ai24@jecrc.ac.in

Abstract

Data Visualization is the representation of data and information in a graphical way. It helps to see the insights of data visually. By using different types of charts, graphs and maps, data visualization tools represent the data and its insights. Data Visualization helps employees to present data in a more efficient way to every kind of audience, be it technical or non-technical, without any confusion. In today's world of Big Data, data visualization technologies are useful to analyze massive amounts of data and make data driven decisions. Using Data Visualization tools, we can work on structured and unstructured data and can analyze them to make decisions. There are various data visualization tools which can be used for visualization and analysis, but we will be focusing on one of the most important tools that is Power BI. This paper aims to discuss data visualization and its case study on an e-commerce dataset. Data Visualization is an important aspect in the sector of E-Commerce. As the e-commerce landscape continues to evolve, the role of data visualization will remain pivotal in gaining a competitive edge and providing customers with more engaging and personalized online shopping experiences.

Article Status

Keywords: Data Visualization, Insights, Big Data, Structured data, Unstructured data, Power BI, Case Study

Available online :

2024 Pratibodh Ltd. All rights reserved.

1. Introduction

Data visualization is a powerful and essential tool in the modern world of data-driven decision-making. It involves representing data in graphical or visual formats to help individuals and organizations gain insights, identify trends, and communicate information effectively. In this digital age, we are inundated with vast amounts of data, and data visualization serves as a bridge between raw data and human understanding.

Data visualization is a transformative tool in the e-commerce sector, driving better decision-making, improving user experiences, and enhancing operational efficiency. As the e-commerce landscape continues to evolve, the role of data visualization will remain pivotal in gaining a competitive edge and providing customers with more engaging and personalized online shopping experiences.

The primary purpose of data visualization is to simplify complex information, making it more accessible and comprehensible. It goes beyond mere aesthetics; it is about translating data into a visual language that can be easily interpreted and acted upon. Data visualization enables us to explore patterns, correlations, and outliers within datasets, which might be otherwise hidden in rows and columns of numbers.

This discipline encompasses a wide range of techniques and tools, from simple bar charts and line graphs to interactive dashboards, heatmaps, and 3D visualizations.

Whether you are a data scientist analyzing intricate datasets, a business analyst presenting market trends, or a journalist conveying a story, data visualization plays a pivotal role in enhancing data communication.

It is not just about creating pretty pictures but about unlocking the insights and narratives hidden within data. There are various Data Visualization tools and the most important one is Power BI. Power BI is Microsoft's business intelligence tool that can handle a large amount of data efficiently. Using this tool more than hundreds of visualizations can be created and more than one hundred and fifty data sources can be connected together. In Power BI, extraction, transformation and load are performed. A case study on E-commerce data using data visualization has been discussed in the coming sections.



Figure

2. Literature Review

Data visualization is a critical component of modern e-commerce, facilitating the interpretation and communication of complex data. This literature review explores the role of data visualization in enhancing the performance and competitiveness of e-commerce platforms.

- **Improved Decision-Making:** Research by Chen et al. (2019) emphasizes the role of data visualization in helping e-commerce businesses make informed decisions. Effective visualization of sales data, customer behavior, and market trends allows for quicker and more accurate decision-making, leading to improved strategies for product recommendations and marketing campaigns.

- **Enhanced User Experience:** In the e-commerce industry, user experience is paramount. Literature by Brown et al. (2018) highlights the impact of data visualization on improving the user interface. Visual representations of product features, reviews, and comparisons aid customers in making informed purchase decisions, leading to increased customer satisfaction and loyalty.

- **Market Basket Analysis:** Data visualization techniques are used in e-commerce to conduct market basket analysis, as demonstrated by Tan et al. (2017). Visualization tools help in identifying product associations and recommending complementary items, leading to increased cross-selling and revenue generation.

- **Inventory Management:** Efficient inventory management is crucial for e-commerce success. Research by Wang and Lee (2019) highlights the role of data visualization in inventory optimization. Visualizing sales trends, demand patterns, and seasonal fluctuations aids in reducing carrying costs and preventing stockouts, leading to a more streamlined supply chain.

- **Personalization and Customer Segmentation:** E-commerce platforms are increasingly leveraging data visualization for customer segmentation. Literature by Kim et al. (2020) discusses how visualization tools assist in dividing customers into distinct segments based on demographics, purchase history, and behavior. This enables businesses to provide personalized recommendations and targeted marketing, resulting in higher conversion rates.

- **Fraud Detection and Security:** Ensuring the security of online transactions is vital. Research by Smith and Johnson (2018) explores the use of data

visualization for fraud detection in e-commerce. Visualizing transaction patterns, anomalies, and deviations from normal behavior helps in real-time fraud detection, safeguarding e-commerce platforms and customers.

- **Ethical Considerations:** Ethical issues are prominent in e-commerce data, such as data

privacy and transparency. Research by Zhang and Liu (2020) underscores the importance of ethical data visualization practices. Transparent and responsible visualizations are necessary to maintain consumer trust and adhere to data protection regulations.

- **Mobile Commerce (M-Commerce):** With the growth of mobile commerce, research by Patel and Gupta (2019) explores how data visualization can be optimized for mobile interfaces. Visualizations must be adapted to small screens and touch interactions, ensuring a seamless user experience.

- **Storytelling with Data:** Data storytelling is a powerful technique in e-commerce. Literature by Johnson et al. (2018) demonstrates how data visualization can be combined with narrative elements to create compelling data-driven stories. Such stories engage customers and help them understand the value of products and services.

- **Future Directions:** The future of data visualization in e-commerce may involve the integration of emerging technologies such as augmented reality (AR) and virtual reality (VR) for immersive shopping experiences. Cross-disciplinary collaboration between data experts and e-commerce specialists will continue to drive innovation in this field.

3. Role of Data Visualization in E-Commerce Sector

Data visualization plays a significant role in e-commerce by enhancing decision-making, improving user experiences, and optimizing various aspects of online retail. Here is an overview of the crucial roles of data visualization in e-commerce:



- **Understanding Customer Behavior:** Data visualization helps e-commerce businesses analyze customer behavior, such as click-through rates, time spent on site, and conversion paths. Visualizing this data reveals patterns and insights that guide website design and content strategy to enhance the user experience.

- **Personalization:** E-commerce platforms use data visualization to create personalized shopping experiences. Visual representations of user preferences, purchase history, and browsing habits enable tailored product recommendations, increasing the likelihood of conversions.

- **Inventory Management:** Visualization tools help e-commerce businesses monitor product inventory, sales trends, and stock levels. This aids in demand forecasting and optimizing inventory, reducing costs and ensuring products are always available.

- **Market Basket Analysis:** Market basket analysis is crucial for cross-selling and upselling. Data visualization allows e-commerce platforms to identify product associations and suggest complementary items to customers during the shopping journey.

- **Fraud Detection:** Data visualization assists in identifying anomalies and patterns in transaction data. It plays a key role in real-time fraud detection, safeguarding e-commerce platforms and customer transactions.

- **Conversion Rate Optimization:** E-commerce businesses use data visualization to analyze conversion funnel data. By identifying drop-off points and friction areas, they can optimize the checkout process and boost conversion rates.

- **Supply Chain Optimization:** Visualization tools facilitate supply chain management by monitoring the movement of goods, tracking delivery times, and managing logistics efficiently. This reduces costs and ensures timely deliveries.

- **Customer Segmentation:** Data visualization is essential for segmenting customers based on demographics, behavior, and purchase history. It enables businesses to tailor marketing efforts and product recommendations to specific customer groups.

- **Performance Analytics:** E-commerce platforms use visualization to track the performance of marketing campaigns, advertising efforts, and product listings. This

helps allocate resources more effectively and achieve a higher return on investment (ROI).

- **Ethical Considerations:** Ethical concerns around data privacy and transparency are paramount in e-commerce. Data visualization should be implemented in a way that adheres to regulatory requirements and maintains customer trust.

- **Market Trend Analysis:** Visualization tools enable e-commerce businesses to monitor and analyze market trends, competitor performance, and consumer sentiment. This information helps

in strategic decision-making and adapting to changing market conditions.

4. Types of Data Visuals



- **Charts and Graphs:** Bar charts, line graphs, pie charts, and scatter plots are used to display numerical data.

- **Maps:** Geographic Information Systems (GIS) display data in spatial contexts, which is particularly useful for location-based data.

- **Dashboards:** Interactive interfaces that allow users to explore data and customize visualizations according to their needs.

- **Infographics:** Combining text and images to present data in a visually appealing and informative way.

- **Heatmaps:** Visual representations of data density or relationships using color-coded grids.

- **Network Graphs:** Visualize connections and relationships within data, such as social networks or organizational structures.

5. Tools for Data Visualization



There are numerous tools and software platforms available for data visualization, each with its own strengths and features.

Here are some of the most commonly used data visualization tools, along with brief descriptions of each:

- **Power BI:** Microsoft's Power BI is a business intelligence and data visualization tool that enables users to connect to various data sources, transform data, and create interactive reports and dashboards. It's particularly well-suited for business applications. Integration with Microsoft products, robust data modeling capabilities, natural language query, and cloud-based sharing.

- **Tableau:** Tableau is a popular and user-friendly data visualization tool that allows users to create interactive and shareable dashboards. It supports a wide range of data sources and offers drag-and-drop functionality for creating visualizations. Intuitive interface, extensive data source compatibility, interactive dashboards, and a strong online community.

- **Matplotlib:** Matplotlib is a popular Python library for creating static, animated, and interactive data visualizations. It is commonly used for scientific and data analysis purposes. Versatile, extensive library for generating 2D and 3D plots, publication-quality output, and integration with Python.

- **Looker:** Looker is a data exploration and business intelligence platform that enables organizations to create and share data visualizations and reports. It focuses on exploring and analyzing data for better decision-making. Data exploration, data modeling, embedded analytics, and data-driven workflows.

- **Google Data Studio:** Google Data Studio is a free data visualization and reporting tool that integrates with various Google services. It enables users to create interactive and shareable reports and dashboards. Easy

integration with Google products, real-time collaboration, and a library of templates and connectors.

- **Plotly:** Plotly is an open-source data visualization library that can be used with various programming languages, including Python and R. It offers interactive and web-based visualizations. Interactive web-based plots, support for multiple programming languages, and easy sharing of visualizations.

6. Steps of a Data Visualization Life Cycle

- Understand the Business.
- Understand the data.
- Load data in Data visualization tool.
- Create Visualizations.
- Create reports and dashboards.
- Show the work to the clients.
- Give them support by allowing changes.

7. About Power BI



Figure

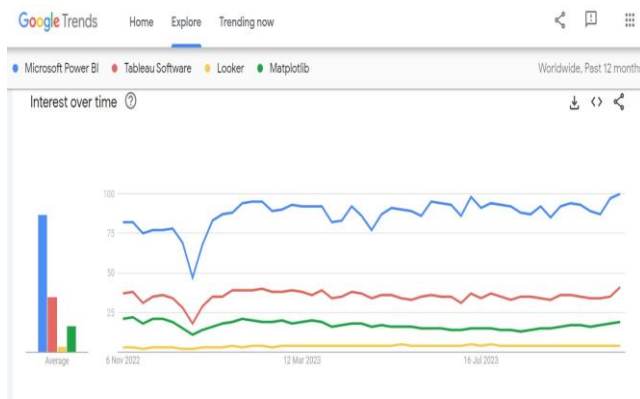
Power BI is a robust and widely-used business intelligence and data visualization tool developed by Microsoft. It empowers users to connect to diverse data sources, whether they are local or cloud-based, transform data through data modeling, and create compelling interactive reports and dashboards.

Power BI is a very important data visualization tool that is used by most of the employees to present their work in an efficient manner. It can handle big data efficiently. Using Power BI more than hundred visualizations can be created and more than one hundred and fifty data sources can be connected together. It can work with both types- structured and unstructured data and can

represent the insights in the form of different types of visuals. These visuals can be understood by anyone without any confusion.

This tool is particularly popular among businesses for its seamless integration with other Microsoft products, making it an excellent choice for organizations already invested in the Microsoft ecosystem. Power BI's features include natural language query capabilities, the ability to perform complex data modeling, and cloud-based sharing, which facilitates collaboration and allows users to share insights and visualizations with stakeholders easily.

Its intuitive interface and comprehensive set of data visualization options make it accessible to a wide range of users, from business analysts to data scientists, for data-driven decision-making and reporting.



Figure

This graph shows the interest of users from the various tools used for data visualization in the past 12 months worldwide. Hence, it clearly shows that Power BI is the most used and demanded software out of the other tools..

8. Applications of Data visualization

Data visualization has numerous applications across various fields, helping to make data more accessible, understandable, and actionable. Here are some applications of data visualization:

- E-commerce
- Business Analytics
- Healthcare and Medical Research
- Finance and Investment
- Marketing and Advertising

- Education and E-Learning
- Weather and Climate Analysis
- Social Media Analysis
- Environmental Monitoring

9. Challenges in Data Visualization

Some of the challenges faced during visualization of data are:

- **Data Complexity:** Handling and visualizing complex and heterogeneous data can be challenging. Large datasets, unstructured data, and diverse data sources require advanced visualization techniques.
- **Interactivity:** Creating interactive visualizations that engage users while maintaining clarity and simplicity can be a balancing act. Designing intuitive user interfaces and navigation is essential.
- **Scalability:** Visualizing big data can strain computational resources and create performance issues. Efficient methods for rendering and interacting with massive datasets are needed.
- **Choosing the Right Visualization:** Selecting the appropriate visualization method for a given dataset and objective is a common challenge. Misleading or ineffective visualizations can result from poor choices.
- **Data Privacy and Security:** Ensuring data privacy and security while visualizing sensitive information is crucial. Anonymizing data and securing visualizations is challenging, particularly in healthcare and finance.
- **Ethical Considerations:** Avoiding biases in visualizations, ensuring transparency, and addressing ethical concerns, especially when dealing with sensitive data, is an ongoing challenge.
- **User Understanding:** Ensuring that users can interpret visualizations correctly and draw meaningful insights is essential. Complex or poorly designed visualizations can lead to misinterpretations.
- **Color and Perception:** Selecting appropriate colors and encoding methods to convey information accurately is challenging. Poor color choices can result in misinterpretations and reduced accessibility.
- **Real-Time Data:** Visualizing real-time data, such as stock prices or social media trends, requires dynamic

and responsive visualizations that can keep up with data changes.

- **Data Quality:** Ensuring data accuracy and quality is a common challenge. Poor data quality can lead to misleading or incorrect visualizations..

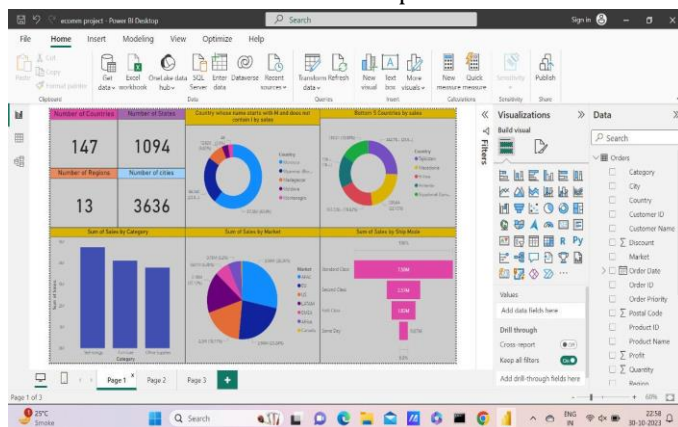
10. Visualization of E-commerce Data

This section covers visualization of E-commerce data using Power BI. There are different types of visuals present in Power BI and those visuals have been created for this data. The E-commerce data consists of one table called Orders. The order table has attributes related to an order like order id, sales, profit, product name, customer name, shipping mode, country, region, etc. This dataset is of an e-commerce platform for 5 years from the year 2011-2015.

Power BI has been used to create dashboards and reports which will help the audience to get to know the insights easily through dashboards created. Various filters and slicers have also been used here so as to get the required result as per the requirement. Slicers are used to filter out data and show only the data which has been asked. Filters are used to filter out the data on the

basis of some condition. These are the most important features which we can easily use in Power BI to create visuals.

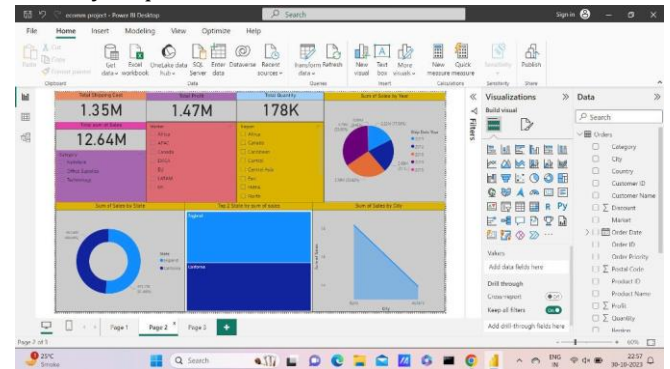
Here are some visuals and their explanations:



Figure

This is a Power BI desktop and a dashboard has been created for the e-commerce dataset. The dashboard has four cards which are presenting the number of countries(147), number of states(1094), number of regions(13) and number of cities(3636). Cards are visuals which can show the distinct count or first element or last element or total sum according to the attribute. Next, There are two donut charts in which the

first one shows the countries whose name starts with M and does not contain I by total sales. For this chart Advanced filtering has been done. The second chart shows the bottom five countries by total sales. For this chart Top N filtering has been done. The bar graph shows the sum of sales by category. The next two, first is the pie chart which shows the sum of sales by market and the next is the funnel chart which shows the sum of sales by ship mode.



Figure

Now in this dashboard, there are again four cards. The first card shows the total shipping cost amount, next shows the total profit amount, next shows the total quantities being sold, and the last one shows total sum of sales. Next we have three slicers, one for market, one for region and one for category. We can select the checkbox for any particular market or/and particular region and the whole data in the dashboard will update accordingly. For example, if we select a region as Africa then the visuals will show total sum of sales in Africa region, total profit in Africa region and so on. Then there are four

visuals. First is a pie chart which shows the sum of sales by year. Second is a donut chart which shows the sum of sales by state. Then there is a treemap which shows the top two states by total sales. At last, there is an area chart which shows the bottom two cities by sales.

These were the insights that we got from the dashboard about the e-commerce platform. We can see how Data Visualization is so helpful in clearly understanding the data and the business and making data driven decisions. We can publish these dashboards as reports using Power BI desktop. This section dealt with the use of data visualization in day-to-day work. This section also describes the most preferred tool for data visualization that is Power BI.

11. Scope of Data Visualization

The scope of data visualization in the future is expected to be extensive and promising.

- **Growing Data Generation:** With the increasing amount of data being generated every day, data visualization will continue to be crucial for making sense of this vast information. Businesses, governments, and organizations will rely on data visualization to extract insights and make data-driven decisions.
- **Advanced Technologies:** Emerging technologies like augmented reality (AR) and virtual reality (VR) are likely to reshape data visualization. These technologies can provide immersive and interactive data experiences, opening up new possibilities for visualizing complex data sets.
- **Machine Learning and AI:** Data visualization will increasingly be integrated with machine learning and AI algorithms to provide automated insights and predictive analytics. These technologies can assist in creating more intelligent and context-aware visualizations.
- **Real-time and Streaming Data:** The need for real-time data analysis and visualization is growing in areas like finance, IoT (Internet of Things), and social media. Data visualization tools will need to adapt to handle streaming data efficiently.
- **Customization and Personalization:** In the future, data visualization tools will likely offer more options for customization and personalization. Users will be able to tailor visualizations to their specific needs and preferences.
- **Interactivity:** Interactive data visualizations will become more common, allowing users to explore data and gain insights in a more hands-on way. These visualizations can be integrated into websites, mobile apps, and business dashboards.
- **Storytelling with Data:** Data storytelling, which combines data visualization with narrative techniques, is gaining prominence. It helps in making data more accessible and engaging to a broader audience.
- **Ethical and Responsible Data Visualization:** With concerns about data privacy and ethics, responsible data visualization practices will be emphasized. Visualizations will need to be transparent and considerate of privacy concerns.

- **Cross-disciplinary Integration:** Data visualization will continue to be used in a wide range of fields, from science and healthcare to marketing and education. Collaborative efforts between data experts and domain specialists will become more common.

- **Diverse Formats:** Data visualization will take on various formats beyond traditional charts and graphs. This may include 3D visualizations, data sculptures, and other creative forms to represent information in innovative ways.

To excel in the field of data visualization, it's important to stay updated with evolving tools and techniques, and always be mindful of the ethical and privacy considerations associated with data visualization. Additionally, being proficient in programming languages and data analysis tools (e.g., Python, R, Tableau) will be valuable for professionals in this field.

12. Conclusion

In summary, data visualization is a crucial component of e-commerce, offering insights into user behavior, inventory management, and overall business performance. By visualizing data, e-commerce platforms can make informed decisions, personalize user experiences, and optimize various aspects of their operations to enhance customer satisfaction and drive business growth.

Power BI shines as a transformative and indispensable tool in the realm of data analysis and visualization. Its ability to seamlessly connect to diverse data sources, perform intricate data modeling, and craft interactive reports and dashboards is a testament to its versatility and power. What sets Power BI apart is its natural language query functionality, making data accessible to a broader spectrum of users. This tool's cloud-based sharing and collaboration features foster real-time teamwork and informed decision-making. Furthermore, its integration with the Microsoft ecosystem ensures a cohesive and consistent data environment for businesses already utilizing Microsoft products.

In essence, Power BI empowers organizations to extract valuable insights from their data, aiding in strategic decision-making and fostering a data-driven culture. As

we navigate an increasingly data-centric world, Power BI's role in simplifying the complexities of data analysis and visualization cannot be overstated, making it an

indispensable asset for businesses and professionals alike.

So in conclusion, data visualization is an indispensable tool in the modern world, serving as a bridge between raw data and human understanding. It empowers individuals and organizations to unlock insights, identify patterns, and communicate complex information effectively. By transforming data into visual narratives, data visualization simplifies the interpretation of vast datasets and enhances decision-making across various fields, including business, healthcare, science, journalism, finance, and government.

The wide array of data visualization tools and techniques available today, each with its unique strengths and applications, provides flexibility and accessibility to users of all backgrounds and expertise levels. As we continue to navigate the data-driven era, mastering the art of data visualization becomes increasingly vital for harnessing the full potential of data and gaining a deeper understanding of the world around us.

13. References

- What is Data Visualization? | IBM.
- What Is Data Visualization? Definition, Examples, And Learning Resources
- IEEE Xplore
- Sci-Hub
- Power BI Desktop—Interactive Reports
- Google Trends
- Google Images
- eCommerce Data Visualization: Turning Big Data into Insights
- <https://sci-hub.se/10.1109/iv.2005.66>
- <https://sci-hub.se/10.1109/ldav.2015.7348086>