A

Project

Report On

"Amazon Sales Data Analysis for E-Commerce Sales And Revenue Trends Using Business Intelligence Tools"

At

"iNeuron"
By

Vivek Satendra Prasad

Under the guidance of

Prof. Pramila Pareek

Submitted to

Savitribai Phule Pune University

In partial fulfillment of the requirement for the award of the degree of

M.B.A

Batch – 2022 - 2024

Through

JSPM's Rajarshi Shahu College of Engineering – Tathawade
Pune 33

ANNEXURE D

DECLARATION OF

STUDENT

(CERTIFICATE OF ORIGINALITY/DECLARATION)

This is to declare that I have carried out this project work myself in partial fulfillment of the MBA Program of Savitribai Phule Pune University.

The work is original, has not been copied from anywhere else and not been submitted to any other University/Institute for an award of any degree/diploma.

Date   Signature

Place   Name:

ANNEXURE E

DECLARATION OF GUIDE

This is to certify that the work incorporated in this Project report **Amazon Sales Data Analysis for E-Commerce Sales And Revenue Trends Using Business Intelligence Tools"** submitted by **Vivek Satendra Prasad** is his original work and completed under my guidance. Material obtained from other sources has been duly acknowledged in the Project Report.

DATE   SIGNATURE OF GUIDE

PLACE

ANNEXURE F

ACKNOWLEDGEMENT

Table of Content

List Of Figures

CHAPTER 1

1.1 EXECUTIVE SUMMARY

This study aims to conduct an extensive examination of revenue and sales trends in the electronic products sector, encompassing gadgets such as phones, headphones, laptops, chargers, batteries, and monitors. The primary objective is to uncover data that facilitates the identification of key markets, allowing the company to more effectively focus its marketing initiatives in order to increase revenue and sales. The project aims to give readers a comprehensive understanding of the dataset created from Amazon sales data for 2022 by using Microsoft PowerBI for data visualisation and Python for data transformation, modification, and cleaning.

Finding patterns and anomalies in the sales data is the primary goal. Using the PowerBI tool, interactive dashboards that convert complex information into actionable insights are then created. By carefully reviewing important factors like orderID, product details, product category, the quantity ordered, price, date of purchase, city, and state, the research looks for subtle trends and linkages.

The project expects to handle problems with data quality and make sure the dataset is ready for analysis by utilising Python's powerful features. With PowerBI selected as the visualisation tool, a visually appealing story will be presented, improving understanding and decision-making.
Essentially, the initiative seeks to offer a revolutionary method for making decisions based on data, in addition to providing a thorough study of Amazon sales for electronic goods. The core of this effort is the combination of PowerBI

and Python as tools with a large dataset, which promises a sophisticated analysis of revenue and sales trends that will steer the business towards strategic expansion in the cutthroat market for electrical items.

1.2 OBJECTIVES OF THE STUDY

1. To perform data wrangling techniques on the given data using python.
2. To identify major markets and regions contributing significantly to sales and revenue using Business intellegence tool- Power BI .
3. To analyze the performance of individual product categories.
4. To uncover patterns and trends in consumer behavior, considering factors like quantity ordered and pricing.
5. To provide actionable insights that can support data-driven decision-making for increasing sales and revenue.

1.1 SCOPE OF THE STUDY

This project's scope is focused on providing a thorough analysis of the Sales and Revenue trends in the electronic products industry, with a particular emphasis on devices like phones, headphones, laptops, chargers, batteries, and monitors. The project intends to offer priceless insights into market dynamics, making it possible to identify important markets in terms of revenue and sales.

The main goal is to use advanced data analysis and visualisation methods to extract useful information from the dataset, which is made up of 2022 Amazon sales data. Significant parameters including orderID, product information, product category, quantity ordered, price, purchase date, city, and state are included in the dataset. There is a chance to identify intricate patterns and connections in consumer behaviour and market trends in this data-rich environment.

The use of Python for data transformation, cleaning, and manipulation is now included in the project's scope. Python is a flexible programming language that is selected because of its capacity to manage advanced data processing tasks and handle complex data structures. To turn unprocessed data into visually appealing dashboards, the PowerBI tool data visualisation is essential. While these dashboards are built, stakeholders will have a comprehensive understanding of the sales and revenue picture, making them effective tools for decision-making.

This study's scope is essentially a detailed analysis of electronic goods sales on Amazon in 2022, with a focus on delving into the dataset's complexities through the use of strong tools and methodologies. Through adherence to the specified goals and utilisation of Python and PowerBI, this project seeks to enable data-driven decision-making, steering the organisation towards strategic endeavours targeted at augmenting revenue and sales in the fiercely competitive electronic products industry.

1.3 NEED OF THE STUDY

The expanding significance of e-commerce platforms and the rising demand for electronic goods are the reasons behind the necessity for this study. It is imperative for businesses to comprehend sales and revenue patterns on a platform such as Amazon in order to optimise their strategy, improve their product offers, and successfully target important markets. The project will support the company's ability to make strategic decisions by allowing it to better match its efforts with consumer preferences and market dynamics.

1.4 LIMITATION OF THE STUDY

1. The analysis is based solely on Amazon sales data for the year 2022, limiting the scope to one platform.
2. The dataset includes only specific product categories, potentially excluding other relevant electronic products.
3. External factors impacting sales, such as economic conditions or competitor activities, are not considered in this analysis.
4. The accuracy of the findings depends on the quality and completeness of the provided dataset.

CHAPTER 2: COMPANY PROFILE/ ORGANISATION PROFILE

2.1 Introduction to Organization

Amazon.com, Inc. is a company that offers services for online retail buying. Jeffrey P. Bezos launched the business in July 1994, and its primary headquarters is in Seattle, Washington. Based on revenue and market capitalization, the multinational technology company is the biggest online retailer globally, ranking second only to Alibaba Group in terms of total sales. For the US, UK and Ireland, France, Canada, Germany, Italy, Spain, Netherlands, Australia, Brazil, Japan, China, India, and Mexico, Amazon maintains distinct retail websites. a few websites, including www.amazon.com, www.amazon.ca, www.amazon.br, etc. Some of Amazon's products can also be shipped internationally to a few additional nations.

2.2 Vision

"To be Earth's most customer-centric company, where customers can find and discover anything they might want to buy online."

The organization's primary goal of becoming the greatest e-commerce business in the world is emphasised in this vision statement. The elements or qualities that Amazon emphasises in its vision statement are as follows:

1. Worldwide scope
2. Prioritisation of customers
3. The widest selection of goods

2.3 Mission

"We strive to offer our customers the lowest possible prices, the best available selection, and the utmost convenience."

This mission statement guarantees a customer-satisfied, eye-catching e-commerce service. The following elements or characteristics can be found in Amazon's mission statement:

1. The lowest costs
2. The best option
3. Maximum convenience

2.4 Key Achievements

Being the most customer-focused organisation on Earth has been our goal from the beginning. It is a privilege for us to be acknowledged for the efforts we make on behalf of our clients, staff, and global communities. These are a few honours we've won recently.
This is an exhaustive list of some of the factors that set Amazon apart from other companies in its field.

1. Price and Loyalty.
2. Decisions that are supportive.
3. Quick and easy to use.
4. Platform-Specific.
5. Service-based content.
6. All-Channel.
7. Curation and unique products.

2.5 Key Products or Services
The following items are available on Amazon: Kindle, Fire TV, Fire OS, Echo, and Fire Tablet. All of these were created with the needs of clients of all ages in mind. Numerous additional goods and services are offered by Amazon, such as:

1. Retail Goods-Amazon product lines include several media (books, DVDs,music CDs,videotapes, and software), apparel, baby products, consumer electronics,beauty products, gourmet food, groceries, health and personal-care items, industrial & scientific supplies, kitchen items,jewelleryand watches, lawn and gardenitems, musical instruments, sporting goods, tools, automotive items and toys & games.
2. Amazon Fresh- Provides grocery and staple products to the customers at affordable prices
3. Amazon Prime Video- Amazon has acquired satellite rights to show movies, web series, original amazon series, kids content, etc
4. Alexa- AI-based virtual assistant that helps with day-to-day things
5. Amazon Drive

6. Kindle Store- A Virtual bookstore that has stored millions of books by millions of writers virtually (on payment)
7. Music- Listen to your favorite song and music from all over the world.
8. Amazon Digital Software & Video Games
9. Amazon Studios
10. AmazonWireless

2.6 SWOT Analysis of the company

Figure 2.6.1 – SWOT Analysis

Strengths

1. Strong brand recognition and reputation: Customers are familiar with both the Amazon name and logo, which highlights the company's great brand recognition and reputation. Its strong position is a result of both its excellent advertising campaigns and its longevity in the market. Due to its client-centric strategy, which includes top-notch customer support and a variety of delivery alternatives, Amazon has gained the confidence and dependability of its customers. A company's reputation and brand awareness are extremely important because they persuade potential customers to buy

its goods or services. A strong brand helps set a business apart in the competitive marketplace of e-commerce and attracts in and keeps customers. Due in significant part to its success and reputation, Amazon has become one of the most powerful corporations in the world.

2. Diverse product offerings- One of Amazon's greatest assets is its capacity to appeal to a wide range of consumers and meet their needs. The company sells a wide range of things, including apparel, books, electronics, household goods, and more. This broad range of products has been crucial to the company's success since it has allowed it to maintain its competitiveness in the face of shifting customer tastes and industry trends and to secure a sizeable portion of the e-commerce market. Amazon's partnerships and collaborations with other businesses are the primary reason for the company's continued success in providing such a wide selection of products. The company has formed partnerships with numerous brands and manufacturers, which allows it to offer a wide range of products to its customers. Amazon has also leveraged its technology and data capabilities to identify and respond to changing consumer preferences, which has allowed the company to adapt its product offerings to meet the needs of its customers. This combination of partnerships and data-driven decision-making has helped Amazon remain a leader in the e-commerce industry and maintain its diverse product offering.

3. Well-established distribution network- One of Amazon's main competitive advantages is its vast distribution network, which enables the company to deliver goods to customers anywhere in the world in a timely and economical manner. The company has established an extensive global network of fulfilment centres, distribution centres, and sorting facilities in order to efficiently store and deliver goods to customers at a reasonable cost. To make sure that goods are delivered to customers in a timely and convenient way, Amazon uses a range of delivery methods in addition to its own facilities, such as its own fleet of delivery trucks and agreements with outside delivery businesses. Amazon's use of technology and data has been critical to the success of its distribution network. The company has invested heavily in technology and data analytics to optimize its fulfillment and delivery operations and to continuously

improve its efficiency and effectiveness. This includes the use of advanced algorithms to predict demand and optimize inventory management, as well as the use of robotics and automation in fulfillment centers. These efforts have allowed Amazon to maintain a well-established distribution network that can efficiently deliver products to customers all over the world, which has played a significant role in the company's success.

4. Cutting-edge technology- Amazon has made extensive use of cutting-edge technology, particularly in the areas of data analytics and artificial intelligence (AI). The business has made significant investments in AI and data analytics to streamline processes and enhance customer satisfaction. For instance, the business employs data analytics to forecast demand, improve inventory control, and target clients with customised recommendations and adverts. In order to enhance customer assistance and service, the business has also deployed chatbots and virtual assistants driven by AI. Amazon has used cutting-edge technology in various parts of its company in addition to data analytics and AI. In order to increase productivity and accuracy, the corporation has integrated sophisticated robots and automation into its fulfilment centres.. It has developed a wide range of innovative products and services, such as its Amazon Web Services (AWS) cloud computing platform and its Amazon Prime streaming service. These initiatives have been crucial to Amazon's success in helping it stay at the forefront of technical innovation.

5. Strong financial performance- One of Amazon's main advantages is its solid financial performance, which shows the corporation can make money and expand its operations. The company has claimed strong financial success, and revenue and earnings have climbed year after year. A number of things contribute to this impressive financial performance, such as the company's wide range of products, established distribution system, and state-of-the-art technology. The business has strong revenue and profit streams to go along with a strong balance sheet that includes significant cash reserves and low debt levels. Due to its financial stability, the business is able to weather obstacles such as economic downturns and undertake significant acquisitions and investments. Being financially strong has enabled Amazon to maintain its position as a

market leader in the technology and e-commerce sectors.

6. Large customer base- Among the main factors contributing to Amazon's success is the enormous range of products it offers. The company's success in the e-commerce industry can be largely credited to the diversity of products it offers, which has helped it draw in and keep clients from a range of demographics. Furthermore, Amazon has been able to draw in and hold on to customers because to its well-known brand, stellar reputation, and customer-first philosophy. Strong product delivery capabilities across the globe are made possible by Amazon's robust distribution network, which has also played a significant role in the company's vast client base. The organisation has been able to draw in and keep consumers because to its dependability and convenience. The size of Amazon's customer base has been crucial to the company's success and has helped the company become a leader in the e-commerce industry.

7. Wide range of services, including e-commerce, cloud computing, and streaming media- Amazon has expanded its services beyond e-commerce by way of its Amazon Web Services (AWS) subsidiary, which offers a range of cloud computing capabilities to businesses and organisations worldwide. As a pioneer in the cloud computing space, AWS has assisted Amazon in reducing its reliance on the e-commerce sector and diversifying its sources of income. One illustration of the company's growth outside of its primary cloud computing sector is Amazon Prime. This service has grown to be a significant player in the streaming media market by providing customers with access to a large selection of films, TV series, and other content. These initiatives have been crucial to Amazon's success in helping the corporation diversify its revenue sources.

8. Strong partnerships and collaborations- This strength is demonstrated by Amazon's ability to establish and maintain fruitful partnerships and alliances with a wide range of companies and organisations. One of the most prominent instances of Amazon's effective utilisation of cooperation and partnerships is the company's ties with its vendors and suppliers. The company offers a wide variety of items to its clients by collaborating with numerous brands and producers. Through these partnerships, Amazon has been able to decrease its reliance on a single source and broaden its product offering. In addition to suppliers and vendors, Amazon has established

relationships with companies in numerous other industries. To enhance its delivery capabilities, for instance, the corporation has partnered with logistics and transportation firms and worked with healthcare institutions to develop innovative healthcare solutions. These partnerships and collaborations have helped Amazon expand its business and have played a significant role in the company's success.

9. Global presence with operations in numerous countries- Amazon has effectively expanded its operations to a global scale, demonstrating its resilience. Amazon's global success can be attributed, in large part, to its e-commerce platform. It allows the business to connect with customers wherever in the world. Amazon has become a dominant force in the online retail space and has captured a significant share of the market thanks to its global reach. Apart from its online store, Amazon has also made a name for itself in a number of other nations through its other business endeavours, such its Prime streaming service and its Amazon Web Services (AWS) division. The company's global footprint and revenue sources have been diversified and expanded thanks to these initiatives.

Amazon Weaknesses

1. Dependence on third-party sellers- A significant amount of Amazon's product offers are provided by third-party merchants, which poses particular hazards for the business. The chance of fraud or the sale of subpar goods on the platform is one of the primary concerns connected to this dependency. Negative consumer experiences with these products could harm Amazon's brand and result in lower sales. Furthermore, Amazon's dependence on independent sellers may provide problems for both customer happiness and product control. Amazon does not have total control over the goods that are offered on its marketplace, even if it does have rules for vendors. This may result in problems with the availability, quality, and delivery of the goods, which may affect the experience of the customer. ensuring third-party vendors adhere to Amazon's standards and provide high-quality products is essential to maintaining customer trust and satisfaction.

2. Data security concerns- Any business handling sensitive data must make sure that

consumer data is secure, and Amazon is no different. Because it manages a lot of consumer data, including financial and personal data, the corporation may become the target of hackers. Identity theft and financial losses could be among the serious problems that would arise for Amazon and its customers if this data were exposed. Amazon has put in place a number of safeguards, such as encryption and secure servers, to protect consumer data in order to address these data security issues. But since the threat of cyberattacks is ever-changing, Amazon needs to keep up its guard to secure client information. If this isn't done, the business may lose the trust of its clients and suffer negative effects.

3. A business model that is easy to copy- The ease with which Amazon's business model can be copied is one of its possible weaknesses. Being the industry leader in online retail, Amazon has created a formula for online product sales that combines a large selection, quick shipment, and a convenient shopping experience. But other businesses may simply copy this strategy and take on Amazon in the marketplace. Due to its lack of distinctiveness, Amazon may find it more challenging to set itself apart from rivals and hold onto its market share. It also implies that in order for the firm to keep ahead of the competition, it must constantly innovate and change its business model.

4. Regulatory issues- Amazon may incur risks and difficulties as a result of regulatory scrutiny it has received in a number of markets. Amazon has faced a number of major regulatory challenges, including questions about its business operations and possible antitrust infractions. For instance, Amazon has come under fire for how it handles independent merchants on its marketplace and for allegedly using user data to give itself an unfair competitive edge. Additionally, the business has been charged with engaging in anticompetitive behaviour, such as squeezing out smaller rivals by abusing its monopoly in the online retail sector. These legal problems may lead to expensive litigation and harm to the company's image, which may have an effect on its expansion and profitability. Amazon needs to manage these regulatory obstacles in order to maintain its compliance and avoid negative consequences.

5. Limited customer loyalty- Despite having a huge customer base, some consumers could be more devoted to particular brands or goods than to the Amazon website. This

low level of client loyalty could be a liability for the business because it might be harder to keep these clients if they have other options.Customers may be more inclined to go to competitors if they believe they are not getting the best deal or service from Amazon, which is one possible effect of low customer loyalty. The company may experience a decline in market share and sales as a result of this. Furthermore, low customer loyalty can make it harder for Amazon to enter new areas because it might encounter more competition from well-known companies with a stronger customer base.

6. Limited product control- Because it functions as a marketplace, Amazon does not have total control over the goods that are offered for sale on its site. The corporation does not guarantee the availability or quality of any goods featured on its website, even if it does set certain criteria for third-party vendors. Regarding customer happiness and product safety, Amazon may face difficulties as a result of this lax product management. For instance, if buyers buy subpar or broken goods from independent sellers on Amazon, it may result in bad feedback and a decline in business. Furthermore, if goods sold on Amazon fail to adhere to safety regulations, it may result in harm to customers and legal ramifications for the corporation. It is imperative for Amazon to guarantee that all products offered on its marketplace adhere to quality and safety requirements.to maintain customer trust and satisfaction.

Amazon Opportunities

1. Expansion into emerging markets- One of Amazon's best opportunities to grow its clientele and revenue is to enter new areas. This can entail branching out into new product categories or consumer segments in addition to new geographic areas, both locally and abroad.

2. Expanding physical stores- By increasing the number of physical locations it has, Amazon can reach more people and offer a more immersive shopping experience. Despite having fewer physical locations than other retailers, the firm has been experimenting with other forms, like its Amazon Go stores, which provide a unique and convenient shopping experience. Increasing the number of physical locations it has

can also help Amazon outperform other shops and take a bigger chunk of the retail industry. In regions where e-commerce penetration is weaker, this might be especially advantageous because it enables Amazon to reach customers who might not feel comfortable making purchases online. To successfully expand its physical store presence, Amazon must carefully evaluate the potential risks and rewards of each market and develop strategies that are tailored to the unique needs and preferences of those customers. This may involve adapting its store formats, offering a wide range of products and services, and leveraging technology to enhance the customer experience.

3. Investment in new ventures- Investing in new projects gives Amazon the chance to expand and diversify its revenue sources. To increase their capabilities and reach, this may entail making investments in or purchasing businesses in related industries like cloud computing, advertising, or healthcare. By making new startup investments, Amazon can also leverage cutting-edge trends and technology, like the Internet of Things and artificial intelligence, to develop new goods and services. This might assist the business in staying one step ahead of the competition and retaining its top spot in the sector. Anticipated Debut into Crypto- Amazon intends to enter the cryptocurrency sector primarily by providing bitcoin payment methods for its goods and services. Customers would then be able to make purchases on the Amazon platform using cryptocurrencies like Bitcoin and Ethereum. By providing bitcoin payment choices, Amazon may be able to draw in new clients who are eager to use cryptocurrencies for online transactions. If more customers start using cryptocurrencies for online payments, it might also assist the business in gaining a bigger portion of the e-commerce market.

4. More acquisitions- Amazon can increase its capabilities, broaden its product range, and spur growth by acquiring adjacent businesses. This may entail purchasing businesses in adjacent sectors, including technology, logistics, or e-commerce, in order to capitalise on their advantages and develop synergies.Purchasing complementary businesses can also help Amazon expand into new areas and clientele, as well as hire fresh staff members with specialised knowledge. This is especially helpful in rapidly evolving industries where it's necessary to constantly innovate and adapt in order to stay ahead

of the competition.

Amazon Threat

1. Intense competition- Amazon's market share and profitability may be impacted by the fierce competition it encounters from both online and physical stores like eBay and Walmart. Rivals might be large, well-established businesses or little, startup businesses that provide comparable goods and services at affordable costs. With the goal to stay ahead of the competition and draw in new clients, Amazon needs to keep coming up with new ideas and improving its business plan.

2. Regulatory issues- Being that Amazon has been subject to regulatory scrutiny in a number of markets, the corporation may be in danger from regulatory difficulties. The corporation may have to pay fines or other penalties or modify its business procedures as a result, which could present risks and obstacles. Moreover, regulatory problems can harm a company's brand and hinder its capacity to function in specific areas.

3. Cybersecurity threats- Due to the handling of a significant quantity of consumer data, Amazon is susceptible to cybersecurity threats, which could pose dangers and present difficulties for the business. Cyberattacks have the ability to damage private client data, including financial and personal information, which could result in a loss of trust from customers and even legal repercussions. Amazon needs to make significant investments in strong security measures, as well as continuously monitor and update its systems to keep ahead of emerging cybersecurity risks.

4. Economic recessions- This poses a risk to Amazon since it may affect consumer spending, which might lower the company's sales and profits. Customers might be less inclined to buy non-essential items when the economy is uncertain, which might hurt Amazon's sales. Economic downturns can also result in more competition as businesses fight for market share, which might have an adverse effect on Amazon's profitability.

5. Natural disasters- Natural disasters like hurricanes or earthquakes can endanger Amazon's supply chain and operations. These catastrophes may impair the business's capacity to complete orders and harm its transportation systems, warehouses, and other

infrastructure. Delivery delays and unhappy customers may result from this, which might hurt Amazon's revenue and profitability.

In conclusion, Amazon is a company with a strong presence in the online retail market. Its strengths include a wide range of products, a strong brand reputation, and a large customer base. However, the company also faces several challenges, including intense competition, regulatory issues, and cybersecurity threats. Amazon's SWOT analysis highlights the company's position as a leader in the online retail market, but also points to the need for ongoing strategic planning to address its weaknesses and threats.

CHAPTER 3: RESEARCH METHODOLOGY

3.1 Research Problem / Statement of Problem

Understanding and assessing the Sales and Revenue patterns of electronic products on Amazon for the year 2022 is the primary goal of the study problem. By delving into the data, this research seeks to identify the important markets in terms of sales and revenue. The study aims to produce useful information that can support data-driven decision-making within the organisation by revealing these insights. By utilising these information, the business can increase sales and revenue by strategizing and making wise decisions. With a better grasp of Amazon's market trends, the business will be able to identify growth and expansion prospects. Furthermore, the study looks closely at the patterns of revenue and sales in order to find any trends or patterns that could be used to obtain an advantage over competitors. As a result, this thorough research will offer the information required to support the business in making decisions that will maximise its performance on Amazon. The business will be in a better position to customise its marketing plans and enhance its market placement by taking use of these insights. Ultimately, using a data-driven strategy is meant to optimise the company's earnings and sales.

3.2 Hypothesis

Given the diverse range of electronic products, the hypothesis may include assumptions such as

1. H0: There is no significant correlation between the sales and revenue of Amazon and states or months.
2. H1: Certain states and months exhibit a stronger correlation with increased sales and revenue.

3.3 Methods of Data Collection

1. Secondary Data- Specifically, we will leverage the huge Amazon sales data for 2022 in order to conduct a thorough study and obtain significant insights. The unique orderID, comprehensive product details including category, amount ordered, price, and purchase date, and geographic information like city and state are all included in this

23

data set. We can find trends, patterns, and correlations that can guide strategic decision-making and spur company growth by examining and evaluating this vast dataset. This extensive dataset presents an unusual opportunity to comprehend consumer preferences, pinpoint well-liked product categories, and even examine sales trends between various states and localities. Adding such extensive data will improve our findings' correctness and dependability significantly. As such, the analysis we conduct will yield valuable insights and facilitate the optimisation of multiple facets of our business operations, ranging from marketing strategies to inventory management.

3.4 Measurement and Scaling Technique

Accurate variable measurement depends heavily on quantitative metrics like pricing, the quantity of orders placed, and sales income. A variety of scaling techniques are used, frequently involving the use of ratio scales, to ensure accurate and thorough measurement. An accurate and instructive comprehension of the variables can be attained by using ratio scales, which allow the development of significant intervals between observations. This makes it easier to find trends, patterns, and connections in the data, which opens up important new information. Furthermore, the measurement and scaling technique's inclusion of ratio scales improves the obtained data's dependability and comparability across various contexts or time periods. As a result, researchers and decision-makers may safely evaluate and understand the data to make defensible judgements and precise findings. All things considered, using these scaling and measurement methods—with a focus on ratio scales in particular—ensures the validity and robustness of the data collected, allowing for a detailed investigation and comprehension of the variables in question.

3.5 Statistical Technique

1. Descriptive Statistics: In descriptive statistics, the key characteristics of a dataset are summed up and described using a variety of metrics, including the mean, median, and mode. These metrics give us crucial information about the data, enabling us to comprehend its distributions and fundamental characteristics more fully. For instance,

the mean provides information about the dataset's average value, whereas the median indicates the middle value regardless of whether the data is sorted in ascending or descending order. Conversely, the mode indicates the value that occurs the most frequently. We may thoroughly examine the dataset and make precise inferences about its properties by utilising these summary statistics. Descriptive statistics can also be used to identify outliers, or data points that dramatically depart from the rest.

2. Exploratory data analysis- As its name suggests, exploratory data analysis, or EDA, is a popular technique used by data specialists. Before using any other statistical analysis techniques, it is typically the initial stage in the data analysis process. EDA is not used in isolation for generalisation or prediction; rather, it provides a preview of the data and helps extract important insights from it. This approach focuses exclusively on looking for patterns in the data to identify possible connections. EDA can be used to investigate assumptions and hypotheses, look for missing data from acquired data and gain the most insights, and find unknown relationships within the data.

CHAPTER 4: THEORETICAL CONCEPTS

4.1 Review of Related Literature

1) Paper Name: An overview and comparison of free Python libraries for data mining and big data analysis.

Description:

Python is becoming more and more popular, particularly in the data science industry. As a result, there are more and more free libraries accessible for use. This review paper's objective is to outline and contrast the features of several Python modules for big data analysis and data mining. There isn't a paper that addresses the topic and lists the advantages and disadvantages of each of these libraries as of yet. Here they evaluate more than 20 libraries and separate them into six groups: core libraries, data preparation, data visualization, machine learning, deep learning and big data. In addition to a library's features, the size of the community and the number of volunteers who build and maintain the library are significant comparative points. Greater communities increase the likelihood of quickly solving a particular issue.

The paper presently recommends: pandas for data preparation; Matplotlib, seaborn or Plotly for data visualization; scikit-learn for machine leraning; TensorFlow, Keras and PyTorch for deep learning; and Hadoop Streaming and PySpark for big data.

2) Paper Name: Human Resource Analytics using Power Bi Visualization Tool.

Description:

Predictive analytics is one of the key components of human resource (HR) management, a large field of study that yields findings such as employee turnover, work performance, and training requirements analysis. Human resource management's primary goals are to assess employees' performance at work, evaluate their contribution to the products or services that the organization benefits from, and track their length of service. The primary goal of human resource analytics is to find highly qualified people who aspire to maximize return on investment for the company by taking into account a number of variables that aid in predictive analysis's ability to comprehend the individual.

For many firms, employee churn is regarded as a serious issue. It is one of the essential challenges to identify because it affects sustainability and also the organization's planning and boosting work culture harmony. As a result, every organization's human resources department works very hard and pays close attention to find the fundamental improvements. By recognizing this need, the research hopes to improve the capacity to use real-time data insights, such as dashboards that run machine learning models like Random Forest and Logistic Regression in the background, to identify staff turnover using POWER BI.

3) Book Name: Data Analysis and Visualization Using Python
Description:
This book takes a data science approach to Python and teaches the reader tried-and-true data visualization methods that are applied to important business choices. The book begins with an introduction to data science using Python, after which it covers the Python environment and introduces you to editors such as the Spyder IDE and Jupyter Notebooks.You will understand the basic Python programming methods utilized in data science after reading through an introduction to the language. Going on, you will discover how data visualization is essential to decision-making and meets the demands of contemporary business. Additionally, you will examine a few well-known Python data visualization frameworks.

This book takes a data science approach to Python and teaches the reader tried-and-true data visualization methods that are applied to important business choices. The book begins with an introduction to data science using Python, after which it covers the Python environment and introduces you to editors such as the Spyder IDE and Jupyter Notebooks.You will understand the basic Python programming methods utilized in data science after reading through an introduction to the language. Going on, you will discover how data visualization is essential to decision-making and meets the demands of contemporary business. Additionally, you will examine a few well-known Python data visualization frameworks.

4) Paper Name: Exploratory data analysis

Description: Any statistical study should, in most cases, start with a casual, exploratory review of the data; this is commonly referred to as exploratory data analysis (abbreviated EDA). Two primary goals—data description and model formulation—are outlined together with a discussion of the components of EDA. The argument is made that it is critical to view EDA as a fundamental component of statistical inference, and various instances are provided to illustrate how EDA is applied when developing models for subjects including queueing, forecasting, stock control, reliability, and regression. A few shortcomings of Tukey's (1977) EDA methodology are examined, and the initial inspection of data (IDA) is proposed as a substitute term.

5) Paper Name: Analysis of book sales prediction at Amazon marketplace in India: a machine learning approach

Description:
An essential component of supply chain management is forecasting customer demand since it helps to prevent overproduction or underproduction and shortens delivery times. Accurately predicting client demand in the context of e-commerce, which is often measured by sales volume, necessitates a comprehensive examination of a number of variables and their interconnections, including product type, country of purchase, price, discount rate, free delivery option, and sentiment of online reviews. This kind of prediction power is particularly crucial for online retailers like Amazon to maintain consumer satisfaction and effectively manage the supply chain. Using a variety of pertinent characteristics and their interactions as predictor variables, this study examines the effectiveness of many modeling techniques, including regression analysis, decision-tree analysis, and artificial neural networks, for forecasting book sales at amazon.in. The polarity of online reviews, which are used as predictors in these models, is measured by sentiment analysis. Based on the results of each model, the significance of each independent predictor variable—discount rate, review sentiment, etc.—is examined to identify the most important factors that marketers may alter to

affect sales. The artificial neural network model outperforms the decision-tree based model in terms of prediction accuracy. Furthermore, similar findings are produced by the regression analysis both with and without sentiment and interaction components. A comparative study of these models yields some important conclusions. First of all, the three models agree that the most significant and relevant predictor of book sales at amazon.in is the volume of reviews. Second, the average ratings, discount rate, and discount amount have little to no impact on sales forecasting. Thirdly, according to regression and decision-tree models, the reviews' positive and negative sentiments are both individually significant predictors; yet, according to neural network models, they are not significant at all.This observation from the neural network model is counter to the existent research which suggests that both negative and positive sentiment are significant with the former having higher influence in forecasting sales. Finally, the interaction effects of review volume with negative and positive sentiment are likewise found to be significant predictors as per all three models. Therefore, review volume, negative and positive sentiment, and their interactions are determined to be the most significant characteristics across all models tried to predict book sales. By making adjustments to these important parameters, online retailers can use the study's results to properly anticipate the amount of sales and successfully manage the supply chain.

4.2 Theoretical Background of the study

1. **Tools used**:

Business Intelligence tools and libraries like Numpy, Pandas, Python, Jupyter and Power BI are used to perform data analysis and building the dashboards.
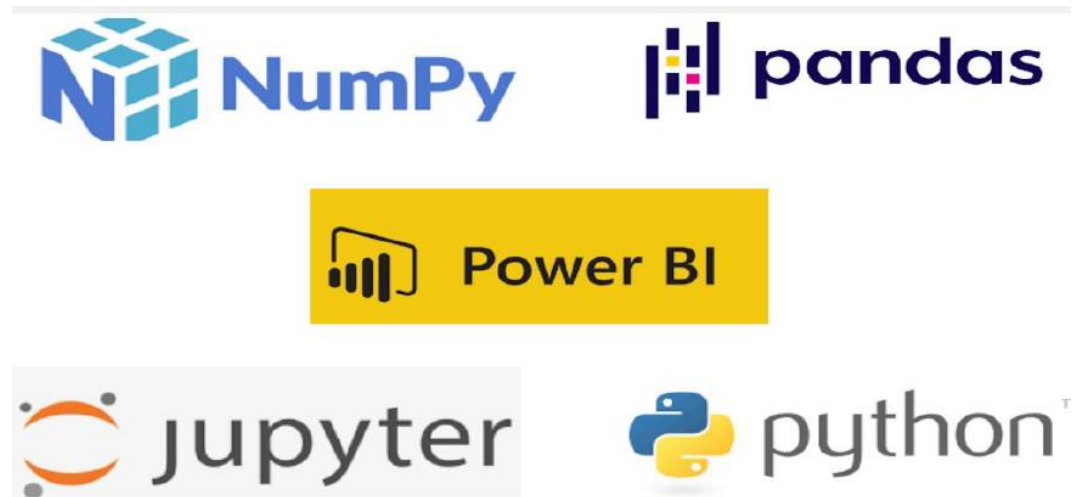
Figure 4.2.1- Tools Used in Business Intelligence

2. **Python** :

a) Definition: Python is a general-purpose, high-level programming language that is well-known for being readable and adaptable. Glob is a package for searching and obtaining file pathnames, while Pandas is a Python library for data processing and analysis.

b) Theories:

   i. Data Manipulation Theory: Pandas provides data structures such as DataFrames and Series for effective data management and analysis. It is built on the concepts of data manipulation.

   ii. File System Theory: Wildcard pattern matching is used by Glob to get file paths, which is based on file system theory.

c) Advantages/Disadvantages:

   i. Advantages:

      1. Versatility: Python's versatility allows it to be used for a wide range of tasks,

including scientific computing and web development.

      2. Open-Source: Python has an extensive and cooperative community due to its open-source nature.

  ii. Disadvantages:

      1. Execution Speed: Python's speed may be slower than that of lower-level languages when it comes to specific computational tasks.

d) Features:

  i. Pandas:

      1. DataFrame: Pandas presents the DataFrame, a two-dimensional table structure that makes it easier to manipulate and analyze data.

      2. Data Cleaning: Pandas provides strong tools for addressing missing values, rearranging datasets, and data cleansing.

  ii. Glob:

      1. File Path Matching: Glob offers an easy-to-use and efficient method of leveraging wildcard patterns to search for file pathnames.

      2. Directory Navigation: Makes it easier to navigate through file directories, which helps with data preparation and loading.

e) Applications/Examples:

  i. Data Cleaning and Preparation: Pandas is widely utilized in data science initiatives for the purpose of cleaning and preparing datasets for analysis.

  ii. File Management: Glob is utilized for jobs that include file management, such locating particular file kinds in directories.

  iii. Data Analysis and Visualization: Python is used for exploratory data analysis in conjunction with Pandas, and it can enhance Power BI by prepping and cleaning data prior to visualization.

f) Flow charts:
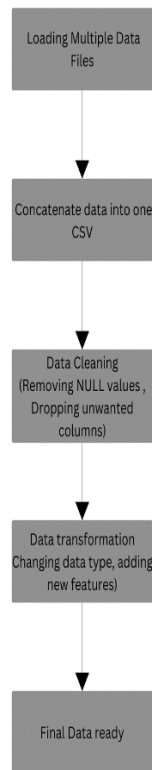
Figure 4.2.2- Workflow Using Python

2) **Power BI**:
   a) Definition: Microsoft offers a business analytics tool called Power BI that lets customers see and share insights from their data. It offers a collection of tools for data analysis, data visualization, and business intelligence. With Power BI, users can build interactive reports and dashboards, connect to multiple data sources, and modify data.
   b) Theories:
      i. Data Visualization Theories: Effective data visualization principles, which prioritize simplicity, clarity, and interactivity, are the foundation of Power BI. Theories that support precise and effective information visualization, such as Edward Tufte's principles of graphical excellence, are relevant.
      ii. Information Design Theories: Charts and dashboards that are both aesthetically pleasing and simple to understand are made possible by information design concepts like the Gestalt principles.

c) Advantages/Disadvantages:

    i. Advantages:

        1. User-Friendly Interface: Power BI's user-friendly interface makes it suitable for users with different levels of technical expertise.

        2. Microsoft Ecosystem: Data exchange and collaboration are improved by a smooth integration with other Microsoft products like Excel and Azure.

        3. Interactive Dashboards: Facilitates the development of dynamic and interactive dashboards that provide a deeper comprehension of data.

    ii. Disadvantages:

        1. Costs: Although there is a free version of Power BI, some sophisticated capabilities could need a subscription, which could be an issue for smaller companies.

        2. Learning Curve: Users who are unfamiliar with business intelligence tools may find advanced functions to have a learning curve.

d) Features:

    a. Data Connectors: Power BI offers an extensive selection of data connectors that make it easier to integrate with different data sources.

    b. Data Transformation: Robust capabilities for data cleaning, transformation, and shaping through a Power Query Editor.

    c. DAX (Data Analysis Expressions): One potent formula language for making unique computations and aggregations in Power BI reports is called DAX (Data Analysis Expressions).

e) Applications/Examples:

    i. Sales and Marketing Analytics: Power BI is widely utilized in sales and marketing analytics to analyze sales data, spot trends, and improve marketing tactics.

    ii. Financial Reporting: It's used for forecasting, budgeting, and analysis of financial data.

    iii. Operational Analytics: By visualizing key performance indicators (KPIs), Power BI is used to monitor and optimize operational processes.

f) PowerBI Architecture:

STEP 1	STEP 2	STEP 3	STEP 4	STEP 5
Data from source systems is integrated and loaded into a data warehouse of other analytics repository.	Data sets are organized into analytics data models or OLAP cubes to prepare them for anlysis.	BI analysts, other analytics professionals and business users run analytical queries against the date.	The query results are built into data visualizations, dashboards, reports and online portals.	Busness executives and workers use the information for decision-making and strategic planning.

Figure 4.2.3– Power BI Architecture

CHAPTER 5: DATA ANALYSIS AND INTERPRETATION

5.1 INTRODUCTION

The dataset being analysed relates to Amazon's 2022 sales statistics for electronic devices. It includes vital information that is necessary to comprehend the dynamics of revenue and sales, support strategic decision-making, and optimise corporate operations.

The following crucial variables are present in the dataset::

   a. Order ID: An individual number linked to every order that facilitates transaction tracking and management.

   b. Product: The brand name of the electronic device that was ordered; this can include a wide range of products, including computers, headphones, phones, batteries, chargers, and monitors.

   c. Quantity Ordered: Indicates client demand by listing the quantity of a certain product ordered in a single transaction.

   d. Price Each: The dollar amount for each individual item, enabling the computation of the total amount of money made per transaction.

   e. Order Date: The date the order was placed, which makes it possible to analyse sales seasonality and temporal trends.

   f. City and State: The geographic data that describes the place where the order came from and serves as a foundation for a regional study..

   g. Category: The category to which the product belongs, offering a means to segment and analyze sales data based on product types.

Importance of Analysing Sales and Revenue of Electronic items: There are a number of reasons why it is crucial to examine the sales and revenue patterns of electronic items.

1. Strategic Decision-Making: Businesses can better align their operations with market trends and demand by having a thorough understanding of sales dynamics. Revenue Optimisation: Companies can optimise their revenue streams and concentrate efforts on high-return possibilities by finding top-performing products and markets.

2. Market Identification: By identifying key markets through the analysis of sales data, companies are able to customise their tactics to target particular geographic areas and

customer preferences.

3. Product Performance Analysis: Improving overall product portfolio management, inventory management, and marketing initiatives are guided by the evaluation of individual product performance.

4. Competitive Advantage: Businesses can get a competitive edge by using insightful sales data analysis to outperform rivals and react quickly to changes in the market.

df

	Order ID	Product	Quantity Ordered	Price Each	Order Date	Purchase Address
0	176558	USB-C Charging Cable	2	11.95	2022-04-19 08:46:00	917 1st St, Dallas, TX 75001
1	NaN	NaN	NaN	NaN	NaN	NaN
2	176559	Bose SoundSport Headphones	1	99.99	2022-04-07 22:30:00	682 Chestnut St, Boston, MA 02215
3	176560	Google Phone	1	600	2022-04-12 14:38:00	669 Spruce St, Los Angeles, CA 90001
4	176560	Wired Headphones	1	11.99	2022-04-12 14:38:00	669 Spruce St, Los Angeles, CA 90001
...	...	...	...	...	...	...
186845	259353	AAA Batteries (4-pack)	3	2.99	2022-09-17 20:56:00	840 Highland St, Los Angeles, CA 90001
186846	259354	iPhone	1	700	2022-09-01 16:00:00	216 Dogwood St, San Francisco, CA 94016
186847	259355	iPhone	1	700	2022-09-23 07:39:00	220 12th St, San Francisco, CA 94016
186848	259356	34in Ultrawide Monitor	1	379.99	2022-09-19 17:30:00	511 Forest St, San Francisco, CA 94016
186849	259357	USB-C Charging Cable	1	11.95	2022-09-30 00:18:00	250 Meadow St, San Francisco, CA 94016

186850 rows × 6 columns

Figure 5.1.1- DataFrame of the dataset created using python

Synopsis of Python and Power BI:

Two potent technologies used in this investigation for thorough data exploration are Power BI and Python:

1. Power BI Power BI is a business analytics solution offered by Microsoft that is popular due to its easy-to-use interface and powerful visualisation features. It makes it easier for consumers to understand complicated datasets by enabling them to construct interactive dashboards and reports. To find patterns and trends in the sales data, Power BI can create dynamic charts, graphs, and geographic maps.

2. Python (Pandas and Glob): Using the Pandas package, Python is a flexible programming language used for data preparation and cleaning. Pandas is a great option for cleaning and shaping datasets since it provides effective data manipulation tools. Furthermore, the Python Glob module assists in file path manipulation, contributing to the seamless

handling of datasets.

5.2 Data Cleaning and Preparation (Python - Pandas)

Data pre-processing and transformation must be done before any model is built in order to feed the right data into the model for analysis and visualisation. The procedure consists of:

1. Managing Missing and Null Values: The dataset was cleansed of the few rows that had only null values. For a number of reasons, removing null values—also referred to as handling missing data—is an essential stage in the data cleansing process.

    a. Avoiding Analytical Errors: Machine learning systems, statistical analysis, and other data-driven procedures may be hampered by null values. Their existence could provide skewed or false results, which would undermine the validity of your conclusions.

    b. Ensuring Accuracy in Descriptive Statistics: Null values have a substantial impact on descriptive statistics, including mean, median, and standard deviation. By deleting them, you may be sure that the central tendency and variability of the available data are appropriately represented in these summary statistics.

    c. Enhancing Visualisation: When null values are present, data visualisations like charts and graphs may not render correctly or may even be deceptive. Eliminating these values contributes to the creation of data visualisations that are more precise and educational.

```
In [7]: # In the dataframe there are some rows which are completely null so dropping them
        Final_data = df.dropna()

In [8]: DESC(Final_data)

        Shape of data : (185950, 6)
        --------------------------------------------------
        <class 'pandas.core.frame.DataFrame'>
        Int64Index: 185950 entries, 0 to 186849
        Data columns (total 6 columns):
         #   Column           Non-Null Count    Dtype
        ---  ------           --------------    -----
         0   Order ID         185950 non-null   object
         1   Product          185950 non-null   object
         2   Quantity Ordered 185950 non-null   object
         3   Price Each       185950 non-null   object
         4   Order Date       185950 non-null   object
         5   Purchase Address 185950 non-null   object
        dtypes: object(6)
        memory usage: 9.9+ MB
        None
        --------------------------------------------------
        Count of null values in columns :
        Order ID          0
        Product           0
        Quantity Ordered  0
        Price Each        0
        Order Date        0
        Purchase Address  0
        dtype: int64
```

Figure 5.2.1- Handling missing values

2. Data Transformation: Some columns that were required for data analysis and dashboard creation had to be extracted from the raw dataset.

    a. After extracting the columns, the Purchase Address column was removed, and the City and State columns were obtained from the raw dataset.

    b. A Category Column was made based on the product name, indicating the type of product (e.g., laptop, phone, TV, etc.).

```
In [20]: def product_category(dataframe, column_name, product_name_ends_with, category_name):
             index_list=dataframe[dataframe[column_name].str.endswith(product_name_ends_with)==True].index
             dataframe.loc[index_list,"Category"]= category_name
             print(f"'Category' column created Successfully for Category '{category_name}'")
```

```
In [21]: product_category(Final_data, "Product", "Charging Cable", "Charging Cable")

         'Category' column created Successfully for Category 'Charging Cable'
```

```
In [22]: product_category(Final_data, "Product", "Headphones", "Headphones")

         'Category' column created Successfully for Category 'Headphones'
```

```
In [23]: product_category(Final_data, "Product", "Phone", "Phone")

         'Category' column created Successfully for Category 'Phone'
```

```
In [24]: product_category(Final_data, "Product", "Laptop", "Laptop")

         'Category' column created Successfully for Category 'Laptop'
```

```
In [25]: product_category(Final_data, "Product", "Monitor", "Monitor")

         'Category' column created Successfully for Category 'Monitor'
```

Figure 5.2.2- Data Transformation

5.3 Exploratory Data Analysis (Power BI)

The first step in the analytical process is called exploratory data analysis (EDA), which helps us identify any underlying patterns, correlations, or abnormalities in the information. We can take advantage of the advantages of both Power BI and Python for this assignment, allowing us to fully comprehend the Amazon sales data for electronic devices in 2022.

1. Power BI for Preliminary Analysis and Visualisation

    a) Overview of the Data: Load and view the dataset using Power BI. Examine the distribution of variables such as Price Each and Quantity Ordered to spot any possible outliers.

    b) Bivariate Analysis: To determine the associations between variables, use correlation matrices and scatter plots. Examine any connections between

Price Each and Quantity Ordered to spot any possible patterns.

c) Time-Series Analysis: Leverage Power BI's time intelligence features to analyze sales and revenue trends over the course of the year. Create line charts and heatmaps to visualize monthly and quarterly variations.


```
1 Dates = DISTINCT(SALES_2022[Order Date])
```

Order Date	Month	Month_fullname	Weekday	Week_name	Month_name
01-01-2022	1	January	7	Sat	Jan
01-01-2022	1	January	7	Sat	Jan
01-01-2022	1	January	7	Sat	Jan
01-01-2022	1	January	7	Sat	Jan
02-01-2022	1	January	1	Sun	Jan
02-01-2022	1	January	1	Sun	Jan
02-01-2022	1	January	1	Sun	Jan
02-01-2022	1	January	1	Sun	Jan
02-01-2022	1	January	1	Sun	Jan
02-01-2022	1	January	1	Sun	Jan
02-01-2022	1	January	1	Sun	Jan
02-01-2022	1	January	1	Sun	Jan
02-01-2022	1	January	1	Sun	Jan
03-01-2022	1	January	2	Mon	Jan
03-01-2022	1	January	2	Mon	Jan
03-01-2022	1	January	2	Mon	Jan
03-01-2022	1	January	2	Mon	Jan
03-01-2022	1	January	2	Mon	Jan
03-01-2022	1	January	2	Mon	Jan
04-01-2022	1	January	3	Tue	Jan
04-01-2022	1	January	3	Tue	Jan
04-01-2022	1	January	3	Tue	Jan
04-01-2022	1	January	3	Tue	Jan
04-01-2022	1	January	3	Tue	Jan
05-01-2022	1	January	4	Wed	Jan
05-01-2022	1	January	4	Wed	Jan

Figure 5.3.1- Time Series analysis

5.4 Market Identification (Power BI)

A thorough analysis of sales and revenue data was conducted as part of the market identification process in order to identify regional preferences and geographic patterns for various electronic product categories on Amazon in 2022. To graphically depict and analyse the data, a number of geographical charts and tables were created using Power BI's features. Significant regional differences in sales and revenue were found by the investigation.
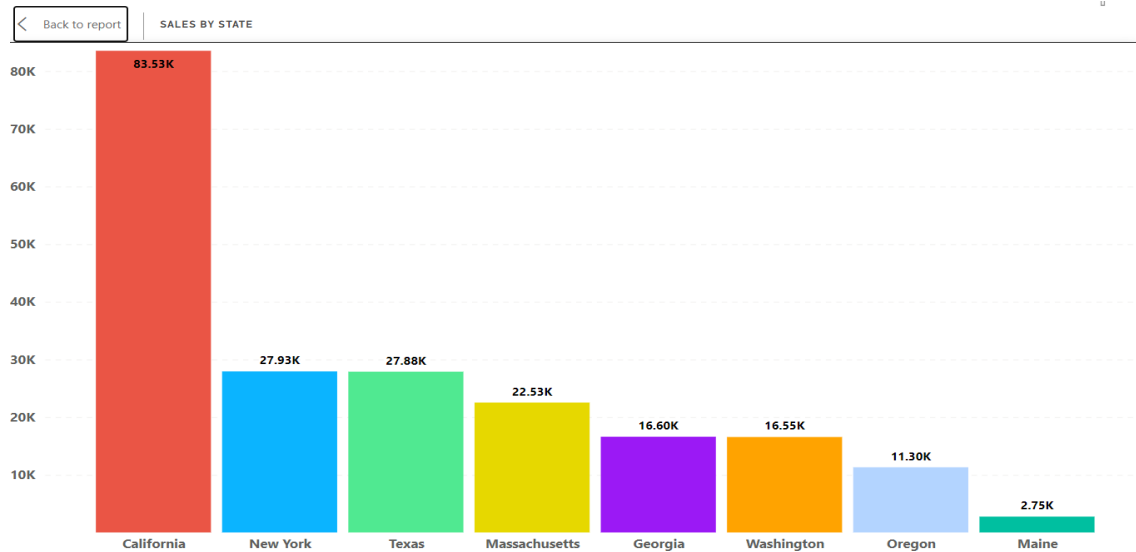
Figure 5.4.1- Sales By State

1. According to the chart, Massachusetts, New York, Texas, and California have the lowest sales. Maine, Oregon, and Washington had the lowest sales.

2. The chart breaks down sales by state. With 83,530 sales, California leads the nation, followed by New York with 27,930 and Texas with 27,880. With sales ranging from 22,530 to 2,750, Massachusetts, Georgia, Washington, Oregon, and Maine had lower numbers.

3. Furthermore, the population of each state is not displayed in the statistics. It's probable that California has higher sales since it has a population that is significantly larger than those of the other states. It's challenging to compare sales data on an apples-to-apples basis without knowing each state's population.
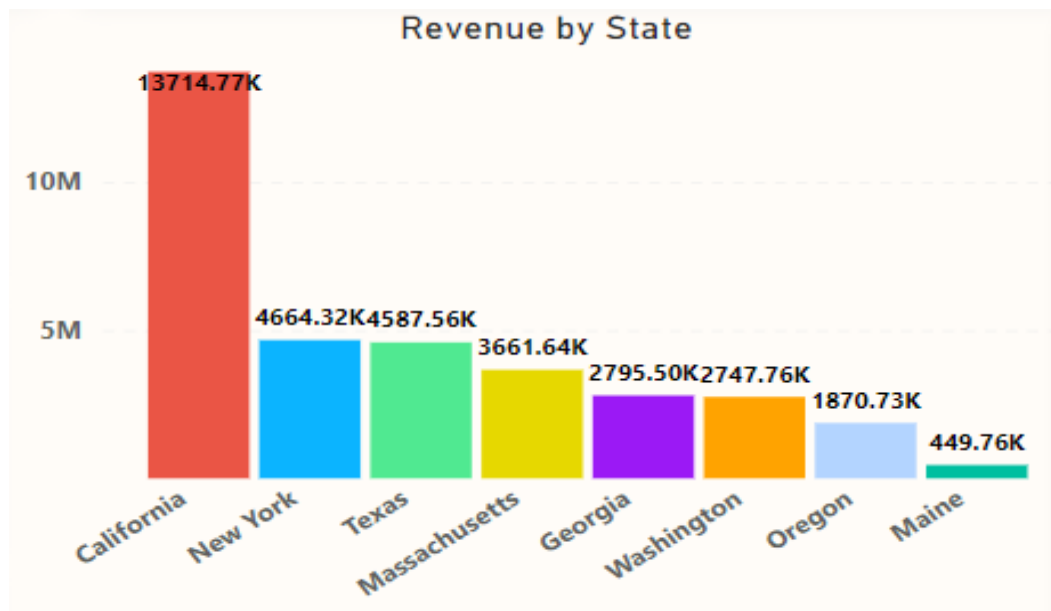
Figure 5.4.2- Revenue By State

The revenue for Amazon's Electronics category for the year 2022 is displayed in a bar chart that breaks down the revenue by state. The eight states with the greatest sales numbers are listed on the x-axis, while the revenue in thousands of dollars is displayed on the y-axis.

These are the chart's main conclusions to take into consideration:

1. With projected sales of $137,147.77K for electronics on Amazon in 2022, California is the state with the highest sales. With $46,643.20K in revenue, New York came in second place, less than three times as much as this state.

2. The top four states Texas, Massachusetts, New York, and California collected $232,365.53K in revenue, or 45.19% of the total. This implies that a limited number of states account for a disproportionate share of Amazon's electronics sales.

3. After the top four states, revenue declines significantly. Georgia, the state in fifth place, brought in less than half as much money as Massachusetts, with only $36,616.40K.

4. The eight states represented on the chart brought in $514,320.76K in total income. It is crucial to remember that this does not account for all of Amazon's US electronics sales. There may be additional revenue from other states as the graphic only displays the top eight states.

5.5 Product Performance Analysis (Power BI)

Using Power BI and Python, a thorough analysis was carried out to look at how each product performed inside the Amazon sales dataset for the year 2022. Finding the best-selling products and comprehending how much of the total money they earned was the main goal.

1. Power BI Analysis: Utilising Power BI, informative visualisations were produced to enable a rapid and simple assessment of product performance. The sales amounts and revenue produced by each product category were displayed using Matrix Table, providing a clear comparison of their respective contributions. Users were able to delve down into specific product details through interactive exploration made possible by Power BI's dynamic features. The figure below shows that Batteries had the largest sales and the best performance in terms of sales. Headphones were next in queue, followed by charging cords. demonstrating that consumers need batteries and charging wires more than they do for laptops and smartphones.

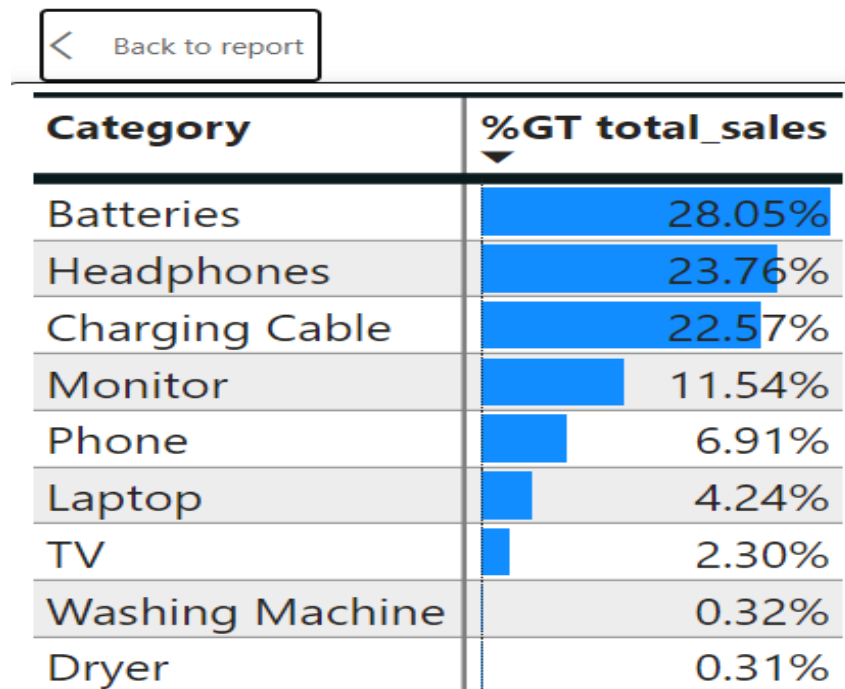| Category	%GT total_sales
Batteries	28.05%
Headphones	23.76%
Charging Cable	22.57%
Monitor	11.54%
Phone	6.91%
Laptop	4.24%
TV	2.30%
Washing Machine	0.32%
Dryer	0.31%

Figure 5.5.1- Percentage of Grand Total Sales by Category

The graph displays the percentage of sales for various product categories on Amazon that

are related to electronics in 2022. These are the chart's main conclusions to remember:

1. Batteries (28.05%), Headphones (23.76%), and Charging Cables (22.57%) are the top three product categories. In 2022, these categories account for 74.38% of all electronics sales on Amazon

2. The next two best-selling categories are phones (6.91%) and monitors (11.54%). Despite being far less well-liked than the top three, these categories nevertheless make up a sizeable share of sales.

3. The remaining categories consist of laptops (4.24%), TVs (2.30%), washing machines (0.32%), and dryers (0.31%). These categories only make up a tiny percentage of sales and are far less common than the others.

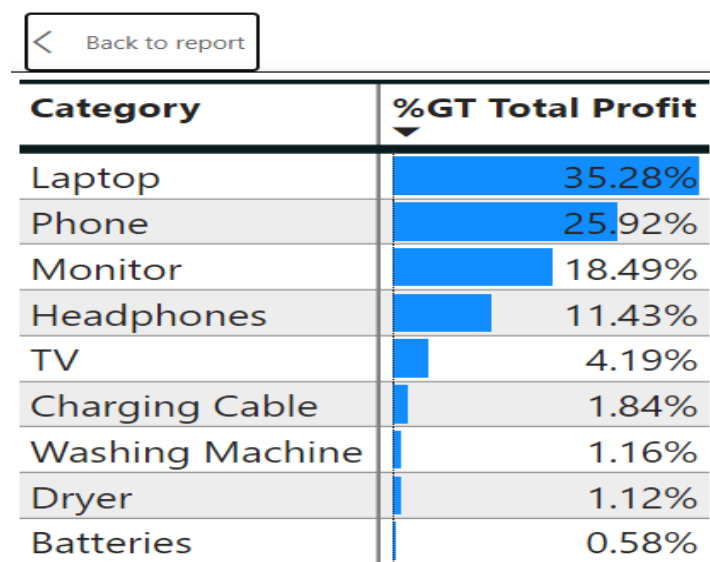| Category	%GT Total Profit
Laptop	35.28%
Phone	25.92%
Monitor	18.49%
Headphones	11.43%
TV	4.19%
Charging Cable	1.84%
Washing Machine	1.16%
Dryer	1.12%
Batteries	0.58%

Figure 5.5.2- Percentage of Grand Total Revenue by Category

The chart displays, for each product category on Amazon, the percentage of total profit made from electronics in 2022. These are the chart's main conclusions to remember:

1. With 35.28% of the total earnings, laptops are the most profitable category. This shows that even while they are the most profitable category for Amazon, they may not be the most popular category in terms of unit sales (as we observed in the earlier graphic you gave).

2. With a profit percentage of 25.92%, phones come in second. In terms of units sold,

phones were the second-highest selling category in the prior chart, which is consistent with the current one.

3. Closely following, monitors take home 18.49% of the overall profit. Considering that they represented only 11.54% of unit sales in the prior chart, this is unexpected. This shows that Amazon may have a big profit margin on displays.

4. The remaining categories are made up of charging cables (1.84%), TVs (4.19%), and headphones (11.43%). Even though they account for a lesser percentage of the overall earnings, these groups nevertheless make a substantial contribution.
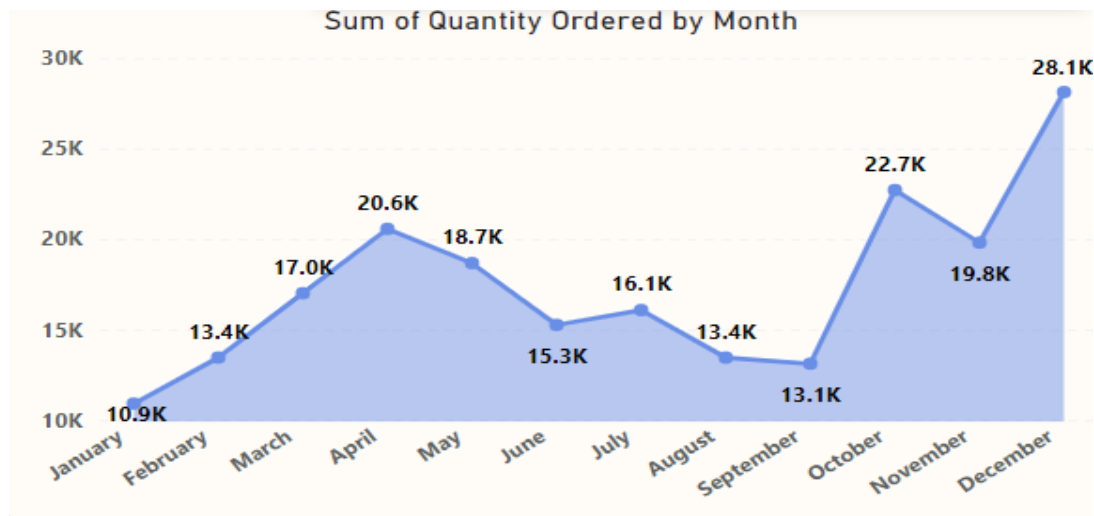
5.6 Temporal Dynamics

Figure 5.6.1- Sum of Quantity Ordered by Month

The line graph displays the total amount of electronic device orders placed on Amazon by month for the year 2022. These are the chart's main conclusions to remember:

1. In general, Amazon saw a rise in electronic goods sales in 2022. The total quantity ordered peaked in December at approximately 30,000, having begun at about 10,000 in January. This indicates a 200% rise in the course of the year.

2. In 2022, there were two distinct sales peaks. Around Mother's Day and Prime Day in May, there was the first peak. During the Christmas shopping season in December, there was a second peak. These peaks imply that holidays and promotional activities

are important factors in Amazon's electrical device sales.

3. From June to September, sales remained largely consistent. This implies that sales of electronic devices might be less seasonal in the summertime.
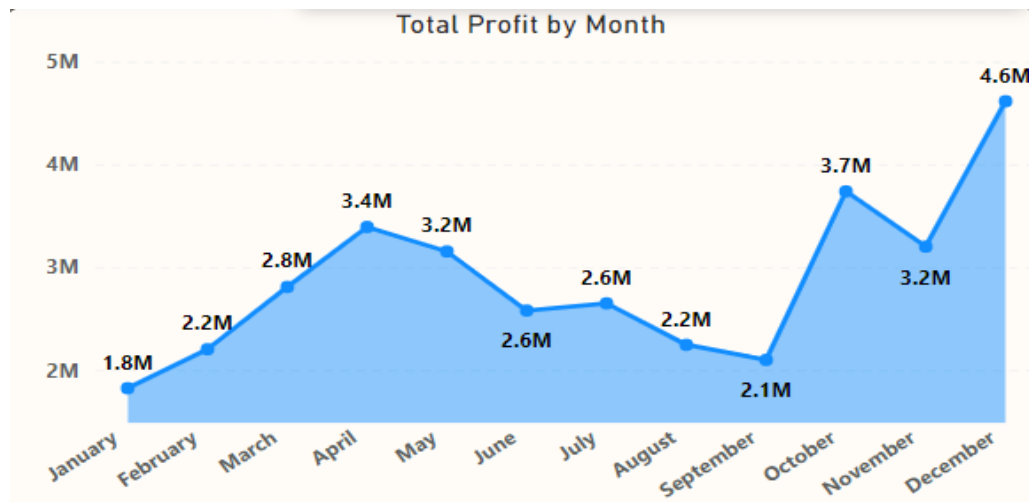

Figure 5.6.2- Total Profit by Month

The line graph displays the overall profit for Amazon's electrical products in 2022 broken down by month. These are the chart's main conclusions to remember:

1. Over the course of 2022, Amazon's profit from electronic devices climbed steadily, peaking in December at approximately $4.6 million, from a starting point of approximately $2.4 million in January. Over the course of the year, this is a 92% gain.

2. In 2022, there were two distinct profit peaks that lined up with the sales peaks shown in the prior graphic you supplied me. Around Mother's Day and Prime Day in May, there was the first peak. During the Christmas shopping season in December, there was a second peak. These peaks imply that vacations and promotional activities are important sources of revenue.

3. From June to September, profit grew more slowly than it did over the rest of the year. This implies that, in relation to sales volume, there might be less seasonality in the profitability of electronic devices over the summer.

5.7 Customer Behavior Analysis (Power BI)

This section focuses on using Power BI to study consumer behaviour. It emphasises analysing revenue by state, revenue by city, and total profit by month. These salient features shed light on regional patterns and temporal fluctuations in consumer purchases, offering insightful information for tactical decision-making.

1. Revenue/Sale by State: The mapping features of Power BI are used to show how sales and revenue are distributed among the various states. We show how each state contributes to the total revenue with bar charts. Finding high-performing areas and adjusting marketing or distribution tactics in accordance with those findings are made possible by this analysis. Furthermore, differences in revenue between states can point up areas that could need more attention or opportunities for expansion.

2. Revenue/Sale by City: By going even closer, we use Power BI to investigate sales and revenue at the city level. To illustrate the revenue from sales of electronic products in different cities, bar charts are utilised. Granular insights are made possible by identifying the best-performing cities. These insights enable localised customer preferences to be understood, marketing efforts to be targeted, and inventory optimisation. This approach also helps identify geographical differences in consumer behaviour and purchasing power.

3. Total Profit by Month: Turning our attention to temporal dynamics, Power BI is used to examine the monthly total profit and sales. The yearly variations in sales and profitability are depicted by time-series line charts. Comprehending monthly fluctuations in profitability is crucial for modifying tactics in accordance with seasonal patterns, refining inventory levels, and scheduling promotional campaigns during periods of high demand. Decisions about the distribution of resources and sales projections are informed by this temporal analysis, which also makes pattern recognition easier.
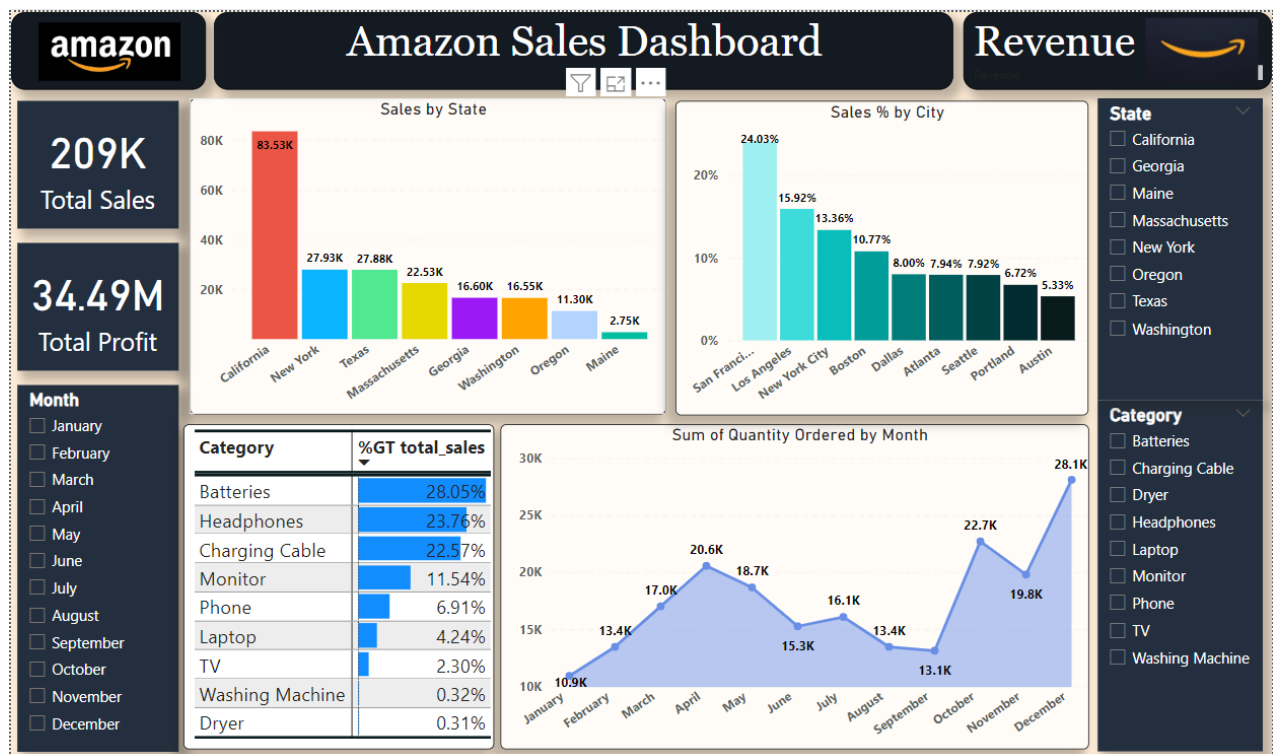
Figure 5.7.1- Sales Analysis

Through the consolidation of these essential elements into Power BI dashboards, we furnish decision-makers with a pictorial account of consumer behaviour patterns. Because Power BI is interactive, users can drill down into particular states, cities, or months, which encourages dynamic data research. This method not only makes it easier to comprehend consumer behaviour in its whole, but it also gives companies the ability to create focused strategies that take into account local and seasonal variations.
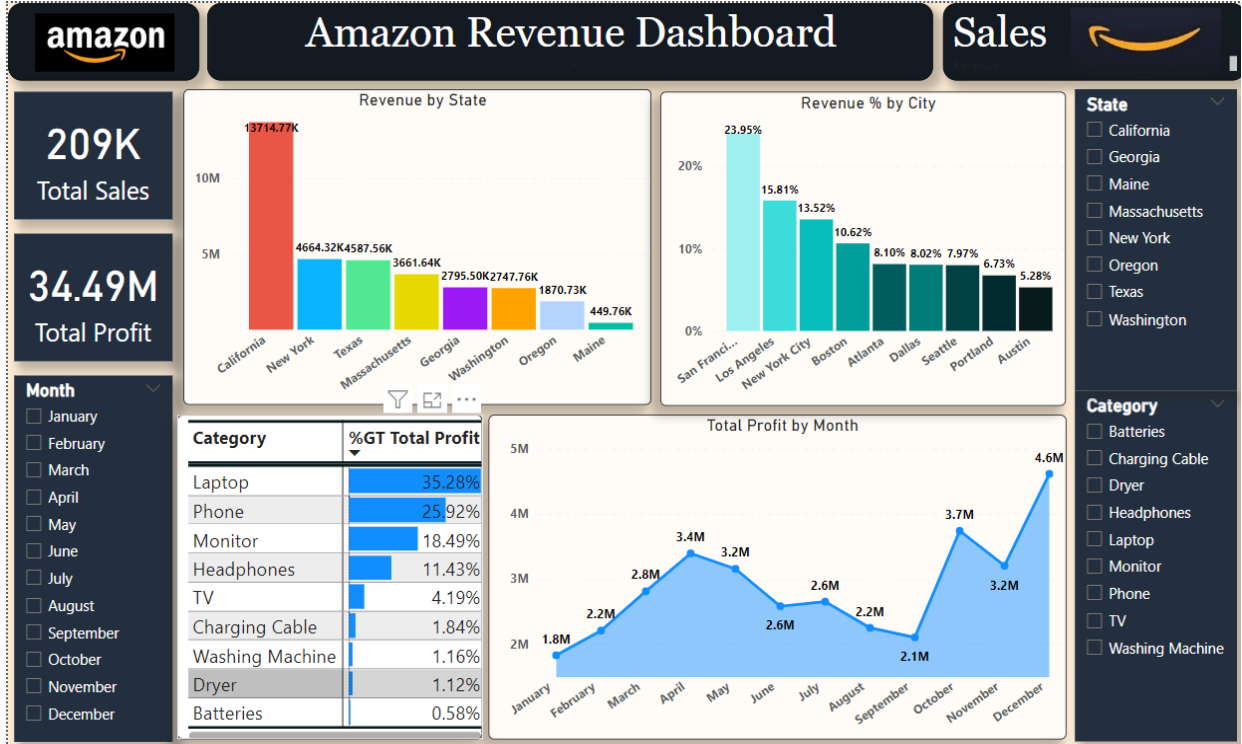
Figure 5.7.2- Revenue Analysis

5.8 Limitations and Future Directions

1. Limitations: Several restrictions were found when analysing Amazon sales data for electronic devices in 2022. First off, a lot of the dataset's completeness and correctness depend on the data that Amazon provides; therefore, any errors or missing data may have an effect on the validity of the findings. Furthermore, the dataset is exclusive to Amazon, which restricts the applicability of the results to the market for all electronic devices. Furthermore, by concentrating just on 2022, the temporal limitation may miss long-term patterns or take into account outside variables that could affect sales. An further constraint pertains to the lack of client demographic information, which impedes a more intricate comprehension of consumer behaviour and inclinations. Finally, the study makes the assumption that the dataset it has been given is an exhaustive list of all electronic devices that are offered on Amazon, which may mean that it ignores niche markets or exclusive product categories.

2. Future Directions: A number of directions for more study and analysis are suggested in order to address the shortcomings found in this analysis. First off, extending the dataset to cover more years may make it easier to spot trends over a longer time frame and offer a more comprehensive picture of the dynamics of the electrical products market. Furthermore, a comparative research using datasets from different e-commerce platforms would provide a more comprehensive picture of the sector, improving the findings' generalizability. Marketing strategies and product development could be more focused with the collection and incorporation of client demographic data. Investigating sentiment analysis in customer reviews may reveal information about preferences and levels of satisfaction with the products. In the future, research may also collaborate with producers to gain more particular product specifications, which would provide a deeper examination of performance aspects. Finally, incorporating outside variables into the analysis—like economic data or technical developments—may provide a deeper comprehension of the dynamics influencing the market for electrical items. In the sphere of electronic product sales, these suggested directions seek to address existing shortcomings and open the door for more thorough and nuanced evaluations.

CHAPTER 6: FINDINGS

4.1 KEY FINDINGS

Dashboards will be used to show and highlight key performance indicators (KPIs) and pertinent indicators for the trends in sales and revenue. Dashboards will be incorporated to show charts over time with progress on various indicators or aspects as and when the system begins to record the historical/periodic data for a user.

Key performance indicators, or KPIs, are measurements that show a summary of sales and revenue along with how they relate to other measures.

For Sales :

- Sales by State
- Sales % by City
- Category-wise Product's Sold (in %)
- Monthly trend of Sales

For Revenue :

- Revenue generation per State
- Distribution of Revenue (in %) by Cities
- Product's Revenue(in %) in various Product Categories
- Month-wise Revenue trend

1. **Top-Performing Products**:
   a) **Power BI Analysis:**
      i. Sales Quantity:
         - The product category with the biggest sales quantity is batteries, which also happens to be the best-performing one.
         - Charging cables and headphones are next in line, suggesting a steady demand for these products.
      ii. Revenue Generation:
         - The primary source of revenue is clearly laptops, which are

more expensive than other products.

- Phones and displays also make a substantial contribution to total revenue.

b) Python Analysis:

Statistical Metrics:

- Batteries are the best-selling goods, according to a thorough statistical analysis done with Python's Pandas module.

- Laptops and phones rank first and second in terms of overall income contributions..

2. Geographical Insights:

a. Power BI Analysis:

    i.   Dominant Market:

- It is determined that North America is the leading market, accounting for a significant portion of total sales.

- Within North America, particular subregional nuances draw attention to differences in consumer preferences and economic reasons.

3. Regional Product Preferences:

a) Power BI Analysis:

    i.   Product Category Distribution:

- The analysis uses bar charts to show how product popularity varies by area.

- Preferences for items such as phones, headphones, batteries, chargers, laptops, and monitors vary between different areas.

4. Temporal Dynamics:

a. Power BI Analysis:

    i.   Monthly and Quarterly Variations:

- Time-series analysis shows annual variations in income and sales.

- Peaks and troughs are found, suggesting the presence of

seasonality and temporal trends.

5. **Customer Behavior Insights:**

    a. **Power BI Analysis:**

        i. Revenue/Sale by State:

            - States contribute significantly to overall revenue, according to geographical mapping; •

            - There are clear regional differences in consumer behaviour and purchasing power.

        ii. Revenue/Sale by City:

            - Recognising high-performing cities through city-level analysis offers valuable insights for targeted marketing campaigns.

            - It is crucial to comprehend regional differences in tastes and buying habits.

        iii. Total Profit by Month:

            - Monthly profit variances are displayed through temporal analysis, which helps strategies be adjusted to seasonal patterns.

            - Based on these temporal insights, resource allocation and sales forecasting can be optimised..

CHAPTER 7 SUGGESTIONS / RECOMMENDATIONS

1.1 **SUGGESTIONS/ RECOMANDATIONS**

1. Product Portfolio Management:
    a. Recommendation: Regularly assess product performance metrics, and optimize the product portfolio by considering factors such as sales quantity, revenue, and profitability.
    b. Implementation Steps:
        i. Conduct periodic reviews of individual product performance.
        ii. Consider introducing new products based on emerging trends.
        iii. Evaluate the feasibility of discontinuing products with consistently low performance.

2. Geographically Targeted Strategies:
    a. Recommendation: Tailor marketing and distribution strategies to align with the unique characteristics and demands of high-performing regions.
    b. Implementation Steps:
        i. Invest in targeted advertising campaigns for top-performing regions.
        ii. Optimize inventory distribution based on regional demand patterns.
        iii. Collaborate with local influencers or partners for regional marketing efforts.

3. Dynamic Pricing:
    a. Recommendation: Implement dynamic pricing strategies based on regional and temporal trends to maximize revenue.
    b. Implementation Steps:
        i. Analyze historical data to identify pricing trends.
        ii. Utilize dynamic pricing tools to adjust prices based on demand fluctuations.
        iii. Monitor competitor pricing strategies for benchmarking.

4. Long-Term Trend Analysis:
    a. Recommendation: Expand the dataset to include multiple years for a more

comprehensive analysis of long-term trends.

    b. Implementation Steps:

        i. Acquire sales data for additional years.

        ii. Analyze trends and patterns over an extended period.

        iii. Identify evolving consumer preferences and market dynamics.

5. Comparative Analysis:

    a. Recommendation: Conduct comparative analyses with sales data from other e-commerce platforms to gain insights into industry-wide trends.

    b. Implementation Steps:

        i. Acquire relevant sales data from other e-commerce platforms.

        ii. Identify commonalities and differences in product preferences.

        iii. Benchmark performance against industry standards.

6. Demographic Data Collection:

    a. Recommendation: Collect and incorporate customer demographic data to tailor marketing strategies.

    b. Implementation Steps:

        i. Include demographic information in customer registration processes.

        ii. Analyze purchasing behavior based on demographics.

        iii. Use demographic insights to personalize marketing content.

7. Sentiment Analysis:

    a. Recommendation: Explore sentiment analysis on customer reviews to understand product satisfaction and preferences.

    b. Implementation Steps:

        i. Utilize natural language processing tools for sentiment analysis.

        ii. Categorize reviews based on sentiment (positive, negative, neutral).

        iii. Extract insights into factors influencing customer satisfaction.

CHAPTER 8 CONCLUSIONS

8.1 Conclusions

In summary, the potent combination of Python and Power BI has completely changed how we analyse Amazon's electronic device sales environment. We have learned a great deal about the underlying patterns and trends that underpin the success of Amazon's sales of electronic products by fusing these two cutting-edge technologies. Businesses may now take decisive action to achieve continuous growth and preserve competitiveness in the constantly changing e-commerce environment thanks to the complete recommendations that have been supplied. These recommendations include portfolio optimisation, focused tactics, and dynamic pricing.

Furthermore, this thorough analysis establishes the foundation for further research in addition to highlighting present trends. This comprehensive strategy guarantees that firms are well-prepared to manage the complexity of the market, capitalise on emerging possibilities, and continuously improve customer satisfaction by resolving constraints and going deeper into the dynamics of electronic product sales.

Businesses may access a plethora of information and get a better grasp of their electrical product sales dynamics by combining the capabilities of Python and capabilities BI. With this fresh knowledge, Amazon and other e-commerce behemoths can strategically place themselves in the market, allowing them to adjust to shifting consumer tastes, stay one step ahead of the competition, and ultimately prosper in the digital era.

REFERENCES

1. https://www.aboutamazon.in/about-us/awards-recognition

2. https://researchci.com/amazon-key-success-factors/

3. https://en.wikipedia.org/wiki/List_of_Amazon_products_and_services

4. https://www.geeksforgeeks.org/amazon-company-overview-history-products/

5. https://www.sellerapp.com/blog/best-selling-products-on-amazon/

6. https://businessmodelanalyst.com/amazon-swot-analysis/

7. https://www.analyticssteps.com/blogs/7-types-statistical-analysis-definition-explanation

8. https://www.simplilearn.com/tutorials/data-analytics-tutorial/data-analytics-with-python

9. https://radacad.com/power-bi-architecture-guidelines