**Ganadipathy Tulsi's Jain Engineering College**

Kaniyambadi - Vellore - 632102

**IBM – PROJECT**

**PROJECT TITLE: PRODUCT SALES ANALYIS**

Phase 5(Final Submission)

**TEAM ID:** 223855

**TEAM MEMBERS:**

Surya G

Gokul Nath K A B

Gowtham R

Dhana Lakshmi G

## 

**INDEX**

**1. INTRODUCTION**

1.1 Problem Definition

1.2 Project Objective

**2. DESIGN THINKING PROCESS**

2.1Analysis Objectives

2.2 Data Collection

2.3 Visualization Strategies

2.4 Actionable Insights

**3. DEVELOPMENT PHRASES**

**4. ANALYSIS OBJECTIVES**

**5. DATA COLLECTION PROCESS**

**6. DATA VISUALIZATION USING IBM COGNOS**

**7. DERIVED ACTIONABLE INSIGHTS**

**8. INSIGHTS FROM THE ANALYSIS GUIDES**

8.1 Inventory Management

8.2 Marketing Strategies

**9. FUTURE SCOPE**

**10. ADVANTAGES**

**11.DISADVANTAGES**

**12.CONCLUSION**

**INTRODUCTION**

Product sales analysis is a critical aspect of business operations, as it helps organizations understand their product performance, customer preferences, and market trends. Analyzing product sales data provides valuable insights for decision-making and can lead to improved sales strategies, inventory management, and profitability.

**PROBLEM DEFINITION:**

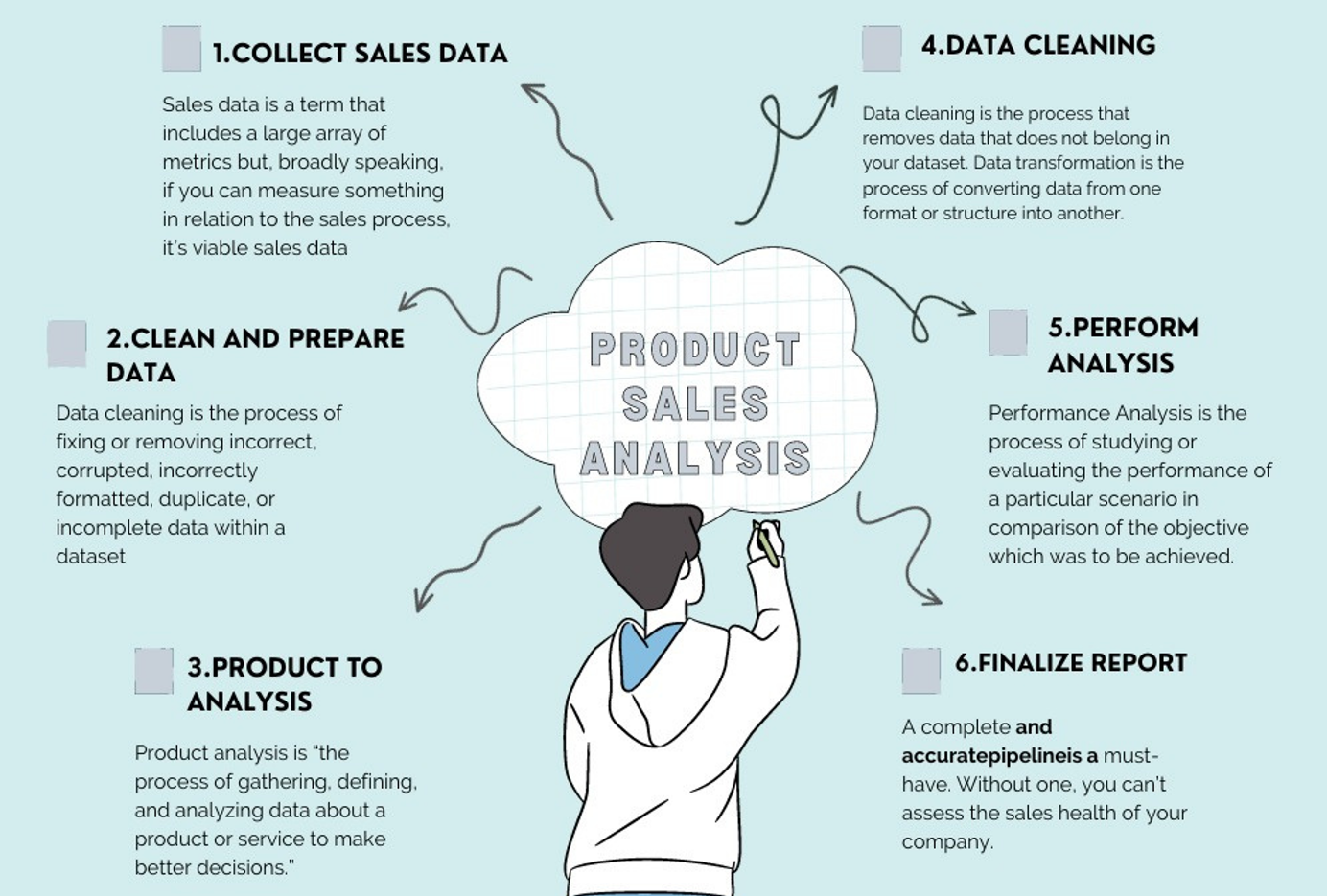
To analyse and improve product sales performance, this project aims to gather, process, and interpret sales data to gain insights into key metrics such as revenue, profit margins, product popularity, and customer behaviour. The objective is to identify trends, opportunities, and areas for optimization in order to enhance overall product sales and profitability.

The primary goal is to understand how well each product is performing in terms of sales, identify factors influencing sales, and develop strategies to increase revenue and profitability.

**PROJECT OBJECTIVE:**

The product sales analysis project aims to comprehensively assess and analyze the sales performance of a company's product over a specific period. This analysis will involve collecting, organizing, and interpreting sales data to provide actionable insights and recommendations for optimizing product sales strategies, identifying top-performing products, and addressing areas of improvement.

Ultimately, the project's goal is to enhance revenue generation and maximize profitability through data-driven decision-making. The Product Sales Analysis project is aimed at gaining valuable insights into a company's sales performance by examining and analyzing the sales data for its products. This analysis will help the organization make informed decisions, optimize sales strategies, and identify opportunities for growth.



**DESIGN THINKING**

Design thinking is a problem-solving approach that emphasizes empathy, creativity, and iterative prototyping to develop innovative solutions. When applied to product sales analysis, it can help you gain a deep understanding of your customers and identify opportunities for improving your sales processes. Here's a step-by-step guide on how to apply design thinking to product sales analysis:

**Analysis Objectives** - Define the specific insights you want to extract from the sale data, such as identifying top-selling products, analyzing sales trends, and understanding customer preferences.

**Data Collection** - Determine the sources and methods for collecting sales data, including transaction records, product information, and customer demographics.

**Visualization Strategy** - Plan how to visualize the insights using IBM cognos to create interactive dashboards and reports.

**Actionable Insights** - Identify how the derived insights can guide inventory management and marketing strategies.

**DEVELOPMENT PHRASES**

**HOW MACHINE LEARNING ALGORITHM CAN HELP YOU PREDICT SALES TRENDS AND CUSTOMER BEHAVIOUR:**

Machine learning algorithms require quality data for training and ongoing monitoring to remain accurate and relevant. It's important to continuously refine models and adapt them to changing market conditions and customer preferences. Additionally, the choice of the right algorithm depends on the specific problem and available data. A combination of various machine learning techniques and domain knowledge is often necessary to gain valuable insights into sales trends and customer behavior.

* The future of product sales analysis lies in the synergy between human expertise and machine learning algorithms. By combining the intuition and experience of sales professionals with the analytical capabilities of machine learning, businesses can unlock new opportunities, identify hidden patterns, and adapt to the ever-evolving market dynamics.
* This innovation document outlines a proposal to enhance the existing product sales analysis project by implementing innovative features to provide deeper insights and improve decision-making.

**DEFINE THE PROBLEM:**

* Before you start collecting data, you have to determine the question you want to answer or what problem you want to solve.

**CLEAN AND PREPARE DATA:**

* You need to collect the relevant data and clean it so it is in the appropriate format for analysis.

**NORMALIZES THE DATA:**

* Normalizes ensure the data is in the consistent model which allow for the analysis the performance.

**ANALYSIS OBJECTIVES**

**TYPES OF MACHINE LEARNING ALGORITHM FOR PREDICT SALES AND TRENDS:**

**1.SUPERVISED LEARNING ALGORITHM:**

* + Train model using labelled data to prediction or classification.

**2.UNSUPERVISED LEARNING ALGORITHM:**

* + Discover patterns and structure in unlabelled data without predefined categories.

**3.REINFORCEMENT LEARNING ALGORITHM:**

* + Learn through interaction with an environment to maximize reward and minimize penalties.

**IMPLEMENTING MACHINE LEARNING ALGORITHMS:**

**1)CHOOSE THE RIGHT ALGORITHM:**

* Based on your specific needs, choose an algorithm that can efficiently analysis.

**2)TRAIN THE MODEL:**

* Feed the algorithm to your data to train the model and analysis the predictions.

**3)MAKE PREDICTION:**

* finally use your trained data to predict the future for sales trends.

**EVALUATING AND REFINING PREDICTION:**

**CONFUSION MATRIX:**

* A table that helps to determine the accuracy of model predictions.

**CROSS VALIDATIONS:**

* A technique for testing a model efficiently divides the data to testing and training.

**FEATURE IMPORTANCE:**

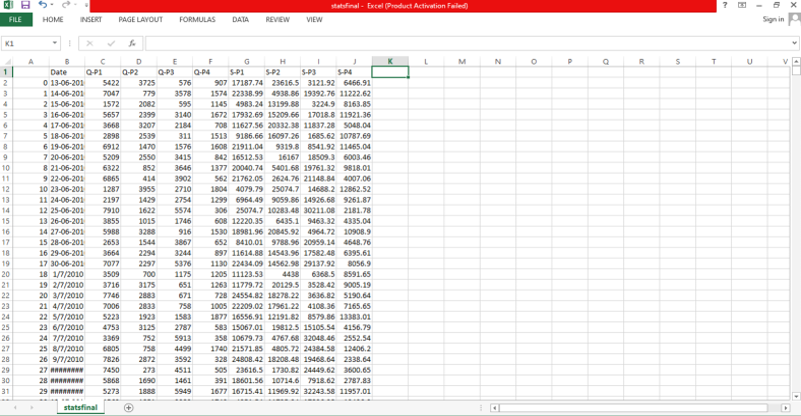
* A feature importance that determines the feature in which data is most important for data set.

**DATA COLLECTION**

Data collection for product sales analysis in data analytics involves gathering and storing relevant information and statistics about product sales. It is a crucial step in the data analysis process that provides the raw material necessary for making informed decisions and deriving insights about your products and sales performance.

**GIVEN DATASET:**

* The dataset for the above mentioned project was obtained from kaggle website.
* Kaggle is an popular dataset providing source where obtained datasets are with high quality and less errors.



**NECESSARY STEPS TO FOLLOW:**

**1.IMPORT THE NECESSARY LIBRARIES:**

* Start by importing the required libraries.

**Program:**

import pandas as pd

import numpy as np

**2. LOAD THE DATASET:**

* Read the dataset from a file, such as a CSV file, into a Pandas Dataframe. Replace ‘statsfinal.csv' with the actual filename of your dataset.

**Program:**

data= pd.read\_csv(‘'C:/Users/asus/Downloads/statsfinal.csv’)

**3.EXPLORE THE DATA:**

* Before preprocessing, it's essential to get an overview of the dataset.

**Program:**

# Display the first few rows of the dataset

print(data.head())

# Get basic information about the dataset

print(data.info())

# Check for missing values

print(data.isnull().sum())

# Summary statistics

print(data.describe())

**4.FEATURE ENGINEERING:**

* Create new features or modify existing ones if needed for your analysis. This step depends on the specifics of your analysis.

**Program:**

data['year'] = pd.to\_datetime(data['date']).dt.year

data['month'] = pd.to\_datetime(data['date']).dt.month

**5.DATA TRANSFORMATION:**

* Transform data to make it suitable for analysis. You might need to encode categorical variables, normalize or scale numerical features, and perform other transformations.

**IMPORTANCE OF LOADING AND PROCESSING DATASET:**

* Loading and preprocessing the dataset is an important first step in building any machine learning model. However, it is especially important for product sales analysis models, as product price datasets are often complex and noisy.
* By loading and preprocessing the dataset, we can ensure that the machine learning algorithm is able to learn from the data effectively and accurately.

**CHALLENGES INVOLVED IN LOADING AND PREPROCESSING A PRODUCT SALES DATASET:**

**DATA VOLUME:**

* Large datasets can be challenging to load and process efficiently, as they may not fit into memory. You may need to implement data loading strategies like data streaming or batch processing.

**MISSING DATA:**

* Sales datasets often contain missing values, which must be dealt with appropriately. Imputing missing values or deciding how to handle records with missing data is an important preprocessing step.

**DATA INTEGRATION:**

* Your sales data comes from multiple sources or systems, integrating the data into a single coherent dataset can be challenging. You'll need to establish data mappings and standardize the data formats.

**DATA VALIDATION:**

* Ensuring data accuracy and consistency is crucial. You may need to implement data validation checks to identify and handle data that doesn't meet predefined criteria.

**HOW TO OVERCOME THE CHALLENGES OF LOADING AND PREPROCESSING A PRODUCT SALES ANALYSIS DATASET:**

There are a number of things that can be done to overcome the challenges of loading and preprocessing a product sales dataset, including:

**Use a data preprocessing library:**

* There are a number of libraries available that can help with data preprocessing tasks, such as handling missing values, encoding categorical variables, and scaling the features.

**Carefully consider the specific needs of your model:**

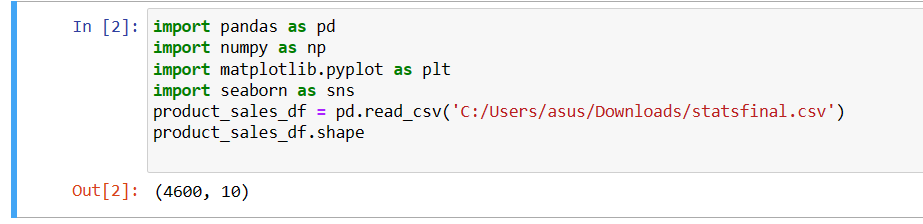
* The best way to preprocess the data will depend on the specific machine learning algorithm that you are using. It is important to carefully consider the requirements of the algorithm and to preprocess the data in a way that is compatible with the algorithm.

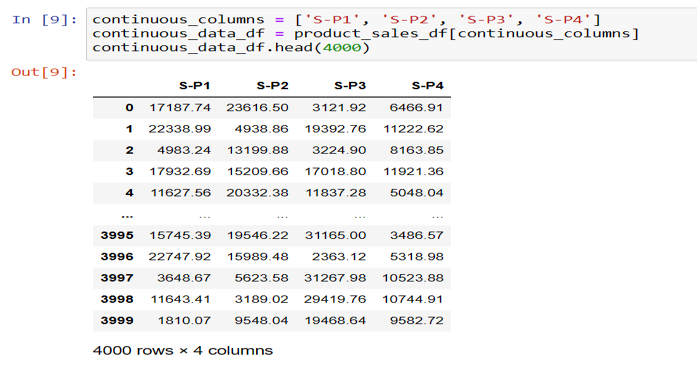
**1.LOADING THE DATASET:**

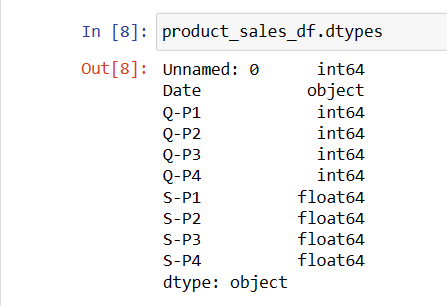
* Loading the dataset using machine learning is the process of bringing the data into the machine learning environment so that it can be used to train and evaluate a model.
* The specific steps involved in loading the dataset will vary depending on the machine learning library or framework that is being used.

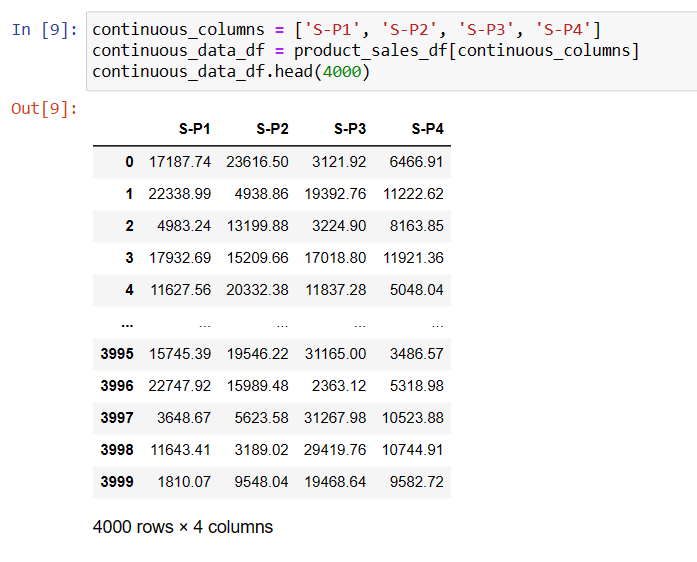
However, there are some general steps that are common to most machine learning frameworks:

1. **Identify the dataset -** The first step is to identify the dataset that you want to load. This dataset may be stored in a local file, in a database, or in a cloud storage service.
2. **Load the dataset -** Once you have identified the dataset, you need to load it into the machine learning environment. This may involve using a built-in function in the machine learning library, or it may involve writing your own code.
3. **Preprocess the dataset -** Once the dataset is loaded into the machine learning environment, you may need to preprocess it before you can start training and evaluating your model. This may involve cleaning the data, transforming the data into a suitable format, and splitting the data into training and test sets.









**2.Preprocessing the dataset:**

* Data preprocessing is the process of cleaning, transforming, and integrating data in order to make it ready for analysis.
* This may involve removing errors and inconsistencies, handling missing values, transforming the data into a consistent format, and scaling the data to a suitable range.

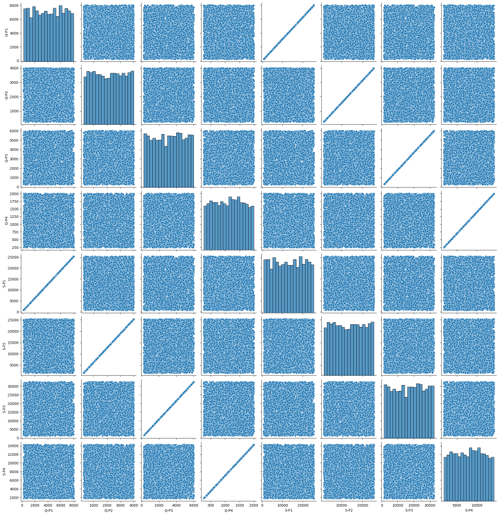
**DATA VISUALIZATION**

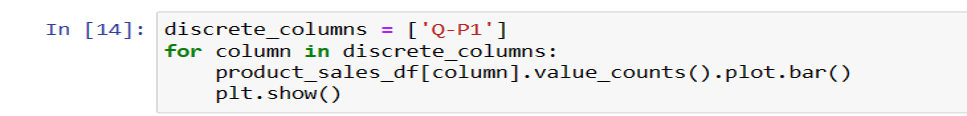
**VISUALISATION AND PRE-PROCESSING OF DATA:**

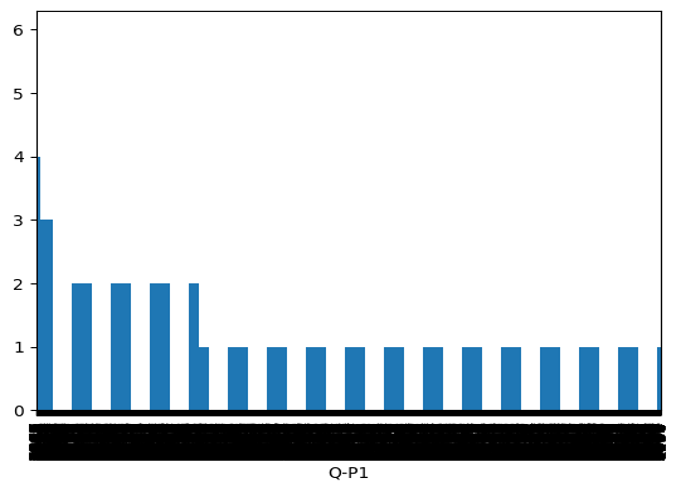
Visualizing and preprocessing data are crucial steps in product sales analysis. Let's start with data visualization using Python and the Matplotlib and Seaborn libraries, and then we'll discuss data preprocessing

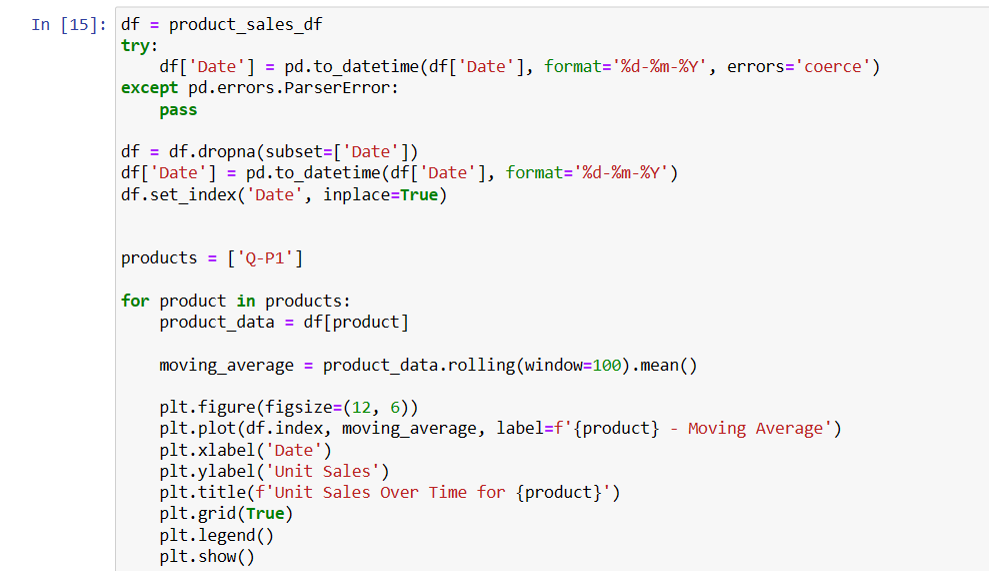
**VISUALIZE CONTINUOUS NUMERIC VARIABLES:**

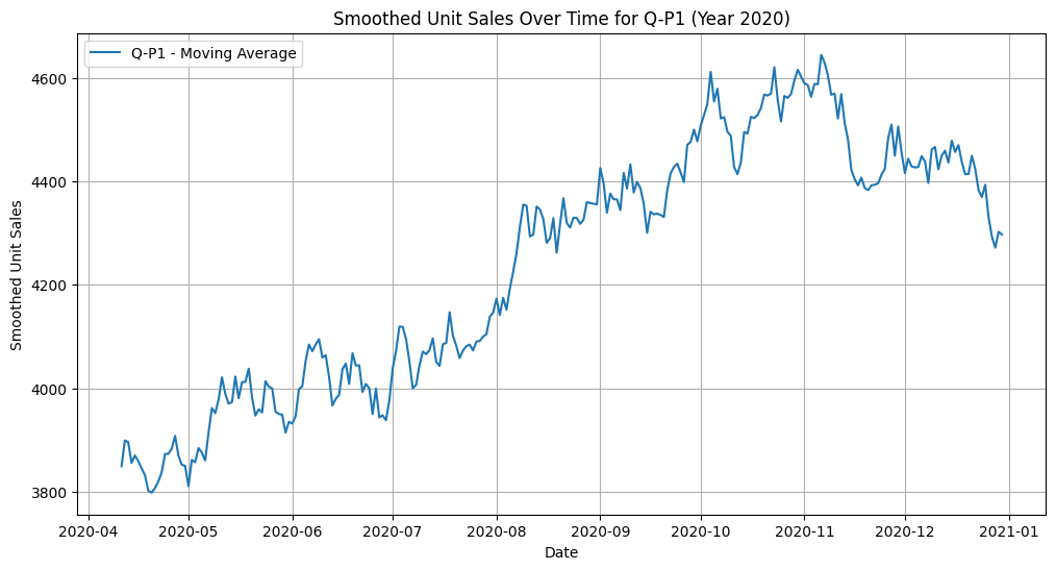


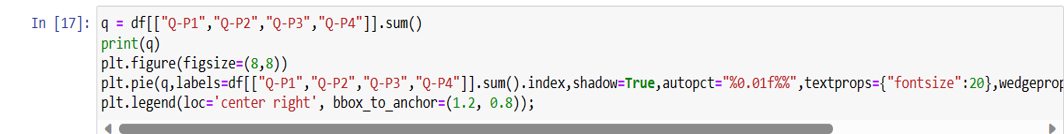


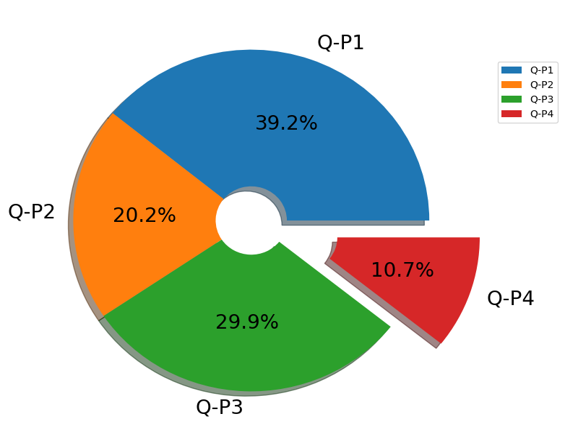








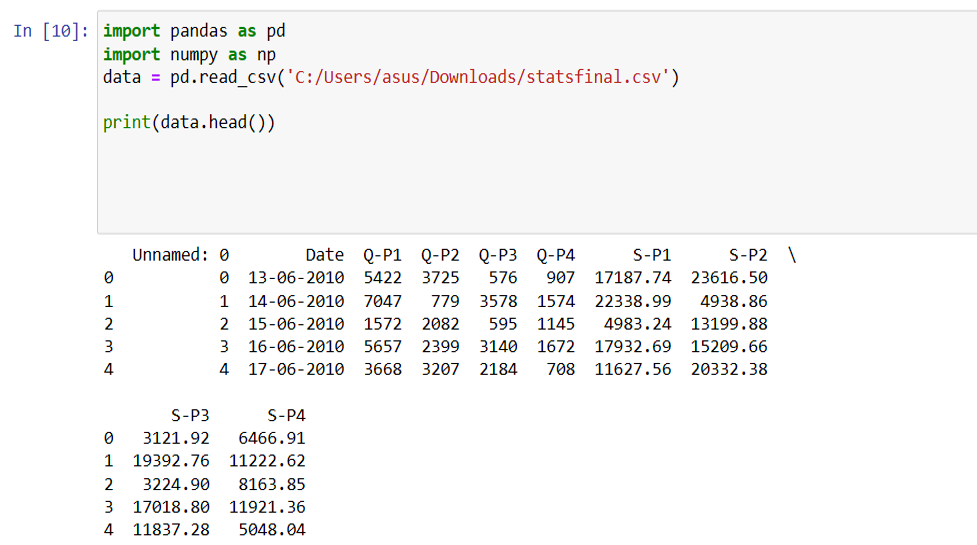


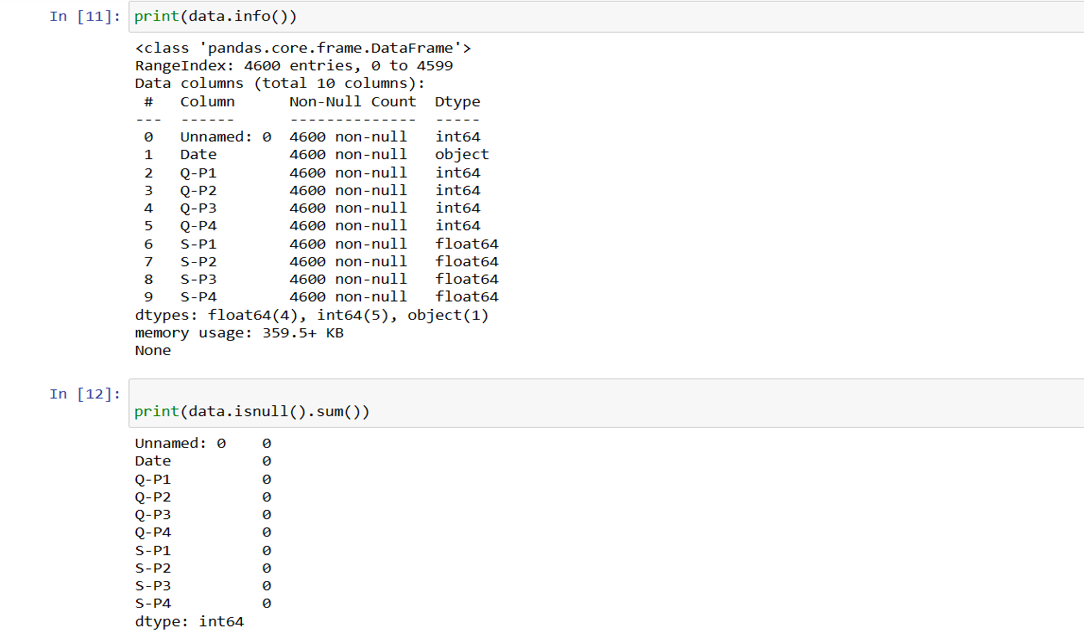


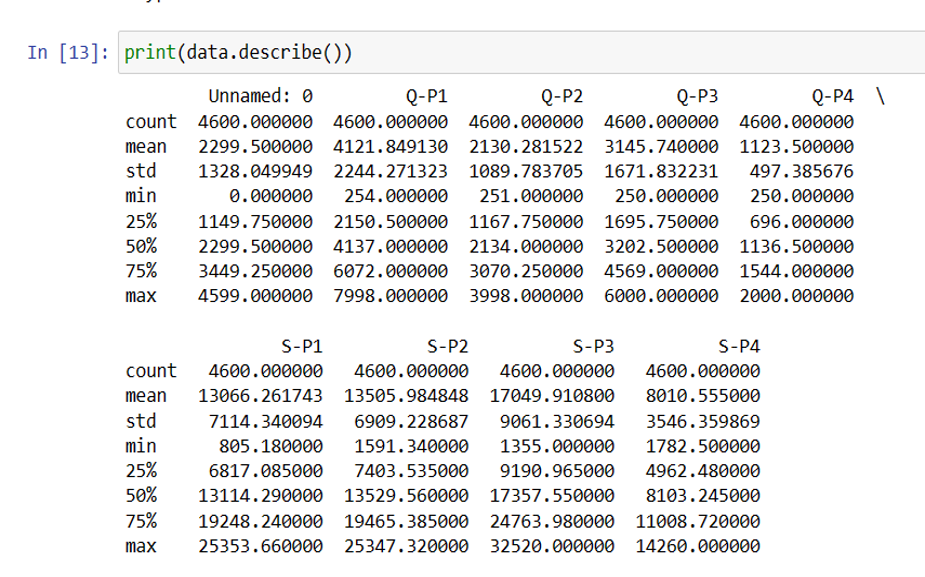
**SOME COMMON DATA PREPROCESSING TASKS INCLUDE:**

* **DATA CLEANING -** This involves identifying and correcting errors and inconsistencies in the data. For example, this may involve removing duplicate records, correcting typos, and filling in missing values.
* **DATA TRANSFORMATION -** This involves converting the data into a format that is suitable for the analysis task. Example, this may involve converting categorical data to numerical data, or scaling the data to a suitable range.
* **FEATURE ENGINEERING -** This involves creating new features from the existing data. For example, this may involve creating features that represent interactions between variables, or features that represent summary statistics of the data.









* Conducting a product sales analysis is a critical component of any business strategy. It provides valuable insights into the performance of products, customer behavior, and market trends.
* Successful product sales analysis often requires cross-functional collaboration among data analysts, domain experts, and decision-makers. Continuous skill development is essential to stay updated with the latest tools and techniques in the field.
* product sales analysis is a multifaceted process that empowers businesses to make informed decisions, optimize their operations, and achieve greater profitability. By leveraging data, visualizations, and insights, companies can adapt to market dynamics and stay competitive in a constantly evolving business environment.

**DATA VISUALIZATION USING IBM COGNOS**

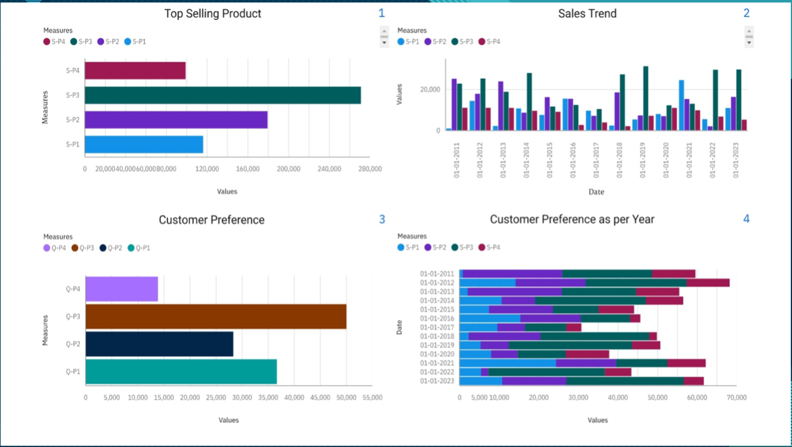
**IBM-COGNOS:**

* Cognos is a business intelligence performance management tools for IBM.
* It allows technical and non-technical employees in any company to analyse, extract and create interactive dashboards that enable the company to take relevant decisions.
* The Cognos tool combines a multitude of products which enables communication with different third parties.
* to visualization and reporting the Dataset, We use IBM Cognos.

**VISUALIZATION AND REPORTING:**

* To perform the three analyze and report
  + - top-selling products
    - customer preferences
    - sales trends
* And we design interactive dashboards
* Dashboards are a type of data visualization, and often use common visualization tools such as graphs, charts, and tables.

**THE FIND OUTS USING IBM COGNOS:**

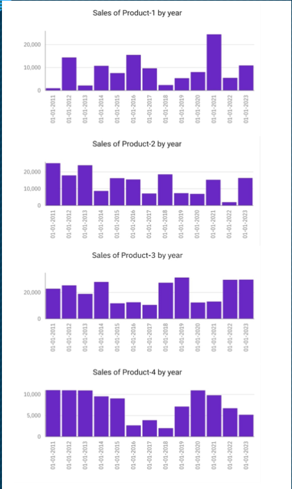
* Top Selling Product
* Sales Trend
* Customer Preference
* Now in the dashboard we have three analysis top selling products , sales trend, customer preferences This visualization are done by using bar chart and filtering is applied in the all analysis. In filtering we using date column for filtering all the sales and quantities by Date.

In the Dashboard we Conclude,

* + Top selling product : S-P3 has highest selling product by Date
  + Sales Trend- S-P3 has the maximum sales in all year
  + Customer preference:Q-P3 has high Quantity and S-P3 has highest sales

**IDENTIFYING PRODUCTS WITH THE HIGHEST SALES:**

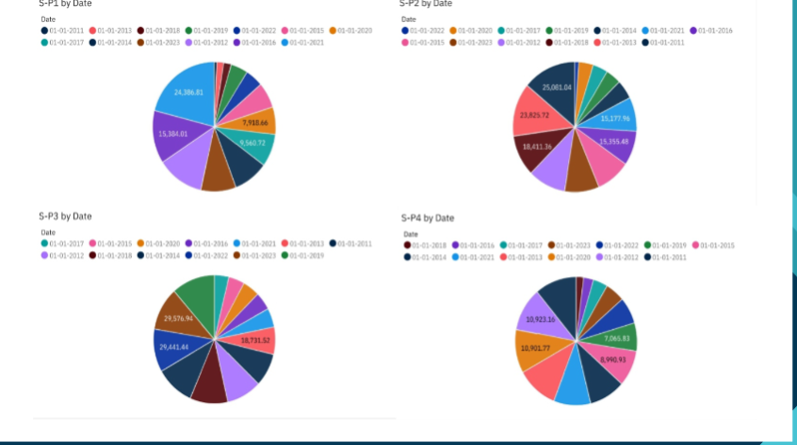
The top-selling product can vary significantly depending on various factors such as time, location, and industry. It's important to note that the top-selling product in one category or market may not be the same in another. In this dataset has 4 products S-P1, S-P2, S-P3, S-P4 then we find the highest sales through the 4 products**.**

****

* In this graphical represent the year wise sales
* we have a 4 column charts (Product 1,product 2,product 3, product 4)
* In this case we apply the date on x-axis
* Products are apply in y-axis
* Dates are filtered as 2011 – 2023

**PEAK SALES PERIODS:**

* In the context of business and data analysis, refers to the long-term movement or direction in which a company's sales figures are changing.
* It involves analyzing historical sales data to identify patterns or tendencies in sales performance over a specific period of time.
* Understanding sales trends is crucial for businesses because it can help them make informed decisions



**CUSTOMER PREFERENCE:**

* Customer preferences refer to the specific choices, tastes, and desires of individual customers or groups of customers when it comes to products, services, or experiences offered by a business.
* Understanding customer preferences is crucial for businesses because it enables them to tailor their offerings and marketing strategies to better meet customer needs and increase customer satisfaction.



**DERIVED ACTIONABLE INSIGHT**

Here are some more examples of actionable insights that can be derived from product sales analysis:

**Identify which products are frequently purchased together:**

* This information can be used to create product bundles and cross-sell products to customers.

**Identify which products are often returned:**

* This information can be used to improve product quality, reduce customer churn, and develop better return policies.

**Identify which products are sold best in which regions or channels:**

* This information can be used to optimize inventory levels and allocate marketing resources more effectively.

**Identify which products are most seasonal:**

* This information can be used to develop seasonal promotions and sales strategies.

**Identify which products are trending:**

* This information can be used to develop new products or improve existing products to meet customer demand.

**Example:**

how to use product sales analysis to derive an actionable insight for a software company:

The software company analyzes its product sales data and finds that customers who purchase its customer relationship management (CRM) software also frequently purchase its marketing automation software. Based on this information, the software company can develop a bundled product offering that includes its CRM software and its marketing automation software. The software company can also promote its marketing automation software to customers who have purchased its CRM software in the past.

By using product sales analysis to derive actionable insights, businesses can increase sales, improve profitability, and reduce customer churn.

**HOW THE INSIGHTS FORM THE ANALYSIS CAN GUIDE INVENTORY MANAGEMENT AND MARKETING STRATEGIES**

The insights from product sales analysis can guide inventory management and marketing strategies in a number of ways.

**INVENTORY MANAGEMENT:**

**Identifying the most popular products:**

* This information can be used to ensure that adequate inventory levels are maintained for popular products to avoid stockouts and lost sales.

**Identifying the least popular products:**

* This information can be used to reduce inventory levels for unpopular products to reduce costs and free up space for more popular products.

**Identifying seasonal trends:**

* This information can be used to adjust inventory levels based on seasonal demand to avoid overstocking and understocking.

**Identifying slow-moving products:**

* This information can be used to develop strategies to sell slow-moving products, such as discounts, promotions, and bundling with other products.

**MARKETING STRATEGIES:**

**Identifying the most profitable products:**

* This information can be used to focus marketing efforts on the most profitable products.

**Identifying the most valuable customer segments:**

* This information can be used to target marketing campaigns to the most likely customers.

**Identifying the key drivers of product sales:**

* This information can be used to develop marketing messages that focus on the benefits that are most important to customers.

**Identifying sales opportunities and threats:**

* This information can be used to develop marketing strategies to capitalize on sales opportunities and mitigate sales threats.

**FUTURE SCOPE**

Here are some of the key trends that will shape the future scope of product sales analysis in data analytics:

**Increased use of artificial intelligence (AI) and machine learning (ML):**

* AI and ML will play an increasingly important role in product sales analysis. AI and ML algorithms can be used to automate tasks such as data collection, cleaning, and analysis. They can also be used to build predictive models that can forecast future sales and identify sales opportunities and threats.

**Real-time analytics:**

* Businesses will demand real-time insights into their product sales data. This will enable them to make faster and more informed decisions about pricing, marketing, and inventory management.

**More granular analytics:**

* Businesses will require more granular analytics that can provide insights into specific product categories, customer segments, and geographic regions. This will enable them to better understand their customers and develop more targeted marketing and sales strategies.

**Integration with other data sources:**

* Product sales analysis will be increasingly integrated with other data sources, such as customer data, marketing data, and financial data. This will provide businesses with a more holistic view of their business and enable them to make better decisions.

**ADVANTAGES**

**Identify the most profitable products and customer segments:**

* Product sales analysis can help businesses identify their most profitable products and customer segments. This information can be used to make better decisions about pricing, marketing, product development, and inventory management.

**Improve marketing campaigns:**

* Product sales analysis can help businesses identify which marketing campaigns are most effective. This information can be used to improve future marketing campaigns and increase sales.

**Identify sales opportunities and threats:**

* Product sales analysis can help businesses identify sales opportunities and threats. This information can be used to develop strategies to capitalize on sales opportunities and mitigate sales threats.

**Make better business decisions:**

* Product sales analysis can help businesses make better decisions about all aspects of their business, from pricing to marketing to product development. By understanding their sales data, businesses can make informed decisions that lead to increased sales and profitability.

**DISADVANTAGES**

**Cost and time:**

* Product sales analysis can be costly and time-consuming. Businesses need to invest in the necessary tools and resources to collect, clean, and analyze their sales data.

**Complexity:**

* Product sales analysis can be complex, especially for businesses with large amounts of data. Businesses need to have the expertise to analyze their sales data and interpret the results.

**Data quality:**

* The accuracy of the insights from product sales analysis depends on the quality of the data. Businesses need to ensure that their sales data is accurate and complete before analyzing it.

**Misinterpretation:**

* It is important to interpret the results of product sales analysis carefully. It is possible to misinterpret the results and draw incorrect conclusions. Businesses need to have the expertise to interpret the results of product sales analysis correctly

**CONCLUSION**

Product Sales Analysis empowers businesses to make informed decisions, optimize their operations, and achieve greater profitability. By leveraging data, visualizations, and insights, companies can adapt to market dynamics and stay competitive in a constantly evolving business environment.