**PRODUCT SALES ANALYSIS**

**Team Member**

**312821243042: VASANTH.B**

**Phase-2 Document Submission**

**Project: Product sales analysis**

**Problem Statement:**



**Problem Statement:**

The project involves using IBM Cognos to analyze sales data and extract insights about top selling products, peak sales periods, and customer preferences. The objective is to help businesses improve inventory management and marketing strategies by understanding sales trends and customer behavior. This project includes defining analysis objectives, collecting sales data, designing relevant visualizations in IBM Cognos, and deriving actionable insights.

**Introduction:**

Product sales analysis is the strategic examination of a company's sales data to uncover insights that drive growth and success. It's the compass that guides businesses in understanding product performance, customer behavior, and market trends, enabling informed decisions for improved competitiveness. In this brief overview, we'll explore the significance of product sales analysis and how it empowers companies to thrive in the ever-evolving marketplace. \

Dataset Link:https://www.kaggle.com/datasets/ksabishek/product-sales-data

**1. Data Preprocessing:**

**Categorical to Numerical Conversion**: Convert categorical variables (e.g., product

categories or customer segments) into numerical values using techniques like one-hot

**Handling Missing Values:**Identify and handle outliers in the data, as extreme values

can skew analysis results. You may choose to remove outliers, transform them, or

**Data Cleaning:**Identify and rectify errors, inconsistencies, and missing values in the

data. This includes removing duplicates, correcting typos, and addressing data entry

errors replace them with more representative values encoding or label encoding

**Python Program:**

**Input:**

**import pandas as pd**

**# Sample product sales data in a DataFrame**

**data = {**

**'ProductID': [1, 2, 3, 1, 2, 3, 1, 2, 3],**

**'ProductName': ['Product A', 'Product B', 'Product C', 'Product A', '**

**Product B', 'Product C', 'Product A', 'Product B', 'Product C'],**

**'SalesAmount': [100, 150, 75, 120, 200, 90, 80, 130, 60]**

**}**

**df = pd.DataFrame(data)**

**# Calculate total sales and average sales price per product**

**product\_sales = df.groupby(['ProductID', 'ProductName']).agg({'**

**SalesAmount': ['sum', 'mean']}).reset\_index()# Rename the columns for clarity**

**product\_sales.columns = ['ProductID', 'ProductName', 'TotalSales', '**

**AveragePrice']**

**# Display the result**

**print(product\_sales)**

**Output:**

**Product ID ProductName TotalSales AveragePrice**

**0 1 Product A 300 100.000000**

**1 2 Product B 480 160.000000**

**2 3 Product C 225 75.000000**

**Data Collection and Preparation**

Collect data from the identified sources. This may involve using tools,

APIs, or manual data entry, depending on the source. Make sure you

have access to the necessary data.

**2. Exploratory Data Analysis:**

Calculate basic statistics (mean, median, etc.) for numeric

variables and frequencies for categorical variables.

Exploratory Data Analysis is a crucial foundation for any product

sales analysis.

It helps you understand your data, identify trends, and formulate

hypotheses, ultimately leading to more targeted and effective

analysis and decision-making.

**3. Feature Engineering:**

If you meant "feature engineering," which is a common technique

in data analysis and machine learning, I can provide some

guidance on that.

Feature engineering is the process of creating new, meaningful

features from existing data to improve the performance of your

analysis or predictive models.

**4. Model Selection:**

Based on your objectives, choose the most suitable modeling techniques.

Options include regression models, time series forecasting, machine

learning, and more, depending on your specific needs.

**5.Model Training and Evaluation:**

Train your selected models using historical data. Ensure that you split your

data into training and testing sets for model validation.

Assess the performance of your models using appropriate evaluation

metrics, such as Mean Absolute Error (MAE) for regression or F1-score for classification.

**6.Result Analysis and Visualization:**

After data processing and modeling, assess the results to identify

key insights and trends.

Use charts, graphs, and interactive dashboards to visually

represent data trends, correlations, and anomalies.

**Conclusion:**

In conclusion, product sales analysis is a vital component of any

business strategy, providing valuable insights to optimize performance

and drive growth. Through data collection, preparation, modeling, and

visualization, organizations can gain a deeper understanding of their

products, customers, and market trends. This understanding allows for

data-driven decision-making, leading to improved marketing

strategies, pricing adjustments, inventory management, and customer

satisfaction. Continuous monitoring, feedback, and adaptation are

crucial to ensuring that the analysis remains relevant and effective in a

dynamic business environment. Ultimately, product sales analysis

empowers businesses to enhance their competitiveness, increase

profitability, and better serve their customers