**KINGSTON ENGINEERING COLLEGE – 5113**

DATA ANALYTICS WITH COGNOS

**PRODUCT SALES ANALYSIS-PROJECT 5**

**TEAM MEMBERS:**

S.PON LAVANYA-au511321205021

S. POORNIMA-au511321205022

J. KAVIYA-au511321205014

S. SRINIDHI-au511321205039

**INTRODUCTION:**

In today's rapidly evolving business landscape, data-driven decision-making is the key to success. As we navigate the intricacies of the market, understanding our product sales performance becomes not only essential but transformative. This Product Sales Analysis Project aims to delve deep into the data, uncover insights, and empower our organization to make informed strategic choices.

In this project, we will comprehensively analyze the sales performance of our product(s) over a specified time period. This analysis will encompass various facets of our sales data, including but not limited to:

* **Revenue**: Examining the total revenue generated by our product(s) and assessing its growth or decline trends.
* **Sales Volume**: Investigating the quantity of units sold and identifying patterns in sales volume.
* **Profit Margins**: Assessing the profitability of our product(s) by examining profit margins.
* **Market Share**: Understanding our product's market share relative to competitors.
* **Customer Insights**: Gaining a deeper understanding of customer behavior, preferences, and demographics.
* **Sales Channels**: Analyzing the effectiveness of different sales channels and distribution methods.
* **Marketing Effectiveness**: Evaluating the impact of marketing campaigns and promotions on sales.
* **Inventory Management**: Optimizing inventory levels for cost-efficiency.

As we embark on this Product Sales Analysis Project, our goal is to equip our organization with the knowledge and insights necessary to thrive in a competitive marketplace. By examining our product sales data from multiple angles, we aim to make strategic decisions that will lead to increased revenue, improved customer satisfaction, and sustained growth.

**ABOUT DATASET:**

Where did we get the dataset?

**Kaggle:**

**.** we got our ‘Product sales Analysis’ dataset from kaggle

[**https://www.kaggle.com**](https://www.kaggle.com)

**.** kaggle is a popular platform for sharing datasets and hosting data science competitions.

**DETAILS:**

A product sales analysis involves a detailed examination of various aspects related to the performance of a product or group of products in the market. To conduct a comprehensive product sales analysis, you'll need to gather, process, and analyze a wide range of data and information. Here are the key details and steps involved in conducting a product sales analysis:

Q1- Total unit sales of product 1  
 Q2- Total unit sales of product 2  
 Q3- Total unit sales of product 3  
 Q4- Total unit sales of product 4  
 S1- Total revenue from product 1  
 S2- Total revenue from product 2  
 S3- Total revenue from product 3  
 S4- Total revenue from product 4

The dataset contains the following example of details:

On 13-06-2010 , product 1 had been brought by 5422 people and INR 17187.74 had been generated in revenue from product

REC corporate needs you to solve the following questions:

1) Is there any trend in the sales of all four products during certain months?  
2) Out of all four products which product has seen the highest sales in all the given years?  
3) The company has all it's retail centers closed on the 31st of December every year. Mr: Hariharan the CEO would love to get an estimate on no: of units of each product that could be sold on 31st of Dec every year , if all their retail centers were kept open.  
4) The CEO is considering an idea to drop the production of any one of the products. He wants you to analyze this data and suggest whether his idea would result in a massive setback for the company.  
5) The CEO would also like to predict the sales and revenues for the year 2024. He wants you to give a yearly estimate with the best possible accuracy.

**COLUMNS THAT WE USED:**

**1.Product Information:**

* Product Name/ID: A unique identifier for each product.
* Category/Type: Categorize products into groups for better analysis.
* SKU (Stock Keeping Unit): A unique identifier for inventory management.
* Description: A brief description of the product.

**2.Sales Data:**

* Date of Sale: The date when the sale occurred.
* Quantity Sold: The number of units sold.
* Sales Price: The price at which each unit was sold.
* Total Sales: The total revenue generated from each sale (Quantity Sold \* Sales Price).
* Discounts: Any discounts applied to the sale.

**3.Customer Information:**

* Customer Name/ID: A unique identifier for each customer.
* Customer Segment: Group customers based on characteristics like demographics or buying behavior.
* Customer Location: The geographic location of the customer.

**4.Order Information:**

* Order Number/ID: A unique identifier for each order.
* Payment Method: The method used for payment (e.g., credit card, cash).
* Shipping Method: The method used for shipping (e.g., standard, express).

**5.Profit and Cost Analysis:**

* Cost per Unit: The cost to produce or purchase each unit.
* Gross Profit: Total sales revenue minus the cost of goods sold (COGS).
* Gross Margin: Gross Profit as a percentage of total sales.
* Net Profit: Gross Profit minus operating expenses.

**6.Time-Based Metrics:**

* Monthly/Quarterly/Annual Sales: Sales data broken down by time periods for trend analysis.
* Seasonal Trends: Identify patterns related to seasons or holidays.
* Average Daily Sales: Average sales per day.

**7.Performance Metrics:**

* Sales Growth Rate: Percentage change in sales compared to a previous period.
* Customer Acquisition Cost (CAC): The cost of acquiring each new customer.
* Customer Lifetime Value (CLV): The value of a customer over their entire relationship with your business.

**8.Customer Feedback and Reviews :**

* Customer Ratings: Average product ratings based on customer reviews.
* Customer Comments: Feedback and comments left by customers.

**LANGUAGE THAT WE USED:**

**1.Python:**

* Python is a versatile programming language with numerous libraries and packages for data analysis, including Pandas, NumPy, Matplotlib, Seaborn, and Plotly. You can use Python to clean, manipulate, and analyze sales data, as well as create visualizations and statistical models.

**2.SQL (Structured Query Language):**

* SQL is essential for database management and querying. You can use SQL to extract, filter, and aggregate sales data from relational databases.

**3.Jupyter Notebooks:**

* Jupyter Notebooks provide an interactive environment for data analysis using Python or R. They are excellent for documenting and sharing your analysis steps and results.

**4.SAS (Statistical Analysis System):**

* SAS is a powerful software suite used for advanced analytics, business intelligence, and data management. It's commonly used in large enterprises for sales forecasting and advanced analytics.

**5.Statistical Software Packages:**

* Depending on the complexity of your analysis, you may use statistical software packages like Stata or Minitab for specialized statistical modeling and hypothesis testing.

**LIBRARIES THAT WE USED:**

**1. Pandas:**

Pandas is a powerful library for data manipulation and analysis. It provides data structures and functions to efficiently handle

and analyze large datasets, making it ideal for sales analysis tasks.

**2. NumPy:**

NumPy is a fundamental library for scientific computing in Python. It provides support for large, multi-dimensional arrays and matrices, as well as a collection of mathematical functions. It can be used for various calculations and statistical analysis in sales analysis.

**3. Matplotlib:**

Matplotlib is a plotting library that allows you to create various types of visualizations, such as line plots, bar charts, scatter plots, and histograms. It is helpful for visualizing sales data and trends.

**4. Seaborn:**

Seaborn is a statistical data visualization library built on top of Matplotlib. It provides a high-level interface for creating attractive and informative statistical graphics. It can be used to create more advanced and visually appealing sales visualizations.

**5. Statsmodels:**

Statsmodels is a Python module that provides statistical models and tests for various types of data analysis. It includes functions for regression analysis, time series analysis, and hypothesis testing, which can be useful for sales analysis tasks.

**SAMPLE CODE:**

**import** pandas **as** pd  
import matplotlib.pyplot **as** plt  
  
# Load the sales data into a pandas DataFrame  
sales\_data = pd.read\_csv('sales\_data.csv')  
  
# Convert the date column to datetime type  
sales\_data['Date'] = pd.to\_datetime(sales\_data['Date'])  
  
# Calculate the total sales for each quarter  
sales\_data['Total\_Q1'] = sales\_data['Q-P1'] + sales\_data['Q-P2'] + sales\_data['Q-P3'] + sales\_data['Q-P4']  
sales\_data['Total\_Q2'] = sales\_data['S-P1'] + sales\_data['S-P2'] + sales\_data['S-P3'] + sales\_data['S-P4']  
  
# Calculate the average sales for each quarter  
sales\_data['Avg\_Q1'] = sales\_data[['Q-P1', 'Q-P2', 'Q-P3', 'Q-P4']].mean(axis=1)  
sales\_data['Avg\_Q2'] = sales\_data[['S-P1', 'S-P2', 'S-P3', 'S-P4']].mean(axis=1)  
  
# Plot the total sales for each quarter  
plt.plot(sales\_data['Date'], sales\_data['Total\_Q1'], label='Total Q1 Sales')  
plt.plot(sales\_data['Date'], sales\_data['Total\_Q2'], label='Total Q2 Sales')  
plt.xlabel('Date')  
plt.ylabel('Total Sales')  
plt.title('Total Sales by Quarter')  
plt.legend()  
plt.show()  
  
# Print the average sales for each quarter  
print("Average Sales for Q1:", sales\_data['Avg\_Q1'].mean())  
print("Average Sales for Q2:", sales\_data['Avg\_Q2'].mean())

**Expected Output:**

Average Sales for Q1: 3322.6666666665  
Average Sales for Q2: 11845.8125

THANK YOU