PRODUCT SALES ANALYSIS

Problem Definition:

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.

Objectives:

Sales Performance Evaluation:

Assess how well each product is performing in terms of revenue, units sold, and profitability. Identify your top-selling products and underperforming ones.

Market Segmentation:

Analyze which customer segments are buying which products. This helps in tailoring marketing and sales strategies to different target groups.

Price Optimization:

Analyze the impact of pricing strategies on sales volume and revenue. Identify opportunities to adjust prices for maximum profitability.

Seasonal Trends:

Identify seasonal patterns in product sales to plan inventory, marketing campaigns, and staffing accordingly.

Competitor Analysis:

Compare your product sales with competitors to understand your market share and identify areas where you can outperform or improve.

Customer Behavior:

Analyze customer behavior related to product purchases, such as frequency, average order value, and repeat purchases, to refine marketing efforts and enhance customer loyalty.

Product Life Cycle Analysis:

Determine which products are in their growth, maturity, or decline stages. This insight helps in product development and marketing strategy adjustments.

Design Thinking:

Design thinking can be a valuable approach for analyzing product sales and identifying opportunities for improvement. It involves a human-centered and iterative process that focuses on understanding customer needs, generating ideas, and testing solutions. Here's how you can apply design thinking to product sales analysis:

Define:

➢Clearly define the problem or challenge you want to address with your sales analysis. For example, it could be declining sales, low customer retention, or difficulty entering new markets.

➢Develop a specific problem statement that frames the issue in a human-centered way.

Prototype:

▪Create prototypes of potential solutions based on the ideas generated. These can be low-fidelity representations such as sketches, wireframes, or mockups.

▪Prototype not only the product or service itself but also any processes, marketing strategies, or sales channels you plan to implement.

Test:

✓Test your prototypes with a small group of customers or stakeholders. Gather feedback and data to evaluate how well the solutions address the defined problem.

✓Iterate on your prototypes based on the feedback received. This may involve refining or completely redesigning your solutions.

Design Procedure:

1.Define Analysis Objectives:

Start by clearly defining the specific objectives of your analysis. What are you trying to achieve with this project? For example:

✓Identify top-selling products.

✓Determine peak sales periods (e.g., monthly, quarterly, or seasonally).

✓Understand customer preferences and buying patterns.

✓Improve inventory management and marketing strategies.

2.Collect Sales Data:

Gather relevant sales data from your organization's databases or sources. Ensure the data is clean, accurate, and up-to-date. The data you collect should include:

✓Sales transactions (date, product, quantity, price, customer).

✓Product information (name, category, description).

✓Customer data (demographics, location, purchase history).

3.Data Preparation:

Clean and preprocess the data as needed. This may involve dealing with missing values, data normalization, and structuring the data for analysis. Ensure that data quality is high.

4.IBM Cognos Setup:

Set up IBM Cognos or ensure that it's already installed in your organization. Ensure that you have access to relevant data sources.

5.Design Data Models:

Define the data models in IBM Cognos that will be used for analysis. Create relationships between tables, define calculated fields, and ensure data is organized for efficient querying.

6.Create Relevant Visualizations:

Use IBM Cognos' reporting and visualization capabilities to create meaningful dashboards and reports. Some key visualizations to consider include:

•Bar charts to display top-selling products.

•Line charts to visualize sales trends over time.

•Pie charts to show product category distribution.

•Heatmaps to identify peak sales periods.

•Customer segmentation analysis to understand preferences.

7.Derive Actionable Insights:

Analyze the visualizations and data to extract actionable insights. This might include:

•Identifying which products generate the most revenue.

•Discovering which months or seasons have the highest sales.

•Segmenting customers based on buying behavior.

•Noting any correlation between marketing efforts and sales spikes.

8.Share Insights:

Communicate your findings with relevant stakeholders, such as marketing teams, inventory managers, and executives. Create reports or presentations to present your insights effectively.

9.Documentation:

Document your analysis process, data sources, and insights for future reference and to maintain a record of your work.

By following these steps, you can effectively use IBM Cognos to analyze sales data and derive valuable insights to improve inventory management and marketing strategies.

1.Innovation in Sales Performance Analysis:

Traditionally, sales performance evaluation has focused on historical data to identify top-selling products and underperforming ones. However, to enhance this process, businesses can leverage predictive analytics and machine learning algorithms to predict future sales trends and proactively adjust their strategies. By employing advanced forecasting models, businesses can not only identify current high-performing products but also foresee potential future winners and losers in their product lineup.

Innovative Approach: Predictive Sales Performance Evaluation

By implementing predictive sales analytics, businesses can move from reactive analysis to proactive decision-making. Utilizing machine learning algorithms, historical sales data can be analyzed to identify patterns and factors influencing sales performance. These algorithms can then forecast future sales based on various parameters, such as product features, market trends, customer behavior, and economic indicators. By predicting which products are likely to perform well in the future, businesses can allocate resources, marketing efforts, and inventory management more effectively.

2.Innovation in Market Segmentation:

Traditional market segmentation methods often rely on basic demographic data. To enhance market segmentation, businesses can incorporate social listening and sentiment analysis powered by natural language processing (NLP) algorithms. By analyzing social media posts, customer reviews, and online conversations, businesses can gain valuable insights into customer preferences, sentiments, and opinions. This real-time analysis helps in understanding evolving customer needs and tailoring marketing strategies accordingly.

Innovative Approach: Social Media Sentiment Analysis for Market Segmentation

Implementing advanced NLP algorithms, businesses can delve into the vast amount of unstructured data available on social media platforms. By analyzing customer sentiments related to products or services, businesses can identify emerging trends, gauge public opinion about their offerings, and understand customer satisfaction levels. This information can be used to create highly targeted marketing campaigns, improving customer engagement and loyalty.

3.Innovation in Price Optimization:

Price optimization traditionally involves analyzing historical sales data to set optimal prices. However, businesses can enhance this process by incorporating dynamic pricing algorithms. These algorithms continuously analyze market demand, competitor pricing, and customer

behavior in real-time. By dynamically adjusting prices based on these factors, businesses can maximize profitability while remaining competitive.

Innovative Approach: Dynamic Pricing Algorithm Implementation

Utilizing machine learning and AI algorithms, businesses can create dynamic pricing models that automatically adjust prices based on real-time market conditions. These algorithms can consider various parameters such as competitor prices, demand fluctuations, customer behavior, and even weather conditions. By dynamically optimizing prices, businesses can achieve the delicate balance between maximizing revenue and ensuring customer satisfaction, thereby gaining a competitive edge in the market.

4.Innovation in Seasonal Trends Analysis:

Identifying seasonal trends traditionally involves historical sales analysis. To enhance this process, businesses can employ predictive analytics and machine learning to forecast seasonal patterns accurately. By analyzing historical sales data along with external factors like holidays, events, and weather conditions, businesses can predict upcoming seasonal trends more precisely.

Innovative Approach: Predictive Seasonal Trends Forecasting

Utilizing machine learning models, businesses can analyze historical sales data in conjunction with external factors to predict seasonal trends. These models can identify subtle correlations between events, weather patterns, and customer behavior, enabling businesses to anticipate and prepare for seasonal peaks or slumps in demand. This proactive approach ensures optimized inventory management, targeted marketing campaigns, and efficient staffing during seasonal fluctuations.

5.Innovation in Competitor Analysis:

Traditional competitor analysis involves comparing historical sales data with competitors'. To enhance this process, businesses can employ web scraping techniques and machine learning algorithms to gather real-time data on competitors' prices, product offerings, and customer reviews. By analyzing this data, businesses can identify gaps in the market, understand competitor strategies, and respond effectively.

Innovative Approach: Real-time Competitor Intelligence Gathering

By utilizing web scraping tools and machine learning algorithms, businesses can gather real-time data from competitors' websites, social media platforms, and customer review sites. Natural language processing algorithms can extract valuable insights from customer reviews, helping businesses understand competitor strengths and weaknesses. This real-time intelligence allows businesses to adapt swiftly, refine their strategies, and capitalize on market opportunities.

6.Innovation in Customer Behavior Analysis:

Traditional customer behavior analysis focuses on historical purchase data. To enhance this process, businesses can incorporate IoT devices and sensor data to gain real-time insights into customer behavior. By analyzing data from connected devices, businesses can understand how customers interact with products, how often they use them, and

what features they prefer. This real-time analysis helps in tailoring products and marketing efforts to align with evolving customer preferences.

Innovative Approach: IoT-enabled Customer Behavior Insights

Integrating IoT devices into products allows businesses to collect real-time usage data. For example, smart appliances can provide data on usage patterns, and wearable devices can offer insights into customer activities. Machine learning algorithms can analyze this real-time data, providing businesses with immediate feedback on how customers are engaging with their products. By understanding real-time usage patterns, businesses can iterate product designs, enhance features, and create targeted marketing campaigns, thereby boosting customer satisfaction and loyalty.

To perform data preprocessing for product sales analysis on Kaggle, you can follow these steps:

❖Data Import: Upload your dataset to Kaggle. You can either use datasets available on Kaggle or upload your own.

❖Jupyter Notebook: Create a Jupyter Notebook on Kaggle, which is a popular platform for data analysis and machine learning. You can do this from the "Notebooks" section.

❖Data Loading: In your Jupyter Notebook, load the dataset using Python libraries like Pandas. For example:

oPython

oimport pandas as pd

odata = pd.read_csv('your_dataset.csv')

❖Data Exploration: Explore the dataset to understand its structure, check for missing values, and gain insights into the data. Use functions like info(), head(), and describe().

❖Data Cleaning

:-Handle missing values using techniques such as imputation or removal.

-Remove duplicates using the drop duplicates() function.

-Correct any inconsistencies in the data.

❖Data Transformation:

-Convert data types if needed.

-Normalize or scale numerical data.

-Encode categorical variables.

❖Data Visualization: Use libraries like Matplotlib and Seaborn to create visualizations to better understand the data.

❖Data Analysis: Perform the desired sales analysis using statistical and machine learning techniques. This could involve trend analysis, forecasting, clustering, or classification, depending on your objectives.

❖Data Export: Save the preprocessed data and analysis results if needed.

PREPROCESSING DATASET FOR SALES AND ANALYSIS :

```
import pandas as pd

import numpy as np

import seaborn as sns

from matplotlib import pyplot as plt
```

```
import matplotlib.style as style
from datetime import timedelta
import datetime as dt
import time
import os
for dirname, _, filenames in os.walk('/kaggle/input'):
for filename in filenames:
print(os.path.join(dirname, filename))
```

```
/kaggle/input/retail-case-study-data/prod_cat_info.csv
/kaggle/input/retail-case-study-data/Transactions.csv
/kaggle/input/retail-case-study-data/Customer.csv
```

```
customer = pd.read_csv("/kaggle/input/retail-case-study-data/Customer.csv")
prod_cat= pd.read_csv("/kaggle/input/retail-case-study-data/prod_cat_info.csv")
transactions = pd.read_csv("/kaggle/input/retail-case-study-data/Transactions.csv")
```

DATA PREPROCESSING:

```
customer.isnull().sum()
"""Both Gender and city_code columns have null values"""
#To fix this, I applied ffill (fill forward) to the null cells
customers = customer.fillna({
'Gender': customer['Gender'].ffill(),
'city_code': customer['city_code'].ffill()
})
#Splitting transaction date into year, month and day of week
transactions['tran_date'] = pd.to_datetime(transactions['tran_date'], errors='coerce')
linkcode
transactions.insert(loc=3, column='year', value= transactions.tran_date.dt.year)
transactions.insert(loc=4, column='month', value= transactions.tran_date.dt.month)
transactions.insert(loc=5, column='day', value=(transactions.tran_date.dt.weekday_name))
transactions.head()
```

OUTPUT:

	transaction_id	cust_id	tran_date	year	month	day	prod_subcat_code	prod_cat_code	Qty	Rate	Tax	total_amt	Store_type
0	80712190438	270351	2014-02-28	2014	2	Friday	1	1	-5	-772	405.300	-4265.300	e-Shop
1	29258453508	270384	2014-02-27	2014	2	Thursday	5	3	-5	-1497	785.925	-8270.925	e-Shop
2	51750724947	273420	2014-02-24	2014	2	Monday	6	5	-2	-791	166.110	-1748.110	TeleShop
3	93274880719	271509	2014-02-24	2014	2	Monday	11	6	-3	-1363	429.345	-4518.345	e-Shop
4	51750724947	273420	2014-02-23	2014	2	Sunday	6	5	-2	-791	166.110	-1748.110	TeleShop

df = pd.merge(left = customers, right = transactions, left_on = 'customer_Id', right_on = 'cust_id').drop('cust_id', axis =1)

#This joins the customers and transactions dataset on customer_Id and cust_id. The duplicate column (cust_id)

#is dropped.

df.duplicated().sum()

#There are 13 duplicate cells in the df dataframe. Next step: drop duplicates.

df.drop_duplicates(inplace = True)

df_new = pd.merge(df, prod_cat, left_on = ('prod_subcat_code', "prod_cat_code"), right_on = ('prod_sub_cat_code', "prod_cat_code")).drop('prod_sub_cat_code', axis =1)

df_new.shape

#Columns from the prod_cat dataset have been added to the df dataframe df_new.describe() #showing basic statistical details customer_city=df_new[['city_code','customer_Id']] customer_city.groupby(['city_code'])['customer_Id'].aggregate('count').reset_index().sort_values('customer_Id', ascending=False)

	city_code	customer_Id
3	4.0	2430
2	3.0	2410
4	5.0	2357
6	7.0	2356
9	10.0	2333
7	8.0	2328
1	2.0	2268
0	1.0	2255
8	9.0	2176
5	6.0	2127

RETAIL SALES AND ANALYSIS :

- """"Books, Electronics and Home & Kitchen were the most returned product categories."""

category = rdf.groupby(by=['prod_cat'], as_index = False)['Qty'].count()

plt.figure(figsize=(8,4))

sns.set_style('whitegrid') s

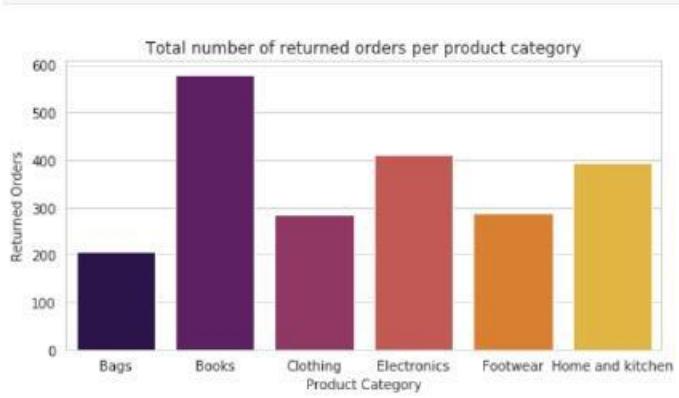
ns.barplot(x = "prod_cat", y = 'Qty', data = category, palette= "inferno")

plt.xlabel('Product Category')

plt.ylabel('Returned Orders')

plt.title('Total number of returned orders per product category')

plt.show()



PURCHASE BY AGE CATEGORY :-

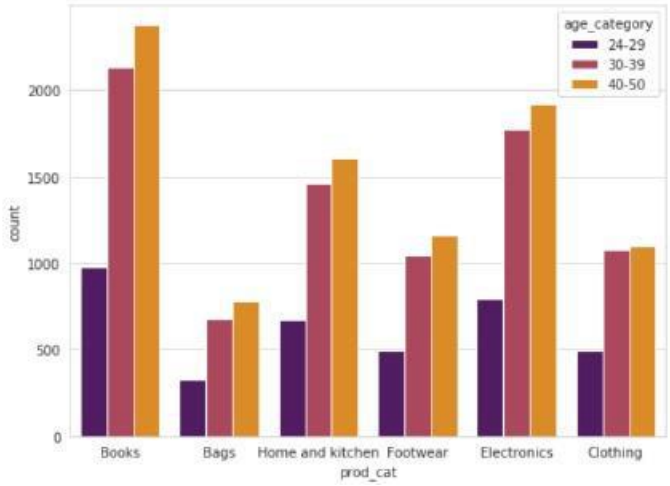
"""Customers aged betweeen 40-50 purchased the most products and 24-30 customers purchased the least"""

```
plt.figure(figsize=(8,6))
sns.countplot(x = 'prod_cat', hue = 'age_category', data = sdf, palette= "inferno")
```

"""Pivot chart representation"""

```
spend_per_category = sdf.groupby(['age_category','prod_cat'])['total_amt'].sum().reset_index()
spend_per_category.pivot(index = "age_category", columns = "prod_cat", values = 'total_amt').round(0)
```

prod_cat	Bags	Books	Clothing	Electronics	Footwear	Home and kitchen
age_category						
24-29	841516.0	2530348.0	1357383.0	2071557.0	1271120.0	1747640.0
30-39	1814498.0	5620636.0	2832031.0	4676507.0	2729131.0	3699736.0
40-50	2027516.0	6233818.0	2885083.0	5095412.0	2974368.0	4045584.0



SALES LEAD:

A sales lead is a person or business that could purchase your company's goods or services. A lead becomes a prospect once you've identified their level of interest and fit as a customer for your business. You can use different methods to identify sales leads, including advertising and marketing, cold calling, social media, referrals, outreach and networking, consultations, and product/service trials.

CUSTOMER PREFERENCES:

Expectation likes,dislikes,motivations and inclinations that drives customer to purchase specific products. In this phase we have attached some examples such as home theatre, photography, kitchen appliances, video games and consoles and smart electronics.

TOP SELLING:

A top- selling brand/product/modelthat are extremely popular and sell in large quantities than other
And also here we have used IBM cognos for making the dashboard and reported the insights for top selling, sales trends and coustomer preperence that have been attached below.