



Data Analytics with Cognos

PRODUCT SALES ANALYSIS

PHASE 5

NAME:PRATHIBA M

REG NO:510421205035

GUIDED BY:MRS.B.MOOHAMBIGAI



Introduction

We will discuss a comprehensive approach to analyzing product sales, including Data Visualization, Analysis Objectives, Data collection. We will also include images to illustrate our points.

Methods

Predicting future sales involves using historical data, statistical algorithms, machine learning algorithms techniques to analyze past sales data patterns and make projections about future sales performance.



Development Phases



- Product Sales Objective
- Data Collection
- Data Analysis
- Data Visualization
- Insights from Visualization
- Sales Trends

Product Sales Objective

- Sales productivity
- Close more deals
- Faster response time
- Increase Revenue
- Customer retention
- Leads
- Provide Customer Analysis
- Improve Sales Strategies



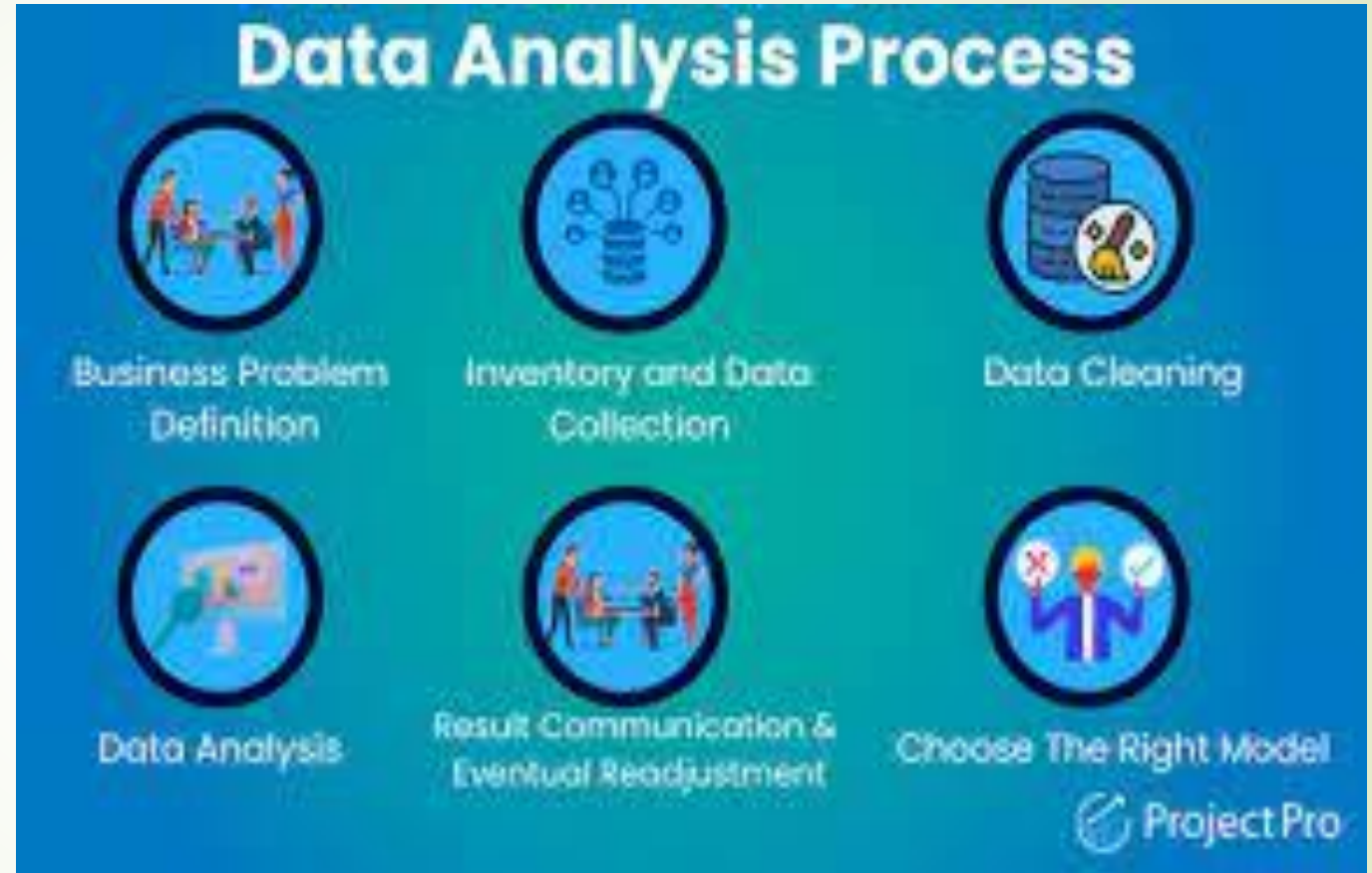
Data Collection

Gather relevant data. This could include historical sales data, customer demographics, website interactions, social media engagement, or any other data points that might influence customer behavior.



Data Analysis

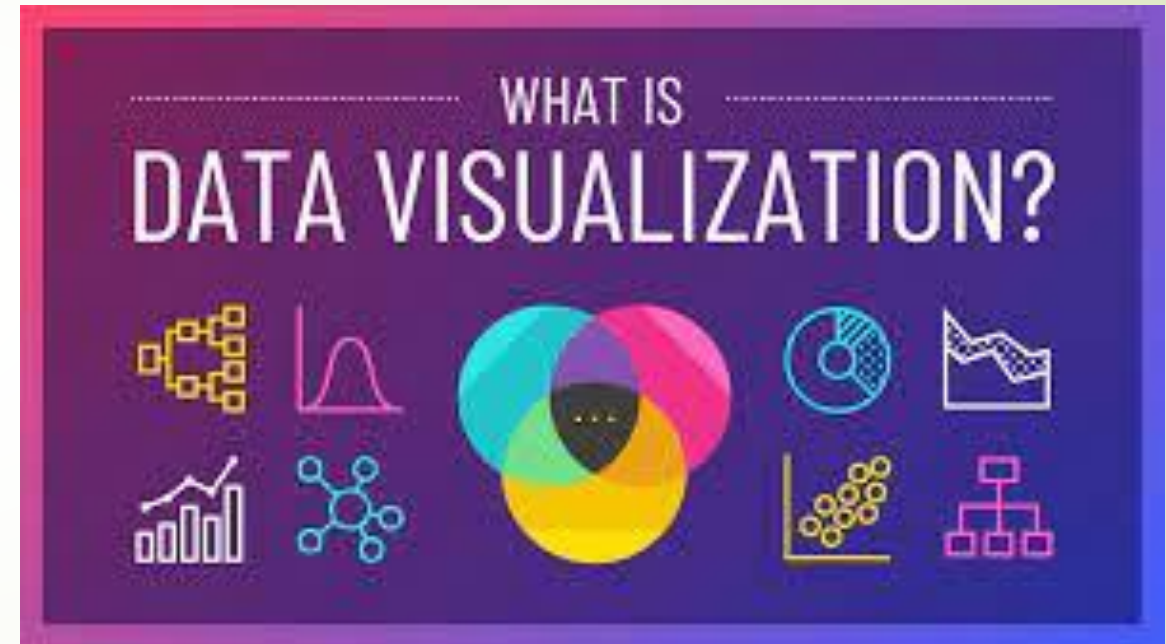
Data analysis is a process for obtaining raw data, and subsequently converting it into information useful for decision-making by users. Data is collected and analyzed to answer questions, test hypotheses, or disprove theories.



Data Visualization

Data visualization is a key component of our approach. By using charts and graphs, we can identify trends and patterns in the sales data that might not

be immediately apparent. This allows us to make informed decisions about product sales and marketing strategies.



Insights From Visualization

Source code

```
import pandas as pd
#Load the dataset
url =
https://www.kaggle.com/datasets/ksabishek/pr
oduct-sales-data
data = pd.read_csv(url)
# Display the first 5 rows of the dataset
print(data.head())
#Print the column names
print(data.columns)
#Check the data types of each column
print(data.dtypes)
# Drop any rows with missing values
data = data.dropna()
# Remove any outliers in the 'quantity' column
data = data[data['quantity'] <= 500]
```

```
import matplotlib.pyplot as plt
import seaborn as sns
# Plot a bar chart of product sales
sns.countplot(data=data, x='product')
plt.title('Product Sales')
plt.show()
# Plot a heatmap of correlations between different
variables
correlation_matrix = data.corr()
sns.heatmap(correlation_matrix, annot=True)
plt.title('Correlation Heatmap')
plt.show()
```

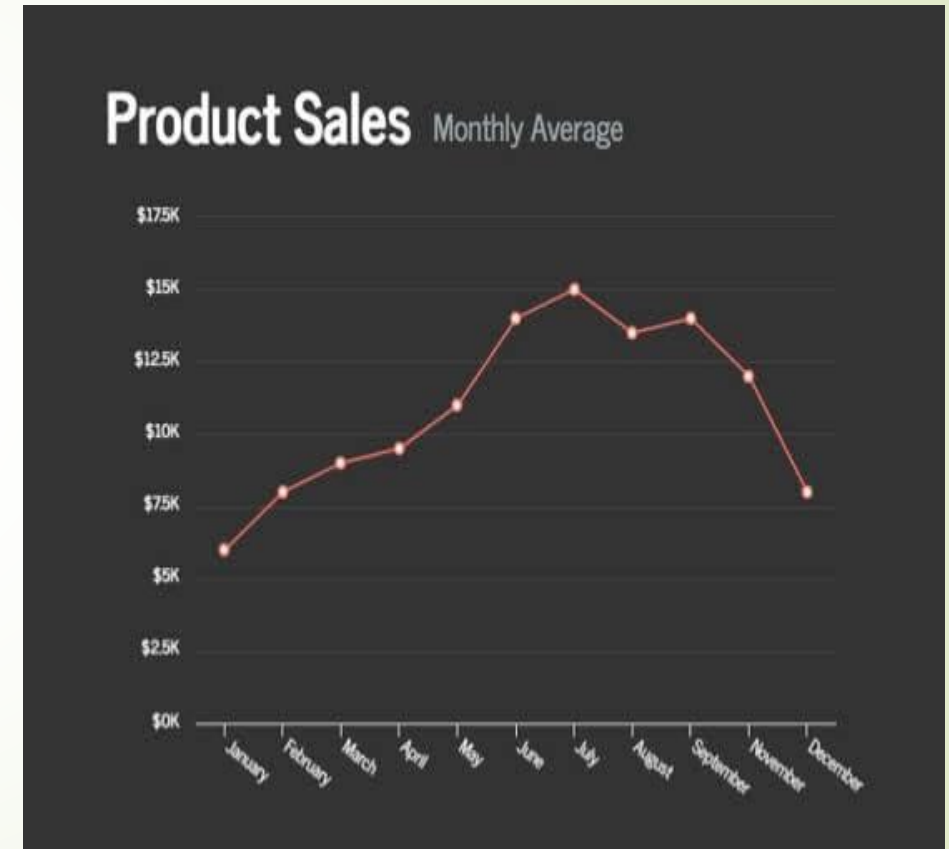

output



Sales Trends

```
# Convert the 'InvoiceDate' column to a
datetime
objectdata['InvoiceDate'] =
pd.to_datetime(data['InvoiceDate'])
# Create a new column called 'MonthYear'
representing the month and year of each
transactiondata
['MonthYear'] =
data['InvoiceDate'].dt.strftime('%Y-%m')
# Group the dataset by 'MonthYear' and
calculate the total quantity sold for each month
sales_trends =
data.groupby('MonthYear')['Quantity'].sum().
reset_index()
```

output



Thank You

