**Case Study 1: Retail Sales Analysis**

1. **Introduction**

Introduction…..case study overview.. problem to be solved…technologies to be used..

**1.1 Purpose**

The purpose of this case study is to analyze retail sales data to identify trends, clean and preprocess data, generate descriptive statistics, and forecast future sales trends. This analysis will provide insights into sales patterns and support decision-making for inventory and marketing strategies.

**1.2 Topics Covered**

* Data Collection – discussion (overview and how did you apply the concept)
* Data Cleaning - discussion
* Descriptive Analytics -discussion
* ……

**1.3 Software and Libraries Used**

* **Python**: Primary programming language for the analysis…
* **Libraries**:
  + **Pandas**: For data manipulation and analysis.
  + **NumPy**: For numerical operations.
  + **Matplotlib**: For data visualization.

1. **Dataset Overview**

Intro and discussion

**2.1.a Description**

The dataset contains retail sales data with the following columns:

* **Date**: Date of the sales transaction.
* **Product\_ID**: Identifier for each product sold.
* **Sales\_Amount**: Total amount of sales in USD.
* **Quantity\_Sold**: Number of units sold in each transaction.

**2.1.b Screenshots/Images**

**2.2 Data Source**

Data can be collected from a CSV file provided by the retailer or downloaded from online resources like Kaggle, containing similar fields.

**2.3 Assumptions**

* Sales data is assumed to be daily or weekly to capture meaningful trends.
* Missing or erroneous values (if any) will be handled as part of the data cleaning process.

**3. Solution Steps**

**3.1 Task 1: Data Collection**

**Steps:**

1. **Import Required Libraries**: Load Pandas, NumPy, and Matplotlib libraries.
2. **Read Data from CSV**: Load the dataset using pd.read\_csv().

**Code:**

python

import pandas as pd

import numpy as np

import matplotlib.pyplot as plt

# Load data

data = pd.read\_csv('sales\_data.csv')

print(data.head())

***[Attach SS here……]***

**3.2 Task 2: Data Cleaning**

**Steps:**

1. **Check for Missing Values**: Identify and handle missing values in the dataset.
2. **Data Type Conversion**: Ensure all columns are of appropriate data types (e.g., Date as DateTime).
3. **Handle Duplicates**: Check for and remove any duplicate entries.

**Code:**

python

# Check for missing values

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

# Fill missing values in 'Sales\_Amount' and 'Quantity\_Sold' with 0 (or mean/median based on context)

data['Sales\_Amount'].fillna(0, inplace=True)

data['Quantity\_Sold'].fillna(0, inplace=True)

# Convert 'Date' column to datetime format

data['Date'] = pd.to\_datetime(data['Date'])

# Drop duplicates if any

data.drop\_duplicates(inplace=True)

# Verify cleaned data

print(data.info())

***[Attach SS here……]***

**3.3 Task 3: Descriptive Statistics**

**Steps:**

1. **Calculate Mean, Median, Mode**: Compute these statistics for Sales\_Amount and Quantity\_Sold.
2. **Interpret Results**: Analyze the statistics to understand average sales and demand.

**Code:**

python

# Calculate mean, median, and mode for 'Sales\_Amount' and 'Quantity\_Sold'

mean\_sales = data['Sales\_Amount'].mean()

median\_sales = data['Sales\_Amount'].median()

mode\_sales = data['Sales\_Amount'].mode()[0]

print(f"Mean Sales: ${mean\_sales}")

print(f"Median Sales: ${median\_sales}")

print(f"Mode Sales: ${mode\_sales}")

# Similarly for Quantity Sold

mean\_quantity = data['Quantity\_Sold'].mean()

median\_quantity = data['Quantity\_Sold'].median()

mode\_quantity = data['Quantity\_Sold'].mode()[0]

print(f"Mean Quantity Sold: {mean\_quantity}")

print(f"Median Quantity Sold: {median\_quantity}")

print(f"Mode Quantity Sold: {mode\_quantity}")

***[Attach SS here……]***

**3.4 Task 4: Visualizing Sales Trends**

**Steps:**

1. **Aggregate Sales by Date**: Sum up daily or weekly sales to observe trends over time.
2. **Plot Sales Trends**: Use line charts to visualize trends.

**Code:**

python

# Aggregate sales by date

data.set\_index('Date', inplace=True)

daily\_sales = data['Sales\_Amount'].resample('D').sum()

# Plot sales trend

plt.figure(figsize=(12, 6))

plt.plot(daily\_sales.index, daily\_sales.values, color='blue', label='Daily Sales')

plt.title('Daily Sales Trend')

plt.xlabel('Date')

plt.ylabel('Sales Amount (USD)')

plt.legend()

plt.show()

***[Attach SS here……]***

**3.5. Other Tasks …**

**4. Results**

**4.1 Descriptive Statistics Analysis**

* **Mean Sales Amount**: The average sales amount is $X (based on calculation).
* **Median Sales Amount**: The median value is $Y, showing the central tendency.
* **Mode Sales Amount**: The mode indicates the most common sales value.
* **Quantity Sold Analysis**: Similar statistics are calculated to analyze quantity trends.

***[Attach SS here……]***

**4.2 Visualization Insights**

* **Sales Trends**: The line chart shows daily sales trends, identifying peak and low sales periods.
* **Trend Patterns**: Seasonal patterns or anomalies (e.g., holiday peaks) can be observed, which are crucial for forecasting.

***[Attach SS here……]***

***4.3 Insights for other Tasks …***

***Description***

***Code***

***SS***

**5. Conclusion and Recommendations**

**5.1 Summary**

The analysis provides a clear view of daily sales trends and general patterns. Descriptive statistics offer insights into average sales and product demand. This information is essential for inventory planning and sales strategy.

**5.2 Future Steps**

* **Further Analysis**: Extend this analysis to weekly and monthly data aggregation.
* **Forecasting**: Build predictive models (e.g., ARIMA) to forecast future sales.
* **Detailed Product Analysis**: Segment data by product to understand individual performance.

**6. Appendix**

**6.1 Resources and References**

* **Documentation**: Pandas Documentation
* **Data Source**: Dataset downloaded from [Kaggle Retail Sales Dataset](https://www.kaggle.com/)

**7. Case Study Presentation/Discussion and Demo Video**

* **Video Link :**
* **Presentation /Slides File :**

Creating a comprehensive guide and rubric for a recorded video demo of a case study ensures consistency in presentation quality and evaluation standards. Here’s a step-by-step guideline for students or presenters, along with a rating rubric.

**Video Demo Guidelines for Case Study Presentation**

**1. Introduction (1-2 minutes)**

* **Objective**: Briefly introduce yourself and provide an overview of the case study topic.
* **Include**:
  + Presenter’s name(s).
  + Case study topic and industry or context.
  + The problem or objective of the case study.
  + Audience relevance (e.g., what problem it addresses for a particular sector or organization).

**2. Background and Context (2-3 minutes)**

* **Objective**: Offer essential background to help viewers understand the case study.
* **Include**:
  + Brief description of the company, organization, or context.
  + Key stakeholders and their roles.
  + Any significant historical data, market trends, or previous attempts to solve the issue.

**3. Problem Definition and Goals (1-2 minutes)**

* **Objective**: Clearly define the problem and specific goals or metrics.
* **Include**:
  + Description of the core problem.
  + Explanation of goals or objectives in quantitative or qualitative terms.
  + Mention of any assumptions, limitations, or scope boundaries.

**4. Data Analysis and Methodology (3-5 minutes) / Demo**

* **Objective**: Describe data sources, preprocessing steps, analytical techniques, and tools.
* **Include**:
  + Sources and type of data used.
  + Data preprocessing steps, including handling of missing values, outlier detection, or data transformations.
  + Methodologies applied (e.g., predictive models, visualization, descriptive analysis).
  + Explanation of why certain methods or tools were chosen (e.g., Python, R, Tableau).

**5. Solution/Analysis (5-7 minutes)**

* **Objective**: Walk through the solution or key insights derived from analysis.
* **Include**:
  + Key findings or insights from data analysis (use visualizations if applicable).
  + Demonstration of the proposed solution and any alternatives considered.
  + Challenges encountered and how they were addressed.
  + Impact analysis or effectiveness of the solution in solving the problem.

**6. Recommendations and Implementation Plan (2-3 minutes)**

* **Objective**: Present actionable recommendations and an implementation plan.
* **Include**:
  + Clear, actionable recommendations based on findings.
  + Suggested steps for implementation.
  + Timeline, resources needed, and possible challenges in implementation.

**7. Conclusion and Reflection (1-2 minutes)**

* **Objective**: Summarize findings and reflect on the project experience.
* **Include**:
  + Key takeaways or the main impact of your findings.
  + Reflection on what went well and what could be improved.
  + Any additional areas for future analysis or limitations of the current study.

**8. Presentation Quality**

* **Video Quality**: Ensure good lighting, clear visuals, and avoid background noise.
* **Slide Design**: Use a consistent theme, avoid clutter, and use visuals or charts to explain points clearly.
* **Delivery**: Speak clearly, maintain a steady pace, and avoid reading directly from slides.

**Rubrics and Scoring Guides:**

**Scoring Guide**

* **40-45 points**: Outstanding
* **30-39 points**: Very Good
* **20-29 points**: Satisfactory
* **Below 20 points**: Needs Improvement

**Rubric for Video Demo Evaluation**

| **Criteria** | **Excellent (5 points)** | **Good (4 points)** | **Average (3 points)** | **Needs Improvement (1-2 points)** |
| --- | --- | --- | --- | --- |
| **Introduction** | Clear, concise intro with strong context setting. | Clear but lacks depth. | Basic introduction with limited context. | Incomplete or missing intro. |
| **Background & Context** | In-depth, relevant, well-articulated background. | Sufficient background provided. | Basic background with minimal context. | Limited background and unclear context. |
| **Problem Definition** | Problem defined in specific, measurable terms with clear goals. | Problem defined but lacks measurable goals. | Vague problem definition and goals. | Problem unclear or goals missing. |
| **Data Analysis & Methods** | Comprehensive, relevant, and clear description of methods. | Relevant methods, minor details missing. | Basic methods, some unclear explanations. | Methods unclear or lacks coherence. |
| **Solution/Analysis** | Clear, thorough analysis with impactful insights and visuals. | Good analysis with minor issues in clarity or depth. | Basic analysis, insights lack depth. | Minimal analysis with unclear insights. |
| **Recommendations** | Actionable, practical, with implementation plan. | Recommendations provided, minor details missing. | Basic recommendations without a clear plan. | Limited or vague recommendations. |
| **Conclusion & Reflection** | Clear summary, insightful reflection, and future directions. | Good summary with minor lack of reflection or insight. | Basic summary with limited reflection or insights. | Weak summary, minimal reflection, or lacking insights. |
| **Presentation Quality** | Professional, engaging, and high-quality visuals and audio. | Good quality, minor issues in engagement or visuals. | Acceptable quality but lacks engagement. | Low-quality visuals, audio issues, or unengaging. |
| **Timing** | Adheres to time limits within 1-2 mins variance. | Slightly over or under time limit (3-4 mins). | Noticeable time variance (5+ mins over/under). | Major time discrepancy; overly lengthy or too brief. |