SNS (DATA ANALYSIS ASSIGNMENT)

INTRODUCTION

Hello, my name is **Ramya Krishnan**, and I am a Data Scientist with a passion for data analysis, machine learning, and data visualization. I have completed my Master's degree in Mathematics and have certifications in data science from prestigious institutions like IITM and IBM. Over the course of my career, I have developed expertise in tools like Python,

Tableau, and Power BI, which have enabled me to deliver data-driven insights and predictive models in various projects.

In this assessment, I have utilized my skills in data cleaning, analysis, and visualization to demonstrate proficiency in tools such as Tableau and Power BI. Each section reflects my approach to solving real-world business problems through effective data analysis and communication.

PART 1: THEORETICAL KNOWLEDGE

1. UNDERSTAING DATA VISUALIZATION:

Question: Explain the importance of data visualization in data analysis. What are the key principles of effective data visualization?

Answer: Data visualization is critical in data analysis because it transforms complex data into visual formats that are easy to interpret. It helps analysts and decision-makers identify patterns, correlations, and trends that may not be apparent in raw data. Key principles of effective data visualization include:

- **Clarity:** Visuals should clearly communicate insights without unnecessary complexity.
- **Accuracy:** Representations must reflect the underlying data accurately.
- **Simplicity:** Avoid overloading visuals with too much information.
- **Context:** Provide enough context for viewers to understand the data.
- **Storytelling:** Effective visualizations tell a story, guiding viewers through the data.
- **Consistency:** Maintain consistent colors, fonts, and formats for ease of interpretation.

2. TABLEAU BASICS

Question: What are the main components of Tableau? Describe the process of creating a basic dashboard in Tableau.

Answer: Main components of Tableau include:

- **Data Source:** Where we import data from (Excel, SQL databases, etc.).
- **Sheets:** These are individual visualizations.
- **Dashboards:** Combine multiple sheets into one interactive view.

- **Stories:** A sequence of dashboards and sheets to tell a story.

Creating a Basic Dashboard:

- **Connect to Data:** Import the dataset using Tableau's data connection options.
- **Create Visualizations:** Use the drag-and-drop interface to create charts like bar charts, line graphs, etc.
- **Create Filters:** Add filters to your visualizations for interactivity.
- **Design Dashboard:** Combine the visualizations by dragging them onto a new dashboard canvas.
- **Publish and Share:** Once satisfied with the design, publish the dashboard for sharing or interactive use.

3. POWER BI FUNDAMENTALS

Question: Discuss the main features of Power BI. How does Power BI differ from Tableau in terms of functionality and use cases?

Answer: Power BI features include:

- **Interactive Reports and Dashboards:** Create and share dynamic visualizations.
- **Natural Language Queries:** Use Q&A feature to ask questions of the data.
- **AI-Powered Analytics:** Leverage built-in AI models for deeper insights.
- **Integration with Microsoft Ecosystem:** Strong integration with Excel, Azure, and other Microsoft tools.

Power BI vs. Tableau:

- **Ease of Use:** Power BI is generally more user-friendly for beginners, especially for those familiar with Microsoft products.
- **Customization:** Tableau provides more advanced customization and flexibility in visualizations.
- **Cost:** Power BI is often more cost-effective for small organizations.
- **Performance:** Tableau handles larger datasets more efficiently than Power BI.
- **Use Case:** Power BI is ideal for business users and organizations using Microsoft services, while Tableau is preferred for advanced analytics and large-scale data analysis.

PART 2: PRACTICLE APPLICATION

4. DATA CLEANING AND PREPARATION

Problem Statement: Given a dataset with missing values and inconsistencies, clean and prepare the data for analysis.

Answer:

- **Handle Missing Values:**
  - Impute missing data with mean/median for numerical columns.
  - Use mode for categorical columns or drop rows with missing critical information.
- **Remove Duplicates:** Identify and remove duplicate rows to prevent skewed analysis.
- **Standardize Formats:** Ensure consistency in date, text, and number formats (e.g., consistent date formats).
- **Feature Engineering:** Create new variables or modify existing ones to improve analysis accuracy.

5. TABLEAU VISUALIZATION

Problem Statement: Create an interactive sales dashboard in Tableau using the provided sales dataset.

Answer:

- **Connect to Sales Dataset:** Import the dataset into Tableau.
- **Create Charts:**
  - **Total Sales:** Use a bar chart to display total sales across regions.
  - **Sales Trends Over Time:** Use a line chart to show monthly/quarterly sales trends.
  - **Sales by Region:** Use a map chart for sales by region.
- **Add Filters:** Create filters for region, time period, and product type.
- **Design Dashboard:** Arrange charts on the dashboard and include a title and interactive filters.
- **Final Review:** Check for interactivity and clarity, then publish.

6. POWER BI REPORT

Problem Statement: Develop a report in Power BI to analyze customer feedback data.

Answer:

- **Import Customer Feedback Data:** Use Power BI to import the dataset.
- **Create Visualizations:**
  - **Customer Satisfaction Levels:** Pie chart to show the distribution of satisfaction ratings.
  - **Common Issues:** Bar chart to display the frequency of issues reported by customers.
  - **Trends Over Time:** Line chart to show feedback trends over months.
- **Add Slicers:** Include slicers to filter data by region, date, and feedback category.
- **Design the Report:** Organize visualizations in a clear, cohesive layout.
- **Finalize and Publish:** Ensure interactivity and insights are clear, then publish the report.

PART 3: ADVANCED ANALYTICS

7. STATISTICAL ANALYSIS

Problem Statement: Perform a statistical analysis on a given dataset to identify significant trends and correlations.

Answer:

- **Correlation Analysis:** Compute the correlation between variables to identify relationships.
- **Hypothesis Testing:** Conduct a t-test to validate significance.
- **Results Summary:** Present key findings (e.g., "Units sold and Total sales have a strong positive correlation").
- **Visualization:** Use heatmaps to display correlations, line charts for test results.

8. PREDICTIVE ANALYTICS

Problem Statement: Build a predictive model to forecast sales for the next quarter.

Answer:

- **Model Selection:** Choose a time series model (e.g., ARIMA) for forecasting.
- **Model Building:** Train the model using Python's statsmodels or R's forecast package.
- **Evaluation:** Use metrics such as RMSE or MAE to evaluate model performance.
- **Prediction:** Forecast sales for the next quarter and present results.

PART 4: SCENARIO-BASED QUESTIONS

9. REAL-WORLD PROBLEM SOLVING

Question: How would you approach identifying key customer segments and their behaviors from a large dataset?

Answer:

- **Data Exploration:** Understand data distribution and key attributes (e.g., demographics, purchase history).
- **Segmentation:** Use clustering algorithms (e.g., K-means) to group customers into segments.
- **Analysis:** Analyze each segment's characteristics and behavior patterns.
- **Tools:** Use Python (pandas, scikit-learn) for analysis and Tableau for visualization.

10. DATA DRIVEN DECISION MAKING

Question: How would you use survey data to help a company make an informed decision about launching a new product?

Answer:

- **Data Analysis:** Analyze customer preferences and segment responses based on demographics.
- **Sentiment Analysis:** Use text analysis techniques to gauge overall sentiment towards the product.
- **Visualization:** Create visual reports (e.g., pie charts, bar graphs) to display key preferences.

- **Recommendation:** Provide insights on customer needs and recommend product features based on survey data.

LINKS

DASHBOARD LINK

Sales Date Link:
https://public.tableau.com/app/profile/ramya.krishnan.a8410/viz/SalesData_17253550967390/Dashboard1#2

Customer Feedback Data Link:

https://public.tableau.com/app/profile/ramya.krishnan.a8410/viz/CustomerFeedbackData_17253576207440/Dashboard1#2

GITHUB LINK

https://github.com/Ramya19rk/SNS-Data-Analyst-Assignment

CONCLUSION

In completing this assignment, I have showcased a comprehensive understanding of both theoretical and practical aspects of data analysis. From discussing key principles of data visualization to applying advanced analytics techniques, I have highlighted my ability to handle data-driven tasks efficiently. Through the use of Tableau, I demonstrated how data can be transformed into valuable insights that help inform decision-making processes.

The key takeaways from this assignment include:

- A thorough understanding of how to clean and prepare datasets for analysis.
- Expertise in creating dynamic, interactive dashboards and reports using Tableau.
- Competence in statistical and predictive analysis to uncover trends and forecast future outcomes.
- The ability to approach real-world problems systematically, leveraging segmentation techniques and decision-making strategies.

Overall, this assignment has reinforced my knowledge and skills, providing a platform to exhibit my proficiency as a Data Analyst, ready to contribute to meaningful data-driven projects.