Assessment for Data Analyst Role

Part 1: Theoretical knowledge

1. Understanding Data Visualization

Data visualization is important because it helps us understand data more easily. Instead of just looking at numbers and spreadsheets, we can use charts and graphs to see patterns, trends, and important insights quickly. This makes it easier to make decisions and communicate findings to others.

Key Principles of Effective Data Visualization:

1. CLARITY: First step is to make sure the visual is clearly and easy to understand.
2. RELEVANCE: Focus on the most important information that helps with decision-making. Don't include unnecessary details.
3. SIMPLICITY: Keep the design straightforward. Too many details can be confusing.
4. CONSISTENCY: Use the same colors and styles throughout so that it's easy to follow.
5. ACCESSIBILITY: Ensure that your visuals are easy to read for everyone, including those with color blindness or other visual impairments.
6. INTERACTIVITY: When possible, allow users to interact with the data (like zooming or filtering) for deeper insights.

2. Tableau Basics

Main Components of Tableau:

1. DATA SOURCE: We collect the data from various sources like spreadsheets, databases, or online services to create a dashboard.
2. WORKBOOKS: These are files where you create your visualizations. A workbook can contain multiple sheets.
3. SHEETS: These are individual pages within a workbook where you build your visualizations. There are three main types:
   - Worksheet: For creating charts, graphs, and other visualizations.
   - Dashboard: A collection of multiple visualizations combined on one screen.
   - Story: A sequence of visualizations that work together to convey a specific message or narrative.

Process of Creating a Basic Dashboard in Tableau:

1. CONNECT TO YOUR DATA: Open Tableau and connect to the data source. This could be a file like Excel or a database.
2. CREATE WORKSHEETS:
    - Go to the **Worksheet** tab.
    - Drag fields from the Data pane to the Columns and Rows shelves to create your charts.
    - Customize our visualizations using the Marks Card and Filters Shelf.
3. BUILD A DASHBOARD:
    - Click on the **Dashboard** tab to create a new dashboard.
    - Drag and drop the worksheets you created onto the dashboard area. You can arrange them as needed.
4. ADD INTERACTIVITY:
    - Use actions like filters or highlights to make our dashboard interactive. For example, clicking on one chart can filter the data shown in other charts.
5. CUSTOMIZE AND FORMAT:
    - Adjust the layout, size, and formatting of your dashboard elements to make it look polished and clear.
    - Add titles, labels, and tooltips to provide context and enhance understanding.
6. SAVE AND SHARE:
    - Save our workbook and dashboard. You can then publish it to Tableau Server or Tableau Online to share it with others.

This is the process helps to create a visual and interactive summary of our data, making it easier to analyze and communicate insights.

3. Power BI fundamentals

Main Features of Power BI:

1. DATA CONNECTIVITY: Power BI connects to a wide range of data sources, including Excel, SQL databases, cloud services like Azure, and many online services.
2. DATA TRANSFORMATION: Power BI has a feature called Power Query that lets you clean and transform data before creating reports. You can filter, merge, and shape data easily.

3. VISUALIZATIONS: Power BI provides a variety of visualizations like charts, graphs, maps, and tables to help you display your data. You can also customize these visuals.
4. DASHBOARDS: You can create interactive dashboards by combining multiple visualizations on one screen. Dashboards can be shared with others for real-time updates and insights.
5. NATURAL LANGUAGE QUERIES: You can ask questions in plain language e.g., "What are my sales this year?" and Power BI will generate the relevant visualizations based on your question.
6. INTEGRATION: Power BI integrates well with other Microsoft tools, such as Excel and Azure, and supports embedding in other applications.

Comparison with Tableau:

1. USER INTERFACE AND EASE OF USE:
    - Power BI: It is user friendly interface to use, especially for users familiar with Microsoft products. It's designed to be straightforward for business users.
    - Tableau: It is powerful and flexible interface to use, and also it offers more advanced customization options.
2. DATA VISUALIZATION:
    - Power BI: It offers a wide range of standard visualizations and it is suitable for creating detailed reports and dashboards and also custom visualizations are available in Power BI.
    - Tableau: It provides a more advance standard visualization and customization options.
3. DATA HANDLING AND PERFORMANCES:
    - Power BI: It is good for handling data and it performance can be optimized with features like DirectQuery for real-time data.
    - Tableau: It also good for handling large datasets efficiently and offers robust options for performance tuning and optimizing data interactions.

In summary, while both Power BI and Tableau are powerful tools for data visualization and analysis, Power BI tends to be more accessible and cost-effective for Microsoft-centric environments, whereas Tableau excels in advanced visualization and handling large, complex datasets.

Part 4: Scenario-Based Questions

Real-World Problem Solving:

These are the steps to identify key customer segments and their behaviors from a large dataset of customer transactions.

1. UNDERSTAND THE DATA:
    - Look at the dataset to understand what information it contains, such as transaction dates, amounts, customer demographics, and purchase categories.
    - Remove any errors or irrelevant data. Ensure that the data is accurate and consistent, such as fixing data types or filling in missing values.
2. PREPARE THE DATA:
    - Arrange the data in a way that makes analysis easier. This might involve creating new variables or aggregating data, like summarizing total spending per customer.
    - Normalize or scale data if necessary. For example, ensure that numerical values are on a similar scale to make comparisons easier.
3. ANALYZE THE DATA:
    - Use statistical methods and visualizations to explore the data. Look for patterns and trends in customer spending, frequency of purchases, and other behaviors using tools like Excel, Power BI, or Tableau for visualizations.
4. SEGMENT THE CUSTOMERS:
    - Use algorithms like K-means clustering to group customers with similar behaviors or characteristics. This technique helps identify natural clusters within the data.
5. INTERPRET THE RESULTS:
    - Create charts and dashboards to visualize the different customer segments and their behaviors. This helps in communicating your findings effectively.
6. APPLY INSIGHTS:
    - Based on these segments create market strategies and product recommendation for each group.
    - Continuously track the performance of your strategies and adjust them as needed based on new data or changes in customer behavior.

Tools Summary:

- Data Cleaning and Preparation: Excel, python (pandas library)

- Exploratory Data Analysis: Excel, Power BI, Tableau
- Segmentation: python (scikit-learn for clustering)
- Visualization and Reporting: Tableau, Power BI

Data-Driven Decision Making:

1. UNDERSTAND THE SURVEY DATA:
    - Examine the survey results to understand what questions were asked and what kind of responses were collected. Look for information on customer preferences, needs, and feedback.
2. PREPARE THE DATA:
    - Check for and correct any errors or inconsistencies in the data. This might involve handling missing values, correcting typos, or standardizing responses.
    - Arrange the data into a structured format that makes it easy to analyze.
3. ANALYZE THE DATA:
    - Identify key metrics from the survey, such as average ratings for different product features, common themes in open-ended responses, or the percentage of respondents interested in the new product.
    - Look for trends in what features or aspects of the new product are most preferred and least preferred.
    - Segment the data by some other factors to identify which groups are most interested in the product.
4. USE VISUALIZATION TOOLS:
    - Use charts and graphs to make the survey results easy to understand. Using tools like Power BI or Tableau to create these visualizations. They allow you to build interactive dashboards that can help explore the data more deeply.
5. DERIVE INSIGHTS:
    - Highlight the key insights from the analysis, such as which product features are most preferred, which customer segments show the most interest, and any common concerns or suggestions from respondents.
6. COMMUNICATE THE RESULTS:
    - Compile the findings and recommendations into a clear, concise report or presentation. Include visualizations to support your conclusions and make the data more accessible.