**ASSIGNMENT DATE-30-12-2023**

**DATA EXPLORATION AND DATA VISUALISATION:**

In Azure, you can perform data exploration and visualization using various services and tools. Below are some commonly used services and tools for these tasks:

1. **Azure Data Explorer (KustoDB):**
   * *Data Exploration:* Azure Data Explorer (ADX) is a fast and highly scalable data exploration service. It allows you to query and analyze large volumes of data in real-time. You can use Kusto Query Language (KQL) to perform ad-hoc queries and exploratory data analysis.
   * *Data Visualization:* You can visualize the results of your queries using various charting libraries and tools. ADX also supports integration with Power BI for creating interactive dashboards and reports.
2. **Azure Databricks:**
   * *Data Exploration:* Azure Databricks is an Apache Spark-based analytics platform. It provides a collaborative environment for big data analytics and machine learning. You can use notebooks for interactive data exploration and analysis.
   * *Data Visualization:* Databricks supports various visualization libraries such as Matplotlib, Seaborn, and Plotly. You can create interactive visualizations directly within the notebooks. Additionally, you can integrate Databricks with Power BI for advanced reporting.
3. **Azure Synapse Analytics (formerly SQL Data Warehouse):**
   * *Data Exploration:* Azure Synapse Analytics is a cloud-based analytics service that brings together big data and data warehousing. You can use T-SQL queries to explore and analyze data stored in Synapse SQL Pools.
   * *Data Visualization:* Synapse Studio provides a workspace for exploring and visualizing data. You can create data visualizations using built-in charts and graphs. Synapse also integrates with Power BI for more advanced visualizations and reporting.
4. **Power BI:**
   * *Data Exploration:* Power BI is a powerful business analytics tool that allows you to connect to various data sources, perform data transformations, and create data models for analysis.
   * *Data Visualization:* Power BI offers a wide range of visualization options, including charts, maps, tables, and custom visuals. You can create interactive dashboards and reports, and share them with others. Power BI can be integrated with various Azure services for seamless data connectivity.
5. **Azure Notebooks:**
   * *Data Exploration:* Azure Notebooks is a Jupyter notebook service that supports multiple programming languages. You can use it for exploratory data analysis, data manipulation, and visualization.
   * *Data Visualization:* Azure Notebooks supports various plotting libraries like Matplotlib and Plotly, allowing you to create visualizations within the notebook environment.

When working with Azure services, it's essential to choose the tool or service that best fits your specific requirements and preferences for data exploration and visualization. Additionally, many of these services can be used together to create end-to-end data workflows.

**Top of Form**

**DATA VISUALISATION STEPS:**

1. **Azure Databricks Workspace:** Make sure you have an Azure Databricks workspace set up.
2. **Sample Dataset:** You can use any dataset for this exercise. For simplicity, we'll use a sample dataset. Databricks comes with some sample datasets, and you can access them via the Databricks workspace.

**Steps:**

1. **Create a Databricks Notebook:**
   * Open your Databricks workspace.
   * Click on the "Workspace" tab in the left sidebar.
   * Navigate to the folder where you want to create the notebook.
   * Click "Create" > "Notebook."
2. **Import Sample Data:**
   * In your notebook, create a new cell and run the following command to import a sample dataset. You can choose a different dataset based on your needs.

pythonCopy code

1. **Explore the Data:**
   * Run a few cells to explore the dataset. For example, you can check the first few rows using **iris\_df.head()**.
2. **Visualize Data with Matplotlib:**
   * Create a new cell and use Matplotlib to create a simple scatter plot. For instance, plot the relationship between the sepal length and width.

pythonCopy code

1. **Enhance the Visualization:**
   * Experiment with different plot types and customizations. For example, you can create a histogram or a bar chart.
2. **Save and Share:**
   * Save your notebook after creating visualizations.
   * You can also export the notebook or share it with others in your workspace.
3. **Advanced Visualizations:**
   * If your dataset contains more complex data, consider using advanced visualization libraries like Plotly or Seaborn.

Remember, this is a basic example to get you started. In a real-world scenario, you might be working with larger datasets or integrating with other Azure services for more advanced analytics and visualizations. Explore the documentation and tutorials provided by Databricks for further guidance on more advanced features and integrations.