**Project Title: E-commerce Customer Analysis Using Azure SQL Server and Power BI**

**Objective**

Analyze customer transaction data to:

* Identify purchasing patterns
* Segment customers based on buying behavior
* Generate revenue insights
* Visualize findings in Power BI for easy, dynamic reporting

**Project Steps**

**1. Data Collection and Storage (Azure SQL Server)**

* **Database Setup on Azure SQL Server**:
  + Set up an Azure SQL Server instance.
  + Create a database for storing customer, product, and transaction data.
* **Define Table Schema**:
  + **Customers** table: customer\_id, name, age, gender, location
  + **Products** table: product\_id, product\_name, category, price
  + **Transactions** table: transaction\_id, customer\_id, product\_id, purchase\_date, quantity, total\_amount
* **Data Ingestion**:
  + Use SQL INSERT statements to load data manually or use Azure Data Factory (ADF) to automate data ingestion from CSVs or other sources.
  + Alternatively, bulk upload data directly using Azure SQL's BULK INSERT or the Azure portal.

**2. Data Extraction and Preparation (Python)**

* **Database Connection**:
  + Use pyodbc or SQLAlchemy libraries in Python to connect to Azure SQL Server.

import pandas as pd

import pyodbc

conn = pyodbc.connect("Driver={ODBC Driver 17 for SQL Server};"

"Server=your\_server.database.windows.net;"

"Database=your\_database;"

"UID=your\_username;"

"PWD=your\_password;")

# Query data from Azure SQL Server

query = "SELECT \* FROM Transactions"

transactions\_df = pd.read\_sql(query, conn)

conn.close()

* **Data Cleaning and Preparation**:
  + Use pandas in Python to clean and preprocess the data.
  + Handle missing values, format dates, and ensure correct data types.

transactions\_df['purchase\_date'] = pd.to\_datetime(transactions\_df['purchase\_date'])

transactions\_df['total\_purchase\_value'] = transactions\_df['quantity'] \* transactions\_df['price']

* **Feature Engineering**:
  + Extract time-based features for analysis, like Year, Month, and Day.

**3. Load Data into Power BI**

* **Connect Power BI to Azure SQL Server**:
  + Open Power BI Desktop.
  + Select **Get Data** > **Azure** > **Azure SQL Database**.
  + Enter your Azure SQL Server credentials and connect to the database.
  + Import the Customers, Products, and Transactions tables.
* **Data Transformation**:
  + Use Power Query within Power BI to perform any additional transformations or data cleaning.
  + Add calculated columns as needed (e.g., Total Purchase Value as Quantity \* Price).

**4. Data Analysis and Visualization (Power BI)**

* **Build a Dashboard**:
  + Use Power BI to create dynamic, interactive visualizations based on the data.

**Visualization Ideas:**

1. **Revenue Trends Over Time**:
   * Use a **Line Chart** to display monthly or quarterly revenue trends.
   * Group data by month and plot total purchase values over time.
2. **Customer Segmentation**:
   * Use **Clustered Bar Charts** or **Pie Charts** to display customer segments based on Recency, Frequency, and Monetary (RFM) scores.
3. **Top Products and Categories**:
   * Use a **Bar Chart** to visualize the top products or categories by revenue.
   * Create filters to enable interactive analysis by category, month, or other dimensions.
4. **Demographics Insights**:
   * Use **Donut Charts** or **Stacked Bar Charts** to analyze demographics like age groups, gender distribution, and geographic location.
5. **Customer Purchase Behavior**:
   * Use **Heat Maps** to show purchasing frequency by product or category for each customer segment.
6. **KPIs and Summary Metrics**:
   * Use **Cards** or **Gauge Charts** to show key metrics such as total revenue, average order value, total number of customers, and other KPIs.

**5. Publish and Share the Report**

* **Publish to Power BI Service**:
  + Once the report is created, publish it to Power BI Service.
  + Set up scheduled data refreshes to keep the data up to date. Azure SQL Server supports scheduled refresh in Power BI.
* **Create Dashboards**:
  + Pin key visualizations to a dashboard for a summary view.
  + Share the dashboard with stakeholders, giving them real-time access to insights.
* **Power BI Workspace and Sharing**:
  + Use Power BI Workspace to organize reports and dashboards.
  + Share with specific team members or groups, and manage access permissions.

**Summary of Tools and Technologies Used**

1. **Azure SQL Server**:
   * For storing structured data in a cloud-based relational database.
   * SQL used for querying and aggregating data.
2. **Python (Optional)**:
   * For any additional data extraction, transformation, and analysis tasks.
3. **Power BI**:
   * Power BI Desktop for data visualization and report creation.
   * Power BI Service for publishing, sharing, and maintaining reports.
4. **Power Query**:
   * For any in-app data transformations needed directly within Power BI.