Tuesday, September 17, 2024

Data Preparation and Segmentation Using SQLSubject

1. Gather Data:

   ○ Assume you have a **Customers** table that includes columns like:

      ▪ customer_id, age, gender, location, purchase_amount, last_purchase_date, etc.

      ▪ Assume you also have a **Sales** or **Transactions** table that includes:

      ▪ customer_id, transaction_id, transaction_date, amount_spent, etc.

2. Segmenting the Customers:

   ○ Use SQL to perform basic customer segmentation based on behavior and demographics. You can use SQL CASE WHEN statements to categorize customers into segments, such as "High Spend," "Frequent Shopper," etc.

```sql
-- Example SQL query to segment customers based on total
purchase amount and frequency
SELECT
    c.customer_id,
    c.age,
    c.gender,
    c.location,
    SUM(s.amount_spent) AS total_spent,
    COUNT(s.transaction_id) AS purchase_frequency,
    CASE
        WHEN SUM(s.amount_spent) > 1000 THEN 'High Value'
        WHEN SUM(s.amount_spent) BETWEEN 500 AND 1000 THEN
'Medium Value'
        ELSE 'Low Value'
    END AS spending_segment,
    CASE
```

```
            WHEN COUNT(s.transaction_id) > 10 THEN 'Frequent
Shopper'
            ELSE 'Infrequent Shopper'
        END AS shopping_frequency
FROM customers c
JOIN sales s ON c.customer_id = s.customer_id
GROUP BY c.customer_id, c.age, c.gender, c.location
ORDER BY total_spent DESC;
```

This query:

- Sums the total spending per customer.

- Counts the number of transactions.

- Segments customers into categories like High Value (spending > 1000), Medium Value (spending between 500 and 1000), and Low Value (spending below 500).

- Segments based on shopping frequency (more than 10 transactions is Frequent Shopper).

Step 2: Import the Data into Power BI for Visualization

1. **Import Data from SQL to Power BI:**

   o Open Power BI Desktop and click on **Home > Get Data > SQL Server**.
   o Enter the server and database details, and load your customer segmentation data into Power BI.

2. **Create Customer Segmentation Report:**

   o Use **Bar Charts** and **Pie Charts** to visualize the number of customers in each segment (e.g., High Value, Medium Value, Low Value).
   o Use **Slicers** for segmentation by age, gender, or location.
   o Visualize metrics such as **total spending**, **average spending**, and **purchase frequency**.

3. **Example of visuals**:

   o **Bar chart**: Total spending by **Spending Segment** (e.g., High, Medium, Low Value).
   o **Pie chart**: Proportion of **Frequent Shopper** vs. **Infrequent Shopper**.
   o **Line chart**: **Trends** in spending by **time** (monthly/yearly).

4. **Customize Visuals:**

o   Create **filters** to drill down by customer location, gender, or age group.
o   Add **KPI cards** to show the total revenue and average spend per segment.

Step 3: Google Sheets for Basic Analysis and Reporting

1.   Import Data into Google Sheets:

   o   You can either manually upload your segmentation results from SQL into **Google Sheets** or connect Google Sheets to Power BI (via a Google Sheets API or export Power BI visuals to CSV).

2.   Create Basic Charts in Google Sheets:

   o   You can create simple **pivot tables** and **charts** directly in Google Sheets to analyze the data:

      ▪   Use **Pivot Tables** to summarize customer data by segments (e.g., total revenue per segment, average spend).

      ▪   Use **Bar Charts** to visualize the distribution of customers in each segment.

3.   Example steps:

   o   Insert a **Pivot Table** (Data > Pivot Table), then drag **Spending Segment** into rows and **Total Spend** into values.

   o   Create **bar charts** or **pie charts** for visual representation.

4.   Customer Segmentation Reporting:

- With **Google Sheets**, you can create summary tables and pivot charts to share insights.

- For example, create a **summary table** that shows:

  - Number of customers in each **spending segment**.

  - Total and average spending in each **location**.

- Share the Google Sheets file with stakeholders or export it for reports.

Step 4: Final Dashboard and Reporting

1. **Combining Insights:**
   - The final result will be a **comprehensive dashboard** in **Power BI**, where you have interactive filters, graphs, and metrics about your customer segments.
   - Google Sheets can be used for additional analysis or sharing summary reports with stakeholders in a more accessible format.

Key Benefits of Using SQL, Power BI, and Google Sheets:

- **SQL**: Efficiently retrieves, processes, and segments large volumes of customer data from relational databases.
- **Power BI**: Provides powerful data visualization tools to make complex data accessible and actionable, ideal for executive dashboards.
- **Google Sheets**: Useful for basic reporting, quick analyses, and sharing insights with others, especially for teams that prefer working with spreadsheets.

By combining these tools, you can create a **robust customer segmentation analysis** that informs marketing strategies, product development, and customer engagement initiatives.