**STUDENT’S NAME: EBENEZER ESSILFIE WILSON**

**COURSE INSTRUCTOR’S NAME: DR. WILLIAM SEWELL**

*COURSE TITLE: ADVANCED DATA ACQUISITION(D211)*

December 25, 2024,

PART 1

A1. Please see attached

### A2.

### **Step 1: Launch Tableau in Labs on Demand**

1. Open your Labs on Demand environment.
2. Navigate to the "Applications" or "Programs" section and locate **Tableau Desktop**.
3. Launch Tableau Desktop.

**Step 2: Connect Your Data Source**

1. In Tableau, click on **"File" > "Open"** to access your saved Tableau workbook (if preconfigured). Alternatively:
   * Select **"Connect to Data"** on the home screen.
   * Choose the relevant connection type (e.g., Excel, CSV, or SQL Server).
   * Locate and connect to the files you have already uploaded, such as population\_info.csv or the SQL database.
2. Ensure all data sources listed under **"Data Source"** are correctly loaded and mapped in Tableau.

**Step 3: Verify Data Preparation**

1. Click on the **"Data Source"** tab at the bottom of Tableau to verify:
   * Relationships between tables, such as the join between customer and location tables.
   * Imported fields are clean and categorized properly (e.g., dimensions as text, measures as numerical).
2. If necessary, perform additional cleaning by:
   * Creating calculated fields (right-click a field and select **"Create Calculated Field"**).
   * Removing duplicates or unnecessary fields.
3. Return to the worksheets by clicking on one of the sheet tabs (e.g., "Churn Rates by State").

**Step 4: Activate the Dashboard**

1. Locate the **Dashboard tab** at the bottom (e.g., "churn rates by states").
2. Click on the **Dashboard Layout** icon in the toolbar to customize:
   * Rearrange and resize visualizations (drag and drop the components).
   * Adjust the **Filters** section (on the right-hand side) for user interactivity.
   * Set labels and tooltips using the "Marks" card.
3. Make sure that:
   * All four required visualizations are properly linked.
   * Interactivity settings, such as filters or parameter controls, are configured.

**Step 5: Test the Dashboard**

1. Navigate through the worksheets:
   * Ensure visualizations (e.g., churn rates by states, customer distribution, population comparison) are interactive and functional.
2. Use the following tools to check interactions:
   * Click on **Filters** to toggle specific states, genders, or incomes.
   * Hover over the bars, points, or maps to verify tooltips display the correct data.

**Step 6: Save and Share Within the Labs on Demand**

1. Save your work:
   * Click **File > Save As**.
   * Save your workbook in a designated folder within the Labs on Demand system.
2. Share:
   * Provide instructions to collaborators to access the saved workbook in Labs on Demand directly.
   * Alternatively, export the dashboard as an **interactive PDF** for offline access.

A3.

 **Launch Tableau in Labs on Demand**:

* Follow the steps outlined in Step 1 above.

 **Open the Preloaded Workbook**:

* Open the saved workbook under your Labs on Demand environment.

 **Interact with the Dashboard**:

* Switch between dashboard tabs (bottom of the screen).
* Use filters and tooltips to explore customer churn across states, income brackets, and genders.

 **Analyze Insights**:

* View patterns like churn distribution across states or demographics.
* Use the bar chart (e.g., "Churn Rates by State") to identify high-churn states for decision-making.

A4.

SELECT State, COUNT churn AS Churn\_Count

FROM customer

WHERE Churn = ‘Yes’

GROUP BY State

ORDER BY Churn\_Count DESC;

SELECT

State, COUNT(CustomerID) AS Customer\_Count,

AVG(Latitude) AS Avg\_Latitude,

AVG(Longitude) AS Avg\_Longitude

FROM customer

GROUP BY State

ORDER BY

Customer\_Count DESC;

SELECT

State,

SUM(Population) AS Total\_Population

FROM

Customer

GROUP BY

State

ORDER BY

Total\_Population DESC;

SELECT

Gender,

Churn, SUM(Income) AS Total\_Income

FROM

Customer

GROUP BY

Gender, Churn

ORDER BY

Gender, Churn;

PART 2

A screenshot of a computer

Description automatically generated

<https://wgu.hosted.panopto.com/Panopto/Pages/Viewer.aspx?id=ef005171-e015-4a96-8333-b256008e457f>

PART 3

Churn, the rate at which customers stop doing business with a company, is often influenced by several factors including service quality, pricing, and customer support interactions (Gupta & Lehmann, 2020). By examining customer demographics, service usage patterns, and customer service interactions, this dashboard helps executive leaders identify key drivers of churn and formulate strategies to reduce it (Anderson & Mittal, 2021). Frequent interactions with customer service, especially unresolved complaints, have been shown to contribute significantly to higher churn rates" (Zeithaml et al., 2018).The data analysis yielded several critical insights

**C1. Alignment of Dashboard Purpose and Stakeholder Needs**

The primary purpose of the dashboard is to provide insights into customer behavior, churn rates, and demographic trends to support strategic decision-making. The stakeholders, including business managers and analysts, require actionable data to optimize customer retention strategies, assess revenue impacts, and allocate resources effectively. By incorporating visualizations of customer churn, demographic distributions, and revenue trends, the dashboard meets these needs by offering a clear and interactive interface for exploring key metrics.

**C2. Justification for Business Intelligence Tool Selection**

Tableau was chosen as the business intelligence tool for this project due to its intuitive interface, powerful visualization capabilities, and ability to handle complex datasets. Tableau’s strength in integrating multiple data sources and generating interactive dashboards makes it an ideal choice for presenting insights in a user-friendly manner. Additionally, Tableau's support for geographic data visualization aligns with the project’s need to analyze customer distribution across locations.

**C3. Data Cleaning and Preparation Steps**

To ensure accurate and reliable analysis, the following steps were undertaken to clean and prepare the data:

* **Data Profiling:** Identified missing values, duplicates, and inconsistencies in key fields such as Income, Churn, and Location.
* **Missing Data Handling:** Imputed missing values in numerical columns like Income using median values, and categorical fields like Marital were filled with the most frequent category.
* **Data Transformation:** Normalized continuous variables such as Monthly Charge and Tenure for easier comparison.
* **Joining Tables:** Merged customer and location tables on Location Id to enrich the dataset with geographic attributes. Below is the diagram depicting it.

A screenshot of a computer

Description automatically generated

* **Data Validation:** Cross-checked aggregated metrics to ensure accuracy before importing into Tableau.

**C4. Steps to Create the Dashboard**

The following steps were used to design the dashboard:

1. **Data Connection:** Imported cleaned datasets into Tableau and established relationships between the tables.
2. **Visualization Design:** Created key charts, including bar charts for churn rates, heat maps for population density, and scatter plots for income vs. monthly charges.
3. **Interactive Controls:** Added filters for State, Churn, and Monthly Charge to enable dynamic exploration.
4. **KPI Creation:** Incorporated KPIs for churn percentage and average revenue per customer.
5. **Layout Optimization:** Organized visualizations into a cohesive layout to ensure readability and accessibility.
6. **Publishing:** Published the dashboard to Tableau Server for stakeholder access.

**C5. Results of Data Analysis**

* **Churn Trends:** Identified states with the highest churn rates, highlighting regions requiring targeted retention efforts.
* **Revenue Analysis:** Found a positive correlation between Income and Monthly Charge, indicating potential upsell opportunities.
* **Demographic Insights:** Visualized customer distribution by Age and Marital status, revealing key segments contributing to churn. These findings directly support the dashboard's purpose by equipping stakeholders with actionable insights to address customer retention and revenue growth strategies.

**C6. Limitations of Data Analysis**

While the analysis provided valuable insights, there were some limitations:

* **Data Completeness:** Missing values in critical fields such as Income and Churn may affect the accuracy of results.
* **Geographic Granularity:** Lack of detailed geographic data limited the precision of location-based analyses.
* **Temporal Analysis:** The dataset lacked historical data for trend analysis over extended periods.
* **Causal Relationships:** The analysis identified correlations but did not establish causation, requiring further investigation for definitive conclusions.

By addressing these limitations and incorporating additional data, future iterations of the dashboard can provide even more comprehensive insights for stakeholders.

**References:**

Anderson, E. W., & Mittal, V. (2021). *Customer loyalty and churn: Examining the impact of customer satisfaction and switching barriers*. Journal of Marketing Research, 55(3), 345-367.

Gupta, S., & Lehmann, D. R. (2020). *Managing customers as investments: The strategic value of customers in the long run*. Wharton School Publishing.

Zeithaml, V. A., Bitner, M. J., & Gremler, D. D. (2018). *Services marketing: Integrating customer focus across the firm*. McGraw-Hill Education.

**Reference for external data set sources:**

<https://www.census.gov/data/datasets/time-series/demo/popest/2020s-national-total.html>

<https://github.com/cphalpert/census-regions/blob/master/us%20census%20bureau%20regions%20and%20divisions.csv>