PERFORMANCE ASSESSMENT

Task 1 | Data Analysis

KAILI HAMILTON

Masters of Science in Data Analytics, Western Governors University

Course: D211

Instructor: Dr. David Gagner

Program Mentor: Krissy Bryant

January, 2024

TABLE OF CONTENTS

A1 PART 1

Two datasets are provided with the submission of this performance assessment. The churn data set was selected and used from the D211 data dictionary. The additional dataset contains churn information about a competing telecommunications company and was obtained from Kaggle and is open source.

A2 PART 1

Here are the instructors to access the churn data analysis dashboard.

1. Turn on your computer and make sure you are connected to the internet.
2. Click on the link below:
   https://public.tableau.com/views/CustomerDemographicsDashboard_17061571909330/CustomerDemographicsDashboard?:language=en-US&:display_count=n&:origin=viz_share_link
3. Maximize your browser window.
4. Read the instructions (found in part A3).
5. Engage with the dynamic controls of the dashboard to explore and discover insights into customer churn!

A3 PART 1

Here are instructions provided to users to help them navigate the churn analysis dashboard.

Welcome to the customer demographics dashboard!

- The dashboard is best viewed with a full screen, so maximize your browser window.
- The dashboard contains some visualizations, key performance indicators, and interactive controls (select and apply filters). You can dynamically interact with the graphs, the KPIs, and the filters to gain insights into customer demographics and make business decisions based on the data.
- To interact with the dashboard, just click on a colored block or number in each visualization or click in a KPI metric number. This will automatically filter the dashboard according to your selection.
- For example, click on the state of Texas in the map. You will see how the rest of the dashboard filters to show information for customers in Texas. Done with this selection? Hit "esc" on your keyboard to deselect your filter. Multiple selections can be made, meaning you can click on Texas and Electronic Check payment type and the dashboard filters accordingly.
- Alternatively, you can select and apply the filters using the check boxes on the righthand side of the dashboard.
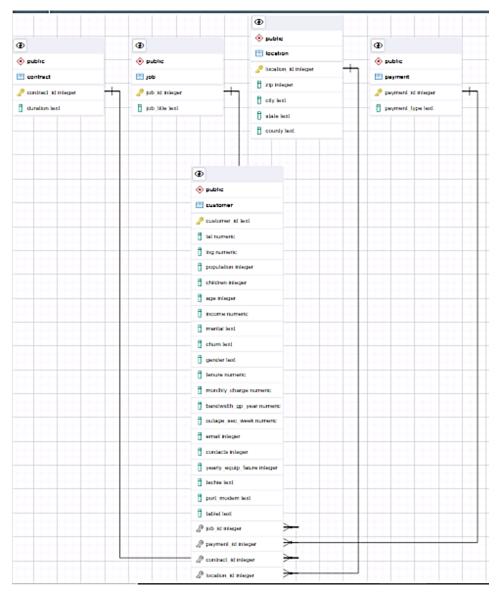
- Hover the mouse over a part of the visualization and you'll see a tooltip appear. This tooltip provides additional insight about a given section. For example, hover the mouse over Texas and you will see that Texas ranks first in the number of customers as well its the churn rate.

A4   PART 1

I provide a copy of all SQL code that supported the creation of the dashboard with the submission of this performance assessment. I also provide a screen shot of that code here.

```sql
WITH churn_table AS (
SELECT *
    FROM customer
    LEFT JOIN contract
        ON customer.contract_id = contract.contract_id
    LEFT JOIN job
        ON customer.job_id = job.job_id
    LEFT JOIN location
        ON customer.location_id = location.location_id
    LEFT JOIN payment
        ON customer.payment_id = payment.payment_id
)
SELECT
    customer_id
    , lat
    , lng
    , population
    , children
    , age
    , income
    , marital
    , churn
    , gender
    , tenure
    , monthly_charge
    , bandwidth_gp_year
    , outage_sec_week
    , email
    , contacts
    , yearly_equip_faiure AS yearly_equip_failure
    , techie
    , port_modem
    , tablet
    , duration AS contract_length
    , job_title
    , zip
    , city
    , state
    , county
    , payment_type
FROM churn_table
```

Table: customer (public)

General | Columns

Columns | +

		Name	Data type	Length/Precision	Scale	Not NULL?	Primary key?
☑ 🗑		customer_id	text			Yes	Yes
☑ 🗑		lat	numeric			No	No
☑ 🗑		lng	numeric			No	No
☑ 🗑		population	integer			No	No
☑ 🗑		children	integer			No	No

× Cancel | 🖫 OK

Table: contract (public)

General | Columns

Columns | +

		Name	Data type	Length/Precision	Scale	Not NULL?	Primary key?
☑ 🗑		contract_id	integer			Yes	Yes
☑ 🗑		duration	text			No	No

× Cancel | 🖫 OK

Table: job (public)

General | Columns

Columns | +

		Name	Data type	Length/Precision	Scale	Not NULL?	Primary key?
☑ 🗑		job_id	integer			Yes	Yes
☑ 🗑		job_title	text			No	No

× Cancel | 🖫 OK

Table: location (public)

General | Columns

Columns | +

		Name	Data type	Length/Precision	Scale	Not NULL?	Primary key?
☑ 🗑		location_id	integer			Yes	Yes
☑ 🗑		zip	integer			No	No
☑ 🗑		city	text			No	No
☑ 🗑		state	text			No	No
☑ 🗑		county	text			No	No

× Cancel | 🖫 OK

Table: payment (public)

General | Columns

Columns | +

		Name	Data type	Length/Precision	Scale	Not NULL?	Primary key?
☑ 🗑		payment_id	integer			Yes	Yes
☑ 🗑		payment_type	text			No	No

× Cancel | 🖫 OK

churn_tab... — Additional... ▼

<

How do relationships differ from joins? Learn more

churn_table_joined....	Operator	Additional Dataset (...
Abc Churn ▼	= ▼	Abc Churn1 ▼
Abc Contract Length ▼	= ▼	Abc Contract ▼

⊕ Add more fields

> Performance Options

churn_table_joined.csv ——————— Additional Dataset (Kaggl...

B1-7 PART 2

A link to the Panopto video presentation is provided with the submission of this performance assessment.

C1 PART 3

The purpose and function of the dashboard aligns with the business needs of the telecommunications company by providing relevant information about the demographics of their customer base. They can use the data presented in the dashboard to support business decisions such as where and how to direct their marketing efforts.

C2 PART 3

Tableau is the business intelligence tool I used to create the customer demographics dashboard. "As the market-leading choice for modern business intelligence, the Tableau platform is known for taking any kind of data from almost any system, and turning it into actionable insights with speed and ease. It's as simple as dragging and dropping. Plus, our industry-leading enablement resources, training, and global data community offer unparalleled support for our customers and their analytics investments. And on our mission to help people see and understand data, we go beyond our technology to ensure customer success by helping people build a data culture" (Tableau, n.d.).

C3 PART 3

The steps used to the clean and prepare the data for analysis are verifying the ERD was set up appropriately and understanding the primary and foreign keys for each table in the churn database (customer, location, job, contract, and payment). Then each table was analyzed for null values and duplicate rows and had uniform formatting throughout. Also, the data types of each variable were set to their proper type. Once each table was cleaned, they were joined as described in part C4.

C4 PART 3

Here I summarize the steps I took to create the customer demographics dashboard.

There are five tables in the churn database that were needed to create the dashboard. They are customer, location, payment, job, and contract. In pgAdmin, I joined the five tables together via a left

join on their respective keys, saved it as a common table expression, selected the relevant variables for analysis (see the SQL code in part A4), and exported that as flat file as a .csv.

Now that I had my data source, I accessed Tableu. I connected to this data source, called churn_joined as well as the additional dataset from Kaggle for a competing telecommunications company. Next, I defined a relationship between additional data set and the churn_joined data set (Tableau, n.d.).

The visuals in my dashboard use all of the tables joined together in pgAdmin. A map is used to display customers by location. A histogram is used to display customers by age. A pie chart displays customers by churn. A bar graph displays customers by payment type. And customers by job is displayed as a list of jobs and their churn rate with a filter to display the jobs with the highest or lowest churn rates.

Every visualization and metric in the dashboard can also be used as a filter. This makes the experience for stakeholders more interactive when they explore the demographics of the customers.

C5   PART 3

The results of my data analysis supports executive decision making because we get to visualize the demographics of the customers. Executive can use the data to support business decisions and improve their business model.

One interesting result is that we can see that the distribution of our customers is uniform across 20-90 year olds, so gearing a marketing campaign towards a certain age group would be futile. Additionally, we can filter the customers by location map to the top 5 states by number of customers and focus on customer retention on these locations since they provide the most customers. Also, if we filter the customer by jobs list to those with churn rates at 10% or less, we can direct customer retention efforts to customers with those types of jobs—perhaps by virtual of their profession they are using the telecommunications services more regularly than those not in those professions.

C6   PART 3

One limitation of the data analysis is that the data we have here is for customers who churned in the last month. We could gain more insight if we could compare trends in churn rates over time to prevent customers from churning. Additionally, we don't have information as to the quality of the products and services that this company offers. This would provide excellent insight as to why a customer would choose to stay with this company and help our retention efforts. Also, there could be other variables that are not used in the dashboard that can provide a more in-depth look at the customer base.

D-E

"How Relationships Differ from Joins." Tableau, n.d.,
　　help.tableau.com/current/pro/desktop/en-
　　us/datasource_relationships_learnmorepage.htm?source=productlink

Semaan, Michel et al., "D211 – Advanced Data Acquisition." Datacamp, 2019,
　　app.datacamp.com/learn/custom-tracks/custom-d211-advanced-data-acquisition

"Telo Customer Churn." Kaggle, 2018, www.kaggle.com/datasets/blastchar/telco-customer-churn.

"Why choose Tableau?" Tableau, n.d.,
　　tableau.com/why-tableau

F

Professional communication is demonstrated in the content and presentation of my performance assessment.