**QUEST FOR SOLVING TELECOMMUNICATION COMPANY CHURN RATE**

**ADVANCED DATA ACQUISITION — D211**

**PRFA — SLM1**

**STUDENT NAME: IBRAHIM SULEIMAN**

**DATA ANALYTICS**

**STUDENT NUMBER: 001429984**

**COLLEGE OF INFORMATION TECHNOLOGY, WESTERN GOVERNORS UNIVERSITY**

**NAME OF INSTRUCTOR: DAVID GAGNER**

**OCTOBER 25th, 2022**

**Part 1: Interactive Data Dashboard**

Section A1: Data Sets [Churn Data and Dictionary Files](https://access.wgu.edu/ASP3/aap/content/f9tjr8djg83jd8c3sdf8.zip) (Churn\_clean.csv) : provided data set located on labs on demand

<https://tasks.wgu.edu/student/001429984/course/23840006/task/3279/overview>

## Additional data set for the dashboard

## Telco-Customer-Churn.csv

<https://www.kaggle.com/code/bhartiprasad17/customer-churn-prediction/data>

Section A2: Installation Instructions

To assist users in installing the dashboard, you can offer detailed instructions. It is not necessary to install this. Observe this link: <https://public.tableau.com/shared/RT325YPKR?:display_count=n&:origin=viz_share_link>

Section A3: Navigation Instructions

For both the telco customer churn data (new) and the churn cleaned data (old), the top data displays the average number of telecom company customers that churned after using their subscriptions to internet service, with the churn difference being used to color the data. Change the "Colorblind" dropdown menu's selection to Yes or No to improve your ability to see the colors.

The arrow on the top right, when clicked, will navigate the user to the next story of the dashboard, and the top left arrow will navigate the user to the previous dashboard story.

**Payment type and phone service**

The Sum of Multiple Lines for each Internet Service. The color shows details about the phone. size shows details about the payment type. The data is filtered on action (payment, payment type), which keeps 1 member. The view is filtered on payment type: bank transfer automatic, credit card automatic, electronic check, and mailed check.

**Internet service and multiple lines**

Secondly, customers with "DSL" and "Optic Fiber" internet service appear to have a high number of multiple lines, whereas those with a "0" appear to have fewer multiple lines. Customers could be persuaded to sign up for DSL as a result, which would lower the churn rate as the color changes if the colorblind is clicked once the user has passed attention while navigating the dashboards.

Section A4: Copy of all SQL code and other codes supporting the dashboards

|  |
| --- |
| SQL CODE FOR ERD-  BEGIN;  CREATE TABLE public.contract  (      contract\_id integer NOT NULL,      duration text,      PRIMARY KEY (contract\_id)  );  CREATE TABLE public.customer  (      customer\_id text NOT NULL,      lat numeric,      lng numeric,      population integer,      children integer,      age integer,      income numeric,      marital text,      churn text,      gender text,      tenure numeric,      monthly\_charge numeric,      bandwidth\_gp\_year numeric,      outage\_sec\_week numeric,      email integer,      contacts integer,      yearly\_equip\_faiure integer,      techie text,      port\_modem text,      tablet text,      job\_id integer,      payment\_id integer,      contract\_id integer,      location\_id integer,      PRIMARY KEY (customer\_id)  );  CREATE TABLE public.job  (      job\_id integer NOT NULL,      job\_title text,      PRIMARY KEY (job\_id)  );  CREATE TABLE public.location  (      location\_id integer NOT NULL,      zip integer,      city text,      state text,      county text,      PRIMARY KEY (location\_id)  );  CREATE TABLE public.payment  (      payment\_id integer NOT NULL,      payment\_type text,      PRIMARY KEY (payment\_id)  );  ALTER TABLE public.customer      ADD FOREIGN KEY (contract\_id)      REFERENCES public.contract (contract\_id)      NOT VALID;  ALTER TABLE public.customer      ADD FOREIGN KEY (job\_id)      REFERENCES public.job (job\_id)      NOT VALID;  ALTER TABLE public.customer      ADD FOREIGN KEY (location\_id)      REFERENCES public.location (location\_id)      NOT VALID;  ALTER TABLE public.customer      ADD FOREIGN KEY (payment\_id)      REFERENCES public.payment (payment\_id)      NOT VALID;  END;  SQL SCRIPT/QUERY FOR CHURN DATASET ALREADY PROVIDED IN pgAdmin4 -  -- Database: churn  -- DROP DATABASE churn;  CREATE DATABASE churn      WITH      OWNER = postgres      ENCODING = 'UTF8'      LC\_COLLATE = 'English\_United States.1252'      LC\_CTYPE = 'English\_United States.1252'      TABLESPACE = pg\_default      CONNECTION LIMIT = -1;  SQL SCRIPT/QUERY FOR EXTERNAL DATASET  -- Table: public.cleaned\_telco\_customer\_churn  -- DROP TABLE public.cleaned\_telco\_customer\_churn;  CREATE TABLE public.cleaned\_telco\_customer\_churn  (      gender text COLLATE pg\_catalog."default",      seniorcitizen integer,      partner integer,      dependents integer,      tenure integer,      phoneservice integer,      multiplelines integer,      internetservice text COLLATE pg\_catalog."default",      onlinesecurity character varying COLLATE pg\_catalog."default",      onlinebackup character varying COLLATE pg\_catalog."default",      deviceprotection character varying COLLATE pg\_catalog."default",      techsupport character varying COLLATE pg\_catalog."default",      streamingtv character varying COLLATE pg\_catalog."default",      streamingmovies character varying COLLATE pg\_catalog."default",      contract integer,      paperlessbilling integer,      paymentmethod integer,      monthlycharge integer,      totalcharges integer,      churn integer  )  TABLESPACE pg\_default;  ALTER TABLE public.cleaned\_telco\_customer\_churn      OWNER to postgres;  select \* from PUBLIC.Cleaned\_Telco\_customer\_churn  COPY PUBLIC.Cleaned\_Telco\_customer\_churn FROM 'C:\Users\LabUser\Desktop\Cleaned\_Telco\_customer\_churn.csv' WITH CSV HEADER;  select \* from Cleaned\_Telco\_customer\_churn |

The ERD tables were created using the SQL codes above

Graphical user interface

Description automatically generated

THE CREATED ON TABLEAU CUSTOM SQL QUERY:

Graphical user interface, text

Description automatically generated

Graphical user interface, application

Description automatically generated

Part II: Panopto Storytelling with Data

Section B: Panopto Demonstration

Advanced Data Acquisition SLM1 | D211.  
  
You can view the session using the following link:  
<https://wgu.hosted.panopto.com/Panopto/Pages/Viewer.aspx?id=71c468e7-f6d7-4771-bef7-af3b013ff552>

Part III: Reflection Paper

Section C1: Dashboard Alignment

The problem of client churn in telecommunications firms is getting worse and worse. A consumer leaving a business is referred to as churning. Customers can select from a variety of service providers in the telecommunications sector and proactively switch from one provider to another. Customer "churn" is the proportion of customers who discontinued utilizing a provider's good or service over a specific period of time.

In this very competitive industry, some telecommunication industries might have average yearly churn rates of up to 25%. Customer retention has now surpassed customer acquisition in importance due to the fact that it is 10 times more expensive to gain new customers than to keep existing ones. Since this is a problem for many telecoms firms, external firms appear to have an excessive amount of turnover annually.

Knowing which consumers are most likely to cancel their contracts within specific time frames would help telecommunications firms plan for and/or avoid the cancellation by providing pre-service customer acquisition and retention on such customers. In order to better understand the data, I decided to concentrate on the churn condition.

The dashboard's goal is to identify those who are most likely to experience and suffer from churn. An intervention for prevention can be used with various consumers at various stages of contracts if we and the stakeholders are aware of who is most likely to churn.

Section C2: Justification of the Selection of the Business Intelligence Tool Used

PostgreSQL

Quick access to data records is made possible using SQL. SQL can offer the data you need with as few as one or two instructions if you're looking for a quick way to locate it.

SQL can also organize and store enormous volumes of data before it is even presented on your screen. Any database application that often processes massive amounts of data can be made more performant with the help of SQL.

Because SQL is standardized, data may be extracted from any kind of database, regardless of the creator or the platform, using it. Because of this, businesses and organizations of all sizes find it to be a very useful tool.

The importance of SQL in data administration can be attributed to a number of factors, just to name a few. When used properly, it can help you save time and trouble while working with large amounts of data. Use SQL in order to maximize the value of your data if you're seeking a practical method. You won't regret doing so!

TABLEAU-

Tableau makes it possible to evaluate data more rapidly and produces visualizations as dashboards and worksheets. Tableau products are always used in virtualized settings. With the help of Tableau, we can create dashboards that provide actionable information and accelerate business growth.

Amazing images: Tableau's built-in functionality allows you to work with more unordered data and create a variety of graphics. Additionally, you will be able to swiftly study the data using a wide range of drilling techniques and excellent context.

Without any long-term goals in mind, Tableau helps businesses analyze upcoming data. You will be able to examine data and study graphics using a variety of techniques. With the aid of hypothetical visualizations, you may create "what-if" queries and work on the data accordingly.

You may integrate many data sources, data warehouses, cloud storage, big data, and other forms of data with Tableau. Tableau can combine many types of data to assist enterprises in creating eye-catching representations. Tableau is more potent than business intelligence and analytics products thanks to this functionality.

Section C3: Data Preparation

Prior to completing the analysis, the data must be ready. Making sure that none of the columns have any missing data is the initial step. The next thing we should do is make sure there are no duplicates of any of the data in the columns. Additionally, we want to confirm that no columns or rows are duplicated, so we will check that and make sure the outcome is (False).

The dataset contains a number of variables that were determined to be useless for the logistic analysis, such as customer demographics that cannot be altered and are linked to the interaction and location of the consumer, so those columns should be removed.

Working with the data is now easier as a result. Any (yes or no) or other categorical alternatives need to be converted to numerical values in order to convert the categorical variables to numerical values. The survey columns also require renaming to provide a clearer comprehension and determination of applicable factors (Peter Grant, 2019).

The categorical target variable will be "churn," as the objective is to determine the factors that impact whether a client will be lost (Yes or No). The initial set of categorical variables will include the following: Online Security, Online Backup, Techie, Contract, Tablet, Marital, Gender, Multiple, Device Protection, Internet Service, Phone, and Tech Support.

The following is a list of the initial continuous variables to be used: Children, age, income, email, contacts, yearly equipment failure, population, and monthly charges should all be taken into account.

Section C4: Interactive Controls

The average values filter for the first dashboard in the story is one of the interactive controls. This enables the consumer to determine which streaming movie selections and which streaming movies have high churn rates. For instance, if the viewer clicks on the bar that says "0," they will see a variety of alternative streaming movie churn ranges. The user will then see the three streaming movie categories if they double-click on the bar.

On the dashboard of the tale, I have given the option to focus on marital status based on gender and income as well as who has yearly equipment failure. Another interactive control I've provided is this one. I did this by showing a single value (list) or several values (list) on the top right of the dashboard to summarize the total, median, average, minimum, maximum income, and yearly equipment failure. I needed the option to be able to only view those particular statistics without the cluster of those without customer turnover because the focus of the tale was on that topic.

Color Blindness

I was able to locate a YouTube video on how to switch between colorblind and colorful dashboards that I used for the first two designs. This demonstrated to me how to offer consumers the choice between a vibrant color palette and a palette that is accessible to colorblind users. As a result, it gives users additional interactive control and makes it available to everyone.

Because it required a variety of colors to match the variety of values, the third graphic was more difficult to create. We selected a palette without a "colorblind" palette that contrasted dark and light colors. In order to help with colors that are quite close to one another on the scale, I have also included percentages in each box.

Since the fourth image focuses on the data's five-number summary, it doesn't actually need a lot of color. To add a little bit, I did use Tableau's "colorblind" palette to display the various streamed movies and churn rate groups.

A summary for dashboard creation Steps;

* Firstly, we connected a Churn dataset and an external dataset to the Tableau tool. After opening it, drag and drop the single connected dataset sheet. To open the tableau worksheet, click sheet 1. When you click Sheet1, you will see a worksheet for work and the entire dataset's properties on the left.
* We generated multiple worksheets with various information and features, such as crosstabs, maps, charts, and graphs, in order to construct a dashboard. To do this, you must take the following actions:
* prepared with the sheets and dashboard work.
* To open a new dashboard, click the plus sign in the square.
* You can drag and drop the sheets however you like.
* Add some amendments like colorblind (YES/NO), which could be optional.

Section C5: Decision-Making Support

The quantity of churn on the internet revealed in the dashboard reveals that most users are rapidly switching between various types of internet services, which is a positive sign. Churn is a difficult, life-challenging circumstance that could occur and influence churn decisions. We can modify the data on the dashboard to just show the users with a high churn rate.

We can observe that the bar for bank transfer automatic is roughly higher than "1600" and the credit card automatic is less than 1600 in relation to the count of phone services. Therefore, we recommend promoting the use of automatic credit card payments and providing prompt assistance when there is equipment breakdown. This can aid senior leaders in making judgments on how to increase client involvement.

Graphical user interface

Description automatically generated

Secondly, customers with "DSL" and "Optic Fiber" internet service appear to have a high number of multiple lines, whereas those with "0" appear to have fewer multiple lines. Customers could be persuaded to sign up for DSL as a result, which would lower the churn rate.

The percentiles for the marital status pie charts are shown below.

Graphical user interface, application

Description automatically generated

The graph below compares marital status and income in terms of how much money each client received along with their chosen method of payment.Graphical user interface, application

Description automatically generated

The dashboard image below shows the count of each gender of senior citizens. in relation to their internet service subscriptions and contracts. It seems that fiber optics has the highest contracts compared to DSL and male senior citizens have a higher count compared to female senior citizens. In order to preserve a long-term client relationship, the objective is to encourage long-term internet service contracts..Chart, line chart

Description automatically generated

However, the population of mailed checks is higher and the population of electronic checks is lower, and the payment type picture provides us with information on the average population of customers based on the payment types of the customer.

Chart, bar chart

Description automatically generated

The Average of Income for each Marital. Color shows details about Marital. The view is filtered on average of Income, which ranges from 39,193.0 to 39,864.0.

Section C6: Limitation of the data analysis

The problems with using external data are brought on by the challenges in comparison brought on by differences in the amount of granularity of data inside and between databases. conflicting or incomplete metadata.

However, there are still a lot of issues with Tableau:

Tableau primarily focuses on visualization and can only work with cleaned data. Before using Tableau efficiently, the underlying database must first undergo sufficient data cleaning.

Data models and data dictionaries include holes that data analysts must fill. This necessitates separate maintenance of our metric definitions elsewhere. Version control and collaboration were lacking when creating data logic and dashboards.

Despite the fact that SQL offers many advantages, there are a few disadvantages. When dealing with the database, few people find SQL's difficult user interface enjoyable. Only programmers can use some versions since they are costly. A concealed set of business rules prevents the database from having complete control.

Reflection Paper

* Audience Analysis

I made the decision to base my story on a typical incident because my audience consists of employees and stakeholders of telecommunication companies. When customers decide to change their subscriptions or memberships, they frequently research their problems, wishes, and possibilities online before becoming anxious about improbable events.

This is demonstrated by the character, who, after enduring yearly equipment failures, looks up her monthly bill on her phone.

Even though churn is something that happens more frequently, the majority of the target audience can connect to the situation and (hopefully) laugh about it. Customers who believe they are entitled to specific perks, savings, and discounts and are confident they will be charged more in the future frequently patronize telecommunications companies, despite the fact that monthly fees and yearly equipment failures are common.

The character in the story tells the customer support agent that she thinks she has excessive monthly rates rather than a more typical telecommunications issue in the winter.

* Universal Access

I made several different attempts to tailor my presentation to a variety of audiences. My tale was first published on Tableau public, where anyone with an internet connection can read it for free. I also kept the number of graphs or images on each dashboard to a manageable level. This makes it easier to read the offered information without being distracted or confused.

To ensure that everyone who can see may comfortably read the text, I chose appropriate colors and provided choices for color blindness. I avoided utilizing complex graphs and statistical analysis techniques because the visualizations should be accessible to all users, regardless of their level of expertise. It was intended to be simple to read and understand for everyone.

Effective Storytelling

In his book Storytelling with Data from 2015, Knafilic writes on page 150, "A good tale grabs your attention and takes you on a journey, generating an emotional response. That was my first instance of using storytelling. No matter who you are, you've all looked up your symptoms online and become alarmed.

In an effort to grab their attention, I made it relatable to my audience. After hearing the tale, I wanted them to reflect on: "Oh! That has been done by others besides just me. " In his book Storytelling with Data from 2015, Knafilic writes on page 154, "The first thing to do is present the plot, building the backdrop for your audience."

I did my best to do this by asking the audience to visualize a price increase or an inclination in monthly charges. I ask them to come up with a variety of difficult scenarios that will cost them barely enough to subscribe to my newsletter. Everyone has encountered a rise in the price of goods and services as well as a telecommunications-related equipment malfunction.

A consumer can undoubtedly picture the difficulties, even if they have never experienced a yearly equipment breakdown. I want to make it interesting and relatable so that the audience will want to keep reading to find out the outcome.

Section D: Sources

Singghak S. (2021). *Advantages and Disadvantages of SQL.*

[**https://www.geeksforgeeks.org/advantages-and-disadvantages-of-sql/**](https://www.geeksforgeeks.org/advantages-and-disadvantages-of-sql/)

*How to Swap Between Colorblind and Colorful Dashboards*. (2016, July 5). YouTube.

<https://www.youtube.com/watch?v=3iNl7KMK8pM>

Knaflic, C. N. (2015). *Storytelling with data: A data visualization guide for business professionals*. Wiley .(Chapter 1, Chapters 3-5, Chapters 7 - 8).

<https://ebookcentral.proquest.com/lib/westerngovernorsebooks/reader.action?docID=4187267&ppg=1>

Story Transcript

Hello, Ibrahim Suleiman is my name. I work as a data analyst for the Board of Health in Chamblee, Georgia's Dekalb County. Since I've been working in the logistics and business sectors for a few years, I've also completed a few projects involving customer retention and telecom data. I've selected the first dataset from the supplied information. The churn clean csv file, which I'll send together with my other docs, displays a wide range of various telecommunication and churn data provided by a telecommunication business about its clients.

I've decided to pay particular attention to turnover, marriage status, contracts, age, dependability, and income. I discovered the second dataset I've chosen on Kaggle. The Telco-Customer-Churn.csv data, which will also be supplied with my other documents, displays a variety of data from a telecom firm on its customers, including information on their yearly equipment failure, internet service, churn rate, and income. I've decided to pay particular attention to gender, phone, internet, and technical support.

Although I have not directly battled with churn, I have grown up witnessing my parents suffer with it, thus that is why I chose these things. It was difficult to watch Dad make difficult choices, and I was always scared that someday I would have to do the same. Therefore, I was interested in the patterns and how likely it was that I could deal with them.

Imagine being charged quickly each month for a subscription. You're strapped for cash, unsteady, and experiencing mild anxiousness. When you have mounting debt and routine equipment breakdowns, it can seem as though the world is spinning out of control. You are aware that something is wrong with you and that you should call the telephone provider. You rise and put your phone in its place. You are not confident in the customer service's ability to respond quickly while it is snowing outdoors.

So, in the midst of a snowstorm, you call your telecommunications provider, which is located approximately 20 miles away, to ask them to take care of your equipment failure since you aren't in a good position to drive yourself.

You make the decision to research the most cost-effective and cutting-edge telecommunications provider for a lower rate as you sit in your living room. The following two sources demonstrate that people with two-year contracts experience fewer equipment problems, a quicker response, and a reduced rate. The telecommunications firm phoned you back when you eventually restored your internet service, asked what the problem was, and offered you a two-year contract at a lower rate with new equipment to make up for the yearly equipment failure. Due to new developments and efforts to preserve positive customer relations, you will receive an annual renewal credit.