CGT 270 Data Visualization
Module 1 ● Week 2
Lab 2: Parsing Data

Name: *Gabby Willard*

The goal of this lab is to understand the structure of data. In this lab you will change data into a format that tags each part of the data with its intended use. After completing this lab every element of the data, you selected (Tableau dataset) and the two (2) additional datasets you acquired in lab last week will be broken into its individual parts.

What you should be able to do (at the end of this lab):

Remember	*Describe* what happens in the **parse** stage.
Understand	*Describe* the data in detail according to the parsing specifications.
Apply	*Demonstrate* the ability to change data into a useful format for future processing.
Evaluate	*Categorize* the data according to parsing specs.
Analysis	*Identify* specific features about the data.
Create	*Generate* a parsed listing of the data.

Part I: Tableau Data Set
In this document:
Answer the following questions about your Tableau data:

1. List the name of the Tableau Dataset you selected in the Acquire Lab:
   Titanic Passenger List
2. How many rows (records) are in the data set?
   1,309 rows are in the data set.
3. How many columns (variables) are in the data set?
   14 columns are in the data set.
4. What assumptions are you making about the data?
   For the survived variable, the only values are either 0 or 1 for each record. I am unsure of which value means survived and which value means didn't survive. I am also unsure of what the integer corresponding to the body variable actually means. The embarked variable consists of the characters S and C which I am not sure what those two characters represent.

Use the following link to provide a parsing of the Tableau Data.
https://tinyurl.com/Parse-Worksheet

For Visualization Assignment select: **Training Data**
Then provide the parsed data and list your assumptions.

Part II: Additional Data Set #1
In this document:
Answer the following questions about your Tableau data:

1. List the name of the Tableau Dataset you selected in the Acquire Lab:
   Titanic Passenger Nationalities

2. How many rows (records) are in the data set?
    418 rows are in the data set.
3. How many columns (variables) are in the data set?
    2 columns are in the data set.
4. What assumptions are you making about the data?
    I do not know what each of the passenger's names are because the Passenger variable only consists of integers. Each passenger is assigned to a number but there is no way of telling any names.

Use the following link to provide a parsing of the first additional data set from Lab 1.
https://tinyurl.com/Parse-Worksheet

For Visualization Assignment select: **Lab assignment**
Then provide the parsed data and list your assumptions.

Part II: Additional Data Set #2

In this document:
Answer the following questions about your Tableau data:

1. List the name of the Tableau Dataset you selected in the Acquire Lab:
    Titanic Crew
2. How many rows (records) are in the data set?
    859 rows are in the data set.
3. How many columns (variables) are in the data set?
    7 columns are in the data set.
4. What assumptions are you making about the data?
    I am unsure of what the integers in the first column represent but if I had to guess, I think it tries to separate the people into groups. The Survived variable consists of the integers 0 and 1. However, it was not specified which integer means survived and which integer means didn't survive.

Use the following link to provide a parsing of the second additional data set from Lab 1.
https://tinyurl.com/Parse-Worksheet

For Visualization Assignment select: **Lab assignment**
Then provide the parsed data and list your assumptions.

What to submit:
1. This document saved as a PDF file: **LastnameFirstInitial-CGT270Fall2021-Lab2Parsing.pdf**
2. Parse Activity Worksheet (.pdf) for the Tableau data set.
3. Parse Activity Worksheet (.pdf) for 1st additional data set.
4. Parse Activity Worksheet (.pdf) for the 2nd additional data set.