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

Name: **Thomas Cluff**

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:
    a. Pokemon
2. How many rows (records) are in the data set?
    a. In the first sheet (Pokemon) there are 1168 rows of data. In the second sheet (Moves) there are 607 rows of data. In the third sheet (Evolution) there are 390 rows of data. In the fourth sheet (TypeChart) there are 324 rows of data.
3. How many columns (variables) are in the data set?
    a. In the first sheet (Pokemon) there are 9 columns of data. In the second sheet (Moves) there are 9 columns of data. In the third sheet (Evolution) there are 5 columns of data. In the fourth sheet (TypeChart) there are 4 columns of data.
4. What assumptions are you making about the data?
    a. I am assuming the names of the Pokemon are the English names. I am assuming the Pokedex number is the national dex number and not the Generation 8 Galar dex numbers. When looking at the stats of pokemon I am assuming the pokemon are not buffed or affected by status effects (paralysis, Dynamaxing). For the Moves sheet, I am assuming that moves that don't have a probability listed cannot miss (same as a 100% chance that can't be lowered).

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:

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

1. List the name of the Tableau Dataset you selected in the Acquire Lab:
    a. Complete Pokémon Dataset
2. How many rows (records) are in the data set?
    a. 1028 rows
3. How many columns (variables) are in the data set?
    a. 51 columns
4. What assumptions are you making about the data?
    a. Again, I am assuming that the Pokedex numbers are for a National dex, not a Galar region dex. I am assuming all type effectiveness is not affected by weather or status effects. I am assuming that the growth rate of Pokémon is base level and not buffed by special events, items, or IV training. I am also assuming that all "against" stats are for current generation (8) balancing and not previous games.

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:
    a. Egg Hatching Steps Dataset
2. How many rows (records) are in the data set?
    a. There are 889 rows
3. How many columns (variables) are in the data set?
    a. There are 8 columns
4. What assumptions are you making about the data?
    a. I am assuming the Pokedex numbers are for the national dex, not Galar dex. I am assuming that cells with the value "NA" are void data types for that row but the whole column is still integer. I am also assuming that the egg hatching steps listed are base values and not affected by any status effects.

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.