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

Name: **Ryoma Okano**

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. Answer the following questions and complete the table for each dataset.

1. List the name of the Tableau Dataset you selected in the Acquire Lab:
    a. U.S. Construction Spending, by value and category, 2002-16
2. How many rows (records) are in the data set?
    a. 50,208
3. How many columns (variables) are in the data set?
    a. 20
4. What assumptions are you making about the data?
    a. I believe that the construction spending has risen each year, especially infrastructure for transportation.

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.

Tableau Data Set

In the table below list each variable and its data type (add more rows as needed):

	Variable	**Data type**
1	**per_idx**	**Integer**
2	**per_name**	**Date**
3	**cat_idx**	**Integer**
4	**cat_code**	**String**
5	**cat_desc**	**String**
6	**cat_indent**	**Boolean**
7	**dt_idx**	**Integer**
8	**dt_code**	**String**
9	**dt_desc**	**String**
10	**dt_unit**	**String**
11	**et_idx**	**Integer**
12	**et_code**	**String**
13	**et_desc**	**String**
14	**et_unit**	**String**

Save this file as: **LastnameFirstInitial-CGT270Fall2021-Lab2Parsing.pdf**

15	**geo_idx**	**Integer**
16	**geo_code**	**String**
17	**geo_desc**	**String**
18	**is_adj**	**Boolean**
19	**val**	**Integer**
20	**serialid**	**Integer**

You may add more rows and attach additional pages if needed.

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

Additional Data Set #1

1. List the name of the first (1st) additional data set you acquired in the Acquire Lab:
    a. PBNRESCONS (Public Non-Residential Construction Spending)
2. How many rows (records) are in the data set?
    a. 4
3. How many columns (variables) are in the data set?
    a. 214
4. What assumptions are you making about the data?
    a. I believe public construction typically costs more to build because of subsidies and general need for public infrastructure.

In the table below list each variable and its data type (add more rows as needed):

	Variable	**Data type**
1	**realtime_start**	**Date**
2	**value**	**Integer**
3	**date**	**Date**
4	**realtime_end**	**Date**

You may add more rows and attach additional pages if needed.

Additional Data Set #2

1. List the name of the second (2nd) additional data set you acquired in the Acquire Lab:
    a. TLPRVCONS (Total Private Construction Spending)
2. How many rows (records) are in the data set?
    a. 4
3. How many columns (variables) are in the data set?
    a. 322
4. What assumptions are you making about the data?
    a. I think private construction spending will be less than public, as they are for the most part not subsidized, and much of the buildings are residential.

In the table below list each variable and its data type (add more rows as needed):

	Variable	**Data type**
1	**realtime_start**	**Date**
2	**value**	**Integer**
3	**date**	**Date**
4	**realtime_end**	**Date**

You may add more rows and attach additional pages if needed.