

Lab 2: Parsing DataName: **Samruddhi Tawade**

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	<i>Describe</i> what happens in the parse stage.
Understand	<i>Describe</i> the data in detail according to the parsing specifications.
Apply	<i>Demonstrate</i> the ability to change data into a useful format for future processing.
Evaluate	<i>Categorize</i> the data according to parsing specs.
Analysis	<i>Identify</i> specific features about the data.
Create	<i>Generate</i> 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. Magnitude 6+ earthquakes
2. How many rows (records) are in the data set?
 - a. 8314
3. How many columns (variables) are in the data set?
 - a. 18
4. What assumptions are you making about the data?
 - a. I am assuming that the unit for the depth is measured in kilometers because the chart doesn't specify it, and that is usually the unit that earthquake depth is measured in. I am also assuming that the gap is referring to the seismic gap, which is measured in years.

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:
 - a. Significant Earthquakes, 1965-2016
2. How many rows (records) are in the data set?
 - a. 23,413

Lab 2: Parsing Data

3. How many columns (variables) are in the data set?
 - a. 21
4. What assumptions are you making about the data?
 - a. Some of the dates are not given, but they are in ascending order, so I'm going to assume that the dates of the earthquakes that were not given fall in between the dates that were given. The units were not given again for the depth of the earthquake, so I am assuming that they are being measured in kilometers.

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. Earthquakes in 1910 - 2017, Turkey
2. How many rows (records) are in the data set?
 - a. 24008
3. How many columns (variables) are in the data set?
 - a. 17
4. What assumptions are you making about the data?
 - a. I'm assuming that the distance and depth is being measured in kilometers because Turkey uses the metric system. I'm also assuming that the direction is referring to the direction of the earthquake in relation to the city it occurred in/near.

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.