

Name: Ayo Obayomi

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:

I selected the Dataset for the 2018 World Cup Squads

1. How many rows (records) are in the data set?

There are 1696 rows in the data set

1. How many columns (variables) are in the data set?

There are 9 columns in the data set

1. What assumptions are you making about the data?

Goals: how many goals each individual player scored for their country to either win the group stage or going all the way to win it for the 2018 World Cup Final

Position: According to the spreadsheet, it shows that there is a significant difference between goals scored as a MF, DF, FW, and GK

DOB: it is possible that age has an influence on their aggression, understanding their position, and being prudent to not cost their country a mistake

Country and Club: If a player showcase their talent, individual skills, and effectiveness to their country, it can some times lead the citizens of the country to idolize them like Portugal idolizes Ronaldo and Argentina Messi

Use the following link to provide 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:

I selected the dataset that includes the results from the World Cup (no year was given)

1. How many rows (records) are in the data set?

There are 853 rows

1. How many columns (variables) are in the data set?

There are 12 columns

1. What assumptions are you making about the data?

Home Team, Away Team, Scoreline, and Observation: This dataset is more detailed to the intended audience which is soccer fans because it provides the scoreline between Home Team and Away Team and which team won as well the observation that perceives more information regarding who won the match after Extra time or something like that. However, it is different to see between it was cut in half.

First Round result between two teams - the Final: The data set also provides what match was played in the first round, the second round and up to the final which also includes the final score.

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:

I selected the dataset for World Cups

1. How many rows (records) are in the data set?

There are 21 rows

1. How many columns (variables) are in the data set?

There are 10 columns

1. What assumptions are you making about the data?

Attendance: It is quite hard to determine which game had what number of people in the stands. It might be the average of every game that happened during the World Cup but I cannot imagine a number that small as 590.549 to be the average. More information should be added next time

Winner, Runners up, Second, Third, and Fourth: Besides that, the dataset provides a ton of information for each year, with a four year interval gap, for when the World Cup was played, where, what country won it, the runners-up, what country came in third, fourth, and second.

Goal scored: it shows the sum of goals scored each year between the qualified teams in their respective matches played each year

Qualified Team: To have a World Cup fixture, teams have to go through like a mini-tournament and whoever ends in the top 3 gets selected to participate in the World Cup

Matches Played: This shows the amount of games played before the final

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.