Pokeman Lab

**Open a new script and save it as a .R file, this will be your lab submission. Save all relevant commands to the .R file**

Try typing the commands, rather than copying and pasting – you may see common error messages, and this will give you the opportunity to understand them.

Download the Pokemon data from:

<https://public.tableau.com/s/resources>

Green Comments inside of an R-script

Blue Executable Code

Red You need to fill in for the lab

Click on sample data sets

**# Read in the Pokemon script, check to make sure you are in your working directory, and that your data is in your working directory, is the correct library loaded for read.xlsx()?**

**install.packages("readxl")**

**library(readxl)**

**pkmn <- read\_excel("pokemon.xlsx", sheetIndex = 1)**

**# Let's explore the dataset a bit**

**# You can click on the 'pkmn' dataframe in the Environment tab of RStudio**

**# You can also use the "head" and "tail" commands...try applying to the pkmn dataframe**

**head(pkmn)**

**tail(pkmn)**

**# Notice when we do head(pkmn) we see duplicated Pokemon, certain Pokemon can have more than one type**

**# For simplicity, let's only look at a Pokemon's first type.**

**# But how can we remove all the duplicated Pokemon?**

**# R has a function called "duplicated” which can detect which names have duplicates**

**duplicated(pkmn$Name)**

**# But we'd like to see which Pokemon are being duplicated.**

**# Let's add another column to the data frame to see which, if any, Pokemon names have duplicates**

**pkmn$pkmn\_dupes <- duplicated(pkmn$Name)**

**# This adds a new column called pkmn\_dupes to the pkmn dataset**

**# Let's check out what it looks like**

**head(pkmn)**

**PAUSE HERE, CHECK ANSWERS WITH CLASS, BRING FORWARD ANY QUESTIONS**

**# Now, let's get rid of the dupes by filtering**

**# We are going to filter out all the duplicates**

**# Try to understand the following:**

**pkmn <- pkmn[pkmn$pkmn\_dupes == "FALSE",]**

**# Now that we have cleaned the data, let's do some analysis.**

**# Let's explore the data**

**# Make a table of the Type of Pokemon**

**table(pkmn$Type)**

**# Show the median, min, max, first and third quartiles of attack points**

**summary(pkmn$Attack)**

**# Summarize attack points of POISON type Pokemon**

**# Try to understand the following:**

**summary(pkmn[pkmn$Type == "POISON", ]$Attack)**

**# Feel free to explore further, using these commands or others to feel more familiar with the dataset**

**# Now, the whole point of analyzing Pokemon is to create a lineup to battle**

**# Let's go with a high attack deck. What are the 6 highest attack Pokemon?**

**# First order the data by the Attack column, descending**

**# Try to understand the following:**

**highattack <- pkmn[order(-pkmn$Attack),]**

**highattack**

**# Then look at only the top 6**

**highattack[1:6,]**

**# Can write that code in a single line, as:**

**highattack <- pkmn[order(-pkmn$Attack),][1:6,]**

**highattack**

**# NOTE YOUR RESULTS IN COMMENTED CODE**

**# Great, these are the top 6 highest attacking pokemon!**

**# But what if you want high defense Pokemon? Or a mixed lineup of high attack, high defense, and high speed?**

**# Grab the 2 highest of each category**

**PAUSE HERE, CHECK ANSWERS WITH CLASS, BRING FORWARD ANY QUESTIONS**

**highattack <- pkmn[order(-pkmn$Attack),][1:2,]**

**highdefense <- YOU WRITE THIS**

**highspeed <- YOU WRITE THIS**

**toppkmn <- rbind(highattack, highdefense, highspeed)**

**toppkmn**

**# NOTE YOUR RESULTS IN COMMENTED CODE**

**# Can you do the same, but with ONLY psychic Pokemon?**

**# Easiest way is to first filter the data to only Psychic Pokemon and do the same commands**

**psychic <- pkmn[pkmn$Type == "PSYCHIC",]**

**highattackpsychic <- YOU WRITE THIS**

**highdefensepsychic <- YOU WRITE THIS**

**highspeedpsychic <- YOU WRITE THIS**

**toppsychic <- YOU WRITE THIS**

**toppsychic**

**# NOTE YOUR RESULTS IN COMMENTED CODE**

**PAUSE HERE, CHECK ANSWERS WITH CLASS, BRING FORWARD ANY QUESTIONS**

**# What are the 6 slowest grass pokemon?**

**grass <- YOU WRITE THIS**

**slow <- YOU WRITE THIS**

**slow**

**# NOTE YOUR RESULTS IN COMMENTED CODE**

**# The Total column is the sum of a Pokemons’ stats. Knowing that, let look into the data.**

**# Which Pokemon is the WORST? WORST Pokemon has lowest total, try using the min() function**

**YOU WRITE THIS**

**# NOTE YOUR RESULTS IN COMMENTED CODE**

**# Which Pokemon are completely mediocre? Define mediocre as MEDIAN, try using the median() function**

**YOU WRITE THIS**

**# NOTE YOUR RESULTS IN COMMENTED CODE**

**# Try messing around with different pokemon types. What's your ideal Pokemon lineup? Why?**

**PAUSE HERE, CHECK ANSWERS WITH CLASS, BRING FORWARD ANY QUESTIONS**

**OPTIONAL ONGOING CHALLENGE (submit whenever you want feedback):**

Read in the other three sheets from the Pokemon file and explore them. How could you use this data to improve your team? You may need to quickly research some information about the game of Pokemon.

How would your team of 6 change if you were limited to 3000 total points?

Create a new team based on this limitation, carefully document your process of exploration and decision making in the .R file

How would your team of 6 change if you knew my team were all DRAGON and GHOST Pokemon (assuming you’re still limited by 3000 total points)?

Create a new team based on this limitation, carefully document your process of exploration and decision making in the .R file

Resource:

http://www.wikihow.com/Create-a-Balanced-Pok%C3%A9mon-Team