Building Bridges

Example Code

Building bridges

Let's run through some of the analyses we will do in this lab with example datasets. We will start with the dataset ChickWeight available in R. ChickWeight contains data on chick weights across time:

```
# Save the first three columns of the ChickWeight dataset
example <- ChickWeight[,-4]
head(example) # Print out the variable names

# Subset these data so we're only looking at 2 timepoints (Time = 0 and Time = 10)
# to resemble the Building Bridges data, which has 2 generations
chickDf <- subset(example, Time == 0 | Time == 10)

# Plot the data
plot(chickDf$Time, chickDf$weight)

# This plotting method doesn't allow us to see which chick is which,
# and how each chick changed across time! Let's fix that.

# Subset data to isolate individual chicks within the dataset
chick1 <- subset(chickDf, Chick == 1)
chick2 <- subset(chickDf, Chick == 2)

# Set up your plotting space
plot(chickDf$Time, chickDf$weight, type="n")

# Add lines for each chick (RUN this at the SAME time as the 'plot' line above)
lines(chick1$Time, chick1$weight, col="blue")
lines(chick2$Time, chick2$weight, col="red")

# If you want to get fancy, you can also add points
points(chick1$Time, chick1$weight, col="blue")
points(chick2$Time, chick2$weight, col="red")
```

It may be easier to see overall trends by comparing timepoints using a boxplot. Boxplots are great because they show us more about the data's distribution (i.e., median and quartiles) than we can see with a scatterplot.

Let's try creating a boxplot with the ChickWeight dataset:

```
# Compare chick weights across the 2 timepoints:
boxplot(chickDf$weight ~ chickDf$Time,
        xlab="Timepoint", ylab="Weight")
```

Another useful analysis is a linear regression, which summarizes the relationship between two continuous variables. Let's practice carrying out a linear regression with the example dataset Iris available in R. Iris contains various measurements from several plant species:

```
head(iris) # Print the variable names

# Plot sepal width by sepal length
```

```r
plot(Sepal.Length ~ Sepal.Width, data=iris)

# Run a linear model to explain sepal length by sepal width
# Save the output as mod1
mod1 <- lm(Sepal.Length ~ Sepal.Width, data=iris)

# How would you find the slope of the line?
summary(mod1)

# Add linear regression line to the plot
plot(Sepal.Length ~ Sepal.Width, data=iris)
abline(mod1)
```

Now, give these analyses a try for your Building Bridges data:

Access the bridge data: [google sheet] (#INSTRUCTORS - ADD IN URL OF GOOGLE SPREADSHEET, INSIDE THESE PARENTHESES)

You'll work with a .csv data file from multiple lab sections. The first part of the assignment will involve analyzing data from *only your lab section.*

Who else was in your group?

your response here

Start by reading in your data:

```r
# Read in your data
# Replace filename with the name of your file
df <- read.csv("filename")

# Subset to get the data from your lab section (subset by day_time)
# IN OUR SPREADSHEET, THAT COLUMN WAS day_time
# Replace labsection with the name of your section (day_time)
df_lab <- subset(df, day_time == "labsection")
```

Prediction 1: Fitness varies among individuals

Prediction 2: Average fitness increases from one generation to the next

Prediction 3: There is a trade-off beween bridge length and load score

Prediction 4: Structural complexity increases from one generation to the next

Let's visualize these data in a way that lets us track how each group's bridge changed across generations (just like we did with the ChickWeight dataset):

```r
# Plot generation time by fitness
plot(df_lab$generation, df_lab$fitness)

# Subset the lab data to isolate the bridges for each group
# e.g. team1 <- subset(df_lab, team_name == "teamname")
# Insert your code here

# Plot each team individually on the same graph to track bridges through time
# Insert your code here
```

Interpret this figure - is there variation in fitness among bridges? Does fitness seem to increase across generations?

your response here

Lets create a boxplot to analyze changes in fitness across generations. Evaluate Prediction 1 and Prediction 2 using your the data from your lab section:

```
# Plot fitness for each generation as a boxplot
# Insert your code here
```

Prediction 1: Was your intial prediction supported?

your figure results and evidence-based claim here

Prediction 2: Was your intial prediction supported?

your figure results and evidence-based claim here

Create a scatterplot to determine whether a tradeoff exists between load score and length and overlay a linear regression line (just like we did with the Iris dataset):

```
# Run a linear model to explain load score by length,
# e.g. lm(y ~ x, data = df_lab)
# Insert your code here


# Find the slope of the line
# Insert your code here


# Add linear regression line to the plot
# Insert your code here
```

Interpret this figure and evaluate prediction 3: Was your intial prediction supported?

your response here – keep the asterisks.

Calculate structural complexity of bridges and compare across generations

The Shannon Index of diversity considers how many different resources were used to build each bridge AND how evenly they were used (e.g., if 90% of the resources used were paper cups, that would not represent an even use of resources). Before running the code below, install the package vegan (install.packages("vegan")) in the R console.

Calculate the Shannon Index for the building bridges data:

```
# Modify the code below as needed to exactly match the data sheet
materials_lab <- subset(df_lab, select = c("paper", "paper_clips", "sticky_pads",
                                           "rubber_bands","straws", "cups","bowls",
                                           "plates","labels","knife"))

# Load the vegan package
library("vegan")

# Calculate the shannon index using the diversity function
shannon <- diversity(materials_lab, index="shannon")
```

```r
# Add the results from the shannon index as a new column in df_lab
df_lab$shannon
```

Great! Now you've got Shannon diversity for each bridge in your lab section. Visualize how this changes across generations using a boxplot, like above:

```r
# Plot shannon index for each generation as a boxplot
# Insert your code here
```

Interpret this figure and evaluate prediction 4: Does complexity increase or decrease across time?

your response here

Self-assessment of group collaboration

Assess whether everyone played a role or whether the contributions were very unbalanced. Give your overall impression of your individual ability to contribute in positive ways towards your group's bridge.

your response here

Bonus section! Evolution is not uniform across populations. . . explore this for yourself by comparing results from your lab section to other lab sections.

```r
# Subset data to get another lab section's data
# Insert your code here

# Use a boxplot to look at how fitness changed across generations in that lab,
# or see if they saw a trade-off in length vs. load of the bridges.
# Compare their results to your results from this lab section!
# Insert your code here
```

Your comparison of results across lab sections here