Part 1. Getting Started

1. Set and check your current working directory. A working directory can be thought of as a "folder" that you would normally find on your computer. Copy and paste your output from the console when running the getwd() function. **(1 pt)**

/cloud/project

2. Save the Elwha estuary dataset in your R environment by following the example code provided. Check the structure using str(), and then preview the first 5 rows and 6 columns. Copy and paste your output from the console when running the str() function. **(1 pt)**

```
  Field_Activity_Number  Dam_Condition Date_Collected Site_Name
1                  <NA> Before removal      6/22/2006       ES1
2                  <NA> Before removal      6/22/2006       ES2
3                  <NA> Before removal      7/20/2006       ES1
4                  <NA> Before removal      7/20/2006       ES2
5                  <NA> Before removal      8/31/2006       ES1
  Latitude Longitude
1 48.14700 -123.5635
2 48.14827 -123.5616
3 48.14700 -123.5635
4 48.14827 -123.5616
5 48.14700 -123.5635
```

3. Provide the longitudinal and latitudinal coordinates for a Field-ID of your choice below. Make sure to include the Field-Activity-Number in your answer! **(1 pt)**

Field-ID 4: 48.14827 -123.5616

Part 2. Descriptive Statistics

4. Walk through generating the descriptive statistics for pH values. Next, create similar R code to generate descriptive statistics (five-point summary) for both temperature and turbidity. Fill in the tables below and add descriptive table title descriptions to each. **(2 pts)**

[1] 13.11597 [1] 5.79 [1] 20.98 [1] 3.289619

[1] 63.23206 [1] 13.9 [1] 0.3 [1] 305.6 [1] 88.97956

5. Calculate the 95% confidence interval for temperature (°C) and write a statement below including these values to assess our confidence in the temperature mean. **(2 pts)**

[1] 12.58419 [1] 13.64775

Part 3. Statistical Analyses in R

6. Walk through the t-test determining if the mean pH values are significantly different before and during dam removal. Next, determine if there is a significant difference in the mean turbidity values before and during dam removal by conducting your own t-test. Report your results in the context of the study by interpreting the p-value and variables used. **(2 pts)**

Before removal: 13.10526 After removal: 71.73571

7. Walk through the example ANOVA determining if the mean pH values are significantly different between different testing sites. Next, conduct your own ANOVA to determine if there are site differences in temperature during dam removal. Report your results in the context of the study by interpreting the p-value and variables used. **(2 pts)**

```
> summary(ANOVA_temp)
             Df Sum Sq Mean Sq F value Pr(>F)
Site_Name     4   83.2   20.80   1.862  0.122
Residuals   113 1262.5   11.17
8 observations deleted due to missingness
```

There was no significant difference in temperature among sampling sites during dam removal (F=1.862; df=4,113; p=0.122).

8. Produce a boxplot to help illustrate your results from AQ-7. Export the image (copy to clipboard), paste below, and include a descriptive figure caption. **(2 pts)**

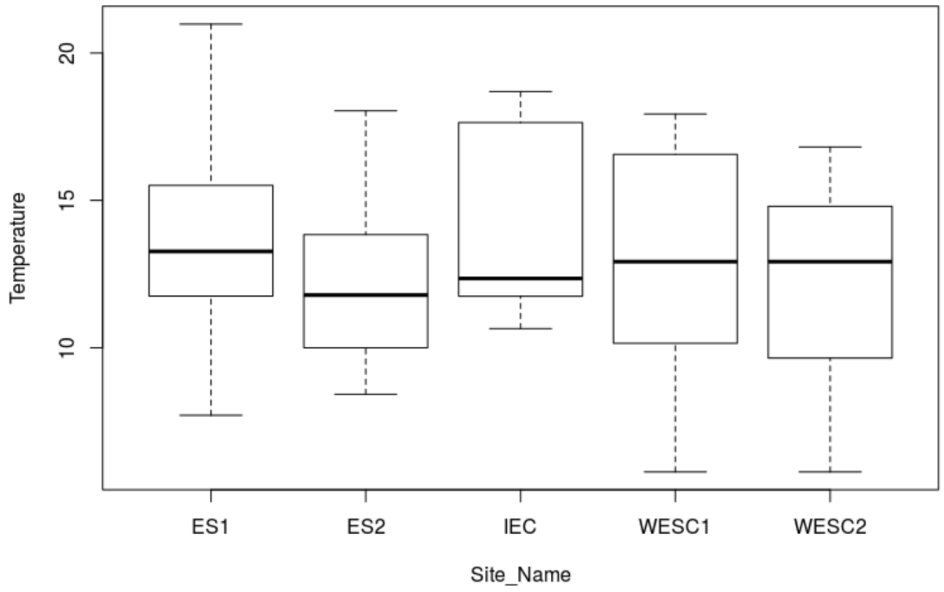

Figure 1: Boxplot that shows outliers (the two open circles in ES1), medians (thick black lines in the middle of the boxes), quartiles (outer bounds of the boxes), and minimum and maximum values (excluding outliers; the lines at the end of the dotted-line arms) for turbidity before and after dam removal.

9. What are the null hypotheses for each of the analyses (t-test; ANOVA) that you conducted above? **(1 pt)**

- There is no significant difference turbidity before and after dam removal.
- There is no significant difference in temperature before and after dam removal.

Part 4. Water quality

10. Why do you think dam removal would affect water quality parameters in a river? **(1 pt)**

Removal of the damn could change water flow which could introduce a wide range of effects, one being the water quality.

11. What is turbidity? Why do you think it is an important parameter to consider when measuring water quality? Can you think of an example of very high turbidity? **(1 pt)**

Turbidity is how clear or transparent the water is. Higher turbidity means increased particles and lower visibility. Turbidity is important to consider because it lets you know how much of the water, isn't actually water. You want lower turbidity. An example of high turbidity would be a river during a flood, where there is an abnormally high amount of additional materiel flowing with the water.

12. Why do you think water quality is important? Find an example from the primary literature to support your argument. What happens when there is poor water quality? What are some of the factors that contribute to poor water quality? **(2 pts)**

- K.E. Wiemer, A. Anderson, B. Stewart, The importance of water quality for media preparation, Human Reproduction, Volume 13, Issue suppl_4, December 1998, Pages 166–172

13. Think about a data-driven question that you would like to answer using RStudio. Find a potential dataset and describe it—what kind of variables it contains, what it measured, and what question you would like to use the data to answer. **(1 pt)**

14. What statistical test from the lab today would you use to analyze your data and why? What would be your null hypothesis? **(1 pt)**