***Answer each question below on separate sheets of paper. Make sure to clearly label each of your answers (e.g., #1 or #4a), put your name on each extra sheet used, and staple these questions to the top of your answer sheets when completed to hand in.***

1. **[4 pts]** A recent Huffington Post/YouGov.com poll asked a sample of 990 voters “Please indicate whether you agree or disagree with the following statement: Sexual harassment against women in the workplace is no longer a problem in the United States.” The percentage of responses are shown in Table 1. Perform a proper EDA from these results.

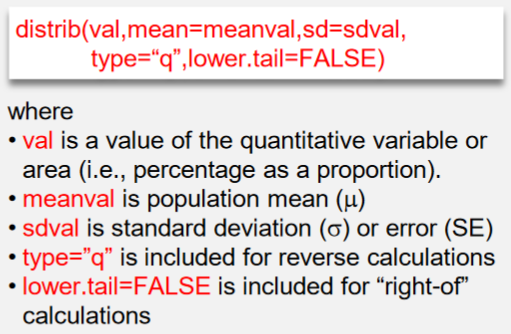
**Table 1.** Percentage of respondents by answer to the sexual harassment question.

Strongly Somewhat Somewhat Strongly Not

Agree Agree disagree disagree Sure

3.9% 7.9% 27.2% 49.6% 11.3%

1. **[22 pts]** Harvest of Lake Trout (*Salvelinus namaycush*) in the Apostle Islands from January to August has a mean of 12,100 fish and a standard deviation of 2,200 fish. If harvest exceeds 15,000 fish during this time then fishing will be closed immediately (end of August). However, if less than 8,000 fish are harvested then the season will be extended through October. Assume that the harvest of Lake Trout during this period is normally distributed. Use this information to answer the questions below. *Please write the R code that you used along with your final numerical answer (to one decimal place).*
2. In what percentage of years will the harvest season be extended through October?
3. In what percentage of years will the harvest season be closed immediately in August?
4. What should the harvest be that would cause the season to close immediately if managers want to close the harvest immediately in only 5% of the years?
5. What is the IQR for harvest?
6. What is the median harvest?



1. **[16 pts]** Researchers at the University of Michigan polled 224 Democrats and 164 Republicans in Spring 2018 about their views on climate change. The responses were labeled as whether they thought climate change was caused by humans, natural processes, or a combination of humans and natural processes; whether they were not sure that the climate was changing; or were sure that it was NOT changing. Use the results in Table 2 to answer the questions below the table *to one decimal place and show your work*.

Table 2. Frequency of responses to the climate change questions by political party.

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Party** | **Human**  **Caused** | | **Natural**  **Processes** | | **Combination** | | **Not Sure**  **Changing** | | **Not**  **Changing** | |
| **Democrat** | | 114 | | 29 | | 59 | | 18 | | 4 | |
| **Republican** | | 25 | | 24 | | 33 | | 25 | | 57 | |

1. What percent of respondents were Democrats and thought climate change was human caused?
2. What percent of Republicans were sure that the climate was not changing?
3. What percent of those that thought that climate change was caused by humans were Democrats?
4. What percent of respondents thought that climate change was caused by natural processes?
5. **[10 pts]** A student mined Christmas Bird Count data to determine the number of bird species recorded in several 15-mile diameter areas on the Demarva Peninsula in 2005. The student hypothesized that the number of bird species in an area was related to the latitude of the area. Use her results in Figure 1 to construct a complete bivariate EDA.

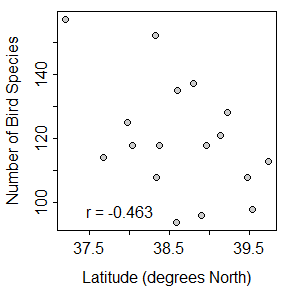


Figure 1. Scatterplot for number of bird species on latitude for Delmarva Peninsula, 2005.

1. **[10 pts]** A student collected data on the average price of a Big Mac (adjusted to U.S. dollars) and the gross domestic product (a measure of relative wealth) for a sample of 51 countries from throughout the world. Use his results in Figure 2 to construct a complete bivariate EDA.

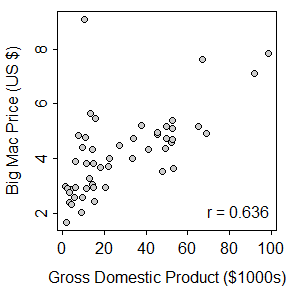


Figure 2. Scatterplot of the price of a Big Mac on gross domestic product for a sample of countries.

1. **[4 pts]** What are the two major goals of regression?
2. **[4 pts]** Define RSS and explain how it is used to find the “best-fit line.”

**[Over for One More Page]**

1. **[30 pts]** Male Gray Tree Frogs (*Hyla versicolor*) make vocal calls to attract females. Gayou (1984) measured the time between mating calls for individual Gray Tree Frogs at different ambient temperatures. His objective was to determine if the time between calls was affected by temperature. Use the results from his study in Figure 3 to answer the questions below. *Show your work as necessary*.

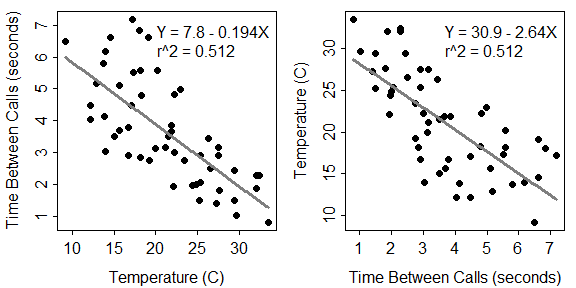


Figure 3. Fitted line plot for the time between calls on temperature (Left) and temperature on the time between calls (Right) for male Gray Tree Frogs.

1. In terms of the variables of this problem, interpret the value of the slope? [*use a complete sentence*]
2. In terms of the variables of this problem, interpret the value of the intercept? [*use a complete sentence*]
3. How much would one expect time between calls to change if temperature increased by 10oC?
4. What percentage of the variability in time between calls is explained by temperature?
5. What is the predicted time between calls if temperature is 45oC?
6. What is the predicted time between calls if temperature is 25oC?
7. What is the residual if the time between calls is 3 seconds and temperature is 20oC?
8. What is the correlation coefficient between time between calls and temperature?
9. Do you have concerns about this regression? [*thoroughly explain your answer, whether you have any concerns or not.*]