**Lab 5: Chi-Square and *t*-Tests**

**Part 1:**

In some Hollywood movies, there will always be a scene where whether Russian or the U.S. President is about to launch a nuclear attack against the other country. Luckily, at the last second, our hero will avoid a catastrophe. Using the data collected by the Pew research group, investigate the question “Is there a similar cold-War era relationship with the US and China?”.

Do you think United States and China will one day confront resembling the Cold War? Millions of people died in Cold War due to proxy wars. Some good things happened as well: many interstate highways, including I-35, were constructed during that era to have a quick national response in case of an atomic bomb attack by Russia. Internet and space explorations are indirect results of the Cold- War era race.

For this assignment, we are going to evaluate the responses from question 4 “How would you rate the likelihood of the current rivalry between China and the United States escalating into a confrontation resembling the Cold War?” to test whether the political view matters in this hypothetical scenario. In other words, we are interested in the research question “Are people with different political views equally likely to say conflict akin to the Cold War will occur between the China and the US?”.

**Analysis Tasks for Research Question 1**

1. Prepare a table of the responses for question 4 by political views as demonstrated in the Chi-Square Analysis handout in Step 4 (except that example showed question 3 by gender). Make sure to include your name in the footer. Failure to complete this in SAS with your name in the footer will result in no credit.
2. Now, select the PEW2020Filtered data set from your library (You created this following the instructions included in the handout accompanying this week’s lab). Prepare a new table of the responses for question 4 from the survey by political views using the recoded variables including the expected values with your first and last name in the footer. (Under options, select suppress plots so that your footer is visible).
3. State the null hypothesis and alternate hypothesis for this analysis.
4. Now, run perform a Ci-Square analysis. Provide a screen shot of the table titled “Statistics for table of Q4 by polview” which also includes your first and last name in the footer. You will not receive any credit for submissions that are not generated by SAS with your name in the same image as the table.
5. What is the critical χ2 value for this problem at the 0.05 significance level?
6. Write out the decision rule.
7. Based on the results do you reject or fail to reject the null hypothesis?
8. What did you used to make your decision in number 7?

In one sentence, interpret the findings (i.e., Do you have statistically significant difference in the percentage of people with different political views on their view about the likelihood of current rivalry between China and the United States turning into a confrontation resembling the Cold War?)

1. Provide an interpretation of the Cramer’s V value (Positive or negative relationship; strong, weak or no relationship).
2. What can we say, based on the Cramer’s V value?

**Part 2:**

The data for this survey was collected in September 2020, which is just two months before the most recent U. S. Presidential election and it happened to be during a global pandemic, when many Americans were instructed to stay home if they held positions that were considered non-essential. Additionally, it is common for people’s differences to come to the surface more in the months surrounding a presidential election, in which the two dominate parties tend to be identified a Liberal and Conservative. This is also a time when new ideas have the potential to make their way to a larger audience as citizens debate various issues and determine which candidate would best represent their beliefs and values. This raises the question of “does the age of voters differ among the two major political groups (Liberals and Conservatives)?”.

1. State the null hypothesis and alternate hypothesis for this research question.
2. Make sure that you only have Liberal and Conservative in to Recoded Polview column. Provide a screen shot of the table including the tests for normality for the PEW2020Filtered data set.
3. Based on the Anderson-Darling test, is age normally distributed and how do you know?
4. Provide the histogram and box plot for an additional test of normality. Do not forget your First and Last name in the footer.
5. Based on the histogram, what can you conclude?
6. Provide a screen shot of the table Equality of Variances table.
7. Which *t* statistic should you use to evaluate the hypothesis?
8. How did you determine which *t* statistics to use?
9. Complete the following sentence:

The \_\_\_\_\_\_\_\_ *t* value is \_\_\_\_\_\_\_\_\_ with a *p* value of \_\_\_\_\_\_\_\_\_, which is \_\_\_\_\_\_\_\_\_ than 0.05. Therefore, or test statistic \_\_\_\_\_\_\_\_\_ statistically significant. We \_\_\_\_\_\_\_\_\_ the null and conclude: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_.