Lyric Readability

In May 2015, data analysts at *SeatSmart*[1] examined the lyrical readability (reading-level) for over 200 songs to examine whether music lyrics are getting "dumber". Each song's reading-level was operationalized using the Flesch–Kincaid score. This score quantifies the reading-level by indicating the grade-level a person would need to comprehend the lyrics. For example, Beyonce's hit song, *Irreplaceable*, has a readability score of 3.9, indicating that a person with slightly less than a fourth-grade reading level would comprehend the vocabulary used in the song's lyrics. In this assignment, you will use the *SeatSmart* data (available in *lyric-readability.tp*) to answer the following research question:

> Is the lyric readability for newer songs (released after 2005) less than that for older songs (released in or prior to 2005)?

Each question is worth 1point unless otherwise indicated.

Explore the Observed Data

Open the *lyric-readability.tp* file in TinkerPlots. Create a single plot of the observed data to compare the reading-levels between recent and older songs. Also include the numerical mean values for each group on the plot.

1. Copy-and-paste (or sketch) the plot of the observed data, with the averages included, into your word-processed document.

Use the groups' mean values to compute the *difference in means* for the two groups. Subtract the recent average from the older average. (Your answer should be a positive number.)

2. Report the observed difference in means.

3. What does the observed result (the difference in means) suggest about the answer to the research question? Explain.

[1] Analyses were reported at https://seatsmart.com/blog/lyric-intelligence/.

4. Why do you need to conduct a simulation in order to determine whether recent songs have a smaller readability than older songs? Explain why you cannot answer the research question using just the observed result. **(2pts)**

Bootstrap Test

5. Write the *null hypothesis* for the bootstrap test.

Set up a TinkerPlots sampler to generate different bootstrap samples of the data under the model specified in the null hypothesis (Hint: You did something similar to this in the *Murderous Nurse* activity).

6. Copy-and-paste (or sketch) a picture of the TinkerPlots sampler window into your word-processed document.

7. What repeat value did you use for your sampler? Explain why you chose this value.

8. Did you use *with* or *without* replacement for the readability scores in the sampling device? Explain your choice.

9. Did you sample *with* or *without* replacement for the group labels (newer and older) in the sampling device? Explain your choice.

Simulate using the Model

Use TinkerPlots to carry out 500 trials of the bootstrap simulation. Create a plot of the 500 bootstrapped differences in means.

Evaluate the Observed Results

10. Describe the shape of the distribution of the 500 simulated results. Also compute and report the mean and standard deviation of this distribution.

11. Explain, by referring to the null hypothesis, why we could have expected the mean you computed in Question #10.

Add a reference line to the plot of the distribution of the 500 simulated results at the value of the difference in the **observed** data.

12. Copy and paste (or sketch) the plot (with the reference line) into your word-processed document.

13. Use the plot from Question #12, to compute the *p*-value. Show your work.

Answer the Research Question

14. Based on the *p*-value you computed, how compatible is the observed difference in means with the results produced by the model specified in the null hypothesis? What does this suggest about the answer to the research question? Explain. **(2pts)**