**Research Question**

A.  Summarize the original real-data research question you identified in task 1. Your summary should include justification for the research question you identified in task 1, a description of the context in which the research question exists, and a discussion of your hypothesis.

Do songs that are released after Spotify comes out have a statistically significantly shorter average length than songs that came out before Spotify (2008)?

This question would benefit from data analysis because it gives us insight into both a cultural aspect of modern society in the seemingly widespread shortening of consumed entertainment and a guideline for musicians. It gives a measure for where the music industry could possibly be going and whether there are outliers that are able to make longer music, or if statistically it is necessary to stay within the bounds of certain song timeframes.

My hypothesis for this research question is that released music has become statistically significantly shorter since after Spotify’s release.

**Data Collection**

B.  Report on your data-collection process by describing the relevant data you collected, discussing one advantage and one disadvantage of the data-gathering methodology you used, and discussing how you overcame any challenges you encountered during the process of collecting your data.

For data collection, I will be using a dataset from Kaggle that has 1.3 million songs with metadata related to answering the hypothesis of this question.

**Data Extraction and Preparation**

C.  Describe your data-extraction and -preparation process and provide screenshots to illustrate each step. Explain the tools and techniques you used for data extraction and data preparation, including how these tools and techniques were used on the data. Justify why you used these particular tools and techniques, including one advantage and one disadvantage when they are used with your data-extraction and -preparation methods.

**Analysis**

D.  Report on your data-analysis process by describing the analysis technique(s) you used to appropriately analyze the data. Include the calculations you performed and their outputs. Justify how you selected the analysis technique(s) you used, including one advantage and one disadvantage of these technique(s).

**Data Summary and Implications**

E.  Summarize the implications of your data analysis by discussing the results of your data analysis in the context of the research question, including one limitation of your analysis. Within the context of your research question, recommend a course of action based on your results. Then propose **two** directions or approaches for future study of the data set.

F.  Acknowledge sources, using in-text citations and references, for content that is quoted.

G.  Demonstrate professional communication in the content and presentation of your submission.