**DSC 530-T302 Data Exploration and Analysis**

**Jake Meyer**

**04/04/21**

**Exercise 2-1:**

**Which summary statistics would you use if you wanted to get a story on the evening news? Which ones would you use if you wanted to reassure an anxious patient?**

To get a story on the evening news, an individual would want to portray that the results of the study were confirmed. This could be achieved by only considering the mean of the two groups (first babies, non-first babies). The difference in the mean indicates that first babies, on average, tend to arrive later than non-first babies. In other words, an individual will want to convey that there is an effect, regardless of the size, that exists for pregnancy length when considering first babies arrive later than non-first babies.

To reassure an anxious patient, an individual or doctor would want to consider showing the context of the differences and the likelihood for the first baby arriving late. An individual could summarize the difference between the pregnancy lengths for first and non-first babies by breaking it down as a 13-hour difference (on average). In addition, an individual could reiterate the small effect of first babies arriving later than non-first babies. The means for both groups still reside below the 39-week pregnancy period.

**“Do first babies arrive late?” Write a paragraph that uses the results in this chapter to answer the questions clearly, precisely, and honestly.**

A study was performed to answer the question, “Do first babies arrive late?” Data was analyzed from the National Survey of Family Growth (NSFG) to help answer this question. The data consists of survey responses from 01/2002 to 03/2003 for individuals ages 15-44. The data was targeted towards Hispanics, African Americans, and teenagers and includes responses for various pregnancy questions. The variable analyzed for this inquiry was pregnancy length. For the analysis, the pregnancy lengths for respondents were broken down into two groups: first babies and others (non-first babies). The average pregnancy length was calculated for each subgroup along with the effect size (Cohen’s d) between the two means. The mean length of pregnancy for first babies was 38.601 weeks, whereas the mean length of pregnancy for other babies was 38.523 weeks. This alludes to a difference in means of approximately 13 hours between the two groups. Cohen’s d was calculated to be 0.029 which indicates a small effect. These characteristics indicate that first babies do, on average, tend to arrive later than non-first babies. Although this is true, the effect size and mean difference are relatively small indicating no major real-world outcomes.