

Name: Christopher Fiah
SIE 507 – Project 2

Q5. After visualizing the data in multiple ways, what are the hidden features, data points, or trends you can find?

Briefly explain your analysis, observations, and findings:

a. What are your overall conclusions?

There is a constant fluctuation of the number of births for every other month of the years. If the preceding month has a high count there is a high chance the next month will experience a drop in the birth count.

b. What are your recommendations or predictions?

From analyzing the of the data, there is a high probability the birth counts at the end of the current year and start of the new year being high. My recommendation is better logistical arrangement should be made period and proper staffing for these peak periods.

c. What are some new things that you learned from the dataset once you visualized the data in different ways? What surprised you?

- In 2007 it was noticed the birth counts are substantially higher on most of the months or close to the highest point. This results in having a very high maximum point and mean point when the data was group into two groups of ten. Increase in teenage births and the occurrence of the Great Recession right after 2007 were notable factors for the spike as reported by the New York Times (<https://www.nytimes.com/2009/03/19/health/19birth.html>) and Market Watch (<https://www.marketwatch.com/story/the-u-s-birth-rate-has-fallen-since-the-2007-great-recession-even-as-growth-and-wealth-have-recovered-what-happened-11645111685>)
- When the data were grouped into months, it was noticed we had outliers in the months of February, August, November.
- The mean values of the months February and July were very low compared to the others.

d. Can you think of any innovative ways to plot the data/findings?

Using a violin plot we can show the mean and median of each month.

Code written in script to illustrate.