**Analyst Position Coding challenge**

Using any programming language(s) you like, answer as many of the following questions as you can. It's okay if you don't do all the questions. If you get stuck, write pseudo-code or explain what you could do or would do.

Examples of languages in order of preference:  
R, python, SQL, Matlab/octave, Stata, SAS, perl, java, Julia, C, C++, javascript (node.js), objective C/Swift

***Don't use a spreadsheet in Excel or calculate by hand. Real datasets are too large for this to be practical.***

**About the data:**

The data come from CDC Wonder’s online open access data portal. The data contains state-level annual fertility information from the United States.

**Coding challenge questions**

1. Read in the data. *Bonus points for using the map package to avoid loops of repetitive code*
2. Dig into the data a bit! Explore the variables, examine missing data. Summarize the data as you see fit. Explain some basic data cleaning steps you took.
3. Create a time series plot of the total number of births. *Bonus points for highlighting the state with the highest average number of births in red*
4. Estimate a regression model looking at how state and year impact birth rates.
   1. Did the birth rate increase or decrease over time?
   2. Interpret the intercept term

*10X bonus points for doing the challenge in RMarkdown and sharing your results in an HTML file*

*100X bonus points for doing the challenge and pushing your work to your Github repo*

*1000X points for presenting the results in a Shiny app*