Diag Davenport

Deliverable 2

# Objective

For the past two years, [commercials](https://www.youtube.com/watch?v=leU-mgKfdcs) have suggested that the cities of the Super Bowl winning team see a rise in births in the nine months or so following the Super Bowl win. This analysis seeks to validate that assertion.

# Data Curation

The key input to this analysis is a CDC-published, county-and-month-level dataset of birth rates. The most notable step of the curation was transforming county-level data into a dataset aggregated to a geographic region appropriate for capturing the effect of a team winning the Super Bowl. This was almost entirely a judgement call. The counties that were rolled into being eligible to pick up the effect of winning can be found explicitly in the code. The principle implemented here, though, was based on the likelihood that people consider themselves from the same place as the locale the team represents. The second kind of aggregation was based on the dimension of time. I aggregated the monthly data into quarters to smooth out some of the noise and analytically recognize that there may be a lagged effect to winning; I’m not strictly interested in an effect nine months after the Super Bowl, rather an effect *around* nine months after the Super Bowl.

# Analysis

The structure of the analysis is a comparison between two quantities: 1) the actual change in birth rates in the final months of the year, and 2) the expected change in birth rates for that period. In cases where actual is greater than expected, we have validated that there is in fact an increase in births after a Super Bowl.[[1]](#footnote-1)

Calculating changes in birth rates is a delicate matter. There is a temptation to evaluate the changes in consecutive months, but this approach is misguided given the consistent cycles throughout the year that we observe in the data. Changes, in this analysis, are evaluated relative to the same period in the previous year. This allows the model to ignore the periodicity in the data and focus on level changes in the cycle. The null hypothesis then is that changes in January births are similar to changes in April births. And that changes in April births are similar those in November. By extension, the expectation is that the average change in birth rates outside of the post-Super Bowl period is similar to the average change within the period.

The graphs available [here](https://github.com/diagdavenport/SuperBowlBabies/blob/master/Results/Year%20over%20year%20differences.pdf) indicate the differences between those averages and how they coincide with Super Bowl participation.

1. Note that no attempt is made to suggest that the Super Bowl win *caused* such an effect, rather that the increase did, in fact, occur post-Super Bowl. [↑](#footnote-ref-1)