

# Week 1 Statistical computing

*MS 276*

*September, 2017*

## Starting with Excel

To start learning R, we'll need to be familiar with Microsoft Excel. The point of today's lecture is to learn basic Microsoft Excel tools for data manipulation and statistical exploration.

Let's start by downloading the Gapminder data set by going here:

<https://www.dropbox.com/s/k50er58pjxnx6k6/gapminder.02.csv?dl=0>

Open the file in Excel and answer the following questions.

1. How many rows and columns? These are the dimensions of the data
2. In cell J2, calculate the average life expectancy for all countries in the data (using typing the command `= average(D2:D143)` in J2).
3. In cell J3, calculate the median gdp per capita for all countries in the data (`= median(F2:F143)`)
4. Identify if there's any correlation between life expectancy and gdp per capita (`CORREL(D2:D143, F2:F143)`) in cell J4.
5. Make a new variable in column G, `life.70`, which is either TRUE or FALSE depending on whether or not average life expectancy in a country is greater than 70. In cell G2, type `= D2 > 70`, and you can drag the bottom right corner down to fill in the other cells.

At this point, double check your answers with the ones provided on the board and we'll learn some additional tools together.