Excel Exercise

<http://kpq.github.io/sherp-31/weeks/week3.html>

<http://kpq.github.io/nyu-data-journalism-fall-2014/classes/git-and-excel/>

<http://shancarter.github.io/ucb-dataviz-fall-2013/classes/kobayashi-maru/>

<http://shancarter.github.io/ucb-dataviz-fall-2013/classes/charts-intro/>

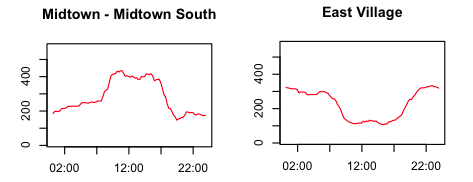
Basic skills

Here is some data on [drunk driving deaths](http://kpq.github.io/sherp-31/assets/data/drunkdrivingdeaths.xls) in 2011.

1. Download the data to your computer, and open it in your spreadsheet program of choice.
2. What were the 3 states with the most deaths? The 3 with the fewest?
3. Here is some data on [population](http://kpq.github.io/sherp-31/assets/data/statepopulation.xls). Add it to your spreadsheet using a function called vlookup.
4. What were the states with the highest rates? The lowest?
5. How many deaths occured in the South? What region of the country had the highest rate of deaths? Here are some [region codes](http://kpq.github.io/sherp-31/assets/data/nytnames.xls) that might help you. For practice, answer this question by using a pivot table. Pivot tables are one way to summarize data.

Basic skills when your data is bigger

A digression: sometimes data is published to the web, but, without context, it’s hard to know whether it’s interesting or not. Here’s [an example](http://schools.nyc.gov/AboutUs/data/Attendance.htm) about [NYC school attendance from WNYC.](http://www.wnyc.org/story/fewer-faces-nyc-schools-after-storm/) It’s easy to pull down this data every day, but, pretty quickly, it’s going to get big enough that you will regret any cheating / manual labor you might have done in our earlier exercises today.

1. See if you can find where the data that powers the [Citibike stations map](http://citibikenyc.com/stations) lives.
2. Here is the same [data](http://kpq.github.io/sherp-31/assets/data/bikes.csv) for every 15 minutes on Wednesday, June 5.
3. Filter the data to look at a single station.
4. What time of day are the fewest bikes available system-wide?
5. Add a column for which [neighborhood](http://kpq.github.io/sherp-31/assets/data/bike-neighborhoods.csv) each station belongs to, and answer the same question for Midtown South, the East Village, and another neighborhood of your choice. 
6. Why might you want to collect this data yourself instead of just asking CitiBike?
7. In the [Google doc](https://docs.google.com/document/d/1QVt45d0JfwLfa0RQtjULr68jZeWFKqEN6Ij4RU5sGVY/edit) we used earlier today, make a list of questions we could answer with this type of data, either by itself, or by joining it to other data.