Group 2: Zika Visualization Proposal

Please prepare a very short (no more than 1 or 2 pages), informal summary of what you plan to do as a group, and upload or email it to your instructor by the week six live session. Address the following in your write-up. This will not be a graded assignment, but it is important for keeping your project group on track. Your instructor will give feedback, and they may ask you to change your scope if they think it's too much or too little for this class.

● Names of students in your group

* Jonathan Landesman
* Rock Baek
* Kevin Davis
* Adam Spitzig

● What is your project concept?

The CDC has created a [github repo](https://github.com/cdcepi/zika) with files scraped from the websites of national health ministries of countries with Zika transmission. There appear to be a dearth of maps and other visualizations of this information that present the data in an intuitive and easy to understand manner.

The CDC provides a single map, indicating the number of cases of zika by state, which can be found [here](http://www.cdc.gov/zika/geo/united-states.html). The map only shows the cumulative total number of cases, providing no time dimension.

On a global basis, the best visualization I have found is at [healthmap.org/zika](http://www.healthmap.org/zika/#timeline). While the time series of Zika distribution is easily captured, the actual map is difficult to see.

Our concept is to create a system that automatically pulls the latest data from the github repo and presents a series of maps and graphs to make the data easy to understand in a single glance.

As a stretch project, we also have a dataset of flights to and from Brazil in the coming months. Hence, under a few assumptions, we can make a forecast of where Zika might go following the summer olympics in Brazil. If we have the time, we will make the projection and incorporate the data into the visualization.

Regarding structure, we plan to create an AWS instance for our data processing that automatically pulls the latest information from the CDC’s github repo, processes the data, and launches the live visualization. Tableau will likely be the tool of choices to create the visualization, though rcharts d3 wrapper is another option as well.