Overview and Motivation:

Zombies are cool. Everyone is fascinated by zombies. Given that zombies are such a big part of our lives (in the entertainment industry; thank God we don't have to actually deal with them yet), we figured this would be a great way to look at how a pandemic, such as a zombie viral/attack, would spread and affect the world.

Related Work:

A recently published paper by researchers at Cornell served as the primary inspiration behind this project. We figured that taking their model and building upon it seemed like a great way to do pandemic mapping in an interesting way. Their visualization for a zombie attack also gave us some further ideas for the project. The nukemap visualization that we had seen before--used to map a nuclear explosion of different yields, anywhere in the world--gave us some other ideas for the interface as well. Our goal is to expand what was done at Cornell and map diseases on an international level.

Questions:

The primary question that we were looking to answer is how a zombie virus would spread throughout the world, based off of different starting points. We generalize this to the idea of disease being spread across the world.

Data:

There are seven types of data that we will, primarily, need:

- [Country population densities]
- [Migrational/International travel data]
- [SIR model variant]
- Country geojson data
- [Worldwide airport data]
- [Country area information]
- [Reverse Geocoding]

To load in data we mostly use D3's csv loading function. We needed to mark the titles / keys for some of the columns like the airport data since d3.csv assumes that the first line of an csv is the keys for the columns. The cleaning and loading are done once the the HTML page is loaded. We load each data values into javascript objects for those. So area goes into an area object, airport data into airport, and so on. The keys are in general names of countries so that if a lat / long point is selected on the map, we can use a reverse geocoder (from Google Maps which I guess is another form of data we're using) to find the country and then the above data to model how many people are there, how many "blocks" of people we can have, and how to model the SIR dynamics within the blocks.

Exploratory Data Analysis:

In the beginning we were thinking about using a simple map and not showing data at a more granular level. However, we realized this was in conflict with some visualization ideas we had, like using Google Maps Street View, which would only be useful if we had very granular data. We decided to switch away from Google Maps for this reason because

Design Evolution:

We will add more details to this as we move into the final phases of the project. As of now, it looks like our design has not changed too much from what we were originally planning. If this changes with time, we will update this section.
We discussed with our TF and realized he wanted multiple views with brushing so we thought about how to add that in. There is a map component which is our main component so we thought it would be natural to brush on that. This does force us to realize the “discrete population blocks” idea that we had for our to model the spread of the disease and visualize it. What we mean by that is we have data that tells us population densities for various countries. We can try to break the country down into one square kilometer blocks where since we say the country had 50 people per square kilometer we can now say that there are 50 people in the block. The block itself will follow the normal SIR dynamics which will be simple to implement but infected individuals will spread into neighboring blocks with a certain probability.

When it comes to spreading the disease outside of a block that it has an SIR interaction in, we tried to think about what would be the most faithful way to do it with the model? Well we have two ways for infections to come into new blocks. They can spread from a neighboring block or they can come from a whole other country through an airport point. For the neighboring block version of spreading the disease we were thinking that we’ll find the frontier of unaffected places around a set of infected blocks. Together, this mega-block (the blocks with infections and the most immediate neighbor blocks without infections) will be the most important base for modeling the spread of the disease. We can run the SIR model for a step on the megablock and any new infected people we might spread probabilistically across the megablock. This way, we can have the disease spread with accordance to the SIR model and also with respect to this understanding we have that diseases spread locally.

We did not specify the need for airport data in our proposal, but we thought our model would be better if a disease enters a new country that does not neighbor any infected countries by way of airports. We already have migrational data so we can model how many people might travel from one country to another, but now we'll only let infected people enter a new country through its airport.

Brushing needs to be implemented and we'll think about ways to incorporate it. One idea is to brush across the map and make a bar chart of percentage of people susceptible, infected, and recovered in the area.

Map projections are something we want to think more about. The map lecture was useful since it helped us realize that the current Mercator projection we are using is not an accurate representation of area in the world. We will use Mercator to establish the base of our code and then fit a more reasonable and accurate project.

Implementation:

We decided that what we were aiming to accomplish is best completed using d3. We will be able to give more details about the implementation once we are further along the project.

Evaluation:

We will be able to better answer this when we are done with the assignment.