Project 2: Giardiasis (Giardia Infection)

By Melinda Wiley, Patricia Escalante, Yanjun Zhou

The George Washington University Data Analytics and Data Visualization Boot Camp

04 January 2019

Proposal

Project 2 will utilize Python Flask powered RESTful API, HTML/CSS, JS, one database (MySQL or SQLite) with user-driver interactions. Data munging and data visualization utilizing CDC provided dataset on Giardiasis (Giardia) from the National Notifiable Diseases Surveillance System (NNDSS). We narrowed the field to this disease caused by parasites to focus on building out our site with a more manageable dataset. Giardia csv and US population by State csv will be read into a SQL database. We will scrape the index from https://www.cdc.gov/parasites/az/index.html#s to pull the list of parasites; working these aspects in jupyter notebook with BeautifulSoup.

Initial layout will include US maps, chart, and possibly some links to the CDC site for more information. The visualizations will show infection cases over the years through mapping and a bar/bubble/line chart with some rad transitions. The 1 JS library not previously covered will be to be determined.

Resources, so far: <https://www.cdc.gov/parasites/az/index.html#s>, <https://data.cdc.gov>, https://www.cdc.gov/parasites/giardia/index.html

Topic: Giardiasis (Giardia Infection)Dataset: National Notifiable Diseases Surveillance System (NNDSS)

Git repo: <https://github.com/Yanjun7/Parasites>

sketch of the design:

Line charts with transitions on raw data, per capita data, and possibly regional data

X=time; y=# of Infections

Bubble chart over time x=population, y=# of infections

D3 charting with transitions and interactions

 