Benjamin Arancibia

10/26/2014

Semester Project

*Please write a paragraph, high level, description of what you think you may want to do for your semester project.  Include what kind of data you want to work with, and any other details you want.  You will not necessarily be bound to what you write, and the specifics can be subject to modification.  This is just to get you started. Due 10/29/2014.*

*It must use some kind of data, such as the data we used in class.  Data can be acquired from public sources such as data.gov.*

*It will need to have a visualization element.  That could be graphs or images.*

*It should be analytic in nature.*

*It should be different than what we have done in class.*

The Semester Project that I am hoping to tackle is an in-depth analysis of foreign aid in Honduras. All known foreign aid projects are accessible via an open data portal found at this website: <http://pgc.sre.gob.hn/>, also I have access to the server that hosts the DB for the foreign aid management system called AMP used by the Government of Honduras. Full disclosure I have access because my company (Development Gateway) built this software and I am currently the Project Manager for our project with the Government of Honduras. The reason why I want to use this data is because I am spearheading the increase in data analytics at my company and trying to convince them that we need a dedicated data analytics team. In terms of the necessary requirements for the project, see below.

Data: Honduras Foreign Aid (already have), Honduras (population density – geospatial data), Other TBD

Visualization: Geospatial, correlation between donor activities

Analytic: My idea right now is to look at donor coordination of foreign aid projects in Honduras. As background, donors tend to do projects in the same administrative levels but do not coordinate with each other on projects. Many times the same type of project is implemented in the same location, i.e. Japan and Germany will both do health projects in a certain administrative level. I want to look at this donor coordination and use other data to do a deeper analysis of the data that the Government of Honduras has in the database. I have foreign project data, locations of projects, population density of Honduras (geospatial), and I will be looking to identify other data sources.

Right now I have a general idea what I want to do, but depending on the data could go different routes for analysis. The amount of data is not “Big Data” only about 450 projects with 3,500 locations, but the data is incredibly complex. The Postgres DB has about 250+ tables in it, with a lot of different fields nested within other fields (i.e. Health Sector, Natal Health within overall Health Sector, and Maternal Natal Health within the Natal Health and the overall health sector).