Benjamin Arancibia

December 1, 2014

IS 607

Project Proposal

For the IS 607 (Data Acquisition and Management) Final Project, three different datasets will be used. These three datasets are financial foreign aid commitments to Honduras, locations of financial foreign aid projects in Honduras, and an indicator level dataset with various types of data. The indicator level dataset has fields such as number of prisons per municipality, hospitals per municipality, and population density. The data science workflow will be the acquire the data (two datasets in csv and one in PostgreSQL), necessary transformations (joining locations and projects based on project ID) and cleanup (removing extra fields), and then performing analysis (combining indicator level data with financial cooperation data). Graphics will be created using ggplot2 and possibly using a geospatial packages such as maps and mapdata. The reason why this data is being used for analysis is to further the development of a data analytics service and data team in my current company Development Gateway.