**Project Proposal: Are Commerce Firms Targeting Poor Areas for Bad Jobs?**

**Motivation:**

Amazon recently placed a fulfillment center in Shakopee, MN. Suburbs have seen recent growth in these company outposts, but are certain neighborhoods with specific demographic or economic traits being targeted? This project will explore the possible correlation between the locations of commerce company distribution centers, and the income, racial breakdown, or immigration status of those areas in 2000, 2010, 2012, 2013, 2014 and 2015. Distribution of facilities will be mapped and spatially analyzed as related to demographic markers. The purpose is to show if, and to what extent, these companies are located in specific types of neighborhoods, and what the traits of those areas are.

**Methods and Timetable:**

The first component will be gathering and sorting a large amount of data sheets. Automation of the download process with be attempted if time allows. Several years of data will be needed, which will allow an analysis of demographic trends through time. IPUMS offers aggregate data on immigration, race, and income at ZCTA geographies. American Community Survey, American Fact Finder, and Census databases have similar data at finer temporal granularity. The Bureaus of Economic Analysis and Labor Statistics offer data on industrial placement values. The next step will be calculating location coefficients (LQ’s) and other density measures for ZIP code/ZCTA geographies. Spatial analyses will be run on those coefficients, and demography measures. All results will be analyzed for trends in demographics and economic markers. Causality will not be proven but a correlative effect can be. The project will include a series of maps showing density measures by geography. The CSV module will be important to sorting downloaded data. A module like Beautiful Soup is required to web-scrape this data. pySAL will help with models. Numpy and SciKit are also useful for calculation heavy processing.

|  |  |  |
| --- | --- | --- |
| **Done** | **Code Elements Completed** | **Paper Section Completed (in syllabus)** |
| 10-17 | Project Proposal | “Introduction and Statement” |
| 10-24 | Data Selection | “Data Gathering and Sources” |
| 10-31 | Data Sorting | “Methods and Code 1” |
| 11-7 | Data Joining | “Methods and Code 2”, “Data Analysis 1” |
| 11-14 | Data Joining | “Data Analysis 2”,“Presentation” |
| 11-21 | LQ, Mapping and Spatial Analysis | “Challenges”, Report Draft |
| 11-28 | LQ, Mapping and Spatial Analysis |  |
| 12-5 | Presentation Due, Revisions | Report Draft 2, Revision Memo |
| 12-12 | Last minute code revisions | Report Complete |

**Deliverables and Results:**

Expected final result is a conclusion about correlation direction and strength between distribution center clusters and low income, immigrant, or minority neighborhoods. A table, maps, and graphs of these correlation measures and regressions are the major deliverables. Maps of distribution center LQs will also be produced. Other necessary deliverables are a report on the project process, clips of 4 major code scripts, and a class presentation of maps/results.