Project Proposal
Felipe Buchbinder – October 10th 2019

OVERVIEW

In Brazil, states and cities are obliged to disclose to the general public information on their financial situation. Brazilian Ministry of Finance requires that these reports follow a standardized accounting plan, which means data from different cities can be directly compared to one another.

I have a hypothesis: I believe that cities copy each other's accounts, just so they don't fail to submit their reports. My goal in this project is to investigate this hypothesis.

I only investigate data on balance sheets. Other financial documents are available but will not be considered in this project.

RESEARCH QUESTIONS

My initial EDA suggests that yes, some cities do copy each other's financial accounts, at least in part. Over 10% of cities copy on average 2 accounts from their peer cities. I am thus led to the following research question: Which variables explain why some cities copy other cities financial reports? Extant literature in the field suggest some variables to be investigated, such as city's size (measured as inhabitants or as GDP) and the distance between copying and copied cities (possibly moderated by whether cities are similar in size to begin with). I add to these variables by noting that some cities outsource the elaboration of their financial reports to accounting firms and that the same firm can sometimes prepare the reports of a handful of different cities. Thus, I theorize that the accounting firm may be behind these copying of financial accounts and, therefore, we would expect such copying practices to be more frequent among cities which share the same accounting firm.

DATA

I'll be using a few different datasets from different sources (all of which ultimately belong to the Brazilian government, so all data is official). Namely, I'll used data from: (1) Brazilian cities balance sheets; (2) Submission dates of the balance sheets; (3) Cities' latitudes and longitudes; (4) Cities' GDP and population; (5) List of the accounting firms who sign the financial reports. My GitHub repo has a document describing precisely the sources of each dataset.

This should give me data to construct the following variables: (1) Which cities copy other cities reports, (2) how many accounts are copied, (3) distance between "copied" and "copier" cities and, importantly, whether they are neighboring cities, (4) accounting firm signing the balance sheet, (4) GDP and population, (5) region.

PROJECT PLAN

I'll explore a logistic model to predict what explains a city copying another city's financial accounts. I may have to adopt a hierarchical model with random effects on the region of the city, to account for the fact that the distance between cities may have different effects depending on the available infrastructure on the region. Alternatively, I may have to adopt random effects on the level of the state. Time permitting, I may also investigate a (possibly zero-inflated) Poisson regression to explain how many accounts are copied by each city, if I find a considerable variability in this variable.