**Capstone Projects**

**David Mayer**

1. **Sectoral Municipal Employment in Mexico**

**Introduction**

Investing in a new business or setting up the local conditions for fostering local business growth requires combining lots of information that is not always available or easy to understand. We seek to provide a service that can be applied to any business sector and work with a well-defined, available data set that can provide sufficient information without the need to engage in costly surveys.

The National Statistical Directory of Economic Units (DENUE is the Spanish acronym) is a listing of businesses in Mexico (5,004,986 for 2016) that provides a particularly useful dataset. It includes information on each firm’s activity, level of employment, geographic location, and local setting. I am planning to use data on 2012 and 2016, which can be used to predict sectoral employment in 2020. While there is a considerable amount of survey data on Mexico, it is generally representative at the state or national level. Since this data set instead approximates a full listing of businesses, it implicitly contains a lot more local information.

We can combine this information to obtain aggregate sectoral employment data at the municipal level (there are about 2,400 municipalities in Mexico). Since employment allocation is a key economic variable, local sectoral employment levels implicitly carry a lot of information on local living standards, production, technological development, tastes and needs, and so on. Since not all determinants of business are local, we can also aggregate the data to the state level to include more extended indicators of business activity. This dataset can be supplemented using other municipal level data on geography, communications, sanitation, health, public policies, etc., to complement the information.

The idea is to train algorithms predicting 2016 employment levels based on the 2012 data. Then we can use the 2016 data to predict 2020 employment levels. We can then see where it is expected that local businesses will grow, new sectors will be introduced, and so on, and this in turn can serve as a guide for investment. It will also be possible to select municipalities according to characteristics such as mainly rural, mainly urban, and so on.

The results can be output as maps or data sets displaying expected levels of growth. A statistical summary can also be made of the main indicators that predict each level of growth.

**Clients**

The application can be useful for entrepreneurs, policy makers and academics. It will allow asking questions about municipal growth of employment sectors or combinations of sectors over different types of geographical locations. For this first approach we will concentrate on macroeconomic questions.

In the case of the academic user, the statistical output could be suitable for verification using other econometric methods. If there is enough time, key lagged municipal data (by for example a decade) can be used as instruments to test the statistical reliability of the main indicators that predict membership in different levels of growth.

**Outputs**

Maps, statistical summary graphics and summaries, relevant datasets.

**Data**

The main DENUE data set includes the following information

* **Firm’s activity** according to a six digit classification with 795,124 entries, which is also meaningful at the 2, 3, and 4 digit levels
* **Number of Employees** in intervals [0, 5], [6, 10], [11, 30], [31, 50], [51, 100], [101, 250], [250, ∞)
* **Location** down to the geostatistical area unit level (AGEB)
* **Type of street/road** in 30 classifications
* **Type of locality** in 43 classifications
* **Type of setting** fixed or semifixed