

# Urban Growth and Unemployment in Washington State

Thomas Nolan

## Summary of Research Questions:

1. *How can we describe the relationship between urban growth and unemployment in Washington State?*

Computation: Plot counties on a graph with unemployment and urban density growth as the axes to display the relationship, and compare to the average urban growth rate for cities across Washington.

Result: Although the county experiencing the highest rate of urban growth between 2000 and 2010 also experienced the highest rate of growth, the plot overall appears random rather than correlated. The Pearson correlation coefficient is 0.26034262481174902, implying a slight positive correlation between rates of urban population growth and unemployment rates.

2. *How could this research be refined in terms of time-scale?*

Computation: Plot the average unemployment for each county for each year in the time frame of 2000 – 2010 to get a sense of where the urban population density data could be disaggregated in order to understand the processes on a more natural time-scale.

Result: The plot tells us that the unemployment rate increased from 2000-2003, decreased from 2003-2007 to its lowest point, and increased again from 2007-2010 to its highest point.

## Motivation and Background:

In times of economic hardship, unemployed individuals seeking jobs may move to cities, which are perceived as centers for employment. It is possible that this perception is bolstered by a city's pattern of growth. Cities with higher rates of growth make be more attractive for

potential jobseekers who perceive a “boom” in a city’s population as an indication of greater economic activity and higher prospects for employment. For urban planners and other concerned parties in the private and public sectors, understanding how the growth of cities affects the livelihoods of citizens allows for decisions to be reached with more informational context. In light of greater understanding of how urban population density affects rates of unemployment, the process by which resources are allocated can not only be improved to distribute based on need, but also to predict future need and plan accordingly so that resources can be ready for deployment when there is likely to be a great need for them.

### **Datasets:**

1. Percent Change in City Population Density Since 2000: Based on WAOFM - Population and Housing, Census 2000 and 2010 (from the State of Washington)

Exported in the form of a .csv file, this dataset contains entries for all cities in the state of Washington for total population in 2000, 2010, and the change between the two years. This will need to be modified before it is used in the program, with a new column added for the county within which each city is situated. I will use other data in order to do this, going through each of the cities listed on the following table and updating the new column with the corresponding county. ([http://en.wikipedia.org/wiki/List\\_of\\_cities\\_in\\_Washington](http://en.wikipedia.org/wiki/List_of_cities_in_Washington))

<https://data.wa.gov/Demographics/Percent-Change-in-City-Population-Density-Since-20/j98i-s49g>

2. Unemployment Rates for Washington State and Counties, 1990 – 2010 (from the State of Washington)

Exported in the form of a .csv file, this dataset contains entries for all years between 1990 and 2010 for every city, county, and metropolitan service area in the state of Washington of the unemployment rate.

<https://data.wa.gov/Economics/Unemployment-Rates-for-Washington-State-and-Counti/hvq3-y2jb>

### **Methodology:**

The first goal of my analysis will be to produce two lists of tuples, where the first list follows the format: (County Name, Change in Unemployment Rate) and the second follows the format: (County Name, Change in Urban Population). The second element in the tuples in both lists is a float that represents a percentage. In order to calculate these values, the program reads the relevant .csv into two lists of dictionaries. In order to calculate the change in unemployment rate, it subtracts the relevant values keyed to the unemployment rates for 2010 and 2000 for each county like so:

$$(\text{Unemployment}_{2010} - \text{Unemployment}_{2000}) / \text{Unemployment}_{2000}$$

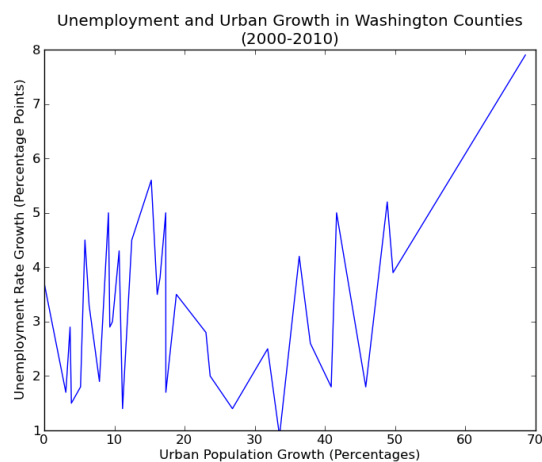
In order to create the second element in the tuples in the second list, the program simply looks up the relevant entry for each county in the list of dictionaries containing the urban population growth data, since there is a field there already containing the percentage change.

Then, once these two lists of tuples have been created, they will be sorted so that the counties appear in the same order in both lists. Once this has been accomplished, the rate of change for both variables will be extracted into two lists of floats which will be used both as x and y points to be plotted on a graph as well as the arguments to be used in a function from the SciPy library (`scipy.stats.pearsonr`) to determine the correlation coefficient between the two

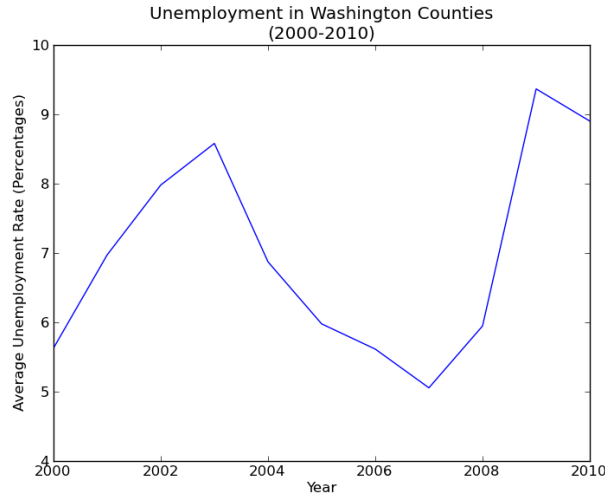
variables. The resultant graph will be saved in the current working directory as a .png file, and the correlation coefficient will be printed onto the console.

In order to answer the second research question, the program reuses some of the functions utilized for the first question in order to first produce a list of tuples mapping each year between 2000-2010 to the unemployment rate (a float) for that year. Then, the two elements of these tuples are extracted into two lists which are used as x and y points to be plotted on a graph, which is then saved as .png into the current working directory.

## Results:



This is the resultant graph from the program's computation of the first research question. Visually, we can discern that there is only a weak positive correlation between urban population growth and increases in the unemployment rate. In addition to this visual estimation, the program produces a Pearson correlation coefficient of 0.26034262481174902, confirming that there is indeed a weak positive correlation between the two variables, but for the purpose of our research this is not sufficiently convincing proof that there is any substance to this correlation. Thus, the results are not strong enough evidence for us to conclude that a faster urban growth rate alone can account for differential unemployment rates across the counties of Washington State.



Our program's computational answer to the second research question can be visually discerned from this graphical output. It is clear that there are significant shifts in the unemployment rate in the years 2003 and 2007, and so we should focus on these mileposts in future study. If urban population growth data can be disaggregated from the 10-year time scale available in the present working data so that we can isolate urban growth for the time frames of 2000-2003, 2003-2007, and 2007-2010 we may be able to find greater evidence to suggest there is a relationship between increases in the growth of urban populations and the average unemployment rate in any particular city. The aggregation level problem likely comes from the fact that the given data is was produced from census count data, which is only collected every ten years, but if we locate some other data source to provide this data we could go forward with further research using the second output of this program.

## **Reproducing the Results:**

In order to reproduce my results, you must load the relevant data (two .csv files uploaded to my dropbox along with the program) and the pop\_employ.py python program together into a directory. Then, after navigating to this directory, enter the following command into the command line: “python pop\_employ.py”. If either of the two .csv files are not in the current working directory, when the program is run it will ask you to load the data into the directory and then re-run the program. If the data files have been renamed anything other than the names they are given by default in the drop-box, the program will not recognize them as being present in the directory. The correlation coefficient referenced in this report will be printed to the console during the course of the program’s runtime, and the graphs referenced will be saved alongside the data and the program in the current working directory.

**Collaboration:** No students or course staff have helped me with my project, but I did consult this paper for guidance: <http://www.census.gov/prod/2003pubs/censr-12.pdf>

**Reflection:** Over the course of this project I have gained valuable experience in abstracting a problem into its base components in order to build the methodology for my program. As well, I have a greater appreciation for the volumes of data available for analyses of this kind, and the importance of research in guiding your understanding about future research based on the results of initial study.