Tianyi Wang
Metadata for ECON 428 final project
Professor O'Hara
Spring 2015

I. China city data

Department of Urban Socio-economic Surveys, National Bureau of Statistics of China. (2004-2011). *China City Statistical Yearbooks 2003-2010*. Beijing: China Statistical Press.

All Chinese city data used in this study are from China City Statistical Yearbooks from 2003 to 2010, compiled by China's National Bureau of Statistics (2004-2011). Because access to these data comes at a fee, I obtained it instead from Professor Zijun Luo at Colgate University on February 27th, 2015.

For year 2003 and 2004, there are three original data files for each year. For year 2005-2010, there are seven original data files each year. This is because that the dataset from 2005 onwards started to come in as a greater number of smaller data files.

- The data file *2003_1_original.xlsx* contains data on the following variables of each city used in this study: total employment level, GDP, GDP per capita, employment level in mining industry, employment level in manufacturing industry, employment level in utilities industry
    - Total employment level is defined as the total number (10,000) of employed persons in the city.
    - GDP is the gross domestic product (10,000 yuan) of the city measured in the current year.
    - GDP per capita is the gross domestic product per person (yuan/person) of the city measured in the current year.
    - Employment level in mining industry is the number (10,000) of people employed in mining industry in the city
    - Employment level in manufacturing industry is the number (10,000) of people employed in manufacturing industry in the city
    - Employment level in utilities industry is the number (10,000) of people employed in power and water generation and supply (i.e. utilities) industry in the city

- The data file *2003_2_original.xlsx* contains data on the following variables used in this study: total industrial output, total industrial employment level, output by domestic industrial firms, output by Hong Kong, Macau, and Taiwan industrial firms, output by foreign industrial firms, net value of total fixed assets
    - Total industrial output is the value (10,000 yuan) of total industrial output in the city

- Total industrial employment level is the total number (10,000) of employees in industrial enterprises in the city
- Output by domestic industrial firms is the value (10,000 yuan) of output by domestic Chinese industrial firms in the city
- Output by Hong Kong, Macau, and Taiwan industrial firms is the value (10,000 yuan) of output by Hong Kong, Macau, and Taiwan industrial firms in the city
- Output by foreign industrial firms is the value (10,000 yuan) of output by foreign (excluding Hong Kong, Macau, and Taiwan) industrial firms in the city
- Net value of total fixed assets is the net value of the average remaining industrial fixed assets in the city

- The data file *2003_3_original.xlsx* contains data only one variable used in this study: industrial sulfur dioxide (SO2) emission
    - industrial sulfur dioxide (SO2) emission is the total amount (tons) of sulfur dioxide emissions from industrial processes in the city

- Original data files of 2004 follow the same structure as those of 2003 described above.

    - The data file *2005_1_original.xlsx* contains data on the following variables used in this study: total employment level
    - *2005_2_original.xlsx* contains data on the following variables used in this study: employment level in mining industry, employment level in manufacturing industry, employment level in utilities industry
    - *2005_3_original.xlsx* contains data on the following variables used in this study: GDP, GDP per capita
    - *2005_4_original.xlsx* contains data on the following variables used in this study: total industrial output, total industrial employment level
    - *2005_5_original.xlsx* contains data on the following variables used in this study: output by domestic industrial firms, output by Hong Kong, Macau, and Taiwan industrial firms, output by foreign industrial firms
    - *2005_6_original.xlsx* contains data on the following variables used in this study: industrial sulfur dioxide emission
    - *2005_7_original.xlsx* contains data on the following variables used in this study: net value of total fixed assets

- Original data files of 2006-2010 follow the same structure as those of 2005 described above.

To convert the original data into importable data files, I made the following changes:

1) Translated names of the variables I am interested in from Chinese to English and erased other unused variable names (in Chinese)
2) Created a column containing the city names in English
3) Deleted other unwanted format features such as empty rows from the original data files

II. GDP deflator

GDP deflator. The World Bank. Downloaded on March 1st. Retrieved from: http://data.worldbank.org/indicator/NY.GDP.DEFL.ZS

To download this data, go to the above website, and then click on "DOWNLOAD DATA" on the top right corner of the page.

- The GDP implicit deflator is the ratio of GDP in current local currency to GDP in constant local currency. The base year varies by country, and for China it is the year 2000.

I used the following formula in my do file *cleaning.do* to deflate nominal values in the dataset to real values: real value=nominal value/GDP deflator