Analysis2:

In this analysis, loop and condition are used instead of panda. The output is the population of the country in the condition of headcount ratio urban is above 25 percentage ordered by the alphabet of country name. This is an analysis which obtain the values separately in two raw data. This can also be a reference value for the former one which’s output is about country’s agriculture, industry and service development.

There are 3 values are used which are country in column 2, headcount ration urban in column 24 and population in column 4. The ‘is\_first\_line’ is used to skip the first line in the CSV when loading the data. What the program do is create a list with the unique key, country, under the condition of headcount ratio urban is above 25 percentage and the value is the product of population in the country and count ratio urban.

Here is the screen shot of the output:

Analysis3:

Loop and condition are used instead of panda in the analysis. The output is the maximum agriculture rate every 2 percentage in the condition of intensity of deprivation rural is above 50 percentage. This analysis produces output that gives the aggregates over multiple subsets. Because there should be a strong relationship between agriculture rate and intensity of deprivation rural which are obtained from different raw data, the output can show a tendency of agriculture rate when intensity of deprivation rural is increasing.

There are 2 values are used which are agriculture in column 19 and intensity of deprivation rural in column 28. The ‘is\_first\_line’ is used to skip the first line in the CSV when loading the data. What’s more, because the data uses ‘,’ to present as decimal point, the function ‘replace()’ is used to change the ‘,’ to ‘.’ for python to convert the data to float. The key 0 means the intensity of deprivation rural is between 50 to 52, key 1 means the intensity of deprivation rural is between 52 to 54 and so on until 8.

Here is the screen shot of the output: