

## Assignment 2.

In this applied data science assignment, we use dataset of world bank that is based on different countries corresponding to years from 1960 to 2019. In this assignment we statistical analyze the data according to specific indicators in dataset. First of all, we review the dataset.

	Country Name	Country Code	Indicator Name	Indicator Code	1960	1961	1962	1963	1964	1965	...
0	Aruba	ABW	Urban population (% of total population)	SP.URB.TOTL.IN.ZS	50.776	50.761000	50.746000	50.730000	50.715000	50.700000	...
1	Aruba	ABW	Urban population	SP.URB.TOTL	27525.000	28139.000000	28537.000000	28763.000000	28922.000000	29080.000000	...
2	Aruba	ABW	Urban population growth (annual %)	SP.URB.GROW	NaN	2.206183	1.404498	0.788835	0.551271	0.544810	...
3	Aruba	ABW	Population total	SP.POP.TOTL	54208.000	55434.000000	56234.000000	56899.000000	57629.000000	57357.000000	...
4	Aruba	ABW	Population growth (annual %)	SP.POP.GROW	NaN	2.238462	1.432843	0.823502	0.580334	0.573488	...
...	...	...	...	...	...	...	...	...	...	...	...
20211	Zimbabwe	ZWE	Rural land area where elevation is below 5 mft.	AG.LND.ELSM.RU.ZS	NaN	NaN	NaN	NaN	NaN	NaN	...

In this analysis we use python library pandas. In this library we make the dataset in frames and easy to use in analysis. We have first made two frames one is countries and another is years.

### Countries dataset.

```
0      Aruba
1      Aruba
2      Aruba
3      Aruba
4      Aruba
...
20211  Zimbabwe
20212  Zimbabwe
20213  Zimbabwe
20214  Zimbabwe
20215  Zimbabwe
Name: Country Name, Length: 20216, dtype: object
```

### Years dataset.

	1960	1961	1962	1963	1964	1965	1966	1967	1968	1969	...
0	50.776	50.761000	50.746000	50.730000	50.715000	50.700000	50.685000	50.670000	50.654000	50.639000	...
1	27525.000	28139.000000	28537.000000	28763.000000	28922.000000	29080.000000	29246.000000	29411.000000	29570.000000	29742.000000	...
2	NaN	2.206183	1.404498	0.788835	0.551271	0.544810	0.569216	0.562594	0.539158	0.579985	...
3	54208.000	55434.000000	56234.000000	56899.000000	57029.000000	57357.000000	57702.000000	58044.000000	58377.000000	58734.000000	...
4	NaN	2.238462	1.432843	0.823502	0.580334	0.573488	0.599694	0.590951	0.572063	0.609680	...
...	...	...	...	...	...	...	...	...	...	...	...
20211	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	...
20212	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	...
20213	NaN	4.872690	5.001939	5.131188	5.260437	5.337986	5.518935	5.648184	5.777433	5.906882	...
20214	NaN	28.396019	28.615743	28.835486	29.055189	29.223213	29.494636	29.714360	29.934083	30.153806	...
20215	NaN	109850.000000	110700.000000	111550.000000	112400.000000	113050.000000	114100.000000	114950.000000	115800.000000	116550.000000	...

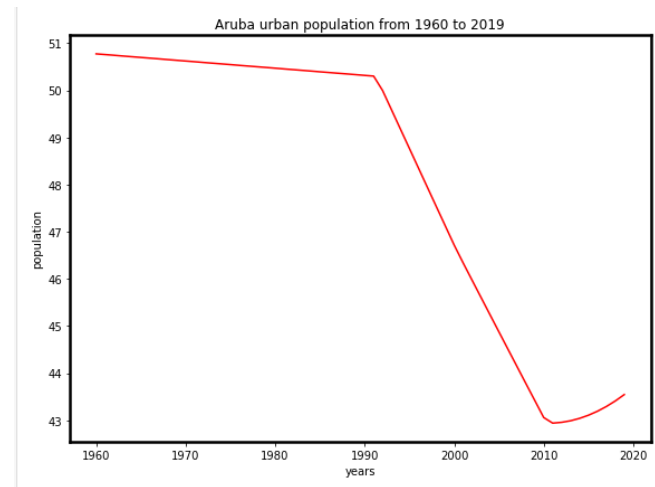
20216 rows × 60 columns

These are the major two dataset that we use in our analysis. First of all, we select one of following indicators. After that we choose different countries and perform analysis on same indicator at the end of analysis we use relation in indicators for same country here, we make analysis on urban population % of

Aruba, Australia and Belgium. We use pandas built in function to search in dataset and get only that row that we need or target specific row in dataset using pandas' function iloc.

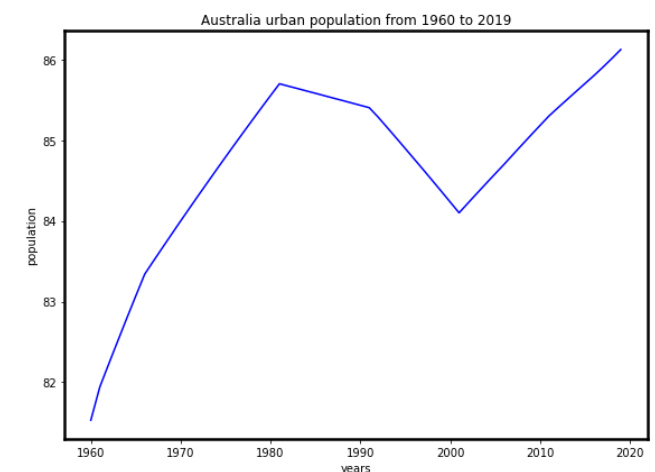
### Analysis on Urban population.

#### Urban population of Aruba.



This is analysis from 1960 to 2019 and we saw in graph that from 1960 to 2019 urban population is gradually decreased. So, it means that Aruba is developing country and most of economy is generated from development in urban sector so, according to this definition of success this is developing country and its economic performance is not so good.

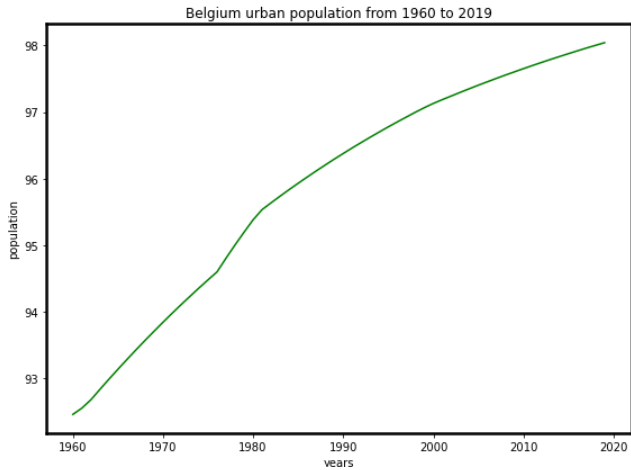
#### Urban population of Australia.



Here we have a country Australia. In Australia we saw from graph from 1960 to 1980 a great

development in urban development and after 1980 may be some issue regarding urban development Australia's urban population decreased to 1997 but gradually. But, after 1998 to 2005 urban population of Australia was decreased very speedily. After, 2005 Australia was a going to be urban country its urban population increased uniformly.

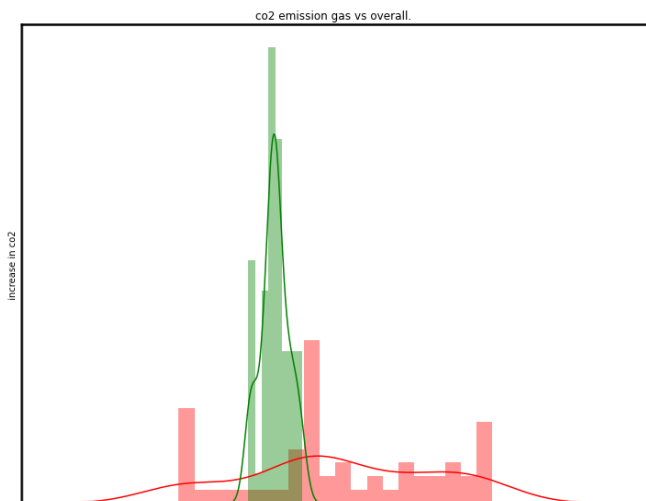
### Urban population of Belgium.



Belgium is developed country and when we saw it according to graphs this thing is cleared that there is now decrement in urban population in Belgium. Belgium is continuously, growing its urban population and this is a developed country in world. This is much for analysis on Belgium urban population. Now, we proceed towards the check correlation in different indicator of some countries.

### Co-relation.

### CO2 emission vs Co2 emission gas.



This is dataset of Belgium; green shows the co2 emission from gas and red shows the co2 emissions

overall. So, an according to this analysis we select only rows of France with indicator CO2 emissions from gaseous fuel consumption (% of total) and CO2 emissions (metric tons per capita). Now we start relating both in single graph. For some years from 19070 to 2001 there is too much co2 emission in Belgium but this is one of the developed country the controlled on co2 emission to save the nature. There is co2 emission is overall in county but decreased in graph and almost none in now a day.

GITHUB LINK.

<https://github.com/Shavitha-9/Statistics-Trends>