|  |
| --- |
|  |
| “The New World Order” Development Index  Statistics, Probabilities and Multivariate Analysis |
| |  |  |  | | --- | --- | --- | | Lazaros-Antonios Chatzilazarou | Universidad Carlos III de Madrid | MIEM, 2019 | |



**Scope of the Project:**

As an initial scope, this project focuses on the general measurement and definition of the development of countries. The variables considered were based on 2 major factors.

1. *Various previous studies which interpreted the development of countries with many ways,*
2. *In respect to an amount of correlation between the variables to be existent.*

Development in countries has been a quite interesting topic in growth and political economy. Geographic factors, GDP per capita as well as institutions are among many variables which are used to determine the development course of a country through time. This project focused on gathering variables in a way to form an index that would indicate **development** **in** **countries**. As there are many interpretations on how this index is formed, this analysis took a slightly different approach keeping though as fundamental pylons the GDP per capita and the population’s characteristics. **As an indirect and final question** this analysis came up to rise, was whether the “new world order” development, taking into account the amount of people moved to urban regions, that elimination in bureaucracy of the start-up procedure steps, as well as the rise of female power in high positions of employing.

**Variables considered and Data sources:**

The variables we considered for this analysis were all acquired form the World Bank Dataset. They are cross-sectional as it was indicated as a preferable choice from the instructions for this project and they are all listed below. Note that after eliminating missing values, the number of the variables considered shrank, as data for the 3rd World countries (majority of African Countries) are not available. Most of the variables collected in the form of percentages so the scale would be equal for the calculations and estimations and plots to be easier to interpret and understand.

Initial Set of Variables considered:

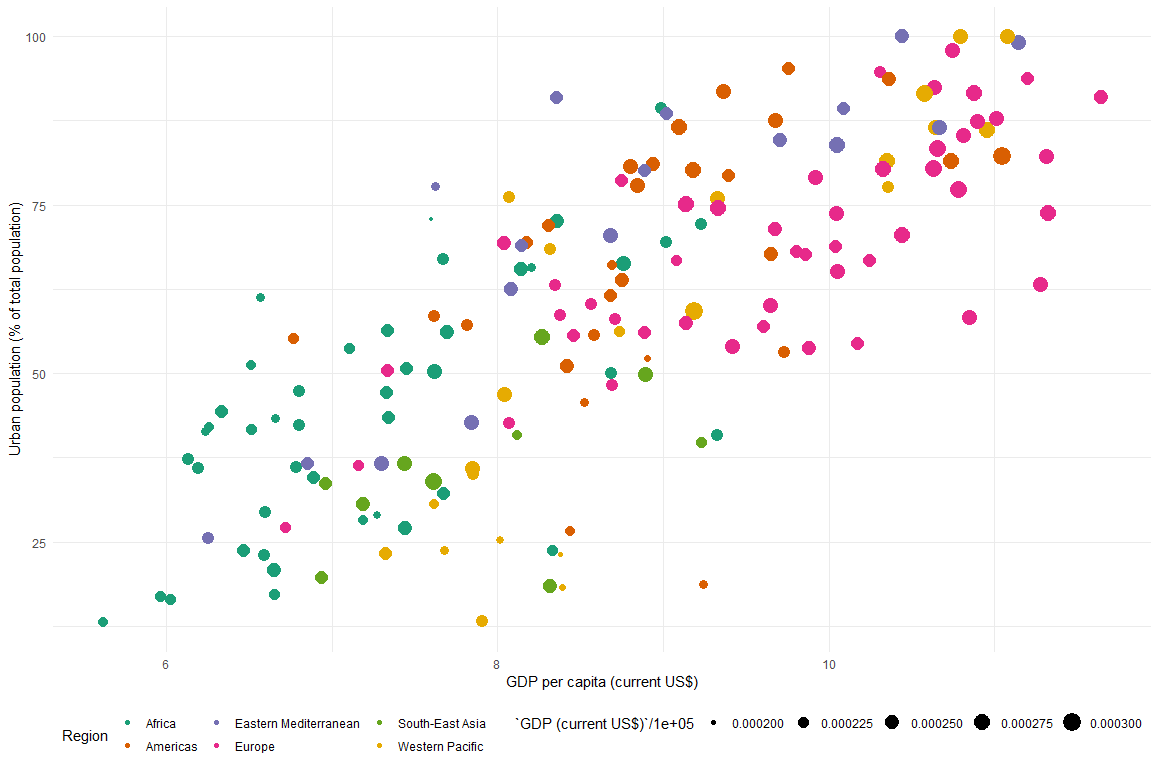
* *Year*
* *Country Name*
* *Region*
* *Employers, female (% of female employment)*
* *Employers, male (% of male employment)*
* *Employment in industry, male (% of male employment)*
* *Employment in industry, female (% of female employment)*
* *GDP (current US$)*
* *GDP per capita (current US$)*
* *GDP per capita growth (annual %)*
* *Labor force participation rate for ages 15-24, total (%) (national estimate)*
* *Life expectancy at birth, female (years)*
* *Life expectancy at birth, male (years)*
* *Literacy rate, adult female (% of females ages 15 and above)*
* *Literacy rate, adult male (% of males ages 15 and above)*
* *Population ages 65 and above, total*
* *Population ages 65 and above (% of total population)*
* *Population ages 65 and above, female (% of female population)*
* *Population ages 65 and above, male (% of male population)*
* *Population growth (annual %)*
* *Pupil-teacher ratio, tertiary*
* *Pupil-teacher ratio, secondary*
* *Pupil-teacher ratio, primary*
* *School enrollment, secondary (% gross)*
* *School enrollment, secondary, female (% gross)*
* *School enrollment, secondary, male (% gross)*
* *Start-up procedures to register a business (number)*
* *Unemployment, male (% of male labor force) (national estimate)*
* *Unemployment, female (% of female labor force) (national estimate)*
* *Urban population growth (annual %)*
* *Urban population (% of total population)*

Final Set of Variables that regard to the plotting and PCA analysis:

* *Year*
* *Country Name*
* *Region*
* *Employers, female (% of female employment)*
* *Employers, male (% of male employment)*
* *Employment in industry, male (% of male employment)*
* *Employment in industry, female (% of female employment)*
* *GDP (current US$)*
* *GDP per capita (current US$)*
* *GDP per capita growth (annual %)*
* *Population ages 65 and above, total*
* *Population ages 65 and above (% of total population)*
* *Population ages 65 and above, female (% of female population)*
* *Population ages 65 and above, male (% of male population)*
* *Population growth (annual %)*
* *Start-up procedures to register a business (number)*
* *Unemployment, male (% of male labor force) (national estimate)*
* *Unemployment, female (% of female labor force) (national estimate)*
* *Urban population growth (annual %)*
* *Urban population (% of total population)*

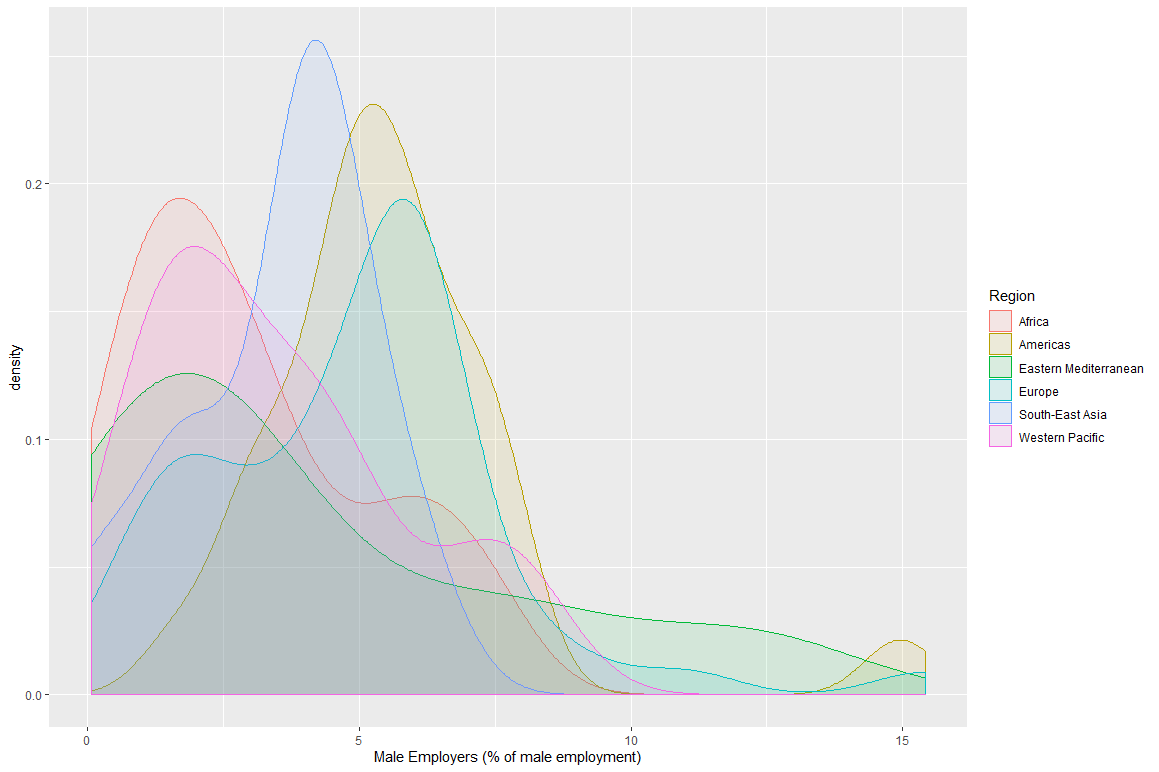
**Descriptive Statistics and Interesting Plots:**

In this section, we will present some basic, intermediate and advanced plots we created after the data cleaning process (it is thoroughly described with comments inside the R code file). *It is very important to state that during the data cleaning process we thoroughly scaled all the variables to be the same and i.e. the total numbers of GPD and GDP per capita etc. were transformed to their logarithmic form.*

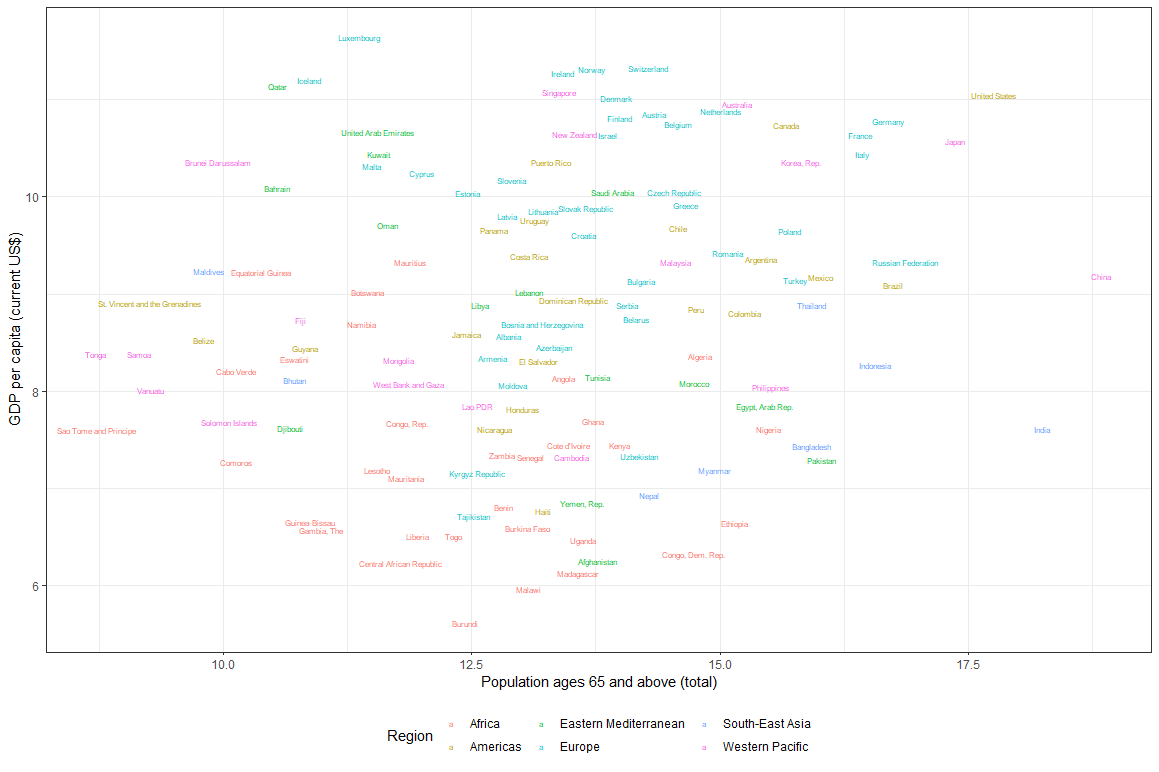
*1st Plot: Correlation between Urban Population and GDP per Capita*

* In this 1st plot we can see a clear correlation between the Urban Population of a nation to it’s GDP per capita. Since GDP per capita is considered by a lot an essential number for growth we compare it to the urban immigration within a country to see whether there are any relations. As we can clearly observe there is a positive relation, meaning that the higher the GDP per capita is the more Urban population the country has. The countries belonging in the European region (pink color), seem to have a positive correlation between urban population and GDP per capita while most African countries are at the bottom left of the chart, thus implying that since GDP per capita is a development indication, developed countries tend to be more urban criticized.

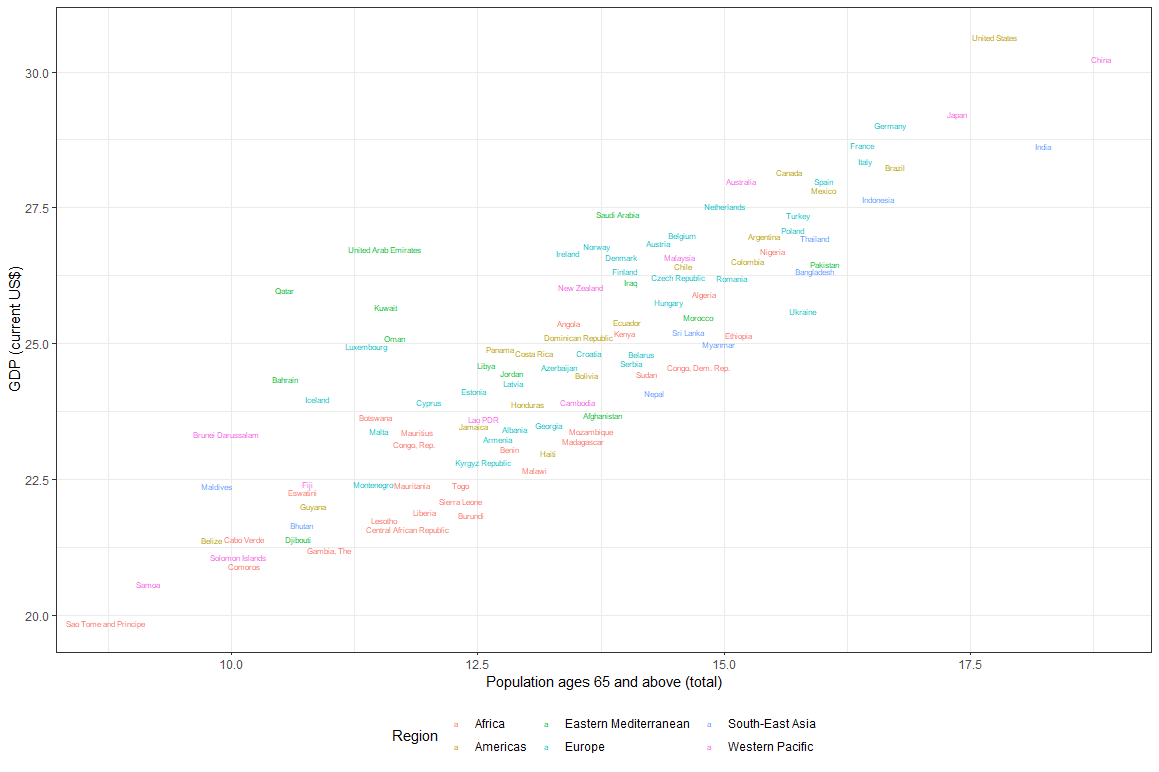
*2nd Plot: Density of Male Employers as % of male employment*

**

*3rd Plot: GDP per capita with Population over 65 years old*

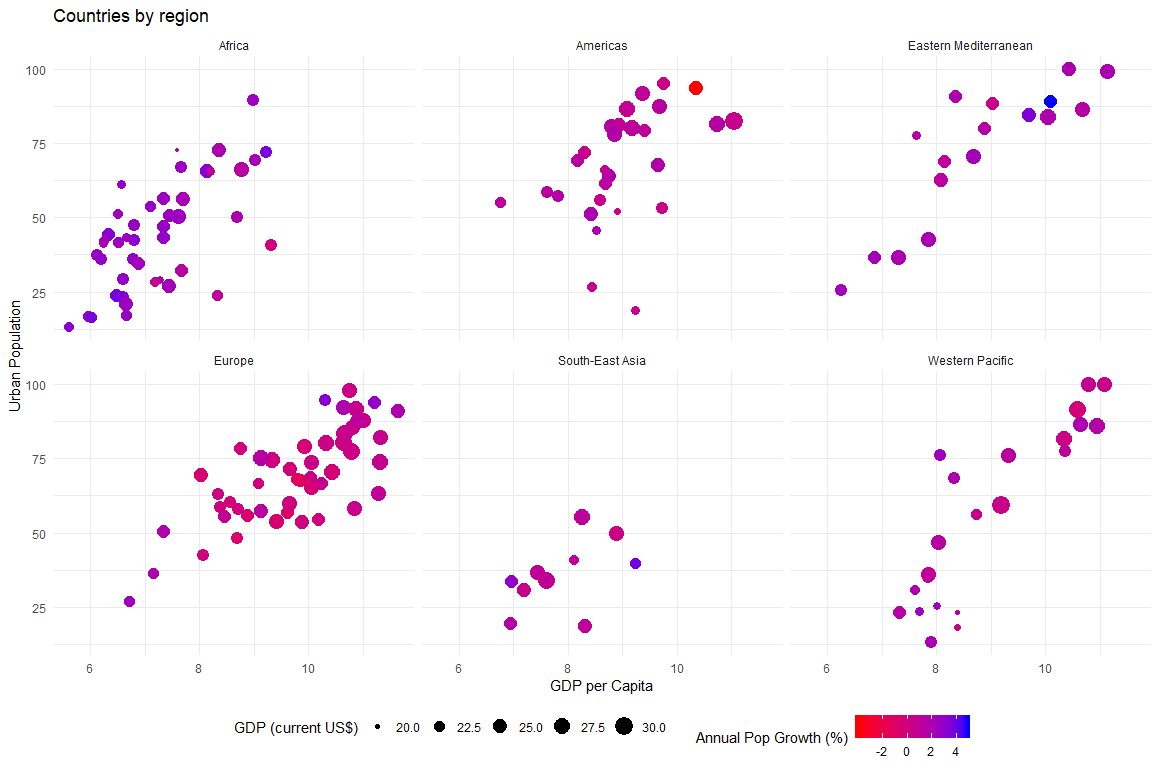
**

*4th Plot: GDP with Population over 65 years old*

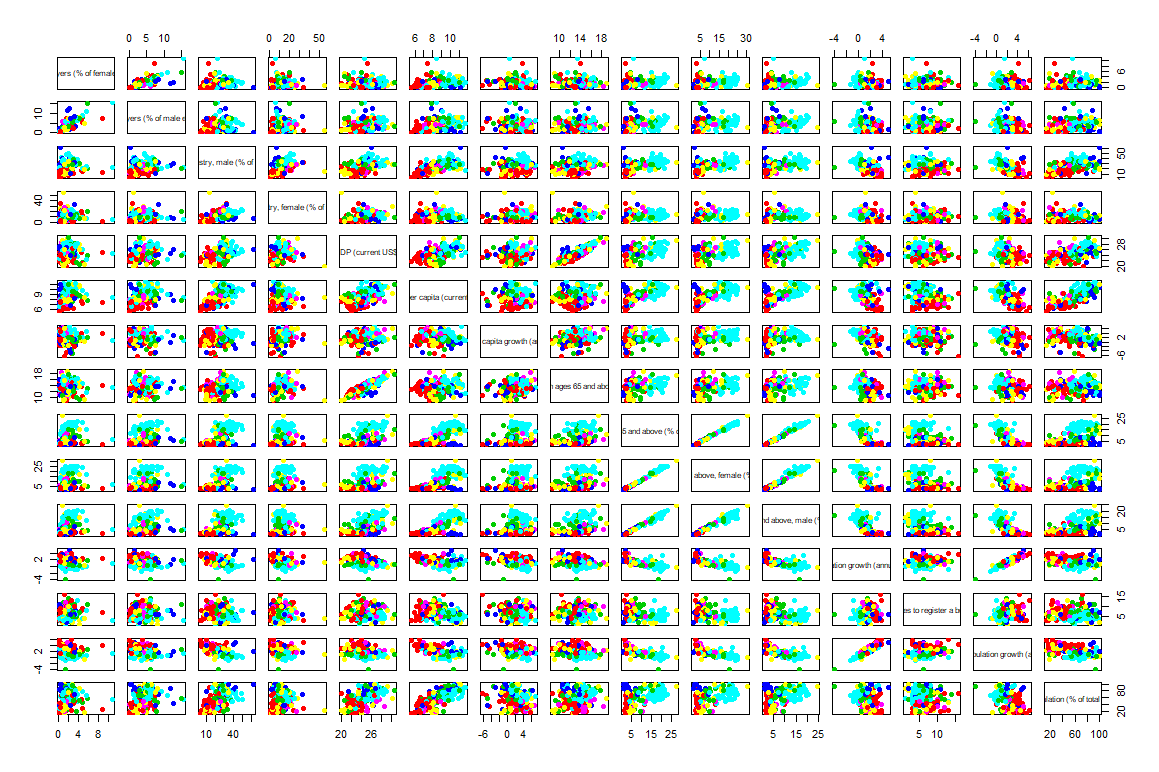
**

* On the 3rd and 4th plot we can see the correlations between GDP per capita and elderly population and GDP (total) with the same elderly population. On the first mentioned graph the correlation is not so clear as the countries are scattered across the 2-axis. On the second graph where we substitute the GDP per capita with total GDP of the country and the relations here are much clearer. We obtain a positive correlation between GDP and the above 65 years of age population, thus observing a lot of high population countries in the right top. In example, India and USA have massive total population and therefore their aggregate GDP is very high. Besides everything, the graph shows that the larger the proportion of the elderly people the larger the overall aggregate GDP.

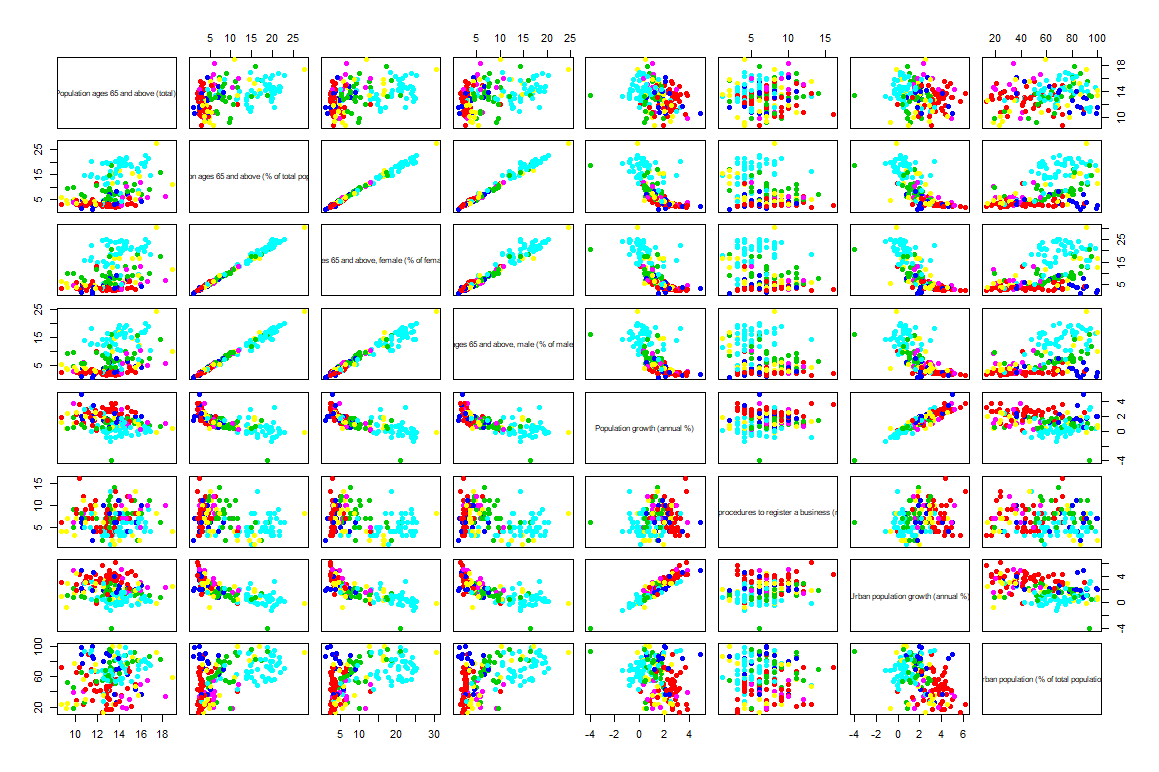
*5th Plot: Urban Population with GDP per capita, Regions*

**

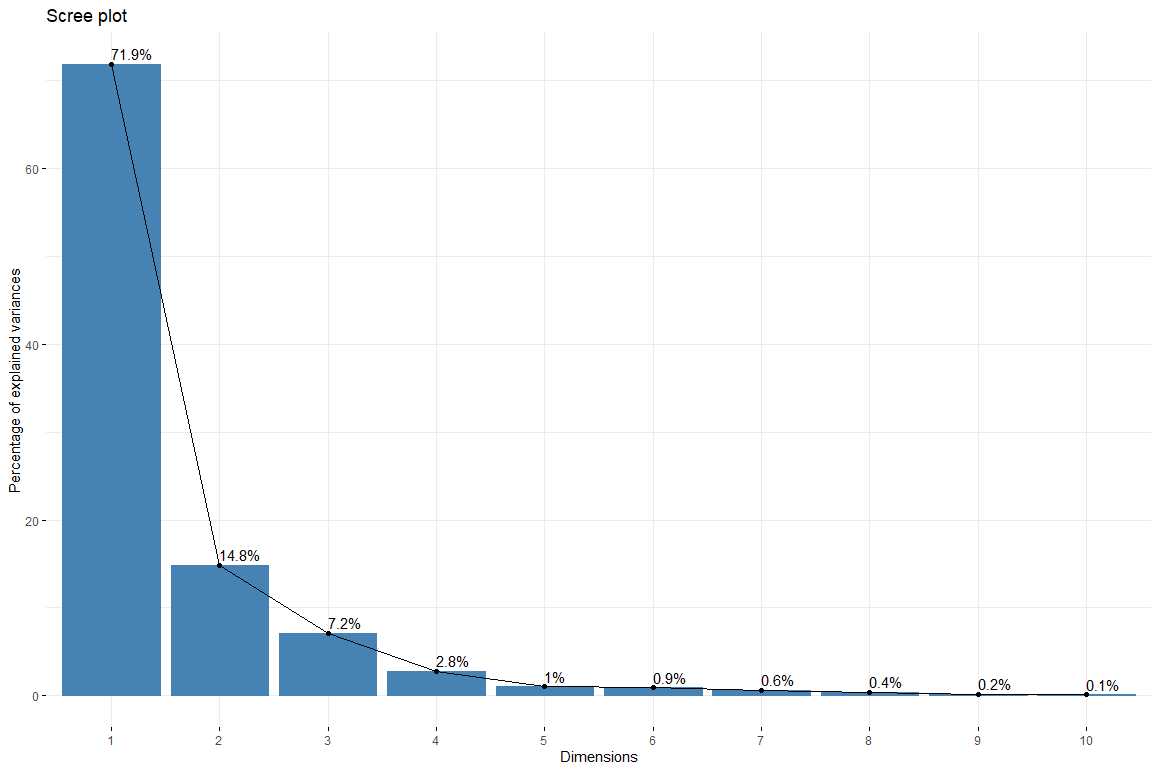
* A very interesting grouped graph showing the correlation between variables of urban population, population growth, GDP and GDP per capita. Very clearly one can see how Africa keeps dealing with massive population growth since no anti-birth controls have been introduced. Significant to say that the axis variables present to have a positive relation*.*

*5th Plot: Pairs*

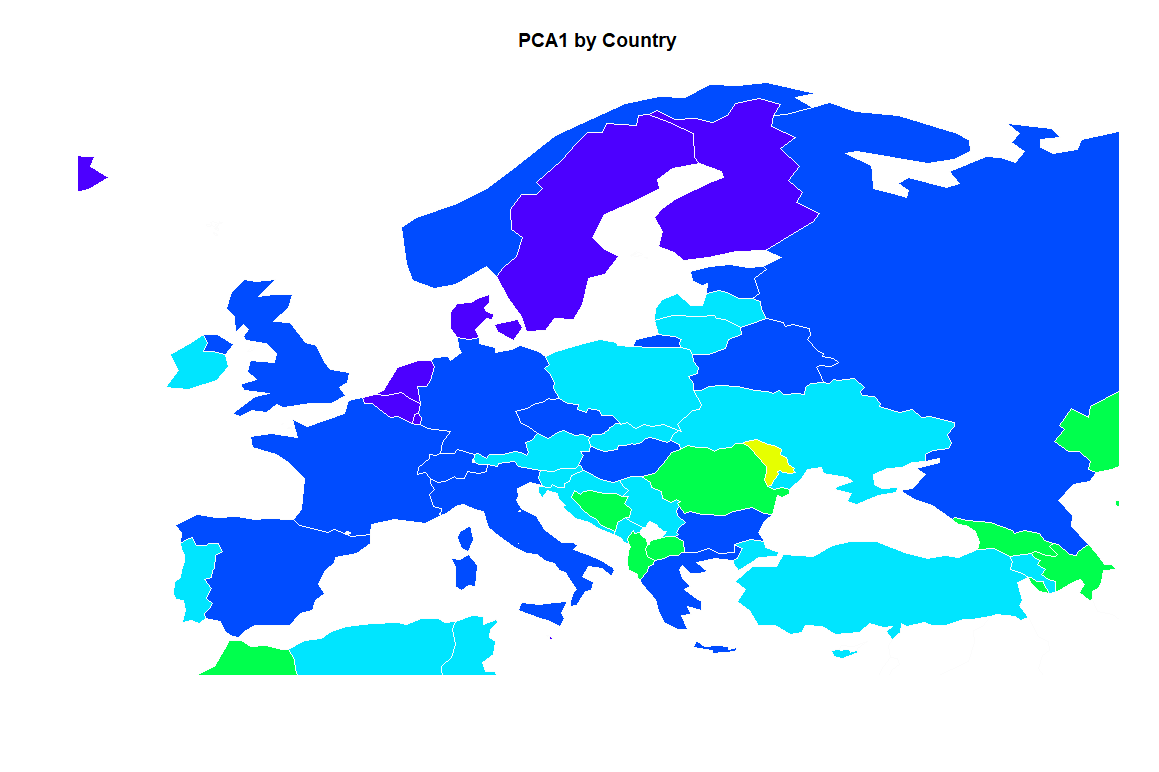
* This graph represents all the pairs that can be depicted in dimension 2 between all 20 variables of the dataframe. As the graph is not so informative due to the massive volume of information containg we plot another set of pairs for the last 8 variables below to show the correlations more clearly.

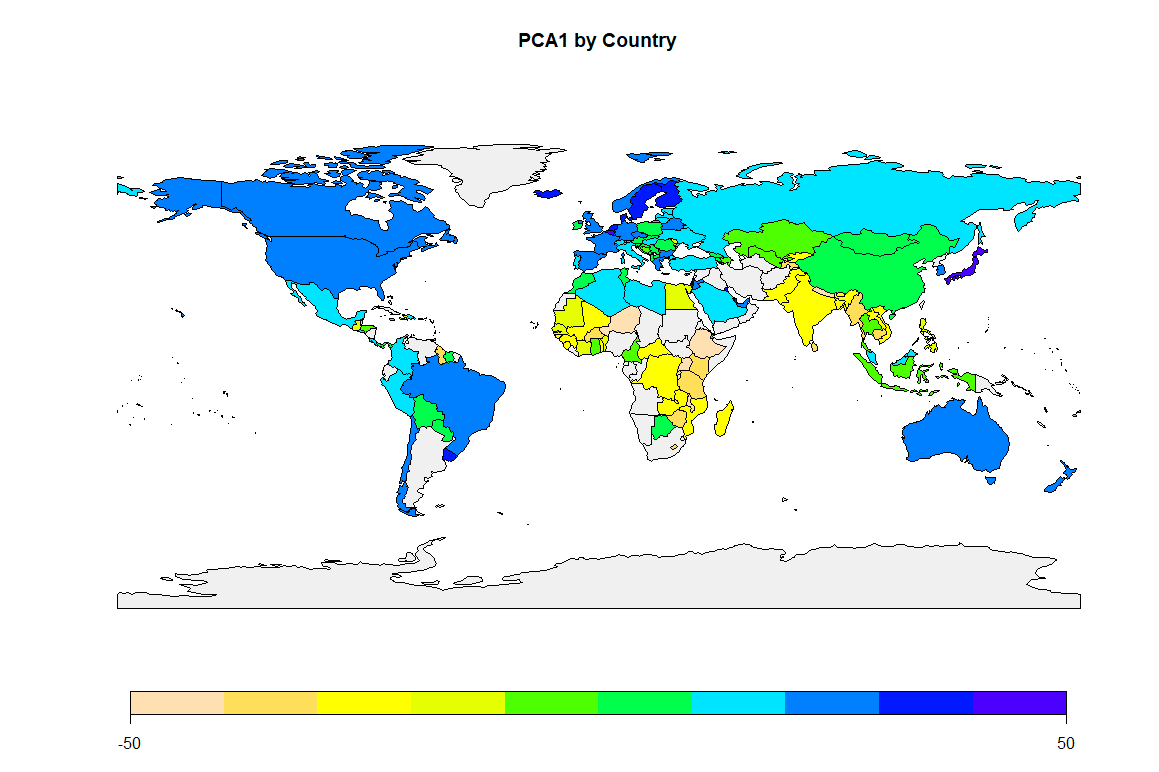
*6th Plot: Pairs*

* The above graph represents, as mentioned before, the pairs that can be depicted in dimension 2 between the last 8 variables of our dataframe. As we can see in quite a few relations there is an obvious relation whether linear or non linear but there are some cases such as the variable of procedures to register a buisness that show no clear pattern of correlation.

**

* Here, we can see a basic screeplot. “A **scree plot** is a line plot of the eigenvalues of factors or principal components in an analysis”. With this plot we can see the number of the PC’s for our analysis. The first PC has captured most of the information, followed by the 2nd and then we can ignore the rest without losing any important information.

**

**

**Final Remarks and Conclusion:**

The maps depict the principal components of our analysis. As we mentioned in the initial scope of the project, the formation of the development index was our main goal with some deviations implementing new perspectives the societies obtain such as the start-up procedure steps for business formation or the high power of women in the employment status.

The map outcome is quite satisfying as the PC explains the development of countries pretty good since the regions of being known to have a high Human Development Index and being developed agree with the color palette. When focused on the continent of Europe we can observe that countries such as France, Germany, UK, Italy and more are colored the maximum blue of the palette indicating that they are fully developed countries who have also implemented the new world order status. By that we mean, that the societies are moving into a direction of equal rights to both sexes and an empowered employing position of a woman as well as the high productivity of the countries which is depicted through startup generations and new ideas of evolving.