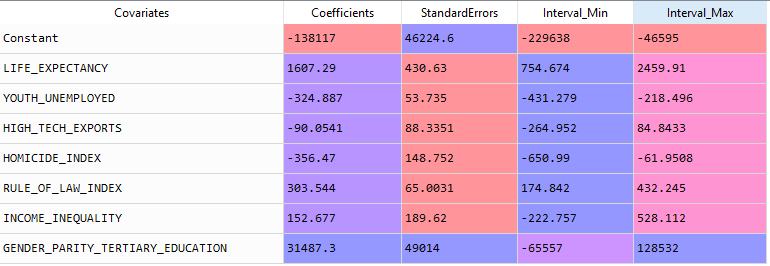
**CEM KIRAÇ - HW2 REPORT – Advanced data analysis in python**

My aim was to predict **gdp\_per\_capita** of countries in this study. I chose some explanatory variables and countries based on my personal judgement, whichI thought would have a correlation with gdp\_per\_capita. I tried to choose them from different fields to avoid multicollinearity, such as health, economics, education and law. The data source is the worldbank API. You can find the variables in the first table below. The 2nd table is the results of statsmodels library of python, just to check for if there are any differences between two linear regressions. I was surprised to see that there is no significant relationship between income\_inequality and gdp\_per\_capita. Similarly I was expecting tertiary education of women would mean a higher gdp\_per\_capita but that wasnt the case. Life expectancy and rule of law, both has a very strong relationship just as expected. Employment\_rate in younger people and homicide index are also significantly related to gdp. High\_tech\_export percentage relationship to gdp was insignificant.

**Linear Regression output**



**Statsmodels Linear Reg. Output**

