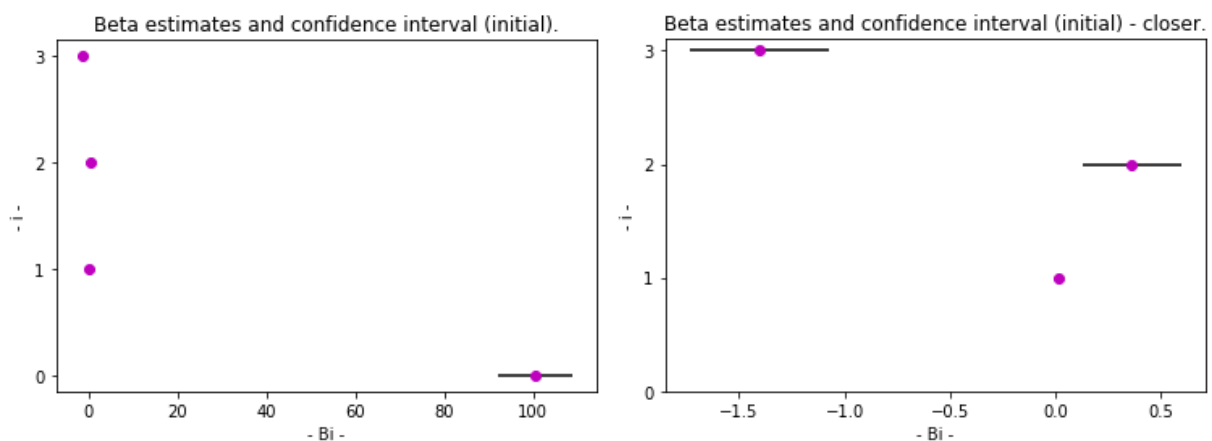# Report

Using the data from World Bank databases, some macroeconomic variables of Tajikistan approximately for the past 30 years were examined. The data for longer period are not available due to the fact that Tajikistan became independent from Soviet Union around 1990s.

In the regression model, the effects of GDP per capita, school enrollment rate, and birth rate of Tajikistan on its life expectancy for the given period was studied. For the period of 1990-2019, 3 out of 30 observations were dropped due to some missing data. The model is as follows:

$$lifexp_t = \beta_0 + \beta_1 gdp_t + \beta_2 school_t + \beta_3 birth_t + \varepsilon_t$$

The regression was run twice: once directly using initially obtained data and once using normalized data too see the difference between both methods. The plots and tables of coefficients with relevant statistics for both runs can be seen below.
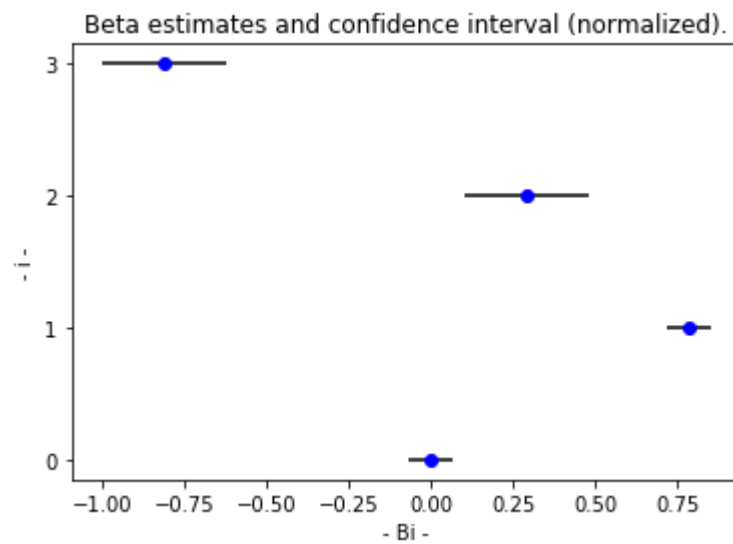

The side effects of not normalizing the data can be seen in the plots of initial regression. This way, the magnitude of one variable can overstate its effects in the calculations. In this example one coefficient is too large, whereas others are too small to be observed easily.

(Initial data)			
Bi	**SEi**	**CI1**	**CI2**
0 100.55341	0.011410	0.361585	-1.404274
1 4.039699	0.000480	0.113904	0.158944
2 92.19666	0.010417	0.125957	-1.733075

(Normalized data)			
Bi	**SEi**	**CI1**	**CI2**
0 2.321189e-15	0.784749	0.292014	-0.812019
1 3.288358e-02	0.033011	0.091988	0.091909
2 -6.802486e-02	0.716460	0.101722	-1.002147

The regression plot of normalized data is smoother and easier to observe. Moreover, the coefficients are more significant. The confidence intervals for each coefficient can be seen below:

Beta estimates and confidence interval (normalized).

To conclude, according to the results of regression using normalized data, for Tajikistan the increase in GDP per capita and schooling rate (but less) positively affects the life expectancy, where the birth rate is inversely proportional with it.