Questions

- My initial question is described in this forum post https://discussions.udacity.com/t/weighted-logistic-regression/196077.

  I have tried different things, but I do not understand which one is appropriate.
  I have tried to run regressions separately for each country (it is described in the post, code is here https://github.com/tommak/PISA2012/blob/master/py/logreg.py, functions calc_countries and calc_countries_reg).
  Then I have read about multilevel regression, but I did not understand how to use it for my case.
  Finally, I did one regression which includes interaction terms country*feature. I fit weighted regression with regularization. Regularization works fine, especially because some variables are very poorly presented for some countries. My code is in this ipython notebook https://github.com/tommak/PISA2012/blob/master/py/Weighted%20Regression%20with%20Interactions.ipynb.
  Visualization for this data is here https://github.com/tommak/PISA2012/blob/master/d3/lp_analysis.html.

- The second question is quite theoretical. Considering regularized regression, is it possible to calculate significance for parameters? Does it make sense or it is assumed that insignificant coefficients will just become very small or even zero (for l1 regularization)?