

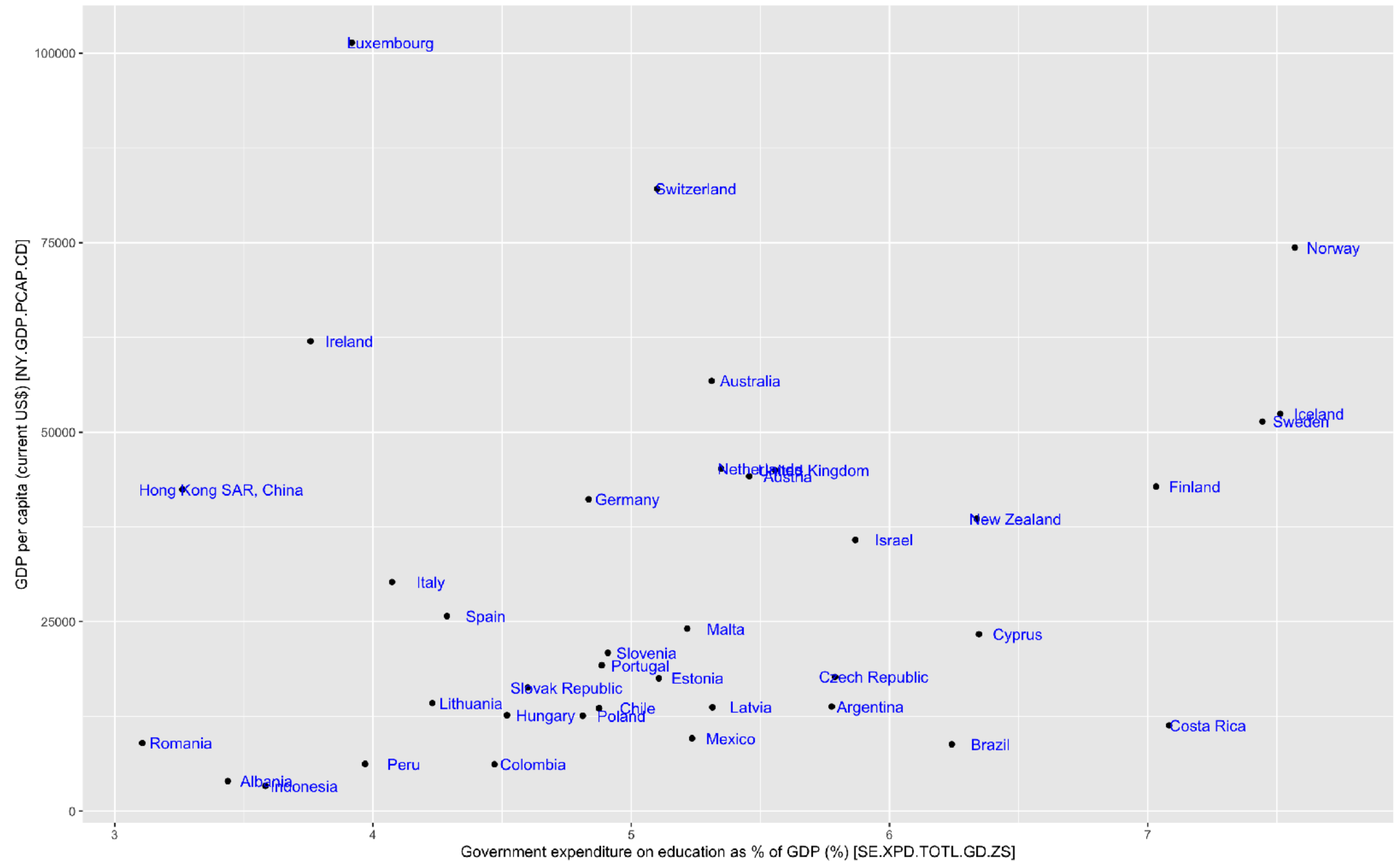
Impact of GDP and Governmental expenditure on education (as a % of GDP) on PISA Science Test Scores

Project for Machine Learning, Statistical Learning, Deep Learning and Artificial Intelligence Module

Aleksandra Łeppek

Background

- UNESCO The Education 2030 Framework for Action Expenditure on Education as % of GDP indicator was chosen as: “ **a key benchmark for a government financing of education, allocating at least 4% to 6% of its GDP.** The indicator also reflects the commitment of a government to invest in human capital development”.
- PISA test, **measures 15-year-olds’ ability to use** their reading, mathematics and **science knowledge** and skills to meet real-life challenges.
-



Supervised Learning: Expenditure on Education and PISA Science Scores



Supervised Learning: Expenditure on Education and PISA Science Scores - Results

	Linear		Polynomial degree = 2			Polynomial degree = 3			
Model	Interce pt	B1	Interce pt	B1	B2	Interce pt	B1	B2	B3
Value	425.403	10.198	417.761 4	13.1811	-0.2783	534.656 4	-56.685 9	13.1316	-0.8274
Std. error	32.446	6.222	132.133 0	50.3473	4.6597	655.366 0	386.774 0	73.7330	4.5397
t value	13.1119	1.639	3.162	0.262	-0.060	0.816	-0.147	0.178	-0.182
P-value	1.79e-1 3	0.112	0.00385	0.79546	0.95281	0.422	0.885	0.860	0.857
Fit metrics									
RSE	37.3291 3		38.0116 2			38.7109 9			
R^2	0.08753 939		0.08765 993			0.08882 394			

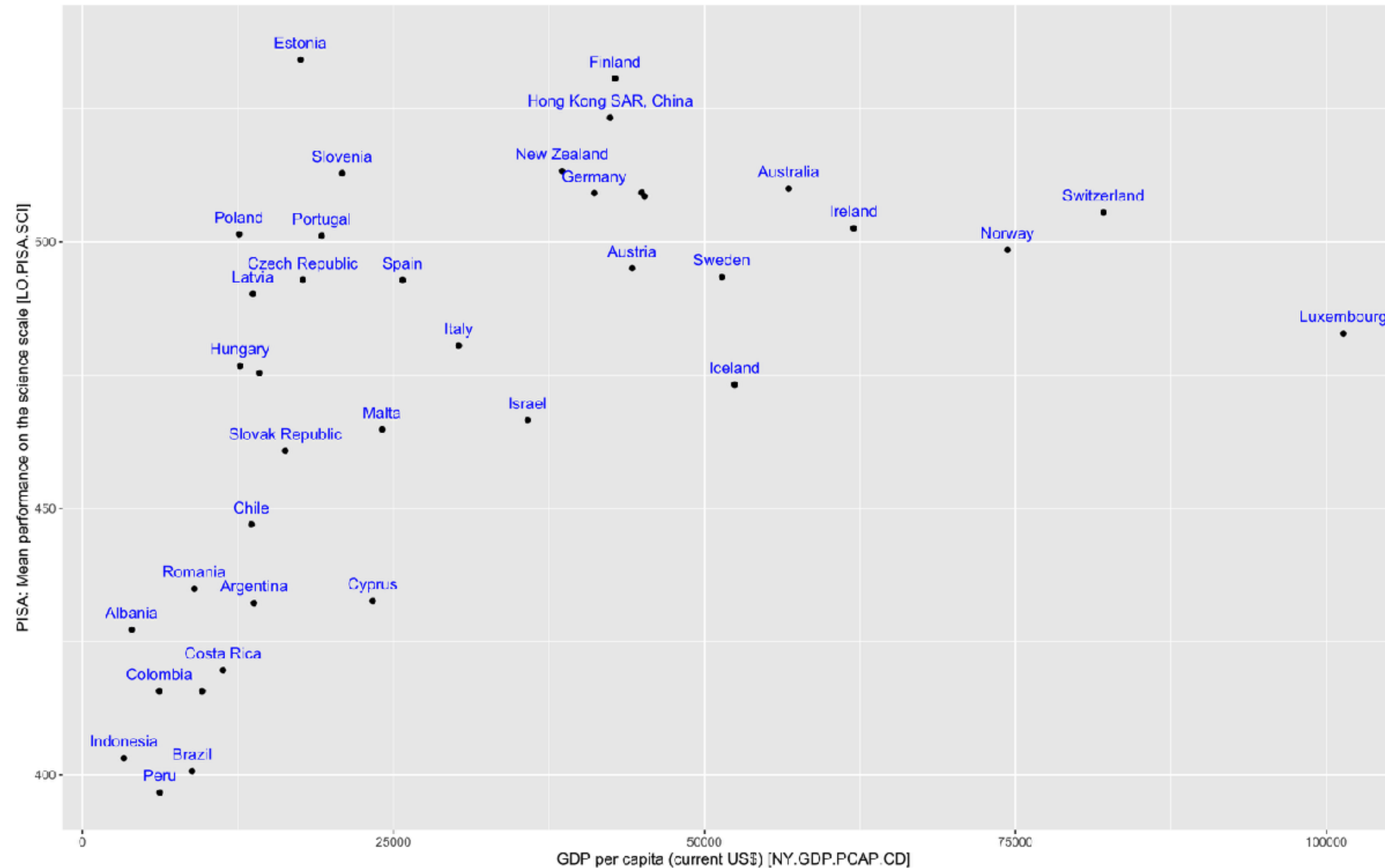
Expenditure on education as a % of GDP impact on PISA Science Test Scores failed to reject null hypothesis for regression models used in the analysis.

Supervised Learning: Expenditure on Education, GDP per capita and PISA Science Scores - Results

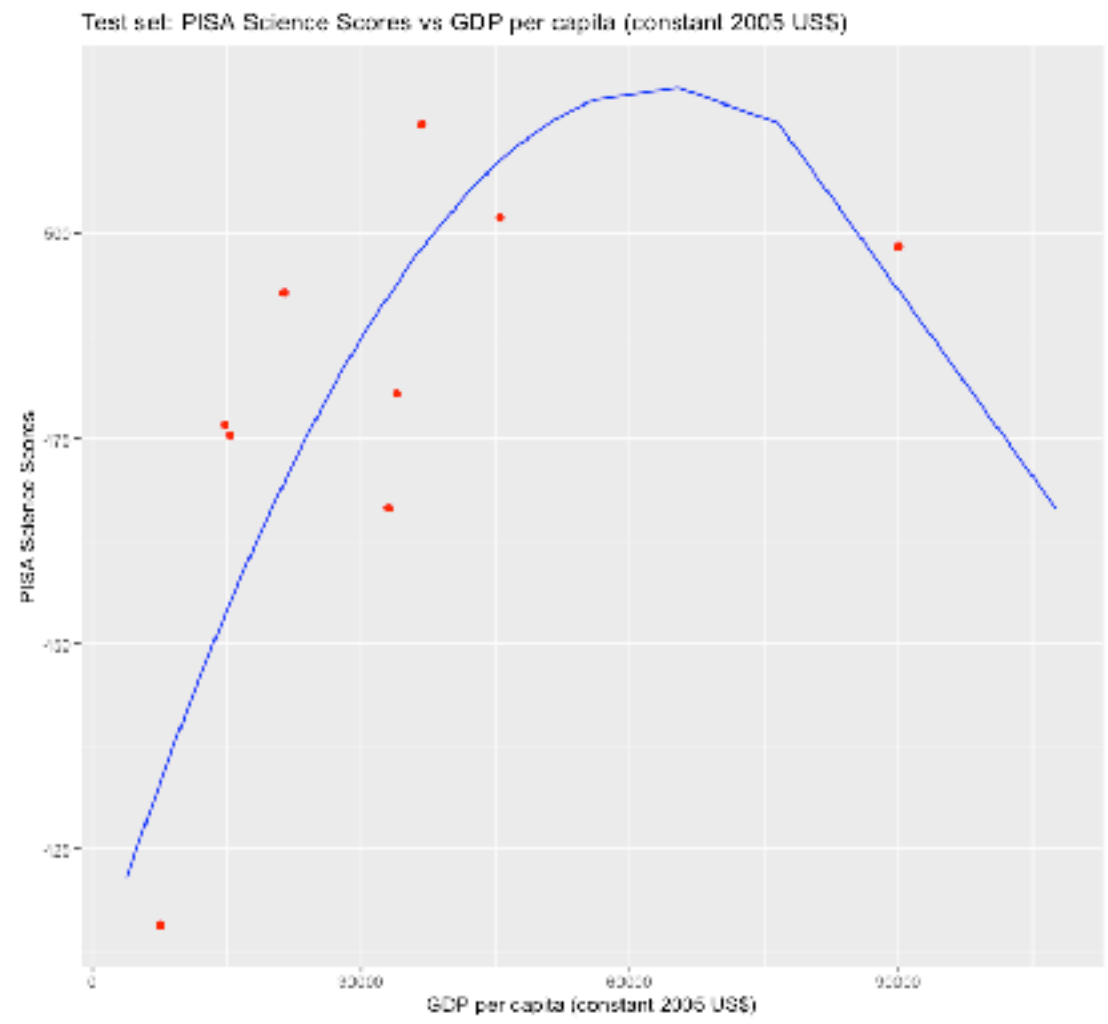
	Multiple linear regression			Interaction terms			
	Intercept	Goverme nt exp.	GDP per capita	Intercept	Goverme nt exp.	GDP per capita	Goverme nt exp. * GDP per capita
Value	475.784	6.400	19.828	477.059	8.880	18.122	4.057
Std. error	5.810	6.447	5.688	6.192	7.553	6.322	6.245
t value	81.891	0.993	3.486	77.041	1.176	2.867	-0.650
P-value	< 2e-16	0.32966	0.00169	< 2e-16	0.25035	0.00811	0.52158
Validation	Training set	Test set		Training set	Test set		
RSE	31.56837			31.91172			
R^2	0.3707424	0.02780928		0.3807956	0.0009537879		
RMSE		41.9403			44.64205		
MAE		35.17922			36.35546		

Combining 2 indicators and using multiple regression and polynomial regression with interaction terms shown to be statistically not significant and failing to reject null hypothesis.

Supervised Learning: GDP per Capita and PISA Science Scores



Supervised Learning: GDP per capita and PISA Science Scores - Results

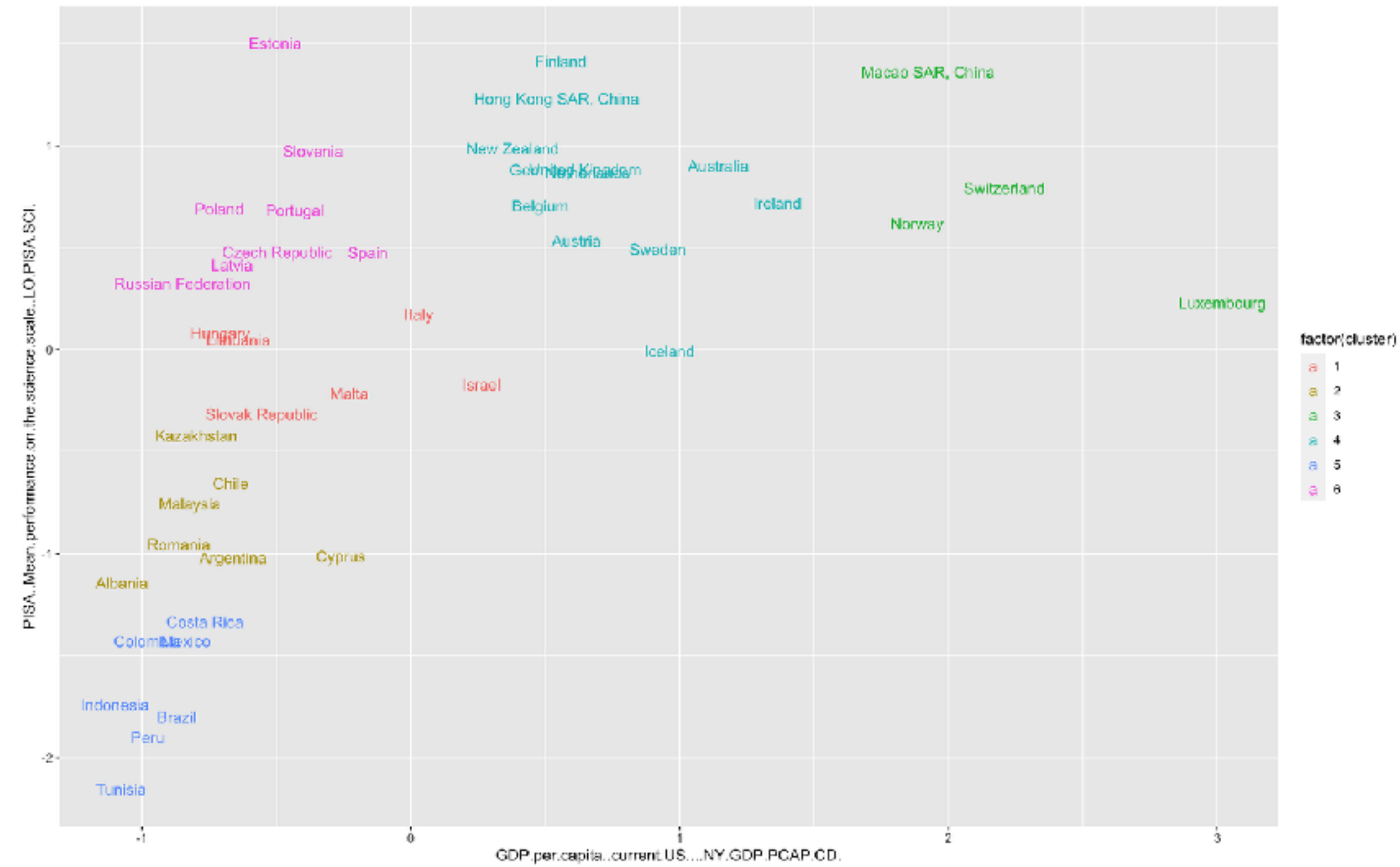


For Polynomial degree = 2 we found **significant relationship between GDP per Capita and PISA Science** test scores (both coefficients $p < 0.001$, $R^2 = 0.6281834 \pm 0.00869$ for a training set), with approx. 3.34 point increase for each additional 1000\$ of GDP per Capita (until 60k \$)

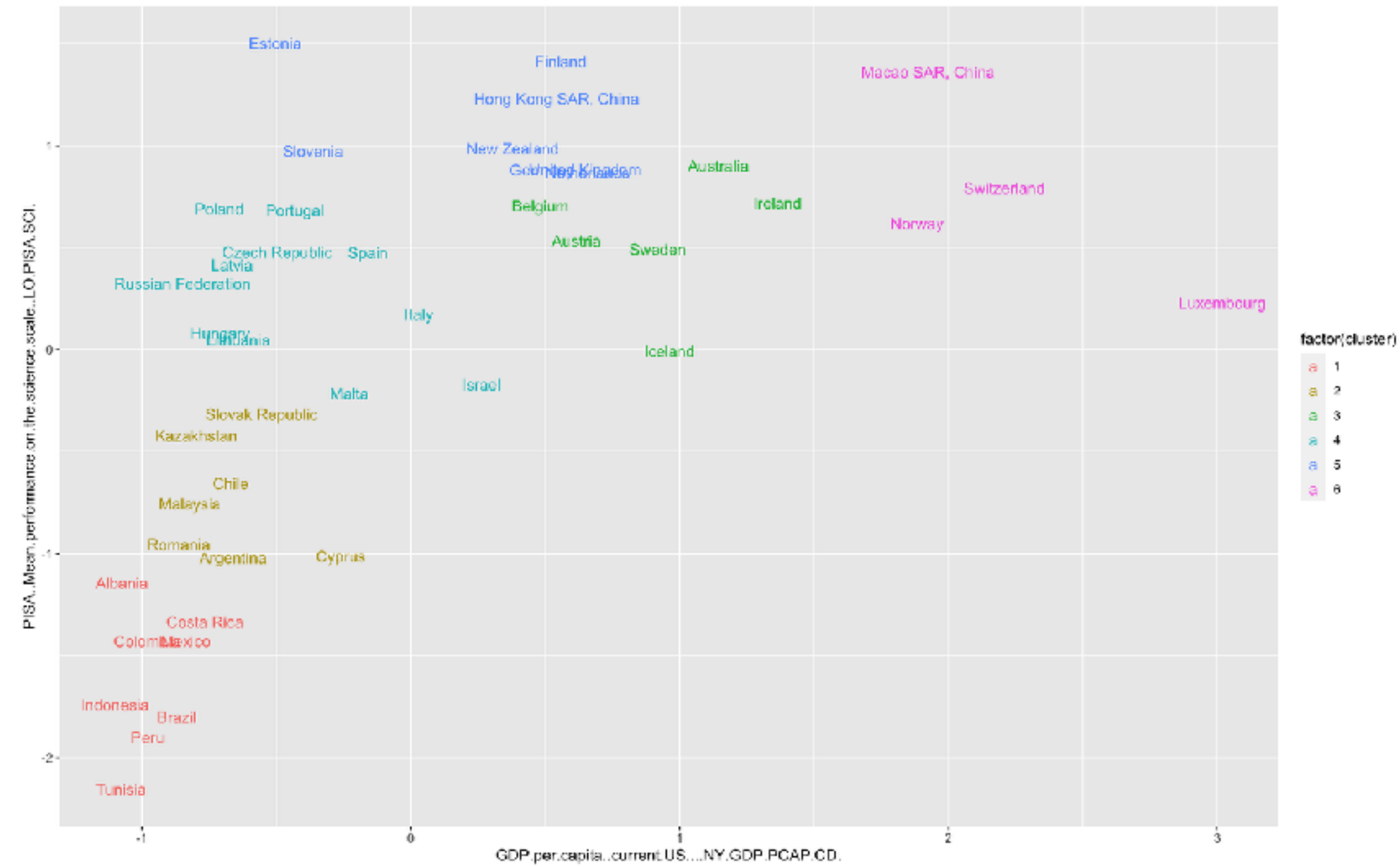
	Linear		Polynomial degree = 2		Polynomial degree = 3	
Validation	Training set	Test set	Training set	Test set	Training set	Test set
RSE	31.56018		24.26623		23.31175	
R ²	0.3477751	0.142833	0.6281834	0.2462836	0.6695671	0.3551297
RMSE		38.33772		36.1408		33.77566
MAE		31.67194		30.46476		29.6111

Unsupervised Learning: GDP per capita and PISA Science Scores - Results

K-Mean Clustering



Hierarchical Clustering



Conclusions

- GDP per capita and PISA test scores shown meaningful relationship for both supervised and unsupervised models.
- Secondary students from countries with higher GDP per capita on average where having higher results on PISA Science Test, no matter the governmental expenditure on education (as % of GDP)