New Challenges for Statistical Software - The Use of R in Official Statistics

April 7-8, 2016, Bucharest

INFLUENCE FACTORS OF THE ECONOMIC DEVELOPMENT LEVEL ACROSS EUROPEAN COUNTRIES

Diana Ioana POPA

National Institute of Statistics Romania
diana.popa@insse.ro

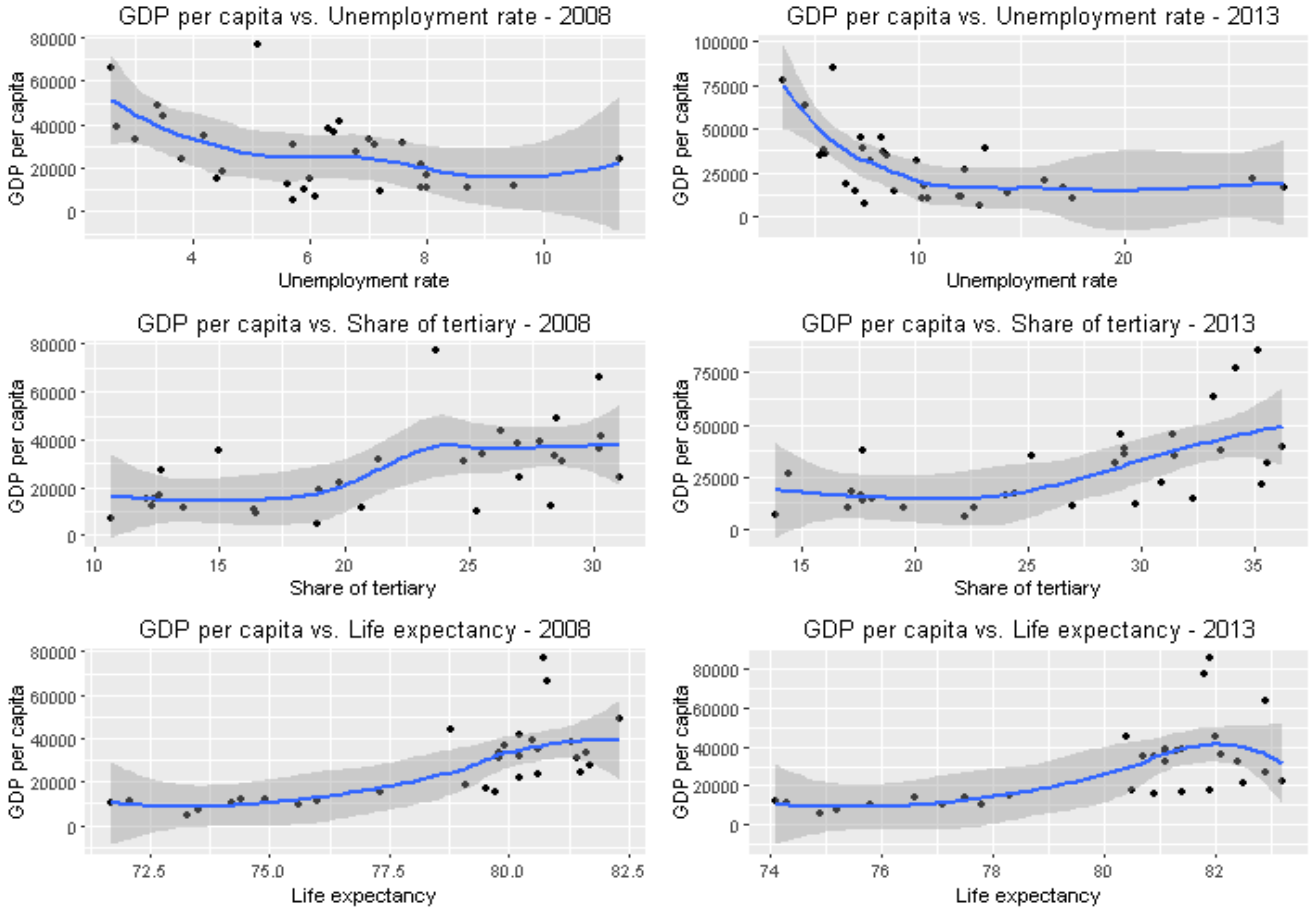
Overview

- GDP per capita - a measure of a country's economic development level
- Factors of influence: various, from social and economical to environmental and government policies
- Aim of research: to discover through statistical tools available in R potential factors
- Data from 31 European countries was used in the interest years 2008 and 2013

Methods

- Multiple Linear Regression models were applied
- Goal: to explain the relationship between GDP per capita and certain independent variables
- 11 independent variables were chosen for testing

The relationship between GDP per capita vs. the independent variables, in 2008 and 2013



Source data: Eurostat

Variables of the model

- Dependent variable: GDP per capita
- The best fitted model contains three independent variables:
  - Unemployment rate (15-64 years)
  - Share of the working age population with tertiary education
  - Life expectancy at birth

Analisys

- The multiple regression equation of the best fitted model:

$$gdp\_cap_i = \beta_0 + \beta_1 * unempl\_rate_i + \beta_2 * sh\_tert_i + \beta_3 * life\_ex_i + \varepsilon_i$$

- Two equations were computed initially, one for each year of interest

- Luxembourg – outlier; two more equations were computed after eliminating it

Coefficient estimates for Model 3 (2008) and Model 4 (2013)

Variables / Year		Coeff. / Std. error	P-value	Confidence interval	
				Lower	Upper
Intercept	2008	-178170.6 (38001.3)	7.65e-05	-256283.3340	-100057.7691
	2013	-231050.8 (50556.4)	0.000105	-334970.9131	-127130.6638
Unemployment rate	2008	-1690.8 (733.8)	0.02945	-3199.0499	-182.5346
	2013	-1294.6 (307.4)	0.000268	-1926.3737	-662.8098
Share of tertiary	2008	736.7 (232.1)	0.00384	259.6531	1213.7082
	2013	812.8 (266.3)	0.005188	265.3323	1360.2085
Life expectancy	2008	2523.5 (485.9)	2.02e-05	1524.8130	3522.2138
	2013	3141.0 (656.8)	5.97e-05	1791.0197	4490.9592

Significance level

- R^2
  - 2008: 0.7495 (p value - 5.606e-08)
  - 2013: 0.7276 (p value - 1.645e-07)

- The significance level and R^2 of both models are similar in both years studied.

Conclusions

- As level, the coefficients estimates are different in the years studied

- The relationship type between the dependent and independent variables - the same in both years of interest

References

- Maindonald, J., Braun, W.J., 2010, Data Analysis and Graphics Using R – an Example-Based Approach, Third Edition, Cambridge University Press, New York, ISBN-13 978-0-511-71286-9
- 2. Caragea, Nicoleta & Alexandru, Ciprian Antoniade & Dobre, Ana Maria, 2012. "**Bringing New Opportunities to Develop Statistical Software and Data Analysis Tools in Romania**," MPRA Paper 48772, University Library of Munich, Germany.
- 3. Dobre Ana Maria & Caragea Nicoleta & Alexandru Ciprian Antoniade, 2013. "R versus Other Statistical Software," Ovidius University Annals, Economic Sciences Series, Ovidius University of Constantza, Faculty of Economic Sciences, vol. 0(1), pages 484-488, May.
- 4. Nicoleta Caragea & Antoniade-Ciprian Alexandru & Ana Maria Dobre, 2014. "R – a Global Sensation in Data Science," Romanian Statistical Review, Romanian Statistical Review, vol. 62(2), pages 7-16, June.
- 5. Caragea, N., 2015, Statistica - Concepte şi metode de analiză a datelor, Editura Mustang, Bucureşti, ISBN 978-606-652-063-8

THANK YOU ！