

R Programming Project Exam

2024-10-26

Groupe F:

- ADOUSSINGANDE Céline Mauricette
- AKONAKPO Romulus
- BOULANKI Loukmane

I- Description of the relationship of interest

1. Describe of the two variables of interest.

- Variable X

In our study, our variable X is the **Number of tertiary graduates**. This variable is calculated as the sum of graduates in the following fields of education (tertiary level (05-07)): F05 Natural Sciences, Mathematics and Statistics; F06 Information and Communications Technologies; and F07 Engineering, Manufacturing and Construction. So, the Variable X represents the total number of graduates at the tertiary education level.

- Variable Y

Our variable Y is the **GDP per capita**. This variable measures the gross domestic product (GDP) per capita, which represents as an indicator of the economic productivity and average income per person within a country. The GDP per capita is derived by dividing the total GDP of a country by its population.

2. Why is it interesting to study the relationship between these two variables ?

We study the relationship between **Number of tertiary graduates** and **GDP per capita** in order to verify some theory of endogenous growth. Indeed, Endogenous Growth Theory attribute growth to external factors like technological advances. It suggests that growth is generated internally and therefore the investments in human capital, innovation, and knowledge, are the factors who foster sustained economic development. As argued by Paul Romer, in his 1986 article “Increasing Returns and Long-Run Growth”, points out that knowledge and innovation are internal drivers of economic growth. However, investments in research and development (R&D) and education stimulate growth because they increase the stock of knowledge that benefits the economy as a whole.

So, in our study to approach human capital, we use our variable X which represents the total number of tertiary level graduates.

II- Description of data

3. How is measured “Number of tertiary graduates”?

The Number of Tertiary Graduates (X) is measured as the annual count (number) of individuals completing tertiary education within specified fields. This measure reflects the yearly output of graduates who have completed tertiary education.

4. How is measured gdp per capita ?

The variable GDP per capita (Y) is measured as the gross domestic product (GDP) per person in each country and calculated annually. It is expressed in current prices in euros per capita per year. It reflects the value of goods and services produced in a country per person in the economy, without adjusting for inflation or currency fluctuations over time.

5. Which column(s) do you use to merge both datasets Number of tertiary graduates and gdp per capita?

To merge the two datasets, Number of Tertiary Graduates and GDP per Capita, the columns “country” and “year” are used as the keys for joining. These columns ensure that data aligns accurately for each country in a specific year, allowing for a consistent comparison across both datasets.

The database obtained from the merger of our two databases will allow us to analyze the relationship between the number of tertiary graduates and GDP per capita by country and by year, which will facilitate the understanding of the effects of human capital on economic performance across different countries and time periods.

6. What is the unit of observation in your final dataset?

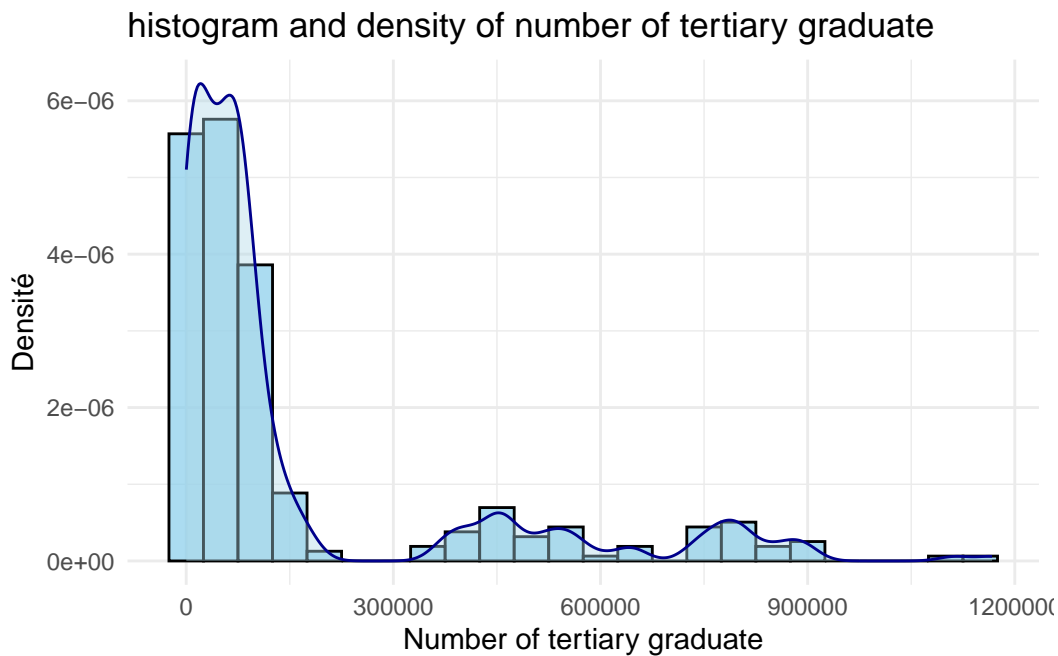
The unit of observation in our final dataset is “**country, year**”. This unit of observation allows for the analysis of trends and relationships between countries and over time, observing how changes in tertiary education graduates can be related to changes in GDP per capita in different countries and years.

7. How many observations are there in your final dataset?

In our final dataset, we have 315 observations. This implies that we have data for 390 unique country-year combinations. Each observation represents a single country and year pair, providing data on the Number of Tertiary Graduates and GDP per Capita for that specific combination.

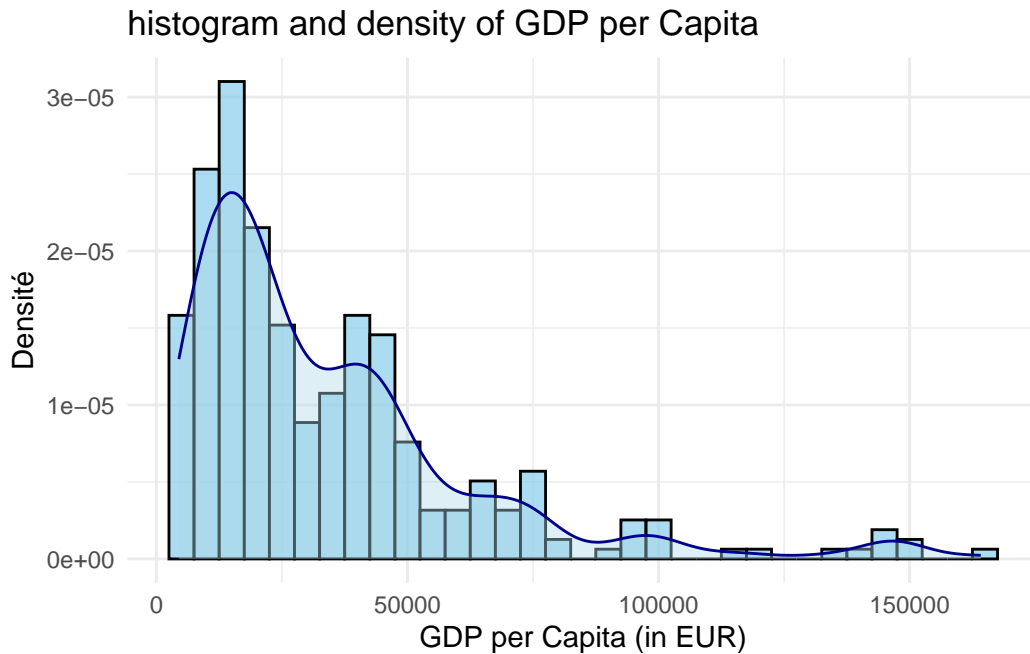
III- Descriptive statistics

8. What is the distribution of the variable “number of tertiary graduate”? Include an histogram and/or a density plot of “number of tertiary graduate”.



The distribution of the graduates by educational level is highly skewed and concentrated towards the lower values. This implies that most observations in our sample are associated with low education levels, while there are fewer observations at the higher education levels.

9. What is the distribution of variable “gdp per capita” ? Include an histogram and/or a density plot of “gdp per capita”.



The GDP per capita distribution is also skewed, though less extreme than graduates by educational level. It has a high density for low GDP values, which gradually decreases as the GDP increases.

10. Create a table of summary statistics with mean, standard deviation, minimum and maximum for both variables “number of tertiary graduate” and “gdp per capita” .

Table 1: Summary statistics for GDP per Capita and Number of Tertiary Graduate

Variable	Mean	Standard Deviation	Minimum	Maximum
GDP per Capita	34051.52	29041.08	4470	164150
Number of Tertiary Graduate	161910.20	243297.45	163	1167119

11. Create a table of summary statistics where each row corresponds to one country and reports the name of the country, the mean of “number of tertiary graduate” (over the entire time period) and standard deviation and Gdp per capita .

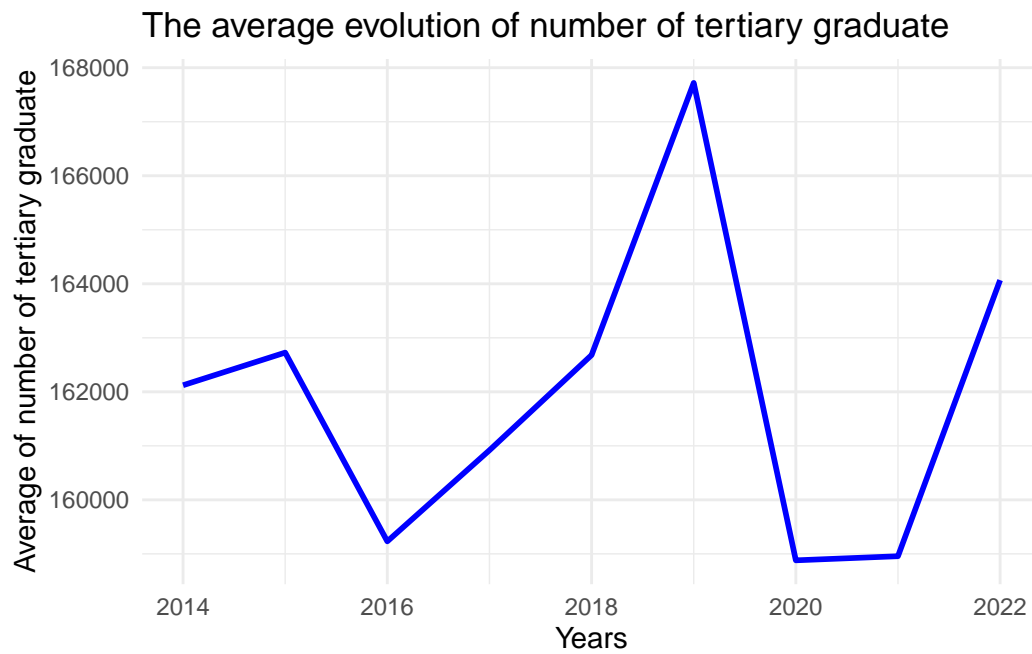
Table 2: Summary statistics by country for X (tertiary graduate education level) et Y (gdp per capita)

Country	Mean of X (Number of tertiary graduate)	SD of X	Mean of Y (GDP per Capita)	SD of Y
Albania	31800.000	1258.65007	5960.000	763.6753
Austria	85281.222	2704.01071	42908.889	3309.0725
Belgium	118101.000	5434.70857	40425.556	3769.0785
Bulgaria	54631.556	6582.55587	8754.444	2303.3297
Croatia	34437.111	1078.20272	12973.333	2158.8423
Cyprus	10059.000	2194.09474	25012.222	3715.9648
Czechia	81527.778	14064.88700	20022.222	3684.3309
Denmark	80522.111	4935.43981	53058.889	5628.6065
Estonia	9574.222	612.91879	20035.556	3860.4472
Finland	58018.889	2799.80073	41871.111	3275.0780
France	808565.333	52353.77744	35005.556	2036.0692
Germany	588036.667	47344.96341	41168.889	3235.0519
Greece	72458.000	6466.52683	16867.778	1179.5632
Hungary	78765.111	37443.75009	13854.444	2238.7112
Iceland	4628.778	418.96234	58574.444	9592.8490
Ireland	82003.333	11198.87648	69923.333	17070.1530
Italy	411600.222	39436.61281	29462.222	2084.0992
Latvia	15290.444	1254.23583	14561.111	2522.9370
Liechtenstein	244.875	41.34329	147141.250	8546.6592
Lithuania	27195.667	3757.55745	16592.222	3738.9396
Luxembourg	1835.111	191.49768	101568.889	8533.1200
Malta	4593.667	622.18888	27320.000	4357.1981
Montenegro	4979.167	205.46768	7241.667	679.6298
Netherlands	160486.750	12564.38077	46583.750	5182.5695
North Macedonia	10306.111	994.43809	5642.222	863.8978
Norway	53044.500	4299.02371	68718.750	5686.6622
Poland	468590.333	56367.39239	13426.667	2256.4519
Portugal	80705.778	6719.35067	19424.444	2046.3022
Romania	129827.333	9824.20065	10526.667	2361.4402
Serbia	46473.444	4400.00415	6646.667	1470.7906
Slovakia	48269.222	10729.05795	16633.333	1968.6925
Slovenia	18436.444	4833.11445	21736.667	2910.7774
Spain	471988.778	39291.50696	25158.889	1931.7249
Sweden	80123.000	6164.49446	47531.111	2653.8295
Switzerland	89582.750	3668.17731	73880.000	3513.7160
Türkiye	895872.000	148141.38617	8961.111	969.2322

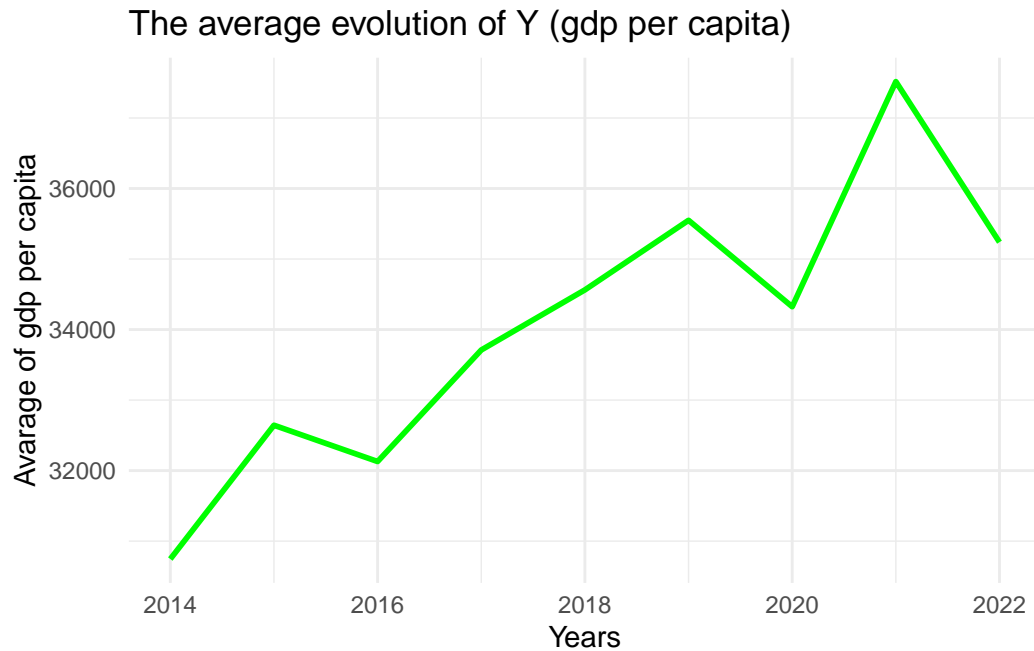
Country	Mean of X (Number of tertiary graduate)	SD of X	Mean of Y (GDP per Capita)	SD of Y
United Kingdom	790636.500	45113.71874	37246.667	1835.7306

12. Plot the average evolution of “Number of tertiary graduate” and “Gdp per capita” across the years into two separated graphs. Include them into your document.

- The average evolution of “number of tertiary graduate”

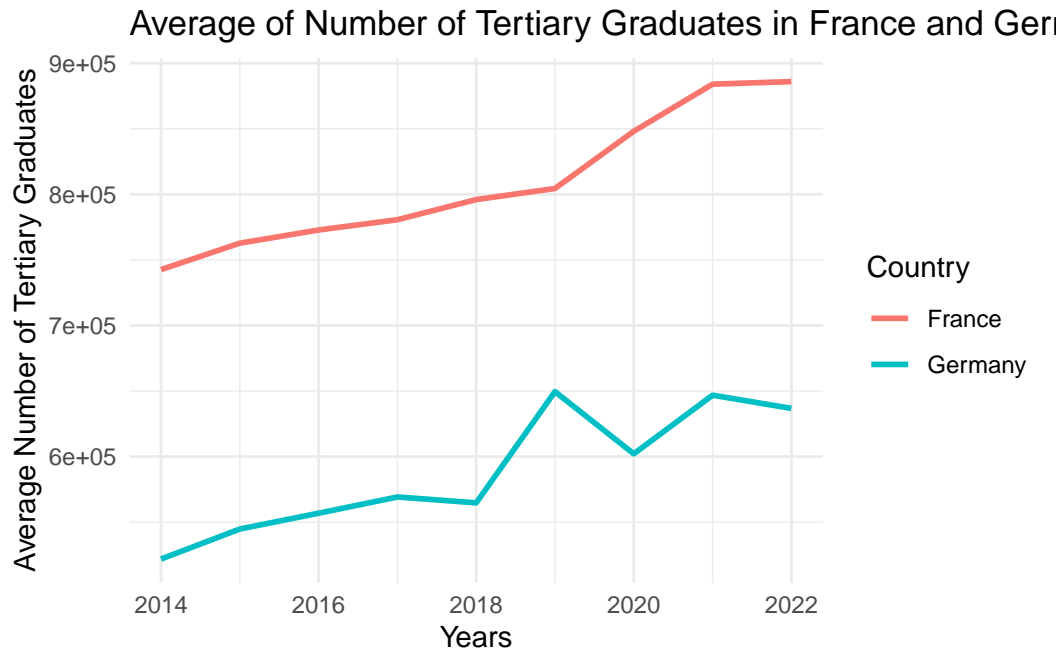


- The average evolution of Y (gdp per capita)

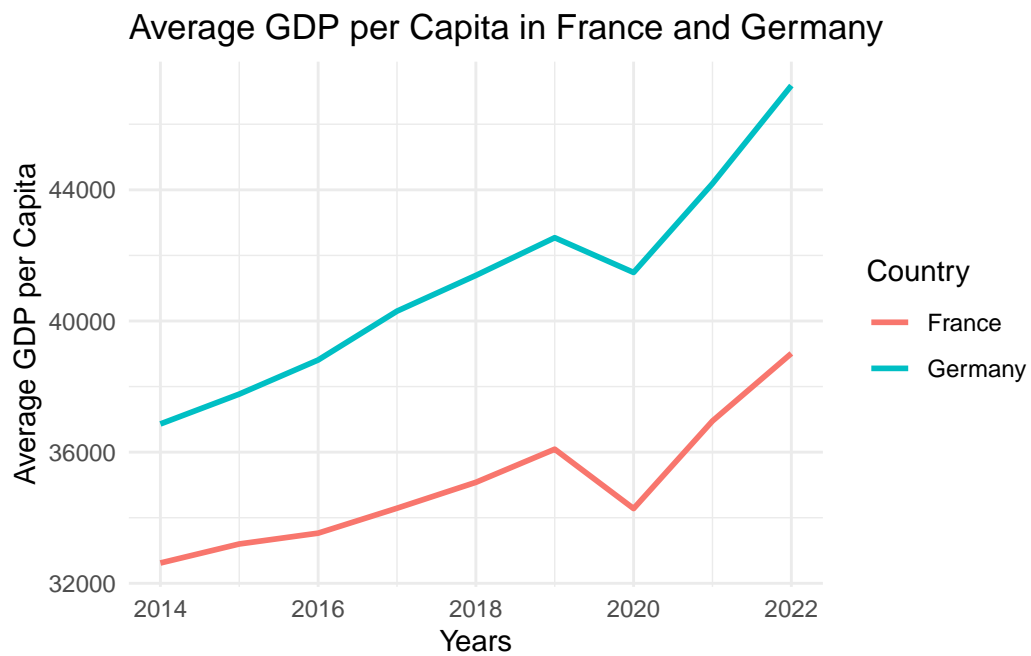


13. Select two countries and plot the average evolution of variables “number of tertiary graduate” and “GDP per capita” across the years for both in two separated graphs.

- Average of number of tertiary graduate in two country (France and Germany)



- The average evolution of Y (gdp per capita)



- Explain why you selected these two countries and why they provide important insights regarding the relationship between “number of tertiary graduate” and “GDP per capita”.

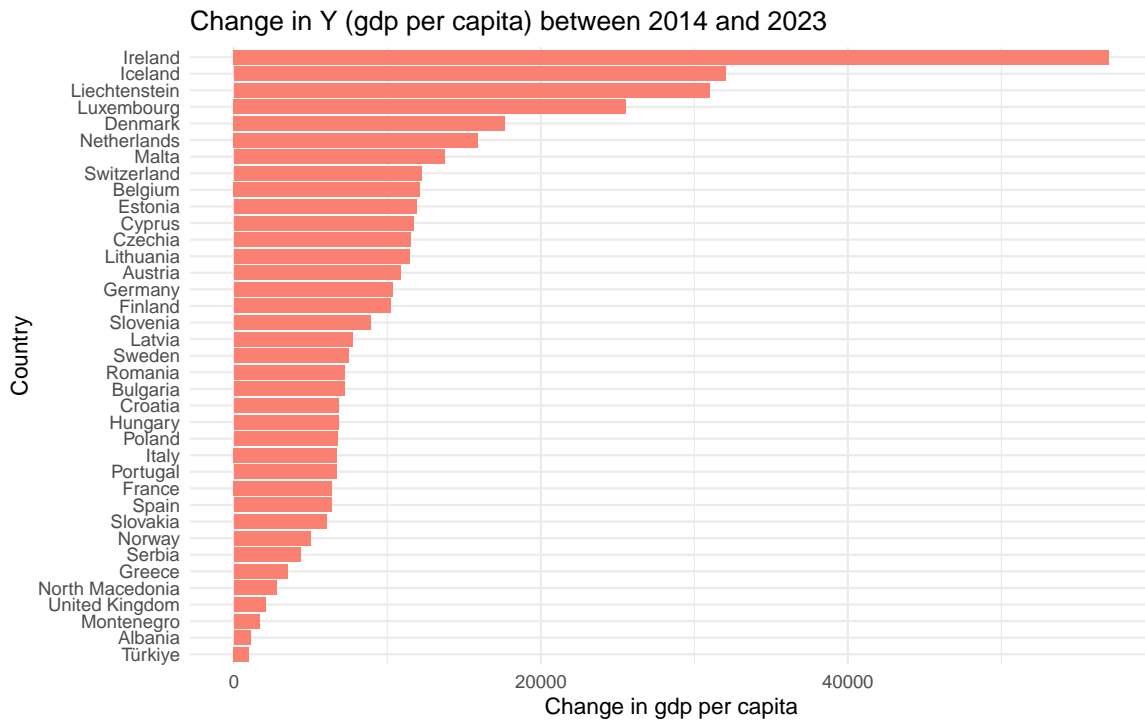
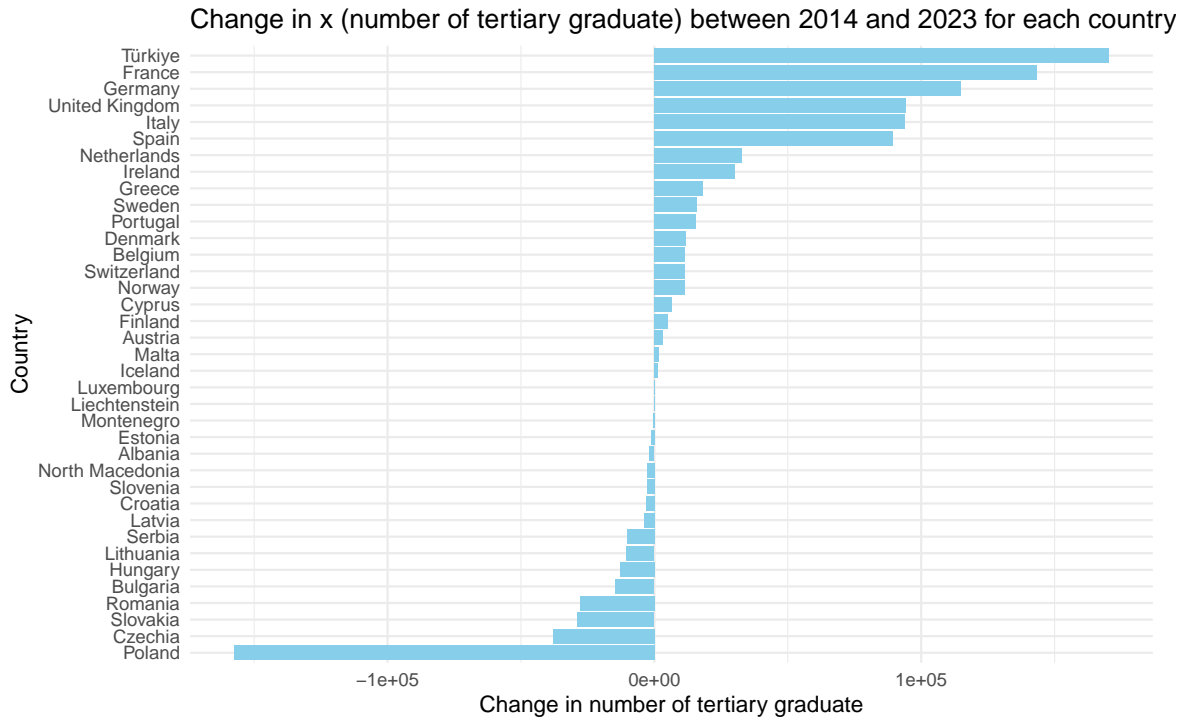
- **France**

France has a diversified economy with strengths in technology, industry, and services. The analysis of France can provide insight into how the education system contributes to economic performance in different sectors. Also, the French state invests massively in education, which makes it interesting to observe how these investments translate into a growth in GDP in relation to the number of graduates.

- **Germany**

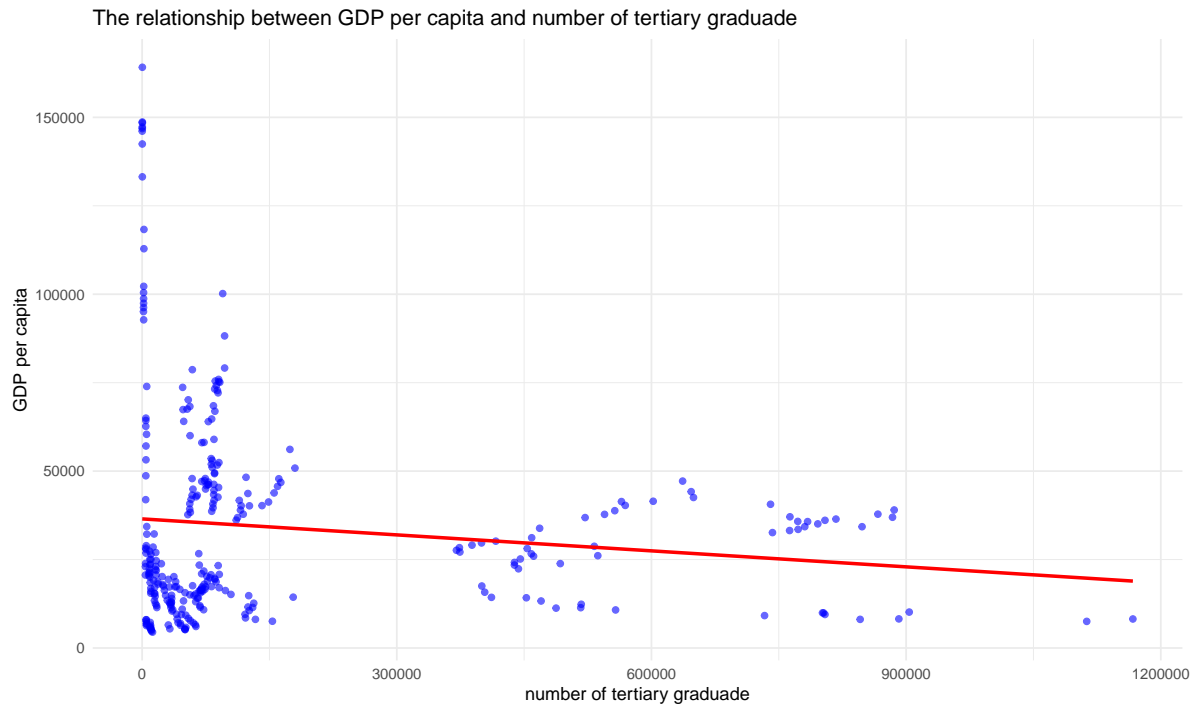
Germany is the largest economy in Europe, known for its industrial base and high level of education. It has a strong higher education system that contributes significantly to its economic growth. Thus, the relationship between the number of tertiary graduates and GDP is often visible in Germany, as higher education is strongly linked to innovation and productivity in various sectors

14. Compute change in “number of tertiary graduate” between first and last year for each country. Rank countries into a barplot in decreasing order. Do the same with variable “GDP per capita” .

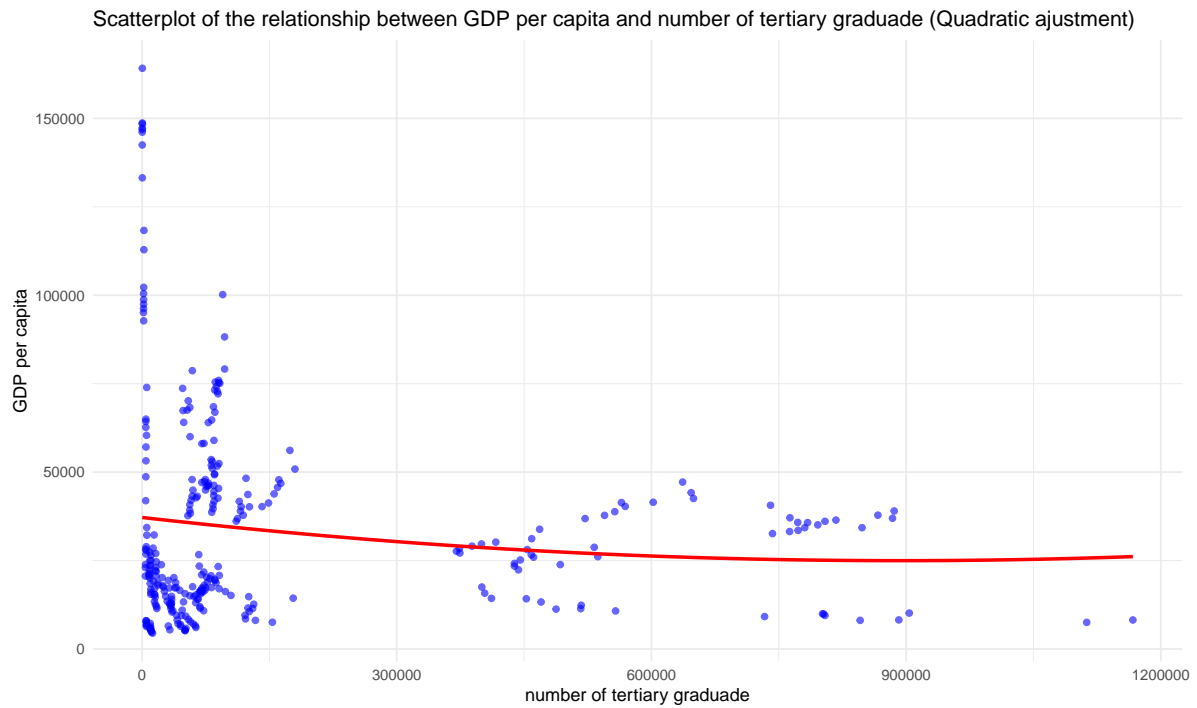


IV. Relationship between “number of tertiary graduate” and “GDP per capita”

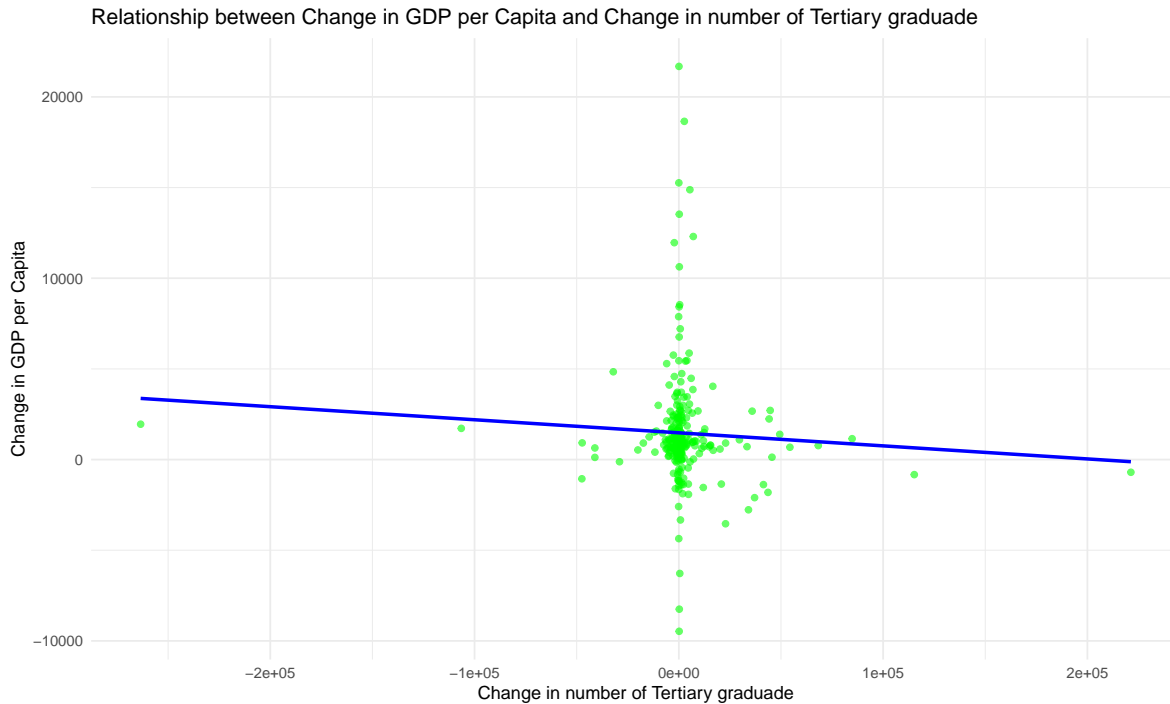
15. Plot a scatterplot for the relationship between “GDP per capita” and “number of tertiary graduate”.



16. Plot another scatterplot of the relationship between “GDP per capita” and “number of tertiary graduate”.



17. Compute within-country changes over the year for variables number of tertiary graduate and GDP per capita . Plot a scatterplot between the within-country change in GDP per capita and the within-country change in number of tertiary graduate.



18. Regress GDP per capita on number of tertiary graduate. Export the results into a table

Table 3: Table of the Results of the Regression of GDP per capita on number of tertiarygraduate

Terms	Estimator	Standard deviation	t-statistic	P-value
(Intercept)	36489.9383604	1950.7416230	18.705675	0.0000000
nb_tertiary_grad	-0.0150603	0.0066823	-2.253764	0.0249004

19. Regress GDP per capita on number of tertiary graduate controlling for country fixed effects.

Table 4: Table of Regression Results of Y on X (with and without fixed effects)

Model	Terms	Estimator	Standard deviation	Statistical t	P-value
without Fixed Effects	(Intercept)	3.648994e+04	1950.7416230	18.7056748	0.0000000
without Fixed Effects	nb_tertiary_grad	- 1.506030e- 02	0.0066823	- 2.2537636	0.0249004
With Fixed Effects	(Intercept)	5.679611e+03	3422.7657871	1.6593631	0.0981703
With Fixed Effects	nb_tertiary_grad	8.817300e- 03	0.0088807	0.9928562	0.3216434
With Fixed Effects	factor(country)Austria	3.647733e+04	3800.8990020	9.5970272	0.0000000
With Fixed Effects	factor(country)Belgium	3.370462e+04	3848.2001455	8.7585406	0.0000000
With Fixed Effects	factor(country)Bulgaria	2.593133e+03	3776.5548827	0.6866398	0.4928821
With Fixed Effects	factor(country)Croatia	6.990081e+03	3771.1806582	1.8535525	0.0648625
With Fixed Effects	factor(country)Cyprus	1.924392e+04	3776.0472926	5.0963128	0.0000006
With Fixed Effects	factor(country)Czechia	1.362376e+04	3796.8778047	3.5881480	0.0003935
With Fixed Effects	factor(country)Denmark	4.666929e+04	3795.8493958	12.2948222	0.0000000
With Fixed Effects	factor(country)Estonia	1.427153e+04	3776.2698703	3.7792654	0.0001925
With Fixed Effects	factor(country)Finland	3.567993e+04	3778.2893644	9.4434092	0.0000000
With Fixed Effects	factor(country)France	2.219662e+04	7861.7229569	2.8233780	0.0050955
With Fixed Effects	factor(country)Germany	3.030441e+04	6214.7075341	4.8762402	0.0000018
With Fixed Effects	factor(country)Greece	1.054929e+04	3788.3542210	2.7846619	0.0057264
With Fixed Effects	factor(country)Hungary	7.480341e+03	3794.1024435	1.9715706	0.0496501
With Fixed Effects	factor(country)Iceland	5.285402e+04	3778.8199903	13.9869113	0.0000000

Model	Terms	Estimator	Standard deviation	Statistical t	P-value
With Fixed Effects	factor(country)Ireland	6.352068e+04	3797.3713324	16.7275391	0.0000000
With Fixed Effects	factor(country)Italy	2.015343e+04	5059.4123396	3.9833531	0.0000868
With Fixed Effects	factor(country)Latvia	8.746680e+03	3773.9570028	2.3176417	0.0211950
With Fixed Effects	factor(country)Liechtenstein	4.144595e+05	3824.0025069	36.9925175	0.0000000
With Fixed Effects	factor(country)Lithuania	1.067282e+04	3771.3296139	2.8299886	0.0049943
With Fixed Effects	factor(country)Luxembourg	9.587310e+04	3780.4853104	25.3599972	0.0000000
With Fixed Effects	factor(country)Malta	2.159989e+04	3778.8399140	5.7160096	0.0000000
With Fixed Effects	factor(country)Montenegro	1.518153e+03	3945.9890865	0.3847332	0.7007293
With Fixed Effects	factor(country)Netherlands	3.948909e+04	3981.2714169	9.9187123	0.0000000
With Fixed Effects	factor(country)North Macedonia	- 1.282606e+02	3775.9357194	- 0.0339679	0.9729271
With Fixed Effects	factor(country)Norway	6.257143e+04	3818.3845144	16.3868860	0.0000000
With Fixed Effects	factor(country)Poland	3.615374e+03	5409.9838149	0.6682781	0.5045106
With Fixed Effects	factor(country)Portugal	1.303323e+04	3796.0356684	3.4333791	0.0006869
With Fixed Effects	factor(country)Romania	3.702334e+03	3870.2861192	0.9566048	0.3395979
With Fixed Effects	factor(country)Serbia	5.572871e+02	3773.3587032	0.1476899	0.8826945
With Fixed Effects	factor(country)Slovakia	1.052812e+04	3773.9431043	2.7896870	0.0056407
With Fixed Effects	factor(country)Slovenia	1.589450e+04	3772.9748861	4.2127226	0.0000341
With Fixed Effects	factor(country)Spain	1.531763e+04	5431.6642738	2.8200622	0.0051470
With Fixed Effects	factor(country)Sweden	4.114503e+04	3795.4470070	10.8406295	0.0000000
With Fixed Effects	factor(country)Switzerland	6.1741051e+04	3848.0890879	17.5179194	0.0000000

Model	Terms	Estimator	Standard deviation	Statistical t	P-value
With Fixed Effects	factor(country)Türkiye	- 4.617634e+03	8550.1366440	- 0.5400656	0.5895842
With Fixed Effects	factor(country)United Kingdom	2.459581e+04	7805.6520137	3.1510257	0.0018047

20. Regress GDP per capita on number of tertiary graduade controlling for year and country fixed effects.

Table 5: Table of Regression Results of Y on X with Fixed Effects (Country and Year)

Model	Terms	Estimator	Standard deviation	Statistical t	P-value
with Fixed Effects (Country and Year))	(Intercept)	- 3.934125e+03	2552.0323283	- 1.541565	0.1243504
with Fixed Effects (Country and Year))	nb_tertiary_grad	- 1.262460e-02	0.0065973	- 1.913585	0.0567306
with Fixed Effects (Country and Year))	factor(country)Austria	4.303688e+04	161.2296951	15.586128	0.0000000
with Fixed Effects (Country and Year))	factor(country)Belgium	1.096788e+04	102.3278246	14.619232	0.0000000
with Fixed Effects (Country and Year))	factor(country)Bulgaria	2.495496e+03	337.7958343	3.103042	0.0021188
with Fixed Effects (Country and Year))	factor(country)Croatia	1.245944e+04	130.4414313	4.563159	0.0000077
with Fixed Effects (Country and Year))	factor(country)Cyprus	2.419057e+04	130.2119209	8.860325	0.0000000
with Fixed Effects (Country and Year))	factor(country)Czechia	2.010283e+04	157.5740204	7.290041	0.0000000
with Fixed Effects (Country and Year))	factor(country)Denmark	5.312680e+04	156.6315101	19.272361	0.0000000
with Fixed Effects (Country and Year))	factor(country)Estonia	1.920778e+04	130.3034288	7.035034	0.0000000
with Fixed Effects (Country and Year))	factor(country)Finland	4.165493e+04	139.6620893	15.204403	0.0000000
with Fixed Effects (Country and Year))	factor(country)France	4.426468e+04	165.1326516	7.547090	0.0000000

Model	Terms	Estimator	Standard deviation	Statistical t	P-value
with Fixed Effects (Country and Year))	factor(country)Germany	4.764394e+00	27.1277435	10.296655	0.0000000
with Fixed Effects (Country and Year))	factor(country)Greece	1.683388e+07	49.6410794	6.122210	0.0000000
with Fixed Effects (Country and Year))	factor(country)Hungary	3.90017e+07	55.0224197	5.045393	0.0000008
with Fixed Effects (Country and Year))	factor(country)Iceland	5.768423e+07	31.4507391	21.118533	0.0000000
with Fixed Effects (Country and Year))	factor(country)Ireland	7.000994e+07	58.0251568	25.384084	0.0000000
with Fixed Effects (Country and Year))	factor(country)Italy	3.370984e+07	50.5166913	8.988053	0.0000000
with Fixed Effects (Country and Year))	factor(country)Latvia	1.380550e+07	29.4626642	5.057955	0.0000008
with Fixed Effects (Country and Year))	factor(country)Liechtenstein	1.470451e+07	70.4885251	53.075502	0.0000000
with Fixed Effects (Country and Year))	factor(country)Lithuania	5.98691e+07	29.3839884	5.857332	0.0000000
with Fixed Effects (Country and Year))	factor(country)Luxembourg	1.006434e+07	32.2708497	36.835078	0.0000000
with Fixed Effects (Country and Year))	factor(country)Malta	2.642935e+07	31.4602762	9.675903	0.0000000
with Fixed Effects (Country and Year))	factor(country)Montenegro	6.470297e+08	58.2887647	2.161537	0.0315328
with Fixed Effects (Country and Year))	factor(country)Netherlands	4.736622e+09	4.6462513	16.307054	0.0000000
with Fixed Effects (Country and Year))	factor(country)North Macedonia	4.823685e+07	30.1667160	1.766810	0.0783899
with Fixed Effects (Country and Year))	factor(country)Norway	6.928914e+07	75.2408373	24.966893	0.0000000
with Fixed Effects (Country and Year))	factor(country)Poland	1.839376e+09	17.7393798	4.578137	0.0000072
with Fixed Effects (Country and Year))	factor(country)Portugal	1.949467e+07	56.8024747	7.071479	0.0000000
with Fixed Effects (Country and Year))	factor(country)Romania	1.121703e+08	20.8993296	3.976402	0.0000899
with Fixed Effects (Country and Year))	factor(country)Serbia	6.284726e+07	34.0457101	2.298691	0.0222866
with Fixed Effects (Country and Year))	factor(country)Slovakia	1.629406e+07	34.7807183	5.958088	0.0000000

Model	Terms	Estimator	Standard deviation	Statistical t	P-value
with Fixed Effects (Country and Year))	factor(country)Sloveni	2.102077e+07	29.2221516	7.702111	0.0000000
with Fixed Effects (Country and Year))	factor(country)Spain	3.016889e+00	34.2223413	7.478241	0.0000000
with Fixed Effects (Country and Year))	factor(country)Sweden	4.759398e+07	56.2618009	17.267584	0.0000000
with Fixed Effects (Country and Year))	factor(country)Switzer	7.491167e+08	3.9749547	26.716240	0.0000000
with Fixed Effects (Country and Year))	factor(country)Türkiy	d.932244e+03	80.5994899	3.028312	0.0026972
with Fixed Effects (Country and Year))	factor(country)United Kingdom	4.805882e+08	56.3819869	8.206231	0.0000000
with Fixed Effects (Country and Year))	factor(year)2015	1.905059e+02	6.6336756	2.304599	0.0219480
with Fixed Effects (Country and Year))	factor(year)2016	2.523163e+03	8.4242920	3.045738	0.0025507
with Fixed Effects (Country and Year))	factor(year)2017	3.797375e+02	2.1424653	4.618877	0.0000060
with Fixed Effects (Country and Year))	factor(year)2018	4.672123e+02	2.5466754	5.680070	0.0000000
with Fixed Effects (Country and Year))	factor(year)2019	5.722678e+02	4.6061823	6.939892	0.0000000
with Fixed Effects (Country and Year))	factor(year)2020	4.733377e+03	8.4398334	5.645458	0.0000000
with Fixed Effects (Country and Year))	factor(year)2021	8.913406e+03	9.5386904	10.617028	0.0000000
with Fixed Effects (Country and Year))	factor(year)2022	1.167776e+04	8.1996768	13.607281	0.0000000