



Norwegian University of
Science and Technology

Erasmus Mobility Analysis

Studying the flow of European mobility participants

Marta Almagro Fuelle (106011)
David Borregón Sacristán (105423)
Aldara Iglesias Sánchez (105734)
Celia Igual Ortiz (105990)
Lucas Redondo Puga (105234)
Lucas Sebastián Peris (106030)

TDT4259 - APPLIED DATA SCIENCE
Norwegian University of Science and Technology
Trondheim, Norway
November 26, 2023

Contents

1	Introduction and problem definition	3
1.1	Context	3
1.2	Motivation	4
1.3	Previous assumptions	4
2	Background	5
2.1	Project objectives	5
2.2	Data strategy	5
3	Method	6
3.1	Our data sources	6
3.2	Tools	6
3.3	Data pre-processing	6
4	Visualizations	9
5	Analysis	14
5.1	Malta	14
5.2	Ukraine	14
5.3	France	14
5.4	Cyprus	15
5.5	Ireland	15
6	Interpretation and recommendations	16
6.1	Recommended measures & implementation plan	16
6.1.1	Factors when choosing a destination	16
6.1.2	How can we promote mobility?	17
6.2	Limitations and future analysis	20
7	Conclusion	21
8	Appendices	22
8.1	Appendix 1: Dot distribution map, colored by receiving/sending ratio	22
8.2	Appendix 2: Arc map, colored my number of mobilities	22
8.3	Appendix 3: Sent versus received individuals	22
8.4	Appendix 4: Country ratios, by country and year	22
8.5	Appendix 5: Computed difference between received and sent participants, by country and year	22
8.6	Appendix 6: Number of sent vs received participants, by population ratio	22
8.7	Appendix 7: Top field of study, represented by labels on a map	22
8.8	Appendix 8: HTML format Jupyter Notebook with the pre-processing and visualization code for dot and arc maps.	22
8.9	Appendix 9: Erasmus web interface prototype	22

List of Figures

1	Dot distribution map, colored by receiving/sending ratio.	9
2	Arc map, colored my number of mobilities.	10
3	Sent versus received individuals.	10
4	Country ratios, by country and year.	11
5	Computed difference between received and sent participants, by country and year.	12
6	Number of sent vs received participants, by population ratio.	12
7	Educational fields in Norway.	13
8	Top field of study, represented by labels on a map.	17
9	Erasmus web interface prototype.	23

1 Introduction and problem definition

1.1 Context

This report is the result of the curiosity and own experiences of six Erasmus+ students. We will analyze the data of the European Union program from 2014 to 2020, highlighting particular cases, proposing solutions to promote mobility and visualizing the future of student ¹ mobility.

The opportunity to support education, formation, youth and sports in Europe, was founded in 1987 by the European Commission, to finance international exchange programs [1]. The students of higher education institutions (universities and professional formation schools) can apply for a studies exchange in one of the 28 countries members of the European Union. Also, they can choose between the four countries of the European Economic Space; Liechtenstein, Norway, Iceland and Sweden; as well as some countries that are expecting to be members of the European Union: North Macedonia, Turkey and Serbia. Covering also, areas like student practices in companies and mobility of staff, such as teachers and professors.

The stay duration may vary between a period of weeks, months or even years, depending on if the student may apply for exchange practices or not. Participants may be able to apply also for scholarships and grants whose quantity will be different depending on the destination country and the economic capabilities of the student.

Besides of exchange students programme, there are also another programmes within Erasmus+ with the purpose of promote mobility in the European Union [2]:

- Master programs and doctorate Erasmus MUNDUS.
- Erasmus for young entrepreneurs.
- Comenius agreements for schools of childhood education, primary education and secondary education.
- Leonardo da Vinci destined to laboral practices.
- Grundtvig destined to adults' education.

These programs are a great experience for students from all over Europe, to continue with their studies or do practices, in addition to experiencing a new culture and traditions of a foreign country [3].

¹We want to clarify that when we are talking about countries or cities receiving or sending “students” we are talking about both students and every kind of staff, such as teachers, that can take this opportunity of taking part of the Erasmus experience.

1.2 Motivation

This project is framed within the subject TDT4259 - Applied Data Science, and arises with the intention of performing a realistic and problem-oriented data analysis, contributing to develop practical skills to configure the relevant tools/technologies, preprocess the data and perform the analysis. In this sense, the topic chosen for the project seemed to us to be a very good choice, since it meets the proposed learning objectives.

The other major motivation has to do with the inequalities and unbalanced rates that we saw in the data after a brief exploratory analysis. Let's analyze some of these inequalities and ratios. We will not come up with solutions to avoid them as that is something that we want to talk about later in this report. Also we will have a look into the factors contributing to these inequalities in section 6.1.1.

- **Ratio 0.0.** We will see a huge amount of cities with ratio 0.0, that is, they do not send or they don't receive students. Most of them are small cities only receiving or sending no more than 10 students, probably because they do not have enough resources to receive new students or they are not as popular and known as other big cities can be in the same country; and they probably will not have many students to send abroad, either.
- **Huge ratios.** On the other hand there will be cities that only receive or send a minor number of students, and as a consequence they have huge ratios. Some of them are also small cities that only send or only receive one student and they get that big ratio due to it. So, they would be in the same situation as the latter mentioned.
- **Ratio 1.0.** Even though this should be the wanted ratio, it could include cities having a very small number of mobilities but the same sent than received. For example, a city sending one student and receiving one, is not what is ideal, but it will have a ratio of 1.0. In Europe we find countries with a huge difference between sent and received participants. This will make us choose what to help first, how to distribute our resources and, for it, we will need to do an exhaustive analysis of what causes these inequalities and who do we want to help more and less.

1.3 Previous assumptions

When embarking on a project of this nature, it is natural to have a number of preconceived ideas about the behavior of the data. These assumptions, although useful to start asking questions, must be checked to avoid erroneous analysis and bias. Some of the assumptions about the data with which we will be working are listed below:

- The best-known European cities are those that receive, in turn, the greatest number of Erasmus+ participants.
- All cities in the same country have a similar behavior regarding the mobility of participants.
- The extent to which a country participates in the Erasmus+ program is determined by its domestic economy, with the richest countries having the highest number of participants.
- The distribution of mobilities of a given city remains relatively constant in time windows as small as the one covered by the dataset (6 years).

2 Background

2.1 Project objectives

Our main objectives will be on the one hand, to **promote mobility**; this is, to encourage both students and university staff to move themselves abroad, maybe showing them the opportunities they will have in different foreign countries, we will explain how this could be done in the following sections. We also want to increase public awareness of the benefits of the Erasmus mobility program, so we encourage people to take part in it.

On the other hand, a key in our project is to **avoid inequalities**, caused by countries sending lots of students but receiving only a few and vice versa. We have already talked about them previously. Also, we have to keep an eye on those countries that both receive and send a small number of students. Finally, we do not want to increase the difference and inequalities between the latter mentioned, and those countries in the opposite situation, that both send and receive lots of students; so, we will see how to do this but we will not help those countries in a great situation as much as those in not as favorable situation. In relation to this, we will provide training and resources to educational institutions to enhance their capacity for hosting and sending students through Erasmus. It would be important also to collaborate with these institutions and governments to provide scholarships, language support and other initiatives to facilitate mobility.

2.2 Data strategy

The strategy followed to manage the data we are working with was to enable easy access to the data, having a copy of the data to make sure not to alter the original dataset. The dataset is big and is never renewed because of being historical data, although each year new data rows should be included in the historical dataset. There is no worry about privacy in this particular case, as the data given is open source and has already been anonymized.

Given this data and our objectives, there is no reason to look for a complex deep-learning model, as our aim is not to predict but to explain and understand, we want to produce valuable insights on the Erasmus program data. That is why the data strategy includes a wide variety of Exploratory Data Analysis (EDA) and interactive visualizations.

3 Method

3.1 Our data sources

The analyses carried out are fed from three main datasets. The first one, aligned with the thematic focus of this project, contains the raw data on Erasmus+ mobility for students (study exchanges and traineeships) and staff (teaching assignments and staff training) in 2014-2020 [4]. The data contains information on various fields: age, gender, duration, special needs, subject area, level of studies, country of origin and destination, etc. It is a Comma-Separated Values (CSV) file containing records corresponding to more than four million different mobilities, although in the framework of this project we will discard all those referring to mobilities involving non-European countries. Because of that, we got rid of the rows where either the receiving or the sending country is not in the European continent. Despite the availability of 20 variables, many of them were discarded because they were not useful for our work. Cases of mobilities belonging to the year 2020 (due to their atypical nature because of COVID-19) and cases associated with countries with a very low number of mobilities (Switzerland, to be specific) were also discarded. The mobilities recorded in the dataset vary widely in duration (weeks, months or even years), but the bulk of the data is made up of stays of a semester or a full academic year. The dataset is written in English, which makes it easier for the reader of this report to understand.

The second dataset provides us with geolocation data for the European cities present in the mobilities set [5]. Specifically, it gives us the opportunity to use the longitude and latitude of each city to plot them on a map and create interactive visualizations. Although we have geolocation data for cities all over the world, we are only interested in intersecting the first dataset with European cities. We know which cities are European thanks to the ISO2 code that accompanies them (this code is a global standard for naming the different countries in the world).

3.2 Tools

The bulk of this project was carried out using Python 3. From the first exploratory analysis to the more advanced analyses, including data pre-processing and visualizations, this programming language proved to be very convenient for the task at hand.

Python was used in a Jupyter environment, as it facilitated code sharing by packaging it in notebooks that can include Markdown editing and quick export to HTML format. In addition, support for third-party widgets made it possible to create interactive visualizations, through the `Keplergl` library. The set of libraries used also includes `Pandas`, `Matplotlib` and `Plotly`.

3.3 Data pre-processing

Before jumping to the visualizations, we have to prepare the data so that they are useful for the purpose we want to give them. We must transform the raw data into data structures suitable for subsequent analysis.

In the case of the first dataset used, which contains the Erasmus mobilities data, the cleaning consisted of the following steps:

- Getting rid of dispensable variables, leaving only those related to the date of the mobility, its duration, the field, educational level and profile of the participant, and the geographical data of the cities involved.
- Converting city names to lowercase (for both Receiving City and Sending City).
- Deleting rows with cities that do not appear in the second dataset, i.e. discard all mobilities from or to cities for which we do not have information on the city.
- The column “Field of studies” has a lot of different categories, around 10.000 different ones, also there were a lot of them that represented the same category but had a spelling mistake or was just written a little different. To solve that we chose a list of the 12 study disciplines and using regular expressions we assigned the original values to one of the 12 categories. The academic fields we ended up having were: “Agriculture and Environment”, “Architecture and Building”, “Creative Arts”, “Education”, “Engineering”, “Food and Hospitality”, “Health”, “Information Technology”, “Management and Commerce”, “Management and Commerce”, “Natural and Physical Sciences”, “Society and Culture”.
- A precomputed dataset is created, based on the first one, which was a grouped dataset by country and year, so that we could know per country and year how many exchanges happened. For this dataset we included the following information: Country, year, sent students, received students, the ratio (received / sent), total exchanges (received + sent), and the difference (received - sent). This information was proven useful to create different exploratory analysis plots and having it precomputed was helpful as we didn’t have to compute it every time.

In the case of the second dataset used, which provides us with geographic coordinates to visualize the cities on a map, the cleaning consisted of the following steps:

- Discarding the dispensable variables, leaving only the name of the city, the name of the country, its ISO2 code (which identifies the country with two characters), latitude and longitude.
- Discarding all rows corresponding to non-European cities. To do this, we contrasted the ISO2 code of each of them with a list found on the Internet containing the ISO2 code for all European countries.
- Converting city names to lowercase.
- Deleting all rows except those where “city” appears in both the “Sending City” column and the “Receiving City” column of the mobilities dataset, i.e. discard the information of cities without mobilities, as they are meaningless for us in this project.

Now that the data from both sources has been dealt with, we proceed to intersect both sets. The goal is to have a single pandas DataFrame, containing the data of the mobilities and the coordinates of their cities. This pandas DataFrame will be different for each of the visualizations, however.

On the one hand, we will show dots on a map representing the European cities participating in the mobilities, colored according to a coefficient calculated as the ratio between received and sent. For this, we will start by creating a dictionary to count the occurrences of each city in ‘sending’ and ‘receiving’, which will be used to create a second dictionary, with the city as key

and a list with the number of sent, the number of received and the ratio calculated as the quotient between received/sent, as value. This dictionary is converted to DataFrame and merged with the geolocation dataset through the “city” column. The result is a DataFrame with the following columns: “city” (city name), “lat” (latitude where the city is located), “lng” (longitude where the city is located), “country” (country where the city is located), “iso2” (code associated to the country), “sending” (integer value representing the number of outgoing participants in that city), “receiving” (integer value representing the number of incoming participants in that city) and “ratio” (float coefficient calculated for that city). A brief exploratory analysis reveals an unexpectedly wide range of ratios, ranging from 0.0 to 1169.0. Further probing reveals that cities with very high ratios are few, and are determined to be anomalous, thus eliminating them from the DataFrame. This small modification reduces the range of ratios to 0.0-32.0, something that seems more reasonable and does not harm the color scale (a very wide range of values would force a color to hold very disparate ratios, making it difficult to observe small differences between cities).

On the other hand, it is of interest to connect on a map the pairs of cities between which mobilities exist. For this, a DataFrame with the following variables “Sending City”, “sending_lat”, “sending_lng”, “Receiving City”, “receiving_lat” and “receiving_lng” is needed. After building it by merging the mobilities DataFrame and the DataFrame containing the ratios, we can count the number of times a city pair is repeated (grouping by “sending_city” and “receiving_city”) and display it in a new column called “mobilities”. Once the repetitions are counted, redundant rows can be removed. The data cleaning does not end here, because as with the data in the visualization explained above, the range of values in the “mobilities” column is very wide, damaging the color palette. The solution is again to eliminate the rows with unusually high values in the “mobilities” column - since there are few of them - thus reducing the range of values for this variable (after this fix, the values range from 69 to 377, which is “manageable”).

4 Visualizations

The nature of this project is largely descriptive, which gives us the opportunity to use visualization as a means of communication. A very convenient way to visualize geospatial data is through the use of maps. We will rely on the Python *keplergl* library for this purpose. We will present two map-based visualizations. The first one shows the cities of European mobilities as dots, colored according to a ratio calculated as the quotient between exchanges received and exchanges sent. A lighter (more yellowish) color indicates higher values of this ratio, and a darker color indicates lower values. The second visualization draws an arc between those cities paired by a mobility, these arcs being colored according to the number of mobilities they represent. If a mobility between two cities has been repeated many times, the arc representing it will have a lighter color than the one used to represent an infrequent mobility.

The library used for the creation of maps, *keplergl*, allows exporting to HTML files, a highly convenient format for interactive visualizations since the user can exploit the functionalities that the designer makes available. For the arcs map, for example, we recommend enabling the 3D view, as it offers a better perspective of the origin and destination of the arcs. Another button on the same right side panel shows the legend, useful in both maps. Also, access to the left side menu, which contains a wide variety of tools for modifying map layers and their appearance, has been enabled for both visualizations. From this menu we can filter to show only the observations within a delimited range of values, or activate the “brush” option to browse the map showing only the data of the areas overflowed with the cursor. Finally, a geocoder has also been included in the dot map, so that you can search for a city by its name or by its geographic coordinates, and the center of the map will move to that point. Below are snapshots of the two interactive visualizations discussed:



Figure 1: Dot distribution map, colored by receiving/sending ratio.



Figure 2: Arc map, colored by number of mobilities.

The ratio used for plotting in the kepler maps was of high importance for the project, as not all countries are in the same situation in the Erasmus program. Some countries send lots of students but barely receive any, and that is something that should be considered for later conclusions. That is why a large part of the exploratory data analysis was conducted under this scope.

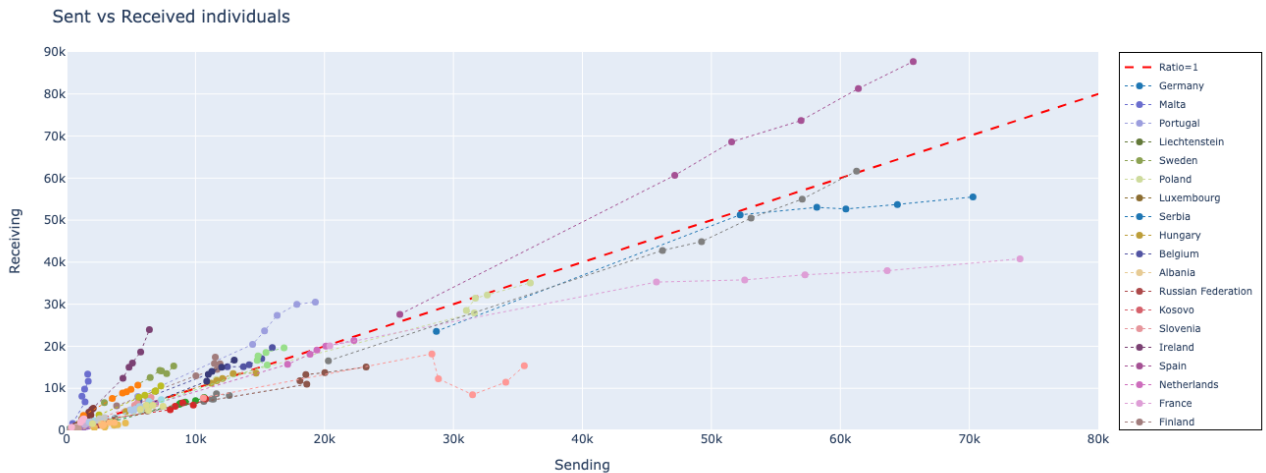


Figure 3: Sent versus received individuals.

As the ratio really represents how many students are received per student sent, a scatter plot was created. In this plot every dot is a country in a year, all years for the same country share color and are joined by a dotted line. The X axis is the number of sent students and the y axis is the number of received students, a dashed diagonal line can be seen representing the space

where the ratio is equal to one. This means that data points on the bottom right corner are countries and years in which more students were sent than received and the opposite for the other corner. This plot can be useful to see the trend for every country, their ratio values and their similarities or differences with other countries.

In order to capture the information about which countries had either too high or too low ratios, a bar plot representing the ratio for every country was done. Here, countries are in the y axis, ratios in the x axis and the plot is faceted by columns per year. A dotted line is used to indicate the ratio equal to one. This visualization provided information on cases where the ratio is high above one, but is not as helpful to detect cases with lower ratio values.

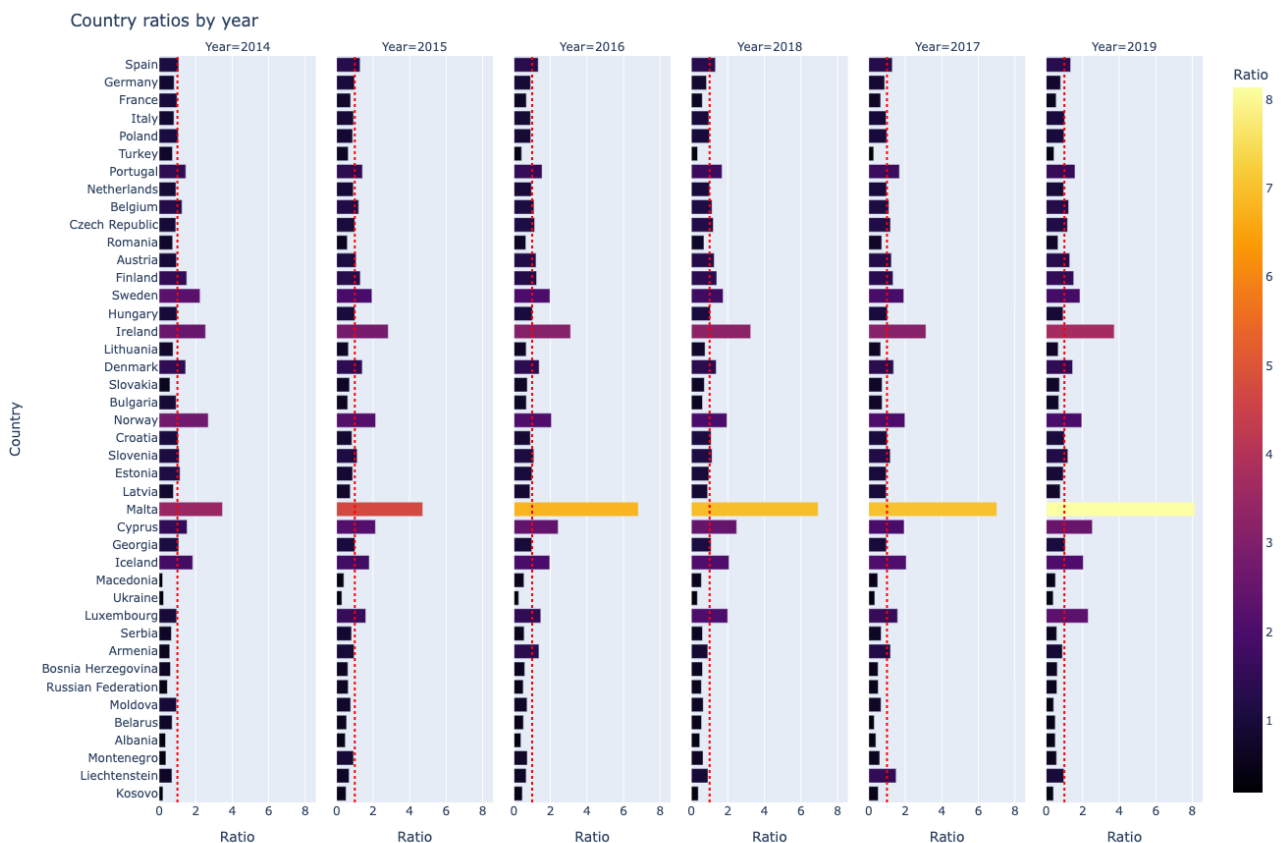


Figure 4: Country ratios, by country and year.

To solve the issue that the previous graph proposed, a similar plot is used. The facet was not used anymore and instead of the ratio, the value represented is the difference between received and sent students, which makes it easier to perceive countries that send more than receive. In particular, cases such as France and Turkey are especially notable. The only problem about this visualization is that bigger countries could appear more relevant than smaller countries.

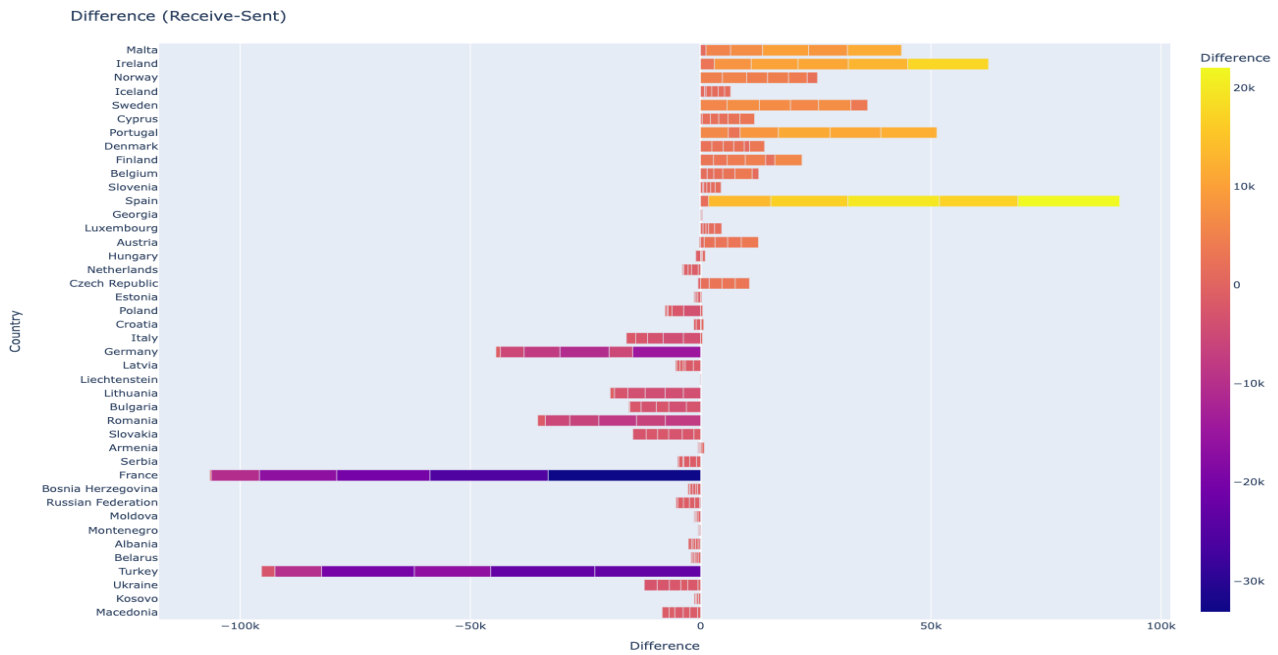


Figure 5: Computed difference between received and sent participants, by country and year.

To finish with the received/sent ratio exploration, it was decided that a problem for the visualization could be the fact that not all countries have the same amount of population, so smaller countries are expected to have more variability in the ratios than countries with more total population. That is the reason why it was decided to include information about the country's population. Two new variables were created: `rec_per` and `sent_per`. These values are what percentage of the country's population does the group of students received or sent represent. For the most part of the countries, this groups represent a very small percentage of the population, especially in bigger countries, but for small and touristic countries this plot is very valuable as you do not only see a simplified version of the ratio, you can also see the total flow of students represents for the totality of the country (e.g., Academic exchanges in Malta have a big importance). If the bar are barely visible, then that means that the country does not participate much in the Erasmus program (e.g., Ukraine).

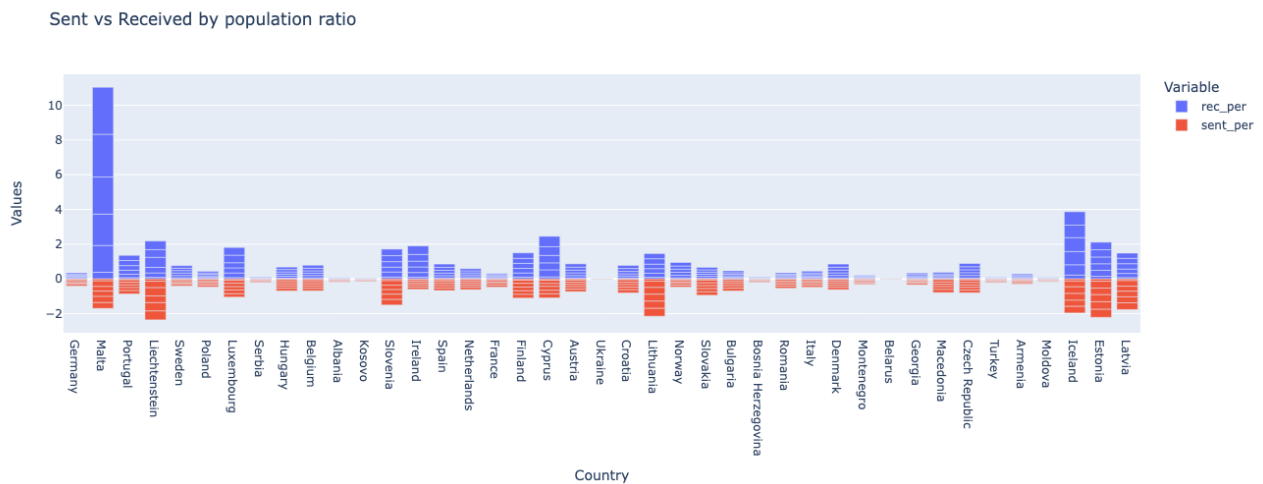


Figure 6: Number of sent vs received participants, by population ratio.

For the academic perspective, a visualization of the variable “Field of Education” was required. To solve that, a grouped bar chart was chosen, as it is really useful to compare numerical and categorical data. Proportions had to be used instead of total values, as not all countries have the same amount of exchanges, also, it had to be precomputed the global mean proportion of that field of studies for all fields. After that, a dropdown menu was implemented to select a country and for each field of studies there is a blue and a red bar. The blue one represents the percentage of exchanges that the field represents (from the total of exchanges) and the red one the mean percentage for all countries. The use of the plot is to check whether the country you are looking is receiving more or less students for each field than the rest of the countries, if you are wondering whether to go to Norway if you are studying Architecture you should, because Norway is receiving more architects than the average European country.

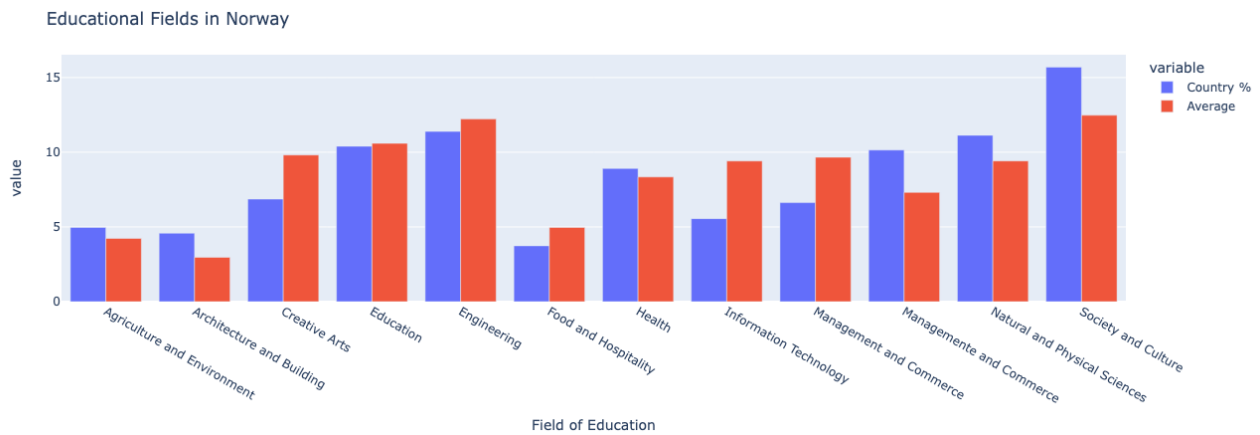


Figure 7: Educational fields in Norway.

Another visualization involving the field of education was done. As the previous plot did not accomplish expressing the information with just a single view, a map was created to cover that aspect. It was expected to view the top field of education for every country, and to know which one, the difference between the country’s percentage and the global average was used. The reason behind not using just the country’s percentage is that the Society and Culture field are much more common so it is always the field with the most amount of received exchanges. In the map, a label with the top field of education can be seen above the capital of every country, providing a quick visualization for all the interested.

5 Analysis

After analyzing all the data as a whole, we have highlighted certain countries that have a peculiar situation, some positively and some negatively.

5.1 Malta

The first one is Malta. As we can see in figure 3, Malta stands out compared to all other countries in terms of its ratio. This country has sent a total of 8231 students, an exaggeratedly small number compared to the 55832 received. The ratio is 6.78 and the difference is 47601 people. The reasons behind this are the following:

- The cost of living in Malta is relatively lower than in other European countries, making it an attractive destination for international students.
- Skilled graduates who stay or move to Malta can benefit from tax breaks, fast-track visas, and free support to help start a new business.
- Located in the Mediterranean region, offers easy access to Europe, North Africa, and the Middle East, making it a popular choice for students worldwide.
- As Malta was a British colony, the majority of locals speak fluent English alongside Maltese, facilitating communication for international students.
- Safety is a notable feature, as Malta is regarded as one of the safest countries globally, with low crime rates and minimal natural disaster occurrences.

It should also be borne in mind that this country has a very small population, approximately 490,000, which is why the ratio is so high.

5.2 Ukraine

This case is the most extreme, even before the war started. They have sent 19188 and have received 6093 students, this is a difference of -13095 people, so the ratio is 0.32. As we have commented, the data we are analyzing is from 2014 to 2020 and here the situation in the country was already very conflicted, Ukraine was facing a series of events that impacted its political and social stability. The Ukrainian Revolution of 2014, known as “Euromaidan”, triggered mass protests and a change of government. Russia’s annexation of Crimea generated international tension, followed by conflicts in eastern Ukraine between separatist groups and the government.

These events have affected the perception of stability in the country, probably influencing the decision of international students to avoid Ukraine as an exchange destination. In turn, Ukrainian students sought opportunities abroad in search of more stable and secure environments.

5.3 France

There is a lot of movement of students but it highlights the fact of the difference between sent and received. They have sent 349295 and received 225106 students, the ratio is 0.64. It has a very low number of incoming students and the main reason for this is the language, it is not a

country where English is very standardized. Moreover, the situation in the city of Paris is also affected by terrorist attacks, and security is often compromise, this causes some rejection and fear for potential exchange students.

5.4 Cyprus

In this case the ratio is also positive 2.31, because they have sent only 9850 students and received 22711. Again, we are talking about a country with a small population, around 870,000, so it does not have as much capacity as other countries to send students. The country is reputed for its high-quality public education. It's one of the top three countries in Europe considering the amount of GDP it spends on education. Also, the cost of education and living are low.

5.5 Ireland

The last country we will look at is Ireland. It has a very high ratio, 3.14. It receives 98849 and sends 31485, a difference of 67364 students. In most cases, this country replaces the choice of the UK as a destination. Although Brexit will happen in the final year of our dataset (end 2019-early 2020), the beginnings of this separation were already an indication of the rise of students in Ireland to replace the UK. Moreover, it is cheaper than US, Canada and the UK.

Highlighting these cases is crucial in understanding the rest of our data. Most countries have similarities with at least one of these cases and could be thought of as the same behavioral group. For example, Germany is a France-like country, they both have one of the biggest flows of participants but their ratio is below one. It would be nice to try to increase the number of participants these countries receive so that their ratio is closer to one.

Another group would be the Ukraine-like countries, in this group we could include cases such as Serbia and Macedonia, where the total number of participants is fairly low and the ratios are also below one. Those would be the countries that should be focused the most, as they barely participate in the Erasmus programme and could benefit more from it. Finally there are the Malta-like group, which are the cases of small countries with high ratios due to being of high interest for touristic, economic or climate factors. Here it would be included places such as Ireland, Cyprus, Iceland or Portugal. As their ratios are really high above one, it could be interesting to promote sending more students from these countries.

6 Interpretation and recommendations

6.1 Recommended measures & implementation plan

6.1.1 Factors when choosing a destination

The main factors when choosing the destination for European mobility tend to vary depending on the preferences of each student. While a high number of students look for destinations with parties and a lot of university life, there is another large majority that seeks to escape from their routine and go to different and exotic destinations that they would not consider at any other time in their lives. However, there are factors that tend to be generic for all students and have a great weight when making the decision. These factors are as follows:

- **Climate:** some students prefer warmer or colder climates. The choice of destination may depend on personal tolerance to climate and seasons. Many students prioritize destinations that do not have extreme climates, i.e., neither too hot nor too cold, so that adaptation to that climate is not hard or difficult to cope with. There is also a group of people who are looking for experiences in climatic environments that challenge them or give them a new perspective.
- **Economic issues:** As we already know, although we are analyzing only European options, each country has its own range of prices, and the monthly expense is different, even in each city. Not everyone can afford all the prices we are talking about, which makes this a crucial issue when making the decision. Although there is the possibility of the scholarship, as students ourselves, we can confirm that it is a great help, but it is far from being enough to finance the whole year with the main expenses of accommodation, food and transportation. Therefore, cities with a low cost of living are usually at the top of the rankings for students. In recent years we could observe how less crowded destinations such as Poland became more famous and increasingly popular for students who could not afford to pay for accommodations in other more well-known university cities.
- **Academic Issues:** The quality of academic institutions at the destination is crucial for students participating in European mobility programs. Global and national reputation, as well as specialization and strengths in specific areas of study, are key factors that students consider when making decisions about their academic destination. In terms of the availability of relevant courses, students face challenges when validating credits earned abroad. Lack of full validation can affect a student's normal progression. Course variety is also essential; the absence of broad and diverse offerings can limit a student's ability to make the most of their academic experience. The inclusion of interactive bar plots addresses these concerns by showing the strength of each educational field by country. Students can evaluate academic offerings based on their area of study and compare European average, for example, to Norway's strength, as seen in figure 7. We believe that offering this information to students could be a great help, so that they can see in a very intuitive and quick way, if it is going to be easy or difficult to find classes of their degree in their country of destination. As well as offering a map like this in which the capital of each country is shown and there is a label that shows the most important field in that country, so that those who really prioritize going to a country with a high interest in their degree, can see it directly:



Figure 8: Top field of study, represented by labels on a map.

- **Language of instruction:** Fluency in the language of instruction is vital. Some students seek to improve their skills in a specific language, while others prefer programs in their native language or in English. For this same reason we can observe how, for example, in the preferences of students of Spanish nationality, Italy is usually at the top of their list of destinations, due to the similarity between the two languages. Although this is not a main reason, it can be an advantage or disadvantage when choosing a destination, especially a few decades ago, when we did not have the facilities we have today to translate texts quickly.
- **Culture:** Exposure to new cultures is often one of the main attractions. Students seek destinations that allow them to immerse themselves in a different culture and broaden their horizons. The availability of cultural events, festivals, and recreational activities can influence the choice. Cities with a rich history and cultural heritage may be attractive to students interested in history and architecture.

6.1.2 How can we promote mobility?

Fostering mobility without positive feedback involves implementing effective promotion efforts and strategic resource allocation. Creating awareness through campaigns that emphasize the benefits of European mobility and sharing success stories of students who have thrived in unconventional destinations is crucial. Providing detailed information about lesser-known destinations, highlighting affordable living costs, unique academic opportunities, and cultural experiences, is essential. Collaboration with government and university institutions to ensure clear access to information on scholarships and financial aid for less popular destinations is important. Building strong academic collaborations between institutions in less-explored locations and students' home countries can motivate consideration of these options. Offering personalized guidance, considering individual preferences and academic goals, is key. Actively participating in educational events, both physical and virtual, to promote unconventional destinations and organizing events with students sharing positive experiences in these places can be impactful. Developing interactive platforms for students to explore and compare destinations, academic offerings, and cultural aspects is beneficial. Leveraging social media and digital marketing to disseminate information and encourage community engagement enhances outreach efforts. By implementing these strategies, student mobility to less conventional destinations can be

encouraged and provide students with the opportunity to enrich their academic and personal experiences. In essence, a comprehensive approach that combines awareness, information, financial support, and community involvement is crucial for promoting mobility without relying on positive feedback mechanisms.

The improvement of the Erasmus programs and all the resources and all the manages that this involves, demands an integral approach that is able to deal not just with individual needs of the students, but also the institutional dynamics that have an influence in the equitable and rewarding participation.

In this section, we will also present some strategic solutions, both on an institutional level and student level, focusing on the development and the improvement of the mobility so that we can reduce the inequalities and guarantee a much more inclusive and beneficial Erasmus experience for all the students, teachers, families, and institutions.

Together, these solutions aim not only to increase participation in the Erasmus programs but also to ensure that this participation is equitable, meaningful, and transformative at both institutional and individual levels.

Institutional level:

On one hand, we have come up with multiple proposals, focused on an institutional level, which they are related to the distribution and management of all the resources, both economic and social, of the different countries, so that they all can contribute to the improvement of the programs from a perspective related to the corresponded institutions.

1. Equitative Incentive

- The Introduction of financial incentives for the educational institutions in the countries with a disproportionately low participation on the Erasmus programs, where they can take some advantages from the collaboration on these programs, as can be “Investigation resources”, “More publicity for the city and University”...
- Another good proposal could be the establishment of clear criteria that could reward the equity in the exchange students and professors, as the implementation of the specific guidelines that could define how will the fairness in the participation of the educational institutions be evaluated and recognized in the Erasmus programs.

2. Strategic Counsel

- To implement strategic advice programs for educational institutions, could be a positive way of reaching the attention of the educational institution, so that they could learn the orientation that these programs could offer them about how to improve their participation in the Erasmus Programs, so that they could notice their weaknesses and fix them.
- Related to this matter, another positive proposal could be to facilitate the collaboration between institutions so that they could exchange good practices and successful experiences. For example, in the case that a university usually receives many students from Erasmus programs, but on the other hand, it sends very few, it could make contact with another university that is in the opposite situation, sends many

students but receives very few. These institutions could find arrangements that could benefit both in order to improve their strategies in these matters.

3. Specific Programs Development

- To promote the creation of specific programs adapted to the unique circumstances and specific challenges that the educational institutions in the counties with low participation rate may have.
- The idea is not to apply generic solutions, but to develop initiatives to promote better inclusion and diversity for these specific institutions for the Erasmus programs from all over Europe.

Student level:

On the other hand, we will also be focusing on the students perspective, taking into account multiple factors in which the students will focus by the time they choose an Erasmus destination. This will contribute in a very useful way when we make a planning for the institutions to improve their programs as well.

1. Consciousness - Raising campaigns

- Launch awareness campaigns targeted at students, highlighting the professional, academic, and personal benefits of participating in Erasmus programs, where the students and their families can visualize all the destinations, with their specific benefits as well.
- Organize presentations coached by students involved in these experiences so that they can give their feedback about their personal experiences in the program, so that they can inspire and motivate others to board into this rewarding experience.

2. Resource access

- One of the most important matters to focus on, is the economic factor for the students, therefore, another initiative that could help, is to facilitate the access to scholarships for students, making sure that, for example, all the basic outlays, like housing, food, transport..., may be covered by this financing, and will not be an obstacle to for the participation in the program.

3. Personalized Mentoring

- In many cases, a lot of individual problems come up during these programs related to institutional matters, such as problems related to the Learning Agreement, Arrival Confirmation, change of subjects..., this is why, a very positive proposal, could be to implement mentoring programs where the students can go to when they find themselves in any of these situations.
- These programs should offer personalized support to fulfill all the problems that the students may have, so that they can have a better and less stressful.

6.2 Limitations and future analysis

There are some limitations involving the objectives of our project. One of the key factors is that the dataset we use is constructed by many different sources and years and it is really common to express the same information in different ways, as there is no standard to follow. A very relevant example is the variable “City Name”, where it is common to find the same city with different names across the dataset, which caused struggles on the data wrangling and made us lose some information that could not be added into the visualizations. Another example is the variable “Field of Education” which not only had the same problem as the cities, but also had too many categories (around tens of thousands) which made it completely unmanageable, that is why we reconverted the field into another one with fewer categories. To solve this and to have more available information, It would be nice to standardize the dataset before beginning the data analysis.

For further work it would be interesting to explore a wider range of information, especially to have the possibility to examine other factors that affect the exchanges such as economic, climatic or touristic and include them into the analysis.

Something that could have also added value to the project would have been to develop a dashboard so that managers of the Erasmus program could have an easier visualization, which could help their decision-making.

7 Conclusion

In any project, the most difficult part is not to create it, but to maintain it. Our model is not predictive, and therefore the way to evaluate its performance is more complicated. The graphics have to be logical and report clearly their conclusions. Each year, normally, the data is going to change and be amplified, and we need a control. It will be analyzed, taking into account that wars, pandemics or any other event that may greatly affect the scale of the data may occur. This analysis has to be very simple, quick and accessible when adding the new dataset.

Another very important step is to check the validity of the information and not to work with erroneous, false or mistaken data.

8 Appendices

8.1 Appendix 1: Dot distribution map, colored by receiving/sending ratio

Through this [link](#) you can download the HTML file for this interactive visualization.

8.2 Appendix 2: Arc map, colored my number of mobilities

Through this [link](#) you can download the HTML file for this interactive visualization.

8.3 Appendix 3: Sent versus received individuals

Through this [link](#) you can download the HTML file for this interactive visualization.

8.4 Appendix 4: Country ratios, by country and year

Through this [link](#) you can download the HTML file for this interactive visualization.

8.5 Appendix 5: Computed difference between received and sent participants, by country and year

Through this [link](#) you can download the HTML file for this interactive visualization.

8.6 Appendix 6: Number of sent vs received participants, by population ratio

Through this [link](#) you can download the HTML file for this interactive visualization.

8.7 Appendix 7: Top field of study, represented by labels on a map

Through this [link](#) you can download the HTML file for this interactive visualization.

8.8 Appendix 8: HTML format Jupyter Notebook with the pre-processing and visualization code for dot and arc maps.

Through this [link](#) you can download the HTML file for this Jupyter Notebook.

8.9 Appendix 9: Erasmus web interface prototype

Here we will be showing a simple example of a website where the students could obtain information about all the things related to Erasmus programs and experiences. Also where they can find information about institutional agreements that might get complicated to complete once you are in the destination or before arriving there:

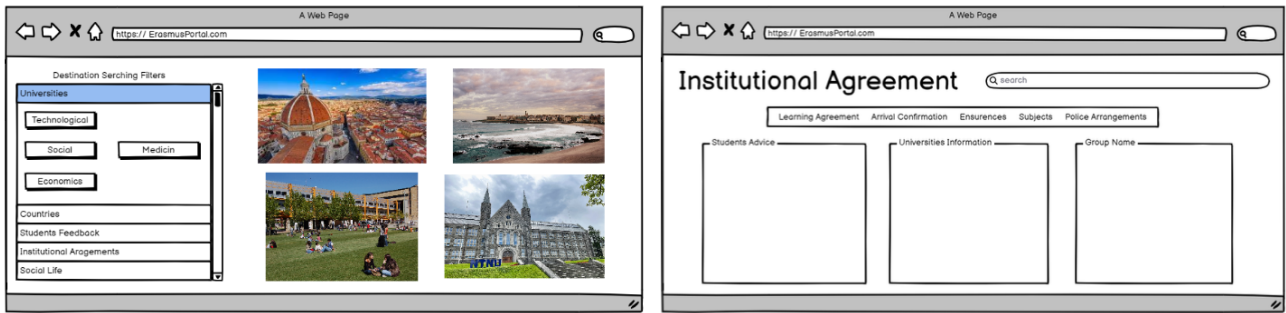


Figure 9: Erasmus web interface prototype.

References

- [1] “Qué es un Erasmus: motivos que te empujarán a hacer uno.” [Online]. Available: <https://www.becas-santander.com/es/blog/que-es-un-erasmus.html>
- [2] “Programa Erasmus+. ¿Cuáles son las ventajas del Programa Erasmus+?” [Online]. Available: <https://www.ef.com.es/erasmus/programa/>
- [3] “¿Qué es Erasmus+? | Erasmus+.” [Online]. Available: <https://erasmus-plus.ec.europa.eu/es/about-erasmus/what-is-erasmus>
- [4] “Estadísticas de movilidad Erasmus 2014-2020 - Data Europa EU.” [Online]. Available: <https://data.europa.eu/data/datasets/erasmus-mobility-statistics-2014-2020?locale=es>
- [5] “World cities database.” [Online]. Available: <https://www.kaggle.com/datasets/juanmah/world-cities>