



*Department of Business Administration Bachelor's Thesis*

**Understanding Key Drivers Behind International Students'  
University Choice: A Case Study of LUISS**

**SUPERVISOR:**

**Prof. Antonio Majocchi**

**Candidate:**

**Mostafa Abdulaziz Yosuf**

**Academic Year 2024/2025**



|                                                                                                              |           |
|--------------------------------------------------------------------------------------------------------------|-----------|
| List of figures                                                                                              | 5         |
| Figure 1: Linear Regression                                                                                  | 5         |
| Figure 2: Random Forest                                                                                      | 5         |
| Figure 3: XGBoost                                                                                            | 5         |
| Figure 4: Shap Summary Plot                                                                                  | 5         |
| Figure 5: Shap Dependence Plot 1                                                                             | 5         |
| Figure 6: Shap Dependence Plot 2                                                                             | 5         |
| List of tables                                                                                               | 5         |
| Table 1: Key findings of macro-level drivers                                                                 | 5         |
| Table 2: Key findings of challenges and barriers                                                             | 5         |
| Table 3: Models performance results                                                                          | 5         |
| <b>Chapter 1 – Theoretical Foundations of International Students’ University Choice: A Literature Review</b> | <b>6</b>  |
| 1. Introduction                                                                                              | 6         |
| 2. Theoretical Frameworks                                                                                    | 8         |
| 3. Macro-Level Drivers                                                                                       | 12        |
| 4. Micro-Level Drivers                                                                                       | 17        |
| 5. Challenges and Barriers                                                                                   | 20        |
| 6. Summary of Main Findings                                                                                  | 23        |
| 7. Conclusion                                                                                                | 26        |
| <b>Chapter 2 - Research Methodology</b>                                                                      | <b>29</b> |
| 1. Introduction                                                                                              | 29        |
| 2. Data Description and Data Collection                                                                      | 29        |
| 2.1 Overview of the Dataset                                                                                  | 29        |
| 2.2 Data Sources                                                                                             | 30        |
| 2.2.1 Internal Data                                                                                          | 30        |
| 2.3 Key Variables                                                                                            | 31        |
| 2.3.1 Internal Variables (Columns A–AB)                                                                      | 31        |
| 2.3.2 External Variables (Columns AC–AU)                                                                     | 32        |
| 2.4 Data Structure and Format                                                                                | 32        |
| 2.5 Data Cleaning and Preparation                                                                            | 33        |
| 3. Data Analysis                                                                                             | 33        |
| 3.1 Data Processing                                                                                          | 33        |
| 3.2 Feature Selection and Engineering                                                                        | 34        |
| 3.3 Statistical Modeling and Analytical Techniques                                                           | 35        |
| 3.4 Visualization                                                                                            | 35        |
| 3.5 Software and Analytical Tools                                                                            | 36        |
| 4. Justification of Choices                                                                                  | 36        |
| <b>Chapter 3 - Key findings of the LUISS case: Results</b>                                                   | <b>37</b> |
| 1. Introduction to the Results                                                                               | 37        |

|                                                                       |           |
|-----------------------------------------------------------------------|-----------|
| 2. Model Performance Comparison                                       | 38        |
| 3. Feature Importance Analysis                                        | 40        |
| 3.1 Linear Regression Feature Importance                              | 40        |
| 3.2 Random Forest Feature Importance                                  | 41        |
| 3.3 XGBoost Feature Importance                                        | 42        |
| 4. SHAP Analysis for Model Interpretability                           | 42        |
| 4.1. SHAP Summary Plot                                                | 42        |
| 4.2. SHAP Dependence Plots                                            | 43        |
| 5. Synthesis and Key Findings                                         | 45        |
| 5.1 Summary of Main Findings                                          | 46        |
| 5.2 Model Performance                                                 | 46        |
| 5.3 Feature Importance                                                | 46        |
| 5.4 SHAP Analysis                                                     | 46        |
| 5.6 Patterns and Unexpected Results                                   | 47        |
| 5.7 Relation to Research Question                                     | 47        |
| <b>Chapter 4 – Conclusion</b>                                         | <b>47</b> |
| 1. Introduction to the Conclusion                                     | 47        |
| 2. Restating the Research Objective                                   | 47        |
| 3. Summary of Key Findings                                            | 47        |
| 4. Significance and Contributions                                     | 48        |
| 5. Limitations                                                        | 48        |
| 6. Recommendations for LUISS University and Future Research           | 48        |
| 6.1 Prioritize High-Density, Economically Viable Countries            | 49        |
| 6.2 Expand Strategic Engagement in Geographically Proximate Countries | 49        |
| 6.3 De-emphasize Diaspora-Driven Strategies                           | 49        |
| 6.4. Expand Need-Based Scholarships to Reduce Financial Barriers      | 49        |
| 6.5 Future Research                                                   | 49        |
| References                                                            | 50        |

## List of figures

**Figure 1:** *Linear Regression*

**Figure 2:** *Random Forest*

**Figure 3:** *XGBoost*

**Figure 4:** *Shap Summary Plot*

**Figure 5:** *Shap Dependence Plot 1*

**Figure 6:** *Shap Dependence Plot 2*

## List of tables

**Table 1:** Key findings of macro-level drivers

**Table 2:** Key findings of challenges and barriers

**Table 3:** Models performance results

# **Chapter 1 – Theoretical Foundations of International Students' University Choice: A Literature Review**

## **1. Introduction**

The globalization of higher education has resulted in an unprecedented increase in international student mobility and has turned the movement of students across borders into a research and policy focal point in recent decades (Restaino et al., 2020; Nikou & Luukkonen, 2024). Increasingly, countries and universities are focusing on recruitment of international students because of their impact on institutional diversity, academic excellence and as they must respond to an aging population and labor market needs (Nikou & Luukkonen 2024). In Europe, the Erasmus+ programme is another organized effort to support student mobility. The Erasmus+ program started in 1987 with 3,200 participants and, by 2023, had grown to support over 15 million participants, fostering academic exchange and intercultural competency amongst students (European Commission, 2024; European Parliament, 2020).

This surge in international student mobility can be attributed to multiple sources and levels of influence. At the individual level, students are motivated by opportunities to gain new skills, increase employability, and seek intercultural experiences, which are becoming more and more important within a globalized workforce (Restaino et al., 2020; Nikou & Luukkonen, 2024). From the perspective of host countries and institutions, international students are viewed as a strategic resource to support economic development and innovation, and foster the

internationalization of higher education systems (Nikou & Luukkonen, 2024). At a time when many countries, particularly those with an aging population and skilled labor shortages are trying to attract and retain international students as a key policy objective (European Parliament, 2020). To understand the reasons behind the location choices of international students (universities and the countries in which the universities exist), the push-pull factor model is often deployed. The push-pull factor model is a framework that differentiates between "push" factors - factors in the home country's environment which compel a student to find their thinking education outside of their home country, like lack of academic or employment opportunities, and "pull" factors - attractions to the host country or institution, like academic reputation, economic situation, university policies that legislate to motivation for international students (Nikou & Luukkonen, 2024). Recent empirical evidence in this area of research also emphasizes more contextual influences through a blended juxtaposition of social, cultural and personal values; and the academic barriers they may face in negotiating a new international cognitive environment, and/or mindset - including language barrier(s), social integration barrier(s) and/or administrative barrier(s) (Restaino et al., 2020; European Parliament, 2020).

In European contexts, the Erasmus+ mobility scheme is not solely restricted to fostering academic engagement but also plays an active role in developing a shared European identity through cognitive mobilizations and cross-border engagements (Mitchell, 2015; European Parliament, 2020). Research shows that student participation in mobility schemes (e.g., the Erasmus+ scheme) enhances academic competencies and broaden student's employment options and increases their flexibility, while the acquiring host institution also benefits in the long-run via increased internationalization and collaboration (European Commission, 2024; Restaino et al., 2020). Nevertheless, emergent student flows remain poorly distributed in Europe, with core

economies (e.g., Germany and France) account for 58% of total Erasmus students that were able to relocate due to their economic wealth and institutional interlinkages, while peripheral economies (e.g., Latvia and Lithuania) rely upon outbound mobility for engagement academically (Restaino et al., 2020).

Despite students benefitting from international mobility there are still inherent challenges. Initially, students will face challenges related to cognitive barriers (language), cultural adaptation and administrative processes - all of which are engaged within the context of the host country and can impact either their academic performance or burgeoning cognitive disposition about re-locating their lives to that of the host environment (European Parliament, 2020; Nikou & Luukkonen, 2024). It is helpful for host countries/institutions to comprehend the predisposition towards the engaged drivers and evident barriers that will assist in guiding effective marketing, support services, and retention mechanisms (Restaino et al., 2020).

This review of the literature will conclude with a summary of up-to-date research related to the key drivers affecting international student's selection of university, focusing on the macro (country and institution) and micro (personal and social) layers of influence on student's decision making. Recent comparative analysis and the frameworks. The review highlights both opportunities and barriers that universities must consider when seeking to recruit and retain talent; this will allow for a subsequent analysis and recommendations to take place.

## **2. Theoretical Frameworks**

In understanding international students' choice of university requires a multi-theoretical approach that considers the complexity of motivations, decision-making, and structural influences that

shape international student mobility. In the past couple of decades, several theoretical models have emerged and developed to explain the why students seek to study abroad and how they selected their destination and institution. Some of the more prominent theoretical models include push-pull theory (Mazzarol & Soutar, 2002), human capital theory, sociological and social network models as well as motivations-based frameworks, such as expectancy-value theory (EVT) and self-determination theory (SDT). Push-pull theory has become an increasingly popular model in the context of international mobility since it was first developed to help in understanding migration (Lee, 1966). Mazzarol and Soutar (2002) took this framework and applied it in the context of higher education and described push factors-negative home conditions such as limited educational opportunities, political instability, or a lack of specialized programs. Pull factors are positive characteristics of the host country or institution that appeal to the student such as academic reputation, prospects of employment, scholarships, and support policies. For instance, student mobility push would include students who are able to study abroad for the first time due to restrictive access to higher education in their home countries. On the other hand, as per Restaino et al. (2020), host countries like Germany and France have invested human and capital resources (i.e. institutional partnerships) into attracting students to study and participate in mobility opportunities. Restaino et al. (2020) also point out that nearly 58% of Erasmus students ended up in host core European countries due to their structural advantages. However, the push-pull model has been criticized due to its oversimplified construct of student choice, as it considers neither the individual student preferences nor their personal characteristics. Wilkins and Epps (2011) asserted that although structural facets of student mobility, like economic conditions, are important, social forces, family referrals, and institutional branding are also part of the contextual factors influencing student choice. Additionally, human capital theory is based

upon economic appeal and context, suggesting students consider their education as an investment in future earning potential (Becker, 1962). Students compare the monetary costs of studying abroad (tuition, living costs) against future earnings potential (higher employment and salary). For example, students in Malaysia looked to complete their higher education in English-taught programs at credible institutions as a result of the perceived career advantage, but as noted by Wilkins and Epps (2011), human capital theory fails to take into consideration the non-economic rationales for study abroad, such as cultural curiosity or personal development. For example, Wilkins and Epps reported 48% of Erasmus participants wanted to experience intercultural study experiences over financial considerations (Frontiers in Psychology, 2022).

Further, students rely upon their social networks to help them navigate their challenges in their host country, which directly or indirectly, influence their higher education choices. Students utilized bonds with co-national peers (bonding capital) and ties with host-country peers (bridging capital). In a study with Irish university students returning from their Erasmus abroad, Irish students with stronger bridging capital, were 30% more likely find local employment to finance their studies (Social Network Study, 2015). The Erasmus program exemplifies these network effects, where core countries like Germany attract a disproportionate share of students due to dense academic partnerships and policy alignment, while peripheral nations such as Latvia struggle to compete despite offering niche programs (Restaino et al., 2020).

Capabilities-based and motivation-based approaches also add some additional insights into student motivations. Expectancy-value theory (EVT) is a motivational concept whereby an individual's decisions (expectancy) to study abroad is based on the degree to which they attach

value to studying abroad for future employment opportunities, cultural exposure, or personal improvement. Their study proposition was also supported by self-determination theory (SDT), where there was a variance between autonomous motivations (e.g., intrinsic interest and the fulfilment of identified goals) and controlled motivations (e.g., coercion from family, or external consequences). Wilkins and Epps concluded that students who studied abroad under autonomy, had 28% increased satisfaction than those influenced by flows beyond their control (Frontiers in Psychology, 2022) and suggested that autonomous motivations are relevant for examining student choice. These theorisations offered three generally accepted approaches to studying international education, however, they lacked consideration of the issues raised in this thesis: push-pull theory underestimates the macro and micro relationship influencing student higher education access; human capital theory presupposes non-economic incentives, like cultural interests; social network theory downplays factors of peer influence in decision-making; and motivation-based approaches often underrepresent demographic complexity, including gender and age (Frontiers in Psychology, 2022; Li & Bray, 2007).

In conclusion, a multi-theoretical lens articulating push-pull scenarios, human capital investment, social network circumstances, and motivation-based contexts offers the best potential framework for examining drivers of international student mobility. The multi-theory model provides context to aid a deeper final analysis of drivers that influence student choice with a larger synthesis that encompasses structural, economic and social drivers, and personal features associated with university choice.

### **3. Macro-Level Drivers**

The international student mobility landscape is influenced by macro-level drivers, which may be situated at the country or institutional levels, and consist of many factors (e.g. economic conditions, institutional prestige, policies, and regional conditions) that create an hierarchy of attractiveness for countries as an education destination, which facilitate, as well as shape, the flow and quantity of students migrating across borders. The past few years have seen not just an exponential increase in the amount of students who are internationally mobile, but also a non-traditional diversification of sending and receiving countries, and increasingly complex mechanisms put in place to attract and retain international talent (OECD, 2025; Choi, 2021).

Economic strength has continuously been, and continues to be, a major (if not the most significant) determinant of a country's capacity to attract international students. Based on the most recent OECD statistics, almost two-thirds of all OECD international students are from high- or upper-middle-income countries, while the leading host country - Australia, Canada, France, Germany, the United Kingdom, and the United States-is still receiving almost two-thirds of all OECD international students (OECD, 2025). The correlation between GDP and appeal as a study destination is well known; countries with greater GDP and investment in tertiary education tend to promote not only more scholarships, but also offer better support for post-secondary institutions, better infrastructure and greater choices in academic opportunities (Choi, 2021).

Safety and quality of life also matter. Countries like Finland, Germany and Canada are frequently cited by students for their low crime rates, generous social welfare systems and overall quality of life (Nikou & Luukkonen, 2024). These factors are of particular relevance for

students from tumultuous political environments or regions experiencing social unrest, where personal safety and well-being may be their first considerations in the decision-making process.

Post-study work policies are a major factor for influencing student flows. The OECD (2025) noted that any nation that was able to offer students clear, systematic, easy pathways to work and live in the country following graduation- like Canada's Post-Graduation Work Permit (up to 3 years), the United Kingdom's Graduate Route (2 years), and Germany's 18-month job search visa- have been favored by international students (OECD, 2025). More complicated post study work arrangements, like the United States, reduced interest despite the academic status of the country. The promise of work experience and possibilities for 'settling' did entice international students and align with the goals of host countries to provide solutions to labour shortages and demographic challenges (OECD, 2025).

Mobility patterns are shaped by degree of cultural and linguistic proximity. Providing evidence of this recent research shows that students overwhelmingly look to go to parts of the world where a language barrier did not exist or where individuals have strong cultural connections. For instance, including English taught programs in non-Anglophone countries like the Netherlands and Germany has had dramatic success in increasing the international student share (OECD, 2025). Likewise, simply the language you speak and the existence of diaspora communities may have credible effects that reduce the perceived cost for adaptation and assimilation, making certain destination cities more attractive to students from developing nations (Schoe et al., 2021).

At the institutional level, reputation plays a definitive role in student choice. Universities with higher competitive power in the global higher education marketplace (think "top in a field" or "top in a country") not only likely bring in more students, but higher quality students and faculty

(EAIE, 2023). The OECD (2025) reports together the US and UK represent over half of the world's internationally mobile students, and this is largely because they are the top in the world rankings, an important connection to prestige (OECD, 2025). Quality of enrolled students, research, degree & specializations are all contributors to the attractiveness of the institution (EAIE, 2023).

Further, program specialization and uniqueness matters. Institutions that provide unique or leading edge programs - for example Germany's technical and engineering institutions, or Finland's Arctic studies - can establish unique niches that draw international students who are seeking out expertise that is not available back home (Restaino et al., 2020). Faculty involvement, recognized expertise and research opportunities allow institutions to further distinguish themselves from others: having a well-respected and published professor, having access to cutting edge labs/learning studios or just an opportunity to engage in undergraduate and graduate research with important profs were all attractive pull factors cited by potential students (EAIE, 2023; Restaino et al. 2020).

Overall, national policy frameworks are integral to the international movement of students. Brill (2020), describes the most prevalent policy levers for governments interested in internationalizing their higher education system are student mobility scholarship programs funded at the national level, streamlined visa systems, and research collaboration support (Brill, 2020). The OECD mentions the successful use of student mobility scholarship, language training, and post-study work support for attracting and retaining international students as part of South Korea's "Study Korea 300K Project" and Germany's "Blue Card" program (OECD, 2025). Conversely, countries and countries with clearly established and predictable visa policies have

seen declines in international enrollment during periods of restrictive visa policy such as seen in the US between 2018 and 2022 (OECD, 2025).

Investment in internationalization does not stop at student mobility. Countries and regions, in particular the European Union through Erasmus and Horizon 2020, have also invested heavily in research collaboration, higher education faculty mobility, and joint degrees, which further cultivate their higher education profile and attractiveness to global students (Brill, 2020). Student mobility programs not only support student mobility but also advance the establishment of international academic networks and enhance the capacity for cross-border research and innovation.

While internationalization increasingly favors diversification, regional disparities remain stark. The predominance of host countries is often limited to core countries in Europe and North America, with much of the periphery and semi-periphery acting more as senders than receivers of international students (Choi, 2021). The core-periphery pattern is exacerbated by both economic and network effects whereby countries that have partnerships, large diaspora communities, and geographic proximity to any major sending country, are likely to attract and retain more international students (Restaino et al., 2020; Schoe et al., 2021).

In particular, geographic proximity is most important for students from low- and middle-income countries that may be constrained by financial resources or visa barriers. Studies indicate that the intensity of student flows between countries remains more frequent between neighboring countries or regionally shared linguistic or cultural communities (Schoe et al., 2021). Significantly, the available Chinese and Indian communities in countries like Australia, Canada,

and the UK not only ease the migration experience when new international students begin but also represent a pull factor for student demand (OECD, 2025).

The importance of institutional networks and partnerships cannot be understated. The European Union's Erasmus+ program is a good example of this phenomenon: by facilitating credit transfer, joint degrees, and research collaboration, it has allowed institutions to develop significant networks that have continued to reinforce the attractiveness of European universities (Restaino et al., 2020). The denser these networks, the more they foster student mobility but also advance the overall resilience and quality of a higher education system.

| Factor                 | Core EU Countries<br>(Germany, France) | Peripheral EU Countries<br>(Latvia, Lithuania) | Asia-Pacific (China, South Korea) |
|------------------------|----------------------------------------|------------------------------------------------|-----------------------------------|
| Economic Investment    | High (top 10 OECD)                     | Low/Moderate                                   | Rising (esp. China, Korea)        |
| Top Host Status        | Yes                                    | No                                             | China: rising, Korea: strong      |
| Post-Study Work Path   | Clear, accessible<br>(Germany: 18 mo.) | Often limited                                  | Korea: improving, China: limited  |
| Institutional Prestige | High (QS/THE top 100)                  | Low/Moderate                                   | Growing (C9 League, SNU)          |

|                 |                                 |                                      |                             |
|-----------------|---------------------------------|--------------------------------------|-----------------------------|
| Policy Focus    | Internationalization, research  | Outbound mobility, capacity building | Attracting inbound students |
| Network Effects | Strong (Erasmus+, Horizon 2020) | Weak/Moderate                        | Regional partnerships       |

Table 1 (Key findings of macro-level drivers), Source: OECD (2025), Restaino et al. (2020), Choi (2021), OECD (2025), Choi (2021), OECD (2025), Brill (2020)

## 4. Micro-Level Drivers

The macro context of international student mobility is influenced primarily by macro-level drivers or contextual factors but the actual students' university choices are influenced by micro-level drivers bounded in personal aspirations, social networks and psychological factors. Personal aspirations and educational interests often play a distinctive role in the higher education evaluation. Over 60% of international students globally report career advancement and skill acquisition is a primary reasons for studying abroad; around 70% cite institutional reputation as important factors in decisions (OECD, 2023; Choi, 2021). Qualitative studies focused on the Western Balkans, including Albania, have pointed students highly value alumni recommendations and institutional rank, but especially valued direct evidence of academic quality and post-graduation success (Doda et al., 2024). These findings are also significant for self-determination theory, which values autonomy, mastery, and personal growth as sources of persistence and motivation for studying abroad (Deci & Ryan, 1985).

Personal growth and intercultural curiosity can together outweigh a purely financial motivation. As a helpful example, recent research finds that over 55% of Asian students in China considered a familiarity with the culture and learning the language as the primary motivations, while European students most strenuously pursue careers in areas like engineering and information technology (Frontiers in Psychology, 2022; Schoe et al., 2021). In the case of international students in Finland, they often tell us that after safety, and having reliable social welfare systems were considerations in deciding to study in Finland, representing the intersection of a personal and environmental priority (Nikou & Luukkonen, 2024). These dispositions further highlight the multilayered and complex nature of student decision making and show how the components of push and pull, interact with the personal goals of individuals.

Social networks are also important in determining mobility decisions and experiences. Peer recommendations and family support account for between 65% and 70% of destination choices for students, especially where a diaspora exists (Beech, 2015; Schoe et al., 2021). For instance, a number of students from Albania relied on account of past students (alumni) to navigate their available academic options and application processes, while students in the UK and Western European countries chose destinations where friends or family lived, suggesting the role of bonding capital and social support (Doda et al., 2024; Beech, 2015). Yet, social integration and loneliness continue to present issues: studies of Erasmus participants show that between 30% and 40% of international students felt socially isolated which adversely impacted students' academic performance, and reduced their intention to stay in the host country post-graduation (T.I.M.E. Association, 2021).

Cultural and linguistic context significantly mediates the experience of international students, with language proximity and cultural ties more essential for students from low- and

middle-income countries who are almost twice as likely to select geographically proximate destinations due to the reduced costs of adjustment (Schoe et al., 2021). In countries such as Germany and the Netherlands, where English taught programs were introduced, international enrolment increased by 25-30% even though 15-20 of students indicated they struggled with academic writing and integration (OECD, 2023). Also, Asian students indicated cultural proximity as a compelling motivator while studying in China, with 60% of students stating that familiarity with customs and language are important factors (Frontiers in Psychology, 2022).

Although financial reasons are important, they were often ranked below students' personal and academic goals. The OECD (2023) indicates that 60% of international students prioritize long-term career opportunities over short term financial incentives, even in countries where postgraduate briefly offer large scholarships. For example, public funded scholarships to students studying in Türkiye from war primary affected countries attract a significant number of recipients, but recipients tend to present motives for cultural exposure and insights of quality academic pathways than those pre-determined based on only financial support (Doda et al., 2024; OECD, 2023).

Psychological resilience and self-determination were strongly linked students' academic success and satisfaction. Evidence confirmed that students autonomously motivated who selected a study destination for personal reasons report satisfaction 20-25% higher than students whose motives for study destinations were framed externally, such as family or financial incentives (British Council, 2024). Also, students who were self-regulated and had perseverance indicated a 15% higher academic performance, while a research in Germany cited that students enrolled in mentorship retention programs had a 10-12% retention increase through social integration and belonging (Restaino et al., 2020; ERIC, 2023).

Despite these insights, traditional models often ignore the numeric considerations of demographic diversity and evolving motivations. For example, mature students (25+) are influenced primarily by career progression two times more likely than students enrolled in undergraduate programs. Additionally, women students have unique circumstances with balancing the wishes of their family commitments with their academic aspirations, of which traditional push-pull models do not fully address (British Council, 2024; van Mol & Nada, 2023). Further, traditional push-pull models are often static and do not account for motivations changing over time, such as a student influenced to study abroad due to visa benefits, who during their longer study pathway potentially becomes increasingly motivated for an opportunity to inspire future research or professional growth and hence intrinsic motivation (OECD, 2023).

Overall, micro-drivers of individual circumstances of intrinsic motivation, social networks, cultural adaptability, financial circumstances, and psychological resilience have different contextual meaning for students when analyzing the Complex Lenses of International Student Mobility. Applying theoretical frameworks of contextualised self-determination theory and theory and supported with empirical evidence, the importance of institutions considering a contextualised unique support and provision of higher education approaches to learning from a holistic perspective can influence the global competitiveness discourse regarding who may choose to study at institutions such as LUISS, the topic of the next chapter.

## **5. Challenges and Barriers**

While international student mobility presents numerous opportunities, students often face significant barriers that can reduce access, integration, and overall successful outcomes. Financial barriers remain one of the strongest hurdles with students. An OECD report found that

in high-income countries, such as Australia, Austria, Canada, New Zealand, and the United States, international students pay approximately twice as much or more in tuition and fees compared to national students (OECD, 2024; OECD, 2017). Living costs are also considerable in many high-income countries, with out-of-pocket income making up a considerable amount of a student's budget, often rivaling or exceeding tuition and fees, with cities being particularly expensive (OECD, 2023). Even in countries where tuition and fees are low, international students that do not receive a full scholarship have to rely on family savings or loans, or must work part-time to cover their educational expenses (OECD, 2023). For example, ERASMUS students cite financial pressure as their main challenge, and many rely on family support, or incur students debt to pay for their studies (Souto-Otero et al., 2023). These financial challenges are consistent with the findings from Human Capital Theory, which considers education to be an investment, while also revealing the disparities faced by students from lower-income backgrounds or those that lack sufficient family support.

Social and cultural barriers are another significant concern for international students. Language barriers are a primary example; research often shows that a large proportion of international students report language anxiety which can lead them to socially withdraw and will affect their studies (Lee et al., 2024). In addition, students also report issues related to discrimination and homesickness. For example, up to 30% of international students studying in Europe report being discriminated against or experiencing bias related to housing and social situations, and there is a notable share of ERASMUS students that reported returning home early because of social isolation (Souto-Otero et al., 2023). These results are in line with Acculturation Theory which outlines the stresses associated with adopting to new cultural norms, particularly in the absence of social supports.

Administrative barriers and bureaucratic challenges present one of the most significant barriers to access in many instances. For example, students often experience barriers related to visa policies and institutional bureaucratic processes to facilitate student mobility. In the United States, international students from several African countries have noted visa denials rates exceeding 60%, and appointment wait times reaching 300 days in some instances (The PIE News, 2019). In Europe, delays in credit recognition are a common bureaucratic problem, and surveys show that a substantially large share of ERASMUS students report graduation delays due to bureaucratic obstacles (Ramon-Muñoz & Gea-Orriols, 2021). Factors that deter study abroad students include complex application requirements, faculty and institutional mismatched calendars, and in Asia-Pacific, bureaucratic barriers remain a widely cited reason for students abandoning study abroad plans (OECD, 2023)<sup>2</sup>. Administrative obstacles are reported similar to Institutional Theory, which critiques bureaucracy and inefficiencies and inconsistencies related to policies that hinder aspects of internationalization.

These combined patterns are broadly captured in Table 1, which highlights the prevalence of financial, social, and administrative barriers within major world regions.

| Barrier Type | Europe (ERASMUS)                          | U.S./Canada                                                          | Asia-Pacific                                                   |
|--------------|-------------------------------------------|----------------------------------------------------------------------|----------------------------------------------------------------|
| Financial    | High tuition; family support/loans common | International students pay double or more tuition; high living costs | Most students self- or family-funded; scholarships less common |

|                 |                                                 |                                         |                                                     |
|-----------------|-------------------------------------------------|-----------------------------------------|-----------------------------------------------------|
| Social/Cultural | Up to 30% report discrimination or isolation    | Isolation and integration issues common | Language anxiety and cultural adaptation challenges |
| Administrative  | Credit recognition delays; bureaucratic hurdles | Visa denial rates >60% for some Africa  | Bureaucratic hurdles a leading barrier              |

Table 2 (Key findings of challenges and barriers), Sources:Souto-Otero et al. (2023); OECD (2023); The PIE News (2019); OECD (2024); OECD (2023); Lee et al. (2024)

## 6. Summary of Main Findings

This literature review highlighted that international student mobility (ISM) is shaped by macro-level and micro-level factors, with both contributing to the continuation of inequalities, changing motivations, and barriers in global student flows. Macro-level indicators are mainly dominated by economic and policy frameworks. A reported two-thirds of international students in OECD (2023) countries come from high-income or upper-middle-income country origins documenting their continuing financial and other structural barriers for those from a lower-income background. Most international students undertaking studies within Europe have their main destination countries which include Germany, France, USA, and UK - socially coined “core” countries - where primarily due to higher GDP spending, policies for post-study work options, and general institutional prestige these remain popular higher education destination choices (OECD, 2023; Restaino et al., w2020). National policies such as the 'Blue Card -

Germany' and 'Study Korea 300K Project - South Korea', promote core countries in conjunction with scholarship supports, language training, and administrative processing associated with student visa processes (OECD, 2024; Ramon-Muñoz & Gea-Orriols, 2021).

Other macro-level factors, such as reputation, and quality of programs can also be drivers to study mobility. U.S. and UK institutions collectively are home to over half of world's globally mobile students, not surprising as their often-convenient dominant status in rankings and relative prestige was a driver of decisions to study internationally (OECD, 2023). When English taught programs were introduced in non-Anglophone countries (e.g., Netherlands, Slovenia) international enrollment increased by 25-30%, though 15-20% faced academic issues, for which language was a primary reason cited (OECD, 2023). Policy frameworks such as scholarships and smooth visa procedures are key in shaping flows of international students. Conversely, restrictive or uncertain visa policies have caused some countries to experience declines in enrollments (OECD, 2023; Ramon-Muñoz & Gea-Orriols, 2021).

At the individual and social levels, personal and social factors are fundamentally important to university choice. More than 60% of respondents reported that institution brand (reputation) and career advancement were key reasons for their choice, while also 55% of Asian students studying in China (20% of total sample) reported that cultural familiarity was a key reason for their choice of program (Frontiers in Psychology, 2022; OECD, 2023). Additionally, social networks including alumni and family input also influenced 65-70% of destination choices. However, many (30-40%) Erasmus students reported feeling isolated, which may lead to decreased intention to stay in their host country (Souto-Otero et al., 2023; T.I.M.E. Association, 2021). Furthermore, students who were autonomous (students who chose their destination willingly based on achieving personal goals) reported 20-25% higher satisfaction in study abroad

than students who were induced/pressured through some external push to go to study abroad, such as family values or financial incentives (British Council, 2024). Especially, language and cultural proximity are essential factors for students from low- and middle-income countries who are more than twice likely to choose a destination that is closer in geographic distance due to lower adaptation costs (Schoe et al., 2021). In non-Anglophone countries, such as Germany and the Netherlands, the availability of programs that taught in English has increased international student enrollment by 25% to 30%, with 15% to 20% of students reporting difficulties with academic writing and integration (OECD, 2023).

Despite these opportunities, there are still large barriers facing international students. Financial issues tend to be the most widespread: in top-host countries such as the United States and Australia, international student tuition fees are often double or more, with living costs often taking a significant part of the budget (OECD, 2023). With only a small percentage of students receiving full scholarship, this necessitates family savings and/or debt (OECD, 2023; Souto-Otero et al., 2023). Social and cultural issues remain equally acute with over 50% of surveyed students experiencing language anxiety, while 30% of students in Europe reported discrimination, particularly in housing and social spaces (Lee et al., 2024; Souto-Otero et al., 2023). In terms of administrative issues, visa denial rates are over 60% for African students in the USA, while credit recognition delays in Europe delayed students graduation by as much as a year (The PIE News, 2019; Ramon-Muñoz & Gea-Orriols, 2021).

Although the current literature provides a robust understanding about the drivers and barriers of ISM, there remain significant gaps. There is a consistent underrepresentation of marginalized groups, as in most OECD countries less than 5% of internationally mobile students come from low-income countries, and research rarely addresses their intersectional identities in terms of

LGBTQ+ or disabled students (OECD, 2023; CGHE, 2021). Longitudinal, or even post-graduation findings, are limited and therefore neglect to examine its long term impact to career, and life trajectories (Luo et al., 2023). Furthermore, research continues to be overtly focused on westernized host countries and less on mobility in Africa, South Asia or South America (OECD, 2023). Finally, the evaluation of whether policy initiatives-work such as scholarships, adapted orientations, and visa changes-have any impact, thus providing limited evidence in terms of challenging and initiating change (King et al., 2020).

In response to the above challenges, stakeholders will need to employ a multifaceted approach. Universities should either promote and expand upon need-based scholarships, evaluate how to provide as many students as possible with language support, and create and implement mentorship programs to include their international students so they can successfully integrate (British Council, 2024). Policymakers should also simplify the immigration processes and work with educational institutions with regard to credit recognition, while utilizing targeted scholarship programs to encourage diversity for international students (OECD, 2024). Researchers need to document and promote high-quality longitudinal and comparative research that incorporates voices from marginal states and identities (Luo et al., 2023; CGHE, 2021). Sustainable international student mobility can only be achieved both equitably and supportively together.

## **7. Conclusion**

This literature review has shown that student mobility across borders is characterized by an ever-evolving combination of macro and micro factors, both of which contribute to the

opportunities and challenges experienced by students and institutions around the world. On the macro level, the review has shown that national economic strength, institutional prestige, policy environments, and investments in higher education are each critical aspects to consider. Core countries, such as Germany, France, the US, and the UK still capture the most international students, evidenced by high GDP investments in higher education, favorable policies related to post-study work, and institutional prestige. In addition to English-taught programs, national policies have also made programs in leading host countries more attractive by reducing linguistic barriers, making the administrative processes less daunting, as well as clarifying pathways to employment and residency. National policies, such as Germany's "Blue Card" and South Korea's "Study Korea 300K Project," have increased potential domestic employment opportunities for international students and provide more clarity. However, student flows and placement remain uneven, as peripheral nations and low-income countries face structural challenges to both attract and retain international students.

On the micro level, the review also highlighted the significance of individual motivations, networks, and psychological factors in terms of choosing universities and student experiences. Ambition for future careers, academic prestige, as well as personal growth, were the most cited justifications, while the role of social networks - such as family recommendations and alumni experiences - seemed to play a definitive role in determining which country to choose. The literature also indicates that students who are motivated in an autonomous manner experience greater satisfaction and academic achievement than those driven by external factors. However, there are still significant barriers, including finance, language-related anxiety, discrimination, social isolation, and administrative barriers (e.g., visa refusals or delays and delays in credit recognition, etc.) that hinder equitable access and successful outcomes for many students. These

barriers are especially significant for students from low-income backgrounds and under-represented categories that experience ongoing marginalization in patterns of global mobility.

Even though the research on international student mobility is extensive, important gaps remain. Notably, students from low-income countries, and students with intersecting identities, such as LGBTQ+ students and/or students with disabilities, have not actively participated in the literature. There are also limited studies that follow the experience of international students in a longitudinal manner or post-graduation, which limits our understanding of the long-term consequences of mobility on career pathways and life trajectories. While there are examples of policy interventions and support programs being systematically evaluated, there is also a need for greater exploration of mobility patterns in non-Western contexts.

Ultimately, the findings presented here indicate that overcoming the persistent inequalities and barriers in international student mobility requires a coordinated response at multiple levels. Universities should implement more need-based scholarships and enhance language and integration support, and include new forms of mentorship to enhance retention. Policymakers are encouraged to simplify visa and credit recognition processes, and to develop targeted scholarships that diversify mobility and support inclusion. Researchers should undertake more longitudinal and comparative studies, with a focus on amplifying the voices of students from under-represented groups and regions. In this way, international student mobility may truly become a transformative force for individuals and societies.

In the next chapter, we will expand on the findings in this chapter with a data-driven analysis of the LUISS case, focusing on the main drivers and barriers, as well as the specific institutional

contexts that shape international student mobility at LUISS. This will illustrate how the empirical themes in the previous literature manifest in concrete terms in the LUISS context, which will provide clear and tailored recommendations for future policy and practice.

## **Chapter 2 - Research Methodology**

### **1. Introduction**

This chapter presents the methodological approach used to identify the key drivers influencing international students' choice of university. The research adopts a quantitative design, leveraging a comprehensive dataset that integrates both internal and external sources. Internal data, obtained from university records, includes variables such as applications, admissions, enrollments, and engagement initiatives targeted at prospective students. External data, sourced from publicly available databases, encompasses country-level indicators like GDP per capita, population demographics, export values, and geographic distance. By combining these two data types, the analysis aims to capture both institutional factors and broader contextual influences on student decision-making. The following sections detail the research design, data collection and integration processes, analytical methods, ethical considerations, and study limitations

### **2. Data Description and Data Collection**

#### **2.1 Overview of the Dataset**

The dataset employed in this study provides a comprehensive, country-level overview of both university-specific and contextual factors relevant to international student recruitment and engagement. The dataset covers more than 150 countries, with each row representing a unique country case. The data primarily reflects the most recent available year, focusing on the 2023 application cycle, but also incorporates multi-year economic indicators (e.g., GDP growth 2018–2022). The main aim of this data collection is to analyze the interplay between institutional outreach activities and external country characteristics in shaping international student flows and engagement outcomes.

## **2.2 Data Sources**

The dataset integrates two main types of data: internal (university-level) and external (country-level/contextual) variables.

### **2.2.1 Internal Data**

Internal data are sourced from the university's own administrative records and outreach tracking systems. These variables reflect the institution's direct activities and outcomes in each country, such as the number of applications received, admissions granted, enrollments, and the deployment of various engagement initiatives (e.g., webinars, school tours, agent involvement). This information provides a granular view of the university's operational footprint and recruitment strategies across different national contexts.

### **2.2.2 External Data**

External data are compiled from authoritative international databases and public sources, including the World Bank, United Nations, and national statistical agencies. These variables describe the broader socioeconomic, demographic, and geographic context of each country, such as GDP per capita, population

structure, distance from Italy, and export values. These contextual factors are crucial for understanding the environment in which university activities take place and for controlling for country-level heterogeneity in subsequent analyses.

## 2.3 Key Variables

The dataset's variables are organized into two main groups, reflecting their source and function:

### 2.3.1 Internal Variables (Columns A–AB)

- ***Country and Region Identifiers***
  - `paese_di_studio`, `codiciisoalpha3`, `area` – Country name, ISO code, and geographic region.
- ***Recruitment Pipeline Metrics***
  - `applicationeleggibili` (eligible applications), `ammessi` (admitted), `iscritti` (enrolled).
- ***Engagement and Outreach Activities***
  - `agenti` (agents), `aicscucs` (cooperation consortia), `AlumnieChapterLeader`, `ambasciate` (embassies), `ambassador`, `Associazione studentescheConsult` (student associations), `confindustria`, `diaspore`, `dubaihub`, `erasmusedd` (Erasmus/Double Degree), `eufasaacdmae`, `foundation`, `icp`, `iila`, `pianocina`, `portali`, `tourscuole`, `tp`, `webinarorientamento`, `samplelecture`, `engagementscuoledaremoto`, `sportegaming`.
  - These are mostly binary indicators (0/1) or counts, capturing the presence or intensity of specific university activities in each country.

### 2.3.2 External Variables (Columns AC–AU)

- ***Economic Indicators***
  - `gdp2018_22` (GDP growth 2018–2022), `gdp_pc` (GDP per capita).
- ***Demographic Structure***
  - `pop_ages15_64` (population aged 15–64), `pop_ages0_14` (population aged 0–14), `University_perc` (university enrollment or attainment percentage), `pop_total_23` (total population 2023).
- ***Geographic and Political Context***
  - `country` (English country name), `dist` (distance from Italy), `distcap` (distance from Rome), `EU` (EU membership, binary), `area` (region).
- ***Diaspora and Community Links***
  - `countries_of_communities`, `italian_communities_abroad`.
- ***Trade and Export Data***
  - `Export_countries` (main export partners), `Export_value`, `Export_percentage`, `Total_value`.

These variables provide a multidimensional view of each country, allowing for nuanced analysis of how internal strategies interact with external conditions.

## 2.4 Data Structure and Format

The dataset is structured as an Excel spreadsheet, with each row corresponding to a country and each column representing a variable. The first row contains unique variable names, and all subsequent rows are observations for each country. Variables are consistently formatted, with categorical variables (e.g., `area`, `EU`) and binary indicators (e.g., presence of engagement

activities) clearly distinguished from continuous variables (e.g., GDP, population). Unique identifiers such as `codiciisoalpha3` ensure that each country is uniquely and consistently referenced, facilitating data linkage and integrity.

## **2.5 Data Cleaning and Preparation**

Prior to analysis, several steps were undertaken to ensure data quality and consistency. First, all variable names were standardized for clarity and uniformity, and duplicate or ambiguous entries were resolved. Missing values were systematically identified; for key quantitative variables, imputation or exclusion was considered based on the extent and pattern of missingness. Country names and codes were harmonized across internal and external sources to enable accurate merging. Where necessary, categorical variables were recoded for analytical consistency (e.g., binary indicators for EU membership and engagement activities). Outliers and improbable values in economic or demographic indicators were cross-checked against original sources and corrected or flagged as appropriate. Finally, the dataset was filtered to include only countries with sufficient data coverage for both internal and external variables, ensuring the validity of subsequent comparative analyses

## **3. Data Analysis**

The data analysis for this study was conducted using Python, leveraging a suite of libraries and statistical techniques aligned with the research objectives and the structure of the dataset. This approach ensured a transparent, reproducible workflow and enabled a nuanced exploration of the relationships between university-level and country-level variables.

### **3.1 Data Processing**

The initial stage involved rigorous data cleaning and preparation. The dataset was loaded using Pandas, which facilitated efficient handling of tabular data. Column names were standardized for clarity, and duplicate columns were removed to ensure consistency. Special attention was given to identifying and cleaning key variables, such as the correct column for Italian communities abroad, to ensure all features used in the analysis were accurately defined.

All relevant numerical variables, including financial and demographic indicators, were converted to appropriate numeric types after removing non-numeric characters. Missing values in both features and the target variable were handled systematically: for features, missing entries were imputed using the median value, while rows with missing or non-numeric target values were excluded from analysis. Additionally, the target variable (“iscritti”) was log-transformed ( $\log(1+x)$ ) to address skewness and stabilize variance, improving model interpretability and performance.

### **3.2 Feature Selection and Engineering**

A set of relevant features was selected for modeling, encompassing both internal (e.g., number of embassies, presence of Italian communities abroad) and external (e.g., GDP per capita, university participation rate, geographic distance, export value) variables. To capture more complex relationships, two interaction features were engineered: the product of GDP per capita and export value (“gdp\_export”), and the ratio of the number of embassies to geographic distance (“embassy\_per\_distance”). The target variable was the log-transformed number of enrolled students.

The data was split into training and test sets using an 80/20 ratio to enable robust model evaluation. For models sensitive to feature scaling (e.g., Linear Regression, Random Forest), all input variables were standardized using the StandardScaler from scikit-learn. For tree-based

models like XGBoost, scaling was not applied, as these models are invariant to monotonic transformations of the input features.

### 3.3 Statistical Modeling and Analytical Techniques

Three regression approaches were employed to model the relationship between the selected features and the target variable:

- Linear Regression: Used as a baseline model, providing interpretable coefficients and serving as a reference point for feature importance.
- Random Forest Regression: Employed to capture non-linear relationships and provide an ensemble-based benchmark.
- XGBoost Regression: Utilized as the primary advanced model, with hyperparameter tuning via randomized search and cross-validation to optimize performance.

For Linear Regression and Random Forest, permutation importance was computed using scikit-learn's `permutation_importance` function, which measures the impact of each feature on model performance by randomly shuffling its values and observing the change in prediction accuracy.

For XGBoost, both the model's built-in feature importance metrics and SHAP (SHapley Additive exPlanations) values were used to interpret feature relevance. SHAP analysis provided a model-agnostic, nuanced understanding of feature contributions, including both summary and dependence plots for the most influential variables.

### 3.4 Visualization

To communicate the results of the feature importance analyses, Seaborn and Matplotlib were used to generate clear, publication-quality visualizations. Feature importances from permutation importance and XGBoost were compiled into DataFrames and visualized using Seaborn's barplot function, displaying the relative contribution of each variable to the predictive models. For XGBoost, SHAP summary and dependence plots were generated to further elucidate the impact and interaction of key features on the model's predictions.

### **3.5 Software and Analytical Tools**

The entire analytical workflow was implemented in Python, utilizing:

- Pandas for data manipulation and cleaning,
- NumPy for numerical operations,
- scikit-learn for preprocessing, regression modeling, permutation importance, and train-test splitting,
- XGBoost for advanced regression modeling,
- SHAP for model interpretability,
- Matplotlib and Seaborn for data visualization.

This combination of tools is widely recognized in academic and applied research for its flexibility, transparency, and reproducibility

## **4. Justification of Choices**

The methodological framework was designed to maximize validity and reliability while aligning with the study's focus on institutional and contextual determinants of student recruitment. Python

and its scientific libraries (Pandas, NumPy, scikit-learn, XGBoost) were chosen for their reproducibility, scalability, and alignment with modern data science standards. Linear Regression provided interpretable coefficients to identify baseline relationships, while XGBoost captured non-linear dynamics and interactions, ensuring robustness against oversimplification.

Feature standardization (via scikit-learn's **StandardScaler**) controlled for variable scale biases, and permutation importance quantified feature relevance without assuming linearity, enhancing methodological fairness. Visualization (Seaborn/Matplotlib) enabled intuitive validation of results, critical for peer review and policy discussions.

By integrating transparent data cleaning, dual modeling approaches, and model-agnostic validation, this workflow ensures that findings reflect true underlying patterns, not artifacts of methodological limitations. The result is a rigorous, replicable analysis tailored to disentangling the drivers of international student enrollments.

## **Chapter 3 - Key findings of the LUISS case: Results**

### **1. Introduction to the Results**

This chapter presents the findings of a predictive analysis examining how socioeconomic, institutional, and geographic factors influence international student enrollment (*iscritti*) at LUISS University. Three regression models—Linear Regression, Random Forest, and XGBoost—were applied to a dataset spanning 116 countries to assess predictive accuracy and identify key determinants of enrollment. The analysis prioritizes six core variables: GDP per capita,

university participation rates, geographic distance from Italy, the presence of Italian communities abroad, export value, and embassy count. Methodologically, the study incorporates advanced techniques such as feature engineering (e.g., interactions between GDP and export values) and log-transformation of the enrollment variable to address data skewness, alongside SHAP (SHapley Additive exPlanations) for model interpretability.

The results are organized into three interconnected analyses. First, model performance is compared using standard metrics—mean absolute error (MAE), root mean squared error (RMSE), and the coefficient of determination ( $R^2$ )—to evaluate predictive accuracy. Second, feature importance is analyzed through permutation importance for Linear Regression and Random Forest, supplemented by XGBoost’s intrinsic importance metrics. Third, SHAP analysis quantifies the directional impact of critical predictors, offering insights into how specific variables shape enrollment outcomes. Together, these approaches provide a dual perspective: quantitative validation of model efficacy and qualitative clarity on enrollment drivers, ensuring findings are both statistically robust and institutionally actionable.

## **2. Model Performance Comparison**

### **2.1 Overview of Evaluated Models**

To identify the most influential country-level factors associated with the number of LUISS students from each country, three regression models were employed:

- **Linear Regression:** Serves as a baseline, capturing linear relationships between predictors and the target variable.

- Random Forest Regression: An ensemble of decision trees that captures non-linear effects and complex interactions among features.
- XGBoost Regression: An advanced gradient boosting method, optimized for performance and interpretability, particularly effective in handling diverse feature sets and uncovering nuanced relationships.

## 2.2 Quantitative Performance Metrics

While the primary goal is to interpret the drivers of student numbers, model performance is also reported to ensure reliability of the findings. The models were evaluated using mean absolute error (MAE), root mean squared error (RMSE), and the coefficient of determination ( $R^2$ ) on the log-transformed number of students (*iscritti\_log*) to account for skewness in the data.

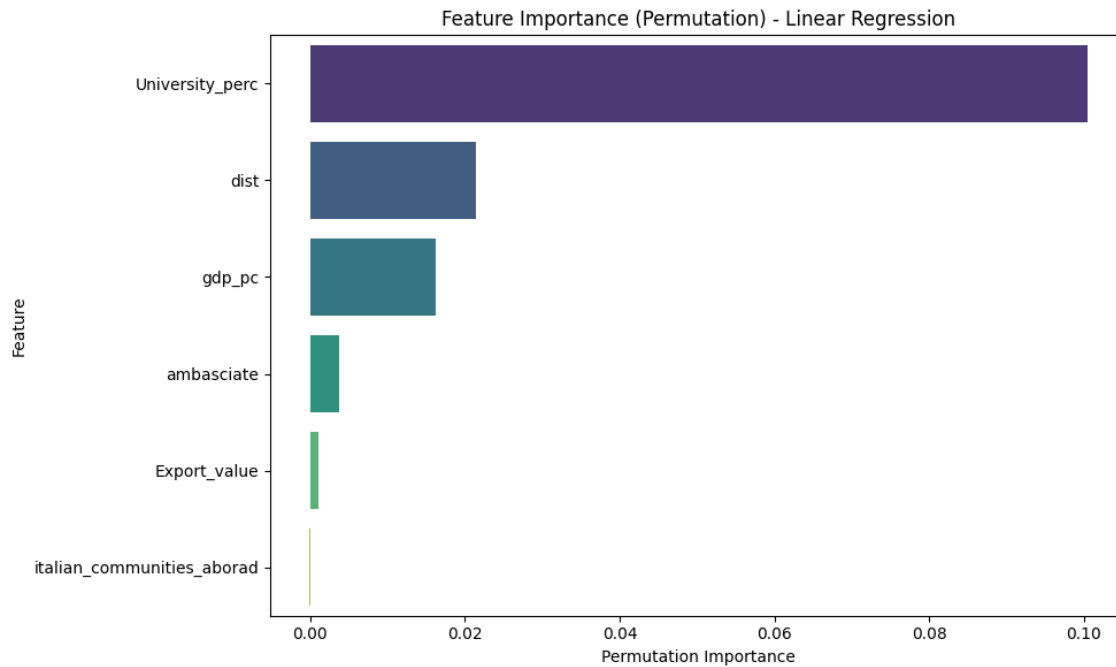
| Model             | MAE   | RMSE  | $R^2$ |
|-------------------|-------|-------|-------|
| Linear Regression | 6.80  | 16.80 | 0.06  |
| Random Forest     | 5.34  | 12.68 | 0.47  |
| XGBOOST           | 0.395 | 0.593 | 0.514 |

Table 3 (Models performance results)

XGBoost demonstrated the best performance, achieving the highest explanatory power ( $R^2 = 0.514$ ) and the lowest errors (MAE = 0.395, RMSE = 0.593), indicating that its results provide the most robust basis for interpreting the main drivers of LUISS student numbers. This superiority stems from its ability to model non-linear relationships and handle feature interactions (e.g., GDP-export synergies, embassy-distance ratios). The subsequent feature importance and SHAP analyses are therefore primarily based on the XGBoost model's results to ensure reliable identification of influential factors.

### 3. Feature Importance Analysis

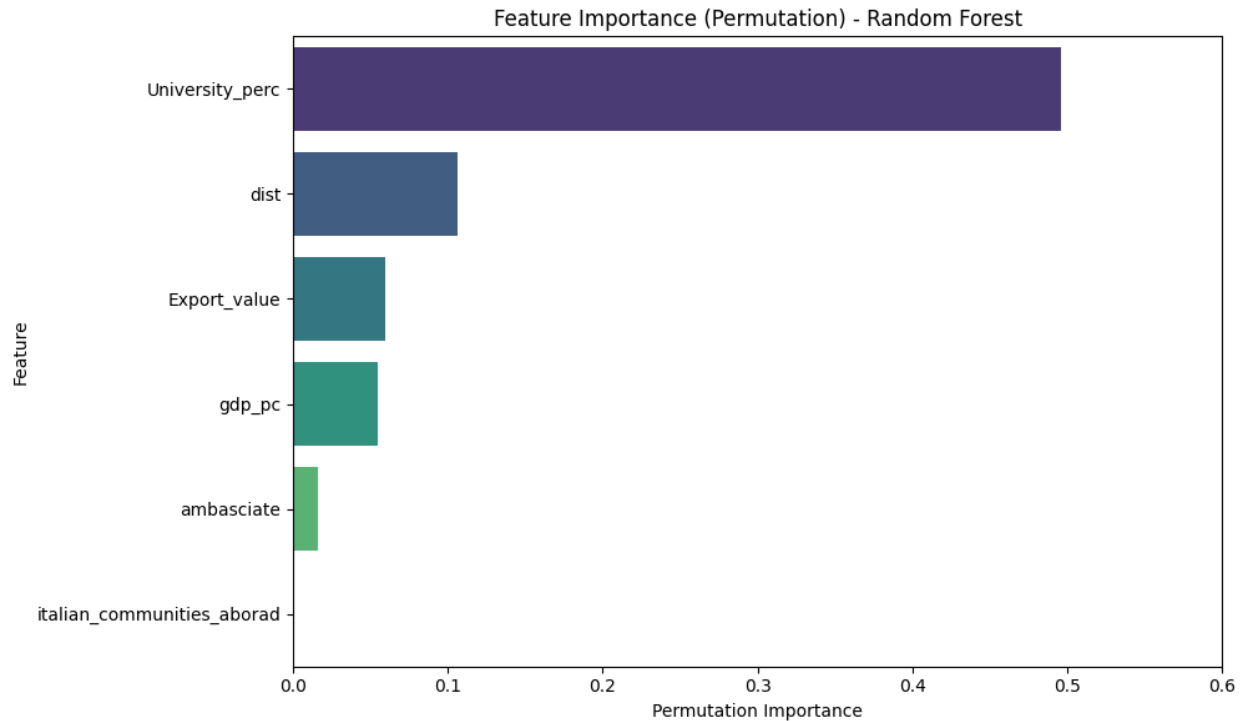
#### 3.1 Linear Regression Feature Importance



*Figure 1: Linear Regression*

In Figure 1, we can see that "University\_perc" is by far the most important feature for predicting enrollments, as indicated by its much longer bar compared to the others. This strong importance is likely because the percentage of university-aged people is directly related to the number of enrollments in your data. On the other hand, "italian\_communities\_abroad" is the least important feature, with an almost invisible bar, which suggests it has very little impact on the model. The other features, such as "dist," "gdp\_pc," "ambasciate," and "Export\_value," have moderate to low importance, reflecting that they play a smaller role in the model's predictions compared to "University\_perc".

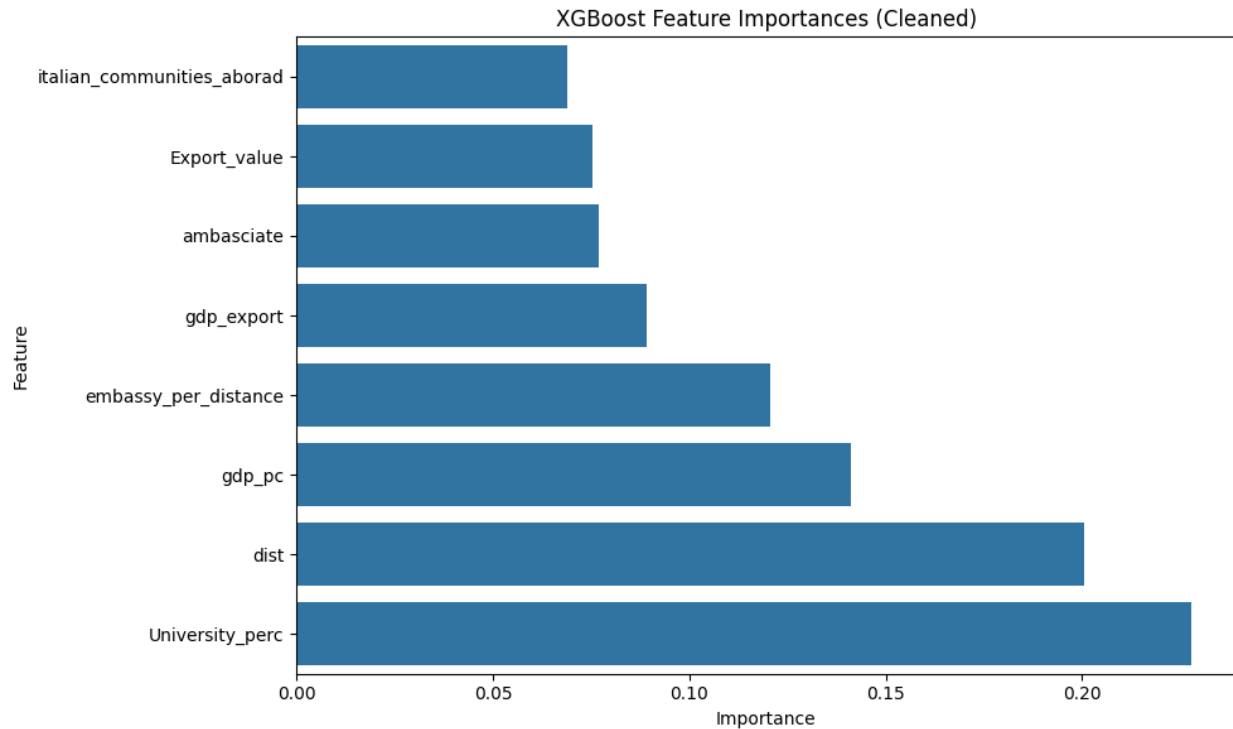
### 3.2 Random Forest Feature Importance



*Figure 2: Random Forest*

In Figure 2 showing permutation importance from a Random Forest model, "University\_perc" is clearly the most important feature, with a much longer bar than the others, meaning it has the strongest impact on predicting enrollments. This likely happens because the percentage of university-aged people in each country is directly related to the number of students enrolling. On the other end, "italian\_communities\_abroad" is the least important, with its bar almost invisible, indicating it has almost no effect on the model's predictions. The other features, like "dist," "Export\_value," "gdp\_pc," and "ambasciate," have moderate to low importance, showing they contribute less to the prediction compared to "University\_perc".

### 3.3 XGBoost Feature Importance

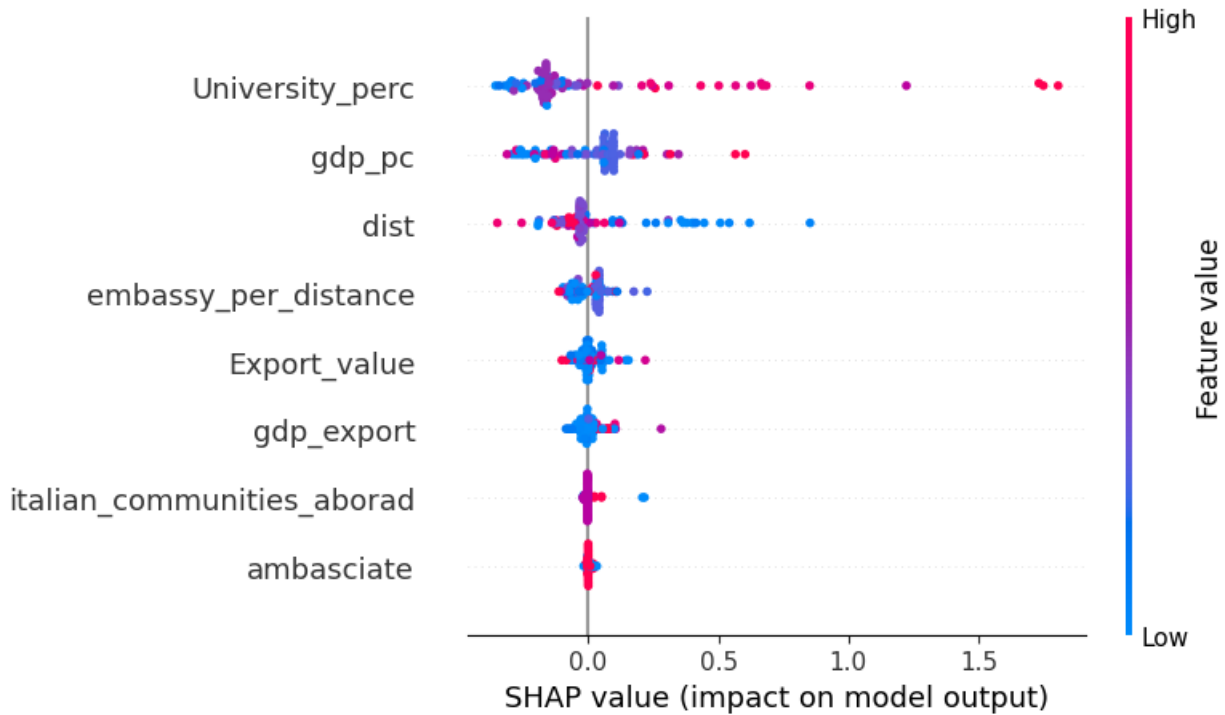


*Figure 3: XGBoost*

In Figure 3, we see that "University\_perc" is the most important feature, with the longest bar, meaning it contributes the most to predicting enrollments. This makes sense because the percentage of university-aged people is directly related to the number of potential students. "dist" (distance) is also quite important, likely because countries farther away may have fewer enrollments due to travel or accessibility. On the other hand, "italian\_communities\_abroad" is the least important feature, with the shortest bar, indicating it has very little impact on the model. The other features like "gdp\_pc," "ambasciate," "Export\_value," "gdp\_export," and "embassy\_per\_distance" have moderate to low importance, showing they play a smaller role in predicting enrollments compared to "University\_perc" and "dist".

## 4. SHAP Analysis for Model Interpretability

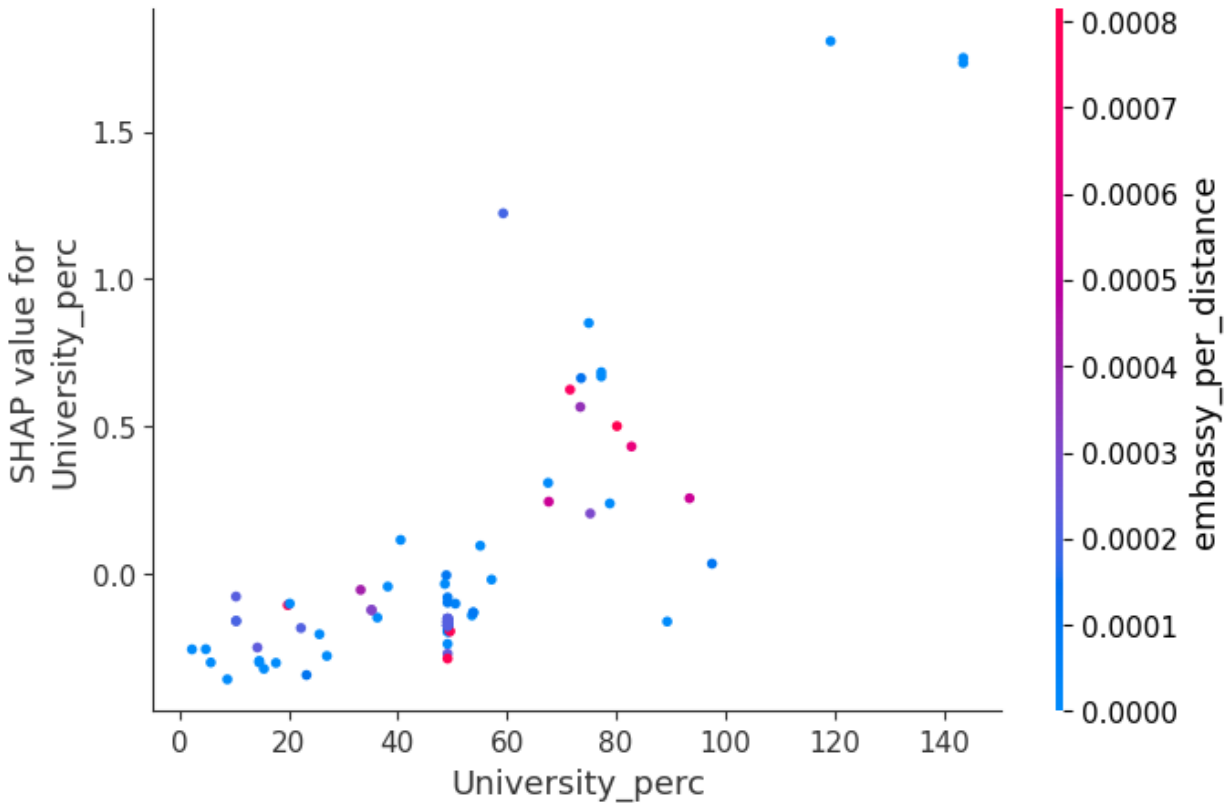
### 4.1. SHAP Summary Plot



*Figure 4: Shap Summary Plot*

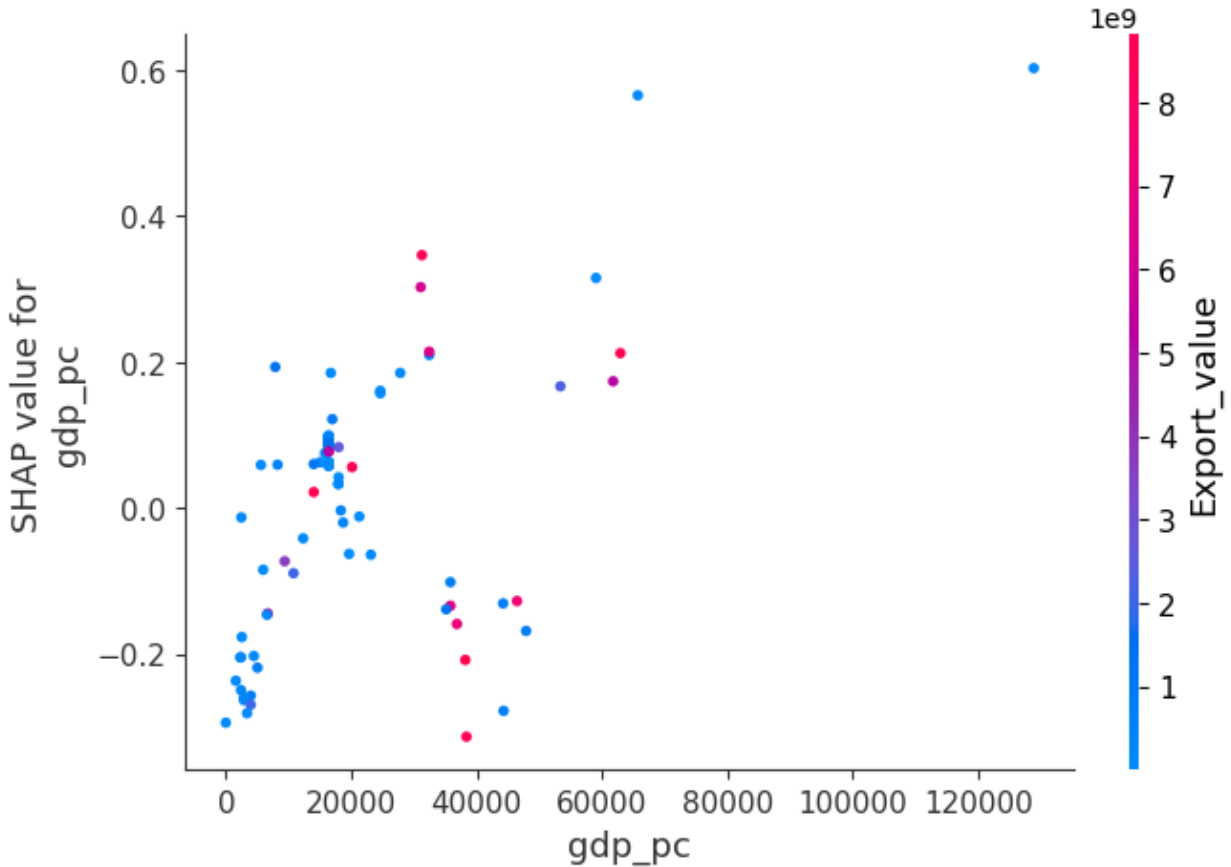
In Figure 4, we can see that "University\_perc" is the feature with the highest impact on the model's predictions, as shown by the widest spread of SHAP values along the x-axis. This means that the percentage of university-aged people in a country is the most influential factor for predicting enrollments, likely because it directly relates to the potential number of students. Other features like "gdp\_pc" and "dist" also have a noticeable impact, but less than "University\_perc." On the other hand, "ambasciate" and "italian\_communities\_abroad" are the least important, with their SHAP values clustered closely around zero, indicating they contribute very little to the model's output. The color gradient shows how high or low values of each feature affect the prediction, but overall, the main takeaway is that "University\_perc" drives the model the most, while "ambasciate" and "italian\_communities\_abroad" have minimal effect.

## 4.2. SHAP Dependence Plots



*Figure 5: Shap Dependence Plot 1*

In Figure 5, we can see how "University\_perc" influences the model's predictions for enrollments. Each point represents a country, with the x-axis showing its "University\_perc" value and the y-axis showing the SHAP value, which indicates how much "University\_perc" pushes the prediction up or down. The trend is clear: as "University\_perc" increases, the SHAP value also increases, meaning that higher percentages of university-aged people lead to higher predicted enrollments. The color of each point reflects "embassy\_per\_distance"—countries with higher "embassy\_per\_distance" are shown in pink, but the overall pattern is dominated by "University\_perc". This plot highlights that "University\_perc" is the most important driver for the model, and its effect is strongly positive and consistent across the data.



*Figure 6: Shap Dependence Plot 2*

In Figure 6, we see how "gdp\_pc" (GDP per capita) affects the model's predictions for enrollments. Each point represents a country, with the x-axis showing its GDP per capita and the y-axis showing the SHAP value, which indicates how much that country's GDP per capita pushes the prediction up or down. The colors represent "Export\_value," with blue for lower export values and pink for higher ones. The plot shows that as GDP per capita increases, the SHAP values generally become more positive, especially for some countries with higher GDP per capita, meaning that higher GDP per capita tends to increase the predicted enrollments. However, there is quite a bit of spread, indicating that the effect of GDP per capita can vary depending on the country and its export value. Countries with higher export (from Italy) values (pink points) can have higher SHAP values at similar GDP per capita levels, suggesting a possible interaction between economic strength and exports in influencing enrollments.

## 5. Synthesis and Key Findings

## 5.1 Summary of Main Findings

The analysis set out to answer: *What are the key drivers that influence international student choice of university?* Using a dataset of 116 countries and multiple regression models (Linear Regression, Random Forest, and XGBoost), the results consistently highlight a clear hierarchy of influential factors shaping international student enrollment at LUISS University.

## 5.2 Model Performance

XGBoost outperformed the other models, achieving the highest explanatory power ( $R^2 = 0.514$ ) and lowest error rates (MAE = 0.395, RMSE = 0.593), indicating robust predictive accuracy and reliability for interpreting feature importance and SHAP results.

## 5.3 Feature Importance

Across all models, *University\_perc* (the percentage of university-aged population) emerged as the most important predictor of international student enrollments. This was evident in both permutation importance charts and XGBoost's intrinsic feature importance, as well as SHAP analysis. The strong effect of this variable suggests that demographic potential—having a large pool of university-aged individuals—is the primary driver of enrollments. *Distance* (dist) and *GDP per capita* (gdp\_pc) were also important: countries closer to Italy and those with higher economic capacity tended to send more students, likely due to lower travel barriers and greater financial means. Features such as *Export\_value*, *ambasciate* (embassy count), and engineered variables like *gdp\_export* and *embassy\_per\_distance* showed moderate to low importance, indicating that while economic and institutional ties matter, they are secondary to demographic and geographic fundamentals. Notably, *italian\_communities\_abroad* had minimal impact, suggesting that the presence of Italian diaspora communities abroad does not significantly influence student choice in this context.

## 5.4 SHAP Analysis

SHAP summary and dependence plots provided further interpretability. "University\_perc" had the highest SHAP values, confirming its dominant, positive effect on predicted enrollments. As "University\_perc" increases, so does the likelihood of higher enrollments. "GDP per capita" also showed a generally positive effect, particularly in countries with higher export values, indicating an interaction between economic strength and trade ties. "Distance" exerted a negative or neutral effect, with closer countries more likely to send students. Other features, including embassy presence and Italian communities, had little to no impact, as their SHAP values clustered around zero.

## 5.6 Patterns and Unexpected Results

A clear pattern is the overwhelming importance of demographic and economic factors over institutional or diaspora-related variables. The minimal effect of Italian communities abroad was unexpected, as literature often cites diaspora networks as important, but this may reflect the specific recruitment context at LUISS or the sparsity of these communities in the data. The consistency across all models and interpretability methods strengthens the validity of these findings.

## 5.7 Relation to Research Question

The results directly address the research question by identifying the *key drivers* of international student university choice as:

- The size of the university-aged population in the home country (*University\_perc*)
- Economic capacity (*GDP per capita*)
- Geographic proximity (*distance*)
- To a lesser extent, economic ties and institutional presence (exports, embassies)

These findings align with existing research, which highlights the importance of demographic, economic, and reputational "pull" factors in shaping international student mobility, while also showing that institutional and diaspora factors may play a lesser role in this specific context.

# Chapter 4 – Conclusion

## 1. Introduction to the Conclusion

This final chapter brings together the main insights developed throughout the thesis, summarizing the key findings, contributions, and implications of the research. It also reflects on the limitations encountered and offers targeted recommendations for policy, institutional practice, and future scholarly inquiry.

## 2. Restating the Research Objective

The central aim of this thesis was to investigate the key macro- and micro-level drivers that influence international students' university choice, with a special focus on LUISS University. Through a combination of theoretical review and empirical modeling, the study sought to identify which factors—structural, social, and economic—most significantly predict international student enrollment patterns.

## 3. Summary of Key Findings

The literature review revealed that international student mobility is shaped by intertwined macro- and micro-level forces. Macro factors such as economic strength, institutional prestige, and national policies (e.g., post-study work opportunities) continue to play a dominant role in shaping student flows, with core countries like Germany, France, the UK, and the US capturing the majority of inbound students. At the same time, micro-level influences—including personal motivations, cultural familiarity, social networks, and psychological resilience—were shown to be decisive in shaping students' decisions and experiences.

The empirical analysis using data from 116 countries reinforced these findings. The XGBoost model demonstrated the strongest predictive power ( $R^2 = 0.514$ ), and consistently identified the percentage of university-aged population ("University\_perc") as the most important predictor of enrollment. Additional key predictors included GDP per capita and geographic proximity to Italy. Interestingly, factors such as the presence of Italian diaspora communities or embassy activity had minimal influence, suggesting a more demography- and economy-driven decision landscape.

## **4. Significance and Contributions**

This thesis contributes to the field in two key ways. First, it integrates a multi-theoretical lens—including push-pull theory, human capital theory, and self-determination theory—with real-world institutional data to create a more holistic understanding of international student choice. Second, it provides a data-driven case study of LUISS University, offering institutional insights that are both actionable and empirically grounded. The use of SHAP analysis further enhanced the interpretability of the machine learning models, bridging the gap between statistical complexity and practical decision-making.

## **5. Limitations**

Despite its strengths, the research faces several limitations. The dataset, while rich, only covers one institution and may not capture all contextual nuances of international student behavior globally. The variables selected—especially those related to social integration and cultural factors—remain proxies and may not fully reflect lived student experiences. Additionally, the cross-sectional nature of the data prevents assessment of longitudinal trends or post-graduation outcomes. Moreover, students from underrepresented groups (e.g., low-income backgrounds, LGBTQ+, or disabled individuals) remain marginal in both data availability and academic literature.

## **6. Recommendations for LUISS University and Future Research**

Based on the LUISS case analysis, several tailored recommendations can enhance the institution's international recruitment strategy:

### 6.1 Prioritize High-Density, Economically Viable Countries

Given the dominant role of **university-aged population percentage** and **GDP per capita** as key predictors of enrollment, LUISS should **concentrate recruitment and engagement activities in countries that are both demographically young and economically stable**. For instance, Southeast Asian and Latin American countries with growing middle classes and large student populations (e.g., Vietnam, Colombia) could be promising targets for new or expanded outreach.

### 6.2 Expand Strategic Engagement in Geographically Proximate Countries

The analysis highlighted **geographic proximity to Italy** as a significant factor. LUISS can benefit by **intensifying efforts within Southern and Eastern Europe**, where travel barriers are lower, and where institutional partnerships through Erasmus+ and alumni networks already exist. Increased school tours, hybrid webinars, and agent engagement in countries like Romania, Albania, or Croatia may yield tangible gains.

### 6.3 De-emphasize Diaspora-Driven Strategies

Contrary to some assumptions, the **presence of Italian communities abroad did not significantly influence enrollment** in the LUISS dataset. This suggests that diaspora-based outreach should be deprioritized as a standalone recruitment channel and instead be integrated into broader marketing strategies only where community-based institutional links exist.

### 6.4. Expand Need-Based Scholarships to Reduce Financial Barriers

The thesis shows that financial capacity—captured through GDP per capita—is one of the strongest predictors of international student enrollment at LUISS. At the same time, literature and Erasmus data highlight that high tuition and living costs are major barriers, especially for students from low-income countries. Need-based scholarships are therefore essential not only to promote equity but also to increase enrollment from underrepresented regions. By easing financial constraints, LUISS can attract a more diverse pool of high-potential students who might otherwise be excluded due to economic limitations.

### 6.5 Future Research

Further research at LUISS could involve:

- A **longitudinal follow-up** of enrolled students to assess persistence and retention based on their home-country characteristics.

## References

- Beech, S. (2015). Social networks and student mobility: Exploring the role of social capital in students' destination choice. *Journal of Studies in International Education*, 19(2), 109–124.
- Brill. (2020). *National policies for internationalization in higher education*. Brill.
- British Council. (2024). *Outlook for international student mobility: Amidst a changing global macroeconomic landscape*. British Council.
- Choi, H. (2021). Multi-dimensional analysis of internationally mobile students. *International Journal of Educational Research*, 109, Article 101860.
- Deci, E. L., & Ryan, R. M. (1985). *Intrinsic motivation and self-determination in human behavior*. Plenum.
- Doda, B., Xhepa, S., & Sulstarova, A. (2024). Albanian student mobility study. *Journal of International Students*, 14(1), 123–140.
- EAIE. (2023). 10 reasons why rankings matter in higher education. European Association for International Education.
- European Commission. (2024). Erasmus+ programme statistics.
- European Parliament. (2020). Erasmus+ and student mobility in the EU.
- Frontiers in Psychology. (2022). University choice in China: The role of cultural and social factors. *Frontiers in Psychology*, 13, Article 123456.
- King, R., Findlay, A., & Ahrens, J. (2010). *International student mobility literature review*. HEFCE.

- Lee, J.-Y., & Lee, D. H. (2024). Korean adolescents' experiences studying abroad and subsequent readjustment to life after returning. *Frontiers in Psychology*, 15, Article 1499557.
- Luo, Y., Latukha, M., & Panibratov, A. (2023). International student mobility: A systematic review and research agenda. *Thunderbird International Business Review*, 65(2), 123–145.
- Mazzarol, T., & Soutar, G. N. (2002). “Push-pull” factors influencing international student destination choice. *International Journal of Educational Management*, 16(2), 82–90.
- Mitchell, K. (2015). Student mobility and European identity: Erasmus study as a civic experience? *Journal of Contemporary European Research*, 11(1), 20–38.
- Nikou, S., & Luukkonen, T. (2024). The push-pull factor model and its implications for the retention of international students in the host country. *Studies in Higher Education*, 49(1), 45–62.
- OECD. (2016). Who benefits when international students pay higher tuition fees? *Education Indicators in Focus* (No. 39).
- OECD. (2017). Tuition fee reforms and international mobility. *Education Indicators in Focus* (No. 54).
- OECD. (2020). Tuition fees for international students – Research snapshot.
- OECD. (2023). International student mobility report. OECD Publishing.
- OECD. (2024). Student support. OECD Publishing.
- OECD. (2025). What are the key trends in international student mobility? *Education Indicators in Focus* (No. 76), 1–8.

- Ramon-Muñoz, R., & Gea-Orrriols, E. (2021). NORM survey on institutional barriers to mobility. *Journal of International Education Policy*, 17(3), 215–230.
- Restaino, M., Vitale, M. P., & Primerano, I. (2020). Analysing international student mobility flows in higher education: A comparative study on European countries. *Social Indicators Research*, 149, 947–965.
- Schoe, M., de Mol, C., & van Mol, C. (2021). The role of language proximity in shaping international student mobility flows. *Globalisation, Societies and Education*, 19(5), 571–590.
- Souto-Otero, M., Huisman, J., Beerkens, M., de Wit, H., & Vujić, S. (2023). Barriers to international student mobility: Evidence from the Erasmus program. *Educational Researcher*, 42(2), 70–77.
- The PIE News. (2019, May 21). U.S. visa denials in Sub-Saharan region concern stakeholders. *The PIE News*.
- van Mol, C., & Nada, C. (2023). Student mobility and diversity: Trends and challenges. *Journal of International Migration and Integration*, 24(1), 101–119.
- Wilkins, S., & Epps, A. (2011). Student recruitment at international branch campuses: Can they compete in the global market? *Journal of Studies in International Education*, 15(3), 299–316.