

National Tourism Promotion

Data Science for Marketing
2024/2025

Professors

Nuno António, Vasco Jesus and Francisco Afonso

Group 42:

Diana Rodrigues 20240110

Laura Tomás 20240116

Matilde Ferreira da Costa 20240115

Miguel Pinto 20240895

Index

Executive Summary	4
Context.....	5
Defining the problem	6
Framing Data Mining Objectives.....	6
Methodology	7
Results and Discussion	8
Timeline Analysis	8
Nationality of Portugal Visitors.....	11
Comparison with main competitors	13
Association Rules for Frequent Patterns	15
RFM Model for User Segmentation	16
Business recommendations	17
Bibliography.....	20

Index of figures

Figure 1 Number of Reviews per Month (2018-2021).....	8
Figure 2 Average Review Rating per Local Name	9
Figure 3 Heatmap of Reviews by Top 10 Countries and Month (Ordered)	11
Figure 4 Total Reviews vs. Average Review Rating by Country.....	13
Figure 5 Top 10 Monuments by Review Rating	14
Figure 6 Number of Customers by RFM Segment	16

Index of tables

Table 1 Holidays per Country in the selected months	12
---	----

Executive Summary

The COVID-19 pandemic reshaped the global tourism landscape, challenging destinations to adapt to shifting visitor behaviors and expectations. For Portugal, a country well-known for its rich history, great weather, breathtaking landscapes, and warm hospitality, the stakes were particularly high. The Portuguese National Tourism Board (NTBO) found itself at a crossroads: how to sustain and grow its reputation as a premier travel destination while facing the competition from neighbors like Spain, Italy, France, and Greece. As satisfaction ratings declined and long-standing issues like infrastructure and crowd management became more pronounced, NTBO turned to data-driven insights to reimagine its strategies and ensure Portugal's continued appeal.

To address these challenges, we applied on our project the CRISP-DM framework to analyze visitor trends from 2018 to 2021. The analysis focused on evaluating the impact of the pandemic on visitation and satisfaction, benchmarking Portuguese attractions against competitors, and uncovering patterns in visitor behavior to inform strategic recommendations. We used data mining techniques, including association rules, timeline analysis, and segmentation, to identify key opportunities for improvement.

Our recommendations include the creation of a “Navigate Portugal Pass” to enhance flexibility and engagement, as well as targeted digital campaigns in the USA, UK, and untapped markets like France and India. These efforts are designed to improve visitor satisfaction, promote lesser-known attractions, and compete effectively with rival tourism destinations.

By implementing these recommendations, NTBO can mitigate the pain points identified during the analysis and position Portugal as a dynamic and sustainable tourism destination. For instance, the “Navigate Portugal Pass” encourages exploration of underrepresented attractions, while targeted campaigns during key holiday periods optimize visitor inflows. Collaborating with neighboring Spain to develop cross-border travel packages offers an additional opportunity to enhance regional tourism and attract a broader audience.

This project emphasizes the critical importance of adapting to shifting visitor preferences and addressing challenges amplified by the pandemic. The insights and innovative strategies outlined aim to equip NTBO with a roadmap for sustaining tourism growth, enhancing visitor experiences, and maintaining Portugal's competitive edge. These efforts align seamlessly with NTBO's mission of establishing Portugal as a globally recognized leader in sustainable and dynamic tourism. For further details, please refer to the complete report and analysis. [\[1\]](#)

Context

Tourism of Portugal (Turismo de Portugal) [\[2\]](#), as the national tourism authority, is committed to enhancing Portugal's global standing as a premier destination for visitors. In light of the COVID-19 pandemic, significant shifts in travel behaviors and perceptions have likely occurred, necessitating a data-driven approach to understand and adapt to these changes. Leveraging the CRISP-DM framework, this project aims to provide insights into visitor patterns and inform strategic decision-making, by means of business recommendations and marketing strategies based on insights we had.

About *Turismo de Portugal*. *Turismo de Portugal* operates under the Ministry of Economy and has a broad mandate to promote and sustain national tourism activities. Its core responsibilities include:

- Promoting Portugal as a leisure, living, studying, or investing destination.
- Supporting infrastructure development and fostering innovation in the tourism sector.
- Developing human resources by providing specialized training through its network of tourism schools.
- Enhancing sustainability and quality in tourism offerings.
- Regulating and monitoring activities such as gambling and responsible tourism practices.

With its "Visit Portugal" platform and numerous campaigns, such as "Can't Skip Portugal" and "It's Not Tourism, It's Futurism," Turismo de Portugal has positioned Portugal as a forward-thinking, sustainable tourism destination. Additionally, collaborates with Regional Tourism Promotion Agencies and international partners to ensure coherent and impactful promotion efforts.

According to the Travel BI [\[3\]](#), in 2021, the tourism accommodation sector in Portugal welcomed 14.46 million guests and recorded 37.33 million overnight stays. This marks a recovery compared to the previous year, with increases of 38.6% in guests and 44.7% in overnight stays. However, when compared to pre-pandemic levels in 2019, the sector experienced a sharp decline, with 46.7% fewer guests and 46.8% fewer overnight stays.

These highlight the significant impact of the pandemic on Portuguese tourism. While there are signs of gradual recovery, the sector still has a long way to go to reach pre-pandemic numbers, reflecting the challenges of rebuilding amidst competitive pressures from other European destinations.

Defining the problem

The aim of this initiative is to support the Portuguese National Tourism Board (NTBO) in the role of promoting, regulating, and supporting tourism, acting as the bridge between public entities, private companies, and international stakeholders, promoting initiatives to strengthen the national economy via tourism related activities.

The problem is about providing NTBO with valuable insights into visitor's trends and preferences about the Portuguese attractions. How they performed in comparison with our main tourism competitors, identified in the Travel & Tourism Competitiveness Report [\[4\]](#), specifically, Spain, Italy, France and Greece, that also have in common the proximity, culture, weather and gastronomy. Complementary the effects of the COVID-19 on the visitors' trends and preferences was also under analysis.

Additionally, insights from the Travel BI [\[5\]](#) report highlight shifting travel preferences among residents. Spain and France remained the primary international destinations, accounting for 41.6% (+3.2 p.p. compared to 2022 and +9.0 p.p. versus 2019) and 10.1% (-0.7 p.p. compared to 2022 and -2.2 p.p. versus 2019) of trips, respectively. Italy maintained its position as the third most popular destination (6.9% of total trips, +0.2 p.p. compared to the previous year and +0.7 p.p. compared to 2019).

Framing Data Mining Objectives

To present solutions to the problem, this project will:

- Evaluate visitor patterns pre- and “post”-pandemic by understanding how visitors' frequency and rating distribution have shifted.
- Identify where these visitors come from and explore whether their reasons for visiting are tied to holidays from their home countries.
- Benchmark Portugal against its competitors to identify similarities between attractions, improve our performance, and enhance market potential.
- Uncover visitor associations, analyze commonalities in visitor attractions and activities.
- Segment visitors based on their behavior and preferences, leveraging metrics like Recency, Frequency, and Ratings (RFR).

Methodology

For our group project we employed the CRISP-DM (Cross-industry standard process for data mining) process model. That has the following phases:

Business understanding

Firstly, we needed to understand what our business was and what it wanted to put forward. To study Turismo de Portugal, we searched its own website to discover their project and campaigns. Afterwards, we defined the problem and framed data mining objectives.

Data understanding

Our base (type: xlsx) and extra (type: csv) dataset was given by our professor. It had their descriptions in a separate notebook. Our job was visualizing the content of the files, understanding what types of data and properties it had and how we could work with them.

Data preparation

In this phase we prepared our files mentioned earlier for modelling. By using python scripts on development environment for notebooks and code, named Jupyter and Visual Studio Code, we selected the data it was relevant to our project. Handled missing values and outliers by correcting or remove values. Saving the result into a cleaned data file (type: csv).

Modeling

For modeling, we used the same development environment as the data preparation to study the data mining objectives. We created different types of charts and tables to improve communication and data visualization.

Evaluation

The results on the previous phase were discussed and complemented with articles and news found online. We included business recommendations and marketing strategies based on insights we had.

Deployment

The outcome of this phase is our business recommendations.

Results and Discussion

Timeline Analysis

The goal of this section is to determine if COVID-19 pandemic had an impact on visits in Portuguese attractions and to analyze how it affected them.

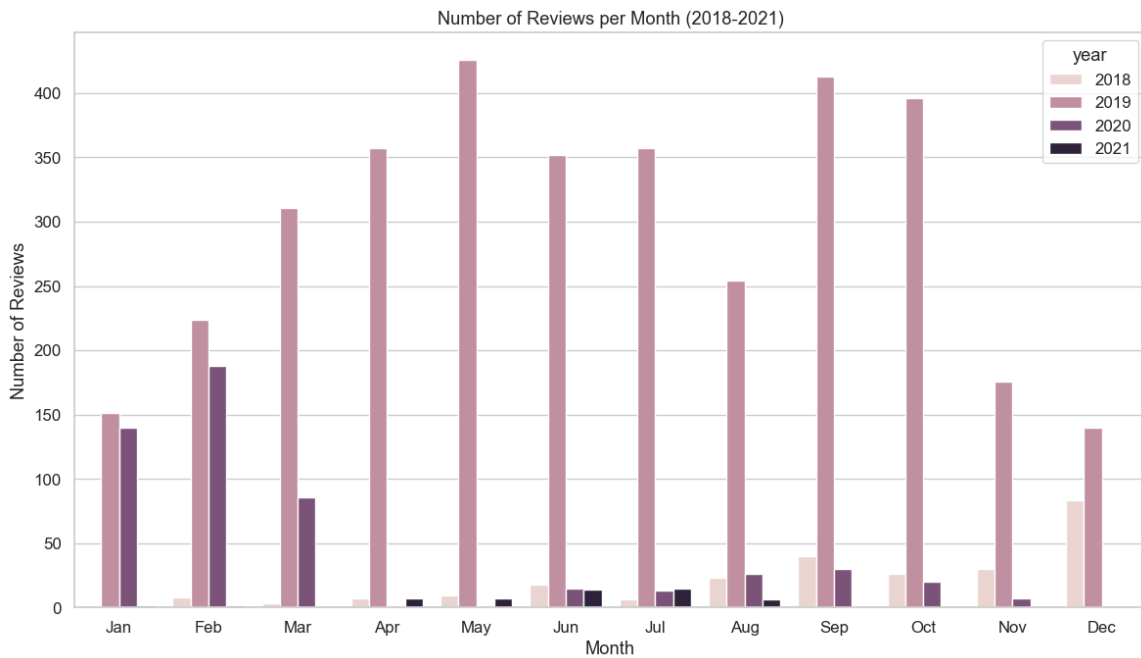


Figure 1 Number of Reviews per Month (2018-2021)

We only have data from visits between February of 2018 and August of 2021. Based on this image, we can see that in 2018 had a slight increase in the number of visits throughout the year. From the beginning of 2019, there was an abrupt and constant increase until November and December with a slightly drop. In 2020, until February we observed the same pattern as in the previous year. However, in March, when the pandemic began, there was a slight drop compared to February and a significant decrease on the remaining of the year and through 2021, those numbers continued to stay as low. Making us conclude that the pandemic had a notable impact on decreasing the number of visits.

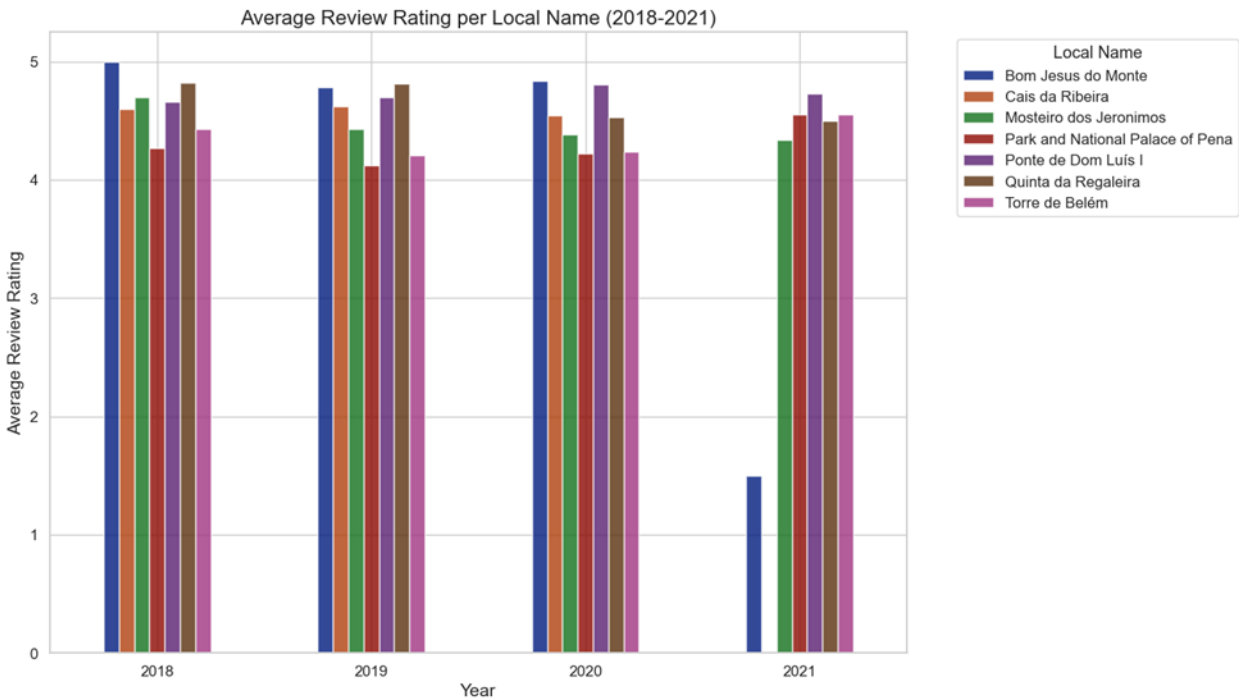


Figure 2 Average Review Rating per Local Name

In order to assess how the pandemic impacted the ratings of Portuguese attractions, we created a bar chart displaying the average rating per attraction for each year. The attractions included in this analysis are those listed in the top 100 attractions file, as referenced in the methodology section.

- **Bom Jesus do Monte (blue bar).** This attraction had a minor drop on average rating between 2018 and 2019 and appears to have an important drop between 2020 and 2021. Nonetheless and there were below three out of five. After reading them we understood that their complaints were about the cleanliness and the number of stairs. [\[6\]](#) Bom Jesus do Monte can be accessed from the bottom or top.
- **Cais da Ribeira (orange bar).** This attraction had almost the same average rating thought the years, except in 2021, we do not take the ratings into consideration because it had no reviews.

- **Mosteiro dos Jerónimos (green bar).** This attraction had a slight drop between 2018 and 2019, maintaining almost the same average rating as 2019 for the rest of the years. Before the pandemic, after a quick analysis on review text, we realized that the long queues and the time spent on them were the major issue reported. In 2020 and 2021, the complaints are different. The areas open for the public and the lack of information inside were most common.
- **Park and National Palace of Pena (red bar).** This attraction had a significantly average rating raise in 2021 compared to previous years. We read some of the lowest rating reviews from those previous years and the negative aspects mentioned were the difficult access and the long queues again. The pandemic lead to a break on visits on this attraction [\[7\]](#) which means that the problem of long queues was solved by itself resulting on better ratings in 2021.
- **Quinta da Regaleira (brown bar).** During pandemic years (2020-2021) the ratings dropped slightly compared to previous years. The main issues reported in those years were that some paths were closed due to covid restrictions or maintenance reasons and people were upset that they were not notified before.
- **Torre de Belém (pink bar).** This attraction had better ratings in 2021 than in previous years. The same issues reported were the same as other attractions, long lines.

Nationality of Portugal Visitors

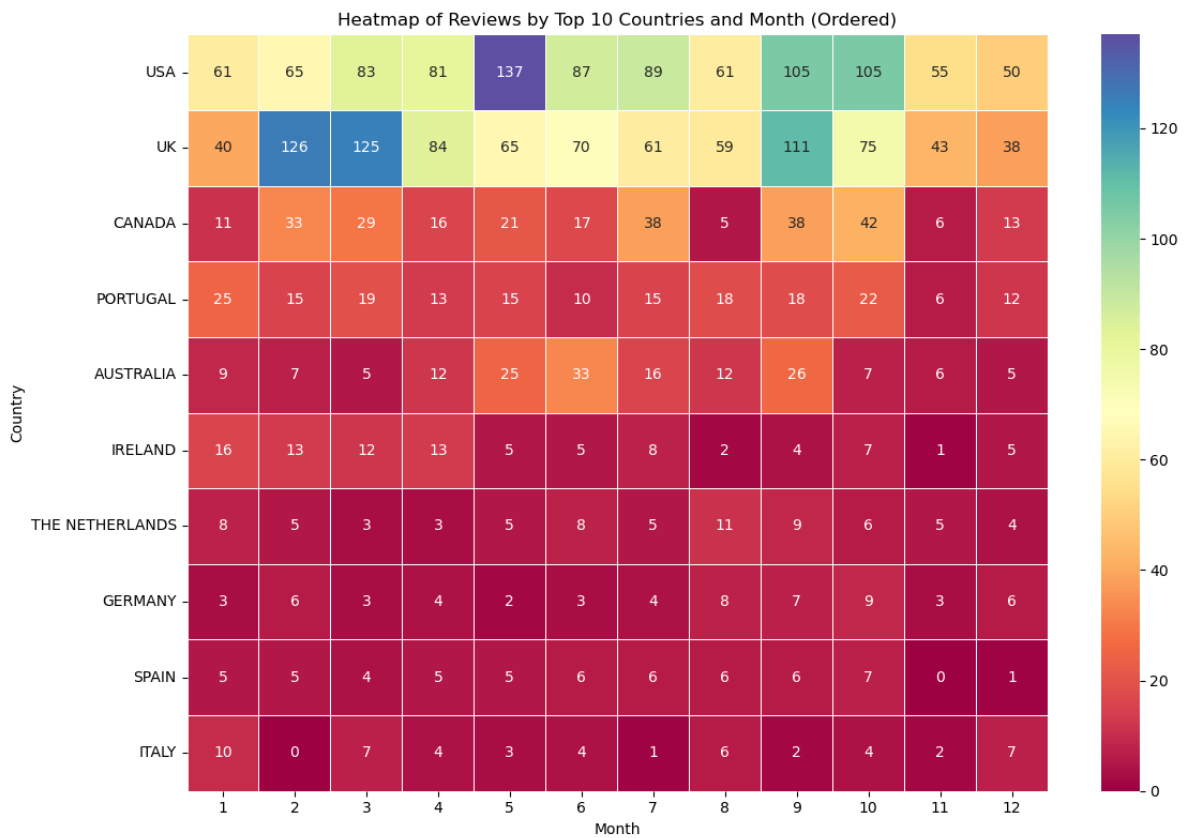


Figure 3 Heatmap of Reviews by Top 10 Countries and Month (Ordered)

Country	Holiday Name	Movable holiday?
United States of America	Memorial Day	No
United States of America	Labor Day	No
United States of America	Columbus Day	No
United States of America	Veterans Day	No
United Kingdom	St. Patrick's Day (Northern Ireland)	No
United Kingdom	Good Friday	Yes
United Kingdom	Easter Monday	Yes
Canada	Labor Day	No
Canada	Thanksgiving	No
Canada	Canada Day	No
Portugal	Ano Novo	No
Portugal	Sexta-feira Santa	Yes

Portugal	Páscoa	Yes
Portugal	Implantação da República	No
Ireland	Easter Monday	Yes
Ireland	New Year's Day	No
Netherlands	Eerste Pinksterdag	Yes
Netherlands	Tweede Pinksterdag	Yes
Netherlands	Hemelvaart	Yes
Germany	Tag der Deutschen Einheit	No
Germany	Reformationstag	No
Spain	Asunción de la Virgen	No
Spain	Día de la Hispanidad	No
Italy	Capodanno	No
Italy	Epifania del Signore	No
Italy	Pasqua di Resurrezione	Yes
Italy	Lunedì dell'Angelo	Yes
Italy	Immacolata Concezione	No
Italy	Natale	No
Italy	Santo Stefano	No

Table 1 Holidays per Country in the selected months

Our goal was to determine where visitors to Portugal's attractions came from and whether there is a relationship between holidays in their home countries and their visits. To achieve this, we started by understanding what countries review Portuguese attractions the most, which are ordered top to bottom in Figure 3 the country with the most reviews to the least. (United States of America, United Kingdom, Canada, Portugal, Australia, Ireland, the Netherlands, Germany, Spain, Italy). Then, with the results of Figure 3, we selected the three months with the most reviews for each country. Regarding ties, proximity to other months was the deciding criterion.

After analyzing our extra file, which has the holidays for each country, we obtained Table 1 which contained the country, the holiday name, and if it is a movable holiday, meaning if the holiday can be celebrated on a different month. For having a relationship between their holidays and visits, we considered only if 70% of the holidays in that month in each country were not movable. That means only the United States of America, Canada, Germany, Spain, and Italy have that relation.

Comparison with main competitors

To compare Portuguese attractions with the main tourism competitors, Spain, Italy, France and Greece, we started by analyzing the total reviews and the average review rating of each country.

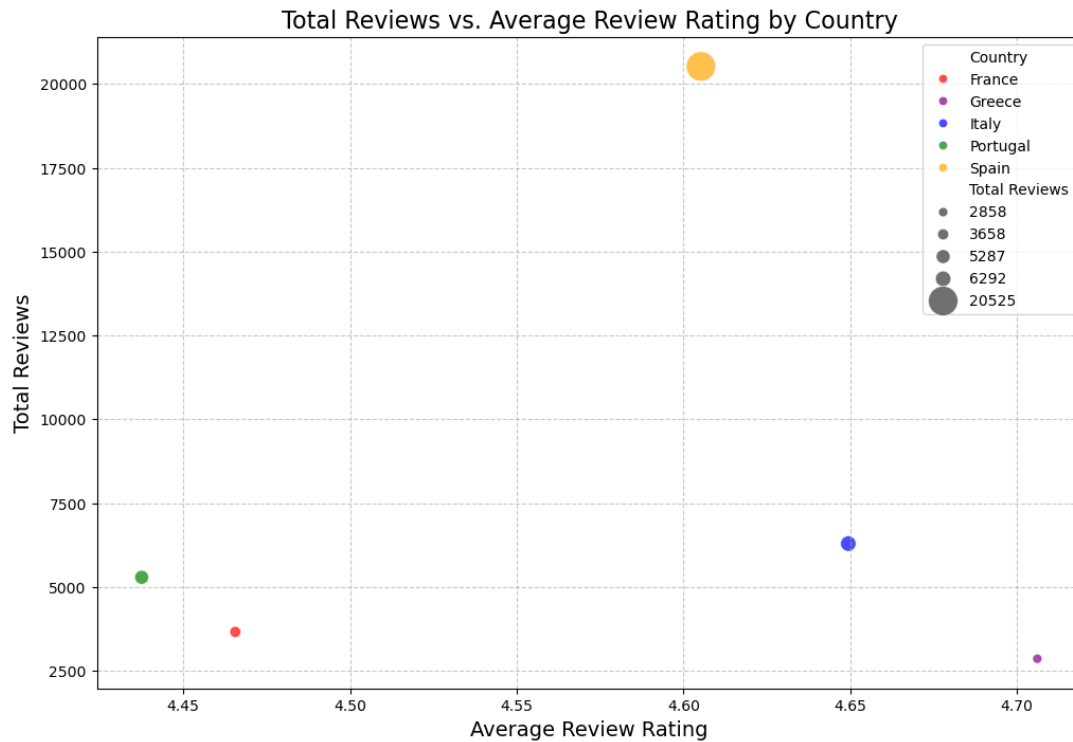


Figure 4 Total Reviews vs. Average Review Rating by Country

We can observe that the average review rating for Portugal (4.44) is lower compared to competing countries. This indicates that those that are coming to see Portuguese attractions are less satisfied compared to attractions in competing destinations.

Then, we generated a bar chart to identify the 10 monuments with the highest review ratings.

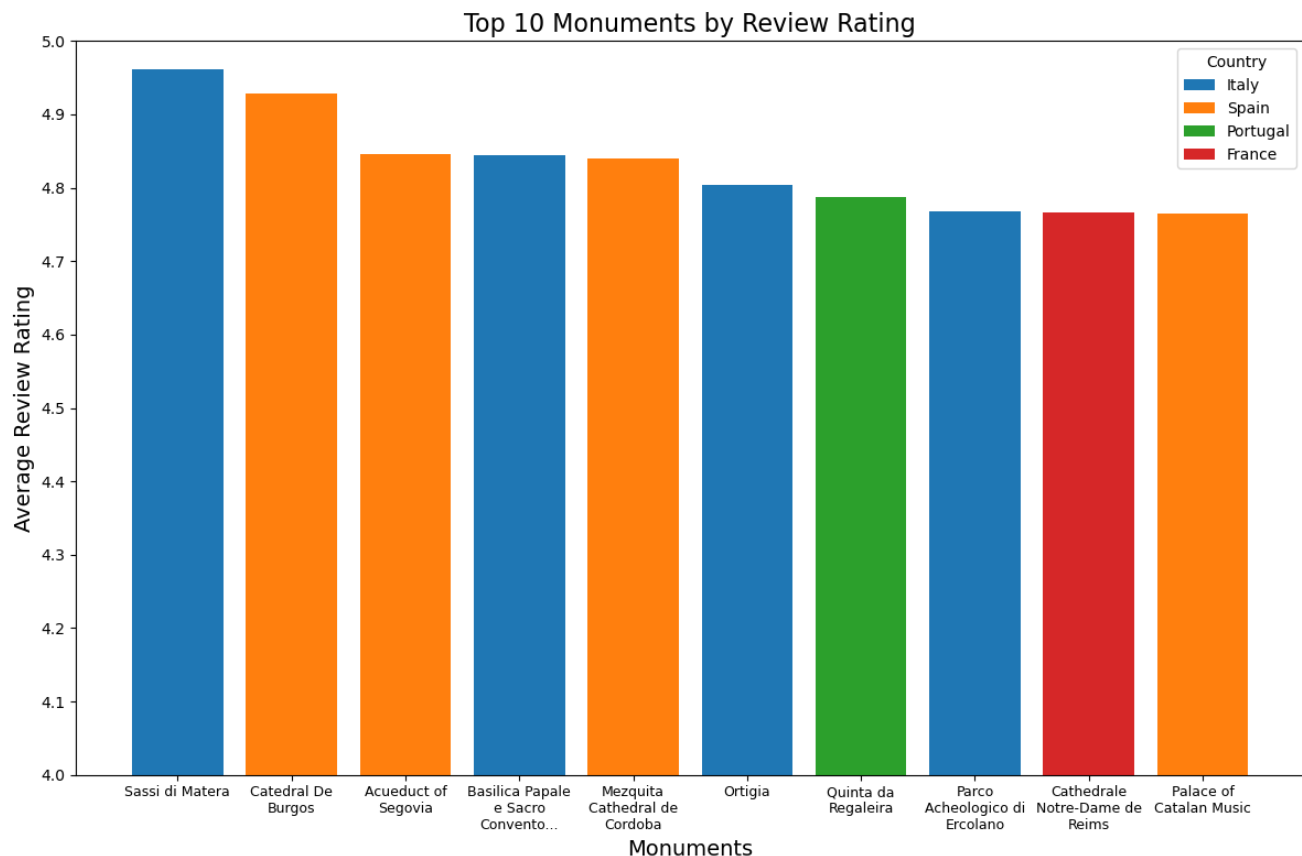


Figure 5 Top 10 Monuments by Review Rating

The chart “**Top 10 Monuments by Review Rating**”, figure 5, shows that Italy and Spain are the countries with the highest number of top-rated monuments, each with four monuments. Quinta da Regaleira ranks as the 7th highest-rated monument and is the only Portuguese monument in the Top 10. This analysis suggests that tourism in Portugal has room for improvement, and we should examine the strategies employed by our main competitors to achieve these superior results.

To conduct a more detailed analysis, we used Quinta da Regaleira to identify attractions with the most similar review ratings. The results predominantly included other Portuguese monuments and several Spanish monuments which shows the significant similarity between the two countries. These results were expected, given that Spain is the closest country to Portugal with similar cultures.

We continue to analyze the countries that made the most reviews in Portugal compared to Spain, using these values as a metric to identify the countries that visit the most each country. We obtained similar results, with the USA and the UK as the top results. The main differences were that Spain includes India and France among the top 10 countries, while Portugal does not.

Association Rules for Frequent Patterns

To better understand visitor behavior and identify patterns in attractions visited together, we applied **association rules to find which attractions** are frequently co-visited. By analyzing the patterns, we observed how attractions naturally group together, often influenced by their geographic proximity or regional connection. From those patterns, we observed that:

- **Torre de Belém** and **Mosteiro dos Jerónimos** were frequently co-visited. This makes sense as both attractions are in Belém, Lisbon, and are iconic landmarks of Portuguese history and architecture.
- **Parque Nacional e Palácio da Pena** and **Quinta da Regaleira** were other frequent pairings, reflecting Sintra's strong draw as a tourist destination, given its cultural heritage and proximity to the sites.
- In Porto, **Ponte de D. Luís I** and **Cais da Ribeira** were commonly visited together, since you can naturally find these landmarks along the Douro River.
- We have also been able to tell that within the Metropolitan Area of Lisbon, in some cases, people who visit Belém also end up visiting Sintra on their visit. So an association that is made is that if one visits **Torre de Belém** and **Palácio de Sintra**, one will probably also visit **Quinta da Regaleira** or **Mosteiro dos Jerónimos**, for example.

RFM Model for User Segmentation

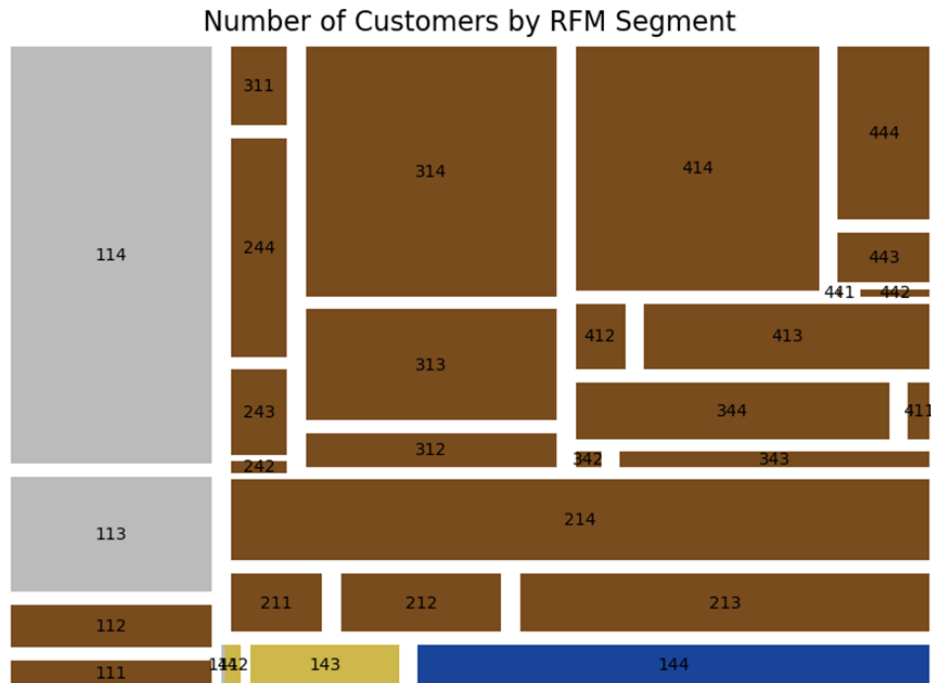


Figure 6 Number of Customers by RFM Segment

In this figure we can observe the number of customers by each RFM segment. RFM reflects the levels of **R**ecency (1st number), how many days have passed since the last visit, **F**requency (2nd number), number of visits made/reviews left, and **M**onetary value (3rd number), average rating given by visitor. The size of each area in the chart represents the number of tourists in that segment, making it easy to identify the most representative groups. The closest to the number 4 the better for Frequency and Monetary, whereas the closest to 1 the better for the Recency, meaning it visited an attraction recently.

The colors identify how ideally the segments are. The blue is the best possible ('144'), next is gold, followed by silver and bronze.

Our figure indicates that we have the most ideal visitors ('144'). However, we have some concerning segments such as '314' and '414' that are bigger than it should be, meaning that besides the customers gave Portuguese attractions good reviews (above 4 out of 5), their last visit was a long time ago and they only left a review on one attraction.

There is also some room for improvement, in gold and silver. In silver, we have visitors that visit Portuguese attractions recently but not frequently but leave a good review. The gold values, we should improve on the monetary value, since visitors can give a better review.

Business recommendations

Timeline Analysis

For each attraction mentioned on this phase we have the following recommendations:

- **Bom Jesus do Monte.** About cleanliness we can only advise the obvious. However, regarding the high number of stairs we can advise tourists about this and promote them to access from the top, with this they can enjoy the views without becoming tired. This could include detailed guides, route suggestions, or even virtual walkthroughs, that you could access through the already existent website.
- **Cais da Ribeira.** The recommendation we have for this attraction is to continue to promote their majors' strengths, free access and its proximity to the historical center.
- **Mosteiro dos Jerónimos.** Our suggestions are to continue investing in measures to maintain lower waiting time on queues, an example of it is through the *Lisboa Card*. [8] If possible, put more information (e.g., On posters) inside the attraction about its history and promote the new feature to tourists.
- **Park and National Palace of Pena.** We advise better management of bookings. Regarding the difficult access, our recommendations are promoting where to park and how to get to the attraction by public transport or by walking.
- **Quinta da Regaleira.** To decrease misunderstandings, we recommend creating a digital map with the places you can visit and the most popular attractions. This way tourists can through a virtual map, access the different places and plan their trip, ahead of time.
- **Torre de Belém.** We have the same recommendation as *Mosteiro dos Jerónimos*, continue promoting fast passes or pre-booked tickets, such as *Lisboa Card*. [8]

Nationality of Portugal visitors and comparison with our main competitors

Building on insights from our analysis of visitor nationalities and comparisons with main competitors, we propose a comprehensive approach to attract more visitors to Portugal by leveraging targeted campaigns, cross-border collaborations, and enhanced digital visibility.

Targeted Digital Campaigns

To capitalize on existing and emerging markets, we recommend launching campaigns tailored to the preferences and peak travel periods of specific countries.

- Promote Portuguese attractions to United States of America, Canada, Germany, Spain and Italy. Use of target ads during the months with the most reviews in each country. For example, for Italy citizens, use specific campaigns to encourage visits during their Christmas holidays, from December 8th and January 6th.
- The goal is to increase tourism in Portugal by promoting our main monuments in the countries that currently visit us the most, USA and UK, as well as targeting countries that visit Spain more frequently but have lower numbers in Portugal, such as India and France, which represent potential markets to explore. We suggest running targeted digital campaigns in current (USA and UK) and potential (India and France) markets, adapting the content to the preferences of these countries.

Cross-Border Tourism Packages

Develop Cross-Border Packages (Spain and Portugal) by establishing partnerships with Spain to offer travel packages focusing on the main monuments of Portugal, as well as Spanish tourist attractions.

Enhanced Digital Visibility

Another measure we suggest is to engage with travel platforms and airlines to increase visibility on platforms like TripAdvisor and Booking through advertisements focused on our monuments. This way we could partner with airlines and travel agencies to feature Portuguese attractions in advertisements and travel bundles. But also we could highlight top-rated monuments and address common tourist pain points to enhance reviews and visibility.

Association Rules

Considering our analysis in frequent patterns in tourists visits we suggest:

Navigate Portugal Pass

We are aiming to create a unique national **pass**. Unlike the Lisboa Card, the Navigate Portugal Pass offers options for all major regions, including Lisbon, Porto, Sintra, and beyond. Integrated with a mobile app linked to the "Visit Portugal" platform, visitors can choose regional bundles tailored to frequently associated attractions, such as the **Belém & Sintra Bundle** (Torre de Belém, Mosteiro dos Jerónimos, Palácio da Pena, and Quinta da Regaleira) or the **Porto Highlights** (Ponte de D. Luís I, Cais da Ribeira, and Bom Jesus do Monte, with Braga as an optional add-on). This nationwide pass also allows for customizable itineraries, encouraging travelers to mix iconic landmarks with hidden

gems while enjoying discounted prices. This approach not only distributes foot traffic more evenly but also introduces visitors to other attractions, enhancing their overall experience while supporting less-visited sites.

To make the experience more engaging, the pass integrates a **gamified element**. Visitors collect **points or stamps** at each attraction they visit, that allows them to unlock a **special reward**. This could be discounts at local restaurants, free souvenirs, or a ticket to another attraction. This playful feature incentivizes exploration and adds a layer of excitement to the journey.

The pass also fosters connectivity by allowing visitors to share their custom itineraries on social media through the pass platform. This feature encourages others to follow similar "paths" and makes the experience more interactive for groups or families planning their visits together.

RFM Model

Considering our analysis in the RFM Model Segmentation we suggest:

Social media campaign and giveaway

The proposed campaign, for Visit Portugal social media, aims to reconnect with past visitors who have only experienced a single attraction in Portugal but left a positive review. By leveraging insights from their travel patterns, the campaign will re-engage these individuals with a giveaway, including the **Navigate Portugal Pass** and a **3 day stay in a location of their choice**. Due to limitations in directly contacting these users—such as privacy concerns and lack of direct access to user data—the strategy combines broader segmentation techniques with strategic partnerships.

To achieve this, we propose using the existing data in the database to identify behavioral patterns, such as visitors who provided high ratings to attractions and only visited one or two locations. This data can be used to create lookalike audiences for targeted social media campaigns on platforms such as Instagram and Facebook. The campaign will focus on specific markets, including the USA, UK, France, and Germany, where visitors have shown interest in Portuguese tourism.

In addition, a partnership with TripAdvisor is recommended to enhance targeting capabilities. By collaborating with TripAdvisor, Turismo de Portugal could use their ad platform to target users who have previously reviewed Portuguese attractions. While privacy laws may limit direct outreach, the partnership would enable highly focused ad campaigns to reach users who are most likely to engage. Cross-posted content between Visit Portugal and TripAdvisor would further maximize visibility and engagement.

Bibliography

- [1] Harvard Kennedy School. (2018). *How to write an executive summary*. Retrieved January 5, 2025, from https://projects.iq.harvard.edu/files/hks-communications-program/files/how_to_write_an_exex_summ_to_use_4_18_18.pdf
- [2] Turismo de Portugal. (n.d.). *Homepage*. Retrieved November 15, 2024, from <https://www.turismodeportugal.pt/pt/Paginas/homepage.aspx>
- [3] Travel and Tourism Competitiveness Report (2015). Retrieved November 15, 2024, from <https://www.turismodeportugal.pt/SiteCollectionDocuments/estrategia/LET-CIT/LET-CIT-ET2027-portugal-competitiveness.pdf>
- [4] Travel BI (2021). Retrieved November 15, 2024, from <https://travelbi.turismodeportugal.pt/turismo-em-portugal/turismo-em-numeros-2021/>
- [5] Travel BI (2023). Retrieved November 15, 2024, from <https://travelbi.turismodeportugal.pt/mercados/portugal/>
- [6] Santuário do Bom Jesus do Monte. (n.d.). *Homepage*. Retrieved December 5, 2024, from <https://bomjesus.pt/>
- [7] Visão. (2021, February 24). *Covid-19: Parques e monumentos de Sintra com quebra de 80% de visitantes em 2020* [Covid-19: Parks and monuments of Sintra with an 80% drop in visitors in 2020]. Retrieved December 18, 2024, from <https://visao.pt/atualidade/cultura/2021-02-24-covid-19-parques-e-monumentos-de-sintra-com-quebra-de-80-de-visitantes-em-2020/>
- [8] Lisboa Card. (n.d.). *Homepage*. Retrieved December 28, 2024, from <https://lisboa-card.com/>