ONLINE FIRST ARTICLE
doi.org/10.26398/IJAS.16199

Statistica Applicata - Italian Journal of Applied Statistics

Received: 19.05.25/ **Accepted:** 25.07.25
Published online: 25.07.25

Understanding Intra-Destination Tourist Behaviors as a Tentative Strategy to Mitigate Overtourism in Florence

Stefano Masini
Dept. of Computer Science, University of Pisa, Pisa, Italy
Silvia Bacci
Dept. of Statistics, Computer Science, Applications "G. Parenti", University of Florence, Florence, Italy
Alessandra Petrucci
Dept. of Statistics, Computer Science, Applications "G. Parenti", University of Florence, Florence, Italy
Carlo Francini
World Heritage and UNESCO relationship Office, Municipality of Florence, Florence, Italy
Bruno Bertaccini[1]
Dept. of Statistics, Computer Science, Applications "G. Parenti", University of Florence, Florence, Italy

Abstract *Tourism is not only a movement of people, but also a key source of income for local and national governments. However, when visitor flows are uncontrolled, they can threaten the sustainability of cities, as is happening in many tourist destinations around the world, and particularly in Italy. In this case, the term "overtourism" has become increasingly relevant to describe the negative impact of tourism on society and the environment. To address the challenges posed by overtourism, policy-makers need to adopt information-based strategies driven by data analysis. In this contribution, we show how statistical tools can help uncover valuable insights from data on tourist behavior, such as those collected automatically through museum pass systems in cultural cities. In particular, we analyze data from Florence, focusing on the sequences of museum visits recorded via the FirenzeCard, the pass that grants access to the city's museums and exhibitions. We apply a latent class item response model to identify homogeneous groups of tourists with similar museum preferences. This approach provides policy-makers with a practical tool to gain a deeper understanding of tourists' decision-making processes and to design targeted promotional strategies tailored to specific tourist profiles.*

Keywords: *Overtourism, intra-destination tourist behaviors, Item response theory, Latent class models, Network analysis*

[1]bruno.bertaccini@unifi.it (**corresponding author**)

1. Introduction

Tourism is a dynamic and multifaceted industry that plays a pivotal role in shaping global destinations across the cultural, economic, and social dimensions. Regarding the cultural dimension, tourism influences a destination's identity, engaging visitors with local traditions, heritage sites, museums, and cultural events, thereby contributing to the preservation and promotion of cultural practices. In the economic context, tourist flows yield economic benefits, affecting local businesses, job creation, and overall development, making tourism revenue a vital income source for many communities and regions. In terms of the social dimension, tourism alters social dynamics by fostering interactions between locals and visitors, promoting cultural exchange, tolerance, and the development of a more diverse and interconnected community (Lindberg and Johnson, 1997; Sharpley and Telfer, 2002; Smith, 2015).

These dimensions have been particularly reinforced in the past two decades due to the advances in digital evolution. The Internet has revolutionized the way people search for information, communicate and do business. This revolution has had a significant impact on the tourism industry, whose businesses use digital marketing strategies to reach potential tourists through targeted advertisements, sponsored content and partnerships with the so called "travel influencers" (Hernández-Méndez and Díaz, 2023). Online travel agencies provide instant access to hotels, flights, experiences and other tourism services, greatly simplifying the travel planning process. In addition, travelers now not only plan their trips online, but often make decisions based on reviews, recommendations, and user-generated content on social media.

Simply put, tourism today is more than just a movement of people. It is without doubt a major source of revenue for local and national governments administering territories with high tourist appeal. But it is also a transformative force, often generated by uncontrolled influxes of visitors, that can undermine the environmental, cultural, economic and social sustainability of cities and areas that are not adequately prepared to receive and host sudden quantities of visitors. Tourist overcrowding can lead to problems such as congestion of roads (especially at particular times of the year) resulting in increased pollution, the need to safeguard the biodiversity of natural habitats, general price increases, and loss of cultural authenticity. For many tourist destinations (particularly many of the most popular Italian ones) today we speak of overtourism rather than tourism when the damages to society and the environment outweigh the economic benefits that tourist flows bring. These destinations therefore need appropriate policies that make tourist presences sustainable, particularly during peak periods when the impact is most severe (emblematic is the media attention given to the surge in tourist numbers in July and August 2024 in the Dolomites UNESCO World Heritage Site).

Over the past decade, the literature has offered a wide range of proposals to address overtourism. However, many of these suggestions relies on theoretical frameworks or visionary approaches that often conflict with the practical needs of local governments whose perspective is to address overtourism as a matter of urgency. This requires strategies that are not only actionable and quickly implementable, but also capable of delivering tangible short-term results, thereby enhancing the policy effectiveness of local government who are committed to maintaining the quality standards of the main public utility services.

In general, the optimal strategy to counteract overtourism is based on changing the way tourists make their choices, both in terms of destination selection and intra-destination itineraries and activities (this holds true except in cases where the choice of a destination intrinsically determines behaviors within the destination, as often happens with cruise tourism). Attempting to influence the way individuals make choices requires a deep understanding of their decision-making processes, that is the factors they consider and the reasoning that guides their behavior. This is especially true in the context of tourism, where choices are increasingly influenced not only by personal preferences and cultural interests, but also - as stated before - by news, trends, and feedback driven by social media platforms.

In this paper, we aim to explore the decisions and behaviors that tourists adopt once they have selected a destination. Unfortunately, there are no datasets specifically designed for this purpose. As a result, it becomes necessary to rely on data sources originally designed for other purposes. For this reason, we argue that valuable insights can be extracted from the data collected through the use of a destination card (today the main cities of art offer one) and its related mobile app. We believe that this type of data can offer a meaningful understanding of how tourists behave, thus leading to the development of effective information tools to influence and potentially reshape their visiting choices. As we will elaborate later, some preliminary attempts in this direction have already been made.

In particular, we focus our attention on the city of Florence and its urgent need to develop reliable policies to manage and mitigate the impact of large tourist flows in its historic center, declared a World Heritage Site by UNESCO in 1982. These policies must be sustainable, concretely actionable and result-oriented: Florence is a city with unique characteristics in the world, with a dense concentration of world-renowned art and architectural masterpieces located within a radius of just under one kilometer. This compactness clearly amplifies the challenges in visitor management. The official destination card of the city of Florence is called FirenzeCard (https://firenzecard.it/en). FirenzeCard is a 72-hour museum pass that provides access to more than 60 museums and exhibitions within the city and surrounding areas. The card collects the sequences of museums visited by its holders (i.e., which museums were visited, when and in what order), thus being a repository of valuable information about tourists' visiting behaviors and preferences and their movements across the city. Thus, the research questions that drive this study are:

- (i) Can this type of data truly be used to understand intra-destination tourist behavior?

- (ii) For the case of the city of Florence, for example, does the sequence of museum visits recorded through the FirenzeCard project represent a seemingly limited dataset, or is it hiding valuable information that could be leveraged to understand and possibly influence the tourist decision-making process?

We show how the analysis of visit sequences recorded through the FirenzeCard can provide a positive response to these research questions, offering valuable informational support to policymakers responsible for urban governance and the development of sustainable strategies to mitigate tourist overcrowding. Specifically, we propose an innovative application of Item Response Theory (IRT; Bartolucci et al., 2015; de Ayala, 2009; Hambleton et al., 1991) models to investigate the sequences of visited museums and, relying on a latent class formulation (Goodman, 1974; Lazarsfeld and Henry, 1968), we identify clusters of tourists who exhibit similar behavior in terms of museum visits (how many and which museums to visit). This allow to characterize museums and exhibitions in terms of tourist preferences and could lead to the development of dedicated (rather than generic) strategies to meet the diverse cultural interests of different tourist segments.

The rest of the article is organized as follows. Section 2 provides some background on the concept of overtourism, with a focus on the city of Florence. Section 3 describes in details the information collected by FirenzeCard project; Section 4 illustrates the statistical approaches with related results to address our research questions. Finally, some conclusive remarks are provided in Section 5.

2. Backgrounds

This section begins by conceptualizing the term "overtourism" and examining its implications from the perspective of a local government. We also introduce the rationale behind selecting Florence as the case study for our research.

2.1. Overtourism

The phenomenon of tourist overcrowding is well synthesized in the neologism "overtourism" that was coined less than ten years ago. Coca-Stefaniak et al. (2016) used the term "overtourism" referring to both congestion and infrastructure failure and the rise of resistance and protest against tourism among marginalized and displaced inhabitants of major tourist cities such as Paris, London and Bangkok. But many of the concepts that contribute to the definition of "overtourism" have been known for at least 60 years (Koens et al., 2018). Forster (1964), in a pioneering study, was among the first to analyze the impact of tourism on societies, including the economic, social and cultural changes that occur as a result of increased tourism activity.

In recent years, an increasing number of scholars have devoted considerable research efforts to analyzing the determinants, implications and consequences of the phenomenon of overtourism, exploring its relationship with the concept of quality of life and the principles of sustainability. Without claiming to be exhaustive and only for the sake of parsimony, only a few of them are summarized here to represent a very wide range of approaches and proposals. Higgins-Desbiolles et al. (2019) propose a conceptual consideration of issues of degrowth in tourism. They argue that a sustainable degrowth will require greater attention on the rights of local communities and a rebuilding of the social capacities of tourism. Milano et al. (2019) examine the intersection between overtourism and degrowth through the lens of social movements: the negative impact of overtourism on the quality of life of residents and on the local environment has given rise to grassroots movements advocating degrowth, i.e. pushing for sustainable tourism practices that prioritize social and environmental well-being over economic growth. Mihalic (2020) conceptualizes overtourism within the context of sustainability, exploring how uncontrolled tourism growth negatively impacts the quality of life of residents and the visitor experience, and proposing a model that can help in identifying possible risks of unsustainable tourism situations. Gretzel (2021a,b) in her works proposes an idealistic approach that encourages future-oriented thinking in the tourism development. This perspective highlights the philosophical aspect of implementing smart technologies in tourism destinations, emphasizing that smart tourism should not just solve current problems, but imagine and work towards a better and more integrated future for both tourists and residents. The presence of reviews and bibliometric studies on the subject is also a clear demonstration of how overtourism is considered both relevant and worrying (Capocchi et al., 2019; Dodds and Butler, 2019; Santos-Rojo et al., 2023, this last review is limited to the years 2018-2021)highlighting the urgency of responsible approaches to managing tourism growth and its negative consequences on destinations and local communities.

2.2. Overtourism from a local perspective

Literature also reveals a particularly wide range of proposals to address overtourism in specific contexts, such as cities or regional areas. Once again for the sake of parsimony, we chose to cite the work of Bastidas-Manzano et al. (2021) who conduct a bibliometric analysis to examine a wide set of smart proposal for tourism and their technological evolution, identifying key themes, authors, and trends. This study can be considered a milestone in understanding how the concept of smart tourism has evolved over time and the directions it may take in the future. However, as outlined in the Section 1, a significant number of such proposals are based on theoretical frameworks and visionary approaches that often conflict with the needs of local governments, whose perspective is to address overtourism as a matter of urgency, making it essential to implement strategies that are concretely and rapidly implementable, as well as results-oriented within a short timeframe. Hospers (2019) lists several readily actionable coping strategies. These include promoting tourism in less-visited areas, implementing tourist taxes to fund infrastructure improvements, and encouraging off-season travels to distribute tourist numbers more evenly throughout the year. Additionally, he emphasizes the importance of involving local communities in tourism planning and decision-making processes to ensure that the benefits of tourism are balanced with the needs and well-being of residents. Some of these strategies could mitigate the managing efforts of mayors of major cities of art who are faced in maintaining the quality standard of core utilities (health, transport, waste disposal) during peak tourist seasons.

This is especially true for almost all of the most famous Italian cities of art which boast a dense concentration of artistic and architectural masterpieces generally located within limited administrative borders. This amplifies the challenges of local authorities in tourism management. Emblematic are the cases of Venice and Florence: a recent study addressing overtourism within the frame of urban liveability through a proxy analysis of tourism-relevant indicators for major European tourist cities finds that Venice is the city with the highest degree of overtourism, immediately followed by Florence (Amore et al., 2020). In these two specific cities, measures like real-time crowd monitoring, predictive analytics, and dynamic pricing tend to be ineffective. UNESCO-protected area of the two cities is crowded with tourists nearly all year, and the concept of high and low season applies only marginally to hotel pricing strategies, with minimal price differences between seasons. Although hotels in peripheral areas have implemented dynamic pricing strategies for years, most tourists still prefer to accommodate near the center to optimize their visiting time. Additionally, museums are not able to adopt dynamic pricing due to a lack of a unified coordination: some collections are privately owned, while public collections are managed either by the central government or the local government.

See Seraphin et al. (2018), Ignaccolo et al. (2020), and Bertocchi et al. (2020) for recent studies about the effects produced by overtourism in Venice. As anticipated in the introduction, in this paper we decided to focus our attention on the city of Florence and its urgent need to develop effective sustainable policies to manage and mitigate the impact of large tourist crowds in its historic center.

2.3. The case of the city of Florence

Initially called "Florentia", Florence was founded a few decades B.C. by the Romans as a fortified settlement strategically located at the confluence of the Arno and Mugnone rivers. Over the centuries, its importance grew to become a flourishing commercial and cultural center. Today it is celebrated as the birthplace of the Renaissance. Its historic center, declared a UNESCO World Heritage Site in 1982, allows to contemplate many world-renowned masterpieces of art and architecture within a radius of just under 1 km. The artistic and cultural importance of Florence, together with its historical and architectural heritage, make it one of the few living museums in the world. As a main consequence, nowadays Florence has to face significant challenges due to overtourism. Namely, the influx of tourists has led to overcrowded city center streets, increasing pollution and noise and rising living costs pushing residents to relocate in peripheral area. These macro-critical issues give rise to collateral problems such as the "foodification" of the historic center, that is its transformation into a food-dominant retail space to the detriment of the typical and traditional businesses which tend to disappear (Loda et al., 2020), as well as the "hostification" of the flats and apartments into the historic center, due to the rapid spread of short-term rental platforms. Such platforms can not be considered simply as a concomitant factor of the overtourism (Celata and Romano, 2020); data about published rentals are crucial to understand how and in which areas critical issues produced by the influx of tourists may arise (Bacci et al., 2020). Tourist illegal accommodation is a phenomenon clearly emerged in recent years, which involves serious critical issues and needs decisive coping housing policies that had to be accompanied by the privilege of creating new hotel facilities to be allocated in suburban or rural areas to support sustainable tourism goals (Shirvani Dastgerdi et al., 2021). Moreover, key cultural sites like the Accademia Gallery, the Uffizi Gallery and the Cathedral of Santa Maria del Fiore experience extreme visitor pressure almost all year round, with very long queues of tourists waiting for their visit in the streets and squares facing these museums and monuments since the early hours of the morning. Sustainable tourism strategies, such as the "Uffizi Diffusi" project, are aimed to mitigate these issues by dispersing tourists and promoting regional attractions (EHT, 2021). Thus, protecting Florence from overtourism ensures the city's historical and cultural integrity while enhancing the quality of life for residents and visitors alike. As anticipated in the introduction and discussed in the previous section, Florence has unique characteristics that make the implementation of coping strategies to mitigate the impact of overtourism essential and urgent. These strategies must be sustainable, concretely implementable and focused on achieving measurable results. For these reasons, since 2006 the Florence Management Plan and its subsequent updates have consistently prioritized the issue of mass tourism, and several research projects (De Luca et al., 2020) have either originated from or been incorporated into the Plan and its updates. In order to assess the exposure to the risk of overtourism, a system of indicators has been developed that measures the carrying capacity of the city and supports local decision-makers in preserving its heritage (Liberatore et al., 2023).

3. The FirenzeCard database

As outlined in the introduction, our objective is to investigate the decisions and behaviors that tourists adopt after selecting a destination. Given the absence of datasets specifically designed for this purpose, it is necessary to rely on data sources originally intended for other functions. Our specific focus is to assess whether meaningful insights can be extracted from data collected through the use of a destination card and its associated mobile app. Focusing our attention on the city of Florence, we refer to data collected as part of the FirenzeCard project (https://firenzecard.it/en). The FirenzeCard is the official 72-hour museum pass for the city of Florence, granting access (often on a priority basis) to 39 museums and other permanent exhibits (collectively referred to as "museums" from now on) within the city and its surrounding areas. The card collects the sequences of museums visited by its holders (i.e., which museums were visited, when and in what order), thus being a vast repository of valuable digital footprints left by tourists while they move across the city.

Figure 1 shows the location of the museums adhering to the FirenzeCard circuit on the map. It is worth noting that almost all the museums of the circuit, as well as many world-famous masterpieces of art and architecture, are included in a circumference of radius 1 Km centered in the forum of the ancient Roman city of Florentia. The card must be activated at the beginning of the first visit and can be extended for additional 48 hours by purchasing the FirenzeCard Restart extension. At each entry, visitors have to show their card, which is swiped on an optical reader, thus the sequence of museums visited and the corresponding entry times are recorded (the card does not allow multiple visits to the same museum). At first glance, such dataset seems quite poor in terms of useful information for the purposes of the research set out in the introduction. In fact, the available information is structured within a relational database, where the main entities include the sold cards — each associated with attributes that categorize tourists by gender, age, and country of origin — the museums — described through attributes such as location, collection type and size — and the sequences of visits recorded for each card. As previously mentioned, we believe that this information could be appropriately exploited to classify tourists according to their preferences for specific types of collections, as well as to categorize museums according to their frequency of selection.

To adequately address our research questions, we consider it sufficient to analyze data from a single year of card sold, thereby limiting the vast amount of information collected through the FirenzeCard project. Specifically, we focus on the visiting behaviors associated with all cards sold in 2018, resulting in a dataset of $125,092$ sequences of visited museums (for a total of $846,510$ visits). Data collected in 2018 might appear a bit outdated: it is worth noting that the last two pre COVID19-pandemic years (2018 and 2019) are those in which the city of Florence reached a record high of around 11 million tourist presences per year. In 2020, tourist presences have plummeted, but recovery began in 2021, with approximately 3.1 million visitors recorded for the entire year, increasing to 7.4 million in 2022. In 2023, there was a further rise to 8.9 million, still not reaching pre-pandemic levels (Città Metropolitana di Firenze, 2024). Thus, FirenzeCard

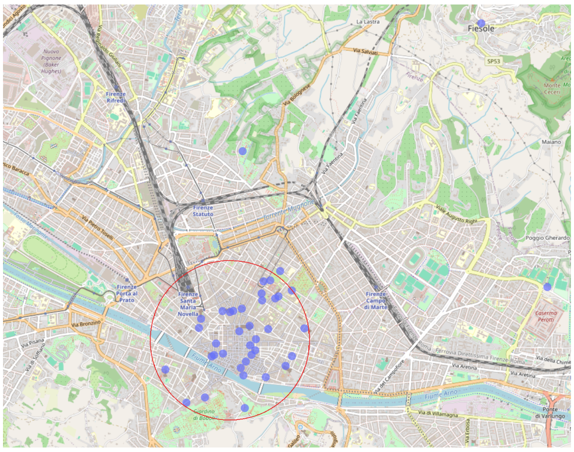

Figure 1: Museums of the FirenzeCard circuit with, superimposed, the circumference of radius 1 Km and centered in the ancient Roman forum of the city.

database referring to 2018 is one of the richest available in terms of cards sold and variety of visiting behaviors: these sequences of visits were deemed more than adequate to comprehensively represent the phenomenon under study. Figure 2 shows the amount of cards sold per month during 2018. The trend appears clearly consistent with the variation in tourist presence in the city over the course of the year: card sales remain substantially constant from March to October (with the exception of peaks observed in April and May) and decrease significantly in Winter.

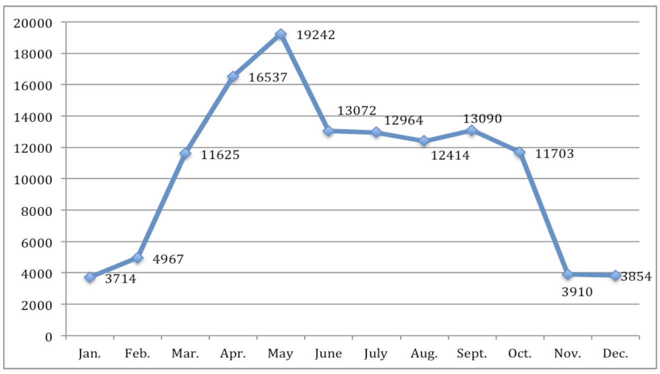

Figure 2: Number of Firenzecards sold in year 2018 per month

Table 1 provides a synthesis of the distribution of the number of museums visited by the FirenzeCard holders (i.e., the length of the sequence of visits recorded by each card). It is evident that the database contains sequences of very different lengths: there are users who visit only a few museums, while others make full use of the available time guaranteed by the card by visiting as many museums as possible. In average, a user visits 6.8 museums (median: 6), with the 25% of users visiting at most 4 museums and the 75% of users having made a maximum of 9 visits. The minimum number of visited museums is 1 and the maximum is 31, both observed only in 9 cases.

Table 1: Summary of the distribution of the number of visited museums: minimum, quartiles (q1, median, q3), maximum mean, and standard deviation.

Min	q1	Median	q3	Max	Mean	St.Dev.
1	4	6	9	31	6.8	3.2

4. Statistical methods and results

This section presents the statistical analysis conducted to address our research questions, along with a discussion of the corresponding results. The analysis is structured in two parts: first, a network analysis of the sequences of visited museums; second, the methodology used to cluster FirenzeCard users into homogeneous groups based on their visiting behavior and choice patterns.

4.1. Network analysis

The sequences of museums visited contribute to the construction of a network, the analysis of which is useful for identifying the most frequently chosen tourist paths. As an example, Figure 3 presents a conditional analysis of the choices made by all tourists who visited the Galleria degli Uffizi. The square node represents the origin of the network — that is, the last visited museum — while the circular nodes indicate the next choice. The size of the nodes is proportional to the corresponding choice rates. Specifically, the size of the square node reflects the proportion of FirenzeCard users who visited the Uffizi during their stay in Florence, while the size of each circular node represents the percentage of tourists who visited a given museum immediately after the Uffizi.

Generalizing the reasoning, all the exhibits and museums that are part of the

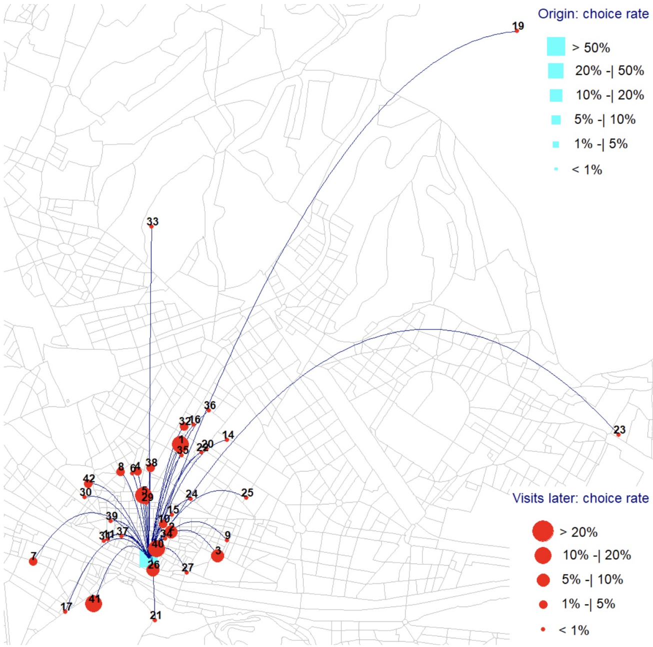

Figure 3: Museums visited after Galleria degli Uffizi by FirenzeCard holders in 2018

FirenzeCard project define the nodes of a directed network whose edges represent the observed paths taken by tourists between museums. For general definitions and techniques related to network analysis, see, among others, Kolaczyk (2009). The number of tourists who follow a given path determines the weight of each edge. Analyzing the distribution of weighted edges in increasing order, we find that the top 10% of edges account for 79% of the total flow between pairs of nodes, while the top 25% account for 94% of flows. These high-weight edges include paths between major attractions such as the Galleria degli Uffizi, Opera del Duomo, Galleria dell'Accademia, Palazzo Vecchio, and the Pitti-Boboli complex (in both directions), followed by the Basilica di San Lorenzo and the Medici Chapels. Overall, the resulting network is highly dense, with a density coefficient — defined as the ratio of observed edges to all possible edges — of 0.849. It is also highly transitive, with a transitivity coefficient — measuring the proportion of connected node pairs that close into triangles — of 0.953. Moreover, tourist flows exhibit strong bidirectionality, as indicated by a reciprocity coefficient (i.e., the ratio of mutual dyads to total dyads) of 0.905. Additionally, 62.7% of all triads form complete bidirectional triangles, highlighting the interconnected and reciprocal nature of tourist movement within the network.

Table 2 reports the list of the museums in the FirenzeCard circuit ordered by number of received visits. For each museum, the table also provides the "hub" and "authority" scores that serve to quantify prominence in the two roles. This idea was proposed by Kleinberg (1999) and developed by him into a centrality algorithm called "hyperlink-induced topic search" (HITS) in the context of web search. Applying these concepts to our case study, a "hub museum" is one whose visitors tend to visit many highly "authority museums" afterward. On the contrary, an "authority museum" is one that attracts visitors who have already visited several prominent "hub museums".

Table 2: Museums adhering to the FirenzeCard circuit: number of visits, hub and authority scores

Museum	n. of visits	Hub	Authority
Uffizi Gallery	104051	0.979	0.964
Accademia Gallery	99909	1.000	0.808
Opera del Duomo	94021	0.708	1.000
Palazzo Vecchio	72703	0.732	0.720
Pitti - Boboli	68607	0.749	0.488
Basilica of Santa Croce	54417	0.585	0.442
Basilica of San Lorenzo	47453	0.420	0.364
Medici Chapels	44278	0.409	0.361
Santa Maria Novella	43980	0.359	0.350
Bargello National Museum	35313	0.374	0.323
Medici Riccardi Palace	33360	0.334	0.289
Galileo Museum	31462	0.382	0.269
San Marco Museum	20803	0.239	0.181
Laurentian Medici Library	16289	0.145	0.126
Dante's House	12038	0.130	0.103
Brancacci Chapel	10539	0.100	0.086
Strozzi Palace	6936	0.057	0.056
Museum of the Innocents	6698	0.061	0.053
Bardini Villa	6654	0.065	0.054
Archaeological Museum	4788	0.045	0.034
Casa Buonarroti	4360	0.037	0.034
Novecento Museum	3143	0.025	0.022
Salvatore Ferragamo Museum	3089	0.029	0.023
Davanzati Palace	3075	0.027	0.023
Synagogue and Jewish Museum	2982	0.028	0.022
NSM - Botanical Garden	2356	0.020	0.015

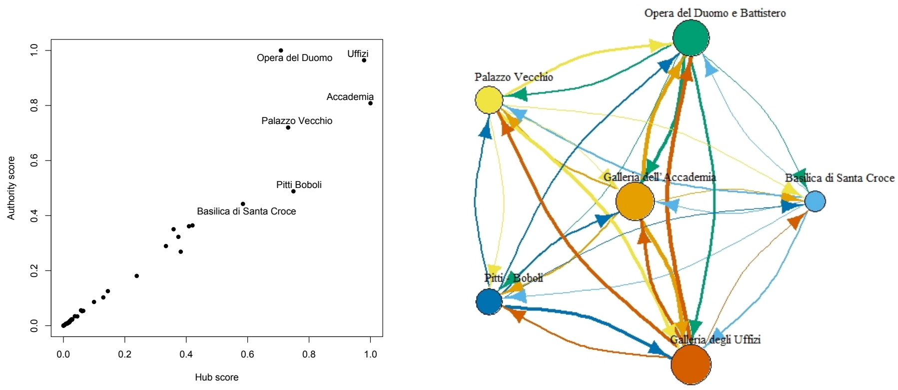

Figure 4: Hub and Authority scores (left panel) and network of museums with medium-high hub score (right panel)

Continuation of Table 2			
Museum	n. of visits	Hub	Authority
NSM - La Specola	2124	0.021	0.017
Opificio delle Pietre Dure	2046	0.021	0.015
NSM - Geology and Paleontology	1680	0.014	0.009
NSM - Anthropology	1584	0.015	0.011
Museum of the Misericordia	1314	0.019	0.012
Fiesole Civic Museums	1251	0.012	0.007
Zeffirelli Museum	1146	0.011	0.008
Horne Museum	947	0.008	0.007
Science and Technology Found.	429	0.003	0.002
Stibbert Museum	392	0.004	0.002
Casamonti Collection	135	0.001	0.001
Football Museum	97	0.001	0.001
Prehistory Museum	61	0.001	0.000

Thus, the hub score measures the importance of a museum based on the number of different museums that visitors choose to visit immediately after it in the analyzed visit sequences, while the authority score measures the importance of a museum based on the number of different museums that visitors visited immediately before it in these sequences. The two scores range from 0 to 1, indicating a spectrum from no importance to maximum importance.

Table 2 clearly identifies the iconic museums of Florence that attract the majority of tourists, that is, Uffizi Gallery, Accademia Gallery, and Opera del

Duomo museum system (which includes the Cathedral of Santa Maria del Fiore, the Brunelleschi's Dome, the Giotto's bell tower and the Baptistery of Saint John, as well as the annexed museum) with more than 90,000 visits each. The central role played by these monuments in the network of Florentine museums is well highlighted by the values of the hub and authority scores that are close to 1 for all the three museums. In particular, Uffizi Gallery and Accademia Gallery emerge as very influential starting point towards other museums (hub scores equal to 0.979 and 1.000, respectively); moreover, Uffizi Gallery and Opera del Duomo show the highest attraction capability (authority scores equal to 0.964 and 1.000, respectively). Unfortunately, the two scores also highlight the presence of relatively unknown museums (see also the left panel of Figure 4 for a graphical representation of these two scores). A synthesis of the above considerations is illustrated in Figure 4 (right panel), which focuses on the network formed by the six museums with the highest hub scores and their associated connections. The size of each node reflects the number of visits to that museum, while the width of each edge represents the volume of tourists who moved along the direction indicated by the arrow. Recalling Figure 1, almost all museums and exhibitions in Florence are located within the UNESCO area, where the leading roles are held by the iconic museums mentioned above. As a result, smaller collections within this area, as well as museums outside of it, are often under-visited, despite housing highly prestigious collections (e.g., the Stibbert Museum).

Other descriptive analyses (not reported here due to lack of space) on the length and content of the sequences immediately reveal the presence of groups of tourists with different behaviors. Among tourists who visit only a few museums, most limit themselves to the iconic ones before leaving, but some arrive with a clear plan of which museums they want to visit. A similar pattern is observed among the most active visitors: some visit many collections without choosing the iconic ones (perhaps because it is not their first time in the city), while others aim to maximize the value of their card by visiting all the most popular museums. These results suggest that it is possible to answer our first research question affirmatively.

4.2. Latent Class Item Response Theory models

To cluster FirenzeCard users in homogeneous groups sharing a common behavior in terms of visiting choices, we rely on the statistical framework of Latent Class Item Response Theory models (LC-IRT; Bartolucci, 2007; Bartolucci et al., 2015; Heinen, 1996; Langheine and Rost, 1988; Rost, 1990; von Davier, 2008) that represent an extension of the standard latent class model (Goodman, 1974; Lazarsfeld and Henry, 1968) to the Item Response Theory setting (Bartolucci et al., 2015; de Ayala, 2009; Hambleton et al., 1991). FirenzeCard holders distinguish for a different propensity to visit museums, which reflects on sequences of visited museums having a different length and a different content in terms of what museums have been visited. Formally, each sequence of visited museums can be represented as a sequence of J values Y_{ij}, with Y_{ij} binary variable equal to 1 when the museum j ($j = 1, \dots, J$) has been visited by user i ($i = 1, \dots, n$), and to 0 otherwise. Each sequence $Y_{i1}, \dots, Y_{ij}, \dots, Y_{iJ}$ may be considered as the individual manifestation of a latent trait Θ_i, denoting the *propensity to visit museums* by FirenzeCard user i: the greater his/her propensity, the more likely observing $Y_{ij} = 1$. The Item Response Theory provides a suitable framework to model data at issue. In virtue of its flexibility, we rely on a Two-Parameter Logistic (2PL; Birnbaum, 1968) parameterization ($i = 1, \dots, n;\ j = 1, \dots, J$)

$$logit\,[p(Y_{ij} = y|\Theta_i)] = \gamma_j(\Theta_i - \beta_j), \quad y = 0, 1, \tag{1}$$

where γ_j and β_j are, according to the IRT definition, the discrimination and difficulty parameters, respectively, associated to museum j. Both γ_j and β_j parameters contribute to assess the attractiveness of museum j: namely, they affect the probability of visiting museum j, conditional on the propensity level Θ_i. Usually, a parametric approach is adopted, assuming that the latent trait Θ_i is continuously (normally) distributed. However, since we are interested in clustering users based on their visit choices, we believe that a semi-parametric approach, based on the discreteness of the latent trait, is more suitable. Thus, we represent Θ_i as a discrete random variable that may assume a finite number of support points, $\xi_1, \dots, \xi_k, \dots, \xi_K$ with probabilities $\pi_1, \dots, \pi_k, \dots, \pi_K$, respectively. Each support point ξ_k identifies a latent class of individuals that share the same propensity to visit museums. Probability π_k denotes the weight of class k in terms of proportion of individuals belonging to this class. The resulting model configures as a Latent Class 2PL (LC-2PL) model. This type of model allows to estimate the probability of visiting a certain museum conditionally on the level of the propensity to visit museums.

The manifest probability of the response vector $\boldsymbol{Y}_i = (Y_{i1}, \dots, Y_{ij}, \dots, Y_{iK})'$ follows as

$$p(\boldsymbol{Y}_i = \boldsymbol{y}_i) = \sum_{k=1}^{K} \prod_{j=1}^{J} p(Y_{ij} = y|\Theta_i = \xi_k)\pi_k, \tag{2}$$

with $p(Y_{ij} = y|\Theta_i = \xi_k)$ defined as in eq. (1).

Model (2) is estimated maximizing the marginal log-likelihood function

$$\ell(\boldsymbol{\eta}) = \sum_{i=1}^{n} \log p(\boldsymbol{Y}_i = \boldsymbol{y}_i),$$

where $\boldsymbol{\eta}$ is the vector of free model parameters including γ_j and β_j; for the model identifiability we set $\gamma_1 = 1$ and $\beta_1 = 0$, as usual in the IRT setting. The maximization of $\ell(\boldsymbol{\eta})$ relies on the EM algorithm (Dempster et al., 1977) applied on the complete data log-likelihood, that is the log-likelihood that would be observed if the latent trait level was known for each class of individuals. The EM algorithm alternates two steps until convergence: (i) step E (expectation) consisting in the computation of the expected value of the complete log-likelihood, given the current values of parameters; step M (maximization) consisting in the maximization of the function obtained at the previous step with respect to the model parameters. As usual in the latent class and mixture models, the algorithm may stop on local maximum solutions; therefore, to contain this risk, it is advisable to run the algorithm several times using different starting values. The LC-IRT model was estimated through functions implemented in the R package `MultiLCIRT` (Bartolucci et al., 2012).

After the model parameters estimation, FirenzeCard users are then clustered in one of the latent classes according to the maximum a posterior probability criterion: all users belonging to the same class share the same propensity to visit museums.

It is worth to be noted that the number K of support points (the number of classes) is not a model parameter, but it has to be provided by the user. To select the optimal value of K we relied on the Akaike's Information Criterion (AIC; Akaike, 1973) and on the Bayesian Information Criterion (BIC; Schwarz, 1978):

$$\begin{aligned} AIC &= -2\hat{\ell} + 2 \cdot ||\boldsymbol{\eta}||, \\ BIC &= -2\hat{\ell} + ||\boldsymbol{\eta}|| \cdot \log(n). \end{aligned}$$

Both these indices are based on a penalization of the estimated log-likelihood $\hat{\ell}$ that accounts for the number of model parameters, thus small values of AIC and BIC are to be preferred.

To select the appropriate number of latent classes, we estimated different models ranging the number of the tentative support points from 1 to 5. Results are shown in Table 3; see also Figure 5.

Table 3: Selection of the number of latent classes (values of AIC, BIC, Δ_{AIC}, Δ_{BIC}).

	$K=1$	$K=2$	$K=3$	$K=4$	$K=5$
AIC	$2,700,623$	$2,547,562$	$2,532,160$	$2,528,177$	$2,527,242$
BIC	$2,701,003$	$2,548,331$	$2,532,948$	$2,528,985$	$2,528,070$
Δ_{AIC}	–	-5.65	-0.60	-0.16	0.04
Δ_{BIC}	–	-5.67	-0.60	-0.16	0.04

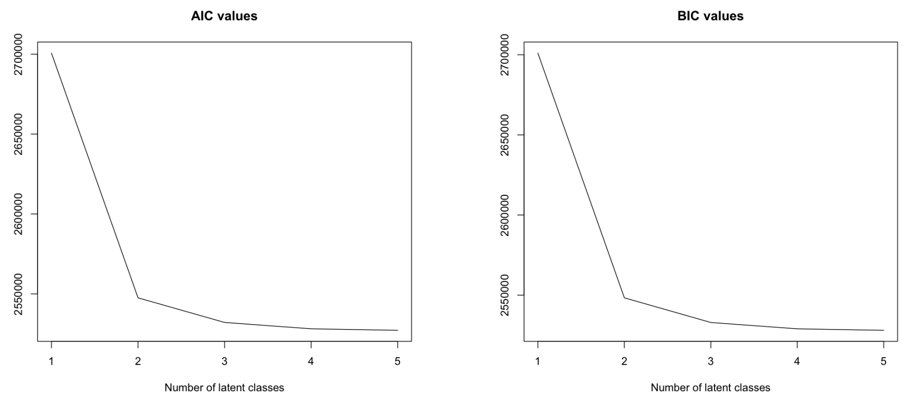

Figure 5: Trend of AIC (left panel) and BIC (right panel) indices

By examining the trend of AIC and BIC indices, we decided to select 3 support points. This choice avoids excessive aggregation of individuals, as it happens with 2 classes, and excessive partitioning, which occurs when the number of classes exceeds 3. Results of the selected LC-2PL model with 3 classes are shown in Table 4. The table reports the estimated levels of latent trait, the class membership probabilities, the average probability of visit any museum and the average number of museums visited by users, per latent class. The computation of the average number of visited museums is obtained after having allocated individuals in the three latent classes, relying on the maximum a posterior criterion.

Looking at the results, we observe that the three classes identified are characterized by incremental values in the propensity to visit museums. Specifically, over one-third of FirenzeCard users are clustered in Class 1, characterized by a 10% probability of visiting any museum and an average of 4.71 museums vis-

Table 4: LC-2PL model with 3 latent classes: estimated propensity to visit, class membership probability, average probability of visit any museum and average number of visited museums, per each latent class.

	Latent Class 1	Latent Class 2	Latent Class 3
Estim. propensity to visit	1.103	1.475	1.759
Class membership prob.	0.357	0.457	0.187
Average prob. of visit	0.100	0.184	0.289
Average visited museums	4.71	7.68	12.09

ited (median: 5, standard deviation: 1.29). Class 2 comprises more than 45% of users, exhibiting an intermediate propensity to visit, with an average probability of 18.4% and an average of 7.68 museums visited (median: 8, standard deviation: 1.75). Finally, the remaining 18.7% of users belongs to Class 3, showing an average visiting probability of 28.9% and an average of 12.09 museums visited (median: 12, standard deviation: 2.41).

Furthermore, for each museum, the adopted model allows to calculate both the marginal (i.e., unconditional) probability and the conditional probability of receiving a new visit given the class membership. Table 5 reports such probabilities for the top and bottom 5 museums and attractions (that is, for those who show an unconditional probability of being visited greater than 0.5 and less than 0.005 respectively). Focusing on the unconditional probabilities of visiting, we observe that three museums (i.e., Uffizi Gallery, Accademy Gallery, and Opera del Duomo) receive the majority of preferences, with a visiting probability exceeding 75 percent. Following behind, but at a significant distance, are Palazzo Vecchio and Pitti Palace (with the annex Boboli garden) with probability exceeding 50 percent. It is worth to be noted that a large number of museums are almost completely ignored: 20 of the 39 museums in the FirenzeCard circuit show a probability of being visited of less than 5%. Analyzing the conditional probabilities allows us to discern specific differences among the latent classes. Users belonging to Class 1 tend to concentrate their visits primarily on the Uffizi Gallery (conditional probability: 79.1%) and the Accademia Gallery (75.1%). Conversely, users in Class 3 exhibit a markedly different profile, showing interest in a wide range of museums. Notably, conditional probabilities associated with the Opera del Duomo and the Basilica of San Lorenzo exceed 90%, making them the top preferred destinations for users in Class 3, surpassing even the Uffizi Gallery and the Accademia Gallery.

Furthermore, museums such as Palazzo Vecchio and the Medici Chapels,

Table 5: Unconditional probability of visit and conditional probability of visit given the latent class.

Museum	Prob. of visit	Prob. of visit, conditional on latent class		
		Class 1	Class 2	Class 3
Uffizi Gallery	0.832	0.791	0.845	0.878
Accademia Gallery	0.799	0.751	0.814	0.853
Opera del Duomo	0.752	0.584	0.816	0.914
Palazzo Vecchio	0.581	0.343	0.661	0.842
Pitti - Boboli	0.548	0.407	0.590	0.717
Basilica of Santa Croce	0.435	0.176	0.502	0.766
Basilica of San Lorenzo	0.379	0.023	0.438	0.919
Medici Chapels	0.354	0.048	0.400	0.827
Santa Maria Novella	0.352	0.098	0.398	0.724
Bargello National Museum	0.282	0.125	0.309	0.518
Medici Riccardi Palace	0.267	0.054	0.283	0.635
Galileo Museum	0.252	0.230	0.257	0.279
San Marco Museum	0.166	0.035	0.165	0.421
Laurentian Medici Library	0.130	0.004	0.084	0.485
Dante's House	0.096	0.037	0.100	0.201
Brancacci Chapel	0.084	0.012	0.074	0.249
Strozzi Palace	0.055	0.021	0.057	0.117
Museum of the Innocents	0.054	0.007	0.045	0.163
Bardini Villa	0.053	0.027	0.056	0.096
Archaeological Museum	0.038	0.013	0.038	0.087
Casa Buonarroti	0.035	0.005	0.029	0.105
Novecento Museum	0.025	0.009	0.025	0.055
Salvatore Ferragamo Museum	0.025	0.015	0.026	0.040
Davanzati Palace	0.025	0.004	0.021	0.074
Synagogue and Jewish Museum	0.024	0.009	0.024	0.051
NSM - Botanical Garden	0.019	0.009	0.020	0.035
NSM - La Specola	0.017	0.012	0.018	0.025
Opificio delle Pietre Dure	0.016	0.003	0.014	0.046
NSM - Geology and Paleontology	0.013	0.007	0.014	0.024
NSM - Anthropology	0.013	0.007	0.013	0.021
Museum of the Misericordia	0.011	0.005	0.011	0.020
Fiesole Civic Museums	0.010	0.007	0.011	0.014
Zeffirelli Museum	0.009	0.005	0.010	0.016
Horne Museum	0.008	0.001	0.006	0.023
Science Tech. Found.	0.003	0.002	0.004	0.006
Stibbert Museum	0.003	0.002	0.003	0.004
Casamonti Collection	0.001	0.000	0.001	0.003
Football Museum	0.001	0.000	0.001	0.002
Prehistory Museum	0.000	0.000	0.000	0.001

which have lower probabilities in Class 1 (34.3% and 4.8%, respectively), boast probabilities higher than 80% in Class 3. Other museums with notably different conditional probabilities between Class 1 and Class 3 include the Basilica of Santa Croce (Class 1: 17.6%, Class 3: 76.6%), Santa Maria Novella (Class 1: 9.8%, Class 3: 72.4%), and the Laurentian Medici Library (Class 1: 0.4%, Class 3: 48.5%). It is worth noting that the Galileo Museum garners similar preferences across all three classes. Users in Class 2 occupy an intermediate position compared to the other two classes: the conditional probabilities increase for all museums from Class 1 to Class 2, continuing to increase until Class 3.

These results clearly demonstrate the potential of a database like the one underlying the FirenzeCard project to generate valuable insights that can be leveraged to understand — and potentially influence — the tourist decision-making process. This, in turn, reinforces the importance of developing effective tools capable of redistributing tourist flows toward lesser-known historical and cultural sites, thereby alleviating pressure on major museums and monuments while promoting a more balanced appreciation of the city's broader artistic heritage.

5. Conclusions

Overtourism has become an increasingly pressing issue in recent years, particularly affecting major tourist destinations that boast a dense concentration of artistic and architectural masterpieces, often confined within narrow administrative boundaries. The academic literature offers a wide array of proposals aimed at addressing overtourism in specific contexts, such as urban centers or regional areas. However, many of these proposals are grounded in theoretical frameworks or visionary approaches that frequently diverge from the practical needs of local governments. From the perspective of public administration, overtourism represents an urgent challenge that requires the adoption of strategies which are actionable, quickly implementable and capable of delivering measurable results in the short term. Mitigating the effects of overtourism requires interventions aimed at reshaping tourist decision-making processes, both at the stage of destination selection and in the planning of intra-destination itineraries and activities. Achieving this objective requires a deep understanding of the cognitive and behavioral mechanisms underlying tourist choices, including the criteria they consider and the reasoning that guides their behavior. This study aims to analyze the decision-making dynamics and behavioral patterns of tourists once a destination has been selected.

Generally, the quality of appropriate decision support tools determines the effectiveness of any policy action. For the case under study (intra-destination tourist behaviors), support must be found in the quantitative analysis of the (big) amount of data generated by tourist flows. Indeed, this data includes not only basic indicators like visitor numbers and origin/destination patterns but also the full range of digital footprints (Golder and Macy, 2014) left by these tourists, due to the significant transformation driven by advances in digital technology that the tourism sector has witnessed in recent years.

In this paper, we argue that all the footprints generated through the use of a destination card system can yield valuable insights into tourist behavior. Our case study focuses on the city of Florence, a highly relevant context given its spatial and cultural characteristics. The historic center hosts an extraordinary concentration of cultural heritage assets—artworks, monuments, and museums — within a very limited geographical area, of about one square kilometer. This spatial density, combined with sustained high visitor volumes, amplifies the urgency of developing evidence-based policies to regulate tourist flows and reduce pressure on key cultural sites. The official destination card of the city, known as FirenzeCard, is supported by an information system capable of recording the sequences of museum visits made by its users, including which museums were visited, in what order, and at what time. As demonstrated through the analysis conducted and illustrated in Section 3, it serves as a valuable repository of data on tourists' visiting behaviors, preferences, and movement patterns within the city. In particular, we propose an innovative application of Latent Class Item Response Theory (LC-IRT) models, which usually are used in the context of multiple item responses, thus enabling the clustering of tourists based on their propensity to visit museums. This latent trait is reflected in both the length and the composition of their museum visit sequences, indicating variability in the number and types of collections chosen. The identification of homogeneous groups of FirenzeCard users allows for the development of dedicated (rather than generic) strategies to meet the cultural needs of different types of tourists.

The sequences of museum visits provide valuable insights to understand intra-destination tourist behavior. This information could be further enriched with additional tourist-level and museum-level characteristics. For instance, at tourist level the data can be integrated with the geolocation of the FirenzeCard holder (obtained by integrating the associated mobile app with a basic geolocation function) and the estimates of visits duration, whereas at the museum level we could exploit museum characteristics such as the thematic content of collections, real-time information on queue lengths and visitor density, and the respective opening hours of each museum.

Integrating the available information with additional data sources opens the possibility of developing a recommendation system (Borràs et al., 2014; Sarkar et al., 2022) based on advanced statistical methods and machine learning algorithms. Such a system would capture - in real time - the preferences of tourists (in terms of content and characterization of the collections they wish to explore) and the criteria they adopt to choose museums to visit (in terms of the visiting priority given to different collections), even if these preferences have never been explicitly stated. The objective of this recommender system would align with broader municipal policy goals: to support strategies for redistributing tourist flows toward lesser-known cultural and historical sites, thereby alleviating congestion at major museums—particularly during peak periods—while optimizing individual itineraries based on tourists' preferences and time constraints. This approach should contribute to reduce overcrowding at iconic attractions and encourages a more balanced and comprehensive engagement with the city's cultural heritage.

References

Akaike, H. (1973). Information theory and an extension of the maximum likelihood principle. In B.N. Petrov and F. Csaki, eds., *Second International symposium of information theory*, 267–281. Akademiai Kiado, Budapest.

Amore, A., Falk, M., and Adie, B. (2020). One visitor too many: assessing the degree of overtourism in established european urban destinations. In *International Journal of Tourism Cities*, 6: 117–137. doi:10.1108/IJTC-09-2019-0152.

Bacci, S., Bertaccini, B., Dugheri, G., Galli, P., Giusti, A., and Sula, V. (2020). Short-term rentals in a tourist town. In A. Pollica, E. Brentari, N. Salvati, and F. Schirripa Spagnolo, eds., *Book of Short Papers SIS2020*, 1525–1530. Pearson, Milano.

Bartolucci, F. (2007). A class of multidimensional IRT models for testing unidimensionality and clustering items. In *Psychometrika*, 72: 141–157.

Bartolucci, F., Bacci, S., and Gnaldi, M. (2012). MultiLCIRT: Multidimensional latent class Item Response Theory models. R package version 1.0, URL http://CRAN.R-project.org/package=MultiLCIRT.

Bartolucci, F., Bacci, S., and Gnaldi, M. (2015). *Statistical Analysis if Questionnaires: a Unified Approach based on R and Stata*. Chapman & Hall, CRC Press, Boca Raton, FL.

Bastidas-Manzano, A.B., Sánchez-Fernández, J., and Casado-Aranda, L.A. (2021). The past, present, and future of smart tourism destinations: A bibliometric analysis. In *Journal of Hospitality & Tourism Research*, 45 (3): 529–552. doi:10.1177/1096348020967062.

Bertocchi, D., Camatti, N., Giove, S., and van der Borg, J. (2020). Venice and overtourism: Simulating sustainable development scenarios through a tourism carrying capacity model. In *Sustainability*, 12 (2).

Birnbaum, A. (1968). Some latent trait models and their use in inferring an examinee's ability. In F.M. Lord and M.R. Novick, eds., *Statistical Theories of Mental Test Scores*, 395–479. Addison-Wesley, Reading, MA.

Borràs, J., Moreno, A., and Valls, A. (2014). Intelligent tourism recommender systems: A survey. In *Expert Systems with Applications*, 41 (16): 7370–7389. doi:https://doi.org/10.1016/j.eswa.2014.06.007.

Capocchi, A., Vallone, C., Pierotti, M., and Amaduzzi, A. (2019). Overtourism: A literature review to assess implications and future perspectives. In *Sustainability*, 11 (12).

Celata, F. and Romano, A. (2020). Overtourism and online short-term rental platforms in italian cities. In *Journal of Sustainable Tourism*, 30: 1–20. doi: 10.1080/09669582.2020.1788568.

Città Metropolitana di Firenze (2024). Statistica del turismo. `https://www.cittametropolitana.fi.it/turismo-2/turismo/statistica-del-turismo/`. [Online; accessed 06-Nov-2024].

Coca-Stefaniak, A., Morrison, A., Edwards, D., Graburn, N., Liu, C., Pearce, P., Ooi, C., Pearce, D., Stepchenkova, S., Richards, G., So, A., Spirou, C., Dinnie, K., Heeley, J., Puczkó, L., Shen, H., Selby, M., bumm Kim, H., and Du, G. (2016). Editorial. In *International Journal of Tourism Cities*, 2 (4): 273–280.

de Ayala, R.J. (2009). *The Theory and Practice of Item Response Theory*. The Guilford Press, New York.

De Luca, G., Shirvani Dastgerdi, A., Francini, C., and Liberatore, G. (2020). Sustainable cultural heritage planning and management of overtourism in art cities: Lessons from atlas world heritage. In *Sustainability*, 12 (9).

Dempster, A.P., Laird, N.M., and Rubin, D.B. (1977). Maximum likelihood from incomplete data via the EM algorithm (with discussion). In *Journal of the Royal Statistical Society, Series B*, 39: 1–38.

Dodds, R. and Butler, R. (2019). The phenomena of overtourism: a review. In *International Journal of Tourism Cities*, 5: 519–528.

EHT (2021). How the Uffizi in Florence plans to deal with overtourism. `https://heritagetribune.eu/italy/how-the-uffizi-in-florence-plans-to-deal-with-overtourism/`. [Online; accessed 22-May-2024].

Forster, J. (1964). The sociological consequences of tourism. In *International Journal of Comparative Sociology*, 5 (2): 217 – 227.

Golder, S.A. and Macy, M.W. (2014). Digital footprints: Opportunities and challenges for online social research. In *Annual Review of Sociology*, 40 (Volume 40, 2014): 129–152. doi:https://doi.org/10.1146/annurev-soc-071913-043145.

Goodman, L.A. (1974). Exploratory latent structure analysis using both identifiable and unidentifiable models. In *Biometrika*, 61: 215–231.

Gretzel, U. (2021a). Conceptualizing the smart tourism mindset: Fostering utopian thinking in smart tourism development. In *Journal of Smart Tourism*, 1: 3–8. doi:10.52255/smarttourism.2021.1.1.2.

Gretzel, U. (2021b). Smart tourism development. In *CABI*, 159–168. doi:10.1079/9781789242812.0013A.

Hambleton, R.K., Swaminathan, H., and Rogers, H.J. (1991). *Fundamentals of Item Response Theory*. Sage Publications, Newbury Park.

Heinen, T. (1996). *Latent class and discrete latent traits models: similarities and differences*. Sage, Thousand Oaks, CA.

Hernández-Méndez, J. and Díaz, N. (2023). Influencer marketing in the promotion of tourist destinations: mega, macro and micro-influencers. In *Current Issues in Tourism*, 27: 1–11.

Higgins-Desbiolles, F., Carnicelli, S., Krolikowski, C., Wijesinghe, G., and Boluk, K. (2019). Degrowing tourism: rethinking tourism. In *Journal of Sustainable Tourism*, 27 (12): 1926–1944.

Hospers, G. (2019). Overtourism in European Cities: From Challenges to Coping Strategies. In *CESifo Forum*, 20 (03): 20–24.

Ignaccolo, C., Levenson, E., Phu, M., Saenger, L., and Zheng, Y. (2020). Unmasking Tourism in Venice. `https://www.overtourismvenice.mit.edu/`. [Online; accessed 21-May-2024].

Kleinberg, J.M. (1999). Authoritative sources in a hyperlinked environment. In *J. ACM*, 46 (5): 604–632. doi:10.1145/324133.324140.

Koens, K., Postma, A., and Papp, B. (2018). Is overtourism overused? understanding the impact of tourism in a city context. In *Sustainability*, 10 (12).

Kolaczyk, E.D. (2009). *Statistical Analysis of Network Data. Methods and Models.* Springer, New York, (NY).

Langheine, R. and Rost, J. (1988). *Latent trait and latent class models*. Plenum, New York.

Lazarsfeld, P.F. and Henry, N.W. (1968). *Latent Structure Analysis*. Houghton Mifflin, Boston.

Liberatore, G., Biagioni, P., Ciappei, C., and Francini, C. (2023). Dealing with uncertainty, from overtourism to overcapacity: a decision support model for art cities: the case of unesco whcc of florence. In *Current Issues in Tourism*, 26 (7): 1067–1081. doi:10.1080/13683500.2022.2046712.

Lindberg, K. and Johnson, R.L. (1997). The economic values of tourism's social impacts. In *Annals of Tourism Research*, 24: 90–116.

Loda, M., Bonati, S., and Puttilli, M. (2020). History to eat. the foodification of the historic centre of florence. In *Cities*, 103: 102746.

Mihalic, T. (2020). Conceptualising overtourism: A sustainability approach. In *Annals of Tourism Research*, 84: 103025. URL `https://www.sciencedirect.com/science/article/pii/S0160738320301699`.

Milano, C., Novelli, M., and Cheer, J.M. (2019). Overtourism and degrowth: a social movements perspective. In *Journal of Sustainable Tourism*, 27 (12): 1857–1875.

Rost, J. (1990). Rasch models in latent classes: an integration of two approaches to item analysis. In *Applied Psychological Measurement*, 14 (3): 271 – 282.

Santos-Rojo, C., Llopis-Amorós, M., and García-García, J.M. (2023). Overtourism and sustainability: A bibliometric study (2018–2021). In *Technological Forecasting and Social Change*, 188 (C).

Sarkar, J., Majumder, A., Panigrahi, C., Roy, S., and Pati, B. (2022). Tourism recommendation system: a survey and future research directions. In *Multimedia Tools and Applications*, 82: 1–45. doi:10.1007/s11042-022-12167-w.

Schwarz, G. (1978). Estimating the dimension of a model. In *Annals of Statistics*, 6: 461–464.

Seraphin, H., Sheeran, P., and Pilato, M. (2018). Over-tourism and the fall of venice as a destination. In *Journal of Destination Marketing & Management*, 9: 374–376.

Sharpley, R. and Telfer, D. (2002). *Tourism and development: Concepts and issues*. Channel View Publications, Clevedon, UK.

Shirvani Dastgerdi, A., De Luca, G., and Francini, C. (2021). Reforming housing policies for the sustainability of historic cities in the post-covid time: Insights from the atlas world heritage. In *Sustainability*, 13 (1). doi:10.3390/su13010174.

Smith, M. (2015). *Issues in Cultural Tourism Studies*. Routledge, London.

von Davier, M. (2008). A general diagnostic model applied to language testing data. In *British Journal of Mathematical and Statistical Psychology*, 61 (2): 287–307.