ISSN 1648-4142 / eISSN 1648-3480

TRANSPORT

2025
Volume 40
Issue 2
Pages 173–196
https://doi.org/10.3846/transport.2025.24650

Original Article

TOURISTS' LOCAL BUSES RIDERSHIP AND PANDEMIC RESILIENCE: A SMART CARD DATA ANALYSIS IN SOUTHERN CATALONIA

Aaron GUTIÉRREZ[1]✉, Leonardo MONTEIRO-FIALHO[2], Sergio TRILLES[2], Benito ZARAGOZÍ[1], Carlos GRANELL[2], Daniel MIRAVET[3]

[1] *Dept of Geography, University of Rovira i Virgili, Tarragona, Spain*
[2] *Institute of New Imaging Technologies, Jaume I University, Castellón de la Plana, Spain*
[3] *Consortium of Public Transport of Camp de Tarragona, Tarragona, Spain*

Highlights:

- data from smart travel cards is used to determine tourist profiles of public transport users in the Camp de Tarragona (Catalonia, Spain);
- a method for identifying and classifying profiles using smart card data, replicable in other case studies, is defined and validated;
- traveller profiles and their patterns in pre-pandemic and pandemic context are compared;
- traveller profiles with the greatest decline in public transport usage and those more resilient during the COVID-19 are identified and characterised;
- lessons are drawn for promoting tourists use of public transport during disruptive periods such as the COVID-19 pandemic.

Article History:
- submitted 7 December 2023;
- resubmitted 6 August 2024, 23 October 2024;
- accepted 5 November 2024.

Abstract. The COVID-19 pandemic's harmful effects have varied across economic sectors and been particularly adverse for the transport and tourism sectors. This article analyses the pandemic's impact on tourists' use of public transport since 2020, including its patterns of change and general decline, using data from more than 40000 smart card holders considered to be summertime users during the peak tourist season in Camp de Tarragona (Catalonia, Spain). 3 model-based clustering analyses of pre-pandemic data from 2019 were performed and used to classify data generated since the pandemic began in 2020. The 1st model included variables of each smart card's volume of activity, the 2nd model analysed the concentration or spatial dispersion of validated uses of each card, and the 3rd model examined the temporal dimension of the use of smart cards depending on the defined objective. Among the major findings, the number of journeys plunged by 92% in summer 2020 – that is, by far more than throughout the year (64%), which suggests a higher loss of travellers linked with tourism activities (e.g., tourists, 2nd-residence owners, and workers in the tourism sector). Regarding the spatial dimension, patterns with minor reductions related to trips taken within cities (45%) or between major cities (78%). By contrast, travellers with sprawled patterns fell the use by 93%. Last, profiles obtained from variables of a temporary nature presented similar percentages of losses; the most significant losses were for use distributed throughout the day (91.81%) and throughout the night (90.12%). This article provides valuable insights into the pandemic's varied effects on the use of public transport during peak season at a tourist destination, insights that could inform policies and actions to ensure a more robust response to future crises.

Keywords: COVID-19, public transport, tourism, smart card data, traveller profile, resilience.

✉Corresponding author. E-mail: *aaron.gutierrez@urv.cat*

Notations

- AIC – Akaike information criterion;
- AFC – automated fare collection;
- ATM – Territorial Mobility Authority (in Catalan: *Autoritat Territorial de la Mobilitat*);
- BIC – Bayesian information criterion;
- CSV – comma-separated value;
- LPA – latent profile analysis;
- SD – standard deviation.

1. Introduction

Developing adequate public transport networks at tourist destinations is essential not only to improving the attractiveness and competitiveness of the destinations (Prideaux 2000; Kim *et al.* 2023) but also to reducing the impact of tourism flows on residents' mobility and quality of life (Miravet *et al.* 2021a). For one, appropriate public transport networks improve tourists' mobility at destinations once they arrive by allowing them to forgo hiring cars or other sorts of motorised vehicles (Gutiérrez, Miravet 2016a). They also increase the number of attractions that can be visited (Leask *et al.* 2013; Zientara *et al.* 2024) and, as a result, make longer stays more likely (Miravet *et al.* 2021a). For another, tourists' use of public transport partly mitigates the impact of negative externalities associated with using private motorised vehicles (Domènech *et al.* 2023) and contributes to the dispersion of visitors across the destination and thus lessens crowding around major centrally located tourist attractions (Albalate, Bel 2010). Furthermore, increasing tourists' use of public transport instead of private motorised vehicles can help destinations to increase the urban space devoted to pedestrians, which affords a more pleasant visiting experience and higher levels of satisfaction among tourists (Ram, Hall 2018). More sustainable mobility is also pivotal to preventing tourists' potential dissatisfaction due to negative externalities stemming from the use of highly polluting modes of transport (Becken *et al.* 2017; Kim *et al.* 2023). Last, from the other direction, public transport can benefit from massive influxes of visitors due to the additional revenue created (Albalate, Bel 2010).

All these trends are far more likely to occur in the absence of a pandemic. Beginning in 2020, the COVID-19 caused a general halt in most economic sectors around the world, and its impact was particularly harmful for the tourism industry, owing to an initial shutdown of all tourism activity in many countries, bans on international travel, and, in some countries, even barriers to national travel (Gössling *et al.* 2021). Much like the pandemic's effects, the recovery of tourism activity across countries has been asymmetrical, for some destinations, especially mature mass coastal destinations, have been far more vulnerable than others to the pandemic's effects on tourism (Duro *et al.* 2021). Public transport has been also severely affected. During the pandemic's 1st months, mobility plummeted due to restrictions imposed by governments to deter the spread of the virus (Jenelius, Cebecauer 2020). Even when mobility began to recover, public transport was perceived as a vector of contagion (Abdullah *et al.* 2021), and the higher perceived risk of becoming infected led to a slower recovery in the use of public transport services than experience by other modes of transport (Eisenmann *et al.* 2021). The use of public transport in mass coastal destinations has been even more sensitive to the pandemic, not only because tourists' journeys are primarily for leisure (Delclòs-Alió *et al.* 2022), but also because residents may fear sharing transport services with visitors, who are generally also perceived as a vector of contagion (Vich *et al.* 2022).

Analysing pandemics and their effects on ridership on public transport is nothing new. In fact, such trends were previously assessed in the context of the SARS and MERS outbreaks (Kim *et al.* 2017; Lau *et al.* 2003; Wang 2014). That said, interest in the present contribution is justified for its insights into an airborne pandemic's impacts on the demand for public transport in a region where tourists play a key role in the configuration of mobility relationships shaped by seasonality (Gutiérrez, Miravet 2016b). To be sure, the COVID-19 pandemic has also exerted considerable pressure on the financial structures supporting public transport service providers (Gutiérrez *et al.* 2021), especially for companies operating in regions where tourists' demand for public transport services forms a substantial part of the annual revenue and where tourist arrivals are more sensitive to disruptive circumstances. In light of those trends, the response to declining demand for public transport could reduce the quality of transport services provided and, in turn, deteriorate destination image (Eusébio, Vieira 2013).

Data from smart travel cards used to purchase public transport services are especially useful for studying the incidence of disruptive events on the operation of such services (Kurauchi, Schmöcker 2017; Tang *et al.* 2016). Indeed, smart card data have already been used to explore the COVID-19 pandemic's effects on the demand for public transport (Almlöf *et al.* 2021; Mützel, Scheiner 2022; Jenelius, Cebecauer 2020; Gramsch *et al.* 2022). The chief advantage provided by such data is the data's total flexibility concerning the dimensions of time and space, which can afford an evolving image of the demand for public transport services (Bagchi, White 2005; Pelletier *et al.* 2011; Tang *et al.* 2019). Beyond that, such flexibility is particularly relevant in contexts where demand is unstable, including seasonal tourist destinations (Domènech *et al.* 2020; Miravet *et al.* 2021b).

This article distinguishes profiles of tourists who used public transport smart cards in 2019 in the Costa Daurada (Catalonia, Spain), a mature, highly popular coastal destination located in the Camp de Tarragona region (Catalonia, Spain). The profiles are constructed based on mobility behaviour during the peak tourist season in 2019 by applying clustering techniques to data from smart cards used by travellers. In turn, smart card data for 2020 are used to explore the effects on each profile identified in 2019. Altogether, the approach provides a clear picture of the effects on different segments of demand for public transport caused by both the crisis in tourist activity due to the spread of SARS-CoV-2 and the limited use of public transport due to fears of becoming infected.

In the rest of this article, Section 2 presents previous studies that are relevant to the present contribution, Section 3 describes the methods used in the study conducted for the article, and Section 4 describes the main findings. Section 5 discusses the results and their implications, after which Section 6 provides the article's conclusions.

2. Background

2.1. The COVID-19 pandemic and public transport

The COVID-19 pandemic has significantly altered the mobility of populations, primarily as a consequence of general constraints in daily activities due to travel restrictions and social distancing policies issued by national, regional, and local governments around the world (De Vos 2020; Wang *et al.* 2024). As a result of such policies and restrictions, the demand for all modes of transport plummeted during the 1st phase of the pandemic in spring 2020 (Anke *et al.* 2021). The subsequent recovery of traffic flows in the aftermath of the lockdown periods was slow for all modes of transport as well (Beck *et al.* 2020) and was marked by fears of contracting the disease while travelling (Campisi *et al.* 2022a, 2022b; Mogaji 2020). Even so, the decline in the use of public transport was more acute than for all other forms of transport (Jenelius, Cebecauer 2020; Eisenmann *et al.* 2021; Cheng *et al.* 2024), and its recovery was also far slower, to the point that while other modes of transport have gradually returned to or been close to recovering pre-pandemic levels of demand (Das *et al.* 2021; Zhang *et al.* 2023; Zaragozí *et al.* 2023), total ridership on public transport and the frequency of using it have remained below 2019 levels (Long *et al.* 2023).

The uneven loss of ridership between public transport and other modes of transport is due to several factors. 1st, from a demographic perspective, gender and especially age have been associated with a higher perception of risk stemming from using public transport (Böcker *et al.* 2023). 2nd, from a socioeconomic perspective, some segments of the population were forced to continue using public transport services (Lizana *et al.* 2024). Having a job requiring an in-person presence and the unavailability of private transport prevented the most disadvantaged sectors of workers from limiting their use of public transport (Jiao, Azimian 2021), and, for that reason, global mobility figures dropped, while preferences between modes of transport changed. Part of the demand for public transport has since shifted to other modes of transport, including private vehicles, despite the pernicious consequences of that modal shift for climate change and air quality. The perceived risk of contagion associated with using public transportation services was the principal driver of the reduction in demand (Barbieri *et al.* 2021; Cheng *et al.* 2024), and that persistent perception has become a chief obstacle to returning to pre-2020 ridership numbers (Abdullah *et al.* 2021; Tan, Ma 2021; Zaragozí *et al*. 2023; Wang *et al.* 2024). Although some works from the pandemic's initial phases revealed the potential of public transport services to be a vector of contagion for SARS-CoV-2 (Harris 2020; Shen *et al.* 2020), those conclusions were not supported by subsequent studies (Severo *et al.* 2021; Moreno *et al.* 2021; Hu *et al.* 2021). The underlying reason for the shift in findings lies in the rapid implementation of measures to prevent contagion on board vehicles (Hanaei, Rezaei 2020; Pradhan *et al.* 2020), measures that have proven their effectiveness at preventing new infections (Ku *et al*. 2021).

Knowledge of the COVID-19 pandemic's evolving impact on public transport ridership in relation to daily mobility is extensive. In fact, evidence from around the world allows comparisons of trends across countries. By contrast, despite an analysis of the pandemic's impact on long-haul travel (Abu-Rayash, Dincer 2020; Korinth 2020), few studies have investigated the pandemic's impact on public transport services at tourist destinations, even though such studies continue to be needed.

1st, the circumstances of tourists' use of public transport services are highly particular and have to be accounted for when analysing the pandemic's impact. After all, in the absence of tourists, the demand for public transport at destinations quickly evaporates. That trend warrants attention, tourism activity has been highly sensitive to the pandemic, and intentions to travel have been severely altered due to not only restrictions on travel and tourist mobilities but also changes in people's perceptions of the risks associated with travel. Avoiding travel altogether is a plausible response to situations when prospective visitors perceive risks associated with spending a holiday at certain destinations (Cahyanto *et al.* 2016). Along those lines, the pandemic amplified negative emotional reactions and perceived risks when planning holidays (Zhang *et al.* 2020) and at once generated subjective conjectures, formed in the collective imaginary, that may have been unrealistic (Lu, Atadil 2021). Even then, perceived risks in travelling often emerge at the individual level and vary from one person to the next depending on external and context factors and the characteristics of individual tourists (Neuburger, Egger 2021).

2nd, the use of public transport at destinations involves journeys to places and attractions that tourists visit during their stays and the journeys back to their accommodations. Tourists' behaviour at destinations and thus patterns in what they visited and modes of transport used to reach those places and attractions were seriously affected by the pandemic. Using data from a survey of the Swiss population's habits on holidays, Thao *et al.* (2024) confirmed the perceived health risk of using shared modes of transport due to the pandemic and their preference for private transport instead. Similarly, using survey and mobile phone data, Östh *et al.* (2023) observed a modal shift from public transport to private cars and micro-mobility when comparing leisure mobility in summer 2020 and summer 2021 with 2019. From another angle, Da Silva Lopes *et al.* (2021) found that the pandemic shortened the time devoted to visiting attractions and shrank the size of visiting areas in the city of Porto (Portugal). Other evidence indicates that private vehicles gained ground as an option for travelling during the holidays (Ivanova *et al.* 2021).

3rd and last, and in complement to the 2 other elements, travelling to tourist destinations and the mobility-oriented decisions made therein, as opposed to decisions in daily commuting journeys, are voluntary. Consequently, individuals might be likely to renounce, whether in part or in full, travelling to tourist destinations and within them in the event that the perceived risk of contagion is too great. In that sense, risk perception has been a cornerstone in examining tourists' decision-making processes under the threat of the pandemic (Rahman *et al.* 2021).

Those 3 elements, combined with a perceived higher risk of contagion associated with public transport services than with private transport, have caused the demand for public transport at tourist destinations to plummet (Delclòs-Alió *et al.* 2022), especially in urban contexts. The drop was particularly severe at coastal destinations characterised by a high density of visitors and a high dependency on international tourism, due to their greater vulnerability to the incidence of COVID-19 (Duro *et al.* 2021). However, despite the pandemic's apparent impact, its effects on the use of public transport services among tourists at destinations remain unclear. In response, studying how the pandemic may have affected tourists' use of public transport, and the determinants of that phenomenon would be valuable. Thus, in the study presented here, smart card data from 2019 and 2020 – the 1st year of the pandemic – were used to compare the use of public transport by tourists across time.

2.2. Profiling travellers using data from smart travel cards

In recent decades, public transport services have adopted AFC systems, which use smart travel cards and guarantee controlled access for users and do so quickly and agilely. AFCs also allow centralising and registering verified transport use. During the 1990s and 2000s, AFCs gained initial popularity and have since spread to major cities worldwide, with examples that now include the *Octopus Ccard* in Hong Kong, *Navigo* in Paris, the *Compass Card* in Vancouver, *Oyster* in London, *Bip!* in Santiago, the *Troika card* in Moscow, and *OV-Chip* in the Netherlands.

Research using data from smart travel cards has supported destination inference, origin–destination matrices, the estimation of demand, and studies on passenger behaviour and trip chains, among others (Cats 2024). Early on, Trépanier *et al.* (2007) used smart card data to determine the destination stops of bus passengers and estimate trip destinations and chains, while Munizaga & Palma (2012) later developed an approach to inferring alighting stops to construct origin–destination matrices. A cluster analysis of passengers based on spatial and temporal behaviour has also been conducted (Morency *et al.* 2007), and, more recently, Foell *et al.* (2015) developed probability models to predict daily bus usage, while Raveau *et al.* (2011) used smart card data to model travel choices. The estimation of the object of the travel activity has been studied as well (Devillaine *et al.* 2012; Kusakabe, Asakura 2014).

In a review on using public data from smart cards, Pelletier *et al.* (2011) pinpointed 3 ways of using data from AFCs in research: (1) strategic-level studies, including long-term network planning, passenger behaviour analysis, and demand forecasting; (2) tactical-level studies, including longitudinal studies oriented towards identifying patterns in travel behaviour in order to adjust transport services; and (3) operational-level studies, which focus primarily on indicators of supply and demand. Considering all 3 purposes, research actions towards analysing, segmenting, and better identifying the travel behaviour of public transport passengers seem to be the common denominator of studies using smart card data.

More recently, Ghaemi *et al.* (2017) classified studies on travel behaviour using data from smart travel cards into 3 domains: (1) studies on understanding the data, which often involves manipulating data to extract significant indicators of what is happening in the transport network analysed; (2) studies to explain travel behaviour, which necessarily implies using external data sources according to the objective and needs of the study; and (3) studies aimed at supporting decision-making, primarily to forecast demand and plan transport. Considering those 3 domains, this article falls into the 2nd and presents research conducted to determine profiles of travellers.

As stated by Zaragozí *et al.* (2021), data from smart travel cards present various opportunities for researchers to seize. For one, the data comprise the whole universe of public transport users in a specific area, in contrast to samples used in traditional surveys. For another, they enable analyses at different territorial and temporal scales because all travels reported are time-stamped and can be geo-referenced. This type of data additionally supports longitudinal studies at the individual level because each transaction is linked to a card. Beyond that, smart card data makes inter-annual studies possible, which allows examining the evolution of the demand for mobility and public transport at different spatial and temporal (or individual) scales. On the downside, smart travel card data also present some difficulties. On that count, the data are continuously collected and due to the possibility of representing large volumes may be regarded as a type of big data. Added to that, because AFC systems are created for collecting fares and managing regional fare integration zones, the data require substantial cleaning, processing, and enrichment in preparation for use in research. Last, sociodemographic data associated with each smart travel card are usually restricted, unavailable or of low quality.

The mentioned studies, including ones measuring the impact of the COVID-19 pandemic on public transport (Fernández Pozo *et al.* 2022; Jenelius, Cebecauer 2020), have focused on the daily mobility of residents and extracted trends in the use of public transport. By contrast, few studies to date have identified the pandemic's impact on tourists or specific groups of travellers among general users of a given public transport network. Nevertheless, the application of advanced classification analysis and the possibility of identifying tourist profiles based on the fare/card type and their behaviour by using smart cards opens the doors to illuminating the resilience of public transport and understanding the diverse behaviours of multiple tourist profiles.

3. Data and methods

3.1. Study area

The area examined in the study was Camp de Tarragona (Figure 1), which in 2021 had a population of 641923 inhabitants (Statistical Institute of Catalonia, *https://www.idescat.cat*). Spatially, most of the distribution of economic activity

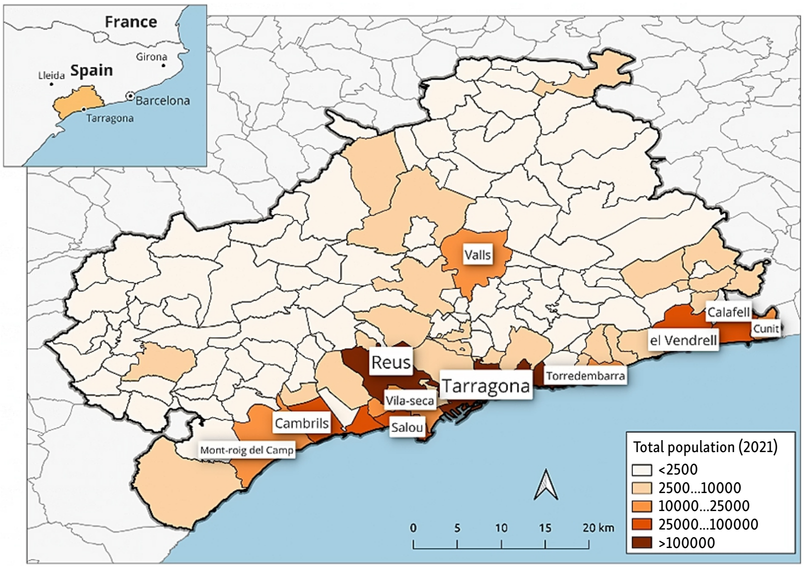

Figure 1. Map of Camp de Tarragona, the study area, by population

and population is concentrated along the coast, especially in the cities of Tarragona (i.e., with 135436 inhabitants) and Reus (i.e., with 106084 inhabitants), together with the 3 coastal municipalities of Cambrils, Salou, and Vila-Seca, with 35064, 28512, and 22522 inhabitants, respectively. Those municipalities form the main tourist destination in south Catalonia – the Costa Daurada – where more than 77% of all hotel accommodations in the region are concentrated, which in 2019 represented more than 20 million overnight stays in regulated accommodations. Camp de Tarragona is also a region where the impact of seasonal tourism on the demand for public transport is particularly evident. Studies in the area have identified the high use of public transport among tourists arriving by plane and train (Gutiérrez, Miravet 20216a). As a result of such use, in the municipalities with the most tourists (i.e., Salou, Cambrils, and Vila-seca), ridership on public transport usually increases sixfold during the summer (Domènech *et al.* 2020). Therefore, the studied area presents an ideal case for exploring the use of public transport by tourists.

Even Camp de Tarragona and beyond, the COVID-19 pandemic significantly disrupted the mobility of the population. Social distancing policies issued by governments, meaning a reduction in daily activities, and travel restrictions had significant impacts worldwide. In Spain, a national lockdown was imposed from 15 March to 13 May 2020, along with various restrictions imposed for several months afterwards. Across the European Union, national borders were closed until 21 June 2020, and only then reopened exclusively countries in the Schengen area, although numerous restrictions and quarantine policies were imposed depending on the traveller's country of origin. The cumulative incidence of COVID-19 in the 14 days prior to the end of the national lockdown was 27.93 per 100000 inhabitants; however, that number increased to 149.75 per 100000 inhabitants on 21 August 2020 (INE 2020). Concerning the study presented herein, it is worth noting the unconventional decrease of 74.5% in the number of visitors to tourist destinations in Spain during the summer of 2020 (Vich *et al.* 2022).

For the study, smart travel card data were obtained from the AFC system managed by the ATM of Camp de Tarragona, the public transport authority responsible for managing the integrated fare system in the studied area. That system, *Sistema de Gestió de la Integració Tarifaria*, stores information about the time at which passengers boarded vehicles, the location of the bus stop where they boarded, the bus line, and the type of transport fare paid (Gutiérrez *et al.* 2020).

Smart card data used in the study were limited not only to the summer (i.e., from 15 June to 15 September) in both 2019 and 2020 but also to the T-10 fare type. T-10 is the unique, multi-person fare in the ATM system that allows groups to travel in which each person uses the same card (i.e., consecutive transactions when boarding). The standard T-10 card covers 10 transactions and can be recharged as many times as desired, although no more than 30 transactions at a time are allowed. The T-10 card is the most-used card by visitors due to the flexibility that it offers and can also be used to identify different profiles of visitors and locals travelling by public transport in Camp de Tarragona (Gutiérrez *et al.* 2020). By contrast, the other transport fare options offered by the ATM system are intended to promote user loyalty by diminishing the unitary price of each journey by means of public transport. They impose some travel conditions that make them unattractive for visitors, including that cards cannot be used by groups because the fares are unipersonal and that they require a minimum number of trips in 30 days or across longer periods. It is also possible to travel by acquiring a single ticket, which is an attractive option for visitors. Single tickets, nevertheless, do not allow tracking users of public transport, for their data remain unconnected.

Table 1. Statistics of smart cards used in the summer (i.e., from June 15 to September 15)

Year	2019		2020	
Fare types	All	T-10	All	T-10
Number of cards issued	36478	34214 (93.79%)	6449	4562 (70.74%)
Number of transactions	656577	588709 (89.66%)	85719	48604 (56.70%)

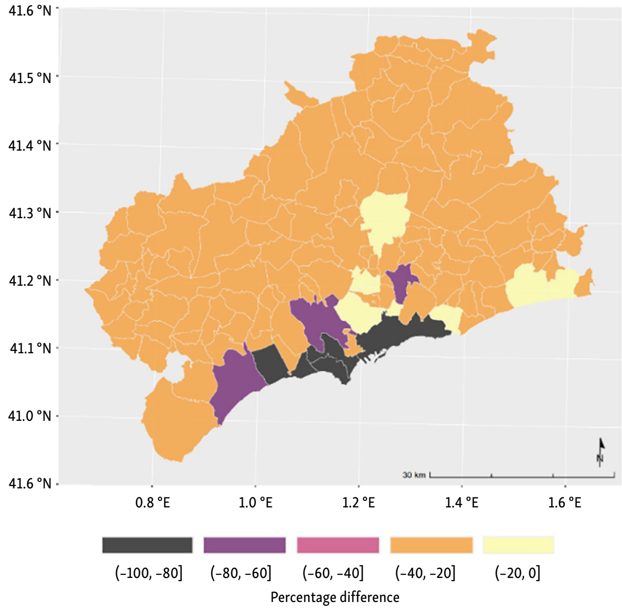

Figure 2. Difference in the percentage of transactions between 2019 and 2020

Table 1 shows global figures representing the COVID-19 pandemic's impact on public transport ridership between 2019 and 2020. The number of transactions made with T-10 cards was 10 times greater in 2019 than in 2020 (i.e., 588,709 transactions vs. 48604 transactions), and the percentage of T-10 card uses across all transactions nearly halved during the study period (i.e., 89.66% in 2019 vs. 56.70% in 2020). The number of T-10 cards issued also reveals a notable decrease from 34214 cards in 2019 to only 4562 cards in 2020. Last, the number of T-10 cards issued among all cards of any type issued fell from 93.79% in 2019 to 70.74% in 2020. Figure 2 shows the difference in the percentage of transactions between 2019 and 2020. A general drop affected all municipalities, although especially coastal ones.

3.2. System architecture

The resilience of the various profiles of public transport use against the pandemic's effects was examined through a series of methods embodied in the software architecture shown in Figure 3. The figure illustrates the analytical methods employed, the software tools and technology used, and the flow of smart card data from raw data to actionable, meaningful information through 3 conceptual layers: content, services, and applications.

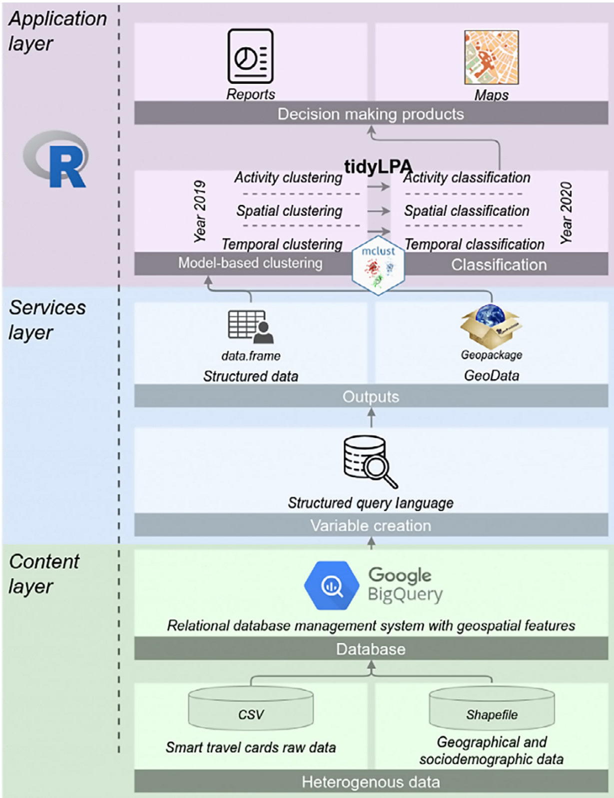

Figure 3. Software architecture for analysing data from smart cards

1st, the content layer was used for collecting and loading the data. After data from the ATM system were anonymised, raw smart card data in CSV format, along with geospatial datasets in ESRI shapefile format for spatial context, were ingested through loading scripts into a *Google BigQuery* database used as backed-as-a-service data storage system. Data models for tables and different views of the data were designed and prepared.

2nd, the service layer was used to oversee data wrangling and cleaning. A series of SQL queries in *Google BigQuery* were performed to compute a set of meaningful variables (see Section 3.3) in preparation for statistical analysis. In the service layer, structured data outputs were generated in the form of data frames and GeoData; the structured data outputs were essential for creating variables and further analysis. Both the structured data and GeoData prepared in the services layer ensured that the data were clean and ready for use in the analysis and classification tasks performed by the tidyLPA in the next layer (i.e., application layer).

3rd and last, the application layer focused on data analysis and visualisation. Using *R* scripts, the processed and structured data were retrieved from *Google BigQuery* to perform model-based clustering and classification for 3 groups of variables in the (1) activity, (2) spatial, and (3) schedule dimensions. Model-based cluster analysis applied to the 2019 data likewise differentiated between activities, the spatial dimension, and scheduling habits. The resulting clusters from the 2019 data were used to classify the

2020 data to compare the pandemic's impact on each of the profiles identified in 2019. Analyses were conducted in the *R* statistical language version 4.2.2 (R Foundation 2025).

3.3. Analysis

The analytical process involved pre-processing the data, filtering the data, creating variables, and classifying variables. Whereas the 1st 2 steps were briefly described in relation to the tools used in the system architecture in Section 3.2, the last 2 steps – creating and classifying the variables – become important especially when describing the analytical methods employed.

To create variables, 25 variables grouped into the activity, spatial, and schedule dimensions were extracted via a series of SQL queries (Table 2). Those variables afforded an overall image of the behaviour of each T-10 card in terms of activity level, spatial distribution, and scheduling habits.

Data classification was divided into 5 steps.

1st, variable selection was performed with attention to the process of selecting a feature subset, which can prevent redundancy. As detailed in Section 4, a set of correlation matrices were created to highlight the most correlated features (i.e., absolute correlation coefficient >0.75).

Table 2. Statistics of the smart cards used in the summer (i.e., June 15 to September 15)

Type	Name	Description (range)	Mean (*SD*)	Median [min, max]
Target	*card*	grouping variable (*N* = 34214)	–	–
Activity	*transactions*	total number of transactions (*N* = 588709)	17.21 (14.34)	11 [1, 19]
	avg_transactions	average number of transactions per day	5.11 (2.76)	4.50 [1, 38]
	active_period	number of days between the 1st and last day the card was used (card lifespan)	9.17 (12.95)	5 [1, 95]
	active_days	number of days the card was used	3.66 (2.99)	3 [1, 58]
	active_months	number of months the card was used	1.24 (0.49)	1 [1, 4]
	avg_group_size	average number of consecutive transactions in any stop	2.78 (1.45)	2.50 [0, 30]
	min_group_size	minimum number of consecutive transactions in any stop	2.19 (1.42)	2 [0, 30]
	max_group_size	maximum number of consecutive transactions in any stop	3.30 (1.85)	3 [0, 30]
	group_transactions	number of transaction chains with more than one transaction	14.95 (13.32)	10 [0, 17]
Spatial	*visited_municipalities*	number of municipalities visited during the entire period	2.76 (1.02)	3 [0, 8]
	used_routes	number of routes used during the entire period	1.89 (0.7)	2 [1, 10]
	main_municipality	percentage of transactions concentrated in the most-visited municipality	56.18 (18.46)	50.00 [0, 100]
	main_two_municipalities	percentage of transactions concentrated in the 2 most-visited municipalities	86.02 (14.87)	89.00 [0, 100]
	main_three_municipalities	percentage of transactions concentrated in the 3 most-visited municipalities	96.23 (8.04)	100.00 [0, 100]
	transactions_tarr_reus	percentage of transactions concentrated in the main cities – Tarragona and Reus – over 50000 inhabitants	28.87 (23.52)	4 [0, 77]
	transactions_cgc	percentage of transactions concentrated in the main touristic cities – Cambrils, Salou, and Vila-seca – between 20000 and 50000 inhabitants	69.60 (25.17)	70.00 [0, 100]
	transactions_urban_municipalities	percentage of transactions concentrated in the main cities over 10000 inhabitants	69.60 (7.73)	70.00 [0,100]
Schedule	*weekdays*	percentage of transactions on weekdays	76.67 (27.80)	83.33 [0, 100]
	weekends	percentage of transactions on weekends	23.33 (27.80)	16.67 [0, 100]
	first_half_day	percentage of transactions on the 1st half of day (7:00–16:00)	51.89 (29.19)	50.00 [0, 100]
	second_half_day	percentage of transactions on the 2nd half of day (16:00–21:00)	43.26 (27.28)	44.86 [0, 100]
	time_morning	Percentage of transactions concentrated in the morning (7:00–12:00)	27.00 (23.63)	25.00 [0, 100]
	time_midday	Percentage of transactions concentrated in midday (12:00–17:00)	31.45 (24.19)	31.03 [0, 100]
	time_afternoon	Percentage of transactions concentrated in the afternoon (17:00–21:00)	36.70 (26.80)	38.71 [0, 100]
	time_night	Percentage of transactions concentrated at night (21:00–6:00)	4.85 (11.80)	0.00 [0, 100]

Next, a meaningful subset was chosen based on the correlations between variables and on the capacity of a feature to describe smart card patterns and underlying models. Last, a new set of correlograms were plotted to compare the level of redundancy in the new subset.

2nd, the data from 2019 were clustered. Model-based clustering assumes that the observed data come from a mixture of distributions and that each group or class is described by a density function, usually a Gaussian distribution. It affords numerous advantages over other clustering methods, including the assessment of the number of clusters and an appropriate model. Finite mixture models, a model-based clustering approach, provide probabilistic clustering in which clusters correspond to model components (Hennig *et al.* 2015).

3rd, a LPA model, a type of finite mixture model, was used to classify the cards into different profiles. LPA is a type of latent variable analysis based on the assumption that the data originate from an unknown distribution arising from a mixture of simpler distributions. Those techniques are often referred to as "Gaussian mixture models", for they typically assume that the data distribution is a mixture of one or more clusters that can be described by normal distributions. The probability density function at each point is given by $f(y;\Psi)=\sum_{i=1}^{g}\pi_i \cdot f_i(y)$, in which y represents the observed data, and g denotes the number of component densities, which in the study was assumed to be multivariate normal components, $f_1(y)$, $f_2(y)$, ..., $f_g(y)$, mixed in unknown proportions, π_1, π_2, ..., π_n. The posterior probability that an observation y_j belongs to the ith component of the model is $\tau_i(y_j)=\frac{\pi_i \cdot f_i(y_j)}{f(y_j)}$ for i = 1, ..., g. Using the expectation–maximisation algorithm, the LPA model seeks maximum likelihood estimators for the parameters of the distributions and π_i. The result provides an estimate of the posterior probabilities that the observed y_j belongs to the ith component of the model – that is, to the cluster C_i for i = 1, ..., g. The *R* package *tidyLPA* (Rosenberg *et al.* 2018), an interface with the *mclust* package also in *R* (Scrucca *et al.* 2016), was used to apply LPA. The package *tidyLPA* not only uses a tidy interface (Wickham *et al.* 2019) but also facilitates the specification of models that are common to LPA.

4th, 3 classifications were made to better understand the features of cards regarding their different activity levels, spatial distribution, and scheduling habits.

5th and finally, the data from 2020 were classified. To illustrate the changes occurring to each of the identified profiles for 2019, the smart card data from 2020 were classified using the same models obtained for 2019. In that way, smart cards were classified by models that characterised their behaviour according to profiles previously identified in 2019, and a between-years comparison could therefore be easily performed.

4. Results

This section describes the clusters of the activity, spatial, and schedule variables. Based on those clusters, different profiles of travellers were detected, followed by an analysis of the evolution of the profiles between 2019 and 2020.

4.1. Activity clusters

1st, a correlogram showcasing the different activity-related features was plotted, as shown in Figure 4a. Given the high correlation grade between the variables, the following features were discarded for their redundancy (i.e., absolute correlation coefficient >0.75): *group_transactions*, *active_months*, *max_group_size*, *avg_transactions*, and *min_group_size*. In light of its limited informational value and skewed distribution – 78% of the data were valued at 1, 20% at 2, and the remaining at predominantly 3 with a few instances of 4 – the variable *active_months* was excluded from the

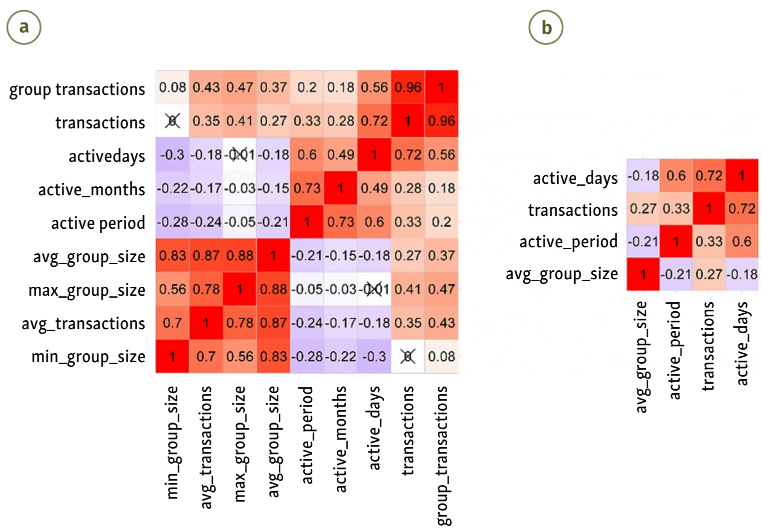

Figure 4. Correlogram showing self-correlation of activity variables: (a) – all variables; (b) – selected variables

analysis. Although its correlation was less than 0.75, the variable adversely affected clustering performance. The selected activity features appear in Figure 4b.

Following the feature selection, clusters were obtained using LPA. Table 3 shows the different models obtained and their fit indices. Based on the fits, the 3 latest models (i.e., with 5, 6, and 7 profiles) were the highest-performing, whereas the models with 6 or 7 profiles presented groups with less than 1% of the total cards and had the lowest minimum posterior probabilities, meaning that individual cards were not classified as accurately in their respective groups as in the model with 5 profiles. Therefore, the model with 5 profiles was chosen.

Table 4 presents descriptive statistics of the activity profiles; additional graphical representations of the clusters appear in Appendix. In general, T-10 cards were clearly used by groups of tourists and not long-lasting, for they had an average group size of 2.78 people and a card lifespan (i.e., *active_period*) of 9.17 days. For a better understanding of the characteristics of the profiles, each has been named accordingly as follows:

- Profile A1: Sporadic (*N* = 1783) had the highest score for card lifespan (i.e., *active_period*), a moderate number of active days, and a low number of transactions, thereby making the profile a good representative of residents and people who own a 2nd-residence in the area and travel sporadically for leisure;
- Profile A2: Continued (*N* = 494) had the highest number of transactions and active days, a high card lifespan (i.e., *active_period*), and one of the lowest average group sizes, all of which made it the perfect fit for seasonal workers who used the T-10 fare;
- Profile A3: Groups (*N* = 407), characterised by its high average group size, represented potential excursionists travelling in groups;
- Profile A4: Long-term (*N* = 4306) had a high number of transactions distributed along a moderate number of days. Thus, the profile might represent tourists having a long (i.e., 2-week) stay on average according to the card's lifespan (i.e., *active_period*);
- Profile A5: Short-term (*N* = 27224) stood out significantly for having accrued the most cards (i.e., 79.56% of all cards). The profile contrasted Profile A4: Long-term, for whereas the latter represented tourists with long stays (i.e., 2 weeks on average), the profile represented the average tourist in the Costa Daurada, whose stays are short (i.e., 5 days on average). The profile subsequently had its activity concentrated in a brief period of individual trips or in small groups and thus depicted the major types of tourism at the destination: tourists travelling in families or in couples.

Table 3. Profiles based on activity variables in 2019, with the selected model in yellow

				Probability		Class						
Class	AIC	BIC	Entropy	min	max	1	2	3	4	5	6	7
2	805265.3	805375.0	0.9863781	0.9708336	0.9979927	31722	2492					
3	789298.2	789450.1	0.9484448	0.8567533	0.9898127	3005	29218	1991				
4	780107.8	780301.9	0.9570076	0.8555235	0.9894300	2962	409	1977	28866			
5	762991.8	763228.2	0.9526662	0.8725295	0.9853362	1783	494	407	4306	27224		
6	759778.5	760057.1	0.8948376	0.6317863	0.9869072	24814	258	1761	2796	483	4102	
7	752365.4	752686.2	0.8831571	0.6745753	0.9805219	290	22509	235	1758	5768	2599	1055

Table 4. Descriptive statistics of the activity profiles in 2019

		A1: Sporadic (*N* = 1783)	A2: Continued (*N* = 494)	A3: Groups (*N* = 407)	A4: Long-term (*N* = 4306)	A5: Short-term (*N* = 27224)	Total (*N* = 34214)
transactions	Mean (*SD*)	17.7 (10.9)	65.7 (28.9)	26.9 (17.9)	40.3 (13.9)	12.5 (7.04)	17.2 (14.3)
	Median [min max]	15.0 [2.0, 69.0]	60.5 [19.0, 189]	20.0 [8.0, 119]	39.0 [10.0, 129]	10.0 [1.0, 40.0]	11.0 [1.0, 189]
active_period	Mean (*SD*)	51.4 (14.6)	47.9 (17.8)	2.64 (3.7)	13.0 (6.8)	5.2 (4.58)	9.17 (12.9)
	Median [min max]	49.0 [28.0, 93.0]	46.0 [11.0, 95.0]	1.0 [1.0, 36.0]	11.0 [3.0, 41.0]	4.0 [1.0, 33.0]	5.0 [1.0, 95.0]
active_days	Mean (*SD*)	6.04 (3.03)	18.3 (6.8)	1.69 (1.05)	7.17 (2.1)	2.71 (1.39)	3.66 (2.99)
	Median [min max]	5.0 [2.0, 16.0]	16.0 [10.0, 58.0]	1.0 [1.0, 7.0]	7.0 [3.0, 18.0]	3.0 [1.0, 10.0]	3.0 [1.0, 58.0]
avg_group_size	Mean (*SD*)	1.88 (0.81)	2.13 (0.881)	10.2 (3.82)	3.10 (1.17)	2.69 (1.11)	2.78 (1.45)
	Median [min max]	1.8 [0, 7.0]	2.11 [1.0, 5.7]	9.0 [6.67, 30.0]	2.92 [1.0, 9.15]	2.5 [0, 7.0]	2.5 [0, 30.0]

Table 5. Descriptive statistics of the activity profiles in 2020

		A1: Sporadic (*N* = 996)	A2: Continued (*N* = 135)	A3: Groups (*N* = 28)	A4: Long-term (*N* = 206)	A5: Short-term (*N* = 3197)	Total (*N* = 4562)
transactions	Mean (*SD*)	13.0 (7.85)	44.7 (24.3)	29.2 (37.0)	27.7 (13.5)	7.23 (5.04)	10.7 (11.1)
	Median [min max]	10.0 [2.0, 52.0]	40.0 [17.0, 170]	18.0 [8.0, 197]	25.5 [9.0, 88.0]	6.0 [1.0, 36.0]	8.0 [1.0, 197]
active_period	Mean (*SD*)	50.9 (14.2)	53.8 (17.8)	2.32 (4.10)	19.0 (7.43)	7.2 (7.75)	18.6 (21.3)
	Median [min max]	49.0 [28.0, 92.0]	53.0 [12.0, 94.0]	1.0 [1.0, 22.0]	19.0 [4.00, 36.0]	4.0 [1.0, 32.0]	9.0 [1.0, 94.0]
active_days	Mean (*SD*)	5.87 (2.97)	21.4 (7.94)	1.46 (1.10)	8.56 (2.49)	2.47 (1.59)	4.04 (4.27)
	Median [min max]	5.0 [2.0, 16.0]	20.0 [9.0, 48.0]	1.0 [1.0, 5.0]	8.0 [4.0, 17.0]	2.0 [1.0, 9.0]	3.0 [1.0, 48.0]
avg_group_size	Mean (*SD*)	1.47 (0.567)	1.38 (0.881)	11.3 (4.37)	2.02 (1.09)	1.96 (1.08)	1.89 (1.29)
	Median [min max]	1.27 [0, 5.0]	1.09 [1.0, 10.0]	9.67 [7.5, 24.0]	1.82 [1.0, 6.0]	2.0 [0, 7.0]	1.67 [0, 24.0]

Last, data from 2020 were classified using the same model of 2019 for comparison. Table 5, presenting descriptive statistics of the profiles for 2020, shows that differences from the 2019 clustering are clearly visible, including a marked increase in the lifespan of cards and decreases in average group size and the number of transactions.

4.2. Spatial clusters

In the same way as activity clustering, Figure 5 shows correlograms representing the initial (Figure 5a) and final (Figure 5b) results of the selection of variables. 4 variables were discarded for their redundancy: *visited_municipalities*, *transactions_tarr_reus*, *transactions_urban_municipalities*, and *main_three_municipalities*.

Next, clusters were formed by applying LPA; Table 6 shows the different models obtained and their fit indices. The same reasoning in forming the activity clusters was followed in forming the spatial clusters – that is, the 3 latest models (i.e., with 5, 6, and 7 profiles) were the highest-performing, but the models with 6 and 7 profiles presented groups with less than 1% of the total cards and had the least minimum posterior probabilities. Therefore, the model with 5 profiles was chosen.

Table 7 shows the descriptive statistics of the spatial profiles. Remarkably, the transactions in the tourist municipalities (i.e., Cambrils, Salou, and Vila-seca) represented

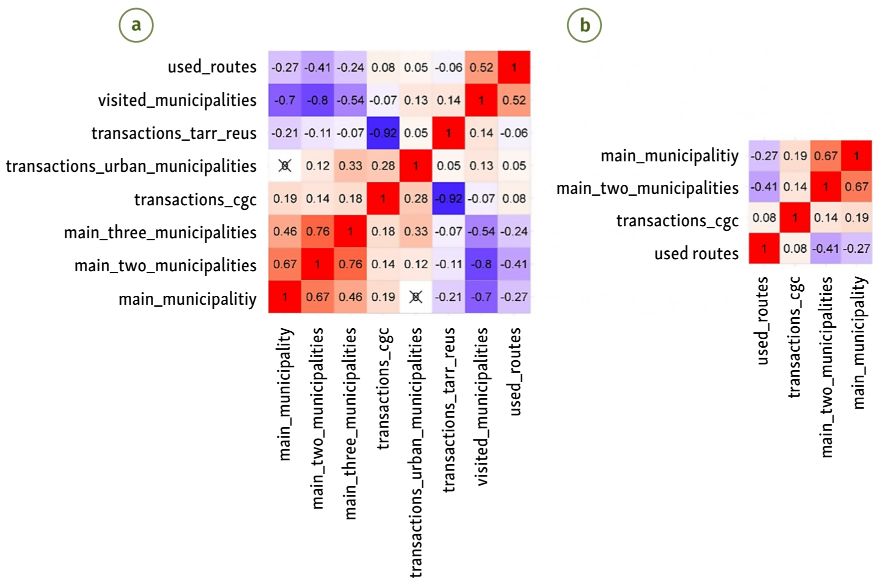

Figure 5. Correlograms showing the self-correlation of spatial variables: (a) – all variables; (b) – selected variables

69.6% of all transactions, and card transactions tended to be concentrated in the 2 main municipalities, at a rate of 86%. Each profile was named accordingly as follows:

- Profile S1: Cities (*N* = 7453) had a lower presence than other profiles in the touristic municipalities (i.e., Cambrils, Salou, and Vila-seca) while maintaining a high spatial concentration in terms of the number of transactions made in the major municipalities visited (i.e., *main_municipality* and *main_two_municipalities*);
- Profile S2: Sprawl (*N* = 13,639) had the most-used routes and the lowest percentage of spatial concentration, with the least number of transactions concentrated in the 2 major municipalities visited;
- Profile S3: Coast (*N* = 8822) had a high number of transactions around the most touristic municipalities and a high percentage of spatial concentration;
- Profile S4: Concentrated in cities (*N* = 1048) stood out by having the lowest percentage of transactions around the most touristic municipalities and one of the highest percentages for spatial concentration (i.e., *main_municipality* and *main_two_municipalities*);
- Profile S5: Concentrated on coast (*N* = 3252) related to profile S4 (i.e., Concentrated in cities) in having one of the highest percentages of spatial concentration (i.e., *main_municipality* and *main_two_municipalities*). At the same time, it also had the highest percentage of transactions around the most touristic municipalities.

The identified profiles share numerous characteristics. In Profiles S1 and S4, most transactions were made in one of the 2 most important cities (i.e., Tarragona and Reus), which makes them good representatives of cultural tourism in urban environments. In Profiles S3 and S5, most transactions were made in the touristic municipalities (i.e., Cambrils, Salou, and Vila-seca), which makes them good representatives of sun-and-sand tourism. However, profiles S4 and S5 were both highly concentrated around only one municipality, which makes them more spatially concentrated than their counterparts S1 and S3.

Based on the classification of data from 2020, Table 8 presents descriptive statistics of the spatial profiles. Compared with transactions in 2019, a generalised decrease in the most touristic municipalities can be observed – namely, from 69.6% in 2019 to 51.8% in 2020 – with S1: Cities and S2: Sprawl being the most-affected profiles. A slight increase in the spatial concentration of cards appeared in the number of transactions in the main municipalities visited by the card holder (i.e., *main_municipality* and *main_two_municipalities*).

Table 6. Profiles based on spatial variables in 2019, with the selected model in yellow

				Probability		Class						
Class	AIC	BIC	Entropy	min	max	1	2	3	4	5	6	7
2	962956.6	963066.3	0.9479918	0.8452416	0.9949068	32602	1612					
3	926153.9	926305.8	0.9001004	0.9472128	0.9684168	10534	15847	7833				
4	914726.0	914920.1	0.9021769	0.9410066	0.9654850	13161	16352	1075	3626			
5	901025.6	901261.9	0.9117115	0.9188214	0.9590967	7453	13639	8822	1048	3252		
6	904711.1	904989.6	0.8782586	0.0052578	0.9621334	11768	5243	2777	7059	7351	16	
7	899473.0	899793.7	0.8907574	0.2712793	0.9662109	631	11486	4963	7324	2795	472	6543

Table 7. Descriptive statistics of the spatial profiles in 2019

		S1: Cities (*N* = 7453)	S2: Sprawl (*N* = 13639)	S3: Coast (*N* = 8822)	S4: Concentrated in cities (*N* = 1048)	S5: Concentrated on coast (*N* = 3252)	Total (*N* = 34214)
transactions_cgc	Mean (*SD*)	45.9 (13.7)	65.9 (15.1)	92.1 (9.98)	4.23 (7.69)	99.3 (3.09)	69.6 (25.2)
	Median [min max]	50.0 [0, 70.6]	66.7 [0, 100]	100 [60.0, 100]	0 [0, 23.5]	100 [80.0, 100]	70.0 [0, 100]
main_two_municipalities	Mean (*SD*)	96.9 (6.16)	70.6 (9.63)	94.0 (7.51)	99.3 (3.47)	99.9 (0.733)	86.0 (14.9)
	Median [min max]	100 [60.0, 100]	72.0 [0, 94.0]	100 [74.0, 100]	100 [67.0, 100]	100 [91.0, 100]	89.0 [0, 100]
main_municipalitiy	Mean (*SD*)	54.7 (6.65)	41.8 (7.90)	60.9 (8.95)	92.0 (11.2)	95.7 (7.30)	56.2 (18.5)
	Median [min max]	50.0 [38.0, 75.0]	40.0 [0, 68.0]	60.0 [42.0, 90.0]	100 [65.0, 100]	100 [78.0, 100]	50.0 [0, 100]
used_routes	Mean (*SD*)	1.43 (0.606)	2.23 (0.676)	1.94 (0.547)	1.26 (0.558)	1.53 (0.546)	1.89 (0.703)
	Median [min max]	1.0 [1.0, 5.0]	2.0 [1.0, 10.0]	2.0 [1.0, 7.0]	1.0 [1.0, 6.0]	2.0 [1.0, 4.0]	2.0 [1.0, 10.0]

Table 8. Descriptive statistics of the spatial profiles in 2020

		S1: Cities (*N* = 1646)	S2: Sprawl (*N* = 937)	S3: Coast (*N* = 840)	S4: Concentrated in cities (*N* = 578)	S5: Concentrated on coast (*N* = 561)	Total (*N* = 4562)
transactions_cgc	Mean (*SD*)	32.4 (22.4)	50.9 (26.5)	93.4 (10.9)	1.86 (5.57)	99.2 (3.3)	51.8 (37.4)
	Median [min max]	40.7 [0, 70.0]	57.6 [0, 100]	100 [60.0, 100]	0 [0, 23.5]	100 [80.0, 100]	50.0 [0, 100]
main_two_municipalities	Mean (*SD*)	96.2 (7.17)	69.8 (14.9)	96.8 (5.98)	99.3 (3.2)	99.9 (0.571)	91.7 (14.1)
	Median [min max]	100 [60.0, 100]	75.0 [0, 100]	100 [78.0, 100]	100 [70.0, 100]	100 [93.0, 100]	100 [0, 100]
main_municipalitiy	Mean (*SD*)	55.0 (7.09)	43.3 (9.65)	62.0 (9.23)	92.1 (12.0)	96.2 (7.02)	63.6 (20.6)
	Median [min max]	50.0 [40.0, 75.0]	45.0 [0, 70.0]	61.0 [43.0, 82.0]	100 [67.0, 100]	100 [78.0, 100]	57.0 [0, 100]
used_routes	Mean (*SD*)	1.64 (0.849)	2.68 (1.35)	1.99 (0.731)	1.30 (0.622)	1.45 (0.568)	1.85 (1.02)
	Median [min max]	1.0 [1.0, 6.0]	2.0 [1.0, 8.0]	2.0 [1.0, 7.0]	1.0 [1.0, 6.0]	1.0 [1.0, 3.0]	2.0 [1.0, 8.0]

4.3. Schedule clusters

Again, the same rationale for activity and spatial clustering was followed for schedule clustering. Figure 6 shows correlograms representing the initial (Figure 6a) and final (Figure 6b) result of variable selection. 4 variables were discarded due to redundancy: *first_half_day*, *weekdays*, *weekends*, and *time_afternoon*.

Table 9 shows the different models obtained and their fit indices. Based on fit, the 3 latest models (i.e., with 5, 6, and 7 profiles) were the highest-performing. Among similarities between the models revealed by deeper analysis, models with 6 and 7 profiles were nearly identical to the 5-profile model but divided some profiles into smaller ones. Because those finer divisions did not provide any additional or meaningful information about the smart cards, the model with 5 profiles was chosen.

As with the previous clustering analyses, Table 10 presents descriptive statistics of the profiles. Each profile was named accordingly and is described in the following:

- Profile T1: Night (*N* = 2774) had the highest concentration of transactions at night (*time_night*), primarily including card transactions from 21:00 to 6:00;
- Profile T2: Distributed (*N* = 11,668) had a balanced concentration of transactions throughout the day, with many during the 1st (i.e., *time_morning* and *time_midday*) and the 2nd halves of the day;
- Profile T3: Evening (*N* = 5252) had the most card transactions during the 2nd half of the day, especially from 16:00 to 21:00;
- Profile T4: Noon (*N* = 8745) represented card transactions during midday, primarily between 12:00 and 17:00;
- Profile T5: Early risers (*N* = 5775) had the highest concentration of card transactions during the morning between 7:00 and 12:00.

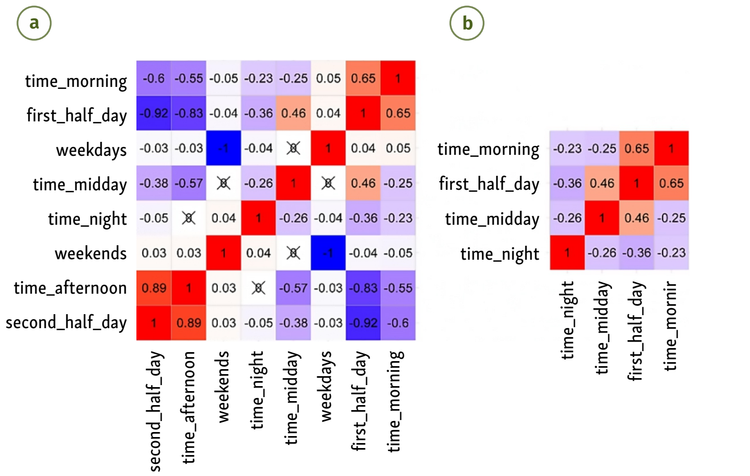

Figure 6. Correlogram showing the self-correlation of schedule variables: (a) – all variables; (b) – selected variables

Table 11 presents descriptive statistics of the profiles for 2020. The total averages did not significantly change from one year to the next, although each profile did experience some changes. Cards tended to be more temporarily concentrated by profile, as shown by the increase of transactions during the 2nd half of the day, morning, and noon for profiles T3: Evening, T4: Noon, and T5: Early Risers, respectively.

Table 9. Profiles based on schedule variables (2019), with the selected model in yellow

				Probability		Class						
Class	AIC	BIC	Entropy	min	max	1	2	3	4	5	6	7
2	1204408	1204518	0.7168136	0.8907952	0.9353218	13676	20538					
3	1197850	1198002	0.7361772	0.7820151	0.9025426	18446	11751	4017				
4	1172515	1172709	0.7757724	0.8046899	0.9391861	4800	13997	12551	2866			
5	1164236	1164472	0.7930217	0.8422535	0.9356653	2774	11668	5252	8745	5775		
6	1153424	1153702	0.8423840	0.8538415	0.9416129	5332	10899	8984	2762	1176	5061	
7	1143467	1143788	0.8629073	0.8468671	0.9541488	2766	6341	1543	10592	1170	7844	3958

Table 10. Descriptive statistics of the schedule profiles in 2019

		T1: Night (*N* = 2274)	T2: Distributed (*N* = 11668)	T3: Evening (*N* = 5252)	T4: Noon (*N* = 8745)	T5: Early Risers (*N* = 5775)	Total (*N* = 34214)
second_half_day	Mean (*SD*)	40.8 (20.9)	44.9 (12.0)	86.5 (13.2)	40.7 (18.9)	5.79 (9.07)	43.3 (27.3)
	Median [min max]	46.7 [0, 76.0]	46.2 [10.0, 75.0]	86.4 [57.7, 100]	44.4 [0, 100]	0 [0, 35.7]	44.9 [0, 100]
time_morning	Mean (*SD*)	11.0 (15.0)	37.9 (12.0)	4.39 (7.46)	9.70 (10.5)	59.4 (19.0)	27.0 (23.6)
	Median [min max]	0 [0, 66.7]	37.5 [8.00, 68.8]	0 [0, 26.3]	7.69 [0, 40.0]	50.0 [30.0, 100]	25.0 [0, 100]
time_midday	Mean (*SD*)	13.3 (15.9)	20.8 (15.0)	15.3 (16.9)	57.7 (18.7)	36.6 (20.0)	31.5 (24.2)
	Median [min max]	8.00 [0, 71.4]	21.4 [0, 61.5]	10.3 [0, 90.0]	50.0 [24.0, 100]	42.9 [0, 70.0]	31.0 [0, 100]
time_night	Mean (*SD*)	37.7 (15.8)	2.68 (5.77)	2.26 (5.54)	1.78 (4.81)	0.471 (2.86)	4.85 (11.8)
	Median [min max]	33.3 [19.7, 100]	0 [0, 25.0]	0 [0, 23.8]	0 [0, 25.0]	0 [0, 30.0]	0 [0, 100]

Table 11. Descriptive statistics of the schedule profiles in 2020

		T1: Night (*N* = 274)	T2: Distributed (*N* = 956)	T3: Evening (*N* = 1012)	T4: Noon (*N* = 1193)	T5: Early Risers (*N* = 1127)	Total (*N* = 4562)
second_half_day	Mean (*SD*)	38.7 (23.6)	45.4 (12.8)	90.7 (12.3)	33.9 (25.1)	4.17 (7.95)	41.9 (34.4)
	Median [min max]	40.0 [0, 75.0]	50.0 [18.2, 75.0]	100 [60.0, 100]	40.0 [0, 100]	0 [0, 33.3]	40.0 [0, 100]
time_morning	Mean (*SD*)	7.20 (12.6)	38.9 (12.1)	2.69 (5.99)	6.52 (9.84)	67.7 (22.8)	27.6 (30.1)
	Median [min max]	0 [0, 66.7]	40.0 [12.5, 66.7]	0 [0, 25.0]	0 [0, 35.7]	61.1 [30.0, 100]	20.0 [0, 100]
time_midday	Mean (*SD*)	15.7 (18.8)	19.1 (15.5)	12.2 (17.1)	67.5 (22.6)	29.2 (22.4)	32.5 (29.4)
	Median [min max]	9.55 [0, 66.7]	20.0 [0, 60.0]	0 [0, 77.8]	60.0 [27.8, 100]	33.3 [0, 70.0]	30.0 [0, 100]
time_night	Mean (*SD*)	41.9 (20.0)	1.69 (4.89)	1.26 (4.15)	1.10 (3.87)	0.294 (2.05)	3.51 (11.5)
	Median [min max]	33.3 [20.0, 100]	0 [0, 25.0]	0 [0, 22.5]	0 [0, 22.2]	0 [0, 22.5]	0 [0, 100]

4.4. Comparison

As described, data collected from smart cards were grouped into 3 categories (i.e., activities, spatial, and schedule) and 5 profiles per category. The results of clustering yielded the 3 respective statistical models, which allowed describing each cluster in detail and classifying data by year. In turn, the chief results of the study are summarised in Figure 7, in which red bars represent the number of cards for each profile in 2019. Therein, the card data for 2020 per profile, classified according to the clustering models, appear in percentages as blue bars. Last, the absolute values of the number of cards for each profile and year are shown as points. Differences between the profiles of one year and the next are easily attributable to the disturbances caused by the pandemic to tourists' arrivals and visitors' willingness to board public transport vehicles.

The 1st and most striking change concerns with the dominant use of T-10 cards, which corresponds to concentrated use by visitors during short periods (i.e., A5: Short-Term). The change can be easily appreciated in absolute terms because the number of cards for the profile decreased from 27240 in 2019 to 3197 in 2020. However, in relative terms, the profile continued to have the most frequent use in terms of activity type. Regarding the most-affected activity profiles, the sporadic use of cards (i.e., A1: Sporadic) quadrupled in use in 2020 due to irregularity in movements during the period and a lower reduction in use than the rest of profiles. Obviously, the group travel profile (A3) with the T-10 card (i.e., consecutive validations at the same stop) practically disappeared in plummeting from 407 to only 28 cards. Although the average size of the groups did not decrease much, throughout the year no card exhausted the maximum number of 30 trips that the T-10 card can carry.

As for the spatial profiles, the decrease in the number of trips with the T-10 was generalised. Even so, it is worth distinguishing profiles that involved travelling to inland municipalities, which experienced drastic drops, from profiles that moved only to coastal municipalities and the rest of profiles that did so in a concentrated, city-oriented way. The behavioural change shows the clear preference of visitors in 2020 to stay in the same location (i.e., S1: Cities, S4: Concentrated in cities, and S5: Concentrated on coast), accompanied by a slight decrease in visitors who moved primarily between major cities (i.e., Reus and Tarragona).

The profiles based on schedules clearly show a decrease in the use of T-10 cards. However, it seems that the decline was far more pronounced in the profiles that distributed journeys with public transport throughout various time slots, going from 11668 to 956 cards in the profile (T2). The rest of the profiles experienced significant falls in absolute terms, but their relative weight increased due to the complementary nature of the groups.

5. Discussion

The study's initial hypothesis was that the impact of the COVID-19 pandemic on the demand for public transport among different segments of tourists was unevenly affected during the summer of 2020, given that people's behavioural responses to travel was directly determined by their subjective perception of risk (Neuburger, Egger 2021). In that vein, the period of the analysis is particularly relevant, for it was characterised by the pairing of tourist activity, just after the end of lockdowns, with severe restrictions imposed to deter the spread of the disease and people's fear of contracting it at the time, when an effective vaccine was not yet available. During that time frame,

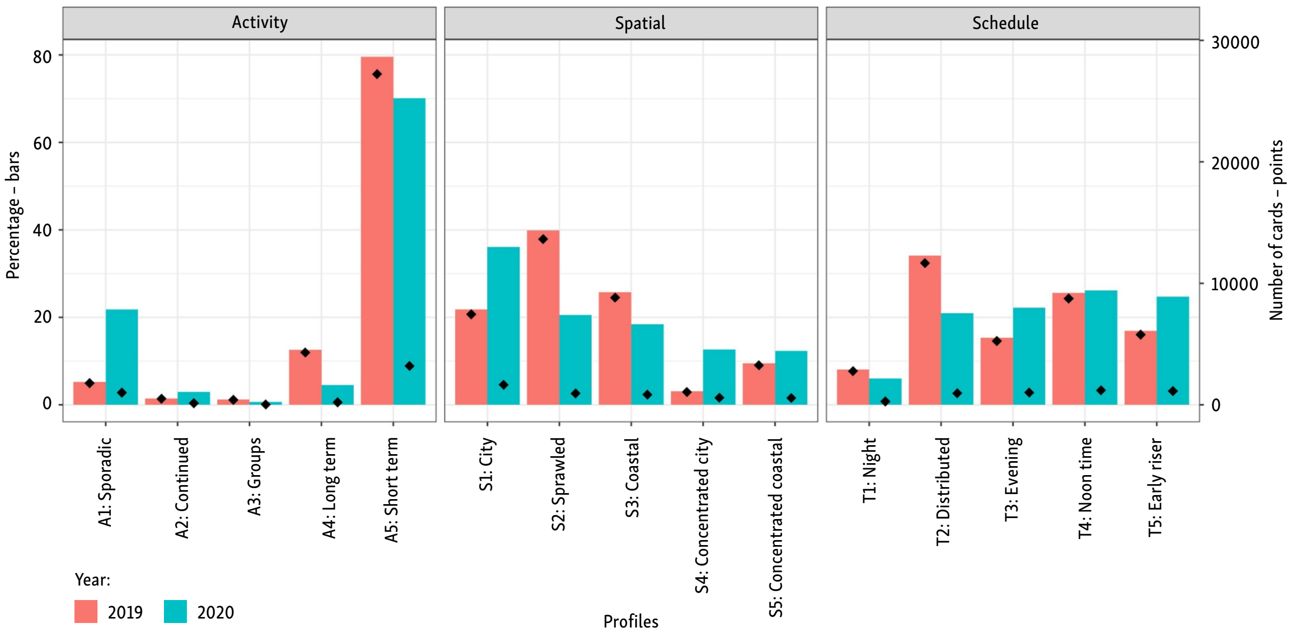

Figure 7. Comparison of the size of activity, spatial, and schedule profiles between 2019 and 2020

several bans and travel restrictions persisted that impacted both the number of visitors and their type. Added to those elements, individual-level fear of the disease conditioned tourists' decisions regarding travelling to a destination and where and how to move around it during their stays (Shin *et al.* 2022).

Given the uncertain circumstances of the pandemic, a time marked by continuously evolving, unpredictable situations, the availability of meaningful information about the era has become especially valuable. In that context, tourists' demand for public transport services was influenced by a range of simultaneously occurring elements that varied from one day to the next. Indeed, the reduction in visitors' ridership under the threat of infection was a consequence of the reduction in tourist arrivals, changes in the profile of visitors, restrictions imposed on certain activities, and the fear of being infected on board or at the attraction or place to be visited. Those factors could also act unevenly across individuals, which would lead to an even more even complex scenario. Empirical evidence related to passenger reduction due to fear of infection has remained rather scarce and rarely highlights factors that resulted in larger or smaller impacts on ridership. In that sense, Delclòs-Alió *et al.* (2022) concluded that some factors (e.g., expenditure and age) could have driven asymmetries in tourists' abandonment of public transport during the pandemic. Such a diversity of elements caused additional uncertainty for the managers of tourist destinations and public transport operators. For that reason, the detection of differences in the use of public transport across profiles might indicate distinct degrees of tourists' sensitivity to boarding public transport vehicles at certain times of day or to visiting specific types of attractions. That information is highly valuable for detecting transport lines, places, and time slots that have become more vulnerable due to the pandemic.

As expected, the combined effect of a high level of vulnerability to COVID-19 among mature, mass coastal destinations (Duro *et al.* 2021) with the pandemic's greater impact on leisure mobility through public transport (Delclòs-Alió *et al.* 2022; Östh *et al.* 2023; Thao *et al.* 2024) caused a significant drop in the demand for public transport compared with reductions that have occurred in urban contexts (Almlöf *et al.* 2021; Jenelius, Cebecauer 2020; Rodríguez González *et al.* 2021; Mützel, Scheiner 2022; Zhang *et al.* 2021; Fernández Pozo *et al.* 2022). The number of smart cards used by tourists fell by 87%, while the number of journeys plunged by 92%. The latter percentage contrasts the percentage of loss of interurban public transport journeys throughout the year that, despite the absolute halt in mobility during the lockdown period, dropped by 64% compared with 2019 (Zaragozí *et al.* 2023).

The evidence obtained not only depicts a scenario in which the pandemic radically reduced the demand for public transport services but also confirms the hypothesis that it shaped how public transport was used at the destination. Regarding changes related to patterns in the validation of transport cards, the greater reduction in trips than in the number of cards signalled a decline in the average intensity of use. That result is backed by the smaller contraction registered by the number of journeys associated with A1: Sporadic (44%). Differences also emerged in the spatial dimension. The type of cards presenting smaller reductions in the demand for public transport were labelled S1: Cities (78%) and S4: Concentrated in cities (45%). By contrast, those within the category representing dispersion (i.e., S3: Sprawl) fell in utilisation by 93%. Opposed to patterns in the spatial and activity dimensions, patterns in schedules presented a more balanced decrease in terms of use. All categories fell between 80% and 90%, with the most significant losses in the profiles T1: Night and T2: Distributed, with drops of 90% and 92%, respectively.

That transport operators are under financial pressure has previously been documented (Gutiérrez *et al.* 2021; Tirachini, Cats 2020; Wasserman *et al.* 2022; Shaheen, Wong 2021). Nonetheless, according to the study's results, public transport operators whose demand is based on the massive influx of visitors to a territory were in a worse position in terms of financial exposure than companies operating within urban environments. The recovery of pre-pandemic levels in ridership on public transport is paramount to the parallel recovery of tourist activity. In that sense, from the perspective of public transport operators, it is essential to provide services that meet visitors' needs, even if the recovery of confidence in the health-related safety aboard public transport vehicles is slow (Vich *et al.* 2022) and even though the range of mobility-related decisions available to them is far higher than for commuters (Zamparini *et al.* 2022). As the study's results show, the spatial reduction in and reduced frequency of tourists' journeys across the destination was uneven, which indicates that tourists were willing to forgo travelling to certain locations instead of others and to modify how they used public transport services. Consequently, the lack of the use of public transport can be detrimental not only for transport operators but also for the attractiveness of the destination, because of the decline of the potential places that can be reached from where tourists' accommodations are located if public transport is not an option for moving around at the destination.

Taking all those elements under consideration, offering high-quality services is essential for visitors and needs to be attractive enough to convince them to move around during their stays. In that light, it is pivotal to cover the whole catchment area, which is configured by the spatial distribution of tourists' attractions and accommodations, and the frequency of trips available (Gronau, Kagermeier 2007). Because the choice of a destination and the location of the accommodation directly depend on the range of desired attractions that are accessible (Lue *et al.* 1993; Paulino *et al.* 2019), public transport is key to developing a destination's competitiveness (Prideaux 2000). Beyond that, for mature mass coastal destinations, it is imperative to rejuvenate the destinations and boost attractiveness to give visitors access to a diversity of attractions beyond the beach (Bujosa *et al.* 2015). In turn, the recovery of overnight stays, visits around influential areas, and levels of public transport ridership are highly dependent on each other. Furthermore, and perhaps more critically, the recovery of tourists' trust is not a homogeneous trend across visitors (Shin *et al.* 2022), for some might hesitate to board a bus or train again. Along the same lines, Ong *et al.* (2024) have posited that changes in the profiles of mobility related to non-commuting trips after the pandemic are likely to become structural and remain in the future depending on individual characteristics and perceptions. Bearing that in mind, and as previously introduced, data provided by smart travel cards can be a useful tool to monitor and manage the evolving scenario caused by asymmetric shocks. The underlying reason is that they allow the segmentation of the demand to the point that comparing ridership before and after a shock for each profile becomes feasible. As a result, it is possible to pinpoint, which profiles remain reluctant to board public transport vehicles, capture passenger behaviours, and identify transfers from one profile to another. It is also possible to disentangle, which tourist attractions reached by public transport have lost visitors.

Under those circumstances and following Miravet *et al.* (2021a), it is necessary to design appropriate tailor-made communication campaigns to focus on the profile of visitors who are targets of the potential gain of using public transport. The success of such campaigns is contingent on the availability of appropriate sources of data that allow the differentiation of behavioural patterns with respect to the use of public transport. Moreover, the sources of information need to be able to accommodate the continuous dynamics of demand inflows derived from uncertain circumstances. The effort needs to be shared between public transport operators and destination managers, for they provide a combined product that includes transport and visits to attractions. On that count, the recovery of the attractiveness of public transport services and the recovery of destinations' potential visits have to be regarded as parallel trends; otherwise, they might be unsuccessful and generate dysfunction.

The study presented in this article was not exempt from limitations. 1st, because smart card travel data usually only indicate the bus stop or station where users board, where tourists disembark remains unknown (Gutiérrez *et al.* 2020; Zaragozí *et al.* 2021). 2nd, restrictions related to data privacy regulations have blurred the precision of certain elements of the data, including bus stops or the exact times of validations. Instead, data at the municipality level and time slots were used. 3rd, a non-negligible part of tourist's journeys is not validated by means of smart travel cards but single transport tickets, and it is impossible to track bus passengers who paid for single tickets. Last, the data provided by the integrated fare system do not contain data beyond the journey; thus, there is a lack of information about the socioeconomic characteristics of the card holders.

6. Conclusions

This article has sought to disentangle the extent to which the demand among different profiles of public transport users was affected by the COVID-19 pandemic during the 2020 tourist season in the Costa Daurada. The characteristics of such a popular coastal destination are especially relevant, not only due to the high sensitivity of its demand to the incidence of the pandemic but also due to the central role that public transport plays in developing tourist activity in the area (Miravet *et al.* 2021a) and the substantial share of transport operators' revenue that directly depends on tourists' use of the services (Gutiérrez, Miravet 2016a).

Following the methodology established by Gutiérrez *et al.* (2020) and Zaragozí *et al.* (2021) for analysing smart travel card data in the context of tourist destinations, LPA was applied to anonymised smart card data from 2019. The analysis distinguished 5 profiles of visitors based on card usage activity: recurrent users, groups, short stays, long stays, and sporadic users. LPA was also used to establish 5 traveller profiles pertinent to a card's spatial utilisation: coast, city, dispersed, coast-concentrated, and city-concentrated. Last, a 3rd classification was explored by considering the temporal use of the cards, which yielded the following profiles: night, distributed, evening, noon, and early risers. Subsequently, smart travel cards used by tourists in 2020 were classified according to the profiles identified in 2019 to compare both years.

Therefore, from a methodological point of view, the study has highlighted the suitability of smart travel cards to exploring the repercussions of shocks to the demand for public transport. Such cards enable researchers to disentangle asymmetric impacts based on comparisons with users' former behaviours, especially considering the dimensions of time and space (Kurauchi, Schmöcker 2017). In fact, analysing AFC data had already been pinpointed as especially valuable in the context of seasonal tourist destinations characterised by continuous fluctuations in transport ridership owing to seasonality (Miravet *et al.* 2021a).

Last, future research should analyse whether the profile of public transport users gradually recovered after the summer of 2020 and which profiles of users did so at a slower pace. Such information is highly valuable for destination managers to ascertain whether a lack of recovery of visits to attractions may be due to the use of public transport. Beyond that, the methodology applied to the study should be extrapolated to other tourist destinations and to both rural and urban environments.

Acknowledgements

The data used in this study was prepared by the Consortium of Public Transport of Camp de Tarragona (Catalonia, Spain).

Access to the microdata was obtained by means of research cooperation agreement between Consortium of

Public Transport of Camp de Tarragona and the Research Group on Territorial Analysis and Tourism Studies (GRATET), University of Rovira i Virgili (Catalonia, Spain).

Funding

This work was supported by the Provincial Council of Tarragona, the "Fondo Supera COVID-19", created by the Santander Bank, CRUE *Universidades Españolas* and the *Consejo Superior de Investigaciones Científicas* (CSIC), the Spanish Ministry of Science, Innovation and Universities (ADAPTOUR/PID2020-112525RB-I00), the AEI/FEDER, UE and the Department of Research and Universities of the Catalan Government (GRATET-2021SGR00657).

Leonardo Monteiro-Fialho has received funding from the INVESTIGO 2022 program of the *Generalitat Valenciana* (INVEST/2022/424) under the "*Plan de recuperación, transformación y resiliencia*" initiated by the Spanish government and the "Next Generation EU" program of the European Union.

Author contributions

Aaron Gutiérrez: conceptualization, project administration, funding acquisition, writing (original draft).

Leonardo Monteiro-Fialho: formal analysis, writing (original draft).

Sergio Trilles: methodology, formal analysis, writing (original draft).

Benito Zaragozí: conceptualization, formal analysis, writing (original draft).

Carlos Granell: supervision, writing (review and editing).

Daniel Miravet: data curation, supervision, writing (original draft).

Disclosure statement

The authors declare no conflict of interest.

Appendix

Figures A1, A2 and A3 represent all 3 different clustering's.

Descriptive statistics, as showed on Section 4, are represented using a boxplot. Each point of the boxplot represents the mean of a profile on such variable, bars represent the confidence interval, and boxes represent the standard deviation enclosing ±64% of all the cards in such profile.

The bivariate scatterplot presents another way of visualising the cluster and easily visualising the clusters. Relations between variables are shown by pairs and the diagonal is composed by a density plot of each variable.

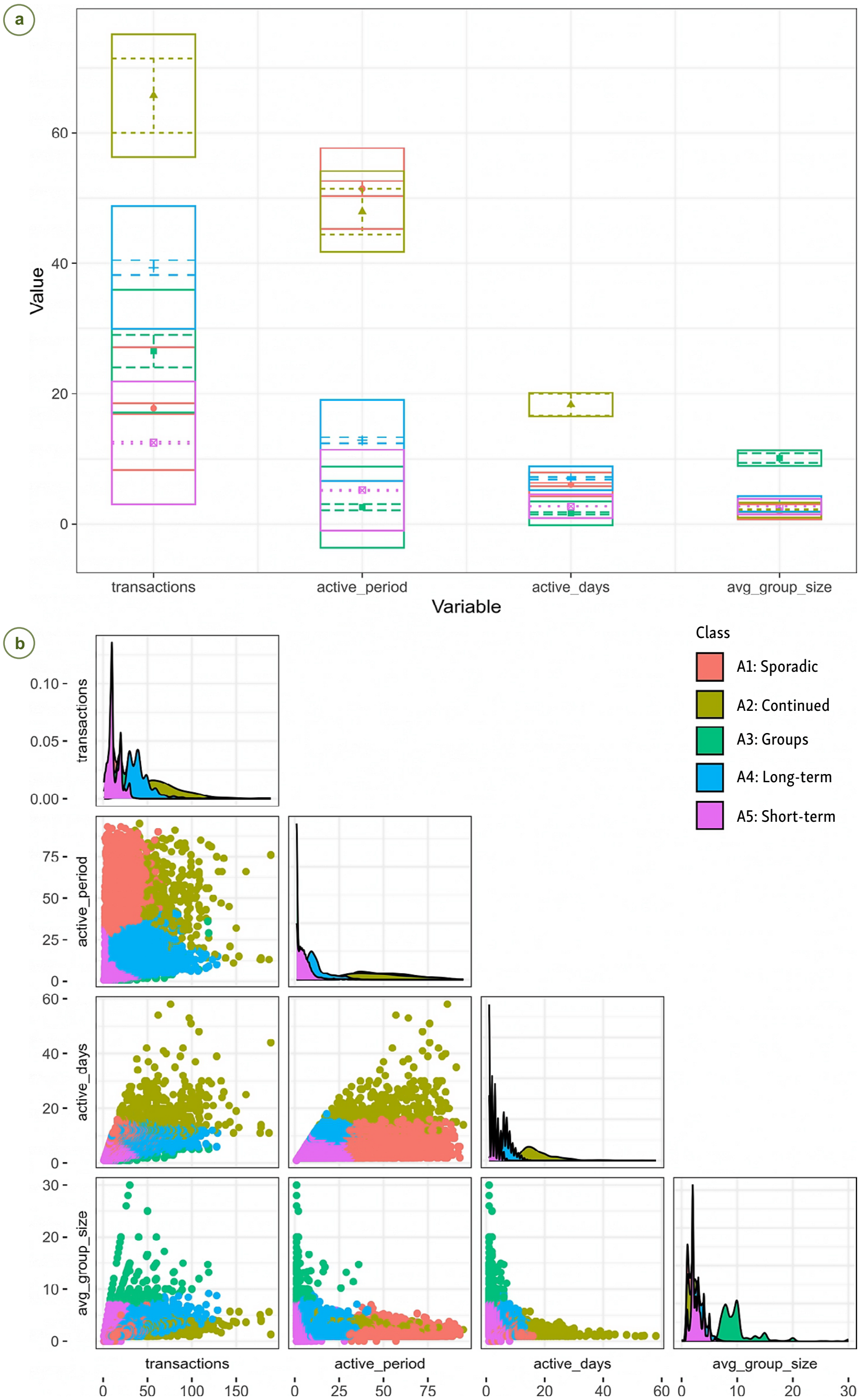

Figure A1. LPA with activity variables:

(a) – boxplot that shows the relation between profiles and variables;

(b) – bivariate scatterplot of the data where clusters and their differences are easily identified

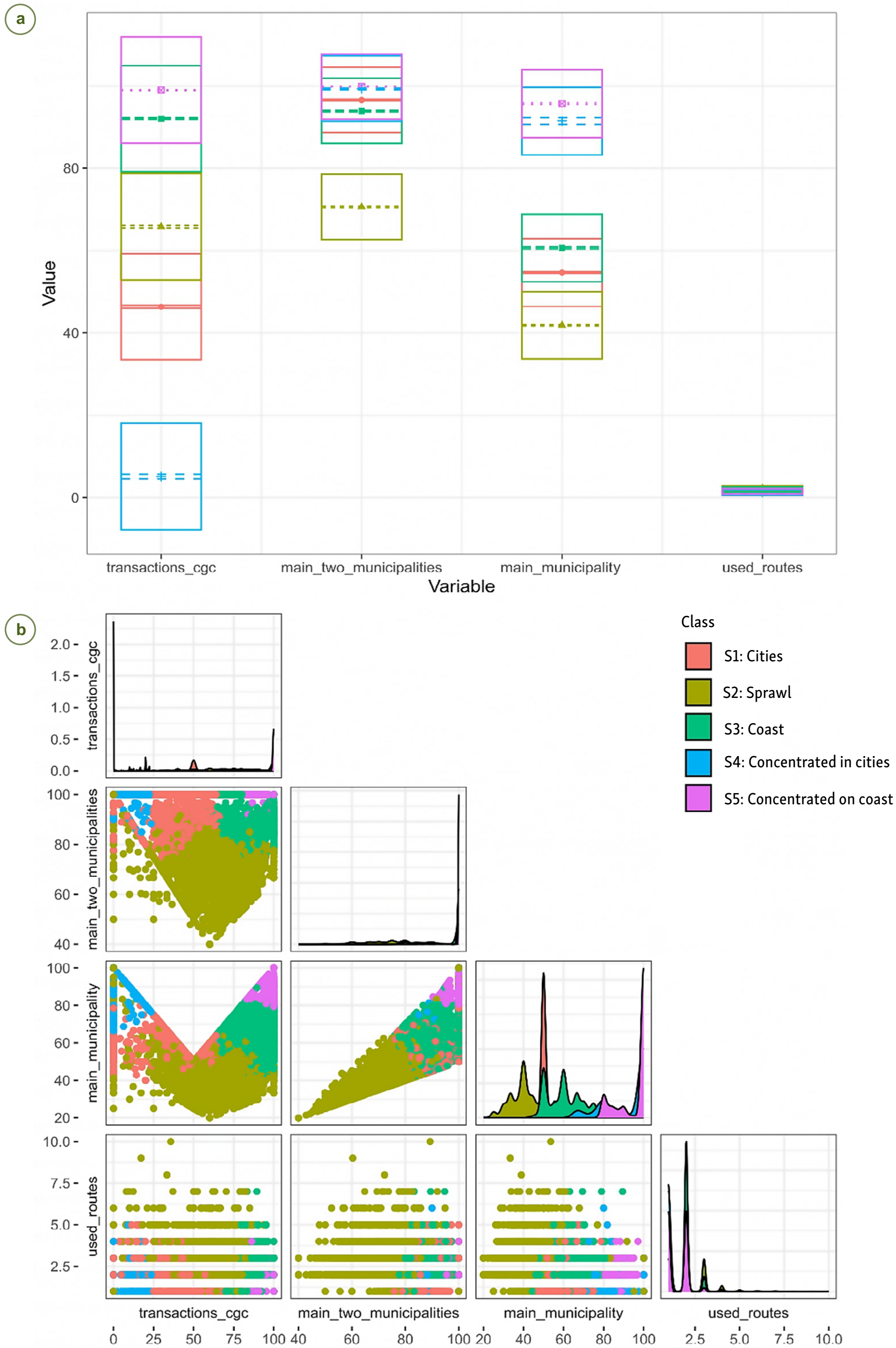

Figure A2. LPA with spatial variables:

a – boxplot that shows the relation between profiles and variables;

b – bivariate scatterplot of the data where clusters and their differences are easily identified

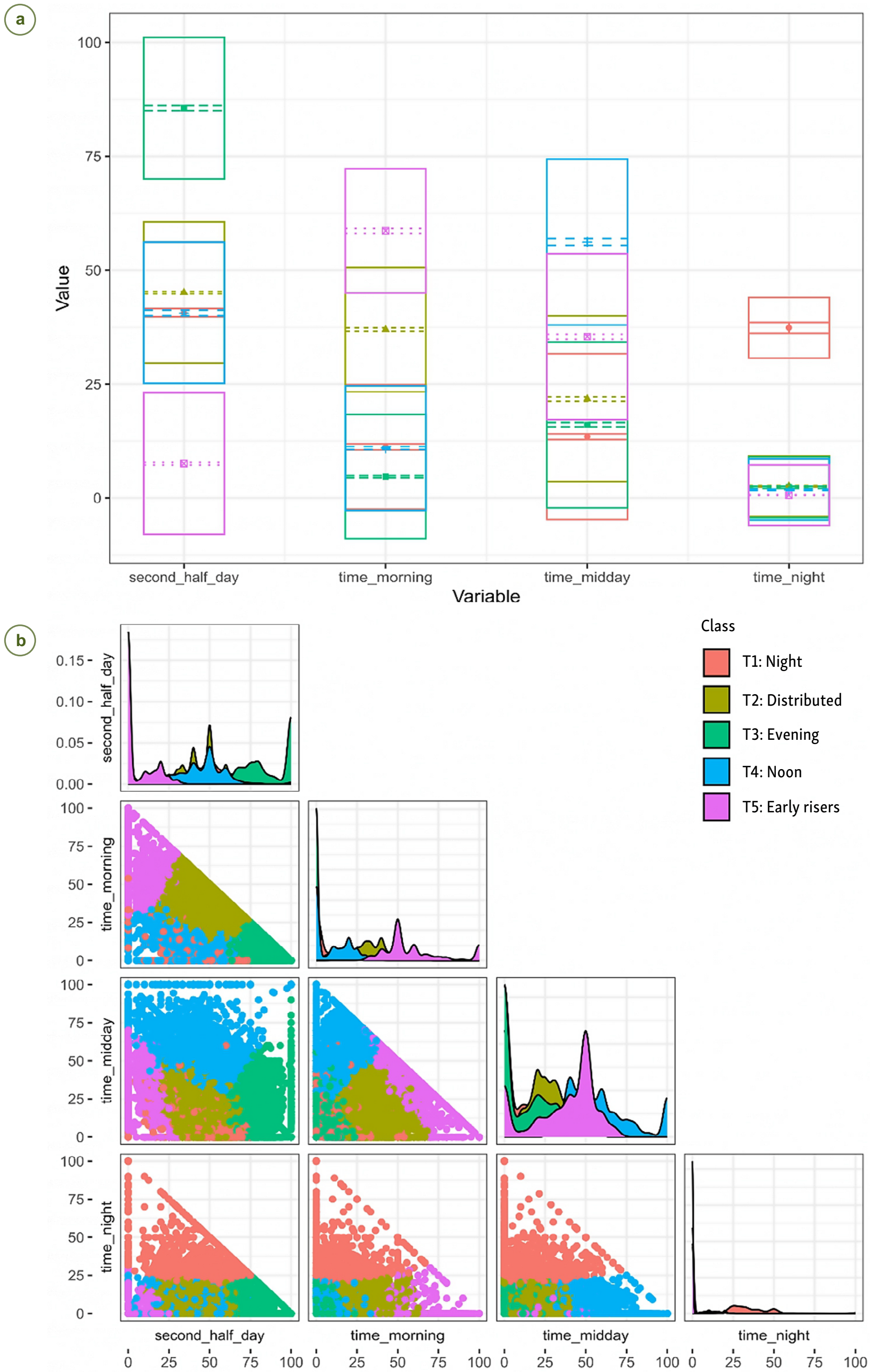

Figure A3. LPA with schedule variables:

a – boxplot that shows the relation between profiles and variables;

b – bivariate scatterplot of the data where clusters and their differences are easily identified

References

Abdullah, M.; Ali, N.; Dias, C.; Campisi, T.; Javid, M. A. 2021. Exploring the traveler's intentions to use public transport during the COVID-19 pandemic while complying with precautionary measures, *Applied Sciences* 11(8): 3630. https://doi.org/10.3390/app11083630

Abu-Rayash, A.; Dincer, I. 2020. Analysis of mobility trends during the COVID-19 coronavirus pandemic: exploring the impacts on global aviation and travel in selected cities, *Energy Research & Social Science* 68: 101693. https://doi.org/10.1016/j.erss.2020.101693

Albalate, D.; Bel, G. 2010. Tourism and urban public transport: holding demand pressure under supply constraints, *Tourism Management* 31(3): 425–433. https://doi.org/10.1016/j.tourman.2009.04.011

Almlöf, E.; Rubensson, I.; Cebecauer, M.; Jenelius, E. 2021. Who continued travelling by public transport during COVID-19? Socioeconomic factors explaining travel behaviour in Stockholm 2020 based on smart card data, *European Transport Research Review* 13: 31. https://doi.org/10.1186/s12544-021-00488-0

Anke, J.; Francke, A.; Schaefer, L.-M.; Petzoldt, T. 2021. Impact of SARS-CoV-2 on the mobility behaviour in Germany, *European Transport Research Review* 13: 10. https://doi.org/10.1186/s12544-021-00469-3

Bagchi, M.; White, P. R. 2005. The potential of public transport smart card data, *Transport Policy* 12(5): 464–474. https://doi.org/10.1016/j.tranpol.2005.06.008

Barbieri, D. M.; Lou, B.; Passavanti, M.; Hui, C.; Hoff, I.; Antunes Lessa, D.; Sikka, G.; Chang, K.; Gupta, A.; Fang, K.; Banerjee, A.; Maharaj, B.; Lam, L.; Ghasemi, N.; Naik, B.; Wang, F.; Foroutan Mirhosseini, A.; Naseri, S.; Liu, Z.; Qiao, Y.; Tucker, A.; Wijayaratna, K.; Peprah, P.; Adomako, S.; Yu, L.; Goswami, S.; Chen, H.; Shu, B.; Hessami, A.; Abbas, M.; Agarwal, N.; Rashidi, T. H. 2021. Impact of covid-19 pandemic on mobility in ten countries and associated perceived risk for all transport modes, *PLoS ONE* 16(2): e0245886. https://doi.org/10.1371/journal.pone.0245886

Beck, M. J.; Hensher, D. A.; Wei, E. 2020. Slowly coming out of COVID-19 restrictions in Australia: implications for working from home and commuting trips by car and public transport, *Journal of Transport Geography* 88: 102846. https://doi.org/10.1016/j.jtrangeo.2020.102846

Becken, S.; Jin, X.; Zhang, C.; Gao, J. 2017. Urban air pollution in China: destination image and risk perceptions, *Journal of Sustainable Tourism* 25(1): 130–147. https://doi.org/10.1080/09669582.2016.1177067

Böcker, L.; Olsson, L. E.; Uteng, T. P.; Friman, M. 2023. Pandemic impacts on public transport safety and stress perceptions in Nordic cities, *Transportation Research Part D: Transport and Environment* 114: 103562. https://doi.org/10.1016/j.trd.2022.103562

Bujosa, A.; Riera, A.; Pons, P. J. 2015. Sun-and-beach tourism and the importance of intra-destination movements in mature destinations, *Tourism Geographies: an International Journal of Tourism Space, Place and Environment* 17(5): 780–794. https://doi.org/10.1080/14616688.2015.1093538

Cahyanto, I.; Wiblishauser, M.; Pennington-Gray, L.; Schroeder, A. 2016. The dynamics of travel avoidance: The case of Ebola in the U.S., *Tourism Management Perspectives* 20: 195–203. https://doi.org/10.1016/j.tmp.2016.09.004

Campisi, T.; Basbas, S.; Tanbay, N. A.; Georgiadis, G. 2022a. Some considerations on the key factors determining the reduction of public transport demand in Sicily during COVID-19 pandemic, *International Journal of Transport Development and Integration* 6(1): 81–94. https://doi.org/10.2495/TDI-V6-N1-81-94

Campisi, T.; Basbas, S.; Trouva, M.; Papas, T.; Akgün-Tanbay, N.; Stamatiadis, N.; Tesoriere, G. 2022b. Anxiety, fear and stress feelings of road users during daily walking in COVID-19 pandemic: Sicilian cities, *Transportation Research Procedia* 62: 107–114. https://doi.org/10.1016/j.trpro.2022.02.014

Cats, O. 2024. Identifying human mobility patterns using smart card data, *Transport Reviews* 44(1): 213–243. https://doi.org/10.1080/01441647.2023.2251688

Cheng, L.; Ning, Z.; Lei, D.; Cai, X.; Chen, X. 2024. Assessing the influence of the COVID-19 pandemic on passengers' reliance on public transport, *Journal of Transport Geography* 119: 103951. https://doi.org/10.1016/j.jtrangeo.2024.103951

Da Silva Lopes, H.; Remoaldo, P. C.; Ribeiro, V.; Martín-Vide, J. 2021. Effects of the COVID-19 pandemic on tourist risk perceptions – the case study of Porto, *Sustainability* 13(11): 6399. https://doi.org/10.3390/su13116399

Das, S.; Boruah, A.; Banerjee, A.; Raoniar, R.; Nama, S.; Maurya, A. K. 2021. Impact of COVID-19: a radical modal shift from public to private transport mode, *Transport Policy* 109: 1–11. https://doi.org/10.1016/j.tranpol.2021.05.005

De Vos, J. 2020. The effect of COVID-19 and subsequent social distancing on travel behavior, *Transportation Research Interdisciplinary Perspectives* 5: 100121. https://doi.org/10.1016/j.trip.2020.100121

Delclòs-Alió, X.; Gutiérrez, A.; Miravet, D.; Tomàs-Porres, J.; Vich, G.; Clavé, S. A. 2022. Impact of the COVID-19 pandemic on tourist public transportation use and on its determinants: evidence from a Catalan coastal destination, *Tourism Management Perspectives* 44: 101003. https://doi.org/10.1016/j.tmp.2022.101003

Devillaine, F.; Munizaga, M.; Trépanier, M. 2012. Detection of activities of public transport users by analyzing smart card data, *Transportation Research Record: Journal of the Transportation Research Board* 2276: 48–55. https://doi.org/10.3141/2276-06

Domènech, A.; Miravet, D.; Gutiérrez, A. 2020. Mining bus travel card data for analysing mobilities in tourist regions, *Journal of Maps* 16(1): 40–49. https://doi.org/10.1080/17445647.2019.1709578

Domènech, A.; Miravet, D.; Gutiérrez, A. 2023. Tourists' transport modal choices in Barcelona, *Research in Transportation Business & Management* 48: 100902. https://doi.org/10.1016/j.rtbm.2022.100902

Duro, J. A.; Perez-Laborda, A.; Turrion-Prats, J.; Fernández-Fernández, M. 2021. COVID-19 and tourism vulnerability, *Tourism Management Perspectives* 38: 100819. https://doi.org/10.1016/j.tmp.2021.100819

Eisenmann, C.; Nobis, C.; Kolarova, V.; Lenz, B.; Winkler, C. 2021. Transport mode use during the COVID-19 lockdown period in Germany: the car became more important, public transport lost ground, *Transport Policy* 103: 60–67. https://doi.org/10.1016/j.tranpol.2021.01.012

Eusébio, C.; Vieira, A. L. 2013. Destination attributes' evaluation, satisfaction and behavioural intentions: a structural modelling approach, *International Journal of Tourism Research* 15(1): 66–80. https://doi.org/10.1002/jtr.877

Fernández Pozo, R.; Wilby, M. R.; Vinagre Díaz, J. J.; Rodríguez González, A. B. 2022. Data-driven analysis of the impact of COVID-19 on Madrid's public transport during each phase of the pandemic, *Cities* 127: 103723. https://doi.org/10.1016/j.cities.2022.103723

Foell, S.; Phithakkitnukoon, S.; Kortuem, G.; Veloso, M.; Bento, C. 2015. Predictability of public transport usage: a study of bus rides in Lisbon, Portugal, *IEEE Transactions on Intelligent Transportation Systems* 16(5): 2955–2960. https://doi.org/10.1109/TITS.2015.2425533

Ghaemi, M. S.; Agard, B.; Trépanier, M.; Partovi Nia, V. 2017. A visual segmentation method for temporal smart card data, *Transportmetrica A*: *Transport Science* 13(5): 381–404. https://doi.org/10.1080/23249935.2016.1273273

Gössling, S.; Scott, D.; Hall, C. M. 2021. Pandemics, tourism and global change: a rapid assessment of COVID-19, *Journal of Sustainable Tourism* 29(1): 1–20. https://doi.org/10.1080/09669582.2020.1758708

Gramsch, B.; Guevara, C. A.; Munizaga, M.; Schwartz, D.; Tirachini, A. 2022. The effect of dynamic lockdowns on public transport demand in times of COVID-19: evidence from smartcard data, *Transport Policy* 126: 136–150. https://doi.org/10.1016/j.tranpol.2022.06.012

Gronau, W.; Kagermeier, A. 2007. Key factors for successful leisure and tourism public transport provision, *Journal of Transport Geography* 15(2): 127–135. https://doi.org/10.1016/j.jtrangeo.2006.12.008

Gutiérrez, A.; Domènech, A.; Zaragozí, B.; Miravet, D. 2020. Profiling tourists' use of public transport through smart travel card data, *Journal of Transport Geography* 88: 102820. https://doi.org/10.1016/j.jtrangeo.2020.102820

Gutiérrez, A.; Miravet, D. 2016a. Estacionalidad turística y dinámicas metropolitanas: un análisis a partir de la movilidad en transporte público en el Camp de Tarragona, *Revista de geografía Norte Grande* 65: 65–89. https://doi.org/10.4067/S0718-34022016000300004 (in Spanish).

Gutiérrez, A.; Miravet, D. 2016b. The determinants of tourist use of public transport at the destination, *Sustainability* 8(9): 908. https://doi.org/10.3390/su8090908

Gutiérrez, A.; Miravet, D.; Domènech, A. 2021. COVID-19 and urban public transport services: emerging challenges and research agenda, *Cities* & *Health* 5(Sup1): S177–S180. https://doi.org/10.1080/23748834.2020.1804291

Hanaei, S.; Rezaei, N. 2020. COVID-19: developing from an outbreak to a pandemic, *Archives of Medical Research* 51(6): 582–584. https://doi.org/10.1016/j.arcmed.2020.04.021

Harris, J. R. 2020. *The Subways Seeded the Massive Coronavirus Epidemic in New York City*. Working Paper 27021. National Bureau of Economic Research, Cambridge, MA, US. 23 p. https://doi.org/10.3386/w27021

Hennig, C.; Meila, M.; Murtagh, F.; Rocci, R. (Eds). 2015. *Handbook of Cluster Analysis*. 773 p. Chapman and Hall/CRC. https://doi.org/10.1201/b19706

Hu, M.; Lin, H.; Wang, J.; Xu, C.; Tatem, A. J.; Meng, B.; Zhang, X.; Liu, Y.; Wang, P.; Wu, G.; Xie, H.; Lai, S. 2021. Risk of coronavirus disease 2019 transmission in train passengers: an epidemiological and modeling study, *Clinical Infectious Diseases* 72(4): 604–610. https://doi.org/10.1093/cid/ciaa1057

INE. 2020. *Censos de población y viviendas*. *2020*. Instituto Nacional de Estadística (INE), Madrid, España. https://www.ine.es (in Spanish).

Ivanova, M.; Ivanov, I. K.; Ivanov, S. 2021. Travel behaviour after the pandemic: the case of Bulgaria, *Anatolia*: *an International Journal of Tourism and Hospitality Research* 32(1): 1–11. https://doi.org/10.1080/13032917.2020.1818267

Jenelius, E.; Cebecauer, M. 2020. Impacts of COVID-19 on public transport ridership in Sweden: analysis of ticket validations, sales and passenger counts, *Transportation Research Interdisciplinary Perspectives* 8: 100242. https://doi.org/10.1016/j.trip.2020.100242

Jiao, J.; Azimian, A. 2021. Exploring the factors affecting travel behaviors during the second phase of the COVID-19 pandemic in the United States, *Transportation Letters*: *the International Journal of Transportation Research* 13(5–6): 331–343. https://doi.org/10.1080/19427867.2021.1904736

Kim, C.; Cheon, S. H.; Choi, K.; Joh, C.-H.; Lee, H.-J. 2017. Exposure to fear: Changes in travel behavior during MERS outbreak in Seoul, *KSCE Journal of Civil Engineering* 21(7): 2888–2895. https://doi.org/10.1007/s12205-017-0821-5

Kim, M. J.; Hall, C. M.; Chung, N.; Kim, M.; Sohn, K. 2023. What makes tourists use public transport? Value-belief-norm theory, environmental, social, and governance factors, and the sustainable development goals, *Journal of Travel Research* 63(6): 1426–1441. https://doi.org/10.1177/00472875231191708

Korinth, B. 2020. The impact of COVID-19 on foreign travel plans of Polish tourists in 2020, *Studia Periegetica* 32(4): 59–69. Available from Internet: https://journals.wsb.poznan.pl/index.php/sp/article/view/24/62

Ku, D.; Yeon, C.; Lee, S.; Lee, K.; Hwang, K.; Li, Y. C.; Wong, S. C. 2021. Safe traveling in public transport amid COVID-19, *Science Advances* 7(43): eabg3691. https://doi.org/10.1126/sciadv.abg3691

Kurauchi, F.; Schmöcker, J.-D. 2017. *Public Transport Planning with Smart Card Data*. 274 p. CRC Press. https://doi.org/10.1201/9781315370408

Kusakabe, T.; Asakura, Y. 2014. Behavioural data mining of transit smart card data: a data fusion approach, *Transportation Research Part C*: *Emerging Technologies* 46: 179–191. https://doi.org/10.1016/j.trc.2014.05.012

Lau, J. T. F.; Yang, X.; Tsui, H.; Kim J. H. 2003. Monitoring community responses to the SARS epidemic in Hong Kong: from day 10 to day 62, *Journal of Epidemiology* & *Community Health* 57(11): 864–870. https://doi.org/10.1136/jech.57.11.864

Leask, A.; Fyall, A.; Garrod, B. 2013. Managing revenue in Scottish visitor attractions, *Current Issues in Tourism* 16(3): 240–265. https://doi.org/10.1080/13683500.2012.667070

Lizana, M.; Choudhury, C.; Watling, D. 2024. Analysing the impacts of individual-level factors on public transport usage during the COVID-19 pandemic: a comprehensive literature review and meta-analysis, *Transport Reviews* 44(2): 434–460. https://doi.org/10.1080/01441647.2023.2295967

Long, A.; Carney, F.; Kandt, J. 2023. Who is returning to public transport for non-work trips after COVID-19? Evidence from older citizens' smart cards in the UK's second largest city region, *Journal of Transport Geography* 107: 103529. https://doi.org/10.1016/j.jtrangeo.2023.103529

Lu, Q.; Atadil, H. A. 2021. Do you dare to travel to China? An examination of China's destination image amid the COVID-19, *Tourism Management Perspectives* 40: 100881. https://doi.org/10.1016/j.tmp.2021.100881

Lue, C.-C.; Crompton, J. L.; Fesenmaier, D. R. 1993. Conceptualization of multi-destination pleasure trips, *Annals of Tourism Research* 20(2): 289–301. https://doi.org/10.1016/0160-7383(93)90056-9

Miravet, D.; Domènech, A.; Gutiérrez, A. 2021a. What prompts tourists to become public transportation users at their destination? The case of a Mediterranean city, *Travel Behaviour and Society* 24: 10–21. https://doi.org/10.1016/j.tbs.2021.01.007

Miravet, D.; Gutiérrez, A.; Domènech, A. 2021b. Sources of data to tackle the challenges of public transport provision in seasonal tourist destination, *Transport and Sustainability* 13: 117–138. https://doi.org/10.1108/S2044-994120210000013011

Mogaji, E. 2020. Impact of COVID-19 on transportation in Lagos, Nigeria, *Transportation Research Interdisciplinary Perspectives* 6: 100154. https://doi.org/10.1016/j.trip.2020.100154

Morency, C.; Trépanier, M.; Agard, B. 2007. Measuring transit use variability with smart-card data, *Transport Policy* 14(3): 193–203. https://doi.org/10.1016/j.tranpol.2007.01.001

Moreno, T.; Pintó, R. M.; Bosch, A.; Moreno, N.; Alastuey, A.; Minguillón, M. C.; Anfruns-Estrada, E.; Guix, S.; Fuentes, C.; Buonanno, G.; Stabile, L.; Morawska, L.; Querol, X. 2021. Tracing surface and airborne SARS-CoV-2 RNA inside public buses and subway trains, *Environment International* 147: 106326. https://doi.org/10.1016/j.envint.2020.106326

Munizaga, M. A.; Palma, C. 2012. Estimation of a disaggregate multimodal public transport origin–destination matrix from passive smartcard data from Santiago, Chile, *Transportation Research Part C*: *Emerging Technologies* 24: 9–18. https://doi.org/10.1016/j.trc.2012.01.007

Mützel, C. M.; Scheiner, J. 2022. Investigating spatio-temporal mobility patterns and changes in metro usage under the impact of COVID-19 using Taipei metro smart card data, *Public Transport* 14(2): 343–366. https://doi.org/10.1007/s12469-021-00280-2

Neuburger, L.; Egger, R. 2021. Travel risk perception and travel behaviour during the COVID-19 pandemic 2020: a case study of the DACH region, *Current Issues in Tourism* 24(7): 1003–1016. https://doi.org/10.1080/13683500.2020.1803807

Ong, F.; Loa, P.; Nurul Habib, K. 2024. A behavioural analysis of post-pandemic modality profiles for non-commuting trips in the greater Toronto area, *Travel Behaviour and Society* 34: 100690. https://doi.org/10.1016/j.tbs.2023.100690

Östh, J.; Toger, M.; Türk, U.; Kourtit, K.; Nijkamp, P. 2023. Leisure mobility changes during the COVID-19 pandemic – an analysis of survey and mobile phone data in Sweden, *Research in Transportation Business* & *Management* 48: 100952. https://doi.org/10.1016/j.rtbm.2023.100952

Paulino, I.; Prats, L.; Schofield, P. 2019. Tourist hub consumption systems: convenient flexibility versus administrative constraint, *Journal of Hospitality and Tourism Management* 41: 69–79. https://doi.org/10.1016/j.jhtm.2019.09.006

Pelletier, M.-P.; Trépanier, M.; Morency, C. 2011. Smart card data use in public transit: A literature review, *Transportation Research Part C*: *Emerging Technologies* 19(4): 557–568. https://doi.org/10.1016/j.trc.2010.12.003

Pradhan, D.; Biswasroy, P.; Kumar Naik, P.; Ghosh, G.; Rath, G. 2020. A review of current interventions for COVID-19 prevention, *Archives of Medical Research* 51(5): 363–374. https://doi.org/10.1016/j.arcmed.2020.04.020

Prideaux, B. 2000. The role of the transport system in destination development, *Tourism Management* 21(1): 53–63. https://doi.org/10.1016/S0261-5177(99)00079-5

R Foundation. 2025. *The R Project for Statistical Computing*. Available from Internet: https://www.r-project.org

Rahman, M. K.; Gazi, A. I.; Bhuiyan, M. A.; Rahaman, A. 2021. Effect of COVID-19 pandemic on tourist travel risk and management perceptions, *PLoS ONE* 16(9): e0256486. https://doi.org/10.1371/journal.pone.0256486

Ram, Y.; Hall, C. M. 2018. Walking tourism in cities: introducing the special issue, *International Journal of Tourism Cities* 4(3): 281–284. https://doi.org/10.1108/IJTC-09-2018-098

Raveau, S.; Muñoz, J. C.; De Grange, L. 2011. A topological route choice model for metro, *Transportation Research Part A*: *Policy and Practice* 45(2): 138–147. https://doi.org/10.1016/j.tra.2010.12.004

Rodríguez González, A. B.; Wilby, M. R.; Vinagre Díaz, J. J.; Fernández Pozo, R. 2021. Characterization of COVID-19's impact on mobility and short-term prediction of public transport demand in a mid-size city in Spain, *Sensors* 21(19): 6574. https://doi.org/10.3390/s21196574

Rosenberg, J. M.; Beymer, P. N.; Anderson, D. J.; Van Lissa, C. J.; Schmidt, J. A. 2018. TidyLPA: an R package to easily carry out latent profile analysis (LPA) using open-source or commercial software, *Journal of Open Source Software* 3(30): 978. https://doi.org/10.21105/joss.00978

Scrucca, L.; Fop, M.; Murphy, T.B.; Raftery, A. E. 2016. Mclust 5: clustering, classification and density estimation using gaussian finite mixture models, *The R Journal* 8(1): 289–317. https://doi.org/10.32614/RJ-2016-021

Severo, M.; Ribeiro, A. I.; Lucas, R.; Leão, T.; Barros, H. 2021. Urban rail transportation and SARS-CoV-2 infections: an ecological study in the Lisbon metropolitan area, *Frontiers in Public Health* 9: 611565. https://doi.org/10.3389/fpubh.2021.611565

Shaheen, S.; Wong, S. 2021. *Future of Public Transit and Shared Mobility*: *Scenario Planning for COVID-19 Recovery*. Research Report No UC-ITS- 2021-58. Institute of Transportation Studies, University of California, Berkeley, CA, US. 78 p. https://doi.org/10.7922/G2NC5ZGR

Shen, Y.; Li, C.; Dong, H.; Wang, Z.; Martinez, L.; Sun, Z.; Handel, A.; Chen, Z.; Chen, E.; Ebell, M. H.; Wang, F.; Yi, B.; Wang, H.; Wang, X.; Wang, A.; Chen, B.; Qi, Y.; Liang, L.; Li, Y.; Ling, F.; Chen, J.; Xu, G. 2020. Community outbreak investigation of SARS-CoV-2 transmission among bus riders in Eastern China, *JAMA Internal Medicine* 180(12): 1665–1671. https://doi.org/10.1001/jamainternmed.2020.5225

Shin, H.; Nicolau, J. L.; Kang, J.; Sharma, A.; Lee, H. 2022. Travel decision determinants during and after COVID-19: the role of tourist trust, travel constraints, and attitudinal factors, *Tourism Management* 88: 104428. https://doi.org/10.1016/j.tourman.2021.104428

Tan, L.; Ma, C. 2021. Choice behavior of commuters' rail transit mode during the COVID-19 pandemic based on logistic model, *Journal of Traffic and Transportation Engineering* (*English Edition*) 8(2): 186–195. https://doi.org/10.1016/j.jtte.2020.07.002

Tang, C.; Ceder, A.; Zhao, S.; Ge, Y.-E. 2016. Determining optimal strategies for single-line bus operation by means of smartphone demand data, *Transportation Research Record*: *Journal of the Transportation Research Board* 2539: 130–139. https://doi.org/10.3141/2539-15

Thao, V. T.; Hüsser, A. P.; Ohnmacht, T. 2024. Collective or Individual? Tourists' travel mode choice during a pandemic, *Transportation Research Interdisciplinary Perspectives* 26: 101150. https://doi.org/10.1016/j.trip.2024.101150

Tang, C.; Ceder, A.; Zhao, S.; Ge, Y.-E. 2019. Vehicle scheduling of single-line bus service using operational strategies, *IEEE Transactions on Intelligent Transportation Systems* 20(3): 1149–1159. https://doi.org/10.1109/TITS.2018.2841061

Tirachini, A.; Cats, O. 2020. COVID-19 and public transportation: current assessment, prospects, and research needs, *Journal of Public Transportation* 22(1): 1–21. https://doi.org/10.5038/2375-0901.22.1.1

Trépanier, M.; Tranchant, N.; Chapleau, R. 2007. Individual trip destination estimation in a transit smart card automated fare collection system, *Journal of Intelligent Transportation Systems*: *Technology*, *Planning*, *and Operations* 11(1): 1–14. https://doi.org/10.1080/15472450601122256

Vich, G.; Gutiérrez, A.; Delclòs-Alió, X.; Tomàs-Porres, J.; Miravet, D. 2022. Presence of tourists and perceived safety from COVID-19 among local bus users: evidence from a Mediterranean city, *Transportation Research Interdisciplinary Perspectives* 14: 100599. https://doi.org/10.1016/j.trip.2022.100599

Wang, K.-Y. 2014. How change of public transportation usage reveals fear of the SARS virus in a city, *PLoS ONE* 9(3): e89405. https://doi.org/10.1371/journal.pone.0089405

Wang, Yu.; Choudhury, C.; Hancock, T. O.; Wang, Ya.; Ortúzar, J. D. 2024. Influence of perceived risk on travel mode choice during COVID-19, *Transport Policy* 148: 181–191. https://doi.org/10.1016/j.tranpol.2024.01.009

Wasserman, J. L.; Rios, N.; King, H.; Siddiq, F.; Bressette, B.; Taylor, B. D. 2022. *Transit*(*ory*) *Finance*: *the Past*, *Present*, *and Future Fiscal Effects of COVID-19 on Public Transit in Southern California*. Research Report. Institute of Transportation Studies, University of California, Los Angeles, CA, US. 82 p. https://doi.org/10.17610/T60G65

Wickham, H.; Averick, M.; Bryan, J.; Chang, W.; McGowan, L. D.; François, R.; Grolemund, G.; Hayes, A.; Henry, L.; Hester, J.; Kuhn, M.; Pedersen, T. L.; Miller, E.; Bache, S. M.; Müller, K.; Ooms, J.; Robinson, D.; Seidel, D. P.; Spinu, V.; Takahashi, K.; Vaughan, D.; Wilke, C.; Woo, K.; Yutani, H. 2019. Welcome to the tidyverse, *Journal of Open Source Software* 4(43): 1686. https://doi.org/10.21105/joss.01686

Zamparini, L.; Domènech, A.; Miravet, D.; Gutiérrez, A. 2022. Green mobility at home, green mobility at tourism destinations: a cross-country study of transport modal choices of educated young adults, *Journal of Transport Geography* 103: 103412. https://doi.org/10.1016/j.jtrangeo.2022.103412

Zaragozí, B.; Gutiérrez, A.; Delclòs-Alió, X.; Miravet, D. 2023. Exploring the public bus ridership loss and recovery during the COVID-19 pandemic: a spatiotemporal analysis using smart card data, *Journal of Map*s 19(1): 2231022. https://doi.org/10.1080/17445647.2023.2231022

Zaragozí, B.; Trilles, S.; Gutiérrez, A.; Miravet, D. 2021. Development of a common framework for analysing public transport smart card data, *Energies* 14(19): 6083. https://doi.org/10.3390/en14196083

Zhang, K.; Hou, Y.; Li, G. 2020. Threat of infectious disease during an outbreak: Influence on tourists' emotional responses to disadvantaged price inequality, *Annals of Tourism Research* 84: 102993. https://doi.org/10.1016/j.annals.2020.102993

Zhang, N.; Jia, W.; Wang, P.; Dung, C.-H.; Zhao, P.; Leung, K.; Su, B.; Cheng, R.; Li, Y. 2021. Changes in local travel behaviour before and during the COVID-19 pandemic in Hong Kong, *Cities* 112: 103139. https://doi.org/10.1016/j.cities.2021.103139

Zhang, Z.; Fu, D.; Liu, F.; Wang, J.; Xiao, K.; Wolshon, B. 2023. COVID-19, traffic demand, and activity restriction in China: a national assessment, *Travel Behaviour and Society* 31: 10–23. https://doi.org/10.1016/j.tbs.2022.11.001

Zientara, P.; Jażdżewska-Gutta, M.; Bąk, M.; Zamojska, A. 2024. Examining the use of public transportation by tourists in ten European capitals through the lens of hierarchical leisure constraints theory, *Journal of Travel Research* 64(4): 888–911. https://doi.org/10.1177/00472875241228499