

Article

# Characterization of Public Transit Mobility Patterns of Different Economic Classes

Priscila Santin <sup>1</sup>, Fernanda R. Gubert <sup>1</sup>, Mauro Fonseca <sup>1</sup>, Anelise Munaretto <sup>1,\*</sup>  
and Thiago Henrique Silva <sup>1,2</sup>

<sup>1</sup> Programa de Pós-Graduação em Engenharia Elétrica e Informática Industrial, Universidade Tecnológica Federal do Paraná, Curitiba PR 80230-901, Brazil; santin.priscila@gmail.com (P.S.); nanda.gubert@gmail.com (F.R.G.); MauroFonseca@utfpr.edu.br (M.F.); th.silva@utoronto.ca (T.H.S.)

<sup>2</sup> School of Cities, University of Toronto, Toronto, ON M5S 0C9, Canada

\* Correspondence: anelise@utfpr.edu.br; Tel.: +55-(41)-3310-4686

Received: 31 October 2020; Accepted: 16 November 2020; Published: 18 November 2020



**Abstract:** This paper analyzes public transit mobility of different economic classes of Curitiba, Brazil, exploring an official smart card dataset provided by the city. With the population divided into subsets corresponding to economic strata, we characterized vital spatial-temporal transit usage patterns, such as departure times and destinations reached by different economic classes. We also constructed a network representing the common origin and destination of public transit users, enabling discovering distinct patterns. Among the results, we observe that with the increase in wealth, the morning activity is postponed (on average for 2 h), and the spatial distribution of the trips becomes more localized compared with lower classes. We also show that our model captures fairly well realistic mobility patterns exploring a cheaper and larger-scale data source by comparing our results with a household travel survey from Curitiba. Understand how people in different economic classes appropriate urban spaces help to provide subsidies for, e.g., more sustainable economic development propositions.

**Keywords:** mobility; characterization; transit; economic stratification; networks

## 1. Introduction

The amount of data generated, collected, and shared by public administrations, companies, non-profit organizations, and even the scientific community has increased considerably in recent years [1]. The use of the knowledge extracted from these digital sources has made a significant difference in businesses and people's lives [2]. In this direction, some studies propose ways in which governments can use strategic data to serve their citizens better, addressing challenges such as health care costs, job creation, natural disasters, and urban mobility (subway [3], taxi [4], and public transit systems [5]).

Urban mobility refers to the movement of the population and goods within the cities' geographical space, constituting a component of the quality of life aspired by its inhabitants [6]. This topic has gained considerable attention internationally and also in Brazil. According to Miranda et al. [5], the leading cause of urban transit mobility problems in Brazil is the increasing use of private transport instead of public transit. However, public transit is often overcrowded in different moments of the day, thus, being, in a way, a paradox, contributing to making this challenge even harder.

While several studies on public transit mobility for different cities worldwide (Shanghai [7], New York [8], Berlin, and San Francisco [9]), Brazilian cities' studies are still scarce. Thus, this study focuses on studying public transit mobility patterns performed by users of different economic classes in Curitiba, Brazil. For that, we use official transit usage based on smart card access. Specifically,

our focus is to answer this question: what are the main usage patterns of different economic classes of Curitiba's public transit?

Similarly to Lotero et al. [10], Zhang et al. [3], and Oviedo et al. [11], our study also focuses on urban public transit mobility. Nevertheless, different from all previous efforts, we use official smart card bus data from a large Brazilian city to characterize urban transit usage. Besides, we created an origin-destination network and enriched this structure with economic Census data, enabling the study of users' flow in the city and the key places associated with these displacements. We validate our origin-destination network with an official household travel survey conducted by a government agency [12] regarding people's origin-destination in the same city, identifying a significant similarity between the two information. I.e., we obtained evidence that our approach captures quite well realistic transit mobility patterns exploring an alternative data source easier to obtain.

The main contributions of this study can be summarized as:

- Approaches for inference of the home and most common destination neighborhoods by each user, and the classification of users in different economic classes exploring Census data;
- Characterization of fundamental transit usage patterns of residents of Curitiba, Brazil, studying, for instance, departure times, distances traveled, and destinations reached by different economic classes;
- Construction and study of a network representing common origin and destination of transit users. This structure enabled uncovering relevant transit mobility patterns from users of distinct economic classes.

The remainder of the study is organized as follows. Section 2 presents the related studies. Section 3 shows some background information regarding the city of study. Section 4 presents the methodology applied in this study. Sections 5 and 6 discuss the results and the validation of our approach, respectively. Finally, Section 7 presents the conclusions.

## 2. Related Work

Urban transit mobility is a concept that can be studied through the lens of the use of private/individual transport (e.g., cars, motorcycle, and bicycles) or the use of public transportation/transit (e.g., buses, and subways). We can also study urban transit mobility at different scales, for instance, global (infectious diseases [13], epidemic scenarios [14]), continental [15], national [16] or regional (Medellin and Manizales in Colombia [10]; Belo Horizonte in Brazil [17]).

Taking our attention to regional urban mobility, we should think about how best to ensure people's access key spots of the city (e.g., business areas, schools, hospitals, and recreation areas). In this direction, Yang et al. [18] describe the integration of smart card data (SCD) of the city's subway and points of interest (POI) to study mobility in the city of Shenzhen, China. The POIs were used to semantically contextualize information derived from the SCD, allowing the purpose of the trip to be deduced. The study showed that it was possible to infer how different groups are associated with varying behaviors of travel, travel purposes, and socioeconomic activities by linking these data.

In a similar direction, Huang et al. [4] analyze the relationship between POIs and mobility networks. Based on two cities' taxi systems in China, the authors built transit networks by assigning urban regions as vertices and their connections as edges weighted by the number of mobility flows. Spatial communities were identified based on the strength of movement between regions. POIs were mapped to vertices in the network and were considered independent variables to classify the spatial community categories. Among the results, the study demonstrates the importance of data integration to a better understanding of urban mobility patterns.

Focusing on the study of the urban mobility during large scale events, Marques-Neto et al. [17] explore mobile phone records on three different types of events: a major football match, a rock concert, and a celebration of New Year, all events in Belo Horizonte, Brazil. By analyzing the spatio-temporal dynamics of the participants' movement patterns, the authors improved the understanding of human

mobility caused by large-scale events. Such analysis helped develop an application designed to help mobile operators plan their infrastructure for large events.

Still using information from mobile phones, the study of Xu et al. [19] proposes an analytical framework to better understand human mobility patterns and their relationship to the population's socioeconomic status by coupling large-scale mobile phones and urban socioeconomic datasets. In a case study conducted in Singapore and Boston, the authors compare mobility patterns' properties and analyze how they vary across socioeconomic classes. The results conclude that generally, wealthier groups of phone users tend to do shorter travels in Singapore and longer in Boston, one of the possible reasons being the location of the two cities' wealthy neighborhoods, which are respectively central and peripheral. For mobility indicators that reflect the diversity of travel patterns and individual activities, the authors concluded that the level of wealth in both cities is not a factor that restricts the way people travel, since different socioeconomic groups presented very similar characteristics. In summary, the authors suggest that the relationship between mobility and socioeconomic status may vary across cities. This relationship is influenced by spatial housing arrangement, employment opportunities, and human activities.

Few studies also investigated public transit usage, being the most similar ones to our present study. For instance, Lotero et al. [10] analyze urban transit mobility between two Colombian cities using an origin-destination survey conducted with their residents. The authors showed that spatial and temporal patterns vary among different socioeconomic groups, concluding that as wealth increases, morning activity is delayed, noon displacement becomes smoother, and travel's spatial distribution becomes localized. Zhang et al. [3] propose an analytical framework to compare urban transit mobility patterns extracted from a transit smart card dataset (bus and subway) and a dataset from Global Positioning System (GPS) taxi trajectories. Using Singapore for this case study, and using the origin-destination matrices extracted from both datasets, they conclude that spatial distributions of travel demand extracted from the two modes of transport exhibit high correlations but with a higher degree of heterogeneity regarding the use of space by public transit. The work of Oviedo et al. [11], on the other hand, uses household socio-demographic and origin-destination surveys to examine the contribution of Lima's Bus Rapid Transit (BRT) system to job accessibility in the city, especially for low-income public transport users. In a comparative analysis of two periods, before and after the implementation of the BRT system, the authors concluded that the BRT line reduced the travel time for users to reach jobs, in comparison with the city's traditional public transport. However, BRT coverage declines in areas with high concentrations of poor populations. An analysis by socioeconomic groups showed positive effects of the BRT system on accessibility for the higher income areas, which did not happen for the lower-income classes.

### 3. Background Information

This section briefly describes relevant information regarding the city studied in this work and how it is administratively structured. Also, we bring information about the economic stratification of Brazil.

#### 3.1. The City of Curitiba

Curitiba was officially founded on 29 March 1693, and is the capital of Paraná, one of the three states that compose the southern region of Brazil. The municipality of Curitiba has, according to the Brazilian Institute of Geography and Statistics (IBGE) (The Brazilian Institute of Geography and Statistics—IBGE is the primary provider of data and information about the Country, website: <https://www.ibge.gov.br/en/institutional/the-ibge.html>) estimates, 1,751,907 inhabitants in a total area of 434,967 km<sup>2</sup>. The North-South extension is 35 km, and the East-West extension is 20 km.

Administratively Curitiba is divided into 10 Regional Administrations covering all its neighborhoods. The integrated transit system serves all these neighborhoods. This division and the main bus routes is shown in Figure 1.

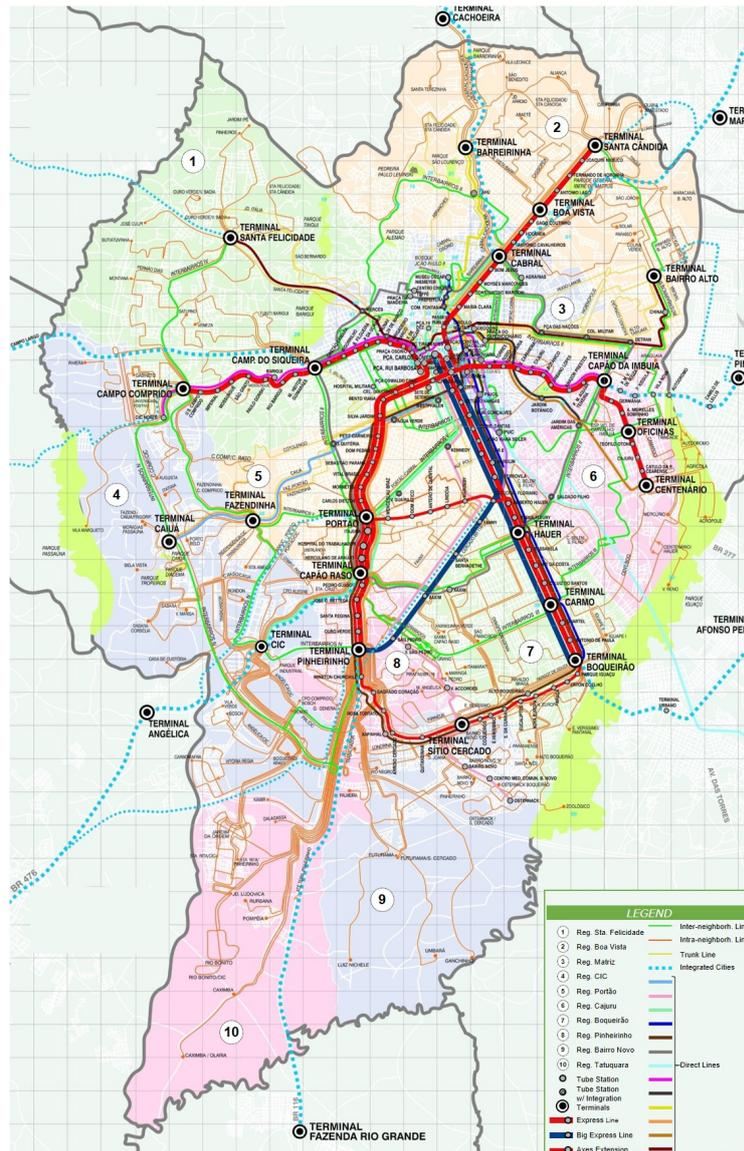


Figure 1. Neighborhoods of Curitiba and the bus routes (figure adapted from [20]).

### 3.2. Public Transit in Curitiba

The city of Curitiba has no subway system. All public transit in the city is based on the bus system, with exclusive commuting corridors where Bus Rapid Transit (BRT) runs. This BRT passes through various integration terminals, which receive local neighborhood buses allowing for system integration. The system also has inter-neighborhood lines enabling users to move from one neighborhood to another without going through city’s central region or riding BRT lines. BRT only stops at designed stations and integration terminals. Local and inter-neighborhood buses stop on specific spots on the streets and in some integration terminals.

Another essential feature of Curitiba’s transit is the integrated fare. Paying only one ticket, the citizen can compose his route, moving around the city. It is also important to mention that users only tap the smart card to enter into the system, not being registered where these users disembark from the system.

Since 2002 Curitiba has been using the smart card for public transit. The implementation of this system was necessary to reduce the cash flow in circulation in the transit system, to speed up the boarding and passage of users through the turnstiles, to discipline and measure the use of the transit

system by categories that enjoy free access and exemptions, in addition to reducing operating system costs [21].

According to Urbanization of Curitiba S.A. (URBS) (URBS is the company responsible for the strategic actions of planning, operation, and supervision involving the public transit service) statistics [22], in 2018, an average of 1,365,615 passengers was transported on weekdays. These passengers are Curitiba residents and the metropolitan area of the city, who travel to Curitiba for some activity. According to the same report, about 60.96% of the fares were paid using the smart card. The number of active smart cards in 2018 was 1,928,184, divided into three of the four existing transport categories, namely: User Card, Exempt Card, Student Card and Single Use Card (for this category, there is no information in the data provided by URBS, the user do not need to associate personal information with this card).

### 3.3. Economic Stratification in Brazil

In Brazil, there are several criteria available to classify society according to the average family income. However, in this study, we used the rule of the Secretary of Strategic Affairs (SAE), an institution from the federal government established in 2012 [23]. The criterion was based on information acquired through the 2010 Brazilian Population Census conducted by IBGE [24]. This national survey performed by IBGE aimed to portray the Brazilian population's socioeconomic characteristics and establish a basis for all public and private planning for the decade between 2010–2020.

SAE classifies groups of population in Brazil according to average monthly family income in eight economic strata. Table 1 presents a summary of this SAE grouping.

**Table 1.** Brazilian economic stratification from SAE.

Group	Average Monthly Family Income
Extremely poor (Class 1)	Up until R\$ 324
Poor but not extremely poor (Class 2)	Up until R\$ 648
Vulnerable (Class 3)	Up until R\$ 1.164
Lower middle class (Class 4)	Up until R\$ 1.764
Middle class (Class 5)	Up until R\$ 2.564
Upper middle class (Class 6)	Up until R\$ 4.076
Low high class (Class 7)	Up until R\$ 9.920
High class (Class 8)	Above R\$ 9.920

According to this classification and the estimated income per neighborhood, Curitiba presents only four of the eight classes: lower middle class, middle class, upper-middle class, and low high class. There may be divergences in intra-neighborhood family incomes. However, the classification considers the average raised during the last Brazilian population Census. The appendix of this paper (Table A1) presents the demography from Curitiba's neighborhoods according to their associated economic class.

## 4. Methodology

In this section, we describe the studied dataset and the essential methodological steps.

### 4.1. Studied Dataset

For the analysis of urban transit mobility, we used URBS data regarding the use of transit using any of the identified bus smart cards. A formal request to URBS is necessary to have access to this dataset. The dataset used in this study has daily records of the bus smart card uses, covering 69 days (53 weekdays and 16 weekend days) from September 5 to November 22 in 2018. For each entry (or use), the dataset provides Line Code, Line Name, Vehicle Code, Card Number, Date of Use, Date of Birth of the User, and Gender.

This dataset has several complementary information files, such as the existing bus lines and bus stops that the lines serve, with GPS and vehicle timetable. Some data are static and are updated once a day. Other data, such as vehicle locations, are updated almost in real-time, as vehicles send their location on average every 5 s. All of this information is made available through *.json* files.

In this study, two complementary files related to Curitiba's public transit were used:

- Vehicles—this file contains the location of all vehicles that circulated in Curitiba on the date before its disclosure. The information contained in this file is distributed in five fields: Vehicle Code, Latitude, Longitude, Usage Date, and Line Code.
- Points—a file that contains the information of all existing bus stops in Curitiba. This file has nine information fields: Point Name, Point Number, Latitude, Longitude, Sequence, Group, Direction, Type, and Line Code.

### 4.2. Dataset Preprocessing

This stage's main objective was to establish the origin and destination information of users of public transit. Since only the entry into the system is recorded, as a result of this preprocessing, we kept only the data from which we were able to infer the origin and destination based on the usage history of each smart card.

Figure 2 shows data preprocessing steps, from obtaining raw data to creating the database used in our study. Each of these steps is detailed below.

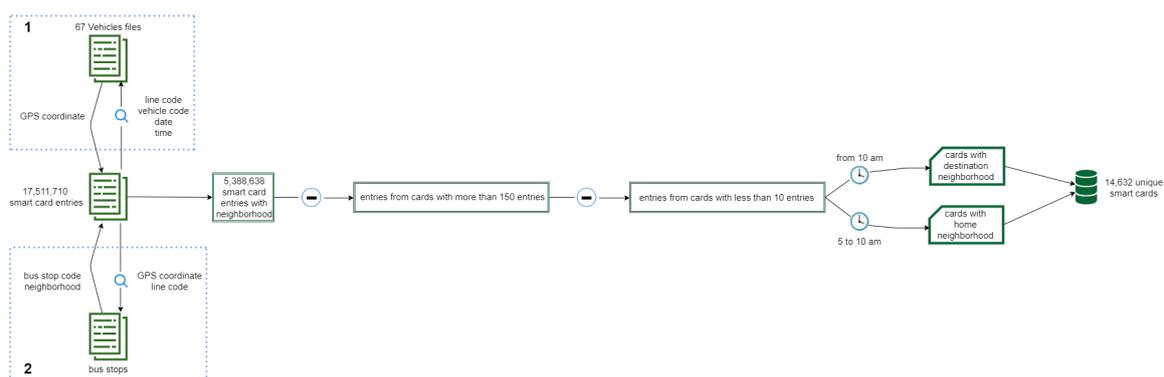


Figure 2. Flowchart of data preprocessing.

We used 58 files about smart card use and 67 Vehicles files. The smart card use represents in total 17,511,710 records, and the vehicles file 129,461,200 records.

To infer the neighborhood where each smart card was used, we used different approaches. When users enter the transit system at designed stations or integration terminals (exclusive bus stop for BRT and some special lines), the bus line code and the vehicle code linked to this entrance are directly related to the bus stop location. Therefore, the identification of the neighborhood happens

with high precision. When user entrance happens at regular bus stops, we search for a match using line code, vehicle code, date and time, to obtain a GPS coordinate for each smart card entry.

Although there are cases where the GPS position recording interval is longer, on average, the GPS coordinates updates of the buses are recorded around every 5 s, making the data more detailed than needed for our purpose: infer the approximate area where the smart card was used. Thus, we keep only one update per minute for each vehicle, since within one minute, it is considered that the bus has not traveled a considerable distance. After this filter, the new vehicle information represents 18,594,495 records, which enable the inference procedure to be executed more efficiently without compromising the quality considerably.

Here it is important to comment that when matching the line code, vehicle code, date and time attributes between the base of the smart cards and the bus records, it was observed that some buses had the line code and/or the vehicle code filled in with a generic value, and it is not possible to guarantee the identification of the GPS location of the card entry. In these cases, those records have been deleted.

After enriching a smart card entry with latitude and longitude, it was associated with the neighborhood information of the particular bus stop. To make this procedure, first, the latitude and longitude inferred were matched with the nearest bus stop of a specific line of the smart card entry. This procedure was done using the *R Tree* data structure. The *R tree* was proposed by Antonin Guttman and is widely used as a space access method, allowing the indexing of multi-dimensional information as geographic coordinates [25]. At the end of this step, we could infer the neighborhood for 5,388,638 smart card entries.

Some people lend their smart cards, with or without financial advantages, to others who are traveling on the same bus line, generating noise in the records. Some entries are made within minutes on the same lines in these cases, generating a lot of entries for the same smart card. To minimize those cards' impact in our analysis, we decided to delete cards with more than 150 entries. The rationality behind this value is that the study period is 69 days (16 days of weekends) and that each user typically uses at least two tickets per day. In a similar reasoning line, cards with less than ten entries were deleted during the analysis period. They represent a small number; on average, one entry per week. By disregarding them, we also aim to make the results more robust. After these exclusions, the number of entries with neighborhood became 5,225,573.

Then we separate the entries into two groups: home neighborhood and destination neighborhood. It was observed in the data analysis a significant growth in the number of entries around 5 a.m. This growth reaches its peak at around 7 a.m., and the phenomenon ends at 10 a.m. Because of this scenario, we considered for the home neighborhood identification entries between 5 a.m. and 10 a.m., as this is a time frame in which users probably use transit to go to work or study. We consider the destination neighborhood only the entries made after 10 a.m. Recall that users in Curitiba only are obliged to use their cards to entry in the system. Thus, one reasonable way to identify the destination point is to assume that a returning bus stop, entries after 10 a.m. in our case, is nearby where the user got off from the home neighborhood.

In possession of the smart card use made between 5 a.m. and 10 a.m., we grouped data by card and selected as the home neighborhood the most frequent one. We also grouped data by card entries after 10 a.m. and applied the same threshold and procedure to identify the most frequent destination neighborhood. By analyzing the number of entries per user in both scenarios, we could verify that these thresholds preserve a significant number of smart cards, not compromising the analysis. After this procedure, we got 14,632 unique smart cards containing home and most typical destination neighborhoods. These neighborhoods' economic classes were added according to the classification presented in the appendix of this work (Table A1).

Despite the possible limitations of our approach, we got evidence that it is a good approximation. We constructed a set of manually classified smart card entries, based on interviews with the smart card owner. In this way, we have got the home neighborhood and the most common visited neighborhoods.

This information was gathered from 8 users, from different economic classes and different city mobility patterns.

Our approach identified all the information correctly for this dataset. The construction of this dataset is very challenging for several reasons, such as volunteer identification and agreement with disclosing private information; however, this experiment is vital to help us get a sense of our approach's quality.

#### 4.3. Network and Metrics

A possible tool for representing and studying mobility patterns is a graph, which consists of the mapping system's elements into vertices and edges to form networks (human urban mobility [26]; social networks [27]). It is an efficient way to model and explain the mechanisms of collective behavior [28]. Several studies on mobility patterns have demonstrated network relevance for mobility studies (identification of communities [29]; city zoning [30]; demand for urban passenger transportation [31]).

To analyze Curitiba's public transit mobility, we use the origin-destination information, described in Section 4.2, to construct urban transit mobility networks. We used as origin the inference of home neighborhood and as the destination the most frequented neighborhood of each smart card. We created a matrix  $M$  of the order  $N \times N$ , where  $N = 74$ , that is, the neighborhoods of Curitiba.

The  $M_{ij}$  elements of the matrix indicate the number of trips observed from origin node  $i$  to destination node  $j$ . With this information, we can calculate the strength of node  $i$ , i.e., the number of trips from node  $i$ . The strength of a node is represented by Equation (1).

$$s_i = \sum_{j=1}^N M_{ij} \quad (1)$$

Another analysis we can do with source-destination information is how connected the city's neighborhoods are. For this analysis, we use the adjacent matrix  $A$ , so  $A_{ij} = 1$  if at least one trip of  $ij$  was observed while  $A_{ij} = 0$ , otherwise. With the matrix  $A$ , we can calculate others characteristics of this network, such as the degree of the nodes, represented by Equation (2), and the shortest path between nodes  $i$  and  $j$ , that in this case, is defined as the minimum sequence of  $n$  nodes  $(i, l_1, l_2, \dots, l_{n-1}, j)$ . So,  $A_{il_1} = A_{l_1 l_2} = \dots = A_{l_{n-1} j} = 1$ , which also defines the distance between  $i$  and  $j$  as  $d_{ij} = n$ .

$$k_i = \sum_{j=1}^N A_{ij} \quad (2)$$

The joining of these two matrices,  $M$  and  $A$ , allows visualization of the established transit mobility network that can be observed in Figure 3, where nodes sizes represent their strength and the thickness of connections represent their degree.

#### 4.4. Discovery of Patterns by Economic Classes

To observe the transit mobility separated by each economic class, we use some averages of measures previously introduced. The first one corresponds to the average degree ( $k$ ) of the transit mobility network:

$$k_\alpha = \frac{1}{N_\alpha} \sum_{i=1}^{N_\alpha} k_{\alpha_i}, \quad (3)$$

where  $\alpha$  ( $\alpha = 4, \dots, 7$ ) represent one of each class or the whole network and  $N_\alpha$  the number of non-isolated nodes in the class  $\alpha$ . The second average is related to strength ( $s$ ):

$$s_\alpha = \frac{1}{N_\alpha} \sum_{i=1}^{N_\alpha} s_{\alpha_i}. \quad (4)$$

Other measure is the average path length ( $d$ ):

$$d_{\alpha} = \frac{1}{N_{\alpha}(N_{\alpha} - 1)} \sum_{i=1}^{N_{\alpha}} \sum_{j>1}^{N_{\alpha}} d_{\alpha_{ij}} \quad (5)$$

To analyze the urban displacement performed by individuals from each of the economic classes, we used the radius of gyration  $r_g$  [32], calculated according to the Equation (6):

$$r_g = \sqrt{\frac{1}{n} \sum_{i=1}^n (r_i - r_{cm})^2} \quad (6)$$

where  $n$  is the number of card entries,  $r_i$  is the geographical location of the entry given by latitude and longitude and  $r_{cm}$  is the center of mass given by the midpoint between all card entries. The Euclidean distance was used to calculate the distance between entry and the center of mass ( $r_i - r_{cm}$ ).

Finalizing the measures for the discovery of patterns for each economic class, we calculate the betweenness centrality to identify the places' centrality. The betweenness centrality of a node  $v$  is the sum of the fraction of all-pairs shortest paths that pass through  $v$  [33]:

$$c_B(v) = \sum_{s,t \in V} \frac{\sigma(s,t|v)}{\sigma(s,t)} \quad (7)$$

where  $V$  is the set of nodes,  $\sigma(s,t)$  is the number of shortest ( $s,t$ ) - paths, and  $\sigma(s,t|v)$  is the number of those paths passing through some node  $v$  other than  $s,t$ . If  $s = t$ ,  $\sigma(s,t) = 1$ , and if  $v \in s,t$ ,  $\sigma(s,t|v) = 0$ .

## 5. Results

To better understand the use of urban space through public transit, some analyzes are necessary. This section presents the analyzes performed in this direction.

### 5.1. Users Economic Information

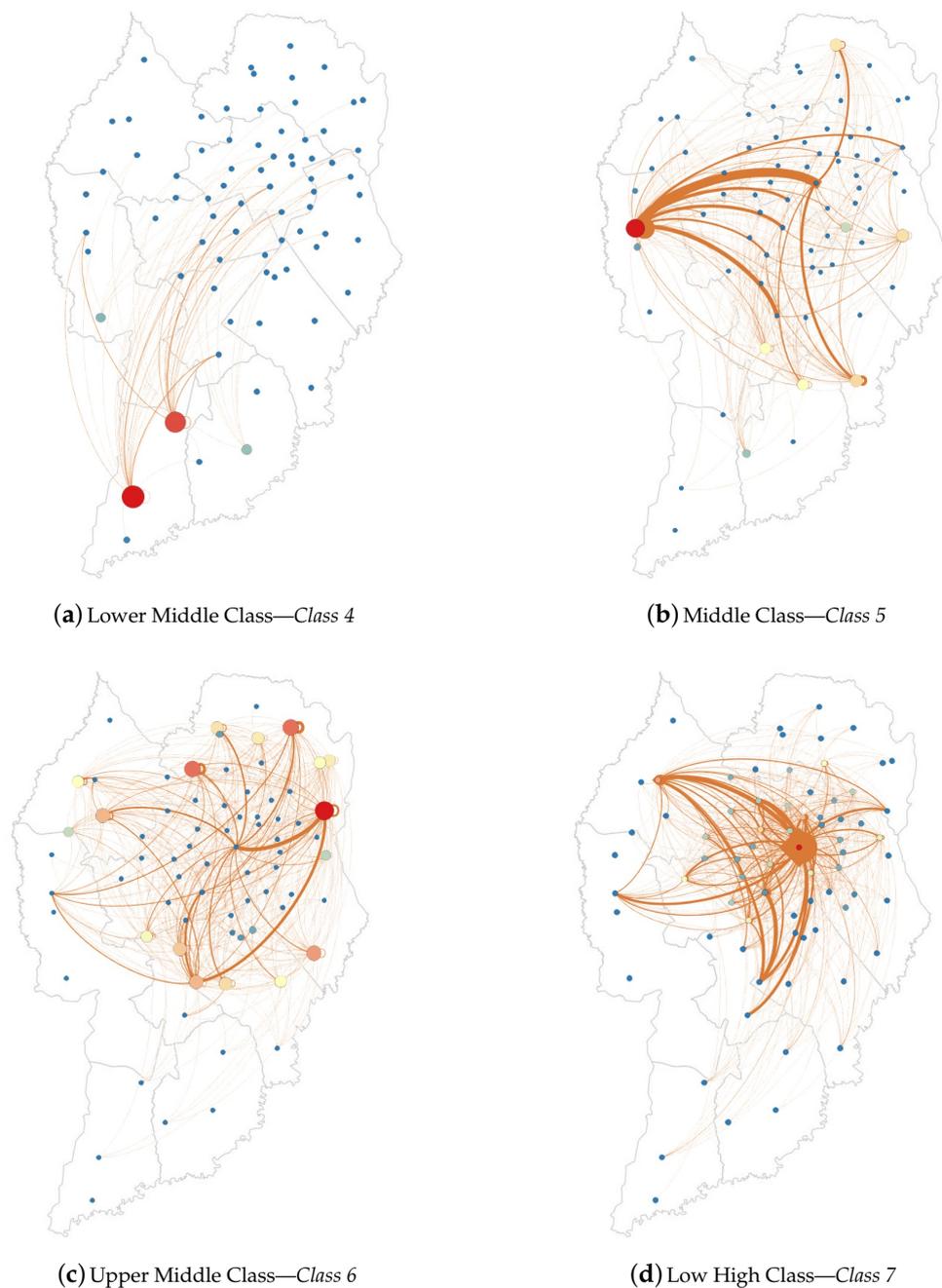
We start our analysis by understanding the economic characteristics regarding the population under study for better modeling of transit mobility performed by these different classes. Table 2 presents in a summarized way the distribution of the population of Curitiba according to their economic class and the number of distinct smart card users per class, with the total trips analyzed in this study. More detailed information on Curitiba's population can be found in the table available in the appendix (Table A1).

**Table 2.** Distributions of the official population and sample used in this study according to their economic class.

	Curitiba	Class 4	Class 5	Class 6	Class 7
Official population	1,751,907	97,910	530,683	602,653	520,661
%		5.59%	30.29%	34.40%	29.72%
Distinct smart card users	14,632	330	3412	3826	7064
%		2.26%	23.32%	26.15%	48.28%
Trips	721,910	14,966	163,776	180,425	362,743
%		2.07%	22.69%	24.99%	50.25%

### 5.2. Stratified Transit Mobility Networks

We now turn our attention to the mobility performed by the studied users. Figure 3 shows our public transit mobility networks corresponding to classes 4, 5, 6, and 7 of Curitiba (the only ones observed in the city according to the 2010 Census). All networks contain all nodes, that is, all Curitiba's neighborhoods, and they are positioned in the images according to their actual positions in the city. However, we do not observe necessarily activities in all neighborhoods for all classes. There are cases where the class  $\alpha$  ( $\alpha = 4, \dots, 7$ ) neither live nor visit other city neighborhoods. In this case, the effective number of non-isolated nodes in the transit mobility network of class  $\alpha$  is  $N_\alpha < N$ . The edge's size reflects their weight, i.e., thicker ones are heavier ones. In each class represented, the nodes of larger sizes are those with greater strength and belong to the studied class.



**Figure 3.** Public transit mobility networks by economic classes.

With the help of Figure 3, we can see very distinct transit mobility patterns among all classes. Thus, making each transit mobility network by economic classes having different structural properties. Some of the main properties are presented in Table 3, considering the aggregated network (observed all transit mobility) and separated by each economic class. We show the total number of non-isolated nodes ( $N$ ), the total number of edges and the average measures introduced above.

**Table 3.** Network properties of public transit mobility in the city of Curitiba.

	Total	Class 4	Class 5	Class 6	Class 7
Number of nodes	74	31	69	65	67
Number of edges	14632	330	3412	3826	7064
Average degree— $k_\alpha$	21.227	1.912	4.028	8.014	10.269
Average strength— $s_\alpha$	195.093	4.400	45.493	51.013	94.187
Aver. path length— $d_\alpha$	1.751	1.000	1.657	1.657	1.767

Analyzing Table 3, we can quantify the apparent differences observed visually between the existing transit mobility networks for the economic classes of Curitiba. Although class 7 has almost twice as many unique users in our dataset compared to class 6, both have a high value of average degree  $k_\alpha$ , proving to be the most connected networks. That is, users in these economic strata tend to visit a higher diversity of neighborhoods. If we compare classes 5 and 6, which have practically the same number of unique users in our dataset and the same average strength  $s_\alpha$ , class 5 has a much less connected network when analyzing its average degree. Looking at the average path length  $d_\alpha$  by the classes, we find that class 4 trips are more concentrated compared to the other classes, especially with class 7, corroborating with what it is observed in Figure 3.

### 5.3. Temporal Patterns of Urban Transit Mobility

This section presents results about the temporal patterns of public transit mobility in Curitiba by different economic classes. For these analyses, we considered all trips made for the 14,632 unique users in our dataset.

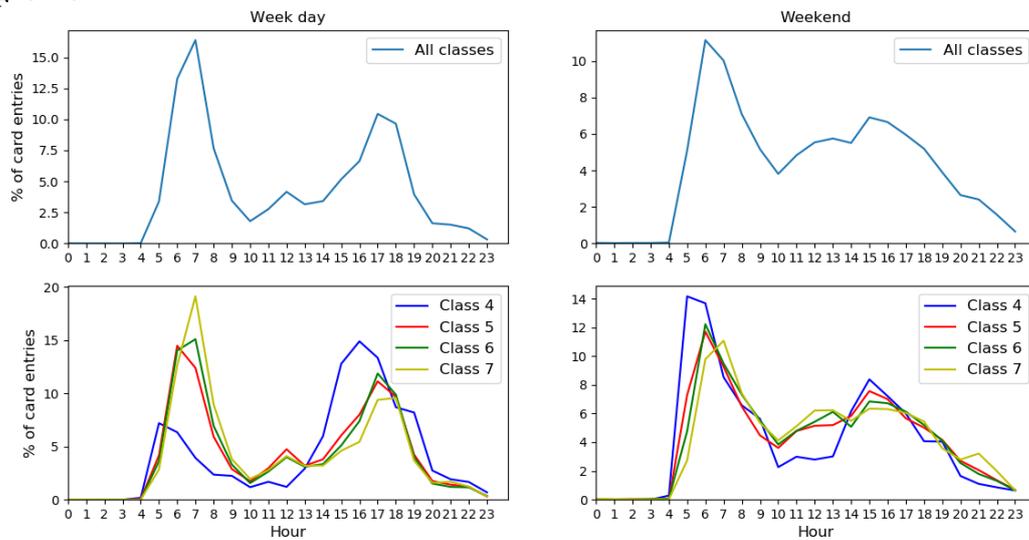
Analyzing Figure 4, which shows the distribution of the number of smart card use throughout the day, it is possible to observe two peaks during the weekdays, one between 5 a.m. and 7 a.m. and another between 4 p.m and 6 p.m. At midday, there is a third peak, but irrelevant compared to the other two. If we look at the weekend, there is a slightly different behavior, with a peak between 5 a.m. and 7 a.m., and then a somewhat uniform distribution in the remainder of the period, except for class 4, which has a very sharp curve.

Looking at the results for individual classes during weekdays in more detail, we observe that the early-morning peak delays as the class increases. Class 4 peaks at 5 a.m., classes 5 and 6 at 6 a.m., and class 7 at 7 a.m. At midday, there is a small peak, but class 4 does not appear in it. Similar behavior in the morning is observed in the afternoon. The lower class has its earlier displacement peak. While in class 4 the peak occurs at 4 p.m., for class 7 this peak is at 6 p.m. It is important to note that the peaks windows for class 4 is wider in this period, showing greater dispersion in transport utilization for this class.

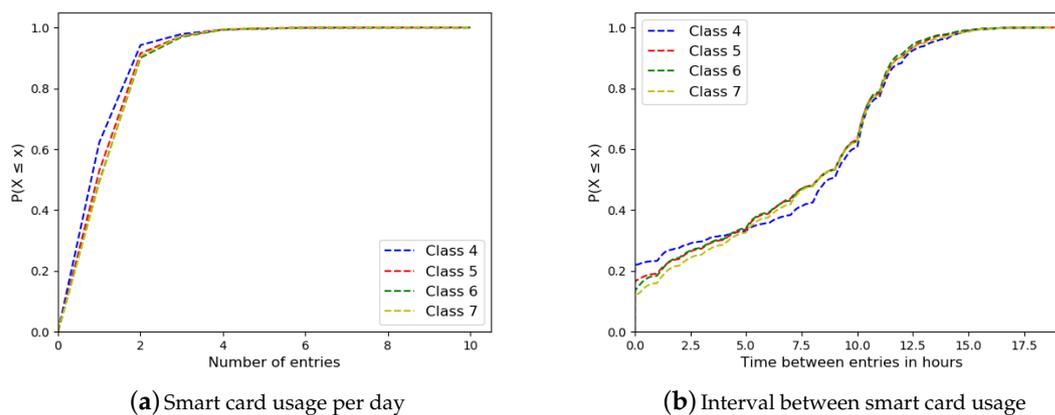
For weekends we can observe a very different behavior for class 4 compared to the other classes. Class 4 behaves similarly on weekdays and weekends but anticipates its afternoon peak to 3 p.m. This characteristic can be attributed to the fact that Saturday is also a working day and labor in industries, for most users of class 4, which may not be the case for users of other classes. Classes 5, 6, and 7 have the same behavior observed during weekdays and what changes are later. From 10 a.m., these three classes make a homogeneous displacement, having a small peak at 3 p.m. as well.

For all classes, the vast majority of smart card users use the system only twice a day (more than 80%—Figure 5a), with an interval of more than 10 h between these uses (Figure 5b). This underscores

the hypothesis that public transit users use the system for displacements for specific purposes, such as going to work.



**Figure 4.** Distribution of the number of smart card use throughout the day. The first row refers to all users, i.e., aggregated results, while the second row shows individual classes results.



**Figure 5.** Smart card usage details.

#### 5.4. Spatio-Temporal Patterns of Urban Transit Mobility

Now we analyze spatio-temporal characteristics of urban movements observed in the studied dataset. We perform aggregation of smart card use in three-time intervals: morning from 5 a.m. to midday, afternoon from noon to 6 p.m. and evening from 6 p.m. to midnight. The period between 0 a.m., and 5 a.m. was disregarded because the number of trips is negligible in this interval, representing only 0.05% of the total. In this period, Curitiba's public transit works with a minimal number of lines in operation, only 20, representing less than 10% of the lines in action during the day.

Figure 6 shows heatmaps of the smart card's geographical location each time interval by individuals from the four economic classes identified in Curitiba. According to the heatmaps, the city's central region receives the largest number of individuals, in line with the fact that it is the most economically active region in the city and has easy access to other regions through public transit. Except for class 4, all other classes have an intra-class similarity between the afternoon and evening periods, which indicates a very similar geographic distribution of permanence for these classes. In the case of class 4, where the neighborhoods that compose this class are geographically more distant from the other regions, there is a large displacement in the morning, with a more uniform distribution in the afternoon and night. This phenomenon can be justified by analyzing that, since it is longer

distances, there is longer departure time, which prevents displacements between origin and destination at intermediate times during the day.



Figure 6. Spatio-temporal patterns of urban transit mobility. Values range from light blue (low values) to red (high values).

### 5.5. Urban Displacement

In this section, we analyze the urban displacement performed by individuals from each of the economic classes. Using the radius of gyration  $r_g$  and calculating the distances in kilometers, according to Figure 7, it is possible to observe that individuals make shorter displacements with the increase in wealth.

By this figure, it is possible to verify that classes 7 and 4 are very distinct. If we take the measure at 50% for each class as a base, we observe that class 7 travels around 4 km, class 4 has a displacement almost double, 8 km. For classes 5 and 6, in 50%, all radius of gyration is approximately up to 4.2 and 4.5 km, respectively, varying considerably less than the other classes.

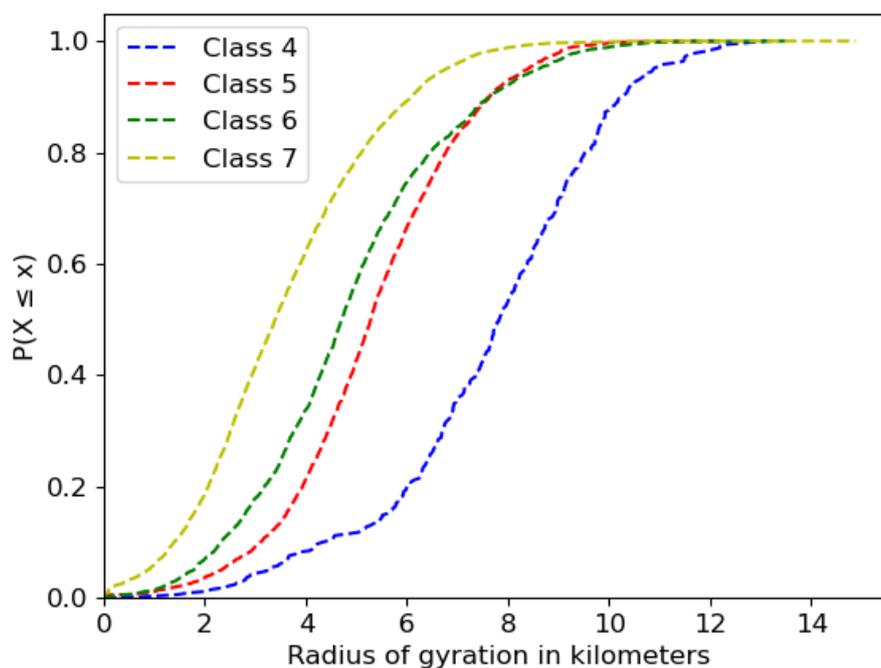


Figure 7. Displacement by class.

### 5.6. Centrality of Places

This section analyzes the centrality measures mentioned above on the transit mobility networks constructed in our study. We can see from Table 4, according to the metric *in-strength* (*i-s*) “Centro” (CE) neighborhood, which is located in the central region of Curitiba, is the neighborhood with the highest network centrality in almost all classes. This is because it is towards this region that occurs the largest displacement of the smart card users studied. However, this is not the case for class 4, where “Centro” is the second most important neighborhood (Top 2), losing to “Pinheirinho”. That is because “Pinheirinho” contains a practical terminal for users of class 4 to move throughout the city. This terminal located in Pinheirinho enables fast access to all regions of Curitiba through the exchange of buses, and, for this reason, it is a popular terminal in the city [22]. As this transfer is done inside this bus terminal, there is no need to pay a new fare for the trip and, consequently, there is no record of the use of the smart card. Note that, for these cases, our approach, which relies on smart card activities, do not capture those buses changes. However, this does not compromise the analysis because we are interested in the main neighborhoods where people perform activities, not just transit through it.

Analyzing the *out-strength* (*o-s*) metric, we observe that for classes 4 and 5, the largest outbound displacement occurs in the neighborhood with the largest population for these classes (as can be seen in the Appendix A—Table A1). For classes 6 and 7, this phenomenon is different. In class 6, the largest outbound displacement neighborhood for these classes is the Top 4 in terms of population, whereas,

in class 7, it is the Top 3. In other words, in both classes, users from these classes who use public transport the most (using smart cards) do not belong to the most populous neighborhoods in these classes; however, they do belong to neighborhoods with lower family income.

**Table 4.** TOP 10 centrality place by class—Abranches (AB), Água Verde (AV), Alto Boqueirão (AL), Augusta (AG), Bairro Alto (BA), Barreirinha (BR), Batel (BT), Butiatuvinha (BU), Cachoeira (CH), Cajuru (CJ), Campo Comprido (CM), Campo de Santana (CS), Capão Raso (CR), Caximba (CA), Centro (CE), Centro Cívico (CC), Cidade Industrial (CI), Fazendinha (FA), Ganchinho (GA), Mercês (ME), Novo Mundo (NM), Pilarzinho (PI), Pinheirinho (PN), Portão (PO), Prado Velho (PV), Rebouças (RE), Riviera (RI), Santa Cândida (SC), Santa Felicidade (SF), Santa Quitéria (SQ), Santo Inácio (IN), São Braz (SB), São Francisco (SA), São Miguel (SM), Sítio Cercado (SI), Tarumã (TR), Tatuquara (TA), Uberaba (UB), Umbará (UM), Vista Alegre (VA), Xaxim (XA).

Class	TOP	1	2	3	4	5	6	7	8	9	10
All Classes	In-Strength	CE	CI	CR	AV	BA	SF	CC	PO	BT	ME
	Out-Strength	CE	CI	SF	BA	AV	CC	SC	AL	PI	CR
	Betweenness	CE	CI	BA	AV	CC	SF	CR	PN	ME	NM
Class 4	In-Strength	PN	CE	CI	RE	SQ	TA	AV	CC	TR	BA
	Out-Strength	TA	CS	SM	GA	CA	PN	CE	CI	RE	SQ
	Betweenness	-	-	-	-	-	-	-	-	-	-
Class 5	In-Strength	CE	CI	CR	AV	AL	PO	BT	BA	CC	PN
	Out-Strength	CI	AL	CJ	CH	SI	PV	PN	UM	AG	RI
	Betweenness	CI	PN	CJ	CH	SI	AB	PV	UM	-	-
Class 6	In-Strength	CE	CR	CI	BA	AV	SC	PI	PO	CC	SF
	Out-Strength	BA	SC	PI	CR	SB	BU	XA	AB	UB	FA
	Betweenness	BA	PI	CR	SC	UB	NM	SB	XA	AB	BR
Class 7	In-Strength	CE	CR	AV	SF	CI	ME	CC	BT	BA	PN
	Out-Strength	CE	SF	AV	CC	VA	CM	ME	PO	BT	IN
	Betweenness	CE	AV	CC	SF	ME	SA	TR	PO	BT	RE

Looking now at the *betweenness* (*bet*) centrality, we can see that for almost all classes, the most relevant neighborhoods are also the neighborhoods that have the largest *out-strength* centrality. Analyzing classes 5 and 6, the neighborhoods with the highest *betweenness* (Top 1) also have bus terminals with high demand [22], enabling transfers without paying a new fare to reach all city regions quickly. This makes these neighborhoods important bridges in urban displacement. For class 7, “Centro” appears again as the central neighborhood, now in the *betweenness* metric. This may be justified by the number of bus lines circulating in the city center, facilitating the movement. It is also interesting to note that, for class 4, all neighborhoods had a *betweenness* equal to zero.

All central places uncovered with different metrics and for different classes can be useful in different ways. The *in-strength* and *out-strength* measures allow a broader understanding of urban transit mobility, as they show the locations (neighborhoods) where users are coming from and where they are going. Note that central places’ indication when looked at different classes is different from the one obtained by the aggregated network. This information helps to understand the mobility demands of different economic classes, contributing, for example, to better urban interventions in the public transit system.

As for *betweenness*, which identifies important bridge places for different areas of the network, it is a useful metric to help implement actions aimed at the city’s population at the strategic points indicated by such measure. These actions could range from strategic information dissemination to

vaccination campaigns, to mention a few. Note that these strategies could be adapted to reach specific economic classes, based on the information obtained by our approach.

## 6. Validation

This section aims to validate our approach with other official data as well as a random graph model.

### 6.1. Household Travel Survey Data

In 2016, the Curitiba Institute for Urban Planning and Research (IPPUC) conducted a household travel survey [12] (IPPUC survey) to draw an overview of transportation displacements and demands in Curitiba and its metropolitan region and allow greater security in defining strategies on the transit mobility of the capital.

The research was carried out through three distinct stages: home interviews, volumetric counting, and vehicle speed measurement, as well as an opinion poll, where residents could assess the quality of public transit and the road system as a whole. We explore this data to create an origin-destination network, as we created by the dataset of smart card access investigated in this study.

### 6.2. Random Graph Model

We also introduced a random graph model that creates a network with the same structure approached until then. This random model has the same nodes as our graph, the 74 neighborhoods. The edges that connect the neighborhoods were generated randomly, respecting the same number of edges in our graph (14,632). To study the similarity between the graphs, 100 random models were generated.

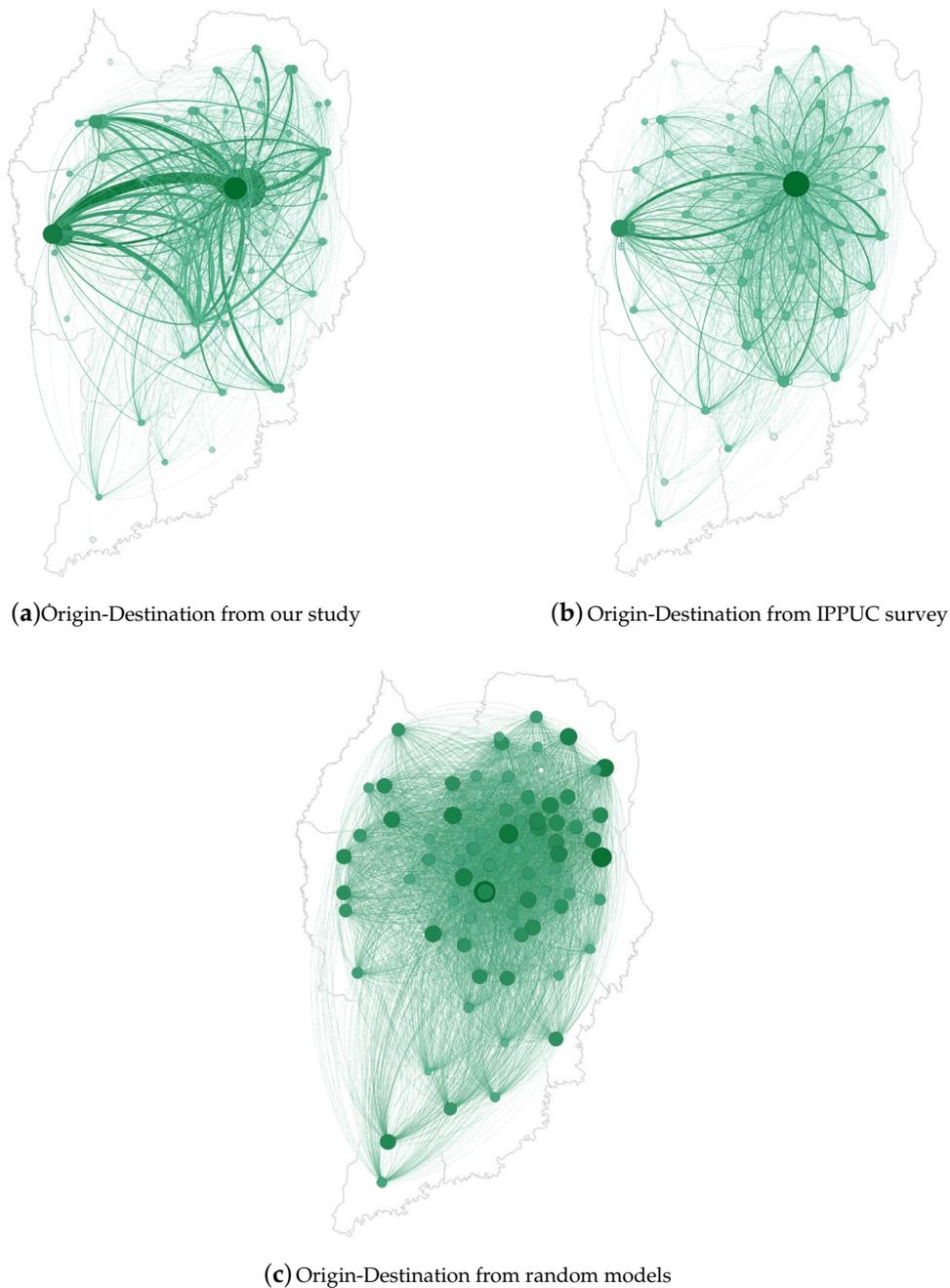
### 6.3. Comparisons

In this section, we compare the three origin-destination networks. Figure 8 provides a visual comparison of the displacement identified by these three networks.

We can see that the network established with the IPPUC survey data is more connected than our study. Analyzing the data, we observed that in the home interviews phase, the interviewees informed that they had to travel using public transit in more than one stage; that is, all connection (steps) used have been reported.

Curitiba has a public transit system that allows the transfer at transit terminals and stations all to different buses paying only one ticket, i.e., registering only once with the smart card. Because our research uses the information from these smart card records in the system, we were unable to observe these steps mentioned in the IPPUC survey (i.e., the average of these steps ranges from 1.02 to 1.92 in the neighborhoods of Curitiba). For our study, the displacement appears in only one step, since the dataset does not present the richness of detail that an interview can extract. In the network produced from the IPPUC survey, all stages from origin to the destination were represented. For a better validation of the results, we performed some quantitative analyzes. Observing the graph of Figure 9a, which shows the distribution of node degree centrality, it is possible to identify this difference: a higher degree for the IPPUC survey network precisely because of the interviewees' details.

From the random model, we observed a very different behavior, demonstrating that what is captured by our approach is far from random models. This difference provides us with the necessary background to conclude that our methodology was well structured.



**Figure 8.** Comparison of Curitiba's transit mobility networks.

Comparing the betweenness presented in Figure 9b further promotes the observed difference between the displacement pattern reported in the IPPUC survey and the pattern observed in our study. The betweenness curve of our study is larger and has higher vertices precisely because it does not demonstrate all the displacement steps reported in the IPPUC survey. Again, from the random model, the behavior of the betweenness was different comparing the other cases, as expected.

To verify the similarity between these three established networks, we used the Pearson's correlation coefficient and the Root Mean Square Error (RMSE). Using the Pearson's correlation coefficient to compare our study and the IPPUC survey, we obtained the value of 0.4582 ( $p$ -value < 0.001), indicating that both studies capture a similar phenomenon and, potentially, can be complementary. Comparing our network and the random model, we obtained a Pearson's correlation of 0.0064 (with a confidence interval of  $\pm 0.0056$ ).

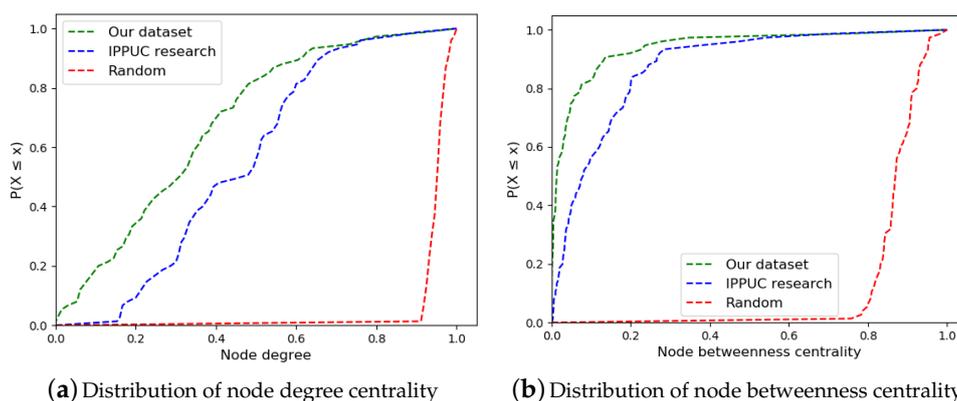


Figure 9. Comparison of urban transit mobility network metrics.

The correlation between the IPPUC survey network and the random model was 0.0114 (with a confidence interval of  $\pm 0.0081$ ). Pearson’s correlation between the random network and the other models showed the lack of linear correlation in both cases.

We now use the RMSE to assess the similarity between the networks. Comparing our approach with the IPPUC survey the value obtained was 17.27. This value indicates, again, that both networks capture a similar phenomenon. In the comparison between our approach and the random network, the RMSE value was 132.0031. In the comparison between the IPPUC survey and the random model, the RMSE value was 130.4139. As with Pearson’s coefficient, the RMSE between the random network and the other models showed the lack of similarity between them.

### 7. Final Discussions and Conclusions

This work presents a study on public transit urban mobility in Curitiba based on data available from the city’s public agencies. We analyze the geographical distribution of travel origin and destination and the temporal pattern of these trips by adding information about smart card users’ economic status. A summary of the results can be seen in Table 5.

Table 5. Summary of analysis results.

Results	Class 4	Class 5	Class 6	Class 7
<b>A. Users Economic Information</b>	2.26%	23.32%	26.15%	48.28%
<b>B. Stratified Transit Mobility Network</b>	trips more concentrated	almost the same number of users comparing with class 6, but less connected network	most connected network	most connected network
<b>C. Temporal Patterns of Urban Transit Mobility</b>	1st access to the public transit system at 5 a.m.	1st access to the public transit system at 6 a.m.	1st access to the public transit system at 7 a.m.	1st access to the public transit system at 7 a.m.
<b>D. Spatio-Temporal Patterns of Urban Transit Mobility</b>	a large displacement in the morning, with a more uniform distribution in the afternoon and night	an intra-class similarity between the afternoon and evening periods	an intra-class similarity between the afternoon and evening periods	an intra-class similarity between the afternoon and evening periods
<b>E. Urban Displacement</b>	80% of users travel more than 9 km	80% of users travel approximately 6 km	80% of users travel approximately 6 km	80% of users travel less than 4 km
<b>F. Centrality of Places</b>	<i>i-s</i> : “Pinheirinho” neighborhood, contains a practical terminal for users of class 4 to move throughout the city; <i>o-s</i> : the largest outbound displacement occurs in the neighborhood with the largest population for this class, “Tatuquara”; <i>bet</i> : all neighborhoods had this metric equal to zero.	<i>i-s</i> : “Centro” neighborhood, region that occurs the largest displacement of the smart card users studied; <i>o-s</i> : the largest outbound displacement occurs in the neighborhood with the largest population for this class, “Cidade Industrial”; <i>bet</i> : “Cidade Industrial” neighborhood has a bus terminal enabling transfers without paying a new fare to reach all city regions quickly.	<i>i-s</i> : “Centro” neighborhood, region that occurs the largest displacement of the smart card users studied; <i>o-s</i> : “Bairro Alto” is not the most populous neighborhood for this class, but it is one of the neighborhoods in this class with the lowest family income; <i>bet</i> : “Bairro Alto” neighborhood has a bus terminal enabling transfers without paying a new fare to reach all city regions quickly.	<i>i-s</i> : “Centro” neighborhood, region that occurs the largest displacement of the smart card users studied; <i>o-s</i> : “Centro” is not the most populous neighborhood in this class, but it is one of the neighborhoods in this class with the lowest family income; <i>bet</i> : the number of bus lines circulating in “Centro” neighborhood facilitates the movement of users.

The transit mobility networks associated with each of the economic classes are quite different. The most connected networks are represented by classes 6 and 7 (the higher-income brackets). Although class 5 has the same strength as class 6, it connects to fewer neighborhoods. Class 4, on the other hand, makes smaller displacements if we compare path length moved by its users.

By observing travel's temporal behavior, approximately 80% of users use the public transit system up to twice a day, with an average interval of 10 h. There are two peaks in demand throughout the day; one very early and one in the late afternoon. During weekdays these peaks are delayed as economic class increases. This behavior shows that higher classes may have greater flexibility in their working hours, which does not happen in a factory production line, for example, where many employees tend to belong to class 4. It is also possible that class 4 has more people doing informal jobs, but these are hypotheses that this work does not capture. For weekends, while class 4 presents the same behavior compared to weekdays, classes 5, 6, and 7 change their transportation use time, being practically constant after 10 a.m.

In a spatio-temporal analysis, it was observed that the city's central region is the most accessed, regardless of economic class. We found that higher classes move less according to the radius of gyration, while 80% of individuals in class 7 travel distances less than or equal to 5 km, the same amount in class 4 reaches 9 km. The other classes (5 and 6) show a significant similarity of displacement in two of the three observed periods (afternoon and evening).

Based on the network centralities, we observe that for classes 4 and 5, the most populous neighborhoods of these classes are also the most important in terms of out-strength and betweenness. In contrast, for classes 6 and 7, this situation is slightly different. The most relevant neighborhoods are not among the Top 3 in terms of population, but they are among the Top 7 in the lowest family income category. In this study, "Centro", a neighborhood in the city's central region, represents the most relevant one in all metrics. A special note for this neighborhood is a high intra-neighborhood displacement, justified by the number of bus lines in this region, facilitating the displacement.

Finally, when comparing our study with the recently conducted research with some city dwellers and a random model, we observed significant similarities between our study and the survey model. It is undeniable that an interview presents a greater detail of the information. Still, if we observe both networks' measures, we show high similarities, indicating that our approach, easier and cheaper to conduct than a survey, could be a complementary tool for the city managers. Comparing the random model with our study and the survey model, we showed the lack of linear correlation in both cases.

Our study shares similar conclusions with the study of Lotero et al. [10]. However, as mentioned earlier, our smart card-based approach is an easier and cheaper methodology to be implemented compared to a household travel survey (the one used by [3,10]). Besides, interviews and questionnaires are expensive to be conducted; it is not uncommon to take a considerable amount of time. We also observe different behavioral patterns between different cities from previous studies. Oviedo et al. [11] found that BRT system's implementation in Lima enabled greater accessibility for the higher income population, unlike the lower-income classes. Our study, on the other hand, observed that users from all economic classes of Curitiba move to all regions of the city, even though they present different characteristics in terms of mobility. Such a conclusion demonstrates the importance of studying particular cities before generalizing results.

The methodology and analysis covered in this study can be adapted for any city or region with transit access data. This information should make it possible, directly or indirectly, to understand urban space use better, enabling comparing particular areas. Using different databases, it was possible to enrich the characterization of public transit mobility patterns of different economic classes, which is essential to help create better urban policies that consider these aspects.

The mobility pattern centered on individual motorized transportation proves to be unsustainable, both in terms of environmental protection and in meeting the displacement demands of current urban life. The traditional response to congestion problems, by increasing road capacity, stimulates the use of the car and generates new congestion, feeding a vicious cycle responsible for the degradation of

air quality, global warming, and compromised quality of life in cities, such as a significant increase in noise levels, loss of time, degradation of public space and stress [34]. Thus, considering different economic classes, a view of urban transit mobility allows us to understand how people in these classes use urban spaces, providing subsidies for more sustainable economic development proposals.

Although the difference in income between classes is very subtle according to Census data, the difference in transit mobility patterns between these classes is noticeable. Besides helping to propose new interventions to promote more sustainable development, there are many other advantages to incorporating socioeconomic differences when studying users' mobility. For example, this could help develop new contagion processes or new mechanisms to disseminate strategic information in urban areas.

**Author Contributions:** Contributions for each authors: conceptualization, P.S., F.R.G., M.F., A.M., and T.H.S.; methodology, P.S., and F.R.G.; software, P.S., and F.R.G.; validation, P.S., M.F., A.M., and T.H.S.; formal analysis, P.S., F.R.G., M.F., A.M., and T.H.S.; investigation, P.S., F.R.G., M.F., A.M., and T.H.S.; data curation, P.S., and F.R.G.; writing—original draft preparation, P.S., F.R.G., M.F., A.M., and T.H.S.; writing—review and editing, P.S., M.F., A.M., and T.H.S.; visualization, P.S., and F.R.G. All authors have read and agreed to the published version of the manuscript.

**Funding:** This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior—Brasil (CAPES)-Finance Code 001. This work is also partially supported by the project URBCOMP (Grant #403260 /2016-7 from CNPq agency) and GoodWeb (Grant #2018/23011-1 from São Paulo Research Foundation—FAPESP). The authors would like to thank also the research agencies CAPES, CNPq and Fundação Araucária.

**Conflicts of Interest:** The authors declare no conflict of interest.

## Appendix A. Complementary Information

Table A1 summarizes some key demographics of Curitiba City.

**Table A1.** General demographics of Curitiba city.

Class	Neighborhoods	Area km <sup>2</sup>	Population	Demographic Density (pop/km <sup>2</sup> )
lower middle	Campo de Santana	21.57	26,657	1235.84
	Caximba	8.17	2522	308.69
	Ganchinho	11.20	11,178	998.04
	São Miguel	7.00	4773	681.86
	Tatuquara	11.23	52,780	4699.91
middle	Alto Boqueirão	12.11	53,671	4431.96
	Augusta	8.84	6598	746.38
	Cachoeira	3.07	9314	3033.88
	Cajuru	11.55	96,200	8329.00
	Cidade Industrial	43.38	172,822	3983.91
	Lamenha Pequena	3.40	1056	310.59
	Pinheirinho	10.73	50,401	4697.20
	Prado Velho	2.43	6077	2500.82
	Riviera	2.36	289	122.46
	Sítio Cercado	11.12	115,525	10,388.94
Umbará	22.47	18,730	833.56	

Table A1. Cont.

Class	Neighborhoods	Area km <sup>2</sup>	Population	Demographic Density (pop/km <sup>2</sup> )
upper middle	Abranches	4.32	13,189	3053.01
	Atuba	4.27	15,935	3731.85
	Bairro Alto	7.02	46,106	6567.81
	Barreirinha	3.73	18,017	4830.29
	Boqueirão	14.80	73,178	4944.46
	Butiatuvinha	10.58	12,876	1217.01
	Capão da Imbuia	3.16	20,473	6478.80
	Capão Raso	5.06	36,065	7127.47
	Fanny	2.00	8415	4207.50
	Fazendinha	3.72	28,074	7546.77
	Guaíra	2.32	14,904	6424.14
	Hauer	4.02	13,315	3312.19
	Lindóia	1.18	8584	7274.58
	Novo Mundo	5.99	44,063	7356.09
	Orleans	5.12	8105	1583.01
	Parolin	2.25	11,554	5135.11
	Pilarzinho	7.13	28,480	3994.39
	Santa Candida	10.33	32,808	3175.99
	São Braz	5.01	23,559	4702.40
	Taboão	1.72	3396	1974.42
Tingui	2.11	12,319	5838.39	
Uberaba	14.09	72,056	5113.98	
Xaxim	8.92	57,182	6410.54	
	Água Verde	4.76	51,425	10,803.57
	Ahú	1.84	11,506	6253.26
	Alto da Glória	0.88	5548	6304.55
	Alto da Rua XV	1.50	8531	5687.33
	Bacacheri	6.98	23,734	3400.29
	Batel	1.76	10,878	6180.68
	Bigorriho	3.50	28,336	8096.00
	Boa Vista	5.14	31,052	6041.25
	Bom Retiro	1.94	5156	2657.73
	Cabral	2.04	13,060	6401.96
	Campina do Siqueira	1.69	7326	4334.91
	Campo Comprido	8.55	28,816	3370.29
	Cascatinha	2.57	2161	840.86
	Centro	3.30	37,283	11,297.88
	Centro Cívico	0.97	4783	4930.93

Table A1. Cont.

Class	Neighborhoods	Area km <sup>2</sup>	Population	Demographic Density (pop/km <sup>2</sup> )
low high	Cristo Rei	1.46	13,795	9448.63
	Guabirota	2.63	11,461	4357.79
	Hugo Lange	1.15	3392	2949.57
	Jardim Botânico	2.77	6172	2228.16
	Jardim das Américas	3.87	15,313	3956.85
	Jardim Social	1.89	5698	3014.81
	Juvevê	1.23	11,582	9416.26
	Mercês	3.28	12,907	3935.06
	Mossunguê	3.38	9664	2859.17
	Portão	5.70	42,662	7484.56
	Rebouças	2.98	14,888	4995.97
	Santa Felicidade	12.27	31,572	2573.11
	Santa Quitéria	2.09	12,075	5777.51
	Santo Inácio	2.72	6494	2387.50
	São Francisco	1.36	6130	4507.35
	São João	3.03	3253	1073.60
	São Lourenço	2.26	6276	2776.99
	Seminário	2.13	6851	3216.43
Tarumã	4.17	8072	1935.73	
Vila Izabel	1.21	11,610	9595.04	
Vista Alegre	3.69	11,199	3034.96	

## References

1. Agarwal, R.; Dhar, V. Big Data, Data Science, and Analytics: The Opportunity and Challenge for IS Research. *Inf. Syst. Res.* **2014**, *25*, 443–448. [[CrossRef](#)]
2. Silva, T.H.; Viana, A.C.; Benevenuto, F.; Villas, L.; Salles, J.; Loureiro, A.; Quercia, D. Urban Computing Leveraging Location-Based Social Network Data: A Survey. *ACM Comput. Surv.* **2019**, *52*, 17:1–17:39. [[CrossRef](#)]
3. Zhang, X.; Xu, Y.; Tu, W.; Ratti, C. Do different datasets tell the same story about urban mobility—A comparative study of public transit and taxi usage. *J. Transp. Geogr.* **2018**, *70*, 78–90. [[CrossRef](#)]
4. Huang, L.; Yang, Y.; Gao, H.; Zhao, X.; Du, Z. Comparing Community Detection Algorithms in Transport Networks via Points of Interest. *IEEE Access* **2018**, *6*, 29729–29738. [[CrossRef](#)]
5. De Freitas Miranda, H.; da Silva, A.N.R. Benchmarking sustainable urban mobility: The case of Curitiba, Brazil. *Transp. Policy* **2012**, *21*, 141–151. [[CrossRef](#)]
6. Araujo, M.; Melo de Oliveira, J.; Jesus, M.; Sa, N.; Santos, P.; Lima, T. Collective public transportation: Discussing accessibility, mobility and quality of life. *Psicol. Soc.* **2011**, *23*, 574–582. [[CrossRef](#)]
7. Ni, Y.; Yi, F.; Ma, S.; Sun, Y. User Journey with Intelligent Public Transportation System in Shanghai. In *HCI in Mobility, Transport, and Automotive Systems*; Krömker, H., Ed.; Springer International Publishing: Cham, Switzerland, 2019; pp. 487–499.
8. Salazar, M.; Rossi, F.; Schiffer, M.; Onder, C.H.; Pavone, M. On the Interaction between Autonomous Mobility-on-Demand and Public Transportation Systems. In Proceedings of the 2018 21st International Conference on Intelligent Transportation Systems (ITSC), Maui, HI, USA, 4–7 November 2018; pp. 2262–2269.

9. Magdolen, M.; Von Behren, S.; Chlond, B.; Hunecke, M.; Vortisch, P. Combining attitudes and travel behavior—A comparison of urban mobility types identified in Shanghai, Berlin and San Francisco. In Proceedings of the Transportation Research Board 98th Annual Meeting, Washington, DC, USA, 13–17 January 2019.
10. Lotero, L.; Hurtado, R.G.; Floría, L.M.; Gómez-Gardeñes, J. Rich do not rise early: Spatio-temporal patterns in the mobility networks of different socio-economic classes. *R. Soc. Open Sci.* **2016**, *3*, 150654. [[CrossRef](#)]
11. Oviedo, D.; Scholl, L.; Innao, M.; Pedraza, L. Do Bus Rapid Transit Systems Improve Accessibility to Job Opportunities for the Poor? The Case of Lima, Peru. *Sustainability* **2019**, *11*, 2795. [[CrossRef](#)]
12. IPPUC. Curitiba Institute of Urban Planning and Research—Origen and Destination Survey. 2017. Available online: [https://ippuc.org.br/visualizar.php?doc=http://admsite2013.ippuc.org.br/arquivos/documentos/D536/D536\\_002\\_BR.pdf](https://ippuc.org.br/visualizar.php?doc=http://admsite2013.ippuc.org.br/arquivos/documentos/D536/D536_002_BR.pdf) (accessed on 10 March 2020).
13. Balcan, D.; Gonçalves, B.; Hu, H.; Ramasco, J.J.; Colizza, V.; Vespignani, A. Modeling the spatial spread of infectious diseases: The Global Epidemic and Mobility computational model. *J. Comput. Sci.* **2010**, *1*, 132–145. [[CrossRef](#)]
14. Van den Broeck, W.; Gioannini, C.; Gonçalves, B.; Quaggiotto, M.; Colizza, V.; Vespignani, A. The GLEAMviz computational tool, a publicly available software to explore realistic epidemic spreading scenarios at the global scale. *BMC Infect. Dis.* **2011**, *11*, 37. [[CrossRef](#)]
15. Cardillo, A.; Zanin, M.; Gómez-Gardeñes, J.; Romance, M.; García del Amo, A.J.; Boccaletti, S. Modeling the multi-layer nature of the European Air Transport Network: Resilience and passengers re-scheduling under random failures. *Eur. Phys. J. Spec. Top.* **2013**, *215*, 23–33. [[CrossRef](#)]
16. Ganin, A.A.; Kitsak, M.; Marchese, D.; Keisler, J.M.; Seager, T.; Linkov, I. Resilience and efficiency in transportation networks. *Sci. Adv.* **2017**, *3*, e1701079. [[CrossRef](#)]
17. Marques-Neto, H.T.; Xavier, F.H.Z.; Xavier, W.Z.; Malab, C.H.S.; Ziviani, A.; Silveira, L.M.; Almeida, J.M. Understanding Human Mobility and Workload Dynamics Due to Different Large-Scale Events Using Mobile Phone Data. *J. Netw. Syst. Manag.* **2018**, *26*, 1079–1100. [[CrossRef](#)]
18. Yang, Y.; Heppenstall, A.; Turner, A.; Comber, A. Who, Where, Why and When? Using Smart Card and Social Media Data to Understand Urban Mobility. *ISPRS Int. J. Geo-Inf.* **2019**, *8*, 271. [[CrossRef](#)]
19. Xu, Y.; Belyi, A.; Bojic, I.; Ratti, C. Human mobility and socioeconomic status: Analysis of Singapore and Boston. *Comput. Environ. Urban Syst.* **2018**, *72*, 51–67. [[CrossRef](#)]
20. IPPUC. Curitiba Institute of Urban Planning and Research—Maps from Curitiba. 2019. Available online: <https://www.urbs.curitiba.pr.gov.br/transporte/rede-integrada-de-transporte/32> (accessed on 7 November 2019).
21. Taniguchi, G.; Duarte, F. Personal Smart Cards: From Transportation to a City Smart Card—The Database Integration of Public Services in Curitiba. In *City Competitiveness and Improving Urban Subsystems: Technologies and Applications*; IGI Global: Pennsylvania, PA, USA, 2012; pp. 217–232. [[CrossRef](#)]
22. URBS. URBS in Numbers. 2018. Available online: <https://www.urbs.curitiba.pr.gov.br/institucional/urbs-em-numeros> (accessed on 7 November 2019).
23. Ninis, A.B.; Paes de Barros, R.; Grosner, D.; Franco, S.; Rosalém, A.; Bugarin, K.; Sayuri Sataka, G.; Jorge, J. Relatório de Definição da Classe Média. 2012. Available online: <https://issuu.com/sae.pr/docs/relatoriocm> (accessed on 17 January 2020).
24. IBGE. Brazilian Institute of Geography and Statistics—2010 Census. 2010. Available online: <https://censo2010.ibge.gov.br/> (accessed on 7 November 2019).
25. Guttman, A. R-trees: A Dynamic Index Structure for Spatial Searching. *SIGMOD Rec.* **1984**, *14*, 47–57. [[CrossRef](#)]
26. Noulas, A.; Scellato, S.; Lambiotte, R.; Pontil, M.; Mascolo, C. A Tale of Many Cities: Universal Patterns in Human Urban Mobility. *PLoS ONE* **2012**, *7*, e0037027. [[CrossRef](#)]
27. Ji, S.; Li, W.; Srivatsa, M.; He, J.S.; Beyah, R. General Graph Data De-Anonymization: From Mobility Traces to Social Networks. *ACM Trans. Inf. Syst. Secur.* **2016**, *18*. [[CrossRef](#)]
28. Newman, M. *Networks: An Introduction*; Oxford University Press: Oxford, UK, 2010.
29. Yildirimoglu, M.; Kim, J. Identification of communities in urban mobility networks using multi-layer graphs of network traffic. *Transp. Res. Part C Emerg. Technol.* **2018**, *89*, 254–267. [[CrossRef](#)]
30. Croce, A.I.; Musolino, G.; Rindone, C.; Vitetta, A. Transport System Models and Big Data: Zoning and Graph Building with Traditional Surveys, FCD and GIS. *ISPRS Int. J. Geo-Inf.* **2019**, *8*, 187. [[CrossRef](#)]

31. Sharov, M. Improving Methods of Estimation of Transport Demand for Urban Passenger Transportation. In Proceedings of the International Conference on Aviaemechanical Engineering and Transport, Irkutsk, Russia, 27 May—1 June 2019; doi:10.2991/aviaent-19.2019.55. [[CrossRef](#)]
32. González, M.C.; Hidalgo, C.A.; Barabási, A.L. Understanding individual human mobility patterns. *Nature* **2008**, *453*, 779–782. [[CrossRef](#)] [[PubMed](#)]
33. Brandes, U. On variants of shortest-path betweenness centrality and their generic computation. *Soc. Netw.* **2008**, *30*, 136–145. [[CrossRef](#)]
34. Litman, T. *Smart Transportation Economic Stimulation*; Victoria Transport Policy Institute: Victoria, BC, Canada, 2009. Available online: [https://www.vtpi.org/econ\\_stim.pdf](https://www.vtpi.org/econ_stim.pdf) (accessed on 17 November 2020).

**Publisher’s Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).