ISSN (PRINT): 1994 - 7658 ISSN (ONLINE): 1314 - 0817

VARNA UNIVERSITY OF MANAGEMENT

European Journal of Tourism Research

http://ejtr.vumk.eu

Large Deluxe or Weekly Stay? Diversity of peer-to-peer accommodation offers at an urban tourism destination

Ewa E. Kiczmachowska [1]*

[1] Kozminski University, Warsaw, Poland, Department of Marketing. Email: ekiczmachowska@kozminski.edu.pl

*Corresponding author

Abstract

The purpose of this study was to explore the diversity of peer-to-peer accommodation (P2PA) offers at an urban tourism destination. For P2PA, where the offers are mostly managed by individual hosts, the level of diversity is high, but it hasn't been investigated so far. To deliver the purpose of this study, K-means clustering algorithm was employed to cluster property level data of 7,780 Airbnb accommodations located in Warsaw, Poland. Employing the hedonic price theory and the research regarding performance determinants in P2PA, 11 variables and 121 amenities were analysed. The study outlined P2PA clusters characterised by the combinations of property and host features, including performance measure of revenues, which carries practical implications for both, current and future short-term tenancy providers. This study added to the P2PA research stream in several ways. First, it used an unsupervised clustering method for the exploration of the diversity of P2PA offers. Second, it investigated a case of Central and Eastern Europe, which was rarely studied in this context. Third, it employed the revenue as the metric of listing performance, in contrast to the price predominantly used so far. Finally, it proposed a typology of P2PA clustering features, based on property-host-guest triangle.

Keywords: peer-to-peer accommodation, diversity of listings, revenue, urban tourism, Airbnb, K-means

Citation: Kiczmachowska, E. (2025). Large Deluxe or Weekly Stay? Diversity of peer-to-peer accommodation offers at an urban tourism destination. *European Journal of Tourism Research*, 39, 3910. https://doi.org/10.54055/ejtr.v39i.3765.

Publication history:

Received: 09/05/2024; Revised: 10/09/2024, 14/10/2024; Accepted: 16/10/2024; Published online: 15/01/2025; Volume date: 01/03/2025

Coordinating editor: Faruk Seyitoglu

1. Introduction

Peer-to-peer accommodation (P2PA) is distinct from other accommodation providers (e.g., hotels) in its high level of differentiation among offers. Specifically, the features of the accommodations (properties), and the features or behaviours of the service providers (hosts) vary widely. In line with the hedonic price theory, which regards the price as the sum of customer valuations of a certain set of product (or service) characteristics (Lancaster, 1966; Chen & Xie, 2017), the customers (guests) assess the offers and choose a mix of features that suits them best.

Although revenue might not be the only reason to get involved in P2PA, as the host's social and emotional motivations might play a significant role, the possibility to obtain additional income remains vital (Sundararajan, 2016; Narasimhan *et al.*, 2018; Kwok & Xie, 2019). Therefore, when searching for ways to explore the diversity of P2PA offers, it seems logical to use the existing work on *P2PA performance determinants*, centred around property or host characteristics impacting the performance. So far, in the P2PA performance determinants research stream, the methods have widely centred on regression modelling (Sainaghi & Baggio, 2020), with price as the predominant dependent variable, whereas relatively few studies have focused on revenue (Sainaghi *et al.*, 2021). The majority of studies have focused on urban destinations (Hernández *et al.*, 2021), as does this study. However, geographically, previous research centred on North American, Asian, and Australian cases (Sainaghi & Baggio, 2020; Akbari *et al.*, 2022), while papers on European destinations have been considerably less represented. Moreover, those studies focused on determining the factors impacting performance rather than exploring the diversity of the offers.

Therefore, it is advisable to develop tools to research the diversity of P2PA listings, with focus on 'attributes of listings, different location patterns, and many other segmentations' (Sainaghi, 2020). As such, the research questions for this study are as follows:

RQ1: What are the applicable factors for clustering P2PA properties?
RQ2: What are the clusters of P2PA properties?
RQ3: What are the characteristics of these clusters?

This study adds to the research stream on P2PA in several ways. First, based on research on determinants of performance, a typology of P2PA offers' features is proposed, employing a property-host-guest triangle. Second, a method of unsupervised machine learning (K-means clustering algorithm) is proposed to analyse the diversity of offers. Third, this study expands the relatively scarce research on revenue as opposed to price. Finally, the study extends European representation in this stream.

In addition to its theoretical contributions, this study also offers practical considerations for current and future P2PA providers. It outlines the high-revenue combinations of property features and host behaviours, which are useful for investors looking for new enterprise options or for current hosts looking for opportunities to improve their performance. For low-revenue combinations, accurate strategies to increase revenue are proposed. While a majority of the findings are generalizable, the methodology is also easily replicable to other datasets.

2. Literature review

As already noted, when searching for ways to explore the diversity of the P2PA offers, researchers should consider the existing work on *P2PA performance determinants*, which is centred around property or host characteristics impacting P2PA performance.

The existing literature captures several typologies of P2PA performance determinants that can be contradictory, leading to confusion. Based on the hedonic price model, Sainaghi *et al.* (2021) group performance determinants into internal (related to the property or the host) and external (related to the location, guests' reviews, social and economic traits, and seasonal patterns). Wang and Nicolau (2017) group performance determinants into five categories: property size, property location, host attributes, rental rules, and advertising. Meanwhile, Hernández *et al.* (2021) propose grouping determinants into structural, host, management, reputation, and location. Finally, Tang *et al.* (2019) split the determinants into site factors (physical attributes, reputation) and situational factors (location-specific factors and marketing conditions). To overcome the complexity and ambiguity of existing typologies, an alternative typology is proposed for this study, employing the triangle of *property* (type, size, location, amenities); *host* (professionalism, management, advertising); and *guest* (reviews, ratings) (see Figure 1). Guest assessments result from property- and host-related determinants, so it is logical to group the determinants into primary (property and host) and secondary (guest).

In terms of *property*-related determinants, *type* (entire home/apartment, private room, or shared room, with entire home/apartment achieving higher prices) and *size* (number of bedrooms, number of bathrooms, and maximum number of guests, with size having a positive relation to performance) proved to be strong performance determinants (Wang & Nicolau, 2017; Perez-Sanchez *et al.*, 2018; Magno *et al.*, 2018; Sainaghi *et al.*, 2021). Regarding *location*, the findings of previous research indicate a significant negative impact of distance to central areas (Gibbs *et al.*, 2018a), suggesting that the farther the property is from the city centre (usually in terms of Euclidean distance; Sainaghi & Chica-Olmo, 2022), the lower is its price (Wang & Nicolau, 2017; Benítez-Aurioles, 2018a). Apart from the city centre (Gyódi & Nawaro, 2021), location has also been operationalized as the distance to the main touristic attraction (Abrate & Viglia, 2019); the nearest airport (Jang & Kim, 2022); and sightseeing, eating, and shopping (Perez-Sanchez *et al.*, 2018). Finally, *amenities* were studied by Chattopadhyay and Mitra (2019), who identified elevator, parking, indoor fireplace, cable TV, and family/kid-friendly as among the most important price determinants.

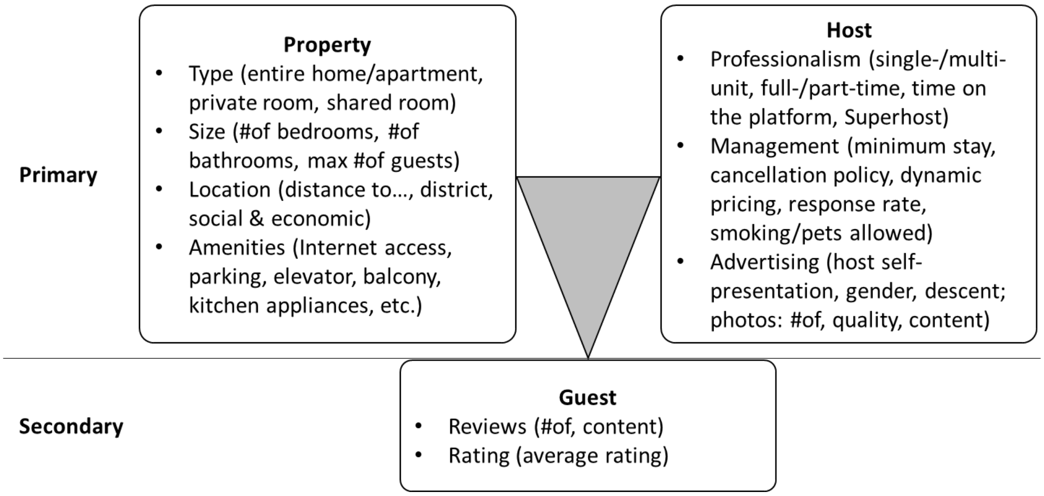


Figure 1. *P2PA performance determinants. Author's own elaboration.*

Regarding *host*-related determinants, host *professionalism* has been studied, operationalized as single-/multi-unit, full-/part-time, experienced/amateur (time on the platform), or Superhost (Chen & Xie, 2017; Xie & Mao, 2017; Gibbs *et al.*, 2018a; Abrate & Viglia, 2019; Deboosere *et al.*, 2019). The results demonstrated that professional hosts are able to command higher prices than non-professionals (Magno *et al.*, 2018). Other determinants include host behaviours related to *management* and *advertising*. Rental rules cover some revenue management tools, like minimum-length-of-stay controls, non-refundable advanced deposits, and cancellation penalties (Talluri & van Ryzin, 2005; Tranter *et al.*, 2014), or other rules, such as rules for smoking and pets. Researchers have studied performance related to the application of dynamic pricing, finding that the listing performance and the average intensity of price variability tend to increase with the degree of professionalism (Gibbs *et al*., 2018b; Abrate *et al.*, 2022; Kiczmachowska, 2022a), and the application of a minimum-stay restriction, finding that amateur hosts underutilized this tool (Kreeger & Smith, 2017). Regarding cancellation policies (CP), the evidence is mixed; on one hand, moving to stricter policies positively affected price as well as revenue (Wang & Nicolau, 2017; Sainaghi *et al.*, 2021), while on the other hand only moderate policy was significantly negatively associated with price (Chen & Xie, 2017). Further, the response rate was assessed as a significant but not particularly strong price and revenue determinant (Sainaghi *et al.*, 2021). Turning to *advertising*, Nieto García *et al.* (2020) studied the effectiveness of host self-presentation and revealed an inverse U-shaped effect of self-presentation length on revenues. Finally, several recent studies concentrated on the role of photos as a performance determinant. Ma *et al.* (2022) found that verified photos had a larger positive effect on the performance of properties that lacked certification and those with fewer customer reviews. According to He *et al.* (2023), the booking rate was boosted by a background photo that featured a living room and showed more interior elements, while bedroom photos tended to decrease it.

In terms of *guest*-related determinants, online reputation and user perception, measured as *number of reviews* and average *rating*, have been identified as drivers of rental price (Liang *et al*., 2017; Ribes *et al*., 2018; Perez-Sanchez *et al.*, 2018; Abrate & Viglia, 2019). Specifically, the number of reviews received by a property was negatively correlated to its price (Liang *et al*., 2017; Magno *et al.*, 2018); meanwhile, customer rating positively affected performance, but its effect was weaker for a multi-listing host and stronger for a full-time host (Xie *et al.*, 2021).

According to Sainaghi (2021), *performance* has been most frequently operationalized as *price* (Chen & Xie, 2017; Wang & Nicolau, 2017; Benítez-Aurioles, 2018b; Gibbs *et al*., 2018a; Magno *et al*., 2018; Perez-Sanchez *et al.*, 2018; Chattopadhyay & Mitra, 2019; Sainaghi *et al.*, 2021; Casamatta *et al.*, 2022; Ferreira *et al.*,2023), while relatively fewer researchers employed *revenue* (Casamatta *et al.*, 2022; Jang *et al.*, 2021; Jang & Kim, 2022; Abrate & Viglia, 2019; Sainaghi *et al*., 2021; Xie *et al.*, 2021), *occupancy rate* (Jang & Kim, 2022; Filieri *et al.*, 2023), *booking rate* (He *et al.*, 2023), or *number of reservation days* (Ferreira *et al.,* 2023).

Methods used have largely centred around regression, usually based on hedonic price models (Sainaghi, 2021; Casamatta *et al.*, 2022). Hernández *et al.* (2021) applied a clustering method (K-means algorithm) to identify groups of significant determinants associated with geographical locations.

3. Methodology

3.1. Data

Data for this study were sourced from AirDNA, which tracks the daily performance of over 10 million properties on Airbnb and Vrbo in 120,000 global markets (AirDNA, 2023). AirDNA data have been successfully used in previous research on the price and revenue of Airbnb properties (Perez-Sanchez *et al.*, 2018; Jang *et al.*, 2021; Koh *et al.*, 2020; Jang & Kim, 2022; Kiczmachowska, 2022b).

The dataset included 28,042 properties based in Warsaw, Poland, which was the second most popular city destination in Central and Eastern Europe in 2021 (excl. Russia), according to Euromonitor (2021). Several exclusions to the main dataset were made, as described below. First, properties of AR with 'zero' or 'empty' cells in the last 12 months (March 2022–February 2023) were removed from the dataset, leaving 9,599 properties. Second, 45 properties described as hotel rooms were excluded as out of scope for this study. Third, as this study focused on short-term rentals, 126 properties with minimum stays longer than 31 days were removed. These exclusions left 9,428 properties, including 7,798 *entire homes/apartments* and 1,630 *private/shared rooms*. Focusing on the *entire home/apartment* property type, 1 property with no *CP* and several outliers were removed: 1 property with number of bedrooms = 27, 1 property with number of photos = 219, 11 properties with AR exceeding 100K USD, and 4 properties with ADR exceeding 2K USD. This left a final sample of 7,780 properties included in the study.

3.2. Variables

Table 1 displays the list of variables. The performance variable was defined as *Annual Revenue (AR)*, which gives the total revenue achieved by the property in the last 12 months recalculated into USD. This measure is calculated by price (Average Daily Rate, ADR), occupancy rate, and the number of days that the property was blocked (unavailable for rent) during the year. This measure was purposefully chosen, as it encompasses all the effects. It thus helps to avoid the bias, as noted by Xie *et al.* (2021), whereby occasional hosts who chose to rent only during peak periods may achieve higher than average ADR results despite having lower AR.

Following the property-host-guest performance determinants typology, this study focused on property- and host-related determinants as underlying primary factors. Guest-related determinants, as secondary factors resulting from the property-host axis, were excluded from the analysis for this study.

As previous studies indicated the significant and positive influence of *entire home/apartment* vs. *private/shared room* property types on performance (Wang & Nicolau, 2017; Magno *et al.*, 2018; Sainaghi *et al.*, 2021), this study focused exclusively on the *entire home/apartment* type.

For the property size, three numerical variables were defined: *maximum-number-of-guests* ranging from 1 to 16, *number-of-bedrooms* ranging from 0 (for 1-room studio with no separate bedroom) to 8, and *number-of-bathrooms* ranging from 0 (e.g., campers) to 8.

Table 1. *Variables.*

Type	Group	Variable	Abbreviation	Source
Performance	Revenue	Annual Revenue	AR	AirDNA
Property	Size	Maximum-number-of-guests	Max Guests	AirDNA
		Number-of-bedrooms	Bedrooms	AirDNA
		Number-of-bathrooms	Bathrooms	AirDNA
	Location	Distance to Central Railway Station	CRS	AirDNA
		Distance to Old Town Square	OTS	latitude/longitude orthodromic distance
Host	Management	Cancellation Policy	CP	AirDNA
		Minimum Stay	Min Stay	AirDNA
		Pets Allowed	Pets	AirDNA
	Professionalism	Superhost	Superhost	AirDNA
	Advertising	Number of Photos	No of Photos	AirDNA

In terms of location, in Warsaw, the Central Railway Station is the centre of the transportation and business district along with the shopping centre. However, the historic district, including the Royal

Castle and the Old Town Square, is located about 2.5 km away. Therefore, two variables for the location were set: *distance to Central Railway Station (CRS)* and *distance to Old Town Square (OTS)*. The distance was calculated using latitude and longitude measures and the great-circle (orthodromic) distance, which is the shortest distance between two points measured along the surface of the sphere. In urban conditions, this measure usually returns results lower than the real-life distance (as one has to stay on the available streets), but it is easy to calculate and a sufficiently close approximation of the distance between two points.

Among host-related determinants, three management-related variables were chosen. Airbnb provides hosts with four standard *CPs* (Airbnb, 2023a), which cover cancellation periods and refund terms. For this study, this variable was coded as 0 = flexible, 1 = moderate, 2 = firm, 3 = strict. Regarding the minimum-length-of-stay tool, the host can set 1 or more days as the minimum rental period for the property. For this study, the *minimum-stay* variable was defined as number of days set by the host, taking values from 1 to 31. The last variable to reflect management-related determinants was created to distinguish between properties that *allow pets* (coded = 1) and those that do not (coded = 0).

The host professionalism characteristic that is clearly visible to potential guest while booking (in the form of a badge icon, , appearing on the listing) is *Superhost* status. This is a combination of certain conditions fulfilled by the host including high activity rate, high response rate, high overall rating, and low cancellation rate (Airbnb, 2023b). For this study, this variable was defined as 1 if the property was managed by a Superhost and 0 otherwise. In terms of the advertising variable, for this study it was operationalized by the *number-of-photos*, defined as a numeric variable ranging from 1 to 113.

Although not included in the clustering procedure because of their considerable complexity, amenities were used as additional descriptors of clusters once the clustering procedure was completed. Each property offered a unique set of amenities, and the number of amenities varied. In total, 121 unique amenities were identified, for which each property was assigned 1 if it offered a given amenity and 0 otherwise. The amenities were ranked in terms of intensity of offering, from extremely popular (e.g., wireless Internet, kitchen, or heating) to extremely rare (e.g., piano, beachfront, or kayak). The parking-related amenities were grouped into free and paid parking. The top 30 amenities and amenities related to parking were analysed per cluster in the form of a heat map table.

3.3. Method

Data mining is a process of extracting knowledge from large amounts of data. This technique makes it possible to identify trends, patterns, correlations, and anomalies in databases, which can aid in making accurate future decisions (Gautam & Kumar, 2022). In clustering, millions of data points are grouped to form a cluster (Dwivedi & P.Bhaiya, 2019), such that the objects in the cluster are as similar as possible, while the objects in different clusters are as dissimilar as possible (Yu *et al.*, 2022; Fernández-Durán & Gregorio-Domínguez, 2021).

Among the non-hierarchical methods, the K-means algorithm, an unsupervised machine learning and clustering algorithm, is very popular because of its speed, efficiency, and simplicity (Chen *et al.*, 2022), as well as its usability for various types of data (Akbar *et al.*, 2020) and its good performance on large datasets (Jain & Dubes, 1988, p. 90; Anitha & Patil, 2022). Thus, it was selected for this study.

The analysis was carried out in several steps (see Figure 2). In Step 1, as advocated by previous research on the K-means algorithm (e.g., Olejniczak, 2021; Sharaf Addin *et al.*, 2022), the main variables were standardised, to take values between -1 and 1. In Step 2, the correlation matrix was calculated using the rho-Spearman test, as all the variables proved not to show normal distribution. The variables with the strongest (> |0.1|) correlation to *Annual Revenue* (AR) were identified, and—together with AR—were included in the K-means clustering procedure.

In K-means clustering, it is necessary to specify the number of initial clusters in advance (Liu *et al.*, 2023); thereafter, the algorithm defines the clusters based on the minimal variation of each data point from its centroid and other centroids (Sharaf Addin *et al.*, 2022). In Step 3, the optimal number of clusters was defined using the hierarchical clustering, the Dendrogram (Jain & Dubes, 1988, p. 97), and the Elbow Method (Sujatha *et al.*, 2023). The Dendrogram was prepared using Ward's method, as it was found to outperform other hierarchical clustering methods (Jain & Dubes, 1988, p. 81). In the Elbow Method, the K-means clustering was performed for a range of k values from 2 to 10, and the within-cluster-sum-of-squares (Cui, 2020; Pradana & Ha, 2021; Chen *et al.*, 2022) for all procedures was calculated. The final presentation of the Elbow Method was a graph, and the most elbow-shaped point was considered as the optimal number of clusters (Larasati *et al.*, 2021; Wu *et al.,* 2020). Based on both tools, the number of clusters for the K-means clustering algorithm was set at 8. In Step 4, the K-means clustering at 8 clusters was performed using SPSS. In Step 5, detailed analysis of the clusters was performed using descriptive analytics, independent samples Kruskal-Wallis tests, and Microsoft Excel 3D Maps. Additionally, descriptive analysis of amenities by cluster was performed using a heat map table. On the basis of these analyses, cluster profiles were proposed.

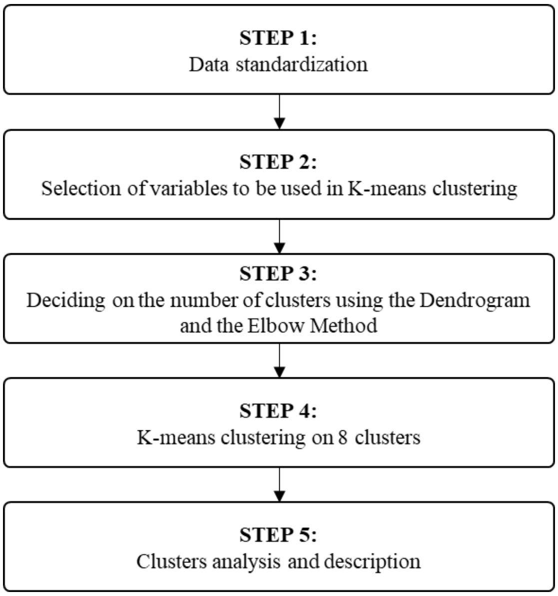


Figure 2. *The research procedure. Author's own elaboration.*

4. Results

4.1. Descriptive statistics and clustering procedure

The descriptive statistics are presented in Table 2. On average, the properties were located closer to *CRS* (M = 3.23 km, SD = 2.61 km) than to *OTS* (M=3.85 km, SD=2.85 km), included 1 *bedroom* (M = 1.07, SD = 0.78) and 1 *bathroom* (M = 1.09, SD = 0.29), and had a maximum capacity of almost 4 *guests* (M = 3.73, SD = 1.58). The *CPs* employed by the hosts were predominantly flexible or moderate (M = 1.02, SD = 1.10), the average *minimum stay* was 2.6 days (M = 2.6, SD = 4.06), and the average *number-of-photos* was almost 19 (M = 18.96, SD = 11.46). Only 8% of the properties were run by *Superhosts*, and 35% *allowed pets*. The average AR was 7,916 USD (SD = 8,401.74), ranging from 10 USD to 75,646 USD. The average *occupancy rate* was 66% (SD = 0.25), and the average ADR was 80.76 USD (SD = 56.92).

Table 2. *Descriptive Statistics (SPSS).*

Variable	N	Min	Max	Mean/ Proportion	Std. Dev.	Median	Skewness	Kurtosis
Central Railway Station	7780	0.14	16.44	3.23	2.61	2.35	1.337	1.641
Old Town	7780	0.01	17.10	3.85	2.85	2.78	1.243	1.111
Average Daily Rate	7780	9.00	1,169.00	80.76	56.92	66.69	5.970	71.932
Annual Revenue LTM	7780	10.00	75,646.00	7,916.00	8,401.74	5,062.00	2.013	6.714
Occupancy Rate LTM	7780	0.026	1.00	0.66	0.25	0.69	-0.536	-0.627
Bedrooms	7780	0	8	1.07	0.78	1.00	1.098	3.181
Bathrooms	7780	0	8	1.09	0.29	1.00	6.052	82.844
Max Guests	7780	1	16	3.73	1.58	4.00	1.373	4.376
Airbnb Superhost	7780	0	1	0.08	0.27	0.00	-	-
Cancellation Policy	7780	0	3	1.02	1.10	1.00	-	-
Minimum Stay	7780	1	31	2.60	4.06	1.00	5.043	28.285
Number of Photos	7780	1	113	18.96	11.46	16.00	1.895	6.201
Pets Allowed	7780	0	1	0.35	0.48	0.00	-	-

All the variables but *occupancy rate* presented skewness and kurtosis >1, indicating non-normal distributions. Variables for the clustering procedure were selected with rho-Spearman correlations (Table 3). AR and the variables with the strongest (> |0.1|) correlations to AR were included in the K-means procedure (marked in grey in Table 3): *AR*, *distance to CRS*, *distance to OTS*, *maximum-number-of-guests*, *minimum-stay,* and *number-of-photos.*

Table 3. *Correlations table (rho-Spearman, SPSS).*

	Variable	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]	[11]	[12]
[1]	C. Railway Station (CRS)												
[2]	Old Town Square (OTS)	0.679**											
		0.000											
[3]	Bedrooms	0.009	-0.023*										
		0.419	0.043										
[4]	Bathrooms	0.031**	0.043**	0.284**									
		0.007	0.000	0.000									
[5]	Max Guests	-0.069**	-0.095**	0.595**	0.208**								
		0.000	0.000	0.000	0.000								
[6]	Airbnb Superhost	-0.012	-0.030**	0.006	-0.002	-0.028*							
		0.277	0.007	0.589	0.826	0.014							
[7]	Cancellation Policy	-0.095**	-0.084**	-0.021	0.011	0.034**	0.068**						
		0.000	0.000	0.070	0.317	0.003	0.000						
[8]	Minimum Stay	0.069**	0.063**	0.071**	0.059**	-0.080**	0.040**	0.032**					
		0.000	0.000	0.000	0.000	0.000	0.000	0.005					
[9]	No of Photos	-0.083**	-0.103**	0.128**	0.101**	0.216**	0.090**	0.067**	0.009				
		0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.418				
[10]	Pets Allowed	-0.031**	-0.016	0.044**	0.036**	0.067**	-0.042**	-0.154**	-0.089**	-0.001			
		0.006	0.166	0.000	0.001	0.000	0.000	0.000	0.000	0.920			
[11]	Average Daily Rate	-0.117**	-0.087**	0.289**	0.163**	0.398**	-0.097**	-0.007	-0.174**	0.223**	0.109**		
		0.000	0.000	0.000	0.000	0.000	0.000	0.543	0.000	0.000	0.000		
[12]	Annual Revenue	-0.164**	-0.153**	0.004	0.037**	0.184**	0.034**	0.035**	-0.170**	0.230**	-0.030**	0.325**	
		0.000	0.000	0.748	0.001	0.000	0.003	0.002	0.000	0.000	0.007	0.000	
[13]	Occupancy Rate	-0.096**	-0.069**	-0.057**	-0.027*	-0.013	0.090**	0.045**	0.132**	0.088**	-0.076**	-0.298**	0.260**
		0.000	0.000	0.000	0.016	0.238	0.000	0.000	0.000	0.000	0.000	0.000	0.000

***p< 0.01 *p<0.05* *N=7,780*

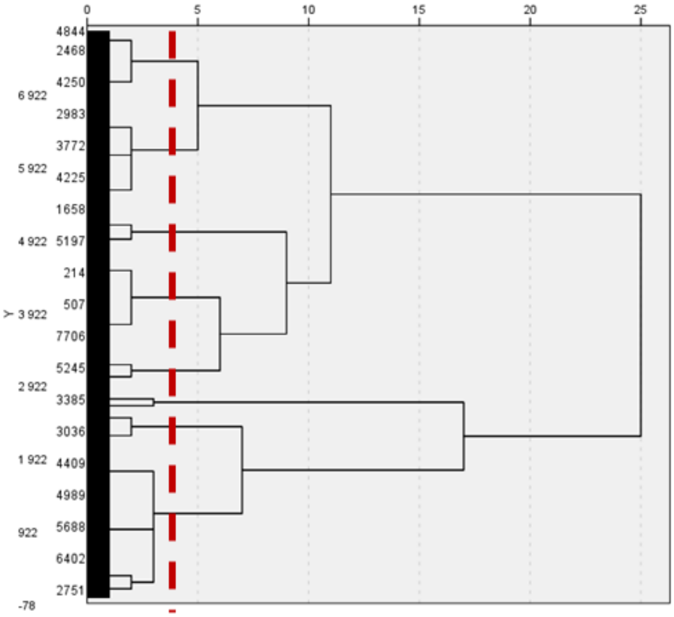


Figure 3. *The Dendrogram. Author's own elaboration (SPSS).*

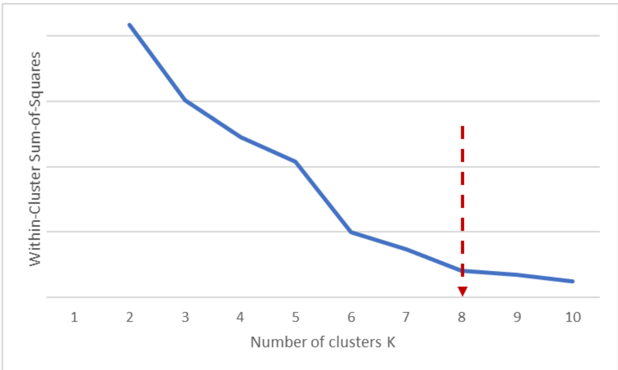


Figure 4. *The Elbow Method. Author's own elaboration.*

Table 4. *ANOVA table (K-means analysis, SPSS).*

Variable	Cluster Mean Square	df	Error Mean Square	df	F	Sig.
Stand (CentralRailwayStation)	794.555	7	0.285	7772	2,785.282	<0.001
Stand (OldTown)	817.891	7	0.264	7772	3,095.118	<0.001
Stand: Max Guests	498.091	7	0.552	7772	901.872	<0.001
Stand: Minimum Stay	844.107	7	0.241	7772	3,507.768	<0.001
Stand: No of Photos	553.206	7	0.503	7772	1,100.588	<0.001
Stand: Annual Revenue	722.277	7	0.350	7772	2,061.479	<0.001

To identify the initial number of clusters, two methods were used: the Dendrogram and the Elbow Method. The Dendrogram results (Figure 3) suggested 8 clusters to be optimal, as this number minimized variations in the clusters' size. In case of the Elbow Method (Figure 4), the inflection points could be identified at 3, 6, and 8 clusters. The 3-cluster solution was assessed as too simplistic, and the

6-cluster solution returned high variations in cluster sizes (from 213 to 3,581 properties), while in the 8-cluster solution the variations were considerably smaller (the biggest cluster was 2,550 properties). Finally, the decision was made to apply the 8-cluster solution, supported by ANOVA tests (a supplemental analysis to the SPSS K-means solution), which returned solid F-values, proving all the variables were significant in forming the clusters (Table 4).

4.2. Description of clusters

The final clusters' centres, as returned by the SPSS K-means algorithm and recalculated from standardised to average values, along with the profile analysis, are presented in Table 5. The clusters differ in size: there are 2 small (2, 4), 3 medium (1, 6, 7) and 3 large clusters (3, 5, 8). The average AR per cluster ranges from 2.2K USD to 35K USD, with 2 clusters with high (2, 3), 3 with medium (1, 7, 6), 2 with low (8, 5) and 1 with very low revenue (4).

Table 6 contains the heat map analysis of the amenities' availability per cluster. The amenities were ranked from the most to the least often offered by the properties at the total level, and the percentages of properties offering a given amenity within a given cluster were calculated. Next, the cells were coloured by row from green to red, where green indicates the highest and red the lowest availability of a given amenity in a given cluster. Reading the heat map by column shows the clusters best equipped with the amenities (clusters 1 and 2) and those worst equipped with the amenities (cluster 4). Average availability of top 30 amenities was presented in line: *Average TOP 30*. Also, the parking-related amenities were added, indicating the highest free parking availability in clusters 6 and 8, located outside the municipal paid parking zone.

Table 5. *Final cluster centres' average values (ranked by AR) (K-means analysis, SPSS).*

Cluster		# of Properties	Annual Revenue	CRS	OTS	Max Guests	Min Stay	No of Photos
2	**Large Deluxe**	231 **2.97%**	Very high 35,121.63	Close 2.26	Close 2.78	Large 6.62	Short 1.49	Many 26.77
3	**Central Mid-size**	1,303 **16.75%**	High 17,203.13	Very close 1.73	Very close 2.31	Medium 3.61	Short 1.49	Medium 19.04
1	**Mid-stay Mid-size**	601 **7.72%**	Medium 8,562.93	Close 2.38	Close 2.71	Medium 3.96	Medium 2.08	Very many 45.57
7	**Standard Large**	594 **7.63%**	Medium 7,718.71	Close 2.44	Close 2.76	Large 6.74	Short/Medium 1.87	Medium 19.79
6	**Distant Mid-size**	830 **10.67%**	Medium 6,018.48	Very distant 8.69	Very distant 9.88	Medium 3.61	Medium 2.14	Fewer 16.09
8	**Distant Small**	1,460 **18.77%**	Low 4,453.60	Distant 4.81	Distant 5.64	Small/Medium 3.14	Medium 2.24	Fewer 15.35
5	**Central Small**	2,550 **32.78%**	Low 3,666.41	Very close 1.77	Very close 2.22	Small/Medium 3.17	Medium 2.16	Fewer 14.95
4	**Weekly Stay**	211 **2.71%**	Very low 2,272.56	Dispersed 3.63	Dispersed 4.40	Small/Medium 3.35	Long 23.71	Fewer 16.52
	Total	**7,780**	**7,916.00**	**3.23**	**3.85**	**3.73**	**2.60**	**18.96**

Table 6. *Amenities: % of listings reporting an amenity by cluster*

Amenities **% of listings reported**	**Total**	**Clusters** **1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
wireless_internet	95.9%	98.8%	99.6%	98.8%	89.1%	95.3%	94.1%	97.5%	93.8%
kitchen	94.8%	93.7%	99.1%	96.8%	93.8%	94.7%	90.1%	98.7%	94.0%
heating	90.3%	98.5%	98.3%	97.5%	82.5%	87.5%	86.5%	92.6%	86.8%
washer	89.0%	92.0%	96.1%	86.3%	93.4%	89.8%	79.3%	95.3%	89.9%
essentials	88.6%	94.0%	97.4%	97.0%	80.1%	85.4%	84.0%	90.2%	86.5%
dryer	87.4%	98.0%	98.3%	97.6%	67.8%	83.5%	81.9%	92.3%	82.7%
hair-dryer	85.3%	97.0%	97.4%	96.4%	62.1%	81.3%	80.4%	90.9%	79.7%
iron	84.7%	95.5%	97.4%	95.2%	67.8%	80.9%	78.2%	89.7%	79.4%
Tv	83.9%	89.9%	95.2%	93.2%	71.1%	78.9%	87.8%	85.9%	79.2%
hangers	82.2%	88.0%	90.0%	89.9%	73.0%	79.3%	79.2%	85.9%	78.3%
hot_water	79.1%	89.7%	87.9%	86.2%	67.3%	76.3%	71.8%	83.2%	76.1%
cooking_basics	78.7%	87.9%	92.2%	88.8%	65.9%	73.0%	73.9%	81.5%	77.0%
dishes_and_silverware	78.6%	90.5%	90.9%	87.0%	65.9%	75.4%	68.7%	82.5%	75.5%
refrigerator	76.1%	90.8%	88.7%	85.2%	69.2%	72.2%	66.0%	80.8%	71.8%
bed_linens	68.9%	82.0%	78.8%	75.6%	58.8%	66.0%	58.3%	73.1%	67.0%
stove	67.3%	82.0%	81.8%	77.7%	57.3%	62.2%	57.8%	71.0%	63.6%
shampoo	58.2%	71.0%	67.1%	68.7%	39.8%	54.7%	55.3%	54.5%	53.9%
air_conditioning	55.8%	65.7%	56.7%	52.0%	46.4%	53.5%	64.6%	53.9%	56.3%
long_term_stays_allowed	54.1%	65.6%	67.1%	59.4%	52.1%	50.1%	52.7%	50.5%	52.1%
elevator	53.6%	63.7%	63.2%	61.2%	54.5%	52.4%	47.3%	46.5%	49.5%
laptop-friendly	51.5%	61.1%	56.3%	40.3%	49.8%	52.4%	55.1%	59.9%	50.3%
hot_water_kettle	48.4%	56.4%	47.2%	48.0%	39.8%	47.2%	46.5%	47.6%	50.2%
oven	47.5%	60.6%	65.4%	50.7%	40.3%	40.2%	46.3%	60.9%	45.6%
dishwasher	45.7%	58.1%	73.6%	52.6%	39.3%	36.8%	43.1%	56.4%	43.6%
lockbox	45.5%	60.2%	66.7%	52.9%	27.5%	40.8%	37.7%	50.5%	42.5%
freezer	40.6%	50.2%	40.7%	39.4%	34.6%	38.2%	40.5%	41.6%	42.7%
wine_glasses	39.9%	50.6%	37.7%	39.8%	24.2%	39.0%	36.5%	37.7%	42.8%
patio_or_balcony	39.7%	47.6%	55.8%	42.6%	37.9%	29.8%	45.3%	43.9%	44.1%
microwave	39.5%	50.4%	57.1%	45.7%	35.5%	35.5%	33.4%	44.3%	36.0%
dining_table	39.4%	47.8%	38.5%	36.5%	32.2%	37.3%	41.3%	42.6%	41.2%
Average TOP 30	**66.3%**	**75.9%**	**76.1%**	**71.3%**	**57.3%**	**63.0%**	**62.8%**	**69.4%**	**64.4%**
free_or_street parking	39.4%	36.8%	30.7%	21.5%	50.2%	20.5%	79.5%	36.2%	67.5%
paid_parking_or_paid parking_on_premises	36.8%	43.6%	41.6%	48.4%	28.9%	46.3%	18.6%	39.7%	16.6%

The detailed cluster descriptions below (ranked from the highest to the lowest revenue) include the conclusions drawn from the cluster characteristics presented in Figure 4, Figure 5, Table 5, and Table 6.

LARGE DELUXE. Cluster 2 (2.97%): Large (6+ people) Luxury Central Location Short stay. The properties are mostly in good locations (close to CRS and/or OTS), the capacity of almost 80% is 6 or more people, they have light minimum-stay restrictions (65% of listings: 1 day), they have relatively many photos of high quality, and they are best equipped with amenities. They are also freshly refurbished (high standard), often in newly built buildings, and full-time rental spots (the lowest of all clusters' average number of blocked days: 51).

CENTRAL MID-SIZE. Cluster 3 (16.75%): Medium (4 people) Best Location Short stay. The properties in this cluster are placed in the best locations (closest to CRS and/or OTS). They are mostly 4-person, 1-bedroom properties with 1-day minimum stay (65%). They are best equipped in terms of amenities, and they are full-time rental spots (the second lowest average number of blocked days: 72).

MID-STAY MID-SIZE. Cluster 1 (7.72%): Medium (4 people) Central Location Medium stay. The listings are similar to CENTRAL MID-SIZE (3) in terms of size, quality, and amenities. However, they are more distant from the central spots and are less open to short stays (45% of listings: 1 day). They have the highest number-of-photos (avg.: 46) and are mostly part-time renters (average number of blocked days: 115).

STANDARD LARGE. Cluster 7 (7.63%): Large (6+ people) Standard Central Location Medium stay. The properties are in good locations; the capacity of 95% is 6 or more people; and they are similar to LARGE DELUXE (2) but with lower standards: they offer fewer amenities, and the kitchens in general are less equipped and smaller. They also have fewer photos, have stricter minimum-stay rules (52% of listings: 1 day), and are mostly part-time renters (average number of blocked days: 113).

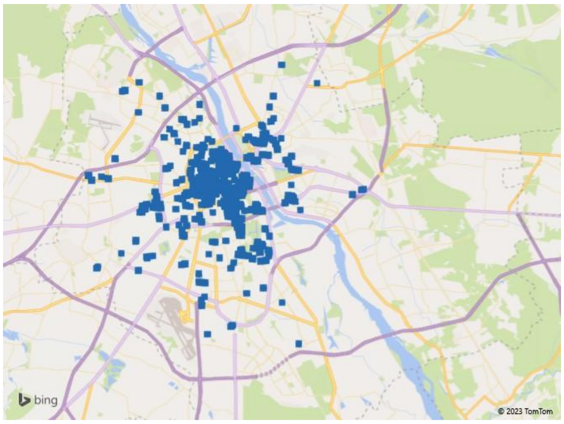
DISTANT MID-SIZE. Cluster 6 (10.67%): Medium (4 people) Very Distant Short stay. This cluster is determined geographically: locations are very distant from the central spots (avg. 8.7 km from CRS and 9.9 km from OTS). The capacity is predominantly 4 people. They are open for short stays (60%) and are equipped with average amenities. Their strength is free parking (64% vs. avg. 29%); their location outside paid-parking municipal zones might be appealing to visitors arriving by car. Their average number of blocked days (116) is similar to those of other medium-revenue clusters, MID-STAY MID-SIZE (1) and STANDARD LARGE (7).

DISTANT SMALL. Cluster 8 (18.77%): Small/Medium (2–4 people) Distant Medium stay. The DISTANT SMALL (8) and CENTRAL SMALL (5) clusters are similar in terms of size but differ in distance to the central points: DISTANT SMALL (8) is between 2.5 and 8.5 km away from both central points, in contrast to the very central locations of CENTRAL SMALL (5). They are equally equipped with average amenities; the differences come in free parking (48.6% vs. 14.8% for CENTRAL SMALL (5) and 'outdoor' amenities: patio_or_balcony (44.1% vs. 29.8%) and garden_or_backyard (17.4% vs. 9.6%). Both have similar structures of minimum-stay restrictions (45% of listings in DISTANT SMALL (8) allow 1-day stays). The number of blocked days (121) is above average but lower than in CENTRAL SMALL (5).

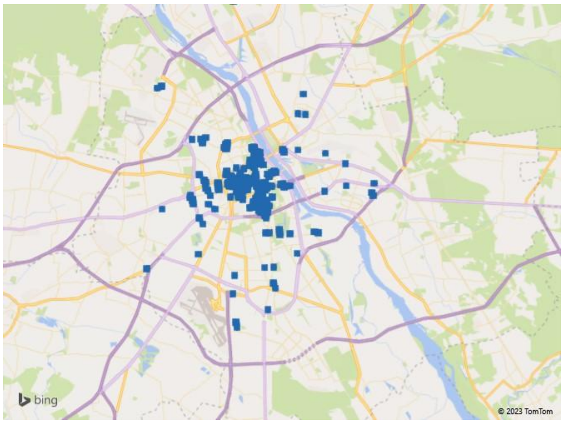
CENTRAL SMALL. Cluster 5 (32.78%): Small/Medium (2–4 people) Best Location Medium stay. The listings have the best locations, in the central circle of the city with around a 5 km radius encompassing both central spots. The size is mostly 2–4 people, and the properties are equipped with average amenities. Half (50%) of the listings allow 1-day stays. For locations in the strict city-centre, 46.3% have paid parking, with free parking available for only 14.6%. The average numberof blocked days is above average (134). Despite better locations, the average revenue is lower than for DISTANT SMALL (8).

Figure 4. *Cluster characteristics (variable % distribution in clusters). Author's own elaboration.*

Cluster 1. MID-STAY MID-SIZE



Cluster 2. LARGE DELUXE



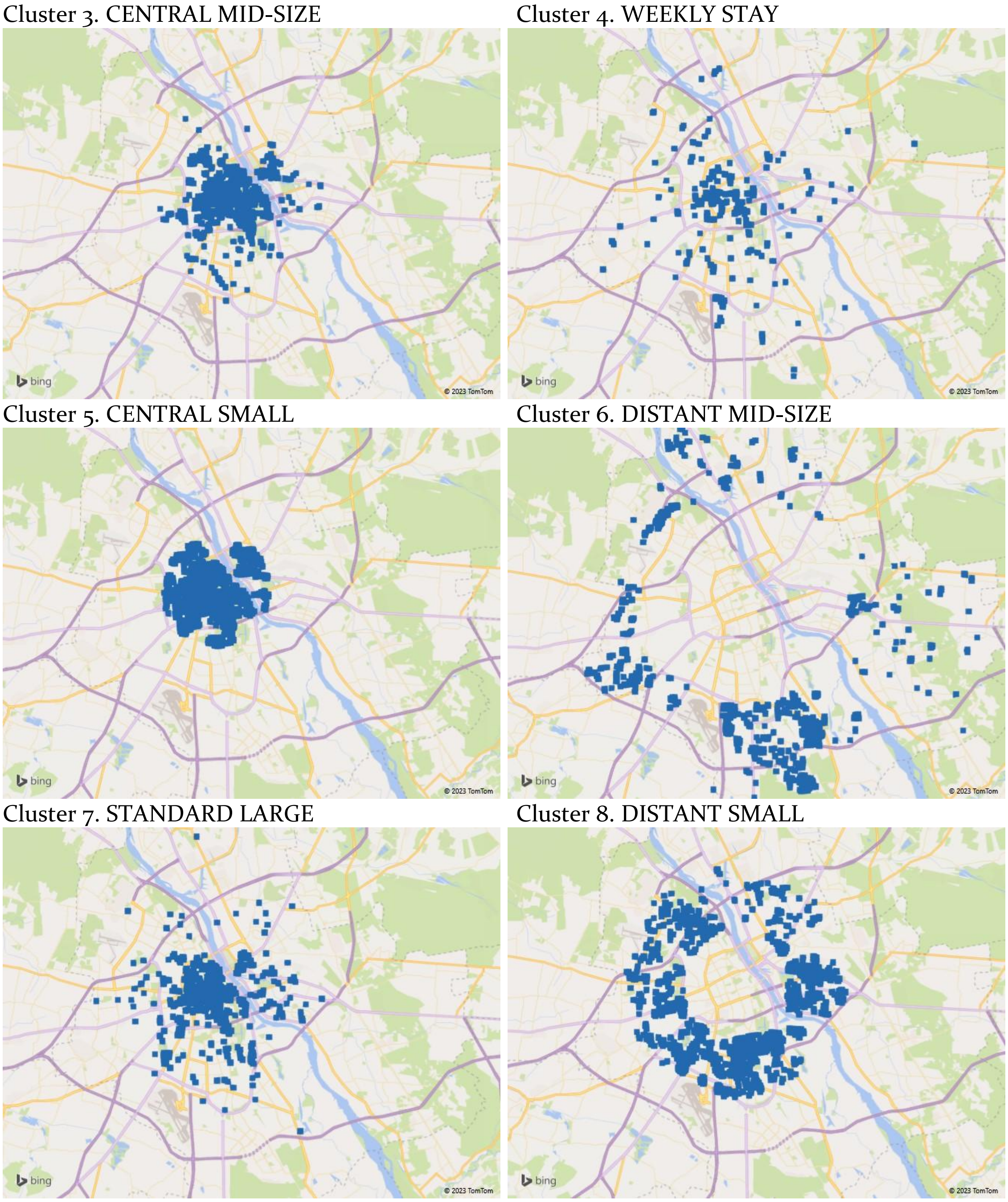


Figure 5. *Location by cluster. Author's own elaboration (Microsoft Excel 3D Maps).*

WEEKLY STAY. Cluster 4 (2.71%): Small/Medium (2–4 people) Weekly stay. These listings are characterised by their weekly stay restriction: all the listings only allow stays longer than 7 days. This cluster has also the strictest CP: more than 40% of listings set firm or strict CP. The properties are dispersed in terms of location and have mainly 2- or 4-person capacity. They are the least equipped with amenities (lowest availability of dryer, hairdryer, iron, TV). In terms of revenue, this is the least favourable option; however, this might be a sign of an opportunistic approach of these hosts, who not only rent for a minimum of 8 days with the strictest CP but also rent only periodically (the highest of all clusters' average numberof blocked days: 182).

In summary, the highest yearly revenues are obtained by LARGE DELUXE (2). It is a very unique group, accounting only for 2.97% of all home/apartment properties. The properties are top quality, are very expensive, and meet the highest expectations of potential guests. The second highest yearly revenue is achieved by CENTRAL MID-SIZE (3). This group is bigger and more mainstream, accounting for 16.5% of all home/apartment properties. Both clusters are characterised by the shortest minimum-stay restrictions, are the best equipped with amenities, and are the most active.

The third highest (but substantially lower) revenues are delivered by MID-STAY MID-SIZE (1). In comparison to CENTRAL MID-SIZE (3), these properties are in worse (but still central) locations, less prepared for 1-day minimum stays, and less equipped with amenities. The next two clusters, STANDARD LARGE (7) and DISTANT MID-SIZE (6), delivering medium revenue, are similar to the best performing clusters in terms of size but differ in terms of quality and minimum stay (STANDARD LARGE, 7) or location (DISTANT MID-SIZE, 6). All three medium-revenue clusters are primarily part-time renters. To increase their revenue, they would need to improve quality, shorten their minimum-stay requirements, or become more active. In the case of DISTANT MID-SIZE (6), it would be advisable to highlight the free parking and quietness of surroundings to attract guests arriving by car and/or families.

The clusters with low revenues (DISTANT SMALL, 8, and CENTRAL SMALL, 5) represent properties of similar size (2–4 people) and minimum-stay restrictions but different location patterns. They are quite numerous (in total, they represent 51.5% of all properties) and thus represent significant growth potential: increasing the number-of-photos, sticking to a 1-day minimum stay requirement, and decreasing the number of blocked days seem to be the easiest paths to grow the revenue. The lowest revenue cluster (WEEKLY STAY, 4) is the smallest and very unique (2.71%), and it contains only properties with more than a 7-day minimum-stay restriction. Although the revenue growth potential for this cluster could be high, the low revenue might be related to a deliberate strategy of the hosts to limit the hassles associated with renting, and thus revenue growth might not be a priority for these hosts.

4.3. Comparison of clusters

The Kruskal-Wallis tests performed to compare the clusters proved that the 8-cluster scenario accurately differentiated the clusters for all the employed variables (Table 7).

However, as indicated earlier, some of the clusters performed similarly in terms of particular features. The pairwise comparisons (Figure 8) indicate that for AR (Figure 8a), two pairs of clusters—1, 7 and 5, 8—did not significantly differ from each other. Clusters 1 and 7 consist of properties with similar *distance to* CRS (Figure 8b) and OTS (Figure 8c) but with different *size* (Figure 8d) and *number-of-photos* (Figure 8f.). Clusters 5 and 8 significantly differ in their *distance to* CRS (Figure 8b) and OTS (Figure 8c), but they achieve similar AR, which might be an indication of customer preferences (e.g., for free parking). Further, examining Figures 8b and 8c, clusters 2, 7, and 1 are similar in terms of *distance to* CRS and OTS (mostly central locations) but differ in terms of *size* (clusters 2 and 7 vs. 1, see Figure 8d), *minimum-stay* requirements (clusters 1 and 7 vs. 2, see Figure 8e) and *number-of-photos* (with all three clusters showing significant pairwise differences, see Figure 8f). Another two clusters with similar *distance to* CRS and OTS are clusters 3 and 5 (Figures 8b and 8c), but they differ on all the remaining variables. For *size* (Figure 8d), three groups of clusters had similar *maximum-number-of-*

guests: 2 to 4 people (clusters 4, 5, and 8); 6 or more people (clusters 2 and 7); and predominantly 4 people (clusters 3 and 6) (for details, see Figure 4). For *number-of-photos* (Figure 8f), there were four similar clusters (clusters 4, 5, 6, and 8) whose average *number-of-photos* ranged from 14.95 to 16.52 (Table 5). Moreover, clusters 3 and 7 presented a similar pattern (Figure 8f) of around 80% of properties having between 11 and 30 pictures (averages of 19.04 and 19.79 respectively, see Table 5). In summary, the clusters proved to be significantly differentiated from each other; however, pairwise, some clusters showed similar distributions of certain features.

Table 7. *Kruskal-Wallis Test results for variables' distribution in clusters.*
Hypothesis Test Summary (Independent-Samples Kruskal-Wallis Test)

	Null Hypothesis	Test statistics	Sig.[a,b]	Decision
1	The distribution of Annual Revenue LTM (USD) is the same across Cluster categories.	3638.577[c]	<0.001	Reject the null hypothesis.
2	The distribution of CentralRailwayStation is the same across Cluster categories.	4428.500[c]	<0.001	Reject the null hypothesis.
3	The distribution of OldTown is the same across Cluster categories.	4451.439[c]	<0.001	Reject the null hypothesis.
4	The distribution of Max Guests is the same across Cluster categories.	2457.796[c]	<0.001	Reject the null hypothesis.
5	The distribution of Minimum Stay is the same across Cluster categories.	936.232[c]	<0.001	Reject the null hypothesis.
6	The distribution of Number of Photos is the same across Cluster categories.	2138.405[c]	<0.001	Reject the null hypothesis.

a. The significance level is 0.050.
b. Asymptotic significance is displayed.
c. The test statistic is adjusted for ties.

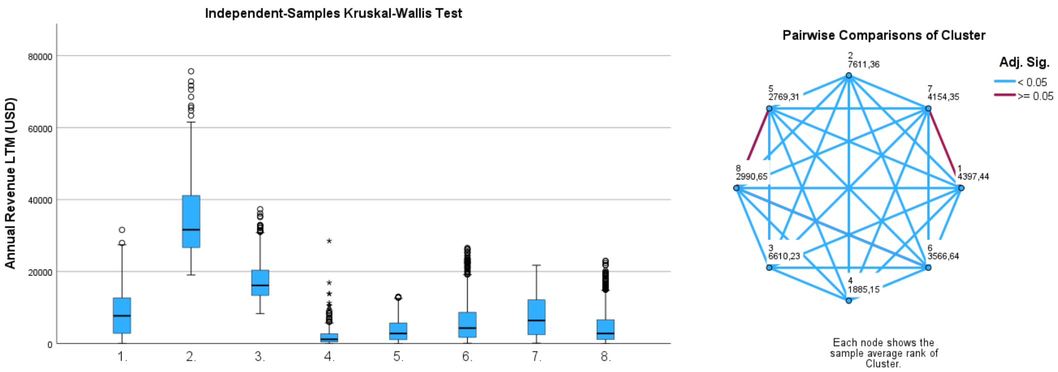


Figure 8a. *Clusters' comparison: Annual Revenue.*

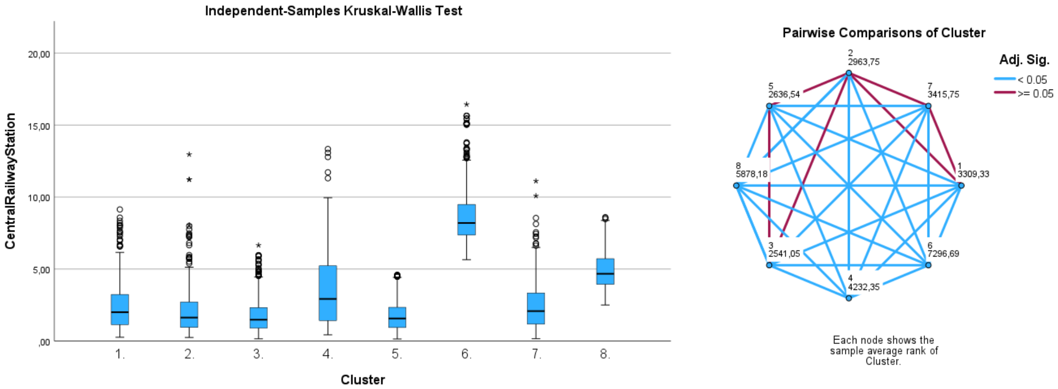


Figure 8b. *Clusters' comparison: distance-to Central Railway Station.*

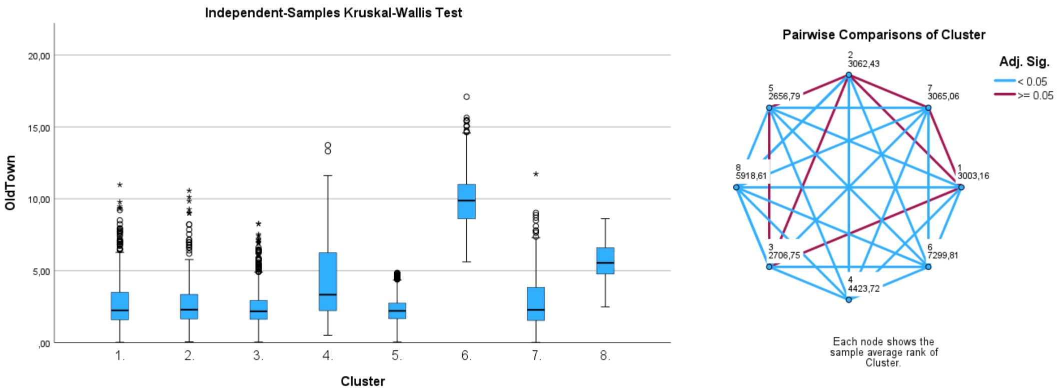


Figure 8c. *Clusters' comparison: distance-to Old Town Square.*

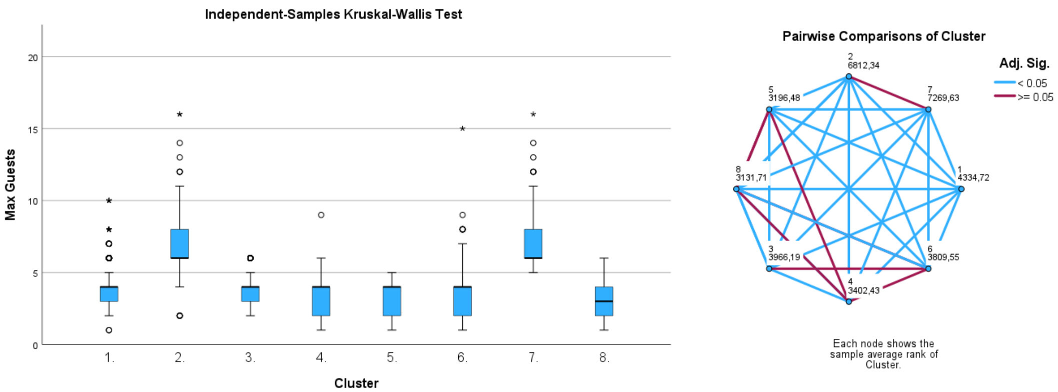


Figure 8d. *Clusters' comparison: maximum number of guests.*

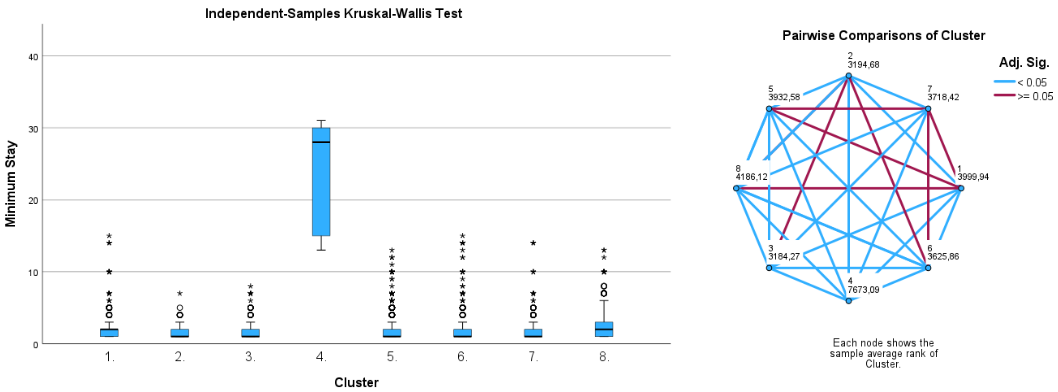


Figure 8e. *Clusters' comparison: minimum stay.*

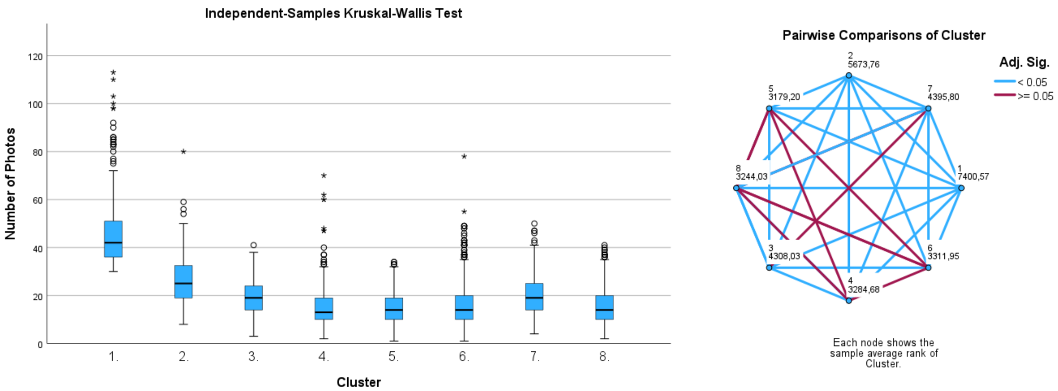


Figure 8f. *Clusters' comparison: number of photos.*

5. Discussion

The purpose of this study was to explore the diversity of offers in P2PA at an urban tourism destination. The proposed method of K-means clustering, employing P2PA performance determinants and revenue as the performance measure, delivered valuable results from both theoretical and practical perspectives. The research questions were addressed by providing a set of factors applicable for clustering of P2PA properties; proposing the clusters based on properties' concrete urban destination location; and outlining the detailed characteristics of those clusters, including thorough testing of their dissimilarities and pairwise similarities. The correlation study, although not the main purpose of this paper, allowed for some directional comparisons to the existing research results on P2PA performance determinants.

The findings of previous research regarding the significant negative impact of properties' distance to central spots suggested that the farther the accommodation is located from the city centre, the lower would be its price (Wang & Nicolau, 2017; Gibbs *et al.*, 2018a; Benítez-Aurioles, 2018a; Abrate & Viglia, 2019; Sainaghi & Chica-Olmo, 2022). The correlation analyses of the current study supported these findings. A significant negative relation of ADR to both the transportation hub and the main touristic highlight was estimated; however, using revenue as the performance indicator resulted in a slightly stronger negative relation compared to ADR. It was the result of the significant negative relation of location features to the occupation rate, confirming that properties closer to the central spots have higher occupation rates.

The findings of this study go further than previous research by revealing the diversity of the offerings in P2PA. For example, the clusters DISTANT SMALL (8) and CENTRAL SMALL (5) are very similar in terms of size, minimum-stay restriction, and amenities; however, on average, the more distant cluster (DISTANT SMALL, 8) achieves revenues that are higher by 21.5% with almost the same ADR level (3% difference). This can be partly explained by the substantially higher availability of free parking, which has been previously found to have a significant positive effect on prices (Wang & Nicolau, 2017), and by the outdoor amenities (patio_or_balcony and garden_or_backyard) in DISTANT SMALL (8) that can attract tourists, especially those arriving by car. Browsing through the guests' comments in CENTRAL SMALL (5), it might also be the city-centre noise and night-life intensity that discourage some guests from the best-located listings and prompt them to redirect to more remote places. Another example is the pair of LARGE DELUXE (2) and STANDARD LARGE (7), which are very similar in terms of location and size; however, due to differences in quality (in terms of amenities and interior design/building), minimum-stay policies, and the level of availability, listings in STANDARD LARGE (7) deliver substantially lower average revenues. Further, the results also indicate that medium-size properties (4 people, MID-STAY MID-SIZE, 1) can achieve higher revenues than large properties (6+ people, STANDARD LARGE, 7), even if their locations are very similar. The explanation here seems to be the quality of the furnishings, interior design, amenities, or ability to meet guests' needs; 4-person properties account for almost 45% of all entire home/apartment listings, of which 44.5% fall into high- or medium-revenue clusters. This also suggests that despite the high revenue potential of large properties (6+ people) in top locations, the scale of demand for such properties might be limited, as they only account for 12.6% of all offerings; still, 94.8% of them fall into medium- or high-revenue clusters.

As shown, looking at the combinations of property attributes may provide valuable insights on how revenue can be developed in a business with highly differentiated offerings like P2PA. Moreover, based on the correlation study, some additional conclusions might be drawn regarding the relation of certain property characteristics to revenue. Of the three indicators of property size (*maximum-number-of-guests, number-of-bedrooms, number-of-bathrooms*), *maximum-number-of-guests* turned out to have the strongest relation to revenue. In fact, *the number-of-bedrooms* did not even show a significant correlation to revenue. This might potentially be a very valuable insight, as it suggests the guests most value the capacity of the property, while the *number-of-bedrooms* is of secondary importance. This contradicts previous research results, where the parameters for both *number-of bedrooms* and *number-of-bathrooms* were significant and also had higher values than *number-of-guests* (Wang & Nicolau, 2017; Perez-Sanchez *et al.*, 2018; Sainaghi *et al.*, 2021). The relationship between minimum-stay requirements and revenue is also significant—in fact, it is negative and stronger than both location indicators—suggesting that in urban settings, lighter minimum-stay restrictions are associated with higher revenue. These results are in line with Sainaghi *et al.* (2021), where having any minimum-stay restriction was significantly negatively related to revenue. Further, considering CP, Wang and Nicolau (2017) found non-flexible CPs ("moderate" to "strict") had a significant positive effect on price; Sainaghi *et al.* (2021) found the same effect on price and revenue; and the significant positive relation to price was also visible in Perez-Sanchez *et al.* (2018). In this research, the importance of CP was limited, which was unexpected in light of the post-COVID-19 situation, where high cancellation flexibility would be expected to be highly appreciated by guests. Regarding Superhost status, only 8% of properties were managed by Superhosts (similar to Wang & Nicolau, 2017, but considerably lower than Sainaghi *et al.*, 2021: 22.34%), and the relation between Superhost status to revenue was significant and positive but still relatively weak (again similar to Wang & Nicolau, 2017; Sainaghi *et al.*, 2021). Similarly, a significant but weak correlation to revenue was observed for pets-allowed variable. However, surprisingly, this relation was negative; this suggests that guests, rather than being prepared to pay a premium for the option of travelling with pets, were instead prepared to pay less for properties that allowed pets, perhaps expecting them to be less clean than those that did not. Finally, the relatively strong, significant, and positive relation of the number-of-photos variable to revenue, in addition to indicating clear communication of the properties' appearance, could be partly a side effect of the property size. All three size-related variables (number-of-bedrooms, -bathrooms, and -guests) showed a significant correlation to number-of-photos, suggesting that the bigger a property was, the more unique photos one could take. The positive relation of number-of-photos to performance was also confirmed by Perez-Sanchez *et al.* (2018), albeit with a weaker relation.

Regarding amenities, previous research found a significantly positive influence of the provision of wireless Internet on price (Wang & Nicolau, 2017). The current study showed that, being provided by 95.9% of properties, wireless Internet has become more of a hygiene factor rather than a price or revenue determinant. Also, it shed some light on the role of free parking in urban areas, showing that more distant locations might better suit guests arriving by car.

6. Conclusions, limitations, and future research

6.1. Theoretical implications

This research has several valuable theoretical implications for the research stream on P2PA. First, based on performance determinants research, a typology of P2PA offers' features is proposed, employing a property-host-guest triangle. This approach offers theoretical implications for future studies on performance determinants by stressing the roles of primary (property-host) and secondary (guest) factors. The guest-related factors (reviews and ratings) stem from the guests' perception and evaluation of certain combinations of property- and host-related features, and therefore their moderating role should be considered in future research regarding P2PA performance determinants. Also, the proposed typology lays a foundation for further research exploring the diversity of P2PA listings, which—in light of the fact that the diversity of the offers is one of the intrinsic features of P2PA—should be further developed. Second, a method of unsupervised machine learning (K-means clustering algorithm) is proposed to explore the diversity of P2PA offers. Previous research mostly focused on regression analysis of price determinants and, recently, on investigating the significance of some developing features, but the diversity of the offerings has not been sufficiently researched. Third, based on correlation analysis, some of the previous research findings regarding revenue determinants in P2PA were either confirmed or challenged by this study. Specifically, it confirmed the strong and significant relation to revenue of *maximum-number-of-guests* (positive), *distance to* central areas (negative), *minimum-stay* (negative), and *number-of-photos* (positive); however, it did not confirm the *number-of-bedrooms* as a significant revenue determinant. It also pointed out that *Superhost* status and *CPs*, although significant, show substantially weaker relations to revenue compared to the aforementioned determinants. Surprisingly, the *pets-allowed* variable showed a weakly but significantly negative relation to revenue, which was nonintuitive. As such, this research adds to the hedonic price theory in P2PA and extends its findings by suggesting that in the environment of P2PA, where the diversity of the offers is high, the performance determinants may not have a unidirectional relation to performance in every location. Fourth, this study expands the relatively scarce research in this area on revenue as opposed to price. Revenue appears to be a broader performance indicator, as it encompasses not only price but also occupancy rate and level of availability (i.e., number of blocked days). Finally, this study extends European representation in the P2PA research stream, using an urban destination in the Central and Eastern Europe region, which is rarely researched in this context.

6.2. Practical implications

This research also offers practical implications for at least two groups of short-term tenancy providers (hosts). First, it can serve as a source of insights on how to rearrange existing listings and where to put emphasis regarding the property features and host behaviours. For example, for a studio in a central location, it is advisable to arrange the property for 4 people (1 double bed, 1 sofa), provide modern furnishings and kitchen equipment, set no minimum-stay restriction, emphasise the short distance to CRS and/or OTS, and (if applicable) highlight the quiet neighbourhood. For a larger listed property (6+ people) in a central location, increased standards (modern interior, freshly refurbished, amenities) could be considered, and a higher number-of-photos and no minimum-stay restrictions could be implemented. In general, as the number-of-guests appeared to have a higher correlation to revenue than number-of-rooms or bathrooms, the potential increase in revenue is relatively easy to achieve: fitting one more sofa into an apartment is much easier than adding a bedroom or bathroom. Therefore, wherever possible, extending the maximum capacity of a property should improve its revenue potential, regardless of its distance to central areas. Another implication for current P2PA providers in urban locations is the negative effect of a long (i.e., weekly) minimum-stay policy. While such a practice could be justifiable for seaside or ski resorts, where people often travel to spend their holidays for a week or longer, urban travel might present a different pattern of shorter stays (e.g. 1 night or weekend). Therefore, although a weekly minimum-stay policy limits the costs and hassle related to cleaning, change-over, and guest reception, it might repel guests coming to the urban location for shorter stays, thus limiting revenue. To minimize the complexity related to guest reception for shorter stays, the solutions of a lockbox, smart-lock, or keypad could be considered. The popularity of these solutions has grown recently due to the COVID-19 pandemic, as they allow for contactless guest reception; however, they also serve the purpose of lowering the host's required engagement in the reception process.

Second, for future hosts searching for new investments, this research could serve as a benchmark of the combinations of property features and renting policies that would potentially lead to high revenue. Based on this particular case, the best option would be either a luxury large (6-person minimum) property in a top location or a 4-person quality property in the best location. Alternatively, a 4-person quality property with outdoor amenities in a distant location could be a viable option if free parking is highlighted to attract guests arriving by car. In any case, a 1-day minimum-stay policy and reasonable number-of-photos (around 27 for a 6-person or 19 for a 4-person property) would be advisable. As indicated earlier, because of the stronger relation to revenue of maximum-number-of-guests vs. number-of-bedrooms and -bathrooms, it is advisable to look for options to increase the maximum-number-of-guests, even if the other size-related features cannot be increased.

6.3. Limitations and future research

The main limitation of this research is that it covers only one urban destination. However, the methodology can be easily replicated in other urban locations—the K-means clustering algorithm is easy and fast to perform and is especially useful for large sets of data (Papamichail & Papamichail, 2007). This would be an interesting future research option. Another limitation is that this research covered a 12-month period. Although this period allows us to capture seasonality, it might be biased due to any exceptional events taking place during this period. However, it is worth stressing that this approach is very flexible and thus can be repeated systematically to trace any trends in the short-term tenancy landscape. As such, it can serve as guidance for investors on how to shape their portfolios in the long term, which would be another advisable future research direction. Another future research path would be further investigation of the location variable. Following previous studies, this paper employed a transportation hub and a touristic highlight as the two reference points for the location factor. It would be useful to carry out an enhanced spatial analysis of other location factors, such as access to public transportation (e.g., distance to underground stations), time to get to the city centre (vs. distance), or natural barriers (e.g., a river) that may affect the attractiveness of the listing.

Acknowledgments
This work was financed by Kozminski University (Warsaw, Poland) Research Fund.

References

Airbnb (2023a). Choose the right cancellation policy for you. Retrieved from https://www.airbnb.co.uk/resources/hosting-homes/a/choose-the-right-cancellation-policy-for-you-19?_set_bev_on_new_domain=1694876528_ZWQ2YzU2MjY3YzEx (Accessed on 2.10.2023).

Airbnb (2023b). How to become a Superhost. Retrieved from https://www.airbnb.co.uk/help/article/829/ (Accessed on 26.10.2023).

AirDNA (2023). About AirDNA. Retrieved from: https://www.airdna.co/about (Accessed on 2.10.2023).

Abrate, G., Sainaghi, R., & Mauri, A. G. (2022). Dynamic pricing in Airbnb: Individual versus professional hosts. *Journal of Business Research*, 141(December 2021), 191–199. https://doi.org/10.1016/j.jbusres.2021.12.012

Abrate, G., & Viglia, G. (2019). Personal or Product Reputation? Optimizing Revenues in the Sharing Economy. *Journal of Travel Research*, 58(1), 136–148. https://doi.org/10.1177/0047287517741998

Akbar, Z., Liu, J., & Latif, Z. (2020). Discovering knowledge by comparing silhouettes using K-means clustering for customer segmentation. *International Journal of Knowledge Management*, 16(3), 70–88. https://doi.org/10.4018/IJKM.2020070105

Akbari, M., Foroudi, P., Khodayari, M., Rahime Zaman, F., Parizi, Z., & Shahriari, E. (2022). Sharing Your Assets: A Holistic Review of Sharing Economy. *Journal of Business Research*, 140(May 2021), 604–625. https://doi.org/10.1016/j.jbusres.2021.11.027

Anitha, P., & Patil, M. M. (2022). RFM model for customer purchase behavior using K-Means algorithm. *Journal of King Saud University - Computer and Information Sciences*. The Authors, 34(5), 1785–1792. https://doi.org/10.1016/j.jksuci.2019.12.011

Benítez-Aurioles, B. (2018a). The role of distance in the peer-to-peer market for tourist accommodation. *Tourism Economics*, 24(3), 237–250. https://doi.org/10.1177/1354816617726211

Benítez-Aurioles, B. (2018b). Why are flexible booking policies priced negatively? *Tourism Management*, 67, 312–325. https://doi.org/10.1016/j.tourman.2018.02.008

Casamatta, G., Giannoni, S., Brunstein, D., & Jouve, J. (2022). Host type and pricing on Airbnb: Seasonality and perceived market power. *Tourism Management*, 88(September 2021), p. 104433. https://doi.org/10.1016/j.tourman.2021.104433

Chattopadhyay, M., & Mitra, S. K. (2019). Do airbnb host listing attributes influence room pricing homogenously? *International Journal of Hospitality Management*, 81(March), 54–64. https://doi.org/10.1016/j.ijhm.2019.03.008

Chen, Y., & Xie, K. L. (2017). Consumer valuation of Airbnb listings: a hedonic pricing approach. *International Journal of Contemporary Hospitality Management*, 29(9), 2405–2424. https://doi.org/10.1108/IJCHM-10-2016-0606

Chen, A. H., Liang, Y. C., Chang, W. J., Siauw, H. Y., & Minanda, V. (2022). RFM Model and K -Means Clustering Analysis of Transit Traveller Profiles: A Case Study. *Journal of Advanced Transportation*, 2022. https://doi.org/10.1155/2022/1108105

Cui, M. (2020). Introduction to the K-Means Clustering Algorithm Based on the Elbow Method. *Accounting, Auditing and Finance*, 1: 5–8. https://doi.org/10.23977/accaf.2020.010102

Deboosere, R., Kerrigan, D. J., Wachsmuth, D., & El-Geneidy, A. (2019). Location, location and professionalization: A multilevel hedonic analysis of airbnb listing prices and revenue. *Regional Studies, Regional Science*, 6(1), 143–156. https://doi.org/10.1080/21681376.2019.1592699

Dwivedi, S., & P.Bhaiya, L. K. (2019). A Systematic Review on K-Means Clustering Techniques. *International Journal of Scientific Research Engineering Trends (IJSRET)*, 5(3), 750–752.

Euromonitor (2021). Top 100 City Destinations. Index 2021. Euromonitor International.

Fernández-Durán, J. J., & Gregorio-Domínguez, M. M. (2021). Consumer Segmentation Based on Use Patterns. *Journal of Classification*, 38(1), 72–88. https://doi.org/10.1007/s00357-019-09360-2

Ferreira, J. P., Bijen Saha, B., Carrero, G. C., Kim, J., & Court, C. (2023). Impacts of red tide in peer-to-peer accommodations: A multi-regional input-output model. *Tourism Economics*, 29(3), 812–834. https://doi.org/10.1177/13548166211068276

Filieri, R., Raguseo, E., & Galati, F. (2023). Negative signals on Peer-to-Peer platforms: The impact of cancellations on host performance across different property types. *International Journal of Hospitality Management*, 114(May), p. 103564. https://doi.org/10.1016/j.ijhm.2023.103564

Gautam, N., & Kumar, N. (2022). Customer segmentation using k-means clustering for developing sustainable marketing strategies. *Business Informatics*, 16(1), 72–82. https://doi.org/10.17323/2587-814X.2022.1.72.82

Gibbs, C., Guttentag, D., Gretzel, U., Morton, J., & Goodwill, A. (2018a). Pricing in the sharing economy: a hedonic pricing model applied to Airbnb listings. *Journal of Travel & Tourism Marketing*, 35(1), 46–56. https://doi.org/10.1080/10548408.2017.1308292

Gibbs, C., Guttentag, D., Gretzel, U., Yao, L., & Morton, J. (2018b). Use of dynamic pricing strategies by Airbnb hosts. *International Journal of Contemporary Hospitality Management*, 30(1), 2–20. https://doi.org/10.1108/IJCHM-09-2016-0540

Gyódi, K., & Nawaro, Ł. (2021). Determinants of Airbnb prices in European cities: A spatial econometrics approach. *Tourism Management*, 86(March). https://doi.org/10.1016/j.tourman.2021.104319

He, J., Li, B., & Wang, X. (Shane) (2023). Image features and demand in the sharing economy: A study of Airbnb. *International Journal of Research in Marketing*, 40(4), 760-780. https://doi.org/10.1016/j.ijresmar.2023.04.001

Hernández, J. M., Bulchand-Gidumal, J., & Suárez-Vega, R. (2021). Using accommodation price determinants to segment tourist areas. *Journal of Destination Marketing and Management*, 21. https://doi.org/10.1016/j.jdmm.2021.100622

Jain, A. K, & Dubes, R. C. (1988). *Algorithms for Clustering Data*, Prentice Hall, Englewood Cliffs, New Jersey.

Jang, S., & Kim, J. (2022). Remedying Airbnb COVID-19 disruption through tourism clusters and community resilience. *Journal of Business Research*, 139(October 2021), 529–542. https://doi.org/10.1016/j.jbusres.2021.10.015

Jang, S., Kim, J., Kim, J., & Kim S. S. (2021). Spatial and experimental analysis of peer-to-peer accommodation consumption during COVID-19. *Journal of Destination Marketing and Management*, 20(January), p. 100563. https://doi.org/10.1016/j.jdmm.2021.100563

Kiczmachowska, E. E. (2022a). Revenue Management Practices in Peer-to-Peer Accommodation: The Case of Airbnb. *Central European Management Journal*, 30(1), 135–152. https://doi.org/10.7206/cemj.2658-0845.78

Kiczmachowska, E. E. (2022b). The impact of the COVID 19 pandemic on peer to peer accommodation businesses: The case of Airbnb. *Journal of Economics and Management*, 44, 286–314. https://doi.org/10.22367/jem.2022.44.12

Koh, Y., Belarmino, A., & Kim, M. G. (2020). Good fences make good revenue: An examination of revenue management practices at peer-to-peer accommodations. *Tourism Economics*, 26(7), 1108–1128. https://doi.org/10.1177/1354816619867579

Kreeger, J. C., & Smith, S. (2017). Amateur innkeepers utilization of minimum length stay restrictions. *International Journal of Contemporary Hospitality Management*, 29(9), 2483–2496. https://doi.org/10.1108/IJCHM-09-2016-0502

Kwok, L., & Xie, K. L. (2019). Pricing strategies on Airbnb: Are multi-unit hosts revenue pros? *International Journal of Hospitality Management*, 82(April 2018), 252–259. https://doi.org/10.1016/j.ijhm.2018.09.013

Lancaster, K.J. (1966). A New Approach to Consumer Theory. *Journal of Political Economy*, 74, 132-157. https://doi.org/10.1086/259131

Larasati, A., Maren, R., & Wulandari, R. (2021). Utilizing Elbow Method for Text Clustering Optimization in Analyzing Social Media Marketing Content of Indonesian e-Commerce. *Jurnal Teknik Industri*, 23(2), 111–120. https://doi.org/10.9744/jti.23.2.111-120

Liang. S., Schuckert. M., Law. R., & Chen, C. C. (2017). Be a Superhost: The importance of badge systems for peer-to-peer rental accommodations. *Tourism Management*, 60, 454–465. https://doi.org/10.1016/j.tourman.2017.01.007

Liu, Y., Du, X., & Ma, S. (2023). Innovative study on clustering center and distance measurement of K-means algorithm: mapreduce efficient parallel algorithm based on user data of JD mall. *Electronic Commerce Research*. https://doi.org/10.1007/s10660-021-09458-z

Ma, S., Cui, X., Xiao, X., & Zhao, X. (2022). The impact of photo verification service on sales performance in the peer-to-peer economy: Moderating role of customer uncertainty. *Journal of Business Research*, 142(December 2021), 45–55. https://doi.org/10.1016/j.jbusres.2021.12.021

Magno, F., Cassia, F., & Ugolini, M. M. (2018). Accommodation prices on Airbnb: effects of host experience and market demand. *TQM Journal*, 30(5), 608–620. https://doi.org/10.1108/TQM-12-2017-0164

Narasimhan, C., Papatla, P., Jiang, B., Kopalle, P., Messinger, P., Moorthy, S., Proserpio, D., Subramanian, U., Wu, C., & Zhu, T. (2018). Sharing Economy: Review of Current Research and Future Directions. *Customer Needs and Solutions*, 5(1–2), 93–106. https://doi.org/10.1007/s40547-017-0079-6

Nieto García M., Muñoz-Gallego P. A., Viglia, G., & González-Benito, Ó. (2020). Be Social! The Impact of Self-Presentation on Peer-to-Peer Accommodation Revenue. *Journal of Travel Research*, 59(7), 1268–1281. https://doi.org/10.1177/0047287519878520

Olejniczak, T. (2021), Innovativeness of Senior Consumers' Attitudes-An Attempt to Conduct Segmentation. *Folia Oeconomica Stetinensia*, 21(1), 76–91. https://doi.org/10.2478/foli-2021-0006

Papamichail, G. P., & Papamichail, D. P. (2007). The k-means range algorithm for personalized data clustering in e-commerce. *European Journal of Operational Research*, 177(3), 1400–1408. https://doi.org/10.1016/j.ejor.2005.04.011

Perez-Sanchez, R., Serrano-Estrada, L., Marti, P., & Mora-Garcia, R. T. (2018). The what, where, and why of airbnb price determinants. *Sustainability (Switzerland)*, 10(12). https://doi.org/10.3390/su10124596

Pradana, M. G., & Ha, H. T. (2021). Maximizing Strategy Improvement in Mall Customer Segmentation using K-means Clustering. *Journal of Applied Data Sciences*, 2(1), 19–25. https://doi.org/10.47738/jads.v2i1.18

Ribes, J. F., Izquierdo, L., Rodríguez, A., & Such Devesa, M. J. (2018). The rental prices of the apartments under the new tourist environment: A hedonic price model applied to the Spanish sun-and-beach destinations. *Economies*, 6(2), 1–9. https://doi.org/10.3390/economies6020023

Sainaghi, R. (2020). The current state of academic research into peer-to-peer accommodation platforms. *International Journal of Hospitality Management*, 89(April), p. 102555. https://doi.org/10.1016/j.ijhm.2020.102555

Sainaghi, R. (2021). Determinants of price and revenue for peer-to-peer hosts. The state of the art. *International Journal of Contemporary Hospitality Management*, Vol. 33 No. 2, 557-586. https://doi.org/10.1108/IJCHM-08-2020-0884

Sainaghi, R., & Baggio, R. (2020). Clusters of topics and research designs in peer-to-peer accommodation platforms. *International Journal of Hospitality Management*, 88(March 2019), p. 102393. https://doi.org/10.1016/j.ijhm.2019.102393

Sainaghi, R., Abrate, G., & Mauri, A. (2021). Price and RevPAR determinants of Airbnb listings: Convergent and divergent evidence. *International Journal of Hospitality Management*, 92(October 2020), p. 102709. https://doi.org/10.1016/j.ijhm.2020.102709

Sainaghi, R., & Chica-Olmo, J. (2022). The effects of location before and during COVID-19: Impacts on revenue of Airbnb listings in Milan (Italy). *Annals of Tourism Research*, 96, p. 103464. https://doi.org/10.1016/j.annals.2022.103464

Sharaf Addin, E. H., Admodisastro, N. I., Mohd Ashri, S. N. S., Kamaruddin, A., & Chew, Y. C. (2022). Customer Mobile Behavioral Segmentation and Analysis in Telecom Using Machine Learning. *Applied Artificial Intelligence*, 36(1). https://doi.org/10.1080/08839514.2021.2009223

Sujatha, V., Najiya, S., Likhitha, T. S., Sravya, M., & Tejaswini, P. (2023). Customer Segmentation Using K-Means Clustering, in: A. B. Reddy *et al.* (eds.). *Proceedings of Third International Conference on Advances in Computer Engineering and Communication Systems*, Lecture Notes in Networks and Systems 612, 457-466, https://doi.org/10.1007/978-981-19-9228-5_38

Sundararajan, A. (2016). *The Sharing Economy: The End of Employment and the Rise of Crowd-Based Capitalism*. The MIT Press, London.

Talluri, K., & van Ryzin, G. (2005). *The Theory and Practice of Revenue Management*. Springer Science & Business Media, New York.

Tang, L. (Rebecca), Kim, J., & Wang, X. (2019). Estimating spatial effects on peer-to-peer accommodation prices: Towards an innovative hedonic model approach. *International Journal of Hospitality Management*, 81(March), 43–53. https://doi.org/10.1016/j.ijhm.2019.03.012

Tranter, K., Stuart-Hill, T., & Parker, J. (2014). *An Introduction to Revenue Management for the Hospitality Industry: Principles and Practices for the Real World*. Pearson, Prentice Hall, New York.

Wang, D., & Nicolau, J. L. (2017). Price determinants of sharing economy based accommodation rental: A study of listings from 33 cities on Airbnb.com. *International Journal of Hospitality Management*, 62, 120–131. https://doi.org/10.1016/j.ijhm.2016.12.007

Wu, J., Shi, L., Lin, W., Tsai, S., Li, Y., Yang, L., & Xu, G. (2020). An Empirical Study on Customer Segmentation by Purchase Behaviors Using a RFM Model and K -Means Algorithm. *Mathematical Problems in Engineering*, 2020(April 2019). https://doi.org/10.1155/2020/8884227

Xie, K., Heo, C. Y., & Mao, Z. E. (2021). Do professional hosts matter? Evidence from multi-listing and full-time hosts in Airbnb. *Journal of Hospitality and Tourism Management*, 47(March), 413–421. https://doi.org/10.1016/j.jhtm.2021.04.016

Xie, K., & Mao, Z. (2017). The impacts of quality and quantity attributes of Airbnb hosts on listing performance. *International Journal of Contemporary Hospitality Management*, Vol. 29 No. 9, 2240-2260. https://doi.org/10.1108/IJCHM-07-2016-0345

Yu, D., Dong, S., & Yao, S. (2022). Improvement of K-Means Algorithm and Its Application in Air Passenger Grouping. *Computational Intelligence and Neuroscience*, 2022. https://doi.org/10.1155/2022/3958423