


Article

Can Location-Based Social Media and Online Reservation Services Tell More about Local Accommodation Industries than Open Governmental Data?

Csaba Sidor ^{1,*}, Branislav Kršák ¹, Ľubomír Štrba ¹, Michal Cehlár ² , Samer Khouri ², Michal Stričík ³, Jaroslav Dugas ⁴, Ján Gajdoš ¹ and Barbora Bolechová ¹

¹ Department of Geo and Mining Tourism, Faculty of Mining, Ecology, Process Control and Geotechnologies, Technical University of Košice, Letná 9, 042 00 Kosice, Slovakia; branislav.krsak@tuke.sk (B.K.); lubomir.strba@tuke.sk (L.S.); jan.gajdos@t-system.sk (J.G.); barbora.bolechova@tuke.sk (B.B.)

² Institute of Earth Resources, Faculty of Mining, Ecology, Process Control and Geotechnologies, Technical University of Košice, Letná 9, 042 00 Kosice, Slovakia; michal.cehlar@tuke.sk (M.C.); samer.khouri@tuke.sk (S.K.)

³ Department of Economics, Faculty of Business Economy with Seat in Košice, University of Economics in Bratislava, Tajovského 13, 041 30 Kosice, Slovakia; michal.stricik@euke.sk

⁴ Department of Information and Language Communication, Faculty of Business Economy with Seat in Košice, University of Economics in Bratislava, Tajovského 13, 041 30 Kosice, Slovakia; jaroslav.dugas@euke.sk

* Correspondence: csaba.sidor@tuke.sk; Tel.: +421-55-602-3298

Received: 17 September 2019; Accepted: 21 October 2019; Published: 24 October 2019



Abstract: The paper follows-up ongoing research focusing on the potential of machine-readable data as additional knowledge in the governance of local tourism and destination management organizations (DMOs) in Slovakia. The current focus is on one classic social media (Facebook), one location-based social media (Foursquare), two hybrid travel-related platforms with partial attributes of reservation services (Google Places, TripAdvisor), and two online reservation services (Booking, Airbnb). The global aim is the usage of extracted data for the identification of additional entities with the obligation of local occupancy taxation, which is the financial backbone of Slovak (DMOs). A set of simple and globally reusable scripts constructed in Python and PostgreSQL were used to extract data on lodging providers from the Google Places application programming interface (API), the Facebook Place Search API and the Foursquare Venue API over grid overlays of districts' spatial representation. For pure scientific purposes in the case of TripAdvisor, Booking, and Airbnb, with no suitable access to open APIs, web scraping methods were used for data extraction. The pilot case was applied in the boundaries of Kosice city (Slovakia), and the aggregations of processed data were compared with official open statistics. Results indicate that the automated continuous monitoring of online platforms could help local public administrations in decreasing occupancy tax evasions and even widen knowledge about online audiences and visitors' satisfaction.

Keywords: destination management; Facebook place search API; google places API; foursquare venue API; booking; TripAdvisor; Airbnb; machine readable data

1. Introduction

Aggregates of local municipalities' hospitality related data at the regional and state levels are widely used for establishing long term local, regional, and national policies, development and marketing strategies and even funding rules. As an essential part of tourism's superstructure, accommodation service providers' data generate an enormous volume of base information, not only in terms of the economic efficiency of tourism but also in terms of gaining knowledge about tourists' (destination

target groups') profile, purchase behavior, satisfaction, etc. Furthermore, lodging providers generate the most basic tourism metrics—the number of tourists and their number of overnight stays in a destination. For example, in terms of public funding opportunities for Slovak destination management organizations (DMOs), the higher number of officially recorded overnight stays means a larger volume of levied occupancy tax [1]. The larger the volume of levied occupancy tax is, the higher is the limit for potential maximum public funding [1]. For this reason and others, official statistics should be open at the local level, and they should also be precise and trustworthy.

The purpose of the paper is to demonstrate on the case of Kosice city (Slovakia) the potential value of machine-readable data extracted from one classic social media (Facebook), one location-based social media (Foursquare), two hybrid travel-related platforms with partial attributes of reservation services (Google Places, TripAdvisor), and two online reservation services (Booking, Airbnb) within the processes of identification of entities with the potential obligation of local occupancy taxation and the capability to generate additional knowledge about local destinations' markets. In terms of significance, the created extraction scripts not only prove the distortion of official data on a more exact level but may also be universally reused for any other Slovak, and, after minor adjustments, for any destination in the world.

In the case of Slovakia, Sidor et al. partially proved the distortion of the Statistical Office of the Slovak Republic's (SOotSR) distributed information about the accommodation the service industry via open machine-readable data extracted from the Registry of Financial Statements' and Registry of Legal Entities' application programming interfaces (APIs) [1]. Following up Sidor et al.'s earlier work, the main aim of this study was to extract and examine the volume and basic attributes of accommodation service providers from the six above mentioned online resources, cross check the results, and estimate the distortion of officially disseminated information by the SOotSR at the district (LAU1) level [1]. The case's results indicate that all of Kosice city's districts contain more open accommodation facilities than official statistics claim. Furthermore, the lodging possibilities offered by private citizens have the largest share as potential reporting units. From the perspective of novelty, the results may instantly be used by city administrators in their efforts to decrease local occupancy tax evasions by easily identifying additional accommodation service providers operating outside the scope of administrators' registries. The main novelty of this study lies in the reusability of the approach by any European municipality with structured results and ultimately by any municipality in the world.

The selected data resources have been in the scope of recent open research focusing on tourism and destination management, but, except a few, the focus has mainly been on gaining and examining user generated content (UGC) for management and business intelligence purposes. Among the most relevant efforts in the domain may be included Silva's et al. ad-hoc routines for Booking and TripAdvisor web data extraction and individual case studies on Airbnb's impact on local housing markets carried out by, e.g., Brauckman, Lima, and Cambell et al. [2–5].

2. Literature Review

Dolan et al. used the NCapture third party application to extract Qantas airline's UGC at Facebook (FB) for identifying, categorizing and examining tourists' complaining practices' potential for value co-creation and co-destruction [6]. Kalinić and Vujičić extracted and analyzed the regional disparities of the above data of only officially registered hotels' audiences' demographics and engagement with FB fan pages' UGC by using the Netvizz third party application [7]. Martinez et al., within an importance–performance analysis of marketing and social media at the most popular Spanish and Italian ski resorts, used non-parametric tests and a cluster analysis of FB pages' UGC related to audiences' engagement extracted via the Lykealyzer web application [8]. Lalicic et al. applied sentimental analysis techniques to the above-self-extracted data covering 6000 FB posts and audiences' feedback from 10 of the most popular destinations for the evaluation of communicated content and UGC as emotional brands [9]. Mariani et al., based on their earlier developed tool for data extraction, preprocessing and visual analytics of Facebook's Graph API, analyzed top 10 of most

visited international destinations' National Tourism Organizations' pages' UGC from the perspective of engagement [10,11]. Wozniak et al. analyzed 150 DMOs' social media returns of investment in Belgium, France, and Switzerland, and they used the Fanpage Karma web application for FB pages' data extraction as outcome variables [12]. Kesorn et al., within their PTIS framework, used the Facebook Graph API extracts covering UGC on check-in data at points of interest (POI) and their feedback for personalized recommendations on visiting POIs in destinations [13].

Silva et al., through ad-hoc routines for web data extraction over aggregations of point-based grids, used Booking and TripAdvisor big data in complement to Eurostat data to identify spatiotemporal patterns in hospitality in the whole of the European Union [2]. Martín et al. used the Python-based Scrapy framework to extract user feedback (comments) from Booking and TripAdvisor and implemented the cleaned data within their sentiment estimators based on conventional neural networks and long short-term memory networks [14]. Martín-Fuentes et al. created a Python-based web browser bot to extract accommodation service providers' rankings from Booking and TripAdvisor for the analysis of about 20,000 providers' (present at both platforms) customer feedbacks' similarity and veracity [15]. Yung-Chun et al., within their framework of data crawlers and visual analytics, used machine learning techniques for sentiment-sensitive tree construction, convolution tree kernel classification, aspect extraction, and category detection of The Hilton hotel brand's UGC at TripAdvisor for supporting decision makers via business intelligence [16]. Sumaronso et al. analyzed TripAdvisor's impact on the actual occupancy in Surakarta (Indonesia) via manual searches and interviews with accommodation facilities' managers [17].

Malazizi et al. analyzed the perception of psychological risks by 221 Airbnb hosts in Northern Cyprus [18]. Brauckmann analyzed possible collisions between the "sharing economy" in hospitality and urban property markets via spatial analytics of Airbnb data and official statistics in Germany [3]. Lima used the Inside Airbnb web application to analyze Airbnb's regional impact on the housing crisis in the Greater Dublin Area [4]. Cambell et al. also analyzed the impacts of Airbnb on local housing markets in New Zealand using data extracted via the combination of Python-based web scrappers and PostgreSQL data integration [5]. Zhang and Chen used the Inside Airbnb web application to obtain Airbnb data covering Chicago, Los Angeles and New York, and then implemented spatial regression models to identify selected social-economic and destination "attractiveness" variables' impact on Airbnb listings' presence and efficiency [19]. Similarly, Deboosere et al. analyzed self-scraped Airbnb data via hedonic regressions to study impacts on price range per night and monthly revenues [20]. Fierro and Aranburu used the Datahippo web application's Airbnb data on Bilbao (Spain) to analyze the relationships between cultural heritage concentration and Airbnb listings [21].

Vu et al. used a combination of Foursquare Venue API's check-in data and Twitter UGC in their exploratory analysis of foreign visitors' activity in Hong Kong [22]. Chen et al. used Foursquare UGC on check-ins for an estimation of users' temporal online behavior via equal-frequency discretization [23]. Aliandu's case in Kupang (Indonesia) focused on a sentimental analysis of hospitality and retail based on a Naïve Bayes classifier probability of Foursquare UGC covering check-ins [24]. Chorley et al. analyzed users' personalities from the perspective of openness, conscientiousness, extraversion, agreeableness, and neuroticism on a sample of 173 globally active Foursquare users' check-ins within relation to visited POI attributes [25].

Munawir et al., within their elaboration of new branding strategies for Bandung's (Indonesia) theme parks, focused on Google Places API's ratings' sentiment analysis [26]. The Dutch case by de Vries et al. developed a Google Places API-based web application called HotSpotMonitor to identify national landscape hot spots [27].

From an application perspective, using third party services like Ncapture, Netvizz, LikeAlyzer, and Karma may be useful for data extraction from platforms without an API or for researchers without basic coding skills; however, the necessity to install them may be discouraging. Ncapture's possible depreciation, LikeAlyzer's and Karma's limited free of charge use, may be also discouraging for strategic use. The Inside Airbnb platform currently offers data covering global cities, and the closest

available city to Slovakia is Prague, Czechia. The same goes for the Dathippo platform's scope focusing on Spain, Portugal, and Andorra. From a subjective point of view, all self-developed data extraction approaches may be useful, not only as inspiration—some of them may also be useful as modifiable and ready to use tools for upscaling the current aims' scope.

3. Materials, Data Extraction, Cleaning and Preprocessing

Willis's comprehensive analysis of the ethical boundaries of online data analysis focused partially on whether Facebook user generated content may be considered as human research or as documentary research, specifically on the necessity of secondary consent for UGC generated by consented users' connections' (friends) [28]. Within the scope of this paper, all data were subjectively considered documentary without the necessity of informed consent. Firstly, data extracted by open APIs (Facebook, Foursquare, Google) are public. Secondly, even though Airbnb, Booking and TripAdvisor do not have an open API, only publicly available data were extracted, which could have also been done manually. Thirdly, the paper is a proof a concept, with the aim of supporting local and national authorities within their efforts of decreasing possible tax evasions. Fourthly, subjectively, all major platforms should have interest on their clients' compliance with local regulations and law. Most importantly, the paper does not point out any private information and does not accuse anyone of tax evasion; it only points out official statistics' scope's limits.

Kosice has a population of over 239,000 and is Slovakia's second largest city; due to its location (Figure 1), it is also considered the metropolitan of the eastern part of the country. According to the SOotSR's official data, the city had the 4th largest number of overnight stays among the top ten most visited cities and local tourism regions in Slovakia in 2018 [29]. The city itself arises from four districts (LAU1), which together arise from 21 city parts (LAU2).

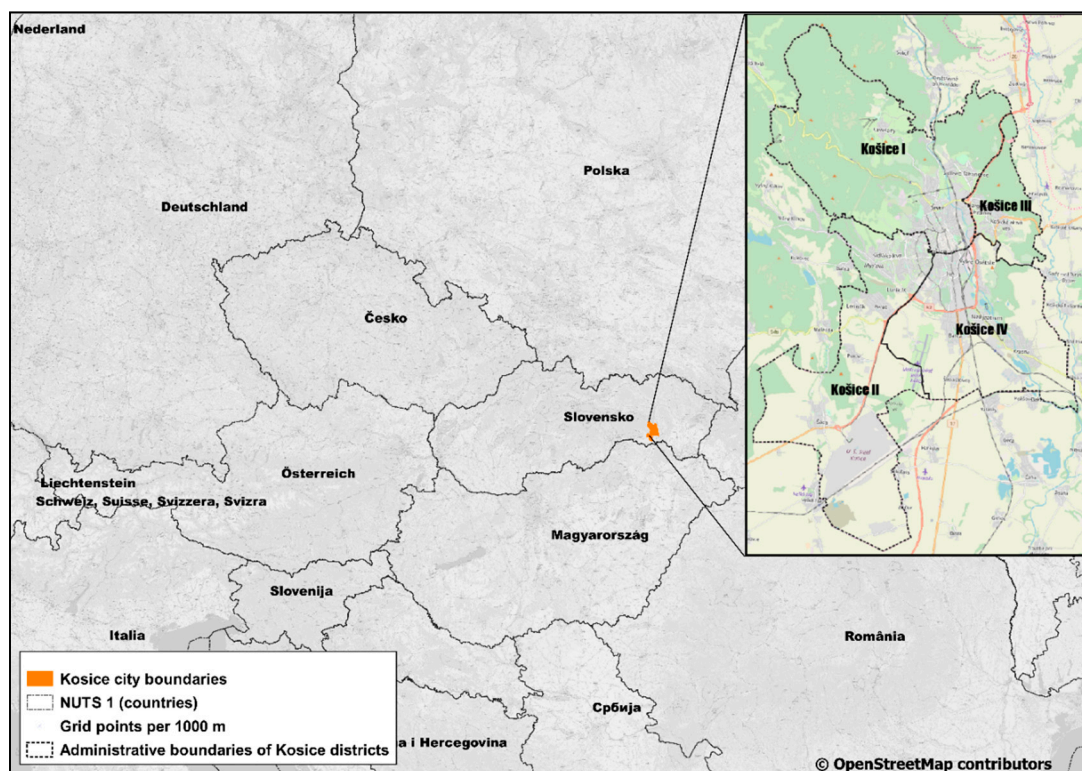


Figure 1. Localization of Kosice in Central Europe with a detail zoom on its districts.

Based on the decree of Hungarian king Béla IV, Kosice has been a city since the 13th century. With almost 800 national monuments, the city is the largest urban conservation area of Slovakia [30,31]. The city has a rich history and has recently received international recognition by becoming the

designated European Capital of Culture (2013), European City of Sport (2016), UNESCO Creative City of Media Arts (2017) and European Volunteering Centre (2019) [32–35]. Even though the number of visiting tourists has grown annually, Kosice with 21.4% of annual occupancy of permanent bed places has not yet reached its full potential as a tourist destination [29].

In terms of the article's scope, Kosice is a medium sized city but at the same time arises from 21 city parts (reporting units); thus, it is a suitable administrative unit for aggregated tests. This way, the extraction scripts may be reused as a proof of concept in both smaller municipalities (LAU2), as well as cities that arise from multiple LAU2s. The suitability of Kosice has been strengthened by a simple analysis of Eurostat's dataset on cities' culture tourism in 2016 (richest records) [36]. The dataset, among others, covered 565 cities with identified capacities (bed places) and 555 cities' basic tourism efficiencies (overnights stays), with an average 7843 bed places per city (BPpC) and 1,086,210 overnight stays per city (OSpC) [37]. Kosice, with 5430 BPpC and 350,145 OSpC, was under average in terms of both metrics, but so were 437 (77%) cities in the case of BPpC and 435 (78%) cities in the case of OSpC [37]. Subjectively, these "under EU (European Union) average" cities could benefit from Kosice's example as a motivation to back check local accommodation services industries.

The lowest administrative unit with data availability for the accommodation services industry covering the entire country of the SOotSR's open database is the district. For this reason, all cleaned results were aggregated at the district level.

From the perspective of data accessibility, the six data resources had to be divided into two groups. The first group's sources (Facebook, Google Places, and Foursquare) have an open representational state transfer (REST) application programming interface. The other three resources either do not provide access to their APIs for pure scientific purposes (Booking and TripAdvisor) or do not have a relevant and usable API (Airbnb).

3.1. Data Extraction with Open API Access

All three APIs' automated search for objects is constructed on point-based spatial querying with radius distances and object type identification. For data extraction convenience and reusability for any other Slovak administrative unit, Slovakia's territorial and administrative units' spatial representation was downloaded in a shapefile and uploaded to a local PostgreSQL database with a PostGIS extension [38–40]. For maximizing the advantages of the API's spatial querying, Siddique's PostgreSQL/PostGIS custom function was used for generating a regular point grid with 1000 m distances in between points over Kosice districts' polygon spatial representation (Figure 2) [41]. With this in place, an individual modifiable extraction script in Python was constructed for each API.

The general logic behind the extraction scripts was to query each created point with regards to the API's content's unique identifiers. Since the point grid's buffer radii overlapped, the scripts contained a section for removing duplicate records. Afterwards the data were extracted to a local database for further processing. Each API's usage requires an individual API key that works as an individual user access token, mainly for security and traffic control.

Within the Facebook Places Search API, all objects with the category identifier "HOTEL_LODGING" and their base attributes (name, location including point spatial representation, website name, status of verification, number of check-ins, number of ratings, and number of user whom liked the object at Facebook) were targeted [42,43]. In the case of Google's Places API, objects with the type identifier "lodging" and their base attributes (id of the object, name, street, the object's average rating, number of ratings, point spatial representation, and id of the place to which the object is related) were targeted [44,45]. Within the Foursquare Venue API, all objects with the main category identifier "Hotel" (identifier includes all available forms of accommodation services) and their base attributes (id, name, related location with point spatial representation, subcategories' id and name, main categories' name, and number of pictures) were targeted [46,47].

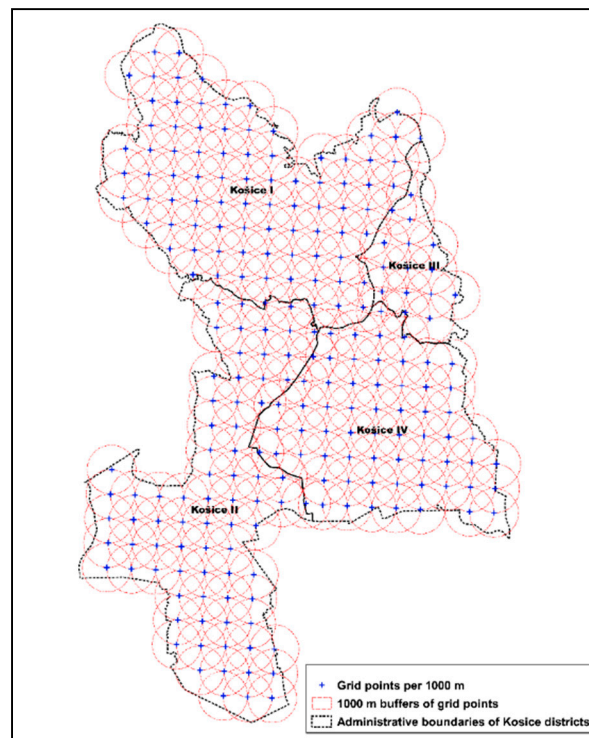


Figure 2. Visualization of the created set of grid points with overlaying buffers over Kosice districts.

3.2. Data Extraction Without an Open Access API

Since Booking, TripAdvisor and Airbnb contain authenticated accommodation service providers at a larger volume but do not provide open API access for scientific purposes, a set of simple Python-based web data scraper scripts were created. Each script took in the uniform resource locator (URL) of the initial user search results at the destination level. Afterwards, the script simulated user clicking, went through each page within the pagination lists and each listing (accommodation service providers' profile page), and then extracted the targeted objects' attributes into a local database.

Within Booking, all objects' listings under the Kosice (Slovakia) destination were searched for and base attributes (name, URL, address, number of reviews, rating, and service start date at Booking) were targeted [48]. In the same fashion TripAdvisor was searched, and base attributes (name, URL, address, number of reviews, and rating) were extracted [49]. In the case of Airbnb, because of overloading the rate limit of user connections, the script had to be separated into two parts: The first extracted each listing's URL [50]; the second part targeted base attributes (name of the "host," city of the service provider, URL, size, number of reviews, rating, and service start date at Airbnb) and had to be scaled into looping over groups of around 100 of listings at one time in a period of two hours [51].

3.3. Data Cleaning, Spatial Representation and Aggregation

The open API-based Booking and TripAdvisor extraction scripts contained sections aimed at duplicate detection and elimination. In the case of Airbnb, cleaning and pre-processing was archived in a separate script [52]. All three sets of web scrapped extracts did not contain listings' spatial coordinates; to find these coordinates, Booking and TripAdvisor extracts' addresses were geocoded via Lyn's Google Maps Geocoding API-based Python bulk geocoder [53–55]. Due to Airbnb's policy of distributing only the accommodation service provider's city (LAU2) before creating a reservation, the extracts were not geocoded to the point of spatial representation. All extracts were then merged with Slovak territorial and administrative units' (LAU2, LAU1, and NUTS3) official numeric identifiers via the PostGIS extension's "ST_Within" function. The cleaned records were then grouped into accommodation facility categories based on the simple text mining of key words within the service provider's name in a

PostgreSQL environment. The next step was crosschecking the record's occurrence in multiple sources via the service providers' name equality respectively similarity in a PostgreSQL environment.

4. Obtained Results' Basic Interpretation

The results were compared with official statistics via spatial interpretation on an OpenStreetMap base map with official territorial borders within the QGIS software's environment; they were graphically using Bostock's open D3 JavaScript library-based radial space-filling and radial stacked visualization methods [29,56–62].

4.1. Facebook Place Search API

Due to duplicate registrations of places (Facebook business or fan pages) related to accommodation service providers (ASPs) and the incorrect usage of the "HOTEL_LODGING," the data extract was additionally cleaned of twelve duplicate registrations, and four service providers not providing lodging [61]. Overall, 47 ASPs within the four districts were found, most of them situated within the city center and the Kosice I district (Figure 3).

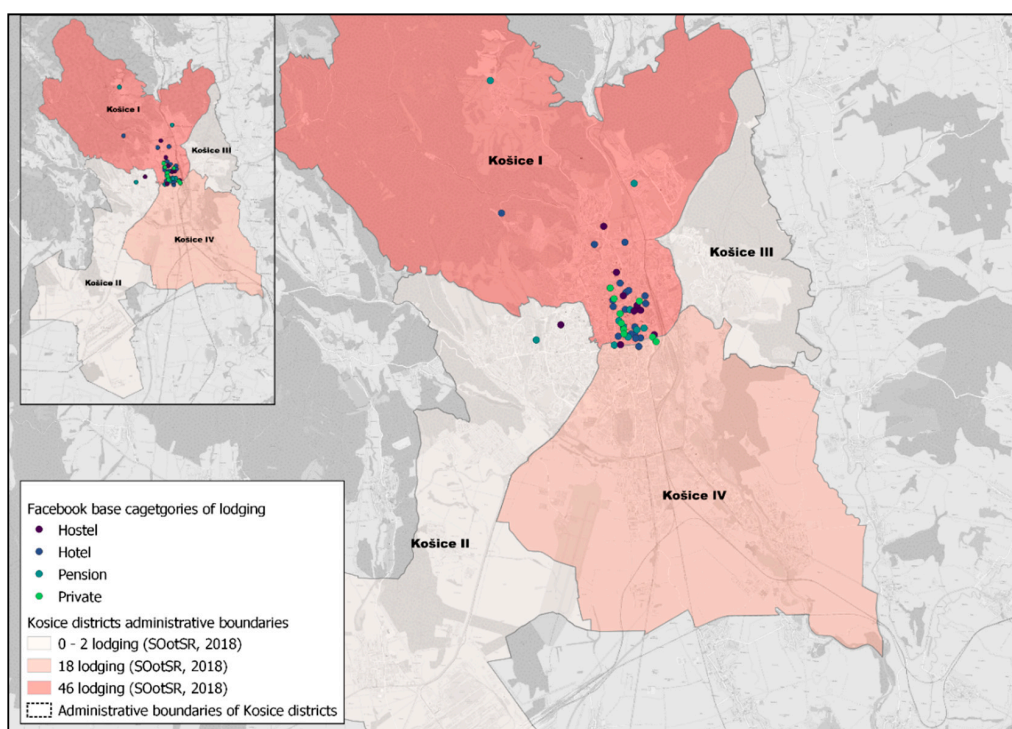


Figure 3. Spatial visualization of the Facebook Place Search application programming (API) data on accommodation providers vs. official number of accommodation providers within Kosice districts.

In comparison to the official SOotSR data at the district level (minimum 70 ASPs), the Facebook Place Search API extract contained overall fewer records (47 ASPs) and individually contained fewer records than the SOotSR in three city districts (Kosice I, III and IV) [29]. In comparison to the SOotSR's other database on ASPs in selected Slovak cities and tourism regions covering Kosice (84 ASPs), FB's volume hypothetically covered only 55.95% of it. The reasons for the fewer records in FB vary, including using incorrect identifiers, not using Facebook as a communication channel, and using Facebook user accounts for the communication of ASP offering instead of business or fan pages. In terms of accommodation facilities' categorization (Figure 4) obtained via ASPs' names' text mining, it should be mentioned that all major hotels and well-known pensions and hostels in the city center are registered and active.

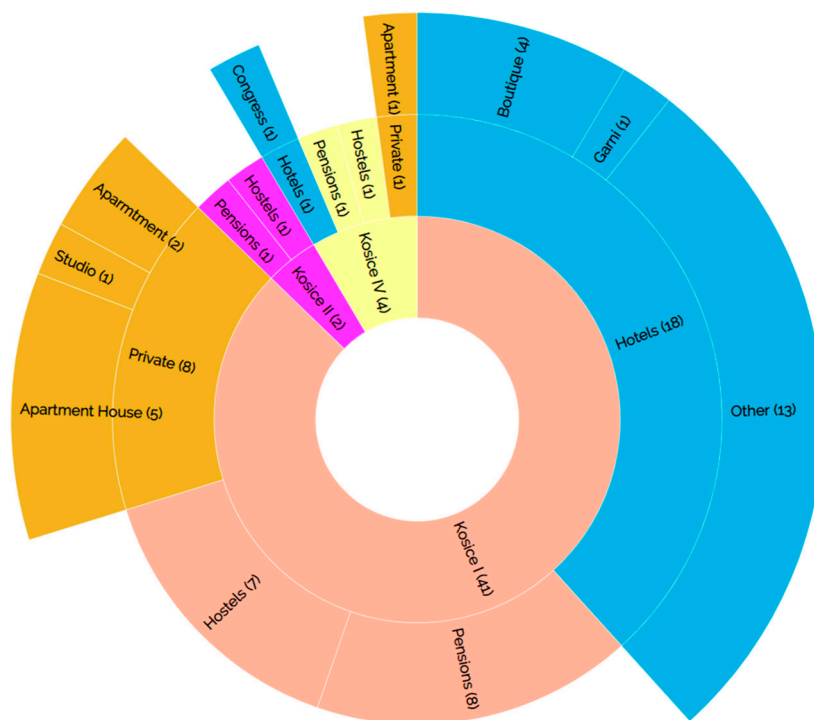


Figure 4. Type and share of accommodation service providers in Kosice districts according to extracted cleaned the Facebook Place Search API data (for full resolution, please click the source) [63].

In terms of user interaction, with an average of 1169 check-ins per ASP (median 69 check-ins), only eight records had zero check-ins. Subjectively, the number of check-ins suggest that 83% of records are actual facilities which have been visited by tourists. From the perspective of public recognition, with an average of 741 user likes per record (median 158), only three records had zero fans and zero check-ins but were promoted by other platforms at the same time. This fact partially proves that the pages actively communicated their content.

4.2. Google Places API

Due to duplicate registrations or incorrect usage of the “lodging” attribute’s identifier 31 records were discarded [64]. Most of irrelevant records were associated with pubs, cafes, and restaurants, but even homeless shelters and the Government Office of the Slovak Republic’s facilities were included.

Overall, 112 ASPs were identified within city boundaries. More than 50% of the ASPs were situated (Figure 5) in the city center (LAU2 Kosice–Old town). In comparison to the SOotSR data, the Google Places API (GP) covered 1.6× more records the SOotSR’s district level database and 1.3× more than the SOotSR’s city level database. Individually, in the Kosice I district, GP contains 1.5× more records, and in the Kosice II district (SOotSR applies its confidentiality rule), the GP contains 9.5× more records than the SOotSR. The Kosice III district’s (one ASP) GP records are hypothetically the same as the SOotSR, and in the Kosice IV district, GP contains one more record than the SOotSR.

In terms user interaction, only 17 records (15%) had no ratings but were active on other platforms. Overall, the records had an average number of 108 ratings per facility (median 32 number of ratings). From the user perspective, (customer) the average rating per unit was 3.4 out of 5 points (median 4.1 points). Nevertheless, it may subjectively be claimed that 85% of identified facilities were rated by users who were potentially actual customers.

In terms of accommodation facilities’ categorization, there were more private home offerings than commercial ASPs within the city center. Most of the private offers appear to be regular flats (13 records), but many them (11 records) appear to be apartment houses with more than one private room (Figure 6). Again, all major commercial hotels and well-known ASPs were identified within GP records.

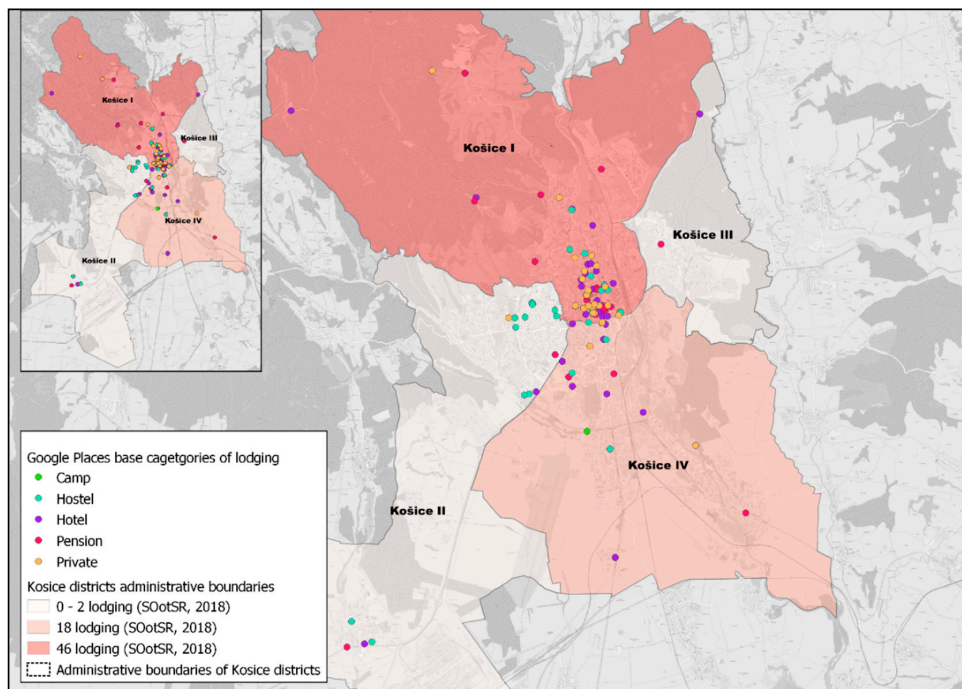


Figure 5. Spatial visualization of the Google Places API data on accommodation providers vs. official number of accommodation providers within Kosice districts.

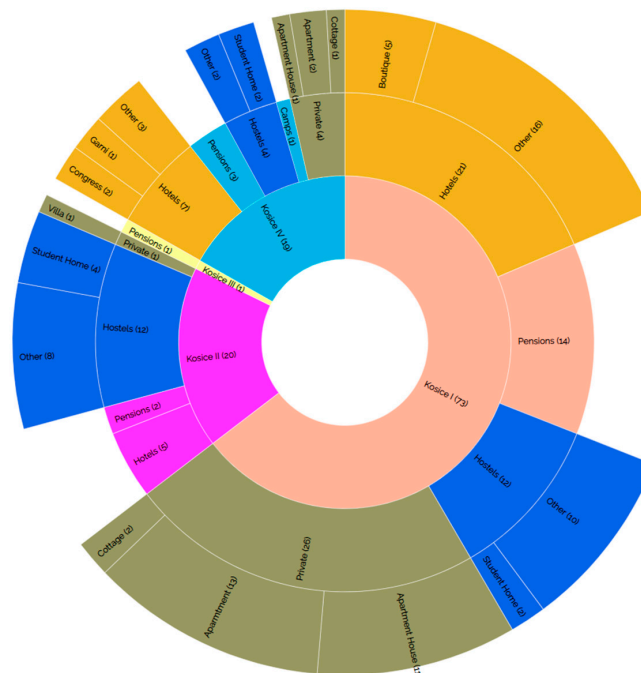


Figure 6. Type and share of accommodation service providers in Kosice districts according to extracted cleaned the Google Places API data (for full resolution, please click the source) [65].

4.3. Foursquare Venue API

More than half of the ASPs (23 records) were concentrated (Figure 7) within the city center (LAU2, Kosice–Old Town). Some medium sized concentrations were situated in the southern (seven ASPs, LAU2, Kosice–South) and northern (six ASPs, LAU2, Kosice–North) parts of the wider city center.

Even though the Foursquare Venue API extract returned only 42 records, only one of them was outside the spatial boundary of Kosice city. In comparison to the SOotSR’s data, except for the Kosice

II district (4 ASPs), Foursquare contained fewer records. In terms of accommodation facility categories, most of commercial hotels are registered within Foursquare as user-reviewed points of interest. The low volume of private home offers in Foursquare may be impacted the application’s user driven nature and main focus on local must-see points of interest (Figure 8). Globally, the small volume of records may have various reasons, but it is probably due to the small traffic share of Slovakia (0.14%) among Foursquare users [66].

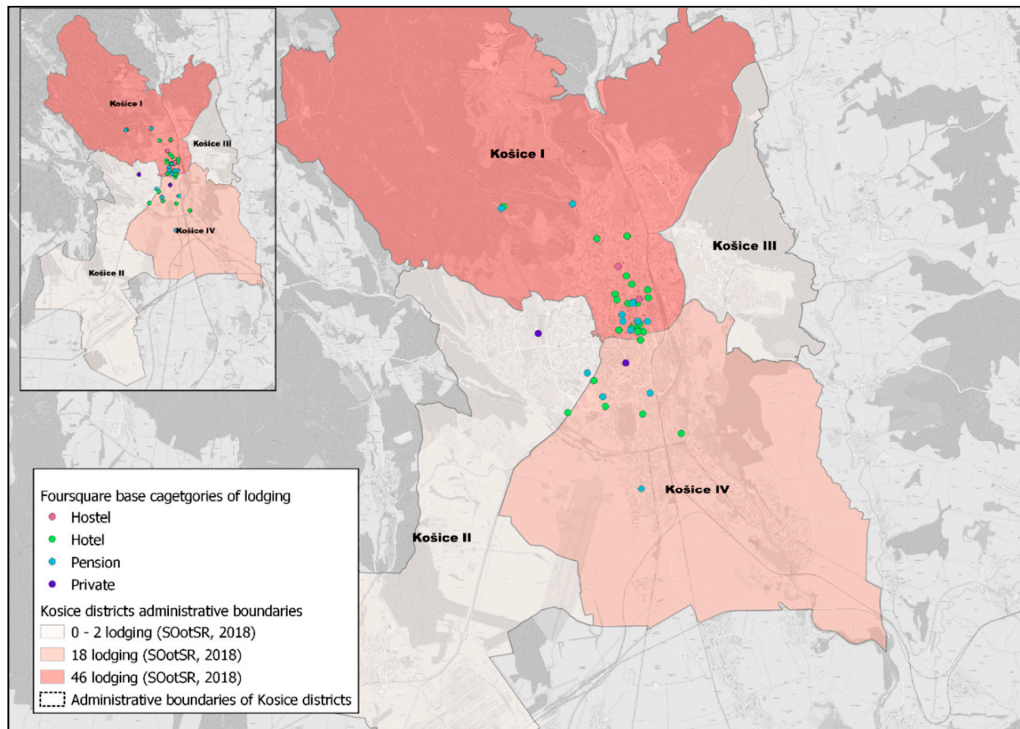


Figure 7. Spatial visualization of the Foursquare Venue API data on accommodation providers vs. official number of accommodation providers within Kosice districts.

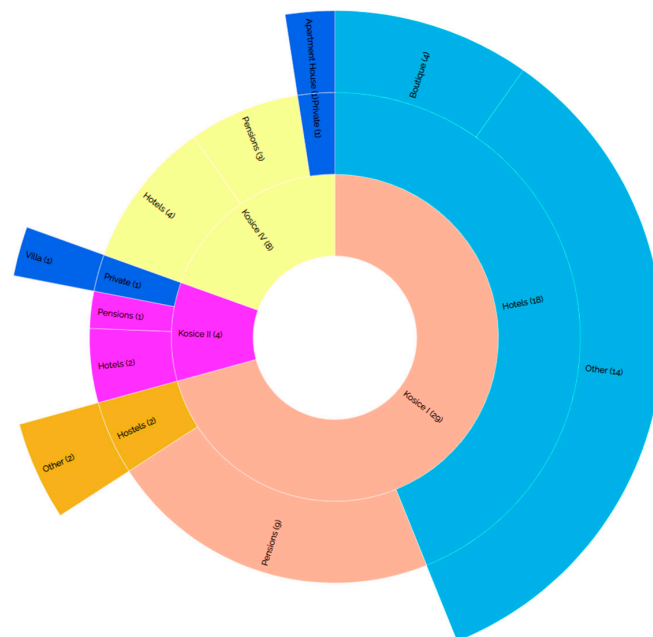


Figure 8. Type and share of accommodation service providers in Kosice districts according to extracted Foursquare Venue API data (for full resolution, please click the source) [67].

4.4. Booking

From 246 records extracted, three were outside of the area of Kosice city, and five duplicates (individually communicated theme-designed rooms of already included ASPs) were excluded. More than 93% of identified the ASPs were located (Figure 9) in the city center (143 ASPs, LAU2 Kosice–Old Town) and its southern (39 ASPs in LAU2, Kosice–South) northern neighborhoods (28 ASPs, LAU2, Kosice–North). It is worth mentioning the concentration of smaller groups of ASPs in residential city parts (Kosice II and III districts) outside of places with tourist attractions.

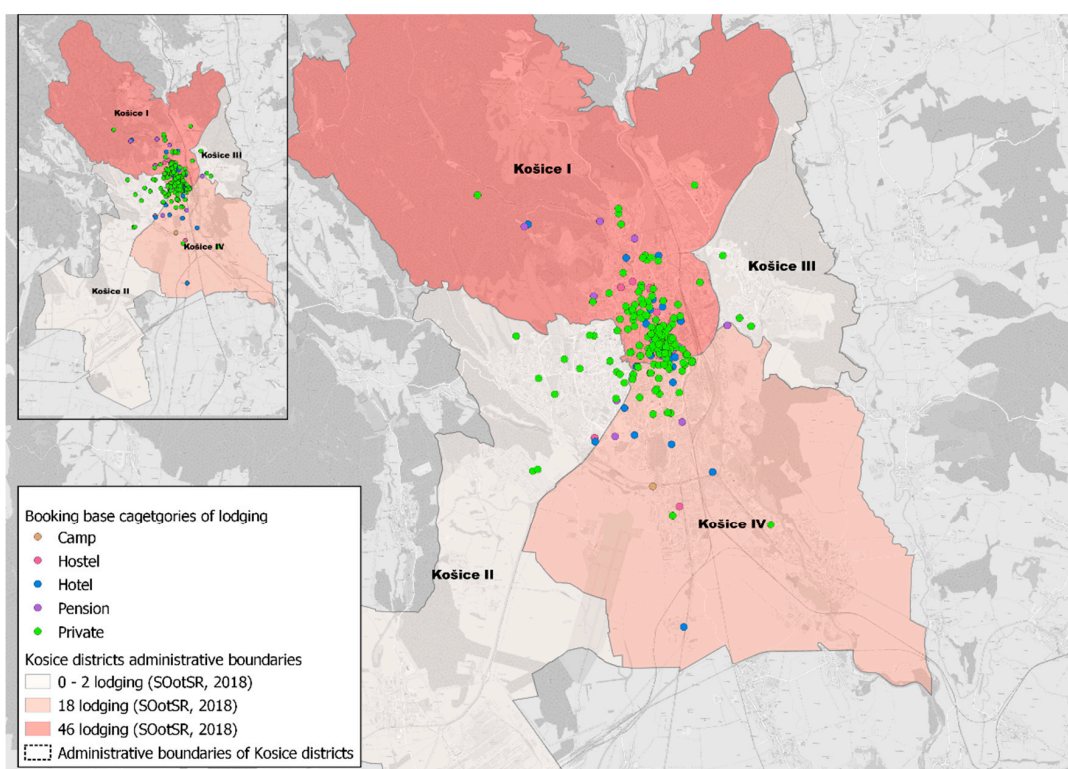


Figure 9. Spatial visualization of Booking data on accommodation providers vs. official number of accommodation providers within Kosice districts.

Booking's user interface contains information on ASPs' service start date, and for this reason, a comparison with the SOotSR data was carried out but only with records active before the 1st of January, 2019, even if the ASPs could have had been active service providers without active operations at Booking. Though 121 records were excluded, Booking still contained a larger volume of ASPs in the Kosice I (1.9× more ASPs), II (at least 5.5× more ASPs) and IV (1.2× more ASPs) districts.

The rise of the ASPs' volume at Booking to almost double (Kosice I had 84 ASPs, Kosice II had 12 ASPs, Kosice III had four ASPs, and Kosice IV had 21 ASPs in 2019) in comparison to 2018 has a simple hypothetic reason. In May 2019, Kosice hosted the IIHF 2019 Ice Hockey Championship. In 2019, private home offers had risen by 116 ASPs (95.87% share of total growth) and by 93 ASPs (76.86% share of total growth) between January 2019 and 9th May 2019 (day before the start of the IIHF Championship).

The above mentioned growth has had impact on the overall share of accommodation facility categories (Figure 10). Private home offers' overall share is 77.86%, and their share is above 70% in every city district. All major commercial hotels were identified with Booking's database, and even three student homes sell their capacity via Booking's channel.



Figure 10. Type and share of accommodation service providers in Kosice districts according to extracted Booking data (for full resolution, please click the source) [68].

In terms of user interaction, 27.5% of records did not contain reviews. Overall, an average of 35 reviews per facility (median 18) was identified, and, in terms of customer satisfaction (reviews at Booking may be conducted only after check-out), an average of 6.5 out of 10 points per facility was identified (median 8.7 points).

4.5. TripAdvisor

From 85 extracted records, two were excluded as irrelevant (one case of an incorrect registration of a municipality center and one case of a facility that was never opened), and four ASPs were outside of Kosice's boundaries. From a spatial distribution perspective (Figure 11), 43% of identified ASPs were located within the city center (34 ASPs, LAU2 Kosice–Old Town), 21% were located in the southern wider city center (17 ASPs, LAU2 Kosice–South), and 15% were located in the northern wider city center (12 ASPs, LAU2 Kosice–North).

In comparison to the SOotSR data, TripAdvisor covered almost the same volume (Kosice I had two ASPs and Kosice IV had two ASPs). A larger disproportion was observed only in the Kosice II district (at least four-and-a-half more ASPs in favor of TripAdvisor). The smaller volume of records may have been (subjective view) caused by the fact that TripAdvisor is primarily a channel used for sharing user experiences in destination and only secondarily as an intermediary sales broker for other reservation services, e.g., Booking or Hotels.com.

Private home offers (39.5%) have a larger share only in the Kosice I district (Figure 12). Similarly, to Booking, some student homes were identified as commercial ASPs in the Kosice I district. Most commercial hotels and well-known accommodation facilities may be identified within TripAdvisor.

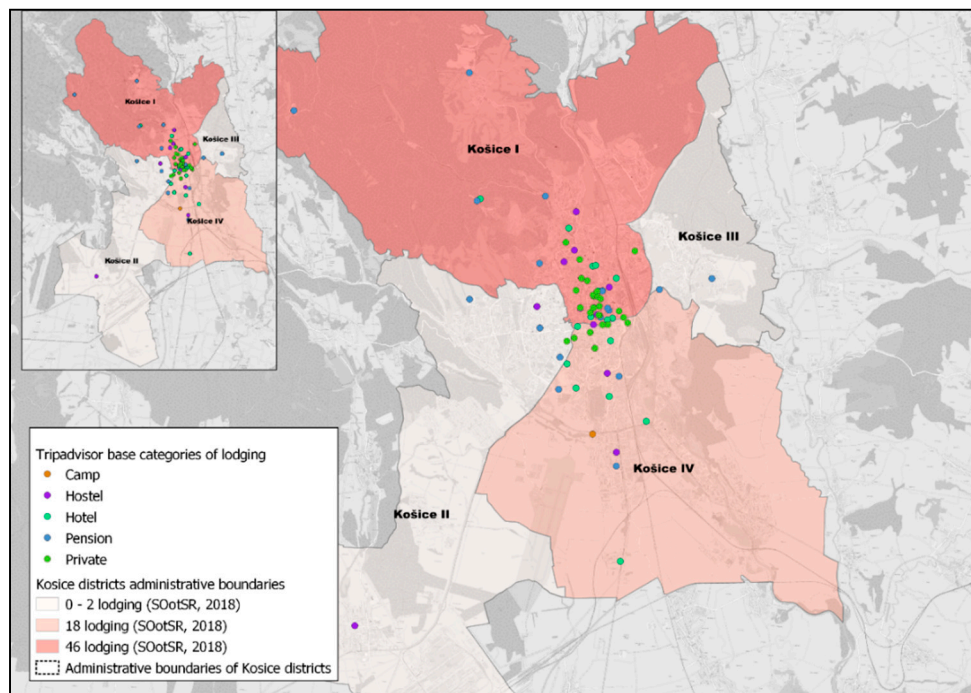


Figure 11. Spatial visualization of TripAdvisor data on accommodation providers vs. official number of accommodation providers within Kosice districts.



Figure 12. Type and share of accommodation service providers in Kosice districts according to extracted TripAdvisor data (for full resolution, please click the source) [69].

In terms of user interaction, 34 records (43%) did not contain any reviews but may have been identified at other platforms. Overall, the records reached an average of 18 reviews per facility (median 2). From the satisfaction perspective, an average rating of 2.2 point per facility (median 3.0) was identified.

4.6. Airbnb

Out of 305 returned records, 29 were excluded due to territorial irrelevance. As already mentioned, Airbnb provides the exact point-based referenceable address only after creating a reservation. For this reason, the extracted results were aggregated by available city parts' name (LAU2). Since many Airbnb records contained the city identifier only as "Kosice" instead of the exact city part's name, 178 ASPs had to be grouped under the spatial unit of the entire city (four districts). Most of the 82 geographic point referenceable ASPs were situated in the city center and its southern and northern neighborhoods, but ASPs were even active in more remote residential city parts within from prefabricated apartment buildings from the 1970s and 1980s (Figure 13).

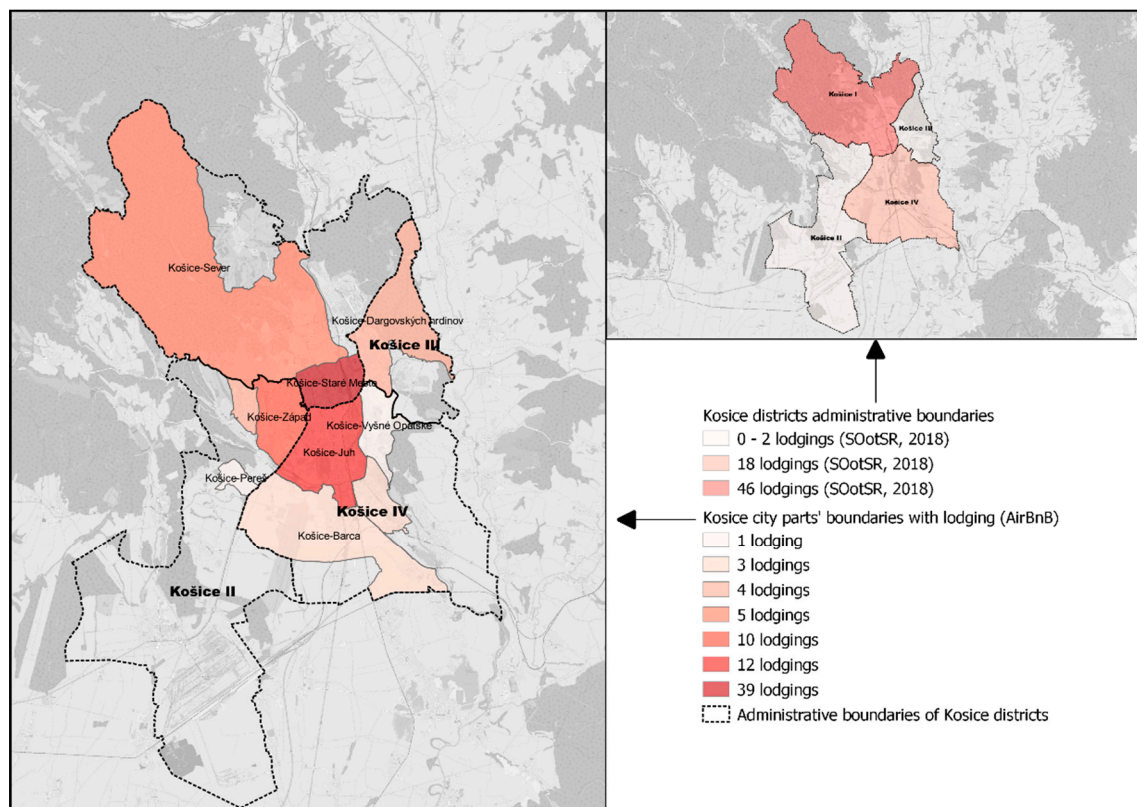


Figure 13. Spatial visualization of Airbnb on accommodation providers with the municipality identifier vs. official number of accommodation providers within Kosice districts.

Due to Airbnb's partial "sharing economy" nature, a comparison with the SOotSR data is arguably relevant. However, under section a), § 2 of Decree no. 277/2008 of the Ministry of Economy of the Slovak Republic, which lays down the classification of accommodation facilities when classifying them into categories and classes (Decree 277/2008), an accommodation facility is a building, space or area in which the public is provided with temporary accommodation and related services all year round for a fee [70]. An accommodation facility is also a seasonal facility that provides accommodation and related services for a maximum of nine months of the year. Another considerable fact is underlined by Section 18 §4 of Decree 277/2008, which states that flats or apartments also fall under the category of private accommodation [70]. A final considerable fact is that Decree 277/2008 was created for Act no. 455/199 Coll. in regards to self-employed businesses;—the act states that, among other taxation obligations, every ASP is within the scope of occupancy and income tax in Slovakia [71].

The Airbnb records contain the ASPs’ service start month and year. At the end of the year 2017, 136 ASPs registered and at the end of year 2018, 184 ASPs (almost double of the SOotSR records) were registered by the platform. While between January and May 2018, the number of ASPs rose by 14 ASPs, between January and May 2019, the number of new ASPs rose by 88. The 26% annual growth between 2017 and 2018 cannot be directly connected to the 2019 IIHF Championship, but the annual difference of growth between January and May may be connected to the higher demand of short-term accommodation within the city. What is important that new ASPs have not yet suspended their Airbnb membership.

In terms of facilities’ categories, 70% of APSs offer individual apartments (Figure 14). Within the ASPs’ names’, the text mining 14% of records not containing any of the targeted key words identifying their category had to be labelled “other,” but they were most likely individual apartments [72]. Only 7% of ASPs’ offering a single room fall theoretically under the “sharing economy” concept.



Figure 14. Type and share of accommodation service providers in Kosice districts according to extracted Airbnb data (for full resolution, please click the source) [73].

In terms of user interaction, 52% of facilities were not reviewed, and 42% of facilities did not have a rating. Overall, the records reached an average of 16 reviews per facility (median 0) and an average rating of 2.8 out of 5 points (median 4.5 points).

4.7. Crosschecked Results

After cross referencing the names of accommodation service providers through data extracts, it was found Airbnb had the largest individual occurrence (Figure 15) of ASPs, almost double that of Booking individual occurrences and seven times more than individual Google occurrences [74]. Airbnb holds more individual occurrences than the SOotSR data on Kosice districts and the individual city data combined.

Larger double occurrences were observed in the Airbnb–Booking (18 ASPs), Booking–TripAdvisor (17 ASPs), Google–TripAdvisor (10 ASPs) and Booking–Google (nine ASPs) combinations. Subjectively, Airbnb–Booking combinations partially prove that at least 18 Airbnb hosts run a regular tax-obligated business operation. Smaller triple occurrences were observed in the Booking–Google–TripAdvisor combination (nine ASPs). Quadruple occurrences were observed in case of hotels and higher standard pensions within the of Booking–Facebook–Foursquare–Google (11 ASPs) and Booking–Foursquare–Google–TripAdvisor (13 ASPs) combinations. A quintuple occurrence was identified only in the case of well-known hotels situated in the city center (LAU2 Kosice–Old Town) within the Booking–Facebook–Foursquare–Google–TripAdvisor combination (8 ASPs).

As can be seen (Figure 15), the SOotSR held fewer records combined for the city and its districts than individual occurrences at Airbnb (SOotSR missing from 171 reporting units at the city level and 187 at the district level) and Booking (SOotSR missing from 50 reporting units at the city level and 66 at the district level) alone. While the individual analysis of Google Places, Booking and Airbnb partially proved the distortion of the number of ASPs recorded by the SOotSR, if we take only into account individual occurrences at Booking and Airbnb (classic reservation systems) and double and multiple occurrences including Booking and Airbnb; the SOotSR data is missing from 404 reporting units at the city level and 420 reporting units at the district level.

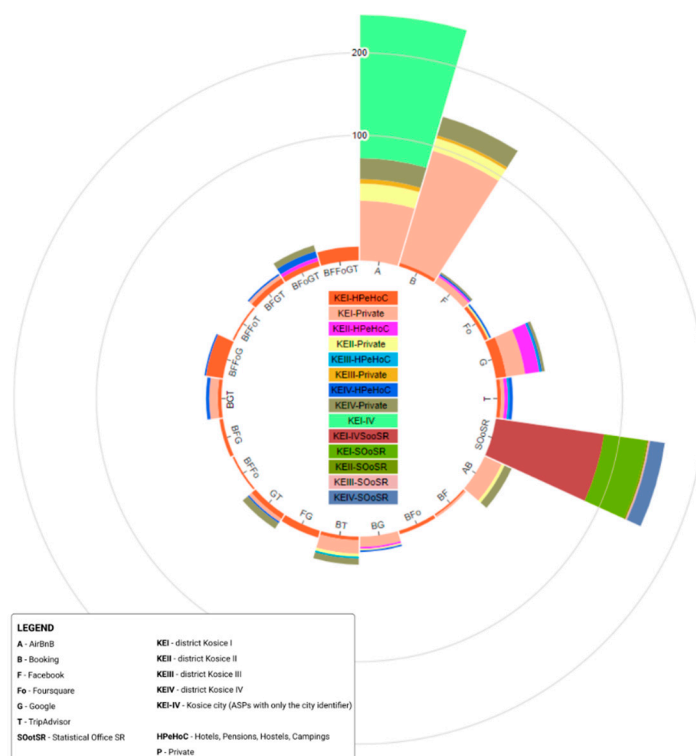


Figure 15. Radial scale of accommodation facilities’ multiplicity cross platform occurrence (for full resolution, please click the source) [75].

5. Conclusions and Future Work

While Sidor et al.’s earlier work proved an error in the system, the current results make the error more tangible and clear [1]. Sidor et al.’s earlier work focused on identifying accommodation service providers within the records of the Registry of Financial Statements (RUZ) and Registry of Legal Entities (RPO); they calculated a 152.24% ratio (RUZ) and a 4400% ratio (RPO) against the SOotSR records; without analyzing whether the entities actually provided accommodation services. However, the current results take into account actual online promoted services. Taking the obtained results into account, the subjective answer to the question within the article’s title is unfortunately a yes, but also a

no. In the case of Kosice city, the Facebook Place Search API is not very suitable; however, reverse querying based on the names of ASPs in official municipality or governmental registries could be beneficial for harvesting user feedback. The same may be said about Foursquare and TripAdvisor, although due to their user driven content, all ASPs could be theoretically integrated into both platforms for generating a larger volume of UGC. The Google Places API, Booking and Airbnb may be clearly considered as essential sources for the regular semi-automatic monitoring of new ASPs occurrence. It should be clearly stated that the authors do not hold these platforms responsible for their “clients” potential unethical actions towards local occupancy taxation. The same goes for the SOotSR, which is probably limited by human and technological resources’ capacities. Though it is true that the line between data protection and protecting the public interest is very thin, the study was subjectively within the boundaries of the public interest. The EU’s General Data Protection Regulation does not apply to legal persons, and the study carried out tasks in the interest of finding a way to decrease occupancy tax evasion [76]. Most importantly, the EU’s Directive on the legal protection of databases states an exception for data extraction and reutilization for the purposes of illustration for teaching and scientific research or processing carried out in the interest of public security [77].

Of course, the current version of extraction scripts and scrappers (as well as extracted data) have their limits, but, from a practical point of view, the extracted results may serve Kosice city’s board as base list for back checking their own registries. The extraction scripts may be easily used for any other Slovak destinations and ultimately any destination in the world. Since the examined data sources are continuously updated and their content continually changed, the used scripts must be maintained.

To exploit the full potential of the analyzed platforms, future scrappers should be automated and extended with content by data on capacity, user feedback (comments), price, and availability. This way, even economic performance could be monitored. This could at least narrow the speculations of some ASPs towards local occupancy taxation. Subjectively, there is no doubt that Booking and Airbnb, in combination with the official SOotSR, could support the sustainability of tourism in terms of additional knowledge. Ultimately, the content of Google Places, the Facebook Place Search API, TripAdvisor and the Foursquare Venue API that cover all types of POIs may be used to draw up the basic nature of any destination or area from a UGC perspective that could help determine the strategic value of tourism compared to other industries or other uses of local resources than for tourism.

Author Contributions: Conceptualization, C.S., B.K. and L.Š.; data curation, C.S.; formal analysis, C.S., B.K., L.Š., M.C., S.K., M.S., J.D., J.G. and B.B.; funding acquisition, B.K., M.C. and S.K.; investigation, C.S., B.K., L.Š., M.C., S.K., M.S., J.D., J.G. and B.B.; methodology, C.S.; project administration, B.K.; resources, C.S., B.K. and L.Š.; software, C.S.; supervision, B.K.; validation, C.S., B.K., L.Š., M.C., S.K., M.S., J.D., J.G. and B.B.; visualization, C.S.; writing—original draft, C.S.; writing—review and editing, C.S., B.K. and L.Š.

Funding: This work was supported by the Slovak Research and Development Agency under contract no. APVV-14-0797; This publication has been supported by the project KEGA No. 036TUKÉ-4/2019. and the project of the Scientific Grant Agency of the Ministry of Education, Science, Research and Sport of the Slovak Republic and the Slovak Academy of Sciences No. VEGA 1/0515/18.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

1. Sidor, C.; Kršák, B.; Štrba, L. Identification of distorted official hospitality statistics’ and their impact on DMOs’ funding’s sustainability: Case notes from Slovakia. *Sustainability* **2019**, *11*, 1084. [[CrossRef](#)]
2. Silva, F.B.E.; Herrera, M.A.M.; Rosina, K.; Barranco, R.R.; Freire, S.; Schiavina, M. Analysing spatiotemporal patterns of tourism in Europe at high-resolution with conventional and big data sources. *Tour. Manag.* **2018**, *68*, 101–115. [[CrossRef](#)]
3. Brauckmann, S. City tourism and the sharing economy—Potential effects of online peer-to-peer marketplaces on urban property markets. *J. Tour. Futures* **2017**, *3*, 114–126. [[CrossRef](#)]

4. Lima, V. Towards an understanding of the regional impact of Airbnb in Ireland. *Reg. Stud. Reg. Sci.* **2019**, *6*, 78–91. [[CrossRef](#)]
5. Campbell, M.; McNair, H.; Mackay, M.; Perkins, H.C. Disrupting the regional housing market: Airbnb in New Zealand. *Reg. Stud. Reg. Sci.* **2019**, *6*, 139–142. [[CrossRef](#)]
6. Dolan, R.; Seo, Y.; Kemper, J. Complaining practices on social media in tourism: A value co-creation and co-destruction perspective. *Tour. Manag.* **2019**, *73*, 35–45. [[CrossRef](#)]
7. Kalinić, Č.; Vujčić, M. A subnational assessment of hotel social media metrics—The case of Serbia. *Geogr. Pannonica* **2019**, *23*, 87–101. [[CrossRef](#)]
8. Martínez, T.L.; Toledo, L.D.; Faraoni, N. Auditing marketing and the use of social media at ski resorts. *Sustainability* **2019**, *11*, 2868. [[CrossRef](#)]
9. Lalicic, L.; Huertas, A.; Moreno, A.; Jabreel, M. Which emotional brand values do my followers want to hear about? An investigation of popular European tourist destinations. *Inf. Technol. Tour.* **2019**, *12*, 63–81. [[CrossRef](#)]
10. Mariani, M.M.; Felice, M.D.; Mura, M. Facebook as a destination marketing tool: Evidence from Italian regional Destination Management Organizations. *Tour. Manag.* **2016**, *54*, 321–343. [[CrossRef](#)]
11. Mariani, M.M.; Mura, M.; Felice, M.D. The determinants of Facebook social engagement for national tourism organizations' Facebook pages: A quantitative approach. *J. Destin. Mark. Manag.* **2018**, *8*, 312–325. [[CrossRef](#)]
12. Wozniak, T.; Stangl, B.; Schegg, R.; Liebrich, A. The return on tourism organizations' social media investments: Preliminary evidence from Belgium, France, and Switzerland. *Inf. Technol. Tour.* **2017**, *17*, 75–100. [[CrossRef](#)]
13. Kesorn, K.; Juraphanthong, W.; Salaiwarakul, A. Personalized attraction recommendation system for tourists through check-in data. *IEEE Access* **2017**, *5*, 26703–26721. [[CrossRef](#)]
14. Martin, C.A.; Torres, J.M.; Aguilar, R.M.; Diaz, S. Using deep learning to predict sentiments: Case study in tourism. *Complexity* **2018**, *2018*, 7408431. [[CrossRef](#)]
15. Martin-Fuentes, E.; Mateu, C.; Fernandez, C. Does verifying uses influence rankings? Analyzing Booking.com and TripAdvisor. *Tour. Anal.* **2018**, *23*, 1–15. [[CrossRef](#)]
16. Yung-Chun, C.; Chih-Hao, K.; Chun-Hung, C. Social media analytics: Extracting and visualizing Hilton hotel ratings and reviews from TripAdvisor. *Int. J. Inf. Manag.* **2019**, *48*, 263–279. [[CrossRef](#)]
17. Sumarsono, D.; Sudardi, B.; Abdullah, W. The influence of TripAdvisor application usage towards hotel occupancy rate in Solo. In *Journal of Physics: Conference Series*; IOP Publishing: Bristol, UK, 2019; Volume 1175. [[CrossRef](#)]
18. Malazizi, N.; Alipour, H.; Olya, H. Risk perceptions of airbnb hosts: Evidence from a Mediterranean island. *Sustainability* **2018**, *10*, 1349. [[CrossRef](#)]
19. Zhang, Z.H.; Chen, R.J.C. Assessing Airbnb logistics in cities: Geographic information system and convenience theory. *Sustainability* **2019**, *11*, 2462. [[CrossRef](#)]
20. Deboosere, R.; Kerrigan, D.J.; Wachsmuth, D.; El-Geneidy, A. Location, location and professionalization: A multilevel hedonic analysis of Airbnb listing prices and revenue. *Reg. Stud. Reg. Sci.* **2019**, *6*, 143–156. [[CrossRef](#)]
21. Fierro, A.; Aranburu, I. Airbnb branding: Heritage as a branding element in the sharing economy. *Sustainability* **2019**, *11*, 74. [[CrossRef](#)]
22. Vu, H.Q.; Li, G.; Law, R.; Zhang, Y.C. Tourist activity analysis by leveraging mobile social media data. *J. Travel Res.* **2018**, *57*, 883–898. [[CrossRef](#)]
23. Chen, W.; Gao, Q.; Xiong, H. Temporal predictability of online behavior in foursquare. *Entropy* **2016**, *18*, 296. [[CrossRef](#)]
24. Aliandu, P. Sentiment analysis to determine accommodation, shopping and culinary location on foursquare in Kupang City. *Procedia Comput. Sci.* **2015**, *72*, 300–305. [[CrossRef](#)]
25. Chorley, M.J.; Colombo, G.B.; Allen, S.M.; Whitaker, R.M. Visiting patterns and personality of foursquare users. In Proceedings of the 2013 IEEE 3rd International Conference on Cloud and Green Computing (CGC 2013), Karlsruhe, Germany, 30 September–2 October 2013; pp. 271–276. [[CrossRef](#)]
26. Koerniawan, M.D.; Dewancker, B.J. Visitor Perceptions and effectiveness of place branding strategies in thematic parks in Bandung City using text mining based on Google Maps user reviews. *Sustainability* **2019**, *11*, 2123. [[CrossRef](#)]

27. de Vries, S.; Buijs, A.E.; Langers, F.; Farjon, H.; van Hinsberg, A.; Sijtsma, F.J. Measuring the attractiveness of Dutch landscapes: Identifying national hotspots of highly valued places using Google Maps. *Appl. Geogr.* **2013**, *45*, 220–229. [CrossRef]
28. Willis, R. Observations online: Finding the ethical boundaries of Facebook research. *Res. Ethics* **2019**, *15*, 1–17. [CrossRef]
29. Statistics of Accommodation Establishments. Available online: <http://datacube.statistics.sk/#!/lang/en> (accessed on 1 July 2019).
30. Hajduová, M.; Bartoš, M. *Košice in the Coordinates of European History*, 1st ed.; The Košice City Archives: Košice, Slovakia, 2013; p. 413. ISBN 978-80-971566-8-8.
31. The Monuments Board of the Slovak Republic. Databases. Available online: <https://www.pamiatky.sk/sk/page/databazy> (accessed on 10 October 2019).
32. European Commission. European Capitals of Culture. Available online: https://ec.europa.eu/programmes/creative-europe/actions/capitals-culture_en (accessed on 10 October 2019).
33. European Capitals and Cities of Sport Federation. European Cities of Sport. Available online: <http://aceseurope.eu/european-cities-of-sport/> (accessed on 10 October 2019).
34. United Nations Educational, Scientific and Cultural Organization. UNESCO Creative Cities Network. Available online: <https://en.unesco.org/news/64-cities-join-unesco-creative-cities-network> (accessed on 10 October 2019).
35. European Volunteer Centre. Košice 2019 European Volunteering Capital. Available online: <https://www.europeanvolunteercentre.org/ev-capital-2019> (accessed on 10 October 2019).
36. Eurostat. Culture and Tourism—Cities and Greater Cities. Available online: <https://ec.europa.eu/eurostat/data/database> (accessed on 10 October 2019).
37. Sidor, C. 12_eurostat_city_tourism.sql. Available online: https://github.com/csabasidor/occupancy_tax_payers_identifier/blob/master/12_eurostat_city_tourism.sql (accessed on 10 October 2019).
38. Geodetic and Cartographic Institute Bratislava. Territorial and Legal Arrangement of the Slovak Republic. Available online: https://www.geoport.sk/sk/zbgis_smd/na-stiahnutie/ (accessed on 1 July 2019).
39. What Is PostgreSQL? Available online: <https://www.postgresql.org/about/> (accessed on 1 July 2019).
40. About PostGIS. Available online: <https://postgis.net/> (accessed on 1 July 2019).
41. Siddique, M.I. I_Grid_Point_Distance.sql. Available online: https://github.com/I1mran/PostGIS-Custom/blob/master/Points_Grid/I_Grid_Point_Distance.sql (accessed on 1 July 2019).
42. Facebook Place Search API. Available online: <https://developers.facebook.com/docs/places/web/search/> (accessed on 1 July 2019).
43. Sidor, C. 01_fb_place_search_downloader.py. Available online: https://github.com/csabasidor/occupancy_tax_payers_identifier/blob/master/01_fb_place_search_downloader.py (accessed on 1 July 2019).
44. Places API. Available online: <https://developers.google.com/places/web-service/search> (accessed on 1 July 2019).
45. Sidor, C. 02_google_places_api_downloader.py. Available online: https://github.com/csabasidor/occupancy_tax_payers_identifier/blob/master/02_google_places_api_downloader.py (accessed on 1 July 2019).
46. Foursquare for Developers. Available online: <https://developer.foursquare.com/docs/api/venues/search> (accessed on 1 July 2019).
47. Sidor, C. 03_fourquare_api_downloader.py. Available online: https://github.com/csabasidor/occupancy_tax_payers_identifier/blob/master/03_fourquare_api_downloader.py (accessed on 1 July 2019).
48. Sidor, C. 04a_scrap_booking_only_for_scientific_purposes.py. Available online: https://github.com/csabasidor/occupancy_tax_payers_identifier/blob/master/04a_scrap_booking_only_for_scientific_purposes.py (accessed on 1 July 2019).
49. Sidor, C. 05a_scrap_tripadvisor_only_for_scientific_purposes.py. Available online: https://github.com/csabasidor/occupancy_tax_payers_identifier/blob/master/05a_scrap_tripadvisor_only_for_scientific_purposes.py (accessed on 1 July 2019).
50. Sidor, C. 06a_airbnb_get_urls_only_for_scientific_purposes.py. Available online: https://github.com/csabasidor/occupancy_tax_payers_identifier/blob/master/06a_airbnb_get_urls_only_for_scientific_purposes.py (accessed on 1 July 2019).

51. Sidor, C. 06b_airbnb_get_acc_only_for_scientific_purposes.py. Available online: https://github.com/csabasidor/occupancy_tax_payers_identifiers/blob/master/06b_airbnb_get_acc_only_for_scientific_purposes.py (accessed on 1 July 2019).
52. Sidor, C. 06c_clean_airbnb.py. Available online: https://github.com/csabasidor/occupancy_tax_payers_identifiers/blob/master/06c_clean_airbnb.py (accessed on 1 July 2019).
53. Lynn, S. Python Batch Geocoding.py. Available online: https://github.com/csabasidor/occupancy_tax_payers_identifiers/blob/master/04b_shane_lynn_booking_google_geocoder.py (accessed on 1 July 2019).
54. Sidor, C. 04b_shane_lynn_booking_google_geocoder.py. Available online: https://github.com/csabasidor/occupancy_tax_payers_identifiers/blob/master/05b_shane_lynn_tripadvisor_google_geocoder.py (accessed on 1 July 2019).
55. Sidor, C. 05b_shane_lynn_tripadvisor_google_geocoder.py. Available online: https://github.com/csabasidor/occupancy_tax_payers_identifiers/blob/master/05b_shane_lynn_tripadvisor_google_geocoder.py (accessed on 1 July 2019).
56. QGIS—A Free and Open Source Geographic Information System. Available online: <https://qgis.org/en/site/> (accessed on 1 July 2019).
57. OpenStreetMap contributors. OpenStreetMap. Available online: <https://d3js.org/> (accessed on 1 July 2019).
58. D3—Data-Driven Documents. Available online: <https://planet.osm.org/> (accessed on 1 July 2019).
59. Bostock, M. Sunburst. Available online: <https://observablehq.com/@d3/sunburst> (accessed on 1 July 2019).
60. Bostock, M. Radial Stacked Bar III. Available online: <https://bl.ocks.org/mbostock/3686329aa6e1f5938df8eef12ec353fe> (accessed on 1 July 2019).
61. Sidor, C. 07_notes_fb_case_sensitive_cleaning.sql. Available online: https://github.com/csabasidor/occupancy_tax_payers_identifiers/blob/master/07_notes_fb_case_sensitive_cleaning.sql (accessed on 1 July 2019).
62. Sidor, C. 08_notes_ASPIs_categorization.sql. Available online: https://github.com/csabasidor/occupancy_tax_payers_identifiers/blob/master/08_notes_ASPIs_categorization.sql (accessed on 1 July 2019).
63. FB Accommodation Facilities Categories. Available online: http://noob.sk.webx4.d2.cz/mdpi19/fb_sun.html (accessed on 1 July 2019).
64. Sidor, C. 09_notes_google_places_case_sensitive_cleaning.sql. Available online: https://github.com/csabasidor/occupancy_tax_payers_identifiers/blob/master/09_notes_google_places_case_sensitive_cleaning.sql (accessed on 1 July 2019).
65. Google Places Accommodation Facilities Categories. Available online: http://noob.sk.webx4.d2.cz/mdpi19/gg_sun.html (accessed on 1 July 2019).
66. Foursquare—Traffic Share by Country. Available online: https://pro.similarweb.com/#/website/audience-geography/foursquare.com/*999/3m (accessed on 1 July 2019).
67. Foursquare Accommodation Facilities Categories. Available online: http://noob.sk.webx4.d2.cz/mdpi19/fq_sun.html (accessed on 1 July 2019).
68. Booking Accommodation Facilities Categories. Available online: http://noob.sk.webx4.d2.cz/mdpi19/bk_sun.html (accessed on 1 July 2019).
69. TripAdvisor Accommodation Facilities Categories. Available online: http://noob.sk.webx4.d2.cz/mdpi19/ta_sun.html (accessed on 1 July 2019).
70. Decree No. 277/2008 of the Ministry of Economy of the Slovak Republic. Available online: <https://www.zakonypreludi.sk/zz/2008-277> (accessed on 1 July 2019).
71. Act No. 455/199 Coll. on Self-Employed Business. Available online: <https://www.epi.sk/zz/1991-455> (accessed on 1 July 2019).
72. Sidor, C. 10_airbnb_categorization_case_sensitive.sql. Available online: https://github.com/csabasidor/occupancy_tax_payers_identifiers/blob/master/10_airbnb_categorization_case_sensitive.sql (accessed on 1 July 2019).
73. AirBnB Accommodation Facilities Categories. Available online: http://noob.sk.webx4.d2.cz/mdpi19/abb_sun.html (accessed on 1 July 2019).
74. Sidor, C. 11_check_cross_platform_case_sensitive.sql. Available online: https://github.com/csabasidor/occupancy_tax_payers_identifiers/blob/master/11_check_cross_platform_case_sensitive.sql (accessed on 1 July 2019).

75. Cross Checked Results. Available online: http://noob.sk.webx4.d2.cz/mdpi19/scale_radial2.html (accessed on 1 July 2019).
76. The European Parliament and the Council of the European Union. Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the Protection of Natural Persons with Regard to the Processing of Personal Data and on the Free Movement of Such Data, and Repealing Directive 95/46/EC (General Data Protection Regulation). Available online: <https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:32016R0679&from=SK> (accessed on 20 October 2019).
77. The European Parliament and the Council of the European Union. Directive 96/9/EC of the European Parliament and of the Council of 11 March 1996 on the Legal Protection of Databases. Available online: <https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:31996L0009&from=EN> (accessed on 20 October 2019).



© 2019 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/>).