

Studying Spatial and Temporal Visitation Patterns of Points of Interest Using SafeGraph Data in Florida

Levente Juhász¹ and Hartwig Hochmair²

¹Florida International University, Miami, FL, USA

²University of Florida, Ft. Lauderdale, FL, USA

Abstract

SafeGraph is a commercial provider of massive Point of Interest (POI) data, including visitation patterns in North America. Although the data source does not share specific travel trajectories, the data available includes daily and monthly POI visitation numbers for over 160 categories, as well as information about where visitors come from and which other POI categories they visit. This allows analysts to gain insight into travel behavior in a geographic region over time. This study analyzes various aspects of visitation patterns that can be derived from the SafeGraph dataset for Florida. Using three major Florida cities, namely Miami, Orlando and Jacksonville, temporal patterns of daily and monthly visit numbers are correlated between various POI categories, and the effect of a short event (Hurricane Irma) on daily visitation numbers around the event is explored. In addition, travel distances from home to POIs are compared between different POI categories, and Ordinary Least Squares (OLS) regression models are used to identify factors associated with increased or decreased distance between home and a specific POI category. The study concludes that the aggregated data provided on the SafeGraph platform helps the GIScience community to learn more about travel patterns in both the spatial and the temporal domains.

Keywords:

Travel behavior, visiting patterns, Point of Interest, Florida, hurricane, urban environment

1 Introduction

Society increasingly utilizes location-based services (LBS) that cover a wide range of functionalities including navigation, social networking, assistive healthcare, customized advertising, event recommendation and participatory decision making. LBS often use spatial information derived from Point of Interest (POI) information, for example when recommending overnight accommodations based on user reviews. The tech industry utilizes POIs in geo-gaming and mapping applications (Juhász & Hochmair, 2017; Juhász, Novack, Hochmair, & Qiao, 2020), and to derive detailed land use/land cover information (Spyratos, Stathakis, Lutz, & Tsinaraki, 2017). Apart from being a static collection of places, POI data combined with visitor patterns can be used to study urban dynamics and user markets. This

added information is commonly applied for location-based advertising or consumer analytics (Baik, Lee, Lee, Kim, & Choi, 2016). For example, Foursquare check-in information can help to predict the types of places a user will visit in the future (Zhuang et al., 2017), or one can provide POI visiting recommendations based on the analysis of visit trajectories (Massimo & Ricci, 2019).

Since POIs can serve as the data foundation for a variety of industry applications and solutions for answering societal questions, there is no single best source of POI in general. That is, different datasets can vary in content, completeness or quality depending on their purpose. For example, POIs compiled from social media services tend to be more abundant than POIs from business- and mapping-oriented sources, but at the same time they tend to have higher positional errors (Hochmair, Juhász, & Cvetojevic, 2018).

SafeGraph is a commercial provider of POI data that compiles its dataset from several sources, such as mobile phone GPS data and governmental open data, to build a comprehensive business listing in the United States and Canada. In addition to the POI data itself, the company derives visitation pattern information and aggregates it to POIs, which enables access to visitor and visit counts and certain aspects of demographics. SafeGraph POI visitation data, which is made available for academic research free of charge in aggregated form, can be a useful source of information for studying certain aspects of urban dynamics and travel behavior. For example, one study used SafeGraph data to reconstruct origin–destination pairs in Milwaukee, Wisconsin, in order to explore the spatial isolation of neighborhoods (Prestby, App, Kang, & Gao, 2019). Different machine learning models applied to SafeGraph data have also been used to predict parking violations (Gao et al., 2019) which depended on certain POI categories, such as retail stores or restaurants. The dataset can also be used to assess the effect of certain events or policy changes on visitation patterns. This has been illustrated in a study that analyzes the change in visitation patterns after Starbucks implemented an open bathroom policy allowing anyone, even without a purchase, to use their bathroom facilities. Results revealed a 6.8 % decline in store visits compared to other nearby restaurants and cafes (Gurun, Nickerson, & Solomon, 2020). More recently, SafeGraph data has been used to assess compliance with guidelines on social distancing in response to COVID-19 (Andersen, 2020) and for building a POI database in conjunction with other data sources for informed decision making (Killeen et al., 2020).

This research contributes to the growing body of literature using SafeGraph data and conducts exploratory analyses in the spatial and temporal domains using three major cities in Florida (Miami, Orlando and Jacksonville) as study areas. More specifically, the study has the following objectives:

1. To compare the temporal characteristics of visitation patterns between different POI categories using monthly and daily correlation analysis
2. To study the effect of short-term events on POI visitation patterns, and
3. To analyze distances between visitors' home locations and POIs visited.

In pursuing these objectives, the study showcases novel analysis approaches applied to the SafeGraph dataset. These include, for example, analyzing the effects of hurricanes on visitation frequency to certain POI types before, during and after the hurricane, reflecting the type of preparation that the population are making for such an event. It also identifies the localness of POI types, i.e. the role of a specific POI type for the local population versus its role for visitors from further away.

2 Study Setup

2.1 POI data collection and pre-processing

SafeGraph’s main product is SafeGraph Places, which consists of three datasets, namely Core POI, Geometry and Patterns. The *Core POI* dataset contains basic information of about 6.1 million POIs in the US and Canada, such as the name, brand association (i.e. if the POI is part of a chain), address, category and opening hours, along with an internal place ID. As well as the point geometry of POIs, the *Geometry* dataset also contains their polygon representation, for example the outline of the buildings that POIs are housed in. The hierarchy of POI locations is also included in this dataset. That is, POIs can be nested within each other, which is often the case when a larger entity, such as a shopping mall, contains multiple individual stores. The *Patterns* dataset describes visitation patterns to over 3.6 million unique POIs. These patterns include monthly aggregated visitor and visitation numbers, daily visits and dwell times. The dataset also includes the number of home and work locations of visitors as well as number of smartphone devices observed, which are aggregated at the level of the US census block group. SafeGraph uses accurate smartphone GPS locations and machine learning to attribute visits to POIs.

SafeGraph provided us with their Florida dataset, which consists of 302,201 POIs. Among these, 258,658 POIs also contain Patterns data for at least one month between January 2017 and August 2019. The dataset is available as a collection of plain text flat files. These files were parsed and inserted into a spatially enabled PostgreSQL database using the Places schema (SafeGraph, 2020) for further processing. Standalone tables can be joined by the common internal place ID that is attached to all places, geometries and patterns. Even though it is stored in a relational database, the schema contains several fields represented as JSON documents. Custom SQL queries were designed to extract information (e.g. from JSON) to complete each analysis step (described in Section 2.2). The study is geographically limited to three major cities in Florida by filtering POI addresses. Further, the analysis was also limited to 17 POI categories. The POI numbers included in this study for the different POI categories and cities are summarized in Table 1.

Table 1: Summary of POI categories and the number of POIs per city included in the study

Code	Category name	# of POIs included		
		Miami	Orlando	Jacksonville
ACC	Accounting, Tax Preparation, Bookkeeping, and Payroll Services	173	150	141
AMU	Amusement Parks and Arcades	19	54	15
AUTO	Automobile Dealers	278	249	163
ALC	Beer, Wine, and Liquor Stores	113	74	71
UNI	Colleges, Universities, and Professional Schools	97	71	77
ELEC	Electronics and Appliance Stores	89	60	61
SCH	Elementary and Secondary Schools	743	445	502
GAS	Gasoline Stations	377	253	356
GROC	Grocery Stores	449	403	427
HOME	Home Health Care Services	102	45	58
LUG	Jewelry, Luggage, and Leather Goods Stores	311	158	101
MUS	Museums, Historical Sites, and Similar Institutions	182	132	146
DOC	Offices of Physicians	1,648	944	859
OTH_AM	Other Amusement and Recreation Industries	579	378	311
POST	Postal Service	37	23	25
REST	Restaurants and Other Eating Places	2,975	2,697	1,994
TRAV	Traveler Accommodation	184	335	146
	Sum	8,356	6,471	5,403
	Total POI included	20,230		

2.2 Analysis Methods

Temporal visiting patterns

Multiple approaches were used to assess the usability of the SafeGraph dataset for analyzing temporal visiting patterns at different time scales. This was achieved by pursuing three tasks: 1) comparison of monthly aggregated visitation patterns between different POI categories; 2) comparison of daily visitation patterns across different POI categories, and 3) exploration of the effects of a short-term event on visitation patterns.

For the first task, seasonal visitation patterns over a year were analyzed. Monthly aggregated visits and monthly unique visitor numbers from the Patterns dataset for each POI were used and correlated between different POI categories. POIs were filtered by category (see Table 1) and city (Miami, Orlando and Jacksonville). To smoothen monthly count data, the average of

the monthly counts from different years was used where available in our study time frame. To compare visitation patterns between POI categories, the average monthly visit and visitor count values were normalized to a range of 0 to 1 for each city and for each category, where the month with the lowest average visit or visitor number was attributed a value of 0, and 1 was assigned to the month with most visits or visitors. In the final step, Pearson correlation coefficients were calculated for each POI category pair. These correlation matrices were created for each city separately. This method of identifying temporal similarity of activity patterns through correlation has been used before, e.g. for comparing weekday and weekend cell phone communication counts (SagI, Delmelle, & Delmelle, 2014).

Daily visit numbers from SafeGraph Patterns data were used to study visitation patterns at a more refined temporal scale. Since seasonal patterns greatly affect visitation numbers, only one month of data (May 2018) was used to analyze daily visits. This month was chosen because it does not include any national US holiday and summer break for schools has not yet started, which means that a typical visitation behavior can be expected. As for task 1, daily visit numbers were extracted and normalized for each city and POI category. Correlation matrices were computed as above, which allows the assessment of the similarity of daily visitation patterns between different types of POIs. The effects of a short-term event on POI visitation rates (task 2) were also explored using daily pattern data. As a showcase, we used Hurricane Irma, which affected the Miami metropolitan area in September 2017 through heavy wind gusts and flooding. The average daily visit per POI category between 25 August and 30 September 2017 was calculated and compared to a reference dataset for the same time period in 2018.

Distance from home

As one of its attributes, the SafeGraph monthly Patterns data includes the median distance between visitors' home locations and a POI. Median distances from June 2018 for the three cities were analyzed with respect to the 17 selected POI categories (see Table 1) in order to identify which types of POIs are associated with shorter or longer travel distances. Comparison of median distances for POIs within a city was conducted through a series of unpaired two-sample Wilcoxon rank sum tests. This type of test is a nonparametric test of the null hypothesis that the medians of two populations are equal. For this analysis, we expected that POIs from categories that provide local services for everyday activities (e.g. grocery stores, gas stations, post offices, or schools) would be closer to home locations than POIs used for recreational or travel-related activities, such as amusement parks, hotels or restaurants, which provide distinct services at specific locations and thus justify longer trips.

The same set of observations was used in a series of multiple linear regression models. Through the use of several explanatory variables, these models predict the median travel distance for a given POI category and city. The predictor variables can be subdivided into those describing spatial characteristics of other POIs in the same category, sociodemographic factors at the US census block group level, and location of a POI in the study area relative to city centers, highway access points, and major airports (Table 2). For each city–category combination, Spearman's rho correlation coefficient was computed between all candidate explanatory variables. To mitigate multicollinearity, predictor combinations with a high

correlation of $|r| > 0.7$ were avoided during the model-building process. All other predictors, even if not significant, were retained in the final models presented here.

Table 2: Candidate explanatory variables for the regression analysis

Variable	Operationalization	Data Source
POI		
Nearest neighbor (NN) distance	Distance to nearest POI in same category (in m)	SafeGraph
POI count	Number of POIs in same category within a 5-km buffer	
Sociodemographic		
Job density	Number of jobs per km ² in census block group	US Census Bureau – LEHD
Population density	Population per km ² in census block group	US Census Bureau
Location		
Distance to CBD	Direct distance between POI and Central Business District (in m)	http://www.city-data.com
Distance to highway	Direct distance to nearest highway access point	HERE NAVSTREETS
Distance to nearest major airport	Direct distance (in m) to Miami/Orlando/Jacksonville/Fort Lauderdale-Hollywood International Airport, whichever is closest	Natural Earth Airports

We hypothesized that, in general, a higher density of POIs in the vicinity of an analyzed POI (as operationalized in the first two variables) would facilitate short trips to a POI of that type. Census block groups with higher job densities mark areas to which people commute and hence where they perform activities away from home. Low density population areas mean less access to activity opportunities and are therefore expected to be associated with longer travel to POIs. Central Business Districts (CBDs) offer a wide range of activities for visitors from outside (tourists, business travelers), so that CBDs are expected to be associated with longer POI–home distances. Access to highways (i.e. shorter distance to highway access points) facilitates larger activity spaces for a given time budget (Parthasarathi, Hochmair, & Levinson, 2015), because of higher speed limits and the absence of intersections on highways. Proximity to highways can therefore be expected to lead to longer distances between home and POI. At least for some POI categories, e.g. gas stations or hotels, POIs near airports are more likely to be visited by air-travelers (many of whom live outside the local area) than other parts of a city, which means that POIs within short distances of airports can be expected to be associated with greater home-to-POI distances.

3 Analysis Results

3.1 Temporal Patterns

Comparison of POI categories

It was expected that different POI categories would show different visitation patterns due to the different nature of their target audiences. To analyze this, correlation matrices of monthly visit and visitor numbers for each city were computed between 17 POI categories based on the normalized values. Figure 1 plots the correlation of monthly visits between POI categories in Orlando. The main diagonal of the plot shows normalized visits between January and December as first and last data points. The lower triangular matrix plots the two normalized visit curves against each other for category pairs, while the upper triangular matrix reports the Pearson correlation coefficients along with their levels of significance. The matrix reveals that in Orlando monthly visitation to POIs generally follows the same pattern. This is indicated by most coefficients between category pairs being positive. Categories that do not follow the usual pattern are highlighted by negative correlation coefficients, in blue. An example of this is Universities, Colleges and Professional Schools, which are negatively correlated with all other categories except elementary schools, electronic stores and accounting services. This can be partially explained by the presence of the University of Central Florida (UCF) in Orlando. This institution has the largest university campus in the US in terms of enrollment, hosting more than 66,000 undergraduate and graduate students. The lower number of students present on campus during summer and winter breaks results in a deviation from the typical POI visit pattern, as indicated by the negative correlation coefficients (which, however, are mostly not significantly different from zero). Miami and Jacksonville show similar monthly visitation patterns with fewer cases of negative correlations between POI categories and some local differences. Two other statistically significant negative correlations are found in Miami between Universities and Amusement Parks, and between Schools and Amusement Parks.

This analysis can be complemented by calculating the same matrix for the number of monthly unique visitors (not shown in the Figures). This would allow the joint interpretation of visit and visitor matrices in order to gain more insights into urban dynamics.

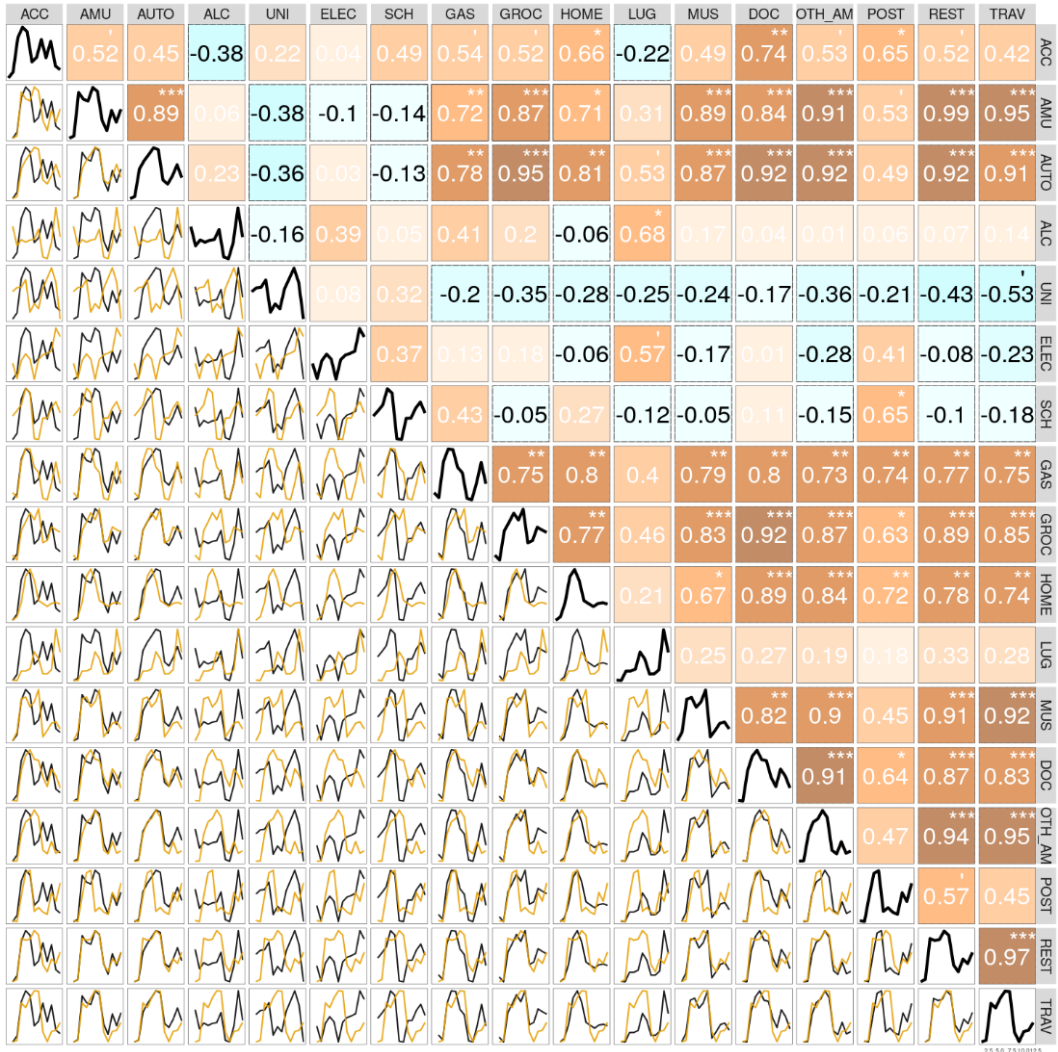


Figure 1: Correlation of monthly visits between selected POI categories in Orlando. (Significance codes: 0 ***, <0.001 **, <0.01 *, <0.05 '). Abbreviations of category labels are explained in Table 1.

Figure 2 shows the correlation plot of daily visits between POI categories in May 2019 for Orlando. Not surprisingly, among other things the plot reveals the negative relationship between visits to universities and amusement parks and hotels (Traveler accommodation). This is because universities are typically attended during weekdays, whereas entertainment facilities tend to be more visited on weekends. Outliers caused by local short-term events (e.g. storm, spike in gas price or athletic games) can influence the results of similar analyses by obfuscating real relationships.

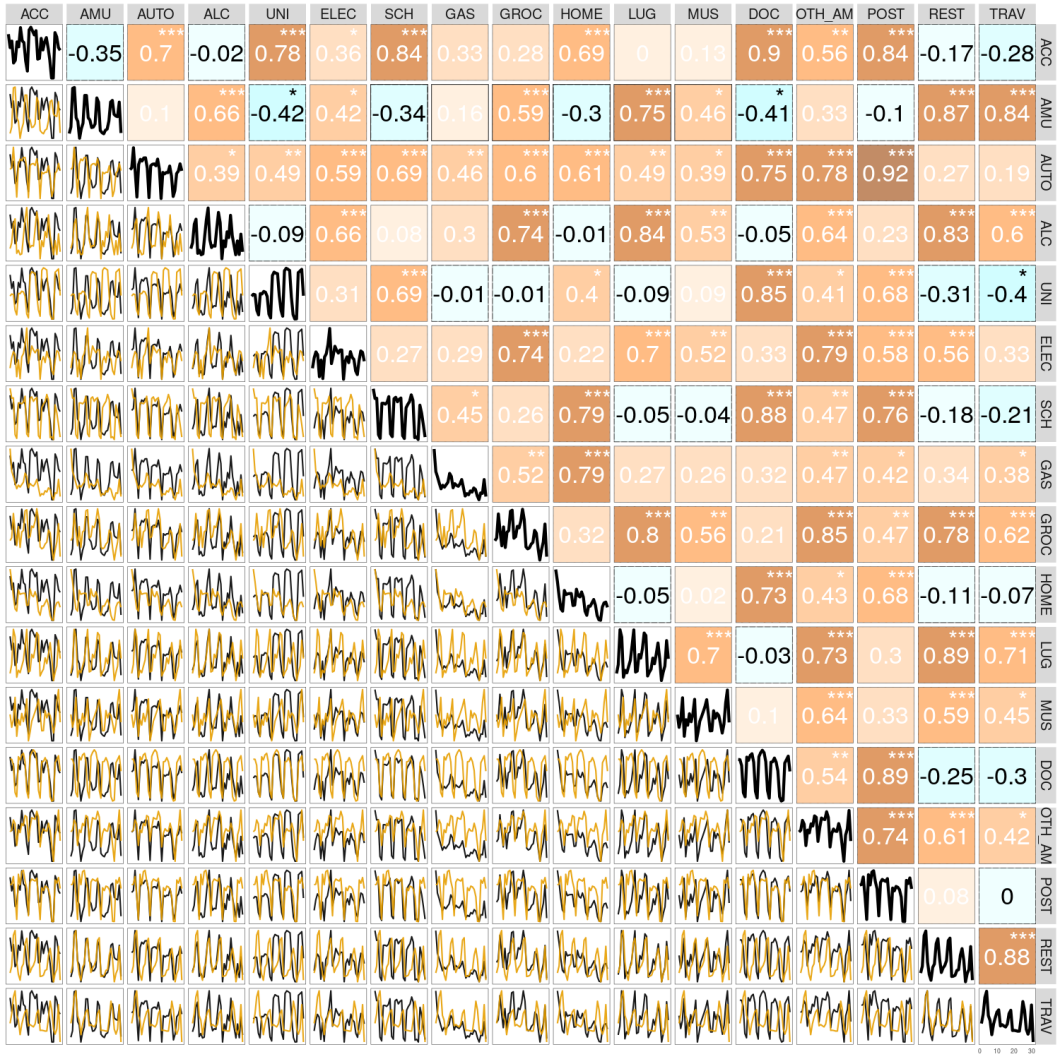


Figure 2: Correlation of daily unique visitors between selected POI categories in Orlando. (Significance codes: 0 ***, <0.001 **, <0.01 *, <0.05 '). Abbreviations of category labels are explained in Table 1.

Event analysis

The effect of short-term events on visitation patterns was explored in more detail by analyzing the effect of Hurricane Irma on the Miami metropolitan area. Hurricane Irma was the most intense hurricane to make landfall in contiguous United States since Hurricane Katrina in 2005. The hurricane had the potential to hit the Miami metropolitan area directly. However, it eventually struck the Florida Keys on September 10, 2017, less than 200 km from Miami. The effects of this hurricane are clearly visible in Figure 3a, which plots the average number of daily visits per POI in a given category between August 25 and September 30, 2017. In the shaded areas, Figure 3a also shows the 95% confidence interval. Yellow vertical bars denote

weekends; the time between September 9 and 11 when the hurricane was closest to Miami is highlighted with an orange vertical bar. On September 7, Miami-Dade County expanded the mandatory evacuation order for residents in all evacuation zones, which affected more than 650,000 people (red vertical line in Figure 3a). The figure shows increased visits to gas stations and grocery stores before the event, which is the usual behavior when a storm is projected to hit an area as residents stock up with non-perishable food and fill up their vehicles with gas. The data also reflects the response of higher educational institutes in the area, as Florida International University, the University of Miami and Miami-Dade College all canceled classes on September 6, with Miami-Dade College remaining closed for the remainder of the week. After the evacuation order, visits to grocery stores and gas stations decreased rapidly. All categories reached their minimum visitation rates on the day of the landfall (September 10). After the storm, visits to grocery stores and gas stations increased more rapidly than visits to universities and colleges. The data shows that it took 4–5 days to reach pre-storm visitation levels in gas stations and grocery stores, while universities and colleges did not resume normal operations until the following week.

In order to compare this pattern to an unaffected time period, average visits per POI in Miami were calculated for the same period in 2018. Slightly different dates (24 August – 29 September, 2017) were used for a pairwise comparison in order to match the days of the week between the two periods (Figure 3b). The difference between the average number of visits per POI in 2018 and 2017 (i.e., year 2018 visits minus year 2017 visits) is shown in Figure 3c. Positive values show an increase in the average number of visits in 2018 compared to the previous year, while negative values mean the opposite. Since the hurricane peaked on a weekend, the observed decrease in visit numbers during the hurricane is smallest for universities (Figure 3c), since these institutions do not generally host many activities on weekends.

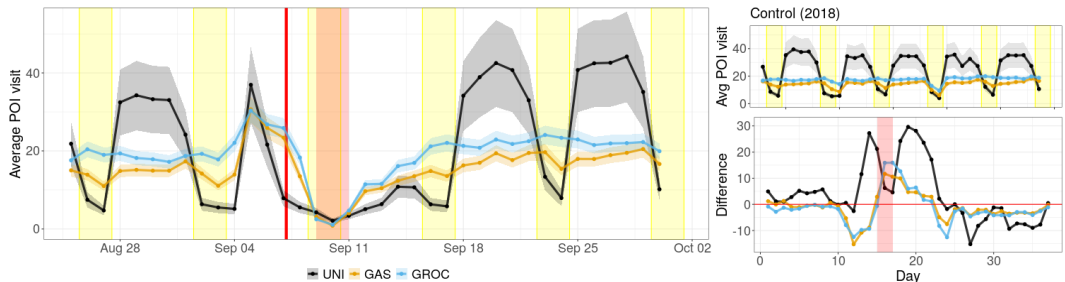


Figure 3: The effect of Hurricane Irma on visitation patterns in three POI categories in Miami, illustrated by (a) the average daily visits per POI in 2017; (b) the average daily visits per POI during the control time period in 2018; (c) the difference in the average daily visits between the control and hurricane periods

3.2 Distance from Home

Spatial and temporal patterns

Figure 4 maps the median distance from home to grocery store (MD = 7.1 km) and restaurant (MD = 9.7 km) POIs in the Jacksonville study area for June 2019. Yellow circles (restaurants)

tend on average to be larger than red ones (grocery stores), suggesting that grocery shopping is a more localized activity than visiting restaurants.

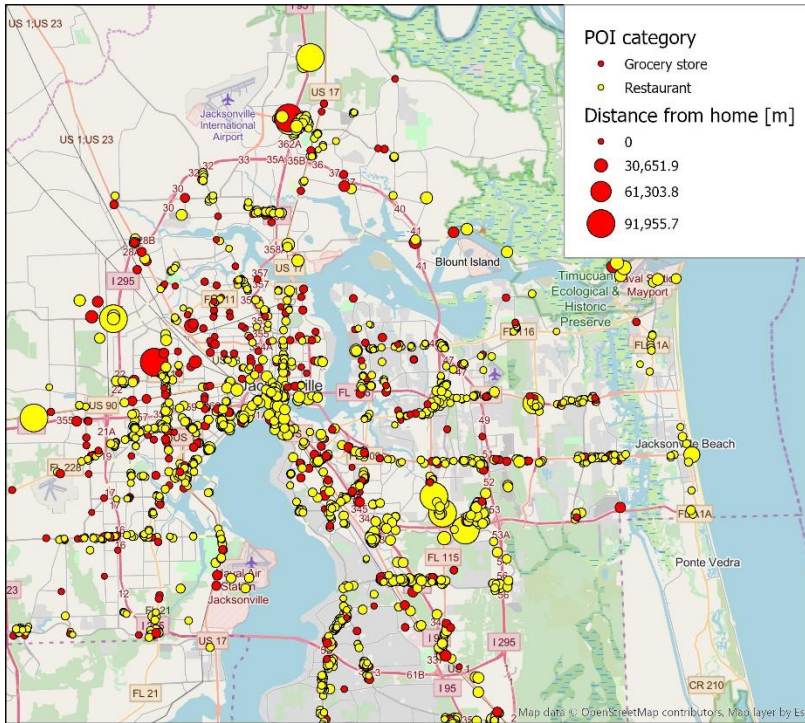


Figure 4: Median distance from home to grocery stores and restaurants for June 2019 POI visits in the Jacksonville study area

Flow maps (Figure 5) were generated by counting the number of residents in US census block groups who visited three Orlando theme parks (Animal Kingdom, Universal Studios, and SeaWorld) during August 2018 and December 2018, based on SafeGraph aggregated Pattern counts. Map comparison reveals a change in visitor patterns towards more people from the south-east and mid-west traveling to Florida in December, possibly to enjoy milder temperatures during the winter season.

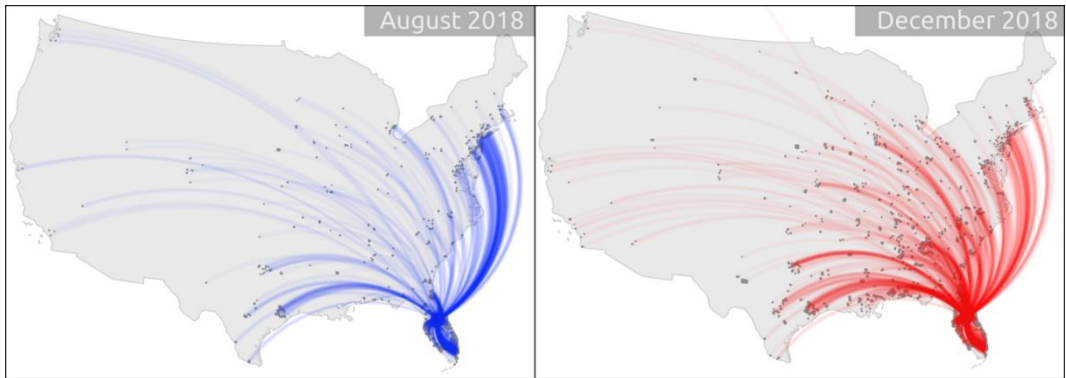


Figure 5: Home locations of Orlando theme park visitors in August and December 2018 at census block group level.

Distances between home and POI vary significantly over the year for some POI categories. As an example, monthly median distances between home and POI visits to Orlando amusement parks and traveler accommodations (hotels) between July 2018 and June 2019 are plotted in Figure 6. Distances to amusement parks (Figure 6a) peak in March, a month which includes the spring break, during which Florida is a traditional travel destination. Smaller peaks can be found in June and July (US holidays) and November (which includes Thanksgiving), all of which appear to contribute to traveling to Orlando theme parks from more distant locations than during the off-season. Distances from home to travel accommodation (Figure 6b) also peak in March, but are generally higher during the winter season (December through April), suggesting more visits from residents outside the Florida region during that time. Longer distances are also observed in June and July, when school vacation allows for extended family vacations and longer-distance travel.

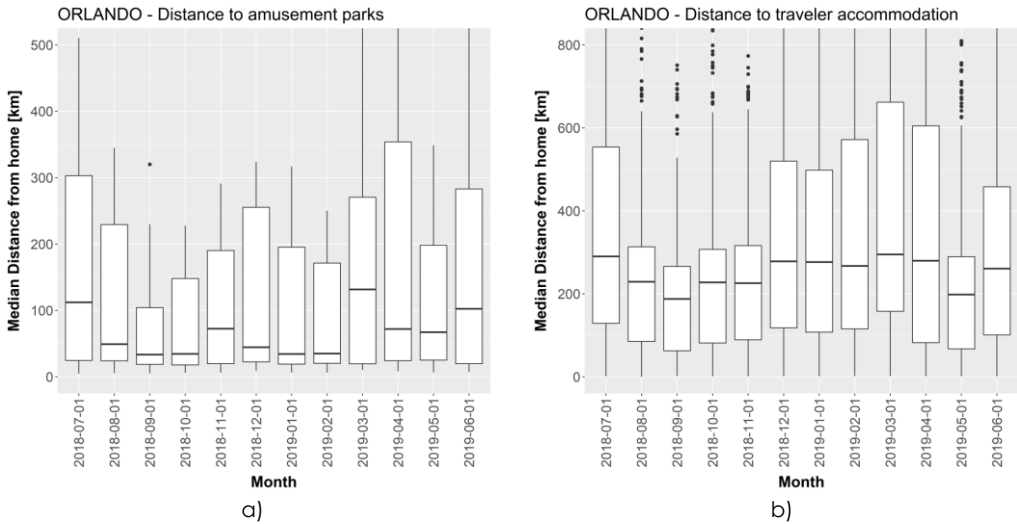


Figure 6: Box plots of median distances, over 12 months, from home to (a) Orlando amusement parks, and (b) travel accommodations

Variation across POI categories

Figure 7 shows box plots of distances from home to POIs for the 17 top POI categories analyzed for (a) Miami, (b) Orlando, and (c) Jacksonville for June 2018. In all three cities, travel accommodation (hotels) attracts those whose homes are farthest away, which can be expected since locals use hotels less frequently than people from outside the region. Another noticeable pattern is greater travel distances for Orlando amusement parks compared to those of Miami and Jacksonville. This can be attributed to the national and worldwide popularity of Orlando's theme parks, whereas Miami is known, rather, for its beaches and nightlife, and Jacksonville is a center for healthcare, retail, marine transportation, and finance. All other distance medians associated with different POI categories vary within only a small range – of about 5 to 15 km across each city.

Wilcoxon rank sum tests, with the Bonferroni correction used for multiple testing, were applied to identify which POI category / distance from home to POI pairs differ at a 5% level of significance. In all three cities, distance to travel accommodation is significantly greater than for all other categories. In Miami, median distances were lowest for postal service (4.7 km), grocery (4.9 km), liquor store (5.4 km), school (5.5 km), and gas station (5.6 km). There was no significant difference between these categories, but the distances were significantly shorter than those of all or most other categories (e.g. physician or museum). This shows that local services allow shorter trips. Orlando reveals a similar pattern, where schools (MD = 7.2 km) have the shortest distance, probably because this type of POI is not affected by tourists who travel from further away. In Jacksonville, distances for visits to grocery stores are shortest (MD = 7.1 km), followed by liquor stores (MD = 7.8 km) and gas stations (MD = 8.4 km); grocery store distances are significantly shorter than distances to all other POI categories.

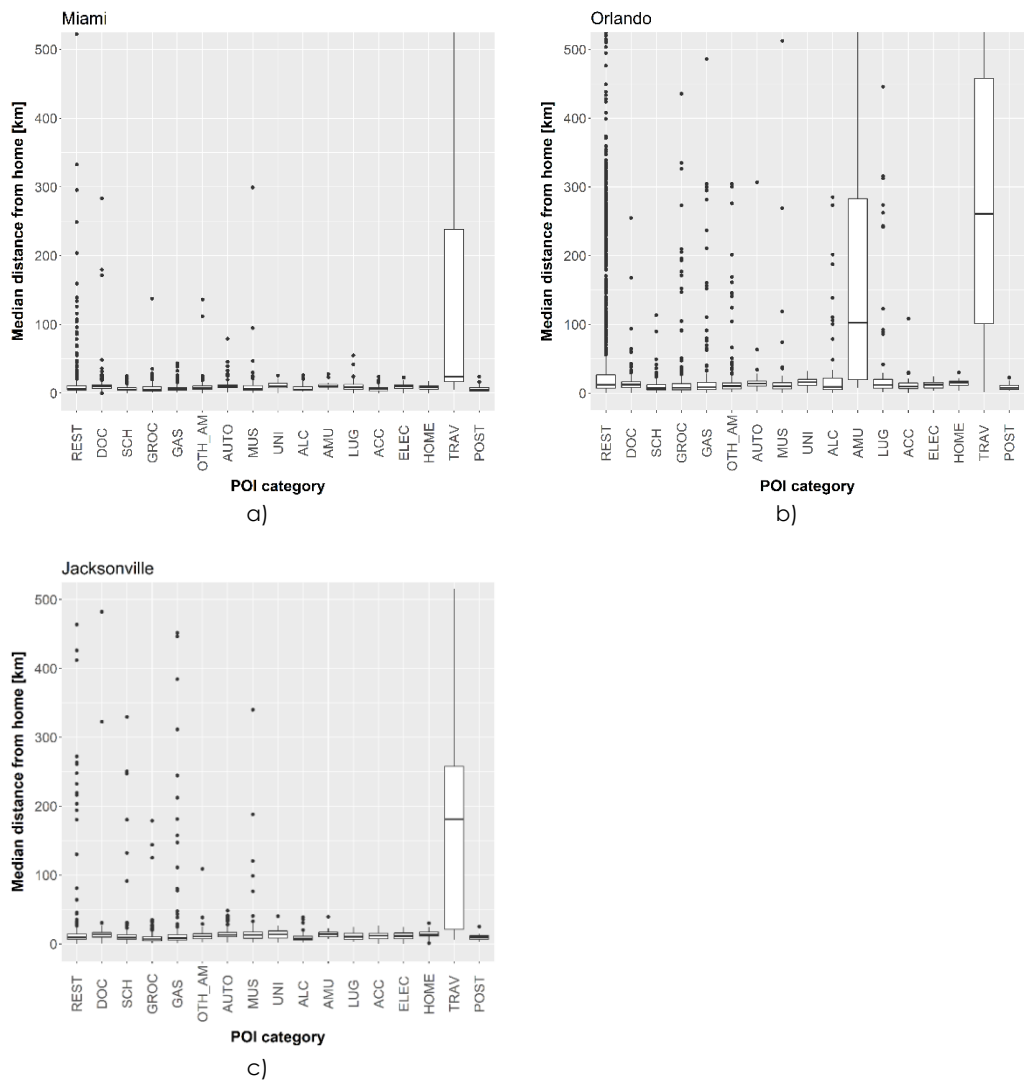


Figure 7: Scatterplots of distance from home for different POI categories in (a) Miami, (b) Orlando, and (c) Jacksonville

Regression analysis

Using the set of seven candidate explanatory variables listed in Table 2, linear regression models that predict the distance between POI and home census block group were constructed for each POI category in each of the three cities. Not every model resulted in significant coefficients. Some of the results appear to be affected by local characteristics of the layout of a city, whereas other findings hold across all three cities analyzed. Table 3 shows the model results for three different POI categories in each of the three cities. Models shown in the table were selected based on the explanatory power (adjusted R-square) and interpretability and usefulness of regression coefficients. Each model description comes with the number of

observations (N) and the adjusted R-square value. Blank cells indicate predictors removed due to collinearity. POIs for everyday services (gas stations, grocery stores) showed the expected effect of POI abundance within a 5-km radius, namely a decrease in distance between home and POI. This suggests that customers of these POIs tend to choose a facility near their home, and that longer trips to another POI which provides similar services would not be justified. For POIs associated with more internal variability (restaurants, museums / historical sites / nature parks), the opposite is the case. This suggests that visitors from further away, such as tourists, tend to visit areas with a higher abundance and more clustering of such facilities. The positive coefficient of POI NN distance for Miami museums and historical sites can be explained by the long driving distances required to get to state and national parks (e.g. Shark Valley Visitor Center) located outside the city boundaries. Arithmetic signs of coefficients for job density (positive) and population density (negative) are as expected across the three cities where significant. Proximity to highway access points was associated with greater travel distances, as expected, for grocery and jewelry stores. However, gas stations further away from highways had greater travel distances, possibly due to generally lower gas prices further away from highways. Longer trips to museums/historical sites away from highways, as shown for Miami, are likely explained by the remoteness of parks from highways (mentioned above). While, as expected, POI proximity to airports (i.e. shorter distance) is associated with longer distances from home (due to visitors coming from out of the state), this is not the case for Orlando restaurants. One reason could be that many restaurants are clustered around the Orlando theme parks, which are themselves located 20 or more km away from the airport. These restaurants, often visited by tourists from far away, could explain the positive regression coefficient for airport distance. The same explanation could hold for the positive signs associated with CBD distance for Orlando restaurants and hotels, since theme parks are between 15 and 30 km from the CBD.

Table 3: Coefficient estimation results for selected OLS models (Significance codes: *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$)

<i>Miami</i>	<i>Gas station</i>	<i>Jewelry store</i>	<i>Museum/historical site</i>
POI NN distance	-0.589	-1.15	6.52E+01***
# POI within 5 km	-9.15E+01***	8.46E01***	3.44E+03***
Job density	-5.73E-03	-1.68E-03	1.19
Population density	-0.254**	-0.597***	-2.38
Highway distance	0.529**	-0.392	4.64E+01***
CBD distance			
Airport distance	-0.338***	5.85E-02	-1.33
Constant	1.45E+04***	7.48E0***	-2.29E+05***
N (Adj. R ²)	347 (0.132)	167 (0.313)	151 (0.650)
<i>Orlando</i>	<i>Grocery</i>	<i>Restaurant</i>	<i>Travel accommodation</i>
POI NN distance	1.24E01	-2.47E+01**	-4.34E+01***
# POI within 5 km	-8.00E+02**	4.43E+02***	4.09E+02
Job density	0.179	-2.11E-03	1.948
Population density	-2.49	-0.429	-2.79
Highway distance	-1.32E+01***	0.780	3.62E+01
CBD distance		1.66E+01***	2.35E+02***
Airport distance	1.28	8.64***	-2.16
Constant	6.28E+05**	-3.56E+05	-2.76E+05
N (Adj. R ²)	353 (0.048)	2302 (0.304)	299 (0.227)
<i>Jacksonville</i>	<i>Gas station</i>	<i>Jewelry store</i>	<i>Car dealer</i>
POI NN distance	-2.42	-1.49*	0.270
# POI within 5 km	-3.48E+02**	2.07E+02*	-1.45E+01
Job density	0.359**	3.43E-02	0.782
Population density	-3.81*	7.16E-03	-1.15*
Highway distance	-1.06	-1.21**	-0.136
CBD distance		-3.76E-02	0.430***
Airport distance	-0.563**	5.43E-02	-0.219**
Constant	4.15E+05***	1.06E+04***	1.52E+04***
N (Adj. R ²)	320 (0.083)	67 (0.322)	152 (0.134)

4 Conclusions

This study analyzed spatial and temporal visitation patterns of POIs in Miami, Orlando and Jacksonville, using SafeGraph POI and Patterns data. It extends the growing body of literature (Andersen, 2020; Gao et al., 2019; Killeen et al., 2020; Prestby et al., 2019) using this data source and provides insights into potential use cases of the data.

Correlation matrices based on aggregated monthly and daily visitation counts identified POI category pairs of similar or dissimilar temporal activity patterns. This type of analysis can be used to reveal whether visitation patterns to a certain POI category deviate from the rest. As an example, it was shown that the visitation pattern to universities in Orlando was different from other categories in the city due to the absence of students during breaks. The ability to analyze the effects of short-term events on visitation patterns was demonstrated by using Hurricane Irma as a case study and by comparing the average number of visits to selected POI categories to a control period in Miami. In the future, this analysis could be extended to measure changes in visitation patterns caused by policy changes or other types of events, such as sporting events or disease outbreaks.

The research demonstrated that the dataset analyzed is a viable source of information for several analysis tasks, although only check-in data but not travel trajectory data is provided. The latter is a limitation compared to other freely available sources, such as tweets, which can be used to derive travel trajectories (Han, Tsou, Knaap, Rey, & Cao, 2019). The methodology presented here, such as co-interpreting visit and visitor correlation matrices, or determining travel distances to POIs for different categories, has the potential to aid urban planners and city managers to better understand the dynamics of a city and to complement data from local visitor surveys on travel and visitation behavior.

Acknowledgements

The authors thank SafeGraph Inc. (www.safegraph.com) for providing free access to Florida POI data, including Core Places, Geometry, and Patterns tables, for this research.

References

- Andersen, M. (2020). Early Evidence on Social Distancing in Response to COVID-19 in the United States. Retrieved from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3569368
- Baik, J., Lee, K., Lee, S., Kim, Y., & Choi, J. (2016). Predicting personality traits related to consumer behavior using SNS analysis. *New Review of Hypermedia and Multimedia*, 22(3), 189-206.
- Gao, S., Li, M., Liang, Y., Marks, J., Kang, Y., & Li, M. (2019). Predicting the spatiotemporal legality of on-street parking using open data and machine learning. *Annals of GIS*, 25(4), 299–312.
- Gurun, U. G., Nickerson, J., & Solomon, D. H. (2020). The Perils of Private Provision of Public Goods. Retrieved from http://www.umatgurun.com/wp-content/uploads/2019/12/PublicProvision_GNS.pdf
- Han, S. Y., Tsou, M.-H., Knaap, E., Rey, S., & Cao, G. (2019). How Do Cities Flow in an Emergency? Tracing Human Mobility Patterns during a Natural Disaster with Big Data and Geospatial Data Science. *Urban Science*, 3(2), 3020051. doi:10.3390/urbansci3020051

- Hochmair, H. H., Juhász, L., & Cvetojevic, S. (2018). Data Quality of Points of Interest in Selected Mapping and Social Media Platforms. In P. Kiefer, H. Huang, N. Van de Weghe, & M. Raubal (Eds.), *Progress in Location Based Services 2018* (Vol. Lecture Notes in Geoinformation and Cartography, pp. 293-313). Springer: Berlin.
- Juhász, L., & Hochmair, H. H. (2017). Where to catch 'em all? – A geographic analysis of Pokémon Go locations. *Geo-spatial Information Science*, 20(3), 241-251.
- Juhász, L., Novack, T., Hochmair, H. H., & Qiao, S. (2020). Cartographic Vandalism in the Era of Geo-Gaming -The Case of OpenStreetMap and Pokémon GO. *ISPRS International Journal of Geo-Information*, 9(4), 197.
- Killeen, B. D., Wu, J. Y., Shah, K., Zapaishchykova, A., Nikutta, P., Tamhane, A., . . . Unberath, M. (2020). A County-level Dataset for Informing the United States' Response to COVID-19. Retrieved from <https://arxiv.org/abs/2004.00756>
- Massimo, D., & Ricci, F. (2019). Clustering Users' POIs Visit Trajectories for Next-POI Recommendation. In J. Pesonen & J. Neidhardt (Eds.), *Information and Communication Technologies in Tourism 2019* (pp. 3-14): Springer.
- Parthasarathi, P., Hochmair, H. H., & Levinson, D. M. (2015). Street Network Structure and Household Activity Spaces. *Urban Studies*, 52(6), 1090-1112.
- Prestby, T., App, J., Kang, Y., & Gao, S. (2019). Understanding neighborhood isolation through spatial interaction network analysis using location big data *Environment and Planning A*.
- SafeGraph. (2020). Places Schema. Retrieved from <https://docs.safegraph.com/docs/places-schema>
- Sagl, G., Delmelle, E., & Delmelle, E. (2014). Mapping collective human activity in an urban environment based on mobile phone data. *Cartography and Geographic Information Science*, 41(3), 272-285. doi:10.1080/15230406.2014.888958
- Spyratos, S., Stathakis, D., Lutz, M., & Tsinaraki, C. (2017). Using Foursquare place data for estimating building block use. *Environment and Planning B, Planning and Design*, 44(4), 693–717.
- Zhuang, Y., Fong, S., Yuan, M., Sung, Y., Cho, K., & Wong, R. K. (2017). Location-based big data analytics for guessing the next Foursquare check-ins. *The Journal of Supercomputing*, 73, 3112-3127.