Assessing the validity of SafeGraph data for visitor monitoring in Yellowstone National Park

Assessing the validity of SafeGraph data for visitor monitoring in Yellowstone National Park

Introduction

In 2019, the United States national parks attracted 327.5 million domestic and international visitors, which marked the fifth consecutive year of more than 300 million visitors to national parks (National Park Service Office of Communications, 2020). A growing visitation may lead to negative impacts on ecosystem and biodiversity and challenge the sustainable utilization of national parks (Kerlinger et al., 2013; Pickering & Hill, 2007). In addition, visitor demographic attributes may influence their demands, such as engaged activities and travel costs, in national parks (Benson et al., 2013; Koontz et al., 2017; Le, 2012). Therefore, monitoring temporal visitation numbers and visitor demographics will be helpful for allocating resources, developing infrastructure, and predicting the demands of visitors (Cessford & Muhar, 2003; Pettebone & Meldrum, 2018; Rice et al., 2019). However, the traditional data collection approaches, such as survey or manual count, are costly, time-consuming, and laborious (Di Minin et al., 2015; Sessions et al., 2016).

With the rapid development of information and communication technology, big data with structured and unstructured types are generated, possessing '3V' characteristics, namely high volume, high value, and high velocity (Laney, 2001). Mobile data, one type of big data sources, has been applied to investigate human mobility at different geolocation levels (Lee et al., 2020), social distancing and inequality (Chiou & Tucker, 2020), factors influencing human mobility (Phithakkitnukoon et al., 2012), mobility and socio-economic indicators (Pappalardo et al., 2015), mobility and event detection (Traag et al., 2011), mobility and social networks (Chang et al., 2021). Also, mobile data has been gradually applied in tourism-related research (Raun et al., 2016; Kubo et al., 2020; Park et al., 2020; Rodríguez et al., 2018). For example, Ma and Kirilenko (2021) validated social media data, mobile data and traditional survey data in terms of tourists' origins of residency. By utilizing mobile positioning data, Raun, Ahas, and Tiru (2016) analyzed foreign visitors in Estonia at three dimensions, namely spatial temporal, and compositional. Therefore, mobile data provides rich information about spatial-temporal patterns and visitor characteristics, it has more potential to be employed in the tourism-related research.

SafeGraph is a commercial company to provide Point of Interest (POI) data and location-based Service data in the United States and Canada (Juhasz & Hochmair, 2020). POI refers to a specific point location that someone may find useful or interesting. The POI data of SafeGraph are compiled from several sources, including mobile phone GPS data and government open data, etc. The main product of SafeGraph data is SafeGraph Places, consisting of three datasets, namely Core Places, Patterns, and Geometry (SafeGraph, 2020). SafeGraph data has been utilized to understand human movement patterns (Chang et al., 2021; Juhasz & Hochmair, 2020; Kang et al., 2020). For example, Kang et al. (2020) utilized SafeGraph data to understand the daily and monthly mobility flow of human cross the United States and validated with other openly available data sources. In addition, SafeGraph data has been applied to explore social distancing response to COVID-19 (Chiou & Tucker, 2020)). In the tourism-related field, Juhasz and Hochmair (2020) investigated temporal visitation patterns, distance from home and event detections in three Florida cities by SafeGraph data.

Although previous studies have argued the effectiveness of using mobile data to investigate human mobility (Kang et al., 2020; Pappalardo et al., 2015; Traag et al., 2011), there are still three research gaps in the current literature. First and most important, previous studies lack the validation of mobile data with traditional survey/count data. To utilize mobile data as a data collection tool instead of or supplementing the traditional count/survey data, it is necessary to assess the similarities and differences between the two data sources. Secondly, numerous studies have investigated spatial-temporal patterns of visitors; limited research explore visitor demographics by mobile data (Ma & Kirilenko, 2021). Thirdly, most research focus on visitation patterns at a national or a city level; rare studies utilize mobile data in a national park context.

Therefore, to fill the research gaps of mobile data, in this study, we validate SafeGraph data, combining with American Community Survey (ACS), with traditional survey/count data in terms of visitor demographics and temporal visitation patterns in a national park context. Yellowstone National Park (YNP) is selected as the context because YNP was designated as the world's first national park and we can access the survey/count data of YNP regarding visitor demographics and temporal visitation patterns.

Methodology

Data

Five datasets were utilized for the data validation, including Yellowstone National Park Summer 2018 Visitor Use Surveys (2018), NPS Stats Recreation Visits by Month (2018-2020), Trails/Gates daily count data in Yellowstone National Park (2018 summer & 2019 summer), SafeGraph data (2018-2020), and American Community Survey (Census Tract level). Most datasets can be found online.

Yellowstone National Park Summer 2018 Visitor Use Surveys (National Park Service, 2019) was employed for assessing the validity of SafeGraph data in terms of visitor demographics, including age, gender, race, educational level, income level, and origins of residency. NPS Stats Recreation Visits by Month (National Park Services Stats, 2021) was utilized for validation of SafeGraph data regarding monthly visitation patterns. Trails/Gates Daily Count Data was direct from a social scientist of YNP, utilizing to validate daily visitation patterns with SafeGraph data.

This study retrieved SafeGraph data from Core Places and Patterns datasets. The Core Places dataset involves about 6.1 million POIs and related information, such as POI's address, category, NAICS CODE, open hours, brands, and unique SafeGraph ID. In addition, each POI has a geographic location (latitude & longitude). The Monthly Patterns dataset contains POIs with unique SafeGraph ID, raw visit counts (monthly), visits by day (daily), visitor home Census Block Groups (CBGs), etc. Currently, SafeGraph Patterns provides visitation data from 2018 January to 2020 November. The two datasets, Core Places and Monthly Patterns, can be linked by the same SafeGraph ID.

The American Community Survey (ACS) is a demographics survey program conducted by the U.S. Census Bureau. It regularly gathers information such as educational attainment, income, language proficiency, migration, disability, employment, and housing characteristics. ACS can be linked with SafeGraph data to extract visitor demographics distributions at Census Tract/Block Groups level. This study utilized ACS 2015-2019 (5-Year Estimates).

SafeGraph POIs Selection and Validation

To extract those POIs that are located within the boundary of Yellowstone National Park, we have performed a geospatial operation, namely point in polygon, where those POIs with coordinates that fall within the polygonal geometry of Yellowstone are kept.

To validate the locations of SafeGraph POIs, Google Maps was employed to assess the location name, latitude & longitude, and address for each POI. The location name of each POI was searched in Google Maps first. If the location name is accurate and can be found in Google Map, the geographic coordinates and detailed addresses can be gained from Google Map. Next, the geographic coordinate and detailed address of the POI gained from Google Map will be compared with latitude & longitude and address provided by SafeGraph. The 80 POIs from the selecting step were assessed through Google Maps manually. Finally, 40 invalid POIs were filtered because of unmatched location with geographic coordinates and inaccurate location names. Figure 1 presents 9 valid POIs in YNP. Next, unique SafeGraph IDs of valid 40 POIs in YNP were utilized to retrieve visitation patterns from SafeGraph Patterns dataset. The authors selected a sample of 9 valid POIs in YNP because of the most visitation among 40 POIs (Table 1).

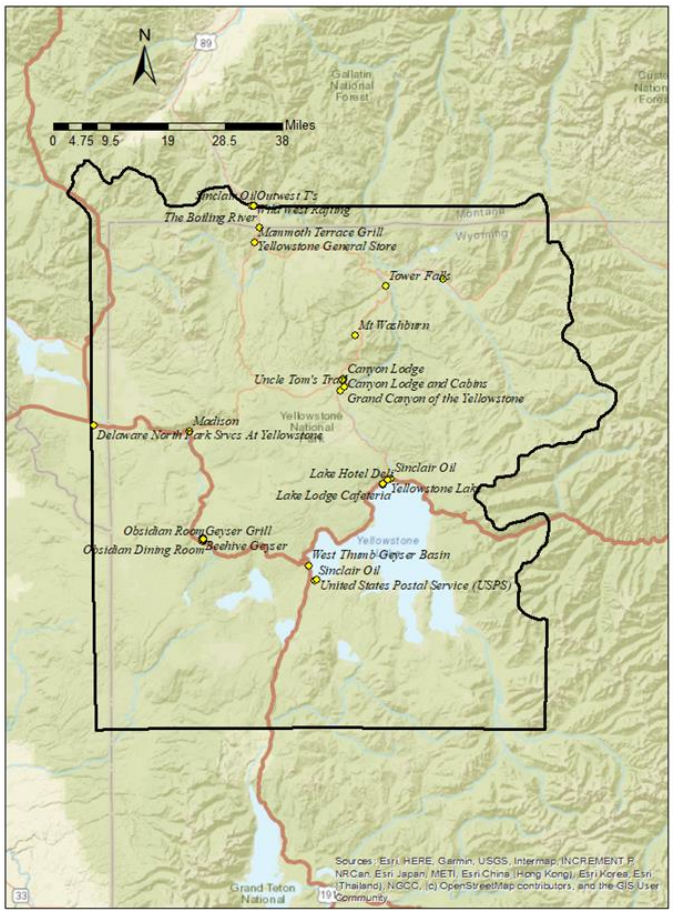

Figure 1. Valid POIs in Yellowstone National Park

Data Analyses

In the SafeGraph Monthly Patterns dataset, each POI provides visitor home CBGs, 12-digit numbers, tracing Census Block Group that visitors come from, and related visitation number. Each 12-digit number has such GEOID structure: 2+3+6+1 (12-digit numbers) = State + County + Tract + Block Group, which means the first two digit of the number represent the State that visitors come from, the first five (2+3) digit of the number indicate the County, the first eleven (2+3+1) represent the Census Tract. Therefore, based on the 12-digit numbers provided by the SafeGraph Monthly Patterns dataset, we can combine them with American Community Survey and extract the related demographic composition of residents at a State/County/Census Tract/Census Block Group level. Considering the workload, we extracted demographic distributions of residents at the Census Tract level from American Community Survey.

To calculate the percentage of each sub-group of each demographic variable of visitors for each POI in the Yellowstone National Park, let X_t (t = 1, 2, …, n) as each sub-group of each demographic variable in different Census Tract, let Y_t (t = 1, 2, …, n) as related visitation number in different Census Tract, therefore, the formula to calculate each sub-group of each demographic variable of a POI is:

$$\text{Percentage} = \Sigma_{t=1}^{n} X_t Y_t \Big/ \Sigma_{t=1}^{n} Y_t$$

Based on Yellowstone National Park Summer 2018 Visitor Use Surveys (National Park Service, 2019), five socioeconomic demographics, including gender, age, race, educational level, and income level, and origins of residency of visitors were collected. The 9 POIs (Table 1) were selected for investigating visitor demographics and residency of visitors based on the results of POI validation and the visitation numbers (the visitation numbers between 2018 May to 2018 September > 1000). After calculating visitor demographic distribution of each POI, the average of visitor demographic of the nine POIs was

Table 1. Selected 9 SafeGraph POI for Validating Visitor Demographics

Location Name	Top Category	Latitude	Longitude	Street Address	City	Region	Postal Code
Yellowstone Art & Photography Center	Other Professional, Scientific, and Technical Services	44.459	-110.827	2 Old Faithful Rd	Wyoming	WY	82190
Old Faithful General Store	Office Supplies, Stationery, and Gift Stores	44.457	-110.828	Old Faithful	Yellowstone National Park	WY	82190
Canyon Lodge Cafeteria	Restaurants and Other Eating Places	44.734	-110.491	1 Grand Loop Rd Canyon Village Yellowstone National	Canyon	WY	82190
Canyon Lodge and Cabins	Traveler Accommodation	44.734	-110.490	Canyon Village North Rim Dr	Yellowstone National Park	WY	82190
Old Faithful Observation Point	Museums, Historical Sites, and Similar Institutions	44.460	-110.828	Old Faithful Village	Yellowstone National Park	WY	82190
Obsidian Room	Restaurants and Other Eating Places	44.457	-110.830	Old Faithful Snow Lodge	Yellowstone National Park	WY	82190
Geyser Grill	Restaurants and Other Eating Places	44.457	-110.829	1000 Old Faithful	Yellowstone National Park	WY	82190
Fishing Bridge General Store	Office Supplies, Stationery, and Gift Stores	44.565	-110.375	1 NE Entrance	Yellowstone National Park	WY	82190
Outwest T's	Clothing Stores	45.030	-110.707	228 Park St	Gardiner	MT	59030

compared with visitor demographics from Yellowstone National Park Summer 2018 Visitor Use Surveys (National Park Service, 2019) by Chi-square test.

To validate the monthly visitation patterns of POIs in YNP, SafeGraph data were compared with the count data from NPS Stats Recreation Visits by Month (National Park Services Stats, 2021). Pearson's r correlation between the two normalized data on the visitation number of each month was used to estimate the match between the two data sources (Tenkanen et al., 2017). In a similar way, to validate the daily visitation patterns of POIs in YNP, SafeGraph data was compared with the count data of daily visitation count for different trails. Pearson's r correlation between the two data with normalization on the visitation number of each day was used to estimate the match between the two data sources (Tenkanen et al., 2017).

Results

Visitor Demographics

Table 2 presents the average visitor demographics of the nine POIs and the results of Chi-square tests with Yellowstone National Park Summer 2018 Visitor Use Surveys. In addition, numbers of origins of residency of visitors for nine POI and the results of comparison of the survey were presented in Table 3. We provided demographics at the national level as a reference line.

Table 2 Comparison of visitor demographics between SafeGraph data (combining with ACS) and Yellowstone National Park Summer 2018 Visitor Use Surveys

Demographics	2018 Visitor Use Study	SafeGraph 9 POIs Average	Chi Square Statistics	P value	National Statistics
Gender					
Male	51%	49.6%	0.39	0.53	49.2%
Female	49%	50.3%			50.8%
Age					
Under 5 Years	NA		NA		
5 to 9 Years	NA		NA		
10 to 14 Years	NA		NA		
15 to 17 Years	NA		NA		
18 to 24 Years	12%	10.7%	0.84	0.36	12.1%
25 to 34 Years	21%	15.9%	8.64	0.0033	18.0%
35 to 44 Years	16%	16.6%	0.13	0.72	16.3%
45 to 54 Years	17%	17.1%	0.0035	0.95	16.8%
55 to 64 Years	17%	17.9%	0.28	0.60	16.7%
65 to 74 Years	14%	13.2%	0.27	0.60	11.8%

75 or older	3%	8.6%	28.70	<0.001	8.4%
Race					
White	82%	82.5%	0.086	0.77	72.5%
Black or African American	0%	4.8%	NA		12.7%
American Indian or Alaska Native	1%	0.9%	0.053	0.82	0.9%
Asian	17%	5.2%	70.55	<0.001	5.5%
Native Hawaiian or Pacific Islander	0%	0.1%	NA		0.2%
Other Race	NA	2.2%	NA		4.9%
Two More Race	NA	2.9%	NA		3.3%
Highest level of former education					
Less than High School	1%	7.0%	4.688	0.0304	12.0%
High School Graduate	8%		Census Tract data and Survey data have different categories in Education Level		
Some College	10%				
Bachelors degree	37%				
Advanced degree	35%				
Vocational/Trade School	2%				
Two-year college degree	7%				
Household Income Level					
Less than $25,000	9%	14.2%	13.18	<0.001	19.3%
$25,000 to $49,999	12%	17.8%	13.27	<0.001	21.3%
$50,000 to $74,999	18%	16.6%	0.69	0.41	17.2%
$75,000 to $99,999	17%	13.3%	5.32	0.021	12.7%
$100,000 to $ 149,999	21%	17.8%	3.27	0.070	15.1%
$150,000 to $199,999	12%	8.7%	5.87	0.015	6.8%
$200,000 or more	12%	11.6%	0.077	0.78	7.7%
df=1					

As the results, there was no difference between the average visitor demographics of POIs and 2018 Visitor Use Study regarding gender. Two age groups were found differences: 15.9% (POI) vs. 21.0% (Survey) of 25~34-year-old ($p < 0.01$) and 8.6% (POI) vs. 3.0% (Survey) of 75 or older ($p < 0.001$). We found a statistically significant difference between average visitor demographics of POIs and 2018 Visitor Use Study regarding Asian, 5.2% (POI) vs. 17.0% (Survey) ($p < 0.001$). 2018 Visitor Use Study did not survey African American and Native Hawaiian or Pacific Islander, while, according to average visitor demographics of the nine POIs, there were 4.8% of African American and 0.1% of Native Hawaiian or Pacific Islander. Furthermore, there was a statistically significant difference of visitors with High School Diploma between the two datasets (7.0% (POI) vs. 1.0% (Survey), $p < 0.05$). Only three income level groups were no differences between the average visitor demographics of POIs and 2018 Visitor Use Study. There were more percentages of less than $25,000 (14.2% (POI) vs. 9.0% (Survey), $p < 0.001$) and $25,000 to $49,999 (17.8% (POI) vs. 12.0% (Survey), $p < 0.001$) of visitors in average visitor demographics of POIs, while less percentages of $75,000 to $99,999 (13.3% (POI) vs. 17.0% (Survey), $p < 0.05$) and $150,000 to $199,999 (8.7% (POI) vs. 12.0% (Survey), $p < 0.05$) of visitors in average visitor demographics of POIs.

Top 10 of States of visitor residency and related frequencies of SafeGraph data and the survey was presented in Table 3. There were slights differences between the ranks of SafeGraph data and the survey data. For example, WY only ranked Top 10 based on SafeGraph data, while NY only ranked Top 10 based on the survey data. In addition, the ranks of States based on the two datasets have different orders.

Table 3. Comparison of origins of residency of visitors between SafeGraph data (combining with ACS) and Yellowstone National Park Summer 2018 Visitor Use Surveys

Total of 9 POIs			2018 Visitor Use Survey		
State	Visitation	Percentage	State	Frequency	Percentage
CA	1925	9.4%	CA	248	13%
WY	1263	6.2%	TX	98	5%
TX	1138	5.6%	FL	84	4%
MT	1008	4.9%	WA	77	4%
UT	908	4.4%	UT	71	4%
FL	901	4.4%	CO	70	4%
ID	844	4.1%	NY	69	4%
WA	670	3.3%	MN	63	3%
CO	617	3.0%	ID	63	3%
MN	608	3.0%	MT	56	3%

Monthly Visitation Patterns

Two POIs, Yellowstone Art & Photography Center and Old Faithful General Store, from SafeGraph data were selected since the two POIs are valid based on the results of SafeGraph POI validation and are the most visited POIs among valid POIs in YNP in 2018 to 2020.

Both Yellowstone Art & Photography Center and Old Faithful General Store showed similar peak periods with official visitation statistics; however, the patterns of the peak seasons of the POIs do not have similar shapes with the official visitation statistics. Figure 2 showed visitation trends through 2018 January to 2020 October for official visitation statistics, Yellowstone Arts & Photography Center data, and Old Faithful General Store data.

Pearson's r correlation was conducted for the visitation numbers of each POI and official visitation statistics (Figure 3). The correlation coefficient of Yellowstone Art & Photography Center and official visitation statistics is 0.88 ($p < 0.001$) and the correlation coefficient of Old Faithful General Store and official visitation statistics is 0.85 ($p < 0.001$), indicating that the SafeGraph data and official statistics have a strong relationship (The Odum Institute, 2015) and can match well at monthly level.

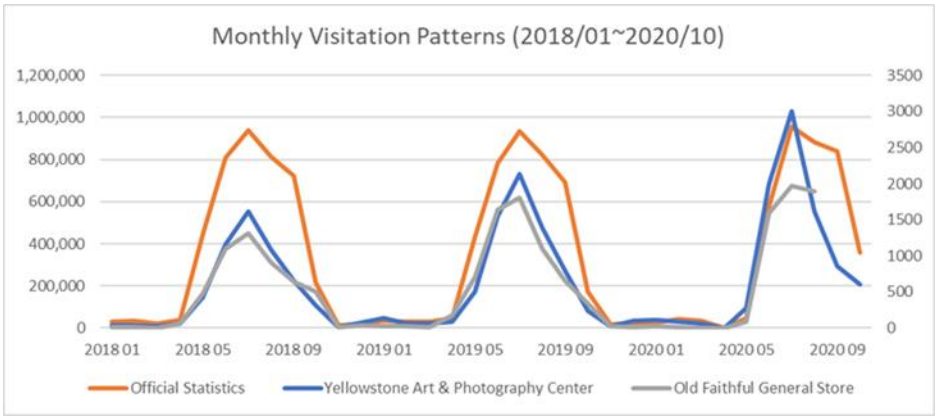

Figure 2. Monthly visitation patterns between two datasets

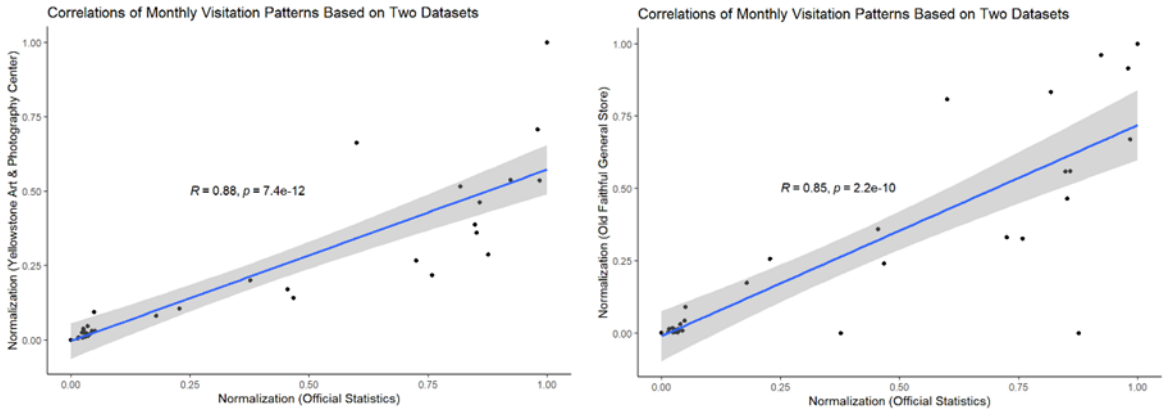

Figure 3. Correlations of monthly visitation patterns between two datasets

Daily Visitation Patterns

Old Faithful Spring Trail vs. Old Faithful General Store

Old Faithful Spring was selected to validate daily visitation patterns between SafeGraph data and official visitation counts based on data availability. The POI, Old Faithful General Store, was

selected for comparison because this POI is close to Old Faithful Spring Trail. The daily visitation patterns of official visitation counts and SafeGraph data were presented in Figure 4. In addition, Pearson's r correlation was conducted for the Old Faithful Spring Trail count data and SafeGraph data. The correlation coefficients were 0.76 ($p < 0.001$) and 0.77 ($p < 0.001$) in 2018 and 2019 separately (Figure 4).

Gates Vehicle Count vs. Yellowstone Arts & Photography Center

The POI, Yellowstone Arts & Photograph Center, was selected to validate daily visitation patterns with gates vehicle count data because this location was the most visited POI in SafeGraph data in YNP and gates vehicle count data can be employed to estimate visits at the entire park level. The daily visitation patterns of official visitation counts and SafeGraph data were presented in Figure 5. Pearson's r correlation was conducted for the Gates Vehicle daily count data and SafeGraph data. The correlation coefficients were 0.79 ($p < 0.001$) and 0.47 ($p < 0.001$) in 2019 and 2020 separately (Figure 5).

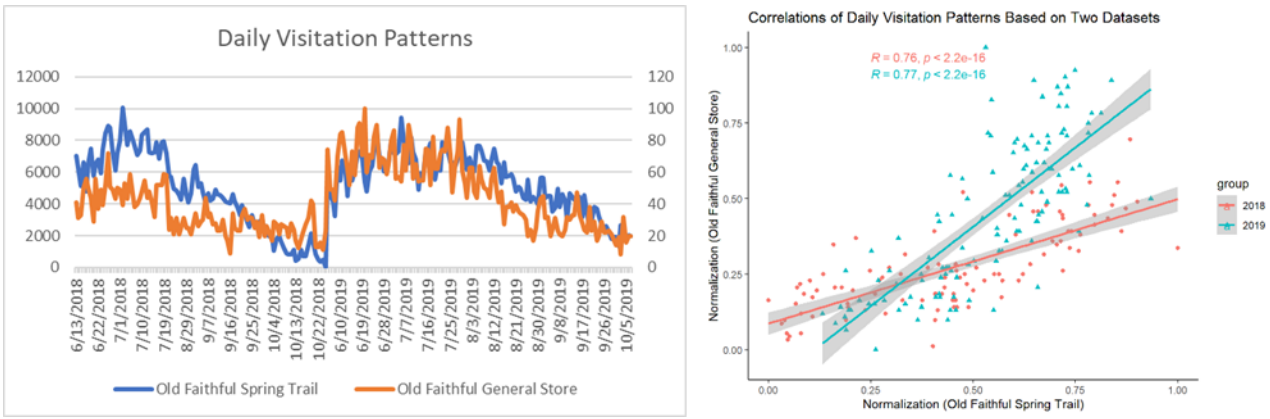

Figure 4. Daily visitation patterns and Correlation between Old Faithful Spring Trail and Old Faithful General Store

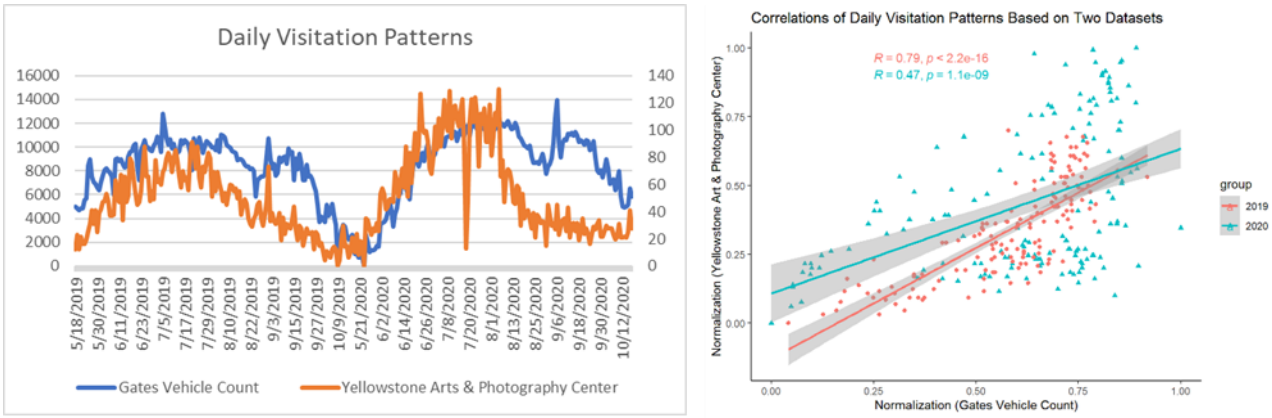

Figure 5. Daily visitation patterns and Correlation between Gates Vehicle Count and Yellowstone Arts & Photography Center

Conclusion and Discussion

This study compares mobile data and visitor use study/count data, validating visitor demographics and temporal visitation patterns in Yellowstone National Park. The similarities and differences of

the two data sources regarding visitor demographics and temporal visitation patterns suggests that SafeGraph data can serve as an additional and complementary source of information to traditional visitor use study and count data. With the potential technology advancement, mobile data may provide more accurate, comprehensive, and timely information related to visitor demographics and temporal visitation patterns for each POI. In addition, future research in national parks could consider the influence of visitor demographic variables and temporal visitation patterns when using mobile data.

Reference

Benson, C., Watson, P., Taylor, G., Cook, P., & Hollenhorst, S. (2013). Who visits a national park and what do they get out of it?: A joint visitor cluster analysis and travel cost model for Yellowstone National Park. *Environmental Management*, 52(4), 917–928. https://doi.org/10.1007/s00267-013-0143-4

Cessford, G. and Muhar, A. 2003. Monitoring Options for Visitor Numbers in National Parks and Natural Areas. Journal of Natural Conservation. 11: 240–250.

Chang, S., Pierson, E., Koh, P. W., Gerardin, J., Redbird, B., Grusky, D., & Leskovec, J. (2021). Mobility network models of COVID-19 explain inequities and inform reopening. *Nature*, 589(7840), 82–87. https://doi.org/10.1038/s41586-020-2923-3

Chiou, L., & Tucker, C. (2020). Social distancing, Internet access and inequality (No. w26982). *National Bureau of Economic Researc*h. https://doi.org/10.3386/w26982

Di Minin, E., Tenkanen, H., & Toivonen, T. (2015). Prospects and challenges for social media data in conservation science. *Frontiers in Environmental Science*, 3. https://doi.org/10.3389/fenvs.2015.00063

Juhasz, L., & Hochmair, H. (2020). Studying Spatial and Temporal Visitation Patterns of Points of Interest Using SafeGraph Data in Florida. GIS Center. https://doi.org/10.1553/giscience2020_01_s119

Kang, Y., Gao, S., Liang, Y., Li, M., Rao, J., & Kruse, J. (2020). Multiscale dynamic human mobility flow dataset in the U.S. during the COVID-19 epidemic. *Scientific Data*, 7(1), 390. https://doi.org/10.1038/s41597-020-00734-5

Kerlinger, P., Burger, J., Cordell, H. K., Decker, D. J., Cole, D. N., Landres, P., O'Shea, T. (2013). *Wildlife and recreationists: Coexistence through management and research*. Washington, DC: Island.

Koontz, L., Thomas, C. C., Ziesler, P., Olson, J., & Meldrum, B. (2017). Visitor spending effects: Assessing and showcasing America's investment in national parks. *Journal of Sustainable Tourism*, 25(12), 1865–1876. https://doi.org/10.1080/09669582.2017.1374600

Kubo, T., Uryu, S., Yamano, H., Tsuge, T., Yamakita, T., & Shirayama, Y. (2020). Mobile phone network data reveal nationwide economic value of coastal tourism under climate change. *Tourism Management*, 77, 104010. https://doi.org/10.1016/j.tourman.2019.104010

Laney, D. (2001). 3D data management: Controlling data volume, velocity and variety. *META Group Research Note*, 6, 70.

Le, L. (2012). Hispanic and white visitors in U.S. national parks: Meta-analysis of visitor use survey. *Journal of Park and Recreation Administration*, 30(4). http://search.proquest.com/docview/1730142366/abstract/DE8E1004056D4B55PQ/1

Lee, M., Zhao, J., Sun, Q., Pan, Y., Zhou, W., Xiong, C., & Zhang, L. (2020). Human mobility trends during the early stage of the COVID-19 pandemic in the United States. *PLOS ONE*, 15(11), e0241468. https://doi.org/10.1371/journal.pone.0241468

Ma, S., & Kirilenko, A. (2021). How reliable is social media data? Validation of TripAdvisor tourism visitations using independent data sources. In W. Wörndl, C. Koo, & J. L. Stienmetz (Eds.), *Information and Communication Technologies in Tourism 2021* (pp. 286–293). Springer International Publishing. https://doi.org/10.1007/978-3-030-65785-7_26

Monz, C., Mitrovich, M., D'Antonio, A., & Sisneros-Kidd, A. (2019). Using mobile device data to estimate visitation in parks and protected areas: An example from the nature reserve of Orange County, California. *Journal of Park and Recreation Administration*, 37(4), Article 4. https://doi.org/10.18666/JPRA-2019-9899

National Park Service (2019, November 07). Yellowstone National Park Summer 2018 Visitor Use Surveys. https://www.nps.gov/yell/learn/management/upload/2018-Yellowstone-Visitor-Use-Surveys-FINAL-REPORT_WEB-RESOLUTION.pdf

National Park Service Office of Communications. (2020, February 27). *National Park Visitation Tops 327 Million in 2019*. https://www.nps.gov/orgs/1207/2019-visitation-numbers.htm#:~:text=WASHINGTON%20%E2%80%93%20America's%20national%20parks%20continue,record%20keeping%20began%20in%201904

National Park Service Stats. (2021, February 27). Recreation Visits by Month. https://irma.nps.gov/STATS/SSRSReports/Park%20Specific%20Reports/Recreation%20Visitors%20By%20Month%20(1979%20-%20Last%20Calendar%20Year)?Park=YELL

Pappalardo, L., Pedreschi, D., Smoreda, Z., & Giannotti, F. (2015). Using big data to study the link between human mobility and socio-economic development. *2015 IEEE International Conference on Big Data (Big Data)*, Santa Clara, CA, USA, 2015, pp. 871-878, doi: 10.1109/BigData.2015.7363835.

Park, S., Xu, Y., Jiang, L., Chen, Z., & Huang, S. (2020). Spatial structures of tourism destinations: A trajectory data mining approach leveraging mobile big data. *Annals of Tourism Research*, 84, 102973. https://doi.org/10.1016/j.annals.2020.102973

Pettebone, D., & Meldrum, B. (2018). The need for a comprehensive socioeconomic research program for the National Park Service. *The George Wright Forum*, 35(1), 22–31.

Phithakkitnukoon, S., Smoreda, Z., & Olivier, P. (2012). Socio-Geography of human mobility: A study using longitudinal mobile phone data. *PLOS ONE*, 7(6), e39253. https://doi.org/10.1371/journal.pone.0039253

Pickering, C. M., & Hill, W. (2007). Impacts of recreation and tourism on plant biodiversity and vegetation in protected areas in Australia. *Journal of Environmental Management*, 85(4), 791–800. https://doi.org/10.1016/j.jenvman.2006.11.021

Raun, J., Ahas, R., & Tiru, M. (2016). Measuring tourism destinations using mobile tracking data. *Tourism Management*, 57, 202–212. https://doi.org/10.1016/j.tourman.2016.06.006

Rice, W. L., & Pan, B. (2020). Understanding drivers of change in park visitation during the COVID-19 pandemic: A spatial application of Big data. *SocArXiv*. https://doi.org/10.31235/osf.io/97qa4

Rice, W. L., Park, S. Y., Pan, B., & Newman, P. (2019). Forecasting campground demand in US national parks. *Annals of Tourism Research*, 75, 424–438. https://doi.org/10.1016/j.annals.2019.01.013

Rodríguez, J., Semanjski, I., Gautama, S., Van de Weghe, N., & Ochoa, D. (2018). Unsupervised hierarchical clustering approach for tourism market segmentation based on crowdsourced mobile phone data. *Sensors*, 18(9), 2972. https://doi.org/10.3390/s18092972

SafeGraph. (2020, December 04). *Places Manual*. https://docs.safegraph.com/docs/places-manual

Sessions, C., Wood, S. A., Rabotyagov, S., & Fisher, D. M. (2016). Measuring recreational visitation at U.S. National Parks with crowd-sourced photographs. *Journal of Environmental Management*, 183, 703–711. https://doi.org/10.1016/j.jenvman.2016.09.018

Tenkanen, H., Di Minin, E., Heikinheimo, V., Hausmann, A., Herbst, M., Kajala, L., & Toivonen, T. (2017). Instagram, Flickr, or Twitter: Assessing the usability of social media data for visitor monitoring in protected areas. *Scientific Reports*, 7(1), 17615. https://doi.org/10.1038/s41598-017-18007-4

The Odum Institute. (2015). Learn about pearson's correlation coefficient in SPSS with data from the global health observatory data (2012). In SAGE Research Methods Datasets Part 1. SAGE Publications, Ltd. https://www.doi.org/10.4135/9781473948167

Traag, V. A., Browet, A., Calabrese, F., & Morlot, F. (2011). Social Event Detection in Massive Mobile Phone Data Using Probabilistic Location Inference. *2011 IEEE Third International Conference on Privacy, Security, Risk and Trust and 2011 IEEE Third International Conference on Social Computing*, Boston, MA, USA, 2011, pp. 625-628, doi: 10.1109/PASSAT/SocialCom.2011.