# Where to catch 'em all? – A geographic analysis of Pokémon Go locations

Levente Juhász
University of Florida
3205 College Ave
Ft, Lauderdale, FL, USA
levente.juhasz@ufl.edu

Hartwig H. Hochmair
University of Florida
3205 College Ave
Ft. Lauderdale, FL, USA
hhhochmair@ufl.edu

Following is a summary of a paper submitted to the VGI-Analytics special issue of Geo-spatial Information Science. The study was also presented at the VGI-Analytics pre-conference workshop at AGILE 2017 in Wageningen, The Netherlands, on May 9, 2017.

## 1 Introduction and research questions

Pokémon Go, released in July 2016, is a location based augmented reality game that quickly became popular. The game allows anyone with a smartphone to collect virtual characters (Pokémons) on the screen which appear to be positioned at the same location as the player. To do so, the player needs to navigate to certain points on a map. These points are often placed at prominent places, such as landmark buildings or statues. Pokémon Go users can directly interact with three different point datasets in the game, which are PokéStops, Pokémon gyms and spawnpoins. PokéStops are points in the geographic space, associated with a landmark, such as buildings or monuments. Players need to visit these locations and perform some actions, such as flipping a coin. In return, players are rewarded with items, such as a Pokéball, which they need to capture Pokémons later. Pokémon gyms are virtual locations that are also associated with landmarks and where Pokémons can be trained. Pokémons can "pop-up" at distinct locations (spawnpoints) during the game for thirty minutes. During this time-window, users have the chance to visit these locations and catch the specific Pokémon, which, upon capturing, will appear in the inventory of the player.

Pokémon Go was in the center of media attention for an extended period of time with several media articles reporting about the game. Residents of certain neighborhoods (often with significant number of minorities) reported that they felt that there was a smaller number of PokéStops in their areas than in other areas of the same city. The lack of PokéStops in these disadvantaged neighborhoods eventually prevented these residents from effectively participating in the game (Akhtar, 2016). Figure 1 illustrates this bias by comparing Pokémon Go related points between a metropolitan downtown area (a) and a neighborhood with a high number of Hispanic population (b).

Pokémon Go was found to motivate people to be physically more active and to go outside. Besides having positive effects on people's health (Nigg et al., 2017, Xu et al., 2017), Pokémon Go encourages players to explore new areas, which can lead to the identification of new species (Nature, 2016). A related example is the recent detection of a new species of the pygmy devil grasshopper through a photo on Facebook (Skejo and Caballero, 2016)

There is also evidence for other services and local businesses trying to utilize the popularity of Pokémon Go by organizing Pokémon related special offers or developing some functionalities in their sites that would be of interest to players. For example, Yelp, a location based service that allows users to rate, review and browse businesses, such as restaurants and stores, added a functionality to its smartphone apps and website where users can mark businesses with a "PokéSstop nearby" attribute. This attribute was incorporated into the Yelp search functionality so that users could find businesses in the proximity of PokéStops and therefore combine two leisure activities, which are Pokémon hunting and eating out in a restaurant.

Based on these recent developments the following research objectives can be formulated:

- R1: For PokéStops, identify the effect of socio-economic factors on the number of PokéStop counts in a US census block group
- R2: For PokéStops, Pokémon gym locations and spawnpoints, quantify the point distributions across different land use categories
- R3: For PokéStop and gym locations, compare their spatial clustering patterns
- R4: Determine the quality of the "PokéStop nearby" attribute in Yelp

To answer these research questions we extracted the locations of over 600 Pokémon gyms, 5,000 PokéStops, and 18,000 spawnpoints within South Florida and in Boston.

## 2 Results

### 2.1 R1 Relationship between PokéStop counts and neighborhood variables

A negative binomial regression model (NBM) that relates PokéStop counts in census block groups with ethnicity, race, and land use based factors was developed. The number of Panoramio photos and the number of Yelp businesses were also included in the analysis as proxies for tourist and economic activities, respectively.

Results show that areas with more business and touristic opportunities and as well as the presence of parks and higher education institutions are positively associated with the number of PokéStops within a neighborhood. As opposed to this, results indicate that areas with a higher percentage of African American and Hispanic population are disadvantaged in terms of PokéStop coverage.

## 2.3 R3 Spatial clustering of Pokémon Go point datasets

Pokémon Go related point sets are not evenly distributed. However, visual inspection suggests that PokéStops and Pokémon Gyms tend to cluster similarly (see Figure 1). To confirm clustering, nearest neighbor distances from each gym to its nearest Pokéstop were calculated. These distance distributions were compared to another set of nearest neighbor distances measured between gyms and a set of random points that had the same number as the number of Pokéstops for each study site. Comparison between the two nearest neighbor distance distributions suggests that Pokéstops are closer to gyms than expected under complete spatial randomness.

Figure 1: Pokémon Go related point datasets in Downtown Miami (a) and in Hialeah, FL (b).


## 2.2 R2 Counts of Pokémon Go points on land use categories

In addition to aggregating PokéStops, Pokémon gyms and spawnpoints on land use categories, an expected count number was determined based on the area covered by each land use category. Results showed that land use categories are significantly associated with point counts. That is, commercial, public and recreational areas have a higher point density than other land use classes. This means that these three land use categories hold more opportunities to participate in Pokémon Go than would be expected under complete spatial randomness. Results show also that Pokémon Go points are underrepresented in agricultural, industrial, natural, residential, and water land use categories in both cities.

To determine whether Gyms and PokéStops are similarly clustered or not, the bivariate version of Ripley's K-function (Cross K-function) was used (Dixon, 2002). Statistical inference of the difference between the observed Cross-K function and a Cross-K function generated by random labeling can be achieved through Monte Carlo simulation. We analyzed the cluster behavior of gyms around PokéStops, using gyms as controls and PokéStops as cases. Within each of the 999 permutations of the Monte Carlo simulation, events were randomly labeled as either PokéStop or gym (while retaining their observed proportions) and the Cross-K function was calculated. This established and upper and lower simulation envelope for random labeling at a 99.9% confidence level.

Results show that the observed Cross-K functions fall within the simulation envelope for the entire distance range for most study sites. This implies that PokéStops and gyms are similarly clustered around each other. There are, however a few study sites where the observed Cross-K function falls slightly below the lower simulation envelope. In these areas, gyms and PokéStops are further apart from each other than excepted under random labeling.

### 2.4 R4 Pokémon related user tagging of Yelp businesses

In our study areas throughout South Florida and in Boston, MA, Yelp users tagged 1,392 businesses with the "PokéStop nearby" attribute out of 21,606 total businesses. To determine whether user tagging of businesses is correct or not, we computed two sets of nearest neighbor distances, namely those measured between businesses tagged with the PokéStop attribute and their nearest PokéStops, and between all Yelp businesses and their nearest PokéStops.

It was found that businesses tagged with the "PokéStop nearby" attribute are indeed situated closer to actual PokéStops than all Yelp businesses, indicating that Yelp users tend to use this tag correctly. A strong geographically linked bias between urbanized and rural areas can also be observed, as tagging is more complete in highly urbanized areas. This can also be the result of the lower density of PokéStops in rural areas or the lack of active crowdsourcing communities in rural areas.

## 3 Summary and conclusions

This paper analyzed point datasets extracted from Pokémon Go, which is a location based augmented reality game for smartphones that engages millions of users worldwide The study confirms the anecdotal experience of players in certain neighborhoods who reported a lack of Pokémon Go point features compared to other neighborhoods. Whereas earlier research already identified that areas with a high percentage of minorities have limited access to PokéStops (Colley et al., 2017) the estimated regression model of this paper presents an extended list of factors which were found to be significantly associated with an increase in PokéStop counts per block group, including presence of parks, higher education institutes, or businesses, or being tourist sites.

Further analysis showed that some land use categories have an overrepresentation of Pokémon Go features. Users will have the highest chance to encounter a Pokémon, or to run into a PokéStop or gym if playing Pokémon Go in commercial areas (e.g. shopping centers), public spaces (e.g. university campuses) or in recreational areas (e.g. parks). As opposed to this, natural, agricultural, residential, and industrial areas, as well as lakes, rivers, and open water have a lower point density and make it therefore more difficult to succeed in this game. Using a Cross K-function, Pokémon gym locations were in most study areas found to cluster similarly to PokéStops, suggesting that those locations were generated with the same methods and derived from the same crowdsourced dataset.

Our study analyzed also the interplay of augmented reality gaming and VGI, suggesting that the Pokémon Go user community participates in crowd-sourcing activities, namely adding information to the "PokéStop nearby" attribute on the Yelp business platform. Nearest neighbor analysis suggests that this tagged information tends to be correct, and can be used by visitors of the Yelp Web site to identify businesses that are located near a PokéStop.

The presented research supports earlier findings of a strong geographic and socio-economic bias in the Pokémon Go dataset. As this bias can affect user experience in location based games negatively, future developments and improvements of location based augmented reality games should address this issue and provide equal access to interactive platforms, such as Pokémon Go, to all user communities.

## References


Akhtar, A. 2016. *Is Pokémon Go racist? How the app may be redlining communities of color* [Online]. USA TODAY. Available: http://www.usatoday.com/story/tech/news/2016/08/09/pokemon-go-racist-app-redlining-communities-color-racist-pokestops-gyms/87732734/.

Colley, A., Thebault-Spieker, J., Lin, A. Y., Degraen, D., Fischman, B., Häkkilä, J., Kuehl, K., Nisi, V., Nunes, N. J., Wenig, N., Wenig, D., Hecht, B. and Schöning, J. (2017) The Geography of Pokémon GO: Beneficial and Problematic Effects on Places and Movement. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, Denver. pp., 2017.

Dixon, P. M. (2002) Ripley's K function. *Encyclopedia of Environmetrics* Wiley Online Library, pp. 1796–1803.

Nature (2016) Gotta name them all: how Pokémon can transform taxonomy (Editorial). *Nature,* (535)**,** 323-324.

Nigg, C. R., Mateo, D. J. and An, J. (2017) Pokémon Go may increase physical activity and decrease sedentary behaviors. *American Journal of Public Health,* 107(1)**,** 37-38.

Skejo, J. and Caballero, J. H. S. (2016) A hidden pygmy devil from the Philippines: *Arulenus miae* sp. nov.—a new species serendipitously discovered in an amateur Facebook post (Tetrigidae: Discotettiginae). *Zootaxa,* 4067(3)**,** 383-393.

Xu, H., Xian, Y., Xu, H., Liang, L., Hernandez, A. F., Wang, T. Y. and Peterson, E. D. (2017) Does Pokemon Go Help Players be More Active? An Evaluation of Pokemon Go and Physical Activity. *Circulation,* 135.