



Identifying training venues for delivering HIV prevention programs

STEPHANIE POSNER

CAPSTONE PROJECT

JUNE 2021

Background

High HIV infection rates on South Africa

- New HIV infections are highest among adolescent girls and young women ages 15-30
- Highest in City of Johannesburg

We run an HIV behavior change prevention program

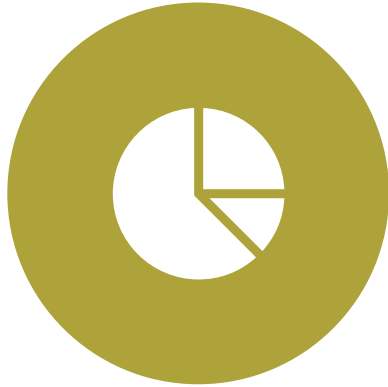
- 14-week program, conducted in 45-min sessions that are held once a week.
- We recruit girls and young women ages 15-30 through government social service sites

Problem Statement

To date, girls and young women of the target age group (15-30) must be recruited to participate in the program and come to the site for participation. Participation dwindles after a few weeks.

To improve recruitment, retention, and completion in the program, we would like to explore conducting these program sessions at venues that are popular with youth, where youth choose to go to hang out, dialogue and meet friends and new people. The logic is that this approach will reach more youth, bring like-minded youth together in different settings, engage the youth on difficult topics right in the venues they frequent, and ultimately increase retention and completion of the program. At the same time, these venues are nearby to additional social services that youth can access for more private sessions.

This requires analysing what and where are viable venues for conducting the training program.



THE PURPOSE:

To identify suitable venues for training that are walking distance from government-funded social services



THE AUDIENCE:

Used by the program staff for planning and implementation



THE USE:

Analyses will produce a clustered list of venues to approach for engagement and participation in health program

Data

Data Source 1:

Location data on Government funded social service sites in City of Johannesburg in South Africa..

Data Source 2:

Venue, address, and category data from the Foursquare API using the version on April 20, 2021

Methodology

Venue data were extracted from Foursquare that were within 1/2 kilometer (i.e. walking distance) of the program sites.

Venues categories were consolidated into one category where they appeared similar, e.g. “café” and Coffee shop). One-hot coding was used to associate all sites with venue types. Sites listed the aggregate number of venues of each category.

Clustering analysis was performed using K-means. Because clustering wasn’t looking for the most efficient clustering, but for an eligible pool of venues, groups were expanded to 6 groups.

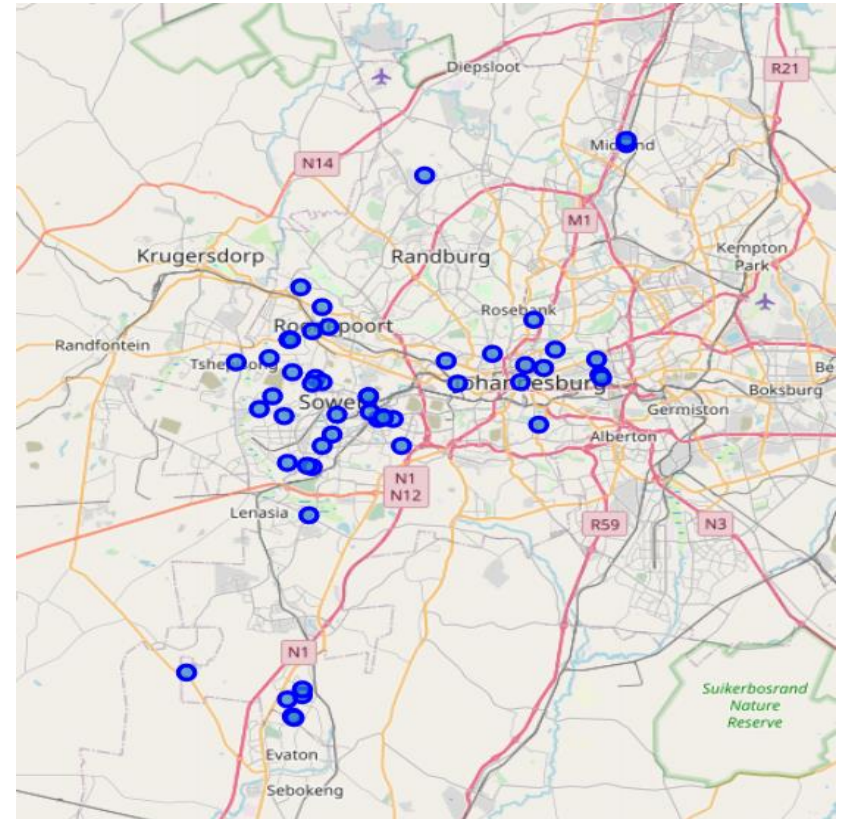
Maps were performed in Folium.

Results

Of 481 sites in South Africa, 435 were geocoded.

Of those, 54 were in Johannesburg – seen in the map to the right.

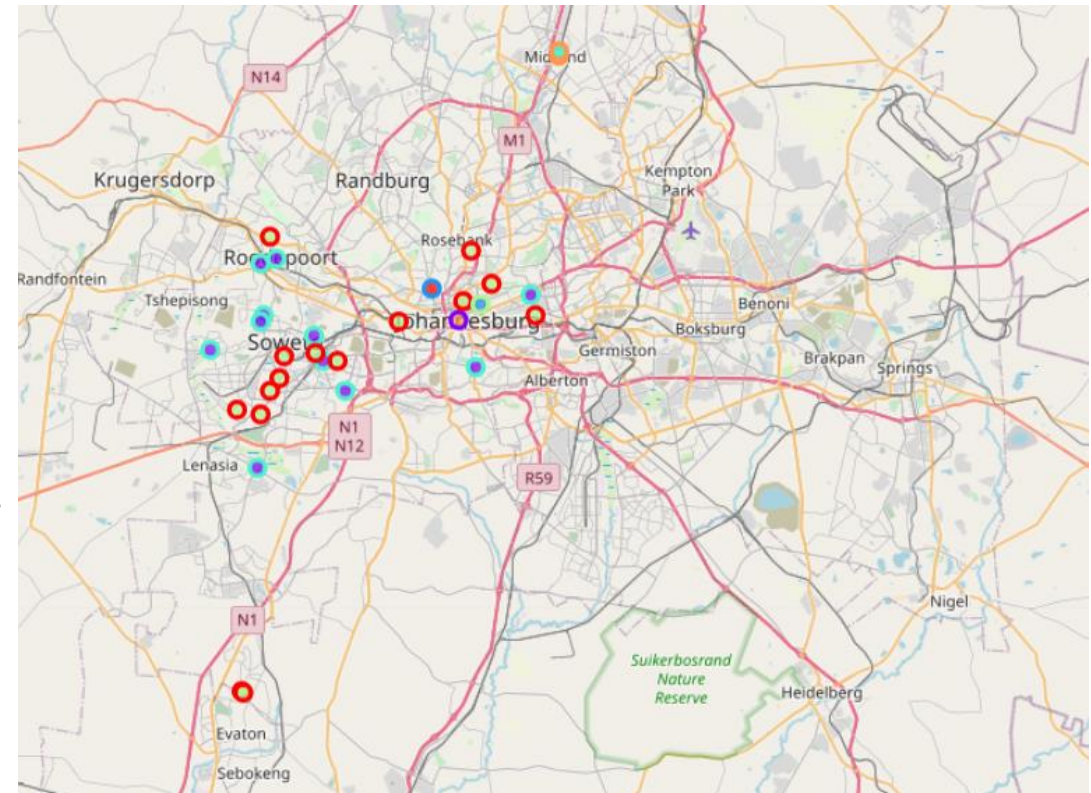
Foursquare returned 117 venues within 500 meters walking from 32 sites. These venues had 44 venue categories.



Results: Clustering

Cluster Labels	1st Most Common Venue
0	Casual Restaurant
1	African Restaurant
	Bar
	Construction & Landscaping
	Convenience Store
	Furniture / Home Store
	Golf Course
	Historic Site
	Home Service
	Italian Restaurant
	Print Shop
	Train Station
	Water Park
2	Coffee Shop

Cluster Labels	1st Most Common Venue
3	Convenience Store
4	Casual Restaurant
	Concert Hall
	Construction & Landscaping
	Grocery Store
	Home Service
	Shopping Mall
5	Casual Restaurant



Discussion

Cluster interpretations:

- Clusters around Midrand and Sayac sites (Clusters 0 and 5) are remote but do have causal dining sites (mostly fast food restaurants) within walking distance. These are suitable sites to cover clients in these three sites.
- 16 sites, most in Soweto (Cluster 1 is near recreational facilities, e.g. water park and golf, which may be suitable venues.
- Alexandria township (Cluster 2), is a large site and can be served by a number of cafes as training sites.
- Guild Cottages (Cluster 3) is in a more up market area with a hotel, although this might be pricier as a training venue.
- Cluster 4 has 11 sites, all of which appear to be served by causal restaurants.

Limitations:

- These data are from all users, not females 15-30. Foursquare does not have age groups similar to the training program.
- Some sites are too remote for venues (either no venues within half of kilometer or none recorded on FourSquare)
- Some sites are in poorer areas where youth may not own a smart phone and thus Foursquare data logging is limited.

Conclusion

This analysis has revealed a number of options that can be pursued as training sites for the social training program for girls and young women ages 15-30.

The clustering reveals that a slightly different approach may be needed for clusters with different types of venues, particularly those with less youth-friendly venues.

Staff can use these results to explore those venues to negotiate participation as a training site with the goal of increasing participation and retention in program.

