Finding convenience stores' optimal locations in Montevideo:

Introduction:

What is a convenience store?

It is an establishment with less than 400m^2, open at least 18 hours per day, all year. Its main daily consumption catalog items focus on drinks, feeding, products of watertight, bazaar, etc.

Location and product offerings of these stores are defined as the population density of the region, and usually have a 300m radius of consumption.

The greatest concentration of convenience stores in Uruguay is given in the capital, Montevideo department, and more specifically in its central and coastal zone.

Background:

According to official sources, the consumption trend through the convenience stores in Uruguay has experienced an increase in their preference of use, from 16% in 2017 to 27% in 2018. The same period shows that those who prefer to buy in supermarkets decreased their proportion from 78% to 64%.

Uruguay, but more particularly Montevideo, is undergoing a paradigm shift in terms of daily consumption, and there are at least two reasons why this situation can be explained. First, Uruguay recently switched to the category of countries with "high income" to be increased its per capita GDP. Under this concept, local consumers now seek time savings and simplicity when making purchases. On the other hand, Uruguay has a zero-growth rate of population and is experiencing growth in the number of households by the breakdown of family units. This leads to the Uruguayans live increasingly in smaller spaces with fewer members cohabiting, which encourages frequent and light purchases against the monthly supply. Under this paradigm, a few big supermarket brands had launched their own convenience stores, according to the new market preferences.

It should be noted that the main convenience stores in Montevideo are “Kinko”, “Devoto Express” y “Frog”, with a total billing of U$S7.7 million in 2015 and a total of 24, 27 and 22 locals respectively.

Idea:

Considering the new trend of consumption of the residents of Montevideo, this report aims to identify potential areas for installation of new convenience stores. The variables to handle with for this identification will be the density of population, purchasing power and criminality per district in the Montevideo department.

The target audience is both those store owners who already have market share, as those potential investors in the field.

Data:

In first place I'll be using a database created by myself using information from different Uruguayan gpvernment organizations, specifically I.N.E (Instituto Nacional de Estadistica), A.N.V (Agencia Nacional de Viviendas) and Ministerio del Interior, in order to have information about common property price per square meter, population density and robbery reports

disaggregated by neighborhood in Montevideo, Uruguay. This information will be used for discard some problematic neighborhoods where any investor would like to install a Convenience Store. Note that a bigger database is loaded in the Jupyter Notebook, and it's modified until get the final one.

That's how the final database looks like:

neighborhood	hct_price	burg_freq	tot_dens
AGUADA	3.529	186	117.63
AIRES PUROS	4.731	186	67.20
ATAHUALPA	6.561	186	67.20
BANADOS DE CARRASCO	3.843	186	19.80
BARRIO SUR	3.843	186	100.80

* 'neighborhood' contains the names of all 62 neighborhoods of Montevideo, Uruguay.

* 'hct_price' stand for "common property" square meter price's mean (2009), in thousands of indexed units. It's used as a purchasing power indicator.

* 'burg_freq' stands for robbery frequency registered in 2018. This will be used as a criminality indicator.

* 'tot_dens' stands for population density per square hectare in Montevideo (2011), and is used as another indicator for the store establishment.

The criteria for decide which are the best neighborhoods consists in create a Price per Square Meter - Robbery Reports ratio. Thus, I will select those neighborhoods where this ratio is higher, combined with those where population density is also higher. The next database shows the result:

	neighborhood	hct_price	burg_freq	tot_dens	Latitud	Longitud	Price_Burglary_Ratio
51	PUNTA CARRETAS	10.763	186	148.6	-34.92030	-56.16154	0.057866
48	POCITOS	9.367	369	148.6	-34.90853	-56.15087	0.025385
43	PARQUE RODO	7.365	186	144.6	-34.91862	-56.16416	0.039597
42	PALERMO	5.437	186	144.6	-34.91150	-56.17860	0.029231
28	LA COMERCIAL	5.495	186	126.1	-34.88732	-56.17040	0.029543

'Latitud' and 'Longitud' columns contains coordinates for each neighborhood, extracted with a specific code shown in Jupyter Notebook.

In second place, a database with the ten most common venues in each one of the best neighborhoods selected is created. This will be useful for k-means clustering. With this tool I'll be able to understand a bit more about the intern characteristics of each zone, and thus, decide which one is the most suitable for a convenience store. The factor that will determine this decision is the venue's frequency: I'll select those neighborhoods where the most frequent

venue is related in some way with convenience stores' target public (e.g. schools, sport clubs, churches).

Lastly, a map will be displayed with the best zone found, the venues related with convenience stores' target public inside this zone and the convenience stores already installed here. With this information, I'll get a relatively specific area for the convenience store installation. Note that the final result does not consist in giving a specific coordinate for the installation, this methodology just finds the best area.

The .JSON file used for print population density per neighborhood with a choropleth map was found on a github repository, and it was modified manually through the different project stages.

Methodology and Analysis:

This part of the report is composed of three 'stages':

In first stage section the objective is discard regions according to three socioeconomic criteria: population density, criminality and purchasing power per neighborhood.

For this, a ratio "common property" square meter price against robbery frequency is created as one way of see "how suitable" is a neighborhood for a convenience store. Note that as larger the ratio, the neighborhood has more purchasing power in relation to criminality. Thus, the main idea consists in select the best five neighborhoods where population density is higher, and, at the same time, where this ratio is also higher.

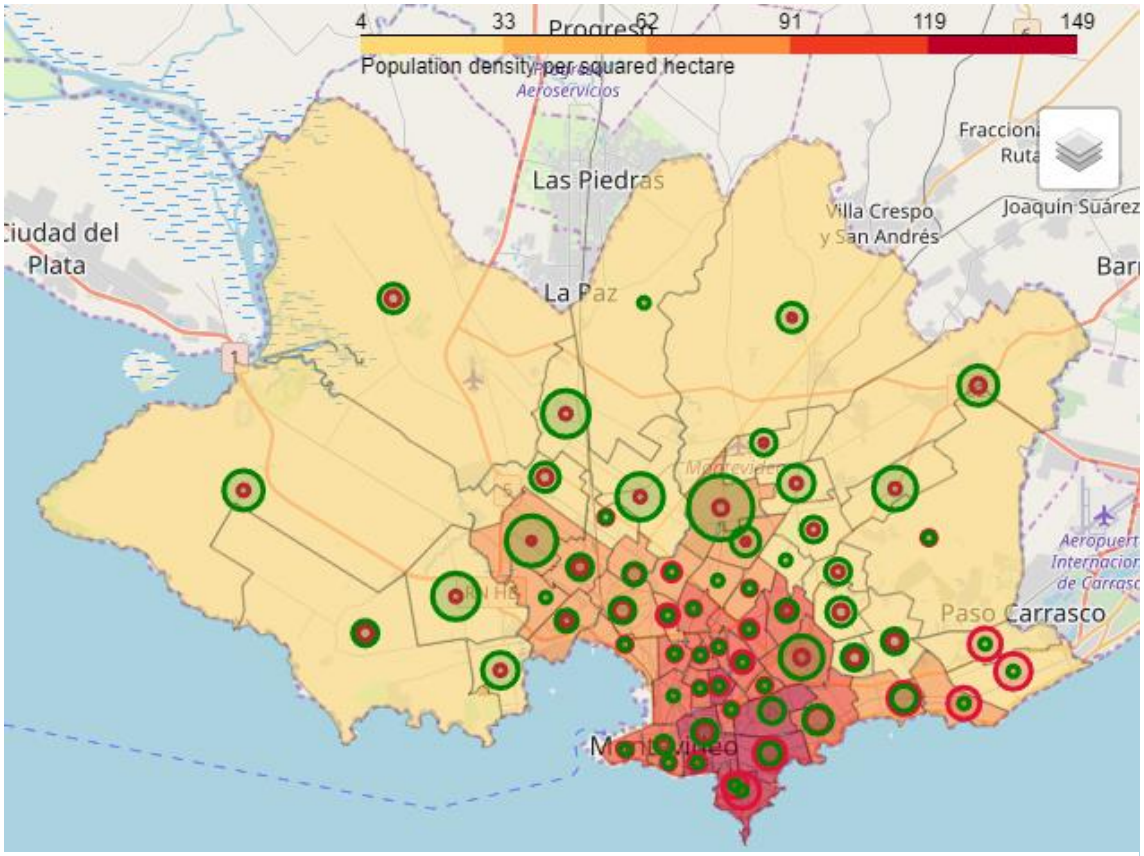

Figure 1: Population density, purchasing power (red circles) and criminality (green circles) per neighborhood, Montevideo.

Visually, we can get a first idea about which neighborhoods are the best by comparing size of red circles and green circles, that represent purchasing power and criminality level respectively. Although both measures are incomparable, this allows comparations between neighborhoods.

As we can see, Montevideo shows a clear population density distribution, concentrated in south central area and decreasing as we move towards the periphery. Moreover, it seems to be a correlation between neighborhood location, purchasing power and criminality rate, in which the more to the periphery, the more crime and the less purchasing power. It is tested below:

	burg_freq	hct_price	tot_dens	a1	a3	a4	a5
burg_freq	1.0***	-0.32*	-0.41**	0.36*	0.02	-0.31	-0.15
hct_price	-0.32*	1.0***	0.42	-0.26	0.03	-0.07	0.45*
tot_dens	-0.41**	0.42	1.0***	-0.87***	0.05	0.49	0.6*
a1	0.36*	-0.26	-0.87***	1.0***	-0.45	-0.47	-0.28
a3	0.02	0.03	0.05	-0.45	1.0***	-0.33	-0.2
a4	-0.31	-0.07	0.49	-0.47	-0.33	1.0***	-0.21
a5	-0.15	0.45*	0.6*	-0.28	-0.2	-0.21	1.0***

*Table 1: Correlations with Statistical Significance (P-Values: *10%, **5%, ***1%)*

'a1', 'a3', 'a4', and 'a5' are categorical variables that represent neighborhoods from the lighter yellow area to the dark purple area or the center zone, being 'a1' the furthest area from the center and 'a5' the dark purple or the center itself. P-value analysis shows only a significative positive correlation between the periphery and criminality rate, and a significative positive correlation between the center and purchasing power.

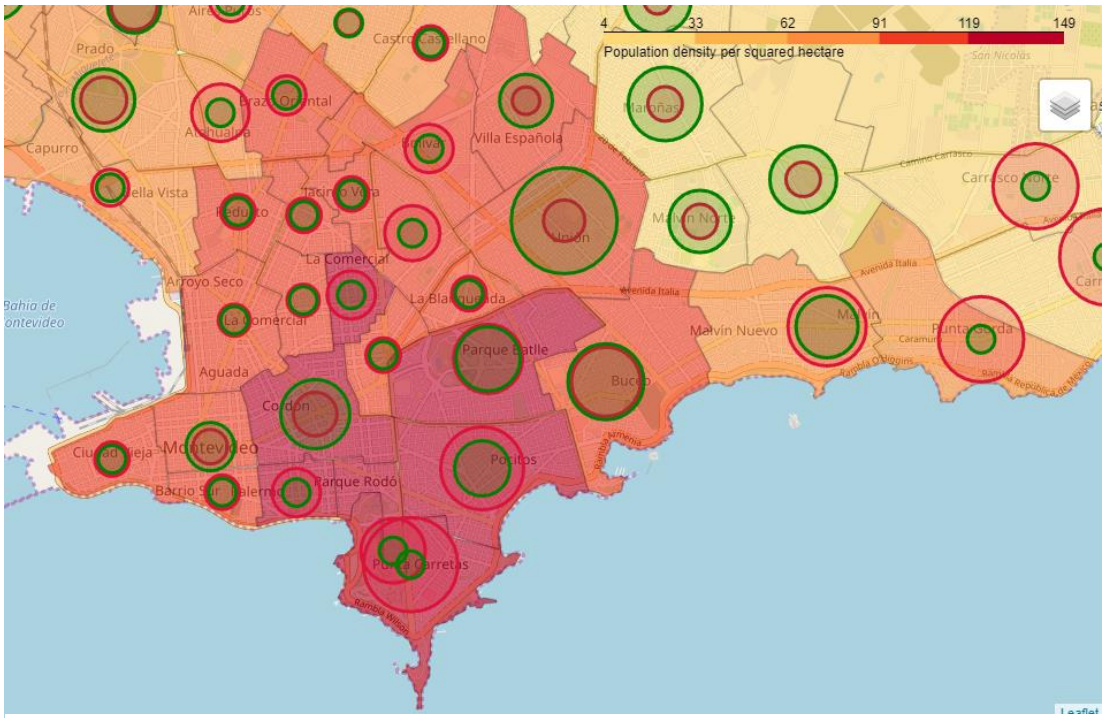

Figure 2: Population density, purchasing power (red circles) and criminality (green circles) per neighborhood, Montevideo (Zoom into interest zone).

This analysis can be considered as one way to justify why the center area is the best zone for establish a convenience store, based on the proposed criteria.

In the second stage, after the best zone is selected, it's explored by getting the top 5 most common venues in each neighborhood. Then, k-means clustering is used to identify similarities between neighborhoods based in those venues, and so select the cluster characterized by having more venues related with convenience stores' target public.

	neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
0	LA COMERCIAL	Restaurant	BBQ Joint	Ice Cream Shop	Sandwich Place	Electronics Store	Latin American Restaurant	Pharmacy	Coffee Shop	Music Venue	Middle Eastern Restaurant
1	PALERMO	Bar	Restaurant	Italian Restaurant	Scenic Lookout	Bakery	Gym	Pub	Coffee Shop	Other Great Outdoors	Hotel
2	PARQUE RODO	Hotel	Steakhouse	Park	Coffee Shop	Restaurant	BBQ Joint	Bakery	Sushi Restaurant	Pizza Place	Café
3	POCITOS	Gym / Fitness Center	Italian Restaurant	Coffee Shop	Restaurant	Dessert Shop	Ice Cream Shop	Brewery	Pizza Place	Café	Vegetarian / Vegan Restaurant
4	PUNTA CARRETAS	Hotel	Deli / Bodega	Restaurant	Steakhouse	Coffee Shop	Pizza Place	Modern European Restaurant	Gym	Gym / Fitness Center	Café

Table 2: Most Common Venues by Neighborhood

Table 2 gives a first idea about what kind of venues are installed in each neighborhood, although it's difficult to conclude something just with this information. That's why the procedure continues with cluster those neighborhoods, in order to find some similarities between them. The number of clusters chosen were two.

"Cluster 0" shows a great frequency of hotels, a mean of 9.064 thousand of indexed units for "common property" square meter price, as well as an average of 186 robbery reports and also an average of 146.6 persons per square hectare.

	neighborhood	hct_price	burg_freq	tot_dens	Latitud	Longitud	Price_Burglary_Ratio	Cluster Labels
51	PUNTA CARRETAS	10.763	186	148.6	-34.92030	-56.16154	0.057866	0
43	PARQUE RODO	7.365	186	144.6	-34.91862	-56.16416	0.039597	0

	neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
51	PUNTA CARRETAS	Hotel	Deli / Bodega	Steakhouse	Restaurant	Modern European Restaurant	Coffee Shop	Gym / Fitness Center	Café	Sushi Restaurant	BBQ Joint
43	PARQUE RODO	Hotel	Steakhouse	Coffee Shop	Park	Restaurant	BBQ Joint	Bakery	Pizza Place	Sushi Restaurant	Café

Table 3: "Cluster 0" insight.

	hct_price	burg_freq	tot_dens	Price_Burglary_Ratio
count	2.000000	2.0	2.000000	2.000000
mean	9.064000	186.0	146.600000	0.048731
std	2.402749	0.0	2.828427	0.012918
min	7.365000	186.0	144.600000	0.039597
25%	8.214500	186.0	145.600000	0.044164
50%	9.064000	186.0	146.600000	0.048731
75%	9.913500	186.0	147.600000	0.053298
max	10.763000	186.0	148.600000	0.057866

Table 4: "Cluster 0" description.

On the other hand, “Cluster 1” doesn’t show a clear pattern about venues distribution. It’s interesting to note that in both clusters it is possible to distinguish a great amount of food places, like restaurants, coffee shops, bakery, etc., but “Cluster 0” has a clear tendency to concentrate more hotels in relation with “Cluster 1”.

About “Cluster 1” description, it’s possible to see a mean of 6.77 thousand of indexed units for "common property" square meter price, as well as an average of 247 robbery reports and also an average of 140 persons per square hectare.

	neighborhood	hct_price	burg_freq	tot_dens	Latitud	Longitud	Price_Burglary_Ratio	Cluster Labels
48	POCITOS	9.367	369	148.6	-34.90853	-56.15087	0.025385	1
42	PALERMO	5.437	186	144.6	-34.91150	-56.17860	0.029231	1
28	LA COMERCIAL	5.495	186	126.1	-34.88732	-56.17040	0.029543	1

	neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
48	POCITOS	Gym / Fitness Center	Italian Restaurant	Coffee Shop	Restaurant	Dessert Shop	Ice Cream Shop	Pizza Place	Brewery	Café	Vegetarian / Vegan Restaurant
42	PALERMO	Bar	Restaurant	Italian Restaurant	Pub	Other Great Outdoors	Bakery	Gym	Coffee Shop	Scenic Lookout	Gym / Fitness Center
28	LA COMERCIAL	Restaurant	BBQ Joint	Ice Cream Shop	Sandwich Place	Electronics Store	Department Store	Pharmacy	Coffee Shop	Music Venue	Middle Eastern Restaurant

Table 5: "Cluster 1" insight.

	hct_price	burg_freq	tot_dens	Price_Burglary_Ratio
count	3.000000	3.000000	3.000000	3.000000
mean	6.766333	247.000000	139.766667	0.028053
std	2.252430	105.655099	12.003472	0.002316
min	5.437000	186.000000	126.100000	0.025385
25%	5.466000	186.000000	135.350000	0.027308
50%	5.495000	186.000000	144.600000	0.029231
75%	7.431000	277.500000	146.600000	0.029387
max	9.367000	369.000000	148.600000	0.029543

Table 6: "Cluster 1" description.

Given this information, it is natural to associate “Cluster 0” as the most suitable for a convenience store installation, judging by its great hotel concentration compared with “Cluster 1”, and assuming that this concentration is related with a significative flow of high/medium high-income level population.

The final stage consists in display in a map all the hotels inside the selected cluster's limits. Also, convenience stores already installed in the selected zone will be located; this will allow to find some area where the distance between hotels and the nearest convenience store exceed three blocks. By the way, this is the convenience store's radius of consumption according to GPA (2019. With all this information, we will be able to determine a relatively specific zone where a convenience store should be installed.

The map below shows “Cluster 0” limits, hotels inside this zone represented by yellow dots and convenience stores already installed represented by red dots.

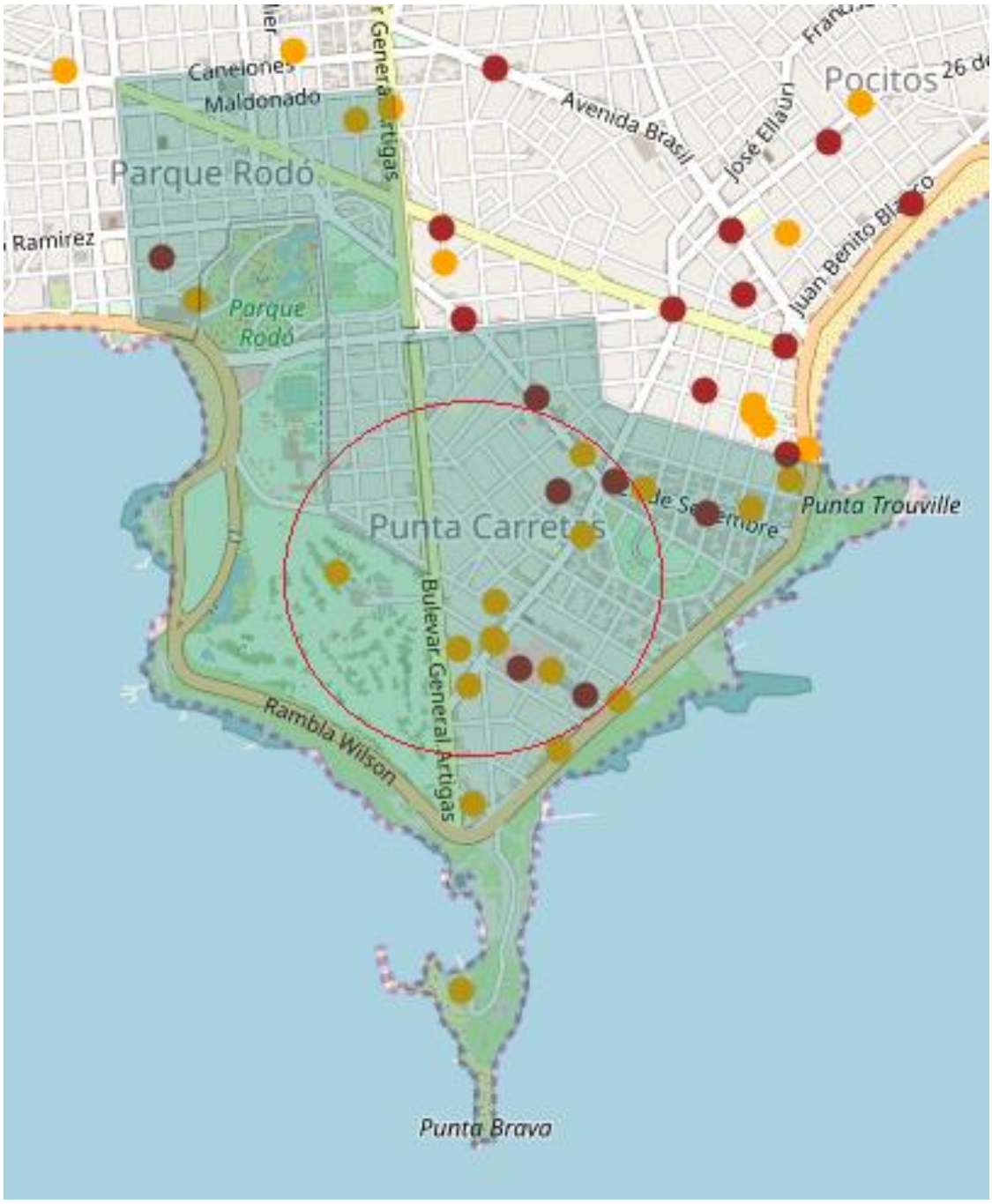

Figure 3: "Cluster 0" map. Yellow dots: Hotels. Red dots: Convenience Stores.

Results and Discussion:

The best area for establish a convenience store seems to be the center or south-center area of Punta Carretas, according to all the standards defined previously and considering Foursquare's accuracy locating venues. It's important to remember that Punta Carretas is characterized for having a great purchase power, combined with a relatively low level of criminality and one of the greatest population densities in Montevideo. Also, this zone concentrates a relatively high number of hotels that, judging by its socioeconomic level, implies an important flow of medium-high/high income level population.

Although the given result is coherent, it's a fact that Foursquare maps aren't very used here in Uruguay and the disposable information may not be actualized, so this report is only useful for have a general idea about where to start.

Also, it's possible to add some others evaluation criteria besides the three considered in this work, like transit flow or closeness to social centers (e.g. Sport Clubs, Schools or Churches).

Furthermore, it's important to note that database created for this report collects information from different ages, being the less actualized indicator from 2009. Related to this, it's also possible to propose others indicators for purchase power or criminality, as well as being less restrictive about the three factors considered. May a modification of this data could extend the interest zone to other neighborhoods.

Conclusion:

This report aimed to find the best locations for install a new convenience store, by imposing different filters stage by stage. After discard the 'worst regions' in Montevideo based on some socioeconomic indicators, the most appropriate was found using k-means clustering and already installed convenience stores' locations. The stakeholders could use this project and change these filters according to their preferences about criminality, purchase power, population density, proximity of other venues, and so on. With this versatility any investor who want to establish a new convenience store could adjust the parameters based on its own standards.

Webograpghy:

http://www.gpa.uy/files.php/archivos-externos/8898/2019-02-12-informe-retail.pdf?dl=false

https://es.wikipedia.org/wiki/Tienda_de_conveniencia

https://www.elobservador.com.uy/nota/las-tiendas-de-conveniencia-arrasan-en-la-oferta-minorista-2016413500

https://raw.githubusercontent.com/vierja/geojson_montevideo/master/barrios.geojson