

Select a borough to set up a productive business in Colombia

APPLIED DATA SCIENCE

Capstone Project
The Battle of Neighborhoods

Introduction

Background

Colombia, my country, is growing in terms of production, industry, agriculture and business; so, more and more foreign companies are betting on investing capital in these lands. Globalization has allowed shortening borders and establishing a diverse trading.



Interest

- Investors looking for projects to fund, and interested in Colombian diversity.
- Someone who wants to identify opportunities in Colombian lands
- New foreign investors who want to establish their headquarters in Colombia and are looking for a place to live similar to their place of origin



Business Problem

For a foreign investor that want to set up business in Colombia what kind of product commercialize with the objective of export to other countries, keeping in mind the Colombian economy and the wide range of climatic variations?

Where the investor could set up his office? In addition, how similar could result the selected neighborhood in Colombia in comparison with the neighborhood he was used to living.



DATA



Data Sources

- <https://www.maro.com.co/>
- [Colombia JSON file](#)
- [Antioquia Boroughs](#)
- [Foursquare location data](#)

Data wrangling and cleaning

Pandas Data Frame

- Sort
- Filter
- Clean
- Join
- Drop

Feature Selection

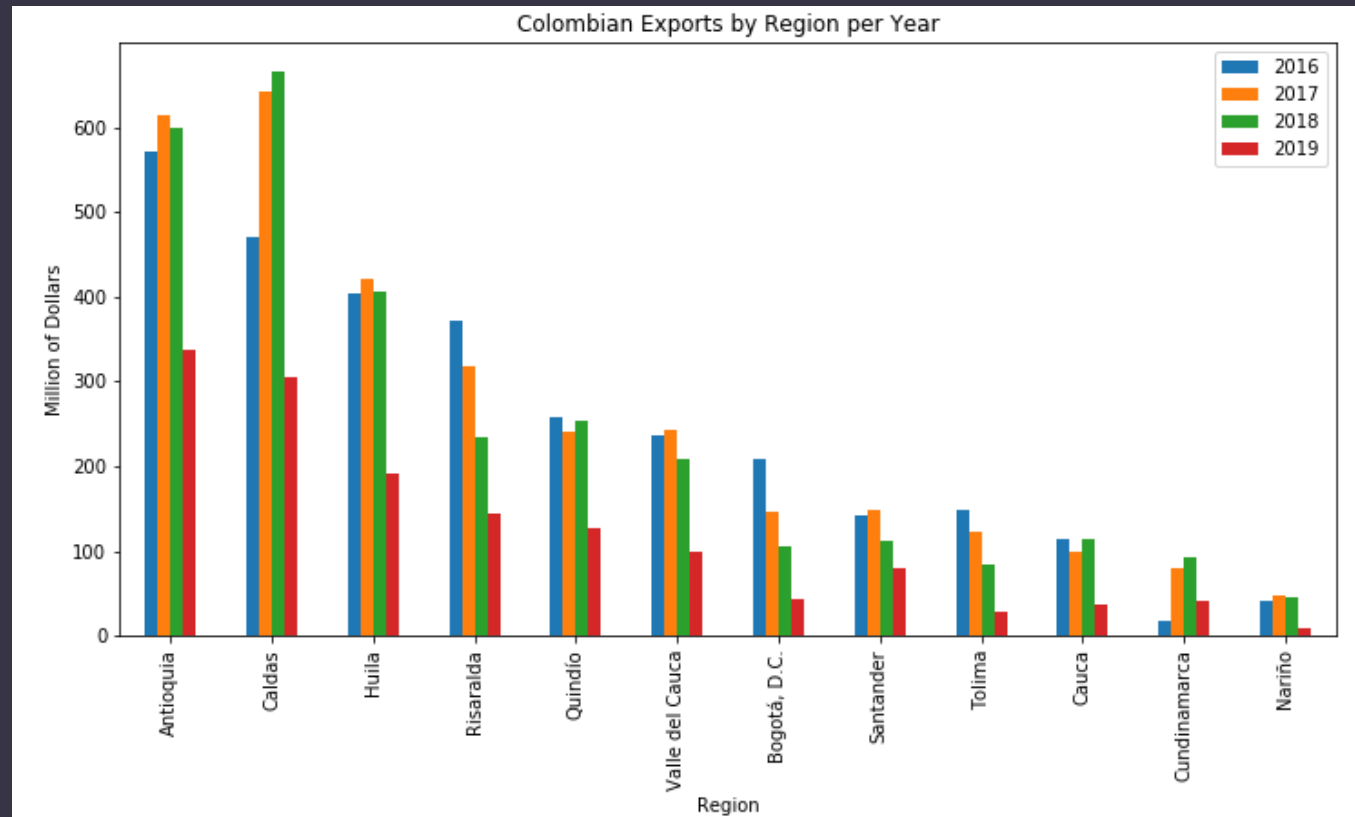
Colombian Exports Data Base

INDEX	Column Name	Description
0	NANDINA	ID Product subcategory
1	Descripción NANDINA	Product subcategory description
2	CIIU Rev.4	ID Product Category
3	Descripción CIIU Rev.4	Product Category Description
4	Destino	Country Export destination Country
5	Departamento	Department in colombia of export origin
6	FOBDO 2010	Value exported in dollars for each year
7	FOBDO 2011	
8	FOBDO 2012	
9	FOBDO 2013	
10	FOBDO 2014	
11	FOBDO 2015	
12	FOBDO 2016	
13	FOBDO 2017	
14	FOBDO 2018	
15	FOBDO ENE-JUN 2018	Value exported in net kilogram
16	FOBDO ENE-JUN 2019	
17	KNETO 2010	
18	KNETO 2011	
19	KNETO 2012	
20	KNETO 2013	
21	KNETO 2014	
22	KNETO 2015	
23	KNETO 2016	
24	KNETO 2017	
25	KNETO 2018	
26	KNETO ENE-JUN 2018	
27	KNETO ENE-JUN 2019	

METHODOLOGY

Exports by department

Top 12 Colombian exports by region per year (Millions of dollars)



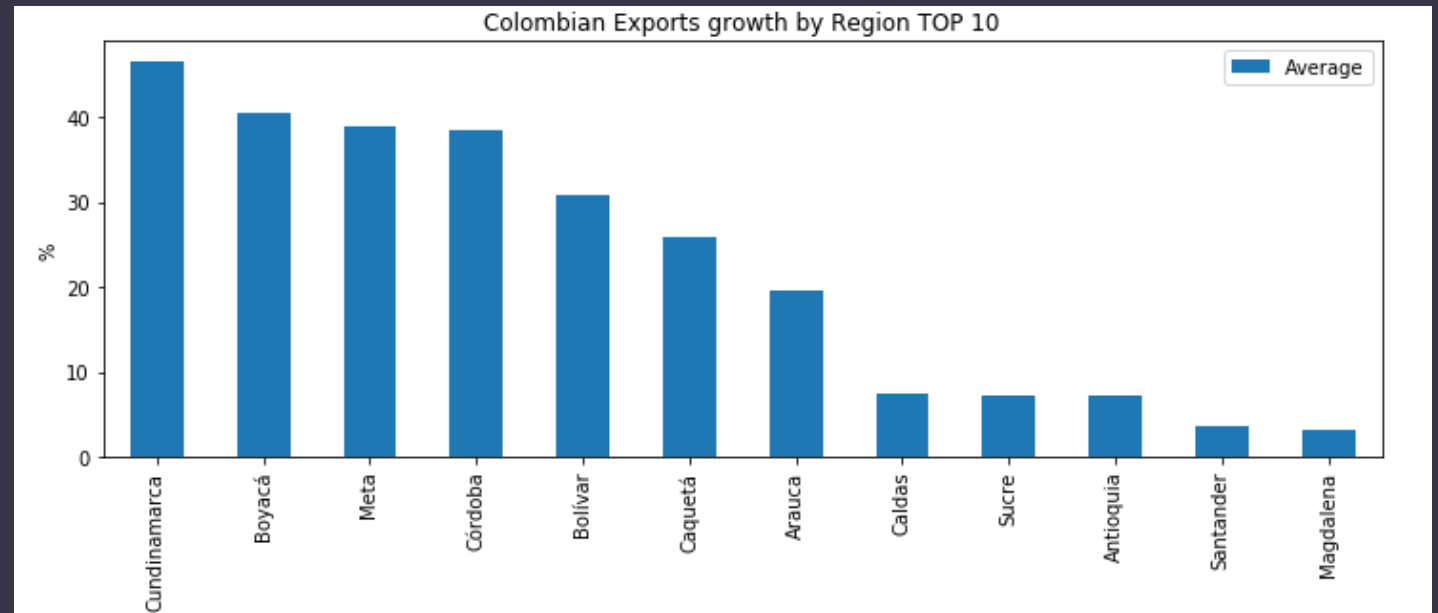
As is evident, the departments of Antioquia and Caldas are the regions leaders in exports

METHODOLOGY

Exports by department

TOP 10 Colombian exports percentage growth between years

	Region	2014-2015	2015-2016	2016-2017	2017-2018	Average
0	Cundinamarca	31.337458	-13.331624	138.710173	29.809296	46.631326
1	Boyacá	0.000000	2.143277	433.493868	-273.417566	40.554895
2	Meta	0.000000	40.135241	4.976893	110.458957	38.892773
3	Córdoba	27.321792	14.460533	98.195294	13.819262	38.449220
4	Bolívar	70.761086	-127.789570	102.224966	78.521597	30.929520



Which departments show the largest percentage growth between years?. To identify that, a new data frame was created calculating the percentage increase or decrease in exports between years: 2014-2015, 2015-2016, 2016-2017, 2017-2018 and then the average growth.

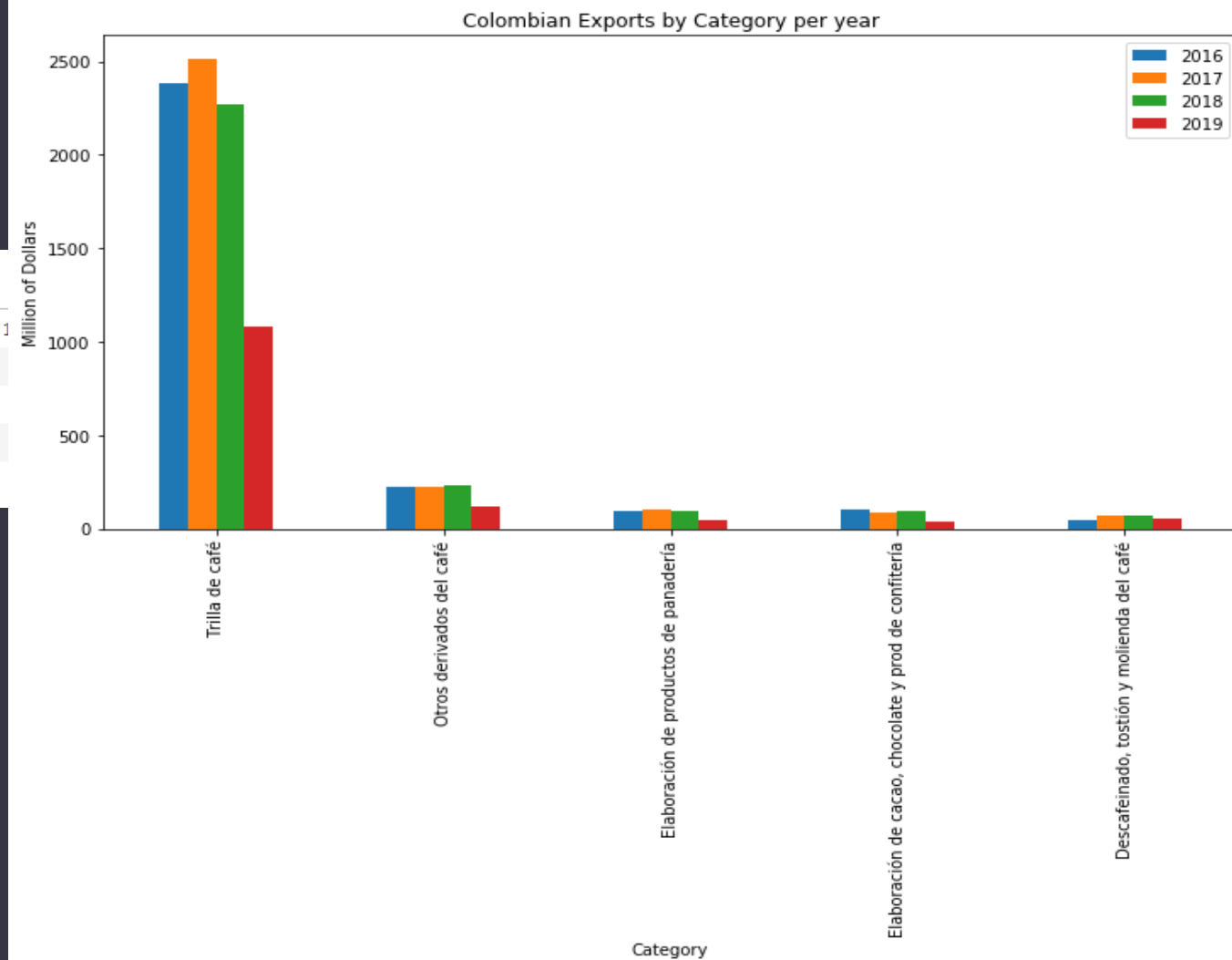
METHODOLOGY

Exports by product category

Top 5 Colombian exports by category per year (Millions of dollars)

Exports by Product Category has 12 Rows and, 9 columns

	Category	2014	2015	2016	2017	2018	2019
0	Trilla de café	2473.260563	2526.470898	2379.267592	2513.090734	2267.459840	1081.603544
1	Otros derivados del café	221.299644	233.621727	220.016536	224.865705	233.153149	116.476565
2	Elaboración de productos de panadería	133.272579	97.685515	92.212480	100.973334	95.496111	42.503701
3	Elaboración de cacao, chocolate y prod de conf...	121.515210	87.200311	105.070720	84.871409	91.606604	39.058075
4	Descafeinado, tostión y molienda del café	43.433770	49.981457	44.801176	68.495834	67.767740	51.836802



By a wide margin, coffee and its derivatives are the star export product in Colombia

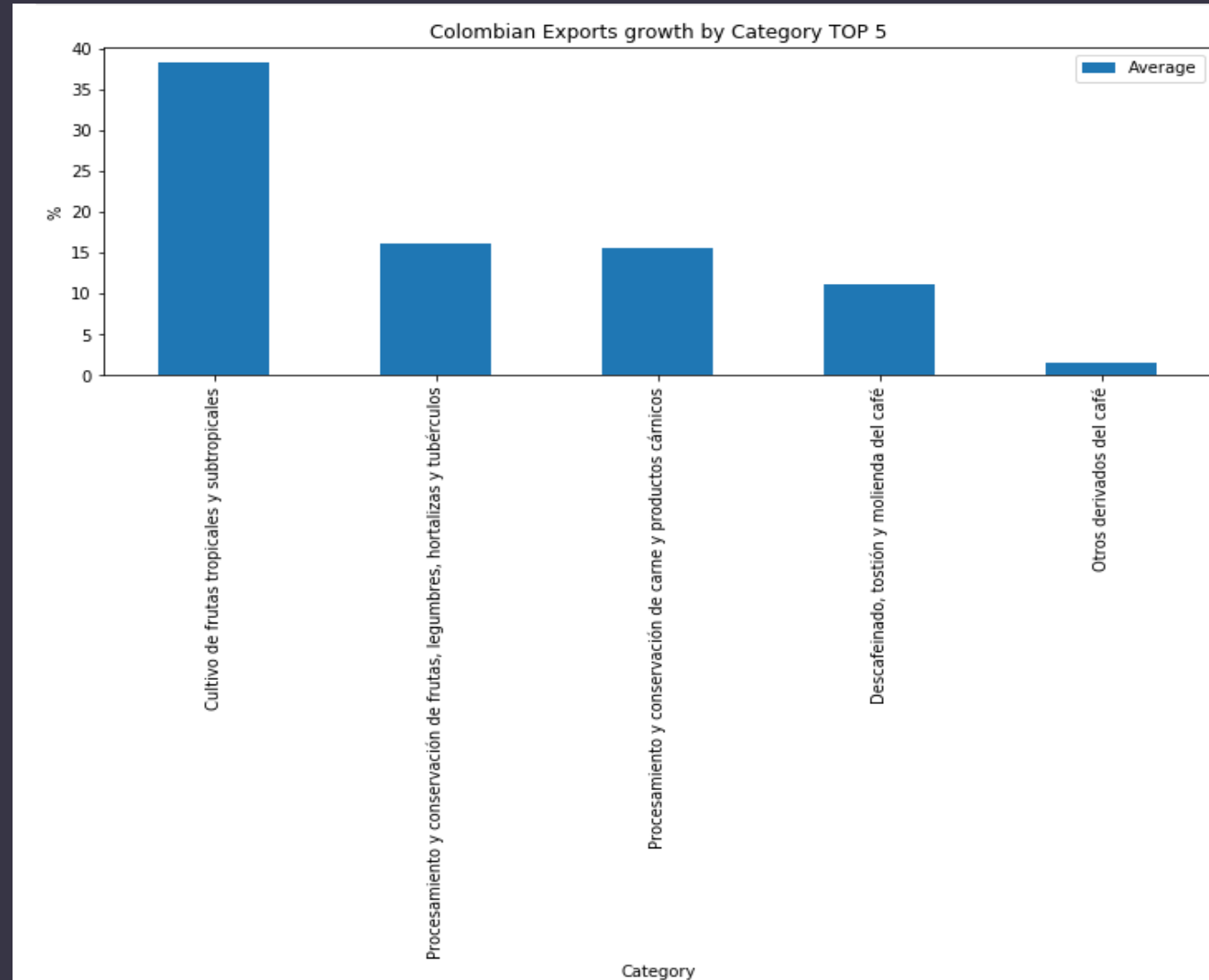
METHODOLOGY

Exports by department

TOP 5 Colombian exports percentage growth between years by category

		Category				
	Category	2014-2015	2015-2016	2016-2017	2017-2018	Average
0	Cultivo de frutas tropicales y subtropicales	18.620080	70.259075	51.366242	12.688275	38.233418
1	Procesamiento y conservación de frutas, legumb...	15.632954	13.455669	20.246361	14.825145	16.040032
2	Procesamiento y conservación de carne y produc...	-23.442304	2.871297	67.222330	15.159912	15.452809
3	Descafeinado, tostión y molienda del café	12.039259	-9.525003	43.567465	-1.338751	11.185743
4	Otros derivados del café	5.917282	-6.533453	2.328657	3.979777	1.423066

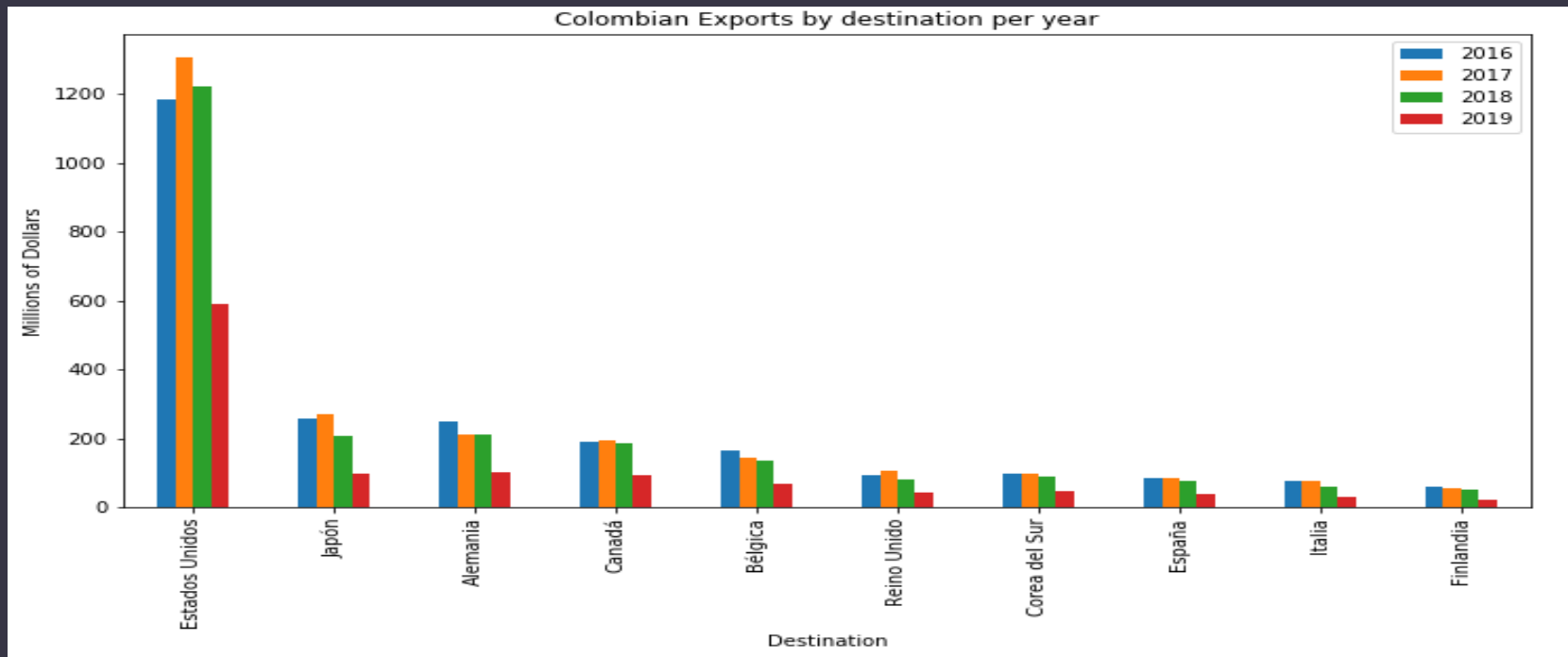
Let's see what happen in terms of growth, which product categories shows the largest percentage growth between years?



METHODOLOGY

Exports by destination.

Top 10 destinations of Colombian exports per year (Millions of dollars)



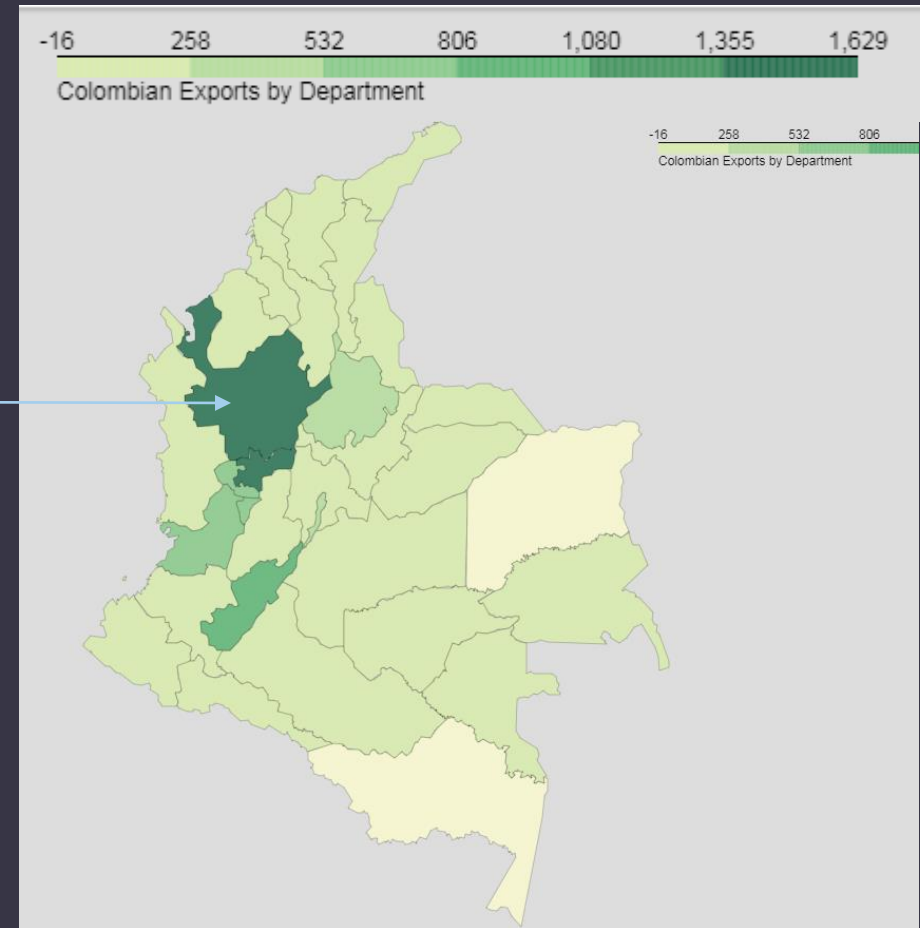
Due to the proximity and current trade agreements, United States is the top export destination of Colombia with about \$11.1B of dollars. Japan, Germany and Canada follow the list.

RESULTS

Find the region with better opportunities.

Region	2017	2018	2019	Total
ANTIOQUIA	613.578035	598.856402	336.743726	1549.178164
ARAUCA	0.140239	0.128348	0.164041	0.432628
ATLANTICO	6.056956	2.230487	1.542750	9.830194
SANTAFE DE BOGOTA D.C	146.757452	106.787289	43.160371	296.705112
BOLIVAR	2.631158	4.348195	0.912072	7.891425
BOYACA	0.053314	0.019853	0.000000	0.073167
CALDAS	641.742743	666.141690	304.714387	1612.598820
CAQUETA	0.228826	0.048947	0.000000	0.277773
CASANARE	0.001371	0.000000	0.002100	0.003471
CAUCA	98.766098	113.592569	37.102860	249.461527

Antioquia



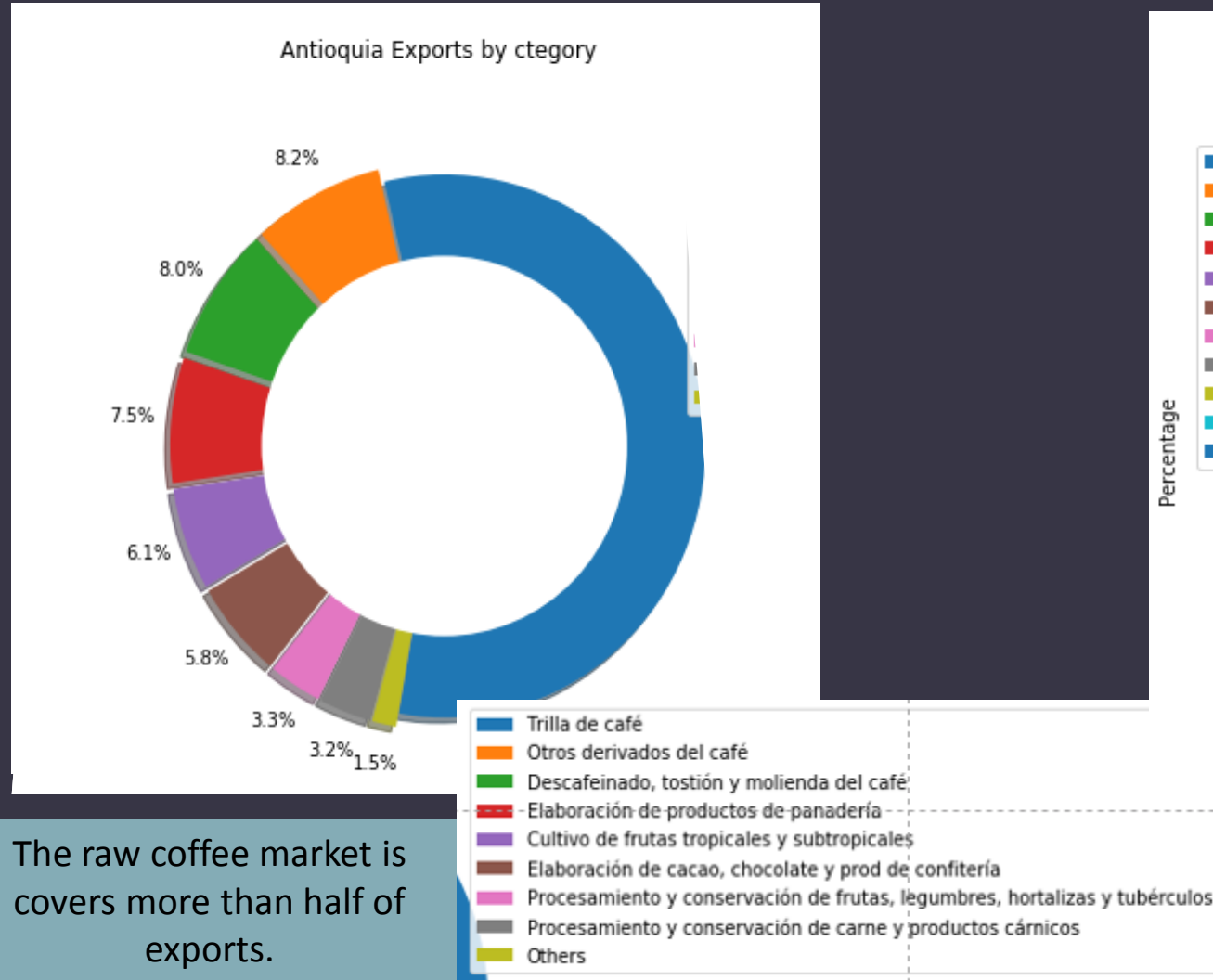
As we can see the darkest regions are those that in the period 2017 – 2019 had a greater value of the exports in millions of dollars:

1. Antioquia
2. Caldas
3. Huila
4. Risaralda
5. Quindio

RESULTS

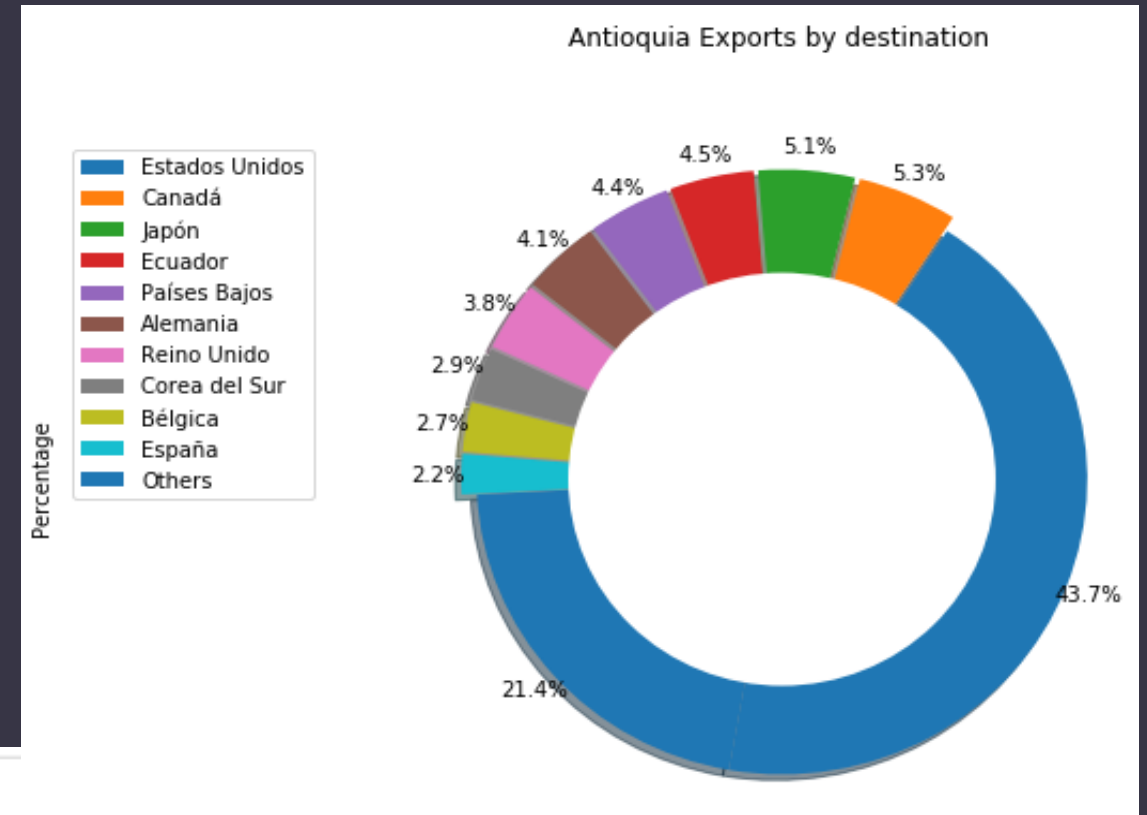
Antioquia's Market.

Product category exported from Antioquia



The raw coffee market is covers more than half of exports.

Antioquia's exports destinations

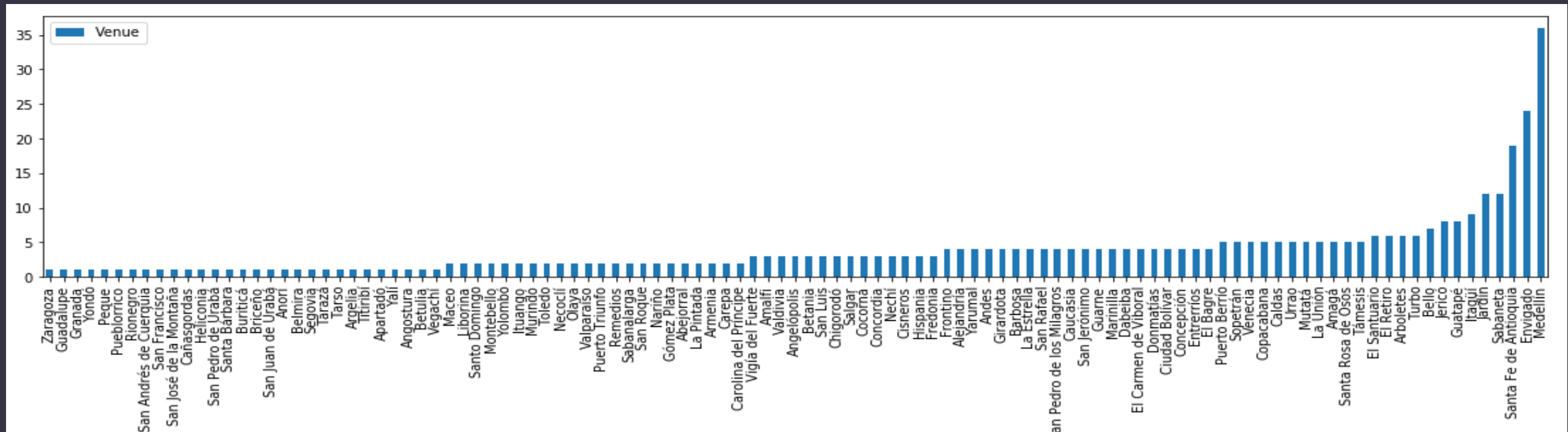


United States is the favorite destination for Antioquia's exports followed by Canada with a 5.3% and Japon with 5.1%

RESULTS

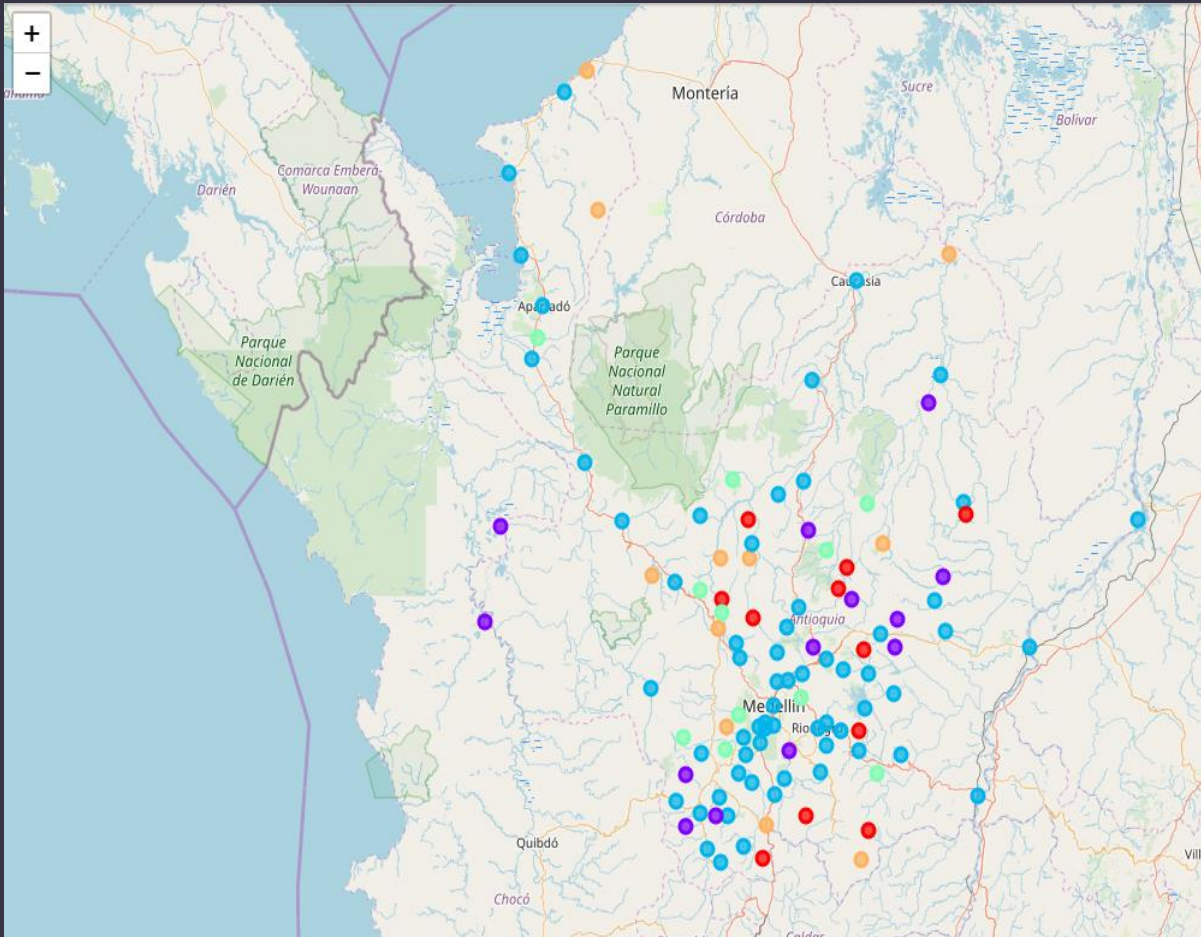
Cluster venues in Antioquia's boroughs

Using the Foursquare location data, a new Data Frame was build including 402 veneus located in Antiquia's boroughs. The figure 17 shows the amount of popular venues that were found by borough



RESULTS

Cluster venues in Antioquia's boroughs



Cluster 1: Boroughs with a Plaza as a first or second most common venue

Cluster 2: Boroughs with Hotels as a first most common venue

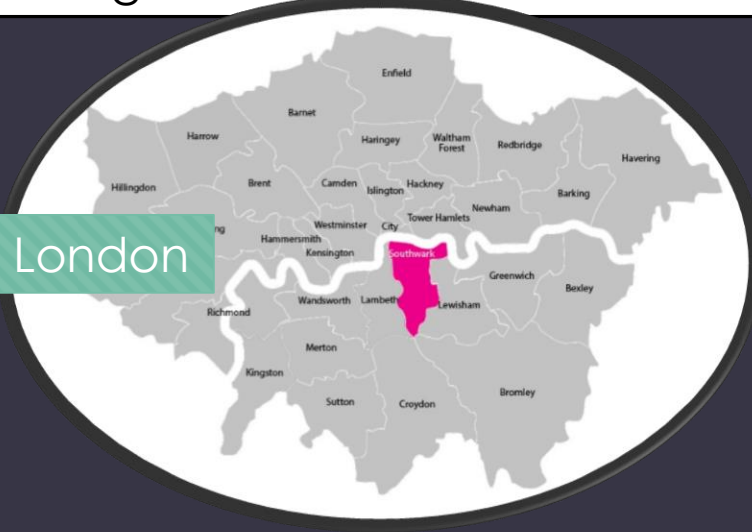
Cluster 3: Boroughs with a Park as a first or second most common venue

Cluster 4: Boroughs with Coffees, bars, farms into their 10 first common venues

Cluster 5: Boroughs with Restaurants, Fast food restaurants, farmers Market, electronic stores into their 10 first common venues

RESULTS

Find a Borough in Antioquia similar to Southwark, London.



Southwark, London

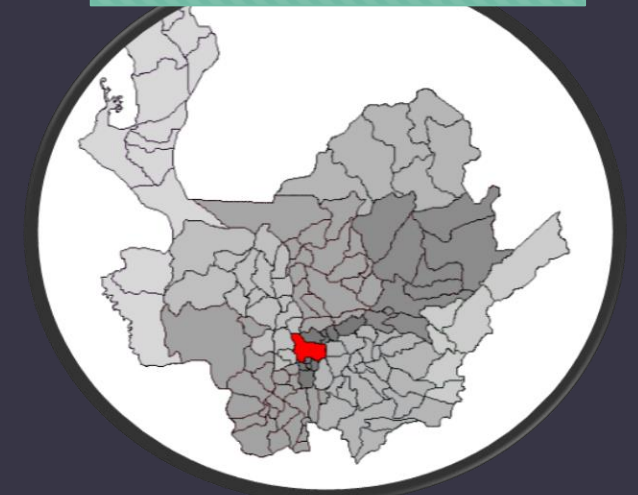
Assuming our investor used to live in Newham, London, and with the Foursquare location data we calculate the same matrix with the categories of common venues in Southwark

Then we calculate the Euclidian distance between the Southwark vector and each borough in Antioquia.

As result, we find Medellin as Borough in Antioquia that result more similar to Southwark, London in terms of popular venues around (The shortest Euclidian distance)

	Borough	Distance
59	Medellín	0.253296
38	Envigado	0.295701
51	Jardín	0.313139
45	Guatapé	0.335857
49	Itagüí	0.338260
...
78	San Francisco	1.014298
16	Betulia	1.014298
43	Guadalupe	1.014298
47	Heliconia	1.014298
18	Buriticá	1.014298
109 rows × 2 columns		

Medellín, Antioquia



CONCLUSION

- Colombian exports data estimated that the department with the greatest export value is Antioquia, where there are variety of products to process and export like coffee byproducts, fruits and vegetables and bakery products.
- The preferred product to export in Colombia is the raw coffee with which the greatest amount of export value has been realized in at least the last 7 years.
- Medellín the capital city of Antioquia is the borough with the greater number of popular venues. Using the Euclidian distance as comparative factor, Medellín is the borough with more similarities to Southwark, London in terms of popular venues around (The shortest Euclidian distance)

