IBM APPLIED DATA SCIENCE CAPSTONE PROJECT

Finding places that maximize the probability of success when opening a coffee shop in Rio de Janeiro, Brazil

Resume

In this report we will define strategic regions combining geospatial data and knowledge of coffee target audience to answer the question: 'If I look to open a new coffee shop in Rio de Janeiro, how can I choose a location that maximizes the probability of success?'

Felipe Vasconcelos dos Passos

February, 2021

Table of Contents

Abstract

This project is looking for the best place to open a coffee shop in the city of Rio de Janeiro based on the location of venues that already exists in the city and their characteristics. We will use clustering methods and machine learn techniques to define strategic neighborhoods and streets to answer the question: “If I look to open a new coffee shop in Rio de Janeiro, how can I choose a location that maximizes the probability of success?”. The project will be runner in two weeks and this is the initial report where the items to submit are:

1 - A description of the problem and a discussion of the background.

2 - A description of the data and how it will be used to solve the problem.

Objective

The two main objectives are:

1 - Describe the problem and discuss the coffee market.

2 - Describe the data requirements (types and formats needed), data collections (which are the data sources) and the planned methodology explaining how the data can be used to solve the problem.

Introduction

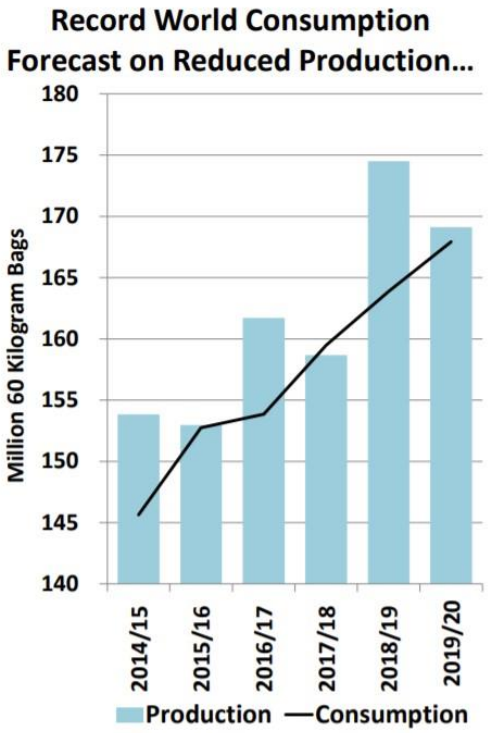

Figure 1. Source: USDA - Coffee: World Markets and Trade, jun 2019

Coffee is one of the most traditional products in Brazil, that produces about a third of all coffee in the world and is by far the world's largest producer. However, despite being what produces the most, when it comes to consumption, it appears in the second position with 6.02 kg per year per capita, behind the United States [1]. This occurs in a context that the market consumption and production have been growing worldwide [2].

According to the Brazilian Coffee Industry Association (ABIC), in the country, 95% of the population consumes the product, both at home and in food services, which represents 13% of world demand [3]. However, only 10% of consumption is represented by specialty coffees, which is the main coffee category sold in food services.

According to a study by Euromonitor consultancy, the Brazilian market for premium coffees has grown at an accelerated rate, as it grows 15% per year, while the traditional coffee market grows 3.5% per year. Despite this, most domestic consumption is still traditional coffee, that is usually sold as ground coffee in supermarkets and consumed at home as pure strained coffee [1]. On the other hand, when we look only at consumption away from home, we notice that consumers have a preference for better quality espresso coffee [4]. Thus, we can look at premium and specialty coffees sold in coffee shops as a promising business case for the coming years.

	Breakfast	Morning	After Lunch	Afternoon	After Dinner
Pure coffee (strained)	32%	28%	26%	21%	20%
Latte	22%	9%	4%	9%	14%
Decaf coffee	5%	5%	4%	4%	9%
Espresso	5%	14%	20%	12%	6%
Coffee capsule	5%	6%	8%	8%	7%
Soluble coffee	12%	5%	6%	7%	10%
Coffee sachet	3%	5%	5%	5%	6%
French Press	2%	5%	5%	4%	5%
Italian (moka)	3%	5%	7%	7%	6%
Machine made Cappuccino	4%	10%	10%	14%	6%
Instant cappuccino (not machine-made)	7%	7%	6%	9%	11%
Others	0%	0%	0%	0%	0%

Figure 2. coffee consumption preference by time of day. Source: Euromonitor consultancy [4]

Away from home, espresso coffee leads the preference in different venues

	Coffee Shops	Restaurants	Bakeries	Kiosks	bars/snack bars
Pure coffee (strained)	8%	16%	16%	11%	16%
Latte	8%	7%	20%	9%	15%
Decaf coffee	4%	4%	3%	4%	4%
Espresso	27%	29%	21%	27%	25%
Coffee capsule	5%	7%	4%	6%	3%
Soluble coffee	2%	3%	3%	3%	3%
Coffee sachet	4%	5%	3%	4%	3%
French press	6%	4%	4%	5%	4%
Italian (moka)	11%	6%	6%	8%	7%
Machine made Cappuccino	18%	14%	15%	15%	14%
Instant cappuccino (not machine-made)	7%	7%	7%	8%	7%
Others	0%	0%	0%	1%	0%

Despite the dominance of espresso coffee; Pure coffee is still of great importance in non-specialized places.

Figure 3. Preference for drinking coffee away from home. Source: Euromonitor consultancy [4]

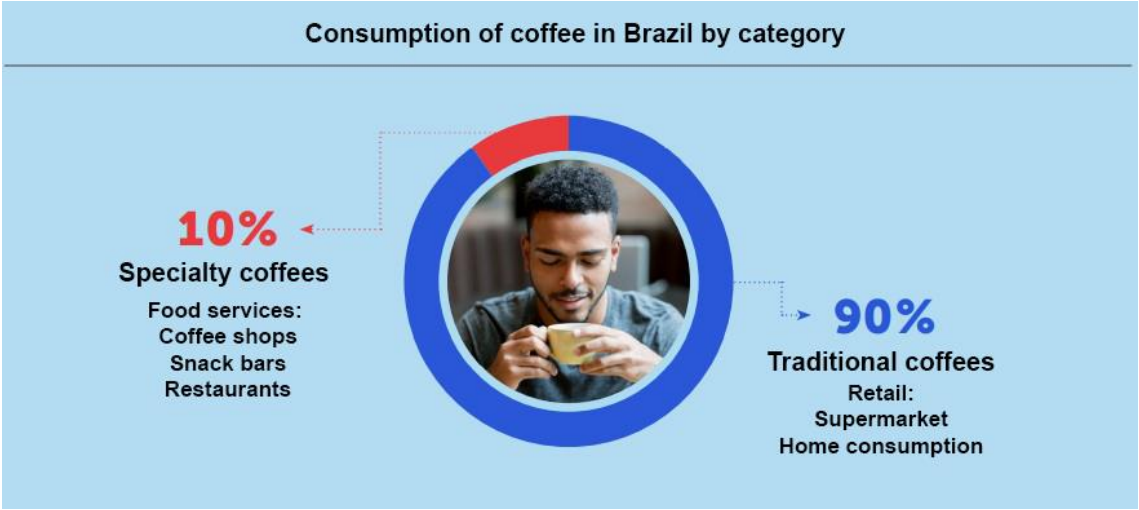

Figure 4. Consumption by coffee category in Brazil. Source: Sebrae-SC [3]

Who is the consumer?

According to a study published by the Federal University of Lavras (UFLA) in partnership with the Agronomic Institute of Campinas (IAC), the consumer profile of specialty coffees is interested in new methods of preparation, concerned with the origin and sustainability of the product [5]. They are divided into two groups:

1 - Coffee lovers - those who go to coffee shops and look for a good environment, with the concept behind the final delivery of the product, and where exchanges of experiences take place.

2 - People who consume specialty coffees, but buy through retail, either in large supermarket chains, or over the internet.

However, in street coffee shops these consumers are not always the majority that drives sales and we should not discard the average consumer, who only works in the neighborhood and does not have much time for a long experience. We must not forget that the main reason why someone on the streets enters a coffee shop is **convenience** [6] and this is strongly related with the business location.

Analytic approach

Looking for good places means finding answer for a lot of important questions about the coffee shop location. Some key questions were listed below. All of them were extracted from the book *"the Daily Grind: How to open and run a coffee shop that makes money"* by Andrew & Claire Bowen [6].

Questions

Q1 – How many competitors within 500 meters from the location?

Q2 – Will you add variety to the area or are you a direct competitor?

Q3 – complementary businesses or services nearby?

Q4 – What is the times of high passing trade? (morning, Lunch, Afternoon, All day)

Q5 – Corner site?

Q6 – On the sunny side of the street?

It is important to mention that busy local cafés are sometimes a good thing [6]. It's not always a bad thing to have a lot of other cafés close to you, as often the area becomes a destination for people when they are hungry or thirsty. Actually, when we look close to United States market, we see that where there is coffee (and donuts), there tends to be a lot of locations [7]. You're not going to find several Pizza Huts in close vicinity to each other, but it's fairly common with Starbucks.

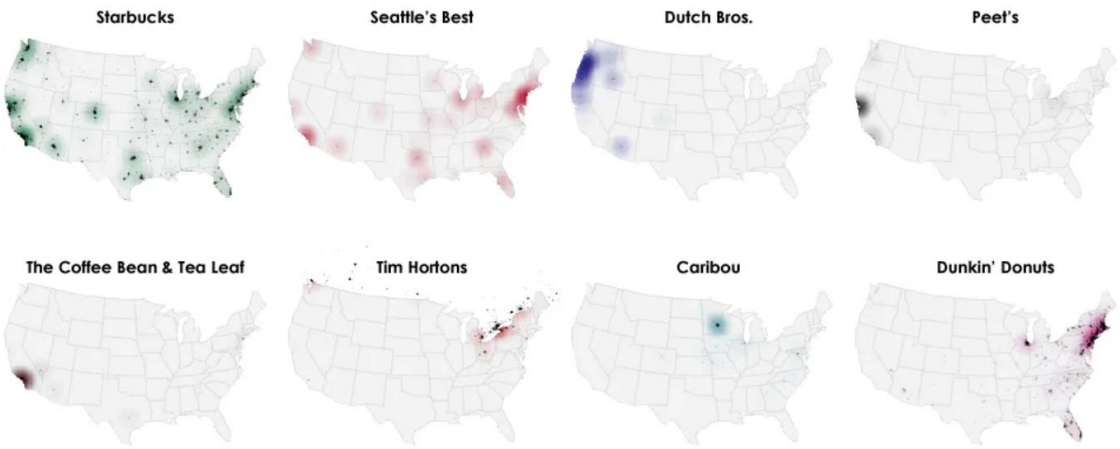

Figure 5. maps of coffee shop places in United States. It shows the nearest coffee place among the popular chains, within a 10-mile radius. Source: [7]

Data Requirements

To find answers to the questions listed above, we first define our data requirements as follow:

1 – Geospatial data from Rio de Janeiro city including neighborhoods and street names.

2 – Geospatial data of existing coffee shops and nearby complementary services with information about popularity, services/menu, ratings and time of high passing trade

Data Collection

The government of Brazil provides a lot of interesting geospatial data in the form of shapefiles through the site of Brazilian Institute of Geography and Statistics (IBGE) where we can download shapefiles of Rio de Janeiro city including neighborhoods and street names.

All geospatial data about coffee shops and complementary services in the city including popularity, services/menu, ratings, and time of high passing trade (trending) will be collected by using Foursquare API or Google Places API.

Planned methodology

All data will be consolidated into a Pandas Dataframe allowing for some quick analysis. We will try to use the sklearn library to group neighborhoods by category of venues and plot them using the Folium library to see distribution, potential service gaps and saturated locations.

We will research if there is a relationship between some kind of venues around a place and its popularity. Try to map and classify the time of high-pass trade through neighborhoods and streets. Check the service and menu of popular coffee shops to see if there is any complementary service that we can do nearby, etc.

These are just a few ideas that have emerged so far. Next week, new ideas are likely to be put on the table and we'll explain why some of the previous ideas were successful (or not).

References

[1] "revistacafeicultura.com.br," Cafeicultura, jun 2019. [Online]. Available: https://revistacafeicultura.com.br/?mat=67988. [Accessed feb 2021].

[2] "https://usda.library.cornell.edu/," U.S. Department of Agriculture (USDA), jun 2019. [Online]. Available: https://downloads.usda.library.cornell.edu/usda-esmis/files/m900nt40f/xk81jw68v/kp78gs60d/coffee.pdf. [Accessed feb 2121].

[3] "atendimento.sebrae-sc.com.br," Sebrae-SC, dec 2019. [Online]. Available: https://atendimento.sebrae-sc.com.br/inteligencia/potencial-de-mercado/Video-o-mercado-de-cafes-e-cafeterias. [Accessed feb 2021].

[4] "www.abic.com.br," Associação Brasileira da indústria de Café (ABIC), jan 2020. [Online]. Available: https://www.abic.com.br/wp-content/uploads/2020/01/Euromonitor_Coffe-Market-Trends-in-Brazil_-Encafe.pdf. [Accessed feb 2021].

[5] "revistaespresso.com.br," Espresso, jan 2018. [Online]. Available: https://revistaespresso.com.br/2018/01/31/pesquisa-busca-compreender-o-comportamento-dos-consumidores-de-cafes-especiais-no-brasil/. [Accessed feb 2021].

[6] C. B. Andrew Bowen, The Daily Grind. How to open & run a coffee shop that makes money, CreateSpace Independent Publishing Platform, 2015.

[7] "flowingdata.com," Flowingdata, mar 2014. [Online]. Available: https://flowingdata.com/2014/03/18/coffee-place-geography/. [Accessed feb 2021].