**Part 1) A description of the problem and a discussion of the background.**

**Market Analysis: A new café in Houston, TX**

**Background of the entrepreneur**

A Brazilian that grow up at a country famous for its coffee (but is a tea lover!), an admirer of the art of baking, cooking and bringing a smile to people. A production engineer that always loved cooking and knows that a successful business starts with good planning.

With the support of the love of her life (and a brilliant business mind), it’s time to put this planning in motion!

A very important part of this business plan is the market analysis, in which we will use real data to validate the hypothesis that the entrepreneur came up after years of observation of cafés around the world.

**Hypothesis**

Although cafes are not a new business and the market might have a big number of them, the idea of this new café is to bring a new experience for your coffee moment, taking it away from the commoditization of the caffeine drink and providing the customers with new flavors and experiences, that won’t be restricted by coffee.

So, let’s start talking about coffee. Being born and bred in Brazil, this is not a new topic for me. You could dream of coffee at any time of the day and my grandma would have it ready and hot for you. But even in a country that had its history written by coffee trading, the behavior of coffee drinkers has been going through deeply transformation. In the past if you wanted a coffee, there you had it: filtered black coffee. Okay, maybe not that simple. You could order a coffee with milk (*Caffe au Lait*) or a *carioca* (which is similar to an americano). It was not common knowledge that the taste from different coffee beans and roasts could bring such a diversity to the once commoditized drink.

Reality is different now. People are not only going for the coffee as a pure habit to get their caffeine dosage, but each day more and more people are learning to savor this drink. This behavior that is far from being restricted to Brazil. The moves of two big international chains – Starbucks and McDonald’s – are proof of this transformation. Starbucks launched their blond roast and created the new brand called Starbucks Reserve that promises *the rarest most extraordinary coffee Starbucks has to offer.* McDonald’s, the seller of simplicity and low prices, realized it needed to step up and started introducing Barista Style coffees at some stores.

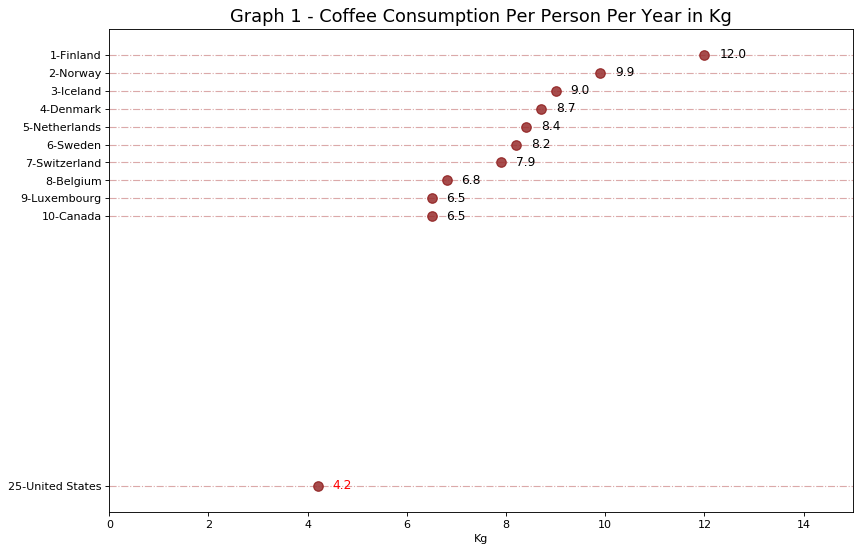
A similar wave is being observed at the tea market. According to the FAO *“Over the next decade, Western countries in general are expected to see lower consumption growth. In the UK, for instance, tea consumption is projected to decrease as black tea struggles to maintain consumers’ interest amid increased competition from other beverages, including coffee.”*

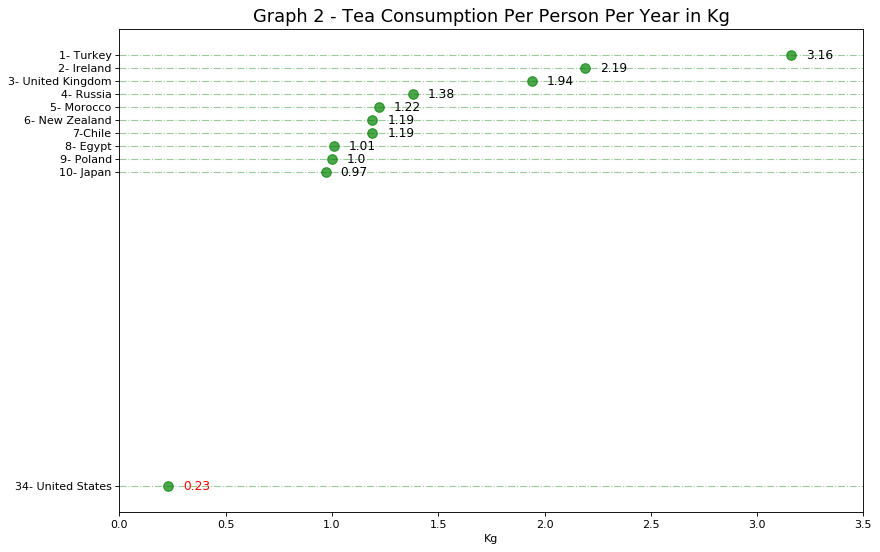
But the FAO’s tea group believes tea companies can stall or even reverse these trends by diversifying into other segments, such as organic and specialty teas, and promoting health and wellbeing benefits.

*“The demand for tea has accelerated due to the ongoing retail revolution and the growing investment into tea education bringing new clientele to know more about tea, where it is sourced, the benefits of drinking tea, and how to properly brew it. Due to this, loose leaf tea has seen a new relevance in the US. Promoting tea culture-based market development and immersion in the cultural identity of societies across the world should be one of the strategies to sustain and expand consumption.”*

But even with the recent changes in consumption, the United States is far from being the top consumer of either coffee or tea. According to the World Atlas (Graph 1), the USA is only the 25th country in coffee consumption in the world, with a North American consuming an average of 4.2 kg per year, lagging behind Finish by 7.8 kg, and consuming 2.3 kg less than a Canadian. And when it comes to tea, the USA is only the 34th in the ranking of tea consumption (Graph 2).

This data indicates that there is room for growth of consumption with the right product and marketing strategies.





More detail on the source of this data can be found at the notebook XXXX

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

**Understanding the target market**

The target market for the café are people between the ages of 25 to 40 and employed, and the data from the USA Census will be the key to answer which of the Houston neighborhoods the target consumers live.

The python package *CensusData* (<https://jtleider.github.io/censusdata/api.html>) will be used to gather data from the following tables of the American Community Survey (ACS) 5-year estimates (2013-2017):

1. **B19326**: Median Income in the Past 12 Months (in 2015 Inflation-Adjusted Dollars) by Sex by Work Experience in the Past 12 Months for the Population 15 Years and Over With Income
2. **B15001**: SEX BY AGE BY EDUCATIONAL ATTAINMENT FOR THE POPULATION 18 YEARS AND OVER

The ACS-5 (2013-2017) is the most recent of the ACS that contains the data for the granularity needed for this analysis: by zip code.

More detail on this data can be found at XXXX

**Understanding the competitors**

An analysis of the competition will be done using the Foursquare location data. By the number and type of venues in each of the Houston’s neighborhoods, it will be possible to build a competitive analysis based on Porter’s Five Forces Model. And that will provide the information needed to make a decision for which of the 3 generic Porter’s strategies for success (cost leadership, differentiation and focus) should be adopted in order to maximize the chance of success of the new business.

**A description of the problem and a discussion of the background. (15 marks)**

*Clearly define a problem or an idea of your choice, where you would need to leverage the Foursquare location data to solve or execute. Remember that data science problems always target an audience and are meant to help a group of stakeholders solve a problem, so make sure that you explicitly describe your audience and why they would care about your problem.*

*This submission will eventually become your Introduction/Business Problem section in your final report. So I recommend that you push the report (having your Introduction/Business Problem section only for now) to your Github repository and submit a link to it.*

**A description of the data and how it will be used to solve the problem. (15 marks)**

*Describe the data that you will be using to solve the problem or execute your idea. Remember that you will need to use the Foursquare location data to solve the problem or execute your idea. You can absolutely use other datasets in combination with the Foursquare location data. So make sure that you provide adequate explanation and discussion, with examples, of the data that you will be using, even if it is only Foursquare location data.*

*This submission will eventually become your Data section in your final report. So I recommend that you push the report (having your Data section) to your Github repository and submit a link to it.*

**2) How many people would be interested in my offering?**

**3) What is the income range and employment rate?**

**4) Where do your customers live?**

**5) How many similar options are already available to consumers?**

**6) What do potential customers pay for these alternatives?**

<https://london.eater.com/2018/7/30/17629998/mcdonalds-mccafe-coffee-advert-hipster-coffee-spoof>

According to data from the XXX, the per capita consumption of coffee in the USA is not among the top 20. And that is seen as a good opportunity for market development.

**Data of tea consumption**

<https://en.wikipedia.org/wiki/List_of_countries_by_tea_consumption_per_capita>

**Data of coffee consumption**

<https://www.worldatlas.com/articles/top-10-coffee-consuming-nations.html>

<http://www.fao.org/3/a-i4480e.pdf>

<http://www.fao.org/economic/est/est-commodities/tea/en/>

<https://www.beveragedaily.com/Article/2018/05/29/Global-tea-consumption-continues-to-rise-Health-and-wellbeing-benefits-are-the-key-drivers-of-future-growth>

<https://www.sba.gov/business-guide/plan-your-business/market-research-competitive-analysis>

<https://www.businessnewsdaily.com/5446-porters-five-forces.html>

After gathering data from foursquare in order to run a competitive analysis.

*Your competitive analysis should identify your competition by product line or service and market segment. Assess the following characteristics of the competitive landscape:  
Market share  
Strengths and weaknesses  
Your window of opportunity to enter the market  
The importance of your target market to your competitors  
Any barriers that may hinder you as you enter the market  
Indirect or secondary competitors who may impact your success*

Porter's Five Forces model

**1. Competitive rivalry**

**2. Bargaining power of suppliers**

**3. Bargaining power of customers**

**4. Threat of new entrants**

**5. Threat of substitute products or services**

BCG Matrix: I am a question mark  
<https://www.businessnewsdaily.com/5693-bcg-matrix.html>

After that it's time to implement a strategy to expand your competitive advantage.

Strategies for success (Porter’s 3 generic strategies)  
Cost leadership  
Differentiation  
Focus

<https://en.wikipedia.org/wiki/List_of_Houston_neighborhoods>

<https://www.houstontx.gov/superneighborhoods/recognized.html>

<http://www.houstontx.gov/planning/Demographics/super_neighborhoods_2.html>

<https://cohgis-mycity.opendata.arcgis.com/datasets/census-2010-by-superneighborhood/data?geometry=-96.458%2C29.615%2C-94.372%2C30.032>

https://opendata.arcgis.com/datasets/700dc2edd41e484da160c882739a981b\_1.geojson

[https://cohegis.houstontx.gov/cohgispub/rest/services/PD/PD\_Census\_Block\_Group\_Demographics\_wm/MapServer/1/query?outFields=\*&where=1%3D1](https://cohegis.houstontx.gov/cohgispub/rest/services/PD/PD_Census_Block_Group_Demographics_wm/MapServer/1/query?outFields=*&where=1%3D1)

<https://hbr.org/2018/05/when-should-entrepreneurs-write-their-business-plans>

<https://www.entrepreneur.com/article/76140>

<https://www.forbes.com/sites/alejandrocremades/2018/07/24/how-to-create-a-business-plan/>

List of countries by coffee consumption per capita

<https://fusiontables.google.com/DataSource?docid=1C-fn6nSe21acP0xJIO1T1x0wohqfMYCQyJjbqdk#rows:id=1>

List of countries by tea consumption per capita

<https://en.wikipedia.org/wiki/List_of_countries_by_tea_consumption_per_capita>

<https://factfinder.census.gov/faces/nav/jsf/pages/searchresults.xhtml?refresh=t#none>

<http://www.city-data.com/city/Houston-Texas.html>

de para neighborhood >> zip code

<http://www.mccannproperties.com/Houston-Neighborhood-ZipCodes.php>

How to develop market? Take coffee away from the simple idea of a caffeine boost to go through the day and bring real flavors to it. It is notable that this is not a new idea with big chains like Starbucks and McDonalds investing in special roasts and barista style coffees. But why do we think we can be different?

Awesome tea and ok coffee. Awesome coffee and ok tea. Why?

Because our idea is to provide people not only with a full of flavor and texture coffee, but serve it together with a few but well chosen options of good food, at an inviting environment with cozy music, attracting both people who wants to take a great coffee to take away and those looking for a convenient place for a social gathering, casual work meeting, working on a computer, reading a book or just passing the time.

Those changes might have changed the quality of the drink, but I am here talking about a good quality of the drink, together with a good variety of

For the second week, the final deliverables of the project will be:

A link to your Notebook on your Github repository, showing your code. (15 marks)

A full report consisting of all of the following components (15 marks):

* Introduction where you discuss the business problem and who would be interested in this project.
* Data where you describe the data that will be used to solve the problem and the source of the data.
* Methodology section which represents the main component of the report where you discuss and describe any exploratory data analysis that you did, any inferential statistical testing that you performed, if any, and what machine learnings were used and why.
* Results section where you discuss the results.
* Discussion section where you discuss any observations you noted and any recommendations you can make based on the results.
* Conclusion section where you conclude the report.

Your choice of a presentation or blogpost. (10 marks)