

Project: Manhattan Coffee Shop

Yin Su

August 16th, 2020

Introduction

Since I live in Manhattan and would like to explore the coffee culture here, the business problem of this project will be “Where to open a coffee shop in Manhattan?”. This project is going to target stakeholders who are interested in investing in a coffee shop in Manhattan. A coffee shop promotes the innovation of the understanding of the coffee and provides an atmosphere that comforts customers in the public sector. However, the cost of opening a coffee shop varies based on store locations in Manhattan. We can use the data science tools that are covered in the Coursera/IBM course Data Science to make our choices, such as data visualization, Foursquare API and machine learning. Finally, we will discuss the results and present the project to investors. Especially, showing investors the recommendations.

Data

List of our data source below:

- Coffee data from the US Department of Agriculture
Use: analyze the US domestic coffee consumption trend
Source: <https://apps.fas.usda.gov/psdonline/app/index.html#/app/downloads>
- Geoinformation accessed from the Foursquare API
Use: access geolocation details of manhattan
Source: <https://foursquare.com/developers/explore>
- Manhattan neighborhood name data from NYU Spatial Data Repository
Use: access neighborhoods location details of manhattan
Source: <https://apps.fas.usda.gov/psdonline/app/index.html#/app/downloads>

Methodology

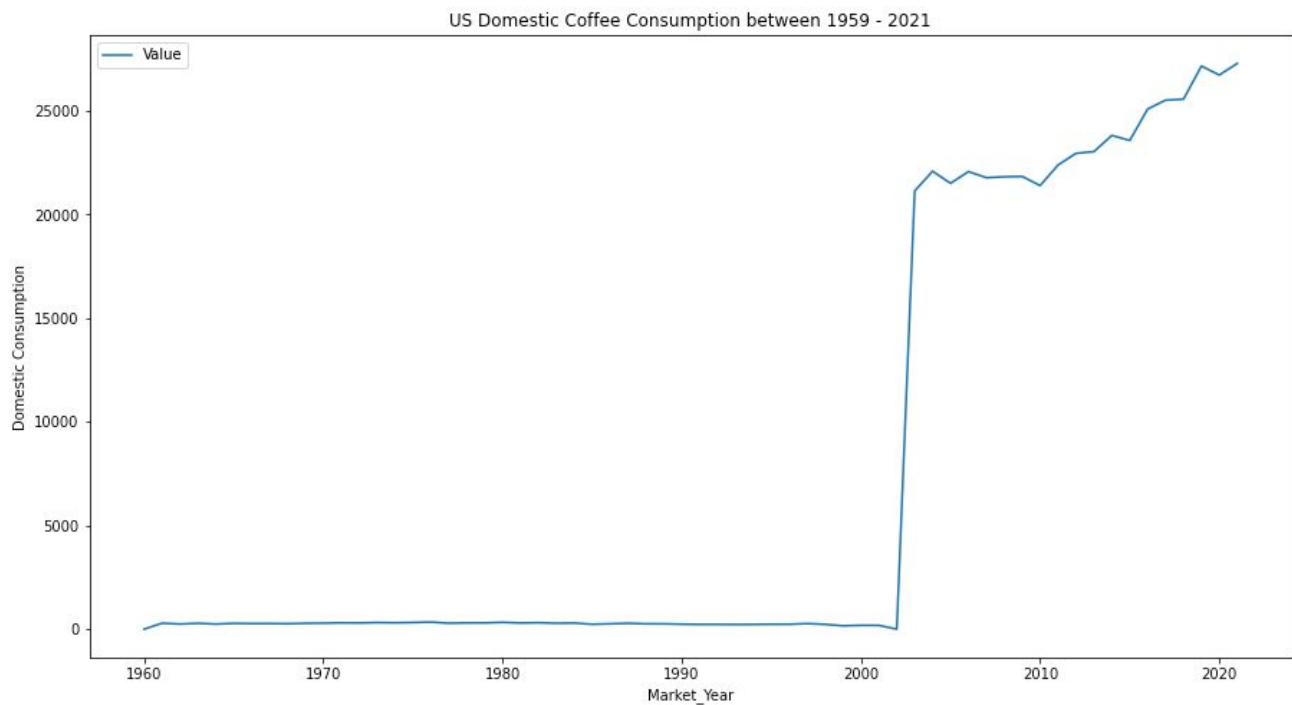
In the methodology session, we will use data visualization, Foursquare API and machine learning. For each method, we will have the details below.

Data visualization

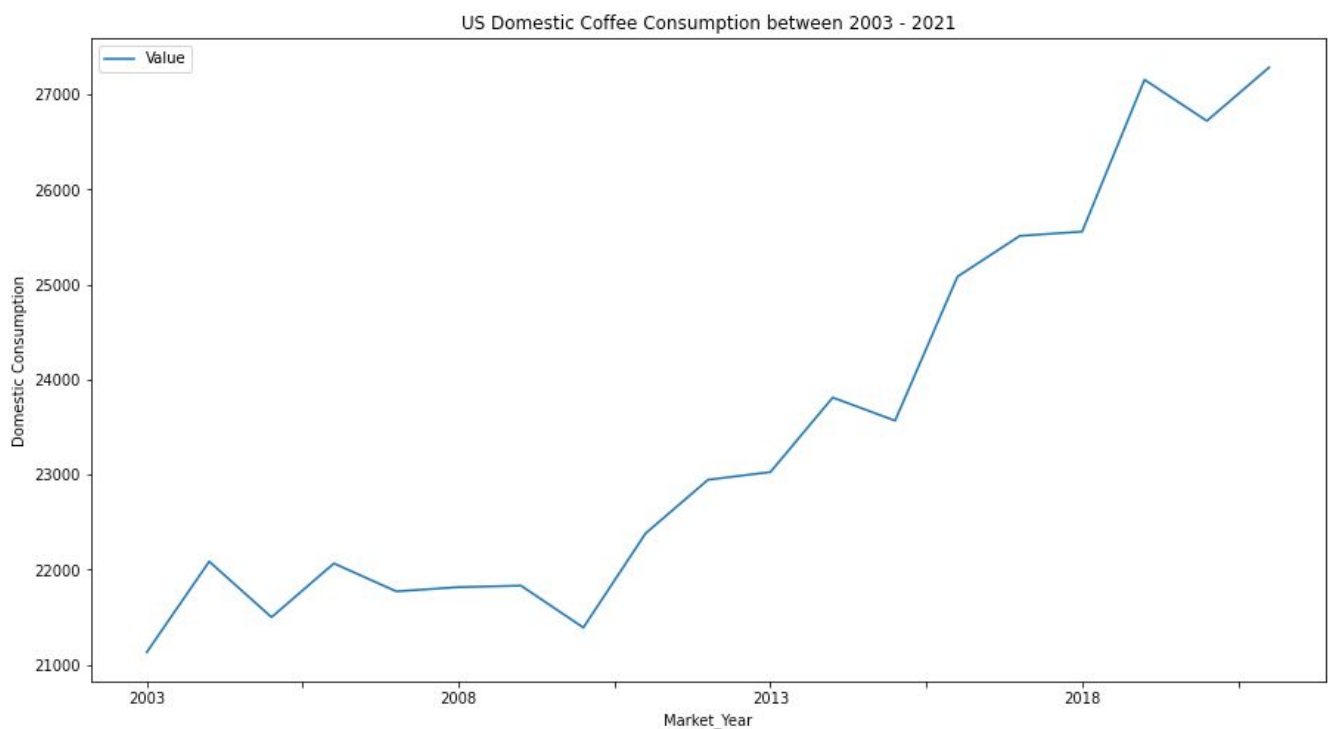
Three data visualization tools are used to show the US domestic coffee consumption, including line pots, horizontal bar chart and pie chart.

- Line Pots

Visualize the US domestic consumption between 1959 -2021

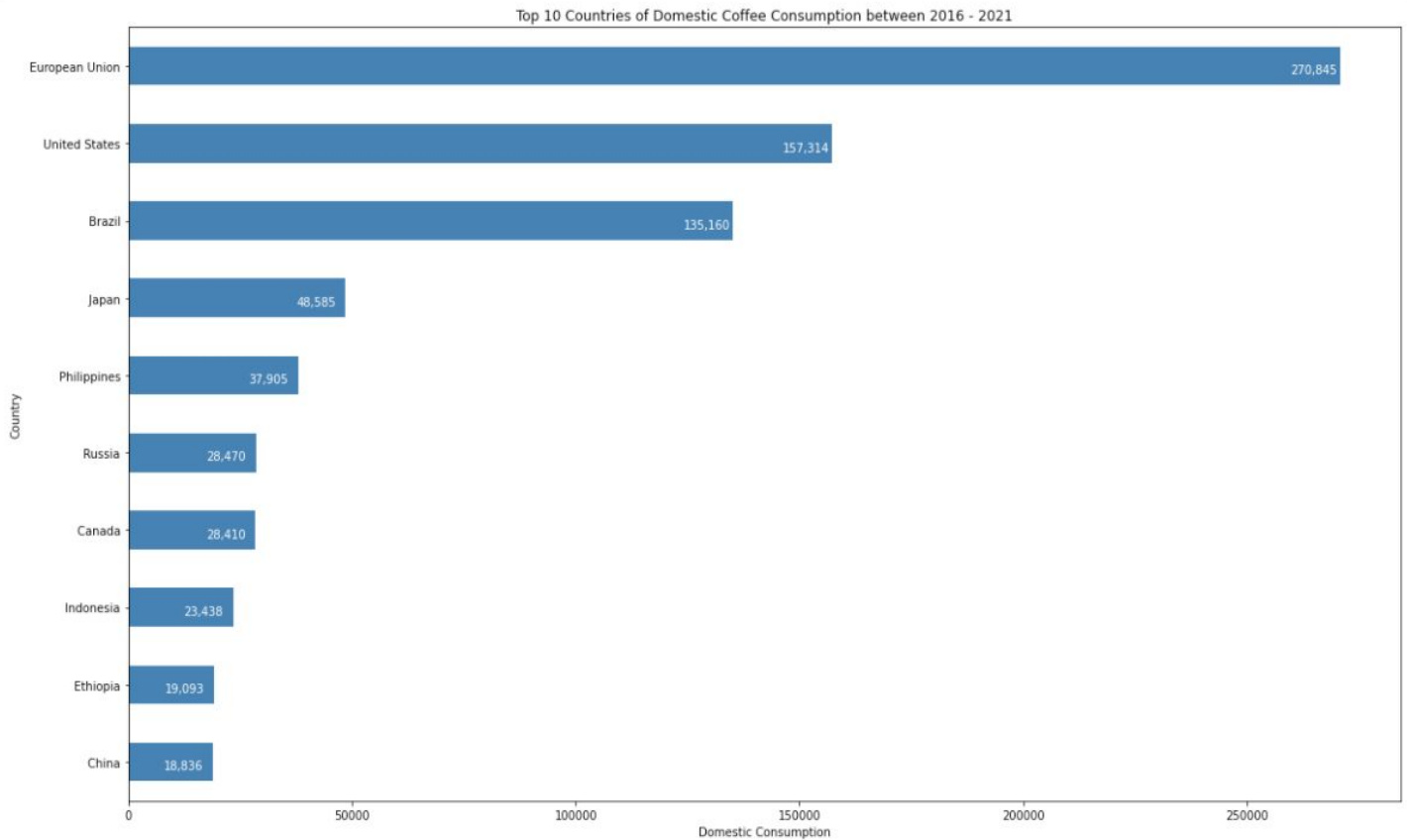


Visualize the US domestic consumption between 2003 - 2021



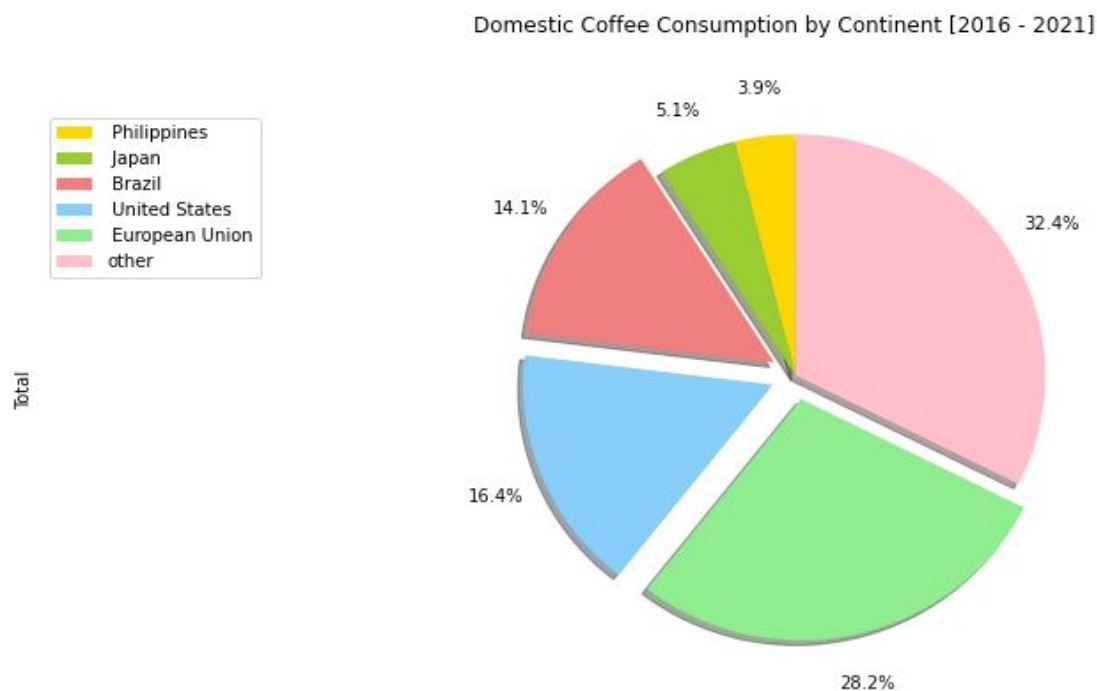
- Horizontal Bar Plot

Visualize the top 10 countries of domestic coffee consumption between 2016 - 2021



- Pie Chart

Visualize the domestic coffee consumption by continent [2016 - 2021]



Foursquare API

The Foursquare API provides the free service for us to use CLIENT_ID, CLIENT_SECRET, and VERSION to get the nearby venues of manhattan. We have the chance to get the data frame of nearby venues of each neighbourhood in manhattan.

```
venues.head()
```

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Marble Hill	40.876551	-73.91066	Arturo's	40.874412	-73.910271	Pizza Place
1	Marble Hill	40.876551	-73.91066	Bikram Yoga	40.876844	-73.906204	Yoga Studio
2	Marble Hill	40.876551	-73.91066	Tibbett Diner	40.880404	-73.908937	Diner
3	Marble Hill	40.876551	-73.91066	Dunkin'	40.877136	-73.906666	Donut Shop
4	Marble Hill	40.876551	-73.91066	Starbucks	40.877531	-73.905582	Coffee Shop

Machine Learning

Two machine learning algorithms are used for clustering coffee shop neighborhoods in manhattan, including one hot encoding and k neighbourhood cluster.

- One Hot Encoding

Encode categorical features as a one-hot numeric array

```
] :
```

	Neighborhood	Accessories Store	Adult Boutique	Afghan Restaurant	African Restaurant	American Restaurant	Antique Shop	Arepa Restaurant
0	Battery Park City	0.0	0.0	0.0	0.000000	0.013889	0.0	0.0
1	Carnegie Hill	0.0	0.0	0.0	0.000000	0.011364	0.0	0.0
2	Central Harlem	0.0	0.0	0.0	0.066667	0.044444	0.0	0.0
3	Chelsea	0.0	0.0	0.0	0.000000	0.040000	0.0	0.0
4	Chinatown	0.0	0.0	0.0	0.000000	0.030000	0.0	0.0

Create the data frame of the top 10 venues of the manhattan neighborhood

```
:
```

	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
0	Battery Park City	Park	Hotel	Gym	Coffee Shop	Shopping Mall
1	Carnegie Hill	Coffee Shop	Café	Yoga Studio	Bakery	Wine Shop
2	Central Harlem	African Restaurant	Chinese Restaurant	French Restaurant	Gym / Fitness Center	American Restaurant
3	Chelsea	Coffee Shop	Art Gallery	American Restaurant	Seafood Restaurant	Hotel
4	Chinatown	Chinese Restaurant	Bakery	Cocktail Bar	Dessert Shop	Vietnamese Restaurant

- K Neighbourhood Cluster

Cluster the neighborhoods in manhattan based on the coffee shop frequency

	Borough	Neighborhood	Latitude	Longitude	Cluster Labels	Coffee Shop
0	Manhattan	Marble Hill	40.876551	-73.910660	2	0.130435
1	Manhattan	Chinatown	40.715618	-73.994279	4	0.020000
2	Manhattan	Washington Heights	40.851903	-73.936900	4	0.024096
3	Manhattan	Inwood	40.867684	-73.921210	4	0.016949
4	Manhattan	Hamilton Heights	40.823604	-73.949688	3	0.064516

Create the map of coffee shop neighbourhoods in manhattan



Results

Results include two parts: US domestic coffee consumption & Manhattan coffee shop

US domestic coffee consumption

1. US domestic coffee consumption jumps from the year 2003 based on the line plots. The trend is that domestic consumption is **ascending**.
2. US domestic coffee consumption between 2016 - 2021 is **ranked 2rd**, based on the horizontal bar plot. The 1st is the European Union.
3. US domestic coffee consumption by continent between 2016 - 2021 is **16.4%**, **ranked 2rd**. The 1st is the European Union, which is 28.2%.

Manhattan coffee shop

1. The category 'Coffee shop' is **a most common venue** among neighborhoods in Manhattan, based on the one-hot encoding.
2. Manhattan neighborhoods are clustered based on the frequency of the coffee shop venue. The map of the neighborhoods shows the clusters in different colors and represents the different popularities of the coffee shop. The Neighbourhood **'Financial District'** on the map shows the significant popularities compared to other neighborhoods. The cluster label 3 shows a large number of neighborhoods cover **shopping areas**, like SoHo, Fifth Avenue, Times Square, etc.,

Discussion

After carefully analyzing the data retrieved from the US Agriculture Department, the US is the second largest country that consumes 157,314 'Thousand 60-Kilogram Bags' green coffee since 2016. Also, Manhattan is so called 'the city that never sleeps'. This project of opening a coffee shop is **with high possibility**.

Recommendations

The location of the coffee shop can be the financial district or shopping areas in Manhattan.

- Financial District
Wall Street must be the best location for a brand opening coffee shop with espressos.
- Shopping Area
SoHo is better than Fifth Avenue because the luxury shops on Fifth Avenue usually provide cocktails and coffee to customers. It's much convenient for customers to step for a cup of coffee among the shops in SoHo.

Conclusion Section

The top 10 most common venues data frame shows that Manhattan is a great place for living and tourism. Coffee shops in Manhattan provide people here a relaxing atmosphere and an optional working place with modern/old fashion design. In conclusion, the project of Manhattan Coffee Shop is designed for the Coursera/IBM course 'Applied Data Science Capstone'. This project helps me not only understand the machine learning algorithms, but also apply them in a real-life project.