Predicting the best place to start your new business

Predicting the best place to start your new business

Business Problem and Stakeholders:

- Someone recently moved to Portugal and this person owns a franchising brand of restaurants. This person wants to know which will be the best city to fit his new restaurant.

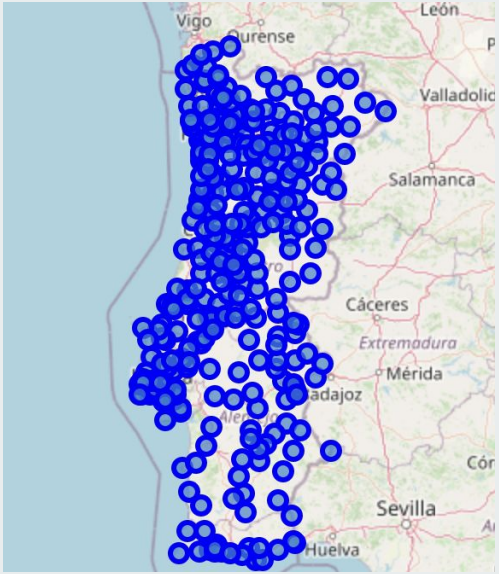
- Besides many other variables that might induct the person to decide where to start this new business, we will try to help him to decide where to start by creating one index which will consist of the ratio between the population of the city versus the number of restaurants that this city already has.

- This index aims to give to the requester which city may be better once that will have an index with more demand and less offering restaurants.

- **Formula**:
  Potential Index (PI) = Population of each city divided by the number of restaurants

DATA

Mapped cities:



Foursquare's venues by category for each city:

	city	City Latitude	City Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Lisbon	38.7452	-9.1604	Malo Clinic	38.746271	-9.163602	Medical Center
1	Lisbon	38.7452	-9.1604	Mariott Executive Lounge	38.747059	-9.164141	Hotel Bar
2	Lisbon	38.7452	-9.1604	Lucimar	38.743135	-9.155755	Portuguese Restaurant
3	Lisbon	38.7452	-9.1604	Escritórios NOS - Café	38.744550	-9.158933	Cafeteria
4	Lisbon	38.7452	-9.1604	Restaurante Santa Maria	38.745503	-9.158635	BBQ Joint

```
pt_venues.groupby('city').count()
```

city	City Latitude	City Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
Abrantes	4	4	4	4	4	4
Agualva	7	7	7	7	7	7
Aguiar da Beira	4	4	4	4	4	4
Alandroal	4	4	4	4	4	4
Albergaria-a-Velha	6	6	6	6	6	6
Albufeira	54	54	54	54	54	54
Alcanena	4	4	4	4	4	4
Alcobaça	12	12	12	12	12	12
Alcochete	14	14	14	14	14	14
Alcoutim	4	4	4	4	4	4

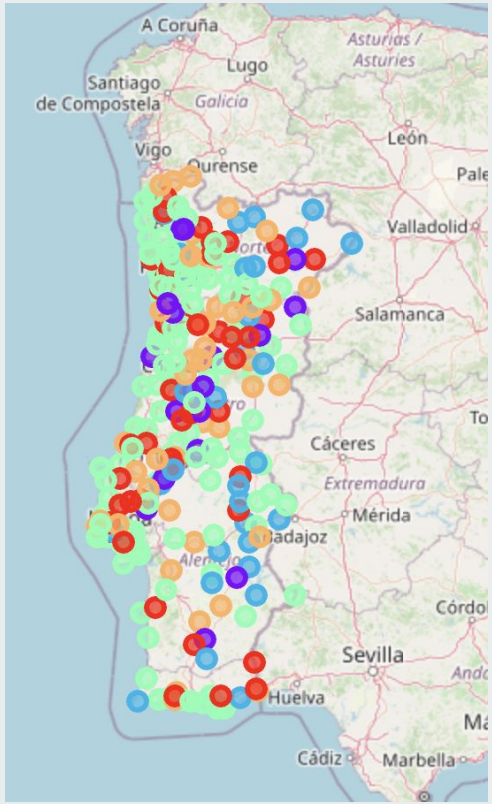
DATA FRAMES GENERATED TO CLUSTER THE VENUES' CATEGORIES

Most common categories for each city:

	city	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	Abrantes	Bar	Historic Site	Supermarket	Grocery Store	Other Nightlife	Mountain	Movie Theater	Multiplex	Museum	Music Venue
1	Agualva	Accessories Store	Supermarket	Café	Bakery	Train Station	Coffee Shop	Electronics Store	Music Venue	National Park	Neighborhood
2	Aguiar da Beira	Portuguese Restaurant	Soccer Field	Food	Bar	Accessories Store	Other Great Outdoors	Mountain	Movie Theater	Multiplex	Museum
3	Alandroal	Hotel	Historic Site	Portuguese Restaurant	Restaurant	Accessories Store	Other Nightlife	Movie Theater	Multiplex	Museum	Music Venue
4	Albergaria-a-Velha	Theater	Garden	Restaurant	Market	Bakery	Café	Other Great Outdoors	Movie Theater	Multiplex	Museum
5	Albufeira	Portuguese Restaurant	Bar	Restaurant	Mediterranean Restaurant	Seafood Restaurant	Hotel	Diner	Grocery Store	Beach	Snack Place

DATA FRAMES GENERATED TO CLUSTER AND TO GET INFO

Clustering the cities by their most common venues category



COUNCURING RESTAURANTS BY CITY

Filtering only to get restaurants that might concur with our future new restaurant

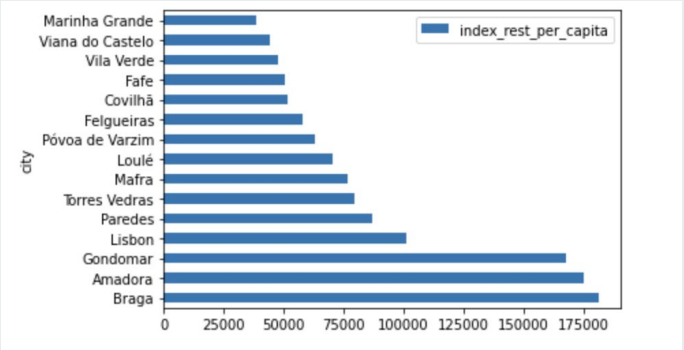
```
pt_venues_filtered = pt_venues[pt_venues['Venue Category'].str.contains('Restaurant', regex=True, na=True)]
pt_venues_filtered.head()
```

	city	City Latitude	City Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
2	Lisbon	38.7452	-9.1604	Lucimar	38.743135	-9.155755	Portuguese Restaurant
7	Lisbon	38.7452	-9.1604	O Táxi	38.743197	-9.156703	Restaurant
8	Lisbon	38.7452	-9.1604	Citrus Restaurant - Marriott	38.746986	-9.164236	Portuguese Restaurant
9	Lisbon	38.7452	-9.1604	Mister Papas	38.745423	-9.163669	Italian Restaurant
12	Lisbon	38.7452	-9.1604	Restaurante o 80	38.743048	-9.157362	Portuguese Restaurant

POTENTIAL INDEX AND CLUSTERS EVALUATION

Getting the most fertile place to start the new business after joining the population data and the amount of restaurants of each city to apply our Potential Index

city	lat	lng	country	iso2	admin_name	capital	population	population_proper	countRests	index_rest_per_capita
Braga	41.5333	-8.4167	Portugal	PT	Braga	admin	181494.0	181494.0	1.0	181494.000000
Amadora	38.7500	-9.2333	Portugal	PT	Lisboa	minor	175136.0	175136.0	1.0	175136.000000
Gondomar	41.1500	-8.5333	Portugal	PT	Porto	minor	168027.0	168027.0	1.0	168027.000000
Lisbon	38.7452	-9.1604	Portugal	PT	Lisboa	primary	506654.0	506654.0	5.0	101330.800000
Paredes	41.2000	-8.3333	Portugal	PT	Porto	minor	86854.0	86854.0	1.0	86854.000000
Torres Vedras	39.0833	-9.2667	Portugal	PT	Lisboa	minor	79465.0	79465.0	1.0	79465.000000
Mafra	38.9333	-9.3333	Portugal	PT	Lisboa	minor	76685.0	76685.0	1.0	76685.000000
Loulé	37.1440	-8.0235	Portugal	PT	Faro	minor	70622.0	70622.0	1.0	70622.000000
Póvoa de Varzim	41.3916	-8.7571	Portugal	PT	Porto	minor	63408.0	63408.0	1.0	63408.000000
Felgueiras	41.3667	-8.1833	Portugal	PT	Porto	minor	58065.0	58065.0	1.0	58065.000000
Covilhã	40.2833	-7.5000	Portugal	PT	Castelo Branco	minor	51797.0	51797.0	1.0	51797.000000



CONCLUSION AND NEXT STEPS

- Try to get new business data

- Apply the hypothesis into these new scenarios:
  - New kind of categories
  - New countries to predict starting a new business

- The clustering cities can have an additional improvement labeling the population size

THANK YOU

Applied Data Science Capstone - Predicting the best place to start your new business