

Finding a best place for own business

Coursera Final Assignment

1. Introduction

When you come to an unfamiliar place, and then many years have passed, you think you already know the city very well, including the names of every street and every store, but also when you want to create your own When you first opened your store, you were still scared because you did n't know how many kilometers away from a location to start your business.

So I hope to help people find the most suitable place to conduct business through models and machine learning

2. Problem

How to choose a location and what type of restaurants and shops are around will be the problem in this part.

3. Data acquisition and cleaning

First I picked the coordinates in Paris and got it to a Chinese town called belle ville. I used the foursquare API to get the restaurants, theaters, bars, nightclubs and shops already around.

As shown in the figure below, a lot of data has not been extracted and processed, including many Nan-type empty data sets and some messy structured data storage.

	categories	id	location.address	location.cc	location.city	location.country	location.crossStreet	location.distance
0	Park	4adccda0af964a520253421e3	47 rue des Couronnes	FR	Paris	France	NaN	30
1	French Restaurant	4b75e3f0f964a520742c2ee3	12 rue des Envierges	FR	Paris	France	NaN	85
2	Pizza Place	5471f9b3498e6e2065dc3ad4	356 rue des Pyrénées	FR	Paris	France	NaN	261
3	French Restaurant	4adccda14f964a5203c3721e3	3 rue Jouye-Rouve	FR	Paris	France	NaN	228
4	French Restaurant	54dd4f8f498e62876939b098	101 rue des Couronnes	FR	Paris	France	NaN	321

By using pandas and a series of python internal functions to extract important messages, classify collections, delete empty subsets, and organize data institutions to facilitate our future data processing tasks.

	categories	id	city	country	distance	address	lat	lng	postalCode	state	name	categories_id
0	Park	4adcda0af964a520253421e3	Paris	France	30	47 rue des Couronnes	48.871457	2.385014	75020	Île-de-France	Parc de Belleville	0
1	French Restaurant	4b75e3f0f964a520742c2ee3	Paris	France	85	12 rue des Envierges	48.871660	2.386255	75020	Île-de-France	Le Vieux Belleville	1
2	Pizza Place	5471f9b3498e6e2065dc3ad4	Paris	France	261	356 rue des Pyrénées	48.873918	2.386359	75020	Île-de-France	Il Posto	2
3	French Restaurant	4adcda14f964a5203c3721e3	Paris	France	228	3 rue Jouye-Rouve	48.873098	2.382767	75020	Île-de-France	Le Baratin	1
4	French Restaurant	54dd4f8f498e62876939b098	Paris	France	321	101 rue des Couronnes	48.872277	2.389393	75020	Île-de-France	Le Jourdain	1

Let's first explore how many business types are around the city, such as bars, restaurants, cafes or shopping malls. We can clearly see through the following analysis that there are more bars and French restaurants in this area.

```
venue_df['categories'].value_counts(normalize=True).head()
```

```
]: Bar                0.195876
   French Restaurant   0.092784
   Café                0.051546
   Italian Restaurant  0.051546
   Pizza Place         0.030928
   Name: categories, dtype: float64
```

Through distance, we can also find some hidden information from it, such as how far is it to open a bar, the best business, or a Chinese restaurant.

distance	address
30	47 rue des Couronnes
85	12 rue des Envierges
261	356 rue des Pyrénées
228	3 rue Jouye-Rouve
321	101 rue des Couronnes