

AN INVESTIGATION TO FIND THE MOST SUITABLE PLACE TO OPEN A COFFEE SHOP IN IBARAKI, JAPAN

IBM Data Science Capstone Project

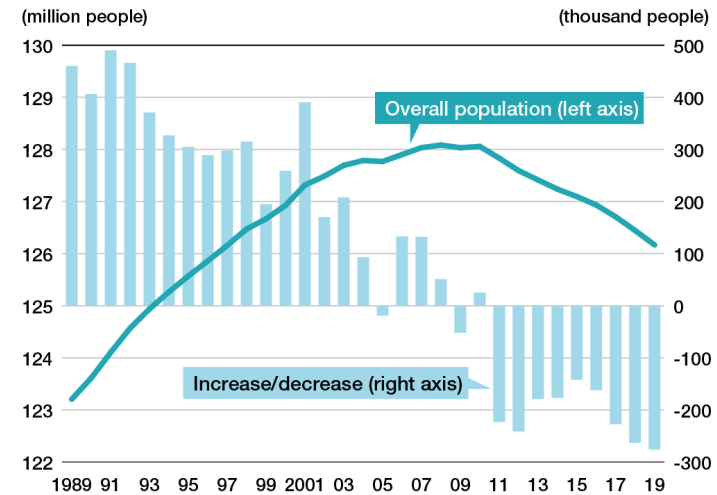
Nguyen Quang Duong
2020, May, 31th

Back Ground



Location of Ibaraki
(near Tokyo)

Population Change in Japan

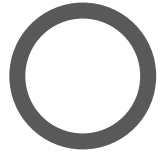


Created by *Nippon.com* based on data from the population estimate of the Ministry of Internal Affairs and Communications (as of October 1, 2019). nippon.com

Population decrease in Japan

Business Problem

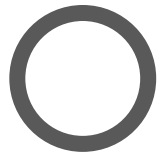
The coffee shop location must clear below criteria



The coffee shop should be located in city which in top populated cities in Ibaraki.



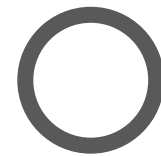
The location should be near Tokyo, so South of Ibaraki is preferable.



The city's population should be stable or increasing (many people coming to this city to cover the natural decrease cause by low birth rate & aging).

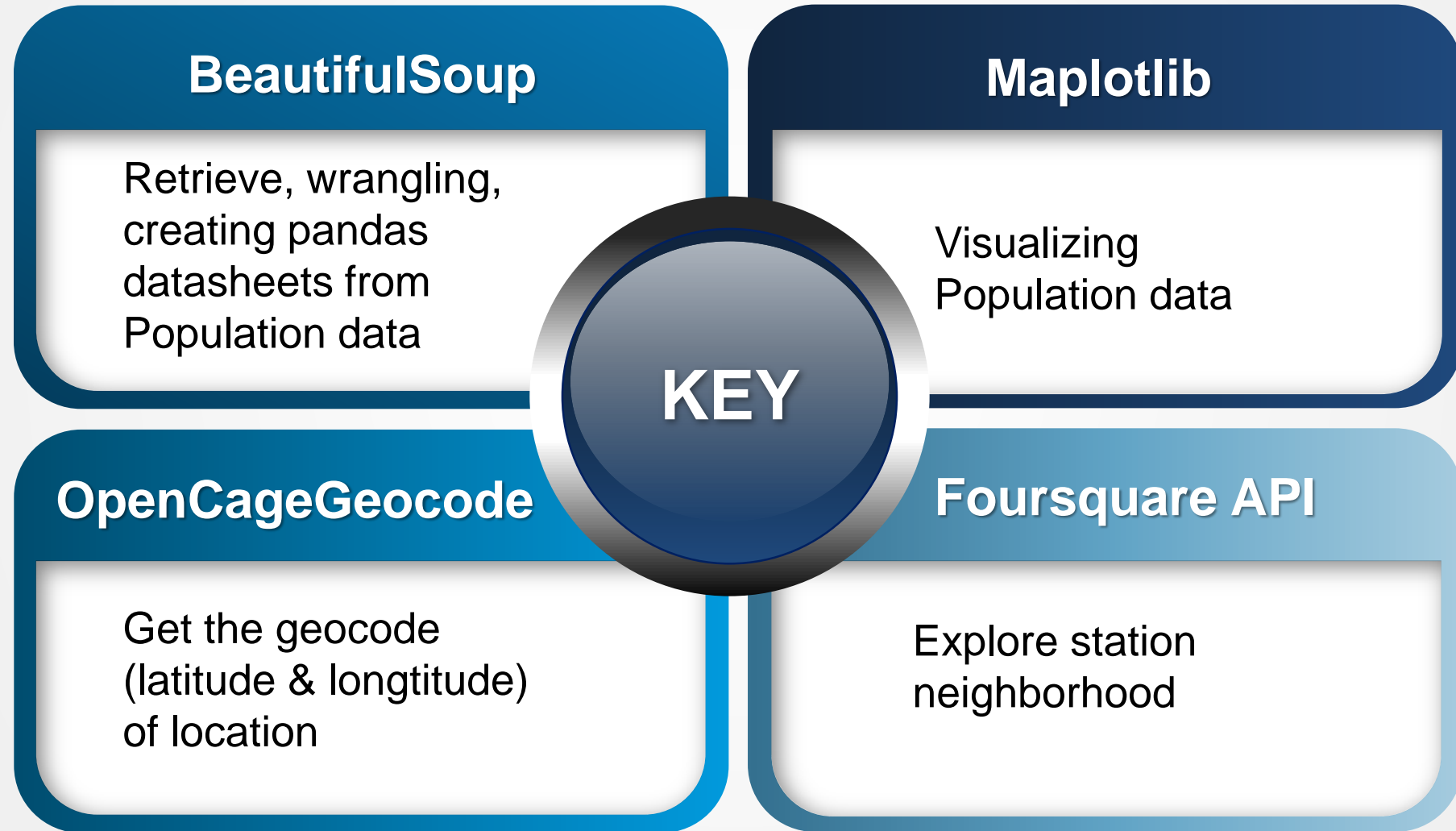


The location should be near some train stations.



There are not too many competitors in the area.

Key Python Library



Step to the conclusion

Step 1

- Data retrieval, exploration and wrangling
- URL:<https://www.pref.ibaraki.jp/kikaku/tokei/fukyu/tokei/betsu/jinko/getsu/jinko2001.html>

Step 2

- Data visualization
- To clarify the most attractive city in Ibaraki

Step 3

- Explore the train stations neighborhood in the city
- To clarify the most appropriate station

Step 1: Data retrieval, exploration and wrangling

Beautifulsoup library was used to retrieve and wrangle the data

Before

Table on the website (Japanese)

年月日	世帯数	人口			前月中の人口移動						
		総数	男	女	増減	自然動態			社会動態		
						自然増減	出生	死亡	社会増減	転入	転出
平成31年1月1日	1,163,338	2,881,020	1,438,291	1,442,729	-1,120	-1,276	1,553	2,829	156	8,057	7,901
平成31年2月1日	1,163,573	2,879,225	1,437,482	1,441,743	-1,795	-2,058	1,667	3,725	263	8,297	8,034
平成31年3月1日	1,163,464	2,877,206	1,436,506	1,440,700	-2,019	-1,576	1,446	3,022	-443	8,451	8,894
平成31年4月1日	1,164,984	2,870,883	1,433,317	1,437,566	-6,323	-1,293	1,410	2,703	-5,030	18,524	23,554
令和元年5月1日	1,168,885	2,871,872	1,434,181	1,437,691	989	-1,200	1,426	2,626	2,189	15,840	13,651
令和元年6月1日	1,170,033	2,871,199	1,433,974	1,437,225	-673	-1,330	1,691	3,021	657	9,623	8,966
令和元年7月1日	1,170,582	2,870,018	1,433,374	1,436,644	-1,181	-811	1,443	2,254	-370	7,904	8,274
令和元年8月1日	1,171,498	2,869,312	1,433,170	1,436,142	-706	-1,071	1,587	2,658	365	9,314	8,949
令和元年9月1日	1,172,116	2,868,148	1,432,670	1,435,478	-1,164	-1,043	1,597	2,640	-121	8,694	8,815
令和元年10月1日	1,173,307	2,868,041	1,432,563	1,435,478	-107	-1,052	1,586	2,638	945	9,332	8,387
令和元年11月1日	1,174,691	2,868,137	1,432,669	1,435,468	96	-1,069	1,591	2,660	1,165	9,688	8,523
令和元年12月1日	1,175,302	2,867,267	1,432,226	1,435,041	-870	-1,172	1,490	2,662	302	8,117	7,815
令和2年1月1日	1,175,894	2,866,325	1,431,725	1,434,600	-942	-1,322	1,509	2,831	380	8,229	7,849

After

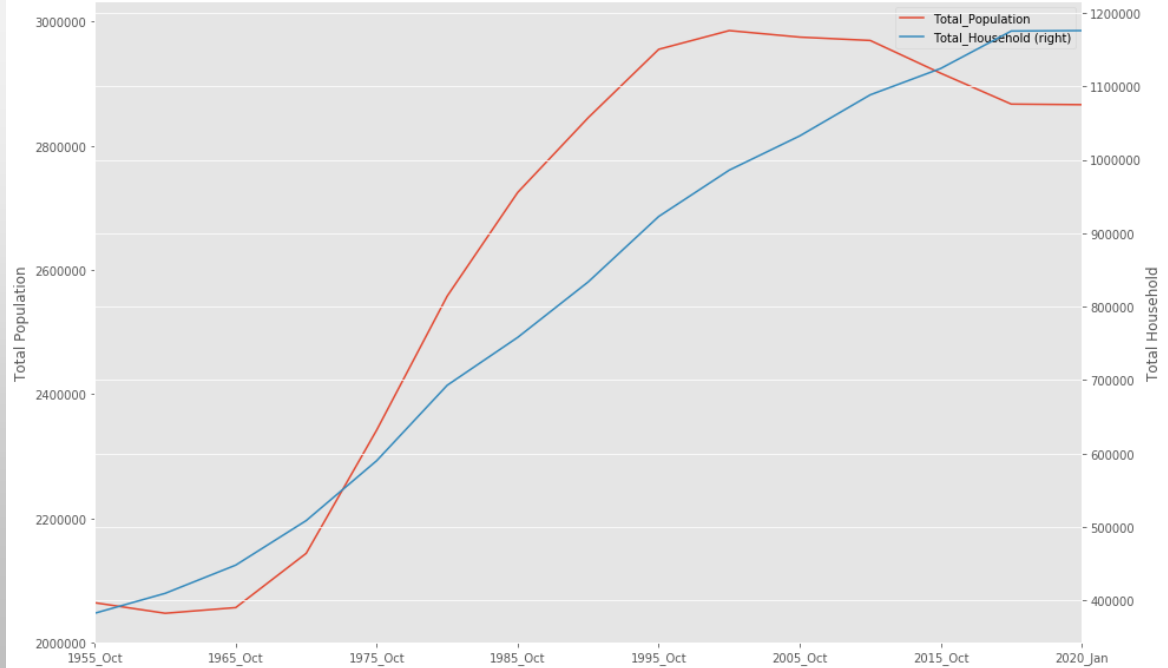
Pandas Data Frame (English)

	Total_Household	Total_Population	Man	Women	Increase_or_Decrease	Natural_Reasons	Birth	Dead	Migration_Reasons	In	Out
2019_Jan	1163338	2881020	1438291	1442729	-1120	-1276	1553	2829	156	8057	7901
2019_Feb	1163573	2879225	1437482	1441743	-1795	-2058	1667	3725	263	8297	8034
2019_Mar	1163464	2877206	1436506	1440700	-2019	-1576	1446	3022	-443	8451	8894
2019_Apr	1164984	2870883	1433317	1437566	-6323	-1293	1410	2703	-5030	18524	23554
2019_May	1168885	2871872	1434181	1437691	989	-1200	1426	2626	2189	15840	13651
2019_Jun	1170033	2871199	1433974	1437225	-673	-1330	1691	3021	657	9623	8966
2019_Jul	1170582	2870018	1433374	1436644	-1181	-811	1443	2254	-370	7904	8274
2019_Aug	1171498	2869312	1433170	1436142	-706	-1071	1587	2658	365	9314	8949
2019_Sep	1172116	2868148	1432670	1435478	-1164	-1043	1597	2640	-121	8694	8815
2019_Oct	1173307	2868041	1432563	1435478	-107	-1052	1586	2638	945	9332	8387
2019_Nov	1174691	2868137	1432669	1435468	96	-1069	1591	2660	1165	9688	8523
2019_Dec	1175302	2867267	1432226	1435041	-870	-1172	1490	2662	302	8117	7815
2020_Jan	1175894	2866325	1431725	1434600	-942	-1322	1509	2831	380	8229	7849

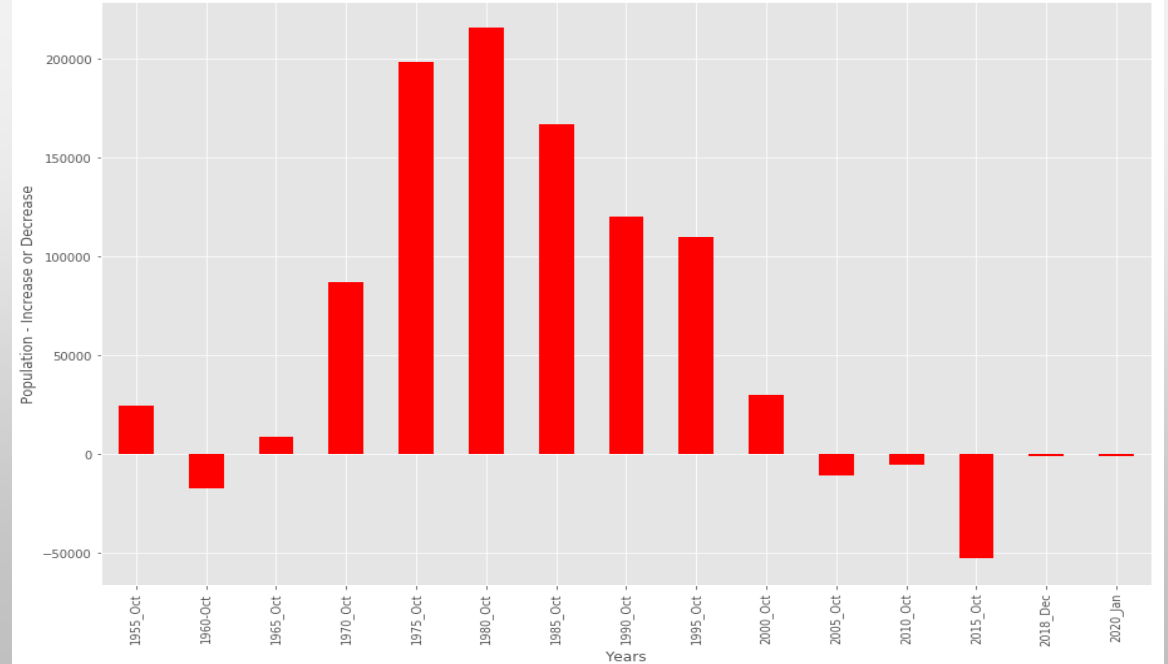
Step 2: Data visualization

The total population in Ibaraki got peak in 2000, after that the population is decreasing lightly and get stable from 2015. May come migration helped covering the decreasing trend caused by low birth rate.

Total Population trending in Ibaraki-prefecture from 1955 to 2019

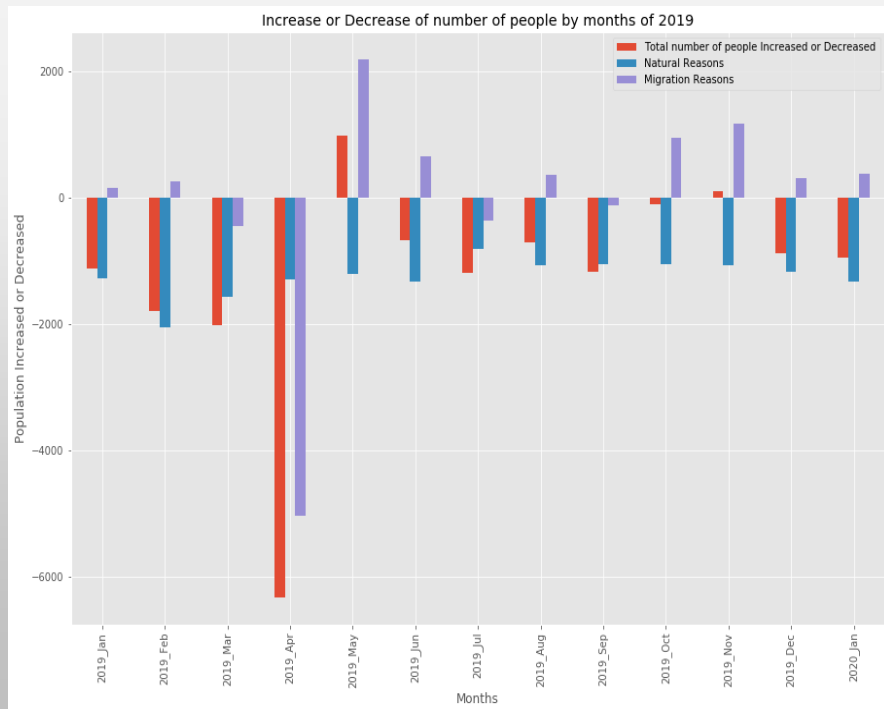


Increase or Decrease of Population in Ibaraki-prefecture from 1955 to 2019

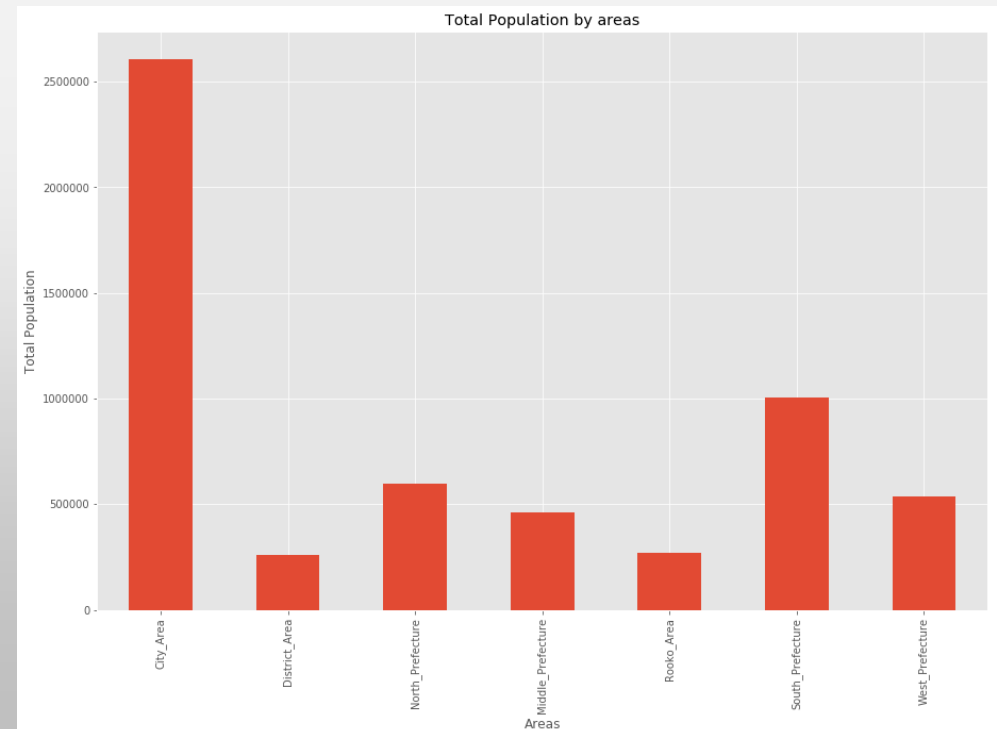


Step 2: Data visualization

- + In every month of the years, population decreased by Natural Reasons.
- + Almost months in the years, people coming to Ibaraki help the population increasing. Only in March & April when starting of fiscal year in Japan, a huge number of people get out.

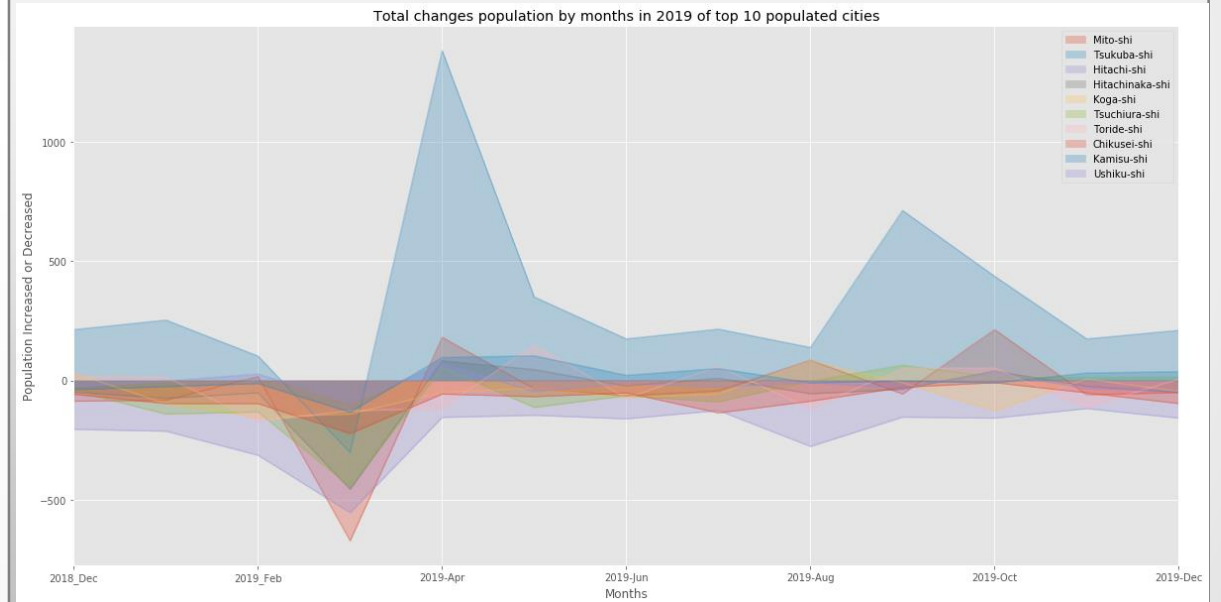
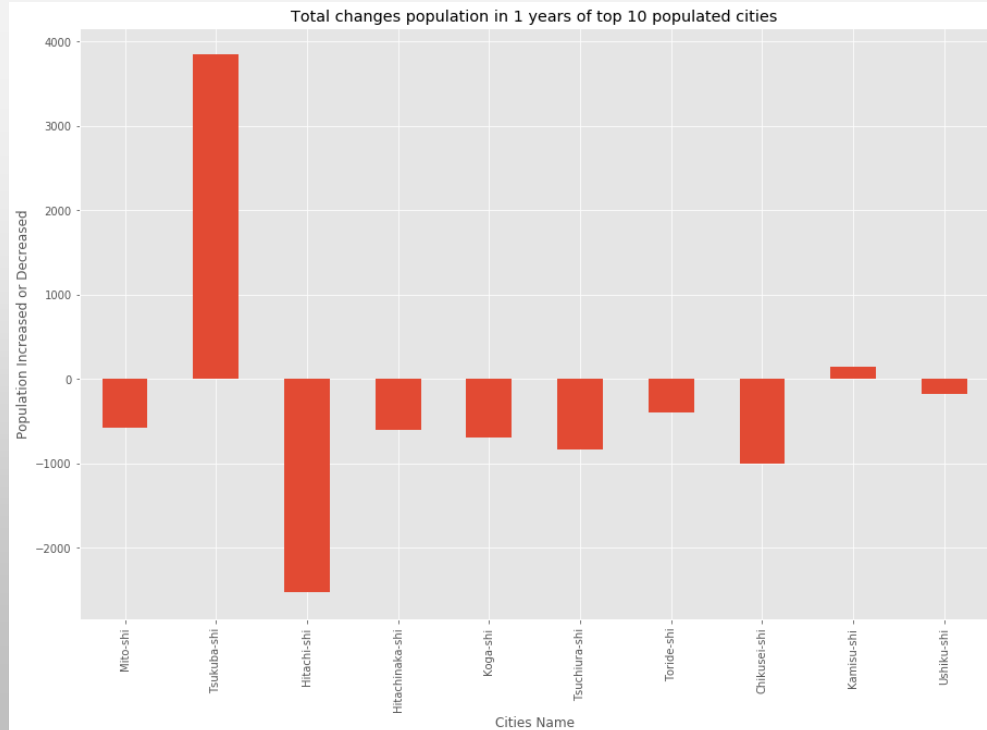


- + Cities areas are more crowded than District areas.
 - + South area is the most crowded area in the prefecture.
- => The coffee shop should be open in a city in South area.



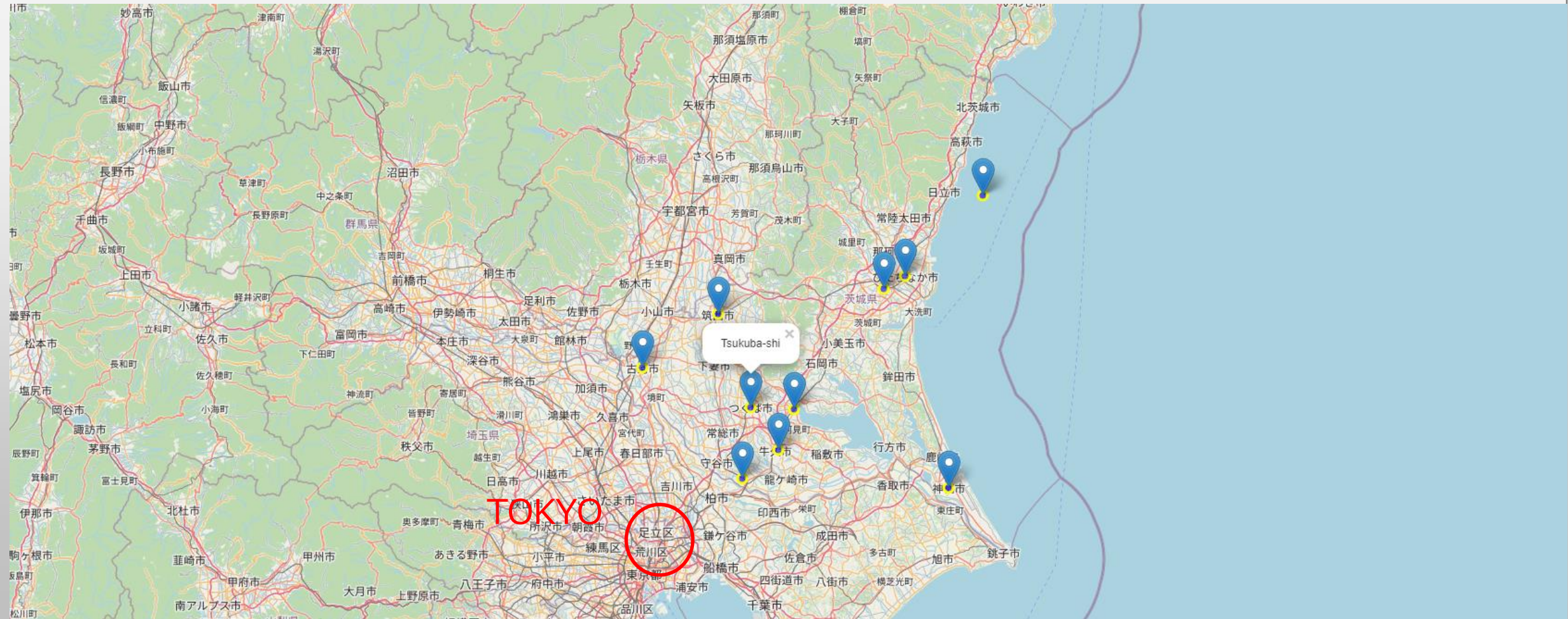
Step 2: Data visualization

Even in the impact of aging population, Tsukuba-shi keeps a very significant Increase during almost months of the year. From that, we can understand that many people come to this city for business last years.



Step 2: Data visualization

Tsukuba-shi location compare to Tokyo



Step 2: Result

Tsukuba-shi meets 3/5 requirements



The coffee shop should be located in city which in top populated cities in Ibaraki.



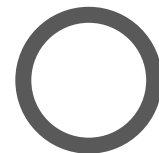
The location should be near Tokyo, so South of Ibaraki is preferable.



The city's population should be stable or increasing (many people coming to this city to cover the natural decrease cause by low birth rate & aging).

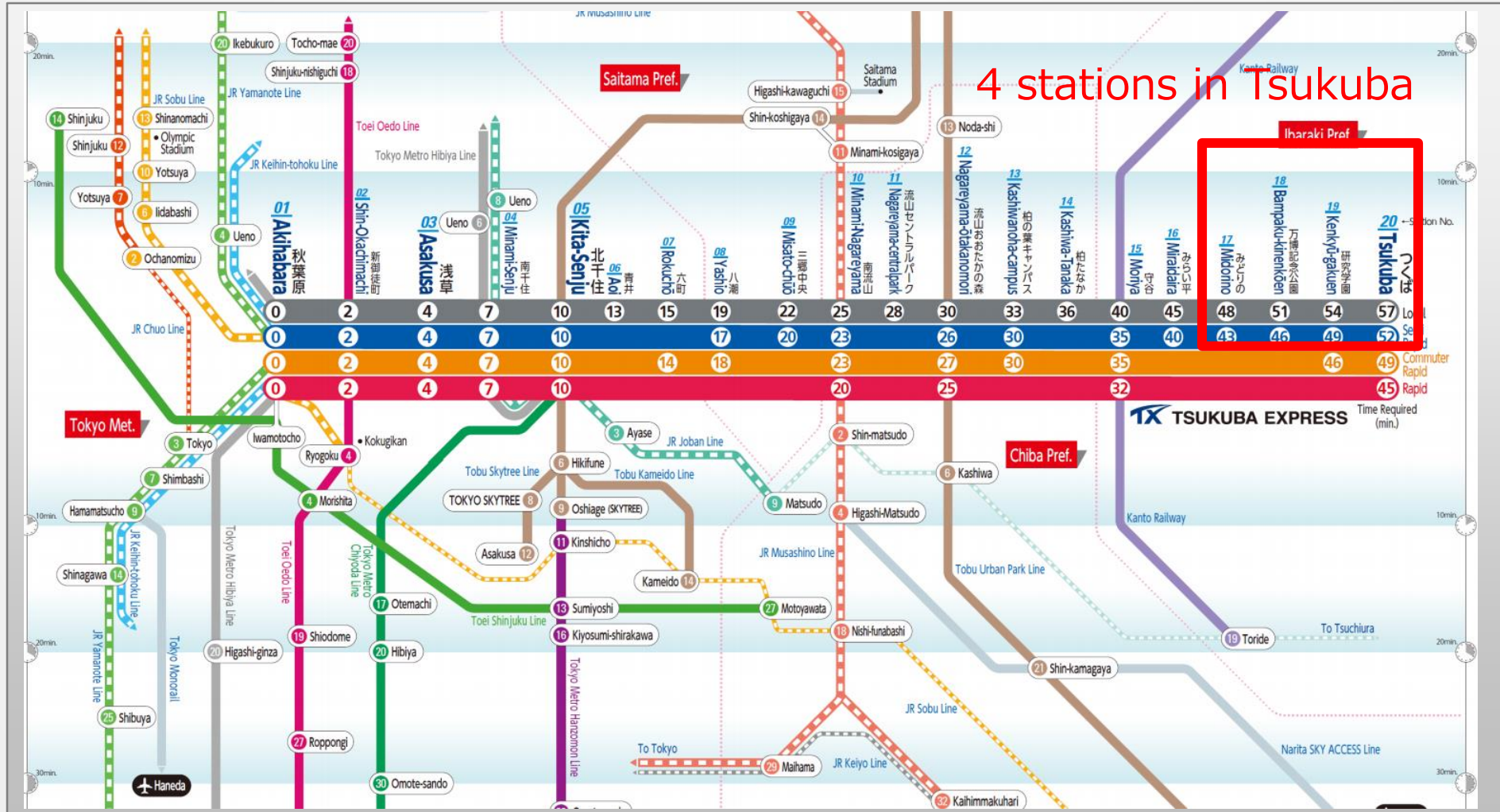


The location should be near some train stations.



There are not too many competitors in the area.

Clarify 4 train stations in Tsukuba-shi



Step 3: Explore the train stations neighborhood

- OpenCageGeocode & API from <https://opencagedata.com> to get the geocode (latitude & longitude) data of the stations
- Foursquare API was used to get the top 100 venues that are near each stations within a radius of 2km

	name	categories	lat	lng
0	SAZA COFFEE つくば駅前店	Coffee Shop	36.081613	140.112569
1	Ton Q (とんQ)	Tonkatsu Restaurant	36.077033	140.110646
2	Starbucks	Coffee Shop	36.081894	140.111416
3	Tsukuba Expo Center (つくばエキスポセンター)	Science Museum	36.086025	140.110567
4	La Cote d'Azur (コート・ダジュール本店)	Dessert Shop	36.076966	140.113734
...
95	珍来 千現店	Noodle House	36.073195	140.125970
96	Tsukuba Creo Square Q't (つくばクレオスクエア Q't)	Shopping Mall	36.081766	140.111592
97	Saizeriya (サイゼリヤ)	Italian Restaurant	36.082084	140.111194
98	Lawson (ローソン つくば駅バスターミナル店)	Convenience Store	36.081788	140.113088
99	FamilyMart (ファミリーマート)	Convenience Store	36.091717	140.106250

100 rows × 4 columns

Group the venues by categories and create a dataframe

```
[44]: group = nearby_venues.groupby('categories').count()  
venues_tsukuba = group.drop(['lat', 'lng'], axis=1)  
venues_tsukuba
```

```
[44]:
```

	name
	categories
	Arcade 2
	BBQ Joint 3
	Bakery 2
	Bar 1
	Bookstore 1
	Botanical Garden 1
	Burger Joint 1
	Café 4
	Chinese Restaurant 2
	Clothing Store 2
	Coffee Shop 4
	Concert Hall 1
	Convenience Store 10
	Dessert Shop 1

Step 3: Explore the train stations neighborhood

Compare the neighborhood in 4 train stations

```
[62]: # Though only "Cafe" % "Coffee Shop" seem to be competitors, remove others index & some these 2 rows.  
      coffe_venues = merge_venues[:2]  
      coffe_venues = coffe_venues.sum(axis=0)  
      coffe_venues
```

```
[62]: Tsukuba-station      8  
      Kenkyugakuen-station 9  
      Banpaku_Kinenkoen-station 4  
      Midorino-station    2  
      dtype: int64
```



Midorino-station has least coffee shop venues

Conclusion: Midorino-Station may the best place to open a new Coffee Shop!

Step 3: Result

Midori-no station in Tsukuba-shi meets 5/5 requirements



The coffee shop should be located in city which in top populated cities in Ibaraki.



The location should be near Tokyo, so South of Ibaraki is preferable.



The city's population should be stable or increasing (many people coming to this city to cover the natural decrease cause by low birth rate & aging).



The location should be near some train stations.



There are not too many competitors in the area.

Conclusion

Midorino-station

is the best place to open a new
Coffee Shop in
Ibaraki-prefecture.