Where to start my business

PYTHON PROJECT

Content

- BUSINESS PROBLEM DEFINITION
- DATA
- METHODOLOGY
- RESULTS

Problem definition

- Target audience: Fresh entrepreneurs, who want to start on their retail business.
- Problem definition : This project is to provide suggestion on where is suitable for fresh entrepreneurs to start their business

Data

- GDP of different countries

→Compare economic power

- Per-Capita GDP of each province in China

→Reflect citizens' purchasing power in a province

- Per-Capita GDP of Prefecture-level city in Guang Dong Province

→Similar to the previous one

- Foursquare location data

→Present the information of shopping mall in Shen Zhen

Methodology

1. Beautiful Soup

→A package to scrap data from HTML

```
# Getting GDP Data among different countries
response=requests.get('https://www.worldometers.info/gdp/gdp-by-country/').text
soup = BeautifulSoup(response, "lxml")
print(soup.prettify())
```

2. Data Frame

→Present data in table format

```
gdpdf = pd.read_html(str(gdptable))[0]
gdpdf.head(5)
```

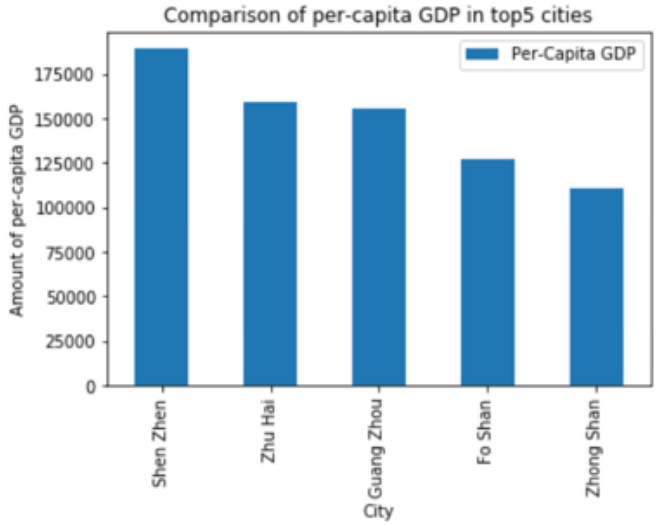
	#	Country	GDP (nominal, 2017)	GDP (abbrev.)	GDP growth	Population (2017)	GDP per capita	Share of World GDP
0	1	United States	$19,485,394,000,000	$19.485 trillion	2.27%	325084756	$59,939	24.08%
1	2	China	$12,237,700,479,375	$12.238 trillion	6.90%	1421021791	$8,612	15.12%
2	3	Japan	$4,872,415,104,315	$4.872 trillion	1.71%	127502725	$38,214	6.02%
3	4	Germany	$3,693,204,332,230	$3.693 trillion	2.22%	82658409	$44,680	4.56%
4	5	India	$2,650,725,335,364	$2.651 trillion	6.68%	1338676785	$1,980	3.28%

Methodology

3. Bar Chart

→Visualize data

```
top5city=top5city.astype(float)
top5city.plot(kind='bar')
plt.xlabel('City')
plt.ylabel('Amount of per-capita GDP')
plt.title('Comparison of per-capita GDP in top5 cities')
plt.show()
```



4. Foursqaure

```
address = 'Shen Zhen, China'

geolocator = Nominatim(user_agent="foursquare_agent")
location = geolocator.geocode(address)
latitude = location.latitude
longitude = location.longitude
print(latitude, longitude)
```

```
22.555454 114.0543297
```

```
search_query = 'Shopping'
url = 'https://api.foursquare.com/v2/venues/search?client_id={}&client_secret={}&ll={},{}&v={}&query={}&limit={}'.format(CLIENT_ID, CLIENT_SECRET, latitude, long
url
```

```
]: 'https://api.foursquare.com/v2/venues/search?client_id=RTACJTCMCJBZ2OHRMPIRZDKKW4MYSMCO1XCV233YCCUF2JXD&client_secret=4HZST1A0TWYN323MHHDTMNEYUX1JJVCFM5XZJIJ
   P3OH2D4EK&ll=22.555454,114.0543297&v=20180604&query=Shopping&limit=30'
```

Results

- In the majority part of my analysis, I rely on economic indicator, such as GDP and per-capita GDP, to finish my analysis.
- In the Foursquare part, I choose to scrap only the ‘shopping mall’ related data in Shen Zhen. It is unfortunate that there is only one tip provided for my target. Therefore, I have to admit that the choice I made may not be the best one in reality
- *To conclude, collecting and analyzing data from a wider aspect help me a lot in narrowing criteria, and the final answer to my question, which is where to start my retail business, is Shen Zhen.*