INTRODUCTION

Thinking of starting a headhunter business in Austin, TX

1. Look for jobs available in the neighborhoods to find out the demand and the lack of services
2. Delve into each neighborhood to understand why and how available jobs existed in the area
3. Compare similarity and dissimilarity among specific neighborhoods to eliminate redundancy

DATASET

1. Search for jobs in Austin to find out what services are provided there

[8]:
```
# assign relevant part of JSON to venues
venues = results['response']['venues']
#venues = results['response']['groups'][0]['items']

# tranform venues into a dataframe
dataframe = json_normalize(venues)
dataframe.head()
```

```
/home/jupyterlab/conda/envs/python/lib/python3.6/site-packages/ipykernel_launcher.py:6: FutureWarning: pandas.io.json.json_normalize is deprecated, use panda
on_normalize instead
```

[8]:

	id	name	categories	referralId	hasPerk	location.lat	location.lng	location.labeledLatLngs	location.distance	location.cc	location.city	l
0	5728f9d4498e4d039ce99626	ServJobs	'4bf58dd8d48988d1ff941735', 'name': 'M...	v-1606084927	False	30.266925	-97.743387	[{'label': 'display', 'lat': 30.26692538455611...	468	US	Austin	

2. Explore jobs in Austin to jobs provided in different neighborhoods

[16]:
```
# assign relevant part of JSON to venues
#venues = results['response']['venues']
venues = results['response']['groups'][0]['items']
#groups = results['response']['groups']
# tranform venues into a dataframe
dataframe = json_normalize(venues)
#dataframe = json_normalize(groups)
dataframe.head()
```

```
/home/jupyterlab/conda/envs/python/lib/python3.6/site-packages/ipykernel_launcher.py:6: FutureWarning: pandas.io.json.json_normalize is deprecated, use pandas.js
on_normalize instead
```

[16]:

	referralId	reasons.count	reasons.items	venue.id	venue.name	venue.location.address	venue.location.crossStreet	venue.location.lat	venue.location.lng	v
0	e-0-49bd2c3af964a52052541fe3-0	0	[{'summary': 'This spot is popular', 'type': '...	49bd2c3af964a52052541fe3	Texas State Capitol	112 E 11th St	at Congress Ave	30.273659	-97.740814	
1	e-0-558796b2498e9a8f0dccb0a5-1	0	[{'summary': 'This spot is popular', 'type': '...	558796b2498e9a8f0dccb0a5	Relief Enterprise Inc	NaN	NaN	30.269668	-97.743226	
2	e-0-50f002abe412530bab143e8b-2	0	[{'summary': 'This spot is popular', 'type': '...	50f002abe412530bab143e8b	Restore Your Door Austin	402 Rio Grande St	NaN	30.271336	-97.747142	

3 rows × 21 columns

3. Look into the neighborhood of Texas State Capitol

[27]:
```
lat = items[0]['venue']['location']['lat']
lng = items[0]['venue']['location']['lng']
lat, lng
```

[27]:
```
(30.27365925589791, -97.7408135475542)
```

There are 3 businesses around this neighborhood

```
[29]: search_query = 'Jobs'
radius = 500
url = 'https://api.foursquare.com/v2/venues/explore?client_id={}&client_secret={}&ll={},{}&v={}&query={}&radius={}&limit={}'.format(CLIENT_ID, CLIENT_SECRET, lat
results = requests.get(url).json()
results
items = results['response']['groups'][0]['items']
items

dataframe = json_normalize(items)
#dataframe = json_normalize(groups)
dataframe.head()
```

```
/home/jupyterlab/conda/envs/python/lib/python3.6/site-packages/ipykernel_launcher.py:9: FutureWarning: pandas.io.json.json_normalize is deprecated, use pandas.js
on_normalize instead
  if __name__ == '__main__':
```

[29]:

	referralId	reasons.count	reasons.items	venue.id	venue.name	venue.location.address	venue.location.crossStreet	venue.location.lat	venue.location.lng	v
0	e-0-49bd2c3af964a52052541fe3-0	0	[{'summary': 'This spot is popular', 'type': '...	49bd2c3af964a52052541fe3	Texas State Capitol	112 E 11th St	at Congress Ave	30.273659	-97.740814	
1	e-0-51ad59b3454af716216ea267-1	0	[{'summary': 'This spot is popular', 'type': '...	51ad59b3454af716216ea267	The Louver Shop Austin	815 Brazos St Ste A	# 185	30.269835	-97.738173	
2	e-0-51fc4dbd498efbf24fb82575-2	0	[{'summary': 'This spot is popular', 'type': '...	51fc4dbd498efbf24fb82575	Trinity Street Players Theatre	901 Trinity St	btwn 9th & 10th St	30.270218	-97.737595	

3 rows × 21 columns

4. Look into the neighborhood of Relief Enterprise Inc

```
[30]: lat = items[1]['venue']['location']['lat']
      lng = items[1]['venue']['location']['lng']
      lat, lng
```

```
[30]: (30.26983497612058, -97.73817300796507)
```

There are 5 businesses around this neighborhood

[32]:

	referralId	reasons.count	reasons.items	venue.id	venue.name	venue.location.address	venue.location.crossStreet	venue.location.lat	venue.location.lng	v
0	e-0-422f8e00f964a520fc1f1fe3-0	0	[{'summary': 'This spot is popular', 'type': '...	422f8e00f964a520fc1f1fe3	Elysium	705 Red River St	at 7th Street	30.267676	-97.736655	
1	e-0-414f6f00f964a520fc1c1fe3-1	0	[{'summary': 'This spot is popular', 'type': '...	414f6f00f964a520fc1c1fe3	The Jackalope	404 E 6th St	btwn Trinity St. & Neches St.	30.267131	-97.739052	
2	e-0-49bd2c3af964a52052541fe3-2	0	[{'summary': 'This spot is popular', 'type': '...	49bd2c3af964a52052541fe3	Texas State Capitol	112 E 11th St	at Congress Ave	30.273659	-97.740814	
3	e-0-51ad59b3454af716216ea267-3	0	[{'summary': 'This spot is popular', 'type': '...	51ad59b3454af716216ea267	The Louver Shop Austin	815 Brazos St Ste A	# 185	30.269835	-97.738173	
4	e-0-51fc4dbd498efbf24fb82575-4	0	[{'summary': 'This spot is popular', 'type': '...	51fc4dbd498efbf24fb82575	Trinity Street Players Theatre	901 Trinity St	btwn 9th & 10th St	30.270218	-97.737595	

5 rows × 21 columns

5. Look into

```
[33]: lat = items[2]['venue']['location']['lat']
      lng = items[2]['venue']['location']['lng']
      lat, lng
```

```
[33]: (30.27365925589791, -97.7408135475542)
```

There are the same businesses as of the first coordinate because their coordinates are very close to each other.

[34]:
```
search_query = 'Jobs'
radius = 500
url = 'https://api.foursquare.com/v2/venues/explore?client_id={}&client_secret={}&ll={},{}&v={}&query={}&radius={}&limit={}'.format(CLIENT_ID, CLIENT_SECRET, lat
results = requests.get(url).json()
results
items = results['response']['groups'][0]['items']
items

dataframe = json_normalize(items)
#dataframe = json_normalize(groups)
dataframe.head()
```

```
/home/jupyterlab/conda/envs/python/lib/python3.6/site-packages/ipykernel_launcher.py:9: FutureWarning: pandas.io.json.json_normalize is deprecated, use pandas.js
on_normalize instead
  if __name__ == '__main__':
```

[34]:

	referralId	reasons.count	reasons.items	venue.id	venue.name	venue.location.address	venue.location.crossStreet	venue.location.lat	venue.location.lng	v
0	e-0-49bd2c3af964a52052541fe3-0	0	[{'summary': 'This spot is popular', 'type': '...	49bd2c3af964a52052541fe3	Texas State Capitol	112 E 11th St	at Congress Ave	30.273659	-97.740814	
1	e-0-51ad59b3454af716216ea267-1	0	[{'summary': 'This spot is popular', 'type': '...	51ad59b3454af716216ea267	The Louver Shop Austin	815 Brazos St Ste A	# 185	30.269835	-97.738173	
2	e-0-51fc4dbd498efbf24fb82575-2	0	[{'summary': 'This spot is popular', 'type': '...	51fc4dbd498efbf24fb82575	Trinity Street Players Theatre	901 Trinity St	btwn 9th & 10th St	30.270218	-97.737595	

3 rows × 21 columns

6. There are many different opportunities for technology and big data.

METHODOLOGY

1. Choose Austin, TX is a popular location in Texas.


```
        'lng': -97.74338722229004}],
      'distance': 468,
      'cc': 'US',
      'city': 'Austin',
      'state': 'TX',
      'country': 'United States',
      'formattedAddress': ['Austin, TX', 'United States']},
     'categories': [{'id': '4bf58dd8d48988d1ff941735',
       'name': 'Miscellaneous Shop',
       'pluralName': 'Miscellaneous Shops',
       'shortName': 'Shop',
       'icon': {'prefix': 'https://ss3.4sqi.net/img/categories_v2/shops/default_',
        'suffix': '.png'},
       'primary': True}],
     'referralId': 'v-1606762193',
     'hasPerk': False}]}}
```


GET VENUES FROM JSON AND TRANSFORM IT INTO PANDAS DATAFRAME

[8]:
```
# assign relevant part of JSON to venues
venues = results['response']['venues']

# tranform venues into a dataframe
dataframe = json_normalize(venues)
dataframe.head()
```

```
/home/jupyterlab/conda/envs/python/lib/python3.6/site-packages/ipykernel_launcher.py:5: FutureWarning: pandas.io.json.json_normalize is depr
ecated, use pandas.json_normalize instead
  """
```

[8]:

	id	name	categories	referralId	hasPerk	location.lat	location.lng	location.labeledLatLngs	location.distance	location.
0	5728f9d4498e4d039ce99626	ServJobs	[{'id': '4bf58dd8d48988d1ff941735', 'name': 'M...	v-1606762193	False	30.266925	-97.743387	[{'label': 'display', 'lat': 30.26692538455611...	468	

2. Filter the data for interesting places

DEFINE INFO OF INTEREST AND FILTER DATAFRAME

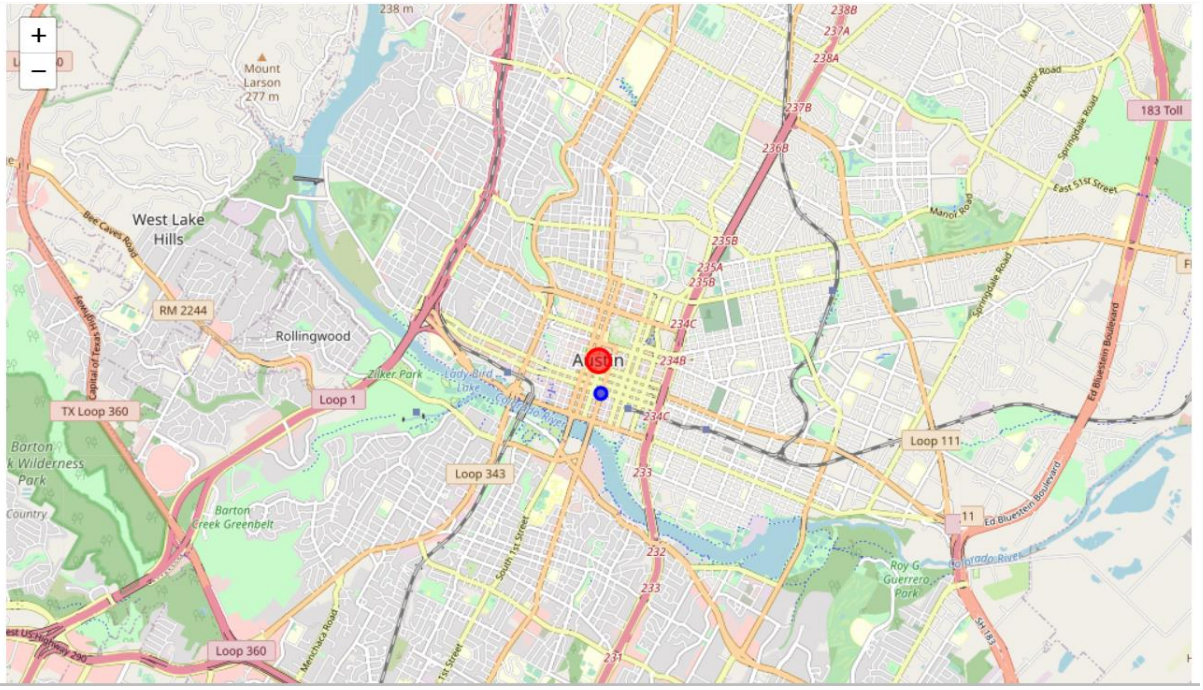
[35]:
```
# keep only columns that include venue name, and anything that is associated with location
filtered_columns = ['name', 'categories'] + [col for col in dataframe.columns if col.startswith('location.')] + ['id']
dataframe_filtered = dataframe.loc[:, filtered_columns]

# function that extracts the category of the venue
def get_category_type(row):
    try:
        categories_list = row['categories']
    except:
        categories_list = row['venue.categories']

    if len(categories_list) == 0:
        return None
    else:
        return categories_list[0]['name']

# filter the category for each row
dataframe_filtered['categories'] = dataframe_filtered.apply(get_category_type, axis=1)

# clean column names by keeping only last term
dataframe_filtered.columns = [column.split('.')[-1] for column in dataframe_filtered.columns]

dataframe_filtered
```

[35]:

	name	categories	lat	lng	labeledLatLngs	distance	cc	city	state	country	formattedAddress	id
0	ServJobs	Miscellaneous Shop	30.266925	-97.743387	[{'label': 'display', 'lat': 30.26692538455611...	468	US	Austin	TX	United States	[Austin, TX, United States]	5728f9d4498e4d039ce99626

3. Visualize the location on the map

[37]:



4. Now, let's explore the neighborhood. There are 30 restaurants around ServJobs, Austin TX

VISUALIZE INTERSTING PLACES NEARBY

[31]:
```
dataframe_filtered.name
#dataframe_filtered.categories
```

[31]:
```
0                     Perry's Steakhouse
1                      Paramount Theatre
2                           Caffé Medici
3                           The Townsend
4                  Upstairs at Caroline
5                          The Iron Bear
6                     Royal Blue Grocery
7                      CU29 Cocktail Bar
8                       The Driskill Bar
9                    Texas State Capitol
10                          The Driskill
11          Firehouse Hostel and Lounge
12                     Slake Cafe & Bar
13                Texas Capitol Grounds
14                   The Roosevelt Room
15               Gloria's Latin Cuisine
16                                 CAVA
17                     Wild About Music
18              Capitol Visitors Center
19                              Eureka!
20                       Here Nor There
21         Lonesome Dove Western Bistro
22                     Maiko Sushi Lounge
23    Moonlight Tower (9th & Guadalupe)
24                               Comedor
25                  Turf N' Surf Poboys
26                      Voodoo Doughnut
27                      Shiner's Saloon
28           Omni Austin Hotel Downtown
29            SoulCycle Downtown Austin
Name: name, dtype: object
```

5. Then visualize places in the neighborhood on map

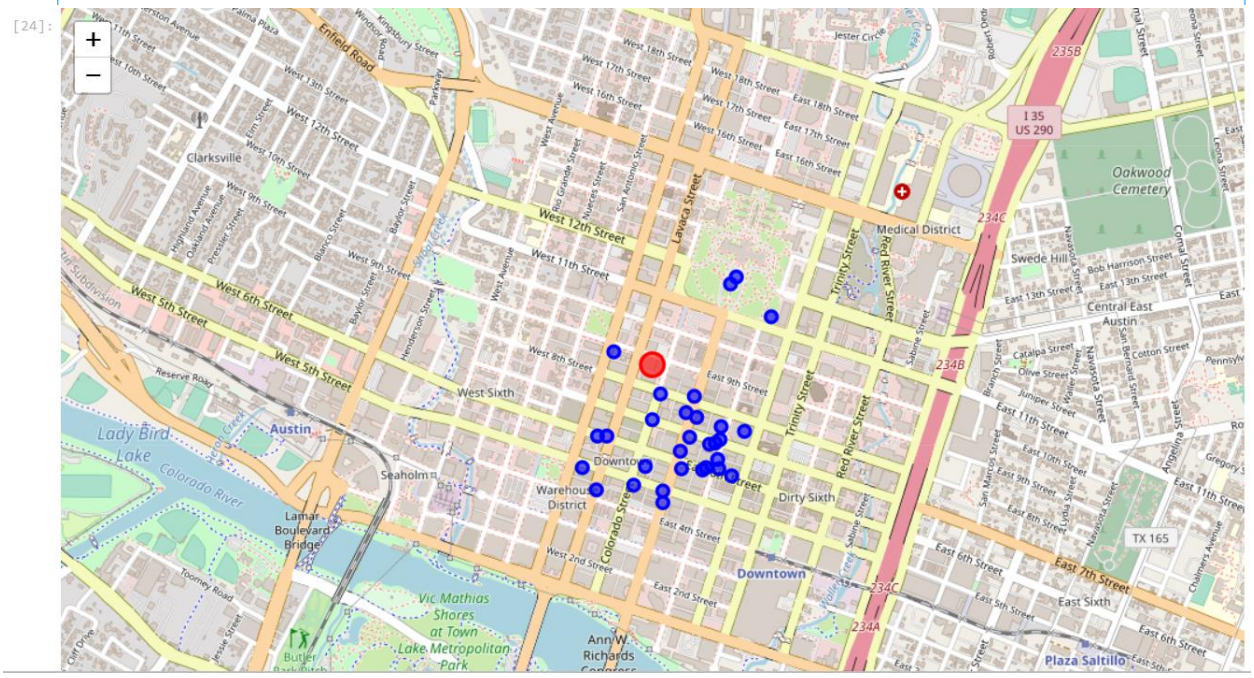
VISUALIZE INTERSTING PLACES NEARBY

[36]:
```
dataframe_filtered.name
#dataframe_filtered.categories
```

[36]:
```
0    ServJobs
Name: name, dtype: object
```

[30]:
```
venues_map = folium.Map(location=[latitude, longitude], zoom_start=13) # generate map centred around the Conrad Hotel

# add a red circle marker to represent the Conrad Hotel
folium.CircleMarker(
    [latitude, longitude],
    radius=10,
    color='red',
    popup='AUSTIN',
    fill = True,
    fill_color = 'red',
    fill_opacity = 0.6
).add_to(venues_map)

# add the Italian restaurants as blue circle markers
for lat, lng, label in zip(dataframe_filtered.lat, dataframe_filtered.lng, dataframe_filtered.categories):
    folium.CircleMarker(
        [lat, lng],
        radius=5,
        color='blue',
        popup=label,
        fill = True,
        fill_color='blue',
        fill_opacity=0.6
    ).add_to(venues_map)

# display map
```

[24]:



RESULTS

At this moment there is not any trends in this neighborhood. That means there are many other opportunities such as tech jobs, beverage stores.

EXPLORE TRENDING VENUES

[25]:
```
# define URL
url = 'https://api.foursquare.com/v2/venues/trending?client_id={}&client_secret={}&ll={},{}&v={}'.format(CLIENT_ID, CLIENT_SECRET, latitude,

# send GET request and get trending venues
results = requests.get(url).json()
results
```

[25]:
```
{'meta': {'code': 200, 'requestId': '5fc547dfadbbaf055559c583'},
 'response': {'venues': []}}
```

CHECK IF THERE ARE ANY TRENDING VENUES AT THIS TIME

[26]:
```
if len(results['response']['venues']) == 0:
    trending_venues_df = 'No trending venues are available at the moment!'

else:
    trending_venues = results['response']['venues']
    trending_venues_df = json_normalize(trending_venues)

    # filter columns
    columns_filtered = ['name', 'categories'] + ['location.distance', 'location.city', 'location.postalCode', 'location.state', 'location.cou
    trending_venues_df = trending_venues_df.loc[:, columns_filtered]

    # filter the category for each row
    trending_venues_df['categories'] = trending_venues_df.apply(get_category_type, axis=1)
```

[27]:
```
# display trending venues
trending_venues_df
```

[27]:
```
'No trending venues are available at the moment!'
```

CONCLUSION

Big Data and Data Scientist opportunities are definitely in high demand in Austin, TX