Best Location for a Vietnamese Restaurant in Austin

Cathy Le

Background

- Austin is known for
  - Tech city
  - Live music capital
  - Nature and landscape
  - Food scene
  - Good employment
- Why businesses have potential in Austin
  - Support for startups
  - Trendy and youthful city
- University of Texas
  - College students
  - Best student life
  - Businesses should cater to this area

Problem and Interest

- Group of stakeholders wanting to open up a Vietnamese restaurant in Austin
- Targeted towards the University of Texas
  - Downtown region
- Factors to consider
  - Other successful Vietnamese restaurants
  - Low competition, high demand
  - Density of restaurants in area
  - High foot traffic
- Vietnamese interest
  - Small but passionate asian and vietnamese community
  - Diversity of culture in Austin
  - Spreading new and authentic food

Data

- Source
  - Foursquare API
  - Geographical location data
- Usage
  - Find Vietnamese restaurants in Austin and visualize them on map
  - Make cluster of restaurants to identify areas
  - Use Foursquare API to get info on venues
  - Use geographic location to cluster restaurants
  - Use explore request to observe areas
  - Pass parameters for Vietnamese restaurants

Exploratory Data Methods

- Get UT address and connect to Foursquare
- Query search for Vietnamese Restaurants near campus

```
#University of Texas address
address = 'Austin, TX 78712'

geolocator = Nominatim()
location = geolocator.geocode(address)
latitude = location.latitude
longitude = location.longitude
print('The geograpical coordinate of The University of Texas are {}, {}.'.format(latituc
```

```
/opt/conda/envs/Python36/lib/python3.6/site-packages/ipykernel/__main__.py:4: Deprec
ationWarning: Using Nominatim with the default "geopy/
ly discouraged, as it violates Nominatim's ToS https:/
policies/nominatim/ and may possibly cause 403 and 42
custom `user_agent` with `Nominatim(user_agent="my-ap
he default `user_agent`: `geopy.geocoders.options.defa
on"`. In geopy 2.0 this will become an exception.

The geograpical coordinate of The University of Texas
```

```
search_query = 'Vietnamese Restaurant'
radius = 5000
print(search_query + ' .... OK!')

url2 = 'https://api.foursquare.com/v2/venues/explore?client_id={}&client_secret={}&ll={
url2
```

```
Vietnamese Restaurant .... OK!

]: 'https://api.foursquare.com/v2/venues/explore?client_id=5BURNXFD4JD3OUXRRQXNRC3XZCK5
HBFEH03N4KSIP0F1CP1H&client_secret=51A4KROSWJOX04TKEEXZSKQAIVNVQTEXC5T4CDLCDEQB1UPR&
ll=30.2711286,-97.7436995&v=201902012&query=Vietnamese Restaurant&radius=5000&limit=
100000'
```

Exploratory Data Methods

- Turn Json file into pandas dataframe

```
results2 = requests.get(url2).json()
```

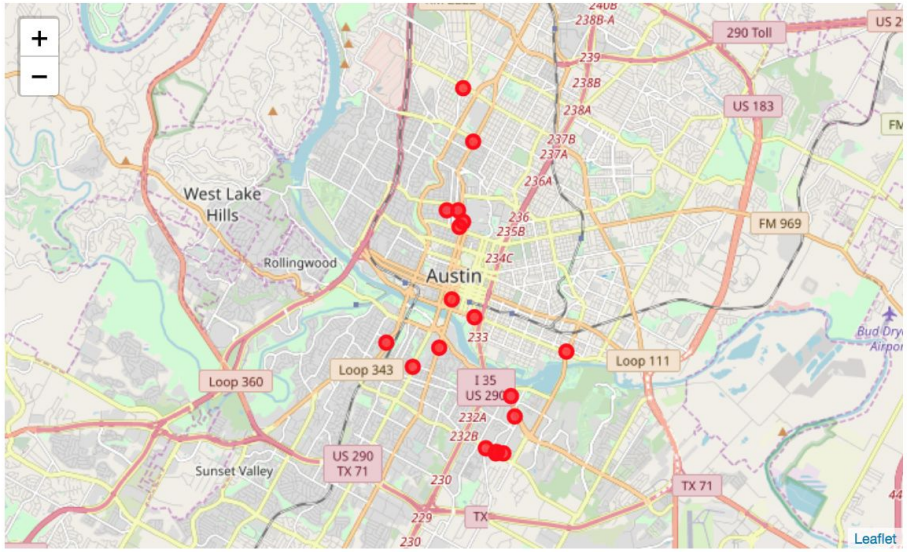
```
venues = results['response']['groups'][0]['items']

dataframe = json_normalize(venues)
```

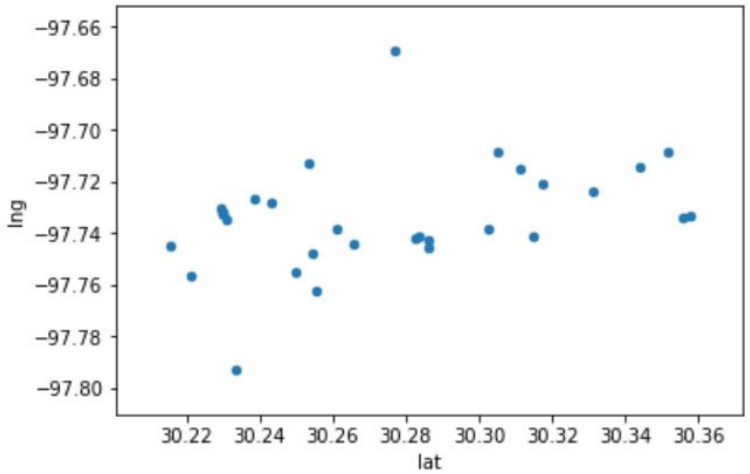
```
df=pd.DataFrame()
df['venue_ID']=dataframe['venue.id']
df['name']=dataframe['venue.name']
df['lat']=dataframe['venue.location.lat']
df['lng']=dataframe['venue.location.lng']
df.head()
```

	venue_ID	name	lat	lng
0	4ef4c212b8f77e0f983c04e4	Elizabeth St. Café	30.249828	-97.754829
1	56a2e11e498e4ab3feb17449	Pho Please	30.243062	-97.728082
2	4a9ebac6f964a520f63a20e3	888 Vietnamese Restaurant	30.229627	-97.730117
3	4a357b73f964a520099d1fe3	Hai Ky	30.230977	-97.735033
4	5816307538fa252712aba520	Heo Eatery	30.331042	-97.723462

Exploratory Data Methods

- Map of Vietnamese Restaurants
- Scatter plot of coordinates



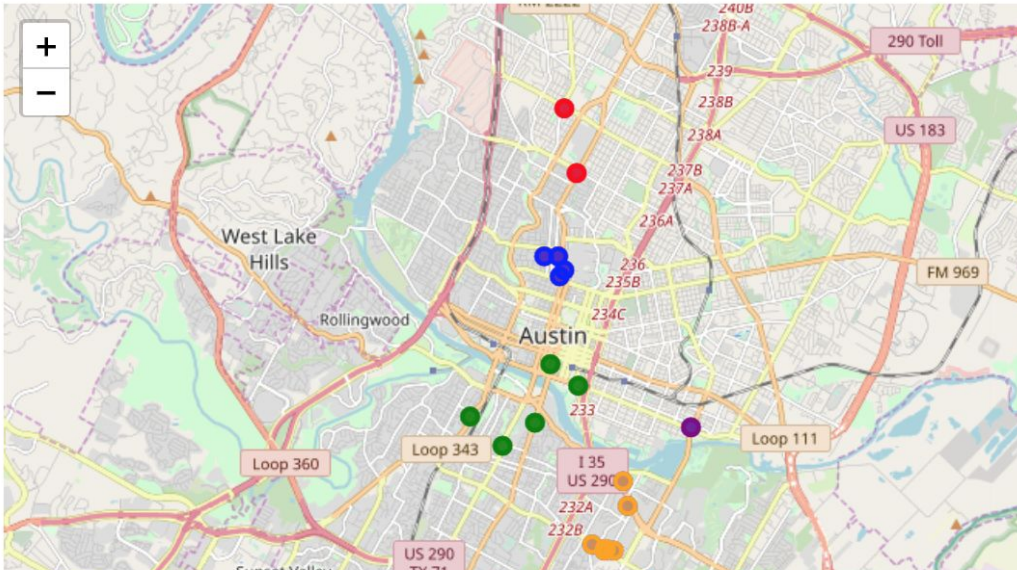


Machine Learning

- Use K-Means to make 5 clusters

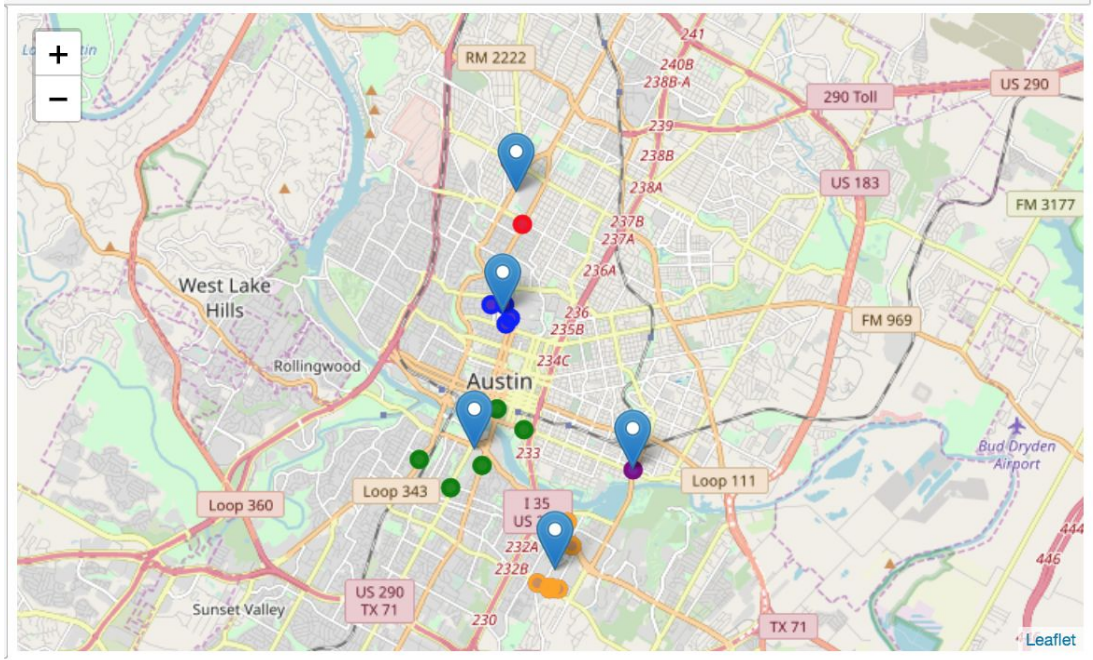
```
k_means = KMeans(init = "k-means++", n_clusters = 5, n_init = 12)
df_array=np.array(df[['lat','lng']])
k_means.fit(df_array)
k_labels=pd.DataFrame(k_means.labels_)
k_labels.info()
df['cluster']=k_labels
df.groupby('cluster').count()
```

cluster	venue_ID	name	lat	lng
0	9	9	9	9
1	10	10	10	10
2	5	5	5	5
3	6	6	6	6
4	1	1	1	1



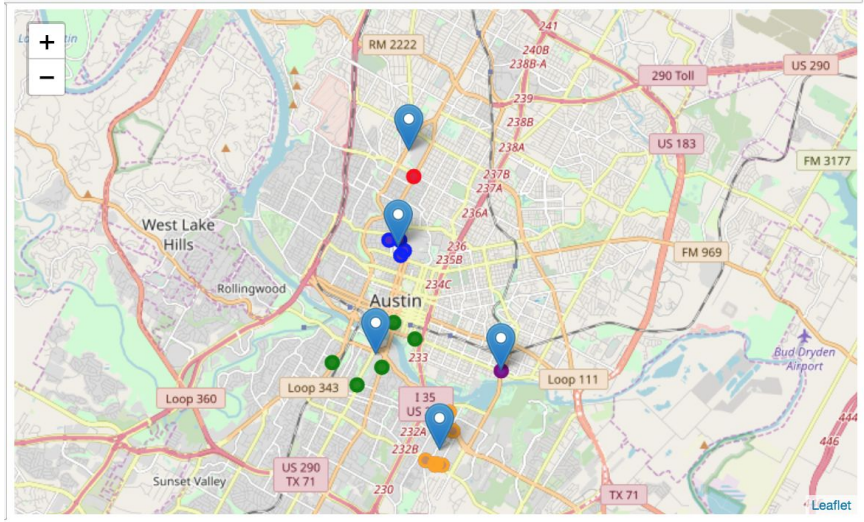
Machine Learning

cluster	lat	lng
0	30.308689	-97.739885
1	30.233123	-97.730973
2	30.257286	-97.749436
3	30.284638	-97.742914
4	30.253432	-97.713051



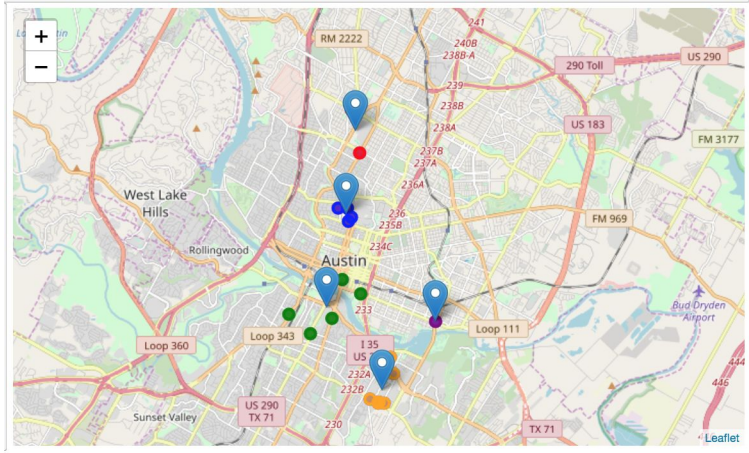
Results

- 18 Vietnamese restaurants 5000 meters from UT
- 5 clusters
- Some areas more dense than others



Observations and Recommendations

- Purple cluster is closest to UT and most dense
- Yellow cluster is also dense
- Green cluster is downtown, more spread out
- Purple and Red cluster are riskier and only have 1 or 2 restaurants - outlier



Conclusion

- Purpose was to identify location close to UT to open Vietnamese restaurant
- 5 locations were found that are good starting points for stakeholders
- Only used geographical data, can be expanded on
- Things to consider for further exploration
  - Number of clusters
  - Proximity to roads
  - Closeness to attractions
  - Noise level
  - Land availability
  - Presence of other venues
  - Prices
  - Rating of the Vietnamese restaurants used