



# ZIP CODE SHOPPING FOR HIGH-END CAFÉ

TAEYOON KIM

---

# PURPOSE

- 
- Opening a high-end dessert café in Houston
  - Catering towards mid to high class of wealthy residents
  - Need to find suitable zip code in Houston that best suits several factors for success



Synergistic with local establishments (restaurants, bars, etc.)



Healthy amounts of human traffic



Consumers with enough financial freedom and buying power to want and to afford upscale goods and services

FACTORS FOR SUCCESS

# DATA ACQUISITION

- Scrape website with `pd.read_html` function to pandas DataFrame
- Edit and clean data to legible dataframe table

#	Zip Code	Location	City	Population	Avg. Income/H/hold	National Rank
1.	<a href="#">77010</a>	29.754310, -95.361109	<a href="#">Houston, Texas</a>	76	\$200,000.00	#1
2.	<a href="#">77094</a>	29.769285, -95.681292	<a href="#">Houston, Texas</a>	7,779	\$123,244.00	#78
3.	<a href="#">77046</a>	29.733084, -95.430659	<a href="#">Houston, Texas</a>	471	\$105,863.00	#181
4.	<a href="#">77059</a>	29.615219, -95.134960	<a href="#">Houston, Texas</a>	16,690	\$104,844.00	#197
5.	<a href="#">77005</a>	29.718435, -95.423555	<a href="#">Houston, Texas</a>	23,338	\$104,035.00	#208
6.	<a href="#">77024</a>	29.771991, -95.515453	<a href="#">Houston, Texas</a>	32,746	\$82,620.00	#706
7.	<a href="#">77068</a>	30.008830, -95.487234	<a href="#">Houston, Texas</a>	9,505	\$77,724.00	#948
8.	<a href="#">77095</a>	29.916055, -95.663077	<a href="#">Houston, Texas</a>	39,275	\$76,814.00	#992
9.	<a href="#">77062</a>	29.575781, -95.134334	<a href="#">Houston, Texas</a>	26,978	\$75,689.00	#1,066
10.	<a href="#">77056</a>	29.749035, -95.469021	<a href="#">Houston, Texas</a>	14,031	\$71,926.00	#1,336
11.	<a href="#">77079</a>	29.774523, -95.601071	<a href="#">Houston, Texas</a>	30,887	\$70,587.00	#1,485
12.	<a href="#">77069</a>	29.986925, -95.524903	<a href="#">Houston, Texas</a>	14,756	\$68,662.00	#1,673
13.	<a href="#">77019</a>	29.752702, -95.407379	<a href="#">Houston, Texas</a>	15,640	\$61,899.00	#2,584
14.	<a href="#">77064</a>	29.924216, -95.537411	<a href="#">Houston, Texas</a>	35,966	\$60,945.00	#2,783
15.	<a href="#">77070</a>	29.979670, -95.572835	<a href="#">Houston, Texas</a>	32,385	\$60,887.00	#2,797
16.	<a href="#">77027</a>	29.744002, -95.443213	<a href="#">Houston, Texas</a>	14,217	\$60,032.00	#3,021
17.	<a href="#">77065</a>	29.926556, -95.603242	<a href="#">Houston, Texas</a>	27,070	\$58,426.00	#3,288
18.	<a href="#">77041</a>	29.857989, -95.585771	<a href="#">Houston, Texas</a>	25,275	\$56,746.00	#3,689
19.	<a href="#">77066</a>	29.958498, -95.500223	<a href="#">Houston, Texas</a>	27,638	\$55,760.00	#3,932
20.	<a href="#">77030</a>	29.704849, -95.404429	<a href="#">Houston, Texas</a>	10,984	\$53,199.00	#4,624

	Zip Code	Population	Avg. Income/H/hold	Latitude	Longitude
1	77010	76	\$200,000.00	29.754310	-95.361109
2	77094	7779	\$123,244.00	29.769285	-95.681292
3	77046	471	\$105,863.00	29.733084	-95.430659
4	77059	16690	\$104,844.00	29.615219	-95.134960
5	77005	23338	\$104,035.00	29.718435	-95.423555

## DATA FILTRATION

- Set data filtration criteria to average household income to remove unnecessary information
- Makes process more efficient and presents better focus on desired dataset

# FOURSQUARE API



Utilized Foursquare API to explore the zip codes and segment them



Unique credentials for access



Used the 5 zip codes (narrowed down earlier) to locate each neighborhood and to find the top 100 venues that are within a 1000 meter radius of each one



Total of 121 unique categories were found

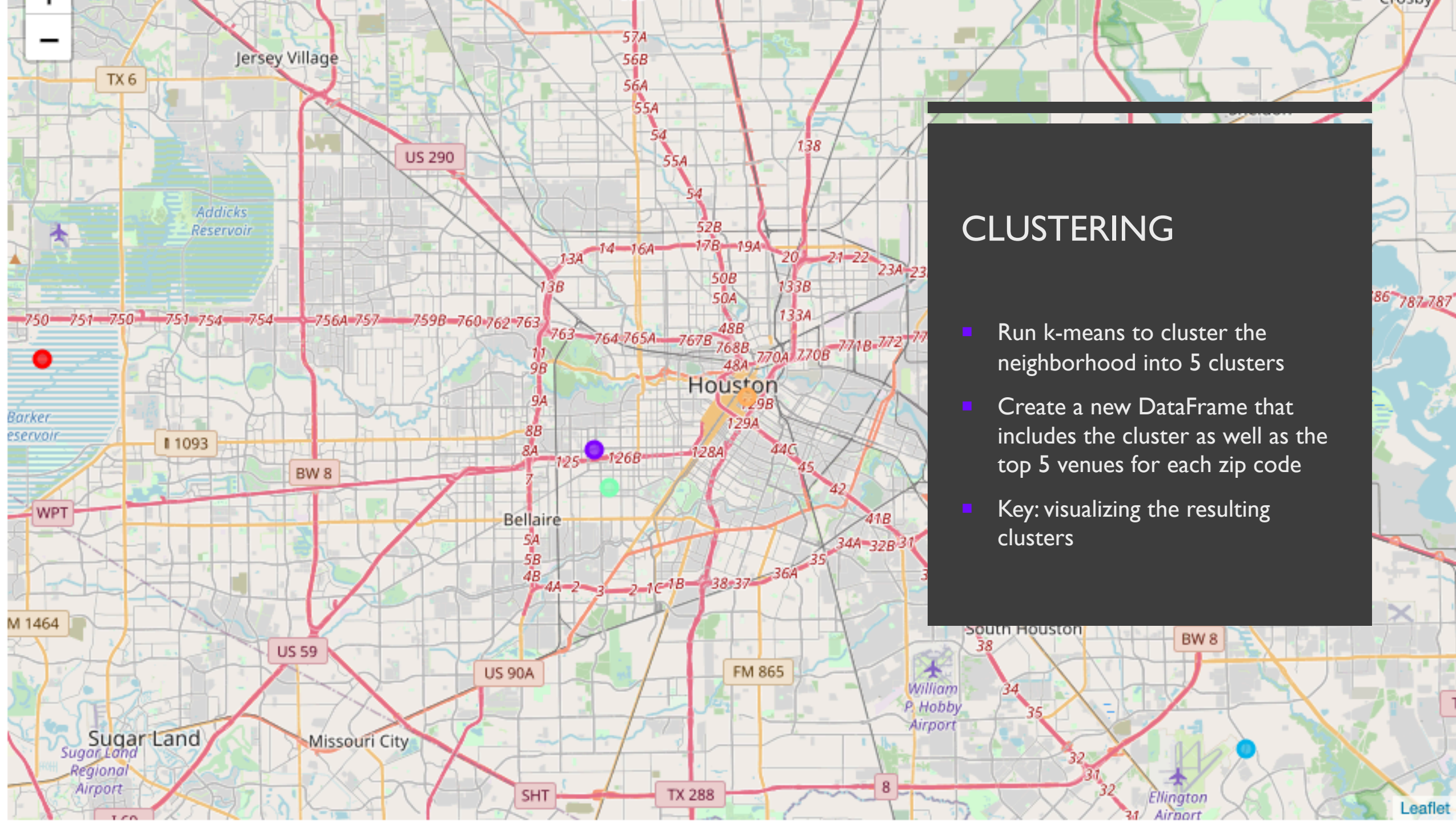
## ANALYZING EACH ZIP CODE

 Grouped rows by zip code and took the mean of the frequency of occurrence of each category

 Organized top 5 most common venues per zip code into a dataframe

	Zip Code	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
0	77005	American Restaurant	Italian Restaurant	Coffee Shop	Burger Joint	Ice Cream Shop
1	77010	Hotel	Bar	Pizza Place	Vietnamese Restaurant	Park
2	77046	Mexican Restaurant	Seafood Restaurant	Burger Joint	Italian Restaurant	New American Restaurant
3	77059	Pizza Place	Spa	Home Service	Wings Joint	Italian Restaurant
4	77094	Construction & Landscaping	Auto Dealership	Other Repair Shop	Gas Station	Fast Food Restaurant





# CLUSTERING

- Run k-means to cluster the neighborhood into 5 clusters
- Create a new DataFrame that includes the cluster as well as the top 5 venues for each zip code
- Key: visualizing the resulting clusters

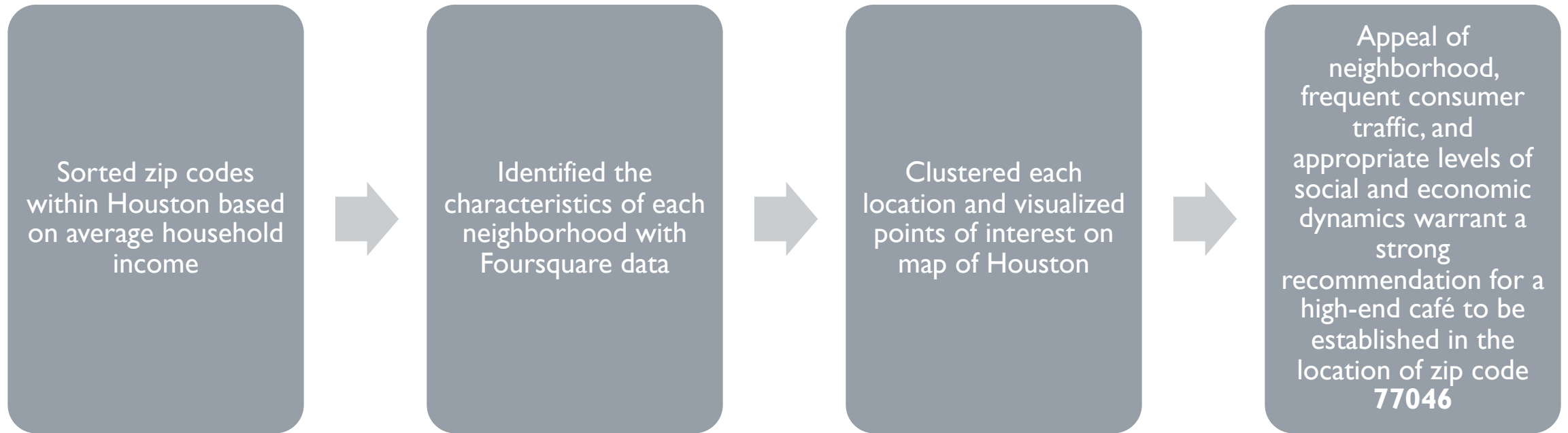


# BRIDGING THE GAP

Merging filtered data based on population and average household income with Foursquare data on surrounding venues

Uniform data set to make a calculated, business decision for strategic placement of upscale café

	Zip Code	Population	Avg. Income/H/hold	Latitude	Longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
1	77010	76	200000	29.754310	-95.361109	4	Hotel	Bar	Pizza Place	Vietnamese Restaurant	Park
2	77094	7779	123244	29.769285	-95.681292	0	Construction & Landscaping	Auto Dealership	Other Repair Shop	Gas Station	Fast Food Restaurant
3	77046	471	105863	29.733084	-95.430659	1	Mexican Restaurant	Seafood Restaurant	Burger Joint	Italian Restaurant	New American Restaurant
4	77059	16690	104844	29.615219	-95.134960	2	Pizza Place	Spa	Home Service	Wings Joint	Italian Restaurant
5	77005	23338	104035	29.718435	-95.423555	3	American Restaurant	Italian Restaurant	Coffee Shop	Burger Joint	Ice Cream Shop



## DRAWING CONCLUSIONS



## FUTURE RECOMMENDATIONS

- Choropleth maps from Folium could be used to better represent the data at hand
- Used in conjunction with geojson datafiles containing zip code boundary coordinates
- Color coding based on numerous features could be possible
  - Population
  - Average household income
  - Real estate prices
- Far more attractive and comprehensive data visualization method to present to stakeholders