Using k-Means Clustering to Determine Ideal Locations to Open an Art Gallery in Research Triangle, North Carolina

IBM Data Science Professional Certificate: Capstone Project

Sam Neubauer

December 15th, 2020

Data (North Carolina Zip Codes)

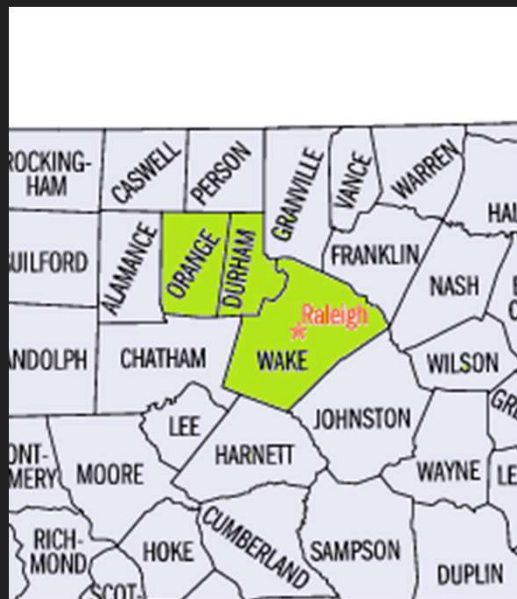○ **CSV file with all zip codes in North Carolina source:**
https://data.opendatasoft.com/explore/dataset/geonames-postal-code@public-us/download/?format=csv&q=&refine.country_code=US&refine.admin_name1=North+Carolina&timezone=America/New_York&use_labels_for_header=true&csv_separator=%3B

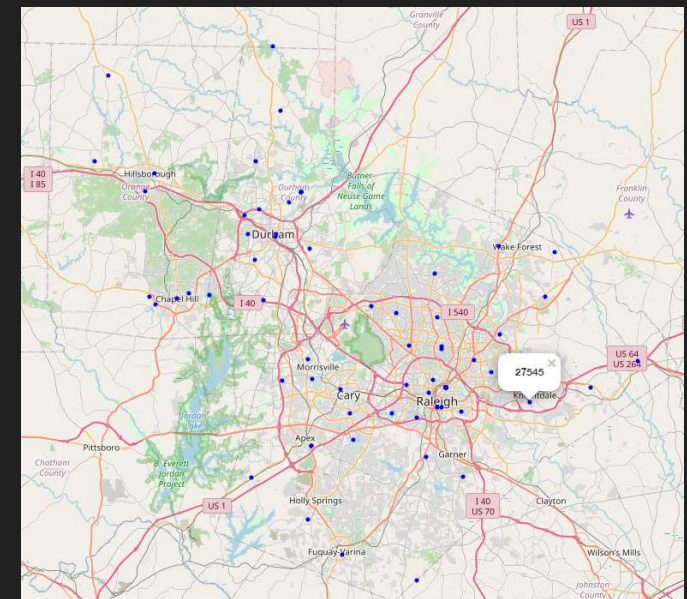|	country code	postal code	place name	admin name1	admin code1	admin name2	admin code2	admin name3	admin code3	latitude	longitude	accuracy	coordinates
0	US	27349	Snow Camp	North Carolina	NC	Alamance	1	NaN	NaN	35.9066	-79.4279	4.0	35.9066,-79.4279
1	US	28678	Stony Point	North Carolina	NC	Alexander	3	NaN	NaN	35.8661	-81.0641	4.0	35.8661,-81.0641
2	US	28675	Sparta	North Carolina	NC	Alleghany	5	NaN	NaN	36.5089	-81.1384	4.0	36.5089,-81.1384
3	US	28626	Fleetwood	North Carolina	NC	Ashe	9	NaN	NaN	36.2814	-81.5140	4.0	36.2814,-81.514
4	US	28640	Jefferson	North Carolina	NC	Ashe	9	NaN	NaN	36.4090	-81.4396	4.0	36.409,-81.4396
...	...	...	...	...	...	...	...	...	...	...	...	...	...
1076	US	27530	Goldsboro	North Carolina	NC	Wayne	191	NaN	NaN	35.3683	-78.0929	4.0	35.3683,-78.0929
1077	US	27863	Pikeville	North Carolina	NC	Wayne	191	NaN	NaN	35.4929	-77.9578	4.0	35.4929,-77.9578
1078	US	28578	Seven Springs	North Carolina	NC	Wayne	191	NaN	NaN	35.2105	-77.9146	4.0	35.2105,-77.9146
1079	US	28651	Millers Creek	North Carolina	NC	Wilkes	193	NaN	NaN	36.2119	-81.2485	4.0	36.2119,-81.2485
1080	US	27894	Wilson	North Carolina	NC	Wilson	195	NaN	NaN	35.7158	-77.9043	4.0	35.7158,-77.9043

1081 rows × 13 columns

Data (North Carolina Zip Codes)

○ Data was then cleaned, removing NaN/unneeded columns, and filtered to only include the counties within the Research Triangle area: Wake, Durham, Orange

	zip	place	county	latitude	longitude
0	27704	Durham	Durham	36.0383	-78.8764
1	27710	Durham	Durham	36.0512	-78.8577
2	27711	Durham	Durham	36.0512	-78.8577
3	27529	Garner	Wake	35.6813	-78.5975
4	27615	Raleigh	Wake	35.8887	-78.6393
...	...	...	...	...	...
86	27523	Apex	Wake	35.7225	-78.8408
87	27560	Morrisville	Wake	35.8344	-78.8466
88	27571	Rolesville	Wake	35.9156	-78.4658
89	27602	Raleigh	Wake	35.7587	-78.6711
90	27695	Raleigh	Wake	35.7977	-78.6253

91 rows × 5 columns

Data (Foursquare API Venues)

Foursquare was utilized to obtain search results for the following queries within a 3-mile radius at each zip code:

- brewery
- cocktails
- distillery
- gallery
- museum
- music venues

	Zip	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category	Venue Code
0	27704	36.0383	-78.8764	Fullsteam Brewery	36.003166	-78.899773	Brewery	0
1	27704	36.0383	-78.8764	Durham Distillery	36.004452	-78.902913	Brewery	0
2	27704	36.0383	-78.8764	Durty Bull Brewing Co.	36.001353	-78.898207	Brewery	0
3	27704	36.0383	-78.8764	Geer Street Garden	36.003601	-78.901057	Gastropub	1
4	27704	36.0383	-78.8764	Motorco Music Hall	36.003599	-78.900325	Concert Hall	2
...	...	...	...	...	...	...	...	...
5868	27712	36.0918	-78.9299	Choir Ready Room	36.063362	-78.926900	Music Venue	30
5869	27712	36.0918	-78.9299	Barry & Claire Ragin's Home	36.071176	-78.961081	Music Venue	30
5870	27588	35.9731	-78.4508	Hater's Ball	35.941644	-78.475700	Music Venue	30
5871	27607	35.8014	-78.6877	Museum Park Theater	35.809629	-78.703345	Music Venue	30
5998	27560	35.8344	-78.8466	NC Violinist	35.810209	-78.816562	Music Venue	30

567 rows × 8 columns