

Data Section:

The data from this project will be sourced from the following places:

(a) data.austintexas.gov (<https://data.austintexas.gov/>) :

A data portal that provides easy access to open data and information about the city government.

This code snippet collects the neighborhood boundaries from the City of Austin website.

```
# Importing data from the Austin Government website
from sodapy import Socrata
# First 2000 results, returned as JSON from API / converted
client = Socrata("data.austintexas.gov",None)
results = client.get("nz5f-3t2e",limit=2000)
# Convert to pandas DataFrame
df_neighbors = pd.DataFrame.from_records(results)
df_neighbors.head()
```

	the_geom	_feature_i	planning_a	shape_area	shape_leng	label
0	['type': 'MultiPolygon', 'coordinates': [[[-9...	1.0	ALLANDALE	65792689.5531	42253.1072105	Allandale
1	['type': 'MultiPolygon', 'coordinates': [[[-9...	2.0	BARTON HILLS	88901714.7947	48353.9339254	Barton Hills
2	['type': 'MultiPolygon', 'coordinates': [[[-9...	3.0	BOULDIN CREEK	33258999.9494	25667.3403756	Bouldin Creek
3	['type': 'MultiPolygon', 'coordinates': [[[-9...	4.0	BRENTWOOD	44207756.068	29612.4036976	Brentwood
4	['type': 'MultiPolygon', 'coordinates': [[[-9...	5.0	CENTRAL EAST AUSTIN	26970983.8606	22198.5298969	Central East Austin

The same portal also provides census tracts along with the demographic data each tract.

```
# Importing data from the Austin Government website
from sodapy import Socrata
# First 2000 results, returned as JSON from API
client = Socrata("austin-energy.data.socrata.com",None)
results = client.get("vi83-tbx3",limit=2000)
# Convert to pandas DataFrame
df_results = pd.DataFrame.from_records(results)
df_results.head()
```

	the_geom	tract10	sum_totpop	sum_white	sum_black	sum_hispan	sum_asian	sum_other	sum_multi	sum_hu_tot	...	sum_age62	sum_age651	sum_age_65	sum_age75	sum_age80	sum_age85	sum_famswc	sum_pop_in	sum_famili	sum_nont
0	['type': 'MultiPolygon', 'coordinates': [[[-9...	000101	3611.0	3130.0	39.0	276.0	90.0	9.0	67.0	2076.0	...	81.0	38.0	151.0	79.0	103.0	303.0	354.0	77.0	805.0	101
1	['type': 'MultiPolygon', 'coordinates': [[[-9...	000102	2552.0	2311.0	23.0	150.0	42.0	3.0	23.0	1153.0	...	126.0	70.0	253.0	77.0	66.0	91.0	299.0	18.0	735.0	34
2	['type': 'MultiPolygon', 'coordinates': [[[-9...	000203	1546.0	1073.0	87.0	213.0	143.0	5.0	25.0	817.0	...	17.0	8.0	25.0	6.0	6.0	2.0	32.0	281.0	122.0	61
3	['type': 'MultiPolygon', 'coordinates': [[[-9...	000204	3009.0	2396.0	43.0	323.0	161.0	8.0	78.0	1750.0	...	77.0	21.0	80.0	27.0	20.0	26.0	234.0	56.0	495.0	110
4	['type': 'MultiPolygon', 'coordinates': [[[-9...	000205	3384.0	2333.0	110.0	676.0	162.0	24.0	89.0	2153.0	...	50.0	35.0	95.0	30.0	28.0	21.0	212.0	0.0	557.0	145

(b) United States Census Bureau (<https://www.census.gov/data.html/>) :

The United States Census Bureau is a principal agency of the U.S. Federal Statistical System, responsible for producing data about the American people and economy. A snippet of the data obtained from the Census Bureau providing the population of Indian living in each Census Tract of Travis County.

GEO_ID	POPGROUP_TTL	NAME	HCT001001
id	Popgroup Label	Geographic Area Name	Total
1400000US48453000302	Asian Indian alone (400-401)	Census Tract 3.02, Travis County, Texas	73
1400000US48453000401	Asian Indian alone (400-401)	Census Tract 4.01, Travis County, Texas	81
1400000US48453000500	Asian Indian alone (400-401)	Census Tract 5, Travis County, Texas	73
1400000US48453000601	Asian Indian alone (400-401)	Census Tract 6.01, Travis County, Texas	15
1400000US48453000603	Asian Indian alone (400-401)	Census Tract 6.03, Travis County, Texas	149
1400000US48453000604	Asian Indian alone (400-401)	Census Tract 6.04, Travis County, Texas	181
1400000US48453001100	Asian Indian alone (400-401)	Census Tract 11, Travis County, Texas	71
1400000US48453001714	Asian Indian alone (400-401)	Census Tract 17.14, Travis County, Texas	260

From Wikipedia: **Travis County** is a [county](#) in south [central Texas](#). As of the [2010 census](#), the population was 1,024,266; the estimated population in 2019 was 1,273,954.^[1] It is the fifth-most populous county in [Texas](#). Its [county seat](#) is [Austin](#),^[2] the capital of Texas. Travis County is part of the Austin-[Round Rock Metropolitan Statistical Area](#).

(c) FOURSQUARE (<https://foursquare.com/>) :

An American technology company that provides location platform for several businesses like UBER, TARGET, Twitter, SUBWAY, and many others. In this project we shall make calls to the Foursquare API to search for venues as well as specific type of venues or restaurants around a given location. With the data available from FOURSQUARE, we shall explore neighborhoods, segment and group them into clusters to find similar neighborhoods in Austin. We shall use a density-based clustering algorithm, called K-means for this unsupervised dataset.