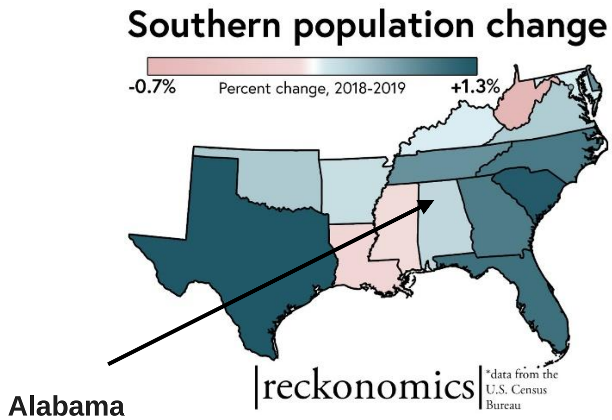
Where to open an American Restaurant in Alabama?

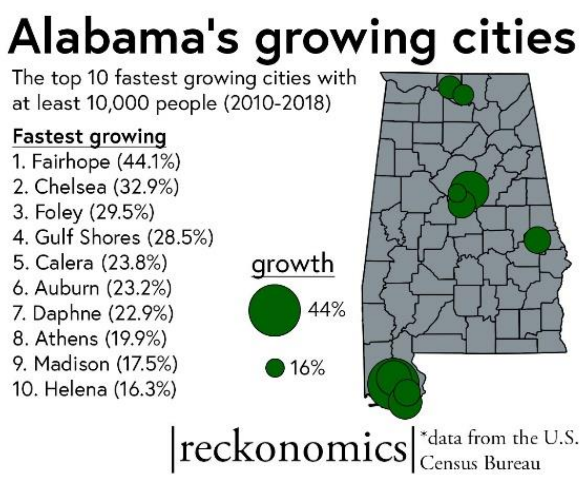
Introduction

A national American Restaurant chain is seeking to add restaurant locations in every state in the Southern United States. The chain is experiencing increasing sales growth in all currently existing restaurant locations and is interested in expanding. This study focuses on the best location to open a restaurant in the state of Alabama.

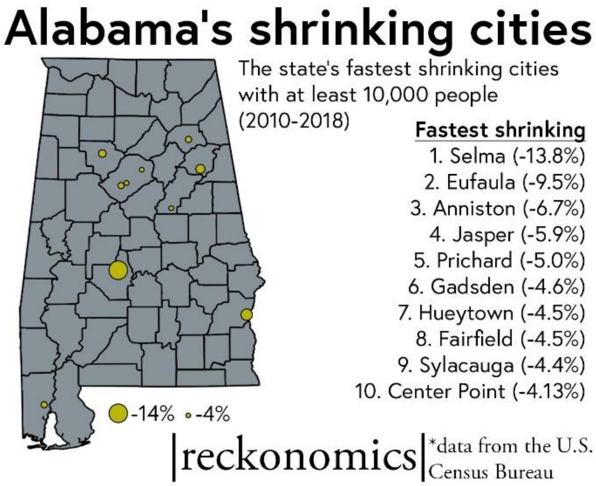
Alabama is the 24th most populous state in the United States of America. It is having relatively stagnant population growth generally, but there is population movement within its respective cities. The graphic below illustrates Alabama’s population growth comparatively with the rest of the American South. In 2019, the population of Alabama increased 0.3%, which was the 26th fastest rate of growth in the United States of America.



As illustrated in the graphic below, the fastest growing cities within Alabama are not concentrated in any particular area. There is population growth in the Northern, Southern, Eastern, and Central regions of Alabama. Further analysis is needed to determine the businesses that are located within these cities.



As illustrated in the graphic below, the slowest growing cities within Alabama are concentrated in the central region of Alabama, outside of any major cities. It would be best to avoid opening up any new businesses in these cities with declining populations. Further analysis is needed to determine the businesses that are located within these cities.



This study will utilize Alabama city location and Foursquare data to evaluate existing business locations to determine the best locations to open a new American restaurant in Alabama.

Data

A list of Alabama cities and their respective Longitude and Latitude coordinates were retrieved from a csv file at: https://simplemaps.com/data/us-cities. This file also contained population information and other descriptive information. There were 584 cities in the dataset, but the data was sorted to include only the top 200 most populous cities in Alabama for this study. To enable retrieval, the .csv file was hosted on a personal Dropbox site at https://www.dropbox.com/s/r823yayzncnqe9q/al_cities_lat_long.csv?dl=0. Screenshots of these actions are below.

Alabama Cities and their Latitudes and Longitudes were retrieved from https://simplemaps.com/data/us-cities

In [19]:
```
!wget -q -O 'al_cities_lat_long.csv' https://www.dropbox.com/s/r823yayzncnqe9q/al_cities_lat_long.csv?dl=0
print('Data downloaded!')
```

```
Data downloaded!
```

Now that the data is downloaded, let's read it into a *pandas* dataframe and display the number of cities in our data set.

In [23]:
```
al_cities = pd.read_csv('al_cities_lat_long.csv')
print(al_cities.shape)
al_cities.head()
```

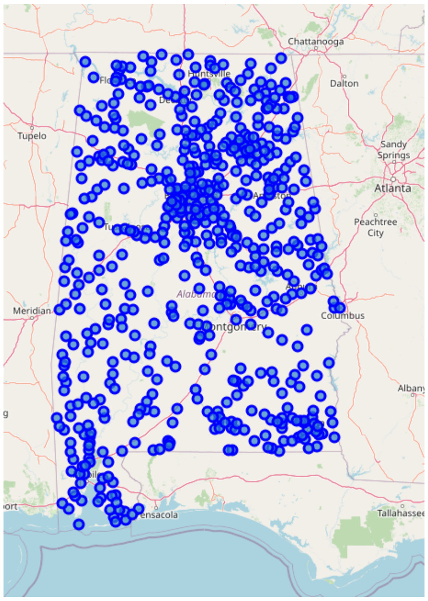
```
(200, 3)
```

Out[23]:

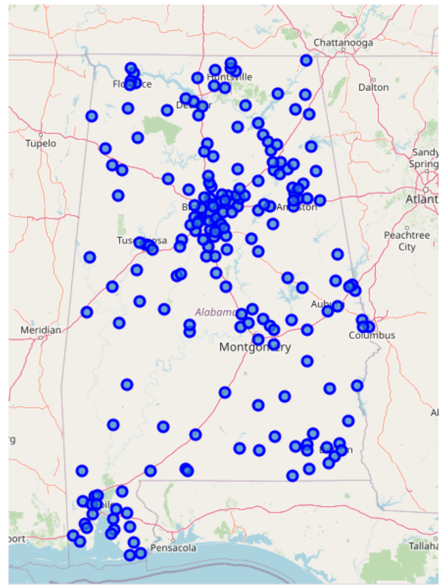
	City	Latitude	Longitude
0	Birmingham	33.5277	-86.7987
1	Huntsville	34.6988	-86.6412
2	Mobile	30.6783	-88.1162
3	Montgomery	32.3473	-86.2666
4	Tuscaloosa	33.2348	-87.5267

Once the data was brought into the Python notebook, Folium was used to display the cities on a map. The images below display a map of Alabama with all 548 cities in the original dataset and a map of Alabama with only the top 200 most populous cities in the dataset.

All 584 Cities in Alabama



200 Most Populous Cities in Alabama



A Foursquare developer account was created to obtain Foursquare credentials that enabled the gathering of Foursquare business type information.

Methodology

After the Alabama City, Latitude, and Longitude data were gathered, a function was created and run to obtain all venues for the 200 most populous cities in Alabama using the Foursquare API (screenshots below).

Let's create a function to obtain all venues for the cities in Alabama.

```
def getNearbyVenues(names, latitudes, longitudes, radius=500):

    venues_list=[]
    for name, lat, lng in zip(names, latitudes, longitudes):
        print(name)

        # create the API request URL
        url = 'https://api.foursquare.com/v2/venues/explore?&client_id={}&client_secret={}&v={}&ll={},{}&radius={}&limit={}'.format(
            CLIENT_ID,
            CLIENT_SECRET,
            VERSION,
            lat,
            lng,
            radius,
            LIMIT)

        # make the GET request
        results = requests.get(url).json()["response"]['groups'][0]['items']

        # return only relevant information for each nearby venue
        venues_list.append([(
            name,
            lat,
            lng,
            v['venue']['name'],
            v['venue']['location']['lat'],
            v['venue']['location']['lng'],
            v['venue']['categories'][0]['name']) for v in results])

    nearby_venues = pd.DataFrame([item for venue_list in venues_list for item in venue_list])
    nearby_venues.columns = ['City',
                  'City Latitude',
                  'City Longitude',
                  'Venue',
                  'Venue Latitude',
                  'Venue Longitude',
                  'Venue Category']

    return(nearby_venues)
```

Let's run the above function on each city in Alabama and create a new dataframe called al_venues.

```
al_venues = getNearbyVenues(names=al_cities['City'],
                                   latitudes=al_cities['Latitude'],
                                   longitudes=al_cities['Longitude']
                                  )
```

```
Birmingham
Huntsville
Mobile
Montgomery
Tuscaloosa
Auburn
Hoover
Florence
Anniston
Dothan
Daphne
Decatur
Gadsden
Madison
Enterprise
Albertville
Foley
Phenix City
Prattville
Vestavia Hills
```

These venue listings were placed in a new dataframe (al_venues) that contained 590 total venues for the 200 most populous cities in Alabama (screenshot below).

Let's check the size of the resulting dataframe.

In [78]:
```
print(al_venues.shape)
al_venues
```

```
(590, 7)
```

Out[78]:

	City	City Latitude	City Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Birmingham	33.5277	-86.7987	The Pit Barbeque	33.530252	-86.800837	BBQ Joint
1	Birmingham	33.5277	-86.7987	Sarris	33.526543	-86.794331	American Restaurant
2	Birmingham	33.5277	-86.7987	Subway	33.526873	-86.801567	Sandwich Place
3	Birmingham	33.5277	-86.7987	Birm/Jeff Bus Barn	33.528248	-86.794528	Bus Station
4	Huntsville	34.6988	-86.6412	Redstone Pool	34.695808	-86.643733	Pool
...	...	...	...	...	...	...	...
585	Bridgeport	34.9495	-85.7243	MeMaw's	34.949116	-85.721802	American Restaurant
586	Greensboro	32.7014	-87.5950	PieLab	32.704189	-87.595226	Café
587	Greensboro	32.7014	-87.5950	CVS pharmacy	32.700775	-87.594643	Pharmacy
588	Greensboro	32.7014	-87.5950	Dollar General	32.705495	-87.595302	Discount Store
589	Greensboro	32.7014	-87.5950	McDonald's	32.699021	-87.595802	Fast Food Restaurant

590 rows × 7 columns

Next, One Hot Encoding was used to create a binary integer column for each venue category to enable K Means clustering analysis.

```
# one hot encoding
al_onehot = pd.get_dummies(al_venues[['Venue Category']], prefix="", prefix_sep="")

# add City column back to dataframe
al_onehot['City'] = al_venues['City']

# move neighborhood column to the first column
fixed_columns = [al_onehot.columns[-1]] + list(al_onehot.columns[:-1])
al_onehot = al_onehot[fixed_columns]

al_onehot.head()
```

12]:

	City	Accessories Store	American Restaurant	Antique Shop	Arcade	Arts & Crafts Store	Asian Restaurant	Athletics & Sports	Automotive Shop	BBQ Joint	...	Thrift / Vintage Store	Toy / Game Store	Trail	Train Station	Transportat Serv
0	Birmingham	0	0	0	0	0	0	0	0	1	...	0	0	0	0	
1	Birmingham	0	1	0	0	0	0	0	0	0	...	0	0	0	0	
2	Birmingham	0	0	0	0	0	0	0	0	0	...	0	0	0	0	
3	Birmingham	0	0	0	0	0	0	0	0	0	...	0	0	0	0	
4	Huntsville	0	0	0	0	0	0	0	0	0	...	0	0	0	0	

5 rows × 150 columns

The One Hot dataframe was next grouped by Alabama city and the mean of the frequency of occurrence of each category was derived. The grouped dataframe was confirmed to contain 123 Alabama cities and 150 unique venue categories. Note: Although 200 Alabama cities were in the source dataset, only 123 of the cities contained Venue data in Foursquare

```
al_onehot.shape
```

3]: (590, 150)

Next, let's group rows by city and by taking the mean of the frequency of occurrence of each category

```
al_grouped = al_onehot.groupby('City').mean().reset_index()
al_grouped
```

4]:

	City	Accessories Store	American Restaurant	Antique Shop	Arcade	Arts & Crafts Store	Asian Restaurant	Athletics & Sports	Automotive Shop	BBQ Joint	...	Thrift / Vintage Store	Toy / Game Store	Trail	Train Station	Transp
0	Abbeville	0.0	0.142857	0.0	0.0	0.0	0.0	0.0	0.0	0.000000	...	0.0	0.0	0.0	0.000	
1	Alabaster	0.0	0.000000	0.0	0.0	0.0	0.0	0.0	0.0	0.000000	...	0.0	0.0	0.0	0.000	
2	Albertville	0.0	0.000000	0.0	0.0	0.0	0.0	0.0	0.0	0.000000	...	0.0	0.0	0.0	0.000	
3	Alexander City	0.0	0.000000	0.0	0.0	0.0	0.0	0.0	0.0	1.000000	...	0.0	0.0	0.0	0.000	
4	Alexandria	0.0	0.000000	0.0	0.0	0.0	0.0	0.0	0.0	0.000000	...	0.0	0.0	0.0	0.000	
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	
118	Tuscumbia	0.0	0.000000	0.0	0.0	0.0	0.0	0.0	0.0	0.000000	...	0.0	0.0	0.0	0.000	
119	Union Springs	0.0	0.000000	0.0	0.0	0.0	0.0	0.0	0.0	0.000000	...	0.0	0.0	0.0	0.000	
120	Vestavia Hills	0.0	0.000000	0.0	0.0	0.0	0.0	0.0	0.0	0.000000	...	0.0	0.0	0.0	0.000	
121	Wetumpka	0.0	0.250000	0.0	0.0	0.0	0.0	0.0	0.0	0.000000	...	0.0	0.0	0.0	0.125	
122	Winfield	0.0	0.000000	0.0	0.0	0.0	0.0	0.0	0.0	0.333333	...	0.0	0.0	0.0	0.000	

123 rows × 150 columns

Let's confirm the new size

```
al_grouped.shape
```

5]: (123, 150)

To take a closer look at the city data, each city was printed along with its top 10 most common venues. As you can see below, Fast Food Restaurants are the most common venue type in the city of Auburn, Alabama making up 22% of all venues.

```
num_top_venues = 10

for hood in al_grouped['City']:
    print("----"+hood+"----")
    temp = al_grouped[al_grouped['City'] == hood].T.reset_index()
    temp.columns = ['venue','freq']
    temp = temp.iloc[1:]
    temp['freq'] = temp['freq'].astype(float)
    temp = temp.round({'freq': 2})
    print(temp.sort_values('freq', ascending=False).reset_index(drop=True).head(num_top_venues))
    print('\n')
```

```
----Auburn----
                     venue  freq
0      Fast Food Restaurant  0.22
1             Deli / Bodega  0.11
2               Coffee Shop  0.11
3        Mexican Restaurant  0.11
4  Mediterranean Restaurant  0.11
5              Liquor Store  0.11
6            Sandwich Place  0.11
7            Discount Store  0.11
8           Piercing Parlor  0.00
9        Photography Studio  0.00
```

To prepare the data for K Means Clustering, a function was written to put the grouped data into a pandas dataframe that displayed the top 10 venues for each Alabama city (screenshots below).

Let's put that into a *pandas* dataframe

```
def return_most_common_venues(row, num_top_venues):
    row_categories = row.iloc[1:]
    row_categories_sorted = row_categories.sort_values(ascending=False)

    return row_categories_sorted.index.values[0:num_top_venues]
```

```
num_top_venues = 10

indicators = ['st', 'nd', 'rd']

# create columns according to number of top venues
columns = ['City']
for ind in np.arange(num_top_venues):
    try:
        columns.append('{}{} Most Common Venue'.format(ind+1, indicators[ind]))
    except:
        columns.append('{}th Most Common Venue'.format(ind+1))

# create a new dataframe
city_venues_sorted = pd.DataFrame(columns=columns)
city_venues_sorted['City'] = al_grouped['City']

for ind in np.arange(al_grouped.shape[0]):
    city_venues_sorted.iloc[ind, 1:] = return_most_common_venues(al_grouped.iloc[ind, :], num_top_venues)

city_venues_sorted
```

6]:

	City	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
0	Abbeville	Gas Station	American Restaurant	Diner	Discount Store	Business Service	Bank	Women's Store	Fish & Chips Shop	Fast Food Restaurant	Farmers Market
1	Alabaster	Pizza Place	Pharmacy	Gas Station	Chinese Restaurant	Fast Food Restaurant	Gym	Convenience Store	Dessert Shop	Diner	Department Store
2	Albertville	Construction & Landscaping	Grocery Store	Discount Store	Performing Arts Venue	Women's Store	Drugstore	Fast Food Restaurant	Farmers Market	Farm	Electronics Store

In the last analytical step, K Means Clustering was run to cluster the cities into 9 clusters and a new dataframe was created that included both the Cluster Labels and the top 10 venues for each city (screenshot below).

Run *k*-means to cluster the cities into 9 clusters.

```
# set number of clusters
kclusters = 9

al_grouped_clustering = al_grouped.drop('City', 1)

# run k-means clustering
kmeans = KMeans(n_clusters=kclusters, random_state=0).fit(al_grouped_clustering)

# check cluster labels generated for each row in the dataframe
kmeans.labels_[0:20]
```

```
7]: array([5, 0, 5, 5, 5, 5, 0, 5, 5, 0, 5, 0, 0, 7, 5, 5, 5, 5, 0, 5],
          dtype=int32)
```

Let's create a new dataframe that includes the cluster as well as the top 10 venues for each city.

```
# add clustering labels

city_venues_sorted = city_venues_sorted.drop(columns=['Cluster Labels'])

city_venues_sorted.insert(0, 'Cluster Labels', kmeans.labels_)

al_merged = al_cities


# merge manhattan_grouped with manhattan_data to add latitude/longitude for each city
al_merged = al_merged.join(city_venues_sorted.set_index('City'), on='City')

al_merged = al_merged.dropna() #drop NaN rows without Cluster Label values

al_merged = al_merged.reset_index(drop=True) #Reset index

al_merged = al_merged.astype({"Cluster Labels":'int'}) #Convert 'Cluster Label' column from Float to Int so the map works

print(al_merged.shape)
al_merged.head() # check the last columns!
```

```
    (123, 14)
```

3]:

	City	Latitude	Longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue
0	Birmingham	33.5277	-86.7987	5	American Restaurant	Bus Station	Sandwich Place	BBQ Joint	Women's Store	Dry Cleaner	Fish & Chips Shop	Fast Food Restaurant	Farmers Market
1	Huntsville	34.6988	-86.6412	5	Pool	Food	Fish & Chips Shop	Fast Food Restaurant	Farmers Market	Farm	Electronics Store	Dry Cleaner	Drugstore

Folium was then used to create a cluster map with each cluster having its own color and city label displayed on a map of Alabama. The cluster map and individual clusters will be discussed in the Results section of the report.

Results

Below is a cluster map of the cities in Alabama clustered by similar venue makeups. There are 9 unique clusters, 123 total clusters, and the highest populated cluster contains 76 clusters. An analysis of each cluster is provided below.

Cluster 1 is primarily composed of Fast Food and Pizza Restaurants.

Cluster 1

```
al_merged.loc[al_merged['Cluster Labels'] == 0, al_merged.columns[[0] + list(range(4, al_merged.shape[1]))]] #fast food/pizza popul
```

	City	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
5	Auburn	Fast Food Restaurant	Deli / Bodega	Discount Store	Liquor Store	Sandwich Place	Mediterranean Restaurant	Mexican Restaurant	Coffee Shop	Doctor's Office	Dog Run
17	Alabaster	Pizza Place	Pharmacy	Gas Station	Chinese Restaurant	Fast Food Restaurant	Gym	Convenience Store	Dessert Shop	Diner	Department Store
31	Muscle Shoals	Fast Food Restaurant	Video Store	Convenience Store	Mexican Restaurant	Medical Center	Women's Store	Dry Cleaner	Fish & Chips Shop	Farmers Market	Farm
32	Hartselle	Pizza Place	Discount Store	Fast Food Restaurant	Video Store	Food	Pharmacy	Asian Restaurant	Gas Station	Supermarket	Women's Store
34	Talladega	Fast Food Restaurant	Mexican Restaurant	Pharmacy	Liquor Store	Café	Pizza Place	Video Store	Diner	Discount Store	Doctor's Office

Cluster 2 has Parks as its most common venue. All 10 of their most common venue types share similar frequencies (e.g. Fast Food Restaurant is either the 5th or 6th most common venue for all cities in Cluster 2).

Cluster 2

```
al_merged.loc[al_merged['Cluster Labels'] == 1, al_merged.columns[[0] + list(range(4, al_merged.shape[1]))]] #parks most popular
```

	City	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
13	Foley	Park	Women's Store	Drugstore	Fish & Chips Shop	Fast Food Restaurant	Farmers Market	Farm	Electronics Store	Dry Cleaner	Donut Shop
59	Montevallo	Park	Women's Store	Drugstore	Fish & Chips Shop	Fast Food Restaurant	Farmers Market	Farm	Electronics Store	Dry Cleaner	Donut Shop
99	Livingston	Bookstore	Park	Women's Store	Drugstore	Fish & Chips Shop	Fast Food Restaurant	Farmers Market	Farm	Electronics Store	Dry Cleaner
115	Taylor	Hardware Store	Park	Women's Store	Drugstore	Fish & Chips Shop	Fast Food Restaurant	Farmers Market	Farm	Electronics Store	Dry Cleaner

In Cluster 3, every city except 1 has Construction & Landscaping as its most common venue.

Cluster 3

```
al_merged.loc[al_merged['Cluster Labels'] == 2, al_merged.columns[[0] + list(range(4, al_merged.shape[1]))]] #construction an
```

	City	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
8	Dothan	Construction & Landscaping	Women's Store	Drugstore	Fish & Chips Shop	Fast Food Restaurant	Farmers Market	Farm	Electronics Store	Dry Cleaner	Donut Shop
10	Decatur	Construction & Landscaping	Fast Food Restaurant	Women's Store	Drugstore	Fish & Chips Shop	Farmers Market	Farm	Electronics Store	Dry Cleaner	Donut Shop
23	Oxford	Construction & Landscaping	Discount Store	Women's Store	Drugstore	Fish & Chips Shop	Fast Food Restaurant	Farmers Market	Farm	Electronics Store	Dry Cleaner
44	Robertsdale	Construction & Landscaping	Women's Store	Drugstore	Fish & Chips Shop	Fast Food Restaurant	Farmers Market	Farm	Electronics Store	Dry Cleaner	Donut Shop
56	Tuscumbia	Construction & Landscaping	Insurance Office	Women's Store	Drugstore	Fish & Chips Shop	Fast Food Restaurant	Farmers Market	Farm	Electronics Store	Dry Cleaner
73	Margaret	Construction & Landscaping	Women's Store	Drugstore	Fish & Chips Shop	Fast Food Restaurant	Farmers Market	Farm	Electronics Store	Dry Cleaner	Donut Shop
82	Priceville	Mexican Restaurant	Construction & Landscaping	Drugstore	Flower Shop	Fish & Chips Shop	Fast Food Restaurant	Farmers Market	Farm	Electronics Store	Dry Cleaner

Cluster 4 contains cities that have Baseball Field as its 1st or 2nd most common venue.

Cluster 4

```
al_merged.loc[al_merged['Cluster Labels'] == 3, al_merged.columns[[0] + list(range(4, al_merged.shape[1]))]] ## baseball fiel
```

	City	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
28	Sylacauga	Baseball Field	Women's Store	Dry Cleaner	Flower Shop	Fish & Chips Shop	Fast Food Restaurant	Farmers Market	Farm	Electronics Store	Drugstore
85	Holtville	Baseball Field	Women's Store	Dry Cleaner	Flower Shop	Fish & Chips Shop	Fast Food Restaurant	Farmers Market	Farm	Electronics Store	Drugstore
96	Heflin	Gym / Fitness Center	Baseball Field	Women's Store	Drugstore	Fish & Chips Shop	Fast Food Restaurant	Farmers Market	Farm	Electronics Store	Dry Cleaner
118	Good Hope	Intersection	Baseball Field	Women's Store	Dry Cleaner	Flower Shop	Fish & Chips Shop	Fast Food Restaurant	Farmers Market	Farm	Electronics Store

In Cluster 5, Discount Stores, Women's Stores, and Drugstores are the most common venues.

Cluster 5 ¶

```
al_merged.loc[al_merged['Cluster Labels'] == 4, al_merged.columns[[0] + list(range(4, al_merged.shape[1]))]] #discount stores,
```

	City	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
51	Forestdale	Grocery Store	Discount Store	Women's Store	Drugstore	Fish & Chips Shop	Fast Food Restaurant	Farmers Market	Farm	Electronics Store	Dry Cleaner
62	Theodore	Discount Store	Women's Store	Drugstore	Fish & Chips Shop	Fast Food Restaurant	Farmers Market	Farm	Electronics Store	Dry Cleaner	Donut Shop
76	Glencoe	Discount Store	Women's Store	Drugstore	Fish & Chips Shop	Fast Food Restaurant	Farmers Market	Farm	Electronics Store	Dry Cleaner	Donut Shop
84	Holt	Fish & Chips Shop	Discount Store	Women's Store	Drugstore	Fast Food Restaurant	Farmers Market	Farm	Electronics Store	Dry Cleaner	Donut Shop
109	Moundville	Discount Store	Video Store	Women's Store	Drugstore	Fish & Chips Shop	Fast Food Restaurant	Farmers Market	Farm	Electronics Store	Dry Cleaner

Cluster 6 is the largest cluster with 76 cities. It also contains the most populous cities in the state of Alabama (e.g. Birmingham, Huntsville, Mobile). All of these cities in this cluster contain a large mix of venues.

Cluster 6

```
al_merged.loc[al_merged['Cluster Labels'] == 5, al_merged.columns[[0] + list(range(4, al_merged.shape[1]))]] #large city mix
```

	City	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
0	Birmingham	American Restaurant	Bus Station	Sandwich Place	BBQ Joint	Women's Store	Dry Cleaner	Fish & Chips Shop	Fast Food Restaurant	Farmers Market	Farm
1	Huntsville	Pool	Food	Fish & Chips Shop	Fast Food Restaurant	Farmers Market	Farm	Electronics Store	Dry Cleaner	Drugstore	Donut Shop
2	Mobile	Breakfast Spot	Intersection	Fast Food Restaurant	Flower Shop	Brazilian Restaurant	Chinese Restaurant	Jewelry Store	Bank	Japanese Restaurant	Automotive Shop
3	Montgomery	Steakhouse	Gym	Women's Store	Drugstore	Fish & Chips Shop	Fast Food Restaurant	Farmers Market	Farm	Electronics Store	Dry Cleaner
4	Tuscaloosa	Food & Drink Shop	Food	Fish & Chips Shop	Fast Food Restaurant	Farmers Market	Farm	Electronics Store	Dry Cleaner	Drugstore	Donut Shop

Cluster 7 contains 2 cities, both with Golf Course as the most common venue. These cities are known as vacation destinations.

Cluster 7

```
al_merged.loc[al_merged['Cluster Labels'] == 6, al_merged.columns[[0] + list(range(4, al_merged.shape[1]))]]#golf course/vacat
```

	City	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
36	Gulf Shores	Golf Course	Women's Store	Donut Shop	Fish & Chips Shop	Fast Food Restaurant	Farmers Market	Farm	Electronics Store	Dry Cleaner	Drugstore
65	Grayson Valley	Golf Course	Candy Store	Women's Store	Drugstore	Fish & Chips Shop	Fast Food Restaurant	Farmers Market	Farm	Electronics Store	Dry Cleaner

The 3 cities in Cluster 8 all share the same venues for their 1st through 10th most common venues.

Cluster 8

```
al_merged.loc[al_merged['Cluster Labels'] == 7, al_merged.columns[[0] + list(range(4, al_merged.shape[1]))]] # american restaurants
```

	City	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
114	Point Clear	American Restaurant	Women's Store	Drugstore	Flower Shop	Fish & Chips Shop	Fast Food Restaurant	Farmers Market	Farm	Electronics Store	Dry Cleaner
119	Midland City	American Restaurant	Women's Store	Drugstore	Flower Shop	Fish & Chips Shop	Fast Food Restaurant	Farmers Market	Farm	Electronics Store	Dry Cleaner
121	Bridgeport	American Restaurant	Women's Store	Drugstore	Flower Shop	Fish & Chips Shop	Fast Food Restaurant	Farmers Market	Farm	Electronics Store	Dry Cleaner

Cluster 9 cities all have Home Service as their most common venue. They also share the same venue categories as their 4th through 9th most common categories.

Cluster 9

```
al_merged.loc[al_merged['Cluster Labels'] == 8, al_merged.columns[[0] + list(range(4, al_merged.shape[1]))]] #home service
```

	City	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
11	Madison	Home Service	Cave	Donut Shop	Fish & Chips Shop	Fast Food Restaurant	Farmers Market	Farm	Electronics Store	Dry Cleaner	Drugstore
47	Rainbow City	Home Service	Drugstore	Flower Shop	Fish & Chips Shop	Fast Food Restaurant	Farmers Market	Farm	Electronics Store	Dry Cleaner	Women's Store
49	Meadowbrook	Home Service	Drugstore	Flower Shop	Fish & Chips Shop	Fast Food Restaurant	Farmers Market	Farm	Electronics Store	Dry Cleaner	Women's Store
86	Mount Olive	Home Service	Drugstore	Flower Shop	Fish & Chips Shop	Fast Food Restaurant	Farmers Market	Farm	Electronics Store	Dry Cleaner	Women's Store
97	Piedmont	Home Service	Café	Drugstore	Fish & Chips Shop	Fast Food Restaurant	Farmers Market	Farm	Electronics Store	Dry Cleaner	Donut Shop

Discussion

Based on the clustering results, I would recommend that the national American Restaurant chain open their restaurant in Auburn, Alabama. Auburn is contained in Cluster 1, which has Fast Food and Pizza Restaurants as their most common venues. It is a popular city for restaurants generally with Deli/Bodega, Sandwich Place, Fast Food, Mediterranean, and Mexican Restaurants ranking among its top 10 most common venues. Also, there currently exists no other competition from other American Restaurants in its top 10 most common venues.

Also, Auburn also has a large population of 62,996 and, as identified in the Introduction section, is the 6^{th} fastest growing city in Alabama. This provides a large and growing customer base that will be interested in dining at a new location. Because it is not contained in Cluster 6, the largest cluster containing the most populous cities, the new restaurant would have little competition from other restaurants that may have an insurmountable presence in these existing large markets.

Conclusion

Alabama is ripe for new growth and the market research present in this report has identified Auburn, Alabama as the best city in Alabama to open a new American Restaurant. Through the use of the Data Science techniques present in this report, further research can be completed to further enable expansion throughout the United States.