How to choose a location for an American restaurant
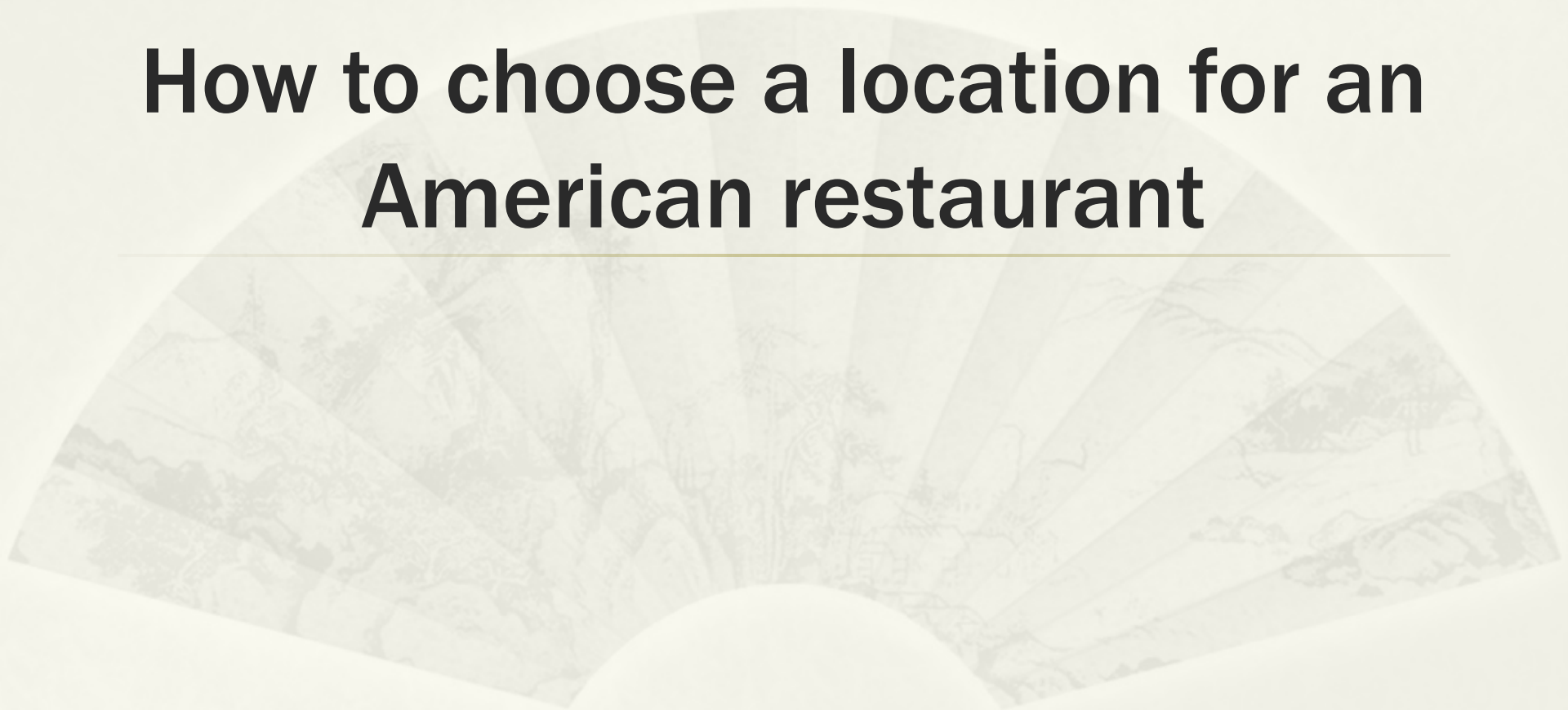
Backgroud

* American flavor of restaurant is very popular in Toronto in Canada.

* Investors are looking to open an American flavor of restaurant.

* They don't know where he should open it in Toronto in Canada.

Problems

* How to choose a location for an American restaurant?

* Are there the least risks of competition in this location?

* Are there the most potential diners or customers in this location?

Solution

* According to experienced businessmen who had or have run restaurants for many years, the following three factors should be considered in the selection of location of this flavor of restaurant.

* 1. The less American flavor of restaurants.

* 2. The less other flavor of restaurants.

* 3. The more entertainment spots.

Investors must find a neighborhood which meets the above three factors at the same time.

Data

* ## 1. The Toronto neighborhood data

  The data is obtained through scraping the Wikipedia page which has all the information we need to explore the neighborhoods in Toronto.

* ## 2. Foursquare location data

  The data is obtained through Foursquare API.

Methodology

* To help investors find the most proper neighborhood, we need to perform the following main work:

* 1. Import and install the required packages

* 2. Scrape the Wikipedia page to get the Toronto neighborhood information and transform the information into a dataframe

* 3. Get the 100 venues for each neighborhood throgh FourSquare API

* 4. Compute the total number of categories of all venues for each neighborhood

* 5. Generate the dummies for all categories of venues

Methodology

* 6. Group according to different neighborhoods and compute the total number of each category for each neighborhood.

* 7. Compute the total number of category whose name contains 'American Restaurant'

*   Compute the total number of category whose name contains 'Restaurant'

*   Compute the total number of category which don't belong to restaurant

* 8. Add three columns 'totals1','totals2','totals3' . 'totals1' is to compute the total of venues whose category is 'American Restaurant' for each neighborhood. 'totals2' is to compute the total of venues whose category is 'Restaurant' for each neighborhood. 'totals3' is to compute the total of venues whose category is 'not restaurant' for each neighborhood

Methodology

* 9. Find those neighborhoods in which there are no American restaurant and restaurants, in other word, find the dataframe which meets 'totals1' equals 0 and 'totals2' equals 0.

* 10. Based on the above obtained dataframe, find the neighorhoods in which there are the most venues which are not restaurants, in other words, find dataframe which meets 'totals3' equals the max.

* 11. Get the latitude and longitude of Toronto city

* 12. Create map of Toronto using latitude and longitude values and add the neighborhoods found by us to the map

Results

* The obtained neighborhoods are "CN Tower,Bathurst Quay,Island airport,Harbourfront West,King and Spadina,Railway Lands,South Niagara" in [Downtown Toronto](#) in [Toronto](#) city in Canada.

Discussion

* We can extend the range of selection through setting the threshold. Those neighborhoods whose total number of entertainment venues are greater than the threshold can also be used as the candidate neighborhoods.

* This concept of the method can be used to select other type of venues like gym, shop, bar, store, and so on.

Conclusion

* This project can help those investors who are going to open an American restaurant in Toronto city in Canada to look for most proper neighborhoods.

* The data used includes the Toronto neighborhood data and Foursquare location data.

* The method has used data analysis tools , data visualization tools and Foursquare API .