Data Science capstone project

Project name: Project_TopUS_Revenue
This is a capstone project for IBM Data Science Professional Certificate.

Introduction/Business Problem:

Target audience: business investors who are interested in investing in the US local businesses

As a business investor, I am interested to know what are the main revenue streams in the US and how those are geographically segregated. This knowledge is important as it guides me on how to invest my money at different locations/cities. Information on what business are booming at different locations can be very useful prior to the investment. Hence, the question is what are the main revenue generating local businesses within the US? And how those top revenue generating businesses geographically change within the US for most affluential cities?

Data Requirement:

First, I need to find main income generators in the US. This requires the knowledge of most affluent cities/regions and the most popular businesses in a specific city. I am going to use Wikipedia for this step. Second, I need to get information/data about the venues and businesses for these cities. I am going to use Foursquare API to extract data about venues/businesses.

- List of affluential neighborhoods in the US (Wikipedia)
- The latitude and longitude of those neighborhoods (Wikipedia)
- The average income for those neighborhoods (Wikipedia)
- Venue data for different neighborhoods (Foursquare API)