Heathiest and unhealthiest

cities in

USA for travelling

Coursera Capstone

IBM Applied Data Science

Namrata Dwivedi (2020)

***Introduction***

Traveling is great for many things, but it's not great for staying healthy. When daily routines are disturbed, indulgent restaurant food prevails, and parks and gyms are few and far between, wellness can easily become an afterthought.  There are some cities where staying healthy, or at least avoiding total gluttony, is surprisingly easy. To [stay healthy](https://www.farandwide.com/s/countries-live-longest-2b1e1e8a7f3045ab), you can walk the city, eat at a restaurant serving vegan-food, do a work out at the local gym and last but not the least indulge in a relaxing spa session.  WalletHub examined 174 cities in the United States, including the 150 most populated cities in the country based on these four dimensions like Fitness, Food, Green space, and Health care. The results are at turns expected and surprising (the healthiest city in the USA, for instance, comes out on top despite being known for its extremely rich cuisine).

So next time you travel you can ask this question “How fit is the city you travel”? Whether it has green spaces, vegan restaurants, yoga and Pilates centers, health-food shops, spas and gyms with sunshine along the average price for a one-night stay factored in as well.

***Business problem:***

The objective of this project was to analyses the top healthiest and unhealthiest city in USA and figure out how the above four factors play a key role in the wellness of the people who wish to travel in these cities. The researchers have calculated overall scores for each city based on a weighted average of all metrics. So, the next time you want to travel across USA and stay healthy the best way is to watch out for these factors to enhance your travelling experience.

***Target Audience:***

Now a days many people want to stay healthy even if they are travelling. The project focuses on people who wish to travel to these cities and find out the difference in the ways these cities are structured.

***Data Collection:***

For the project we needed data from there sources:

1.The data for the healthiest and unhealthiest cities in the USA will be obtained from wallet hub. The latitudes and longitudes are required to build an interactive map of the cities.

<https://wallethub.com/edu/healthiest-cities/31072/>

The healthiest city is San Francisco, we will need the neighborhood of the city along with the latitude and longitude. It will be obtained by <https://en.wikipedia.org/wiki/List_of_neighborhoods_in_San_Francisco> and python Geocoders package.

Similarly, the unhealthiest city is Brownsville in Texas. The data will be obtained in the similar manner as mentioned above for San Francisco

***Data Extraction:***

The wallet hub contains data from 174 cities in the country from healthiest to unhealthiest. We will use web scraping method to obtain the data. The interactive map will also be created by getting the latitude and longitude of the cites using python Geocoders library. Similarly, the data for the healthiest and unhealthiest cities will be extracted using web scarping from Wikipedia. The geographical coordinates can be obtained by python Geocoders library.

The Foursquare API is used to get the venues data for those cities. Foursquare lets users search for restaurants, gyms, spas, parks and other places in a location. The app displays personalized recommendations based on factors that include the time of day, a user's check-in history, their taste and the venue ratings.

This project will use the data science skills like web scarping, data cleaning, data wrangling , data visualization which we learned in the course to display our creativity in the project. In the project we will we present the methodology for the project.

**Methodology:**

The first thing is getting all the data for analysis. The wallet hub published the healthiest and unhealthiest cities in USA. We extracted the table for the healthiest and unhealthiest cities in Jupiter notebook using web scarping. The file is stored as a csv file. The data frame is divided between the top healthiest and non-healthiest cities. The Latitude and longitude of the cities were also obtained by python geocoders package. The interactive map was then plotted to show the location of the cities in the USA map. We colored the map in five different colors to difference the cities. A bar graph for the top five and bottom five cities are plotted using parameters like fitness, food and green space rank. The health care rank was not considered because of lack of data.

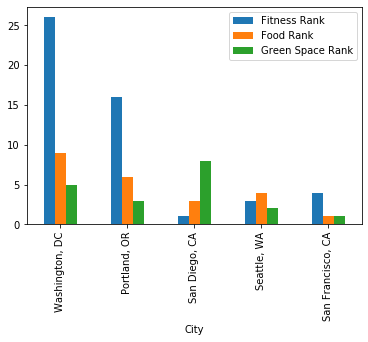
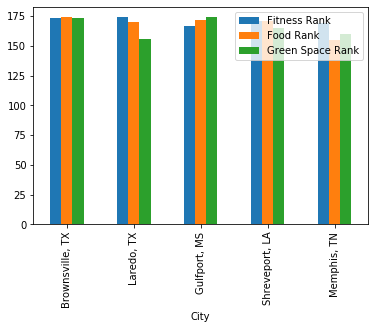
Healthiest cities

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Overall Rank** | **City** | **Total Score** | **Health Care Rank** | **Food Rank** | **Fitness Rank** | **Green Space Rank** | **Longitude** | **Latitude** |
| 5 | Washington, DC | 63.87 | 47 | 9 | 26 | 5 | -77.036571 | 38.894985 |  |
| 4 | Portland, OR | 65.66 | 61 | 6 | 16 | 3 | -122.674195 | 45.520247 |  |
| 3 | San Diego, CA | 70.01 | 25 | 3 | 1 | 8 | -117.162771 | 32.717421 |  |
| 2 | Seattle, WA | 70.62 | 19 | 4 | 3 | 2 | -122.330062 | 47.603832 |  |
| 1 | San Francisco, CA | 73.99 | 29 | 1 | 4 | 1 | -71.274327 | 46.844325 |  |

Unhealthiest Cities

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Overall Rank** | **City** | **Total Score** | **Health CareRank** | **Food Rank** | **Fitness Rank** | **Green Space Rank** | **Longitude** | **Latitude** |
|  | 174 | Brownsville, TX | 21.41 | 174 | 174 | 173 | 173 | -97.489086 | 25.914026 |
|  | 173 | Laredo, TX | 24.06 | 151 | 170 | 174 | 156 | -99.495376 | 27.519984 |
|  | 172 | Gulfport, MS | 24.82 | 171 | 172 | 167 | 174 | -89.092816 | 30.36742 |
|  | 171 | Shreveport, LA | 27.42 | 165 | 171 | 171 | 165 | -93.765194 | 32.522183 |
|  | 170 | Memphis, TN | 29.64 | 166 | 155 | 169 | 160 | -90.051629 | 35.149022 |

Healthy cities Unhealthy cities

The next step was to use foursquare API to get the 100 venues that are within 5000 meters radius. We need client ID and password to get those venues. We will then make an API calls to foursquare using geographical coordinates for our selected cities. It will return venue details like venue name, category, venue latitude longitude as a JASON file. We can also make a specific call using either the category id or name. Since we are interested in places like vegan restaurants, spas, gym and parks we can make those calls through four square API. We will then analyses top and bottom city for place like vegan restaurants, spas, gym and parks. We know that the healthiest city has more green spaces, gyms spas and vegan restaurants that is why the place is often tourist hotspot. These cities enhance the traveling experience of the people.

***Result:***

After the analysis of the city San Francisco and Brownsville Texas we figured that there is a lot of difference in the way these cities are structured.

San Francisco, which is the healthiest city, has more green spaces like parks to cater to people’s health. The vegan and vegetarian food choices are unlimited. We can also see that there are many spas and gyms in its neighborhood overall, it’s a perfect travelling destination.

Brownsville Texas is the unhealthiest city in the USA. We can hardly find parks in the area. The vegan and vegetarian food choices are rare. Though we can see that there are many spas and gyms in its neighborhood, but these two factors are not the only factor which need to be considered while travelling

***Discussion:***

Brownsville had the lowest percentage of physically active adults of any US city. It also has the fewest number of healthy restaurants. Brownsville also lacks walking trails, running trails per capita it has third-fewest walking trails per capita in the country. San Francisco had the most. When a city doesn’t take care of these issues, it can be hard to keep up good health. Brownsville scored an overall health score of 21.41, which just so happens to be about the lbs. of lard the average Brownsville resident consumes in a day. The city ranked dead last in health care and fruit/vegetable consumption because Hot Cheetos with cheese is Brownsville’s signature dish.

San Francisco beat out all other cities for having the most walking trails per capita, tied for first for having the most running trails per capita, finished in fourth for having the most healthy restaurants per capita and tied for fourth for having the lowest percentage of adults not eating enough fruits and vegetables per day, according to the report.

A lot of cities can boast parks, hills, restaurants, and other amenities, but the weather forces their citizens to stay inside for half of the year,

***Limitations:***

The project has certain limitations like we only compared the top and the bottom city. The foursquare API has limited access for non-paid users. We also did not consider the health care rank as the data for health care system was unavailable to us. We can further investigate the heath care parameter to further authenticate our finding. If we can compare all the cities in the USA, then our project would be more meaningful. We can further investigate our finding from Foursquare paid account to bypass all these limitations.

***Conclusion:***

In this project we have made an effort to summarize the travelling experience in the healthiest and unhealthiest city in the USA. map. Various location like restaurants spas parks and gyms are explored in the neighborhood of the cities and compared with each other. The findings of these projects will help the travelers explore the scope and limitation of the region with respect to various parameters.