Vibe of Boston

—

Marc Bacvanski and Vedant Rautela

Q: Where should I stay when I visit Boston?

(I have unique preferences)

We were inspired by the question, "Where should I stay when I visit Boston?" When we visit a new city, we do not know what to expect, but we are looking for a place to stay that matches our "vibe" – our individual preferences in terms of environment, attitude, and emotional state. No single "Guide to Visiting Boston" will suffice for all of us, since our preferences are very individual and unique.

19 Million

People visit Boston every year.

Where would I enjoy staying?

Boston's neighborhoods mean the city has a diverse set of environments to stay in. Where should visitors stay, to maximize the enjoyment of their experience?

19 million people visit Boston every year, with 1.3 million of those people being from overseas. While Boston is smaller than many similarly visited cities, Boston's set of neighborhoods mean that the city has a very diverse set of places to stay. Each neighborhood has its own personality, environment, and overall "vibe" than other neighborhoods, and people may prefer different qualities of the place they live in. We want to answer the question of where a visitor would enjoy staying, to maximize their enjoyment.
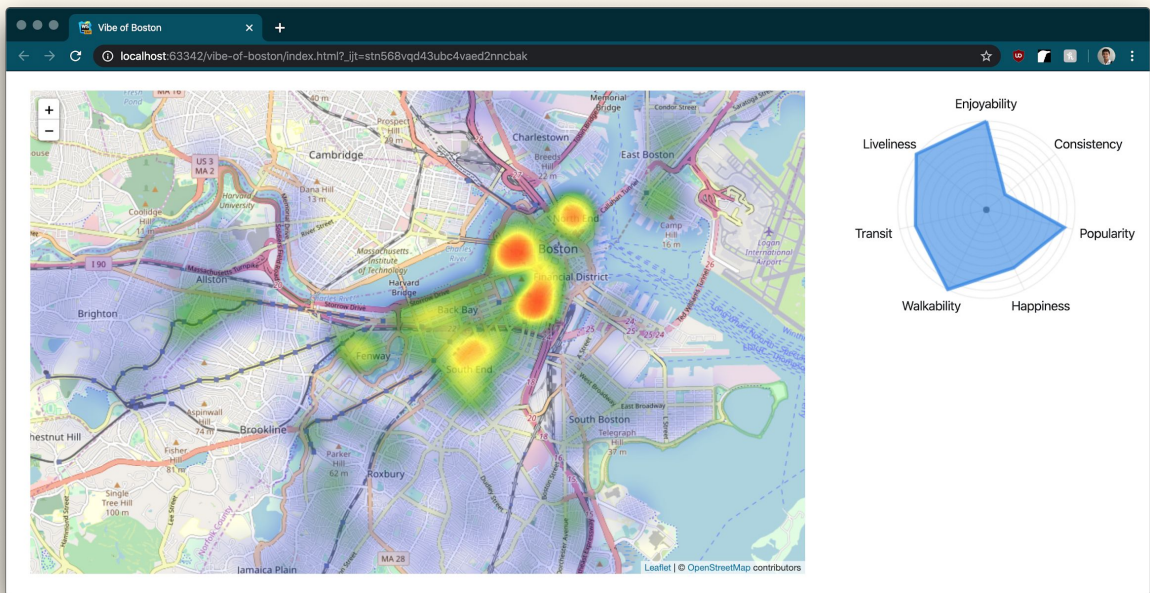
It looks like I'll enjoy **Beacon Hill** and **Financial District** the most.

The tool that we developed to allow users to answer this question is a heatmap which lets users see for themselves which areas of Boston they would enjoy the most. Here is a screenshot of our tool in action. Here, the user has inputted their preferences, giving high weights to Enjoyability, Liveliness, Walkability, and Popularity. To this user, Consistency does not matter so much, nor overall Happiness of the neighborhood. Given this individualized weighting of factors, our application has generated a custom heatmap of the areas of Boston that most match this vibe. The user can move the map around and zoom in, observing how Beacon Hill and Financial District are highlighted in red, and may be the best places for them to stay. Other areas that the user may find interesting have less intensity, such as North End, South End, and Back Bay.

Airbnb Dataset

- AirBnb listings, reviews, and bookings in the Boston area
- 2015-2020, sampled yearly
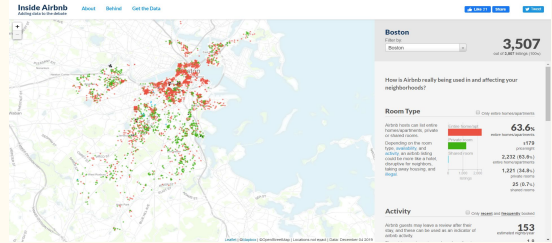
- 806,857 data points
  - 79,4377 reviews
 - 12,480 unique listings
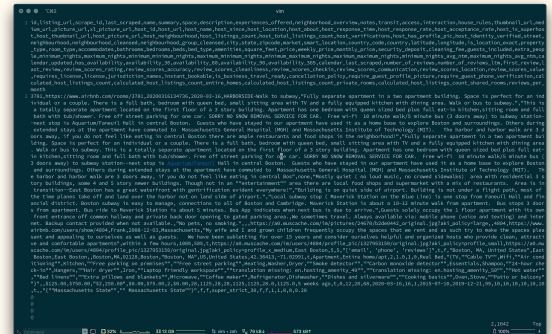
We performed our analysis on the Airbnb dataset of listings, reviews, and bookings in the Boston area. We used the publicly available data from 2015 to 2020, sampled yearly. This yielded us 806,857 unique data points, of which the majority were reviews. We chose the Airbnb dataset since it includes data from both residents of Boston (who write listings) and the visitors to Boston (who write reviews). We hope that the combination of this data will yield insight into how it is to live in Boston, from a diverse and inclusive perspective.

Vibe Factors

We calculate a vector of scores for each listing based on data in their ratings and reviews, according to these metrics:

- Enjoyability
- Consistency
- Popularity
- Happiness
- Walkability
- Transit
- Liveliness
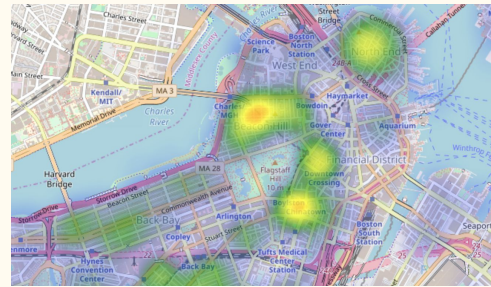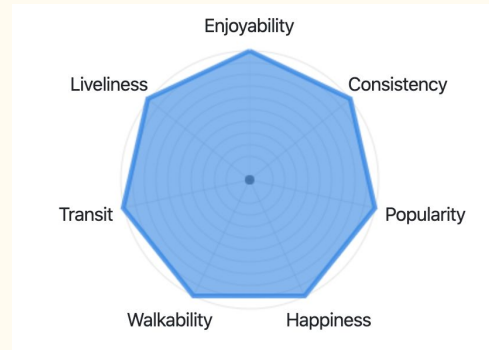
"Vibe" is a very vague word, and in the context of analyzing the "vibe of Boston", we wanted to offer the user the ability to quantify this emotional response. For each listing, we calculate a vector comprised of our analysis on both listings and reviews – this includes natural language processing, frequency analysis, numerical calculations, and correlating between reviews and listings. The seven vibe factors we've chosen to include are Enjoyability, Consistency, Popularity, Happiness, Walkability, Transit, and Liveliness. This is by no means a comprehensive list of all the features of Boston, but we decided that these factors are all relatively common ones that people look for in determining where to stay.
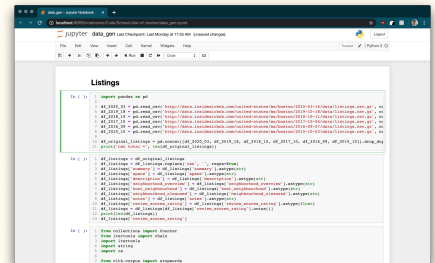
Data Analysis

To process and analyze the data we used:

- **Pandas** - Python data manipulation library
- **nltk** - Python Natural Language Processing library

We publish a Jupyter notebook containing all our data analysis techniques in Python.

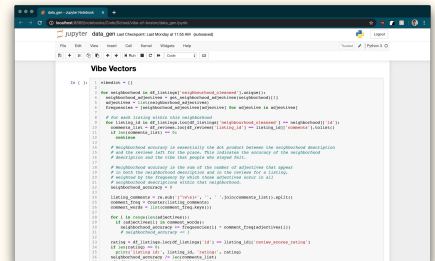Data is exported as JSON, which is loaded by the visualization tool

We performed our data analysis in Python, using the Pandas and nltk library. Our Jupyter Notebook highlights our approach to cleaning, analyzing, and preparing the data for visualization. As mentioned earlier, we use NLP to analyze the listings and review data, specifically by extracting the most common adjectives from the listings and comparing them with the most common adjectives from the reviews. This is how we calculate a vibe vector for a given listing and by performing this process for all the listings we are able to gather data about the vibe of the different areas of Boston as a whole. This data is exported as a large JSON file, which is then loaded by our user-facing visualization tool.

Visualization

We developed a web page that loads the data outputted from Python

- **leaflet-heat.js** - OpenStreetmaps plugin for custom layers
- **spidergraph.js** - JavaScript library for creating the vibe alignment chart

Adjustable weights: Prioritize the factors that are most important to the user

```
102      $('#spidergraphcontainer').spidergraph('setactivedata', {
103          'strokecolor': 'rgba(46,131,230,0.8)',
104          'fillcolor': 'rgba(46,131,230,0.6)',
105          'data': [defaultWeights['rating'],
106              defaultWeights['n_accuracy'],
107              defaultWeights['reviewcount'],
108              defaultWeights['happiness'],
109              defaultWeights['walkability'],
110              defaultWeights['transit'],
111              defaultWeights['liveliness'],
112          ],
113          'linear': true,
114      });
```

```
125      function processVibeTable(data, weights, map) {
126          let biggestVibe = 0;
127          let formattedData = [];
128          Object.keys(data).forEach(key => {
129              const point = data[key];
130              const vibe = calcVibe(point, weights);
131              const arr = [point['lat'], point['lon'], vibe];
132              if (vibe > biggestVibe) {
133                  biggestVibe = vibe;
134              }
135              formattedData.push(arr);
136          });
```

We have developed a web page that features our interactive heatmap and vibe alignment customizer. We utilized OpenStreetmaps as our base map, with the Leaflet library to create our custom heatmap layer atop the map. We use the spidergraph library to create our interactive alignment charts, allowing the user to prioritize the factors most important to them.

Impact

Airbnb data sheds light on the intangible aspects of living in Boston, both from the perspective of visitors and residents.

Our project presents findings from Airbnb data in a directly actionable format.

The tool we created will be useful for visitors making decisions on where to stay in Boston, as well as residents to better understand the complex dynamics of their city.

With this project, we hope to have shed light on the intangible aspects of living in Boston, that cannot be found through reviews on Google Maps or "How to visit Boston" guides. We were able to analyze and present findings from Airbnb in a directly actionable format, guiding people's decisions on where to stay in Boston. We hope that this tool will be useful not just to potential visitors of Boston, but to college students and residents of Boston, helping them better understand the complex dynamics of a great city.

Future Work

- Add more vibe factors
  - More cultural aspects (i.e. art, music, sports, food, etc.)
  - Allow users to input vibe factors of their choice
- Display change in vibe over time: four-dimensional heatmap
- Include data from other cities: global vibe map
- More user adjustability of individual vibe factor calculations
- Process data using a faster language than Python, enabling user-side computation

Future work on this project would include adding more vibe factors, such as culture, and allowing users to programmatically add vibe factors of their choice. Given the amount of data available from Airbnb, it would be interesting to display the change in vibe over time. Airbnb has locations all over the world, and so expanding to create a global vibe map would be exciting. Additionally, we'd like to integrate data from sources other than just Airbnb (i.e. Twitter, Reddit, etc.). These other sources would provide us a more comprehensive set of data regarding the vibe of Boston and sample certain sections of the Boston population that may be missing from the Airbnb data alone. Finally, rewriting the data processing portion using a faster language such as Golang or Rust would enable user-side vibe computation, which would open the door to extensibility and greater scale.