Vacation Recommendations

Problem

Vacation planning under time pressure:

- Not doing research on what to see
- Not checking best time of the year to visit an attraction
- Picking a wrong season to visit a city
- Picking a wrong city to visit on vacation

Objective

- Build a vacation recommender for a tourist with limited travel time flexibility.
  - Propose a city for vacation based on the month when the vacation is planned
  - Propose top N attractions to visit in a given city in a given month.

WorkFlow

Data Acquisition

- TripAdvisor “Things to Do” using Selenium
- Top 10 attractions in Austin, Chicago, and New York City
- Hotels, restaurants, and landmarks are excluded

Data Preprocessing

- Pandas and Numpy
- NLTK

Vectorization

- CountVectorizer
- TFIDF

Topic Modeling

- LSA
- CorEx

Sentiment Analysis and Seasonality Score

- VADER
- CorEx
- Linear Regression

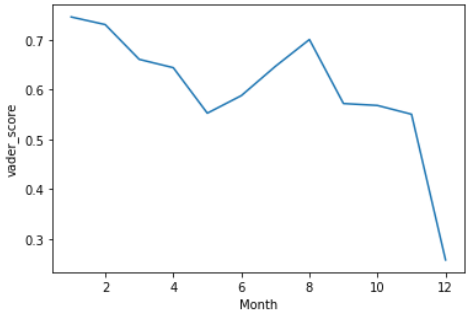
Recommendations

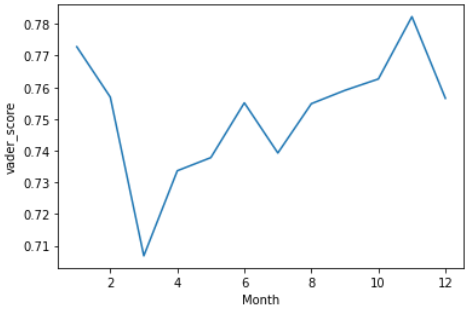
- Propose a city for vacation based on month
- Propose top N attractions in a city based on month

Sentiment Analysis

Austin

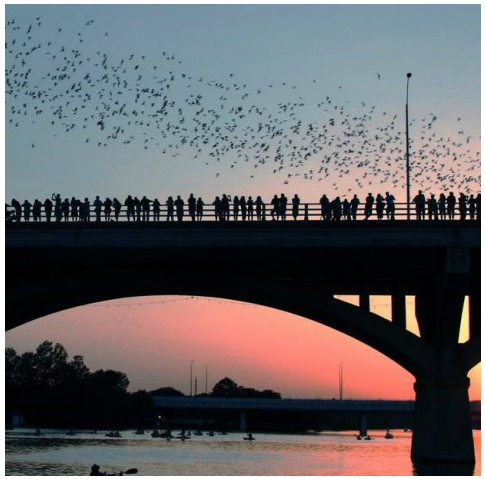
Congress Ave Bridge

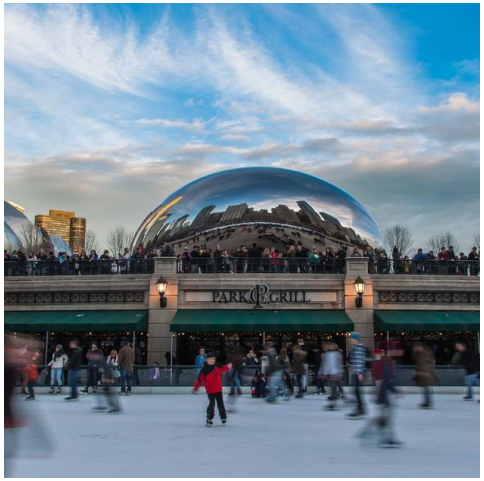




Chicago

Millenium Park



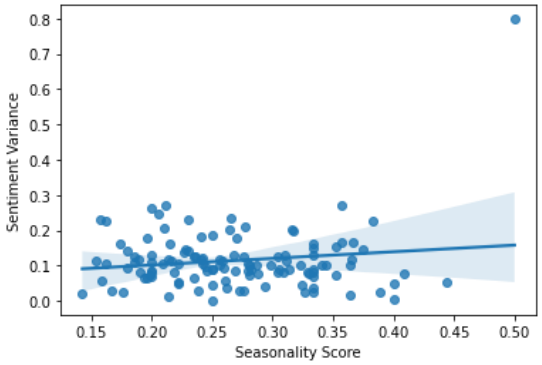


LSA and CorEx Models

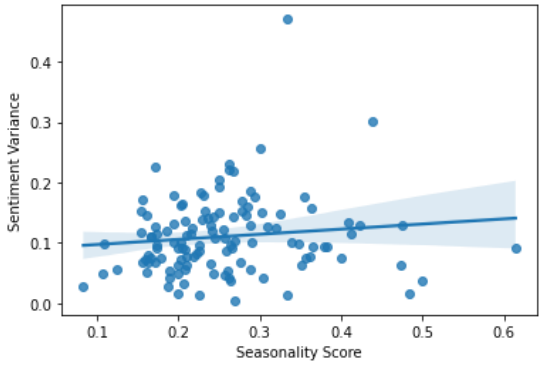
- Explored frequently discussed topics in each cities' reviews.
- Created topics based on anchors: 'time', 'cold', 'freeze', 'weather', 'hot', 'warm'
- Constructed a seasonality score based on the percentage of reviews discussing the season related topics
- Explored whether the variation in tourist sentiments can be attributed to weather or season related concerns.

Results: CorEx Model and Linear Regression Analysis

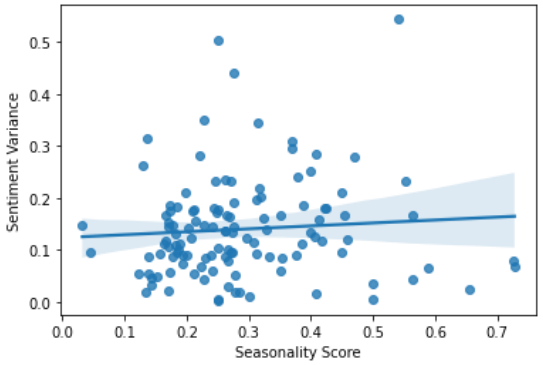
Chicago



Austin



New York City



Recommendations

Recommender based on sentiment score:

Users choose:

- a month when their vacation is planned
- number of top attraction they would like to visit

Example:

- Top city to visit in May is New York City
- Top 5 attractions to visit in May are:
  - Empire State Building
  - Bryant Park
  - Central Park
  - Radio City Hall
  - Metropolitan Museum of Art

Future Work

- Use spacy for text preprocessing
- Complete integration of ratings to create two-dimensional recommendations
- Incorporate seasonality score into the recommender.