

# **A Recommendation System for Museum Selection**

Anne Chen

09/18/16

# Project Overview

## Why Built this App

- There is no existing solution to locate museums based on one's preference
- Finding a right museum is time consuming

## Data Collection and Visualization

- Scraped ~1600 data from TripAdvisor

## Recommendation System Construction

- Engineered new features including sentiment analysis scores
- Imputed missing values with KNN
- Developed content-based recommendation system
- Implemented system into Shiny App

# The App is Built for People to Locate Museums of Interest

## 1. Explore museums based on users' preferences

Museums Finder(Beta) Get Recommendation Museum Map About This App

1. Search or Scroll Down to Select Museum(s) of Interest:  
The Metropolitan Museum of Art

2. Hit the Button to Get Recommendation  
>> Recommend <<

	1. Chrysler Museum of Art	2. Kelvingrove Art Gallery and Museum	3. Tenement Museum	4. The Museum of Modern Art (MoMA)	5. Philadelphia Museum of Art
Similarity Score:	0.99900	0.99888	0.99840	0.99816	0.99807
Rating:	5	4.5	4.5	4.5	4.5
Type:	Art Museums	Art Museums	Specialty Museums	Art Museums	Art Museums
Tags in Common:	Special Exhibits	Van Gogh few Hours		Special Exhibits Van Gogh	Special Exhibits Impressionists Armour
Address:	245 W Olney Rd, Norfolk, VA 23510-1509	Kelvingrove Art Galleries Argyle Street, Glasgow G3 8AG, Scotland	103 Orchard St, Lower East Side, New York City, NY 10002-3102	11 West 53rd Street, New York City, NY 10019	2600 Benjamin Franklin Pkwy, Philadelphia, PA 19130-2399
Description:	N/A	Kelvingrove Art Gallery and Museum is Scotland's most visited free attraction. With 22 themed, state-of-the-art galleries displaying an astonishing 8000 objects, the collections are extensive, wide-ranging and internationally-significant.	The Tenement Museum preserves the history of immigration through the personal accounts of those who built lives in the Lower East Side. Visitors can view restored apartments from the 19th and 20th centuries, walk the historic neighborhood, and interact with	Located in the heart of midtown Manhattan, The Museum of Modern Art is home to the world's greatest collection of modern and contemporary art. The collection includes Vincent van Gogh's "The Starry Night" and Andy Warhol's "Gold Marilyn Monroe," along with	As one of the largest museums in the United States, the Philadelphia Museum of Art invites visitors from around the world to explore its renowned collections, acclaimed special exhibitions, and enriching programs, both in person and

## 2. Allow users to filter museums that meet their needs

Museums Finder(Beta) Get Recommendation Museum Map About This App

Museum Filter

Museum Types

Good For:

- ☐ All Ages
- ☐ Family
- ☐ Children
- ☐ Rainy Day

Suggested Length of Visit:

- ☐ <1 hour
- ☐ 1-2 hours
- ☐ 2-3 hours
- ☐ >3 hours

Other Options:

- ☐ Free Admission
- ☐ Has Gift Shop
- ☐ Has Audio guide

Search

# Scraped Data of ~1600 Museums from TripAdvisor

## Museum Info

- Museum name
- Address\*
- Category
- Relative Popularity\*\*
- Times being featured
- Recommended length of visit
- Fee
- Description
- Displayed rating (0-5)
- Image

## Traveler & Review Info

- Number of reviews
- Review tags
- Traveler rating
- Traveler types
- Quotes
- Review contents

\* Addresses were passed to Google Geocoding API to convert them to latitude and longitude  
 \*\* Relative Popularity is later computed as X/Y where rank X of Y things to do in the same city

# Generate New Features with Museum Categories, Review Tags, and Textural Data

## 83 Distinct Categories

- Art Museums
- Children's Museums
- History Museums
- Natural History Museums
- Science Museums

## 100 Most Frequent Review Tags

- Gift shop
- All ages
- Special exhibits
- Few hours
- Well worth a visit
- Permanent collection
- Free admission
- Audio guide
- Beautiful building

## Textural Data

10 Quotes

10 Reviews

Museum Description

## Sentiment Analysis

Polarity & Subjectivity

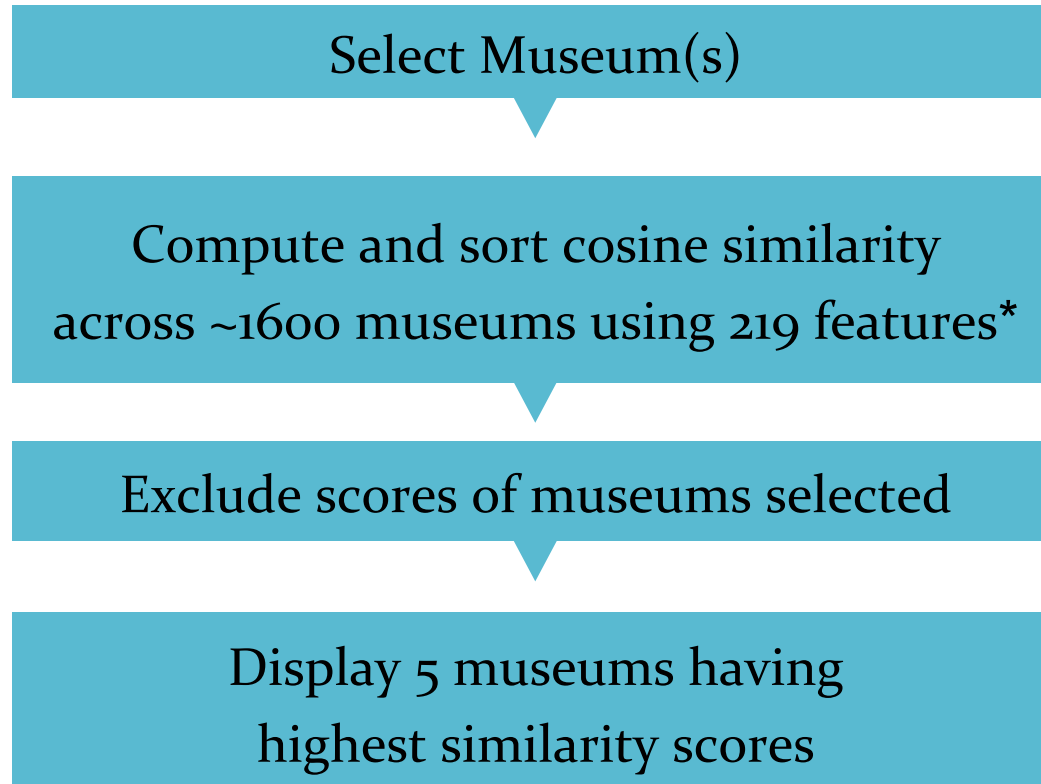
Max

Average

Min

Variance

# Constructed a Recommendation System that Based on Cosine Similarity



\*"ReviewCount" and "TotalThingsToDo" were removed

**Demo Time!**

**Questions?**



# Other Slides

# If I got more time...

- Predict if a museum will be featured or not
- Cluster museums based on review polarity or subjectivity
- Apply association rules to see if there are rules for combinations of review tags

# Imputed Missing Values Using KNN and Dummified Categorical Variables

- Imputed missing values with KNN using  $k = 40$
- Created dummy variables for “Country”, “Fee”, and “Recommended Length of Visit”

# **Rank, Review Polarity, and Quote Subjectivity are Important Features in Predicting Rating**

- Random Forest (Classification)
- Some important features:
  - Rank & Relative Popularity
  - Review Polarity
  - Quote Subjectivity
  - Review Count