Weather & Happiness 

Machine Vision Final - Spring 2017

Chris Dalke and Nate Conroy

Introduction

Project Overview

Is there a relationship between climate features and happiness?

We seek to:

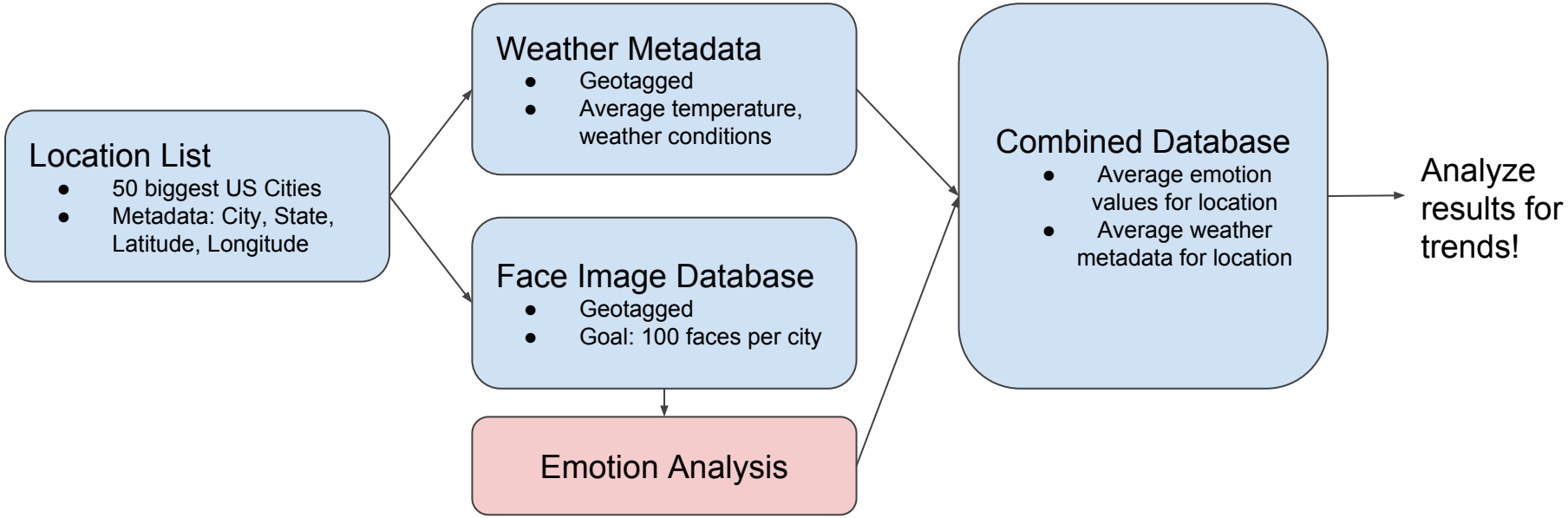
- Gather a large dataset of geotagged images containing faces
- Evaluate their emotion ratings
- Plot them with data obtained from climate databases
- Look for trends

Disclaimer

- Machine and human vision both make judgements from limited data
- The perceived emotion of a person in an image doesn’t necessarily correlate with their actual happiness.
- This is a limitation with the medium of photography; analyzing perceived emotion is still a valuable machine vision project.

Process Summary

Process Summary

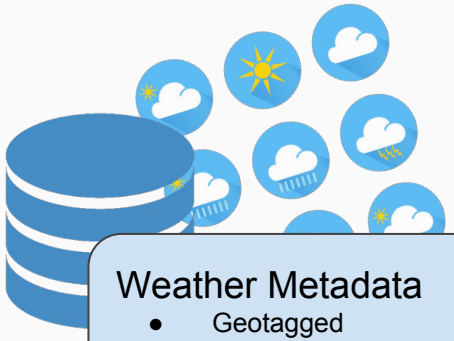


Process Summary - Data Sets

We'll need to collect 3 data sets, as shown below.

Location List

- 50 biggest US Cities
- Metadata: City, State, Latitude, Longitude



Weather Metadata

- Geotagged
- Average temperature, weather conditions

Face Image Database

- Geotagged
- Goal: 100 faces per city

Data Collection

Data Collection: Location List

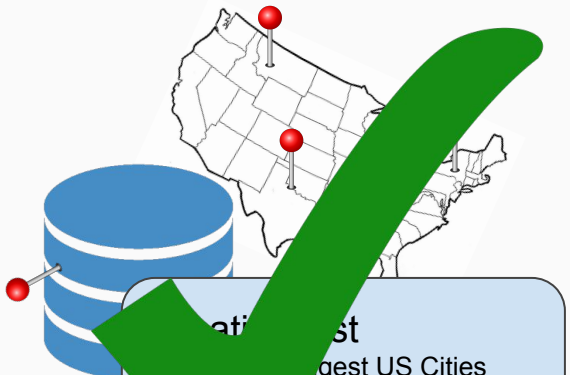
Which locations are we going to look at?

Rural areas are unlikely to have a significant amount of geotagged photos. Look at cities; large amount of data

Compiling a list of 50 cities and their locations is easy to do by hand.

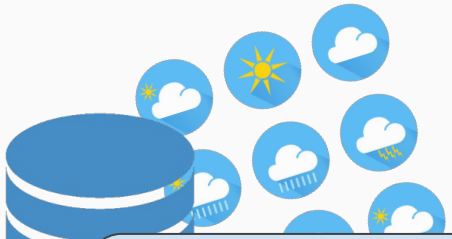
We built a database of the 50 largest US cities, along with their latitude and longitude values for use in weather and face collection.

Process Summary - Data Sets



...st

...gest US Cities
...data: City, State,
...tude, Longitude



Weather Metadata

- Geotagged
- Average temperature, weather conditions

Face Image Database

- Geotagged
- Goal: 100 faces per city

Data Collection: Weather Metadata

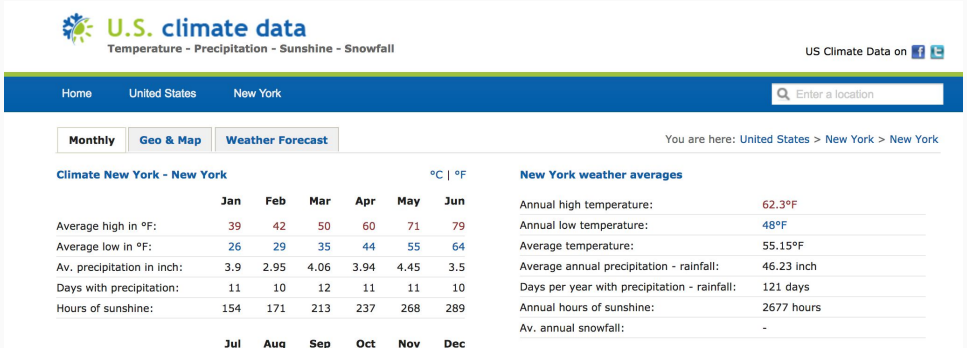
Localized weather conditions are not indicative of overall weather conditions, therefore analyzing weather in photos will not give us meaningful long-term weather data.

Filters and image editing will also cause inaccuracies.

Solution: Use databases for weather data trends fused with location data

Data Collection: Weather Metadata

- Aggregated climate data from USClimateData.com
- Data from each city
  - Average temperature
  - Average days with precipitation per year
  - Average hours of sunshine per year



U.S. climate data

Temperature - Precipitation - Sunshine - Snowfall

US Climate Data on

Home | United States | New York | Enter a location

Monthly | Geo & Map | Weather Forecast

You are here: United States > New York > New York

Climate New York - New York °C | °F

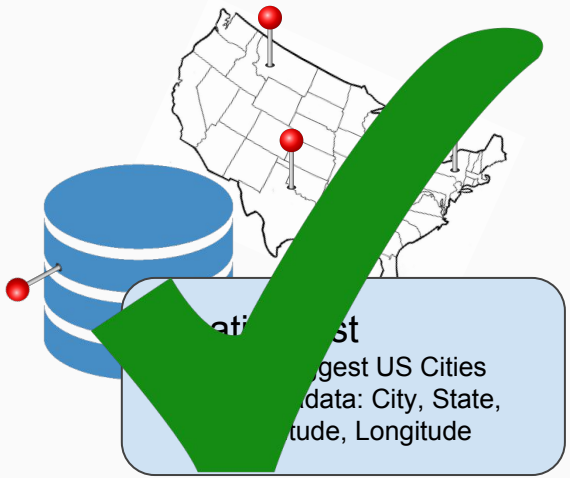
	Jan	Feb	Mar	Apr	May	Jun
Average high in °F:	39	42	50	60	71	79
Average low in °F:	26	29	35	44	55	64
Av. precipitation in inch:	3.9	2.95	4.06	3.94	4.45	3.5
Days with precipitation:	11	10	12	11	11	10
Hours of sunshine:	154	171	213	237	268	289

	Jul	Aug	Sep	Oct	Nov	Dec

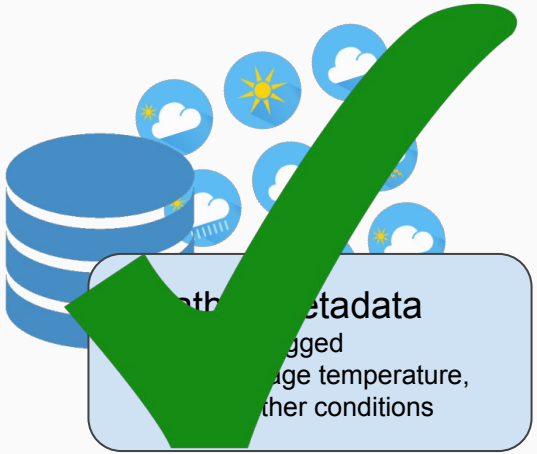
New York weather averages

Annual high temperature:	62.3°F
Annual low temperature:	48°F
Average temperature:	55.15°F
Average annual precipitation - rainfall:	46.23 inch
Days per year with precipitation - rainfall:	121 days
Annual hours of sunshine:	2677 hours
Av. annual snowfall:	-

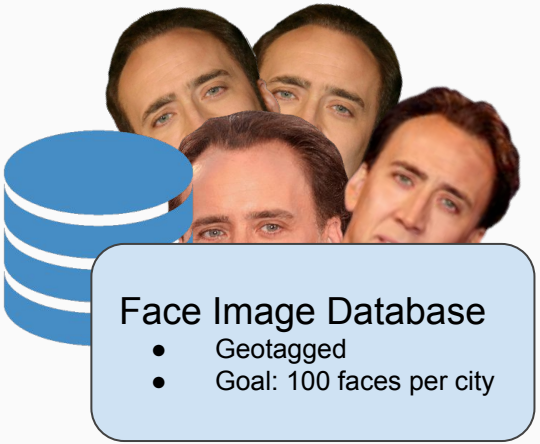
Process Summary - Data Sets



st
gest US Cities
data: City, State,
tude, Longitude



etadata
gged
ge temperature,
ther conditions



Face Image Database

- Geotagged
- Goal: 100 faces per city

Data Collection: Face Image Database

Use Flickr APIs to search for publically available photos geotagged near each city.

Eliminating landscape photography: add the “person” tag to searches.

- Decreases false positives (Photos we download but discard)
- Increases false negatives (Photos with faces we missed).
- An acceptable tradeoff. Doesn’t waste time / power to miss photos.

Sort by “Interestingness” (Flickr API setting) to avoid chronological photo albums.

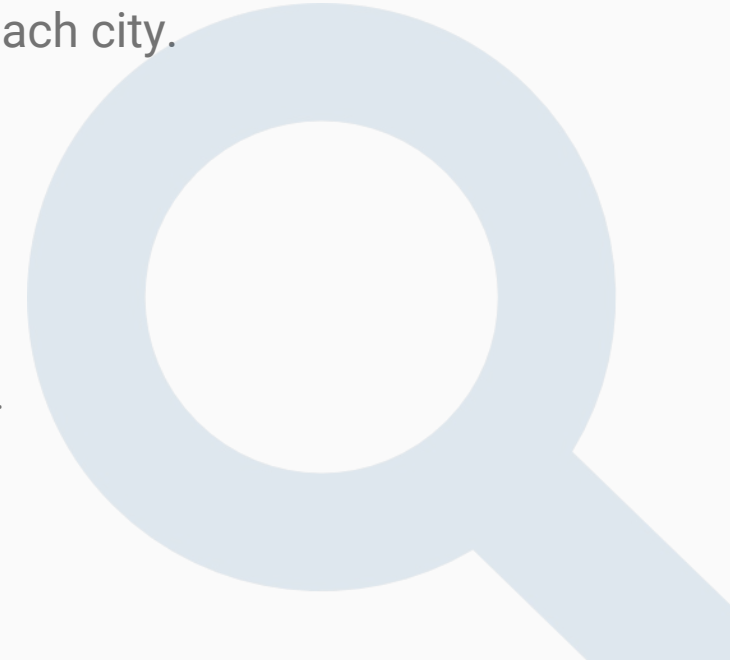
Data Collection: Face Image Database

Goal is to obtain 100 photos containing faces for each city.

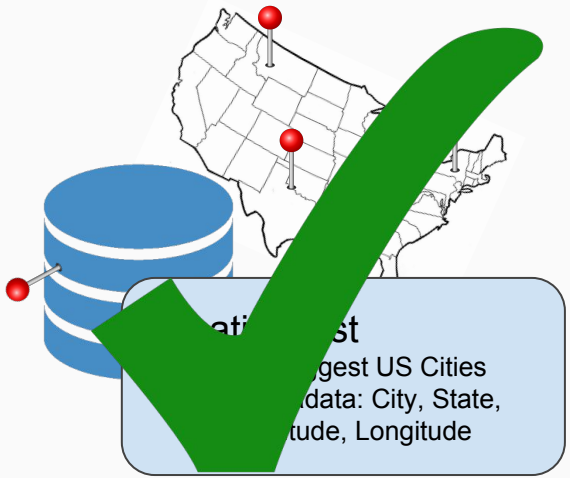
Algorithm:

- Download new batch of images from Flickr
- Perform facial analysis on photos in this batch
- Discard images that don't contain a face
- Continue until we've reached 100 face images for each city.

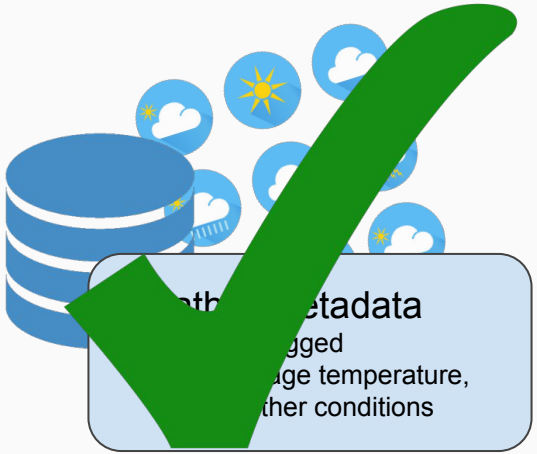
Note: This takes time! *Rate limiting* for API calls.



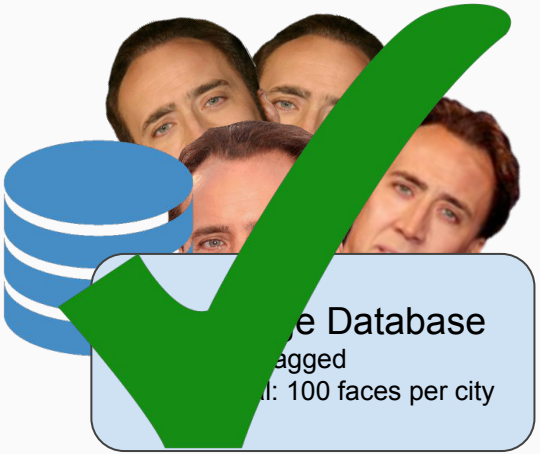
Process Summary - Data Sets



st
gest US Cities
data: City, State,
tude, Longitude



etadata
gged
ge temperature,
ther conditions



e Database
agged
l: 100 faces per city

Emotion Analysis

Emotion Analysis

Goal data set of 5000 face images (50 cities, 100 images per city)

Huge amount of calculation to perform locally!

Local Method: Download each Flickr image, use openCV and machine learning algorithm to identify facial emotions. HARD + SLOW!

Cloud Method: Microsoft Emotion API lets you send an image URL and it will return JSON data about any faces it finds in the image. EASY + FAST!

Emotion Analysis

We chose to use the Microsoft Emotion API since it is easy to use and fits within our time constraints.

Inputs: Image URL, API subscription key

Outputs: Bounding box coordinates, emotion data for each face found in the image in JSON form (each value represented as a number between 0 and 1)

- Emotions identified: Anger, Contempt, Disgust, Fear, Happiness, Neutral, Sadness, Surprise

Emotion Analysis

Example of a response:

- Returns **bounding box**
- Returns **normalized probabilities** for each emotion
- NOT an absolute score; represents the model's confidence across a set of emotions.


```
{
    "faceRectangle": {
      "left": 488,
      "top": 263,
      "width": 148,
      "height": 148
    },
    "scores": {
      "anger": 9.075572e-13,
      "contempt": 7.048959e-9,
      "disgust": 1.02152783e-11,
      "fear": 1.778957e-14,
      "happiness": 0.9999999,
      "neutral": 1.31694478e-7,
      "sadness": 6.04054263e-12,
      "surprise": 3.92249462e-11
    }
  }
```


Microsoft

Emotion Analysis

- Loop through all images we have downloaded and saved metadata for.
- For each image, use Emotion API to calculate the emotion probability matrix. Save these values into Photo object in database.
- Repeat between this and data collection until we have a sufficiently large database.

Emotion Analysis

Upon gathering all of our individual emotion data for each image, we compiled the data to obtain average emotion values for each city in our set

We did this by summing the decimal fractions for each emotion value found in images in each city, and then normalizing those sums within each of the cities

The resulting values are, roughly, the percentages of each emotion

Note: Still subject to drawback discussed in disclaimer before: “average emotion value” is really just “average perceived emotion in images selected”

Data Pruning

Data Pruning

Various Housework:

- Compile emotion probabilities for each location, normalize values
- Trim remaining faceless images from database
- Save complete data set for photos and locations to a single JSON file which can later be used in visualization tool.

Resulting Data Set

Final Data Set

- 50 US Cities with Metadata (Latitude, Longitude, City, State)
- Weather Data for 50 US Cities
- 5000 Face Images with Emotion data (50 cities, 100 images per city)
    - *Note: Data set is currently 376 images, still collecting full data set.*

Ready for Analysis! ✓

Results & Analysis

Visualization

Data Visualization

We need a way to analyze the data we have produced and display it

Main objectives:

- Ensure the program is identifying all faces in images
- Ensure the face bounding boxes match actual faces
- Ensure the emotion tags (mostly) match the real emotions of the faces

Data Visualization

Building a visualization system will also help with analysis. Allow navigating through each image, sorting by city, etc. Should display charts of the city scores.

Good system for creating this: HTML/JS. Our database is saved into JSON which can easily be read and traversed in JS, and an HTML page is a good framework for producing user interfaces quickly.

Data Visualization

Resulting Program: Web-based Data Visualization Tool

Features:

- Navigate through complete data set of images
- View face bounding box and emotion probabilities for each image
- View interactive chart of summed emotion value for all cities

Live Demo

chrisdalke.com/csc249

Information

CSC 249 Final Project - Data Visualization Tool

Chris Dalke & Nate Conroy

Image Navigation

Num Images: 376
Current Image: 76
Next Photo
Previous Photo
Next Photo With Face

Image Metadata

Dallas,TX
(32.7757,-96.7967)
No Faces

[Face 0]

Emotion	Value
contempt	0.00009684803
surprise	0.0000031540203
happiness	0.9969788
neutral	0.0028259505
sadness	0.00006325432
disgust	0.000015422369
anger	0.000016335222
fear	1.9200019e-7

[Face 1]

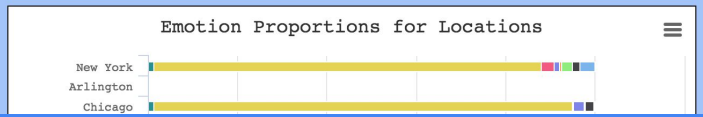
Emotion	Value
contempt	0.11510897
surprise	0.001103519
happiness	0.0032744803
neutral	0.8496792
sadness	0.019857801
disgust	0.0013917868
anger	0.009517134
fear	0.00006715171

Image Viewer

Location Metadata

Num Locations: 50
Num Locations with Emotion Data: 46

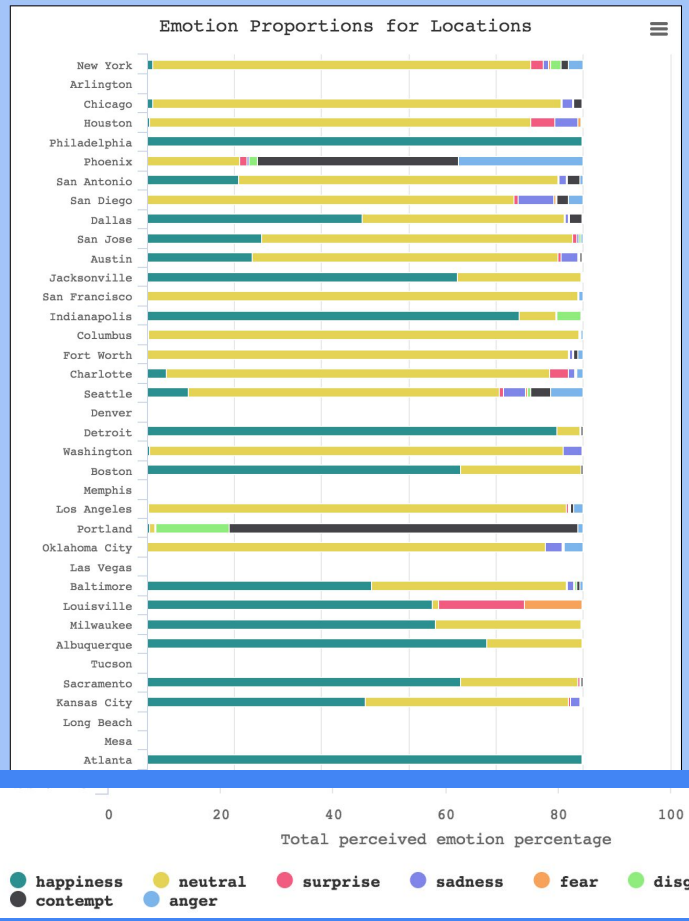
Location Emotions



Location Metadata

Num Locations: 50
Num Locations with Emotion Data: 46

Location Emotions



Additional Visualizations (Todo)

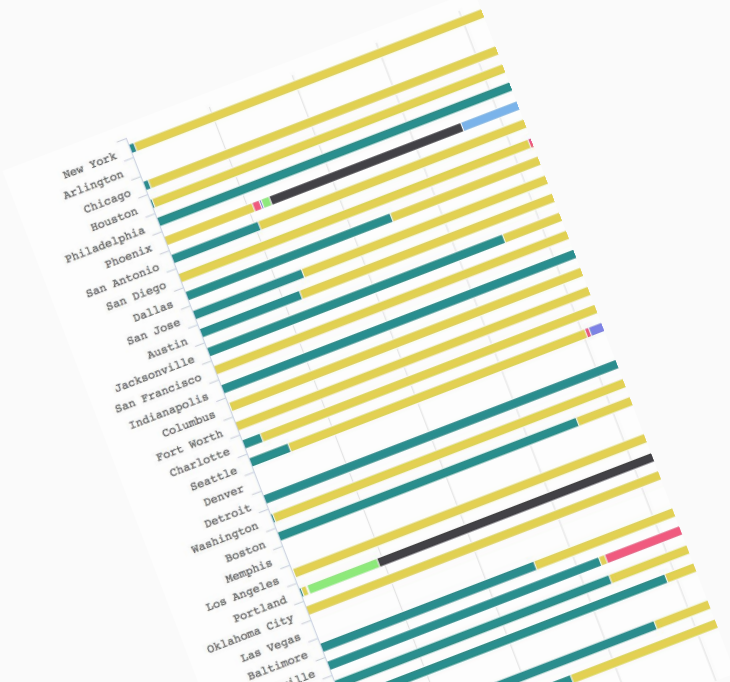
Location-based heatmap showing happiness score on US map as color.

If the heatmap demonstrates a gradient, there is a correlation between location and happiness.

This visualization would take advantage of displaying the data spatially to clearly show location-based trends.

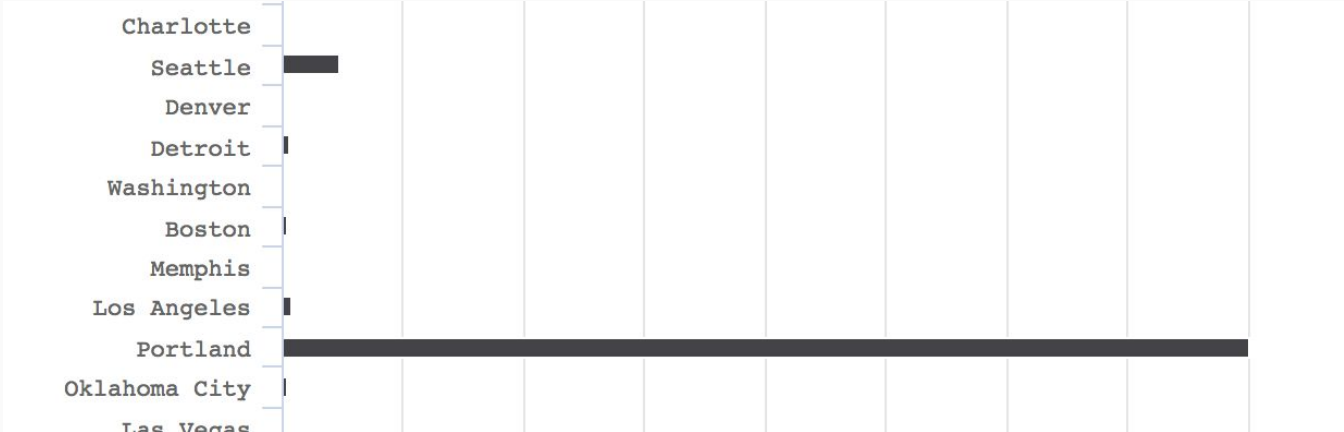
Results

- Resulting system can display emotion distributions for the photos from each of our cities.



Failure Cases

- Our system relies on large data sets to be reliable.
- Example failure case in current data: Portland is 80% contemptful (It only has a single photo, and the person is not smiling!)



Conclusion

Conclusion

We were able to successfully access a public database of images, form a dataset of images geotagged within our targeted locations, filter out images without faces, and compile emotion data using Microsoft's Emotion API

We found that a very large dataset is required in order to stabilize the data.

Moving forward we plan to collect a dataset of 5,000 images (100 from each city) to more accurately form conclusions involving relationships between climate and emotion

Works Cited & Previous Work

Works Cited & Previous Work

1. Roser, Martin, and Frank Moosmann.”Classification of weather situations on single color images.” IEEE Intelligent Vehicles Symposium (2008):798-803. IEEE Xplore Digital Library. Web. 27 Mar. 2017.
2. Saeed Abdullah, Elizabeth L Murnane, Jean MR Costa, and Tanzeem Choudhury. Collective smile: Measuring societal happiness from geolocated images. In Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing, pages 361–374. ACM, 2015.
3. Suchitra, Suja P. and S. Tripathi, ”Real-time emotion recognition from facial images using Raspberry Pi II,” 2016 3rd International Conference on Signal Processing and Integrated Networks (SPIN), Noida, 2016, pp. 666-670
4. Yuchen Wu, Jianbo Yuan, Quanzeng You, Jiebo Luo, The Effect of Pets on Happiness: A Data-Driven Approach via Large-Scale Social Media, In Big Data (Big Data), 2015 IEEE International Conference on. IEEE, 2016

The End