

Identifying and Avoiding Areas At Risk for COVID-19 and Other Natural Disasters

Sophia Scarano

Ogenevwede A Agboro-Jimoh

Shiffraw Dagnechaw

PROBLEM STATEMENT

It is critical to mitigate and tame the diffusion of COVID-19 among people. Increased Social distancing is associated with lower rates of COVID- 19. Providing a navigation tool that will equip people with information in regards to the density of COVID cases, in their vicinity, can potentially decrease the general exposure to COVID-19. Additionally, the navigation tool will provide functionality to view the areas currently affected by wildfires, whether these fires are contained or ongoing so that the user can avoid these areas accordingly.

Peer Reviewed Research

Social distancing to slow the US COVID-19 epidemic: Longitudinal pretest–posttest comparison group study

Mark J. Siedner  , Guy Harling , Zahra Reynolds, Rebecca F. Gilbert, Sebastien Haneuse, Atheendar S. Venkataramani, Alexander C. Tsai

Published: August 11, 2020 • <https://doi.org/10.1371/journal.pmed.1003244>

Conclusions

Statewide social distancing measures were associated with a decrease in the COVID-19 case growth rate that was statistically significant. Statewide social distancing measures were also

Primary Goal

The underlying goal is to construct an application that locates proximal gas stations within California, upon user request, by mechanism of minimizing exposure to covid hotspots. The application will additionally display location of fires that are ongoing in order to reduce the risk of the user.

General Overview of App

Application Functionality

User Inputs their location in terms of longitude and latitude. User also inputs maximum distance willing to travel to gas station



The output will yield gas stations in their proximity and will rank each gas station on a scale from 0 to 4 in order of increased probability of COVID exposure

Data Acquisition

Data

- COVID-19 data was acquired from the Los Angeles Times Data and Graphics Department
- The dataset that was utilized was “latimes-place-totals”

california-coronavirus-data

The Los Angeles Times' independent tally of coronavirus cases in California.

 [launch binder](#)  [Jupyter Notebook tests](#) [passing](#)

Table of contents

- [latimes-agency-totals.csv](#)
- [latimes-county-totals.csv](#)
- [latimes-state-totals.csv](#)
- [latimes-agency-websites.csv](#)
- [latimes-place-totals.csv](#)
- [cdph-state-totals.csv](#)
- [cdph-positive-test-rate.csv](#)
- [cdph-age.csv](#)
- [cdph-race-ethnicity.csv](#)
- [cdph-skilled-nursing-totals.csv](#)
- [cdph-adult-and-senior-care-totals.csv](#)
- [cdph-skilled-nursing-facilities.csv](#)
- [cdph-adult-and-senior-care-facilities.csv](#)
- [cdph-nursing-home-county-totals.csv](#)
- [cdph-hospital-patient-county-totals.csv](#)
- [cdph-reopening-tiers.csv](#)
- [cdph-reopening-metrics.csv](#)
- [cdcr-state-totals.csv](#)
- [cdcr-prison-totals.csv](#)
- [latimes-project-roomkey-totals.csv](#)
- [latimes-beach-closures-county-list.csv](#)
- [latimes-beach-closures-area-list.csv](#)
- [los-angeles-countywide-statistical-areas.json](#)

Data

field	type	description
date	date	The date when the data were retrieved in ISO 8601 format.
county	string	The name of the county where the city is located.
fips	string	The FIPS code given to the county by the federal government. Can be used to merge with other data sources.
place	string	The name of the city, neighborhood or other area.
confirmed_cases	integer	The cumulative number of confirmed coronavirus case at that time.
note	string	In cases where the <code>confirmed_cases</code> are obscured, this explains the range of possible values.
x	float	The longitude of the <code>place</code> .
y	float	The latitude of the <code>place</code> .

Data Cleaning

- The column that contains covid cases was normalized by mechanism of logarithmic transform
- A column was added that mapped each logarithmic value of a covid case to an ascending scale from 0 to 4 (0 representing a relatively small amount of COVID-19 cases and 4 representing a large amount of COVID-19 cases)
- The features that were utilized to predict the COVID-19 ordinal rank were: longitude and latitude, county, and place.

Predictive Modeling

Models Tested

Transformers	Estimators
One Hot Encoder, PCA, LDA	Logistic Regression
One Hot Encoder, PCA, LDA	Support Vector Machine
One Hot Encoder, PCA, LDA	Random Forests Classifier
One Hot Encoder, PCA, LDA	Ada Boost Classifier
One Hot Encoder, PCA, LDA	Stochastic Gradient Descent Classifier
One Hot Encoder, PCA, LDA	K-Nearest Neighbors Classifier

PCA = Principal Component Analysis

LDA = Linear Discriminant Analysis

For rows that have both PCA and LDA, note that they were implemented separately

Best Model

1. Accuracy of best model: 0.48 score on Testing Set
2. Best Model: Support Vector Machine
3. Parameters of the Best Model: Kernel - Gaussian Radial Basis Function, $C = 0.01$, $\Gamma = 1.0$

Classification Metrics for Best Model

Accuracy: 0.32

Micro Precision: 0.41

Micro Recall: 0.41

Micro F1-score: 0.41

Macro Precision: 0.39

Macro Recall: 0.29

Macro F1-score: 0.27

Weighted Precision: 0.42

Weighted Recall: 0.41

Weighted F1-score: 0.36

Classification Report

	precision	recall	f1-score	support
Class 1	0.86	0.16	0.27	37
Class 2	0.26	0.09	0.13	67
Class 3	0.42	0.52	0.46	98
Class 4	0.40	0.68	0.50	88
Class 5	0.00	0.00	0.00	12
accuracy			0.41	302
macro avg	0.39	0.29	0.27	302
weighted avg	0.42	0.41	0.36	302

Mapping

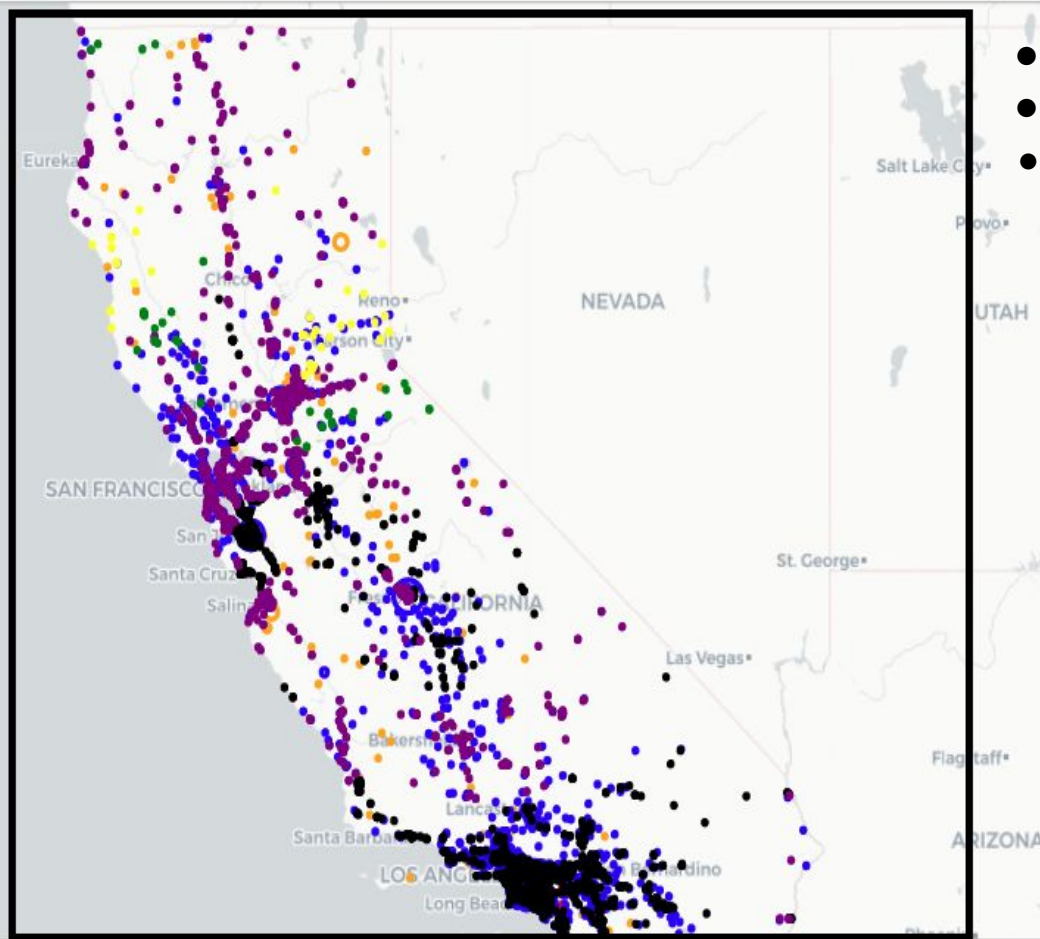
Arnold

METHOD OF APPROACH



- The data is very recent COVID-19, Fire data , Gas Station data circa 2020
- Covid-19 cases were plot alongside Fire cases as the Gas station data

MAP OF CALIFORNIA



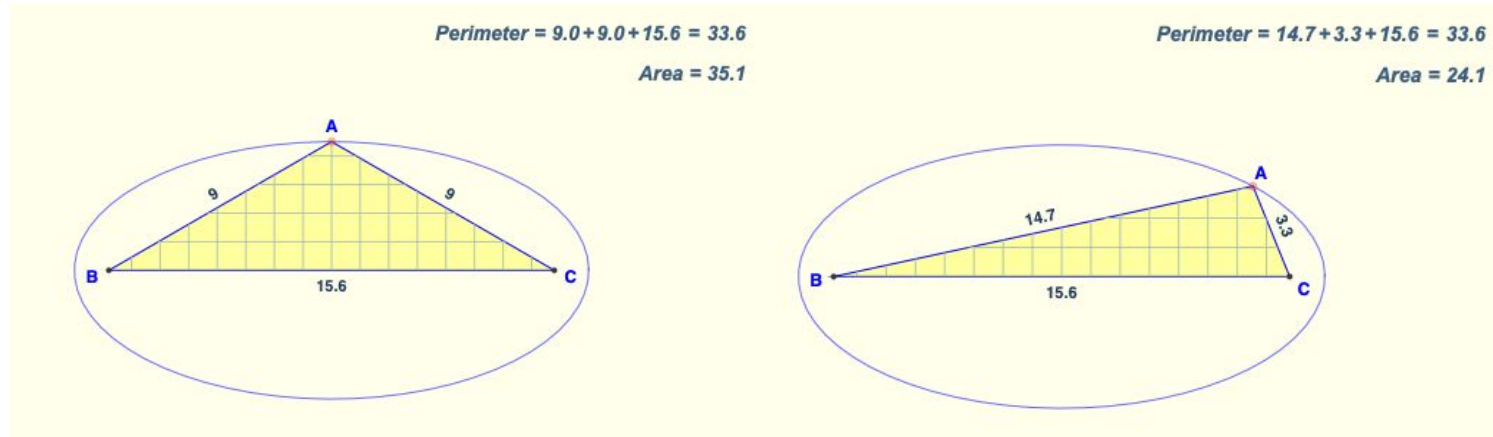
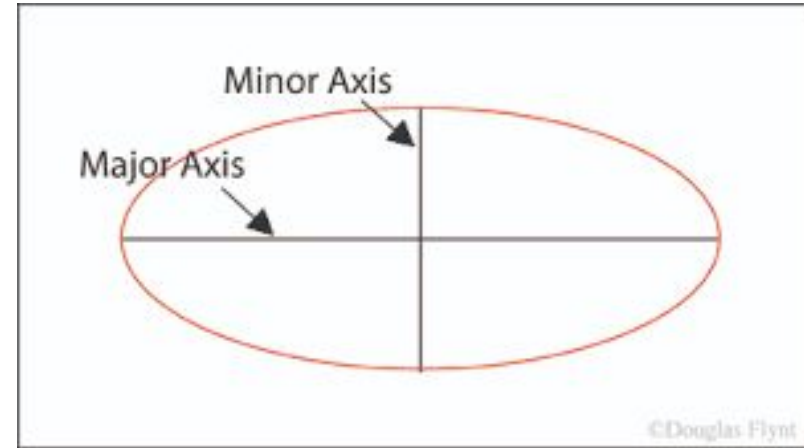
- Static California Map (Prototype)
- Points indicate varying levels of interests
- Color coded according to severity of risk

Mapping: Ellipse

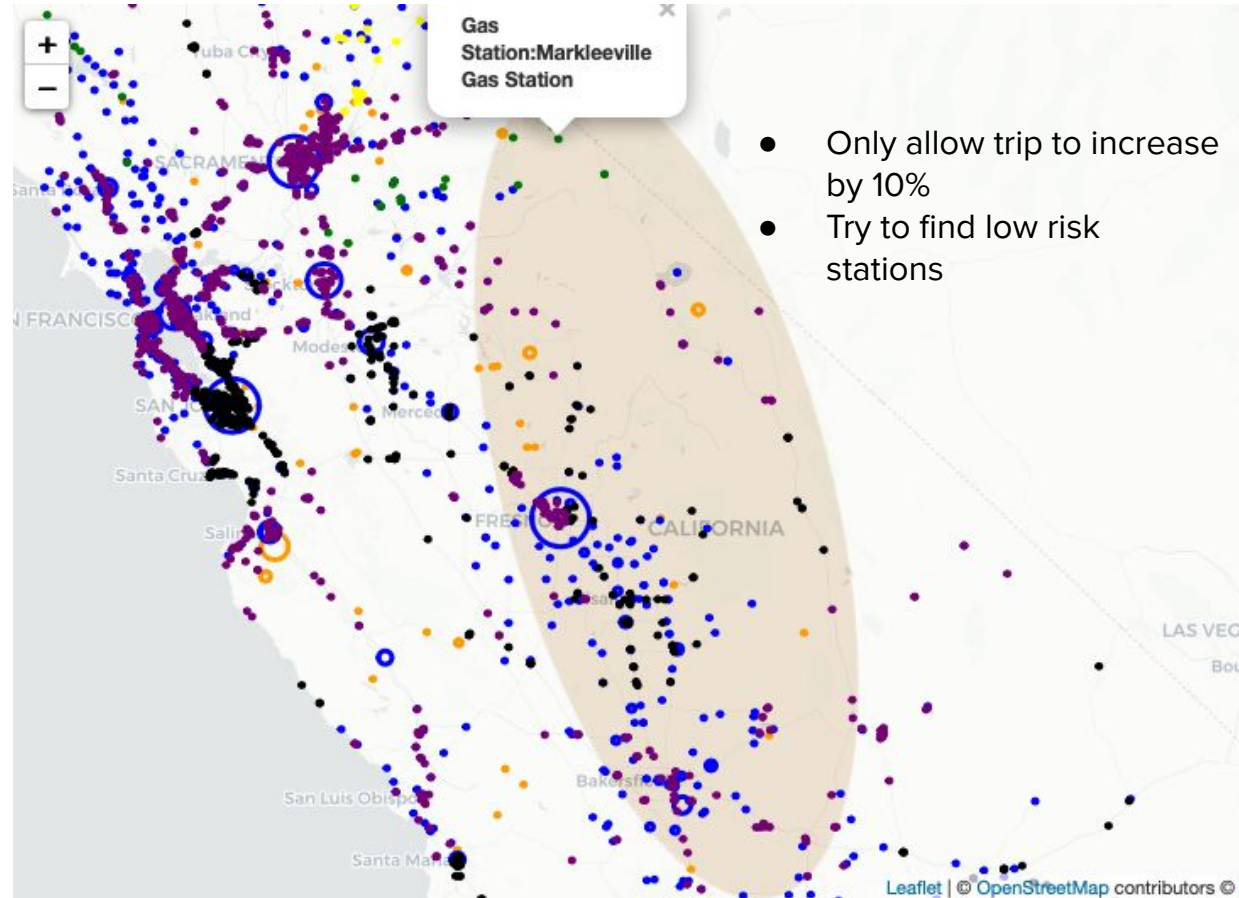
Sophia

Trip Boundaries: Ellipse

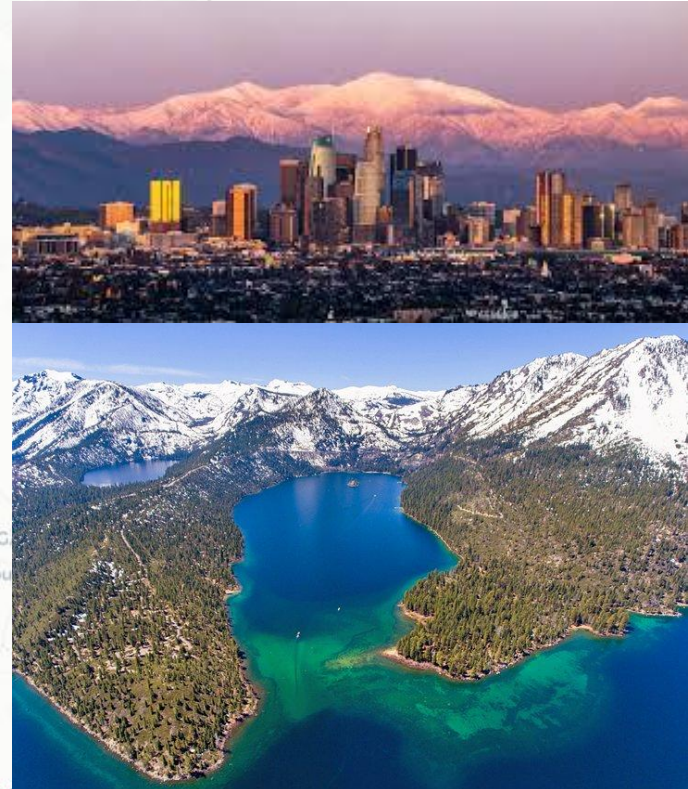
- The 'range' is whatever is in the ellipse shown here, with the focal distance representing the distance from your start and ending points
- In this ellipse, the perimeter represents the points at which your path has the same length.



Which Gas Station?

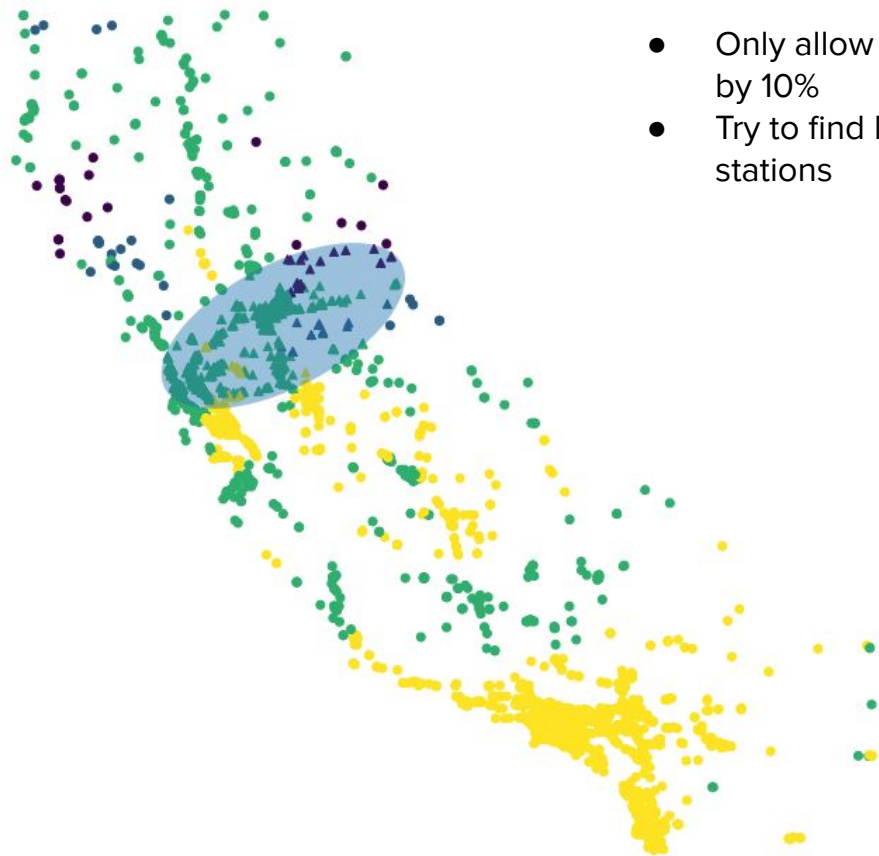


From LA to Tahoe



Easier to see:

- Only allow trip to increase by 10%
- Try to find low risk stations



From SF to Tahoe



Future Directions

1. Map functionality
2. Interactive input and output
3. Live Covid-19 data updates

Citations

- Wildfire data: <https://www.fire.ca.gov/incidents/>
- COVID-19 data: <https://github.com/datadesk/california-coronavirus-data/blob/master/latimes-place-totals.csv>
- Map data copyrighted OpenStreetMap contributors and available from <https://www.openstreetmap.org>
- Folium: <https://python-visualization.github.io/folium/modules.html>
- Image Overlay: <http://qingkaikong.blogspot.com/2016/06/using-folium-5-image-overlay-overlay.html>
- Ellipse Images: <https://www.mathopenref.com/ellipseaxes.html>
- Ellipse animation: <https://nickcharlton.net/posts/drawing-animating-shapes-matplotlib.html>