

On The Road Again

*Using Deep Learning and Regional Analysis to
Remap Touring Entertainment after COVID-19*

Nick Wilders

Metis | December 8th, 2020

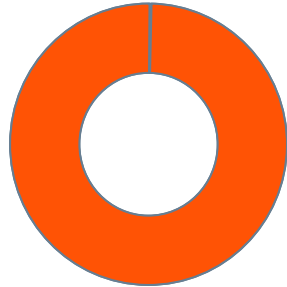


WHERE WE ARE

Dozens of industries have been devastated by COVID-19, but ***none*** have been quite as dramatically muted as the ***performing arts***, particularly ***national tours***.

\$1.6 billion
in revenue (B-way nat'l
tours)

(2018–2019)



WHERE WE ARE

Dozens of industries have been devastated by COVID-19, but ***none*** have been quite as dramatically muted as the ***performing arts***, particularly ***national tours***.

\$1.6 billion

in revenue (B-way nat'l tours)

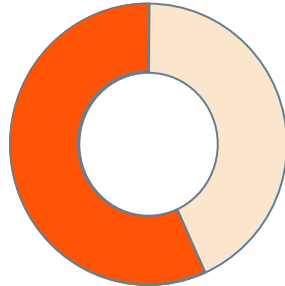
(2018-2019)

**~ 43%
decrease**

~ \$900 million

in revenue (B-way nat'l tours)

(2019-2020)



WHERE WE ARE

Dozens of industries have been devastated by COVID-19, but ***none*** have been quite as dramatically muted as the ***performing arts, particularly national tours.***

\$1.6 billion

in revenue (B-way nat'l tours)

(2018-2019)

**~ 43%
decrease**

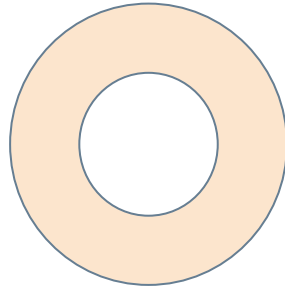
in revenue (B-way nat'l tours)

(2019-2020)

\$0

in revenue (B-way nat'l tours)

(2020-2021)




WHERE WE'RE HEADING

- ***Long Short-Term Memory RNN*** model trained on previous national tour routes to auto-generate new routes across America




WHERE WE'RE HEADING

- **Long Short-Term Memory RNN** model trained on previous national tour routes to auto-generate new routes across America
- Incorporate predictions into **human decision**:
 - **COVID-19 sentiment** in a geographic region based on Twitter sentiment from 33,000 Tweets, produced through NLP sentiment analysis
 -  **Google Trends** search term data (560k frequencies) for relevant terms



WHERE WE'RE HEADING

- **Long Short-Term Memory RNN** model trained on previous national tour routes to auto-generate new routes across America
- Incorporate predictions into **human decision**:
 - **COVID-19 sentiment** in a geographic region based on Twitter sentiment from 33,000 Tweets, produced through NLP sentiment analysis
 -  **Google Trends** search term data (560k frequencies) for relevant terms

TOOLS



Twitter API
for Beginners



tweepy

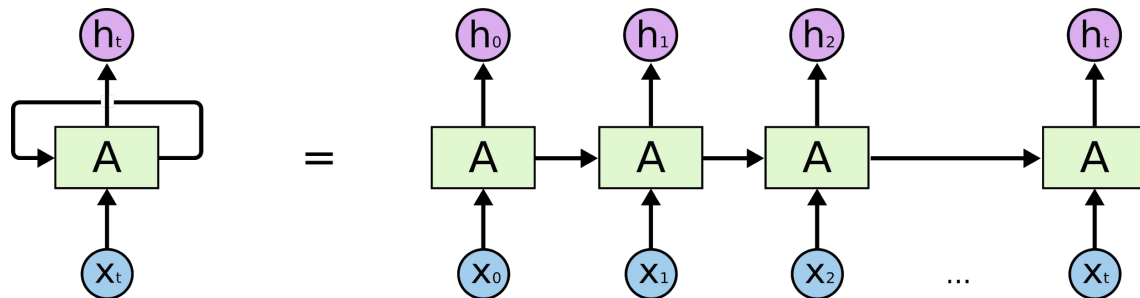
GeoPy



Py trends

LONG SHORT-TERM MEMORY MODEL

- **Recurrent neural networks** employ loops, inputting information as a sequence of numbers
- Long Short-Term Memory models learn **long-term dependencies** and were developed for Natural Language Processing



LONG SHORT-TERM MEMORY MODEL

TRAINING

Enter sequenced coordinates
from **100+ tours** (2003-2020)
for trend analysis



LONG SHORT-TERM MEMORY MODEL

TRAINING

Enter sequenced coordinates
from **100+ tours** (2003-2020)
for trend analysis



TUNING

Adjusting number of
epochs, neurons, and
dense/hidden layers

LONG SHORT-TERM MEMORY MODEL

TRAINING

Enter sequenced coordinates
from **100+ tours** (2003-2020)
for trend analysis



PREDICTING

Use NN to generate coordinates
for next tour stop



TUNING

Adjusting number of
epochs, neurons, and
dense/hidden layers

LONG SHORT-TERM MEMORY MODEL

TRAINING

Enter sequenced coordinates from **100+ tours** (2003-2020) for trend analysis



TUNING

Adjusting number of epochs, neurons, and dense/hidden layers

PREDICTING

Use NN to generate coordinates for next tour stop



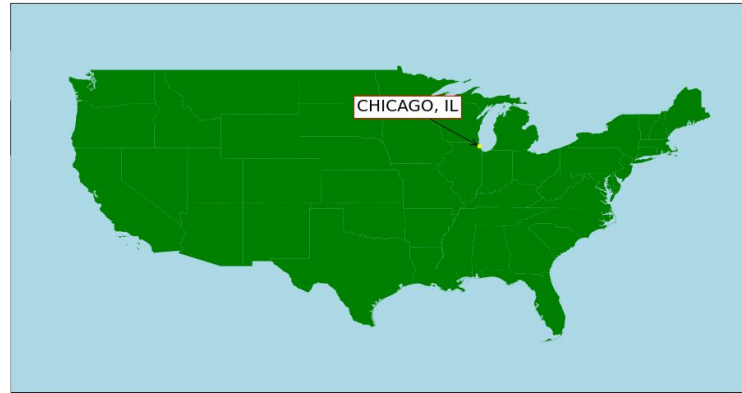
MAPPING

Assign to closest known area, iterate, and observe

AI MAP GENERATOR DEMO

- **Chicago ---> Seattle**
 - With *94 stops* along the way
- Model can choose from **100 most popular cities**, plus a list of specialized target added cities
- Large distances represent **flights** - shorter distance represent **bus trips**

LSTM-GENERATED TOUR ROUTE
Starting Point: CHICAGO



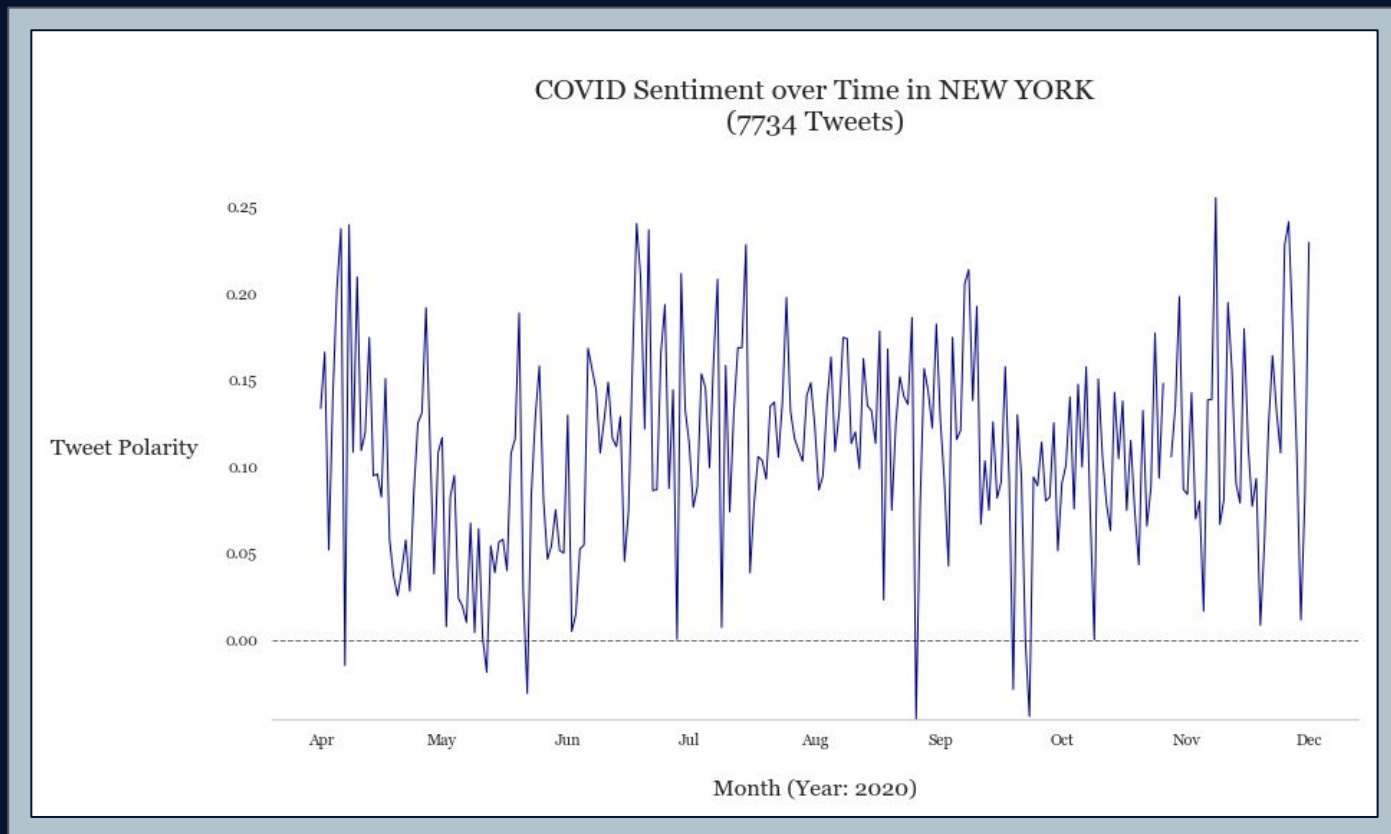


But how do we *improve* these routes?

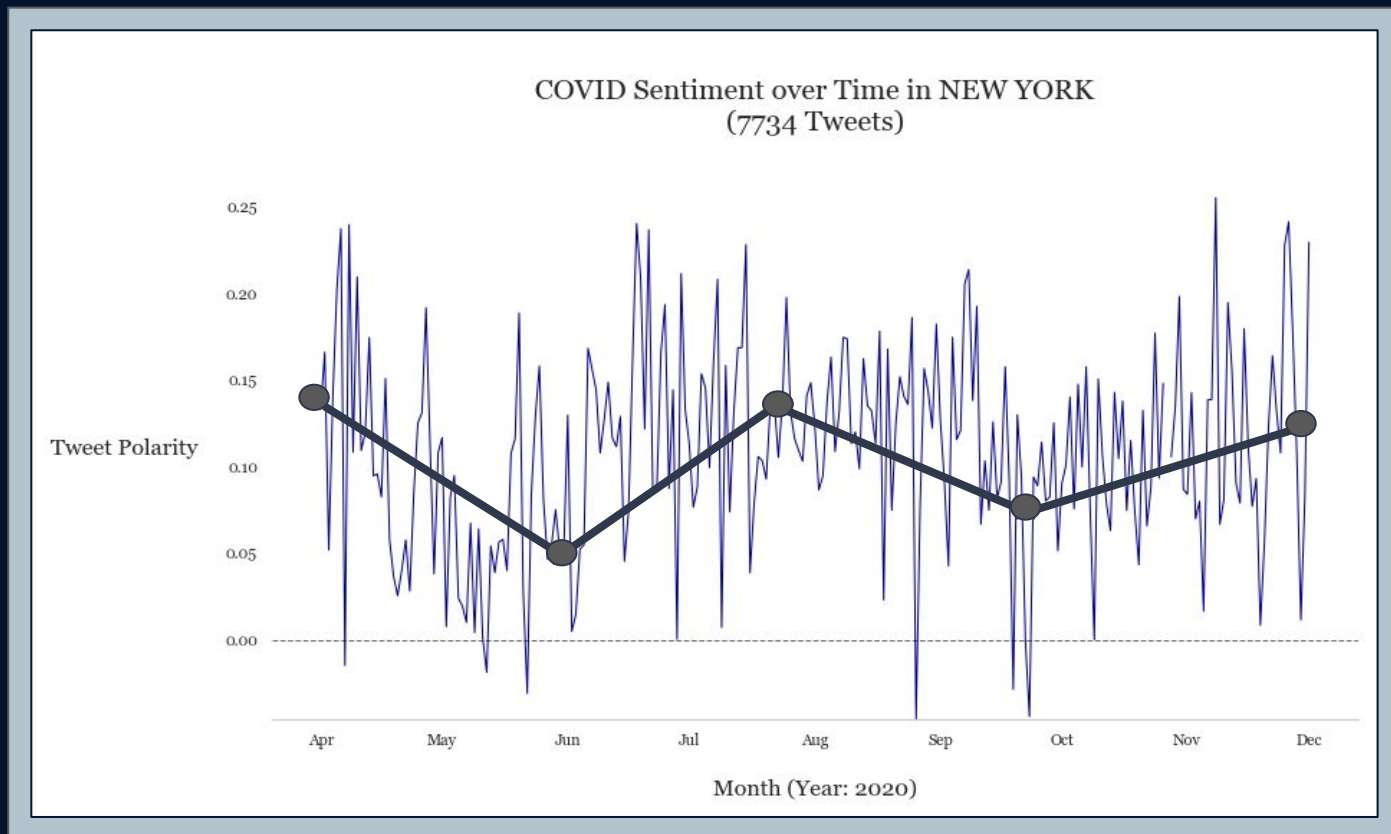


But how do we *improve* these
routes? ...*human guidance*

COVID-19 Sentiment



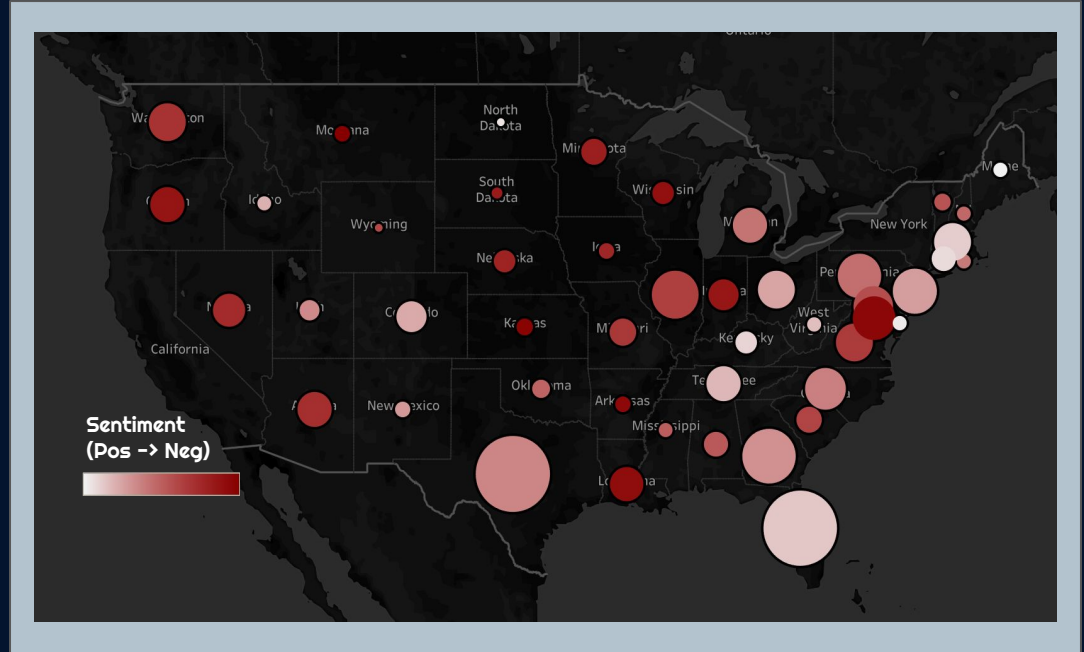
COVID-19 Sentiment



COVID-19 Sentiment

- Collected from **34,000** geotagged U.S. Tweets
- *Collection and sentiment analysis by IEEE Dataport*

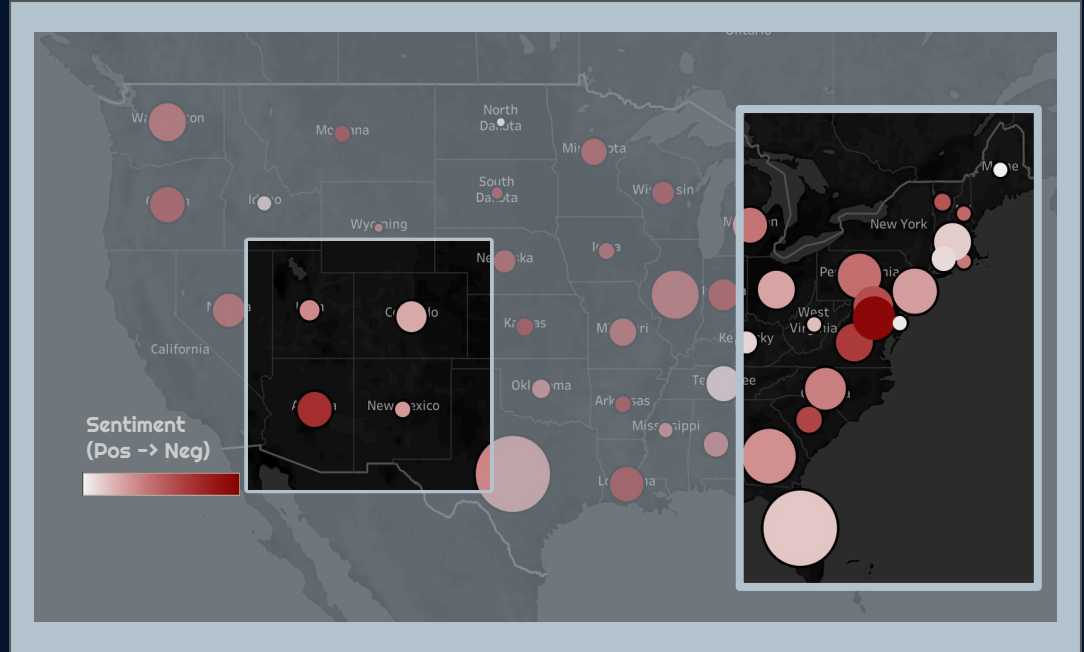
Tweet Polarity across United States



COVID-19 Sentiment

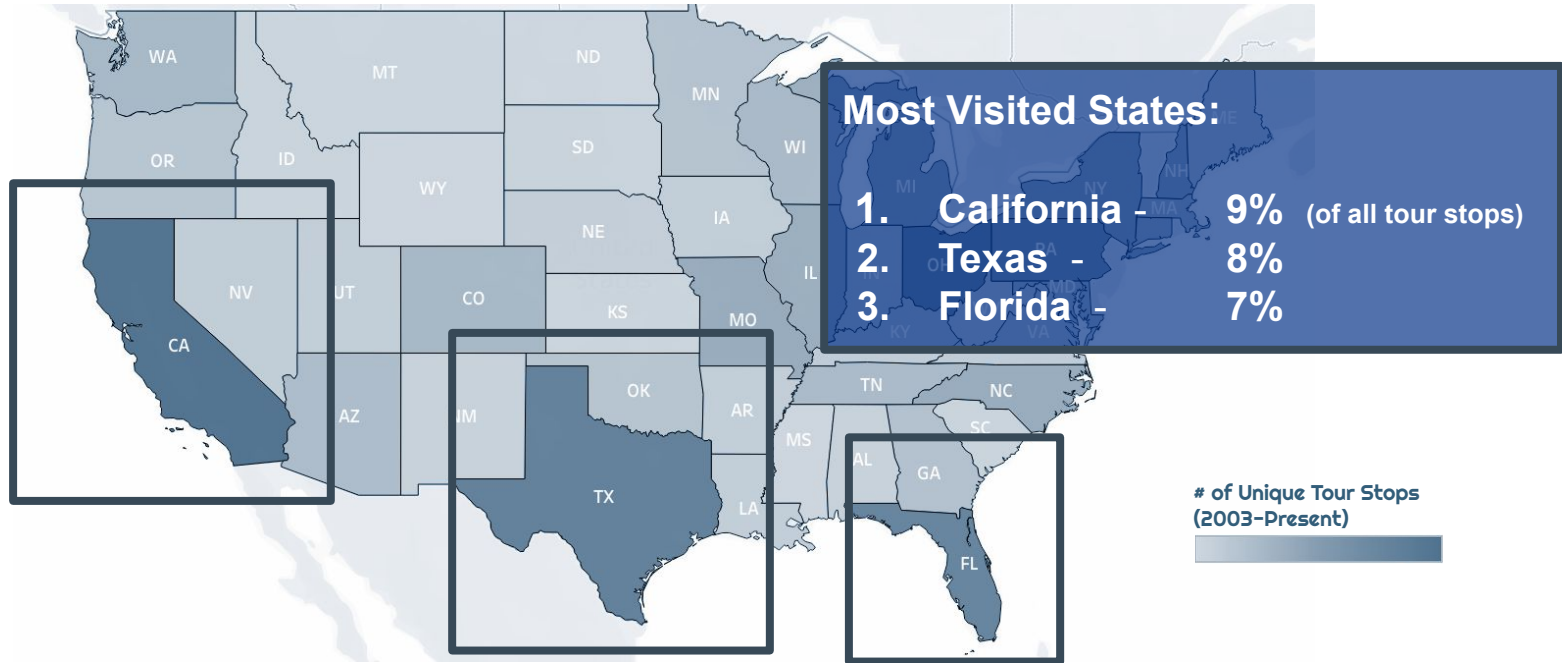
- Collected from **34,000** geotagged U.S. Tweets
 - *Collection and sentiment analysis by IEEE Dataport*
- Higher sentiment Tweets could indicate:
 - Higher “supportive” sentiment for victims
 - Markets willing to gather

Tweet Polarity across United States

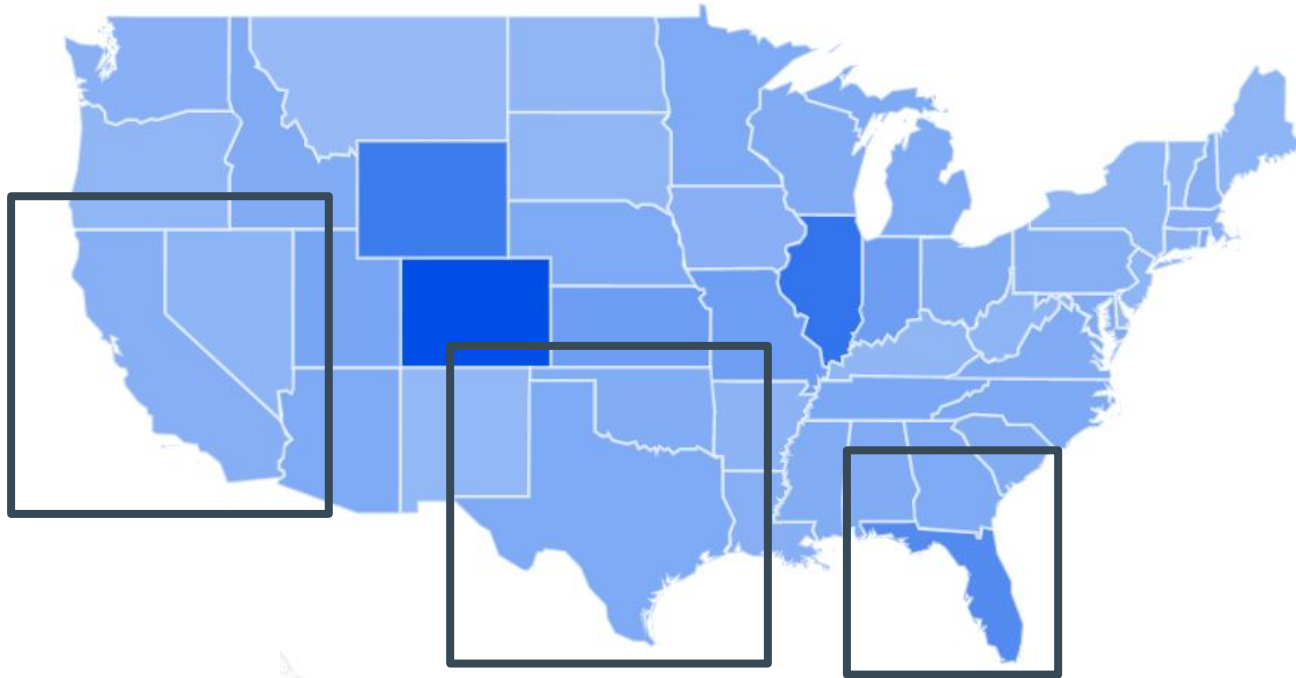


Three Southern states dominate the touring market, accounting for 25% of city visits.

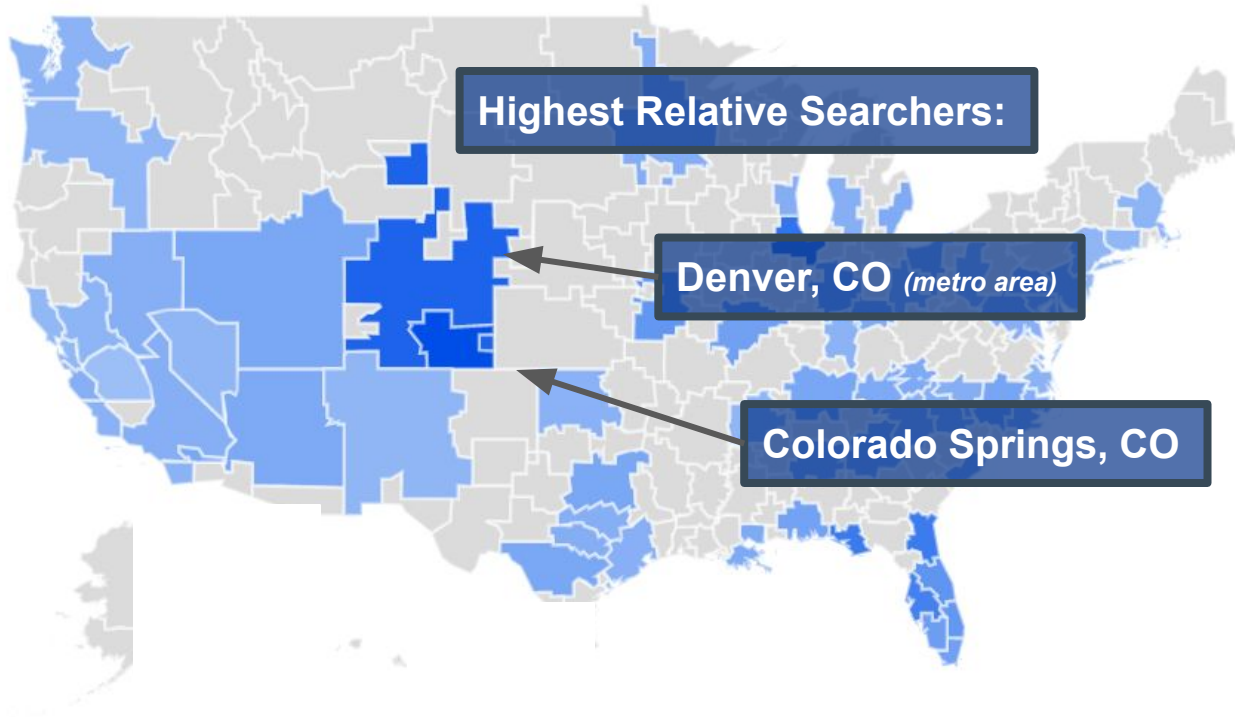
Number of Visiting Tours per State (2003 - Present)



These states are not frequent searches with a Google Trends analysis of “*Theatre Near Me*”...



... but, again, the Four Corners states emerge as an undernourished market, with the highest relative amount of searches for “*Theatre Near Me*”.



Conclusions

FOUR CORNERS = *Optimum Market*

- Given COVID--19 sentiment levels and pre-existing search term interest, **the Four Corners should be prioritized for exposure to new and interested audiences**

Conclusions

FOUR CORNERS = *Optimum Market*

- Given COVID--19 sentiment levels and pre-existing search term interest, **the Four Corners should be prioritized for exposure to new and interested audiences**

YourRoute Tour Planner

- App uses **LSTM model** to make **predictions**, then gives user **three options for their “next stop”**
- ***Check it out!***
 - YourRoutePlanner.herokuapp.com
 - Link in student bio

Conclusions

FOUR CORNERS = *Optimum Market*

- Given COVID--19 sentiment levels and pre-existing search term interest, **the Four Corners should be prioritized for exposure to new and interested audiences**

YourRoute Tour Planner

- App uses **LSTM model** to make **predictions**, then gives user **three options for their “next stop”**
- ***Check it out!***
 - [YourRoutePlanner.herokuapp.com](https://yourrouteplanner.herokuapp.com)
 - Link in student bio

Future Work

FOUR CORNERS Expansion

- **Vibrant arts communities and available venues in**
 - ***Colorado Springs, CO***
 - ***Provo, UT***
 - ***Yuma, AZ***

Conclusions

FOUR CORNERS = *Optimum Market*

- Given COVID--19 sentiment levels and pre-existing search term interest, **the Four Corners should be prioritized for exposure to new and interested audiences**

YourRoute Tour Planner

- App uses **LSTM model** to make **predictions**, then gives user **three options for their “next stop”**
- ***Check it out!***
 - [YourRoutePlanner.herokuapp.com](https://yourrouteplanner.herokuapp.com)
 - Link in student bio

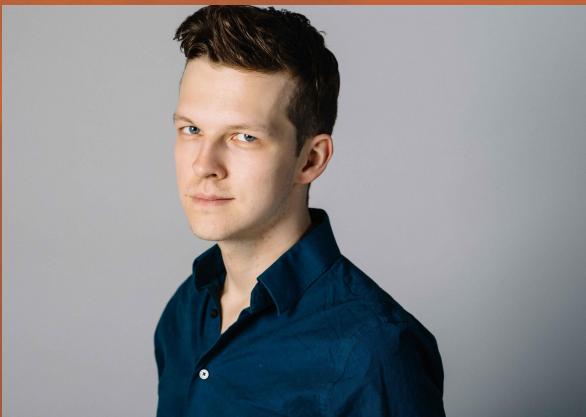
Future Work

FOUR CORNERS Expansion

- **Vibrant arts communities** and **available venues** in
 - ***Colorado Springs, CO***
 - ***Provo, UT***
 - ***Yuma, AZ***

COVID-19

- Investigate regionalized COVID-19 Twitter data and connect to case rates (inspired by **Google Flu Trends**)



Nick Wilders

Data Scientist. Musician. Problem Solver.
New York, NY

nawilders@gmail.com
717-609-2637



[https://www.linkedin.com/in/
nick-wilders-7a7555b/](https://www.linkedin.com/in/nick-wilders-7a7555b/)



[*nickwilders*.medium.com](https://nickwilders.medium.com)



[*nickwilders*](https://github.com/nickwilders)



www.nickwilders.com

CREDITS: This presentation template was created by Slidesgo, including icons by Flaticon, infographics & images by Freepik

DATA VISUALIZATIONS created using Tableau Desktop and Google Trends



APPENDIX



APPENDIX A

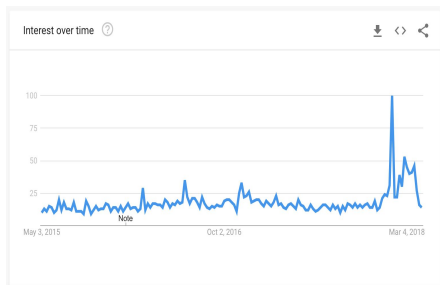
Search Term Engagement Prediction

Project workflow was web scraping intensive, working with high volumes of Google and Twitter data

1

WEB SCRAPE

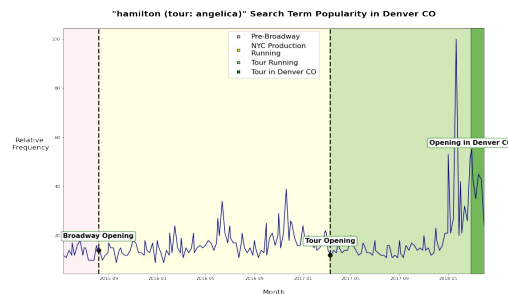
Capture **560k** values for Google search terms relative to geographic region + **300k** COVID-related Tweets



2

DATA VIZ

Recontextualize data and what contributes to maximum engagement spike

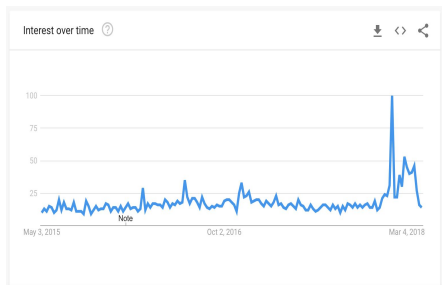


Project workflow was web scraping intensive, working with high volumes of Google and Twitter data

1

WEB SCRAPE

Capture **560k** values for Google search terms relative to geographic region + **300k** COVID-related Tweets



3

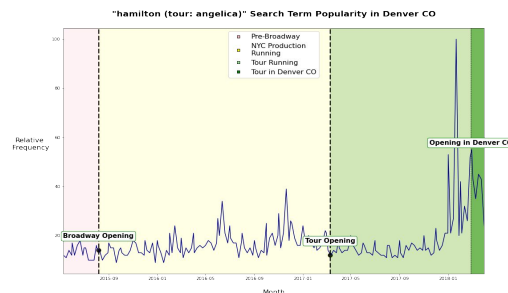
MODEL DEVELOPMENT

Create **Random Forest** model to predict engagement + **LTSM Time Series** to predict COVID sentiment

2

DATA VIZ

Recontextualize data and what contributes to maximum engagement spike

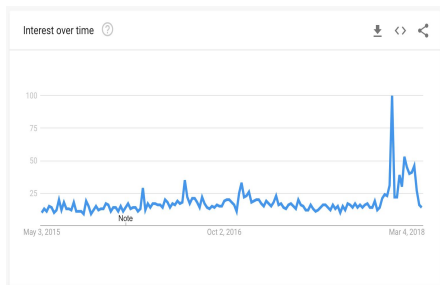


Project workflow was web scraping intensive, working with high volumes of Google and Twitter data

1

WEB SCRAPE

Capture **560k** values for Google search terms relative to geographic region + **300k** COVID-related Tweets



3

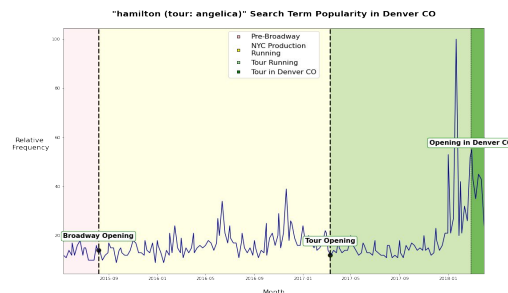
MODEL DEVELOPMENT

Create **Random Forest** model to predict engagement + **LTSM Time Series** to predict COVID sentiment

2

DATA VIZ

Recontextualize data and what contributes to maximum engagement spike

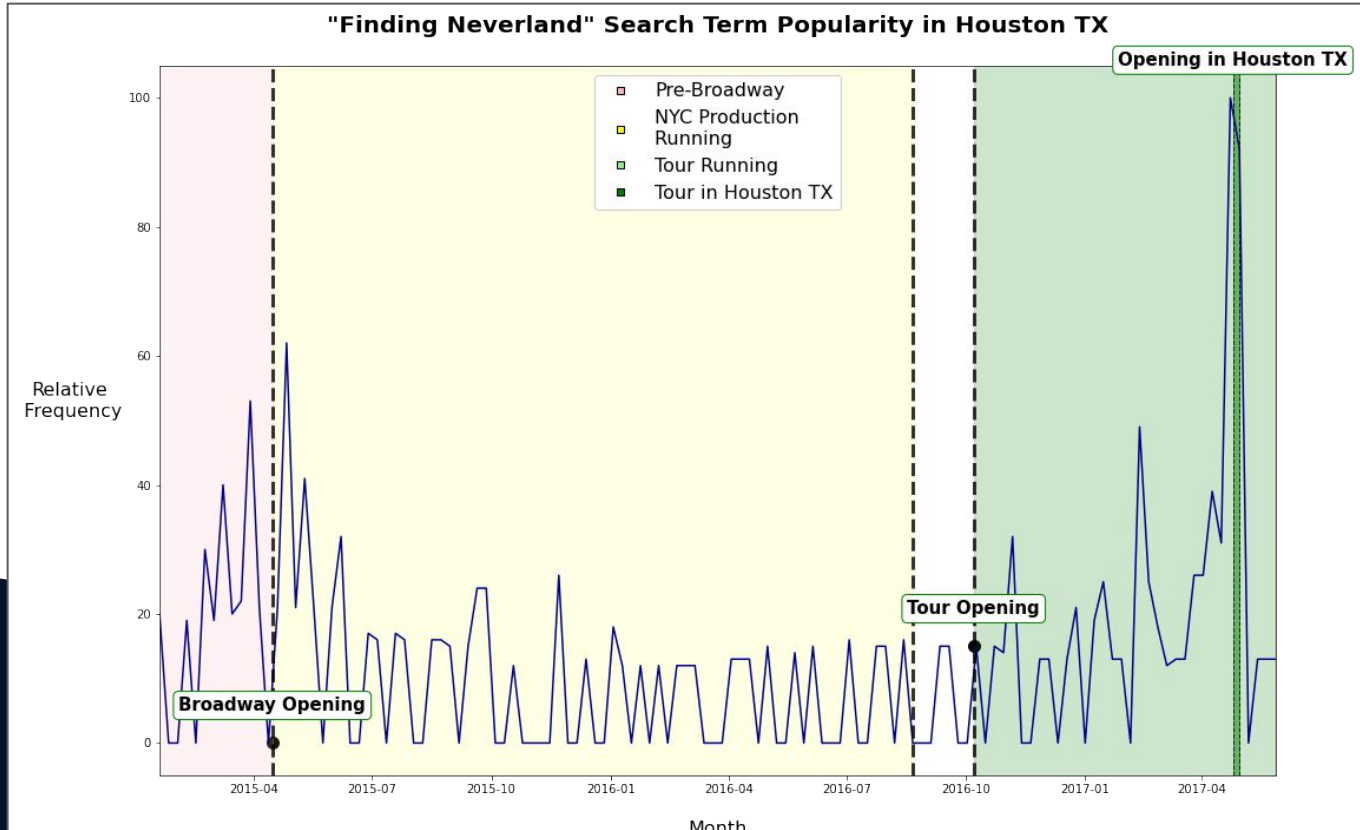


4

SUGGEST CITY ROUTE

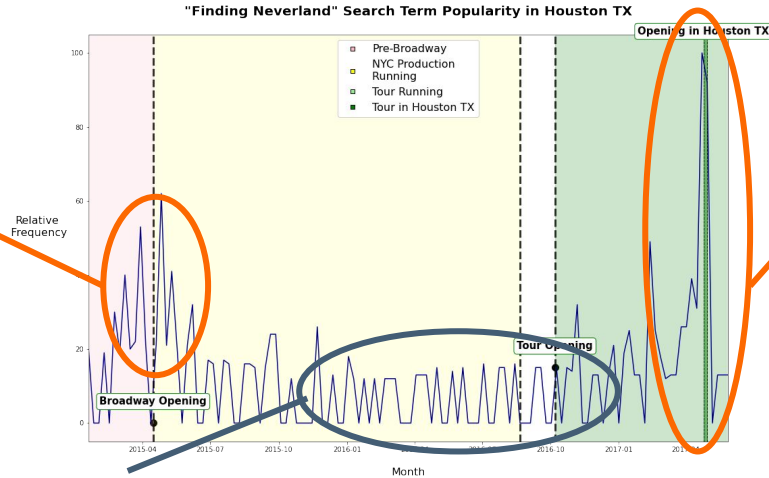
Find optimized route between stops to reach full potential audience

Visualization of a show's "life" before it visits a specific city was necessary to identify key trends in the data.



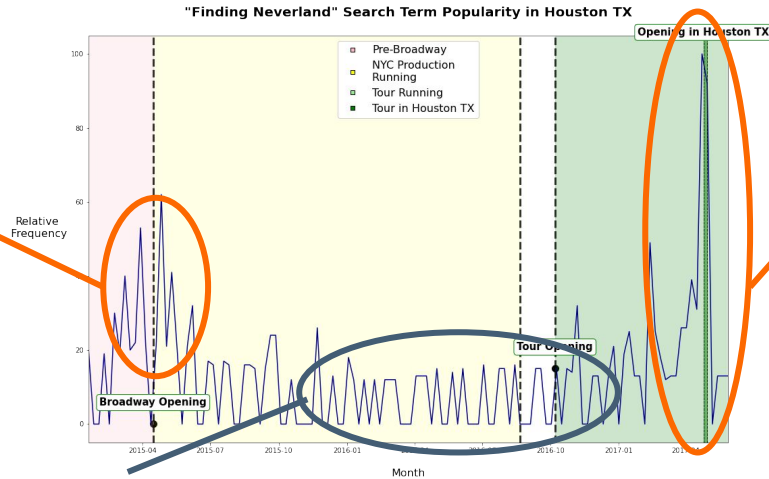
Visualization of a show's "life" before it visits a specific city was necessary to identify key trends in the data.

Engagement spikes surrounding a "pre-show" event (show opening in NYC, announcement that show is coming to specific city) are promising results for a strong spike when show is in-town.



Visualization of a show's "life" before it visits a specific city was necessary to identify key trends in the data.

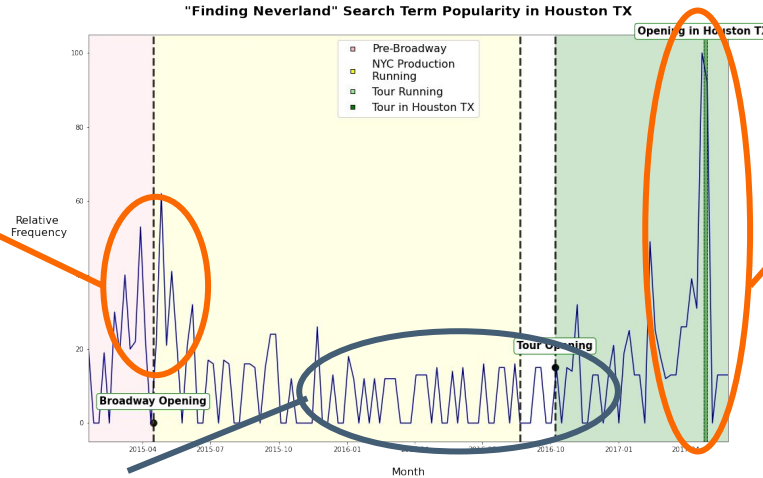
Engagement spikes surrounding a "pre-show" event (show opening in NYC, announcement that show is coming to specific city) are promising results for a strong spike when show is in-town.



Biggest predictor of steady spike is **static engagement** - the engagement after the "pre-show event" - particularly the *range of searches*

Visualization of a show's "life" before it visits a specific city was necessary to identify key trends in the data.

Engagement spikes surrounding a "pre-show" event (show opening in NYC, announcement that show is coming to specific city) are promising results for a strong spike when show is in-town.



In-Town Engagement indicates maximum engagement with the title's intellectual property and time-sensitive interest

Biggest predictor of steady spike is **static engagement** - the engagement after the "pre-show event" - particularly the *range of searches*



APPENDIX B

Search Term Engagement for Show Titles

...but undersaturated markets like the **Four Corners** showing consistent search term engagement, despite having only five venues combined.

