Finding Coronavirus Hotspots in Texas

...

Introduction/Problem

- Coronavirus is a global pandemic that has forced millions of people into quarantine and it is vital to stop the spread of the disease
- The general public as well as lawmakers would benefit from this study as it provides valuable insights into causes of spread
- Analyzing the types of venues present in different areas and exploring how they relate to infection rates can help:
    - Provide more insight into causes of spread of disease
    - Help to develop preventative measures
    - Give the general public more guidance on places to avoid

Data

- Foursquare API - Location data used to search for popular venues in every county

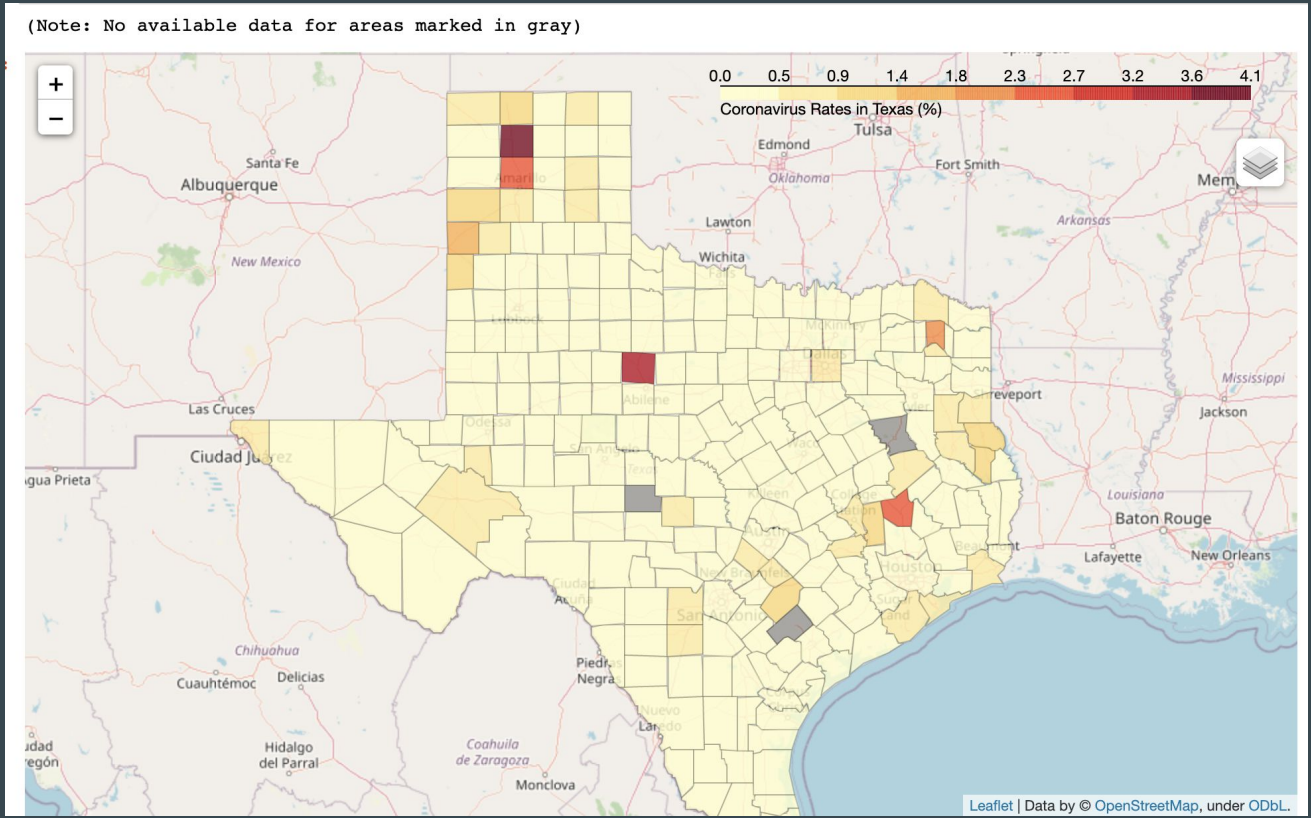
- Texas Coronavirus Cases - dataset updated every day that provides current case totals for every county (from *dshs.texas.gov*)

- Texas Boundaries and Centroid Data - provides location data for visualization as well as feeding locations into the Foursquare API (from *data.texas.gov* and gis-txdot.opendata.arcgis.com)

Methodology

- Infection rates calculated using the coronavirus case data and visualized
- Popular venues found using the Foursquare API for each county and were ranked based on frequency
- K-means clustering was done based on infection rates and popular venues
- Venues were ranked based on a weighted score function considering both infection rates and popularity of venues

Choropleth map of infection rates of Texas Counties

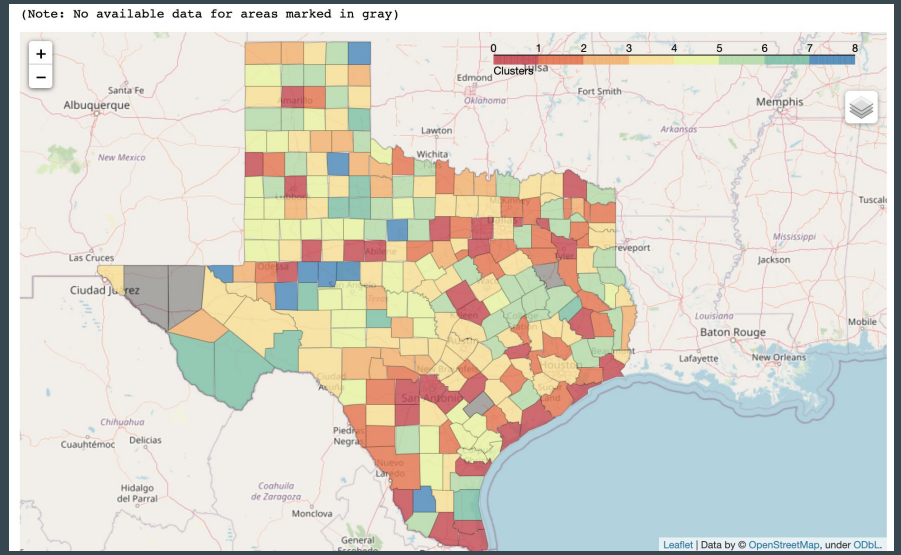


Ranking venues for each county

	County	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue
0	Anderson	Pizza Place	Mexican Restaurant	Burger Joint	Clothing Store	BBQ Joint	Discount Store
1	Andrews	Convenience Store	American Restaurant	Deli / Bodega	Fast Food Restaurant	Grocery Store	Coffee Shop
2	Angelina	Discount Store	Fast Food Restaurant	Sandwich Place	Mexican Restaurant	Gas Station	Burger Joint
3	Aransas	Convenience Store	Hotel	Seafood Restaurant	Harbor / Marina	Park	Bar
4	Archer	Ice Cream Shop	Post Office	Spa	Other Repair Shop	Outdoors & Recreation	BBQ Joint

- Using the Foursquare API, venue categories were ranked for each county based on frequency/popularity

Clustering



- Venues were clustered into 8 groups using venue data combined with infection rate data for each county

Cluster Labels	Infection Rate	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue
0	0.408361	Mexican Restaurant	Burger Joint	Convenience Store	American Restaurant	Fast Food Restaurant	Grocery Store	Pizza Place	Coffee Shop
1	0.204659	Mexican Restaurant	Fast Food Restaurant	American Restaurant	Burger Joint	Sandwich Place	Ice Cream Shop	American Restaurant	Grocery Store
2	0.286926	American Restaurant	Mexican Restaurant	Ice Cream Shop	Campground	Discount Store	Zoo	Fish & Chips Shop	Factory
3	0.274147	Coffee Shop	Mexican Restaurant	American Restaurant	Fast Food Restaurant	Pizza Place	Sandwich Place	Discount Store	Fast Food Restaurant
4	0.185823	Convenience Store	Discount Store	Ice Cream Shop	Sandwich Place	Discount Store	Factory	Falafel Restaurant	Farm
5	0.471894	Fast Food Restaurant	Discount Store	Sandwich Place	Pizza Place	Discount Store	Sandwich Place	Hotel	Mexican Restaurant
6	0.073975	Ice Cream Shop	Home Service	Zoo	Zoo	Factory	Falafel Restaurant	Farm	Farmers Market
7	0.117144	Convenience Store	Zoo	Zoo	Fishing Spot	Falafel Restaurant	Farm	Farmers Market	Fast Food Restaurant

Ranking Venues

	Average Weighted Score
Chinese Restaurant	7.391056
Video Store	5.485156
Discount Store	5.459786
Hotel	5.405881
Mexican Restaurant	5.269406
Airport	5.175274
Convenience Store	5.147593
Fast Food Restaurant	4.999459
Department Store	4.976926
Pizza Place	4.972640

- Weighted score calculated based on infection rates for each county combined with how popular a type of venue
- Venues were ranked based on score (higher score means higher contribution to infection rate)

Results

- Clustering - Clusters with higher infection rates tended to have more venues such as restaurants and convenience stores, while clusters with lower rates had more venues such as farms and zoos

- Ranking - The top ten contributors calculated were mostly venues such as restaurants and stores, which supports the cluster observations

Discussion

- External Factors - Many external factors could also play into determining infection rates such as testing rates and income level which merits further research

- Implications - Though not comprehensive, the results of this study suggest that more urbanized areas containing fast food restaurants and convenience stores are more susceptible to the spread of disease than rural areas where farms are more common

Conclusion

This study gives a good insight into the types of venues that could possibly be dangerous for the average person to go to when trying to stay safe from the coronavirus. Though not a comprehensive result, there is evidence to suggest that more urbanized areas may be more susceptible to disease, and thus people may want to avoid venues like restaurants and convenience stores.

This project's results could be further validated or built upon by considering regions where reopening policies have not been implemented and comparing the results, or by including more regions on a national level and considering more factors than just type of venue.