Slide 1:

Hello, I'm Mason Jerome, and I'm here with my partner, Mayur Hooli to talk about our paper, COVID-19 Management: preparation and response. This project was done during the early stages of the virus and focuses on categorizing different areas based upon the severity of the virus in the state of Texas in the United States.

Slide 2:

Earlier this year, COVID-19 has turned into the deadliest pandemic in over 100 years ever since the Spanish Flu of 1918. With a shortage of medical supplies, these supplies need to be put to the best use possible. To do this, Suppliers and policymakers need to know how to best allocate these resources in order to lessen the spread of the virus through lockdowns or other measures of lesser magnitude. For instance, a mandatory public mask wearing policy or limiting the capacity that certain businesses can entertain at once. To help instruct those in charge of making these crucial decisions, we came up with the COVID-19 Management System.

Slide 3:

So, what does this management system do? The COVID-19 Management System is an intelligent system that identifies the critical areas where certain actions need to be taken. For this, we examined Texas in the United States. Not all areas in the state will necessarily need the same mandates put in place or they may have varying degrees of aid required. To help inform those using this system, the counties around the state will be ranked based upon the risk they pose of spreading the virus to neighboring counties. This clearly shows which affected zones should have focus put upon them to help curb the spread of the virus.

Slide 4:

How is this done? The idea implements clustering alongside the PageRank algorithm. For the clustering, we identify 4 distinct clusters based upon the dataset that we prepared. The K-means algorithm is implemented to achieve the clusters where centroids are placed among the data points that group and associate the counties with their respective clusters for classification. For the PageRank, we first determined a value to associate with each county by applying a custom equation to dataset using several parameters before normalizing the values. We do this to get values for all of the counties from the highest risk one all the way down to the lowest risk county. With this value, PageRank then can determine how the viral spread of each county affects other, neighboring counties. This can help inform future decisions and show where efforts to stop the spread of the virus will be most impactful.

Slide 5:

Why is all of this done? When this project was completed, there was a significant shortage of medical gear needed to tackle this virus. Whether it was: N-95 masks or personal protective equipment (PPE) to help protect the doctors and nurses from getting infected by the virus from the patients they were treating, testing kits to get accurate data on the number of cases, or ventilators for those with severe cases of the virus, all of these were lacking. Certain hospitals were also inundated with patients, making it so not all of them could be effectively treated. This lead to drastic measures in those highly infected areas like full lockdowns of the affected cities.

The categorization of counties helps with the effective allocation of resources. With this in mind, the prediction of future affected counties would have helped in better logistical preparation to mitigate the virus which could have reduced the overall number of cases and deaths.

Slide 6:

Here, you can see the general diagram of the code used to create the COVID-19 Management System. We started with data collection, compiling multiple datasets to come up with the parameters that we used in our classification. There are some examples up here such as the total population in the county, the population density, health index. The travel index which we came up with by calculating the correlation coefficients between passports and median income, economic output, and degree of education. Other parameters include the number of cases and deaths, and whether a stay-at-home order is in place or not. With these parameters collected, we could then cluster the counties together to come up with four separate classes. They ranged from counties with nearly no risk of people there contracting the COVID-19 virus to dangerous counties where the risk is very high in comparison. We then used the PageRank algorithm determine the inter-county spread of the virus in the immediate future based upon the current situation.

Slide 7:

Here are some the results that we ended up with. You can see a sampling of some of the counties in each labeled classification, which ranges from the safest county in Texas, Wheeler county which has very few confirmed cases of COVID-19, with few people traveling to and from this county, along with a very low population density of 3 people per square mile, making contact with other people less likely. The PageRank algorithm took into account the danger presented to Wheeler county by its neighboring counties as well. On the other end of the spectrum there is the most dangerous county, Dallas county, which has a high population with a high population density. It is also surrounded by high population counties with many cases in them. Also, 2 major airports along with another small one within the county lead to a high degree of travel, and the number of cases within the county is already high, so there is a high degree of danger.

Slide 8

With everyone in the world interested methods of fighting the virus, it is important to propose possible solutions. And we took a novel approach of evaluating the virus by combining clustering and PageRank together to help us better understand where the risk lies. We hope that following the results of this system would help us make informed decisions to fight against the virus and save lives.

Slide 9:

We had some ideas for the future work of this project to help make it better. Right now, this project only looked at Texas counties, making it only useful in one state in the U.S. at the moment. Upscaling this COVID-19 Management System would be something that we would be interested in working on to include the entire United States or even other countries. We could do this by simply collecting data from all counties and accounting for state-to-state spread or

even country-to-country spread. Another way we could improve the project would be using the Geographic Information System framework for better visualization of the project. An obvious improvement would be collecting more parameters to identify the spread of the virus. I'm sure we could find some relevant parameters that could be added by combing over county datasets. With COVID-19 vaccines on the horizon, the end of this pandemic is hopefully insight. So, creating similar models for future pandemic and disaster responses to help manage similar situations in the future would be a good way to pivot this system and make it applicable for other work. And, in general we hope that something like this system can help keep people healthy.