Abstract Section:

Currently, there are 26.2 million COVID-19 cases in the US due to people taking lower precautions to reduce transmission at public venues. This project aims to create a tool that provides individuals with information to make a responsible decision about visiting an establishment to prevent unnecessary SARS-CoV-2 cases. A probabilistic model that predicts the risk of an individual contracting COVID-19 (transmission risk) during a visit to an establishment was created. This model was integrated into a web app built using the StreamLit framework in Python3.8. It acquired data through parsing COVID-19 databases and obtaining venue-specific data through user-inputs. The transmission risk model worked by multiplying the likelihood an individual will encounter a virus carrier (exposure risk) and the risk the individual will contract the virus if one carrier is present (contract risk). The exposure risk model was tested through the Spearman's Rank Correlation by calculating the correlation between the model's risk prediction for 200 random counties to new cases two weeks later of the aforementioned counties. Contract risk was tested using four different scenarios. Transmission risk was tested using these four scenarios factoring three counties with ranging incidence rates. The exposure risk model averaged a Spearman Rank correlation of 0.81, placing it in the "very strong" category. The contract and transmission risk provided sensible predictions for the scenarios provided. This model can be easily expanded to other databases and adapted to high-incidence countries. Since virus-specific aspects can apply to other illnesses, the model can be adjusted easily for other viruses.

Introduction:

A pandemic is an international disease outbreak that affects millions of people over a vast geographic area. In modern society, extensive global travel and trade can exacerbate the effects of a pandemic. The severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), also known as coronavirus disease of 2019 (COVID-19), originated from a patient from Wuhan, China, in December of 2019. The virus then spread rapidly around the globe, eventually reaching the United States in January. Although cases in the United States were stagnant for two months, in mid-March of 2020, cases increased swiftly at densely populated localities along the east and west coasts of the United States. Due to effective social distancing measures over five months, these areas had a comparatively low rise in new cases than the reopening months. Assuming the peak had already passed, the United States took on the daunting challenge of reopening the country and transitioning back to “normal’ life in the autumn of 2020. Cases began to sky-rocket after many people started attending public establishments, such as malls, grocery stores, and recreational events.

According to public health officials, the reopening of socially interactive venues, such as bars and theaters, and relaxation of social distancing protocol, allowing more people to attend a venue were significant factors for the spike seen (Gamio, 2020). The rise in cases leads to a multitude of detrimental effects for the United States and its denizens. Most notably, there is a clear correlation between COVID-19 cases and deaths in a community. With a mortality rate of about 70.07 to 100,000 people, the death toll has risen to 417,000 people as of January 2021 (Mortality Analyses, 2020). Deaths are sobering for the entire country and affect minorities, elders, and people with compromised immune systems to worst (Cooper & Williams, 2020). On top of that, a rise in cases leads to increased hospitalizations, straining the community’s health care infrastructure and hindering them from treating COVID-19 patients and others in need of medical assistance. Growth in cases affects the economy on a national and personal level due to travel and import restrictions implemented to reduce transmission of the virus (Oh, 2020). The low economy led to a recession and a loss of over 30 million jobs (insert source). Overall, the pandemic has disrupted the lives of many individuals medically, socially, and financially. Hence, people must be cautious and make the right decision when deciding to visit an establishment.

There are currently two types of public tools to make decisions: COVID-19 dashboards and risk estimator tools. Like the Johns Hopkins COVID-19 dashboards, dashboards are great sources to visualize trends in various geographical areas. They allow users to understand the pandemic's current state by viewing how incidence has changed over time. This type of tool is also accurate since they source their data from trusted databases, such as the (insert a formal name for JH database) and New York Times. Dashboards simply provide users with general data without analysis. They are not best suited to provide users with a risk associated with visiting a venue. Risk estimator tools, such as the *Community Social Risk Estimator (CSORE)* and *COVID-19 Event Risk Assessment Planning Tool,* are web apps that use data-driven probabilistic models, COVID-19 incidence data, and venue population to provide users with the probability that they will encounter a carrier during their visit the establishment. That likelihood is known as exposure risk, and it allows individuals to make a more informed decision through a direct metric. However, they are not entirely accurate since they do not account for the virus's transmissivity and the population density inside the establishment.

To eliminate unnecessary cases during the country's reopening period, individuals need to make decisions about visiting an establishment regarding the risk it poses to themselves. Many tools provide people with an estimated risk; however, they do not include the transmissive properties of COVID-19, which is an essential aspect when predicting risk. This project aims to implement a web app that utilizes a data-driven transmission risk model to give users a better understanding of the risk associated with visiting an establishment. With this knowledge, individuals in every community can take the right steps towards lowering COVID-19 cases and avoid the detriments of high COVID-19 incidence rates.

Methods and Materials:

A probabilistic model that predicts transmission risk was created and tested on Python3.8 using the PyCharm IDE. Additionally, results from testing were stored on an MS Excel Workbook. Multiple Python Modules were used during the creation and testing of the probabilistic model. The Pandas module was used to parse data from databases. The Datetime module was used to get the current date and travel between different dates. The inbuilt Python Math module was used to perform advanced calculations. During the testing process, the Barnum module was used to create 200 random counties. Matplotlib was used to create graphs based on the data collected. The SciPy module was used to calculate the Spearman's Rank Correlation. The probabilistic model initially obtained data from the NY-Times COVID-19 database for COVID-19 related information but switched to the COVID-19 Data Repository by the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University. County population data was initially acquired from the United States' Census Bureau's 2010 Census but switched United States' Census Bureau's 2019 population predictions. A 2018 US Zip Code to County State to FIPS Look Up data set was used to covert zip codes to FIPS codes.

Transmission Risk Model:

The transmission risk model works using a simple equation that has three parts (Sun et al., 2020):

$$R_{transmission} = \frac{R_{exposure} \times R_{contract}}{I_{immunity}}$$

The model predicts the likelihood that an individual would be infected with Cov-SARs-2 ($R_{transmission}$) when visiting a venue. $R_{exposure}$ is the probability that one or more carriers of the virus would be present in an establishment during the visit. Essentially, $R_{exposure}$ models an individual's exposure to a carrier of the virus in a community. This part of the model is affected by factors such as incidence rates and the number of people at the venue. $R_{contract}$ is the likelihood that an individual contracts the virus after interacting with a carrier. This value will be low the more protection – face mask, air purifiers, etc. – that the individual and carrier use (Sun et al., 2020). Finally, $I_{immunity}$ represent the individual's ability to be immune to a virus. It is known that previously infected individuals develop anti-bodies and provides the individual with a short-term period of immunity; however, estimates for this period vary greatly from one week to six weeks ("Immunity passports" in the context of COVID-19., n.d; Lumley et al., 2020, p. 9). In addition, lower incidence rates among children suggests their immune system offers some immunity (Center for Disease Control and Prevention, 2020). More definitive research is needed in this area which is why $I_{immunity}$ is assumed to be same for every individual in this model. To make this model, $R_{exposure}$ and $R_{contract}$ were created and tested independently and then combined and tested.

Exposure Risk General Model:

The likelihood that one or multiple carriers of the virus are present in a population of *n* people can be modeled using this equation, where *p* is the probability that a randomly selected individual in the venue's population is a carrier and *n* is the venue's population:

$$R_{exposure} = 1 - (1-p)^n$$

The model works by first predicting probability that each of the *n* individuals are not infected by raising $1-p$ to the power of *n* ($(1-p)^n$), which is then subtracted from one to the get the final exposure risk (Chande et al, 2020).

Circulating Case Estimate:

To calculate the probability (*p)* that a randomly selected individual in the venue's population (*n)* is a carrier, the circulating case estimate and the total population of the county is needed because *p* is simply the circulating cases divided by the total population.

The circulating case estimate is calculated by summing all of the new cases from the previous 10 days and then multiplying it by the ascertainment bias. New cases are summed from 10 days prior to follow CDC guidelines of infectiousness (Chande et al., 2020). The new cases are then multiplied by the ascertainment bias to account for unreported cases. An ascertainment bias of 10 was chosen because only 1/10 of cases are reported in reality (1). In the future, ascertainment bias may lower as reporting gets more accurate.

Exposure Risk Protype #1:

$$R_{exposure} = 1 - \frac{(p - Ab * \sum_{\Delta d=0}^{10}[Nc_{d-\Delta d}])_!^{\frac{n}{Ar}}}{p_!^{\frac{n}{Ar}}}$$

Based off of the general model, this was the first model created to predict $R_{exposure}$. This prototype aimed to incorporate circulating case estimate to calculate the cases in a community, which was done through the $Ab * \sum_{\Delta d=0}^{10}[Nc_{d-\Delta d}]$ section of the model, where Ab is the ascertainment bias that is multiplied to sum all of the new cases Nc at day number d for 10 days. The new cases *c* for each day was obtained from the New York Times COVID-19 Database. Another addition this model made was incorporating population density $\frac{n}{Ar}$, with *n* being the venue's currently population and *Ar* being the venue's area in meters squared. *p* is the 2010 census population of county in which the venue is located. Population data was acquired from the US 2010 census.

Exposure Risk Model Prototype #2:

$$R_{exposure} = 1 - \left(\frac{p - Ab * [Tc_d - Tc_{d-10}]}{p}\right)^n$$

In the second iteration of the model, efficiency was prioritized. All unnecessary calculation - factorials and summation of new cases - were eliminated. Population density ($\frac{n}{Ar}$) was implemented incorrectly in the first prototype and thus eliminated as well; it was factored back into the model in the $R_{contract}$ aspect of the model. Since the New York Times COVID-19 data bases also provided total cases (Tc_d) at specific day *d*, total cases data was used instead of new cases data to find the total number of new cases between the current day (*d)* and the day 10 days later (*d-10)*. All other factors are kept the same.

Exposure Risk Model Prototype #3 (Final Model):

$$R_{exposure} = 1 - \left(1 - \frac{Ab * [Tc_d - Tc_{d-10}] * 0.2}{p}\right)^n$$

The third and final iteration of the model is more efficient since the population *p* of the county does not need to be retrieved twice from the data base. Circulating case estimate was updated to factor in a finding that 20% of cases are asymptomatic, meaning that only about 80% of the population quarantines since the other 20% is not aware they are a carrier. This iteration obtained its data from the COVID-19 Data Repository by the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University.

Contract Risk Model General Model:

$$R_{contract} = 1 - (1 - P_{RNA})^{(D_{episode} \times n)}$$

Contract risk model predicts the probability that an individual will contract the virus after interacting with a COVID-19 carrier. Created by Lelieveld, J., Helleis, F., Borrmann, S., Cheng, Y., Drewnick, F., Haug, G., Klimach, T., Sciare, J., Su, H., & Pöschl, U. (2020), the model factors in three types of values: individual related, virus related, and venue related factors. To calculate the contract risk of any virus the infective dose D50, which is the number of virus RNA copies needed to infect 50% of subjects, is needed. D50 of COVID-19 was not available since it requires human testing, so it was calculated using data from Cov-SARS-1. The D50 for Cov-SARS-2 was estimated to be 316 virus RNA copies (Lelieveld et al. 2020). Then, D50 is used to calculate the probability that one virus RNA copy (P_{RNA}) can infect a subject, which was calculated to be 0.22%. Then, the model uses $D_{episode}$, the viral RNA copies inhaled by an individual, and the number to individuals at the venue (*n*) to calculate the final contract risk (Lelieveld et al. 2020). $D_{episode}$ is calculated by using factors such as the volume of the establishment, efficiency of masks, number of hours during the visit, concentration of viral RNA copies, the number of aerosols emitted during speaking and breathing, room ventilation, and lifespan of the virus in an aerosol (Appendix #1).

Transmission Risk Model:

$$R_{transmission} = \frac{\left[1 - \left(1 - \frac{Ab * [Tc_d - Tc_{d-10}] * 0.2}{p}\right)^n\right] \times \left[1 - (1 - P_{RNA})^{(D_{episode} \times n)}\right]}{I_{immunity}}$$

After combining $R_{exposure}$ and $R_{contract}$, the final $R_{transmission}$ model predicts the probability that an individual would be transmitted the virus based on the county's estimated 2019 population (*p*); circulating case estimate ($Ab * [Tc_d - Tc_{d-10}]$), where *Ab* is the ascertainment bias, Tc_d is the total cases in the county on the current day, and Tc_{d-10} is the total number of cases 10 days prior; COVID-19's P_{RNA}, the probability that one viral RNA copy can infect an individual; $D_{episode}$, the amount of viral RNA copies and individual will inhale during their visit; and *n*, the total number of people at the establishment.

Testing the Models:

Exposure Risk

The exposure risk model was testing by generating 200 random counties (appendix #2) using the Barnum Python Module. The exposure risk for a venue with 50 individuals was calculated for all of these counties for 18 different dates (list dates). The exposure risks for each date were plotted against the number of new cases per 1000 people for a date two weeks later. Using the Matplotlib Module, for each date, a "New Cases per 1000 people vs Exposure Risk Two Weeks Prior" scatter plot was created. The Spearman's Rank Correlation statistical test was used to correlate these two variables and was calculated using the scipy.stats module's "spearmanr" function. It outputted the Spearman Correlation ratio as well as a P-value. To find a general trend, the Spearman Correlation ratios and p-values were arithmetically averaged.

Contract Risk

The contract risk model was scenario tested to examine its viability before being used in the in the transmission risk model. Four scenarios were created, each one in a different venue – a high school classroom, house party, grocery store, and small office. These settings were chosen to test a wide range of factors (table 2). The contract risk was calculated for three different types of mask wearing habits – no mask, normal mask, and surgical mask. This information was stored on an Excel file.

Table 1. Characteristics of Venue #1

Venue #1: High School Classroom			
Room Properties		Event Properties	
Ventilation: (0.35 = no vent, 2 = ventilation once, 6 = public space)	2	Duration (hours):	7
Area:	60 meters squared	Population:	25
Height:	3 meters	Fraction of Speaking (% of time):	50%

Table 2. Characteristics of Venue #2

Venue #2: House Party			
Room Properties		Event Properties	
Ventilation:	0.35	Duration (hours):	4
Area:	60 meters squared	Population:	10
Height:	3 meters	Fraction of Speaking:	80%

Table 3. Characteristics of Venue #3

Venue #3: Large Grocery Store			
Room Properties		Event Properties	
Ventilation:	6	Duration (hours):	1
Area:	12929 meters squared	Population:	120
Height:	5 meters	Fraction of Speaking:	10%

Table 4. Characteristics of Venue #4

Venue #4: Small Office			
Room Properties		Event Properties	
Ventilation:	2	Duration (hours):	15
Area:	100	Population:	20
Height:	4	Fraction of Speaking:	60%

Transmission Risk

Transmission risk model was tested on the same 4 venues with three different levels of mask usage, but the transmission risk was calculated for scenarios where venues were located in three different counties. The counties Allegheny, Pennsylvania; Wyandot, Ohio; and Los Angeles County, California were grouped into low, medium, and high exposure risk, respectively. A hypothesis was made about which model would have the highest transmission risk based on venue factors (Tables 1 to Table 4) and the county exposure risk group.

Results:

Exposure Risk:

The exposure risk model averaged a spearman correlation of 0.81, which puts the correlation between the exposure risk prediction and new cases two weeks after in the “very strong” category. The average p-value of the statistical test was 1.57E-24.

Table 5: Spearman Correlation and P-Values for Each Date

Test #	Initial Date	Spearman Correlation	P-value
1	5/15/20	0.81480425	3.03E-44
2	6/15/20	7.53E-01	1.61E-35
3	7/15/20	0.84837689	1.18E-54
4	8/15/20	0.88068476	3.25E-64
5	9/15/20	0.81446534	8.54E-47
6	10/15/20	0.80981234	1.21E-45
7	11/15/20	0.80601881	3.80E-45
8	12/15/20	0.63733672	2.82E-23
9	1/15/21	0.81750437	2.08E-47
10	5/30/20	0.88634784	4.23E-63
11	6/30/20	0.79395872	7.67E-42
12	7/30/20	0.91893914	2.70E-78
13	8/30/20	0.84222822	1.60E-51
14	9/30/20	0.78671714	1.87E-41
15	10/30/20	0.83801217	7.11E-52
16	11/30/20	0.83859006	5.21E-52
17	12/30/20	0.66992323	2.28E-26

18	1/27/21	0.83283032	1.09E-50

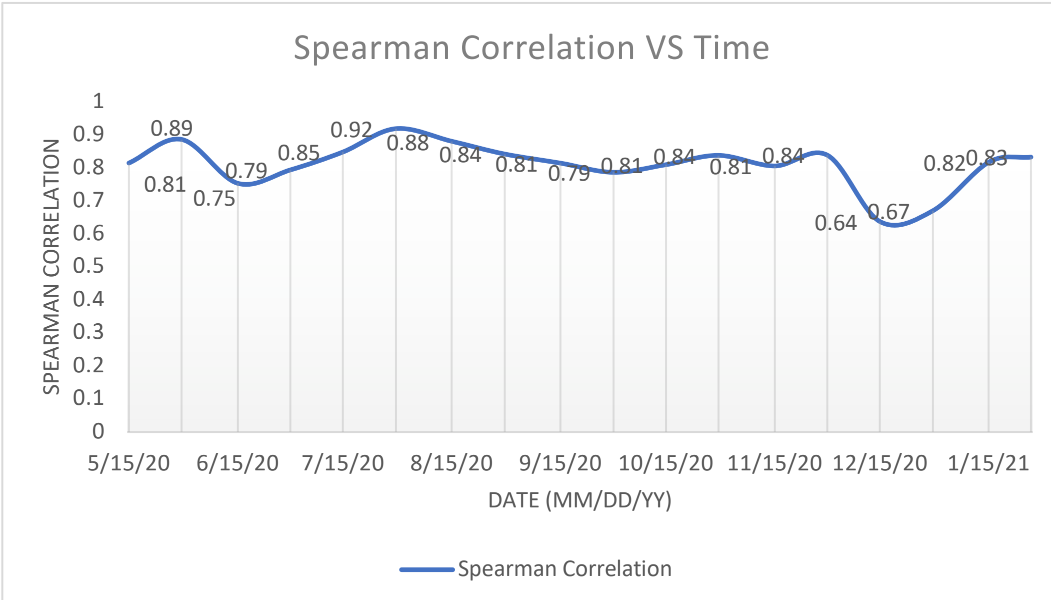

Figure 1: This chart visualized the exposure risk model's Spearman Rank Correlation ratio for each month from 5/15/2020 to 1/15/2021.

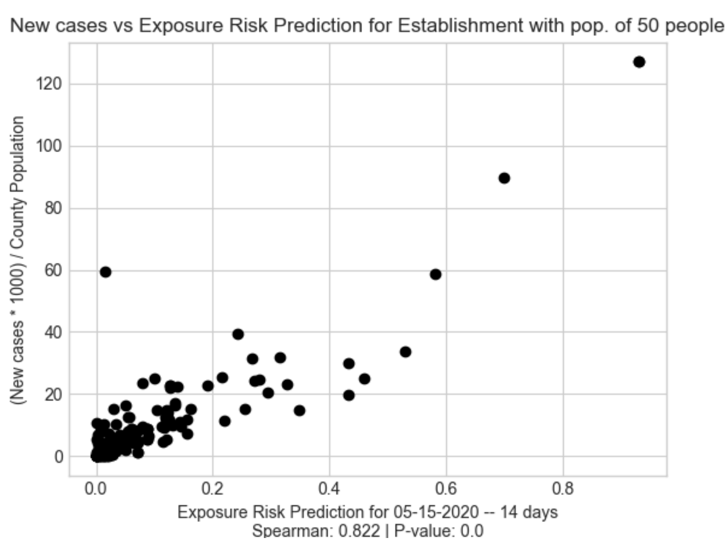

Figure 2: Exposure Risk for 05/01/2020 vs new cases from 5/01/2020 to 5/15/2020.

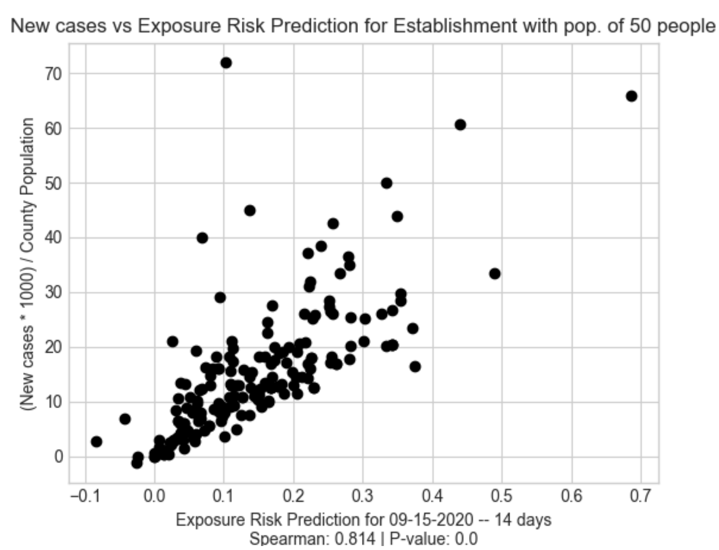

Figure 3: Exposure Risk for 09/01/2020 vs new cases from 09/01/2020 to 09/15/2020.

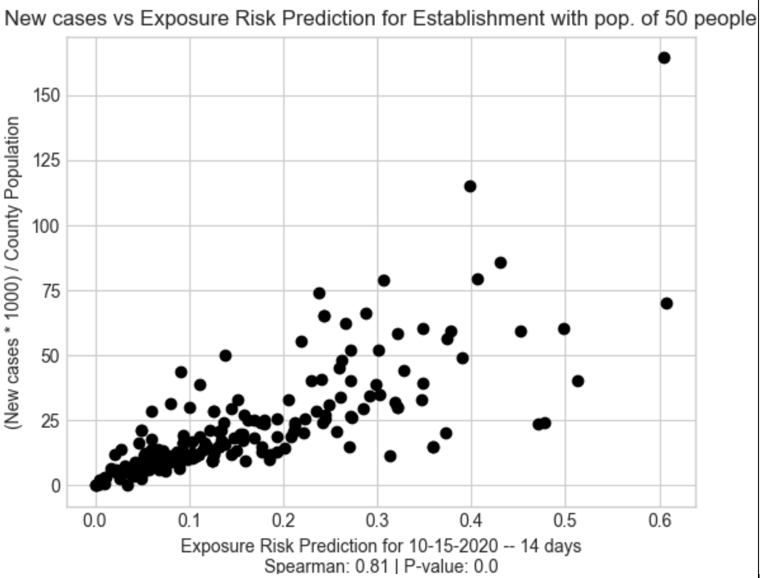

Figure 4: Exposure Risk for 10/01/2020 vs new cases from 10/01/2020 to 10/15/2020.

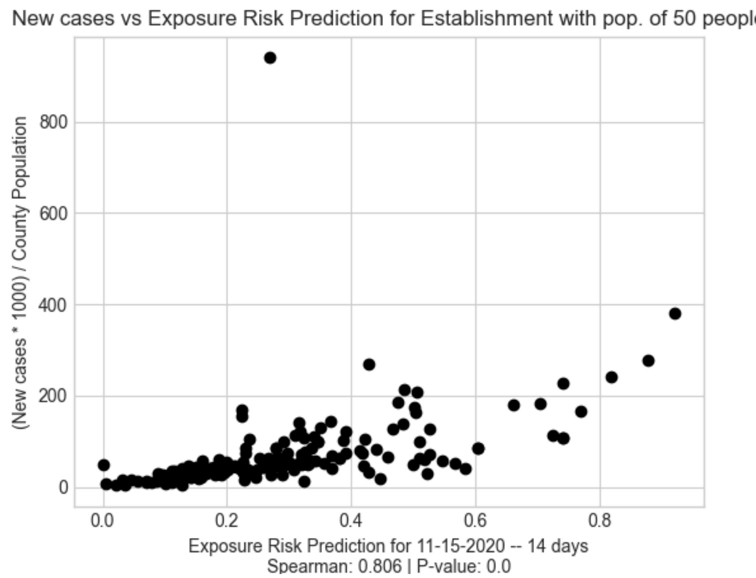

Figure 5: Exposure Risk for 11/01/2020 vs new cases from 11/01/2020 to 11/15/2020.

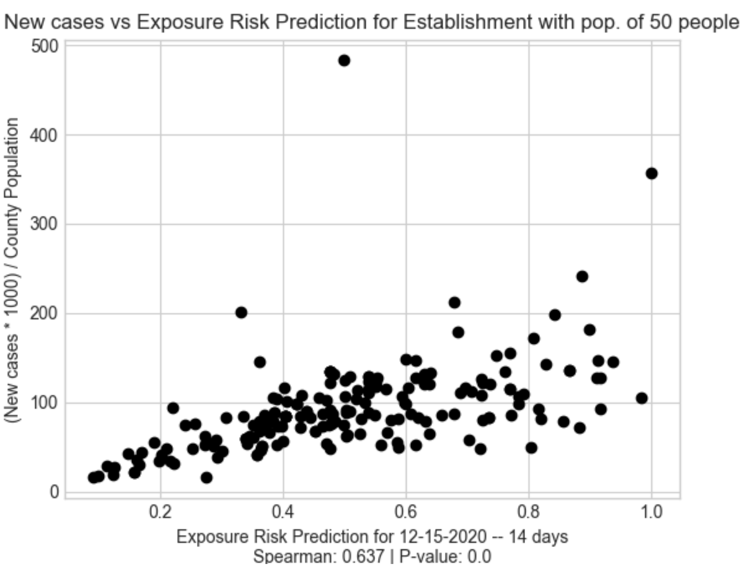

Figure 6: Exposure Risk for 12/01/2020 vs new cases from 12/01/2020 to 12/15/2020.

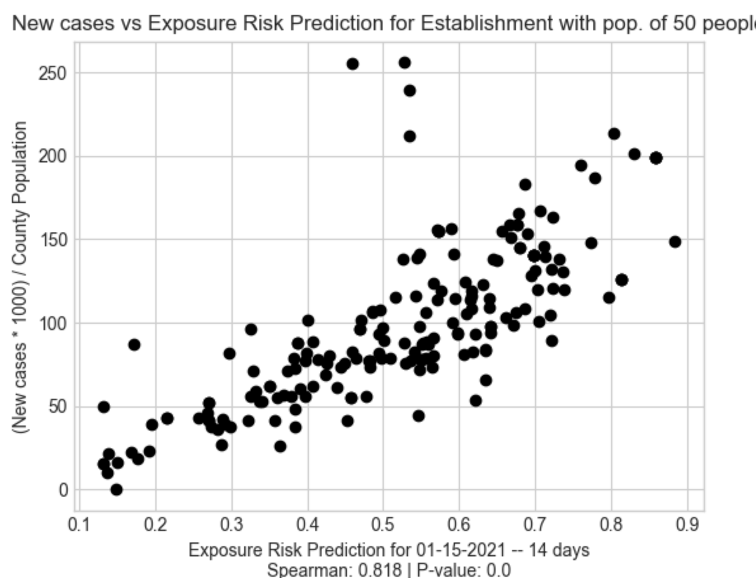

Figure 7: Exposure Risk for 01/01/2021 vs new cases from 01/01/2021 to 01/15/2021.

Transmission Risk Testing:

Table 6: Sample Exposure risk values based on venue population for low, mid, and high-risk counties used for transmission risk calculation in table 8.

Venue 1:	Exposure risk
Low Exposure: Allegheny, Pennsylvania	0.208094786
Mid Exposure: Wyandot, Ohia	0.332207167
High Exposure: Los Angeles County, California	0.487816321

Venue 2:	
Low Exposure: Allegheny, Pennsylvania	0.089102981
Mid Exposure: Wyandot, Ohia	0.149142754
High Exposure: Los Angeles County, California	0.234808223

Venue 3:	
Low Exposure: Allegheny, Pennsylvania	0.673689257
Mid Exposure: Wyandot, Ohia	0.856027211
High Exposure: Los Angeles County, California	0.959705708

Venue 4:	
Low Exposure: Allegheny, Pennsylvania	0.17026662
Mid Exposure: Wyandot, Ohia	0.276041947
High Exposure: Los Angeles County, California	0.414481545

Table 7: Transmission risk values based on mask efficiency and county exposure risk

Venue 1:	High School Class		
Conditions	No Mask	normal Mask	surgical Mask
low Exposure	0.198327055	0.190085886	0.163010267
mid exposure	0.316613743	0.303457356	0.260233233
high exposure	0.464918782	0.445599812	0.382129078

Venue 2:	House - Party		
Conditions	No Mask	normal Mask	surgical Mask
low Exposure	0.084079258	0.08017406	0.067945743
mid exposure	0.140733924	0.134197307	0.113729252
high exposure	0.22156948	0.211278324	0.17905371

Venue 3:	Grocery Store		
Conditions	No Mask	normal Mask	surgical Mask
low Exposure	6.23721E-05	5.34618E-05	2.67309E-05
mid exposure	0.0001044	8.94857E-05	4.47428E-05

high exposure	0.000164366	0.000140885	7.04425E-05

Venue 4:	Office		
Conditions	No Mask	normal Mask	surgical Mask
low Exposure	0.137485759	0.124692091	0.095557274
mid exposure	0.222896517	0.202154994	0.154920653
high exposure	0.334682803	0.303539064	0.232615919

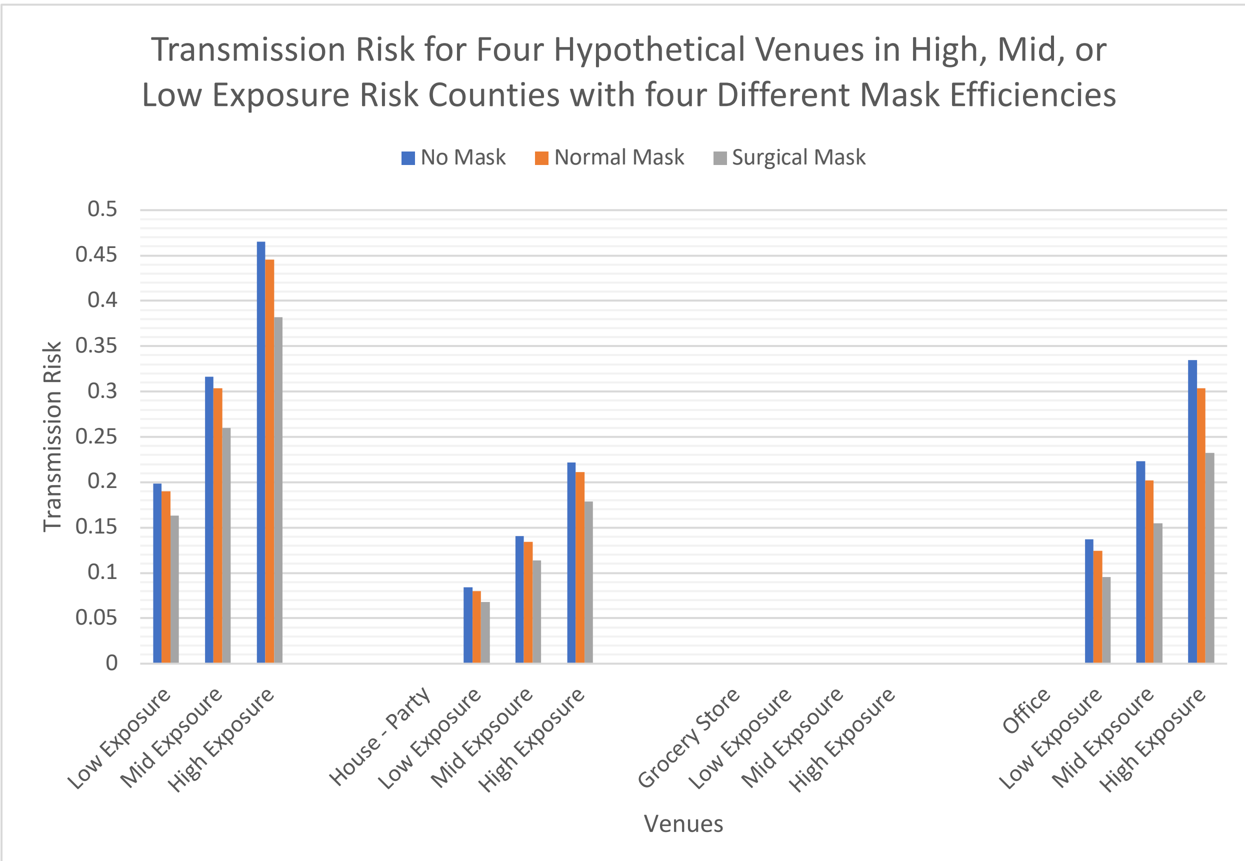

Figure 8: Transmission Risk for four venues - High School Classroom, House Party, Grocery Store, and Office - in high, mid, and low incidence counties and different mask situations**.**

Discussion:

The tests conducted on these models were done to shows that they are more accurate than models already in the market and they can provide their respective risk predictions accurately.

Exposure Risk

Firstly, the exposure risk model testing gave an average Spearman Rank Correlation of 0.81, which means that the model had a “very strong” positive correlation according to the statistical test. This shows that the Exposure Risk Model’s risk predictions are strongly correlated to the number of new case 2 weeks later. In addition, since the model was tested on a random set of 200 counties, it means these results would hold true for most of the United States. Because of the strong correlation, this model serves as an accurate model to use in the larger Transmission Risk Model.

After an analysis done between the dates and the Spearman Rank Correlation ratio it was found that the exposure risk model had a weaker correlation (about 0.2 lower ratio) for both dates in December (Figure 1). This can most likely be attributed to the unpredictable manner that new cases grew after period of vacation (Figure 6), where many individuals were in contact with outsiders. This suggests that the exposure risk model is not well suited to handle unexpected jumps in cases. Instead, as seen by the steady correlation between months seven to eleven, the model is more accurate during longer periods of fairly steady increase.

Some errors could have happened during the testing of the exposure risk model. Since this model was tested using random zip-codes generated through the Barnum Python module. One of the 200 counties may have been duplicates. Although this is extremely unlikely since there are 3006 counties in the US, it is still a possibility (source 2). This would impact how well the sample counties were analogous to all of the United States, which could provide biased results for a certain type of counties.

Finally, this model prove to be more accurate than other exposure risk models developed and tested using similar methods. A model created by Sun et all (2020), averaged a Spearman Correlation of 0.6, while this paper's exposure risk modeled averaged a ratio of 0.81. Both were tested using the similar methods – by correlating exposure risk to new cases two weeks later. This paper's model performed better because, when calculating circulating cases in a community, it found new cases between two weeks prior and the current date, while the other exposure risk model found circulating cases by summing new cases for ten days. In addition, my model's circulating case estimate factored in unreported cases and asymptomatic cases through the Ascertainment Bias and multiplication by 0.2. These factors made my probabilistic model more realistic and allowed to provide a stronger exposure risk prediction.

Since the hypothesis was that the there would a positive correlation between new cases and the model's exposure risk prediction, the hypothesis was confirmed by the average Spearman Rank Correlation of 0.81. The 18 tests conducted for the 18 different have a mean p-value of 1.57E-24, confirming that the data was significant.

Transmission Risk:

The transmission risk model provided sensible results relative to the other hypothetical venues, mask scenarios, and county types. There were two hypotheses for this test. The first one was that the transmission risk prediction for each combination of mask efficiency and country exposure would go in this order for each venue: Surgical Mask + Low Exposure County < Surgical Mask + Mid Exposure County < Surgical Mask + High Exposure County < Normal Mask + Low Exposure County < Normal Mask + Mid Exposure County < Normal Mask + High Exposure County < No Mask + Low Exposure < No Mask + Mid Exposure < No Mask + High Exposure. This hypothesis was created based on the simple fact that High Exposure Risk counties present a high risk of coming in contact with a carrier and more efficient face masks provided better protection against transmission of the virus.

Based off of the results in Figure 8, the model's results were exactly the same as the hypothesis. This means that the model's transmission risk prediction make sense in context with the other venues; however, this test does show that the model's probabilities are accurate. It only shows that that, relative to other venues, the probabilities (risk) make sense. To test if the actual risk prediction is accurate, on-site testing would need to be done in an actual community and real venue. Because of that, the transmission the risk prediction outputted by the model should be viewed as score between 1 and 100 instead of a probability. The score is helpful to compare different venues with each other.

The second hypothesis was that the average transmission risk prediction for each venue would be in this order: Grocery Store (Venue 3) < House Party (Venue 4) < High School (Venue 1) < Small Office (Venue 2). I hypothesized that the estimate would be lowest for the grocery store because of its lowest fraction of speaking (Table 3), its low event duration (Table 3), and most importantly its large volume compared to its population (Table 3). Since it has a low fraction of speaking and low event duration, an individual would have less instances and less time to inhale doses of the virus, leading to a lower transmission risk. Also, the larger volume of the venue violates he assumption of homogenous mixture of air, which is what Lelieveld et al. (2020) had assumed when creating the contract risk model used in my transmission risk model.

Overall, the model made the predictions similar to the hypothesis; however, the high school classroom (Venue #1) achieved the highest average transmission risk, instead of the hypothesized office (Figure 8). This is because the office actually had a larger volume and a lower amount of people inside the venue, so the result makes sense. In accordance with the hypothesis, this venue with the lowest score was the grocery store, mainly because of its lower population density and its large room size which violates the assumption made in the contract risk model of the homogenous mixture of air. This essentially means that the model cannot make accurate prediction for larger venues sizes; a potential solution to combat this would be to view the venue as smaller sections, such as rooms or hallways. It would be more accurate to get the transmission risk prediction for a specific room in a venue than the whole venue itself. For example, a school could be sectioned into different classrooms or a grocery store could be sectioned into different aisles.

Overall, the sub models (exposure risk and contract risk) and the final combined model (transmission risk) proved to accurate predictions for their respective risks. The discussion section only discussed the exposure and contract risk, since those were the models created by the paper. Contract risk was adapted from Lelieveld et al (2020). This model was implemented into a web-app that allowed individuals to input their zip code, personal factors, event factors, and venue factors to be presented with the transmission risk of attending their venue. Using these models along with the web-app individuals can assess the risk associated with visiting an establishment and take the right decision for their safety, hindering the virus to spread.

References:

Cooper, L. A., & Williams, D. R. (2020, October 20). Excess Deaths From COVID-19, Community Bereavement, and Restorative Justice for Communities of Color. Health Disparities | JAMA | JAMA Network. https://jamanetwork.com/journals/jama/fullarticle/2771762

Gamio, L. (2020, July 9). How coronavirus cases have risen since states reopened. The New York Times. https://www.nytimes.com/interactive/2020/07/09/us/coronavirus-cases-reopening-trends.html

Mortality risk of covid-19—Statistics and research. (n.d.). Our World in Data. Retrieved November 4, 2020, from https://ourworldindata.org/mortality-risk-covid

Oh, S., Watts, W. (2020, October 26). Dow ends 650 points lower as rising COVID-19 cases, stalled stimulus efforts highlight recovery jitters. MarketWatch. Retrieved November 4, 2020, from https://www.marketwatch.com/story/dow-futures-fall-260-points-as-covid-19-cases-surge-stimulus-remains-stalled-11603712231

Sun, Z., Di, L., Sprigg, W., Tong, D., & Casal, M. (2020). Community venue exposure risk estimator for the COVID-19 pandemic. Health & Place, 66, 102450.

Lumley, S. F., O'Donnell, D., Stoesser, N. E., Matthews, P. C., Howarth, A., Hatch, S. B., Marsden, B. D., Cox, S., James, T., Warren, F., Peck, L. J., Ritter, T. G., de Toledo, Z., Warren, L., Axten, D., Cornall, R. J., Jones, E. Y., Stuart, D. I., Screaton, G., … Eyre, D. W. (2020). Antibodies to SARS-CoV-2 are associated with protection against reinfection. MedRXiv, 9. https://doi.org/10.1101/2020.11.18.20234369

Center for Disease Control and Prevention. (2020, February 11). Information for Pediatric Healthcare Providers. https://www.cdc.gov/coronavirus/2019-ncov/hcp/pediatric-hcp.html#:%7E:text=While%20children%20infected%20with%20SARS,the%20same%20in%20adults.

Chande, A., Lee, S., Harris, M. et al. Real-time, interactive website for US-county-level COVID-19 event risk assessment. Nat Hum Behav (2020). https://doi.org/10.1038/s41562-020-01000-9

Lelieveld, J., Helleis, F., Borrmann, S., Cheng, Y., Drewnick, F., Haug, G., Klimach, T., Sciare, J., Su, H., & Pöschl, U. (2020). Model calculations of aerosol transmission and infection risk of covid-19 in indoor environments. MedRxiv, 2020.09.22.20199489. https://doi.org/10.1101/2020.09.22.20199489

Morath, E. (2020, June 3). How many u. S. Workers have lost jobs during coronavirus pandemic? There are several ways to count. Wall Street Journal. https://www.wsj.com/articles/how-many-u-s-workers-have-lost-jobs-during-coronavirus-pandemic-there-are-several-ways-to-count-11591176601

Sources:

1) https://jamanetwork.com/journals/jamainternalmedicine/fullarticle/2768834
2) https://www2.census.gov/geo/pdfs/reference/GARM/Ch4GARM.pdf

Appendix and Acknowledgments: