Project 3

Shelbie Burchfield, Frank Gattorno, Andres Makarem, Priyanka Shrestha

8/3/2020

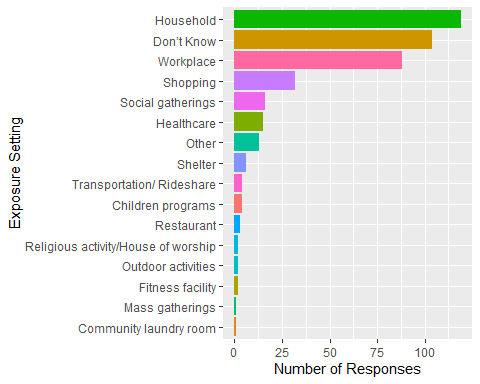
Table of Contents

# Introduction

This is an R Markdown document that works with extended contact tracing data collected by the Department of Health (DOH) of Miami-Dade County. This data set contains the data for a total of 302 patients who have tested positive for the novel coronavirus known as SARS-CoV-2 which causes COVID-19, and where they might have possibly contracted the virus as their exposure setting. The list of possible exposure settings were Workplace, Household, Transportation/ Rideshare, Hotels, Healthcare, Children Programs (e.g., daycare, summer camp, summer school), Shelter, Restaurant, Fitness facility (e.g., gym, studios), Religious activity/house of worship (e.g., services, choir practice, bible study), Personal care services (e.g., hair or nail salon, massage, tattoos), Social gatherings (e.g., parties, weddings, funerals), Mass gathering (e.g., demonstrations, fitness event, fundraising event), Community laundry room, Shopping (e.g. grocery shopping, malls), Movie theaters, concert halls, Outdoor activities (e.g., pool, beach, park, outdoor sports), Social club (e.g., book club, hobbyists), Don’t Know, and Other. This data set also included the ZIP codes and the occupation for most of those who stated “Workplace” as their exposure setting. Furthermore, most of the patients have also provided their relationship to the person from whom they might have contracted the virus. In this study we identify which of the listed exposure settings are the most prominent for contracting the virus. We also investigate whether or not there is a co-occurrence of reported exposure settings.

# Self-Reported Exposure to COVID-19

We sought to investigate the frequency of locations respondents identified as potential places of exposure. We hypothesized that locations where it is difficult to socially distance would show higher frequencies, especially in places that are necessary for regular life such as households, workplaces, and grocery stores. Since our graph shows the frequency and not rate, it can only help us elucidate where the majority of cases can occur, but it should not be used to determine high-risk places where transmission is easiest, as the number of people engaging in each activity can differ significantly.



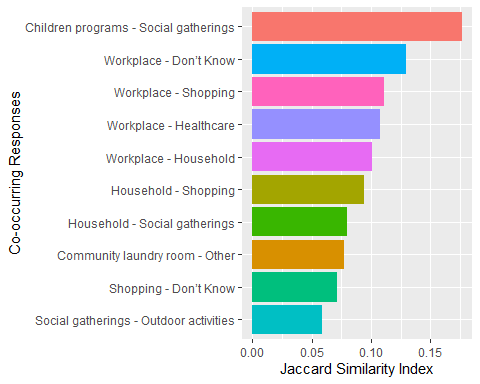
## Discussion of COVID-19 Exposure Settings

The bar graph shows an arrangement of the frequency of self-identified places of exposure in descending order. The following exposure settings were not reported by respondents in this data set and were therefore excluded from the graph: Social club, Personal care services, Movie theaters/Concert halls, and Hotels. We can see that over a third of total responses (119) state that they believe they became infected at their household, while 107 individuals did not know where they were exposed and infected. Completing the top 3 responses is the workplace with 88 total responses. The top responses show a pattern in accordance with our hypothesis. “Household” transmission appears to have the highest frequency of respondents, due to close-quarter interactions, difficulty socially distancing, and ease in identifying a source. The household is the one location during this pandemic where most people freely associate with others because they feel safe. Once one individual in a household becomes infected, there are usually no measures in place to prevent the spread to everyone else. Outside of the household, it is more difficult to identify the potential source, which is reflected in the high frequency of “Don’t Know” responses, which has the second highest frequency of respondents. “Workplace” transmissions had the third highest response, which again confirms our suspicion of likely places of transmission. It seems that keeping customer-interacting jobs open, or jobs where there is no ability to social distance, is what is causing further spread of COVID-19. Those individuals who contract the virus at their place of employment then go to their homes and spread the virus to their families. This reinforces the fact that better workplace policies need to be enacted to reduce the risk to essential workers keeping our society afloat.

It is important to note that while high-risk activities such as mass gatherings and fitness facilities rank lower, the graph shows us the absolute frequency of self-reported cases and not a rate; therefore, it is not indicative of the likelihood of infection for an individual if they were to partake in these activities. These values are further confounded by enacted policies such as curfews, or closures of mass social gatherings. Further information about the number of people engaging in high-risk behavior and infection is necessary.

# Exposure Setting Co-occurrence

We also sought to investigate the co-occurrence of locations respondents identified as potential places of exposure. We hypothesized that locations that co-occur in everyday routines, such as workplace and household or household and shopping, would also have greater co-occurrence as reported exposure settings.



## Discussion of Exposure Setting Co-occurrence

When observing the Jaccard Similarity Index for the provided data set, we can see that children programs and social gatherings have the highest Jaccard Similarity with each other at 17.6%. This means that of all the choices that were provided to individuals for where they believe they became infected with COVID-19, 17.6% of those who responded children programs also reported social gatherings as the possible location of infection. A possible explanation for this occurrence is that individuals with children are more likely to attend a gathering with other parents so their children can be entertained with each other. Living in isolation is difficult for everyone, but it is especially difficult for children who crave social contact. Children are also less likely to obey guidelines (social distancing and mask usage) for containing the spread of the virus. We can also theorize that individuals who are comfortable sending their children to child care programs are perhaps also more comfortable with attending social gatherings. However, it should be noted that this is a very small proportion of the total responses recorded. Of 302 respondents, only 16 responded that they believe they became infected at a social gathering and only 4 state that they became infected at a child care facility.

The second, third, fourth, and fifth highest Jaccard Similarities all include the workplace with another co-occurring place of exposure. This co-occurrence further emphasizes that going to the workplace is often a high risk activity which can lead to a high number of infections if proper distancing/mask usage is not able to be followed. The most notable co-occurrences are the Workplace-Don’t know, Workplace-Shopping, and Workplace-Household with 12.7%, 11.1%, and 10.1% similarities, accordingly. The Workplace-Don’t Know co-occurrence interestingly had the second highest Jaccard Similarity value, indicating that a large portion of those who think they may have gotten the virus at work are not entirely sure if this truly was where their exposure came from. Given that Workplace, Household, and Shopping had the three highest number of responses from individuals polled, it is not surprising to see that they all bear similarities with one another and are frequently listed together. It should also be noted that these are all locations which individuals must frequent in their day-to-day lives in order to sustain themselves and their families.

# Conclusions

The responses analyzed in this data set highlight the challenges faced by individuals during this pandemic to stay safe while continuing to do everyday necessities such as going to work or grocery shopping. Not all workplaces are able to support working from home, and many workplaces involve interactions with customers or clients. This puts employees at risk of coming in contact with individuals that are infected with coronavirus. Additionally, essential activities such as grocery shopping also put individuals at risk of coming in contact with individuals that are infected. Abstaining from these essential activities is difficult if not impossible when trying to sustain oneself and/or a family, and because social distancing is difficult to practice at home, anyone living in a given household is at risk of becoming infected if one individual in a household becomes infected. This may explain why Household, Workplace, and Shopping are some of the most commonly reported exposure settings, and it may also explain the co-occurrences seen in these exposure settings.