a. A different system that could be utilized to verify that the business license inputted by the user is valid for short term rentals in San Francisco could be for Airbnb to collaborate with the San Francisco planning office to cross-check their data base. As previously described in the project, we had to manual check with the San Francisco Planning Office to see if the imputed listing id was fraudulent or not. Therefore, if there was an ability to have access to all policy number ids, we could compare the imputed ids from the Airbnb lists with the registered policy number ids from the city. One argument that that solution might bring about is the protection of data that is relatively private to San Francisco. With a large organization having access to secure data from the city, some individuals could have concerns for the increased risk of a data breach. Additionally, some could have concerns about how realistic this cross checking the data can be. If listeners on the Air bnb website were to become aware of this process of cross-checking the data, then they could list their policy ID as pending so they could still list their location.

b. The research question that could be answered by the data set we have created could be: What is the ratio of affordable housing in San Francisco as listed on the Airbnb website compared to the rented household in San Francisco facing housing insecurity. The get_detailed_listing_database() function returns the cost of renting from Airbnb with the amount of rooms in the rental. If we got the cost and room breakdown, it could be compared to the percentage of individuals facing housing insecurity in San Francisco. We could use this ratio to help determine if some rental locations of Airbnb could be redistributed to fight the rising rate of individuals who rent facing housing insecurity, or, it could show that there is not an excess amount of rental locations being offered from Airbnb in San Francisco.

c. I believe that web-scraping is a complex issue that is directly correlated to data and data breaches. One one hand, collecting data can help platforms perform better based on user interaction and feedback. However, collecting people's data without their knowledge and using it for malicious purposes is extremely unethical. In this case, I believe that the law needs to take precedent and provide stricter regulations around data collection and privacy. If companies were going to web-scrape, I believe that the law needs to enforce stricter oversight and force these companies to state how they were going to collect and use this data. Additionally, I think the law needs to enforce stricter regulations on companies to protect users' data.

d. As I previously stated, web scraping is a complex issue. While there might be positive outcomes from web scraping, the act of web scraping is a complex ethical issue. Therefore, one guideline I would consider when deciding to use public data is if I know the data I am taking is collected with context and cannot be skewed to cause harm to the individual it originated from. This is so the data is not twisted to cause misinformation

and the spread of harmful lies. Additionally, another guideline that I would consider is if the person has any idea that their data is in the public domain. For example, if the user has a public account and understands the terms and conditions of the website then I would feel more comfortable using their data if it was allowed in the terms and condition of the website. However, if it was a private account on a platform that did not allow web scraping in their terms and conditions, then I believe it to be unethical to web scrape.