Project 2 - Questions

a. A different system to verify San Fransisco Airbnb policy numbers would be to require verification from the city before allowing a listing to go public. For example, when a lister goes to post a new Airbnb listing, they would first have to insert their listing number. Then the listing would sit on a private page of the website only available to the lister until the city could verify the policy number with Airbnb. One argument against this system would be that that it wastes city officials' time and the resources could be put toward something else. Another argument against it could be that this system would cost Airbnb property owners' money because their listings wouldn't be able to post right away and would cause fewer people to see them which could lead to less rentals.

b. One research question that could be answered by access to data on housing insecurity in the US is, how does the number of Airbnb rentals in an area correlate with the percent of the population facing housing insecurity?

c. One factor that I think is important to consider while discussing the legality of web scraping is the extent of adverse effects on the website owner. For example, if web scraping requests cause a website to crash and loss of sales, I think there should be consequences for the web scraper. For example, QVC's website crashed due to Resultly web scraping item price data which prevented QVC from making sales for two days. This had monetary consequences that were unjustified. However, a person who wishes to retrieve data from a public website should not face consequences if the actual act of web scraping caused no harm to the website. Civilians should be able to access this data and should be able to hold big companies accountable.

d. One guideline is that the data must be repurposed or thoughtfully construed if it is going to be shared. If the data is just plainfully presented or monetized exactly how it was found, this would be an infringement of someone's intellectual property and would be a copyright infringement. In addition, I don't think web scraping should be allowed on social media accounts. If someone wants to gather the information from the website of a social media page they should not be allowed to get personal information from people's accounts. The person who has made the account is most likely unaware of this possibility and certain information like address, phone number, etc should not be extracted with web scraping.