Reflection

a. If becoming a software developer at Airbnb, I would recommend a collaboration between Airbnb and the San Francisco Planning Office (SFPO) to make sure to validate business licenses / policy number of new listings prior to posting them on the Airbnb site. I would also recommend for the SFPO to request Airbnb to send all the policy numbers of current listings that have already been posted. The first argument that I will expect to get would be that Airbnb must already expect all the listings posted on their website must already be valid as it is the duty of SFPO to grant people their business licenses. The second argument would be that it would need a lot of capitals to work with SFPO and will definitely take a lot of time before being able to validate just one listings as the process must be complicated from either end.

b. From the data scraped from Airbnb, it can be seen that all of the listings are located in the Mission District. Since we only web-scraped a small portion of the Airbnb website, there should be more listings that we should have included in Mission District itself. Mission District is just one district in San Francisco, meaning there should be countless Airbnb listings just in the city itself. San Francisco also has a really high cost of living and high rent. With people listing different places around the city for rent instead of leasing them, it would decrease the supply of apartments or places in general for people to lease. However, listings without business license like listing id '16204265' are driving the situation worse as they operate without having a license which is illegal. Hence, this would drive prices up creating housing insecurity as people are having a harder time finding affordable housing with even basic amenities in the city.

c. Though it is sort of illegal to do web scraping as technically the act copyrights the web, violate the CFAA, and trespass to chatel, a website usually has its own terms of service which lists out violation acts when scraping the site. According to the article about legal issues with CFAA, an act that violates the terms of service of a web will be criminalized. Even though terms of services of webs are usually hidden or lengthy, what if the act of web scraping only scraps the section of the website that is not prohibited or listed under the terms of service? What if the act of web scraping is helping the business itself for catching fraud or illegal activities like the listing id '16204265' as they are able to post their listing without having their policy number or business license verified. I believe that if

the intention and act of web scraping a certain section of the web should always lie under the terms of service of the web.

d. One guideline that is really important would be reviewing the public data that a company or a website publishes. If a public data is published, the company or web developer must be expecting viewers and users to be able to access the data as well as interpreting it. So, by just web scraping the public data available from the official company/website is better than finding ways to find public or private data of a certain company from unofficial sites. Another guideline would be the ethical framework of using the public data. As research after web scraping is done and published, anyone can judge them wrongly as people think differently. Someone could review our work and immediately think that Airbnb is not a great company as it allows illegal listings which in the end leads to lawsuit etcc. This would then lead to the invasion of privacy of Airbnb and even worse. Hence, I feel like when using public data for research like web scraping, a good explanation and justification of the research itself and interpretation of the result obtained should be clearly defined.