Lea Srebrenik

a. As a software developer at Airbnb, a new system that I would recommend being put in place is one where each listing posted on the Airbnb website is not officially uploaded until the policy number is checked with the San Fransico Planning Office. This means that every time someone creates a listing, the policy number they enter is sent to the planning office while their listing is awaiting approval. The planning office can then check if the number is valid, which can prompt the Airbnb website to either post or reject the listing. One argument that I may hear from Airbnb is not wanting to adopt this system because of the waiting period, which could potentially reduce sales. The Planning Office may take a while to validate numbers, which could lead to very long waiting periods. Another argument against this system could be that the Planning Office might not want to be bombarded with requests from Airbnb and would not agree to do this.

b. One research question I could explore is "Have rental prices in the Mission District led to housing insecurity?" The data I collected through this project includes rental costs for 20 listings in the Mission District. This data could be used to explore this research question. It can be compared to people's salaries in that area to see if rent prices are affordable, and to what extent rental prices have an effect on housing insecurity.

c. One factor that I believe is important to consider when discussing the legality of web scraping is testing for discrimination. Oftentimes, bias is embedded unintentionally or even intentionally into algorithms. Through web scraping, researchers are able to test algorithms for discrimination. If web scraping is legal, there can be algorithmic transparency and accountability. We can work towards creating a more just and equitable society by testing for bias in algorithms and preventing discrimination. I think this factor is very important to consider when advocating for web scraping's legal standing.

d. One guideline that is very important to consider when deciding whether to use public data is user safety. If this data is being shared, it is important that the personal information gathered is not going to put a user at risk of danger. Even though the data is publicly provided, users may not be aware of how their data is going to be used, and it is very important to protect their safety and privacy. For example, using data that contains personal addresses could expose users to safety risks. Another guideline to consider is transparency surrounding the purpose of using public data. Openness about the usage of data should be a guideline. Knowing why and for what reason data is being collected and analyzed will make people more comfortable and allow them to feel informed as to why their data is being collected. I think that if a researcher does not want to be transparent with what they are using public data for, it is an indicator that they

should not be using it at all.