

1. As a software developer at Airbnb, I would recommend a system that cross-references business license data from local governments, specifically San Francisco in this case. To collect this data, the SFPO could compile a list of short term rental licenses that is updated regularly that Airbnb could utilize to ensure that all users are validated before listing a property. By cross-referencing short-term rental licenses with the SFPO, Airbnb could uphold a higher system of accountability and minimize the complaints from third parties. One argument against this could be the resources it takes to implement. Although it is one minor tweak in the system, it could cost a lot of time and money to put in place and even more time ensuring the legality of implementing a system like this. Another possible complaint is that for the accountability it is providing Airbnb, a system like this could potentially lose profits for the company by discouraging un-licensed individuals from uploading listings, which may be a large group of users.
2. One question we could explore using the web-scraping data is whether housing affordability has a correlation with the number of Airbnb listings in the area. In this project, we found the prices of listings as well as the number of listings given the HTML data, so with it we could analyze how Airbnb is impacting the housing crisis in major cities like San Francisco, where the homeless population has skyrocketed.
3. One factor to consider is the degree which web scraping is affecting a site owner. Although the laws surrounding 'trespass to chattel' are somewhat vague and ambiguous, most web scraping is still considered legal. It crosses a line once the trespasser adversely affects the site owner and the purpose of the site. This is an important factor to consider because not only are courts more lenient towards protecting proprietary information but without this protection, concerns can be raised about privacy while using the internet for both web developers, site owners, and everyday users.
4. One factor to consider when deciding to use public data is whose information it is we are accessing and what type of information it is. When knowingly accessing highly personal information, such as medical records or criminal backgrounds, it raises an issue of ethics and privacy. Web scraping is a way to keep entities accountable and improve visibility, but when used to access information that an individual would not want to be shared, we are defeating the purpose of it. Additionally, understanding whether the personal data was collected with the consent of the individual is incredibly important. It raises an issue already if the information provided is not consensual, but going further to look through this data is even worse. As users of public data and human beings, we should be aware of the ethical issues surrounding web scraping and be sure to use data for the good of others.