

Vikram Reddy

November 18th, 2022

SI 206

### SI 206 Final Project 2

a.) A different system that verifies whether a business license is valid for short term rentals in San Francisco is to use programs such as an API. API's can hold large amounts of data and allows the user to web scrape the entire website for all the content. The government could provide some state-sponsored API's for the software company to check if the policy number is valid. Every house sharing company like Airbnb can use the API and if they find that the policy number is not valid, then the applicant's application would be taken off. I believe that Airbnb should take ownership and be in charge of validating their policy numbers and therefore, this should be integrated into their website.

Two arguments against it would be some problem with the API system. Someone could create a fake policy number and that could lead to a data leak and a privacy problem. Everyone can access the website and that would leave sensitive data at risk. The second argument is that if they do end up using an API base, the database has to be updated at all times. That means that the government would have to put a lot of money into the system, in order to update privacy and keep up storage, and would lead to a lot of tax money being used on updating the storage area, and fitting it with security and brand new technology.

b). One research question that I could explore using this data is the level of housing affordability in the country. In my data, I am able to see the prices of all the listings of houses on the Airbnb

site. Using my data, we can find out where each house is located, how prices differ per location and be able to determine which regions of the country are dealing with unaffordable housing. Housing activists can then go to those specific regions and petition for change, ask the local governments to pass rent control laws and even protest if that is the only way to get the message through. Housing activists can also go to these places and help out with people who may face homelessness due to rising costs, and provide temporary food, support and shelter.

c.) One factor that is important to consider when discussing the legality of web scraping is the ownership of content and the right to access it. Many web scraping companies have run into problems of violating a company's terms of service, in order to find data on a variety of topics such as whether a company discriminates against certain genders or how fast a novelty item sells in the first 24 hours. I believe that web scraping should be legal, if helps prevents any form of discrimination. However, if web scraping is used for other means such as buying the entire stock of new limited edition shoes, then there should be some legal actions to protect companies from this invasion of privacy and to give other customers a fair chance at owning that product.

d.) One guideline we must consider when deciding to use public data, is whether the data is a benefit or a drain on society. If someone is web scraping data, in order to create a website of the cheapest places to shop for people with low income, then I think this would be a good and ethical use of web scraping. However, if someone was web scraping data, in order to buy the newest graphics cards and then to resell them on the market for double the price, then I would think this would fall into an unethical guideline

Another guideline to consider would be the type of information on the internet and the privacy guidelines in place. For example, if someone was web scraping data of hospital records and then found private information of an individual, I would think that would be a breach of privacy and highly unethical, as the person never gave consent to have their records analyzed. However, if the information that is getting web scraped is public information, such as a Washington D.C Metro schedule, then I think that would be fine, as it is public transport and is funded through public taxes.