Andrew Nakamura
SI 206
Dr. Barbara Ericson
November 15, 2022

Project 2 Questions

A. I would largely keep the system the same, but I wouldn't allow AirBnb users to make their properties publicly available until their business license is fully validated, rather than allowing listers to set their policy number as "pending." I think that the lapse between having a license creates unwanted variables in the quality control process, and allows listers to avoid accountability for their practices. Arguments against this policy could be that it would deter listers from using AirBnb to list and display fewer options for renters, both losing the company profits and creating a false idea of scarcity in the AirBnb renting market, potentially driving up prices for the existing properties.

B. My home state of Hawaii has among the highest rates of owned households facing housing insecurity, rented households facing housing insecurity, and total housing insecurity according to the data presented on the [Housing Insecurity in the United States](#) Wikipedia page. Our homeless population is so high in large part due to property owners who are not residents of the state keeping houses vacant to rent, especially during AirBnb's rise in popularity. Because these property owners often drive up the rent of these properties, they are only accessible to tourists and short-term renters, and not to local residents. I would use the database explored in this project to compare the prices of AirBnbs in my local area to other property values in the area, as well as tracking the rental price increases of both the AirBnb properties and the local rental properties overtime, in order to see the effect of the mass rental uprising on the affordability and standard of living for local residents.

C. A website's terms of use has been used to argue against the use of public web scraping in the court cases of both Cvent, Inc. v. Eventbrite, Inc and QVC v. Resultly. When discussing the legality of web scraping, it is important to keep in mind the terms of use of the websites being scraped. However, this raises a few questions about terms of use itself. How should these individual websites set their terms of use in an appropriate way that doesn't fully restrict users? Should these terms of use be standardized to prevent these types of exploits, and who should set them? With web scraping being technically legal, how should these websites reconcile with setting terms of use preventing it? While the court cases presenting these issues are quite egregious, the ability for a company to set their own terms of use brings tension between the company and its users, and without proper regulation set, could be abused to let companies scam their customers if left unchecked by governmental entities and the people themselves.

D. Two guidelines for ethical web scraping could be: explicitly ask to scrape data, state your purposes before scraping, and show exactly what data has been scraped and only send a reasonable amount of requests per second to avoid rendering a site inoperable. Just as we would like public entities to be open with their data, we should be open with ours. While this may not allow us to find the information we are looking for and leaves lapses

in the ability to hold an entity accountable, it's more important that this data is collected ethically and protects the privacy and rights of peripheral individuals whose data could be collaterally scraped. It's also important to ensure the safety of the site itself by not DDosing the website through an excessive amount of requests.