

Bhousel

Bobby Housel

54626941

## SI 206 Project 2 Written Responses

1. One idea would be to get our list from the SFPO. Instead of our original method, we could persuade the SFPO to give us a list of policy numbers which we could then use to reference against each policy number and find out if it's legitimate. One classic problem with marketplace businesses like Airbnb is that you will inevitably always encounter fraudsters. Our new system thus still does not solve for people that are fraudulently classifying their property as pending on the site. Another obvious hiccup with our new plan is that it is not very agile, and rather static. The list we get from SFPO may not update with information as time passes. With all the fast-moving parts of the industry and licenses, it is likely our new system would not be as up to date as we would hope, and thus remain somewhat static.
2. A research question that could be explored using this data could be: What is the relationship between average rent in a county and that county's predisposition to housing insecurity? Throughout the years as rent has skyrocketed across the US, there seems to be some correlation between the economic living costs of a place and housing insecurity in that area. There was a ton of interesting data on the Wikipedia page that could be analyzed, but this was just one thing that stood out to me given how prevalent housing insecurity across the nation is becoming, in addition to concerns over increasing inflation.
3. I think one important factor to consider when discussing the legality of web scraping is the value that is being received via the scraping, and whether or not any reciprocal harm is being caused to the web host (said in another way, is the web scraping a "zero sum game" or not?). This ties into one of the theses primarily described in the [Electronic Frontier Foundation's article](#), in that web scraping may break a website's terms of services, but not necessarily be a crime. Furthermore, web scraping could be unethical at times, without necessarily being illegal. The second that web scraping starts to "steal" potential IP while breaking terms of service though, is when a seemingly harmless act can turn criminal quite quickly. It's very important to remember why you are scraping and whether your scraping is ethical, breaking the terms of service, the law, or perhaps living in some muted grey area.

4. I think one of the most widely acknowledged concerns about web scraping in the US has to do with privacy concerns. I believe it is important to remember privacy concerns when discussing the legality of web scraping because so many users of the internet are unknowingly having their data scraped and having that data used without their express consent every day. Just because you posted something on your public Instagram story, does not mean you agreed to having your entire profile scraped, sorted, analyzed, and interpreted. Thus, I believe there is a fine line to thread when talking about web scraping and privacy. The two major guidelines I'd consider would thus be: Is the creator of the content and everyone involved in the content being scraped over 18 years of age? Secondly, was the content approved for scraping, or was it to be expected/relatively known that the content would be scraped? If the answers to these guidelines are both "yes," then scrape away!