Lauren Snerson

Project 2 Reflection

1. Currently, Airbnb is running on a complaint service looking at general complaints made by the public. Because of this, there is an extreme accountability issue, as illegal rentals are complicated to report and hard to understand. In order to assure that business licenses are real and accurate, a new system needs to be put into place at Airbnb. One thing Airbnb could implement is an automation service to verify the policy number on an individual listing to see if it follows the correct format. This service can incorporate inputting the policy number directly and matching the information with factors such as time, numbers, and any unique characters the number may include. Further, there can be a face ID requirement when inputting the policy number that must match the face sent in with the business license. With any new service or validation system, there will be backlash and opposing arguments. For instance, listers may say that this brings upon a feeling of distrust. Airbnb is based on trust, as people rent out their homes and allow people the privacy of staying in their personal space. If listers do not feel trusted by Airbnb, why would they feel comfortable renting out their homes? Further, this will allude that Airbnb does not trust its listers, making customers less likely to rent from them. Another possible argument is that going through each policy number and possibly requiring face ID technology could be a great added time commitment and expense for employees.

2. One of the things I enjoyed most about this project is that the database we created is extremely useful and important. Data scientists scrape data and look into data like this every day in order to better understand platforms. Housing insecurity is defined by looking at various factors including income, housing quality, neighborhoods, overcrowding, and homelessness. There are many questions that come to mind from my data when thinking about housing insecurity and the corruption in this field. One research question that I could explore using my data is how affordable housing is when looking at it in terms of the number of rooms there are in each household. In other words, our data will help us analyze whether or not Airbnb is the most affordable housing option

compared to other competitors, such as Vrbo. With our database, we can look into each listing on Airbnb, focusing on how many rooms the house has compared to the price of the unit as a whole.

3. There are many important factors to take into consideration when discussing the legality of web scraping. What I have learned through this project and through research is that web scraping is completely legal, as long as one is scraping data that is publicly available on the internet. The reason this is so important is that it is very easy to abuse this power and violate laws of privacy and data security. When thinking about legality, it is most important to think about why the scraper is doing this action. This includes their intent and what the information will be used for. The intent behind the scraping helps us understand if one is trying to find information that could be used against someone or something, or if they simply just want to find public information easier. Take for example eBay v. Bidder's Edge, an example from the Wikipedia article. In this case, the intent of the scraper is malicious, as they are trying to gain special treatment and information. This is a prime example of when a scraper uses a computer system as a way to trespass, and in this case perform auction sniping. This is when one will place a bid that will exceed the current highest bid as late as possible, giving other bidders no time to come back. This intent is not ethical and proves why it is important to look at the intent of the scraper. In another example, one ethical intent of web scraping would be websites or companies that help users compare prices of the same product across different retailers. These websites use web scrapers to scrape data and prices that are publicly available, to make it easier to and more efficient to compare prices.

4. Although web scraping is legal with public information, there are many issues in regards to ethics and privacy. It is easy to violate someone's privacy on the web, if you do not follow the ethical guidelines when scraping. When deciding whether ot not to use public data when web scraping, it important to remember to give credit when it is due. When scraping, the data one collects is not their own, and it is illegal to allow others to use it. While not all websites will demand credit, in order to avoid legal allegations and be ethical, remember to give credit and respect copyright. Further, another guideline to

consider is not using personal information that isn't publicly posted on the website. It is important to remember not to scrape private data. I learned through research that scrapers should refer to the website's robots.txt and analytics to avoid scraping data from sensitive areas. In terms of personal information, names and biography information that is public is fair game, however any financial information, personal records, or hidden information that is found is unethical to report and should be left private. One should only web scrape if the information they are gathering is being used for beneficial purposes, whether it be to benefit them or society as a whole.