

1. Currently, listers and Airbnb are only held accountable through only complaints made to the government. This means that there are likely short-term rentals that are illegal, but are still available if no one has made a complaint about them. One alternative system for validating short term rentals could be to make each rental go through a process of validation before being put on the market. This would ensure that only valid and legal rentals are listed. However, this method would be very time consuming and costly. It would require either hiring additional staff to validate each prospective rental or require development of new software that can validate each rental, both of which are costly options. Another issue with this system is that even if it were to be implemented today, all existing rentals would have to be validated or otherwise stay on the market even if they are not legal rentals. Validating each existing rental would be even more costly and would require an immense amount of time and labor in order to check.
2. Looking at solutions for housing insecurity in the US, one solution mentioned are starter homes. One research question data scientists might have is, "Is renting short-term rentals an efficient alternative to owning a smaller home, or in the long run, is a better solution to combating housing insecurity owning smaller homes?" The database we have created provides a list of properties and their details, including price per night and the number of bedrooms. This information can be compared against the cost and size of smaller homes that are currently available to buy rather than rent. These comparisons can then be used to individually assist families looking for housing with a certain budget and size, and determine whether it would be best to get a short-term rental or invest in owning a smaller house. The data can also be used overall to compare the availability of different size and price housing options to determine what housing types are in abundance and which are scarce.
3. One factor that seems to be important when considering the legality of web scraping is the intention behind the action. I think if the web scraping is being done with the intention of helping others, holding a website and its algorithms accountable, or for research or journalism purposes, then these cases should be evaluated with a different standard than instances where bots or others are using web scraping to undermine a website or use the information in a way that would harm the website/owner. For example, in *Sandvig v. Barr*, the goal was to research how algorithms unlawfully discriminate based on characteristics such as race or gender. This type of research was conducted in order to hold algorithms accountable and test them for discrimination, whether it is intentional or unintentional. This is an example of the type of research that should not only be allowed, but is needed in order to ensure algorithms do not discriminate, and should not be in violation of the Computer Fraud and Abuse Act.
4. One factor that is important to consider when discussing the ethics of web scraping is transparency in using user's data. A user should explicitly be made aware of how their data would be used, by who, and for what. Additionally, if data is going to be used by a company for more intensive purposes than what would be considered typical, a user should be informed in a clearer way than just through terms and conditions, which are

often lengthy, hard to read, and often skipped by most users. When a company lacks transparency with how they handle their user's data is when unethical practices begin to come into play. Another factor is anonymity and privacy of the data being used. If a user's data is going to be used, especially for something that will eventually be made public, it is important that the user's identity is kept anonymous and cannot be traced back to them. Additionally, if any data is used for research purposes but will not be made public if it is confidential information, a company has the responsibility to ensure that their user's data is highly protected and that the company can be trusted to keep this data secure.