

**Project 2 Reflection Questions**

1. Another approach to ensuring that listers have an official business license is creating a feature that performs this verification for us using the official database of business license numbers. The feature would be created so that a lister cannot post a listing on the AirBnb website unless the policy number that they enter matches one in an official database under their name. This database ideally holds all official business license numbers along with the corresponding name of the person who it belongs to. Under this system, we could catch users who attempt to post short term listings without obtaining an official business license number. Therefore with this feature, third parties would not need to check the AirBnb listings to file complaints, because the verification process would be carried out before the listing is posted.  
One argument against this system is how to access the official database. There may be many regulations and privacy rules against AirBnb accessing the official database, and as a result this system would not work. Additionally, this feature does not make sure that the name of the lister matches the name of the number and name in the official database. As a result, people who do not have an official policy number but have access to one could post a short term listing on AirBnb.
2. One aspect that could be further explored with our data is housing affordability. Question: What is the cost of an AirBnb per night compared with the average review rating given (positive or negative)? Is there a correlation between price and rating? Do higher rated places have a larger price than those with lower ratings?
3. One factor I believe is important to consider when discussing the legality of web scraping is the motivation of the web scraper. What is the purpose of web scraping in this case? What is the data scientist using the data to analyze? I believe that the legality of web scraping should be determined on a case by case basis. If the data is being used with malicious intent to unrightfully expose a company or organization, this is different from the data being scraped to create a chart or statistical analysis. Therefore, the intent behind the web scraping should be taken into account when discussing legality issues.
4. Two guidelines to consider when deciding to use or not use public data: If people's data is used without their consent, their name should be kept out of the publishing of the data results. Using this guideline, the data can still be scraped but anonymity will be respected. Additionally when using public data, the user should make sure they are not changing or altering the data. Therefore, the public data remains protected and in the same state as it was before being scraped.