a.) Throughout this project, we acted as investigators to uphold the system of accountability created by the San Francisco lawmakers: listers must register with the city’s planning office and put the business license’s number on Airbnb’s website, Airbnb must display some effort in validating these policy numbers, and third parties can register a complaint of illegal short-term rentals with the city planning office. We used web-scraping to do the latter using several hours of our personal time. Imagine you’re a software developer at either the San Francisco Planning Office (SFPO) or Airbnb.com. Describe a different system that verifies that the business license is valid for short term rentals in San Francisco and list at least two arguments you might hear at your organization (either SFPO or Airbnb.com) against adopting your system.

- If I was a software developer at either of these establishments, I would require that a proper business license be provided before allowing the listing to go up on the website instead of allowing options of “Pending” to show up as an option. One argument I might hear could possibly be that this would limit the number of listings and thus limit the customers as we are turning away people who are in the process of obtaining these licenses. Another argument might be that if the license is pending, that means the license will be obtained in the near future so it should be acceptable.   
   
b.) The database we’ve created through web-scraping is a great data source of information for data scientists in order to answer and explore research questions. Skim through the Housing Insecurity in the US Wikipedia page and describe at least one research question that you could answer or explore using this data if you were a data scientist working with a housing activist organization to fight against housing insecurity.

- One research question that could be explored using web scraping is if certain areas of the country are more Housing Insecure than others. One way I would do this is by assessing data regarding housing affordability across the US. According to the Wikipedia page, housing affordability is equal to annualized housing cost divided by annual income. If someone spends more than 30% or 50% of their income on housing, they are considered housing insecure. By applying web-scraping, I would be able to easily retrieve this data and visualize housing affordability across the US and then further compare that data with demographics/geographic data about the country to see if there is any overlap.

c.) As discussed in the introduction, the legality of web scraping is still uncertain in the US. Skim through the Legal Issues section of Web Scraping in the US on Wikipedia and this article about the legal issues with the Computer Fraud and Abuse Act , and describe at least one factor you believe is important to consider when discussing the legality of web scraping and why.

* In the information about the legality of web scraping, it was mentioned that website owners can use copyright infringement as an avenue to prevent unwanted web scraping. I believe this is an important factor to consider in the discussion of the legality of web scraping because under copyright law, there is a fair use doctrine that basically states that under certain conditions, permission is not needed to use copyrighted material. I think this is very important because, at least in the case of this project, the web scraping uses the data from the website to further critique or comment on the original data, therefore it should be protected under fair use.

d.) Scraping public data does not always lead to positive results for society. While web scraping is important for accountability and open access of information, we must also consider issues of privacy as well. Many argue that using someone’s personal data without their consent (even if publicly provided) is unethical. Web scraping requires thoughtful intervention, what are two or more guidelines that must we consider when deciding to use or not to use public data?

* One guideline to consider when scraping public data from the web is the legality of doing so. As we explored in the previous question, the legality of web scraping is uncertain so not fully understanding the existing laws or rules surrounding data scraping could end up being harmful. Additionally, another guideline to consider is where the data you are scraping is coming from. If scraping from an untrustworthy or suspicious source, you could be using bad data and it could end up harming your computer.