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.

I would suggest that the city of San Francisco should create a database of all the business licenses in the city. Then, whenever a new listing is created, the system can check if the business license number is valid. If it is, then the listing will be approved. If it is not, then the listing will be rejected. This would be a much more efficient system than manually checking the business license numbers. It would also be more accurate, as the system would be able to check if the business license is valid, rather than manually checking the website for the business license number.

Arguments against this system would be that it would be too difficult to create a database of all the business licenses in the city. It would also be difficult to keep this database updated. I disagree with these arguments, as I believe that this system would be relatively easy to implement. The city of San Francisco has a large IT department, and they could easily implement a database to store all the business licenses. The city could also update this database whenever a new business license is issued. This would be a much more efficient system than manually checking the business license numbers.

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 I could answer or explore using this data is the relationship between the quality of the housing and the cost of the housing. I could use the data to see if the quality of the housing is related to the cost of the housing. I could also use the data to see if the location of the housing is related to the cost of the housing. Shown in the data, is a graph showing home prices in the United States in 2021 by county. I could use this data to see if the cost of the housing is related to the location of the housing. I could also use this data to see if the quality of the housing is related to the location of the housing. I can also look at the total housing insecure households by state and see if the total housing insecure households by state is related to the cost of the housing.

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.**

One factor that I believe is important to consider when discussing the legality of web scraping is the purpose of the scraping. If the purpose of the scraping is to harm the website, then it is illegal. However, if the purpose of the scraping is to help the website, then it is legal. For example, if a website is experiencing a lot of traffic, then it is legal to scrape the website to help the website. However, if a website is experiencing a lot of traffic, then it is illegal to scrape the website to harm the website. The purpose of the scraping is important to consider when discussing the legality of web scraping because it determines whether the scraping is legal or not. The degree of protection of the website if not settled and will depend on the type of access made by the scraper, the amount of information scraped, and the purpose of the scraping.

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?**

1. The data should be publicly available. If the data is not publicly available, then it is not ethical to scrape it.
2. The data should be available in a format that can be easily scraped. If the data is not available in a format that can be easily scraped, then it is not ethical to scrape it.
3. The data should not be protected by a license that forbids scraping. If the data is protected by a license that forbids scraping, then it is not ethical to scrape it.
4. The data should not be protected by a license that requires permission to scrape. If the data is protected by a license that requires permission to scrape, then it is not ethical to scrape it without permission.

All of these guidelines are important to consider when deciding to use or not because if you do not follow them, then you are violating the privacy of the people whose data you are scraping.