Allana Tran

Project 2 Responses

a. If I was a software developer at either the San Francisco Planning Office or Airbnb, I would probably propose a system that involves having the two organizations working together to verify the validity of Airbnb rentals. For example, when a lister puts in their policy numbers it can be cross referenced to databases in the SFPO before listings can show up on Airbnb. I think this will allow for better accountability for all parties. Although I think this is a great idea, I think 2 arguments against it would be that it may take too long for verification (assuming there is wait time) and that will be harmful for Airbnb's business since no lister would want to wait to put up their listing. Another argument could be that sharing information with another organization may not be as easy and it can cost a lot if there has to be a totally new system that is compatible with both organizations.

b. I think with the type of data that we pulled from Airbnb we can look at housing affordability along with the corresponding number of bedrooms. Although the data we get will be different since when talking about housing affordability it usually is looking at rent and for Airbnb the price is per night and the listings are usually different from just regular housing units since they are like a hotel room. We could however cross reference different areas and see how the prices compare per night per area per bedroom as I'm sure that can give data scientists an insight into housing prices and affordability in an area as well as cost of living.

c. I think one factor that is important to consider when discussing the legality of web scraping is intention. What are the intentions of whoever is doing the web scraping? Like many other things that have grey areas, it's important to find out the intention of the action and what the effects are. For example, in the text it mentions someone using web scraping to auction snipe with bots which is unfair and negatively impacts others, whereas I believe that web scraping can definitely be used for good purposes that help people – such as collecting information from various sites to inform people about their decisions. Personally, I use a site that pulls information on video game prices from various websites. So I definitely think it comes down to the intent of web scraping and what it is used for.

d. I think two more factors on web scraping that we should consider include types of information as well as collection methods. Recently, a class I had talked about scraping data on social media platforms to sell to advertisers and this sort of topic came up. I think it's really important to consider the types of data being collected since there are definitely privacy concerns. A lot of people will post their personal information online including what they like, if they're dating, where they went to school, etc. So I think it's important to be mindful of what sort of data you are collecting about people and whether they can be identified and also what you will do with the data (such as companies selling user data without their knowledge :-( ). Data collection methods are also important for privacy concerns because I think it's also good to consider how you can collect data without compromising someone's privacy (being able to identify certain persons).