

1. Our idea for a slightly differing system would be to get a list from the San Francisco Planning Office that we would be able to loop through and cross reference each policy number to verify whether or not it is valid. One argument against our system is the static nature of this list. If the Planning Office were to send us a list, it may not reflect updated information as time goes on. New properties might be given licenses, and current properties may have their licenses revoked or expired in real time, so we would not be able to ensure that the list is up to date. Additionally, our system still fails to account for people who fraudulently classify their property as pending on the Airbnb website.
2. We could research the question: How does crime rate correlate to median home price in US counties? We could cross reference the data for each and compare!
3. One example we found in the Wikipedia article showcasing how web scraping may be harmful was the case with QVC and Resultly. Resultly search QVC's website for deals 200-300 times per minute which ultimately completely crashed QVC's website. This is a case where the web scraping may be illegal because it completely disrupted QVC's sales for a long period of time..
4. Web scrapers should always ensure that the data they are scraping is accurate. Often, especially in the presence of AI scrapers, scraped data is used to (deliberately or otherwise) spread false information. Furthermore, scrapers should know the conditions under which the data they scrape was collected. Many people volunteer their data to use an app but are not aware they are doing so. Some companies even change their privacy policies while users are already using the app. Web scrapers should know the source of their data, and if even a shred of evidence points to a company using unscrupulous data collection methods, they should avoid using the data.