0 Overview

Instructions

Due Date: Sunday May 4 @ 11:59PM

Total points: 30

Your written solution may be either handwritten and scanned, or typeset. Either way, you must produce a PDF that is legible and displays reasonably on a typical PDF reader. This PDF should be submitted via autolab as WA5. You may submit as many times as you like, but only your last submission will be graded and should include the entire submission. You should view your submission after you upload it to make sure that it is not corrupted or malformed. Submissions that are rotated, upside down, or that do not load will not receive credit. Illegible or incomplete submissions will lose credit depending on what can be read. Ensure that your final submission contains all pages.

You are responsible for making sure your submission went through successfully.

Written submissions may be turned in up to one day late for a 50% penalty.

No grace day usage is allowed.

1 Motivation

After working with health and voter record datasets, you've learned that publically available data can sometimes be deanonymized. Another use case we'll highlight briefly is using publically available tax data to deanonymize data on citations for various civic violations. In some cases, like with violations for not cutting your grass, this might not seem like a big deal. After all, if you're not mowing your lawn, your neighbors (i.e. the public) already know, and must have reported you in order to receive a violation. However, thanks to the scale and access to data that we currently experience, minor infractions can add up to major stigma.

Activists in Detroit have written an Equitable Open Data Report in response to this very problem. As they claim, open data can have positive outcomes like increasing government transparency on a local level, but it can also lead to harm for residents. As examples, they list how "blight tickets" can reinforce stigma about "bad" neighborhoods and negatively influence property values or contribute to red-lining practices; how block level crime incident reports can be deanonymized and subject crime victims to re-victimization through future harassment; and how attempts to improve Detroit with the "Open Issues" queue can actually lead to loss of property by allowing metal scrappers to pinpoint temporarily abandoned vehicles faster than previously possible.

Unfortunately, no similar Equitable Open Data Report or movement exists in Buffalo just yet. But that provides this class with an excellent learning opportunity. For this reflection, please review Detroit's Equitable Open Data Report and familiarize yourself with its general purpose and aims. Then, explore the data available at Open Data Buffalo (some of which you've already worked with in this class).

2 Questions

Once you've checked out both the open data report and Buffalo's publically available data, answer the following questions:

1. Do any of the data sets available at Open Data Buffalo carry the potential to cause harm to residents, especially if they can be deanonymized with other publically available data? And if so, what is the nature of the potential harm?
2. What, if any, changes could be made to reduce the risk of harm for residents?

Please remember that the purpose of a reflection is to carefully consider the causes and potential results of a problem. This is a thought experiment, and you will be graded on how deeply you considered the problem and the connections you make to the course materials. You are not expected to solve all potential problems that arise from open data everywhere. Just take some time to think about the principles outlined in Detroit's report and compare it with the level of accessibility in Buffalo.

Your reflection should be between 300-500 words in length. You are required to reference both Open Data Buffalo and the Equitable Open Data Report at least once. You do not need citations, as we are all working with the same resources. Simply note where the information came from in the body of your reflection. You are allowed to write the reflection on data you've already used for prior assignments, but keep in mind that simply reproducing statements already made on this (or previous) assignment sheets will not indicate very deep thinking on your part. Excellent reflections will go beyond the surface level observations listed here and bring their own perspectives into consideration.

3 Revision History/Credits

- Summer 2021 - Initial project draft developed as part of the Mozilla Responsible Computer Science Challenge by:
  - Joshua Caskie (jmcaskie@buffalo.edu)
  - Alexander Fernandez (adfernan@buffalo.edu)
  - Garegin Grigoryan (grigoryan@alfred.edu)
  - Andrew Hughes (ahughes6@buffalo.edu)
  - Macy McDonald (macymcdo@buffalo.edu)
- Fall 2021 - Adapted for CSE-250 by Oliver Kennedy (okennedy@buffalo.edu)