Background and Motivation

Although none of us has a specific prior interest or involvement in this topic,
all 3 group members found it fascinating and were impressed with the depth of
the available data. We read new articles about a recent surge in heroin's popularity
in the United States, and wanted to explore the geographic and demographic areas
in which this was most prevalent.

Project Objectives

The questions we are trying to answer are:
1. Where is heroin use most prevalent in the United States?
2. What is the relationship between heroin use and income level?
3. What is the relationship between heroin use and other criminal activity?
4. How do the answers to the above questions change as we filter by
age, race, and sex?

The motivation is to gain a better understanding of who is using heroin in
America today.

Data

Our dataset is from the National Survey on Drug Use and Health, 2012. The
survey was conducted by the US Department of Health and Human Services. The
data, which is available for free download, can be found by following this link:

http://www.icpsr.umich.edu/icpsrweb/ICPSR/studies/34933

Data Processing

The data cannot be downloaded in JSON form, so we expect to convert it. The final
form of our data will be an array of JSON objects representing individuals, with
keys for age, sex, race, state, heroin use, income level, and criminal activity.
Due to the massive size of the dataset, we will also need to filter it sustantially.
Data processing will be the first step in our project.

Data Display

We plan on using several interactive views to illustrate the interplay among a number of
demographic relationships for heroin users, including
gender, age, socioeconomic status, and race. The views we are currently considering include: (1)
a map of the United States, which shows hotspots

of cocaine use via a gradient scale, (2) a bar chart relating-- more specifically-- socioeconomic status and heroin use, (3) a pie chart that

depicts the type/severity of criminal activity for the currently selected demographics of heroin users, and (4) a New York Times connected

news feed that provides headlines re: national heroin use or specific state heroin use if a state is selected on the map (incorporating NYT Linked Data Application into our app)

So, currently all of our views are connected, which we feel is important in illustrating who is using heroin in the US today and how environments

have affected that use.

Must-Have Features

The above views are all ones that we feel are necessary for our project.

Optional Features

We are currently trying to figure out what other views (possibly in the form of scatterplots) could be useful in helping illustrate relationships

about heroin use that will be helpful to our audience (in so far as additional views would not be extraneous).

Project Schedule

Proposal due Friday, April 3
April 10, Cleaned up and processed dataset ready for filtering/incorporation into views, as well as research on NYT linked application within our website
April 17, Project Milestone with our core working views
April 24, Inclusion of Optional features
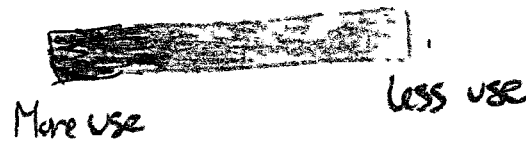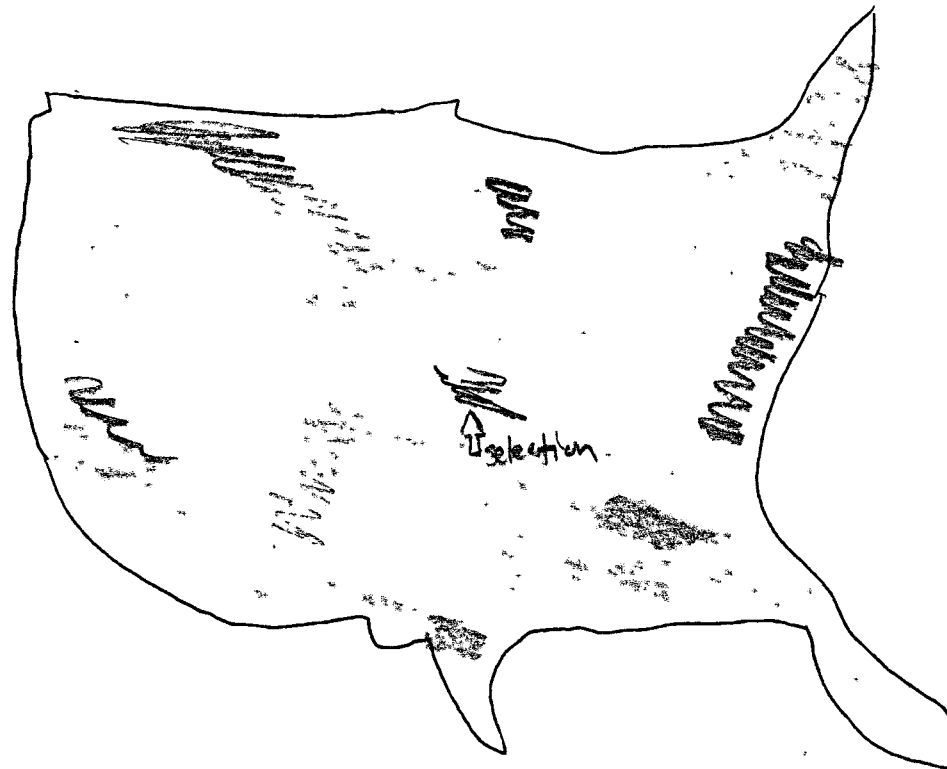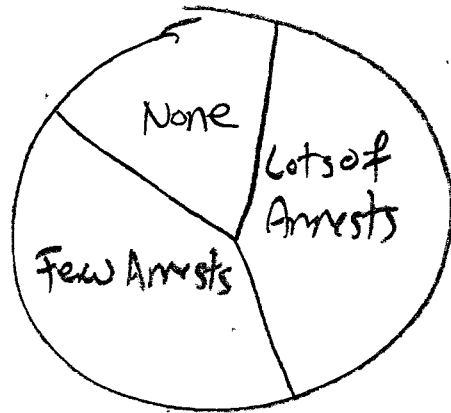May 5, Project ready for submission
(Throughout work process, update Process Book)

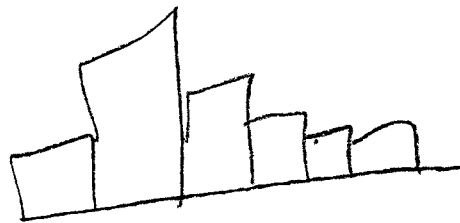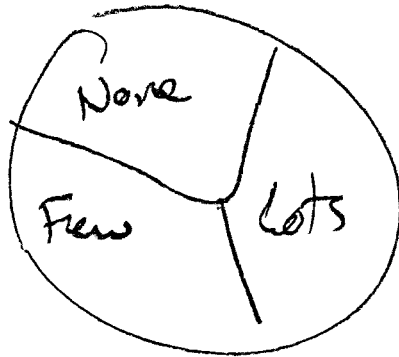National Survey on Drug Use and Health, 2012
www.icpsr.umich.edu

CRIME

None

Lots of Arrests

Few Arrests

Income ⟶

selection

More use   Less use

- o ——— ———
- o ——— ———
- o ——— ———
- o ——— ———

*crime, income + headlines based on selection

TIME

CRIME

None

Few   Lots
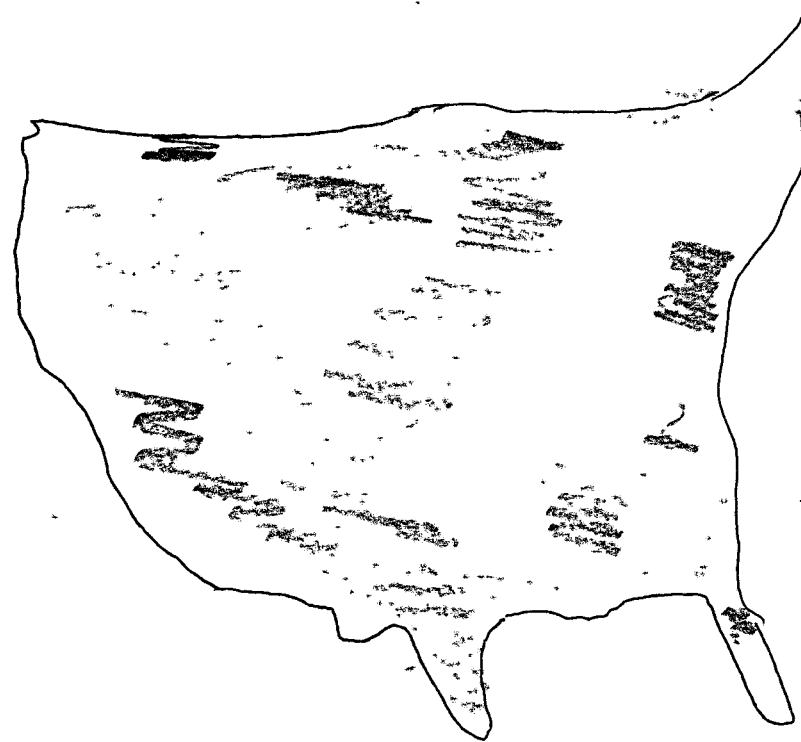
Filter By Sex
Male

Filter By Age
18-25

Income →

Usage Rate OR Crime

Population Density of Different Centres
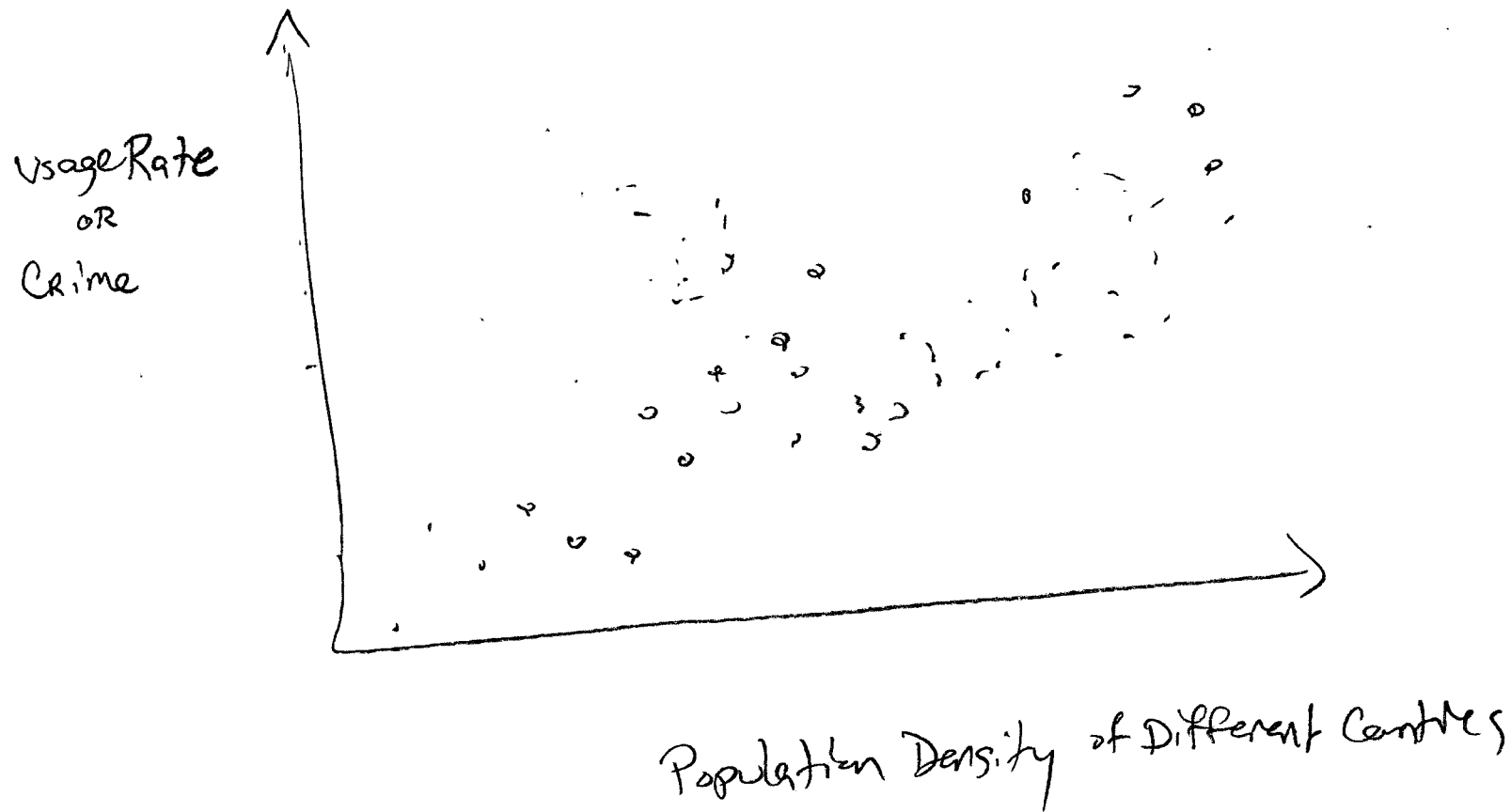
* Note, this would probably require separate census data, which is why it's not one of our core views