

# CRIME OVER TIME

An Investigation of Crime in the US from 1960 - 2014

Jacob Kapan and Kit Zellerbach

Process Book

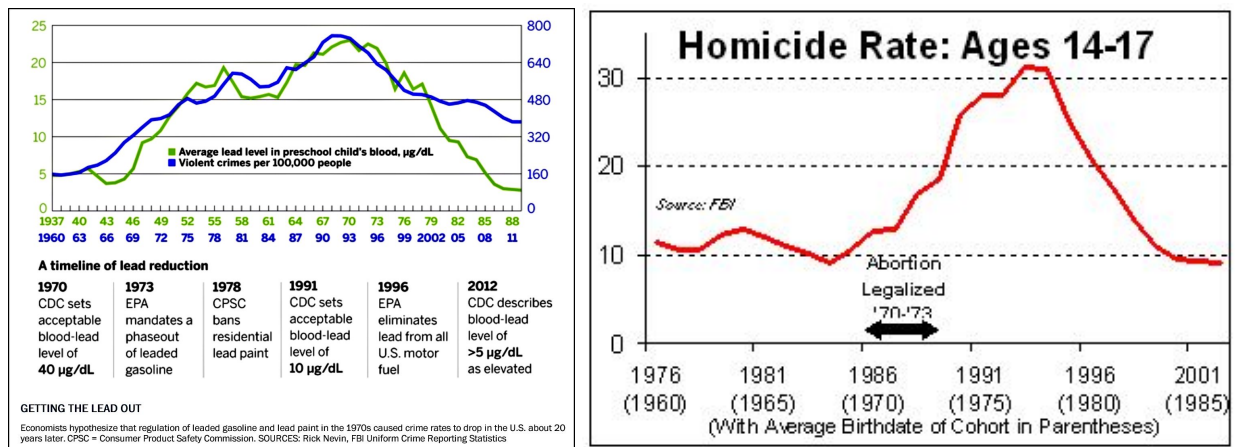
February 28th 2019

## Overview and Motivation

It is interesting to see what external factors drive human behavior, especially when it drives someone to do something drastic (commit crimes). We want to create a visualization which allows users to explore the co-occurrence of reduced crime rates with birth control and unleaded gasoline.

## Related Work

There are several articles that explore the lead-crime hypotheses and the legalized abortion-crime effect. We were inspired by this phenomenon to create an interactive visualization that would help users explore these.



We were also so impressed by the visualization on MBTA data (<http://mbtaviz.github.io/>) and wanted to make something that looked even just a little bit as good.

## Questions

What variables affect crime rates in the US?

How easy is it to detect correlation?

Are results immediate?

## Data

The government and local police departments put historical records online. Likewise, there are good timelines established for TEL being phased out, and for birth control (and abortion) being phased in.

Our main data sources were:

- <https://www.ucrdatatool.gov/Search/Crime/Local/LocalCrime.cfm>
- <https://www.ucrdatatool.gov/Search/Crime/State/StatebyState.cfm>
- <http://www.johnstonsarchive.net/policy/abortion/graphusabrate.html>
- <https://www.eia.gov/state/seds/>
- <https://www.guttmacher.org/>

Data always needs to be cleaned up and transformed. This project was no exception. Many of the data files required aggregation by US state. Also, some files were so big that we couldn't load them without crashing the file reader.

```
measure_name,datum,state_id,state_name,first_year
Abortion rate (the no. of abortions per 1000 women aged 15-44) by state of residence,20.7,AL,Alabama,1979
Abortion rate (the no. of abortions per 1000 women aged 15-44) by state of residence,23.6,AK,Alaska,1979
Abortion rate (the no. of abortions per 1000 women aged 15-44) by state of residence,24.4,AZ,Arizona,1979
Abortion rate (the no. of abortions per 1000 women aged 15-44) by state of residence,15.6,AR,Arkansas,1979
Abortion rate (the no. of abortions per 1000 women aged 15-44) by state of residence,44.1,CA,California,1979
Abortion rate (the no. of abortions per 1000 women aged 15-44) by state of residence,29.3,CO,Colorado,1979
Abortion rate (the no. of abortions per 1000 women aged 15-44) by state of residence,28.6,CT,Connecticut,1979
Abortion rate (the no. of abortions per 1000 women aged 15-44) by state of residence,26.3,DE,Delaware,1979
Abortion rate (the no. of abortions per 1000 women aged 15-44) by state of residence,83.3,DC,District of Co,1979
Abortion rate (the no. of abortions per 1000 women aged 15-44) by state of residence,33.8,FL,Florida,1979
Abortion rate (the no. of abortions per 1000 women aged 15-44) by state of residence,27.8,GA,Georgia,1979
```

```
{"Year":1961,"Population":3302000,"Violent crime total":5564,"Murder and nonnegligent Manslaughter":427,"
{"Year":1962,"Population":3358000,"Violent crime total":5283,"Murder and nonnegligent Manslaughter":316,"
{"Year":1963,"Population":3347000,"Violent crime total":6115,"Murder and nonnegligent Manslaughter":340,"
{"Year":1964,"Population":3407000,"Violent crime total":7260,"Murder and nonnegligent Manslaughter":316,"
{"Year":1965,"Population":3462000,"Violent crime total":6916,"Murder and nonnegligent Manslaughter":395,"
{"Year":1966,"Population":3517000,"Violent crime total":8098,"Murder and nonnegligent Manslaughter":384,"
{"Year":1967,"Population":3540000,"Violent crime total":8448,"Murder and nonnegligent Manslaughter":415,"
{"Year":1968,"Population":3566000,"Violent crime total":8280,"Murder and nonnegligent Manslaughter":421,"
{"Year":1969,"Population":3531000,"Violent crime total":8842,"Murder and nonnegligent Manslaughter":485,"
{"Year":1970,"Population":3444165,"Violent crime total":10185,"Murder and nonnegligent Manslaughter":404,"
{"Year":1971,"Population":3479000,"Violent crime total":10835,"Murder and nonnegligent Manslaughter":524,"
{"Year":1972,"Population":3510000,"Violent crime total":10994,"Murder and nonnegligent Manslaughter":496,"
{"Year":1973,"Population":3539000,"Violent crime total":12390,"Murder and nonnegligent Manslaughter":468,"
{"Year":1974,"Population":3577000,"Violent crime total":13338,"Murder and nonnegligent Manslaughter":536,"
{"Year":1975,"Population":3614000,"Violent crime total":14201,"Murder and nonnegligent Manslaughter":577,"
{"Year":1976,"Population":3665000,"Violent crime total":14248,"Murder and nonnegligent Manslaughter":553,"
{"Year":1977,"Population":3690000,"Violent crime total":15293,"Murder and nonnegligent Manslaughter":524,"
{"Year":1978,"Population":3742000,"Violent crime total":15682,"Murder and nonnegligent Manslaughter":499,"
{"Year":1979,"Population":3769000,"Violent crime total":15578,"Murder and nonnegligent Manslaughter":496,"
{"Year":1980,"Population":3861466,"Violent crime total":17320,"Murder and nonnegligent Manslaughter":509,"
{"Year":1981,"Population":3916000,"Violent crime total":18423,"Murder and nonnegligent Manslaughter":465,"
{"Year":1982,"Population":3943000,"Violent crime total":17653,"Murder and nonnegligent Manslaughter":417,"
{"Year":1983,"Population":3959000,"Violent crime total":16471,"Murder and nonnegligent Manslaughter":364,"
{"Year":1984,"Population":3990000,"Violent crime total":17204,"Murder and nonnegligent Manslaughter":374,"
```

In order to calculate the lead in gas values, we had to eyeball lead levels from a graph since we couldn't find any tabular data online.

```

"state": "AK", "msn": "AVTXP", "1960": 1032.00, "1961": 1276.00, "1962": 1030.00, "1963": 648.00,
"state": "AL", "msn": "AVTXP", "1960": 280.00, "1961": 292.00, "1962": 321.00, "1963": 373.00,
"state": "AR", "msn": "AVTXP", "1960": 177.00, "1961": 165.00, "1962": 347.00, "1963": 400.00,
"state": "AZ", "msn": "AVTXP", "1960": 699.00, "1961": 460.00, "1962": 748.00, "1963": 673.00,
"state": "CA", "msn": "AVTXP", "1960": 5383.00, "1961": 3438.00, "1962": 4943.00, "1963": 4366.00,
"state": "CO", "msn": "AVTXP", "1960": 1125.00, "1961": 797.00, "1962": 1457.00, "1963": 1335.00,
"state": "CT", "msn": "AVTXP", "1960": 104.00, "1961": 103.00, "1962": 124.00, "1963": 172.00,
"state": "DC", "msn": "AVTXP", "1960": 0.00, "1961": 0.00, "1962": 0.00, "1963": 0.00, "1964": 0.00,
"state": "DE", "msn": "AVTXP", "1960": 19.00, "1961": 18.00, "1962": 42.00, "1963": 69.00, "1964": 100.00,
"state": "FL", "msn": "AVTXP", "1960": 4517.00, "1961": 3926.00, "1962": 5459.00, "1963": 5580.00,
"state": "GA", "msn": "AVTXP", "1960": 262.00, "1961": 221.00, "1962": 424.00, "1963": 494.00,
"state": "HI", "msn": "AVTXP", "1960": 2640.00, "1961": 3068.00, "1962": 1852.00, "1963": 1100.00,
"state": "IA", "msn": "AVTXP", "1960": 366.00, "1961": 225.00, "1962": 328.00, "1963": 315.00,
"state": "ID", "msn": "AVTXP", "1960": 133.00, "1961": 124.00, "1962": 211.00, "1963": 233.00,
"state": "IL", "msn": "AVTXP", "1960": 3733.00, "1961": 2488.00, "1962": 819.00, "1963": 546.00,

```

## Exploratory Data Analysis

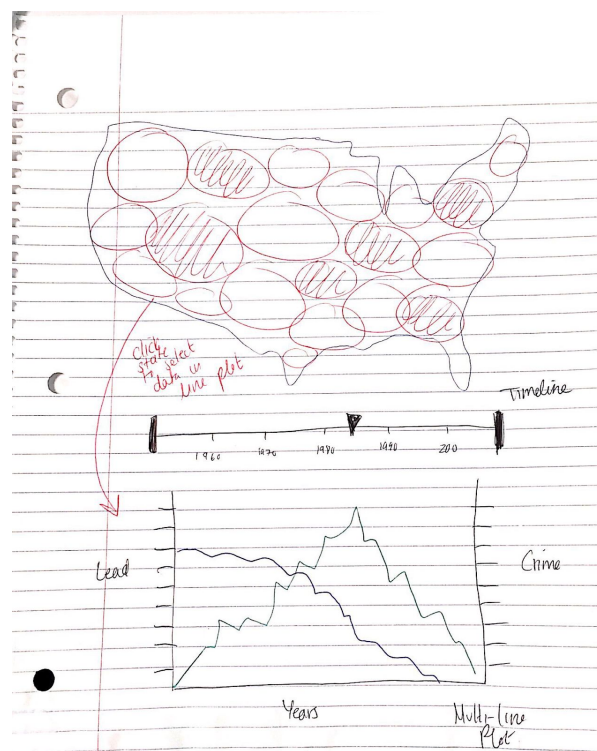
We began plotting crime on a scatter plot to look at the general trend. Then, we quickly moved on to testing prototypes of our design ideas. Some of the websites where we obtained data offered online chart creator tools, which we could use for a first glance at the data. We began the project knowing that reduction of lead levels and the legalization of abortion affected crime rates, but only until we looked at the data did we realise there was a 20 year process in-between.

## Design Evolution

The first designs started out on paper.

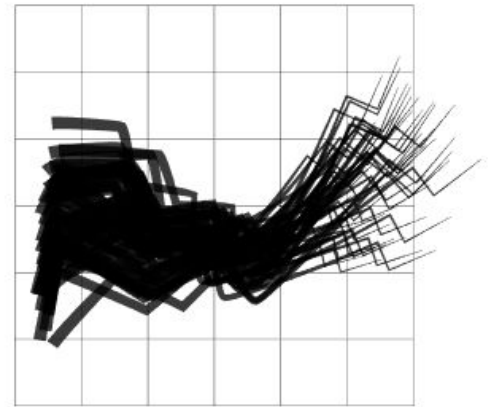
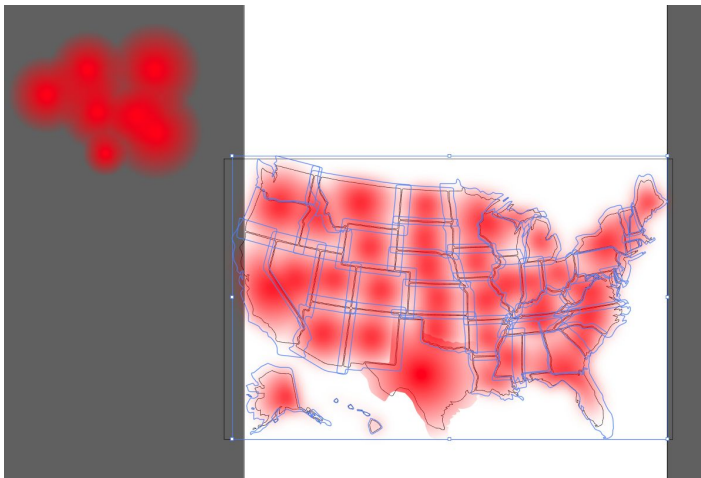
We knew we wanted a map that would have an interactive timeline, as well as other multi-views to complement the map and allow for a more in-depth exploration of the data.

Since we were warned to stay away from choropleth maps, we looked at alternative map



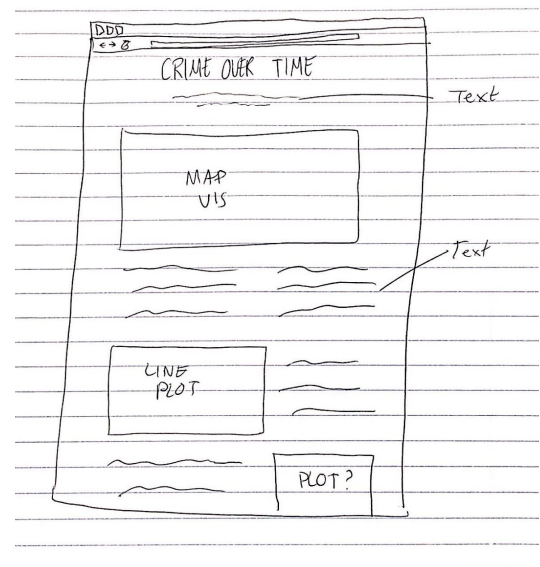


options, but weren't satisfied with any. Therefore, we continued to go with a choropleth map and decided to use rate as an attribute to help combat some pitfalls. Then we transferred our ideas over to Adobe Illustrator.



These images show the design process on Illustrator, where we made sure that the design looked good before we implemented it.

We also wanted to make the website similar to an article with lots of relevant text and so sketched out some potential layouts. We were going for something similar to a NY Times article. We tried out different fonts and styles until we found a consistent theme we were satisfied with. We looked at alternative text placements and debated whether the charts should take up the width of the screen, or if smaller charts to the left/right of the text were also acceptable. Eventually, after trying out several options we came to a decision that we could both agree on.



Also, initially we were going to have it so that when the user hovers over a state on the map, the corresponding state line on the multi-line plot is highlighted. However, we realized that both views are rather far apart and that it didn't feel necessary to implement that feature. It would have taken a couple of minutes to add that functionality, but as a design choice, we decided to keep the views unlinked.

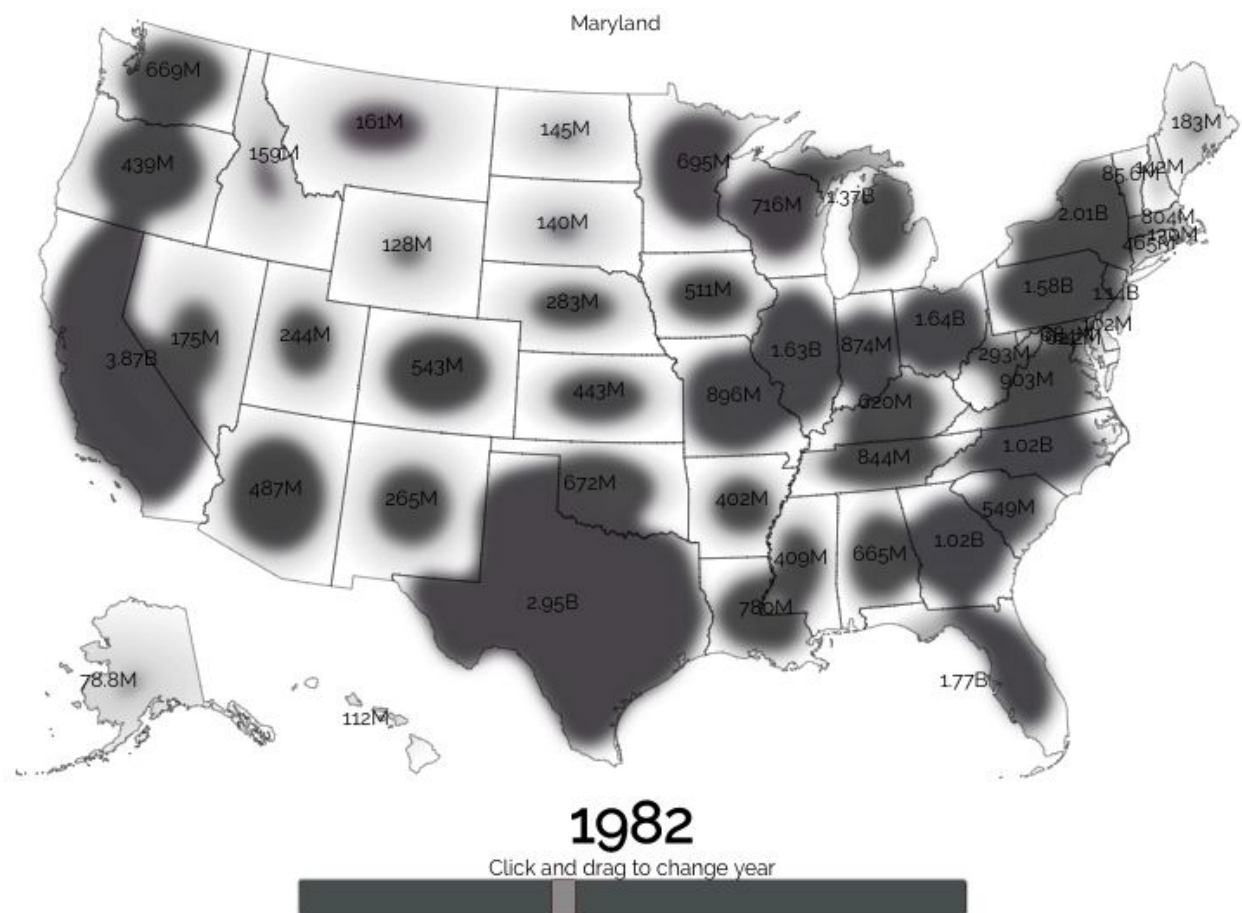
The cloud map also went through several iterations. Different blur matrices and opacity functions were created to make the clouds move in very specific ways. Namely, the idea was to have “gooey” movements, where clouds would spill over into other the other states. This took several attempts at tweaking gradients and svg image effects.

## Implementation

We ended up implementing the website using d3.js and React. This would allow us to structure our website in structured modules.

The first visualization is a cloud map of the US. The size and darkness of the clouds is relative to the  $\log_{10}$  of how many grams of lead were purchased (in the form of gasoline) for motor vehicles (M and B for M/Billion grams, respectively).

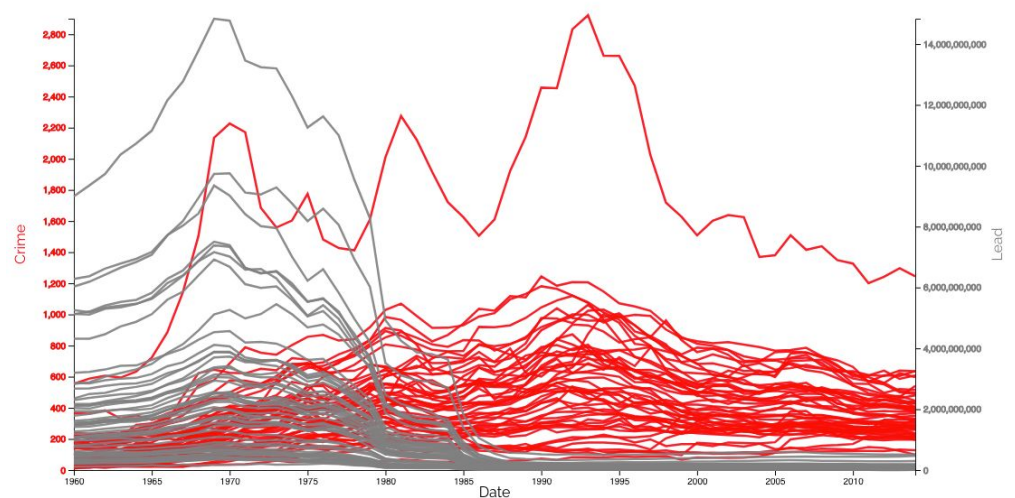
Our formula is  $\log_{10}(\text{state.year}/\text{US.allTimeMin})/\log_{10}(\text{US.allTimeMax}/\text{US.allTimeMin})$ . We chose this formula, as our values have a great range. Our data takes into account the different limits of grams of lead per gallon of fuel through the years, as well as how much of the purchased gasoline was leaded vs. unleaded. The user can click and drag the bar below the graph to see how the clouds grow and shrink through the years. Hovering over a state will also display its name at the top, as seen below with Maryland.

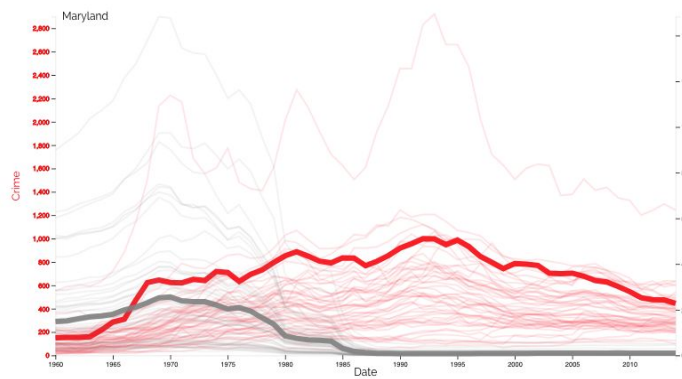


Below the cloud map is a Multiline Chart that compares crime and lead consumption for each state over the span of 55 years. Crime rates are in red and lead rates are in grey. The crime rate is per 100,000 people, while the lead rate has been calculated as mentioned above.

The chart has some extra functionality. Hovering over any line will highlight both crime and lead lines for that particular state.

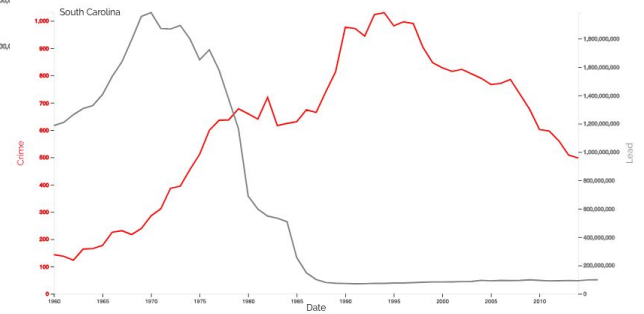
Furthermore, clicking any line will remove the lines from all the





other states and rescale the axis to fit the lines from the selected state.

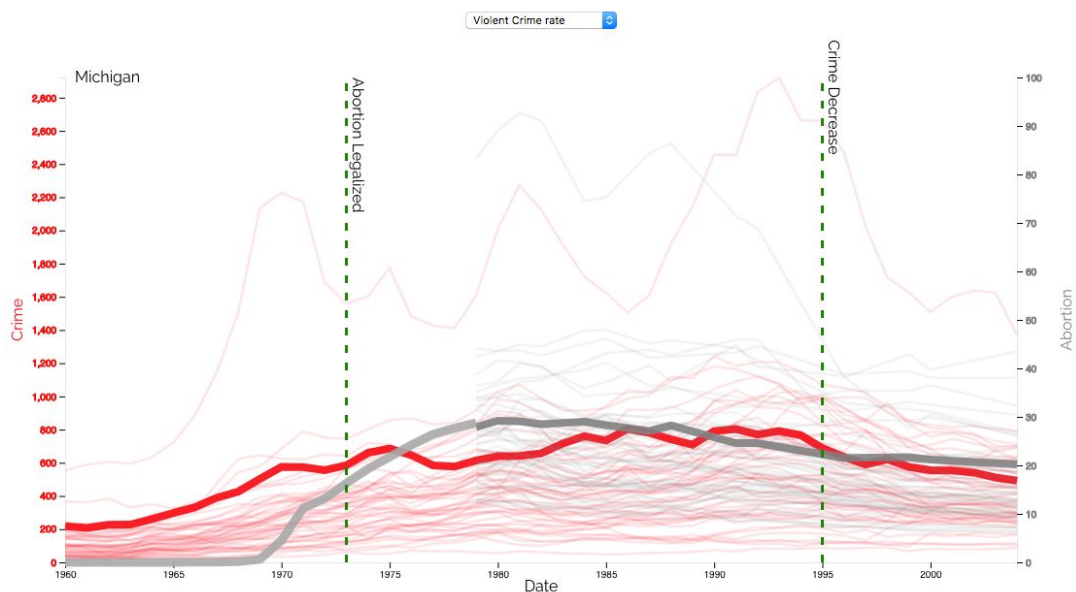
The graph on the left is the result of hovering, while the graph below is the result of clicking on a state line.



Next, we wanted to implement data visualizations on the effect of the legalization of abortion. We created another multi-line plot to show the crime rate against the abortion rate per state.

Crime rates are in red and abortion rates are in grey. The crime rate is per 100,000 people, while the abortion rate is the number of abortions per 1000 women aged 15-44 by state of residence.

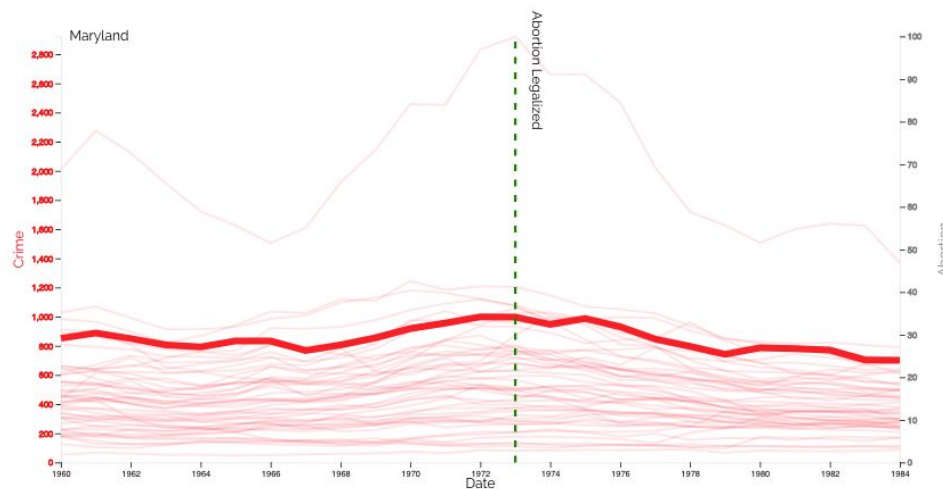
The plot also has two extra vertical green lines. The one on the left represents the year that abortion was legalized, while the line on the right marks 18 years later and highlights the decrease in crime rate at that time. There is also a dropdown box that allows the user to select which type of crime they want to plot in the visualization.





There is a rift in the data at the 1979 mark since there was no data on abortion rates in individual states before then. However, there was data on US total abortion rates from 1960 to 1979. Therefore, we decided to include this data as a singular line so that the overall increase in abortion rates is understood, as 1970 to 1979 have the highest rate of change.

The last feature is double-clicking on the visualization. When the user double-clicks, the crime rate lines are shifted back 18 years to show the exact change in crime rates at the 18-year mark. The abortion rate lines are removed for clarity.



## Evaluation

We learned a lot about crime from these visualizations. We were able to clearly see the correlation between lead levels, abortion legalization, and crime. We also gained a deeper understanding of what external factors can cause humans to commit crimes.

We think our visualization works well to explore some hypothesized explanations of the drop in US crime rates. We definitely answered the questions we were exploring. However, there is always room for further improvement. Especially as there are other factors that affect crime rates that could be added and explored. Additionally, the cloud map could be improved on and detailed with crime overlays.