

Process Book

U.S. Campus Safety

Link: <https://danielwwong.github.io/DataVisFinal/>

Team:

Name	Email	GitHub ID
Zhanfeng Huang	zhuang4@wpi.edu	danielwwong
Qiankun Li	qli6@wpi.edu	Qiankun0103
Yujun Tian	ytian5@wpi.edu	duoshyyigirl

- Overview and Motivation

Motivation:

When people select college, campus safety is a very important factor to consider. When we foreign students planned to study abroad, we really want to know how safe the college is, the states is. Fortunately, we found data from *Campus Safety and Security*, which provides detail US college crime data.

Overview:

We tried to plot the data by two parts. First, the overview of our map is an US map, surrounded by many bar charts. The US states map are marked by color code, when the states are marked by darker color, it means that the states are less safe than states with light color. When viewers mouse over specific state, the corresponding state would be highlighted. If viewers further click it, the map would shrink and detail information about the state would show by bar charts below. The bar charts will draw all the states information as an x axis, and if viewers hover the mouse over specific state in the map, the corresponding bar would be highlighted. The bar charts would cover information including gender rate, number of crime and schools' attributes. Second, we planned to add another part, which can provide us with several filters. The filter contains gender rate, years, whether the school is private or public, how many years the school offers, states choice. Once the user has made his choice, the US city map would map top 5 safest and top 5 dangerous schools' place on the US map. We want to help parents and students to study in their satisfied collage. And hope this website could provide useful information.

- Related Work

We watched a TV program, it said there have 35 Chinese students died on America in 2016. It shocked us, so we decided to study the college crime data.

In light of visualization part, professor in class introduced interactive multiply chart, which made us very interested. When user click the big map, it shrinks to a corner, and present three bar charts near it. These bar charts can interact together. When hover your mouse on a particular state, all the bar charts will highlight the specific bar of the state.

- Questions

- Which area in the U.S. has more campus crime records than others?
- Can I know more details about the school which has crime records?(e.g gender ratio, institution's attributes)
- Can I filter out the institutions by state or ideal gender ratio?

- Data

Source: <https://ope.ed.gov/campussafety/#/datafile/list>

Data scraping:

As so far, we used the “noncampuscrime121314.xls” sheet.

Data cleanup:

We have done substantial data cleanup, but we may do more in future. What we have done include:

1. Drop all null value records by Python’s pandas.

2. Zip code subtract by “awk” command:

In the original data, the record in Zip code column includes 8 digits, 9 digits, 5 digits and 4 digits. To make it corresponding to our “cities.json” file, we subtract first 5 digits of the 9digits zip, first 4 digits of the 8 digits zip, and add leading 0 to the 4 digits. We achieved this by “awk” command.

3. Split one column into two by Python’s pandas:

The “Sector_desc” column used to contain information like “public, 4 year or above”. We think this can convey two kinds of information, so we split it into two columns, one shows “public” and another one shows years.

4. Add the information up to states information by Python’s pandas:

Since not all the schools have crime records, so we aggregated all the schools’ information up by states’ categories. Thus, we could present states information in the data visualization.

- **Exploratory Data Analysis**

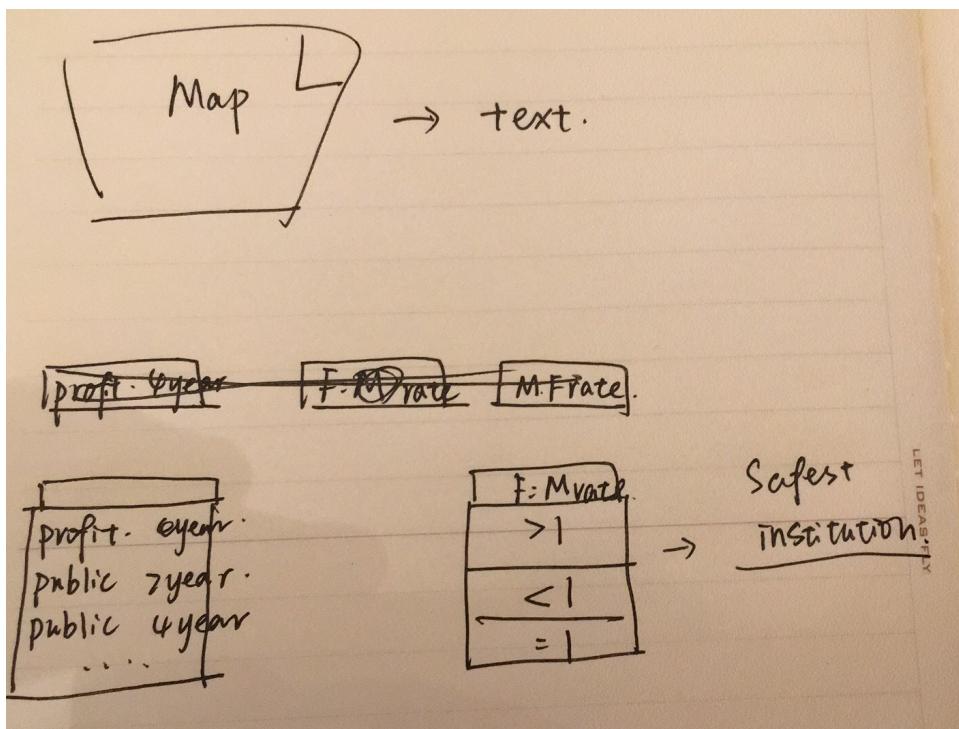
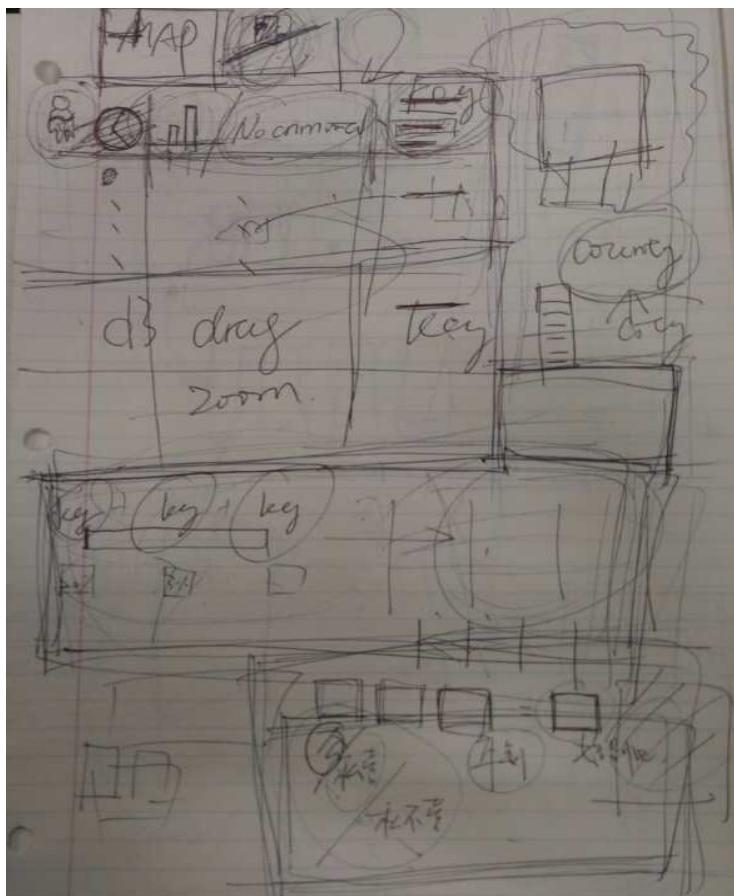
Initially, at the first sight of the data, we tried to map all the college information on to the U.S map. But in light of the number of the colleges of U.S, all the points will be assembled, which was a mess. So, we further informed our design by categorizing the data by states. States data offer benefits, such as crime numbers are more obvious, the whole overview is more concise and beautiful. This is why we had the map as states information. Also, viewers may show curiosity to which school is safer or how dangerous it is. From that point, we would provide a filter to satisfy viewers need of checking the school's information.

- **Design Evolution**

When the mouse click the state, the color could be changed It is convenient to observe. At the same time, the map would shrink in the top of left. When hover the mouse, the bar chart would interactive with the state. For example, when the mouse over the MA, the bar which indicates the MA states' data would be highlighted. And our filter is very useful, it could search safety collage very fast in the different state. In our visualizations, most of them are bar chart, because the bar chart is straightforward and easier to compare. We pay more attention to the color, we want to design the image more beautiful, which would attract user eyes. Because we change a new data, and spent a lot of time to find this campus safety data, but now our achievement is not deviate second proposal.

- **Implementation**

We hope our project could help people to find a satisfied college, provide useful information and convenience to select safety college.



[Overview](#)[Campus Information](#)[Campus Choice](#)[Team](#)

U.S. Campus Security

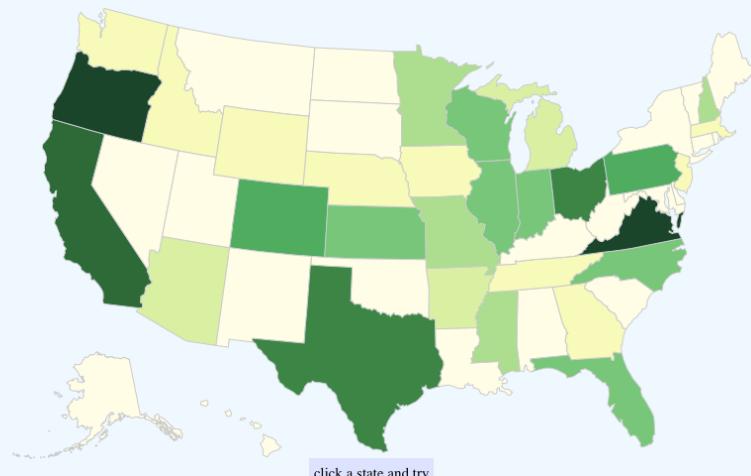
When people consider which college to attend, safety is always the great concern. By a statistical report, it says there were 36 Chinese students died in U.S at the year of 2016. Given the data of campus crimes' records in the year from 2012 to 2014, we summed up all the different types' crime records in each campus, and grouped it by states.

In the visualization work, we will provide two parts. In the first part, viewers can have an overview of the U.S. map. Darker color represents more serious crime rates. In the second part, we further offered filter function, so that viewers could choose their own best campuses.

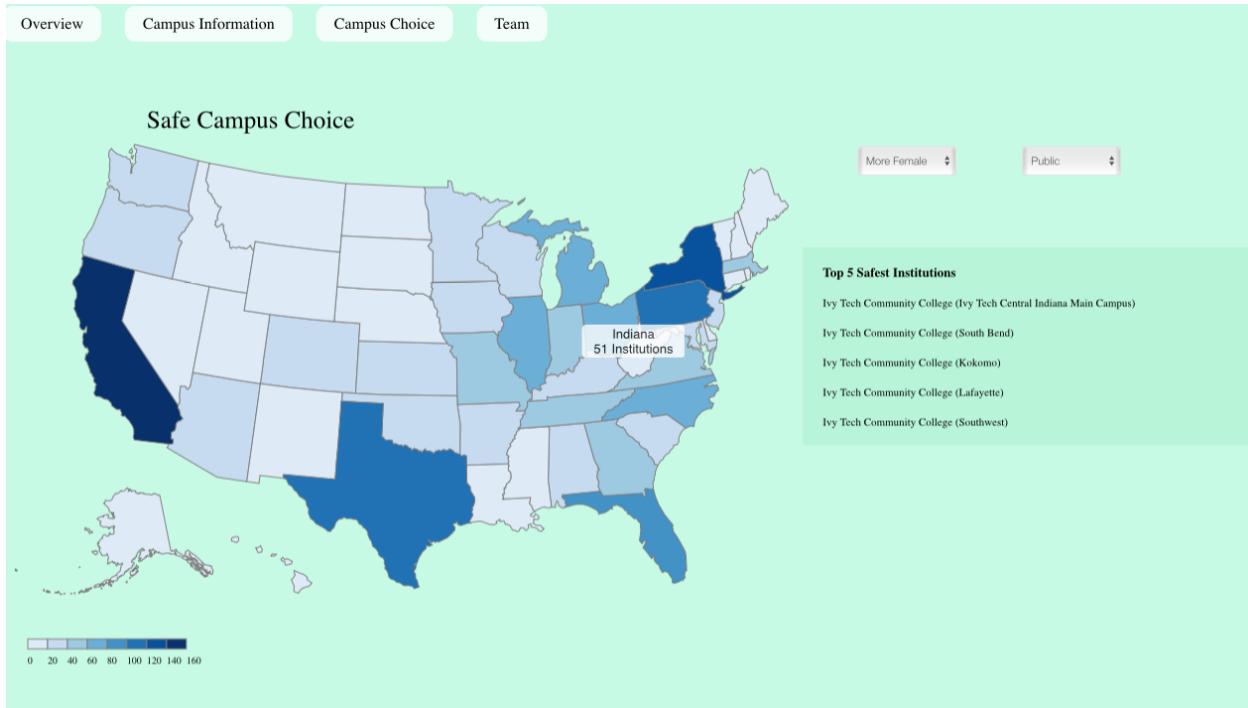
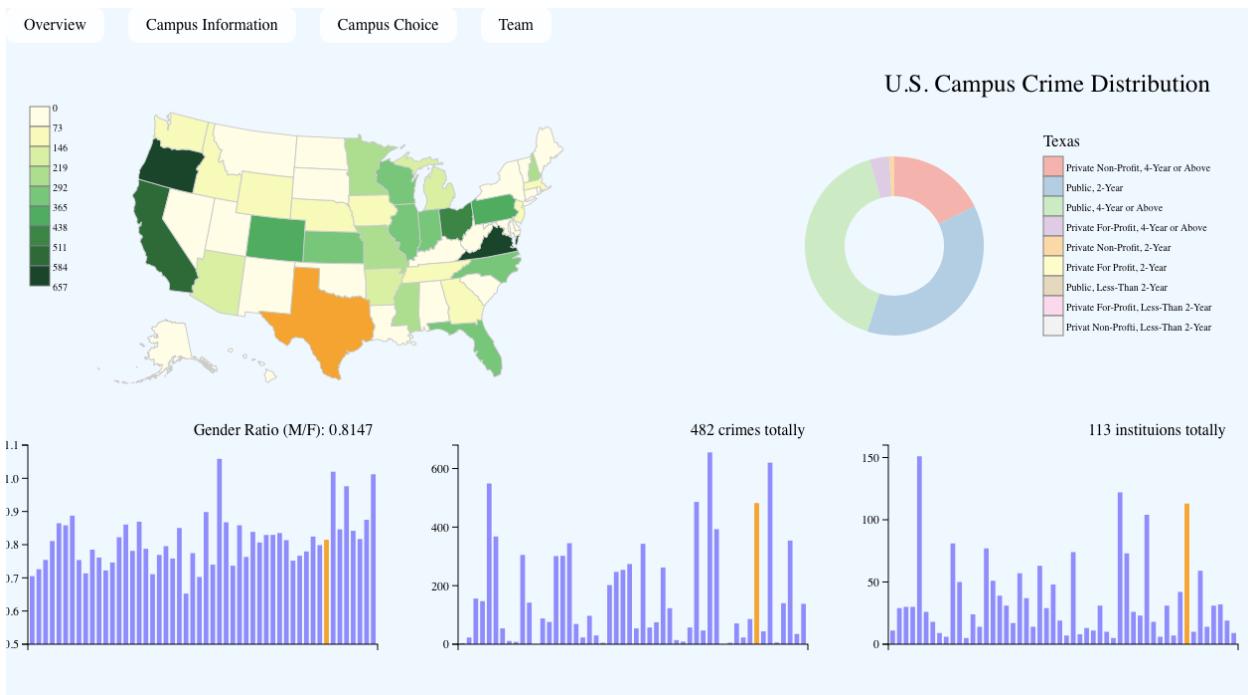
Enjoy your safe campus life!

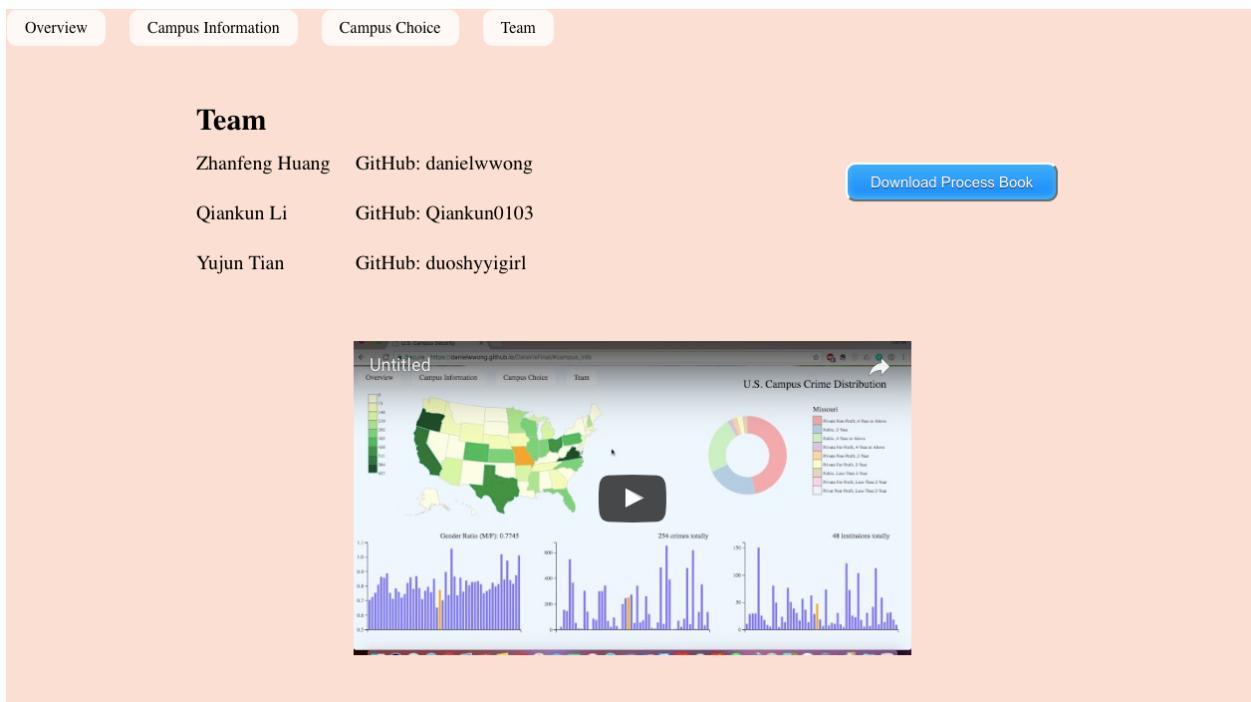
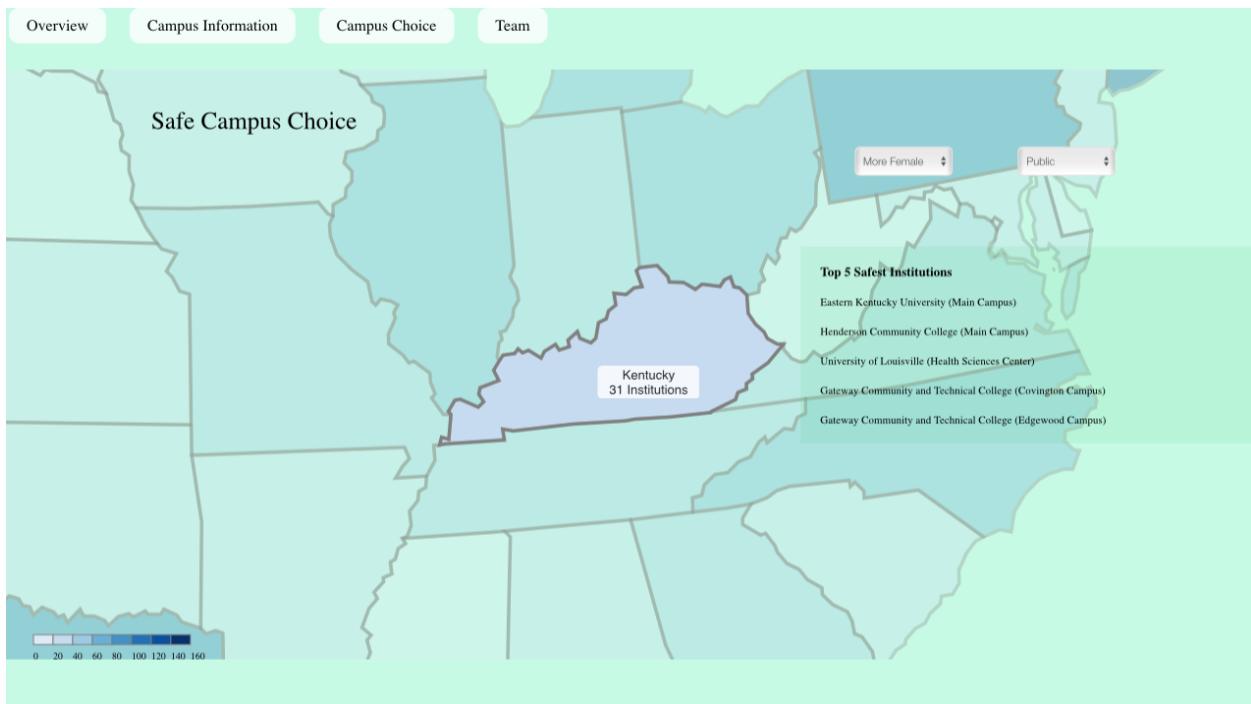
[Overview](#)[Campus Information](#)[Campus Choice](#)[Team](#)

U.S. Campus Crime Distribution



click a state and try





- [Evaluation](#)

- Which area in the U.S. has more campus crime records than others?

According to our data visual work, it is pretty obvious both in the “Campus Information” and “Campus Choice” parts that darker colored area has more campus crimes than the lighter ones. More serious campus crimes appeared in Virginia, Oregon, California, Texas and etc.

And from the “Campus Information” part, we can conclude that campus crime difference is huge between areas, since it can be up to 655 records in Oregon, and the least crime records owner, Alaska, just has two crimes happened as a sum of crimes from 2012 to 2014.

- Can I know more details about the school which has crime records?(e.g gender ratio, institution's attributes)

Definitely yes from our data visualization result. The “Campus Information” part offers details about the state’s campus data. After having chosen a state, viewers can see the all the college students’ gender ratio in that state, what types of school that state has and the proportion issue and most importantly, the total campus crime records in that state from 2012 to 2014.

- Can I filter out the institutions by state or ideal gender ratio?

We designed to give a name list of top safest institutions for students, where “the safest” means least campus crime records in that institutions occurred. Our “Campus Choice” part is designed to give viewers selection facility to filter out the top 5 safest institutions given state’s name, sex ratio and school’s attributes.

Even if viewer is indecisive, which means he hasn’t made up his mind where to go, what kind of school he wants to choose, just hovering his mouse over the map, the name list of top 5 safest institutions would also show up. Also, how many institutions that state owns would also be presented with sliding the mouse.

- What should we improve?

Our data story based on the show part of crime distribution in the U.S. state map. It could gives people clear idea of how it looks like, but didn’t dig out the reason.

So we wants to improve from the reason analysis, say, why Virgnia has such high campus crime records. We will try to connect this dataset with the U.S. GDP, unemployment rate in different state, since we assumed that great difference of

citizens' income and high unemployment would lead to high crime rate. Moreover, we will map institutions by cities' location, so more detail information will provide.