

CSCI572 HW3 Report Team 17

1. How easy to use was D3?

D3 is actually a simple but powerful library to represent the frontend interface. Although we are not completely familiar with all the functions and feature in D3, we still can easily make many beatiful charts and graphs by referencing online D3 examples to represent our data result. In sum, it's easy to show your data in a pretty way, but if you would like to customize your own graph to show your data perfectly, you have to learn the way how D3 draw graph.

2. How easy to use was Banana?

The easiest part to use Banana is that it can get the connection to solr automatically, which means we don't need to deal with data since solr guarantee its functionality. We can simply add useful panels and find the useful information to get. Banana is an AngularJS-based framework and can also integrate with D3 code. Using the MVC structure may be a little bit complicated during the first time of leveraging. It, however, provides a standard way to create customer panels.

The only problem with using Banana is when the data format using by the panels is different from the original data inherited from solr. For example, when we want to show the point of document in the us map, we find that the data format in the map panel only accept "2-letter state code". However, the only accurate location our source is the latitude and longitude. In this case, we need to manage the data format in the early beginning again, put the state code into json by retrieving coordinate and map it to state location. Besides this there's no such big issue to deal with in showing and illustrate the information in the data by Banana.

3. How easy was it to use FacetView?

FacetView is the easiest visualization tool that we use in this project. Most of components have already been covered in the installation package, few codes that we have to revise in index.html are:

1. Connect FecetView to Solr:

Change "search_url" to our local solr server and set "search_index" to "solr"

2. Set the classification:

In this part, we will classify all queried results to keyword, manufacturer, and weapon category.

3. Show what to display in the description area

In the description area, we will show title, time, detailed description for each result.

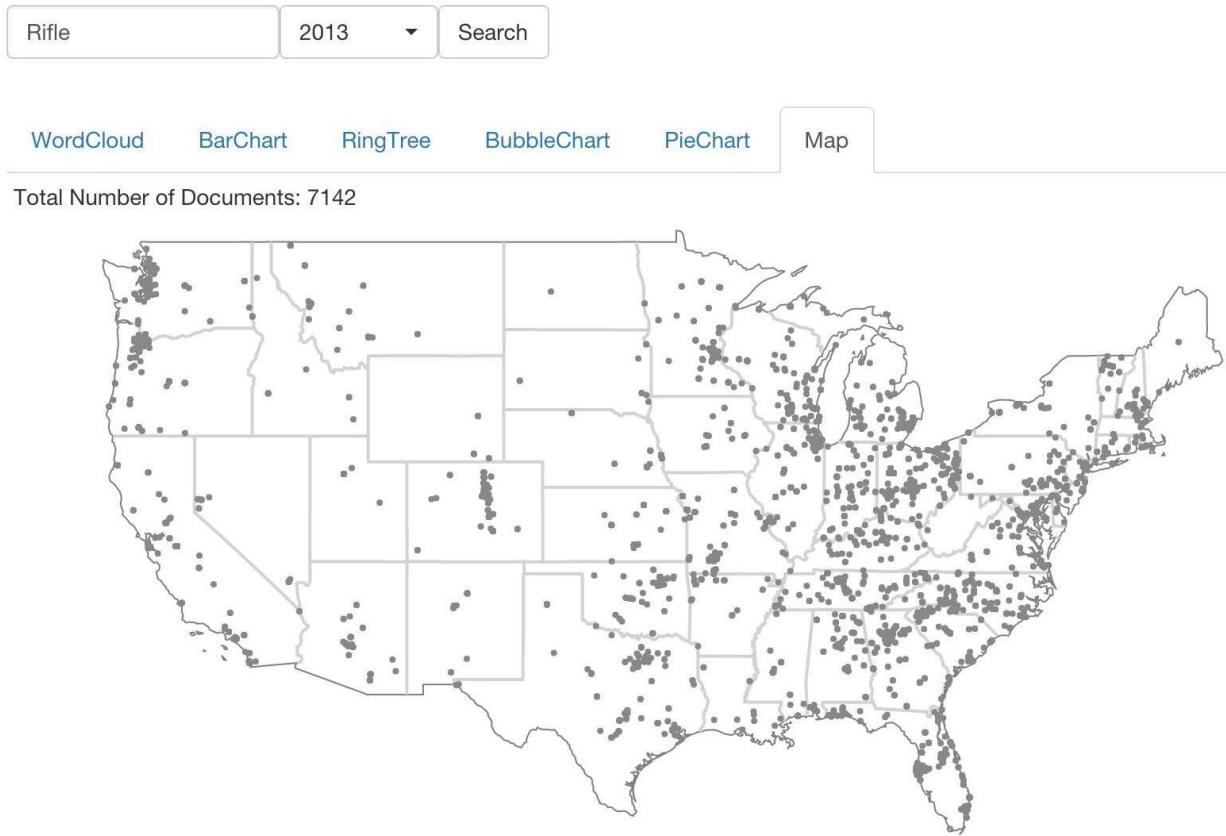
After the modification, we can request a query in index.html; FacetView will directly send AJAX call, extract JSON from solr, and display the results. We also can make the advanced modification by editing the config in jquery.facetview.js.

4. What was the hardest part, loading data, or visualizing it?

	Hardest Part	Explanation
D3	Both are hard	We develop our d3 web with nodejs, expressjs and d3js. Each part is equally hard and we already put on much effort on it. To visualize the data we have to pick several types of graph format which match our data result and implement it. Backend part is to load data from solr, we take advantage of backend nodejs to request the solr json result

		and then transform the json format to ours. Then the D3 frontend make a REST call to get the json conforming to the format to show the graph/chart/map.
Banana	Visualizing	Using the default panel in Banana is easy, only if the data column can map to the panel requirement then it should be displayed automatically. The hard part of visualizing in Banana is when user trying to makes a customer panel. It's like building a layout, we need to deal with the style and also integrate with D3 if necessary. Whereas, MVC structure provides a well developed way to do all the work. In the other hand, loading data in Banana is quite easy that we nearly don't need to do anything. As long as the data is correct, it can be loading from solr automatically.
FacetView	Both are easy	How nice the tool is! It provides a very simple way to extract and analyze the solr data. What we have to do is to configure the setting such as the host and some facet fields you want to analyze, and All are set up!

Task #3: Mapping Queries to Visualized Results



Maps shown above is the way we use to represent all queried results. In addition to input the keywords that we want to search via search bar, we can also retrieve the data based on time trends and region zones. To get the data which is based on time trends, we can select the particular Year in drop-down menu. To obtain the data which is based on region zone, we can directly enter the area in search bar. Each point on the map represents a queried data, which is related to the particular state. The total number of results is shown on the top. Following are details about mapping each answer to visualized results:

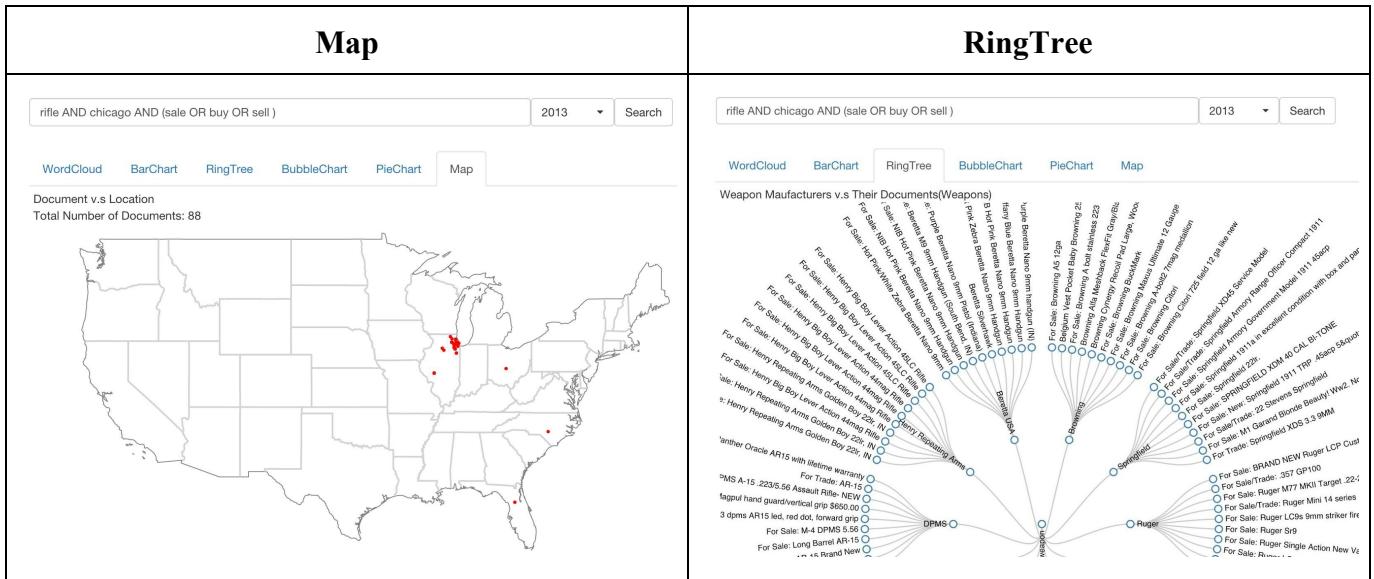
a. Identify the Texas unauthorized purchase of gun in 2015.

- Query: Texas AND (illegal OR prohibit OR ban OR (official AND lack))
- Select Year: 2015
- Keywords: illegal, prohibit, ban, (lack, official)



b.1) Identify all rifles which are sold in Chicago in 2013.

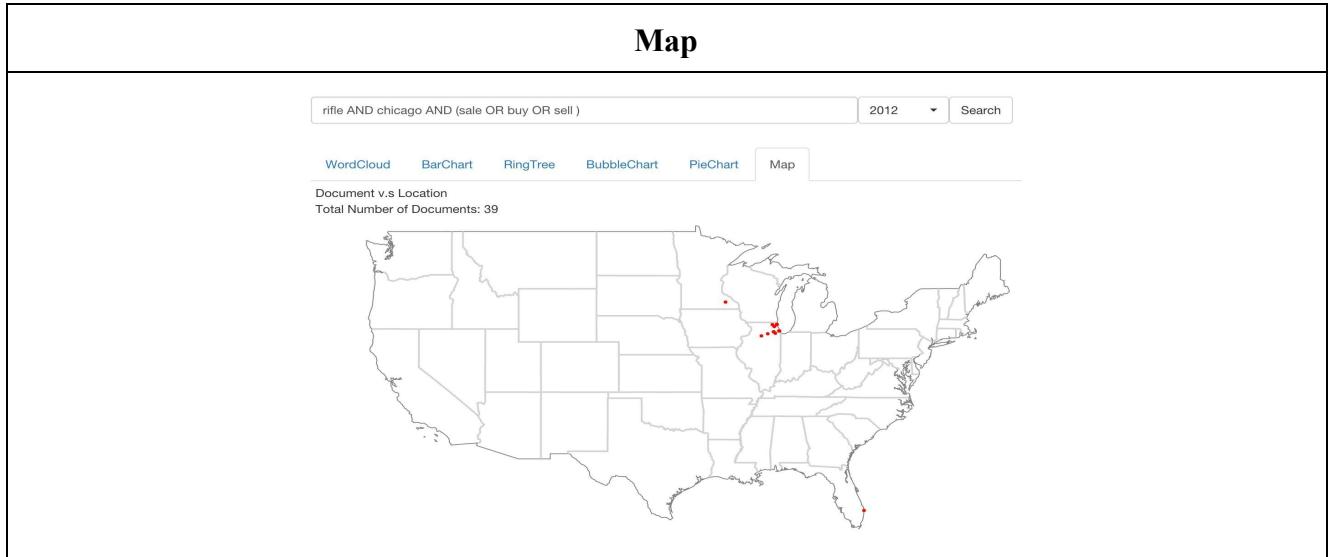
- Query: rifle AND chicago AND (sale OR buy OR sell)
- Select Year: 2013
- Keywords: sell, buy, sale



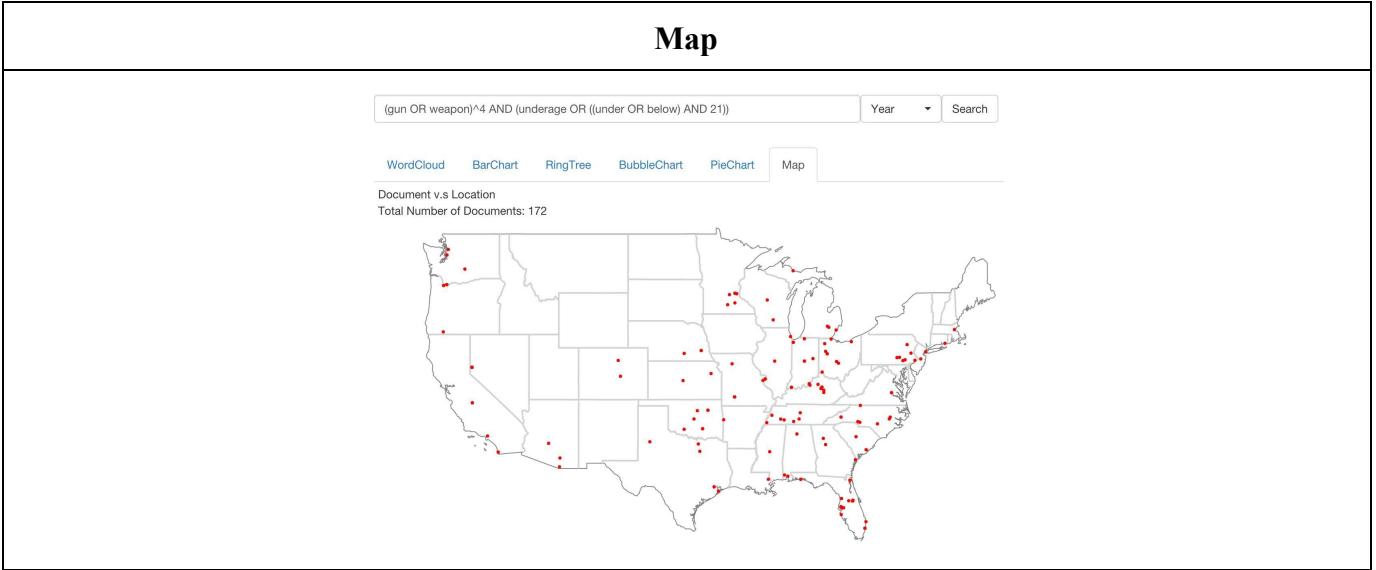
2) Determine whether rifles that is queried via b(1) are stolen goods or not.

- Query: rifle AND chicago AND (sale OR buy OR sell)
- Select Year: 2012
- Explanation:

- > In order to know whether the rifles that we queried in b(1) are stolen goods or not, we have to compare the number of rifles that are sold in 2013 with the number of rifles that are sold earlier than 2013 (In our case, the year we selected is 2012). If the former is larger than the later, then we conclude that all rifles that we retrieve are related to stolen goods.
- > Based on our queries for b(1) and b(2), the number of queries that we found in b(1) is 88 and the number of queries in b(2) is 39. Thus, all rifles that we get in b(1) are related to stolen goods.

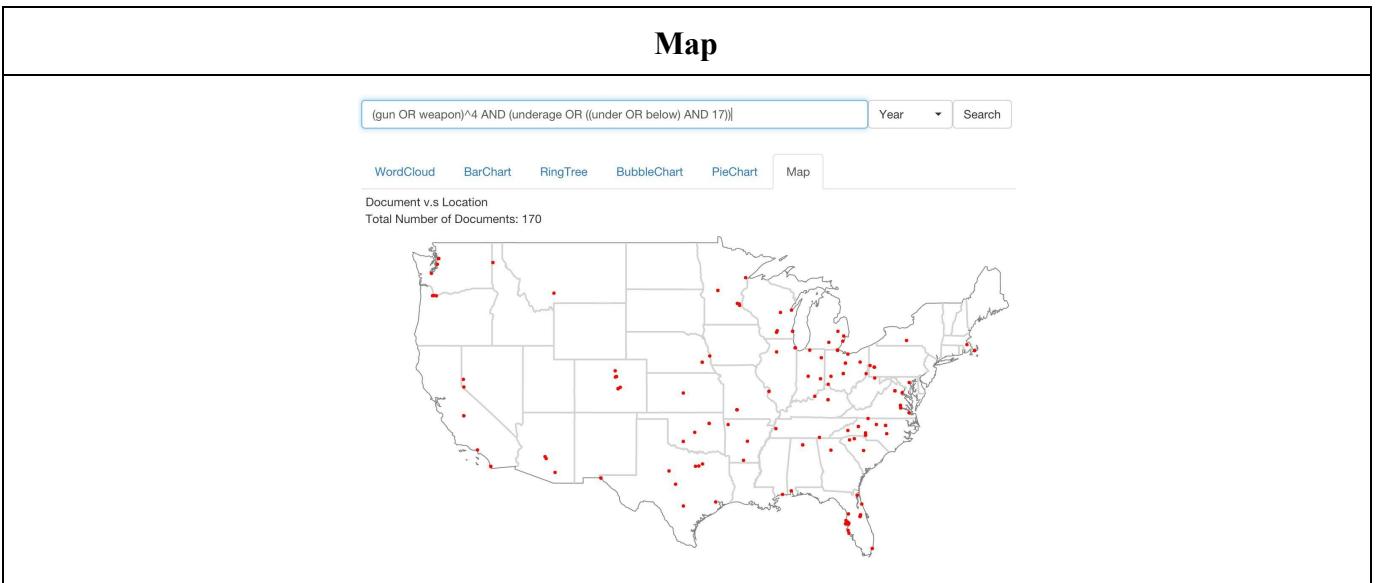


- c. Identify whether the rifles which are sold in Chicago in 2013 are stolen.
 - Query: rifle AND chicago AND (sale OR buy OR sell)
 - Select Year: 2013
 - Keywords: sell, buy, sale
 - Explanation:
 - > To decide if the rifles which are sold in 2013 are related to stolen goods, we have to know the number of rifles that is sold earlier than 2013 (In our case, the year we selected is 2012).
 - + Query: content: rifle AND chicago AND (sale OR buy OR sell)
 - + Select Year: 2012
 - > Since the number of former queries that we found is 88 and the number of later queries is 39. Thus, all rifles that we get in the former are related to stolen goods, we conclude that all rifles which are sold in Chicago in 2013 are stolen.
- d. Identify gun and weapon ads that are posted by person whom are underage.
 - 1) If underage means the age under 17:
 - Query: (gun OR weapon)^4 AND (underage OR ((under OR below) AND 17))
 - keywords: underage, under 17, below 17



2) If underage means the age under 21:

- Query: (gun OR weapon)^4 AND (underage OR ((under OR below) AND 21))
- keywords: underage, under 21, below 21



e. Identify the unlawful transfer, sale, possession of explosives, and WMD devices

- Query: (violate OR violation OR illegal OR prohibit OR ban OR criminal)^4 AND (explore OR explosive OR nuclear OR chemical OR transfer OR (sale OR buy OR sell))
- keywords: violate, violation, illegal, prohibit, ban, criminal, sell, buy, sale, explore, explosive, nuclear, chemical

Map

(gun OR weapon)^4 AND (underage OR ((under OR below) AND 17))

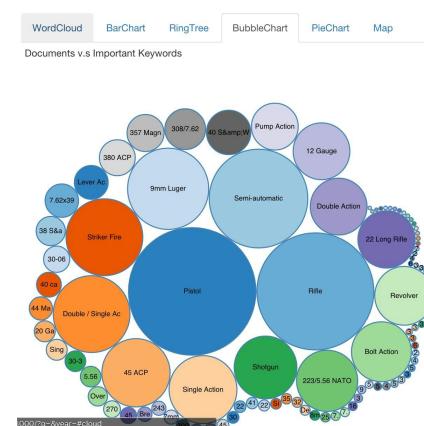
WordCloud BarChart RingTree BubbleChart PieChart Map



D3 Demonstration

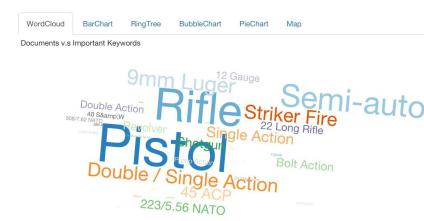
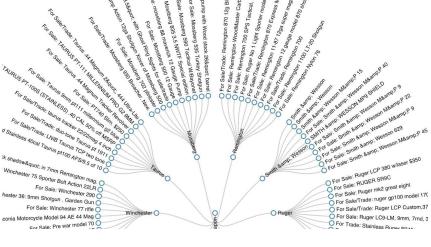
(gun OR weapon)^4 AND (underage OR ((under OR below) AND 21))

WordCloud BarChart



WordCloud BarChart RingTree BubbleChart PieChart Map

Weapon Manufacturers v.s Their Documents(Weapons)



Video Link: <https://youtu.be/An44Q7xFq9g>