Assignment 3

Visual Search and Interaction with Weapons Datasets

I. D3 Visualization –

- This is a visualization technology to interact and visualize the search engine data.
  An XMLHttpRequest is issued to /solr/search by appending a query string consisting of free text and params:
  wt=json
  &rows=2000
  indent=true
- We have created a REST API which is communicating with the SOLR server to get the JSON response, that has been passed to our D3 widgets using JavaScript, AJAX and Java servlets.
- CORS was enabled in JavaScript to interact with the SOLR server.
- We picked the following six widgets –
  - US State Map
  - Line Chart
  - Word Cloud
  - Bar Chart
  - Calendar View
  - Bubble Chart

- **How easy was it to use D3?**
  The documentation provided for D3 visualization online was initially very hard to interpret but the Github repositories and the open source contributions helped us a great deal in overcoming all the challenges that we faced while implementing this assignment. This was a great learning curve for us.

- **What was the hardest part?**
  - To manipulate our data in such a way that we could visualize them in our D3 widgets.
  - Selection of the most appropriate D3 widgets to accurately represent our data for the challenge questions.

- **Advantages:**
  - This technology was a great way to represent our data results in an easy to comprehend format. We could understand the content in a more straightforward format.
  - D3 helped us reveal hidden patterns in our data and find connections between elements that we would have otherwise overlooked.
  - It was easy to integrate D3 with AngularJS and VanillaJS.

- **Disadvantages:**
  - Shortage of D3 visualization examples that correctly visualize all kinds of data.
  - It was a risk that the resultant visualization we made, might not fully illustrate the significant implications of our datasets. A certain amount of the important information was getting

overlooked so we used another visualization for better representation.

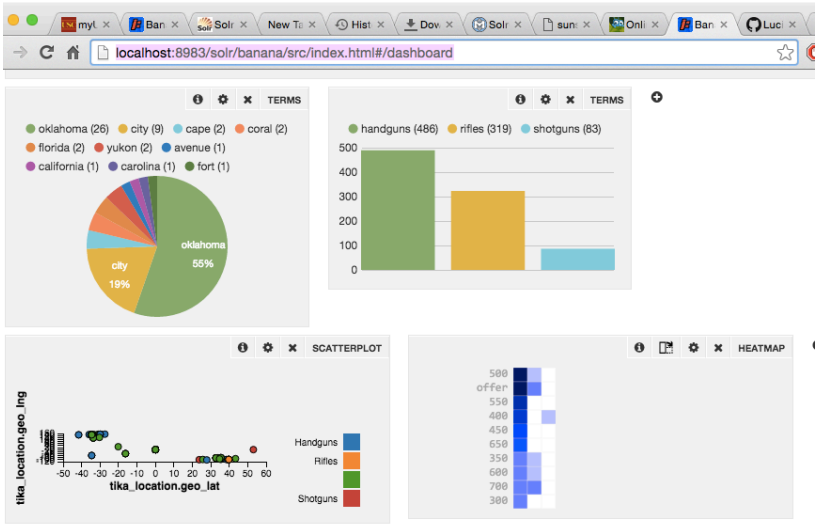
II. Banana Technology:

- Be it time series data or non-time series data, Banana can be run on any kind of data, not just with logs.
- It uses Kibana's powerful dashboard configuration capabilities and integrates it in SOLR.
- An easy-to-use and flexible user interface makes it approachable for any non-technical user as well.
- The six panels that we used were:
  - HITS Panel
  - Query Panel
  - Terms Panel
  - ScatterPlot
  - Table Panel
  - Tag Cloud

- **How easy was it to use Banana?**
  - We effortlessly connected to SOLR using Banana as it was extremely well documented.
  - The pre-defined panels were very undemanding as it took our values as inputs and gave us a corresponding, easy to understand visualization.
  - We got multiple visualizations for each of our queries. This lead to more elaborate understanding of the query results and hence, our weapons dataset.

- **What was the hardest part?**
  - Banana requires that changes need to be made to include the field-set every time it is loaded onto another system. If changes could be made so that it reads or suggests an alternative to that of the configured field it would add up to its usability.
  - Getting values on the map was very tough as we had to have a specific format for which we had to manipulate the data.

- The following is a screen-shot of one of the results that we got after implementing Banana –



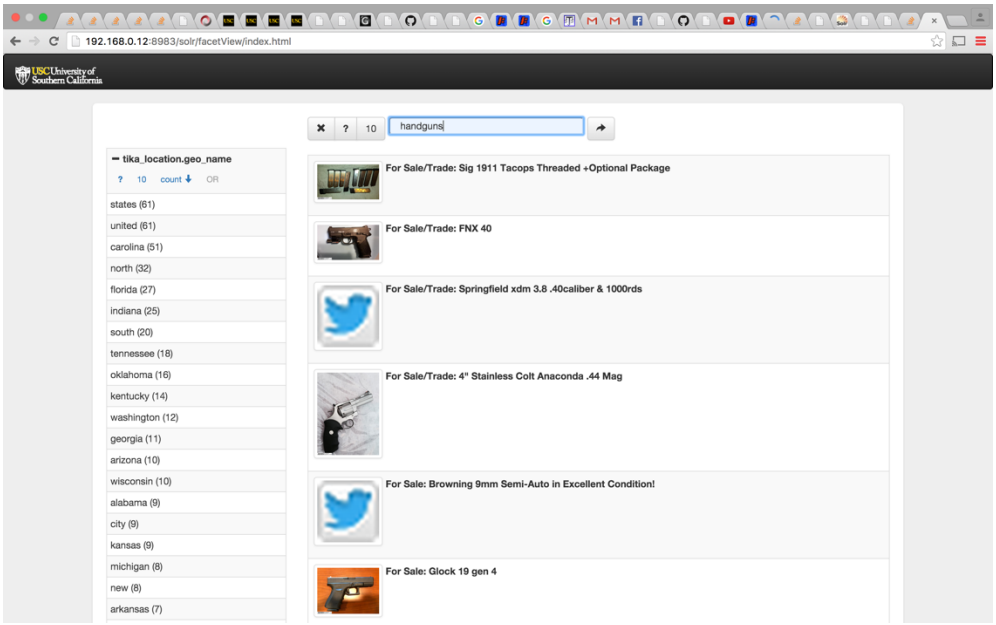
III. FacetView–based Search Interface:

- This is a user-defined interface with pure JavaScript at its front end.
- It is used for searching indexed data from SOLR. It also provides a micro-framework you can build on when creating user interfaces to SOLR.

- **How easy was it to use FacetView?**
  - FacetView took us a long time to figure out but once we did, it was not very hard to integrate it into our dashboard.
  - It helped us visualize the data accurately by grouping it on pre-defined categories.
  - Querying to SOLR from FacetView was relatively simple.

- **What was the hardest part?**
  - FacetView has very limited documentation hence, it was very difficult to load and visualize the data.
  - The JavaScript present on the Open Source Repository is not fully compatible with SOLR.
  - Creating a custom JS and integrating it with Facetview and SOLR was a challenge.
  - It was hard to integrate FacetView with the old Bootstrap region.

- The following is a screenshot of our FacetView Visualization –

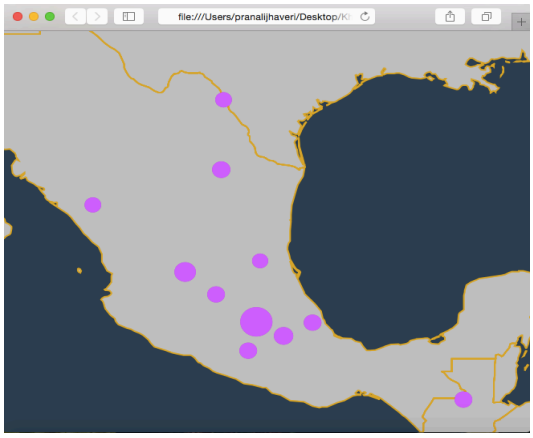


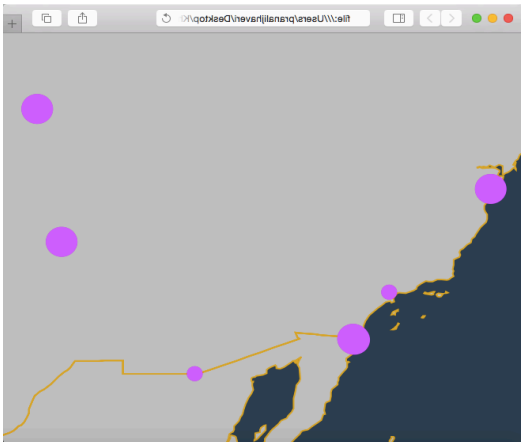
IV. Map Visualization:

- **How to use the Map to visualize the results?**
  - We did Map Visualization using the Geospatial Khooshe tool using the sample_tsv file which contains a list of our latitudes and longitudes.
  - We fetched the latitude and longitude by firing a query to SOLR and recording the

required data of latitudes and longitudes.

- Sample_points.tsv contains a list of the latitudes and longitudes which we fed to Khooshe obtained from SOLR.





➢ **Answering challenge questions presented in Task #3 *(Queries addressed by the Map):***

- In Task #3, we fired a query to each of the challenge questions and found out relevant data, with which we plotted a Khooshe Map and visualized the spatial points based on the data we received.

  a. **Unauthorized Purchase of Guns:**

  We observed several points in the Florida and in the years 2014–2015. We learn that there are lot of points in the map that signifies these purchases.

  b. **Similar Firearm Types found in the Same Region and Time:**
  Florida and its nearby regions as it can be noticed in the map, can be identified as those regions that sells the maximum number of similar firearm types for example, shotguns. This indicates an influx of stolen goods in the Florida region.

  c. **Stolen goods increase Rate of Sales for a Particular Make/Model:**
  The result we are getting is of handguns and rifles, which goes to show that there is an increase in the sales of handguns and rifles if the bulk of the shipment is stolen in the Mexico area. The Mexico region has the maximum of points according to the map.

  d. **Weapons Sold by Underage Persons or De-identified Weapons:**
  We see that the Bubble Density on the map is the largest in the Mexico and Florida Regions, so that might indicate a large number of weapons sold by underage persons as anyone would be allowed to sell/purchase firearms.

  e. **Unlawful Transfer, Sale or Possession of explosives, WMD Devices and Precursors:**
  For this query, we did a search on explosives, WMD devices and precursors and in respect to that we can locate the points on the map. For example, the location of the explosives and other such weapons or devices.

Link to our project video : https://youtu.be/MNk39_eehQA