Hands-on Elasticsearch & Twitter: data visualization & text classification by Mary Loubele PhD

0. Needed software

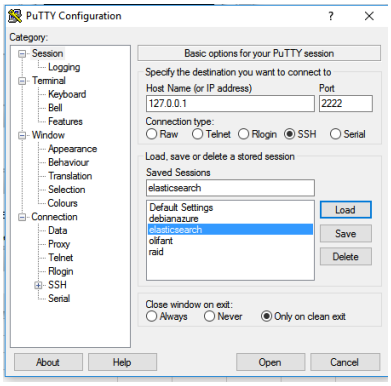
Virtualbox https://www.virtualbox.org/
putty https://www.chiark.greenend.org.uk/~sgtatham/putty/latest.html
A webbrowser like Chrome, firefox

1. Testing the Elasticsearch installation

Download the VM http://loubele.org/Elastic.ova or get it from the USB key
Open Virtualbox
Go to File, Import Appliance and select the (virtual machine) VM from the USB
Right click on the imported VM and select Start, Normal Start

Accessing the VM for Windows users

Open putty and make sure the putty window looks as follows:



Next login to the VM. The **username** is elastic and the login is **search.**

Accessing the VM for Mac and Linux users

Open a terminal and type ssh -p 2222 elastic@localhost and press enter
Next type the password search

Checking whether Elasticsearch is alive

curl 127.0.0.1:9200

Checking for the data available in Elasticsearch

curl 127.0.0.1:9200/_cat/indices?v

Checking whether Kibana is alive

Go to your web browser and type the following url

http://127.0.0.1:5601

2. Hello snowy Canada

Loading the data

Go back to the console of your vm

cd ~/code

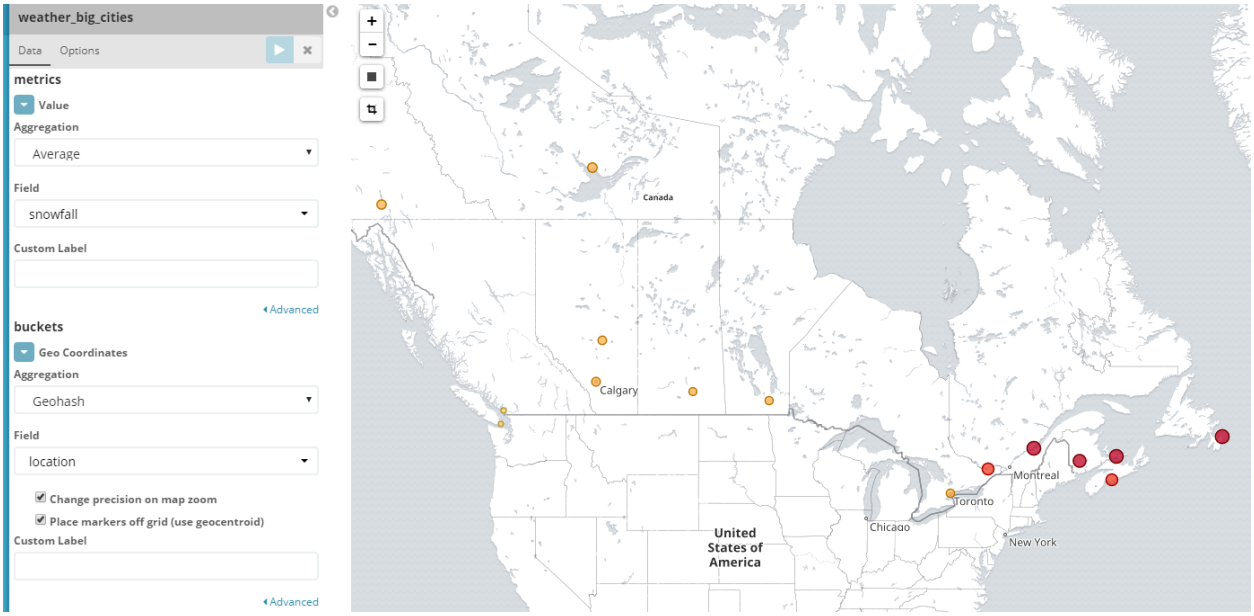
python create_index_for_weather_big_cities.py
python load_weather_big_cities.py

Making the coordinate map in Kibana

Left column: select Management
In Index pattern type weather_big_cities, hit create

Left column: select Visualize, Create a visualization, Select Coordinate Maps

Select weather_big_cities and let it map with the following. Save the visualization.

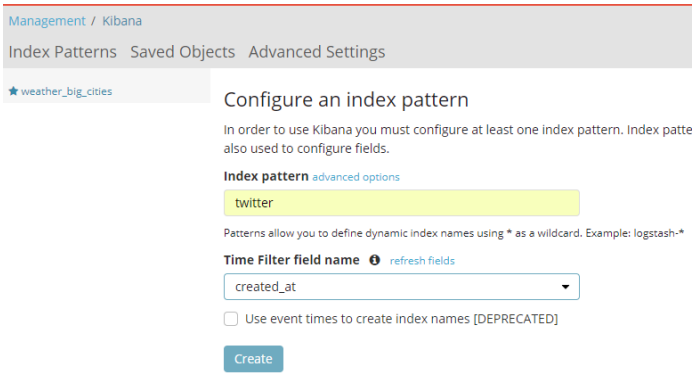


Extra exercises

Which other fields can be visualized in the map. Build a visualization with them.
Build a dashboard with these three visualizations
Add in an extra city in the load_weather_big_cities.py and execute the script again.

3. Visualizing Twitter data

Create index pattern twitter



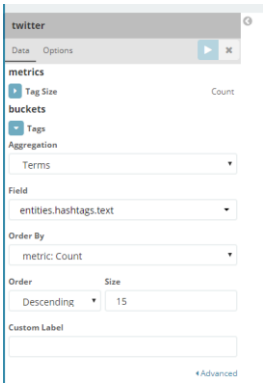
Type Discover and select time picker and select last 7 days

Language visualization

Go in middle left column to lang and type Visualize

Make a word cloud

Go to visualization, Select Tag Cloud, twitter



Extra exercises

Make two extra visualizations of your choice

Build a dashboard

Add a filter and check the influence for example from the user language, words in the text, …

Working with the sentiment_analyzer

From the command line

Go back to the console and type the following command. You can play around with some extra sentences

python get_sentiment_for_tweet.py "I had a wonderful day"

Adding the tweets_with_analysis index

Add the index in the same way and make a pie chart for the “sentiment” field.
Add this visualization to the dashboard
Play again with different visualizations

4. Working with your own twitter data

Getting API keys

1. Make sure your phone number is registered with your twitter account
2. Go to htts://apps.twitter.com
3. Create New app
4. Fill in all the information
5. Go to Key and Access Tokens and select **“Create my access tokens”**

Adding in the different keys and loading the data

1. Go to command line type nano load_tweets_final.py
2. Fill in all the different keys Control O to stop
3. python load_tweets_final.py word_to_analyze
4. Analyze the data from your keywords