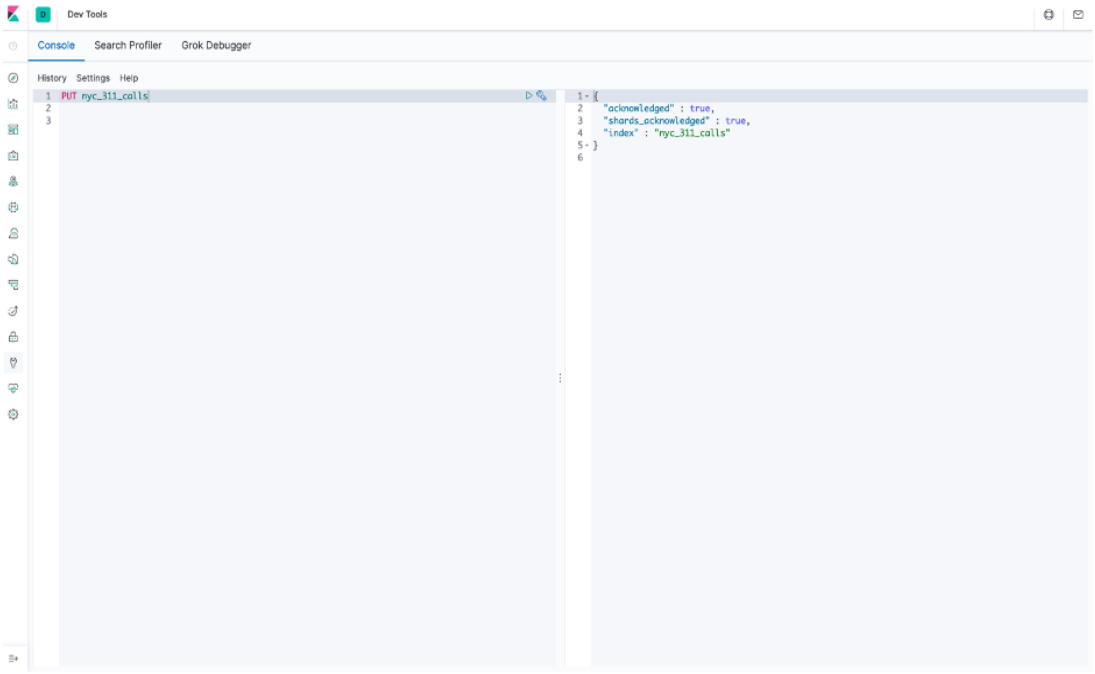
Group 1
Group members: Goodnews Agbadu (ID: 200532679)
Yi-Chen Hsaio (ID: 200540377)
Cassandra Forlani (ID: 200544446)
Due Date: Sunday, December 18th, 2022

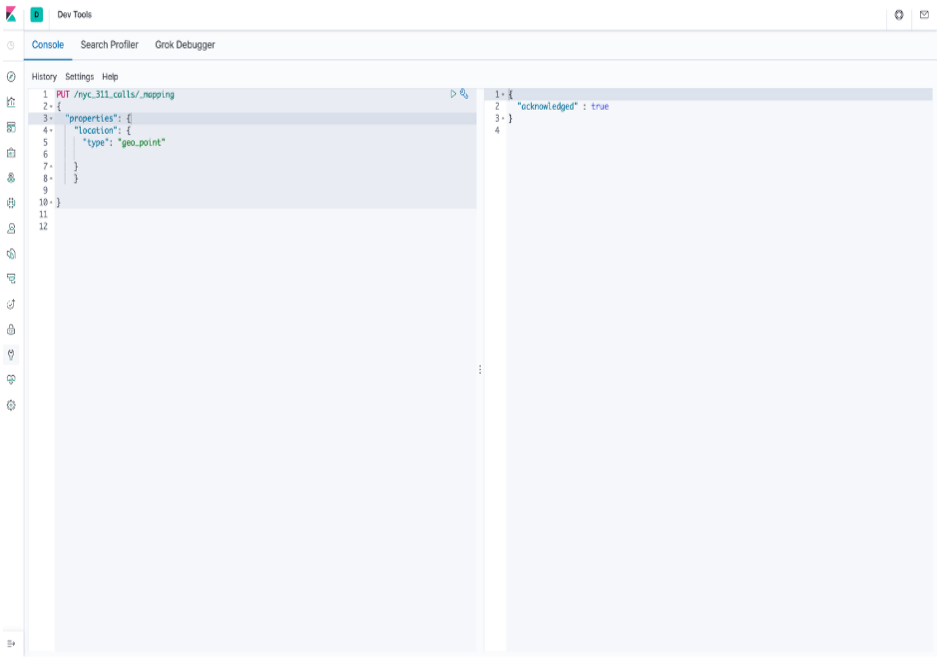
BDAT1002-02 Final Project

Part 1: Beginning Process

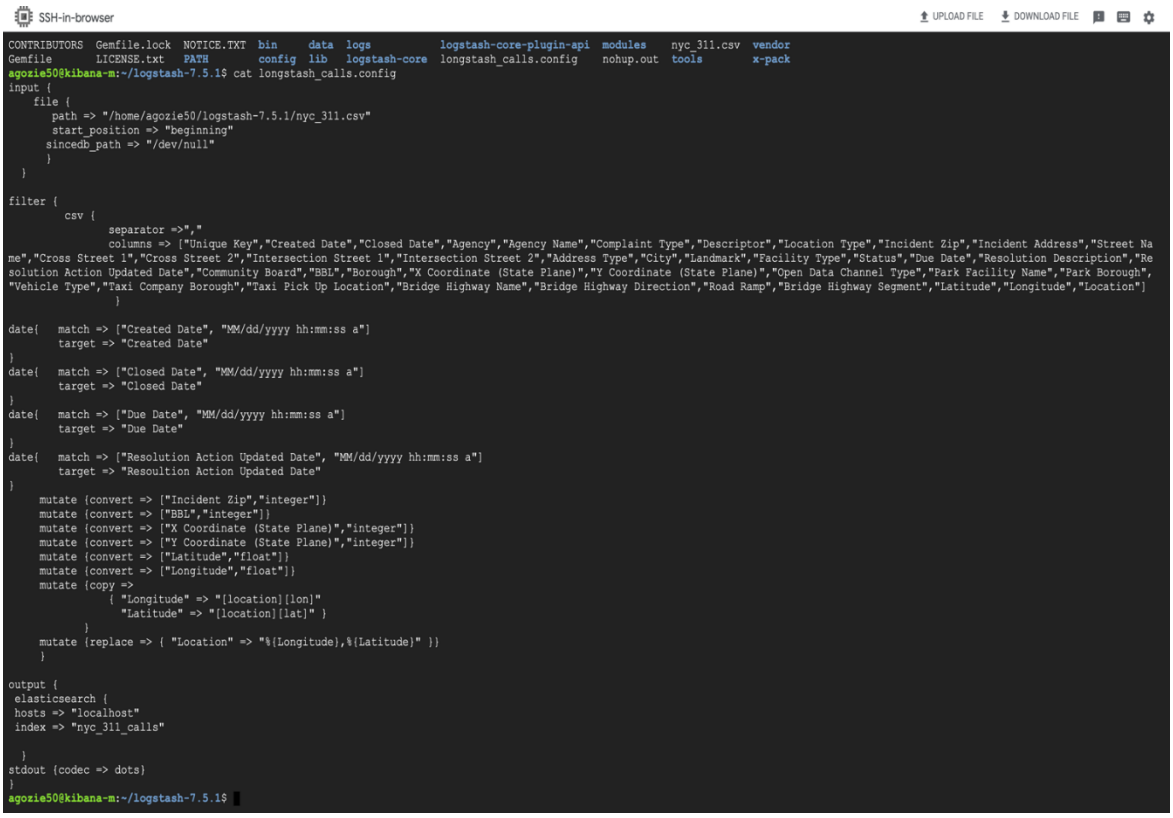
Creating an Index in Kibana



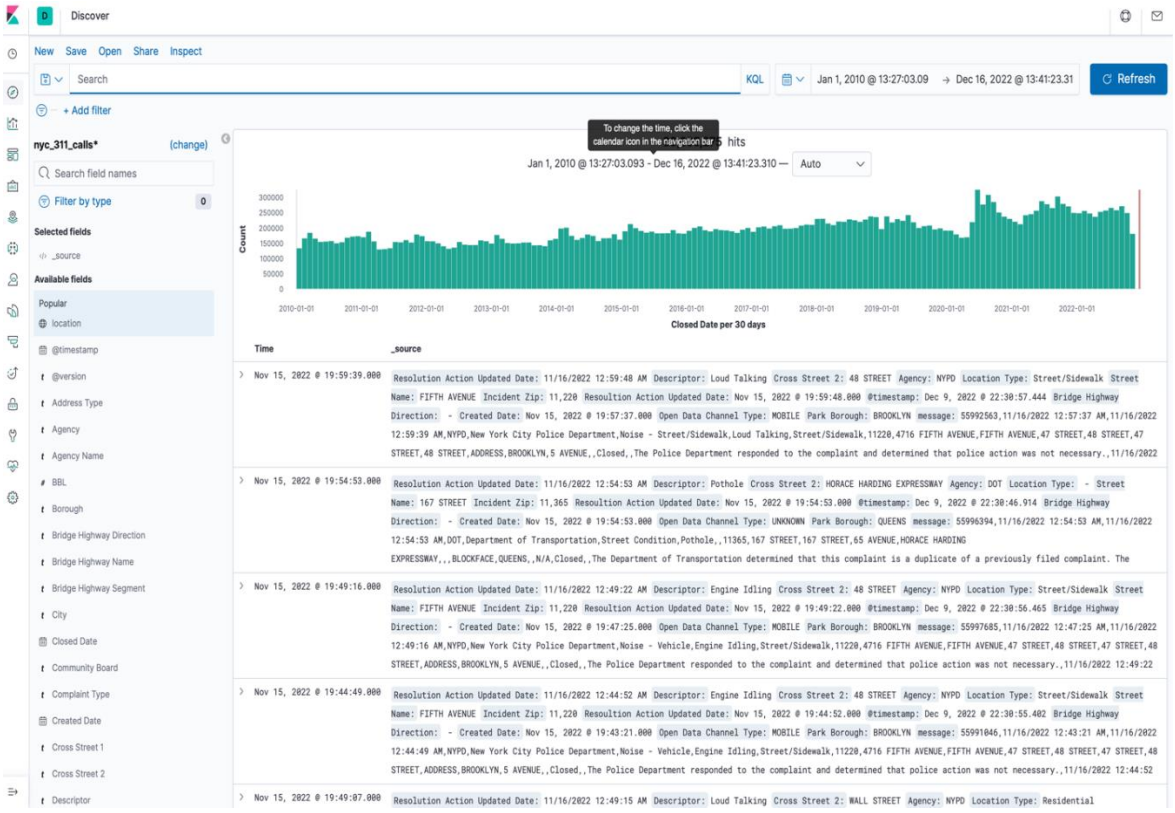
Mapping to add location type as Geo point



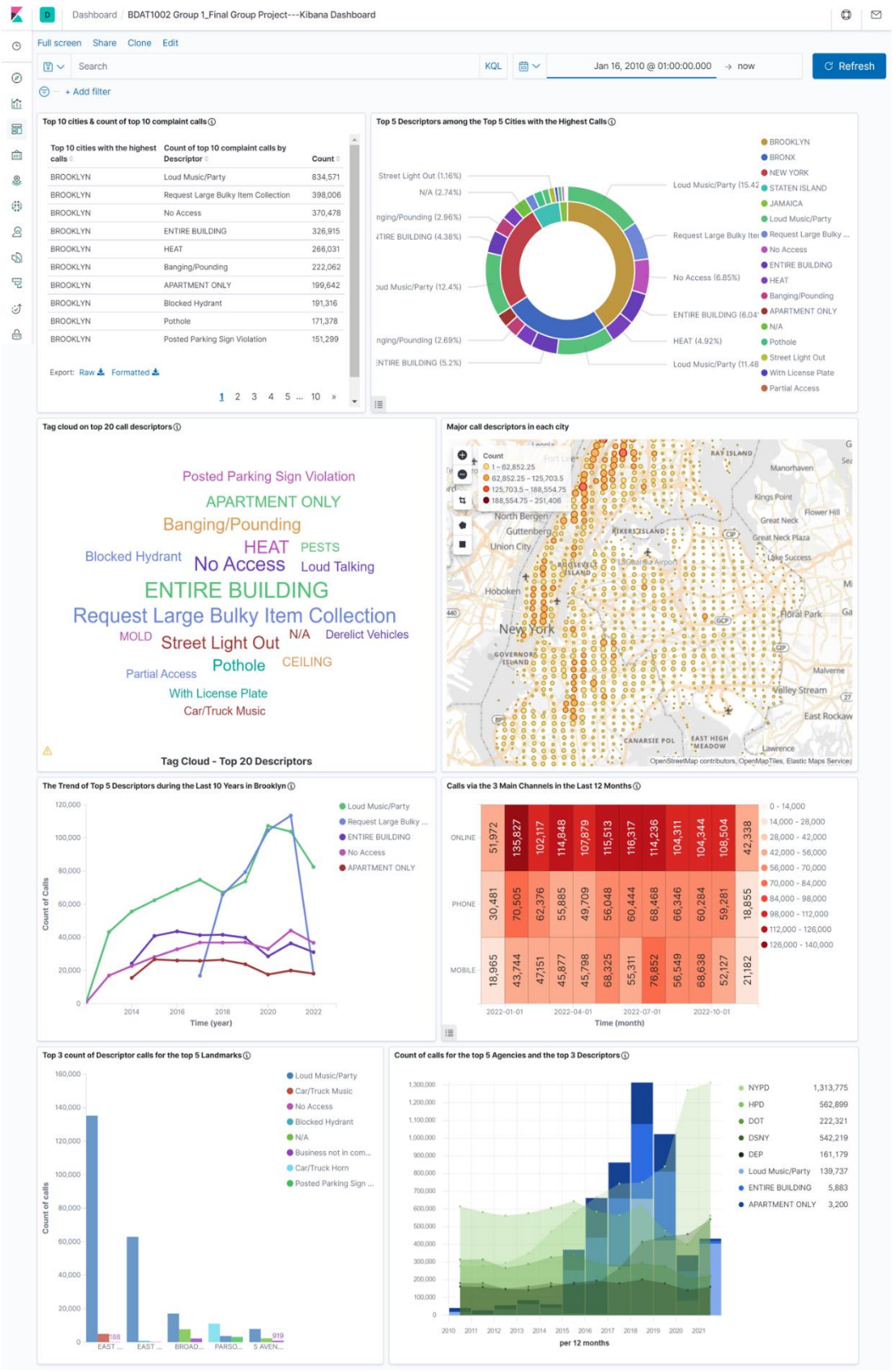
Checking the content of Logstash configuration file to ensure that index is probably named



Checking to confirm that the data is loaded



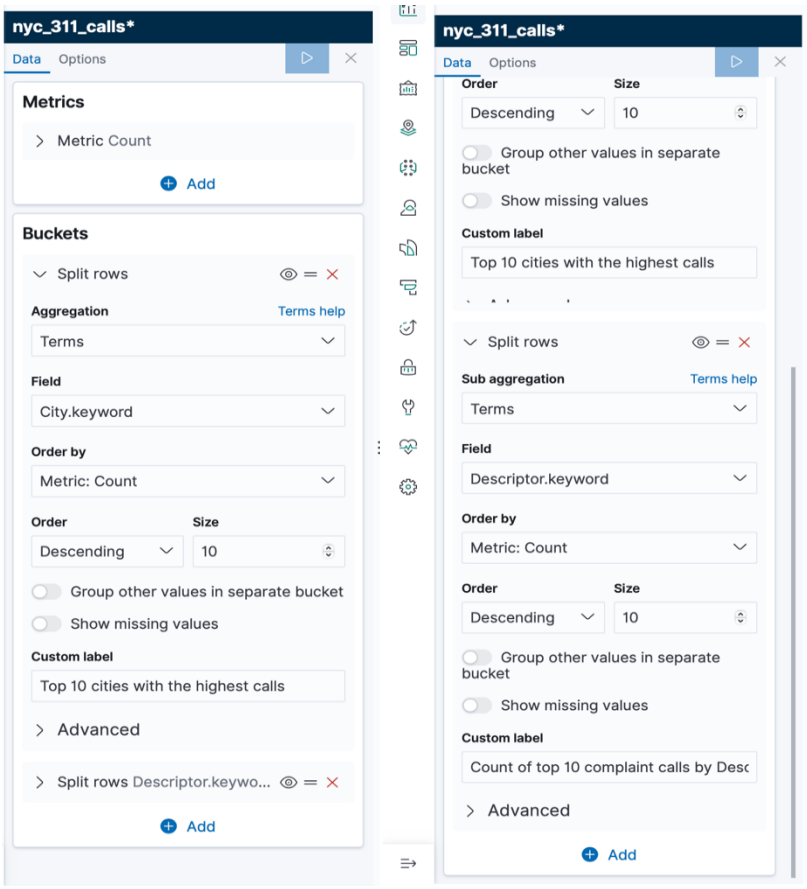
Dashboard



Part 2: Analytical Questions

Question 1: Create a table showing the top 10 cities with the highest calls alongside the count of top 10 complaint calls (by Descriptor) in each city

How to get the visual



Actual visual

Top 10 cities & count of top 10 complaint calls

Top 10 cities with the highest calls	Count of top 10 complaint calls by Descriptor	Count
BROOKLYN	Loud Music/Party	834,695
BROOKLYN	Request Large Bulky Item Collection	398,006
BROOKLYN	No Access	370,543
BROOKLYN	ENTIRE BUILDING	326,915
BROOKLYN	HEAT	266,894
BROOKLYN	Banging/Pounding	222,106
BROOKLYN	APARTMENT ONLY	199,642
BROOKLYN	Blocked Hydrant	191,328
BROOKLYN	Pothole	171,477
BROOKLYN	Posted Parking Sign Violation	151,306

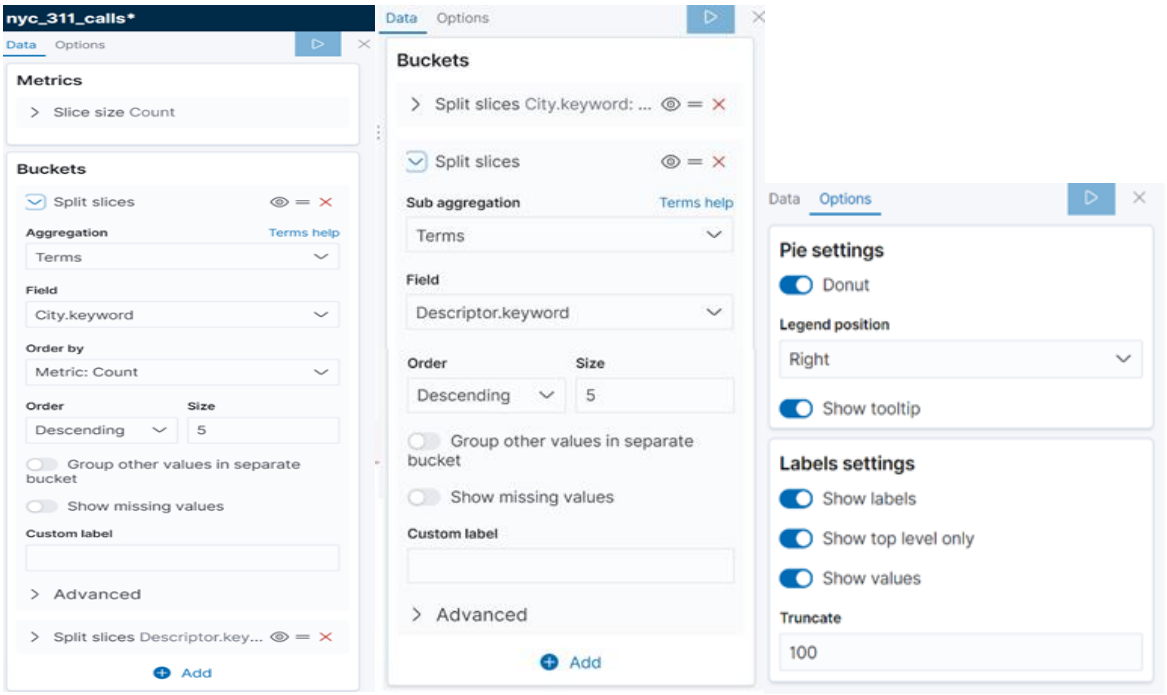
Export: Raw Formatted

1 2 3 4 5 ... 10 »

Analysis: The table above shows that Brooklyn recorded the highest calls among the cities while the highest calls complaint by descriptor was Loud Music/Party.

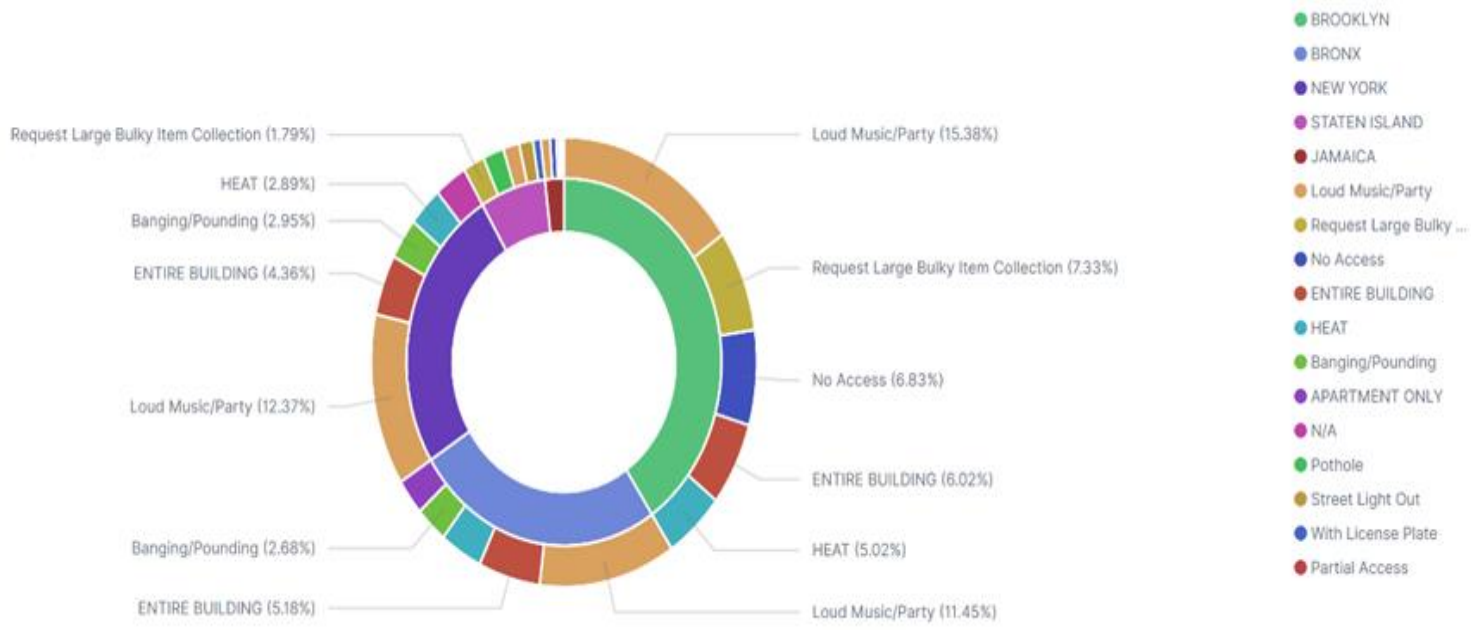
Question 2: Create a pie chart showing the top 5 cities with the highest calls alongside the top five calls (Descriptor) in each city

How to get the visual



Actual visual

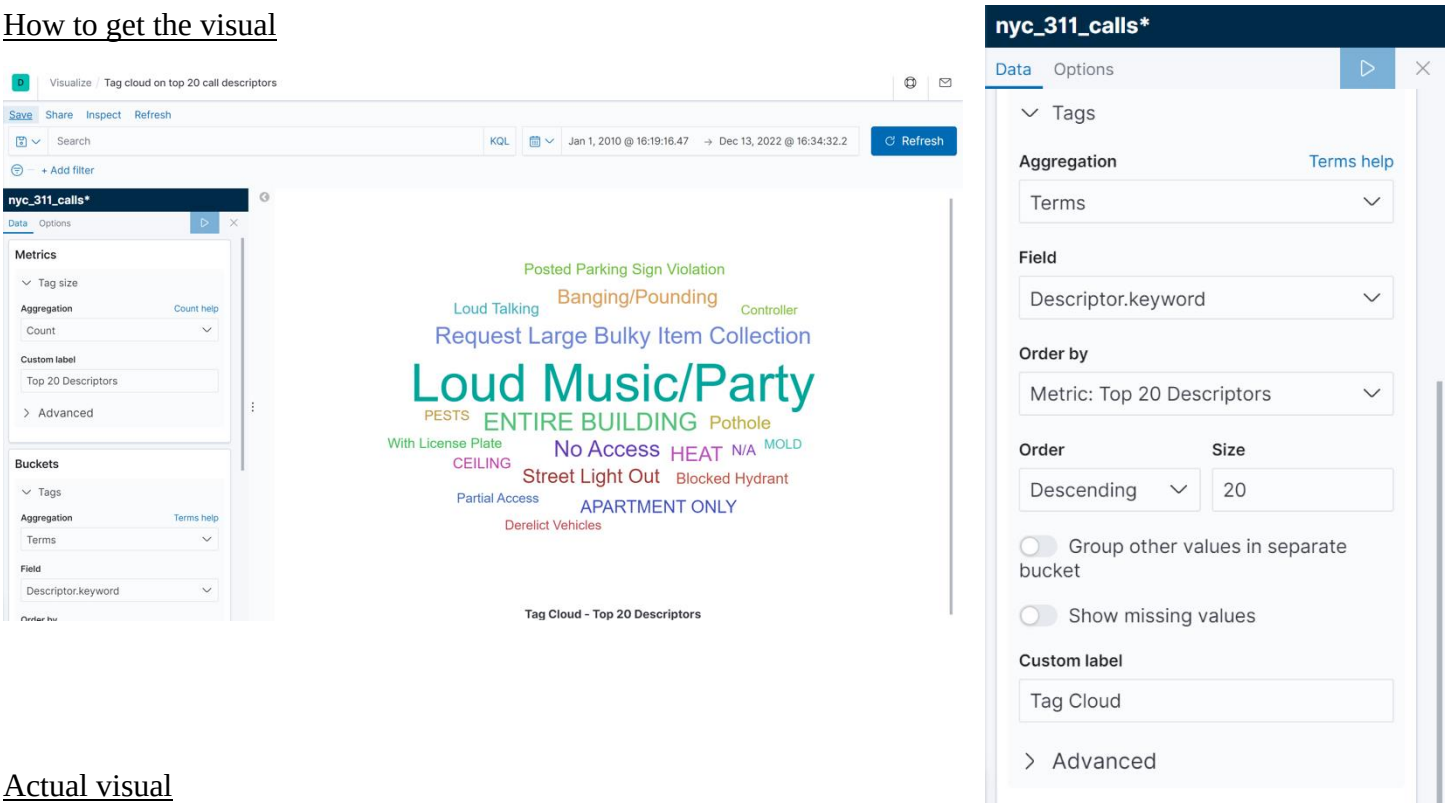
Title: Top 5 Descriptors among the Top 5 Cities with the Highest Calls



Analysis: The above pie chart shows the top five descriptors in the top five cities with the highest number of calls. As shown in the chart, Brooklyn (9,057,926 calls) has the highest number of calls, followed by Bronx (5,685,664 calls), New York (5,646,162 calls), Staten (1,499,554 calls), and Jamaica (431,885 calls). The chart reveals that 'Loud music/party' is the top descriptor across the five cities except for Staten.

Question 3: Create a tag cloud representing the top 20 call descriptors.

How to get the visual

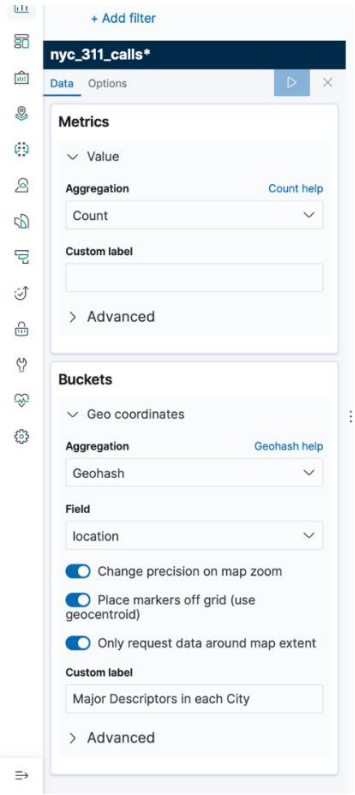


Actual visual

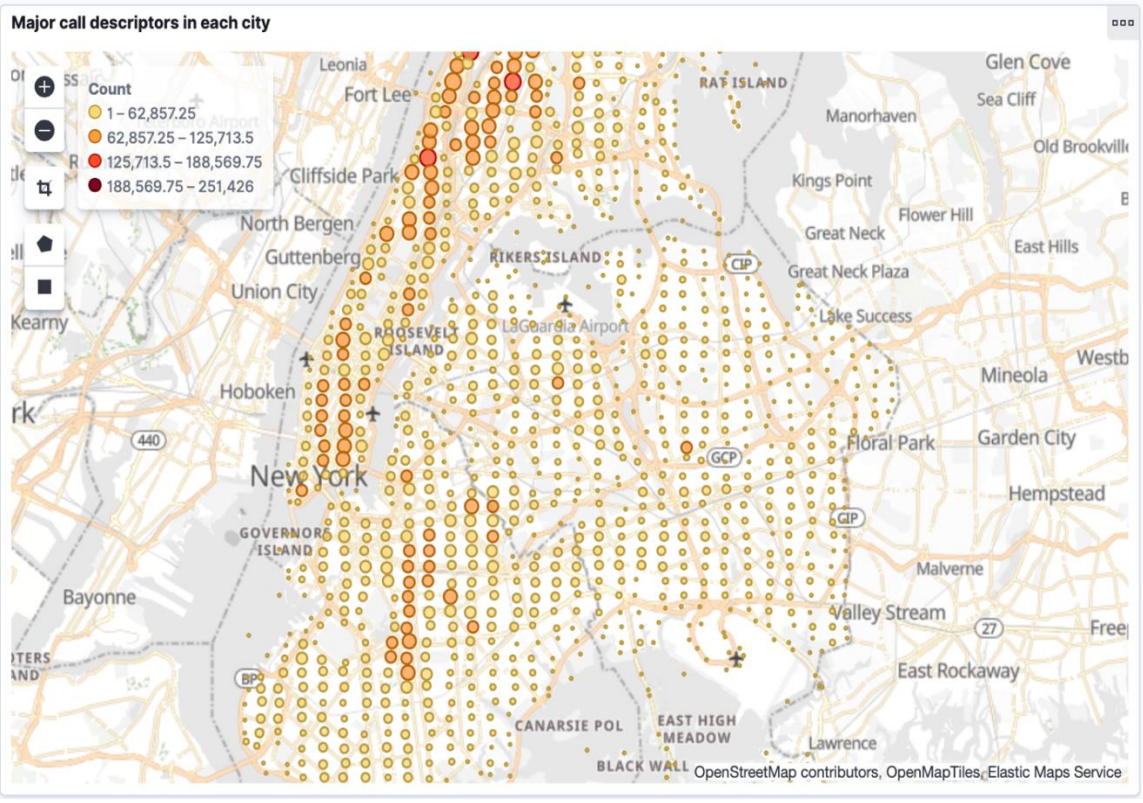
Analysis: The above image shows the top 20 descriptors. Their size is based on their count value. For example, Loud Music/Party is in the largest font which shows it is the highest descriptor being called in.

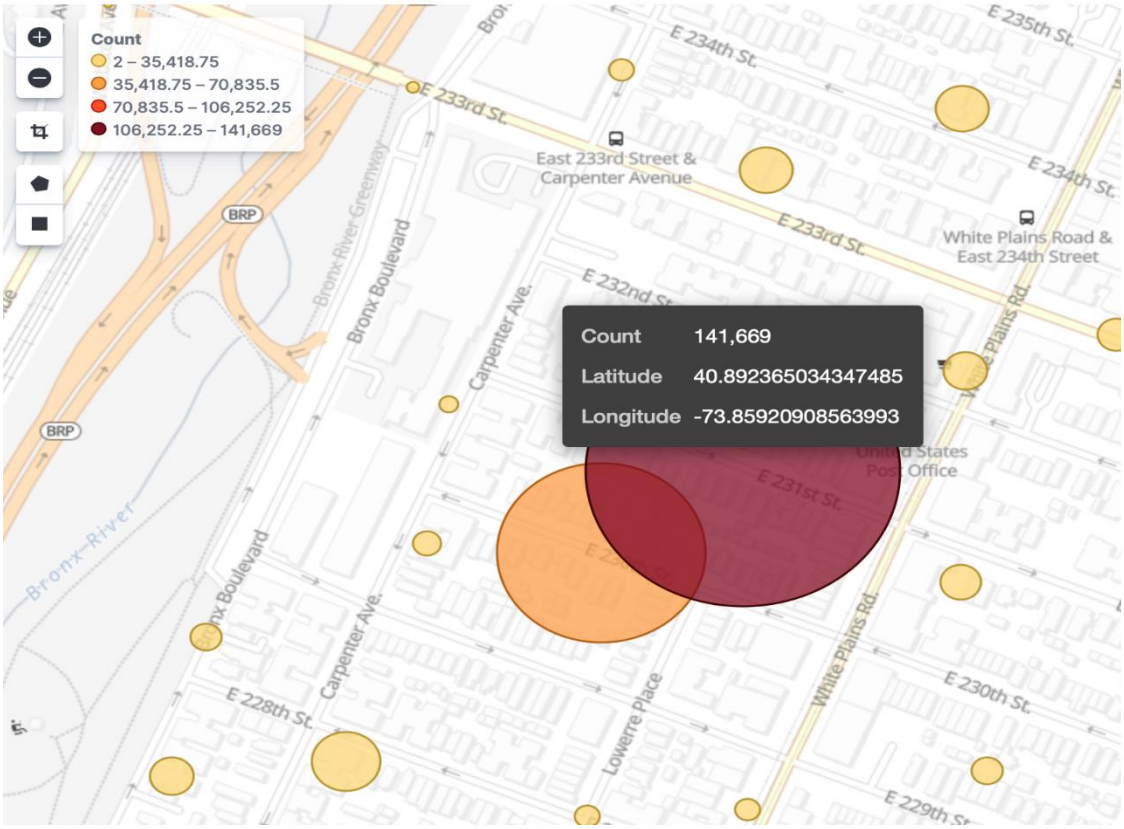
Question 4: Create a coordinated map of all the major call descriptors in each city

How to get the visual



Actual visual





Analysis: The above map shows all major call descriptors in major cities, New York had the highest call concentration. In addition, carpenter Ave had the highest call descriptors by street.

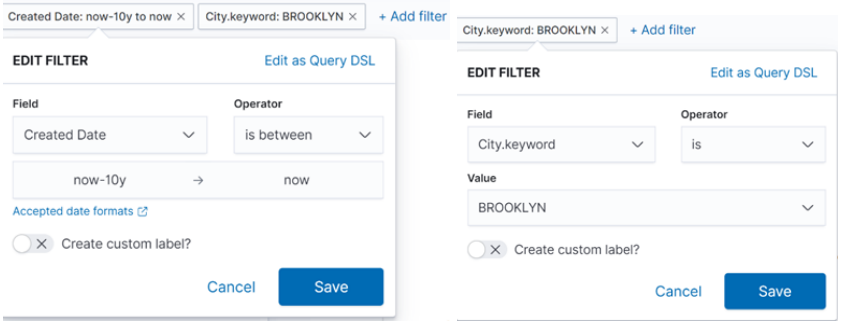
Part 3: Extra Visualizations

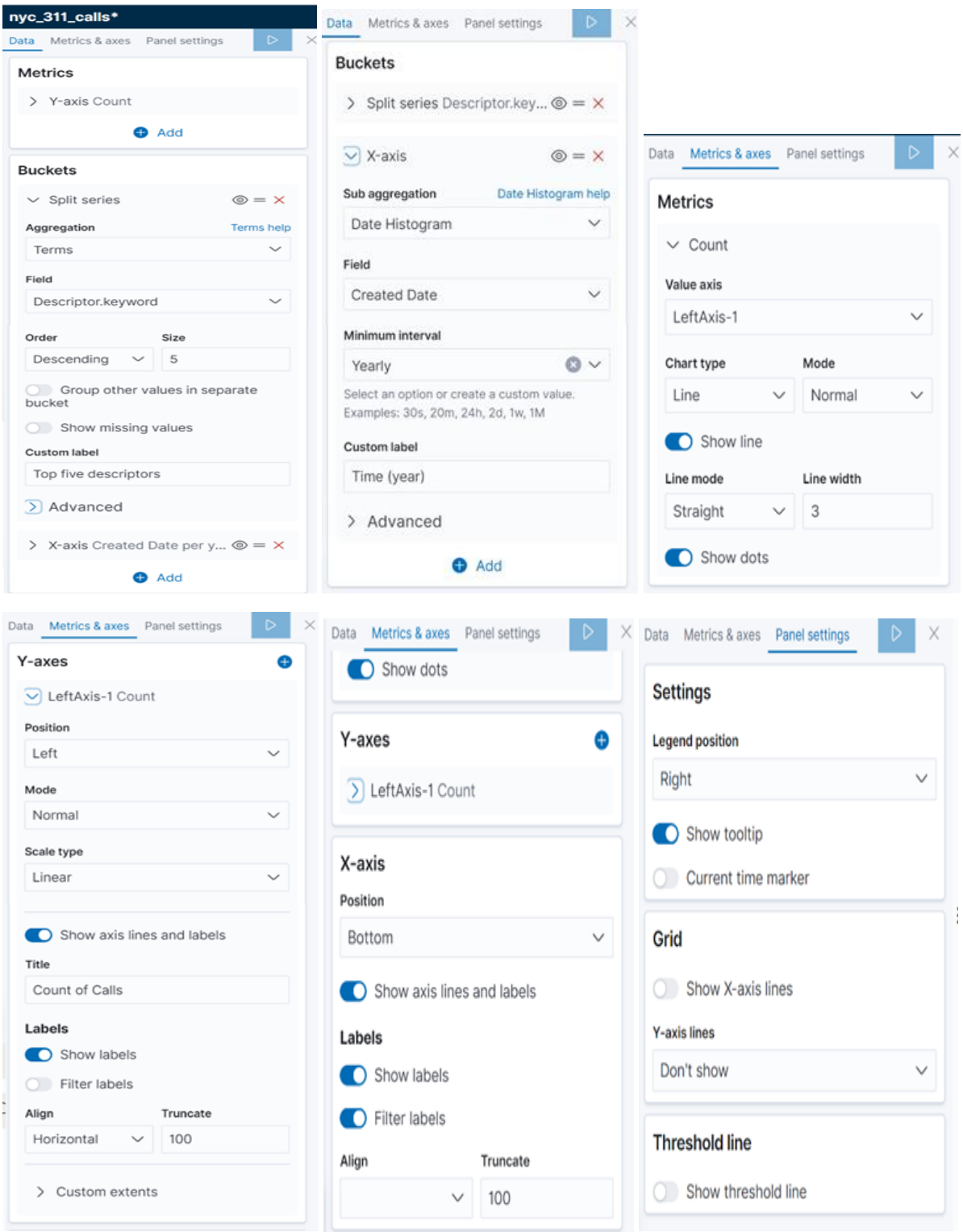
Extra Visualization 1:

Idea of Visual

Brooklyn is the city with the most significant number of calls. The visual aims to take a closer look and examine how the top five descriptors of calls in Brooklyn have changed over time during the past 10 years.

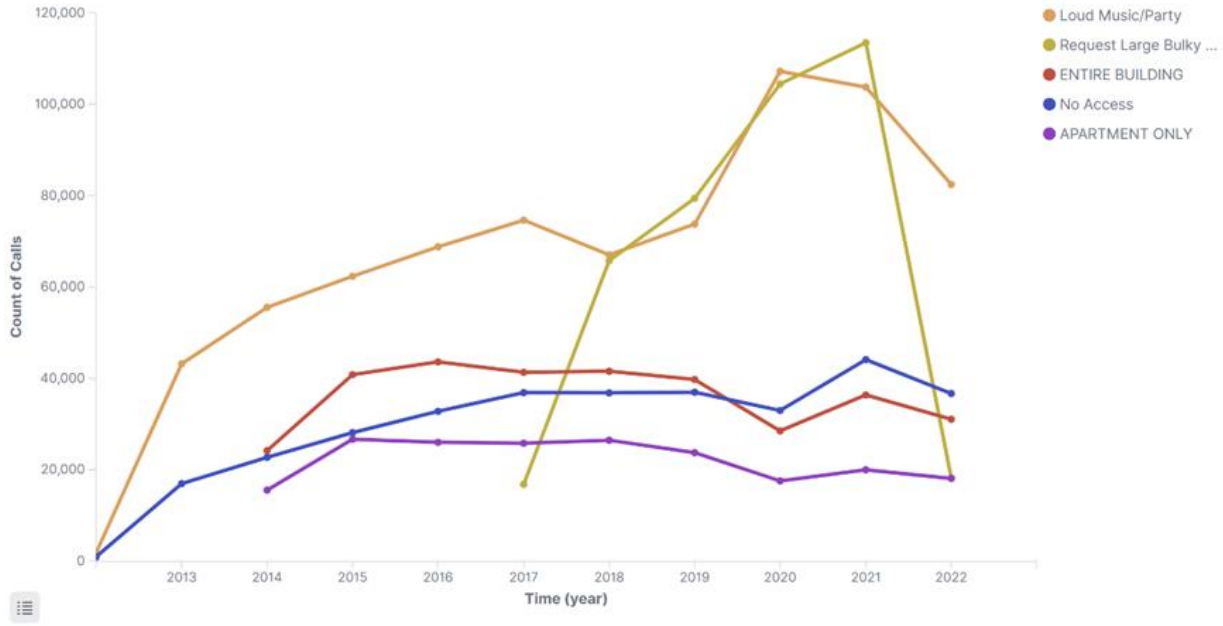
How to get the visual




nyc_311_calls*
Data
Metrics & axes
Panel settings
Metrics
Y-axis Count
Add
Buckets
Split series
Aggregation
Terms help
Terms
Field
Descriptor.keyword
Order
Size
Descending
5
Group other values in separate bucket
Show missing values
Custom label
Top five descriptors
Advanced
X-axis Created Date per y...
Add
Buckets
Split series Descriptor.key...
X-axis
Sub aggregation
Date Histogram help
Date Histogram
Field
Created Date
Minimum interval
Yearly
Select an option or create a custom value.
Examples: 30s, 20m, 24h, 2d, 1w, 1M
Custom label
Time (year)
Advanced
Add
Metrics
Count
Value axis
LeftAxis-1
Chart type
Mode
Line
Normal
Show line
Line mode
Line width
Straight
3
Show dots
Y-axes
LeftAxis-1 Count
Position
Left
Mode
Normal
Scale type
Linear
Show axis lines and labels
Title
Count of Calls
Labels
Show labels
Filter labels
Align
Truncate
Horizontal
100
Custom extents
Show dots
Y-axes
LeftAxis-1 Count
X-axis
Position
Bottom
Show axis lines and labels
Labels
Show labels
Filter labels
Align
Truncate
100
Settings
Legend position
Right
Show tooltip
Current time marker
Grid
Show X-axis lines
Y-axis lines
Don't show
Threshold line
Show threshold line

Actual visual

Title: The Trend of Top 5 Descriptors during the Last 10 Years in Brooklyn



Analysis:

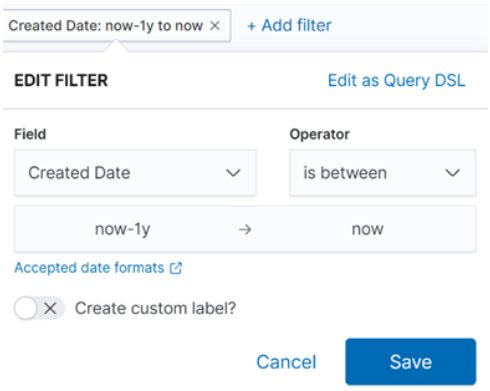
This line chart shows that 'Loud music/party' has been the primary descriptor in Brooklyn during the past 10 years. One interesting finding from this visual is that 'Request large bulky item collection' showed up in 2017, which surged significantly during the following 5 years and then dropped suddenly. In addition, the rest of the three descriptors ('Enter building', 'No access', and 'Apartments only') have remained relatively stable from 2014 to 2022.

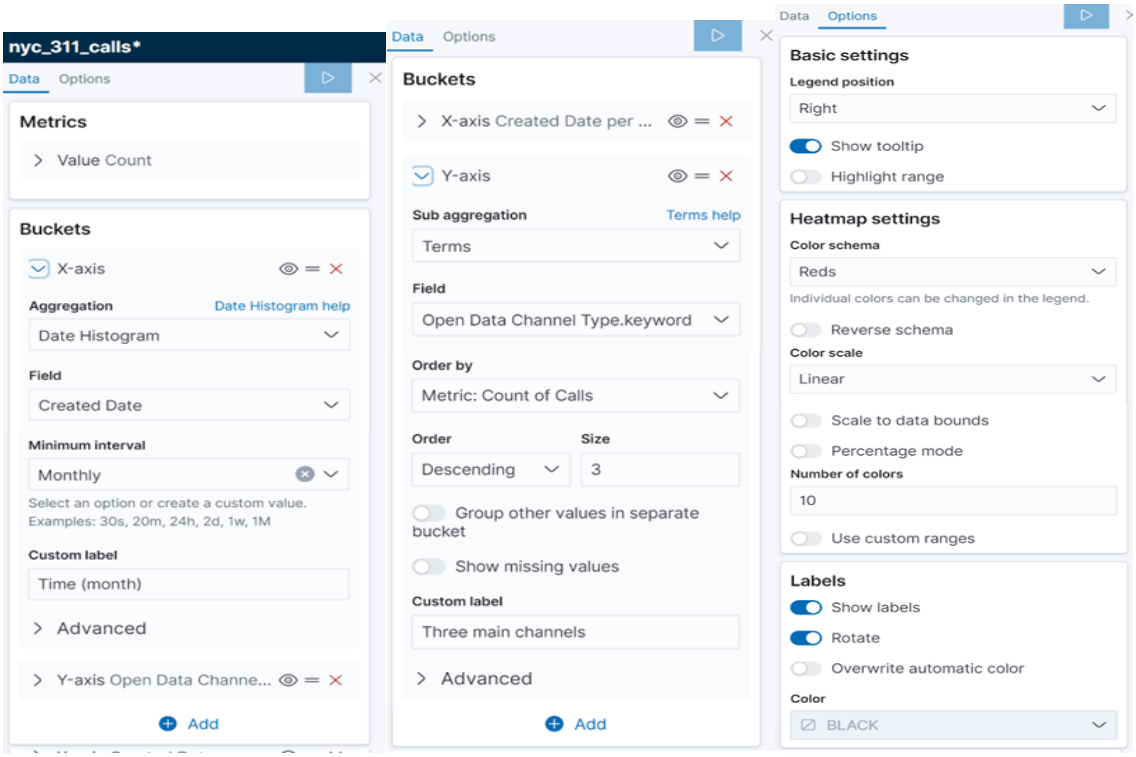
Extra Visualization 2:

Idea of Visual

This visual aims to examine the traffic fluctuation of calls from three main channel types during the past 12 months.

How to get the visual





Actual visual

Title: Calls via the 3 Main Channels in the Last 12 Months

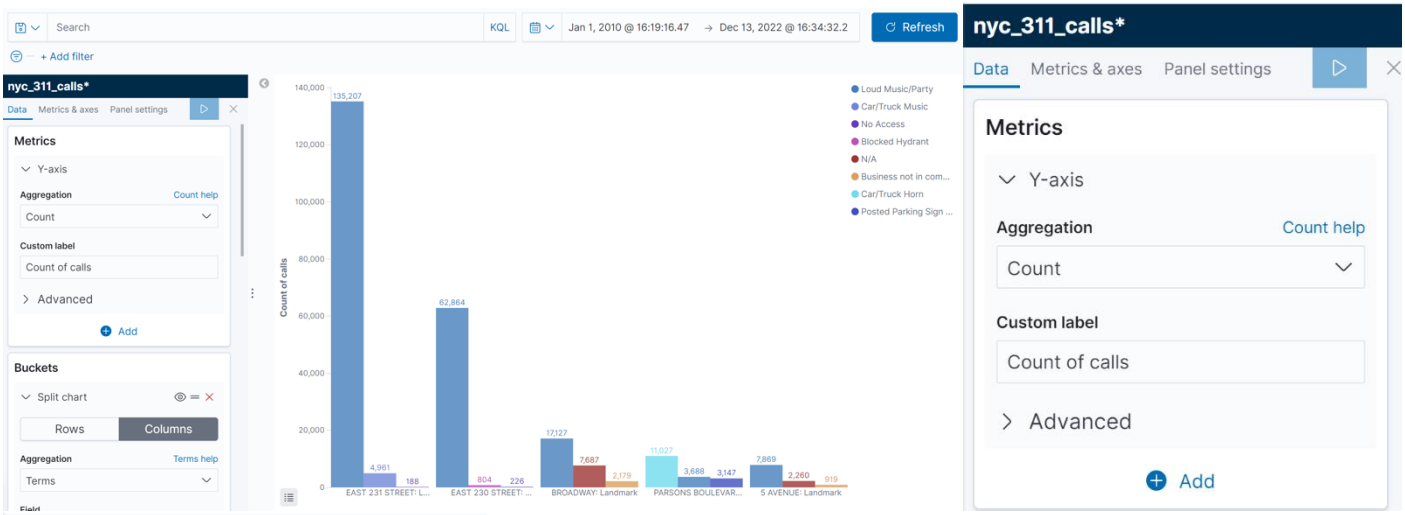
Analysis: The above heat map demonstrates the number of calls via the three main channels by color. The deepest red color means the highest number of calls, whereas the lightest red means the opposite. According to the visual, the 'Online' channel accounts for the most calls, followed by 'Phone' and 'Mobile' When examining the traffic change within a single channel, it shows January is the month with the highest number of calls via the 'Online' and 'Phone' channels in the past 12 months, whereas July is that of the 'Mobile' channel.

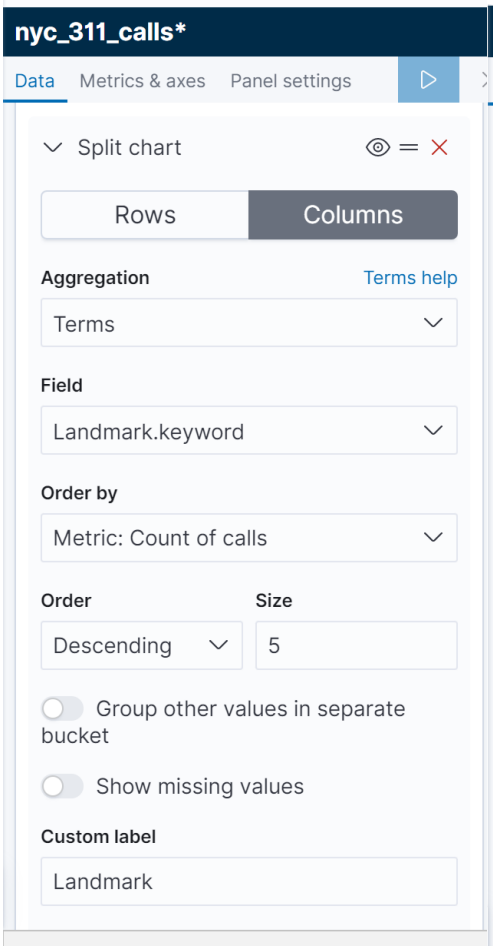
Extra Visualization 3:

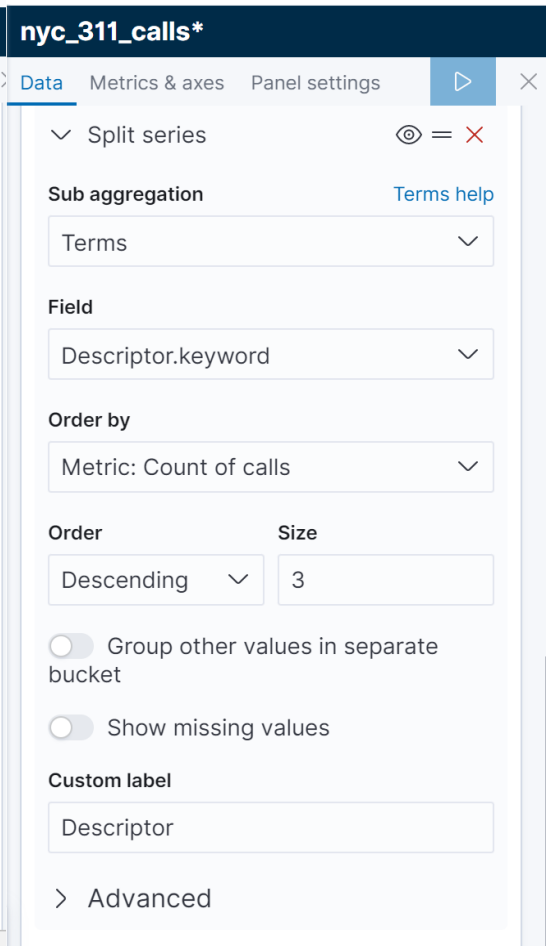
Idea of Visual

This visual aim is to look at the top 5 landmarks and their top 3 descriptors.

How to get the visual

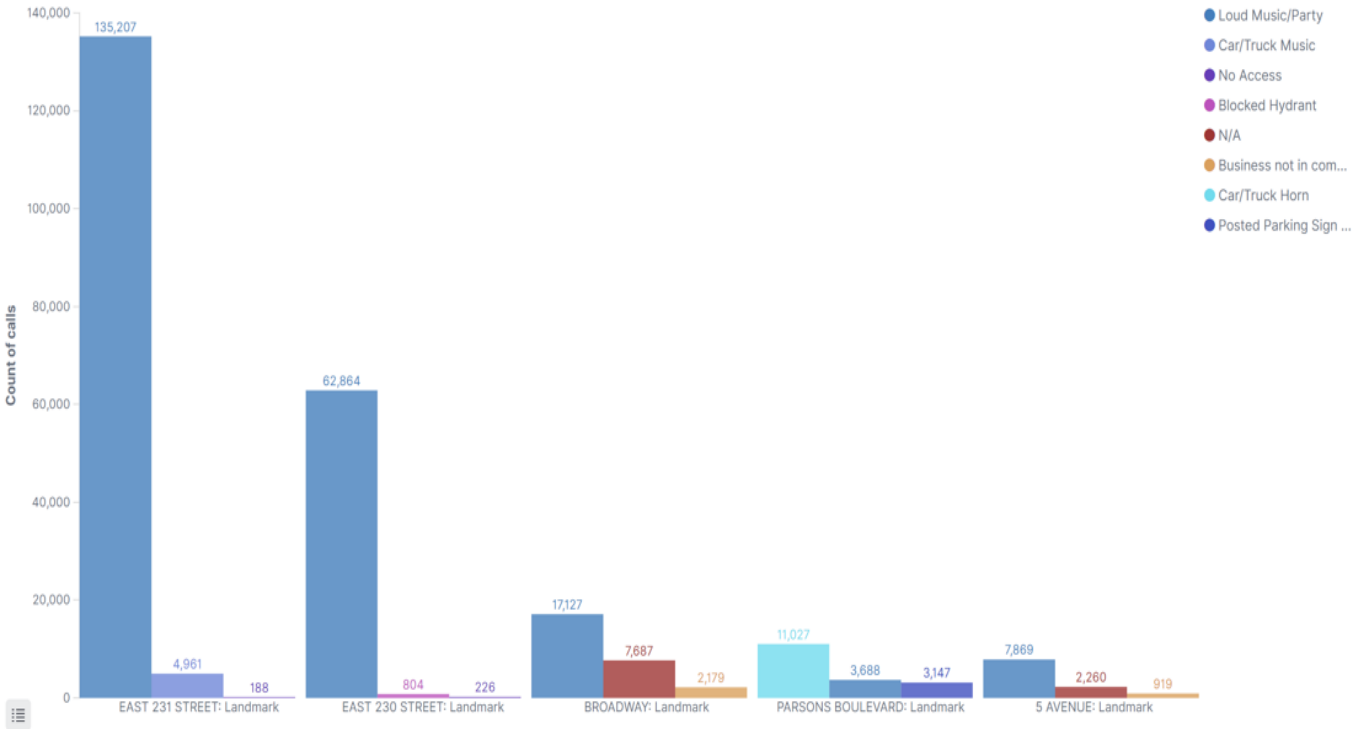






Actual visual

Title: Top 5 landmarks and their top 3 descriptors



Analysis: The top 5 landmarks are East 231 street, East 230 street, Broadway, Parsons Boulevard and 5 Avenue. For East 231 street we can see the top 3 descriptors for this location was loud music/party, car/truck music and no access. East 230 street top 3 descriptors were loud music/party, blocked hydrant and no access. Broadway's top 3 were loud music/party, N/A and Business not in compliance. Parsons boulevards top 3 were car/truck horn, loud music/party and posted parking sign violation. Lastly, 5 Avenue top 3 were loud music/party, N/A and Business not in compliance.

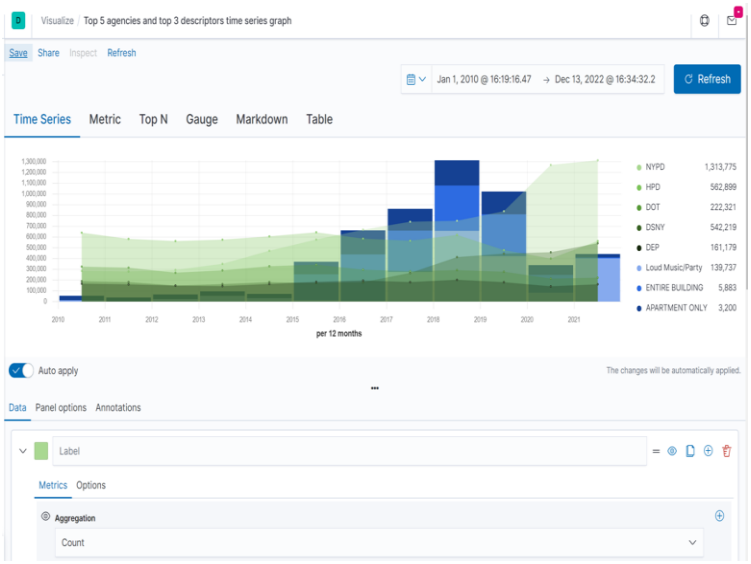
From this we can see that loud music/party was in all locations.

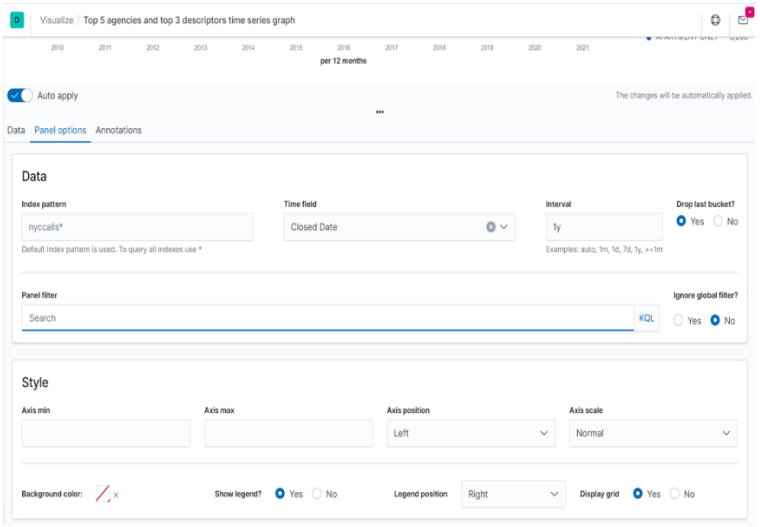
Extra Visualization 4:

Idea of Visual

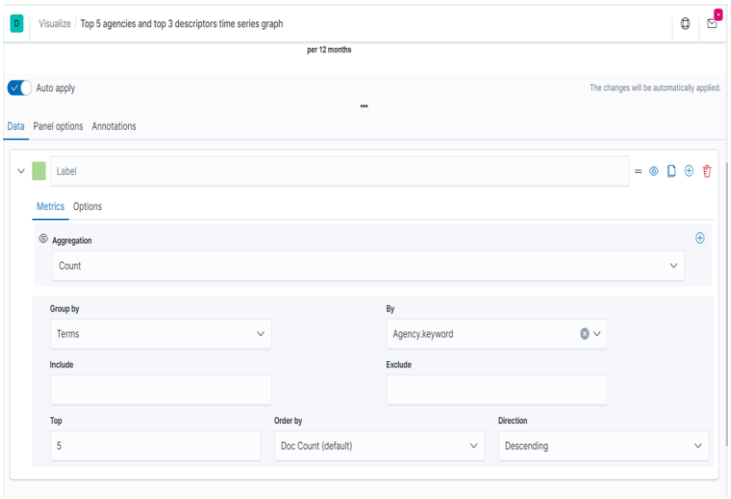
This visual aim is to look at the top 5 agencies over time with the top 3 descriptors.

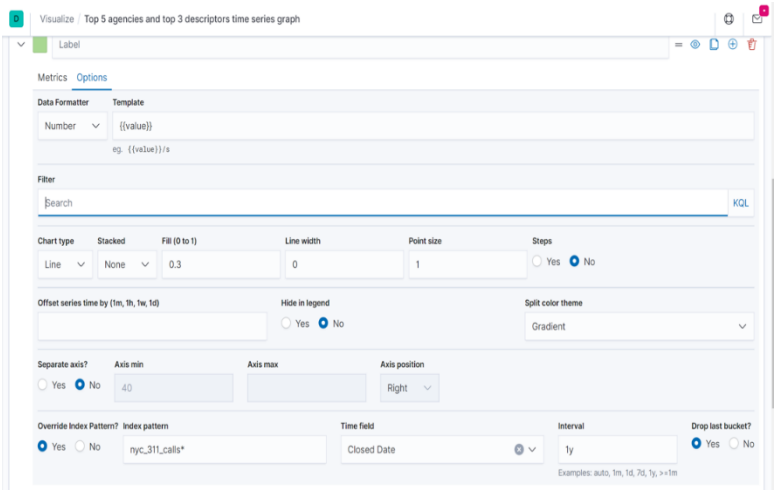
How to get the visual



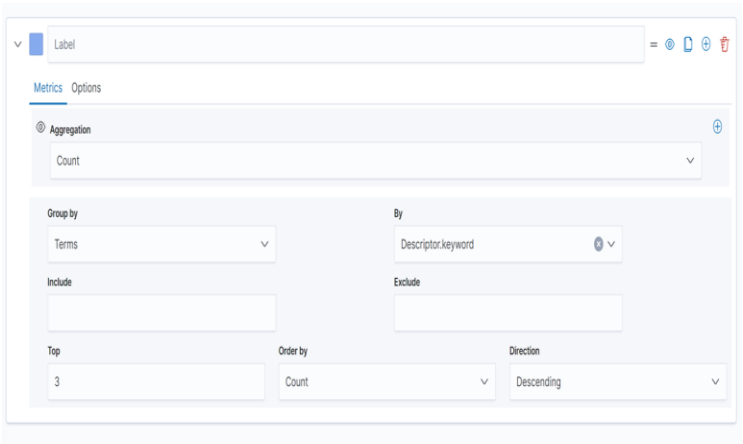


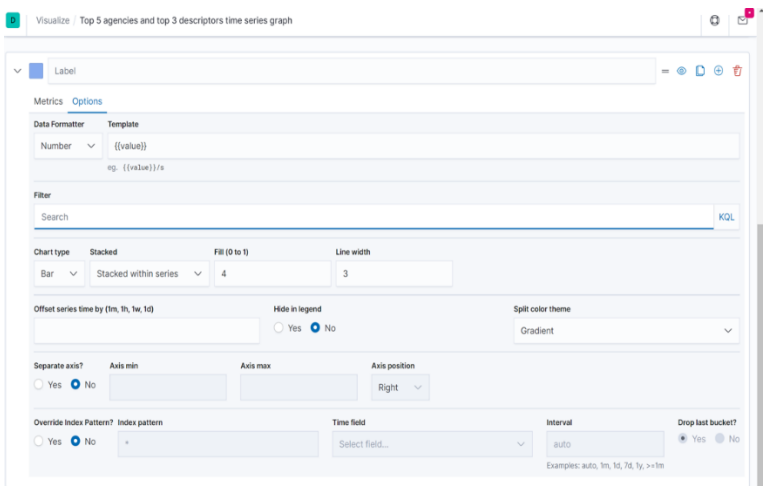
Agency settings





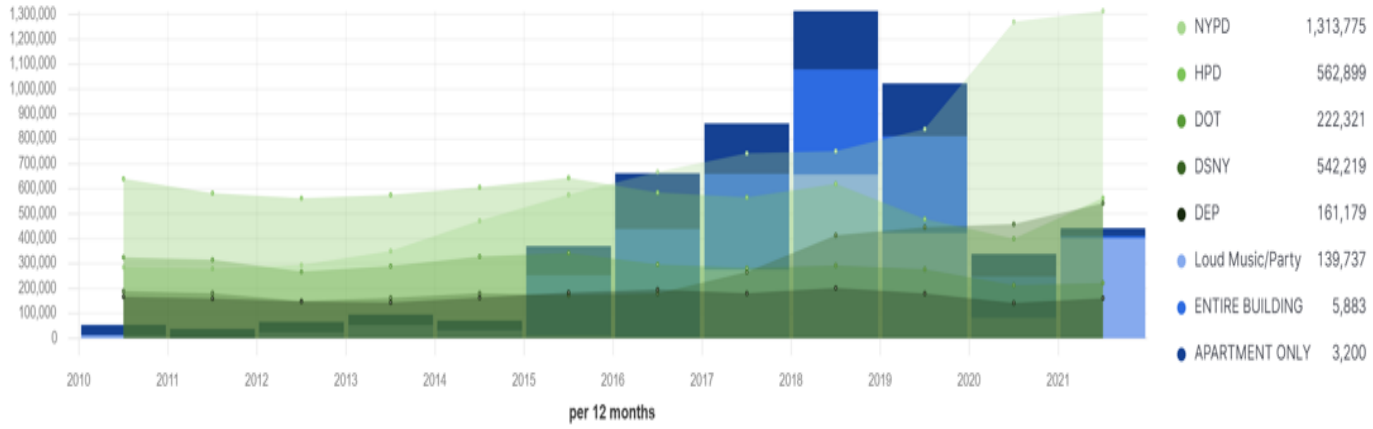
Descriptor settings





Actual visual

Title: Top 5 agencies with the top 3 descriptors over the last 11 years



Analysis: Over the last 11 years we can see that NYPD has been taking the most calls. Furthermore, most calls seem to come from Loud Music/Party (which has been shown in the above analysis as well).