

## Getting started on Elastic Cloud with a Sample Dataset “”

### Objective:

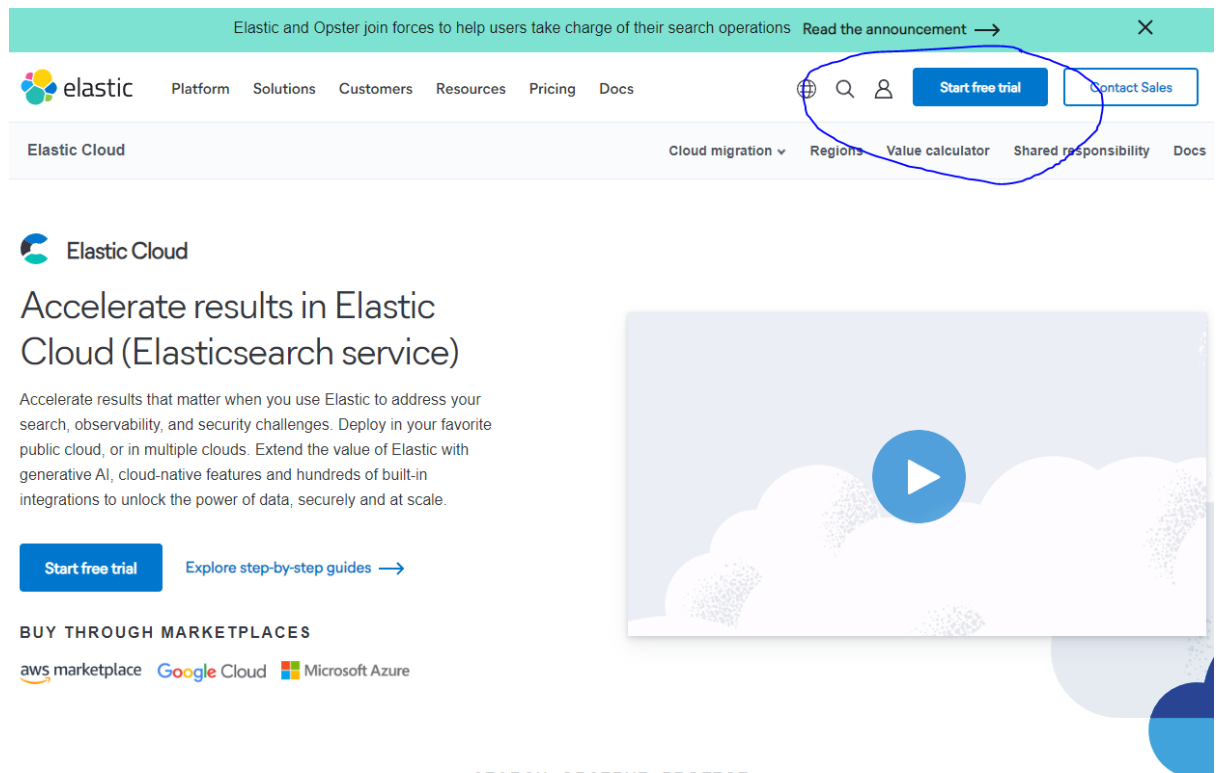
- Visit and create an account for Elastic Cloud
- Implement data sets
- Create Simple Graphs Visualizations

### Platform Specifications:

- ElasticSearch & Kibana
- Set size to 1 GB RAM
- Set nodes per zone to 1

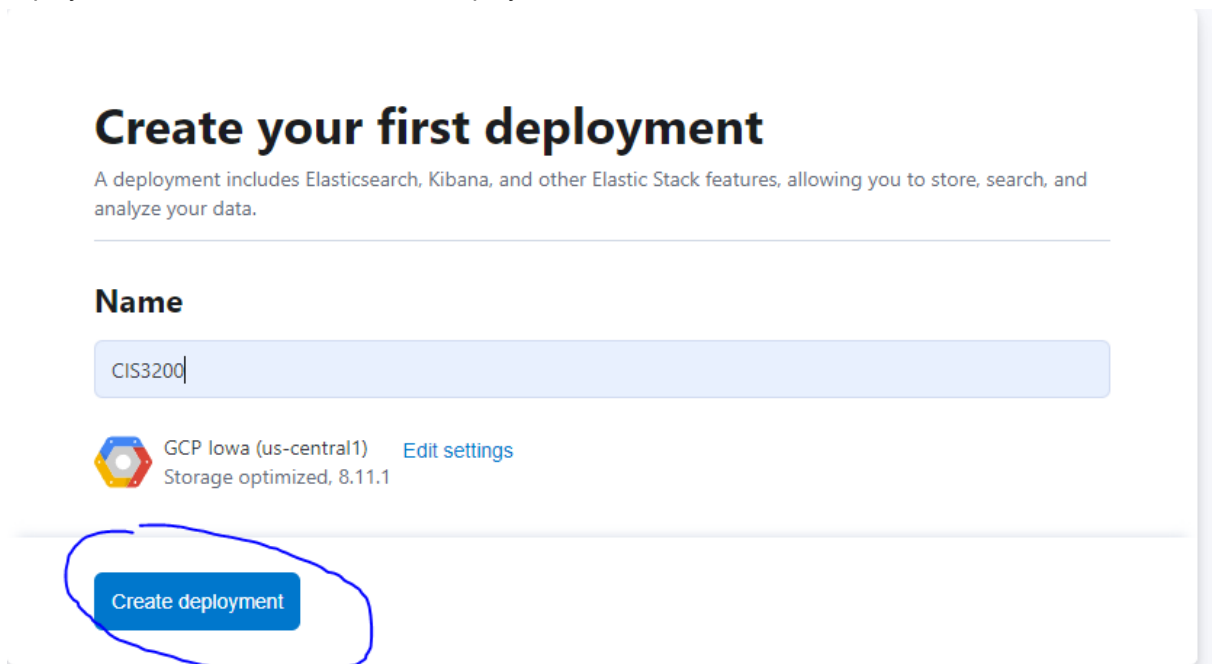
### Step 1 Create an account for Elastic Cloud:

1. Go to <https://www.elastic.co/cloud/elasticsearch-service/signup>
2. Click the option said **Start free trial**

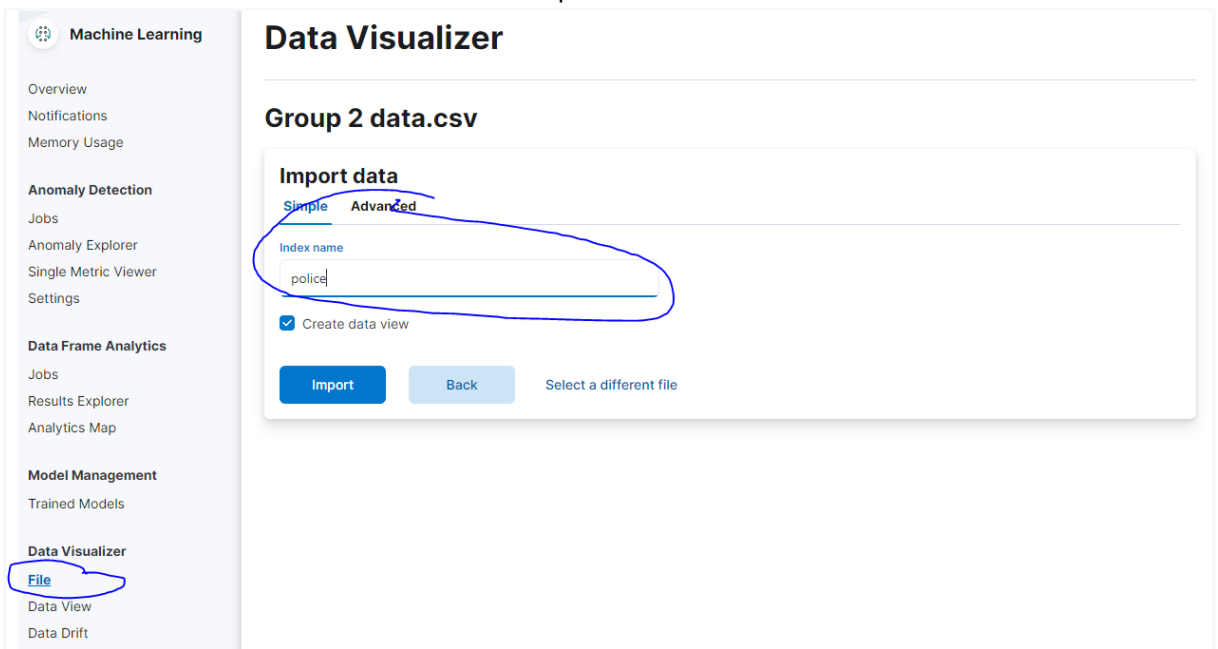


3. Next, you can register and Log into your Elastic Cloud account by using your school email.

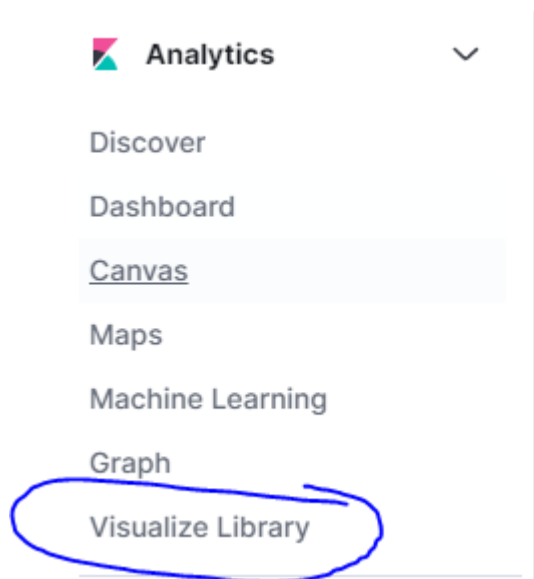
- Once you have signed in you will see an page asking your name and company. Name your deployment and click on the "Create Deployment" button.



- Next, you will need to download the data sets from [https://github.com/ZehaoChen3200/CIS-3200-group-2/blob/main/Q2\\_2019\\_SFPD\\_Data.csv](https://github.com/ZehaoChen3200/CIS-3200-group-2/blob/main/Q2_2019_SFPD_Data.csv) and <https://github.com/ZehaoChen3200/CIS-3200-group-2/blob/main/SFPolice-2020Q2.csv>
- Go back to the home page of Elastic Cloud and click the **Machine Learning and go to the Data visualizer->Files**. Set the Index name to police.



- Click on "Visualize" tab to start creating the Visualization for the data, then click the Create visualization and choose Vertical Bar in the Aggregation based.



## New visualization



### Lens

Create visualizations with our drag and drop editor. Switch between visualization types at any time. *Recommended for most users.*



### Maps

Create and style maps with multiple layers and indices.



### TSVB

Perform advanced analysis of your time series data.



### Custom visualization

Use Vega to create new types of visualizations. *Requires knowledge of Vega syntax.*



### Aggregation based

Use our classic visualize library to create charts based on aggregations.

[Explore options →](#)

### Tools



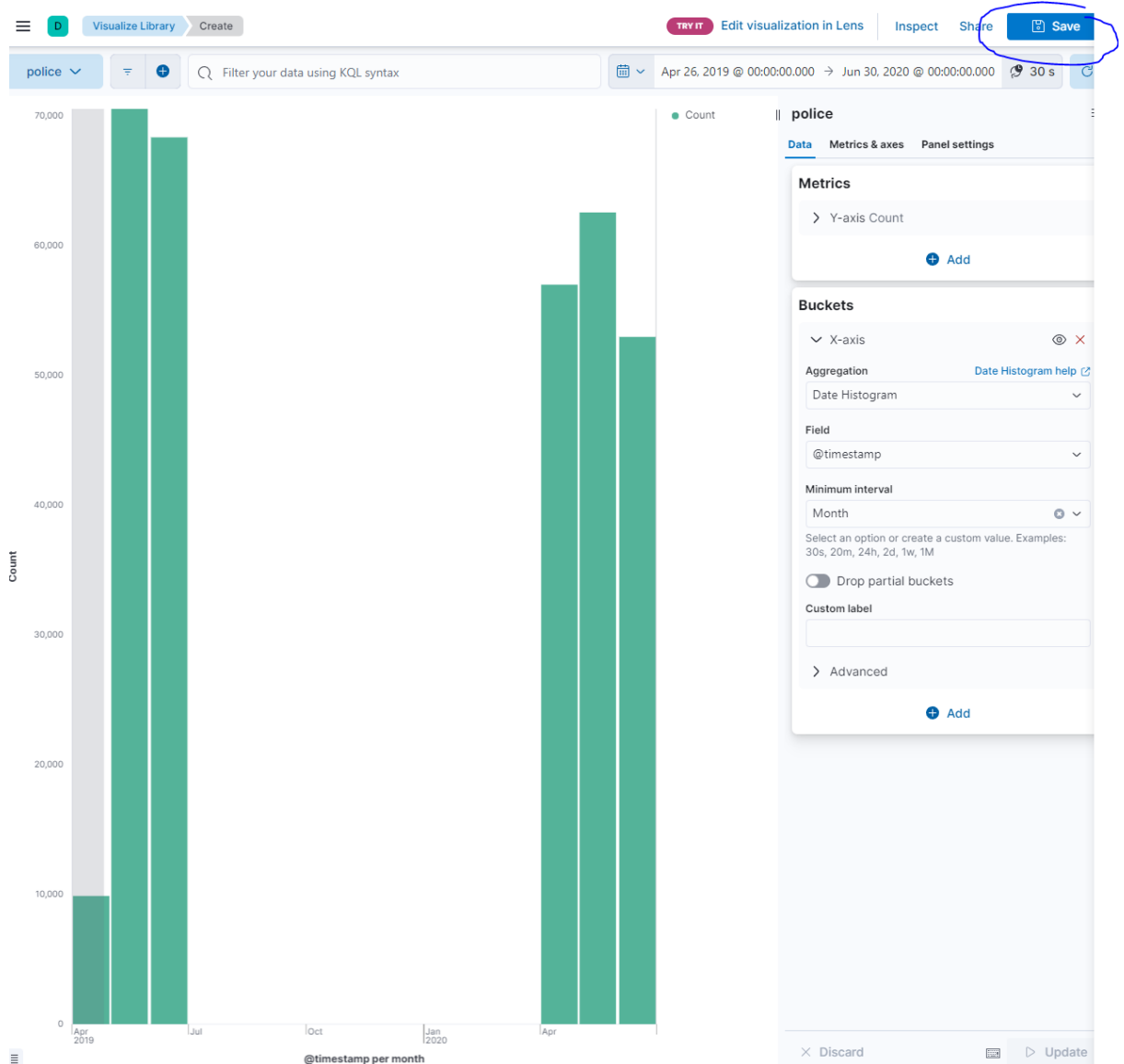
#### Text

Add text and images to your dashboard.

Want to learn more? [Read documentation](#) [↗](#)

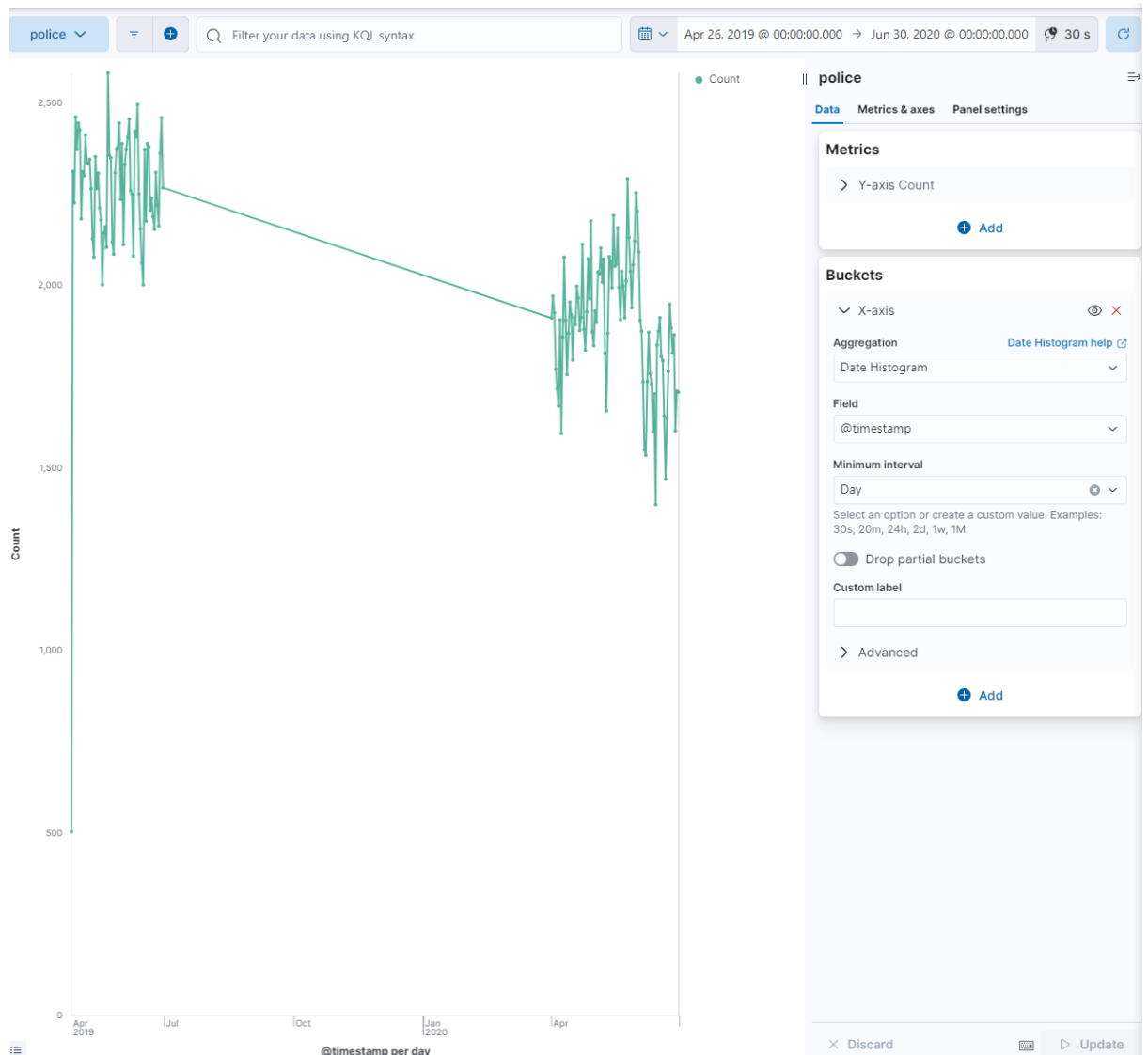
8. At the New Vertical Bar, click on the index pattern name "police".
9. Next, set the time range from Apr.24 2019, 00:00:00, to Jun 30 2020, 00:00:00. and then add a X-axis into bucket and set Aggregation->Date Histogram, Field->timestamp, ->minimum

interval ->Month. You can see the difference between Q2 2019 and Q2 2020. Save it.



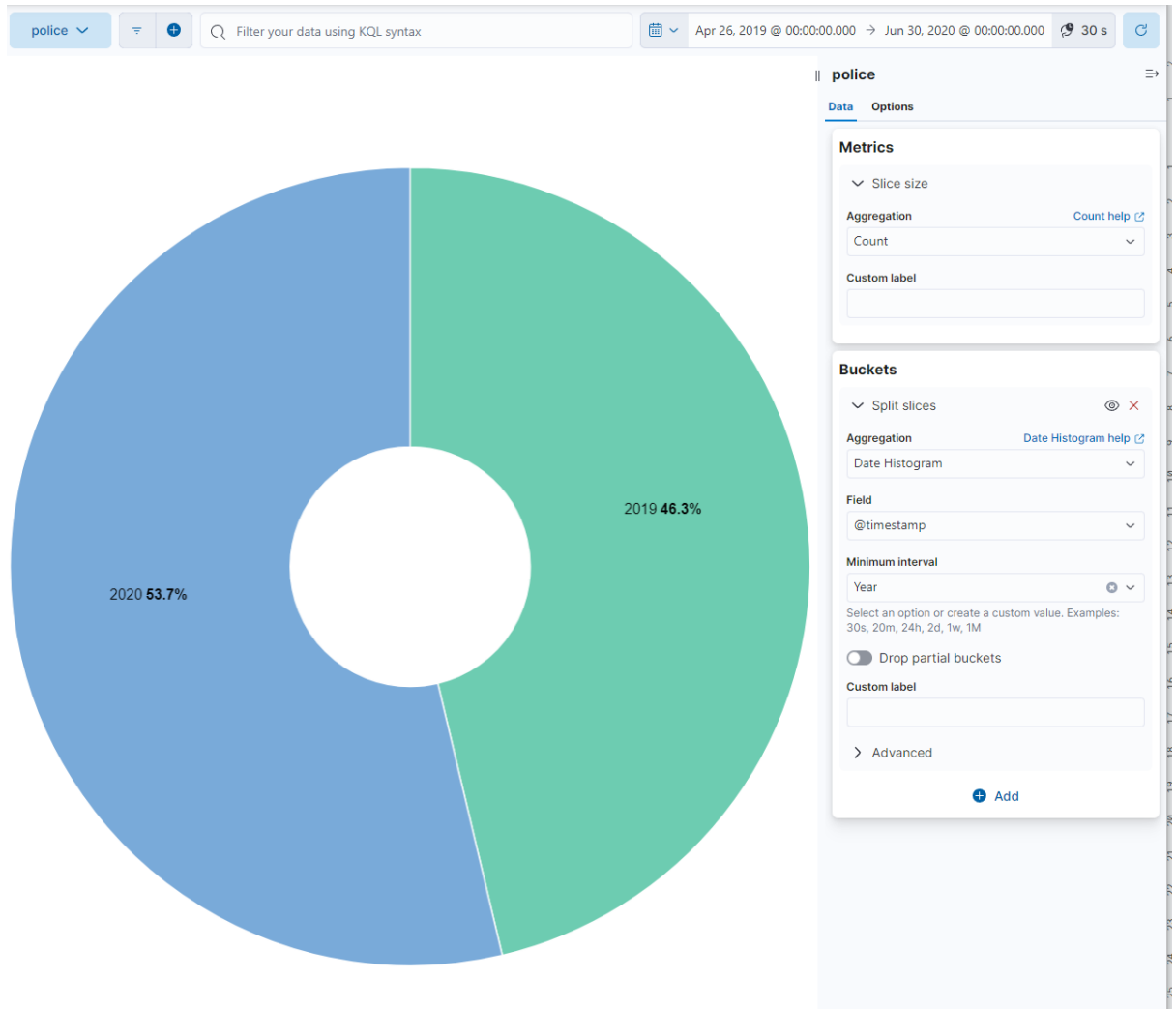
10. Next, go back to the visualize library and click the Create visualization again. Choose Line in the Aggregation based. Then click on the index pattern name "police".
11. Set the time range from Apr.24 2019, 00:00:00, to Jun 30 2020, 00:00:00. and then add a X-axis into bucket and set Aggregation->Date Histogram, Field->timestamp, ->minimum

interval ->Day. Save it.

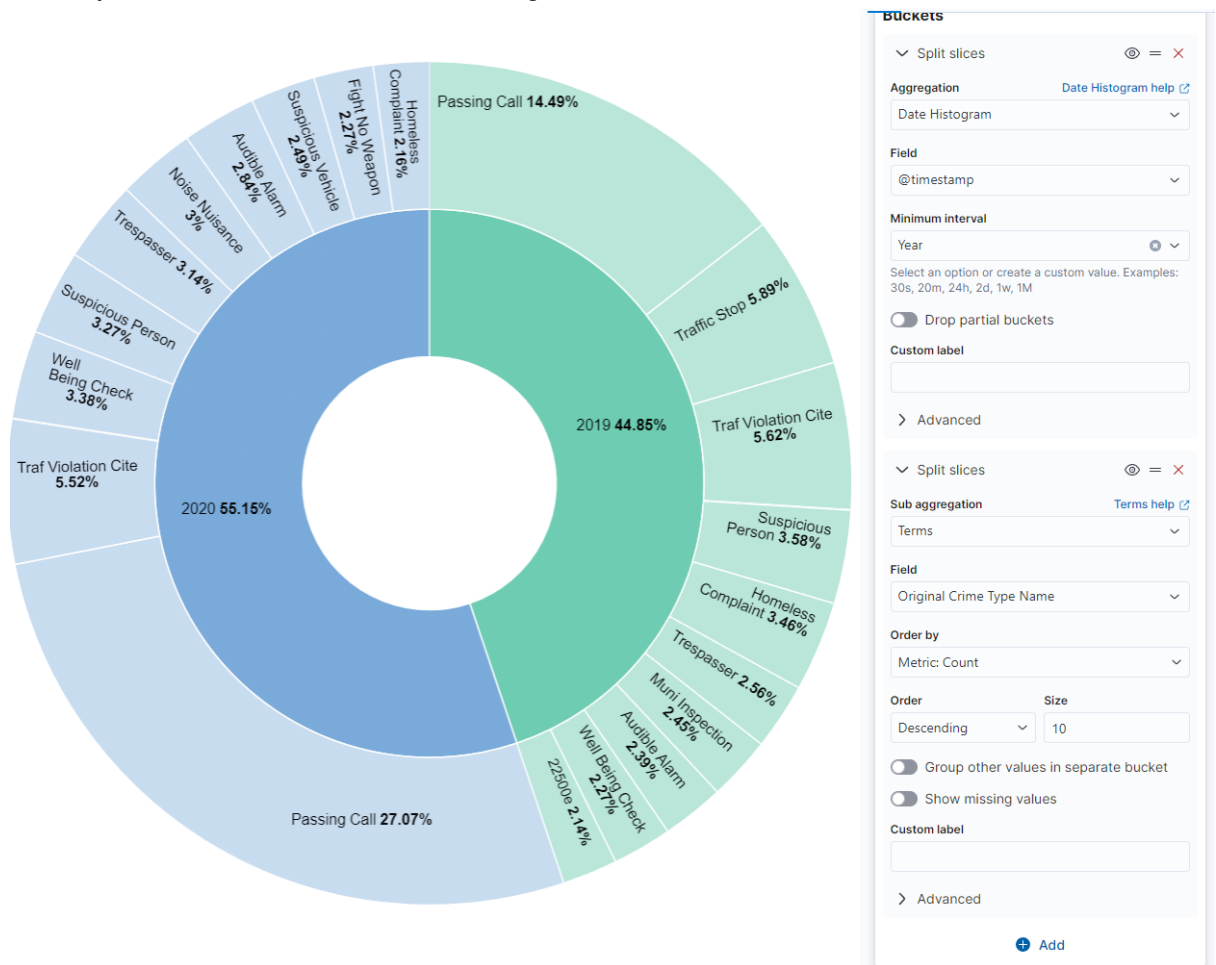


12. Next, go back to the visualize library and click the Create visualization again. Choose Pie in the Aggregation based. Then click on the index pattern name "police".
13. Set the time range from Apr.24 2019, 00:00:00, to Jun 30 2020, 00:00:00. and then add a Bucket for Split slices. Set Aggregation->Date Histogram, Field->timestamp, ->minimum

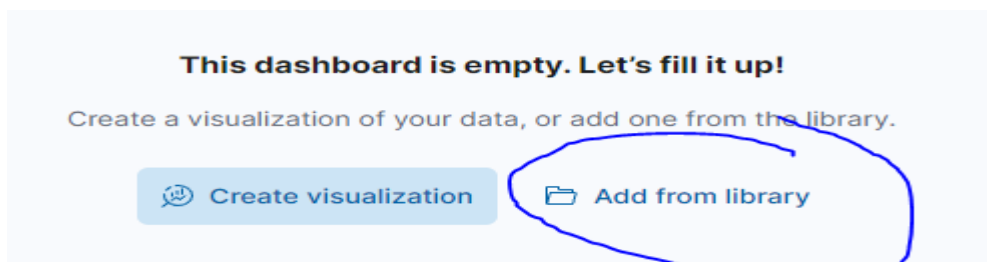
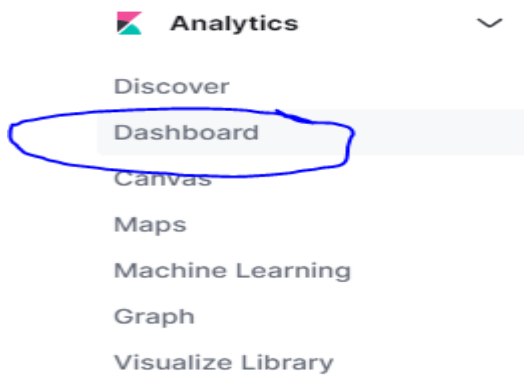
interval ->year.



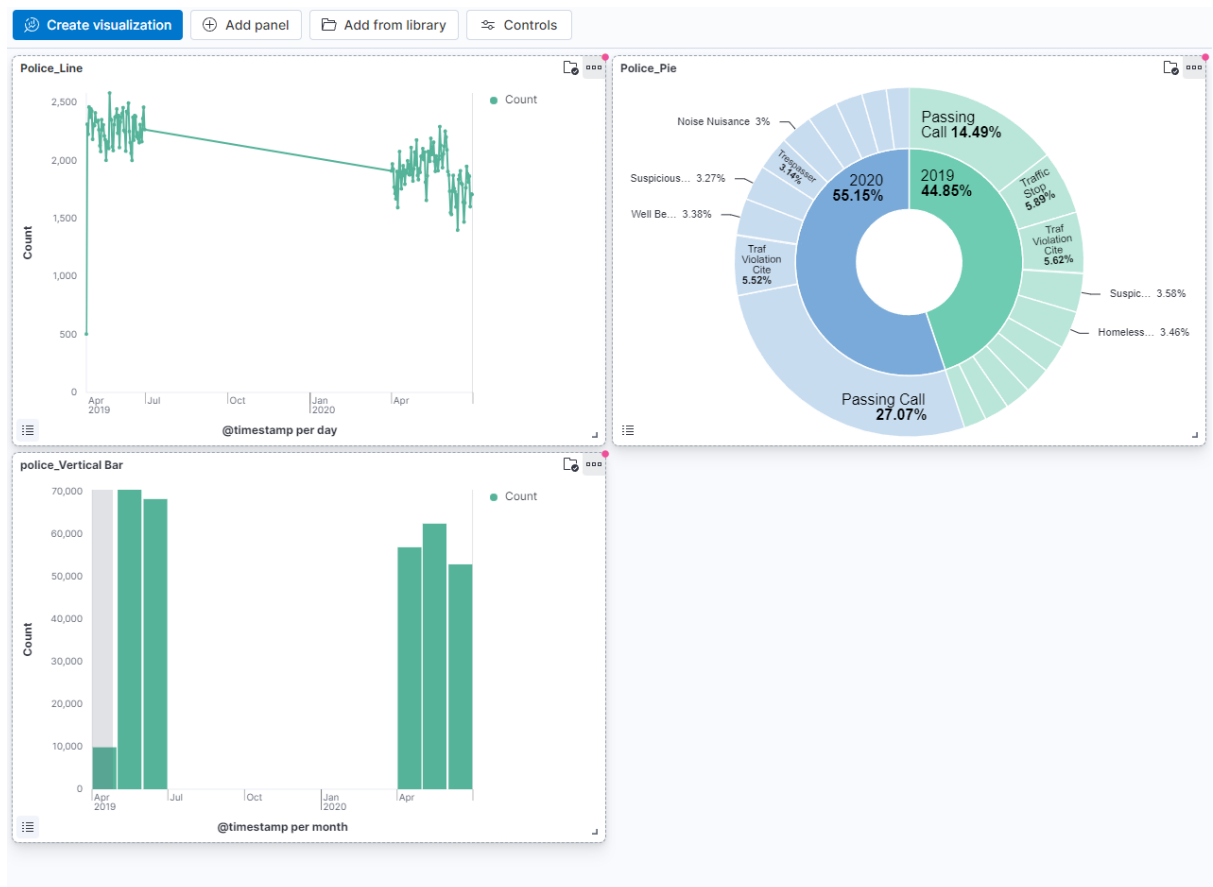
14. Add another split slices. Set sub aggregation->Terms, Field->Original Crime Type Name, Order by-> Metric:Count, Order->Descending, Size->10.Save it.



15. Click the dashboard and click the Create dashboard. Choose add from library and select all the chart you made in previous steps.



16. You will see the page similar to the following: Click Save




17. Go back to Machine Learning. Data Frame Analytics->job, click Create job, choose “police” which is the data we input in previous step.



18. Click Classification and set the Dependent variable to “Address type”. Included fields are address type, call time and city. Click continue


1 Configuration



### Outlier detection

Identify unusual data points in the data set.


Select



### Regression

Predict numerical values in the data set.

Select



### Classification

Predict classes of data points in the data set.

✓ Selected

Query

🔍 Search for e.g. method : "GET" or status : "404"

Runtime fields

No runtime field

🔍 Edit runtime fields

police

5 columns hidden Sort fields Histogram charts

@timestamp	Address	Address Type	Agency Id	Call Date	Call Date Time	Call Time
Jun 30, 2020 @ 00:00:0...	100 Block Of 34th Av	Common Location		1 Jun 29, 2020 @ 17:00:0...	Jun 30, 2020 @ 16:58:0...	23:58
Jun 30, 2020 @ 00:00:0...	400 Block Of Paris St	Premise Address		1 Jun 29, 2020 @ 17:00:0...	Jun 30, 2020 @ 16:51:0...	23:51
Jun 30, 2020 @ 00:00:0...	1300 Block Of Polk St	Premise Address		1 Jun 29, 2020 @ 17:00:0...	Jun 30, 2020 @ 16:57:0...	23:57
Jun 30, 2020 @ 00:00:0...	Fell St/clayton St	Intersection		1 Jun 29, 2020 @ 17:00:0...	Jun 30, 2020 @ 16:56:0...	23:56
Jun 30, 2020 @ 00:00:0...	16th St/market St	Intersection		1 Jun 29, 2020 @ 17:00:0...	Jun 30, 2020 @ 16:55:0...	23:55

Rows per page: 5

< 1 2 3 4 5 ... 2000 >

Dependent variable

Address Type

Included fields

3 fields included in the analysis

🔍 is\_included:true

Is included Is not included

Field name	Mapping	Is included	Is required	Reason
<input type="checkbox"/> Address	text	Yes	No	
<input checked="" type="checkbox"/> Address Type	keyword	Yes	Yes	
<input type="checkbox"/> Agency Id	long	Yes	No	
<input checked="" type="checkbox"/> Call Time	keyword	Yes	No	
<input checked="" type="checkbox"/> City	keyword	Yes	No	

Rows per page: 5

< 1 2 >

19. Set the model memory limit to 153mb. Set jobID to policemachinelearning and the index name should be same as your jobID. Click create.

1

Configuration

Source index  
police

Job type  
classification

Dependent variable  
Address Type

Query  
--

Training percent  
80

Included fields  
Address Type, Call Time, City

Edit

2

Additional options

Advanced configuration

Top feature importance values  
0

Prediction field name  
Address Type\_prediction

Model memory limit  
153mb

Maximum number of threads  
1

Randomized seed  
--

Top classes  
All classes

Hyperparameters

Lambda  
--

Max trees  
--

Gamma  
--

Eta  
--

Feature bag fraction  
--

Edit

3

Job details

Job ID  
policemachinelearning

Job description

Destination index  
policemachinelearning

Edit

4

Validation

Successful checks  
3 ✓

Warnings  
1 ⚠

View

5

Create

☒ Start immediately  
If unselected, job can be started later by returning to the jobs list.

☒ Create data view

Create

20. You will got the result like the following:

