

Analysing Cars Data using Elastic Search, Log Stash and Kibana

1. Start the Elastic Search

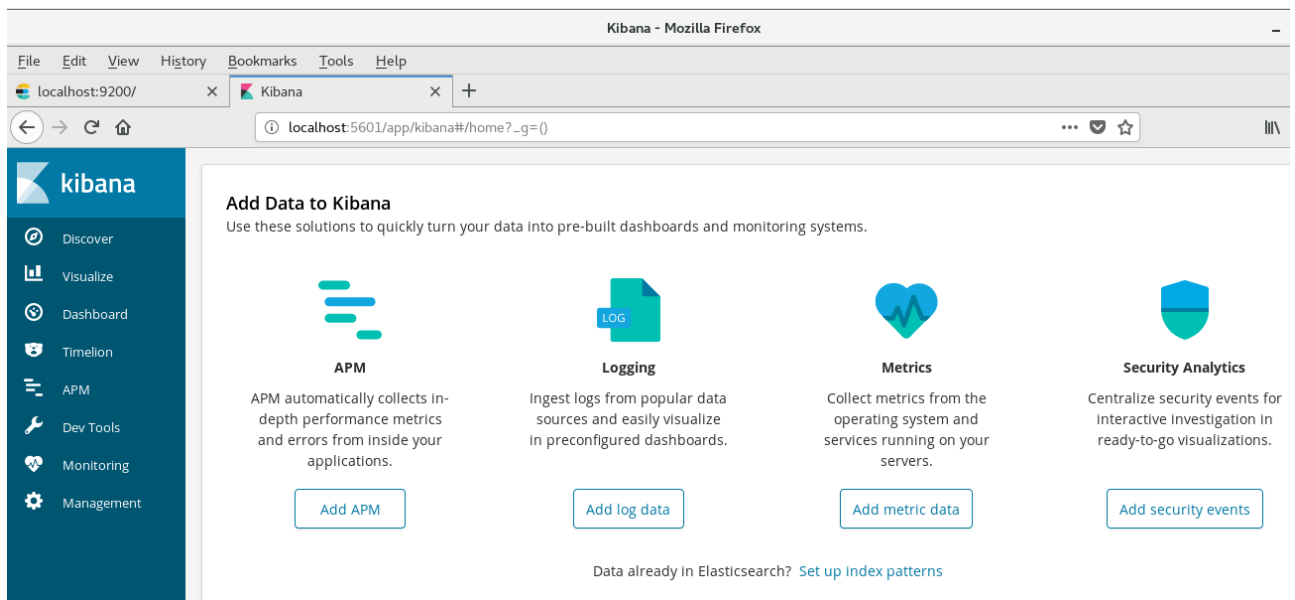
```
$ /usr/local/elasticsearch-6.3.2/bin/elasticsearch
```



Elastic search is up and running at <http://localhost:9200>

2. Start the Kibana

```
$ /usr/local/kibana-6.3.2-linux-x86_64/bin/kibana
```



Kibana is up and running at <http://localhost:5601>

3. Input Data:

Prepare some data in CSV. For this example data is stored in the following file.

```
/home/ashok/Desktop/Examples/Data/Cars.csv
```

```
ashok@learning:~/Desktop/Examples/Data
File Edit View Search Terminal Help
[ashok@learning Data]$ head -15 Cars.csv
maker,model,mileage,manufacture_year,engine_displacement,engine_power,body_type,color_slug,stk_year,transmission,door_count
,seat_count,fuel_type,date_created,date_last_seen,price_eur
ford,galaxy,151000,2011,2000,103,,None,man,5,7,diesel,2015-11-14 18:10:06.838319+00,2016-01-27 20:40:15.46361+00,10584.75
skoda,octavia,143476,2012,2000,81,,None,man,5,5,diesel,2015-11-14 18:10:06.853411+00,2016-01-27 20:40:15.46361+00,8882.31
bmw,,97676,2010,1995,85,,None,man,5,5,diesel,2015-11-14 18:10:06.861792+00,2016-01-27 20:40:15.46361+00,12065.06
skoda,fabia,111970,2004,1200,47,,None,man,5,5,gasoline,2015-11-14 18:10:06.872313+00,2016-01-27 20:40:15.46361+00,2960.77
skoda,fabia,128886,2004,1200,47,,None,man,5,5,gasoline,2015-11-14 18:10:06.880335+00,2016-01-27 20:40:15.46361+00,2738.71
skoda,fabia,140932,2003,1200,40,,None,man,5,5,gasoline,2015-11-14 18:10:06.894643+00,2016-01-27 20:40:15.46361+00,1628.42
skoda,fabia,167220,2001,1400,74,,None,man,5,5,gasoline,2015-11-14 18:10:06.915376+00,2016-01-27 20:40:15.46361+00,2072.54
bmw,,148500,2009,2000,130,,None,auto,5,5,diesel,2015-11-14 18:10:06.924123+00,2016-01-27 20:40:15.46361+00,10547.74
skoda,octavia,105389,2003,1900,81,,None,man,5,5,diesel,2015-11-14 18:10:06.936239+00,2016-01-27 20:40:15.46361+00,4293.12
,,301381,2002,1900,88,,None,man,5,5,diesel,2015-11-14 18:10:06.954319+00,2016-01-27 20:40:15.46361+00,1332.35
,,202136,2002,1400,55,,None,man,5,5,gasoline,2015-11-14 18:10:06.962458+00,2016-01-27 20:40:15.46361+00,740.19
,,263840,1998,1900,81,,None,man,5,5,diesel,2015-11-14 18:10:06.993167+00,2016-01-27 20:40:15.46361+00,999.26
,,105394,2000,1360,55,,None,man,3,5,gasoline,2015-11-14 18:10:07.036951+00,2016-01-27 20:40:15.46361+00,1665.43
skoda,favorit,41250,1990,1300,44,,None,man,5,5,gasoline,2015-11-14 18:10:07.051147+00,2016-01-27 20:40:15.46361+00,370.1
[ashok@learning Data]$
```

4. Write the Log Stash Configuration file.

```
File Edit View Search Terminal Tabs Help
Confluent x Elasticsearch x Kibana x Logstash x ashok@learning:~/Desktop... x
input {
  file {
    path => "/home/ashok/Desktop/Examples/Data/Cars.csv"
    # The below two are needed to read the data from the start of the file.
    # By default the file plugin points to the end of the file
    start_position => "beginning"
    sincedb_path => "/tmp/null"
  }
}
filter {
  csv {
    separator => ",",
    columns => [ "maker", "model", "mileage", "manufacture_year", "engine_displacement", "engine_power", "body_type", "color_slug", "stk_year", "t
ransmission", "door_count", "seat_count", "fuel_type", "date_created", "date_last_seen", "price_eur" ]
  }

  mutate { convert => ["mileage", "integer"] }
  mutate { convert => ["manufacture_year", "integer"] }
  mutate { convert => ["engine_power", "integer"] }
  mutate { convert => ["stk_year", "integer"] }
  mutate { convert => ["seat_count", "integer"] }
  mutate { convert => ["price_eur", "float"] }
}
output {
  elasticsearch {
    hosts => ["localhost:9200"]
    index => "cars"
    document_type => "used_cars"
  }
  stdout { codec => rubydebug }
}
}
```

Logstash cars.config

```
input {

  file {

    path => "/home/ashok/Desktop/Examples/Data/Cars.csv"

    # The below two are needed to read the data from the start of the file.

    # By default the file plugin points to the end of the file

    start_position => "beginning"

    sincedb_path => "/tmp/null"

  }

}

filter {
```

```

csv {

  separator => ","

  columns => [ "maker", "model", "mileage", "manufacture_year", "engine_displacement",
"engine_power", "body_type", "color_slug", "stk_year", "transmission", "door_count", "seat_count",
"fuel_type", "date_created", "date_last_seen", "price_eur" ]

}

mutate { convert => ["mileage", "integer"] }

mutate { convert => ["manufacture_year", "integer"] }

mutate { convert => ["engine_power", "integer"] }

mutate { convert => ["stk_year", "integer"] }

mutate { convert => ["seat_count", "integer"] }

mutate { convert => ["price_eur", "float"] }

}

output {

  elasticsearch {

    hosts => ["localhost:9200"]

    index => "cars"

    document_type => "used_cars"

  }

  stdout { codec => rubydebug }

}

```

5. List of Indices in Elastic Search before running Log Stash

The screenshot shows the Kibana console interface in a Mozilla Firefox browser. The console is displaying a list of indices in Elasticsearch. The left sidebar shows the Kibana navigation menu with options like Discover, Visualize, Dashboard, Timelion, APM, Dev Tools, Monitoring, and Management. The main console area shows a list of indices with columns for index name, status, type, and size. The indices listed are:

Index Name	Status	Type	Size
megacorp	open	yellow	13.6kb
logstash-2018.07.28	open	yellow	17.9kb
user	open	green	19.2kb
accounts	open	yellow	5.4kb
.kibana	open	green	24kb
job_status	open	yellow	73.3kb

The console also shows the following commands and their results:

```

GET user/profile/_search
{
  "query": {
    "match": {
      "last_name": "smith"
    }
  }
}

GET job_status/job/1533237539

GET _cat/indices

```

6. Run the Log Stash as follows:

Checking the configuration file:

```
$ logstash -f logstash_cars.config --config.text_and_exit
```

Running:

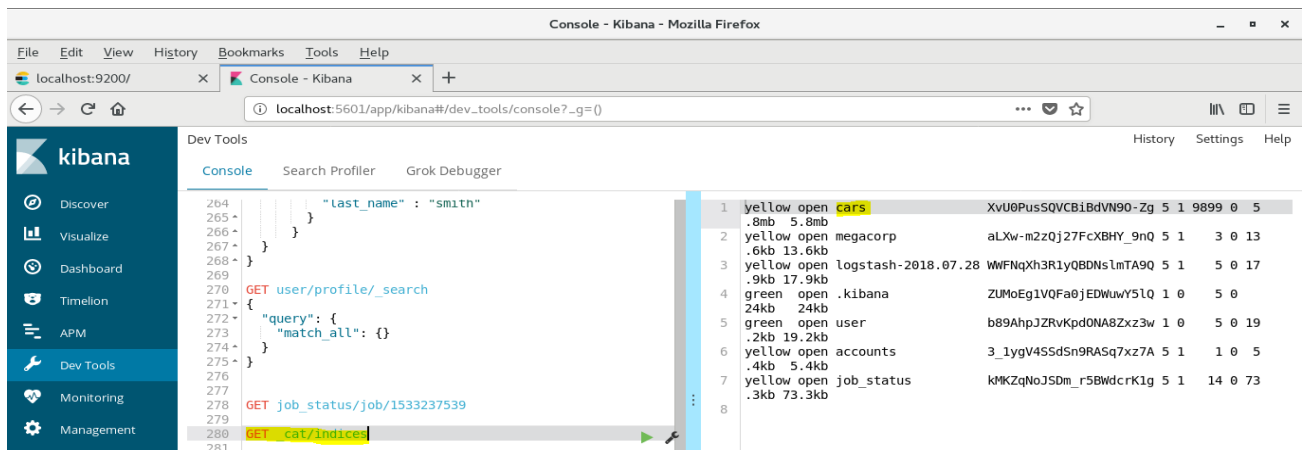
```
$ logstash -f logstash_cars.config
```

```
File Edit View Search Terminal Tabs Help
Elasticsearch x Kibana x Logstash x ashok@learning:~/Desktop/Examples... x
[ashok@learning LogStash Files]$
[ashok@learning LogStash Files]$ logstash -f logstash_cars.config
^C[ashok@learning LogStash Files]$ less logstash_cars.config
[ashok@learning LogStash Files]$ logstash -f logstash_cars.config
Sending Logstash's logs to /usr/local/logstash-6.3.2/logs which is now configured via log4j2.properties
[2018-08-11T19:32:53,256][WARN ][logstash.config.source.multilocal] Ignoring the 'pipelines.yml' file because modules or command line options are specified
[2018-08-11T19:32:54,316][INFO ][logstash.runner] Starting Logstash {"logstash.version"=>"6.3.2"}
[2018-08-11T19:32:58,816][WARN ][logstash.outputs.elasticsearch] You are using a deprecated config setting "document type" set in elasticsearch. Deprecated settings will continue to work, but are scheduled for removal from logstash in the future. Document types are being deprecated in Elasticsearch 6.0, and removed entirely in 7.0. You should avoid this feature. If you have any questions about this, please visit the #logstash channel on freenode irc. {:name=>"document type", :plugin=><LogStash::Outputs::ElasticSearch hosts=>[//localhost:9200], index=>"cars", document_type=>"used cars", id=>"1439ab10d0505cee487d65d53946cb93496b137d48205b1953ffea025c3445f8", enable_metric=>true, codec=><LogStash::Codecs::Plain id=>"plain_20c6465d-85b3-4d13-98ba-df706110b850", enable_metric=>true, charset=>"UTF-8">, workers=>1, manage_template=>true, template_name=>"logstash", template_overwrite=>false, docs_upsert=>false, script_type=>"inline", script_lang=>"painless", script_var_name=>"event", scripted_upsert=>false, retry_initial_interval=>2, retry_max_interval=>64, retry_on_conflict=>1, action=>"index", ssl_certificate_verification=>true, sniffing=>false, sniffing_delay=>5, timeout=>60, pool_max=>1000, pool_max_per_route=>100, resurrect_delay=>5, validate_after_inactivity=>10000, http_compression=>false>}
[2018-08-11T19:33:01,738][INFO ][logstash.pipeline] Starting pipeline {:pipeline_id=>"main", "pipeline.workers"=>1, "pipeline.batch.size"=>125, "pipeline.batch.delay"=>50}
[2018-08-11T19:33:02,522][INFO ][logstash.outputs.elasticsearch] Elasticsearch pool URLs updated {:changes=>{:removed=>[], :added=>[http://localhost:9200/]}}
[2018-08-11T19:33:02,529][INFO ][logstash.outputs.elasticsearch] Running health check to see if an Elasticsearch connection is working {:healthcheck_url=>http://localhost:9200/, :path=>"/"}
[2018-08-11T19:33:02,907][WARN ][logstash.outputs.elasticsearch] Restored connection to ES instance {:url=>"http://localhost:9200/" }
[2018-08-11T19:33:03,006][INFO ][logstash.outputs.elasticsearch] ES Output version determined {:es_version=>6}
[2018-08-11T19:33:03,009][WARN ][logstash.outputs.elasticsearch] Detected a 6.x and above cluster: the 'type' event field won't be used to determine the document type {:es_version=>6}
[2018-08-11T19:33:03,096][INFO ][logstash.outputs.elasticsearch] New Elasticsearch output {:class=>"LogStash::Outputs::ElasticSearch", :hosts=>[//localhost:9200/]}
```

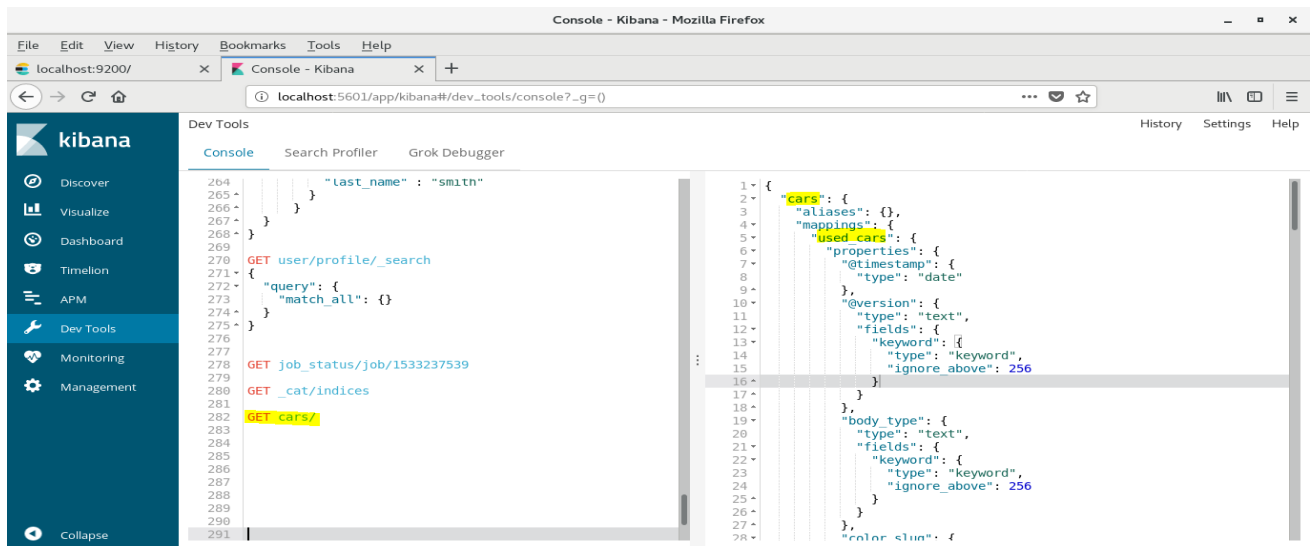
Sample record loading into Elastic Search:

```
Logstash
File Edit View Search Terminal Tabs Help
Elasticsearch x Kibana x Logstash x ashok@learning:~/Desktop/Examples... x
{"stk_year" => nil,
 "transmission" => nil,
 "fuel_type" => "gasoline",
 "date_created" => "2017-03-16 17:27:17.781459+00",
 "maker" => "peugeot",
 "door_count" => nil,
 "date_last_seen" => "2017-03-16 17:27:17.781459+00",
 "price_eur" => 1295.34
}
{
  "@version" => "1",
  "@timestamp" => 2018-08-11T15:17:50.336Z,
  "engine_power" => 80,
  "seat_count" => nil,
  "body_type" => "other",
  "color_slug" => nil,
  "path" => "/home/ashok/Desktop/Examples/Data/Cars.csv",
  "host" => "learning.bigdata.com",
  "engine_displacement" => nil,
  "message" => "peugeot,307,,2003,,80,other,,,,,electric,2017-03-16 17:27:18.511154+00,2017-03-16 17:27:18.511154+00,1295.34\r",
  "manufacture_year" => 2003,
  "model" => "307",
  "mileage" => nil,
  "stk_year" => nil,
  "transmission" => nil,
  "fuel_type" => "electric",
  "date_created" => "2017-03-16 17:27:18.511154+00",
  "maker" => "peugeot",
  "door_count" => nil,
  "date_last_seen" => "2017-03-16 17:27:18.511154+00",
  "price_eur" => 1295.34
}
```

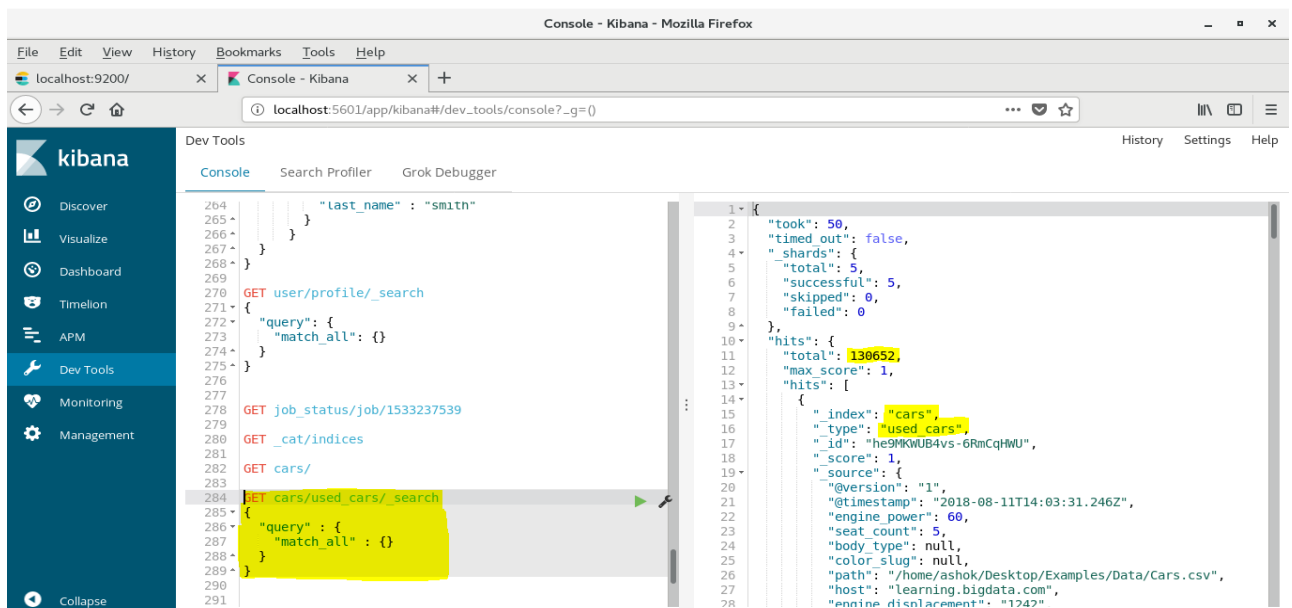
7. After launching Log Stash, a new Index is created in Elastic search



Description of the Index (CARS)



Able to the see some records in Elastic Search



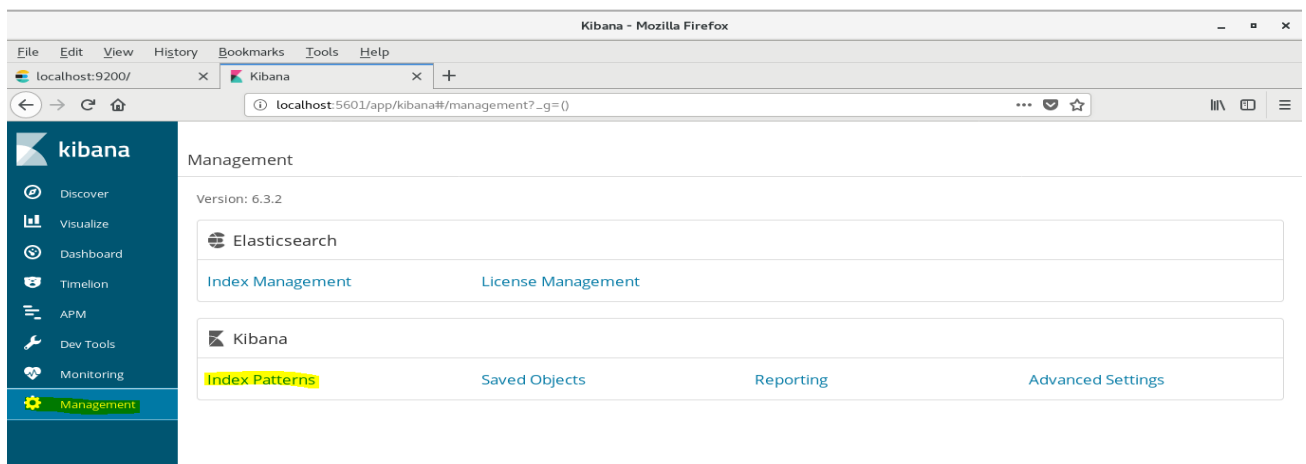
NOTE: Log Stash is still running and loading data into Elastic Search

Count on index CARS:

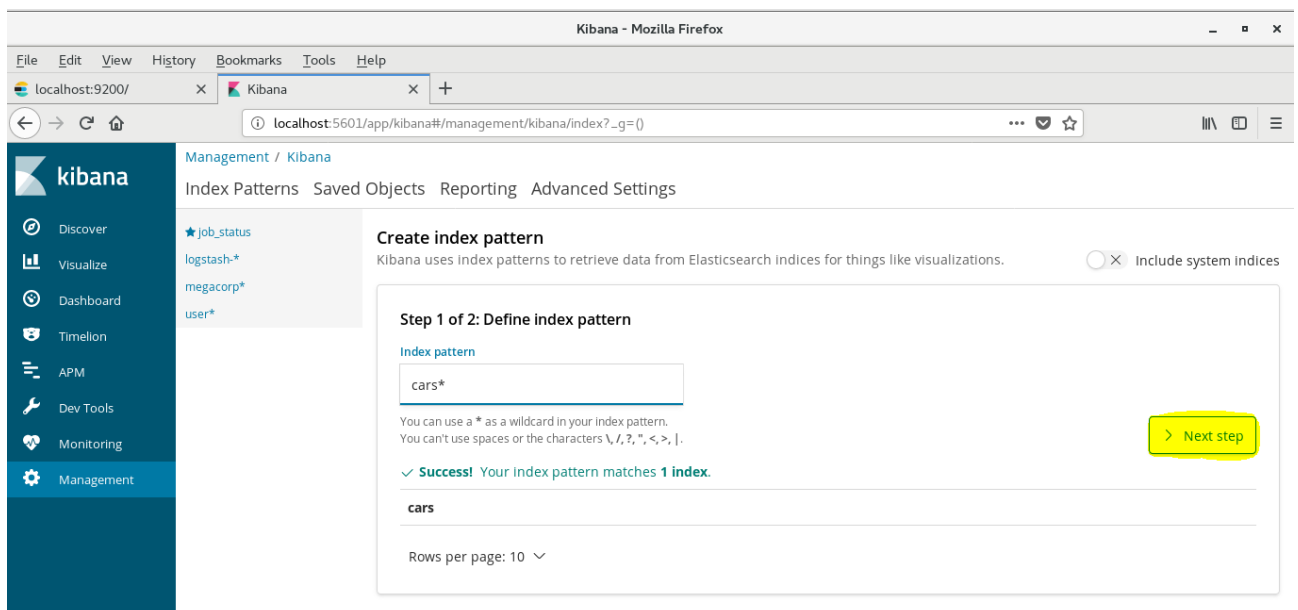


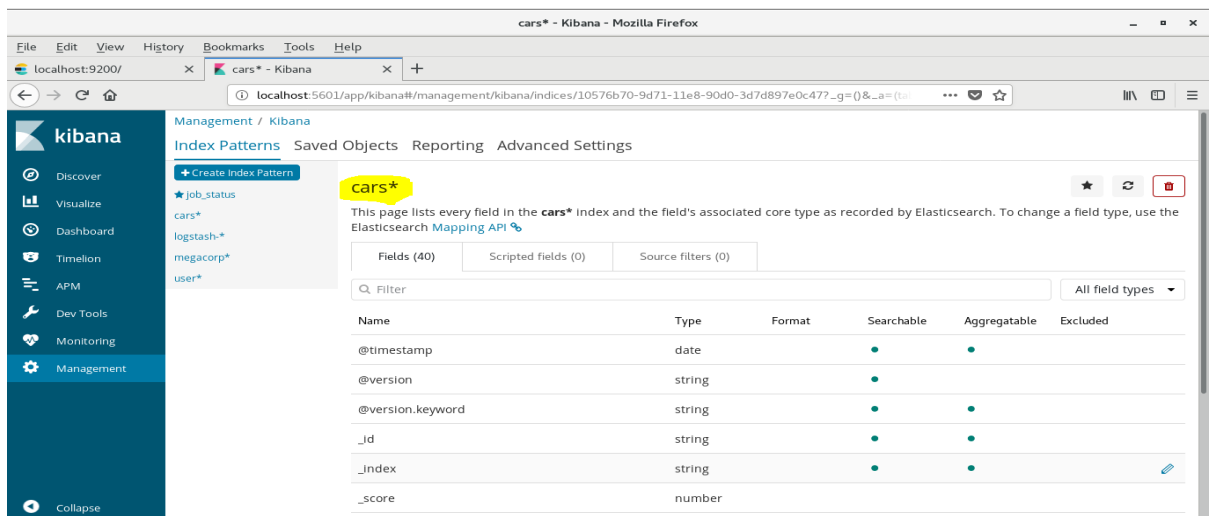
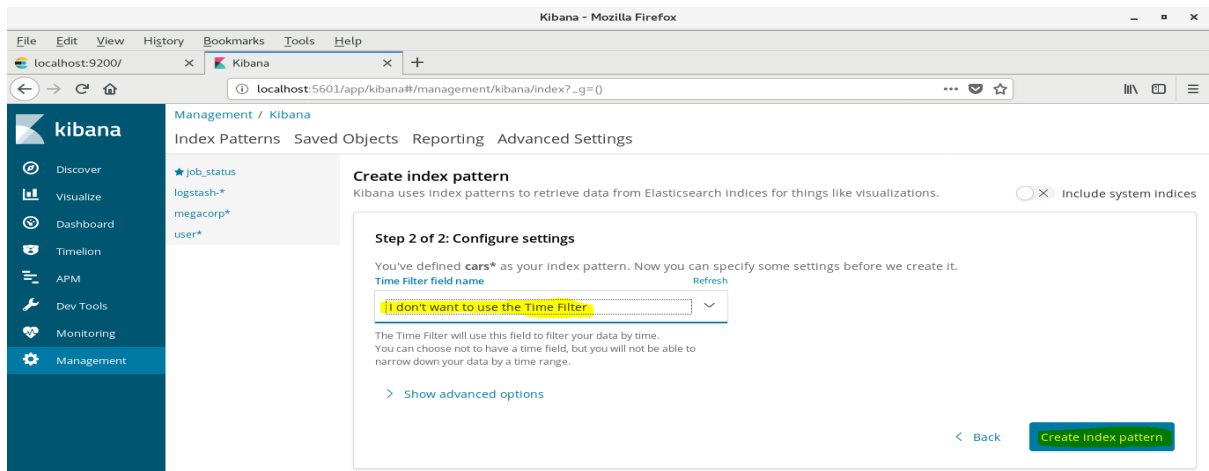
8. Creating an Index Pattern in Kibana:

In Monitoring Section, create the Index Pattern



Creating Index Pattern:

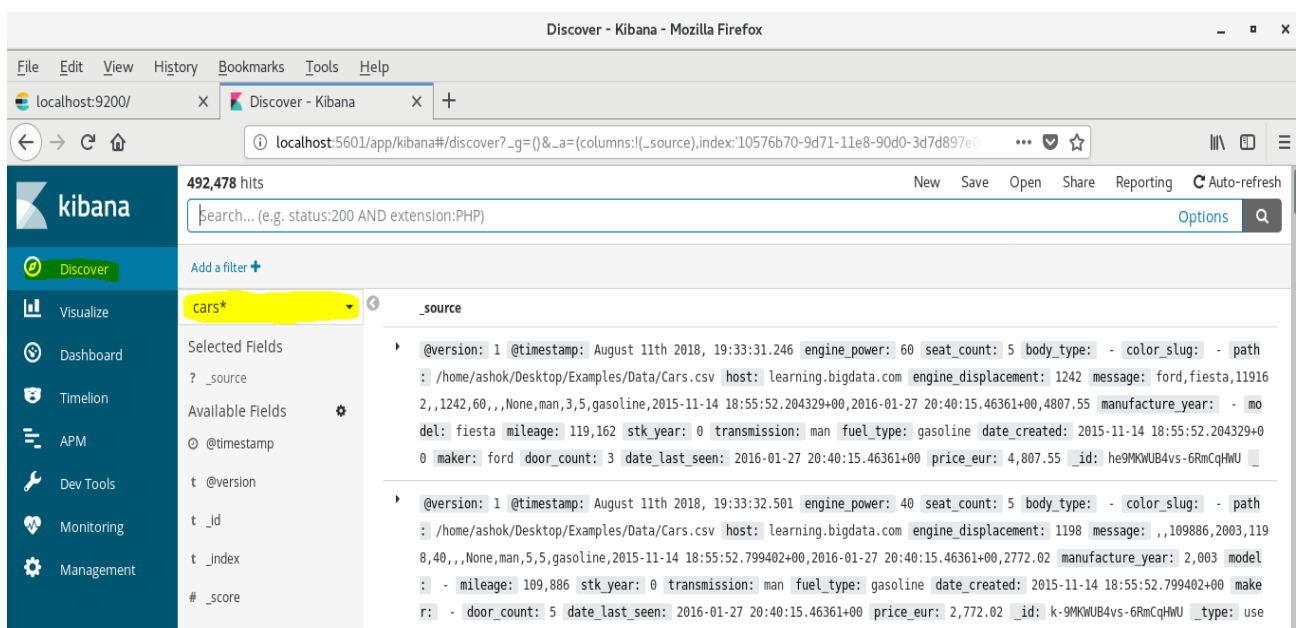




CARS Index Pattern is created.

9. Discovering Data in Kibana:

Then the data can be seen in Discover tab by selecting the Index Pattern



Can see each record as tabular format or JSON

The screenshot shows the Kibana Discover interface in a Mozilla Firefox browser. The address bar shows the URL: `localhost:5601/app/kibana#/discover?_g=()&_a=(columns:!(._source),index:'10576b70-9d71-11e8-90d0-3d7d897e0...')`. The left sidebar contains the Kibana navigation menu with options like Discover, Visualize, Dashboard, Timelion, APM, Dev Tools, Monitoring, and Management. The main panel is divided into three sections: a search bar with the query `cars*`, a list of available fields, and a table of search results. The table has columns for various fields including `@timestamp`, `@version`, `_id`, `_index`, `#_score`, `_type`, `body_type`, `color_slug`, `date_created`, `date_last_seen`, `door_count`, `engine_displacement`, and `engine_power`. The first row of data shows a car record with details like `engine_power: 60`, `seat_count: 5`, `body_type: -`, `color_slug: -`, `path: /home/ashok/Desktop/Examples/Data/Cars.csv`, `host: learning.bigdata.com`, `engine_displacement: 1242`, `message: ford,fiesta,119162,,1242,60,,,None,man,3,5,gasoline,2015-11-14 18:55:52.204329+00,2016-01-27 20:40:15.46361+00,4807.55`, `manufacture_year: -`, `model: fiesta`, `mileage: 119,162`, `stk_year: 0`, `transmission: man`, `fuel_type: gasoline`, `date_created: 2015-11-14 18:55:52.204329+00`, `maker: ford`, `door_count: 3`, `date_last_seen: 2016-01-27 20:40:15.46361+00`, `price_eur: 4,807.55`, and `_id: he9MKWUB4vs-6RmCqHwU`. Below the table, there is a button to 'View single document'.

The screenshot shows the Kibana Discover interface in a Mozilla Firefox browser, similar to the first one but with the view set to JSON. The address bar shows the same URL. The left sidebar is identical. The main panel shows the search results in a JSON format. The first row of data is a JSON object representing the car record, with fields like `"index": "cars"`, `"type": "used_cars"`, `"id": "he9MKWUB4vs-6RmCqHwU"`, `"version": 1`, `"score": 0`, and `"source": {` followed by the same data as the first row in the tabular view. Below the JSON view, there is a button to 'View single document'.

Can select only the required columns for all the records as follows:

Click on the symbol for the required fields:

Discover - Kibana - Mozilla Firefox

localhost:9200/ Discover - Kibana

localhost:5601/app/kibana#/discover?_g=()&_a=(columns:!((_source),index:'10576b70-9d71-11e8-90d0-3d7d897e0...))

kibana

- Discover
- Visualize
- Dashboard
- Timelion
- APM
- Dev Tools
- Monitoring
- Management
- Collapse

Field	Value
# _score	0
t _type	used_cars
t body_type	-
t color_slug	-
t date_created	2015-11-14 18:55:52.204329+00
t date_last_seen	2016-01-27 20:40:15.46361+00
t door_count	3
t engine_displacement	1242
# engine_power	60
t fuel_type	gasoline
t host	learning.bigdata.com
t maker	ford
# manufacture_year	-
t message	ford,fiesta,119162,,1242,60,,,None,man,3,5,gasoline,2015-11-14 18:55:52.204329+00,2016-01-27 20:40:15.46361+00,4807.55
# mileage	119,162
t model	fiesta
t path	/home/ashok/Desktop/Examples/Data/Cars.csv
# price_eur	4,807.55
# seat_count	5
# stk_year	0

Then

Discover - Kibana - Mozilla Firefox

localhost:9200/ Discover - Kibana

localhost:5601/app/kibana#/discover?_g=()&_a=(columns:!(maker,manufacture_year,model,price_eur),index:'10576b70-9d71-11e8-90d0-3d7d897e0...))

kibana

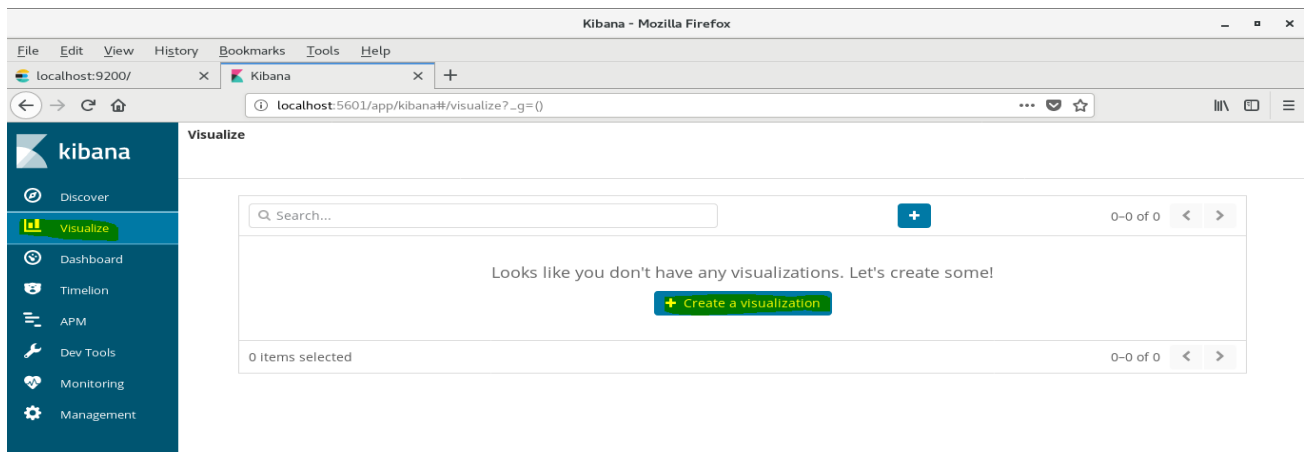
- Discover
- Visualize
- Dashboard
- Timelion
- APM
- Dev Tools
- Monitoring
- Management
- Collapse

Add a filter +

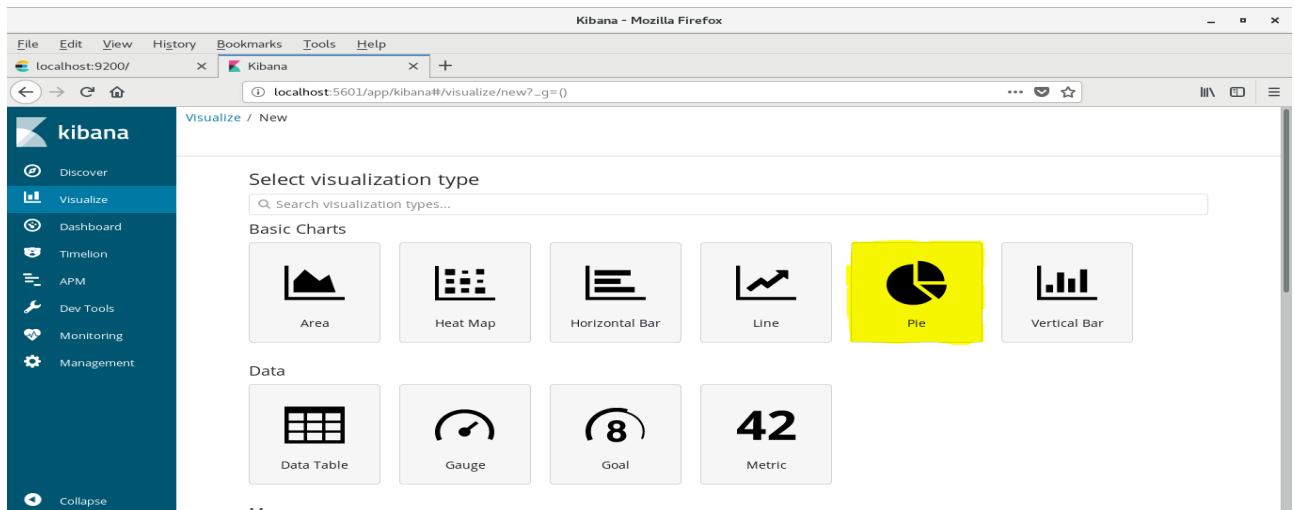
cars* Selected Fields Available Fields

maker	manufacture_year	model	price_eur
ford	-	fiesta	4,807.55
-	2,003	-	2,772.02
-	2,011	-	16,728.35
skoda	2,010	octavia	15,136.97
audi	2,007	a6	6,457.11
opel	2,002	combo	2,586.94
skoda	2,015	superb	35,325.68
-	2,002	-	1,957.81
-	2,005	-	3,997.04
-	-	-	5,162.84
suzuki	2,015	vitara	20,721.69
audi	1,998	a3	1,476.68
ford	2,010	focus	5,917.8

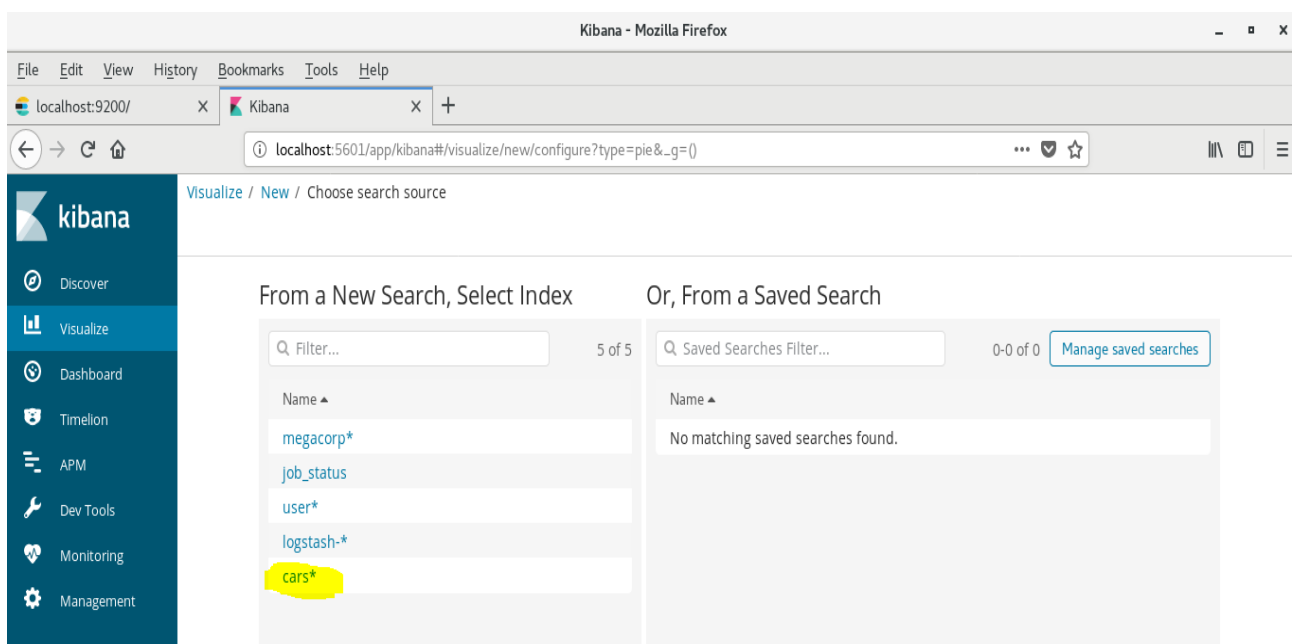
10. Creating Visualization in Kibana:

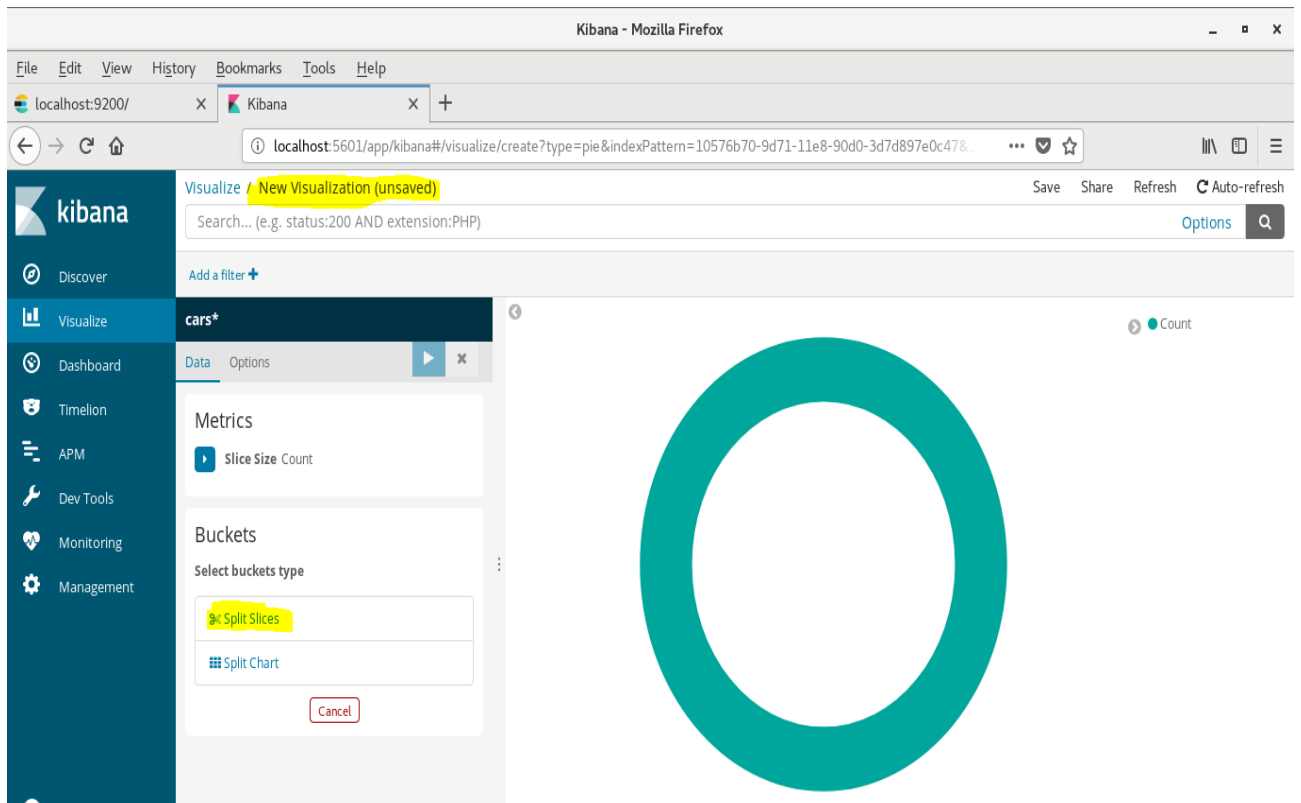


Selecting the PIE Chart:

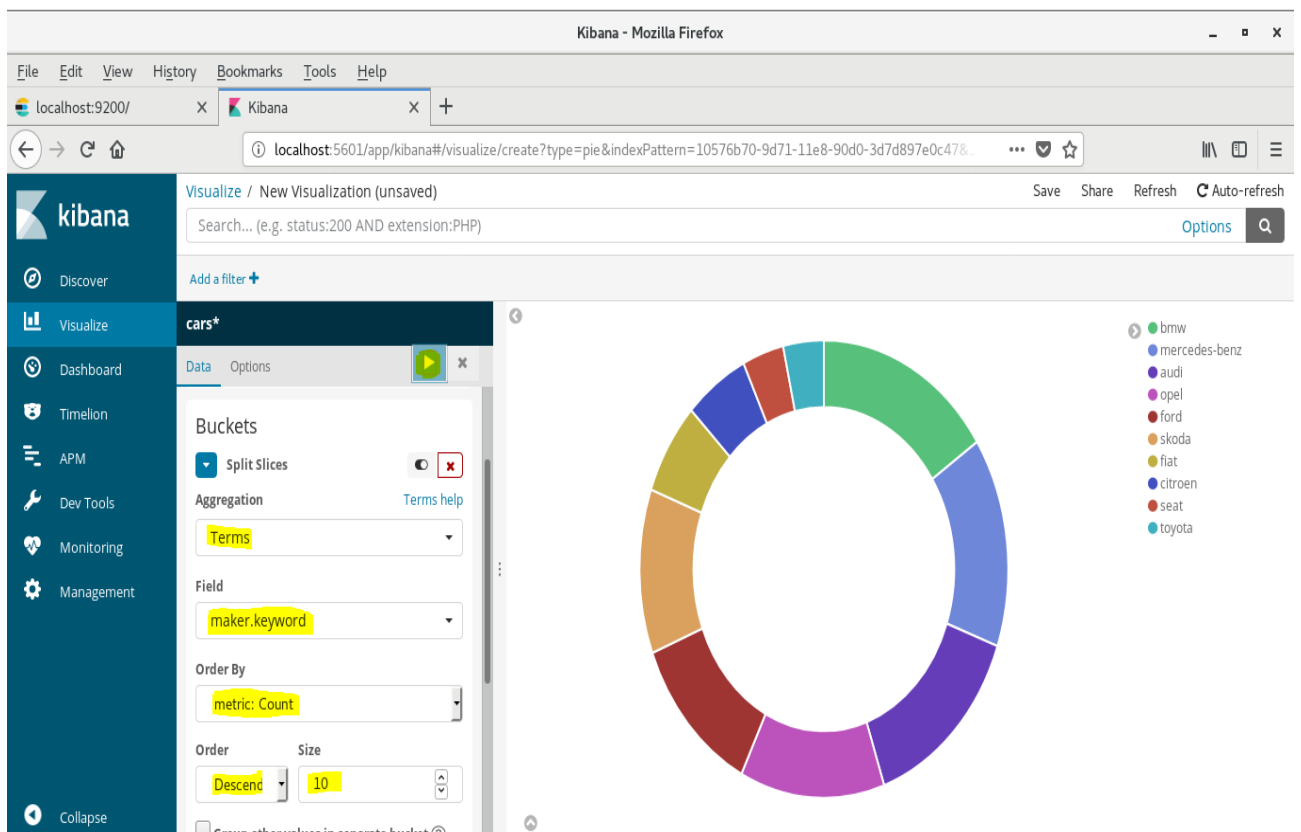


Select the CARS Index Pattern:

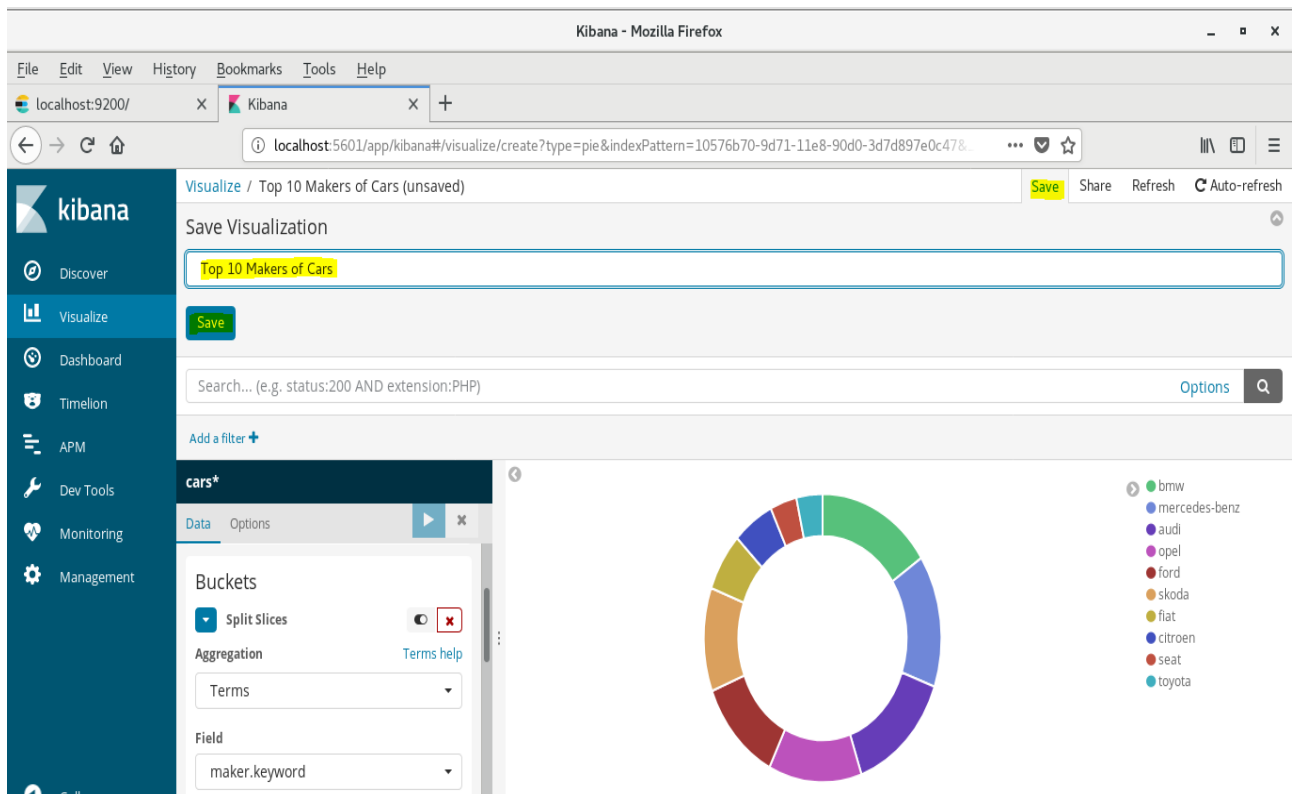




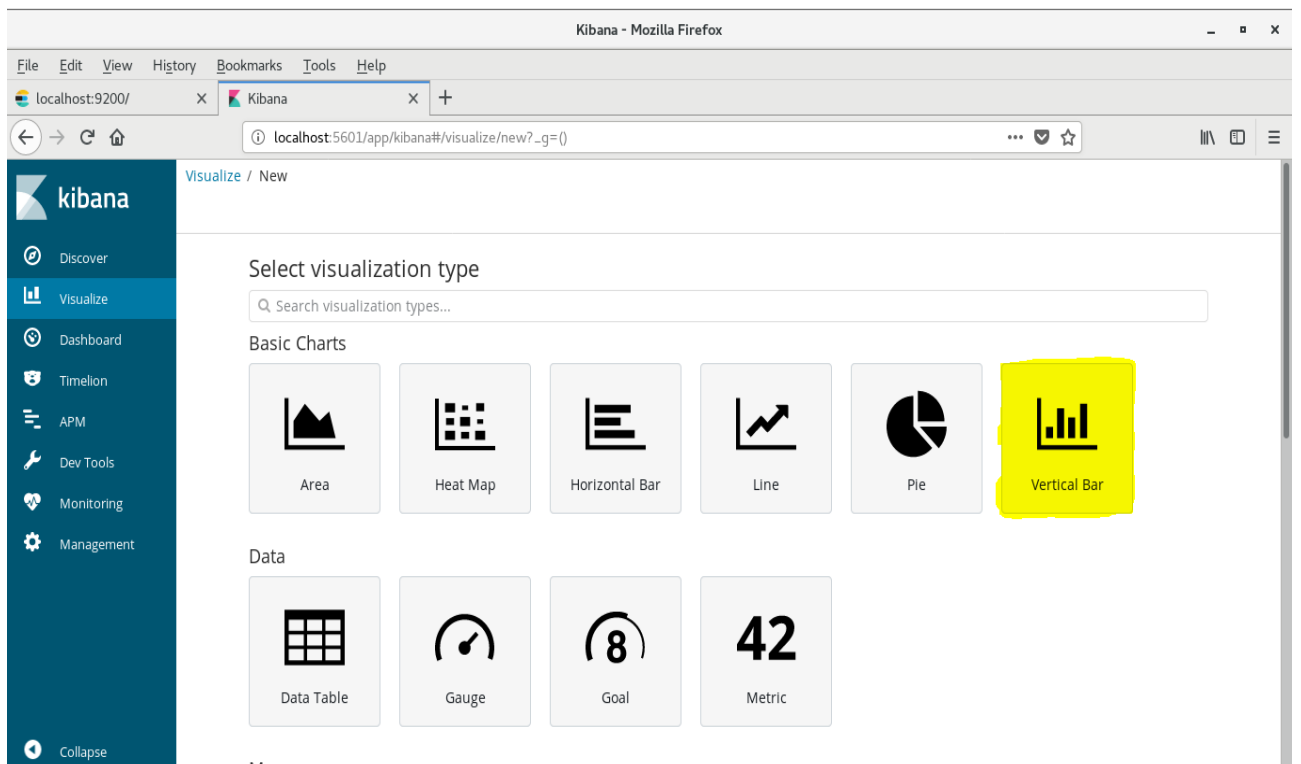
Top 10 Makers of the cars and this is real time, whenever the data inserts into Elastic Search, the PIE chart varies.

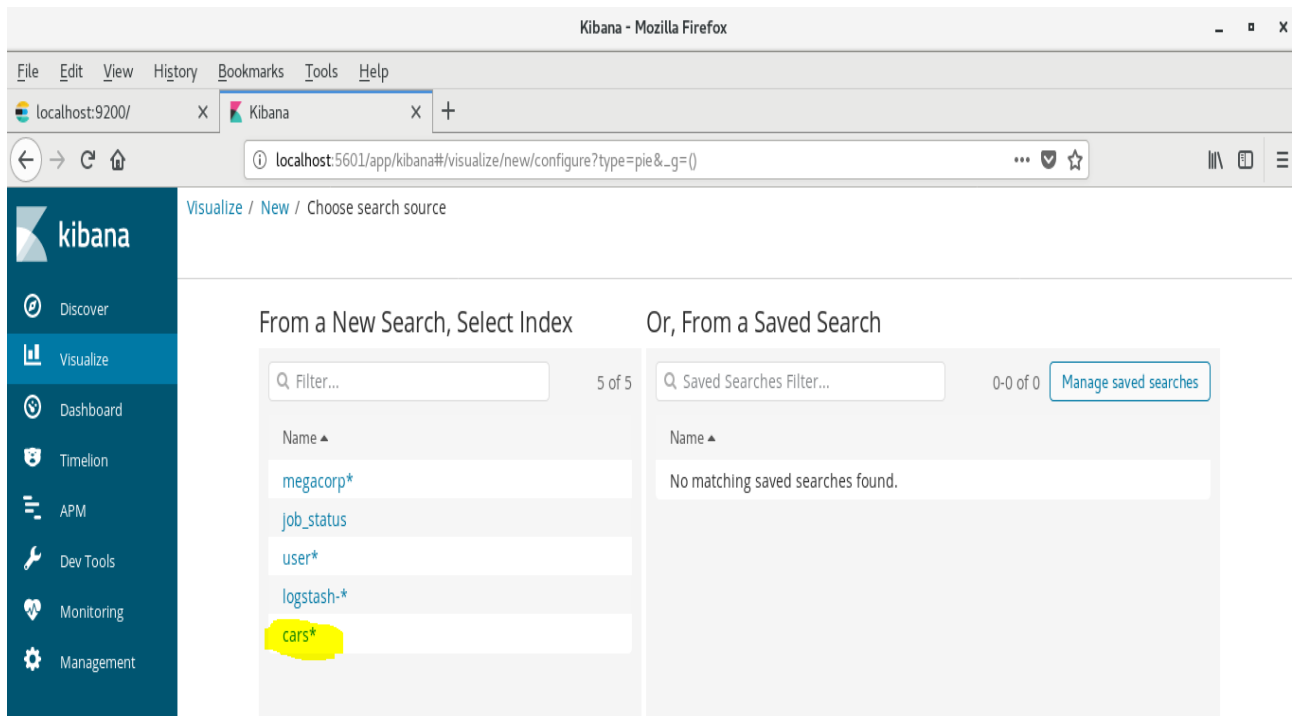


Save the Chart:

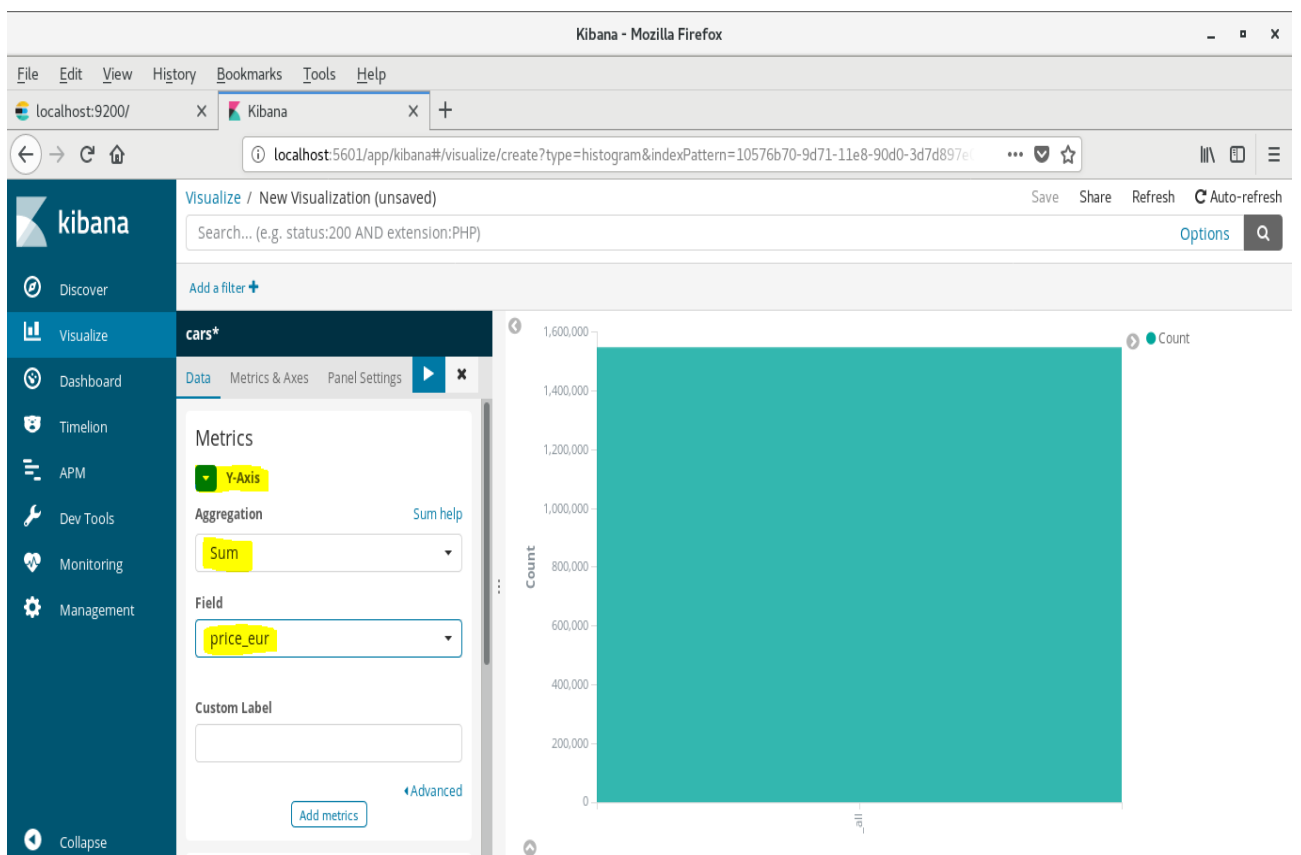


Creating other Visualization Chart:

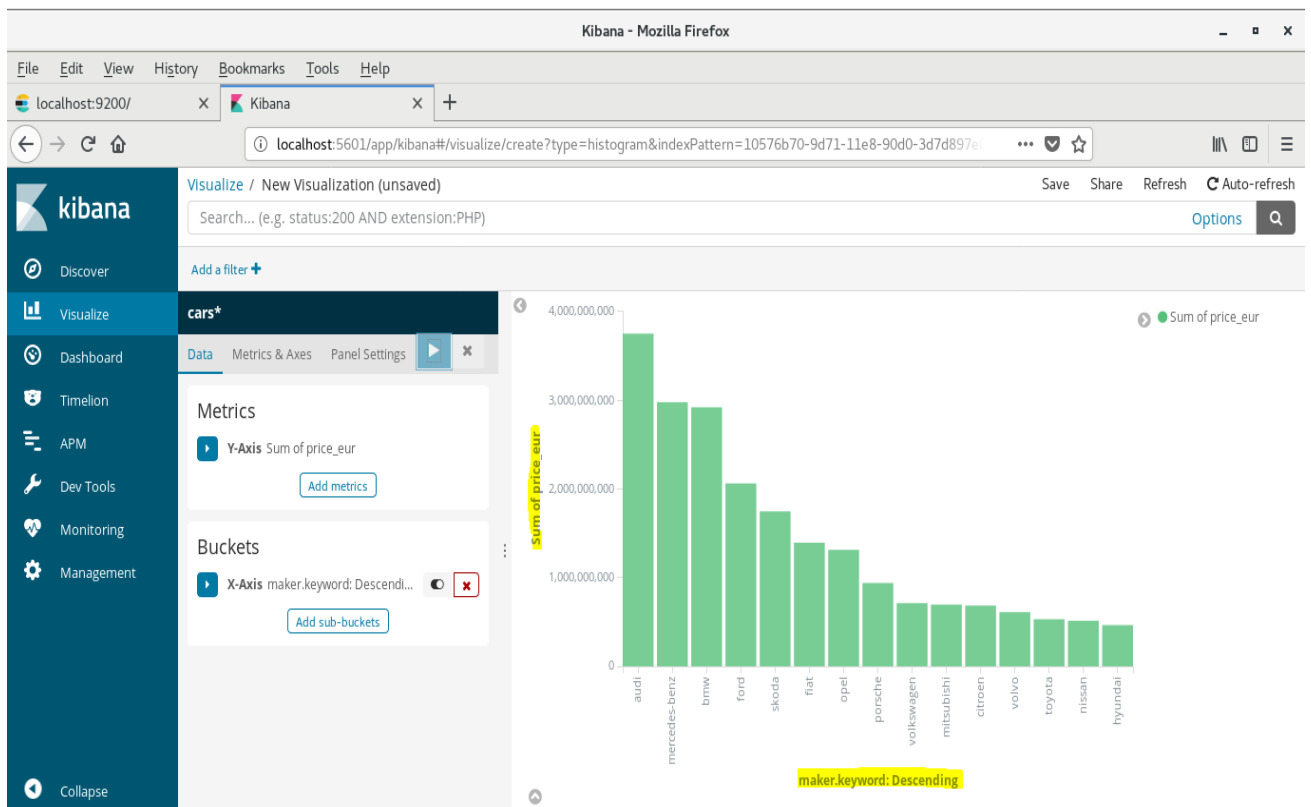
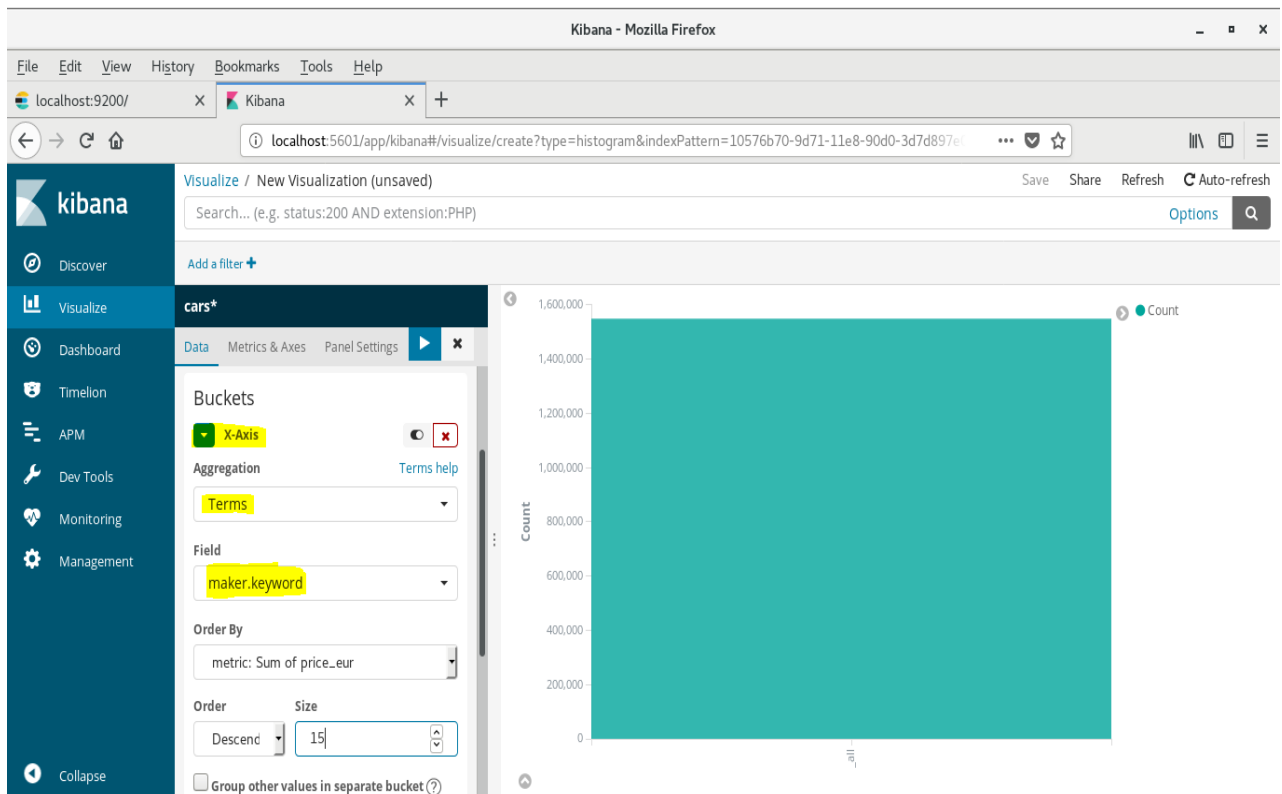




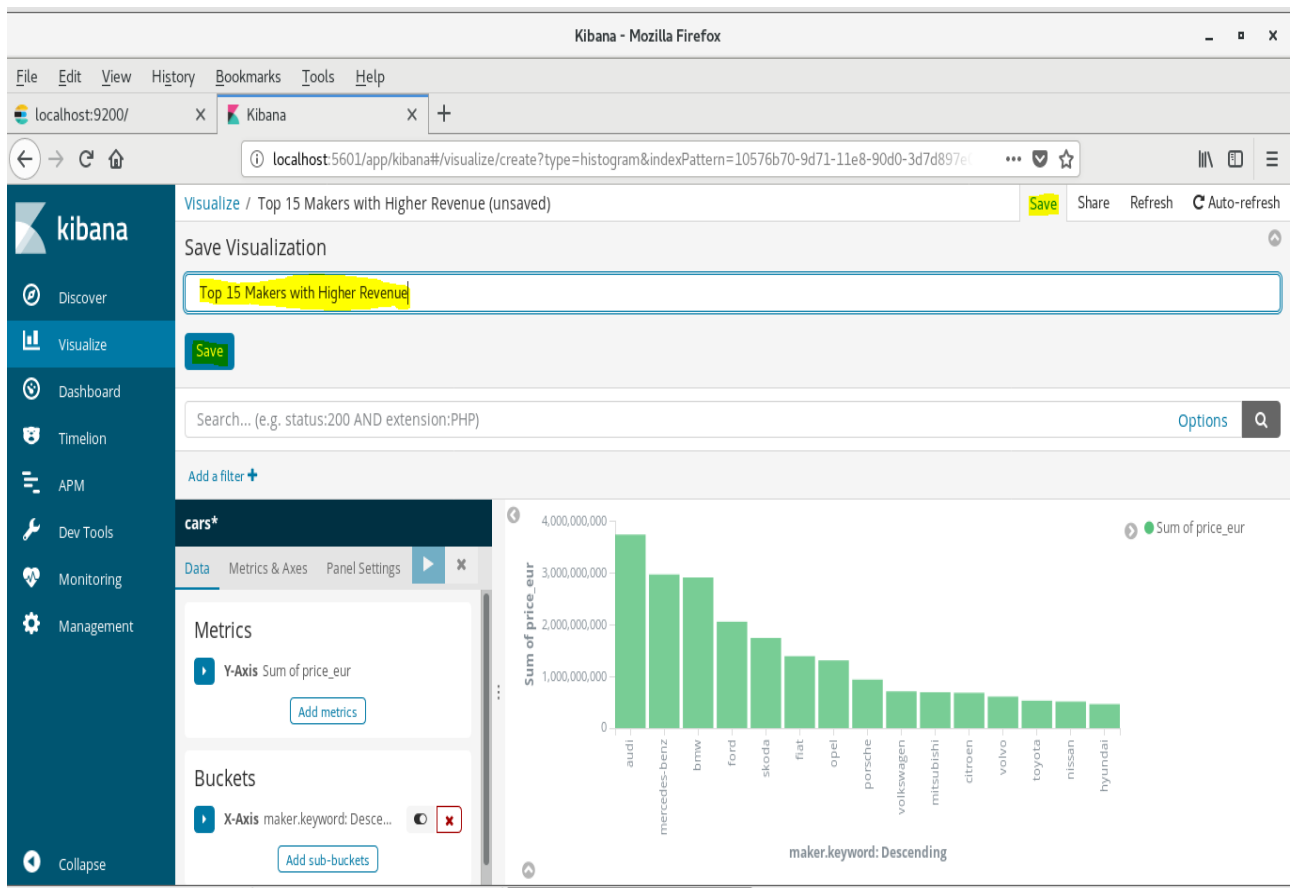
Y-Axis: Sum of Price of the Cars for each Maker



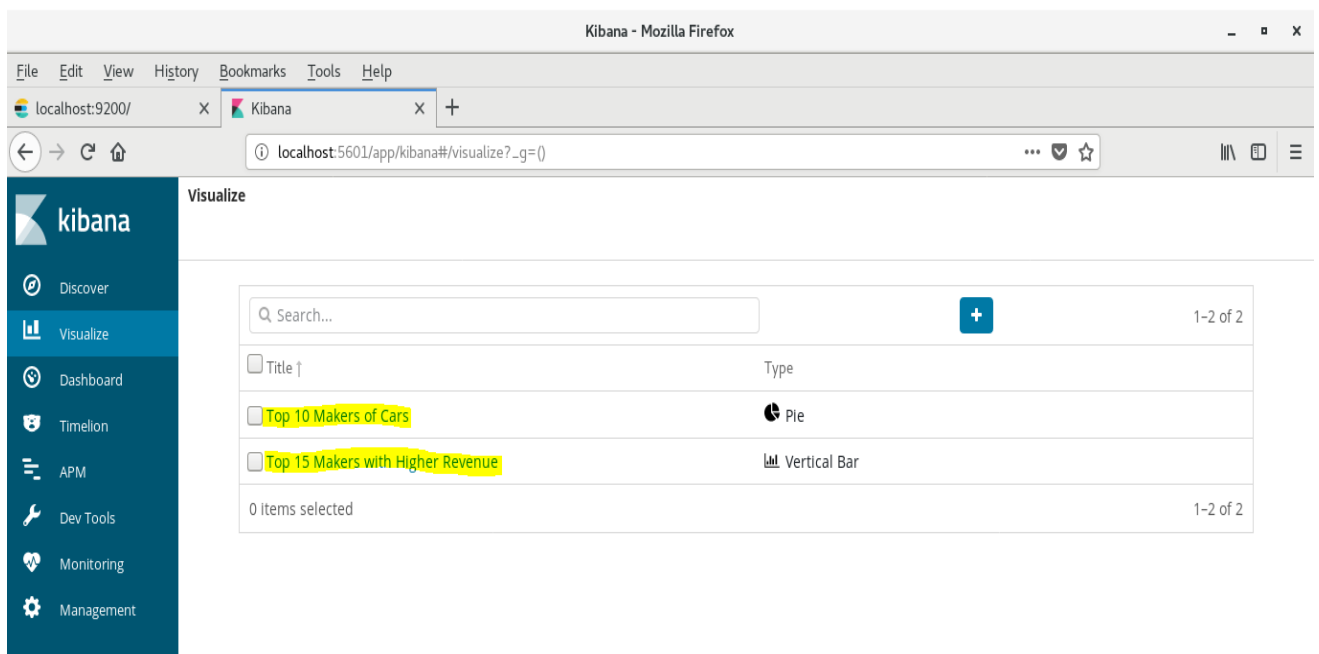
X-Axis: Top 15 Makers which make more revenue



Save bar plot.



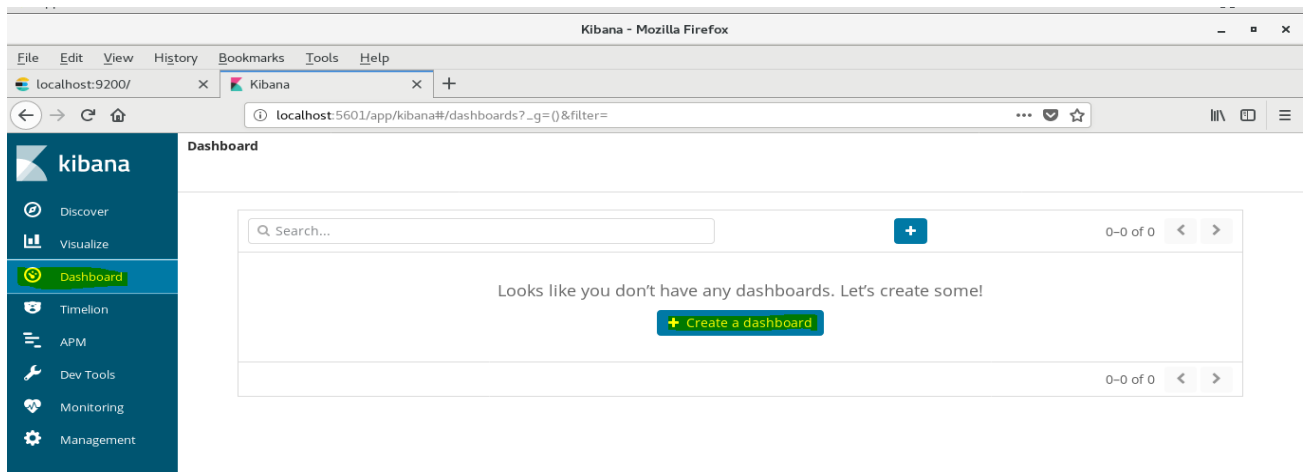
The list of Visualizations can be seen in the Visualize tab.



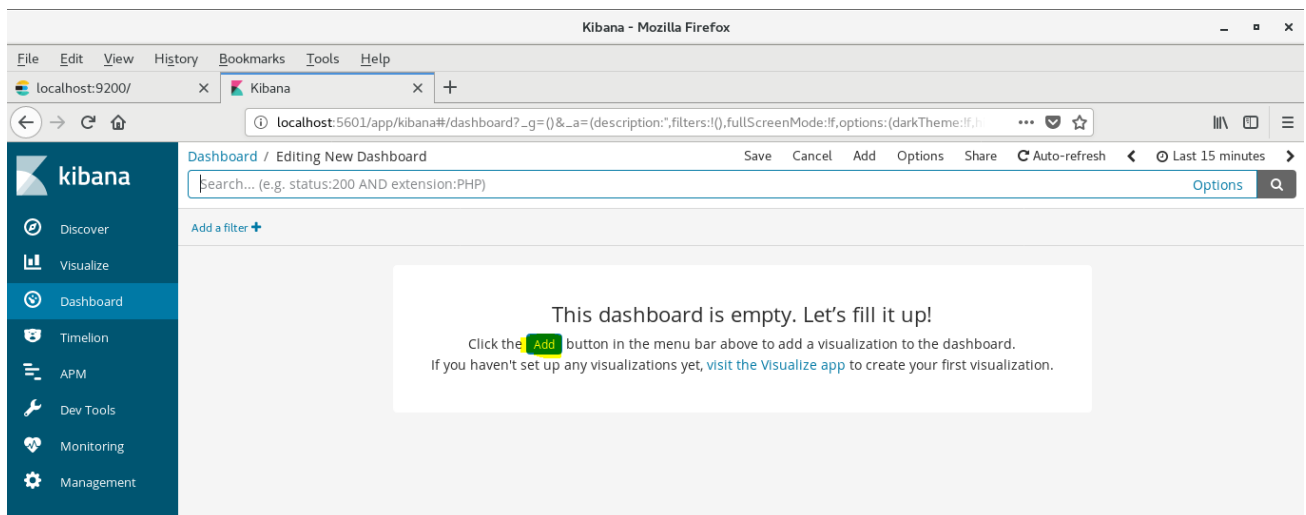
Select the Visualizations to create the Dash Board.

11. Creating Dashboard in Kibana:

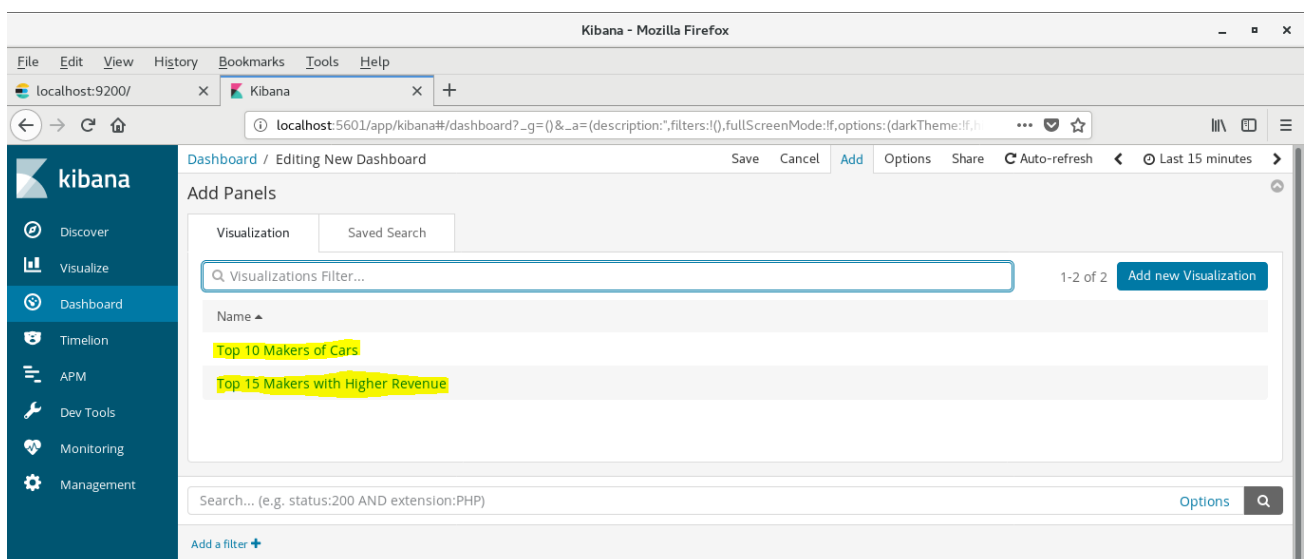
Dashboards can be created from the Dashboard tab.



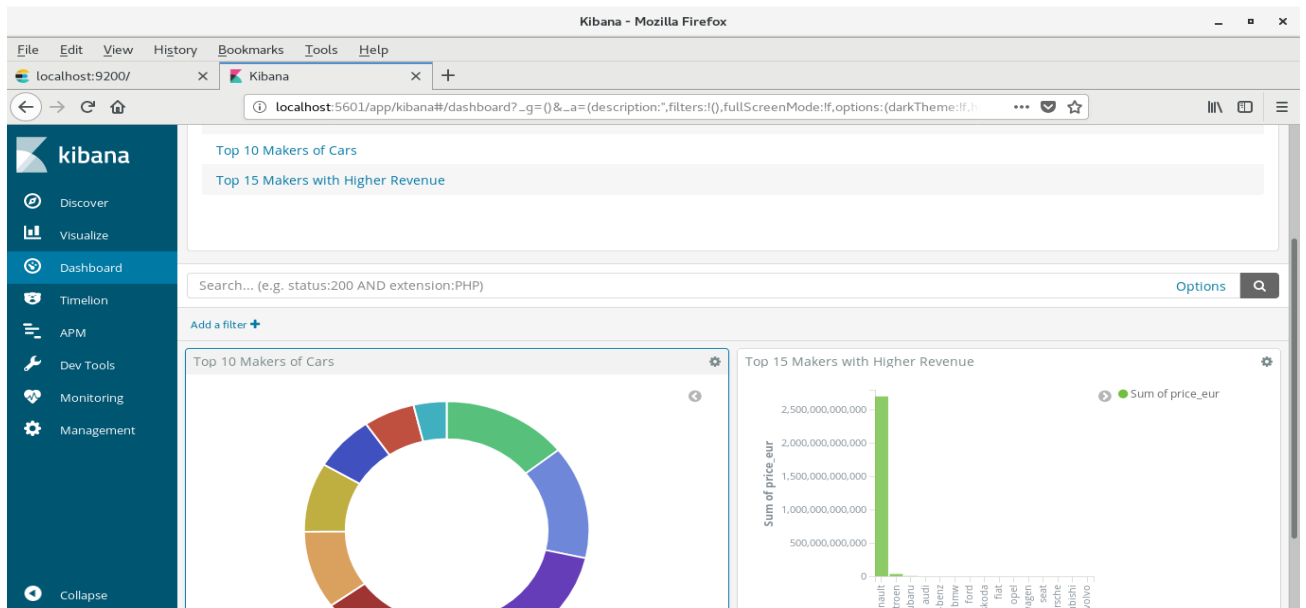
Add the Visualization.



Select the Visualizations:



After Selecting Visualizations, can view the Visualizations as follows:



The screenshot shows the Kibana 'Save dashboard' interface. The title is 'Top Makers' and the description is 'Top Makers by Cars and Revenue'. There is a checkbox for 'Store time with dashboard' which is currently unchecked. Below the description, there is a search bar and a 'Save' button. The main area displays the same two visualizations as the previous screenshot: 'Top 10 Makers of Cars' and 'Top 15 Makers with Higher Revenue'.

End of the tutorial.