비정형 빅데이터 분석의 응용과 실습

Week-07. Elastic Stack

서중원 2020.11.07

Elasticsearch Mapping

Mapping이란?

- 관계형 데이터의 스키마와 동일
- Mapping 없이 데이터를 엘라스틱서치에 삽입?
  - 가능
  - 하지만, Mapping을 없이 데이터를 넣는 것은 데이터의 사용성이 떨어질 수 있음
    - 예를들어 2020-11-07이라는 값을 Mapping없이 넣는다면?
      - String으로 인지할 수 있음
      - 날짜별 정렬, 월별 정렬/필터링과 같은 연산을 사용할 수 없음
    - 숫자를 입력했지만, 문자열로 인식했다?
      - Min, max, mean, median과 같은 연산을 사용할 수 X
- 가능하다면! 항상 Mapping을 먼저 지정해 놓고 데이터를 삽입!
  - 데이터 먼저 넣고, Mapping을 후에 지정해줄 수도 있음

Elasticsearch Search

Search 방식

- request_body에 json 형식으로 조건을 작성!

```
body = {
    "query": {
        "term": {
            "points":30
        }
    }
}
res = es.search(body=body,index=INDEX_NAME)
pprint.pprint(res)
```

Elasticsearch Aggregation

Aggregation 이란?

- Search Query 사용시 수치적 값들의 다양한 값들을 얻어 낼 수 있다.

```
body = {
    "size" : 0,
    "aggs" : {
        "avg_score" : {
            "avg" : {
                "field" : "points"
            }
        }
    }
}
res = es.search(body=body,index=INDEX_NAME)
pprint.pprint(res)
```

Elasticsearch Bucket Aggregation

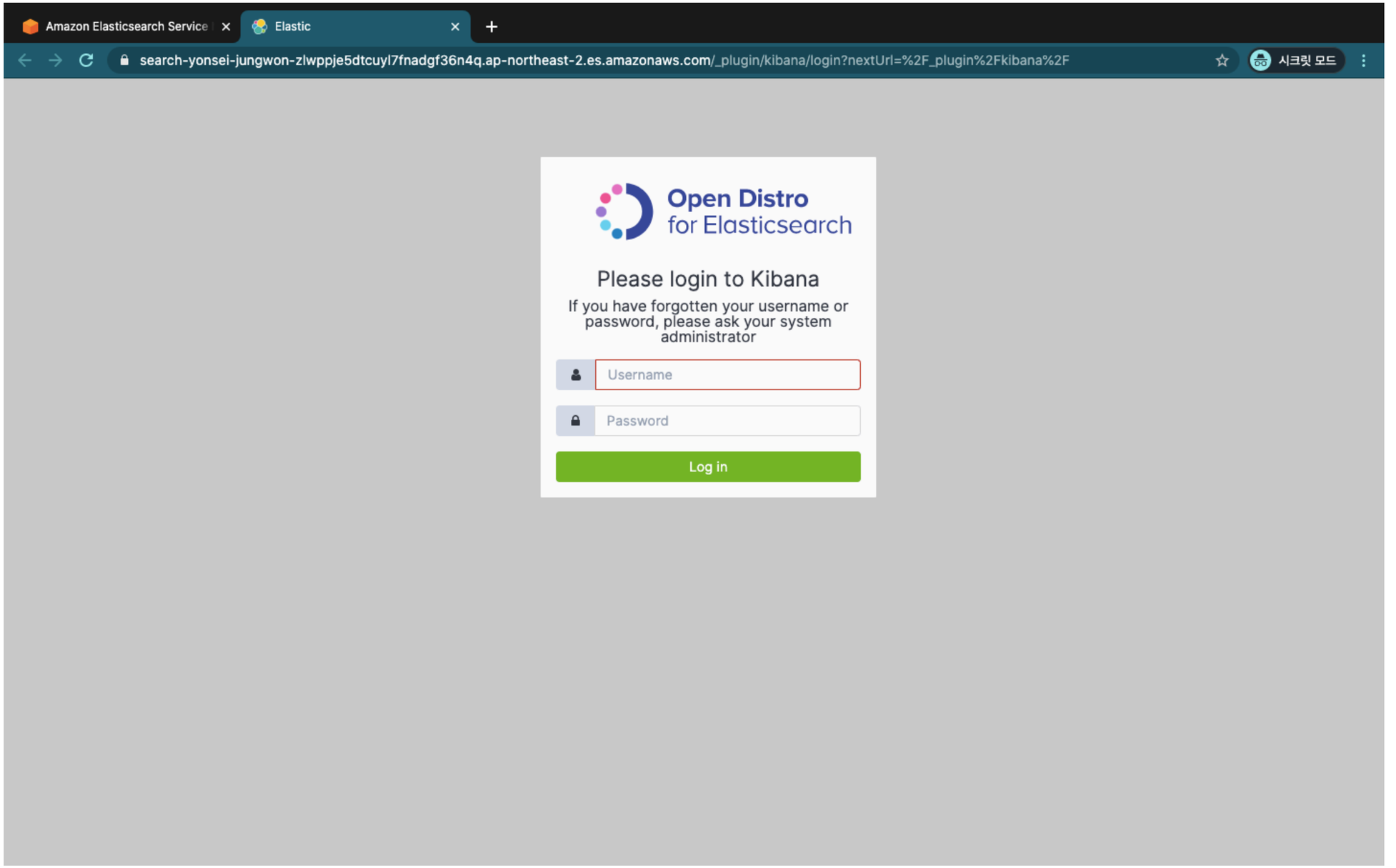
Bucket Aggregation 이란?

- RDB의 group by와 유사한 기능
- “document의 bucket을 만든다”

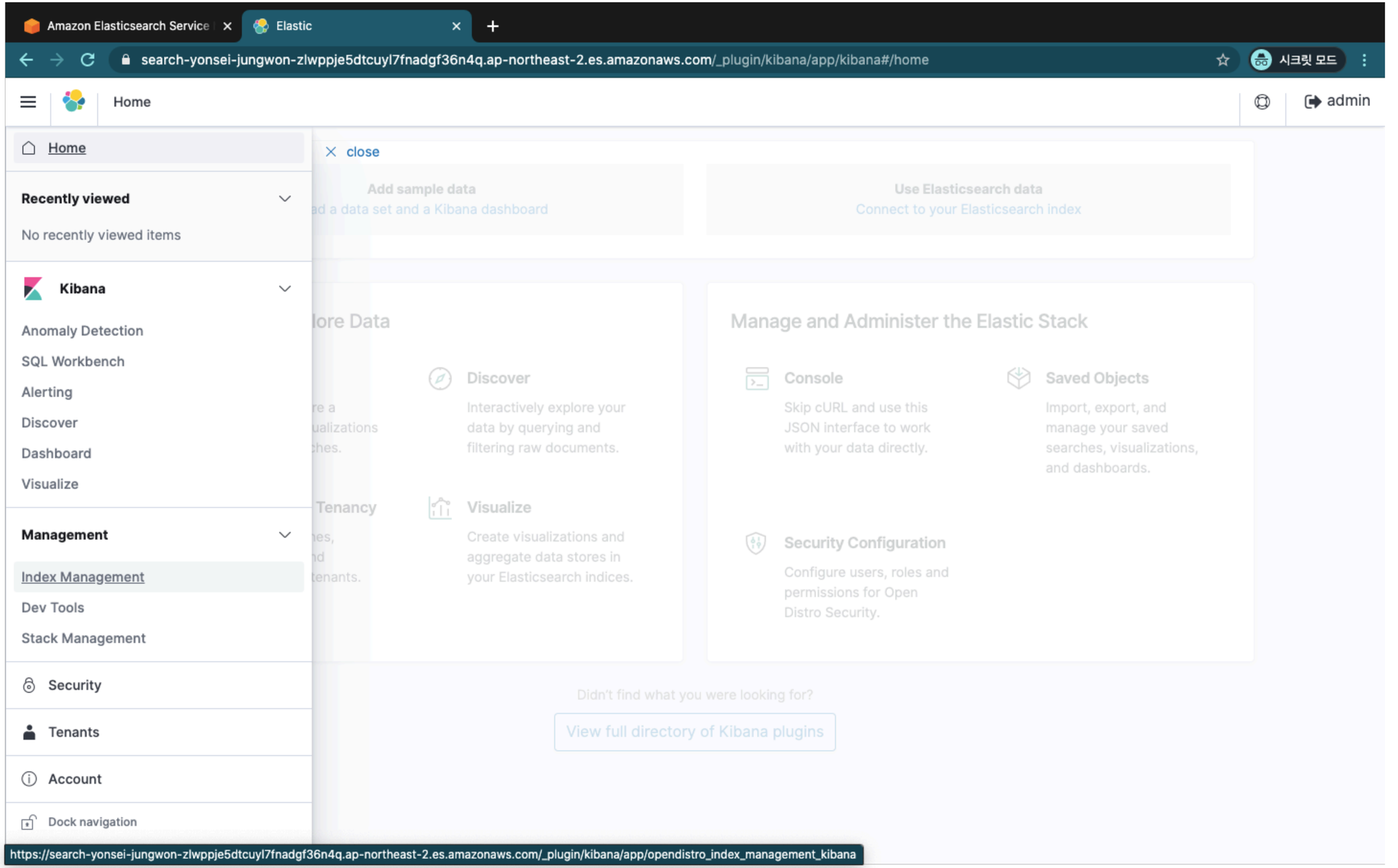
```
body = {
    "size" : 0,
    "aggs" : {
        "team_stats" : {
            "terms" : {
                "field" : "team"
            },
            "aggs" : {
                "stats_score" : {
                    "stats" : {
                        "field" : "points"
                    }
                }
            }
        }
    }
}
res = es.search(body=body,index=INDEX_NAME)
pprint.pprint(res)
```

Kibana

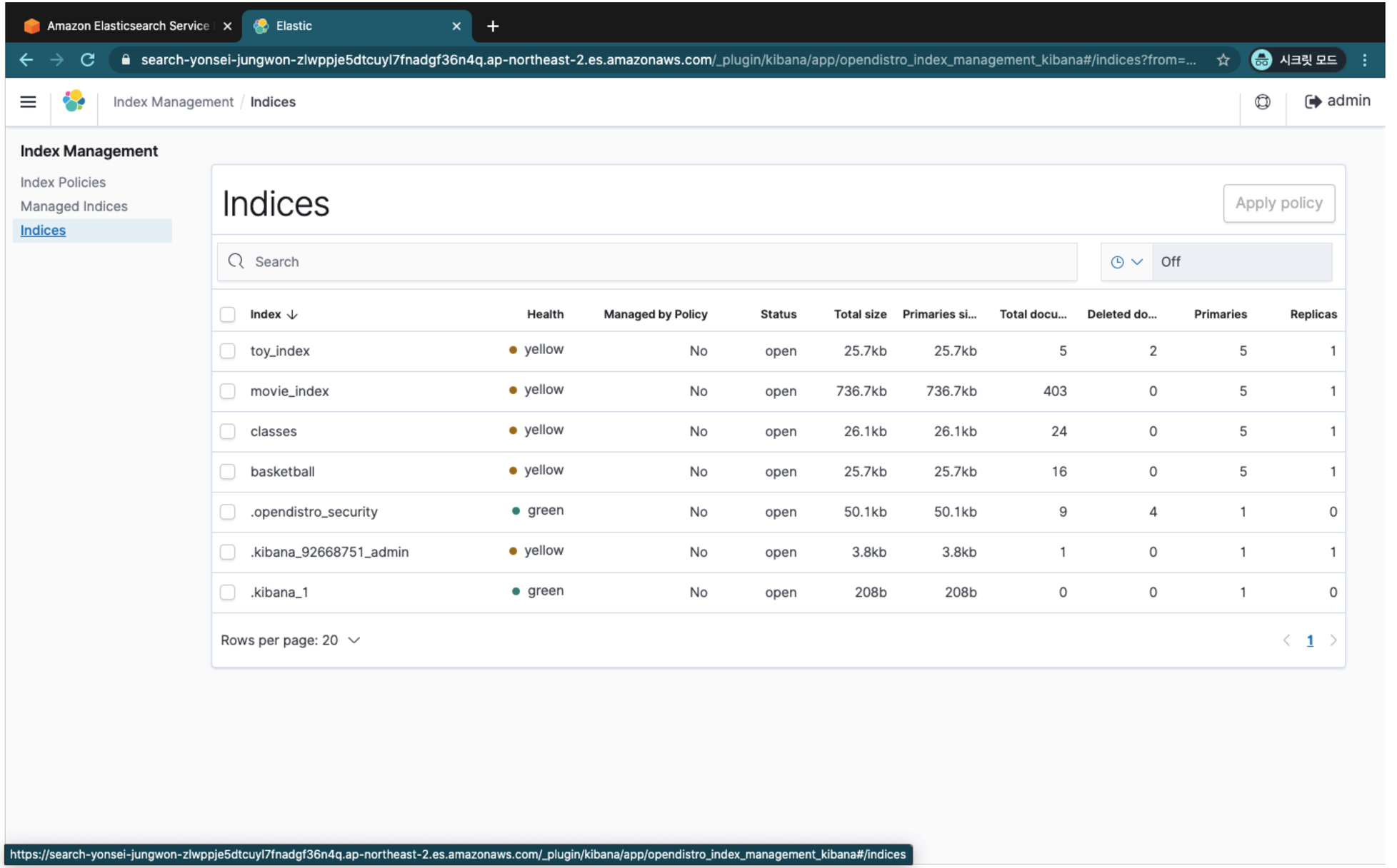
AWS 엘라스틱 서치 콘솔로 이동후 Kibana 부분의 링크 클릭!



좌측의 메뉴화면에서 Index Management 클릭후, 그동안 생성한 인덱스들이 잘 들어갔는지 확인



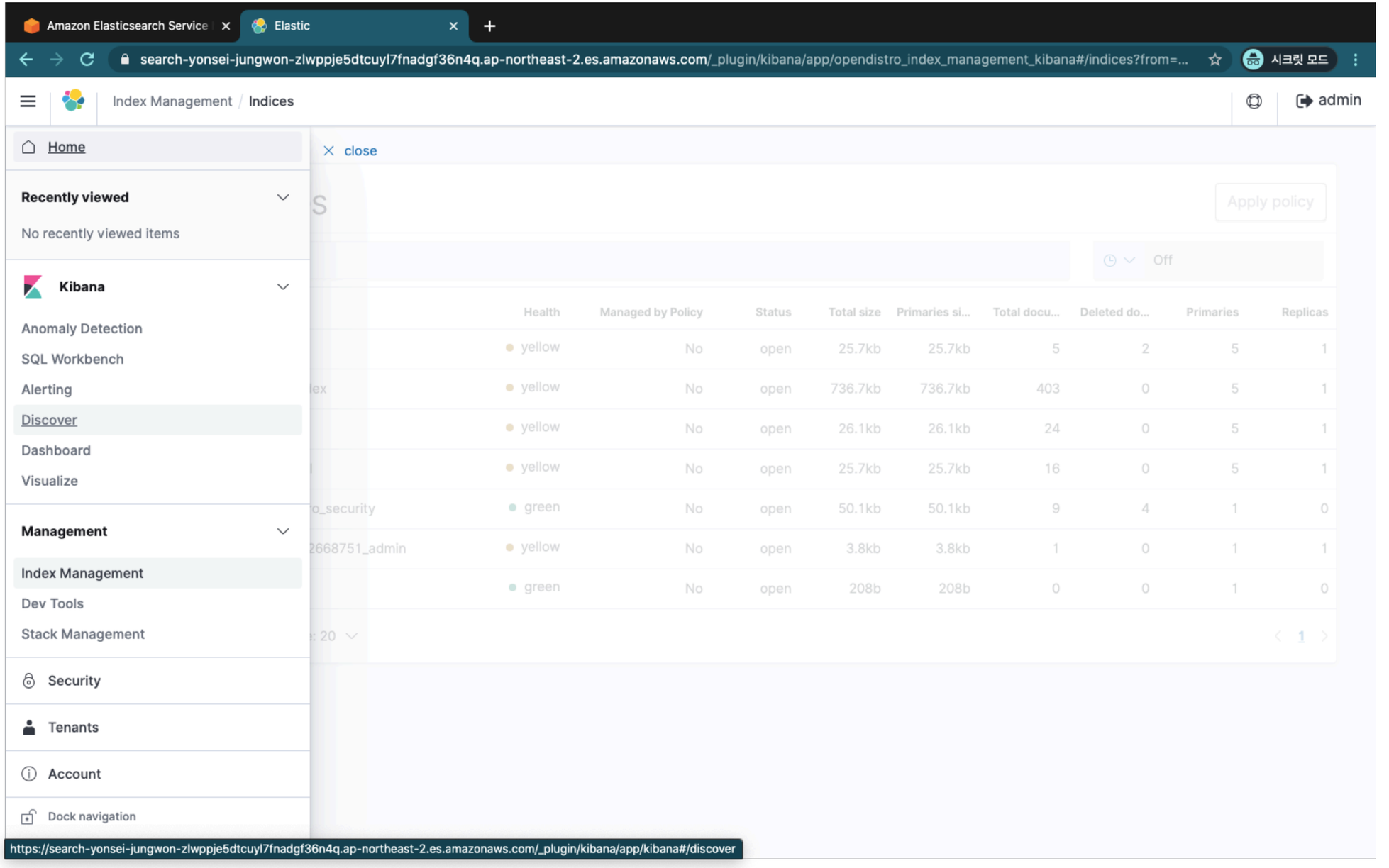
잘들어갔네요!


Amazon Elasticsearch Service
Elastic
search-yonsei-jungwon-zlwppje5dtcuyl7fnadgf36n4q.ap-northeast-2.es.amazonaws.com/_plugin/kibana/app/opendistro_index_management_kibana#/indices?from=...
시크릿 모드
Index Management / Indices
admin
Index Management
Index Policies
Managed Indices
Indices
Indices
Apply policy
Search
Off
Index
Health
Managed by Policy
Status
Total size
Primaries si...
Total docu...
Deleted do...
Primaries
Replicas
toy_index yellow No open 25.7kb 25.7kb 5 2 5 1
movie_index yellow No open 736.7kb 736.7kb 403 0 5 1
classes yellow No open 26.1kb 26.1kb 24 0 5 1
basketball yellow No open 25.7kb 25.7kb 16 0 5 1
.opendistro_security green No open 50.1kb 50.1kb 9 4 1 0
.kibana_92668751_admin yellow No open 3.8kb 3.8kb 1 0 1 1
.kibana_1 green No open 208b 208b 0 0 1 0
Rows per page: 20
1
https://search-yonsei-jungwon-zlwppje5dtcuyl7fnadgf36n4q.ap-northeast-2.es.amazonaws.com/_plugin/kibana/app/opendistro_index_management_kibana#/indices

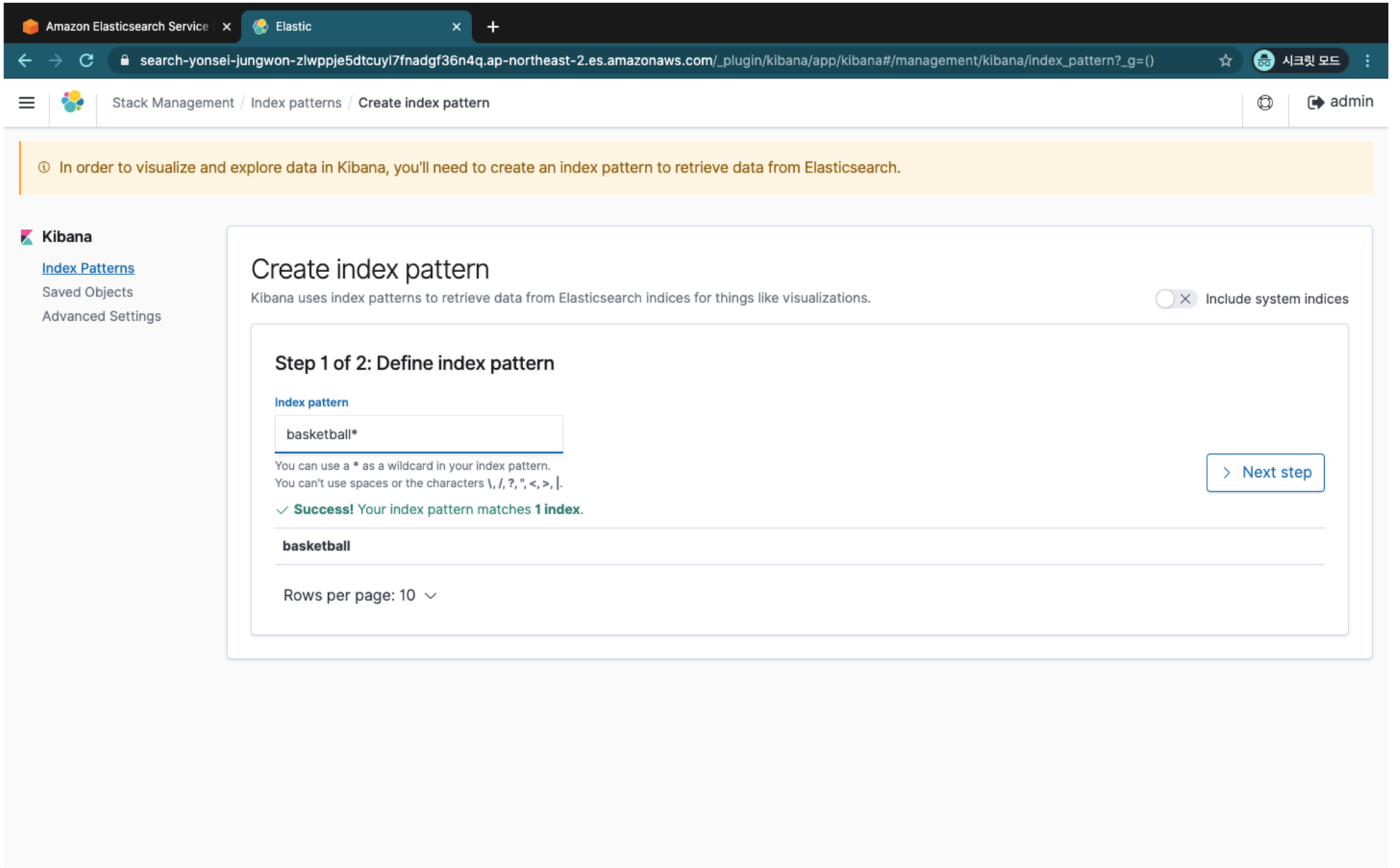
Kibana Discover

가장 기본이 되는 View

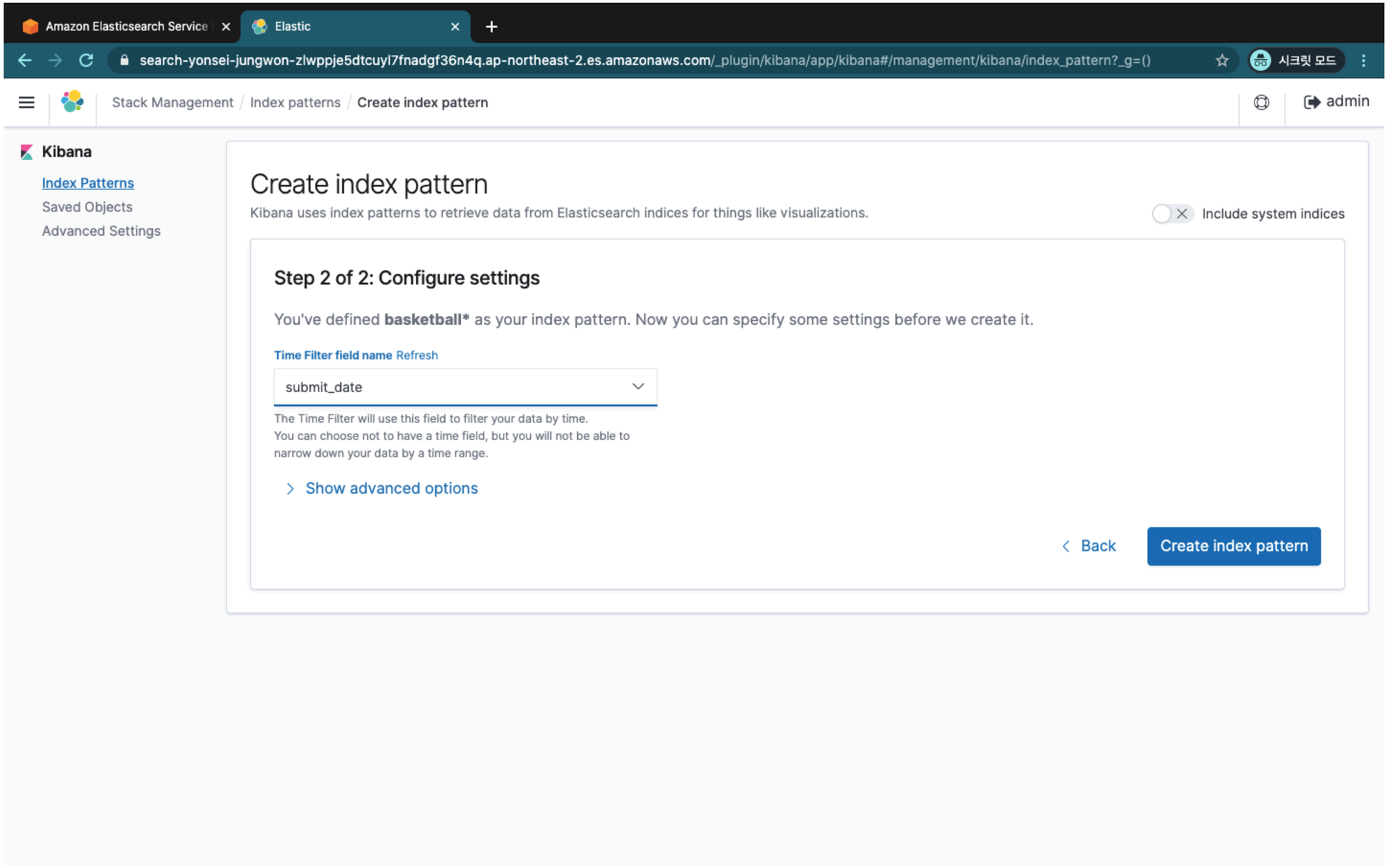
기존에 Elastic Search에 등록된 Index를 Kibana에 등록, Stack Management 클릭



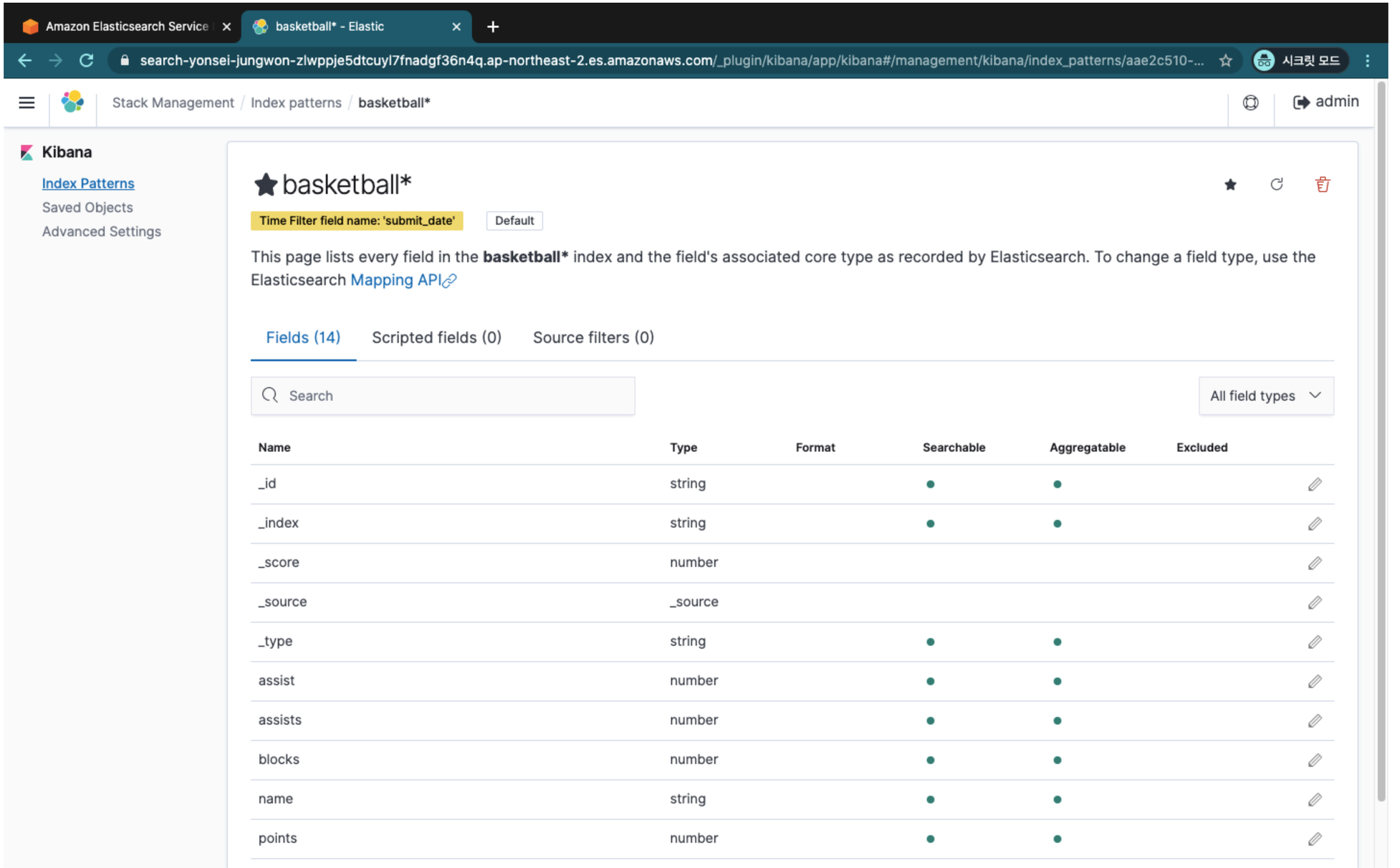
Basketball 검색



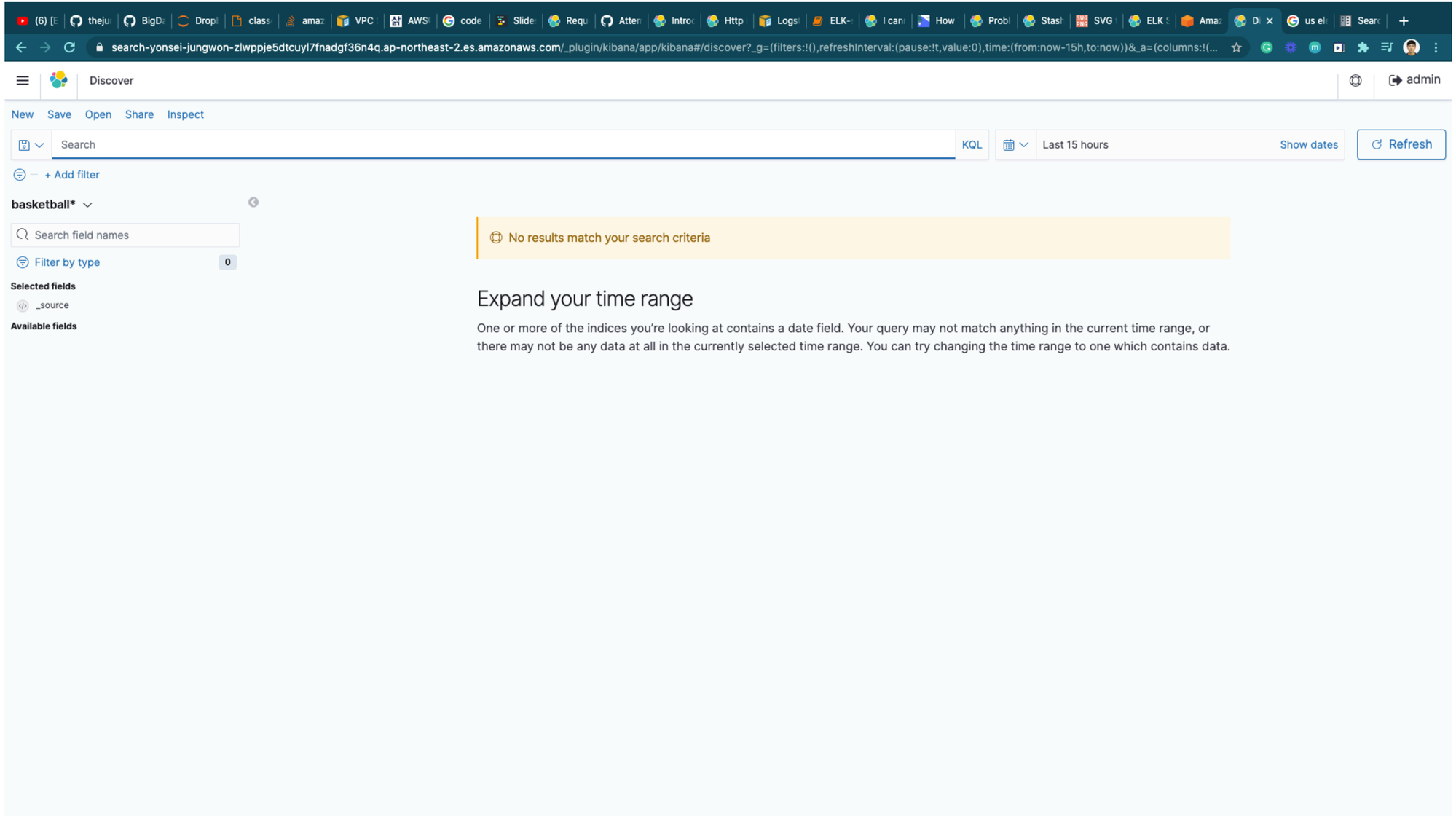
submit_date 선택



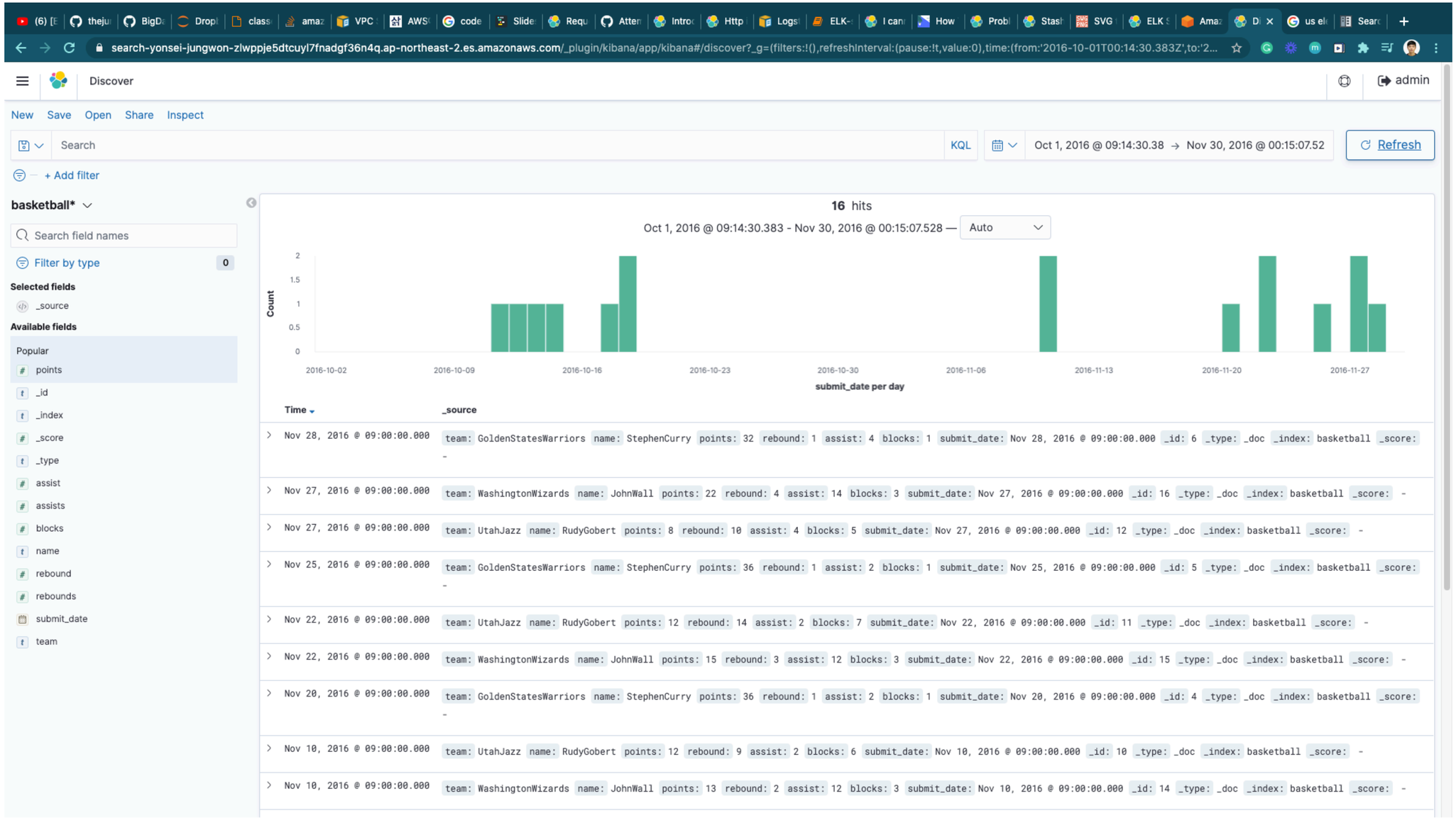
각각의 field가 잘 들어갔는지 확인



우상단에서 날짜 조정후 Basketball 데이터 확인

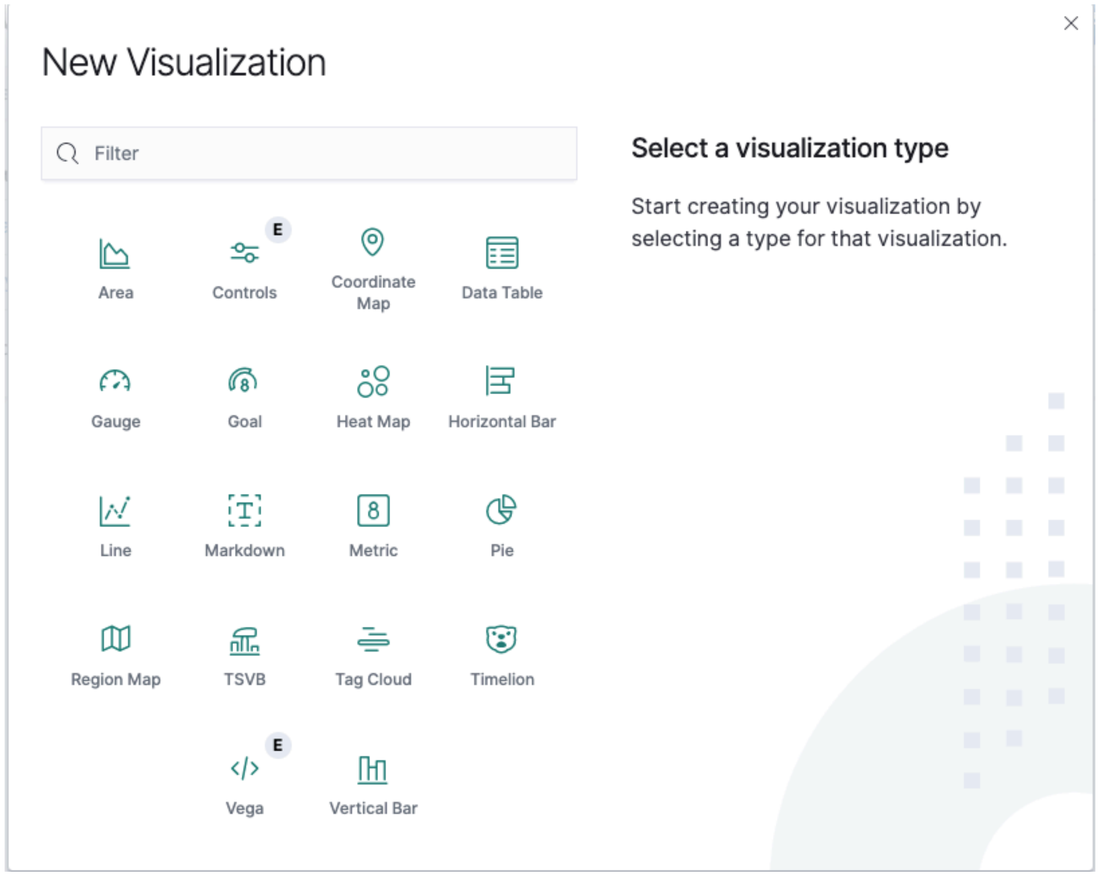


2016년 10월 1일 부터 11월 30일까지로 설정

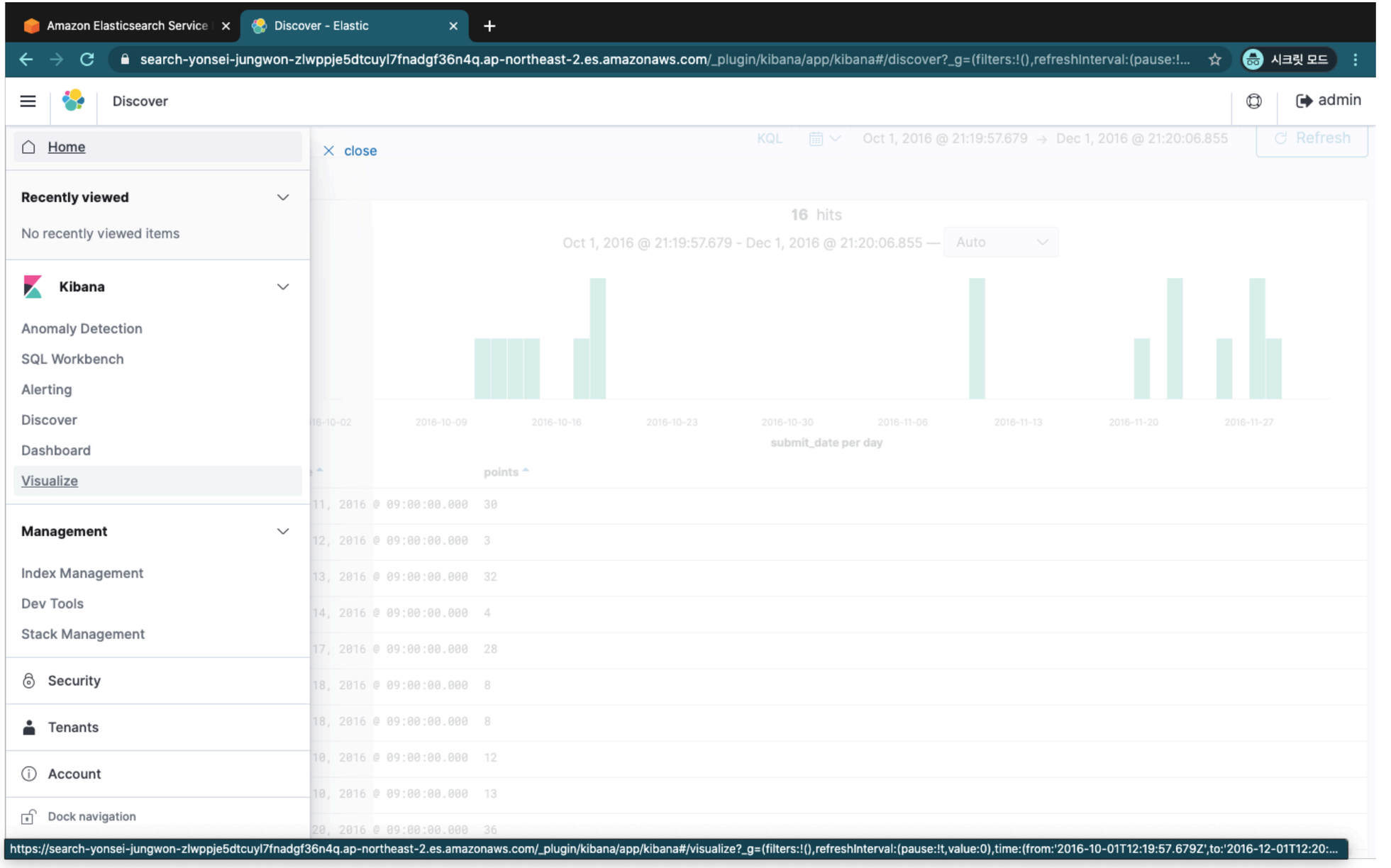


Kibana Visualize

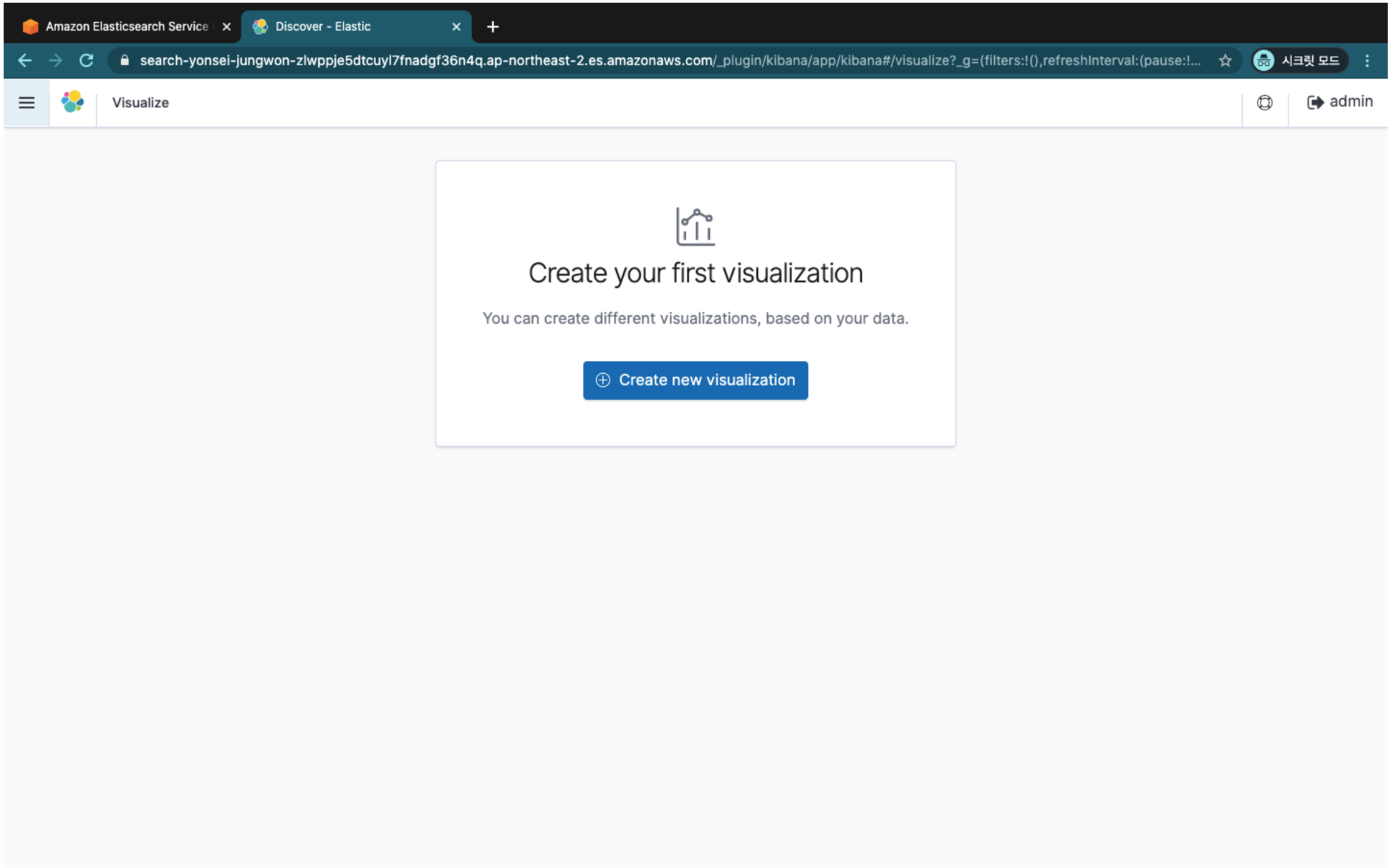
Bar Chart, Pie Chart



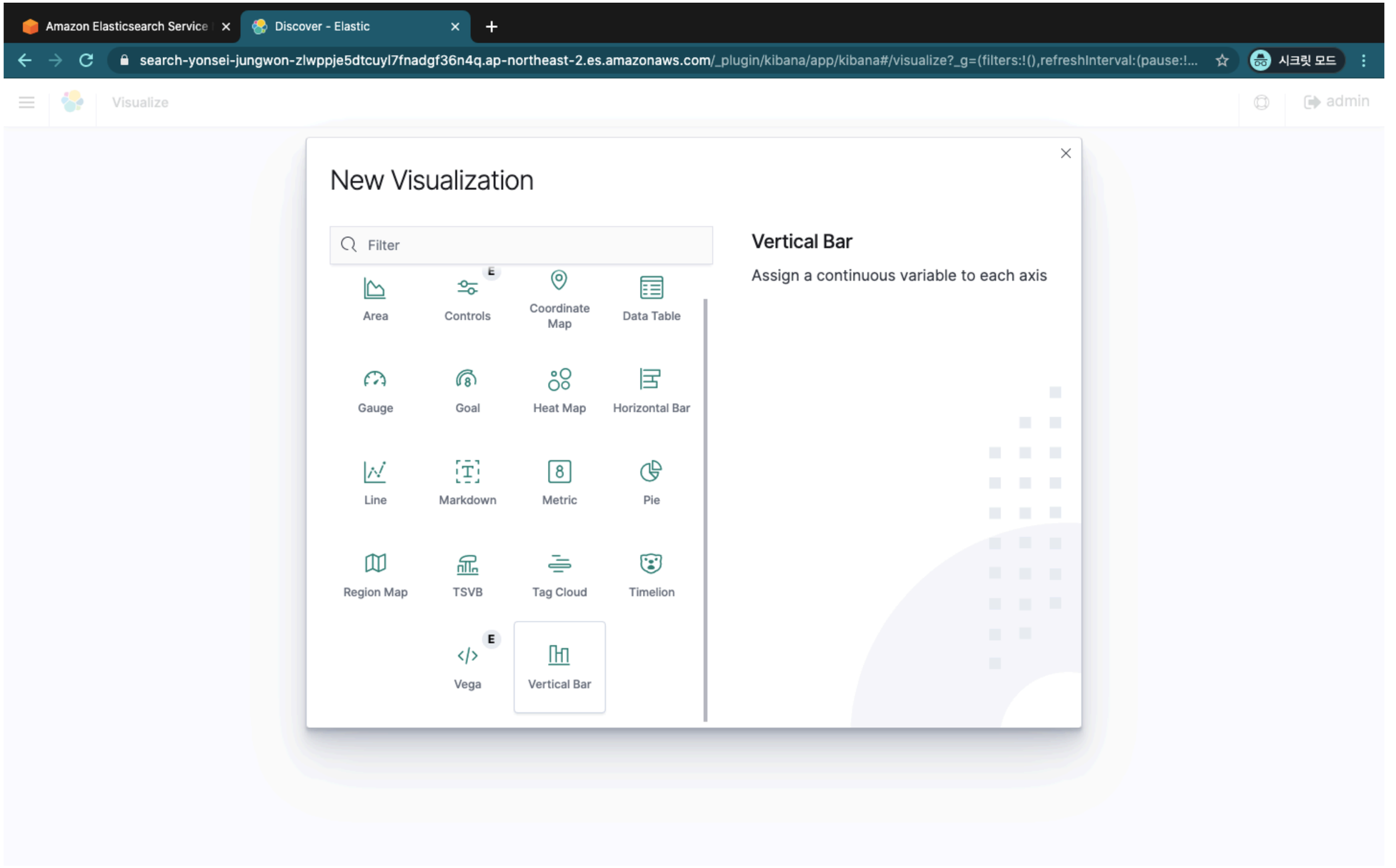
메뉴에서 Visualize 클릭



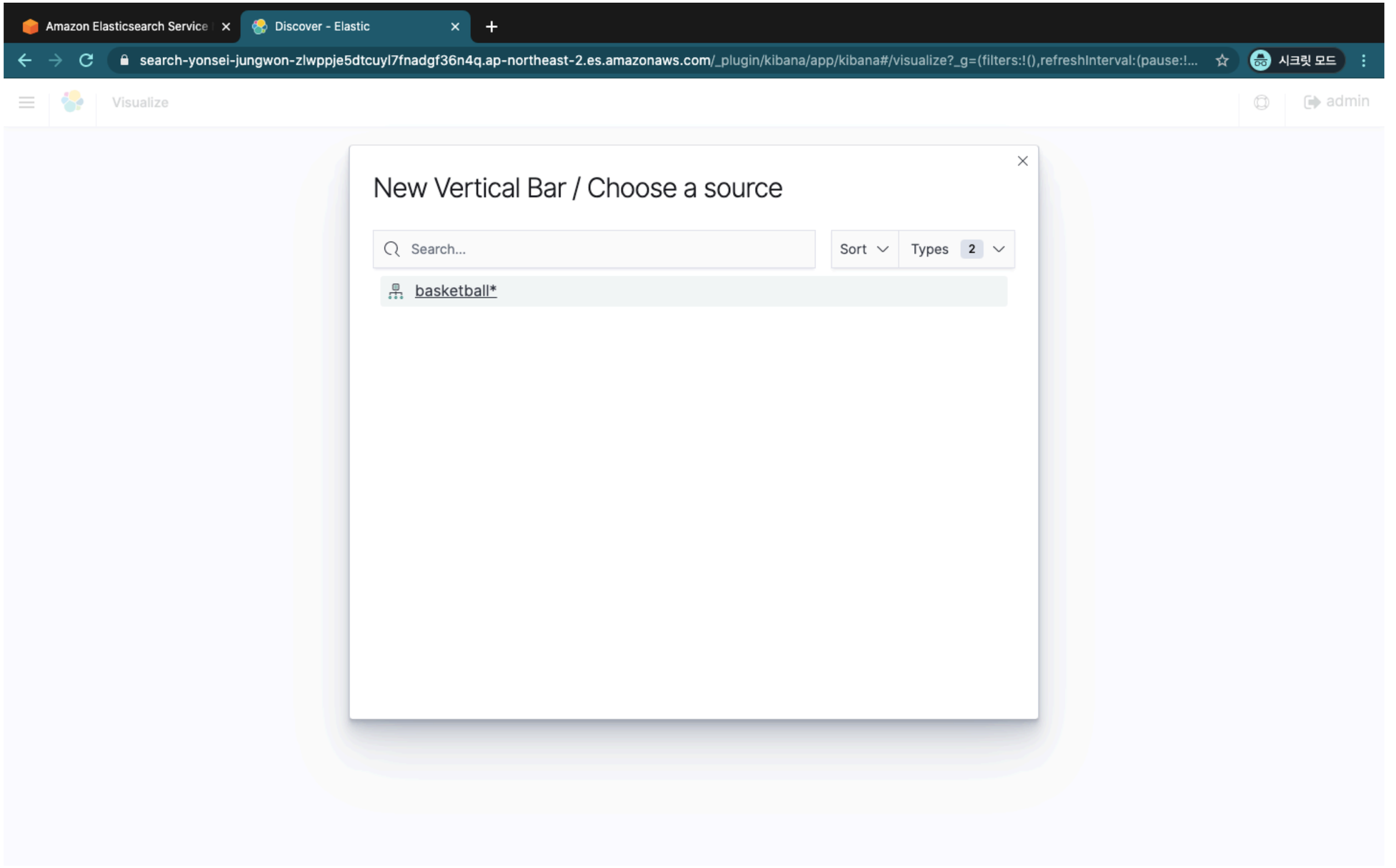
첫 시각화 만들기!


Amazon Elasticsearch Service
Discover - Elastic
search-yonsei-jungwon-zlwppje5dtcuyl7fnadgf36n4q.ap-northeast-2.es.amazonaws.com/_plugin/kibana/app/kibana#/visualize?_g=(filters:!(),refreshInterval:(pause:!...
시크릿 모드
Visualize
admin
Create your first visualization
You can create different visualizations, based on your data.
Create new visualization

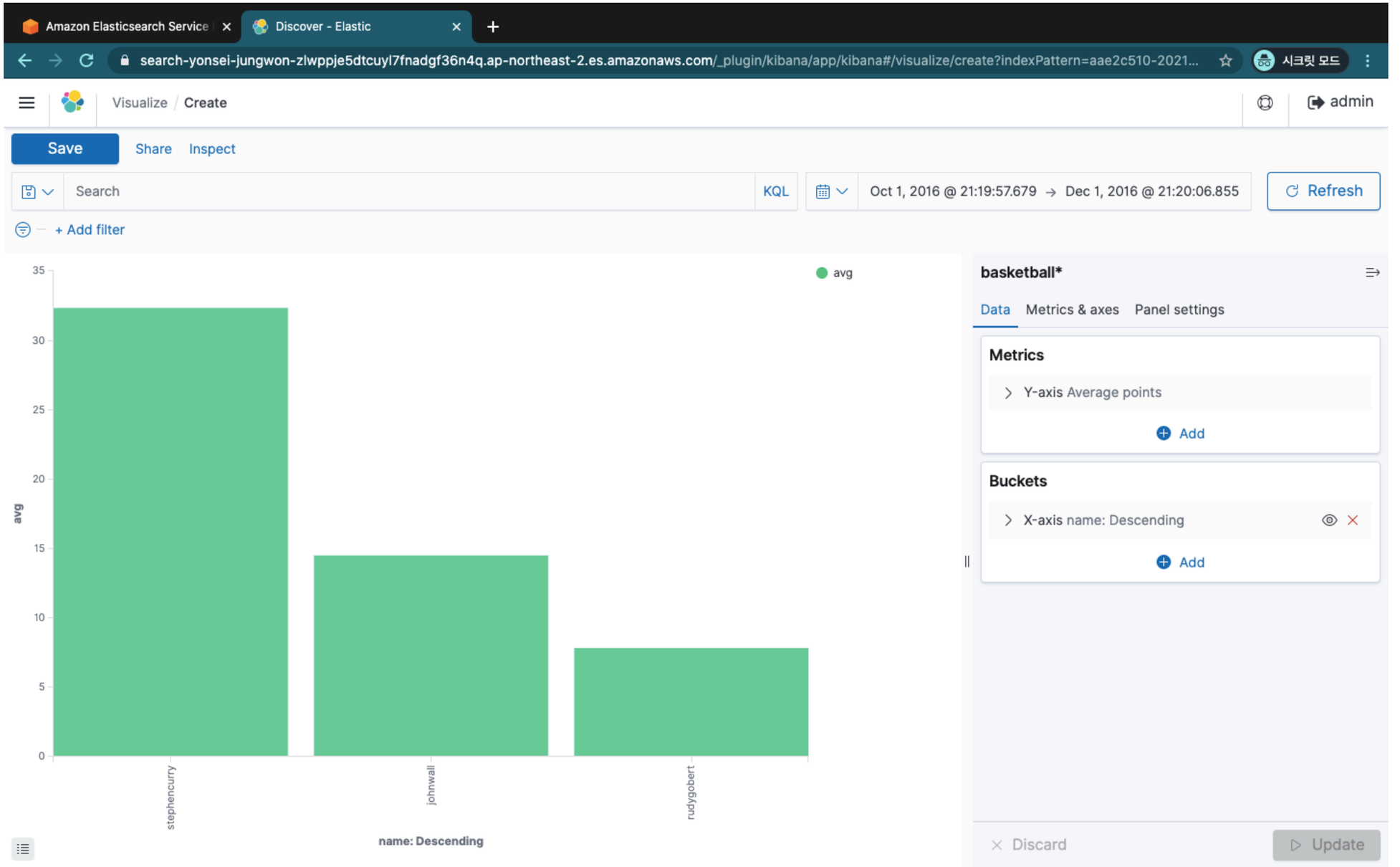
Vertical Bar 고르기



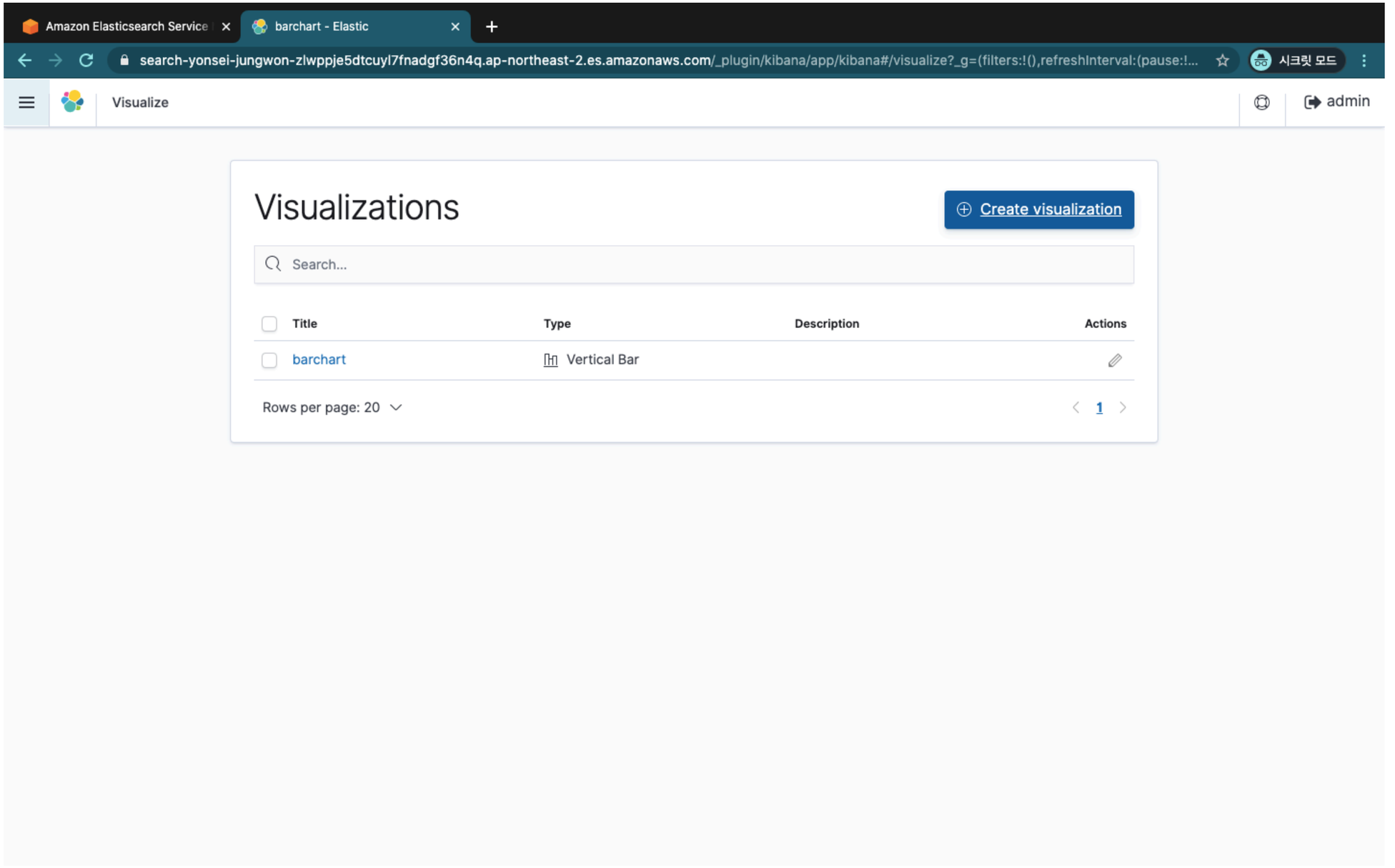
Basketball 인덱스 고르기



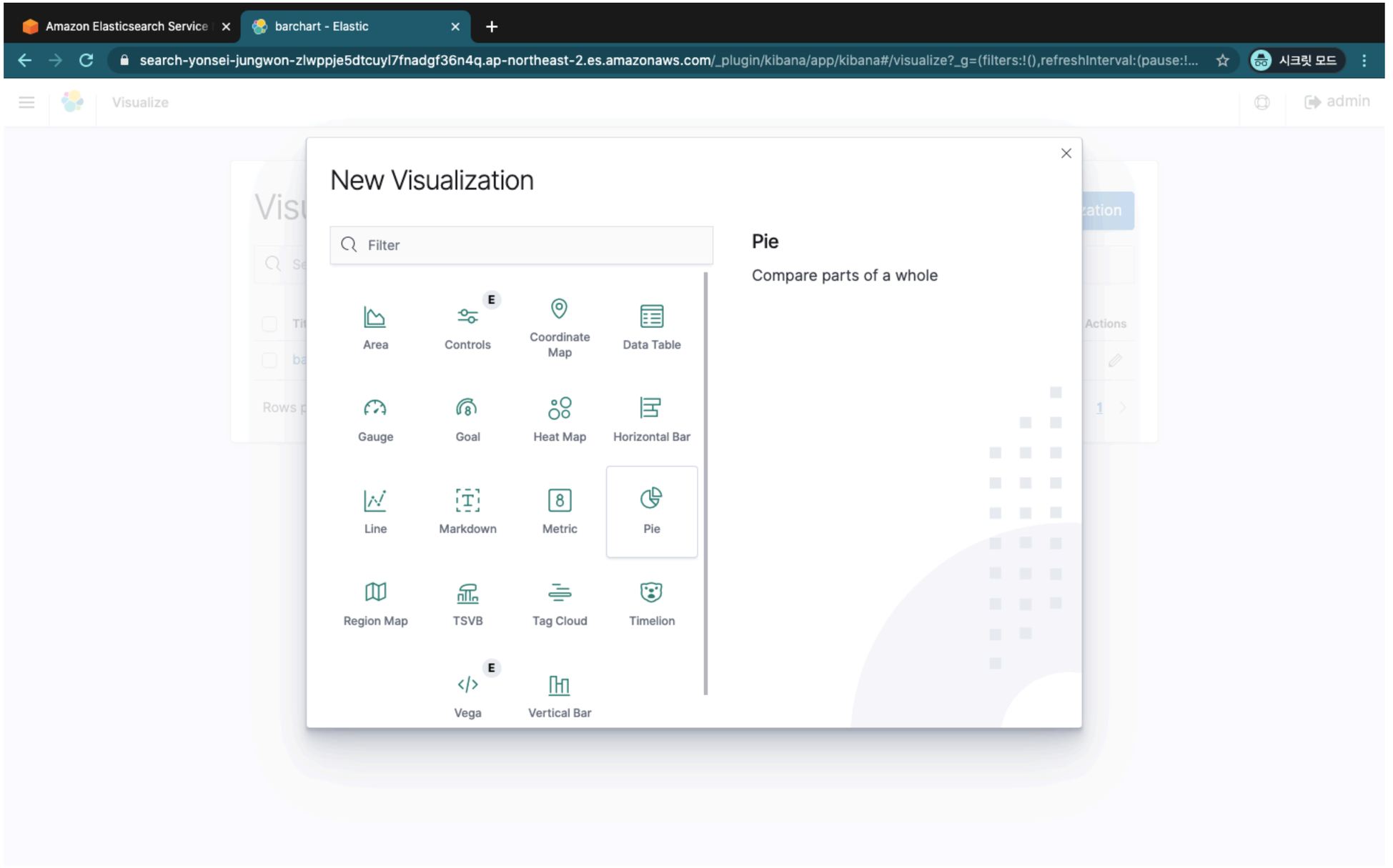
Metric과 Buckets 설정



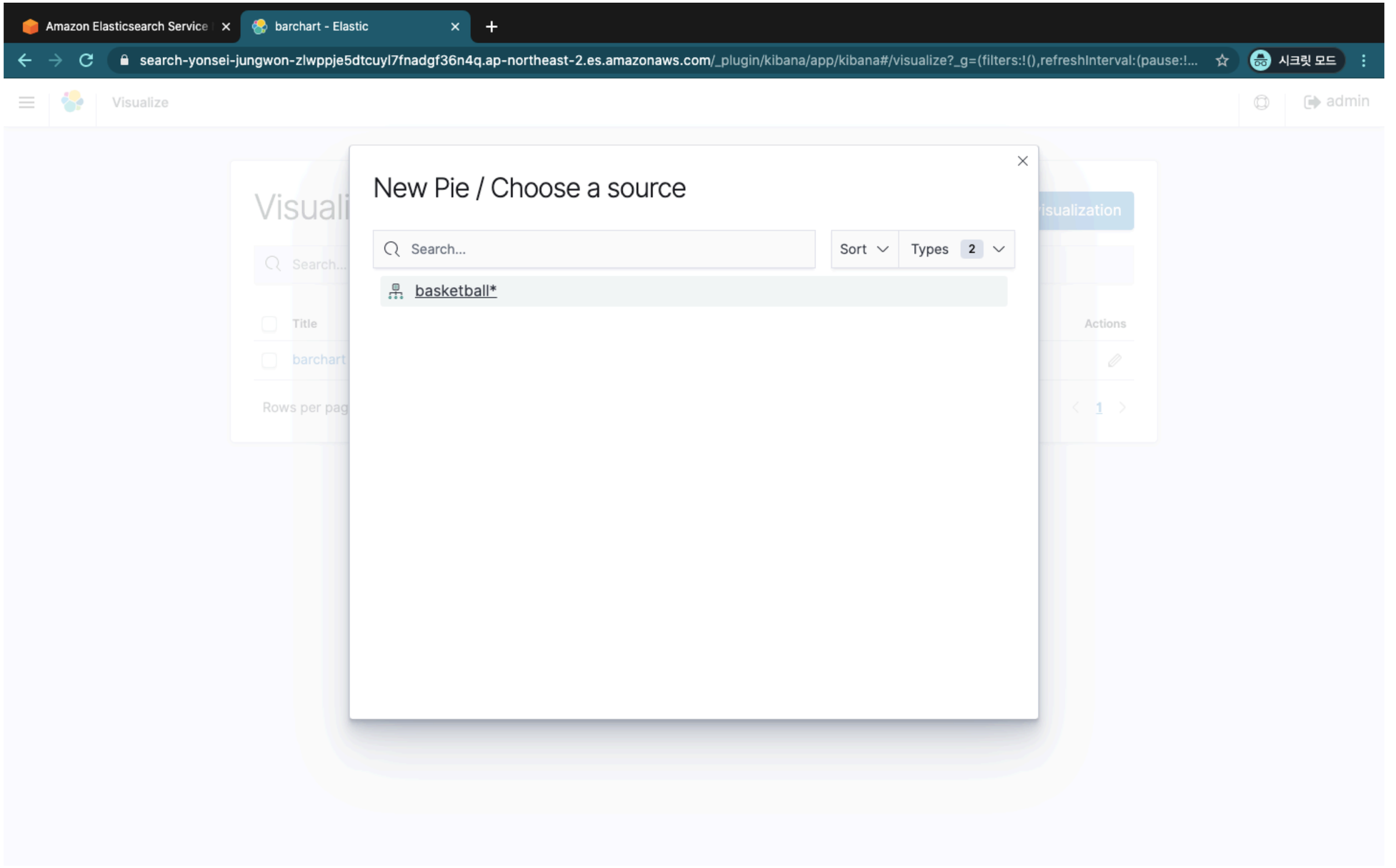
새로운 차트 하나 더 만들기



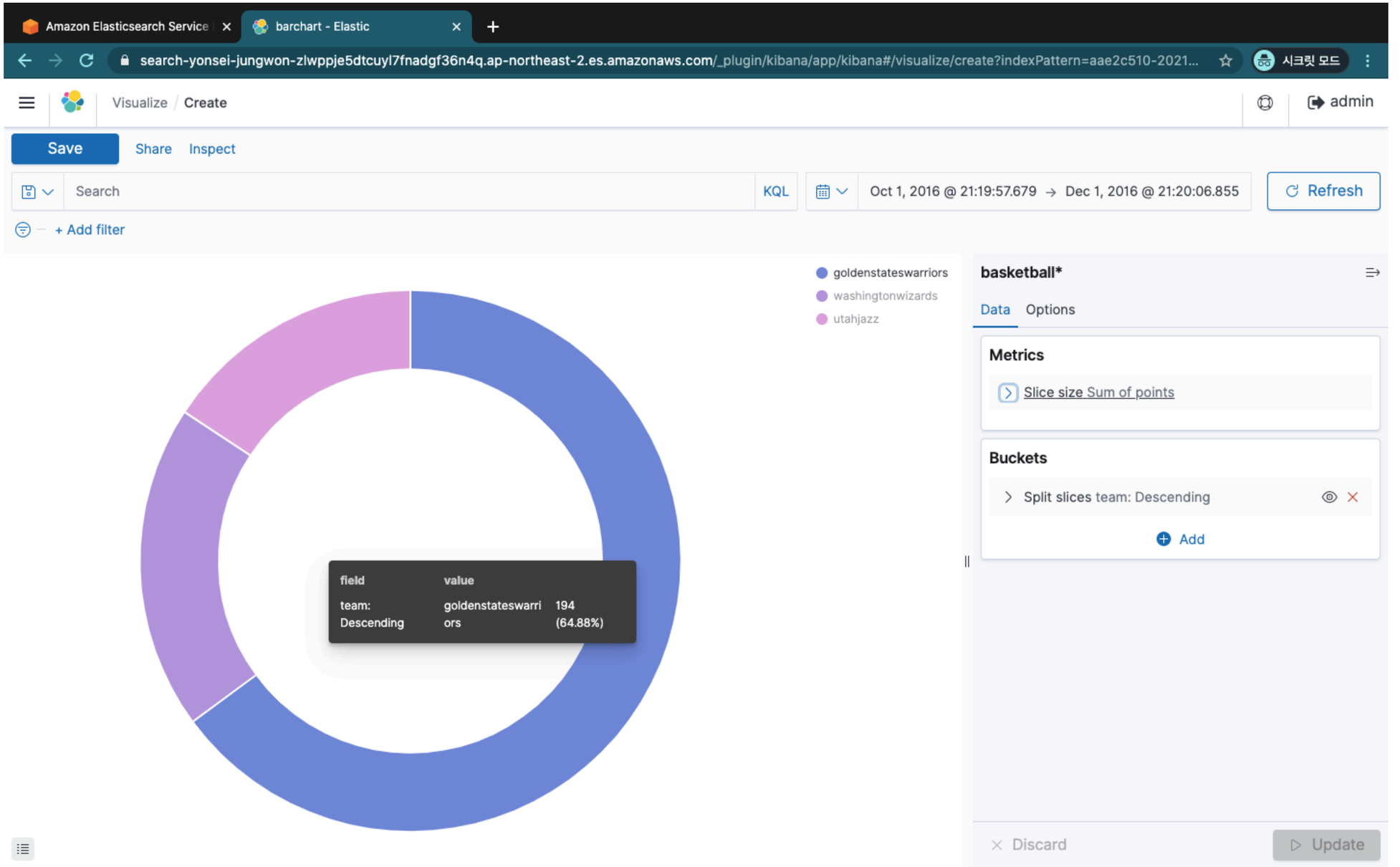
Pie Chart 고르기



마찬가지로 Basketball 고르기



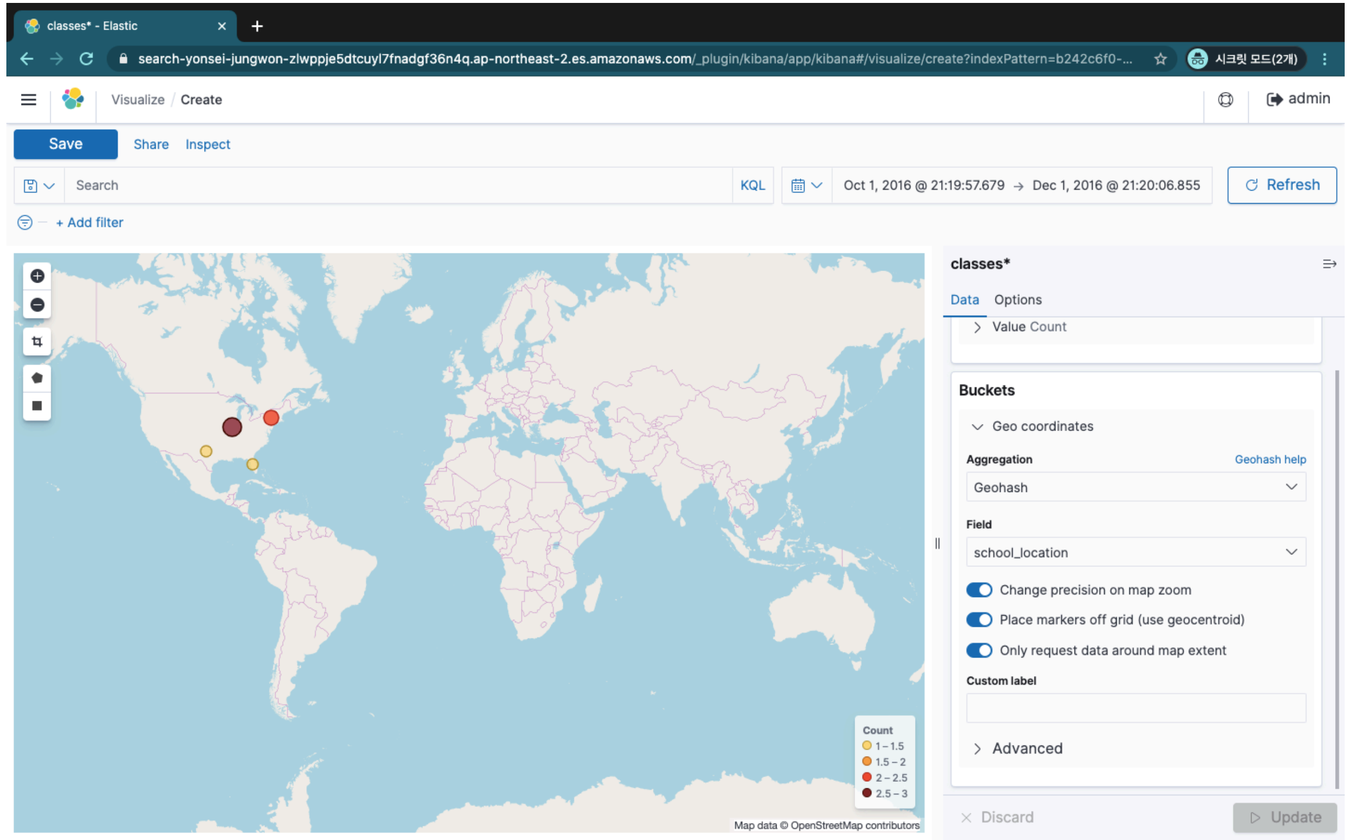
Metric과 Bucket 설정!



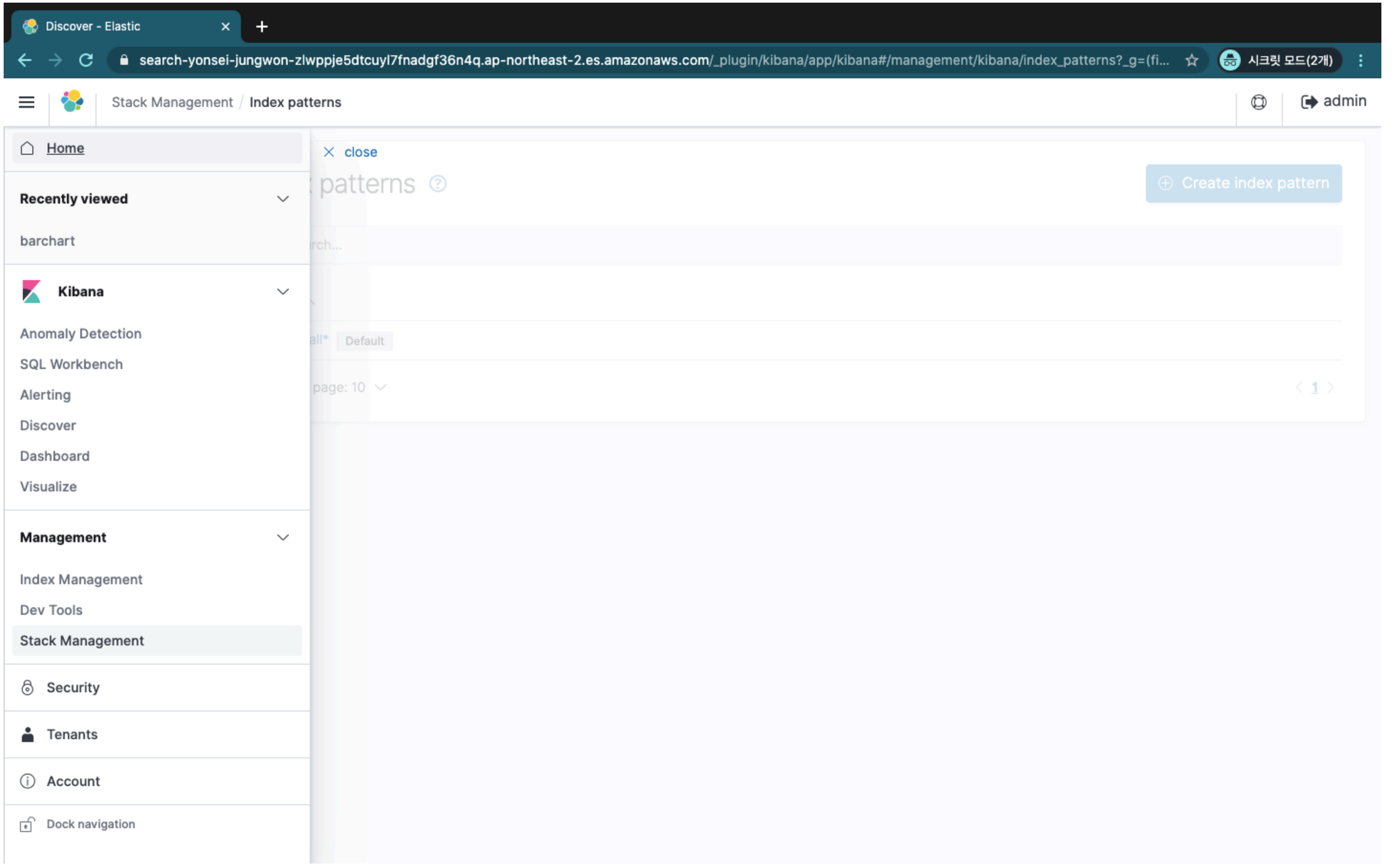
Kibana Visualize

Coordinated Map

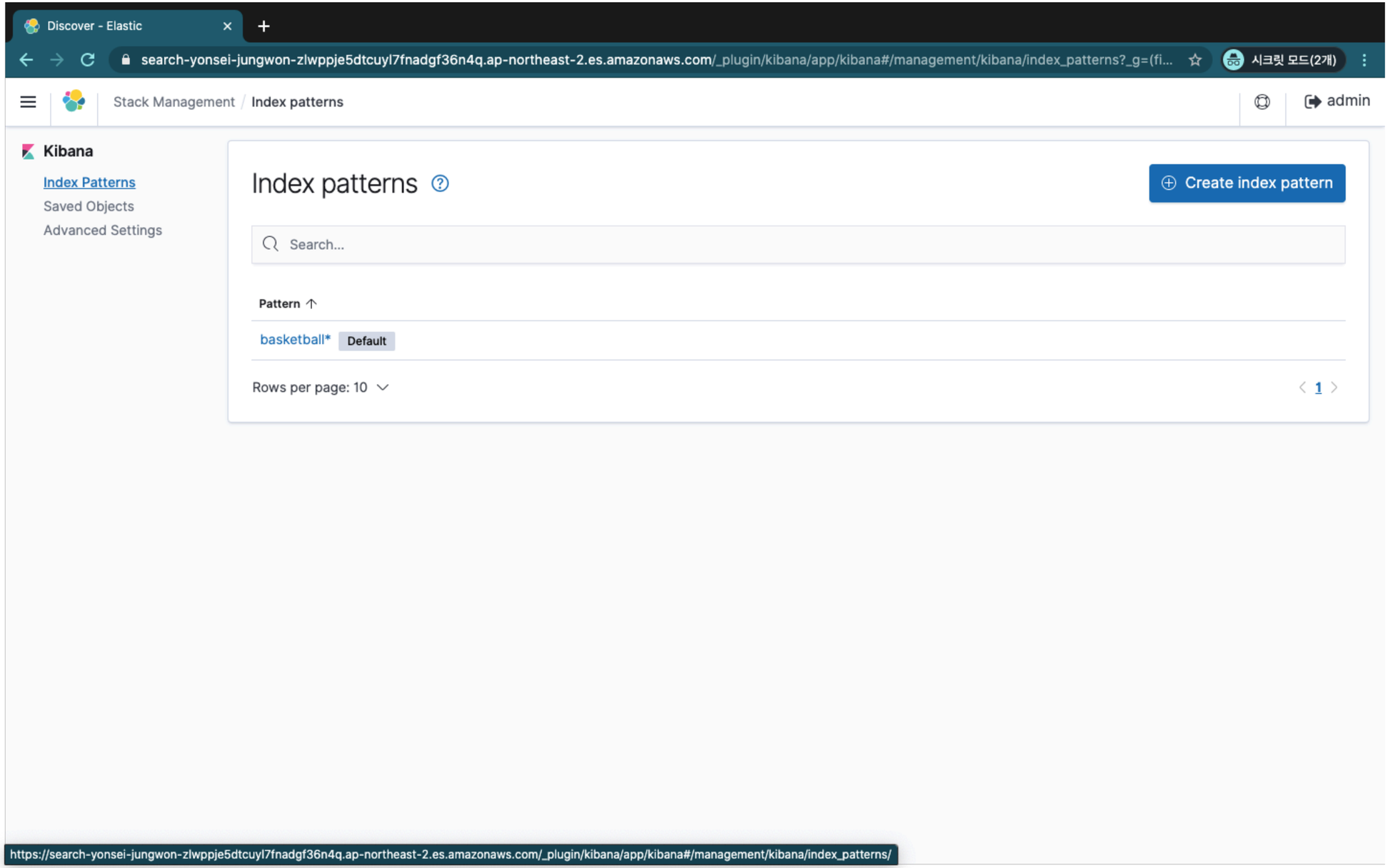
- 위도와 경도 데이터를 활용하여, 지도상에 해당하는 데이터 포인트를 출력!



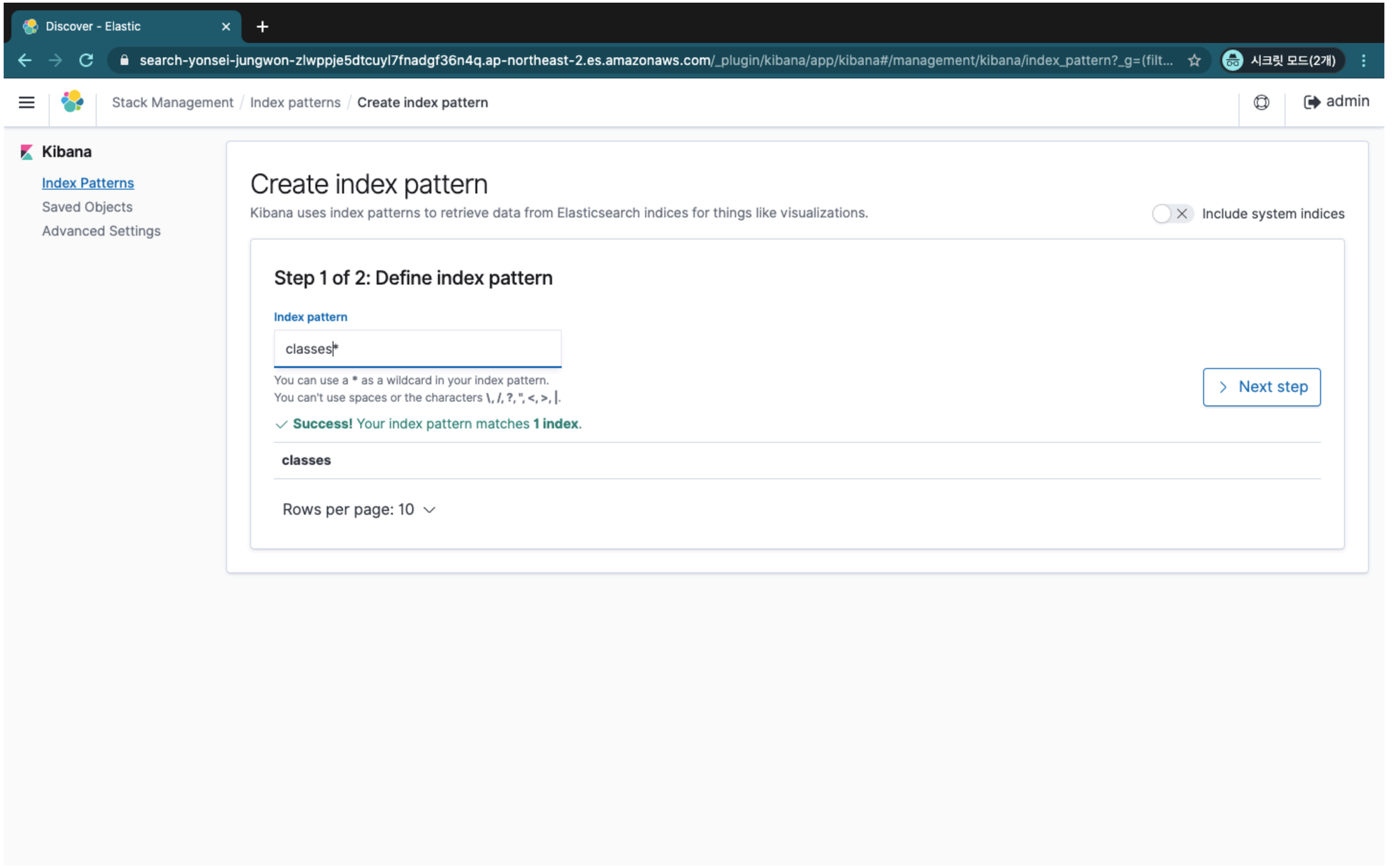
Stack Management 클릭



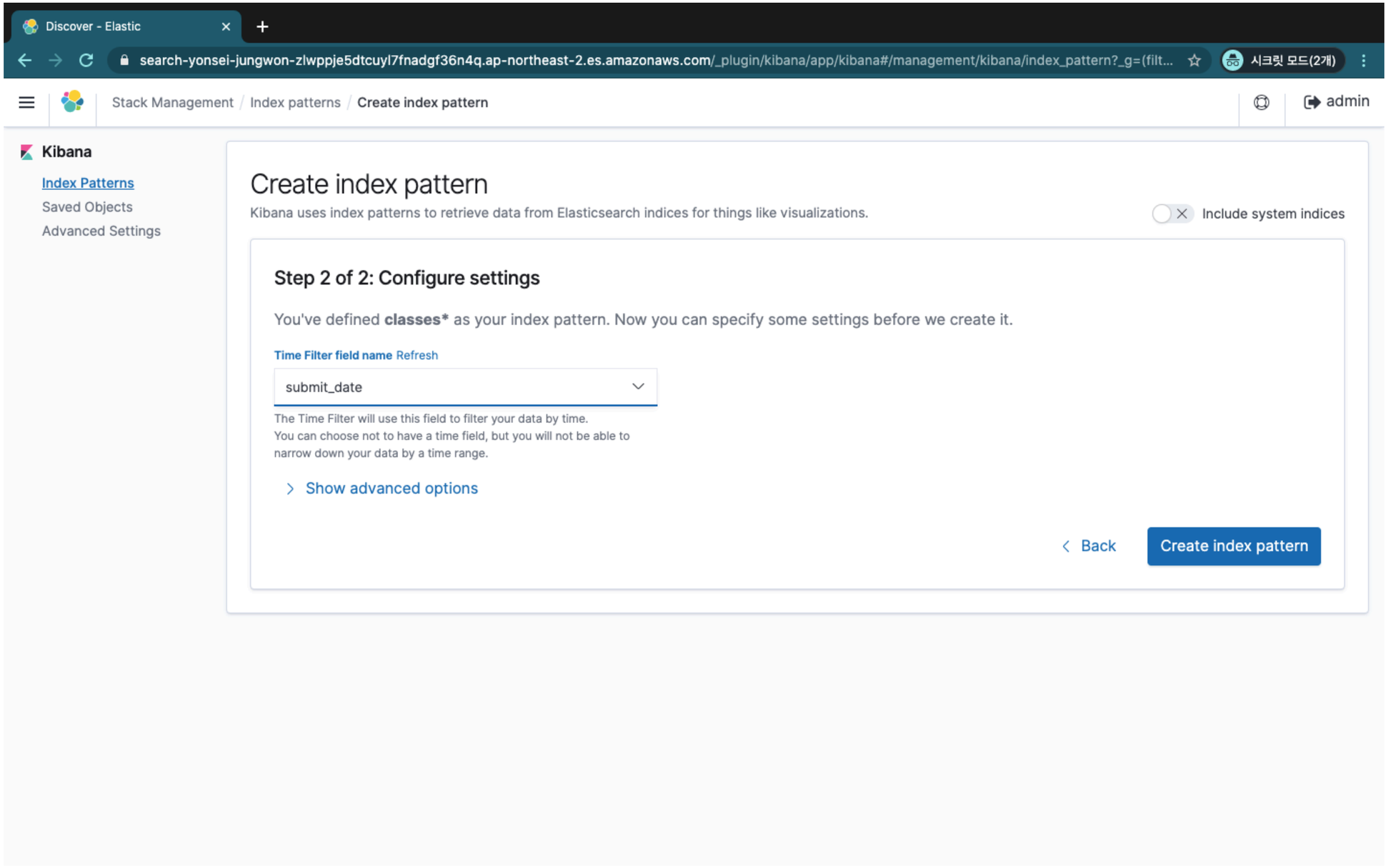
인덱스 패턴 생성!


Discover - Elastic
search-yonsei-jungwon-zlwppje5dtcuyl7fnadgf36n4q.ap-northeast-2.es.amazonaws.com/_plugin/kibana/app/kibana#/management/kibana/index_patterns?_g=(fi...
시크릿 모드(2개)
Stack Management / Index patterns
admin
Kibana
Index Patterns
Saved Objects
Advanced Settings
Index patterns
Create index pattern
Search...
Pattern
basketball*
Default
Rows per page: 10
1
https://search-yonsei-jungwon-zlwppje5dtcuyl7fnadgf36n4q.ap-northeast-2.es.amazonaws.com/_plugin/kibana/app/kibana#/management/kibana/index_patterns/

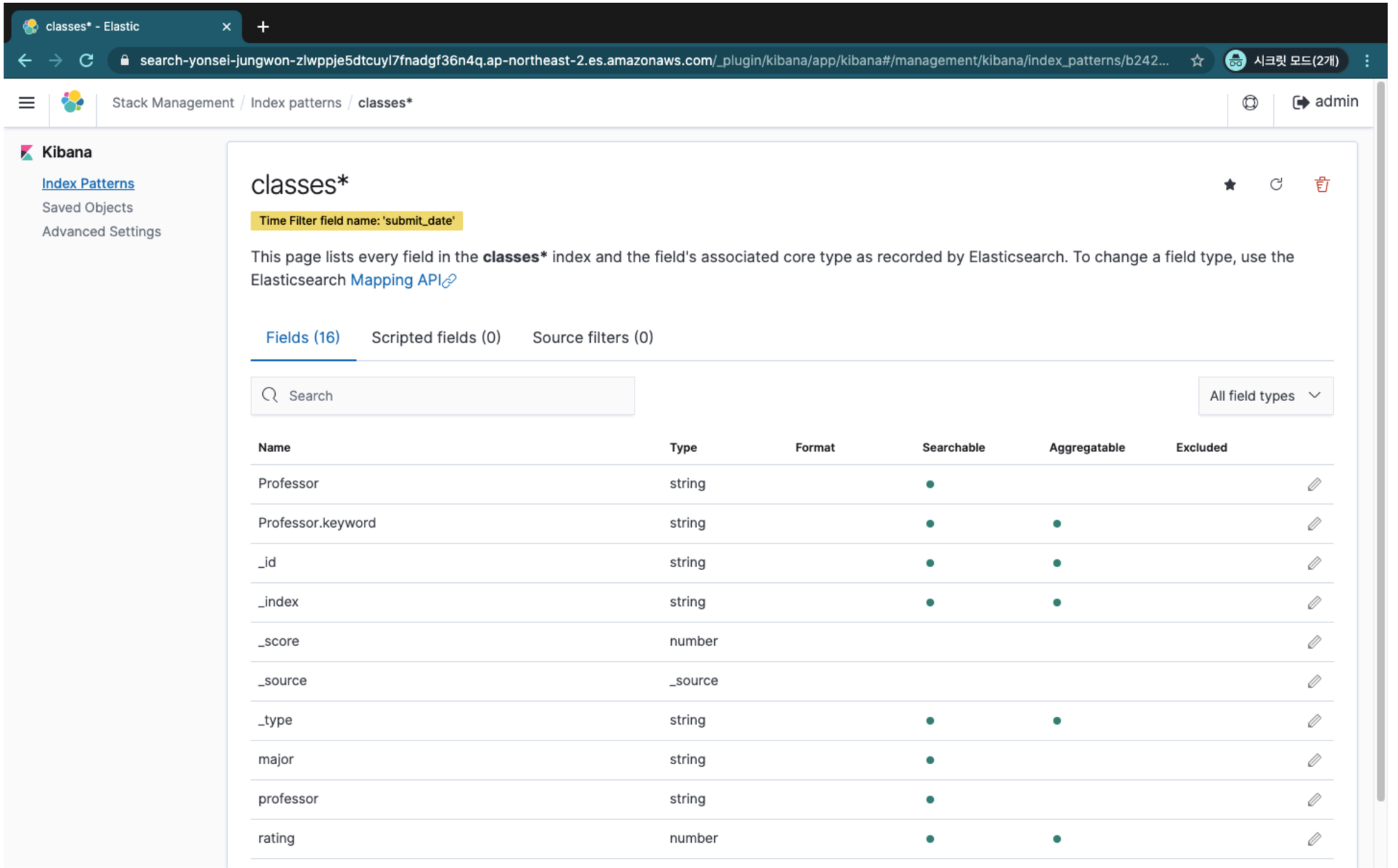
Classes 검색!


Discover - Elastic
search-yonsei-jungwon-zlwppje5dtcuyl7fnadgf36n4q.ap-northeast-2.es.amazonaws.com/_plugin/kibana/app/kibana#/management/kibana/index_pattern?_g=(filt...
시크릿 모드(2개)
Stack Management / Index patterns / Create index pattern
admin
Kibana
Index Patterns
Saved Objects
Advanced Settings
Create index pattern
Kibana uses index patterns to retrieve data from Elasticsearch indices for things like visualizations.
Include system indices
Step 1 of 2: Define index pattern
Index pattern
classes*
You can use a * as a wildcard in your index pattern.
You can't use spaces or the characters \, /, ?, ", <, >, |.
Next step
Success! Your index pattern matches 1 index.
classes
Rows per page: 10

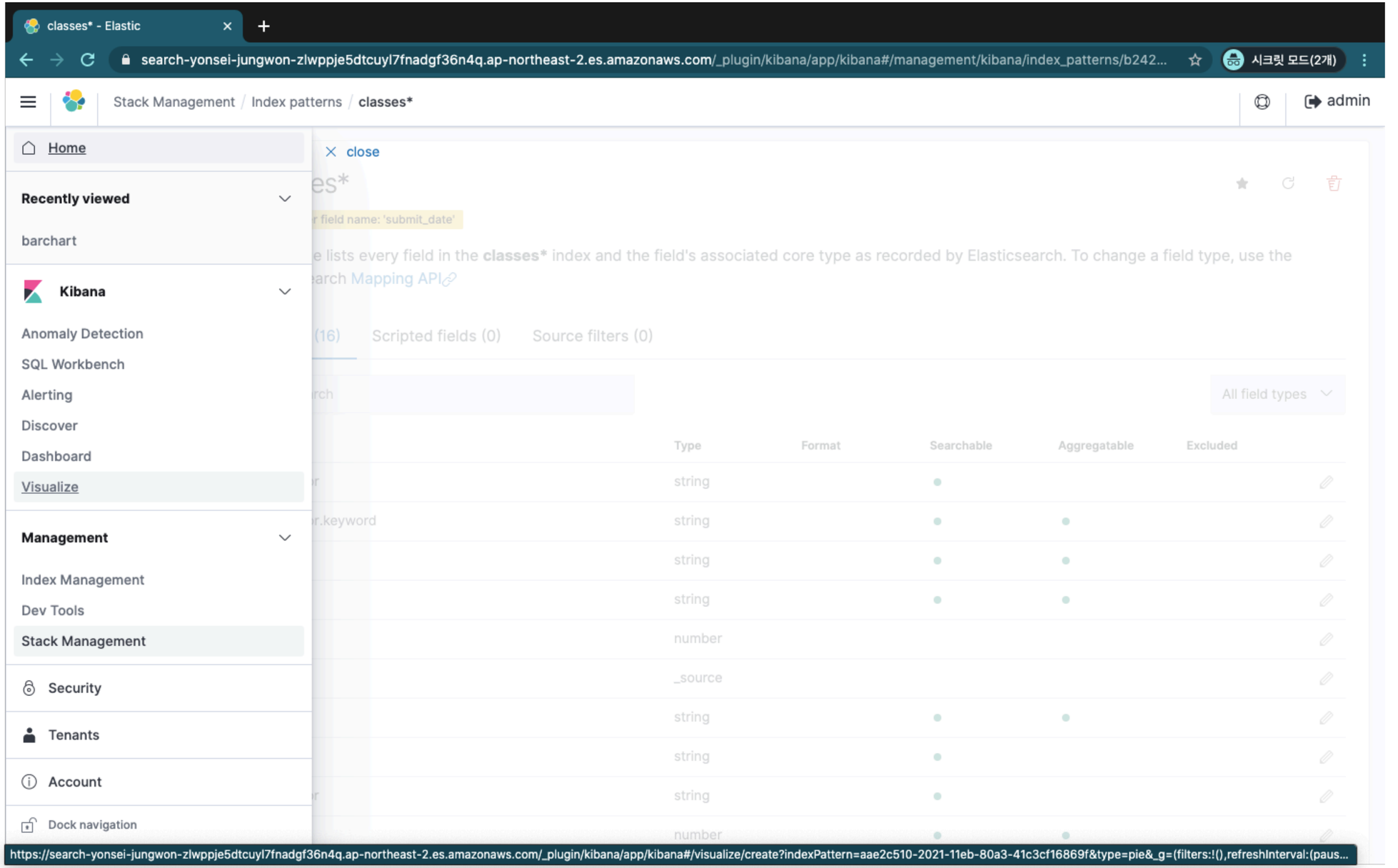
마찬가지로 Time Filter는 submit date



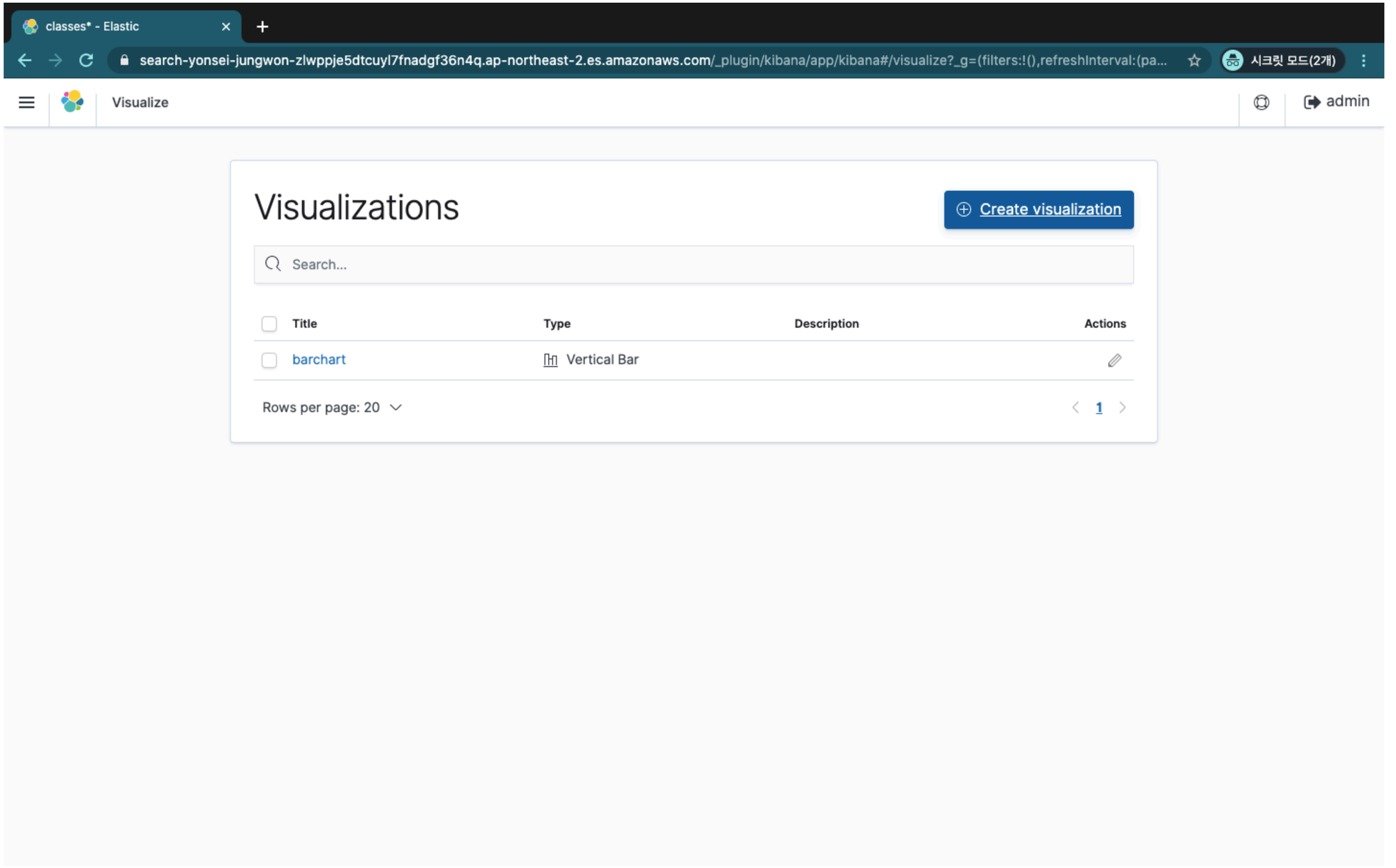
잘 들어갔는지 확인!


classes* - Elastic
search-yonsei-jungwon-zlwppje5dtcuyl7fnadgf36n4q.ap-northeast-2.es.amazonaws.com/_plugin/kibana/app/kibana#/management/kibana/index_patterns/b242...
시크릿 모드(2개)
Stack Management / Index patterns / classes*
admin
Kibana
Index Patterns
Saved Objects
Advanced Settings
classes*
Time Filter field name: 'submit_date'
This page lists every field in the classes* index and the field's associated core type as recorded by Elasticsearch. To change a field type, use the Elasticsearch Mapping API
Fields (16)
Scripted fields (0)
Source filters (0)
Search
All field types
Name
Type
Format
Searchable
Aggregatable
Excluded
Professor string
Professor.keyword string
_id string
_index string
_score number
_source _source
_type string
major string
professor string
rating number

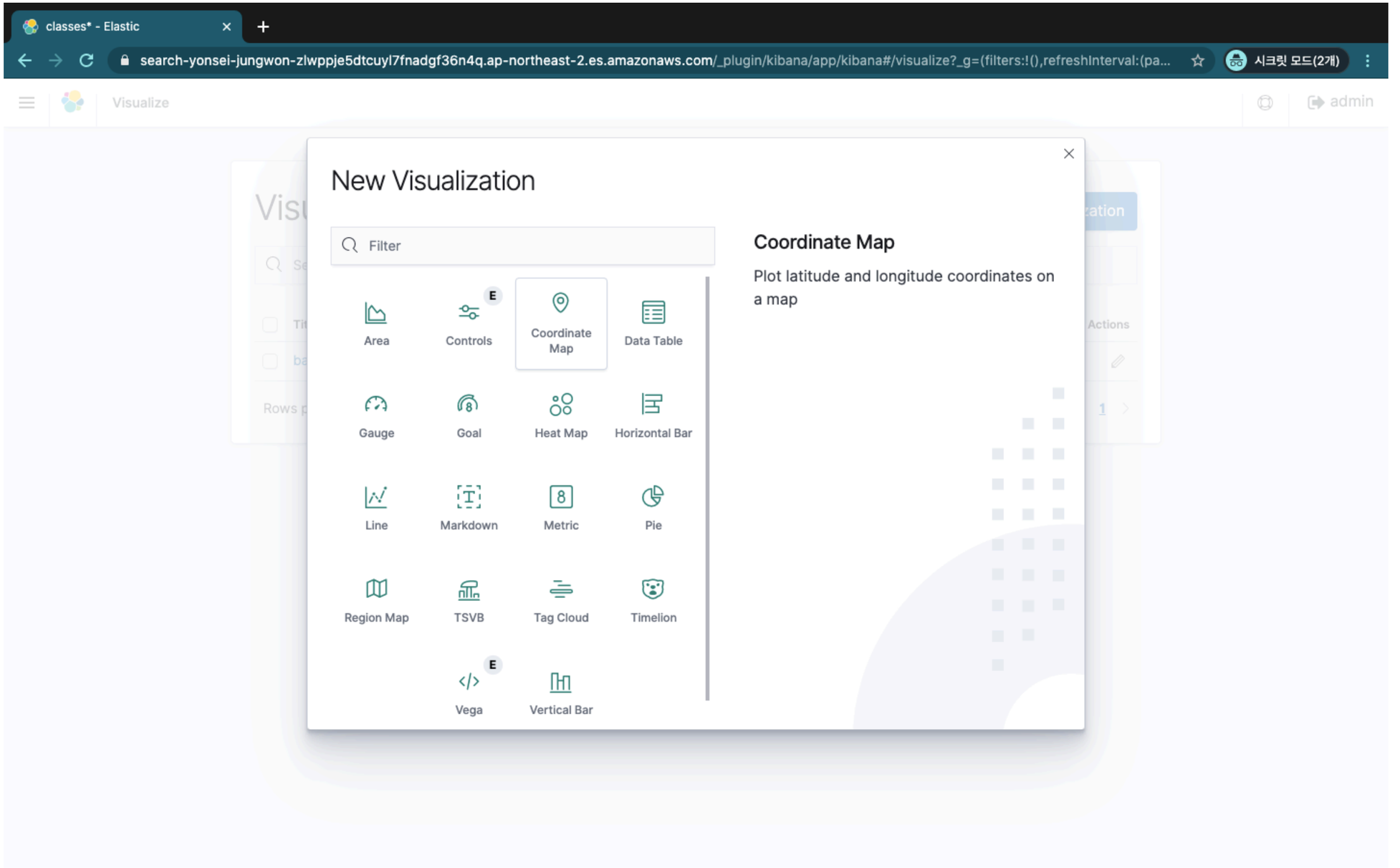
시각화 선택!


classes* - Elastic
search-yonsei-jungwon-zlwppje5dtcuyl7fnadgf36n4q.ap-northeast-2.es.amazonaws.com/_plugin/kibana/app/kibana#/management/kibana/index_patterns/b242...
시크릿 모드(2개)
Stack Management / Index patterns / classes*
admin
Home
Recently viewed
barchart
Kibana
Anomaly Detection
SQL Workbench
Alerting
Discover
Dashboard
Visualize
Management
Index Management
Dev Tools
Stack Management
Security
Tenants
Account
Dock navigation
close
Scripted fields (0)
Source filters (0)
All field types
Type
Format
Searchable
Aggregatable
Excluded
https://search-yonsei-jungwon-zlwppje5dtcuyl7fnadgf36n4q.ap-northeast-2.es.amazonaws.com/_plugin/kibana/app/kibana#/visualize/create?indexPattern=aae2c510-2021-11eb-80a3-41c3cf16869f&type=pie&_g=(filters:!(),refreshInterval:(paus...

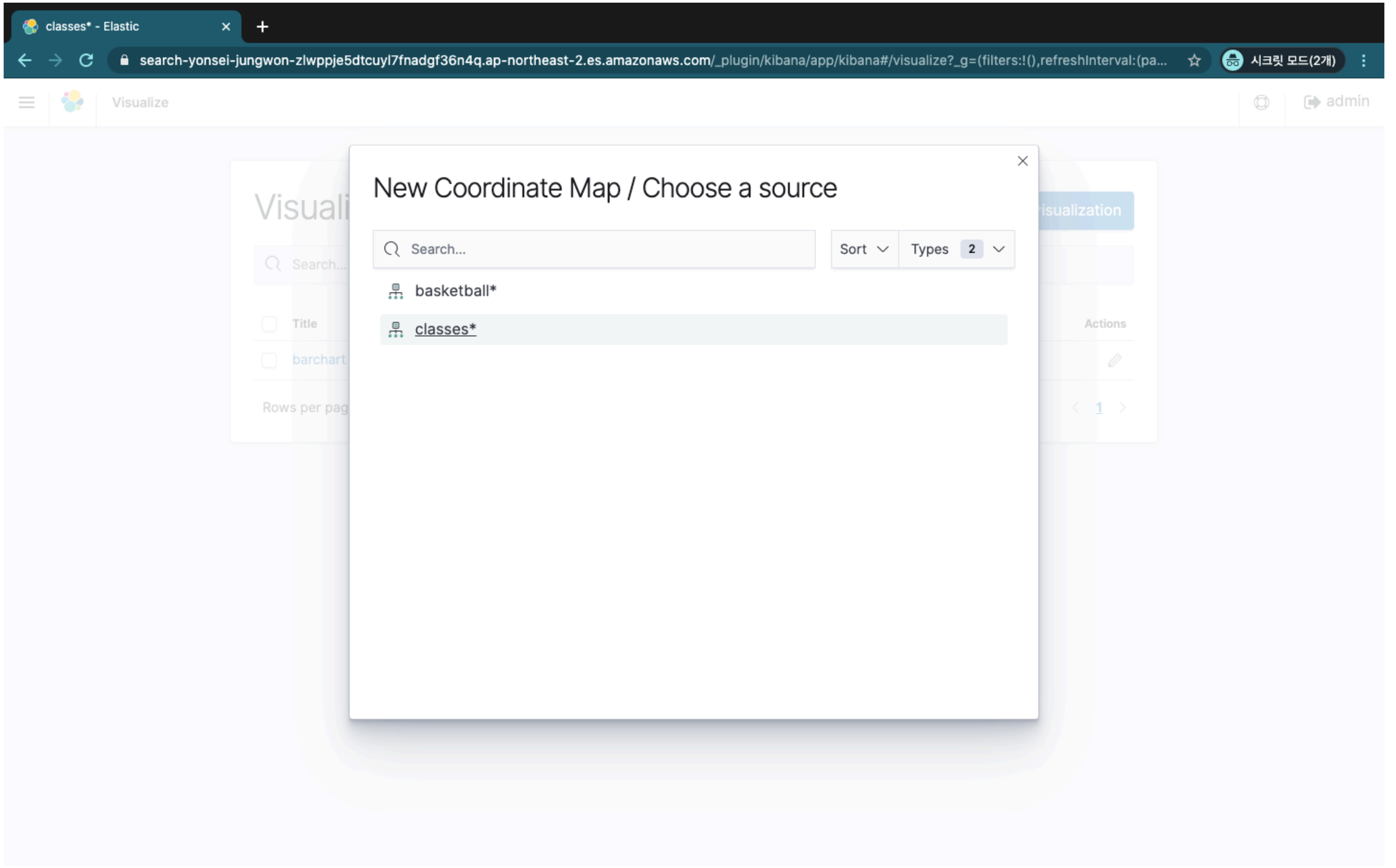
새 시각화 생성!


classes* - Elastic
search-yonsei-jungwon-zlwppje5dtcuyl7fnadgf36n4q.ap-northeast-2.es.amazonaws.com/_plugin/kibana/app/kibana#/visualize?_g=(filters:!(),refreshInterval:(pa...
시크릿 모드(2개)
Visualize
admin
Visualizations
Create visualization
Search...
Title
Type
Description
Actions
barchart
Vertical Bar
Rows per page: 20
1

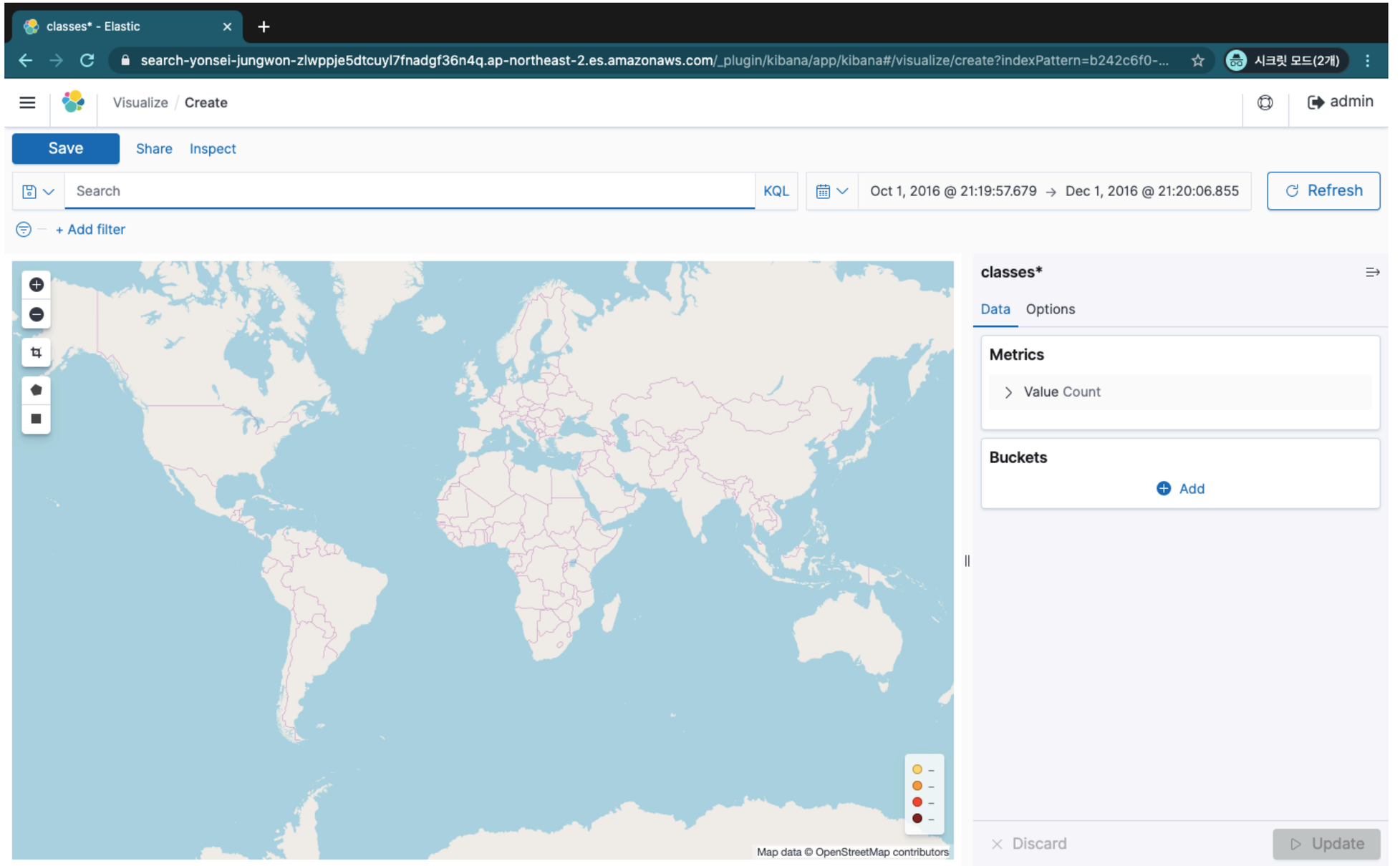
Coordinate Map 선택



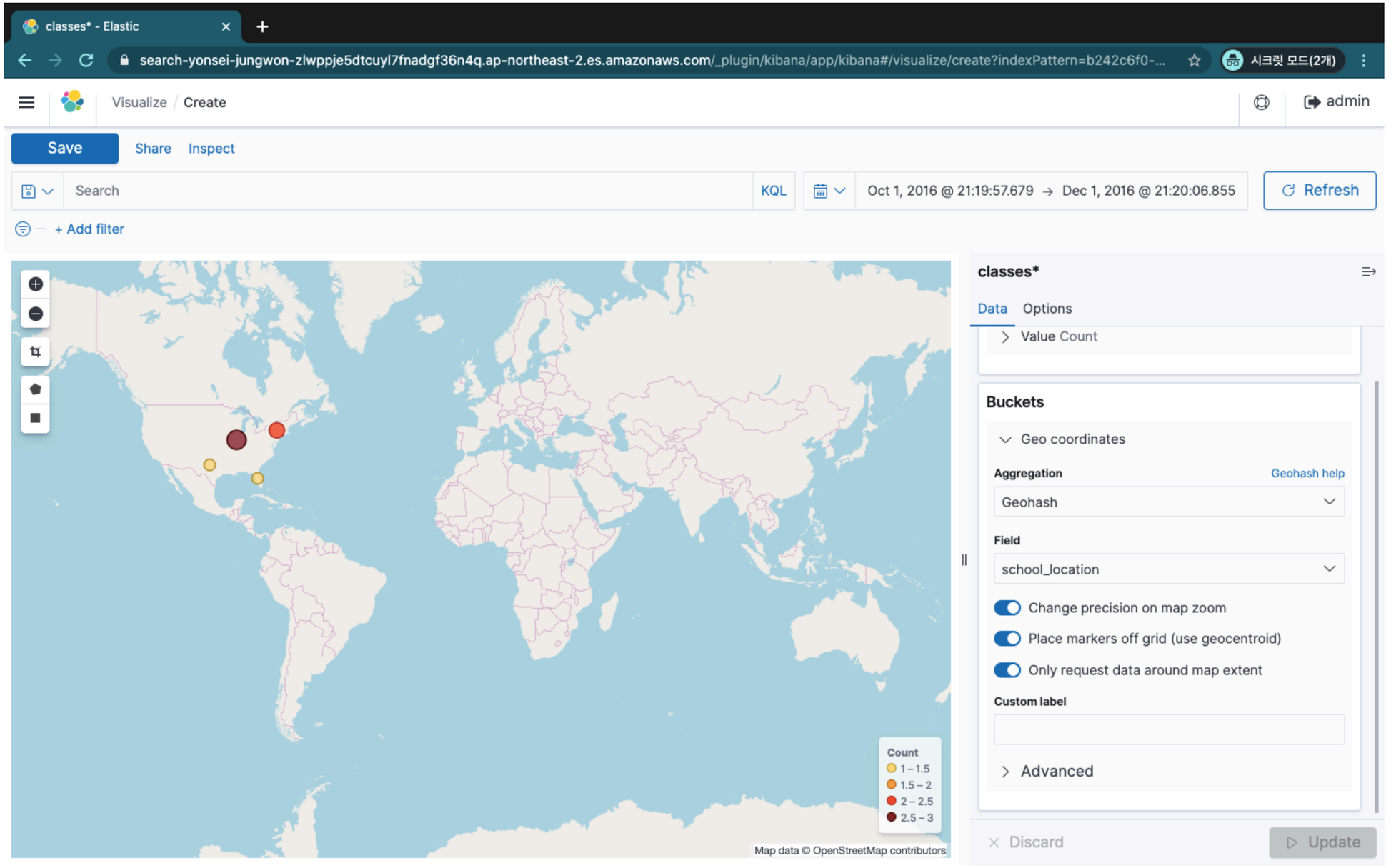
Classes 선택


classes* - Elastic
search-yonsei-jungwon-zlwppje5dtcuyl7fnadgf36n4q.ap-northeast-2.es.amazonaws.com/_plugin/kibana/app/kibana#/visualize?_g=(filters:!(),refreshInterval:(pa...
시크릿 모드(2개)
Visualize
admin
New Coordinate Map / Choose a source
Search...
Sort
Types
2
basketball*
classes*

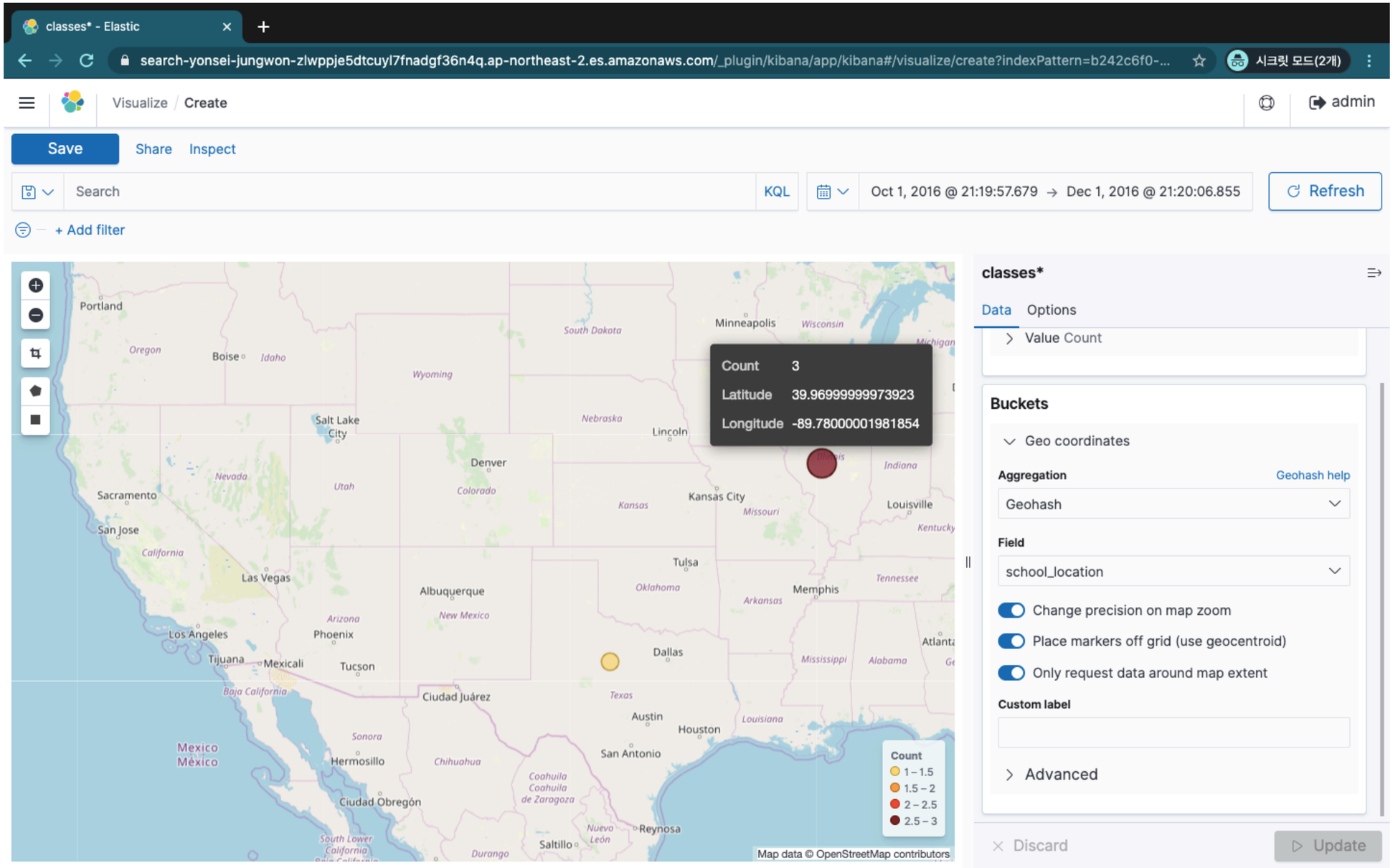
Metric 설정


classes* - Elastic
search-yonsei-jungwon-zlwppje5dtcuyl7fnadgf36n4q.ap-northeast-2.es.amazonaws.com/_plugin/kibana/app/kibana#/visualize/create?indexPattern=b242c6f0-...
시크릿 모드(2개)
Visualize / Create
admin
Save
Share
Inspect
Search
KQL
Oct 1, 2016 @ 21:19:57.679 → Dec 1, 2016 @ 21:20:06.855
Refresh
+ Add filter
Map data © OpenStreetMap contributors
classes*
Data
Options
Metrics
Value Count
Buckets
Add
Discard
Update

Buckets 설정

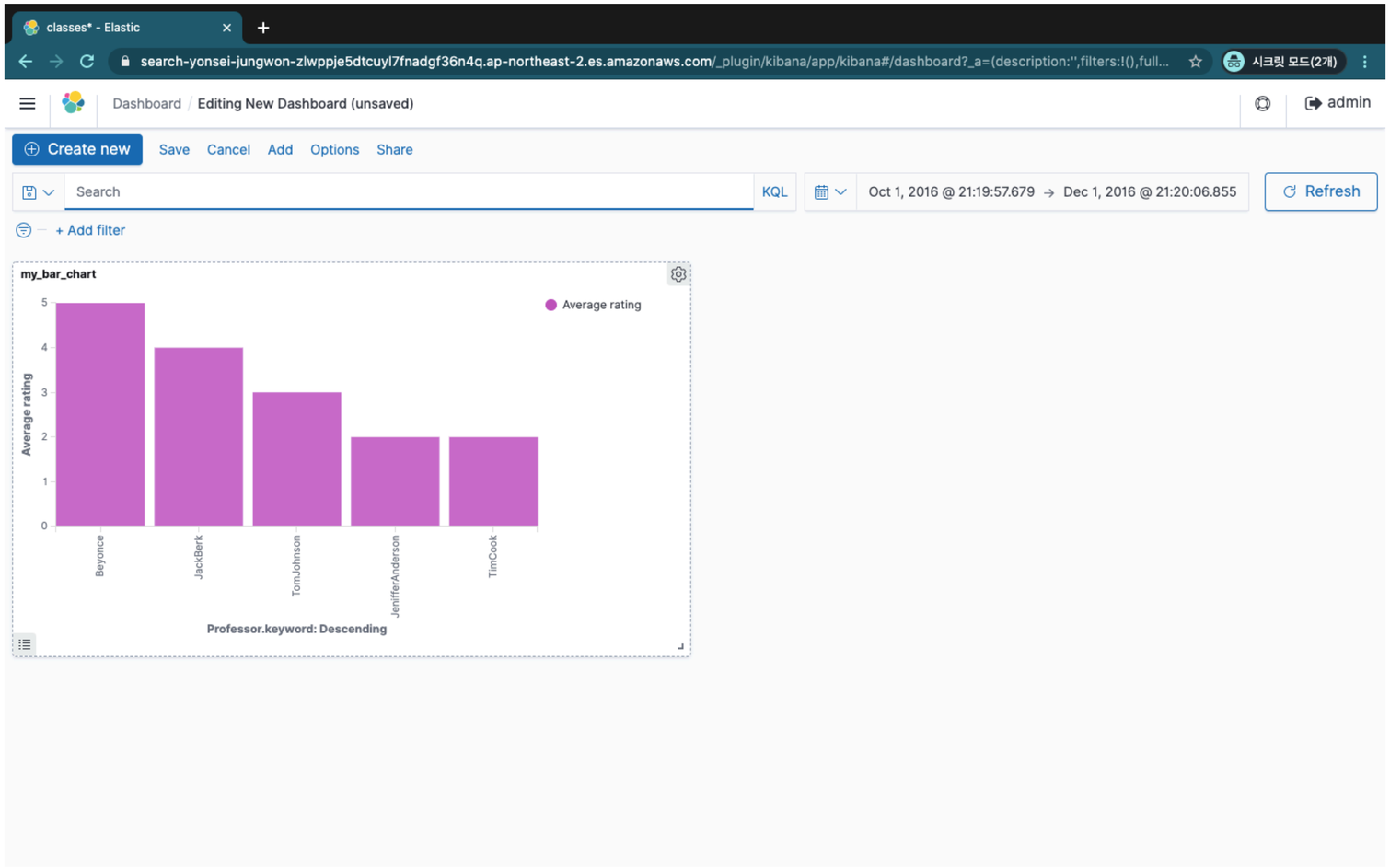

classes* - Elastic
search-yonsei-jungwon-zlwppje5dtcuyl7fnadgf36n4q.ap-northeast-2.es.amazonaws.com/_plugin/kibana/app/kibana#/visualize/create?indexPattern=b242c6f0-...
시크릿 모드(2개)
Visualize / Create
admin
Save
Share
Inspect
Search
KQL
Oct 1, 2016 @ 21:19:57.679 → Dec 1, 2016 @ 21:20:06.855
Refresh
+ Add filter
Count
1 – 1.5
1.5 – 2
2 – 2.5
2.5 – 3
Map data © OpenStreetMap contributors
classes*
Data
Options
Value Count
Buckets
Geo coordinates
Aggregation
Geohash help
Geohash
Field
school_location
Change precision on map zoom
Place markers off grid (use geocentroid)
Only request data around map extent
Custom label
Advanced
Discard
Update

짜잔!

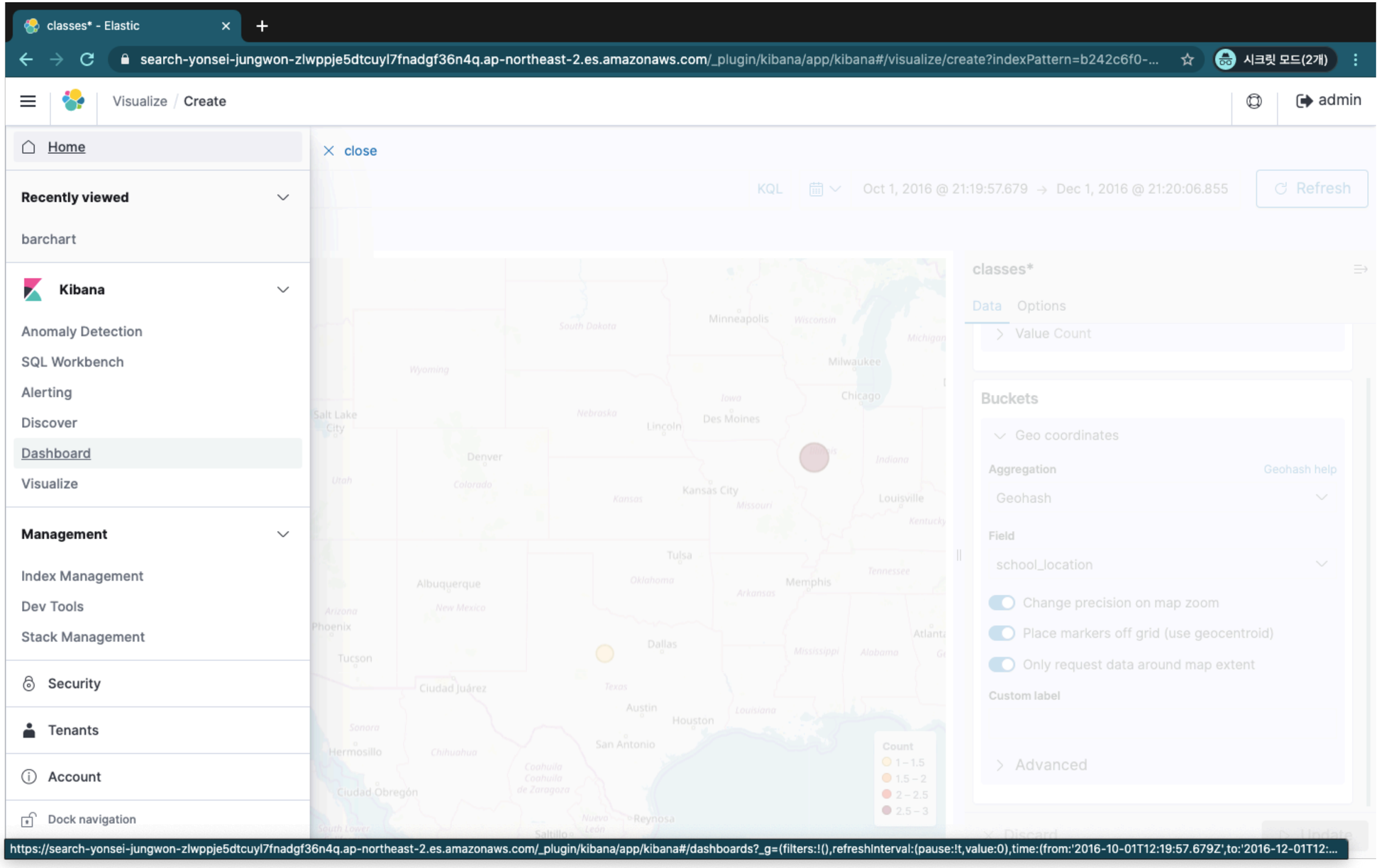

classes* - Elastic
search-yonsei-jungwon-zlwppje5dtcuyl7fnadgf36n4q.ap-northeast-2.es.amazonaws.com/_plugin/kibana/app/kibana#/visualize/create?indexPattern=b242c6f0-...
시크릿 모드(2개)
Visualize / Create
admin
Save
Share
Inspect
Search
KQL
Oct 1, 2016 @ 21:19:57.679 → Dec 1, 2016 @ 21:20:06.855
Refresh
+ Add filter
Count 3
Latitude 39.96999999973923
Longitude -89.78000001981854
Portland
Oregon
Boise
Idaho
Wyoming
South Dakota
Minneapolis
Wisconsin
Salt Lake City
Nebraska
Lincoln
Denver
Colorado
Nevada
Utah
Sacramento
San Jose
California
Kansas
Kansas City
Missouri
Indiana
Louisville
Kentucky
Las Vegas
Albuquerque
New Mexico
Arizona
Phoenix
Tulsa
Oklahoma
Memphis
Arkansas
Tennessee
Los Angeles
Tijuana
Mexicali
Tucson
Baja California
Dallas
Texas
Ciudad Juárez
Austin
Houston
Louisiana
Mississippi
Alabama
Atlanta
Sonora
Hermosillo
Chihuahua
San Antonio
Coahuila
Coahuila de Zaragoza
Ciudad Obregón
Nuevo León
Reynosa
Saltillo
Durango
South Lower California
Mexico
México
Count
1 – 1.5
1.5 – 2
2 – 2.5
2.5 – 3
Map data © OpenStreetMap contributors
classes*
Data
Options
Value Count
Buckets
Geo coordinates
Aggregation
Geohash help
Geohash
Field
school_location
Change precision on map zoom
Place markers off grid (use geocentroid)
Only request data around map extent
Custom label
Advanced
Discard
Update

Kibana Dashboard

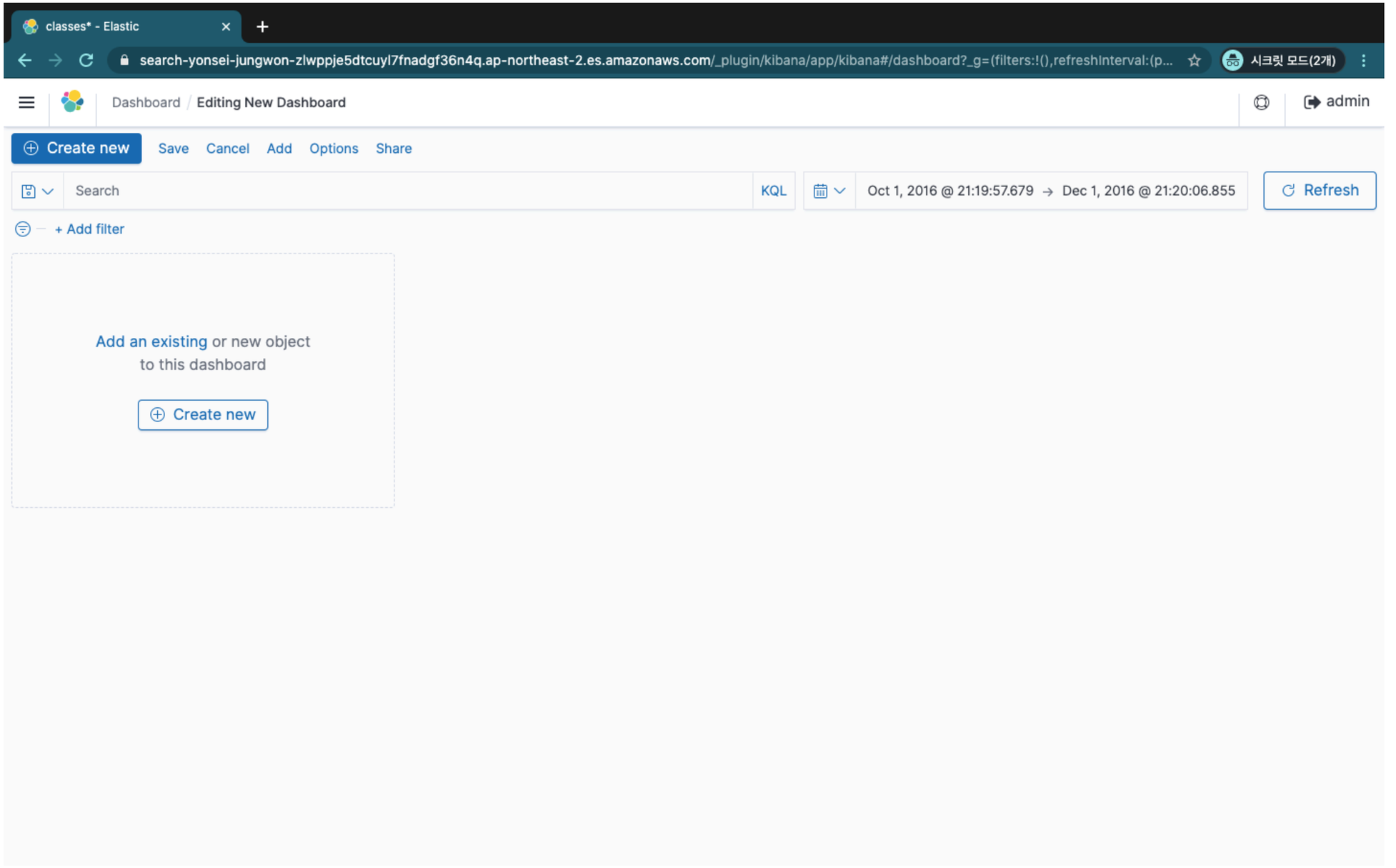
생성한 시각화나 각종 수치들을 한눈에!



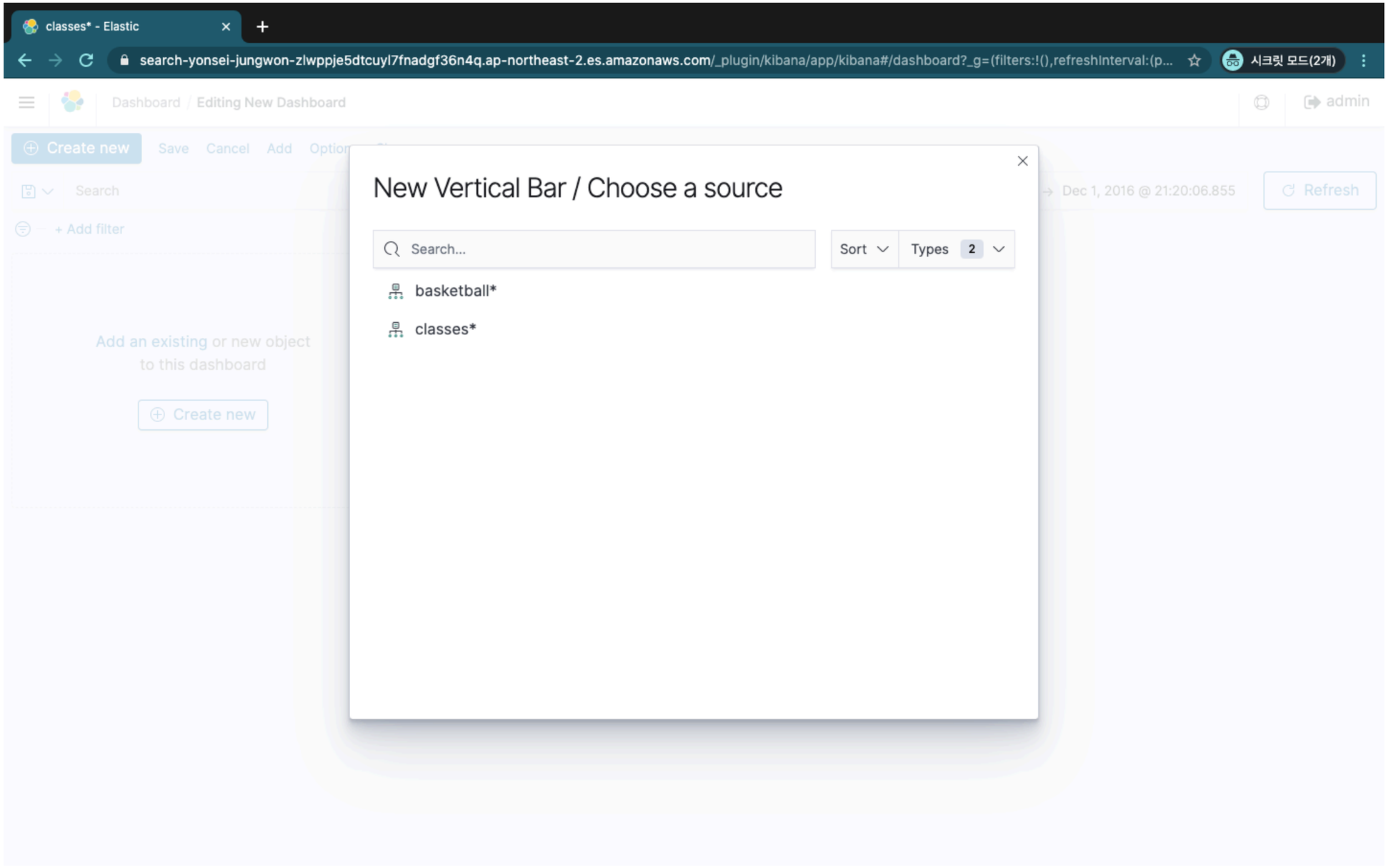
Dashboard 클릭!



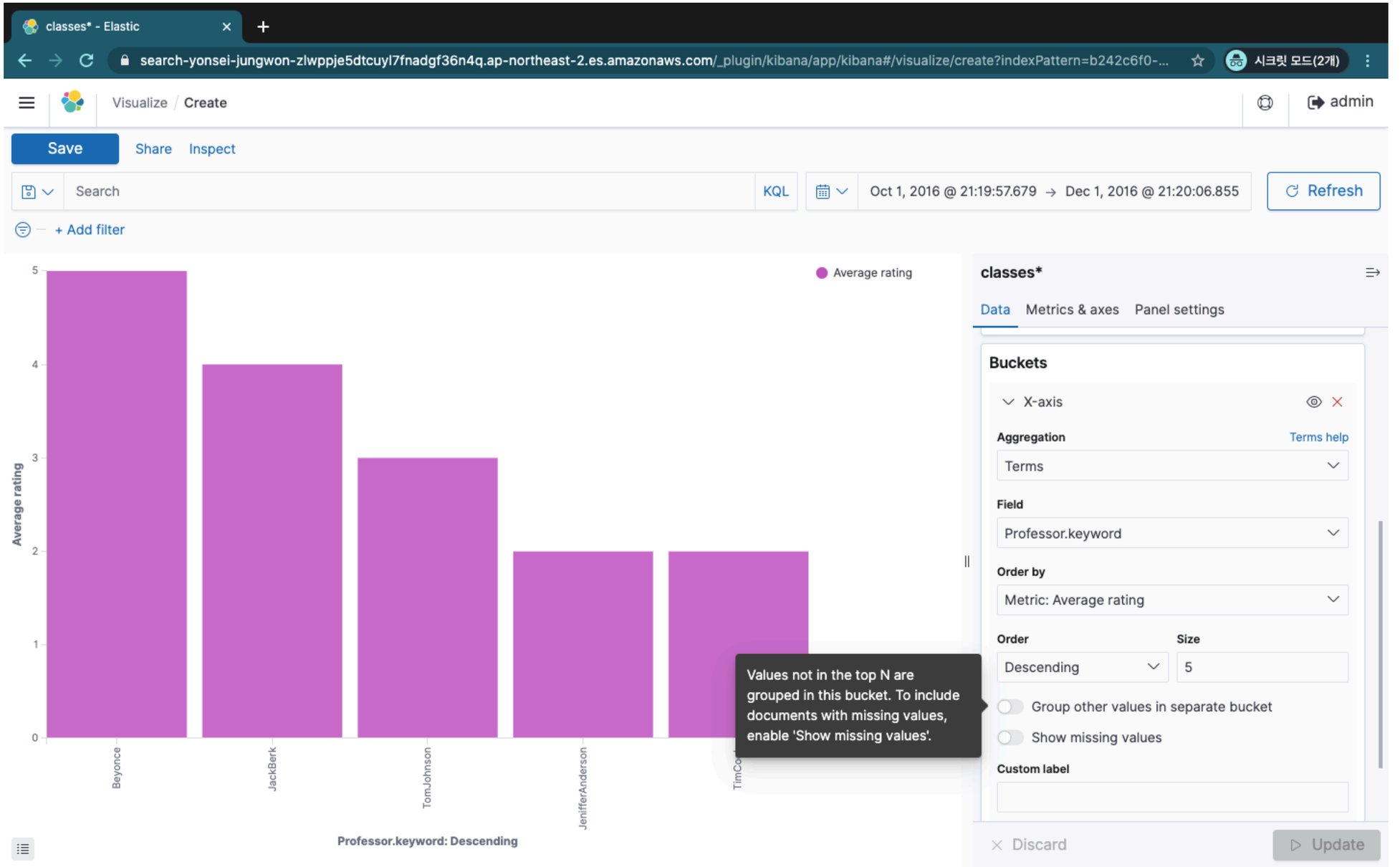
Create New 또는 Add an existing 클릭 (여기선 Create New)



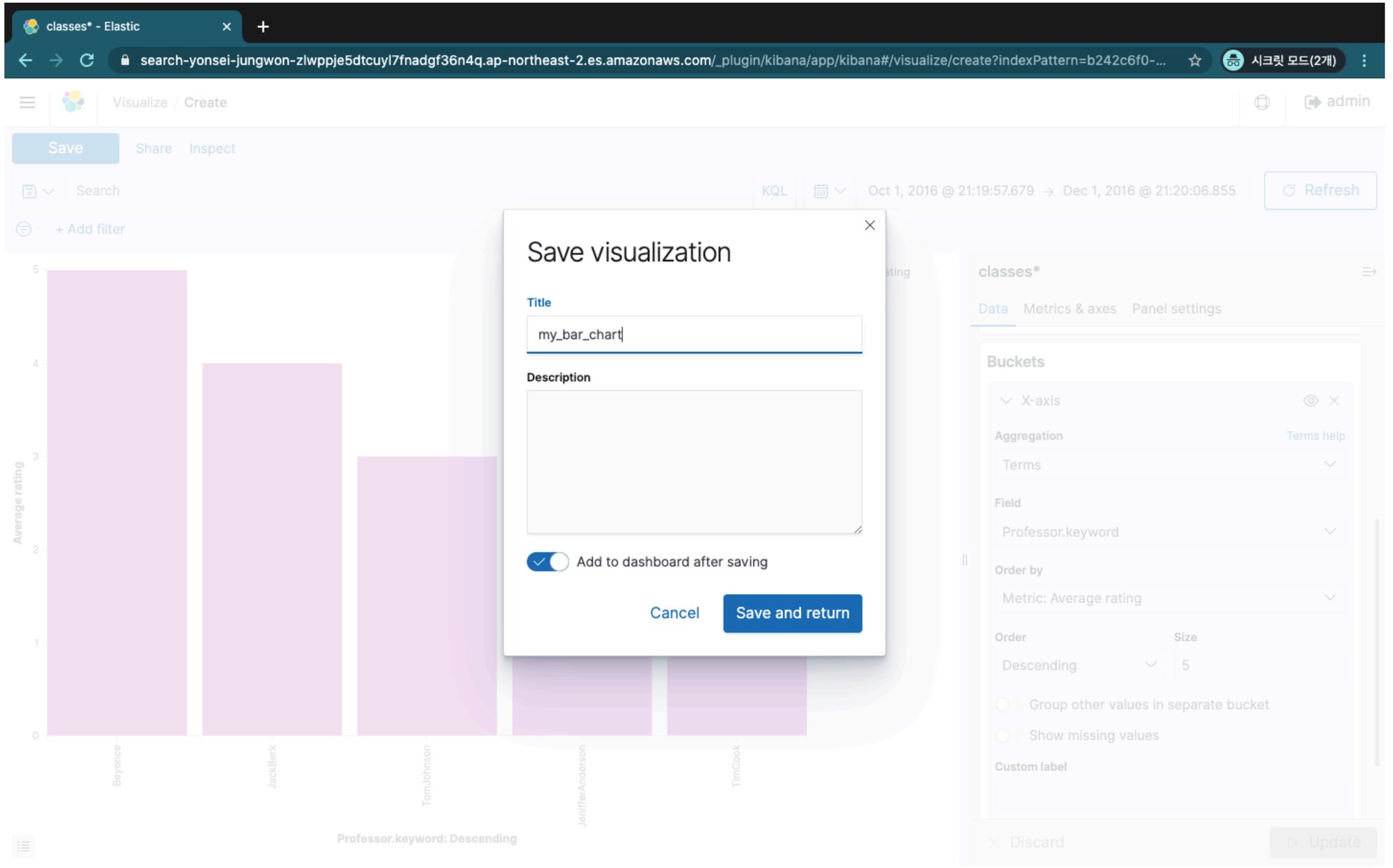
Classes 선택


classes* - Elastic
search-yonsei-jungwon-zlwppje5dtcuyl7fnadgf36n4q.ap-northeast-2.es.amazonaws.com/_plugin/kibana/app/kibana#/dashboard?_g=(filters:!(),refreshInterval:(p...
시크릿 모드(2개)
Dashboard / Editing New Dashboard
admin
Create new
Save
Cancel
Add
Search
Dec 1, 2016 @ 21:20:06.855
Refresh
+ Add filter
Add an existing or new object to this dashboard
Create new
New Vertical Bar / Choose a source
Search...
Sort
Types 2
basketball*
classes*

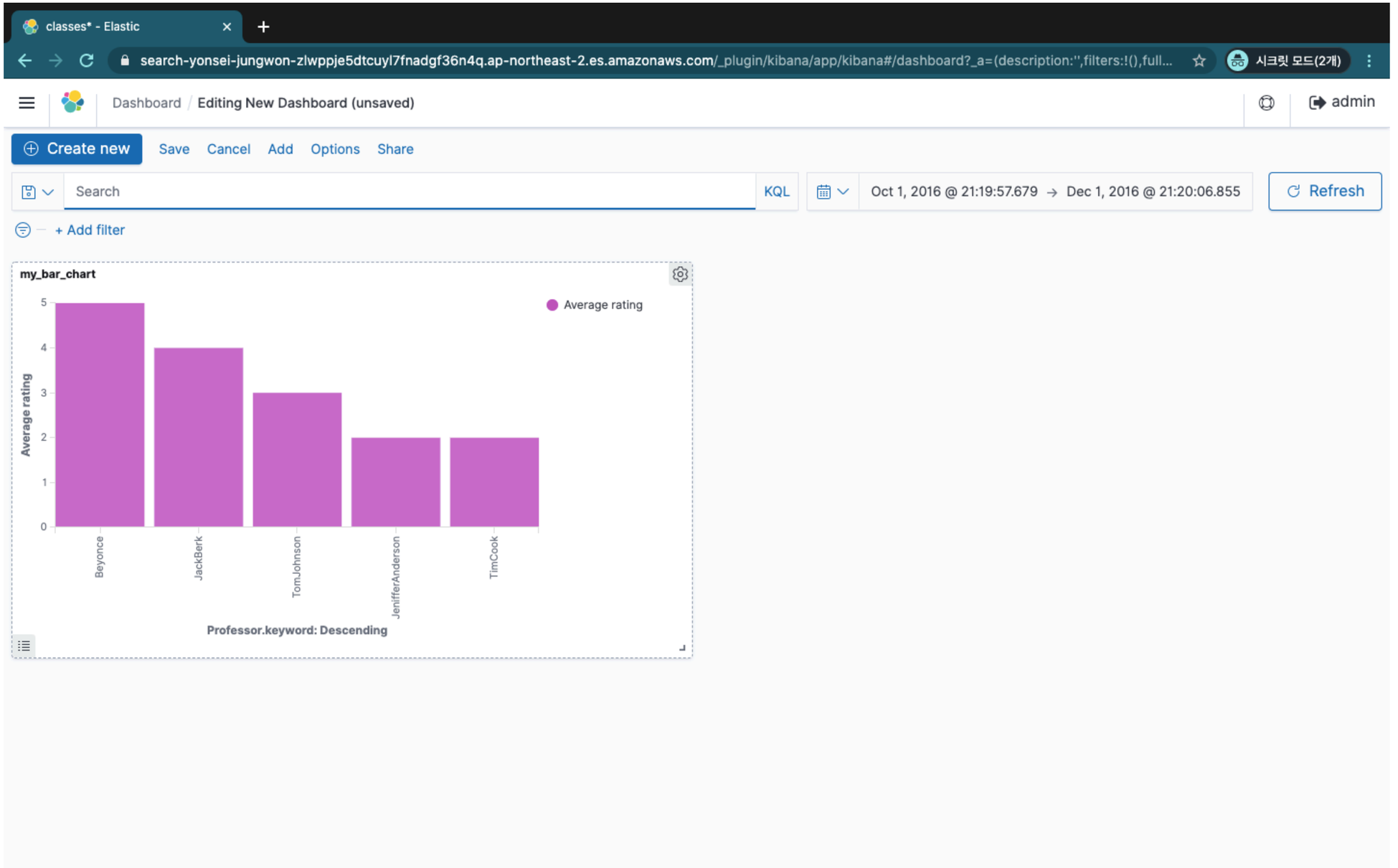
적당히 Metric과 Bucket 설정으로 왼쪽과 같은 Vertical Chart 만들기!



저장!


classes* - Elastic
search-yonsei-jungwon-zlwppje5dtcuyl7fnadgf36n4q.ap-northeast-2.es.amazonaws.com/_plugin/kibana/app/kibana#/visualize/create?indexPattern=b242c6f0-...
시크릿 모드(2개)
Visualize / Create
admin
Save
Share
Inspect
Search
KQL
Oct 1, 2016 @ 21:19:57.679 → Dec 1, 2016 @ 21:20:06.855
Refresh
+ Add filter
Save visualization
Title
my_bar_chart
Description
Add to dashboard after saving
Cancel
Save and return
Average rating
Beyonce
JackBerk
TomJohnson
JenifferAnderson
TimCook
Professor.keyword: Descending
classes*
Data
Metrics & axes
Panel settings
Buckets
X-axis
Aggregation
Terms help
Terms
Field
Professor.keyword
Order by
Metric: Average rating
Order
Descending
Size
5
Group other values in separate bucket
Show missing values
Custom label
Discard
Update

짜잔! 같은 방식으로 다양한 차트들을 한눈에 볼 수 있습니다!

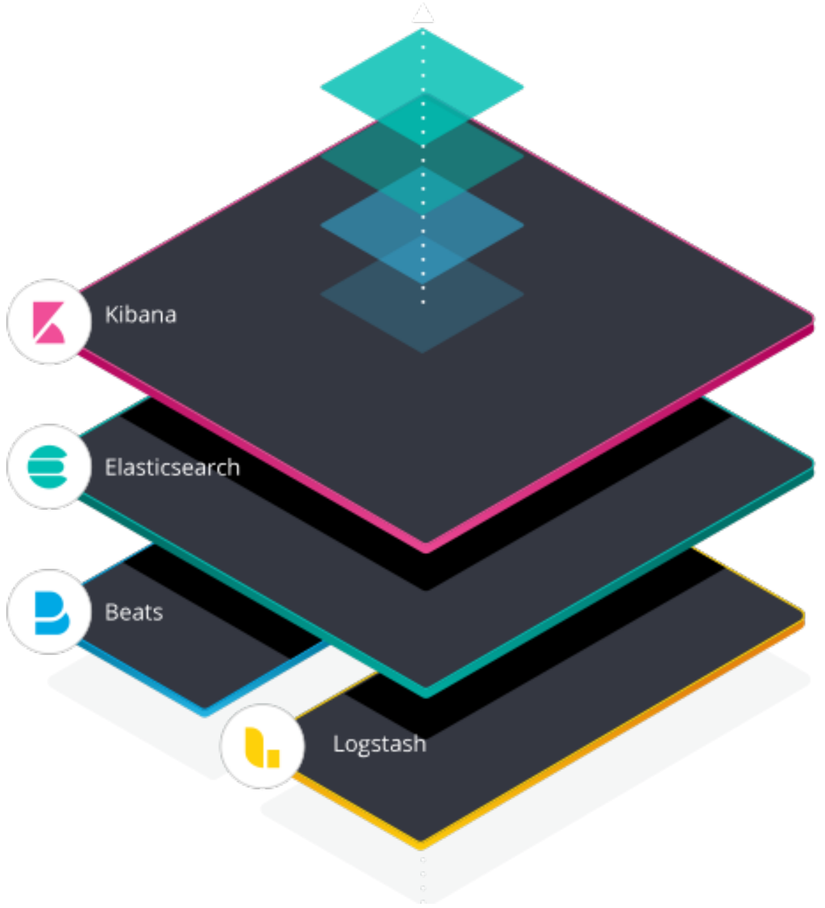


LOGSTASH

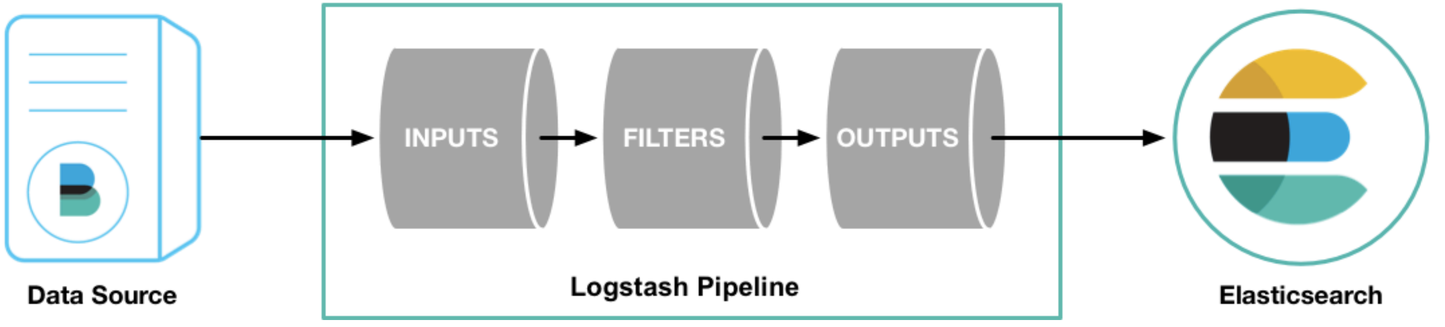
ELK Stack

Elasticsearch, Logstash, Kibana

- Elasticsearch
  - 검색 및 분석 엔진
- Logstash
  - 데이터 처리 파이프라인
  - 여러 소스에서 동시에 데이터를 수집 및 변환
- Kibana
  - 차트와 그래프를 이용해 데이터 시각화
- Beat
  - 파일 추적
  - Beat가 추가됨으로써 ELK Stack에서 Elastic Stack으로!



https://www.elastic.co/kr/what-is/elk-stack



https://www.elastic.co/guide/en/logstash/current/first-event.html

Logstash

Things to do

- Educate 계정을 이용해서 EC2 t2.medium으로 인스턴스 하나 생성
- Security Group에 8080 포트 열어두기
- 강의자료에서 제공된 명령어 입력
- stdinput and stdoutput test
- http input test
- 강의자료에서 제공된 logstash.conf파일을 이용하여, ELK 스택 완성!

E.O.D