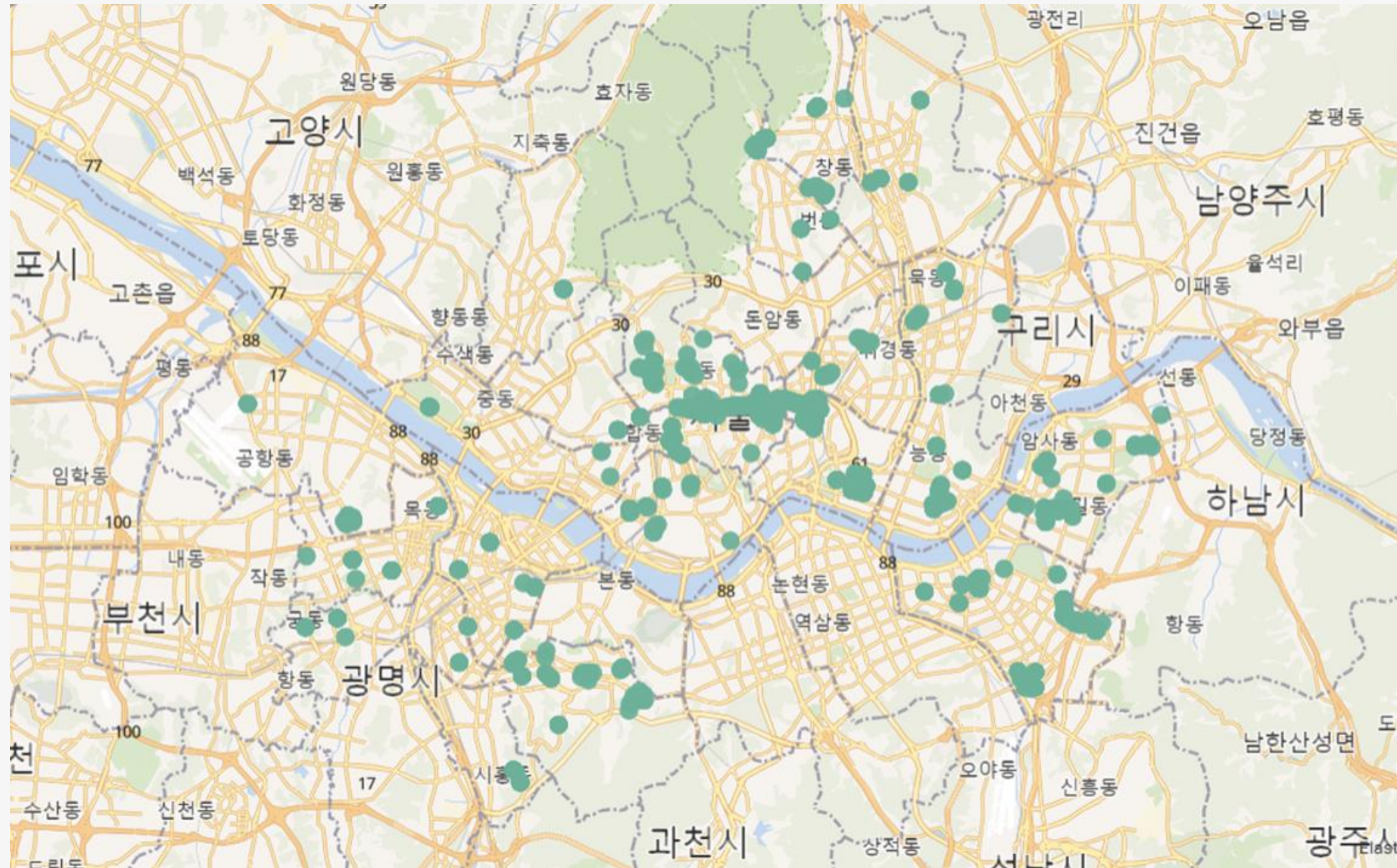


# IR Term project

201812167 이진형

# API를 이용한 서울시 공공와이파이 지도



데이터광장

# 주제 선택 이유 :

엘라스틱서치의 Maps 기능이 있는 것을 저번 과제에 보았고  
이를 사용해보고 싶었다.

최근 API 공부를 하면서 API와 연계해서 할만한 게 없을까 해서 살펴보다가  
WIFI 지도가 있는 것을 확인  
->이를 만들어보고자 했다.

# API 확보

REST API를 이용해서  
위도 경도 및 주소와 구까지 추출

# 설계 순서

## API 호출 및 전송

- > Elastic Search 인덱스 사전 매핑
- > 데이터를 엘라스틱으로 전송
- > 지도에 표기

# Python 클라이언트

API를 이용해서 Elastic Search에 데이터를 삽입해야하는 핵심 Part

```
import urllib.request
from xml.etree.ElementTree import fromstring, ElementTree
from elasticsearch import Elasticsearch, helpers

es = Elasticsearch()

docs = []
```



# Python 클라이언트 2

생략 코드는 REST API를 부르는  
API url/인증키 구간

```
response = urllib.request.urlopen(url)
xml_str = response.read().decode("utf-8")

tree = ElementTree.fromstring(xml_str)
root = tree.getroot()

✓ for row in root.iter("row"):
    gu_nm = row.find("X_SWIFI_WRDOFC").text
    place_nm = row.find("X_SWIFI_MAIN_NM").text
    place_x = float(row.find("LAT").text)
    place_y = float(row.find("LNT").text)
    ✓ doc = {
        "gu_nm": gu_nm,
        "place_nm": place_nm,
        "instl_xy": {"lat": place_y, "lon": place_x},
    }
    res = es.index(index="seoul_wifi", doc_type="_doc", body=doc)
    print(res)
    # print("END", iStart, "~", iEnd)
    print("END")

# print(json.dumps(data, indent=4, ensure_ascii=False))
```

# 데이터 사전 매핑

GEO\_POINT는 까다로운 타입

엘라스틱서치의 공식 레퍼런스에서도 Python에 대해서 GEO\_POINT 코드 예시가 없다

해결을 위해서 구글링 하던 중 데이터 삽입 이전에 미리 매핑을 끝내고 삽입을 하면 해결



```

8 PUT seoul_wifi
9 {
10   "settings":{
11     "analysis":{
12       "analyzer":{
13         "korean":{
14           "tokenizer":"nori_tokenizer"
15         }
16       }
17     }
18   },
19   "mappings":{
20     "properties": {
21       "gu_nm":{"type":"keyword"},
22       "place_nm":{"type":"text","analyzer":"korean"},
23       "instl_xy":{"type":"geo_point"}
24     }
25   }
26 }

```

```

PUT seoul_wifi
{
  "settings":{
    "analysis":{
      "analyzer":{
        "korean":{
          "tokenizer":"nori_tokenizer"
        }
      }
    },
    "mappings":{
      "properties": {
        "gu_nm":{"type":"keyword"},
        "place_nm":{"type":"text","analyzer":"korean"},
        "instl_xy":{"type":"geo_point"}
      }
    }
  }
}

```

# Data Type 문제

지도 내에 Data를 삽입하기 위해서는 Elasticsearch에서 인식할 수 있는 Geotype이 필요 -> 굉장한 시간 소요

Mapping 실수로 float type이 들어가고 지도에서 인식할 수 없는 상황이 발생

# 클라이언트로 데이터 전송

Logstash??? Beats?? 다른 방법은 없을까??

Logstash 원인모를 이유로 작동불가

잘 모르는 Beats로 만들기에는 시간 부족

구현 당시 Python 클라이언트에서 mapping이 잘 되지 않아서

GEO\_POINT 타입이 아닌 float 타입

```
-12-03 23:05:33.556 [main] Logstash - Logstash stopped processing because of an error: (ArgumentError)
ence in US-ASCII
```

# GEO\_POINT란?

```
PUT my-index-000001/_doc/1
{
  "text": "Geopoint as an object using GeoJSON format",
  "location": { ❶
    "type": "Point",
    "coordinates": [-71.34, 41.12]
  }
}

PUT my-index-000001/_doc/2
{
  "text": "Geopoint as a WKT POINT primitive",
  "location" : "POINT (-71.34 41.12)" ❷
}

PUT my-index-000001/_doc/3
{
  "text": "Geopoint as an object with 'lat' and 'lon' keys",
  "location": { ❸
    "lat": 41.12,
    "lon": -71.34
  }
}
```

# GEO\_POINT란?

```
PUT my-index-000001/_doc/4
{
  "text": "Geopoint as an array",
  "location": [ -71.34, 41.12 ] ④
}
```

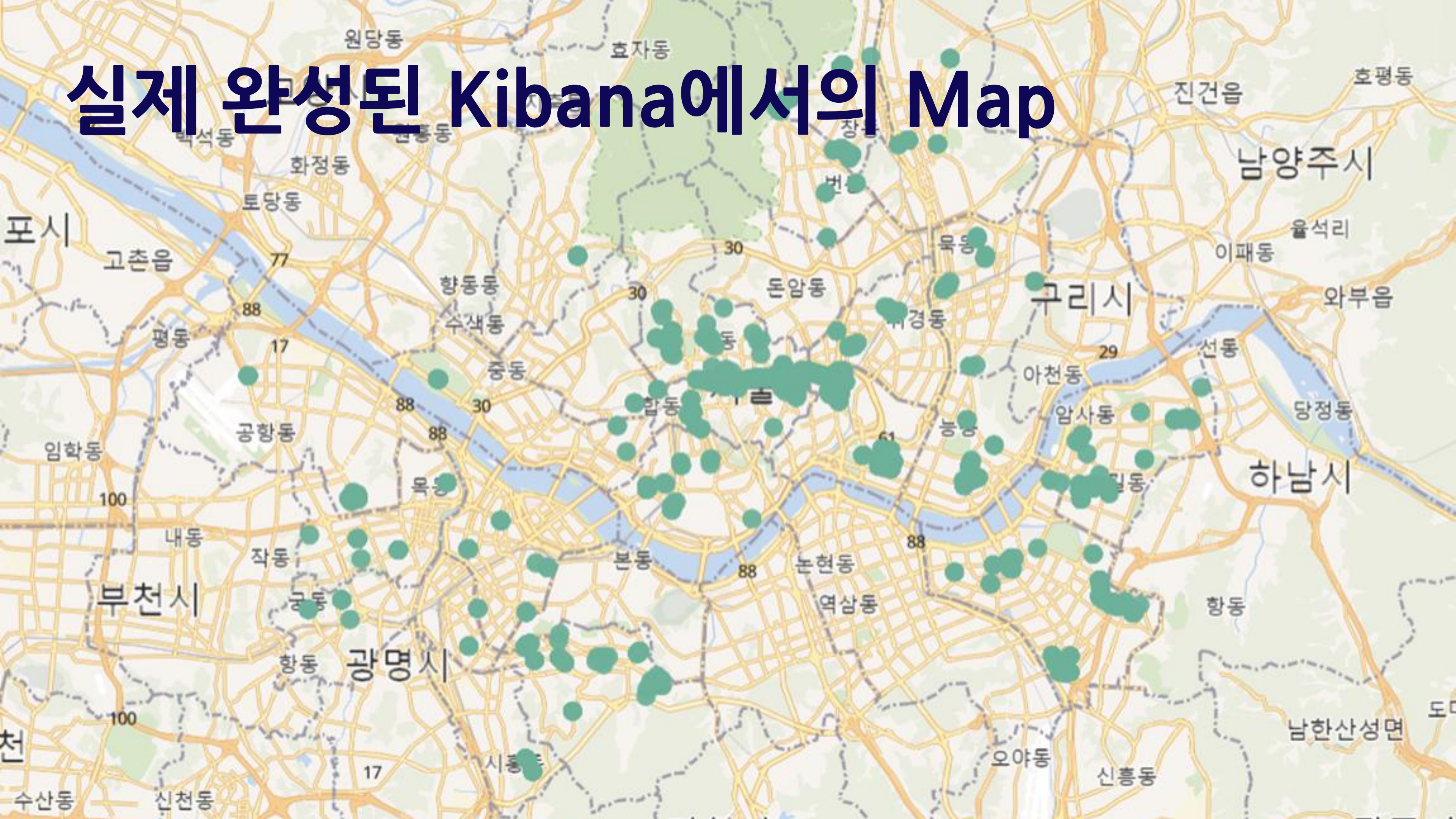
```
PUT my-index-000001/_doc/5
{
  "text": "Geopoint as a string",
  "location": "41.12,-71.34" ⑤
}
```

```
PUT my-index-000001/_doc/6
{
  "text": "Geopoint as a geohash",
  "location": "drm3btev3e86" ⑥
}
```

```
GET my-index-000001/_search
{
  "query": {
    "geo_bounding_box": { ⑦
      "location": {
        "top_left": {
          "lat": 42,
          "lon": -72
```



# 실제 완성된 Kibana에서의 Map





# 들인 시간 : 10일

토론하고 싶은 것을 간략하게 서술하세요.