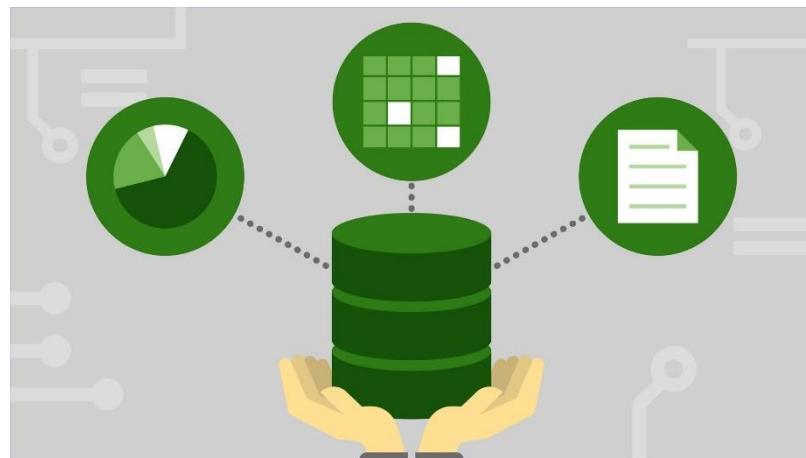


به نام خدا



دانشگاه تهران
پردیس دانشکده‌های فنی
دانشکده برق و کامپیوتر



آزمایشگاه پایگاه داده

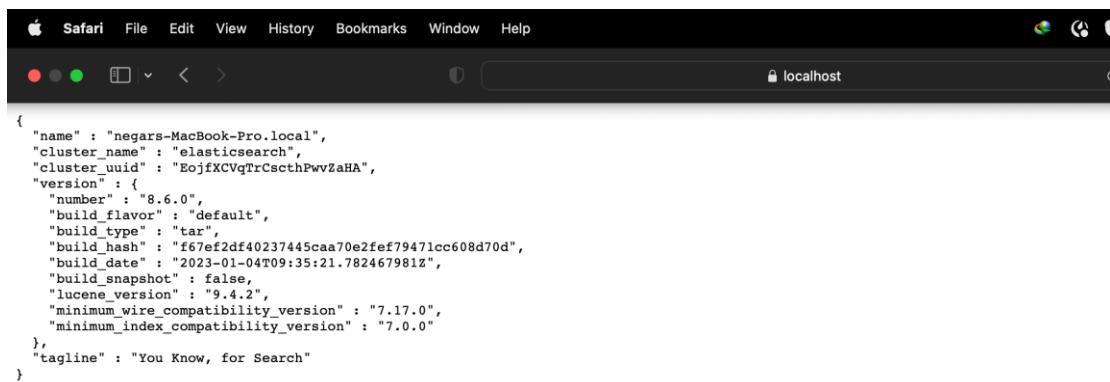
دستور کار شماره 5

نگار مرادی
810198543

1401 مهر

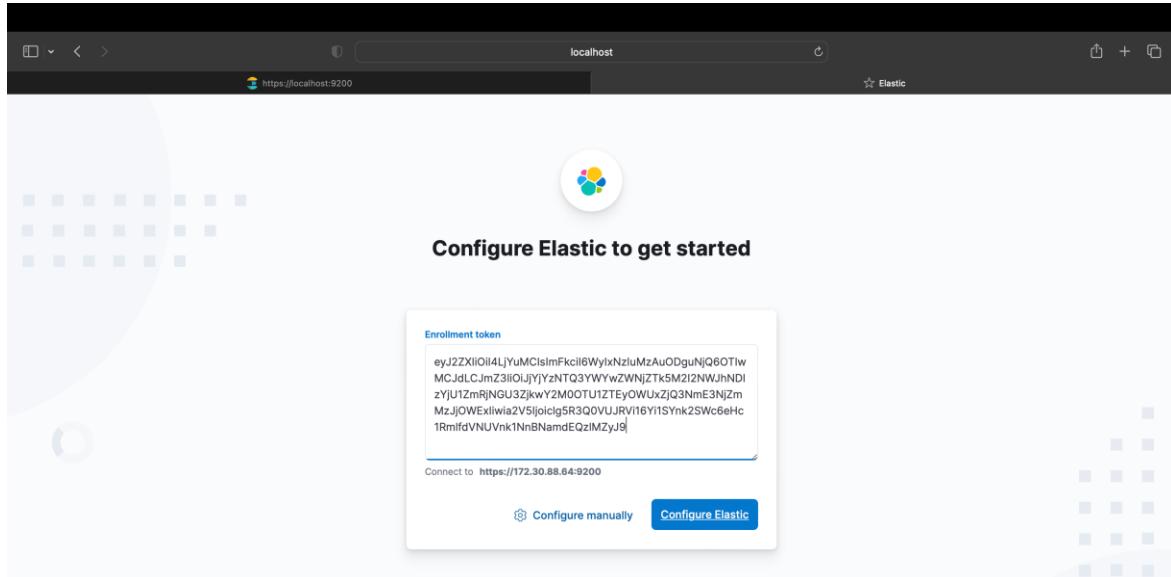
گزارش دستورکار انجام شده

گام اول: ابتدا elasticsearch را دانلود و نصب میکنیم. در شکل زیر آدرس localhost:9200 را مشاهده میکنید.

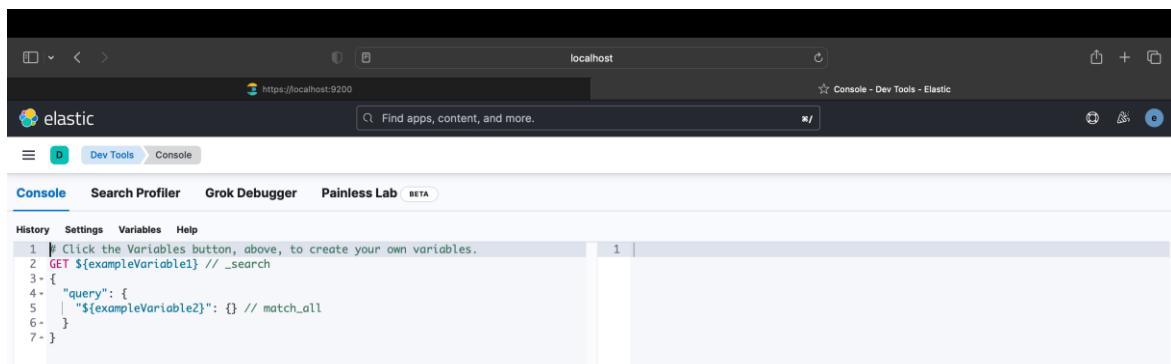


```
{
  "name" : "negars-MacBook-Pro.local",
  "cluster_name" : "elasticsearch",
  "cluster_uuid" : "EojfXCVqTrCscthPvwZaHA",
  "version" : {
    "number" : "8.6.0",
    "build_flavor" : "default",
    "build_type" : "tar",
    "build_hash" : "f67ef2df40237445caa70e2fef79471cc608d70d",
    "build_date" : "2023-01-04T09:35:21.782467981Z",
    "build_snapshot" : false,
    "lucene_version" : "9.4.2",
    "minimum_wire_compatibility_version" : "7.17.0",
    "minimum_index_compatibility_version" : "7.0.0"
  },
  "tagline" : "You Know, for Search"
}
```

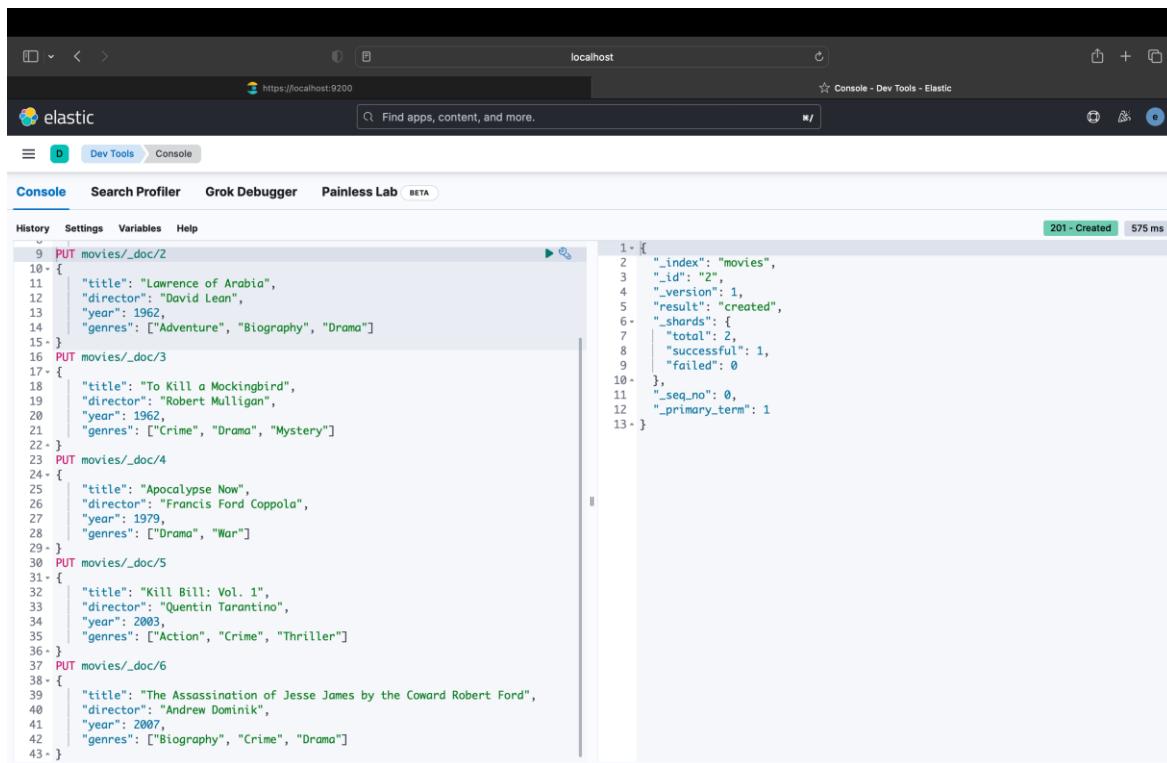
همین کار را برای کیانا هم انجام میدهیم. در شکل زیر مراحل نصب کیانا را مشاهده میکنید.



محیط کیانا



حال ابتدا داده ها را اضافه میکنیم.



```

9 PUT movies/_doc/2
10 {
11   "title": "Lawrence of Arabia",
12   "director": "David Lean",
13   "year": 1962,
14   "genres": ["Adventure", "Biography", "Drama"]
15 }
16 PUT movies/_doc/3
17 {
18   "title": "To Kill a Mockingbird",
19   "director": "Robert Mulligan",
20   "year": 1962,
21   "genres": ["Crime", "Drama", "Mystery"]
22 }
23 PUT movies/_doc/4
24 {
25   "title": "Apocalypse Now",
26   "director": "Francis Ford Coppola",
27   "year": 1979,
28   "genres": ["Drama", "War"]
29 }
30 PUT movies/_doc/5
31 {
32   "title": "Kill Bill: Vol. 1",
33   "director": "Quentin Tarantino",
34   "year": 2003,
35   "genres": ["Action", "Crime", "Thriller"]
36 }
37 PUT movies/_doc/6
38 {
39   "title": "The Assassination of Jesse James by the Coward Robert Ford",
40   "director": "Andrew Dominik",
41   "year": 2007,
42   "genres": ["Biography", "Crime", "Drama"]
43 }

```

201 - Created 575 ms

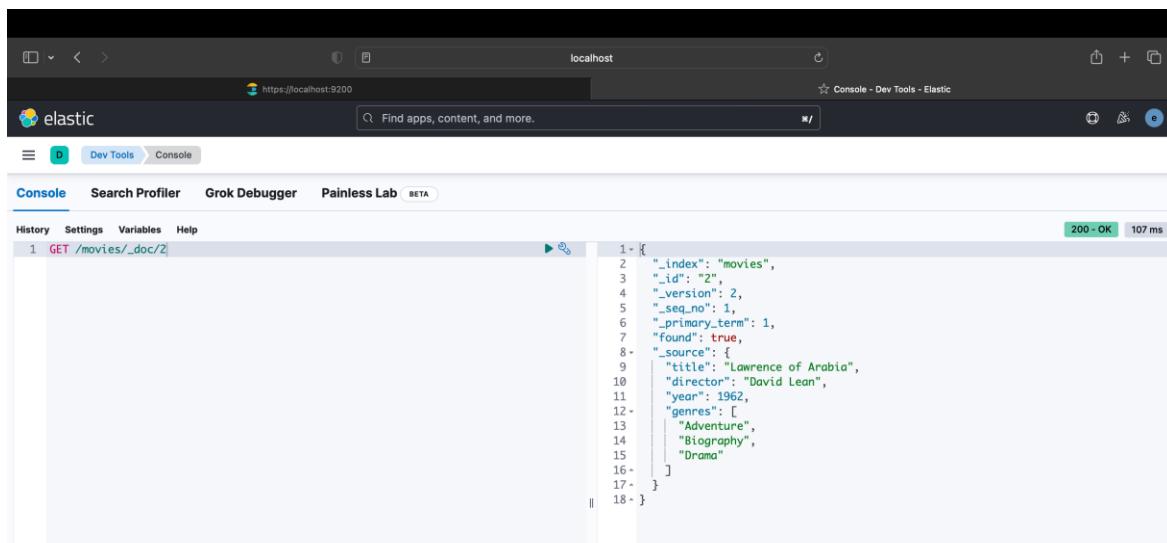
```

1- [
2   "_index": "movies",
3   "_id": "2",
4   "_version": 1,
5   "result": "created",
6   "_shards": {
7     "total": 2,
8     "successful": 1,
9     "failed": 0
10 },
11   "_seq_no": 0,
12   "_primary_term": 1
13 ]

```

دستورات:

اطلاعات فیلم با شناسه 2 اگر فیلم با شناسه مورد نظر وجود داشته باشد آن را نشان میدهد در غیر این صورت تصویر پایین تر را نشان میدهد.



```

1 GET /movies/_doc/2

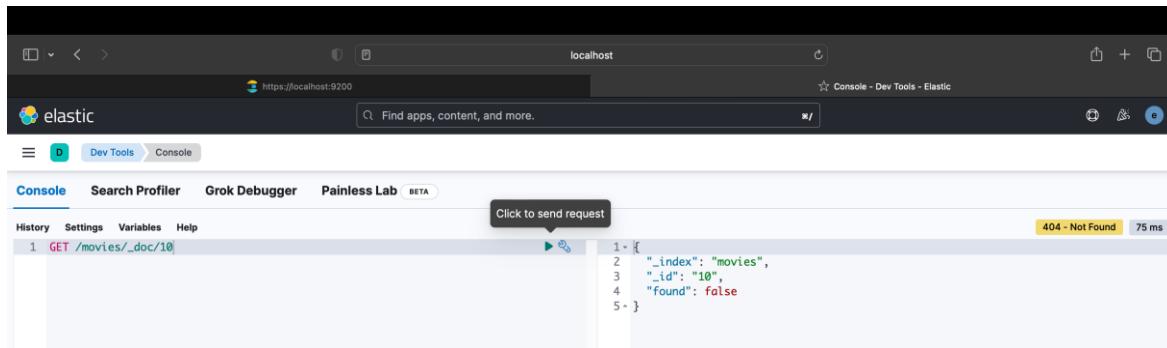
```

200 - OK 107 ms

```

1- [
2   "_index": "movies",
3   "_id": "2",
4   "_version": 2,
5   "_seq_no": 1,
6   "_primary_term": 1,
7   "found": true,
8   "_source": {
9     "title": "Lawrence of Arabia",
10    "director": "David Lean",
11    "year": 1962,
12    "genres": [
13      "Adventure",
14      "Biography",
15      "Drama"
16    ]
17  }
18 ]

```



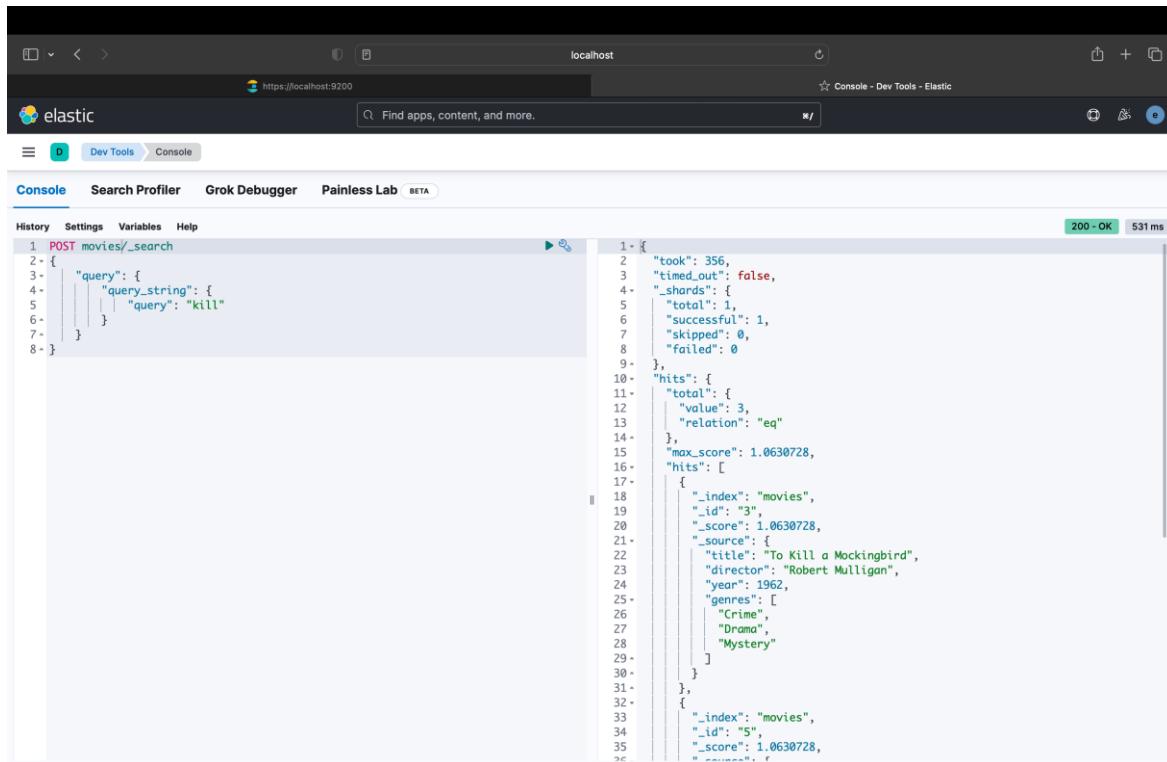
The screenshot shows the Elasticsearch Dev Tools Console at <https://localhost:9200>. A failed GET request is shown in the history:

```
1 GET /movies/_doc/10
```

The response status is 404 - Not Found with a duration of 75 ms. The response body is:

```
1+ [{  
2+   "_index": "movies",  
3+   "_id": "10",  
4+   "found": false  
5+ }]
```

هر فیلمی که در آن کلمه kill به کار رفته است.



The screenshot shows the Elasticsearch Dev Tools Console at <https://localhost:9200>. A successful POST search request is shown in the history:

```
1 POST movies/_search
```

The response status is 200 - OK with a duration of 531 ms. The response body is a detailed search result:

```
1+ [{  
2+   "took": 356,  
3+   "timed_out": false,  
4+   "_shards": {  
5+     "total": 1,  
6+     "successful": 1,  
7+     "skipped": 0,  
8+     "Failed": 0  
9+   },  
10+   "hits": {  
11+     "total": {  
12+       "value": 3,  
13+       "relation": "eq"  
14+     },  
15+     "max_score": 1.0630728,  
16+     "hits": [  
17+       {  
18+         "_index": "movies",  
19+         "_id": "3",  
20+         "_score": 1.0630728,  
21+         "_source": {  
22+           "title": "To Kill a Mockingbird",  
23+           "director": "Robert Mulligan",  
24+           "year": 1962,  
25+           "genres": [  
26+             "Crime",  
27+             "Drama",  
28+             "Mystery"  
29+           ]  
30+         },  
31+       },  
32+       {  
33+         "_index": "movies",  
34+         "_id": "5",  
35+         "_score": 1.0630728,  
36+         "_source": {  
37+           "title": "The Godfather",  
38+           "director": "Francis Ford Coppola",  
39+           "year": 1972,  
40+           "genres": [  
41+             "Crime",  
42+             "Drama",  
43+             "Thriller"  
44+           ]  
45+         },  
46+       }  
47+     ]  
48+   }]  
49+ }
```

فیلم هایی که در عنوانشان کلمه ford به کار رفته است.

```

1 GET /movies/_search
2 {
3   "query": {
4     "query_string": {
5       "query": "ford",
6       "fields": ["title"]
7     }
8   }
9 }

```

Response:

```

1 [ {
2   "took": 5,
3   "timed_out": false,
4   "_shards": {
5     "total": 1,
6     "successful": 1,
7     "skipped": 0,
8     "failed": 0
9   },
10  "hits": {
11    "total": {
12      "value": 1,
13      "relation": "eq"
14    },
15    "max_score": 1.0035701,
16    "hits": [
17      {
18        "_index": "movies",
19        "_id": "6",
20        "_score": 1.0035701,
21        "_source": {
22          "title": "The Assassination of Jesse James by the Coward Robert Ford",
23          "director": "Andrew Dominik",
24          "year": 2007,
25          "genres": [
26            "Biography",
27            "Crime",
28            "Drama"
29          ]
30        }
31      }
32    ]
33  }
34 }

```

تمام فیلم هایی که در نام آن ها کلمه `ford` و یا `kill` به کار رفته است.

```

1 GET /movies/_search
2 {
3   "query": {
4     "query_string": {
5       "query": "title:ford OR kill"
6     }
7   }
8 }

```

Response:

```

1 [ {
2   "genres": [
3     "Crime",
4     "Drama",
5     "Mystery"
6   ],
7   {
8     "_index": "movies",
9     "_id": "5",
10    "_score": 1.0630728,
11    "_source": {
12      "title": "Kill Bill: Vol. 1",
13      "director": "Quentin Tarantino",
14      "year": 2003,
15      "genres": [
16        "Action",
17        "Crime",
18        "Thriller"
19      ]
20    }
21  },
22  {
23    "_index": "movies",
24    "_id": "6",
25    "_score": 1.0035701,
26    "_source": {
27      "title": "The Assassination of Jesse James by the Coward Robert Ford",
28      "director": "Andrew Dominik",
29      "year": 2007,
30      "genres": [
31        "Biography",
32        "Crime",
33        "Drama"
34      ]
35    }
36  }
37 }

```

تمام فیلم هایی که عنوان kill داشته و کارگردان آنها Tarantino است.

```

1 GET /movies/_search
2 {
3   "query": {
4     "query_string": { "query": "(title:kill) AND (director:Tarantino)" }
5   }
6 }

1 [
2   "took": 26,
3   "timed_out": false,
4   "_shards": {
5     "total": 1,
6     "successful": 1,
7     "skipped": 0,
8     "failed": 0
9   },
10  "hits": {
11    "total": {
12      "value": 1,
13      "relation": "eq"
14    },
15    "max_score": 2.3646588,
16    "hits": [
17      {
18        "_index": "movies",
19        "_id": "5",
20        "_score": 2.3646588,
21        "_source": {
22          "title": "Kill Bill: Vol. 1",
23          "director": "Quentin Tarantino",
24          "year": 2003,
25          "genres": [
26            "Action",
27            "Crime",
28            "Thriller"
29          ]
30        }
31      }
32    ]
33  }
34 ]

```

تمام فیلم های زانر درام در سال ۱۹۶۲

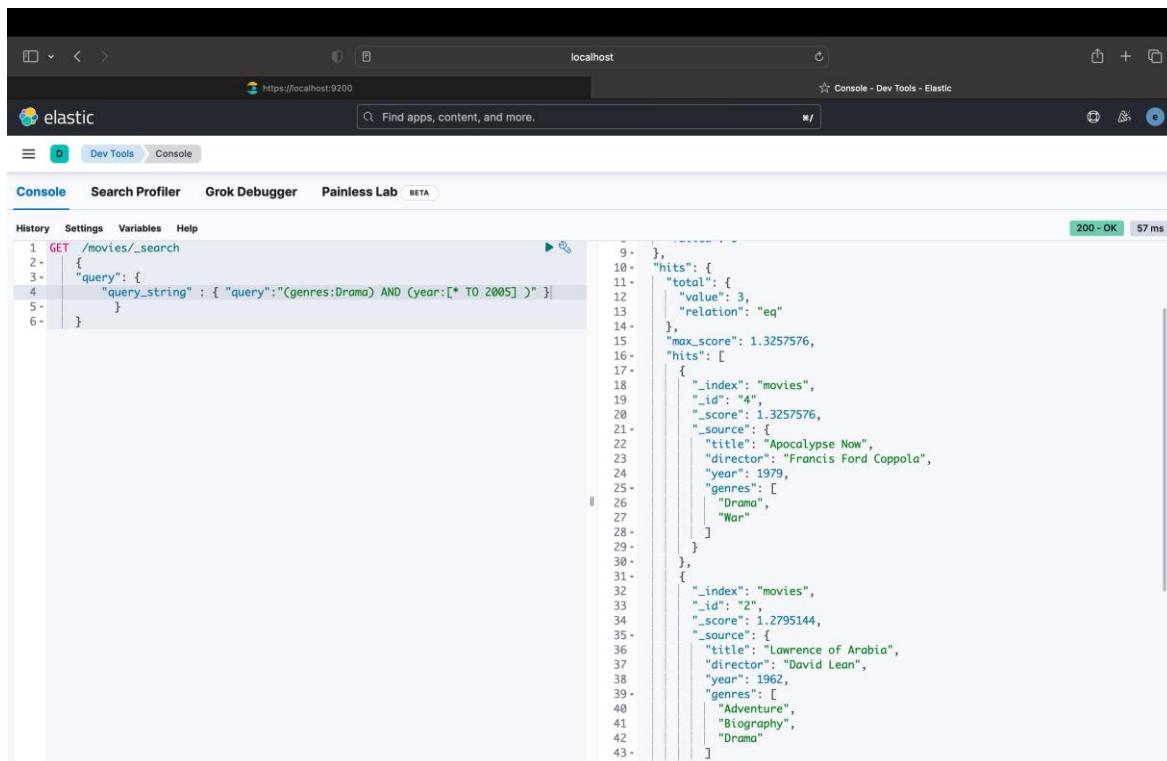
```

1 GET /movies/_search
2 {
3   "query": {
4     "query_string": { "query": "(genres:Drama) AND (year:1962 )" }
5   }
6 }

9 ],
10  "hits": {
11    "total": {
12      "value": 2,
13      "relation": "eq"
14    },
15    "max_score": 1.2795144,
16    "hits": [
17      {
18        "_index": "movies",
19        "_id": "2",
20        "_score": 1.2795144,
21        "_source": {
22          "title": "Lawrence of Arabia",
23          "director": "David Lean",
24          "year": 1962,
25          "genres": [
26            "Adventure",
27            "Biography",
28            "Drama"
29          ]
30        }
31      },
32      {
33        "_index": "movies",
34        "_id": "3",
35        "_score": 1.2795144,
36        "_source": {
37          "title": "To Kill a Mockingbird",
38          "director": "Robert Mulligan",
39          "year": 1962,
40          "genres": [
41            "Crime",
42            "Drama",
43            "Mystery"
44        }
45      }
46    ]
47  }
48 ]

```

تمام فیلم های ژانر درام تا سال ۲۰۰۵

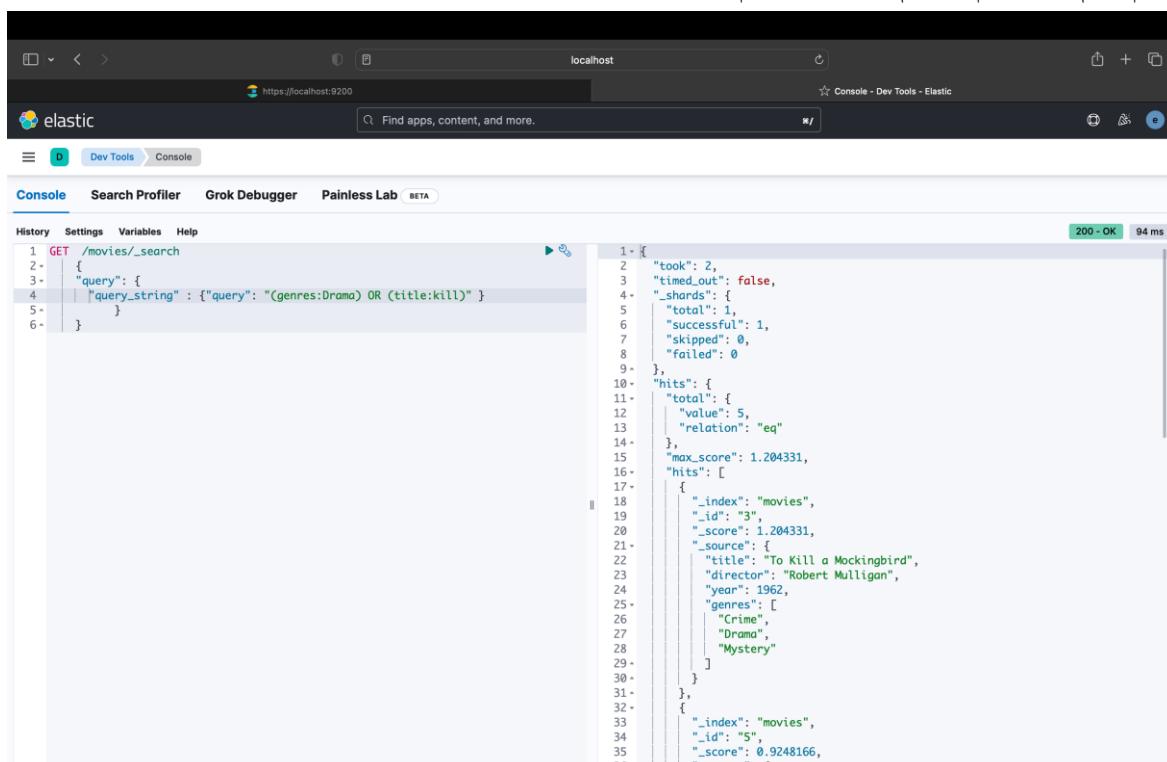


```

1 GET /movies/_search
2 {
3   "query": {
4     "query_string": { "query": "(genres:Drama) AND (year:[* TO 2005])" }
5   }
6 }

9 },
10 "hits": {
11   "total": {
12     "value": 3,
13     "relation": "eq"
14   },
15   "max_score": 1.3257576,
16   "hits": [
17     {
18       "_index": "movies",
19       "_id": "4",
20       "_score": 1.3257576,
21       "_source": {
22         "title": "Apocalypse Now",
23         "director": "Francis Ford Coppola",
24         "year": 1979,
25         "genres": [
26           "Drama",
27           "War"
28         ]
29       }
30     },
31     {
32       "_index": "movies",
33       "_id": "2",
34       "_score": 1.2795144,
35       "_source": {
36         "title": "Lawrence of Arabia",
37         "director": "David Lean",
38         "year": 1962,
39         "genres": [
40           "Adventure",
41           "Biography",
42           "Drama"
43         ]
44       }
45     }
46   ]
47 }
48
49 
```

تمام فیلم های درام یا فیلم هایی که در نام آنها کلمه kill به کار رفته باشد



```

1 GET /movies/_search
2 {
3   "query": {
4     "query_string": { "query": "(genres:Drama) OR (title:kill)" }
5   }
6 }

1 {
2   "took": 2,
3   "timed_out": false,
4   "_shards": {
5     "total": 1,
6     "successful": 1,
7     "skipped": 0,
8     "failed": 0
9   },
10  "hits": {
11    "total": {
12      "value": 5,
13      "relation": "eq"
14    },
15    "max_score": 1.204331,
16    "hits": [
17      {
18        "_index": "movies",
19        "_id": "3",
20        "_score": 1.204331,
21        "_source": {
22          "title": "To Kill a Mockingbird",
23          "director": "Robert Mulligan",
24          "year": 1962,
25          "genres": [
26            "Crime",
27            "Drama",
28            "Mystery"
29          ]
30        }
31      },
32      {
33        "_index": "movies",
34        "_id": "5",
35        "_score": 0.9248166,
36        "_source": {
37          "title": "The Killers",
38          "director": "Elia Kazan",
39          "year": 1946,
40          "genres": [
41            "Crime",
42            "Drama",
43            "Thriller"
44          ]
45        }
46      }
47    ]
48  }
49 }
50 
```

فیلم هایی که هم ژانر درام باشد و هم در عنوان آن kill به کار رفته باشد. با عملگر boost اولویت بیشتری به عنوان kill میدهیم.

```

1 GET /movies/_search
2 {
3   "query": {
4     "query_string": {"query": "(genres:Drama) OR (title:kill^2)"}
5   }
6 }

```

```

1 [
2   "took": 5,
3   "timed_out": false,
4   "_shards": {
5     "total": 1,
6     "successful": 1,
7     "skipped": 0,
8     "failed": 0
9   },
10  "hits": {
11    "total": {
12      "value": 5,
13      "relation": "eq"
14    },
15    "max_score": 2.1291475,
16    "hits": [
17      {
18        "_index": "movies",
19        "_id": "3",
20        "_score": 2.1291475,
21        "_source": {
22          "title": "To Kill a Mockingbird",
23          "director": "Robert Mulligan",
24          "year": 1962,
25          "genres": [
26            "Crime",
27            "Drama",
28            "Mystery"
29          ]
30        }
31      },
32      {
33        "_index": "movies",
34        "_id": "5",
35        "_score": 1.8496332,
36        "_source": {
37          "title": "The Godfather"
38        }
39      }
40    ]
41  }
42 }

```

فیلم های ژانر درام. با دستور زیر یک درخت جست و جو میسازیم و فیلم هایی که ژانر درام دارند را برミگردانیم.

```

1 GET /movies/_search
2 {
3   "query": {
4     "match": {
5       "genres": "drama"
6     }
7   }
8 }

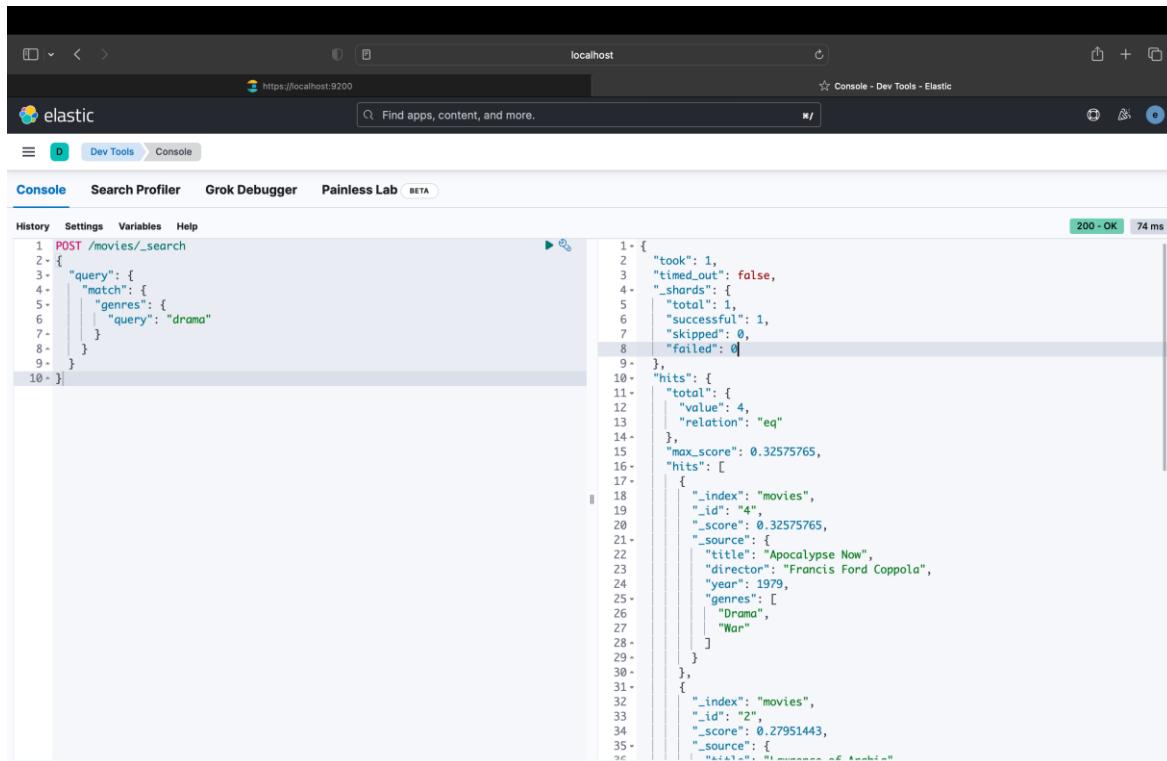
```

```

1 [
2   "took": 5,
3   "timed_out": false,
4   "_shards": {
5     "total": 1,
6     "successful": 1,
7     "skipped": 0,
8     "failed": 0
9   },
10  "hits": {
11    "total": {
12      "value": 4,
13      "relation": "eq"
14    },
15    "max_score": 0.32575765,
16    "hits": [
17      {
18        "_index": "movies",
19        "_id": "4",
20        "_score": 0.32575765,
21        "_source": {
22          "title": "Apocalypse Now",
23          "director": "Francis Ford Coppola",
24          "year": 1979,
25          "genres": [
26            "Drama",
27            "War"
28          ]
29        }
30      },
31      {
32        "_index": "movies",
33        "_id": "2",
34        "_score": 0.27951443,
35        "_source": {
36          "title": "The Godfather"
37        }
38      }
39    ]
40  }
41 }

```

در دستور زیر با ساخت یک درخت جست وجو تمامی فیلمها را بررسی کرده و فیلد genres آنها را با کلمه Drama مقایسه می کنیم.

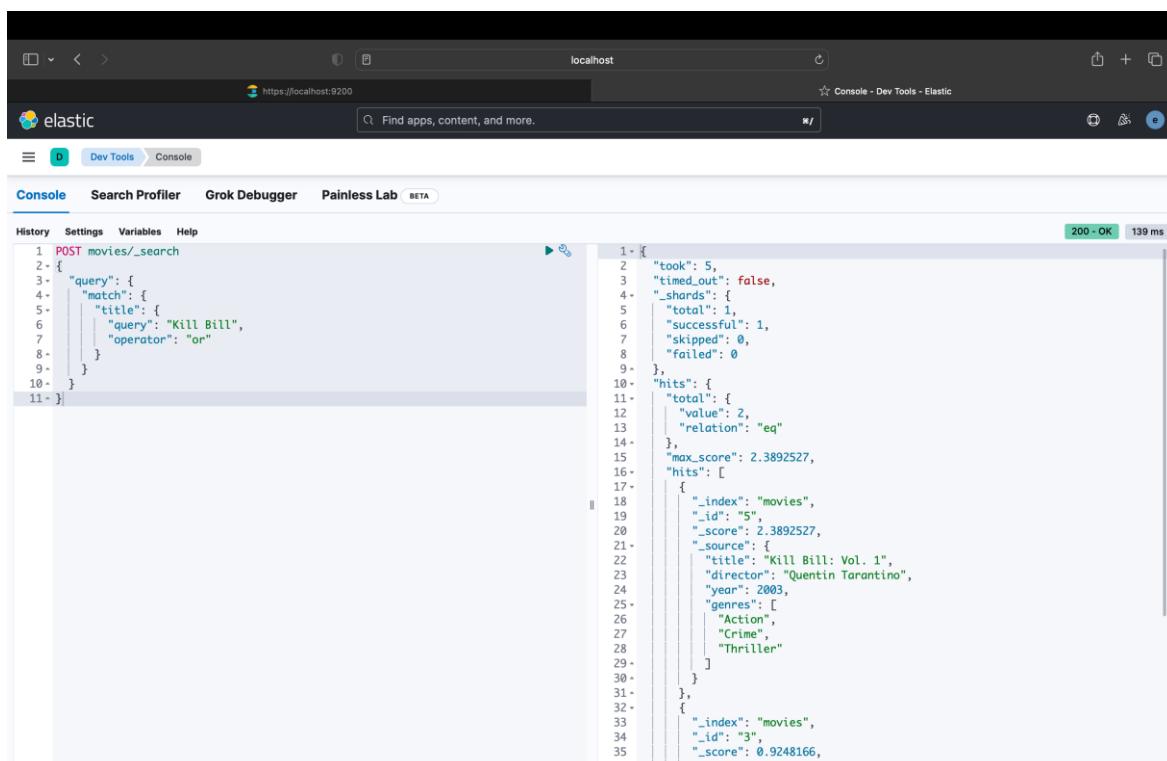


```
POST /movies/_search
{
  "query": {
    "match": {
      "genres": {
        "query": "drama"
      }
    }
  }
}
```

The response shows the search results for movies with the genre "drama". Two movies are found:

- Movie ID 1: Title - Apocalypse Now, Director - Francis Ford Coppola, Year - 1979, Genres - Drama, War
- Movie ID 2: Title - Kill Bill: Vol. 1, Director - Quentin Tarantino, Year - 2003, Genres - Action, Crime, Thriller

با استفاده از عملگر or در کوئری زیر فیلمهایی بازگردانده میشود که در نام آنها کلمه Bill یا Kill به کار رفته باشد.

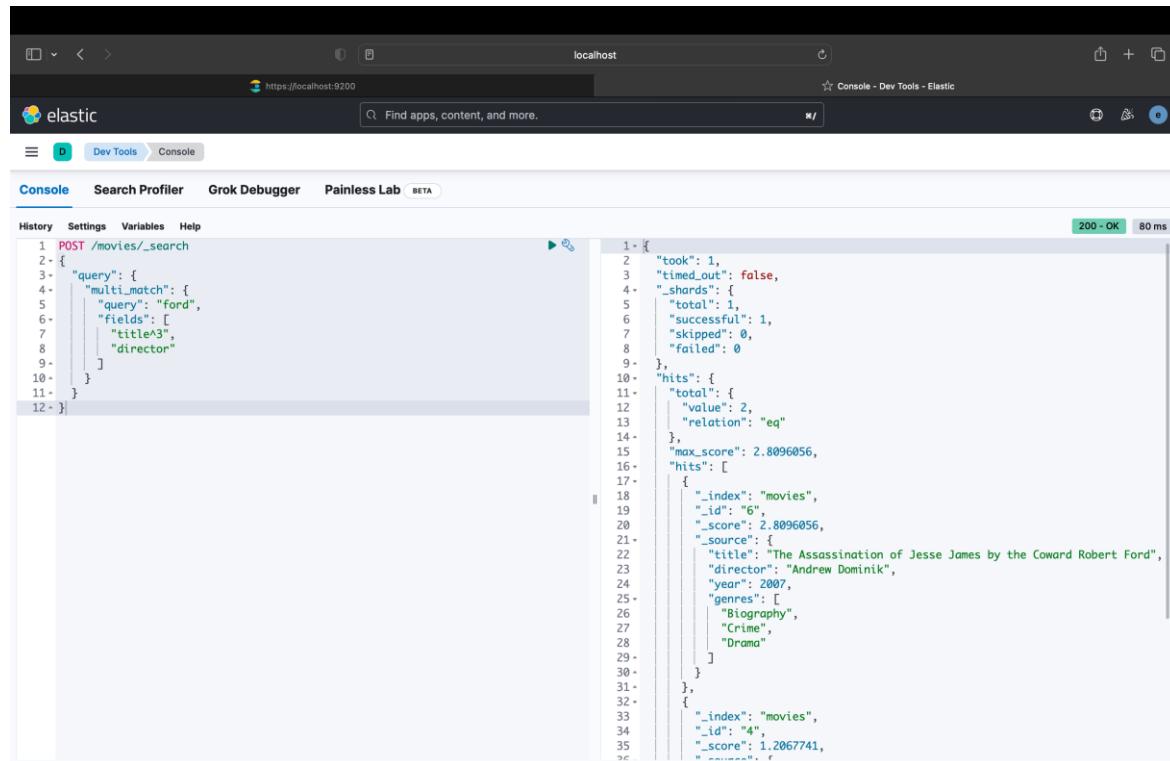


```
POST movies/_search
{
  "query": {
    "match": {
      "title": {
        "query": "Kill Bill",
        "operator": "or"
      }
    }
  }
}
```

The response shows the search results for movies with titles containing either "Kill Bill" or "Bill". Three movies are found:

- Movie ID 1: Title - Kill Bill: Vol. 1, Director - Quentin Tarantino, Year - 2003, Genres - Action, Crime, Thriller
- Movie ID 2: Title - Kill Bill, Director - Unknown, Year - 2003, Genres - Action, Crime, Thriller
- Movie ID 3: Title - Bill, Director - Unknown, Year - 2003, Genres - Action, Crime, Thriller

اگر بخواهیم همزمان بر روی دو فیلد یا بیشتر جستجو انجام دهیم، می توانیم از عبارت `multi_match` استفاده کنیم. در کوئری زیر با استفاده از دستور `^` اولویت عنوان فیلم را سه برابر نام کارگردان، تعیین کرده ایم.

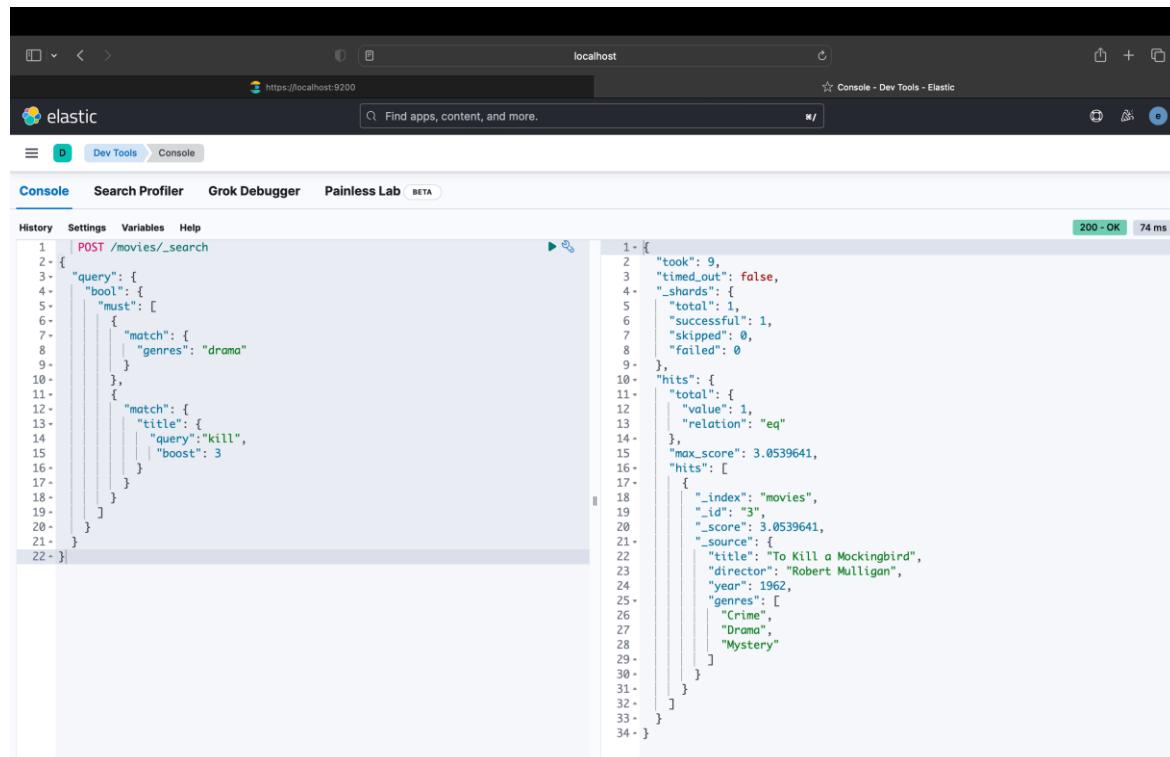


```
POST /movies/_search
{
  "query": {
    "multi_match": {
      "query": "ford",
      "fields": [
        "title^3",
        "director"
      ]
    }
  }
}
```

The response shows two hits:

- Movie: The Assassination of Jesse James by the Coward Robert Ford, Director: Andrew Dominik, Year: 2007, Genres: Biography, Crime, Drama. Score: 2.8096056.
- Movie: To Kill a Mockingbird, Director: Robert Mulligan, Year: 1962, Genres: Crime, Drama, Mystery. Score: 1.2067741.

در دستور زیر، تمامی رکوردهای جدول فیلم بررسی می شود و دنبال فیلم های درام با عنوانی شامل `kill` می گردد. می توانیم عبارت `must` را به وسیله `and`, `or` پیاده سازی کنیم.

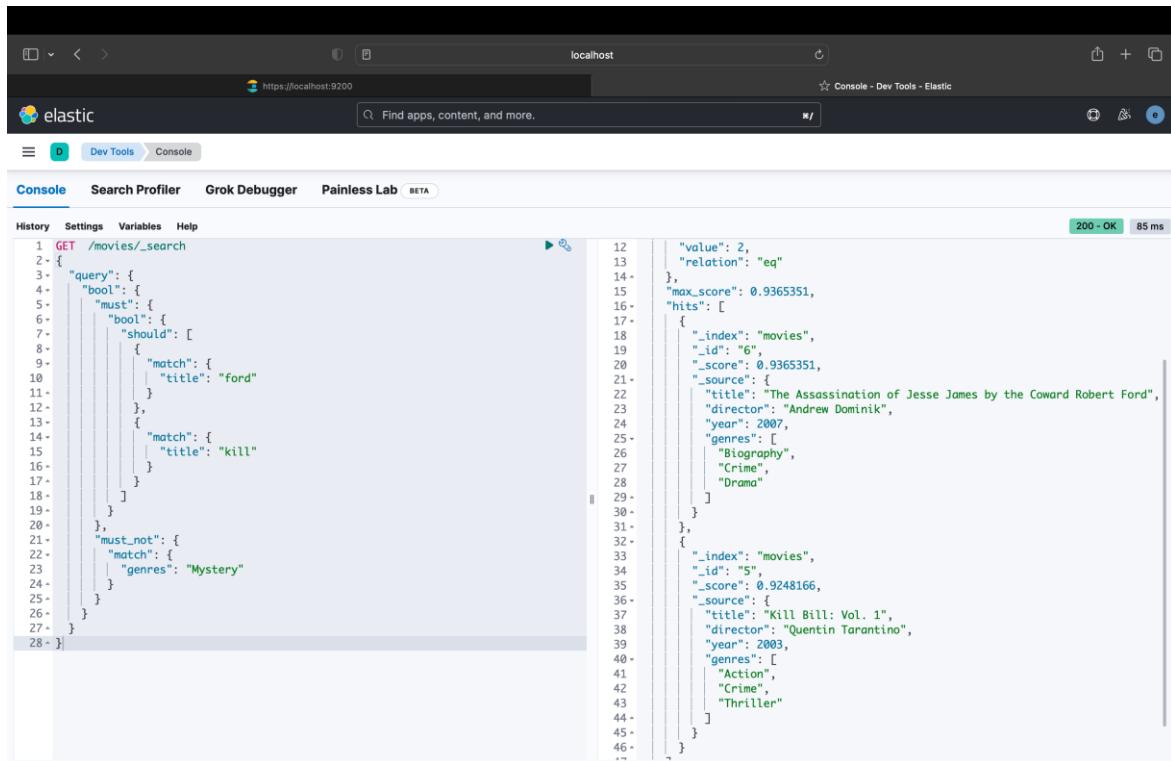


```
POST /movies/_search
{
  "query": {
    "bool": {
      "must": [
        {
          "match": {
            "genres": "drama"
          }
        },
        {
          "match": {
            "title": {
              "query": "kill",
              "boost": 3
            }
          }
        }
      ]
    }
  }
}
```

The response shows one hit:

- Movie: To Kill a Mockingbird, Director: Robert Mulligan, Year: 1962, Genres: Crime, Drama, Mystery. Score: 3.0539641.

فیلم هایی که در عنوان آنها کلمه Ford یا Kill به کار رفته است و زانر آنها اسرارآمیز نیست.



```

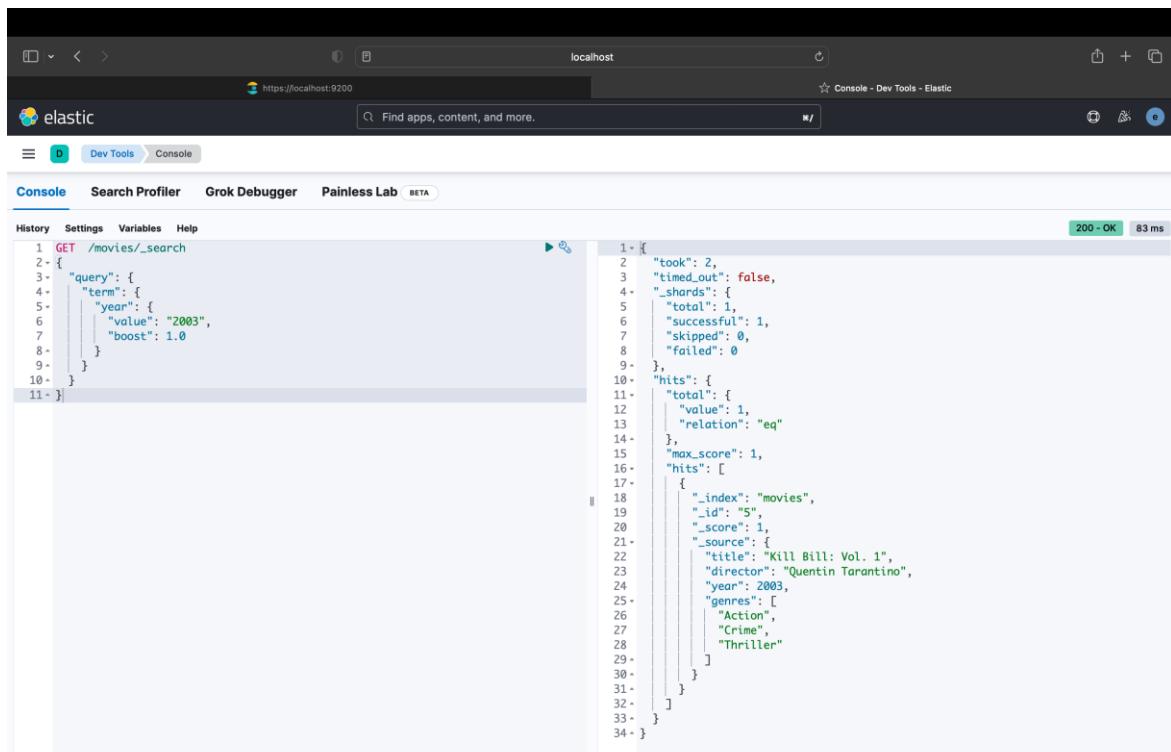
1 GET /movies/_search
2 {
3   "query": {
4     "bool": {
5       "must": [
6         {
7           "bool": {
8             "should": [
9               {
10                 "match": {
11                   "title": "ford"
12                 }
13               },
14               {
15                 "match": {
16                   "title": "kill"
17                 }
18               }
19             ],
20           }
21         },
22         "must_not": {
23           "match": {
24             "genres": "Mystery"
25           }
26         }
27     }
28   }
29 }
```

The search results show two movies: "The Assassination of Jesse James by the Coward Robert Ford" and "Kill Bill: Vol. 1". Both have a score of approximately 0.9365351.

```

12   "value": 2,
13   "relation": "eq"
14 },
15   "max_score": 0.9365351,
16   "hits": [
17     {
18       "_index": "movies",
19       "_id": "6",
20       "score": 0.9365351,
21       "source": {
22         "title": "The Assassination of Jesse James by the Coward Robert Ford",
23         "director": "Andrew Dominik",
24         "year": 2007,
25         "genres": [
26           "Biography",
27           "Crime",
28           "Drama"
29         ]
30       }
31     },
32     {
33       "_index": "movies",
34       "_id": "5",
35       "score": 0.9248166,
36       "source": {
37         "title": "Kill Bill: Vol. 1",
38         "director": "Quentin Tarantino",
39         "year": 2003,
40         "genres": [
41           "Action",
42           "Crime",
43           "Thriller"
44         ]
45       }
46     }
47   ]
48 }
```

برای جستجوی فیلدهای غیر متنی از نوع کوئری Term استفاده میکنیم. در کوئری زیر فیلم هایی که سال ساخت آن ها 2003 را میابیم.



```

1 GET /movies/_search
2 {
3   "query": {
4     "term": {
5       "year": {
6         "value": "2003",
7         "boost": 1.0
8       }
9     }
10   }
11 }
```

The search results show one movie: "Kill Bill: Vol. 1". It has a score of 1.0.

```

1 {
2   "took": 2,
3   "timed_out": false,
4   "_shards": {
5     "total": 1,
6     "successful": 1,
7     "skipped": 0,
8     "failed": 0
9   },
10   "hits": [
11     {
12       "total": {
13         "value": 1,
14         "relation": "eq"
15       },
16       "max_score": 1,
17       "hits": [
18         {
19           "_index": "movies",
20           "_id": "5",
21           "score": 1,
22           "source": {
23             "title": "Kill Bill: Vol. 1",
24             "director": "Quentin Tarantino",
25             "year": 2003,
26             "genres": [
27               "Action",
28               "Crime",
29               "Thriller"
30             ]
31           }
32         }
33       ]
34     }
35   ]
36 }
```

برای بالا بردن سرعت از فیلترینگ استفاده میکنیم. فیلترینگ، سرعت جستجو را بسیار بالا می برد چون تعداد سندهای لازم برای جستجو را کاهش می دهد. اعمال محدودیت روی رکوردهای مورد جستجو را فیلترینگ می نامیم و به صورت زیر آنرا اعمال می کنیم. ابتدا تمام فیلمهای تولید شده بعد از سال ۱۹۶۰ و از ژانر درام را فیلتر کرده سپس فیلم هایی را پیدا میکنیم که عنوان آنها شامل کلمه kill باشد.

```

1 GET /movies/_search
2 {
3   "query": {
4     "bool": {
5       "must": [
6         {
7           "match": {
8             "title": "kill"
9           }
10        }
11      ],
12      "filter": {
13        "bool": {
14          "must": [
15            {
16              "range": {
17                "year": {
18                  "gte": 1960
19                }
20              }
21            },
22            {
23              "term": {
24                "genres": {
25                  "value": "drama"
26                }
27              }
28            ]
29          }
30        }
31      }
32    }
33  }
34 }
```

1 [{
2 "took": 2,
3 "timed_out": false,
4 "_shards": {
5 "total": 1,
6 "successful": 1,
7 "skipped": 0,
8 "failed": 0
9 },
10 "hits": {
11 "total": {
12 "value": 1,
13 "relation": "eq"
14 },
15 "max_score": 0.9248166,
16 "hits": [
17 {
18 "_index": "movies",
19 "_id": "3",
20 "_score": 0.9248166,
21 "_source": {
22 "title": "To Kill a Mockingbird",
23 "director": "Robert Mulligan",
24 "year": 1962,
25 "genres": [
26 "Crime",
27 "Drama",
28 "Mystery"
29]
30 }
31 }
32]
33 }
34 }

می توان کوئری بالا را به صورت خلاصه زیر هم نوشت.

```

1 GET /movies/_search
2 {
3   "query": {
4     "bool": {
5       "must": [
6         {
7           "match": {
8             "title": "kill"
9           }
10        },
11        {
12          "range": {
13            "year": {
14              "gte": 1960
15            }
16          }
17        },
18        {
19          "term": {
20            "genres": {
21              "value": "drama"
22            }
23          }
24        ]
25      }
26    }
27  }
28 }
```

1 [{
2 "took": 6,
3 "timed_out": false,
4 "_shards": {
5 "total": 1,
6 "successful": 1,
7 "skipped": 0,
8 "failed": 0
9 },
10 "hits": {
11 "total": {
12 "value": 1,
13 "relation": "eq"
14 },
15 "max_score": 2.204331,
16 "hits": [
17 {
18 "_index": "movies",
19 "_id": "3",
20 "_score": 2.204331,
21 "_source": {
22 "title": "To Kill a Mockingbird",
23 "director": "Robert Mulligan",
24 "year": 1962,
25 "genres": [
26 "Crime",
27 "Drama",
28 "Mystery"
29]
30 }
31 }
32]
33 }
34 }

The screenshot shows the Elasticsearch Dev Tools Console interface. At the top, it says "localhost" and "https://localhost:9200". Below that is a search bar with "elastic" and a "Find apps, content, and more." placeholder. The main area has tabs for "Console", "Search Profiler", "Grok Debugger", and "Painless Lab (BETA)". The "Console" tab is selected. It displays a code editor with the following query:

```

1 GET /movies/_search
2 {
3   "query": {
4     "bool": {
5       "must": [
6         {
7           "range": {
8             "year": {
9               "gte": 2000
10            }
11          }
12        },
13        {
14          "term": {
15            "genres": {
16              "value": "drama"
17            }
18          }
19        }
20      ]
21    }
22  }
23

```

Below the code editor is a "Click to send request" button. To its right, the response is shown in a JSON format:

```

1 {
2   "took": 3,
3   "timed_out": false,
4   "_shards": {
5     "total": 1,
6     "successful": 1,
7     "skipped": 0,
8     "failed": 0
9   },
10  "hits": {
11    "total": {
12      "value": 1,
13      "relation": "eq"
14    },
15    "max_score": 1.2795144,
16    "hits": [
17      {
18        "_index": "movies",
19        "_id": "6",
20        "_score": 1.2795144,
21        "_source": {
22          "title": "The Assassination of Jesse James by the Coward Robert Ford",
23          "director": "Andrew Dominik",
24          "year": 2007,
25          "genres": [
26            "Biography",
27            "Crime",
28            "Drama"
29          ]
30        }
31      }
32    ]
33  }
34

```

The response includes status codes "200 - OK" and "93 ms".

می توانیم کلا بخش پرس و جو را حذف کنیم و جستجوی ما بر اساس مقادیر فیلدها مانند سال و رشته و ژانر و ... باشد.

This screenshot is identical to the one above, except the "source" field has been removed from the "hits" section of the search response. The "source" field was present in the previous screenshot but is now absent, demonstrating how to exclude specific fields from the returned results.

با استفاده از "fuzziness": "AUTO" حتی اگر غلط املایی داشتیم بتوانیم جست و جو کنیم. در کوئری زیر با اینکه godfather غلط املایی دارد آن را پیدا میکند.

```

1 GET /movies/_search
2 {
3   "query": {
4     "match": {
5       "title": {
6         "query": "godfater",
7         "fuzziness": "AUTO"
8       }
9     }
10  }
11 }

```

```

1 [ {
2   "took": 47,
3   "timed_out": false,
4   "_shards": {
5     "total": 1,
6     "successful": 1,
7     "skipped": 0,
8     "failed": 0
9   },
10  "hits": {
11    "total": {
12      "value": 0,
13      "relation": "eq"
14    },
15    "max_score": null,
16    "hits": []
17  }
18 }
19 ]

```

کوئری زیر به الاستیک اعلام می کند که به دنبال یک عبارت bill هستیم و باید دقیقاً کلمات ذکر شده پشت سرهم ظاهر شده باشند و نهایتاً سه فضای خالی بین هر کلمه تا بعدی را می توانیم قبول کنیم.

```

1 GET /movies/_search
2 {
3   "query": {
4     "match_phrase": {
5       "title": {
6         "query": "Kill Bill",
7         "slop": 2
8       }
9     }
10  }
11 }

```

```

1 [ {
2   "took": 18,
3   "timed_out": false,
4   "_shards": {
5     "total": 1,
6     "successful": 1,
7     "skipped": 0,
8     "failed": 0
9   },
10  "hits": {
11    "total": {
12      "value": 1,
13      "relation": "eq"
14    },
15    "max_score": 2.3892527,
16    "hits": [
17      {
18        "_index": "movies",
19        "_id": "5",
20        "_score": 2.3892527,
21        "_source": {
22          "title": "Kill Bill: Vol. 1",
23          "director": "Quentin Tarantino",
24          "year": 2003,
25          "genres": [
26            "Action",
27            "Crime",
28            "Thriller"
29          ]
30        }
31      }
32    ]
33  }
34 }

```

در دستور زیر به دنبال کارگردانی با نام **Francis Ford Coppola** هستیم. برای این کل متن را ایدنکس کنیم و نه فقط کلمه keyword از **Francis Ford Coppola** استفاده میکنیم.

```

PUT /movies/_mapping
{
  "movie": {
    "properties": {
      "director": {
        "type": "text",
        "fields": {
          "original": {
            "type": "keyword"
          }
        }
      }
    }
  }
}

```

The response shows an error message: "Root mapping definition has unsupported parameters: [movie : {properties={director={type=text, fields={original={type=keyword}}}}}]".

اگر بخواهیم نام کامل کارگردان را جستجو کنیم، به صورت زیر عمل میکنیم.

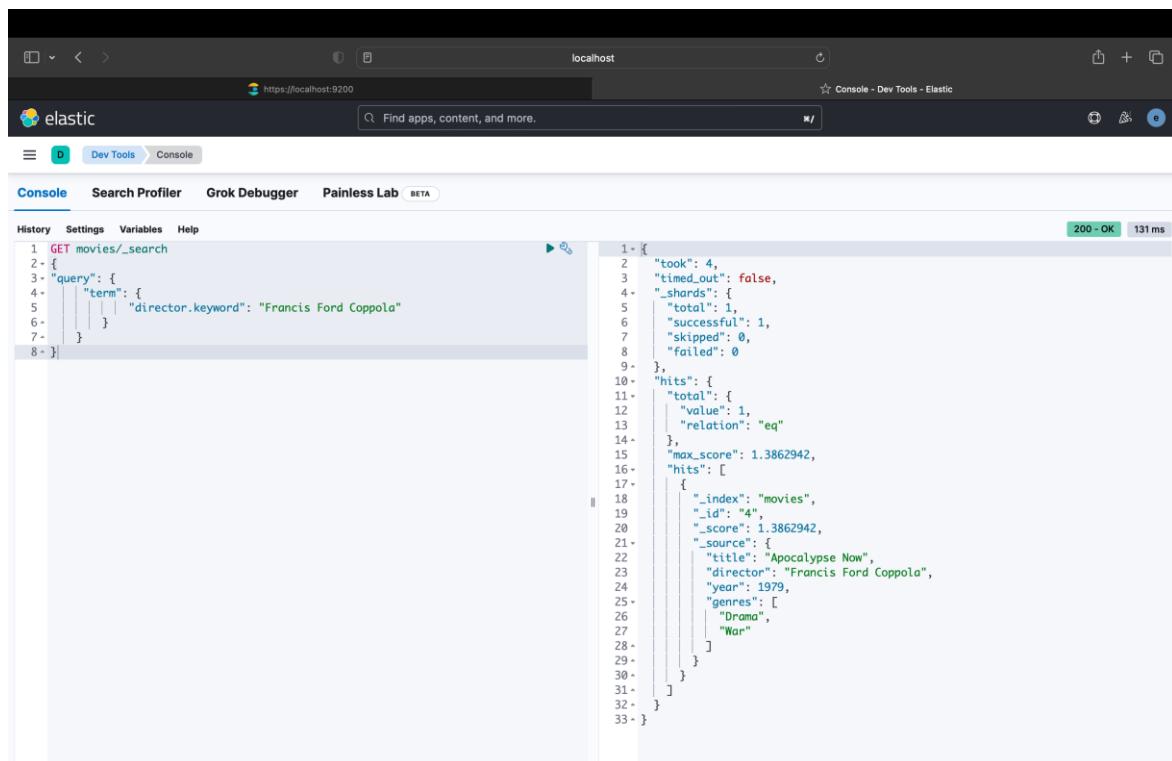
```

GET /movies/_search
{
  "query": {
    "term": {
      "director.original": "Francis Ford Coppola"
    }
  }
}

```

The response shows the search results, indicating 2 hits found for the query.

راه ساده‌تر این کار استفاده از field.keyword بود که برای همه فیلدها در دسترس است و نیاز به استفاده از فیلدهای چندمقداره برای کارگردان هم نبود. یعنی در مثال فوق کافی بود به جای "director": "Francis Ford" را "director.keyword": "Francis Ford Coppola" و بدون تغییر در نوع داده، عبارت "Coppola" بنویسیم.



The screenshot shows the Elasticsearch Dev Tools Console interface. In the left panel, a search query is typed:

```
1 GET movies/_search
2 {
3   "query": {
4     "term": {
5       "director.keyword": "Francis Ford Coppola"
6     }
7   }
8 }
```

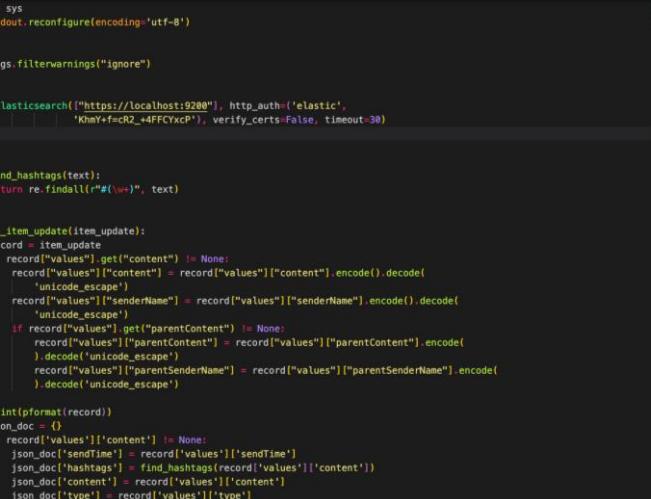
In the right panel, the response is displayed:

```
1 [
2   {
3     "took": 4,
4     "timed_out": false,
5     "_shards": {
6       "total": 1,
7       "successful": 1,
8       "skipped": 0,
9       "failed": 0
10    },
11    "hits": {
12      "total": {
13        "value": 1,
14        "relation": "eq"
15      },
16      "max_score": 1.3862942,
17      "hits": [
18        {
19          "_index": "movies",
20          "_id": "4",
21          "_score": 1.3862942,
22          "_source": {
23            "title": "Apocalypse Now",
24            "director": "Francis Ford Coppola",
25            "year": 1979,
26            "genres": [
27              "Drama",
28              "War"
29            ]
30          }
31        }
32      ]
33    }
34  ]
35 }
```

The response indicates a single hit for the movie "Apocalypse Now" directed by Francis Ford Coppola, with a score of 1.3862942. The source object contains the movie's title, director, year, and genres (Drama, War).

گام دوم:

تغییرات تابع update item on در شکل زیر آورده شده است.



```
pipe_line_connector.py -- sahamyab:4
pipe_line_connector.py > ...
pipe_line_connector.py > sahamyab.py

7 import sys
8 sys.stdout.reconfigure(encoding='utf-8')
9
10
11 warnings.filterwarnings("ignore")
12
13
14 es = Elasticsearch(["https://localhost:9200"], http_auth=('elastic',
15                   'KhmY+f=rcR2+4FFCYxcP'), verify_certs=False, timeout=30)
16 i = 0
17
18
19 def find_hashtags(text):
20     return re.findall("#(\w+)", text)
21
22
23 def on_item_update(item_update):
24     record = item_update
25     if record["values"]["get("content")"] == None:
26         record["values"]["content"] = record["values"]["content"].encode().decode(
27             "unicode_escape")
28     record["values"]["senderName"] = record["values"]["senderName"].encode().decode(
29         "unicode_escape")
30     if record["values"].get("parentContent") == None:
31         record["values"]["parentContent"] = record["values"]["parentContent"].encode(
32             "unicode_escape")
33     record["values"]["parentSenderName"] = record["values"]["parentSenderName"].encode(
34         "unicode_escape")
35
36     print(record)
37     json_doc = {}
38     if record["values"]["content"] != None:
39         json_doc["sendTime"] = record["values"]["sendTime"]
40         json_doc["hashtags"] = find_hashtags(record["values"]["content"])
41         json_doc["content"] = record["values"]["content"]
42         json_doc["type"] = record["values"]["type"]
43         json_doc["senderName"] = record["values"]["senderName"]
44         es.index(index="twitter1", body=json.dumps(json_doc))
45         global i
46         i += 1
47         print("{} tweets fetched".format(i))
```

حال برنامه را به اجرا در می آوریم تا 300 توییت دریافت کند.

ما 378 توبیت استخراج کردیم. در گام بعد از آن ها استفاده میکنیم.

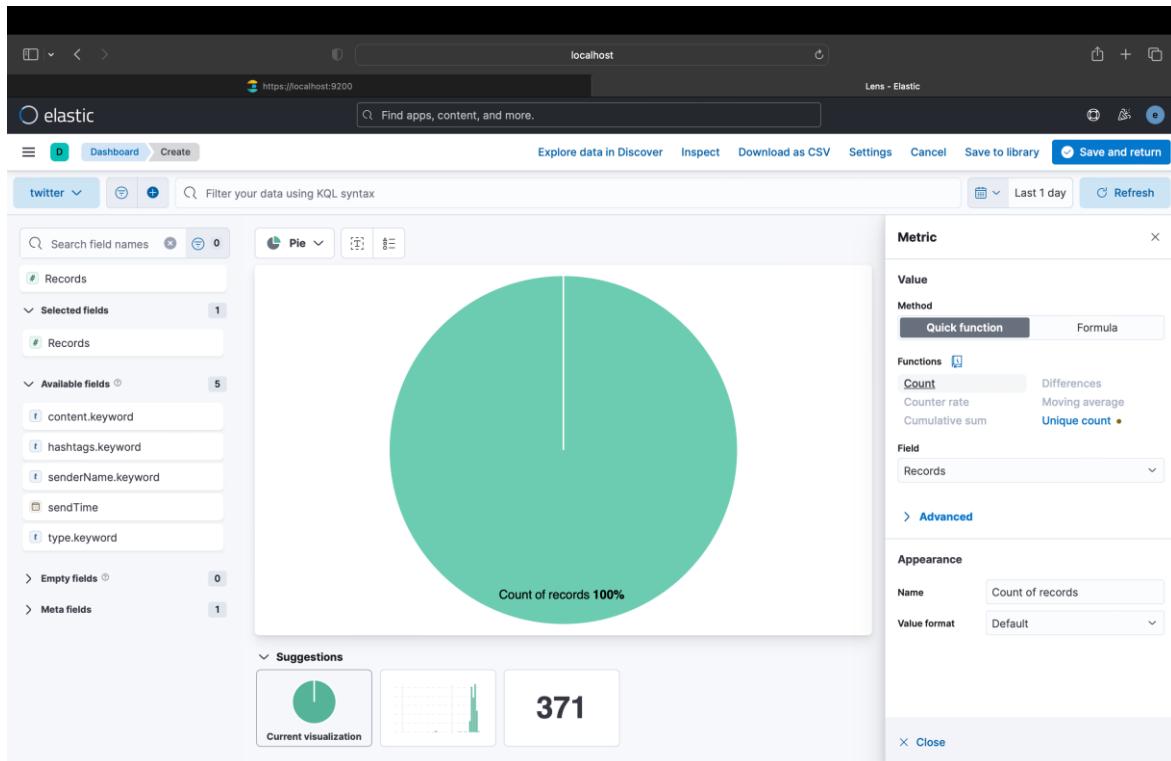
گام سوم:

ابتدا یک index pattern با data view میسازیم.

The screenshot shows the Elasticsearch Index Management interface. On the left, there's a sidebar with various management options like Ingest Pipelines, Index Management (selected), and Data Views. The main area is titled 'Index Management' and shows a table of indices. One row is selected, labeled 'twitter1', which has a yellow status indicator. A modal window titled 'twitter1' is open, showing detailed information about the index, including its health (yellow), primary shards (1), document count (380), storage size (402.55kb), and aliases (none). A 'Manage' button is at the bottom right of the modal.

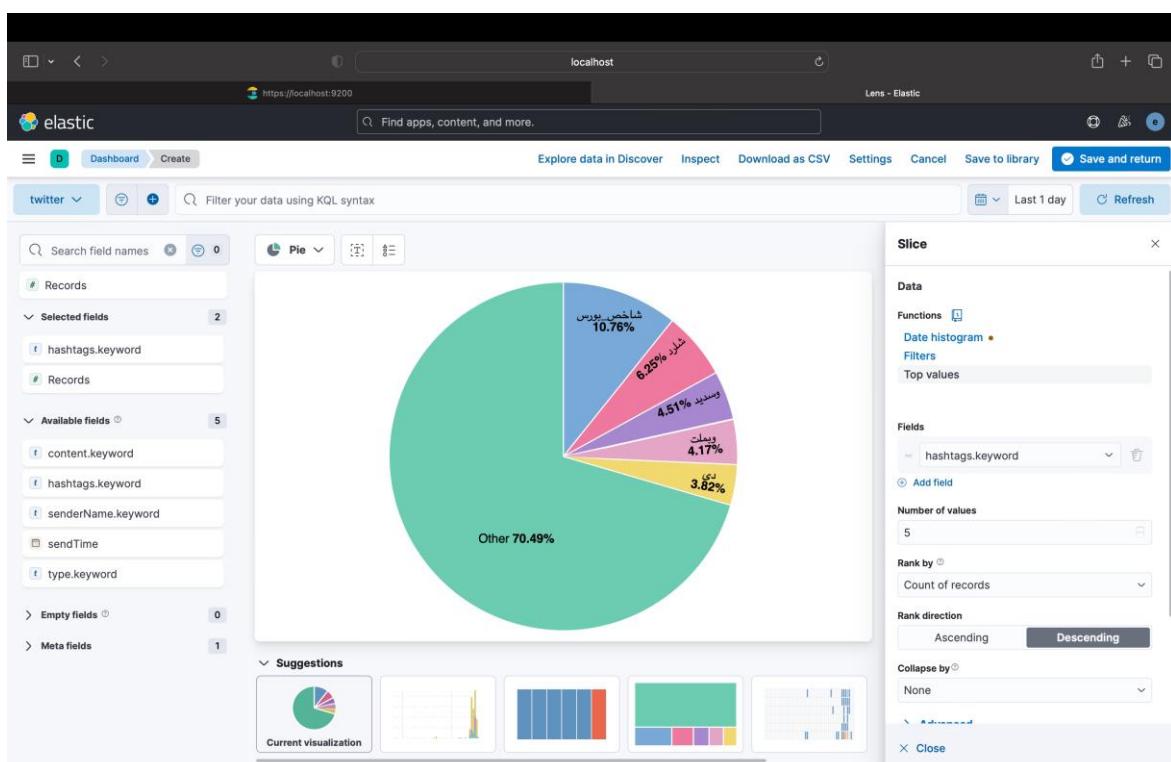
The screenshot shows the Elasticsearch Data Views interface. The left sidebar includes options like Rollup Jobs, Transforms, and Data Views (selected). A data view named 'twitter' is being edited. The configuration page shows the index pattern 'tw*' and time field 'sendTime'. Below this, there's a table of fields with 13 entries. The table columns include Name, Type, Format, Searchable, Aggregatable, and Excluded. Fields listed include '_id', '_index', '_score', '_source', 'content', 'content.keyword', 'hashtags', 'hashtags.keyword', 'sendTime', and 'senderName'. A 'Set as default' button is visible at the top right of the data view configuration area.

پس از visualization و dashboard را میسازیم.

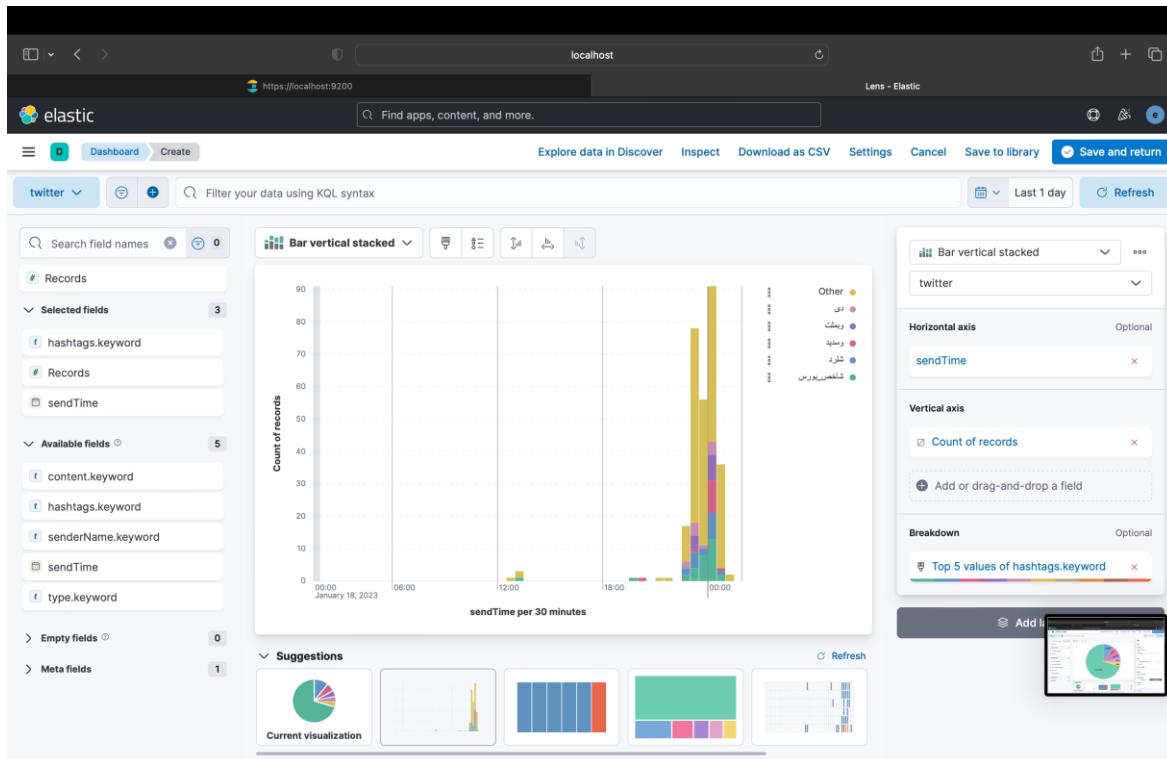


در نسخه‌ی کیبانای من فیلتری به نام tag cloud وجود نداشت در نتیجه در نمودار‌های دیگر استفاده استفاده کردم.

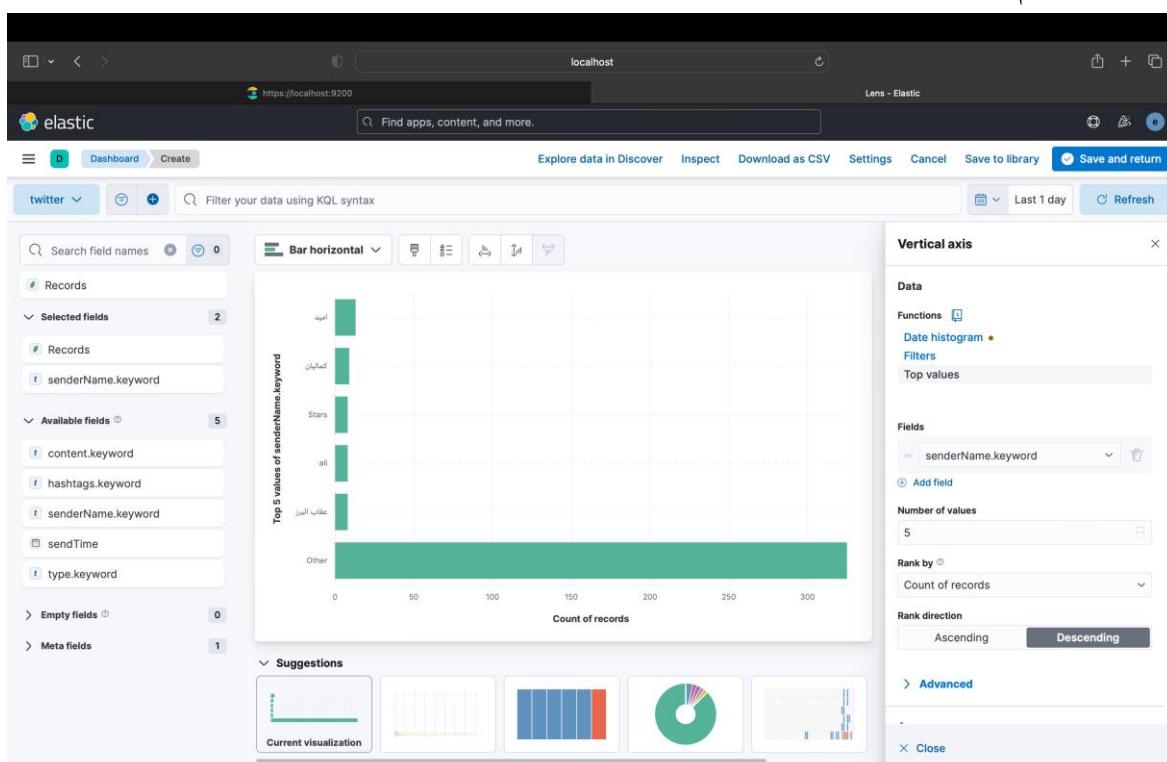
نمودار 5 هشتگ بیشتر



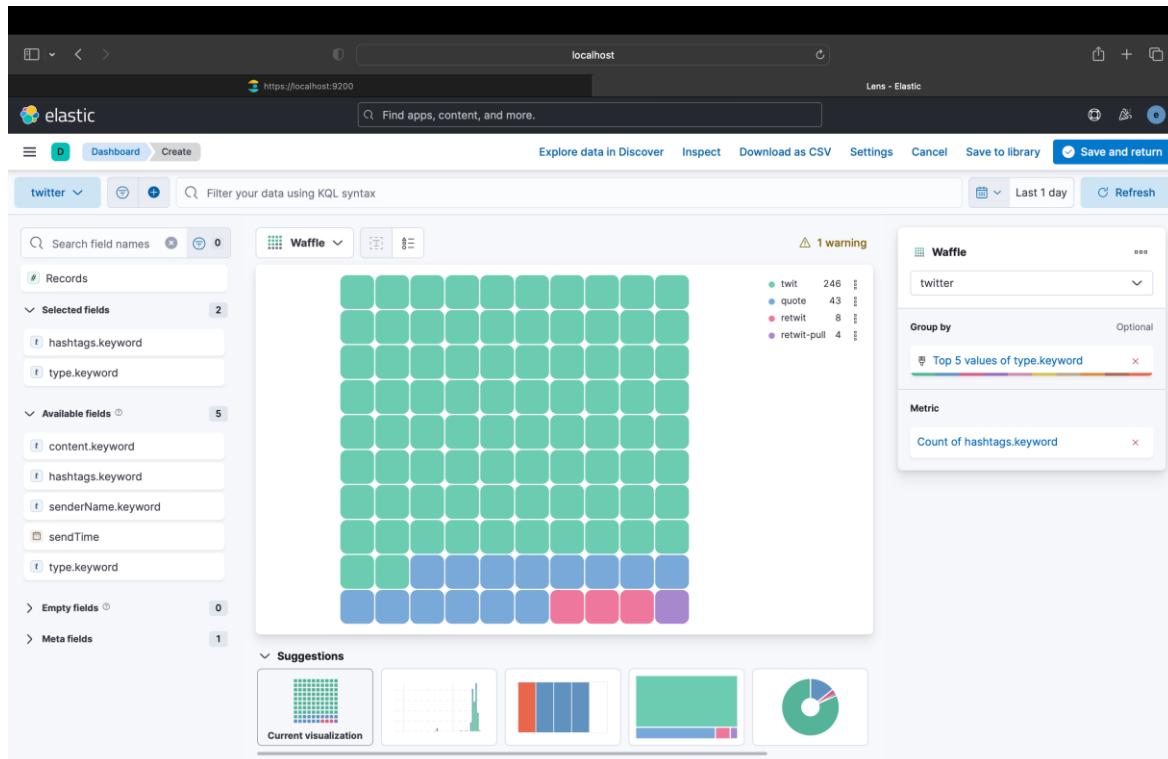
تایم هشتگ ها Bar plot



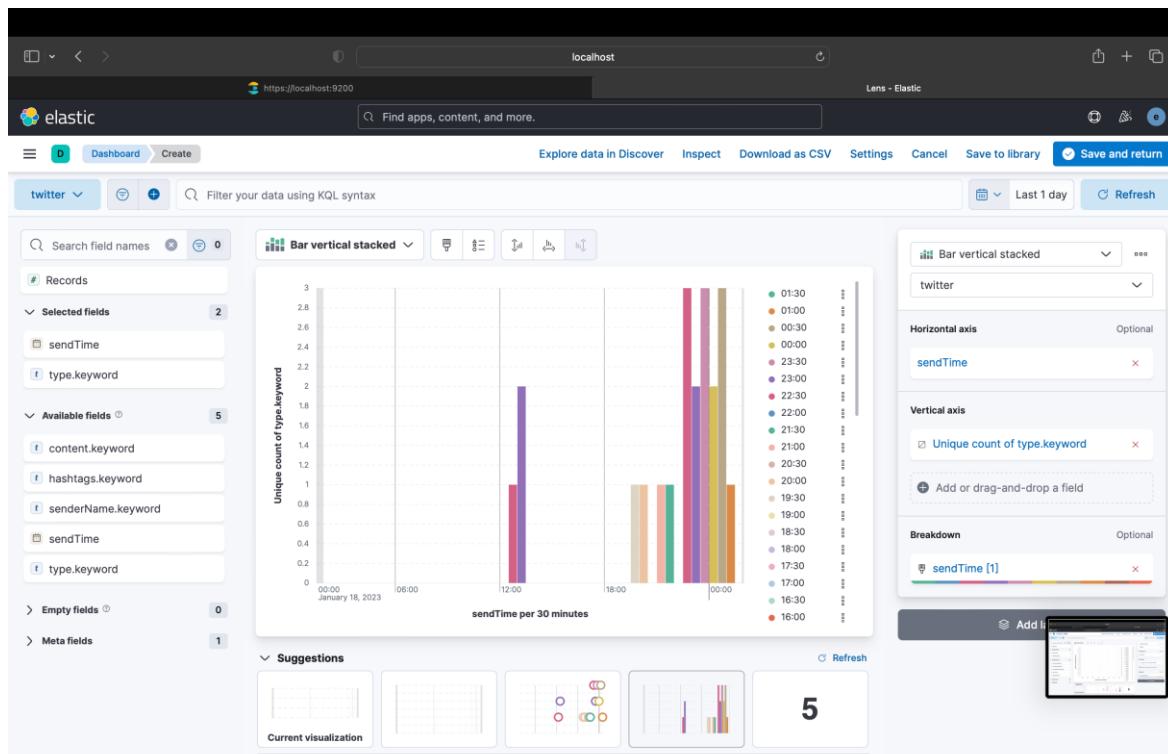
اسم 5 نفری که بیشتر توییت گذاشتند.



تایپ توییت ها Waffle



نمودار زمان توییت ها



آنچه آموختم / پیشنهادات

در این دستور کار کار با elasticsearch, kibana را آموختم ولی مشکلاتی سر نصب هردوی آن ها داشتم و مجبور شدم آخر از یک لپ تاپ دیگر استفاده کنم با سیستم عامل متفاوت. ورژن کیانا جدید تر شده بود و بعضی قسمت ها را نمیدونستم چطور پیدا کنم در آخر هم قسمت tag cloud کلا نبود. برای اسکریپت پایتون از دوستانم کمک گرفتم.