

Exerc-Aula05-Elastic - Analyzer

Dica: Para setar a vm.max_map_count

```
marco@DESKTOP-G2455QH:/mnt/c/Windows/system32$ sudo sysctl -w  
vm.max_map_count=262144
```

[sudo] password for marco:

```
vm.max_map_count = 262144
```

```
marco@DESKTOP-G2455QH:/mnt/c/Windows/system32$
```

1. Criar os Analyzer simple, standard, brazilian e portuguese para a seguinte frase:

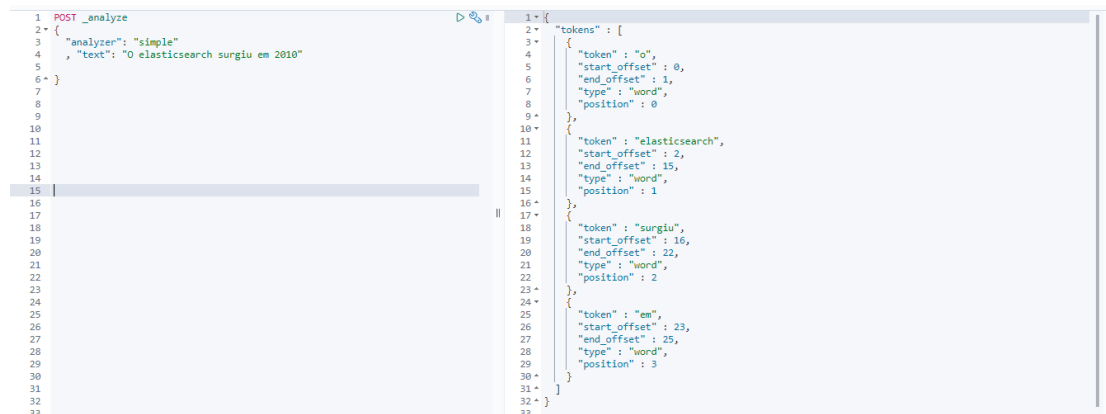
Dica: O Simple :Remove os numeros, deixa tudo em minusculo e retira os espaços.

O elasticsearch surgiu em 2010

**** Resposta:**

POST _analyze

```
{  
  "analyzer": "simple"  
  , "text": "O elasticsearch surgiu em 2010"  
}
```



XX

Dica: O standard deixa o token exatamente igual, retirando apenas os espaços.

POST _analyze

```
{
```

}

```

1 POST _analyze
2 {
3   "analyzer": "standard"
4   , "text": "O elasticsearch surgui em 2010"
5 }
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33

```

```

8 "position" : 0
9 },
10 {
11   "token" : "elasticsearch",
12   "start_offset" : 2,
13   "end_offset" : 15,
14   "type" : "<ALPHANUM>",
15   "position" : 1
16 },
17 {
18   "token" : "surgui",
19   "start_offset" : 16,
20   "end_offset" : 22,
21   "type" : "<ALPHANUM>",
22   "position" : 2
23 },
24 {
25   "token" : "em",
26   "start_offset" : 23,
27   "end_offset" : 25,
28   "type" : "<ALPHANUM>",
29   "position" : 3
30 },
31 {
32   "token" : "2010",
33   "start_offset" : 26,
34   "end_offset" : 30,
35   "type" : "<NUM>",
36   "position" : 4
37 },
38 ]
39 }
40

```

XX

 $\{$

}

```
1 POST _analyze
2 {
3   "analyzer": "brazilian"
4   , "text": "O elasticsearch surgiu em 2010"
5 }
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
```

```
1 {
2   "tokens" : [
3     {
4       "token" : "elasticsearch",
5       "start_offset" : 2,
6       "end_offset" : 15,
7       "type" : "<ALPHANUM>",
8       "position" : 1
9     },
10    {
11      "token" : "surg",
12      "start_offset" : 16,
13      "end_offset" : 22,
14      "type" : "<ALPHANUM>",
15      "position" : 2
16    },
17    {
18      "token" : "2010",
19      "start_offset" : 26,
20      "end_offset" : 30,
21      "type" : "<NUM>",
22      "position" : 4
23    }
24  ]
25 }
26
```

[illegible]

Dica retirou o "O" e o em,

POST _analyze

```
{  
  "analyzer": "portuguese"  
  , "text": "O elasticsearch surgiu em 2010"  
}
```



```
1 POST _analyze  
2 {  
3   "analyzer": "portuguese"  
4   , "text": "O elasticsearch surgiu em 2010"  
5 }  
6 ^  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28
```

```
1 {  
2   "tokens" : [  
3     {  
4       "token" : "elasticsearch",  
5       "start_offset" : 2,  
6       "end_offset" : 15,  
7       "type" : "<ALPHANUM>",  
8       "position" : 1  
9     },  
10    {  
11      "token" : "surgiu",  
12      "start_offset" : 16,  
13      "end_offset" : 22,  
14      "type" : "<ALPHANUM>",  
15      "position" : 2  
16    },  
17    {  
18      "token" : "2010",  
19      "start_offset" : 26,  
20      "end_offset" : 30,  
21      "type" : "<NUM>",  
22      "position" : 4  
23    }  
24  ]  
25 }  
26
```

2. Realizar os passos no índice produto

Dica: 1º Precisamos verificar o mapping

GET produto/_mapping

Oservamos que não temos o analyze e que não podemos usar o PUT direto, pois já existe.A solução é fazer o processo de reindexação.

```

1 GET produto/_mapping
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33

```

```

1 {
2   "produto": {
3     "mappings": {
4       "properties": {
5         "data": {
6           "type": "date"
7         },
8         "descricao": {
9           "type": "text",
10          "fields": {
11            "keyword": {
12              "type": "keyword",
13              "ignore_above": 256
14            }
15          }
16        },
17        "nome": {
18          "type": "text",
19          "fields": {
20            "keyword": {
21              "type": "keyword",
22              "ignore_above": 256
23            }
24          }
25        },
26        "qtd": {
27          "type": "long"
28        }
29      }
30    }
31  }
32 }
33

```

a) Criar um analyzer brazilian para o atributo descricao

PUT produto1

```

{
  "settings": {
    "index": {
      "number_of_shards": 1,
      "number_of_replicas": 0
    }
  },
  "mappings": {
    "properties": {
      "descricao": {
        "type": "text",
        "analyzer": "brazilian"
      }
    }
  }
}

```

OBs: Onde está produto99 leia-se produto1

```
1 PUT produto99
2 {
3   "settings": {
4     "index": {
5       "number_of_shards" : 1,
6       "number_of_replicas" : 0
7     }
8   },
9   "mappings": {
10    "properties": {
11      "descricao": {
12        "type": "text",
13        "analyzer": "brazilian"
14      }
15    }
16  }
17 }
18
19
20
```

```
1 {
2   "acknowledged" : true,
3   "shards_acknowledged" : true,
4   "index" : "produto99"
5 }
6
```

Agora precisamos reindexar o nosso produto para produto1

POST _reindex

```
{
  "source": {
    "index": "produto"
  },
  "dest": {
    "index": "produto1"
  }
}
```

```
1 POST _reindex
2 {
3   "source": {
4     "index": "produto"
5   },
6   "dest": {
7     "index": "produto1"
8   }
9 }
10
11
12
13
14
15
16
17
18
19
20
21
22
```

```
1 {
2   "took" : 174,
3   "timed_out" : false,
4   "total" : 5,
5   "updated" : 0,
6   "created" : 5,
7   "deleted" : 0,
8   "batches" : 1,
9   "version_conflicts" : 0,
10  "noops" : 0,
11  "retries" : {
12    "bulk" : 0,
13    "search" : 0
14  },
15  "throttled_millis" : 0,
16  "requests_per_second" : -1.0,
17  "throttled_until_millis" : 0,
18  "failures" : [ ]
19 }
20
```

Para verificarmos :

GET produto1/_mapping

```

1 GET produto1/_mapping
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28

```

```

1 {
2   "produto1" : {
3     "mappings" : {
4       "properties" : {
5         "data" : {
6           "type" : "date"
7         },
8         "descricao" : {
9           "type" : "text",
10          "analyzer" : "brazilian"
11        },
12        "nome" : {
13          "type" : "text",
14          "fields" : {
15            "keyword" : {
16              "type" : "keyword",
17              "ignore_above" : 256
18            }
19          }
20        },
21        "qtd" : {
22          "type" : "long"
23        }
24      }
25    }
26  }
27 }
28

```

OBs: Para deixar com o nome "produto" precisaríamos fazer o reindex de produto1 para produto.

b) Para o atributo descricao aplicar o analyzer brazilian para o tipo de campo text e criar o atributo descricao.original com o dado do tipo keyword

PUT produto1

```

{
  "settings": {
    "index": {
      "number_of_shards": 1,
      "number_of_replicas": 0
    }
  },
  "mappings": {
    "properties": {
      "descricao": {
        "type": "text",
        "analyzer": "brazilian",
        "fields": { "original": {
          "type": "keyword"
        } }
      }
    }
  }
}

```

```

}
}

1 PUT produto1
2 {
3   "settings": {
4     "index": {
5       "number_of_shards" : 1,
6       "number_of_replicas" : 0
7     }
8   },
9   "mappings": {
10    "properties": {
11      "descricao": {
12        "type" : "text",
13        "analyzer": "brazilian",
14        "fields": { "original" : {
15          "type": "keyword"
16        } }
17      },
18      "nome": {
19        "type" : "text",
20        "fields": {
21          "keyword": {
22            "type" : "keyword",
23            "ignore_above" : 256
24          }
25        }
26      },
27      "qtd": {
28        "type" : "long"
29      }
30    }
31  }
32 }

```

Para renomear produto1 para produto

1ª) deletar o produto

```

1 DELETE produto
2
3
4
5

```

2ª)

POST _reindex

```

{
  "source": {
    "index": "produto1"
  },
  "dest": {
    "index": "produto"
  }
}

```

c) Buscar a palavra “compativei” no campo descricao.original (hits = 0)

Dica sempre que for usar o analyse, fazer as 3 fases abaixo

```
test
cria indice analyzer
reindex
```

** Resposta (Corretamente não irá trazer nenhum dado, pois no texto está com acento e no analyse sem acento e provavelmente apenas o radical.)

GET produto/_search

```
{
  "query": {
    "match": {
      "descricao.keyword": "compativel"
    }
  }
}
```



XX

d) Buscar a palavra “compativel” no campo descricao

Dica sempre que for usar o analyse, fazer as 3 fases abaixo

```
test
cria indice analyzer
reindex
```

** Resposta (Dessa forma achou)

GET produto/_search

```
{
  "query": {
    "match": {
```



```

    "descricao": "compativel"
  }
}
}

```

```

1 GET produto/_search
2 {
3   "query": {
4     "match": {
5       "descricao": "compativel"
6     }
7   }
8 }
9
1 {
2   "took" : 0,
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" : 1.1129161,
16    "hits" : [
17      {
18        "_index" : "produto",
19        "_type" : "_doc",
20        "_id" : "1",
21        "_score" : 1.1129161,
22        "_source" : {
23          "nome" : "mouse",
24          "qtd" : 50,
25          "descricao" : "com fio USB, compativel com Windows,
26                        Mac e Linux"
27        }
28      }
29    ]
30  }
}

```

3. Clicar no botão de Enviar Tarefa, e enviar o texto: ok