

~/HashAgile.py

```
1 # THIS WORK IS DONE BY PRIYANSH CHHABRA
2 # COLLEGE : GLAGOTIAS COLLEGE OF ENGINEERING AND TECHNOLOGY
3 # DOMAIN : FULL STACK DEVELOPER
4 # CODE IS WRITTEN BY PRIYANSH CHHABRA
5
6
7
8
9 from elasticsearch import Elasticsearch, helpers
10 import pandas as pd
11
12 # In this way i am connecting it to elasticsearch
13
14 es = Elasticsearch([{'host': 'localhost', 'port': 9200, 'scheme': 'http'}])
15
16
17 def createCollection(p_collection_name):
18     if not es.indices.exists(index=p_collection_name):
19         es.indices.create(index=p_collection_name)
20         print(f"Collection {p_collection_name} created.")
21     else:
22         print(f"Collection {p_collection_name} already exists.")
23
24 def indexData(p_collection_name, p_exclude_column):
25
26     df = pd.read_csv('/Users/priyansh/Downloads/Employee Sample Data 1.csv')
27     df = df.drop(columns=[p_exclude_column])
28
29
30     actions = [
31         {
32             "_index": p_collection_name,
33             "_id": str(row['Employee_id']),
34             "_source": row.to_dict()
35         }
36         for _, row in df.iterrows()
37     ]
38
39
40     helpers.bulk(es, actions)
41     print(f"Data indexed into {p_collection_name}, excluding {p_exclude_column}
42     .")
43
44 def searchByColumn(p_collection_name, p_column_name, p_column_value):
45     body = {
46         "query": {
47             "match": {
48                 p_column_name: p_column_value
49             }
50         }
51     }
52     res = es.search(index=p_collection_name, body=body)
53     return res['hits']['hits']
```

```
54
55 def getEmpCount(p_collection_name):
56     res = es.count(index=p_collection_name)
57     return res['count']
58
59
60 def delEmpById(p_collection_name, p_employee_id):
61     es.delete(index=p_collection_name, id=p_employee_id)
62     print(f"Employee {p_employee_id} deleted from {p_collection_name}.")
63
64 def getDepFacet(p_collection_name):
65     body = {
66         "aggs": {
67             "by_department": {
68                 "terms": {
69                     "field": "Department.keyword"
70                 }
71             }
72         }
73     }
74     res = es.search(index=p_collection_name, body=body)
75     return res['aggregations']['by_department']['buckets']
76
77
78 v_nameCollection = 'Hash_Priyansh'
79 v_phoneCollection = 'Hash_2234'
80
81
82 createCollection(v_nameCollection)
83 createCollection(v_phoneCollection)
84
85 print(f"Initial employee count: {getEmpCount(v_nameCollection)}")
86
87 indexData(v_nameCollection, 'Department')
88 indexData(v_phoneCollection, 'Gender')
89
90 delEmpById(v_nameCollection, 'E02003')
91
92
93 print(f"Employee count after deletion: {getEmpCount(v_nameCollection)}")
94
95
96 print(searchByColumn(v_nameCollection, 'Department', 'IT'))
97 print(searchByColumn(v_nameCollection, 'Gender', 'Male'))
98 print(searchByColumn(v_phoneCollection, 'Department', 'IT'))
99
100
101 print(getDepFacet(v_nameCollection))
102 print(getDepFacet(v_phoneCollection))
103
```