HASH AGILE – SECOND CHALLENGE

ASSIGNMENT:

1. Read about Elasticsearch: Getting Started with Elasticsearch.
2. Install Elasticsearch on your local machine.
3. Create an index (collection) in Elasticsearch.
4. Index the Employee data
   from https://www.kaggle.com/datasets/williamlucas0/employee-sample-data

SOLUTION:

Step by Step Process:

1. Install Elasticsearch:

- First, download and install Elasticsearch on your local machine:

➢ For Windows visit the Elasticsearch download page.
➢ Extract the files and follow the installation steps.
➢ Start Elasticsearch by navigating to the extracted folder and running:
- Windows: bin\elasticsearch.bat

➢ Elasticsearch will be running at http://localhost:9200.

To sign in localhost: username:elastic **, password:** s_A0pGy+sJY-Z5OvyO8-

2. Index Employee Data:

I decided to interact with Elasticsearch using Python because for its simplicity.

Install Elasticsearch Python client:

- **Run:**

  pip install elasticsearch

Download Employee Data:

- Go to the Kaggle link you were provided: Employee Data and download the dataset.
- Extract the CSV file for use.

Create Python Script to Index Data:

Filename: index_employee_data.py

```
# Import necessary libraries

import csv

from elasticsearch import Elasticsearch


# Initialize Elasticsearch client

es = Elasticsearch(

    [{'scheme': 'http', 'host': 'localhost', 'port': 9200}],

    basic_auth=('elastic', 's_A0pGy+sJY-Z5OvyO8-')  # Replace 'your_password' with the actual password

)


# Create an index (collection) in Elasticsearch

index_name = 'employee_data'

if not es.indices.exists(index=index_name):
```

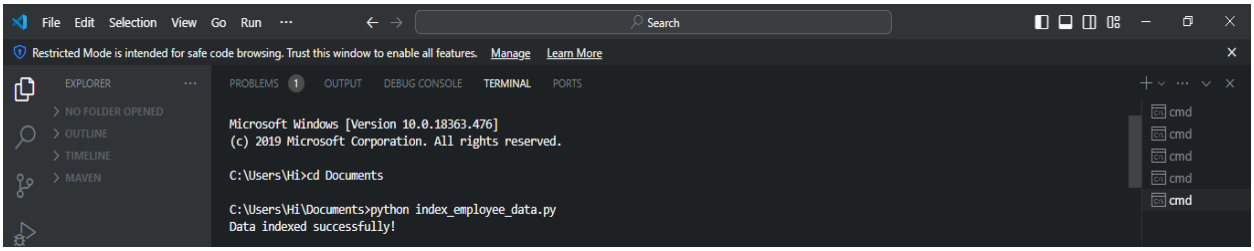
```
    es.indices.create(index=index_name)


# Read employee data from CSV and index it
def index_employee_data(csv_file):
    with open(csv_file, newline='') as file:
        reader = csv.DictReader(file)
        for i, row in enumerate(reader):
            es.index(index=index_name, id=i+1, document=row)


# Path to your CSV file
csv_file_path = r'C:\Users\Hi\Downloads\Employee Sample Data 1.csv'


# Index the data
index_employee_data(csv_file_path)


print("Data indexed successfully!")
```

3. Run the Python Script:

Run:

```
python index_employee_data.py
```

After running the above command, the successful indexing message shown in terminal



4. Verify Indexed Data:

We can verify that the data is successfully indexed by using a GET request to Elasticsearch

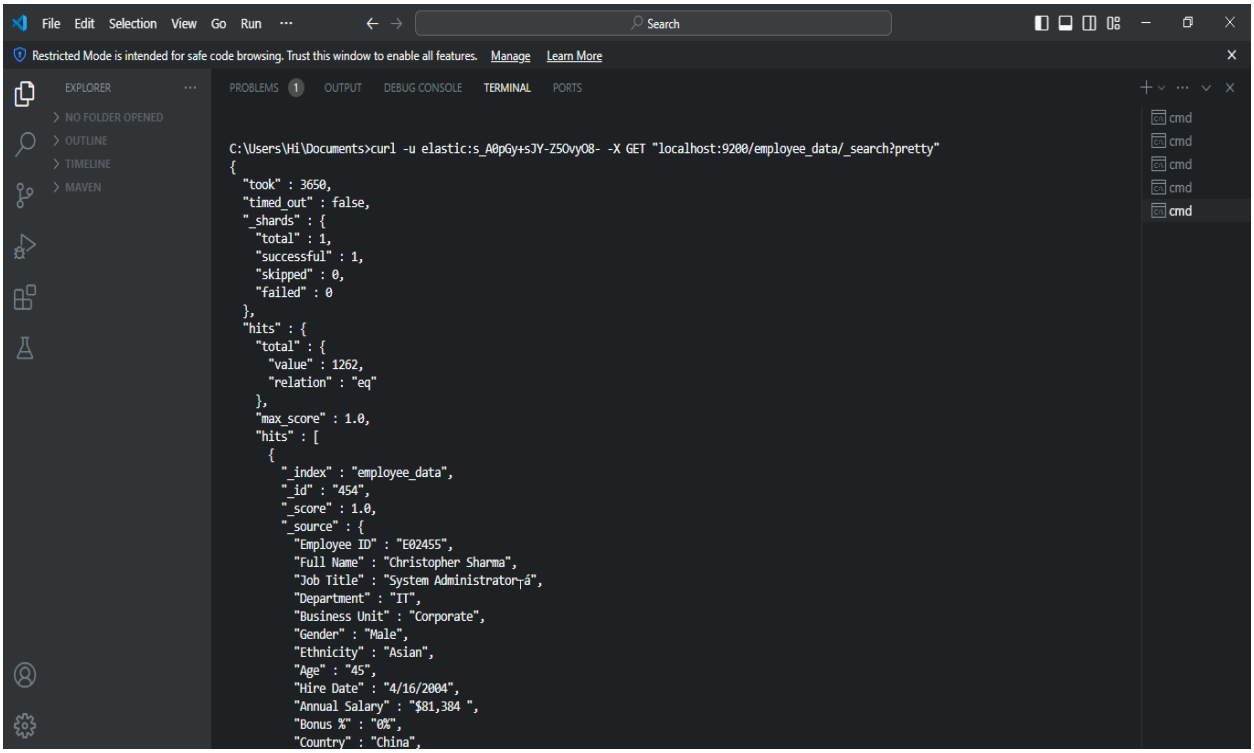
Run:

curl -u elastic:s_A0pGy+sJY-Z5OvyO8- -X GET "localhost:9200/employee_data/_search?pretty"

Breakdown of the Command:

- curl: The command-line tool for transferring data.
- -u elastic:s_A0pGy+sJY-Z5OvyO8-: This part specifies the username (elastic) and password (s_A0pGy+sJY-Z5OvyO8-) for basic authentication.
- -X GET: This specifies the HTTP method to use, which is GET in this case.
- "localhost:9200/employee_data/_search?pretty": This is the URL for the search request, with pretty formatting for easier reading.

Output:



Terminal Output:

```
C:\Users\Hi\Documents>curl -u elastic:s_A0pGy+sJY-Z5OvyO8- -X GET
"localhost:9200/employee_data/_search?pretty"
{
 "took" : 3650,
 "timed_out" : false,
 "_shards" : {
  "total" : 1,
  "successful" : 1,
  "skipped" : 0,
  "failed" : 0
 },
 "hits" : {
  "total" : {
   "value" : 1262,
   "relation" : "eq"
  },
  "max_score" : 1.0,
  "hits" : [
   {
    "_index" : "employee_data",
    "_id" : "454",
    "_score" : 1.0,
    "_source" : {
     "Employee ID" : "E02455",
     "Full Name" : "Christopher Sharma",
     "Job Title" : "System Administrator┬á",
     "Department" : "IT",
     "Business Unit" : "Corporate",
     "Gender" : "Male",
     "Ethnicity" : "Asian",
     "Age" : "45",
     "Hire Date" : "4/16/2004",
     "Annual Salary" : "$81,384 ",
     "Bonus %" : "0%",
     "Country" : "China",
```

```
    "City" : "Beijing",
    "Exit Date" : ""
   }
  },
  {
   "_index" : "employee_data",
   "_id" : "455",
   "_score" : 1.0,
   "_source" : {
    "Employee ID" : "E02456",
    "Full Name" : "Lucas Grant",
    "Job Title" : "Sr. Business Partner",
    "Department" : "Human Resources",
    "Business Unit" : "Manufacturing",
    "Gender" : "Male",
    "Ethnicity" : "Caucasian",
    "Age" : "53",
    "Hire Date" : "3/28/2012",
    "Annual Salary" : "$97,218 ",
    "Bonus %" : "0%",
    "Country" : "United States",
    "City" : "Austin",
    "Exit Date" : ""
   }
  },
  {
   "_index" : "employee_data",
   "_id" : "456",
   "_score" : 1.0,
   "_source" : {
    "Employee ID" : "E02457",
    "Full Name" : "Delilah Washington",
    "Job Title" : "Analyst",
    "Department" : "Sales",
    "Business Unit" : "Research & Development",
    "Gender" : "Female",
    "Ethnicity" : "Caucasian",
```

```
    "Age" : "46",
    "Hire Date" : "9/17/2006",
    "Annual Salary" : "$42,944 ",
    "Bonus %" : "0%",
    "Country" : "United States",
    "City" : "Miami",
    "Exit Date" : ""
   }
  },
  {
   "_index" : "employee_data",
   "_id" : "457",
   "_score" : 1.0,
   "_source" : {
    "Employee ID" : "E02458",
    "Full Name" : "Eloise Molina",
    "Job Title" : "Analyst",
    "Department" : "Finance",
    "Business Unit" : "Research & Development",
    "Gender" : "Female",
    "Ethnicity" : "Latino",
    "Age" : "48",
    "Hire Date" : "1/18/2015",
    "Annual Salary" : "$43,980 ",
    "Bonus %" : "0%",
    "Country" : "United States",
    "City" : "Columbus",
    "Exit Date" : ""
   }
  },
  {
   "_index" : "employee_data",
   "_id" : "458",
   "_score" : 1.0,
   "_source" : {
    "Employee ID" : "E02459",
    "Full Name" : "Clara Desai",
```

```
    "Job Title" : "Manager",
    "Department" : "Marketing",
    "Business Unit" : "Corporate",
    "Gender" : "Female",
    "Ethnicity" : "Asian",
    "Age" : "33",
    "Hire Date" : "11/20/2014",
    "Annual Salary" : "$109,533 ",
    "Bonus %" : "5%",
    "Country" : "China",
    "City" : "Chengdu",
    "Exit Date" : ""
   }
},
{
  "_index" : "employee_data",
  "_id" : "459",
  "_score" : 1.0,
  "_source" : {
    "Employee ID" : "E02460",
    "Full Name" : "Audrey Campbell",
    "Job Title" : "Vice President",
    "Department" : "Accounting",
    "Business Unit" : "Research & Development",
    "Gender" : "Female",
    "Ethnicity" : "Caucasian",
    "Age" : "56",
    "Hire Date" : "7/30/2008",
    "Annual Salary" : "$226,952 ",
    "Bonus %" : "39%",
    "Country" : "United States",
    "City" : "Miami",
    "Exit Date" : ""
   }
},
{
  "_index" : "employee_data",
```

```
  "_id" : "460",
  "_score" : 1.0,
  "_source" : {
    "Employee ID" : "E02461",
    "Full Name" : "Delilah Alvarez",
    "Job Title" : "System Administrator┬á",
    "Department" : "IT",
    "Business Unit" : "Manufacturing",
    "Gender" : "Female",
    "Ethnicity" : "Latino",
    "Age" : "59",
    "Hire Date" : "11/27/2007",
    "Annual Salary" : "$71,580 ",
    "Bonus %" : "0%",
    "Country" : "United States",
    "City" : "Seattle",
    "Exit Date" : ""
  }
},
{
  "_index" : "employee_data",
  "_id" : "461",
  "_score" : 1.0,
  "_source" : {
    "Employee ID" : "E02462",
    "Full Name" : "Luke Luna",
    "Job Title" : "Analyst II",
    "Department" : "Finance",
    "Business Unit" : "Manufacturing",
    "Gender" : "Male",
    "Ethnicity" : "Latino",
    "Age" : "55",
    "Hire Date" : "11/16/2012",
    "Annual Salary" : "$66,172 ",
    "Bonus %" : "0%",
    "Country" : "Brazil",
    "City" : "Manaus",
```

```
    "Exit Date" : ""
   }
},
{
  "_index" : "employee_data",
  "_id" : "462",
  "_score" : 1.0,
  "_source" : {
    "Employee ID" : "E02463",
    "Full Name" : "Elizabeth Tan",
    "Job Title" : "Vice President",
    "Department" : "Marketing",
    "Business Unit" : "Research & Development",
    "Gender" : "Female",
    "Ethnicity" : "Asian",
    "Age" : "46",
    "Hire Date" : "7/27/2018",
    "Annual Salary" : "$237,489 ",
    "Bonus %" : "30%",
    "Country" : "United States",
    "City" : "Chicago",
    "Exit Date" : ""
  }
},
{
  "_index" : "employee_data",
  "_id" : "463",
  "_score" : 1.0,
  "_source" : {
    "Employee ID" : "E02464",
    "Full Name" : "Sophia Wang",
    "Job Title" : "Analyst",
    "Department" : "Marketing",
    "Business Unit" : "Specialty Products",
    "Gender" : "Female",
    "Ethnicity" : "Asian",
    "Age" : "50",
```

```
            "Hire Date" : "5/20/2000",
            "Annual Salary" : "$48,733 ",
            "Bonus %" : "0%",
            "Country" : "United States",
            "City" : "Seattle",
            "Exit Date" : ""
          }
          }
        }
      ]
    }
     ]
    }
    }
}
```


Note:

Reset the Password (if needed): If you don't know the password for the elastic user, you can reset it using the command line:

Open a command prompt and navigate to your Elasticsearch bin directory, then **Run**:

elasticsearch-reset-password -u elastic