UE17CS412-AIR Assignment

Dataset

- The given dataset link consists of 418 csv files.
- Consider each CSV file as an input document. CSV file is a comma separated values file.
- Do the indexing for each csv file.
- Search can be done on one csv or all the csv files.
- It should display the searched row and the document in which it was found.
- Look each csv file as table in a database and URL, MatchDate, Station, Show, IAShowID, IAPreview, Snippet are the columns.

Note:

- Don't combine all the csv files into a single file.
- There is no need for redirection to transcript. URL is just a column in the csv.
- Snippet is different for each row and no need to combine snippet also to form a single document.

Here is a sample output screenshot for indian_foods.csv using elasticsearch

Fig 2: Search name=Vineetha* in elasticsearch

```
{
  "took" : 17,
  "timed_out" : false,
  "_shards" : {
    "total" : 1,
    "successful" : 1,
    "skipped" : 0,
    "failed" : 0
  },
  "hits" : {
    "total" : {
      "value" : 1,
      "relation" : "eq"
    },
    "max_score" : 6.0996456,
    "hits" : [
      {
        "_index" : "welcome",
        "_type" : "_doc",
        "_id" : "Fb169HQBV24UwIrTTNry",
        "_score" : 6.0996456,
        "_source" : {
          "host" : "vijay-Lenovo-Erazer-Y50-70",
          "ingredients" : "Rose syrup, falooda sev, mixed nuts, saffron, sugar",
          "cook_time" : "25",
          "path" : "/home/vijay/opt/indian_food.csv",
          "name" : "VineethaPES",
          "prep_time" : "45",
          "@version" : "1",
          "message" : "VineethaPES,\"Rose syrup, falooda sev, mixed nuts, saffron, sugar\",vegetarian,45,25,sweet,dessert,-1,-1",
          "state" : "-1",
          "diet" : "vegetarian",
          "flavor_profile" : "sweet",
          "region" : "-1",
          "course" : "dessert",
          "@timestamp" : "2020-10-04T16:38:08.418Z"
        }
      }
    ]
  }
}
```

Fig 2: Search name=Vineethapes in elasticsearch