ADBMS ASSIGNMENT REPORT
A case study related to Covid-19

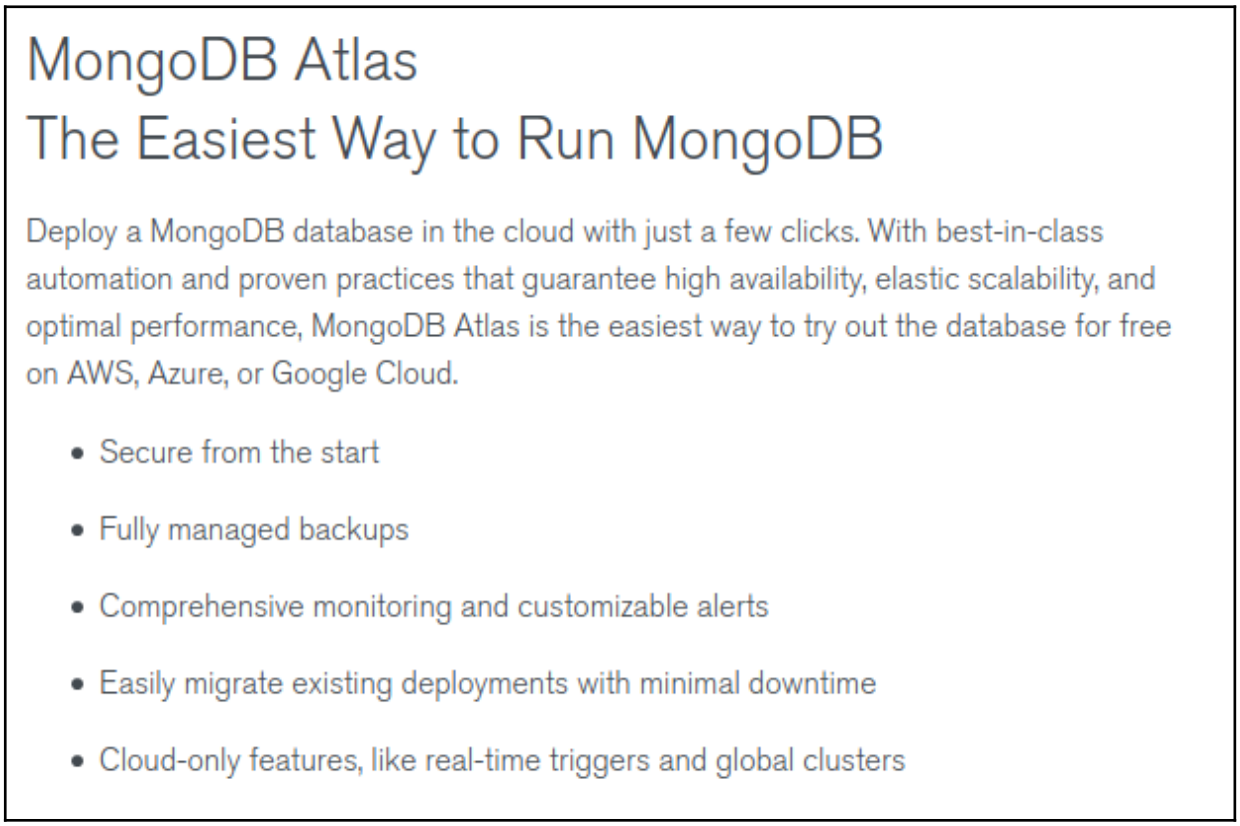
Submitted to:

Dr.Shelly Sachdeva
HoD CSE Dept.
NIT Delhi

Submitted By:

G.D.V.Sujith Sagar
171210023
CSE Dept.
Btech 3rd year
NIT Delhi

Database used



Start Mongo Server in Terminal

```
sujith@sujith-HP-Pavilion-Notebook:~$ service mongodb status
● mongodb.service - An object/document-oriented database
     Loaded: loaded (/lib/systemd/system/mongodb.service; disabled; vendor preset: enabled)
     Active: inactive (dead)
       Docs: man:mongod(1)
sujith@sujith-HP-Pavilion-Notebook:~$ sudo service mongodb start
[sudo] password for sujith:
sujith@sujith-HP-Pavilion-Notebook:~$ service mongodb status
● mongodb.service - An object/document-oriented database
     Loaded: loaded (/lib/systemd/system/mongodb.service; disabled; vendor preset: enabled)
     Active: active (running) since Fri 2020-05-22 10:42:17 IST; 3s ago
       Docs: man:mongod(1)
   Main PID: 3020 (mongod)
      Tasks: 1 (limit: 9345)
     Memory: 36.8M
     CGroup: /system.slice/mongodb.service
             └─3020 /usr/bin/mongod --unixSocketPrefix=/run/mongodb --config /etc/mongodb.conf

May 22 10:42:17 sujith-HP-Pavilion-Notebook systemd[1]: Started An object/document-oriented database.
```

Start Mongo Shell

```
sujith@sujith-HP-Pavilion-Notebook:~$ mongo
MongoDB shell version v3.6.8
connecting to: mongodb://127.0.0.1:27017
Implicit session: session { "id" : UUID("df26a154-5aee-43bf-bb15-7f8093883a82") }
MongoDB server version: 3.6.8
Server has startup warnings:
2020-05-22T14:51:48.968+0530 I STORAGE  [initandlisten]
2020-05-22T14:51:48.968+0530 I STORAGE  [initandlisten] ** WARNING: Using the XFS f
e
2020-05-22T14:51:48.968+0530 I STORAGE  [initandlisten] **          See http://doch
2020-05-22T14:51:52.602+0530 I CONTROL  [initandlisten]
2020-05-22T14:51:52.602+0530 I CONTROL  [initandlisten] ** WARNING: Access control
2020-05-22T14:51:52.602+0530 I CONTROL  [initandlisten] **          Read and write
2020-05-22T14:51:52.602+0530 I CONTROL  [initandlisten]
>
```

Database Status:

```
> show dbs
admin          0.000GB
classroomPlus  0.000GB
config         0.000GB
covidDB        0.000GB
local          0.000GB
test           0.000GB
> use covidDB
switched to db covidDB
```

Collections:

```
> db.getCollectionNames()
[
        "age_group_details_db",
        "covid_india_db",
        "hospital_beds_india_db",
        "icmr_testing_details_db",
        "icmr_testing_labs_db",
        "individual_details_db",
        "population_india_db",
        "statewise_testing_db"
]
```

Converting CSV to DataFrame, Then DataFrame to Dictionaries, Then Dictionaries to JSON:

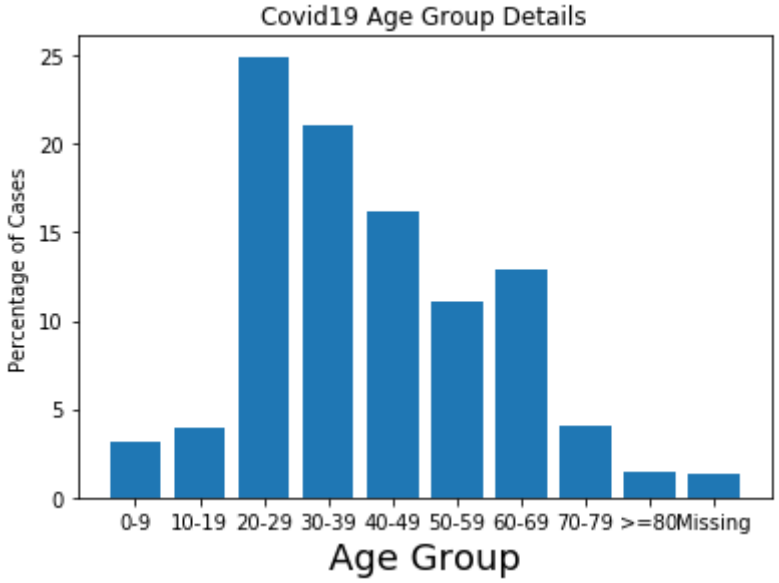
```
import json
age_group_details.to_json('age_group_details.json')                    # saving to json file
jdf = open('age_group_details.json').read()                    # loading the json file
data = json.loads(jdf)
```

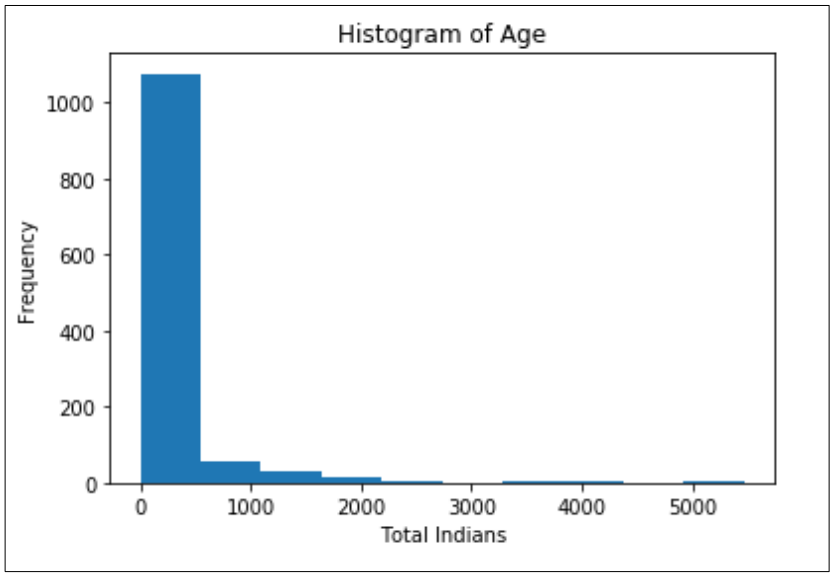
```
def convert(d):
    cols=[]
    for i in d.keys():
        cols.append(i)

    start=cols[0]
    length=len(d[start].keys())
    print(length)
    ans=[]
    for i in range(length):
        temp={}
        for j in cols:
            temp[j]=0
        for j in cols:
            temp[j]=d[j][str(i)]
        ans.append(temp)
    return ans
```

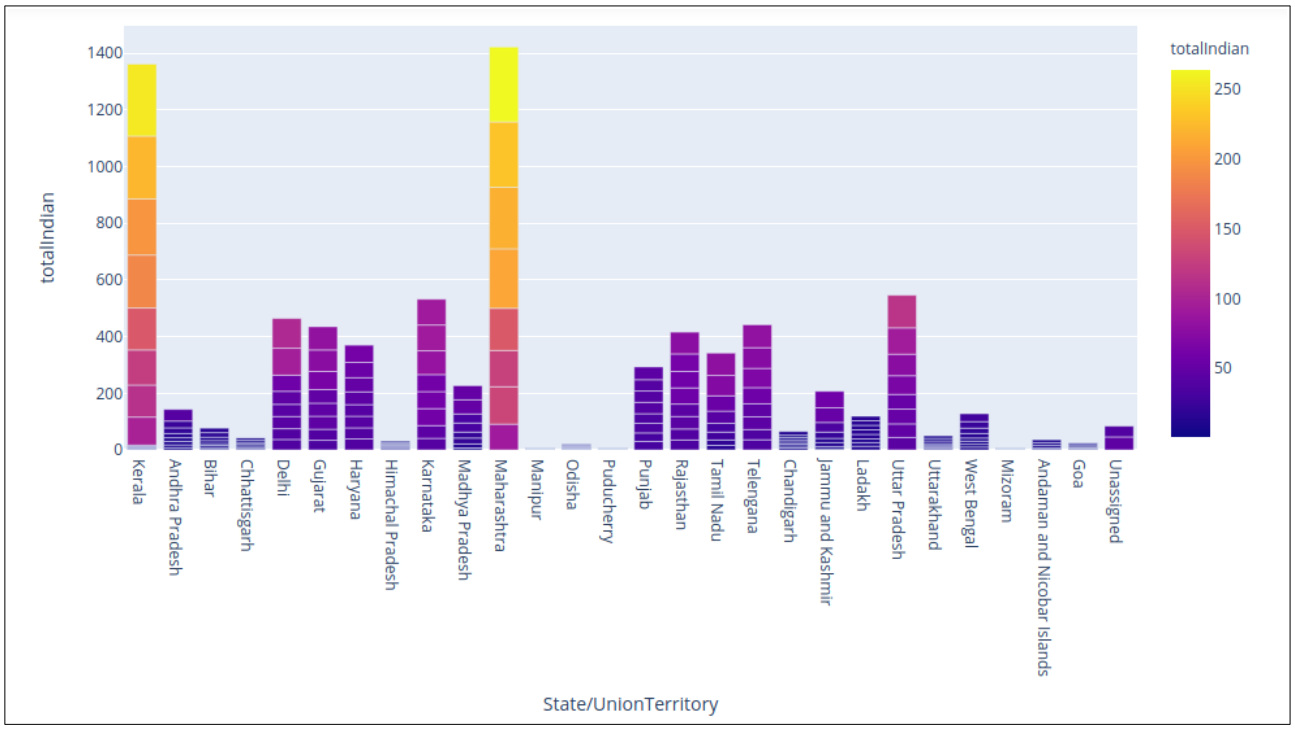
```
c=age_group_details_db.insert_many(convert(data))
print(c.inserted_ids)
```

```
10
[ObjectId('5ecbefa44c76ae568d8ef691'), ObjectId('5ecbefa44c76ae568d8ef692'), ObjectId('5ecbefa44c76ae568d8ef693'),
ObjectId('5ecbefa44c76ae568d8ef694'), ObjectId('5ecbefa44c76ae568d8ef695'), ObjectId('5ecbefa44c76ae568d8ef696'),
```

Analysis of Data




Histogram of Age
Frequency
Total Indians
0
200
400
600
800
1000
0
1000
2000
3000
4000
5000


totalIndian
State/UnionTerritory
0
200
400
600
800
1000
1200
1400
Kerala
Andhra Pradesh
Bihar
Chhattisgarh
Delhi
Gujarat
Haryana
Himachal Pradesh
Karnataka
Madhya Pradesh
Maharashtra
Manipur
Odisha
Puducherry
Punjab
Rajasthan
Tamil Nadu
Telengana
Chandigarh
Jammu and Kashmir
Ladakh
Uttar Pradesh
Uttarakhand
West Bengal
Mizoram
Andaman and Nicobar Islands
Goa
Unassigned
totalIndian
50
100
150
200
250

Querying Data from Database

```
query= { "TotalCases": { "$gt": 20 } }
mydoc = age_group_details_db.find(query)

for x in mydoc:
  print(x)
```

```
{'_id': ObjectId('5ecbefa44c76ae568d8ef691'), 'Sno': 1, 'AgeGroup': '0-9', 'TotalCases': 22, 'percent': 3.18}
{'_id': ObjectId('5ecbefa44c76ae568d8ef692'), 'Sno': 2, 'AgeGroup': '10-19', 'TotalCases': 27, 'percent': 3.9}
{'_id': ObjectId('5ecbefa44c76ae568d8ef693'), 'Sno': 3, 'AgeGroup': '20-29', 'TotalCases': 172, 'percent': 24.86}
{'_id': ObjectId('5ecbefa44c76ae568d8ef694'), 'Sno': 4, 'AgeGroup': '30-39', 'TotalCases': 146, 'percent': 21.1}
{'_id': ObjectId('5ecbefa44c76ae568d8ef695'), 'Sno': 5, 'AgeGroup': '40-49', 'TotalCases': 112, 'percent': 16.18}
{'_id': ObjectId('5ecbefa44c76ae568d8ef696'), 'Sno': 6, 'AgeGroup': '50-59', 'TotalCases': 77, 'percent': 11.13}
{'_id': ObjectId('5ecbefa44c76ae568d8ef697'), 'Sno': 7, 'AgeGroup': '60-69', 'TotalCases': 89, 'percent': 12.86}
{'_id': ObjectId('5ecbefa44c76ae568d8ef698'), 'Sno': 8, 'AgeGroup': '70-79', 'TotalCases': 28, 'percent': 4.05}
```

```
query= { "totalIndian": { "$gt": 1000 } }
mydoc = covid_india_db.find(query)

for x in mydoc:
  print(x)
```

```
{'_id': ObjectId('5ecbf14c4c76ae568d8ef990'), 'Sno': 758, 'Date': '08/04/20', 'Time': '5:00 PM', 'State/UnionTerr
itory': 'Maharashtra', 'ConfirmedIndianNational': '-', 'ConfirmedForeignNational': '-', 'Cured': 79, 'Deaths': 6
4, 'Confirmed': 1018, 'totalIndian': 1161}
{'_id': ObjectId('5ecbf14c4c76ae568d8ef9af'), 'Sno': 789, 'Date': '09/04/20', 'Time': '5:00 PM', 'State/UnionTerr
itory': 'Maharashtra', 'ConfirmedIndianNational': '-', 'ConfirmedForeignNational': '-', 'Cured': 117, 'Deaths': 7
2, 'Confirmed': 1135, 'totalIndian': 1324}
{'_id': ObjectId('5ecbf14c4c76ae568d8ef9ce'), 'Sno': 820, 'Date': '10/04/20', 'Time': '5:00 PM', 'State/UnionTerr
itory': 'Maharashtra', 'ConfirmedIndianNational': '-', 'ConfirmedForeignNational': '-', 'Cured': 125, 'Deaths': 9
7, 'Confirmed': 1364, 'totalIndian': 1586}
{'_id': ObjectId('5ecbf14c4c76ae568d8ef9ed'), 'Sno': 851, 'Date': '11/04/20', 'Time': '5:00 PM', 'State/UnionTerr
itory': 'Maharashtra', 'ConfirmedIndianNational': '-', 'ConfirmedForeignNational': '-', 'Cured': 188, 'Deaths': 1
10, 'Confirmed': 1574, 'totalIndian': 1872}
{'_id': ObjectId('5ecbf14c4c76ae568d8efa01'), 'Sno': 871, 'Date': '12/04/20', 'Time': '5:00 PM', 'State/UnionTerr
itory': 'Delhi', 'ConfirmedIndianNational': '-', 'ConfirmedForeignNational': '-', 'Cured': 25, 'Deaths': 19, 'Con
firmed': 1069, 'totalIndian': 1113}
{'_id': ObjectId('5ecbf14c4c76ae568d8efa0c'), 'Sno': 882, 'Date': '12/04/20', 'Time': '5:00 PM', 'State/UnionTerr
itory': 'Maharashtra', 'ConfirmedIndianNational': '-', 'ConfirmedForeignNational': '-', 'Cured': 208, 'Deaths': 1
27, 'Confirmed': 1761, 'totalIndian': 2096}
{'_id': ObjectId('5ecbf14c4c76ae568d8efa13'), 'Sno': 889, 'Date': '12/04/20', 'Time': '5:00 PM', 'State/UnionTerr
```

```
query= { "TotalPositiveCases": { "$gt": 5000 } }
mydoc = icmr_testing_details_db.find(query)

for x in mydoc:
  print(x)
```

```
{'_id': ObjectId('5ecbf1554c76ae568d8efb7c'), 'SNo': 23, 'DateTime': '08/04/20 21:00', 'TotalSamplesTested': 12791
9.0, 'TotalIndividualsTested': None, 'TotalPositiveCases': 5114.0}
{'_id': ObjectId('5ecbf1554c76ae568d8efb7d'), 'SNo': 24, 'DateTime': '09/04/20 21:00', 'TotalSamplesTested': 14491
0.0, 'TotalIndividualsTested': 130792.0, 'TotalPositiveCases': 5705.0}
{'_id': ObjectId('5ecbf1554c76ae568d8efb7e'), 'SNo': 25, 'DateTime': '10/04/20 21:00', 'TotalSamplesTested': 16133
0.0, 'TotalIndividualsTested': 147034.0, 'TotalPositiveCases': 6872.0}
{'_id': ObjectId('5ecbf1554c76ae568d8efb7f'), 'SNo': 26, 'DateTime': '11/04/20 21:00', 'TotalSamplesTested': 17937
4.0, 'TotalIndividualsTested': 164773.0, 'TotalPositiveCases': 7703.0}
{'_id': ObjectId('5ecbf1554c76ae568d8efb80'), 'SNo': 27, 'DateTime': '12/04/20 21:00', 'TotalSamplesTested': 19574
8.0, 'TotalIndividualsTested': 181028.0, 'TotalPositiveCases': 8312.0}
{'_id': ObjectId('5ecbf1554c76ae568d8efb81'), 'SNo': 28, 'DateTime': '13/04/20 21:00', 'TotalSamplesTested': 21755
4.0, 'TotalIndividualsTested': 202551.0, 'TotalPositiveCases': 9341.0}
{'_id': ObjectId('5ecbf1554c76ae568d8efb82'), 'SNo': 29, 'DateTime': '14/04/20 21:00', 'TotalSamplesTested': 24489
3.0, 'TotalIndividualsTested': 229426.0, 'TotalPositiveCases': 10307.0}
{'_id': ObjectId('5ecbf1554c76ae568d8efb83'), 'SNo': 30, 'DateTime': '15/04/20 21:00', 'TotalSamplesTested': 27459
9.0, 'TotalIndividualsTested': 258730.0, 'TotalPositiveCases': 11297.0}
{'_id': ObjectId('5ecbf1554c76ae568d8efb84'), 'SNo': 31, 'DateTime': '16/04/20 21:00', 'TotalSamplesTested': 30295
6.0, 'TotalIndividualsTested': 286714.0, 'TotalPositiveCases': 12581.0}
{'_id': ObjectId('5ecbf1554c76ae568d8efb85'), 'SNo': 32, 'DateTime': '17/04/20 21:00', 'TotalSamplesTested': 33512
3.0, 'TotalIndividualsTested': 318449.0, 'TotalPositiveCases': 14098.0}
{'_id': ObjectId('5ecbf1554c76ae568d8efb86'), 'SNo': 33, 'DateTime': '18/04/20 21:00', 'TotalSamplesTested': 37212
3.0, 'TotalIndividualsTested': 354969.0, 'TotalPositiveCases': 16365.0}
{'_id': ObjectId('5ecbf1554c76ae568d8efb87'), 'SNo': 34, 'DateTime': '19/04/20 21:00', 'TotalSamplesTested': 40158
6.0, 'TotalIndividualsTested': 383985.0, 'TotalPositiveCases': 17615.0}
```

```
query= { "Population": { "$gt": 100000000 } }
mydoc = population_india_db.find(query)

for x in mydoc:
  print(x)
```

```
{'_id': ObjectId('5ecbf15b4c76ae568d8efc95'), 'Sno': 1, 'State / Union Territory': 'Uttar Pradesh', 'Population':
199812341, 'Rural population': 155317278, 'Urban population': 44495063, 'Area': '240,928\xa0km2\xa0(93,023\xa0sq\x
a0mi)', 'Density': '828/km2\xa0(2,140/sq\xa0mi)', 'Gender Ratio': 912}
{'_id': ObjectId('5ecbf15b4c76ae568d8efc96'), 'Sno': 2, 'State / Union Territory': 'Maharashtra', 'Population': 11
2374333, 'Rural population': 61556074, 'Urban population': 50818259, 'Area': '307,713\xa0km2\xa0(118,809\xa0sq\xa0
mi)', 'Density': '365/km2\xa0(950/sq\xa0mi)', 'Gender Ratio': 929}
{'_id': ObjectId('5ecbf15b4c76ae568d8efc97'), 'Sno': 3, 'State / Union Territory': 'Bihar', 'Population': 10409945
2, 'Rural population': 92341436, 'Urban population': 11758016, 'Area': '94,163\xa0km2\xa0(36,357\xa0sq\xa0mi)', 'D
ensity': '1,102/km2\xa0(2,850/sq\xa0mi)', 'Gender Ratio': 918}
```

Program link:
https://github.com/sujithsagar217/covid19_data_analysis/blob/master/Covid_Data_interpretation.ipynb

Github Repo Link:
https://github.com/sujithsagar217/covid19_data_analysis

References

- *https://specify.io/how-tos/find-documents-in-mongodb-using-the-mongo-shell*
- https://www.w3schools.com/python/python_mongodb_getstarted.asp