original data source (<https://www1.nyc.gov/site/nypd/stats/crime-statistics/citywide-crime-stats.page> )

Steps for setup mongodb and import json file :

1. Install mongodb in mac (google it)
2. Write **mongod** on terminal, it will start the server
3. Open new terminal and write mongo, **mongo shell** will open
4. In mongo shell, you can see and create dbs and collections
5. Now, first create a json file (if your data is not in json) by following code :

Here, I have imported data of complaint through api provide by NYC on its site (

<https://dev.socrata.com/foundry/data.cityofnewyork.us/5uac-w243> )

Now call this API by code provided in the above site.

import pandas as pd

from sodapy import Socrata

import json

import io

client = Socrata("data.cityofnewyork.us", None) // write this once in jupyter notebook to get the data once ( it should be for storing data).

results = client.get("5uac-w243", limit=2000) // taking first 2000 rows only.

Save the this above results( which is json data) into a json file :

with open('complaints.json', 'w') as f:

json.dump(results, f)

1. Now , open terminal and write following to import json file and store in into mongodb

mongoimport --db crime --collection complaint --file complaints.json –jsonArray

1. Now connect your database and fetch record from mongodb in python

#Connection with the database

connection = MongoClient('localhost', 27017)

db = connection.crime

#Handle collections

data = db.complaint

complaintData = data.find()

#fetching particular column collections -> just for analyzing (need to remove later)

data = db.complaint

complaintData = data.find()

complaint\_number =[]

victim\_race =[]

victims\_age=[]

for i in complaintData:

victim\_race.append(i["vic\_race"])

complaint\_number.append((i["cmplnt\_num"]))

victims\_age.append(i["vic\_age\_group"])

#print(complaint\_number)

df1 = pd.DataFrame(complaint\_number)

df1