1. **Create a table attribute dataset and dress dataset**

**import** **mysql.connector** **as** **connection**

**try**:

mydb = connection.connect(host="localhost", database = 'test1', user="root", passwd="mysql",use\_pure=**True**)

cur=mydb.cursor()

query="create table Attribute\_dataset1(Dress\_ID int(15),Style char(10), Price char(10), Rating float(8,2), Size char(10), Season char(10), NeckLine char(20), SleeveLength char(20), waiseline char(20), Material char(20), FabricType char(20), Decoration char(20), PatternType char(20), Recommendation int(2))"

query1="create table dress\_sale(Dress\_ID int(15),29aug2013 int(5), 31aug2013 int(5), 2oct2013 int(5), 4oct2013 int(5), 6oct2013 int(5), 8oct2013 int(5), 10oct2013 int(5), 12oct2013 int(5), 14oct2013 int(5), 16oct2013 int(5), 18oct2013 int(5), 20oct2013 int(5), 22oct2013 int(5), 24oct2013 int(5), 26oct2013 int(5), 28oct2013 int(5), 30oct2013 int(5), 2Nov2013 int(5), 4Nov2013 int(5), 6Nov2013 int(5), 8Nov2010 int(5), 10Nov2013 int(5), 12Nov2013 int(5))"

cur.execute(query)

cur.execute(query1)

print('table created')

mydb.close()

**except** **Exception** **as** e:

print(str(e))

1. **Do a bulk load for these two table for respective dataset ?**

First I have converted excel to csv file.

**import** **pandas** **as** **pd**

df=pd.read\_excel('G:/ineuron/dataset/data fsds/Attribute DataSet.xlsx',encoding='utf8')

df.to\_csv("attribute\_dataset.csv", index=**False**)

df1=pd.read\_excel('G:/ineuron/dataset/data fsds/Dress Sales.xlsx',encoding='utf8')

df1.to\_csv("Dress\_sales.csv",index=**False**)

mydb = connection.connect(host="localhost", database = 'test1', user="root", passwd="mysql",use\_pure=**True**)

cur=mydb.cursor()

ad=pd.read\_csv('attribute Dataset.csv',index\_col=**False**,delimiter=',')

**for** i,row **in** ad.iterrows():

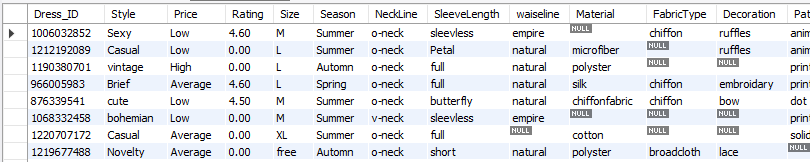
query= "insert into Attribute\_dataset1 (Dress\_ID, Style, Price, Rating, Size, Season, NeckLine, SleeveLength, waiseline, Material, FabricType, Decoration, PatternType, Recommendation) values (**%s**,**%s**,**%s**,**%s**,**%s**,**%s**,**%s**,**%s**,**%s**,**%s**,**%s**,**%s**,**%s**,**%s**)"

cur.execute(query, tuple(row))

print("Record inserted")

mydb.commit()

mydb.close()



mydb = connection.connect(host="localhost", database = 'test1', user="root", passwd="mysql",use\_pure=**True**)

cur=mydb.cursor()

ad=pd.read\_csv('Dress\_sales.csv',index\_col=**False**,delimiter=',')

**for** i,row **in** ad.iterrows():

query= "insert into dress\_sale(Dress\_ID, 29aug2013, 31aug2013, 2oct2013, 4oct2013, 6oct2013, 8oct2013, 10oct2013, 12oct2013, 14oct2013, 16oct2013, 18oct2013, 20oct2013, 22oct2013, 24oct2013, 26oct2013, 28oct2013, 30oct2013, 2Nov2013, 4Nov2013, 6Nov2013, 8Nov2010, 10Nov2013, 12Nov2013) values (**%s**,**%s**,**%s**,**%s**,**%s**,**%s**,**%s**,**%s**,**%s**,**%s**,**%s**,**%s**,**%s**,**%s**,**%s**,**%s**,**%s**,**%s**,**%s**,**%s**,**%s**,**%s**,**%s**,**%s**)"

cur.execute(query, tuple(row))

print("Record inserted")

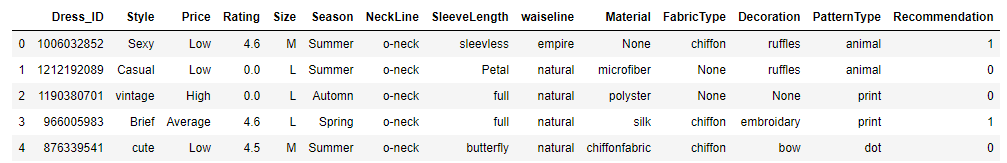
mydb.commit()

mydb.close()

1. **Read these dataset in pandas as a dataframe?**

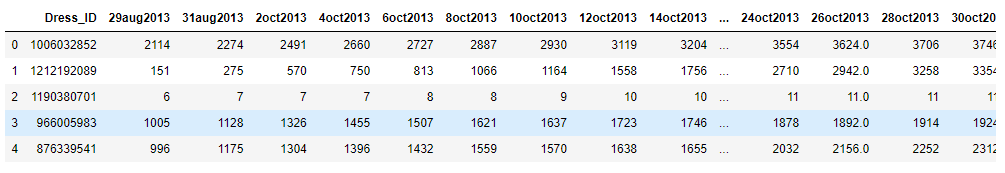
mydb = connection.connect(host="localhost", database = 'test1', user="root", passwd="mysql",use\_pure=**True**)

df=pd.read\_sql('select \* from attribute\_dataset1',mydb)

****

df1=pd.read\_sql('select \* from dress\_sale',mydb)

df1.head()



**4. Convert attribute dataset in json format?**

df.to\_json('attribute\_dataset\_json.json')

**5. Store this dataset into mongodb?**

**import** **pymongo**

**import** **json**

client = pymongo.MongoClient("mongodb+srv://sumit:LJ7ypyaWYLqCedo2@cluster0.cuurr7n.mongodb.net/?retryWrites=true&w=majority")

db = client.task

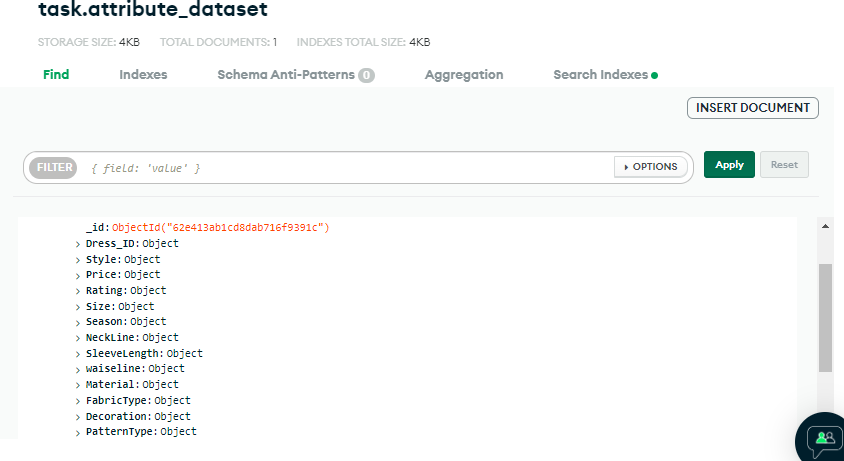
db=client['task']

collection=db['attribute\_dataset']

**with** open('attribute\_dataset\_json.json') **as** file:

file\_data = json.load(file)

collection.insert\_one(file\_data)

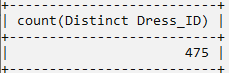


**6. Try to perform left join operation with attribute dataset and dress dataset on column Dress\_ID?**

select \* from attribute\_dataset as ad left join dress\_sale as ds on ad.Dress\_ID=ds.Dress\_ID;

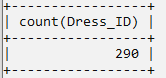
**7. Write a sql query to find out how many unique dress that we have based on dress id**

Select count(Distinct Dress\_ID) from attribute\_dataset1;



**8. Try to find out how mnay dress is having recommendation 0?**

select count(Dress\_ID) from attribute\_dataset1 where Recommendation=0;



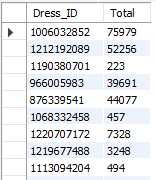
**9. Try to find out total dress sale for individual dress id ?**

select Dress\_ID, (29aug2013 + 31aug2013 + 2oct2013 + 4oct2013 + 6oct2013 + 8oct2013 +

10oct2013 + 12oct2013 + 14oct2013 + 16oct2013 + 18oct2013 + 20oct2013 +

22oct2013 + 24oct2013 + 26oct2013 + 28oct2013 + 30oct2013 + 2Nov2013 +

4Nov2013 + 6Nov2013 + 8Nov2010 + 10Nov2013 + 12Nov2013) as Total from dress\_sale



**10. Try to find out a third highest most selling dress id?**

select Dress\_ID, (29aug2013 + 31aug2013 + 2oct2013 + 4oct2013 + 6oct2013 + 8oct2013 +

10oct2013 + 12oct2013 + 14oct2013 + 16oct2013 + 18oct2013 + 20oct2013 +

22oct2013 + 24oct2013 + 26oct2013 + 28oct2013 + 30oct2013 + 2Nov2013 +

4Nov2013 + 6Nov2013 + 8Nov2010 + 10Nov2013 + 12Nov2013) as Total from dress\_sale order by Total Desc

limit 2,1;

