**Session 13: assignment 1**

**Problem Statement**

Read the following data set:

https://archive.ics.uci.edu/ml/machine-learning-databases/adult/adult.data

Task:

1. Create an sqlalchemy engine using a sample from the data set

2. Write two basic update queries

3. Write two delete queries

4. Write two filter queries

5. Write two function queries

* Task 1: Create an sqlalchemy engine using a sample from the data set

Code:

import random

import numpy as np

import pandas as pd

import sqlite3

from IPython.display import display

### Read data set

adult\_data\_df = pd.read\_csv('https://archive.ics.uci.edu/ml/machine-learning-databases/adult/adult.data')

### Change column names

adult\_data\_df.columns =["age","workclass","fnlwgt","education","education\_num","marital\_status","occupation","relationship","race","sex","capital\_gain","capital\_loss","hours\_per\_week","native\_country","income"]

display(adult\_data\_df.head(5))

###Task 1. Create an sqlalchemy engine using a sample from the data set

import sqlalchemy

from sqlalchemy import create\_engine

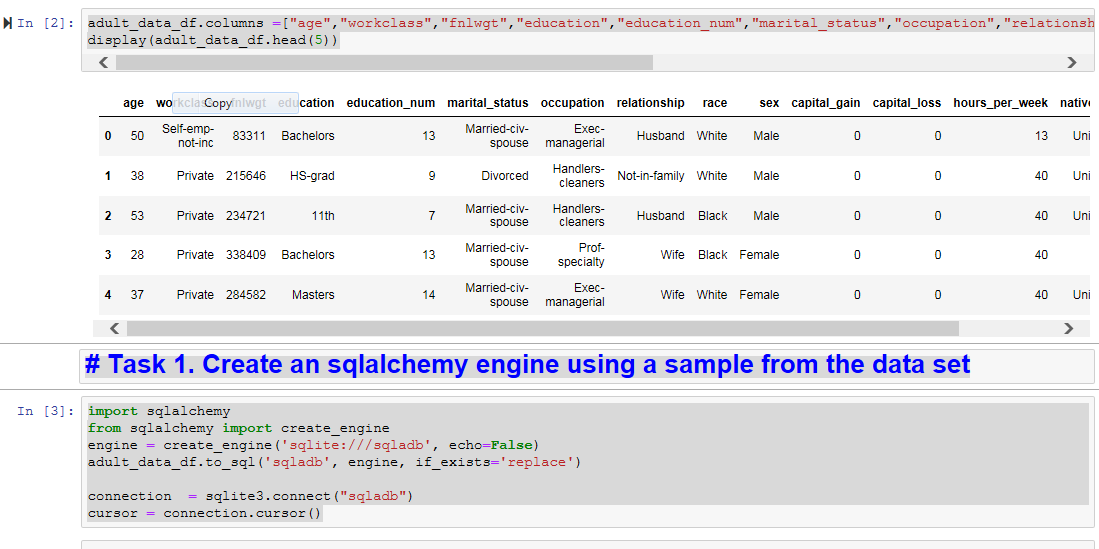
engine = create\_engine('sqlite:///sqladb', echo=False)

adult\_data\_df.to\_sql('sqladb', engine, if\_exists='replace')

connection = sqlite3.connect("sqladb")

cursor = connection.cursor()

Output:



* 2) Write two basic update queries

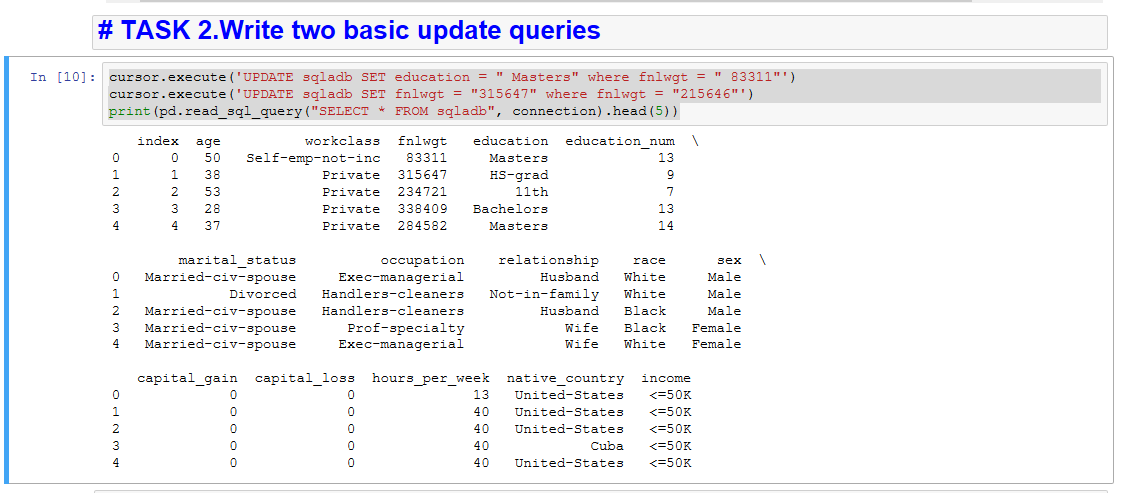
Code:

cursor.execute('UPDATE sqladb SET education = " Masters" where fnlwgt = " 83311"')

cursor.execute('UPDATE sqladb SET fnlwgt = "315647" where fnlwgt = "215646"')

print(pd.read\_sql\_query("SELECT \* FROM sqladb", connection).head(5))

Output:

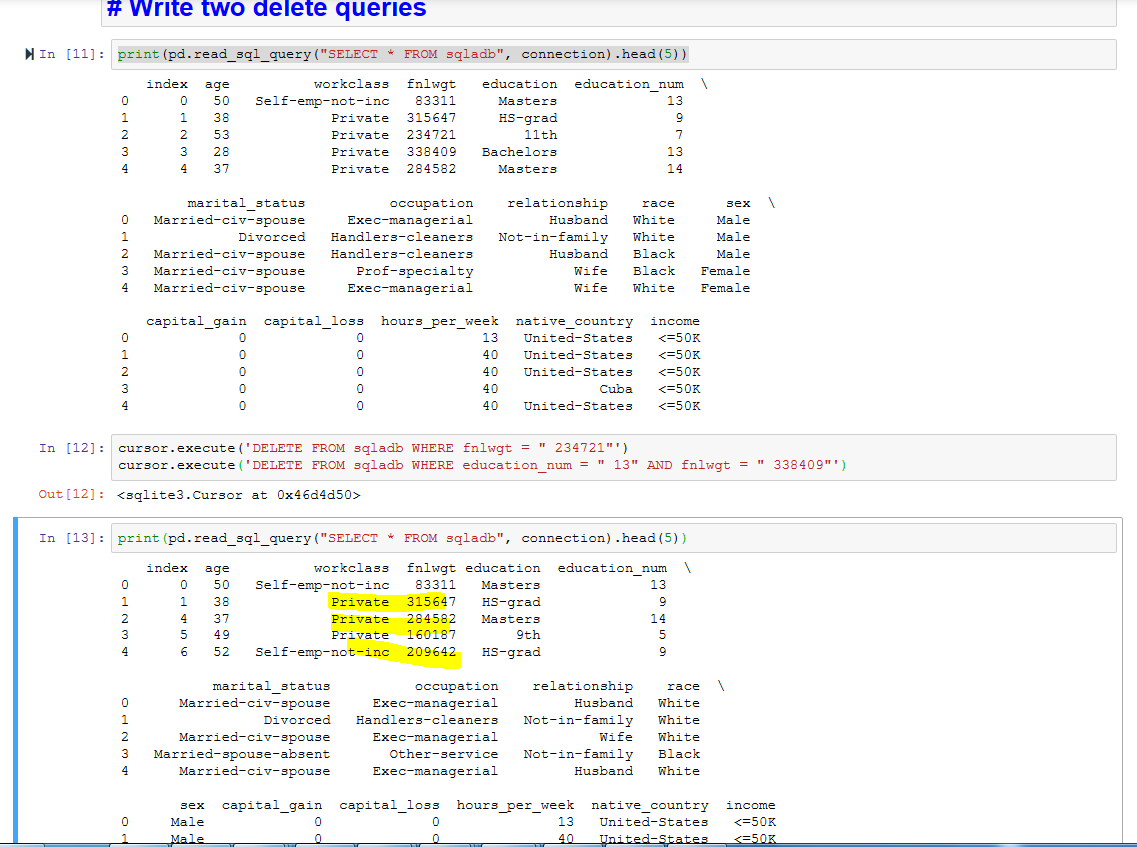


* 3: Write two delete queries

Code:

cursor.execute('DELETE FROM sqladb WHERE fnlwgt = " 234721"')

cursor.execute('DELETE FROM sqladb WHERE education\_num = " 13" AND fnlwgt = " 338409"')



* 4: Write two filter queries

Code:

cursor.execute('SELECT DISTINCT \* FROM sqladb WHERE education = " Masters" AND age < 25 AND fnlwgt = " 166851"')

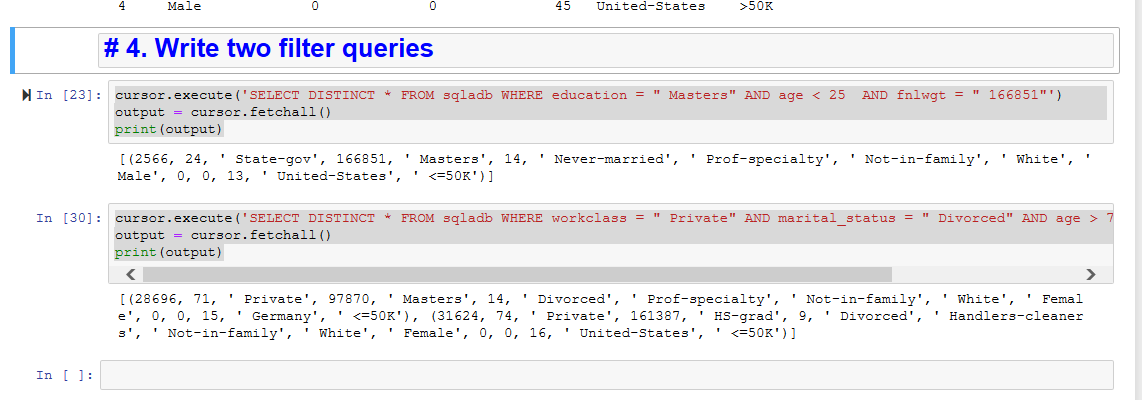
output = cursor.fetchall()

print(output)

cursor.execute('SELECT DISTINCT \* FROM sqladb WHERE workclass = " Private" AND marital\_status = " Divorced" AND age > 70 group by native\_country ')

output = cursor.fetchall()

print(output)



* 5. Write two function queries

Code:

def new\_entry(db\_file, new\_data):

query = "INSERT INTO sqladb VALUES (?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?);"

cursor.execute(query, list(new\_data))

cursor.close()

connection.close()

new\_entry('sqladb',('72190', '54','Private', '123456', 'Masters', '10', 'Married', 'Front\_End', 'In\_Family', 'Indian', 'Male', '0', '0', '40', 'United-States', '>=50K'))

def age\_check():

connection = sqlite3.connect("sqladb")

cursor = connection.cursor()

cursor.execute('SELECT avg(age) FROM sqladb WHERE education=" Masters" AND workclass=" Private"')

output = cursor.fetchall()

print(output)

connection.close()

age\_check()

