

Connect to Mysql using Python (Create,Insert,Read)

- This notebook helps you connect to mysql using python and import the csv file into mysql and create a table and fetch all record from mysql database.

Objective

- In the big data scenarios , we use MySQL as a relational database at times to store transactional data and Sqoop the data into HDFS based environments for further analysis to gain business insights out of the data .We interact with mysql database to store and retrieve data, sometimes we create tables through the csv files

1. Create And Insert CSV To Mysql DataBase

- I m going to insert this csv file in to mysql database by using puthon

Step1: ReadCSV By Pandas

In [1]:

```
import pandas as pd
empdata = pd.read_csv('us-500.csv')
empdata.head()
```

Out[1]:

	first_name	last_name	company_name	address	city	county	state	zip	phone
0	James	Butt	Benton, John B Jr	6649 N Blue Gum St	New Orleans	Orleans	LA	70116	504 621 892
1	Josephine	Darakjy	Chanay, Jeffrey A Esq	4 B Blue Ridge Blvd	Brighton	Livingston	MI	48116	810 292 938
2	Art	Venere	Chemel, James L Cpa	8 W Cerritos Ave #54	Bridgeport	Gloucester	NJ	8014	856 636 874
3	Lenna	Paprocki	Feltz Printing Service	639 Main St	Anchorage	Anchorage	AK	99501	907 385 441
4	Donette	Foller	Printing Dimensions	34 Center St	Hamilton	Butler	OH	45011	513 570 189

Step2: Create Database 'employee'

In [2]:

```
import mysql.connector as mysql
from mysql.connector import Error
try:
    conn = mysql.connect(host='localhost', user='root',password='')#give ur username, passwo
    if conn.is_connected():
        cursor = conn.cursor()
        cursor.execute("CREATE DATABASE employee")
        print("Database is created")
except Error as e:
    print("Error while connecting to MySQL", e)
```

Database is created

Step3: Create a table 'employee_data' & Insert the CSV data into the 'employee_data' table

In [3]:

```
import mysql.connector as mysql
from mysql.connector import Error
try:
    conn = mysql.connect(host='localhost', database='employee', user='root', password='')
    if conn.is_connected():
        cursor = conn.cursor()
        cursor.execute("select database();")
        record = cursor.fetchone()
        print("You're connected to database: ", record)
        cursor.execute('DROP TABLE IF EXISTS employee_data;')
        print('Creating table....')
    # in the below line please pass the create table statement which you want #to create
    query = '''CREATE TABLE employee_data(first_name varchar(255),last_name varchar(255)
    company_name varchar(255),address varchar(255),city varchar(255),county varchar(255)
    zip int,phone1 varchar(255),phone2 varchar(255),
    email varchar(255),web varchar(255))'''
    cursor.execute(query)
    print("Table is created....")
    #Loop through the data frame
    for i,row in empdata.iterrows():
        #here %S means string values
        sql = "INSERT INTO employee.employee_data VALUES (%s,%s,%s,%s,%s,%s,%s,%s,%s,%s,%s,%s)"
        cursor.execute(sql, tuple(row))
        # the connection is not auto committed by default, so we must commit to save ou
        conn.commit()
        print("Record inserted")
except Error as e:
    print("Error while connecting to MySQL", e)
```

You're connected to database: ('employee',)
Creating table....
Table is created....
Record inserted

2. Read That Record Using Python

- Here I am using Two method

1. By mysql
2. By sqlalchemy

Read Method 1 By mysql

In [4]:

```
# !pip install mysql-connector-python  
# !pip install mysql
```

In [5]:

```
import pandas as pd  
import mysql.connector as sql
```

In [6]:

```
username = "root"  
host = "localhost"  
password = ""  
database_name = "employee"
```

In [7]:

```
db_connect = sql.connect(host=host, database=database_name, user = username, password = pas
```

In [8]:

```
df1 = pd.read_sql('SELECT * from employee_data', con=db_connect)
```

In [9]:

```
df1.head()
```

Out[9]:

	first_name	last_name	company_name	address	city	county	state	zip	phone
0	James	Butt	Benton, John B Jr	6649 N Blue Gum St	New Orleans	Orleans	LA	70116	504 621 8921
1	Josephine	Darakjy	Chanay, Jeffrey A Esq	4 B Blue Ridge Blvd	Brighton	Livingston	MI	48116	810 292 9381
2	Art	Venere	Chemel, James L Cpa	8 W Cerritos Ave #54	Bridgeport	Gloucester	NJ	8014	856 636 8741
3	Lenna	Paprocki	Feltz Printing Service	639 Main St	Anchorage	Anchorage	AK	99501	907 385 4411
4	Donette	Foller	Printing Dimensions	34 Center St	Hamilton	Butler	OH	45011	513 570 1891

In [10]:

```
df1.shape
```

Out[10]:

```
(500, 12)
```

Read Method 2 By sqlalchemy

In [11]:

```
# !pip install SQLAlchemy
# !pip install PyMySQL
```

In [12]:

```
import pandas as pd
import sqlalchemy
```

In [13]:

```
engine = sqlalchemy.create_engine('mysql+pymysql://root:@localhost:3306/employee')
```

mysql+pymysql://username:password@host:port/database_name

Read entire table in a dataframe using read_sql_table

In [14]:

```
df2 = pd.read_sql_table('employee_data',engine)
df2.head()
```

Out[14]:

	first_name	last_name	company_name	address	city	county	state	zip	phone
0	James	Butt	Benton, John B Jr	6649 N Blue Gum St	New Orleans	Orleans	LA	70116	504 621 8921
1	Josephine	Darakjy	Chanay, Jeffrey A Esq	4 B Blue Ridge Blvd	Brighton	Livingston	MI	48116	810 292 9381
2	Art	Venere	Chemel, James L Cpa	8 W Cerritos Ave #54	Bridgeport	Gloucester	NJ	8014	856 636 8741
3	Lenna	Paprocki	Feltz Printing Service	639 Main St	Anchorage	Anchorage	AK	99501	907 385 4411
4	Donette	Foller	Printing Dimensions	34 Center St	Hamilton	Butler	OH	45011	513 570 1891

In [15]:

```
df2.shape
```

Out[15]:

```
(500, 12)
```

Read only selected columns

In [16]:

```
df2 = pd.read_sql_table('employee_data',engine,columns=['company_name'])  
df2
```

Out[16]:

	company_name
0	Benton, John B Jr
1	Chanay, Jeffrey A Esq
2	Chemel, James L Cpa
3	Feltz Printing Service
4	Printing Dimensions
...	...
495	Inner Label
496	Hermar Inc
497	Simonton Howe & Schneider Pc
498	Warehouse Office & Paper Prod
499	Affiliated With Travelodge

500 rows × 1 columns

fetch by read_sql_query

In [17]:

```
query = 'SELECT * from employee_data'
df3 = pd.read_sql_query(query,engine)
df3.head()
```

Out[17]:

	first_name	last_name	company_name	address	city	county	state	zip	phone
0	James	Butt	Benton, John B Jr	6649 N Blue Gum St	New Orleans	Orleans	LA	70116	504 621 8921
1	Josephine	Darakjy	Chanay, Jeffrey A Esq	4 B Blue Ridge Blvd	Brighton	Livingston	MI	48116	810 292 9381
2	Art	Venere	Chemel, James L Cpa	8 W Cerritos Ave #54	Bridgeport	Gloucester	NJ	8014	856 636 8741
3	Lenna	Paprocki	Feltz Printing Service	639 Main St	Anchorage	Anchorage	AK	99501	907 385 4411
4	Donette	Foller	Printing Dimensions	34 Center St	Hamilton	Butler	OH	45011	513 570 1891

Now you can do all preprocessing, model building step by using this dataframe. 