Yingna Lu

1.

```python
import pymysql.cursors
import pandas as pd
from sqlalchemy import create_engine, types
connection=pymysql.connect(host='127.0.0.1',
                user='root',
                password='Icedcocoa030.',
                db='hw2',
                cursorclass=pymysql.cursors.DictCursor)

engine = create_engine('mysql+pymysql://root:Icedcocoa030.@localhost/hw2') # enter your pass
col=['Job #','Doc #','Borough','House #','Street Name','Job Type','Job Status','Job Status Descrp','Latest Action Date
data=pd.read_csv('DOB_Job_Application_Filings.csv',usecols=col,nrows=100,encoding='utf8')

sql="create table if not exists example(Job varchar(25),Doc varchar(100),Borough varchar(25),House varchar(10),Street
 Type varchar(100),Job_Status varchar(100),Job_Status_Descrp varchar(100),Latest_Action_Date datetime);"
data.to_sql('example',con=engine,index=False,if_exists='append',chunksize=100) # Replace Table_name with your sql tabl
```

```
inna030030@instance-1:~$ vi problem1.py
inna030030@instance-1:~$ python3 problem1.py
inna030030@instance-1:~$ mysql -u root -p --host 127.0.0.1
Enter password:
Welcome to the MariaDB monitor.  Commands end with ; or \g.
Your MySQL connection id is 105456
Server version: 5.7.25-google-log (Google)

Copyright (c) 2000, 2018, Oracle, MariaDB Corporation Ab and others.

Type 'help;' or '\h' for help. Type '\c' to clear the current input statement.

MySQL [(none)]> use hw2;
Reading table information for completion of table and column names
You can turn off this feature to get a quicker startup with -A

Database changed
MySQL [hw2]> show tables;
+---------------+
| Tables_in_hw2 |
+---------------+
| example       |
| work          |
+---------------+
2 rows in set (0.001 sec)

MySQL [hw2]> select * from example;
```

```
Database changed
MySQL [hw2]> select * from example limit 5;
+-----------+-------+-----------+---------+----------------+----------+------------+-------------------------------+--------------------+
| Job #     | Doc # | Borough   | House # | Street Name    | Job Type | Job Status | Job Status Descrp             | Latest Action Date |
+-----------+-------+-----------+---------+----------------+----------+------------+-------------------------------+--------------------+
| 110083524 |     2 | MANHATTAN | 152     | WEST   57 STREET | A2      | X          | SIGNED OFF                    | 2009-04-08         |
| 110083533 |     1 | MANHATTAN | 444     | EAST   52 STREET | A2      | X          | SIGNED OFF                    | 2009-06-04         |
| 110083542 |     1 | MANHATTAN | 2       | WARDS ISLAND   | A1       | J          | PLAN EXAM - DISAPPROVED       | 2008-02-14         |
| 110083542 |     2 | MANHATTAN | 2       | WARDS ISLAND   | A1       | J          | PLAN EXAM - DISAPPROVED       | 2008-02-14         |
| 110083542 |     3 | MANHATTAN | 2       | WARDS ISLAND   | A1       | D          | APPLICATION PROCESSED - ENTIRE | 2008-02-15        |
+-----------+-------+-----------+---------+----------------+----------+------------+-------------------------------+--------------------+
5 rows in set (0.003 sec)
```

2.

```python
import requests
from io import import StringIO
import pymysql.cursors
import pandas as pd

from sqlalchemy import create_engine,types
connection=pymysql.connect(host='127.0.0.1',
                           user='root',
                           password='Icedcocoa030.',
                           db='hw2',
                           cursorclass=pymysql.cursors.DictCursor)
engine=create_engine('mysql+pymysql://root:Icedcocoa030.@localhost/hw2')
r="https://data.cityofnewyork.us/resource/ic3t-wcy2.csv"


s=requests.get(r).text
col=['job__','doc__','borough','house__','street_name','job_type','job_status','job_status_descrp','latest_action_date']
data=pd.read_csv(StringIO(s),usecols=col,nrows=100)
sql="create table if  not exists problem2(Job varchar(100),Doc varchar(100),Borough varchar(100),house varchar(100),stre
et_name varchar(100),job type varchar(100),job_status varchar(100),job_status_descrp varchar(100),latest_action_date dat
etime);"
data.to_sql('problem2',con=engine,index=False,if_exists='append',chunksize=100)
~
~
```

```
MySQL [hw2]> select * from problem2 limit 5;
+-----------+-------+-----------+---------+-----------------+----------+------------+----------------------------+---------------------+
| job__     | doc__ | borough   | house__ | street_name     | job_type | job_status | job_status_descrp          | latest_action_date  |
+-----------+-------+-----------+---------+-----------------+----------+------------+----------------------------+---------------------+
| 110083524 |     2 | MANHATTAN | 152     | WEST   57 STREET | A2       | X          | SIGNED OFF                 | 2009-04-08          |
| 110083533 |     1 | MANHATTAN | 444     | EAST   52 STREET | A2       | X          | SIGNED OFF                 | 2009-06-04          |
| 110083542 |     1 | MANHATTAN | 2       | WARDS ISLAND    | A1       | J          | PLAN EXAM - DISAPPROVED    | 2008-02-14          |
| 110083542 |     2 | MANHATTAN | 2       | WARDS ISLAND    | A1       | J          | PLAN EXAM - DISAPPROVED    | 2008-02-14          |
| 110083542 |     3 | MANHATTAN | 2       | WARDS ISLAND    | A1       | D          | APPLICATION PROCESSED - ENTIRE | 2008-02-15      |
+-----------+-------+-----------+---------+-----------------+----------+------------+----------------------------+---------------------+
5 rows in set (0.003 sec)
```

3.

```python
import requests
from io import StringIO
import pymysql.cursors
import pandas as pd
from sqlalchemy import create_engine,types
connection=pymysql.connect(host='127.0.0.1',
                           user='root',
                           password='Icedcocoa030.',
                           db='hw2',
                           cursorclass=pymysql.cursors.DictCursor)
engine=create_engine('mysql+pymysql://root:Icedcocoa030.@localhost/hw2')
r="https://data.cityofnewyork.us/resource/ic3t-wcy2.csv"
s=requests.get(r).text
col=['job__','doc__','borough','house__','street_name','job_type','job_status','job_status_descrp','landmarked','adult_estab','latest_action_date','existing_occupancy','propos
ed_occupancy','owner_type','job_description','initial_cost']
df=pd.read_csv(StringIO(s),usecols=col)
df['initial_cost']=df['initial_cost'].str.replace('$','')
df['initial_cost']=pd.to_numeric(df['initial_cost'])
df['initial_cost']=(df.initial_cost/1000).apply(lambda x:'${:,.2f}'.format(x))


df=df.dropna(subset=['job_description','landmarked','adult_estab','initial_cost'])
df=pd.DataFrame(df)

sql="create table if not exists problem3(Job# varchar(100),doc varchar(100),borough varchar(100),house varchar(100),streetname varchar(100),jobtype varchar(100),jobstatus varc
har(100),job_status_descrp varchar(100),landmard varchar(100),adult_estab varchar(100),latest_action_data datetime,existing_occupancy varchar(100),proposed_occuoancy varchar(1
00),owner_type varchar(100),jobdescr varchar(100),initial_cost varchar(100));"
df.to_sql('problem3',con=engine,index=False,if_exists='append',chunksize=100)
```

```
Database changed
MySQL [hw2]> select * from problem3 limit 5;
+-----------+-------+-----------+---------+-----------------+----------+------------+--------------------------------+---------------------+------------+-------------+--------
------+--------------------+--------------------+-------------+----------------------------------------------------------------------------------------------+
| job__     | doc__ | borough   | house__ | street_name     | job_type | job_status | job_status_descrp              | latest_action_date  | landmarked | adult_estab | initia
l_cost | existing_occupancy | proposed_occupancy | owner_type  | job_description                                                                              |
+-----------+-------+-----------+---------+-----------------+----------+------------+--------------------------------+---------------------+------------+-------------+--------
------+--------------------+--------------------+-------------+----------------------------------------------------------------------------------------------+
| 110083524 |     2 | MANHATTAN | 152     | WEST   57 STREET | A2       | X          | SIGNED OFF                     | 2009-04-08          | N          | N           | $160.0
       | NULL               | NULL               | PARTNERSHIP | Mechanical and Plumbing modifications as indicated on plans filed herewith. No change in use, egress or occupa
cy.    |
| 110083533 |     1 | MANHATTAN | 444     | EAST   52 STREET | A2       | X          | SIGNED OFF                     | 2009-06-04          | N          | N           | $100.0
       | RES                | NULL               | CORPORATION | Apartment renovation work to include partition, ceiling, door and plumbing work as indicated on plans filed he
ewith. No change in use, egress or occupancy. |
| 110083542 |     1 | MANHATTAN | 2       | WARDS ISLAND    | A1       | J          | PLAN EXAM - DISAPPROVED        | 2008-02-14          | N          | N           | $400.0
       | J-1                | J-1                | OTHER       | Herewith filing alteration type 1 to change use.                                            |
| 110083542 |     2 | MANHATTAN | 2       | WARDS ISLAND    | A1       | J          | PLAN EXAM - DISAPPROVED        | 2008-02-14          | N          | N           | $300.0
       | NULL               | NULL               | OTHER       | Filing herewith  Plumbing and Mechanical work in conjunction with ALT 1                      |
| 110083542 |     3 | MANHATTAN | 2       | WARDS ISLAND    | A1       | D          | APPLICATION PROCESSED - ENTIRE | 2008-02-15          | N          | N           | $0.00
       | NULL               | NULL               | CORPORATION | FILING HEREWITH, STRUCTURAL WORK                                                             |
+-----------+-------+-----------+---------+-----------------+----------+------------+--------------------------------+---------------------+------------+-------------+--------
------+--------------------+--------------------+-------------+----------------------------------------------------------------------------------------------+
```