create_sql_file_from_dataset

March 9, 2023

[3]:
```
import pandas as pd
```

[4]:
```
d = {'col1': [1, 2], 'col2': [3, 4]}
df = pd.DataFrame(data=d)
```

[ ]:
```
df.head()
```

[ ]:
```
   col1  col2
0     1     3
1     2     4
```

[ ]:
```
df=pd.read_csv('test.csv')
```

[ ]:
```
df.head(10)
```

[ ]:
```
    x          y
0  77  79.775152
1  21  23.177279
2  22  25.609262
3  20  17.857388
4  36  41.849864
```

[47]:
```
sql_table="Data"
sql_create_table=pd.io.sql.get_schema(df,name=sql_table)
print(sql_create_table)
```

```
CREATE TABLE "Data" (
"x" INTEGER,
  "y" REAL
)
```

[73]:
```
file=open("dataframe.sql",'a')
file.write(sql_create_table+";\n")
file.close()
file=open("dataframe.sql",'a')
for i in range(0,300):
    sql_cmd=""" INSERT INTO {SQL_TABLE} VALUES ({X},{Y})""".
 ↪format(SQL_TABLE=sql_table,X=df['x'][i],Y=df['y'][i])+";\n"
```

```
    file.write(sql_cmd)
file.close()
```

[37]:
```
df.head(5)
```

[37]:
```
    x          y
0  77  79.775152
1  21  23.177279
2  22  25.609262
3  20  17.857388
4  36  41.849864
```

[74]:
```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 300 entries, 0 to 299
Data columns (total 2 columns):
 #   Column  Non-Null Count  Dtype  
---  ------  --------------  -----  
 0   x       300 non-null    int64  
 1   y       300 non-null    float64
dtypes: float64(1), int64(1)
memory usage: 4.8 KB
```