MySQL 导入CSV

直接导入 LOAD DATA INFILE

```
-- 显示local_infile这个全局变量的值,用于确定是否允许导入本地数据文件
SHOW global variables like 'local_infile';

-- 将local_infile变量设置为1,允许导入本地数据文件
SET global local_infile=1;

SHOW variables like '%secure%';


USE my_project;

CREATE TABLE IF NOT EXISTS starbucks(
uniqueid         INT           NOT NULL PRIMARY KEY,
gender           VARCHAR(255),
age              INT,
id               VARCHAR(255),
became_member_on TEXT,
income           INT
);

LOAD DATA INFILE 'C:/ProgramData/MySQL/MySQL Server 8.0/Uploads/profile.csv' INTO TABLE starbucks
FIELDS TERMINATED BY ','
LINES TERMINATED BY '\n'
IGNORE 1 LINES
(uniqueid, gender, age, id, became_member_on, @income)
SET income = NULLIF(@income, '');
-- 将income字段设置为@income,除非@income为空,在这种情况下将income设置为NULL。
```

利用pandas导入

1. 利用python处理csv文件

1. 导入csv文件

In [1]:
```
import pandas as pd
```

In [4]:
```
df = pd.read_csv('Bike Sales.csv')

df.head()
```

Out[4]:

	Date	Day	Month	Year	Customer_Age	Age_Group	Customer_Gender	Country	State	Product_Category	Sub_Category	Product	Order_Quantity	U
0	2013-11-26	26	November	2013	19	Youth (<25)	M	Canada	British Columbia	Accessories	Bike Racks	Hitch Rack - 4-Bike	8	
1	2015-11-26	26	November	2015	19	Youth (<25)	M	Canada	British Columbia	Accessories	Bike Racks	Hitch Rack - 4-Bike	8	
2	2014-03-23	23	March	2014	49	Adults (35-64)	M	Australia	New South Wales	Accessories	Bike Racks	Hitch Rack - 4-Bike	23	
3	2016-03-23	23	March	2016	49	Adults (35-64)	M	Australia	New South Wales	Accessories	Bike Racks	Hitch Rack - 4-Bike	20	
4	2014-05-15	15	May	2014	47	Adults (35-64)	F	Australia	New South Wales	Accessories	Bike Racks	Hitch Rack - 4-Bike	4	

2. 数据清理

2.1 删除无用列

In [5]:
```
cols = ['Day', 'Month', 'Year', 'Age_Group']

df.drop(cols, inplace=True, axis=1)

df.head()
```

Out[5]:

	Date	Customer_Age	Customer_Gender	Country	State	Product_Category	Sub_Category	Product	Order_Quantity	Unit_Cost	Unit_Price	Profit	Cost
0	2013-11-26	19	M	Canada	British Columbia	Accessories	Bike Racks	Hitch Rack - 4-Bike	8	45	120	590	360
1	2015-11-26	19	M	Canada	British Columbia	Accessories	Bike Racks	Hitch Rack - 4-Bike	8	45	120	590	360
2	2014-03-23	49	M	Australia	New South Wales	Accessories	Bike Racks	Hitch Rack - 4-Bike	23	45	120	1366	1035
3	2016-03-23	49	M	Australia	New South Wales	Accessories	Bike Racks	Hitch Rack - 4-Bike	20	45	120	1188	900
4	2014-05-15	47	F	Australia	New South Wales	Accessories	Bike Racks	Hitch Rack - 4-Bike	4	45	120	238	180

2.2 检查缺失值

In [6]:
```
df.isnull().sum()
```

Out[6]:
```
Date                0
Customer_Age        0
Customer_Gender     0
Country             0
State               0
Product_Category    0
Sub_Category        0
Product             0
Order_Quantity      0
Unit_Cost           0
Unit_Price          0
Profit              0
Cost                0
Revenue             0
dtype: int64
```

2.3 检查数据类型

In [8]:
```
df.dtypes
```

Out[8]:
```
Date                object
Customer_Age         int64
Customer_Gender     object
Country             object
State               object
Product_Category    object
Sub_Category        object
Product             object
Order_Quantity       int64
Unit_Cost            int64
Unit_Price           int64
Profit               int64
Cost                 int64
Revenue              int64
dtype: object
```

2.4 转换数据类型

In [9]:
```
df['Date'] = pd.to_datetime(df['Date'])
```

In [10]:
```
df = df.astype({'Unit_Cost':'float', 'Unit_Price':'float', 'Profit':'float', 'Cost':'float', 'Revenue':'float'})
```

3 导出数据文件

3.1 转换所有行为元组

In [12]:
```
y = []

for i in range(len(df)):
    x = tuple(df.iloc[i])
    y.append(x)

y
```

Out[12]:
```
[(Timestamp('2013-11-26 00:00:00'),
  19,
  'M',
  'Canada',
  'British Columbia',
  'Accessories',
  'Bike Racks',
  'Hitch Rack - 4-Bike',
  8,
  45.0,
  120.0,
  590.0,
  360.0,
  950.0),
 (Timestamp('2015-11-26 00:00:00'),
  19,
  'M',
  'Canada',
  'British Columbia',
```

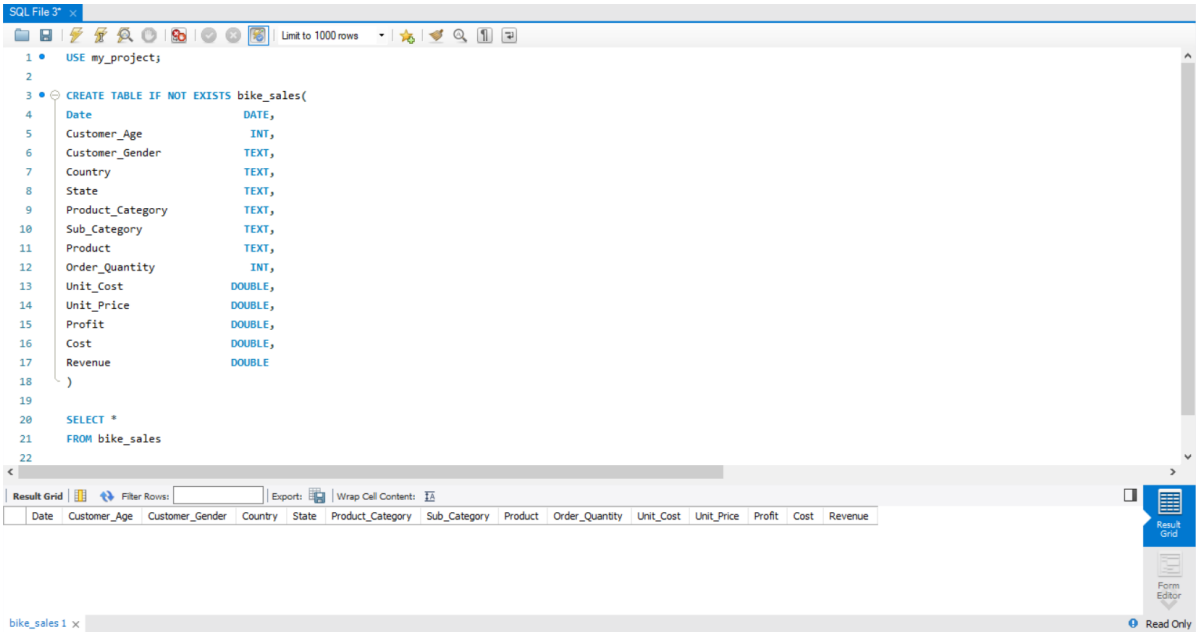
3.2 将元组保存为txt文件

In [13]:
```
file = open('Bike Sales.txt', 'w')

for tuple in y:
    file.write(str(tuple) + ',' + '\n')

file.close()
```

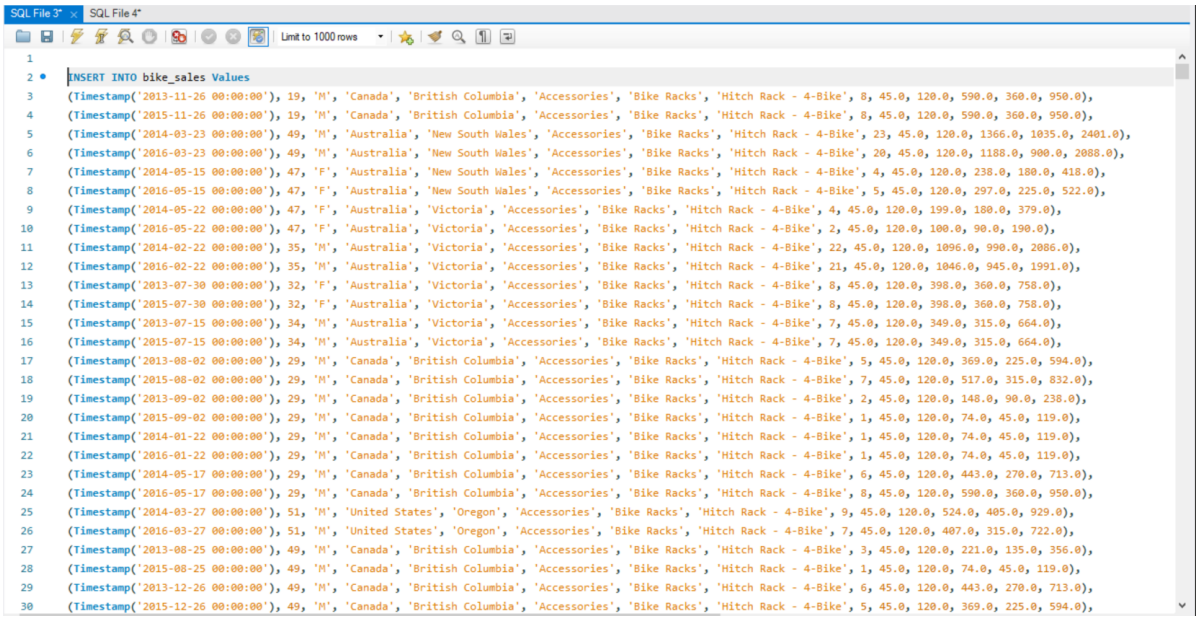
In [17]:
```
df.columns
```

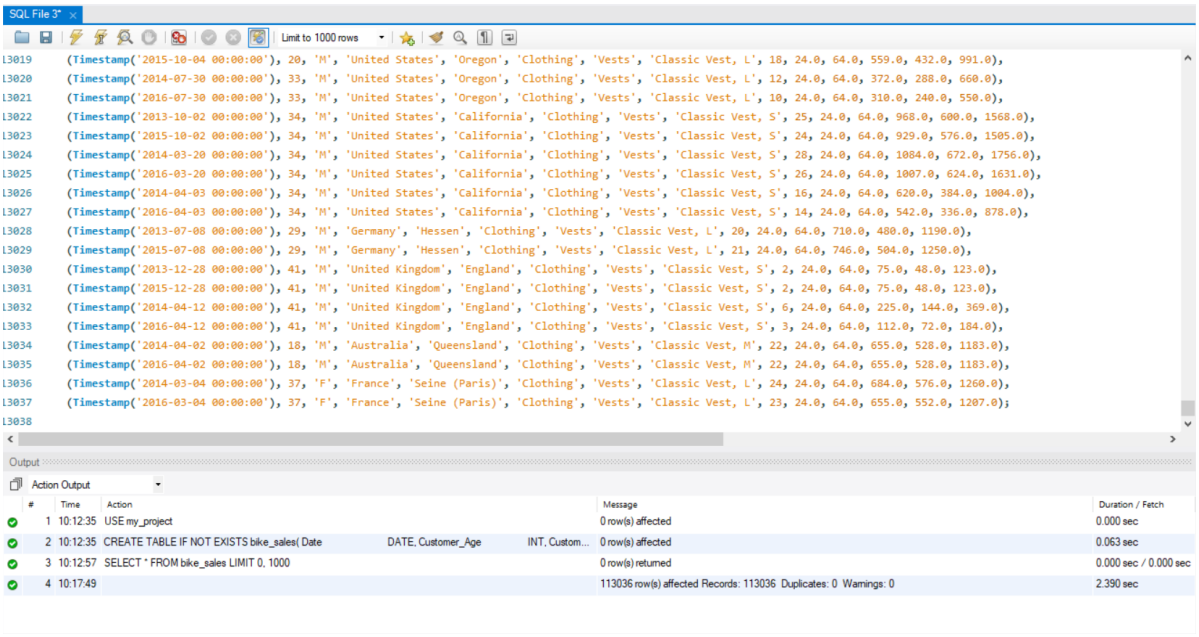
Out[17]:
```
Index(['Date', 'Customer_Age', 'Customer_Gender', 'Country', 'State',
       'Product_Category', 'Sub_Category', 'Product', 'Order_Quantity',
       'Unit_Cost', 'Unit_Price', 'Profit', 'Cost', 'Revenue'],
      dtype='object')
```

2. MySQL创建表格并插入数据

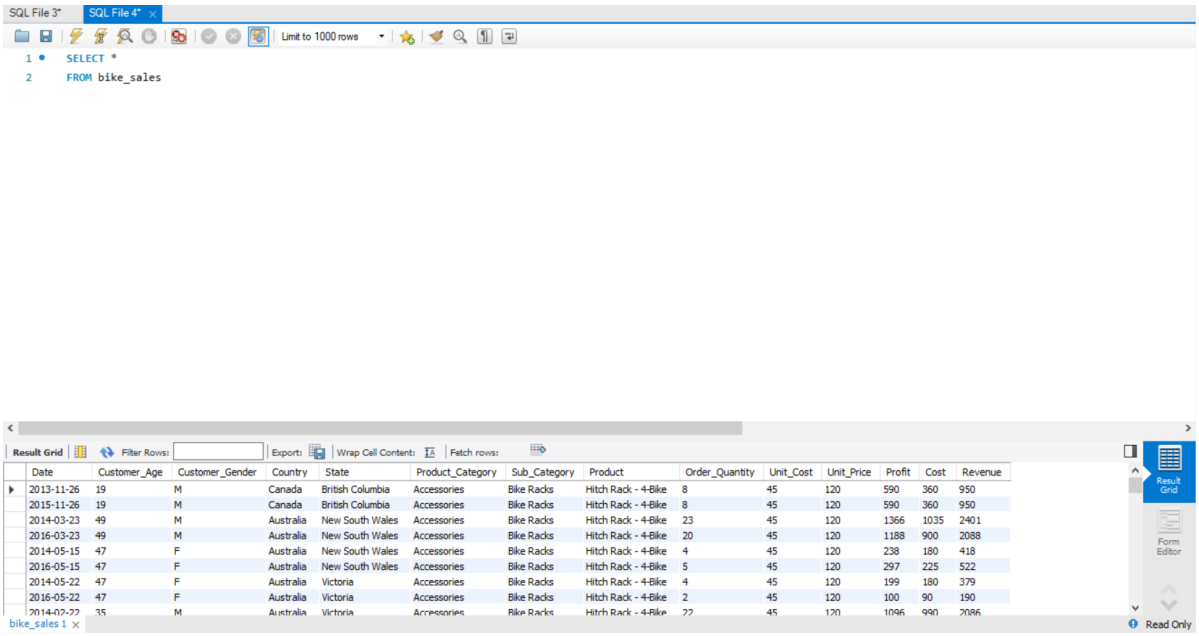


3. 直接将生成的txt文件一键复制粘贴




SQL File 3*
Limit to 1000 rows
(Timestamp('2015-10-04 00:00:00'), 20, 'M', 'United States', 'Oregon', 'Clothing', 'Vests', 'Classic Vest, L', 18, 24.0, 64.0, 559.0, 432.0, 991.0),
(Timestamp('2014-07-30 00:00:00'), 33, 'M', 'United States', 'Oregon', 'Clothing', 'Vests', 'Classic Vest, L', 12, 24.0, 64.0, 372.0, 288.0, 660.0),
(Timestamp('2016-07-30 00:00:00'), 33, 'M', 'United States', 'Oregon', 'Clothing', 'Vests', 'Classic Vest, L', 10, 24.0, 64.0, 310.0, 240.0, 550.0),
(Timestamp('2013-10-02 00:00:00'), 34, 'M', 'United States', 'California', 'Clothing', 'Vests', 'Classic Vest, S', 25, 24.0, 64.0, 968.0, 600.0, 1568.0),
(Timestamp('2015-10-02 00:00:00'), 34, 'M', 'United States', 'California', 'Clothing', 'Vests', 'Classic Vest, S', 24, 24.0, 64.0, 929.0, 576.0, 1505.0),
(Timestamp('2014-03-20 00:00:00'), 34, 'M', 'United States', 'California', 'Clothing', 'Vests', 'Classic Vest, S', 28, 24.0, 64.0, 1084.0, 672.0, 1756.0),
(Timestamp('2016-03-20 00:00:00'), 34, 'M', 'United States', 'California', 'Clothing', 'Vests', 'Classic Vest, S', 26, 24.0, 64.0, 1007.0, 624.0, 1631.0),
(Timestamp('2014-04-03 00:00:00'), 34, 'M', 'United States', 'California', 'Clothing', 'Vests', 'Classic Vest, S', 16, 24.0, 64.0, 620.0, 384.0, 1004.0),
(Timestamp('2016-04-03 00:00:00'), 34, 'M', 'United States', 'California', 'Clothing', 'Vests', 'Classic Vest, S', 14, 24.0, 64.0, 542.0, 336.0, 878.0),
(Timestamp('2013-07-08 00:00:00'), 29, 'M', 'Germany', 'Hessen', 'Clothing', 'Vests', 'Classic Vest, L', 20, 24.0, 64.0, 710.0, 480.0, 1190.0),
(Timestamp('2015-07-08 00:00:00'), 29, 'M', 'Germany', 'Hessen', 'Clothing', 'Vests', 'Classic Vest, L', 21, 24.0, 64.0, 746.0, 504.0, 1250.0),
(Timestamp('2013-12-28 00:00:00'), 41, 'M', 'United Kingdom', 'England', 'Clothing', 'Vests', 'Classic Vest, S', 2, 24.0, 64.0, 75.0, 48.0, 123.0),
(Timestamp('2015-12-28 00:00:00'), 41, 'M', 'United Kingdom', 'England', 'Clothing', 'Vests', 'Classic Vest, S', 2, 24.0, 64.0, 75.0, 48.0, 123.0),
(Timestamp('2014-04-12 00:00:00'), 41, 'M', 'United Kingdom', 'England', 'Clothing', 'Vests', 'Classic Vest, S', 6, 24.0, 64.0, 225.0, 144.0, 369.0),
(Timestamp('2016-04-12 00:00:00'), 41, 'M', 'United Kingdom', 'England', 'Clothing', 'Vests', 'Classic Vest, S', 3, 24.0, 64.0, 112.0, 72.0, 184.0),
(Timestamp('2014-04-02 00:00:00'), 18, 'M', 'Australia', 'Queensland', 'Clothing', 'Vests', 'Classic Vest, M', 22, 24.0, 64.0, 655.0, 528.0, 1183.0),
(Timestamp('2016-04-02 00:00:00'), 18, 'M', 'Australia', 'Queensland', 'Clothing', 'Vests', 'Classic Vest, M', 22, 24.0, 64.0, 655.0, 528.0, 1183.0),
(Timestamp('2014-03-04 00:00:00'), 37, 'F', 'France', 'Seine (Paris)', 'Clothing', 'Vests', 'Classic Vest, L', 24, 24.0, 64.0, 684.0, 576.0, 1260.0),
(Timestamp('2016-03-04 00:00:00'), 37, 'F', 'France', 'Seine (Paris)', 'Clothing', 'Vests', 'Classic Vest, L', 23, 24.0, 64.0, 655.0, 552.0, 1207.0);
Output
Action Output
Time Action Message Duration / Fetch
1 10:12:35 USE my_project 0 row(s) affected 0.000 sec
2 10:12:35 CREATE TABLE IF NOT EXISTS bike_sales(Date DATE, Customer_Age INT, Custom... 0 row(s) affected 0.063 sec
3 10:12:57 SELECT * FROM bike_sales LIMIT 0, 1000 0 row(s) returned 0.000 sec / 0.000 sec
4 10:17:49 113036 row(s) affected Records: 113036 Duplicates: 0 Warnings: 0 2.390 sec

4. 导入成功


SQL File 3*
SQL File 4*
Limit to 1000 rows
SELECT *
FROM bike_sales
Result Grid
Filter Rows:
Export:
Wrap Cell Content:
Fetch rows:
Date Customer_Age Customer_Gender Country State Product_Category Sub_Category Product Order_Quantity Unit_Cost Unit_Price Profit Cost Revenue
2013-11-26 19 M Canada British Columbia Accessories Bike Racks Hitch Rack - 4-Bike 8 45 120 590 360 950
2015-11-26 19 M Canada British Columbia Accessories Bike Racks Hitch Rack - 4-Bike 8 45 120 590 360 950
2014-03-23 49 M Australia New South Wales Accessories Bike Racks Hitch Rack - 4-Bike 23 45 120 1366 1035 2401
2016-03-23 49 M Australia New South Wales Accessories Bike Racks Hitch Rack - 4-Bike 20 45 120 1188 900 2088
2014-05-15 47 F Australia New South Wales Accessories Bike Racks Hitch Rack - 4-Bike 4 45 120 238 180 418
2016-05-15 47 F Australia New South Wales Accessories Bike Racks Hitch Rack - 4-Bike 5 45 120 297 225 522
2014-05-22 47 F Australia Victoria Accessories Bike Racks Hitch Rack - 4-Bike 4 45 120 199 180 379
2016-05-22 47 F Australia Victoria Accessories Bike Racks Hitch Rack - 4-Bike 2 45 120 100 90 190
bike_sales 1
Result Grid
Form Editor
Read Only