

Lab 2 - Spark SQL Lab

- Họ và tên:
- MSSV:

1. Thiết lập SQLSession

Chú ý: Để thực thi được các đoạn mã kết nối đến Northwind SQLite database ở bên dưới bạn cần đặt file `sqlite-jdbc-3.50.3.0.jar` cùng thư mục với file `.ipynb`.

- Tham khảo: <https://spark.apache.org/docs/latest/sql-data-sources-jdbc.html>

```
In [1]: import findspark
findspark.init()

import pyspark
findspark.find()

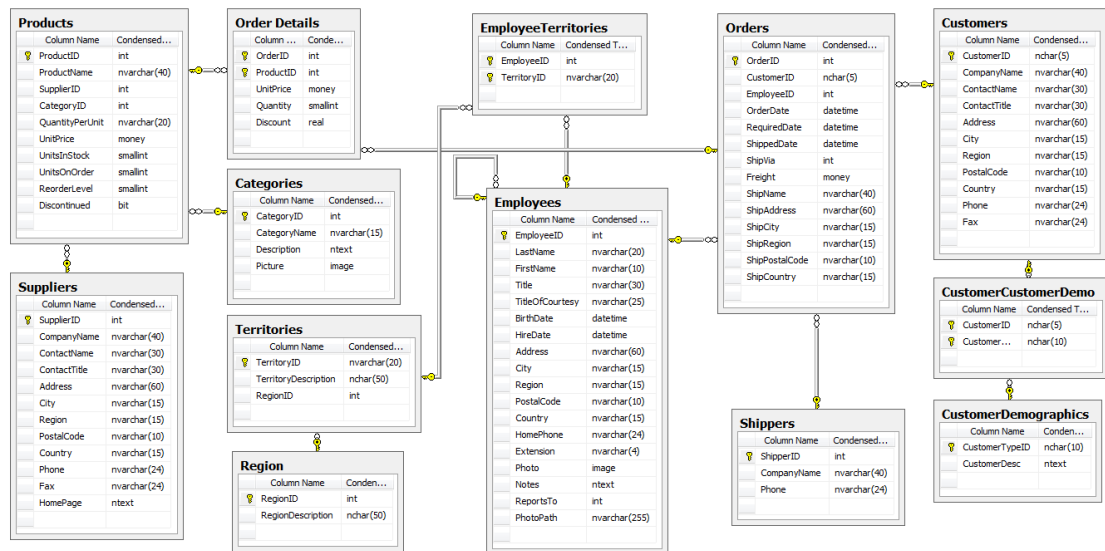
import os
from pyspark.sql import SparkSession

spark = (SparkSession.builder
        .master("local")
        .appName("SQLite JDBC")
        .config(
            "spark.jars",
            "{}sqlite-jdbc-3.50.3.0.jar".format(os.getcwd()))
        .config(
            "spark.driver.extraClassPath",
            "{}sqlite-jdbc-3.50.3.0.jar".format(os.getcwd()))
        .getOrCreate())
```

2. Tạo các dataframe từ các table

Ở bài lab này ta sẽ làm việc với CSDL Northwind. Ta sẽ kết nối đến Northwind SQLite database và tạo ra các dataframe từ các table của CSDL này.

Bạn có thể xem lược đồ CSDL Northwind ở hình sau.



Chú ý: tên bảng và một số thuộc tính trong hình **không giống hoàn toàn** với CSDL trong file **Northwind.sqlite**.

```
In [2]: emp = spark.read.format('jdbc').\
        options(url='jdbc:sqlite:Northwind.sqlite',\
        dbtable='employee',driver='org.sqlite.JDBC').load()
```

```
In [3]: emp.printSchema()
```

```
root
|-- Id: integer (nullable = true)
|-- LastName: string (nullable = true)
|-- FirstName: string (nullable = true)
|-- Title: string (nullable = true)
|-- TitleOfCourtesy: string (nullable = true)
|-- BirthDate: string (nullable = true)
|-- HireDate: string (nullable = true)
|-- Address: string (nullable = true)
|-- City: string (nullable = true)
|-- Region: string (nullable = true)
|-- PostalCode: string (nullable = true)
|-- Country: string (nullable = true)
|-- HomePhone: string (nullable = true)
|-- Extension: string (nullable = true)
|-- Photo: binary (nullable = true)
|-- Notes: string (nullable = true)
|-- ReportsTo: integer (nullable = true)
|-- PhotoPath: string (nullable = true)
```

```
In [4]: ord = spark.read.format('jdbc').\
        options(url='jdbc:sqlite:Northwind.sqlite',\
        dbtable='[order]',driver='org.sqlite.JDBC').load()
```

```
In [5]: ord.printSchema()
```

```

root
|-- Id: integer (nullable = true)
|-- CustomerId: string (nullable = true)
|-- EmployeeId: integer (nullable = true)
|-- OrderDate: string (nullable = true)
|-- RequiredDate: string (nullable = true)
|-- ShippedDate: string (nullable = true)
|-- ShipVia: integer (nullable = true)
|-- Freight: decimal(38,18) (nullable = true)
|-- ShipName: string (nullable = true)
|-- ShipAddress: string (nullable = true)
|-- ShipCity: string (nullable = true)
|-- ShipRegion: string (nullable = true)
|-- ShipPostalCode: string (nullable = true)
|-- ShipCountry: string (nullable = true)

```

```

In [6]: cat = spark.read.format('jdbc').\
        options(url='jdbc:sqlite:Northwind.sqlite',\
               dbtable='category',driver='org.sqlite.JDBC').load()

```

```

In [7]: cat.printSchema()

```

```

root
|-- Id: integer (nullable = true)
|-- CategoryName: string (nullable = true)
|-- Description: string (nullable = true)

```

```

In [8]: cus = spark.read.format('jdbc').\
        options(url='jdbc:sqlite:Northwind.sqlite',\
               dbtable='customer',driver='org.sqlite.JDBC').load()

```

```

In [9]: cus.printSchema()

```

```

root
|-- Id: string (nullable = true)
|-- CompanyName: string (nullable = true)
|-- ContactName: string (nullable = true)
|-- ContactTitle: string (nullable = true)
|-- Address: string (nullable = true)
|-- City: string (nullable = true)
|-- Region: string (nullable = true)
|-- PostalCode: string (nullable = true)
|-- Country: string (nullable = true)
|-- Phone: string (nullable = true)
|-- Fax: string (nullable = true)

```

```

In [10]: od = spark.read.format('jdbc').\
         options(url='jdbc:sqlite:Northwind.sqlite',\
                dbtable='orderdetail',driver='org.sqlite.JDBC').load()

```

```

In [11]: od.printSchema()

```

```
root
|-- Id: string (nullable = true)
|-- OrderId: integer (nullable = true)
|-- ProductId: integer (nullable = true)
|-- UnitPrice: decimal(38,18) (nullable = true)
|-- Quantity: integer (nullable = true)
|-- Discount: double (nullable = true)
```

```
In [12]: reg = spark.read.format('jdbc').\
         options(url='jdbc:sqlite:Northwind.sqlite',\
         dbtable='region',driver='org.sqlite.JDBC').load()
```

```
In [13]: reg.printSchema()
```

```
root
|-- Id: integer (nullable = true)
|-- RegionDescription: string (nullable = true)
```

```
In [14]: sp = spark.read.format('jdbc').\
         options(url='jdbc:sqlite:Northwind.sqlite',\
         dbtable='shipper',driver='org.sqlite.JDBC').load()
```

```
In [15]: sp.printSchema()
```

```
root
|-- Id: integer (nullable = true)
|-- CompanyName: string (nullable = true)
|-- Phone: string (nullable = true)
```

```
In [16]: sup = spark.read.format('jdbc').\
         options(url='jdbc:sqlite:Northwind.sqlite',\
         dbtable='supplier',driver='org.sqlite.JDBC').load()
```

```
In [17]: sup.printSchema()
```

```
root
|-- Id: integer (nullable = true)
|-- CompanyName: string (nullable = true)
|-- ContactName: string (nullable = true)
|-- ContactTitle: string (nullable = true)
|-- Address: string (nullable = true)
|-- City: string (nullable = true)
|-- Region: string (nullable = true)
|-- PostalCode: string (nullable = true)
|-- Country: string (nullable = true)
|-- Phone: string (nullable = true)
|-- Fax: string (nullable = true)
|-- HomePage: string (nullable = true)
```

```
In [18]: ter = spark.read.format('jdbc').\
         options(url='jdbc:sqlite:Northwind.sqlite',\
         dbtable='territory',driver='org.sqlite.JDBC').load()
```

```
In [19]: ter.printSchema()
```

```
root
|-- Id: string (nullable = true)
|-- TerritoryDescription: string (nullable = true)
|-- RegionId: integer (nullable = true)
```

```
In [20]: empter = spark.read.format('jdbc').\
         options(url='jdbc:sqlite:Northwind.sqlite',\
         dbtable='employeeterritory',driver='org.sqlite.JDBC').load()
```

```
In [21]: empter.printSchema()
```

```
root
|-- Id: string (nullable = true)
|-- EmployeeId: integer (nullable = true)
|-- TerritoryId: string (nullable = true)
```

```
In [22]: prod = spark.read.format('jdbc').\
         options(url='jdbc:sqlite:Northwind.sqlite',\
         dbtable='product',driver='org.sqlite.JDBC').load()
```

```
In [23]: prod.printSchema()
```

```
root
|-- Id: integer (nullable = true)
|-- ProductName: string (nullable = true)
|-- SupplierId: integer (nullable = true)
|-- CategoryId: integer (nullable = true)
|-- QuantityPerUnit: string (nullable = true)
|-- UnitPrice: decimal(38,18) (nullable = true)
|-- UnitsInStock: integer (nullable = true)
|-- UnitsOnOrder: integer (nullable = true)
|-- ReorderLevel: integer (nullable = true)
|-- Discontinued: integer (nullable = true)
```

3. Tạo các temp view cho các dataframe

```
In [24]: cat.createOrReplaceTempView("category")
         cus.createOrReplaceTempView("customer")
         emp.createOrReplaceTempView("employee")
         empter.createOrReplaceTempView("employeeterritory")
         ord.createOrReplaceTempView("order")
         od.createOrReplaceTempView("orderdetail")
         prod.createOrReplaceTempView("product")
         reg.createOrReplaceTempView("region")
         sp.createOrReplaceTempView("shipper")
         sup.createOrReplaceTempView("supplier")
         ter.createOrReplaceTempView("territory")
```

4. Xem vài dòng đầu của các table

```
In [25]: query = """
         select *
         from category
         """
```

```
spark.sql(query).show(n=5, truncate=10)
```

```
+---+-----+-----+
| Id|CategoryName|Description|
+---+-----+-----+
| 1|   Beverages| Soft dr...|
| 2|  Condiments| Sweet a...|
| 3|  Confect...| Dessert...|
| 4|  Dairy P...|   Cheeses|
| 5|  Grains/...| Breads,...|
+---+-----+-----+
only showing top 5 rows
```

```
In [26]: query = """
        select *
        from customer
        """
        spark.sql(query).show(n=5, truncate=10)
```

```
+-----+-----+-----+-----+-----+-----+-----+
+-----+-----+-----+-----+-----+-----+-----+
|  Id|CompanyName|ContactName|ContactTitle|  Address|      City|      Region|Post
alCode|Country|      Phone|      Fax|
+-----+-----+-----+-----+-----+-----+-----+
+-----+-----+-----+-----+-----+-----+-----+
|ALFKI| Alfreds...| Maria A...|  Sales R...|Obere S...|    Berlin|Western...|
12209|Germany|030-007...|030-007...|
|ANATR| Ana Tru...| Ana Tru...|      Owner|Avda. d...|México ...|Central...|
05021| Mexico|(5) 555...|(5) 555...|
|ANTON| Antonio...| Antonio...|      Owner|Matader...|México ...|Central...|
05023| Mexico|(5) 555...|      NULL|
|AROUT| Around ...| Thomas ...|  Sales R...|120 Han...|    London|British...|  W
A1 1DP|    UK|(171) 5...|(171) 5...|
|BERGS| Berglun...| Christi...|  Order A...|Berguvs...|    Luleå|Norther...|
S-958 22| Sweden|0921-12...|0921-12...|
+-----+-----+-----+-----+-----+-----+-----+
+-----+-----+-----+-----+-----+-----+-----+
only showing top 5 rows
```

```
In [27]: query = """
        select *
        from employee
        """
        spark.sql(query).show(n=5, truncate=10)
```

```

+---+-----+-----+-----+-----+-----+-----+
+-----+-----+-----+-----+-----+-----+-----+
+-----+-----+-----+
| Id| LastName|FirstName|      Title|TitleOfCourtesy| BirthDate| HireDate| Addr
ess|      City|      Region|PostalCode|Country| HomePhone|Extension|Photo|      Notes|
ReportsTo| PhotoPath|
+---+-----+-----+-----+-----+-----+-----+
+-----+-----+-----+-----+-----+-----+-----+
+-----+-----+-----+
| 1|  Davolio|    Nancy|Sales R...|      Ms.|1980-12-08|2024-05-01|507 -
2...| Seattle|North A...|    98122|    USA|(206) 5...|    5467| NULL|Educat
i...|      2|http://...|
| 2|  Fuller|   Andrew|Vice Pr...|      Dr.|1984-02-19|2024-08-14|908 W.
...| Tacoma|North A...|    98401|    USA|(206) 5...|    3457| NULL|Andrew ...|
NULL|http://...|
| 3|Leverling|   Janet|Sales R...|      Ms.|1995-08-30|2024-04-01|722 Mo
s...|Kirkland|North A...|    98033|    USA|(206) 5...|    3355| NULL|Janet
h...|      2|http://...|
| 4| Peacock| Margaret|Sales R...|      Mrs.|1969-09-19|2025-05-03|4110 O
l...| Redmond|North A...|    98052|    USA|(206) 5...|    5176| NULL|Margar
e...|      2|http://...|
| 5| Buchanan|  Steven|Sales M...|      Mr.|1987-03-04|2025-10-17|14 Gar
r...| London|British...|   SW1 8JR|    UK|(71) 55...|    3453| NULL|Steven
...|      2|http://...|
+---+-----+-----+-----+-----+-----+-----+
+-----+-----+-----+-----+-----+-----+-----+
+-----+-----+-----+
only showing top 5 rows

```

```

In [28]: query = """
        select *
        from employeeterritory
        """
        spark.sql(query).show(n=5, truncate=10)

```

```

+-----+-----+-----+
|      Id|EmployeeId|TerritoryId|
+-----+-----+-----+
|1/06897|          1|      06897|
|1/19713|          1|      19713|
|2/01581|          2|      01581|
|2/01730|          2|      01730|
|2/01833|          2|      01833|
+-----+-----+-----+
only showing top 5 rows

```

```

In [29]: query = """
        select *
        from order
        """
        spark.sql(query).show(n=5, truncate=10)

```

```

+-----+-----+-----+-----+-----+-----+-----+
+-----+-----+-----+-----+-----+-----+-----+
+-----+
| Id|CustomerId|EmployeeId| OrderDate|RequiredDate|ShippedDate|ShipVia|   Freigh
ht|  ShipName|ShipAddress|  ShipCity|ShipRegion|ShipPostalCode|ShipCountry|
+-----+-----+-----+-----+-----+-----+-----+
+-----+-----+-----+-----+-----+-----+-----+
+-----+
|10248|      VINET|          5|2012-07-04|  2012-08-01| 2012-07-16|      3|32.380
0...|Vins et...| 59 rue ...|    Reims|Western...|    51100|    France|
|10249|      TOMSP|          6|2012-07-05|  2012-08-16| 2012-07-10|      1|11.610
0...|Toms Sp...| Luisens...|   Münster|Western...|    44087|    Germany|
|10250|      HANAR|          4|2012-07-08|  2012-08-05| 2012-07-12|      2|65.830
0...|Hanari ...| Rua do ...|Rio de ...|South A...|    05454-876|    Brazil|
|10251|      VICTE|          3|2012-07-08|  2012-08-05| 2012-07-15|      1|41.340
0...|Victuai...| 2, rue ...|    Lyon|Western...|    69004|    France|
|10252|      SUPRD|          4|2012-07-09|  2012-08-06| 2012-07-11|      2|51.300
0...|Suprême...| Bouleva...| Charleroi|Western...|    B-6000|    Belgium|
+-----+-----+-----+-----+-----+-----+-----+
+-----+-----+-----+-----+-----+-----+-----+
+-----+
only showing top 5 rows

```

```

In [34]: query = """
        select *
        from orderdetail
        """
spark.sql(query).show(n=5, truncate=10)

```

```

+-----+-----+-----+-----+-----+-----+
|      Id|OrderId|ProductId| UnitPrice|Quantity|Discount|
+-----+-----+-----+-----+-----+-----+
|10248/11|  10248|      11|14.0000...|    12|    0.0|
|10248/42|  10248|      42|9.80000...|    10|    0.0|
|10248/72|  10248|      72|34.8000...|     5|    0.0|
|10249/14|  10249|      14|18.6000...|     9|    0.0|
|10249/51|  10249|      51|42.4000...|    40|    0.0|
+-----+-----+-----+-----+-----+-----+
only showing top 5 rows

```

```

In [35]: query = """
        select *
        from product
        """
spark.sql(query).show(n=5, truncate=10)

```



```

+---+-----+-----+-----+-----+-----+-----+
+-----+-----+-----+
| Id|ProductName|SupplierId|CategoryId|QuantityPerUnit| UnitPrice|UnitsInStock|Un
itsOnOrder|ReorderLevel|Discontinued|
+---+-----+-----+-----+-----+-----+-----+
+-----+-----+-----+
| 1|      Chai|      1|      1|    10 boxe...|18.0000...|      39|
0|      10|      0|
| 2|      Chang|      1|      1|    24 - 12...|19.0000...|      17|
40|      25|      0|
| 3| Aniseed...|      1|      2|    12 - 55...|10.0000...|      13|
70|      25|      0|
| 4| Chef An...|      2|      2|    48 - 6 ...|22.0000...|      53|
0|      0|      0|
| 5| Chef An...|      2|      2|    36 boxes|21.3500...|       0|
0|      0|      1|
+---+-----+-----+-----+-----+-----+-----+
+-----+-----+-----+
only showing top 5 rows

```

```

In [32]: query = """
          select *
          from region
          """
          spark.sql(query).show(n=5, truncate=10)

```

```

+---+-----+-----+
| Id|RegionDescription|
+---+-----+-----+
| 1|      Eastern|
| 2|      Western|
| 3|      Northern|
| 4|      Southern|
+---+-----+-----+

```

```

In [33]: query = """
          select *
          from shipper
          """
          spark.sql(query).show(n=5, truncate=10)

```

```

+---+-----+-----+
| Id|CompanyName|      Phone|
+---+-----+-----+
| 1| Speedy ...|(503) 5...|
| 2| United ...|(503) 5...|
| 3| Federal...|(503) 5...|
+---+-----+-----+

```

```

In [36]: query = """
          select *
          from supplier
          """
          spark.sql(query).show(n=5, truncate=10)

```

```

+---+-----+-----+-----+-----+-----+-----+
+-----+-----+-----+-----+-----+
| Id|CompanyName|ContactName|ContactTitle| Address| City| Region|Postal
Code|Country| Phone| Fax| HomePage|
+---+-----+-----+-----+-----+-----+-----+
+-----+-----+-----+-----+-----+
| 1| Exotic ...| Charlot...| Purchas...|49 Gilb...| London|British...| EC1
4SD| UK|(171) 5...| NULL| NULL|
| 2| New Orl...| Shelley...| Order A...|P.O. Bo...|New Orl...|North A...| 7
0117| USA|(100) 5...| NULL|#CAJUN...|
| 3| Grandma...| Regina ...| Sales R...|707 Oxf...| Ann Arbor|North A...| 4
8104| USA|(313) 5...|(313) 5...| NULL|
| 4| Tokyo T...| Yoshi N...| Marketi...|9-8 Sek...| Tokyo|Eastern...|
100| Japan|(03) 35...| NULL| NULL|
| 5| Coopera...| Antonio...| Export ...|Calle d...| Oviedo|Souther...| 3
3007| Spain|(98) 59...| NULL| NULL|
+---+-----+-----+-----+-----+-----+-----+
+-----+-----+-----+-----+-----+
only showing top 5 rows

```

```

In [37]: query = """
        select *
        from territory
        """
        spark.sql(query).show(n=5, truncate=10)

```

```

+---+-----+-----+
| Id|TerritoryDescription|RegionId|
+---+-----+-----+
|01581| Westboro| 1|
|01730| Bedford| 1|
|01833| Georgetow| 1|
|02116| Boston| 1|
|02139| Cambridge| 1|
+---+-----+-----+
only showing top 5 rows

```

5. Câu hỏi

Viết truy vấn SQL để trả lời các câu hỏi sau đồng thời thực thi các truy vấn đó để hiển thị kết quả.

Q1. Liệt kê name , address , city , và region của các employee sống ở 'USA' .

In []:

```

+-----+-----+-----+-----+-----+
|firstname| lastname|          address|    city|    region|
+-----+-----+-----+-----+-----+
|   Nancy|  Davolio|507 - 20th Ave. E...| Seattle|North America|
|  Andrew|   Fuller| 908 W. Capital Way| Tacoma|North America|
|   Janet|Leverling| 722 Moss Bay Blvd.|Kirkland|North America|
| Margaret| Peacock|4110 Old Redmond Rd.| Redmond|North America|
|   Laura| Callahan|4726 - 11th Ave. ...| Seattle|North America|
+-----+-----+-----+-----+-----+

```

Q2. Liệt kê title và name của các employee đã bán ít nhất một product có tên 'Gravad Lax' hoặc 'Mishi Kobe Niku' .

In []:

```

+-----+-----+
|          Title|  EmployeeName|
+-----+-----+
|Sales Representative|    Robert King|
|Vice President, S...|  Andrew Fuller|
|Sales Representative| Anne Dodsworth|
|Sales Representative|Janet Leverling|
+-----+-----+

```

Q3. Liệt kê name và title của các employee và name và title của các employee khác mà họ phải báo cáo (ReportsTo) (hoặc null nếu họ không phải báo cáo với ai).

In []:

```

+-----+-----+-----+-----+-----+
+-----+
|EmployeeID|      EmployeeTitle|      EmployeeName|ManagerId|      ManagerTitle|
ManagerName|
+-----+-----+-----+-----+
+-----+
|          2|Vice President, S...|      Andrew Fuller|      NULL|      NULL|
NULL|
|          6|Sales Representative|      Michael Suyama|          5|      Sales Manager|
Steven Buchanan|
|          7|Sales Representative|          Robert King|          5|      Sales Manager|
Steven Buchanan|
|          9|Sales Representative|      Anne Dodsworth|          5|      Sales Manager|
Steven Buchanan|
|          1|Sales Representative|      Nancy Davolio|          2|Vice President, S...|
Andrew Fuller|
|          3|Sales Representative|      Janet Leverling|          2|Vice President, S...|
Andrew Fuller|
|          4|Sales Representative|Margaret Peacock|          2|Vice President, S...|
Andrew Fuller|
|          5|      Sales Manager|      Steven Buchanan|          2|Vice President, S...|
Andrew Fuller|
|          8|Inside Sales Coor...|      Laura Callahan|          2|Vice President, S...|
Andrew Fuller|
+-----+-----+-----+-----+
+-----+

```

Q4. Liệt kê customer name, product name và supplier name cho các customer sống ở 'London' và các supplier có tên 'Pavlova, Ltd.' hoặc 'Karkki Oy'.

In []:

```

+-----+-----+-----+
|      CompanyName|      ProductName|      CompanyName|
+-----+-----+-----+
|      North/South|      Outback Lager|Pavlova, Ltd.|
|Eastern Connection|          Maxilaku|      Karkki Oy|
|      Around the Horn|Valkoinen suklaa|      Karkki Oy|
|      B's Beverages|      Outback Lager|Pavlova, Ltd.|
|      Around the Horn|      Outback Lager|Pavlova, Ltd.|
|Seven Seas Imports|          Maxilaku|      Karkki Oy|
|Seven Seas Imports|          Pavlova|Pavlova, Ltd.|
|Seven Seas Imports|      Alice Mutton|Pavlova, Ltd.|
|      B's Beverages|          Maxilaku|      Karkki Oy|
+-----+-----+-----+

```

Q5. Liệt kê name của các employee có tuổi lớn hơn **một employee nào đó** sống ở 'London'.

In []:

```

+-----+-----+
|FirstName| LastName|
+-----+-----+
|    Nancy|  Davolio|
|  Andrew|   Fuller|
|   Janet|Leverling|
|Margaret|  Peacock|
|  Steven| Buchanan|
| Michael|   Suyama|
|  Robert|    King|
|   Laura| Callahan|
+-----+-----+

```

Q6. Liệt kê name của các employee có tuổi lớn hơn **tất cả** employee sống ở 'London' .

In []:

```

+-----+-----+
|FirstName| LastName|
+-----+-----+
|    Nancy|  Davolio|
|  Andrew|   Fuller|
|Margaret|  Peacock|
+-----+-----+

```

Q7. Liệt kê name của các employee làm việc lâu năm hơn **tất cả** employee ở 'London' .

In []:

```

+-----+-----+
|FirstName| LastName|
+-----+-----+
|    Nancy|  Davolio|
|  Andrew|   Fuller|
|   Janet|Leverling|
|Margaret|  Peacock|
+-----+-----+

```

Q8. Liệt kê name của các employee và city nơi họ sống đối với các employee đã bán hàng cho các customer ở cùng thành phố với họ.

In []:

```

+-----+-----+-----+
|FirstName| LastName| City|
+-----+-----+-----+
| Nancy| Davolio|Seattle|
| Steven| Buchanan| London|
| Anne|Dodsworth| London|
| Michael| Suyama| London|
| Laura| Callahan|Seattle|
| Robert| King| London|
+-----+-----+-----+

```

Q9. Liệt kê name của các customer không đặt bất cứ product nào.

In []:

```

+-----+
| ContactName|
+-----+
|Antonio Moreno|
|Janine Labruno|
| Diego Roel|
|Yvonne Moncada|
|Marie Bertrand|
|Lúcia Carvalho|
+-----+

```

Q10. Liệt kê name của các category và giá trung bình (average price) của các product trong mỗi category .

In []:

```

+-----+-----+
| CategoryName| AvgUnitPrice|
+-----+-----+
| Beverages|37.979166666666666...|
| Condiments|23.062500000000000...|
| Confections|25.160000000000000...|
| Dairy Products|28.730000000000000...|
| Grains/Cereals|20.250000000000000...|
| Meat/Poultry|54.006666666666666...|
| Produce|32.370000000000000...|
| Seafood|20.682500000000000...|
+-----+-----+

```

Q11. Liệt kê id và name của các supplier cung cấp hơn 3 product .

In []:

Id	CompanyName
2	New Orleans Cajun...
7	Pavlova, Ltd.
8	Specialty Biscuit...
12	Plutzer Lebensmit...

Q12. Liệt kê id , name , và số lượng order của các employee sắp xếp theo employee id .

In []:

Id	FirstName	LastName	OrderNum
1	Nancy	Davolio	123
2	Andrew	Fuller	96
3	Janet	Leverling	127
4	Margaret	Peacock	156
5	Steven	Buchanan	42
6	Michael	Suyama	67
7	Robert	King	72
8	Laura	Callahan	104
9	Anne	Dodsworth	43

Q13. Với mỗi employee , liệt kê id , name , và số loại product khác nhau họ đã bán, sắp xếp employee id .

In []:

Id	FirstName	LastName	ProductNum
1	Nancy	Davolio	72
2	Andrew	Fuller	68
3	Janet	Leverling	74
4	Margaret	Peacock	75
5	Steven	Buchanan	52
6	Michael	Suyama	57
7	Robert	King	67
8	Laura	Callahan	70
9	Anne	Dodsworth	53

Q14. Liệt kê id , name , và tổng doanh thu (sau khi trừ đi chiết khấu - Discount) của mỗi employee , sắp xếp theo employee id .

In []:

```
+---+-----+-----+-----+
| Id|FirstName| LastName|      TotalSales|
+---+-----+-----+-----+
| 1|   Nancy|   Davolio|192107.60449999993|
| 2|  Andrew|    Fuller|   166537.755|
| 3|   Janet|Leverling|   202812.843|
| 4| Margaret|   Peacock|232890.84599999996|
| 5|   Steven|Buchanan|   68792.2825|
| 6| Michael|   Suyama|   73913.1295|
| 7|   Robert|    King|   124568.235|
| 8|   Laura|Callahan|126862.27750000004|
| 9|    Anne|Dodsworth|   77308.0665|
+---+-----+-----+-----+
```

Q15. Liệt kê id, name, và tổng doanh thu (có trừ đi chiết khấu - Discount) của các employee đã bán được hơn 70 product khác nhau, sắp xếp theo employee id.

In []:

```
+---+-----+-----+-----+
| Id|FirstName| LastName|      TotalSales|
+---+-----+-----+-----+
| 1|   Nancy|   Davolio|192107.60450000007|
| 3|   Janet|Leverling|202812.84299999996|
| 4| Margaret|   Peacock|232890.84600000005|
+---+-----+-----+-----+
```

Q16. Liệt kê name của các employee đã bán các product của hơn 28 supplier.

In []:

```
+---+-----+-----+
| ID|FirstName| LastName|
+---+-----+-----+
| 8|   Laura|Callahan|
| 3|   Janet|Leverling|
| 5|   Steven|Buchanan|
| 7|   Robert|    King|
| 4| Margaret|   Peacock|
| 2|  Andrew|    Fuller|
| 1|   Nancy|   Davolio|
+---+-----+-----+
```


Q17. Liệt kê customer name và product name sao cho số lượng product đó được mua bởi customer đó trong một order lớn hơn 5 lần số lượng trung bình product đó được mua trong một order bởi **tất cả** các customer .

In []:

```
+-----+-----+
| ContactName| ProductName|
+-----+-----+
| Roland Mendel|Manjimup Dried Ap...|
| Jose Pavarotti|Jack's New Englan...|
| Roland Mendel|Wimmers gute Semm...|
+-----+-----+
```

Q18. Liệt kê name của các customer đã mua **tất cả** product với giá < 5 .

In []:

```
+-----+
| ContactName|
+-----+
| Aria Cruz|
| Horst Kloss|
| Giovanni Rovelli|
| Karl Jablonski|
| Peter Franken|
| Maurizio Moroni|
| Jose Pavarotti|
| Jonas Bergulfsen|
| Pascale Cartrain|
| Philip Cramer|
| Mario Pontes|
| Carlos Hernández|
| Elizabeth Lincoln|
| Roland Mendel|
| Palle Ibsen|
| Annette Roulet|
| Christina Berglund|
|Miguel Angel Paolino|
+-----+
```

Q19. Liệt kê name của các product được bán bởi **mọi** employee .

In []:

```

+-----+
|      ProductName|
+-----+
|   Gorgonzola Telino|
|Louisiana Fiery H...|
|   Lakkalikööri|
|   Zaanse koeken|
|           Chai|
|           Pavlova|
|   Boston Crab Meat|
|   Tourtière|
|Teatime Chocolate...|
|Wimmers gute Semm...|
|Jack's New Englan...|
|Mozzarella di Gio...|
|   Steeleye Stout|
|Raclette Courdavault|
|   Chartreuse verte|
|   Gudbrandsdalsost|
|Original Frankfur...|
|           Ikura|
|   Guaraná Fantástica|
|   Sir Rodney's Scones|
+-----+
only showing top 20 rows

```

Q20. Liệt kê name của các customer đã mua **tất cả** các product được đặt hàng bởi customer có id là 'LAZYK' .

In []:

```

+-----+
|      ContactName|
+-----+
|   Carlos Hernández|
|Guillermo Fernández|
|   Renate Messner|
|   Philip Cramer|
|   Roland Mendel|
|   Patricia McKenna|
|   Elizabeth Lincoln|
|Henriette Pfalzheim|
+-----+

```