

Using Pyodbc and Pandas to execute SQL query.

Sixian Chen

11/24/2020

LinkedIn: <http://linkedin.com/in/seashane-sixian-chen> (<http://linkedin.com/in/seashane-sixian-chen>)

Background of this database

This database was downloaded from *MS-Access2010 Tutorials* .

According to database's creator, Profsr Financial Services is a privately-owned broker for insurance, mutual funds and other financial instruments. The database was designed for helping the firm to track sales and commissions for 30 or so sales professionals. As is usually done in the business, sales persons are paid only on commission based on their performance. Their sales have to be analyzed constantly over time (weekly, monthly, annually) and by product and product line because certain persons specialize in certain product lines or specific products that may or may not be more profitable than others.

The commission rate is different for different products. It may be calculated on the premium paid in the case of insurance or on the invested amount in the case of funds. There is also a second commission level which represents the amount paid to the agency.

Sales and commissions reports will be produced as required, usually based on dates.

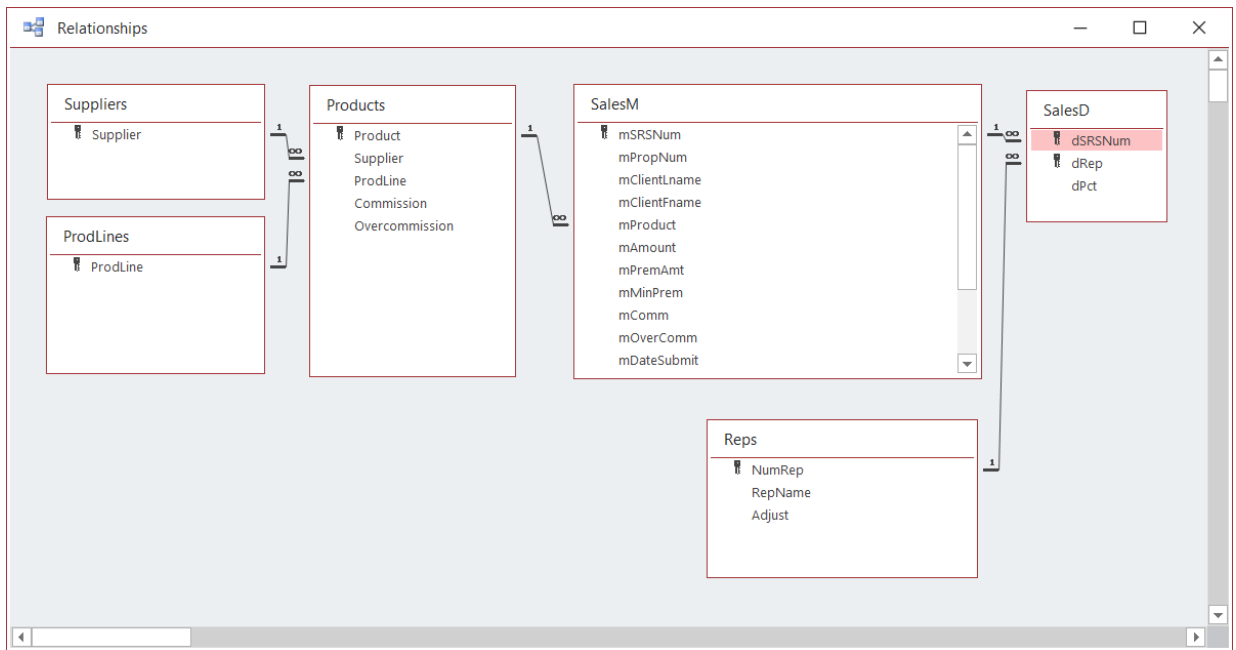
Different groupings will be needed to analyze products, product lines, special promotions and so on.

Download Source: <http://www.ms-access2010.com/tutorials/download.html> (<http://www.ms-access2010.com/tutorials/download.html>)

For more information: <http://www.ms-access2010.com/tutorials/project.html> (<http://www.ms-access2010.com/tutorials/project.html>)

The following shows the relationships of the dataset's tables.

Out[587]:



This database has the following tables:

ProdLines
Products
ReportDates
Reps
SalesD
SalesM
Suppliers
Switchboard Items

PREVIEW OF ALL TABLES IN THIS DATABASE

Table: ProdLines

Out[589]:

	ProdLine
0	Annuity
1	Disability Ins.
2	Life Ins.
3	Mutual funds

Table: Products

Out[590]:

	Product	Supplier	ProdLine	Commission	Overcommission
0	10 Payments	Excellence	Life Ins.	45.0	56.0
1	20 Payments	Excellence	Life Ins.	45.0	56.0
2	Accent-Life 10	Zurich	Life Ins.	35.0	40.0
3	Advantage Life Plus	Transamerica Life	Life Ins.	65.0	58.0
4	AGF funds	AGF	Mutual funds	5.0	100.0

Table: ReportDates

Out[591]:

	dateFrom	dateTo
0	2010-12-01	2010-12-19

Table: Reps

Out[592]:

	NumRep	RepName	Adjust
0	1057	Adam Carson	33.0
1	1197	Jean Campo	33.0
2	1232	Guy Tripper	28.0
3	1266	Ronald Charles	33.0
4	1320	Ken Dye	33.0

Table: SalesD

Out[593]:

	dSRSNum	dRep	dPct
0	1	1430	25.0
1	1	2623	75.0
2	2	3396	100.0
3	3	1197	100.0
4	4	1197	100.0

Table: SalesM

Out[594]:

	mSRSNum	mPropNum	mClientLname	mClientFname	mProduct	mAmount	mPremAmt	mM
0	1	013962	DOE	RICHARD	Perspecta (V.U)	50000.0	1020.000000	818
1	2	014049	DOE	GABRIEL	Horizon Progression: Pol.	112000.0	1080.000000	617
2	3	021834	DOE	MARCEL	Protector and Protector improved accelerated	50000.0	447.480011	447
3	4	021836	DOE	JOANNE	Protector and Protector improved accelerated	50000.0	319.140015	319
4	5	023205	DOE	DIANE	Protector and Protector improved accelerated	50000.0	569.510010	569

Table: Suppliers

Out[595]:

	Supplier
0	ACA
1	AGF
2	AIG
3	AVDL
4	C.I.

Table: Switchboard Items

Out[596]:

	SwitchboardID	ItemNumber	ItemText	Command	Argument
0	1	0	Main Switchboard	NaN	Default
1	1	1	Forms	1.0	2
2	1	2	Reports	1.0	3
3	1	3	Exit Application	6.0	
4	2	0	Forms Switchboard Page	0.0	None

SQL QUERY EXAMPLES

1. Execute a table which shows the data of the product 'Protector and Protector improved accelerated' in table SalesM sorting by the investment amount in descending order.

The query is:

```
SELECT * FROM SalesM WHERE mProduct='Protector AND Protector improved accelerated' ORDER BY mAmount DESC ;
```

Out[597]:

	mSRSNum	mPropNum	mClientLname	mClientFname	mProduct	mAmount	mPremAmt	mN
0	73	010954622	DOE	JOHN	Protector and Protector improved accelerated	60210.0	4744.470215	4744
1	5	023205	DOE	DIANE	Protector and Protector improved accelerated	50000.0	569.510010	569
2	4	021836	DOE	JOANNE	Protector and Protector improved accelerated	50000.0	319.140015	319
3	3	021834	DOE	MARCEL	Protector and Protector improved accelerated	50000.0	447.480011	447
4	50	010952954	DOE	MAURICE	Protector and Protector improved accelerated	45119.0	2966.699951	2966
5	39	023196	DOE	NOAH	Protector and Protector improved accelerated	23160.0	277.790009	277
6	77	010950989	DOE	DAPHNE	Protector and Protector improved accelerated	20000.0	302.980011	302
7	22	207871	DOE	JUDITH	Protector and Protector improved accelerated	15000.0	297.089996	297
8	6	023206	DOE	ROGER	Protector and Protector improved accelerated	10000.0	520.510010	520

2. What is average investment amount of 'Protector and Protector improved accelerated'?

The query is:

```
SELECT avg(mAmount) as [average investment amount] FROM SalesM WHERE mProduct = 'Protector AND Protector improved accelerated' ;
```

Out[598]:

	average investment amount
0	35943.222222

3. Execute the client list whose representative is 'Adam Carson'.

The query is:

```
SELECT mClientFname, mClientLname FROM SalesM, SalesD, Reps WHERE Reps.NumRep=SalesD.dRep AND SalesD.dSRSNum=SalesM.mSRSNum AND RepName='Adam Carson' ;
```

Out[599]:

	mClientFname	mClientLname
0	PATRICIA	DOE
1	JAM	DOE
2	TONY	DOE
3	MIKE	DOE
4	JOAN	DOE

4. Execute a list of representatives whose client had an investment amount over 100000 in year 2010, with the corresponding client names.

The query is:

```
SELECT RepName, mClientFname, mClientLname, mDateSubmit,mAmount FROM SalesM, SalesD, Reps WHERE Reps.NumRep=SalesD.dRep AND SalesD.dSRSNum=SalesM.mSRSNum AND mAmount>100000 AND YEAR(mDateSubmit) = 2010 ;
```

Out[600]:

	RepName	mClientFname	mClientLname	mDateSubmit	mAmount
0	Carl Brunet	MELANIE	DOE	2010-10-19	110000.0
1	Carl Brunet	CHRIS	DOE	2010-12-03	110000.0
2	Mike Jordan	CLAUDE	DOE	2010-10-28	164299.0
3	Jean Lefrank	CLARISSA	DOE	2010-12-12	550000.0
4	Mark McDuff	LUCY	DOE	2010-11-06	255276.0

5. Execute a list of representatives and their managed total investment amount in 2011 sorting by the total investment amount in descending order.

The query is:

```
SELECT * FROM (SELECT RepName, sum(mAmount) as sumAmount FROM SalesM, SalesD,
Reps WHERE Reps.NumRep=SalesD.dRep AND SalesD.dSRSNum=SalesM.mSRSNum AND YEAR(m
DateSubmit)=2011 GROUP BY RepName) ORDER BY sumAmount DESC ;
```

Out[601]:

	RepName	sumAmount
0	Peter Carpenter	3000000.0
1	Adam Carson	3000000.0
2	Danielle Marcotte	2000000.0
3	Sylvia Mason	902000.0
4	Francine Robinson	585000.0
5	Robert Lepage	420000.0
6	Ben St-Pierre	347040.0
7	Jean Campo	306160.0
8	Jack Boise	250000.0
9	Mark McDuff	250000.0
10	Paul Scott	250000.0
11	Germaine Greene	225000.0
12	Ken Dye	200000.0
13	Jerry Turpin	200000.0
14	Ron Dubcheck	175000.0
15	Gabby Dorey	152000.0
16	John Sur	150000.0
17	Daniel Wymer	110000.0
18	Raymond Lacelle	102000.0
19	Mike Devries	100000.0
20	George Mercer	60000.0
21	Scott Dawes	50000.0
22	Daniel Charland	25000.0
23	Guy Tripper	10000.0

6. Execute a list of clients with mPropNum whose minimum premium is greater than the average minimum premium in ascending order.

The query is:

```
SELECT mPropNum, mClientFname, mClientLname, mMinPrem from SalesM WHERE mMinPrem > (SELECT avg(mMinPrem) FROM SalesM) ORDER BY mMinPrem ASC ;
```

Out[602]:

	mPropNum	mClientFname	mClientLname	mMinPrem
0	4176026R	BRENDA	DOE	2450.000000
1	208689	STAN	DOE	2615.760010
2	4272246R	DONALD	DOE	2803.000000
3	4275202R	FRANK	DOE	2803.000000
4	010952534	JACQUE	DOE	2966.669922
5	010952954	MAURICE	DOE	2966.699951
6	010952535	NICOLE	DOE	2966.699951
7	4328543R	MIKE	DOE	3000.000000
8	L10689751	LUCILLE	DOE	3353.280029
9	4311470R	DENISE	DOE	3474.000000
10	4316286R	LORRAINE	DOE	4300.000000
11	4277566R	POLLY	DOE	4458.000000
12	010952971	CLARISSA	DOE	4489.000000
13	4273870R	GRANT	DOE	4500.000000
14	010951882	CLAUDE	DOE	4744.450195
15	010954622	JOHN	DOE	4744.470215
16	207949	CLAUDE	DOE	4887.000000
17	4308569R	WAYNE	DOE	5000.000000
18	4297630R	JEAN	DOE	5330.000000
19	356475476	Mike	Doe	5555.000000
20	202568	ANDRÉ	DOE	6380.000000
21	4261740R	RICHARD	DOE	6862.000000
22	4277604R	CLEMENT	DOE	10000.000000
23	4368952R	JOAN	DOE	12000.000000
24	3119975R	MARIE	DOE	12175.000000
25	024322	PATRICIA	DOE	23190.000000
26	4368936R	VITA	DOE	26185.000000

7. Execute a list to show the average investment amount, the average, min, max of premium and commission rate, grouping by product line then by suppliers, and sorting by the average investment amount with the highest on top.

The query is:

```
SELECT ProdLine, supplier, avg(mAmount) as avgAmount, avg(mPremAmt) as avgPremium, min(mPremAmt) as minPremium, max(mPremAmt) as maxPremium, avg(mComm) as avgCommission, min(mComm) as minCommission, max(mComm) as maxCommission from SalesM, Products WHERE SalesM.mProduct=Products.Product GROUP BY ProdLine, Supplier ORDER BY avg(mAmount) DESC ;
```

Out[603]:

	ProdLine	supplier	avgAmount	avgPremium	minPremium	maxPremium	avgCommission
0	Life Ins.	Transamerica Life	250000.000000	459.000000	459.000000	459.000000	37.500000
1	Life Ins.	RBC	250000.000000	334.799988	334.799988	334.799988	35.000000
2	Life Ins.	AVDL	170171.934426	2713.521643	113.040001	50000.000000	49.491803
3	Mutual funds	AGF	100000.000000	0.000000	0.000000	0.000000	5.000000
4	Life Ins.	Zurich	100000.000000	477.359985	477.359985	477.359985	42.500000
5	Life Ins.	Standard Life	100000.000000	2310.000000	1020.000000	3600.000000	65.000000
6	Life Ins.	AIG	100000.000000	5555.000000	5555.000000	5555.000000	32.500000
7	Annuity	AVDL	14755.161696	5624.833844	903.000000	28803.000000	1.500000
8	Life Ins.	ACA	11500.000000	493.860001	415.079987	572.640015	60.000000

8. Execute a list of suppliers who are in the Suppliers table but not in the Products table .

The query is:

```
SELECT Supplier FROM Suppliers WHERE Suppliers.Supplier NOT IN (SELECT DISTINCT Supplier FROM Products) ;
```

Out[604]:

	Supplier
0	Fidelity
1	Sun Life

9. Execute a list of product with the total counts of cases and sum of investment amount with Submission date in February 2011, order by the sum of investment amount descending.

The query is:

```
SELECT mProduct, count(mProduct) as ProductCounts, sum(mAmount) as SumAmount FROM SalesM WHERE YEAR(mDateSubmit)=2011 AND MONTH(mDateSubmit)=2 GROUP BY mProduct ORDER BY sum(mAmount) DESC ;
```

Out[605]:

	mProduct	ProductCounts	SumAmount
0	Horizon Liberty accelerated	7	2475000.0
1	Quietude +T65 accelerated	2	525000.0
2	Horizon 10 : 100 000 to 499 999 \$	2	300000.0
3	Temporary 10 Plus: 50,000 - 999,999\$	1	250000.0
4	T - 10	1	250000.0
5	Horizon Progression accelerated	2	225000.0
6	Horizon 20 accelerated : 100 000 to 499 999 \$	2	200000.0
7	Perspecta (V.U)	1	150000.0
8	Horizon Progression: Pol.	1	112000.0
9	Qualiterm 10	1	100000.0
10	AGF funds	1	100000.0
11	Priviledge health accelerated	1	100000.0
12	Protector and Protector improved accelerated	3	88160.0
13	Horizon Habitation	1	59040.0
14	Quietude +T75	1	50000.0
15	Solo	4	39000.0
16	Horizon Security accelerated	2	35000.0
17	Life Certified	2	23000.0

10. Execute a list of products which are in top 5 Premium over Investment Amount rate, sorting by the rate in descending order.

The query is:

```
SELECT TOP 5 mProduct, avg(mPremAmt/mAmount)*100 as [% of Premium over Amount]  
FROM SalesM GROUP BY mProduct ORDER BY avg(mPremAmt/mAmount) DESC ;
```

Out[606]:

	mProduct	% of Premium over Amount
0	IPS 3	64.429221
1	Solo	22.908540
2	ART 75	5.555000
3	Protector Future 20 accelerated	4.982868
4	Life Certified	4.503050