Excel Lab Guide

The essence of multi-dimensional data modeling is to prepare the tables and hence the associated data in a format suitable for analysis. Once the database is multi-dimensionally modeled, we can start analyzing the data from various perspectives.

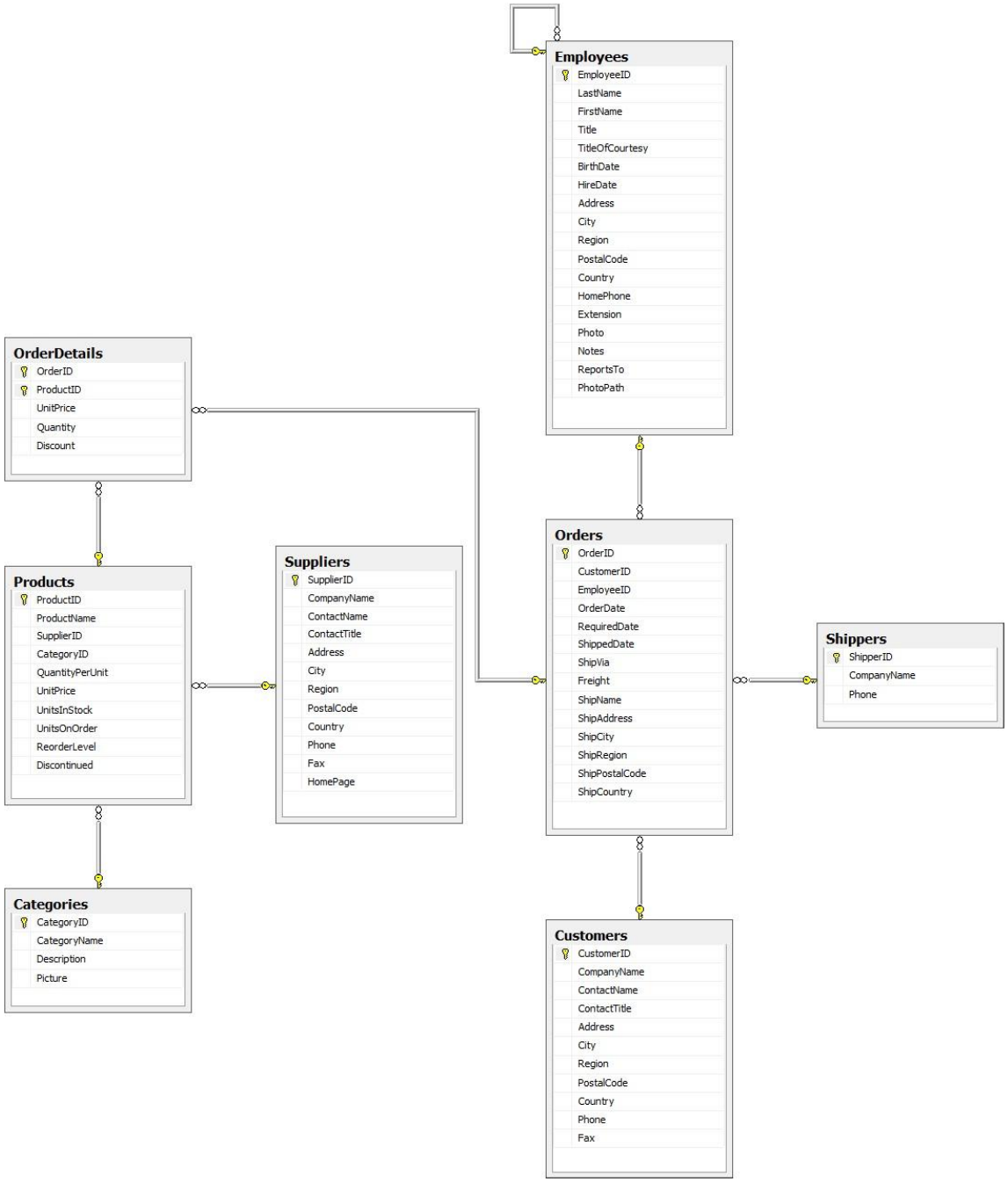
Understanding the Need for Data Analysis

Northwind is an online shopping facility, which deals with supplies and ships various products, the orders for which are placed online. It sells various categories of products such as beverages, dairy products, grains/cereals, meat/poultry, seafood, etc. to over 21 different countries including Canada, France, Germany, Italy, Portugal, Spain, UK and USA.

The customer places an order via a phone call, which is attended to by an employee. Every order consists of various products belonging to various categories and being supplied by different suppliers. Once the orders have been placed, all the products corresponding to the respective order are shipped to the respective customer via the various shippers.

The database design is explained below:

1. The **Employees** table contains details of all the employees present in the company.
2. The **Customers** table contains details of all the customers who order the various products.
3. The **Shippers** table contains details of various shippers who transport the products from the warehouse to the customer.
4. The **Suppliers** table contains details of various suppliers who supply the products to the warehouse.
5. The **Categories** table contains descriptions of the categories to which a product belongs.
6. The **Products** table contains the information of the various products being sold along with the categories they belong to and the suppliers who supply the corresponding products.
7. The **Orders** table keeps a track of a complete transaction including the customer who placed the order, the employee who took the order, the shipper who would be making the delivery, the dates of ordering and shipping, etc.
8. The **OrderDetails** table keeps a detailed list of all the products that were bought in a single order including the total quantity ordered, and the possible available discount.



The table structures are as given below:

EMPLOYEES TABLE

Column	Data Type	
EmployeeID	int	NOT NULL, PRIMARY KEY
LastName	nvarchar(20)	NOT NULL
FirstName	nvarchar(10)	NOT NULL
Title	nvarchar(30)	
TitleOfCourtesy	nvarchar(25)	
BirthDate	datetime	

HireDate	datetime	
Address	nvarchar(60)	
City	nvarchar(15)	
Region	nvarchar(15)	
PostalCode	nvarchar(10)	
Country	nvarchar(15)	
HomePhone	nvarchar(24)	
Extension	nvarchar(4)	
Photo	image	
Notes	ntext	
ReportsTo	int	
PhotoPath	nvarchar(255)	

CUSTOMERS TABLE

Column	**Data Type**	
CustomerID	nchar(5)	NOT NULL, PRIMARY KEY
CompanyName	nvarchar(40)	NOT NULL
ContactName	nvarchar(30)	
ContactTitle	nvarchar(30)	
Address	nvarchar(60)	
City	nvarchar(15)	
Region	nvarchar(15)	
PostalCode	nvarchar(10)	
Country	nvarchar(15)	
Phone	nvarchar(24)	
Fax	nvarchar(24)	

SHIPPERS TABLE

Column	**Data Type**	
ShipperID	int	NOT NULL, PRIMARY KEY
CompanyName	nvarchar(40)	NOT NULL
Phone	nvarchar(24)	

SUPPLIERS TABLE

Column	**Data Type**	
SupplierID	int	NOT NULL, PRIMARY KEY
CompanyName	nvarchar(40)	NOT NULL
ContactName	nvarchar(30)	
ContactTitle	nvarchar(30)	
Address	nvarchar(60)	
City	nvarchar(15)	
Region	nvarchar(15)	
PostalCode	nvarchar(10)	
Country	nvarchar(15)	
Phone	nvarchar(24)	
Fax	nvarchar(24)	
HomePage	ntext	

CATEGORIES TABLE

Column	Data Type	
CategoryID	int	NOT NULL, PRIMARY KEY
CategoryName	nvarchar(15)	NOT NULL
Description	ntext	
Picture	image	

PRODUCTS TABLE

Column	Data Type	
ProductID	int	NOT NULL, PRIMARY KEY
ProductName	nvarchar(40)	NOT NULL
SupplierID	int	SUPPLIER(SupplierID)
CategoryID	int	CATEGORY(CategoryID)
QuantityPerUnit	nvarchar(20)	
UnitPrice	money	
UnitsInStock	smallint	
UnitsOnOrder	smallint	
ReorderLevel	smallint	
Discontinued	bit	NOT NULL

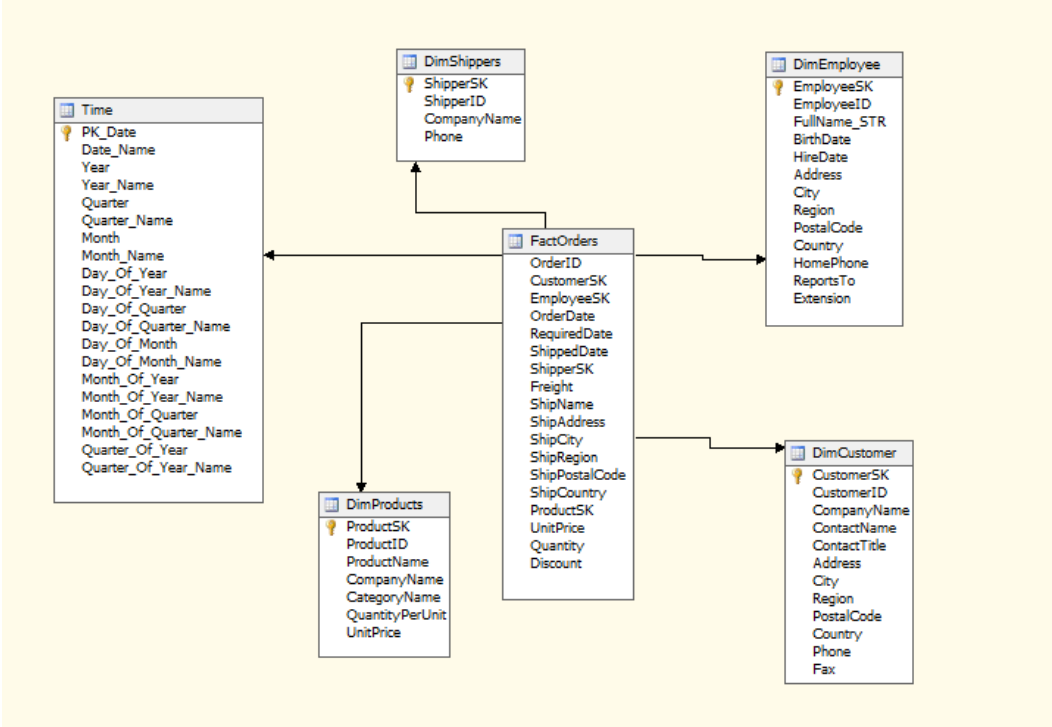
ORDERS TABLE

Column	Data Type	
OrderID	int	NOT NULL, PRIMARY KEY
CustomerID	nchar(5)	CUSTOMER(CustomerID)
EmployeeID	int	EMPLOYEE(EmployeeID)
OrderDate	datetime	
RequiredDate	datetime	
ShippedDate	datetime	
ShipVia	int	SHIPPER(ShipperID)
Freight	money	
ShipName	nvarchar(40)	
ShipAddress	nvarchar(60)	
ShipCity	nvarchar(15)	
ShipRegion	nvarchar(15)	
ShipPostalCode	nvarchar(10)	
ShipCountry	nvarchar(15)	

ORDERDETAILS TABLE

Column	Data Type	
OrderID	int	NOT NULL, PRIMARY KEY, ORDER(OrderID)
ProductID	int	NOT NULL, PRIMARY KEY, PRODUCT(ProductID)
UnitPrice	money	NOT NULL
Quantity	smallint	NOT NULL
Discount	real	NOT NULL

Since we need to analyze the data using multiple scenarios, this data model which is a relational model is converted to a dimensional model using the dimension modeling steps so that it is better suited for analysis. The dimensional model is as shown below:

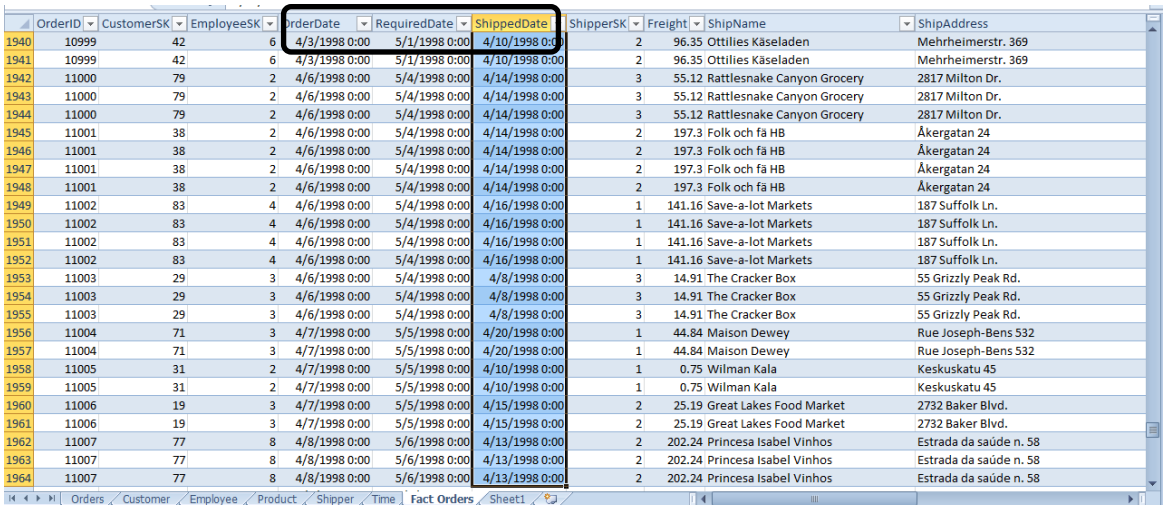


The data associated with the above multi-dimensional model is shared in an excel sheet (Northwind_Traders data.xlsx) in the folder MS Excel Lab Guide and the same data is used for analysis in this lab guide.

Note: The OrderDate, ShippedDate and the RequiredDate which are columns in the Orders table in the relational model have been mapped to a column named PK_Date of the Time dimension. If you observe the columns in the Time table we see that the PK_Date has the corresponding month, quarter and year to which it belongs. This Time dimension plays a major role in monthly, quarterly and annual data analyses.

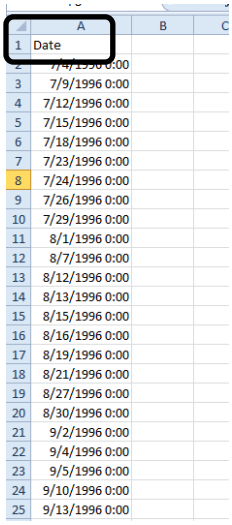
The Time dimension has been created from the available date columns using a tool called SQL server analysis services 2008. However, we can create a Time table in excel using the date columns in the following way:

Go to the Fact Orders sheet of the Northwind_Traders data workbook. Individually select all values in the columns OrderDate, RequiredDate and ShippedDate and paste them in one single column in a new sheet. The below 2 snapshots will assist you in doing the same.



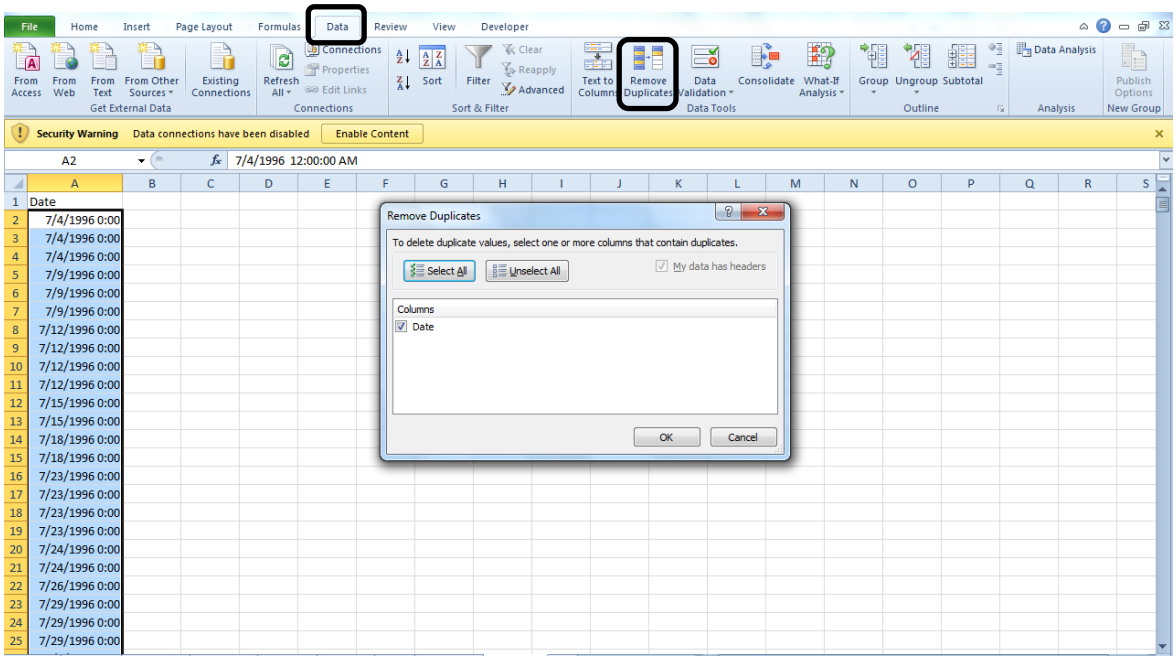
	OrderID	CustomerSK	EmployeeSK	OrderDate	RequiredDate	ShippedDate	ShipperSK	Freight	ShipName	ShipAddress
1940	10999	42	6	4/3/1998 0:00	5/1/1998 0:00	4/10/1998 0:00	2	96.35	Ottilies Käseladen	Mehrheimerstr. 369
1941	10999	42	6	4/3/1998 0:00	5/1/1998 0:00	4/10/1998 0:00	2	96.35	Ottilies Käseladen	Mehrheimerstr. 369
1942	11000	79	2	4/6/1998 0:00	5/4/1998 0:00	4/14/1998 0:00	3	55.12	Rattlesnake Canyon Grocery	2817 Milton Dr.
1943	11000	79	2	4/6/1998 0:00	5/4/1998 0:00	4/14/1998 0:00	3	55.12	Rattlesnake Canyon Grocery	2817 Milton Dr.
1944	11000	79	2	4/6/1998 0:00	5/4/1998 0:00	4/14/1998 0:00	3	55.12	Rattlesnake Canyon Grocery	2817 Milton Dr.
1945	11001	38	2	4/6/1998 0:00	5/4/1998 0:00	4/14/1998 0:00	2	197.3	Folk och fä HB	Åkergatan 24
1946	11001	38	2	4/6/1998 0:00	5/4/1998 0:00	4/14/1998 0:00	2	197.3	Folk och fä HB	Åkergatan 24
1947	11001	38	2	4/6/1998 0:00	5/4/1998 0:00	4/14/1998 0:00	2	197.3	Folk och fä HB	Åkergatan 24
1948	11001	38	2	4/6/1998 0:00	5/4/1998 0:00	4/14/1998 0:00	2	197.3	Folk och fä HB	Åkergatan 24
1949	11002	83	4	4/6/1998 0:00	5/4/1998 0:00	4/16/1998 0:00	1	141.16	Save-a-lot Markets	187 Suffolk Ln.
1950	11002	83	4	4/6/1998 0:00	5/4/1998 0:00	4/16/1998 0:00	1	141.16	Save-a-lot Markets	187 Suffolk Ln.
1951	11002	83	4	4/6/1998 0:00	5/4/1998 0:00	4/16/1998 0:00	1	141.16	Save-a-lot Markets	187 Suffolk Ln.
1952	11002	83	4	4/6/1998 0:00	5/4/1998 0:00	4/16/1998 0:00	1	141.16	Save-a-lot Markets	187 Suffolk Ln.
1953	11003	29	3	4/6/1998 0:00	5/4/1998 0:00	4/8/1998 0:00	3	14.91	The Cracker Box	55 Grizzly Peak Rd.
1954	11003	29	3	4/6/1998 0:00	5/4/1998 0:00	4/8/1998 0:00	3	14.91	The Cracker Box	55 Grizzly Peak Rd.
1955	11003	29	3	4/6/1998 0:00	5/4/1998 0:00	4/8/1998 0:00	3	14.91	The Cracker Box	55 Grizzly Peak Rd.
1956	11004	71	3	4/7/1998 0:00	5/5/1998 0:00	4/20/1998 0:00	1	44.84	Maison Dewey	Rue Joseph-Bens 532
1957	11004	71	3	4/7/1998 0:00	5/5/1998 0:00	4/20/1998 0:00	1	44.84	Maison Dewey	Rue Joseph-Bens 532
1958	11005	31	2	4/7/1998 0:00	5/5/1998 0:00	4/10/1998 0:00	1	0.75	Wilman Kala	Keskuskatu 45
1959	11005	31	2	4/7/1998 0:00	5/5/1998 0:00	4/10/1998 0:00	1	0.75	Wilman Kala	Keskuskatu 45
1960	11006	19	3	4/7/1998 0:00	5/5/1998 0:00	4/15/1998 0:00	2	25.19	Great Lakes Food Market	2732 Baker Blvd.
1961	11006	19	3	4/7/1998 0:00	5/5/1998 0:00	4/15/1998 0:00	2	25.19	Great Lakes Food Market	2732 Baker Blvd.
1962	11007	77	8	4/8/1998 0:00	5/6/1998 0:00	4/13/1998 0:00	2	202.24	Princesa Isabel Vinhos	Estrada da saúde n. 58
1963	11007	77	8	4/8/1998 0:00	5/6/1998 0:00	4/13/1998 0:00	2	202.24	Princesa Isabel Vinhos	Estrada da saúde n. 58
1964	11007	77	8	4/8/1998 0:00	5/6/1998 0:00	4/13/1998 0:00	2	202.24	Princesa Isabel Vinhos	Estrada da saúde n. 58

Orders | Customer | Employee | Product | Shipper | Time | Fact Orders | Sheet1



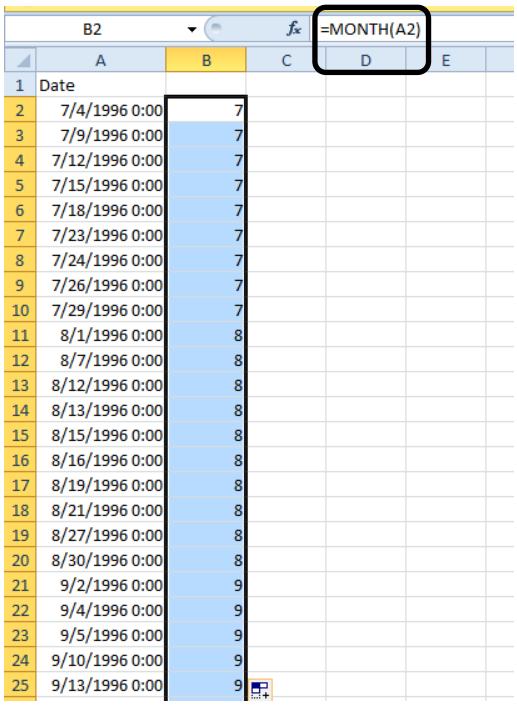
	A	B	C
1	Date		
2	7/4/1996 0:00		
3	7/9/1996 0:00		
4	7/12/1996 0:00		
5	7/15/1996 0:00		
6	7/18/1996 0:00		
7	7/23/1996 0:00		
8	7/24/1996 0:00		
9	7/26/1996 0:00		
10	7/29/1996 0:00		
11	8/1/1996 0:00		
12	8/7/1996 0:00		
13	8/12/1996 0:00		
14	8/13/1996 0:00		
15	8/15/1996 0:00		
16	8/16/1996 0:00		
17	8/19/1996 0:00		
18	8/21/1996 0:00		
19	8/27/1996 0:00		
20	8/30/1996 0:00		
21	9/2/1996 0:00		
22	9/4/1996 0:00		
23	9/5/1996 0:00		
24	9/10/1996 0:00		
25	9/13/1996 0:00		

Once all the dates are copied, select the data tab, select the Remove Duplicates option and click OK twice.

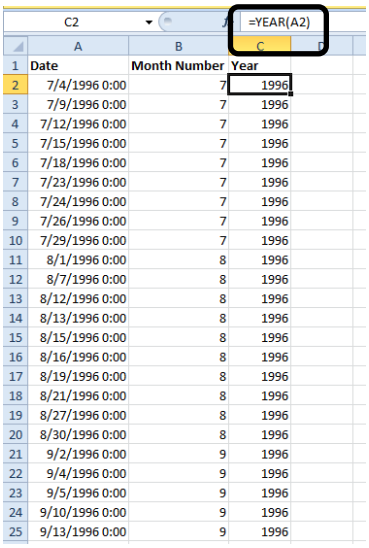


	A
1	Date
2	7/4/1996 0:00
3	7/4/1996 0:00
4	7/4/1996 0:00
5	7/9/1996 0:00
6	7/9/1996 0:00
7	7/9/1996 0:00
8	7/12/1996 0:00
9	7/12/1996 0:00
10	7/12/1996 0:00
11	7/12/1996 0:00
12	7/15/1996 0:00
13	7/15/1996 0:00
14	7/18/1996 0:00
15	7/18/1996 0:00
16	7/23/1996 0:00
17	7/23/1996 0:00
18	7/23/1996 0:00
19	7/23/1996 0:00
20	7/24/1996 0:00
21	7/24/1996 0:00
22	7/26/1996 0:00
23	7/29/1996 0:00
24	7/29/1996 0:00
25	7/29/1996 0:00

Use the Month and Year Excel functions to fetch the corresponding month numbers and years from the dates. Make sure all the dates have the respective month numbers and years in the adjacent columns as shown below in the 2 snapshots.



	A	B	C	D	E
1	Date				
2	7/4/1996 0:00	7			
3	7/9/1996 0:00	7			
4	7/12/1996 0:00	7			
5	7/15/1996 0:00	7			
6	7/18/1996 0:00	7			
7	7/23/1996 0:00	7			
8	7/24/1996 0:00	7			
9	7/26/1996 0:00	7			
10	7/29/1996 0:00	7			
11	8/1/1996 0:00	8			
12	8/7/1996 0:00	8			
13	8/12/1996 0:00	8			
14	8/13/1996 0:00	8			
15	8/15/1996 0:00	8			
16	8/16/1996 0:00	8			
17	8/19/1996 0:00	8			
18	8/21/1996 0:00	8			
19	8/27/1996 0:00	8			
20	8/30/1996 0:00	8			
21	9/2/1996 0:00	9			
22	9/4/1996 0:00	9			
23	9/5/1996 0:00	9			
24	9/10/1996 0:00	9			
25	9/13/1996 0:00	9			

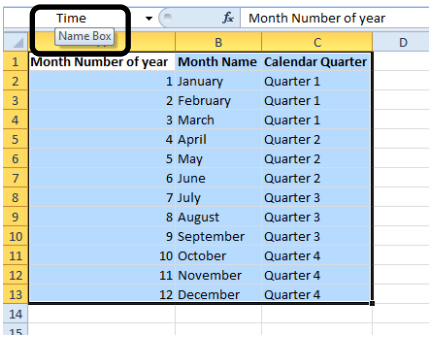


	A	B	C
1	Date	Month Number	Year
2	7/4/1996 0:00	7	1996
3	7/9/1996 0:00	7	1996
4	7/12/1996 0:00	7	1996
5	7/15/1996 0:00	7	1996
6	7/18/1996 0:00	7	1996
7	7/23/1996 0:00	7	1996
8	7/24/1996 0:00	7	1996
9	7/26/1996 0:00	7	1996
10	7/29/1996 0:00	7	1996
11	8/1/1996 0:00	8	1996
12	8/7/1996 0:00	8	1996
13	8/12/1996 0:00	8	1996
14	8/13/1996 0:00	8	1996
15	8/15/1996 0:00	8	1996
16	8/16/1996 0:00	8	1996
17	8/19/1996 0:00	8	1996
18	8/21/1996 0:00	8	1996
19	8/27/1996 0:00	8	1996
20	8/30/1996 0:00	8	1996
21	9/2/1996 0:00	9	1996
22	9/4/1996 0:00	9	1996
23	9/5/1996 0:00	9	1996
24	9/10/1996 0:00	9	1996
25	9/13/1996 0:00	9	1996

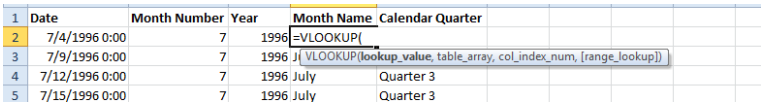
Create a dataset in a new excel sheet as shown below:

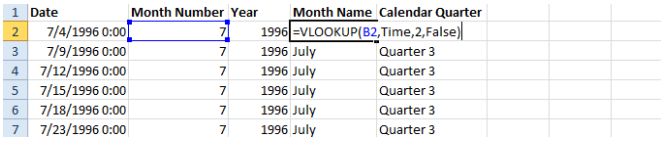
	Month Number of year	Month Name	Calendar Quarter
2	1	January	Quarter 1
3	2	February	Quarter 1
4	3	March	Quarter 1
5	4	April	Quarter 2
6	5	May	Quarter 2
7	6	June	Quarter 2
8	7	July	Quarter 3
9	8	August	Quarter 3
10	9	September	Quarter 3
11	10	October	Quarter 4
12	11	November	Quarter 4
13	12	December	Quarter 4
14			

Select the entire dataset and name it as “Time” in the Name Box as shown below. Do not forget to hit Enter once name is typed in the name box.



In the sheet where we have the Date, Month Number and Year, add a new column – “Month Name” and type the vlookup formula as shown below;





Follow the same procedure for Calendar Quarter and use the formula =vlookup(B2,Time,3,False).

Note: Here is how the vlookup works: It takes the lookup_value which is the Month Number in this case, matches it with the values in the column whose number is equal to the col_index_num (2 in this case) in the Time table. Since the second column in the Time table has the month names they get populated in the cell where the vlookup function is executed. Double clicking on the right corner of the cell will populate the month names for all the month numbers as shown above. The same vlookup concept is used in converting the not so user friendly “Fact Orders” sheet to the user-friendly “Orders” sheet in the Northwind_Traders data workbook.

Organizing and Analyzing the Data

When we have a large amount of data, we want to organize it, analyze it, get summary information and then graph it. We need analytical tools for this, and the PivotTable and PivotChart tools in Microsoft Excel are some of the most popular tools. The data does not need to be in one workbook. We may analyze data from multiple workbooks without too much trouble.

Basics of PivotTables

PivotTables allow us to consolidate huge amounts of data with similar fields & analyze the consolidated data or just make a summary of the consolidated data. The PivotTable in Excel gives us a simple way to create a PivotTable for our data. Please note that the data should have at least 1 field in common, or else the consolidation will not work and any spelling error in the data will produce incorrect PivotTables. PivotTables in Excel are synonymous to Cubes in other analytical tools.

Basics of Charts

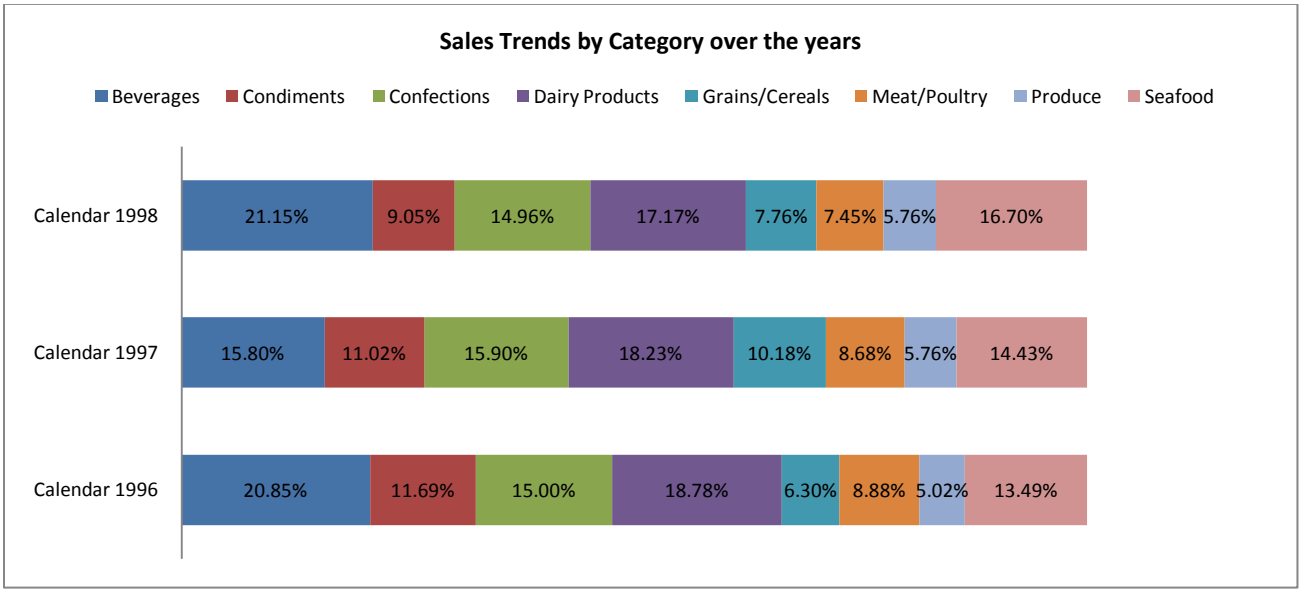
Charts are used to graphically represent data. The PivotChart tool in Excel provides a simple way to create a PivotTable and an accompanying chart. Remember, a chart is only as good as the data or the summary table (PivotTable). If we try to cram too many fields into a chart, we will end up with a non-informative chart. We must always try and keep it simple and informative.

Let us see some scenarios which explore the different ways of analyzing the Northwind data.

Scenario 1: Graph the percentage sales over time to see the trends.

Here is the answer:

Sum of Quantity	Column Labels								
Row Labels	Beverages	Condiments	Confections	Dairy Products	Grains/Cereals	Meat/Poultry	Produce	Seafood	Grand Total
Calendar 1996	20.85%	11.69%	15.00%	18.78%	6.30%	8.88%	5.02%	13.49%	100.00%
Calendar 1997	15.80%	11.02%	15.90%	18.23%	10.18%	8.68%	5.76%	14.43%	100.00%
Calendar 1998	21.15%	9.05%	14.96%	17.17%	7.76%	7.45%	5.76%	16.70%	100.00%
Grand Total	**18.38%**	**10.45%**	**15.45%**	**17.95%**	**8.78%**	**8.29%**	**5.65%**	**15.07%**	**100.00%**



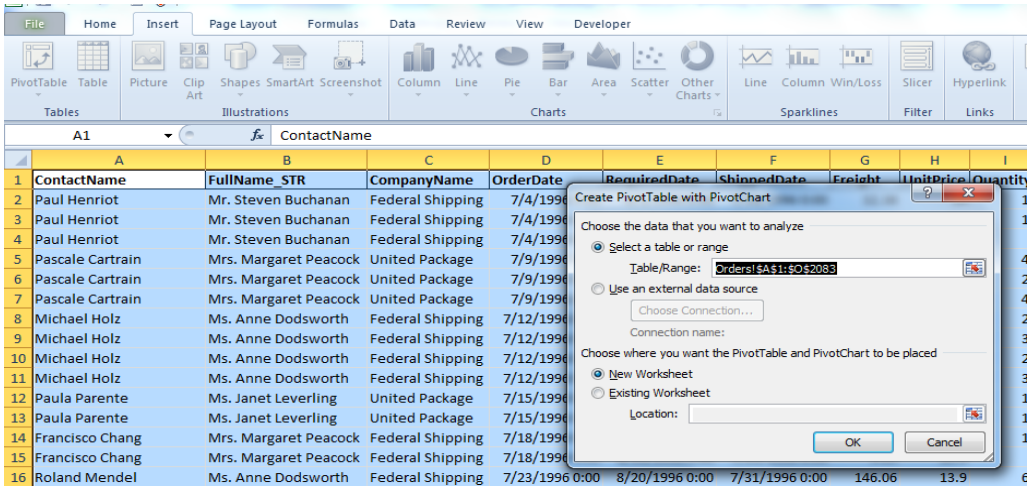
Now how do we get to this?

Here is how:

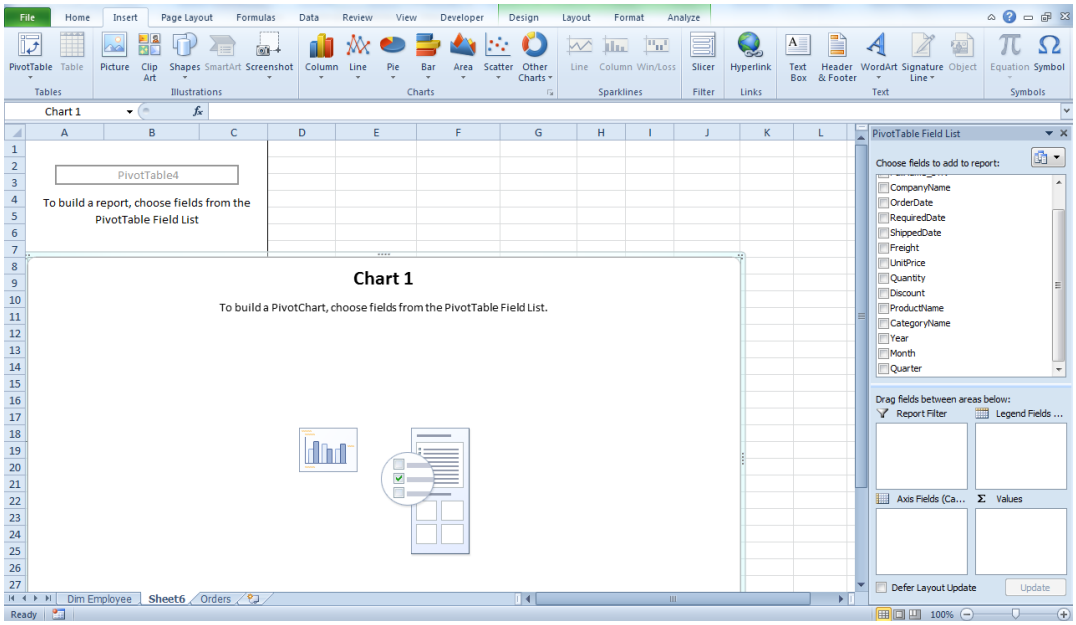
Creating a PivotTable and PivotChart

1. Open the Northwind_Traders data.xlsx workbook.
2. Go to the **Insert** tab on the ribbon.
3. Look for the **Tables** group, and select **PivotTable**. Choose **PivotChart** from the drop-down options.

4. You should now see the *Create PivotTable with PivotChart* dialog box.

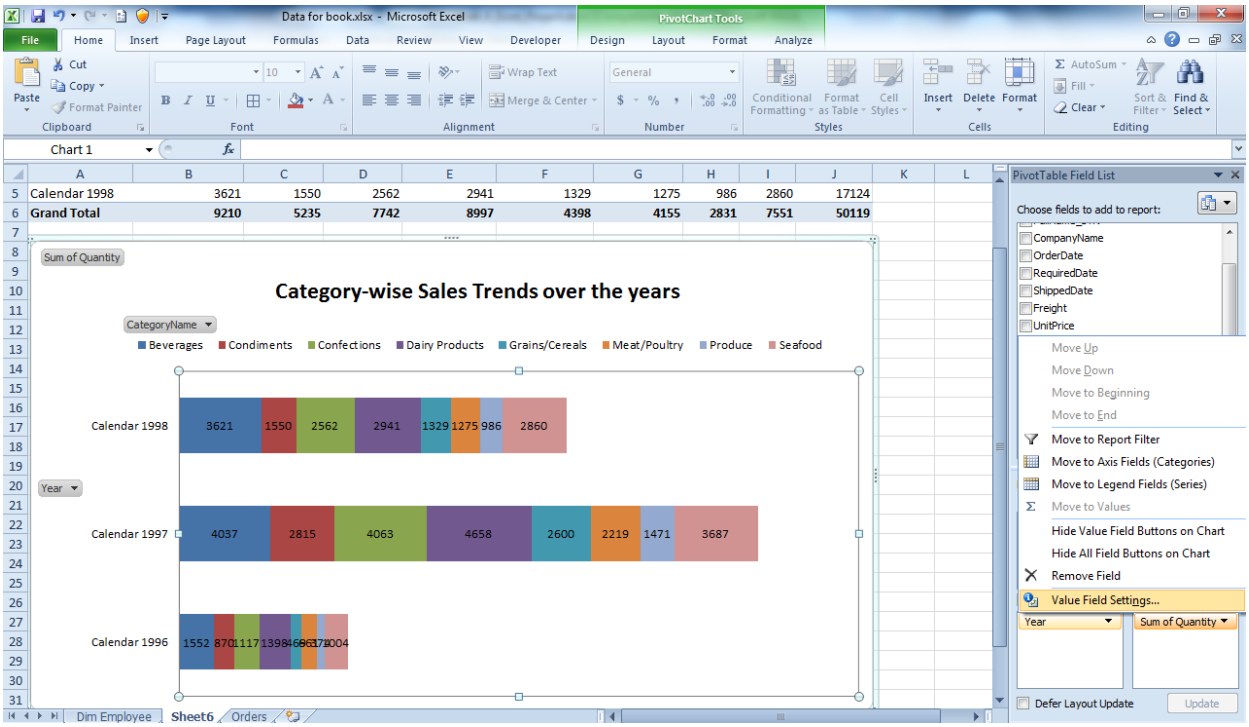


5. There are a couple of things you'll need to define in order to create your PivotTable. First, you need to tell Excel where to get the data from. In the *Choose the data you want to analyze* section of the dialog box, you will define where Excel is going to get the data for the PivotTable from. Since the data for this example is within the workbook, we will choose *Select a table or range*.
   a. Under *Table/Range*, Excel might be smart enough to automatically select the proper range of data from the worksheet.
   b. If it *does not,* click on the litt**l**e button on the right of the range box. Select the *range* of data. In the example, we have data in cells **A1 to O2083**. This means that we have selected fifteen columns (*A thro*ugh O) and 2083 **r**ows.
6. In the *Choose where you **want the Pi**votTable and PivotChart to be placed* section of the dialog box, you can choose to either create the PivotTable and the accompanying PivotChart on a new worksheet in the Excel workbook, or you can place both on the current worksheet. For this exercise, we will choose to create a new worksheet for the PivotTable and PivotChart. To do this, select *New Worksheet*, and click OK. You should now be taken to the new worksheet.

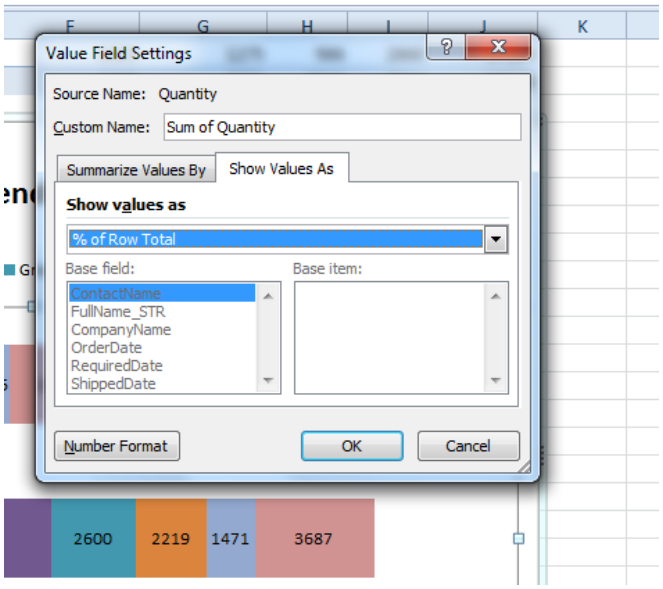


1. While on the new worksheet, you should see four new tabs on the ribbon:
   a. **Design**, **Layout**, **Format**, and **Analyze**.

   b. You should also see the **PivotTable Field List** and the **PivotChart Filter Pane** (if you don't see the panes, click on the **Analyze** tab in the ribbon, and, in the **Show/Hide** group, make sure that at least **Field List** is selected).
2. You now need to set up the PivotTable so that you can get the data summary you desire. The new tabs, as well as the two panes, will be used to help format your PivotTable and PivotChart to do the necessary analysis. For this exercise, let's see the total number of products sold by category in each year. In the **PivotTable Field List**, select **Quantity, Year** and **CategoryName** to add to the PivotTable report. Place the CategoryName in the Legends field, Year in the Axis Field and Quantity in the Value field. You should see the PivotTable and PivotChart update appropriately. Select the chart type as "Stacked List" to get the chart as shown below



You can use different functions on the field placed under *Values*. For example, if you wanted to see average quantity for each category, you would click on **Sum of Quantity** and choose **Value Field Settings**. In the **Value Field Settings** dialog box, choose the "Show Values As" tab and select the *% of Row Total* function to apply to **Sum of Quantity** (see the screenshot below). A number of other functions are available here to use in your PivotTable.



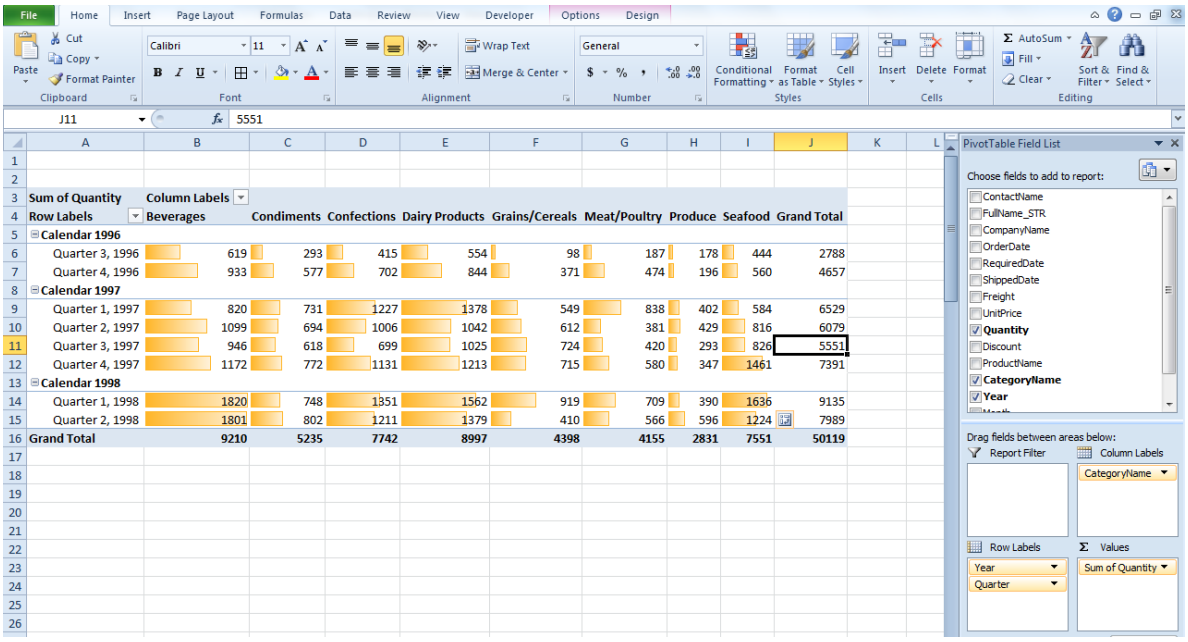
Once this is done, you would have graphed the percentage sales over time to see the trends.

Scenario 2: Pivot the data to see total sales by quarter and category. Are there any highs? Are there any lows that need to be addressed?

Here is the solution:

(The data has been pivoted in terms of Year and Quarter in columns and Category in Rows.)

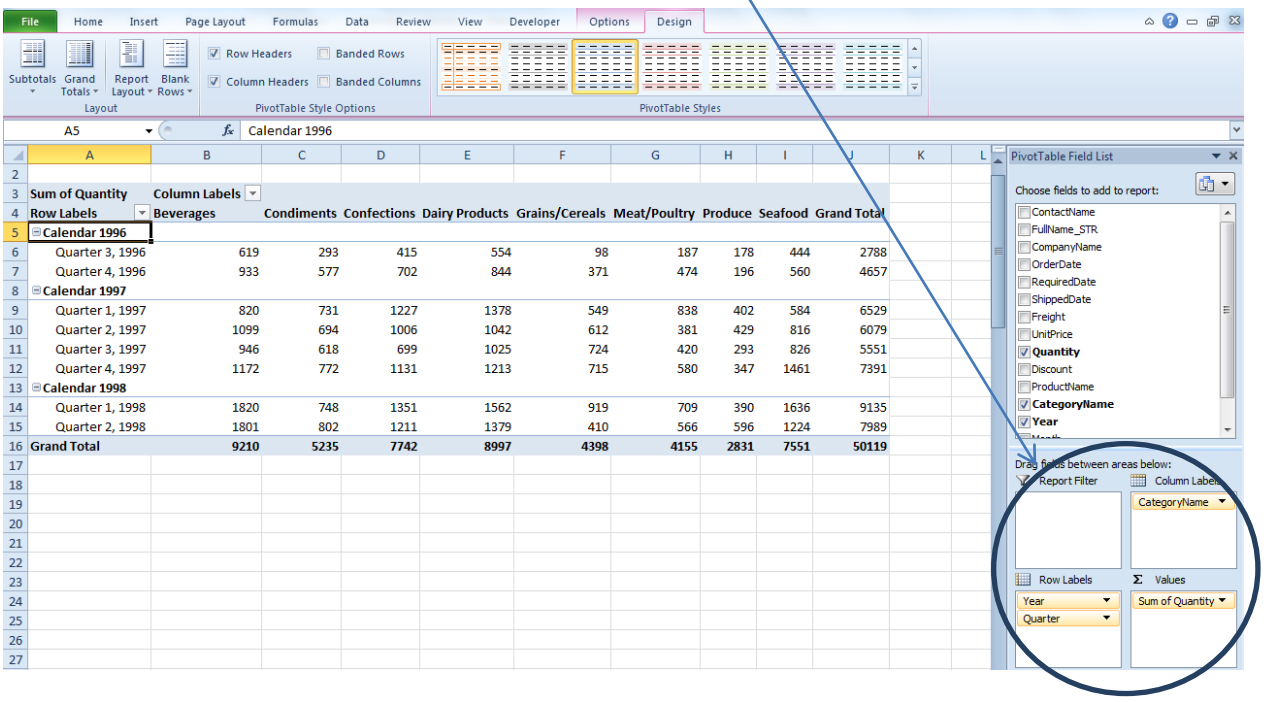
NOTE: *The Year, Quarter and the Month provided in the excel sheet are mapped to the required date and NOT year and quarter of the Orderdate or ShippedDate.*



Sum of Quantity	Column Labels								
Row Labels	Beverages	Condiments	Confections	Dairy Products	Grains/Cereals	Meat/Poultry	Produce	Seafood	Grand Total
Calendar 1996									
Quarter 3, 1996	619	293	415	554	98	187	178	444	2788
Quarter 4, 1996	933	577	702	844	371	474	196	560	4657
Calendar 1997									
Quarter 1, 1997	820	731	1227	1378	549	838	402	584	6529
Quarter 2, 1997	1099	694	1006	1042	612	381	429	816	6079
Quarter 3, 1997	946	618	699	1025	724	420	293	826	5551
Quarter 4, 1997	1172	772	1131	1213	715	580	347	1461	7391
Calendar 1998									
Quarter 1, 1998	1820	748	1351	1562	919	709	390	1636	9135
Quarter 2, 1998	1801	802	1211	1379	410	566	596	1224	7989
Grand Total	9210	5235	7742	8997	4398	4155	2831	7551	50119

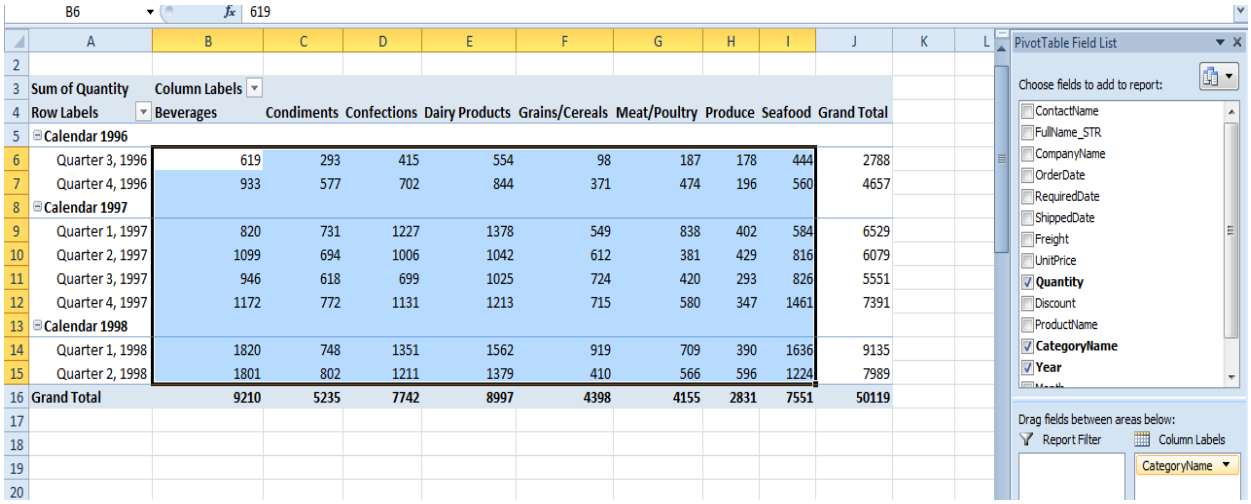
Here is how we do it:

1. Select the the entire table in the Orders worksheet.
2. Go to the Insert menu, select PivotTable.
3. Select the columns as highlighted in the circle below (make sure Quantity is represented as Sum of Quantity. Else use the previously used Value Field Settings option to sum it).



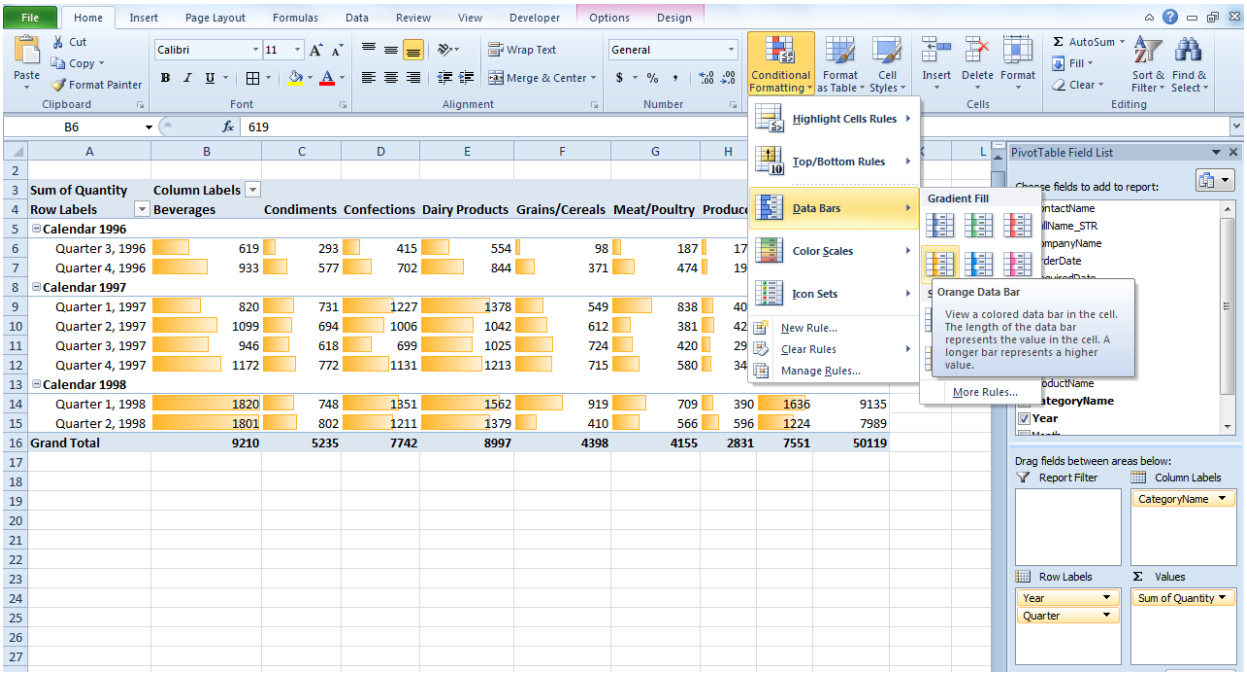
Sum of Quantity	Column Labels								
Row Labels	Beverages	Condiments	Confections	Dairy Products	Grains/Cereals	Meat/Poultry	Produce	Seafood	Grand Total
Calendar 1996									
Quarter 3, 1996	619	293	415	554	98	187	178	444	2788
Quarter 4, 1996	933	577	702	844	371	474	196	560	4657
Calendar 1997									
Quarter 1, 1997	820	731	1227	1378	549	838	402	584	6529
Quarter 2, 1997	1099	694	1006	1042	612	381	429	816	6079
Quarter 3, 1997	946	618	699	1025	724	420	293	826	5551
Quarter 4, 1997	1172	772	1131	1213	715	580	347	1461	7391
Calendar 1998									
Quarter 1, 1998	1820	748	1351	1562	919	709	390	1636	9135
Quarter 2, 1998	1801	802	1211	1379	410	566	596	1224	7989
Grand Total	9210	5235	7742	8997	4398	4155	2831	7551	50119

4. Select the data as shown in the screenshot below



Sum of Quantity	Column Labels								
Row Labels	Beverages	Condiments	Confections	Dairy Products	Grains/Cereals	Meat/Poultry	Produce	Seafood	Grand Total
Calendar 1996									
Quarter 3, 1996	619	293	415	554	98	187	178	444	2788
Quarter 4, 1996	933	577	702	844	371	474	196	560	4657
Calendar 1997									
Quarter 1, 1997	820	731	1227	1378	549	838	402	584	6529
Quarter 2, 1997	1099	694	1006	1042	612	381	429	816	6079
Quarter 3, 1997	946	618	699	1025	724	420	293	826	5551
Quarter 4, 1997	1172	772	1131	1213	715	580	347	1461	7391
Calendar 1998									
Quarter 1, 1998	1820	748	1351	1562	919	709	390	1636	9135
Quarter 2, 1998	1801	802	1211	1379	410	566	596	1224	7989
Grand Total	9210	5235	7742	8997	4398	4155	2831	7551	50119

5. Apply the conditional formating as shown below.
(Make sure you are in the Home menu to view the Conditional Formatting tab)

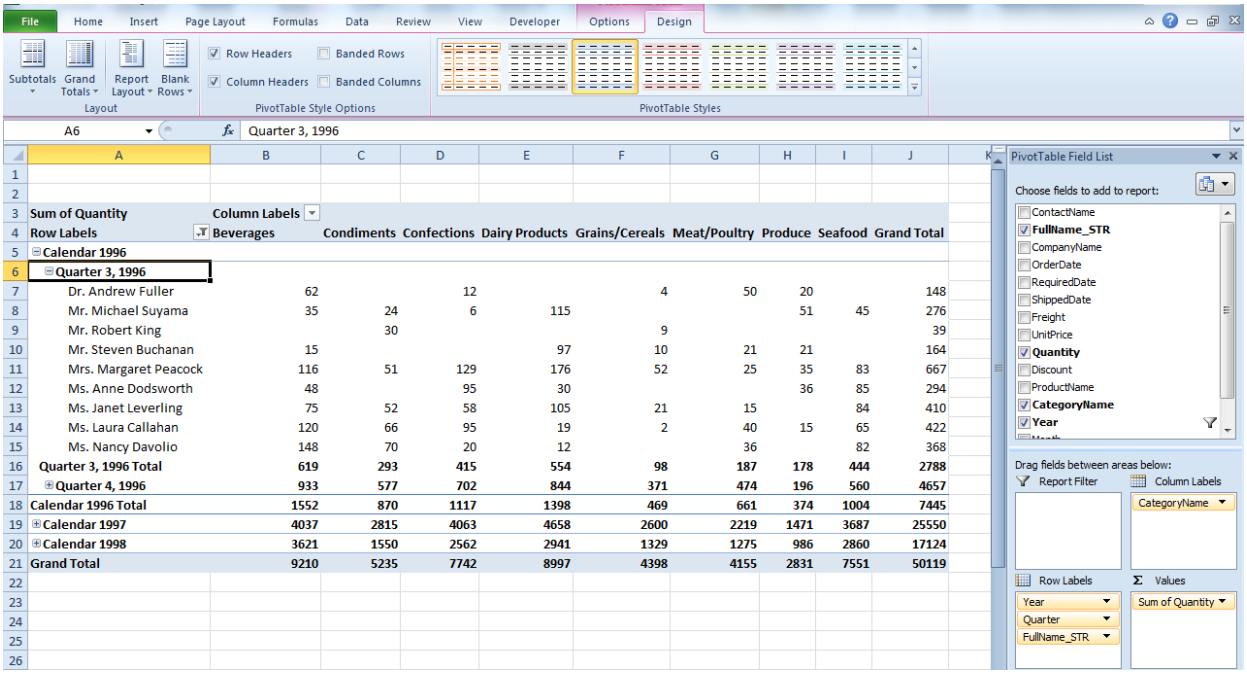


Sum of Quantity	Column Labels							
Row Labels	Beverages	Condiments	Confections	Dairy Products	Grains/Cereals	Meat/Poultry	Produc	
Calendar 1996								
Quarter 3, 1996	619	293	415	554	98	187	17	
Quarter 4, 1996	933	577	702	844	371	474	19	
Calendar 1997								
Quarter 1, 1997	820	731	1227	1378	549	838	40	
Quarter 2, 1997	1099	694	1006	1042	612	381	42	
Quarter 3, 1997	946	618	699	1025	724	420	29	
Quarter 4, 1997	1172	772	1131	1213	715	580	34	
Calendar 1998								
Quarter 1, 1998	1820	748	1351	1562	919	709	390	1636	9135
Quarter 2, 1998	1801	802	1211	1379	410	566	596	1224	7989
Grand Total	9210	5235	7742	8997	4398	4155	2831	7551	50119

By applying conditional formatting we are able to achieve better visualisation of data. It can be observed that the 2nd quarter of 1998 had the highest total sales and also the beverages category had the highest total sales for all years, closely followed by dairy products.

Scenario 3: How are quarterly sales totals by salesperson? Subtotal the data.

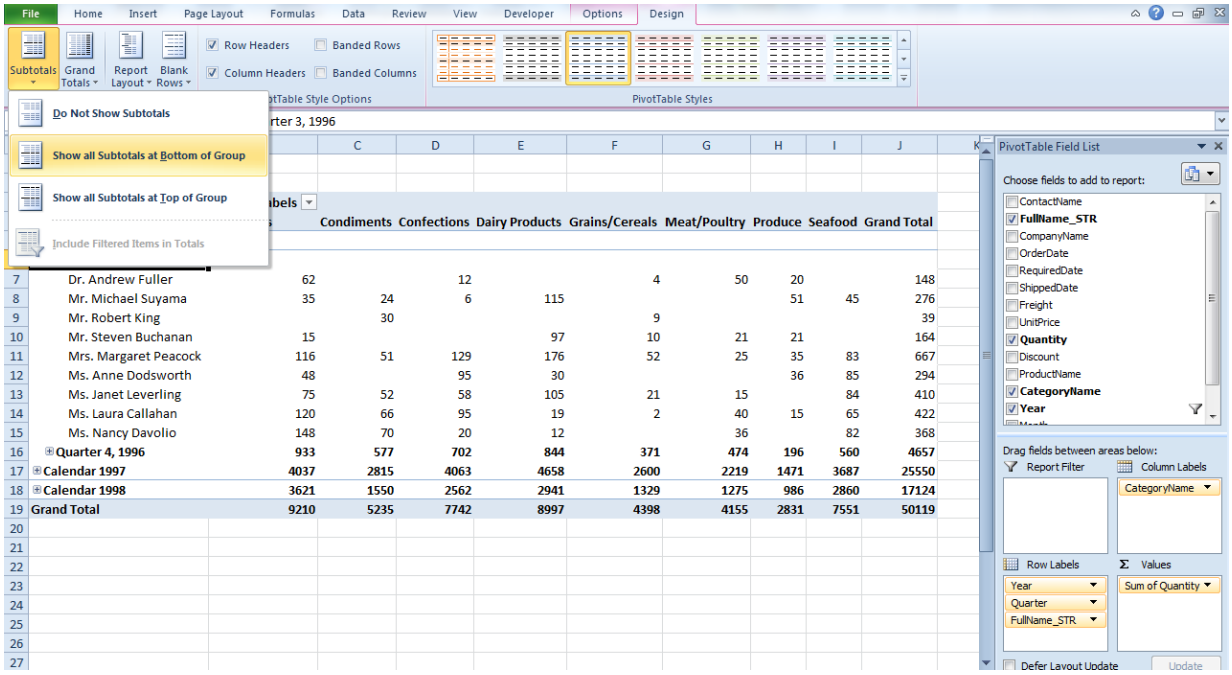
Here is the answer:



Sum of Quantity	Column Labels								
Row Labels	Beverages	Condiments	Confections	Dairy Products	Grains/Cereals	Meat/Poultry	Produce	Seafood	Grand Total
Calendar 1996									
Quarter 3, 1996									
Dr. Andrew Fuller	62		12		4	50	20		148
Mr. Michael Suyama	35	24	6	115			51	45	276
Mr. Robert King		30			9				39
Mr. Steven Buchanan	15			97	10	21	21		164
Mrs. Margaret Peacock	116	51	129	176	52	25	35	83	667
Ms. Anne Dodsworth	48		95	30			36	85	294
Ms. Janet Leverling	75	52	58	105	21	15		84	410
Ms. Laura Callahan	120	66	95	19	2	40	15	65	422
Ms. Nancy Davolio	148	70	20	12		36		82	368
Quarter 3, 1996 Total	619	293	415	554	98	187	178	444	2788
Quarter 4, 1996	933	577	702	844	371	474	196	560	4657
Calendar 1996 Total	1552	870	1117	1398	469	661	374	1004	7445
Calendar 1997	4037	2815	4063	4658	2600	2219	1471	3687	25550
Calendar 1998	3621	1550	2562	2941	1329	1275	986	2860	17124
Grand Total	9210	5235	7742	8997	4398	4155	2831	7551	50119

Here is how it is done:

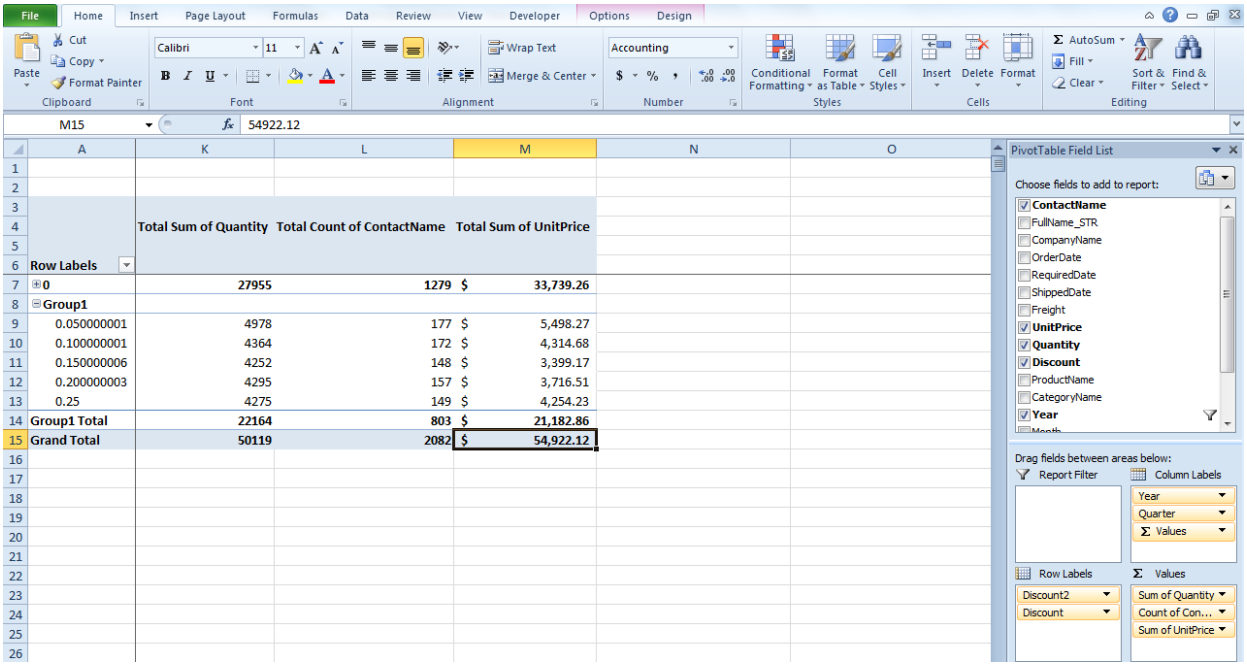
Just drag and drop FullName_STR (Name of the salesperson) to the Row Label. Then, in the PivotTable tools, select the Design menu, in the Design menu select the Subtotals tab and then select the option-Show all Subtotals at Bottom of Group as shown below.



	Condiments	Confections	Dairy Products	Grains/Cereals	Meat/Poultry	Produce	Seafood	Grand Total
Dr. Andrew Fuller	62		12		4	50	20	148
Mr. Michael Suyama	35	24	6	115			51	45	276
Mr. Robert King		30			9				39
Mr. Steven Buchanan	15			97	10	21	21		164
Mrs. Margaret Peacock	116	51	129	176	52	25	35	83	667
Ms. Anne Dodsworth	48		95	30			36	85	294
Ms. Janet Leverling	75	52	58	105	21	15		84	410
Ms. Laura Callahan	120	66	95	19	2	40	15	65	422
Ms. Nancy Davolio	148	70	20	12		36		82	368
Quarter 4, 1996	**933**	**577**	**702**	**844**	**371**	**474**	**196**	**560**	**4657**
Calendar 1997	**4037**	**2815**	**4063**	**4658**	**2600**	**2219**	**1471**	**3687**	**25550**
Calendar 1998	**3621**	**1550**	**2562**	**2941**	**1329**	**1275**	**986**	**2860**	**17124**
Grand Total	**9210**	**5235**	**7742**	**8997**	**4398**	**4155**	**2831**	**7551**	**50119**

Scenario 4: Is there any increase in sales when the products are sold at a discounted rate?

Here is the answer:

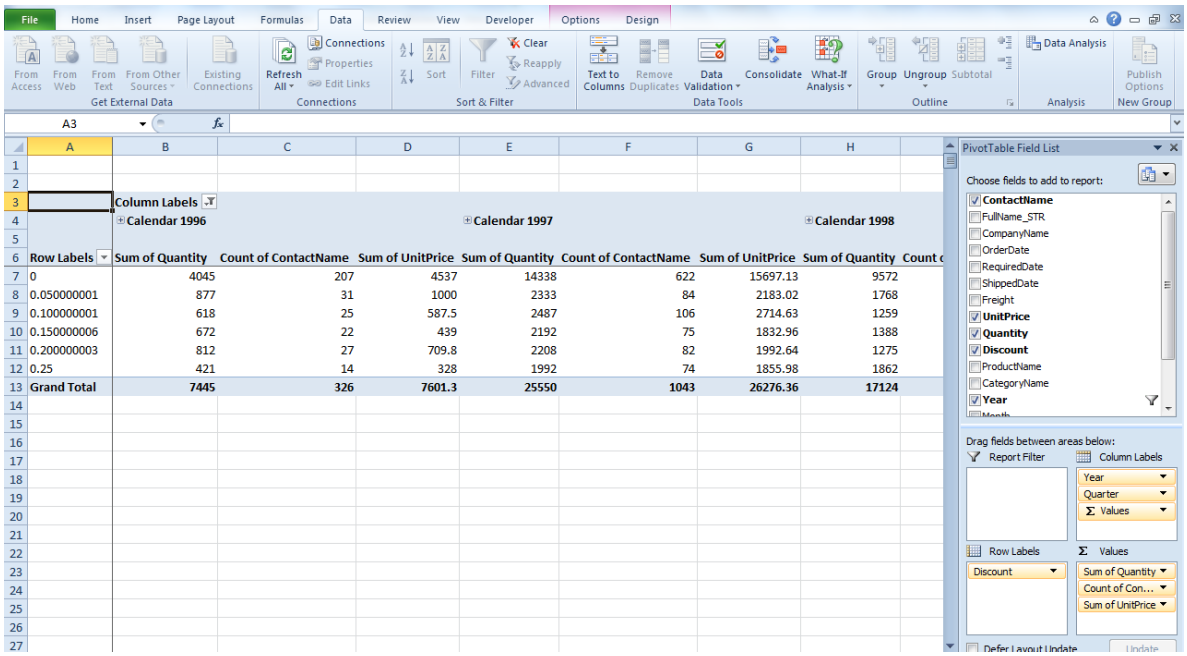


Row Labels	Total Sum of Quantity	Total Count of ContactName	Total Sum of UnitPrice
0	27955	1279	$ 33,739.26
Group1			
0.050000001	4978	177	$ 5,498.27
0.100000001	4364	172	$ 4,314.68
0.150000006	4252	148	$ 3,399.17
0.200000003	4295	157	$ 3,716.51
0.25	4275	149	$ 4,254.23
Group1 Total	**22164**	**803**	**$ 21,182.86**
Grand Total	**50119**	**2082**	**$ 54,922.12**

When you observe the above screenshot, you will realize that even though the total price is higher for products with no discount, the number of units sold per customer is more for products with discount.

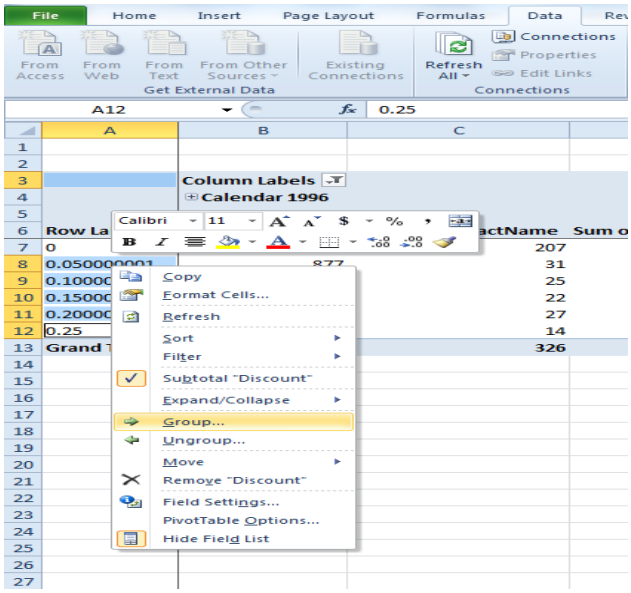
Let us see how we get to this.

In the Pivot Table, drag the Year and the Quarter as Column Labels, Discount as the Row Labels and Quantity and Unit Price as Values. To see how many customers bought something at different discount rates drag the ContactName to Values. You will see the ContactName automatically changing to Count of ContactName. The screenshot is shown below for reference.



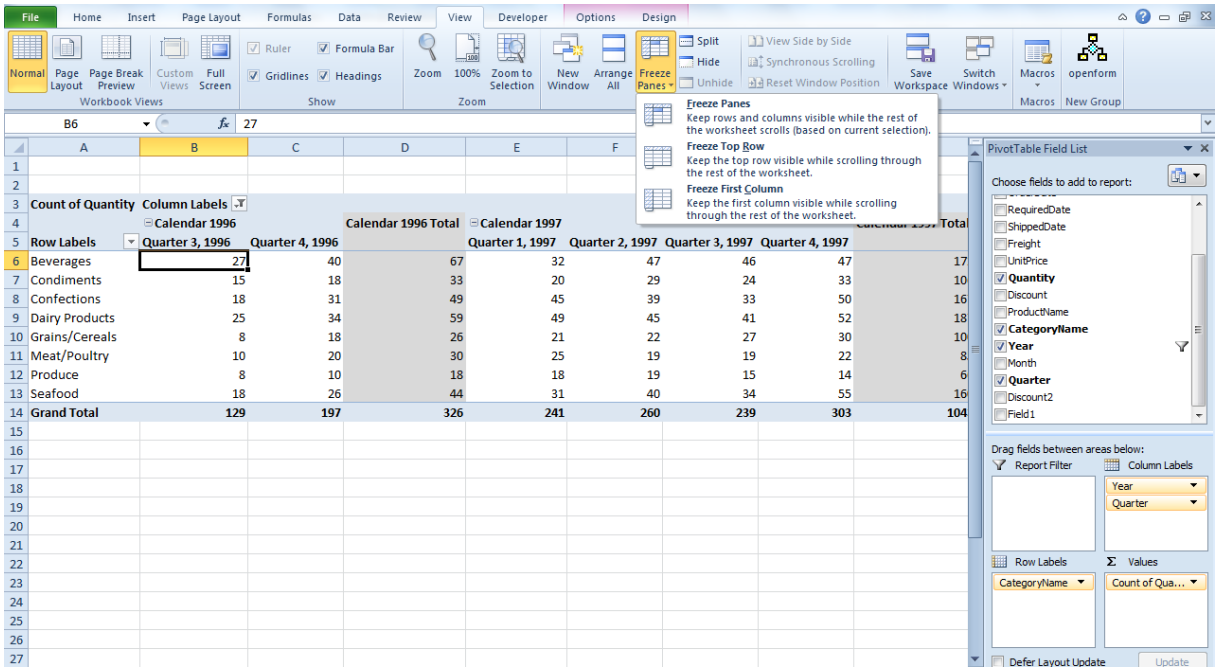
Row Labels	Calendar 1996 Sum of Quantity	Calendar 1996 Count of ContactName	Calendar 1996 Sum of UnitPrice	Calendar 1997 Sum of Quantity	Calendar 1997 Count of ContactName	Calendar 1997 Sum of UnitPrice	Calendar 1998 Sum of Quantity	Calendar 1998 Count
0	4045	207	4537	14338	622	15697.13	9572	
0.050000001	877	31	1000	2333	84	2183.02	1768	
0.100000001	618	25	587.5	2487	106	2714.63	1259	
0.150000006	672	22	439	2192	75	1832.96	1388	
0.200000003	812	27	709.8	2208	82	1992.64	1275	
0.25	421	14	328	1992	74	1855.98	1862	
Grand Total	**7445**	**326**	**7601.3**	**25550**	**1043**	**26276.36**	**17124**	

Except for 0, ctrl select the row label and right click. You will see an option called Group in the menu that appears. Select that option as shown below. Once it is selected, you will be able to see the performance of products with and without discounts.



Note: In the screenshot above, you can also see Sort, Filter, etc. along with other options. You can explore the sort and filter options yourselves and see the results of using them. Some of the other options will be covered later in the chapter.

Tip: Whenever we have a huge number of columns and rows, we will realize that the column labels and row labels will dissappear as we scroll vertically or horizontally or both. Here is how you can avoid it:

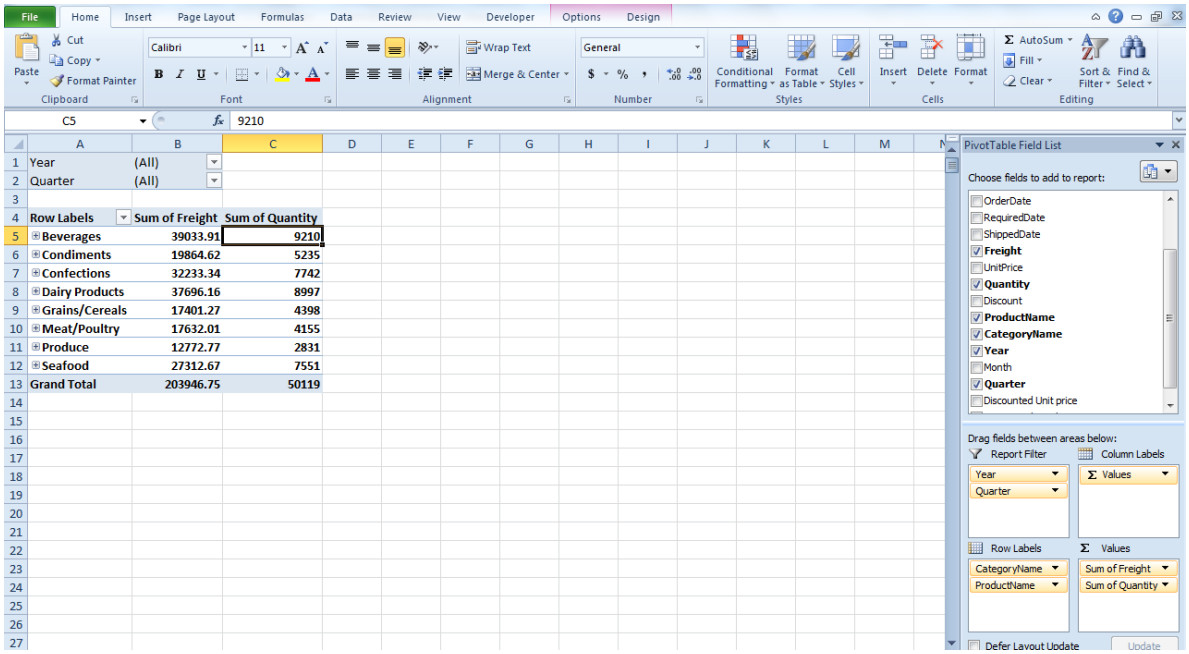


In the screenshot above, if you want to make sure that both CategoryName and Year and Quarter have to stay visible at all times, select the cell that is right below the column label and to the right of the row label (Cell containing the number 27 in this case). Go to the View menu, select the Freeze Panes tab and select the option Freeze Panes.

If you have to freeze only the columns, select the entire row right below the column labels and select the Freeze Panes option. To freeze only the rows, select the entire first right column having the row label and select the Freeze Panes option.

Scenario 5: Report the sales by category and the corresponding freight charges. Filtering should be enabled in the Year and Quarter columns, and the selected Year and Quarter need to be visible.

We will take a slightly different approach to solve this scenario. Please take a look at the below screenshot.

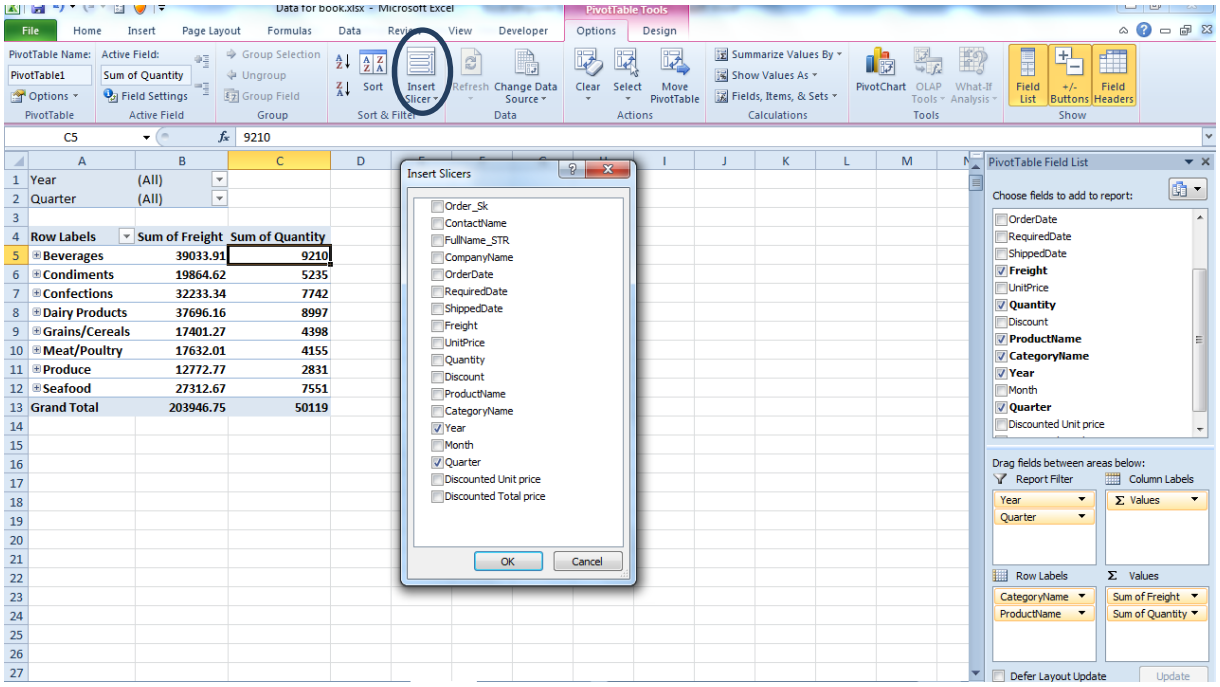


Row Labels	Sum of Freight	Sum of Quantity
Beverages	39033.91	9210
Condiments	19864.62	5235
Confections	32233.34	7742
Dairy Products	37696.16	8997
Grains/Cereals	17401.27	4398
Meat/Poultry	17632.01	4155
Produce	12772.77	2831
Seafood	27312.67	7551
Grand Total	203946.75	50119

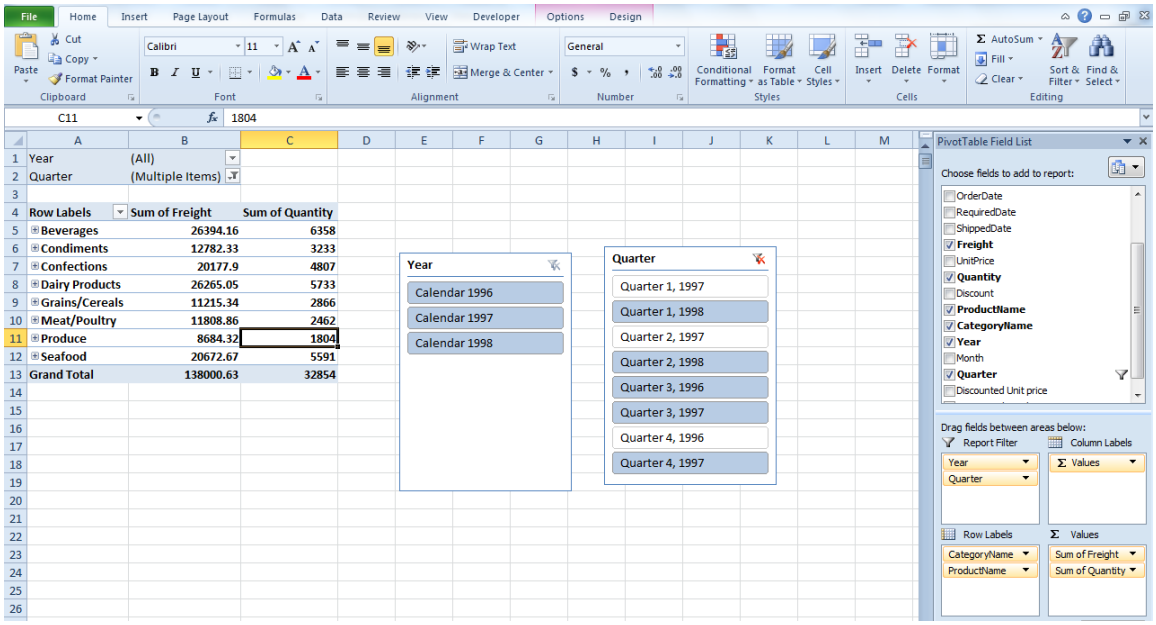
By using CategoryName and ProductName in the Row Labels, Freight and Quantity as values and Year and Quarter as Report Filters we will be able to get the answer. But if we filter either Year or Quarter we will not be able to see which Year/Quarter have been selected and which are not.

To see the individual Years and Quarters, the concept of slicers is used. Lets see how it is done:

Click on the Options menu which gets highlighted under PivotTable Options. Click on the Insert Slicer. You will get to see the Insert Slicers window as shown. Select the Year and Quarter and click OK.



If you see the below screenshot, you will notice that I have selected all the years and I have selected Quarter 4, 1996, Quarter 1, 1997, Quarter 2, 1997. You may also notice the data has changed and that data is the data for the years and quarters we have selected.

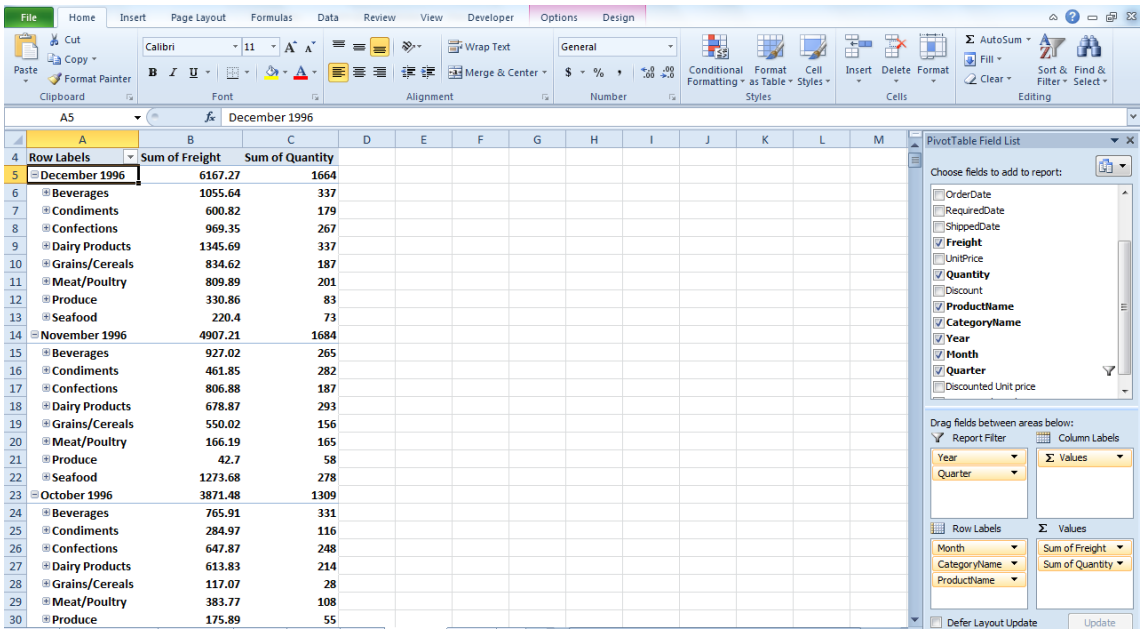


Year	(All)	
Quarter	(Multiple Items)	
Row Labels	Sum of Freight	Sum of Quantity
Beverages	26394.16	6358
Condiments	12782.33	3233
Confections	20177.9	4807
Dairy Products	26265.05	5733
Grains/Cereals	11215.34	2866
Meat/Poultry	11808.86	2462
Produce	8684.32	1804
Seafood	20672.67	5591
Grand Total	138000.63	32854

Note: The Slicer option is available only in Excel 2010

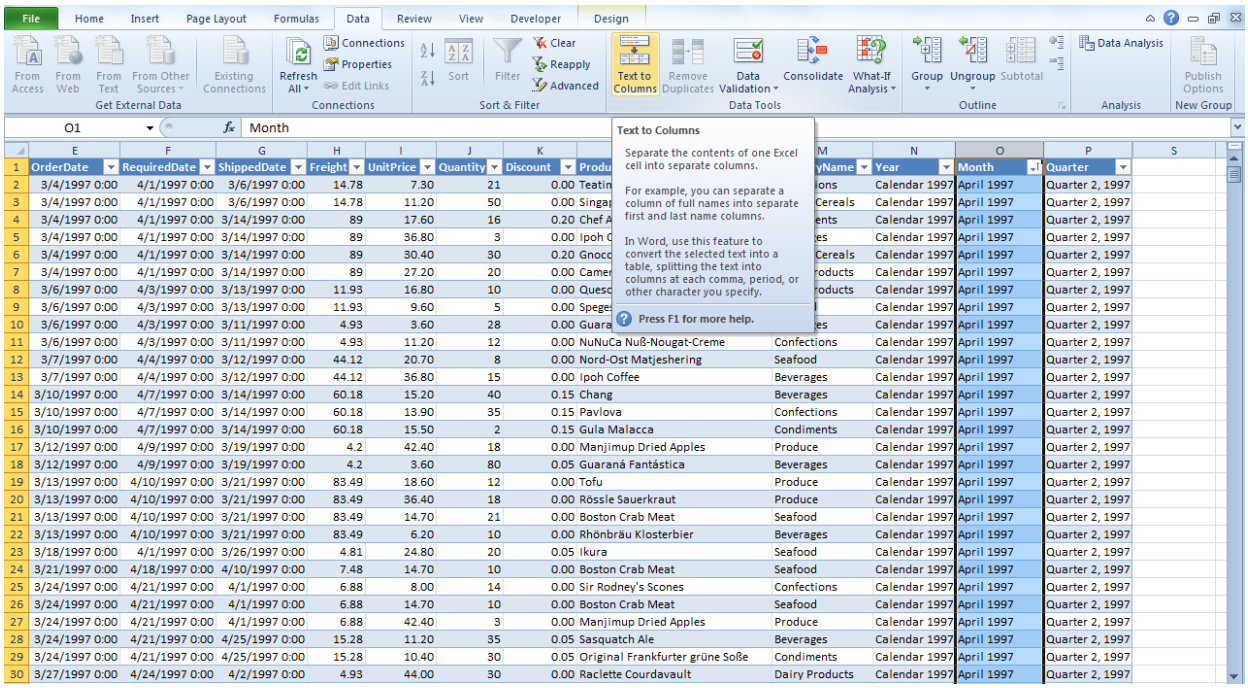
Scenario 6: Sort the Sales data in terms of Year, Quarter and Month.

This sounds very easy to achieve But it is not that easy. Just observe the below screenshot. We will see that the months are alphabetically sorted which is logically incorrect. Let us sort it correctly.

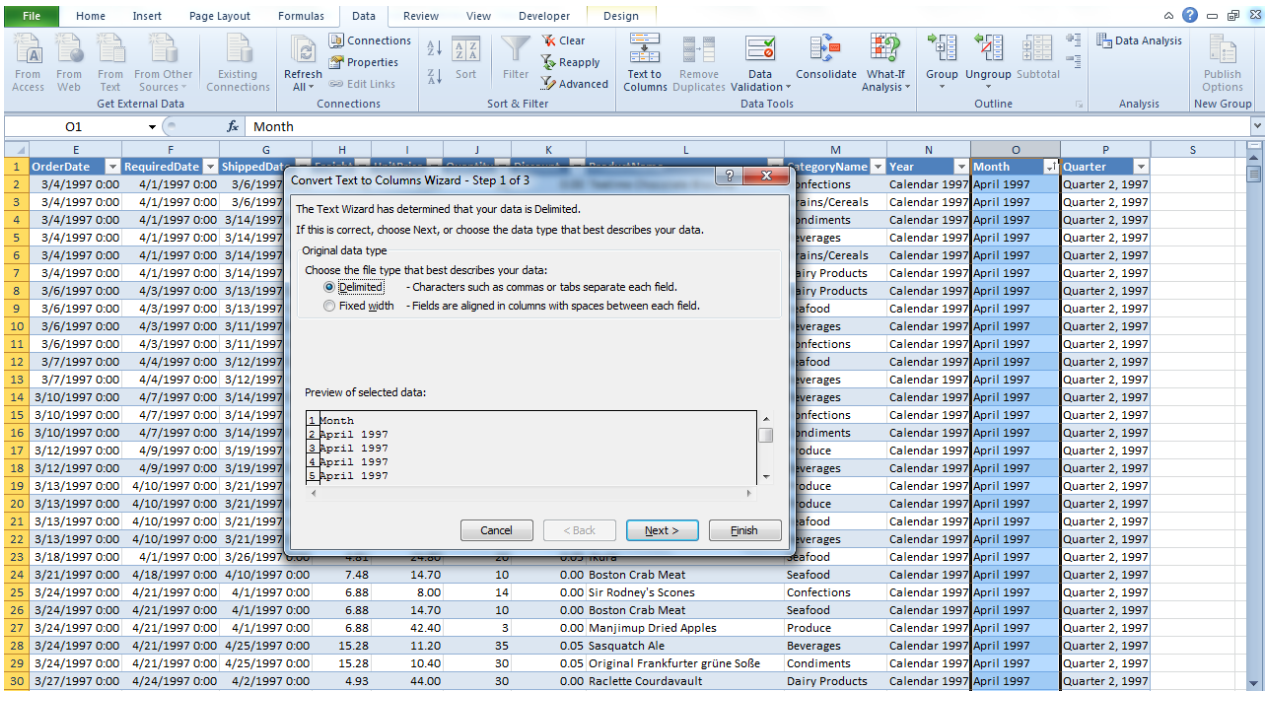


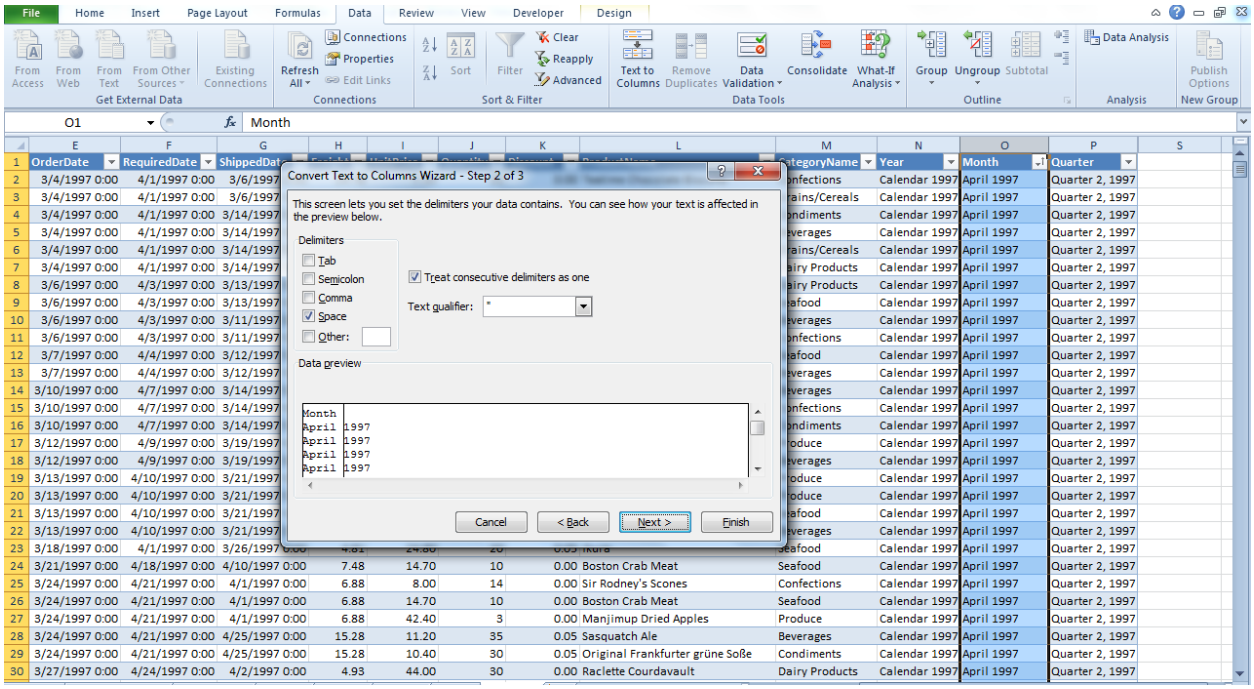
Row Labels	Sum of Freight	Sum of Quantity
December 1996	6167.27	1664
Beverages	1055.64	337
Condiments	600.82	179
Confections	969.35	267
Dairy Products	1345.69	337
Grains/Cereals	834.62	187
Meat/Poultry	809.89	201
Produce	330.86	83
Seafood	220.4	73
November 1996	4907.21	1684
Beverages	927.02	265
Condiments	461.85	282
Confections	806.88	187
Dairy Products	678.87	293
Grains/Cereals	550.02	156
Meat/Poultry	166.19	165
Produce	42.7	58
Seafood	1273.68	278
October 1996	3871.48	1309
Beverages	765.91	331
Condiments	284.97	116
Confections	647.87	248
Dairy Products	613.83	214
Grains/Cereals	117.07	28
Meat/Poultry	383.77	108
Produce	175.89	55

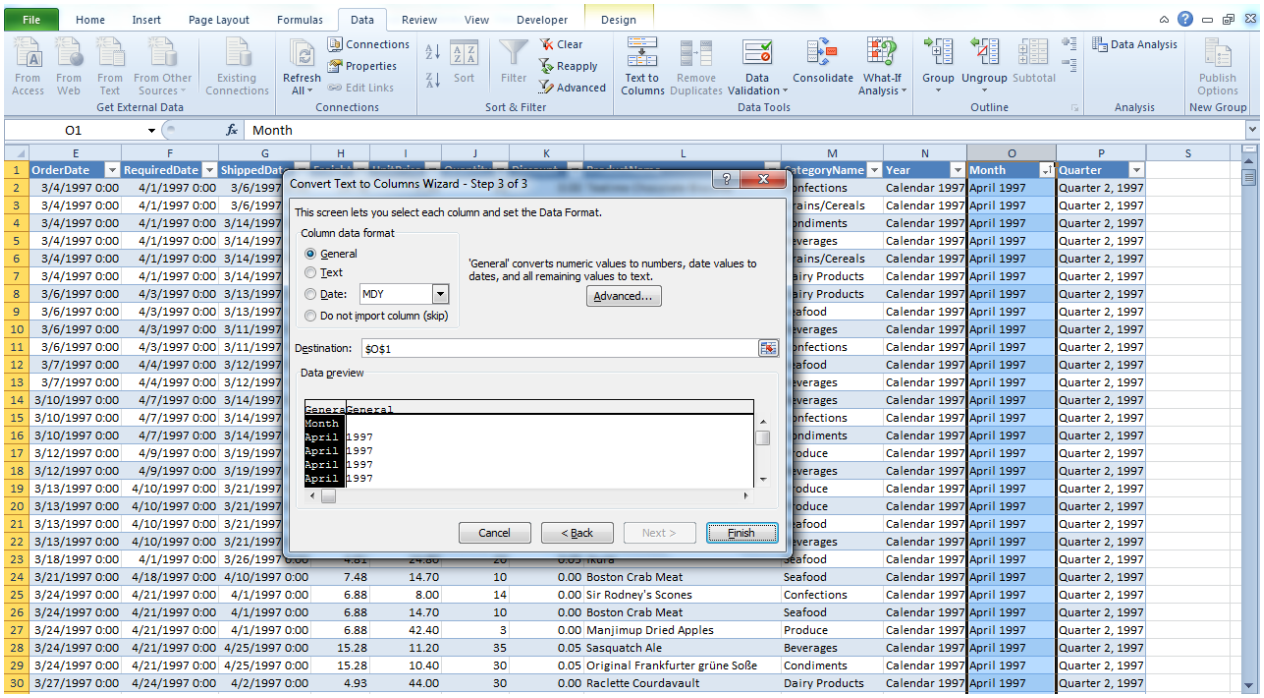
Go to the Orders sheet and select the Month column. Then select the Data menu and click on the Text to Columns tab as shown below.



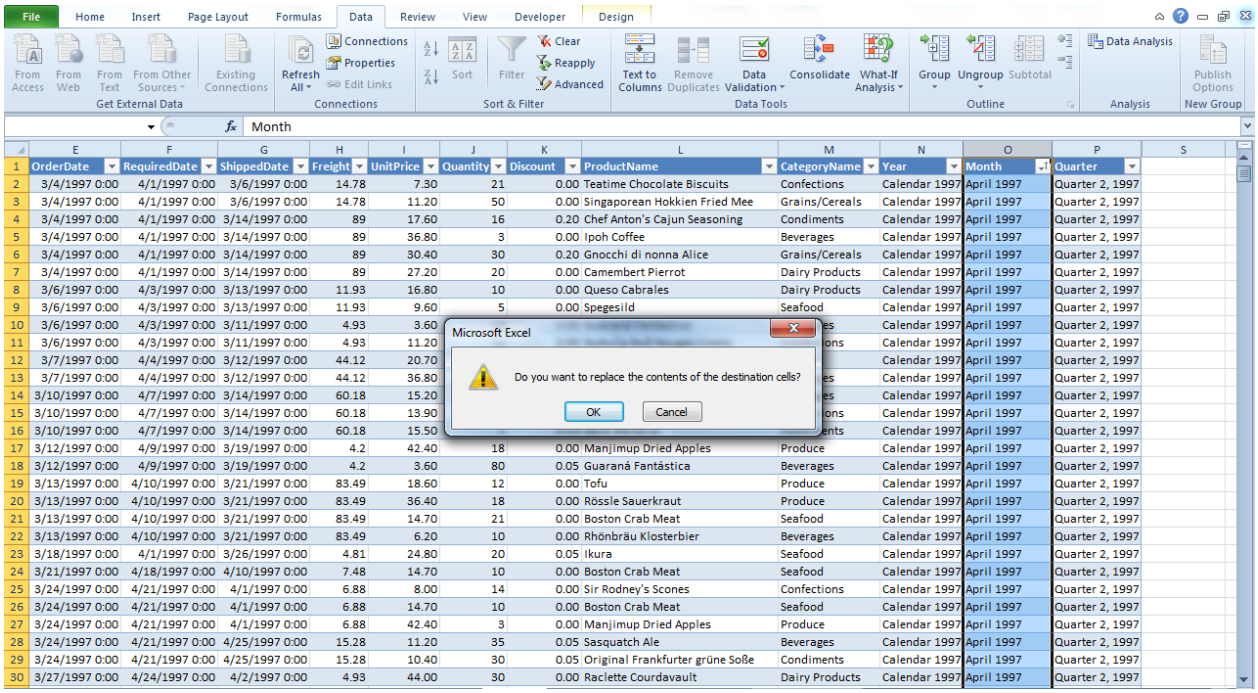
Follow the next 3 steps as shown in the below 3 screenshots:



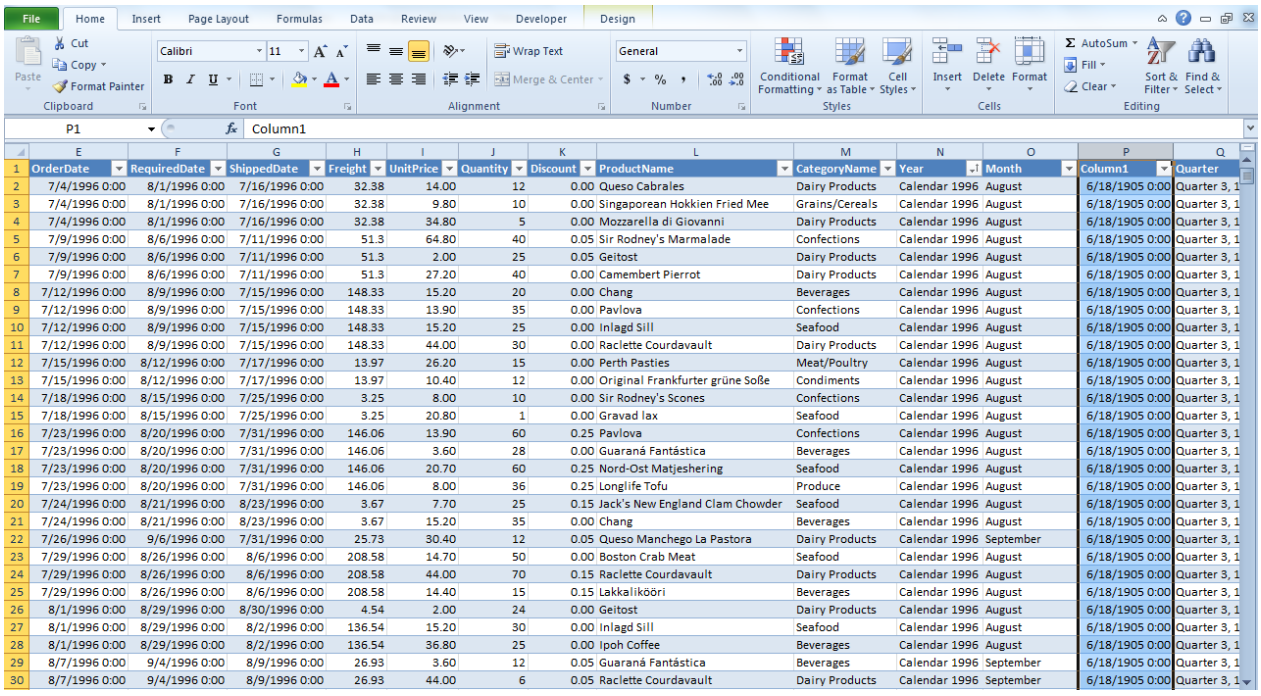




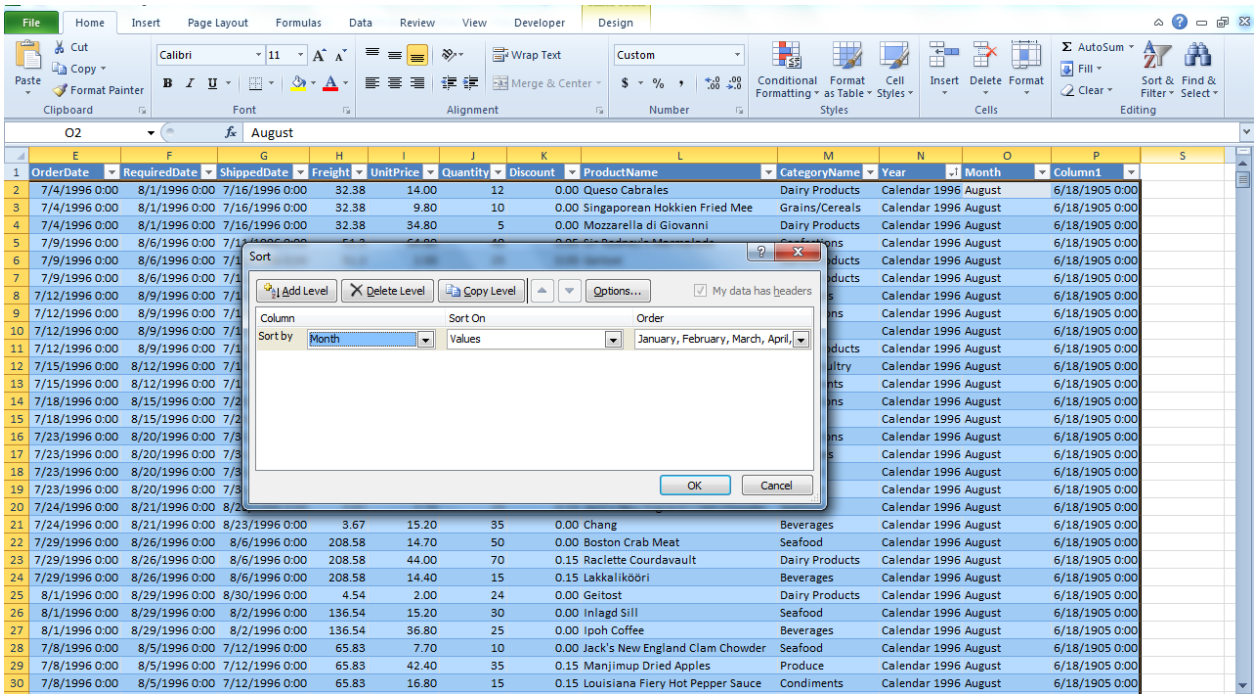
The message asked if the destination column already has the data is shown in the screenshot below.



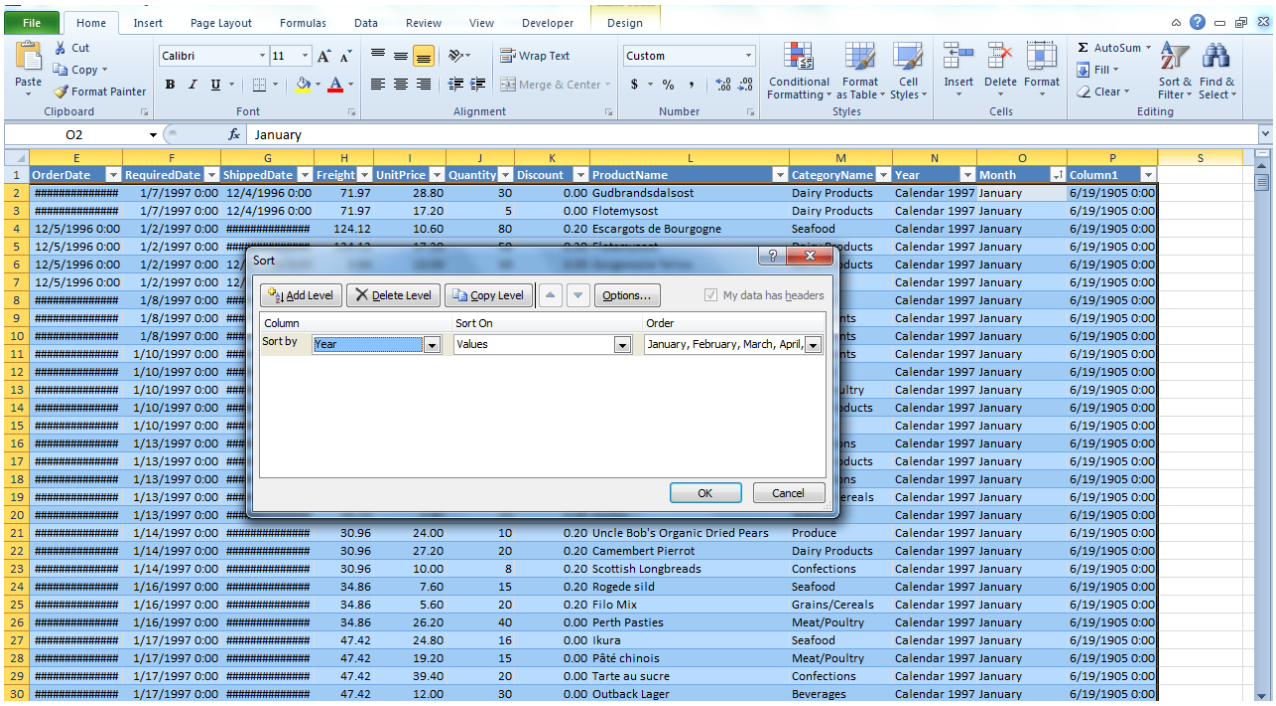
While splitting the components of a cell make sure that the Quarter column is cut and pasted in the neighboring column because the two outputs of the split will occupy the adjacent columns and if the Quarter coulmn is not moved as described, it will be overwritten by the data as shown.



Once the above steps are achieved, sort the Month column by Month first as shown below.



Then sort the same Month column by Year as shown below.



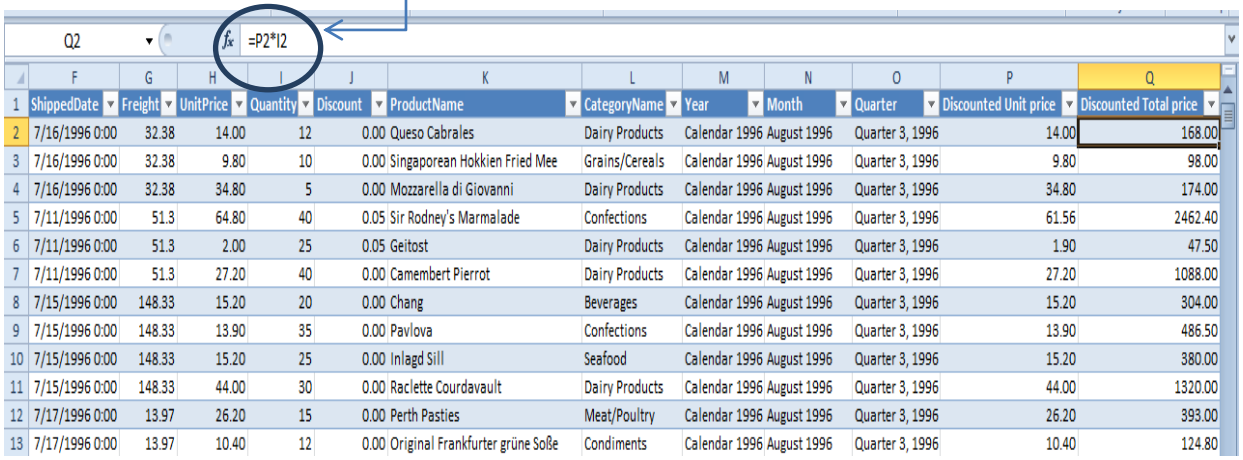
If analysis based on the discounts is required, add 2 columns as explained below and create a new Pivot Table on this sheet.

First go to the Orders sheet and create 2 new columns as shown:

Column 1: Discounted Unit Price: This column will contain the unit prices after discount.

So the formula would be (1-Discount)(Unit Price). In terms of excel columns the formula would be (1-J2)*(H2).

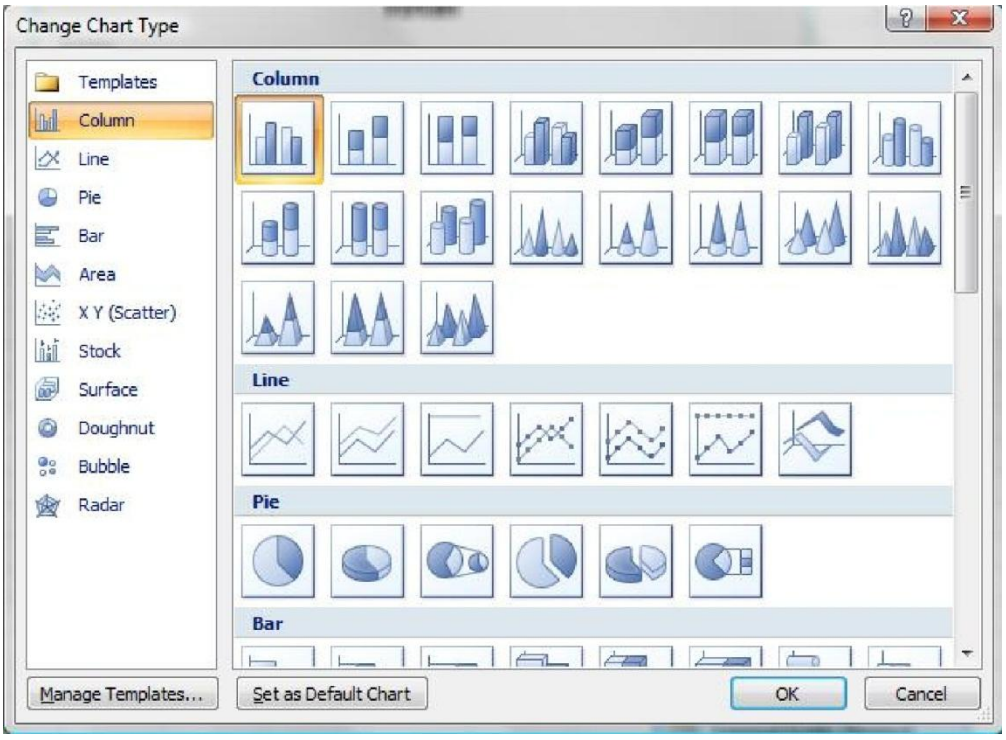
Column 2: Discounted Total Price: This column would contain discounted unit price multiplied by quantity. In terms of excel columns, it would be P2*I2.



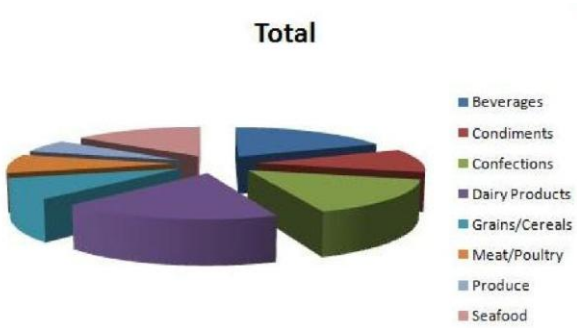
	F	G	H	I	J	K	L	M	N	O	P	Q
1	ShippedDate	Freight	UnitPrice	Quantity	Discount	ProductName	CategoryName	Year	Month	Quarter	Discounted Unit price	Discounted Total price
2	7/16/1996 0:00	32.38	14.00	12	0.00	Queso Cabrales	Dairy Products	Calendar 1996	August 1996	Quarter 3, 1996	14.00	168.00
3	7/16/1996 0:00	32.38	9.80	10	0.00	Singaporean Hokkien Fried Mee	Grains/Cereals	Calendar 1996	August 1996	Quarter 3, 1996	9.80	98.00
4	7/16/1996 0:00	32.38	34.80	5	0.00	Mozzarella di Giovanni	Dairy Products	Calendar 1996	August 1996	Quarter 3, 1996	34.80	174.00
5	7/11/1996 0:00	51.3	64.80	40	0.05	Sir Rodney's Marmalade	Confections	Calendar 1996	August 1996	Quarter 3, 1996	61.56	2462.40
6	7/11/1996 0:00	51.3	2.00	25	0.05	Geitost	Dairy Products	Calendar 1996	August 1996	Quarter 3, 1996	1.90	47.50
7	7/11/1996 0:00	51.3	27.20	40	0.00	Camembert Pierrot	Dairy Products	Calendar 1996	August 1996	Quarter 3, 1996	27.20	1088.00
8	7/15/1996 0:00	148.33	15.20	20	0.00	Chang	Beverages	Calendar 1996	August 1996	Quarter 3, 1996	15.20	304.00
9	7/15/1996 0:00	148.33	13.90	35	0.00	Pavlova	Confections	Calendar 1996	August 1996	Quarter 3, 1996	13.90	486.50
10	7/15/1996 0:00	148.33	15.20	25	0.00	Inlagd Sill	Seafood	Calendar 1996	August 1996	Quarter 3, 1996	15.20	380.00
11	7/15/1996 0:00	148.33	44.00	30	0.00	Raclette Courdavault	Dairy Products	Calendar 1996	August 1996	Quarter 3, 1996	44.00	1320.00
12	7/17/1996 0:00	13.97	26.20	15	0.00	Perth Pasties	Meat/Poultry	Calendar 1996	August 1996	Quarter 3, 1996	26.20	393.00
13	7/17/1996 0:00	13.97	10.40	12	0.00	Original Frankfurter grüne Soße	Condiments	Calendar 1996	August 1996	Quarter 3, 1996	10.40	124.80

Formatting of PivotCharts

The PivotChart can be changed to any chart type offered by Excel. In the ribbon, click on the **Design** tab, and then select **Change Chart Type** from the *Type* group.



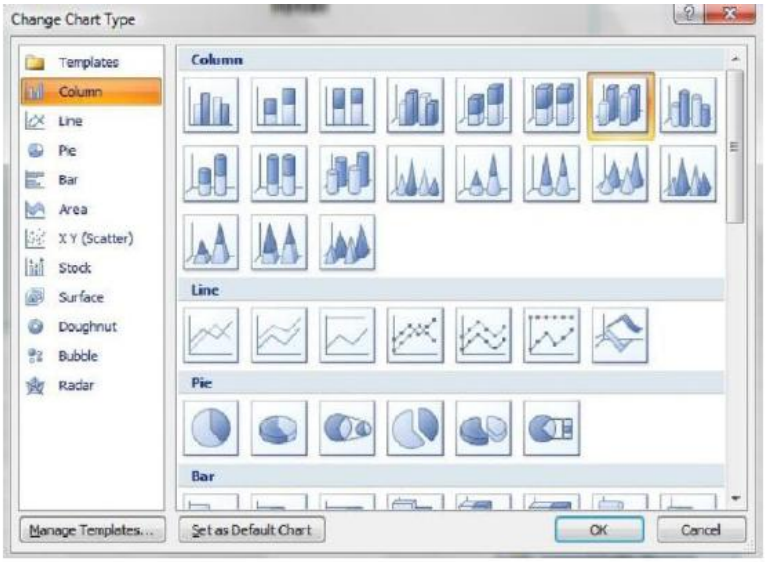
Below is an example of the PivotChart using a pie chart.



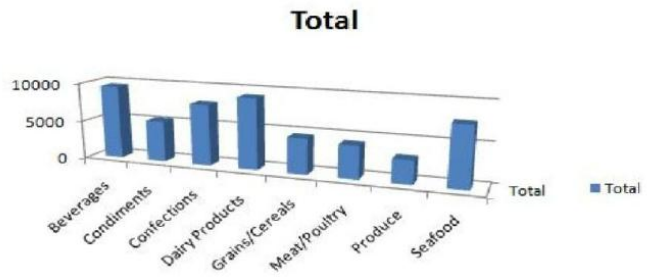
In the next section, we'll explore some other customizations that can be used for a PivotChart that may be useful.

Customizing a PivotChart

1. Go to the worksheet containing the PivotChart, and select the PivotChart. You should be able to see the **Design**, **Layout**, **Format**, and **Analyze** tabs on the ribbon.
2. Click on the **Design** tab, and select **Change Chart Type** from the *Type* group. Select the **3-D Column** chart.



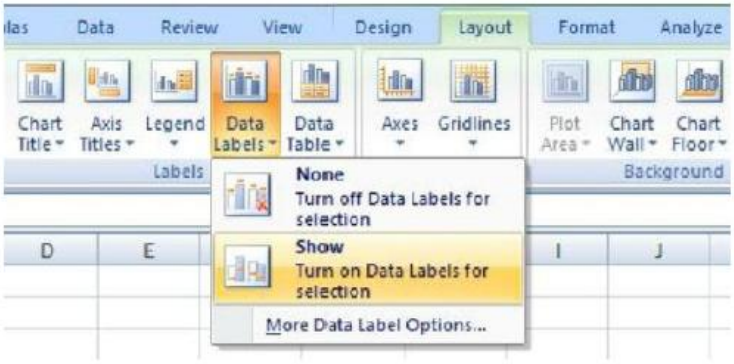
3. The PivotChart should automatically update itself using the new chart type.



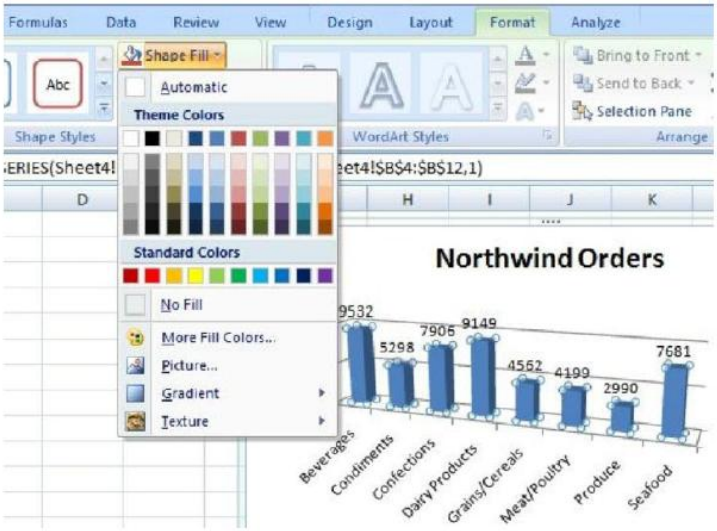
4. Update the title on the chart by clicking on the chart title, and the replacing the current title of

Total with **Northwind Orders**.

5. The numeric values that each column represents above the respective column can be shown by displaying the data labels. To do this, go to the **Layout** tab on the ribbon, select **Data Labels** from the *Labels* group, and then choose **Show**.



6. Finally, change the color of the columns of the chart. To do this, first click on one of the columns within the PivotChart. Then, go to the **Format** tab on the ribbon, select **Shape Fill** from the *Shape Styles* group, and choose the desired color. For the purpose of this exercise, let's choose red.



7. Our PivotChart should now look similar to the one below

Unsolved Exercises

1. The manager wants to see the percentage of difference of product sales from quarter to quarter in a year and also from year to year.
2. He also wants to analyze whether the discount on products increased the sales of the products as compared to their sales without discounts.
3. The CEO of Northwind wants to analyze the growth trend of sales to check whether the sales have increased or decreased over time.
4. He also wants to forecast the volume of products that are going to be sold in the next 3 years.
5. The CEO needs to analyze the shipper's efficiency to offer incentives to the best performing shippers.
6. He has to prepare a list of "Top 10 Customers" depending on the volume of sales they have provided to the company over the years.