

## Lab 17 More on joins and introduction to LOD expressions

Learning objectives:

- one-to-many vs. many-to-many inner joins
- default aggregation is not always desirable
- use FIXED and INCLUDE expression to control the level of aggregation
- create groups
- create parameters

1. One-to-many inner join
2. Many-to-many inner join
3. How do we join Bibitor's Sales and Purchases tables?

### 4. Connecting to mock sales database

Let's connect Tableau to a toy excel dataset called mock sales. It has two tables: sales and stores. You can see the two tables below.

#	sales	sales	#	#	#	#	#	#	#
Product ID	Classification	Store ID	Price	Quantity	Unit Cost of Goods	Store ID	Sq Ft	Rent	
1	Wine	1	10	10	5	1	100	50	
2	Wine	1	12	5	5	1	100	50	
1	Wine	1	15	1	5	1	100	50	
1	Wine	2	12	10	6	2	200	80	
3	Liquor	2	12	5	6	2	200	80	

Below is what the underlying data looks like, if we connect the two tables via *Store ID*:

#	sales	sales	sales	#	#	#	#	#	#
Product ID	Classification	Store ID	Price	Quantity	Unit Cost of Goods	Store ID (stor...	Sq Ft	Rent	
1	Wine	1	10	10	5	1	100	50	
2	Wine	1	12	5	5	1	100	50	
1	Wine	1	15	1	5	1	100	50	
1	Wine	2	12	10	6	2	200	80	
3	Liquor	2	12	5	6	2	200	80	

### 5. Calculating gross profit

Let's calculate *Revenue* as Price\*Quantity, and *Cost of Goods* as Unit Cost of Goods\*Quantity. We can calculate *Gross Profit* as *Revenue* minus *Cost of Goods*. We can plot gross revenue against store square footage. What is SUM or AVG the appropriate level of aggregation for store square footage?

Store ID	Gross Profit	Sq Ft	Store ID	Gross Profit	Avg. Sq Ft
1	95.0	300.0	1	95.0	100.0
2	90.0	400.0	2	90.0	200.0

## 6. Calculating net profit

Suppose we wanted to calculate net profit which is gross profit minus store rent. Let's try creating Net Profit as Gross Profit - rent. Here is what we get when we use SUM and when we use AVG aggregation for Net Profit. Are either of these correct?

Store ID	Gross Profit	Net Profit	Store ID	Gross Profit	Avg. Net Profit
1	95.0	-55.0	1	95.0	-18.3
2	90.0	-70.0	2	90.0	-35.0

The problem with calculating net profit is that some of the observations we want to aggregate using SUM and others using AVG. We can accomplish this using a level of detail expression (LOD). Tableau's LOD expressions include three keywords: FIXED, INCLUDE, and EXCLUDE. INCLUDE forces Tableau to include a dimension and do a specified aggregation, before doing the aggregation specified in the view.

$$\text{Net Profit} = \{INCLUDE [Store ID]: SUM([Gross Profit]) - AVG([Rent])\}$$

Here is what we get when we aggregate the new Net Profit using SUM and AVG:

Store ID	Gross Profit	Net Profit	Store ID	Gross Profit	Avg. Net Profit
1	95.00	45.00	1	95.00	45.00
2	90.00	10.00	2	90.00	10.00

What happens when we remove the Store ID dimension and keep SUM aggregation?  
 What happens when we remove the Store ID dimension and use AVG aggregation?

## 7. Calculating a product share in its category

Suppose we wanted to calculate share of each product in its category (wine or liquor). We would need the sum of sales across all products in each category. We can accomplish this with an LOD expression:

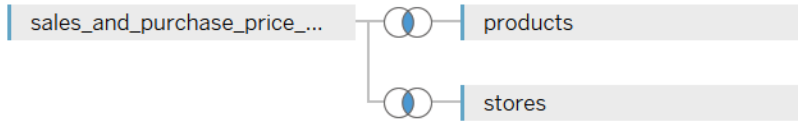
$$\text{Share in Category} = [Revenue] / \{FIXED [Classification]: SUM([Revenue])\}$$

Classific..	Product ID	Share in Category
Liquor	3	1.0000
Wine	1	0.7966
	2	0.2034

Can we turn the table into a viz?

## 8. Create groups

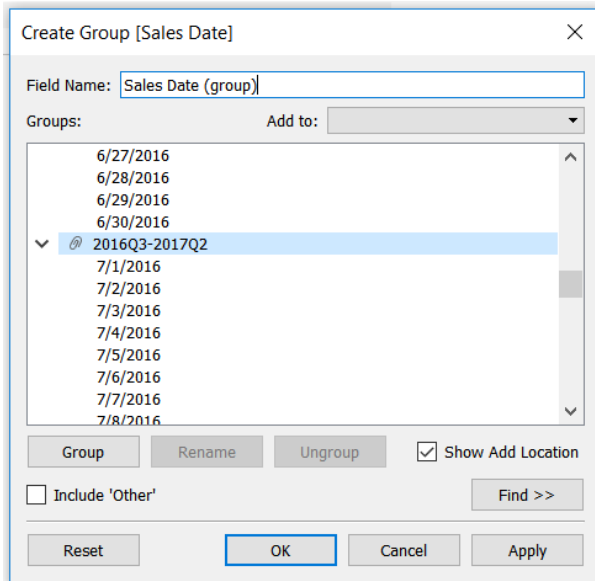
Let's connect to the bibitor database and sale and purchase price table connected to product and store tables.



Let's create some profitability measures: *Cost of Goods* and *Gross Profit*, and *Gross Margin*. Let's put them in a table by date:

Year of S..	Sales Dollars	Cost of Goods	Gross Profit
2015	43,973,848	30,528,709	13,445,139
2016	83,733,035	58,068,337	25,664,699
2017	38,676,074	26,750,985	11,925,089

The problem with the above table is that we only have parts of years 2015 and 2017 and thus these profits are not comparable to 2016. We can fix it by creating groups for the date variable. We can create two groups 2015Q3 through 2016Q2 and 2016Q3 through 2017Q2. We can do this by clicking on Date and 'Create' and 'Group'. It will open a window below listing all dates. We can highlight the dates we want to group and then name each of the two groups. This will give us two full years of data.



Sales Date (group)	Sales Dollars	Cost of Goods	Gross Profit
2015Q3-2016Q2	80,031,757	55,199,264	24,832,492
2016Q3-2017Q2	86,351,201	60,148,767	26,202,434

## 9. Create Parameters

Our data does not have information on Rent paid for each store. However, we do have information on the square footage of each store and we can estimate rent per square foot of commercial retail space. We could hard code a certain figure into the viz, but that would make it difficult to see how profitability changes with alternative assumptions for rent. Therefore, we will create a parameter that contains the rent and Tableau will allow us to us to change that parameter. We can click on “show parameter control” so that we can change this value on the fly in the viz.

The screenshot shows the 'Edit Parameter [Annual Rent per SQFT]' dialog box. The 'Name' field contains 'Annual Rent per SQFT'. Under 'Properties', 'Data type' is set to 'Float', 'Current value' is '60', and 'Display format' is 'Automatic'. The 'Allowable values' section has 'Range' selected. Under 'Range of values', 'Minimum' is 10, 'Maximum' is 100, and 'Step size' is 10. There are 'Set from Parameter' and 'Set from Field' buttons, and 'OK' and 'Cancel' buttons at the bottom.

Let’s create Annual Rent measure as an LOD expression of average square footage for each store times the parameter stored in *Annual Rent per SQFT* .:

*{INCLUDE [Store id]:AVG([Sq Ft])\*[Annual Rent per SQFT]}*

Let’s include it into our profitability table:

Sales Date (group)	Sales Dollars	Cost of Goods	Gross Profit	Annual Rent
2015Q3-2016Q2	80,031,757	55,199,264	24,832,492	7,320,000
2016Q3-2017Q2	86,351,201	60,148,767	26,202,434	7,320,000

What happens when we add Store ID dimension to the rows shelf?

## 10. Calculating net profit

We are now ready to calculate Net Profit using the following LOD expression.

*{INCLUDE [Store id]: SUM([Gross Profit])- AVG([Annual Rent])}*

Sales Date (group)	Sales Dollars	Cost of Goods	Gross Profit	Annual Rent	Net Profit
2015Q3-2016Q2	80,031,757	55,199,264	24,832,492	7,320,000	17,512,492
2016Q3-2017Q2	86,351,201	60,148,767	26,202,434	7,320,000	18,882,434

## Illustrating the difference between FIXED and INCLUDE

Date	Store ID	Product ID	Price
1	NY	A	2
2	NY	A	3
2	NY	B	10
1	LA	A	7
2	LA	B	12

### No Dimension

Avg. Price	Avg. Price INCLUDE Product ID	Avg. Price FIXED Product ID
6.8000	7.5000	7.5000

### Product ID

Product ID	Avg. Price	Avg. Price FIXED Product ID	Avg. Price INCLUDE Product ID
A	4.000	4.000	4.000
B	11.000	11.000	11.000

### Store ID

Store ID	Avg. Price	Avg. Price INCLUDE Product ID	Avg. Price FIXED Product ID
NY	5.000	6.250	7.500
LA	9.500	9.500	7.500

### Date

Date	Avg. Price	Avg. Price INCLUDE Product ID	Avg. Price FIXED Product ID
1	4.500	4.500	4.000
2	8.333	7.000	7.500

11. dkfjdl

How do we merge sale and purchases exercise:  
gross margin at the product level.

12. f

13. f

#### Bibitor part I

1. Download and install Tableau on your computer. You can request one-year free student licence [here](#).
2. Watch [this "Getting Started" video](#) on Tableau's training website. You may need to sign-in with your student Tableau account. Watch the video very carefully. The tasks I ask you to do below follow very closely what is happening in the video.
3. You can connect to it by clicking in Microsoft SQL Server.
  - a. Drag tables "sales" and "products" onto the canvas. Keep the default "inner" join.
4. Click on "Sheet 1" to go to the worksheet.
5. What is the total dollar value of all sales in our data?
6. Create a sheet that shows monthly sales over the two years? What were the sales in June 2017? What is the overall trajectory?
7. The variable classification is 1 for liquor and 2 for wine. Create a new sheet that shows the seasonality of wine compared to liquor.

#### Bibitor Part II

Create a dashboard showing an analysis of Bibitor's sales (use only sales, products and stores tables). Upload your Tableau workbook to Nexus.

- aliases
- show exclude
- show q to q calc
- split by team member

1. Which product has the highest sales?
2. Which product has the highest sales growth from Q2 2016 to Q2 2017?
3. Does "Description" uniquely identify products?

4. Notice that our bibitor database also has a table on purchases. It includes records of all the invoices paid by bibitor for its product. Explain how this additional information might be useful. Explain how you would combine it with our sales data.
- 5.

Bibitor Part III