

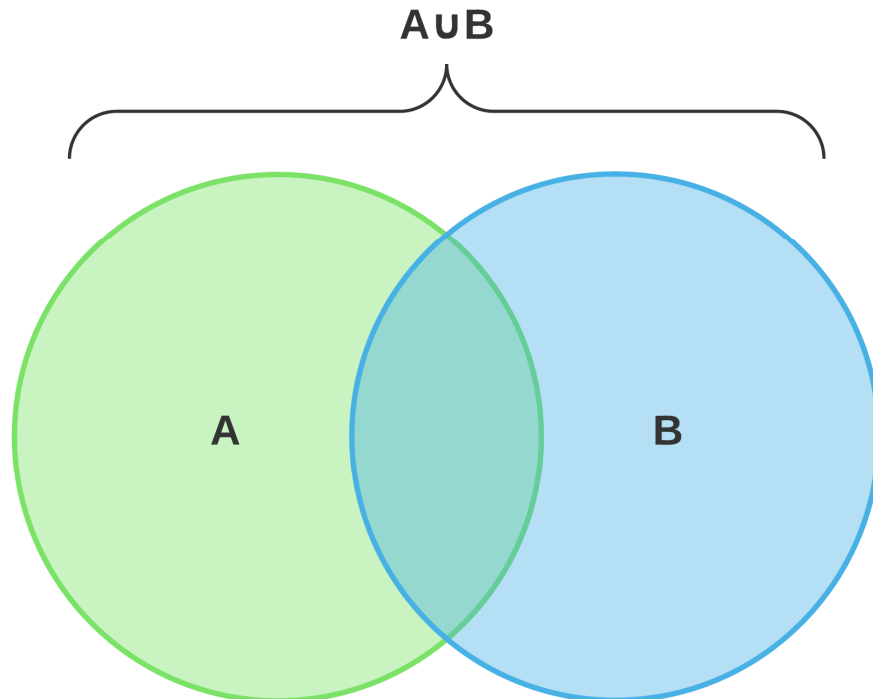


LEAD WITH DATA™

# Qlik® Sense

(September 2019 Edition)

## Set Analysis - Hands on Workshop



# Table of Contents

Introduction .....	4
It was the best of times, it was the worst of times.....	4
Basics of Set Analysis .....	4
States.....	5
Section 1 – Just getting started.....	6
Let’s try it .....	6
{ \$ }.....	6
{ 1 }.....	6
{ Green, Grey and White }.....	7
Section 2 – Now we are getting warmer .....	8
Forcing Field Values .....	8
"Oh those Users".....	9
Multiple Values .....	9
Section 3 – Making progress now .....	11
Multi Field Set Analysis .....	11
Beyond KPI's.....	12
Section 4 - Visualizing Context.....	14
Let’s think together.....	15
Section 5 – Working with text values.....	16
Set Analysis with Text Values.....	16
Multiple Text Values .....	17
Wildcards .....	18
A little practice goes a long way .....	18
Section 6 – Coding for flexibility .....	20
Put Your Thinking Caps On.....	20
Flexibility in Assignment .....	20
Relationship to User Input .....	20
Validating Results.....	21
Real World.....	21
Section 7 – You don't know what you don't you know .....	22
Seriously we don't always know .....	22
Crazy Syntax .....	24
Pop Quiz Superstar.....	24
E(xcluded).....	25
Possible based on Selections .....	25

Section 8 – Advanced material ..... 26

- Alternate States ..... 26
- { StateName } ..... 27
- Bar Chart ..... 29
- Visualizing Variance ..... 30
- Let’s think together..... 30

Section 8 – Analysis of Sets ..... 32

- Intersection ..... 32
- Difference between sets ..... 35
- Exclusion ..... 36
- Total Spend for Set Operations..... 37

Section 9 – The easy button..... 38

Appendix ..... 39

Set Analysis Cheat Sheet..... 40

## Introduction

### It was the best of times, it was the worst of times

One of the most awesome features of Qlik is that you can do analysis on "sets" of data.

One of the most feared features of Qlik is that you can do analysis on "sets" of data.

By that I mean that you start out seeing these ridiculous, Greek looking, formulas that make no sense at all. Which in turn leads to fear. But the great news is that they are possible. Now it's just a matter of learning in small pieces so that you get comfortable. To that you are encouraged NOT TO READ AHEAD. Let the learning process take place without instilling fear into your brain.

### Basics of Set Analysis

At the very heart of Set Analysis is the simple math concept of a "set." In math terms a set is a collection of items denoted with set "{ }" brackets. If I had 5 values for ages, 1, 7, 19, 23 OR 33 I would denote that set as:

**Ages = {1, 7, 19, 23, 33}**

While we don't often think about it these sets are at the very core of how Qlik operates. Seriously!

When you click on any filter or search and find any values, you are in every sense of the word creating a set of values. For instance if you see a list of values for a field called REGION and you click on East and West you have told Qlik "I want to use the set of values where REGION includes East OR West." In other words:

**Region = { "East", "West" }**

That's not too hard. Now try to interpret the following line thinking of SET terminology:

**Organizations = { "Department of Defense", "NASA" }**

If you translated that line in your mind to "Organizations equals the set of values Department of Defense OR NASA" you are well on your way to lighting up this workshop.



OR -- Notice that the word OR was intentionally upper cased when grammatically it shouldn't be. Typically, when you read a set of values you would use the word AND. East and West. 1, 7, 9, 23 AND 33. But when you are filtering things you are really asking for any of the values you've filtered. Meaning you want any rows of data where the Region is East OR West. Where a patients age is 1, 7, 19, 23 OR 33. This is an important concept to remember as you proceed. In the Advanced section of this workshop that will especially come into play when you purposely want the opposite. Meaning you only want something where the sales person sold something in the East AND sold something in the West. Ruh-roh. That's not how Qlik works is it? Filters provide the OR so what can you possibly do to override that behavior?

Set Analysis of course. That's why you are taking this workshop.

How would you read the following expression?

**Sum( Sales )**

If you said, "Sum of Sales" you get partial credit.

For full credit you should have said "The Sum of Sales using the "Default selection State" as the Set.

## States

The previous question was sort of a trick to help you transition from just blindly thinking that every single expression has to accept the default selection state as the set of values for your expressions.

The bare bones of Set Analysis would involve putting one of the two following snippets inside one of your expressions:

**{ \$ } or { 1 }**

That "Sum( Sales )" expression can literally be written out with complete syntax to help you read the complete meaning like this:

**Sum ( { \$ } Sales)**

The { \$ } literally means the Set of Values based on the DEFAULT selection State.

What if you don't want that default selection state because you know the users often do silly things like filter to something then complain that your overall KPI shows a value that is too small? In that case you could write your expression like this:

**Sum ( { 1 } Sales)**

The { 1 } literally means the Set of ALL Values in the system ignoring any selection state. In other words "I want to show a Sum of all of the Sales values no matter whether or not the user has filtered any of the fields."

You will learn more about other "states" later. But for now, think "Set of all values ignoring what the user selected" OR "Set of all values honoring the selections of the user."

## Section 1 – Just getting started

### Let's try it

Open the application "Qlik US Contract Spend (Set Analysis)"



Go to the "Start here" sheet. You will find a KPI at the top of the screen that says, "Total Spend." The expression for that KPI is: `Sum( SpendAmt )`

Start here

Total Spend  
147.4B

{ \$ }

Copy the "Total Spend" KPI and change the expression to explicitly use the { \$ } as follows and label the new KPI as "Total Spend Default".

`Sum( { $ } SpendAmt )`

Notice that the value is the same.

Now select some values and you will see that in fact when expressions are executed they assume that you want to use the "set of values based on the default selection state from the user." That also happens to be the default that is applied when you write expressions but don't declare anything specifically. In other words, you've been doing Set Analysis all along, but didn't know it.

{ 1 }

What if I'm a big cheese and I'm looking at a KPI that says "Total Spend" and I assume that it is our "total spend" but I've clicked a filter. In that case having an expression use my selection is a bad thing.

That's where the { 1 } comes in handy. It tells the expression "The big cheese will more than likely click a filter and then call us wondering why the number doesn't match what he read in a report somewhere so please ignore anything they've filtered."

Copy the Total Spend KPI and change the expression to use the { 1 } as follows and label the new KPI as "Total Spend Ignore User":

`Sum( { 1 } SpendAmt )`

Ensure you have clicked on some filters and notice that it ignores everything. It continues to calculate the SpendAmt from the entire "data set" for the application.

## { Green, Grey and White }

Independently each of these is nice. Are you ready to have some fun by using both "states" at the same time?

Well ready or not here we go. We talk about the Green/Grey/White coloring and the fact that you get to see what is left behind. If you click on East and West those appear in Green. Any sales people that sold to those areas are in white and any sales people that did not sell to those areas appears in dark grey.

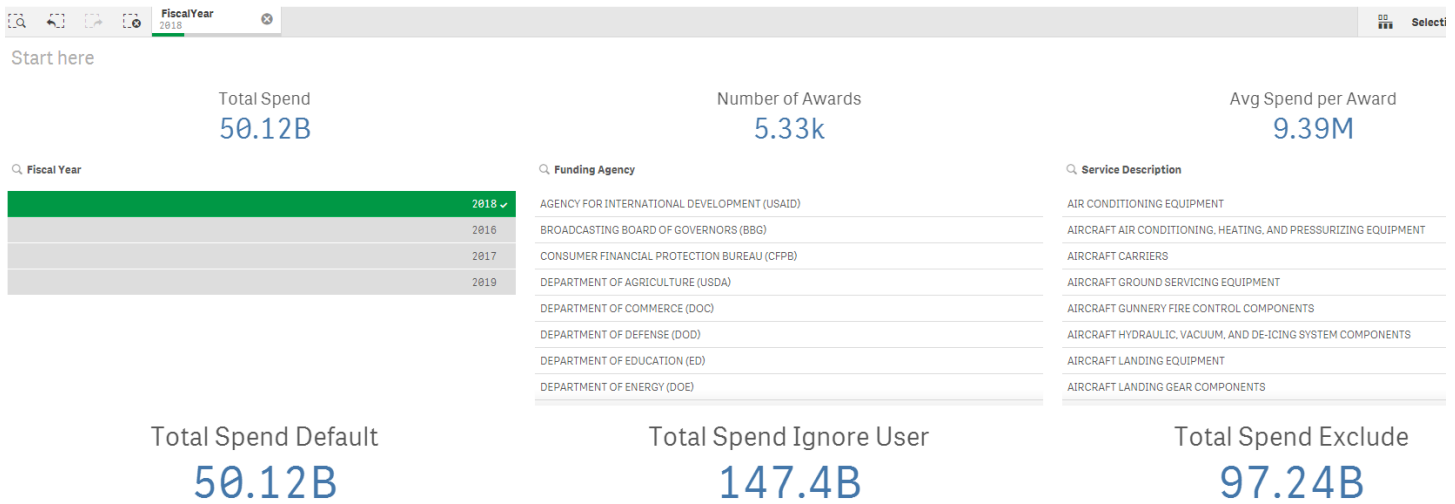
Similarly, what if in our application we want to see all of the spend for things that are excluded by your selections.

Copy the Total Spend KPI and change the expression as follows and label it "Total Spend Exclude":

$\text{Sum}( \{ 1 - \$ \} \text{ SpendAmt } )$

You barely broke a sweat and you just did some really cool *analysis* of the two *sets*. You told the expression to start with all of the data, and then exclude the items that were filtered. Or even simpler ... you calculated the Total Spend of the items in the Dark Grey.

You now have 3 KPI's on the screen. Validate that this formula is showing the difference between your KPI with the { 1 } and the { \$ }.



Hopefully you have a renewed sense of the power of how Qlik operates. The ability to operate on SETS, and your ability control and manipulate them, not just accept the default will prove to super powerful. Hopefully, the simple stuff thus far has piqued your curiosity to see how much more we can do with it. Because there is so much more.

## Section 2 – Now we are getting warmer

### Forcing Field Values

If you are feeling pretty good about the power of this Set Analysis stuff you should.

The great news is we are just getting started. We can also use it to force a desired "set" of values. Meaning, who needs the end users anyway. If we want an expression to represent only the set of values for the year 2018, then let's just say that.

To do that we need a way to define to the expression that we are defining some field values that we want to "force." We do that using that by modifying the basic Set Analysis syntax that you've learned so far as follows:

```
{ $/1 < field(s) = {values} > }
```

Notice that we have introduced the < > as a way to identify the field(s) we want to force.

Why do you think we need to use the { } brackets around the values we want to force the fields to?

Because we are always, *always*, always, **always** working with sets. Fields aren't filtered to a value, they are filtered to a SET of values. A set of values that may contain 1 value or it may be multiple values. But it's always a SET.

Go to the sheet "Forcing Field Values" and clear all of the selections.

Filter to the year 2018 and jot down the Total Spend: 50.12B

Clear all of the selections again.

Add a KPI with the following expression and label it "Spending for 2018":

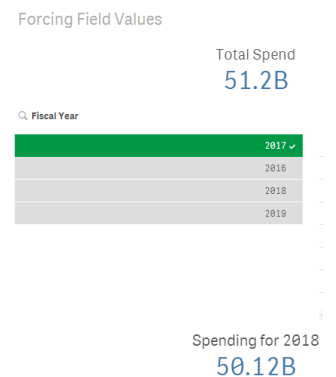
```
Sum( { < FiscalYear = { 2018 } > } SpendAmt)
```

The KPI should show the value 50.12 B which is the amount of spending for the year 2018 even though the user didn't have to select anything.

Now select 2018, what does the new KPI show?

Now select 2017, what does the new KPI show?

It always shows 50.12B because we have told the expression that we don't care what the big cheese selected, we want that field to use the set of values that we have provided.





## "Oh those Users"

Now clear all of the selections and select "Air Conditioning Equipment" in the Service Description. What does the new KPI show?

It shows 5.41M because our Set Analysis expression is using the "default" filter state for fields, except for the Year which we have overridden.

Clear all of the selections and select "Access Board" as the Funding Agency.

Spending for 2018

0

Why does our KPI show 0?

Although we want to ignore the user filters for the Fiscal Year our problem now is that the year we want, 2018, isn't even available as an option. Notice that all of the Fiscal Years are Dark Gray when we select that Funding Agency. Our KPI can't calculate spending for 2018 because there is never any spending for that Funding Agency.

Can we get around the user selections if we want our KPI to always, always, always show the total spend for Fiscal Year 2018? You better believe we can.

Add a new KPI using the following expression and label it "Total Spend FY2018 Ignore User"

**Sum( { 1 < FiscalYear = { 2018 } > } SpendAmt )**

## Multiple Values

Again, we are using the set brackets because fields are always assigned a set of values. Let's create a new KPI that uses multiple values.

Create a KPI with the following expression and label it "Spend for 2018 and 2019":

**Sum( { 1 < FiscalYear = { 2018, 2019 } > } SpendAmt )**

There are times that is practical and times that it isn't. If we want everything since 2017 we could list { 2017, 2018, 2019 }. But what happens next year when we have 2020 data?

I remember now, it's the next persons problem. They will need to answer the phone with all of the complaints and have to figure it out. Just kidding. Set Analysis can include sets of values, this is so cool, that includes a formula and not just the expressly written out set of values.

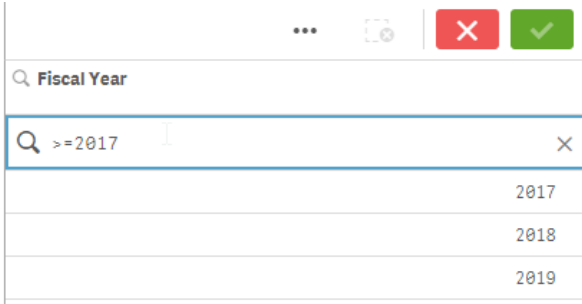
No kidding. We don't need to be specific. We can assign a set of values that is determined on the fly.

Create a KPI with the following expression and label it "Spend from 2017 and on":

**Sum( { 1 < FiscalYear = { ">= 2017" } > } SpendAmt )**

The new KPI we added should show 103.4B. Why?

Clear all of the selections. Click in the search box for the Fiscal Year and enter the following text: **>=2017** and press the Green Arrow to accept the search selections.



See if the top KPI (using default set) confirms that the number should be 103.4B.

### Forcing Field Values

Total Spend

103.4B

Fiscal Year



It's not part of this workshop, but if you didn't know you could input expressions like that into the search box you do now. Qlik provides a tremendous online resource to help you learn all that you can do inside a search pane. You will find that you can do lots of complicated things in the search windows. Check it out:

[https://help.qlik.com/en-US/sense/April2019/Subsystems/Hub/Content/Sense\\_Hub/Selections/SelectionsToolbar/search-selections.htm](https://help.qlik.com/en-US/sense/April2019/Subsystems/Hub/Content/Sense_Hub/Selections/SelectionsToolbar/search-selections.htm)

## Section 3 – Making progress now

### Multi Field Set Analysis

What you've already learned allows you to do some powerful things. Your customers will sing your praises as they walk down the hallways. But once you've helped them see the light, they will start asking much more complicated questions that don't involve something as easy as "accept the selections or ignore them." They will ask you to do things like honor some field requests but not honor others. Of course, they won't say it that way. They will say "I need to know the total spend for 2018 for the Service Description that I choose, but I don't care what Funding Agency it is for."

Create a new KPI, label it "User wants to trick me" and use this expression:

**Sum( { < FiscalYear = { 2018 }, [Funding Agency]= > } SpendAmt )**

Clear all of your selections.

Select "CLOTHING, SPECIAL PURPOSE" in the Service Description field. Your new KPI should show 181.5M.

Select 2017 as the Fiscal year

Select "DEPARTMENT OF STATE (DOS)" as the Funding Agency.

Notice that your new KPI still shows 181.5M. You have told the system you want to ignore what the user selects in the Funding Agency field, force the year to be 2018 while honoring the other filters that the user has selected like Service Description. That's kind of cool.

#### Forcing Field Values

The screenshot displays a dashboard with three filter panels and two KPI cards. The 'Total Spend' KPI shows 6.02M, and the 'Number of Awards' KPI shows 1. The filters are: Fiscal Year (2017), Funding Agency (DEPARTMENT OF STATE (DOS)), and Service Description (CLOTHING, SPECIAL PURPOSE). Below the filters, the KPI 'User wants to trick me' is shown with a value of 181.5M.

Filter	Selected Value
Fiscal Year	2017
Funding Agency	DEPARTMENT OF STATE (DOS)
Service Description	CLOTHING, SPECIAL PURPOSE

**Total Spend**  
6.02M

**Number of Awards**  
1

**User wants to trick me**  
181.5M

## Beyond KPI's

Life starts getting a whole lot more fun when we use Set Analysis in visuals beyond KPI's. For simplicity let's imagine that we want to show a Year to Year comparison. We want to compare 2017 to 2018 rather than "just" seeing either.

Create a bar chart titled "2017 and 2018 Spend"

Select the field "Funding Agency" as the Dimension

Create a Measure with this expression and label it "2017 Spend"

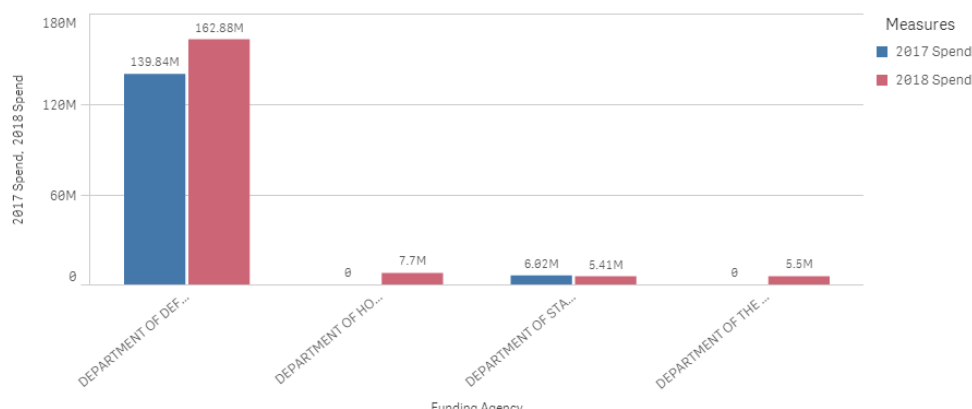
**Sum( { < FiscalYear = { 2017 } > } SpendAmt )**

Create a second Measure with this expression and label it "2018 Spend"

**Sum( { < FiscalYear = { 2018 } > } SpendAmt )**

Be sure to set the Sort Order to sort by Funding Agency.

Clear all of your selections and select "CLOTHING, SPECIAL PURPOSE" in the Service Description field.



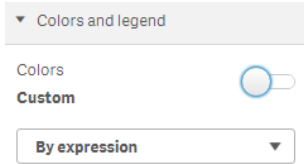
Our Set Analysis expressions honored the Dimensional field inside of our bar chart. The problem I have, and your users have is ... "I can't really tell what's better versus what's not." Our eyes will have to bounce back and forth to the legend. Our minds are rattled when we see DEPARTMENT OF HOUSING and DEPARTMENT OF THE INTERIOR but the "blue" bar is missing. So, let's do those poor end users a favor and do the comparison for them to make the chart a little easier to understand. In layman's terms, let's compare the values for them instead of asking them to compare the values in their heads.

Copy the bar chart you created, retitle it "Comparison of 2018 Spend to 2017 Spend" and remove the 2 measures.

Create a Measure with the following expression labeled "2017-2018 YOY":

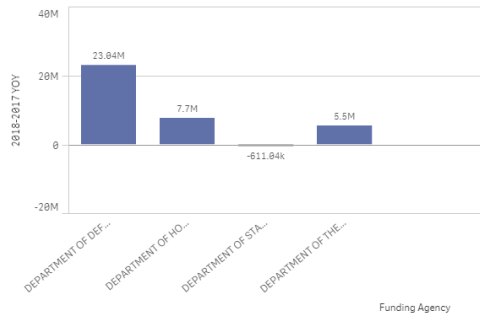
**Sum( { < FiscalYear = { 2018 } > } SpendAmt ) -  
Sum( { < FiscalYear = { 2017 } > } SpendAmt )**

Change the Colors property not to be Custom and utilize the following expression



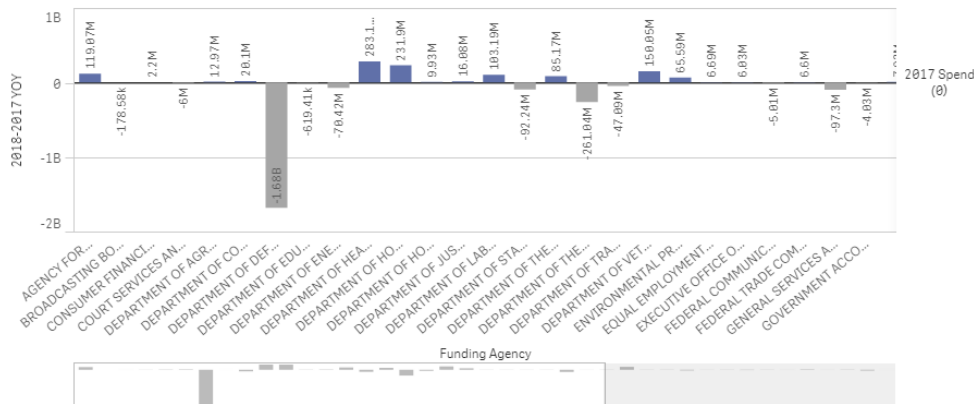
**If( Sum( { < FiscalYear = { 2018 } > } SpendAmt ) - Sum( { < FiscalYear = { 2017 } > } SpendAmt ) > 0, QlikTechBlue(), QlikTechGray() )**

Your chart should look like the following image. (You may need to expand the size of your bar chart to clearly see it.)



Instead of asking the user to do all of the work mentally you are now visualizing the "comparison" for them. You are simply showing them the Year over Year Spending change.

Now clear all of the selections and let's see how it looks:



Notice how easy it is to identify the Funding Agencies that spent more money in 2018 than they did in 2017 vs those that spent more in 2017.

For bonus points edit your bar chart and add the Reference Line label like appears above to really make it clear what the user is seeing.

## Section 4 - Visualizing Context

Click on the giant gray bar which is for DEPARTMENT OF DEFENSE (DOD) which shows the huge 1.68B reduction in spend between 2017 and 2018.



Oh no. All of the other bars disappeared. Is that a good thing or a bad thing?

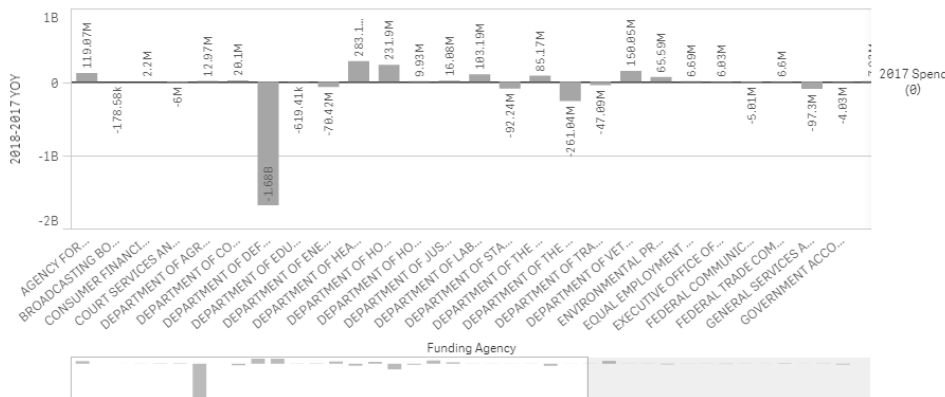
It's a good thing. Qlik is immediately responsive to the user's selection and you purposely selected 1 bar, which means our "default filter state" is to only show that specific Funding Agency.

It's a bad thing. The user can visualize the YOY comparison, but now they've lost the context to the other bars to understand if 1.68B is a big deal or not.

Edit the Measure and use the following expression:

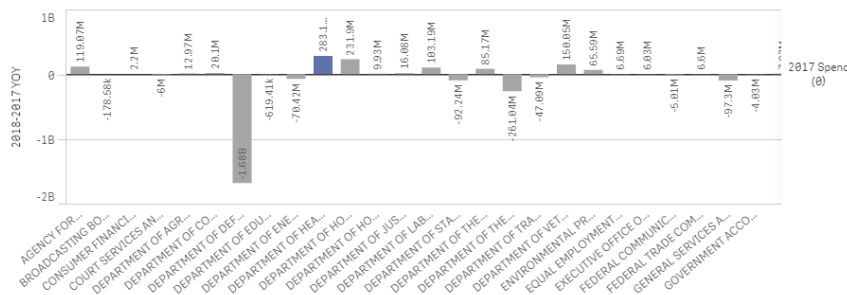
**Sum( { < FiscalYear = { 2018 }, [Funding Agency]= > } SpendAmt ) -  
Sum( { < FiscalYear = { 2017 }, [Funding Agency]= > } SpendAmt )**

Well now this is a head scratcher isn't it? All of the Funding Agencies show up, but they are all gray.



There is an explanation. But first ...

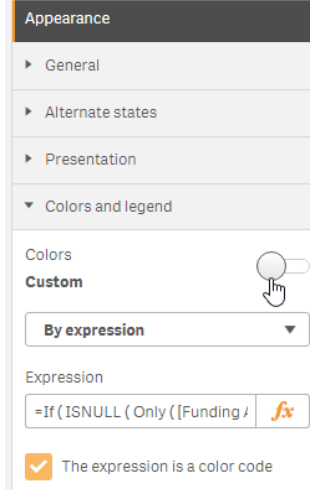
Clear your selections. Then click on 1 of the Funding Agencies that had more spend in 2018 than in 2017.



Notice the Funding Agency you selected that was blue, remains blue, while all of the other bars are now gray. Can you explain why? Hint – Think about the Color expression you used as well as the Measure expression that we used to show Funding Agencies that were not selected.

Now we get to see all of the Funding Agencies despite having selected some. But we still have a slight problem. All of the bars are the exact same color. That makes it difficult to do any type of "comparison" we might like to see.

Edit the Bar Chart and change the Colors and Legend to use the following expression:



**=If ( ISNULL ( Only ( [Funding Agency] ) ) , LightGray() , RGB (141, 170, 203) )**

Combining the Set Analysis to allow us to visualize all the values, including those excluded by the user, while still helping them reference the context of what they've chosen is called Brush stroking.

### Let's think together



While not specifically related to Set Analysis this visual "Brush stroking" or "Brushing" concept enables you to provide for contextual comparisons. The expression above compares each dimensional value to the "set" of values selected by the user for the Funding Agency field. If the value is NULL, not selected, then it is colored in LightGray. If the value is not null, meaning it was selected, then it is colored using that specific RGB color.

Do you think this technique is helpful in conveying context?  
Do you think this technique is intuitive for end users or will it confuse them that they selected something?

## Section 5 – Working with text values

### Set Analysis with Text Values

Move to the sheet titled "Department of Defense"

Clear all of the selections

Select "DEPARTMENT OF DEFENSE (DOD)". The total spend should be 91.61B.

Now clear all of the selections.

Set Analysis also provides for the ability to use textual values. The only difference between text values and numeric values is that they need to be wrapped in either single or double quotes.

Create a KPI using the following expression and label it "Department of Defense Spend"  
{Notice that because the field "Funding Agency" has a space in it we are wrapping the field in brackets}

**Sum( { < [Funding Agency] = { 'DEPARTMENT OF DEFENSE (DOD)' } > } SpendAmt )**

Select "ACCESS BOARD (USAB)"

Our new KPI should still show the 91.61B figure?

When we tried that before we were in big trouble because no years of data were present. How come it shows the correct value even though there are no years that are valid?

Our Set Analysis is overriding the Funding Agency field now. Thus, the years aren't excluded from the user's selection. Don't miss that. It's important to understand the ramifications of overriding existing user selection.

Change the selection for Funding Agency to be "FEDERAL COMMUNICATIONS COMMISSION (FCC)" and also select the year 2017.

Our new KPI will now show 33.04B? Hopefully by now you understand why and could it explain it. While our expression forces the set to use the specific Funding Agency we are still honoring the user's desire to see a specific Fiscal Year. Again, we can override certain fields but still honor the others.

Remember we can use multiple fields in our Set Analysis if we want to if you didn't want to honor the user's choice in Fiscal Year.

Create a KPI that will show us total spend for "DEPARTMENT OF DEFENSE (DOD)" but only for the year 2018 using the following expression and label it DOD 2018:

**Sum( { < [Funding Agency] = { 'DEPARTMENT OF DEFENSE (DOD)' }, FiscalYear= {2018} > } SpendAmt )**





Our DOD KPI will show 31.36B. How would you validate that number?

## Multiple Text Values

Create a new KPI so that we can see the total spend for both the DOD and the DOC for 2018 as follows and label it "DOD and DOC 2018":

**Sum( { < [Funding Agency] = { 'DEPARTMENT OF DEFENSE (DOD)' , 'DEPARTMENT OF COMMERCE (DOC)' }, FiscalYear= {2018} >} SpendAmt )**

Notice that when dealing with textual fields when we have multiples the field values we assign to the set each individual value is wrapped in quotes but the commas separating the values are not.

Now and then we find ourselves want to exclude something specific. Earlier we simply excluded whatever the user selected in the default selection state. We created a very specific KPI to show only the spend by the Department of Defense (DOD) so perhaps we want to create another KPI to show the total spend for all agencies except that one.

Create a new KPI using the following expression and label it "Non DOD Spend":

**Sum( { <[Funding Agency] -= { 'DEPARTMENT OF DEFENSE (DOD)' } > } SpendAmt )**



If you are working ahead in the group, or working on your own and want to earn a gold star copy the KPI you created for DOD and DOC for 2018. Modify the expression to exclude the set for Funding Agency and the FiscalYear. Your total should be 36.05B.

## Wildcards

Clear all of your selections. Notice that there are a lot of Funding Agencies that start with the name "DEPARTMENT OF." Don't freak out you don't have to create a KPI that includes all of them individually. Instead we can take advantage of the fact that Set Analysis will allow us to use wildcards. That enables our KPI to be ready when new DEPARTMENTS are created to spend money.

Create a new KPI using the following expression and label it "Spend for all Departments":

**Sum( { < [Funding Agency]={ 'DEPARTMENT OF\*' } > } SpendAmt)**

Notice that it shows a value of 0. The wildcard search isn't a value. It's an expression itself. We need to wrap "expressions" like that inside of double quotes (") instead of single quotes. *If you look back to the earlier exercise when we wanted the set of values >= 2017, it was wrapped in double quotes.*

Change the KPI Expression to use Double Quotes (") instead of single quotes around our expression.

**Sum( { < [Funding Agency]={ "DEPARTMENT OF\*" } > } SpendAmt)**

*I started the explanation for using text values by saying that we could use single or double quotes. Given that you may want to rapidly do something like this in the future, get in the habit now of using double quotes whenever you use textual fields. It will save you a lot of headaches.*

Just as we wanted to specify more than 1 actual value, there may be times we need a set to include more than 1 "expression."

Create a KPI using the following expression and label it "Spend for Dev and Dept"

**Sum( { < [Funding Agency] = { "\*"DEVELOPMENT\*" , "DEPARTMENT OF\*" } > } SpendAmt)**

## A little practice goes a long way

We've been focusing on the Spend for all of our KPI's. Take some time on your own now to test what you've learned.

Move to the sheet labeled "Awards"

Copy the "Number of Awards" KPI to create a new one. Try on your own to update the expression to only view the number of awards for the Fiscal Year 2018. The value should be 5.33K. (Answer is on the next page but don't cheat unless you run into problems.)

Hopefully you were able to make it work. But if not you might have wondered where the Set Analysis syntax needs to go when you are using an expression like "Count( Distinct". The great news is that it doesn't matter whether you use Distinct before or after your Set Analysis expression. Both of the following will work.

**Count( { < FiscalYear= {2018} >} Distinct award\_id\_piid)**

**Count( Distinct { < FiscalYear = {2018} > } award\_id\_piid)**

Spend the next 5-10 minutes trying to create some other KPI's with fixed field values for Awards and/or Avg Spend per Award.



Many organizations have a focus on the buzzwords Data Democratization. Which is an admirable goal for sure. But can you imagine every visual designer and end user trying to build these Set Analysis expressions on their own?

Scary thought isn't it. While we are building these awesome, custom expressions inside each of our KPI's, imagine the value we can add if we build them as Master Measures instead.

We put the brain power into defining them 1 time, and every person that comes after us is then able to drag/drop/select them for use in any type of visual that they want without knowing how to program. That is called "Governance." It's like guard rails for those that want to do more, but can't code.

Don't worry you won't be tested on this thought, but since there white space on this page it made sense to fill it.

## Section 6 – Coding for flexibility

### Put Your Thinking Caps On

When I was in school and a teacher said that, I knew it was time to focus. My friends, it is now time to focus. The next several things we attempt might make your head hurt, but I assure you that it's critical you push through. So "put your thinking caps on", reread explanations if necessary, or walk away and come back if need be. Here we go.

### Flexibility in Assignment

We created multiple forms of KPI's that were hard coded with years, 2018, 2017 etc. But often you want to be a little more flexible than that because you don't want your KPI's to "time out" in the future.

Move the sheet titled "Flexibility"

Create a "Text & Image" object

Create a Measure for it with the following expression:

**Year( Today() )-1**

It should show you the Year that is prior to the year of the day you do this in.

No biggie yet. But here is what's important. You didn't have to type the year, it was calculated for you. Now we want to create a KPI that would show us the total spend for the previous year.

Create a KPI using the following expression and label it "Spend for Last Fiscal year"

**Sum( { < FiscalYear = { \$(= Year( Today() )-1 ) } > } SpendAmt )**

Notice the special syntax that we had to use to tell the system to "evaluate" the expression for us. We don't want to assign the FiscalYear to the formula itself, we want the value of what the formula will result in. If we needed to use variables in the definition of a set of values we would use the same "dollar sign expansion" syntax.

### Relationship to User Input

There are times when we will need to assign a value that is related to something the user has selected and not just force it to what we want. Imagine that we want to allow the user to select a year, and then we provide the total spend for the previous year to their selection.

Create a KPI using the following expression and label it "Spend for Year Prior to selection":

**Sum( {< FiscalYear = { \$(= ONLY(FiscalYear) -1) } > } SpendAmt)**

You can find more explanation of the ONLY function by clicking here. [https://help.qlik.com/en-US/qlikview/April2019/Subsystems/Client/Content/QV\\_QlikView/ChartFunctions/BasicAggregationFunctions/only.htm](https://help.qlik.com/en-US/qlikview/April2019/Subsystems/Client/Content/QV_QlikView/ChartFunctions/BasicAggregationFunctions/only.htm)

In simplest terms it says, "Hey go grab the value that the user has selected and subtract 1 from it."

Clear all of the selections if there are any.

Select 2019 in the Fiscal year and your KPI should show 50.12B

Select 2018 in the Fiscal year and your KPI should show 51.1B

## Validating Results

Notice that on all of these sheets that the same root KPI's were on the top. We didn't refer to them often so can you guess why they are there?

If you guessed "so that we could validate results" you are either very smart, or very perceptive since that is the section title.

You should always ensure that you double check any results to validate that your Set Analysis KPI's are showing exactly what you think that they should.

In this case the numbers won't match because your new KPI's are forcing a different year. If you click 2019 the top KPI will show the value for 2019, while your new, prior year, KPI will be showing the value for the year prior. So you will need to write some values down as you double check in situations like this.

## Real World

Let's put a couple of these things together now and imagine that we want to display a KPI for the total spend amount just for the DEPARTMENT OF DEFENSE (DOD) for the prior year.

Create a new KPI using the following expression labeled "Spend for DOD Last year."

```
Sum( { < [Funding Agency]='DEPARTMENT OF DEFENSE (DOD)' } , FiscalYear={  
$(=Year(Today()-1) } > } SpendAmt )
```

Now copy that KPI and create a new one that is flexible enough to look at the year prior to the one that you select. [No scrolling down. No help. This is a pop quiz.] You can do it.

## Section 7 – You don't know what you don't you know

### Seriously we don't always know

Not much of a professional header but it's a very real-world use case. We don't always know what values that we want to assign to field. For instance, what if we want our Set Analysis to show all spend for "any" agency that spent money on the specific service "AIRCRAFT CARRIERS". We could manually create the list after selecting AIRCRAFT CARRIERS but that list could change next year so we need a way to literally perform that Set Analysis on the fly.

Clear all of your selections

Create a KPI using the following expression and label it "Spend by agencies who buy AC":

```
Sum( { < [Funding Agency] = P({1<[Service Description] = {'AIRCRAFT CARRIERS'} > } ) > } SpendAmt)
```

We will get into the syntax of that expression in a minute, but first let's figure out how to verify the 91.61B that it should show you. Try and verify that number on your own and if you can't then scroll down and read.

First thing we would want to do to verify is first select "AIRCRAFT CARRIERS" in the Service Description filter:

Q Service Description



Notice that our KPI, and the top KPI for Total Spend both show 61.97M. Hey that doesn't match. And it's not supposed to. That's the spend for AIRCRAFT CARRIERS, not all spend for agencies that buy AIRCRAFT CARRIERS. What we need to is filter to the agency(s) that are in White. Those that are "associated" with purchases of AIRCRAFT CARRIERS. For this first exercise, it's only DEPARTMENT OF DEFENSE (DOD). Select the DOD and remove the filter on AIRCRAFT CARRIERS.

Q Funding Agency

DEPARTMENT OF DEFENSE (DOD)

Q Service Description

AIRCRAFT CARRIERS

Q Funding Agency

DEPARTMENT OF DEFENSE (DOD)

ACCESS BOARD (USAB)

Q Service Description

AIR CONDITIONING EQUIPMENT

AIRCRAFT AIR CONDITIONING, HEATING, AND PRESSU

Now the top KPI for Total Spend validates against our KPI for 91.61B.  
 Let's try that again with something a little more complicated because there is no point if the only agency is the DOD and we've already built a ton of formulas for them. This time let's try with "HOUSEKEEPING- GUARD"

Create a new KPI with the following expression and label it "Spend by Agencies with HK":  
 \*\*\*\*\* **Be sure to notice that there is a space between the – and GUARD.**

**Sum({<[Funding Agency] = P({1<[Service Description]='HOUSEKEEPING- GUARD'})>}) > SpendAmt)**

Your new KPI should show you 125.2B. Again, let's try and verify.

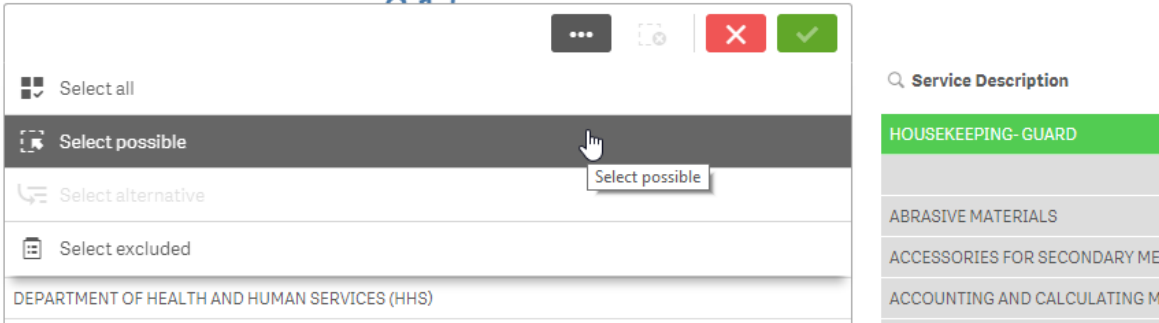
Start that process by selecting HOUSEKEEPING- GUARD as the Service Description. You will see that there are many Funding Agencies that have spent money on that area. If you try to hand select the items you might miss them. Right?

So how can we be sure that you select all of the POSSIBLE items that are in white?

By clicking the search button by the Funding Agency

Q Funding Agency

And then clicking on the Elipsis (...) and choosing "Select Possible" of course



Notice that it immediately selects all 15 agencies. Pretty handy dandy huh.



Now remove the selection for HOUSEKEEPING- GUARD so that we are only filtering to the Funding Agency(s) so that we can see their total spend in the upper left. What do you know ... it's 125.2B.



If you are beginning to feel like you really understand how Qlik works behind the scenes, that's not an accident. Understanding how our "state engine" works and how calculations are based on "sets" is the key to blasting off in your ability to provide amazingly powerful analytics.

## Crazy Syntax

When I presented the first expression, I said "we will get into the syntax soon but let's verify first." Probably drove you crazy then, but there is a reason for that. The process of verification will help you understand that syntax. So now let's take a look at the expression and start breaking it down.

The entire expression was this:

**Sum({<[Funding Agency] = P({1<[Service Description]='HOUSEKEEPING- GUARD'})>}) SpendAmt)**

We are going to focus on 1 piece at a time ... what does this piece do?

**{ 1 <[Service Description] = {'HOUSEKEEPING- GUARD'} > }**

If you said "it's a Set Analysis expression that says ignore what the user has selected and cause the set to include the Service Description for a value of HOUSEKEEPING- GUARD then you have been paying attention.

Now let's look at the next piece. It is the part that wraps that simple Set Analysis formula inside of a function P. The P stands for POSSIBLE. You know, just like the Search you did in order to verify the number we got. It "selects the POSSIBLE values." You can find more on the P() function here: [https://help.qlik.com/en-US/sense/April2019/Subsystems/Hub/Content/Sense\\_Hub/ChartFunctions/SetAnalysis/set-modifiers-implicit-field-value-definitions.htm](https://help.qlik.com/en-US/sense/April2019/Subsystems/Hub/Content/Sense_Hub/ChartFunctions/SetAnalysis/set-modifiers-implicit-field-value-definitions.htm)

**P({1<[Service Description]='HOUSEKEEPING- GUARD'})>})**

Well now that expression isn't so intimidating, is it? The first part of it is simply saying I want the Funding Agency(s) that are that set of possible values.

**Sum({<[Funding Agency] =**

## Pop Quiz Superstar

The one problem we had with our KPI was that if we selected a value then the KPI is limited to showing spend only for that selection. We don't want that in this case. Instead we want to allow the user to select the Service Description(s) only to identify which Federal Agencies spent money on it/them. But we want to know their total spending not just spend for that/those Service Descriptions. Think you can adjust the expression to do that?

Remember earlier when we used the Set Analysis expression to say, "ignore the value" of a certain field? Hopefully you do. Because you need to do that now.

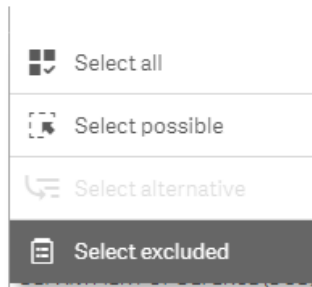
Modify your "Spend by Agencies with HK" KPI to the following expression:

**Sum({<[Funding Agency] = P({1<[Service Description]='HOUSEKEEPING- GUARD'})>}),  
[Service Description]= >} SpendAmt)**



## E(xcluded)

Similar to the Search function for selecting Possible values you can also ask it to select the Excluded values. Those Dark Grey values.



That might be of value to us. If we want to show "the total spend for all Funding Agency(s) that have NOT purchased any of the Service Descriptions. If the Possible values were handled through a simple function P(), would you care to guess what the function is to do the same for the Excluded values?

If you guessed E() you are wicked awesome and deserve a raise. Here is the expression so go ahead and create a new KPI using this expression and label it "Agency Spend Exclude":

**Sum({<[Funding Agency] = E({1<[Service Description]='HOUSEKEEPING- GUARD'>}), [Service Description]= >} SpendAmt)**

## Possible based on Selections

Before you think our tasks are over, consider the fact that we may need our Set Analysis to be even a little more flexible. Instead of hardcoding it to use "HOUSEKEEPING- GUARD" or anything else, what if we wanted to allow the end user to see the total spend for all Federal Agencies that ever-spent money on whatever and however many categories the end user has selected?

Before you lose your mind, guess what. It's actually easier than the previous ones.

**Sum({<[Funding Agency] = P([Funding Agency]), [Service Description]= >} SpendAmt)**

A Natural Language Translation of that formula would read: "Hey Set Analysis Engine you awesome thing you, I want to assign the Funding Agency to whatever the "currently" Possible set of values for that field based on whatever the user has selected, and then I want to ensure that I get the spend for all of the Service Description items, and not just the ones that they might have filtered." Gotta love Natural Language Processing. But that's an entirely different topic.

The workshop started with "It was the best of times, it was the worst of times." The syntax can be intimidating if you see one of the complicated formulas that we built at the end of the workshop. But if you followed along from the beginning hopefully it all makes more sense and isn't quite as intimidating.

## Section 8 – Advanced material

To be sure you understand the purpose of this section, here is the definition for the word Advanced.



*adjective*

far on or ahead in development or progress.  
"negotiations are at an advanced stage"

Nothing about that implies easy.

Nothing about that implies that everyone will get through it in the basic workshop.

### Alternate States

We've used the term "default selection state" multiple times now. Might have seemed out of place to be that elaborate, but the truth is that's exactly what it is. It's only the default selection and you can absolutely create more "selection states."

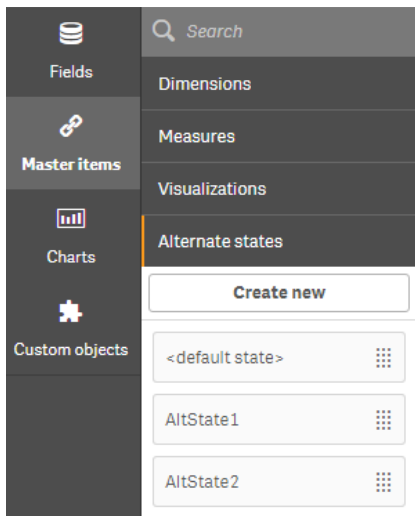
You might want to compare values between things. For instance, instead of saying "Ok Bob make a selection, now write the KPI value down, now make a different selection, now write that KPI value down. Now do the math in your head or on your toes." Kind of silly when it's written out that way, but that's kind of what happens in organizations around the world.

Qlik allows you to provide these Alternate States so that Bob can create both selection sets at the same time, and you can present the variances for him. Because you are super cool like that. If Set Analysis can refer to the Default State, then you have guessed correctly that it will allow us to refer to any Alternate States you create.

You can create these Alternate States using the Master Items tab.

Click on Master Items and scroll down to "Alternate States" and you will see that in this workshop there is a <default state>. Hey we refer to that as { \$ }.

In addition, you will see 2 Alternative or Alternate States that have been created: [AltState1](#) and [AltState2](#)



{ StateName }

If you want to use the filters from these alternate states in a formula you use the general syntax.

**{ StateName }**

Move to the "Comparisons" sheet

Select "2018" and "Department of Commerce" in Comparison 1 window

Select "2017" and "Department of Commerce" in Comparison 2 window

Comparison 1

Q Fiscal Year

2018 ✓
2016
2017
2019

Comparison 2

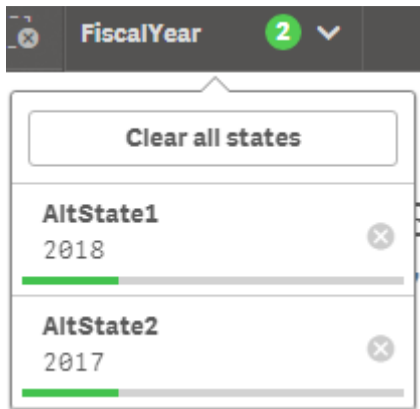
Q Fiscal Year

2017 ✓
2016
2018
2019

Notice nothing at the top of the window changes, Total Spend remains at 147.4B. Why?

Because the KPI's at the top of the screen are based on the "default" state selection and you haven't made any. 😊

Click on the FiscalYear filter in the breadcrumb trail. Notice that it shows the value you selected for that alternate state. Notice that there is a little gray X to the right of the value for each state so that you can clear only the value(s) from either state.



Now let's go ahead and build some expressions that use these 2 new states so that we can do some comparisons.

Create a KPI and label it "Spend AltState1" using the following expression (which literally means Sum of SpendAmt for the Set of Values selected from the State that is named AltState1):

**Sum( { AltState1 } SpendAmt)**

Create a KPI and label it "Spend AltState2" using the following expression:

**Sum( { AltState2 } SpendAmt)**

Create a KPI and label it "Spend Variance" using the following expression:

**Sum( { AltState1 } SpendAmt ) - Sum( { AltState2 } SpendAmt )**

How good are you feeling right now? You just provided your end user the ability to do comparisons and show them the variance without making them take off their shoes to count for themselves.



The ability to show CONTEXT and COMPARISON in applications is often the difference between adoption or shelf ware.

## Bar Chart

These Set Analysis expressions aren't limited to just KPI's. Let's see how much more powerful it might be using a bar chart.

Create a Bar Chart at the bottom of the sheet so that we can show the variance between selections across the different Funding Agencies.

The Dimension for your bar chart should be the field "Funding Agency"

The Measure is the variance formula you just created. Label the measure "Spending Variance":  
**Sum( { AltState1 } SpendAmt ) - Sum( { AltState2 } SpendAmt )**

Clear all of the selections

Filter 2018 in Comparison 1 and 2017 in Comparison 2

Do you understand what the bar chart is showing?

*If you said "It's showing the difference in spend between the 2 years that were selected you are really getting this. If not, this would be a good time to review before moving on with the rest of the exercises if you are going through this on your own or ask the workshop leader if you are in a group.*

Which Funding Agency had the biggest drop in spend between 2018 and 2017 (Enlarge the bar chart if you need to in order to see the Funding Agency names)?

The section on Alternate States discussed the fact that you can create additional "states" and allow the user to create various "sets" of filters. We used our 2 states individually. We allowed the user to pick 2 different years, then we used 1 of the years as the context to visualize the other. But there are many more, mathematical things we can do with sets.

## Visualizing Variance

Remember that cool comparison chart we did using the alternate states? Let's recreate that but without asking the user to lift a finger.

Create a bar chart at the bottom of the sheet

Dimension is the field "Funding Agency"

Using the following expression as the measure:

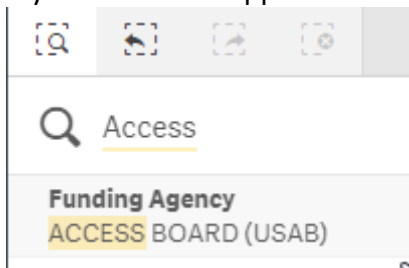
```
Sum( { < FiscalYear = { 2018 } > } SpendAmt ) -  
Sum( { < FiscalYear = { 2017 } > } SpendAmt)
```

Clear all of the selections. What is the chart visualizing now?

Same as before, it's visualizing the difference in spend between the 2 years and it's doing it without any human interaction.

## Let's think together

What do you think will happen if we search for "ACCESS BOARD (USAB)" and select it?



*Chart will be blank because all years are unavailable including the 2 we need.*

How could we fix it?

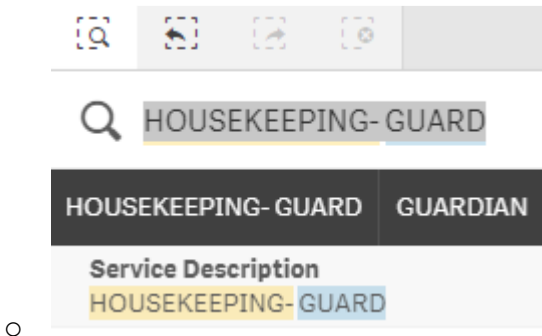
*Use 1 in our Set Analysis to ignore user selections*

Why did we need to Search for "ACCESS BOARD (USAB)" instead of using either of the Funding Agency filter panes already on the screen?

*Because the filter panes on the screen are tied to Alternate States 1 and 2, but our new bar chart isn't. We don't have a filter pane for the "default" state on this screen so we need to search for it.*

Clear all of the selections.

Search for the value " HOUSEKEEPING- GUARD " in the Service Description Field and select it.



Notice that there are multiple Funding Agencies that have spent money on that Service Description.

Now select Department of Defense (DOD)

What happens to our bar chart?

Only the Department of Defense (DOD) is available because that's what user filtered

Is that reasonable to expect?



By "default" Qlik is immediately responsive to what the user selects. That's great the vast majority of the time. But as you can see already there are questions that users will need to answer, and yet they themselves will be the ones causing the problems. Set analysis provides us the way to get answers regardless what we need to solve for. Refer to the earlier section where we touched on the concept of Brush Stroking.

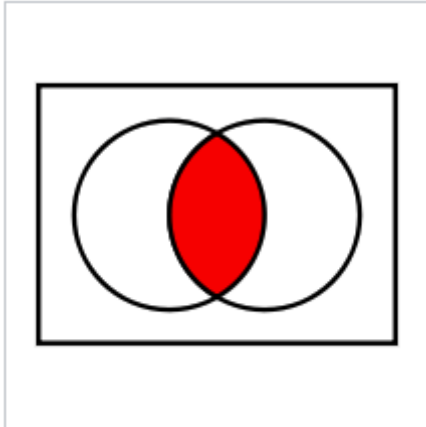
## Section 8 – Analysis of Sets

Student: Oh no not Venn Diagrams.

Instructor: Oh yes Venn Diagrams.

### Intersection

We might want to find the Intersection of the two sets. Meaning ... where they overlap.



Intersection of two sets  $A \cap B$

Given 2 sets:

$A = \{1, 2, 3\}$

$B = \{1, 3, 5\}$

The intersection of the two sets would be 1 and 3.

You might be saying "who cares."

Remember the section on AND/OR and we made a big deal of the words. It was to ensure you understood that out of the box when a user filters 2 values the filters are applied as OR. In the following image, the user has filtered to two values.

🔍 Service Description

ABRASIVE MATERIALS	✓
ADDITIVE METAL MATERIALS	✓
ACCESSORIES FOR SECONDARY METAL...	
ACCOUNTING AND CALCULATING MACHL...	

Any expression would be including rows that had a Service Description of ABRASIVE MATERIALS – OR – ACCESSING FOR SECONDARY METAL.

But what if the user wants a way to find agencies that have spent money on both ABRASIVE MATERIALS – AND – ACCESSING FOR SECONDARY METAL.



If that were the case it would be super nice if we had a way to find the intersection of the two Sets. Alternate States allows us to enable the two distinct sets and we do have the syntax to allow us to find the intersection between them.

Move to the sheet titled **Intersection**.

In Comparison 1 select AIRCRAFT, FIXED WING

**Comparison 1**

Fiscal Year	Funding Agency	Service Description
2016	DEPARTMENT OF AGRICULTURE (USDA)	AIRCRAFT, FIXED WING ✓
2017	DEPARTMENT OF DEFENSE (DOD)	ABRASIVE MATERIALS
2018	DEPARTMENT OF STATE (DOS)	ACCESSORIES FOR SECONDARY METAL...
2019	NATIONAL AERONAUTICS AND SPACE AD...	ACCOUNTING AND CALCULATING MACHI...

In Comparison 2 select AIRCRAFT, ROTARY WING

**Comparison 2**

Fiscal Year	Funding Agency	Service Description
2016	DEPARTMENT OF DEFENSE (DOD)	AIRCRAFT, ROTARY WING ✓
2017	DEPARTMENT OF JUSTICE (DOJ)	ABRASIVE MATERIALS
2018	DEPARTMENT OF STATE (DOS)	ACCESSORIES FOR SECONDARY METAL...
2019	ACCESS BOARD (USAB)	ACCOUNTING AND CALCULATING MACHI...

Notice that in Comparison 1 there are 4 Funding Agencies, and Comparison 2 there are 3 Funding Agencies. You can visually see that 2 of them "intersect", DOD and DOS.

Our first objective is to identify those Funding Agencies visually. Meaning the 2 that have spent money on both types of aircraft.

Create a new Table object

Dimension will be the field Funding Agency

Measure will be

```
Count(
{
< [Funding Agency] = P({AltState1}) > *
< [Funding Agency] = P({AltState2}) >
} [Funding Agency])
```

Visually you will see:

Funding Agency	Count( { < [Funding Agency]=P({AltState1}) > * < [Funding Agency]=P({AltState2}) > } [Funding Agency])
<b>Totals</b>	<b>22580</b>
DEPARTMENT OF DEFENSE (DOD)	22400
DEPARTMENT OF STATE (DOS)	180

Let's break down the syntax to ensure you understand how it works. The outside of the expression is a simple Set Analysis expression that will Count the number of Funding Agency records that match "whatever" the "set" analyses to.

```
Count(  
{ .....  
} [Funding Agency])
```

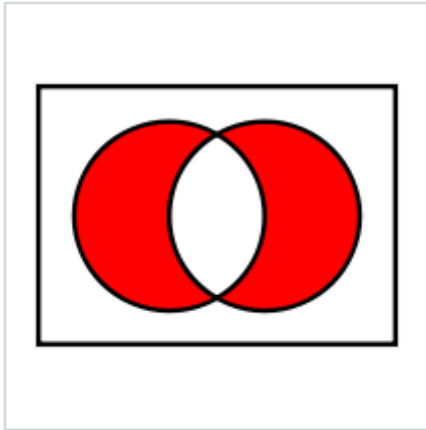
Our set analysis expression simply says find the possible Funding Agency values based on the set of values selected for the AltState1 alternate state. The \* symbol indicates that we want to find the intersection with the possible values for the Funding Agency values based on the set of values selected for the AltState2 alternate state.

```
< [Funding Agency] = P({AltState1}) > *  
< [Funding Agency] = P({AltState2}) >
```

You never knew it could be so easy, did you?

## Difference between sets

Oh no. I just heard from the customer and they didn't mean that they wanted the intersection of the two sets. They meant that they wanted the Funding Agency values that the two sets did not have in common with each other.



Symmetric difference of two sets

$$A \triangle B$$

Not to worry my friend. Qlik can handle that easily. That "operation" can be handled by simply changing the operator from "\*" to "/".

Modify the expression to the following:

```
Count(
{
< [Funding Agency]=P({AltState1}) > /
< [Funding Agency]=P({AltState2}) >
} [Funding Agency])
```

You should now see the three values for Funding Agency that the two alternate states do not have in common:

Funding Agency	Count( { < [Funding Agency]=P({AltState1}) > / < [Funding Agency]=P({AltState2}) > } [Funding Agency])
<b>Totals</b>	<b>3696</b>
DEPARTMENT OF AGRICULTURE (USDA)	2259
DEPARTMENT OF JUSTICE (DOJ)	1411
NATIONAL AERONAUTICS AND SPACE ADMINISTRATION (NASA)	26

## Exclusion

Well that figures. The customer just called and said that they were wrong again. What they really wanted was the values in the first set that were excluded from the second set.

Check out the following snippet that comes from [Qlik's online help regarding Set Analysis and the Operators](#) you have at your disposal to figure out how to modify your expression to see the Funding Agency values that exist in AltState1 but are excluded from AltState2:

## Operators

Operators are used to include, exclude, or intersect parts of or whole data sets. All operators use sets as operands and return a set as result.

This table shows operators that can be used in set expressions.

Operator	Description
+	Union. This binary operation returns a set consisting of the records that belong to any of the two set operands.
-	Exclusion. This binary operation returns a set of the records that belong to the first but not the other of the two set operands. Also, when used as a unary operator, it returns the complement set.
*	Intersection. This binary operation returns a set consisting of the records that belong to both of the two set operands.
/	Symmetric difference (XOR). This binary operation returns a set consisting of the records that belong to either, but not both of the two set operands.

If you modified your expression correctly you should see the following:

Funding Agency	Set Operation Hidden
<b>Totals</b>	<b>2285</b>
DEPARTMENT OF AGRICULTURE (USDA)	2259
NATIONAL AERONAUTICS AND SPACE ADMINISTRATION (NASA)	26



The real value in building a table as we have is in the fact that you can use it to validate your set analysis expression is correct when you want to perform these "operations" on your alternate states. Especially as you will see in the next section.

## Total Spend for Set Operations

In the last exercise we identified the Funding Agency values that matched the operations we needed to perform on our alternate states based on the user requirements. Given they kept changing their mind, I think you did a great job.

In this exercise we need to create a KPI to show the total SpendAmt.

Don't panic. It's super easy as you've already done the heavy lifting. All we need to do is change the outside of the expression so that instead of COUNTing the number of Funding Agency values we SUM the SpendAmt.

Create a KPI with the following expression:

```
Sum(  
{ < [Funding Agency]=P({AltState1}) > -  
< [Funding Agency]=P({AltState2}) >} SpendAmt)
```

The KPI should show 7.43B. But if you've learned anything in this workshop, hopefully it is that you need to validate the result. But how? We are doing all of this crazy Set Analysis stuff.

Well that table that we created is showing the 2 Funding Agency values that AltState1 has that are excluded from AltState2. Simply click on both so that they are selected.

Funding Agency	Set Operation Hidden
<b>Totals</b>	<b>2285</b>
DEPARTMENT OF AGRICULTURE (USDA)	2259
NATIONAL AERONAUTICS AND SPACE ADMINISTRATION (NASA)	26

The Total Spend value in the upper left shows ... hey that's cool ... it shows the same 7.43B value.

Total Spend  
7.43B

## Section 9 – The easy button

Move to the sheet titled Set Wizard

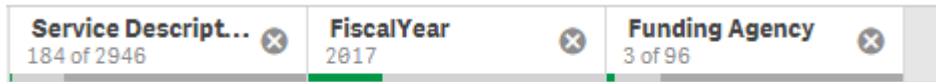
Clear all of the selections

Select 2017 for the Fiscal Year

Select DEPARTMENT OF COMMERCE (DOC), DEPARTMENT OF DEFENSE (DOD) and DEPARTMENT OF ENERGY (DOE) for the Funding Agency

Use the Search for Service Description and type "\*CONSTRUCTION\*" and press Enter

Your breadcrumb trail should now look like this



Total Spend is 1.06B

We want to create a Set Analysis expression that would implement everything we just selected. You could create this by hand now that you are such a Set Analysis whiz. But you don't have to. There is an easy button.

Add a KPI, and open the expression editor to input the formula

Select SpendAmt as the Field, and choose Sum as the Aggregation function. Then check the box for Set Expression and press insert

Fields

Filter by table  
All tables

Field  
SpendAmt

Aggregation functions  
Sum

Set Expression  
 Distinct  
 Total

Insert

Add expression

```
1 Sum({<FiscalYear={'2017'},[Funding Agency]={'DEPARTMENT OF COMMERCE (DOC)','DEPARTMENT OF DEFENSE (DOD)','DEPARTMENT OF ENERGY (DOE)'},[Service Description]='*CONSTRUCTION*'} SpendAmt)
```

Press Apply then press Done to exit the edit mode.

Clear all of your selections, and your new KPI will show that 1.07B figure.

By golly that seems like the proverbial Easy Button. Did we forget to mention that to you earlier? Our bad.

It doesn't do everything that you have learned in this workshop, or that you will ever need, but it sure is handy dandy for getting started so you can avoid typing a lot of long field values etc.

## Appendix

No training material would be complete with the obligatory list of additional resources. Seriously the importance of checking out the resources is that your learning can often times be limited if the examples don't make sense to you. Calculating federal spend might mean nothing to you so it's hard to understand. But the moment you see an example related more closely to what you do care about/support daily and the light bulbs might go off in your head.

For instance, if you are in the healthcare field and you see a Set Analysis example that shows patients who have taken this drug, that drug and have had this diagnosis (Intersection) suddenly you will be laser focused. While you may have thought who cares what department spent money. Or if you are in retail sales and you see an example of customers who purchased specific products in the previous 2 years, but have yet to purchase in the current year alarm bells go off in your head.

### Set analysis and set expressions

- <http://help.qlik.com/en-US/sense/September2017/Subsystems/Hub/Content/ChartFunctions/SetAnalysis/set-analysis-expressions.htm>

### Set modifiers with implicit field value definitions

- <https://help.qlik.com/en-US/sense/April2018/Subsystems/Hub/Content/ChartFunctions/SetAnalysis/set-modifiers-implicit-field-value-definitions.htm>

### Set modifiers with advanced searches

- <https://help.qlik.com/en-US/sense/September2017/Subsystems/Hub/Content/ChartFunctions/SetAnalysis/set-modifiers-advanced-searches.htm>

### Set analysis – an introduction

- <http://help.qlik.com/en-US/sense/September2017/Content/Videos/Videos-set-analysis-introduction.htm>

### The Magic of Set Analysis – Syntax and Examples

- <https://aftersync.com/blog/the-magic-of-set-analysis-syntax-and-examples>
- <https://aftersync.com/blog/the-magic-of-set-analysis-point-in-time-reporting>

### Set Analysis Cheat Sheet

- <http://cdn.aftersync.com/Blog/QlikView-Set-Analysis-Cheat-Sheet.pdf>

### Set Analysis Webinar on Demand (Healthcare focused examples)

- <https://bit.ly/2Jy9NwG>

## Set Analysis Cheat Sheet

A Set can include: **Identifier** + **Operator** + **Modifier**

### Identifier

- 0 = Empty Set
- 1 = Complete Application (all dimensions "reset")
- \$ = Current Selection (\$selection – default)
- \$1 = Previous Selection
- \$\_1 = Next Selection
- Bookmark01 = Name/ID of a bookmark

### Operator

- + : Union of sets (AND)
- \* : intersection of sets (BOTH)
- - : Exclusion of sets (NEITHER)
- / : Belonging to only one of the sets (OR)
- += OR -=: all values in current selection +/- additional set

### Modifiers

#### 1. ALL

- a. `< Dimension1 = { * } >` (*{ \* }* for numeric, *{ "\*" }* for text – note: will not include NULL data)
- b. `< Dimension1 = >` (*will also include NULL values*)

#### 2. KNOWN

- a. `{ <Dimension1 = { member1, member2, ... }, Dimension2 = { ... } >`
  - i. Numbers are not enclosed in single/double quotes. Text is enclosed by 'TEXT' , "TEXT", or [TEXT]

#### 3. SEARCH STRINGS

- a. `{ <Dimension1 = { "*"search string*" , "fixed text" , " ?earch2*" } >`
  - i. (*use \* for unlimited characters, ? for 1 character*)
  - ii. *Can do multiple searches at once!*

#### 4. BOUNDARIES

- a. `{ <TIME_KEY = { ">=12 <=18" } >`
  - i. Code (*>=<*) is BETWEEN double quotes

#### 5. VARIABLES

- a. **Standard:**
  - i. `< Dimension = { $(vVariable) } >` OR `< Dimension = { $(=vVariable) }`



1. = will interpret the variable
2. If member names contain spaces, need to enclose in " "

**b. Storing a search string**

- i. < Dimension = { \$(vSearchVariable) } >
  1. vSearchVariable must be in " " if string
  2. **NOTE:**if vSearchVariable is an expression that returns a string, you need to write <DIMENSION = \$(= \$(vSearchVariable))>, which tells Qlik to (a) evaluate the expression and (b) expand the variable. Alternatively, if the variable contains an "=", you only need one \$ (it will expand the evaluated expression)

**c. Storing an integer**

- i. < Dimension = { "<\$(vTimeVariable)" } > OR < Dimension = { "<\$(=vTimeVariable)" }
  1. Need to enclose variable name and code (<=>) between double quotes ""

**d. Storing a dimension**

- i. < [\$(vDim)] = { "\*" } >
  1. Safe to add brackets in case dimension name includes spaces

**e. Storing a Set**

- i. Sum ( < \$(vSet) > Sales )
  1. vSet could be ...
    - a. DIM = {"\*"}, TIME = { " < 10 > 24 " }
    - b. YEAR = { " \$(=max( {1} Year ) ) " }, TIME =

**6. FUNCTIONS**

**a. Numeric Functions**

- i. < Dimension = { " \$(=f(args)) " } >
  1. \$ is used to expand or interpret the function
  2. For example, for max year -1 **YEAR = { " \$(=max( {1} Year ) -1) " }**
    - a. Max takes into account current selection, so {1} means the whole application, disregarding selections
    - b. Must use TWO = signs
  3. For example, for rolling 12 periods { <Month =, Year =, TIME = { " > **\$(=max(TIME)-12) <= \$(=max(TIME)) " } > }**
  4. As this can become tedious, recommended to use FLAGS

**b. Function returning members**

- i. { < Dimension = { " =f(parameters) " } > }
  1. Text function is between quotes → no \$
  2. WHERE set analysis, FOR EXAMPLE, we want to aggr the stores whose value sales are greater than \$100,000 based on current selection
    - a. → sum ( { < STORE = { " =sum(SALES)>100000 " } > } SALES)
  3. INCEPTION set analysis, FOR EXAMPLE, we want to aggr the stores whose value sales are greater than \$100,000 for cheese cake and ACC during January 2013
    - a. → sum ( { < STORE = { " =sum( { 1 < TIME = { "01/2013" } , PRODUCT = { 'Cheese cake' , 'ACC' } > } SALES)>100000 " } > } SALES)

4. RANKING set analysis, FOR EXAMPLE, top 20 products

a.  $\text{Sum} ( \{ < \text{PRODUCT} = \{ \text{"=rank(sum(Sales),4) <= 20 " } > \} \text{ SALES} )$

7. USING P() and E()

a.  $\{ < \text{Dimension1} = \text{P} ( \{ \text{set to select } \} \text{Dimension1} ) > \}$

i. Returns members that have (or don't have) data, based on associative model. Think associated (white/light grey) versus not associated (dark grey)

ii. FOR EXAMPLE, show sales for ALL items, but only for customers who have purchased croissants at some time (aka croissant is a possible/associated selection)

1.  $\rightarrow \text{sum} ( \{ < \text{CATEGORY} = \{ "*" \} , \text{CUSTOMER} = \text{P} ( \{ 1 < \text{CATEGORY} = \{ \text{'Croissant'} \} > \} \text{CUSTOMER} ) > \} \text{ SALES} )$

8. TWO FIELDS

a. ... refer to master guide ....

9. IF STATEMENT

a.  $\text{Sum} ( \{ < \$ ( \text{vSet} ) > \} \text{ SALES} )$  where  $\text{vSet} = \text{if}(\text{GetSelectedCount}(\text{'PRODUCT'})=1, \text{'PRODUCT} = \text{' ; \$ '}$

i. The IF statement can't be in the set analysis but CAN be stored in a variable

ii. The statement above says, if one product selected, show total for all products, but if multiple selected, show for just selection