**Business Analyst learning Path**

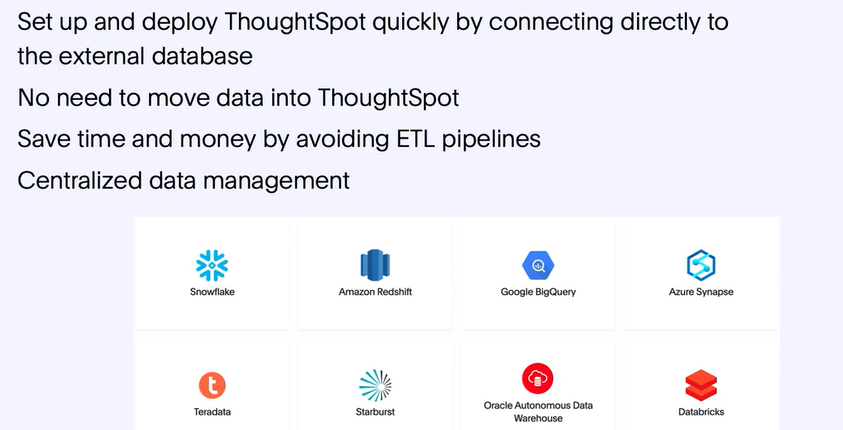
# Getting started with Thoughtspot Cloud Analytics

* Connecting to Your Data
* Setting up Joins
* Creating Worksheets
* Searching Your Data
* Using the ThoughtSpot Cloud Admin Console

## Connecting to data:

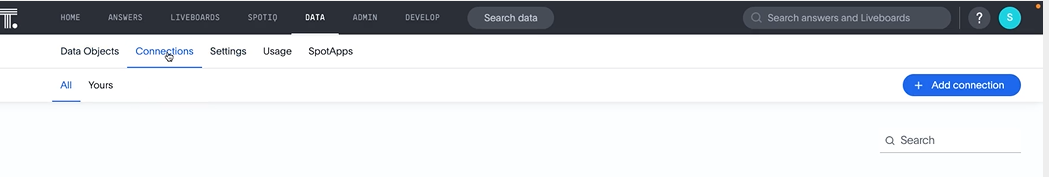
Use case: connection to snowflake

* DIRECT connection to external database
* With snowflake, we can create a linked connection
* Joins, pk are important in thoughtspot
* Some datasources below…



**How to create a connection?**

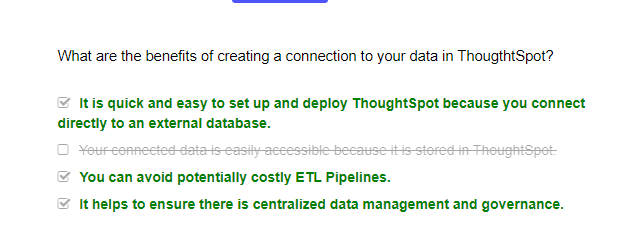
**Goto data> connection> add**

****

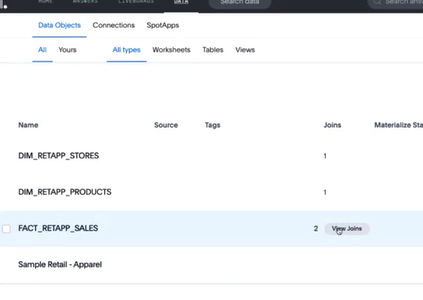
**Note: If tables are joined in snowflake, we can directly search by selecting the tables**

If not, we have to setup joins in Thoughtspot

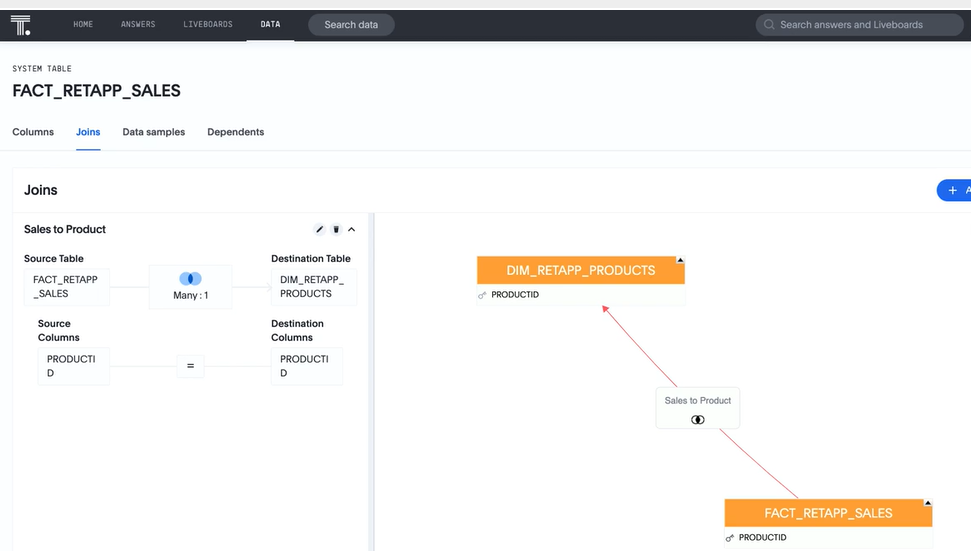
### Quiz:

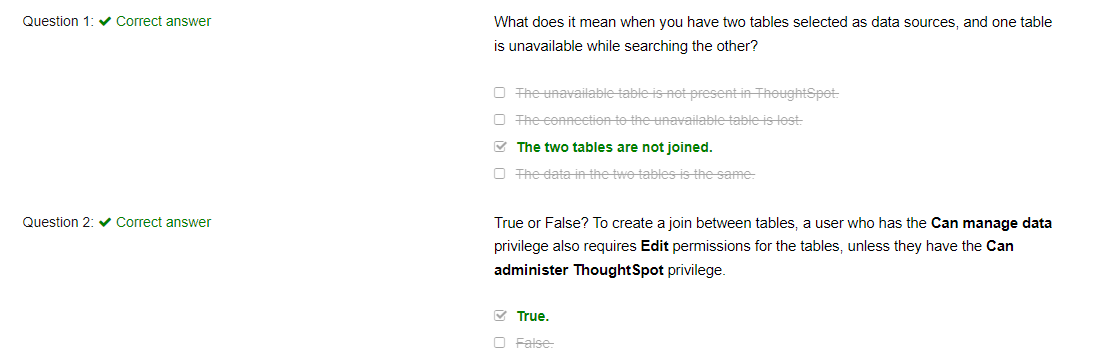


## How to create JOIN

* Join tables enables search to be conducted using both the tables
* How can we find if join exist :
* Option#1: if you select one table in dta source and other table is greyed out, joins donot exist
* Option#2: under data tab, > Join > select view joins\* 

1. **Access: can administer /can manage data priviledge to create own join (In addition to can manage , also need can edit on table to join)**
2. Start with fact table for join





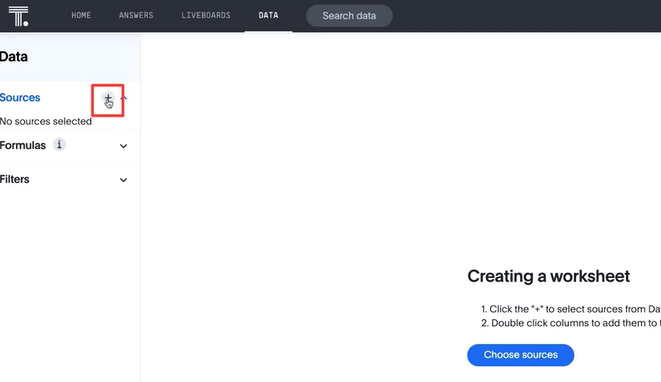
## Creating worksheet

Why?   make the data more user-friendly by changing the names of columns, removing underscores, adding formulas, and so on.Also, access control

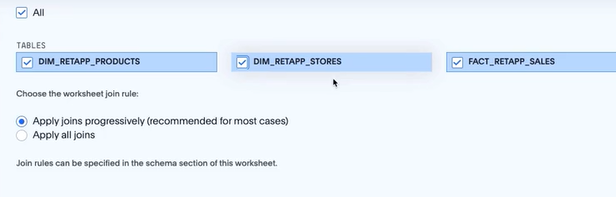


**Worksheets**

1. Group tables in logical way (tables or imported files can be used to create worksheet)
2. Allows user to search without users knowing the underlying layout
3. **How?** Data > elipses > create worksheet

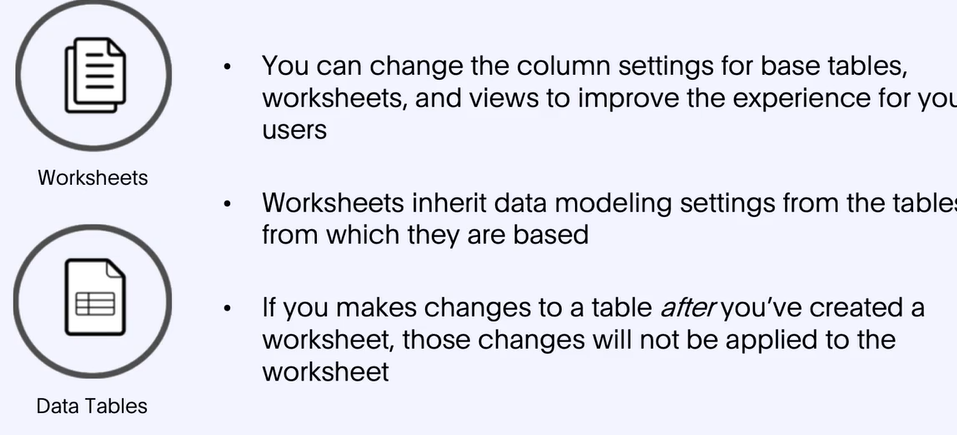
****

1. **“Join table progressively”** – means table will be added to the query only when pulled in the query (join all tables) – all tables are incuded irrespective of the fields selected

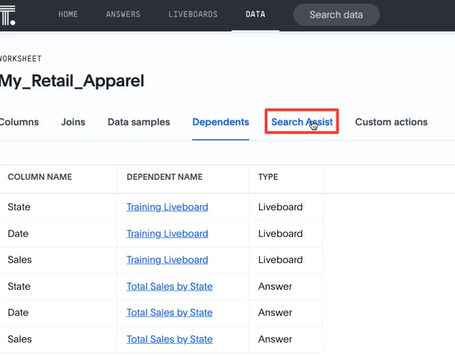
****

1. **COLUMN SETTINGS**

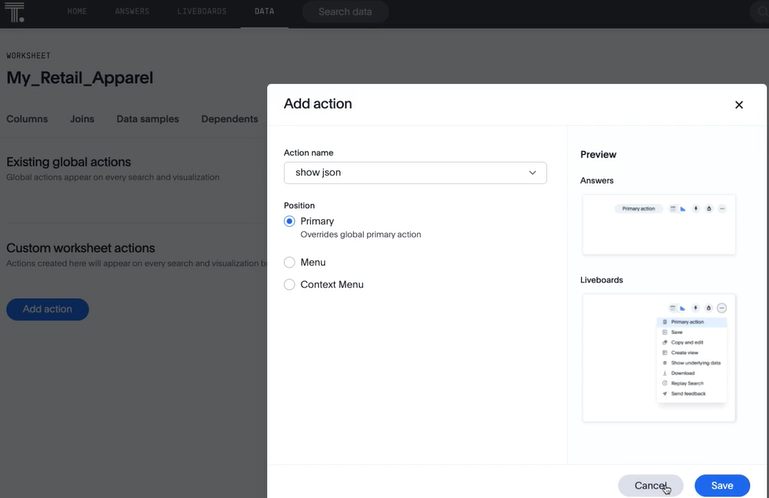
* Column setting applied to tables, work sheet and views



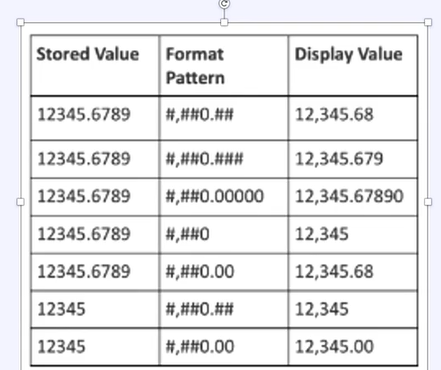
1. **We can see “dependants” of a worksheet from Data > All types> worksheet > Dependants**
2. **We cannot delete worksheet if Answers are present on it and we cannot delete a table if worksheet is based on it**

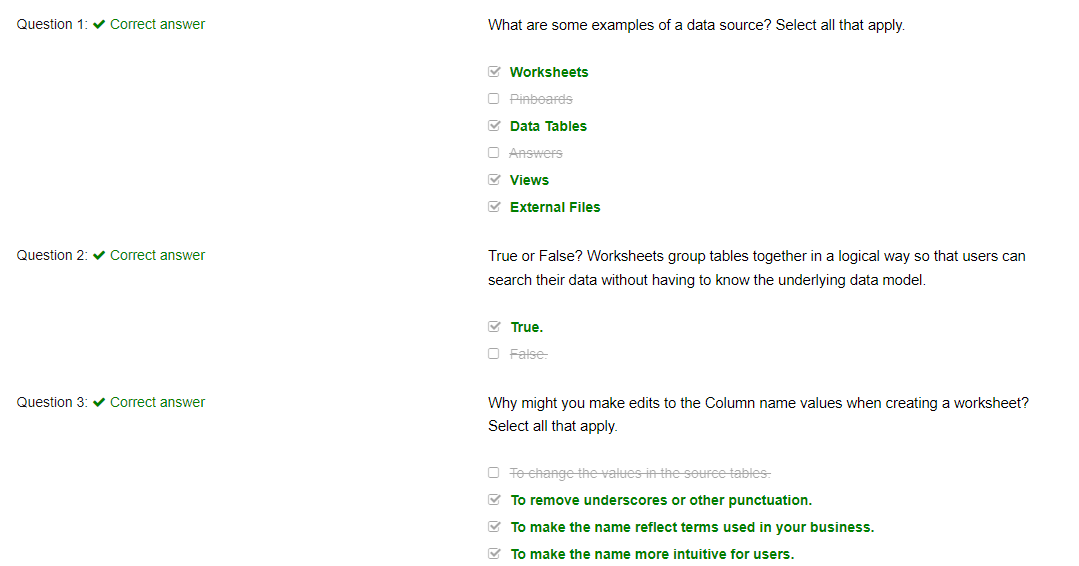
****

1. **We can “custom action”** to triggeran action

****

1. **“Synonym**” to a column helps to search a project by “product” or when typing “item” (separate by comma)
2. **“Format pattern”**



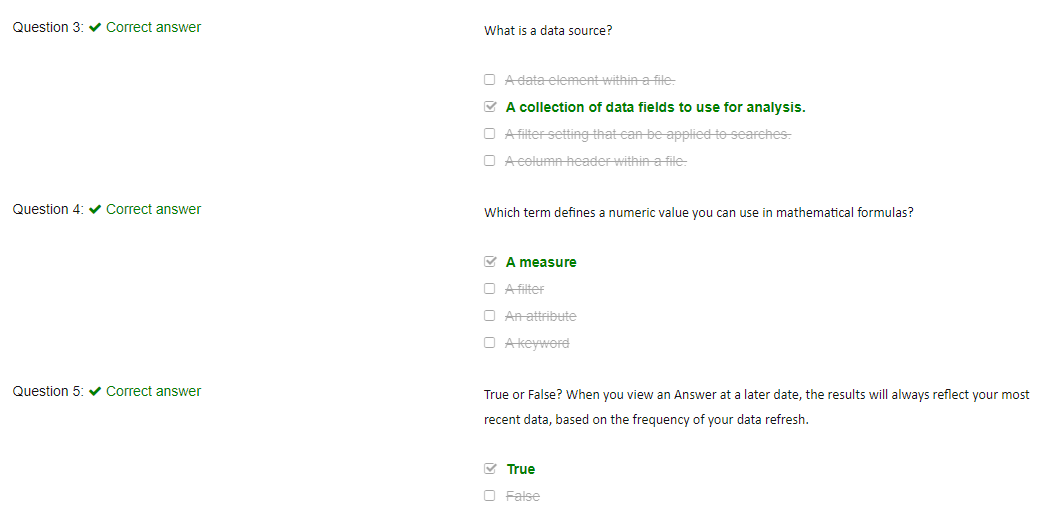


## Search, answers, pinboards

* To pin KPI, use **column Summary**(from settings), Not the table (use pin icon over column summary to pin it)

### Quiz

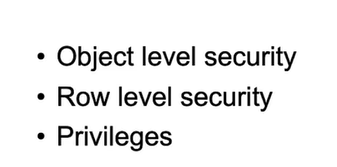




## 5.Admin console

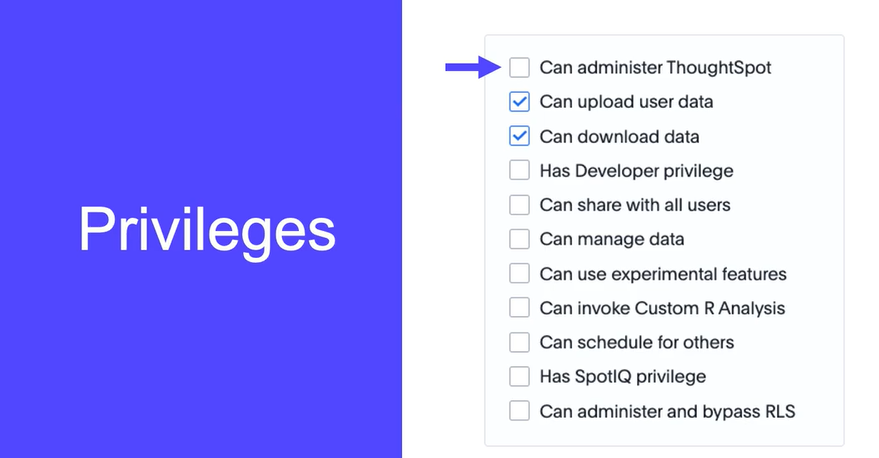
**What is covered:** In this lesson, we will cover the Administration page available for ThoughtSpot Cloud. The cloud admin console provides administrators with an intuitive user-friendly interface to accomplish necessary tasks for administering ThoughtSpot. This lesson covers user and group management, which allows administrators to create, edit, or delete users are groups as well as authentication which allows administrators to configure local or SAML authentication. The predefined system activity Liveboards will be reviewed, as well as how to customize your ThoughtSpot environment.

### User and groups:

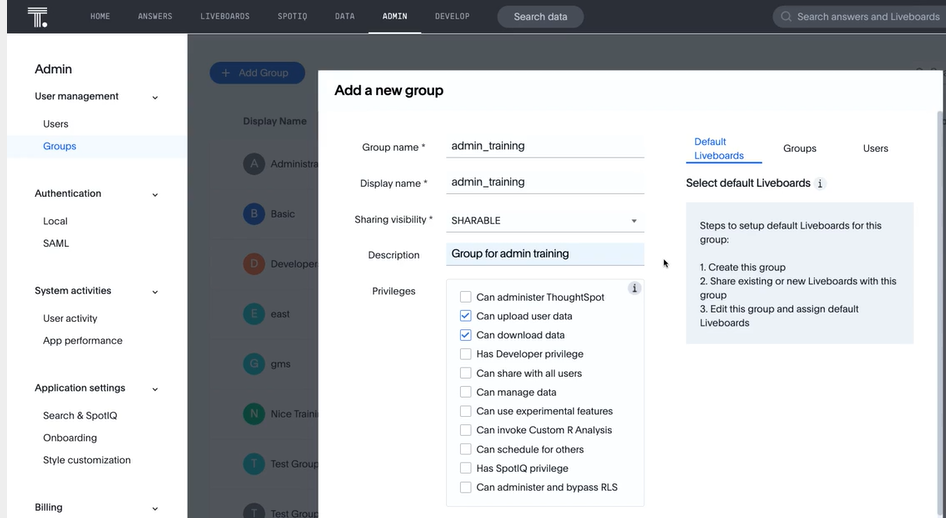
****

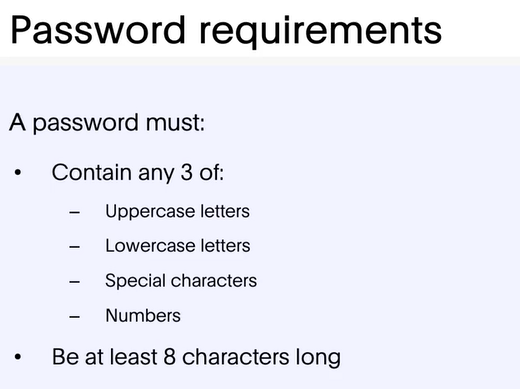
**\*NOTE : Privileges are set at group level, not user level**

**\*NOTE: If user is part of one group, highest privileges apply(unline BOBJ)**

Privileges

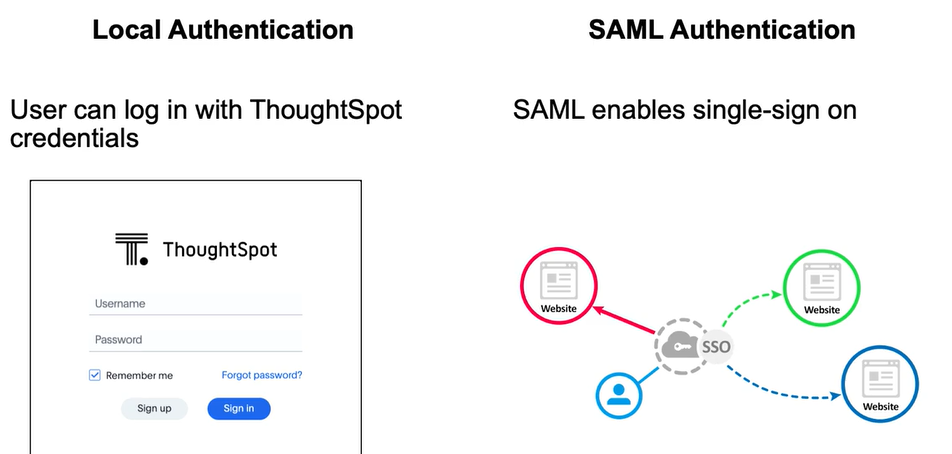
Default priviledges:





### Authentication:

1. SAML for Single sign on
2. Local authentication is local to TS



Some example: one login,azure, okta,

**SAML:**

To have the configuration persist with version, use the metadata XML File

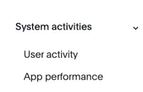
**NOTE:** when user is added using SAML, they can be added to TS user. By default, they are assigned to “ALL groups” users which doesn’t have any prioviledges.



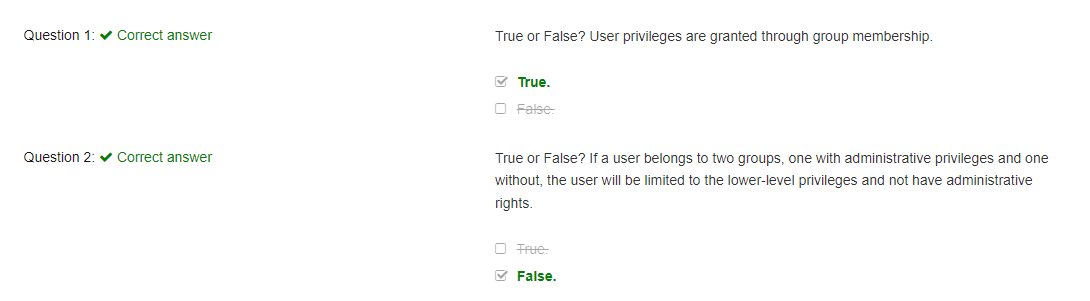
### System Activities

-only admin can access

* 1) User activity- understand users’ adoption of thoughtspot -active./inactive users, popular dashboard, more..
* 2) App performance –clusters’ performance,query ;latenct, user traffic,

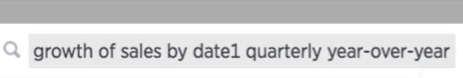


### Quiz



# Advanced Searches

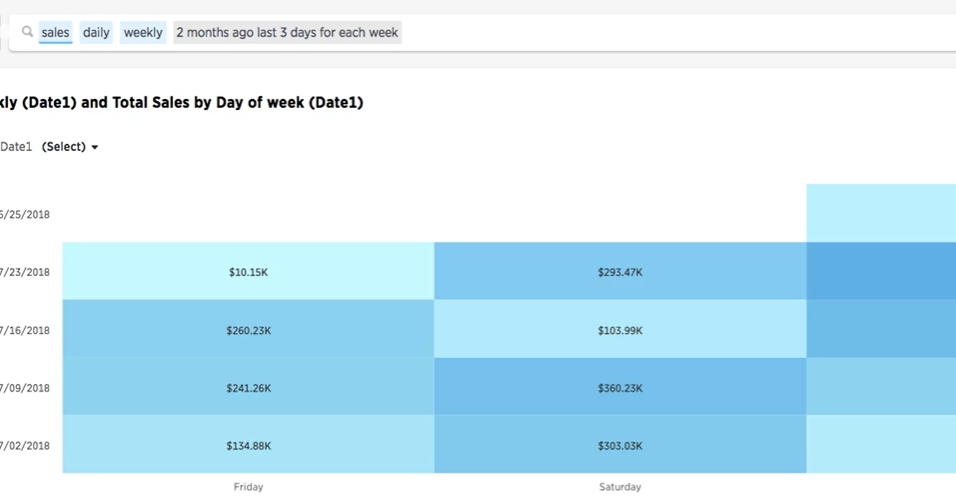
* Growth to compare past and current value
* **Keyword:”growth of” ,say ..Growth of <measures> by <date>**

****

Keyword : “growth of”; “by”; “year-over-year”..growth shows %

**Top/Bottom : Default is 10**

* How to get **4th quarter last year vs. this year (Ago keyword)**

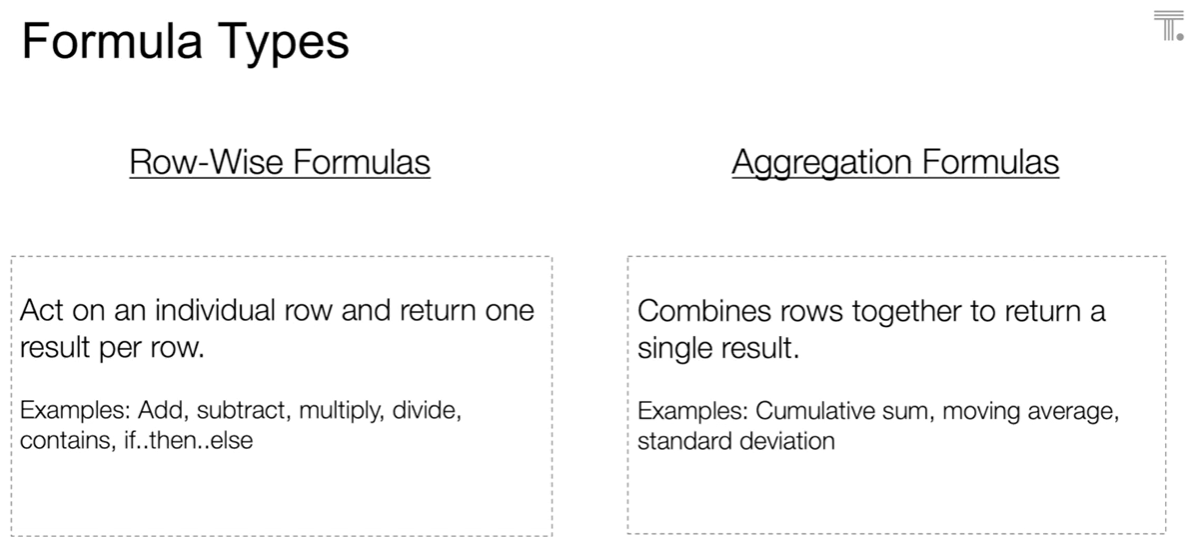


# Formula

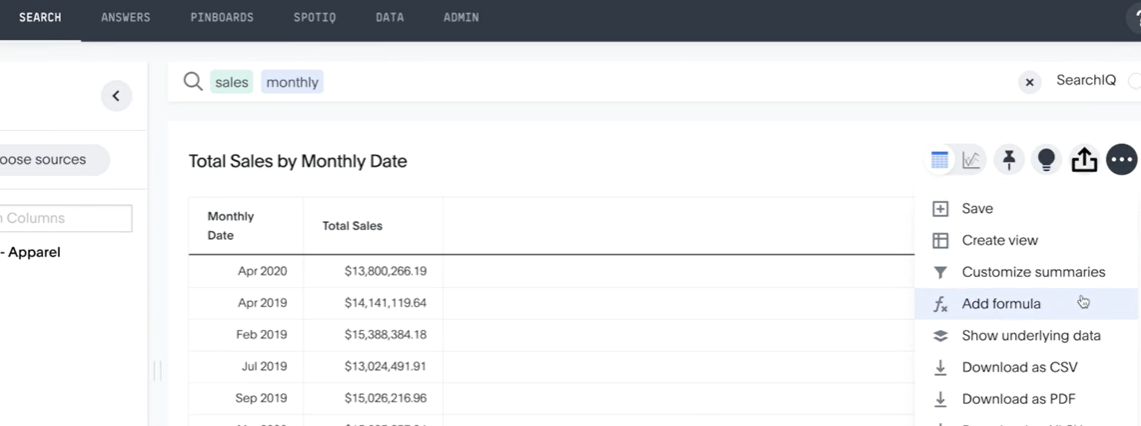
-derived columns using formula

**-keywords are in blue..purple means the field from the table**

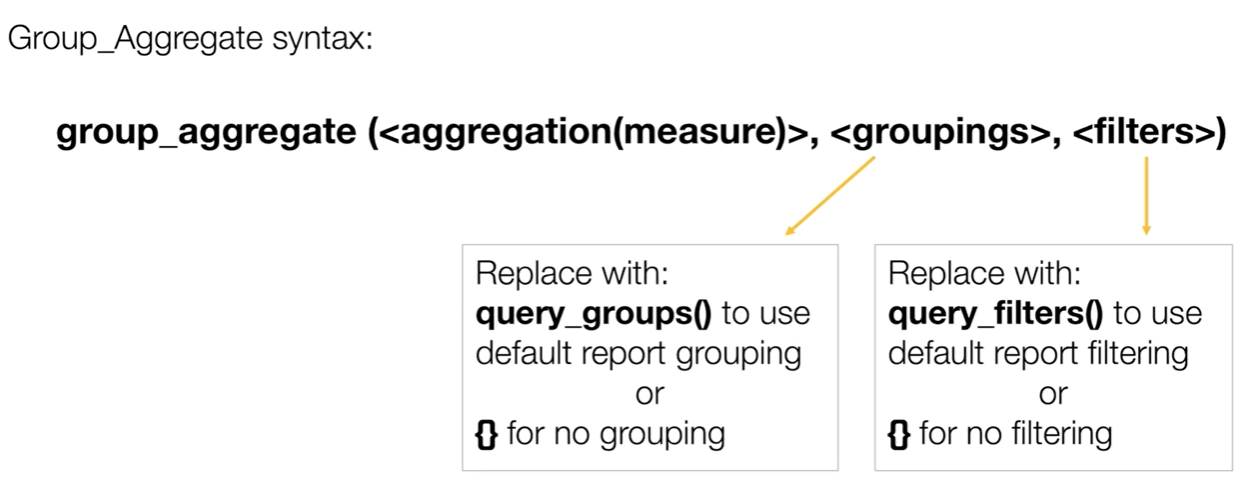
-can create formula on **search and worksheet**



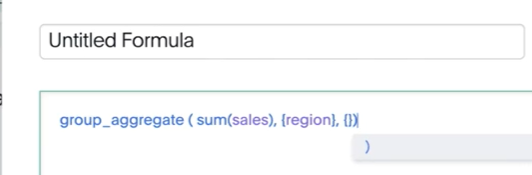
### Formula on search



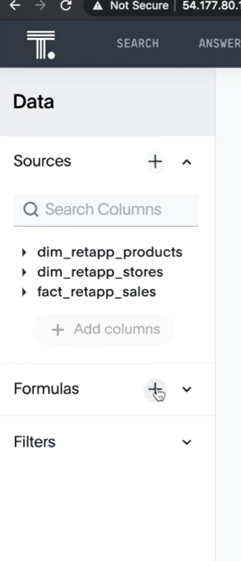
**Group**: used when we can define measure at different granularity



Use case: sales of a store to sales of a region…the search is by sales and store..we have to add a formula to get sales by region(higher granularity)



### Formula on worksheet



# Formula Challenges

**Dataset:** In downloads

### Challenge#1:

You are a sales manager interested in boosting sales by promoting the most expensive items sold recently.

**Hint:**

This data set does not include a unit price column. Therefore, we cannot use the average function. The function will require the Sales and Quantity columns.

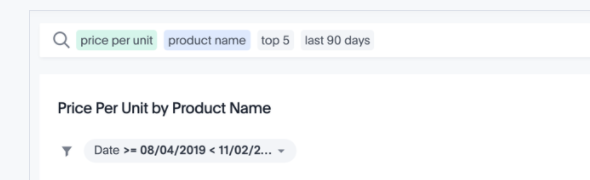
In the formula editor, click **Formula assistant** and review the **Aggregate** section.

Show me the Solution

**Steps to Follow:**

1. Create a formula named Price Per Unit: sum(sales)/sum(quantity)
2. Type the following search into the Search bar: price per unit product name top 5 last 90 days

**Expected Answer:** SanTop Telephoto Lens, Mother’s Choice Heart Medication, Portland’s Best Fishing Pole, Simplebuy Watch, Portland's Best Greeting Cards



### Challenge#2:

You are a sales manager and want to determine which state had the most profitable stores in the last 180 days.

**Ques:** What store state had the largest gross margin % in the last 180 days?

**Hint:**

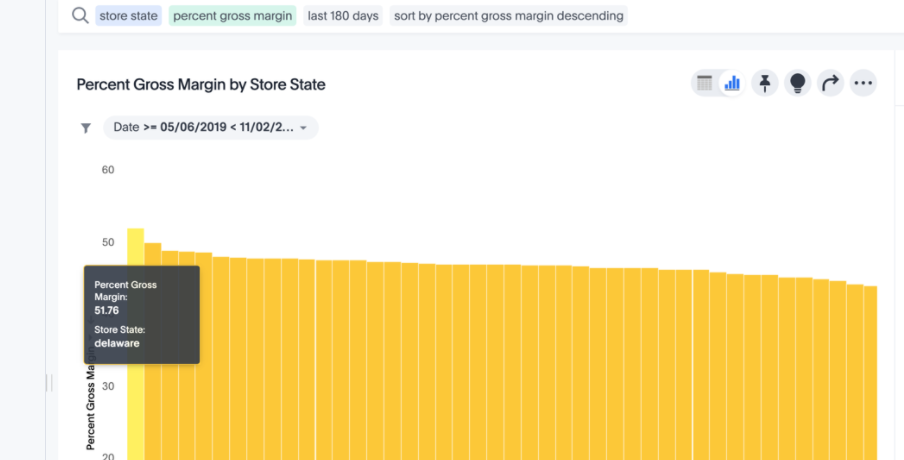
Create a formula for calculating gross margin percentage, where gross margin percentage is calculated as gross margin in dollars divided by sales in dollars. Multiply it by 100 to present it as a percentage. Remember to use sum to calculate all values in a column instead of the individual values.

In the formula editor, click **Formula assistant** and review the **sum** function in the **Aggregate** section.

**Steps to Follow:**

1. Create a formula named Percent Gross Margin: sum(gross margin ($))/sum(sales)\*100
2. Type the following search into the Search bar: store state percent gross margin last 180 days sort by percent gross margin descending

**Expected Answer:** Delaware 51.76%



### Challenge#3:

You are a purchasing agent who wants to determine how many unique products are sold across the country for each of your four product categories.

**Ques:** How many unique products exist for each category?

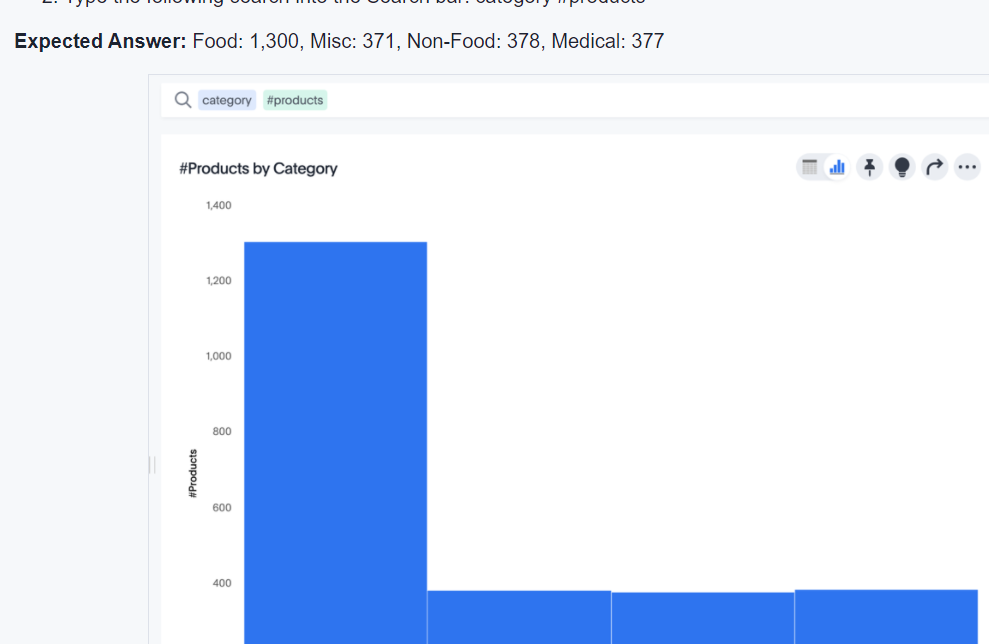
**Hint:**

Your formula should count each product, but must do so using the unique qualifier.

In the formula editor, click **Formula assistant** and review the **unique count** function in the **Aggregate** section.

**Steps to Follow:**

1. Create a formula named #Products: unique count(product name)
2. Type the following search into the Search bar: category #products



**Additional Notes:** If this calculation is only going to be performed occasionally, then the search can be performed directly in the search bar as: unique count product name category.

If the calculation is going to be performed often, then the formula would be added into the Worksheet to make it easier for all of the business users.

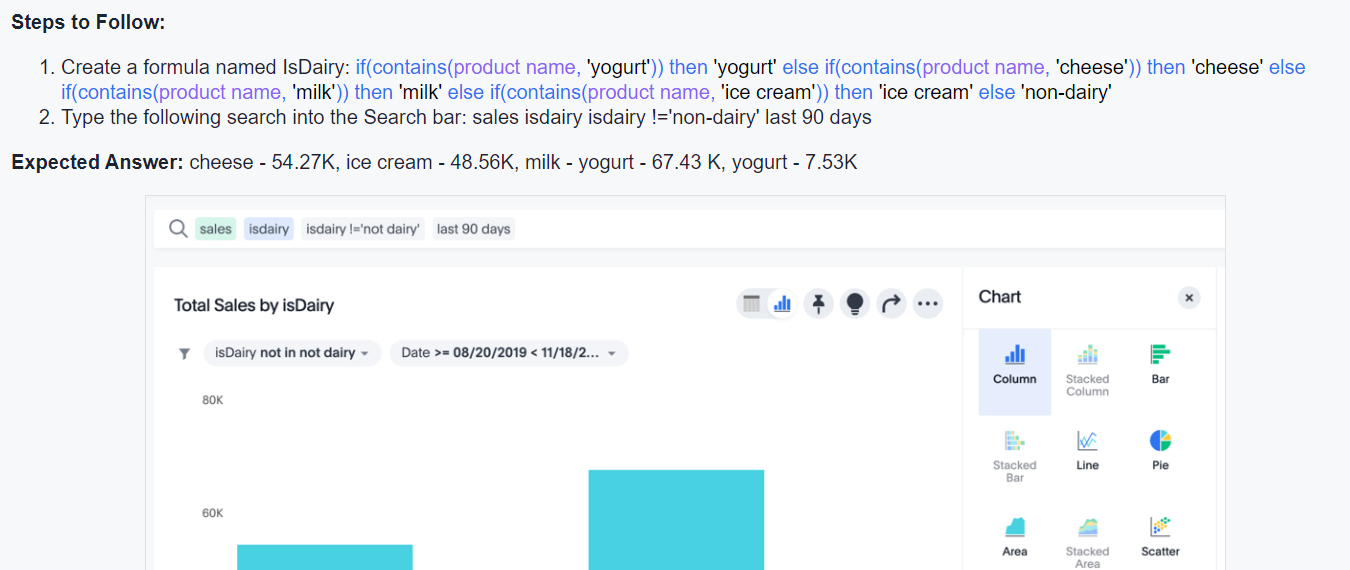
### Challenge#4:

You are a sales manager interested in determining sales for a group of dairy products.

What are the sales of dairy, including milk, yogurt, cheese and ice cream, over the last 90 days?

**Hint:**

In the formula editor, click **Formula assistant** and review the **if** function in the **Operators** section and **contains** function in the **Text** section.



### Challenge#5:

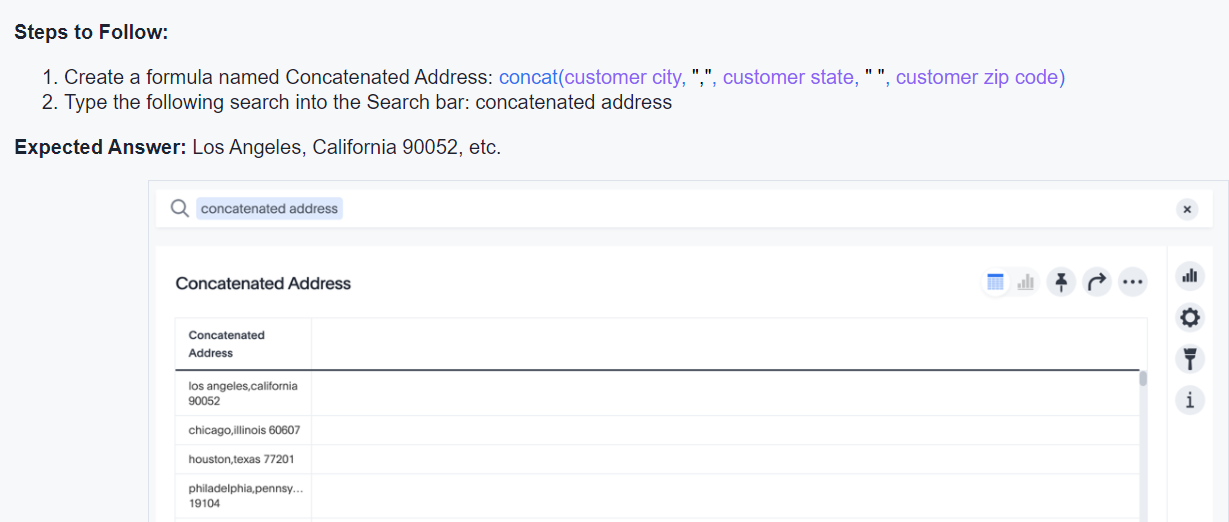
You are a business analyst and require an attribute in a specific format.

**Ques:**Create a new attribute (City, State Zip) in the following format (Sunnyvale, California 94085) using customer city, customer state, and customer zip code.

**Hint:**

You can join text together using the concat command.

In the formula editor, click **Formula assistant** and review the **concat** function in the **Text** section.



### Challenge#6:

You are a marketing manager interested in finding out which customer age-range contributed the most to sales.

**Ques:** Given sales and customer demographics, show me which age group (e.g. a) 18-29, b) 30-49, c) 50-64, d) 65+) contributed to the highest sales.

**Hint:**

Multiple if statements can be nested using the if() then() else if() then() else() format.

In the formula editor, click **Formula assistant** and review the **if...then...else** function in the **Operators** section.

****

**Challenge#7:**

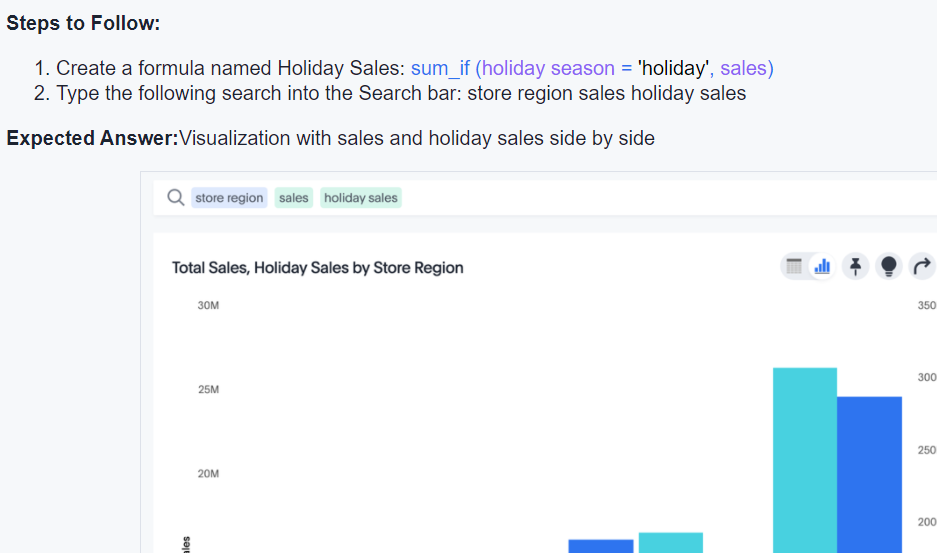
As a general manager, you want to determine how sales on holidays compare from store region to store region.

**Ques:** Show sales for a store region during holidays, along with total sales.

**Hint:**

The sum\_if aggregation can be used to add data together based on meeting a specified criteria.

In the formula editor, click **Formula assistant** and review the **sum\_if** function in the **Aggregate** section.

****

### Challenge#8:

You are an administrator, and want to clean up your data by replacing empty fields with text.

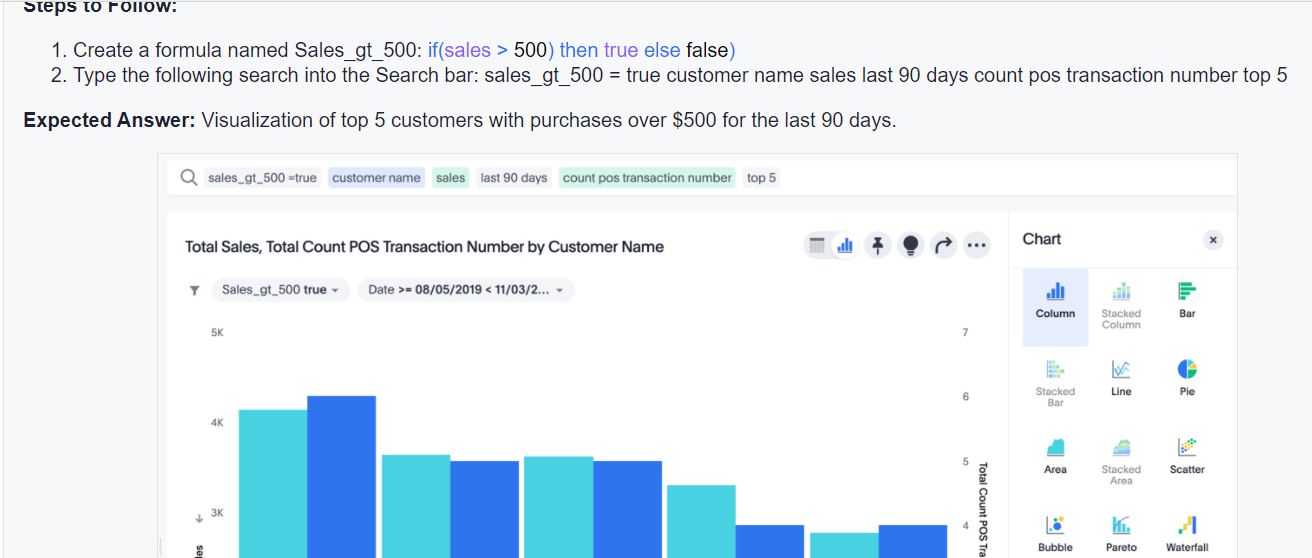
Ques: If a null value exists for Occupation, replace it with “unknown”.



### Challenge#9:

You are a marketing manager looking to run a targeted campaign by determining your top customers by dollar amount.

**Ques:** Find the top 5 customers making purchases of over $500 in the last 90 days.



### Challenge#10(Intermediate):

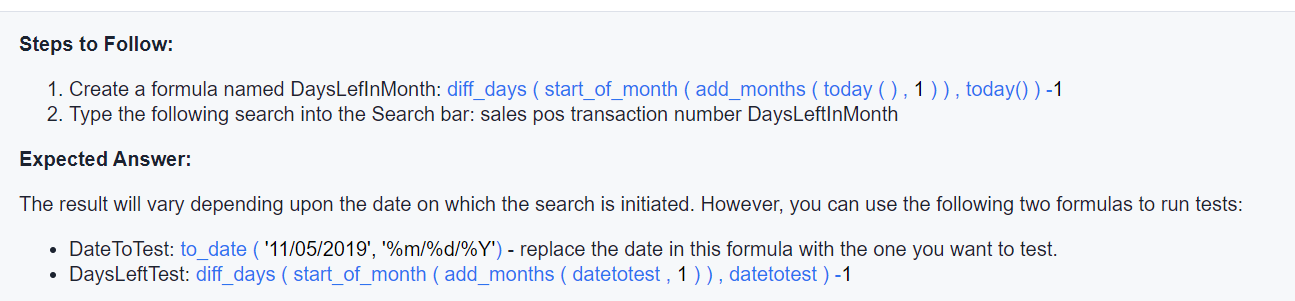
You are a manager interested in determining the projected sales for the current month.

**Ques:** Calculate the remaining days in the month.

Hint:

Add one month from today, get the first day of the next month, find the difference in days from the start of next month to today's date, and subtract 1.

In the formula editor, click **Formula assistant** and review the **add\_months**, **add\_days**, **diff\_days**, and **Today()** functions in the **Date** section.



### Challnege#11(Intermediate)

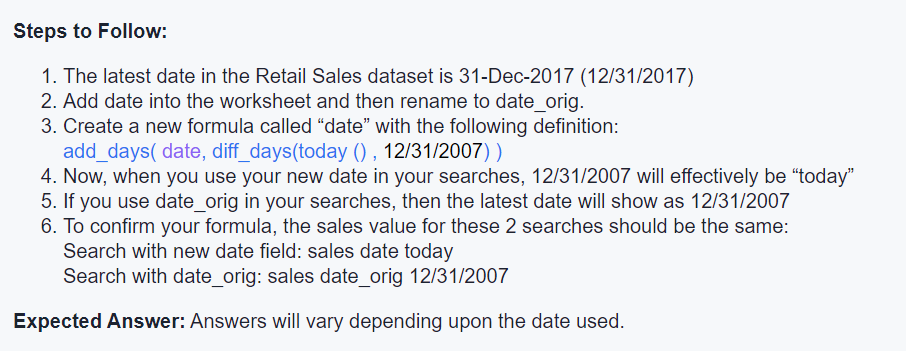
You are a BI team leader and want to use a demo dataset to demonstrate ThoughtSpot. You want to be able to use date keywords like last 7 days, last month, this month, this quarter.

Create a formula that will adjust the dates in the dataset to always appear “current” even though the latest dates in your dataset are from several years ago.

**Hint:**

Identify the latest date in the fact table. Use a diff\_days formula to calculate the # of days between the latest date in the dataset and today. Then adjust the date by that number of days.

In the formula editor, click **Formula assistant** and review the **add\_months**, **add\_days**, **diff\_days**, and **Today()** functions in the **Date** section.



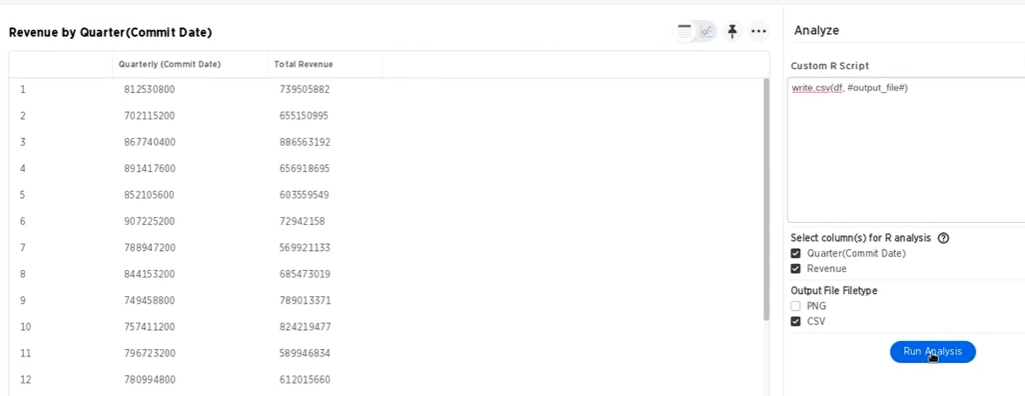
## Advanced Challenge

1. **Scenario:** The user wants to see how many days, on average there are between sales of a given product by different slices (e.g. region, state, etc.)
2. **Solution:**
3. (diff\_days (group\_aggregate (max (Date ), {Product} + query\_groups (), query\_filters ()), group\_aggregate (min(Date ), {Product} + query\_groups(), query\_filters ())) + 1)/ group\_aggregate (unique count (Date ) , {Product} + query\_groups(), query\_filters ()

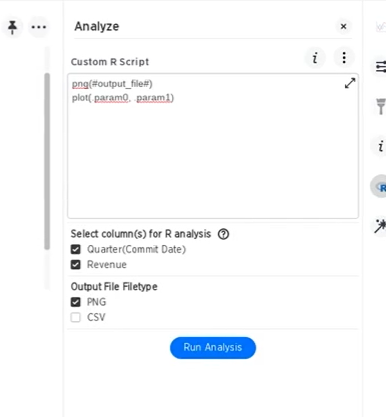
# Datascience in Thoughtspot

**Use R in Thoughtspot:**

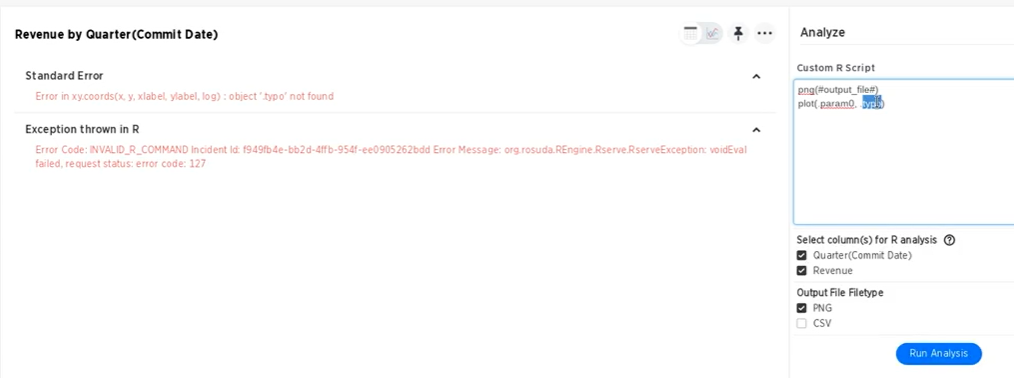
**R script as csv**

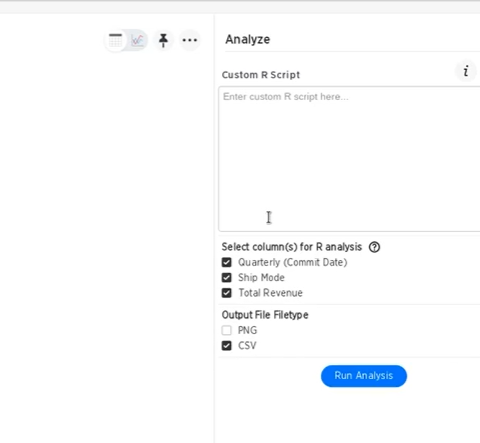
****

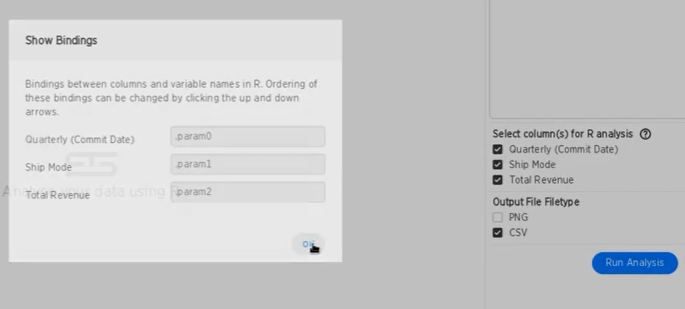
**R script as png:**

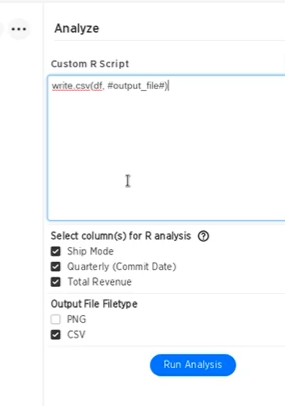
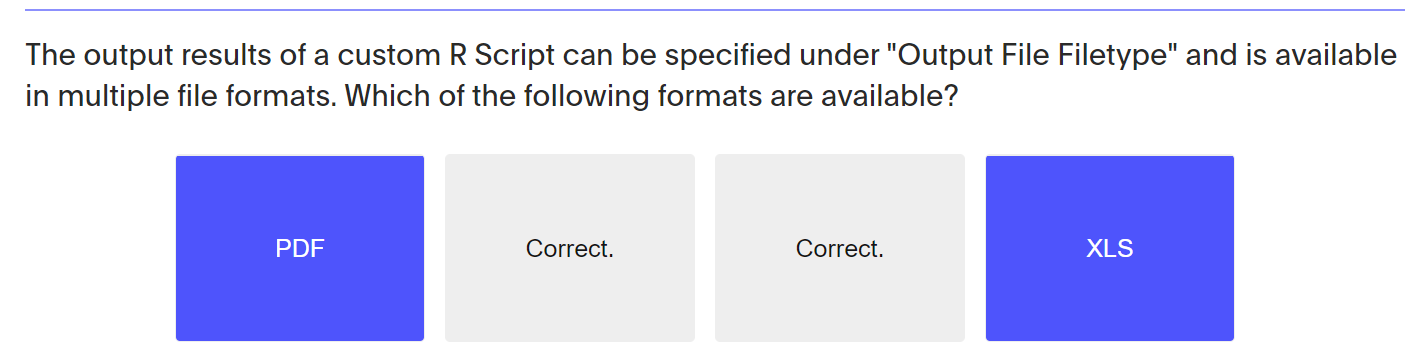
****

**To show error message in the generation(use “Typo”)**

****

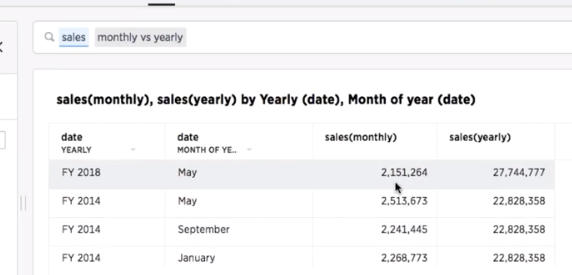
* **R script can be saved as “R” templates by using eclipse icon**
* The fields send from data source is listed under “select column(s) for R analysis”.
* The first column quarterly commt date can be reference in R vector using param0 , second column ship mode referenced by param1, ..so on
* ****
* **Question mark icon will give the bindings**

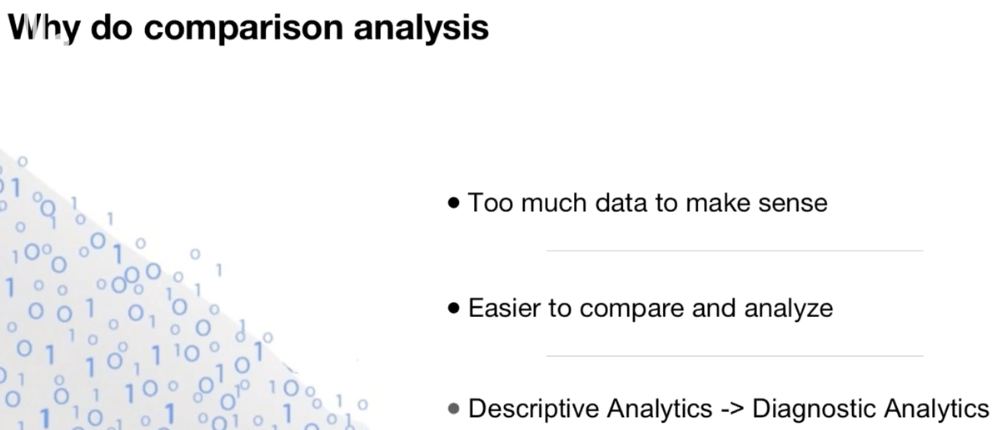
****

* **R automatically creates a data frame object for all selected columns**
* **Default system generated dataframe name is “df”**
* **Below code simply write the R’ dataframe “df” to output file refereced as “#output\_file#”(this is thoughtspot syntax integrated in R)**
* **\_**
* ****

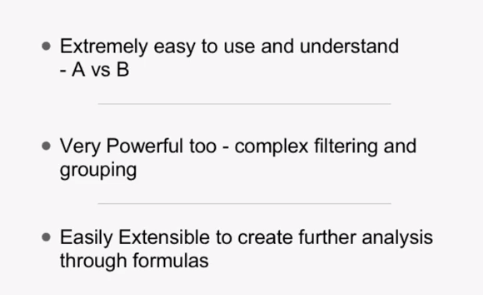
# 6. Comparison Analysis with Versus

4.5 version (complex analysis easy)

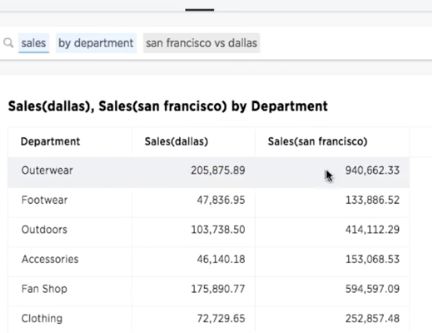




Descriptive analysis : what is happeneing, diagnostic analysis => why is it happening

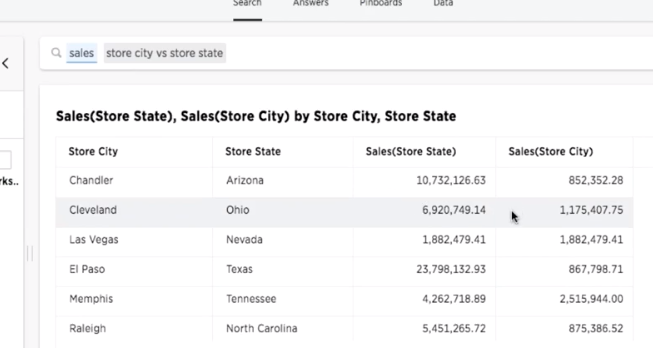


## Example: (compare same measure for same attributes)

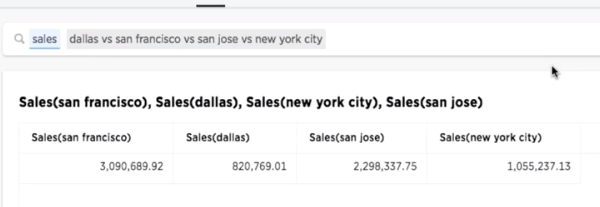
****

## Example: (compare different attributes)

**-sales is common measure…this shows sales for “store” and sales for “city”**

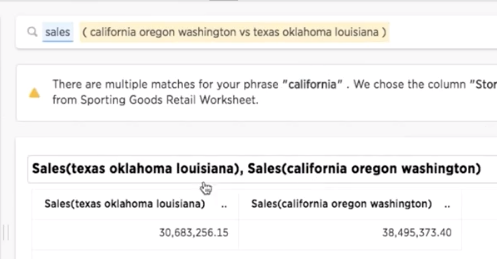
****

## Example: more than one vs.



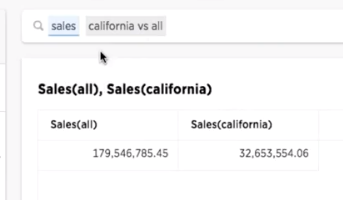
## Example: group different states into region

This compares set of states to other set of states

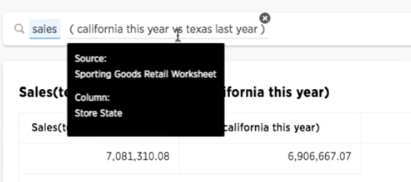


## Example: “ALL”

Sales at California vs. all



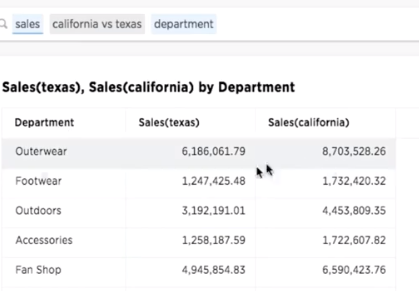
## Example multiple filters on each side of vs. clauses

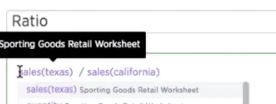


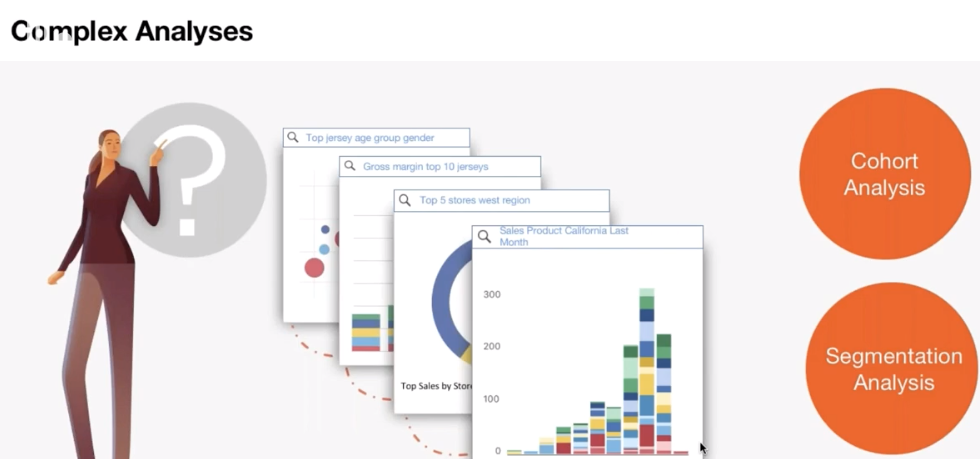
## Example:use search keywords and variables in the formula.

In same below, we say California vs. texas shows sales(texas) and sales(California) variables.

This variables can be used in formula to create ratio





****

## Cohort analysis

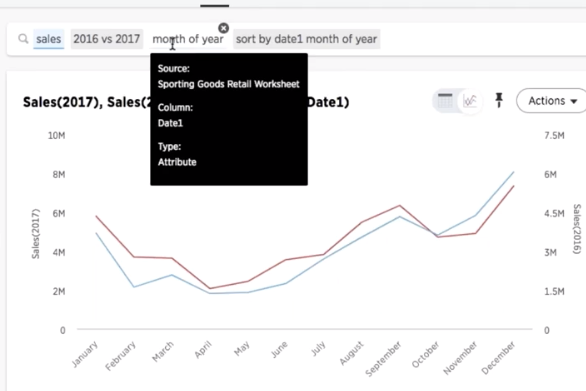
Number of months since the customer started, how is the revenue?

## Vs. affect performance?

Say, texas runs as one query and californai runs as separate query (more vs. means more query and so impacts performance)

More measures with vs. (more queries)

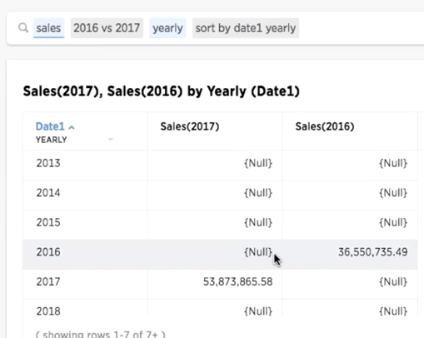
## Vs. used with date (YOY comparison)

****

**Don’t use “monthly” because monthly uses month+year and “month of year” uses just “month”**

## Syntax?

Measures – attribute –filter



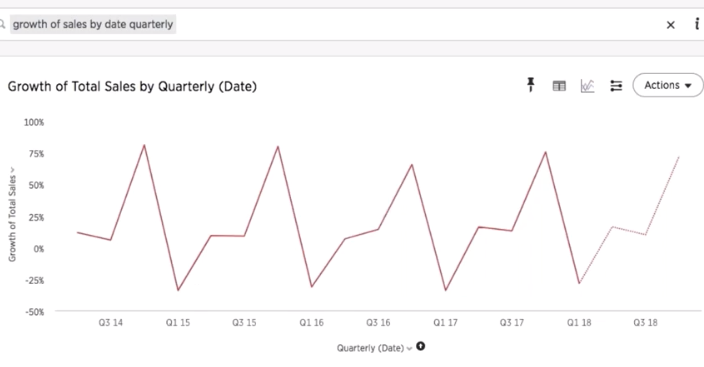
# 7. Time series Analysis

**Why do time series analysis?**

-study of data over varies period of time



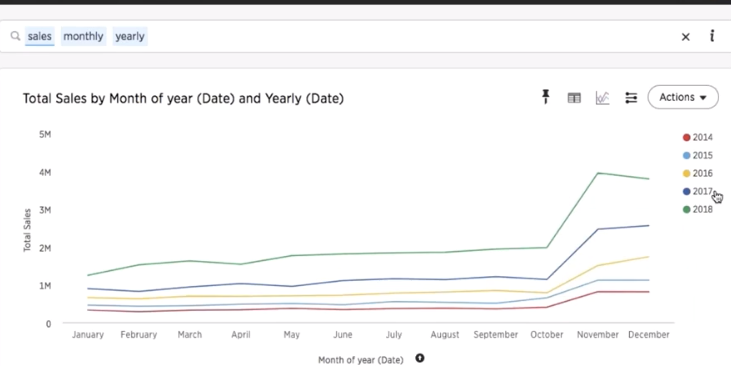
* Data from this year is more actionable than last year data
* **Time buckets : granutlairy of the date (month, daily etc..)[[1]](#footnote-1)**
* **Filters: absolute filter ; relative filter**
* **Growth : it shows % of growth**

****

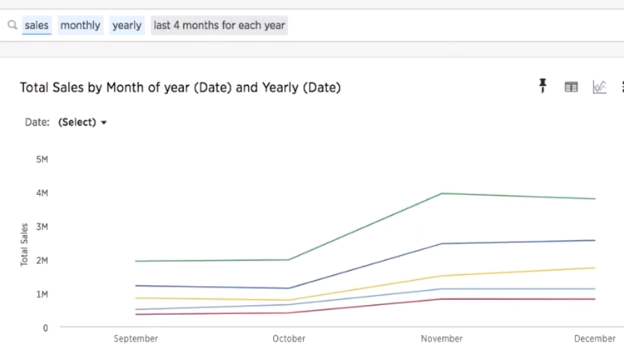
**New features in 4.4.2**

****

**Multiple date buckets:**

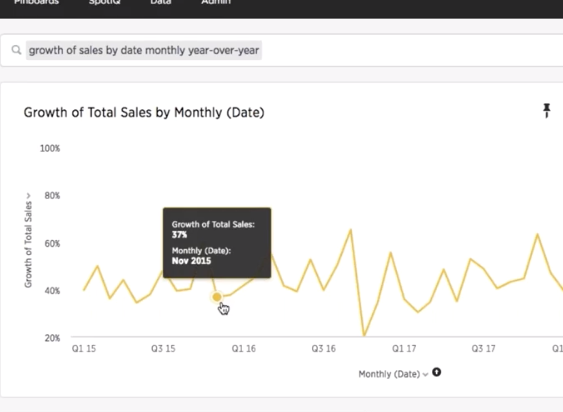
****

**Keyword :”FOR”**

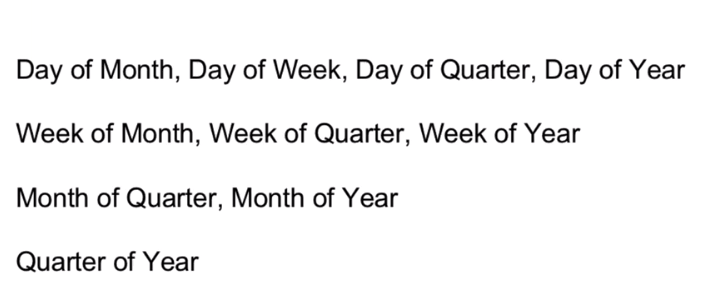
****

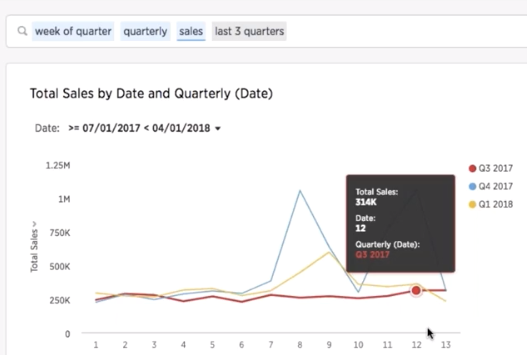
**Growth of one period over another period**

**Keyword : year-over-year ( meaning growth based on period of 2015 vs. nov of 2016)**

****

**Date buckets**

****

****

# \*\*\*Investigate : How to do this?

1. How to create join in snowflake and import the same in thoughtspot?
2. Is it better to use join in snowflake or thoughtspot?

– can we use both…main model. We can kee in snowflake and the crete alias in thoughtspot?

1. Can alias table be created to avoid loop? (fantrap?)
2. How is refresh frequency set up in TS?
3. If user using SAML is disabled, what happens to his answers, liveboards and schedules? Will it be automatically tagged to administrator?

# \*\*\*Features I like

Study cohort and document for TS use case

Can use developer tool to use TS answers and liveboards to share with a web application

\*vs. use dataset in “C:\Preethi\ThoughtSpot\Business Analyst- Webinar vs. materials” and try below

**Question to work#1**

We'd also like to give you the chance to build on what you learned. So, try using the practice data to answer these additional questions:

* How were the sales on each day of the week for the last 2 weeks of all the Novembers?
* What was my sales growth on a monthly basis, quarter over quarter for the last 8 quarters?
* What were the sales of my bakery department for the entire year so far? How does it compare to the sales for the same department till the same day last year? (Hint: Use formulas for the second part.)

**Question to work#2**

\* We'd also like to give you the chance to build on what you learned. So, try using the practice data to answer these additional questions:

* How were the sales on each day of the week for the last 2 weeks of all the Novembers?
* What was my sales growth on a monthly basis, quarter over quarter for the last 8 quarters?
* What were the sales of my bakery department for the entire year so far? How does it compare to the sales for the same department till the same day last year? (Hint: Use formulas for the second part.)

# \*\*\*Tableau vs. Thoughtspot:

**\*NOTE: If user is part of one group, highest privileges apply(unlike BOBJ)**

-custom sql can be done in tableau..TS? right now, we cannot edit the sql behind the answers

**To check later**

(webinar on how to customize data science

How to use worksheets to model intuitive searches)

<https://training.thoughtspot.com/webinar-comparison-analysis-with-versus/388061>

# \*\*\*Before Exam

1. [Formula Challenges (thoughtspot.com)](https://training.thoughtspot.com/formula-challenges-embed/1057649)

1. [↑](#footnote-ref-1)