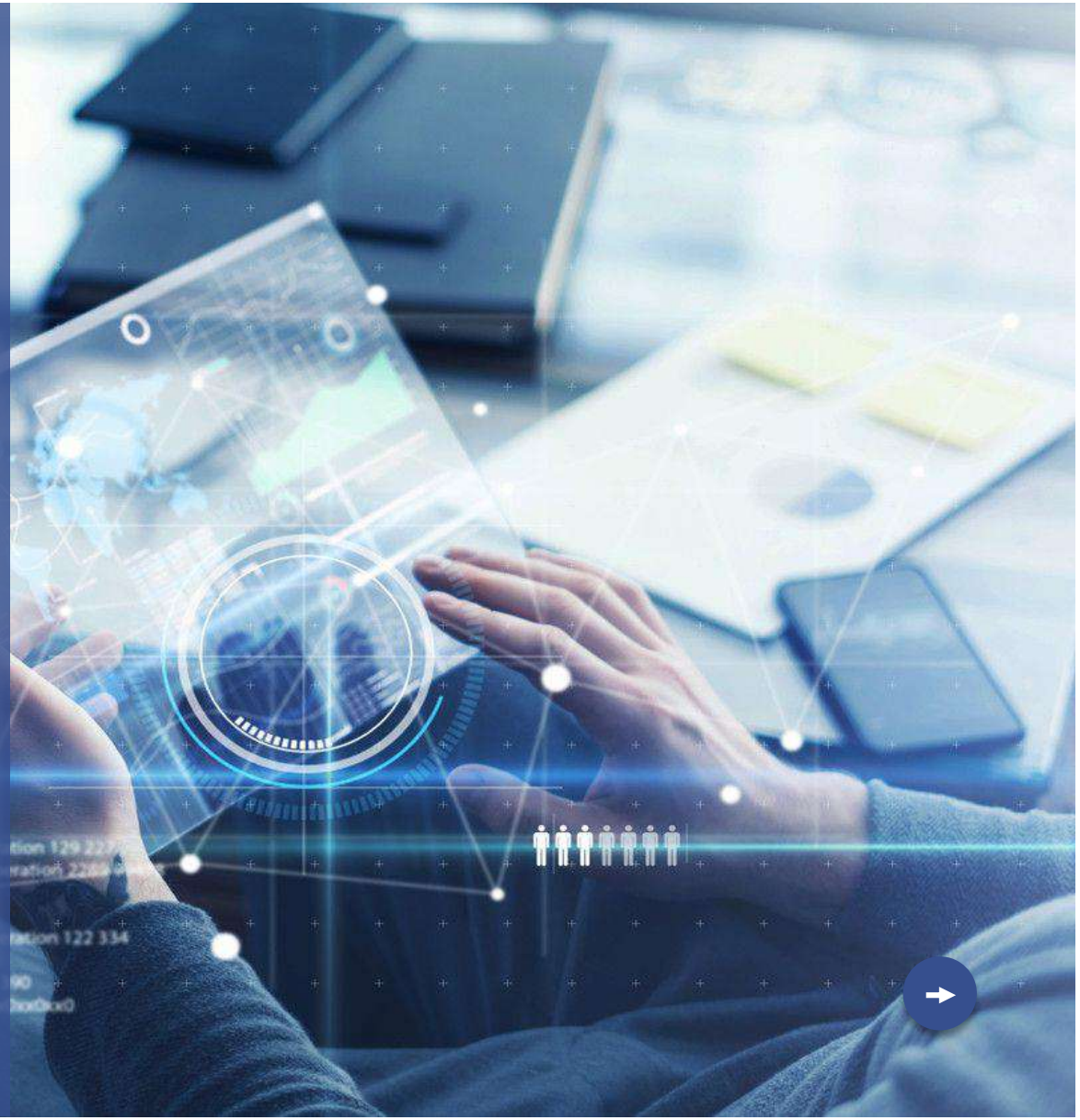
THOUGHTSPOT TECHIE TUESDAY'S

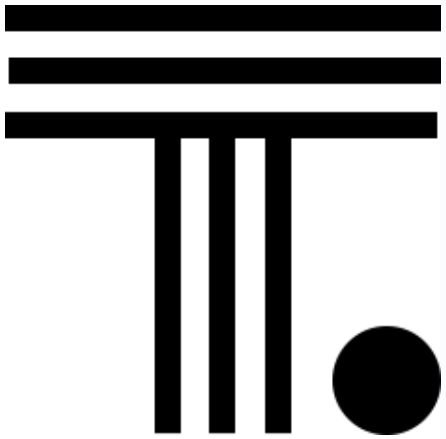
Want to know how to democratize access to insights in minutes and transform the way you uses data. Get your weekly dose of search and AI-Driven Analytics every Tuesday in 'Techie Tuesday' and ***unleash the power of 1,000 analysts*** *in your hand.*



ThoughtSpot

Techie Tuesday's

Introduction to ThoughtSpot



Once upon a time, a team of engineers who have worked for companies like Google, Oracle, Microsoft, Yahoo, etc. identified the future inclination of self-service analytics towards cloud-driven and software-driven storage.

In 2012, they founded a technology company that produces business intelligence analytics search software known as ThoughtSpot Inc. The company is based in Sunnyvale, California with additional offices in London, Bangalore, Seattle and Tokyo.

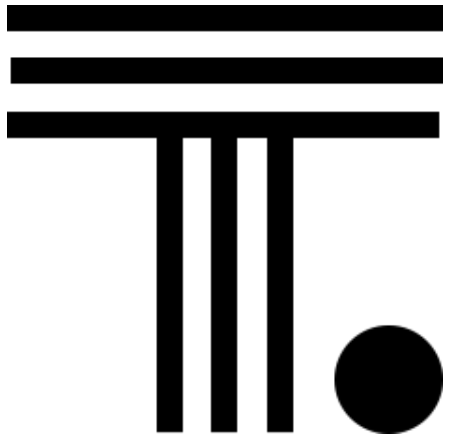
ThoughtSpot allows for non-technical individuals to conduct a self-service data analysis search. In 2016, ThoughtSpot was named a "Cool Vendor in Analytics" by Gartner. In 2017, the company announced that it was included on Gartner Magic Quadrant for Business Intelligence and Analytics Platforms report.

In 2020, ThoughtSpot has moved to the LEADERS position in the Gartner Magic Quadrant.

Watch out for our next flyer to know some interesting facts about ThoughtSpot.

Techie Tuesday's

Introduction to ThoughtSpot



As the fact goes, 'Data' is the lifeblood of any enterprise. Companies are experimenting with the technology-driven processes for analysing data and presenting actionable information.

While old-school BI capabilities like visualization and dashboards are still well-liked, the ever-growing mountain of data now needs a new interface. And that's exactly where 'ThoughtSpot' comes into the picture.

ThoughtSpot is a 'Search-Driven Analytics' platform that enables you to view and analyse your data through a search-based user interface. You can create your searches quickly and easily, just by typing them into a search bar, like you do when using an Internet search engine.

ThoughtSpot brings a Google-like search interface called 'SearchIQ' and combines it with 'SpotIQ', an AI-driven analytics engine, that makes the job of finding anomaly and pattern-matching as simple as it can be.

We will discuss more about '**Search IQ**' and '**Spot IQ**' in the next section.

Techie Tuesday's

SpotIQ | Bringing Artificial Intelligence to Business Intelligence

Previously in Techie Tuesday's:

- What is ThoughtSpot
- ThoughtSpot search interface (SpotIQ)

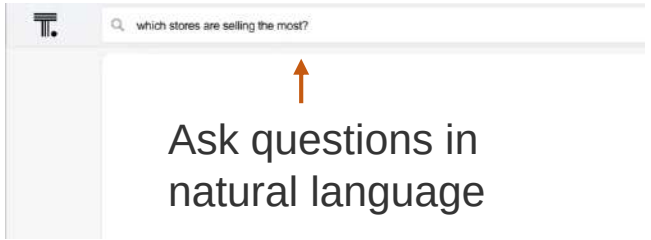
SpotIQ, starts with analytics features such as **anomaly detection**, **trend detection** and **cross-correlation** that could drive the most impact for BI professionals and the business users.

Anomaly or outlier detection: Enables users to find the 'unexpected' insights hidden in data. SpotIQ computes z-scores and linear regression model to find high and low outliers in a data set. The z-score method can be further customized with settings for min standard deviation, adaptively reducing the min std-dev if no outlier detected, min number of data points etc.

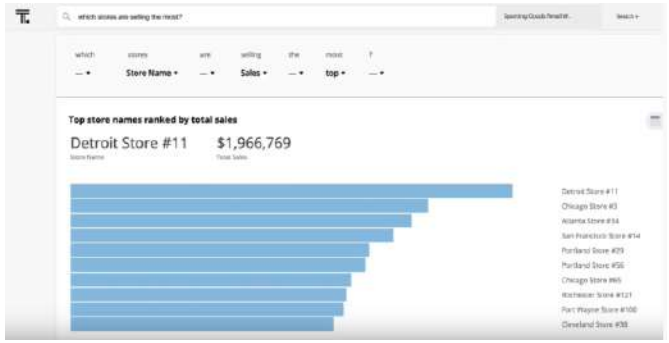
Trend detection: SpotIQ leverages the linear regression algorithm to detect trends, which can be further customized by specifying the minimum gradient/slope and the max p-value.

Cross-correlation helps get to the 'why' behind an insight by detecting other factors influencing a specific metric. These algorithms can be customized by specifying the minimum correlation coefficient and the max lag.

With SpotIQ



Ask questions in natural language



Get visualization to analyze data, build reports and dashboards – all in seconds.

Coming next in Techie Tuesday's:

- Going Under the Hood: SpotIQ's Analysis Methodology

nice Software Solutions Data Analytics Services Go Beyond Insights

Techie Tuesday's

Going Under the Hood | SpotIQ's Analysis Methodology

Previously in Techie Tuesday's:

- How SpotIQ is bringing Artificial Intelligence to Business Intelligence

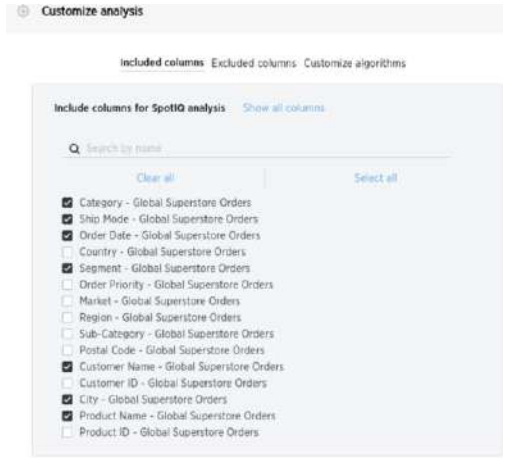
So how exactly does SpotIQ use capabilities such as Anomaly, Trend detection, and Cross-correlation to analyse data?

For a given query and its result set data, the analysis is first performed by drilling on the query attributes and measures. For example, for "Revenue by State" - the drill attribute would be "State" and the measure would be "Revenue". Any states with very high or very low revenue relative to the mean/median - would be flagged as outliers.

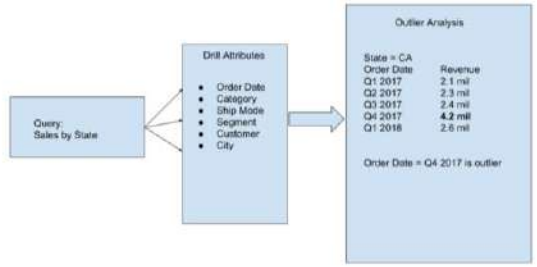
After selecting your first attribute, up to 10 additional attributes are added, one at a time, for deeper drill down. These attributes are obtained from the involved data tables and ordered by ranking algorithm based on factors such as contextual usage, static ranking, and overall usage.

For each of the datasets, any steady up or down trend is detected using linear regression.
The extent and magnitude of comprehensive analysis described here is what we like to call "bringing the power of 10,000 analysts to your fingertips".

The process of drill column selection



Outlier Analysis for the query "Sales by State"



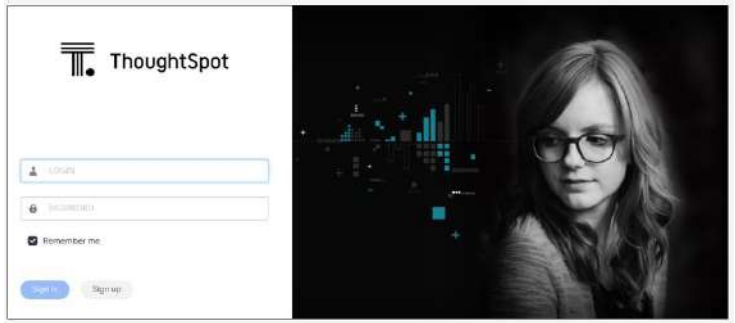
Coming next in Techie Tuesday's:

- Understand ThoughtSpot Architecture - Architectural Components

Techie Tuesdays

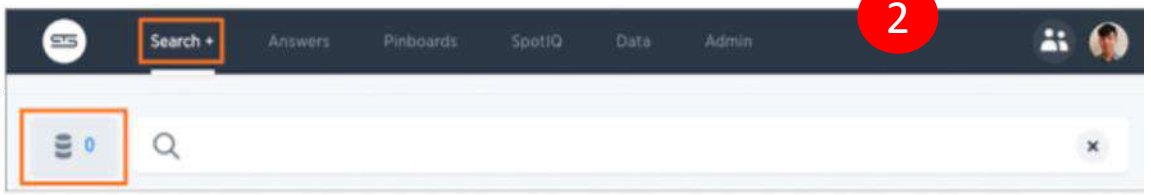
Search | Starting a new ThoughtSpot search is simple, like starting a new Google search

Like any other BI tool the rule is same here - Before you start a new search, make sure you have chosen the right data sources



1

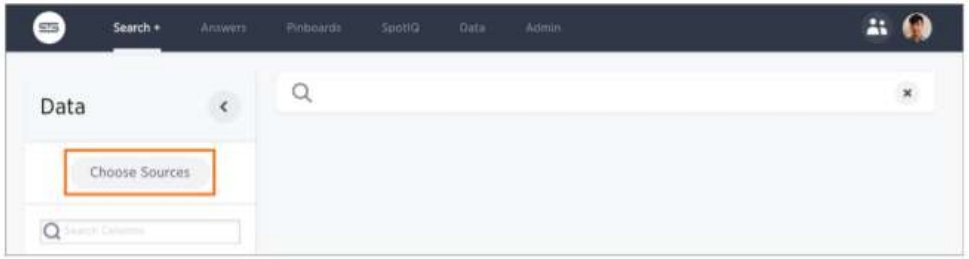
Log-in to your ThoughtSpot account



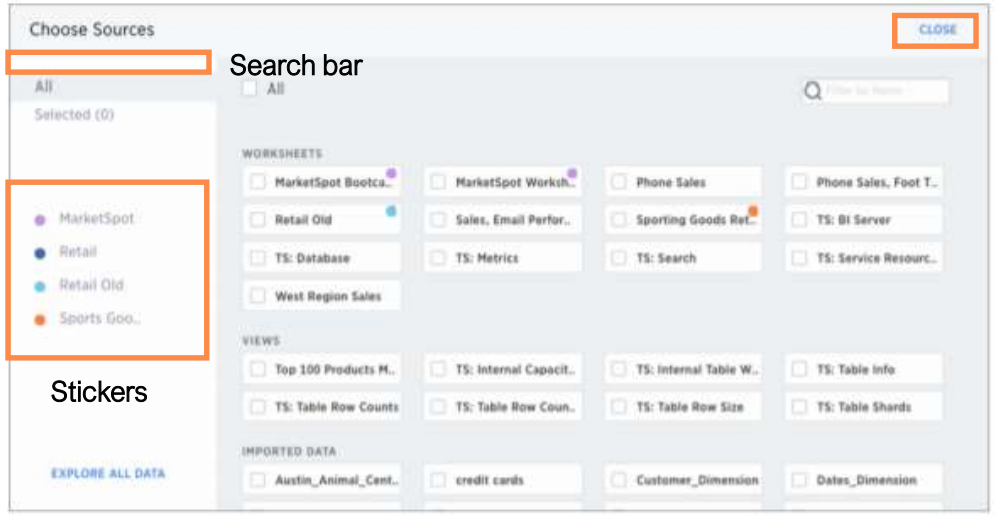
2

To begin a new search, you must first select your data sources by clicking the '**Data**' button.

3

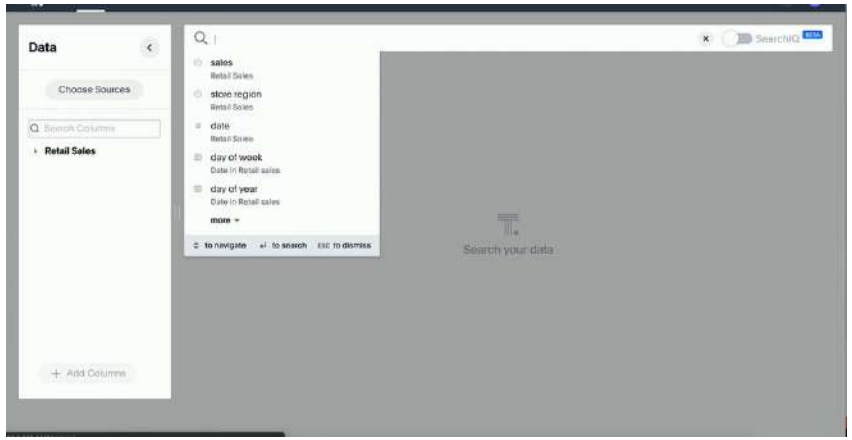


Now, click **'Choose Sources'** You will see a list of data sources that have been shared with you. The data sources are usually created by your administrator, though you can also upload your own data. (We will discuss 'How to upload your own data' in one of the coming sections)



4

Now, Filter through all the available sources by using the search bar or stickers. Select your sources and Click CLOSE.



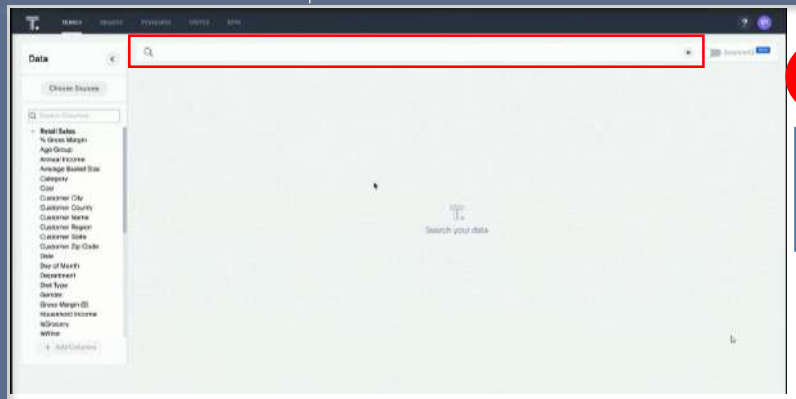
5

When your data has been uploaded, you can immediately append an existing table, search or explore your data.

nice Software Solutions
Data Analytics Services
Go Beyond Insights

Techie Tuesdays

Search | Starting a new ThoughtSpot search is simple, like starting a new Google search

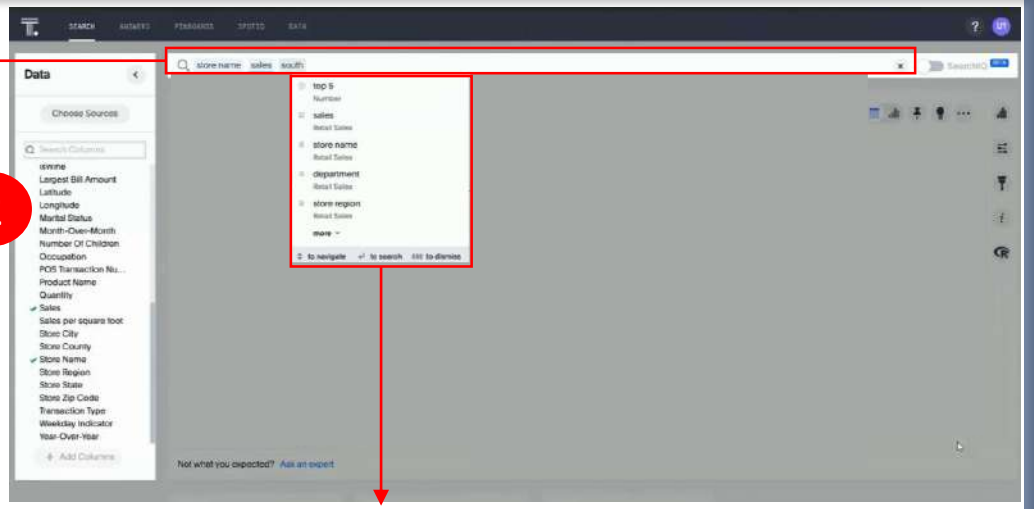


1

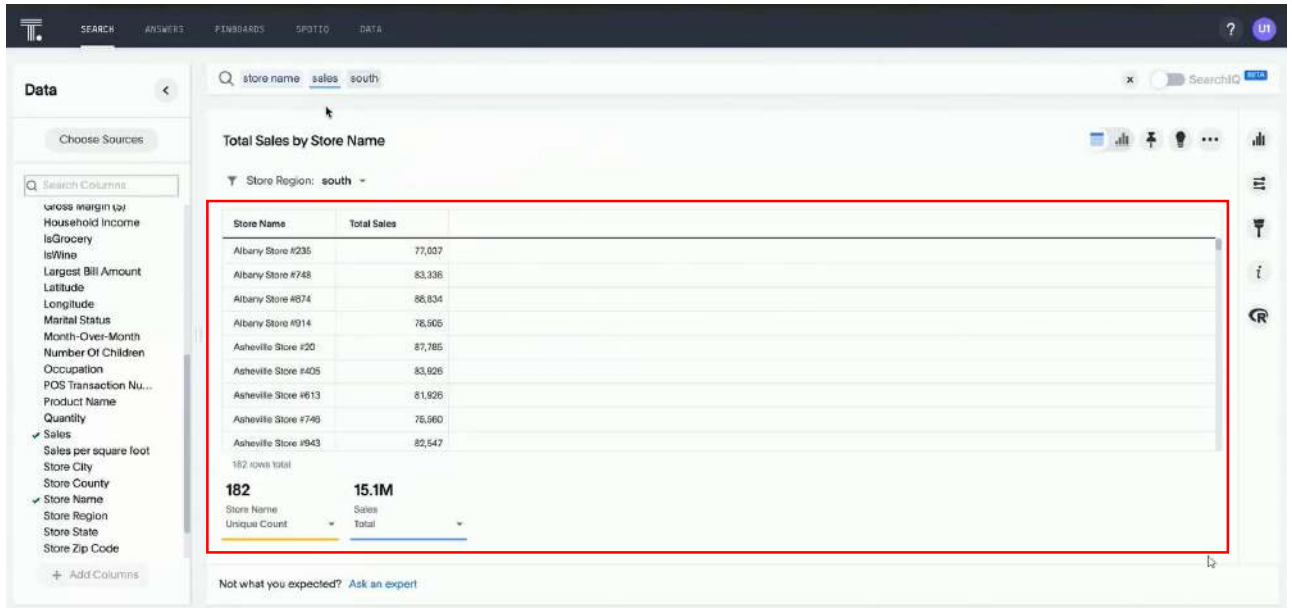
Here you can perform 'Keyword Search'

The search bar shows boxes around each search phrase, so you can easily see where it begins and ends. ThoughtSpot identifies search columns as either attributes or measures. The columns you choose impact your search results.

2



When you use the search functionality, your drop-down will be flooded with the suggestions the moment you type the first letter of any attribute or measure. These suggestions are based on 'Usage-based ranking. Search suggestions are relevant to the data and personalized to your search behavior. ThoughtSpot learns over time what columns are most important to you and to your company as a whole. Then, it uses this knowledge to rank the search terms it offers.

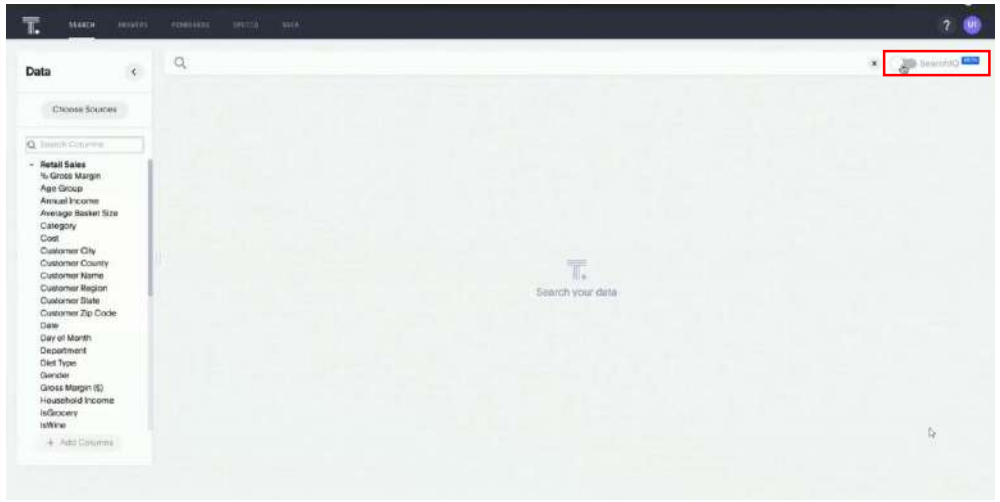


3

Based on the keywords entered, ThoughtSpot will display your data in table format or in graph whichever suits best.

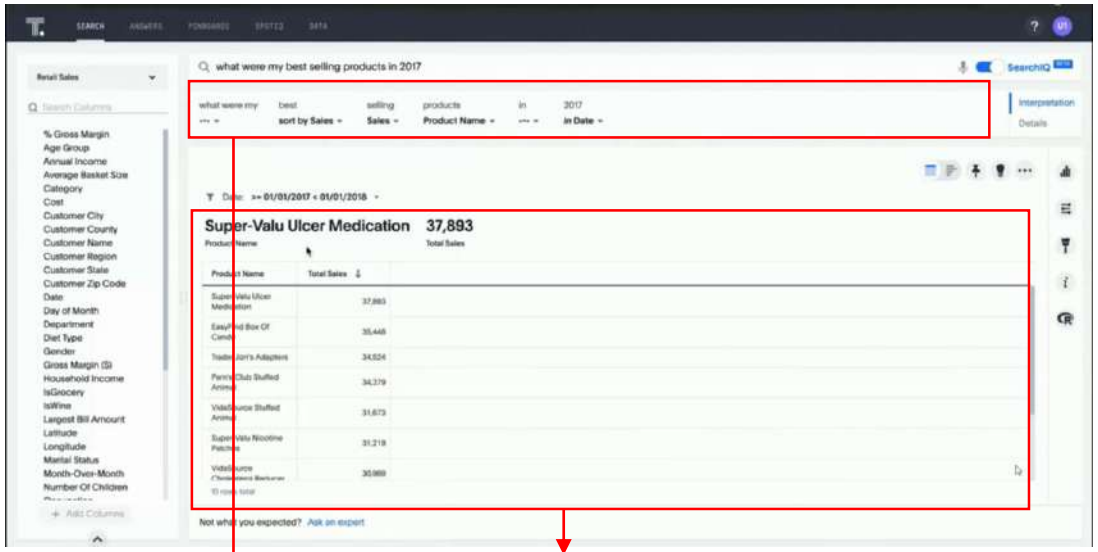
Techie Tuesdays

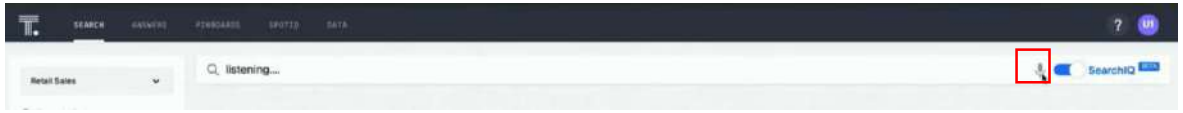
Search | With SearchIQ, you can search your data through natural language, just like speaking



1

Turn on Search IQ with this toggle





2

Click the microphone and ask your question. SearchIQ will translate your plain English into a search.

Note: You can type in a search using plain, speech-like language.

3

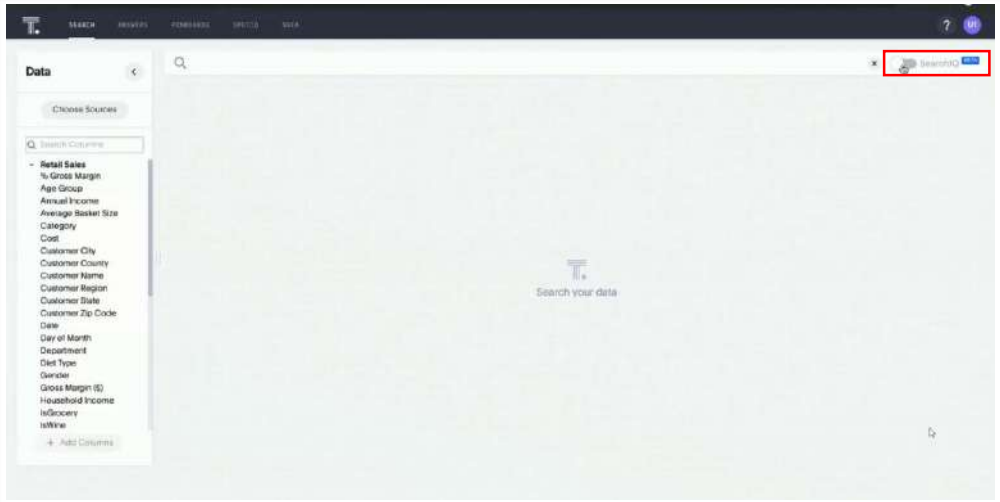
You can now see the results on your screen. This approach is similar to what we learned in the previous session. The only difference is, earlier we used keywords to create visualization and here we simply asked the question and get the relevant output.

This is useful for those who don't know much about data but want answers to their questions quickly

Unlike the previous search here you can see an additional section on the top. This section explains how ThoughtSpot perform search on your question. It simply ignores some generic words such as 'what were my' and 'in' and took best as '**sort by sales**', selling as '**sales**', products as '**product name**', and 2017 as '**in Date**'

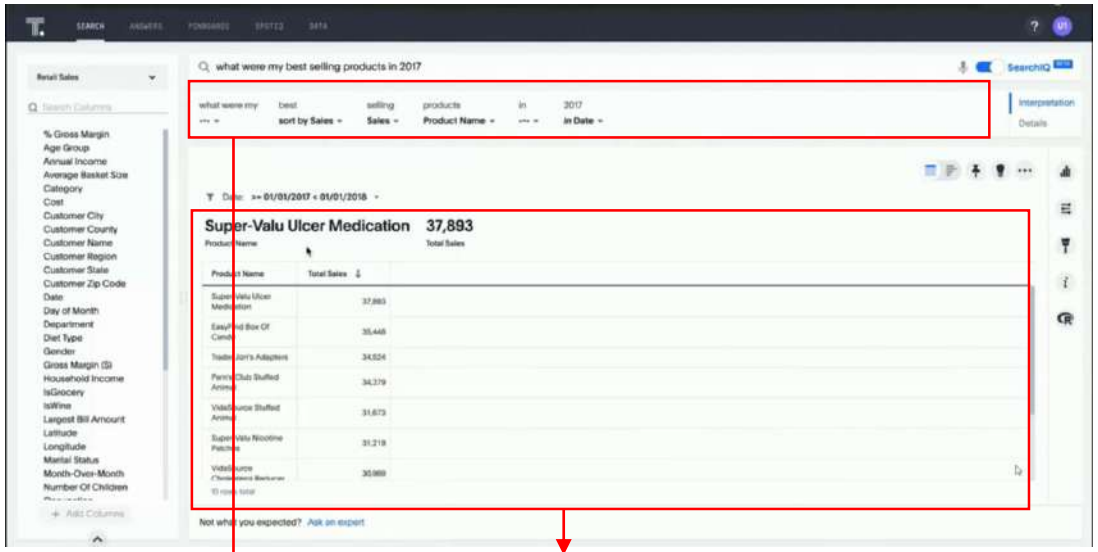
Techie Tuesdays

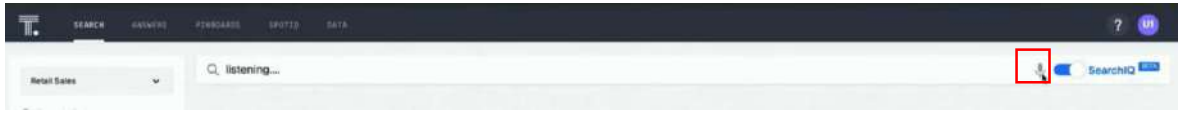
Search | With SearchIQ, you can search your data through natural language, just like speaking



1

Turn on Search IQ with this toggle





2

Click the microphone and ask your question. SearchIQ will translate your plain English into a search.

Note: You can type in a search using plain, speech-like language.

3

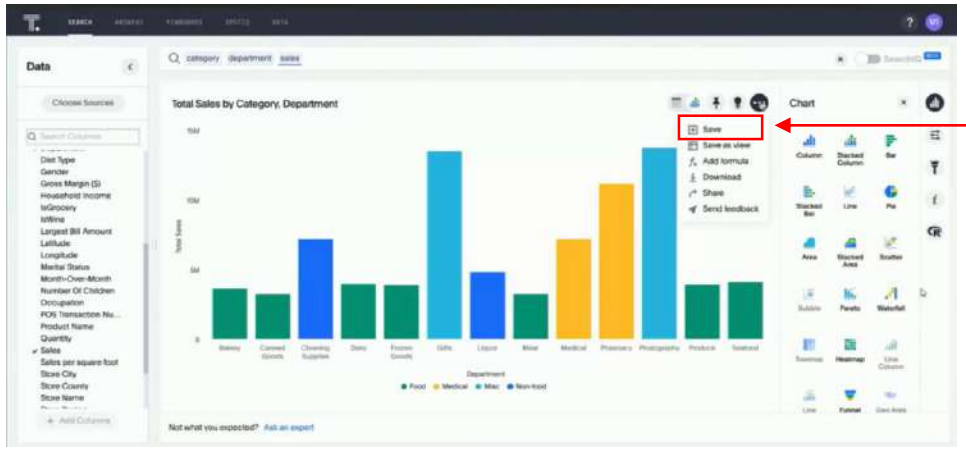
You can now see the results on your screen. This approach is similar to what we learned in the previous session. The only difference is, earlier we used keywords to create visualization and here we simply asked the question and get the relevant output.

This is useful for those who don't know much about data but want answers to their questions quickly

Unlike the previous search here you can see an additional section on the top. This section explains how ThoughtSpot perform search on your question. It simply ignores some generic words such as 'what were my' and 'in' and took best as '**sort by sales**', selling as '**sales**', products as '**product name**', and 2017 as '**in Date**'

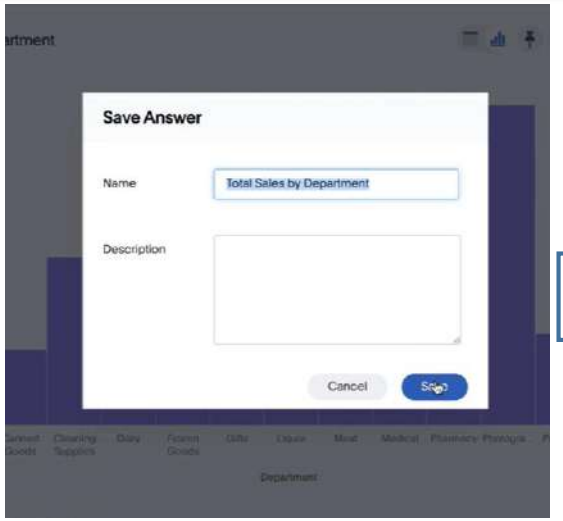
Techie Tuesdays

Answer



1

Once you complete your search, Save your Answer by selecting Save from the dropdown list of the ellipsis icon.

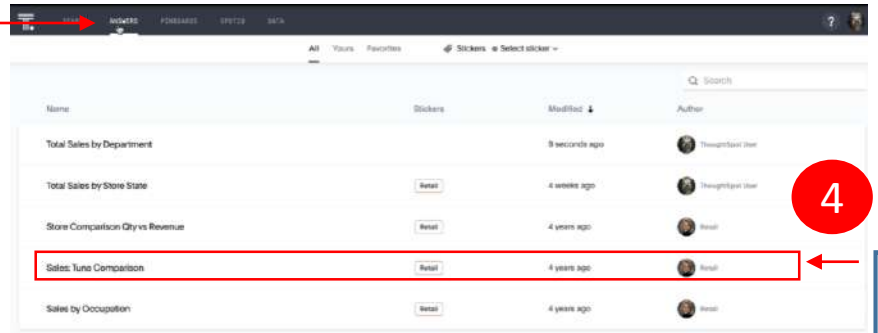


2

Name your answer and click save

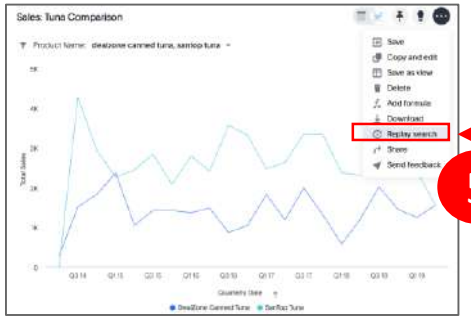
3

You can find the Answers that have been shared with you, and Answers that you have saved, in the Answers tab.



4

Click any of the previously saved answers



5

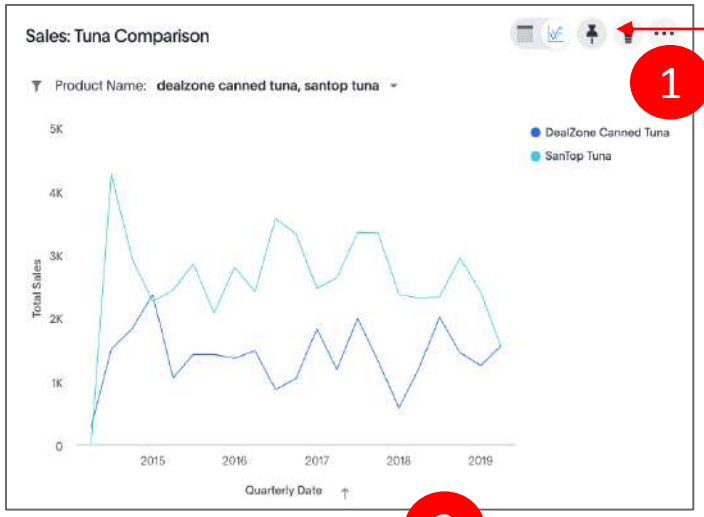
From all the available options, click replay search

The replay feature shows the step-by-step process of how to create the chart or table you are viewing. You will be able to see the steps used by another user to create a particular view or table.

nice Software Solutions
Data Analytics Services
Go Beyond Insights

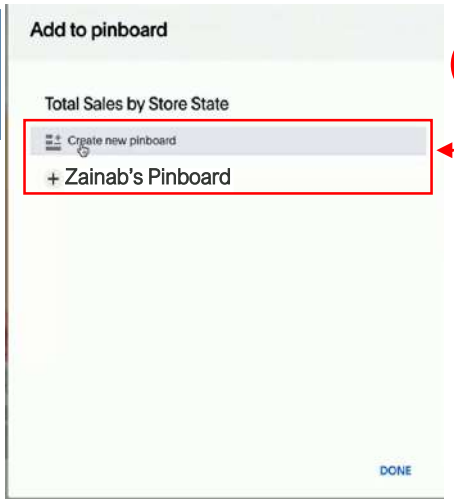
Techie Tuesdays

Pinboards



1

While viewing an answer, click the Pin icon at the top right of the answer

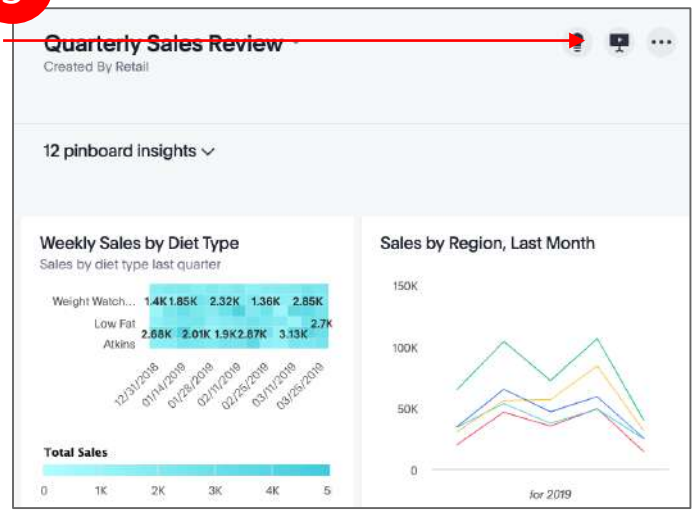


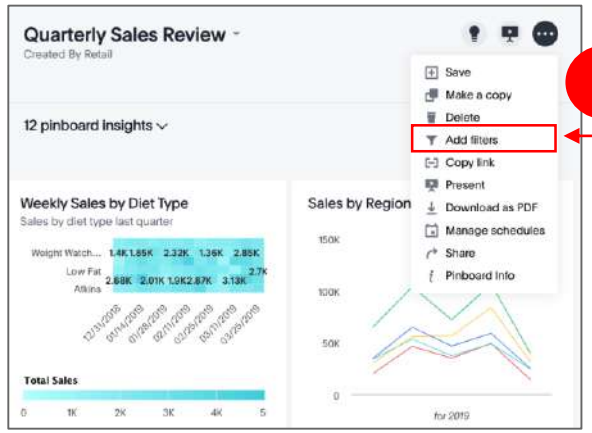
2

Select an existing pinboard, or Create a New Pinboard.

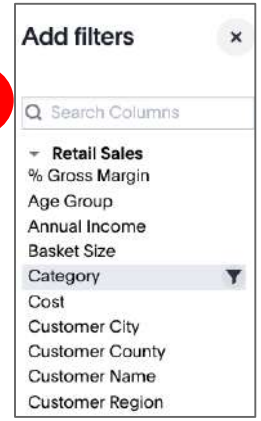
3

Once your Pinboard is saved. You can present your Pinboard as a slideshow, and turn your Pinboard into an interactive presentation by clicking this button.





4



Filters, including bulk filters and exclude filters, can be applied to Pinboards just as with tables and charts.

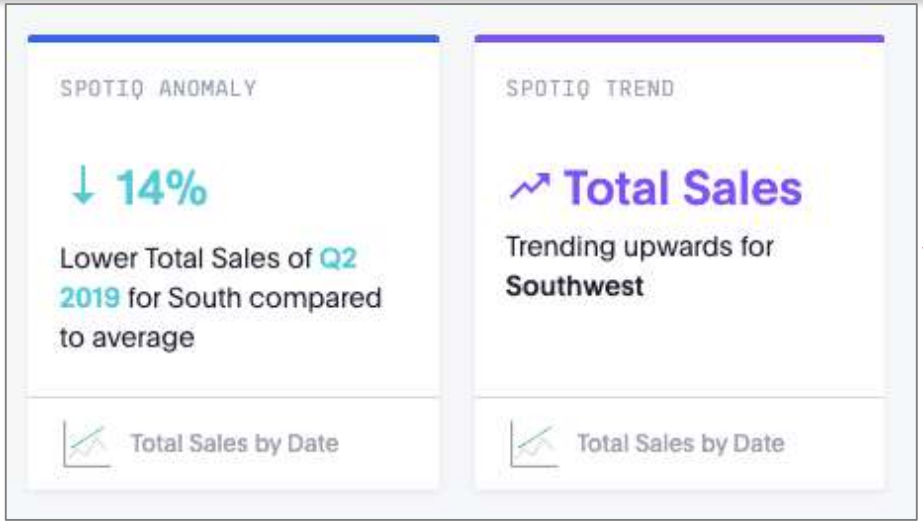
These kinds of filters apply to an entire pinboard, making it easy to see only the data that you are interested in across the tables and charts within a pinboard.

Pinboard filters can be very useful when you want to apply the same filters to more than one related visualization.

You can narrow the focus of your pinboard for specific purposes or audiences.

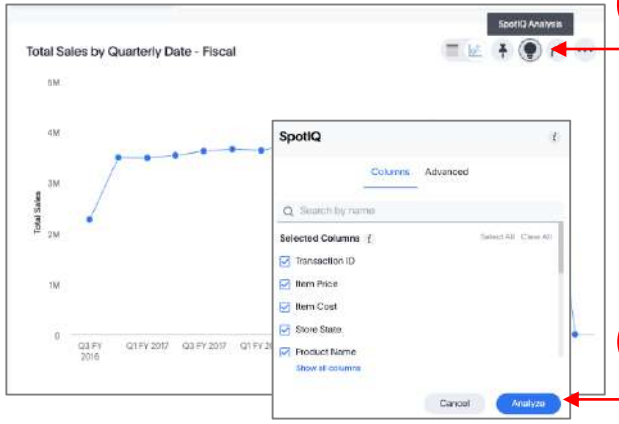
Techie Tuesdays

SpotIQ



1

You can find SpotIQ insights located just below answers. SpotIQ's automatically generated insights can answer interesting answers in your data that you may not have found on your own.



2

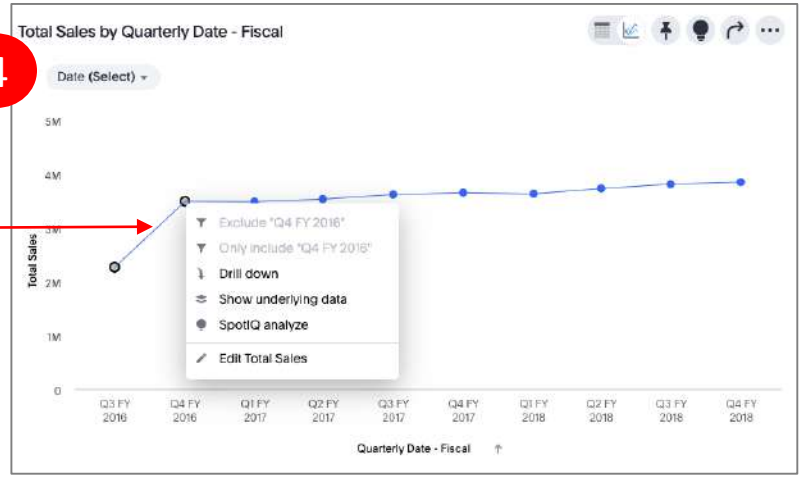
Click the SpotIQ icon above a chart to launch automatic analysis for a search. Customize analysis by explicitly indicating which columns to include or exclude

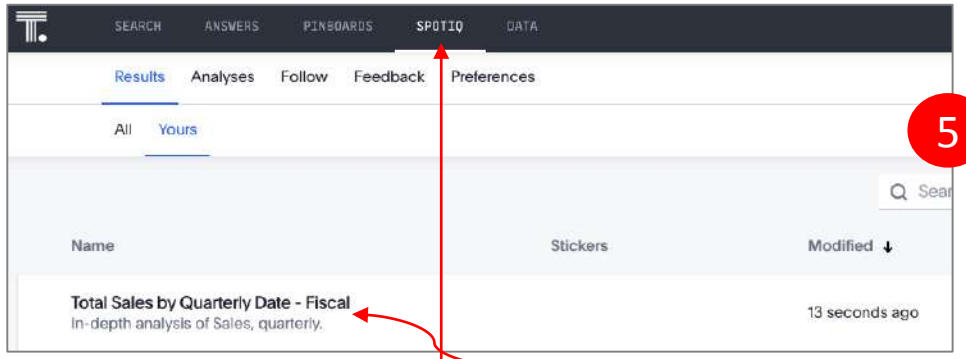
3

The Advanced tab allows you to select additional parameters for your SpotIQ analysis.

4

You can even compare two data points by running SpotIQ on only those data points that are selected. Highlight the data points of interest, then right click to select SpotIQ.





5

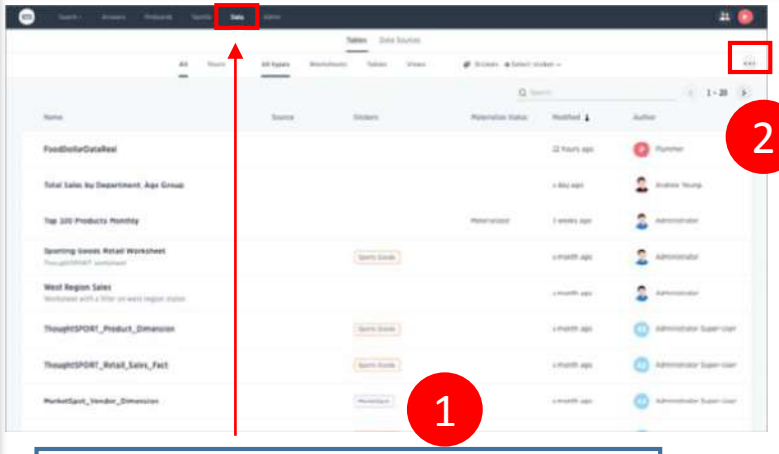
Click SpotIQ from the top menu bar to view the results of your analysis and analysis that is in progress

Once SpotIQ analysis is initiated, the results are saved in the SpotIQ page.

Please note that ThoughtSpot saves the results of SpotIQ analysis for 24 hours.

Techie Tuesdays

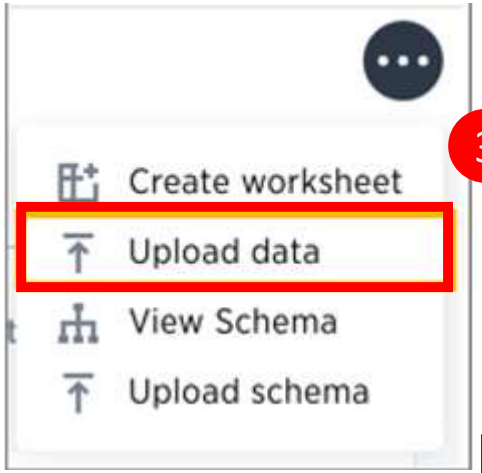
Data



1

Click Data. The Data tab lists all of the tables and data sources available for you to search on. Not all of them may be searchable, if you do not have access to underlying data in all tables.

2 Click ellipsis



3 Any user who belongs to a group that has administration privileges or Can upload user data, can upload their 'own data' from the browser

There are three types of data sources that you may see in the data list. They are tables, worksheets, and user uploaded data. You usually see only worksheets and user-uploaded sources. When to click to 'Choose Data'. These are the most commonly used data sources for searching. You can see 'View' in data sources only if you save your search as 'materialized view'

4

Name	Definition	Created By	Used By
Table	Raw tables loaded by an Administrator	Administrators	Administrators
Worksheet	Collection of related tables, optimized for searching	Anyone	Anyone it's shared with
User Imported	Table uploaded by a user through the Web browser	Anyone	Anyone it's shared with
View	A prior search saved as a data source	Anyone	Anyone it's shared with

Table or Raw table is where the admin originally build the ts load. User don't usually search on raw data.

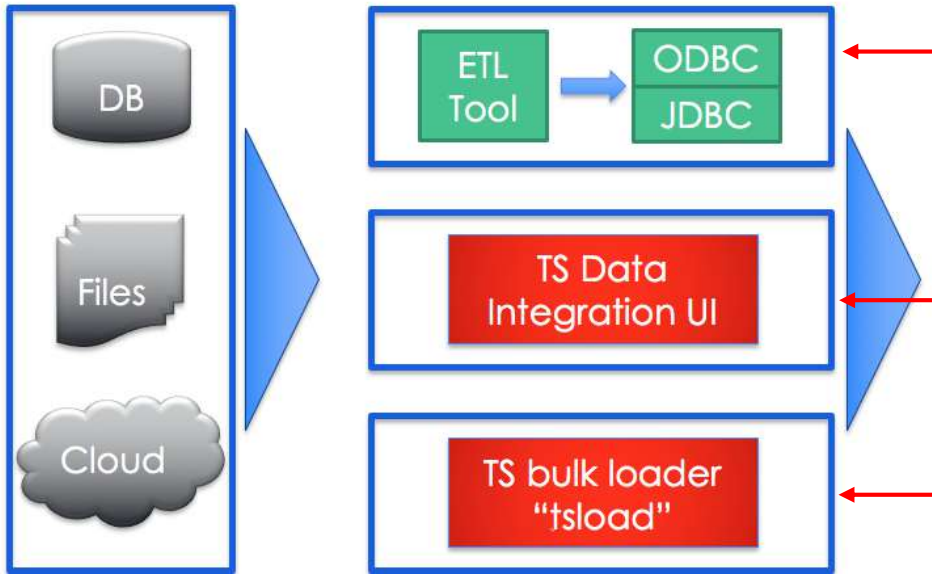
Worksheet is; When you take the raw tables and join them together to indicate which columns you want to use from which table. In worksheet the columns are added that are useful for users.

View is: Table saved as a materialized view by a user through the web browser through Save as view option on a search. Usually Save a search as a view is used so user can do searches on top of that view.

Techie Tuesdays

Data Loading (Caching)

There are several ways to load data into ThoughtSpot, depending on your goals and where the data is located. Consider your requirements for recurring loads when planning how best to bring your data into ThoughtSpot.



1. You can make use of an established extract, transform, load (ETL) process and tool(s), and connect to ThoughtSpot using third party tools.

2. Another fast and easy way to load a new table is by importing it using the web browser.

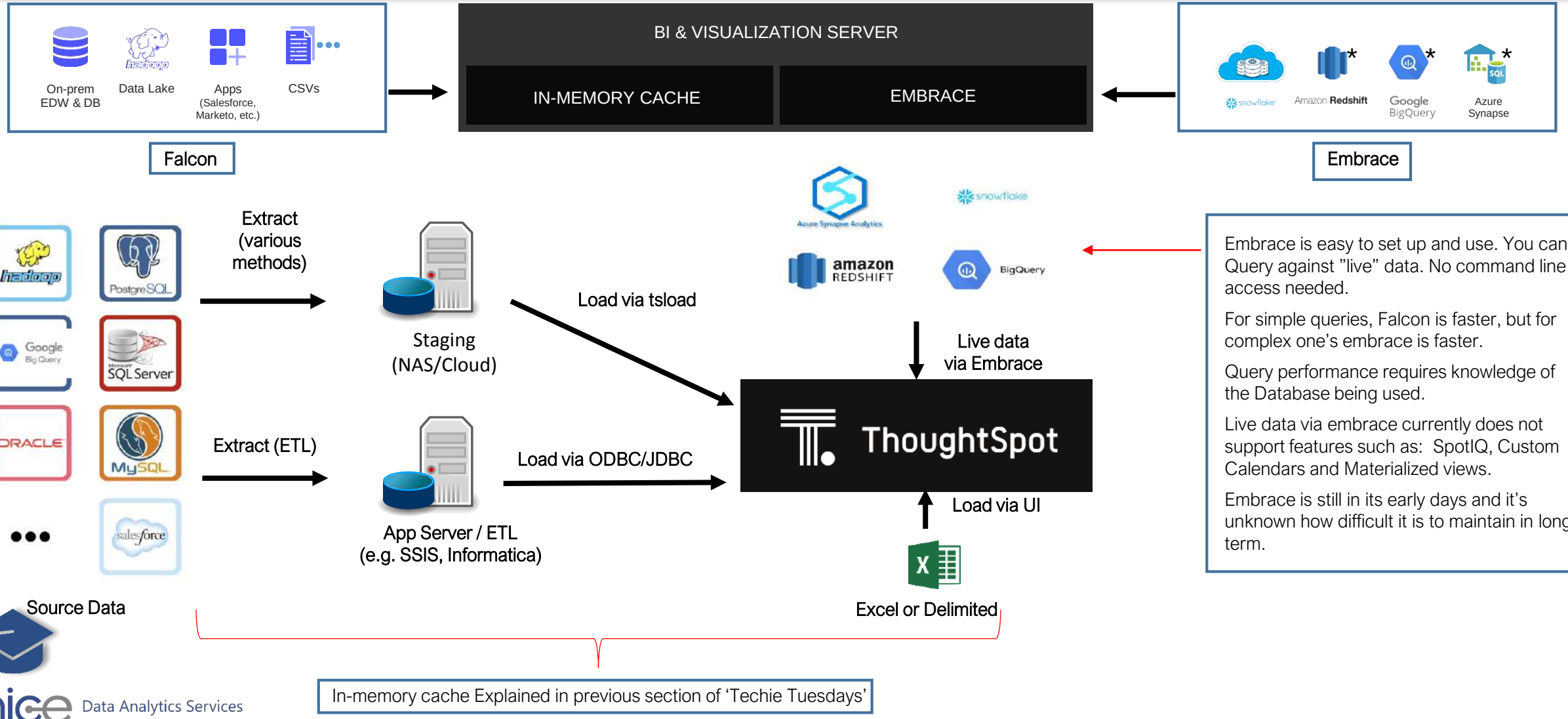
   This is best for one-time data loads of small tables which do not have complex relationships to other tables. This method is limited to tables that are under 50 MB (megabytes) in size.

3. You can also import data with the ThoughtSpot bulk loader, tsload. With this method, you would use TQL and tsload to load data directly into the back-end database that ThoughtSpot uses.

   This is the best way to load large amounts of data or a schema with multiple tables. Tsload can be scripted and used for recurring data loads. It can be integrated with an ETL tool for automation. This is the fastest method of loading data, and is approximately 10x faster than using ODBC.

Techie Tuesdays

Data Loading (Embrace)



Embrace is easy to set up and use. You can Query against "live" data. No command line access needed.

For simple queries, Falcon is faster, but for complex one's embrace is faster.

Query performance requires knowledge of the Database being used.

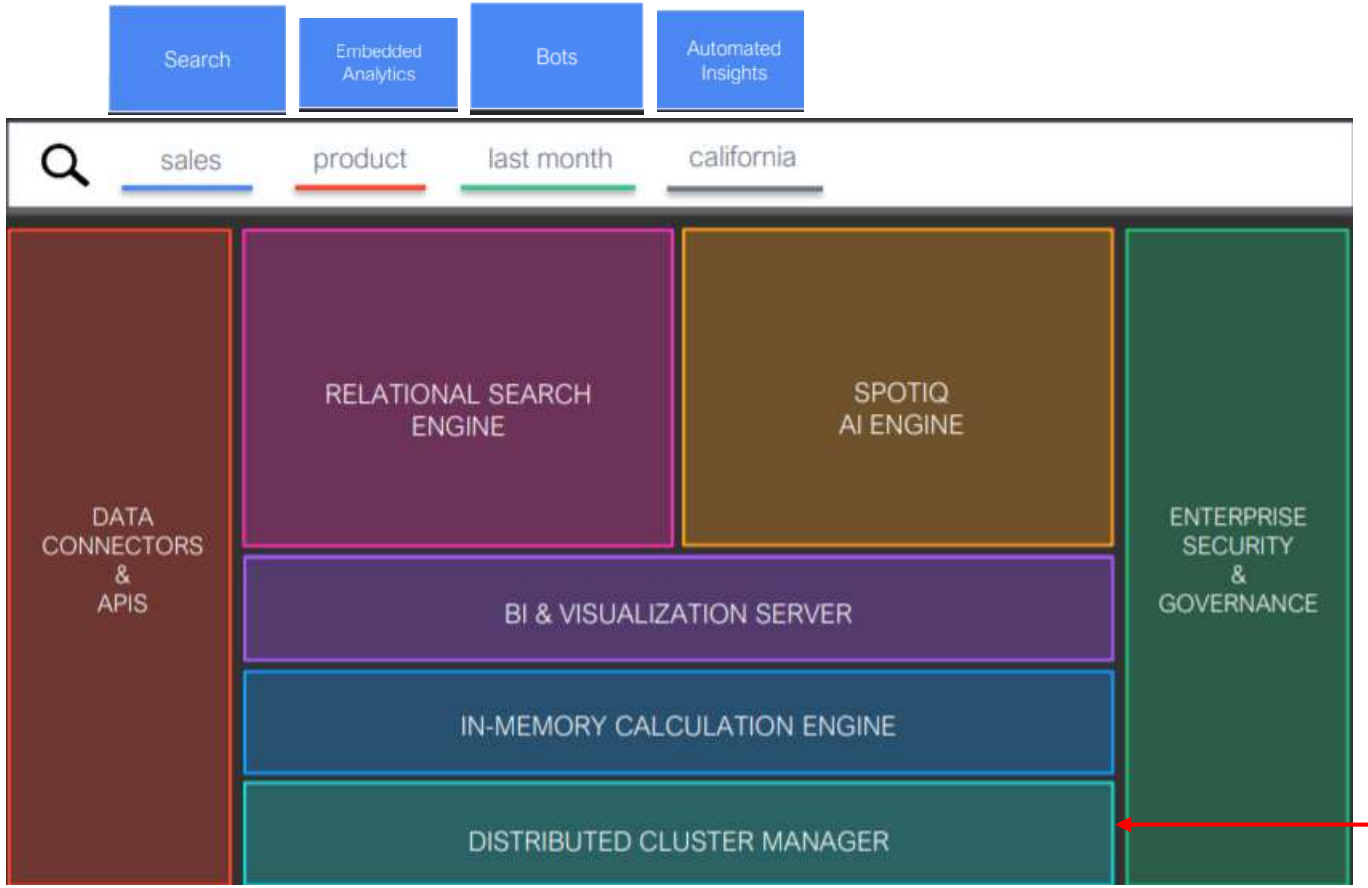
Live data via embrace currently does not support features such as: SpotIQ, Custom Calendars and Materialized views.

Embrace is still in its early days and it's unknown how difficult it is to maintain in long term.

In-memory cache Explained in previous section of 'Techie Tuesdays'

Techie Tuesdays

ThoughtSpot Architecture

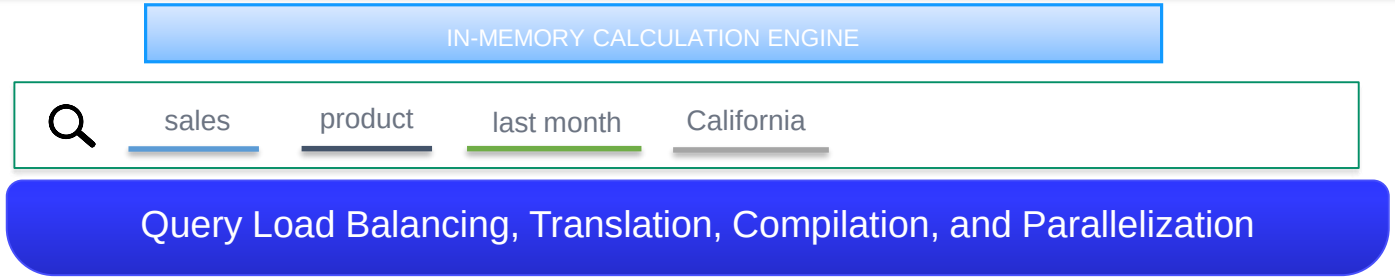


We will discuss these topics one by one in the coming sections

Discussed in '**Techie Tuesdays**' 15 & 16

Techie Tuesdays

ThoughtSpot's In-Memory Calculation Engine | In-memory, parallel data processing at scale

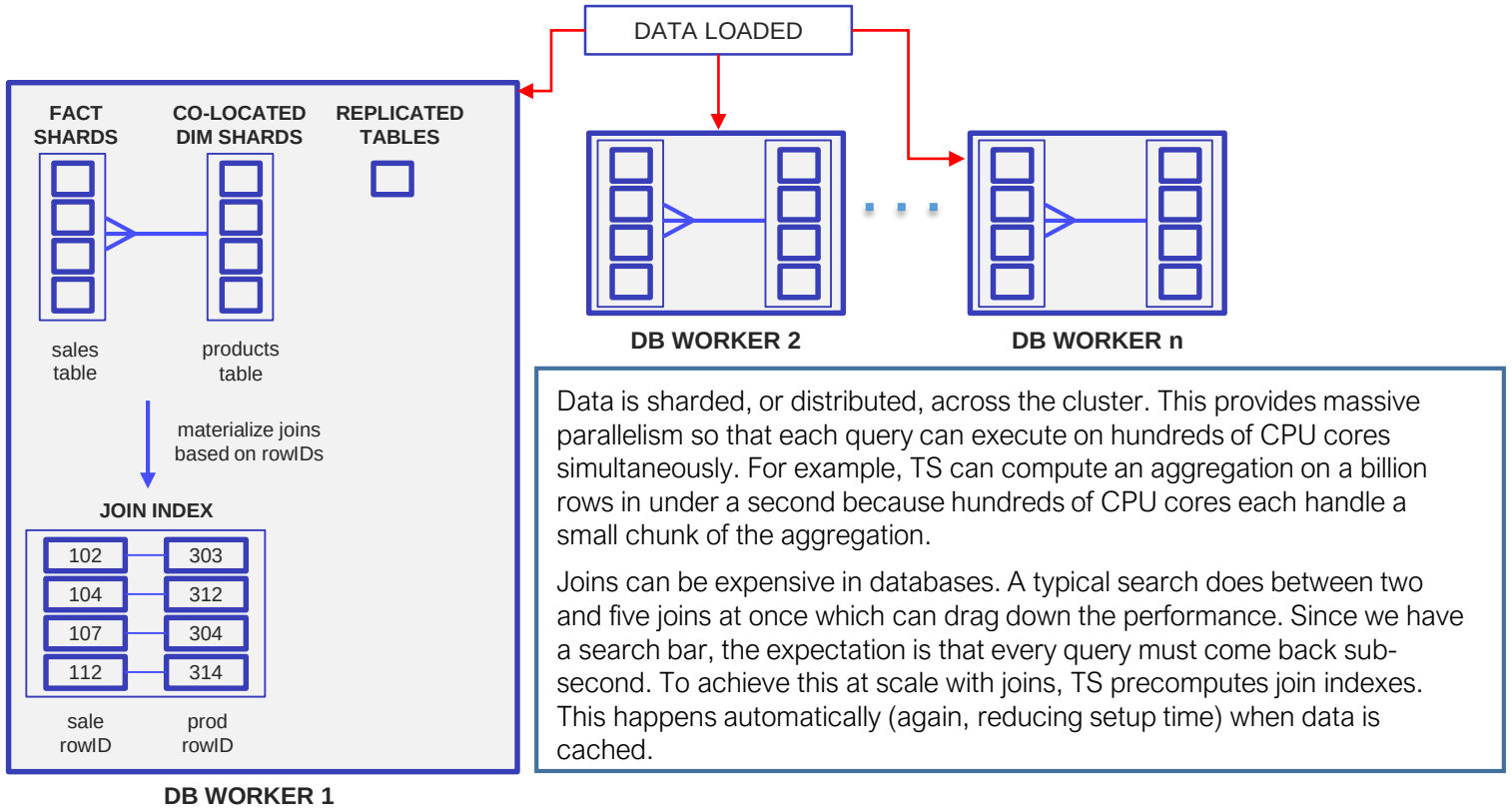


ThoughtSpot has an in memory, columnar, distributed database that makes its database very fast.

In-memory database has advanced features like JIT and results cache that you can find in more mature databases. The DB is purpose built which allows you to do things like push growth queries natively into the database vs generating generic SQL which can be very slow.

Large scale joins are fast:

- Fact data is sharded
- Dimension data is also sharded, co-located with fact data
- Join index is precomputed automatically when data is loaded
- At query time, join is done in parallel



Data is sharded, or distributed, across the cluster. This provides massive parallelism so that each query can execute on hundreds of CPU cores simultaneously. For example, TS can compute an aggregation on a billion rows in under a second because hundreds of CPU cores each handle a small chunk of the aggregation.

Joins can be expensive in databases. A typical search does between two and five joins at once which can drag down the performance. Since we have a search bar, the expectation is that every query must come back sub-second. To achieve this at scale with joins, TS precomputes join indexes. This happens automatically (again, reducing setup time) when data is cached.

THANK YOU

Securing your future with quality data and analysing with precision.

+91-712-6644571

21, IT Park, Gayatri Nagar, behind VNIT, Nagpur – 440022

info@nicesoftwaresolutions.com

www.nicesoftwaresolutions.com