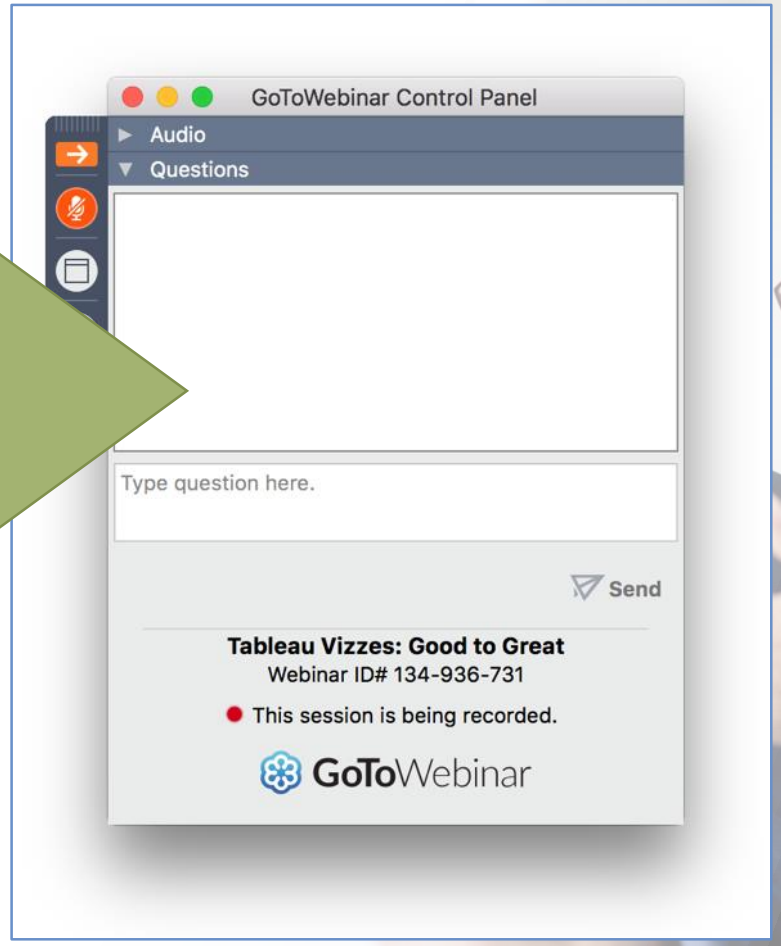
10 Reasons Snowflake Is Great for Analytics

GoToWebinar control panel

Click arrow to restore full control panel

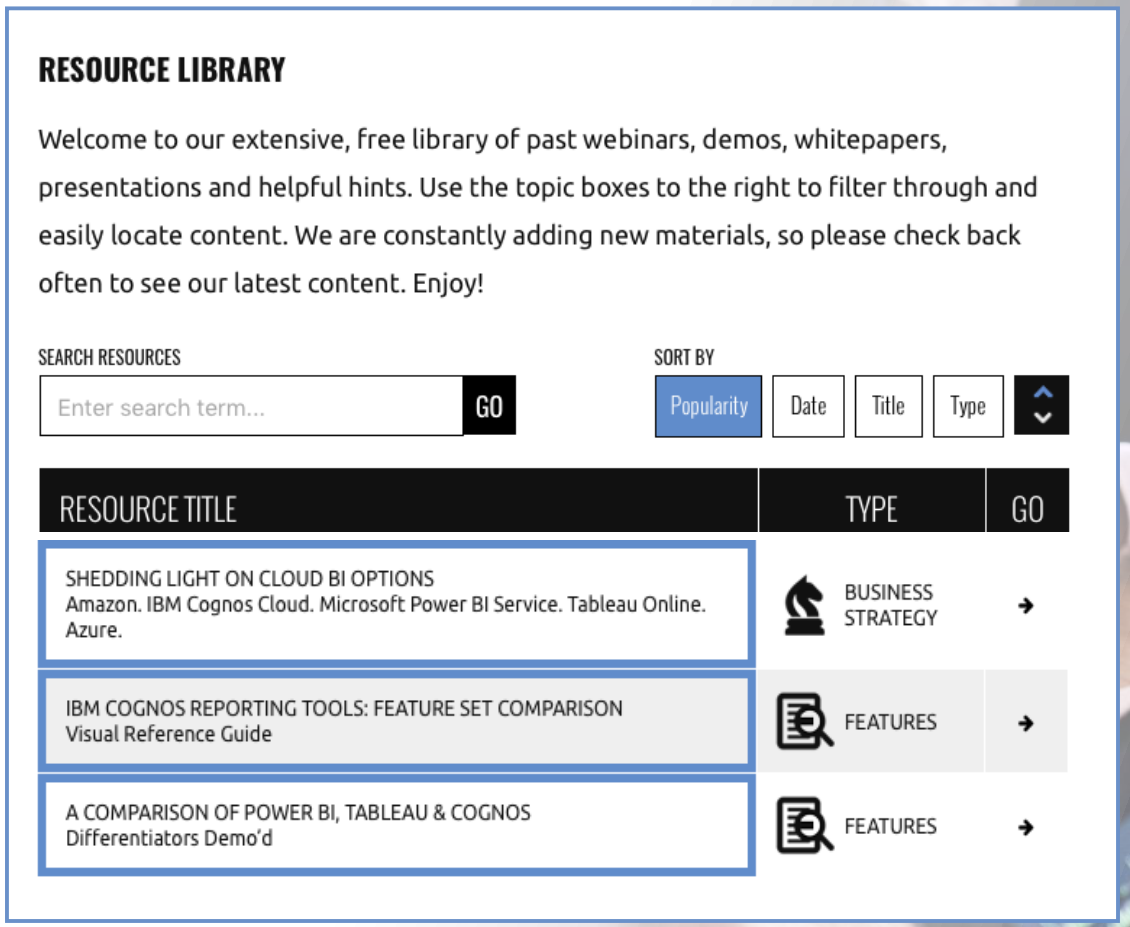
Submit questions in this section



To obtain this presentation

Visit the Resource Library
on the Senturus website
to download this presentation
and explore other assets

senturus.com/resources



Agenda

Introduction

Snowflake overview

10 reasons Snowflake is built for analytics

Senturus overview

Additional resources

Q&A

Introductions

Bob Looney

VP Software & Architecture

Senturus, Inc.

Reeves Smith

Principal Snowflake Architect

Senturus, Inc.

Michael Weinhauer

Director

Senturus, Inc.

Poll

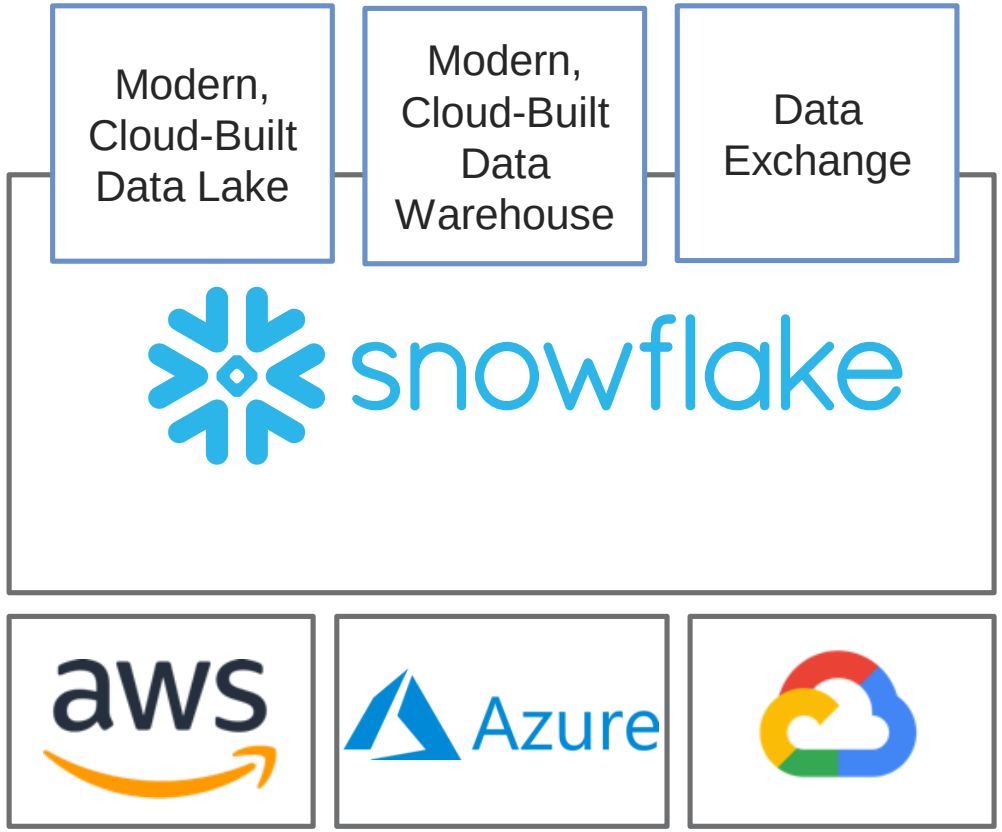
Why did this webinar pique your interest?

- I like to keep up-to-date on new technologies
- Our organization is evaluating data platforms
- Our organization is evaluating Snowflake
- Our organization recently purchased Snowflake
- Other

Snowflake overview

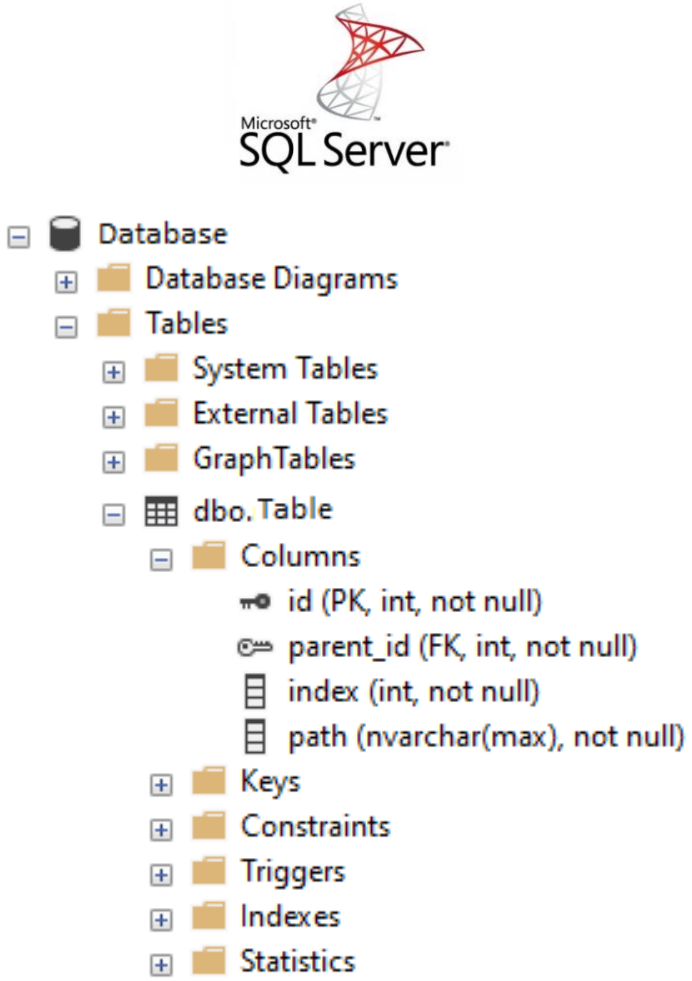
Snowflake overview

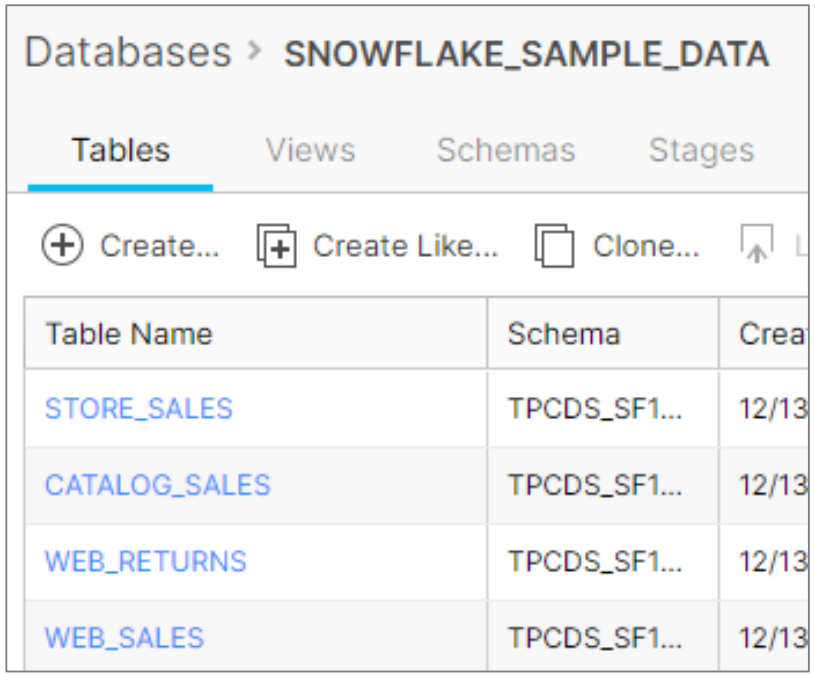
- SaaS, not PaaS
  - Low administration
  - Start a trial online for free
- Consumption based pricing
  - Compute & storage
- Cloud agnostic
  - Abstracted on top of AWS, Azure or GCP
- Analytics focus, not transactional
  - OLAP, not OLTP



Consistent database concepts

- Database
- Schema
- Table
- Column
- View
- User
- Role
- ANSI SQL compliant





Unique concepts

Snowflake (company and product)

- Does **not** refer to a preference for the snowflake data model
- *“Our founders just really love skiing and Snowflakes are made in the cloud.”*
- *Also a reference to each client being unique and the flexibility of the platform to fit each use case*

Cloud data platform

- SaaS product that Snowflake sells consisting of storage and compute resources

Unique concepts

Procedures

- A mix of JavaScript (loops, logic) and SQL (data access)
- Unlike database procedures, don't return queries of data


```
CREATE PROCEDURE stproc1 (@FLOAT_PARAM1 float)
AS
BEGIN

    BEGIN TRY
        INSERT INTO stproc_test_table1 (num_col1) VALUES (@FLOAT_PARAM1);

        SELECT 'Succeeded.';

    END TRY
    BEGIN CATCH

        SELECT 'Failed:' + ERROR_MESSAGE();

    END CATCH

END
```


```
create or replace procedure stproc1(FLOAT_PARAM1 FLOAT)
    returns string
    language javascript
    strict
    execute as owner
    as
    $$
    var sql_command =
     "INSERT INTO stproc_test_table1 (num_col1) VALUES (" + FLOAT_PARAM1 + ")";
    try {
        snowflake.execute (
            {sqlText: sql_command}
            );
        return "Succeeded.";   // Return a success/error indicator.
        }
    catch (err)  {
        return "Failed: " + err;   // Return a success/error indicator.
        }
    $$
    ;
```

Unique concepts

True decoupled compute & storage

- Multiple, independent compute resources access the same database
- *"virtually unlimited number of concurrent workloads against the same, single copy of your data"*

Warehouse

- Sized **compute** capacity
- Acts on a database
- Start/pause/stop, scale up & down

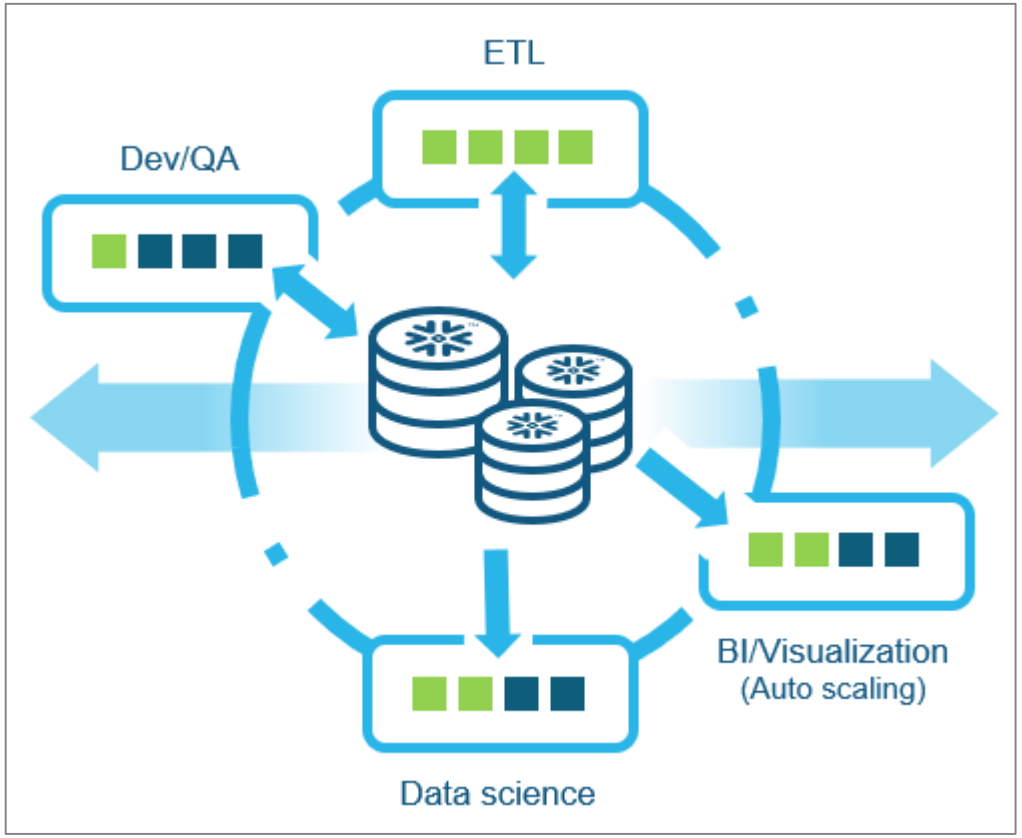
10 reasons Snowflake is built for Analytics

10 reasons Snowflake is built for analytics

1. Large data volumes
2. Data loading flexibility
3. Broad BI tools support
4. Supports “Analysis Ready” data models
5. Minimized administration
6. Performance scalability
7. Semi-structured data
8. Cloning
9. Time Travel
10. Data sharing

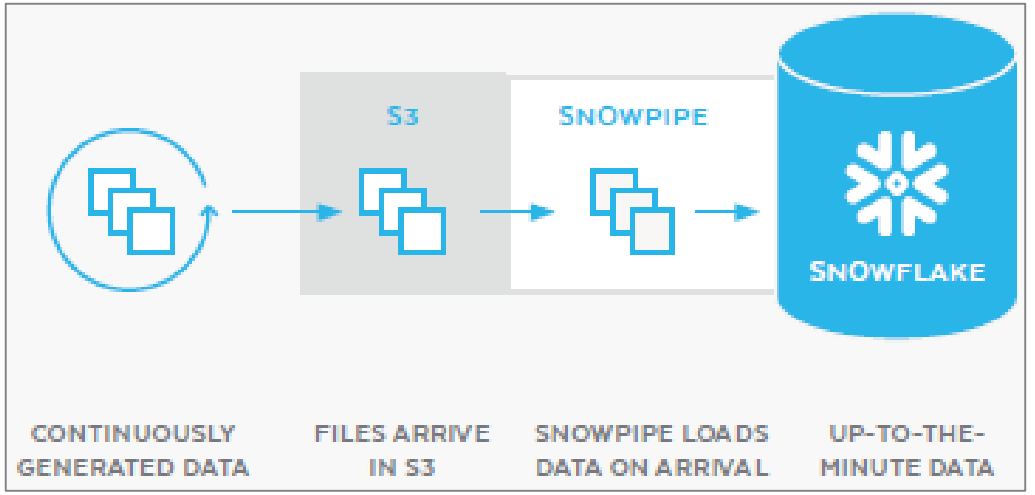
1) Large data volumes

- Highly reliable & scalable
  - Storage backed by cloud providers
- Flexible staging
  - **Internal stage:** part of the Snowflake tenant
  - **External stage:** Amazon S3, Azure Blob, Google cloud storage
- Fast data loading of large data sets



2) Data loading flexibility

- File based data loading
  - Structured and unstructured
- Snowpipe event driven loading
- Tool support
  - Enables rapid lift and shift of on-prem SQL based EDW
  - Integrate web applications



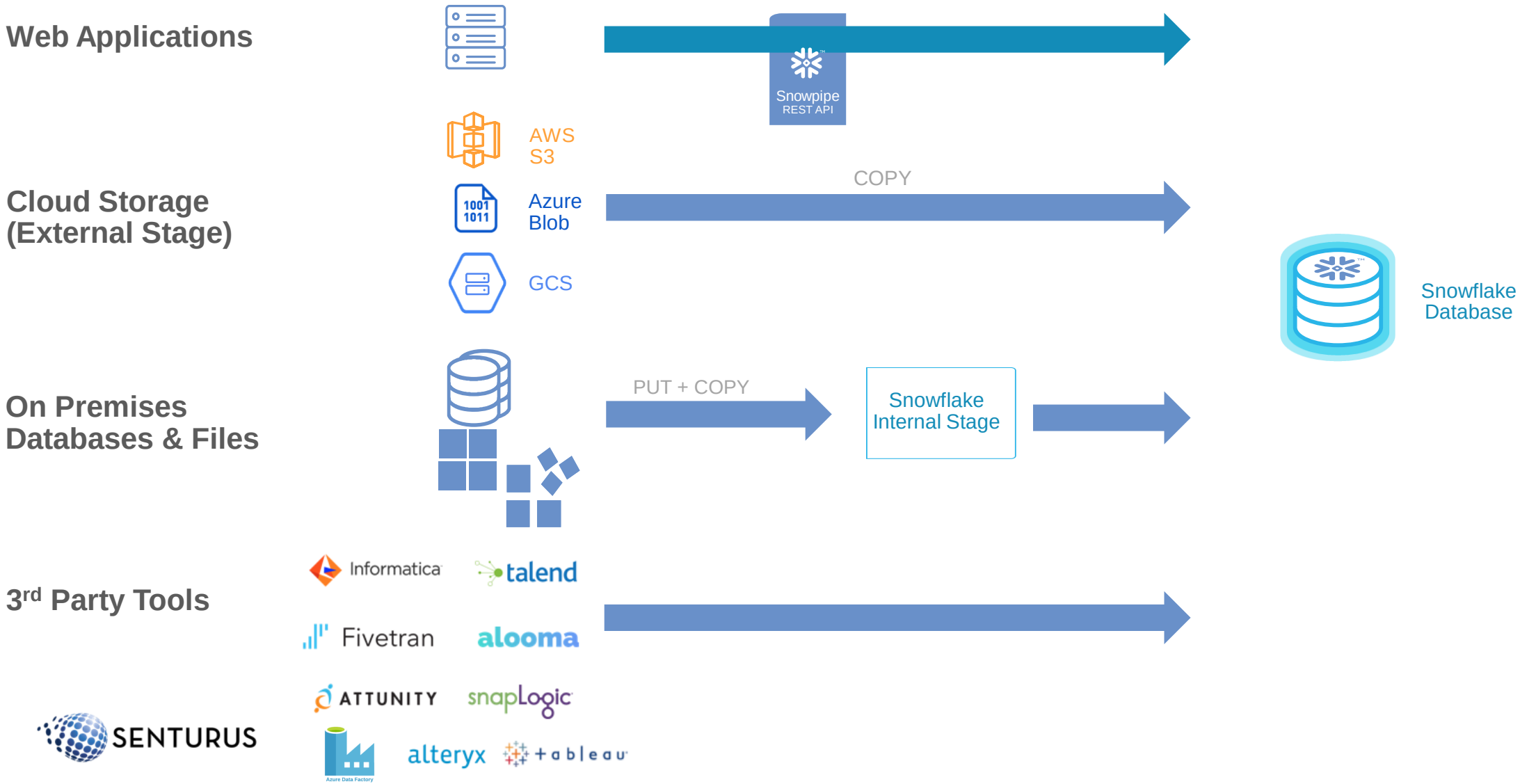
Fivetran Informatica

MATILLION segment

Azure Data Factory Stitch talend

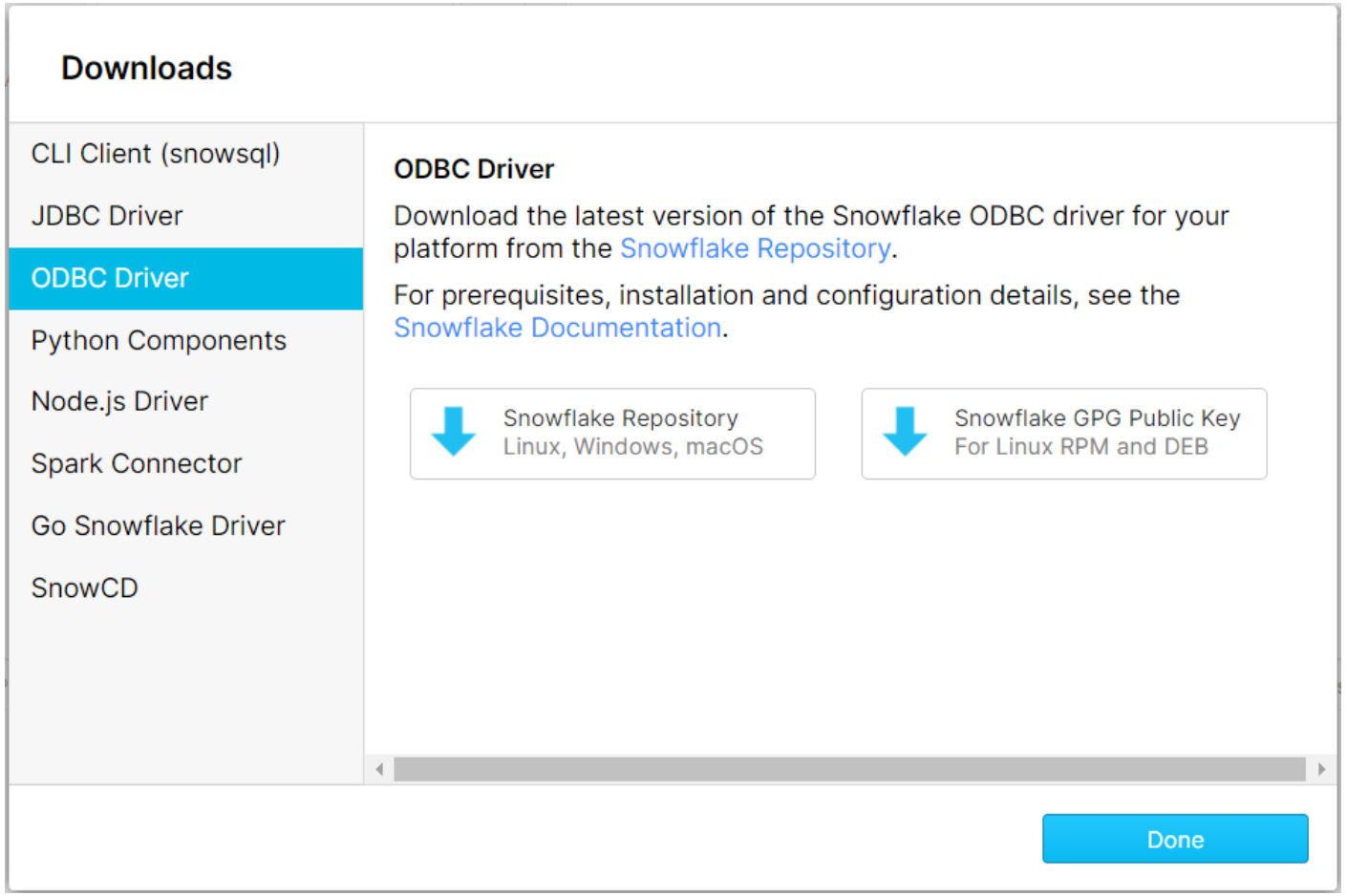
tableau prep alteryx

Data loading options

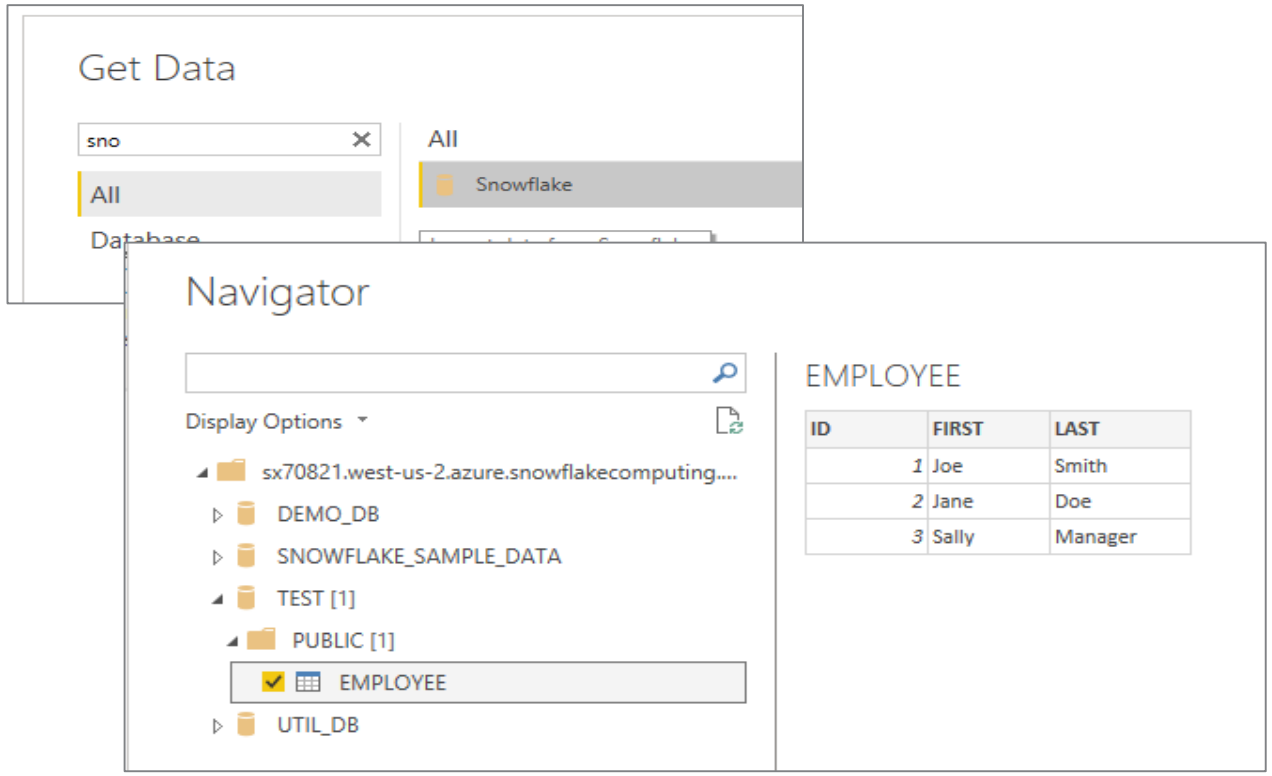


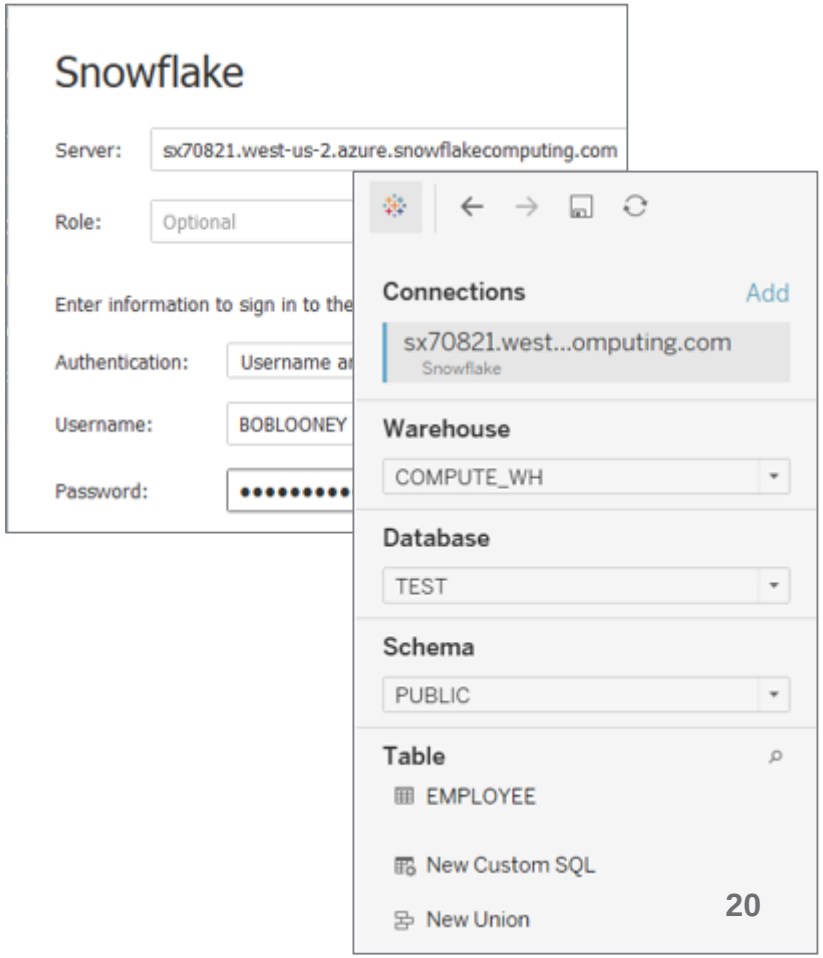
3) Broad BI tool support

- Chartio
- Cognos
- Domo
- Looker
- MicroStrategy
- Power BI
- Qlik
- QuickSight
- Sisense
- Tableau
- Tibco
- ThoughtSpot
- …



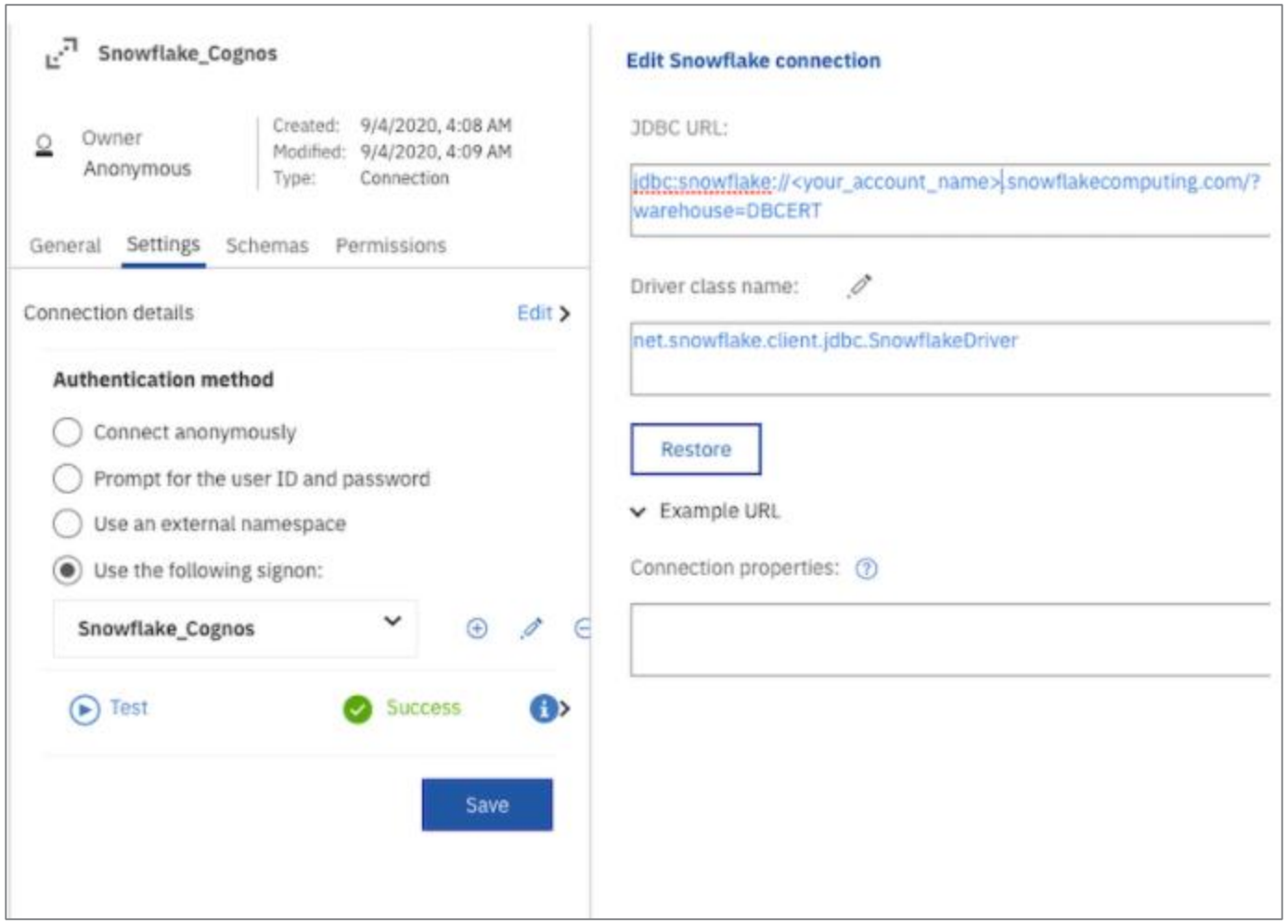
BI tool demos - Power BI & Tableau





SENTURUS

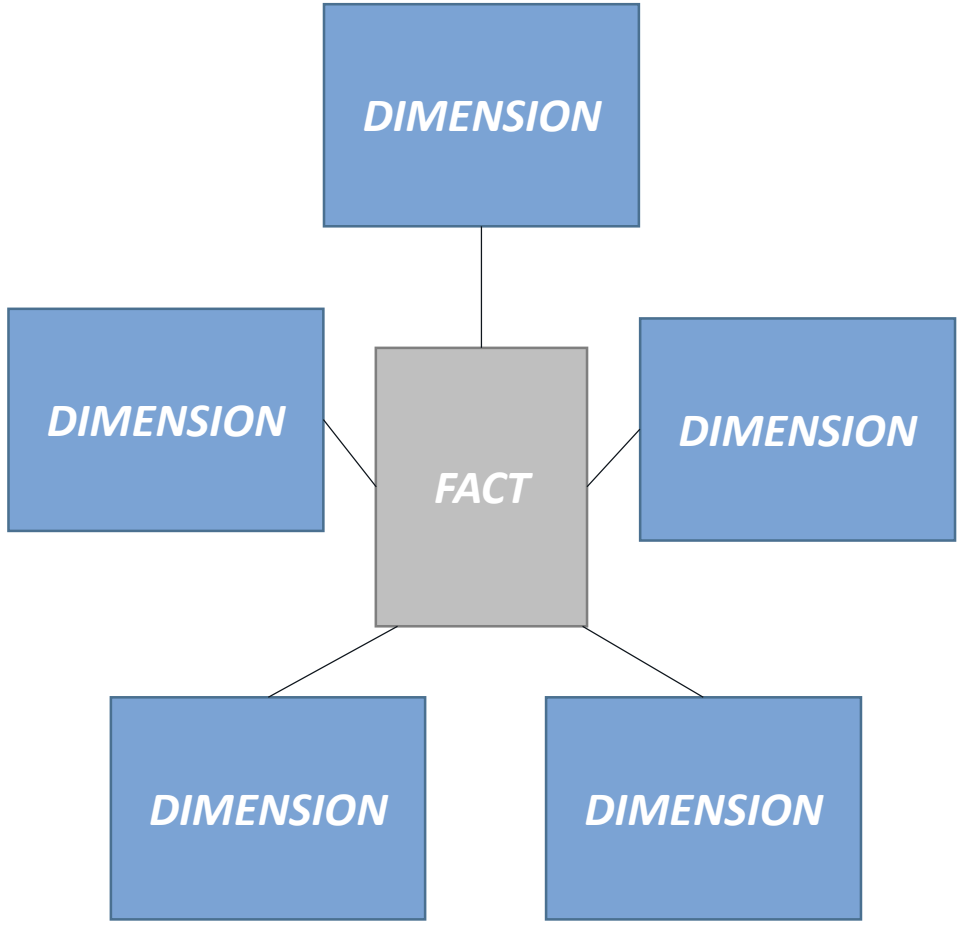
BI tool - Cognos



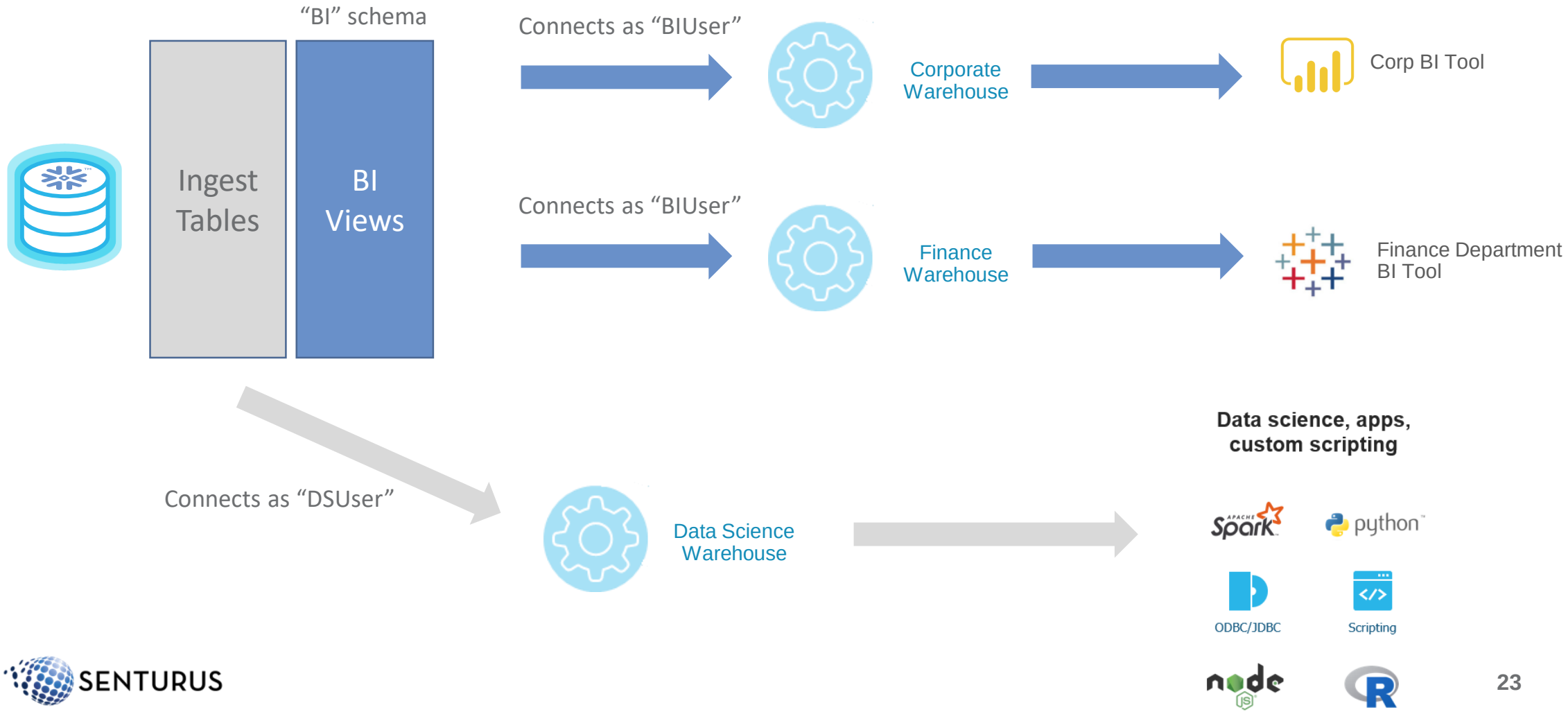
Ref: https://www.ibm.com/support/pages/how-set-snowflake-data-source-connection-cognos-analytics

4) Supports “Analysis Ready” data models

- Modern BI tools work best with star or Snowflake data models
- Create a schema with BI views on top of ingestion tables
- Create security so that BI users only see BI views



BI data modeling in Snowflake



5) Minimized administration

TRADITIONAL PLATFORMS

Infrastructure

- Initial Setup
- Upgrading
- Patching
- Capacity Planning
- Storage
- Security

Physical Design

- Partitioning
- Indexing
- Ordering
- Vacuuming

Data Collaboration

- Loading
- Moving
- Transforming
- Copying
- Securing

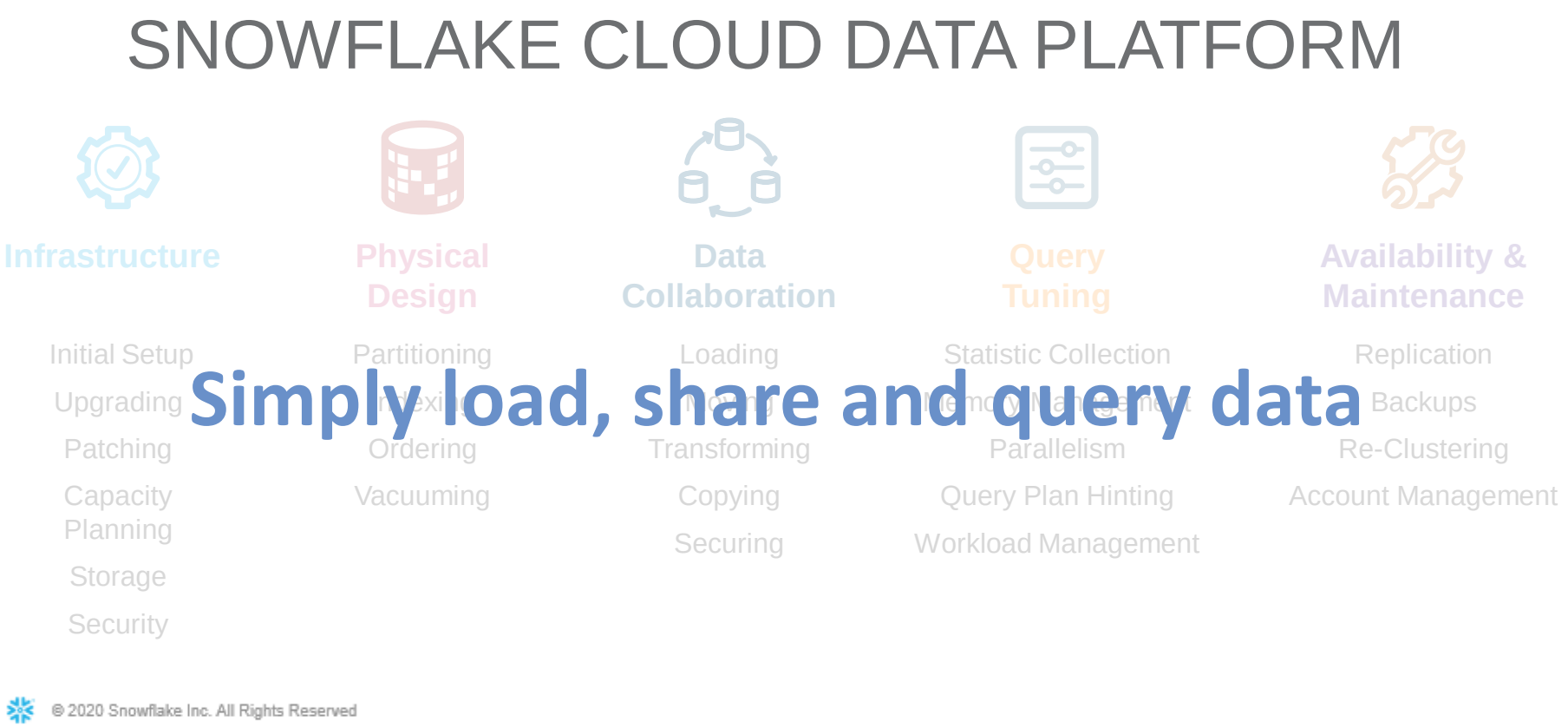
Query Tuning

- Statistic Collection
- Memory Management
- Parallelism
- Query Plan Hinting
- Workload Management

Availability & Maintenance

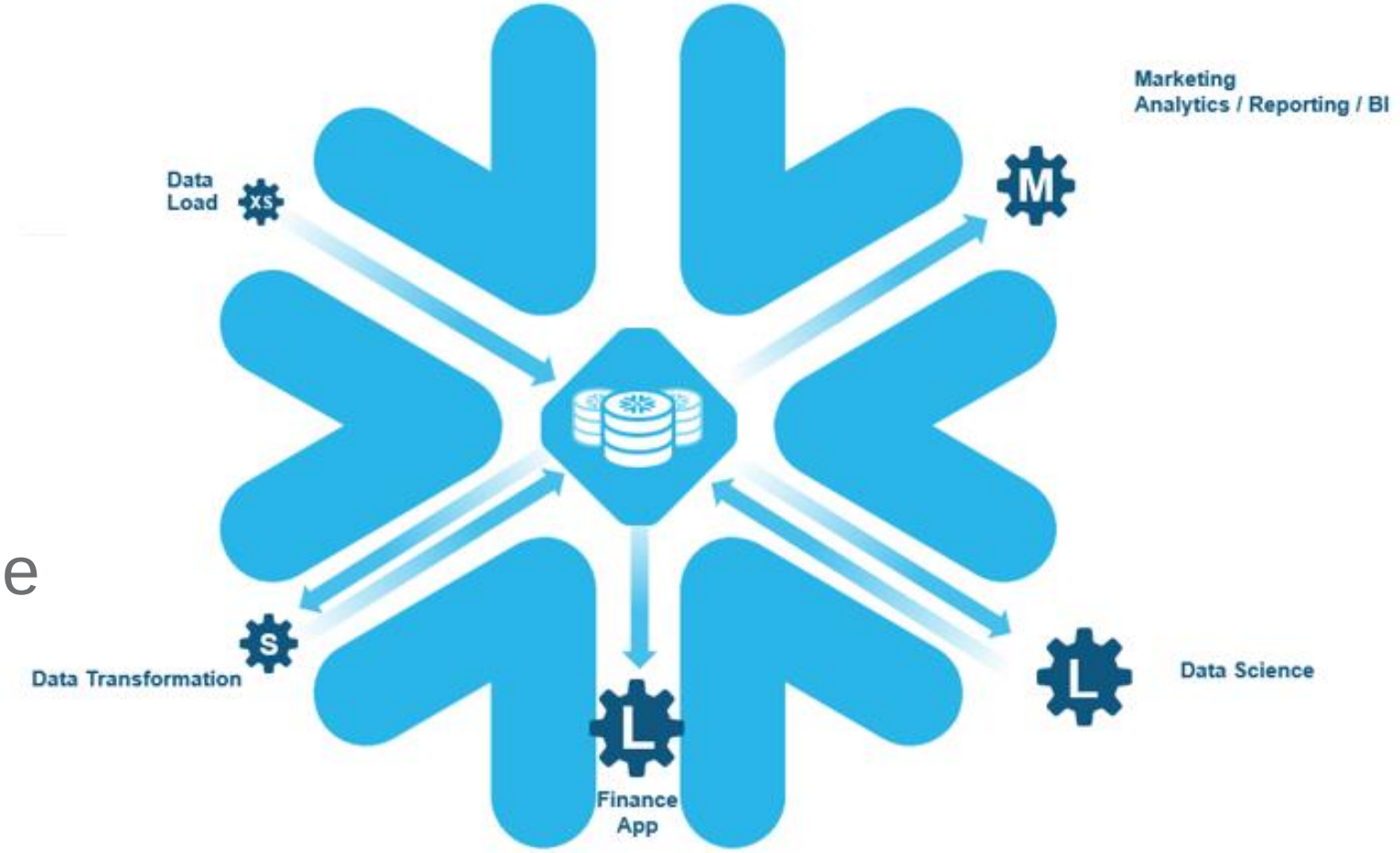
- Replication
- Backups
- Re-Clustering
- Account Management

Snowflake minimized administration



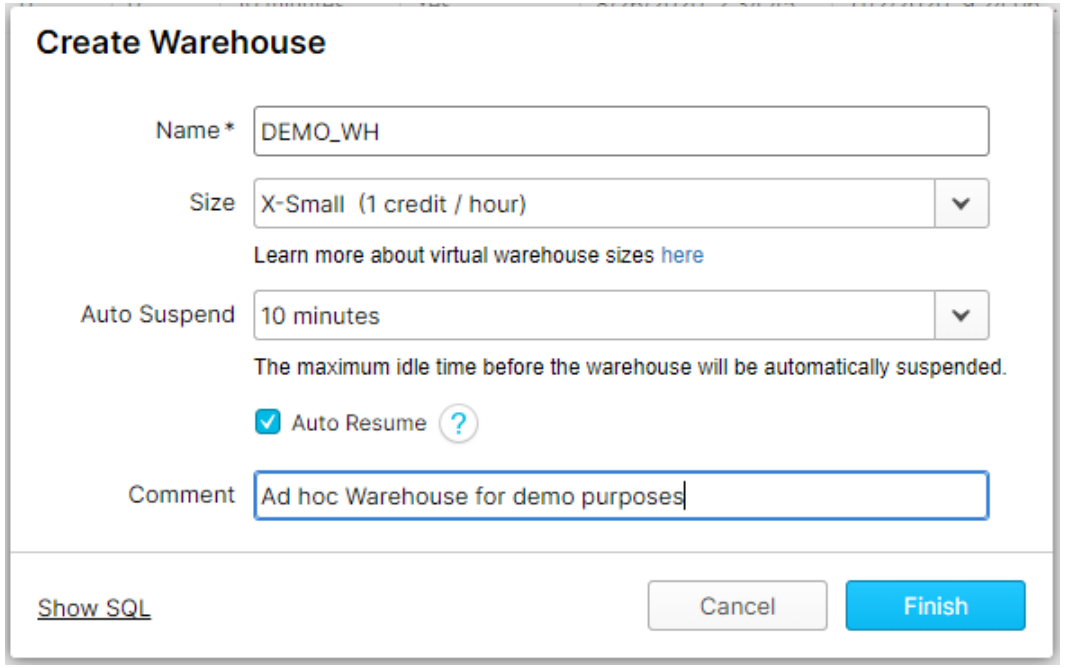
6) Performance scalability

- Workload isolation
  - Data loading
  - Data query
  - Data science
- Auto-scale out
- Scale up/down
- Ad hoc warehouse use cases
- Auto-pause

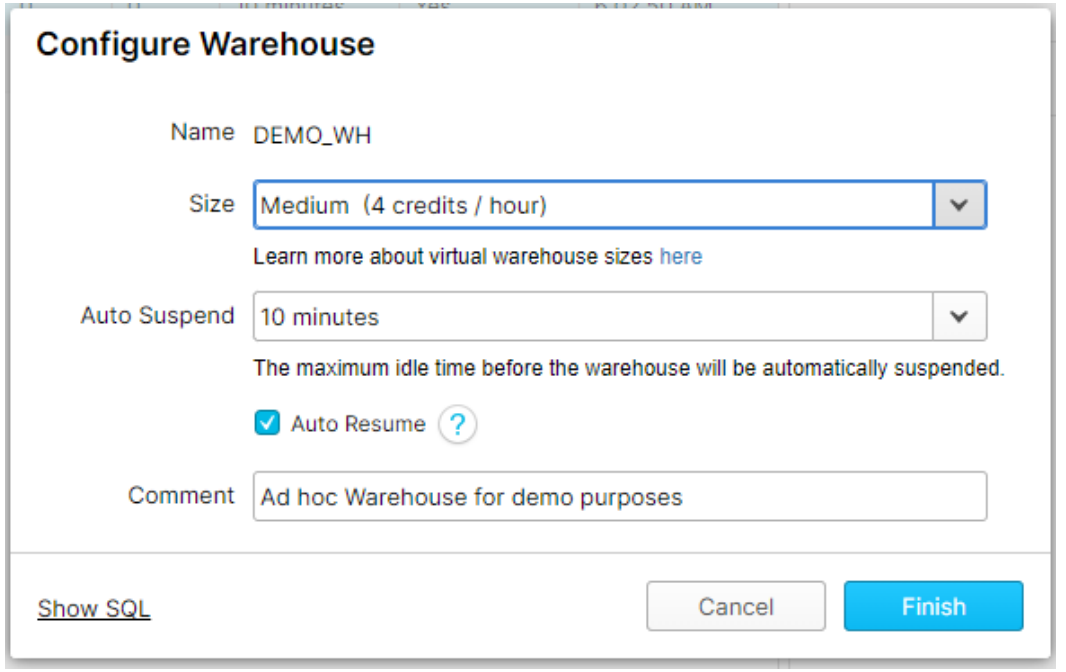


Performance scalability demo

Create warehouse

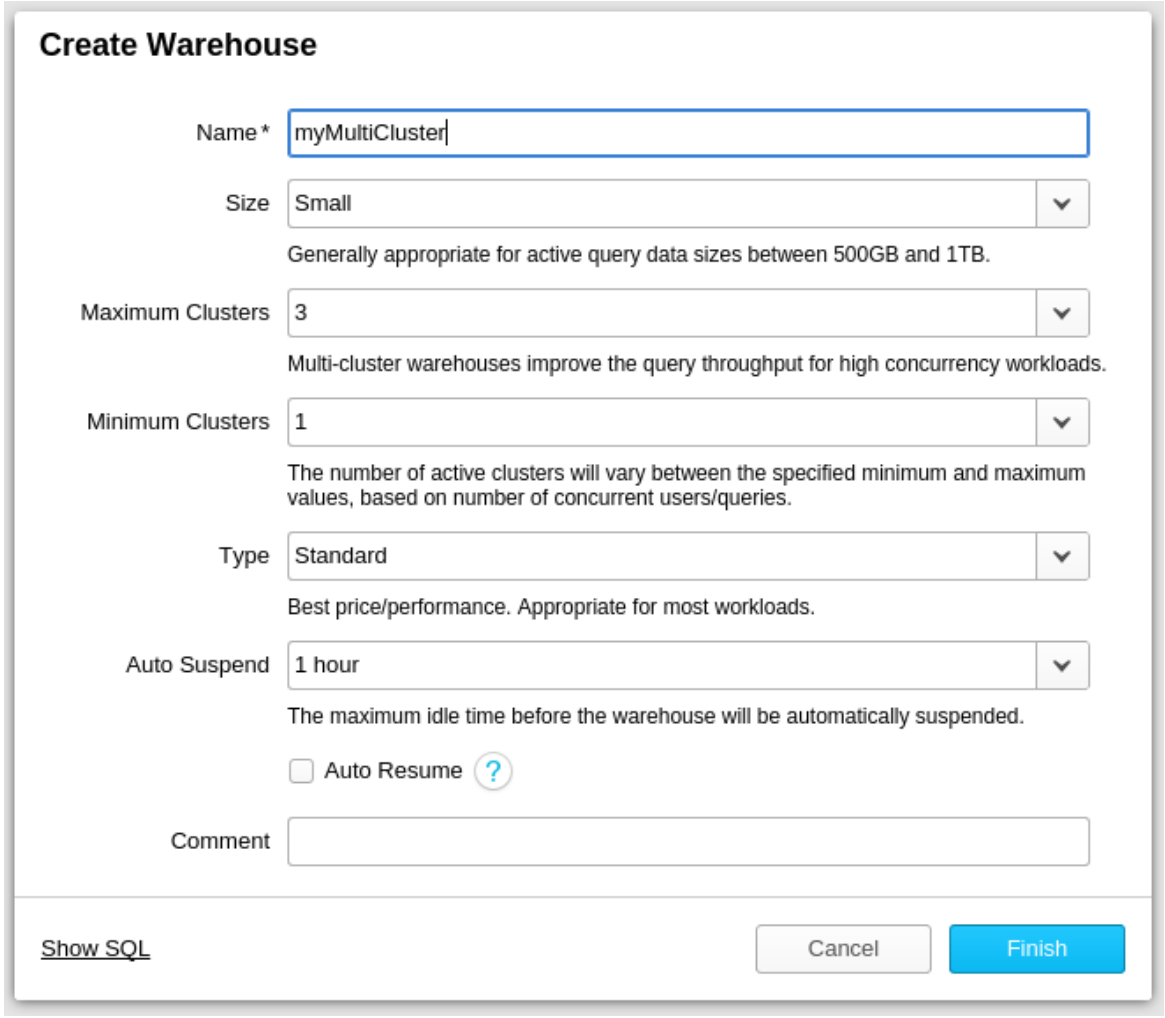


Scale up/down warehouse



Performance scalability demo

Multi-cluster warehouse – scale out



7) Semi-structured data

- Query directly from XML, JSON, and other semi-structured data
- Define a table with a VARIANT column

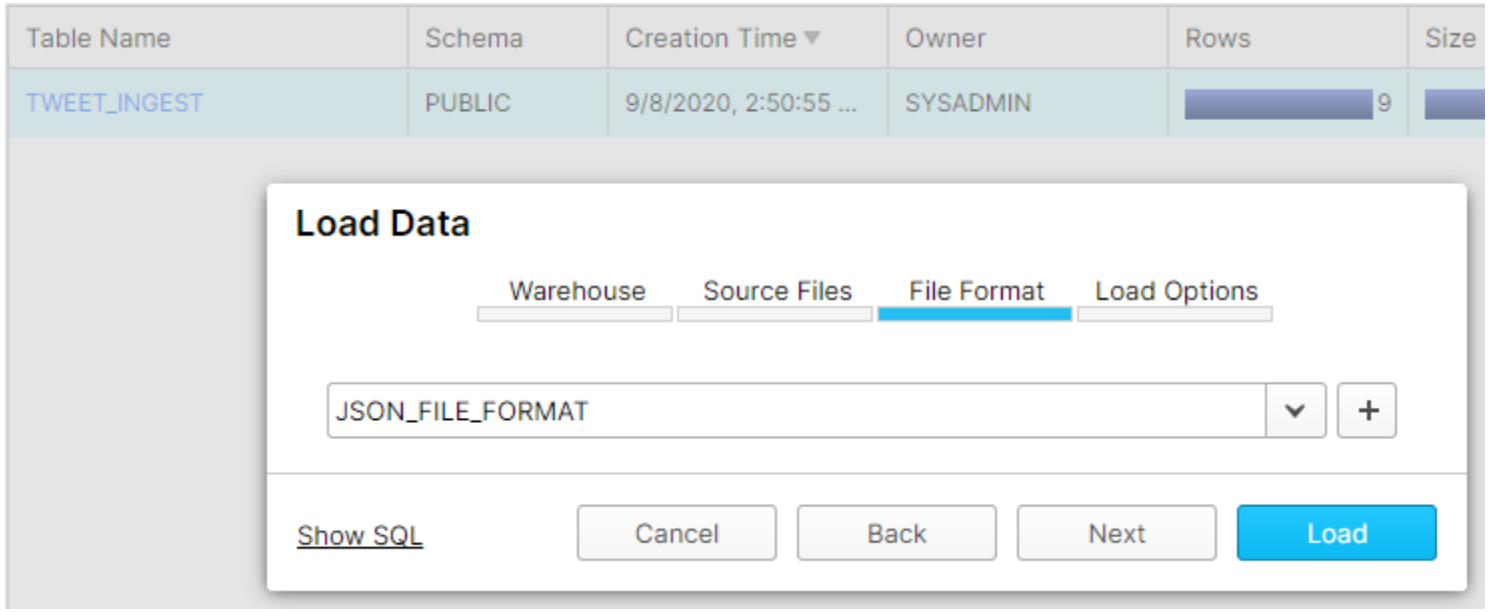
```
//Create a table in the new database
CREATE TABLE "SOCIAL_MEDIA_FLOODGATES"."PUBLIC"."TWEET_INGEST"
("RAW_STATUS" VARIANT)
COMMENT = 'Bring in tweets, one row per tweet or status entity';
```

- Create a file format that aligns with the data being imported (JSON)

```
//Create a JSON file format in the new database
CREATE FILE FORMAT "SOCIAL_MEDIA_FLOODGATES"."PUBLIC".JSON_FILE_FORMAT
TYPE = 'JSON'
COMPRESSION = 'AUTO'
ENABLE_OCTAL = FALSE
ALLOW_DUPLICATE = FALSE
STRIP_OUTER_ARRAY = TRUE
STRIP_NULL_VALUES = FALSE
IGNORE_UTF8_ERRORS = FALSE;
```

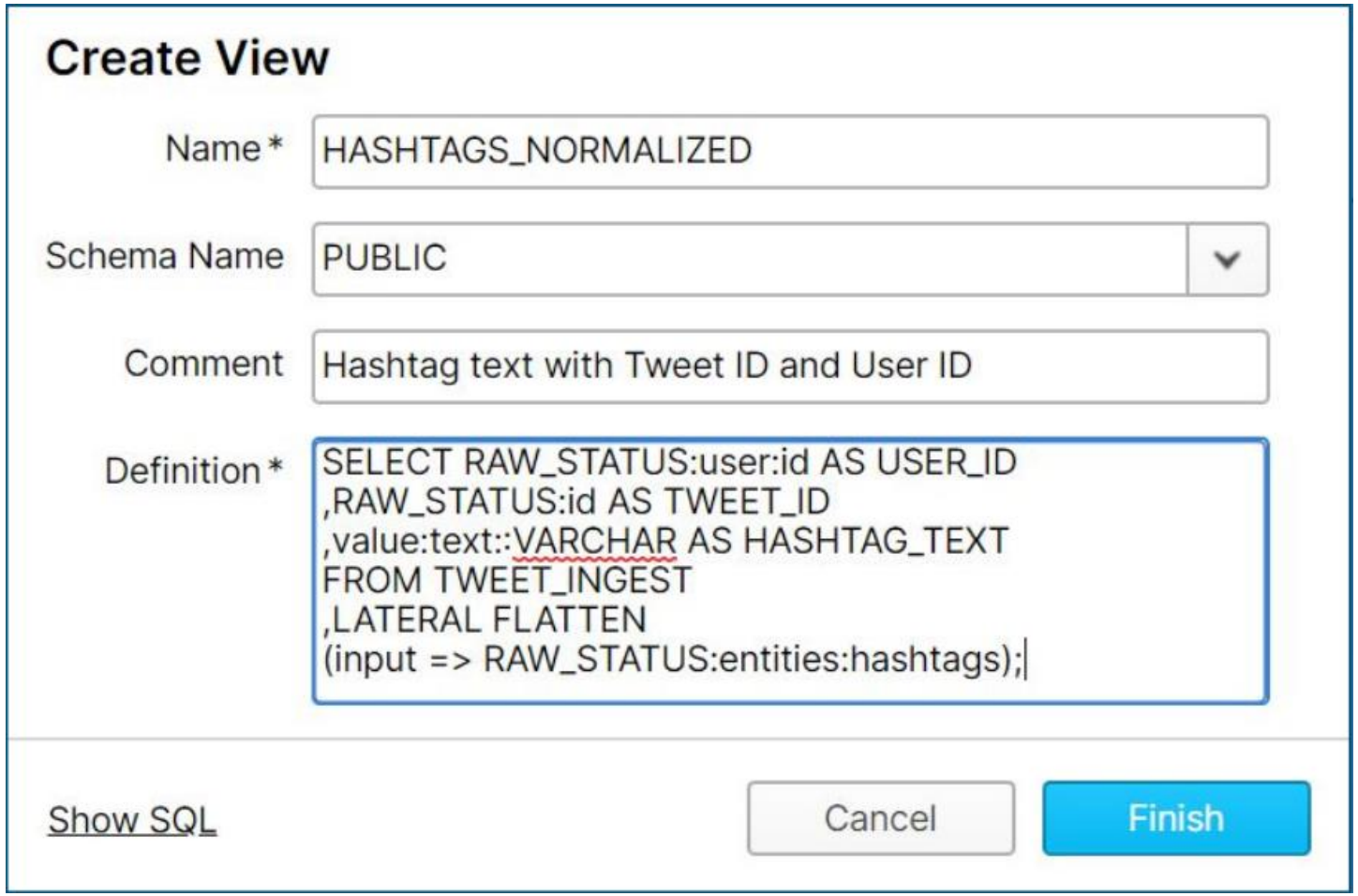
Load data from file

Load data from a JSON file into table using the JSON file format.



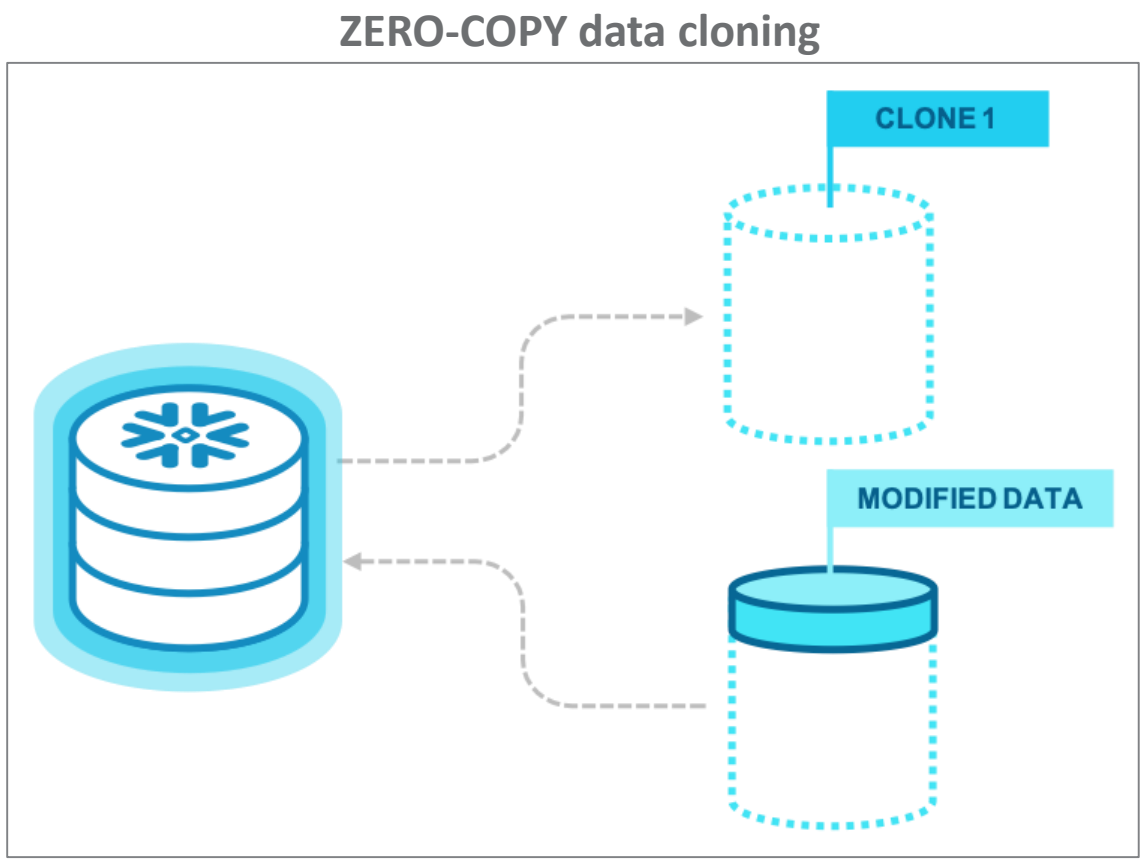

```
PUT file://<file path>/nutrition_tweets.json @TWEET_INGEST/ui1604406151312

COPY INTO "SOCIAL_MEDIA_FLOODGATES"."PUBLIC"."TWEET_INGEST" FROM
@/ui1604406151312 FILE_FORMAT =
'"SOCIAL_MEDIA_FLOODGATES"."PUBLIC"."JSON_FILE_FORMAT"' ON_ERROR =
'ABORT_STATEMENT' PURGE = TRUE;
```

Create structured view



8) Cloning

- Copy a huge database very quickly without consuming additional storage (cost)
- DevOps implications
  - Promoting table changes through Dev → Test → Prod
  - Copy Prod database back to Dev or Test quickly



Ref: https://www.snowflake.com/blog/saving-time-space-simplifying-devops-fast-cloning/

Cloning demo

Clone Prod to a Dev database

```
CREATE DATABASE Dev CLONE Prod;
```

Promote a new table from Dev to Test or Prod

- With data…

```
CREATE TABLE C CLONE Dev.public.C;
```

- Without data…

```
CREATE TABLE C LIKE Dev.public.C;
```


Ref: https://www.snowflake.com/blog/saving-time-space-simplifying-devops-fast-cloning/

9) Time travel

- Query data back in time
- SQL extensions for “AT” and “BEFORE” keywords
- Automatically enabled with a 1-day retention
  - 90-day max
  - Impacts storage costs
- Benefits
  - Troubleshoot data loading and transformations
  - Don’t have to worry as much about making data mistakes

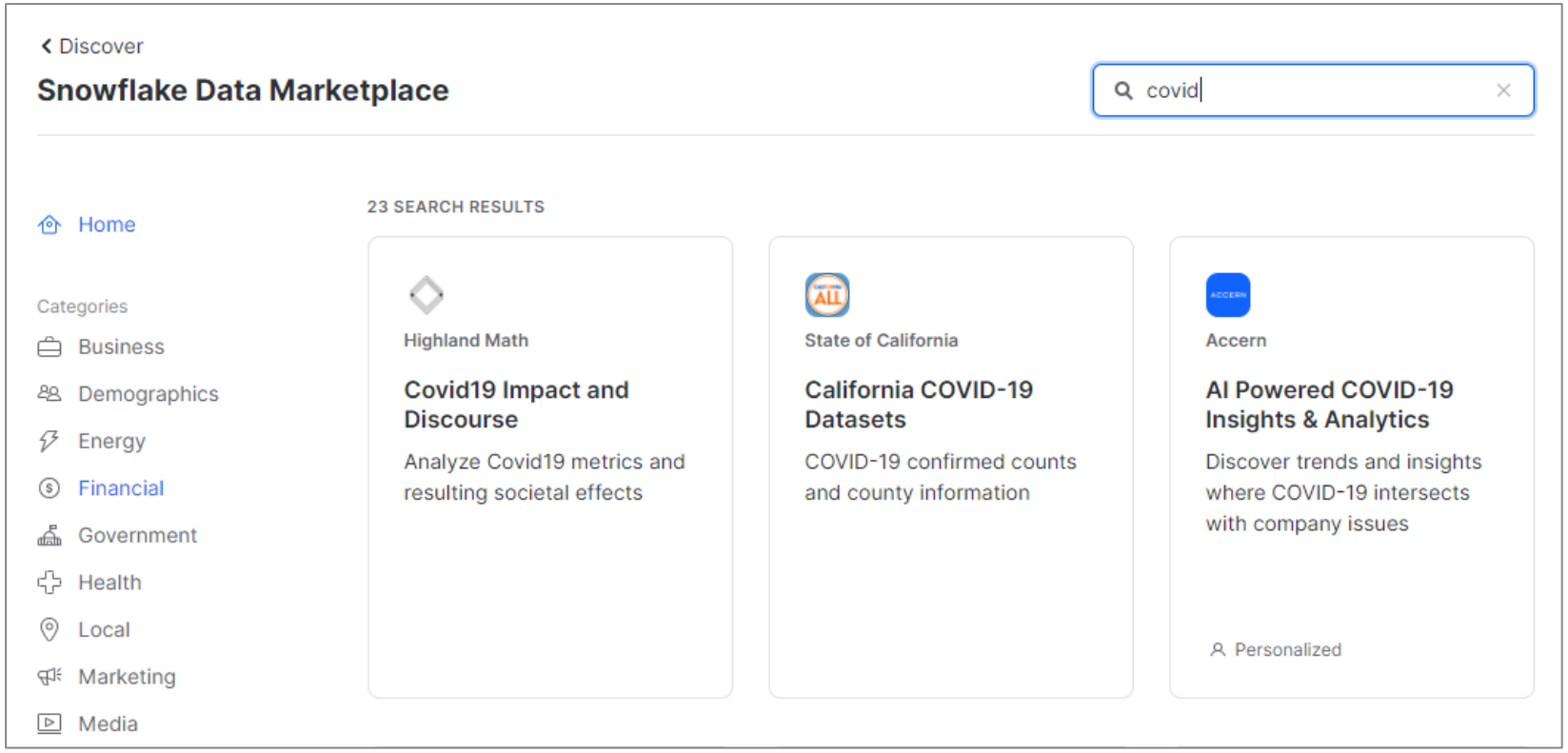
Time Travel demo

Example: select what the data was 5 minutes ago.

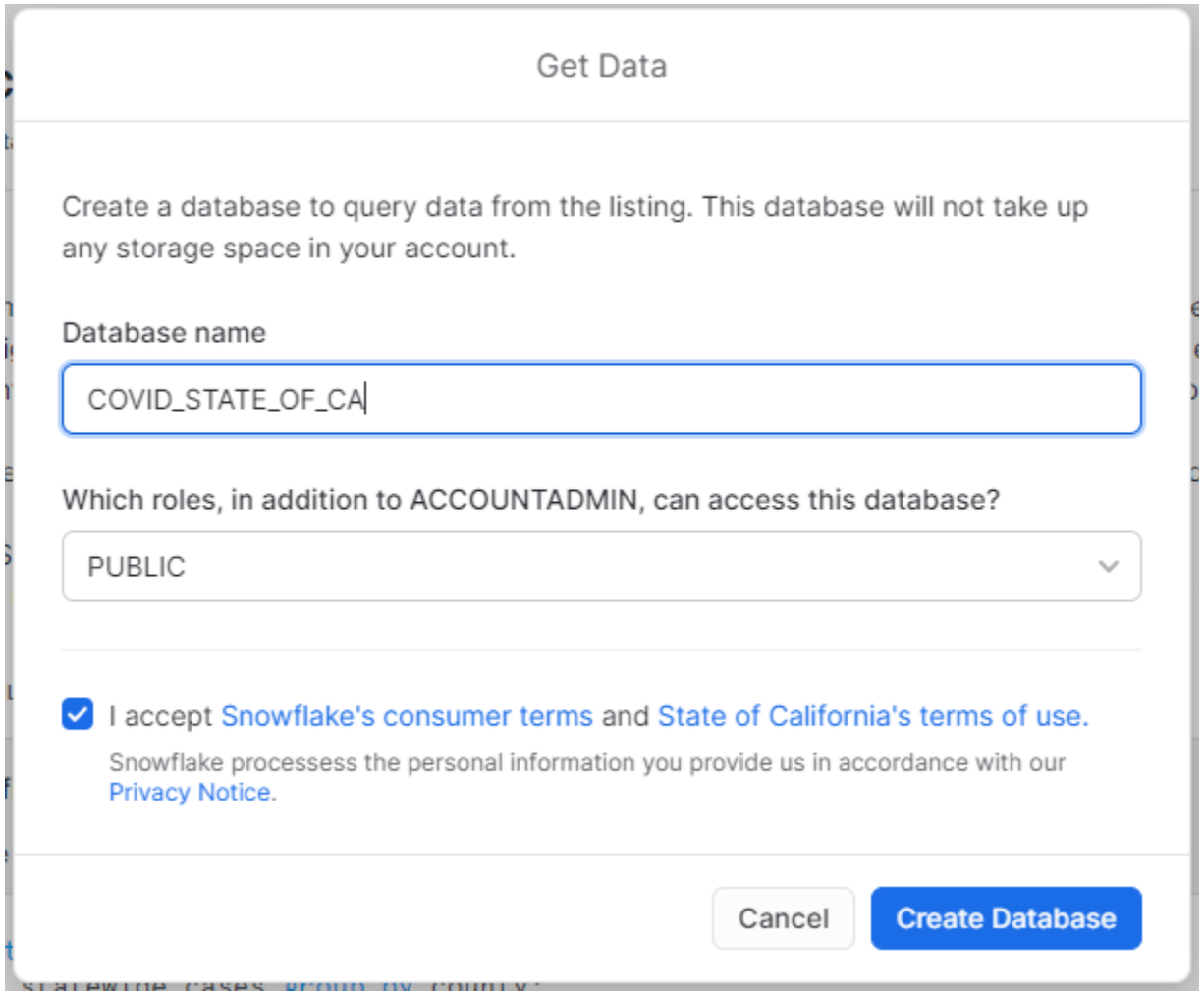
```
select * from my_table at(offset => -60*5);
```

10) Data sharing

- Pull in curated data sets
- Share your data with partners, safely, securely, efficiently



Data sharing demo



Data sharing demo

```
1  select top 10 * from "COVID_STATE_OF_CA"."OPEN_DATA"."CASE_DEMOGRAPHICS_AGE" ORDER BY Date DESC, AGE_GROUP ASC
```

Results | Data Preview

Query ID | SQL | 238ms | 10 rows

Filter result... | Copy

Row	AGE_GROUP	TOTALPOSITIVE	DATE	CASE_PERCENT	DEATHS	DEATHS_PERCENT
1	0-17	99679	2020-11-01	10.7	2	0
2	18-49	555933	2020-11-01	59.7	1274	7.2
3	50-64	175508	2020-11-01	18.9	3368	19
4	65+	98643	2020-11-01	10.6	13037	73.7
5	Missing	865	2020-11-01	0.1	2	0
6	0-17	99123	2020-10-31	10.7	2	0
7	18-49	553578	2020-10-31	59.7	1273	7.2
8	50-64	174737	2020-10-31	18.9	3365	19.1
9	65+	98240	2020-10-31	10.6	13020	73.7
10	Missing	856	2020-10-31	0.1	2	0

10 reasons Snowflake is built for analytics

1. Large data volumes
2. Data loading flexibility
3. Broad BI tools support
4. Supports “Analysis Ready” data models
5. Minimized administration
6. Performance scalability
7. Semi-structured data
8. Cloning
9. Time Travel
10. Data sharing

Questions?

Schedule a complimentary call to address your specific questions regarding using Snowflake for analytics

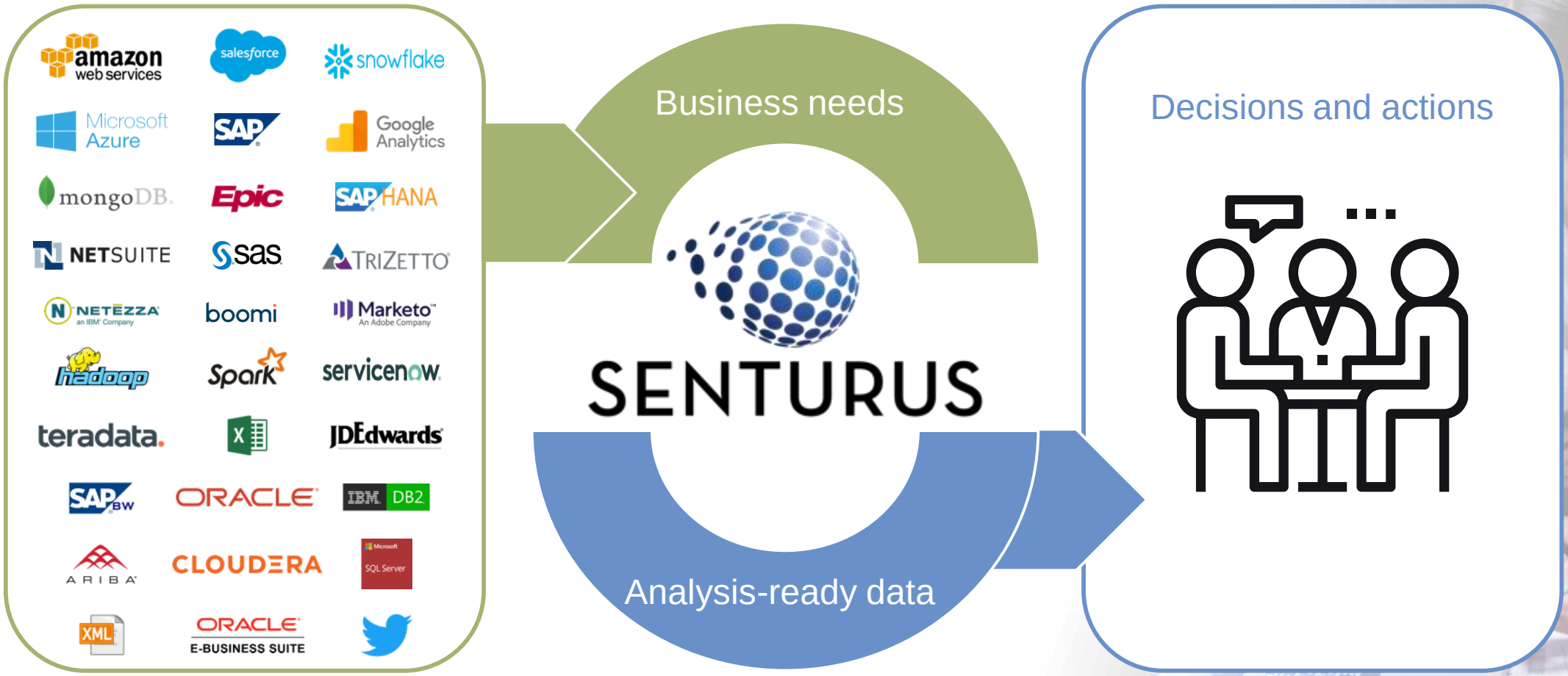
- Migration
- Performance
- Architecture

info@Senturus.com | 888 601 6010

The authority in Business Intelligence

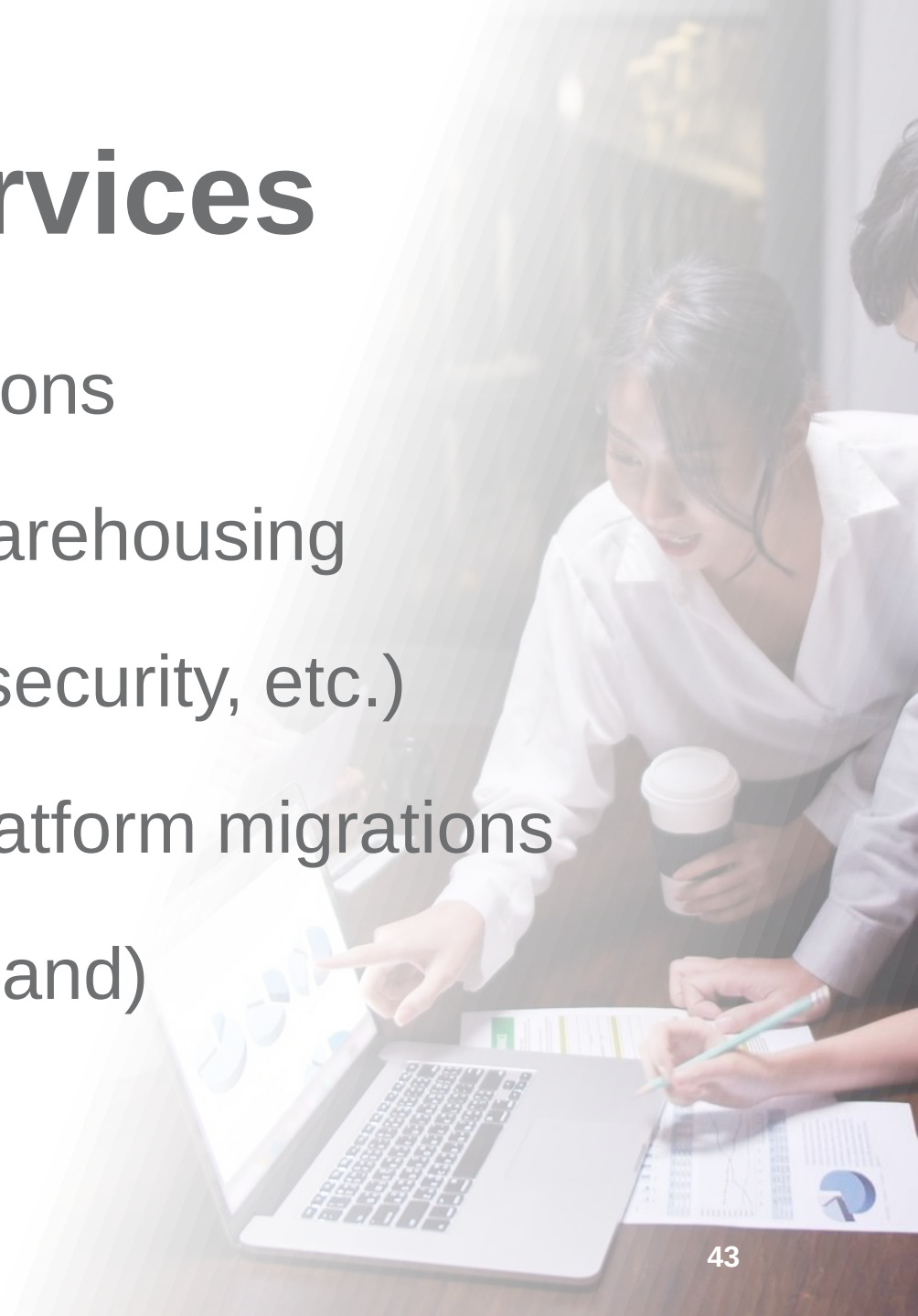
Exclusively focused on BI, Senturus is unrivaled in its expertise across the BI stack

Bridging the data and decisioning gap



Full spectrum of BI services

- Dashboards, reporting and visualizations
- Data preparation and modern data warehousing
- Hybrid BI environments (migrations, security, etc.)
- Software to enable bimodal BI and platform migrations
- BI support retainer (expertise on demand)
- Training and mentoring



A long, strong history of success

- 19+ years
- 1300+ clients
- 2500+ projects

Expand your knowledge

Find more resources
on the Senturus website

senturus.com/senturus-resources

Upcoming event

Data Integration Options for Microsoft Power BI

Choosing the right tool for the job

Thursday, Nov. 19, 2020, 11am PT/2pm ET

SENTURUS

Complete BI training offerings

IBM Cognos Analytics

Power BI

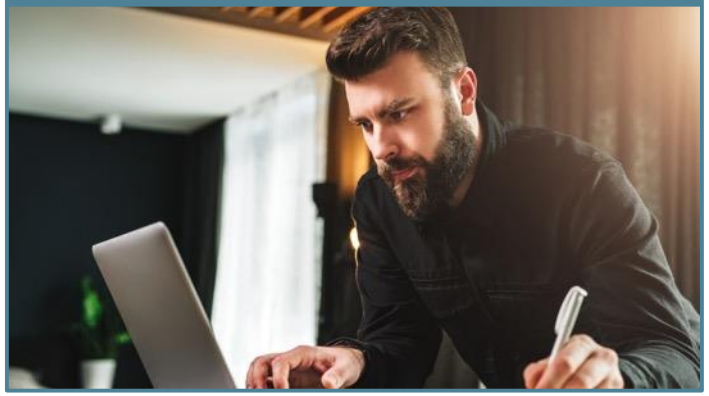
tableau

Tailored group sessions

Mentoring

Instructor-led online courses

Self-paced learning



Additional resources from Senturus

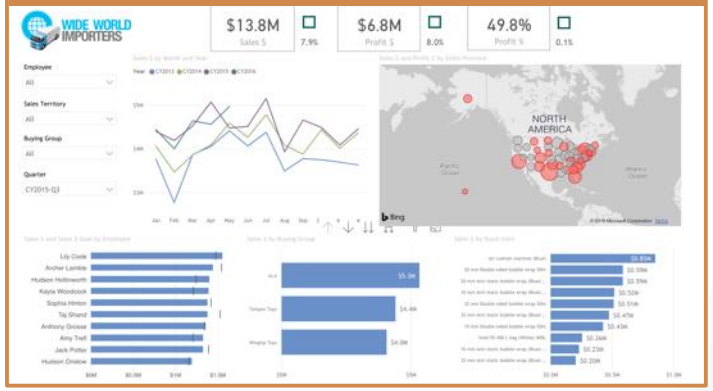
Unbiased product reviews

Technical tips

Insider viewpoints

More on this subject

Product demos



Upcoming events

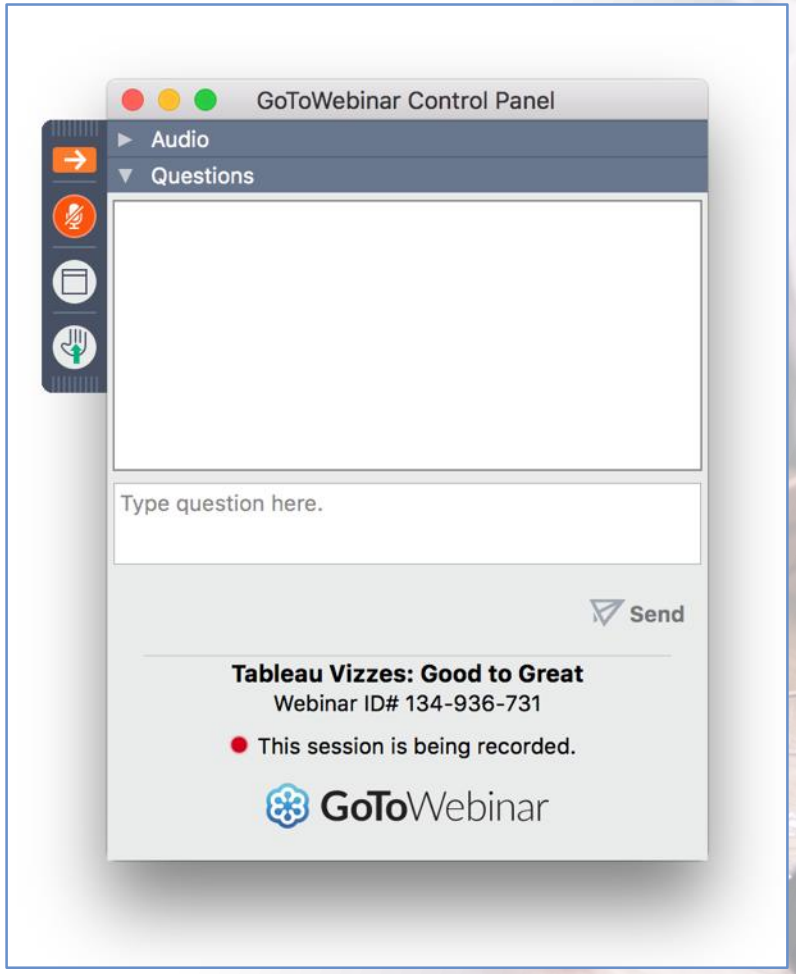
SENTURUS

Q & A

If your question or issue is broader than what we are able to answer today, contact us at

info@senturus.com

and we will set up a free consultation.



Thank You

www.senturus.com 888 601 6010 info@senturus.com