Snowflake Fundamentals

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake
2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

INTRODUCTIONS

- Name
- Company and Role
- Where is your company with its Snowflake implementation?
- What is your experience with Snowflake?
  - What do you already know?
  - What do you want to learn?
- Tell us something about yourself

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

REMOTE ATTENDEE CONTROLS

- Hover over your Zoom interface to display the control panel at the bottom

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

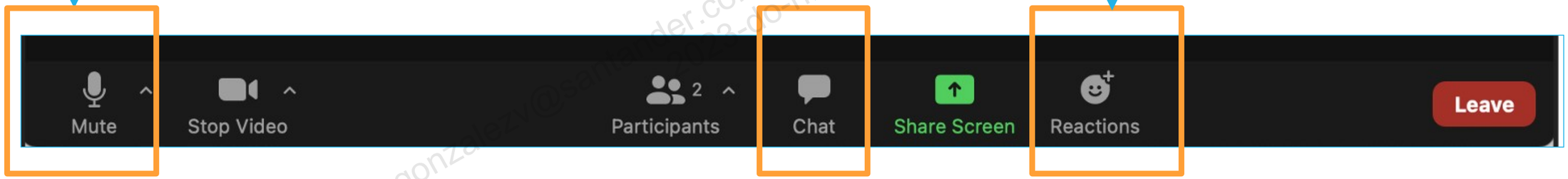
© 2023 Snowflake Computing Inc. All Rights Reserved

CONTROL PANEL

Mute/Unmute

Open the Chat window

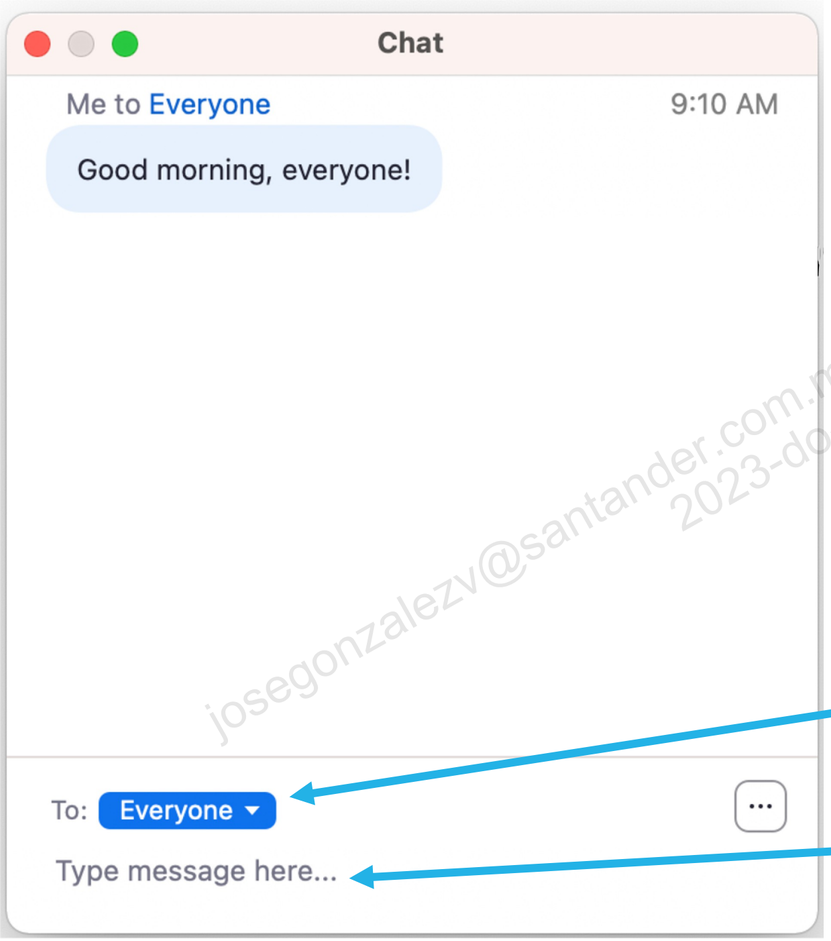
Reactions and nonverbal feedback



josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake
2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

CHAT INTERFACE



josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

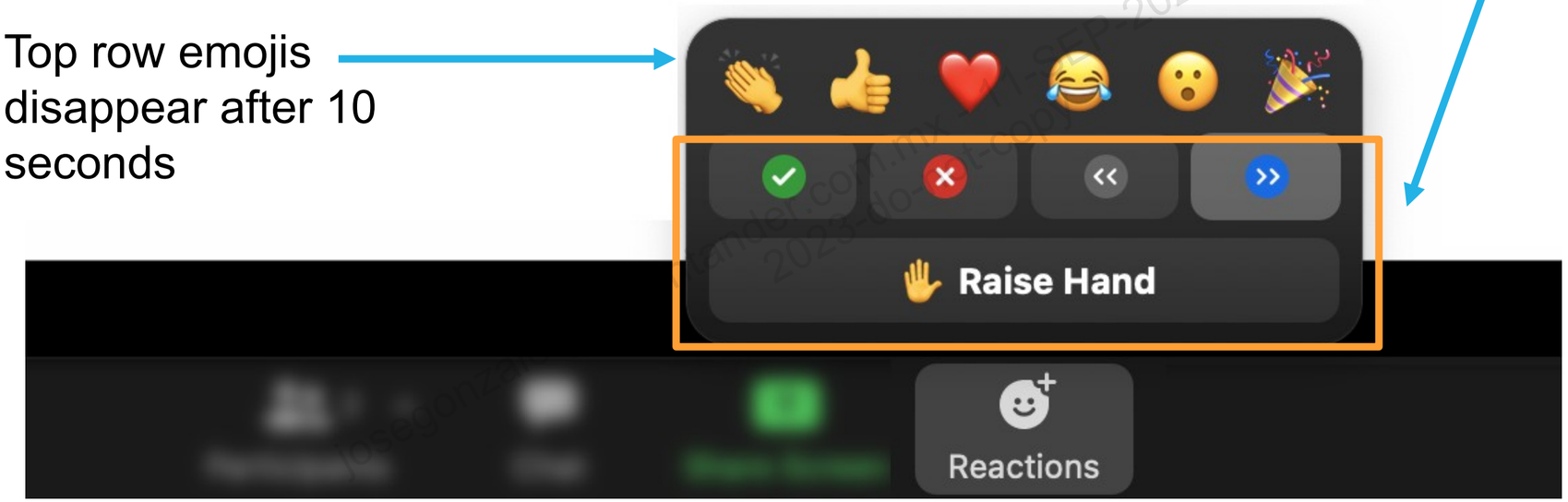
1. Use drop-down to choose recipient
2. Click to type your message and press Enter

© 2023 Snowflake Computing Inc. All Rights Reserved

REACTIONS AND NONVERBAL FEEDBACK

Use bottom two rows to raise your hand, click Yes/No, etc. (click again to cancel)

Top row emojis disappear after 10 seconds



Click **Reactions** to display panel

© 2023 Snowflake Computing Inc. All Rights Reserved

COURSE AGENDA

Overview and Architecture

Connecting to Snowflake

Data Protection Features

SQL Support in Snowflake

Caching in Snowflake

Query Performance

Data Loading and Unloading

Functions and Procedures

Using Tasks

Managing Security

Access Control and User Management

Semi-Structured Data

Introduction to Data Sharing

Virtual Warehouse Scaling

Financial Governance

Releases and Preview Features

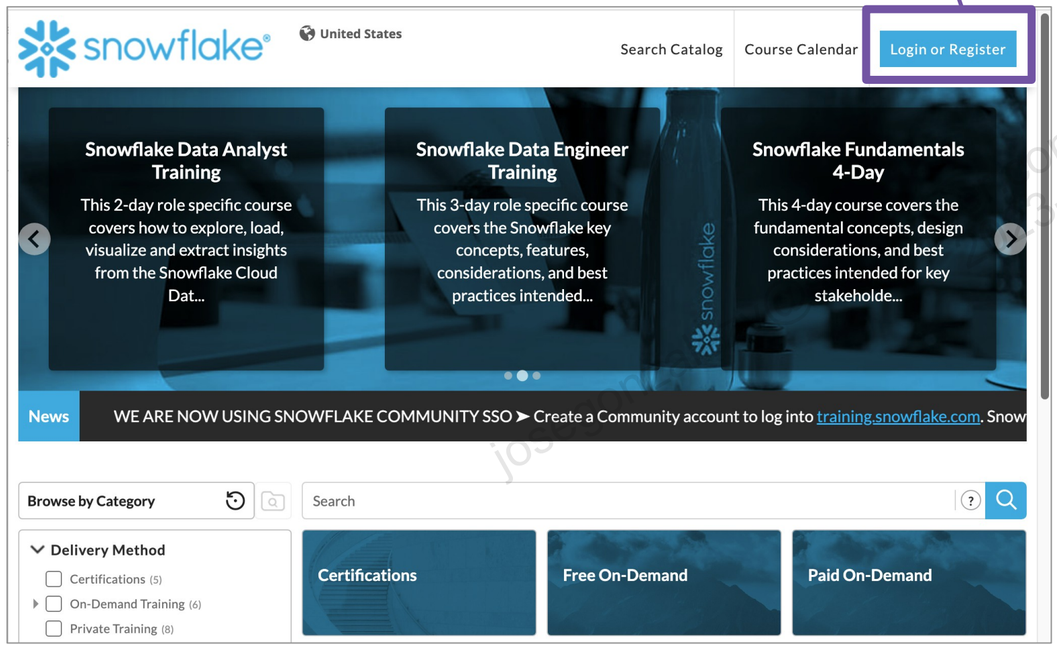
josegonzalezv@santander.com.mx - 11-SEP-2023 - Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

LOG IN TO YOUR COURSE DASHBOARD

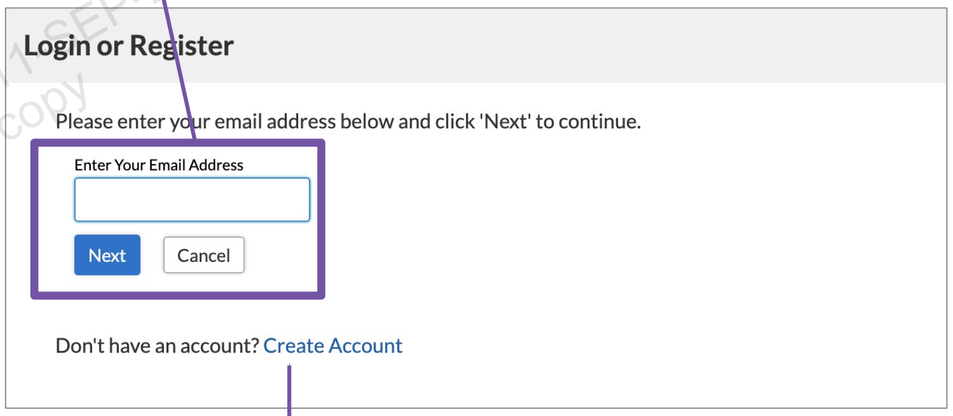
Go to https://training.snowflake.com
and click the "Login or Register" Button

Click to log in



Log in with your
Snowflake Community credentials

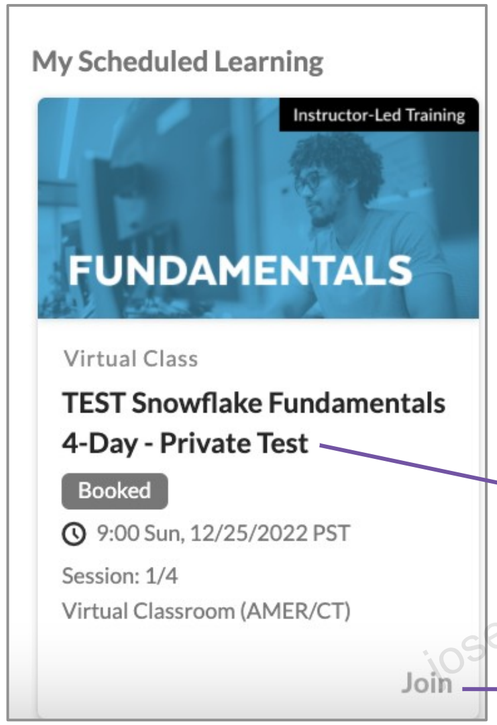
Type email address here



Don't have a Snowflake Community login?
Click "Create Account."

© 2023 Snowflake Computing Inc. All Rights Reserved

CLICK TO JOIN SESSION



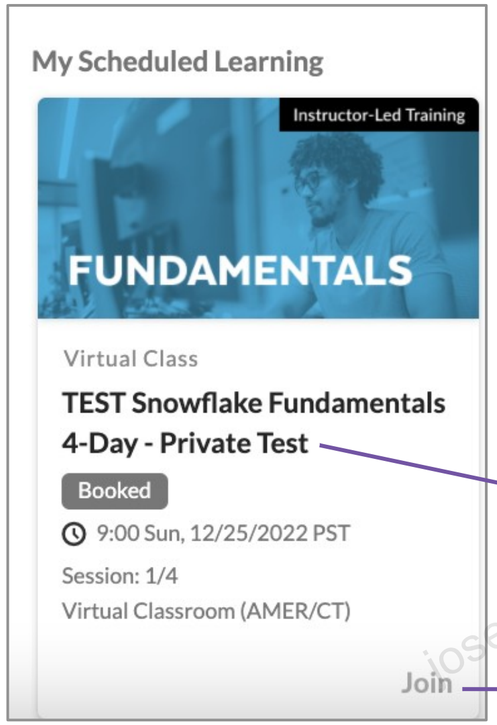
Your course name will be **different** from what is shown here

Join here each day of the course (link active about 30 minutes ahead)

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

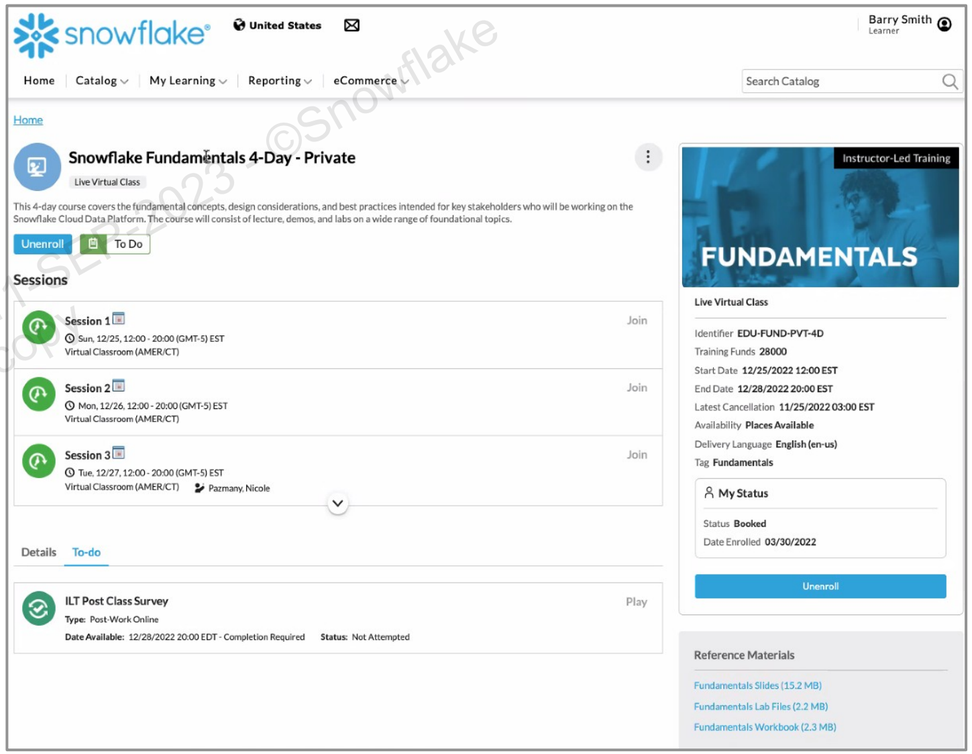
NAVIGATE TO YOUR COURSE PAGE



Click tile to view details

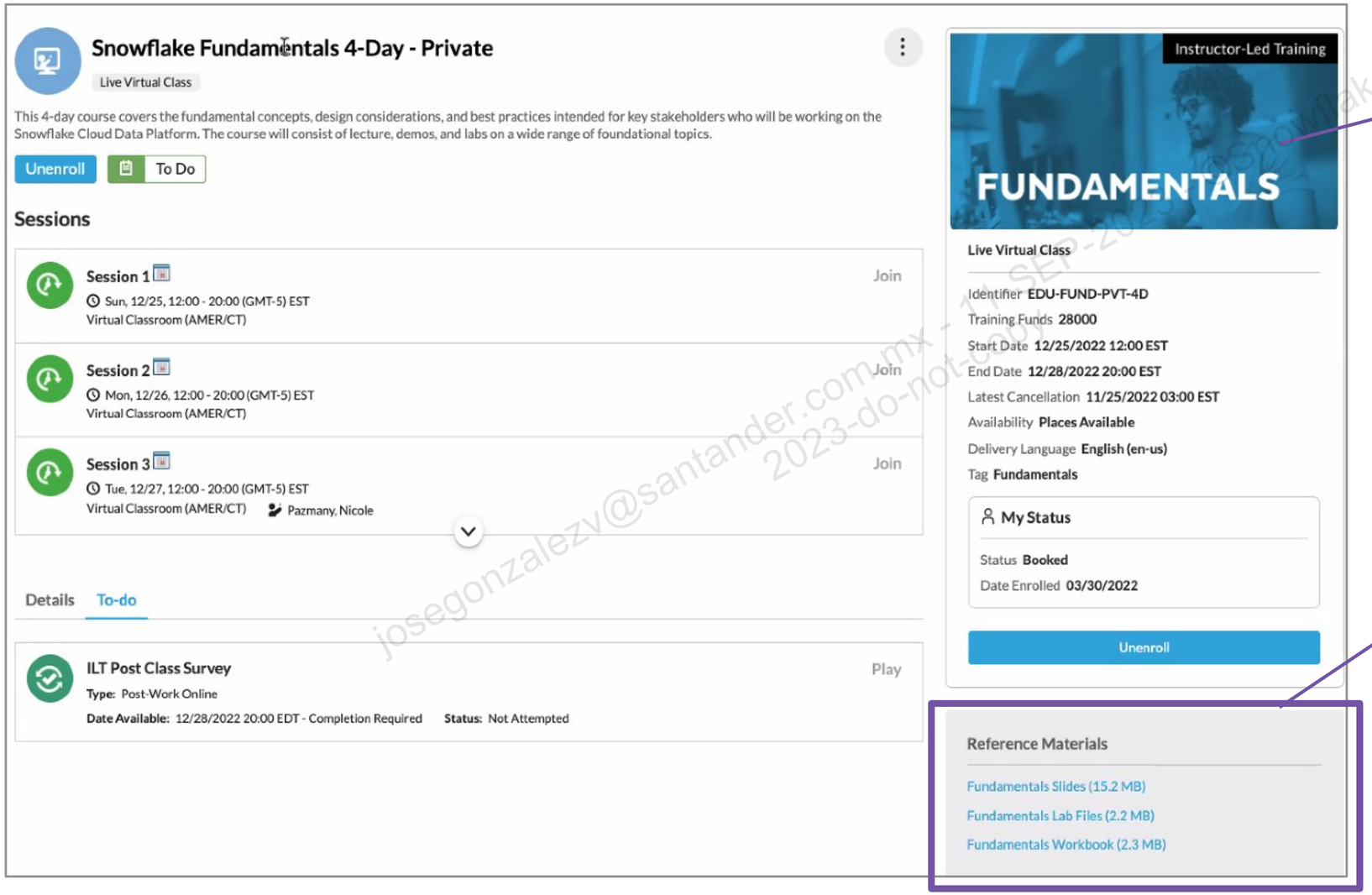
Your course name will be **different** from what is shown here

Join here each day of the course (link active about 30 minutes ahead)



© 2023 Snowflake Computing Inc. All Rights Reserved

ACCESS THE COURSE MATERIALS



Your course title may be different from what is shown

Use links to access materials (file names will vary from what is shown here)

© 2023 Snowflake Computing Inc. All Rights Reserved

LAB EXERCISE

10 minutes

Introduction - Prepare for Class

- Log in to Snowflake Training Portal
  - https://training.snowflake.com/login.html
- Instructor will provide you access to a Snowflake account
  - Lab access is available for 30 days from the first day of class.

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake
2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

OVERVIEW AND ARCHITECTURE

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake
2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

MODULE AGENDA

- Overview
- Snowflake Structure
- Using Snowsight
- Snowflake Architecture
- Storage Layer
- Compute Layer
- Cloud Services Layer

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

OVERVIEW

josegonzalezv@santander.com - 11-SEP-2023 - ©Snowflake
2023-do-not-copy

SNOWFLAKE - THE CLOUD DATA PLATFORM

DATA SOURCES

OLTP DATABASES

ENTERPRISE APPLICATIONS

THIRD-PARTY

WEB/LOG DATA

IoT

DATA ANALYSTS

DATA MONETIZATION

OPERATIONAL REPORTING

AD HOC ANALYSIS

REAL-TIME ANALYTICS

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©snowflake
2023-dc-1 - copy

© 2023 Snowflake Computing Inc. All Rights Reserved

PLATFORM REQUIREMENTS

FAST FOR ANY WORKLOAD

Run any number or type of jobs across all users and data volumes quickly and reliably.

IT JUST WORKS

Replace manual with automated to operate at scale, optimize costs, and minimize downtime.

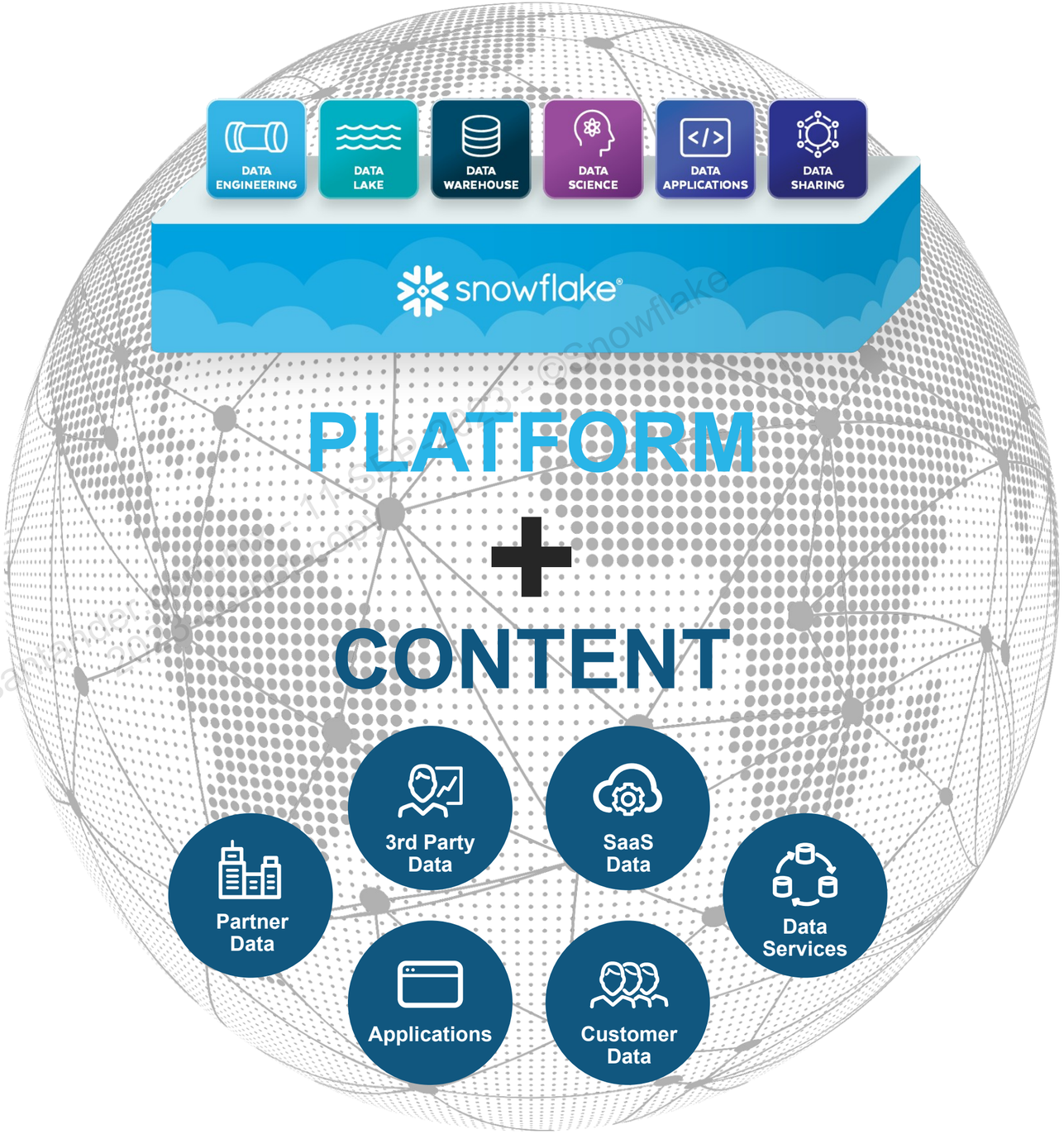
CONNECTED TO WHAT MATTERS

Extend access and collaboration across teams, workloads, clouds, and data, seamlessly and securely.

jsegonzalezv@santander.com - 11-SEP-2023 - ©Snowflake
2023 draft-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

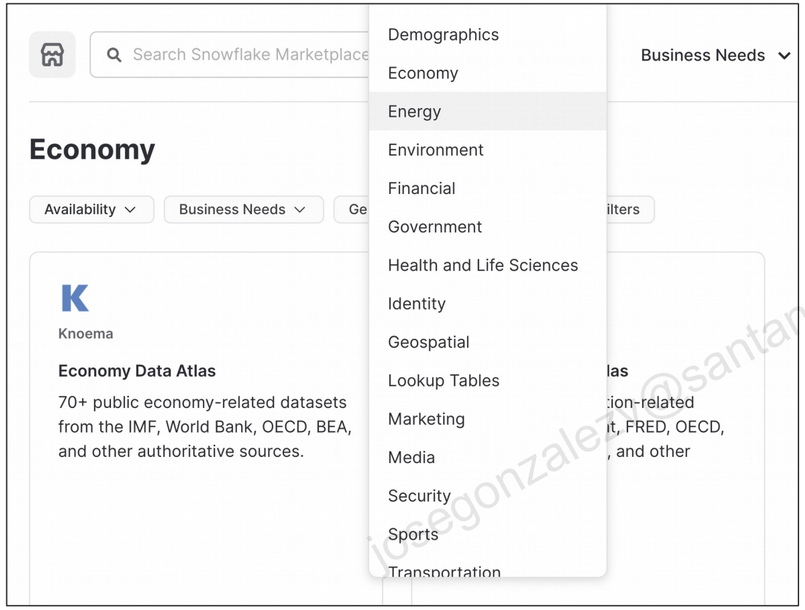
ELEMENTS OF THE DATA CLOUD



© 2023 Snowflake Computing Inc. All Rights Reserved

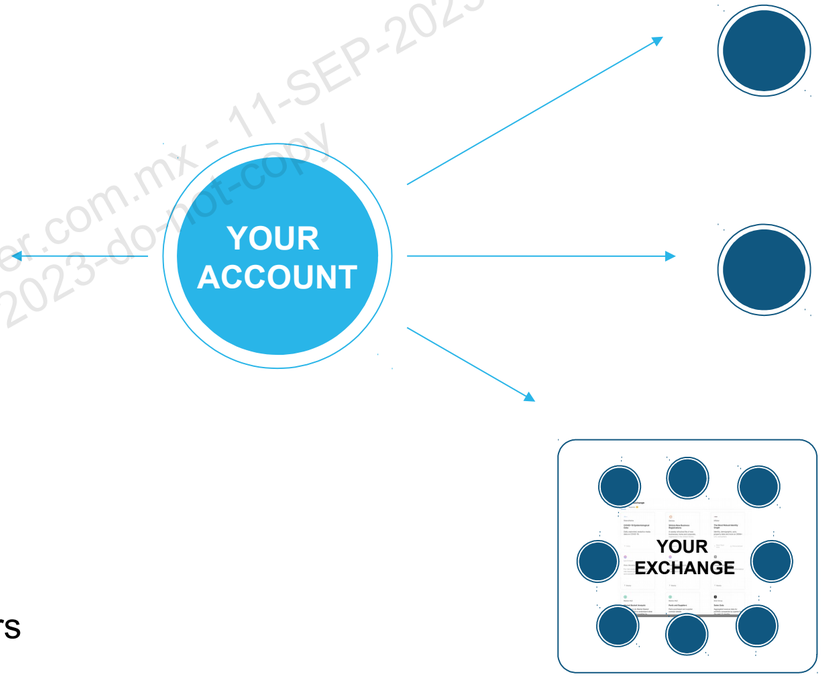
SHARE AND COLLABORATE IN THE DATA CLOUD

SNOWFLAKE MARKETPLACE



Access data and services from hundreds of providers

SHARE ACROSS YOUR BUSINESS ECOSYSTEM



DIRECT SHARE

Share with other Snowflake customers

READER ACCOUNTS

Share with companies not yet on Snowflake

DATA EXCHANGE

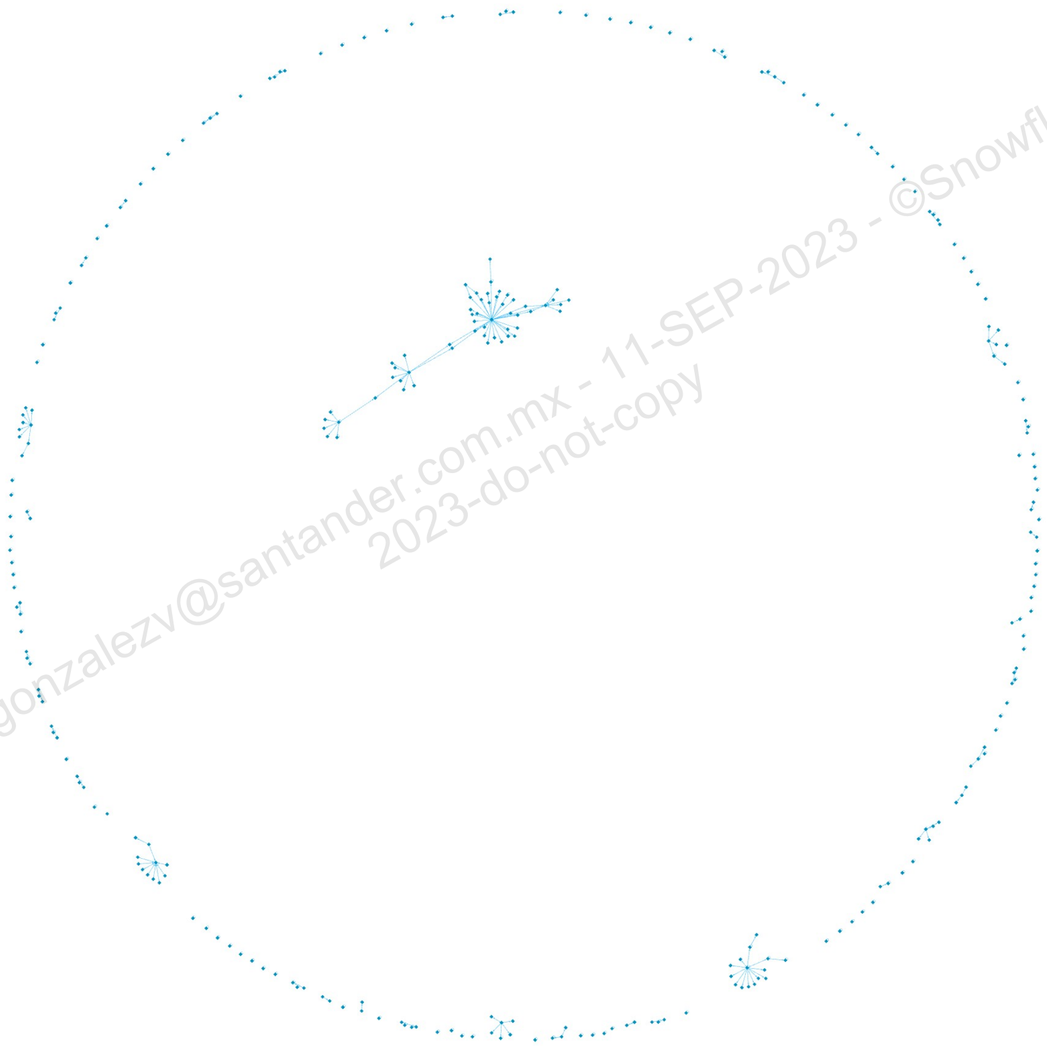
Administer group sharing and data discovery across business units

josegonzalez@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

DATA CLOUD GROWTH

APRIL 2020



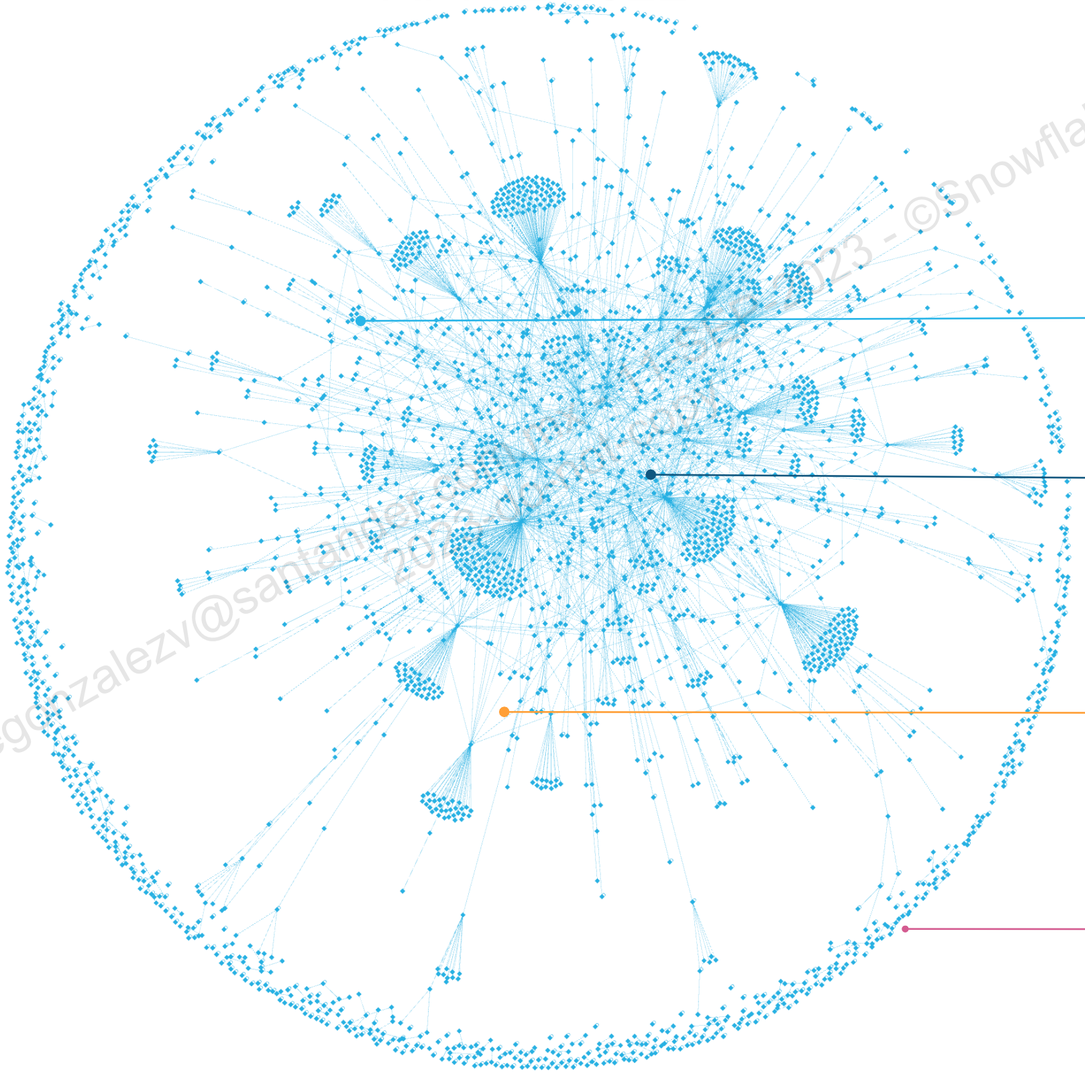
josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

THE DATA CLOUD TODAY

APRIL 2022

THOUSANDS OF ORGANIZATIONS ARE SHARING DATA WITH THEIR ECOSYSTEM



A software company shares terabytes of data with hundreds of customers

COVID-19 data is available live on Snowflake Marketplace from a US State, and other organizations

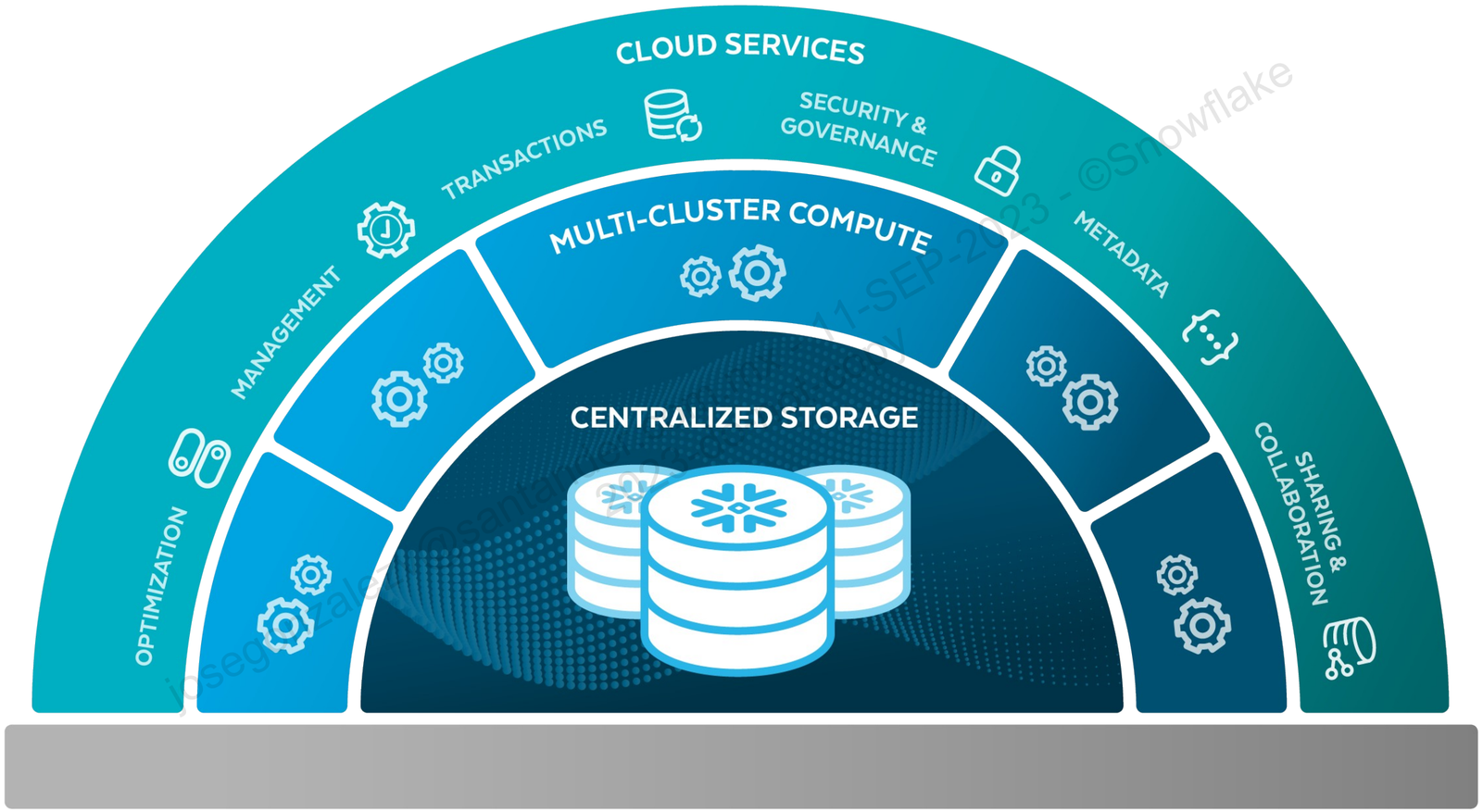
Today's financial data is accessible immediately without data pipelines

Thousands of companies share data with suppliers, partners, or other business units

© 2023 Snowflake Computing Inc. All Rights Reserved

*Visualization based on actual Data Cloud sharing activity as of April 2022

SNOWFLAKE ARCHITECTURE



Google Cloud aws Azure

© 2023 Snowflake Computing Inc. All Rights Reserved

FUNCTIONAL ARCHITECTURE

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

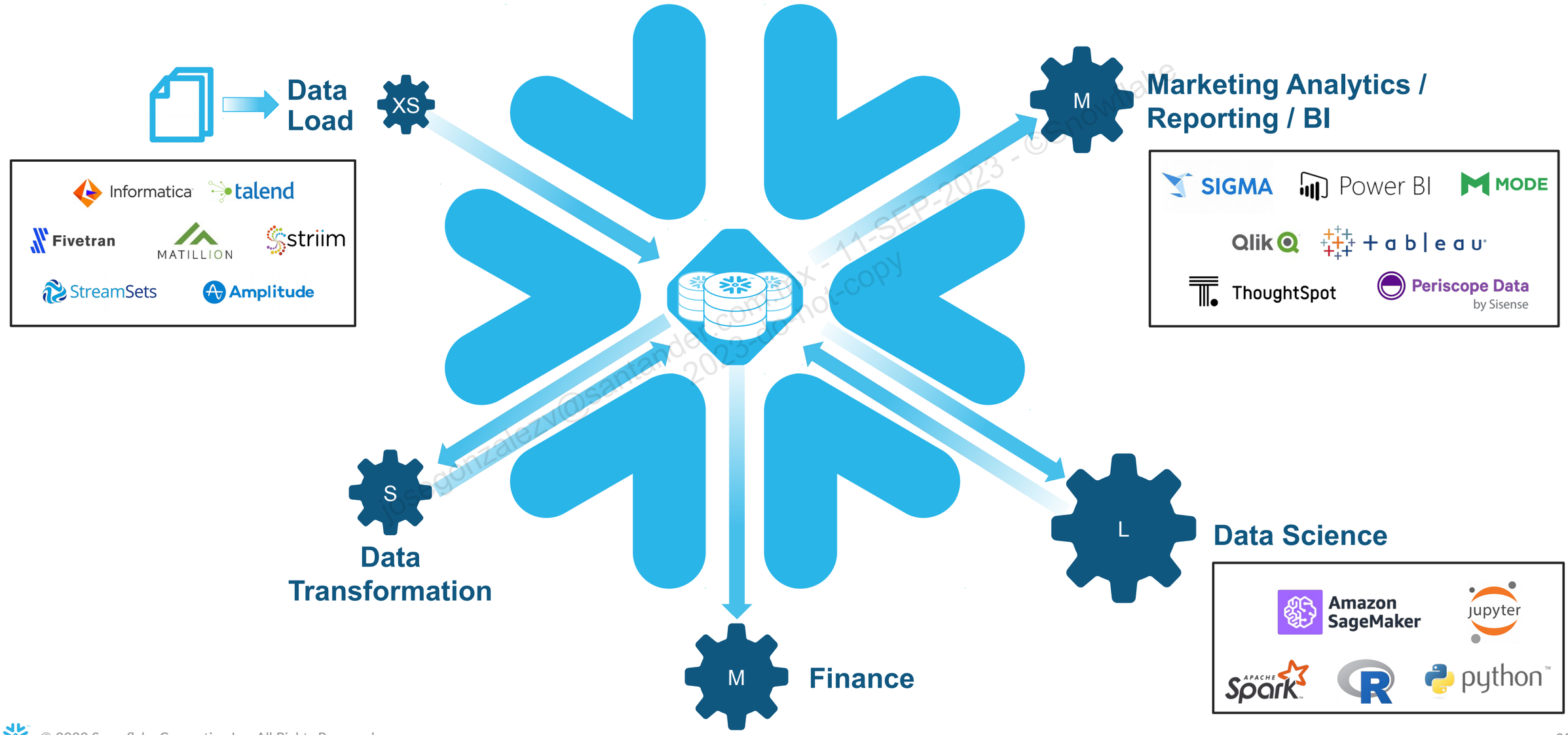
© 2023 Snowflake Computing Inc. All Rights Reserved

FUNCTIONAL ARCHITECTURE

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

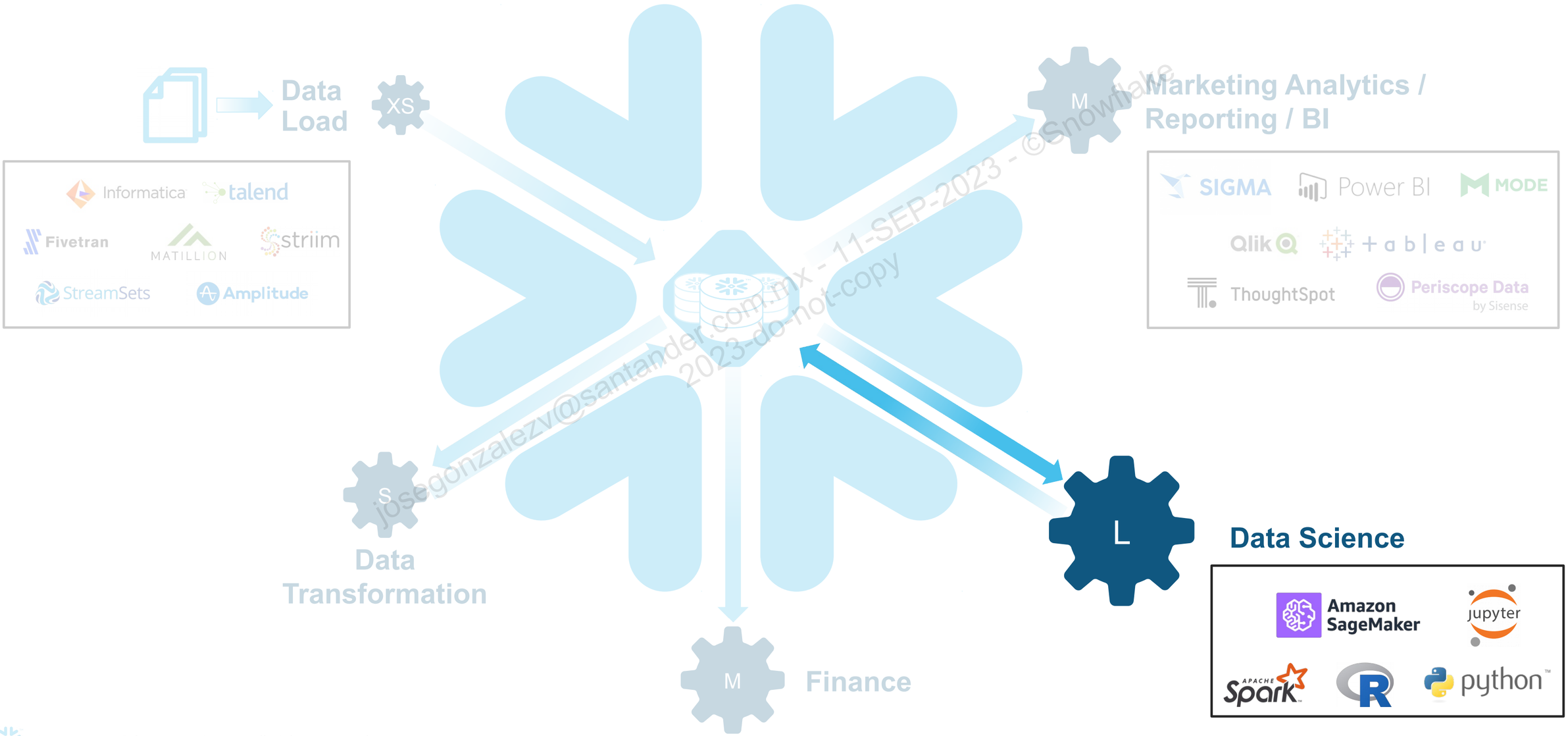
FUNCTIONAL ARCHITECTURE



jose.gonzalez.v@santander.com.mx - 11-SEP-2023 - © Snowflake 2023 - do not copy

© 2023 Snowflake Computing Inc. All Rights Reserved

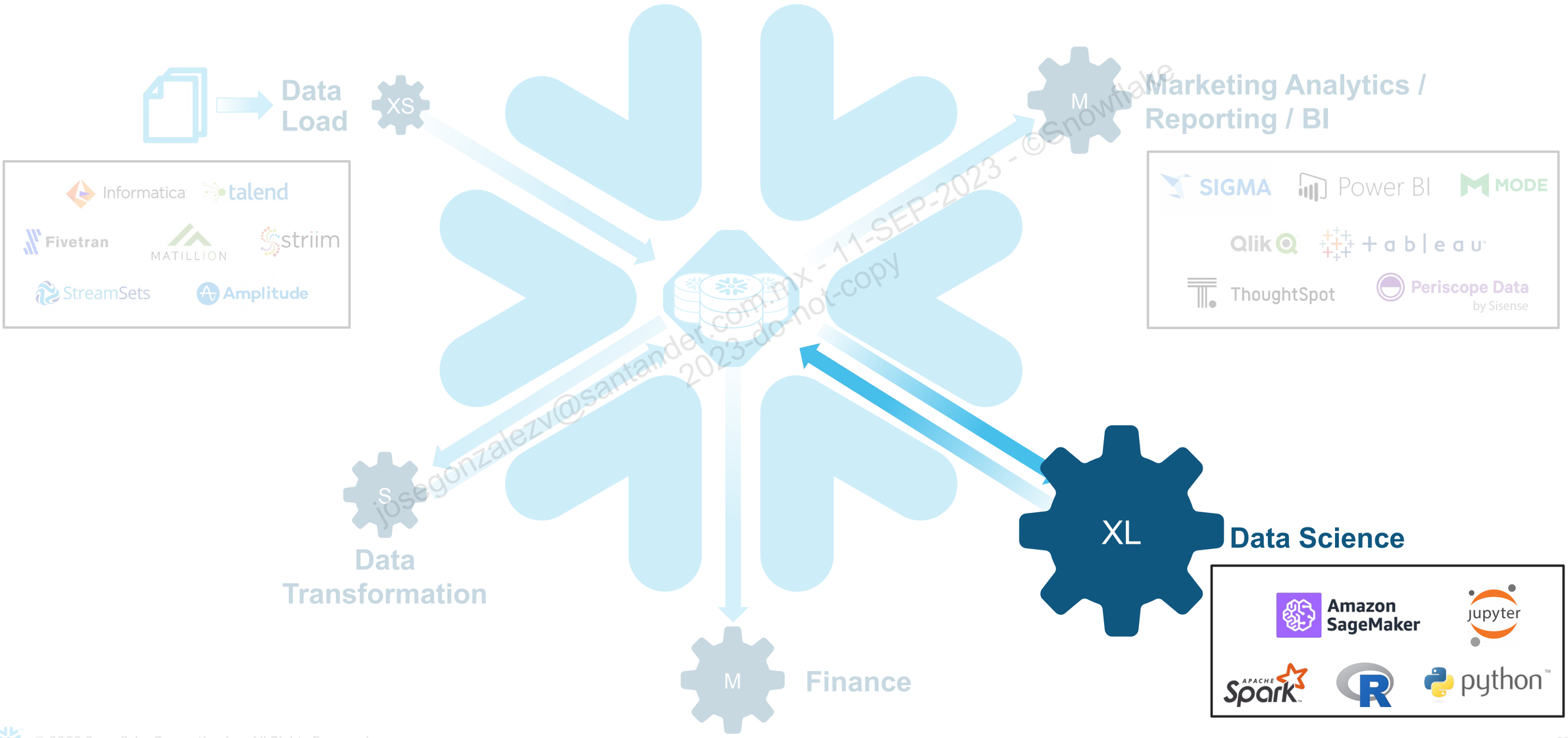
FUNCTIONAL ARCHITECTURE



josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

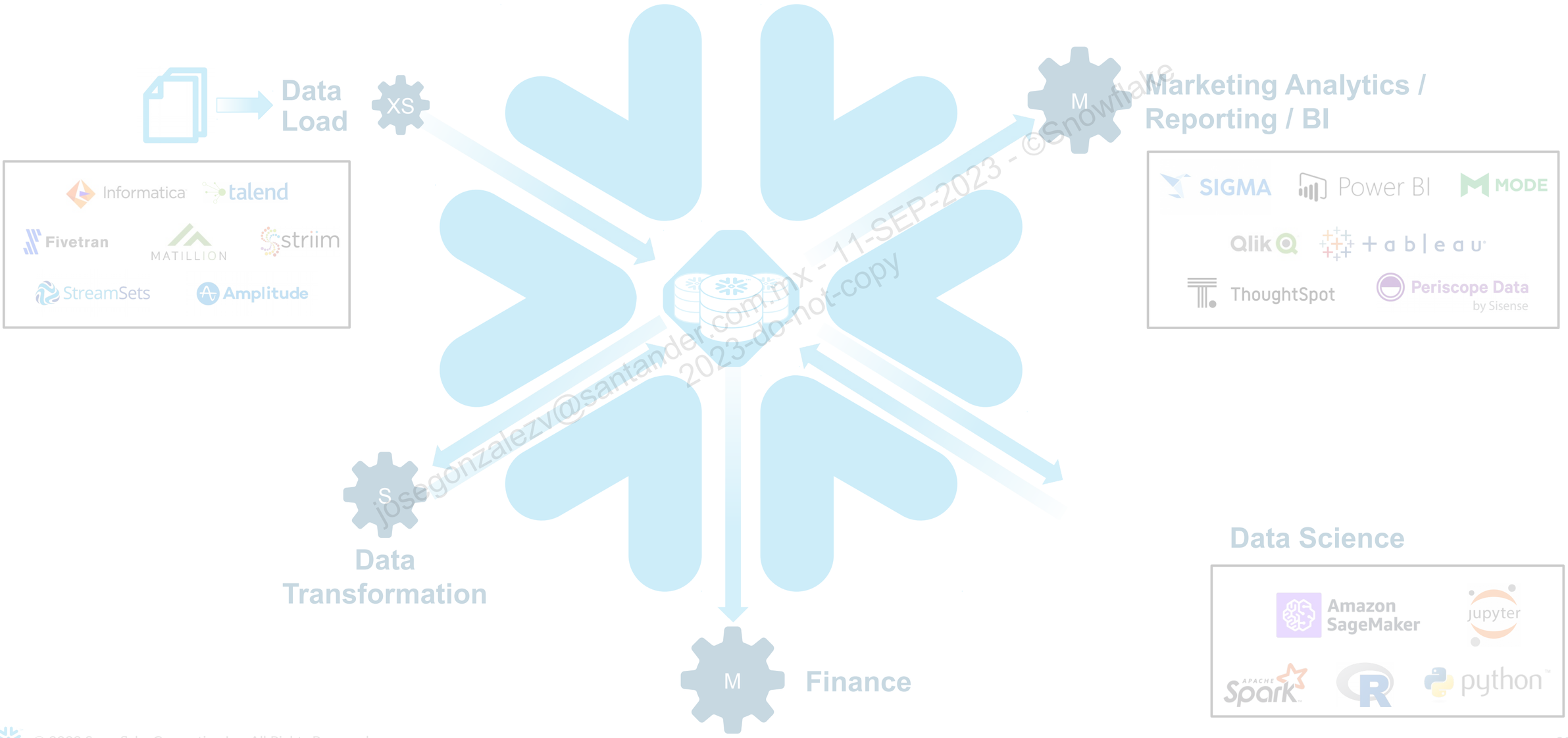
FUNCTIONAL ARCHITECTURE



josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

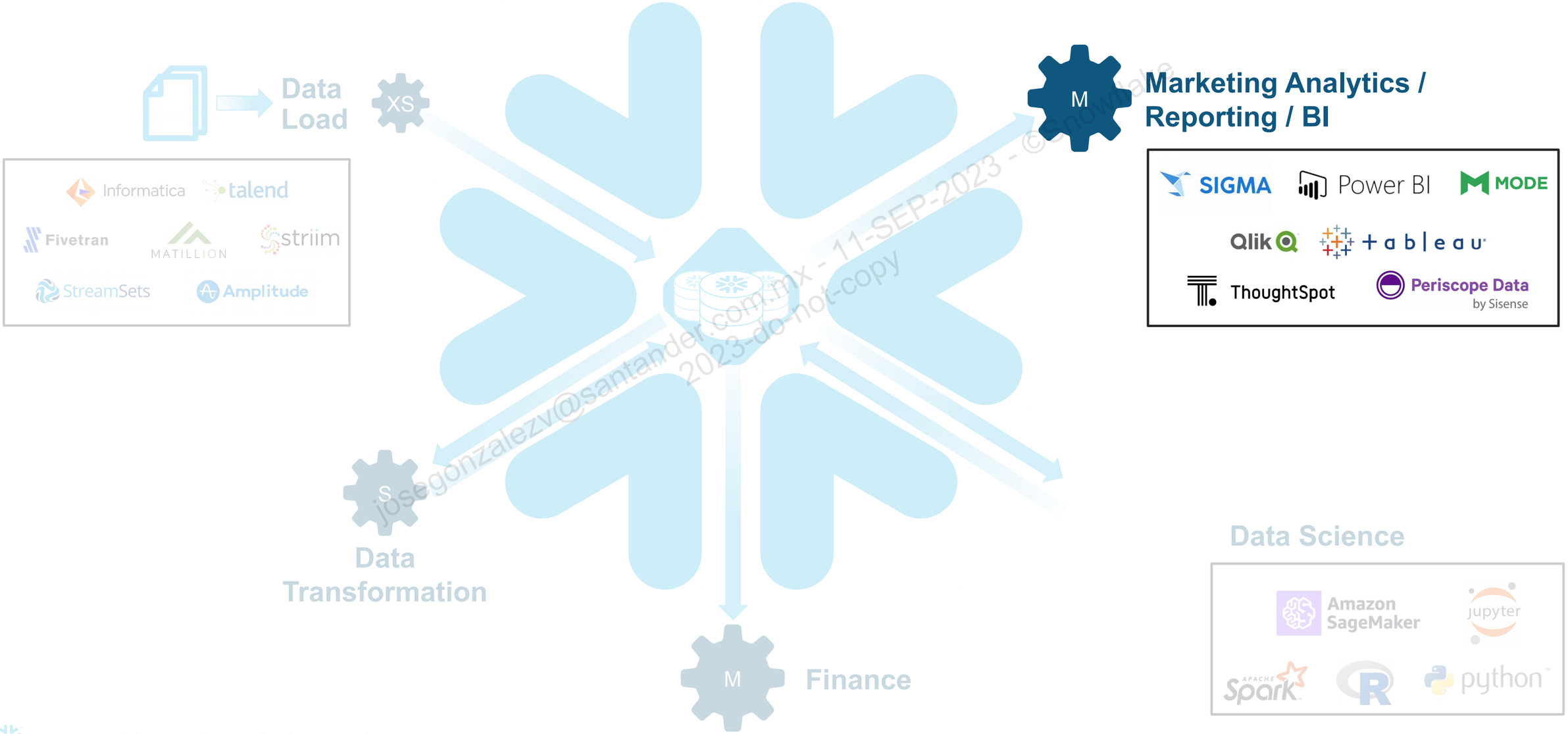
FUNCTIONAL ARCHITECTURE



josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

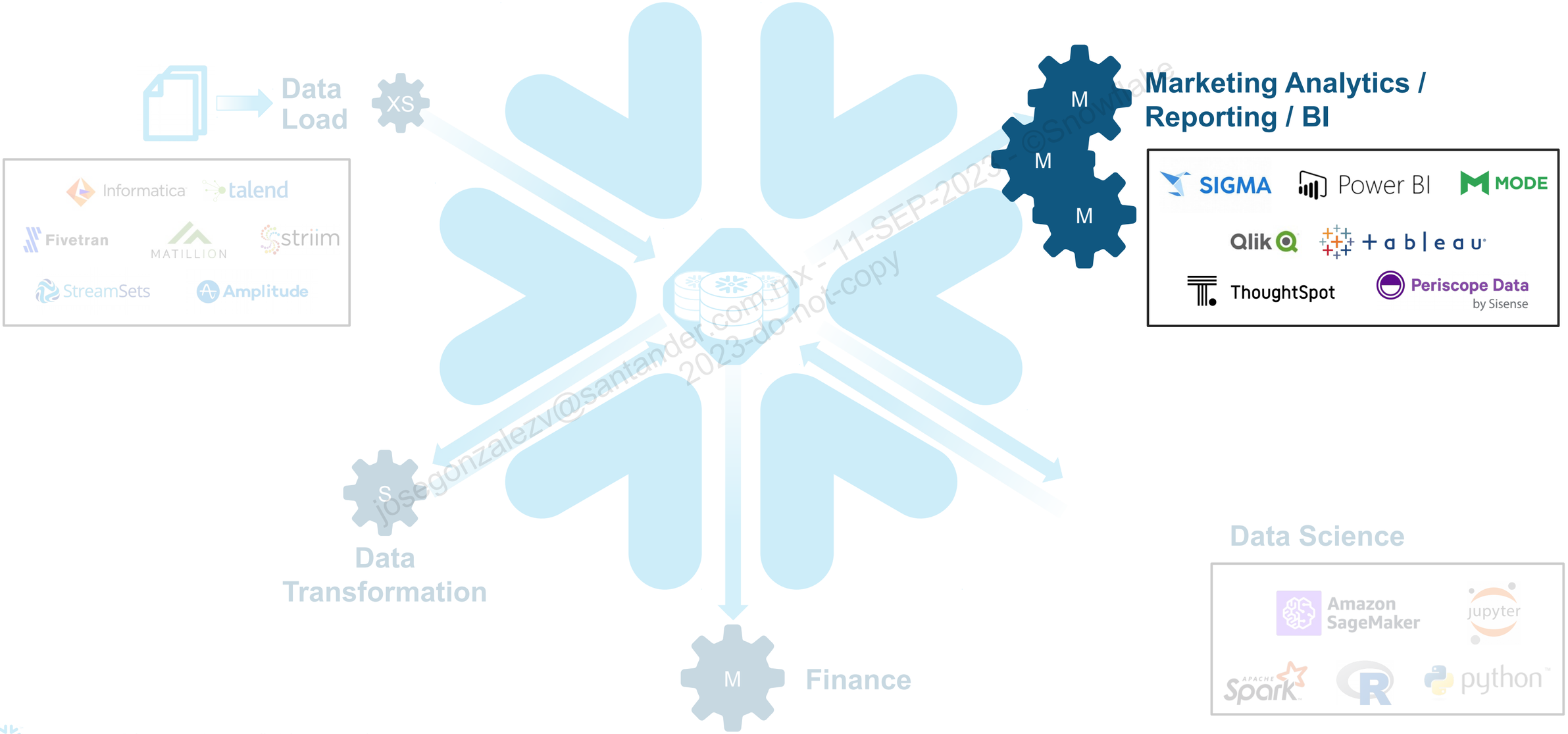
FUNCTIONAL ARCHITECTURE



josegonzalezv@santander.com.mx - 11-SEP-2023 - © Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

FUNCTIONAL ARCHITECTURE



josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

SNOWFLAKE EDITIONS

STANDARD

Complete SQL data warehouse

Secure Data Sharing across regions / clouds

Premier Support 24 x 365

1 day of time travel

Always-on enterprise grade encryption in transit and at rest

Customer-dedicated virtual warehouses

Federated authentication

Database replication

External Functions

Snowsight

Create your own Data Exchange

Data Marketplace access

ENTERPRISE

Standard +

Multi-cluster warehouse

Up to 90 days of time travel

Annual rekeying of all encrypted data

Materialized views

Search Optimization Service

Dynamic Data Masking

External Data Tokenization

BUSINESS CRITICAL

Enterprise +

HIPAA support

PCI compliance

Tri-Secret Secure using customer-managed keys

AWS PrivateLink support

Azure Private Link support

Google Cloud Private Service Connect support

Database failover and failback for business continuity

External Functions - AWS API Gateway Private Endpoints support

VIRTUAL PRIVATE SNOWFLAKE (VPS)

Business Critical +

Customer-dedicated virtual servers wherever the encryption key is in memory

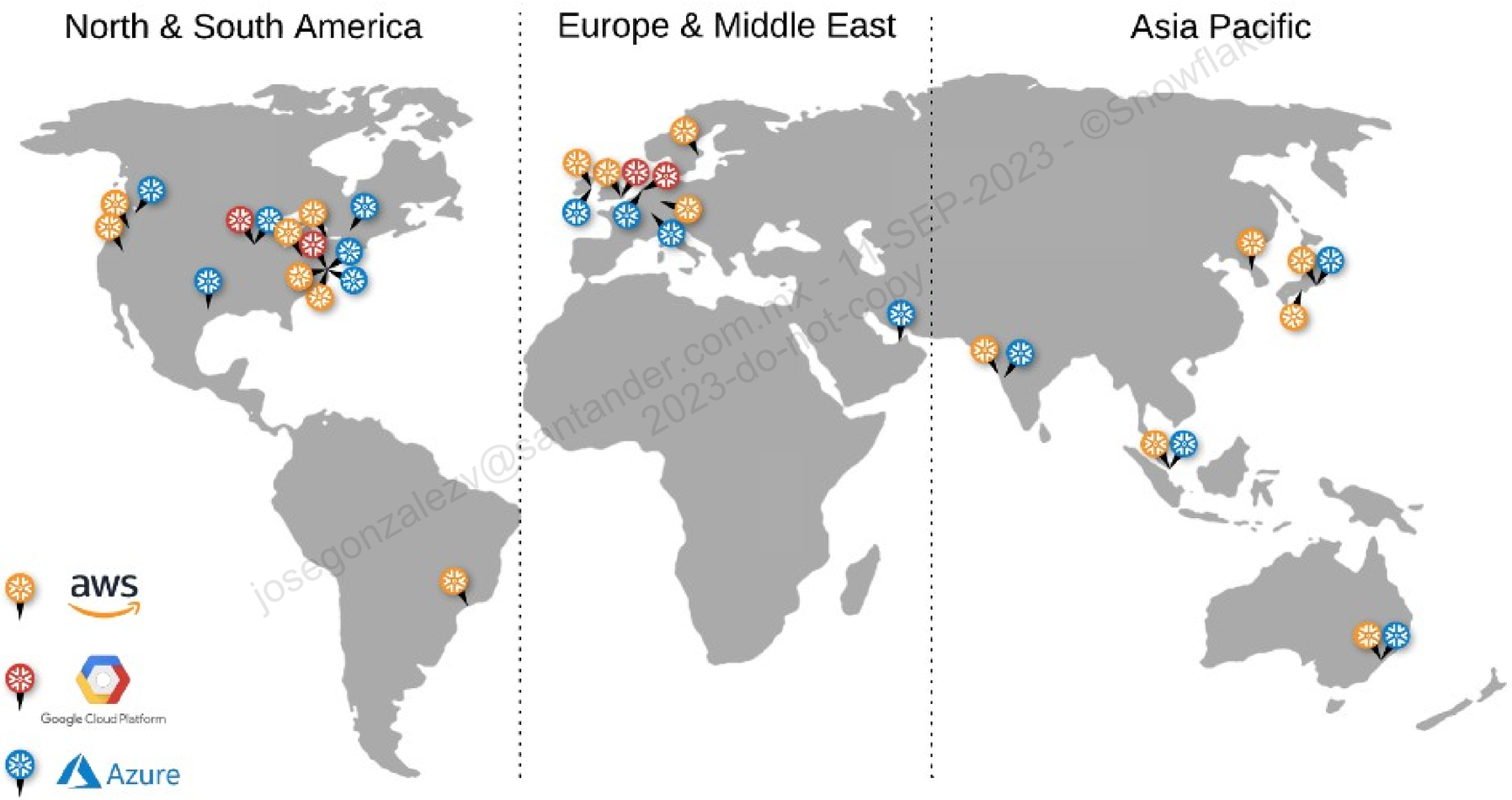
Customer-dedicated metadata store

jusegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

GLOBAL SNOWFLAKE

CURRENTLY SUPPORTED REGIONS



josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

KNOWLEDGE CHECK

What is a virtual warehouse?

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

KNOWLEDGE CHECK

What is a virtual warehouse?

A collection of compute resources (CPU, memory, and SSD) used to run queries or to load/unload data

josegonzalezv@santander.com.mx - 11-SEP-2023 - @Snowflake
2023-do-not-copy

KNOWLEDGE CHECK

Complete the following sentence:

Snowflake separates _____ and _____ to eliminate resource contention.

A. Databases and schema
B. Compute and storage
C. Large queries and small queries
D. Production data and test data

jose.gonzalez@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

KNOWLEDGE CHECK

Complete the following sentence:

Snowflake separates _____ and _____ to eliminate resource contention.

A. Databases and schema
B. Compute and storage
C. Large queries and small queries
D. Production data and test data

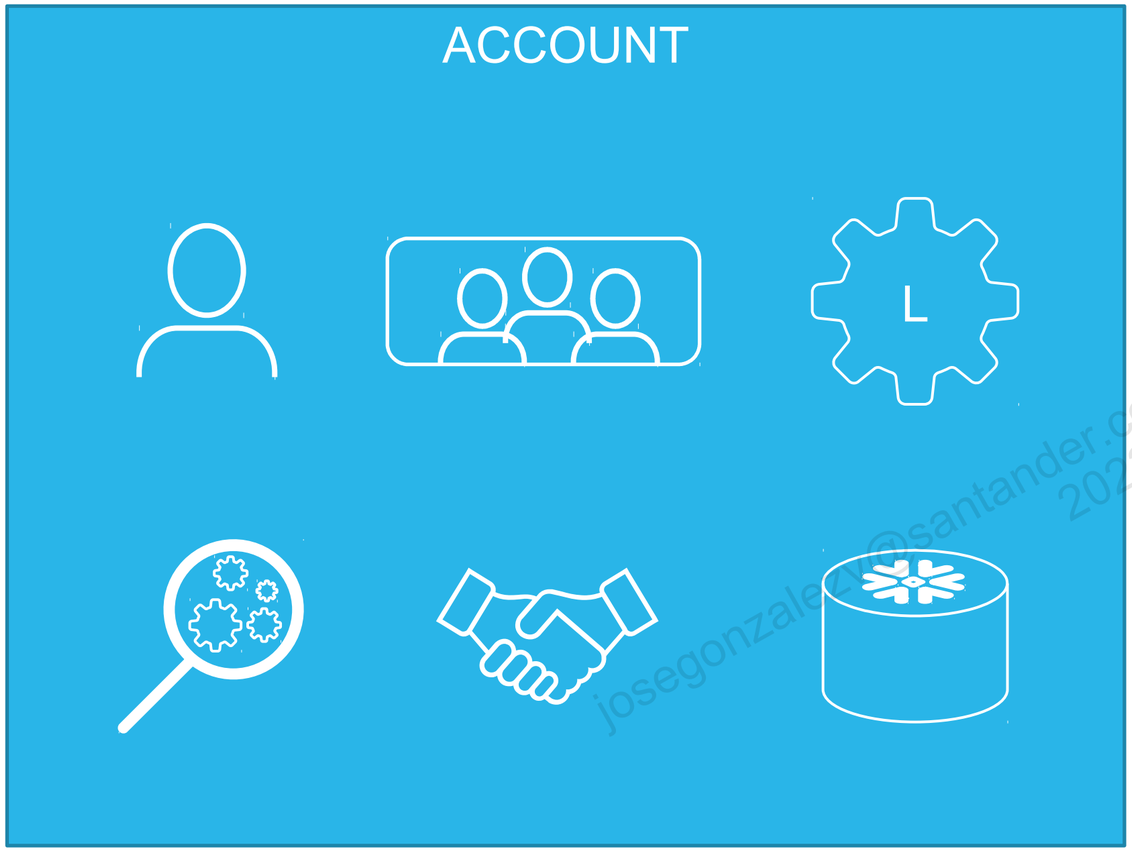
B. Compute and storage

jose@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

SNOWFLAKE STRUCTURE

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023 - do-not-copy

ACCOUNT OBJECTS

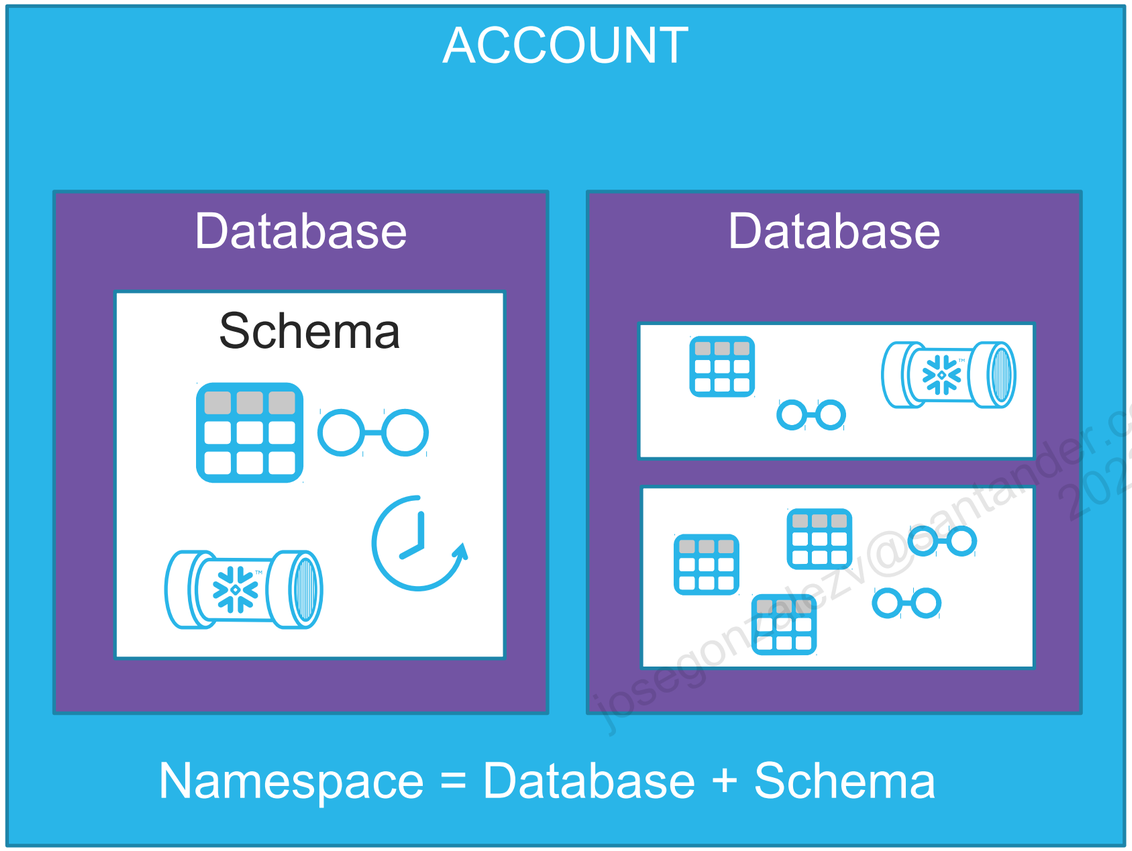


- Account-level objects include:
  - Users
  - Roles
  - Virtual warehouses
  - Resource monitors
  - Integrations
  - Databases

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

DATABASE AND SCHEMA OBJECTS

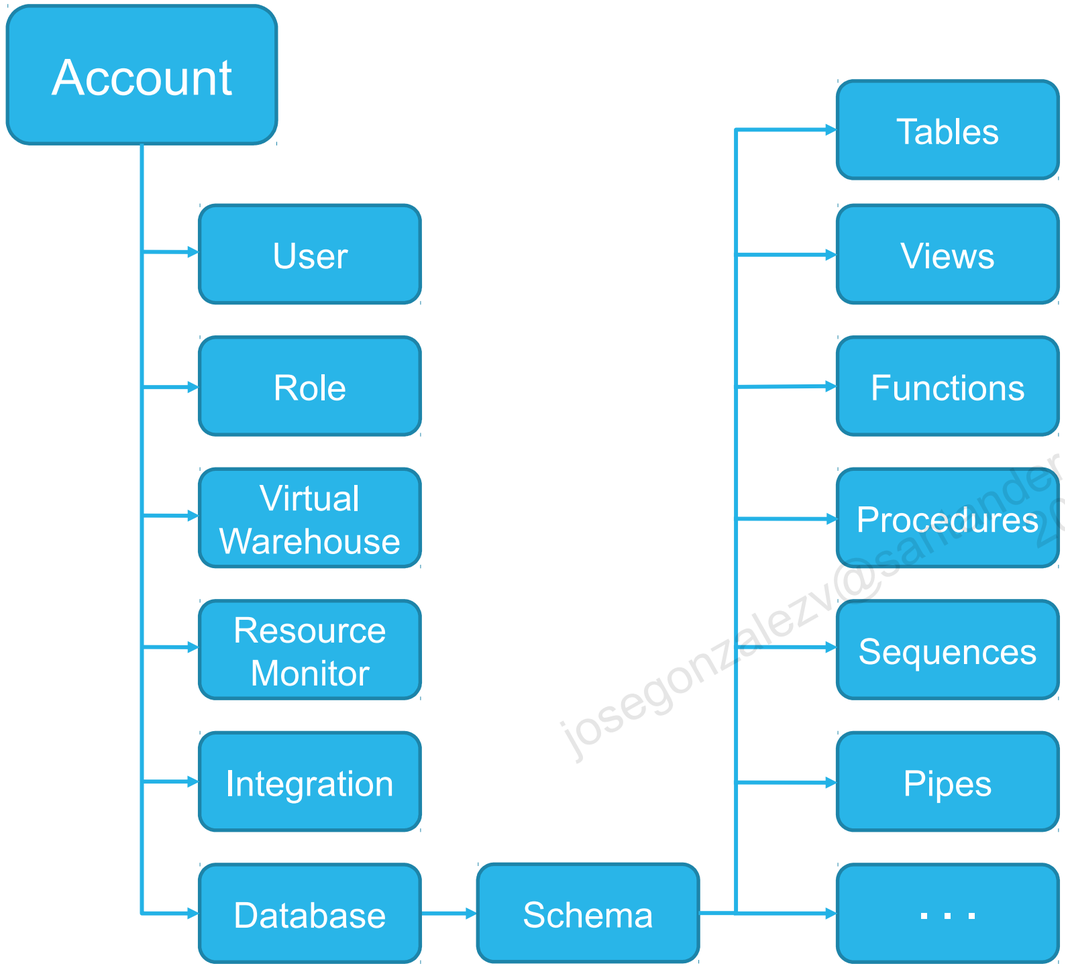


- Each Snowflake customer has one or more accounts
- An account has many databases, used to logically organize data
  - Production, Test, Development
  - Finance, Engineering, Marketing
- A database contains schemas
  - Each database belongs to a single account
  - Each schema belongs to a single database
- A schema is a logical grouping of objects (tables, views, functions, etc.)

josegonzalezv@santander.com.mx - 11-SEP-2023 @Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

OBJECT HIERARCHY

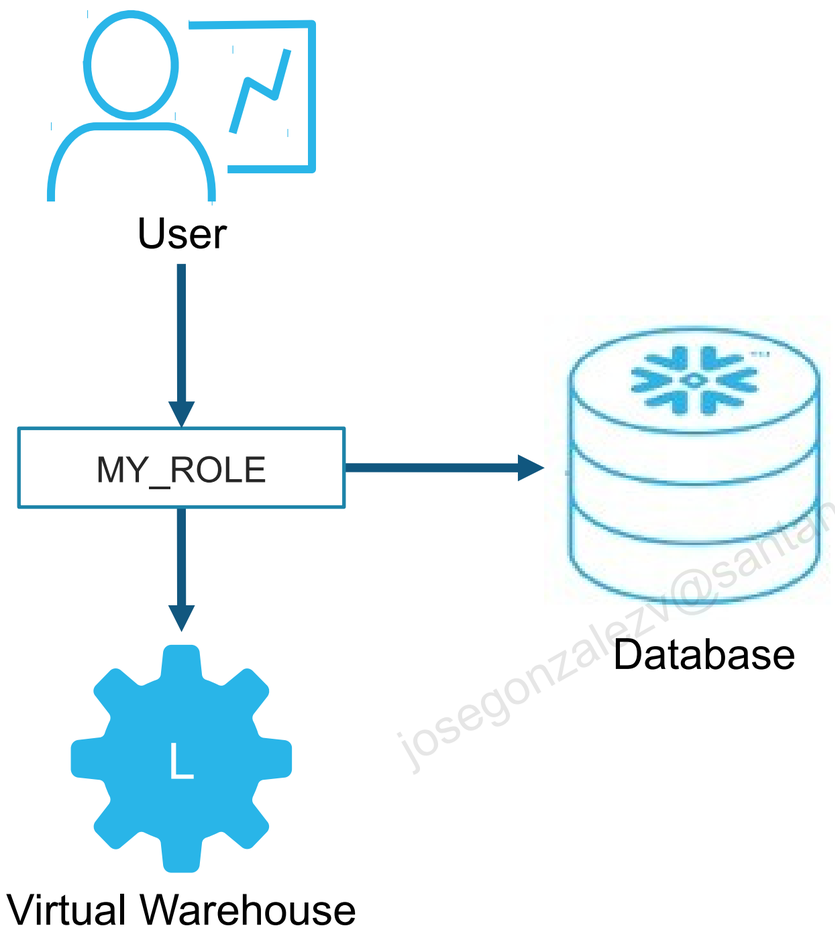


- All Snowflake objects reside within logical containers
  - Organization
  - Account
  - Database
  - Schema
- All objects are individually securable

josegonzalezv@santander.com.mx - 11-SEP-2023-do-not-copy ©Snowflake

© 2023 Snowflake Computing Inc. All Rights Reserved

USERS AND ROLES



- Every user is assigned to one or more roles
- Privileges are granted to roles
  - What databases the role can see
  - What warehouses the role can use
- The role you are using determines what you can do and see
- Set your role first!

josegonzalezv@santander.com.mx - 11-SEP-2023 @Snowflake
2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

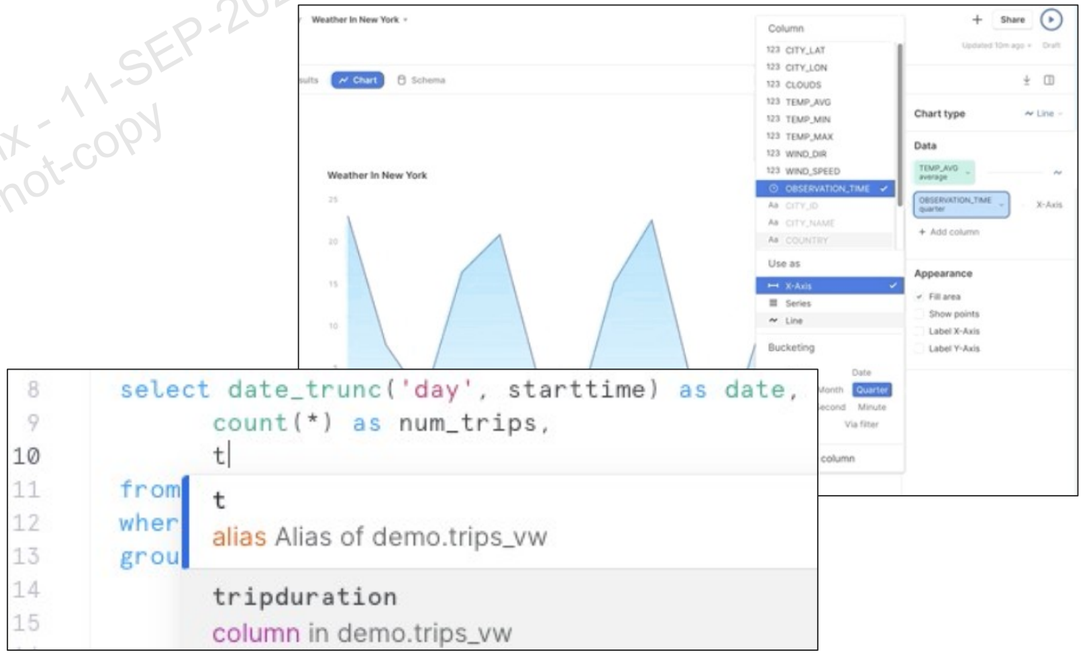
USING SNOWSIGHT

josegonzalezv@santander.com mar 11-SEP-2023 - ©Snowflake
2023-do not copy

SNOWSIGHT

ANALYST EXPERIENCE

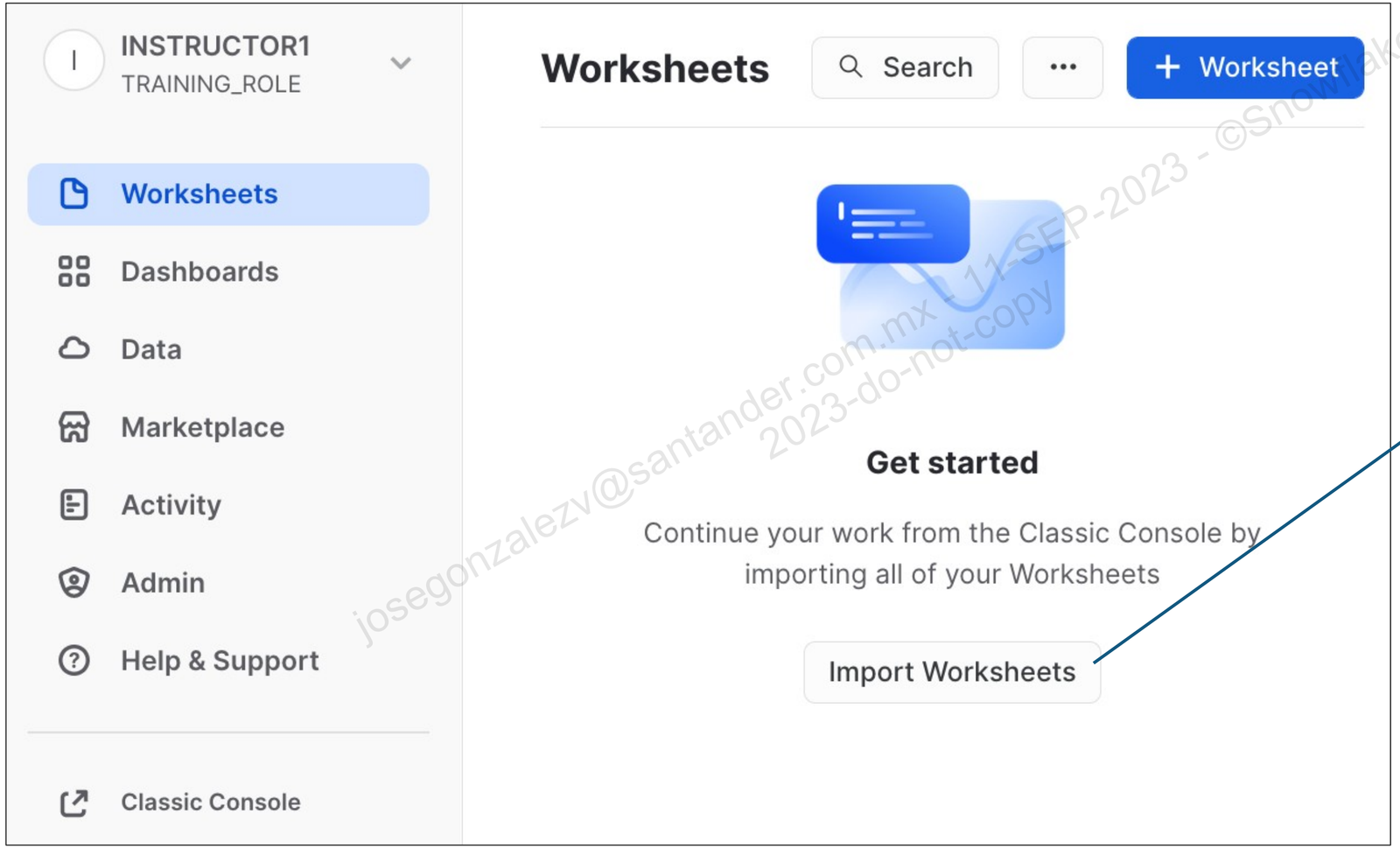
- Fast and responsive querying with autocomplete
- Instant schema search
- Interactive results exploration
- Beautiful visualizations
- Modern dashboards
- Sharing and collaboration
- **NOT** a replacement for your BI tool



josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake
2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

HOME PAGE (FIRST TIME LOGGING IN)

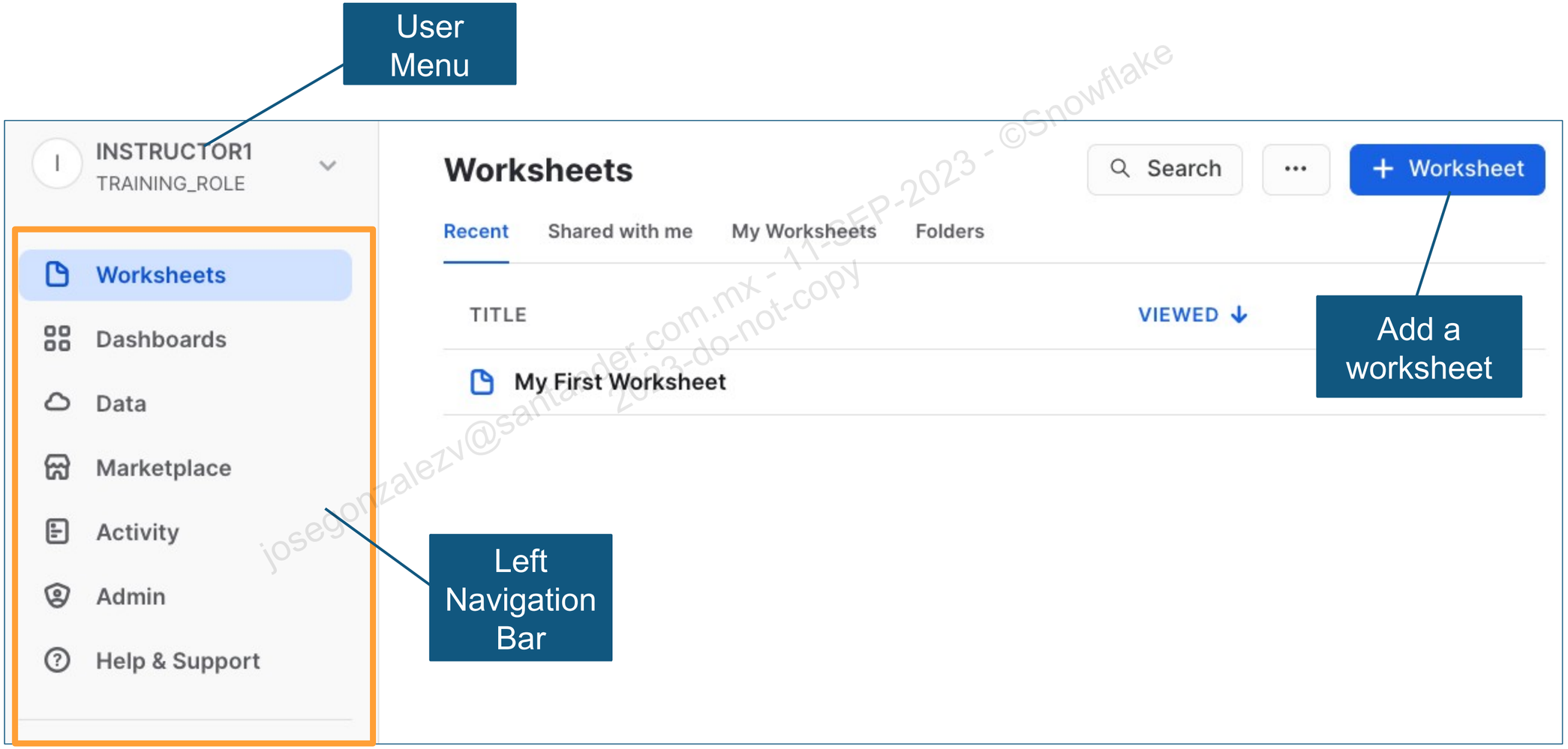


Optional:
Import worksheets from Classic UI

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

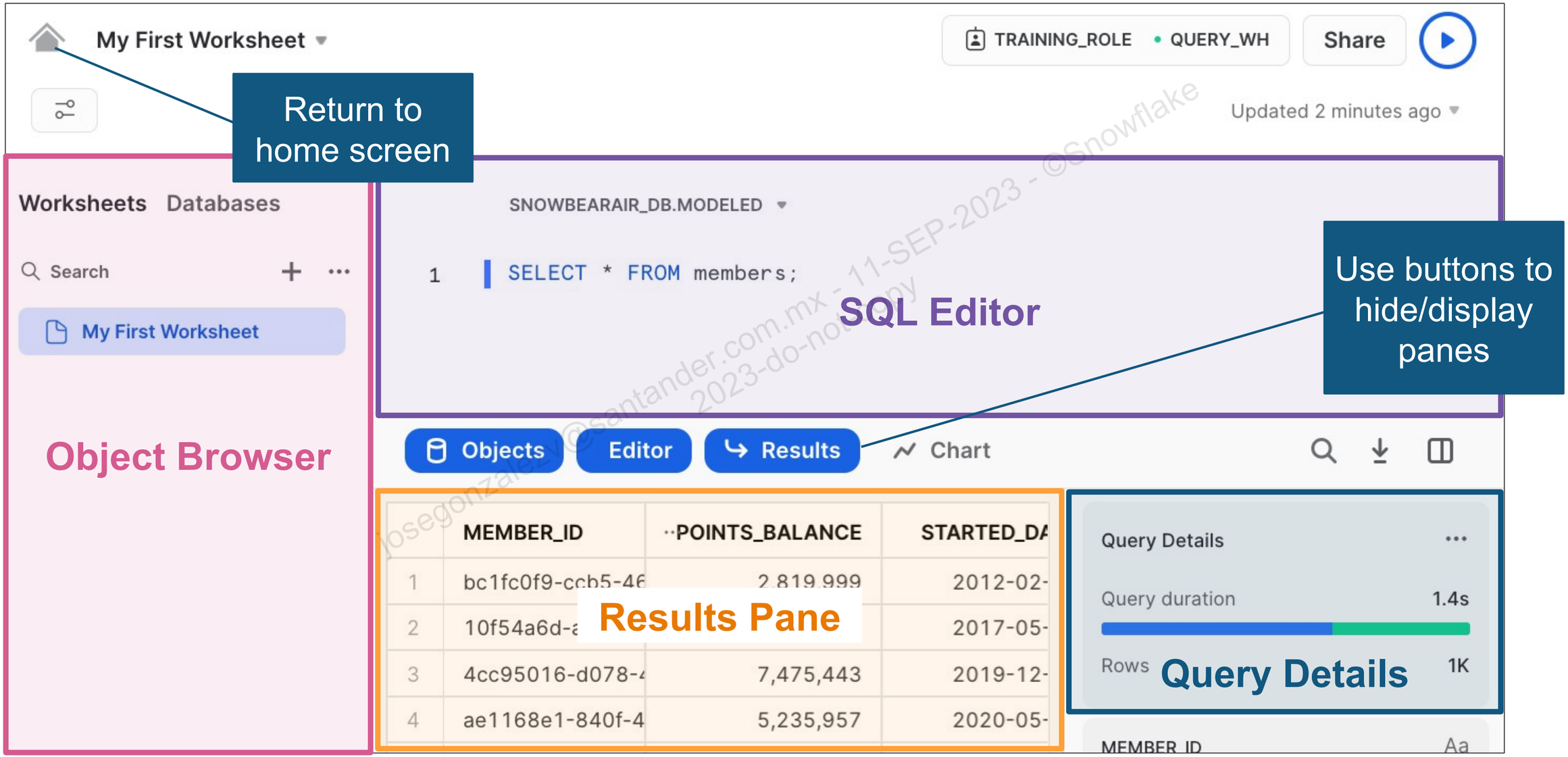
HOME PAGE (SUBSEQUENT LOGINS)



josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

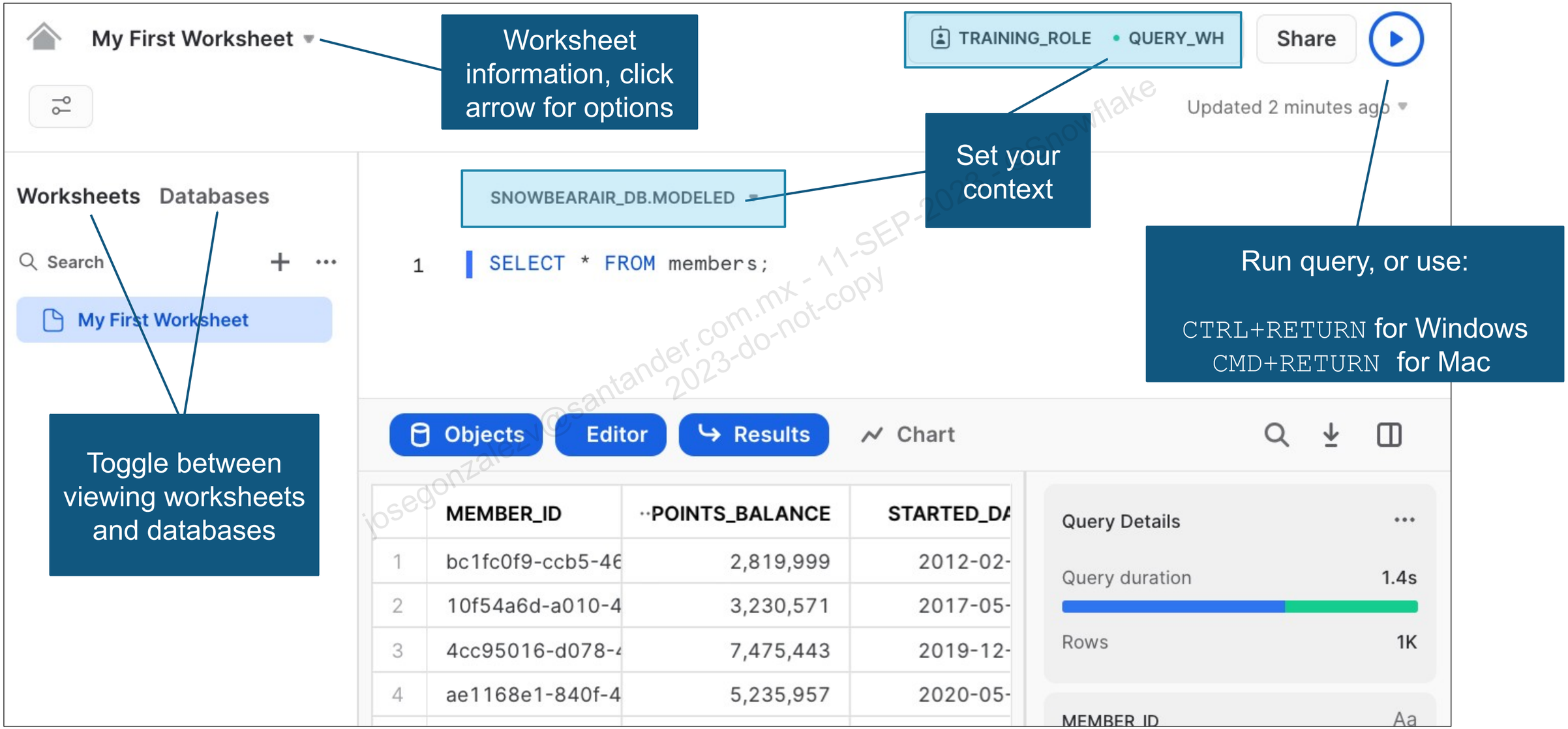
WORKSHEET PANES



josegonzalez@santander.com.mx - 11-SEP-2023 - ©Snowflake
2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

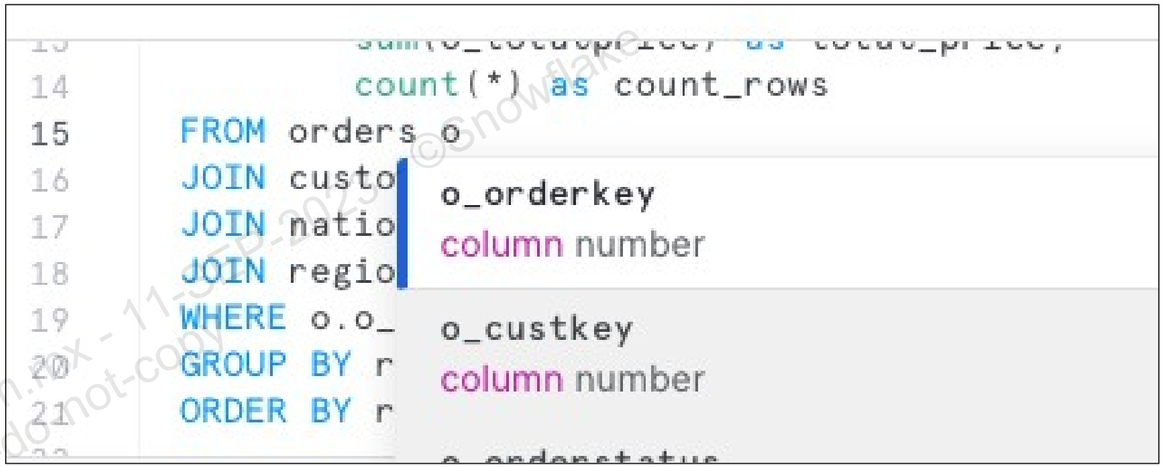
WORKSHEET CONTROLS



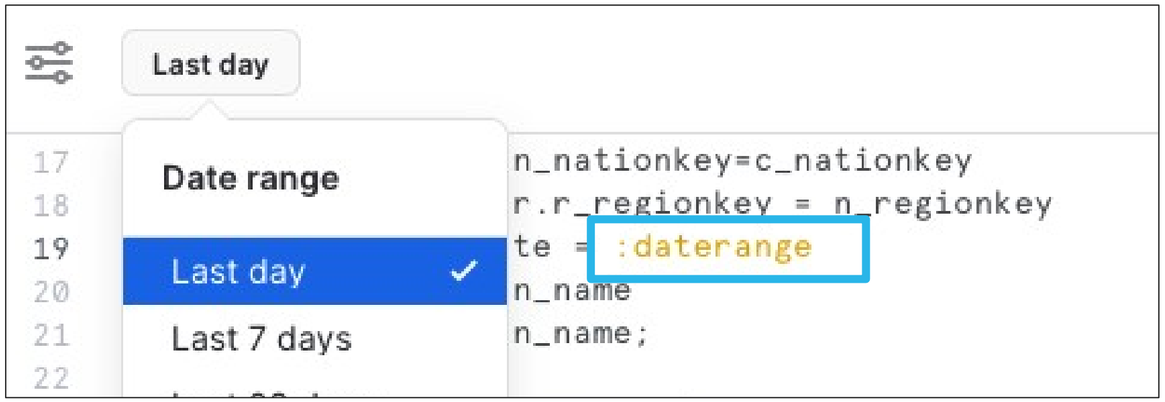
© 2023 Snowflake Computing Inc. All Rights Reserved

EDITING WORKSHEETS

- Use autocomplete (with a link to Snowflake docs) to quickly write queries



- Use pre-defined or custom filters to simplify syntax and adjust queries without editing the SQL



josegonzalezv@santander.com.mx - 11-27-FP-2023 - 2023-do-not-copy ©Snowflake

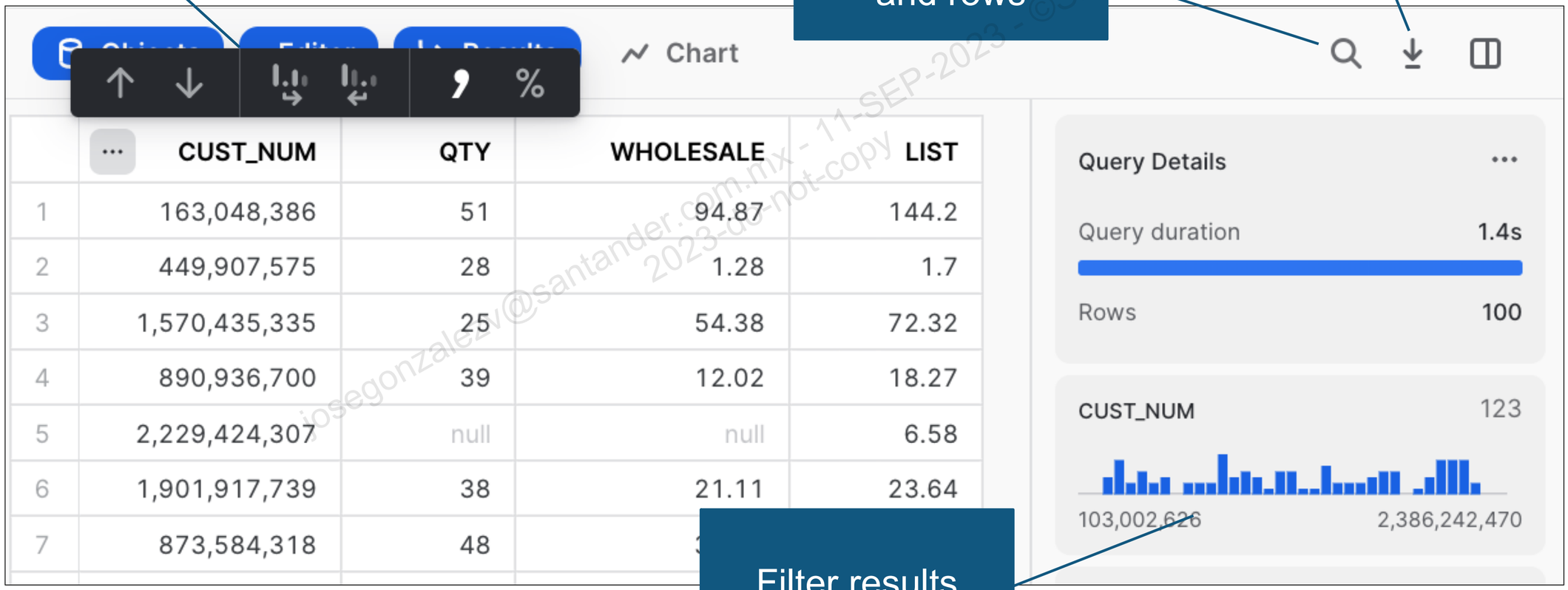
© 2023 Snowflake Computing Inc. All Rights Reserved

RESULTS

View, sort, and format results

Search columns and rows

Download results as .csv



Filter results

josegonzalez@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

CREATE CHARTS

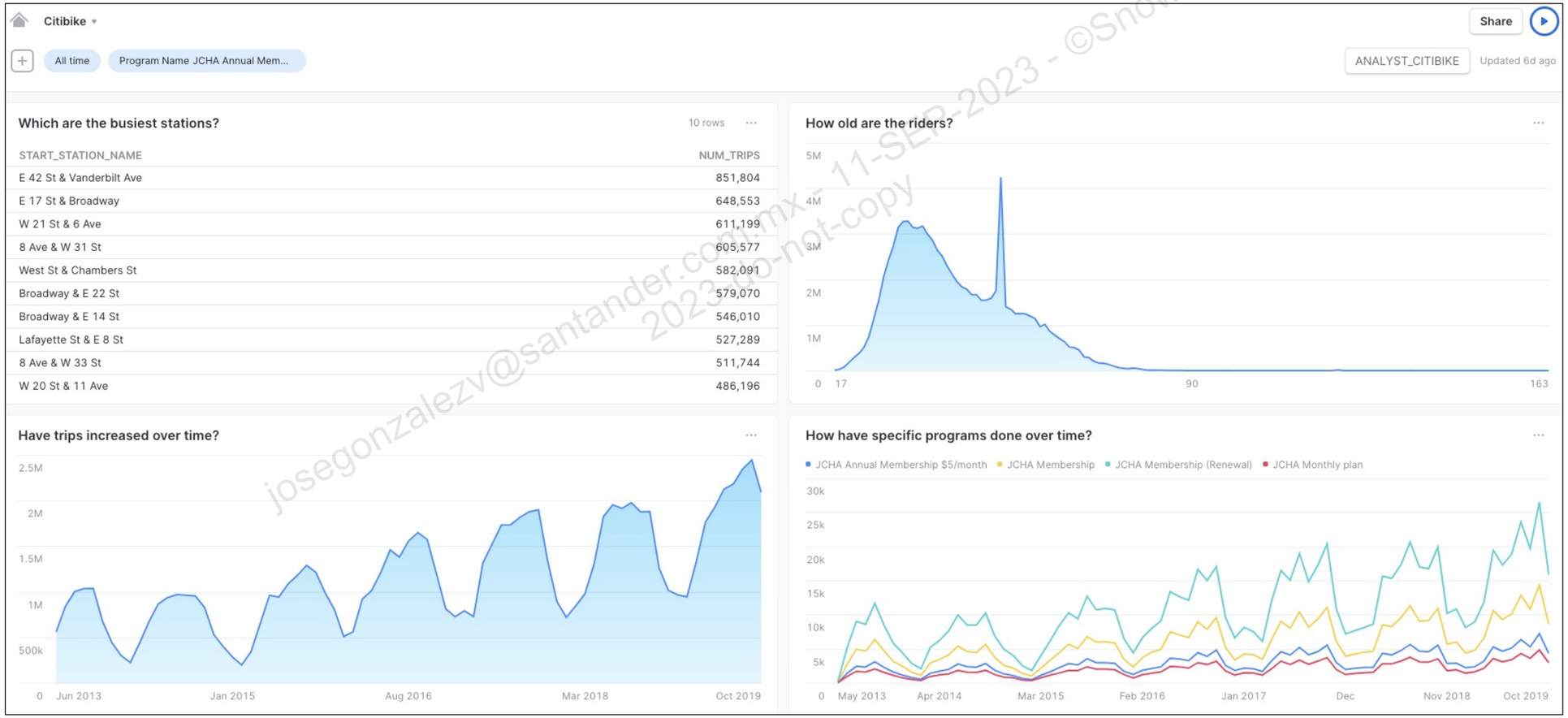
- Visualize and interact with your data
- Download charts as .png
- Choose between multiple chart types
- Configure data sources, axes, aggregations, and bucketing

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake
2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

DASHBOARDS

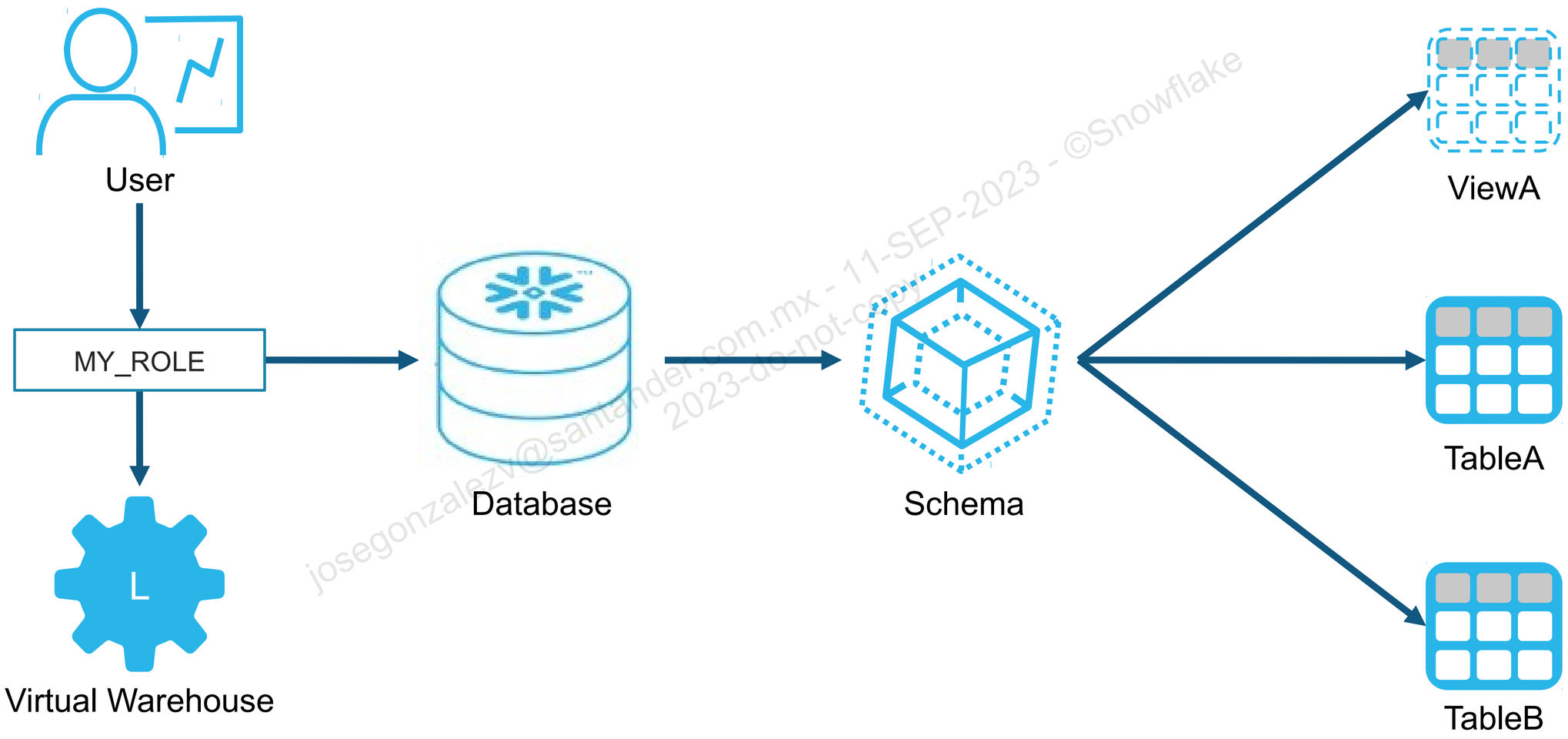
- Create and share dashboards using standard and custom filters



josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023 do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

SET YOUR CONTEXT



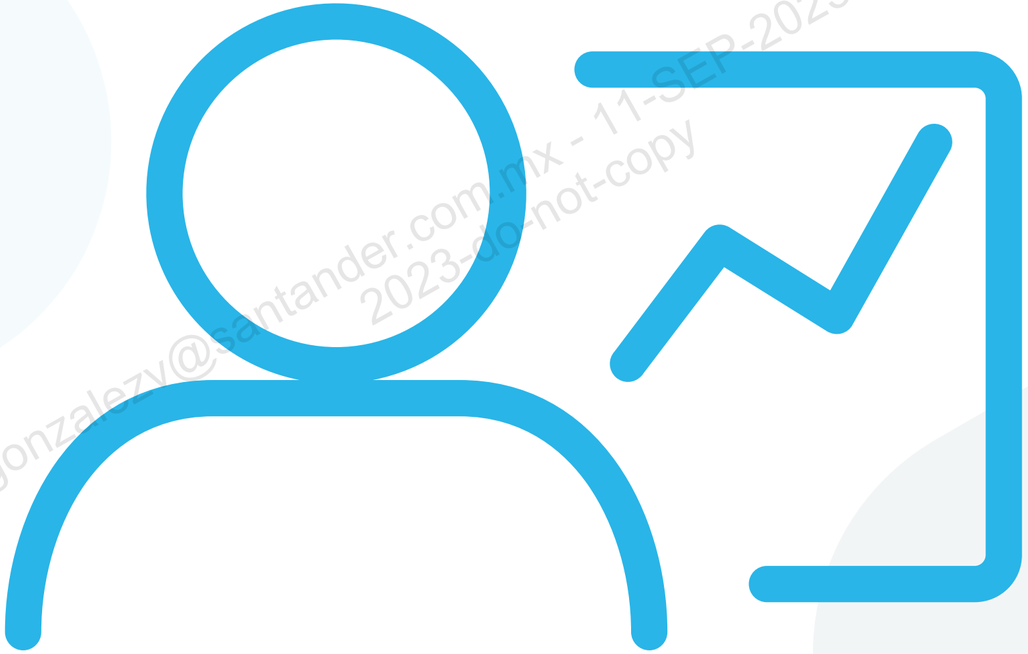
josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

INSTRUCTOR DEMO

Using Snowsight

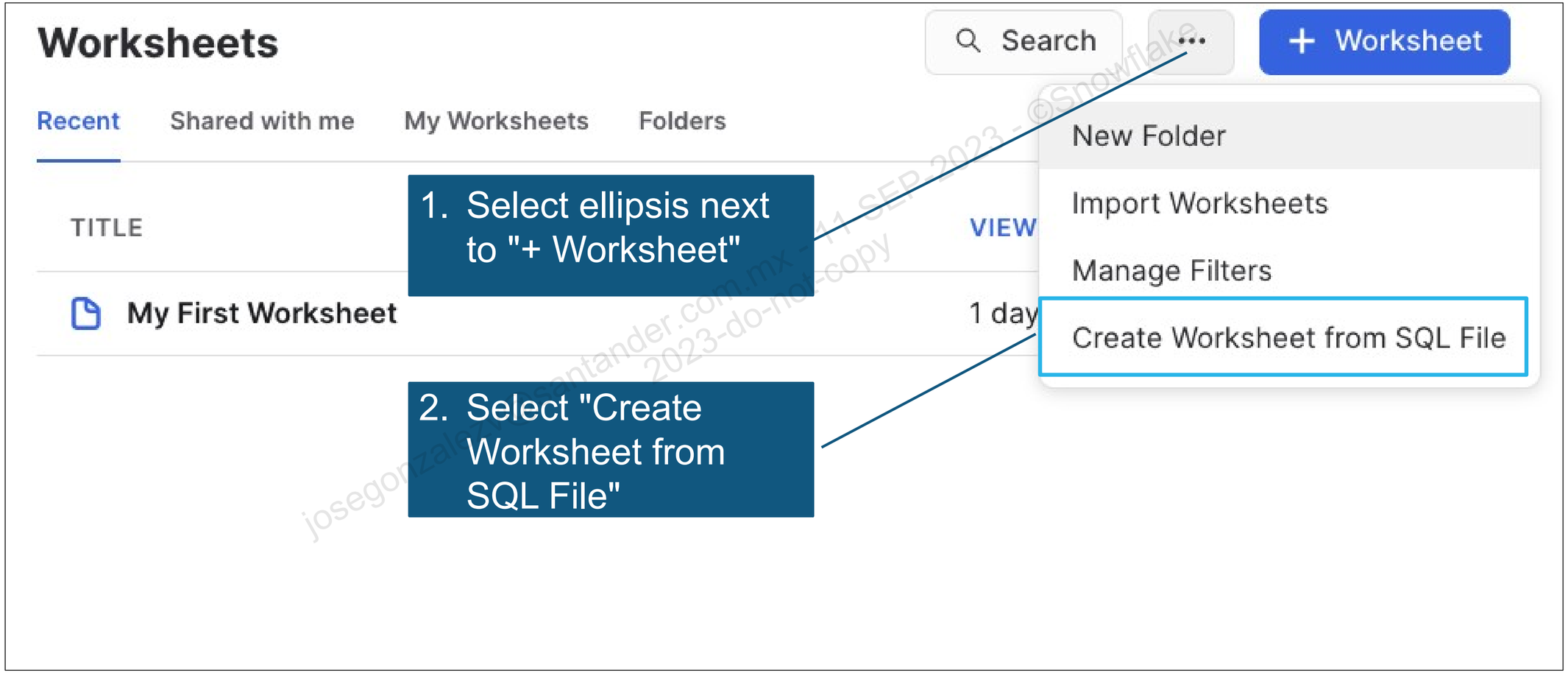
- Overview
- Worksheet Context
- Tips and Tricks



josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

LOAD SCRIPTS INTO SNOWSIGHT



josegonzalez@santander.com.mx - 11 SEP 2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

LAB EXERCISE: 1

Introduction to Snowflake and Snowflake Objects

40 minutes

Learning Objectives:

- How to navigate Snowsight to find the tools you'll need
- How to create and manage folders worksheets
- How to set the context
- How to create warehouses, databases, schemas and tables
- How to run a simple query

Note:

- **Internet Explorer is not supported!**

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023 do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

KNOWLEDGE CHECK

Put these objects in order from outer to inner position:

Account
Database
Organization
Table
Schema
View

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

KNOWLEDGE CHECK

Put these objects in order from outer to inner position:

Account
Database
Organization
Table
Schema
View

Organization

Account

Database

Schema

Table and view are at the same level!

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake
2023-do-not-copy

KNOWLEDGE CHECK

What elements make up your context?

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

KNOWLEDGE CHECK

What elements make up your context?

Database

Schema

Role

Warehouse

josegonzalezv@santander.com.mx - 11-SEP-2023 @Snowflake
2023-do-not-copy

KNOWLEDGE CHECK

What does your default context do?

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

KNOWLEDGE CHECK

What does your default context do?

It automatically sets the context for every new worksheet you create.

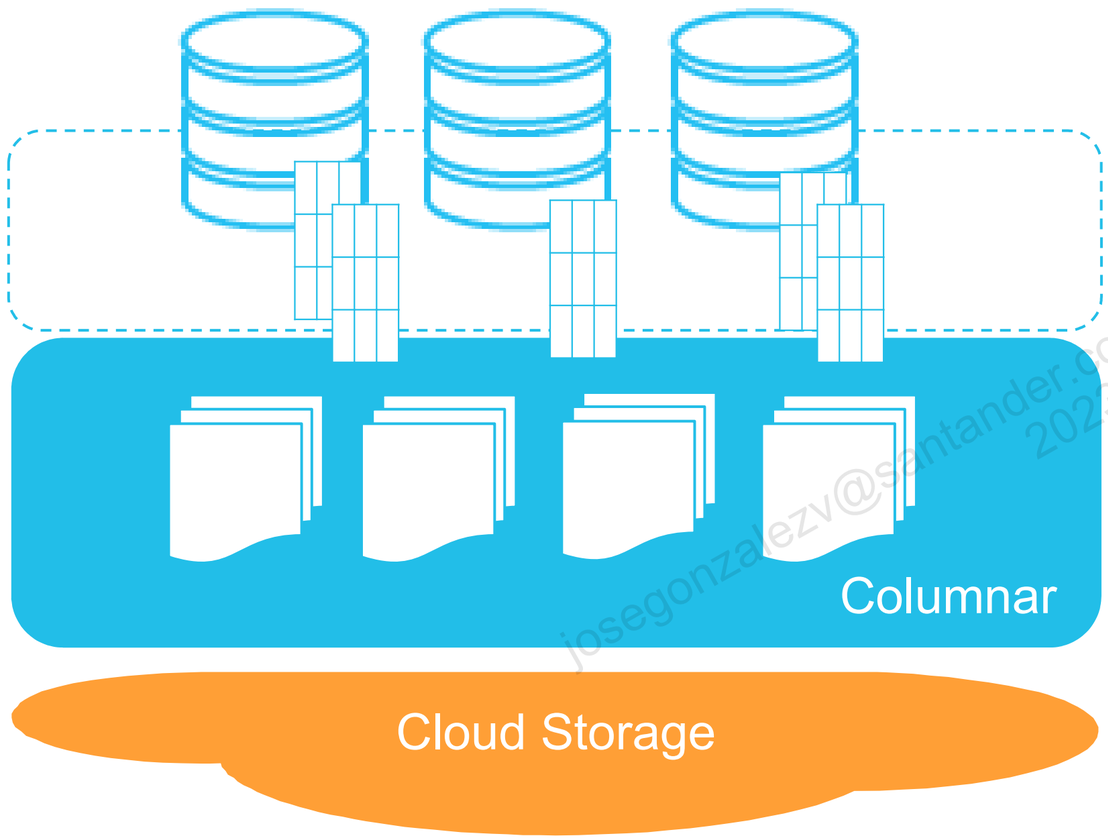
Once in a worksheet, you can change any of your context elements – using either the USE command, or the drop-down menus for each context item.

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake
2023-do-not-copy

STORAGE LAYER

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

STORAGE LAYER

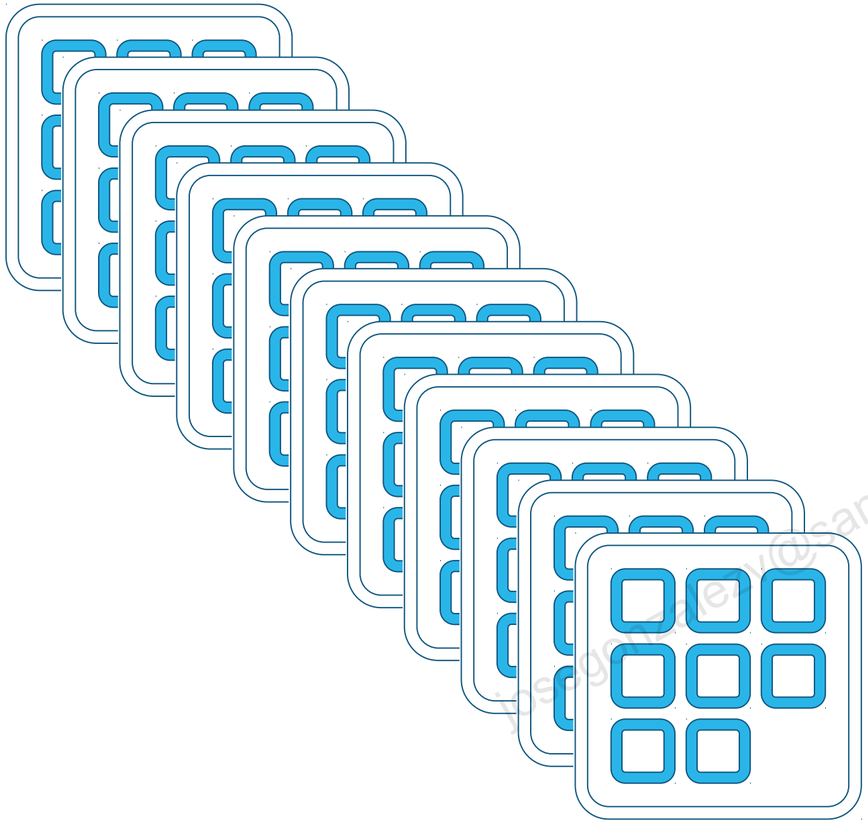


- Hybrid-Columnar storage (sets of rows stored in columns)
- Automatic micro-partitioning
- Semi-Structured data support and optimization

josegonzalezv@santander.com.mx - 11-SEP-2023 - @Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

MICRO-PARTITIONS



- Contiguous units of storage that hold table data
  - Generally 10MB to 16MB (Compressed)
- Many micro-partitions per table
- IMMUTABLE !!!!!
- Services layer stores metadata about every micro-partition
  - MIN/MAX (Range of values in each column)
  - Number of distinct values
  - NULL count

issgonzalez@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

MICRO-PARTITIONING

- As data is loaded, Snowflake preserves the natural order of the data

Input Files

MONTH	LNAME	ORDER_TOT
Jan	Wiggins	1034.22
Jan	Brown	99.95
Jan	Nguyen	2044.56
Feb	Belehrad	100.34
Feb	Wilson	2134.45
Feb	Lopez	3308.24
Mar	Cooper	34.98
Mar	Zelen	1566.90
Mar	Aziz	804.28

jose.gonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

MICRO-PARTITIONING

- As data is loaded, Snowflake preserves the natural order of the data
  - Data collected by date will be naturally ordered by date

Input Files

MONTH	LNAME	ORDER_TOT
Jan	Wiggins	1034.22
Jan	Brown	99.95
Jan	Nguyen	2044.56
Feb	Belehrad	100.34
Feb	Wilson	2134.45
Feb	Lopez	3308.24
Mar	Cooper	34.98
Mar	Zelen	1566.90
Mar	Aziz	804.28

Data stored in micro-partitions, with column data together

```
MONTH: Jan Jan Jan
LNAME: Wiggins Brown Nguyen
TOTAL: 1034.22 99.95 2044.56

MONTH: Feb Feb Feb
LNAME: Belehrad Wilson Lopez
TOTAL: 100.34 2134.45 3308.24

MONTH: Mar Mar Mar
LNAME: Cooper Zelen Aziz
TOTAL: 34.98 1566.90 804.28
```

josegonzalez@santander.com.mx - 11 SEP 2023 - ©Snowflake
2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

MICRO-PARTITIONING

- We collect metadata on each micro-partition that can reduce scan times

Input Files

MONTH	LNAME	TOTAL
Jan	Wiggins	1034.22
Jan	Brown	99.95
Jan	Nguyen	2044.56
Feb	Belehrad	100.34
Feb	Wilson	2134.45
Feb	Lopez	3308.24
Mar	Cooper	34.98
Mar	Zelen	1566.90
Mar	Aziz	804.28

Data stored in micro-partitions, with column data together

```
MONTH:  Jan Jan Jan
LNAME:  Wiggins Brown Nguyen
TOTAL:  1034.22 99.95 2044.56

MONTH:  Feb Feb Feb
LNAME:  Belehrad Wilson Lopez
TOTAL:  100.34 2134.45 3308.24

MONTH:  Mar Mar Mar
LNAME:  Cooper Zelen Aziz
TOTAL:  34.98 1566.90 804.28
```

Metadata

Min	Max
Jan	Jan
Brown	Wiggins
99.95	2044.56

Min	Max
Feb	Feb
Belehrad	Wilson
100.34	3308.24

Min	Max
Mar	Mar
Aziz	Zelen
34.98	1566.90

josegonzalezv@sanu.edu.mx - 11-SEP-2023 - ©Snowflake 2023-do not copy

© 2023 Snowflake Computing Inc. All Rights Reserved

MICRO-PARTITION ELIMINATION (PRUNING)

```
SELECT * FROM sales WHERE month = 'May';
```

Min	Max	Min	Max	Min	Max	Min	Max
Jan	Jan	Apr	Apr	Jul	Jul	Oct	Oct
Brown	Wiggins	Clark	Patel	Adams	Wang	Adamos	Yamamoto
99.95	2044.56	103.22	2204.33	1103.22	21990.23	35.12	10335.22
Feb	Feb	May	May	Aug	Aug	Nov	Nov
Belehrad	Wilson	Bigliani	Tessay	Byrne	Visser	Barac	Shevchenko
100.34	3308.24	99.25	14002.34	104.45	2210.34	228.34	1992.45
Mar	Mar	Jun	Jun	Sep	Sep	Dec	Dec
Aziz	Zelen	Blanchet	Volkov	Accola	Zhang	Castillo	Palakiko
34.98	1566.90	9.92	1002.34	12.34	998.22	436.22	1099.45

josegonzalezv@santander.com.mx 17-SEP-2023 ©Snowflake
2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

MICRO-PARTITION ELIMINATION (PRUNING)

```
SELECT * FROM sales WHERE month = 'May';
```

Min	Max	Min	Max	Min	Max	Min	Max
Jan	Jan	Apr	Apr	Jul	Jul	Oct	Oct
Brown	Wiggins	Clark	Patel	Adams	Wang	Adamos	Yamamoto
99.95	2044.56	103.22	2204.33	1103.22	21990.23	35.12	10335.22
Feb	Feb	May	May	Aug	Aug	Nov	Nov
Belehrad	Wilson	Bigliani	Tessay	Byrne	Visser	Barac	Shevchenko
100.34	3308.24	99.25	14002.34	104.45	2210.34	228.34	1992.45
Mar	Mar	Jun	Jun	Sep	Sep	Dec	Dec
Aziz	Zelen	Blanchet	Volkov	Accola	Zhang	Castillo	Palakiko
34.98	1566.90	9.92	1002.34	12.34	998.22	436.22	1099.45

josegonzalezv@santander.com.mx 17-SEP-2023 ©Snowflake
2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

MICRO-PARTITION ELIMINATION (PRUNING)

```
SELECT * FROM sales WHERE lname = 'Thompson';
```

Min	Max	Min	Max	Min	Max	Min	Max
Jan	Jan	Apr	Apr	Jul	Jul	Oct	Oct
Brown	Wiggins	Clark	Patel	Adams	Wang	Adamos	Yamamoto
99.95	2044.56	103.22	2204.33	1103.22	21990.23	35.12	10335.22
Feb	Feb	May	May	Aug	Aug	Nov	Nov
Belehrad	Wilson	Bigliani	Tessay	Byrne	Visser	Barac	Shevchenko
100.34	3308.24	99.25	14002.34	104.45	2210.34	228.34	1992.45
Mar	Mar	Jun	Jun	Sep	Sep	Dec	Dec
Aziz	Zelen	Blanchet	Volkov	Accola	Zhang	Castillo	Palakiko
34.98	1566.90	9.92	1002.34	12.34	998.22	436.22	1099.45

josegonzalezv@santander.com.mx 17-SEP-2023 ©Snowflake
2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

MICRO-PARTITION ELIMINATION (PRUNING)

```
SELECT * FROM sales WHERE lname = 'Thompson';
```

Min	Max	Min	Max	Min	Max	Min	Max
Jan	Jan	Apr	Apr	Jul	Jul	Oct	Oct
Brown	Wiggins	Clark	Patel	Adams	Wang	Adamos	Yamamoto
99.95	2044.56	103.22	2204.33	1103.22	21990.23	35.12	10335.22
Feb	Feb	May	May	Aug	Aug	Nov	Nov
Belehrad	Wilson	Bigliani	Tessay	Byrne	Visser	Barac	Shevchenko
100.34	3308.24	99.25	14002.34	104.45	2210.34	228.34	1992.45
Mar	Mar	Jun	Jun	Sep	Sep	Dec	Dec
Aziz	Zelen	Blanchet	Volkov	Accola	Zhang	Castillo	Palakiko
34.98	1566.90	9.92	1002.34	12.34	998.22	436.22	1099.45

josegonzalezv@santander.com.mx 11-SEP-2023 ©Snowflake
2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

COLUMNAR FORMAT BENEFITS

- Columnar format improves read times and compressibility

We can jump directly to any column offset, to read columns without scanning all the data

Data stored in micro-partitions, in columnar format

```
MONTH: Jan Jan Jan
LNAME: Wiggins Brown Nguyen
TOTAL: 1034.22 99.95 2044.56

MONTH: Feb Feb Feb
LNAME: Belehrad Wilson Lopez
TOTAL: 100.34 2134.45 3308.24

MONTH: Mar Mar Mar
LNAME: Cooper Zelen Aziz
TOTAL: 34.98 1566.90 804.28
```

Columnar storage increases compressibility

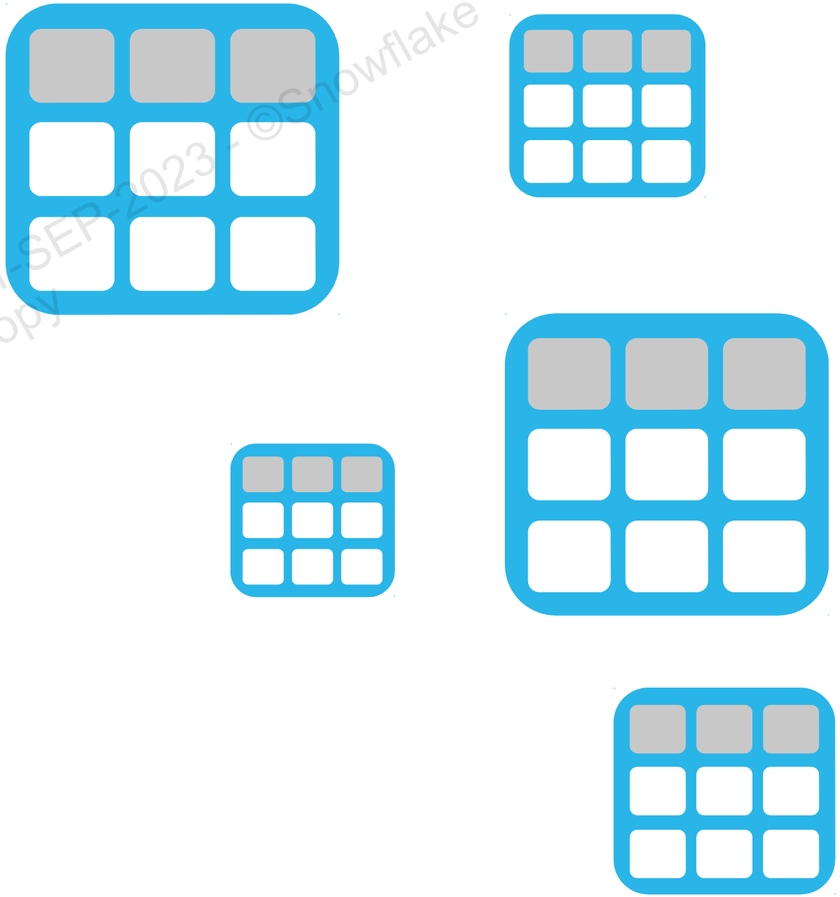
"Feb Feb Feb" compresses better than "Feb Belehrad 100.34"

josegonzalezv@sandbox.com.mx - 11 SEP 2023 - ©Snowflake
2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

DATA STORAGE BILLING

- Billed for actual storage use
  - Daily average Terabytes per month
- On-demand pricing
  - Billed in arrears for storage used
  - Minimum monthly charge of $25
  - Example for AWS in Singapore: $46/TB/month
- Pre-Purchased Capacity
  - Billed up front with commitment to a certain capacity
  - Price varies on amount and cloud platform
  - Customer is notified at 70% of capacity
  - Example for AWS in Singapore: $25/TB/month



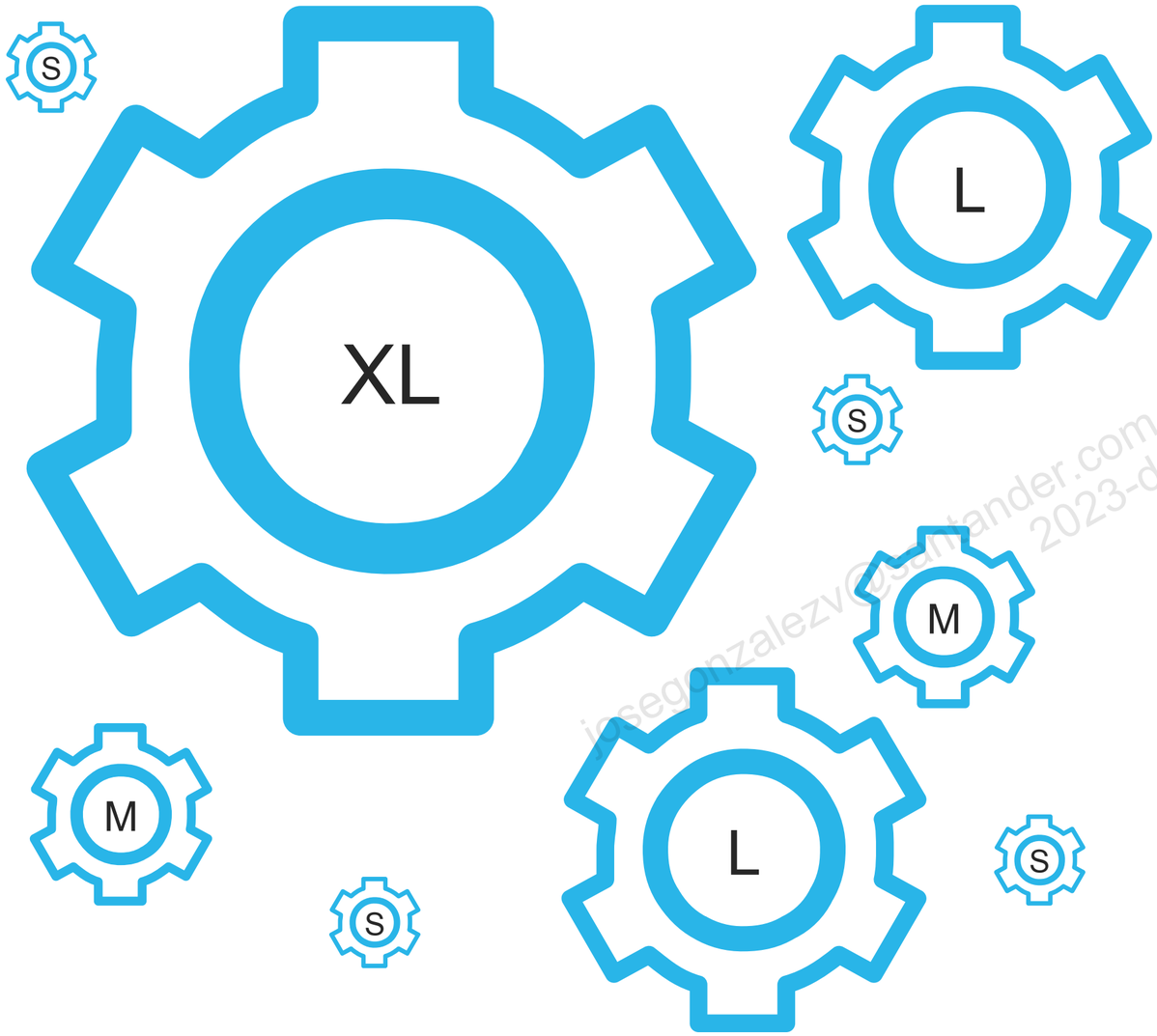
josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

COMPUTE LAYER

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake
2023 do-not-copy

VIRTUAL WAREHOUSES

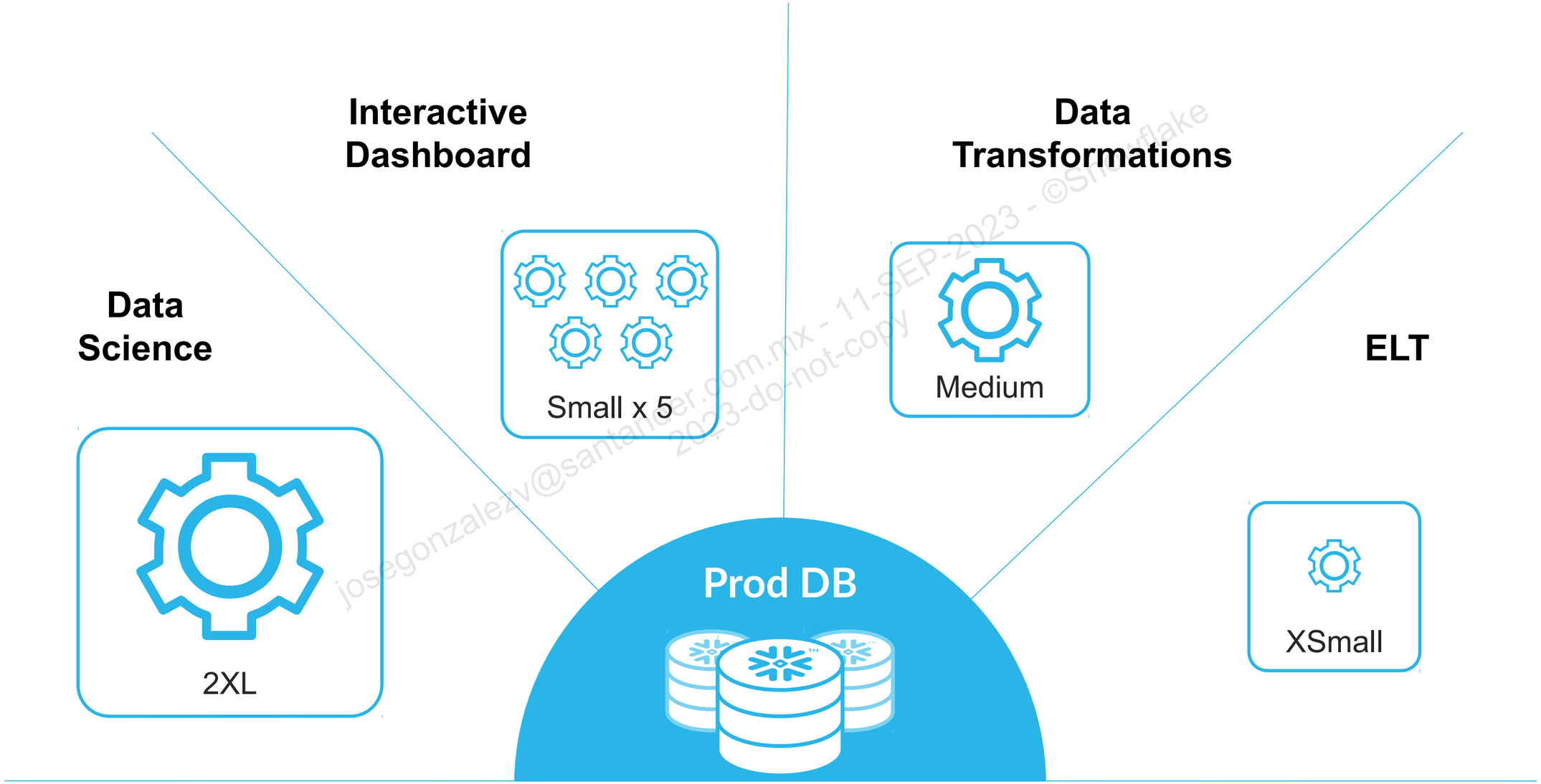


- Clusters of compute resources with CPU, memory, and disk
- Compute resources double with each size increase
- Jobs requiring compute run on virtual warehouses
- While running, a virtual warehouse consumes Snowflake credits
  - How you are charged for compute

jose.gonzalezv@santander.com.mx - 11-SEP-2023 ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

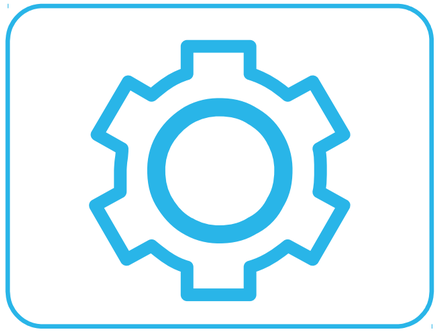
WORKLOAD SEGMENTATION



josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

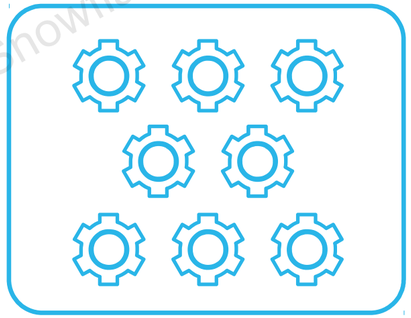
© 2023 Snowflake Computing Inc. All Rights Reserved

VIRTUAL WAREHOUSE TYPES



Standard

- Always a single cluster
- Scale up to a larger virtual warehouse for more compute



Multi-Cluster Warehouse (MCW)

- One or more clusters
- Automatically scaled by Snowflake to match traffic (for high concurrency)
- Enterprise Edition feature

jose.gonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

COMPUTE CREDITS

Warehouse Size	Credits / Hour	Credits / Second
XS	1	0.0003
S	2	0.0006
M	4	0.0011
L	8	0.0022
XL	16	0.0044
2XL	32	0.0089
3XL	64	0.0178
4XL	128	0.0356
5XL	256	0.0711
6XL	512	0.1422

- How you are billed for compute usage
- Charged by the number of running clusters, their sizes, and how long they are running
- Compute is charged per-second, with a one-minute minimum
- 5XL and 6XL virtual are generally available in Amazon Web Services regions, and in preview in US Government and Azure regions

josegonzalezv@santander.com.mx - 11-SEP-2023 @Snowflake
2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

SCALING UP VS. SCALING OUT

CREDITS PER HOUR

NUMBER OF CLUSTERS →

VIRTUAL WAREHOUSE SIZE ↑

	1	2	3	4	5	6	7	8	9	10
4XL	128	256	384	512	640	768	896	1024	1152	1280
3XL	64	128	192	256	320	384	448	512	576	640
2XL	32	64	96	128	160	192	224	256	288	320
XL	16	32	48	64	80	96	112	128	144	160
L	8	16	24	32	40	48	56	64	72	80
M	4	8	12	16	20	24	28	32	36	40
S	2	4	6	8	10	12	14	16	18	20
XS	1	2	3	4	5	6	7	8	9	10

XL x1 = 16 credits per hour
Scale up for complexity

M x4 = 16 credits per hour
Scale out for concurrency

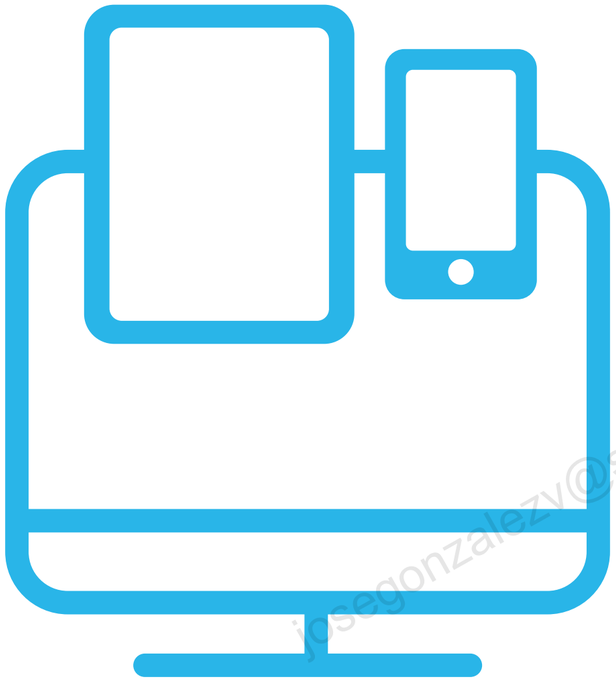
josegonzalezv@santander.com.mx - 14-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

CLOUD SERVICES LAYER

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023 do-not-copy

MANAGEMENT

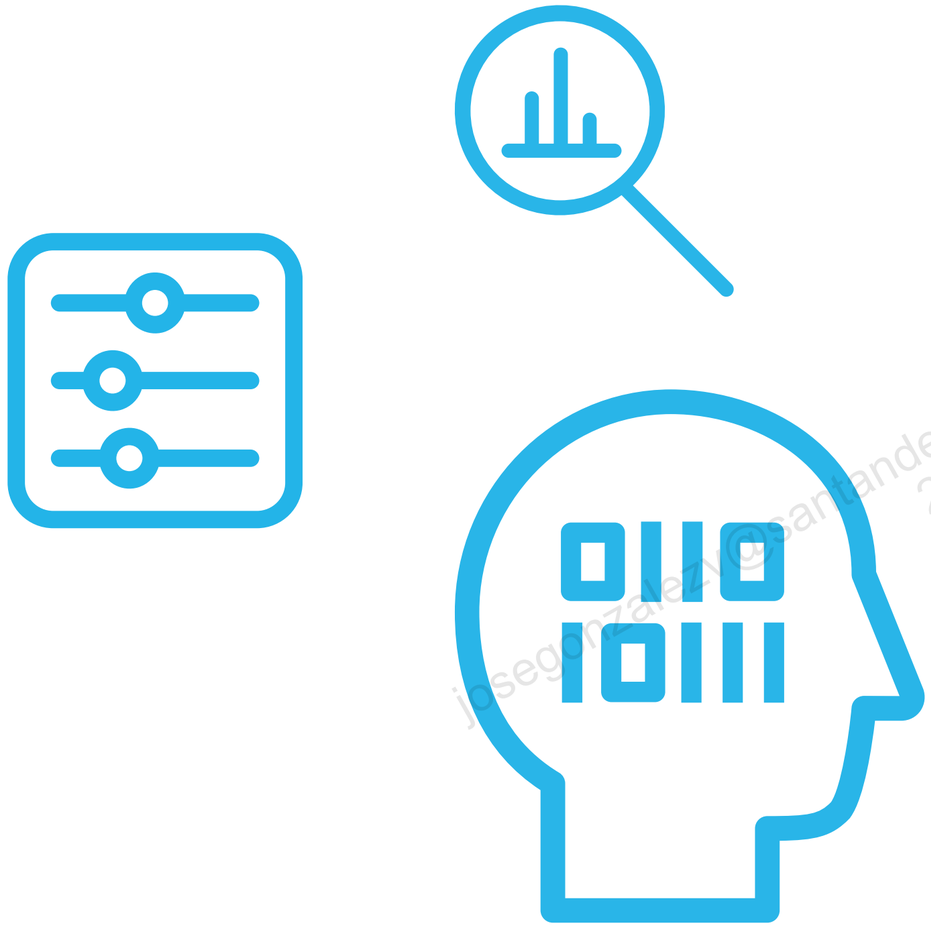


- Centralized management for all storage
- Manages the compute that works with the storage
- Transparent, handling online updates and patches

josegonzalezv@santander.com.mx - 11-SEP-2023 - @Snowflake
2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

QUERY HANDLING



- SQL Optimizer
  - Cost-based optimization (CBO)
- Automatic JOIN order optimization
  - No user input or tuning required
- Automatic statistics gathering
- Pruning using metadata about micro-partitions

jose.gonzalez@santander.com.mx - 11-SEP-2023 @Snowflake
2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

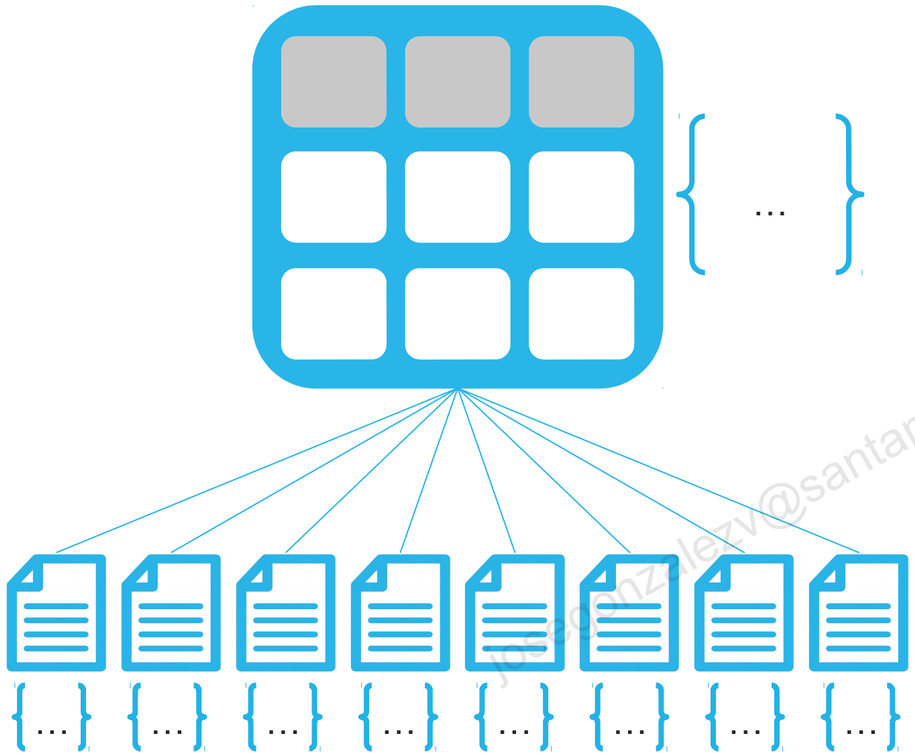
SECURITY

- Authentication
- Authorization
- Share management and security
- Encryption and key management

josegonzalezv@santander.com.mx - 11-SEP-2023 - @snowflake
2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

METADATA MANAGEMENT



- Stores metadata as data is loaded into the system
- Handles queries that can be processed completely from metadata or from the query result cache
- Almost every aspect of Snowflake architecture leverages metadata, including:
  - Time travel
  - Cloning
  - Replication

jsegonzalezv@santander.com.mx - 11-SEP-2023 @Snowflake
2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

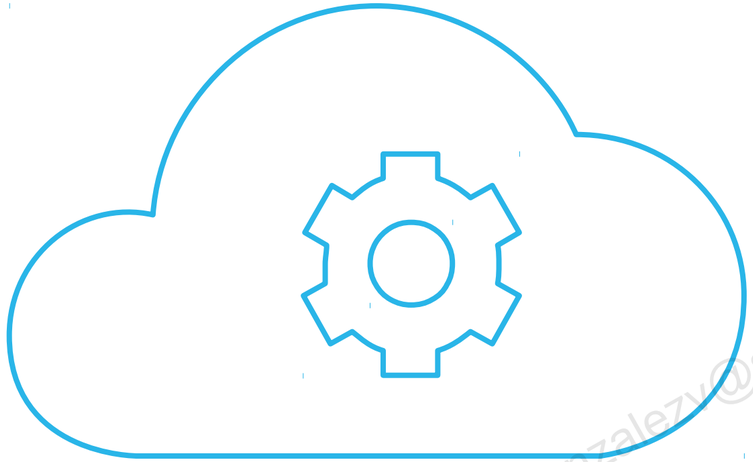
SERVERLESS FEATURES COMPUTE

- Some features have compute allocated by Snowflake, rather than running a virtual warehouse
- Charged per second
  - Cost/credit usage varies by feature. Some may include overhead charges.
  - Example: Snowpipe adds .06 Snowflake Credits per 1000 Files processed
- Used for:
  - Materialized view maintenance
  - Database replication
  - Snowpipe
  - Others

joseg.gonzalezv@santander.com.mx - 01-SEP-2023 - ©Snowflake 2023-do not copy

© 2023 Snowflake Computing Inc. All Rights Reserved

CLOUD SERVICES COMPUTE



- Short amounts of compute used for object creation and queries
  - Cache lookups
  - Query optimizer
  - Object creation or deletion
- Customers are charged for cloud services compute that exceeds 10% of the total daily compute used from all sources
- Provided by the cloud services layer, not the compute layer

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

KNOWLEDGE CHECK

What are the three layers of the Snowflake architecture?

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

KNOWLEDGE CHECK

What are the three layers of the Snowflake architecture?

Storage

Compute

Cloud Services

josegonzalezv@santander.com.mx - 11-SEP-2023 ©Snowflake
2023-do-not-copy

KNOWLEDGE CHECK

What are two kinds of virtual warehouses?

A. Basic
B. Standard
C. Concurrent
D. Multi-cluster
E. Auto

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

KNOWLEDGE CHECK

What are two kinds of virtual warehouses?

A. Basic
B. Standard
C. Concurrent
D. Multi-cluster
E. Auto

B. Standard

D. Multi-cluster

josegonzalezv@santander.com.mx - 11-SEP-2023 @Snowflake
2023-do-not-copy

KNOWLEDGE CHECK

What is a micro-partition?

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

KNOWLEDGE CHECK

What is a micro-partition?

A small file (generally ~ 16MB) that holds some number of rows in your table.

A single row is always stored completely in one micro-partition; we will never break rows across micro-partitions.

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

KNOWLEDGE CHECK

Which of the following are characteristics of a micro-partition?

A. Micro-partition data is uncompressed
B. There is one micro-partition per table
C. Column data is stored together
D. They are updated when data changes
E. You cannot change their size

jusegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

KNOWLEDGE CHECK

Which of the following are characteristics of a micro-partition?

A. Micro-partition data is uncompressed
B. There is one micro-partition per table
C. Column data is stored together
D. They are updated when data changes
E. You cannot change their size

C. Column data is stored together

E. You cannot change their size

jusegonzalezv@santander.com.mx - 11-SEP-2023 - @Snowflake 2023-do-not-copy

KNOWLEDGE CHECK

What are some of the things that the cloud services layer is responsible for?

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

KNOWLEDGE CHECK

What are some of the things that the cloud services layer is responsible for?

Storing cache

Managing compute and storage

Query optimization

Authentication and authorization

Encryption key management

Software updates

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

CONNECTING TO SNOWFLAKE

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake
2023 - do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

MODULE AGENDA

- Connection Options
- SnowSQL

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

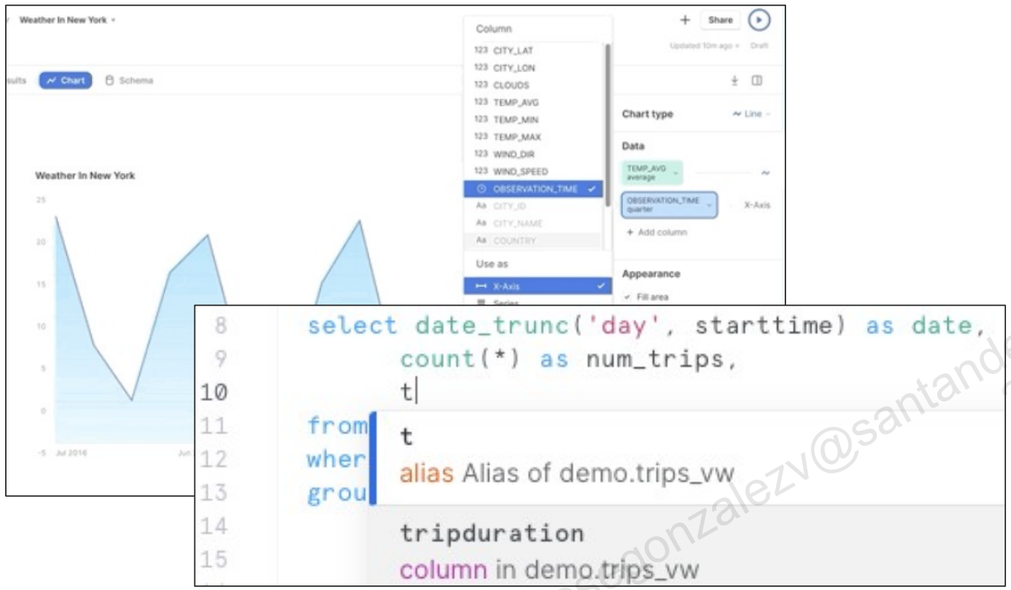
© 2023 Snowflake Computing Inc. All Rights Reserved

CONNECTION OPTIONS

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake
2023-do not copy

© 2023 Snowflake Computing Inc. All Rights Reserved

SNOWSIGHT



- Web-based UI
- Perform Snowflake operations, including:
  - Account management
  - Loading data
  - Querying data
  - Building visualizations and dashboards
  - Managing Snowflake objects
  - Sharing data
  - Monitoring account activities

jose.gonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

SNOWSQL

```
dlittlefield — snowsql ◂ snowsql -a sh32710 -u deborah — 55×16
deborah#DRL_WH@DRL_DB.PUBLIC>select current_role();

+----------------+
| CURRENT_ROLE() |
|----------------|
| TRAINING_ROLE  |
+----------------+
1 Row(s) produced. Time Elapsed: 0.221s
deborah#DRL_WH@DRL_DB.PUBLIC>
```

- Snowflake's command-line client
- Run as interactive shell or in batch mode
- Configure a default connection, and one or more named connections

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake
2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

THIRD-PARTY SOLUTIONS

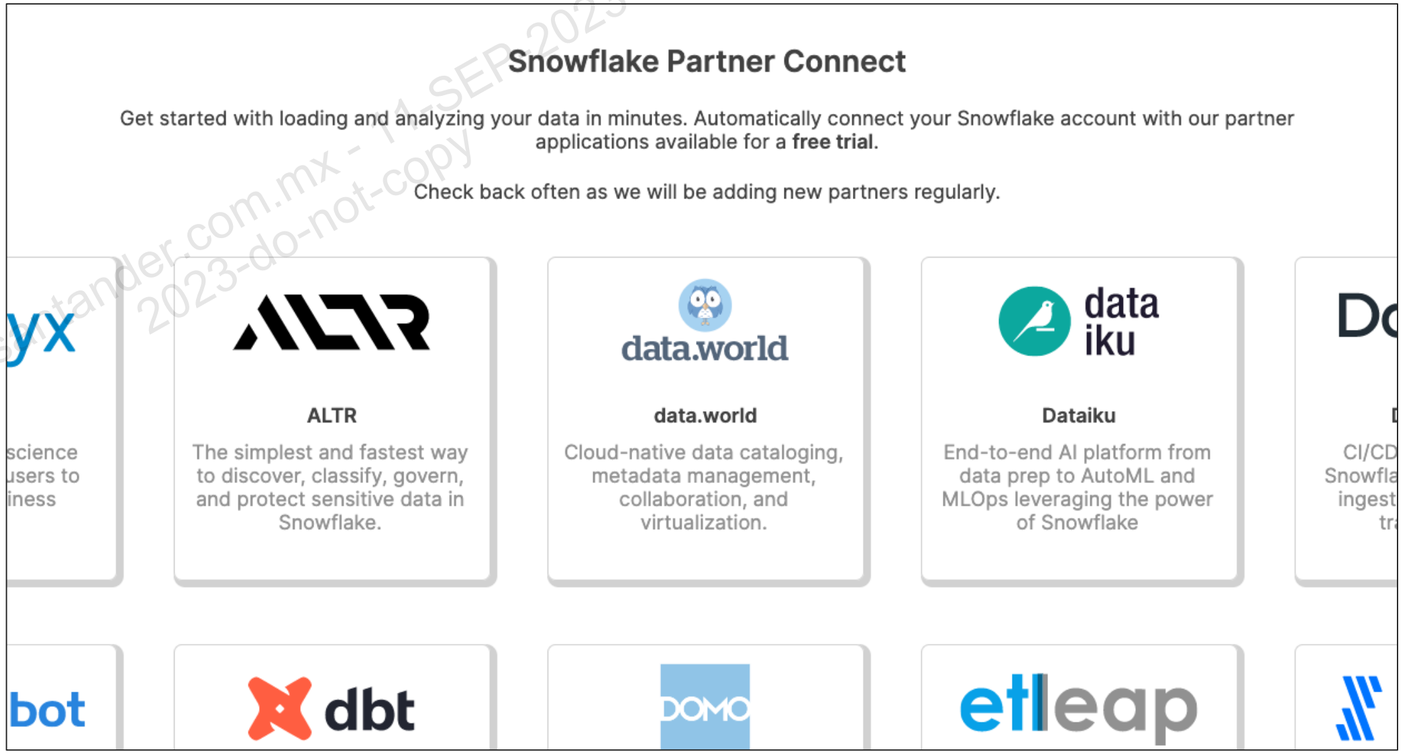
- 3rd-party partners and technologies certified to provide connectivity to Snowflake

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

PARTNER CONNECT

- Create trial account with selected Snowflake partners
- Use limited to `ACCOUNTADMIN`



josegonzalezv@santander.com.mx - 14-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

DRIVER SUPPORT

- GO
- JDBC
- Microsoft .NET
- Node.js
- ODBC
- PHP PDO

josegonzalezv@santander.com.mx - 14-SEP-2023 ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

SNOWFLAKE CONNECTORS

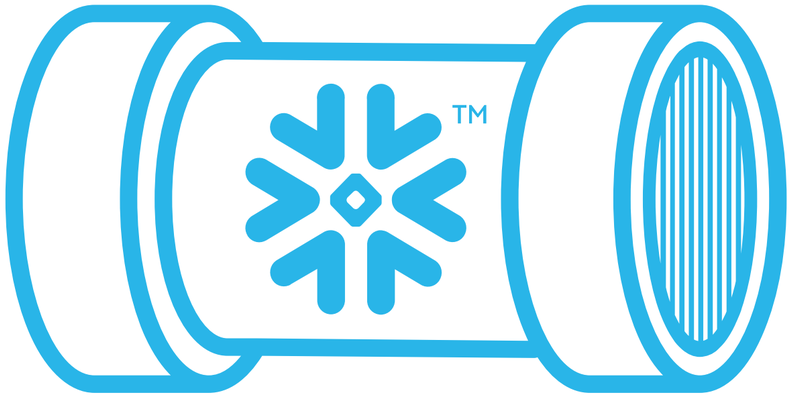
- Snowflake connector for Python
- Snowflake connector for Spark
- Snowflake connector for Kafka

jose.gonzalezv@santander.com.mx - 11-SEP-2023 - Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

APPLICATION PROGRAMMING INTERFACES

- SQL REST API
  - Execute standard queries and most DDL/DML statements
  - Manage your deployment
- Snowpipe REST API
  - Used for continuous data loading pipelines
- Javascript API
  - For User-Defined Functions and Stored Procedures
- REST API for unstructured data support



josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

SNOWPARK

- Build in the language of your choice
  - Scala
  - Java
  - Python (in preview)
- Supports pushdown for all operations
- Does not require a cluster outside of Snowflake
- Define and call User-Defined Functions in your custom code
  - Run code directly in Snowflake, without moving the data to the application

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake
2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

ADDITIONAL INFORMATION

- For information on other ways to connect to Snowflake:
  - Development guides in our user documentation:

    Development guide for Scala

    Development guide for Python

    Development guide for Java

- 1 day add-on course, *Snowpark DataFrame Programming*

- On-demand course, *Data Application Builders Workshop*

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

SNOWSQL

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake
2023-08-Not to copy

© 2023 Snowflake Computing Inc. All Rights Reserved

SNOWSQL

- Command-line client for connecting to Snowflake
- Developed using the Snowflake connector for Python
- Versions for Windows, MacOS, Linux

Desktop — snowsql ◂ Snowsql -a edsvcs1 -u SNOWDEBORAH — 80×36

```
FVFYD0G8HV2G:DESKTOP dlittlefield$ Snowsql -a edsvcs1 -u SNOWDEBORAH
Password:
* SnowSQL * v1.1.82
Type SQL statements or !help
SNOWDEBORAH#(no warehouse)@(no database).(no schema)>USE ROLE TRAINING_ROLE;
+-----------------------------------+
| status                            |
|-----------------------------------|
| Statement executed successfully.  |
+-----------------------------------+
1 Row(s) produced. Time Elapsed: 0.305s
```

joseguz_ibm@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

SNOWSQL

- Run as an interactive shell, or in batch mode
- Can leverage a configuration file to pass parameters
- Can load/unload data to/from local file systems

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

Desktop — snowsql ◂ Snowsql -a edsvcs1 -u SNOWDEBORAH — 80×36

```
FVFYD0G8HV2G:DESKTOP dlittlefield$ Snowsql -a edsvcs1 -u SNOWDEBORAH
Password:
* SnowSQL * v1.1.82
Type SQL statements or !help
SNOWDEBORAH#(no warehouse)@(no database).(no schema)>USE ROLE TRAINING_ROLE;
+-----------------------------------+
| status                            |
|-----------------------------------|
| Statement executed successfully.  |
+-----------------------------------+
1 Row(s) produced. Time Elapsed: 0.305s
```

© 2023 Snowflake Computing Inc. All Rights Reserved

SNOWSQL CONFIGURATION FILE

- Set defaults to be used in SnowSQL
- Can set up multiple connectors

.snowsql — vi config — 80×24

```
#accountname = defaultaccount
#region = defaultregion
#username = defaultuser
#password = defaultpassword
#dbname = defaultdbname
#schemaname = defaultschema
#warehousename = defaultwarehouse
#rolename = defaultrolename
#proxy_host = defaultproxyhost
#proxy_port = defaultproxyport
```

CAUTION: Password, if specified, is stored in clear text

- Locations:
  - Linux/Mac: `~/.snowsql/config`
  - Windows: `%USERPROFILE%\.snowsql\config`

josegonzalez2@santander.com.mx - 17-SEP-2023 - ©Snowflake

© 2023 Snowflake Computing Inc. All Rights Reserved

SNOWSQL CONFIGURATION FILE

- Configure multiple connections

```
dlittlefield — vi .snowsql/confi
[connections.edsvcs3]
accountname = edsvcs3
username = SNOWDRL
dbname = SNOWFLAKE_SAMPLE_DATA
schemaname = WEATHER
warehousename = DRL_WH
rolename = TRAINING_ROLE
```

- Connect with `-c` option

```
dlittlefield — snowsql ◂ snowsql -c edsvc
FVFYD0G8HV2G:~ dlittlefield$ snowsql -c edsvcs3
Password:
* SnowSQL * v1.1.84
Type SQL statements or !help
SNOWDRL#DRL_WH@SNOWFLAKE_SAMPLE_DATA.WEATHER>
```

jsegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

CHECK CURRENT ROLE IN SNOWSQL

dlittlefield — snowsql ◂ snowsql -c edsvcs3 — 80×24

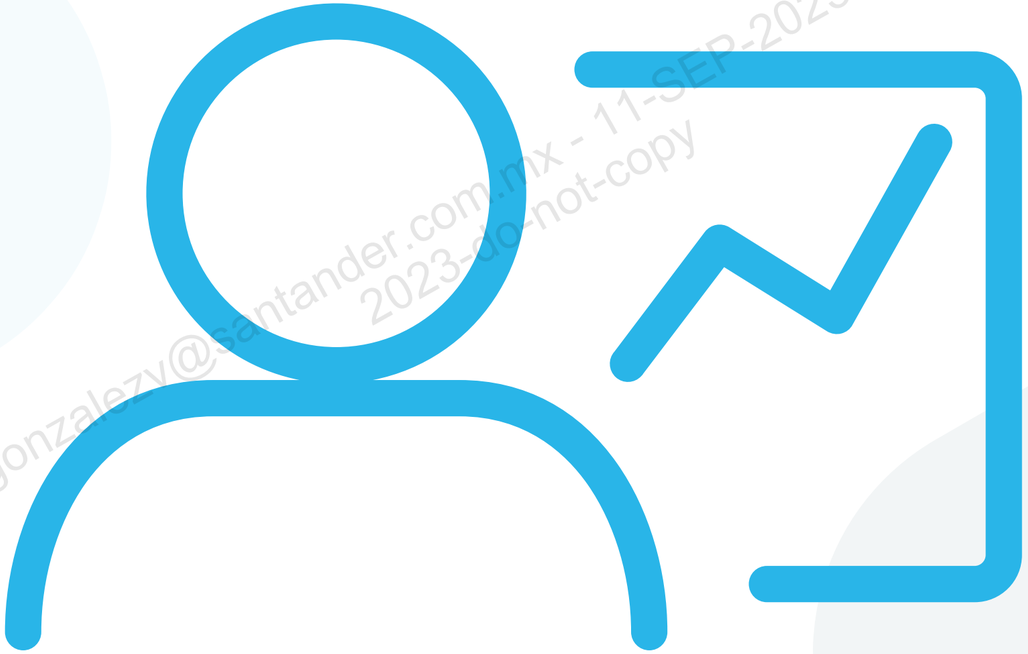
```
FVFYD0G8HV2G:~ dlittlefield$ snowsql -c edsvcs3
Password:
* SnowSQL * v1.1.84
Type SQL statements or !help
SNOWDRL#(no warehouse)@SNOWFLAKE_SAMPLE_DATA.WEATHER>SELECT CURRENT_ROLE();
+----------------+
| CURRENT_ROLE() |
|----------------|
| TRAINING_ROLE  |
+----------------+
1 Row(s) produced. Time Elapsed: 0.125s
SNOWDRL#(no warehouse)@SNOWFLAKE_SAMPLE_DATA.WEATHER>
```

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

INSTRUCTOR DEMO

Visualizations in Snowsight



josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

LAB EXERCISE: 2

Visualizations in Snowsight

20 minutes

Learning Objectives:

- How to create a dashboard from a worksheet
- How to use ad-hoc filters to get insights into data
- How to add tiles to a dashboard
- How to share a dashboard

Note: In this lab you will need to use the instructions in the workbook PDF rather than relying on the instructions in the SQL file.

jose.gonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

LAB EXERCISE: 3

Challenge Lab: Creating a Dashboard

60 minutes

Learning Objectives:

- How to create a dashboard
- How to add tiles to a dashboard
- How to create a bar chart

NOTE: In this lab you will need to use the instructions in the workbook PDF rather than relying on the instructions in the SQL file.

Also, a solution is presented on the last page of this lab in the workbook PDF.

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

DATA PROTECTION FEATURES

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake
2023 do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

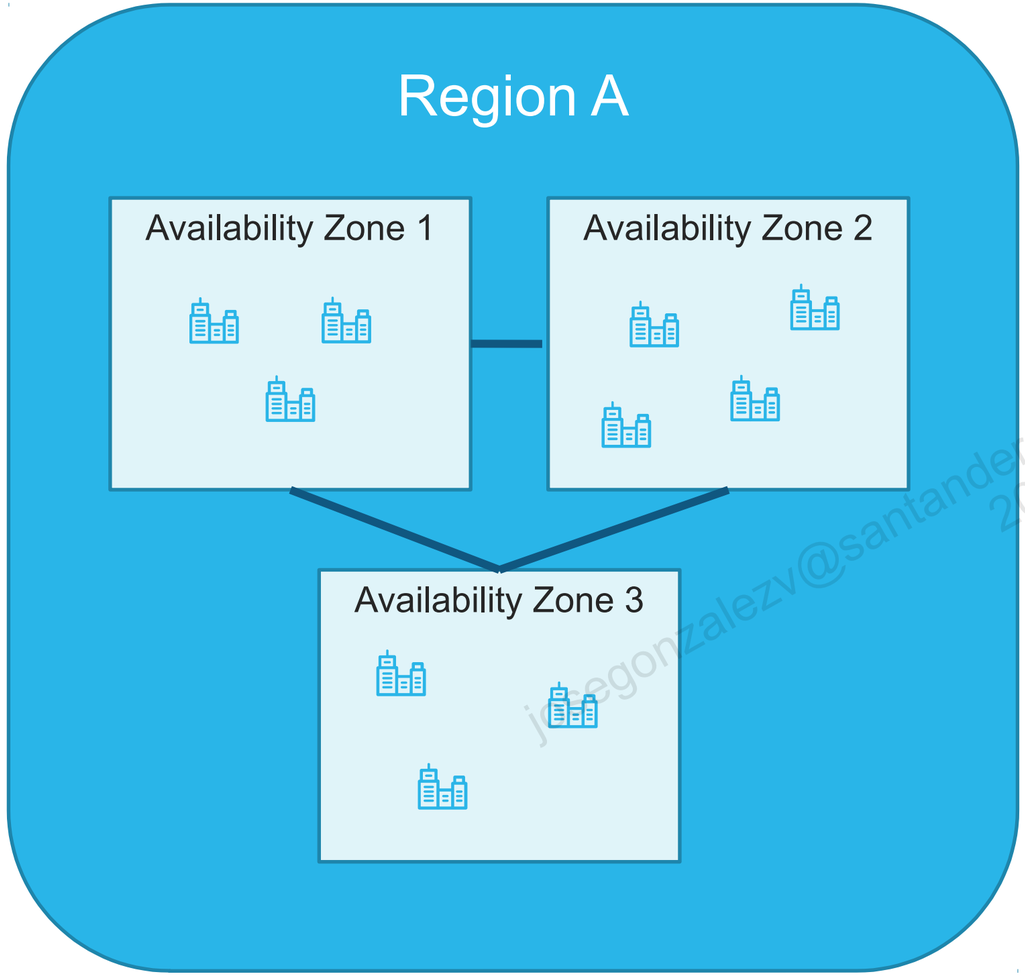
MODULE AGENDA

- Cloning
- Time Travel
- Fail-safe
- Introduction to Replication

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

BUILT-IN CONTINUOUS AVAILABILITY



- A region has multiple availability zones
  - Geographically separate from other zones
  - Each contains one or more data centers
- Data centers in each availability zone have redundant components
- Data is transparently replicated to three different availability zones in a region
- Query processing is done in virtual warehouses, on copies of the data
  - If a VWH goes down, no data is lost
  - VWHs can self-repair

jesegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

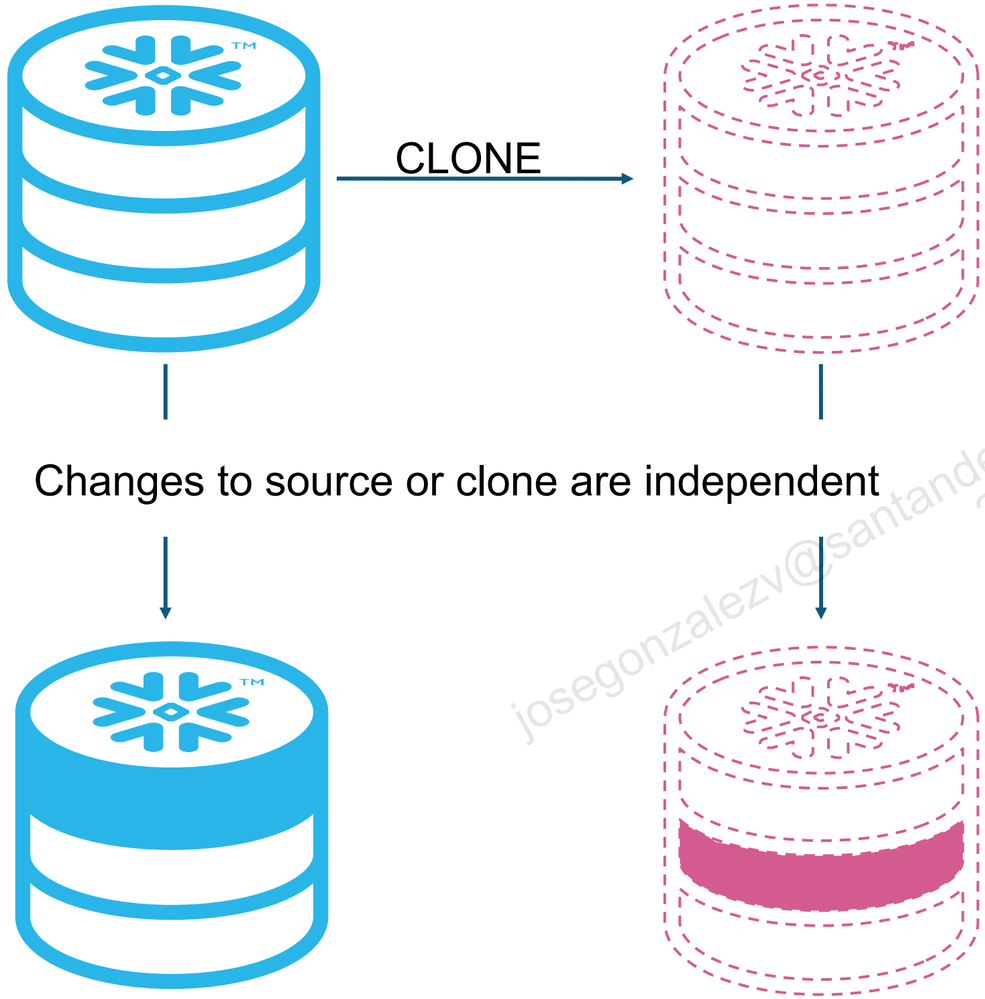
© 2023 Snowflake Computing Inc. All Rights Reserved

CLONING

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake
2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

ZERO-COPY CLONING



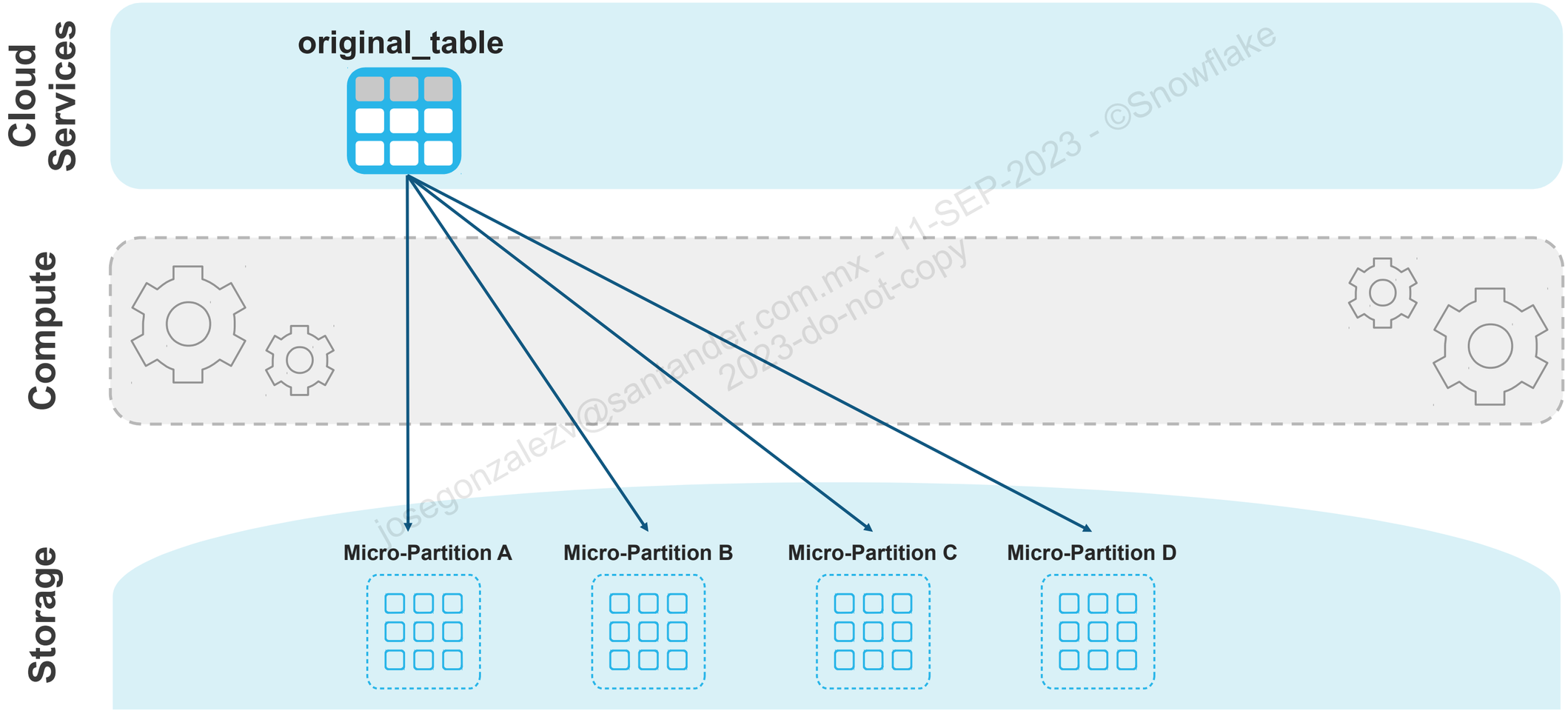
- Quickly take a “snapshot” of any table, schema, or database
  - When the clone is created, micro-partitions are fully shared
- No additional storage costs until changes are made to the original or the clone
- Often used to quickly spin up Dev or Test environments
- Effective “backup” option as well

josegonzalezv@santander.com.mx - 11-SEP-2023 @Snowflake
2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

CLONING EXAMPLE

TABLE TO BE CLONED

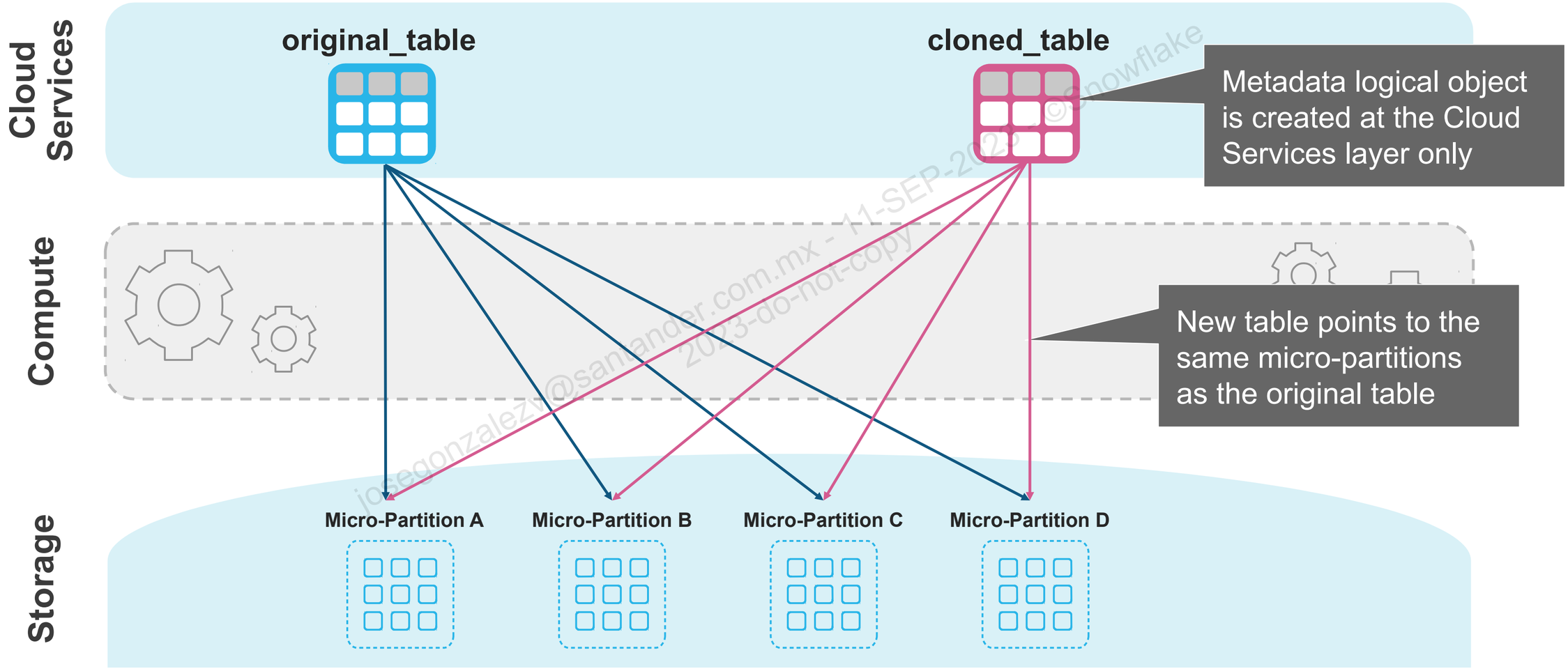


josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake
2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

CLONING EXAMPLE

CLONE TABLE

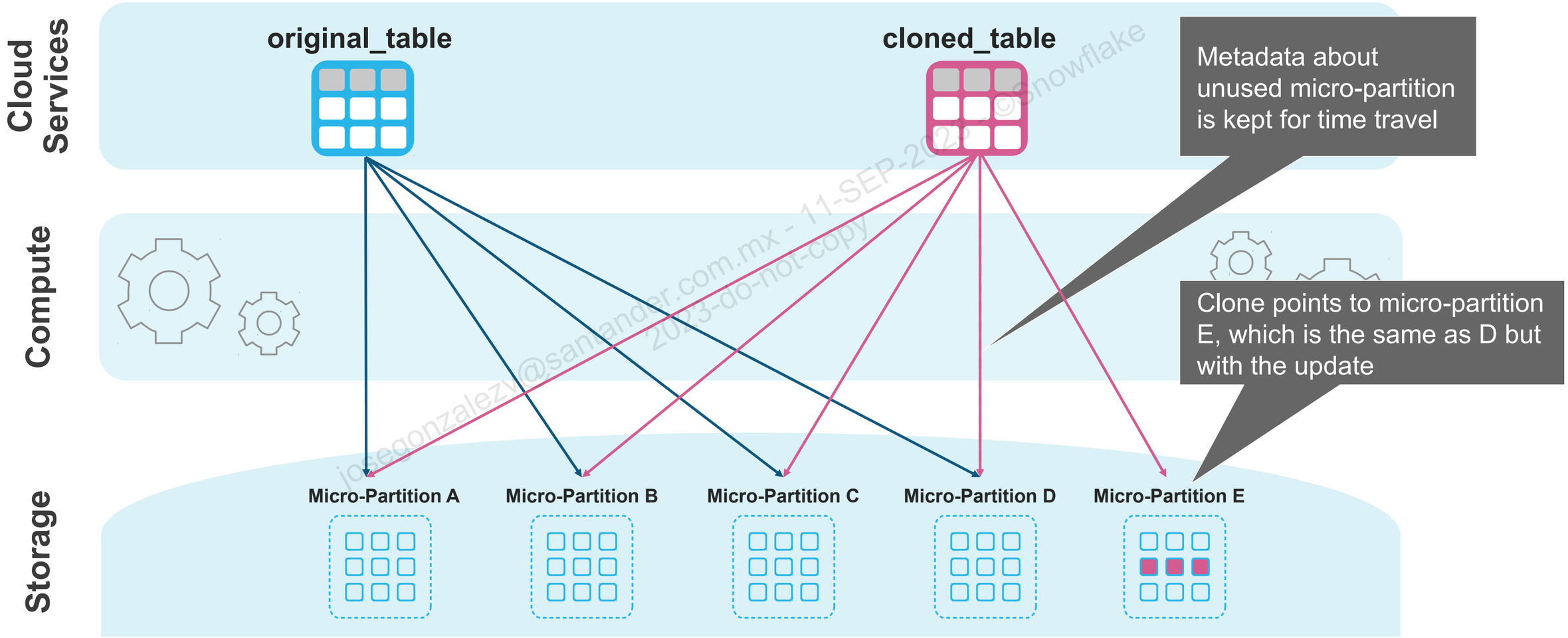


josegonzalezv@santander.com.mx - 11-SEP-2023-do-not-copy © Snowflake

© 2023 Snowflake Computing Inc. All Rights Reserved

CLONING EXAMPLE

UPDATE CLONED TABLE



josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

TIME TRAVEL

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake
2023-do not copy

© 2023 Snowflake Computing Inc. All Rights Reserved

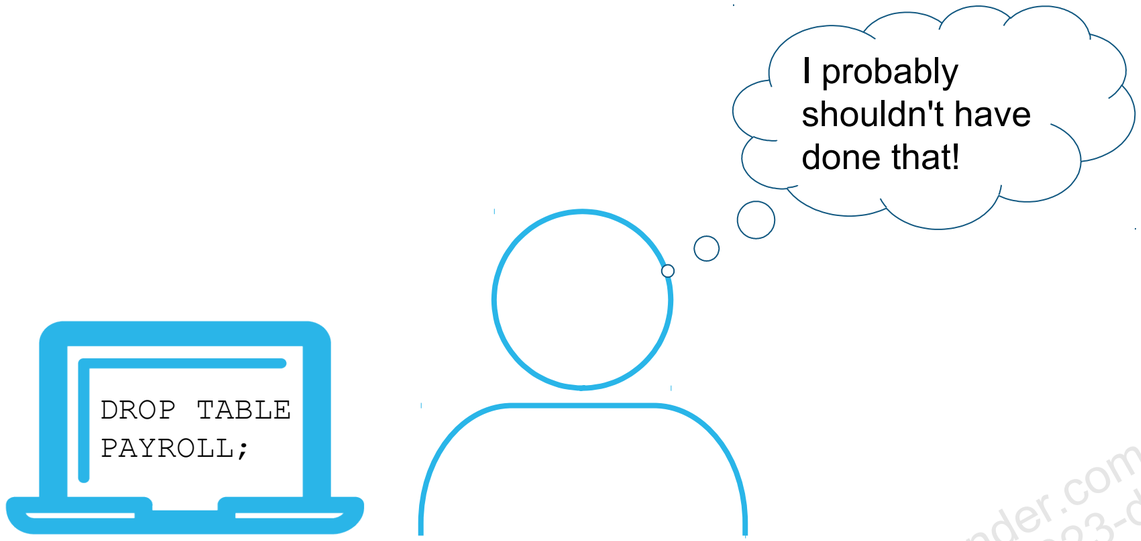
TIME TRAVEL

- Access historical data at any point within a defined retention period
- UNDO common mistakes
- Protect against accidental or intentional modification, removal, or corruption
  - Fix drops, deletes, edits
- Backup/Restore from time or ID
- Time interval depends on value of `DATA_RETENTION_TIME_IN_DAYS`

josegonzalezv@santander.com.mx - 11-SEP-2023 - @Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

TIME TRAVEL SQL EXTENSIONS




```
UNDROP <object>

CREATE <object> <name> CLONE... AT|BEFORE

SELECT... FROM <table> AT|BEFORE

DATA_RETENTION_TIME_IN_DAYS
```

- Instantly bring back deleted tables, schemas, databases
- Restore or duplicate data from key points in the past:
  - Point-in-time
  - Prior to a specific query ID
- Set retention times at the table, schema, database, or account level

josegonzalezv@santander.com.mx - 11-SEP-2023 - @Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

CREATE TABLE WITH TIME TRAVEL

- With default retention period:

```
CREATE TABLE my_table (c1 int);
```

- Customizable retention period:

```
CREATE TABLE my_table (c1 int)
   DATA_RETENTION_TIME_IN_DAYS=90;

ALTER TABLE my_table
   SET DATA_RETENTION_TIME_IN_DAYS=30;
```

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

QUERYING WITH TIME TRAVEL

Query Clauses to support Time Travel Actions

- AT a specific time

```
SELECT * FROM my_table1
   AT(TIMESTAMP => 'Mon, 01 May 2020 16:20:00'::timestamp);
```

- AT an OFFSET

```
SELECT * FROM my_table1
   AT(OFFSET => -60 * 5);
```

- BEFORE a specific query

```
SELECT * FROM my_table1
   BEFORE(STATEMENT => '8e5d0ca9-005e-44e6-b858-a8f5b37c5726');
```

josegonzalezv@santander.com.mx - 01-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

TIME TRAVEL AND CLONING EXAMPLES

- Restoring Objects

```
UNDROP TABLE/SCHEMA/DATABASE
```

- Recovering Historical Objects

```
CREATE TABLE restored_table
   CLONE my_table1
   AT(TIMESTAMP => 'Mon, 09 May 2020 01:01:00'::timestamp);

CREATE DATABASE restored_db
   CLONE my_db
   BEFORE(STATEMENT => '8e5d0ca9-005e-44e6-b858-a8f5b37c5726');
```

josegonzaezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

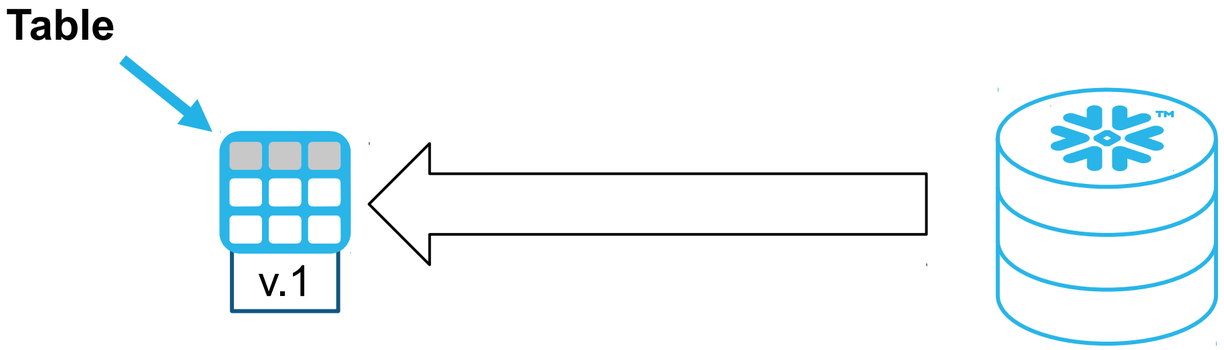
HOW DOES IT WORK?

- Micro-partitions!
- Micro-partitions are immutable
- When data is changed, new micro-partitions are created
- We keep the old versions for the specified retention time

josegonzalezv@santander.com.mx - 1-SEP-2023 - ©Snowflake
2023 - do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

TIME TRAVEL OVERVIEW

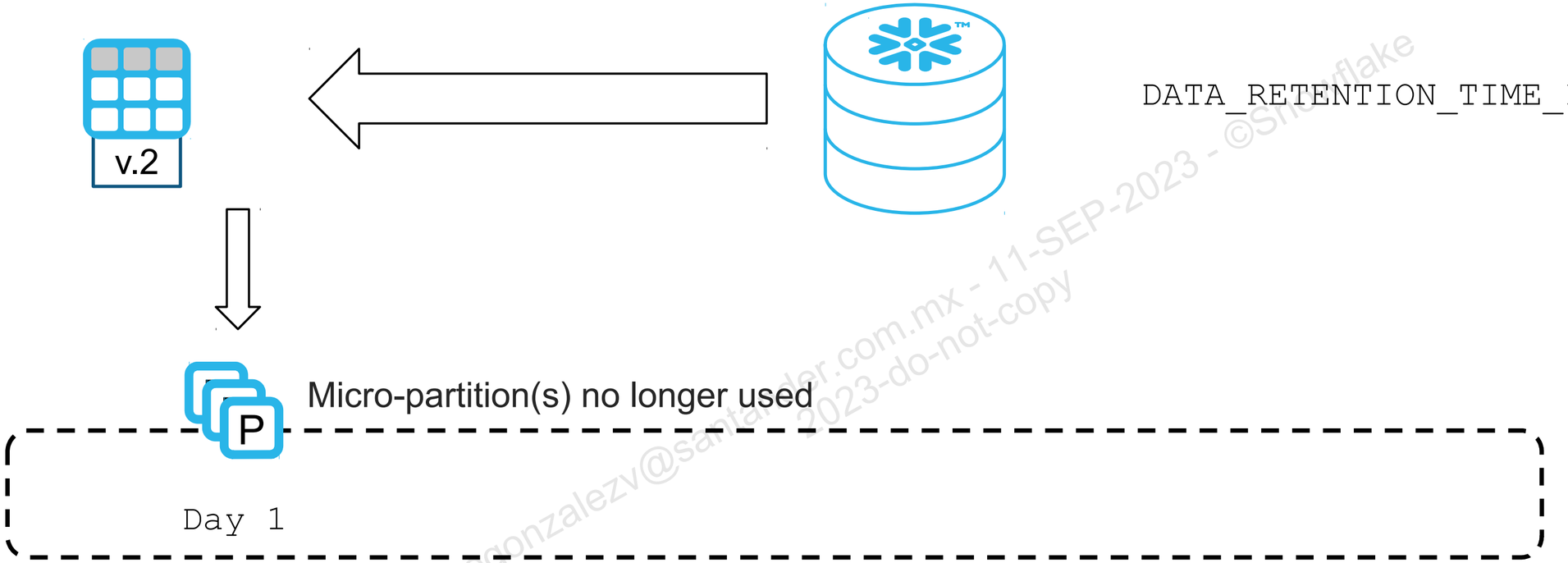



```
DATA_RETENTION_TIME_IN_DAYS = 3
```

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake
2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

TIME TRAVEL OVERVIEW

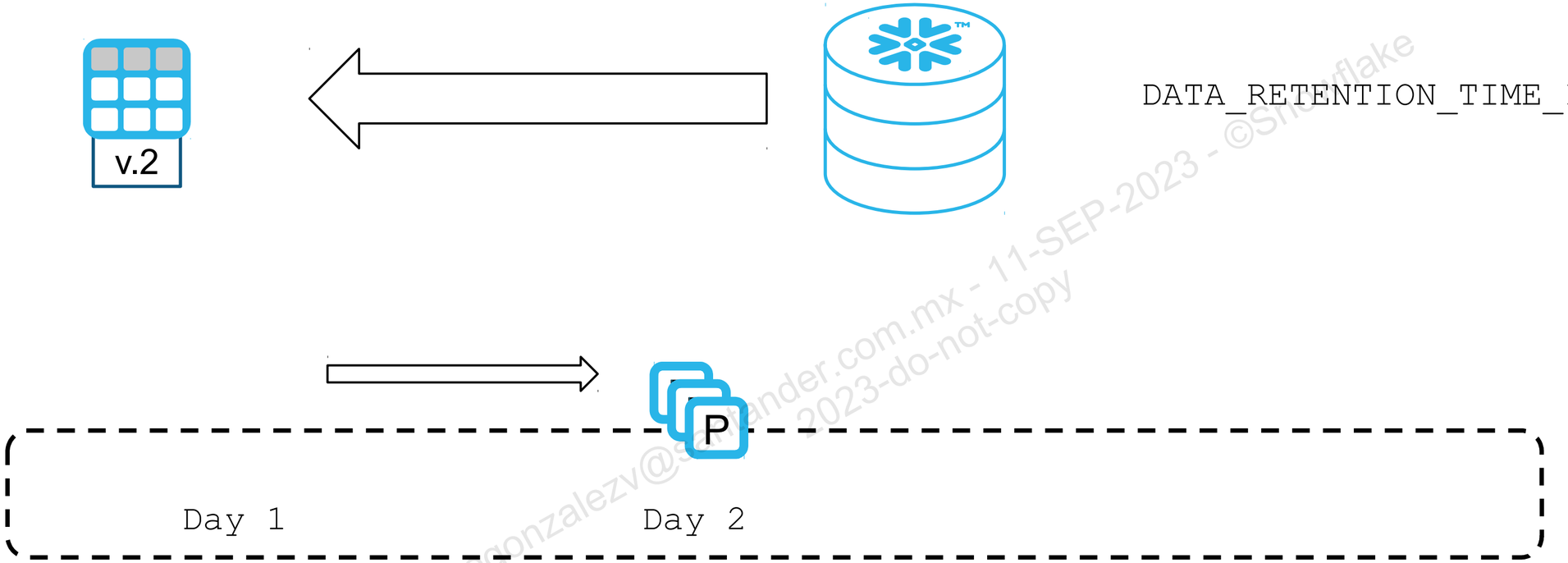



```
DATA_RETENTION_TIME_IN_DAYS = 3
```

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

TIME TRAVEL OVERVIEW

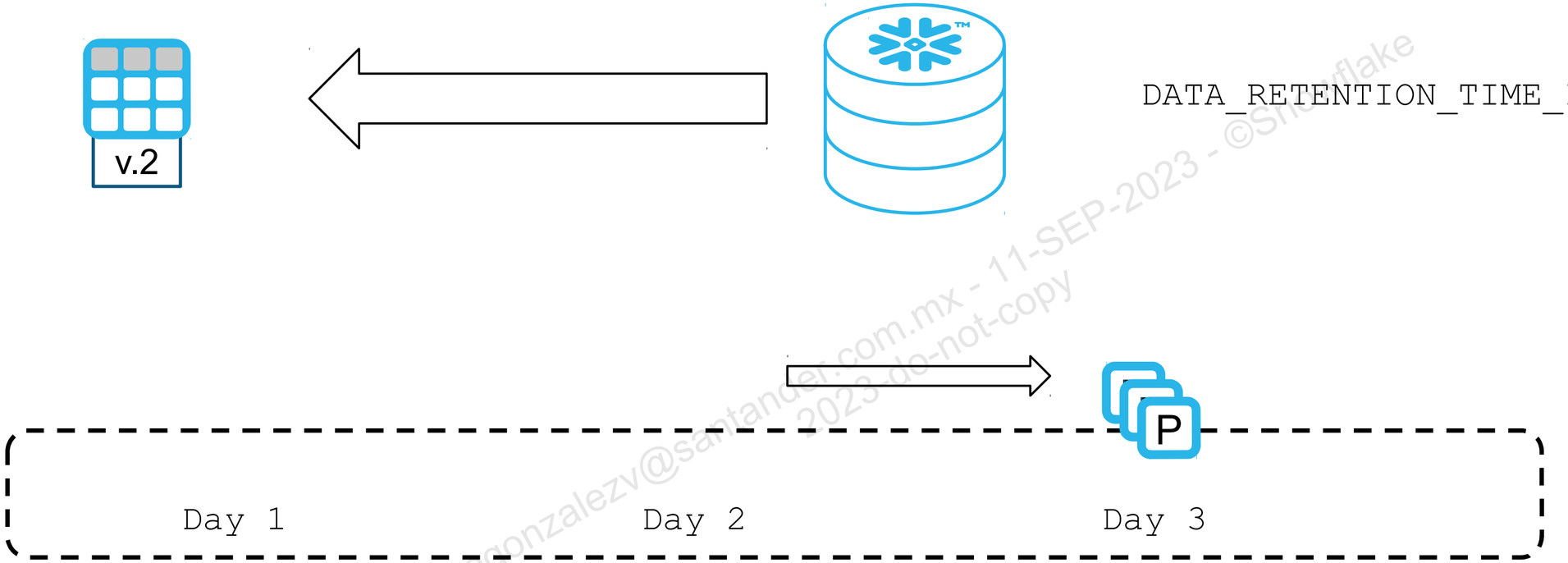



```
DATA_RETENTION_TIME_IN_DAYS = 3
```

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

TIME TRAVEL OVERVIEW

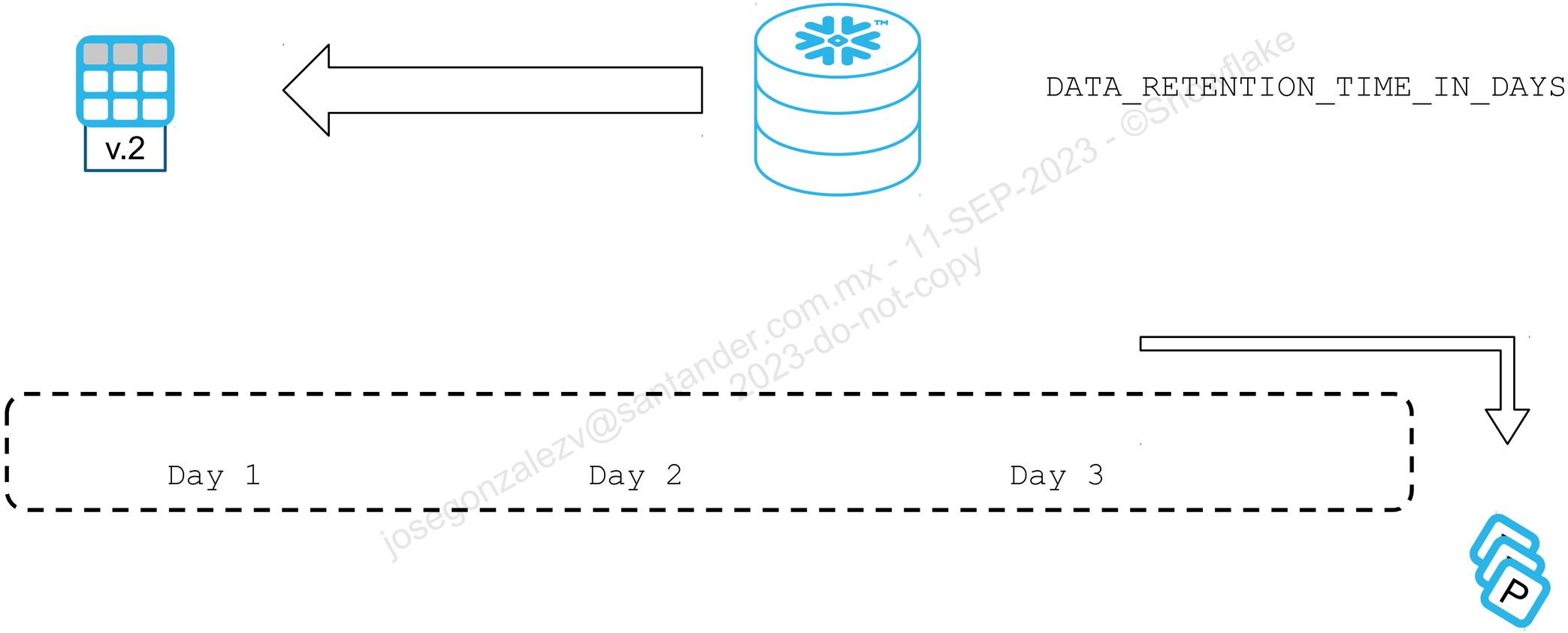



```
DATA_RETENTION_TIME_IN_DAYS = 3
```

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023 do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

TIME TRAVEL OVERVIEW



josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

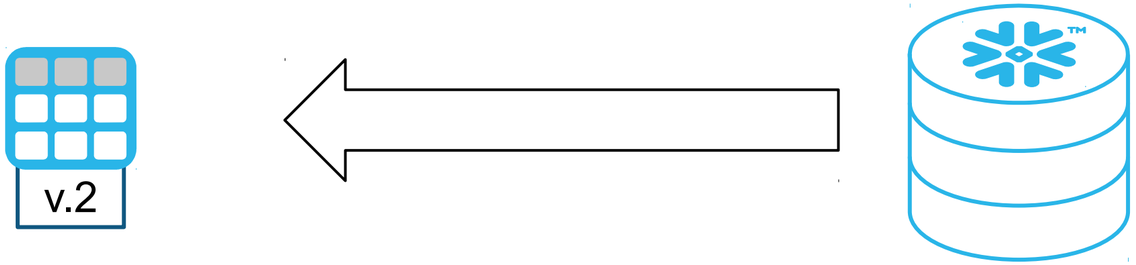
© 2023 Snowflake Computing Inc. All Rights Reserved

FAIL-SAFE

josegonzalezv@santander.com - 11-SEP-2023 - ©Snowflake
2023-do-not-copy

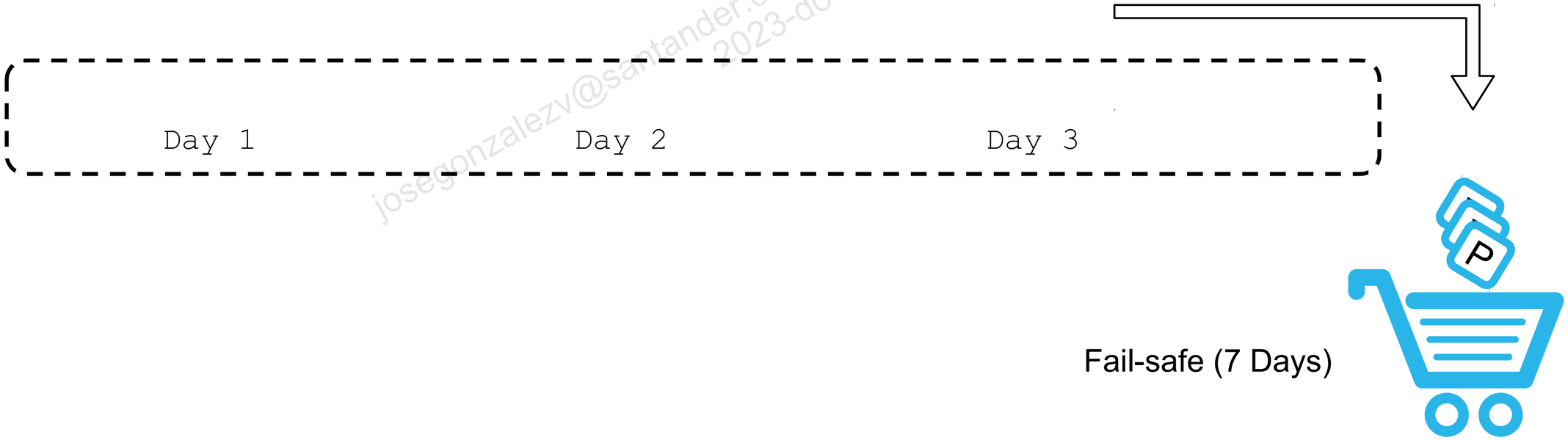
© 2023 Snowflake Computing Inc. All Rights Reserved

FAIL-SAFE OVERVIEW



DATA_RETENTION_TIME_IN_DAYS = 3

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake
2023-do-not-copy



© 2023 Snowflake Computing Inc. All Rights Reserved

FAIL-SAFE STORAGE

- Non-configurable, 7-day retention for historical data after Time Travel expiration
- Only accessible by Snowflake personnel
- Admins can view Fail-safe use in the Snowflake Web UI under *Account* > *Billing and Usage*

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

DATA PROTECTION OPTIONS BY EDITION

Snowflake Edition	Time Travel (1 day)	Time Travel (90 days max)	Fail-safe (7 days)
Standard	✓		✓
Enterprise		✓	✓
Business Critical		✓	✓
VPS		✓	✓

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

LAB EXERCISE: 4

Querying Data with Time Travel

15 minutes

Learning Objectives:

- How to clone a table
- How to write query clauses that support Time Travel actions
- How to fetch and utilize the ID of the last query you ran
- How to use Time Travel to restore data

NOTE: Various steps in this lab intentionally generate errors. The purpose of these errors is to demonstrate and reinforce time travel concepts. So, proceed slowly through the lab and read the instructions carefully.

jose.gonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

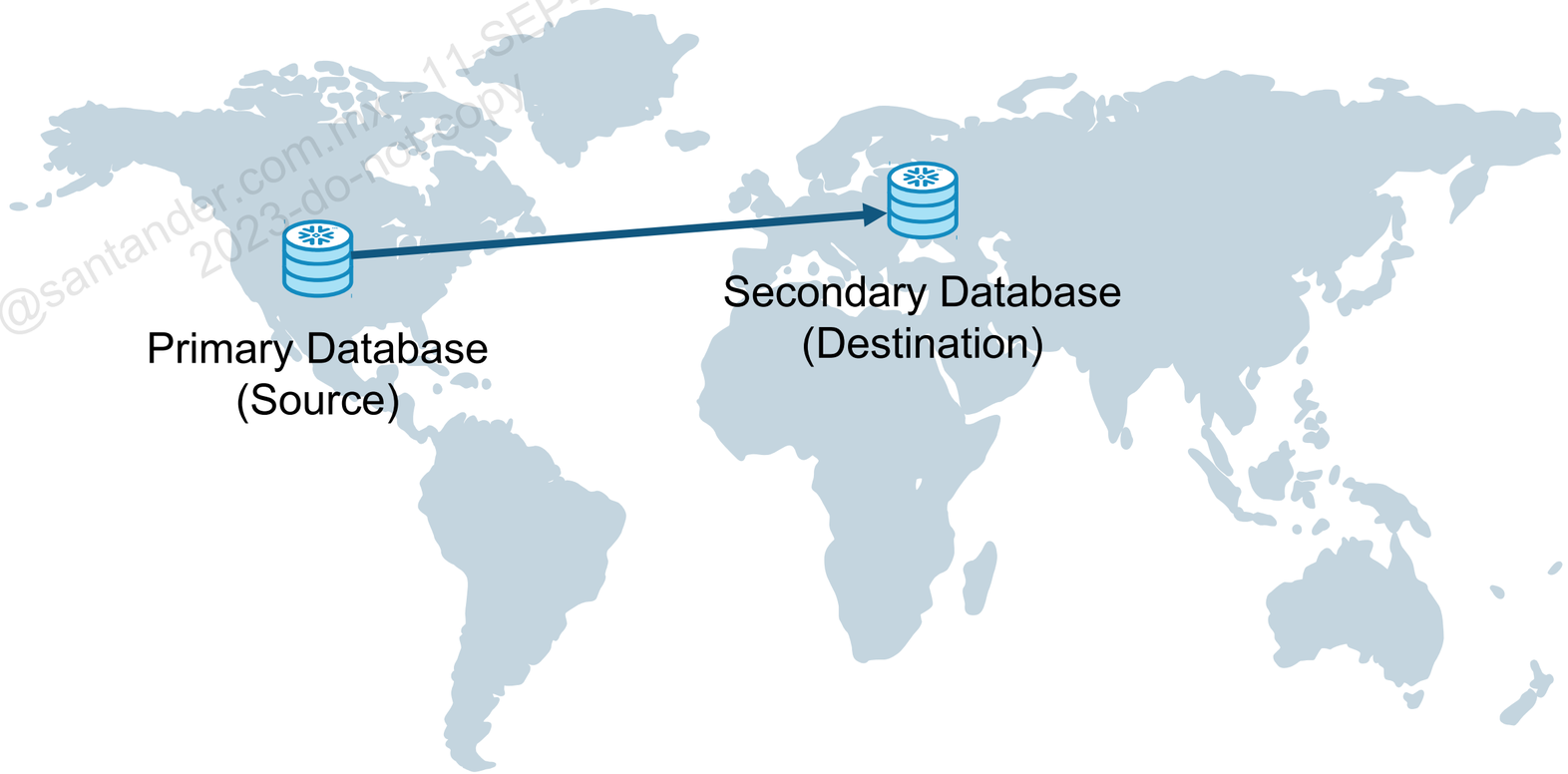
INTRODUCTION TO REPLICATION

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023 do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

WHAT IS DATABASE REPLICATION?

- Available on all editions of Snowflake
- Keeps data synchronized between one or more accounts within the same organization
- Unit of replication is a database (permanent or transient)
- Secondary database is read-only



josegonzalezv@santander.com.mx 11-SEP-2023 @Snowflake
2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

REPLICATION USE CASES

Business Continuity
and Disaster Recovery

Secure Data Sharing
Across Regions / Clouds

Data Portability for
Account Migrations

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023 - do not copy

© 2023 Snowflake Computing Inc. All Rights Reserved

DATABASE REPLICATION CONSIDERATIONS

- Source and destination accounts must be in the same organization
- After initial replication operation, refresh on a scheduled basis
  - Frequency depends on your Recovery Point Objective
- Database replication currently excludes these database objects:
  - Temporary tables, temporary stages, and external tables
  - Pipes, streams, and tasks
- Privileges on database objects are not replicated to a secondary database
- Enable client redirect to maintain client connections after failover (Business Critical feature)

josegonzalezv@santander.com.mx 11-SEP-2023 - @snowflake
2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

ACCOUNT REPLICATION (IN PREVIEW)

- Enables replication of account-level objects, such as:
  - Databases
  - Security & API integrations
  - Network policies
  - Account parameters
  - Resource monitors
  - Users and roles
  - Shares
  - Virtual warehouses

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

REPLICATION FEATURES BY EDITION

Feature	Standard Edition	Business Critical Edition
Database replication	YES	YES
Share replication	YES	YES
Replication groups	YES	YES
Account object replication		YES
Failover groups		YES
Client redirect on failover		YES

josegonzalezv@santander.com.mx - 11-SEP-2023 - @snowflake 2023 - not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

REPLICATION COSTS

- Compute
  - Uses serverless compute
- Storage
  - Storage is duplicated on the target account
  - Clones on the source are physically replicated to the secondary account
- Data transfer/egress
  - Charged by the cloud service provider
  - Amount depends on source and destination

josegonzalez1@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

ADDITIONAL INFORMATION

- Database replication is covered in more detail in the Snowflake documentation (docs.snowflake.com), and in the Snowflake Advanced course

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

KNOWLEDGE CHECK

When you clone a table, what happens initially?

A. The clone does not consume additional storage.

B. The micro-partitions that make up the original table are copied and allocated to the new table.

C. The clone is read-only.

D. Storage for all the micro-partitions is allocated to the original table.

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

KNOWLEDGE CHECK

When you clone a table, what happens initially?

A. The clone does not consume additional storage.

B. The micro-partitions that make up the original table are copied and allocated to the new table.

C. The clone is read-only.

D. Storage for all the micro-partitions is allocated to the original table.

A. The clone does not consume additional storage.

D. Storage for all the micro-partitions is allocated to the original table.

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

KNOWLEDGE CHECK

Which of the following is true about cloned tables?

A. Cloned tables are read-only.

B. Cloned tables can access the time travel data of the original table.

C. Cloned tables can, in turn, be cloned.

D. Cloned tables must retain the same parameter values as the source (such as `DATA_RETENTION_TIME_IN_DAYS`).

josegonzalez@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

KNOWLEDGE CHECK

Which of the following is true about cloned tables?

A. Cloned tables are read-only.

B. Cloned tables can access the time travel data of the original table.

C. Cloned tables can, in turn, be cloned.

D. Cloned tables must retain the same parameter values as the source (such as `DATA_RETENTION_TIME_IN_DAYS`).

C. Cloned tables can, in turn, be cloned.

josegonzalez@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

KNOWLEDGE CHECK

Which of these statements will clone `T_A` to `T_B`?

A. `CLONE TABLE T_A TO T_B`

B. `CLONE TABLE T_B FROM T_A`

C. `CREATE TABLE T_B CLONE T_A`

D. `CREATE TABLE T_B FROM T_A`

E. `CLONE TABLE T_A CREATE T_B`

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

KNOWLEDGE CHECK

Which of these statements will clone `T_A` to `T_B`?

A. `CLONE TABLE T_A TO T_B`

B. `CLONE TABLE T_B FROM T_A`

C. `CREATE TABLE T_B CLONE T_A`

D. `CREATE TABLE T_B FROM T_A`

E. `CLONE TABLE T_A CREATE T_B`

C. `CREATE TABLE T_B CLONE T_A`

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowfla
2023-do-not-copy

KNOWLEDGE CHECK

With the Enterprise edition of Snowflake, what is longest setting for the parameter `DATA_RETENTION_TIME_IN_DAYS`?

A. 1 day
B. 10 days
C. 90 days
D. 120 days

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

KNOWLEDGE CHECK

With the Enterprise edition of Snowflake, what is longest setting for the parameter `DATA_RETENTION_TIME_IN_DAYS`?

A. 1 day
B. 10 days
C. 90 days
D. 120 days

C. 90 days

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

KNOWLEDGE CHECK

What causes micro-partitions to go into Time Travel?

A. Inserting data to a table

B. Deleting data from a table

C. Updating rows in a table

D. Truncating a table

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

KNOWLEDGE CHECK

What causes micro-partitions to go into Time Travel?

A. Inserting data to a table

B. Deleting data from a table

C. Updating rows in a table

D. Truncating a table

B. Deleting data from a table

C. Updating rows in a table

D. Truncating a table

josegonzalezv@santander.com.mx - 11-SEP-2023 ©Snowflake 2023-do-not-copy

KNOWLEDGE CHECK

What are common use cases for database replication?

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

KNOWLEDGE CHECK

What are common use cases for database replication?

Disaster recovery

Sharing data outside your region

Data migration

josegonzalezv@santander.com.mx - 11-SEP-2023 ©Snowflake
2023-do-not-copy

KNOWLEDGE CHECK

With database replication, which of the following objects will be replicated?

A. Schemas

B. Views

C. Resource Monitors

D. Tables

E. Users

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

KNOWLEDGE CHECK

With database replication, which of the following objects will be replicated?

A. Schemas

B. Views

C. Resource Monitors

D. Tables

E. Users

A. Schemas

B. Views

D. Tables

josegonzalezv@santander.com.mx - 11-SEP-2023 ©Snowflake
2023-do-not-copy

SQL SUPPORT IN SNOWFLAKE

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake
2023-06-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

MODULE AGENDA

- Tables, Views, and Data Types
- Standard SQL and Snowflake
- Transactions
- Multi-Table Inserts
- Collations
- Query Tags
- Working with Parameters

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

TABLES, VIEWS, AND DATA TYPES

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake
2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

TABLE CHARACTERISTICS

- **Permanence** – does a user need to drop the table, or is it dropped automatically?
- **Visibility** – who can see the table?
- **Time Travel** – what is the maximum time travel setting?
  - Time travel is a feature that allows you to view or restore a table as it existed in the past
  - This feature potentially consumes additional storage space
- **Fail-safe** – is the table covered by fail-safe?
  - Fail-safe is a feature that saves deleted data for an additional 7 days
  - This feature consumes additional storage space

© 2023 Snowflake Computing Inc. All Rights Reserved

TABLE TYPES

PERMANENT

- Persists until dropped
- Visible to all in the role (or others with privileges)
- For data that requires the highest level of data protection and recovery

Time Travel	Up to 90 days with Enterprise
Fail-safe	✔

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

TABLE TYPES

	PERMANENT	TRANSIENT*
	• Persists until dropped • Visible to all in the role (or others with privileges) • For data that requires the highest level of data protection and recovery	• Persists until dropped • Visible to all in the role (or others with privileges) • Does not need the same level of data retention as a permanent table
Time Travel	Up to 90 days with Enterprise	0 or 1 days
Fail-safe	✓	✗

*Can have transient databases, schemas, and tables

josegonzalezvocantandar.com.mx - 11-SEP-2023 - ©Snowflake 2023 - do not copy

© 2023 Snowflake Computing Inc. All Rights Reserved

TABLE TYPES

	PERMANENT	TRANSIENT*	TEMPORARY
	• Persists until dropped • Visible to all in the role (or others with privileges) • For data that requires the highest level of data protection and recovery	• Persists until dropped • Visible to all in the role (or others with privileges) • Does not need the same level of data retention as a permanent table	• Dropped when the session that created it is closed • Visible only in the session where it was created • Used for temporary data (for example, ETL/ELT)
Time Travel	Up to 90 days with Enterprise	0 or 1 days	0 or 1 days
Fail-safe	✔	✘	✘

*Can have transient databases, schemas, and tables

josegonzalezvocantandes.com.mx - 11-SEP-2023 - ©Snowflake 2023 - do not copy

© 2023 Snowflake Computing Inc. All Rights Reserved

TABLE TYPES

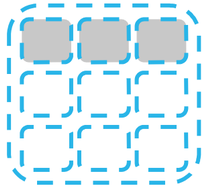
	PERMANENT	TRANSIENT*	TEMPORARY	EXTERNAL
	• Persists until dropped • Visible to all in the role (or others with privileges) • For data that requires the highest level of data protection and recovery	• Persists until dropped • Visible to all in the role (or others with privileges) • Does not need the same level of data retention as a permanent table	• Dropped when the session that created it is closed • Visible only in the session where it was created • Used for temporary data (for example, ETL/ELT)	• Persists until dropped • Visible to all in the role (or others with privileges) • Snowflake “over” an external data lake • Read-only
Time Travel	Up to 90 days with Enterprise	0 or 1 days	0 or 1 days	✗
Fail-safe	✓	✗	✗	✗

*Can have transient databases, schemas, and tables

josegonzalezvocantandar.com.mx - 11-SEP-2023 - ©Snowflake 2023 do not copy

© 2023 Snowflake Computing Inc. All Rights Reserved

VIEW TYPES



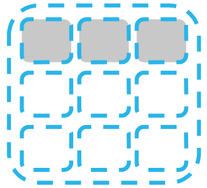
Standard View

- Default view type
- Named definition of a query
- Executes as owning role
- Underlying DDL available to any role with access to the view.

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

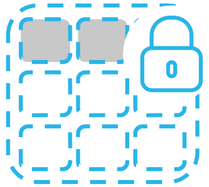
© 2023 Snowflake Computing Inc. All Rights Reserved

VIEW TYPES



Standard View

- Default view type
- Named definition of a query
- Executes as owning role
- Underlying DDL available to any role with access to the view.



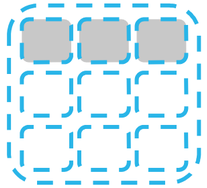
Secure View

- Definition and details only visible to authorized users
- Executes as owning role
- Snowflake query optimizer bypasses optimizations used for regular views

jsegonzalez@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023 do-not-copy

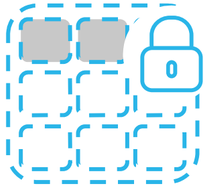
© 2023 Snowflake Computing Inc. All Rights Reserved

VIEW TYPES



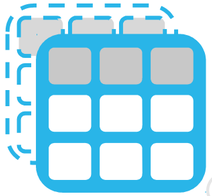
Standard View

- Default view type
- Named definition of a query
- Executes as owning role
- Underlying DDL available to any role with access to the view.



Secure View

- Definition and details only visible to authorized users
- Executes as owning role
- Snowflake query optimizer bypasses optimizations used for regular views



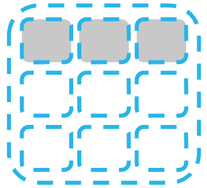
Materialized View

- Behaves more like a table
- Results of underlying query are stored
- Auto-refreshed

jsegonzalez@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023 do-not-copy

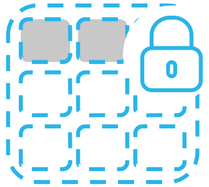
© 2023 Snowflake Computing Inc. All Rights Reserved

VIEW TYPES



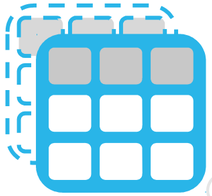
Standard View

- Default view type
- Named definition of a query
- Executes as owning role
- Underlying DDL available to any role with access to the view.



Secure View

- Definition and details only visible to authorized users
- Executes as owning role
- Snowflake query optimizer bypasses optimizations used for regular views



Materialized View

- Behaves more like a table
- Results of underlying query are stored
- Auto-refreshed

Secure Materialized View

- Definition and details only visible to authorized users
- Executes as owning role
- Snowflake query optimizer bypasses optimizations used for regular materialized views

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023 do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

NUMERIC DATA TYPES

Type defined in column…	Stored as this in Snowflake…
NUMBER(38,0), DECIMAL, NUMERIC	NUMBER(38,0) – unless you specify a different scale and precision
INT, INTEGER, BIGINT, SMALLINT	NUMBER(38,0) – cannot specify scale or precision
DOUBLE, DOUBLE PRECISION, FLOAT, FLOAT4, FLOAT8, REAL	DOUBLE

josegonzalezy@santander.com.mx - 11-DEC-2023 - ©Snowflake
2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

STRING, BINARY, AND LOGICAL DATA TYPES

Type defined in column…	Stored as this in Snowflake…
VARCHAR, STRING, TEXT	VARCHAR – max size is 16,777,216 bytes
CHAR, CHARACTER	VARCHAR(1) – though you can specify a different length
BINARY, VARBINARY	BINARY
BOOLEAN	BOOLEAN

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake
2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

DATE AND TIME DATA TYPES

Type defined in column…	Stored as this in Snowflake…
DATE	DATE
TIME	TIME
TIMESTAMP_LTZ	TIMESTAMP with Local Time Zone - If incoming timestamp has no time zone specified, its loaded with the local time zone
TIMESTAMP_NTZ, DATETIME	TIMESTAMP with no time zone – time zone, if provided, is not stored
TIMESTAMP_TZ	TIMESTAMP with time zone – UTC time with associated time zone offset (session time zone offset used if it is not provided)
TIMESTAMP	Alias for one of the TIMESTAMP variations – TIMESTAMP_NTZ by default; set with TIMESTAMP_TYPE_MAPPING parameter

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake
2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

OTHER DATA TYPES

Type	Notes
GEOSPATIAL	Includes GEOGRAPHY and GEOMETRY types
VARIANT	Used for storing raw semi-structured data – maximum length of a VARIANT column is 16 MB (compressed)
OBJECT	A collection of key-value pairs, where each value is of type VARIANT
ARRAY	A collection of 0 or more values, indexed starting at [0] – all values in the array are of type VARIANT

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake
2023 do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

OBJECT IDENTIFIER CASE SENSITIVITY

Table name specified	Result	Displayed as:
`CREATE TABLE My_Table`	Case-insensitive	`MY_TABLE`
`CREATE TABLE my_table`	Case-insensitive	`MY_TABLE`
`CREATE TABLE "My_Table"`	Case-sensitive	`My_Table`
`CREATE TABLE "my_table"`	Case-sensitive	`my_table`

jcsegonzalez@santander.com.mx - 11-SEP-2023 - Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

CONSTRAINTS

- Snowflake provides support for defining constraints
- The only constraint Snowflake enforces is NOT NULL
- Primarily for data modeling purposes, and to support client tools that use constraints
  - Example: Tableau supports using constraints for join culling, which can improve performance

josegonzalezv@santander.com.mx - 01-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

KNOWLEDGE CHECK

What are the four types of tables supported in Snowflake?

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

KNOWLEDGE CHECK

What are the four types of tables supported in Snowflake?

Permanent tables

Temporary tables

Transient tables

External tables

josegonzalezv@santander.com.mx - 11-SEP-2023 - @Snowflake 2023-do-not-copy

KNOWLEDGE CHECK

You are loading data into a landing table, and several users will be taking data from the table, transforming it, and storing it into different permanent tables.

If you wanted to minimize the storage used by the landing table – what type of table would you use?

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

KNOWLEDGE CHECK

You are loading data into a landing table, and several users will be taking data from the table, transforming it, and storing it into different permanent tables.

If you wanted to minimize the storage used by the landing table – what type of table would you use?

A transient table, because it is not covered by fail-safe (and up to only one day of time travel).

A temporary table is also not covered by fail-safe – but since multiple users will access the landing table, it cannot be a temporary table (which is specific to a user and a session).

josegonzalezv@santander.com.mx - 11-SEP-2023 - @Snowflake
2023-do-not-copy

KNOWLEDGE CHECK

You create a table with this statement:

```
CREATE TABLE my_table (c1
INT);
```

When you `DESCRIBE` the table, what data type will you see listed for c1?

A. `INTEGER`

B. `NUMBER`

C. `NUMBER(38,0)`

D. `INT`

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

KNOWLEDGE CHECK

You create a table with this statement:

```
CREATE TABLE my_table (c1
INT);
```

When you `DESCRIBE` the table, what data type will you see listed for c1?

A. `INTEGER`

B. `NUMBER`

C. `NUMBER(38,0)`

D. `INT`

C. `NUMBER(38,0)`

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

KNOWLEDGE CHECK

You create a table with this statement:

```
CREATE TABLE tbl1 (c1 CHAR);
```

When you DESCRIBE the table, what data type will you see listed for c1?

A. `VARCHAR(1)`

B. `VARCHAR`

C. `VARCHAR(16777216)`

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

KNOWLEDGE CHECK

You create a table with this statement:

```
CREATE TABLE tbl1 (c1 CHAR);
```

When you DESCRIBE the table, what data type will you see listed for c1?

A. `VARCHAR(1)`

B. `VARCHAR`

C. `VARCHAR(16777216)`

A. `VARCHAR(1)`

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

KNOWLEDGE CHECK

Which of the following constraints does Snowflake maintain?

A. UNIQUE

B. PRIMARY KEY

C. FOREIGN KEY

D. NOT NULL

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

KNOWLEDGE CHECK

Which of the following constraints does Snowflake maintain?

A. UNIQUE

B. PRIMARY KEY

C. FOREIGN KEY

D. NOT NULL

A. UNIQUE

B. PRIMARY KEY

C. FOREIGN KEY

D. NOT NULL

Snowflake maintains all these constraints, but the only one it **enforces** is NOT NULL

josegonzalezv@santander.com.mx - 01-SEP-2023 - ©Snowflake 2023-do-not-copy

TRANSACTIONS

josegonzalezv@santandertecnologia.com.mx - 11-SEP-2023 - ©Snowflake
2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

SNOWFLAKE TRANSACTIONS

- Snowflake supports ACID transactions
  - **Atomicity** – each transaction either succeeds completely, or fails completely
  - **Consistency** – every transaction will result in a consistent, or valid, state
  - **Isolation** – the state of the database after multiple transactions will be the same whether the transactions are completed concurrently, or sequentially
  - **Durability** – once a transaction is completed, it retains its state even if there is some system failure

- Readers and writers do not block each other
  - Support Read Committed transaction isolation level

joseagonzalezv@santander.com.mx - 14-SEP-2023 - ©Snowflake
2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

IMPLICIT TRANSACTIONS

```
INSERT INTO
  my_table
VALUES
   ('Deborah', 'Mark',  12),
   ('Jenn',    'Keith', 25),
   ('Sharon',  'Steve', 'two');
```

Numeric value 'two' is not recognized

- Snowflake's default is AUTOCOMMIT
  - Each statement, by default, is implicitly treated as a transaction
- Everything is committed, or everything is rolled back
- In the example, there is a mixture of data types for the third column - the entire statement will fail, not just the entry with the error

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

EXPLICIT TRANSACTIONS

```
BEGIN TRANSACTION;
  <statement one>;
  <statement two>;
  ...
  <statement n>;
COMMIT;

BEGIN TRANSACTION;
  <statement one>;
  <statement two>;
  ...
  <statement n>;
ROLLBACK;
```

- Define explicit transactions to group two or more statements into a transaction
- If a transaction is already active, any `BEGIN TRANSACTION` statements are ignored
- Include only DML statements or query statements
  - DDL statements implicitly commit active transactions

jose.gonzalezv@santander.com.mx - 11-SEP-2023 - @Snowflake
2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

LAB EXERCISE: 5

Working with Transactions

20 Minutes

Learning Objectives:

- Summarize the transaction model and concurrency control.
- Explore autocommit default values.
- Change session values for the autocommit parameter.
- Explore multi-statement transactions.
- Monitor the SHOW LOCKS and LOCK_TIMEOUT parameters for concurrency control.
- Use SYSTEM$ABORT_TRANSACTION to release locks and terminate transactions.

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

STANDARD SQL AND SNOWFLAKE

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake
please-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

STANDARD SQL

- Snowflake supports most DDL and DML defined in SQL:2003
  - Also supports some constructs from later ANSI SQL standards
- DDL and DML for tables:
  - `INSERT`, `UPDATE`, `MERGE`, `DELETE`, `TRUNCATE`, `CREATE`, `DROP`, `ALTER`
- DDL and DML for views, schemas, and databases:
  - `CREATE, DROP, ALTER`

josegonzalezv@santander.com.mx - 14-SEP-2023 - ©Snowflake 2023 - do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

QUERY SYNTAX

- Snowflake SELECT supports all the standard syntax options including:
  - All standard JOIN types (inner, left/right/full outer, natural, cross, lateral)
  - All standard function types (scalar, aggregate, window, table, system, user-defined)
  - All standard query syntax (`WITH`, `INTO`, `FROM`, `WHERE`, `GROUP BY`, `QUALIFY`, `ORDER BY`, `LIMIT`…)
- Most options supported by other standard RDBMS will work
- See the documentation (docs.snowflake.com) for details
  - https://docs.snowflake.com/en/sql-reference.html
  - https://docs.snowflake.com/en/sql-reference-commands.html
  - https://docs.snowflake.com/en/sql-reference/constructs.html

jpsegovialez@santander.com.mx - 11-SEP-2023 ©Snowflake
2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

SHOW OBJECTS

- Lists and shows information on the existing objects of the specified type

```
SHOW DATABASES;
```

··· created_on	name	is_default	is_current	origin	owner	comment	options	retention_time
12:54:48.076 -0700	ADMIN_ONLY	N	N		ACCOUNTADMI			1
09:11:14.789 -0700	DRL_DB	N	N		DBA			1
14:17:27.895 -0700	DWEELEA_DB	N	N		SYSADMIN			1
13:44:16.277 -0700	HIDDEN_DB	N	N		FLOATER			1

- Use `LIKE <pattern>` to filter output by object name

```
SHOW TABLES LIKE '%test%';
```

joseg.gonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

DESCRIBE OBJECT

- Describes the components of a specified object

```
DESCRIBE DATABASE my database;
```

... created_on	name	kind
2022-02-16 14:38:32.843 -0800	GARDENING	SCHEMA
2022-09-27 09:40:32.495 -0700	INFORMATION_SCHEMA	SCHEMA
2022-07-16 13:59:53.120 -0700	PAROUET	SCHEMA

```
DESCRIBE TABLE members;
```

name ...	type	kind	null?	def
MEMBER_ID	VARCHAR(16777216)	COLUMN	Y	nul
POINTS_BALANCE	NUMBER(38,0)	COLUMN	Y	nul
STARTED_DATE	DATE	COLUMN	Y	nul

jose.gonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

GET_DDL

- Returns a DDL statement that can be used to recreate the specified object

```
SELECT GET_DDL('SEQUENCE', 'my_seq');
```

GET_DDL('SEQUENCE', 'MY_SEQ')
create or replace sequence MY_SEQ start with 5 increment by 2;

- Works with most database objects
- Does not work with secure views unless you are in the owning role

Object does not exist, or operation cannot be performed.

josegonzalez@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

SUPPORTED FUNCTION TYPES

Type	Description
Scalar	Takes 0 or more parameter values, returns a single value per input.
Aggregate	Performs a calculation on a set of values, and returns a single value.
Window	Aggregate functions that operate on a subset of rows within the input rows.
Table	Produces a collection of rows (either a nested table or a vararray) that can be queried like a physical database table. Use in the FROM clause of a query.
System	Used to execute action in the system, or return information about the system.
User-Defined	Written to modularize queries or portions of queries.

jiosegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake
Do not copy

© 2023 Snowflake Computing Inc. All Rights Reserved

SOME SYSTEM FUNCTIONS

CONTROL

- CANCEL_QUERY
- CANCEL_ALL_QUERIES
- GLOBAL_ACCOUNT_SET_PARAMETER
- TASK_DEPENDENTS_ENABLE
- WAIT

INFORMATION

- ALLOWLIST
- ALLOWLIST_PRIVATELINK
- GET_TAG
- TYPEOF
- LAST_CHANGE_COMMIT_TIME

josegonzalezv@santander.com.mx - 14-SEP-2023 - ©Snowflake
2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

LAB EXERCISE: 6

Snowflake Functions

35 minutes

Learning Objectives:

- Apply scalar functions
- Apply conditional functions such as IFF() and CASE
- Apply aggregate functions such as MIN, MAX, SUM, AVG and MEDIAN
- Apply window functions
- Apply date/context functions
- Apply table functions
- Apply table functions

Note: This lab has some exercises where you'll apply functions to queries you write from scratch. Solutions are presented on the last page of this lab in the workbook.

josegonzalezv@santander.com.mx 11-SEP-2023 - ©Snowflake
2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

ESTIMATION FUNCTIONS

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake
2023 do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

APPROXIMATION

- Snowflake uses *HyperLogLog* for a set of aggregate functions to estimate the approximate number of distinct values in a data set

- *HyperLogLog* is a state-of-the-art cardinality estimation algorithm, capable of estimating distinct cardinalities of trillions of rows with an average relative error of a few percent

```
SELECT COUNT(DISTINCT o_custkey) AS actual,
APPROX_COUNT_DISTINCT(o_custkey) AS estimated
FROM orders;
```

Objects | Editor | Results | Chart

	ACTUAL	ESTIMATED
1	99,999,998	99,619,656

Estimate is off by .38%

josegonzalezv@santander.com.mx - 1 SEP 2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

COUNTING LARGE DISTINCT DATASETS

COUNT DISTINCT

- Exact
- Memory-intensive
- May spill to disk or remote storage

```
SELECT COUNT(DISTINCT l_orderkey)
FROM lineitem;
```

	··· COUNT(DISTINCT L_ORDERKEY)
1	1,500,000,000

Query duration 2m 17s

APPROX_COUNT_DISTINCT

- Approximate with small margin of error
- Requires less memory

```
SELECT APPROX_COUNT_DISTINCT(l_orderkey)
FROM lineitem;
```

	APPROX_COUNT_DISTINCT(L_ORDERKEY)
1	1,491,111,415

Query duration 25s

josegonzalezv@santander.com.mx - 11-SEP-2023 - @Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

ESTIMATION VS EXACT

MEDIAN

- Exact
- Memory-intensive

```
SELECT MEDIAN(ss_sales_price)
```

	MEDIAN(SS_SALES_PRICE)
1	27.59

Query duration 22m 37s

APPROX_PERCENTILE

- Small margin of error
- Uses a constant amount of memory regardless of the size of the input

```
SELECT
  APPROX_PERCENTILE(ss_sales_price, 0.5)
FROM store_sales;
```

	APPROX_PERCENTILE(SS_SALES_PRICE, 0.5)
1	27.593369198

Query duration 4m 28s

josegonzalezv@santander.com.mx - 14 SEP 2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

FREQUENCY ESTIMATION

```
APPROX_TOP_K ( <expr> [ , <k> [ , <counters> ] ] )
```
Returns estimated frequency of most frequent values

```
APPROX_TOP_K_ACCUMULATE( <expr> , <counters> )
```
Skips the final estimation step and returns the Space-Saving state at the end of an aggregation

```
APPROX_TOP_K_COMBINE( <state> [ , <counters> ] )
```
Combines (i.e., merges) input states into a single output state

```
APPROX_TOP_K_ESTIMATE( <state> [ , <k> ] )
```
Computes a cardinality estimate of a Space-Saving state produced by `APPROX_TOP_K_ACCUMULATE` and `APPROX_TOP_K_COMBINE`

josegonzalezv@santander.com.mx - 14-SEP-2023 - ©Snowflake
2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

SIMILARITY ESTIMATION

MINHASH(<k> , [distinct] expr+)
Estimates the approximate similarity between two or more data sets.

MINHASH_COMBINE([distinct] <state>])
Combines input MinHash states into a single MinHash output state.
This Minhash state can then be input to the `APPROXIMATE_SIMILARITY` function to estimate the similarity with other MinHash states.

APPROXIMATE_SIMILARITY ([distinct] <expr> [,...])
Returns an estimation of the similarity (Jaccard index) of inputs based on their MinHash states.
Returns a value between 0.0 (no similarity) to 1.0 (identical).

Equivalent to **APPROXIMATE_JACCARD_INDEX**.

jose.gonzalezv@santander.com - SEP-2023 - ©Snowflake
2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

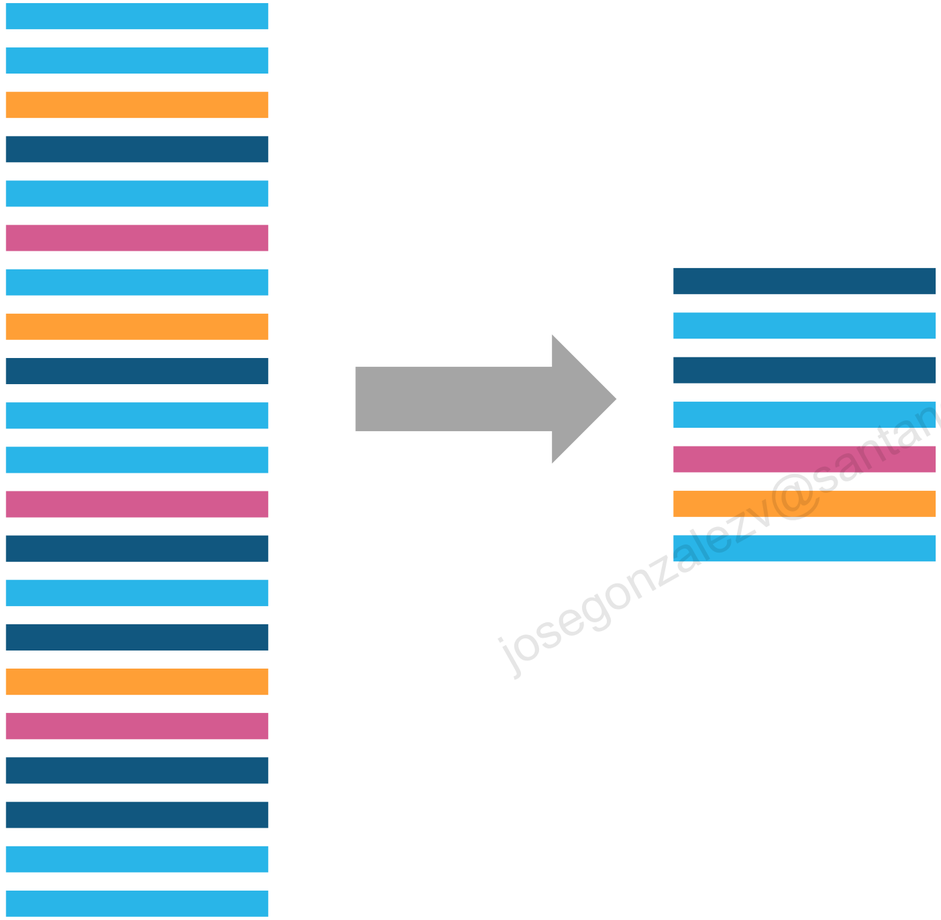
SAMPLING

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

SAMPLING

REPRESENTATIVE SAMPLE OF DATA SET



- Use the `SAMPLE` or `TABLESAMPLE` clause to create a sample data set
- Preferred over `RANDOM` or `LIMIT` as a sampling method
  - Creates a more representative sample
  - Faster than `RANDOM` in block mode
- Has knowledge of sampling and built-in optimizations
- Sampling the result of a JOIN is possible under certain conditions

josegonzalezv@santander.com.mx - 11-SEP-2023 - @Snowflake
2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

SAMPLE CLAUSE PARAMETERS

```
SELECT *
FROM usa_housing SAMPLE [<type>] (<size>) [SEED (<value>)];
```

Parameter	Description	Default
`<type>`	Specify `ROW` to evaluate each row individually Specify `BLOCK` to evaluate a block of rows at once	`ROW`
`<size>`	Either enter a number that specifies the percentage chance each row or block has of being in the sample, or specify the number of rows to return in the sample	none
`SEED (<value>)`	Enter a seed value to make the results deterministic	none

jose.gonzalez@santander.com.mx 11-SEP-2023 - ©Snowflake 2023 - do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

SAMPLING EXAMPLES

```
SELECT * FROM lineitem SAMPLE (10);

SELECT * FROM lineitem SAMPLE BLOCK (10);

SELECT * FROM orders SAMPLE (5) SEED (4444);

SELECT * FROM customer SAMPLE (10000 rows);

SELECT AVG(price) FROM (SELECT price FROM lineitem SAMPLE (5));
```

josegonzalez@santander.com.mx - 14-SEP-2023 - ©Snowflake 2023 - do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

LAB EXERCISE: 7

Using Estimation Functions

10 minutes

Learning Objectives:

- How to leverage HyperLogLog to get an approximate count
- How to leverage APPROX_PERCENTILE to get an approximate median

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

COLLATIONS

josegonzalezv@santander.com - 11-SEP-2023 - ©Snowflake 2023-do not copy

© 2023 Snowflake Computing Inc. All Rights Reserved

COLLATION

- Collate: to collect and combine in a specific order
  - Desired order is open to interpretation
- With Snowflake, strings are stored internally in UTF-8
  - UTF-8 uses the numeric representation of characters to sort
  - Upper-case letter have a lower numeric representation than lower-case
  - Example:

    Anne

    Bob

    Jerry

    charles

    deborah

    klaus

josegonzalezv@santander.com.mx - 14-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

SUPPORTED COLLATIONS

SPECIFY HOW TEXT SHOULD BE SORTED/COMPARED

Collation	Use	Examples
Language locale	Language- and country-specific rules to apply	`en, en_US, fr, fr_CA, etc.`
Case sensitivity	Whether to consider case when comparing values	`cs` or `ci`
Accent sensitivity	Whether to consider accented characters equal to, or different from, their base characters	`as` or `ai`
Punctuation sensitivity	Whether non-letter characters matter	`ps` or `pi`
First letter preference	Whether to sort upper-case or lower-case letters first	`fl` or `fu`
Case conversion	Convert to upper or lower case before comparisons	`upper` or `lower`
Space trimming	Remove leading spaces, trailing spaces, or both	`trim, ltrim, rtrim`

josegonzalezv@santander.com for 11-SEP-2023 - ©Snowflake 2023-do not share or copy

© 2023 Snowflake Computing Inc. All Rights Reserved

SPECIFYING COLLATIONS IN ORDER BYs

```
CREATE TABLE test (col1 VARCHAR);

INSERT INTO test VALUES
   ('Apple’),
   ('apple'),
   ('AB CD'),
   ('ABC'),
   ('pino'),
   ('piñata'),
   ('ab\'cd'),
   ('abc');
```

```
SELECT col1 FROM test;
```

unsorted

COL1
Apple
apple
AB CD
ABC
pino
piñata
ab'cd
abc

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

SPECIFYING COLLATIONS IN ORDER BYs

```
CREATE TABLE test (col1 VARCHAR);

INSERT INTO test VALUES
   ('Apple’),
   ('apple'),
   ('AB CD'),
   ('ABC'),
   ('pino'),
   ('piñata'),
   ('ab\'cd'),
   ('abc');
```

```
SELECT col1 FROM test;
```

```
SELECT col1 FROM test ORDER BY col1;
```

unsorted

COL1
Apple
apple
AB CD
ABC
pino
piñata
ab'cd
abc

sorted by utf8

COL1
AB CD
ABC
Apple
ab'cd
abc
apple
pino
piñata

jsegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

SPECIFYING COLLATIONS IN ORDER BYs

```
CREATE TABLE test (col1 VARCHAR);

INSERT INTO test VALUES
   ('Apple’),
   ('apple'),
   ('AB CD'),
   ('ABC'),
   ('pino'),
   ('piñata'),
   ('ab\'cd'),
   ('abc');
```

```
SELECT col1 FROM test;
```

```
SELECT col1 FROM test ORDER BY col1;
```

```
SELECT col1 FROM test ORDER BY COLLATE(col1, 'fr');
```

unsorted	sorted by utf8	sorted by french
COL1	COL1	COL1
Apple	AB CD	AB CD
apple	ABC	ab'cd
AB CD	Apple	abc
ABC	ab'cd	ABC
pino	abc	apple
piñata	apple	Apple
ab'cd	pino	piñata
abc	piñata	pino

jsegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake
2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

SPECIFYING COLLATIONS IN ORDER BYs

```
CREATE TABLE test (col1 VARCHAR);

INSERT INTO test VALUES
   ('Apple’),
   ('apple'),
   ('AB CD'),
   ('ABC'),
   ('pino'),
   ('piñata'),
   ('ab\'cd'),
   ('abc');
```

```
SELECT col1 FROM test;
```

```
SELECT col1 FROM test ORDER BY col1;
```

```
SELECT col1 FROM test ORDER BY COLLATE(col1, 'fr');
```

```
SELECT col1 FROM test ORDER BY COLLATE(col1, 'es');
```

unsorted	sorted by utf8	sorted by french	sorted by spanish
COL1	COL1	COL1	COL1
Apple	AB CD	AB CD	AB CD
apple	ABC	ab'cd	ab'cd
AB CD	Apple	abc	abc
ABC	ab'cd	ABC	ABC
pino	abc	apple	apple
piñata	apple	Apple	Apple
ab'cd	pino	piñata	pino
abc	piñata	pino	piñata

jsegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake
2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

SPECIFYING COLLATIONS ON COLUMNS

```
CREATE TABLE test (c_fr VARCHAR COLLATE 'fr', c_es VARCHAR COLLATE 'es');

INSERT INTO test VALUES
   ('Apple',  'Apple'),
   ('apple',  'apple'),
   ('AB CD',  'AB CD'),
   ('ABC',    'ABC'),
   ('pino',   'pino'),
   ('piñata', 'piñata'),
   ('ab\'cd', 'ab\'cd'),
   ('abc',    'abc');
```

```
SELECT * FROM test
ORDER BY c_fr;
```

C_FR	C_ES
AB CD	AB CD
ab'cd	ab'cd
abc	abc
ABC	ABC
apple	apple
Apple	Apple
piñata	piñata
pino	pino

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

SPECIFYING COLLATIONS ON COLUMNS

```
CREATE TABLE test (c_fr VARCHAR COLLATE 'fr', c_es VARCHAR COLLATE 'es');

INSERT INTO test VALUES
   ('Apple',  'Apple'),
   ('apple',  'apple'),
   ('AB CD',  'AB CD'),
   ('ABC',    'ABC'),
   ('pino',   'pino'),
   ('piñata', 'piñata'),
   ('ab\'cd', 'ab\'cd'),
   ('abc',    'abc');
```

```
SELECT * FROM test
ORDER BY c_fr;
```

C_FR	C_ES
AB CD	AB CD
ab'cd	ab'cd
abc	abc
ABC	ABC
apple	apple
Apple	Apple
piñata	piñata
pino	pino

```
SELECT * FROM test
ORDER BY c_es;
```

C_FR	C_ES
AB CD	AB CD
ab'cd	ab'cd
abc	abc
ABC	ABC
apple	apple
Apple	Apple
pino	pino
piñata	piñata

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

SPECIFYING COLLATIONS IN COMPARISONS

```
CREATE TABLE test (c1 VARCHAR, c2 VARCHAR);

INSERT INTO test VALUES
   ('its', 'it\'s'),
   ('log in', 'login'),
   ('MyTable', 'mytable');
```

C1	C2
its	it's
log in	login
MyTable	mytable

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

SPECIFYING COLLATIONS IN COMPARISONS

```
CREATE TABLE test (c1 VARCHAR, c2 VARCHAR);

INSERT INTO test VALUES
   ('its', 'it\'s'),
   ('log in', 'login'),
   ('MyTable', 'mytable');
```

C1	C2
its	it's
log in	login
MyTable	mytable

```
SELECT * FROM test
WHERE c1=c2;
```

C1	C2

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not copy

© 2023 Snowflake Computing Inc. All Rights Reserved

SPECIFYING COLLATIONS IN COMPARISONS

```
CREATE TABLE test (c1 VARCHAR, c2 VARCHAR);

INSERT INTO test VALUES
   ('its', 'it\'s'),
   ('log in', 'login'),
   ('MyTable', 'mytable');
```

C1	C2
its	it's
log in	login
MyTable	mytable

```
SELECT * FROM test
WHERE c1=c2;
```

C1	C2

```
SELECT * FROM test
WHERE c1=c2 COLLATE 'en-ci';
```

C1	C2
MyTable	mytable

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake
2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

SPECIFYING COLLATIONS IN COMPARISONS

```
CREATE TABLE test (c1 VARCHAR, c2 VARCHAR);

INSERT INTO test VALUES
   ('its', 'it\'s'),
   ('log in', 'login'),
   ('MyTable', 'mytable');
```

C1	C2
its	it's
log in	login
MyTable	mytable

```
SELECT * FROM test
WHERE c1=c2;
```

C1	C2

```
SELECT * FROM test
WHERE c1=c2 COLLATE 'en-ci';
```

C1	C2
MyTable	mytable

```
SELECT * FROM test
WHERE c1=c2 COLLATE 'en-ci-pi';
```

C1	C2
its	it's
log in	login
MyTable	mytable

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

PERFORMANCE IMPACTS OF USING COLLATION

- Micro-partition pruning can be less efficient
- Operations involving comparisons might be slower
- `WHERE` clauses, joins, sorts and `GROUP BY`s can be impacted

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

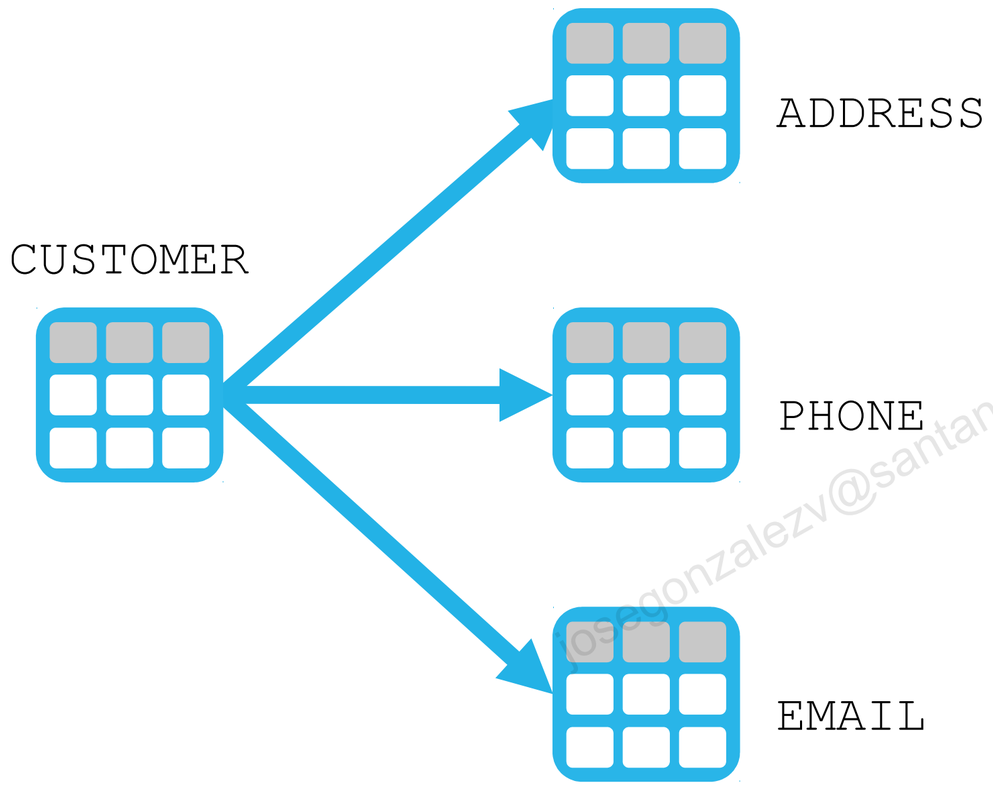
© 2023 Snowflake Computing Inc. All Rights Reserved

MULTI-TABLE INSERTS

josegonzalezv@samanet.com.mx - 14-SEP-2023 - ©Snowflake
2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

MULTI-TABLE INSERT



- Update multiple tables from a single table, in a single query
- Inserts can be unconditional or conditional

josegonzalezv@santander.com.mx - 11-SEP-2023 @Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

UNCONDITIONAL INSERT

- INTO statements designate target tables
- INSERT OVERWRITE truncates tables prior to insert
- Scenario: Writing customer info into separate address, phone, and email tables

```
INSERT ALL
INTO address (city, state, zip) VALUES (cust_city, cust_st, cust_zip)
INTO phone (area_code, phone) VALUES (cust_area_code, cust_phone)
INTO email (city, email) VALUES (cust_city, cust_email)
SELECT cust_city, cust_st, cust_zip, cust_area_code, cust_phone, cust_email
FROM customer_details;
```

jcgonzalezv@santander.com.mx - SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

CONDITIONAL INSERT

- INSERT ALL: evaluate all WHEN clauses
- INSERT FIRST: executes only on the first WHEN clause that evaluates to true
- Scenario: Separating reward members into platinum, gold, and silver

```
INSERT FIRST
  WHEN points >= 500000 THEN
    INTO platinum_members VALUES (member_id, points)
  WHEN points >= 100000 THEN
    INTO gold_members VALUES(member_id, points)
  ELSE
    INTO silver_members VALUES(member_id, points)
SELECT member_id, points
FROM members;
```

jose.gonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

LAB EXERCISE: 8

Multi-Table Inserts

25 minutes

Learning Objectives:

- How to use sequences to create unique values in a primary key column
- How to use unconditional multi-table insert statements
- How to use ALTER TABLE <table_name> SWAP WITH to swap table content and metadata
- How to use MERGE statements to add new rows to a table

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake
2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

QUERY TAGS

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

USE QUERY TAGS

1. Set up a query tag
   - `ALTER SESSION SET QUERY_TAG = 'ProjectA';`

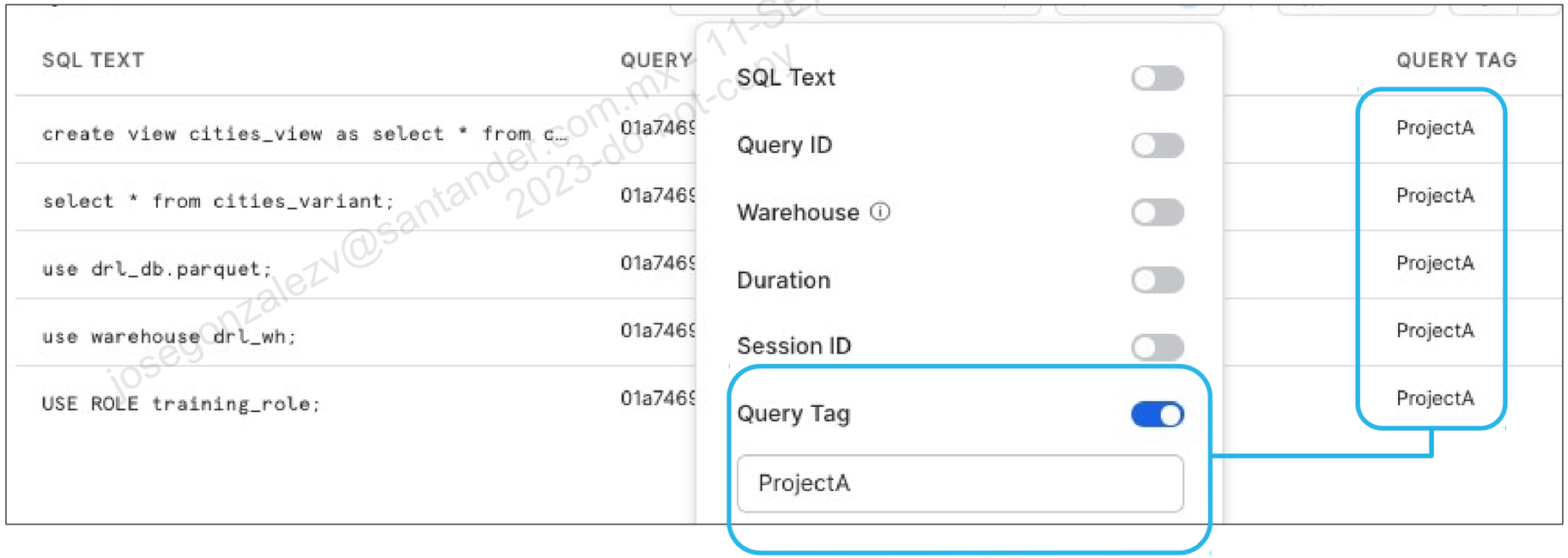
josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

USE QUERY TAGS

1. Set up a query tag
   - `ALTER SESSION SET QUERY_TAG = 'ProjectA';`

2. Filter History by query tag:



josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

LOCATE TAGGED QUERIES

```
SELECT query_id, query_type, user_name, query_tag
FROM SNOWFLAKE.ACCOUNT_USAGE.QUERY_HISTORY
WHERE query_tag = 'ProjectA';
```

QUERY_ID	QUERY_TYPE	USER_NAME	QUERY_TAG
01a74694-0604-dcad-0000-71e9004b5046	SHOW	DEBORAH	ProjectA
01a74694-0604-e137-0000-71e9004b6042	SHOW	DEBORAH	ProjectA
01a74694-0604-e137-0000-71e9004b6046	SHOW	DEBORAH	ProjectA
01a74694-0604-e137-0000-71e9004b604a	SHOW	DEBORAH	ProjectA
01a74694-0604-e137-0000-71e9004b604e	SHOW	DEBORAH	ProjectA

josegonzalezv@salesforce.com.mx - 11-SEP-2023 - ©Snowflake
2023-do-not-copy

- Clear tag with `ALTER SESSION UNSET query_tag;`

© 2023 Snowflake Computing Inc. All Rights Reserved

WORKING WITH PARAMETERS

josegonzalezv@santander.com.mx - 14-SEP-2023 - ©Snowflake
2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

PARAMETERS AND LEVELS

- Parameters can be set to control certain functionality
- There are different levels at which parameters can be set
  - Account
  - Session
  - Object
- Some parameters can be set at multiple levels
  - Example: `DATA_RETENTION_TIME_IN_DAYS` for Time Travel

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

MULTI-LEVEL PARAMETERS

- All parameters can be set at the account level
  - Defaults set by `ACCOUNTADMIN` (some by `SECURITYADMIN` as well)
- Most parameters can be set at multiple levels
- The lowest level wins
  - A parameter set at the table level takes precedence over the same parameter set at the schema level
- The lowest value wins
  - FOR `STATEMENT_TIMEOUT_IN_SECONDS` and `STATEMENT_QUEUED_TIMEOUT_IN_SECONDS` both can be set at the session level and/or on a warehouse. If it is set on multiple levels, the lowest **value** wins.

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake
2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

PARAMETER HIERARCHY

Default values for all parameters are set at the account level

ACCOUNT parameters

Cannot be overridden

`NETWORK_POLICY`

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

PARAMETER HIERARCHY

Default values for all parameters are set at the account level

ACCOUNT parameters

Cannot be overridden

`NETWORK_POLICY`

SESSION parameters

Overridden at the USER level

Overridden as the SESSION level

`AUTOCOMMIT`

josegonzalez@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

PARAMETER HIERARCHY

Default values for all parameters are set at the account level

ACCOUNT parameters

Cannot be overridden

`NETWORK_POLICY`

SESSION parameters

Overridden at the USER level

Overridden as the SESSION level

`AUTOCOMMIT`

VIRTUAL WH parameters

Overridden at the WAREHOUSE level

`STATEMENT_TIMEOUT_IN_SECONDS`

josegonzalez@santander.com.mx - 11-SEP-2023 ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

PARAMETER HIERARCHY

Default values for all parameters are set at the account level

ACCOUNT parameters

Cannot be overridden

`NETWORK_POLICY`

SESSION parameters

Overridden at the USER level

Overridden as the SESSION level

`AUTOCOMMIT`

VIRTUAL WH parameters

Overridden at the WAREHOUSE level

`STATEMENT_TIMEOUT_IN_SECONDS`

DB/SCHEMA/TABLE parameters

Overridden at the DATABASE level

Overridden at the SCHEMA level

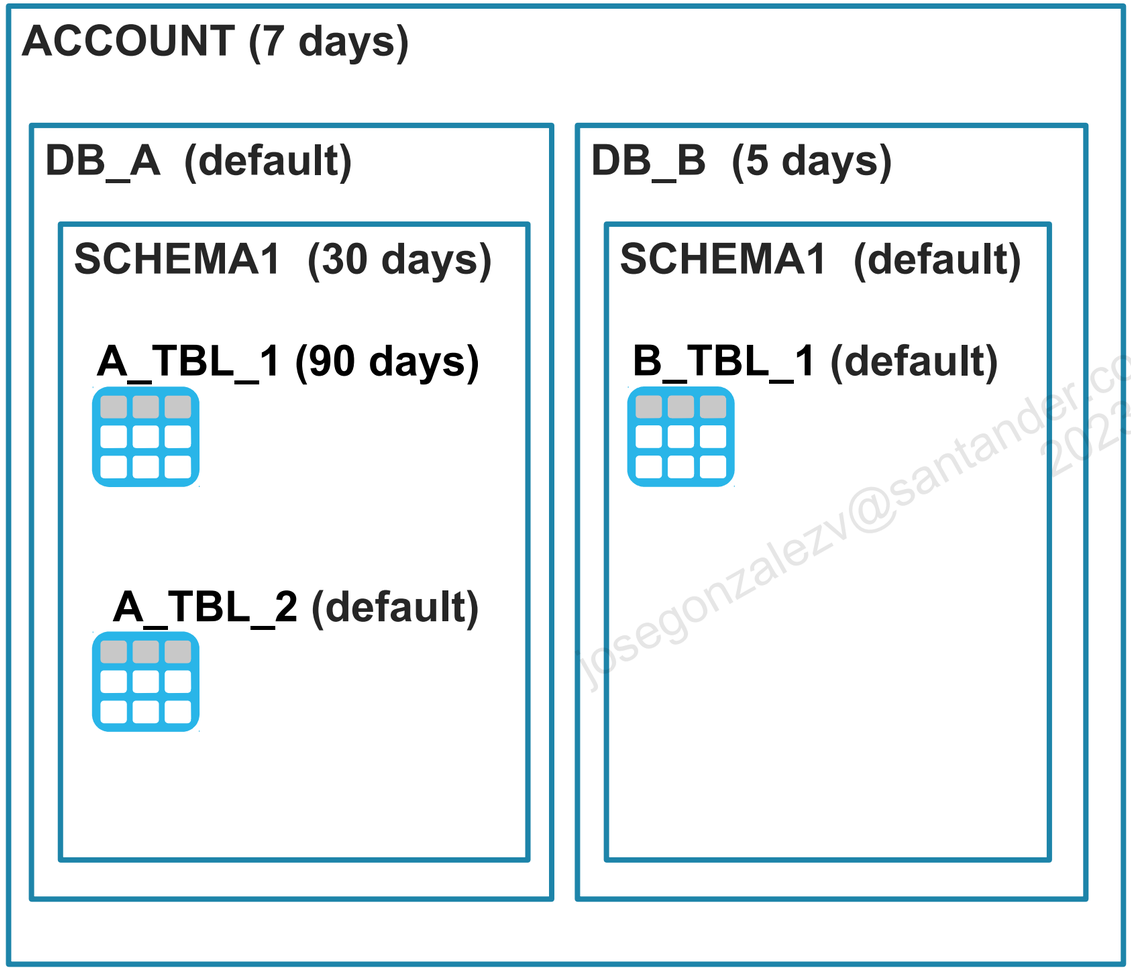
Overridden at the TABLE level

`DATA_RETENTION_TIME_IN_DAYS`

josegonzalez@santander.com.mx - 11-SEP-2023 ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

EXAMPLE:
TIME TRAVEL SETTING



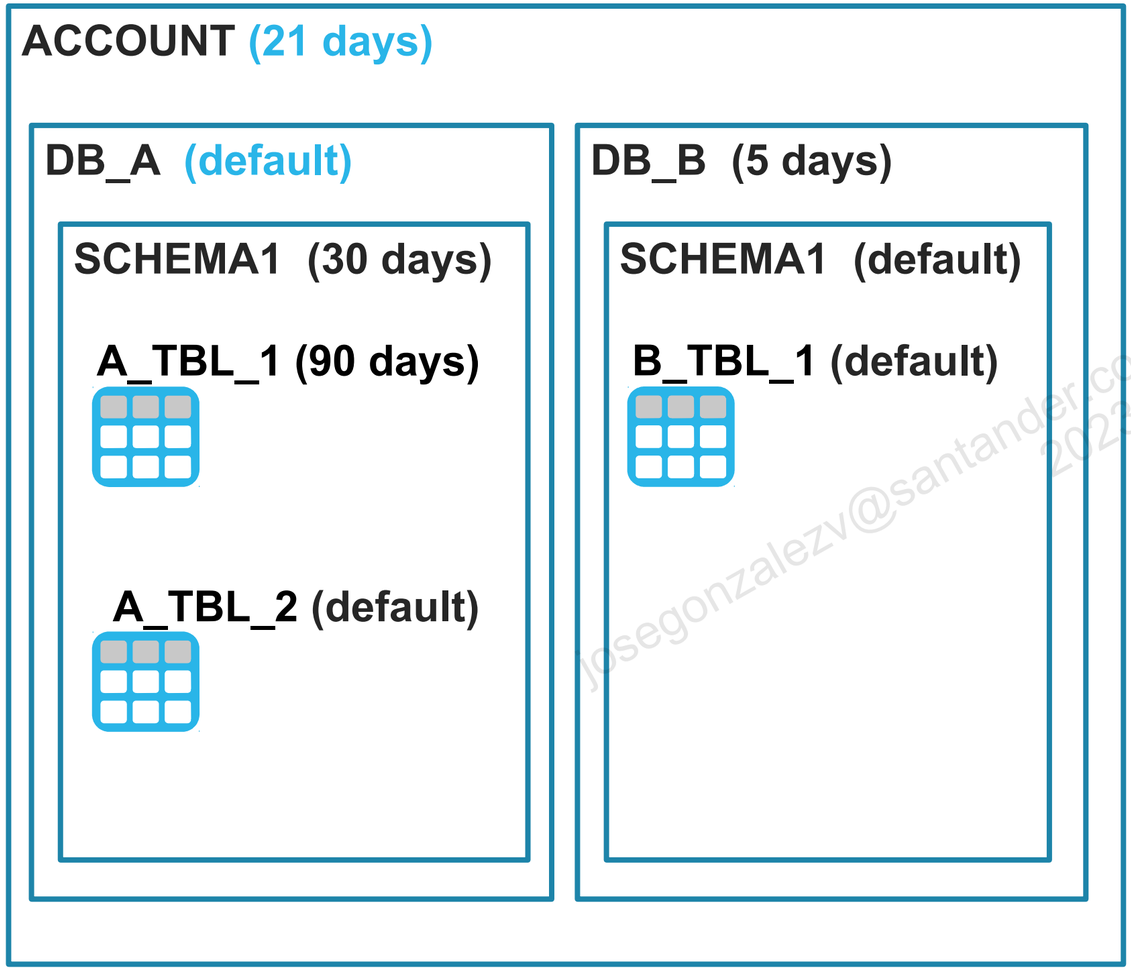
- `DATA_RETENTION_TIME_IN_DAYS` can be set at the account, database, schema, or table level
- If set at multiple levels, the setting closest to the object will be used
- In the example at left –
  - `A_TBL_2` will inherit its time travel setting from `SCHEMA1` (30 days)
  - `B_TBL_1` will inherit its time travel setting from `DB_B` (5 days)

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake
2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

EXAMPLE: TIME TRAVEL SETTING

DATA_RETENTION_TIME_IN_DAYS



- Changing a default value will affect all objects below it that are using the default
- In this example, if you change the account-level parameter to 21 days, only the setting for DB_A will be affected
  - A_TBL_1 is specifically set
  - A_TBL_2 inherits its value from SCHEMA1
  - B_TBL_1 Inherits its value from DB_B

josegonzalezv@santander.com.mx - 11-SEP-2023 @Snowflake
2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

VIEWING PARAMETERS

```
SHOW PARAMETERS FOR <object>;
```

```
5  SHOW PARAMETERS FOR SESSION;
```

Objects | Query | Results | Chart

	key	value	default
1	ABORT_DETACHED_QUERY	false	false
2	AUTOCOMMIT	true	true
3	AUTOCOMMIT_API_SUPPORTED	true	true
4	BINARY_INPUT_FORMAT	HEX	HEX

jsegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake
2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

CHANGING PARAMETERS

```
ALTER <object> SET <parameter> = <value>;
```

Examples:

```
ALTER SESSION
SET use_cached_result = false;

ALTER DATABASE my_db
SET data_retention_time_in_days = 5;

ALTER ACCOUNT
SET network_policy = account_policy;
```

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

POPULAR SESSION-LEVEL PARAMETERS

- Session parameters determine how date, time, and timestamp data is passed into/out of Snowflake
- Parameters can be set at the account, user, or session levels
- Input formats for date, time, and time stamp (default is AUTO):
  - `DATE_INPUT_FORMAT`
  - `TIME_INPUT_FORMAT`
  - `TIMESTAMP_INPUT_FORMAT`

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

POPULAR SESSION-LEVEL PARAMETERS

- Output Formats
  - `DATE_OUTPUT_FORMAT`
  - `TIME_OUTPUT_FORMAT`
  - `TIMESTAMP_[LTZ|NTZ|TZ]_OUTPUT_FORMAT`
- `TIMESTAMP_TYPE_MAPPING`
  - specifies which variation of `TIMESTAMP_*` is mapped to the `TIMESTAMP` data type
- `TIMEZONE` parameter
  - specifies the time zone (e.g., Europe/London, America/Los Angeles)

jsegongalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

POPULAR OBJECT-LEVEL PARAMETERS

- `DATA_RETENTION_TIME_IN_DAYS:`
  - Set at the account level
  - Set at the hierarchical object level
  - Table overrides schema, schema overrides database, and database overrides account

- `STATEMENT_TIMEOUT_IN_SECONDS, STATEMENT_QUEUED_TIMEOUT_IN_SECONDS:`
  - Set at the account level
  - Set at the warehouse level for individual warehouses

josegonzalezy@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023 - do not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

KNOWLEDGE CHECK

`DATA_RETENTION_TIME_IN_DAYS` is set to 1 at the account level, 30 on `DB_A`, and 15 on `DB_A.SchemaA`.

You create a table in `DB_A.SchemaA` without specifying a retention time. What will the table use?

A. 1 day

B. 15 days

C. 30 days

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

KNOWLEDGE CHECK

`DATA_RETENTION_TIME_IN_DAYS` is set to 1 at the account level, 30 on `DB_A`, and 15 on `DB_A.SchemaA`.

You create a table in `DB_A.SchemaA` without specifying a retention time. What will the table use?

A. 1 day

B. 15 days

C. 30 days

B. 15 days

If no time travel value is set on an object, it will inherit the value used by the closest enclosing object.

josegonzalezv@santander.com.mx - 14 SEP 2023 - ©Snowflake 2023-do-not-copy

KNOWLEDGE CHECK

Using the default collation method, in what order will the following values be listed in the results if you do an `ORDER BY` on the column?

Hot Dogs
Hot dogs
hamburgers
pizza
Mac 'n Cheese
anchovies

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

KNOWLEDGE CHECK

Using the default collation method, in what order will the following values be listed in the results if you do an `ORDER BY` on the column?

Hot Dogs
Hot dogs
hamburgers
pizza
Mac 'n Cheese
anchovies

Hot Dogs
Hot dogs
Mac 'n Cheese
anchovies
hamburgers
pizza

With UTF-8, upper-case letters will always be sorted before lower-case.

josegonzalezv@santander.com.mx - 11-SEP-2023 - @snowflake
2023-do-not-copy

CACHING IN SNOWFLAKE

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake
2023 do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

MODULE AGENDA

- Overview
- Metadata Cache
- Query Result Cache
- Data Cache

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

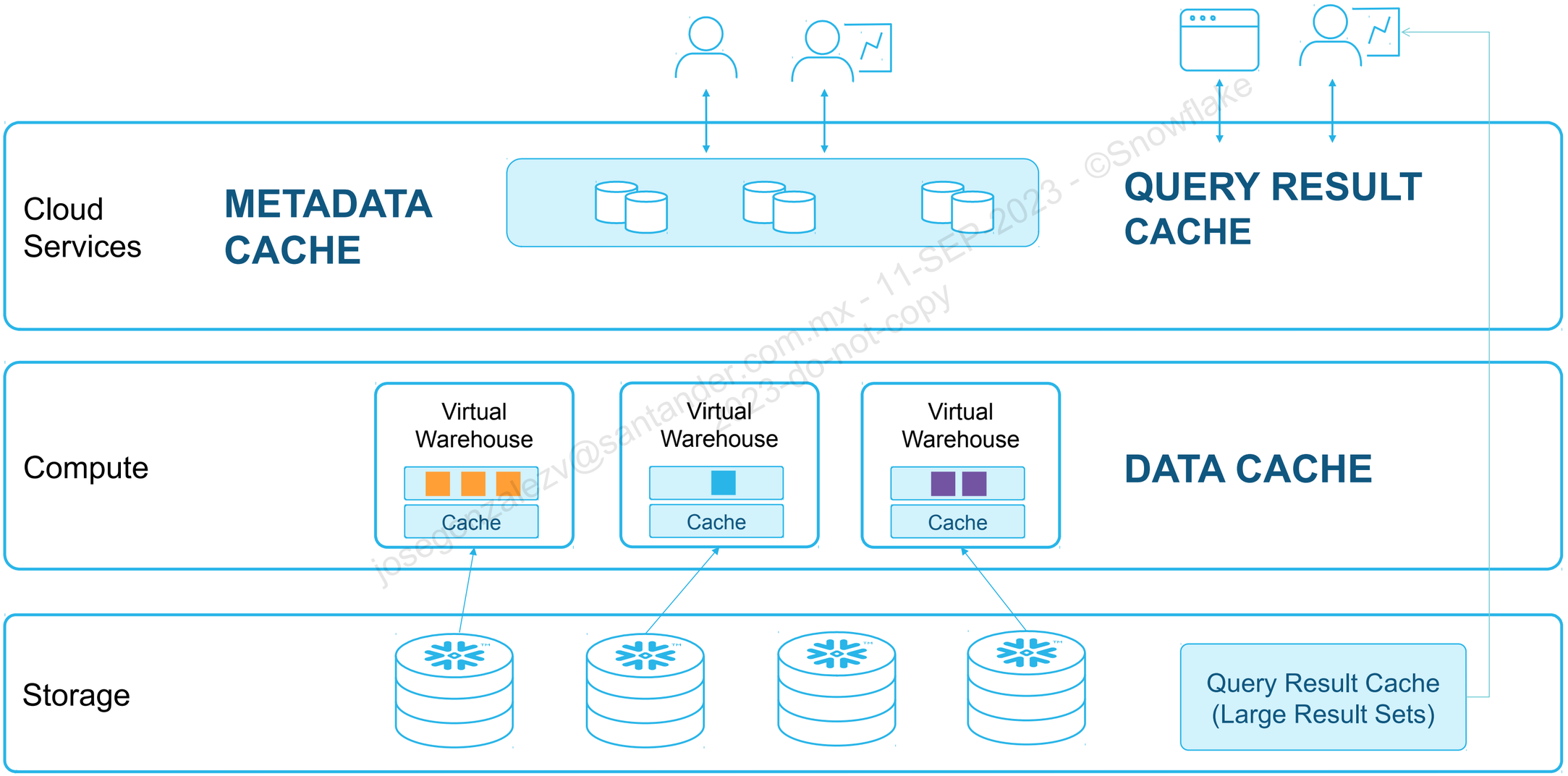
© 2023 Snowflake Computing Inc. All Rights Reserved

OVERVIEW

josegonzalezv@santander.com.br - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

CACHING IN SNOWFLAKE



josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

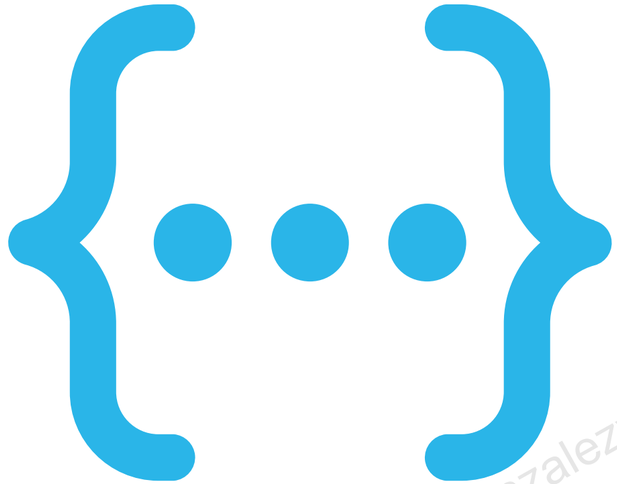
© 2023 Snowflake Computing Inc. All Rights Reserved

METADATA CACHE

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake
2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

METADATA CACHE



- Metadata stored in cloud services layer
- Micro-partition level metadata
  - Row count
- Micro-partition column-level metadata
  - MIN/MAX values
  - Number of DISTINCT values
  - Number of NULL values
- Table versions and references to physical files (.fdn)

josegonzalezv@santander.com.mx - 11-SEP-2023 ©Snowflake
2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

MICRO-PARTITION METADATA CACHE

- Used by SQL optimizer to speed up query compilation
- Used for queries that can be answered completely by metadata
  - SHOW commands
  - MIN, MAX (only for integer and date data types)
  - COUNT
- Fast
- No virtual warehouse used
  - NOTE: Cloud Services charges may still apply!

jose.gonzalezv@santander.com.mx - 11-SEP-2023 - @Snowflake
2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

METADATA CACHE - QUERY PROFILE

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023 do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

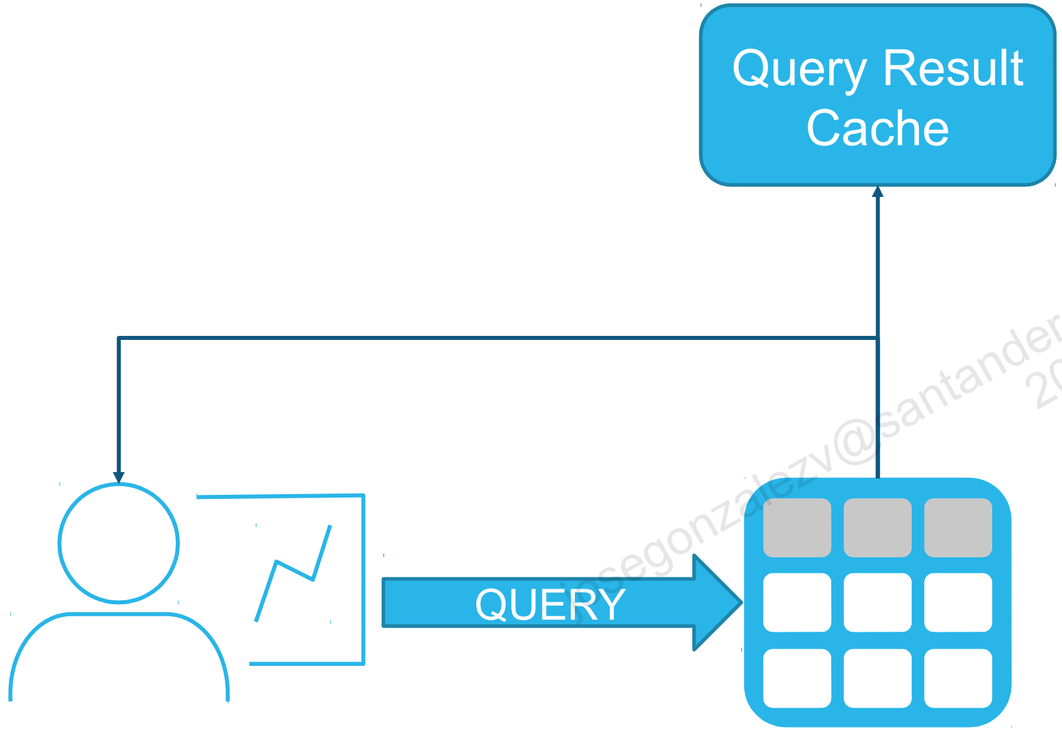
QUERY RESULT CACHE

josegonzalezv@santander.com.mx - 14-SEP-2023 - ©Snowflake
2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

QUERY RESULT CACHE

PERSISTED QUERY RESULTS



- Query results are stored and managed by the Cloud Services layer
- Used if the identical query is run, and base tables have not changed
- Available to other users:
  - SELECT: Any role that has SELECT permission on all the tables involved can use the QR cache

jesegonzalezv@santander.com.mx - 11-SEP-2023 - @Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

HOW IT WORKS

- Result sets are cached for 24 hours
  - Counter resets each time matching query is re-used
- Result reuse controlled by `USE_CACHED_RESULT`
- Eligibility requirements for query to use result set cache:
  - Exact same SQL query (except whitespace)
  - Result must be deterministic (no random function)
  - Changes can be made to source table(s) if they do not affect any micro-partitions relevant to query
  - User must have the right permissions to the objects in the query

jesgonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

QUERY RESULT CACHE USE CASES

- Any query which is run repeatedly but does not often change
  - Example: Static dashboards
- Refine the output of another query
  - Use TABLE function `RESULT_SCAN(<query id>);`
- Benefits
  - Fast
  - Will never give stale results
  - No virtual warehouse used (unless filtering results when using `RESULT_SCAN`)

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

QUERY RESULT CACHE EXAMPLE

Query profile confirms re-use of query result cache

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023 do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

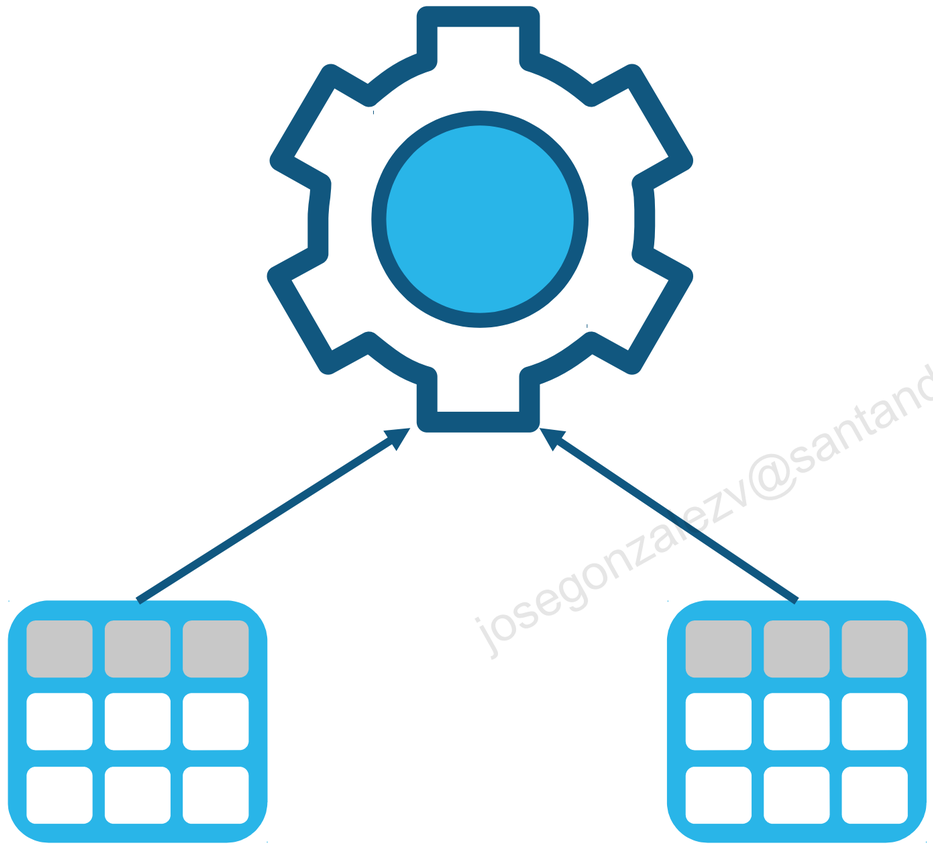
DATA CACHE

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake
2023-do not copy

© 2023 Snowflake Computing Inc. All Rights Reserved

DATA CACHE

DATA CACHED ON VIRTUAL WAREHOUSE



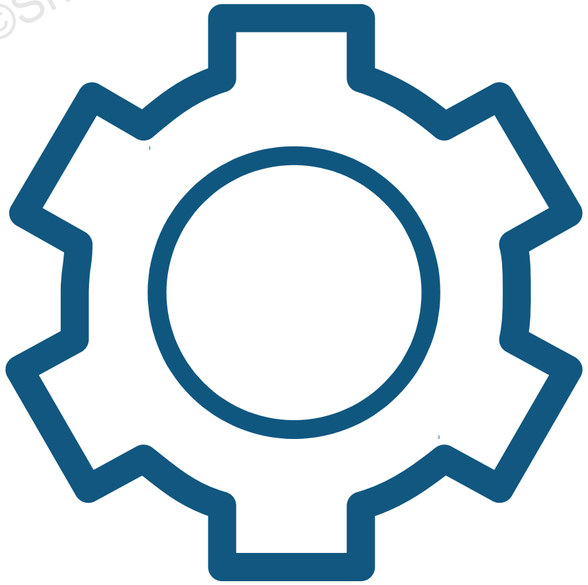
- Stores file headers and column data from queries
  - Stores the data, not the result
- Stored SSD in virtual warehouse
- When a similar query is run, Snowflake will use as much data from the cache as possible
- Available for all queries run on the same virtual warehouse

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

HOW DATA CACHE WORKS

- When a query is run, relevant file headers and column data retrieved are stored on SSD
- Virtual warehouse will first read any locally available data, then read remainder from remote cloud storage
- Data is flushed out in a Least Recently Used (LRU) fashion when cache fills



josegonzalezv@santander.com.mx - 14-SEP-2023 - ©Snowflake 2023 - do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

WHAT GETS CACHED?

Query Example:

```
SELECT firstname FROM customer WHERE state= 'CA';
```

1. Complier will determine what micro-partitions are needed for the query

```
SELECT firstname FROM customer WHERE state= 'CA';
```

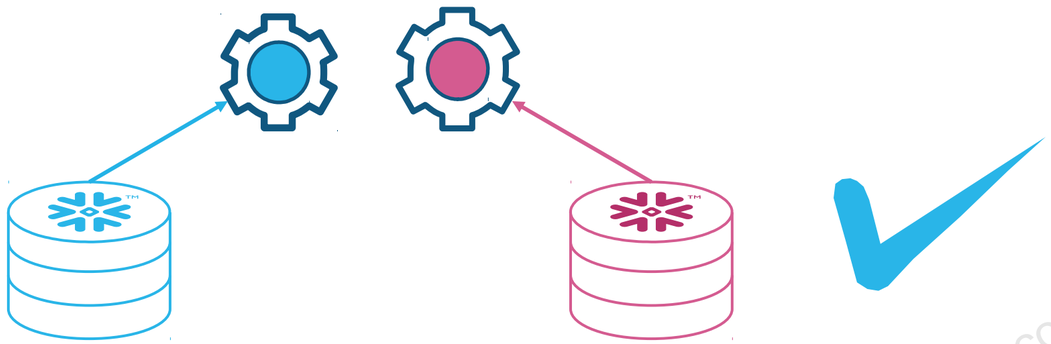
2. Only the required columns from those micro-partitions are copied to the virtual warehouse

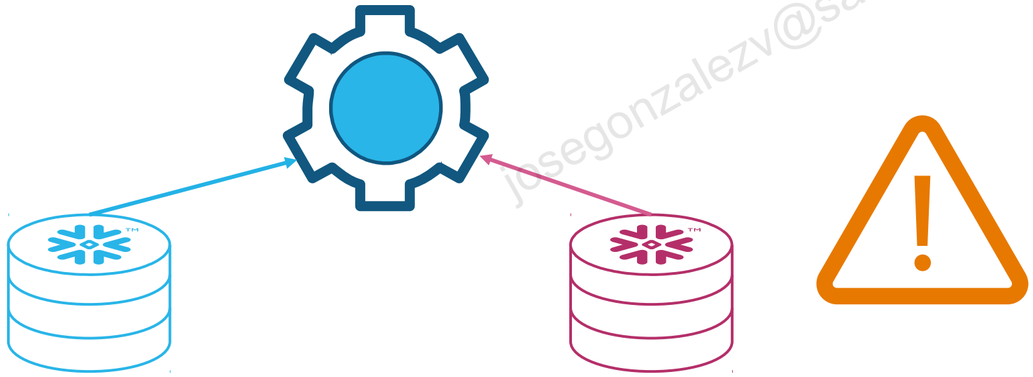
```
SELECT firstname FROM customer WHERE state= 'CA';
```

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

BEST PRACTICES





- Execute similar queries on the same virtual warehouse to maximize data cache reuse, for performance and cost optimization
- Do not suspend or auto-suspend virtual warehouses too quickly, as that will clear the data cache

josegonzalezv@santander.com.mx - 11-SEP-2023 - Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

DATA CACHE: COLD VS. WARM

Cold Virtual Warehouse

Profile Overview (Finished)

Total Execution Time	(2m 16s) 100.0%
Processing	78.1%
Local Disk I/O	0.3%
Remote Disk I/O	8.7%
Synchronization	0.1%
Initialization	12.7%

Statistics

Scan progress	100.00%
Bytes scanned	15.46GB
Percentage scanned from cache	0.00%

Warm Virtual Warehouse

Profile Overview (Finished)

Total Execution Time	(2.4s) 100.0%
Processing	10.3%
Synchronization	74.1%
Initialization	15.5%

Statistics

Scan progress	0.98%
Bytes scanned	916.43MB
Percentage scanned from cache	100.00%

2m16s versus 2.4s

Remote Disk I/O versus No Remote Disk I/O

0% in cache to 100% in cache

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do not copy

© 2023 Snowflake Computing Inc. All Rights Reserved

SUMMARY OF CACHE OPTIONS

	Metadata Cache	Query Result Cache	Data Cache
Where is it stored?	Cloud services layer	Cloud services and storage layers	Warehouse SSD
What does it store?	Metadata and statistics for micro-partitions and tables	Results set for each query	Data used in query
When is it used?	The optimizer uses this for commands like `MIN, MAX, COUNT`	Identical query run again and the base table data has not changed	Query using some or all the same data is run and the base table data has not changed
How long does it last?	Continuously updated	24 hours (reset after every use of a result in the cache)	Until the warehouse is suspended
Who can use it?	Everyone	For SELECT: user with SELECT permissions on all tables	Anyone who uses the same warehouse
Does it use a virtual warehouse?	No	No	Yes

josegonzalezv@santander.com.mx - 4-SEP-2023 - ©Snowflake
2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

KNOWLEDGE CHECK

What are the three types of cache?

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

KNOWLEDGE CHECK

What are the three types of cache?

Metadata cache

Query result cache

Data (warehouse) cache

josegonzalezv@santander.com.mx - 11-SEP-2023 - @Snowflake
2023-do-not-copy

KNOWLEDGE CHECK

You issue this query:

```
SELECT * FROM my_table;
```

Followed by this query:

```
SELECT * from my_table;
```

Which cache will help increase the speed of this query?

josegonzalez@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

KNOWLEDGE CHECK

You issue this query:

```
SELECT * FROM my_table;
```

Followed by this query:

```
SELECT * from my_table;
```

Which cache will help increase the speed of this query?

The data cache. The query result cache would not be used in this instance, because the queries are not identical (`FROM` versus `from`).

josegonzalez@santander.com.mx - 11-SEP-2023 @Snowflake
2023-do-not-copy

KNOWLEDGE CHECK

You run this query on a cold warehouse:

```
SELECT
   col1, col2, col3
FROM
   my_table
WHERE
  col1 > 128.02
```

What cache, if any, will be used?

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

KNOWLEDGE CHECK

You run this query on a cold warehouse:

```
SELECT
   col1, col2, col3
FROM
   my_table
WHERE
  col1 > 128.02
```

What cache, if any, will be used?

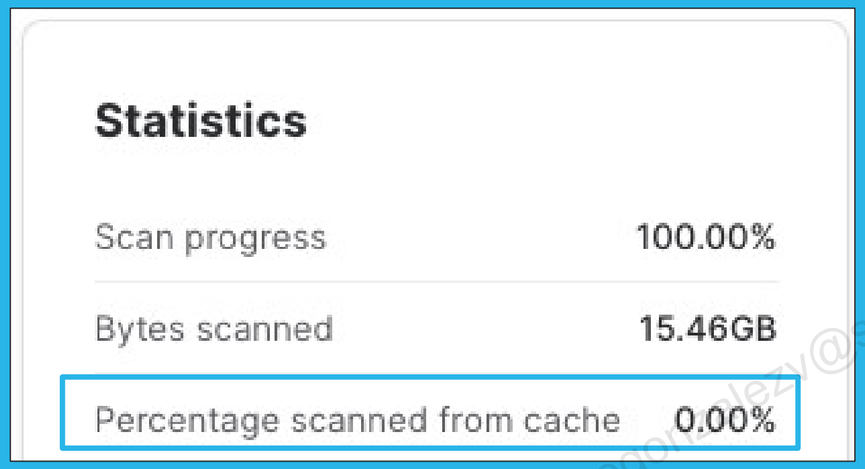
The metadata cache will be used by the query optimizer, to prune out micro-partitions that are not relevant to the query.

However, if the query was run previously, none of the micro-partitions relevant to the query were changed, the persisted result is still available, and the user's role has the right privileges to the objects in the query, the query results cache will be used instead.

josegonzalezv@santander.com.mx - 11-SEP-2023 - @SnowflakeDB
2023-do-not-copy

KNOWLEDGE CHECK

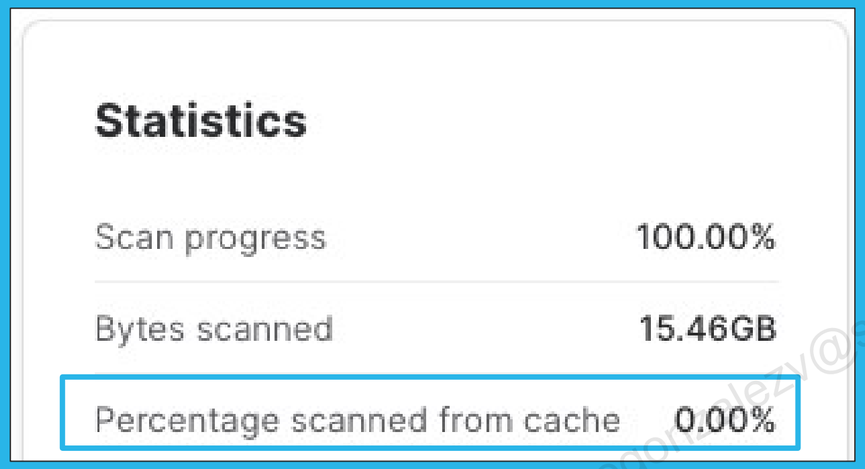
In the query profile shown below, which cache is being referred to?



josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

KNOWLEDGE CHECK

In the query profile shown below, which cache is being referred to?



"Percentage scanned from cache" refers to the data cache.

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

KNOWLEDGE CHECK

Rate each of the following statements as TRUE or FALSE:

A. Virtual warehouse costs are always decreased by setting a very short auto-suspend time.

B. The metadata cache is used in every query that has a `WHERE` clause.

C. Data cache use is all-or-nothing.

D. The timer on the query result cache resets every time someone uses the result.

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

KNOWLEDGE CHECK

Rate each of the following statements as TRUE or FALSE:

A. Virtual warehouse costs are always decreased by setting a very short auto-suspend time.

B. The metadata cache is used in every query that has a `WHERE` clause.

C. Data cache use is all-or-nothing.

D. The timer on the query result cache resets every time someone uses the result.

A. FALSE – shutting down the virtual warehouse clears the data cache, so shutting a warehouse down too quickly can reduce use of that cache

B. TRUE – even if the query cannot be pruned, the metadata is examined

C. FALSE – you can use as much of the data cache as is relevant to your query

D. TRUE – it resets to 24 hours

josegonzalezv@santander.com.mx - 11-SEP-2023 - @Snowflake 2023-do-not-copy

QUERY PERFORMANCE

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

MODULE AGENDA

- Using EXPLAIN
- Query Profile
- SQL Performance Tips

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

USING EXPLAIN

josegonzalezv@santander.com - MAT 11-SEP-2023 - ©Snowflake
2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

WHAT IS “EXPLAIN”?

```
EXPLAIN
SELECT
    l_orderkey
FROM
    lineitem
LIMIT 10;
```

- Displays the logical execution steps without executing
- Results show in JSON, text or tabular format
- Helpful for evaluating micro-partition pruning and joins

josegonzalezv@santander.com.mx - 11-SEP-2023 @Snowflake
2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

WHAT'S IN A PLAN?

KEY POINTS

- Partition Pruning
- Join Ordering
- Join Types

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

WHAT IS IN A PLAN?

MICRO-PARTITION PRUNING

operation		partitionsTotal	partitionsAssigned
GlobalStats		367	92
Result		null	null
Sort		null	null
InnerJoin		null	null
TableScan		1	1
InnerJoin	. . .	null	null
JoinFilter		null	null
TableScan		1	1
InnerJoin		null	null
Filter		null	null
TableScan		291	16
JoinFilter		null	null
TableScan		74	74

Total shown at top

One row for each table scan

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

WHAT IS IN A PLAN?

JOIN ORDERING AND JOIN TYPES

operation	objects
GlobalStats	null
Result	null
Sort	null
Aggregate	null
InnerJoin	null
TableScan	SNOWFLAKE_SAMPLE_DATA.TPCH_SF100.CUSTOMER
CartesianJoin	null

josegonzale2v@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

TOP USE CASE

VALIDATING PREDICATE (WHERE) MICRO-PARTITION PRUNING

```
SELECT
  COUNT(1)
FROM
  orders
WHERE
  o_orderdate >= to_date('2022-06-26');
```

Must scan only 154 of 3242 micro-partitions: predicate eliminates 95% of micro-partitions

ons	partitionsTotal	partitionsAssigned
	3,242	154
)	null	null
: [COUNT(*)]	null	null
O_ORDERDATE >= '1998-01-01'	null	null
DATE	3,242	154

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

QUERY PROFILE

josegonzalezv@santander.com DAY - 11-SEP-2023 - ©Snowflake
2023-do-not-copy

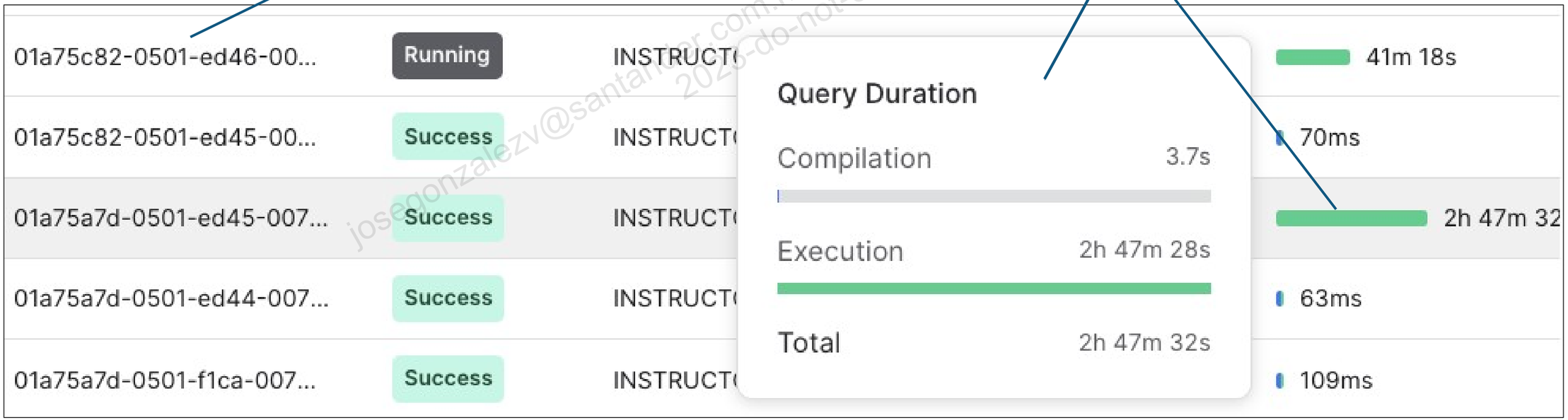
© 2023 Snowflake Computing Inc. All Rights Reserved

QUERY HISTORY

- **Activity > Query History**

Click Query ID to access Query Profile

Roll mouse over duration bar to see breakdown



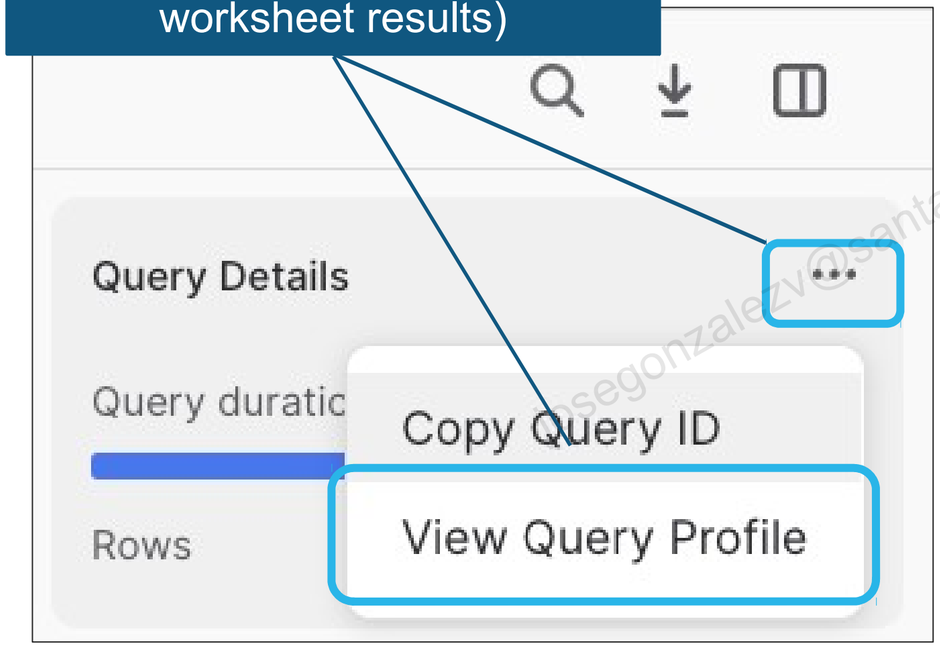
josegonzalezv@santander.com.mx - 11-SEP-2023 - Snowflake 2026 - do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

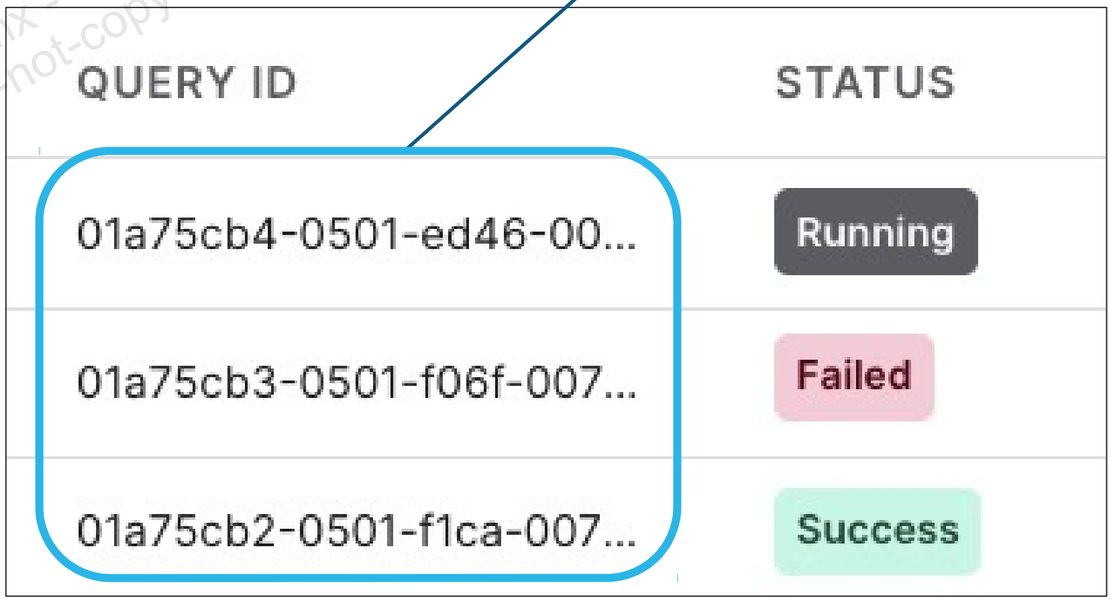
ACCESS THE QUERY PROFILE

- Multiple ways to access the Query Profile

Click ellipsis and select View Query Profile (to right of worksheet results)



Click Query ID from History



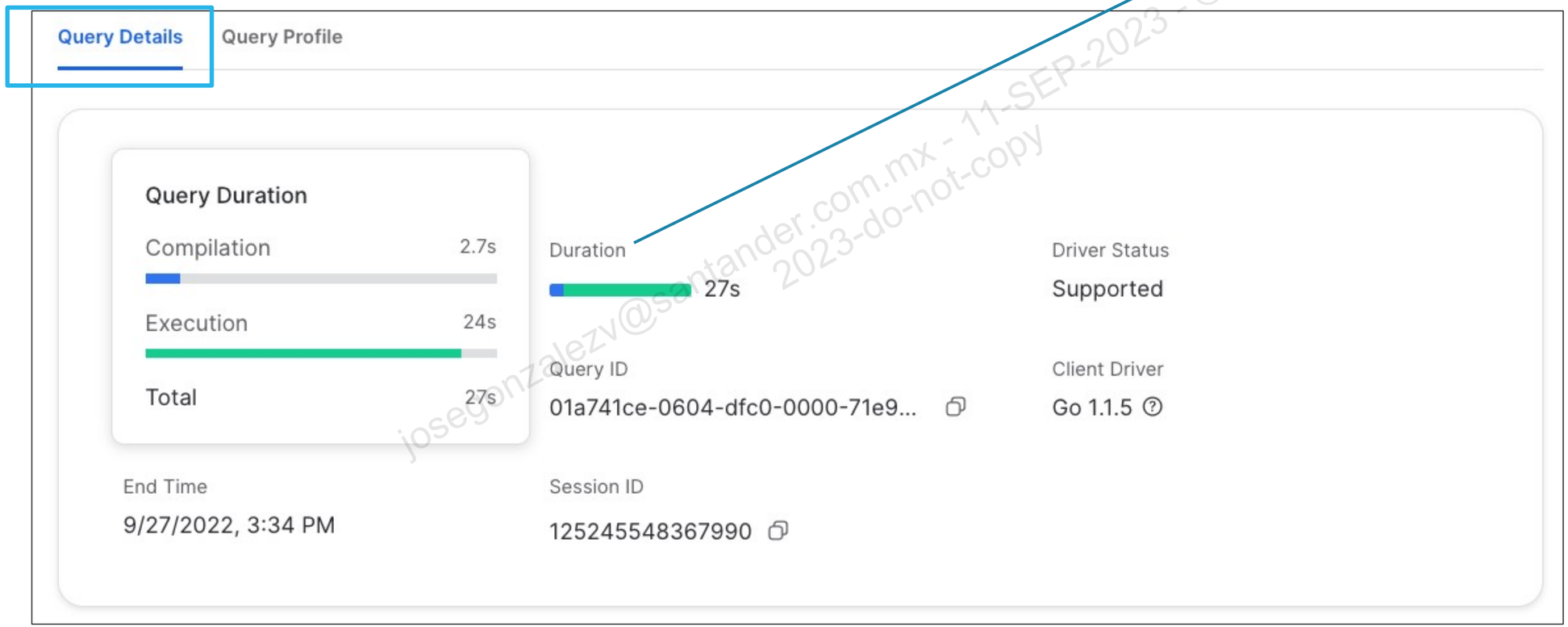
jsegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

QUERY DETAILS

- Selecting the Query Profile brings up this view:

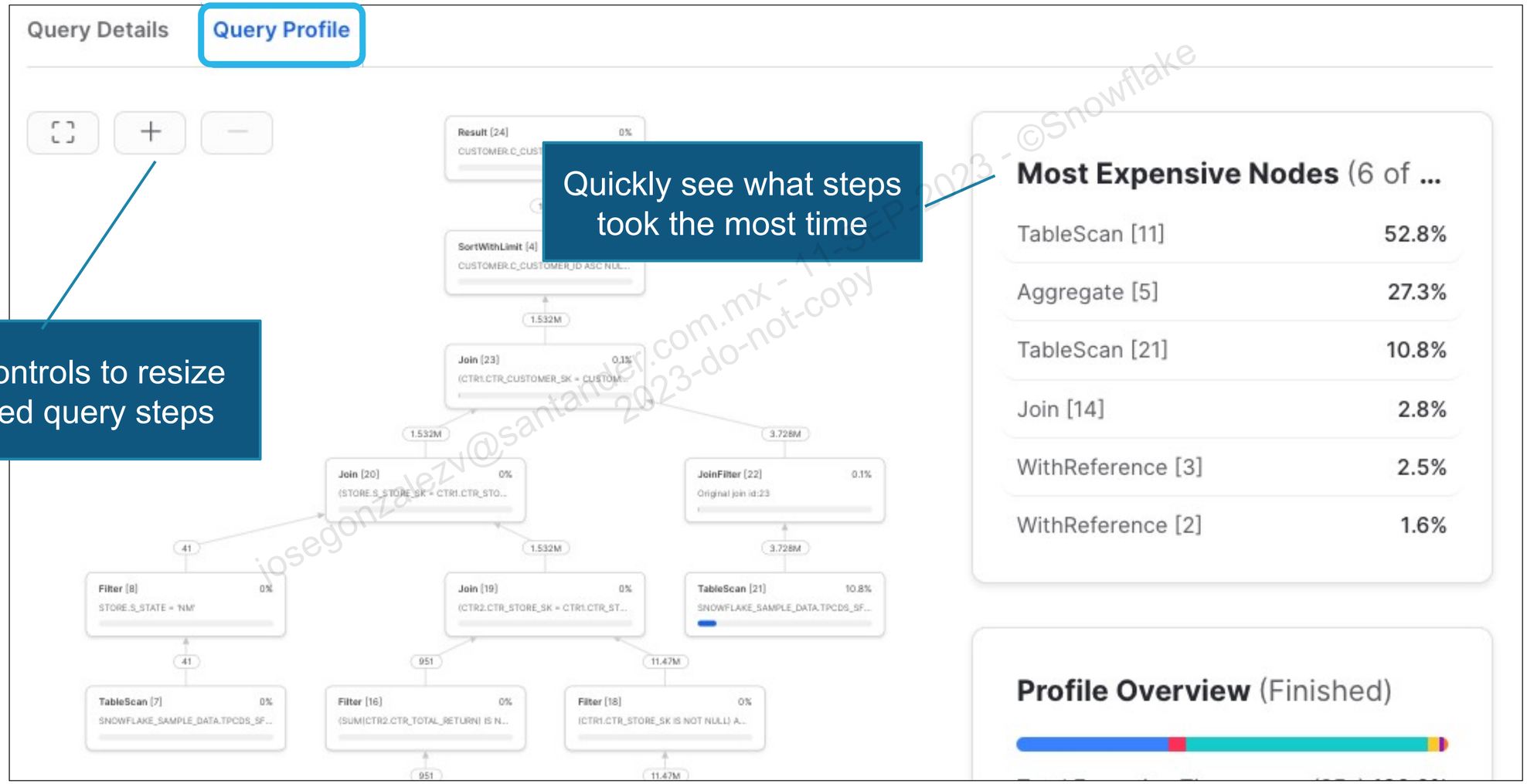
Roll the mouse over Query Duration to see details



josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

QUERY PROFILE



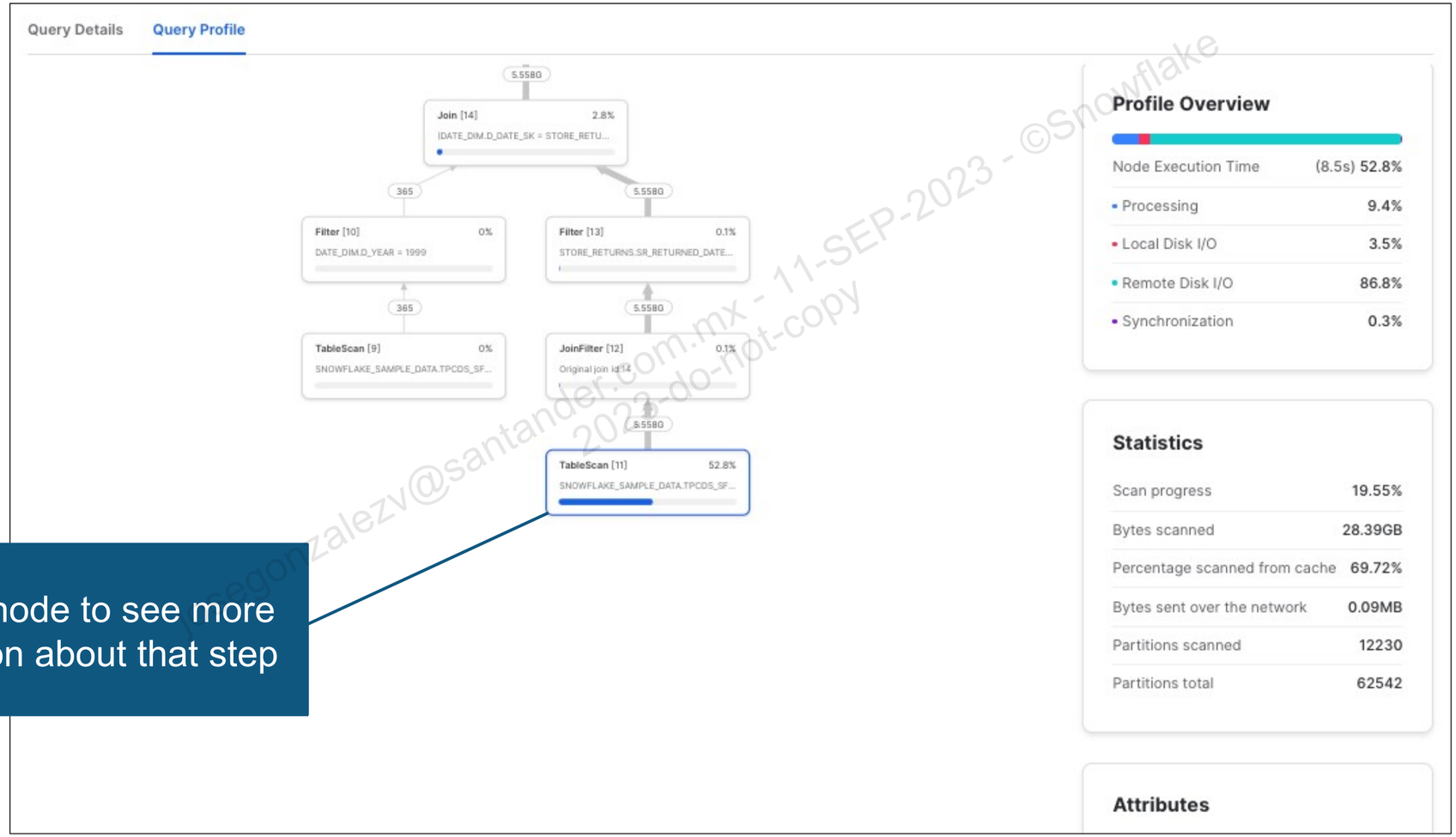
Quickly see what steps took the most time

Use the controls to resize the detailed query steps

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

NODE-SPECIFIC DETAILS

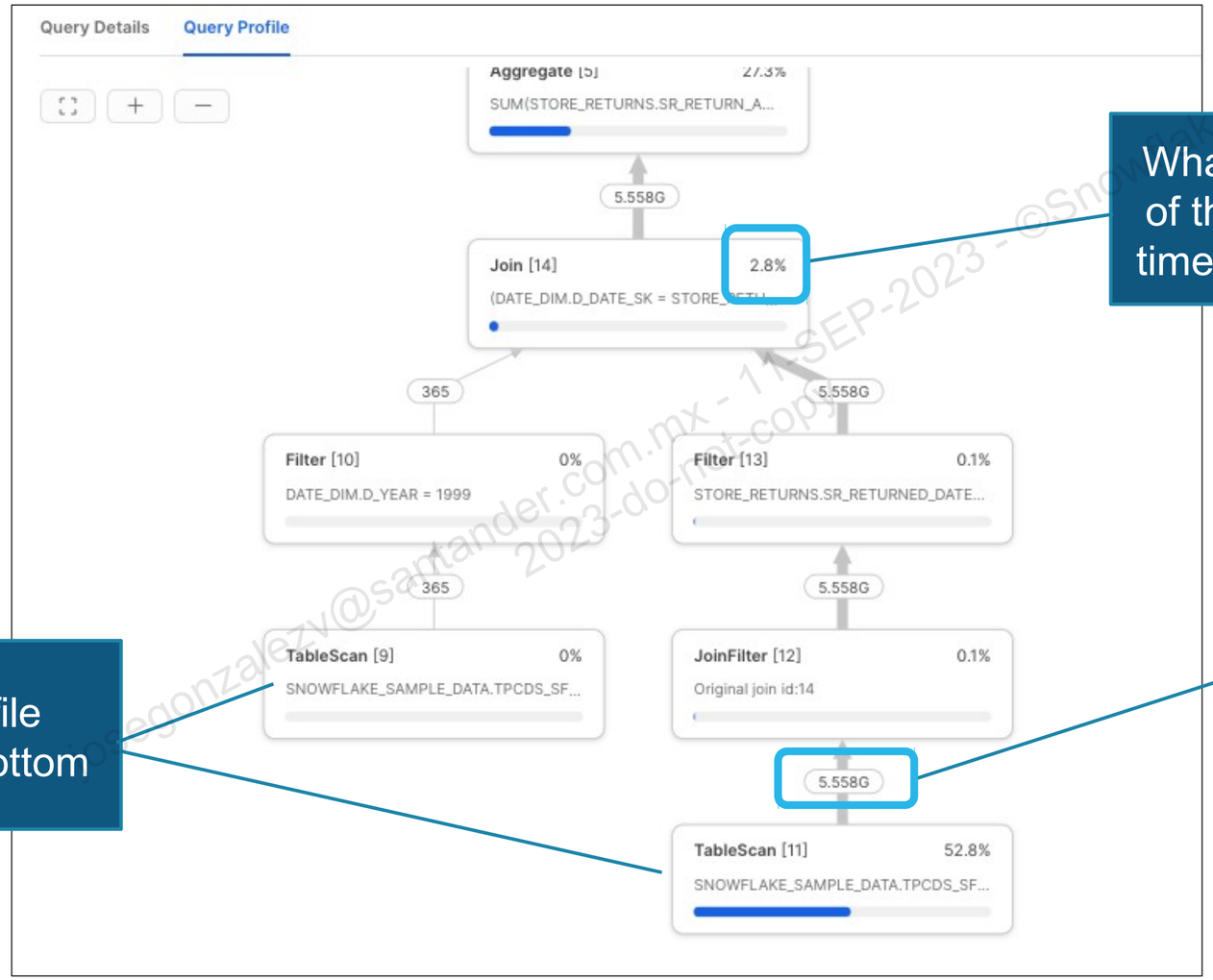


Select a node to see more information about that step

jsegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

READING A QUERY PROFILE



What percentage of the total query time the step took

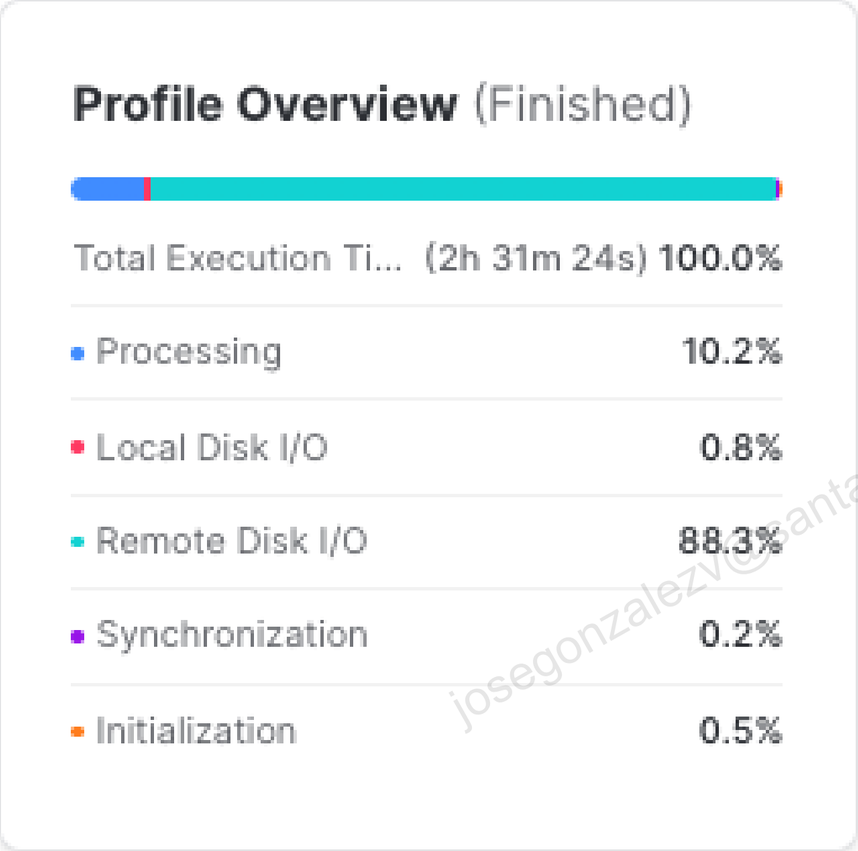
The number of rows passed from one step to the next

Read the profile starting at the bottom

jorgegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

PROFILE OVERVIEW



- **Processing:** data processing by CPU
- **Local Disk I/O:** processing blocked by local disk access
- **Remote Disk I/O:** processing blocked by remote disk access
- **Network Communication:** processing waiting for network data transfer
- **Synchronization:** managing activities between processes
- **Initialization:** setting up query processing

josegonzalezv@santander.com.mx - 11-SEP-2023 - @Snowflake
2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

STATISTICS

Statistics

Scan progress	100.00%
Bytes scanned	1.13TB
Percentage scanned from cache	0.00%
Bytes written to result	20.29GB
Partitions scanned	721507
Partitions total	721507
Bytes spilled to local storage	373.89GB
Bytes spilled to remote storage	105.81GB

- Percentage scanned from cache
  - Percentage of the required data that was already on the warehouse cache
- Partitions scanned vs Partitions total
  - Indication of how much data the optimizer was able to prune
- Bytes spilled to local storage
  - Insufficient memory
- Bytes spilled to remote storage
  - Insufficient memory and insufficient SSD

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

SQL PERFORMANCE TIPS

josegonzalezv@santander.com.mx - SEP-2023 - ©Snowflake
2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

TYPICAL DATABASE ORDER OF EXECUTION

- Row filtering operations are performed first

Rows

- FROM
- JOIN
- WHERE

→

Groups

- GROUP BY
- HAVING

→

Result

- SELECT
- DISTINCT
- ORDER BY
- LIMIT

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023 - do not copy

© 2023 Snowflake Computing Inc. All Rights Reserved

TOP PERFORMANCE TIP

- Row filtering operations are performed before GROUP BY operations
- Check row operations first
  - First, check `FROM` and `WHERE` clauses
  - Then, check `GROUP BY` and `HAVING` clauses
- Use appropriate filters, as early as possible
  - The Optimizer uses the filters to prune out unnecessary data
  - Best performance if you filter on well-ordered columns
- Avoid `SELECT *`

josegonzalez@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

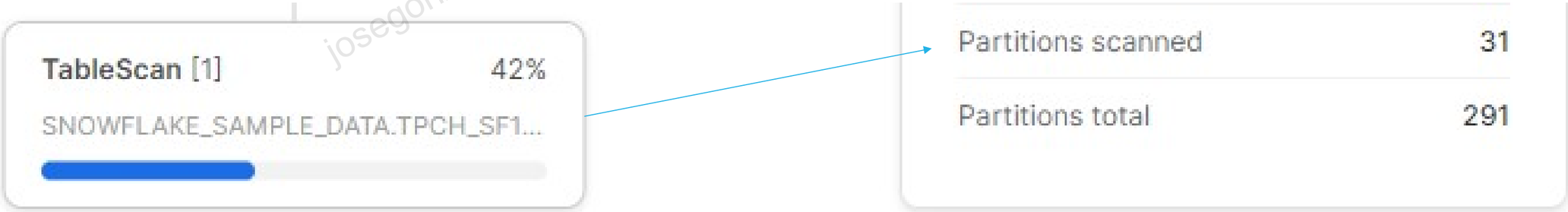
© 2023 Snowflake Computing Inc. All Rights Reserved

FILTER ON WELL-ORDERED COLUMNS

- Effective pruning: query's filter column is well-ordered
- In many cases, date columns are well-ordered

```
SELECT
    *
FROM
    order
WHERE
    o_orderdate >= '1993-10-01';
```

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy



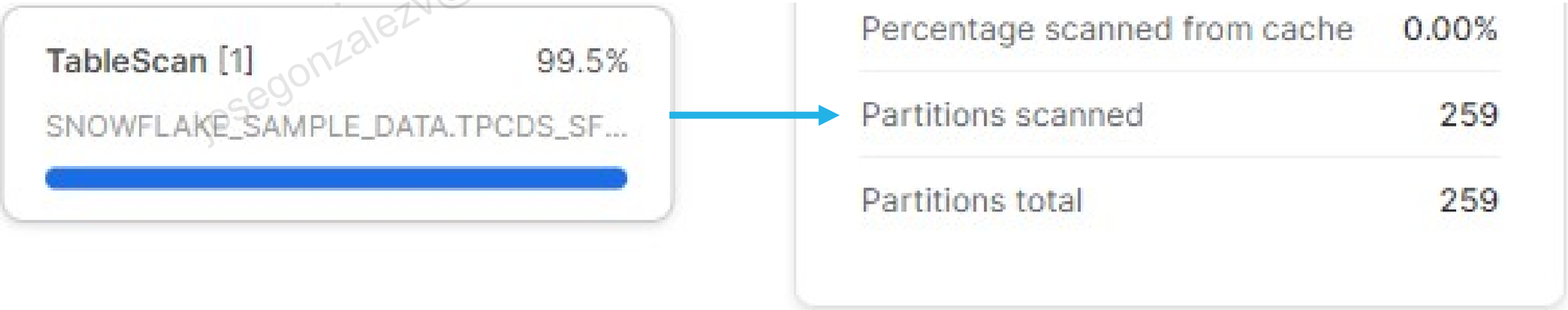
© 2023 Snowflake Computing Inc. All Rights Reserved

AVOID NON-PERFORMING PREDICATES

- Some WHERE clauses provide no opportunity to optimize pruning

```
SELECT
   c_customer_id, c_last_name
FROM
   tpcds.customer
WHERE
   UPPER (c_last_name) LIKE '%KROLL%';
```

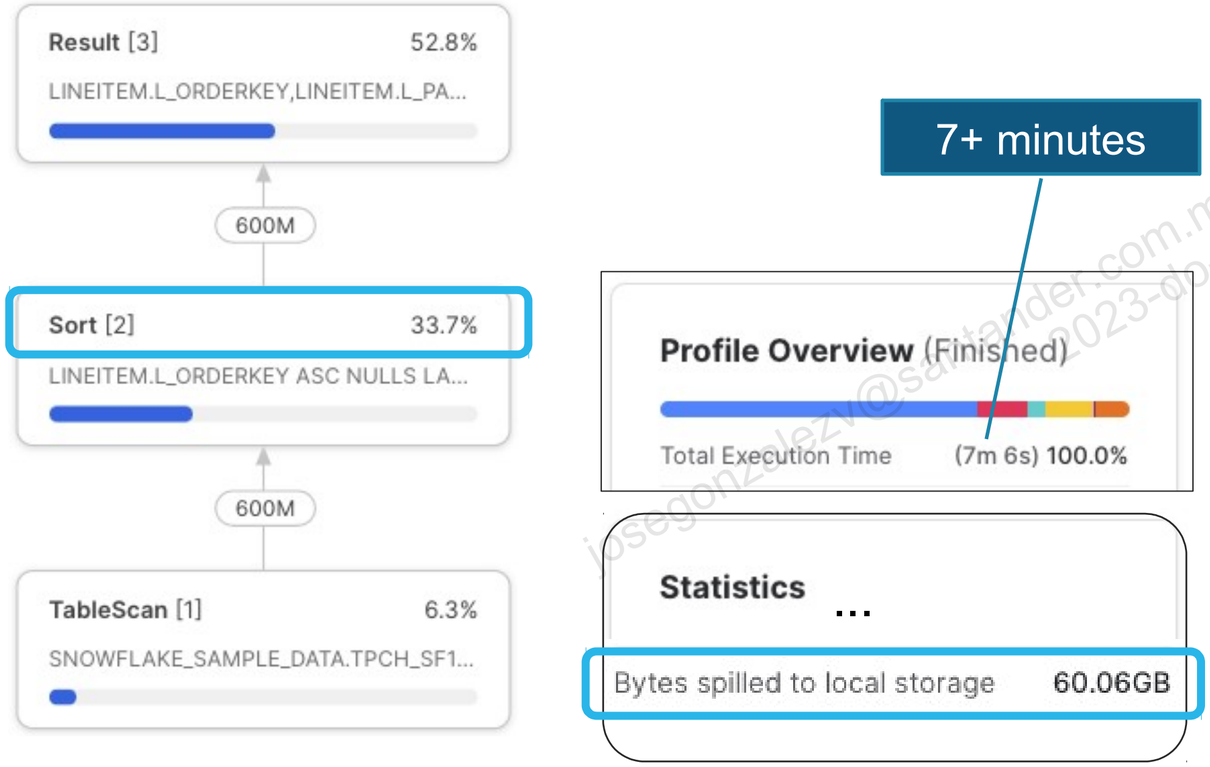
jsegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy



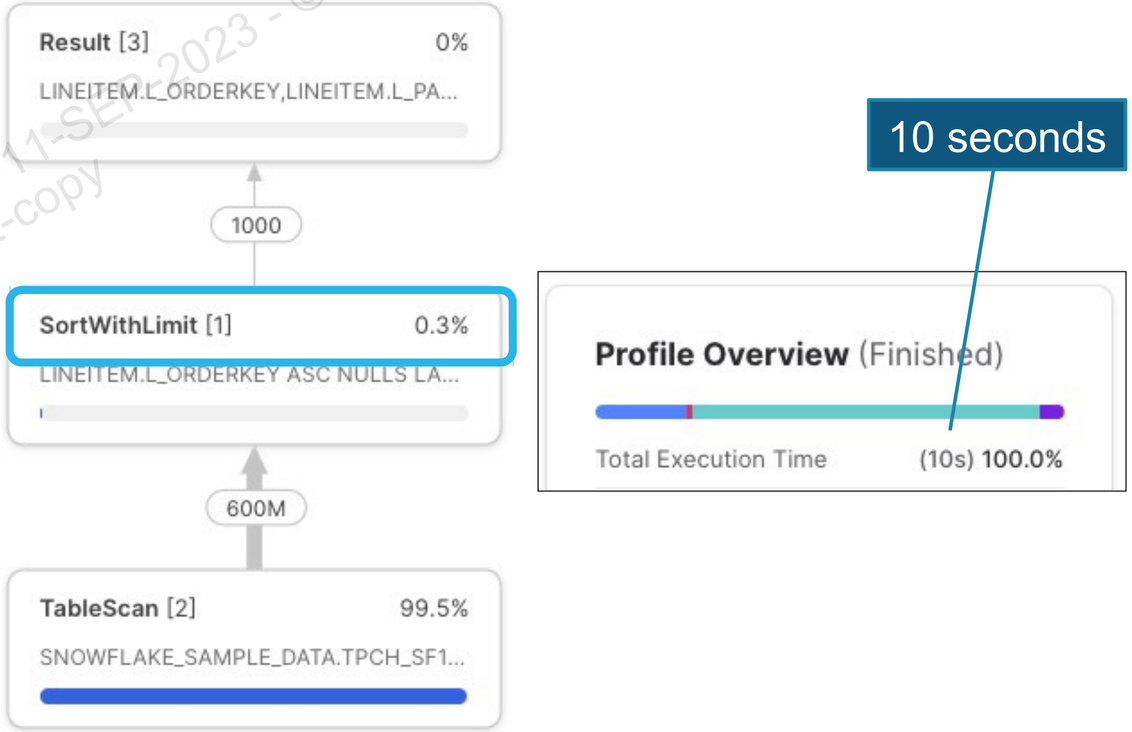
© 2023 Snowflake Computing Inc. All Rights Reserved

ADD LIMIT TO LARGE ORDER BY

```
SELECT * FROM LINEITEM
ORDER BY L_QUANTITY;
```




```
SELECT * FROM LINEITEM
ORDER BY L_QUANTITY LIMIT 1000;
```



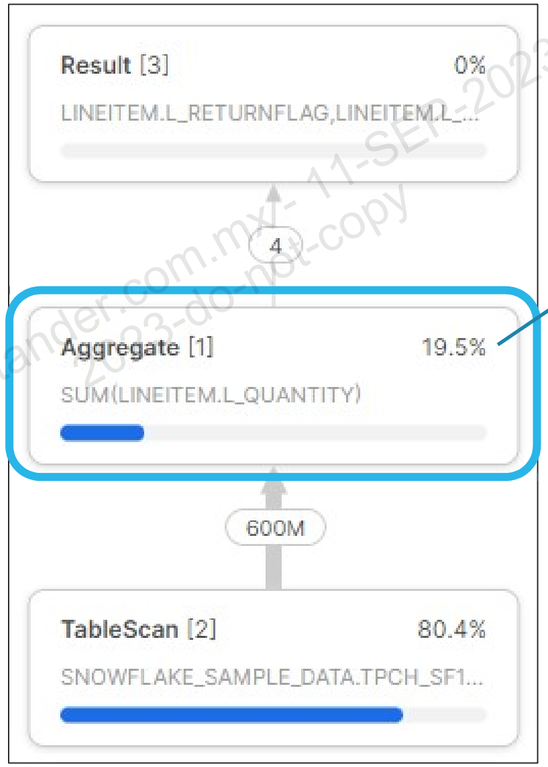
jsegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

GROUP BY WITH FEW DISTINCT VALUES

- `GROUP BY` can perform better on low-cardinality columns than on high-cardinality ones

```
SELECT
    l_returnflag,
    l_linestatus,
    SUM(l_quantity)
FROM
    lineitem
GROUP BY
    l_returnflag,
    l_linestatus;
```



Limit query times spent on aggregations

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

JOIN ON UNIQUE KEYS

- Understand relationships between your tables before joining
- Ensure keys are distinct
  - Joining on non-unique keys can explode your data output
  - Each row in table1 matches multiple rows in table2
- Avoid many-to-many join
- Avoid unintentional cross join

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

USE TEMPORARY TABLES

- Great for long calculations that need to be referenced multiple times in same query
- Great for materializing intermediate result
- Backed by micro-partitions and may assist in pruning
- Exist only during current session
- Disappear when you disconnect

josegonzalez@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

USE FUNCTIONS IN FILTER CAUTIOUSLY

- Sometimes the compiler can eliminate micro-partitions when a function is used to filter
  - The query optimizer compares the timestamp value to partition MIN/MAX date values; you get pruning.

```
SELECT count(*)
FROM orders
WHERE to_timestamp(o_orderdate) = '1995-02-04T00:00:00Z';
```

- Sometimes, it's not possible
  - Data type produced by function is **different** from the data type of the column; no pruning.

```
SELECT count(*) FROM orders
WHERE TO_CHAR(o_orderdate, 'yyyy-mm-dd') = '2022-02-04';
```

jsegonzalez@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

UDF IN FILTER

REDUCED PERFORMANCE

Statistics

IO

Scan progress	100.00 %
Bytes scanned	1.24 GB
Percentage scanned from cache	0.00 %

Network

Bytes sent over the network	0.04 MB

Pruning

Partitions scanned	2,315
Partitions total	2,315

Attributes

Full table name

```
SAMPLE_DATA.TPCH_SF1000.PARTSUPP
```

Columns (1 / 6)

PS_SUPPKEY

```
SELECT
    count(1)
FROM
    partsupp
WHERE
    NUM_TEST(ps_suppkey) = 65840;
```

- No Partition Pruning

josegonzalez@santander.com.mx - 11-SEP-2023 @Snowflake
2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

ADDITIONAL INFORMATION

- Snowflake Advanced discusses specific performance tuning capabilities, such as Query Acceleration, Search Optimization, and Clustering
- Snowflake's *Performance Workshop* provides in-depth strategies for optimizing query performance in Snowflake

josegonzalezv@santander.com.mx - 19-SEP-2023 - @Snowflake
2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

LAB EXERCISE: 9

Caching and Query Performance

45 minutes

Learning Objectives:

- How to access and navigate the Query Profile.
- Understand the differences between metadata cache, query result cache and data cache.
- How to determine when and why the query result cache is being used or not.
- How to determine if partition pruning is efficient, and if not, how to improve it.
- How to determine if spillage is taking place.
- How to use EXPLAIN to determine how to improve your queries.

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake

© 2023 Snowflake Computing Inc. All Rights Reserved

KNOWLEDGE CHECK

The root cause of "Bytes spilled to local storage" is what?

A. Insufficient memory

B. Insufficient disk space

C. Insufficient CPU

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

KNOWLEDGE CHECK

The root cause of "Bytes spilled to local storage" is what?

A. Insufficient memory

B. Insufficient disk space

C. Insufficient CPU

A. Insufficient memory

If a node runs out of memory, it spills first to local storage (the SSD on the node). If the SSD fills up, it will spill to remote storage (the storage layer).

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

KNOWLEDGE CHECK

TRUE or FALSE?

You should use the EXPLAIN command to determine how a query will run, so you can give hints to the compiler for a more efficient plan.

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

KNOWLEDGE CHECK

TRUE or FALSE?

You should use the EXPLAIN command to determine how a query will run, so you can give hints to the compiler for a more efficient plan.

FALSE

You cannot give hints to Snowflake's compiler. The primary use case for `EXPLAIN` is to examine micro-partition pruning, `JOIN` ordering, and `JOIN` types.

josegonzalezv@santander.com.mx - 11-SEP-2023 ©Snowflake
2023-do-not-copy

KNOWLEDGE CHECK

Which of these are potential causes of poor micro-partition pruning?

A. Not using a `WHERE` clause.

B. Filtering on any column other than the date.

C. Using a `WHERE` clause with a wildcard on a text column.

D. Using `SELECT` in the `FROM` clause.

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

KNOWLEDGE CHECK

Which of these are potential causes of poor micro-partition pruning?

A. Not using a `WHERE` clause.

B. Filtering on any column other than the date.

C. Using a `WHERE` clause with a wildcard on a text column.

D. Using `SELECT` in the `FROM` clause.

A. Not using a `WHERE` clause.

C. Using a `WHERE` clause with a wildcard on a text column.

josegonzalezv@santander.com.mx - 11-SEP-2023 @Snowflake
2023-do-not-copy

DATA LOADING AND UNLOADING

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake
2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

MODULE AGENDA

- Data Loading Objects
- Data Loading Process
- Transformations and Copy Options
- Data Loading Recommendations
- Continuous Data Loading
- Unloading Data

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

DATA LOADING OBJECTS

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake
2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

HIGH LEVEL DATA LOADING PROCESS

Output Data from System of Record to Files

Stage Files to Cloud Storage

Load Data from Cloud Storage Into Snowflake

jose.gonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

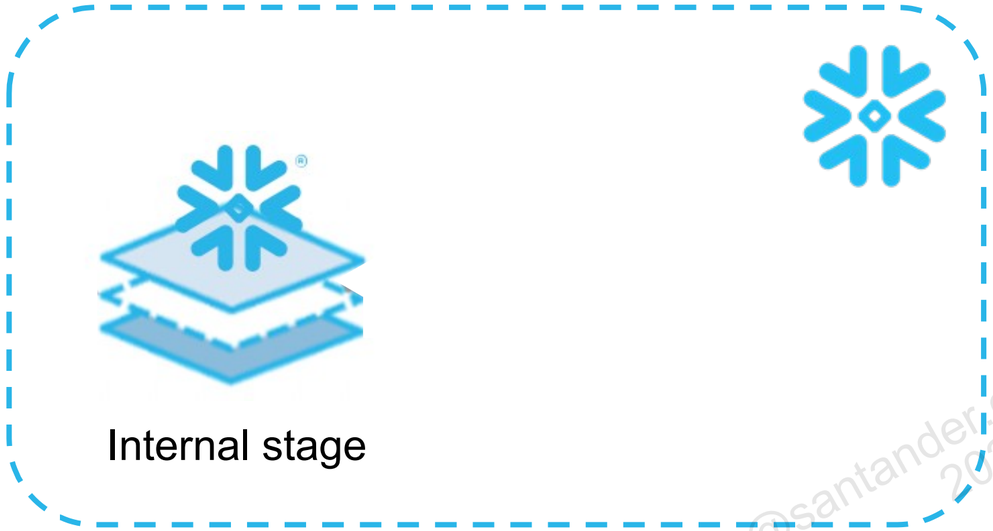
WHAT IS A STAGE?

- Cloud file repository that holds files to be loaded
- Can be internal or external
  - Internal: stored internally in Snowflake
  - External: stored in an external location
- Data can be queried directly from a stage without loading it first

josegonzalezv@santander.com.mx - 11-SEP-2025 - ©Snowflake
2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

NAMED STAGES



External stage

- Can be internal (managed by Snowflake) or external (managed by customer)
- Create internal stage:
  - `CREATE STAGE <name>;`
- Create external stage:
  - "Pointer" to external cloud location
  - `CREATE STAGE <name> <cloud-provider-information>;`
- Referenced as `@<stage name>`

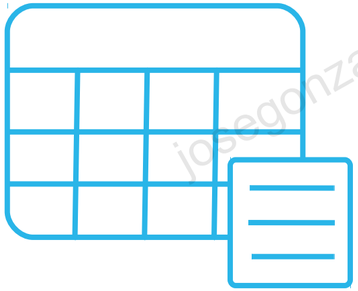
josegonzalezv@santander.com.mx - 11-SEP-2023 @Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

AUTOMATICALLY-CREATED STAGES

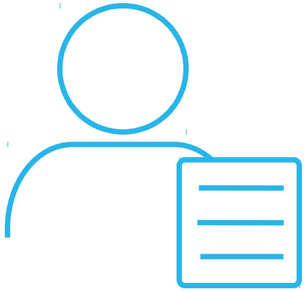
Table Stage

- Automatically created by Snowflake for every permanent table
- Reference as `@%<table name>`



User Stage

- Automatically created by Snowflake for every user on the system
- Reference as `@~`



josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

WHAT IS A FILE FORMAT?

```
Header line 1
Header line 2
Col1|Col2|Col3|Col4
123|ABC|987|FED
234|BCD|876|EDC
345|CDE|765|DCB
```

```
CREATE FILE FORMAT my_csv
  TYPE = CSV
  SKIP_HEADER = 3
  FIELD_DELIMITER = '|';
```

Base type

Modifiers

- Object that describes the structure of files that are read or written
- Includes a base type:
  - CSV (for all delimited data)
  - JSON, Avro, ORC, Parquet, XML (for semi-structured data)
- Optional modifiers to further refine the base type

jessegonzalezv@santander.com.mx - 11-SEP-2023 @Snowflake
2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

FILE FORMAT SYNTAX

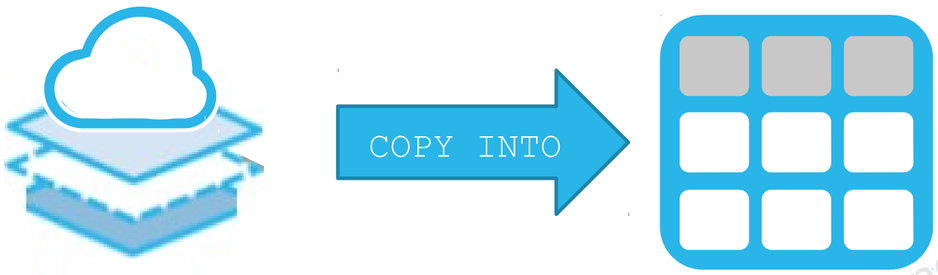
- CSV is the default if a base type is not included
- See documentation for modifiers for all base types
- Sample syntax for CSV base type:

```
CREATE FILE FORMAT new_csv
  TYPE = CSV
  FIELD_DELIMITER = '<character>' | NONE
  RECORD_DELIMITER = '<character>' | NONE
  SKIP_BLANK_LINES = TRUE
```

jesgonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

COPY INTO



- Copies data from staged files to a table
- All files loaded in a single operation must use the same stage and file format
- Can auto-detect compression mode, date and time formats
- Requires an active virtual warehouse

josegonzalezv@santander.com.mx - 11-SEP-2023 ©Snowflake
2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

BASIC LOAD SYNTAX

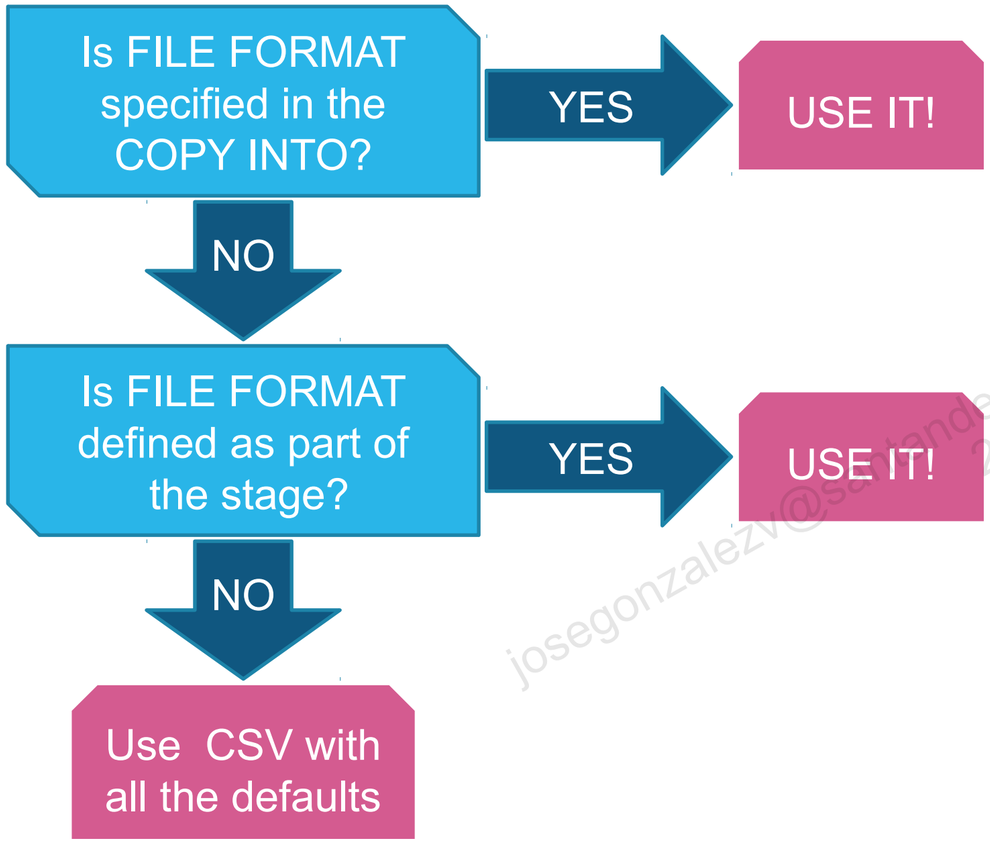
- Simplest form: `COPY INTO <table> FROM <stage>;`
- Options can specify files to load, file format, error handling, etc:

```
COPY INTO
   <table>
FROM
   <stage>
FILE_FORMAT = 'my_csv'
FILES = ('file1', 'file2'...) | PATTERN = '<regular expr>'
ON_ERROR = <error-handling method>
```

jgsegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake
2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

SPECIFYING THE FILE FORMAT



- Every `COPY INTO` command must know the file format
- Three ways the file format can be specified:
  - In the COPY INTO command
  - In the definition of the stage (either a named stage, or a table stage)
  - Default is CSV with all defaults

josegonzalezv@santander.com.mx - 11-SEP-2023 - @Snowflake
2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

QUERYING STAGED DATA FILES

```sql
SELECT n.$1 AS n_key,
   CASE
      WHEN n.$3 = 0 THEN 'AFRICA'
      WHEN n.$3 = 1 THEN 'AMERICA'
      WHEN n.$3 = 2 THEN 'ASIA'
      WHEN n.$3 = 3 THEN 'EUROPE'
      ELSE 'MIDDLE EAST'
   END AS region,
   SUBSTR(region, 1, 2) AS region_code,
   n.$2 AS nation,
   n.$4 AS comments
FROM @training_db.traininglab.ed_stage/load/lab_files/nation.tbl
   (FILE_FORMAT => 'MYPIPEFORMAT') n
WHERE region = 'ASIA';
```

Use CASE statements and functions

File format

Add filters

Objects | Editor | Results | Chart

N_KEY	REGION	REGION_CODE	NATION	COMMENTS
8	ASIA	AS	INDIA	ss excuses cajole slyly across the packages. de
9	ASIA	AS	INDONESIA	slyly express asymptotes. regular deposits hag
12	ASIA	AS	JAPAN	ously. final, express gifts cajole a
18	ASIA	AS	CHINA	c dependencies, furiously express notornis slee

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

KNOWLEDGE CHECK

- You are loading a table named `my_table`
- The files you want to load are on the table stage
- You've created a file format called `my_csv` that you want to use

What `COPY INTO` command will you use?

jose.gonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

KNOWLEDGE CHECK

- You are loading a table named `my_table`
- The files you want to load are on the table stage
- You've created a file format called `my_csv` that you want to use

What `COPY INTO` command will you use?

```
COPY INTO
  my_table
FROM
  @%my_table
FILE_FORMAT = 'my_csv';
```

jose.gonzalezv@santander.com.mx - 11-SEP-2023 ©Snowflake
2023-do-not-copy

KNOWLEDGE CHECK

What does a file format do?

A. Format the data as it is being written to the table.

B. Describe how to read the input file.

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

KNOWLEDGE CHECK

What does a file format do?

A. Format the data as it is being written to the table.

B. Describe how to read the input file.

B. Describe how to read the input file.

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

KNOWLEDGE CHECK

TRUE OR FALSE?

Metadata for micro-partitions is created only when you load data from an internal stage.

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

KNOWLEDGE CHECK

TRUE OR FALSE?

Metadata for micro-partitions is created only when you load data from an internal stage.

FALSE

Metadata for micro-partitions is always created, even if you load data from an external stage, or use other methods (such as INSERT or CTAS) to add data to a table.

josegonzalezv@santander.com.mx - 11-SEP-2023 ©Snowflake
2023-do-not-copy

KNOWLEDGE CHECK

Which of the following statements about stages are TRUE?

A. A stage is a cloud file repository.

B. A stage can be internal or external.

C. Stages disappear as soon as all their files are loaded.

D. You can manually create named stages, user stages, or table stages.

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

KNOWLEDGE CHECK

Which of the following statements about stages are TRUE?

A. A stage is a cloud file repository.

B. A stage can be internal or external.

C. Stages disappear as soon as all their files are loaded.

D. You can manually create named stages, user stages, or table stages.

A. A stage is a cloud file repository.

B. A stage can be internal or external.

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

KNOWLEDGE CHECK

Which of the following statements about file formats are TRUE?

A. File formats store data of a particular type.

B. A file format stores information needed to parse input files for loading.

C. You must specify a file format as part of every COPY INTO command.

D. The default file format is CSV.

jsegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

KNOWLEDGE CHECK

Which of the following statements about file formats are TRUE?

A. File formats store data of a particular type.

B. A file format stores information needed to parse input files for loading.

C. You must specify a file format as part of every COPY INTO command.

D. The default file format is CSV.

B. A file format stores information needed to parse input files for loading.

D. The default file format is CSV.

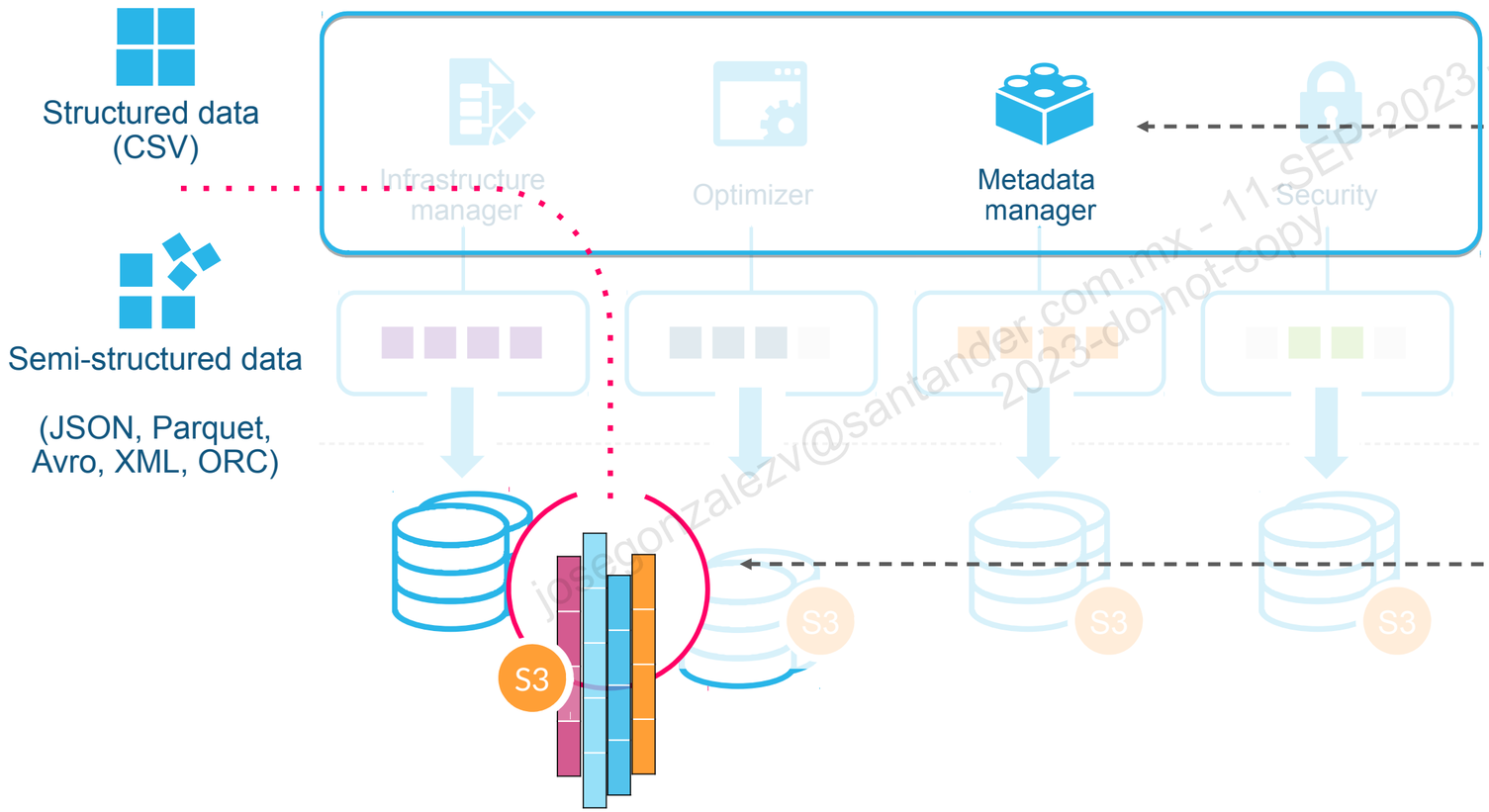
jsegonzalezv@santander.com.mx - 11-SEP-2023 ©Snowflake 2023-do-not-copy

DATA LOADING PROCESS

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake
2023-do not copy

© 2023 Snowflake Computing Inc. All Rights Reserved

WHAT HAPPENS DURING LOAD



Metadata created

Statistics for databases, tables, columns, and files stored in the cloud services layer

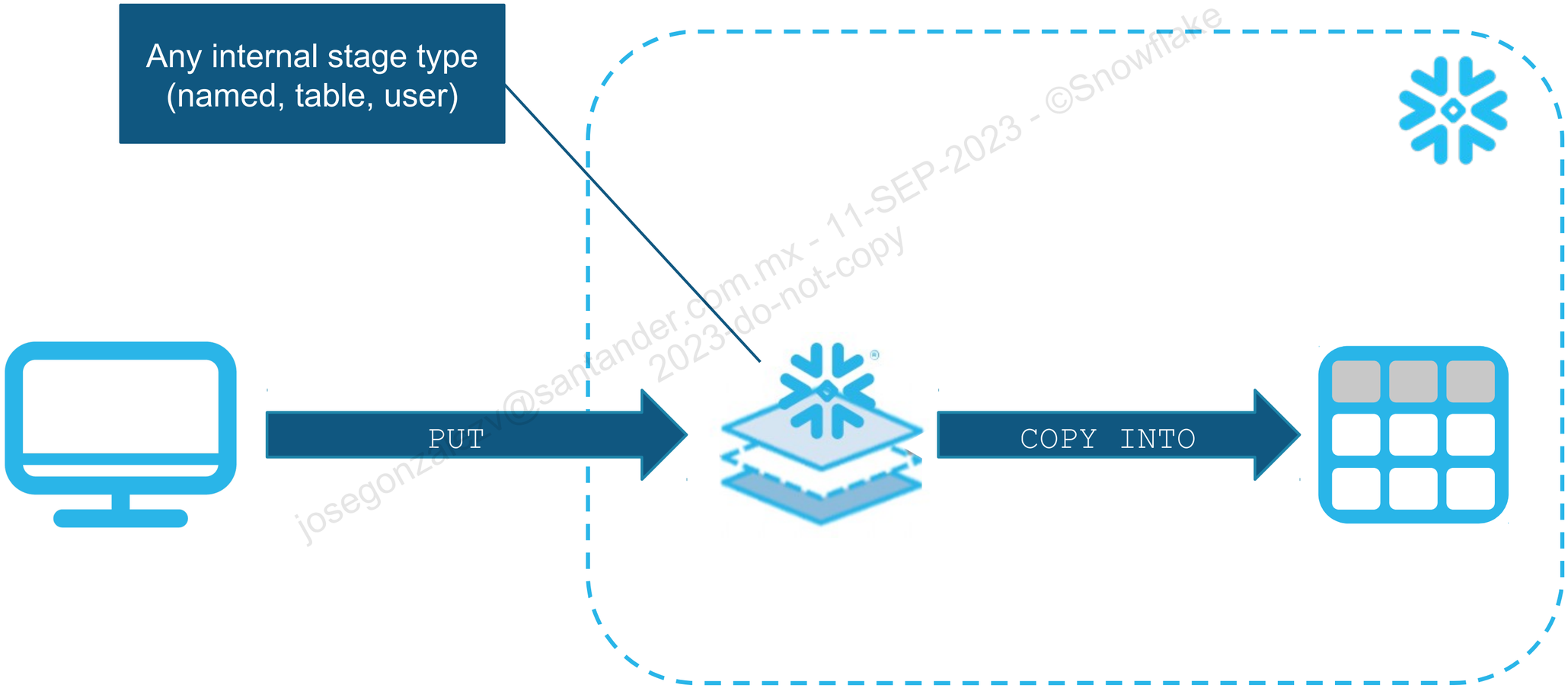
Database data stored

- Data stored in Snowflake-managed cloud storage
- Optimized, proprietary file format
- Automatically compressed and encrypted

jose.gonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

LOAD FROM LOCAL FILE SYSTEM



josegonzalez@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023 do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

EXAMPLE FROM LOCAL STORAGE

1. Create internal stage (or use table/user stage):

```
CREATE STAGE my_internal_stage;
```

2. Use a Snowflake connector like SnowSQL to `PUT` files onto stage:

```
PUT file://C:\User1\Desktop\File* @my_internal_stage;
```

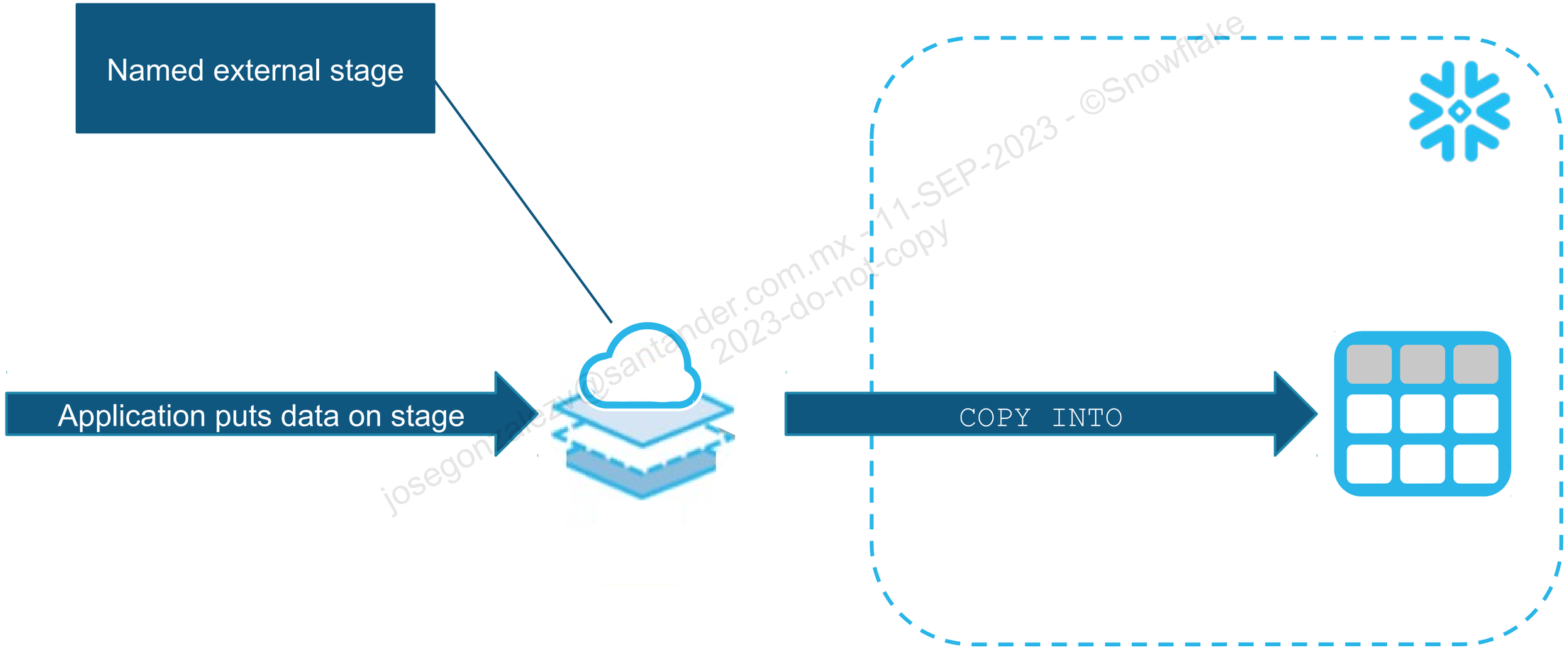
3. Load data into table:

```
COPY INTO
    my_table
FROM
    @my_internal_stage;
```

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023 - do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

LOAD FROM CLOUD STORAGE



josegonzalez@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

EXAMPLE FROM CLOUD STORAGE

1. Create external stage (example shows AWS):

```
CREATE STAGE my_s3_stage
  URL = 's3://mybucket/encrypted_files/'
  CREDENTIALS = (aws_key_id='******' aws_secret_key='******')
  ENCRYPTION = (TYPE = 'AWS_SSE_S3')
  FILE_FORMAT  =  my_csv_format;
```

2. Load data:

```
COPY INTO
   my_table
FROM
   @my_s3_stage;
```

josegonzalezv@santander.com.mx - 14 SEP 2023 - ©Snowflake
2023 do not copy

© 2023 Snowflake Computing Inc. All Rights Reserved

KNOWLEDGE CHECK

Does this command use an internal stage, or an external stage?

```
COPY INTO
  my_table
FROM
  @my_stage
```

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

KNOWLEDGE CHECK

Does this command use an internal stage, or an external stage?

```
COPY INTO
  my_table
FROM
  @my_stage
```

You can't tell. It's a named stage, which could be either external or internal

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

KNOWLEDGE CHECK

TRUE OR FALSE?

This command will load data from a named stage:

```
COPY INTO my_table
FROM @%my_table;
```

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

KNOWLEDGE CHECK

TRUE OR FALSE?

This command will load data from a named stage:

```
COPY INTO my_table
FROM @%my_table;
```

FALSE

The command is loading data from a table stage. You can tell from the prefix:

`@<stage name>` = named stage
`@%<table name>` = table stage
`@~` = user stage

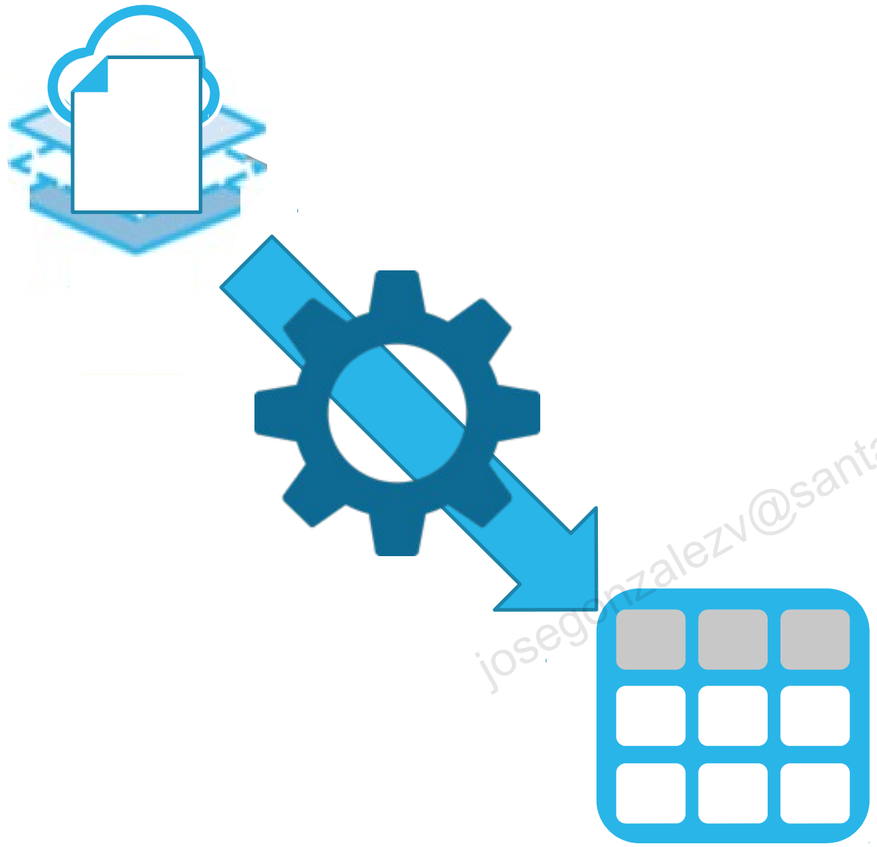
josegonzalezv@santander.com.mx - 11-SEP-2023 ©Snowflake
2023-do-not-copy

TRANSFORMATIONS AND COPY OPTIONS

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake
2026 - do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

TRANSFORMING DATA DURING LOAD



- The COPY command supports column reordering, column omission, and CAST using a SELECT statement
  - NOT SUPPORTED: Joins, filters, aggregations
- Can include SEQUENCE columns, current_timestamp(), or other column functions during data load

josegonzalezv@santander.com.mx - 11-SEP-2023 - @Snowflake
2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

TRANSFORMING DATA DURING LOAD EXAMPLES

```
COPY INTO
   home_sales (city, zip, sale_date, price)
FROM
   (SELECT SUBSTR(t.$2,4), t.$1, t.$5, t.$4 FROM @my_stage t)
FILE_FORMAT = 'mycsvformat';
```

```
COPY INTO
   casttb(col1, col2, col3)
FROM
   (SELECT TO_BINARY(t.$1, 'utf-8'),
      TO_DECIMAL (t.$2, '99.9', 9, 5),
      TO_TIMESTAMP_NTZ(t.$3)
    FROM @~/datafile.csv.gz t)
FILE_FORMAT = (TYPE = CSV);
```

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

COPY COMMAND OPTIONAL PARAMETERS

```
FILES = ( '<file name>' [ , '<file name>'] [, ...])

PATTERN = '<regex pattern>'

FILE_FORMAT = '<format name>' | (TYPE = <type> [<type options>])


PURGE = TRUE | FALSE

ENFORCE_LENGTH = TRUE | FALSE   OR   TRUNCATECOLUMNS = TRUE | FALSE
```

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake
2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

ERROR HANDLING METHODS

SELECTED MODE	ACTION TAKEN
ABORT_STATEMENT (default)	Stop and roll back if any error is encountered.
CONTINUE	Attempt to load every record, no matter how many errors are found.
SKIP_FILE	Skip a file when an error is found. Buffers the file contents until the entire file is read, so slower than CONTINUE.
SKIP_FILE_<num>	Skip a file when the number of error rows found in the file is equal to or exceeds the specified number. Buffers file contents while loading.
SKIP_FILE_<num>%	Skip a file when the percentage of error rows found in the file exceeds the specified percentage. Buffers file contents while loading.

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do not copy

© 2023 Snowflake Computing Inc. All Rights Reserved

VALIDATE THE LOAD

- The `VALIDATE` table function lists all errors found during a `COPY INTO`

ERROR	FILE	LINE	CHAR	REJECTED_RECORD
Numeric value 'Zero' is not recogni:	Plants_2_Errs.csv.gz	45	81	Leucanthemum Double Angel Daisy,Daisy,Late
User character length limit (3) exce	Plants_2_Errs.csv.gz	26	108	Dropmore Purple Wand Loosestrife,Loosestrife,
User character length limit (3) exce	Plants_Line1_Err.csv.gz	2	105	Aguilegia Double Dark Blue and White,Columbir
Number of columns in file (5) does	Plants_Good.xlsx.gz	3	1	⍰D⍰⍰⍰Z ⍰ H ⍰ ⍰ [Ø9s ⍰⍰⍰1⍰⍰@⍰

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

- Unload rejected records into a text file or table, fix, and re-load

© 2023 Snowflake Computing Inc. All Rights Reserved

MONITOR LOAD COMMANDS

- Monitor `COPY` commands from Query History
- Use the `INFORMATION_SCHEMA.LOAD_HISTORY` view to retrieve `COPY` history

```
SELECT
    table_name,
    file_name,
    last_load_time,
    status
FROM
    load_history
WHERE
    schema_name=CURRENT_SCHEMA();
```

TABLE_NAME	FILE_NAME	LAST_LOAD_TIME	STATUS
PLANT_LIST	stages/1bbdca8e-eea4-4fb4-b6cb-f	2023-02-02 10:49:11.034 -0800	PARTIALLY_LOADED
PLANT_LIST	stages/1bbdca8e-eea4-4fb4-b6cb-f	2023-02-02 10:49:11.034 -0800	PARTIALLY_LOADED
PLANT_LIST	stages/1bbdca8e-eea4-4fb4-b6cb-f	2023-02-02 10:49:11.034 -0800	LOADED
PLANT_LIST	stages/1bbdca8e-eea4-4fb4-b6cb-f	2023-02-02 10:49:11.034 -0800	LOADED
PLANT_LIST	stages/1bbdca8e-eea4-4fb4-b6cb-f	2023-02-02 10:49:11.034 -0800	LOADED

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

DATA LOADING RECOMMENDATIONS

josegonzalezv@santander.com.mx - 15-SEP-2023 - ©Snowflake
2024-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

FILE SIZE AND NUMBER

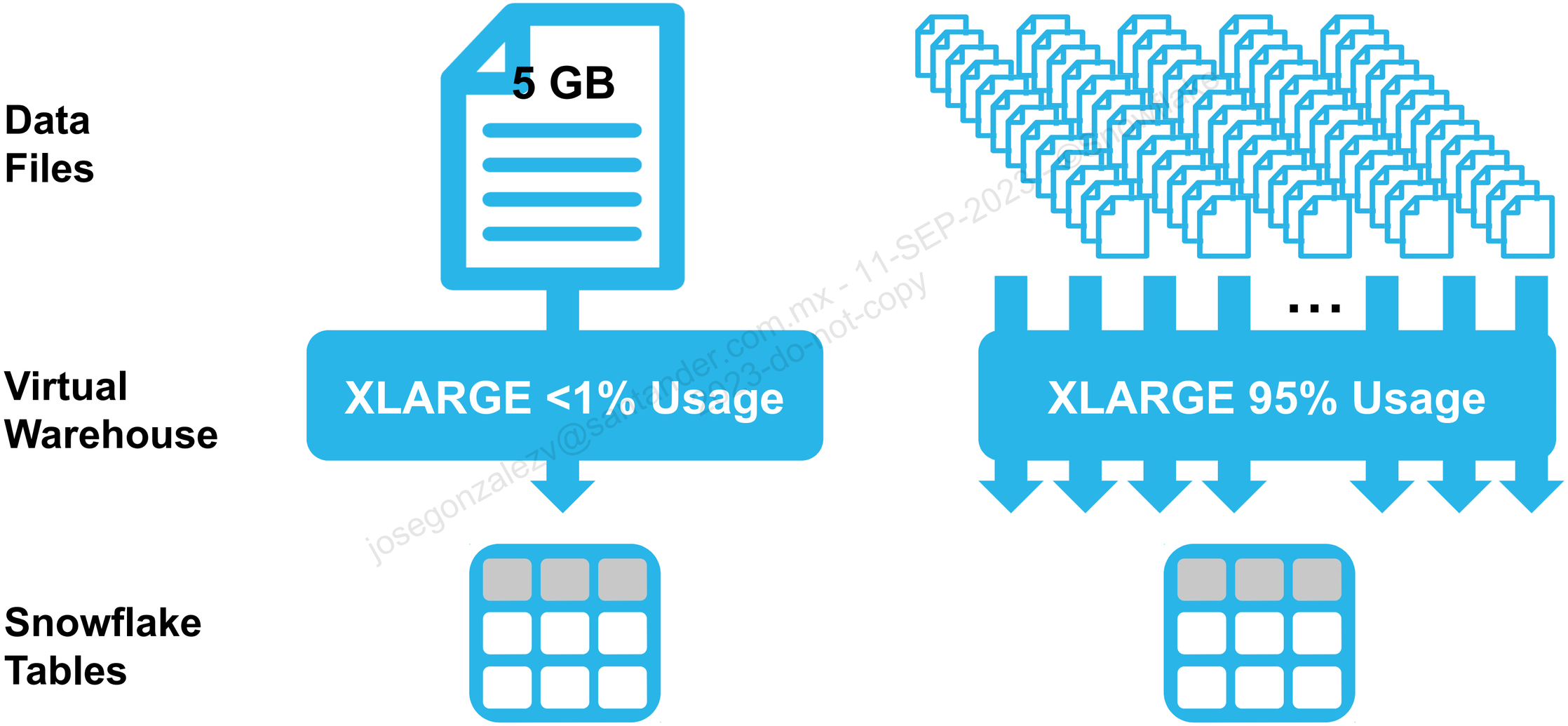
- File size and number of files are crucial to optimizing load performance
- Split large files before loading into Snowflake
- Recommended: 100MB to 250MB (compressed)

Warehouse Size	# files in parallel
XS	8
S	16
M	32
L	64
XL	128

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

SERIAL COPY VS PARALLEL COPY



Note: Single COPY Command

josegonzalezv@santander.com.mx - 11-SEP-2023 - 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

FILE ORGANIZATION

- Organize data in logical paths (e.g., subject area and create date)

```
/system/market/daily/2018/09/05/
```

- Use pattern matching to only load specified files, thus reducing scanning:

```
COPY INTO
   my_table
FROM
   @mystage/system/market/daily/2018/
PATTERN = '<regex_pattern>';
```

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

SAMPLE PATTERN MATCHING

- Load only uncompressed CSV files whose names include the string "sales":

```
COPY INTO
   mytable
FROM
   @mystage/system/market/daily/2018/
FILE_FORMAT = 'myformat'
PATTERN = '.*sales.*[.]csv';
```

josegonzalez@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

LAB EXERCISE: 10

Loading Structured Data

20 minutes

Learning Objectives:

- How to load data from a file in an external stage into a table using the COPY INTO command.
- How to define a GZIP file format.
- How to review the properties of a stage.
- How to load a GZipped file from an external stage into a table.
- How to validate data prior to loading
- How to handle data loading errors

Note: Towards the end of this lab you'll get a chance to write a COPY INTO statement on your own. A solution is presented on the last page of this lab in the workbook.

jose.gonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

LAB EXERCISE: 11

Challenge Lab: Loading Structured Data

40 minutes

Learning Objectives:

- How to create a table
- How to create a file format
- How to load data from a file in an external stage into a table using the COPY INTO command.

Note: A solution is presented on the last page of this lab in the workbook.

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

CONTINUOUS DATA LOADING

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake
2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

BATCH VS. CONTINUOUS LOADING

COPY command (BATCH)

- Migration from traditional data sources
- Transaction boundary control support
- Independently scale compute resources for different ingestion workloads

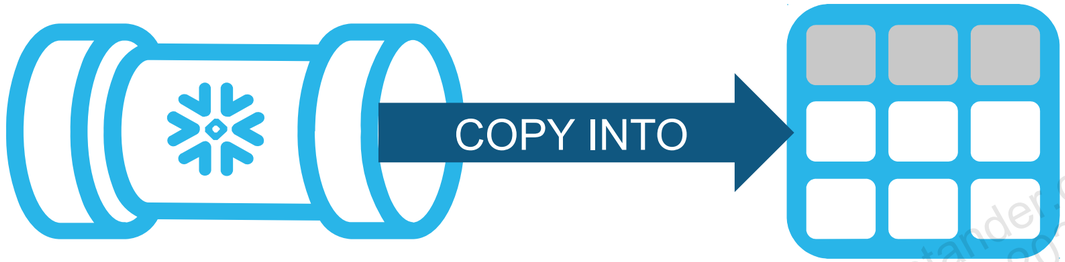
Snowpipe (CONTINUOUS)

- Ingestion from modern data sources
- Data is available near-real-time
- No scheduling (with auto-ingest)
- Serverless model needs no user-managed virtual warehouse

josegonzalezv@santander.com.mx - 11-SEP-2023 ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

WHAT IS A PIPE?



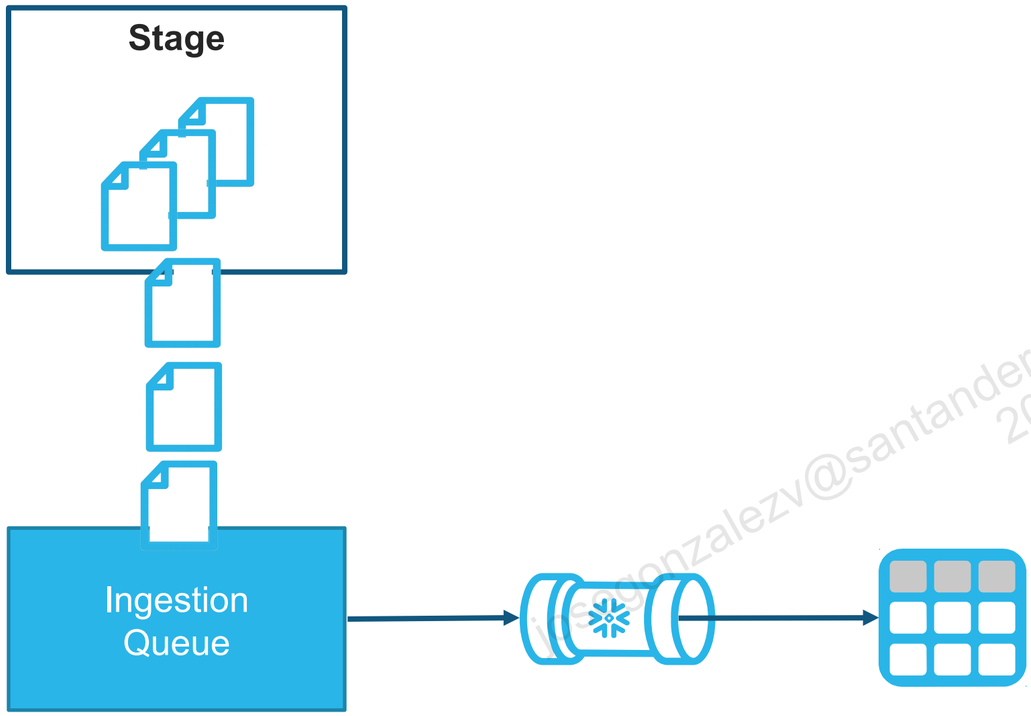

```
CREATE PIPE load_iot
AS
   COPY INTO
      my_table
   FROM
      @iot_stage
   FILE_FORMAT = (TYPE = JSON);
```

- Schema-level object
- Named definition of a COPY INTO
- Loads data into tables continuously from an ingestion queue
- Can be suspended/resumed, return status

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake
2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

WORKFLOW



1. Data files are collected onto a stage
2. When triggered, files are moved into an ingestion queue
3. The pipe processes the files using `COPY INTO`

jesgonzalezv@santander.com.mx - 11-SEP-2023 - @Snowflake
2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

WHAT TRIGGERS A PIPE?

REST API

- Application calls REST endpoint and passes a list of files to load
- Works with internal or external stages

AUTO_INGEST

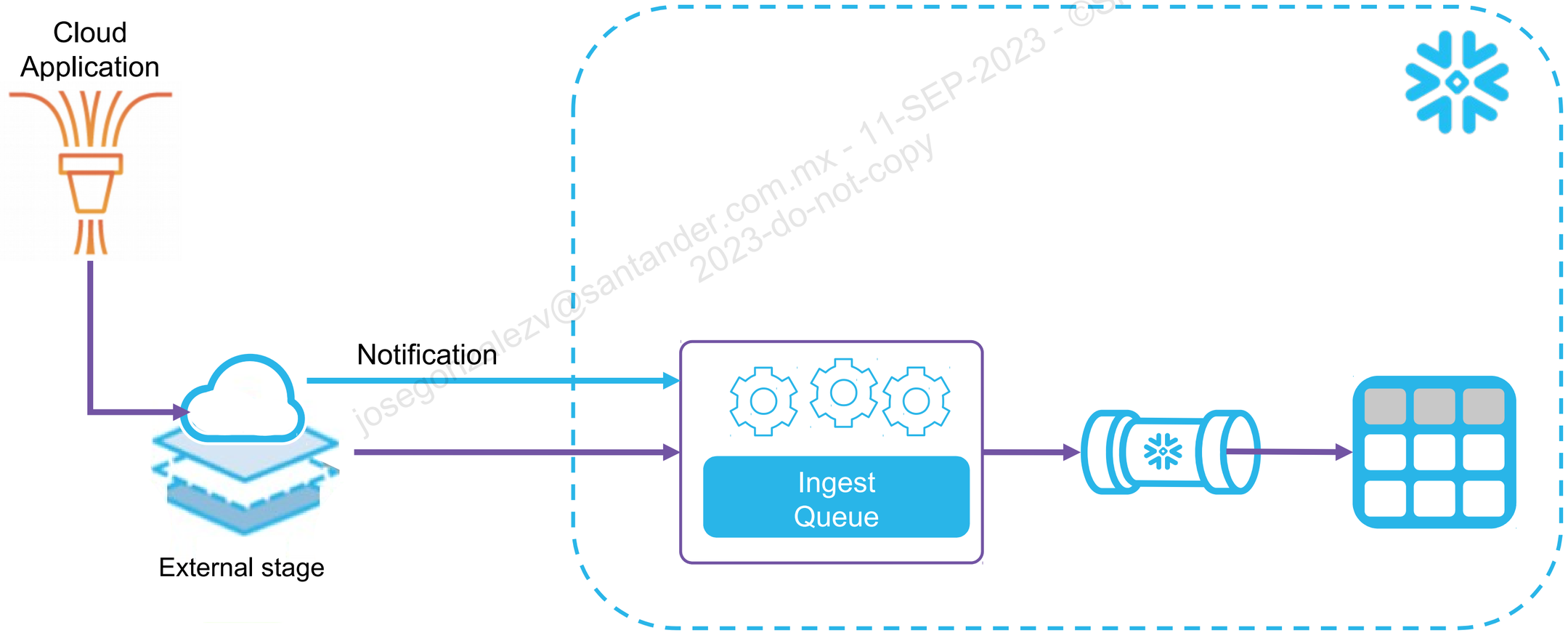
- Notifications from cloud provider alerts pipe to look for files on stage
- Only works with external stages

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

SNOWPIPE AUTO-INGEST

```
CREATE PIPE mypipe AS COPY INTO sales FROM @mystage AUTO_INGEST=TRUE;
```

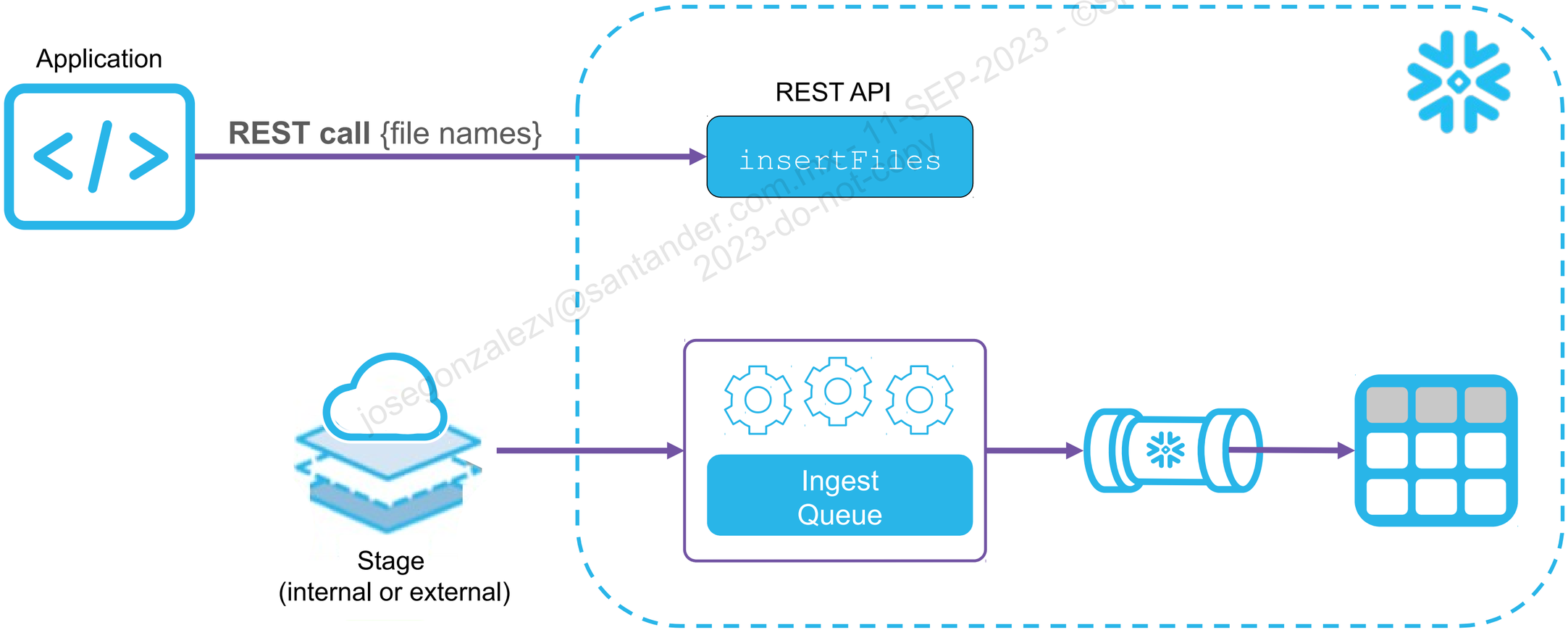


josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

SNOWPIPE REST API

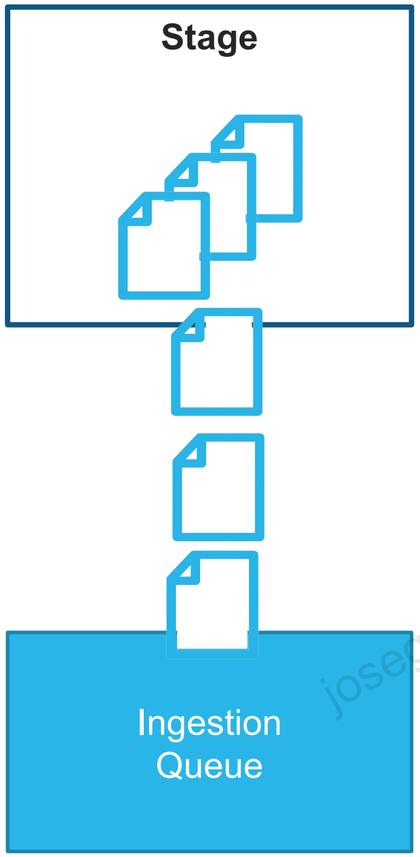
```
CREATE PIPE mypipe AS COPY INTO sales FROM @mystage;
```



joseconzalezv@santander.com.mx 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

FILE LOAD ORDER



- Staged files are copied into an ingest queue
- Files that appear in the stage later are appended to the queue
- Multiple processes pull from the queue
- Older files are generally loaded first, but files are not guaranteed to be loaded in the same order they are staged

josegonzalezv@santander.com.mx - 11-SEP-2023 - @Snowflake
2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

SNOWPIPE RECOMMENDATIONS

- File sizes 100-250MB compressed
- To decrease overhead to manage files, stage files no more than once per minute

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake
2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

SNOWPIPE BILLING

- Serverless model: does not require a virtual warehouse
- Snowflake provides and manages compute resources
  - Capacity grows and shrinks depending on load
- Charged per-second, per-core
- Utilization cost of 0.06 credits per 1000 files notified to Snowpipe
  - Charged even if load fails

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

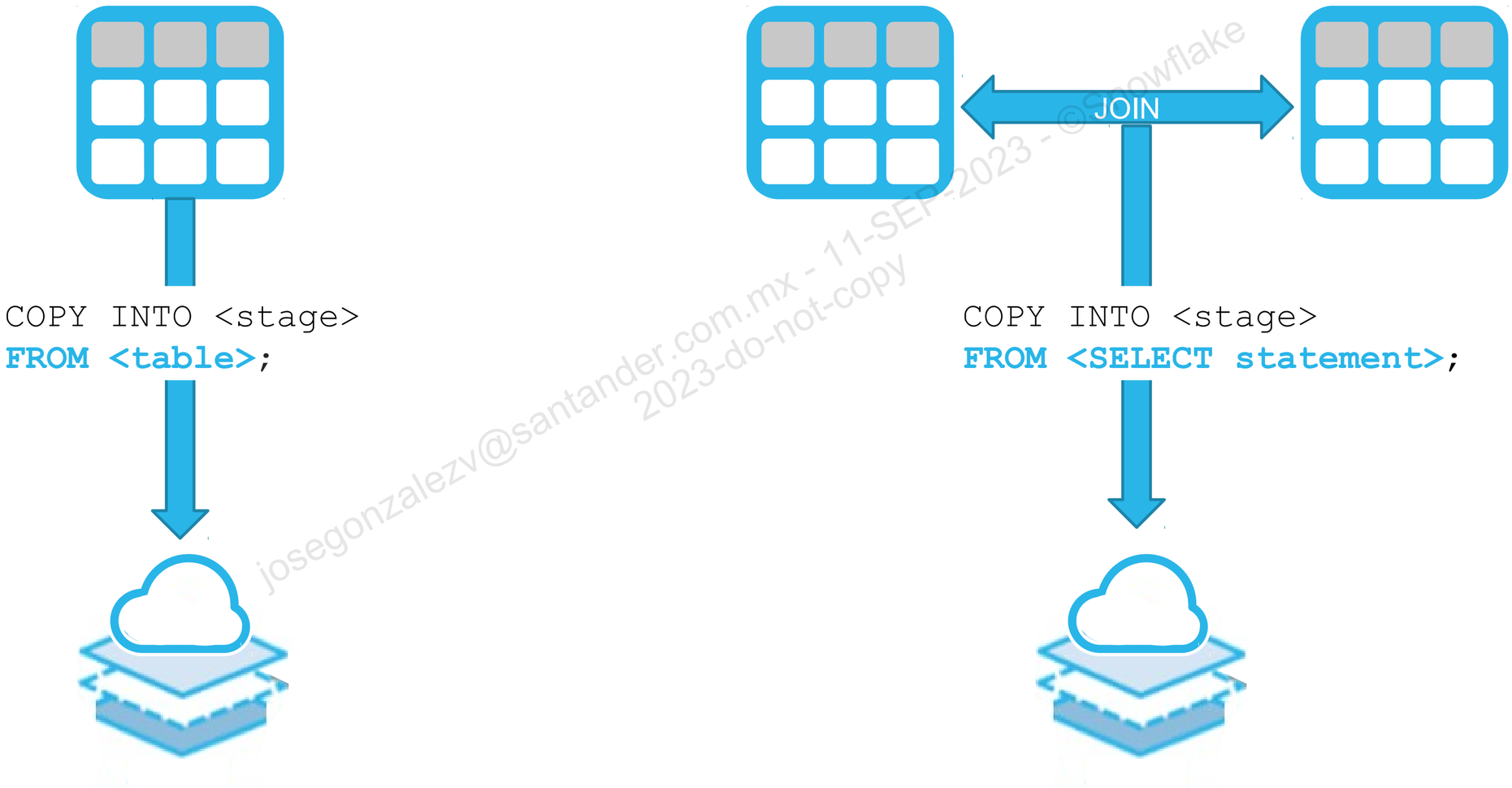
© 2023 Snowflake Computing Inc. All Rights Reserved

UNLOADING DATA

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake
2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

HIGH-LEVEL DATA UNLOADING PROCESS



josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake
2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

UNLOAD SYNTAX EXAMPLES

- Unload entire contents of table into named stage:

```
COPY INTO @my_stage
FROM my_table
FILE_FORMAT = 'my_format';
```

- Unload result of `SELECT` into table stage:

```
COPY INTO @%my_table
FROM (SELECT column1, column2
    FROM my_table m
      JOIN your_table y ON m.id = y.id
    ORDER BY column2);
```

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

OPTIONAL UNLOAD PARAMETERS

Parameter	Description
FILE_FORMAT	CSV (default), JSON, or Parquet supported
VALIDATION_MODE = RETURN_ROWS	Test the unload before performing it
HEADER	TRUE or FALSE (default) – specifies whether to include table column headings in the output files
SINGLE	TRUE or FALSE (default) – specifies whether to unload the result set into a single output file (will impact performance if TRUE); generally used with MAX_FILE_SIZE parameter
INCLUDE_QUERY_ID	TRUE or FALSE (default) – specifies whether to include the query ID of the COPY statement in the output file name

josegonzalezv@santander.com - mar 11-SEP-2023 - Snowflake
2023-do-not-cut

© 2023 Snowflake Computing Inc. All Rights Reserved

FILE PATHS AND NAMES

- Default file names take the form `data_0_0_0.<extension>`
  - Numbers after the prefix ("data") change to make sure the file names are unique

- To set the file **path** for the files, add the path after the stage name:

```
COPY INTO @mystage/CSVFiles/
FROM mytable
```

- To change the file name prefix of "data," add the desired prefix after the path:

```
COPY INTO @mystage/CSVFiles/testfile
FROM mytable
```

- Changes above yield files names: `@mystage/CSVFiles/testfile_x_y_z.<ext>`

jcgegonzalez@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

ADVANCED UNLOAD TOPICS

Additional topics are covered in the Snowflake Advanced course, such as:

- Unloading structured data into JSON or Parquet format
- Unloading semi-structured data into CSV format
- Automatically partitioning unloaded rows into Parquet files
  - Automatically creates a meaningful path

josegonzalezv@santander.com.mx - 11-SEP-2023 - @Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

LIST COMMAND

- Use the LIST command to list the files stored in a stage object:

```
LIST @%monthly_sales_agg;
```

name	size	md5	last_modified
data	144	055c5ee1d38271580a6a925ad1a69d45	Sun, 2 Dec 2018 17:48:22 GMT
data_0_0_0.csv	320	aefac11b7e3b543809988ea3f90924c5	Sun, 2 Dec 2018 18:15:06 GMT
data_0_0_0.csv.gz	144	0f01461eab63268987260ae6622c3ec6	Sun, 2 Dec 2018 18:15:19 GMT
data_0_0_0.json.gz	208	72e0f0d577b590233cda48464184204b	Sat, 1 Dec 2018 16:44:43 GMT
monthly_sales_dec_0_0_0.json.gz	208	37888eaf7698a6ac9eee178eaa509e57	Sat, 1 Dec 2018 16:47:26 GMT

jose.gonzalez.ic@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023 do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

REMOVE COMMAND

- Use the REMOVE command to delete files stored in an internal stage:

```
REMOVE @%monthly_sales_agg PATTERN='.*data.*';
```

name	result
data_0_0_0.json.gz	removed
data_0_0_0.csv.gz	removed
data	removed
data_0_0_0.csv	removed

josegonzalezy@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

```
LIST @%monthly_sales_agg;
```

name	size	md5	last_modified
monthly_sales_dec_0_0_0.json.gz	208	37888eaf7698a6ac9eee178eaa509e57	Sat, 1 Dec 2018 16:47:26 GMT

© 2023 Snowflake Computing Inc. All Rights Reserved

LOAD/UNLOAD DIFFERENCES

LOAD	UNLOAD
File formats: CSV, JSON, Avro, ORC, Parquet, XML	File formats: CSV, JSON, Parquet
No joins, filters, or aggregations	No limitations on SELECT features
Into TABLE from STAGE	Into STAGE from TABLE
Data goes into only a single table	Data can come from any number of tables

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake
2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

LAB EXERCISE: 12

Unloading Structured Data

10 minutes

Learning Objectives:

- How to unload table data into a Table Stage in Pipe-Delimited File format
- How to write a SQL statement containing a JOIN to unload a table into an internal stage

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake
2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

KNOWLEDGE CHECK

TRUE OR FALSE?

Using `AUTO_INGEST`, Snowpipe can load files either from an internal stage or an external stage.

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

KNOWLEDGE CHECK

TRUE OR FALSE?

Using `AUTO_INGEST`, Snowpipe can load files either from an internal stage or an external stage.

FALSE

`AUTO_INGEST` must use an external stage, because it relies on notification from the cloud provider (which has no visibility to internal stages).

josegonzalezv@santander.com.mx - 11-SEP-2023 @Snowflake
2023-do-not-copy

FUNCTIONS AND PROCEDURES

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake
2023 - do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

MODULE AGENDA

- User-Defined Functions
- JavaScript Stored Procedures
- Snowflake Scripting

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

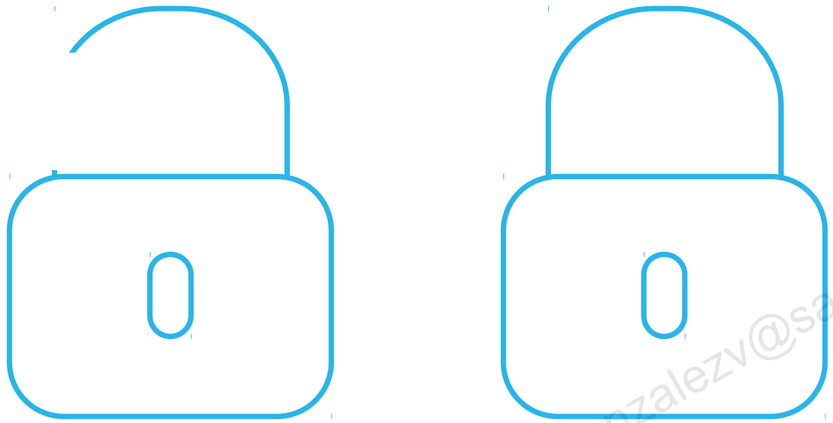
© 2023 Snowflake Computing Inc. All Rights Reserved

USER-DEFINED FUNCTIONS

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake
2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

USER-DEFINED FUNCTIONS



- Perform custom operations that are not available through the built-in functions
- Can be written in:
  - JavaScript
  - SQL
  - Java
  - Python
- No DDL/DML support
- Can be standard or secure
- Return a singular scalar value or, if defined as a table function, a set of rows

josegonzalezv@santander.com.mx - 11-SEP-2023 @Snowflake
2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

SQL UDF EXAMPLE

```
CREATE  FUNCTION order_cnt(custkey number(38,0))
RETURNS number(38,0)
AS
$$
  SELECT COUNT(1)
  FROM SNOWFLAKE_SAMPLE_DATA.TPCH_SF1.ORDERS
  WHERE o_custkey = custkey
$$;


SELECT C_name, C_address, order_cnt(C_custkey)
FROM SNOWFLAKE_SAMPLE_DATA.TPCH_SF1.CUSTOMER;
```

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

SQL TABLE UDF EXAMPLE

```
CREATE FUNCTION get_members_for_state(state_name VARCHAR)
RETURNS TABLE (fullname VARCHAR, state VARCHAR)
AS
$$
  SELECT DISTINCT( firstname || '' || lastname) fullname, state
  FROM members
  WHERE state = state_name
$$;

SELECT * FROM TABLE (get_members_for_state('California'));
```

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

JAVASCRIPT UDF

```
CREATE FUNCTION convert2fahrenheit(kelvin double)
RETURNS double
LANGUAGE JAVASCRIPT
AS
$$
   return (KELVIN * 9/5 - 459.67)
$$;


SELECT convert2fahrenheit(290);
```

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

JAVASCRIPT STORED PROCEDURES

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake
2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

STORED PROCEDURES

JS

JAVASCRIPT

- JavaScript provides the control structures
- SQL is executed within the JavaScript by calling functions in an API
- Argument names are:
  - Case-insensitive in the SQL portion of stored procedure code
  - Case-sensitive in the JavaScript portion

josegonzalezv@santander.com.mx - 11-SEP-2023 @Snowflake
2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

COMPONENTS OF A STORED PROCEDURE

Name

```
CREATE PROCEDURE stproc1(float_param1 FLOAT)
RETURNS STRING
LANGUAGE JAVASCRIPT
STRICT
AS
$$  // marks beginning and end of the code
try {
     snowflake.execute (
      {sqlText: "INSERT INTO stproc_test_table1 (num_col1)
       VALUES (" +FLOAT_PARAM1 + ")"} );
       return "Succeeded"; // Return status
       } catch (err) {
          return "Failed" + err; // status
       }
$$ ;

call stproc1(5.14::FLOAT);
```

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023 - Do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

COMPONENTS OF A STORED PROCEDURE

Data type to return

```
CREATE PROCEDURE stproc1(float_param1 FLOAT)
RETURNS STRING
LANGUAGE JAVASCRIPT
STRICT
AS
$$  // marks beginning and end of the code
try {
     snowflake.execute (
      {sqlText: "INSERT INTO stproc_test_table1 (num_col1)
       VALUES (" +FLOAT_PARAM1 + ")"} );
       return "Succeeded"; // Return status
       } catch (err) {
          return "Failed" + err; // status
       }
$$ ;

call stproc1(5.14::FLOAT);
```

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023 - Do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

COMPONENTS OF A STORED PROCEDURE

```
CREATE PROCEDURE stproc1(float_param1 FLOAT)
RETURNS STRING
LANGUAGE JAVASCRIPT
STRICT
AS
$$  // marks beginning and end of the code
try {
     snowflake.execute (
      {sqlText: "INSERT INTO stproc_test_table1 (num_col1)
       VALUES (" +FLOAT_PARAM1 + ")"} );
       return "Succeeded"; // Return status
       } catch (err) {
          return "Failed" + err; // status
       }
$$ ;

call stproc1(5.14::FLOAT);
```

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023 - Do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

COMPONENTS OF A STORED PROCEDURE

Statements to execute

```
CREATE PROCEDURE stproc1(float_param1 FLOAT)
RETURNS STRING
LANGUAGE JAVASCRIPT
STRICT
AS
$$  // marks beginning and end of the code
try {
     snowflake.execute (
      {sqlText: "INSERT INTO stproc_test_table1 (num_col1)
      VALUES (" +FLOAT_PARAM1 + ")"} );
      return "Succeeded"; // Return status
      } catch (err) {
         return "Failed" + err; // status
      }
$$ ;

CALL stproc1(5.14::FLOAT);
```

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023 - Do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

COMPONENTS OF A STORED PROCEDURE

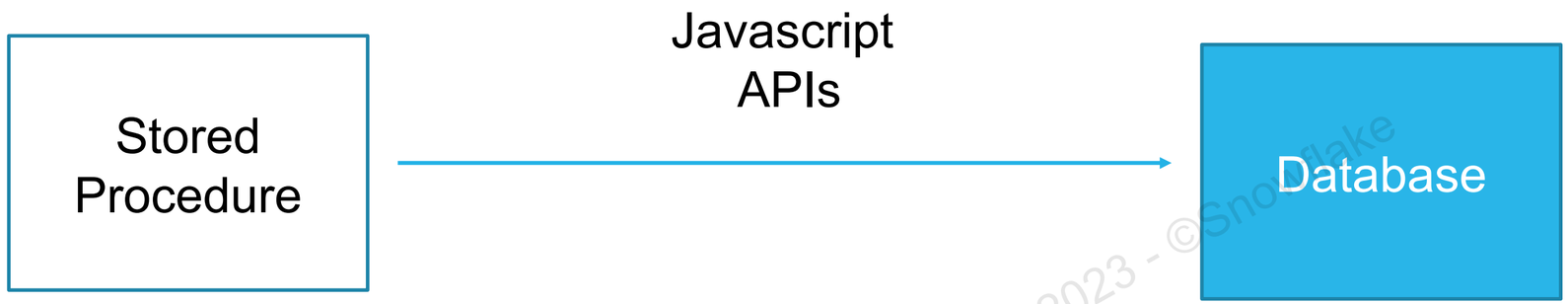
```
CREATE PROCEDURE stproc1(float_param1 FLOAT)
RETURNS STRING
LANGUAGE JAVASCRIPT
STRICT
AS
$$  // marks beginning and end of the code
try {
     snowflake.execute (
      {sqlText: "INSERT INTO stproc_test_table1 (num_col1)
       VALUES (" +FLOAT_PARAM1 + ")"} );
       return "Succeeded"; // Return status
       } catch (err) {
          return "Failed" + err; // status
       }
$$ ;
```

Invoke the stored procedure → `CALL stproc1(5.14::FLOAT);`

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023 - Do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

HOW IT WORKS



- Snowflake-provided JavaScript APIs
  - JavaScript objects and methods
  - Error handling and procedural logic
- SQL and database access/operations
  - Embed SQL code in the JavaScript
  - The SQL code executes database operations
  - Migrate other databases' stored procedure (SQL) code by embedding the SQL in JavaScript

jrsegomezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

STORED PROCEDURE SQL

EXECUTED THROUGH API OBJECTS

Object Class	Description
Snowflake Object	Contains the methods in the stored procedure API. Accessible by default to JavaScript code (you do not need to create the object).
Statement Object	Provides the methods for executing a query statement, and accessing metadata (for example, information about columns and rows) about the statement
ResultSet Object	Contains the results returned by a query (zero or more rows, each with one or more columns). You iterate through a result set by repeating next() and taking some action
SfDate Object	JavaScript does not have a built-in data type that corresponds to Snowflake SQL TIMESTAMP data types. Used when you want to retrieve a TIMESTAMP and stored it as a variable.

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake
2023-demo-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

STORED PROCEDURES VS UDFS

STORED PROCEDURE

- MAY return a value
- DDL and DML permitted
- Supports transactions inside procedure
- Can run using owning or calling role's privileges
  - Owning role is the default
  - Specified at creation time

USER-DEFINED FUNCTION

- MUST return a value
- DDL and DML not permitted
- No transaction support
- Runs using owning role's privileges

josegonzalez@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

LAB EXERCISE: 13

User-Defined Functions and User-Defined Table Functions

10 minutes

Learning Objectives:

- How to create a JavaScript user-defined function
- How to create a SQL user-defined function
- How to create a SQL user-defined table function

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

LAB EXERCISE: 14

Challenge Lab: Writing User-Defined Functions and User-Defined Table Functions

25 minutes

Learning Objectives:

- How to write a SQL user-defined function (UDF) that returns a single value.
- How to write a SQL user-define table function (UDTF) that returns a set of rows.

Note: A solution is presented on the last page of this lab in the workbook.

josegonzalezv@santander.com.mx - 14-SEP-2023 - ©Snowflake
2023 do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

SNOWFLAKE SCRIPTING

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake
2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

WHAT IS SNOWFLAKE SCRIPTING?

```
CREATE PROCEDURE my_test
   (my_arg INT)
RETURNS VARIANT
LANGUAGE SQL
AS
$$
DECLARE
   my_var NUMBER;
BEGIN
   LET my_cur CURSOR FOR
   SELECT c1 FROM T1;
...
$$
;
```

An extension to Snowflake SQL that adds support for procedural logic

- Create and save complex SQL flows, including SQL stored procedures
- Write standalone scripts and stored procedures that combine SQL statements and expressions with procedural code
- Similarities with Oracle's PL/SQL and Microsoft's T-SQL

josegonzalezv@santander.com.mx - 11-SEP-2023 @Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

SNOWFLAKE SCRIPTING FEATURES

- Execution blocks
- Control flow – branching and looping
- Scoped variables
- Exception handling
- Cursors
- Dynamic SQL
- Table return types
- Built-in status codes
- Caller's and owner's rights for stored procedures
- Secure stored procedures

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

BLOCKS

```
DECLARE
   <variable/cursor declarations, etc.>

BEGIN
   <Snowflake scripting and SQL statements>

EXCEPTION
   <Statements for handling exceptions>

END;
```

- Blocks are the basic units of construction
  - Blocks can be nested
  - Blocks can be executed via SnowSQL
- `BEGIN` / `END` required
- `DECLARE` and `EXCEPTION` optional
- Stored procedure: block wrapped in a definition and saved inside Snowflake
- Anonymous block: ad-hoc code block scripts executed in-line (not in Stored Procedure)

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

SNOWFLAKE SCRIPTING

- In-line statements can be wrapped with procedural logic

```
EXECUTE IMMEDIATE $$
DECLARE
    return_msg := 'code ran successfully';
BEGIN
      USE SCHEMA training_db.raw;
      CREATE TABLE PARENT (id INTEGER);
      CREATE TABLE CHILD (id INTEGER, parent_id INTEGER);
    RETURN return_msg;
EXCEPTION
WHEN OTHER THEN
  RETURN 'Failure ' || SQLERRM;
END;
$$;
```

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake
2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

VARIABLE DECLARATION AND ASSIGNMENT

```
DECLARE
   var1 INT;
   var2 DEFAULT 2;
   var3 INT DEFAULT 3;
   var4 VARCHAR := (SELECT CURRENT_USER());

BEGIN
   LET var5 INT DEFAULT 5;
   LET var6 INT :=6;
   ...
   INSERT INTO tbl1 (id) VALUES (:var6);
   ...
   SELECT COUNT(*), MIN(id), MAX(id)
      INTO var1, var2, var3
      FROM tbl1 WHERE id IS NOT NULL;
   ...
```

- Variables must be declared before use, in the `DECLARE` and/or `BEGIN` sections
  - Types can be specified or inferred
  - Different syntax for `DECLARE` and `BEGIN`
- Can use bind variables in your SQL
- Set variables to column values specified in a SELECT clause with INTO
- Variables are local to a block
  - Nested blocks can access variables in outer blocks
  - Local named objects take precedence

jsogonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

BRANCHING

Construct	Syntax

IF/ELSE	IF <condition> THEN...
	  ELSEIF <condition> THEN...
	  ELSE...
	END IF;

CASE (simple)	CASE (<expression_to_evaluate>)
	  WHEN <evaluates_to_value1> THEN...
	  WHEN <evaluates_to_value2> THEN...

CASE (searched)	CASE
	  WHEN <condition> THEN...
	  WHEN <condition> THEN...

EXCEPTION	WHEN STATEMENT_ERROR THEN…

josegonzalez@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023 - do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

LOOPING

Construct	Description
`LOOP`	Continues iterating until a `BREAK` or `RETURN` occurs inside the loop
`WHILE`	Iterates while a condition is `TRUE`
`REPEAT`	Iterates until a condition is `TRUE`
`FOR`	Repeats a sequence of steps a specific number of times
`CONTINUE`	Enables skipping loop iterations
`BREAK`	Exits a loop completely

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake
2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

LAB EXERCISE: 15

Snowflake Stored Procedures and Snowflake Scripting

25 minutes

Learning Objectives:

- Write and call stored procedures
- Use Snowflake Scripting to add procedural logic to a stored procedure
- Write an IF-THEN statement and a CASE statement
- Open a cursor and iterate through rows
- Apply a concatenation operator to join two or more strings
- Write a stored procedure in JavaScript

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

LAB EXERCISE: 16

Challenge Lab: Snowflake Stored Procedures and Snowflake Scripting

35 minutes

Learning Objectives:

- Create a stored procedure that can insert rows into another table
- Implement a CASE statement in a stored procedure
- Implement a FOR loop in a stored procedure
- Successfully iterate through rows in a cursor

Note: A solution is presented on the last page of this lab in the workbook.

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake
2023 do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

KNOWLEDGE CHECK

Indicate whether each of the following statements describe a user-defined function or a stored procedure.

A. Can run with either owner's rights or caller's rights

B. Must return a value

C. Can use DDL or DML statements

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

KNOWLEDGE CHECK

Indicate whether each of the following statements describe a user-defined function or a stored procedure.

A. Can run with either owner's rights or caller's rights

B. Must return a value

C. Can use DDL or DML statements

A. A **stored procedure** can use either owner's rights, or caller's rights. A UDF always uses owner's rights.

B. A **UDF** must return a value.

C. A **stored procedure** can include DDL or DML statements; a UDF cannot.

josegonzalezv@santander.com.mx - 11-SEP-2023 @Snowflake
2023-do-not-copy

KNOWLEDGE CHECK

How do you invoke a stored procedure?

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

KNOWLEDGE CHECK

How do you invoke a stored procedure?

You invoke a stored procedures with a **CALL** statement. For example:

```
CALL <my_proc>(<variable>);
```

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

KNOWLEDGE CHECK

What Snowflake feature allows you to write stored procedures in SQL?

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

KNOWLEDGE CHECK

What Snowflake feature allows you to write stored procedures in SQL?

Snowflake Scripting

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

KNOWLEDGE CHECK

What kind of function does the following statement use?

```
SELECT
  SUM(col1)
FROM
  mytable;
```

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

KNOWLEDGE CHECK

What kind of function does the following statement use?

```
SELECT
  SUM(col1)
FROM
  mytable;
```

This is an example of an aggregate function. An aggregate function performs an operation on a group of rows and returns a single value.

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

USING TASKS

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

MODULE AGENDA

- Overview
- Creating Tasks
- Managing Tasks

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

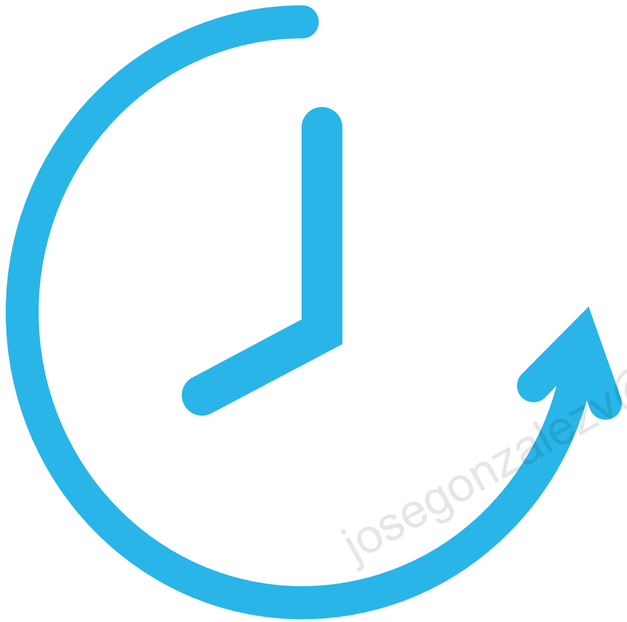
© 2023 Snowflake Computing Inc. All Rights Reserved

OVERVIEW

josegonzalezv@santander.com.br - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

WORKING WITH TASKS

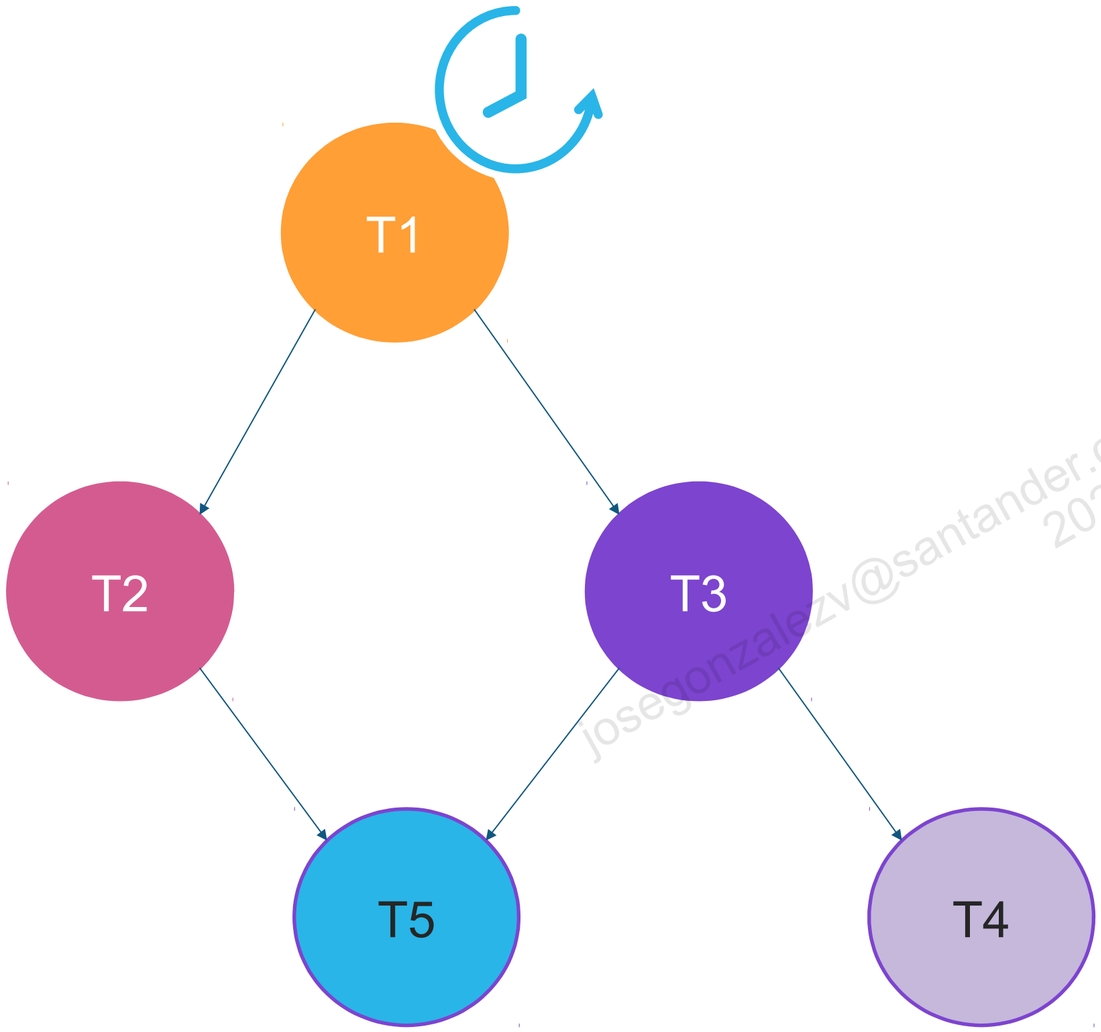


- Schedule recurring work
- Execute a:
  - DDL/DML/SQL statement
  - Stored procedure call
  - Snowflake Scripting code
- Run on a defined schedule, or after completion of a previous task
- Use Cases:
  - Keep aggregates up-to-date
  - Generate periodic reports
  - Load data
  - Others

josegonzalezv@santander.com.mx - 11-SEP-2023 - @Snowflake
2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

TASK COMPLEXITY

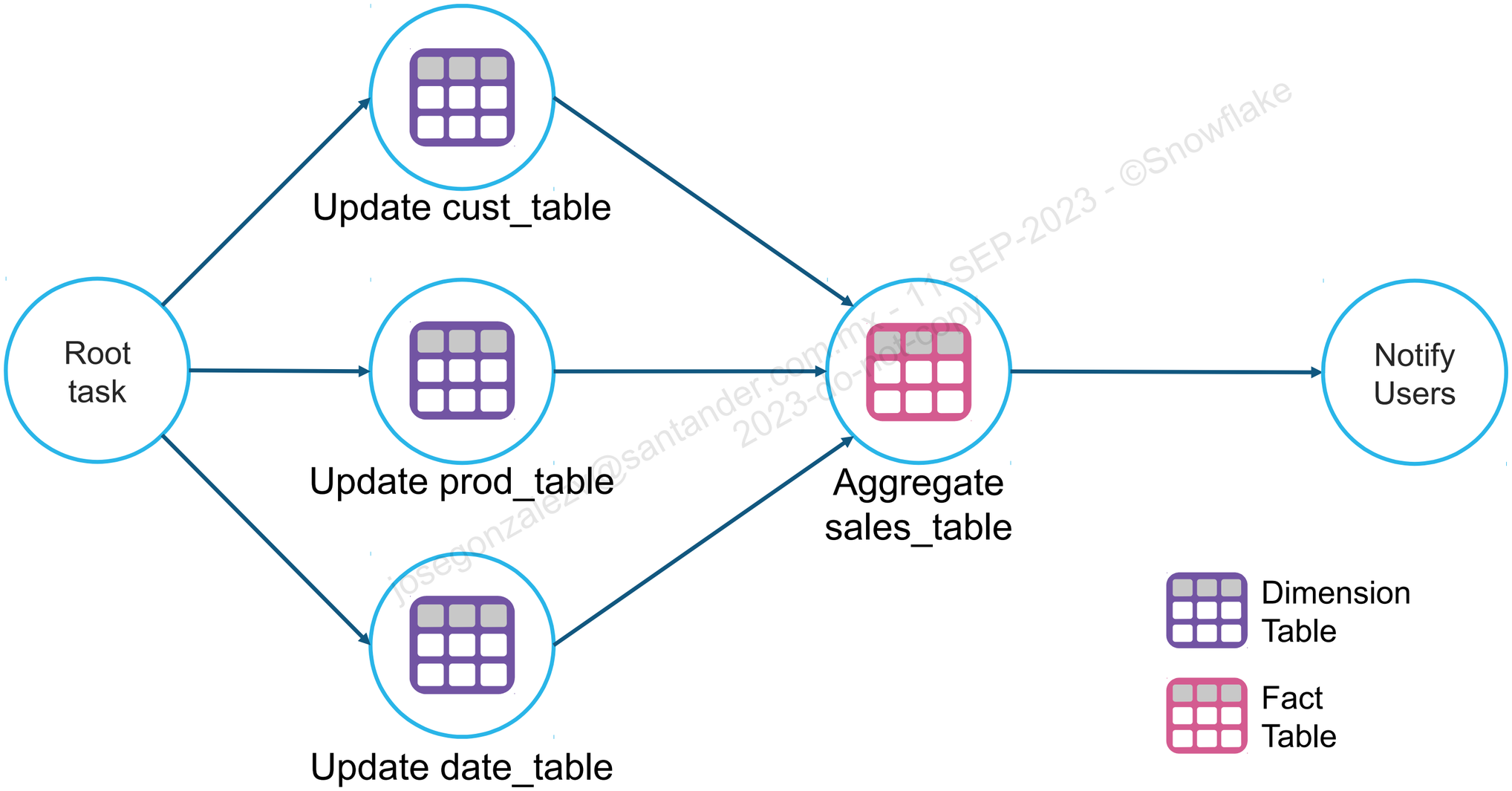


- Single task – always runs on a schedule
- Cascading tasks – a series of tasks organized by dependencies
  - DAG (Directed Acyclic Graph) of tasks
  - Single root task runs on a schedule
  - Child tasks triggered by successful completion of one or more predecessors
- Each DAG limited to 1000 total tasks
- Each task can have up to 100 children, and 100 predecessors

josegonzalezv@santander.com.mx - 11-SEP-2025 - Snowflake
2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

SAMPLE USE CASE



josegonzalez@santander.com.mx - 11-SEP-2023 - ©Snowflake
2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

COMPUTE FOR TASKS

Virtual Warehouse

- User-designated virtual warehouse
- Make sure warehouse is large enough for task to finish before next run

Serverless

- Compute allocated and dynamically scaled by Snowflake
- Will use task history to determine compute required
- Optionally, set warehouse size for initial task run

`USER_TASK_MANAGED_INITIAL_WAREHOUSE_SIZE`

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

CHOOSING A COMPUTE MODEL

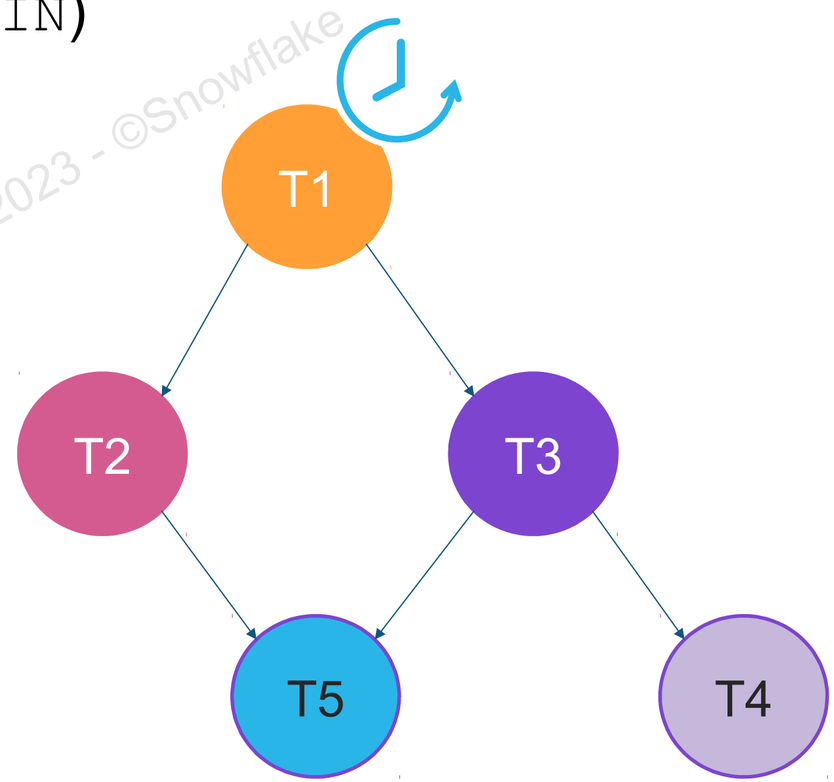
Task(s) will run to completion quickly (under 1 minute)	**Serverless.** A virtual warehouse will be charged for a full minute plus any auto-suspend time, which will likely cost more than serverless compute.
It is critical that a task completes before the next run	**Serverless.** Snowflake dynamically adds resources to ensure the task completes before the next scheduled run.
Task requires more compute than an XXLarge warehouse can provide	**Virtual Warehouse.** The maximum compute allocated using serverless compute is equivalent to an XXLarge virtual warehouse.
You want the task to finish as quickly as possible	**Virtual Warehouse.** Serverless will allocate enough compute to complete before the next scheduled run, but you can specify a virtual warehouse to complete the task more quickly.
Spiky, unpredictable loads	**Virtual Warehouse (possibly multi-cluster).** Since serverless compute bases initial resources on task history, it may not be as efficient if the work changes significantly between runs. You may want to test both models.
Many tasks run concurrently	**Virtual Warehouse.** If you can fully utilize a single virtual warehouse with concurrent tasks, specifying a virtual warehouse may be more efficient.

josegonzalez@santander.com.mx - 11-SEP-2025 - Do-not-copy - Snowflake

© 2023 Snowflake Computing Inc. All Rights Reserved

WORKFLOW

1. Create a Task Administrator Role (or use `ACCOUNTADMIN`)
2. Create the task(s)
   - Run serverless or specify a virtual warehouse
   - Define when the task will run
   - Define any condition
   - Indicate the SQL/stored procedure to run
3. Alter the tasks to resume them and start the schedule
   - Or execute immediately with `EXECUTE TASK`


T1
T2
T3
T5
T4 ovalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

CREATING TASKS

josegonzalezv@santander.com.mx 11-SEP-2023 - ©Snowflake 2023-do not copy

© 2023 Snowflake Computing Inc. All Rights Reserved

CREATE A ROOT TASK

```
CREATE TASK <task name>
   [WAREHOUSE = <wh name>]

   SCHEDULE = '<n> minute' |
   SCHEDULE = 'USING CRON <expr>'

   [WHEN <boolean condition>]
AS
   <SQL or stored procedure call>;
```

- Tasks run on serverless compute, unless you specify a virtual warehouse
- Root tasks can be scheduled every <n> minutes, or using CRON syntax
- Optionally specify a Boolean condition to check before running
  - If false, the task does not run
- Provide SQL statement or stored procedure call

jose.gonzalezv@santander.com.mx - 11-SEP-2023 @Snowflake
2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

SPECIFYING A SCHEDULE

Schedule can be expressed in two ways:

- `SCHEDULE = '<num> minute'`
  - Task will run every `<num>` minutes

- `SCHEDULE = 'USING CRON <expression> [<time_zone>]'`
  - Specifies a cron expression and time zone for running the task
  - Supports a subset of standard cron utility syntax

```
# ___________ minute (0-59)
# | _________ hour (0-23)
# | | _______ day of month (1-31, or L)
# | | | _____ month (1-12, JAN-DEC)
# | | | | _ day of week (0-6, SUN-SAT, or L)
# | | | | |
# | | | | |
  * * * * *
```

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023 do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

CREATE A CHILD TASK

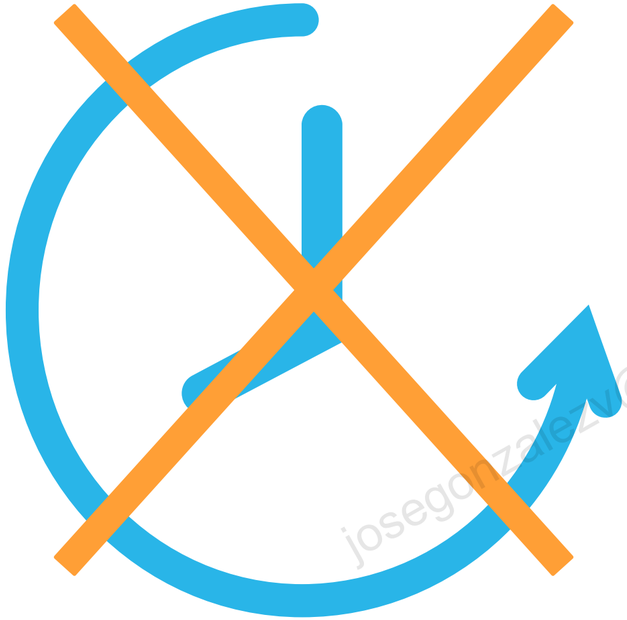
```
CREATE TASK <task name>
   [WAREHOUSE = <wh name>]

   AFTER <task> [, <task>…]

   [WHEN <boolean condition>]
AS
   <SQL or stored procedure call>;
```

- Only difference between root tasks and child tasks is how they are scheduled
- Child tasks run AFTER one or more predecessor tasks
  - Predecessors can be the root or child tasks
  - All predecessor tasks must complete successfully, or the child will not run
  - A child task cannot have two root tasks as a predecessor

josegonzalezv@santander.com.mx - 11-SEP-2023 - @Snowflake
2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

SERVERLESS TASK LIMITATIONS



- Serverless tasks do not support invoking:
  - UDFs with Java or Python code
  - Stored procedures that call UDFs with Java or Python code
  - Stored procedures written in Scala
  - `EXTRACT_SEMANTIC_CATEGORIES`

josegonzalez@santander.com.mx - 11-SEP-2023 ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

TASKS EXAMPLE #1

- Runs an ELT task every 60 minutes using a virtual warehouse:

```
CREATE TASK elt_employee
    WAREHOUSE = hr_wh
    SCHEDULE = '60 minute'
AS
    COPY INTO employee_raw_data FROM @my_stage;
```

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

TASKS EXAMPLE #2

- Runs a quarterly report using CRON syntax

```
CREATE TASK previous_quarter_results
   SCHEDULE = 'USING CRON 0 0 1 1,4,7,10 *'
AS
   CREATE OR REPLACE TABLE prev_quarter CLONE finance_data;
```

josegonzalezv@santander.com.mx - 1 - SEP-2023 - @Snowflake
2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

TASKS EXAMPLE #3

- Simple tree with a root task using a virtual warehouse, and a serverless child task

```
CREATE TASK elt_employee
    WAREHOUSE = hr_wh
    SCHEDULE = '60 minute'
AS
    COPY INTO employee_raw_data FROM @my_stage;

CREATE TASK process_employee
    AFTER elt_employee
AS
INSERT INTO employees(first_name, last_name)
SELECT firstname, lastname
FROM employee_raw_data
WHERE status = 'new';
```

josegonzalez@santander.com.mx - 11-SEP-2023 - ©Snowflake
2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

SOME TASK OPTIONS

OPTION	USE
`<session_parameter> = <value> [, <session_parameter> = <value>…]`	• Any **session** parameter can be set in the definition of a task
`ALLOW_OVERLAPPING_EXECUTION`	• FALSE (default) or TRUE • Only set on the root task • Applies to all tasks in the DAG
`USER_TASK_TIMEOUT_MS`	• Default: 3600000 (1 hour) • Can be set from 0 to 864000000 (1 day)
`SUSPEND_TASK_AFTER_NUM_FAILURES`	• Number of consecutive failures before a task will be suspended • Default is 0 (will never suspend automatically) • Does not include tasks that are skipped or canceled • Does include tasks that time out • Root task suspended after any single task in a DAG fails the specified number of times

josegonzalezv@santander.com.mx - 11 SEP 2023 - ©Snowflake
2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

MANAGING TASKS

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake
2023-db-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

EXECUTE TASKS

- `ALTER TASK [ resume | suspend ]`
  - New tasks must be resumed before they will run
- `EXECUTE TASK`
  - To run a task immediately
- `SYSTEM$TASK_DEPENDENTS_ENABLE` function
  - Recursively resumes all tasks in a specified DAG

josegonzalez@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

RETRIEVE TASK INFORMATION

- `SYSTEM$CURRENT_USER_TASK_NAME`
  - Invoked from the statement or stored procedure defined by the task
  - Returns the name of the task currently executing
- `TASK_DEPENDENTS` table function
  - Returns a list of child tasks for a given root task
  - By default, lists only direct child tasks (set `RECURSIVE => TRUE` to list all tasks in the DAG)
- `CURRENT_TASK_GRAPHS` table function
  - Returns the status of a task or DAG scheduled withing 8 days or currently executing

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

RETRIEVE TASK INFORMATION

- **`SERVERLESS_TASK_HISTORY`** view
  - Useful for tracking credits consumed by serverless tasks
  - In `SNOWFLAKE.ACCOUNT_USAGE`
  - Latency up to 3 hours

- **`TASK_VERSIONS`** view
  - Retrieves the history of task versions (changes to tasks over time)
  - In `SNOWFLAKE.ACCOUNT_USAGE`
  - Latency up to 3 hours

josegonzalez@salesfender.com.mx - 11-SEP-2023 - ©Snowflake
2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

VIEW TASK HISTORY

- `TASK_HISTORY()` table function
  - Returns task history within the last 7 days or the next scheduled execution within the next 8 days
- `TASK_HISTORY` view in Snowflake database
  - Returns task history within the last year
- `COMPLETE_TASK_GRAPHS` table function
  - Returns details for task runs in the past hour
  - Can specify to only return tasks with errors
  - Limited to 10,000 rows
- `COMPLETE_TASK_GRAPHS` view
  - Avoids the 10,000 row limitation of the table function of the same name

josegonzalez@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

TROUBLESHOOT TASKS

- Task fails or does not run:
  - Check `TASK_HISTORY` for error code and message
  - Verify the predecessor task(s) completed successfully
  - Verify the task state was `RESUMED`
  - Verify the task owner has all required privileges
  - If a condition is specified, verify that it evaluates to `TRUE`

- Task times out or gets skipped
  - Check `TASK_HISTORY` to determine if task timed out, or exceeded scheduled window
  - Check timeout and (if necessary) increase timeout limit (default is one hour)
  - Verify warehouse size, it may need to be larger
  - Timeout or warehouse size may not resolve the issue if the problem is query parallelization

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake
2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

LAB EXERCISE: 17

Creating, Starting and Stopping a Task

20 minutes

Learning Objectives:

- Create tasks.
- Run tasks on a schedule, or immediately.
- Use functions to view task successes and failures.
- Troubleshoot task failures.
- Create root tasks and child tasks.

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

LAB EXERCISE: 18

Challenge Lab: Creating, Starting and Stopping a Task

25 minutes

- Create a task
- Start a task
- Use table function `information_schema.task_history()` to monitor tasks
- Stop a task

Note: A solution is presented on the last page of this lab in the workbook.

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

KNOWLEDGE CHECK

Will the following create a root task, or a child task?

```
CREATE TASK my_task
  WAREHOUSE = xsmall_wh
  AFTER my_other_task
AS
  CALL my_stored_proc();
```

josegonzalez@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

KNOWLEDGE CHECK

Will the following create a root task, or a child task?

```
CREATE TASK my_task
  WAREHOUSE = xsmall_wh
  AFTER my_other_task
AS
  CALL my_stored_proc();
```

This creates a child task, since it is schedule to run `AFTER` another task.

josegonzalez@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

KNOWLEDGE CHECK

What are the two ways you can schedule a root task?

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

KNOWLEDGE CHECK

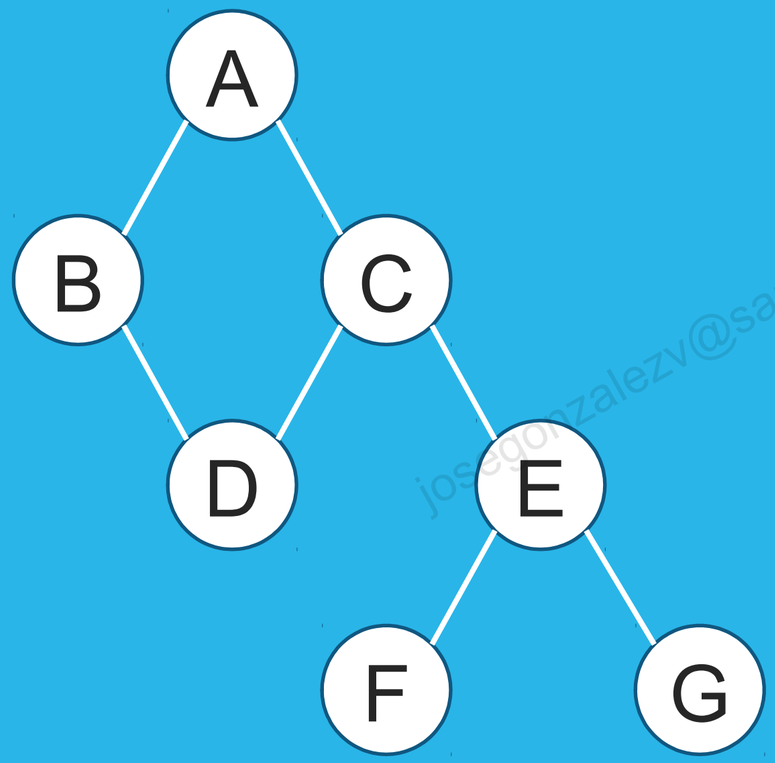
What are the two ways you can schedule a root task?

```
SCHEDULE = '<n> minute'

SCHEDULE = 'USING CRON <expr>'
```

josegonzalezv@santander.com.mx - 11-SEP-2023 ©Snowflake 2023-do-not-copy

KNOWLEDGE CHECK

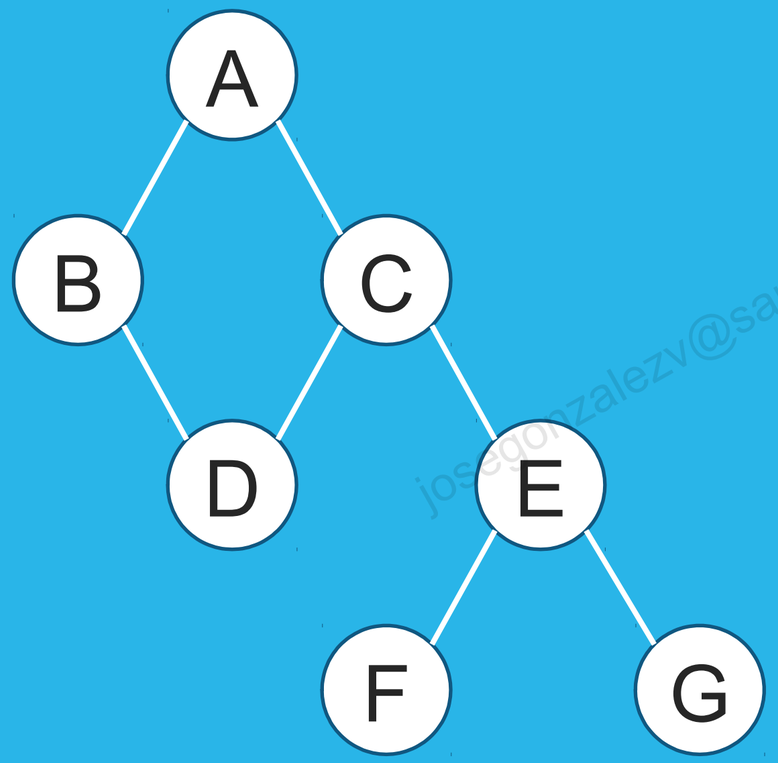
In the following tree of tasks: if B fails, which of the other tasks will run?



josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

KNOWLEDGE CHECK

In the following tree of tasks: if B fails, which of the other tasks will run?



If task B fails, then task D will not run.
All of the other tasks will run.

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

MANAGING SECURITY

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake
2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

MODULE AGENDA

- Security Overview
- Access
- Authentication
- Authorization
- Data Protection

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

SECURITY OVERVIEW

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake
2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

SECURITY AT A GLANCE

ALL EDITIONS COMBINED

Access

Authentication

Authorization

Data Protection

Snowflake operational controls

- NIST 800-53
- SOC2 Type 2
- SOC1 Type II
- ISO/IEC 27001:2013
- HIPAA
- PCI
- FedRAMP

josegonzalezv@santander.mx - 11-SEP-2023 - @Snowflake
2023 - do not copy

© 2023 Snowflake Computing Inc. All Rights Reserved

SNOWFLAKE LAYERED SECURITY

Data Protection – Protect customer data using AES 256-bit encryption, key rotation, and the option of periodic rekeying

Access	Authentication	Authorization
Can you access the account?	Who are you?	What can you do?

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

ACCESS

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake
2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

ACCESS

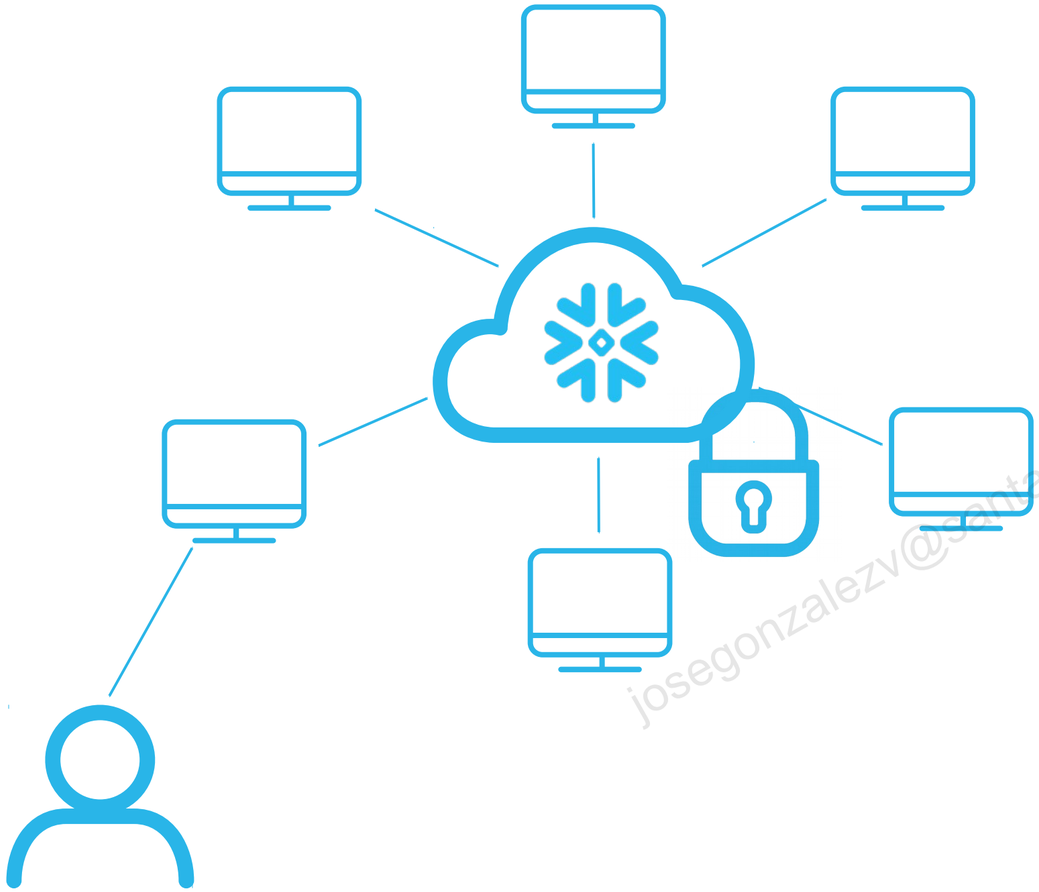
CAN YOU GET TO THE SNOWFLAKE ACCOUNT?

- IP lists allow or disallow access
  - Can be set at the account and user levels
- All client communications encrypted with TLS 1.2
- Support for cloud-provider private network solutions
  - Business Critical edition and above

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake
2023-do not copy

© 2023 Snowflake Computing Inc. All Rights Reserved

NETWORK POLICIES



- Define IP addresses that are allowed to connect to your Snowflake account
- `ALLOWED_LIST` contains a list or block of address that can connect
- Optional `BLOCKED_LIST` blocks some IP addresses in the `ALLOWED_LIST`
- Can be assigned at the account level, and for specific users

josegonzalezv@santander.com.mx - 11-SEP-2023 ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

CONFIGURE NETWORK POLICY

- User must be `ACCOUNTADMIN` or `SECURITYADMIN`
- `ALLOWED_IP_LIST` is required, `BLOCKED_IP_LIST` is optional

```
CREATE NETWORK POLICY <policy name>
   ALLOWED_IP_LIST = (<IP address or range>, ...)
   BLOCKED_IP_LIST = (<IP address(es) in allowed range>)
   COMMENT = '<string>';
```

- Example:
```
CREATE NETWORK POLICY account_policy
        ALLOWED_IP_LIST = ('200.50.1.0/24')
        BLOCKED_IP_LIST = ('200.50.1.40');
```

jvgonzalezv@santander.com.mx - 11-SEP-2025 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

SET NETWORK POLICY

```
ALTER ACCOUNT SET NETWORK_POLICY = <policy_name>;

ALTER USER <name> SET NETWORK_POLICY = <policy_name>
```

- Only one network policy can be active for an account or user
  - Different policies can be applied to different users
- The `BLOCKED_IP_LIST` is applied before the `ALLOWED_IP_LIST`
- If a user-level policy is applied while a user is logged in, their IP address is verified
- A user-level policy takes precedence over the account-level policy

jose.gonzalezv@santander.com.mx - 14-SEP-2023 ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

ENSURING CONNECTIVITY

- Allow the following ports:
  - 443: Required for general Snowflake traffic
  - 80: Required for OCSP cache server, which listens for all Snowflake client communication
- Hostname Allow:
  - All Snowflake clients require access to cloud storage
  - Allow hostname patterns for the required hosts (see documentation for details)
  - Use `SYSTEM$ALLOWLIST()` to obtain required hostnames for your Snowflake account

josegonzalez@santander.com.mx - 11-SEP-2023 - ©Snowflake
2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

AUTHENTICATION

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

AUTHENTICATION

WHO ARE YOU?

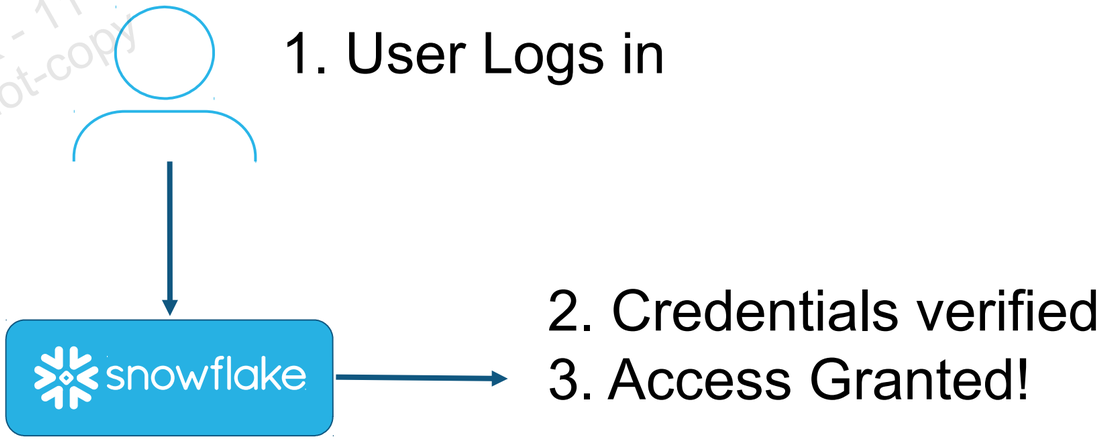
- Username and password
- Multi-factor authentication
- Federated authentication (SAML 2.0)
- Other authentication methods (OAuth, Key-Pair)

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023 - do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

USERNAME AND PASSWORD

- Verify the password given matches the one associated with the user
- Default method
- Default requirements:
  - 8-256 characters
  - At least one upper-case letter
  - At least one lower-case letter
  - At least one number



josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

PASSWORD POLICY

- Change password requirements and set for the account
- `CREATE PASSWORD POLICY` options:

Parameter	Min Value	Max Value	Default
PASSWORD_MIN_LENGTH	8	256	8
PASSWORD_MAX_LENGTH	8	256	256
PASSWORD_MIN_UPPER_CASE_CHARS	0	256	1
PASSWORD_MIN_LOWER_CASE_CHARS	0	256	1
PASSWORD_MIN_NUMERIC_CHARS	0	256	1
PASSWORD_MIN_SPECIAL_CHARS	0	256	1
PASSWORD_MAX_AGE_DAYS	0	999	0 (never require change)
PASSWORD_MAX_RETRIES	1	10	5
PASSWORD_LOCKOUT_TIME_MINS	1	999	15
COMMENT	-	-	None

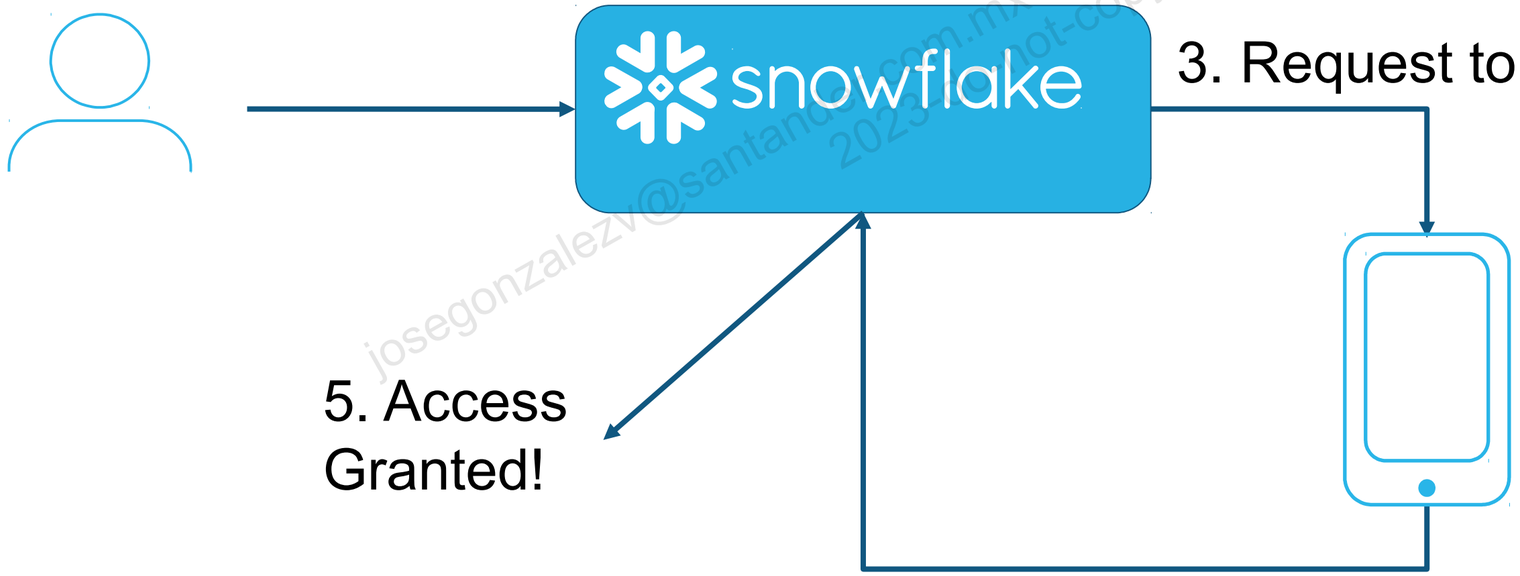
jsegovzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake
2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

MULTI-FACTOR AUTHENTICATION

MFA is an **opt-in** feature for users, it cannot be enabled for users by `SECURITYADMIN`

1. User enables MFA
2. User Logs in



3. Request to Duo

4. Confirm on smartphone, or enter passcode

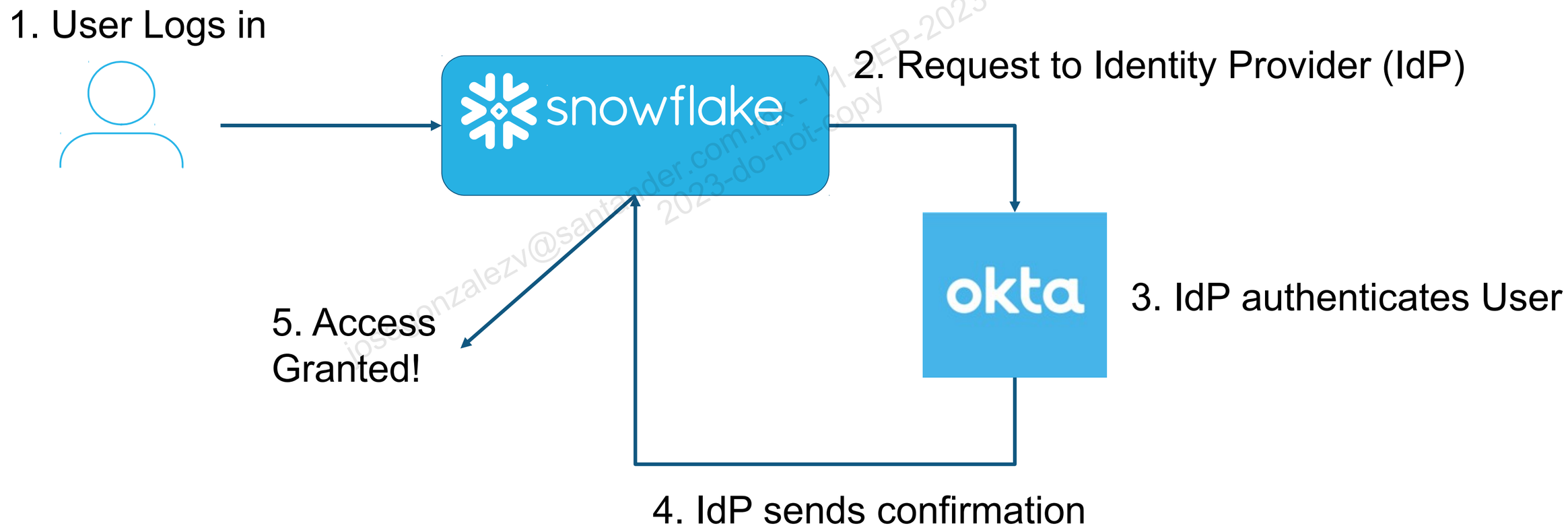
5. Access Granted!

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023 - Do not copy

© 2023 Snowflake Computing Inc. All Rights Reserved

FEDERATED AUTHENTICATION (SSO)

Authenticate through an external identity provider, such as Okta



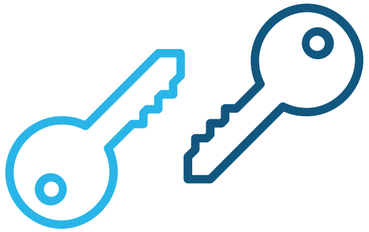
jose.gonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

OTHER AUTHENTICATION METHODS

- **OAuth**: Authorizations on behalf of a user from a 3rd party service or application

- **KeyPair**: JSON Web Token (JWTs) signed using a public/private key pair with RSA encryption



josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

AUTHORIZATION

josegonzalezv@santander.com - 11-SEP-2023 - ©Snowflake
2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

AUTHORIZATION

WHAT CAN YOU DO?

- Flexible and granular authorization
- Secure views and UDFs to protect information access
- Data Masking and Row Access Polices

josegonzalezv@santander.com.mx - 04-SEP-2023 @Snowflake
2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

AUTHORIZATION CONTROL

- Combination of Role-Based (RBAC) and Discretionary Control
  - Role-based: Access privileges are assigned to roles
  - Discretionary: Each object has an owner, which can grant others access to that object
- Privileges that can be set depend on the securable object
- Privileges can be broad or granular

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake
2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

SECURE VIEWS AND UDFs

- Prevent data from being indirectly exposed through programmatic methods
- Only authorized users can see the view or UDF definition
- Secure view and secure UDFs bypass some query optimizations to achieve greater security

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

ROW- AND COLUMN-LEVEL PROTECTION

- Protect row-level access
  - Row access policies
  - Context functions (such as `CURRENT_USER`, `CURRENT_ROLE`, etc.)
  - Secure views
- Protect column-level access
  - Dynamic data masking
  - External tokenization
  - Secure views

josegonzalezv@santander.com.mx - 11-SEP-2025 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

ADDITIONAL INFORMATION

- Snowflake has many data governance features that can protect your data beyond what can be done with role-based access control
- These are covered in Snowflake's 1-day Data Governance course that can be added on to any Snowflake course

josegonzalez@santander.com.mx - 11-SEP-2023 @Snowflake
2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

DATA PROTECTION

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

DATA PROTECTION

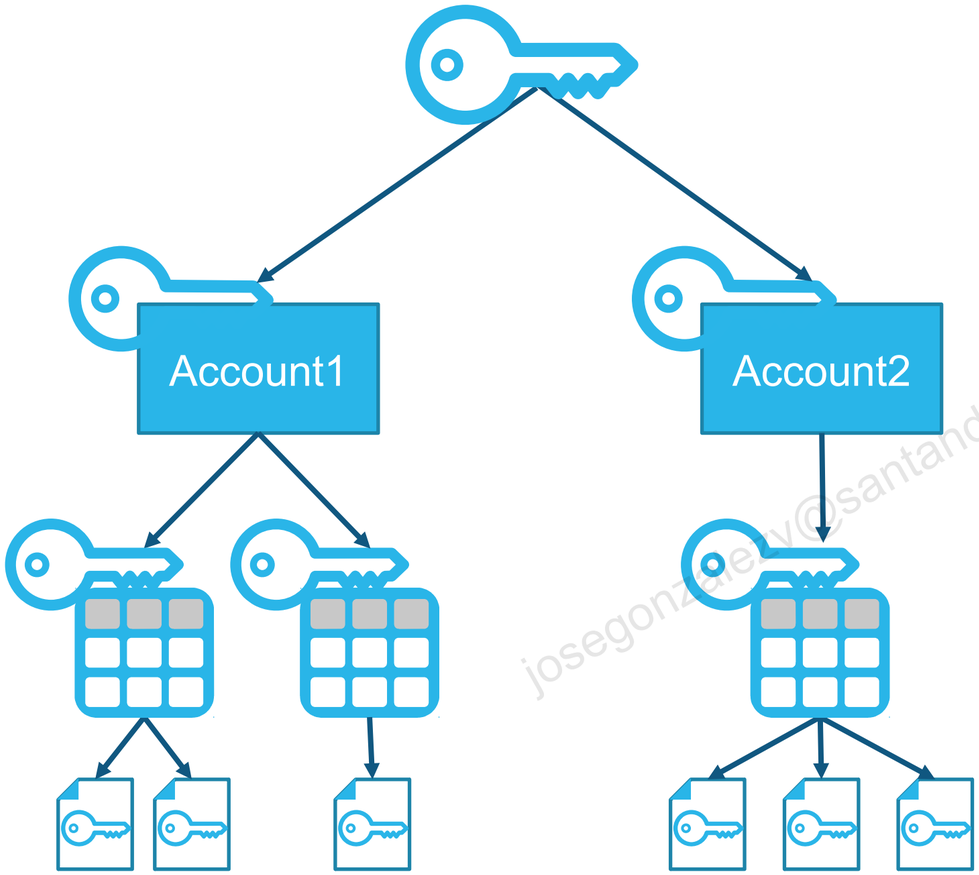
IS YOUR DATA SAFE?

- Encrypted at rest and in motion
- Replication across availability zones
- Time Travel, Fail-safe, and Replication Protection

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

ENCRYPTION

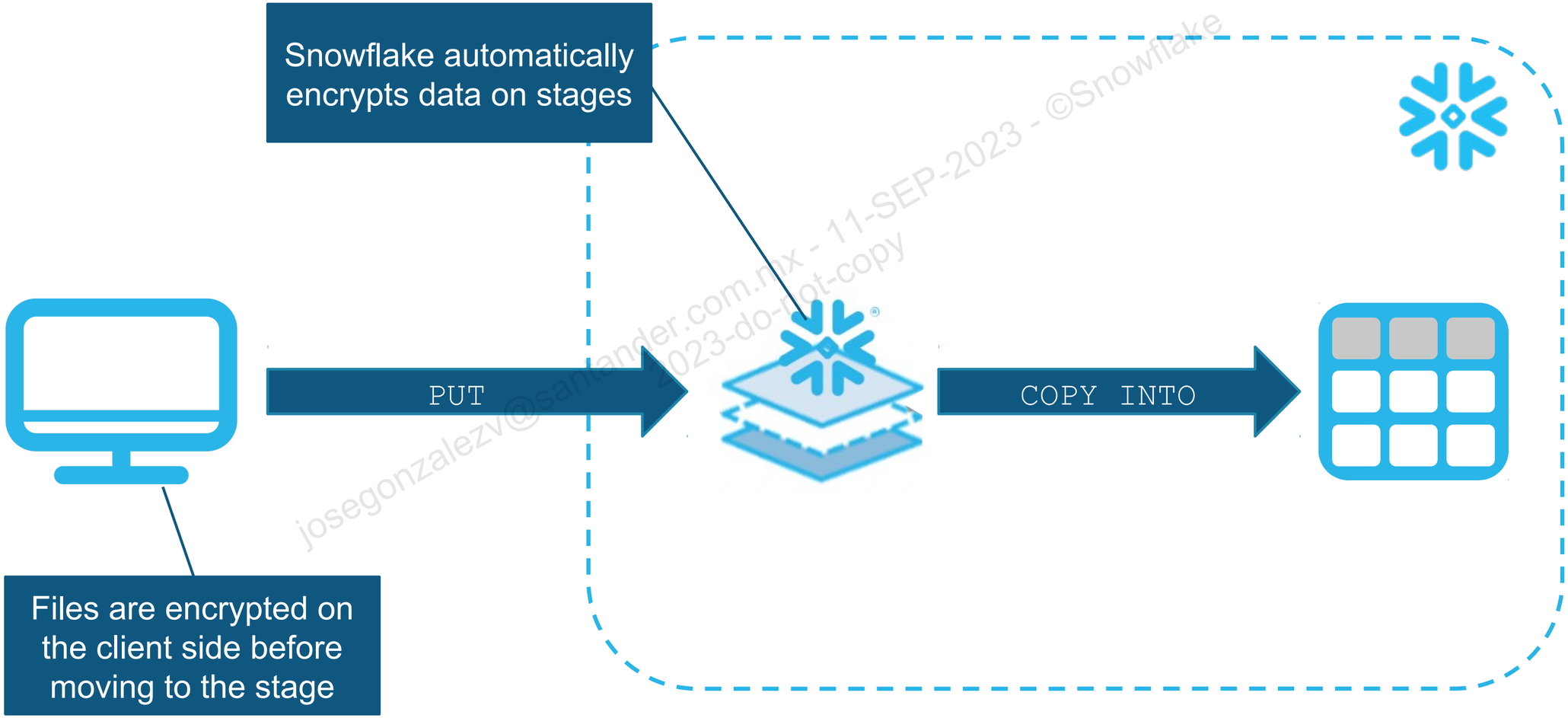


- Encryption at rest and in motion
- Hierarchical key model
- Automatic key rotation every 30 days
- Optional periodic rekeying (Enterprise edition and above)
- Bring Your Own Key (BYOK)

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

ENCRYPTION ON INTERNAL STAGES

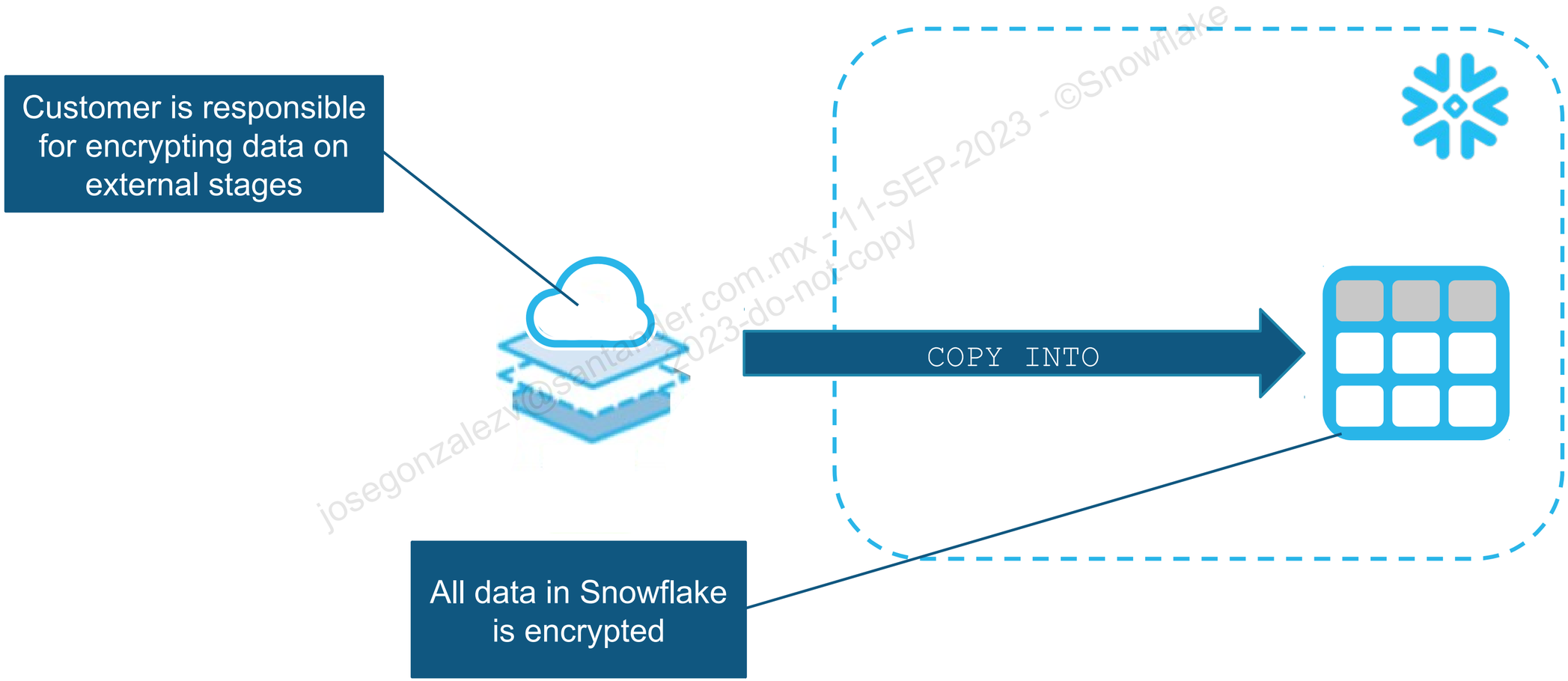


josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake
2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

ENCRYPTION ON EXTERNAL STAGES

ENCRYPTION/SECURITY MANAGED BY CUSTOMER



josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake
2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

KNOWLEDGE CHECK

What is the term for each of the following descriptions?

A. Can you attempt to log into the account?

B. Are you who you say you are?

C. What objects can you access?

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

KNOWLEDGE CHECK

What is the term for each of the following descriptions?

A. Can you attempt to log into the account?

B. Are you who you say you are?

C. What objects can you access?

A. Access

B. Authentication

C. Authorization

josegonzalezv@santander.com.mx - 11-SEP-2023 @Snowflake
2023-do-not-copy

KNOWLEDGE CHECK

TRUE of FALSE?

The `SECURITYADMIN` can set up Multi-Factor-Authentication (MFA) on any user in the system.

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

KNOWLEDGE CHECK

TRUE of FALSE?

The `SECURITYADMIN` can set up Multi-Factor-Authentication (MFA) on any user in the system.

FALSE

Users must opt into MFA and set it up through their profile.

However, the `SECURITYADMIN` can see who has MFA enabled and who does not, and disable users who are not enrolled in MFA.

josegonzalezv@santander.com.mx - 11-SEP-2023 ©Snowflake
2023-do-not-copy

KNOWLEDGE CHECK

What are some features of Snowflake that can add extra access controls on top of our Role-Based Access Control (RBAC) model?

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

KNOWLEDGE CHECK

What are some features of Snowflake that can add extra access controls on top of our Role-Based Access Control (RBAC) model?

- Row access policies
- Dynamic data masking
- External tokenization
- Context functions
- Secure views

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

KNOWLEDGE CHECK

How often does Snowflake rotate encryption keys?

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

KNOWLEDGE CHECK

How often does Snowflake rotate encryption keys?

All Snowflake-managed keys are automatically rotated by Snowflake when they are more than 30 days old.

josegonzalezv@santander.com.mx - 11-SEP-2023 @Snowflake
2023-do-not-copy

KNOWLEDGE CHECK

When is Snowflake responsible for encrypting files on a stage?

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

KNOWLEDGE CHECK

When is Snowflake responsible for encrypting files on a stage?

When it is an internal stage.

Customers must make sure files on their external stages are encrypted while at rest.

josegonzalezv@santander.com.mx - 11 SEP 2023 - ©Snowflake 2023-do-not-copy

ACCESS CONTROL AND USER MANAGEMENT

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake
2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

MODULE AGENDA

- Concepts
- Types of Roles
- Ownership
- View Grants

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

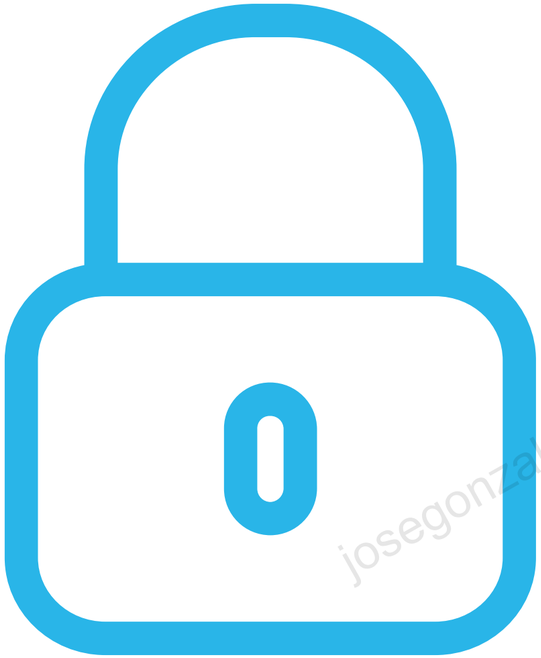
CONCEPTS

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake
2023-do not copy

© 2023 Snowflake Computing Inc. All Rights Reserved

ACCESS CONTROL

What can you do?

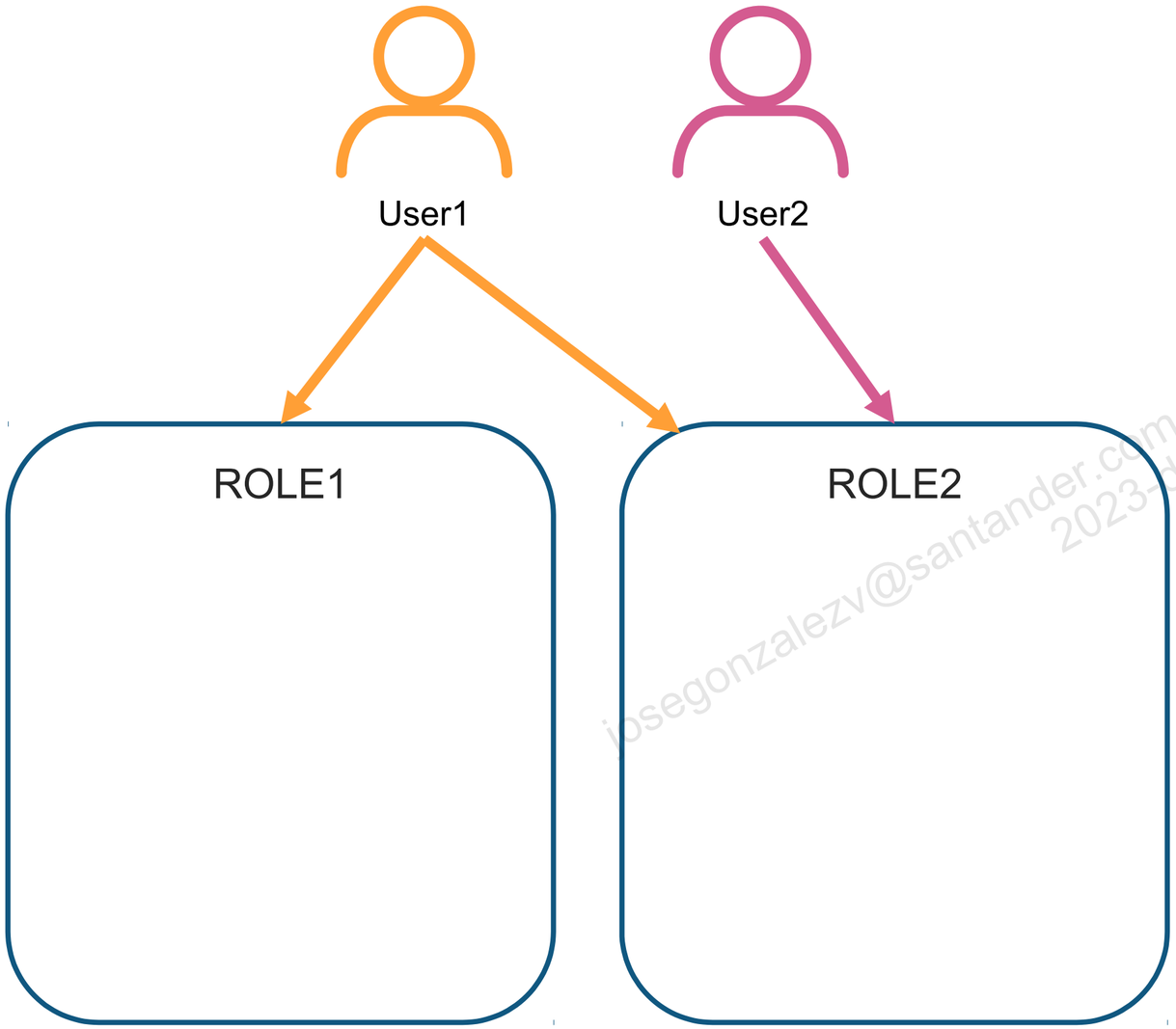


- Access Control is part of Snowflake's authorization model
- Access Control defines:
  - Who can use which roles
  - Which role can access which objects
  - What operations can be performed on those objects
  - Which role can create or alter access control policies

josegonzalezv@santander.com.mx - 11-SEP-2025 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

USERS AND ROLES

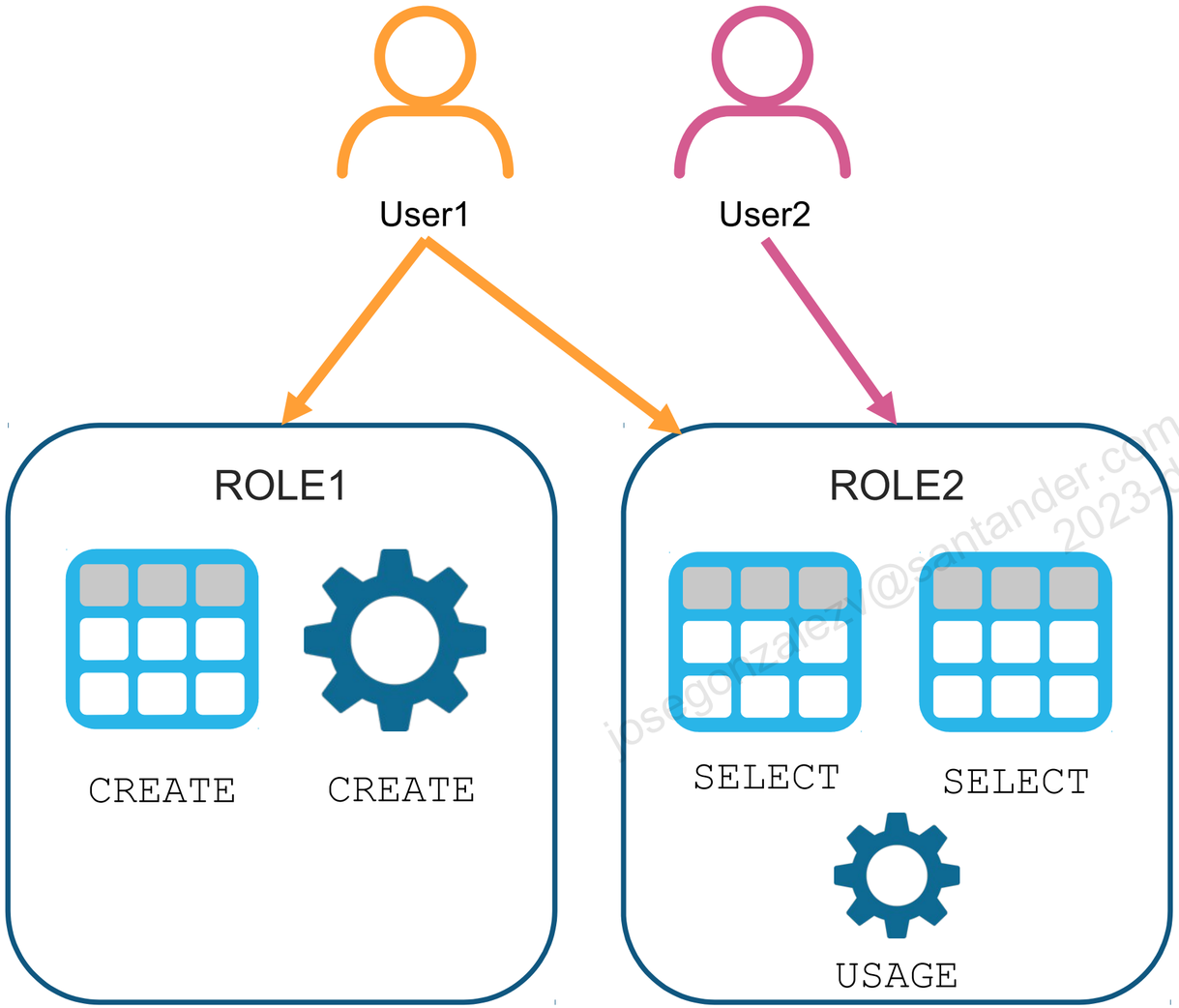


- **Users** are granted access to one or more roles by the `USERADMIN` or the `SECURITYADMIN`

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

USERS AND ROLES



- **Users** are granted access to one or more roles by the `USERADMIN` or the `SECURITYADMIN`
- Privileges on objects are granted to **roles** by the `SYSADMIN`
  - Create table
  - Use warehouse
  - Select from view
- Privileges are not granted directly to users
  - Except for the right to use a role

jose.gonzalez@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

USING ROLES

- You will be given access to one or more roles
- The **role** you are using determines what data you can see, and what you can do with it
- Privileges granted to roles allow users with those roles to do specific things
  - `GRANT USAGE ON WAREHOUSE elt_wh TO ROLE elt;`
  - `GRANT CREATE DATABASE ON ACCOUNT TO ROLE object_mgr;`
  - `GRANT SELECT ON ALL TABLES IN DATABASE main TO ROLE main_analyst;`

josegonzalezv@santander.com.mx - 14-SEP-2023 ©Snowflake
2023 do-not-copy

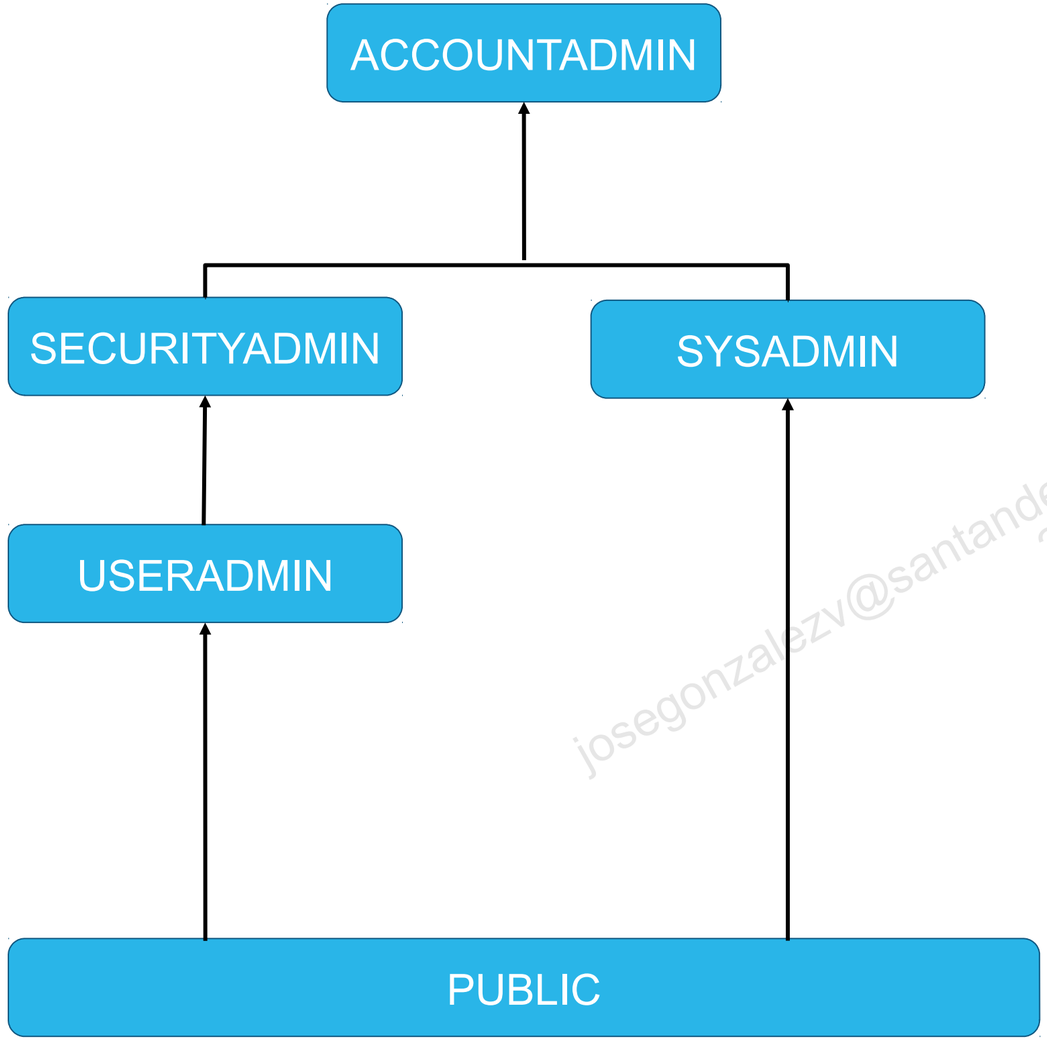
© 2023 Snowflake Computing Inc. All Rights Reserved

TYPES OF ROLES

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

SYSTEM ROLES

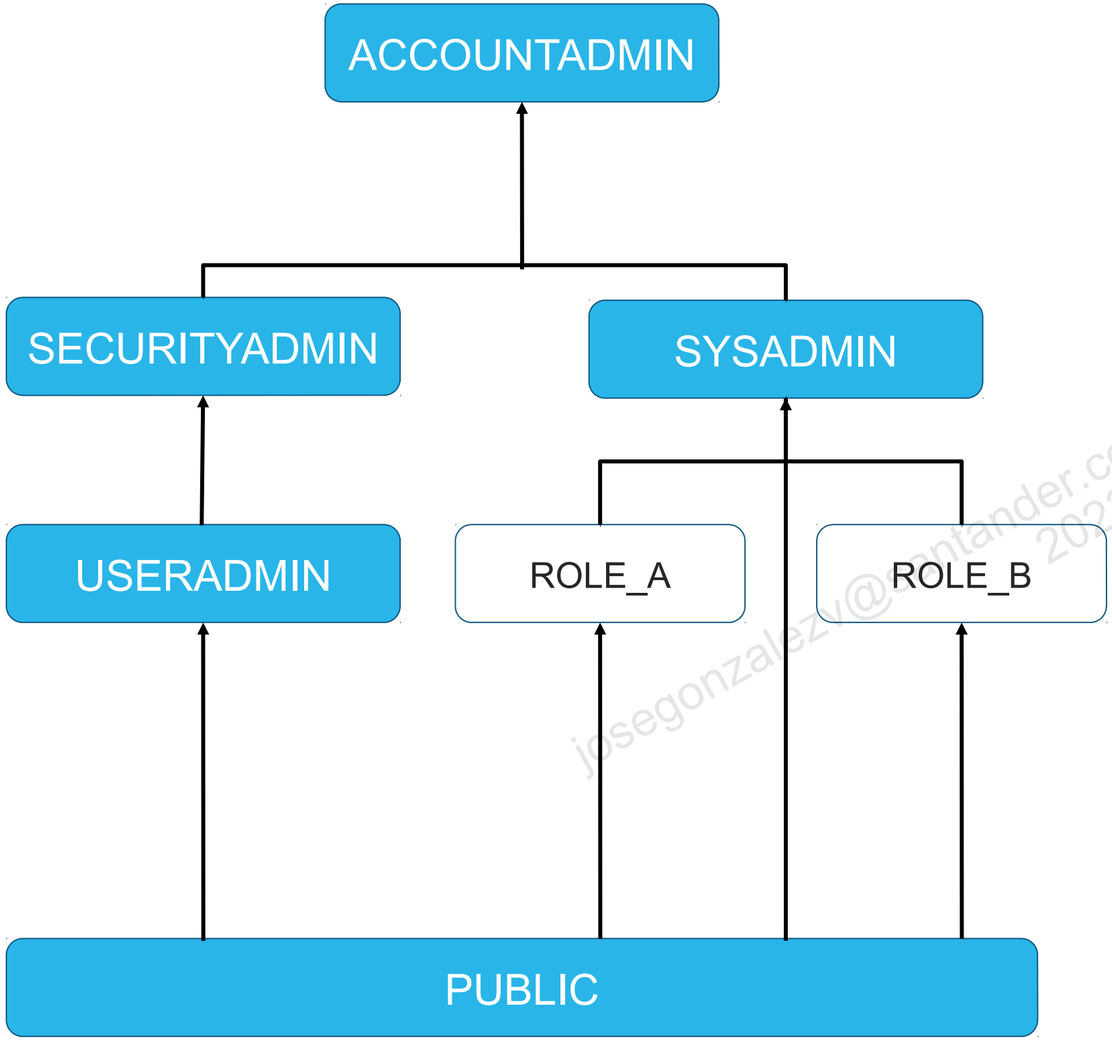


- `ACCOUNTADMIN` performs administrative functions
- `SECURITYADMIN` creates and manages users and roles, and manages grants
- `USERADMIN` creates users and roles
- `SYSADMIN` creates and manages databases, schemas, and warehouses
- `PUBLIC` is available to all users, but (by default) cannot do anything but log in

josegonzalezv@santander.com.mx - 11-SEP-2023 - @Snowflake
2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

CUSTOM ROLES



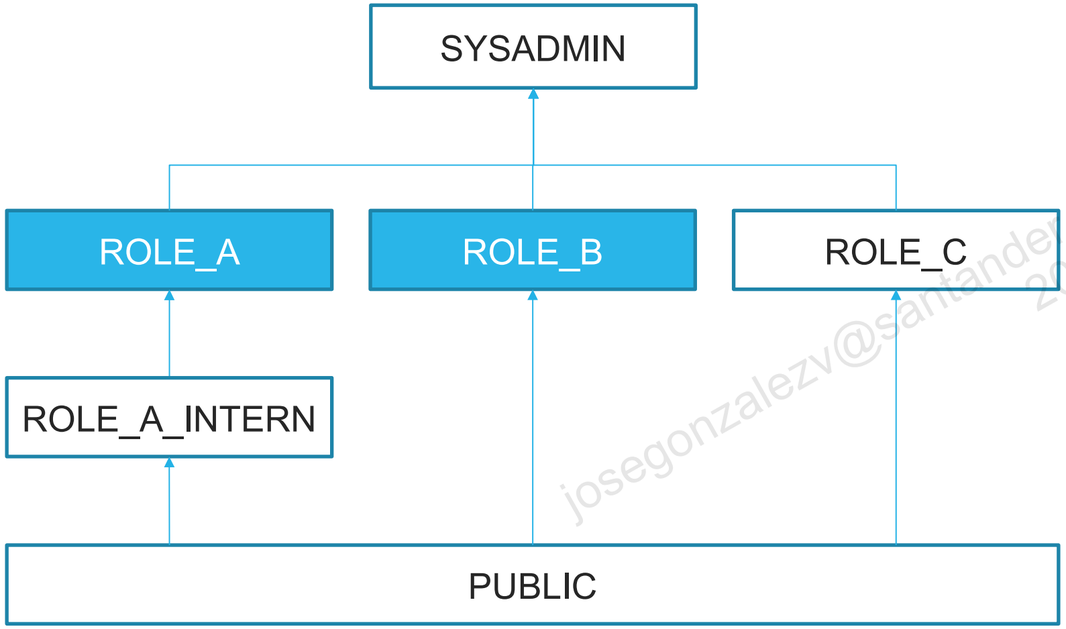
- Created by the `SECURITYADMIN` or `USERADMIN` based on a company's desired hierarchy
- Custom roles generally "roll up" to the `SYADMIN` role
- Users are assigned to one or more roles, which may be custom roles
- All privileges roll up the hierarchy

josegonzalezy@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

GRANTED ROLES

```
GRANT ROLE ROLE_A, ROLE_B
TO USER EMP123;
```



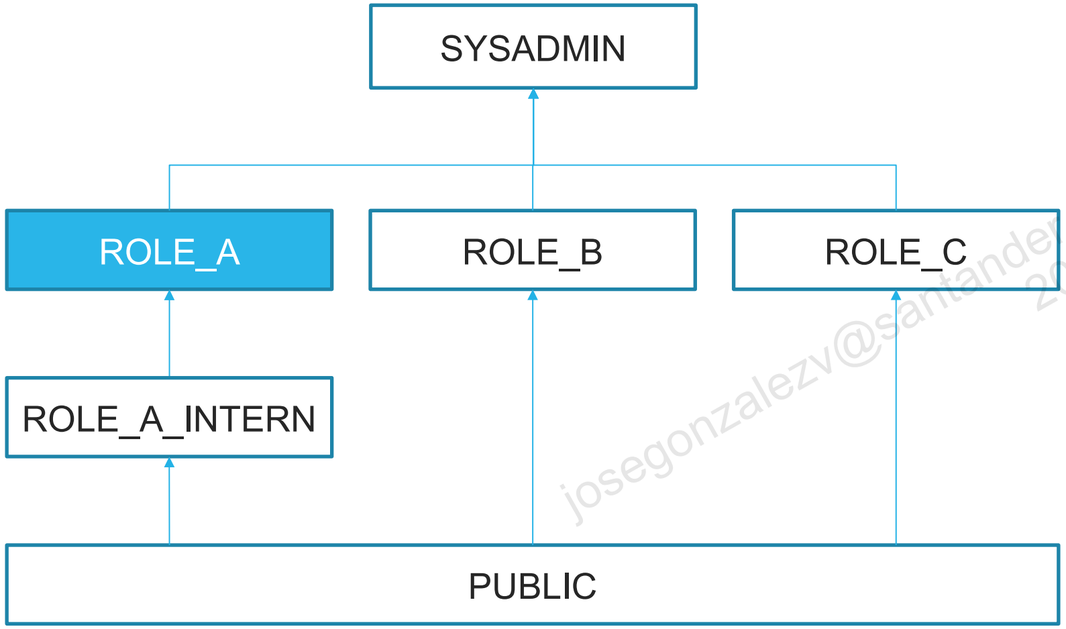
- Roles that are specifically granted to a user
- `EMP123` has been granted the roles `ROLE_A` and `ROLE_B`
  - They can decide which role to use for a session

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

CURRENT (PRIMARY) ROLE

```
USE ROLE ROLE_A;
```



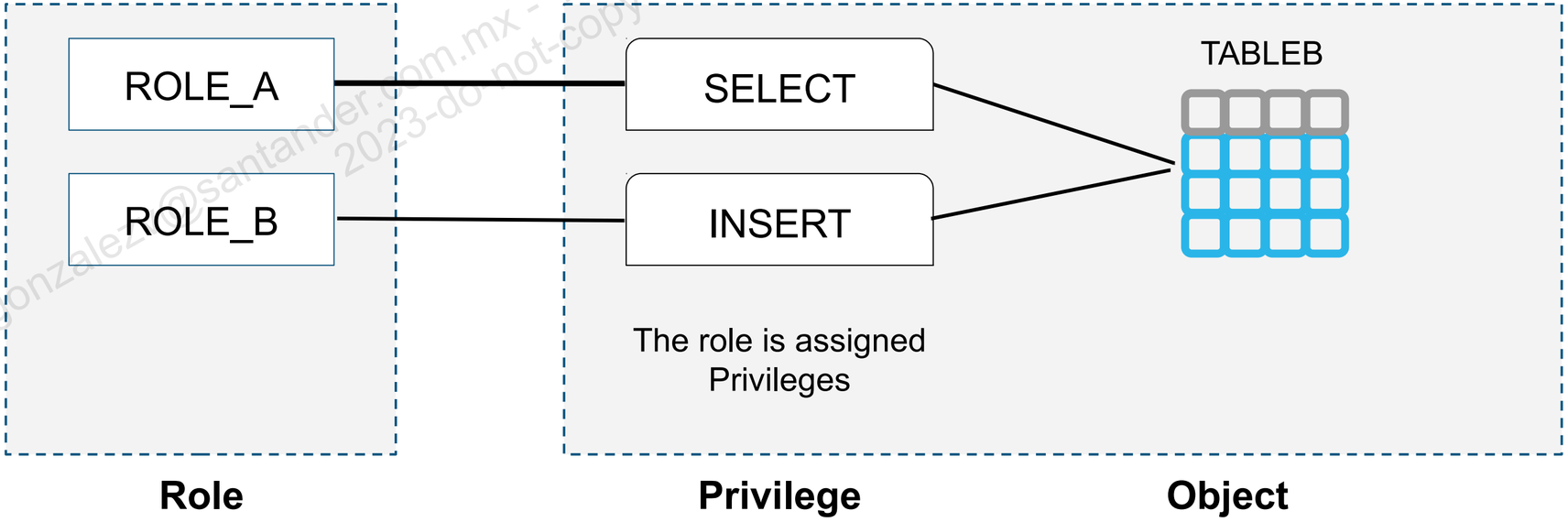
- The role a user is currently acting under (has chosen to `USE`)
- `EMP123` has been granted the `ROLE_A` and `ROLE_B` roles
  - Here, they have decided to use `ROLE_A`
- Each session can have a different primary role

josegonzalezv@santander.com.mx - 11-SEP-2023 - @Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

USERS, ROLES, AND GRANTS

- Privileges on objects are granted to roles

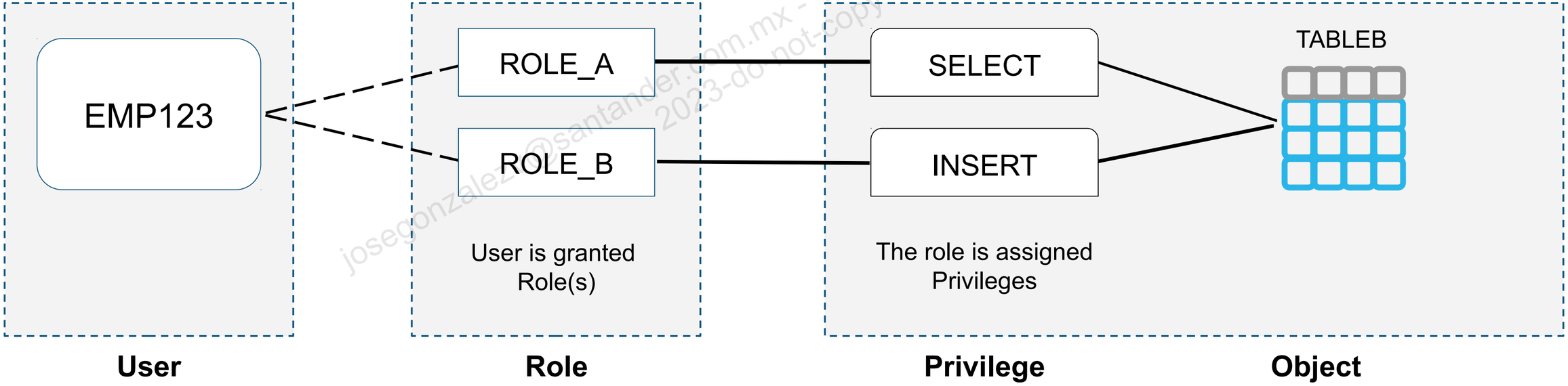


joseconzale2@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do not copy

© 2023 Snowflake Computing Inc. All Rights Reserved

USERS, ROLES, AND GRANTS

- Privileges on objects are granted to roles
- Users are granted roles which they can use



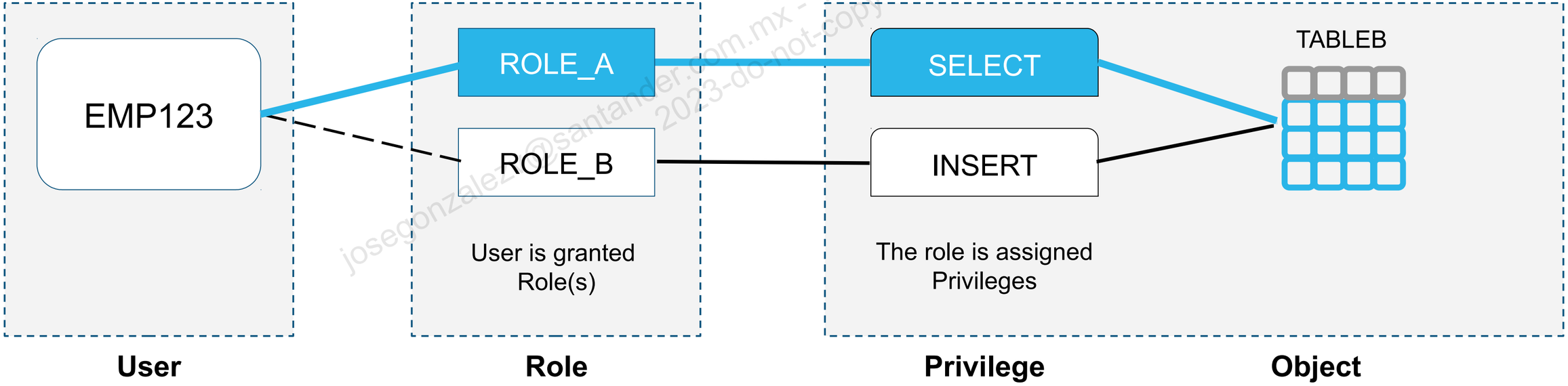
joseconzale2@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do not copy

© 2023 Snowflake Computing Inc. All Rights Reserved

USERS, ROLES, AND GRANTS

- Privileges on objects are granted to roles
- Users are granted roles which they can use
- Users `USE` a role, and get the role's privileges while using the role

joseconzale2@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

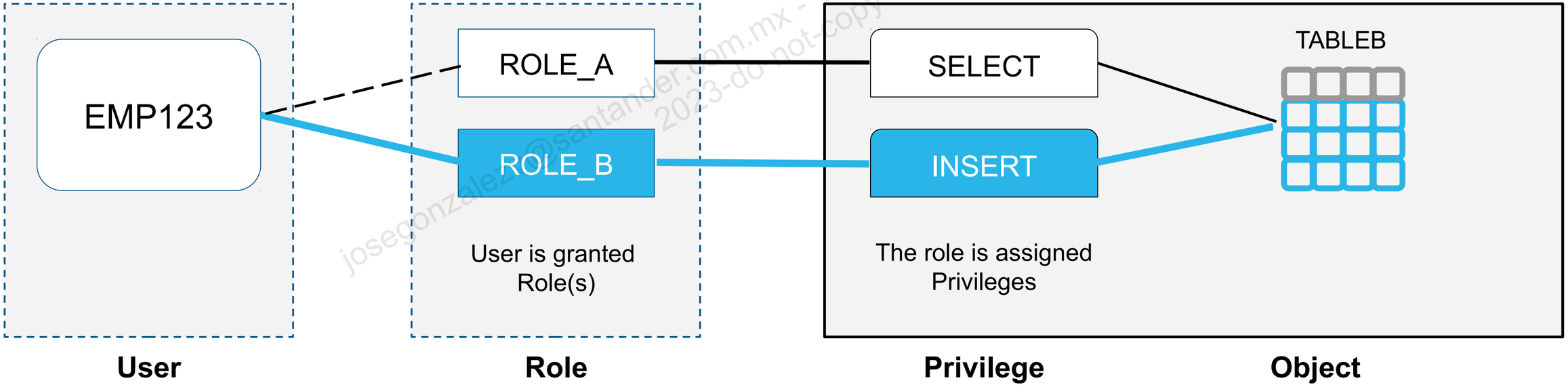


When `EMP123` uses **`ROLE_A`**, they can **`SELECT`** from `tableb`.

© 2023 Snowflake Computing Inc. All Rights Reserved

USERS, ROLES, AND GRANTS

- Privileges on objects are granted to roles
- Users are granted roles which they can use
- Users `USE` a role, and get the role's privileges while using the role
- Users can be assigned to multiple roles, but only use one (current role) at a time



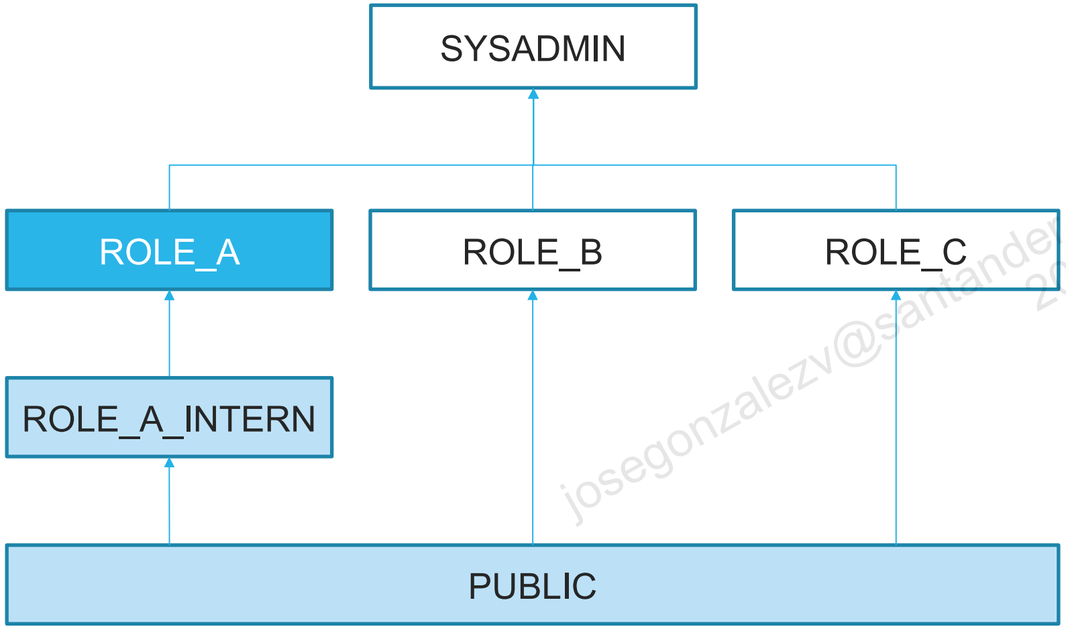
joseconzalez@santander.com.mx - 11-SEP-2023 @Snowflake 2023-do-not-copy

`EMP123` has privileges to **`INSERT`** into `tableb` while using **`ROLE_B`** (but they cannot `SELECT`)

© 2023 Snowflake Computing Inc. All Rights Reserved

INHERITED ROLES

```
USE ROLE ROLE_A;
```

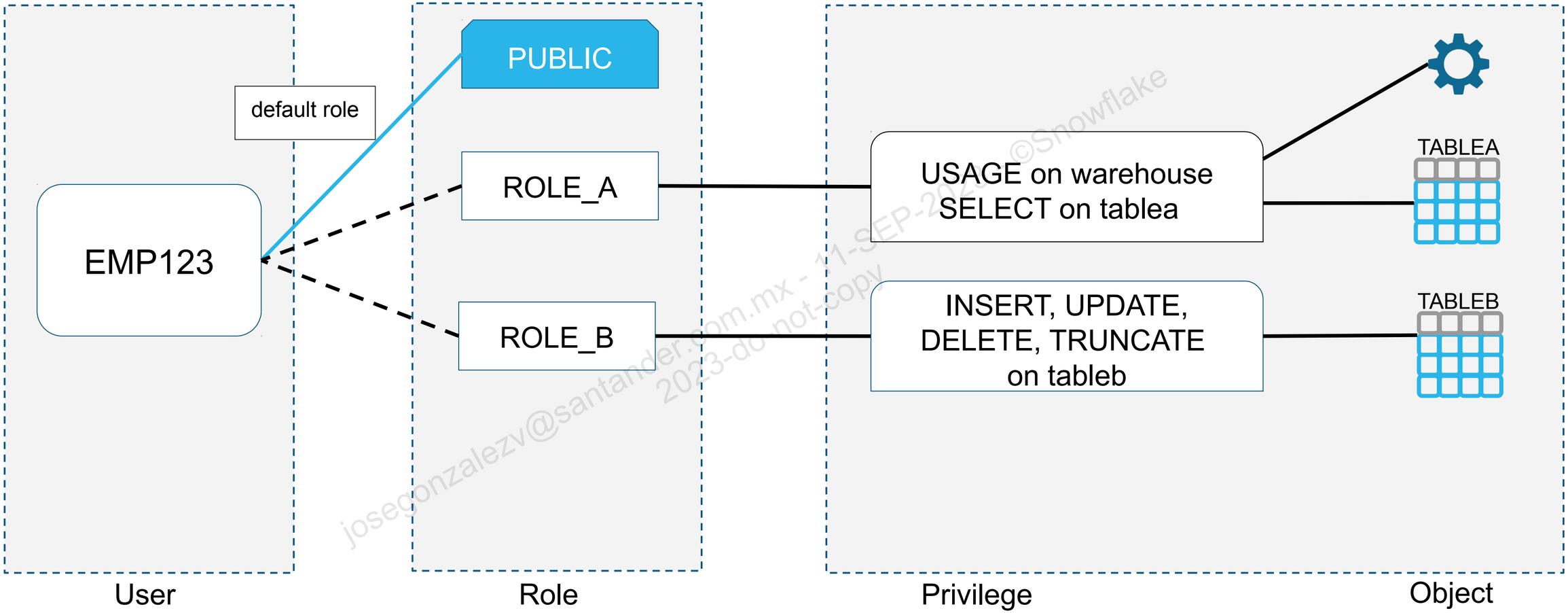


- Roles connected **below** the primary role in the hierarchy
- Privileges granted to these roles are also available to the primary role, through inheritance

josegonzalezv@santander.com.mx - 11-SEP-2023 @Snowflake
2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

ROLES AND INHERITANCE

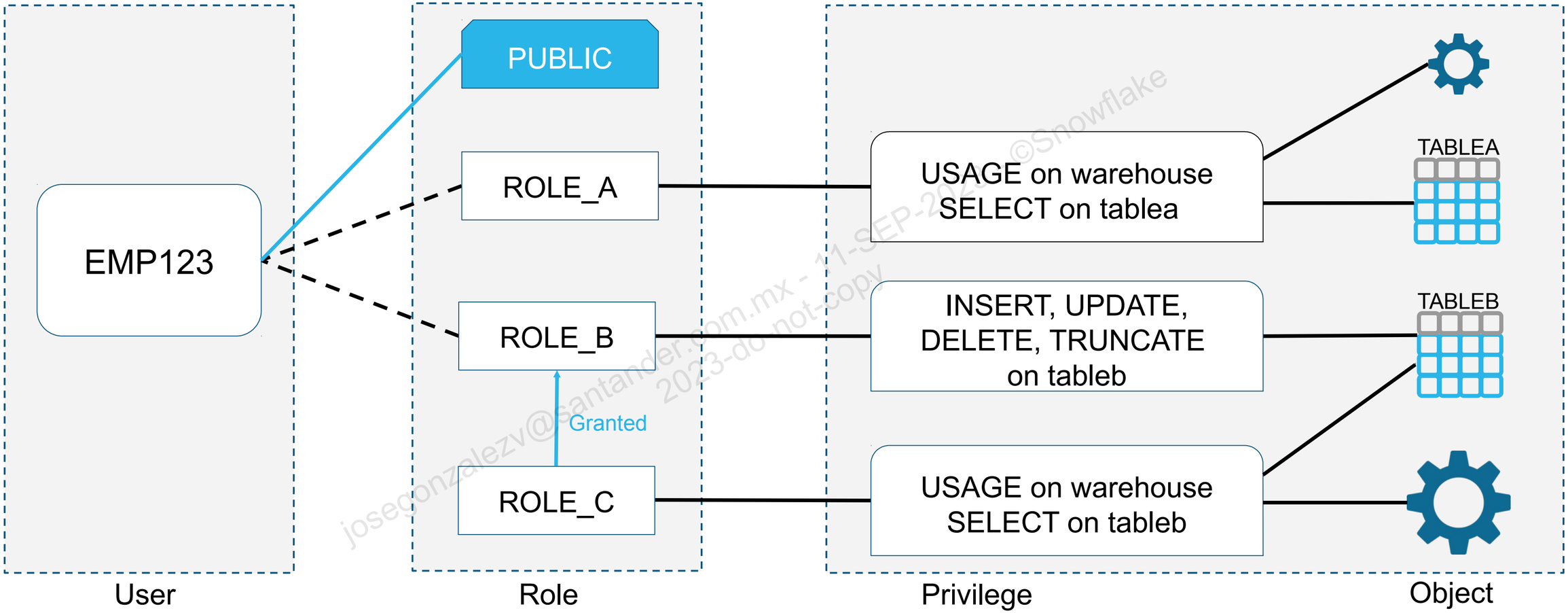


joseconzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do not copy

Upon connecting to Snowflake, a user takes on their default role

© 2023 Snowflake Computing Inc. All Rights Reserved

ROLES AND INHERITANCE

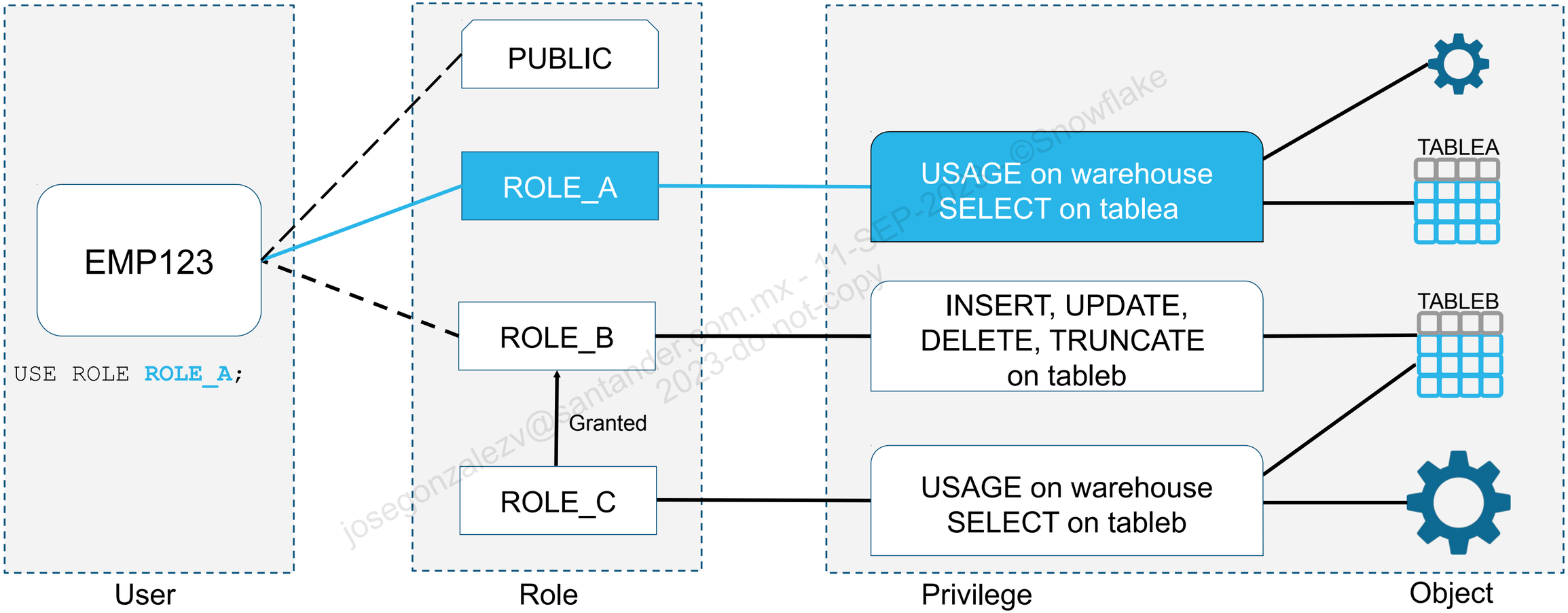


joseconzalezv@santander.com.mx - 11-SEP-2023 ©Snowflake
2023-do not copy

The role **ROLE_C** has been **granted** to the role **ROLE_B**

© 2023 Snowflake Computing Inc. All Rights Reserved

ROLES AND INHERITANCE

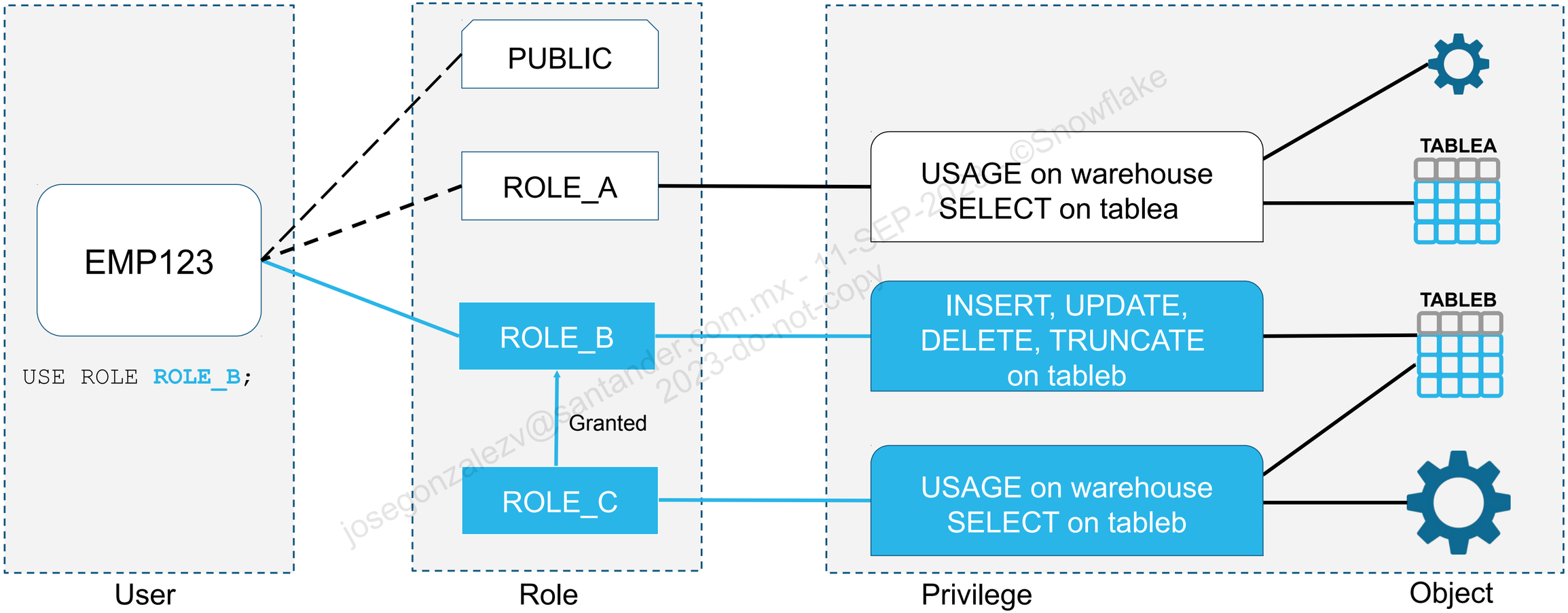


joseconzalezv@santander.com.mx - 11-SEP-2023 ©Snowflake
2023-do not copy

When using **ROLE_A**, EMP123 can **SELECT** from **TABLEA** and **USE** a **warehouse**.

© 2023 Snowflake Computing Inc. All Rights Reserved

ROLES AND INHERITANCE

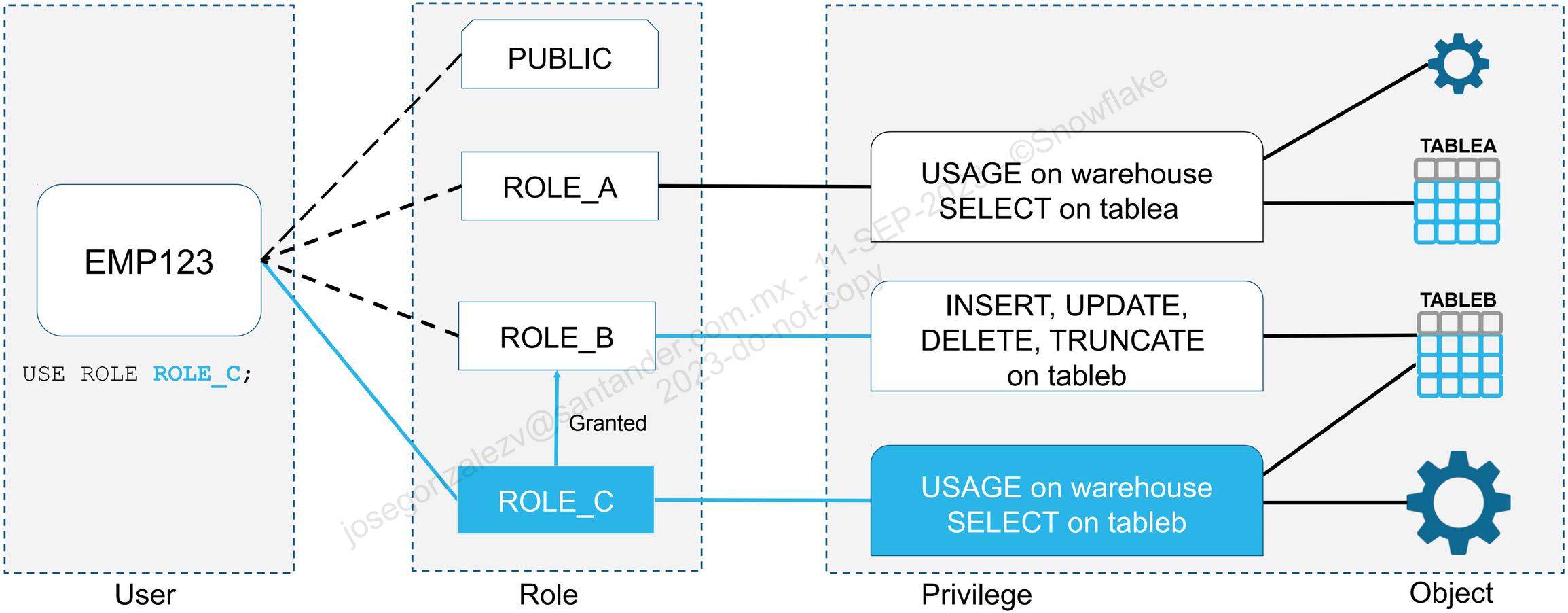


joseconzalezv@santander.com.mx - 11-SEP-2023 ©Snowflake
2023-do-not-copy

Using **ROLE_B**, they can **INSERT**, **UPDATE**, **DELETE**, and **TRUNCATE TABLEB**.
They also inherit the right to **SELECT** from **TABLEB** and **USE** a **warehouse**.

© 2023 Snowflake Computing Inc. All Rights Reserved

ROLES AND INHERITANCE



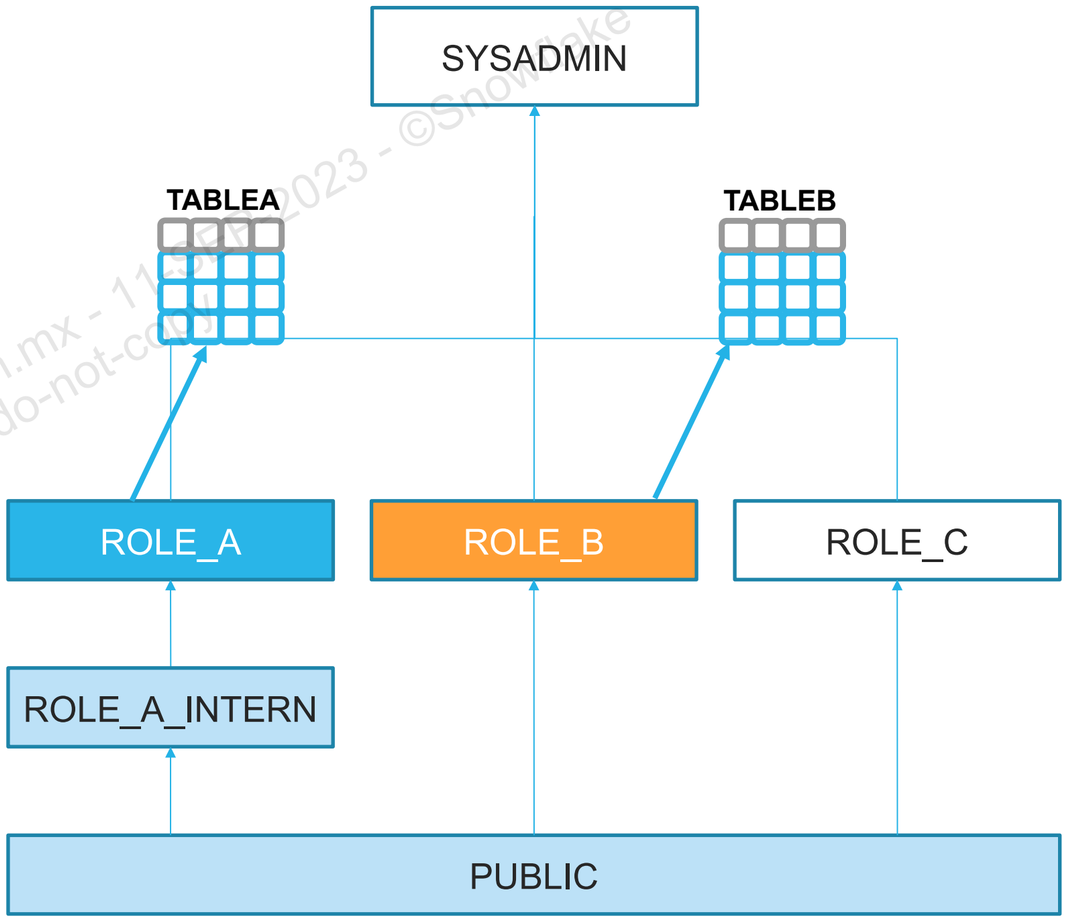
joseqonzalezv@santander.com.mx - 11-SEP-2023 ©Snowflake
2023-do-not-copy

They can also **USE ROLE_C**, since it was granted to a role they have the right to use.
As **ROLE_C**, they can only **SELECT** from **TABLEB** and **USE** a **warehouse**.

© 2023 Snowflake Computing Inc. All Rights Reserved

PROBLEM

- EMP123 has been granted the roles `ROLE_A` and `ROLE_B`
- `ROLE_A` role can `SELECT` from `tablea`, but not `tableb`
- `ROLE_B` role can `SELECT` from `tableb`, but not `tablea`
- `EMP123` needs to join `tablea` to `tableb`; how can they do this?

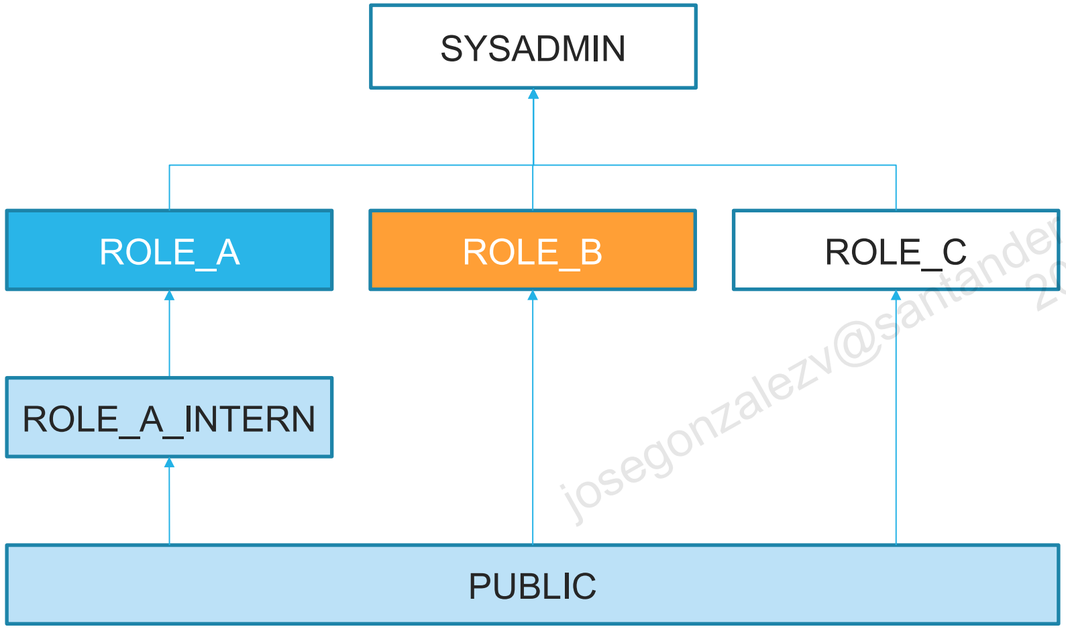


jose.gonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

SECONDARY ROLES

```
USE ROLE ROLE_A;
```



- Roles that have been granted to a user, but are not the primary role for the session

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

SOLUTION

- Secondary roles can be activated in a session, to give a user access to privileges across multiple roles they have been granted
- Any **create** privileges granted to secondary roles are NOT activated; only the create privileges granted to the **primary** role are available
- `USE SECONDARY ROLES { ALL | NONE }`

josegonzalezv@santander.com.mx - SEP 2023 - ©Snowflake 2023 do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

OWNERSHIP

josegonzalezv@santander.com mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

OWNERSHIP

- When you create an **object**, your **current role** becomes the **owner**
- Every role or object is owned by a single role
  - All users in that role share ownership when they are using that role
- The owner can do anything with the object
- The owner can grant privileges to itself or other roles

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

OWNERSHIP

- Warning: Object ownership may cause confusion
  - Be aware of which role you're using when you create an object
  - When you change roles, your object may no longer be accessible
  - Check with a user who has access to the role that created the object (this may include yourself)

- Example:

```
GRANT CREATE DATABASE ON ACCOUNT TO ROLE ROLE_A;
USE ROLE ROLE_A;
CREATE DATABASE my_db1;
CREATE SCHEMA my_db1.myschema; -- privilege acquired through ownership
```

josegonzalezv@santander.com.mx - 11-SEP-2025 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

OWNERSHIP VS. GRANTS

- **GRANTs** give a role the right to do something with an object or to use a role
- **OWNERSHIP** is granted to the role that creates an object
- The **OWNING ROLE** can do anything with an object it owns
  - Including **GRANT**ing access to other roles

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake
2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

VIEW GRANTS

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake
2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

ABOUT GRANTS

- Grants are used by the `SYSADMIN` to assign privileges to roles
- Privileges can be granted or revoked
- Examples:

```
GRANT USAGE ON WAREHOUSE finance_wh TO ROLE finance;
REVOKE USAGE ON WAREHOUSE finance_wh FROM ROLE finance;

GRANT SELECT ON ALL TABLES IN DATABASE prod TO ROLE prod;
GRANT USAGE on
REVOKE SELECT ON TABLE prod.public.sensitive FROM ROLE prod;
```

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

VIEW CURRENT PRIVILEGES

- View the current roles granted to a user:

```
SHOW GRANTS TO USER <user_name>;
```

- View the current set of privileges granted to a role:

```
SHOW GRANTS TO ROLE <role_name>;
```

- View which users a role has been granted to:

```
SHOW GRANTS OF ROLE <role_name>;
```

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

SHOW GRANTS ON OBJECTS

- Lists all the account-level (i.e., global) privileges that have been granted to roles:

```
SHOW GRANTS ON ACCOUNT;
```

- Lists all privileges that have been granted on an object:

```
SHOW GRANTS ON <object_type> <object_name>;
```

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

ADDITIONAL INFORMATION

- Security Administrators who will be designing and/or implementing a security plan for their account(s) can find more information in the user documentation, or in Snowflake's course for Administrators

josegonzalezv@santander.com.mx - 11-SEP-2023 - @Snowflake
2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

LAB EXERCISE: 19

Access Control and User Management

20 minutes

Learning Objectives:

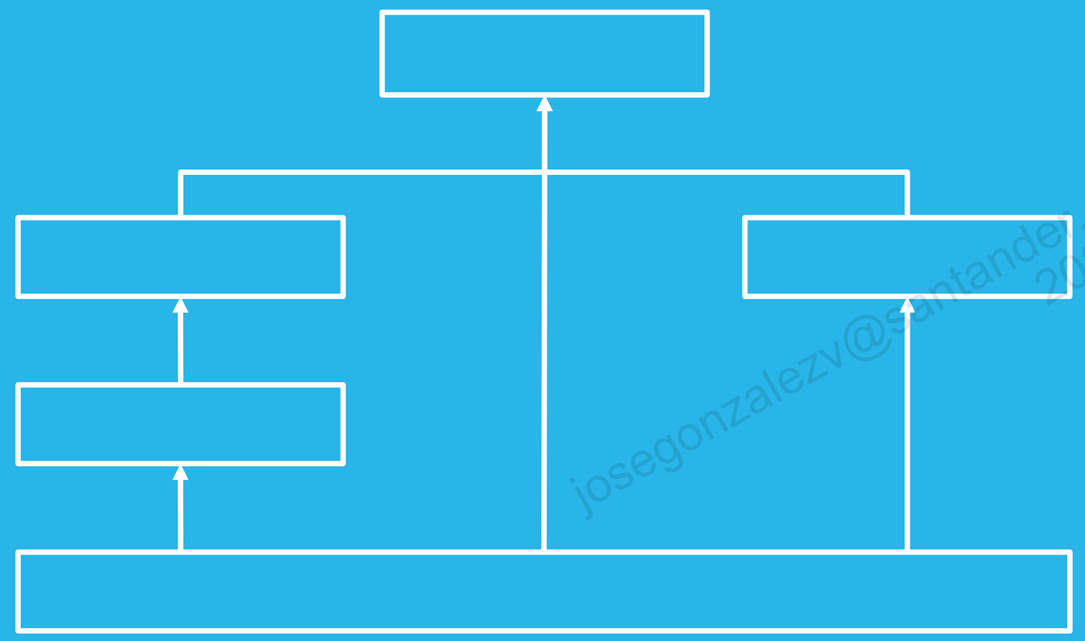
- How to show grants to users and roles
- How to grant usage on objects to roles
- How to use Secondary Roles to aggregate permissions from more than one role

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do not copy

© 2023 Snowflake Computing Inc. All Rights Reserved

KNOWLEDGE CHECK

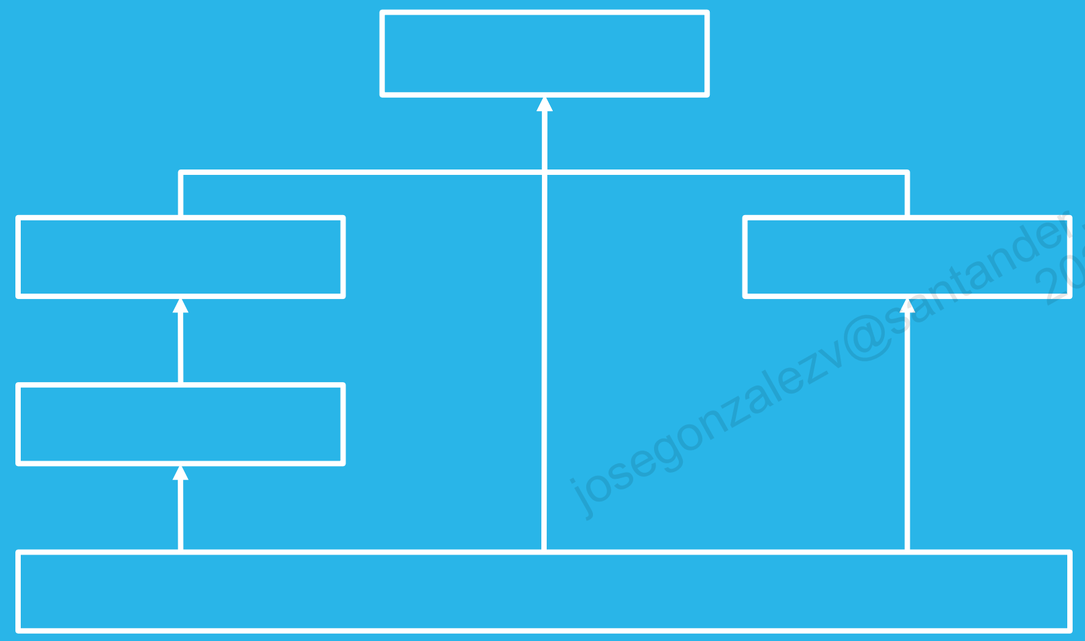
Fill out the following to identify the hierarchy of system roles:

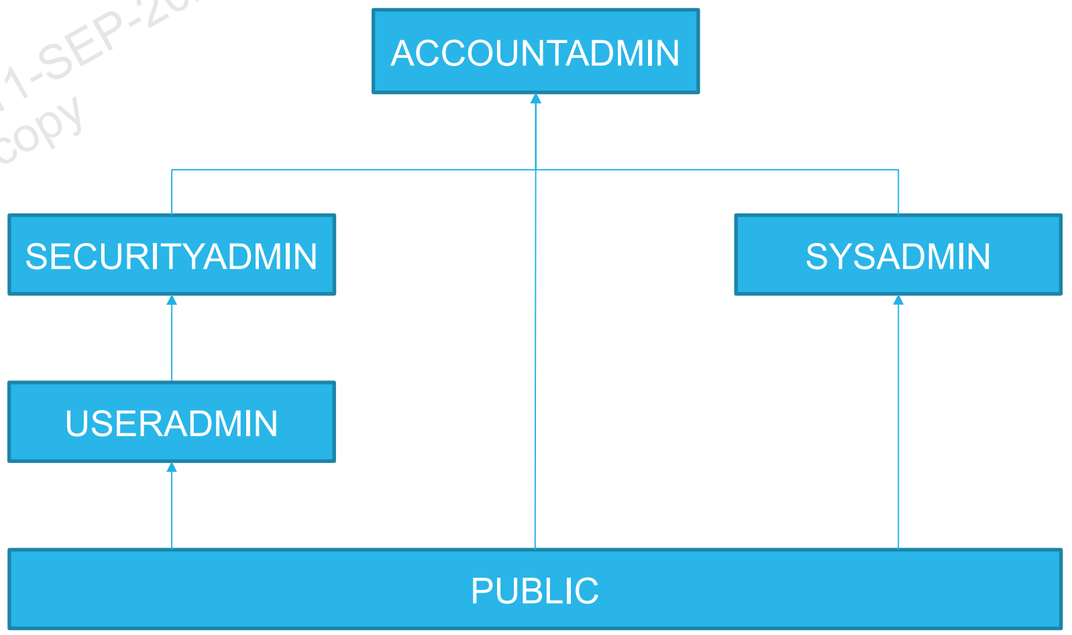


josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

KNOWLEDGE CHECK

Fill out the following to identify the hierarchy of system roles:





josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

KNOWLEDGE CHECK

Complete the following sentence?

Privileges are granted to _____.

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

KNOWLEDGE CHECK

Complete the following sentence?

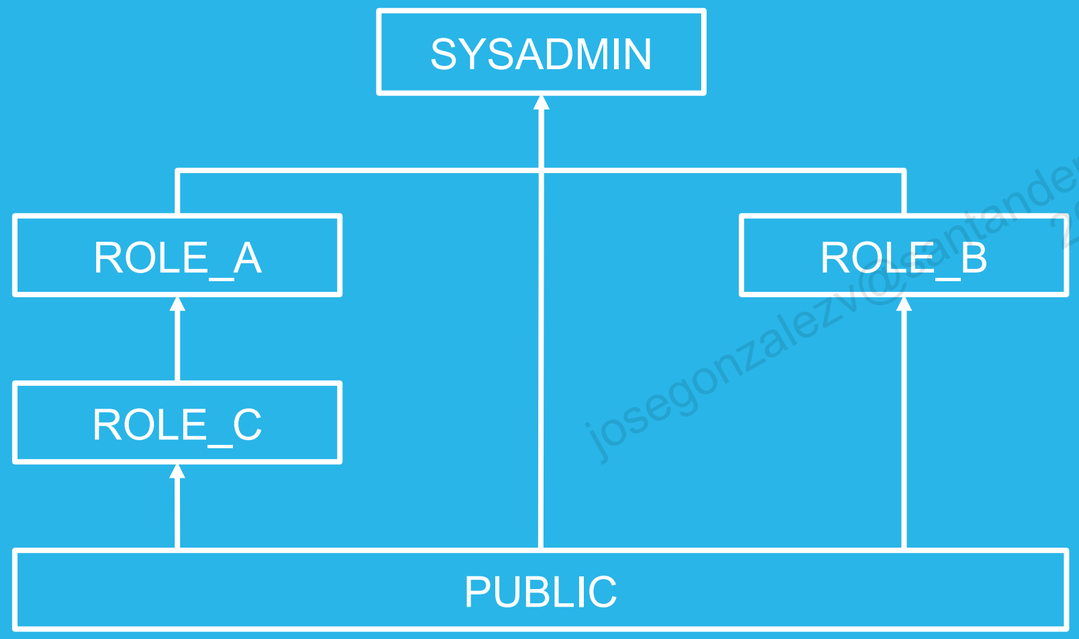
Privileges are granted to _____.

Privileges are granted to **roles**.

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

KNOWLEDGE CHECK

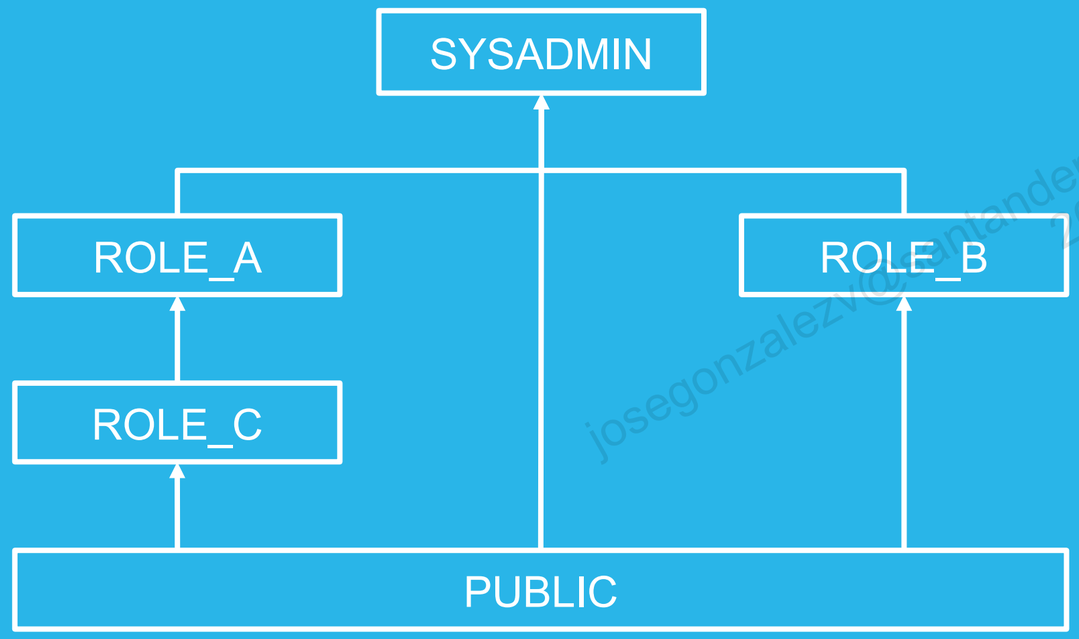
In the hierarchy shown below, `User1` has been granted `ROLE_A`. What other roles can they `USE`?



josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

KNOWLEDGE CHECK

In the hierarchy shown below, `User1` has been granted `ROLE_A`. What other roles can they `USE`?



`ROLE_C` and `PUBLIC`

When you are granted a role, you can `USE` that role, and any others that are connected below it in the hierarchy.

`User1` can `USE ROLE_A`, `ROLE_C`, or `PUBLIC`.

josegonzalezv@santander.com.mx - 11 SEP 2023 ©Snowflake 2023-do-not-copy

KNOWLEDGE CHECK

You log into Snowflake and run this command:

```
SELECT * FROM tablea;
```

Snowflake tells you that the object does not exist, or you are not authorized – but you know it exists because you created it.

What should you check?

josegonzalez@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

KNOWLEDGE CHECK

You log into Snowflake and run this command:

```
SELECT * FROM tablea;
```

Snowflake tells you that the object does not exist, or you are not authorized – but you know it exists because you created it.

What should you check?

- Make sure you are using a **role** that has access to the object. If you created the object under one role, and now you are using another, you may not see it.
- Check your **context** to see if you are using the correct database and schema for that object.
- Check for **typos**!

josegonzalez@santander.com.mx - 11-SEP-2023 @Snowflake
2023-do-not-copy

KNOWLEDGE CHECK

- `ROLE_A` and `ROLE_B` are not connected in the hierarchy
- `ROLE_A` can create tables in `PROD_DB.SCHEMA1`
- `ROLE_B` can create tables in `DEV_DB.SCHEMA1`

You have access to both roles. What objects can you create after this:

```
USE ROLE ROLE_A;
USE SECONDARY ROLES ALL;
```

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

KNOWLEDGE CHECK

- `ROLE_A` and `ROLE_B` are not connected in the hierarchy
- `ROLE_A` can create tables in `PROD_DB.SCHEMA1`
- `ROLE_B` can create tables in `DEV_DB.SCHEMA1`

You have access to both roles. What objects can you create after this:

```
USE ROLE ROLE_A;
USE SECONDARY ROLES ALL;
```

You will be able to create tables in **`PROD_DB.SCHEMA1`**.

When you enable secondary roles, you only get the create privileges of your primary role (the one you have chosen to `USE`).

josegonzalezv@santander.com.mx - 11 SEP 2023 - ©Snowflake 2023-do-not-copy

LAB EXERCISE: 20

Capstone Lab: Putting it All Together

60 minutes

Learning Objectives:

- How to create a table
- How to load data from a file in an internal named stage into a table
- How to create and call a user-defined function
- How to create and call a user-defined table function
- How to create a dashboard
- How to add tiles to a dashboard
- How to create a bar chart in a dashboard tile

Note: A solution is presented on the last page of this lab in the workbook.

jose.gonzalez@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023 do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

SEMI-STRUCTURED DATA

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

MODULE AGENDA

- Overview
- Query Semi-Structured Data

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

OVERVIEW

josegonzalezv@santander.com.br - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

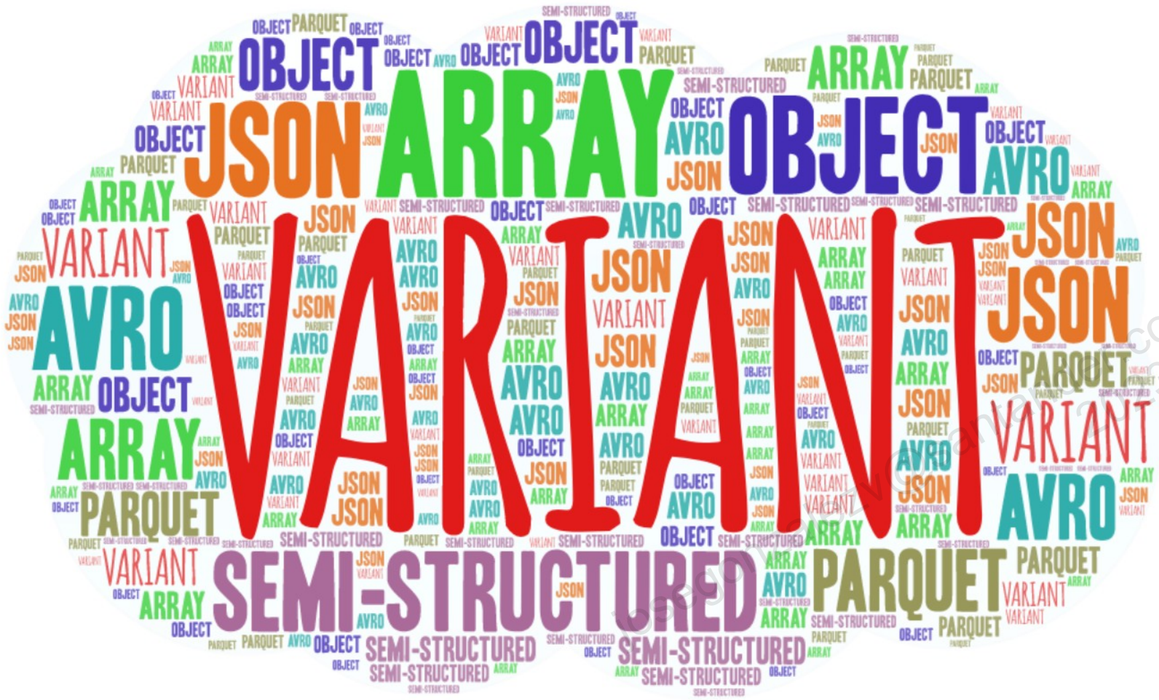
BUILT-IN SUPPORT

- Built-in support for all your data
  - Data Lakes with machine-generated data
  - Sensor IoT type data
  - Semi-Structured origin formats (JSON, Avro, ORC, Parquet, or XML)
- Statistics are gathered and maintained, including subfields for quicker query access and effective partition pruning for performance
- Support for all SQL operations (joins, group by, order by, …)
  - Built-in syntax for accessing data even in semi-structured fields
- Size of single record limited to 16 MB (compressed)

jcgonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake
2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

SEMI-STRUCTURED DATA TYPES

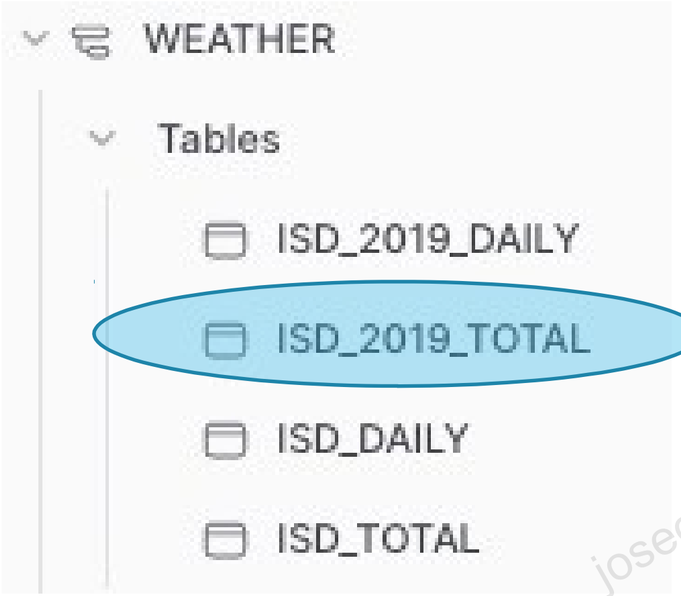


- VARIANT – holds values of standard SQL type, arrays, and objects
  - Used for raw semi-structured data
- OBJECT – a collection of key-value pairs
  - The key is a non-empty string
  - The value is a VARIANT
- ARRAY – Arrays of varying sizes
  - The index is a non-negative integer
  - The value is a VARIANT

© 2023 Snowflake Computing Inc. All Rights Reserved

VARIANT DATA TYPE EXAMPLE

SAMPLE WEATHER DATA SET IN TRAINING_DB



josegonzalezv@santander.com.mx - 11-SEP-2023 - Snowflake
2023-do-not-copy


```
{
  "data": {
    "observations": [
      {
        "air": {
          "dew-point": 10.6,
          "dew-point-quality-code": "1",
          "temp": 11.5,
          "temp-quality-code": "1"
        },
        "atmospheric": {
          "pressure": 10180,
          "pressure-quality-code": "1"
        },
        "dt": "2019-02-28T06:00:00",
        "sky": {
          "ceiling": 450,
          "ceiling-quality-code": "1"
        },
        "visibility": {
          "distance": 8000,
          "distance-quality-code": "1"
        },
        "wind": {
          "direction-angle": 140,
          "direction-quality-code": "1",
          "speed-quality-code": "1",
          "speed-rate": 10
        }
```


© 2023 Snowflake Computing Inc. All Rights Reserved

JSON DATA SAMPLE:

Key

Value

```
{
  "observations": [
    {
      "air": {
        "dew-point": 15.7,
        "dew-point-quality-code": "1",
        "temp": 18.1,
        "temp-quality-code": "1"
      },
      "atmospheric": {
        "pressure": 99999,
        "pressure-quality-code": "9"
      },
      "dt": "1984-04-11T00:00:00",
      "sky": {
        "ceiling": 9000,
        "ceiling-quality-code": "1"
      },
      "visibility": {
```

- Click on a semi-structured record ("row") to see its structure.
- **Simple key-value pair**: both key and value are scalar values.

jose.gonzalezv@santander.com.mx - 11-SEP-2023 - @Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

JSON DATA SAMPLE

Value

Key

```
{
  "observations": [
    {
      "air": {
        "dew-point": 15.7,
        "dew-point-quality-code": "1",
        "temp": 18.1,
        "temp-quality-code": "1"
      },
      "atmospheric": {
        "pressure": 99999,
        "pressure-quality-code": "9"
      },
      "dt": "1984-04-11T00:00:00",
      "sky": {
        "ceiling": 9000,
        "ceiling-quality-code": "1"
      },
      "visibility": {
```

- Click on a semi-structured record ("row") to see its structure.
- Simple key-value pair: both key and value are scalar values.
- **Objects**: Value consists of one or more simple key-value pairs enclosed in curly braces.

josegonzalezv@santander.com.mx - 11-SEP-2023 @Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

JSON DATA SAMPLE

Value

Key

```
{
  "observations": [
    {
      "air": {
        "dew-point": 15.7,
        "dew-point-quality-code": "1",
        "temp": 18.1,
        "temp-quality-code": "1"
      },
      "atmospheric": {
        "pressure": 99999,
        "pressure-quality-code": "9"
      },
      "dt": "1984-04-11T00:00:00",
      "sky": {
        "ceiling": 9000,
        "ceiling-quality-code": "1"
      },
      "visibility": {
```

- Click on a semi-structured record ("row") to see its structure.
- Simple key-value pair: both key and value are scalar values.
- **Objects**: Value consists of one or more simple key-value pairs enclosed in curly braces.
- Objects and arrays can be nested inside each other

jose.gonzalezv@santander.com.mx - 11-SEP-2023 @Snowflake
2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

SEMI-STRUCTURED FILE FORMATS

Other semi-structured file formats can be loaded into a VARIANT column and queried, using similar commands and functions as JSON data.

- Avro: Open-source framework originally developed for use with Apache Hadoop
- ORC (Optimized Row Columnar): binary format used to store Hive data
- Parquet: binary format designed for projects in the Hadoop ecosystem
- XML (Extensible Markup Language): consists primarily of tags <> and elements

jsegovia@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

QUERY
SEMI-STRUCTURED DATA

josegonzalezv@santander.com.mx - 14-SEP-2023 - ©Snowflake
2023-info-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

QUERY RICH HIERARCHICAL STRUCTURES IN SQL


```
{[ V
{
  "data": {
    "observations": [
      {
        "air": {
          "dew-point": 19.2,
          "dew-point-quality-code": "1",
          "temp": 19.7,
          "temp-quality-code": "1"
        },
        "atmospheric": {
          "pressure": 99999,
          "pressure-quality-code": "9"
        },
        "dt": "2019-08-23T06:00:00",
        "sky": {
          "ceiling": 780,
          "ceiling-quality-code": "1"
        },
        "visibility": {
          "distance": 6000,
          "distance-quality-code": "1"
        },
        "wind": {
          "direction-angle": 300,
          "direction-quality-code": "1",
          "speed-quality-code": "1",
          "speed-rate": 15
        }
      },
```


```
SELECT
   v:data.observations[0].air.temp AS temp,
   v:data.observations[0].air."dew-point"
     AS dew_point,
   v:data.observations[0].wind."speed-rate"
     AS wind_speed
FROM
   weather.isd_2019_total
LIMIT 10;
```

	TEMP	DEW_POINT	WIND_SPEED
1	19.7	19.2	15
2	-17.1	-19.7	30
3	-5.4	-6.5	0
4	-1.5	-6.2	36
5	8.6	5.7	9999
6	23.7	21.8	9999
7	-3.2	-3.6	20
8	-5.7	-5.8	0
9	17.2	999.9	5
10	10.5	9.1	15

josegonzalezv@santander.com.mx - 11-SEP-2023 - @snowflake
2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

ACCESSING VALUES IN NESTED DATA

Access a value in a VARIANT column for a particular key

Dot Syntax

```
SELECT v:station.elev
  AS elevation
FROM
  isd_2019_total
…
```

```
ELEVATION
=========
    763
```

Bracket Syntax

```
SELECT v['station']['elev']
  AS elevation
FROM
  isd_2019_total
…
```

```
ELEVATION
=========
    763
```

josegonzalezv@santander.com.mx - 13-SEP-2023 - ©Snowflake
2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

CASTING DATA FROM VARIANTS

- VARIANTs are often just strings, arrays or numbers
- Without CASTing, they remain VARIANT object types
- Cast them to SQL data types using the **::** operator

```
SELECT
    v:station.elev::NUMBER(6,2)
    AS elevation
FROM
    weather.isd_2019_total;
```

	ELEVATION
1	7
2	182
3	513.3
4	99.7
5	115.8

josegonzalezv@santander.com.mx - 14-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

DATA FUNCTIONS

Array Creation and Manipulation

- ARRAY_AGG
- ARRAY_APPEND
- ARRAY_CAT
- ARRAY_COMPACT
- ARRAY_CONSTRUCT
- ARRAY_CONSTRUCT_COMPACT
- ARRAY_CONTAINS
- ARRAY_INSERT
- ARRAY_POSITION
- ARRAY_PREPEND
- ARRAY_SIZE
- ARRAY_SLICE
- ARRAY_TO_STRING
- ARRAYS_OVERLAP

Object Creation and Manipulation

- OBJECT_AGG
- OBJECT_CONSTRUCT
- OBJECT_DELETE
- OBJECT_INSERT

JSON and XML Parsing

- CHECK_JSON
- CHECK_XML
- PARSE_JSON
- PARSE_XML
- STRIP_NULL_VALUE

Extraction

- FLATTEN
- GET
- GET_PATH , :
- XMLGET

Conversion

- TO_ARRAY
- TO_JSON
- TO_OBJECT
- TO_VARIANT
- TO_XML

Casting

- AS_<object_type>
  - AS_ARRAY
  - AS_BINARY
  - AS_CHAR , AS_VARCHAR
  - AS_DATE
  - AS_DECIMAL , AS_NUMBER
  - AS_DOUBLE , AS_REAL
  - AS_INTEGER
  - AS_OBJECT
  - AS_TIME
  - AS_TIMESTAMP_*

Type Checking

- IS_<object_type>
  - IS_ARRAY
  - IS_BOOLEAN
  - IS_BINARY
  - IS_CHAR , IS_VARCHAR
  - IS_DATE , IS_DATE_VALUE
  - IS_DECIMAL
  - IS_DOUBLE , IS_REAL
  - IS_INTEGER
  - IS_NULL_VALUE
  - IS_OBJECT
  - IS_TIME
  - IS_TIMESTAMP_*

josegonzalezv@santander.com.mx - 11-SEP-2023 - @Snowflake
2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

FLATTENING DATA

- VARIANTs may contain nested elements (arrays and objects that contain the data)
- FLATTEN() extracts data from nested elements
- Often used with a LATERAL join (to refer to columns from other tables, views, or table functions)

```
LATERAL FLATTEN(input=> expression [ options ] )
```

- input => The expression or column that will be unseated into rows
  - The data must be of type VARIANT, OBJECT, or ARRAY

josegonzalezv@santander.com.mx - 14-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

COLUMNS RETURNED BY FLATTEN

SEQ	Unique sequence number associated with input record; not guaranteed to be gap-free or ordered any particular way
KEY	For maps or objects, this column contains the key to the exploded value
PATH	Path to the element within a data structure that needs to be flattened
INDEX	Index of the element, if it is an array; otherwise NULL
VALUE	Value of the element of the flattened array or object
THIS	The element being flattened (useful in recursive flattening)

jose.gonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

FLATTENING DATA

```
SELECT * FROM TABLE (FLATTEN(input => PARSE_JSON('{
   "species":"Canus familiaris",
   "name":"Izzy",
   "traits": ["Cute", "Loving", "Smart"]}'), RECURSIVE => TRUE));
```

SEQ	KEY	PATH	INDEX	VALUE	THIS
1	name	name	null	"Izzy"	{ "name": "Izzy", "species": "C
1	species	species	null	"Canus familiaris"	{ "name": "Izzy", "species": "C
1	traits	traits	null	["Cute", "Loving", "Smart"]	{ "name": "Izzy", "species": "C
1	null	traits[0]	0	"Cute"	["Cute", "Loving", "Smart"]
1	null	traits[1]	1	"Loving"	["Cute", "Loving", "Smart"]
1	null	traits[2]	2	"Smart"	["Cute", "Loving", "Smart"]

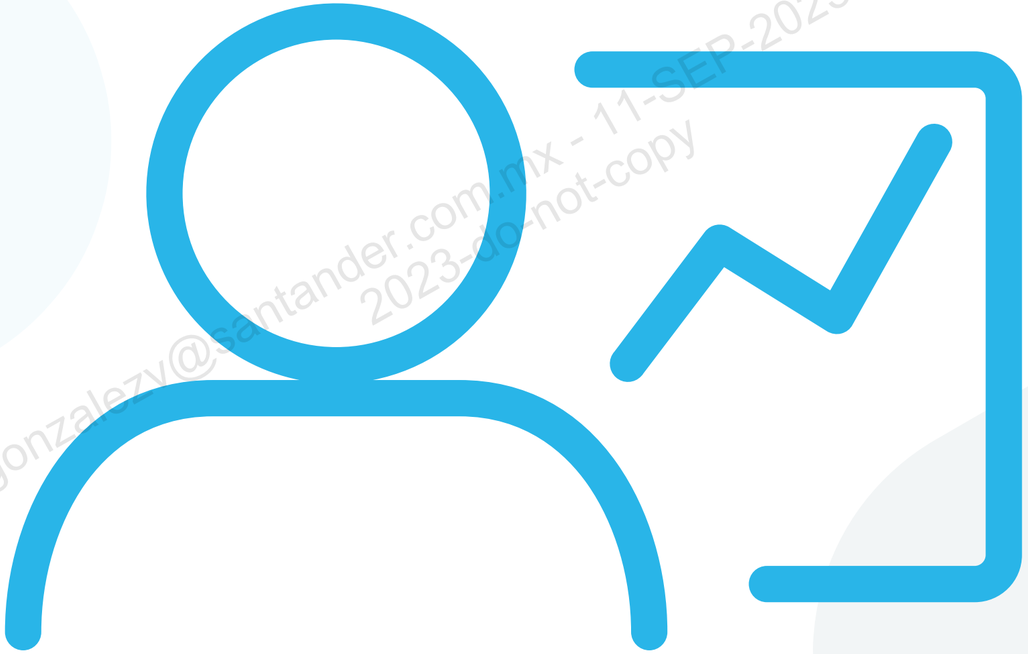
josegonzalezv@santander.com.mx - 11-SEP-2023 ©Snowflake
2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

INSTRUCTOR DEMO

View, Query, and Flatten Semi-Structured Data

10 minutes



josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

LAB EXERCISE: 21

Exploring Semi-Structured JSON Data

45 minutes

Learning Objectives:

- How to query a column containing semi-structured data without using the FLATTEN function.
- How to write a query that uses the FLATTEN table function to query semi-structured data.
- How to query a column containing semi-structured data flattened with the FLATTEN function.
- How to leverage the PATH and KEY output values of the FLATTEN function to get the data you want.
- How to leverage DESCRIBE RESULT to determine the data types of a query's columns.

josegonzalez@santander.com.mx - 1-SEP-2023 - ©Snowflake
2023 do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

LAB EXERCISE: 22

Challenge Lab: Querying Semi-Structured Data

50 minutes

Learning Objectives:

- Analyze semi-structured JSON data.
- Write queries that extract data from nested arrays and objects in semi-structured JSON data.

Note: A solution is presented on the last page of this lab in the workbook.

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

KNOWLEDGE CHECK

What are the six columns returned by the `FLATTEN` function?

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

KNOWLEDGE CHECK

What are the six columns returned by the `FLATTEN` function?

- SEQ
- KEY
- PATH
- INDEX
- VALUE
- THIS

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

KNOWLEDGE CHECK

What are the five semi-structured data formats supported by Snowflake?

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

KNOWLEDGE CHECK

What are the five semi-structured data formats supported by Snowflake?

- JSON
- Parquet
- Avro
- ORC
- XML

josegonzalezv@santander.com.mx - 11-SEP-2023 ©Snowflake
2023-do-not-copy

KNOWLEDGE CHECK

How is the `OBJECT` data type structured?

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

KNOWLEDGE CHECK

How is the `OBJECT` data type structured?

An object is a **collection of key-value pairs**.

Each value in the key-value pair can either be a simple value, another object, or an array.

josegonzalezv@santander.com.mx - 11 SEP 2023 - ©Snowflake 2023-do-not-copy

KNOWLEDGE CHECK

In the `SELECT` command below – what is the name of the column from which you are selecting data?

```
SELECT
    a:b.c
FROM
  my_table;
```

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

KNOWLEDGE CHECK

In the `SELECT` command below – what is the name of the column from which you are selecting data?

```
SELECT
    a:b.c
FROM
  my_table;
```

You are selecting from column `a`.

Dot syntax for querying semi-structured data has the form:

```
SELECT
    <column>:<key>.<key>
```

josegonzalezv@santander.com.mx - 11-SEP-2023 ©Snowflake 2023-do-not-copy

INTRODUCTION TO DATA SHARING

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake
2023 do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

MODULE AGENDA

- Snowflake Data Sharing Overview

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

SNOWFLAKE DATA SHARING OVERVIEW

josegonzalezv@santander.com.mx - 11-SEP-2023 ©Snowflake
2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

DATA SHARING

One Platform,
One Copy of Data,
Many Workloads

Secure & Governed
Access to All
Data

Virtually Unlimited
Performance
and Scale

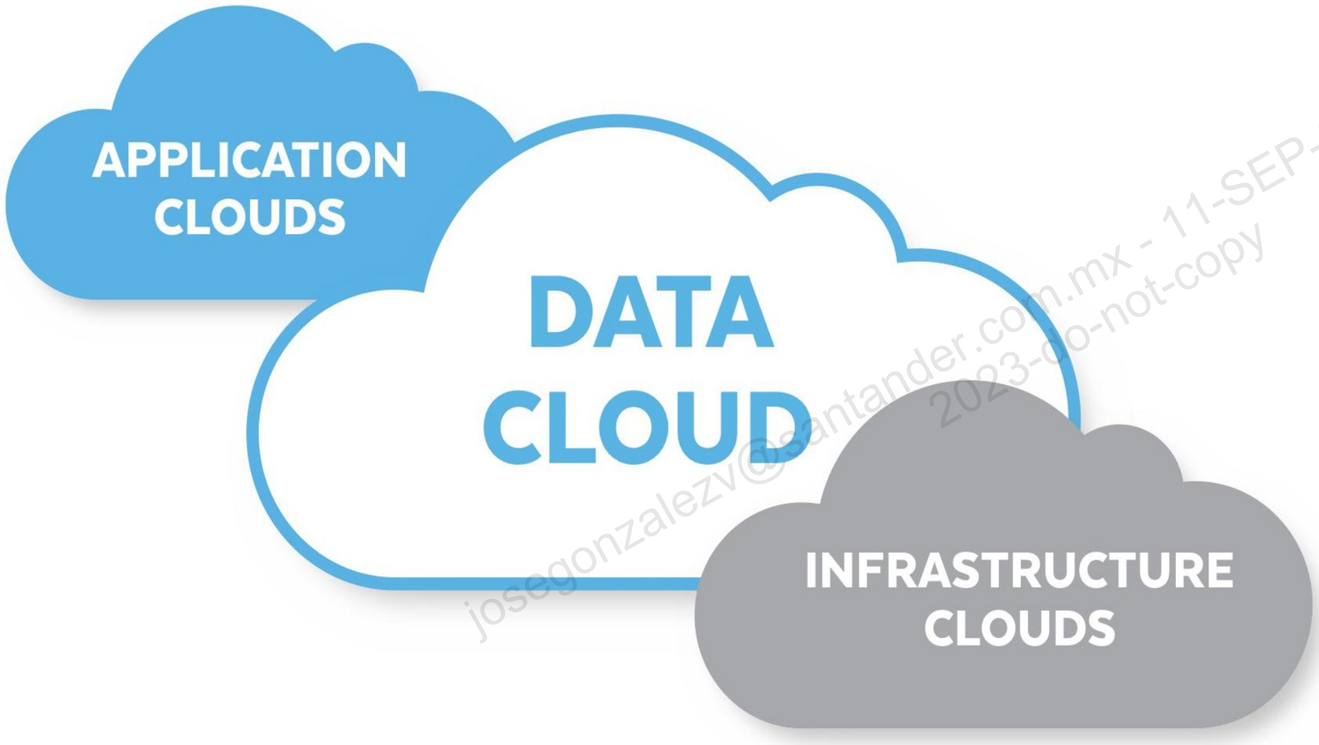
Near-Zero
Maintenance,
as a Service

© 2023 Snowflake Computing Inc. All Rights Reserved

THE DATA CLOUD: A GLOBAL ECOSYSTEM

© 2023 Snowflake Computing Inc. All Rights Reserved

WHAT IS THE DATA CLOUD?



josegonzalezv@santander.com.mx - 11-SEP-2023 ©Snowflake 2023-do-not-copy

THE SOLUTION:

The Data Cloud

WHAT IS IT?

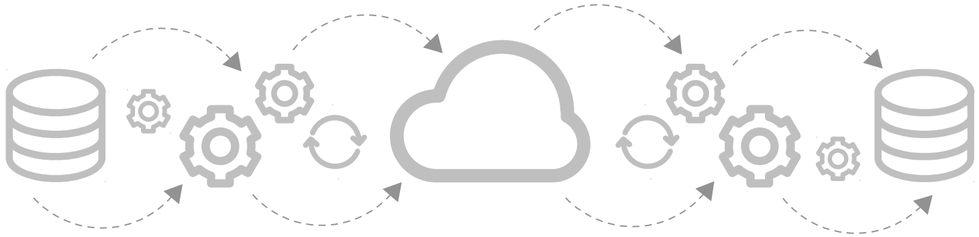
Snowflake Cloud Data Platform technology **plus the data itself**.

In the Data Cloud, every Snowflake customer can connect and share data in a completely governed and secure manner

© 2023 Snowflake Computing Inc. All Rights Reserved

TRADITIONAL WAYS OF SHARING DATA

Traditional Methods



- Copying files in FTP/ cloud buckets
- Building, Maintaining and Calling APIs
- ETL pipelines
- Data marts

Gaps

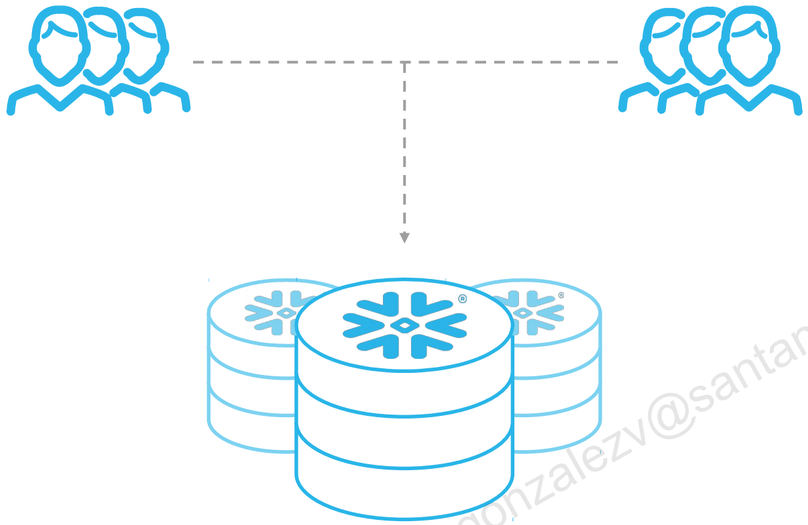
- ✗ Copy and move data
- ✗ Costly to maintain
- ✗ Data is delayed
- ✗ Error-prone
- ✗ Unsecure, once data is moved

josegonzalez@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

SECURE DIRECT DATA SHARING

Secure Data Sharing



Single live, copy of the data

Differentiators

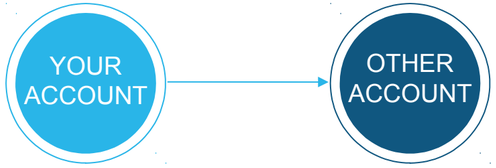
- No moving, no copying
- Live data, no delays
- Share personalized, secure views
- Governed, revocable access
- Globally available, cross-cloud

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

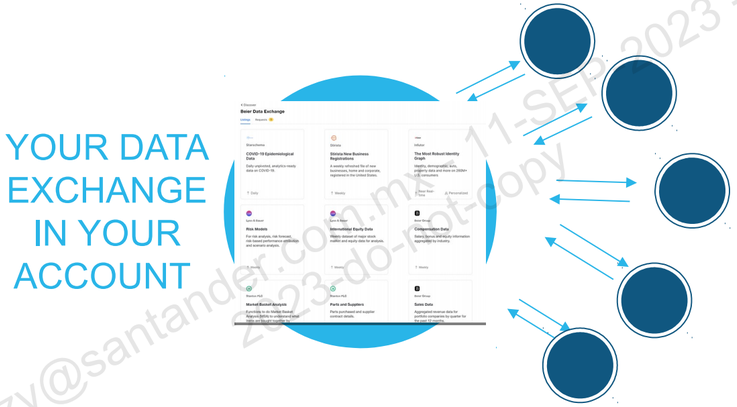
PRODUCT OFFERINGS

SECURE DIRECT DATA SHARING



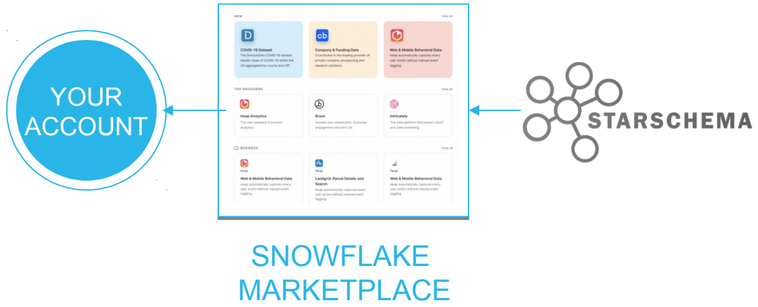
Direct Data sharing
1-to-1/many

SNOWFLAKE DATA EXCHANGE



Your Data *<u>privately published</u>*
to your consumers
<u>many</u> parties, <u>many</u> data sets

SNOWFLAKE MARKETPLACE

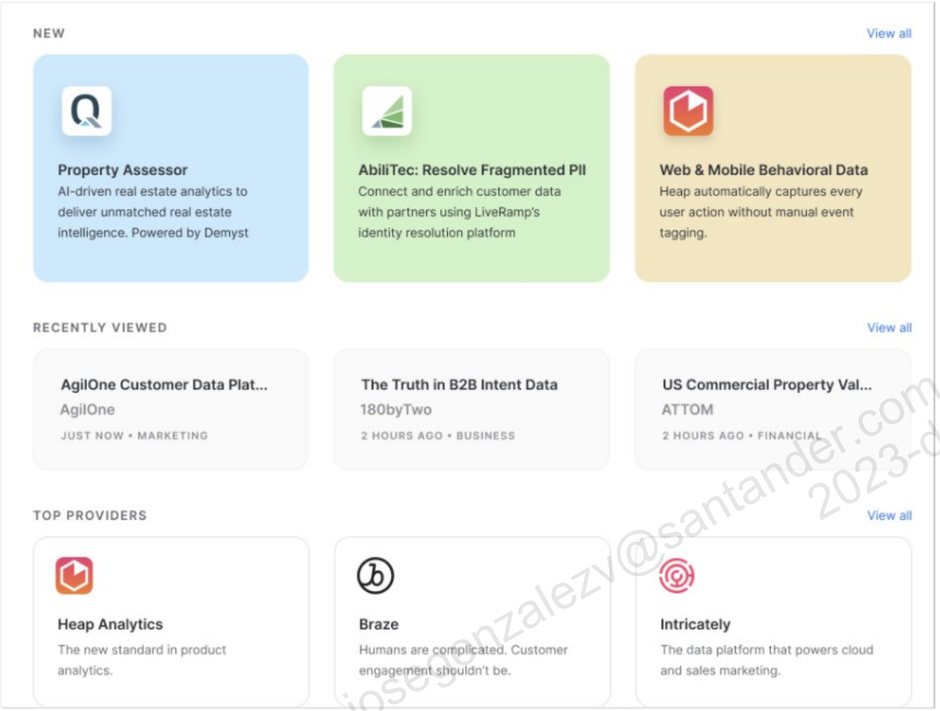


Publicly publish/consume data from 3rd parties:
Access data from hundreds of providers

josegonzalezv@santander.com.mx 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

SNOWFLAKE MARKETPLACE



- Live, ready-to-query data; no copying or moving
- Only data marketplace with personalized data
- Globally available, across clouds

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

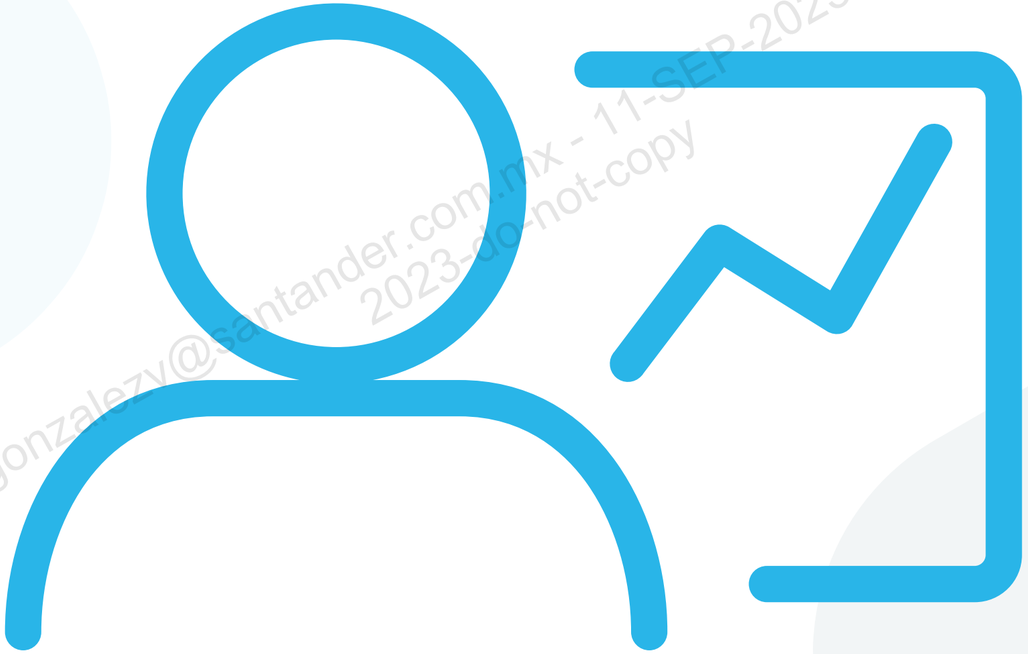
Financial | Marketing | Demographic | Macroeconomic | Government | Healthcare | Business

© 2023 Snowflake Computing Inc. All Rights Reserved

INSTRUCTOR DEMO

The Snowflake Marketplace and Data Sharing

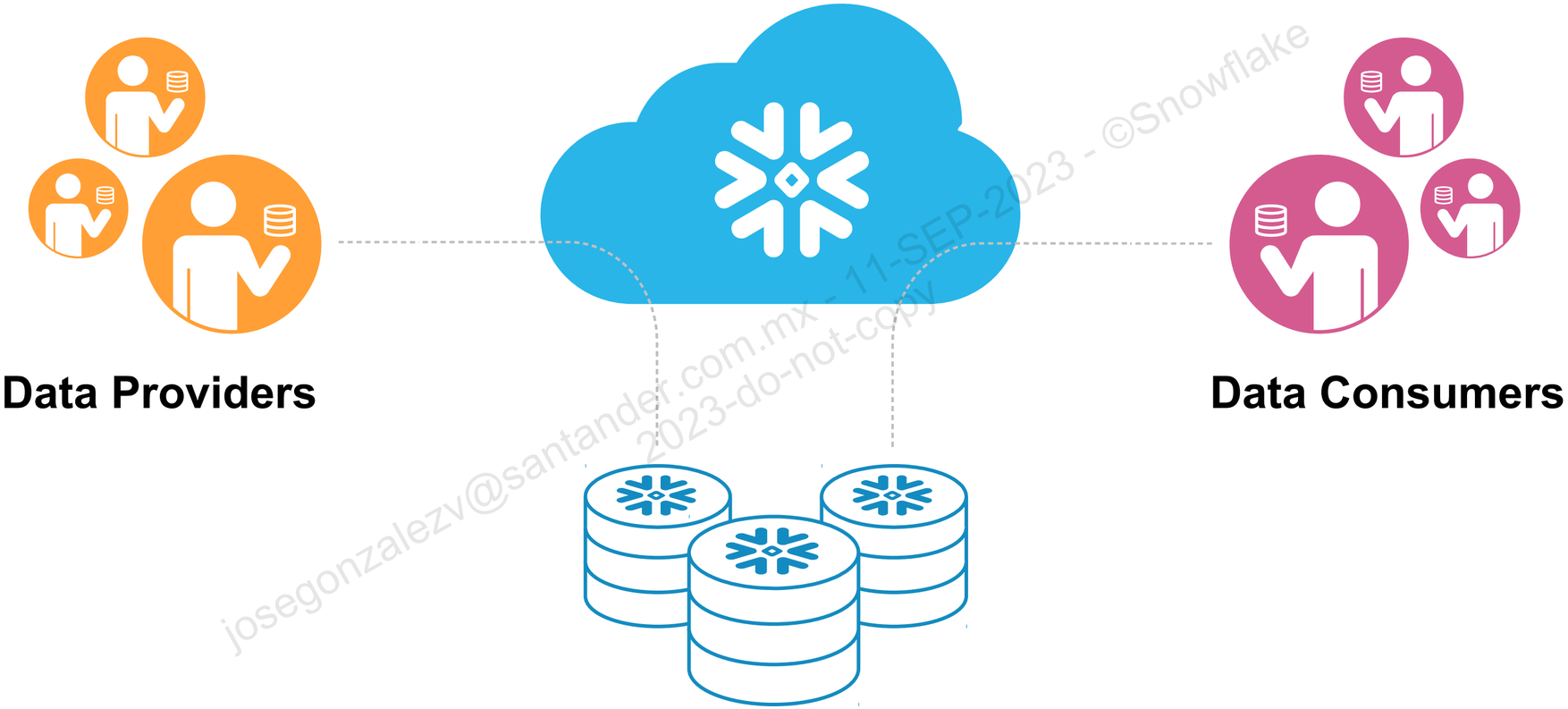
15 minutes



josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

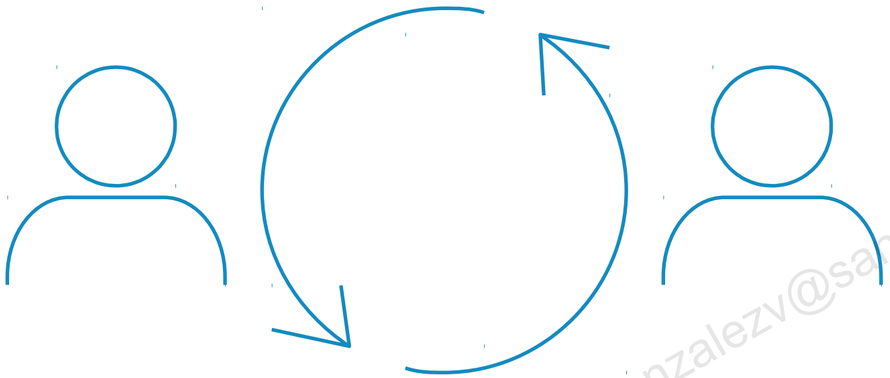
A BETTER WAY TO SHARE DATA



josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

INSTANT, LIVE DATA SHARING

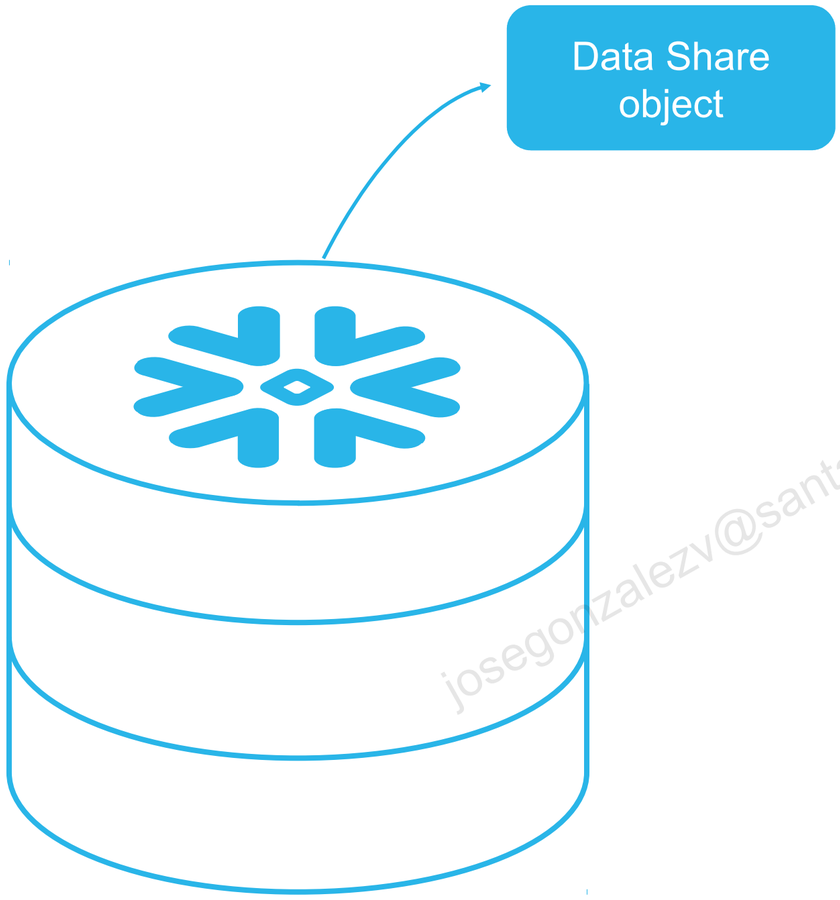


- No data movement; nothing is copied
- Data consumers immediately see all updates
- Share with an unlimited number of consumers

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

WHAT IS A SHARE?



- A named Snowflake object that encapsulates all the information required to share objects within a database
- Share Contains:
  - Privileges that grant access to the database, schema, or specific objects
  - Data Consumer or Reader account name
- Shares do not contain any data

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake
2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

ABOUT SHARES

- Shares are read-only
- Shares cannot be cloned or re-shared
- You can share tables(including external tables), secure views, secure materialized views, and secure UDFs
- Access to a share can be revoked at any time
- Shares are limited to a single region and cloud provider
  - Database replication can be used to share to another region or cloud provider

jose.gonzalezv@santander.com.mx - 14-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

PROVIDERS AND CONSUMERS

- Providers have data they are willing to share with others
  - Can create unlimited shares
  - Can share with unlimited consumers
  - Can create "reader accounts" to share data with others who have no Snowflake account

- Consumers want to access data shared by a provider
  - Can consume unlimited shares
  - Only the most current version of data is visible (consumers cannot use Time Travel on a share)

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023 - do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

ADDITIONAL INFORMATION

- For more information on data sharing:
  - Snowflake's Advanced instructor-led course
  - Snowflake's on-demand training course "*Sharing, marketplace, & exchanges workshop*"
  - Docs.snowflake.com

jgonzalezv@santander.com.mx - 11-SEP-2023 ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

VIRTUAL WAREHOUSE SCALING

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake
2023 do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

MODULE AGENDA

- Types of Warehouse Scaling
- Auto-Scaling Policies

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

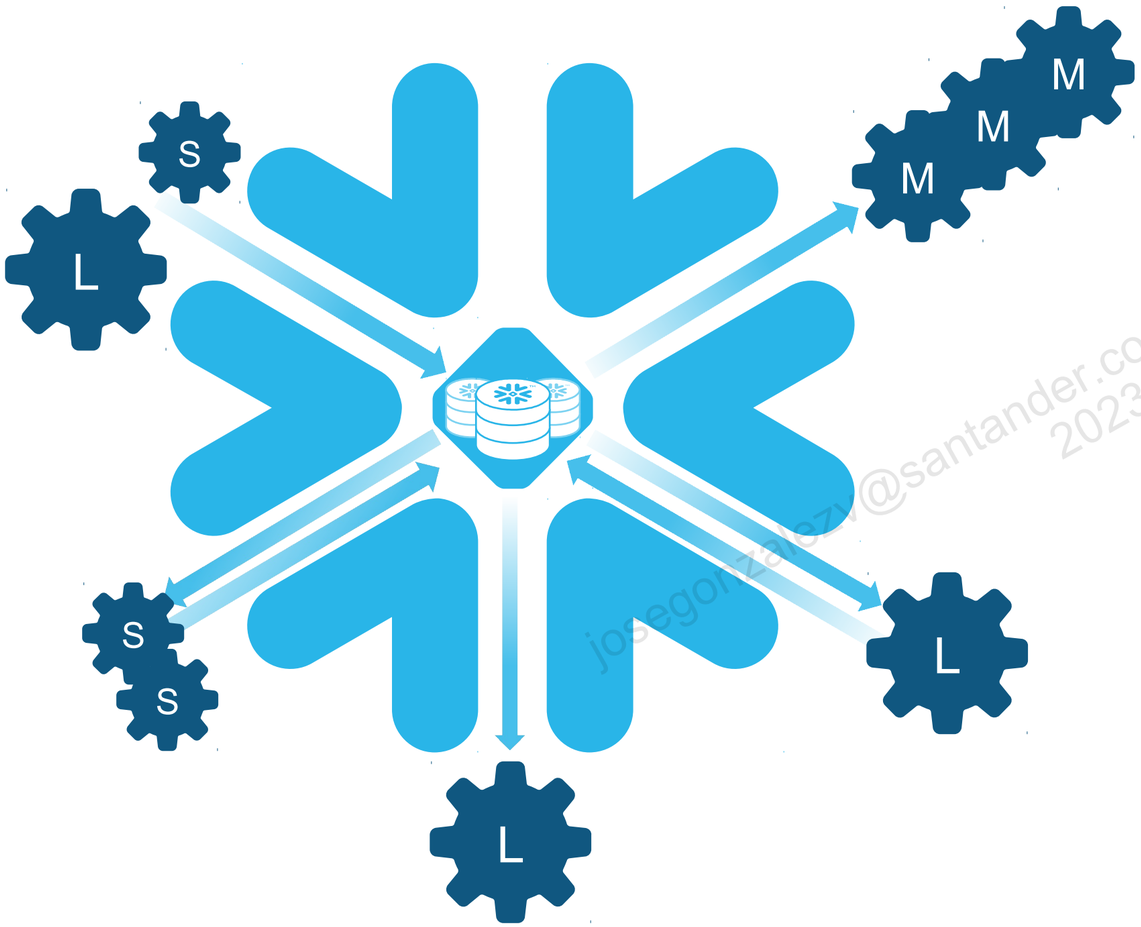
TYPES OF WAREHOUSE SCALING

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake
2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

SCALE ACROSS

ELIMINATE RESOURCE CONTENTION

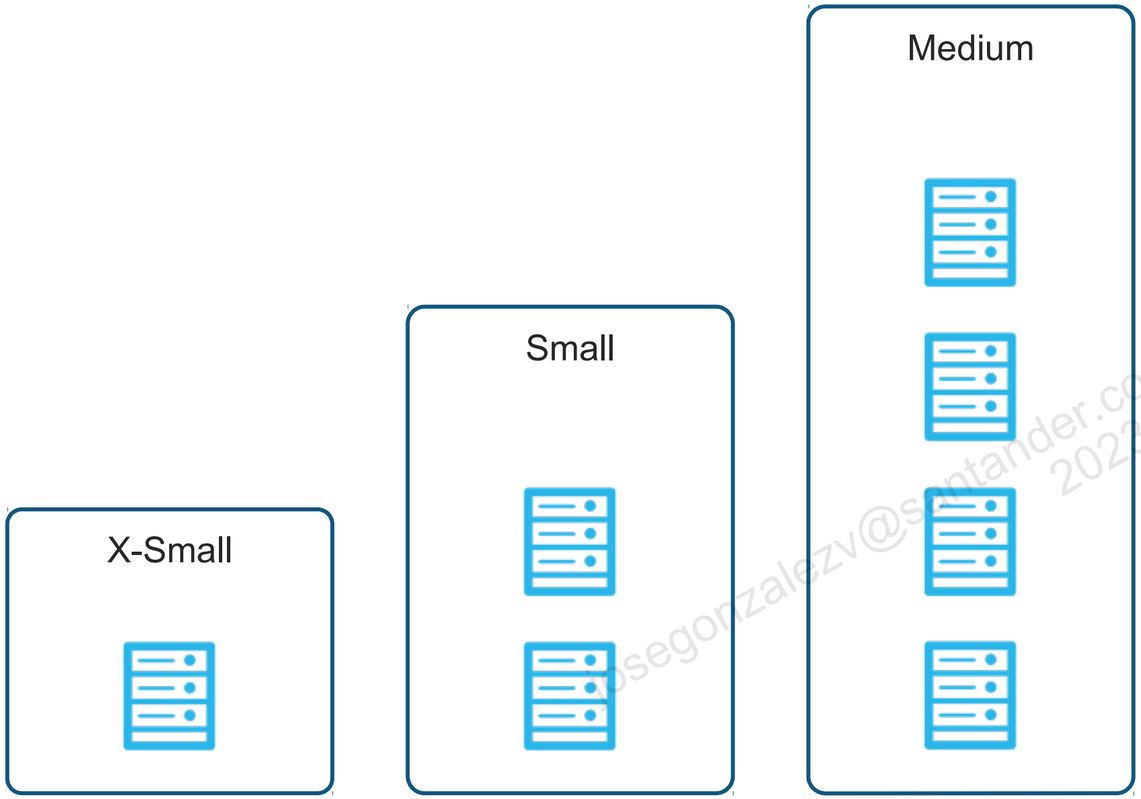


- Assign virtual warehouses by workload to eliminate resource contention and maximize the use of the data cache.

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

SCALE UP (OR DOWN)

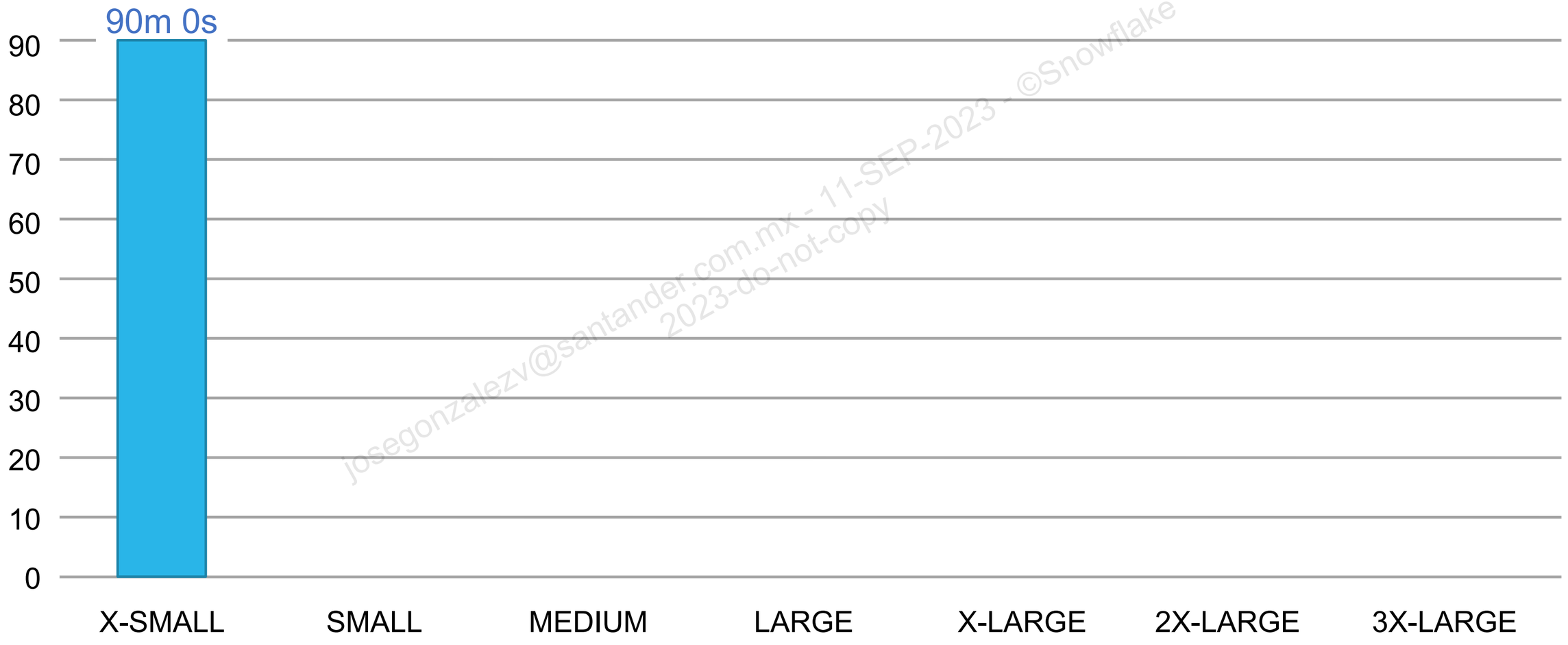


- Scale up to increase performance
- Make sure the warehouse is the right size for the workload
  - Each size up doubles the resources (and doubles the cost per second)
- Scaling is generally linear, to a point
  - Smaller warehouses are not necessarily more cost-efficient
- Find the sweet spot

jsegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

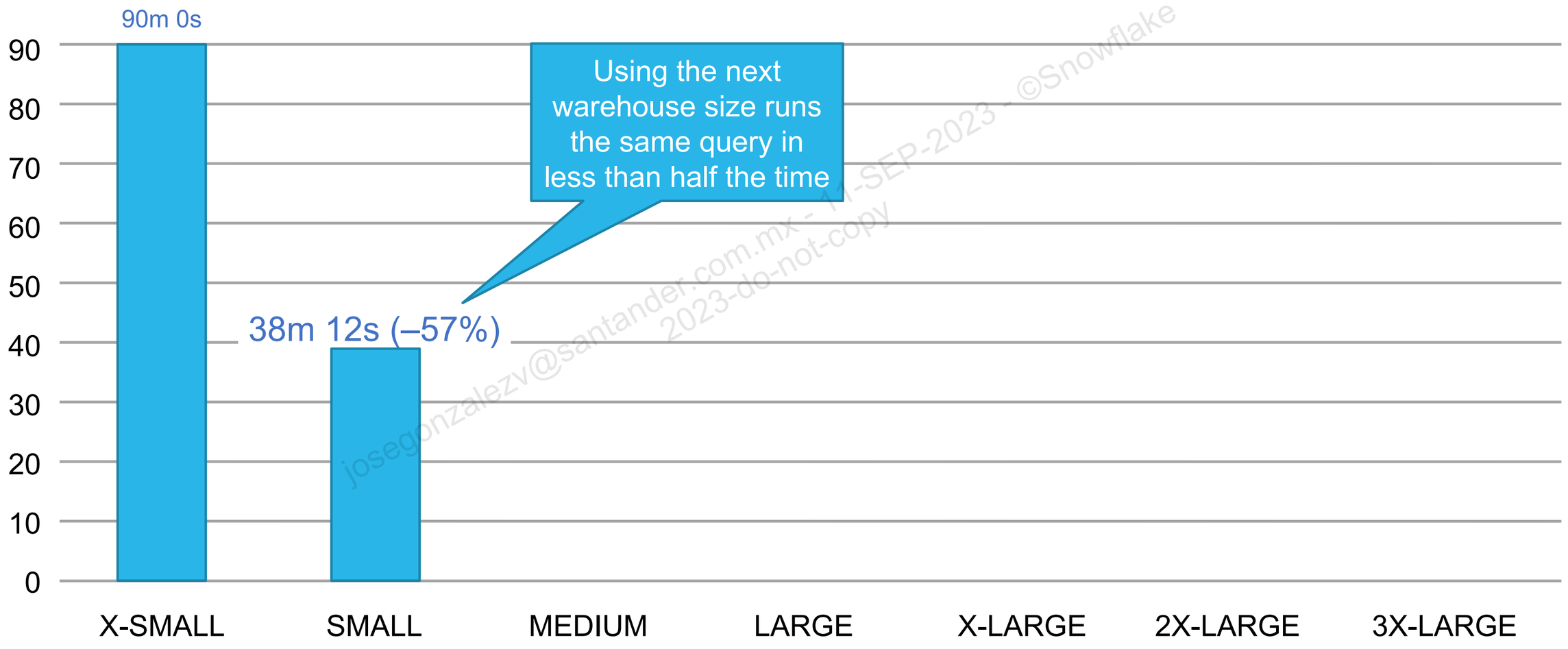
QUERY TIME BY WAREHOUSE SIZE



josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

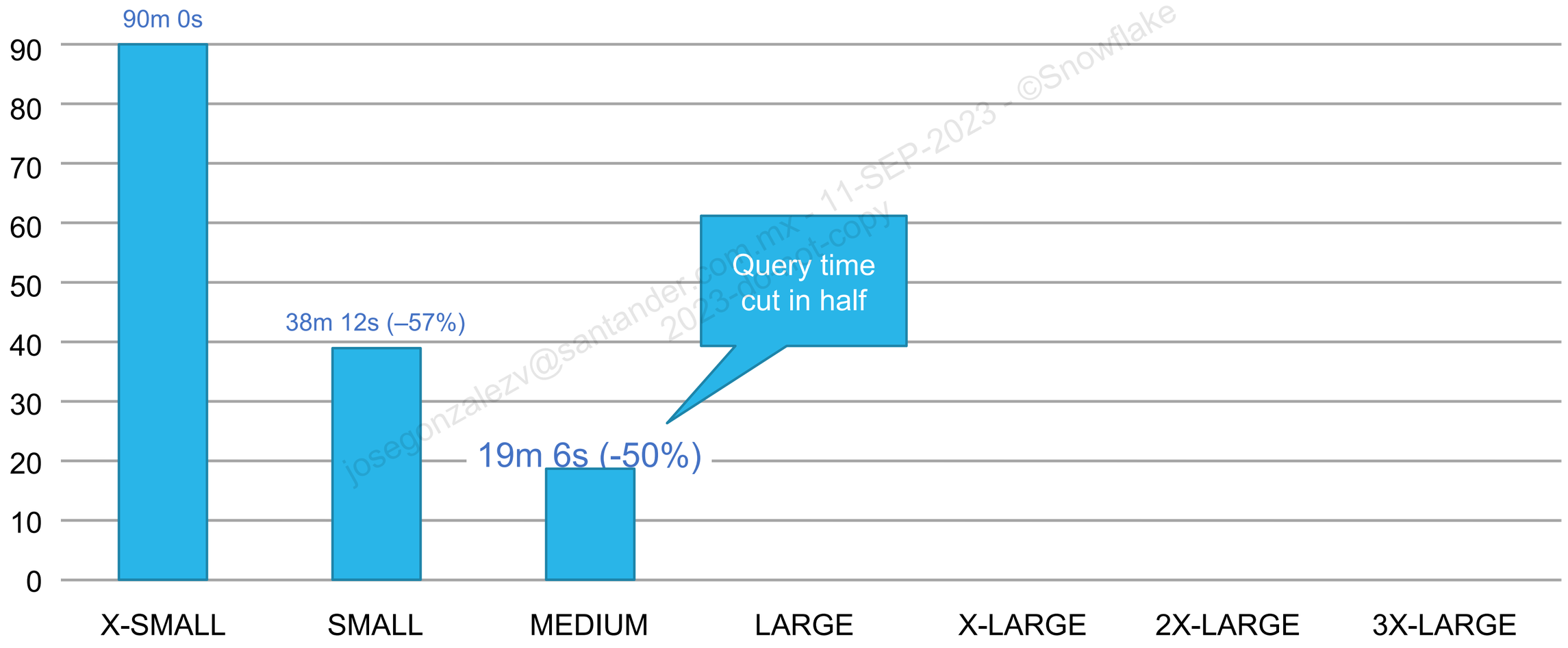
QUERY TIME BY WAREHOUSE SIZE



josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

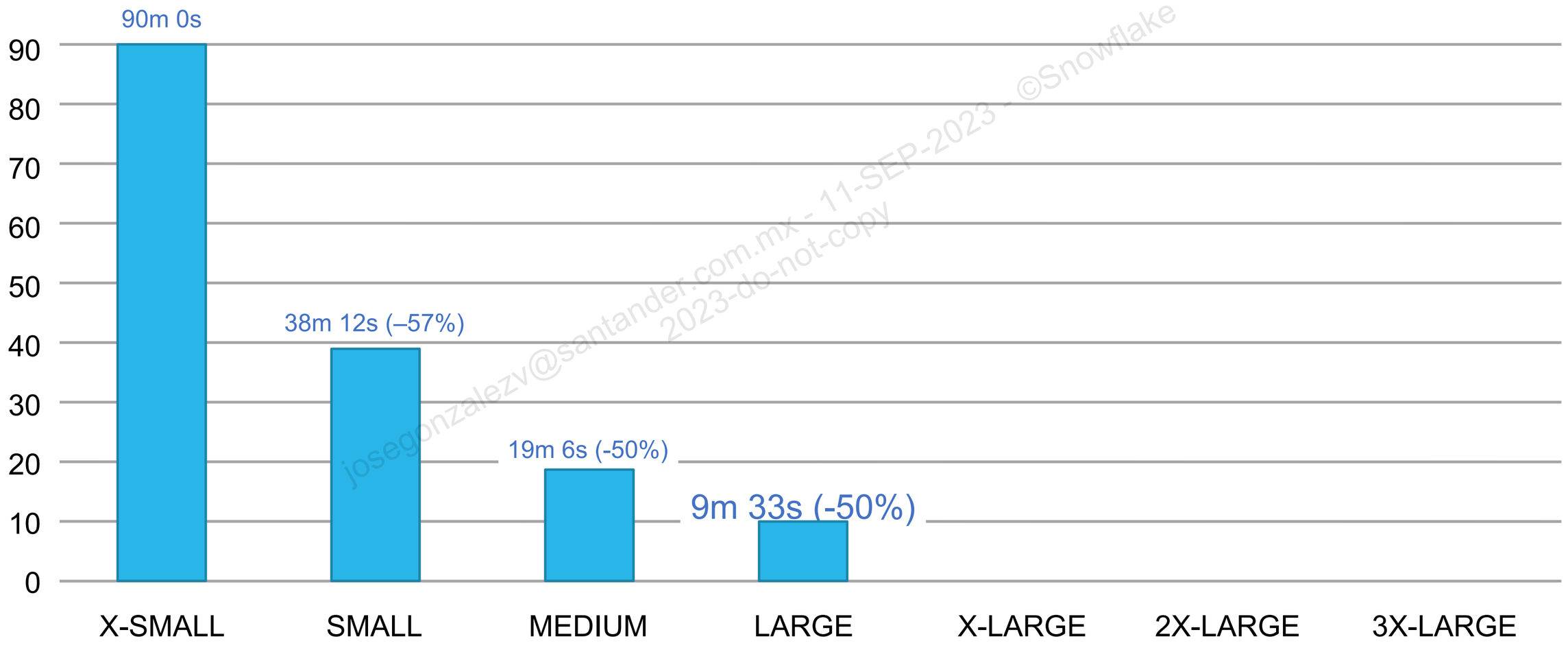
QUERY TIME BY WAREHOUSE SIZE



josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake
2023-00 - not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

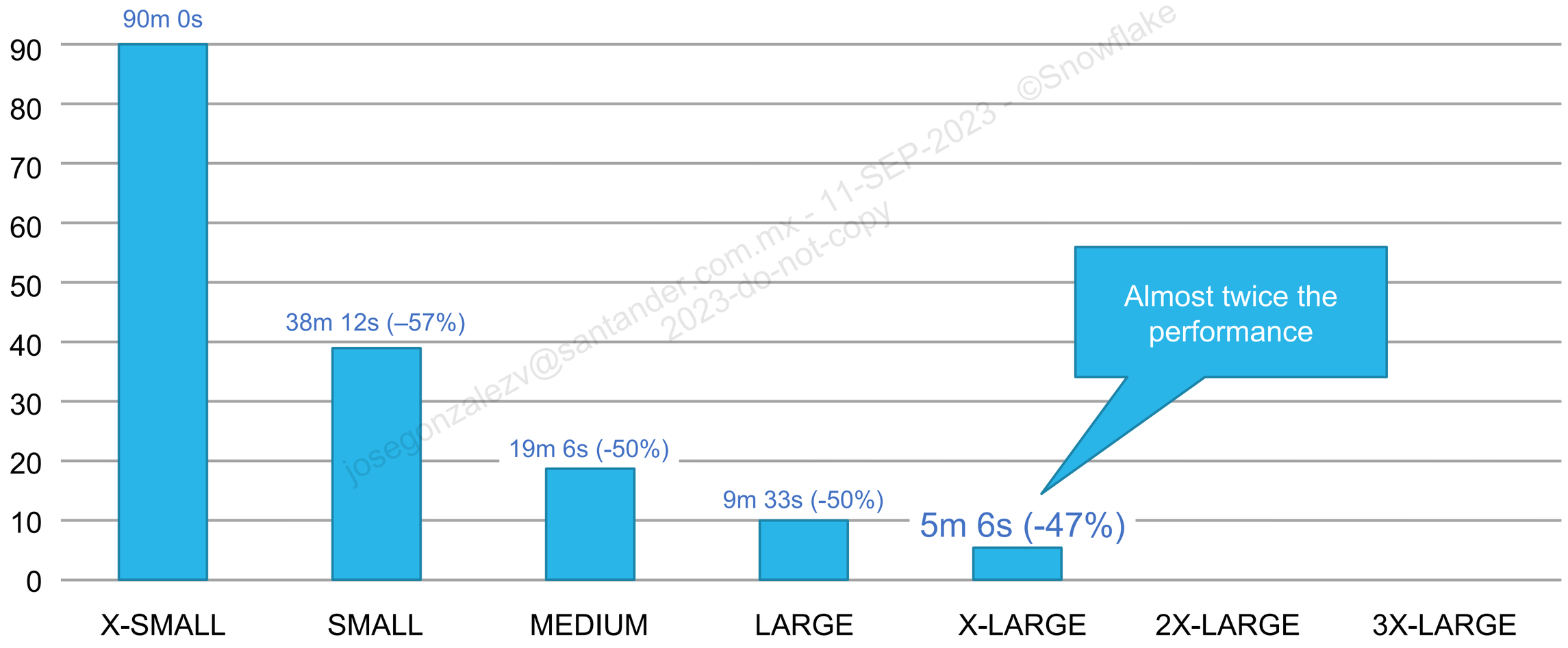
QUERY TIME BY WAREHOUSE SIZE



josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

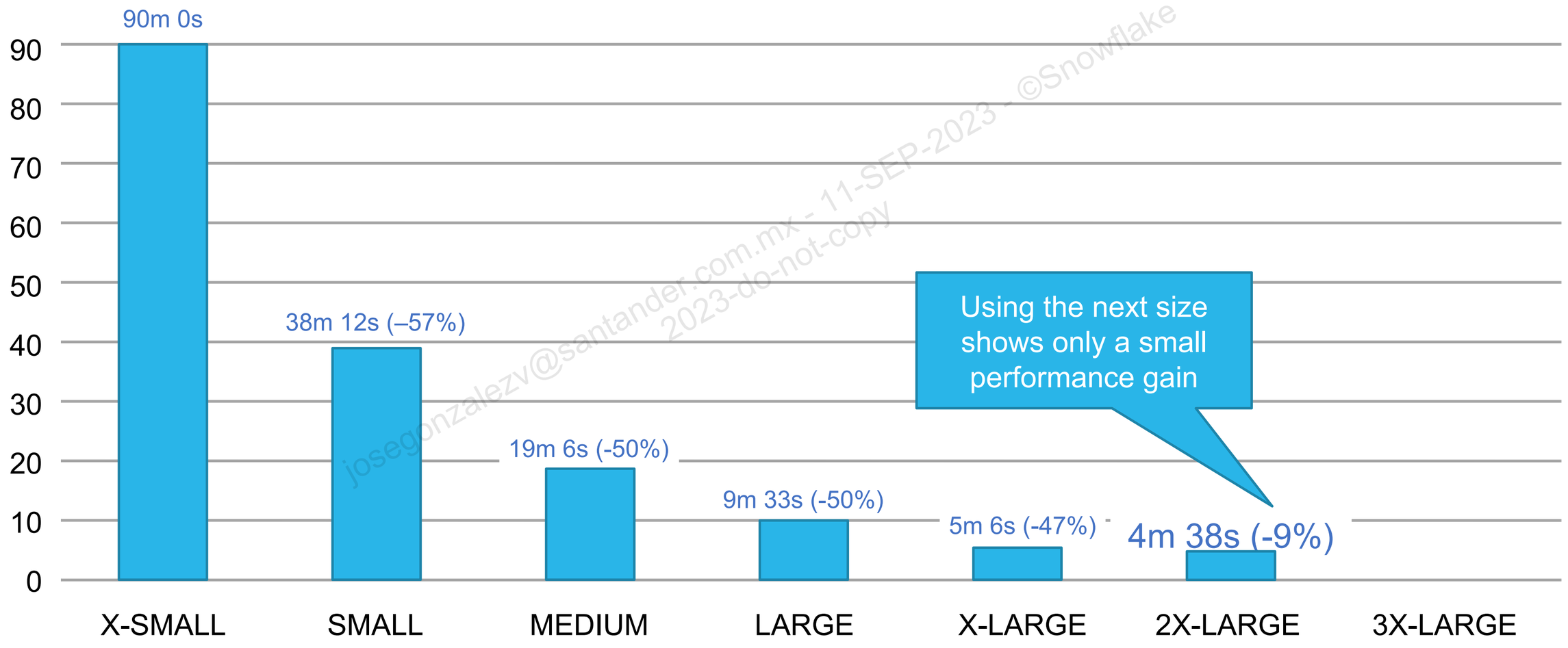
QUERY TIME BY WAREHOUSE SIZE



josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

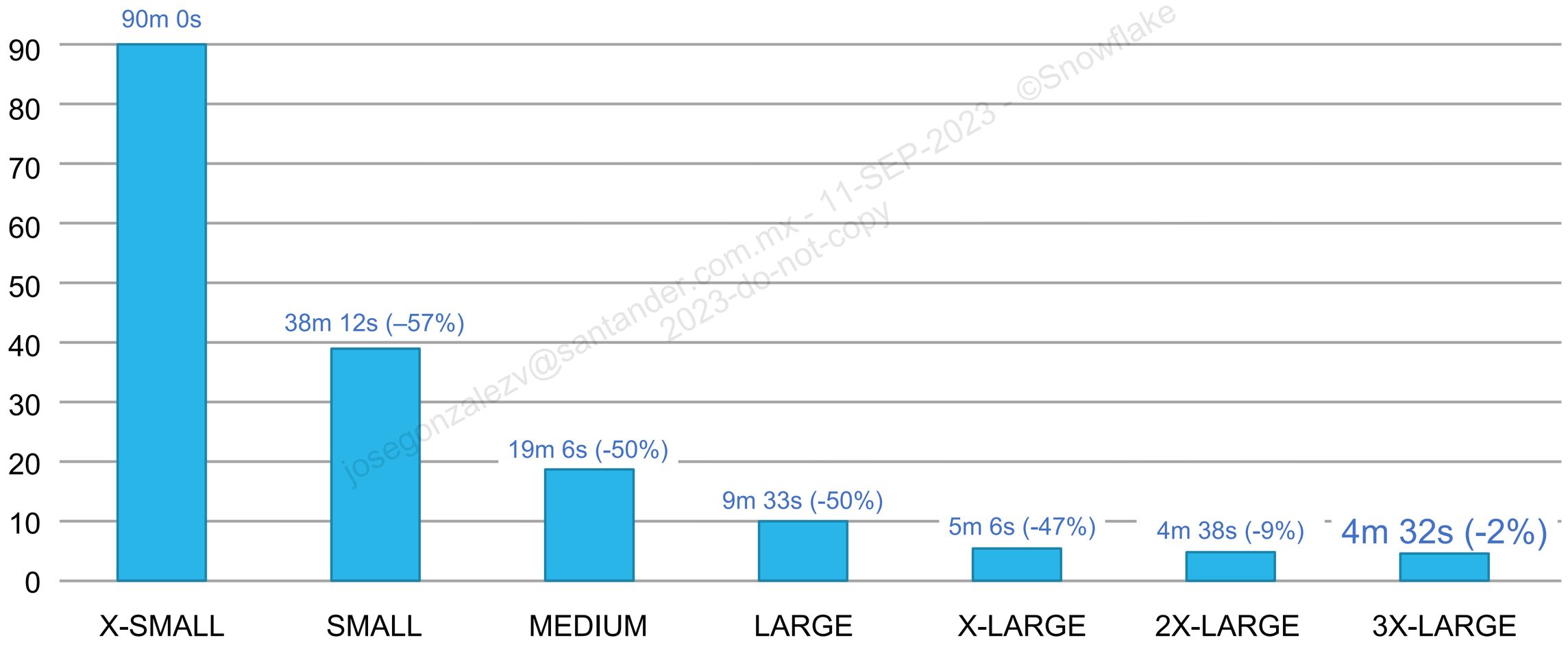
QUERY TIME BY WAREHOUSE SIZE



josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

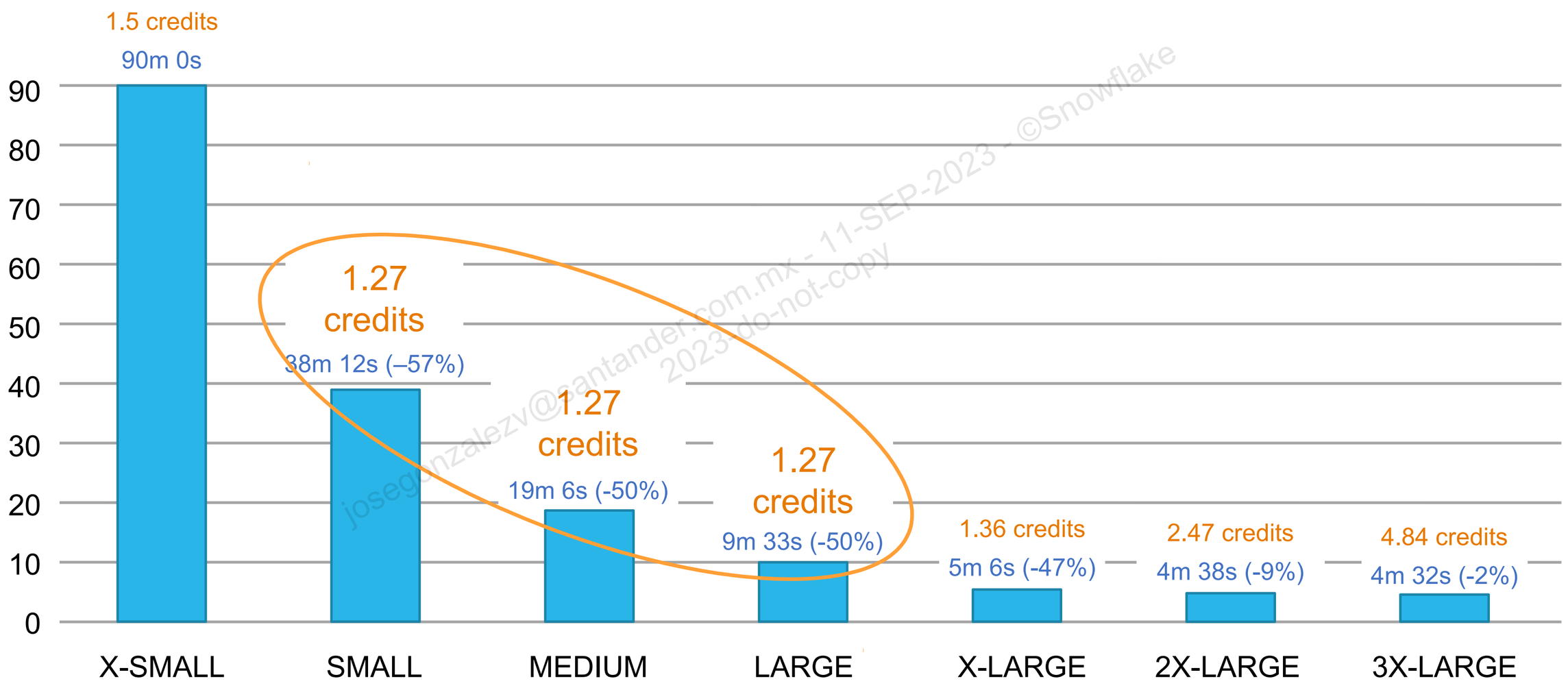
QUERY TIME BY WAREHOUSE SIZE



josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

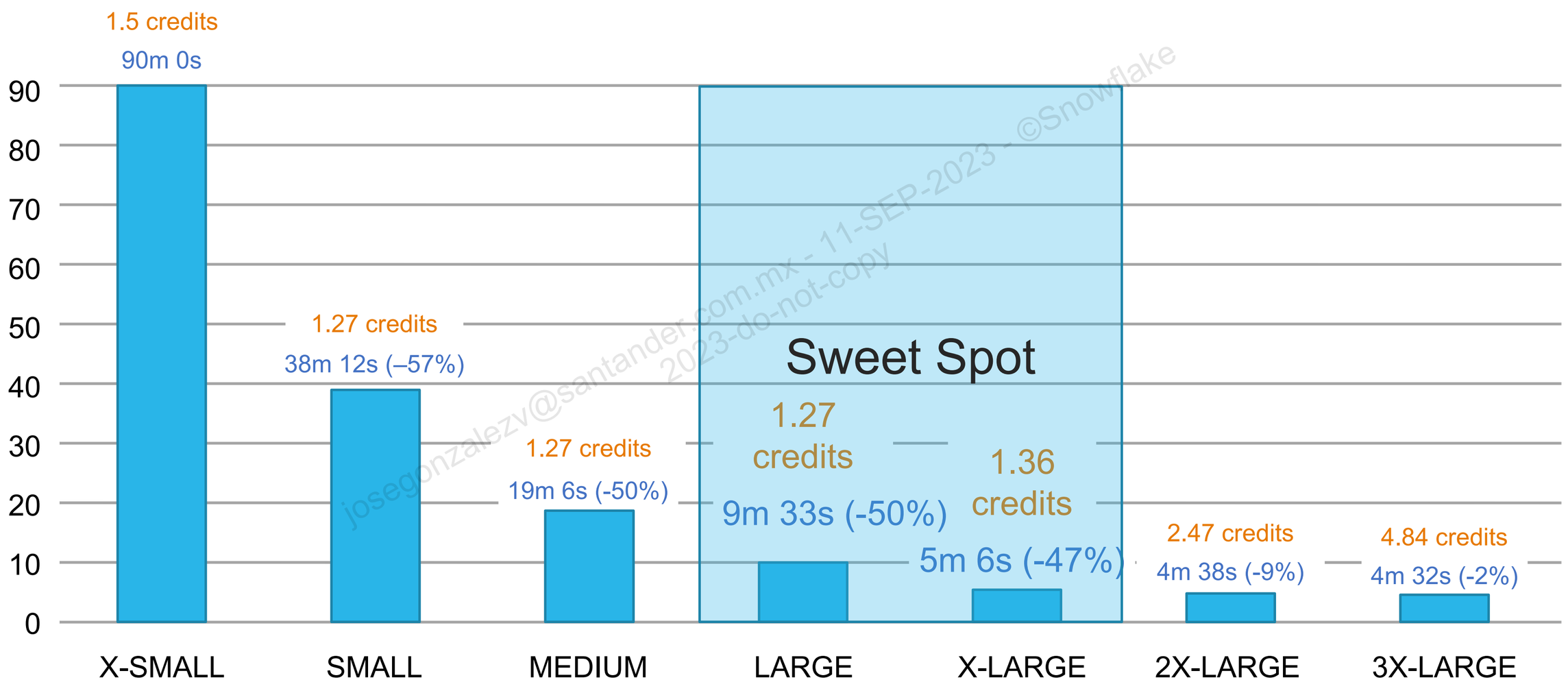
CREDITS PER QUERY



joseglonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023 do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

CREDITS PER QUERY

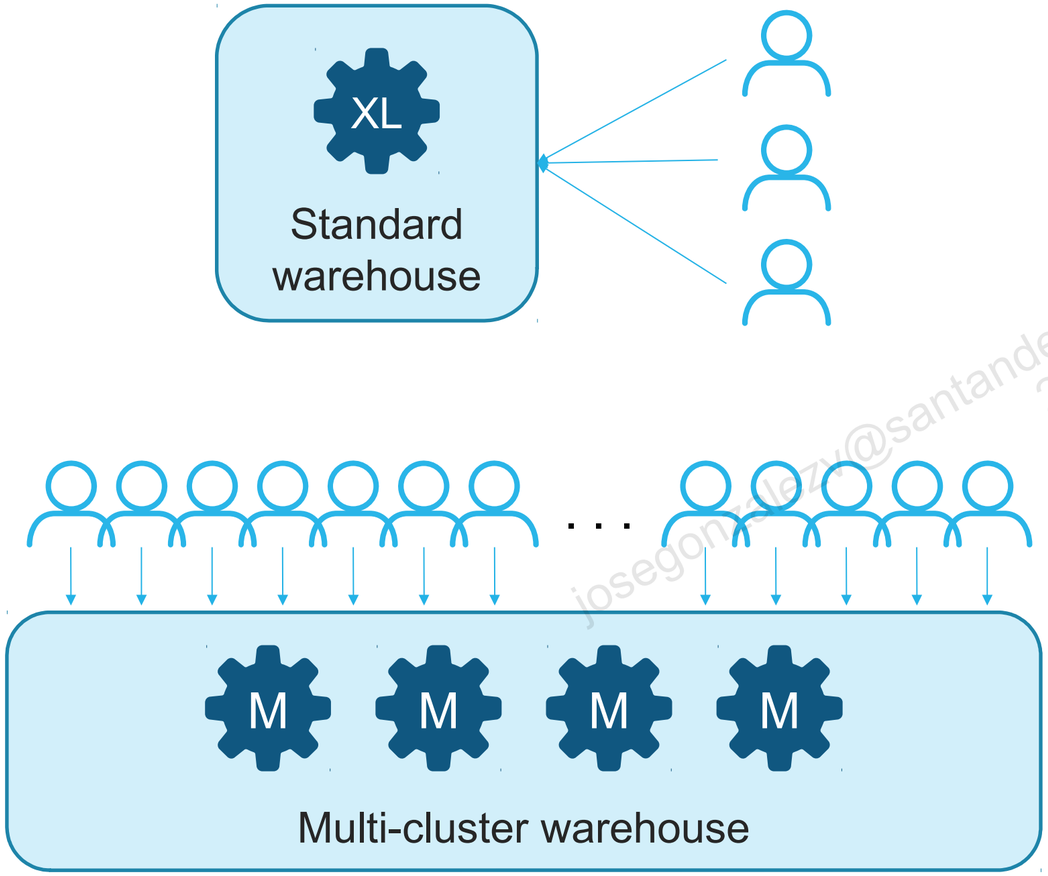


josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

SCALE OUT (OR BACK)

HANDLE GREATER CONCURRENCY



- Use a multi-cluster warehouse when many users are hitting the same warehouse
  - Example: reporting tools
- Scaling up will help to some extent with concurrency, but scaling out will help much more
- Each cluster added is the same size as the original
- Snowflake automatically scales out and back based on usage

josegonzalezv@santander.com.mx - 11-SEP-2023 @Snowflake
2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

SCALE UP VS SCALE OUT

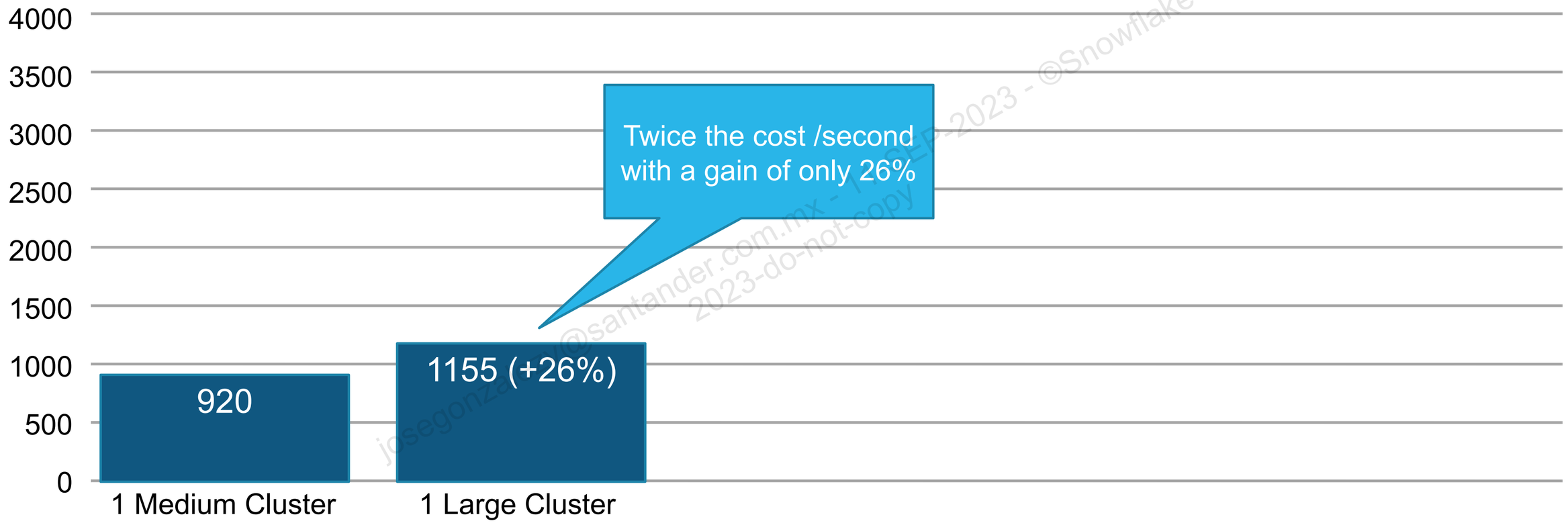
RECORDS PROCESSED BY 50 CONCURRENT USERS

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

SCALE UP VS SCALE OUT

RECORDS PROCESSED BY 50 CONCURRENT USERS

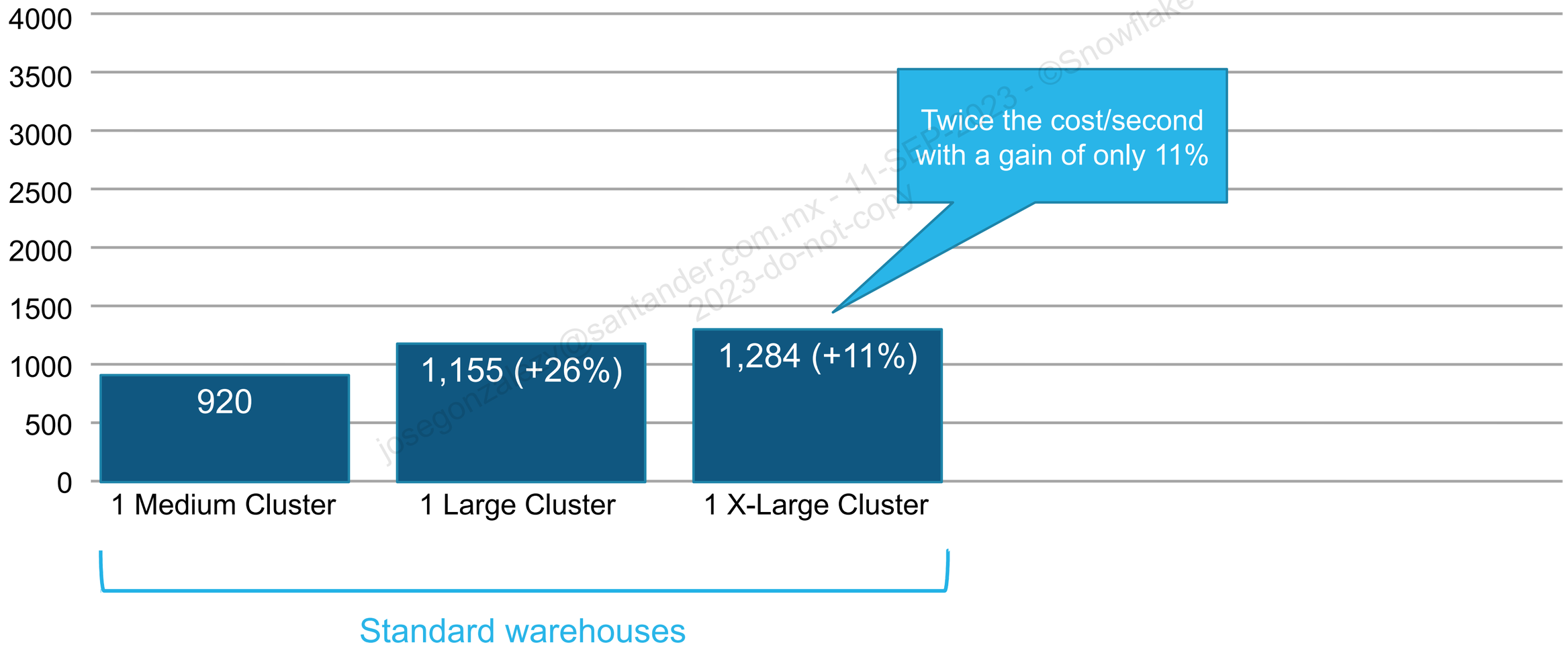


josegonzalez@santander.com.mx - 14-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

SCALE UP VS SCALE OUT

RECORDS PROCESSED BY 50 CONCURRENT USERS

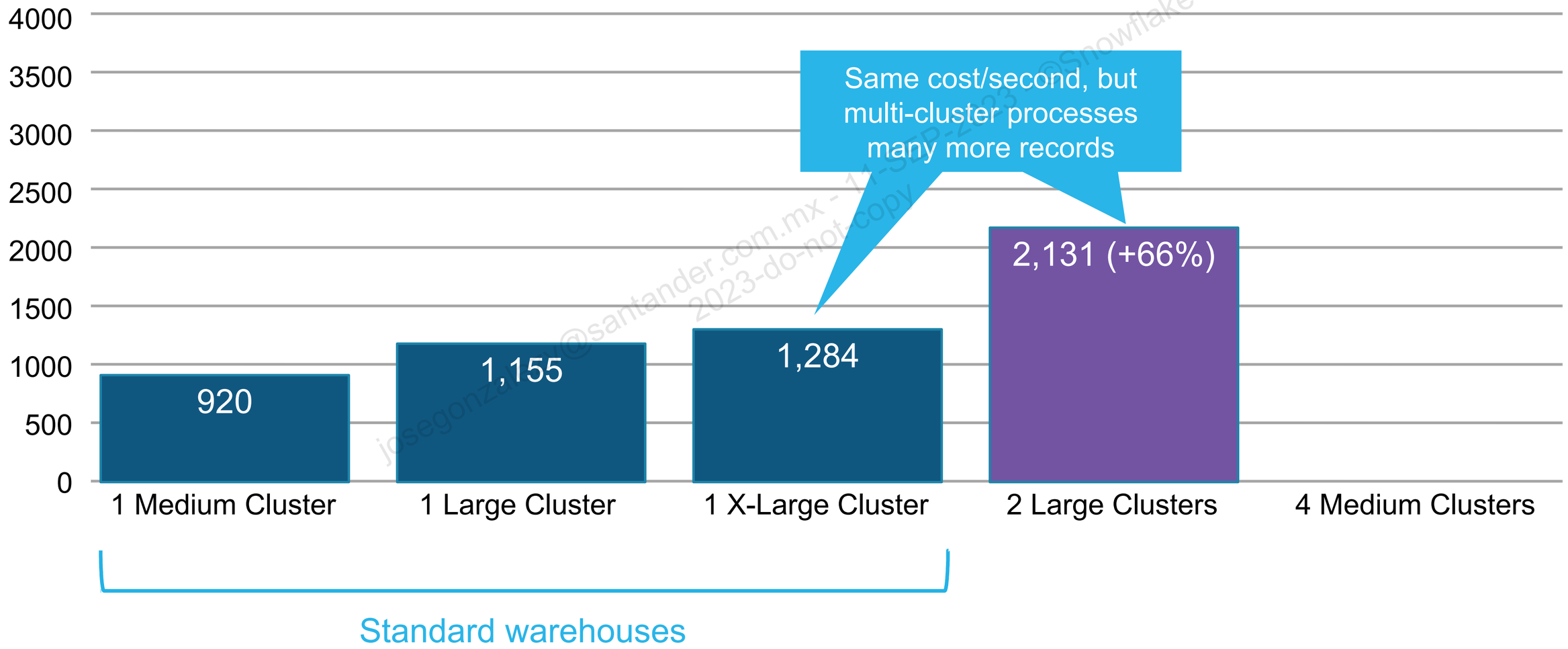


josegonzalez@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

SCALE UP VS SCALE OUT

RECORDS PROCESSED BY 50 CONCURRENT USERS

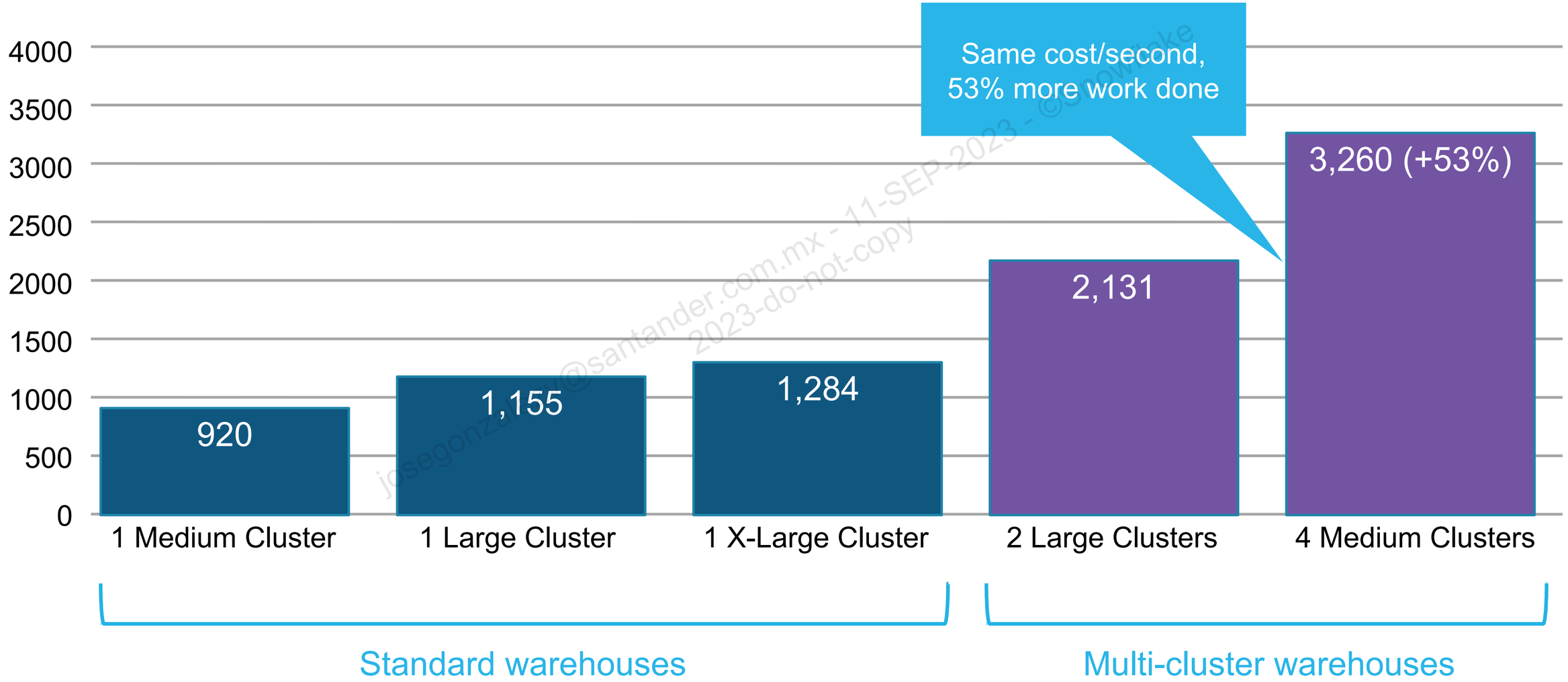


josegonzalez@santander.com.mx - 11-SEP-2023 @Snowflake
2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

SCALE UP VS SCALE OUT

RECORDS PROCESSED BY 50 CONCURRENT USERS

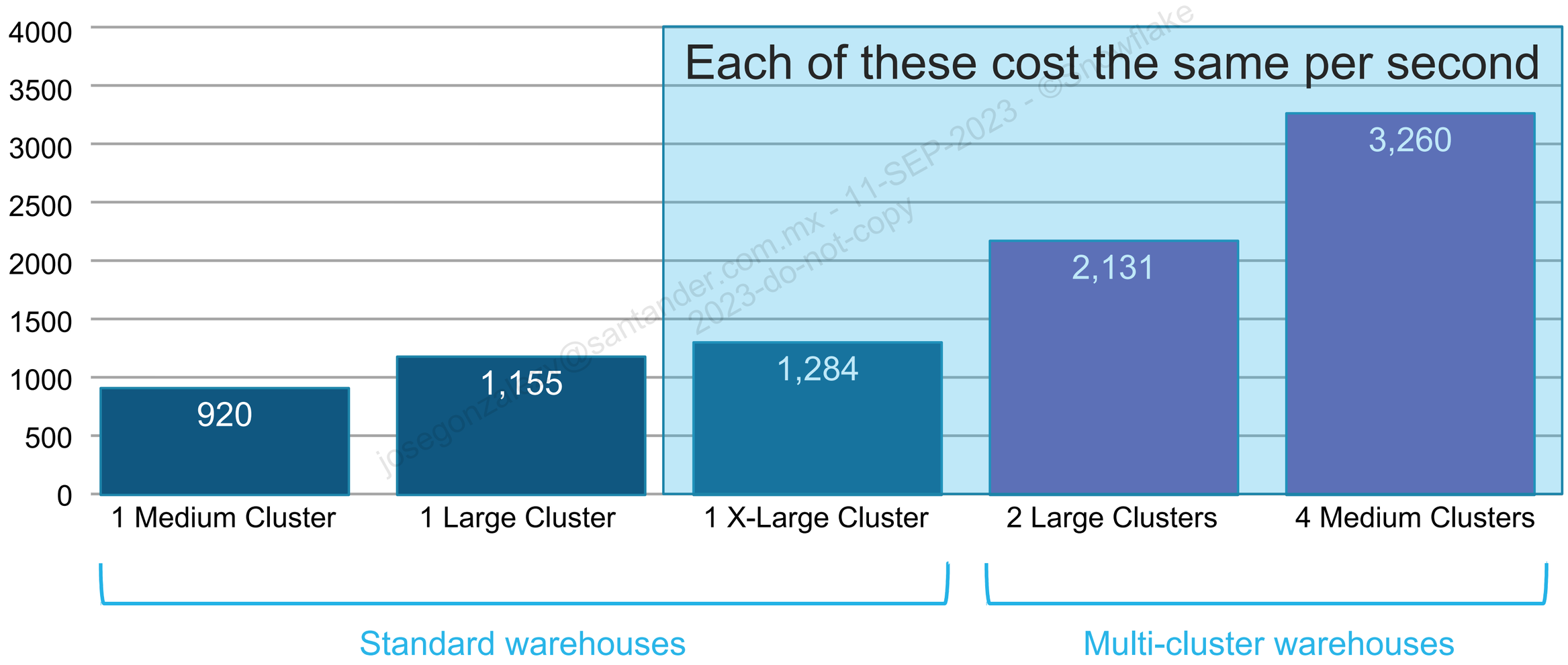


josegonzalez@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

SCALE UP VS SCALE OUT

RECORDS PROCESSED BY 50 CONCURRENT USERS



© 2023 Snowflake Computing Inc. All Rights Reserved

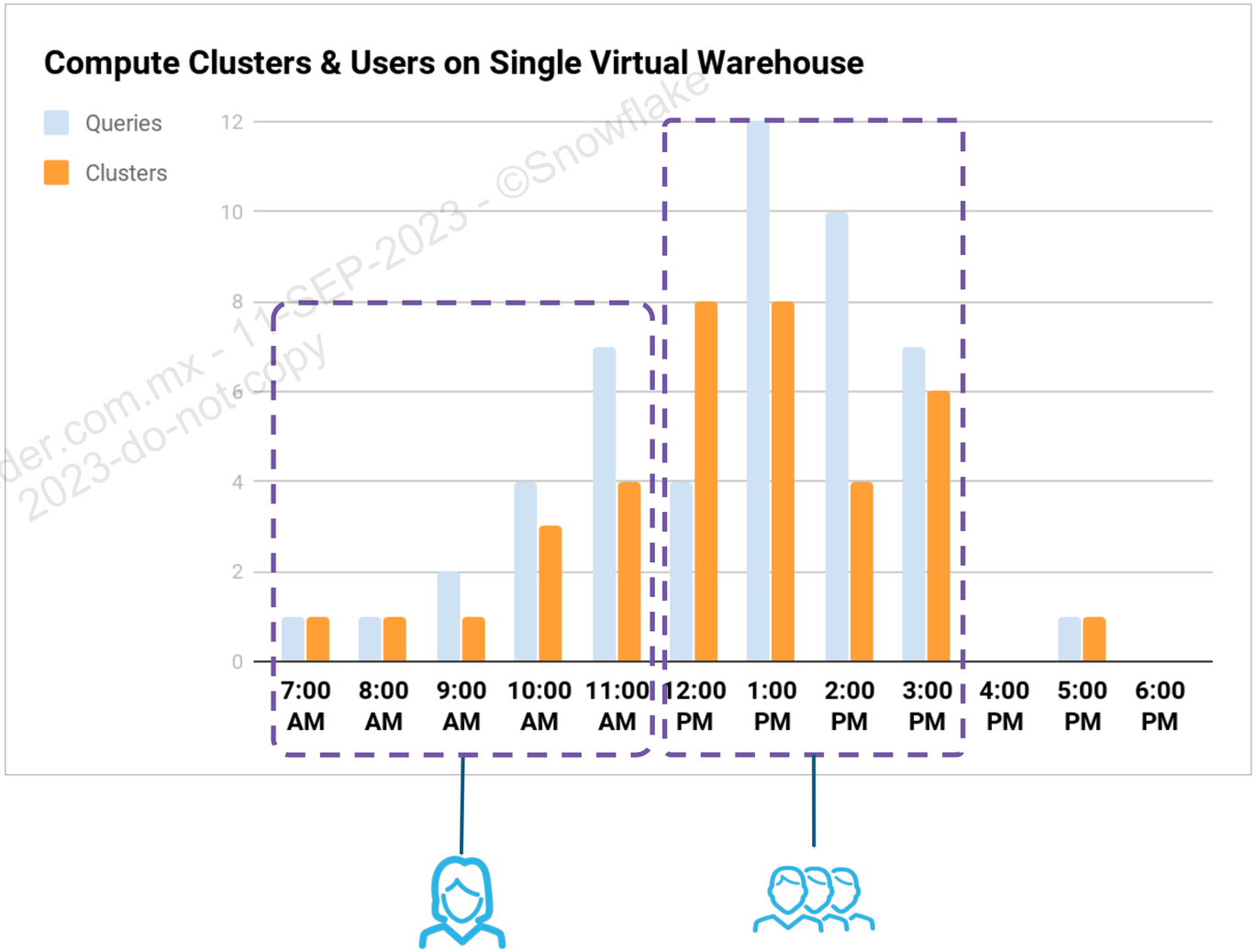
AUTO-SCALING POLICIES

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake
2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

SCALING OUT: AUTO-SCALING

Allows Snowflake to start and stop clusters as needed to dynamically manage the load on the warehouse



josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

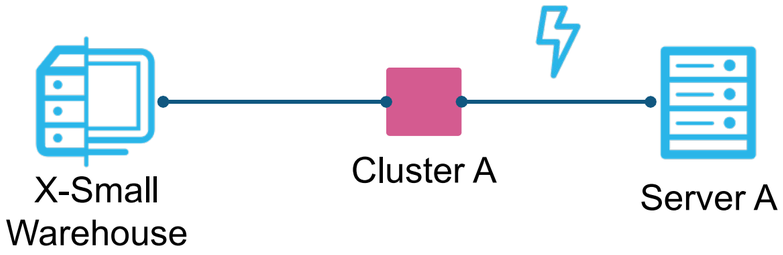
© 2023 Snowflake Computing Inc. All Rights Reserved

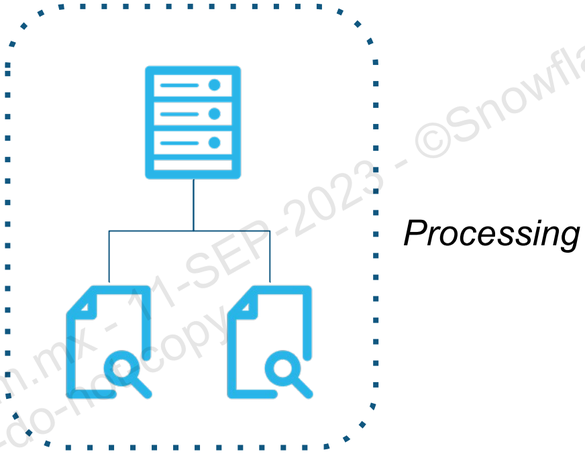
AUTO-SCALING: SCALING POLICY

Policy	Description	Cluster Starts...	Cluster Shuts down...
Standard	Minimizes queuing and starts additional clusters.	Immediately when either a query is queued or the system detects that there's one more query than the currently-running clusters can execute.	After 2-3 consecutive checks (performed at 1 minute intervals) determine that the load on the least-loaded cluster could be redistributed to the others without spinning up the cluster again.
Economy	Favors keeping running clusters fully-loaded rather than starting additional clusters; could result in jobs being queued rather than starting additional Clusters.	**Only** if the system estimates there's enough query load to keep the cluster busy for at least 6 minutes.	After 5-6 consecutive checks (performed at 1 minute intervals), determine that the load on the least-loaded cluster could be redistributed to the others without spinning up the cluster again.

© 2023 Snowflake Computing Inc. All Rights Reserved

SCALING OUT POLICY: STANDARD

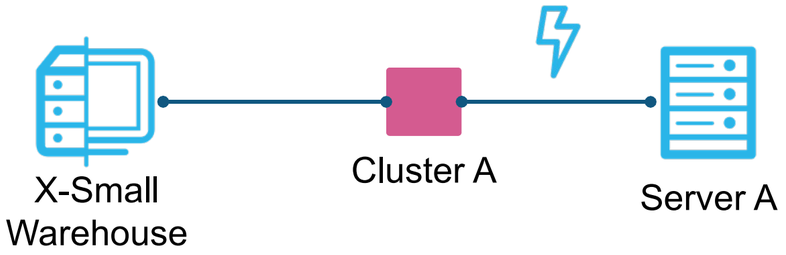


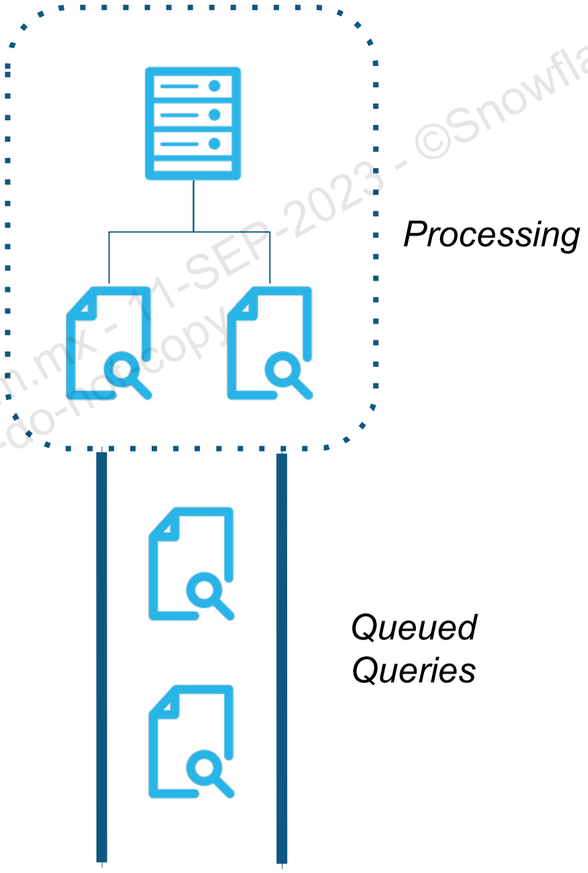


josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake
2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

SCALING OUT POLICY: STANDARD

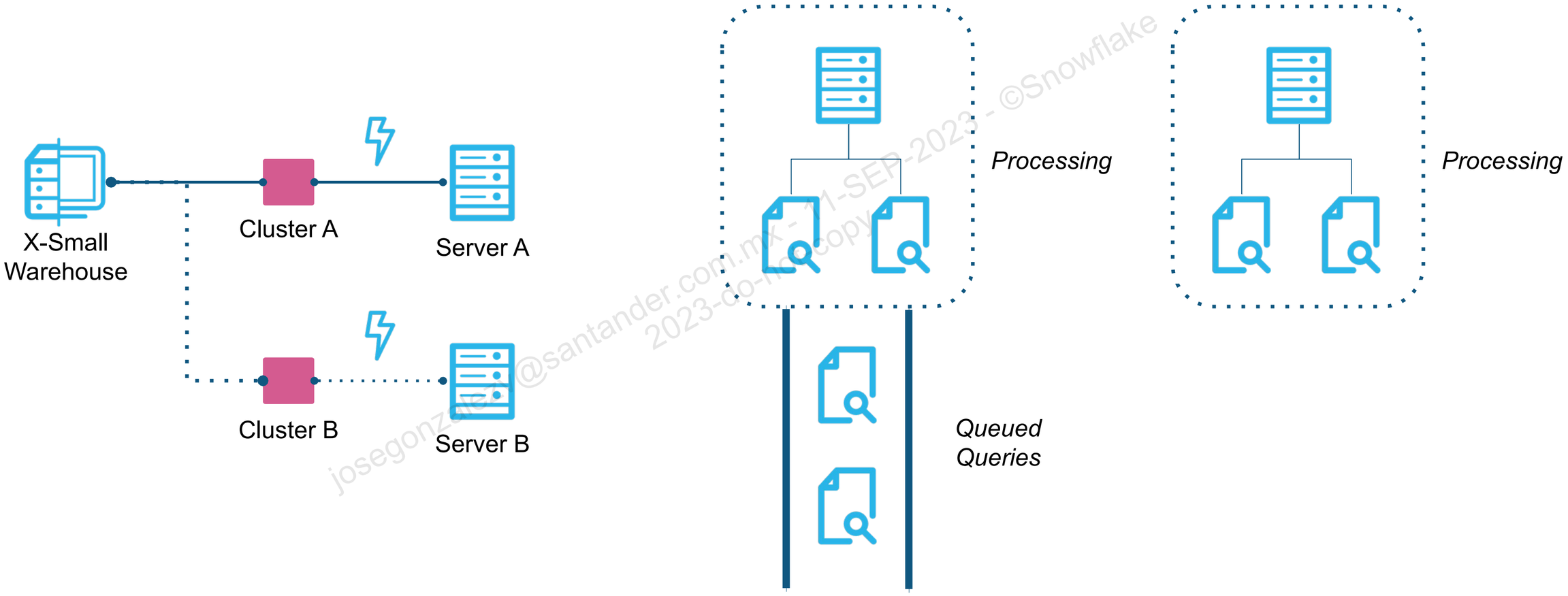




josegonzalezv@santander.com.mx - 21-SEP-2023 - ©Snowflake
2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

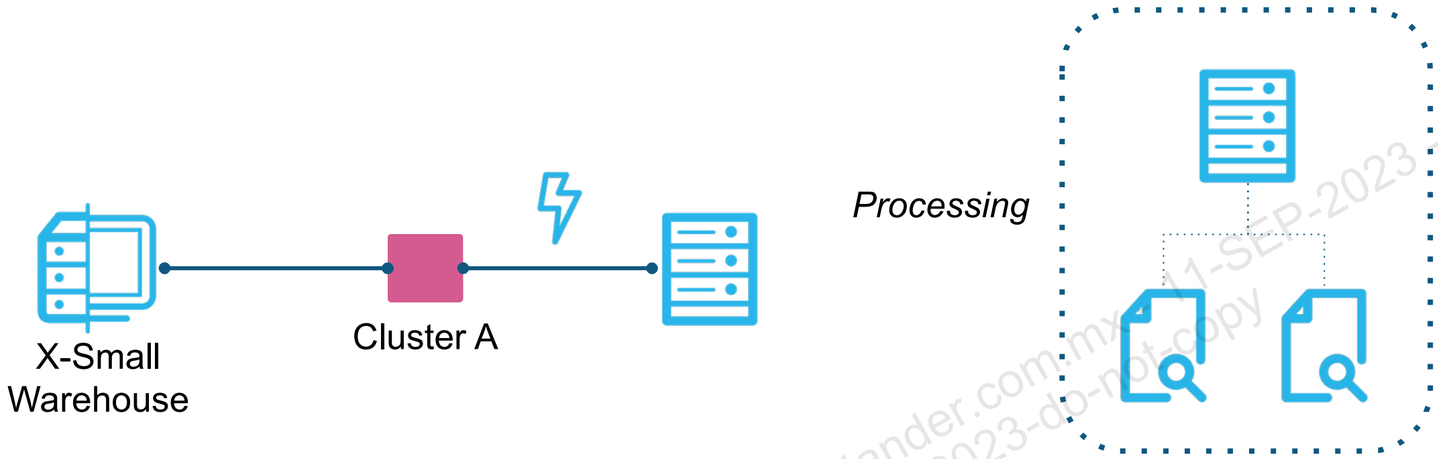
SCALING OUT POLICY: STANDARD



josegonzalez@santander.com.mx - 11-SEP-2023 - ©Snowflake
2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

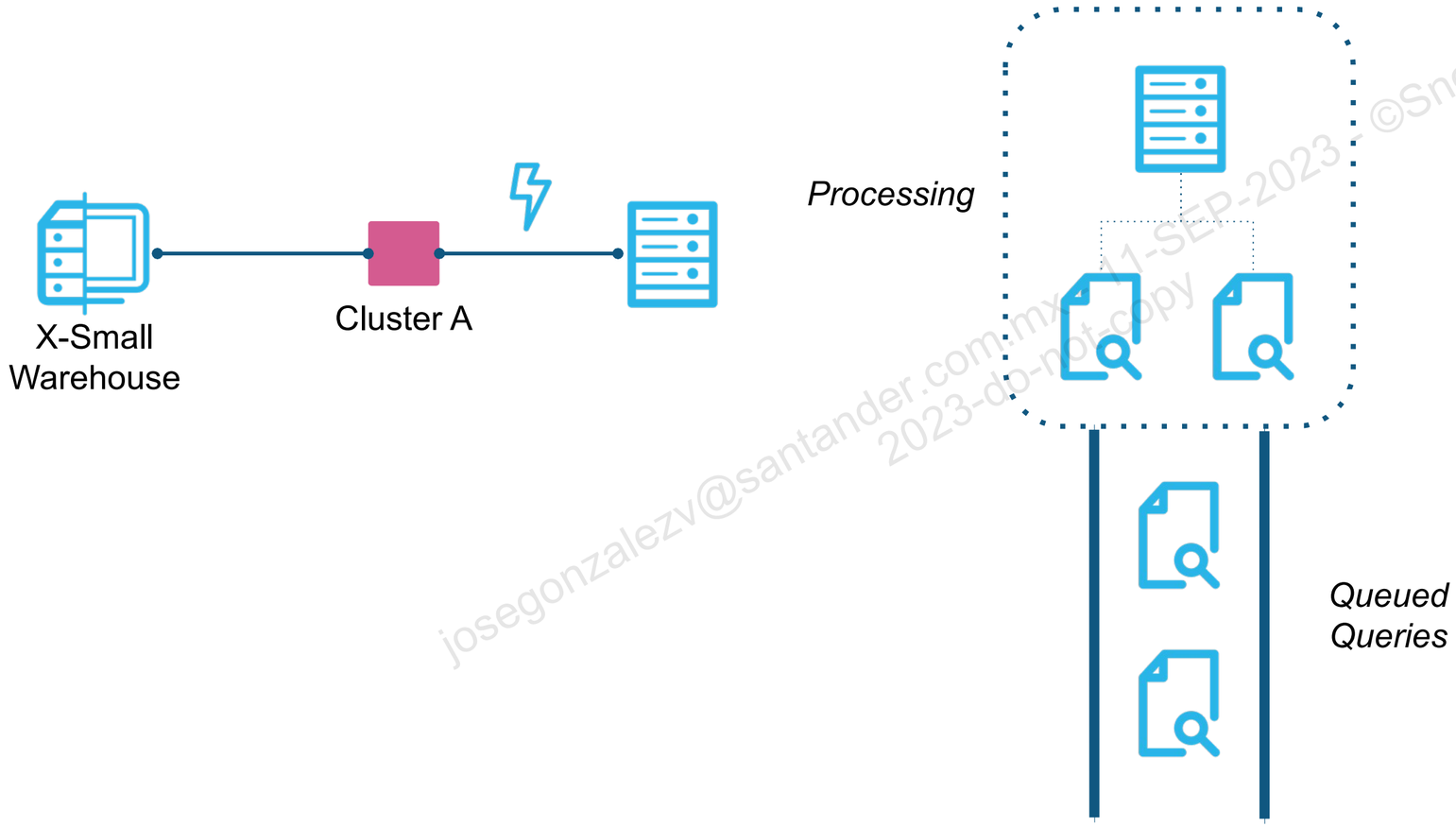
SCALING OUT POLICY: ECONOMY



josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake
2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

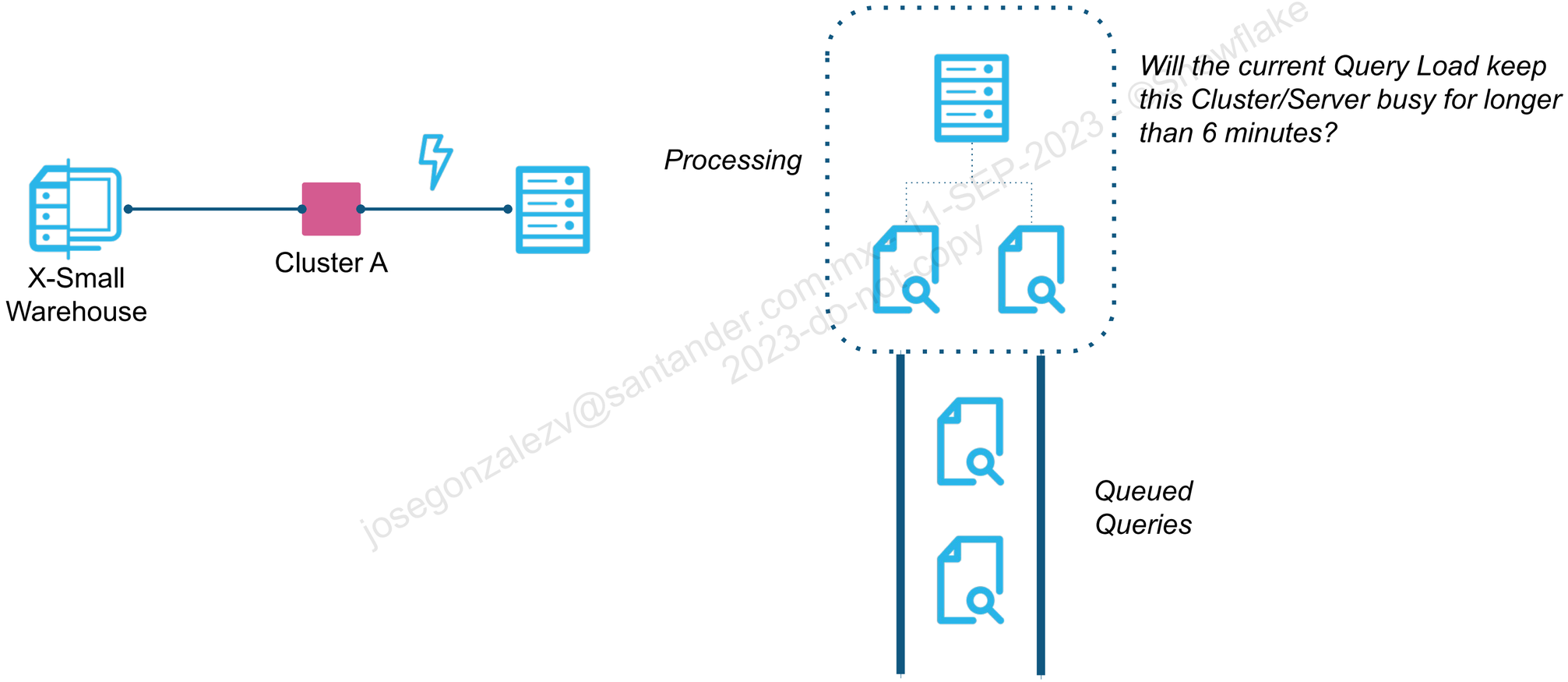
SCALING OUT POLICY: ECONOMY



josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake
2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

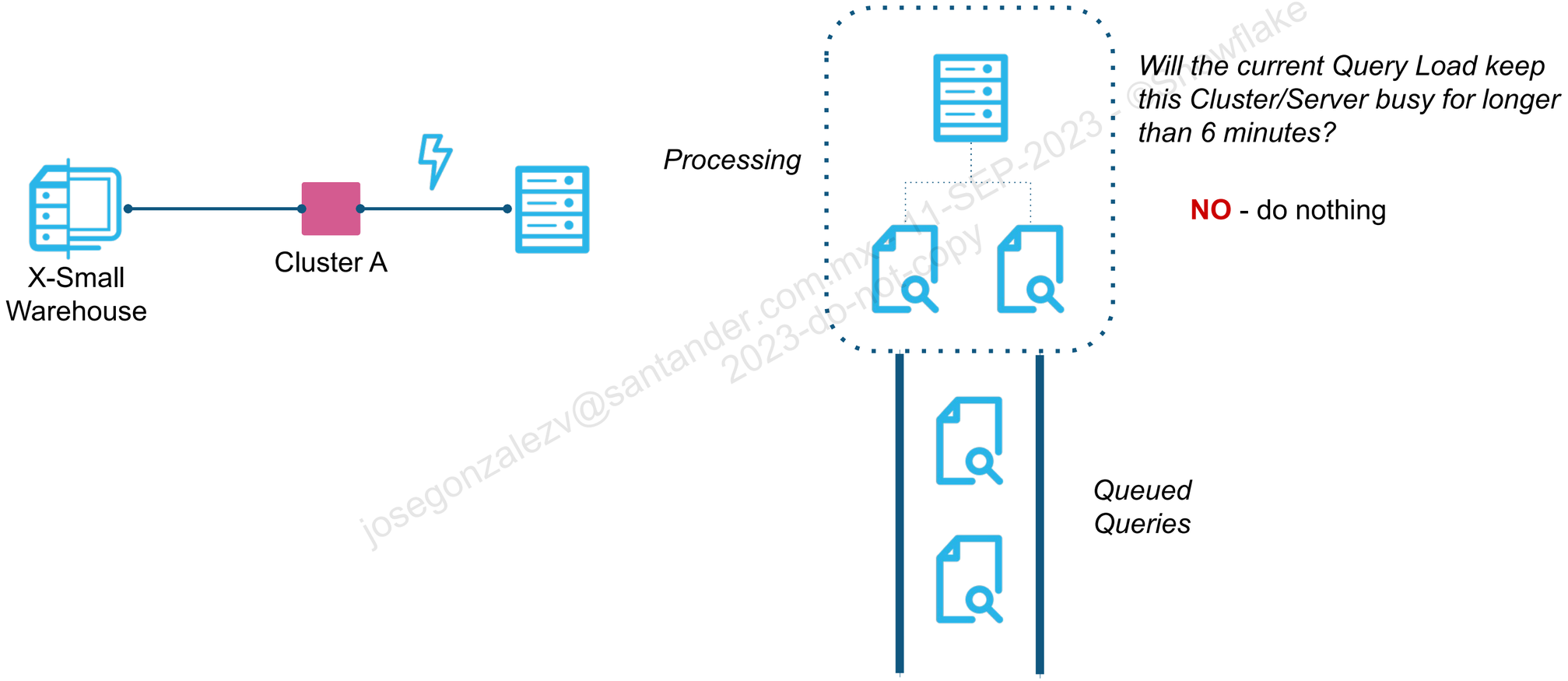
SCALING OUT POLICY: ECONOMY



josegonzalezv@santander.com.mx 11-SEP-2023 @Snowflake
2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

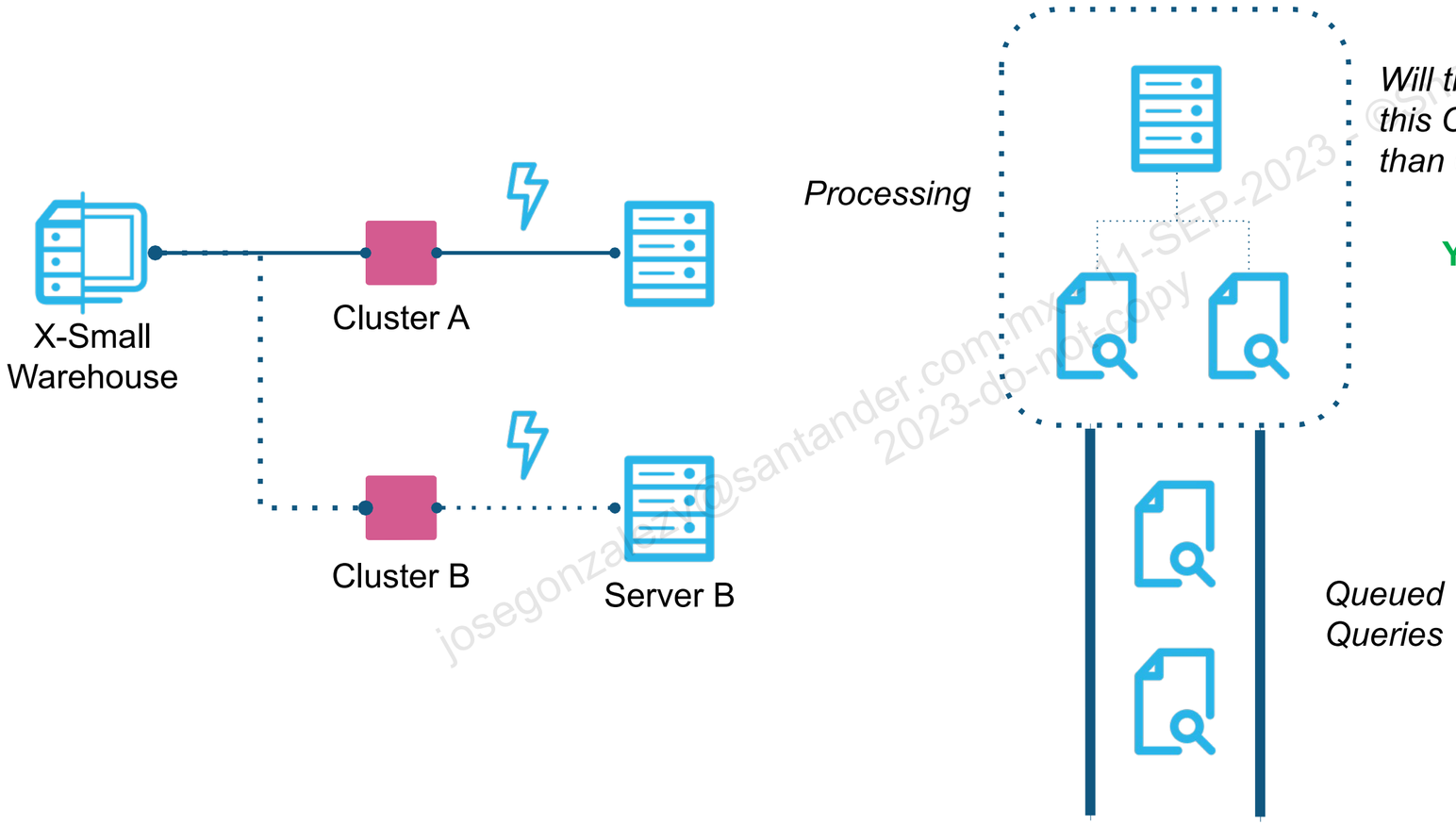
SCALING OUT POLICY: ECONOMY



josegonzalezv@santander.com.mx - 11-SEP-2023 - @Snowflake
2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

SCALING OUT POLICY: ECONOMY



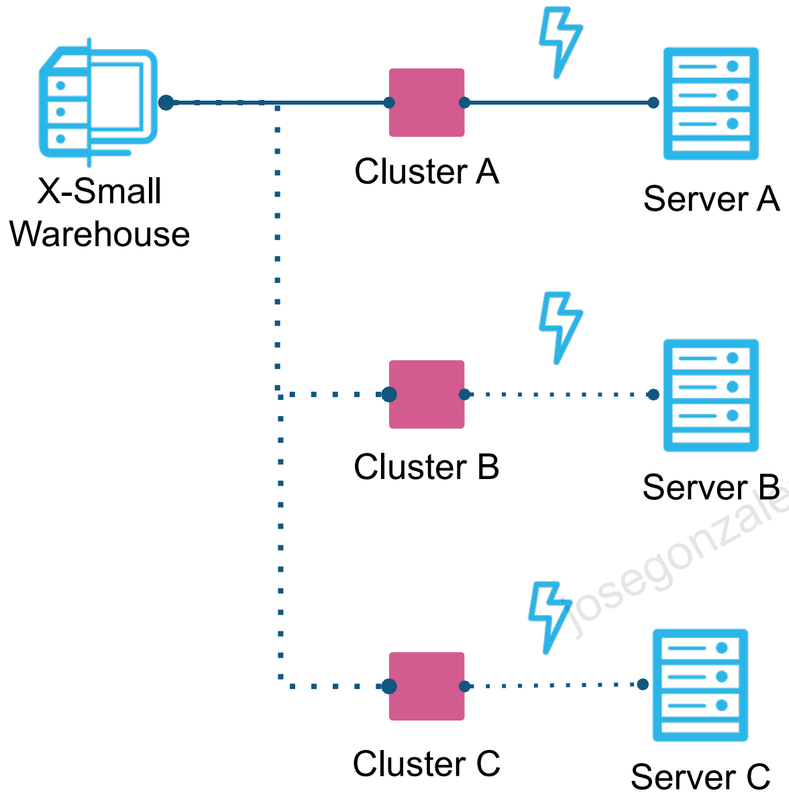
Will the current Query Load keep this Cluster/Server busy for longer than 6 minutes?

YES – spin up another cluster

josegonzalez@santander.com.mx 11-SEP-2023 @Snowflake
2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

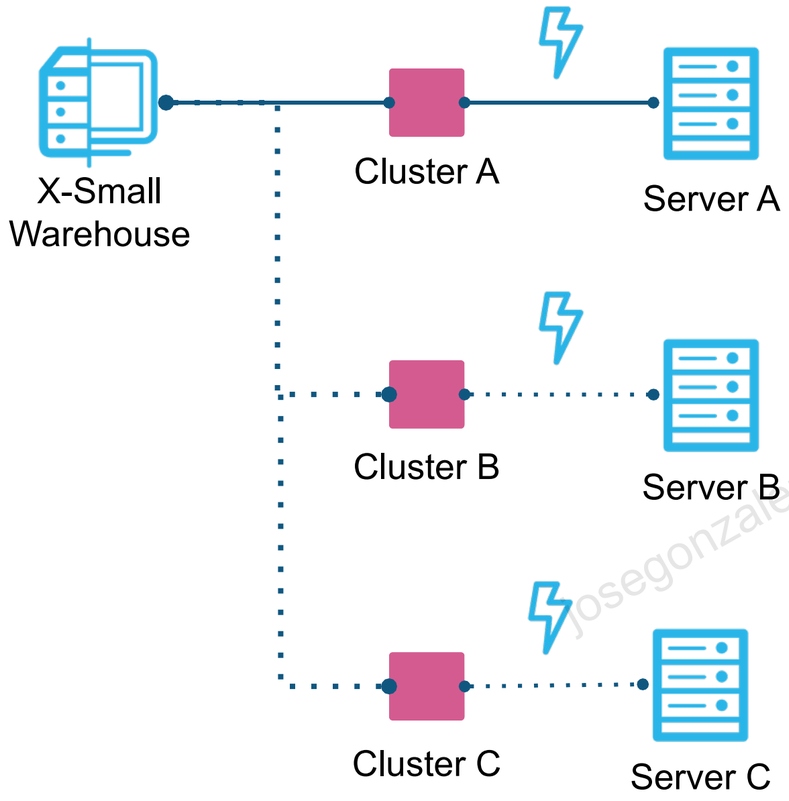
SCALING BACK POLICY: STANDARD



josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake
2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

SCALING BACK POLICY: STANDARD

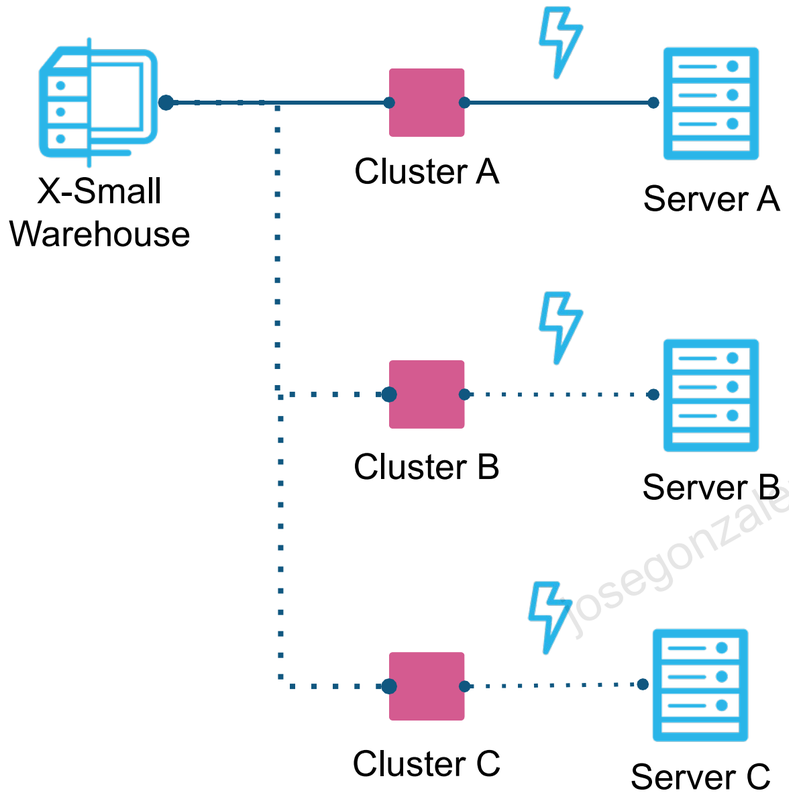


T0	*Could the load on the least-busy cluster be redistributed to the other clusters?*

josegonzalezv@santander.com.mx - 11-SEP-2023 - © Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

SCALING BACK POLICY: STANDARD



T0	*Could the load on the least-busy cluster be redistributed to the other clusters?*
T0 + 1 minute	*Could the load on the least-busy cluster be redistributed to the other clusters?*

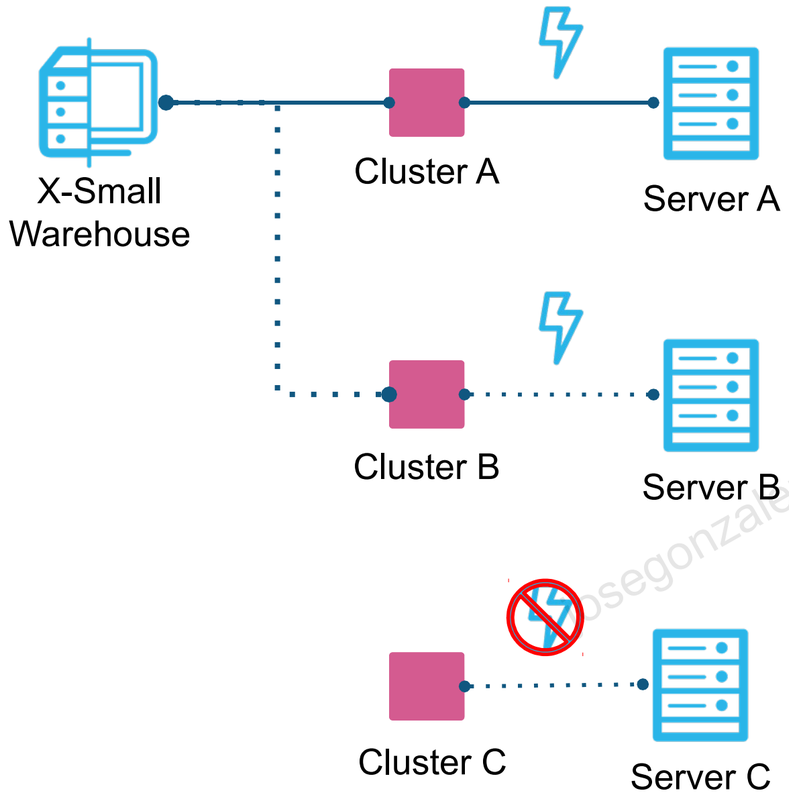
NO

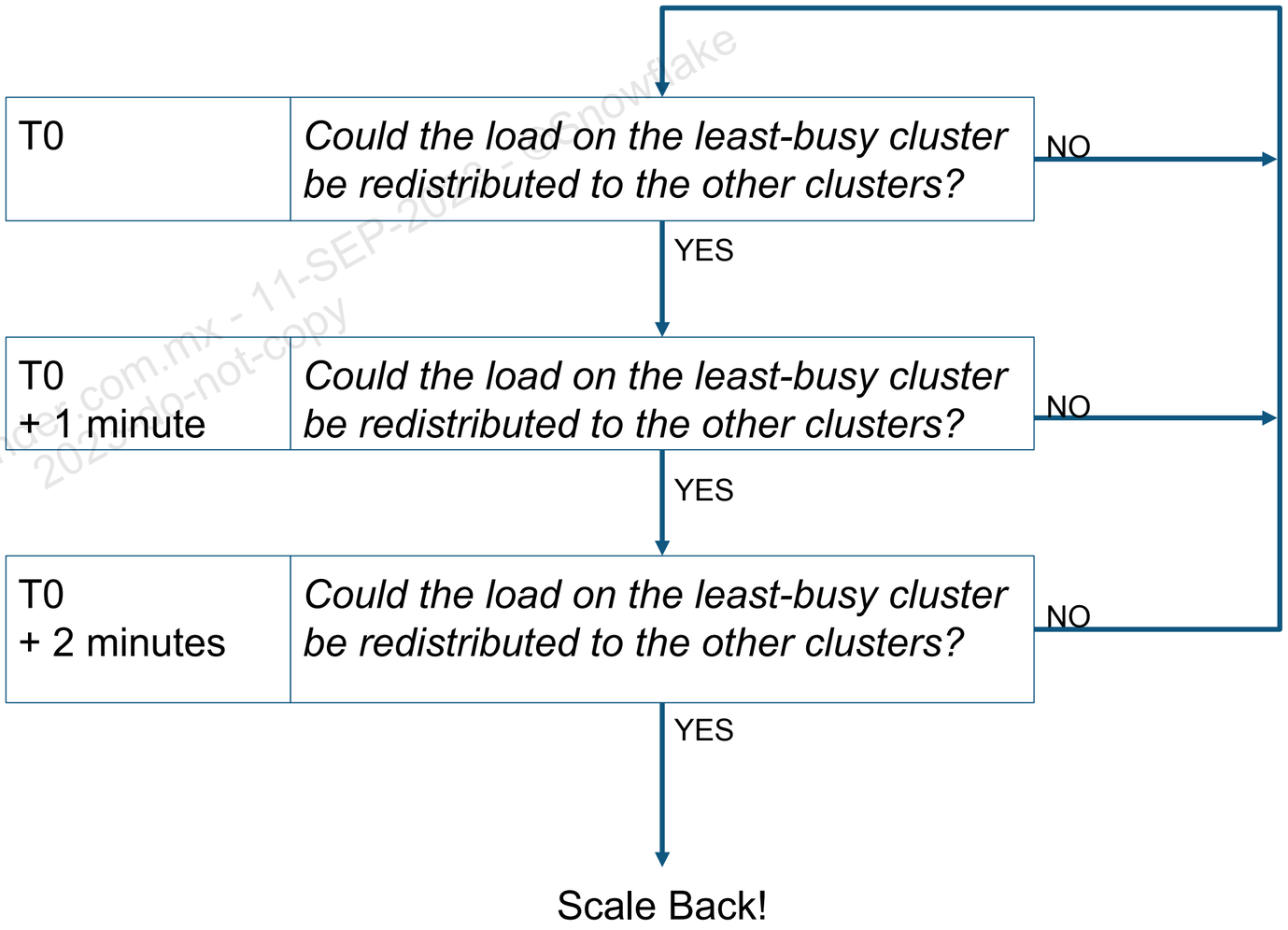
YES

josegonzalezv@santander.com.mx - 11-SEP-2025 - © Snowflake 2023 - do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

SCALING BACK POLICY: STANDARD

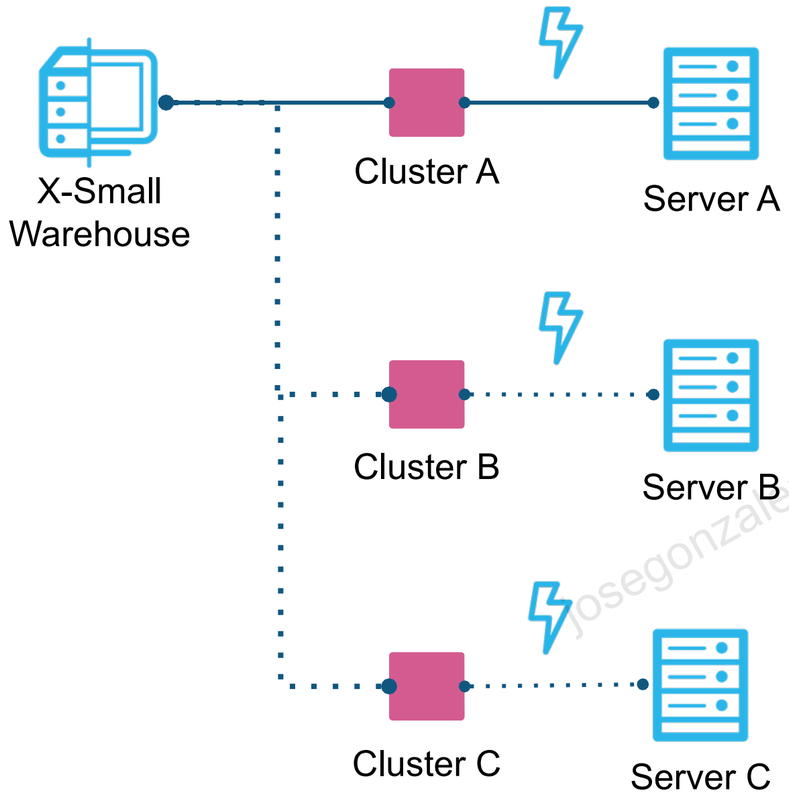




josegonzalezv@santander.com.mx - 11-SEP-2023 - © Snowflake 2023 do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

SCALING BACK POLICY: ECONOMY



T0	*Could the load on the least-busy cluster be redistributed to the other clusters?*
T0 + 1 minute	*Could the load on the least-busy cluster be redistributed to the other clusters?*

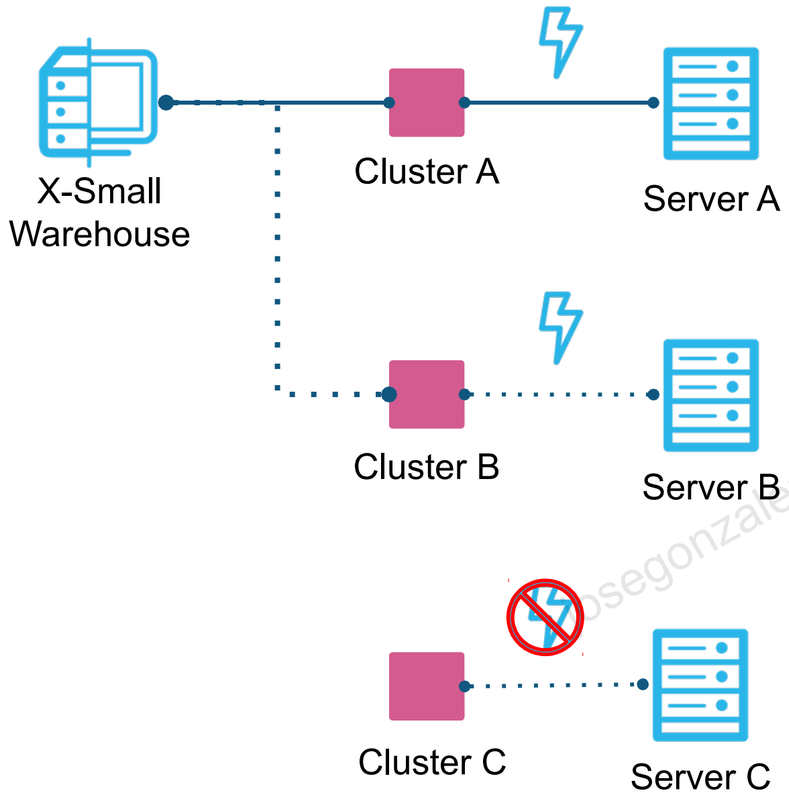
NO

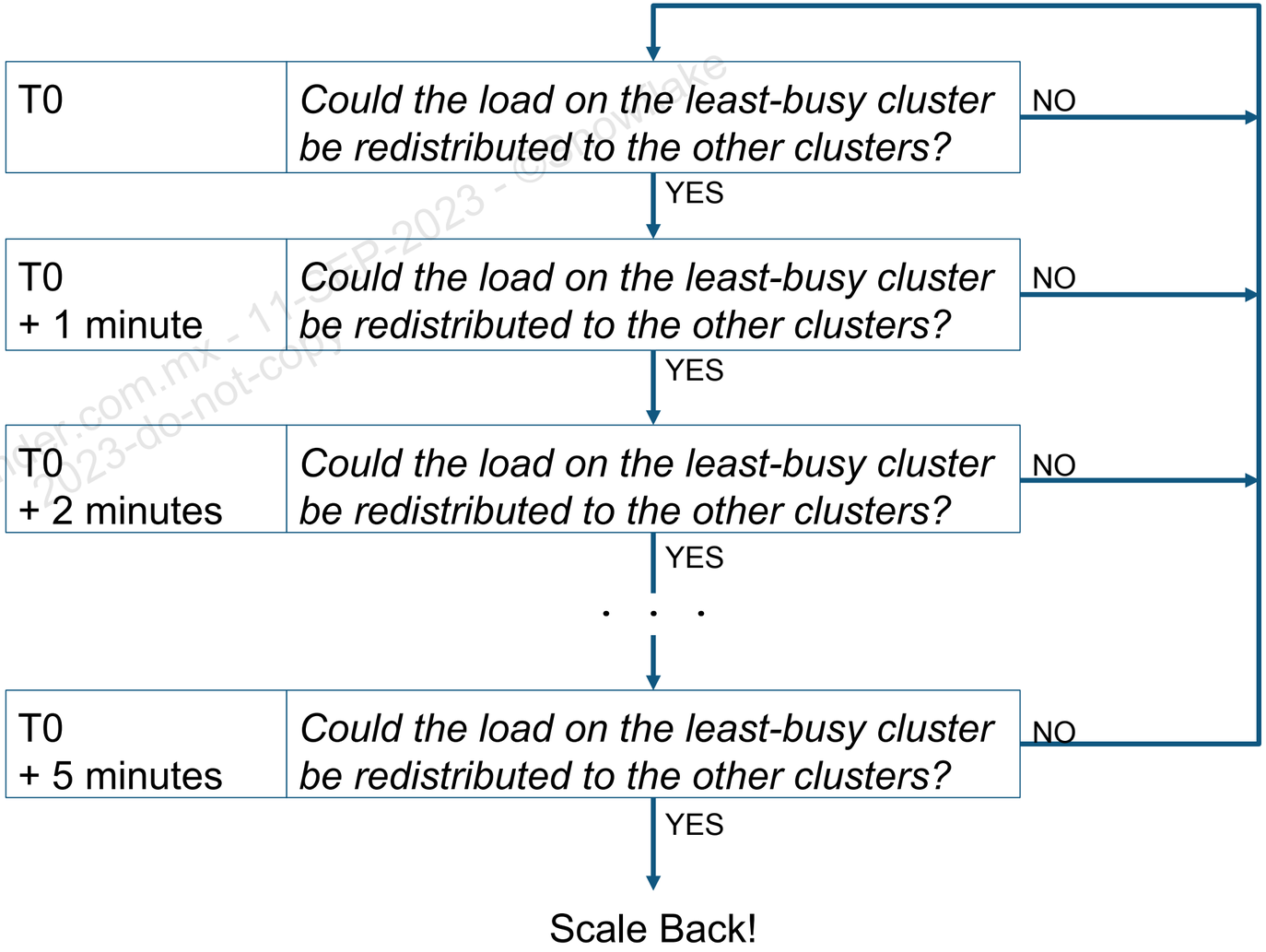
YES

josegonzalezv@santander.com.mx - 11-SEP-2023 - @snowflake
2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

SCALING BACK POLICY: ECONOMY

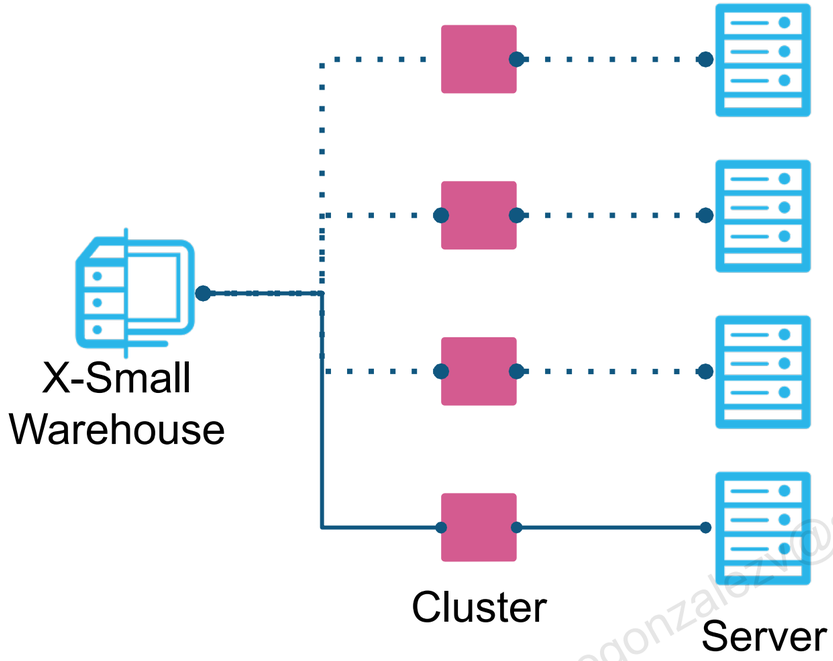


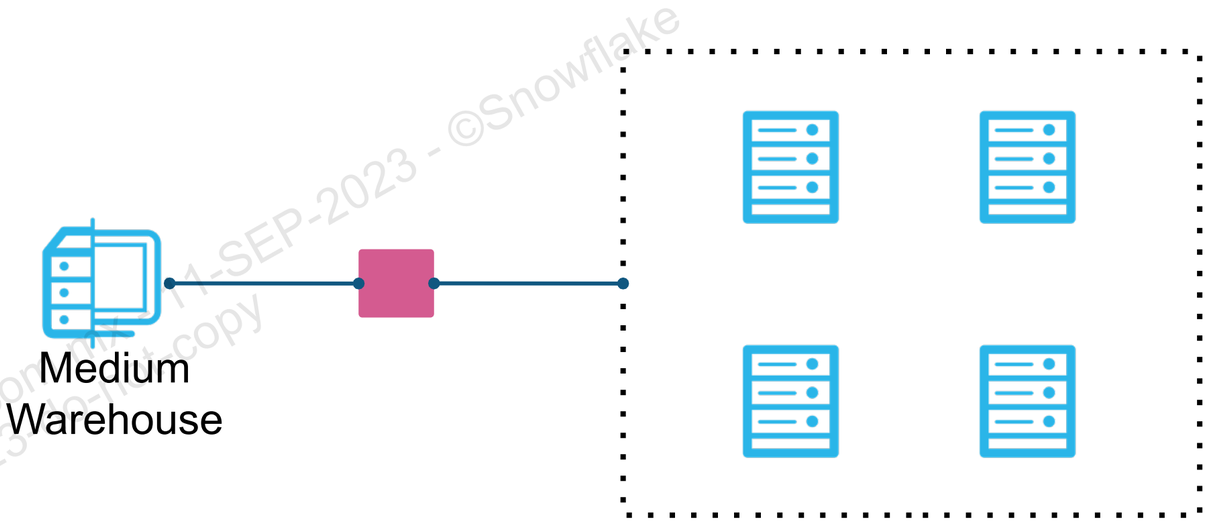


josegonzalezv@santander.com.mx - 11-SEP-2023 - @snowflake
2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

ARE THESE WAREHOUSES EQUIVALENT?





josegonzalezv@santander.com - 11-SEP-2023 - ©Snowflake
2023 do not copy

Auto-Scale, Multi-Cluster X-Small Warehouse

- Minimum of 1
- Maximum of 4

Medium Warehouse

- One cluster of 4 servers

© 2023 Snowflake Computing Inc. All Rights Reserved

KNOWLEDGE CHECK

Using the **standard** scaling policy, at what point will Snowflake start another cluster?

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

KNOWLEDGE CHECK

Using the **standard** scaling policy, at what point will Snowflake start another cluster?

Snowflake will start another cluster **as soon as something gets queued**.

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

KNOWLEDGE CHECK

When Snowflake decides to scale back, how long will it take to shut down one of the clusters?

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

KNOWLEDGE CHECK

When Snowflake decides to scale back, how long will it take to shut down one of the clusters?

It depends – Snowflake will begin the process of scaling back by making sure no more queries are allocated to the least-busy cluster. But the cluster cannot actually be suspended until **all its queries have completed**.

josegonzalezv@santander.com.mx - 11 SEP 2023 - @Snowflake
2023-do-not-copy

KNOWLEDGE CHECK

Using the **economy** scaling policy, how long might a query be queued?

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

KNOWLEDGE CHECK

Using the **economy** scaling policy, how long might a query be queued?

6 minutes

With the economy scaling policy, Snowflake will only start another cluster if the load on the current cluster(s) will keep them busy for more than 6 minutes.

josegonzalezv@santander.com.mx - 11 SEP 2023 - ©Snowflake
2023-do-not-copy

COURSE DISCUSSION

- Below are the results from running the same query against various warehouses sizes.
- Which warehouse size would you recommend, and why?

Warehouse Size	Query Time
Small	24m 30s
Medium	10m 00s
Large	5m 00s
XLarge	2m 40s
XXLarge	1m 40s

jose.orizaba@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

COURSE DISCUSSION

- Below are the results from running the same query against various warehouses sizes.
- Which warehouse size would you recommend, and why?

Warehouse Size	Query Time	Credits
Small	24m 30s	.882
Medium	10m 00s	.660
Large	5m 00s	.660
XLarge	2m 40s	.704
XXLarge	1m 40s	.890

jose.orizaba@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

COST GOVERNANCE

josegonzalezv@santander.com.mx 11-SEP-2023 - ©Snowflake
2023do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

MODULE AGENDA

- Overview
- Visibility
- Control
- Optimization

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

OVERVIEW

josegonzalezv@santander.com.br - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

FRAMEWORK

Visibility

Understand, attribute, and monitor costs

Control

Limit and control costs

Optimization

Identify and optimize costs

jose.gonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

TOOLS

Visibility

- Built-in usage charts
- Built-in views
- Custom dashboards

Control

- Configurable virtual warehouse sizes
- Auto suspend/resume
- Resource monitors
- Role-based access control

Optimization

- Continual platform improvements
- Query profile
- `QUERY_HISTORY` view
- `WAREHOUSE_LOAD` and other views

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023 do not copy

© 2023 Snowflake Computing Inc. All Rights Reserved

VISIBILITY

josegonzalezv@santander.com - 11-SEP-2023 - ©Snowflake
2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

WHY VISIBILITY?

You can't manage what you don't measure.

- Peter F. Drucker

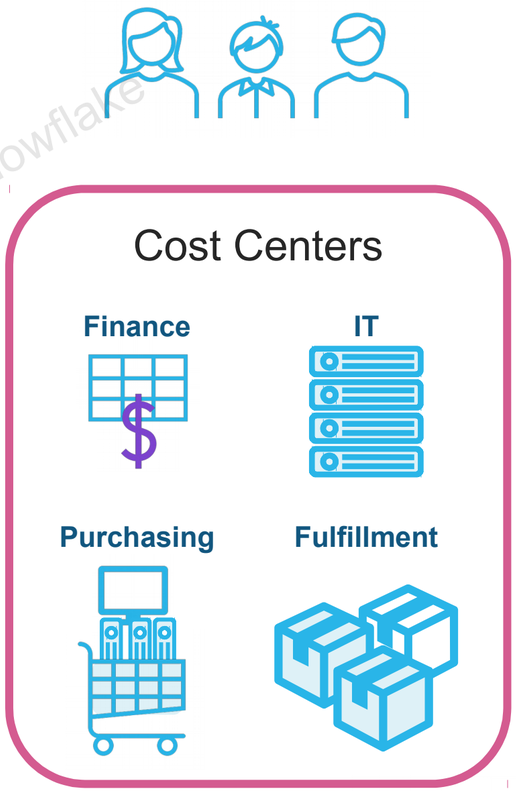
- Timely visibility into cost is critical to managing the elastic nature of the Snowflake Data Cloud
- Establish a baseline
- Avoid surprises
- Identify opportunities for control and optimization

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

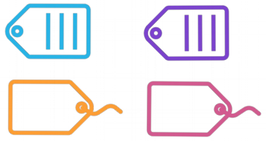
© 2023 Snowflake Computing Inc. All Rights Reserved

PROVIDE VISIBILITY INTO USAGE

- Identify the team that is going to manage and monitor resource consumption
- Establish your baseline requirements
- Identify the cost centers in your organization – who needs to be billed for consumption
- Tag objects to identify the cost center they belong to



Tag	Tag Value
Cost Center	PROD
Cost Center	REPORT
Cost Center	TEST

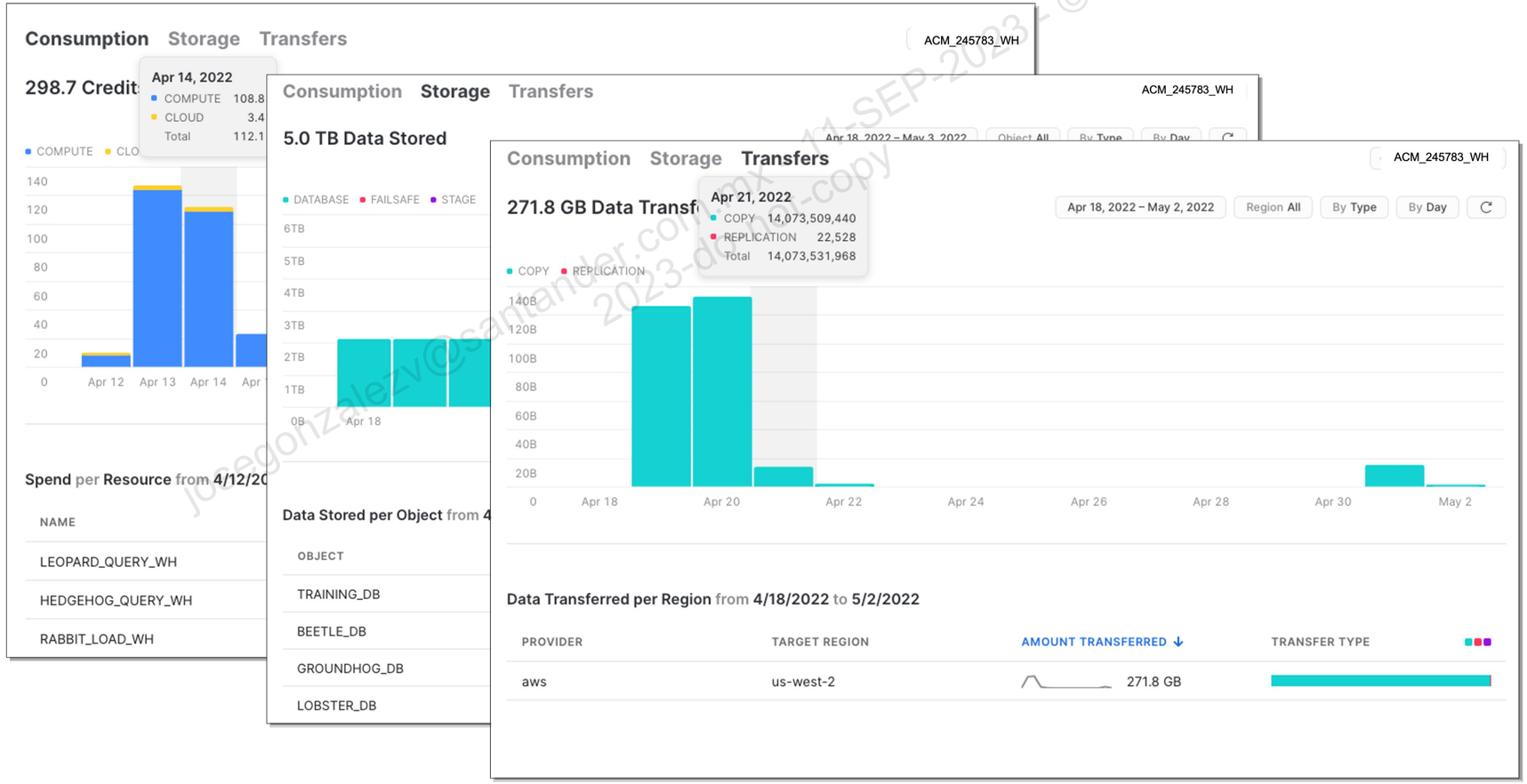


josegonzalez@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

BUILT-IN USAGE CHARTS

- Navigate to `Admin > Usage` (requires `ACCOUNTADMIN` or designated privileges)
- Toggle between compute, storage, and data transfer



© 2023 Snowflake Computing Inc. All Rights Reserved

BUILT-IN VIEWS AND TABLE FUNCTIONS

INFORMATION_SCHEMA

- Available in every database
- Available to anyone who has access to a database
- Holds information up to 6 months back
- Does not include dropped objects
- Low latency

ACCOUNT_USAGE

- Available in SNOWFLAKE database
- Available to ACCOUNTADMIN only by default (can be granted to other roles)
- Holds information up to 1 year back
- Includes dropped objects
- Latency up to 3 hours

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

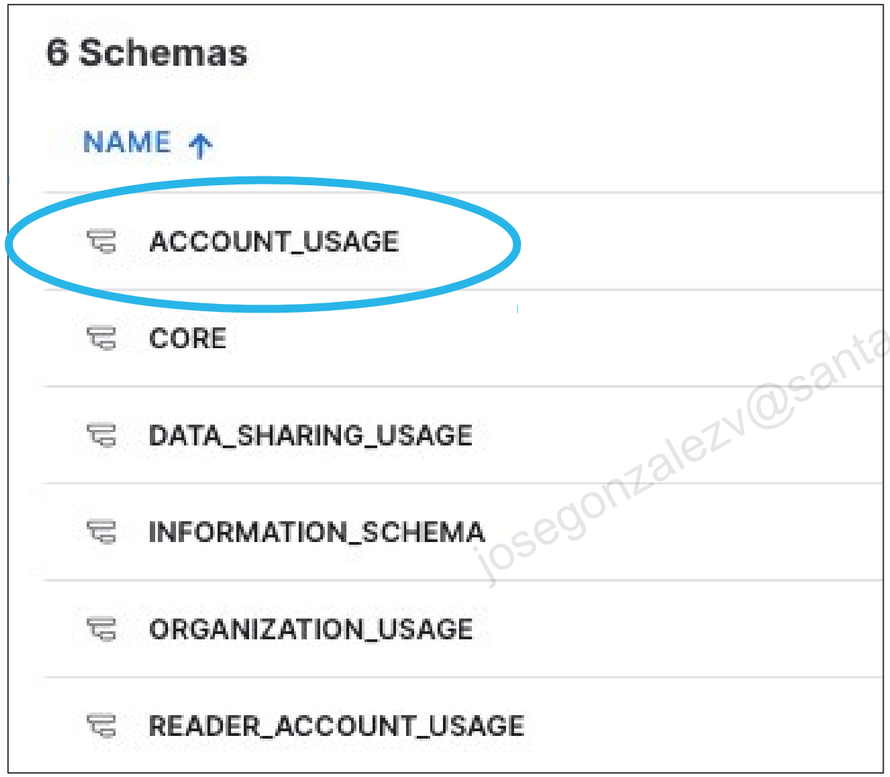
INFORMATION_SCHEMA

- ANSI standard views for databases and objects relevant to Snowflake
- Snowflake-specific views for non-standard objects that Snowflake supports
  - Stages, file formats, etc.
- Queries return only objects to which the current role has access
- Views are listed in the object browser
- Table functions are listed in the documentation

jose.gonzalezv@santander.com.mx - 11-SEP-2023 - Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

SNOWFLAKE DATABASE

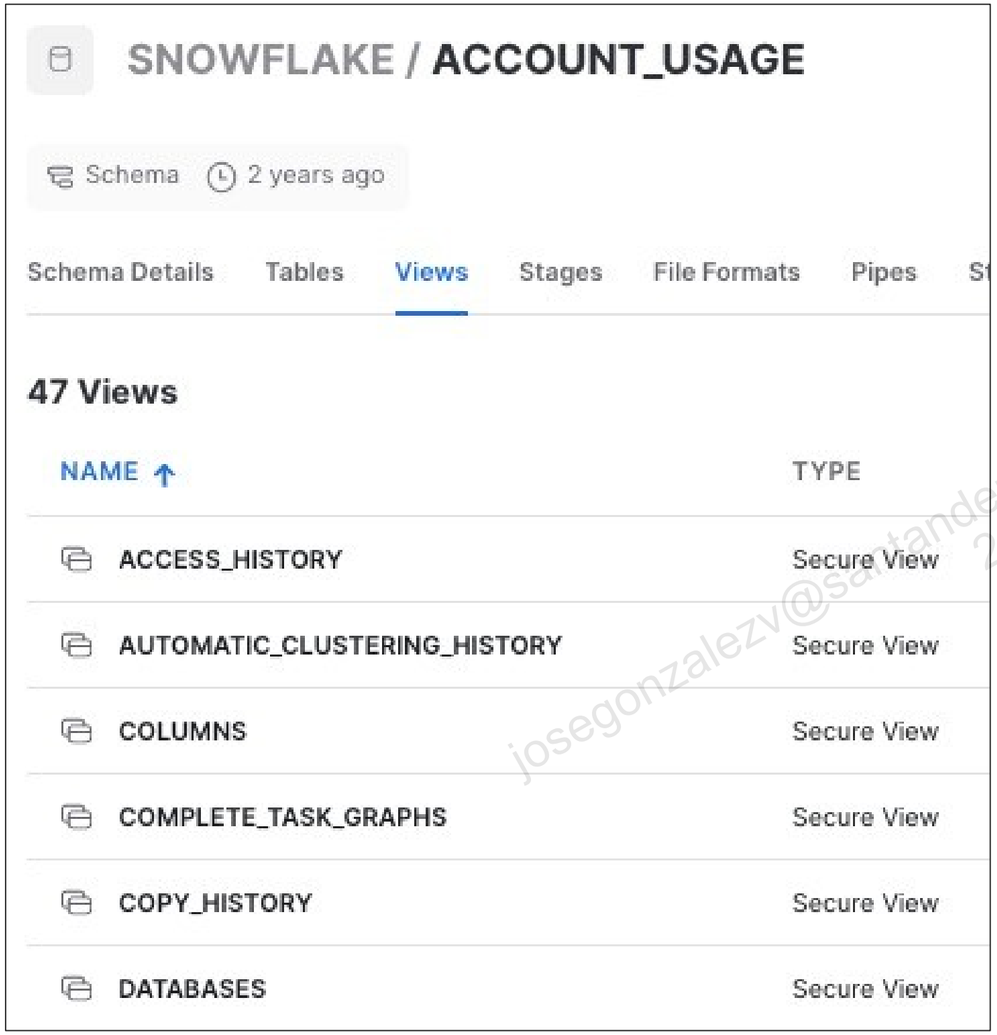


- Shared by Snowflake to all accounts
  - Example of secure data sharing
- Contains multiple schemas
- `ACCOUNT_USAGE` primarily used in account monitoring

josegonzalezv@santander.com.mx - 11-SEP-2023 @Snowflake
2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

ACCOUNT_USAGE



- Contains views that display object metadata and usage metrics
- Generally, `ACCOUNT_USAGE` mirrors the corresponding views and table functions in `INFORMATION_SCHEMA`, with some differences

josegonzalezv@santander.com.mx - 11-SEP-2023 - @Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

INFORMATION_SCHEMA EXAMPLE

Check average queries queued, by warehouse

```
SELECT warehouse_name, AVG(avg_queued_load) AS avg_queued
FROM TABLE(information_schema.warehouse_load_history())
GROUP BY warehouse_name
ORDER BY warehouse_name;
```

WAREHOUSE_NAME	AVG_QUEUED
PROD_SM	1.15
PROD_LG	0.27
REPORT_MC_SM	2.47
REPORT_MC_MED	0
TEST01	0.8
TEST01	12.45

josegonzalezv@santander.com.mx - 11-SEP-2023 - © Snowflake
2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

ACCOUNT_USAGE EXAMPLE

Check for warehouses with insufficient memory

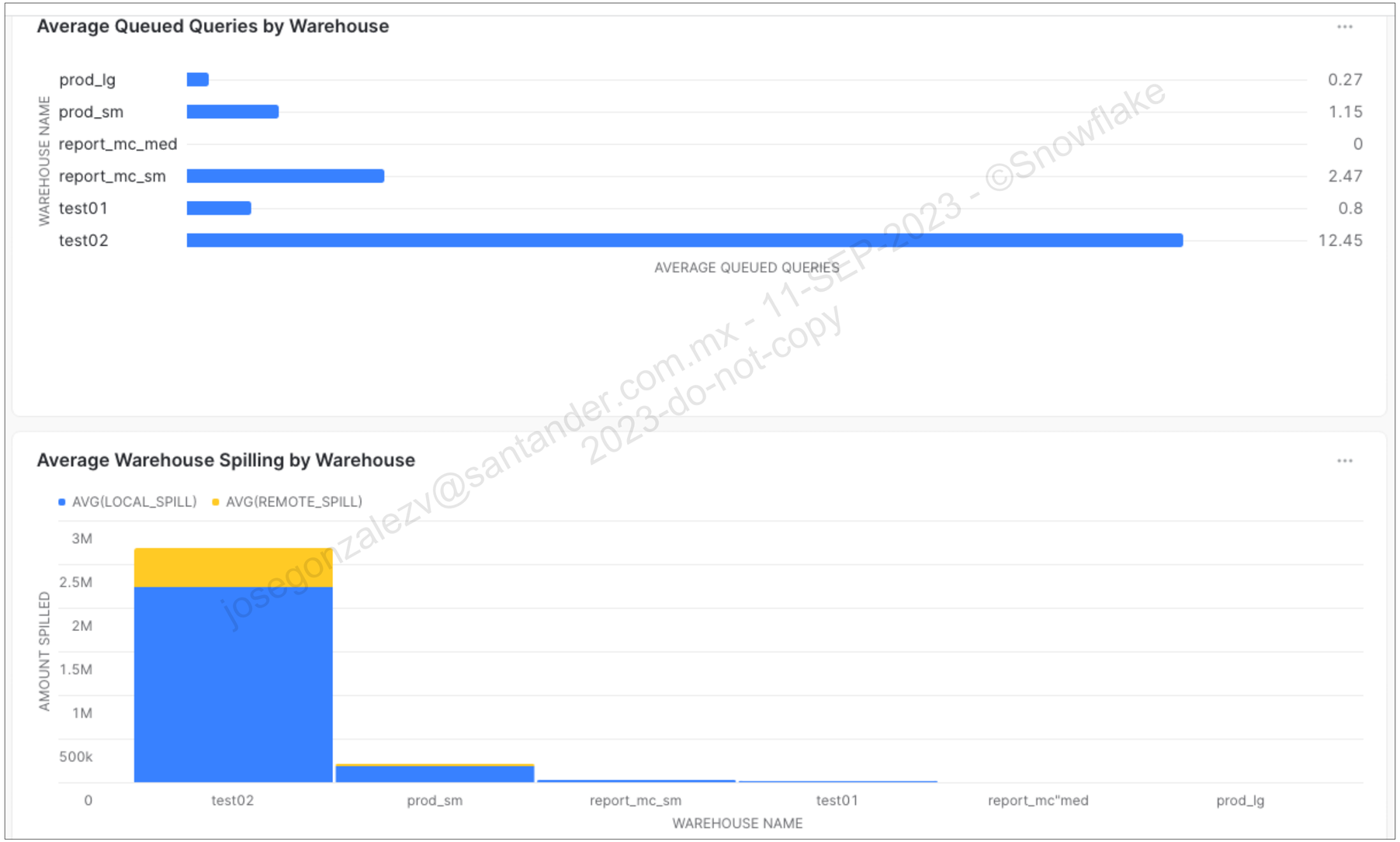
```
SELECT
   warehouse_name AS wh,
   AVG(bytes_spilled_to_local_storage)  AS local_spill,
   AVG(bytes_spilled_to_remote_storage) AS remote_spill
FROM account_usage.query_history
GROUP BY warehouse_name;
```

WH	LOCAL_SPILL	REMOTE_SPILL
PROD_SM	188456	22567
PROD_LG	0	0
REPORT_MC_SM	22433	0
REPORT_MC_MED	1442	0
TEST01	15047	1228
TEST02	2234556	442335

josegonzalezv@santander.com.mx - 14-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

CUSTOM DASHBOARDS



josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

KNOWLEDGE CHECK

Which schema has longer latency for updating its views: `ACCOUNT_USAGE` or `INFORMATION_SCHEMA`?

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

KNOWLEDGE CHECK

Which schema has longer latency for updating its views: `ACCOUNT_USAGE` or `INFORMATION_SCHEMA`?

ACCOUNT_USAGE

`INFORMATION_SCHEMA` generally has very little latency. `ACCOUNT_USAGE` could have latency of several hours.

josegonzalezv@santander.com.mx - 11 SEP 2023
2023-do-not-copy ©Snowflake

LAB EXERCISE: 23

Introduction to Monitoring Usage and Billing

20 minutes

Learning Objectives:

- How to monitor Usage and Billing with the ACCOUNT_USAGE schema
- How to determine warehouses without resource monitors
- How to determine the most expensive queries from the last 30 days
- How to determine the top 10 spillers to remote storage

josegonzalezv@santander.com.mx 11-Oct-2023 - ©Snowflake
2023-document-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

LAB EXERCISE: 24

Challenge Lab: Monitoring Billing and Usage

30 minutes

Learning Objectives:

- Create a dashboard
- Add tiles to a dashboard
- Add filters to a dashboard
- Create a bar chart
- Write basic queries using the QUERY_HISTORY view in the SNOWFLAKE.ACCOUNT_USAGE schema
- Apply an existing custom filter to a query in a dashboard

Note: A solution is presented on the last page of this lab in the workbook.

jaime.gorzowski@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

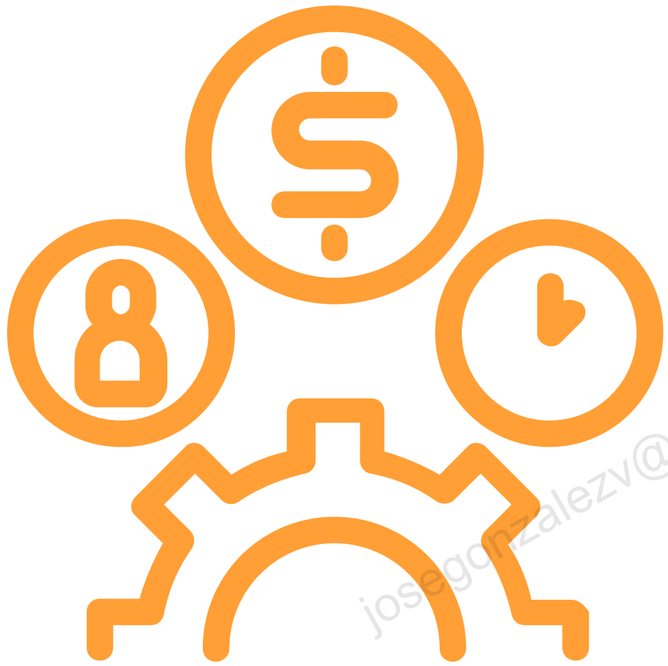
© 2023 Snowflake Computing Inc. All Rights Reserved

CONTROL

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

CONTROL



- Configurable virtual warehouse sizes
- Auto suspend and auto resume
- Resource monitors
- Role-based access control

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

SET UP CONTROLS

- Review settings for your virtual warehouses
- Review query timeout limits
- Review resource monitor settings and establish limits for specific workloads

Virtual Warehouses

- Size
- Minimum clusters
- Maximum clusters
- Auto suspend
- Auto resume

Resource Monitor

- Credit quota
- Monitor level
- Schedule
- Actions

Queries

- Queued timeout
- Query statement timeout

jose.gonzalezv@santander.com.mx - 11-SEP-2023 - @Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

VIRTUAL WAREHOUSE CONTROLS

- Auto suspend turns off a warehouse after a specified amount of idle time
  - Tailor auto-suspend times to patterns for each workload
  - Most users cannot suspend warehouses, so auto-suspend is critical
- Auto resume restarts a virtual warehouse when new queries are received
- Multi-cluster auto-scaling enables Snowflake to add and remove clusters to dynamically manage high-concurrency workloads
- Restrict who can resize virtual warehouses
- Have users running similar queries share a virtual warehouse
  - Makes better use of the data cache

juanconzalezv@santander.com.mx 11-SEP-2023 ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

SET QUERY LIMITS

- Query limits can be set at the account, warehouse, or session level
- Set parameters according to needs of each workload

PARAMETER	DEFAULT	DESCRIPTION
STATEMENT_TIMEOUT_IN_SECONDS	172800	Timeout in seconds for statements: statements are automatically canceled if they run for longer; if set to zero, max value (604800) is enforced.

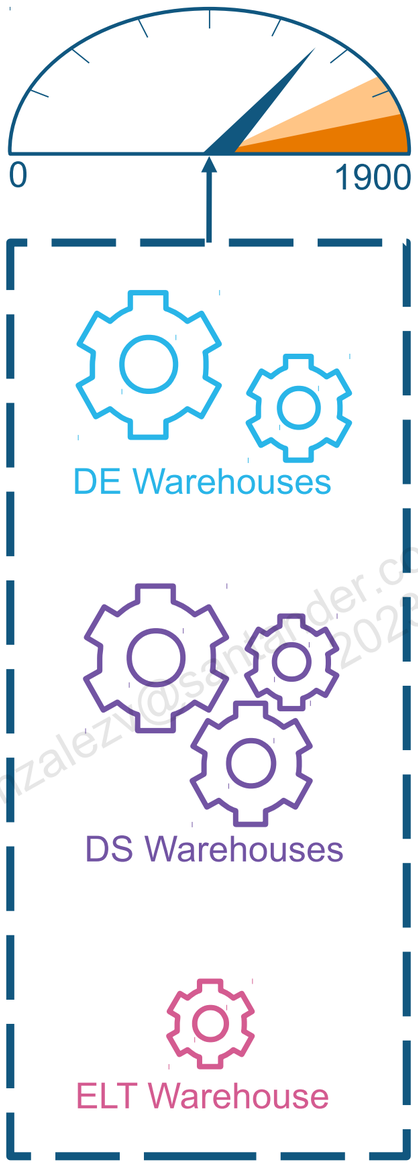
PARAMETER	DEFAULT	DESCRIPTION
STATEMENT_QUEUED_TIMEOUT_IN_SECONDS	0	Timeout in seconds for queued statements: statements will automatically be canceled if they are queued on a warehouse for longer than this amount of time; disabled if set to zero.

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake
2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

RESOURCE MONITORS

Account Monitor
max 1900 credits

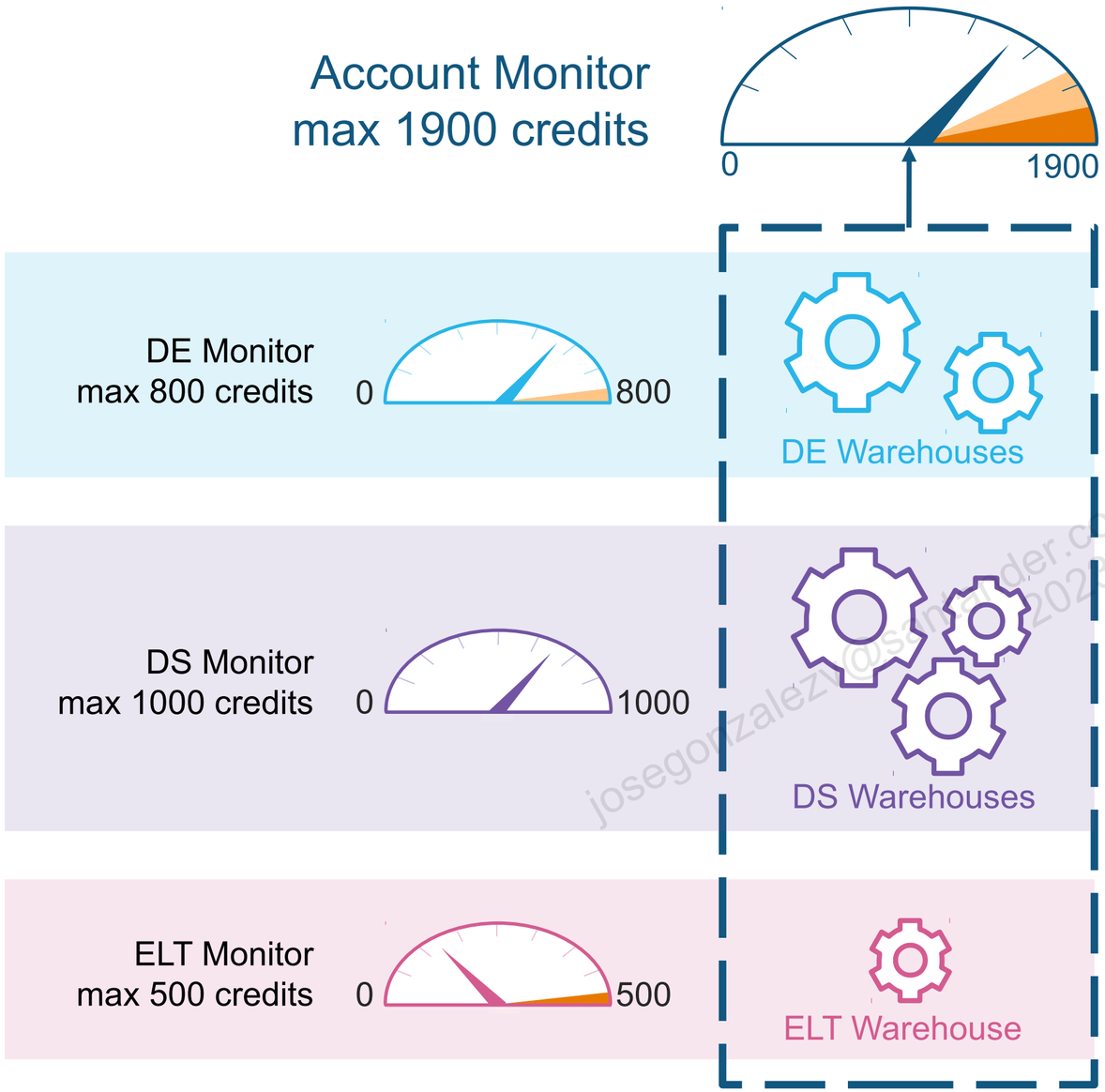


- Help control costs and avoid unexpected warehouse credit usage
- Tracks warehouse use **<u>and</u>** cloud computing
- Can trigger various actions
  - Sending alert notifications
  - Suspending the warehouse(s)

josegonzalez@sandbox.com.mx - 11-SEP-2023 @Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

RESOURCE MONITORS



- Help control costs and avoid unexpected warehouse credit usage
- Tracks warehouse use **and** cloud computing
- Can trigger various actions
  - Sending alert notifications
  - Suspending the warehouse(s)
- Set at the account or warehouse level
- By default, only `ACCOUNTADMIN` can create

josegonzalezv@samsalder.com.mx - 11-SEP-2023 @Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

RESOURCE MONITOR ATTRIBUTES

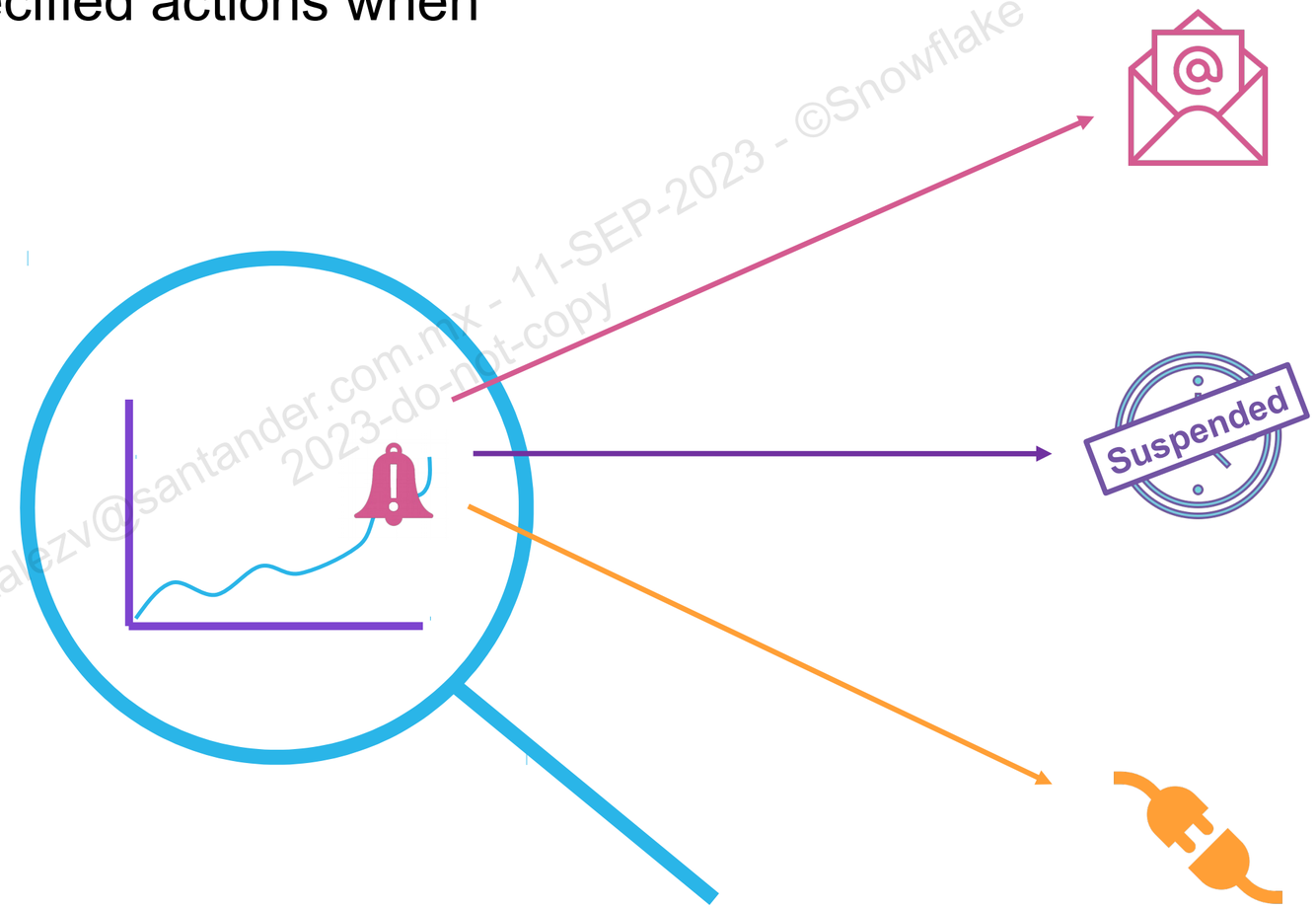
- **Name**: Descriptive name given to the monitor
- **Credit quota**: Number of credits granted to the resource monitor each cycle
- **Monitor type**: Account or warehouse
- **Range (cycle)**: How frequently the resource monitor resets (daily, weekly, etc.)
- **Schedule**: When the monitor should start (and optionally end) tracking
- **Actions** to take at various thresholds:
  - Not intended to strictly control usage to the second
  - Quota may be exceeded before action is triggered

jesegonzale@santander.com.mx - 11-SEP-2025 ©Snowflake 2025 do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

RESOURCE MONITOR ACTIONS

- Resource monitors take specified actions when thresholds are reached
  - Notify
  - Suspend and notify
  - Suspend immediate



josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

RESOURCE MONITOR NOTIFICATIONS

- `ACCOUNTADMIN` members alerted if they have email configured and notifications enabled
  - Must verify email after setting
- `NOTIFY_USERS` list specifies non-admins who should be notified
- Users in the list must have a verified email address, and have notifications enabled

jsegonzalez@santander.com - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

```
CREATE RESOURCE MONITOR <name>
    …
    NOTIFY_USERS = (User1, User2, "Space in Name");
```

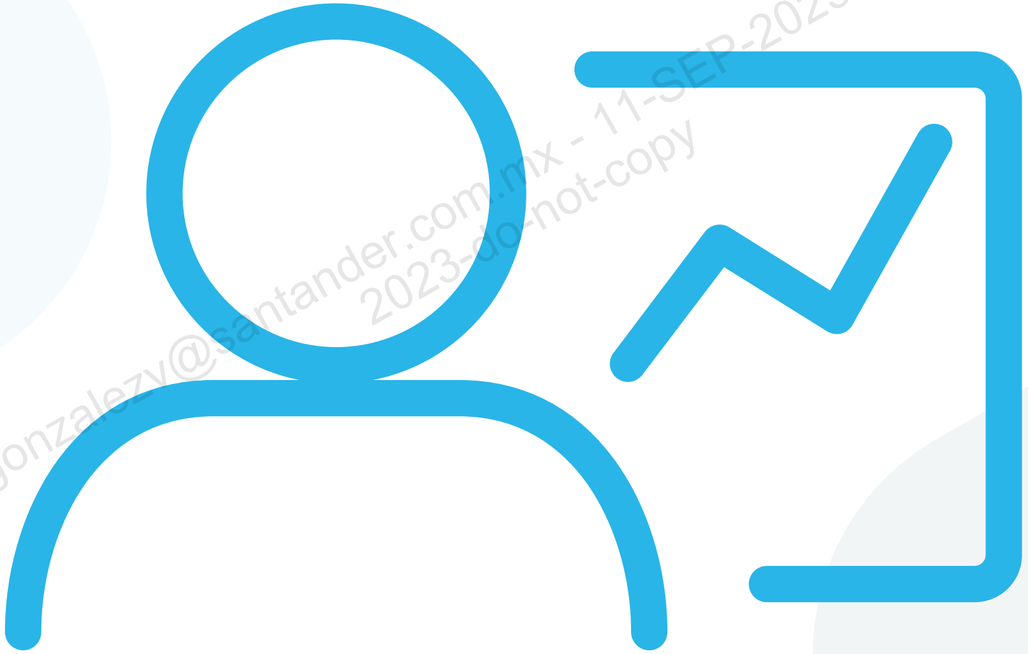
Use double quotes if there are spaces in a username

© 2023 Snowflake Computing Inc. All Rights Reserved

INSTRUCTOR DEMO

Set Up Resource Monitors

10 minutes



josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

ADDITIONAL INFORMATION

- More information, and additional exercises, is available in Snowflake's on-demand Cost Governance course

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

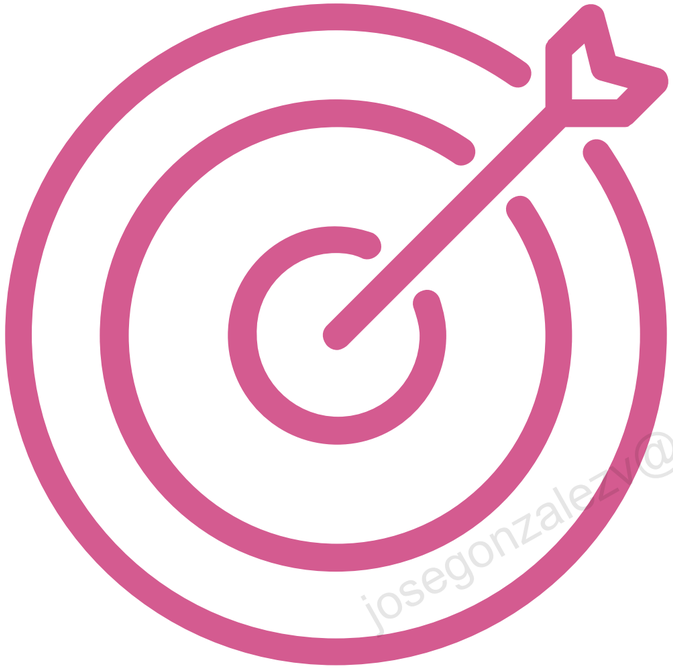
© 2023 Snowflake Computing Inc. All Rights Reserved

OPTIMIZATION

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake
2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

OPTIMIZATION

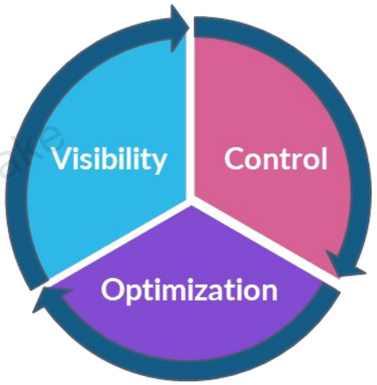


- Snowflake is continually improving its platform to provide faster and more efficient compute
- Use the query profile to optimize individual queries
- Use views in `ACCOUNT_USAGE` and `INFORMATION_SCHEMA` to evaluate warehouse efficiency and make changes to improve cost/performance

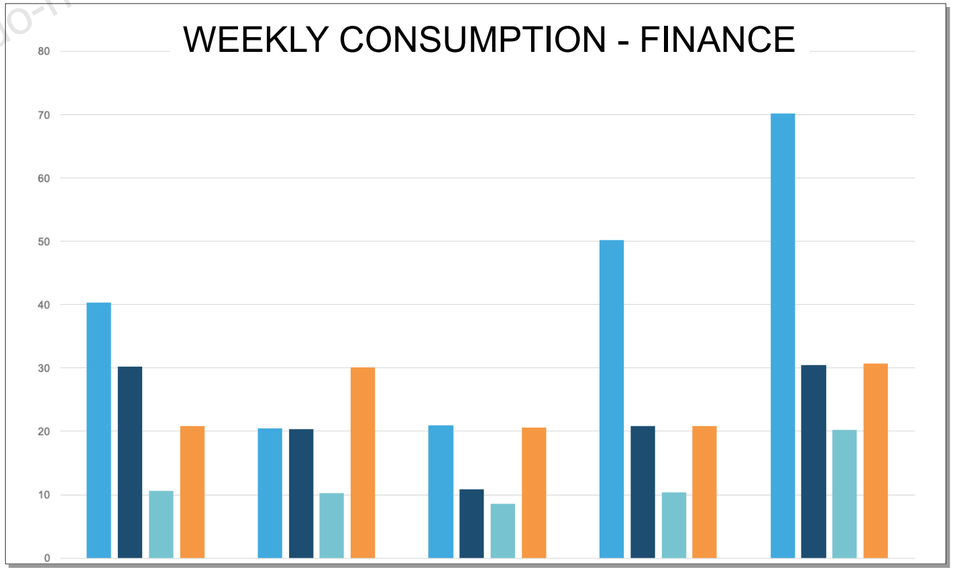
josegonzalez@santander.com.mx - 11-SEP-2023 - ©Snowflake
2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

PREPARE TO OPTIMIZE



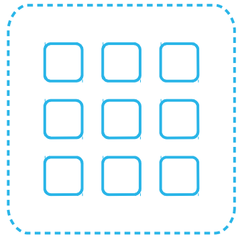
- Identify anomalies in consumption
- Loop back to visibility and control
- Work with the governance team to identify and improve:
  - Long-running queries
  - Warehouse sizing and usage
  - Storage usage
  - Data transferred from your account



josegonzalezv@santander.com.mx - 4 SEP-2023 - ©Snowflake 2023-do-not-copy

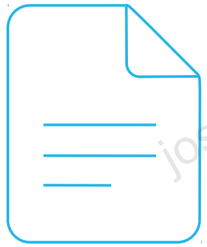
© 2023 Snowflake Computing Inc. All Rights Reserved

STORAGE USAGE



Database Tables

(Active, Time Travel, Fail-safe)



Internal Stages

(User, Table, Named)

- Billed monthly
  - Calculated based on daily bytes of all data stored each day in account
- Cost varies based on type of account, cloud provider, and region
- Storage costs include:
  - Active storage – tables and objects
  - Time travel storage
  - Fail-safe storage
  - Internal stages

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

USER-MANAGED COMPUTE

Warehouse Size	Credits / Hour	Credits / Second
X-Small	1	0.0003
Small	2	0.0006
Medium	4	0.0011
Large	8	0.0022
X-Large	16	0.0044
2X-Large	32	0.0089
3X-Large	64	0.0178
4X-Large	128	0.0356

- Only billed when running
- Billed per-second, with a 60-second minimum
- Can be limited and controlled via resource monitors
- Usage can be viewed via the UI or Information Schema/Account Usage

josegonzalezy@santander.com.mx - 11-SEP-2023 - ©snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

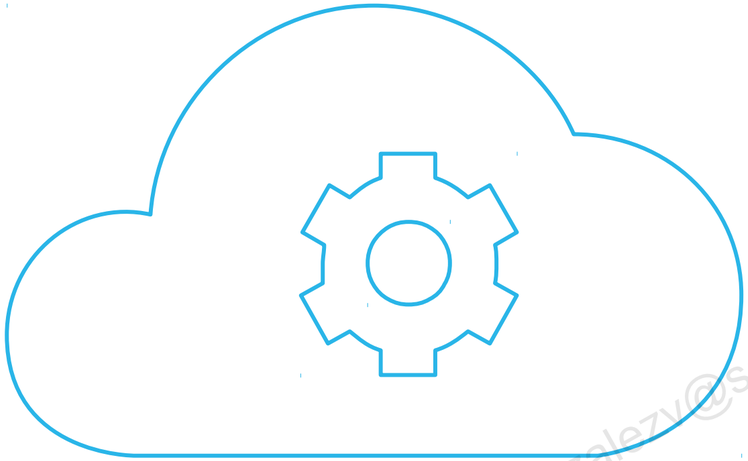
SERVERLESS FEATURES

- Billed per second, per core
- No virtual warehouses needed
- Serverless features include:
  - Database replication
  - Materialized view refresh
  - Tasks (serverless option)
  - Others
- Some serverless features have a multiplier for compute costs, so monitor those features closely!

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake
2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

CLOUD SERVICES COMPUTE

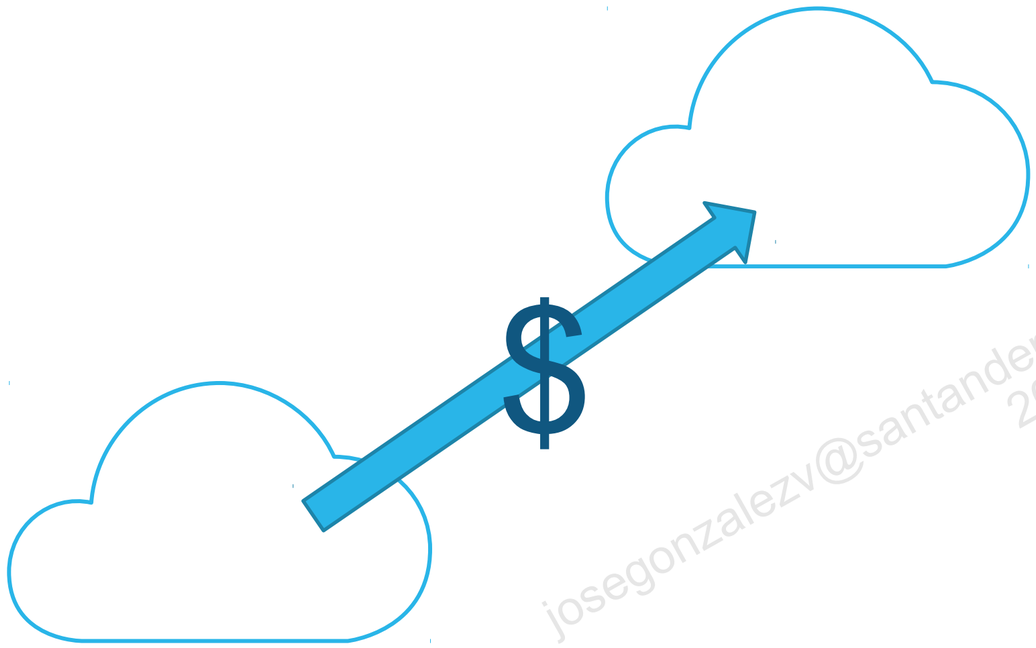


- Only charged for cloud services compute that exceeds 10% of total compute usage
- Operations like:
  - `CREATE` object
  - `TRUNCATE` table
  - `DELETE` object
  - Cache lookups

josegonzalezv@santander.com.mx - 11-SEP-2023 @Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

DATA TRANSFER CHARGES



- Database replication
  - Initial replication
  - Refreshes
- Unloading data to external stages in another region/cloud provider
- Varies based on source and destination
  - Region to region transfer within a cloud provider are less expensive than transfers from one cloud provider to another

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake
2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

FIND OPTIMIZATION OPPORTUNITIES

Virtual Warehouse-Related

- Explore adjusting warehouse size to reduce contention or spilling
- Analyze cache usage to tune the auto-suspend time for your warehouses
- Compare warehouse usage across a defined timeframe to see if adjustments are required

Query-Related

- Look for long-running queries
- Find queries that spill
- Identify users who most frequently use full table scans in their queries
- Identify users who process large amounts of data

jose.gonzalez@santander.com.mx - 11-SEP-2023 - @Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

OPTIMIZE COMPUTE COSTS

- Right-size your virtual warehouses
- Monitor warehouse efficiency on a regular basis
- Configure **resource monitors** to take action when compute exceeds given limits
- Use multi-cluster warehouses for high concurrency
- Set warehouses to auto-suspend and auto-resume
  - Different warehouses may need different suspend times
- Use role-based access control to limit who can monitor or modify virtual warehouses

josegonzalezv@santander.com.mx - 14-SEP-2023 - ©Snowflake
2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

ACCOUNT_USAGE.QUERY_HISTORY

- Identify inefficient warehouses when aggregated by warehouse:
  - WAREHOUSE_NAME
  - WAREHOUSE_SIZE
  - PERCENTAGE_SCANNED_FROM_CACHE
  - PARTITIONS_SCANNED vs. PARTITIONS_TOTAL
  - BYTES_SPILLED_TO_LOCAL_STORAGE
  - BYTES_SPILLED_TO_REMOTE_STORAGE
  - QUEUED_OVERLOAD_TIME
  - CREDITS_USED_CLOUD_SERVICES
  - QUERY_LOAD_PERCENT

jose.gonzalezv@santander.com.mx 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

OPTIMIZE STORAGE COSTS

- Use temporary or transient tables for short-lived data
- Use transient tables for permanent high-churn data, and limit time travel retention period
- Remove transient objects when they are no longer needed
- Remove files in stages
  - Or use the `PURGE` option with `COPY INTO` with internal stages
- Use appropriate time travel settings, at the right level (table, schema, or database)
- Batch single-statement DML operations into a single transaction
- Set up a storage capacity plan, rather than using on-demand

jose.gonzalezv@santander.com.mx - 14-SEP-2023 - @Snowflake
2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

KNOWLEDGE CHECK

What are some things you should do to maximize the efficiency of your virtual warehouses?

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

KNOWLEDGE CHECK

What are some things you should do to maximize the efficiency of your virtual warehouses?

- Make sure they are not too large OR too small.
- Make sure that roles that are sharing a virtual warehouse are querying similar data.
- Set warehouses to auto_suspend, but not too quickly – you don't want to lose the benefit of the data cache.
- Use multi-cluster warehouses for high-concurrency workloads.

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

KNOWLEDGE CHECK

Which way to pay for storage is less expensive: on-demand, or capacity?

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

KNOWLEDGE CHECK

Which way to pay for storage is less expensive: on-demand, or capacity?

Capacity

With on-demand, you use whatever storage you need and we bill you for it at the end of the month.

With a capacity contract, you pre-pay for storage and Snowflake will bill against that pre-paid amount. A capacity contract is typically about 60-70% of the cost of on-demand storage.

josegonzalezv@santander.com.mx - 11 SEP 2025 - ©Snowflake
2023-do-not-copy

RELEASES AND PREVIEW FEATURES

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake
2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

MODULE AGENDA

- Releases
- Preview Features

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

RELEASES

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake
2023-do not copy

© 2023 Snowflake Computing Inc. All Rights Reserved

RELEASE SCHEDULE

- Snowflake generally releases weekly, with releases typically completed on Fridays
  - Some weeks, typically around the end of the year, releases are not rolled out
- In the case of bug fixes/hotfixes, releases will be rolled out as needed
- Releases are seamless to the customer--they require no downtime, but take effect as soon as they are committed

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake
2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

RELEASE NOTES

- Snowflake publishes release notes detailing new functionality and behavior changes to existing functionality; these notes are available via two different locations and formats depending on your needs:
  - Do you want the detail behind each individual release and behavior change? See the Lodge Community Announcements
  - Do you want a summary, by month, of new features and behavior changes? See the Documentation Release Notes

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

PENDING BEHAVIOR CHANGES

- Snowflake publishes upcoming behavior changes in advance of their release so you may adequately prepare
- The Pending Behavior Change Log will detail the upcoming change and its release date: https://support.snowflake.net/s/article/Pending-Behavior-Change-Log
- This Log will also detail recently released behavior changes

josegonzalezv@santander.com.mx - SEP 2023 - ©Snowflake 2023 - do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

PREVIEW FEATURES

josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake
2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

WHAT IS A PREVIEW FEATURE?

- Implemented and tested in Snowflake
- Full usability and corner-case handling may not be complete
- Use of preview features are not warranted against defects that may produce undesired results
  - Behavior may change between Preview and Release
- Not recommended for production use

josegonzalezv@santander.com.mx - 11-SEP-2023 ©Snowflake
2023 do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

PREVIEW TYPES

- Private Preview / On Request
  - Typically invite-only at first
  - Generally, in the early stages of preview
  - Work with your account team if access is required
- Public Preview / Open
  - Enabled by default for all accounts
  - Free to use, though still carry some risk and are not for full production use

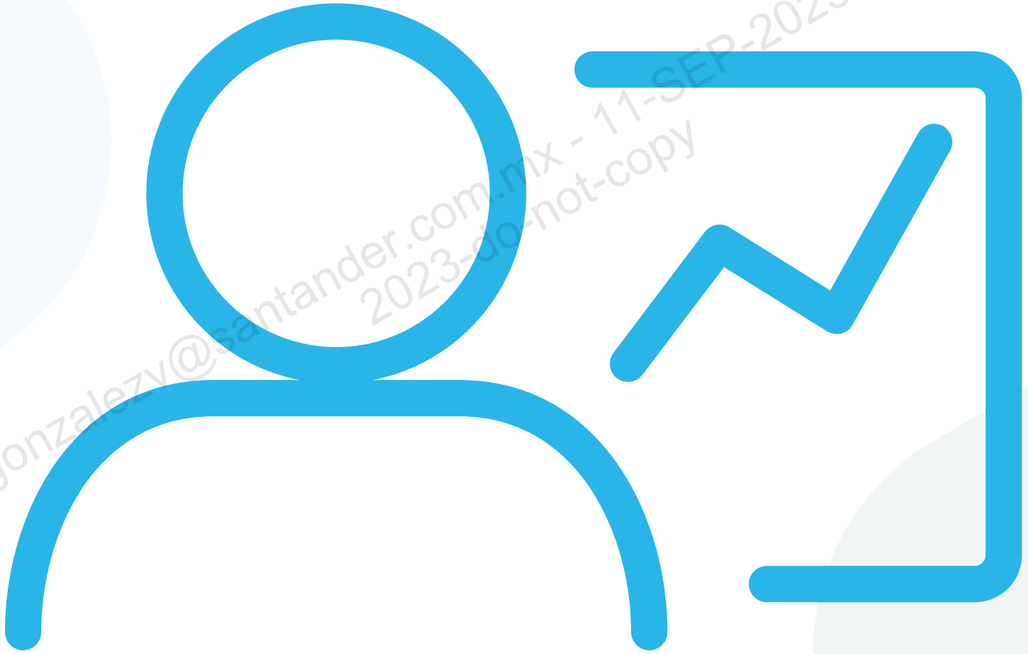
josegonzalez@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

INSTRUCTOR DEMO

Current Preview Features

15 minutes



josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

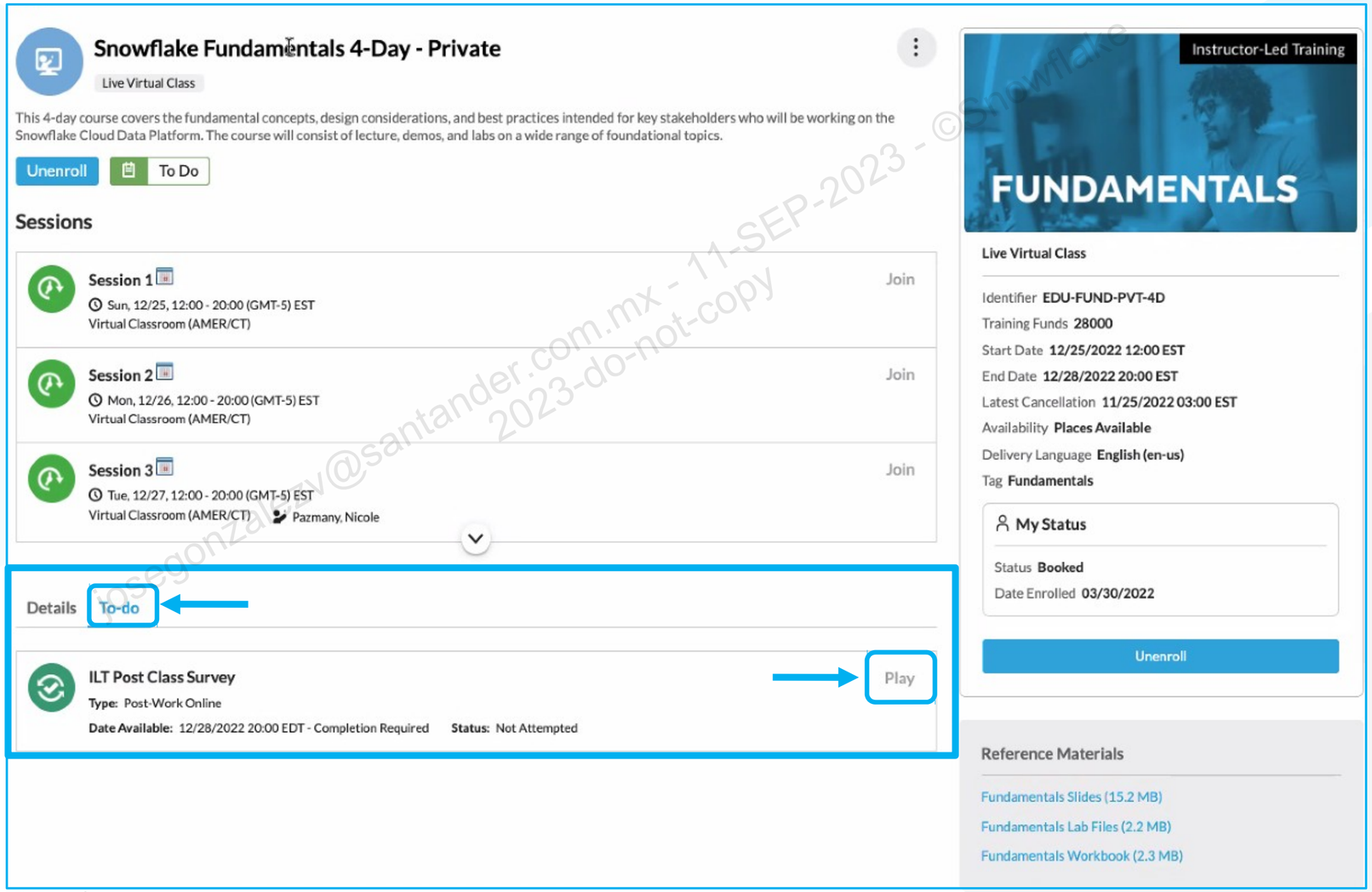
© 2023 Snowflake Computing Inc. All Rights Reserved

LAB EXERCISE

Complete the Course Survey

10 minutes

1. Navigate to the course page
2. Click “To-do” in the section at the bottom of the page
3. Click “Play”



josegonzalezv@santander.com.mx - 11-SEP-2023 - ©Snowflake 2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved

SNOWPRO CERTIFICATIONS

SNOWFLAKE TECHNICAL CERTIFICATIONS

josegonzalezv@santander.com.mx - 14-SEP-2024 - 2025-do-not-copy - ©Snowflake

© 2023 Snowflake Computing Inc. All Rights Reserved

THANK YOU

josegonzalezv@santander.com.mx - 14-SEP-2023 - ©Snowflake
2023-do-not-copy

© 2023 Snowflake Computing Inc. All Rights Reserved