Snowflake managed

Data storage : Aws s3 azure blob

Compute : Ec2 ---> snowflake managed

User managed ;

External stage :

aws s3 azure blob gcp

Recap : semi structure data :

Json

Parquet

Avro

Orc

Xml

===============variant ===============

Max -16 mb -compressed

Variant ===c1 : pid

Csv ⇒ encoded file format

Compression codec =gzip /deflate /

===========storage optimization ===============

Parquet /orc ⇒ columnar / metadata

Src:customer:name ===>

====================================

External table :

Schema design ;

1. Upfront schema =create/load
2. Schema on read
3. Schema on write

Snowtable

1 . create table

Local ===put =stage =copy ====snowtable

Query ====snowflake store--remote

2. External stage ;

Create external table ;

Aws s3 → all data ----------read ------ snow ext table

Query -----> read ------> external source

Delete ext table ===> no impact on data

=====data type ===> value : variant

User defined column on external table ==>virtual

=======================================

Data loading : internal/stage =========snowflake

External table =======data

First data :

lti935/employee/emp.csv ---copy ---emp

Emp1.csv

Emp2.csv =============snowflake --copy

Automate --data loading

File watcher :

CDP : continuos data protection -TT/FS

CDP --continuous data pipeline

**Snowpipes :**

Automation of manual copy process from external stage to snowflake table

**Iam user --trust --snowflake user arn**

Snowpipes

Event notification at your bucket/external stage

External stage → storage integration and aws key id and secret

=================share

Full account ==existing snowflake -billing is individual the account

Account name :

Reader account ⇒ new snowflake ---billing included in their source provider

Reader get url and login from share