Recap : Architecture

Micropartition : immutable

header : range , distinct ,null

Zonemap :

Column Pruning

Cluster key --sorted ---> minimize no of mp to be scanned

Snowflake : automatic clustering ==reclustering ---expensive

100 mp ---cluster ---> new 100 mp

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

Compute /query processing : warehouse

Node : multicluster

Shared nothing

Dattaaaaaaaaaaaaaaaaaaaaaaaaa : remote

node 1 node 2 node 3 : local /ssd

MPP :

AWS : EC2 ---> Virtual machine

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

Stages :

Local ===== put == stage ==copy == snowflake table

Compressed optimized columnar micro partitioned encrypted

Stages

1. User stage - list @~; single user/multiple tables
2. Table stage ===> load data into same tablename

Multiple users/ single table

drop/delete/renaming -not posible

Internal stage : multiple user/multiple tables

1. Internal named stage -snowflake managed
2. External named stage

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

COPY INTO "LTIDB"."LTISCHEMA"."EMP" FROM @%emp FILE\_FORMAT = '"LTIDB"."LTISCHEMA"."CSVFILEFORM"'

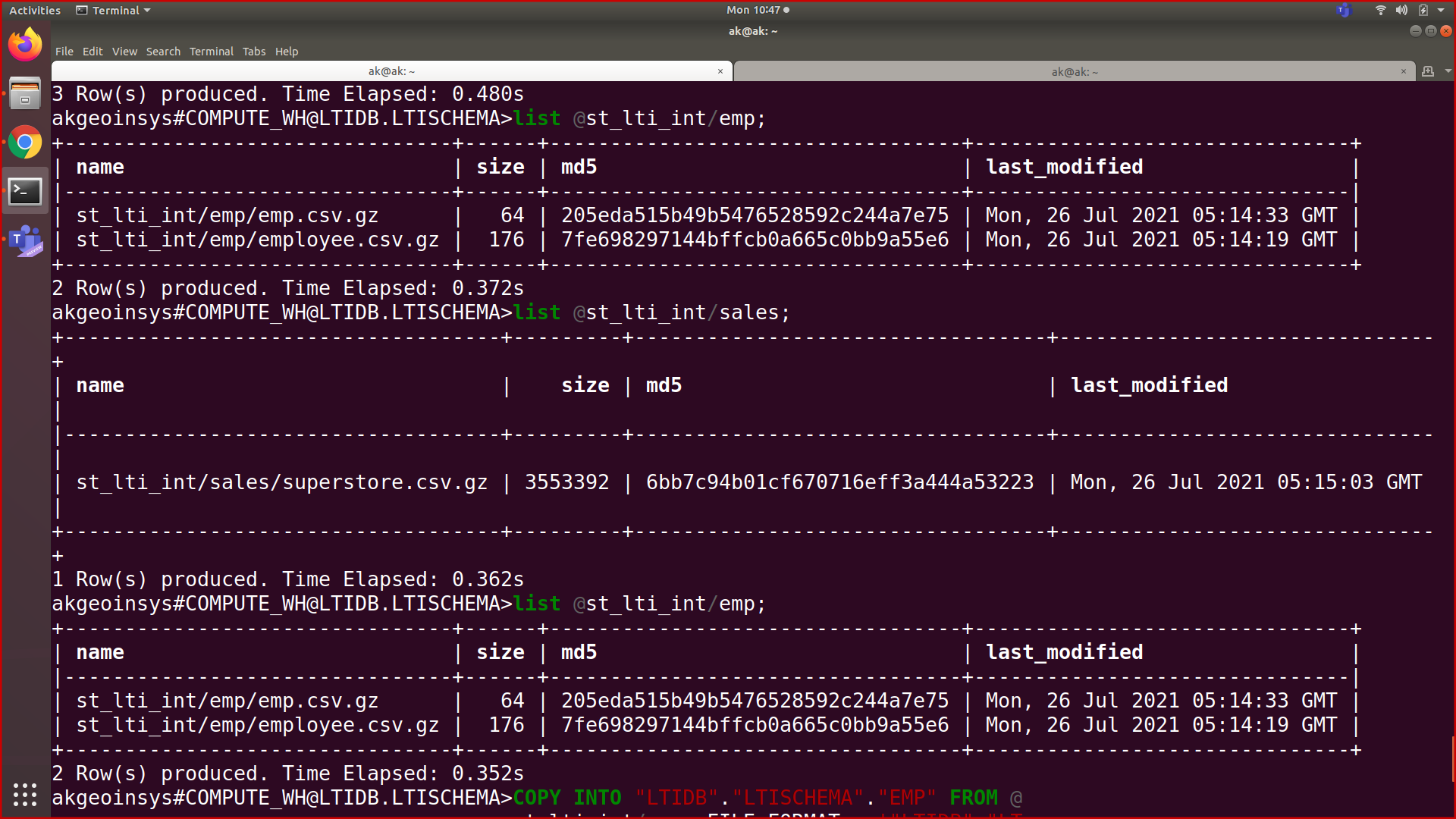
ON\_ERROR = 'CONTINUE' PURGE = FALSE ;

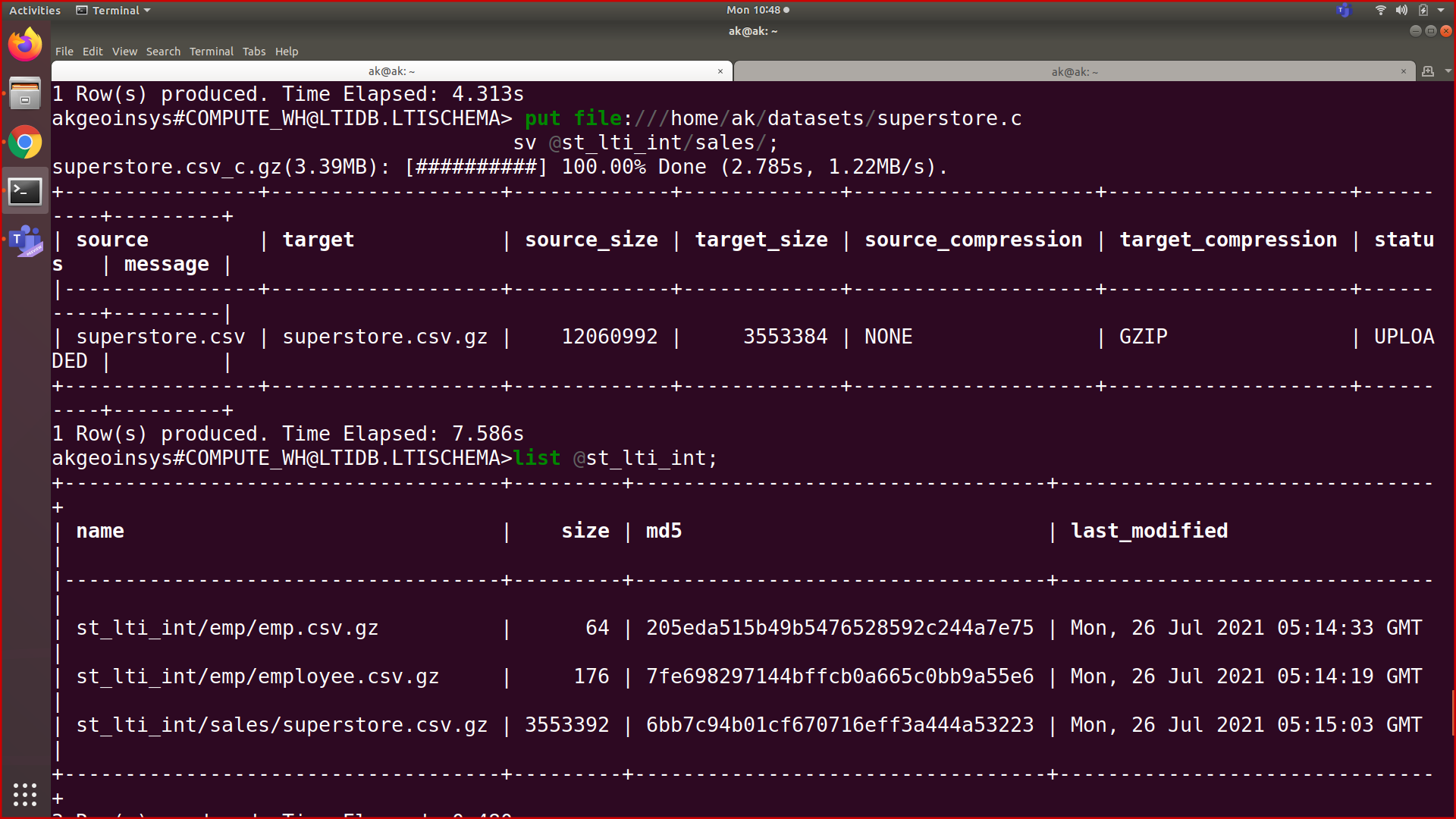
CREATE TABLE EMP1 LIKE EMP;

SELECT \* FROM EMP1;

put file:///home/ak/datasets/empext/employee.csv @st\_lti\_int/sales/2021/jul/28/;

put file:///home/ak/datasets/empext/employee.csv @st\_lti\_int/emp/;





External stage :

Cloud services :

AWS azure GCP

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

AWS public service :

AWS s3 :

1. External stage

2. External table

3. Snow pipes

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

AWS Root and IAM

IAM : access type

Programmatic access : access key id and secret access key

Aws management console : direct

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

AWS s3 ;

Region : mumbai

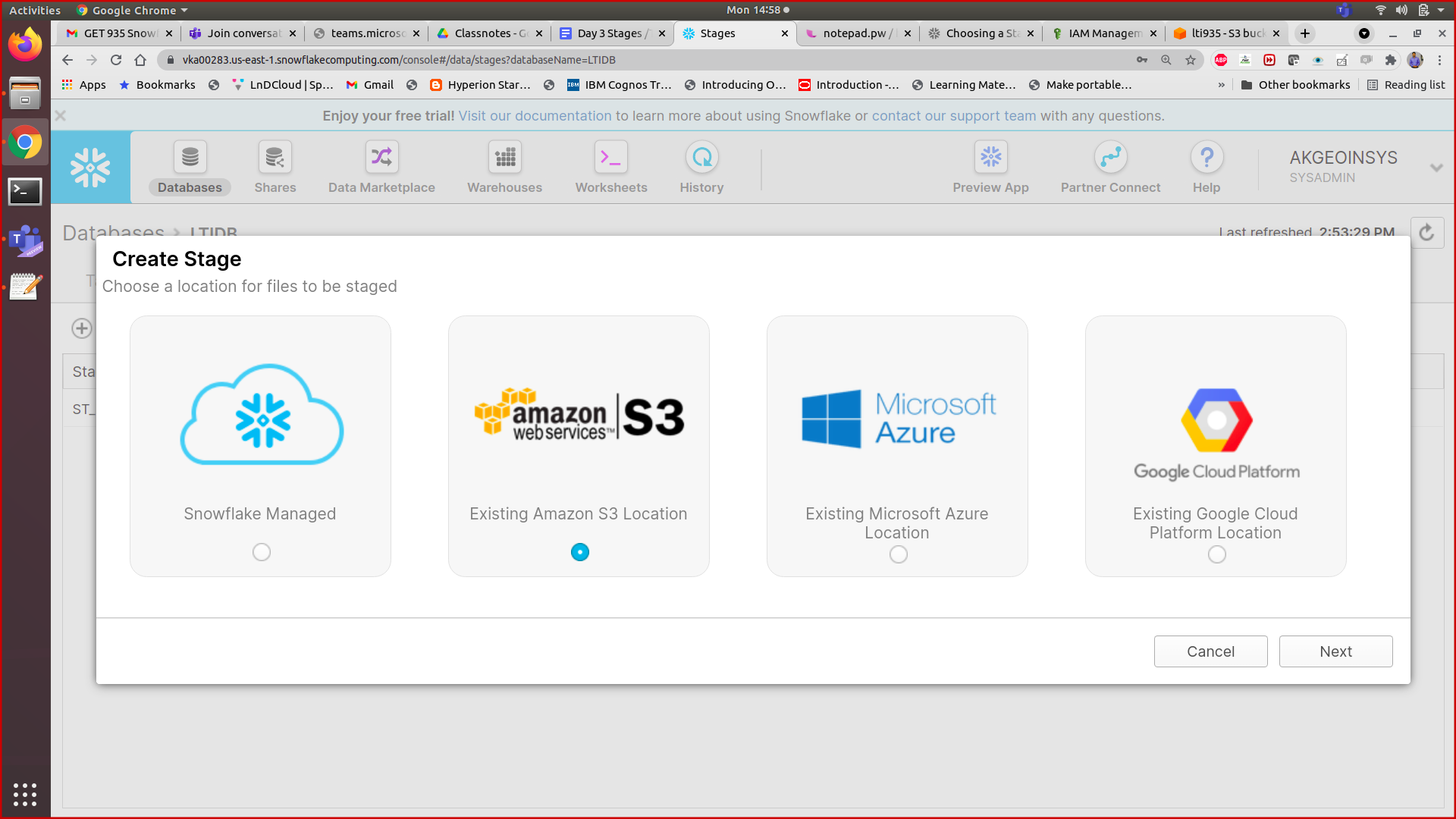
availability zone : 3 --within same region ---100 km

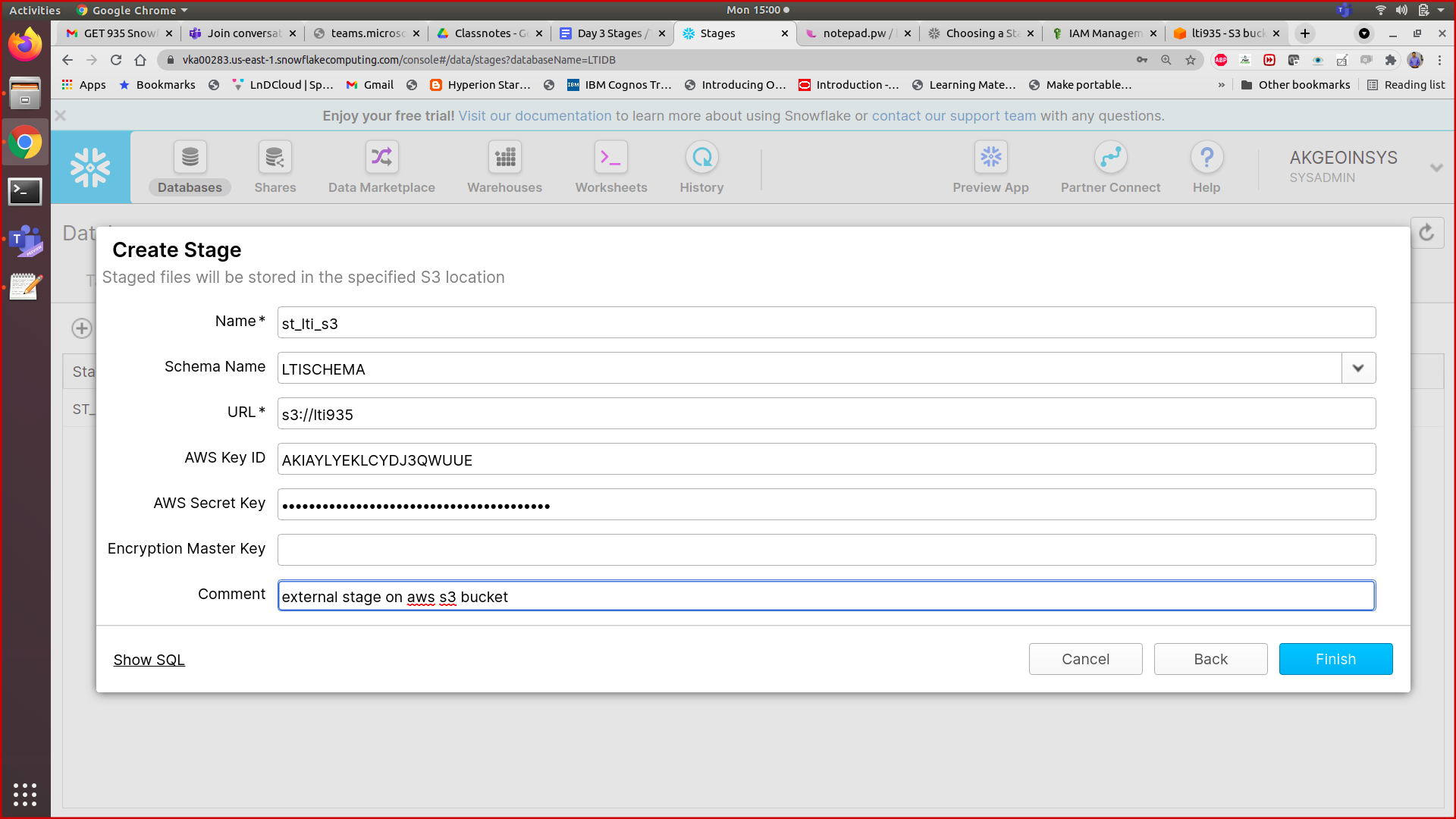
===Bucket ===>

Write once Read many → Wo RM

Local ==> Put ⇒ stage

External stage -s3 ⇒ext stage ==> copy ==>





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

Types of tables ;

1. Permanent table
2. Temporary table - temp -->user/session
3. Transient table
4. External table

1 Temporary database / schema /tables

2. Transient and permanent ;

Feature /storage cost

Time travel

Failsafe =7 days

Continuous data protection in snowflake (CDP)

1. Retention time :

Trans ==0 or 1

Perm ==0,1 or 90 days

==========TTL ===>

History data ⇒

26/6/2021 ⇒

26/7/2021

Time travel ;

Time stamp at/before

Interval /offset

Query id ⇒ statement

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

Drop table;

Undrop table ;

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

days Sal

1-----7000

5--30000

10--35000

20-40000

90- 70000

100-80000 =========table data current available