---------ETL JOB RUN Validation----------

Check whether job/pakage run is successfully executed.

---------Count validation----------

select count(\*) from employees

----------Data validation---------

select \* from employees order by employee\_id

---Duplicate

1)

select \* ,employee\_id ,first\_name ,last\_name ,email, phone\_number, hire\_date, job\_id ,salary, commission\_pct, manager\_id ,

department\_id, etl\_insert\_date, etl\_update\_date ,etl\_insert\_run\_id from employees

group by employee\_id ,first\_name ,last\_name ,email, phone\_number, hire\_date, job\_id ,salary, commission\_pct, manager\_id ,

department\_id, etl\_insert\_date, etl\_update\_date ,etl\_insert\_run\_id having count(\*)>1

2)

select count(\*) from employees

select count(distinct(employee\_id)) from employees

select \* from employees

3)

with cte as

(

select \*,DENSE\_RANK() over (partition by employee\_id ,first\_name ,last\_name ,email, phone\_number, hire\_date, job\_id ,salary, commission\_pct, manager\_id ,

department\_id, etl\_insert\_date, etl\_update\_date ,etl\_insert\_run\_id order by employee\_id ,first\_name ,last\_name ,email, phone\_number, hire\_date, job\_id ,salary, commission\_pct, manager\_id ,

department\_id, etl\_insert\_date, etl\_update\_date ,etl\_insert\_run\_id) as rnk from employees

)

select \* from cte where rnk>1

insert into employees values (102, 'Lex' ,'De Haan', 'LDEHAAN' , '515.123.4569', '1993-01-13', 'AD\_VP' ,50000 ,0 ,100 ,90, '2022-07-16', '2022-07-16' ,555666 )

insert into employees values (103, 'Alexander','Hunold', 'AHUNOLD', '590.423.4567', '1990-01-03', 'IT\_PROG', 31186, 0, 102, 60, '2022-07-16', '2022-07-16', 555666)

----Null value validation-------

select \* from employees

select \*

from INFORMATION\_SCHEMA.COLUMNS

where TABLE\_NAME='employees'

select count(\*)-count(first\_name) from employees;

select count(\*)-count(last\_name) from employees;

select count(\*)-count(phone\_number) from employees;

select count(\*)-count(hire\_date) from employees;

select count(\*)-count(employee\_id) from employees;

----control column validation----

select count(\*) from employees

select count(\*) as TGT\_TABLE\_COUNT ,etl\_insert\_date, etl\_update\_date, etl\_insert\_run\_id

from employees group by etl\_insert\_date, etl\_update\_date, etl\_insert\_run\_id

create table employees(

employee\_id decimal(10,0) primary key,

first\_name varchar(20),

last\_name varchar(25) ,

email varchar(25) ,

phone\_number varchar(20) ,

hire\_date date ,

job\_id varchar(10),

salary decimal(10,2),

commission\_pct decimal(10,2),

manager\_id decimal(10,0),

department\_id decimal(10,0),

START\_DATE date,

END\_DATE date,

CRNT\_FLAG varchar(1),

etl\_insert\_date date ,

etl\_update\_date date,

etl\_insert\_run\_id char(50),

)

select \* from employees

truncate table employees

drop table employees

update employees

set START\_DATE='2022-7-16'

update employees

set END\_DATE='2099-12-31'

update employees

set CRNT\_FLAG='Y'

---------ETL JOB RUN Validation----------

Check whether job/pakage run is successfully executed.

---------Count validation----------

select \* from employees

----------Data validation---------

select \* from employees order by employee\_id

---Duplicate

1)

select \* ,employee\_id ,first\_name ,last\_name ,email, phone\_number, hire\_date, job\_id ,salary, commission\_pct, manager\_id ,

department\_id, etl\_insert\_date, etl\_update\_date ,etl\_insert\_run\_id from employees

group by employee\_id ,first\_name ,last\_name ,email, phone\_number, hire\_date, job\_id ,salary, commission\_pct, manager\_id ,

department\_id, etl\_insert\_date, etl\_update\_date ,etl\_insert\_run\_id having count(\*)>1

2)

select count(\*) from employees

select count(distinct(employee\_id)) from employees

select \* from employees

3)

with cte as

(

select \*,DENSE\_RANK() over (partition by employee\_id ,first\_name ,last\_name ,email, phone\_number, hire\_date, job\_id ,salary, commission\_pct, manager\_id ,

department\_id, etl\_insert\_date, etl\_update\_date ,etl\_insert\_run\_id order by employee\_id ,first\_name ,last\_name ,email, phone\_number, hire\_date, job\_id ,salary, commission\_pct, manager\_id ,

department\_id, etl\_insert\_date, etl\_update\_date ,etl\_insert\_run\_id) as rnk from employees

)

select \* from cte where rnk>1

----Null value validation-------

select \* from employees

select \*

from INFORMATION\_SCHEMA.COLUMNS

where TABLE\_NAME='employees'

select count(\*)-count(first\_name) from employees;

select count(\*)-count(last\_name) from employees;

select count(\*)-count(phone\_number) from employees;

select count(\*)-count(hire\_date) from employees;

select count(\*)-count(employee\_id) from employees;

----control column validation----

select count(\*) from employees

select count(\*) as TGT\_TABLE\_COUNT ,etl\_insert\_date, etl\_update\_date, etl\_insert\_run\_id

from employees group by etl\_insert\_date, etl\_update\_date, etl\_insert\_run\_id

select \* from employees

select \* from INFORMATION\_SCHEMA.TABLES

select \* from INFORMATION\_SCHEMA.TABLES

select \*

from INFORMATION\_SCHEMA.COLUMNS

where TABLE\_NAME='employees'

create table employees(

employee\_id decimal(10,0) ,

first\_name varchar(20) NOT NULL,

last\_name varchar(25) NOT NULL,

email varchar(25),

phone\_number varchar(20) NOT NULL ,

hire\_date date NOT NULL,

job\_id varchar(10) ,

salary decimal(10,0),

commission\_pct decimal(10,0),

manager\_id decimal(10,0),

department\_id decimal(10,0),

etl\_insert\_date date ,

etl\_update\_date date,

etl\_insert\_run\_id char(50),

)

drop table employees

select \* from employees

select COL\_LENGTH(employee\_id,employees)

desc employees

update employees

set etl\_insert\_date='2022-7-16'

update employees

set etl\_update\_date='2022-7-16'

update employees

set etl\_update\_date='2022-7-16'

update employees

set etl\_insert\_run\_id=555666

truncate table employees

-----------------Landing Layer-----------------

----First layer

----First Database where data is landed

create table employees [employee\_id]

,[first\_name]

,[last\_name]

,[email]

,[phone\_number]

,[hire\_date]

,[job\_id]

,[salary]

,[manager\_id]

,[department\_id]

FROM [employees\_hr].[dbo].[employees]

select

select \* from employees

create table employees

(

EMPLOYEE\_ID NOT NULL int(6) ,

FIRST\_NAME VARCHAR2(20) ,

LAST\_NAME NOT NULL VARCHAR2(25) ,

EMAIL NOT NULL VARCHAR2(25) ,

PHONE\_NUMBER VARCHAR2(20) ,

HIRE\_DATE NOT NULL DATE ,

JOB\_ID NOT NULL VARCHAR2(10) ,

SALARY NUMBER(8,2) ,

COMMISSION\_PCT NUMBER(2,2) ,

MANAGER\_ID NUMBER(6) ,

DEPARTMENT\_ID NUMBER(4) )

Type 2 data for ELAQA

select \* from employees

alter table employees add ETL\_FLAG varchar

alter table employees add CRNT\_FLAG varchar

alter table employees add START\_DATE date

alter table employees add END\_DATE date

--alter table employees add CRNT\_FLAG varchar

sp\_help employees

create table employees(

employee\_id decimal(10,0) ,

first\_name varchar(20) ,

last\_name varchar(25) ,

email varchar(25),

phone\_number varchar(20) ,

hire\_date date ,

job\_id varchar(10) ,

salary decimal(10,0),

commission\_pct decimal(10,0),

manager\_id decimal(10,0),

department\_id decimal(10,0),

etl\_flag varchar,

etl\_insert\_date date ,

etl\_update\_date date,

etl\_insert\_run\_id char(50),

START\_DATE date,

END\_DATE date,

CRNT\_FLAG varchar

)

drop table employees

create table dummy(

employee\_id decimal(10,0) ,

first\_name varchar(20) ,

last\_name varchar(25) ,

email varchar(25),

phone\_number varchar(20) ,

hire\_date date ,

job\_id varchar(10) ,

salary decimal(10,0),

commission\_pct decimal(10,0),

manager\_id decimal(10,0),

department\_id decimal(10,0),

etl\_flag varchar,

etl\_insert\_date date ,

etl\_update\_date date,

etl\_insert\_run\_id char(50),

START\_DATE date,

END\_DATE date,

CRNT\_FLAG varchar

)

update employees

set CRNT\_FLAG='Y'

update employees

set etl\_flag='I'

update employees

set START\_DATE='2022-07-17'

where employee\_id in (209,210)

update employees

set END\_DATE='2022-07-17'

where employee\_id in (101, 102 ,103)

select \* from employees order by employee\_id

update employees

set etl\_insert\_date='2022-07-17'

where employee\_id in (100) and salary=150000

update employees

set etl\_flag='U'

where employee\_id in (100) and salary=150000

update employees

set etl\_insert\_run\_id=555777

where employee\_id in (100) and salary=150000

update employees

set CRNT\_FLAG='N'

where employee\_id in (101, 102 ,103)

insert into employees

values

(100,'Steven','King','SKING','515.123.4567','1987-06-17','AD\_PRES',95000,NULL,NULL ,90 ,'I','2022-07-16','2022-07-17',666777, '2022-07-16' ,'2022-07-17','N')

insert into employees

values(

107 ,'Diana', 'Lorentz', 'DLORENTZ', '590.423.5567', '1999-02-07','IT\_PROG', 14546, 0, 103, 60, 'I' ,'2022-07-16', '2022-07-17',666777,'2022-07-16','2022-07-17','Y')

select \* from employees order by employee\_id

update employees

set etl\_flag='I'

where employee\_id in (100) and salary=150000

update employees

set etl\_insert\_run\_id=555777

where employee\_id in (100) and salary=150000

update employees

set START\_DATE='2022-07-17'

where employee\_id in (100) and salary=150000

update employees

set etl\_flag='U'

where employee\_id in (100) and salary=95000

update employees

set etl\_insert\_run\_id=666777

where employee\_id in (100) and salary=95000

update employees

set START\_DATE='2022-07-17'

where employee\_id in (100) and salary=95000

insert into employees

values(

107 ,'Diana','Lorentz', 'DLORENTZ', '590.423.5567', '1999-02-07','IT\_PROG', 14546,

0, 103, 60, 'I' ,'2022-07-16', '2022-07-17',666777,'2022-07-16','2022-07-17','Y')

select \* from employees where employee\_id in (100,107) order by employee\_id

update employees

set etl\_flag='U'

where employee\_id in (107) and last\_name='Lorentz'

update employees

set etl\_insert\_date='2022-07-17'

where employee\_id in (107) and last\_name='Patel'

update employees

set etl\_insert\_run\_id=555777

where employee\_id in (100) and salary=150000

update employees

set START\_DATE='2022-07-17'

where employee\_id in (100) and salary=150000

update employees

set etl\_flag='U'

where employee\_id in (100) and salary=95000

update employees

set etl\_insert\_run\_id=666777

where employee\_id in (100) and salary=95000

update employees

set START\_DATE='2022-07-17'

where employee\_id in (100) and salary=95000

Type 1

select \* from employees

alter table employees drop column START\_DATE

alter table employees drop column END\_DATE

alter table employees drop column CRNT\_FLAG

alter table employees add ETL\_FLAG varchar

sp\_help employees

update employees

set salary=150000

where employee\_id=100

update employees

set ETL\_FLAG='U'

where employee\_id in (107)

update employees

set last\_name='Patel'

where employee\_id in (107)

update employees

set etl\_insert\_run\_id=666777

where employee\_id in (100,101,102,103,107)

update employees

set etl\_update\_date='2022-07-17'

where employee\_id in (107)

SR\_AD\_PRES 150000

create table dummy(

employee\_id decimal(10,0) primary key,

first\_name varchar(20),

last\_name varchar(25) ,

email varchar(25) ,

phone\_number varchar(20) ,

hire\_date date ,

job\_id varchar(10),

salary decimal(10,2),

commission\_pct decimal(10,2),

manager\_id decimal(10,0),

department\_id decimal(10,0),

START\_DATE date,

END\_DATE date,

CRNT\_FLAG varchar(1),

etl\_insert\_date date ,

etl\_update\_date date,

etl\_insert\_run\_id char(50),

)

209 Julia Bull JRUSSEL 011.44.1344.429268 10/1/1996 SA\_MAN 48505.27 0.4 100 80

210 Samuel Dellinger KPARTNER 011.44.1344.467268 1/5/1997 SA\_MAN 40671.73 0.3 100 80

211 Winston Fay PFAY 603.123.6666 35659 MK\_REP 15716.69 0.01 120 50

select count(\*) from employees where crnt\_flag='Y'

select count(\*) from employees --where crnt\_flag='Y'

select

emp.employee\_id ,

emp.first\_name,

emp.last\_name,

first\_name+' '+last\_name as full\_name,

case when salary<=20000 then 'LOW'

when salary>20000 and salary<=60000 then 'AVG'

when salary>60000 then 'GOOD' end

as salary\_ind,

DATEDIFF(yyyy,emp.hire\_date,GETDATE()) Experience\_year,

case when job\_id like 'SA%' then 'Sales'

when job\_id like 'SA%' then 'Other' end

as sales\_indicator,

email,

phone\_number,

hire\_date,

job\_id,

salary,

commission\_pct,

null as Marriage\_Status,

null as Employee\_performance\_out\_of\_10,

null as EMP\_GENDER,

emp.manager\_id,

emp.department\_id,

loc.city,

dp.department\_name

from employees emp left outer join departments dp

on emp.department\_id=dp.department\_id

left outer join locations loc on dp.location\_id=loc.location\_id

where CRNT\_FLAG='Y' and salary>=20000 and hire\_date>'1998-12-31'

select \* from locations

create table locations

(

LOCATION\_ID decimal ,

city varchar(30))

insert into locations

values(1500,'Tokyo')

sp\_help employees

create table departments (

DEPARTMENT\_ID decimal ,

DEPARTMENT\_NAME varchar(40) ,

MANAGER\_ID decimal ,

LOCATION\_ID decimal )

drop table departments

select \* from departments

insert into departments values (80 ,'Sales', 145,2500)

20 201 1800

30 114 1700

40 203 2400

50 121 1500

60 IT 103 1400

70 Public Relations 204 2700

80 145 2500

90 Executive 100 1700

100 Finance 108 1700

110 Accounting 205 1700

create table Emp\_Information (

employee\_id decimal(10,0),

first\_name varchar(50),

last\_name varchar(50),

full\_name varchar(50),

salary\_ind varchar(50),

Experience\_year varchar(10)

,sales\_indicator varchar(50), email varchar(50)

,phone\_number varchar(50),

hire\_date date ,job\_id varchar(50),

salary decimal(10,2),

commission\_pct decimal(10,2) ,

Marriage\_Status varchar(50),

Employee\_performance\_out\_of\_10 varchar(50), EMP\_GENDER varchar(50),

manager\_id decimal(10,0),

department\_id decimal(10,0), city varchar(50), department\_name varchar(50))

drop table Emp\_Information

create table dummy(

employee\_id decimal(10,0) primary key,

first\_name varchar(20),

last\_name varchar(25) ,

email varchar(25) ,

phone\_number varchar(20) ,

hire\_date date ,

job\_id varchar(10),

salary decimal(10,2),

commission\_pct decimal(10,2),

manager\_id decimal(10,0),

department\_id decimal(10,0),

START\_DATE date,

END\_DATE date,

CRNT\_FLAG varchar(1),

etl\_insert\_date date ,

etl\_update\_date date,

etl\_insert\_run\_id char(50),

)

select \* from Emp\_Information

truncate table emp\_information

insert into emp\_information(full\_name ,salary\_ind, Experience\_year, sales\_indicator ,city ,deprtment\_name)

values

(Steven Markle GOOD 22 NULL

Gerald Cambrault AVG 23 Sales

Eleni Zlotkey AVG 22 Sales

Oliver Tuvault AVG 23 Sales

Danielle Greene AVG 23 Sales

Mattea Marvins AVG 22 Sales

David Lee AVG 22 Sales

William Smith AVG 23 Sales

Elizabeth Bates AVG 23 Sales)

drop table Emp\_Information

insert into Emp\_Information values

(

128, 'Steven', 'Markle', 'Steven Markle', 'GOOD',

22, NULL, 'SMARKLE' ,'650.124.1434', '2000-03-08',

'ST\_CLERK', 72600 ,0 ,NULL, NULL, NULL, 120 ,50, 'Tokyo', 'Shipping')

(148, 'Gerald' ,'Cambrault' ,'Gerald' ,

'Cambrault', 'AVG', '23', 'Sales', 'GCAMBRAU', 011.44.1344.619268, 10/15/1999,

SA\_MAN ,33143 ,0, NULL ,NULL, NULL, 100, 80, NULL, 'Sales')

149 Eleni Zlotkey Eleni Zlotkey AVG 22 Sales EZLOTKEY 011.44.1344.429018 1/29/2000 SA\_MAN 31628 0 NULL NULL NULL 100 80 NULL Sales

155 Oliver Tuvault Oliver Tuvault AVG 23 Sales OTUVAULT 011.44.1344.486508 11/23/1999 SA\_REP 21093 0 NULL NULL NULL 145 80 NULL Sales

163 Danielle Greene Danielle Greene AVG 23 Sales DGREENE 011.44.1346.229268 3/19/1999 SA\_REP 28621 0 NULL NULL NULL 147 80 NULL Sales

164 Mattea Marvins Mattea Marvins AVG 22 Sales MMARVINS 011.44.1346.329268 1/24/2000 SA\_REP 21692 0 NULL NULL NULL 147 80 NULL Sales

165 David Lee David Lee AVG 22 Sales DLEE 011.44.1346.529268 2/23/2000 SA\_REP 20482 0 NULL NULL NULL 147 80 NULL Sales

171 William Smith William Smith AVG 23 Sales WSMITH 011.44.1343.629268 2/23/1999 SA\_REP 22290 0 NULL NULL NULL 148 80 NULL Sales

172 Elizabeth Bates Elizabeth Bates AVG 23 Sales EBATES 011.44.1343.529268 3/24/1999 SA\_REP 21997 0 NULL NULL NULL 148 80 NULL Sales

----------------------------------------------------------------------------------------------

select \* from Emp\_Information

alter table Emp\_Information add etl\_insert\_date date

alter table Emp\_Information add etl\_update\_date date

alter table Emp\_Information add etl\_run\_id varchar(50)

alter table Emp\_Information add etl\_flag varchar(5)

update Emp\_Information

set etl\_flag ='I'