**EDA & DATA CLEANING IN SQL**

select \*

from `first-447405.Product\_Analytics.Orders` limit 10

1. **Check missing data**

SELECT count(\*) as total\_rows,

        countif(order\_id is null) as null\_order\_ids,

        countif(user\_id is null) as null\_user\_ids,

        countif(product\_id is null) as null\_product\_ids,

        countif(order\_date is null) as null\_orderdate,

        countif(quantity is null ) as null\_qunatity

FROM `first-447405.Product\_Analytics.Orders`

1. **How many users never made a purchase**

SELECT user\_id

from `first-447405.Product\_Analytics.Orders`

group by user\_id

having count(order\_date)=0

1. **count orders by product category**

select p.category,count(o.order\_id)

from `first-447405.Product\_Analytics.Orders` o

Join `first-447405.Product\_Analytics.Product` p

on p.product\_id=o.product\_id

group by category

1. **For Order date Data Range**

SELECT

  MIN(PARSE\_DATE('%d-%m-%Y', order\_date)) AS first\_order\_date,

  MAX(PARSE\_DATE('%d-%m-%Y', order\_date)) AS last\_order\_date

from `first-447405.Product\_Analytics.Orders`

WHERE SAFE.PARSE\_DATE('%d-%m-%Y', order\_date) IS NOT NULL;

1. **How many users signed up per month**

SELECT user\_id,

      min(order\_date) as earliest\_purchase

from `first-447405.Product\_Analytics.Orders`

group by 1

1. **Monthly signups**

SELECT format\_date('%Y-%m',Date(signup\_date)) as signup\_month,

        count(\*) as users\_signed\_up

From

group by 1

order by 1

1. **Query for finding the  signup date and order date**

SELECT u.user\_id, u.signup\_date, o.order\_date

FROM `first-447405.Product\_Analytics.Orders`o

Join `first-447405.Product\_Analytics.Users` u

on u.user\_id=o.user\_id

 LIMIT 1000

1. **Clean orders Table**

create or replace table `first-447405.Product\_Analytics.Orders\_cleaned` as

select \*

from `first-447405.Product\_Analytics.Orders`

where user\_id is not null

and order\_id is not null

and Safe.Parse\_date('%d-%m-%Y',order\_date) is not null;