Case Study: Paintings

KPI’s

1. Fetch all the paintings which are not displayed in any museums?

SELECT name

FROM work

WHERE museum\_id IS NULL

LIMIT 10;

A screenshot of a computer

Description automatically generated

2. Are there museums without any paintings?

select \* from museum m

where not exists (select 1 from work w

where w.museum\_id=m.museum\_id)



3. How many paintings have an asking price of more than their regular price?

SELECT COUNT(\*) AS Num

FROM product\_size

WHERE sale\_price > regular\_price

A screenshot of a computer

Description automatically generated

4. Identify the paintings whose asking price is less than 50% of its regular price

SELECT DISTINCT name

FROM product\_size ps

INNER JOIN work w

ON w.work\_id = ps.work\_id

WHERE sale\_price < regular\_price \* .5

A screenshot of a computer

Description automatically generated

5. Which canva size costs the most?

SELECT cs.label, ps.sale\_price

FROM(

SELECT \*,

RANK() OVER(ORDER BY sale\_price DESC) AS rnk

FROM product\_size) ps

JOIN canvas\_size cs ON cs.size\_id::TEXT = ps.size\_id

WHERE ps.rnk = 1

A screenshot of a computer

Description automatically generated

6. Delete duplicate records from work, product\_size, subject and image\_link tables?

DELETE FROM work

WHERE ctid not in (select min(ctid)

FROM work

GROUP BY work\_id );

DELETE FROM product\_size

WHERE ctid NOT IN (SELECT MIN(ctid )

FROM product\_size

GROUP BY work\_id, size\_id

) 

DELETE FROM image\_link

WHERE ctid NOT IN (SELECT MIN(ctid)

FROM image\_link

GROUP BY work\_id)



DELETE FROM subject

WHERE ctid NOT IN (SELECT MIN(ctid)

FROM subject

GROUP BY work\_id, subject)



7. Identify the museums with invalid city information in the given dataset

SELECT \* FROM museum

WHERE city ~ '^[0-9]'

A screenshot of a computer

Description automatically generated

8. Museum\_Hours table has 1 invalid entry. Identify it and remove it.

SELECT DISTINCT day FROM museum\_hours;

A screenshot of a calendar

Description automatically generated

SELECT \* FROM museum\_hours

WHERE day NOT IN ('Tuesday','Saturday','Friday','Wednesday','Monday','Sunday','Thursday')

A screenshot of a calendar

Description automatically generated

UPDATE museum\_hours

SET day = 'Thursday'

WHERE day = 'Thusday'

--clean

DELETE FROM museum\_hours

WHERE ctid NOT IN (SELECT MIN(ctid)

FROM museum\_hours

GROUP BY museum\_id, day)



9. Fetch the top 10 most famous painting subject

select \*

from (

select s.subject,count(1) as no\_of\_paintings

,rank() over(order by count(1) desc) as ranking

from work w

join subject s on s.work\_id=w.work\_id

group by s.subject ) x

where ranking <= 10;

A screenshot of a computer

Description automatically generated

10. Identify the museums which are open on both Sunday and Monday. Display

museum name, city.

-- CTE

WITH museum\_sun\_mon AS(

SELECT museum\_id FROM museum\_hours

WHERE day IN ('Sunday', 'Monday')

GROUP BY museum\_id

HAVING COUNT(\*) = 2 )

-- Query

SELECT m.museum\_id, name

FROM museum m

JOIN museum\_sun\_mon msm

ON msm.museum\_id = m.museum\_id

A screenshot of a computer

Description automatically generated

11. How many museums are open every single day?

-- CTE

WITH museum\_sun\_mon AS(

SELECT museum\_id FROM museum\_hours

GROUP BY museum\_id

HAVING COUNT(\*) = 7 )

-- Query

SELECT COUNT(\*)

FROM museum\_sun\_mon

A screenshot of a computer

Description automatically generated

12. Which are the top 5 most popular museum? (Popularity is defined based on most no of paintings in a museum)

SELECT top\_5.museum\_id, name

FROM (

SELECT museum\_id, COUNT(\*) num\_of\_paintings

FROM work

WHERE museum\_id IS NOT NULL

GROUP BY museum\_id

ORDER BY num\_of\_paintings DESC

LIMIT 5) top\_5

JOIN museum m

ON top\_5.museum\_id = m.museum\_id

A screenshot of a computer

Description automatically generated

13. Who are the top 5 most popular artist? (Popularity is defined based on most no of paintings done by an artist)

SELECT top\_5.artist\_id, full\_name

FROM (

SELECT artist\_id, COUNT(\*) num\_of\_paintings

A screenshot of a computer

Description automatically generatedFROM work

WHERE artist\_id IS NOT NULL

GROUP BY artist\_id

ORDER BY num\_of\_paintings DESC

LIMIT 5) top\_5

JOIN artist a

ON top\_5.artist\_id = a.artist\_id

14. Display the 3 least popular canva sizes

SELECT \*

FROM (

SELECT \*,

RANK() OVER(ORDER BY height\*width ) rnk

FROM canvas\_size

WHERE width IS NOT NULL AND height IS NOT NULL) cs

WHERE cs.rnk < 4

A screenshot of a computer

Description automatically generated

15. Which museum is open for the longest during a day. Display museum name, state and hours open and which day?

SELECT name, state, day, open, close, duration

FROM(

SELECT \*,

to\_timestamp(open, 'HH:MI AM') time\_open,

to\_timestamp(close, 'HH:MI PM') time\_close,

to\_timestamp(close, 'HH:MI PM') - to\_timestamp(open, 'HH:MI AM') AS duration,

RANK() OVER(ORDER BY to\_timestamp(close, 'HH:MI PM') - to\_timestamp(open, 'HH:MI AM') DESC) rnk

FROM museum\_hours ) time1

JOIN museum m

ON m.museum\_id = time1.museum\_id

WHERE time1.rnk = 1

LIMIT 1

A screenshot of a computer

Description automatically generated

16. Which museum has the most no of most popular painting style?

with pop\_style as

(select style

,rank() over(order by count(1) desc) as rnk

from work

group by style),

cte as

(select w.museum\_id,m.name as museum\_name,ps.style, count(1) as no\_of\_paintings

,rank() over(order by count(1) desc) as rnk

from work w

join museum m on m.museum\_id=w.museum\_id

join pop\_style ps on ps.style = w.style

where w.museum\_id is not null

and ps.rnk=1

group by w.museum\_id, m.name,ps.style)

select museum\_name,style,no\_of\_paintings

from cte

where rnk=1;

A screenshot of a computer

Description automatically generated

17. Identify the artists whose paintings are displayed in multiple countries

WITH cte AS (

SELECT distinct a.full\_name artist\_name, w.name painting\_name, m.name museum\_name, m.country museum\_country

FROM work w

JOIN artist a

ON a.artist\_id = w.artist\_id

JOIN museum m

ON w.museum\_id = m.museum\_id

WHERE w.museum\_id IS NOT NULL)

SELECT artist\_name, COUNT(DISTINCT museum\_country) num\_countries

FROM cte

GROUP BY artist\_name

ORDER BY num\_countries DESC

A screenshot of a computer

Description automatically generated

18. Display the country and the city with most no of museums. Output 2 seperate

columns to mention the city and country. If there are multiple value, seperate them with comma.

with cte\_country as

(select country, count(1)

, rank() over(order by count(1) desc) as rnk

from museum

WHERE country IS NOT NULL

group by country),

cte\_city as

(select city, count(1)

, rank() over(order by count(1) desc) as rnk

from museum

WHERE city IS NOT NULL

group by city)

select string\_agg(distinct country.country,', '), string\_agg(city.city,', ')

from cte\_country country

cross join cte\_city city

where country.rnk = 1

and city.rnk = 1

A screenshot of a computer

Description automatically generated

19. Identify the artist and the museum where the most expensive and least expensive painting is placed. Display the artist name, sale\_price, painting name, museum name, museum city and canvas label

WITH table1 AS (

SELECT a.full\_name artist\_name, ps.sale\_price, w.name painting\_name, m.name museum\_name, m.city, cs.label,

RANK() OVER(ORDER BY ps.sale\_price DESC) rnk\_desc

FROM work w

JOIN artist a ON a.artist\_id = w.artist\_id

JOIN museum m ON m.museum\_id = w.museum\_id

JOIN product\_size ps ON ps.work\_id = w.work\_id

JOIN canvas\_size cs ON cs.size\_id::TEXT = ps.size\_id

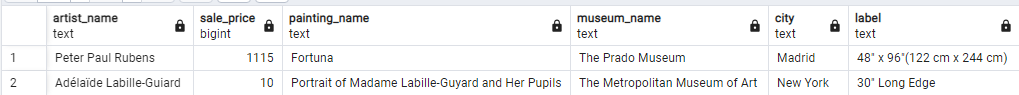
)

SELECT artist\_name, sale\_price, painting\_name, museum\_name, city, label

FROM table1

WHERE rnk\_desc = 1 OR rnk\_desc = (SELECT MAX(rnk\_desc) FROM table1)

LIMIT 2



20. Which country has the 5th highest no of paintings?

WITH cte AS(

SELECT m.country, COUNT(\*) num,

RANK() OVER(ORDER BY count(\*) DESC) rnk

FROM work w

JOIN museum m

ON m.museum\_id = w.museum\_id

WHERE w.museum\_id IS NOT NULL

GROUP BY country

)

SELECT country, num FROM cte

WHERE rnk = 5

A screenshot of a computer

Description automatically generated

21. Which are the 3 most popular and 3 least popular painting styles?

WITH cte AS(

SELECT style, COUNT(\*) count\_num,

RANK() OVER(ORDER BY COUNT(\*) DESC) rnk,

COUNT(\*) OVER() AS counting

FROM work

WHERE style IS NOT NULL

GROUP BY style

)

SELECT style, count\_num,

CASE

WHEN rnk <= 3 THEN 'MOST POPULAR PAINTING'

ELSE 'LEAST POPULAR PAINTING'

END "Status"

FROM cte

WHERE rnk <= 3

OR rnk > counting – 3

A screenshot of a computer

Description automatically generated

22. Which artist has the most no of Portraits paintings outside USA?. Display artist

name, no of paintings and the artist nationality.

WITH cte AS (

SELECT a.full\_name, nationality, COUNT(\*) no\_of\_portraits,

RANK() OVER(ORDER BY COUNT(\*) DESC) rnk

FROM work w

JOIN subject s

ON w.work\_id = s.work\_id

JOIN museum m

ON m.museum\_id = w.museum\_id

JOIN artist a

ON a.artist\_id = w.artist\_id

WHERE subject = 'Portraits' AND country != 'USA'

GROUP BY a.full\_name, nationality

)

SELECT full\_name, nationality, no\_of\_portraits

FROM cte

WHERE rnk = 1

A screenshot of a computer

Description automatically generated