USE sql\_tasks;

SELECT \* FROM insurance;

-- 1. What are the top 5 patients who claimed the highest insurance amounts?

SELECT \*,DENSE\_RANK() OVER(ORDER BY claim DESC) FROM insurance LIMIT 5;

-- 2. What is the average insurance claimed by patients based on the

-- number of children they have?

SELECT children,avg\_claim FROM (SELECT \*,

AVG(claim) OVER(PARTITION BY children) AS avg\_claim,

ROW\_NUMBER() OVER(PARTITION BY children) AS row\_num

FROM insurance) t

WHERE t.row\_num = 1;

-- 3. What is the highest and lowest claimed amount by patients in each region?

SELECT region,min\_claim,max\_claim FROM (SELECT \*,

MIN(claim) OVER(PARTITION BY region) AS min\_claim,

MAX(claim) OVER(PARTITION BY region) AS max\_claim,

ROW\_NUMBER() OVER(PARTITION BY region) AS row\_num

FROM insurance) t

WHERE t.row\_num = 1;

-- 4. What is the percentage of smokers in each age group?

SELECT \* FROM insurance;

SELECT age, percentage FROM( SELECT age, ROW NUMBER() OVER (PARTITION BY age ORDER BY age) AS ‘ROW’, (VALUE/TOTAL) \* 100 AS ‘PERCENTAGE’

FROM ( SELECT LAST\_VALUE(M) OVER (ORDER BY time ROWS BETWEEN UNBOUNDED PRECEDING AND UNBOUNDED FOLLOWING) AS ‘TOTAL’

FROM (SELECT COUNT(smoker) OVER (PARTITION BY age ORDER BY age) AS ‘VALUE’,ROW NUMBER() OVER (ORDER BY age)) AS ‘M’ FROM INSURANCE;

-- 5. What is the difference between the claimed amount of each patient and the claimed

-- amount of the first patient?

SELECT \*,

claim - FIRST\_VALUE(claim) OVER() AS diff

FROM insurance;

-- 6. For each patient, calculate the difference between their claimed amount

-- and the average claimed amount of patients with the same number of children.

SELECT \*,

claim - AVG(claim) OVER(PARTITION BY children)

FROM insurance;

-- 7. Show the patient with the highest BMI in each region and their

-- respective overall rank.

SELECT \* FROM (SELECT \*,

RANK() OVER(PARTITION BY region ORDER BY bmi DESC) AS group\_rank,

RANK() OVER(ORDER BY bmi DESC) AS overall\_rank

FROM insurance) t

WHERE t.group\_rank = 1;

-- 8. Calculate the difference between the claimed amount of each patient

-- and the claimed amount of the patient who has the highest BMI

-- in their region.

SELECT \*,

claim - FIRST\_VALUE(claim) OVER(PARTITION BY region ORDER BY bmi DESC)

FROM insurance;

-- 9. For each patient, calculate the difference in claim amount

-- between the patient and the patient with the highest claim amount

-- among patients with smoker status,

-- within the same region. Return the result in descending order difference.

SELECT \*,

(MAX(claim) OVER(PARTITION BY region,smoker) - claim) AS claim\_diff

FROM insurance

ORDER BY claim\_diff DESC;

-- 10. For each patient, find the maximum BMI value among their next three

-- records (ordered by age).

SELECT \*,

MAX(bmi) OVER(ORDER BY age ROWS BETWEEN 1 FOLLOWING AND 3 FOLLOWING)

FROM insurance;

-- 11. For each patient, find the rolling average of the last 2 claims.

SELECT \*,

AVG(claim) OVER(ROWS BETWEEN 2 PRECEDING AND 1 PRECEDING)

FROM insurance;

-- 12. Find the first claimed insurance value for male and female patients,

-- within each region order the data by patient age in ascending order,

-- and only include patients who are non-diabetic and have a bmi value

-- between 25 and 30.

WITH filtered\_data AS (

SELECT \* FROM insurance

WHERE diabetic = 'No' AND bmi BETWEEN 25 AND 30

)

SELECT region,gender,first\_claim FROM (SELECT \*,

FIRST\_VALUE(claim) OVER(PARTITION BY region,gender ORDER BY age) AS first\_claim,

ROW\_NUMBER() OVER(PARTITION BY region,gender ORDER BY age) AS row\_num

FROM filtered\_data) t

WHERE t.row\_num = 1