Healthcare Analysis

M VINOD KUMAR

PATIENTS

patient_id	patient_name	age	gender	city
1	John Smith	45	Male	Seattle
2	Jane Doe	32	Female	Miami
3	Mike Johnson	50	Male	Seattle
4	Lisa Jones	28	Female	Miami
5	David Kim	60	Male	Chicago
NULL	NULL	NULL	NULL	NULL

PATIENT_ID

SYMPTOMS

symptom_id	symptom_name
1	Fever
2	Cough
3	Difficulty Breathing
4	Fatigue
5	Headache
NULL	NULL

SYMPTOM_ID

VISITS

visit_id	patient_id	symptom_id	diagnosis_id	visit_date
1	1	1	2	2022-01-01
2	2	2	1	2022-01-02
3	3	3	3	2022-01-02
4	4	1	4	2022-01-03
5	5	2	5	2022-01-03
6	1	4	1	2022-05-13
7	3	4	1	2022-05-20
8	3	2	1	2022-05-20
9	2	1	4	2022-08-19
10	1	2	5	2022-12-01
NULL	NULL	NULL	NULL	NULL

DIAGNOSIS_ID

EMPLOYEES

diagnosis_id	diagnosis_name
1	Common Cold
2	Influenza
3	Pneumonia
4	Bronchitis
5	COVID-19
NULL	NULL

DATA MODEL

```sql
-- 1. Write a SQL query to retrieve all patients who have been diagnosed with COVID-19

SELECT PATIENT_NAME, DIAGNOSIS_NAME FROM VISITS JOIN DIAGNOSES
USING (DIAGNOSIS_ID)
JOIN PATIENTS
USING(PATIENT_ID)
WHERE DIAGNOSIS_NAME = "COVID-19"
ORDER BY PATIENT_ID;
```

Grid | Filter Rows: | Export: | Wrap Cell Content:

PATIENT_NAME	DIAGNOSIS_NAME
John Smith	COVID-19
David Kim	COVID-19

```sql
-- 2. Write a SQL query to retrieve the number of visits made by each patient, ordered by the number of visits in descending order.

SELECT PATIENT_ID,PATIENT_NAME, COUNT(PATIENT_ID) AS TOTAL_VISITS
FROM VISITS
JOIN PATIENTS
USING(PATIENT_ID)
GROUP BY PATIENT_ID
ORDER BY TOTAL_VISITS DESC;
```

Grid | Filter Rows: | Export: | Wrap Cell Content:

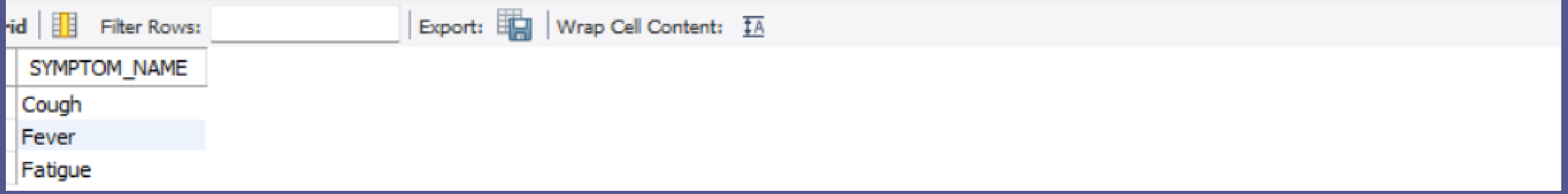
PATIENT_ID	PATIENT_NAME	TOTAL_VISITS
1	John Smith	3
3	Mike Johnson	3
2	Jane Doe	2
4	Lisa Jones	1
5	David Kim	1

```sql
-- 3. Write a SQL query to calculate the average age of patients who have been diagnosed with Pneumonia.

SELECT DIAGNOSIS_NAME, AVG(AGE) FROM VISITS
JOIN PATIENTS USING(PATIENT_ID)
JOIN DIAGNOSES USING(DIAGNOSIS_ID)
WHERE DIAGNOSIS_NAME = "PNEUMONIA" ORDER BY DIAGNOSIS_ID;
```

rid | Filter Rows: | Export: | Wrap Cell Content:

DIAGNOSIS_NAME	AVG(AGE)
Pneumonia	50.0000

```sql
-- 4. Write a SQL query to retrieve the top 3 most common symptoms among all visits.

WITH TEMP AS(
SELECT SYMPTOM_ID, COUNT(SYMPTOM_ID) AS MORE_FREQUENT_SYMPTOM,
RANK() OVER(ORDER BY COUNT(SYMPTOM_ID) DESC) AS RN
FROM VISITS GROUP BY SYMPTOM_ID)

SELECT SYMPTOM_NAME FROM TEMP JOIN SYMPTOMS USING(SYMPTOM_ID) WHERE RN <= 3 ORDER BY MORE_FREQUENT_SYMPTOM DESC;
```



SYMPTOM_NAME
Cough
Fever
Fatigue

```sql
-- 5. Write a SQL query to retrieve the patient who has the highest number of different symptoms reported.

WITH TEMP AS (
SELECT PATIENT_ID, COUNT(SYMPTOM_ID) AS TOTAL_DIFF_SYMPTOMS,
DENSE_RANK() OVER(ORDER BY COUNT(SYMPTOM_ID) DESC) AS RN
FROM VISITS GROUP BY PATIENT_ID)

SELECT PATIENT_NAME, TOTAL_DIFF_SYMPTOMS FROM TEMP JOIN PATIENTS USING (PATIENT_ID)
WHERE RN = 1;
```

id | Filter Rows: | Export: | Wrap Cell Content:

PATIENT_NAME	TOTAL_DIFF_SYMPTOMS
John Smith	3
Mike Johnson	3

```sql
-- 6. Write a SQL query to calculate the percentage of patients who have been diagnosed with COVID-19 out of the total number of patients.

WITH TEMP AS (
SELECT COUNT(*) AS COVID19_PATIENTS,(SELECT COUNT(*) AS TOTAL_PATIENTS FROM VISITS) AS TOTAL_PATIENTS
FROM VISITS
JOIN DIAGNOSES
USING(DIAGNOSIS_ID)
WHERE DIAGNOSIS_NAME = "COVID-19")

SELECT CONCAT(ROUND((COVID19_PATIENTS/TOTAL_PATIENTS)*100,0),"%") AS "% COVID-19 PATIENTS" FROM TEMP;
```

Grid | Filter Rows: | Export: | Wrap Cell Content:

% COVID-19 PATIENTS
20%

```sql
-- 7. Write a SQL query to retrieve the top 5 cities with the highest number of visits, along with the count of visits in each city.

WITH TEMP AS (
SELECT CITY, COUNT(*) AS NO_OF_VISITS, RANK() OVER(ORDER BY COUNT(*) DESC) AS RN
FROM VISITS
JOIN PATIENTS
USING(PATIENT_ID)
GROUP BY CITY)

SELECT * FROM TEMP WHERE RN<=5;
```

rid | Filter Rows: | Export: | Wrap Cell Content:

CITY	NO_OF_VISITS	RN
Seattle	6	1
Miami	3	2
Chicago	1	3

```sql
-- 8. Write a SQL query to find the patient who has the highest number of visits in a single day, along with the corresponding visit date.

SELECT PATIENT_NAME, VISIT_DATE, NO_OF_VISITS FROM(
SELECT PATIENT_NAME, VISIT_DATE, COUNT(*) AS NO_OF_VISITS, RANK() OVER(ORDER BY COUNT(*) DESC) AS RN
FROM PATIENTS
JOIN VISITS
USING(PATIENT_ID)
GROUP BY PATIENT_ID, VISIT_DATE) AS SUB WHERE RN = 1;
```

id | Filter Rows: | Export: | Wrap Cell Content:

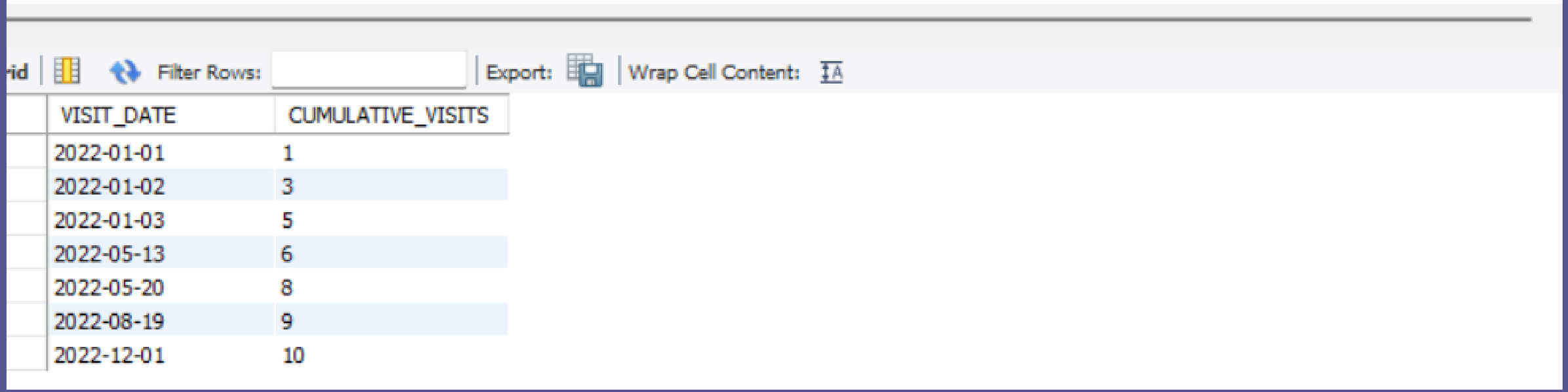
PATIENT_NAME	VISIT_DATE	NO_OF_VISITS
Mike Johnson	2022-05-20	2

```sql
-- 9. Write a SQL query to retrieve the average age of patients for each diagnosis, ordered by the average age in descending order.

WITH TEMP AS(
SELECT DISTINCT DIAGNOSIS_ID, PATIENT_ID, AGE FROM VISITS
JOIN PATIENTS USING(PATIENT_ID) ORDER BY DIAGNOSIS_ID)

SELECT DIAGNOSIS_NAME, ROUND(AVG(AGE),2) AS AVG_PATIENT_AGE FROM TEMP
JOIN DIAGNOSES USING(DIAGNOSIS_ID) GROUP BY DIAGNOSIS_ID ORDER BY 2 DESC;
```

Filter Rows: Export: Wrap Cell Content:

DIAGNOSIS_NAME	AVG_PATIENT_AGE
COVID-19	52.50
Pneumonia	50.00
Influenza	45.00
Common Cold	42.33
Bronchitis	30.00

```sql
-- 10. Write a SQL query to calculate the cumulative count of visits over time, ordered by the visit date.

SELECT VISIT_DATE, SUM(count(*)) OVER(ORDER BY VISIT_DATE) CUMULATIVE_VISITS FROM VISITS GROUP BY VISIT_DATE;
```



VISIT_DATE	CUMULATIVE_VISITS
2022-01-01	1
2022-01-02	3
2022-01-03	5
2022-05-13	6
2022-05-20	8
2022-08-19	9
2022-12-01	10