MIMIC Dataset SQL Query

MIMIC-II SQL Queries

1. Count the number of patients

```
SELECT COUNT(*)
FROM d_patients;
```

2. List unique patient genders

```
SELECT DISTINCT sex
FROM d_patients;
```

3. Count female patients

```
SELECT COUNT(*)
FROM d_patients
WHERE sex = 'F';
```

4. Count patients by gender

```
SELECT sex, COUNT(*)
FROM d_patients
GROUP BY sex;
```

5. Count hospital deaths

```
SELECT hospital_expire_flg, COUNT(*)
FROM d_patients
GROUP BY hospital_expire_flg;
```

6. Link patients with their admissions

```
SELECT p.subject_id, p.dob, a.hadm_id, a.admit_dt, p.hospital_expire_flg
FROM admissions a, d_patients p
WHERE p.subject_id = a.subject_id;
```

7. Get the first admission date for each patient

```
SELECT DISTINCT p.subject_id, p.dob, a.hadm_id, a.admit_dt, p.hospital_expire_flg,
        MIN(a.admit_dt) OVER (PARTITION BY p.subject_id) AS first_adm_dt
```

```
FROM admissions a, d_patients p
WHERE p.subject_id = a.subject_id
AND p.dob IS NOT NULL
ORDER BY a.hadm_id, p.subject_id;
```

8. Calculate age at first admission and classify into groups

```
WITH first_admission_date AS (
   SELECT DISTINCT p.subject_id, p.dob, p.sex,
            a.hadm_id, a.admit_dt,
            MIN(a.admit_dt) OVER (PARTITION BY a.hadm_id, p.subject_id) AS first_adm_dt
   FROM admissions a, d_patients p
   WHERE p.subject_id = a.subject_id
   AND p.dob IS NOT NULL
   ORDER BY a.hadm_id, p.subject_id
),
age AS (
   SELECT subject_id, hadm_id, dob, sex, first_adm_dt,
            ROUND(MONTHS_BETWEEN(first_adm_dt, dob) / 12, 2) AS first_adm_age,
            CASE
               WHEN (MONTHS_BETWEEN(first_adm_dt, dob) / 12) >= 15 THEN 'adult'
               WHEN MONTHS_BETWEEN(first_adm_dt, dob) <= 1 THEN 'neonate'
               ELSE 'middle'
            END AS age_group
   FROM first_admission_date
   ORDER BY subject_id, hadm_id
)
SELECT * FROM age;
```

9. ICU stay movement (from censusevents table)

```
SELECT *
FROM censusevents;
```

10. Join ICU movements with careunit names

```
SELECT ce.census_id, ce.subject_id, ce.intime, ce.outtime,
        ce.careunit, cu.label AS careunit_name,
        ce.destcareunit, dcu.label AS destunit_name,
        ce.dischstatus, ce.los
FROM censusevents ce
JOIN d_careunits cu ON ce.careunit = cu.cuid
JOIN d_careunits dcu ON ce.destcareunit = dcu.cuid;
```

11. Explore ICU stay summary view (icustay_detail)

```
SELECT *
FROM icustay_detail;
```

12. Select adult ICU patients

```
SELECT *
FROM icustay_detail
WHERE icustay_age_group = 'adult';
```

13. Calculate adult ICU mortality

```
SELECT icustay_expire_flg, COUNT(icustay_expire_flg)
FROM icustay_detail
WHERE icustay_age_group = 'adult'
GROUP BY icustay_expire_flg;
```

14. Challenge Problem: Find patients with MAP < 60 mmHg

```
SELECT COUNT(DISTINCT subject_id)
FROM vital_signs_raw
WHERE parameter = 'MAP'
AND value < 60;
```

MIMIC-III SQL Queries

1. Count the number of patients

```
SELECT COUNT(*)
FROM patients;
```

2. List unique patient genders

```
SELECT DISTINCT gender
FROM patients;
```

3. Count female patients

```
SELECT COUNT(*)
FROM patients
WHERE gender = 'F';
```

4. Count patients by gender

```
SELECT gender, COUNT(*)
FROM patients
GROUP BY gender;
```

5. Count hospital deaths

```
SELECT expire_flag, COUNT(*)
FROM patients
GROUP BY expire_flag;
```

6. Link patients with their admissions

```
SELECT p.subject_id, p.dob, a.hadm_id, a.admittime, p.expire_flag
FROM admissions a
INNER JOIN patients p ON p.subject_id = a.subject_id;
```

7. Get the first admission date

```
SELECT p.subject_id, p.dob, a.hadm_id, a.admittime, p.expire_flag,
        MIN(a.admittime) OVER (PARTITION BY p.subject_id) AS first_admittime
FROM admissions a
INNER JOIN patients p ON p.subject_id = a.subject_id;
```

8. Calculate age at first admission and classify

```
WITH first_admission AS (
   SELECT p.subject_id, p.dob,
```

```
        MIN(a.admittime) AS first_admittime
    FROM admissions a
    INNER JOIN patients p ON p.subject_id = a.subject_id
    GROUP BY p.subject_id, p.dob
)
SELECT subject_id, dob, first_admittime,
        ROUND((CAST(first_admittime AS DATE) - CAST(dob AS DATE)) / 365.242, 2) AS
first_admit_age,
        CASE
          WHEN ROUND((CAST(first_admittime AS DATE) - CAST(dob AS DATE)) / 365.242, 2) >= 15
THEN 'adult'
          WHEN ROUND((CAST(first_admittime AS DATE) - CAST(dob AS DATE)) / 365.242, 2) <= 1
THEN 'neonate'
          ELSE 'middle'
        END AS age_group
FROM first_admission;
```

9. ICU movements

```
SELECT *
FROM transfers;
```

10. ICU stay basic info

```
SELECT *
FROM icustays;
```

11. Adult ICU patients

```
SELECT *
FROM icustays
WHERE los >= 1;
```

12. ICU mortality calculation

```
SELECT CASE WHEN adm.deathtime BETWEEN icu.intime AND icu.outtime THEN 1 ELSE 0 END AS
died_in_icu,
        COUNT(*)
FROM icustays icu
```

```
JOIN admissions adm ON icu.hadm_id = adm.hadm_id
GROUP BY died_in_icu;
```

MIMIC-IV SQL Queries

1. Count the number of patients

```
SELECT COUNT(*)
FROM mimiciv_hosp.patients;
```

2. List unique patient genders

```
SELECT DISTINCT gender
FROM mimiciv_hosp.patients;
```

3. Count female patients

```
SELECT COUNT(*)
FROM mimiciv_hosp.patients
WHERE gender = 'F';
```

4. Count patients by gender

```
SELECT gender, COUNT(*)
FROM mimiciv_hosp.patients
GROUP BY gender;
```

5. Count hospital deaths

```
SELECT hospital_expire_flag, COUNT(*)
FROM mimiciv_hosp.admissions
GROUP BY hospital_expire_flag;
```

6. Link patients with their admissions

```
SELECT p.subject_id, p.dob, a.hadm_id, a.admittime, a.hospital_expire_flag
FROM mimiciv_hosp.admissions a
INNER JOIN mimiciv_hosp.patients p ON p.subject_id = a.subject_id;
```

7. Get the first admission date

```
SELECT p.subject_id, p.dob, a.hadm_id, a.admittime, a.hospital_expire_flag,
        MIN(a.admittime) OVER (PARTITION BY p.subject_id) AS first_admittime
FROM mimiciv_hosp.admissions a
INNER JOIN mimiciv_hosp.patients p ON p.subject_id = a.subject_id;
```

8. Calculate age at first admission and classify

```
WITH first_admission AS (
   SELECT p.subject_id, p.dob,
          MIN(a.admittime) AS first_admittime
   FROM mimiciv_hosp.admissions a
   INNER JOIN mimiciv_hosp.patients p ON p.subject_id = a.subject_id
   GROUP BY p.subject_id, p.dob
)
SELECT subject_id, dob, first_admittime,
        ROUND((CAST(first_admittime AS DATE) - CAST(dob AS DATE)) / 365.242, 2) AS
first_admit_age,
        CASE
          WHEN ROUND((CAST(first_admittime AS DATE) - CAST(dob AS DATE)) / 365.242, 2) >= 15
THEN 'adult'
          WHEN ROUND((CAST(first_admittime AS DATE) - CAST(dob AS DATE)) / 365.242, 2) <= 1
THEN 'neonate'
          ELSE 'middle'
        END AS age_group
FROM first_admission;
```

9. ICU movements

```
SELECT *
FROM mimiciv_hosp.transfers;
```

10. ICU stay basic info

```
SELECT *
FROM mimiciv_icu.icustays;
```

11. Adult ICU patients

```
SELECT *

FROM mimiciv_icu.icustays

WHERE los >= 1;
```

12. ICU mortality calculation

```
SELECT CASE WHEN adm.deathtime BETWEEN icu.intime AND icu.outtime THEN 1 ELSE 0 END AS
died_in_icu,

        COUNT(*)

FROM mimiciv_icu.icustays icu

JOIN mimiciv_hosp.admissions adm ON icu.hadm_id = adm.hadm_id

GROUP BY died_in_icu;
```

MIMIC-CXR SQL Queries

1. Count the number of patients

```
SELECT COUNT(DISTINCT subject_id)

FROM mimic_cxr_meta;
```

2. List unique patient genders

```
SELECT DISTINCT p.gender

FROM mimic_cxr_meta m

JOIN mimiciv_hosp.patients p ON m.subject_id = p.subject_id;
```

3. Count female patients

```
SELECT COUNT(DISTINCT m.subject_id)

FROM mimic_cxr_meta m

JOIN mimiciv_hosp.patients p ON m.subject_id = p.subject_id

WHERE p.gender = 'F';
```

4. Count patients by gender

```
SELECT p.gender, COUNT(DISTINCT m.subject_id) AS num_patients

FROM mimic_cxr_meta m
```

```
JOIN mimiciv_hosp.patients p ON m.subject_id = p.subject_id
GROUP BY p.gender;
```

5. Count hospital deaths

```
SELECT a.hospital_expire_flag, COUNT(DISTINCT m.subject_id)
FROM mimic_cxr_meta m
JOIN mimiciv_hosp.admissions a ON m.subject_id = a.subject_id
GROUP BY a.hospital_expire_flag;
```

6. Get first chest X-ray study date

```
SELECT subject_id, MIN(study_date) AS first_cxr_date
FROM mimic_cxr_meta
GROUP BY subject_id;
```

7. Calculate age at first chest X-ray

```
WITH first_cxr AS (
   SELECT subject_id, MIN(study_date) AS first_study_date
   FROM mimic_cxr_meta
   GROUP BY subject_id
)
SELECT f.subject_id,
        ROUND(DATETIME_DIFF(f.first_study_date, p.dob, DAY) / 365.242, 2) AS age_at_cxr
FROM first_cxr f
JOIN mimiciv_hosp.patients p ON f.subject_id = p.subject_id;
```

8. Adult mortality (age ≥ 15)

```
WITH first_cxr AS (
   SELECT subject_id, MIN(study_date) AS first_study_date
   FROM mimic_cxr_meta
   GROUP BY subject_id
),
age_info AS (
   SELECT f.subject_id,
```

```
        ROUND(DATETIME_DIFF(f.first_study_date, p.dob, DAY) / 365.242, 2) AS age,
        a.hospital_expire_flag
    FROM first_cxr f
    JOIN mimiciv_hosp.patients p ON f.subject_id = p.subject_id
    JOIN mimiciv_hosp.admissions a ON f.subject_id = a.subject_id
)
SELECT hospital_expire_flag, COUNT(*)
FROM age_info
WHERE age >= 15
GROUP BY hospital_expire_flag;
```

9. Count number of patients with chest X-rays

```
SELECT COUNT(DISTINCT subject_id)
FROM mimic_cxr_meta;
```

10. Total number of X-ray studies

```
SELECT COUNT(DISTINCT study_id)
FROM mimic_cxr_meta;
```

11. Number of X-rays labeled with pneumonia

```
SELECT COUNT(*)
FROM mimic_cxr_chexpert
WHERE label = 'Pneumonia';
```

12. Patients with at least one pneumonia-positive X-ray

```
SELECT COUNT(DISTINCT m.subject_id)
FROM mimic_cxr_chexpert c
JOIN mimic_cxr_meta m ON c.study_id = m.study_id
WHERE c.label = 'Pneumonia';
```

13. Average number of X-rays per patient

```
SELECT AVG(xr_count) AS avg_xrays_per_patient
FROM (
   SELECT subject_id, COUNT(*) AS xr_count
   FROM mimic_cxr_meta
   GROUP BY subject_id
) AS sub;
```

14. Most frequent X-ray view positions

```
SELECT view_position, COUNT(*) AS num_studies
FROM mimic_cxr_meta
GROUP BY view_position
ORDER BY num_studies DESC;
```

15. First chest X-ray date per patient

```
SELECT subject_id, MIN(study_date) AS first_cxr_date
FROM mimic_cxr_meta
GROUP BY subject_id;
```

16. Distribution of CheXpert labels across all studies

```
SELECT label, COUNT(*) AS count
FROM mimic_cxr_chexpert
GROUP BY label
ORDER BY count DESC;
```

17. Average age of patients at first chest X-ray

```
WITH first_cxr AS (
```

```
    SELECT subject_id, MIN(study_date) AS first_date
    FROM mimic_cxr_meta
    GROUP BY subject_id
)
SELECT ROUND(AVG(DATETIME_DIFF(f.first_date, p.dob, DAY) / 365.242), 1) AS avg_age_at_cxr
FROM first_cxr f
JOIN mimiciv_hosp.patients p ON f.subject_id = p.subject_id;
```

18. Link study, image, and view metadata

```
SELECT i.subject_id, i.study_id, i.dicom_id, m.study_date, m.view_position
FROM mimic_cxr_images i
JOIN mimic_cxr_meta m ON i.study_id = m.study_id
LIMIT 100;
```

MIMIC-IV-ED SQL Queries

1. Count the number of patients

```
SELECT COUNT(DISTINCT subject_id)
FROM mimiciv_ed.edstays;
```

2. List unique patient genders

```
SELECT DISTINCT p.gender
FROM mimiciv_ed.edstays e
JOIN mimiciv_hosp.patients p ON e.subject_id = p.subject_id;
```

3. Count female patients

```
SELECT COUNT(DISTINCT e.subject_id)
FROM mimiciv_ed.edstays e
JOIN mimiciv_hosp.patients p ON e.subject_id = p.subject_id
WHERE p.gender = 'F';
```

4. Count patients by gender

```
SELECT p.gender, COUNT(DISTINCT e.subject_id) AS num_patients
FROM mimiciv_ed.edstays e
JOIN mimiciv_hosp.patients p ON e.subject_id = p.subject_id
GROUP BY p.gender;
```

5. Count hospital deaths

```
SELECT a.hospital_expire_flag, COUNT(DISTINCT e.subject_id)
FROM mimiciv_ed.edstays e
JOIN mimiciv_hosp.admissions a ON e.hadm_id = a.hadm_id
GROUP BY a.hospital_expire_flag;
```

6. Get first ED visit datetime

```
SELECT subject_id, MIN(intime) AS first_ed_visit
FROM mimiciv_ed.edstays
GROUP BY subject_id;
```

7. Calculate age at first ED visit

```
WITH first_ed AS (
    SELECT subject_id, MIN(intime) AS first_ed_time
    FROM mimiciv_ed.edstays
    GROUP BY subject_id
)
SELECT f.subject_id,
        ROUND(DATETIME_DIFF(f.first_ed_time, p.dob, DAY) / 365.242, 2) AS age_at_ed
FROM first_ed f
JOIN mimiciv_hosp.patients p ON f.subject_id = p.subject_id;
```

8. Adult mortality (age ≥ 15)

```
WITH first_ed AS (
    SELECT subject_id, MIN(intime) AS first_ed_time
    FROM mimiciv_ed.edstays
    GROUP BY subject_id
),
age_info AS (
    SELECT f.subject_id,
            ROUND(DATETIME_DIFF(f.first_ed_time, p.dob, DAY) / 365.242, 2) AS age,
            a.hospital_expire_flag
    FROM first_ed f
    JOIN mimiciv_hosp.patients p ON f.subject_id = p.subject_id
    JOIN mimiciv_hosp.admissions a ON f.subject_id = a.subject_id
)
SELECT hospital_expire_flag, COUNT(*)
FROM age_info
WHERE age >= 15
GROUP BY hospital_expire_flag;
```

9. Number of unique ED patients

```
SELECT COUNT(DISTINCT subject_id) AS unique_ed_patients
FROM mimiciv_ed.edstays;
```

10. Average length of ED stay (in hours)

```
SELECT ROUND(AVG(TIMESTAMP_DIFF(outtime, intime, MINUTE)) / 60, 2) AS avg_los_hours
FROM mimiciv_ed.edstays;
```

11. Distribution of ED diagnoses

```
SELECT diagnosis, COUNT(*) AS count
FROM mimiciv_ed.edstays
GROUP BY diagnosis
ORDER BY count DESC
LIMIT 20;
```

12. Vital sign: proportion of patients with hypotension (MAP < 60)

```
SELECT COUNT(DISTINCT stay_id) AS hypotensive_cases
FROM mimiciv_ed.vitalsign
WHERE map < 60;
```

13. Triage level distribution

```
SELECT triage_level, COUNT(*) AS count
FROM mimiciv_ed.triage
GROUP BY triage_level
ORDER BY count DESC;
```

14. Number of patients admitted to hospital from ED

```
SELECT COUNT(DISTINCT e.subject_id)
FROM mimiciv_ed.edstays e
```

```
JOIN mimiciv_hosp.admissions a ON e.hadm_id = a.hadm_id;
```

15. Patients returning to ED within 72 hours

```
WITH ed_visits AS (
   SELECT subject_id, intime,
          LEAD(intime) OVER (PARTITION BY subject_id ORDER BY intime) AS next_intime
   FROM mimiciv_ed.edstays
)
SELECT COUNT(*) AS return_visits_within_72h
FROM ed_visits
WHERE TIMESTAMP_DIFF(next_intime, intime, HOUR) <= 72;
```

16. Hour-of-day distribution of ED arrivals

```
SELECT EXTRACT(HOUR FROM intime) AS hour,
       COUNT(*) AS visits
FROM mimiciv_ed.edstays
GROUP BY hour
ORDER BY hour;
```

17. Age distribution at first ED visit

```
WITH first_ed AS (
   SELECT subject_id, MIN(intime) AS first_ed_time
   FROM mimiciv_ed.edstays
   GROUP BY subject_id
)
SELECT ROUND(DATETIME_DIFF(f.first_ed_time, p.dob, DAY)/365.242, 0) AS age,
       COUNT(*) AS num_patients
FROM first_ed f
```

```
JOIN mimiciv_hosp.patients p ON f.subject_id = p.subject_id

GROUP BY age

ORDER BY age;
```