



# SQL Queries for Analyzing Diabetes Using the MIMIC Database

Jesus Minjares

UT AUSTIN: AI in HEALTHCARE

# Query 1: Diabetes-Related Hospital Admissions

## Query

```
SELECT COUNT(*) AS diabetes_admissions
FROM ADMISSIONS
JOIN DIAGNOSES_ICD ON ADMISSIONS.hadm_id = DIAGNOSES_ICD.hadm_id
WHERE DIAGNOSES_ICD.icd9_code LIKE '250%';
```

## Result

diabetes_admissions
11742

## Meaning

- Counts the number of hospital admissions related to diabetes.

## Explanation

- This query filters hospital admissions where the diagnosis code starts with 250, the ICD-9 code for diabetes.
- The output provides the total count of diabetes-related admissions.

# Query 2: Most Common Diabetes Diagnoses

## Query

```
SELECT D_ICD_DIAGNOSES.long_title, COUNT(*) AS diagnosis_count
FROM DIAGNOSES_ICD
JOIN D_ICD_DIAGNOSES ON DIAGNOSES_ICD.icd9_code = D_ICD_DIAGNOSES.icd9_code
WHERE DIAGNOSES_ICD.icd9_code LIKE '250%'
GROUP BY D_ICD_DIAGNOSES.long_title
ORDER BY diagnosis_count DESC
LIMIT 5;
```

## Result

	LONG_TITLE	diagnosis_count
0	Diabetes mellitus without mention of complication, type II or unspecified type, not stated as uncontrolled	9058
1	Diabetes with neurological manifestations, type II or unspecified type, not stated as uncontrolled	1138
2	Diabetes with renal manifestations, type II or unspecified type, not stated as uncontrolled	954
3	Diabetes mellitus without mention of complication, type II or unspecified type, uncontrolled	454
4	Diabetes with other specified manifestations, type II or unspecified type, not stated as uncontrolled	447

## Meaning

- Identifies the most frequently diagnosed diabetes conditions.

## Explanation

- The query joins diagnosis codes with their descriptions.
- It counts occurrences of each diagnosis and returns the top 5 most common ones.

# Query 3: Average Length of Stay for Diabetes Patients

## Query

```
SELECT AVG(EXTRACT(DAY FROM (disctime - admittime))) AS avg_length_of_stay
FROM ADMISSIONS
JOIN DIAGNOSES_ICD ON ADMISSIONS.hadm_id = DIAGNOSES_ICD.hadm_id
WHERE DIAGNOSES_ICD.icd9_code LIKE '250%';
```

## Result

```
avg_length_of_stay
0                9.245784
```

## Meaning

- Calculates the average length of hospital stays for diabetes patients.

## Explanation

- The time difference between discharge and admission is extracted in days.
- The average length of stay for diabetes patients is computed.

# Query 4: Gender Distribution of Diabetes Patients

## Query

```
SELECT PATIENTS.gender, COUNT(*) AS count
FROM PATIENTS
JOIN DIAGNOSES_ICD ON PATIENTS.subject_id = DIAGNOSES_ICD.subject_id
WHERE DIAGNOSES_ICD.icd9_code LIKE '250%'
GROUP BY PATIENTS.gender;
```

## Result

	GENDER	count
0	F	7280
1	M	9174

## Meaning

- Shows the gender distribution of diabetes patients

## Explanation

- This query groups diabetes patients by gender and counts the occurrences.
- It helps understand if diabetes is more common in a specific gender.



# Query 5: Top 3 Prescribed Medications for Diabetes

## Query

```
SELECT PRESCRIPTIONS.drug, COUNT(*) AS prescription_count
FROM PRESCRIPTIONS
JOIN ADMISSIONS ON PRESCRIPTIONS.hadm_id = ADMISSIONS.hadm_id
JOIN DIAGNOSES_ICD ON ADMISSIONS.hadm_id = DIAGNOSES_ICD.hadm_id
WHERE DIAGNOSES_ICD.icd9_code LIKE '250%'
GROUP BY PRESCRIPTIONS.drug
ORDER BY prescription_count DESC
LIMIT 3;
```

## Result

	DRUG	prescription_count
0	Insulin	89803
1	Potassium Chloride	39150
2	0.9% Sodium Chloride	36670

## Meaning

- Identifies the most prescribed medications for diabetes patients.

## Explanation

- The query counts how often each drug is prescribed for diabetes patients.
- The top 3 most commonly prescribed medications are listed.

# Query 6: Hospital Mortality Rate for Diabetes Patients

## Query

```
SELECT (SUM(hospital_expire_flag) * 100.0 / COUNT(*)) AS mortality_rate
FROM ADMISSIONS
JOIN DIAGNOSES_ICD ON ADMISSIONS.hadm_id = DIAGNOSES_ICD.hadm_id
WHERE DIAGNOSES_ICD.icd9_code LIKE '250%';
```

## Result

```
mortality_rate
0          9.930165
```

## Meaning

- Calculates the percentage of diabetes patients who died in the hospital.

## Explanation

- The query checks how many diabetes patients had **hospital\_expire\_flag = 1**.
- The percentage of diabetes patients who passed away during hospitalization is calculated.

# Query 7: Most Common Insurance Types for Diabetes Patients

## Query

```
SELECT insurance, COUNT(*) AS count
FROM ADMISSIONS
JOIN DIAGNOSES_ICD ON ADMISSIONS.hadm_id = DIAGNOSES_ICD.hadm_id
WHERE DIAGNOSES_ICD.icd9_code LIKE '250%'
GROUP BY insurance
ORDER BY count DESC;
```

## Result

	INSURANCE	count
0	Medicare	7462
1	Private	2906
2	Medicaid	1062
3	Government	251
4	Self Pay	61

## Meaning

- Lists the most common insurance types among diabetes patients.

## Explanation

- The query groups diabetes patients by insurance type and counts occurrences.
- It helps in understanding the coverage patterns for diabetes patients.



# Query 8: ICU Admissions for Diabetes Patients

## Query

```
SELECT COUNT(*) AS icu_admissions
FROM ICUSTAYS
JOIN ADMISSIONS ON ICUSTAYS.hadm_id = ADMISSIONS.hadm_id
JOIN DIAGNOSES_ICD ON ADMISSIONS.hadm_id = DIAGNOSES_ICD.hadm_id
WHERE DIAGNOSES_ICD.icd9_code LIKE '250%';
```

## Result

	icu_admissions
0	12378

## Meaning

- Counts how many diabetes patients were admitted to the ICU.

## Explanation

- The query checks how many diabetes-related hospitalizations led to ICU stays.
- This helps assess the severity of diabetes cases requiring intensive care.

# Query 9: Most Common Procedures for Diabetes Patients

## Query

```
SELECT D_ICD_PROCEDURES.long_title, COUNT(*) AS procedure_count
FROM PROCEDURES_ICD
JOIN D_ICD_PROCEDURES ON PROCEDURES_ICD.icd9_code = D_ICD_PROCEDURES.icd9_code
JOIN DIAGNOSES_ICD ON PROCEDURES_ICD.hadm_id = DIAGNOSES_ICD.hadm_id
WHERE DIAGNOSES_ICD.icd9_code LIKE '250%'
GROUP BY D_ICD_PROCEDURES.long_title
ORDER BY procedure_count DESC
LIMIT 5;
```

## Result

	LONG_TITLE	procedure_count
0	Venous catheterization, not elsewhere classified	5129
1	Insertion of endotracheal tube	2660
2	Transfusion of packed cells	2357
3	Continuous invasive mechanical ventilation for less than 96 consecutive hours	2335
4	Extracorporeal circulation auxiliary to open heart surgery	2333

## Meaning

- Identifies the most common medical procedures performed on diabetes patients.

## Explanation

- This query retrieves the most frequently conducted medical procedures for diabetes patients.
- The output ranks the top 5 procedures performed on hospitalized diabetes patients.

# Query 10: Patients with Multiple Diabetes-Related Hospitalizations

## Query

```
SELECT diabetes_patients.subject_id, COUNT(diabetes_patients.hadm_id) AS readmission_count
FROM (
  SELECT ADMISSIONS.subject_id, ADMISSIONS.hadm_id
  FROM ADMISSIONS
  JOIN DIAGNOSES_ICD ON ADMISSIONS.hadm_id = DIAGNOSES_ICD.hadm_id
  WHERE DIAGNOSES_ICD.icd9_code LIKE '250%'
) AS diabetes_patients
GROUP BY diabetes_patients.subject_id
HAVING COUNT(diabetes_patients.hadm_id) > 1
ORDER BY readmission_count DESC;
```

## Result

	SUBJECT_ID	readmission_count
0	13033	66
1	11861	56
2	7809	47
3	58163	32
4	8799	31

## Meaning

- Identifies patients with multiple diabetes-related hospital admissions.

## Explanation

- The query finds diabetes patients who have been hospitalized multiple times.
- It counts how many times each patient was admitted and returns those with more than one admission.

# Conclusion

**11,742 diabetes-related admissions**, with an **average hospital stay of 9.25 days**.

**High-risk cases:** 12,378 ICU admissions and **9.93% mortality rate**.

**Insulin** is the most prescribed medication, indicating its central role in diabetes treatment.

**Most common procedures** involve catheterization, intubation, and blood transfusions.

**Gender split:** More males (9,174) than females (7,280) admitted.

**Frequent readmissions**, with some patients hospitalized up to **66 times**