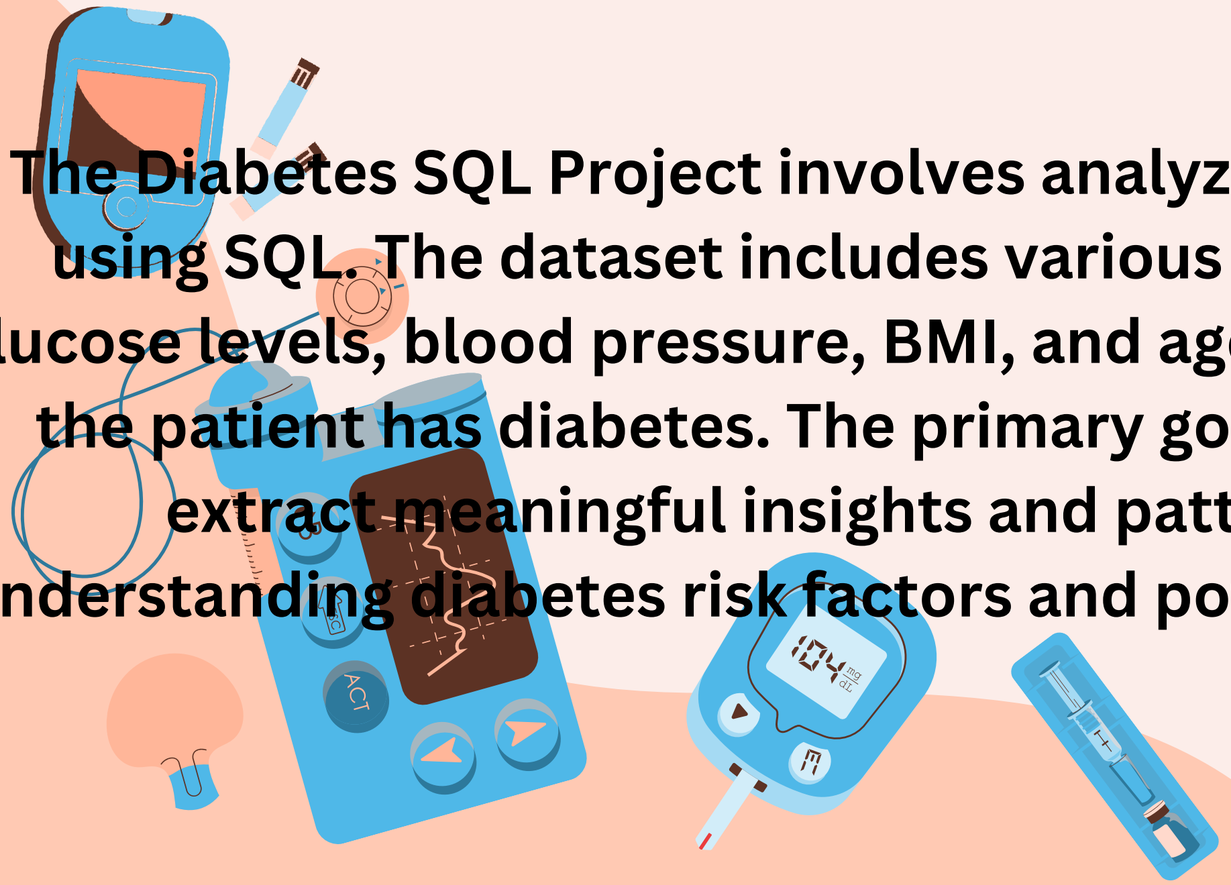
Introduction to Diabetes SQL Project

Project Overview

The Diabetes SQL Project involves analyzing a dataset related to diabetes patients using SQL. The dataset includes various health metrics for each patient, such as glucose levels, blood pressure, BMI, and age, along with an indicator of whether or not the patient has diabetes. The primary goal of this project is to use SQL queries to extract meaningful insights and patterns from the data, which can aid in understanding diabetes risk factors and potentially improve diagnosis and treatment.



Count the Total Number of Records

Write a query to count the total number of records in the table.

```
SELECT COUNT(*) AS total_records
FROM diabetes;
```

	total_records
1	768

Average Age of Patients

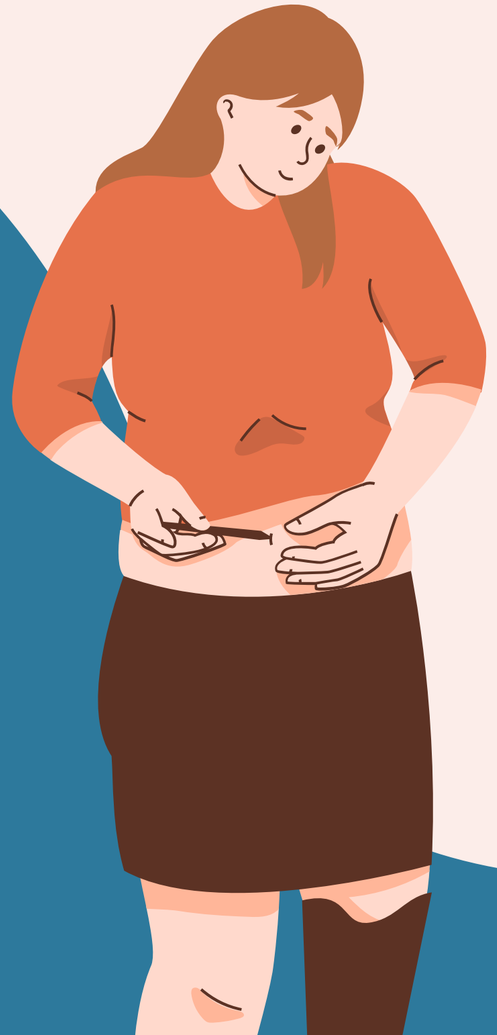
```sql
SELECT AVG(Age) AS average_age
FROM diabetes;
```

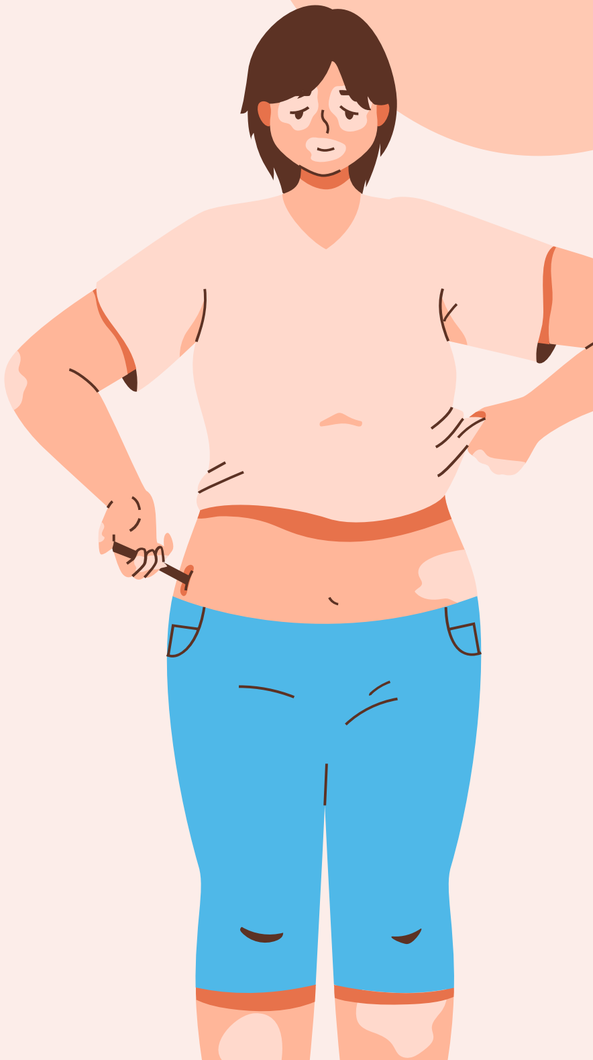
	average_age
1	33

Maximum Glucose Level

```
SELECT MAX(Glucose) AS max_glucose
FROM diabetes;
```

	max_glucose
1	199





Number of Patients with Diabetes

```sql
SELECT COUNT(*) AS num_diabetes_patients
FROM diabetes
WHERE Outcome = Outcome;
```

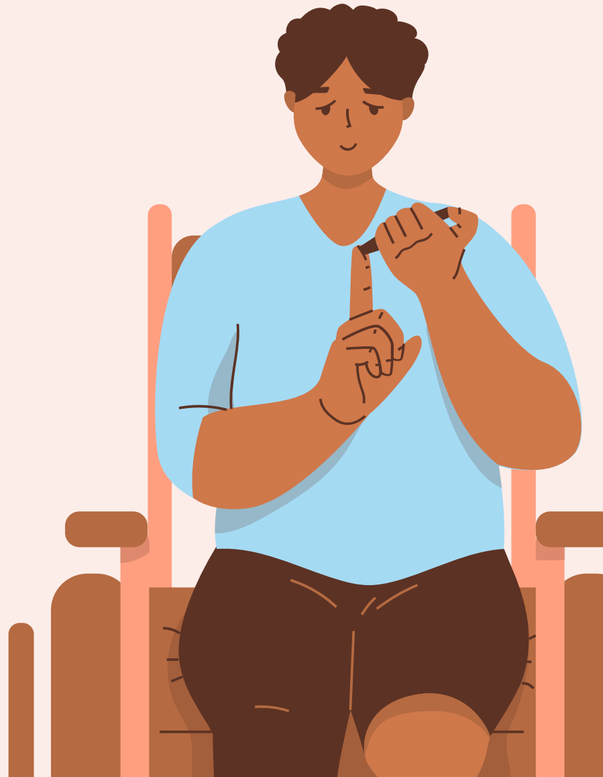
	num_diabetes_patients
1	768

Average BMI by Outcome

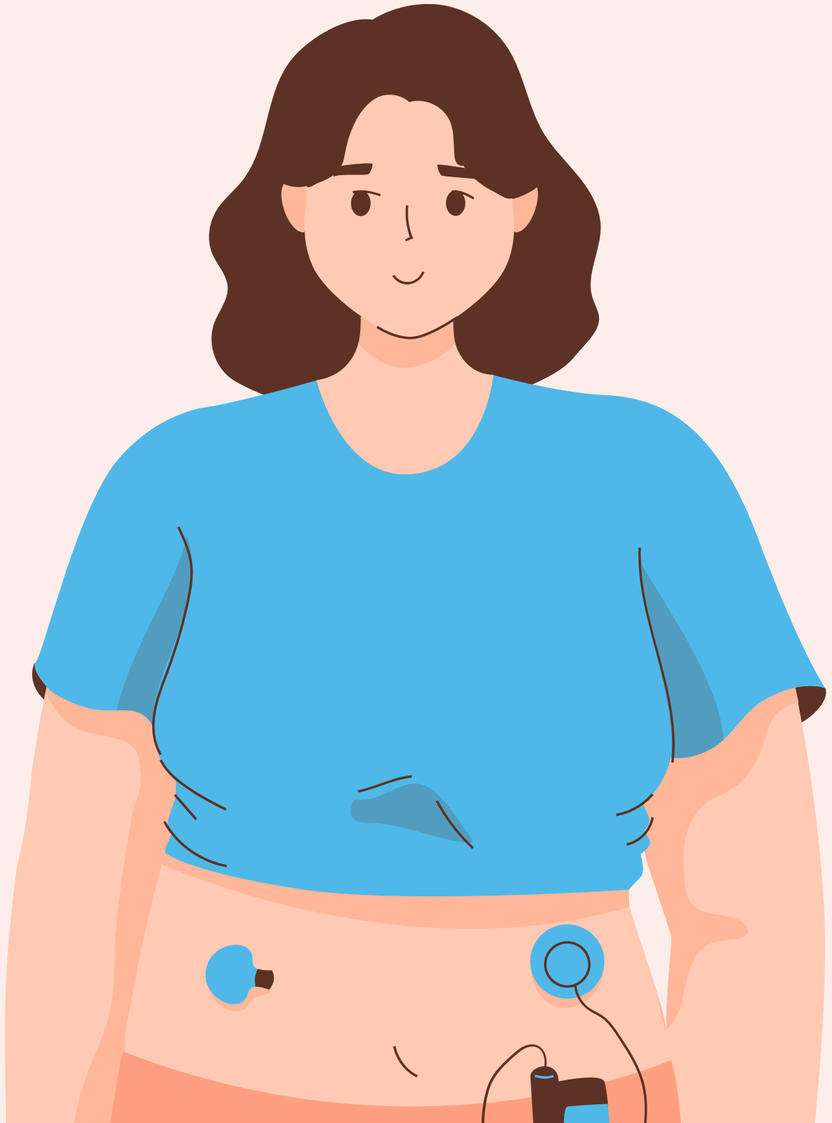
Write a query to find the average BMI for patients with and without diabetes.

```
SELECT Outcome, AVG(BMI) AS average_bmi
FROM diabetes
GROUP BY Outcome;
```

	Outcome	average_bmi
1	Non Diabetic	30.3042000083923
2	Diabetic	35.1425373376305



Patients Aged 30 and Above



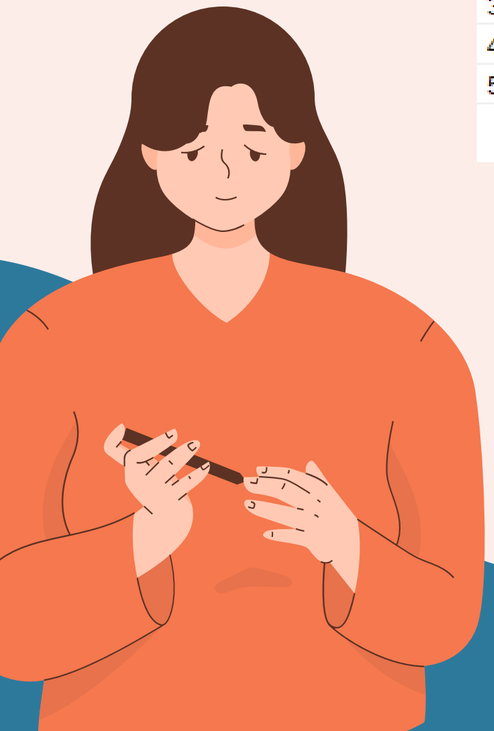
```sql
SELECT COUNT(*) AS patients_above_30
FROM diabetes
WHERE Age >= 30;
```

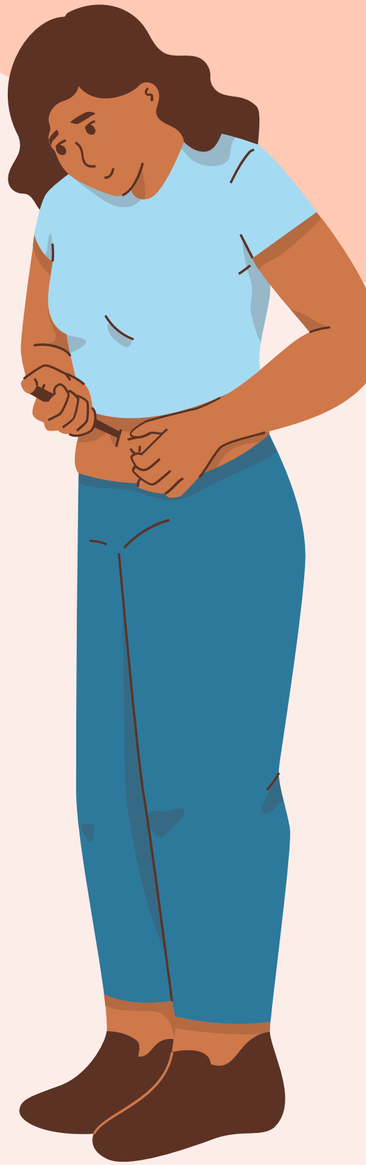
	patients_above_30
1	372

Top 5 Patients with Highest Insulin Levels

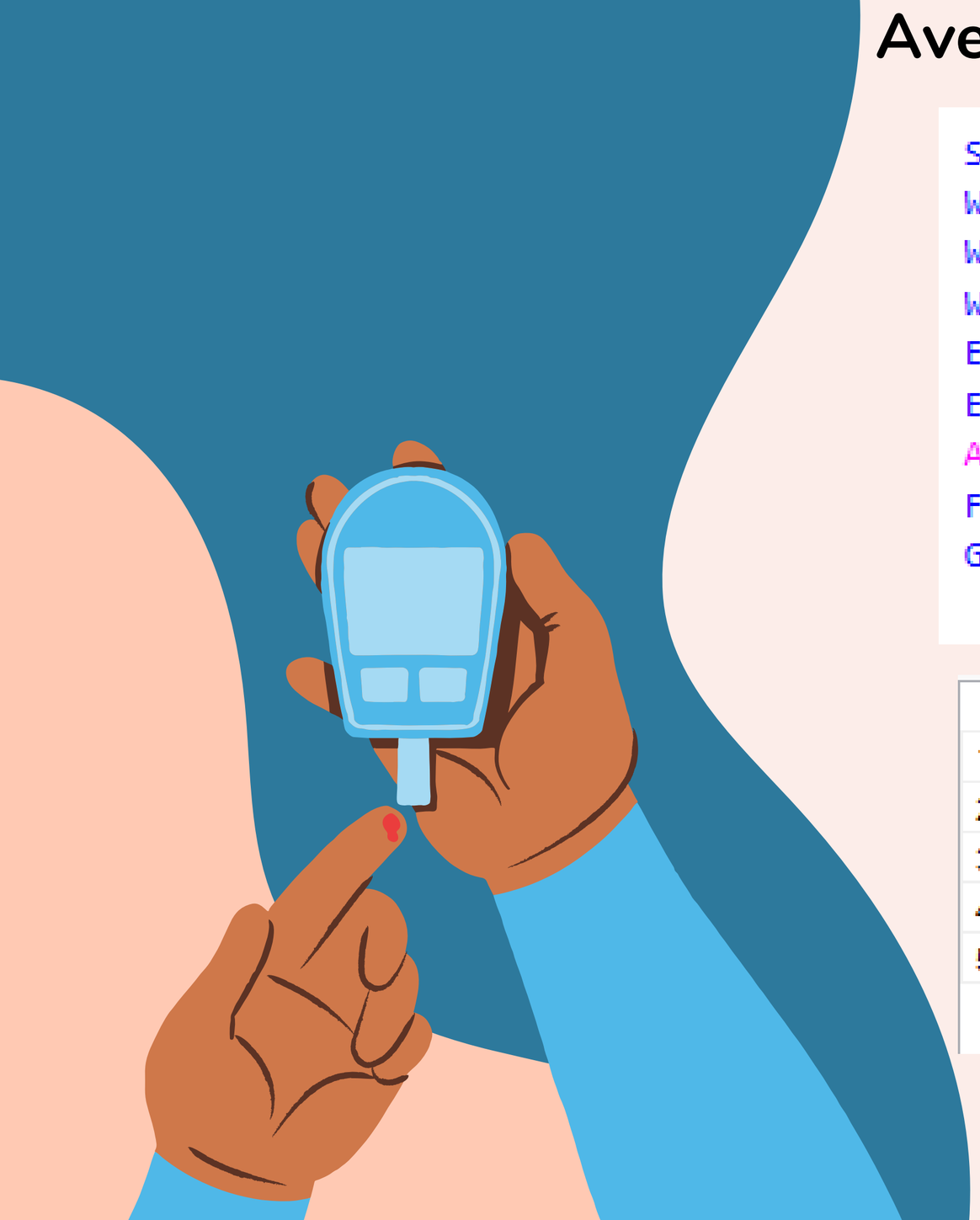
```
SELECT top 5 *
FROM diabetes
ORDER BY Insulin DESC;
```

	Pregnancies	Glucose	BloodPressure	SkinThickness	Insulin	BMI	DiabetesPedigreeFunction	Age	Outcome
1	1	189	60	23	846	30.1000003814697	0.398000001907349	59	Diabetic
2	4	197	70	39	744	36.7000007629395	2.3289999961853	31	Non Diabetic
3	0	165	90	33	680	52.2999992370605	0.426999986171722	23	Non Diabetic
4	8	124	76	24	600	28.7000007629395	0.686999976634979	52	Diabetic
5	1	172	68	49	579	42.4000015258789	0.702000021934509	28	Diabetic





Average Glucose Level by Age Group



```sql
SELECT top 5 CASE
WHEN Age BETWEEN 20 AND 29 THEN '20-29'
WHEN Age BETWEEN 30 AND 39 THEN '30-39'
WHEN Age BETWEEN 40 AND 49 THEN '40-49'
ELSE '50+'
END AS age_group,
AVG(Glucose) AS average_glucose
FROM diabetes
GROUP BY Age;
```

	age_group	average_glucose
1	20-29	111
2	40-49	105
3	50+	134
4	20-29	127
5	50+	133

Patients with High Glucose and Low BMI

```
SELECT * FROM diabetes
WHERE Glucose > 150 AND BMI < 25;
```

	Pregnancies	Glucose	BloodPressure	SkinThickness	Insulin	BMI	DiabetesPedigreeFunction	Age	Outcome
1	8	183	64	0	0	23.2999992370605	0.671999990940094	32	Diabetic
2	0	161	50	0	0	21.8999996185303	0.254000008106232	65	Non Diabetic
3	6	194	78	0	0	23.5	0.12899999320507	59	Diabetic
4	1	167	74	17	144	23.3999996185303	0.446999996900558	33	Diabetic
5	9	156	86	0	0	24.7999992370605	0.230000004172325	53	Diabetic
6	2	175	88	0	0	22.8999996185303	0.32600000500679	22	Non Diabetic
7	6	162	62	0	0	24.2999992370605	0.178000003099442	50	Diabetic