## 1. Retrieve the Patient_id and ages of all patients

```
SELECT Patient_id, DATEDIFF(YEAR, D#O#B, GETDATE()) AS Age
FROM diabetes_prediction.dbo.Dataset;
```

| | Patient_id | Age |
|---|---|---|
| 1 | PT2499 | 25 |
| 2 | PT2500 | 25 |
| 3 | PT2501 | 25 |
| 4 | PT2502 | 25 |
| 5 | PT2503 | 25 |
| 6 | PT2504 | 25 |
| 7 | PT2505 | 25 |
| 8 | PT2506 | 25 |
| 9 | PT2507 | 25 |
| 10 | PT2508 | 25 |
| 11 | PT2509 | 25 |
| 12 | PT2510 | 25 |
| 13 | PT2511 | 25 |
| 14 | PT2512 | 25 |
| 15 | PT2513 | 25 |
| 16 | PT2514 | 25 |
| 17 | PT2515 | 25 |
| 18 | PT2516 | 25 |
| 19 | PT2517 | 25 |
| 20 | PT2518 | 25 |
| 21 | PT2519 | 25 |
| 22 | PT2520 | 25 |
| 23 | PT2521 | 25 |
| 24 | PT2522 | 25 |
| 25 | PT2523 | 25 |
| 26 | PT2524 | 25 |
| 27 | PT2525 | 25 |
| 28 | PT2526 | 25 |
| 29 | PT2527 | 25 |

## 2. Select all female patients who are older than 25.

```
SELECT *FROM diabetes_prediction.dbo.Dataset
WHERE gender = 'Female' AND DATEDIFF(YEAR, [D#O#B], GETDATE()) > 25;
```

| | EmployeeName | Patient_id | gender | D#O#B | hypertension | heart_disease | smoking_history | bmi | HbA1c_level | blood_glucose_level | diabetes |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | NATHANIEL FORD | PT101 | Female | 1992-11-05 00:00:00.000 | 0 | 1 | never | 25.19 | 6.6 | 140 | 0 |
| 2 | GARY JIMENEZ | PT102 | Female | 1992-11-11 00:00:00.000 | 0 | 0 | No Info | 27.32 | 6.6 | 80 | 0 |
| 3 | CHRISTOPHER CHONG | PT104 | Female | 1992-12-05 00:00:00.000 | 0 | 0 | current | 23.45 | 5 | 155 | 0 |
| 4 | DAVID SULLIVAN | PT106 | Female | 1989-01-05 00:00:00.000 | 0 | 0 | never | 27.32 | 6.6 | 85 | 0 |
| 5 | ALSON LEE | PT107 | Female | 1989-01-23 00:00:00.000 | 0 | 0 | never | 19.31 | 6.5 | 200 | 1 |
| 6 | DAVID KUSHNER | PT108 | Female | 1989-02-05 00:00:00.000 | 0 | 0 | No Info | 23.86 | 5.7 | 85 | 0 |
| 7 | JOANNE HAYES-WHITE | PT110 | Female | 1989-03-09 00:00:00.000 | 0 | 0 | never | 27.32 | 5 | 100 | 0 |
| 8 | ARTHUR KENNEY | PT111 | Female | 1989-03-19 00:00:00.000 | 0 | 0 | never | 27.32 | 6.1 | 85 | 0 |
| 9 | PATRICIA JACKSON | PT112 | Female | 1989-04-01 00:00:00.000 | 0 | 0 | former | 54.7 | 6 | 100 | 0 |
| 10 | EDWARD HARRINGTON | PT113 | Female | 1989-04-14 00:00:00.000 | 0 | 0 | former | 36.05 | 5 | 130 | 0 |
| 11 | JOHN MARTIN | PT114 | Female | 1989-04-21 00:00:00.000 | 0 | 0 | never | 25.69 | 5.8 | 200 | 0 |
| 12 | DAVID FRANKLIN | PT115 | Female | 1989-04-26 00:00:00.000 | 0 | 0 | No Info | 27.32 | 5 | 160 | 0 |
| 13 | SEBASTIAN WONG | PT118 | Female | 1989-04-30 00:00:00.000 | 0 | 0 | never | 24.48 | 5.7 | 158 | 0 |
| 14 | MARTY ROSS | PT119 | Female | 1989-05-10 00:00:00.000 | 0 | 0 | No Info | 27.32 | 5.7 | 80 | 0 |
| 15 | GEORGE GARCIA | PT123 | Female | 1989-06-14 00:00:00.000 | 0 | 0 | never | 21.24 | 4.8 | 85 | 0 |
| 16 | VICTOR WYRSCH | PT124 | Female | 1989-06-17 00:00:00.000 | 0 | 1 | former | 27.94 | 6.5 | 130 | 0 |
| 17 | JOSEPH DRISCOLL | PT125 | Female | 1989-06-24 00:00:00.000 | 0 | 0 | No Info | 13.99 | 4 | 140 | 0 |
| 18 | HARLAN KELLY-JR | PT131 | Female | 1989-07-06 00:00:00.000 | 0 | 0 | No Info | 31.75 | 4 | 200 | 0 |
| 19 | GARY AMELIO | PT133 | Female | 1989-07-22 00:00:00.000 | 0 | 0 | current | 22.01 | 6.2 | 126 | 0 |
| 20 | JOHN TURSI | PT134 | Female | 1989-07-24 00:00:00.000 | 0 | 0 | never | 22.19 | 3.5 | 100 | 0 |
| 21 | JOSE VELO | PT135 | Female | 1989-07-30 00:00:00.000 | 0 | 0 | never | 23.55 | 5 | 85 | 0 |
| 22 | SUSAN CURRIN | PT137 | Female | 1989-08-04 00:00:00.000 | 0 | 0 | No Info | 21.76 | 4.5 | 130 | 0 |
| 23 | JAMES BOSCH | PT138 | Female | 1989-08-04 00:00:00.000 | 0 | 0 | never | 21.22 | 6.6 | 200 | 0 |
| 24 | BRENDAN WARD | PT140 | Female | 1989-08-14 00:00:00.000 | 0 | 0 | never | 56.43 | 6.2 | 200 | 0 |
| 25 | THOMAS SIRAGUSA | PT143 | Female | 1989-09-16 00:00:00.000 | 1 | 1 | never | 32.02 | 5 | 159 | 0 |
| 26 | MICHAEL THOMPSON | PT144 | Female | 1989-10-18 00:00:00.000 | 0 | 0 | No Info | 29.3 | 4.8 | 159 | 0 |
| 27 | SHARON MCCOLE WI... | PT145 | Female | 1989-10-25 00:00:00.000 | 0 | 0 | No Info | 27.32 | 3.5 | 160 | 0 |
| 28 | EDWIN LEE | PT146 | Female | 1989-10-26 00:00:00.000 | 0 | 0 | never | 24.93 | 6.1 | 100 | 0 |
| 29 | BRYAN RUBENSTEIN | PT147 | Female | 1989-10-31 00:00:00.000 | 0 | 0 | never | 19.95 | 5 | 90 | 0 |
| 30 | TRENT RHORER | PT148 | Female | 1989-11-08 00:00:00.000 | 0 | 0 | never | 18.03 | 4 | 159 | 0 |

## 3. Calculate the average BMI of patients.

```
SELECT AVG(bmi) AS Average_BMI
FROM diabetes_prediction.dbo.Dataset ;
```

100 %

Results | Messages

| | Average_BMI |
|---|---|
| 1 | 27.3512737929384 |

## 4. List patients in descending order of blood glucose levels.

```
SELECT *
FROM diabetes_prediction.dbo.Dataset
ORDER BY blood_glucose_level DESC;
```

100 %

Results | Messages

| | EmployeeName | Patient_id | gender | D#O#B | hypertension | heart_disease | smoking_history | bmi | HbA1c_level | blood_glucose_level | diabetes |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | TADAO YAMAGUCHI | PT2522 | Male | 1999-04-22 00:00:00.000 | 0 | 0 | former | 27.32 | 7 | 300 | 1 |
| 2 | BROCK WELLS | PT2635 | Female | 1999-04-26 00:00:00.000 | 0 | 0 | never | 31.26 | 6.6 | 300 | 1 |
| 3 | MARK TRIERWEILER | PT2639 | Female | 1999-04-26 00:00:00.000 | 1 | 0 | never | 27.32 | 6.5 | 300 | 1 |
| 4 | PATRICK BRYAN | PT2658 | Female | 1999-04-27 00:00:00.000 | 1 | 0 | former | 28.73 | 7.5 | 300 | 1 |
| 5 | KEVIN LYONS | PT2662 | Female | 1999-04-27 00:00:00.000 | 0 | 1 | never | 32.36 | 7.5 | 300 | 1 |
| 6 | DAMON O'BRIEN | PT2809 | Female | 1999-04-30 00:00:00.000 | 0 | 0 | never | 21.79 | 9 | 300 | 1 |
| 7 | JOHN MORANVILLE | PT2834 | Female | 1999-05-01 00:00:00.000 | 0 | 0 | No Info | 37.79 | 9 | 300 | 1 |
| 8 | REX HALE | PT195 | Female | 1997-05-27 00:00:00.000 | 0 | 0 | never | 27.32 | 7.5 | 300 | 1 |
| 9 | GERALD DARCY | PT243 | Female | 1997-07-13 00:00:00.000 | 0 | 0 | former | 21.97 | 7 | 300 | 1 |
| 10 | LORI BORGHI | PT300 | Female | 1997-08-19 00:00:00.000 | 0 | 0 | never | 26.71 | 6.5 | 300 | 1 |
| 11 | ROBERT DOSS | PT847 | Male | 1999-01-14 00:00:00.000 | 0 | 0 | not current | 32.19 | 5.8 | 300 | 1 |
| 12 | BOAZ MARILES | PT1037 | Male | 1999-02-10 00:00:00.000 | 0 | 0 | never | 27.32 | 6.5 | 300 | 1 |
| 13 | BRIDGET CULLINANE | PT1145 | Male | 1999-02-20 00:00:00.000 | 0 | 0 | current | 24.2 | 5.7 | 300 | 1 |
| 14 | THOMAS CULLINAN | PT1183 | Female | 1999-02-24 00:00:00.000 | 1 | 0 | never | 41.76 | 6.8 | 300 | 1 |
| 15 | CURTIS CHAN | PT1222 | Male | 1999-03-01 00:00:00.000 | 1 | 0 | never | 23.55 | 5.7 | 300 | 1 |
| 16 | DANIEL DECOSSIO | PT1319 | Male | 1999-03-08 00:00:00.000 | 1 | 0 | former | 22.06 | 9 | 300 | 1 |
| 17 | WILLIAM GARCIA | PT1321 | Male | 1999-03-08 00:00:00.000 | 1 | 0 | former | 57.17 | 5.8 | 300 | 1 |
| 18 | KIRK EDISON JR | PT1461 | Female | 1999-03-17 00:00:00.000 | 0 | 0 | never | 36.06 | 7.5 | 300 | 1 |
| 19 | RICHARD JONES | PT1466 | Male | 1999-03-17 00:00:00.000 | 0 | 0 | former | 29.4 | 8.8 | 300 | 1 |
| 20 | GREGORY KNIGHT | PT26430 | Female | 1999-12-25 00:00:00.000 | 0 | 0 | never | 33.72 | 7 | 300 | 1 |

## 5. Find patients who have hypertension and diabetes.

```
SELECT *FROM diabetes_prediction.dbo.Dataset
WHERE hypertension = 1 AND diabetes = 1;
```

00 %

Results | Messages

| | EmployeeName | Patient_id | gender | D#O#B | hypertension | heart_disease | smoking_history | bmi | HbA1c_level | blood_glucose_level | diabetes |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | LOUIS GALARCE | PT2597 | Female | 1999-04-25 00:00:00.000 | 1 | 0 | former | 28.89 | 6.1 | 145 | 1 |
| 2 | CHRISTOPHER YOCK | PT2619 | Female | 1999-04-25 00:00:00.000 | 1 | 0 | No Info | 31.99 | 6.5 | 260 | 1 |
| 3 | MARK TRIERWEILER | PT2639 | Female | 1999-04-26 00:00:00.000 | 1 | 0 | never | 27.32 | 6.5 | 300 | 1 |
| 4 | KHAE SAEPHAN | PT2641 | Female | 1999-04-26 00:00:00.000 | 1 | 0 | never | 28.73 | 6.2 | 220 | 1 |
| 5 | PATRICK BRYAN | PT2658 | Female | 1999-04-27 00:00:00.000 | 1 | 0 | former | 28.73 | 7.5 | 300 | 1 |
| 6 | JAMES CUSTER | PT2697 | Female | 1999-04-28 00:00:00.000 | 1 | 1 | not current | 31.29 | 8.8 | 280 | 1 |
| 7 | SUSAN NANGLE | PT2717 | Male | 1999-04-28 00:00:00.000 | 1 | 0 | never | 33.5 | 6.5 | 140 | 1 |
| 8 | STEVEN GARCIA | PT2723 | Female | 1999-04-28 00:00:00.000 | 1 | 0 | current | 45.13 | 8.8 | 280 | 1 |
| 9 | TRACY BARRAZA | PT2746 | Female | 1999-04-29 00:00:00.000 | 1 | 0 | current | 26.23 | 6.2 | 220 | 1 |
| 10 | PATRICK SULLIVAN | PT2785 | Female | 1999-04-30 00:00:00.000 | 1 | 0 | never | 37.3 | 5.8 | 130 | 1 |
| 11 | MICHAEL WOLF | PT2795 | Male | 1999-04-30 00:00:00.000 | 1 | 0 | No Info | 27.32 | 7.5 | 145 | 1 |
| 12 | PETER ALBERT | PT2918 | Male | 1999-05-03 00:00:00.000 | 1 | 0 | former | 25.53 | 6.6 | 145 | 1 |
| 13 | JONES WONG | PT139 | Male | 1989-08-09 00:00:00.000 | 1 | 0 | current | 27.32 | 5.7 | 260 | 1 |
| 14 | PATRIC STEELE | PT205 | Female | 1997-06-04 00:00:00.000 | 1 | 0 | never | 27.32 | 6.8 | 280 | 1 |
| 15 | ARTHUR STELLINI | PT343 | Male | 1997-09-07 00:00:00.000 | 1 | 1 | not current | 27.77 | 6.6 | 160 | 1 |
| 16 | CHAD LAW | PT355 | Male | 1997-09-12 00:00:00.000 | 1 | 0 | ever | 35.06 | 5.8 | 200 | 1 |
| 17 | CATHERINE JAMES | PT451 | Female | 1997-10-21 00:00:00.000 | 1 | 0 | never | 50.3 | 6.6 | 155 | 1 |
| 18 | JOHN HART | PT565 | Male | 1997-11-10 00:00:00.000 | 1 | 0 | current | 36.12 | 6.8 | 140 | 1 |
| 19 | JOHN BARKER | PT567 | Female | 1997-11-11 00:00:00.000 | 1 | 0 | former | 27.32 | 6.5 | 159 | 1 |
| 20 | ROBERT BONNET | PT632 | Female | 1997-12-01 00:00:00.000 | 1 | 0 | not current | 36.93 | 8.8 | 155 | 1 |

## 6. Determine the number of patients with heart disease.

```
SELECT COUNT(*) AS Heart_Disease_Patients
FROM diabetes_prediction.dbo.Dataset
WHERE heart_disease = 1;
```

100 %

Results | Messages

| | Heart_Disease_Patients |
|---|---|
| 1 | 1334 |

## 7. Group patients by smoking history and count how many smokers and non-smokers there are.

```
SELECT smoking_history, COUNT(*) AS Count
FROM diabetes_prediction.dbo.Dataset
GROUP BY smoking_history;
```

100 %

Results | Messages

| | smoking_history | Count |
|---|---|---|
| 1 | current | 3170 |
| 2 | ever | 1350 |
| 3 | former | 3295 |
| 4 | never | 12196 |
| 5 | No Info | 12273 |
| 6 | not current | 2180 |

## 8. Retrieve the Patient_ids of patients who have a BMI greater than the average BMI.

```
SELECT Patient_id
FROM diabetes_prediction.dbo.Dataset
WHERE bmi > (SELECT AVG(bmi) FROM diabetes_prediction.dbo.Dataset);
```

100 %

Results | Messages

| | Patient_id |
|---|---|
| 1 | PT2499 |
| 2 | PT2500 |
| 3 | PT2501 |
| 4 | PT2502 |
| 5 | PT2504 |
| 6 | PT2509 |
| 7 | PT2510 |
| 8 | PT2511 |
| 9 | PT2512 |
| 10 | PT2513 |
| 11 | PT2515 |
| 12 | PT2524 |
| 13 | PT2526 |
| 14 | PT2529 |
| 15 | PT2530 |
| 16 | PT2536 |
| 17 | PT2538 |
| 18 | PT2541 |
| 19 | PT2547 |

Query executed successfully. DESKTOP-TMQLL98\SQLEXPRESS ... DESKTOP-TMQLL98\DELL (60) master 00:00:00 11,631 rows

## 9. Find the patient with the highest HbA1c level and the patient with the lowest HbA1clevel.

```
SELECT TOP 1 WITH TIES Patient_id, HbA1c_level
FROM diabetes_prediction.dbo.Dataset
ORDER BY HbA1c_level DESC;

SELECT TOP 1 WITH TIES Patient_id, HbA1c_level
FROM diabetes_prediction.dbo.Dataset
ORDER BY HbA1c_level ASC;
```

100 % | Results | Messages

| | Patient_id | HbA1c_level |
|---|---|---|
| 1 | PT2642 | 9 |
| 2 | PT2809 | 9 |
| 3 | PT2834 | 9 |
| 4 | PT2839 | 9 |
| 5 | PT2925 | 9 |
| 6 | PT141 | 9 |
| 7 | PT156 | 9 |
| 8 | PT236 | 9 |

| | Patient_id | HbA1c_level |
|---|---|---|
| 1 | PT2519 | 3.5 |
| 2 | PT2537 | 3.5 |
| 3 | PT2546 | 3.5 |
| 4 | PT2554 | 3.5 |
| 5 | PT2555 | 3.5 |
| 6 | PT2557 | 3.5 |
| 7 | PT2579 | 3.5 |
| 8 | PT2583 | 3.5 |
| 9 | PT2589 | 3.5 |
| 10 | PT2603 | 3.5 |

Query executed successfully. DESKTOP-TMQLL98\SQLEXPRESS ... DESKTOP-TMQLL98\DELL (60) master 00:00:50 2,913 rows

## 10. Calculate the age of patients in years (assuming the current date as of now).

```
SELECT Patient_id, DATEDIFF(YEAR, D#O#B, GETDATE()) AS Age
FROM diabetes_prediction.dbo.Dataset;
```

100 % | Results | Messages

| | Patient_id | Age |
|---|---|---|
| 1 | PT2499 | 25 |
| 2 | PT2500 | 25 |
| 3 | PT2501 | 25 |
| 4 | PT2502 | 25 |
| 5 | PT2503 | 25 |
| 6 | PT2504 | 25 |
| 7 | PT2505 | 25 |
| 8 | PT2506 | 25 |
| 9 | PT2507 | 25 |
| 10 | PT2508 | 25 |
| 11 | PT2509 | 25 |
| 12 | PT2510 | 25 |
| 13 | PT2511 | 25 |
| 14 | PT2512 | 25 |
| 15 | PT2513 | 25 |
| 16 | PT2514 | 25 |
| 17 | PT2515 | 25 |
| 18 | PT2516 | 25 |
| 19 | PT2517 | 25 |

Query executed successfully. DESKTOP-TMQLL98\SQLEXPRESS ... DESKTOP-TMQLL98\DELL (60) master 00:00:00 34,464 rows

## 11. Rank patients by blood glucose level within each gender group.

```
SELECT *,
RANK() OVER (PARTITION BY gender ORDER BY blood_glucose_level DESC) AS Glucose_Level_Rank
FROM diabetes_prediction.dbo.Dataset;
```

100 % | Results | Messages

| | EmployeeName | Patient_id | gender | D#O#B | hypertension | heart_disease | smoking_history | bmi | HbA1c_level | blood_glucose_level | diabetes | Glucose_Level_Rank |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | BROCK WELLS | PT2635 | Female | 1999-04-26 00:00:00.000 | 0 | 0 | never | 31.26 | 6.6 | 300 | 1 | 1 |
| 2 | MARK TRIERWEILER | PT2639 | Female | 1999-04-26 00:00:00.000 | 1 | 0 | never | 27.32 | 6.5 | 300 | 1 | 1 |
| 3 | PATRICK BRYAN | PT2658 | Female | 1999-04-27 00:00:00.000 | 1 | 0 | former | 28.73 | 7.5 | 300 | 1 | 1 |
| 4 | KEVIN LYONS | PT2662 | Female | 1999-04-27 00:00:00.000 | 0 | 1 | never | 32.36 | 7.5 | 300 | 1 | 1 |
| 5 | DAMON O'BRIEN | PT2809 | Female | 1999-04-30 00:00:00.000 | 0 | 0 | never | 21.79 | 9 | 300 | 1 | 1 |
| 6 | JOHN MORANVILLE | PT2834 | Female | 1999-05-01 00:00:00.000 | 0 | 0 | No Info | 37.79 | 9 | 300 | 1 | 1 |
| 7 | REX HALE | PT195 | Female | 1997-05-27 00:00:00.000 | 0 | 0 | never | 27.32 | 7.5 | 300 | 1 | 1 |
| 8 | GERALD DARCY | PT243 | Female | 1997-07-13 00:00:00.000 | 0 | 0 | former | 21.97 | 7 | 300 | 1 | 1 |
| 9 | LORI BORGHI | PT300 | Female | 1997-08-19 00:00:00.000 | 0 | 0 | never | 26.71 | 6.5 | 300 | 1 | 1 |
| 10 | THOMAS CULLINAN | PT1183 | Female | 1999-02-24 00:00:00.000 | 1 | 0 | never | 41.76 | 6.8 | 300 | 1 | 1 |
| 11 | KIRK EDISON JR | PT1461 | Female | 1999-03-17 00:00:00.000 | 0 | 0 | never | 36.06 | 7.5 | 300 | 1 | 1 |
| 12 | GREGORY KNIGHT | PT26430 | Female | 1999-12-25 00:00:00.000 | 0 | 0 | never | 33.72 | 7 | 300 | 1 | 1 |
| 13 | JUDITH TENORIO | PT26773 | Female | 1999-12-27 00:00:00.000 | 0 | 0 | former | 26.23 | 8.8 | 300 | 1 | 1 |
| 14 | BILAL LOUBIE | PT26831 | Female | 1999-12-28 00:00:00.000 | 0 | 0 | never | 29.4 | 9 | 300 | 1 | 1 |
| 15 | WILLIAM MURPHY | PT26908 | Female | 1999-12-28 00:00:00.000 | 0 | 0 | not current | 40.71 | 5.8 | 300 | 1 | 1 |
| 16 | STEFANO MORONI | PT26989 | Female | 1999-12-29 00:00:00.000 | 1 | 0 | never | 27.53 | 6.1 | 300 | 1 | 1 |
| 17 | BRENDA TAYLOR | PT27314 | Female | 1999-12-30 00:00:00.000 | 1 | 0 | never | 43.42 | 8.2 | 300 | 1 | 1 |
| 18 | SACHI MANALISAY | PT27338 | Female | 1999-12-30 00:00:00.000 | 0 | 0 | never | 32.09 | 6.6 | 300 | 1 | 1 |
| 19 | GREGORY HANSEN | PT27460 | Female | 1999-12-31 00:00:00.000 | 0 | 0 | current | 45.59 | 6.1 | 300 | 1 | 1 |

Query executed successfully. DESKTOP-TMQLL98\SQLEXPRESS ... DESKTOP-TMQLL98\DELL (60) master 00:00:36 34,464 rows

## 12. Update the smoking history of patients who are older than 50 to "Ex-smoker."

```
UPDATE diabetes_prediction.dbo.Dataset
SET smoking_history = 'former'
WHERE DATEDIFF(YEAR, D#O#B, GETDATE()) > 50;
```

100 %

Messages

```
(0 rows affected)

Completion time: 2024-04-26T13:58:33.1350014+05:30
```

## 13. Insert a new patient into the database with sample data.

```
INSERT INTO diabetes_prediction.dbo.Dataset (EmployeeName, Patient_id, gender, D#O#B, hypertension, heart_disease, bmi, HbA1c_level, blood_glucose_level, diabetes)
VALUES ('Advaika Jhon', 'SP001', 'Male', '1973-01-05', 0, 0, 25.5, 6.2, 110, 0);
```

100 %

Messages

```
(1 row affected)

Completion time: 2024-04-26T14:03:54.9183718+05:30
```

## 14. Delete all patients with heart disease from the database

```
DELETE FROM diabetes_prediction.dbo.Dataset
WHERE heart_disease = '1';
```

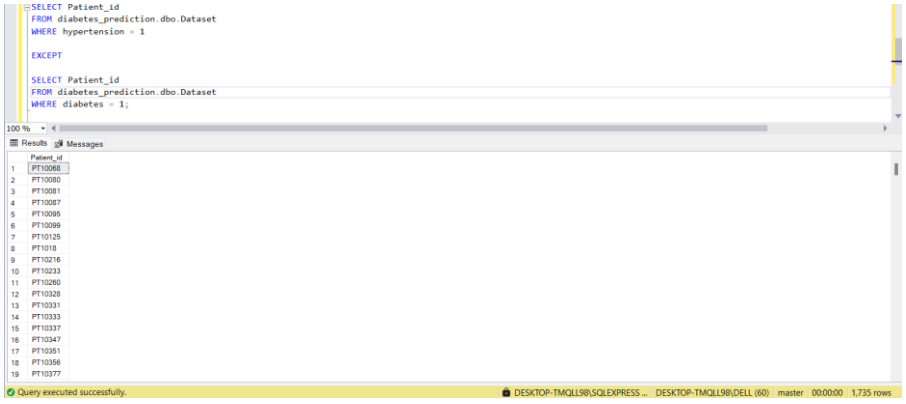
100 %

Messages

```
(1334 rows affected)

Completion time: 2024-04-26T14:13:17.8167530+05:30
```

## 15. Find patients who have hypertension but not diabetes using the EXCEPT Operator.



## 16. Define a unique constraint on the "patient_id" column to ensure its values are unique.

```
ALTER TABLE diabetes_prediction.dbo.Dataset
ADD CONSTRAINT UC_Patient_id UNIQUE (Patient_id);
```

100 %

Messages

```
Commands completed successfully.

Completion time: 2024-04-26T14:18:33.8186386+05:30
```

## 17. Create a view that displays the Patient_ids, ages, and BMI of patients.

```
CREATE VIEW Patient_Info AS
SELECT Patient_id, DATEDIFF(YEAR, D#O#B, GETDATE()) AS Age, bmi AS BMI
FROM diabetes_prediction.dbo.Dataset;
```

100 %

Messages

```
Commands completed successfully.

Completion time: 2024-04-26T14:21:01.3613396+05:30
```

## 18. Suggest improvements in the database schema to reduce data redundancy and improve data integrity?

To reduce data redundancy and improve data integrity in the database schema, consider the following improvements:

1) Normalization

2) Master Data Management (MDM)

3) Data Cleansing

4) Enforce Unique Constraints

5) Normalization Levels

6) Denormalization

## 19. Explain how you can optimize the performance of SQL queries on this Dataset?

To optimize the performance of SQL queries on the dataset, several strategies can be

1) Implemented
2) Indexing
3) Query Optimization
4) Stored Procedures

5) Normalization
6) Database Tuning
7) Use of Indexes
8) Caching
9) Optimize Joins