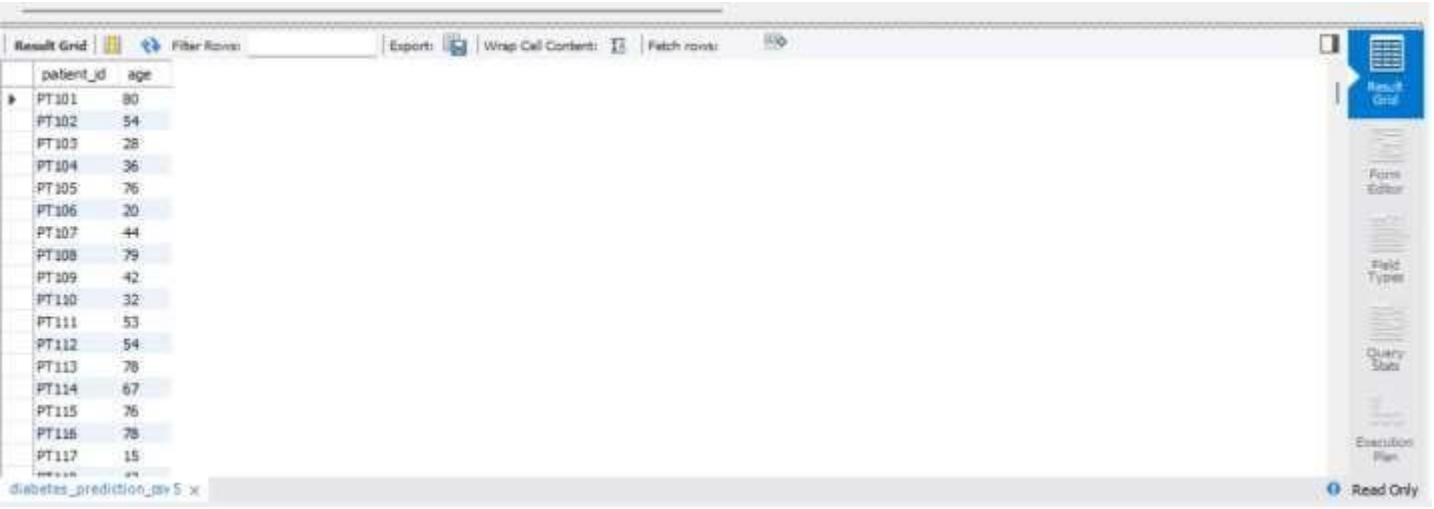
SQL Project
Diabetes Prediction

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

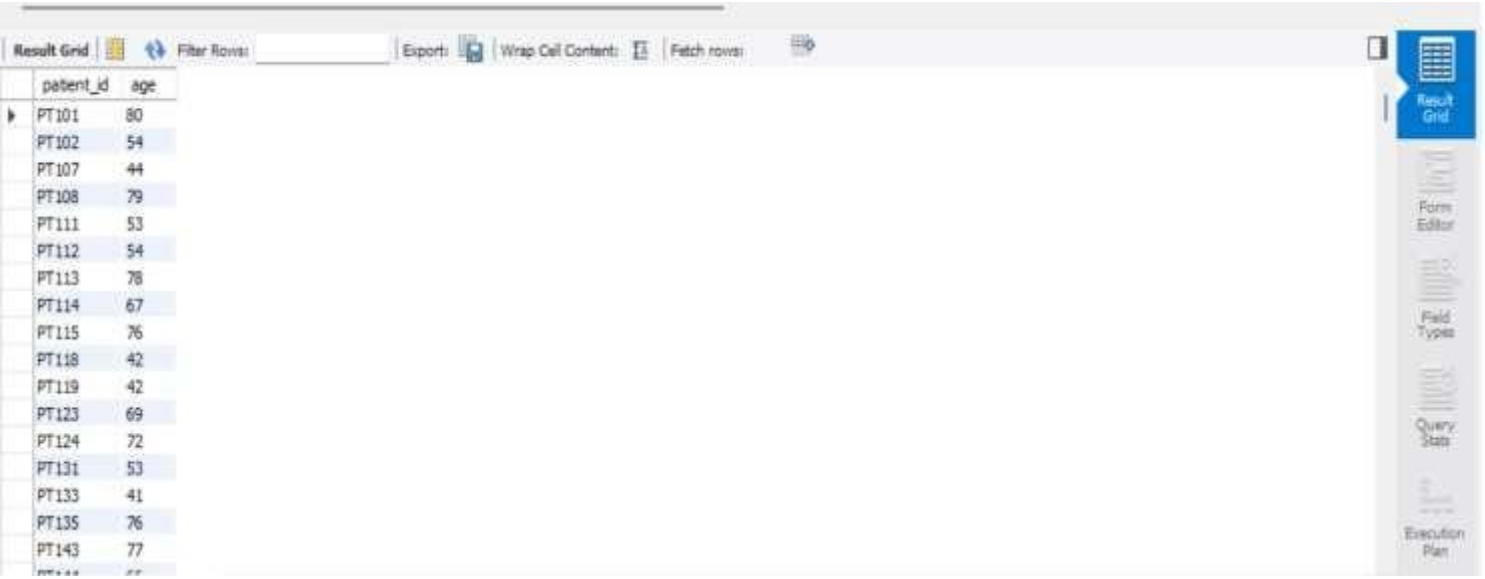
```
select patient_id,age from diabetes_prediction_csv;
```



patient_id	age
PT101	80
PT102	54
PT103	28
PT104	36
PT105	76
PT106	20
PT107	44
PT108	79
PT109	42
PT110	32
PT111	53
PT112	54
PT113	78
PT114	67
PT115	76
PT116	78
PT117	15

#2. Select all female patients who are older than 40.

```
select patient_id,age from diabetes_prediction_csv
where gender='Female' and age > 40;
```



patient_id	age
PT101	80
PT102	54
PT107	44
PT108	79
PT111	53
PT112	54
PT113	78
PT114	67
PT115	76
PT118	42
PT119	42
PT123	69
PT124	72
PT131	53
PT133	41
PT135	76
PT143	77

#3. Calculate the average BMI of patients;

```
select avg(bmi) as AVG_BMI from diabetes_prediction_csv;
```

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

AVG_BMI
27.32076709999422

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

```
select employeename,patient_id,blood_glucose_level from diabetes_prediction_csv
order by blood_glucose_level desc;
```

employeename	patient_id	blood_glucose_level
Gilbert J Fragoso	PT99638	300
Amado A Lumas Jr	PT99663	300
Shanice M Guidry	PT99672	300
Angelica J Young	PT99764	300
Flor D Roman	PT99809	300
Clyde L Woods	PT99927	300
Josephine C Cabrera	PT99968	300
Marquis D Walker	PT100039	300
Brenda G Velasquez	PT89459	300
Haroon Ahmad	PT88934	300
Marc S Slaven	PT93637	300
Silvia Woo	PT91896	300
Terese M Bonilla	PT91863	300
Nigel L Hicks	PT91250	300
Cliff E Bell	PT89546	300

#5 Find patients who have hypertension and diabetes.

```
select * from diabetes_prediction_csv where hypertension =1 and diabetes=1;
```

Employeename	Patient_id	gender	age	hypertension	heart_disease	smoking_history	bmi	HbA1c_level	blood_glucose_level	diabetes
BARRY WONG	PT3429	Male	48	1	0	current	27.32	8.2	155	1
SHELDA JOHNSON	PT3450	Female	58	1	0	never	38.52	6.6	130	1
CYNTHIA GAABUC...	PT3454	Female	28	1	0	never	42.44	5.8	130	1
ALISON YEE	PT3576	Female	48	1	1	former	50.86	6	159	1
SOCORRO SISON	PT3581	Female	73	1	0	never	22.8	9	126	1
JASON JACKSON	PT3638	Female	74	1	0	never	29.74	5.8	220	1
ANTHONY SAN JOSE	PT3721	Male	26	1	0	current	20.04	7.5	155	1
MA GRACIA LOPEZ	PT3773	Male	47	1	0	current	46.61	8.2	145	1
ERNA FACULTAD	PT3810	Male	65	1	0	former	33.9	5.8	160	1
RYAN MAK	PT3913	Female	59	1	0	never	38.1	6.6	140	1
JOCELYN MADAMBA	PT3914	Female	65	1	0	never	37.2	6.8	160	1
FIDEL GONZALEZ	PT3919	Male	61	1	0	former	37.88	9	200	1
FERN EBELING	PT3955	Female	61	1	0	never	38.95	6	240	1
ROBERT CAPPA	PT4103	Female	73	1	0	never	27.8	6.8	280	1
THERESA CONWAY	PT4172	Female	56	1	0	not current	28.22	7.5	140	1

#6 Determine the number of patients with heart disease.

```
select count(*) as Patients_With_HeartDisease from diabetes_prediction_csv
where heart_disease=1;
```

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

Patients_With_HeartDisease
3942

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

```
select smoking_history,count(*) from diabetes_prediction_csv
group by smoking_history;
```

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

smoking_history	count(*)
never	35095
No Info	35816
current	9286
former	9352
ever	4004
not current	6447

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

```
select * from diabetes_prediction_csv
where bmi > (select avg(bmi) from diabetes_prediction_csv);
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content: | Fetch rows:

EmployeeName	Patient_id	gender	age	hypertension	heart_disease	smoking_history	bmi	HbA1c_level	blood_glucose_level	diabetes
MICHAEL MORRIS	PT109	Male	42	0	0	never	33.64	4.8	145	0
PATRICIA JACKSON	PT112	Female	54	0	0	former	54.7	6	100	0
EDWARD HARRINGTON	PT113	Female	78	0	0	former	36.05	5	130	0
AMY HART	PT117	Male	15	0	0	never	30.36	6.1	200	0
VENUS AZAR	PT121	Male	40	0	0	current	36.38	6	90	0
VICTOR WYRSCH	PT124	Female	72	0	1	former	27.94	6.5	130	0
GREGORY SUHR	PT126	Male	30	0	0	never	33.76	6.1	126	0
RAYMOND GUZMAN	PT128	Male	40	0	0	former	27.85	5.8	80	0
HARLAN KELLY-JR	PT131	Female	53	0	0	No Info	31.75	4	200	0
BRENDAN WARD	PT140	Female	34	0	0	never	56.43	6.2	200	0
THOMAS SIRAGUSA	PT143	Female	77	1	1	never	32.02	5	159	0
MICHAEL THOMPSON	PT144	Female	66	0	0	No Info	29.3	4.8	159	0
JAMES DUDLEY	PT149	Female	38	0	0	never	28.27	6.2	155	0
MICHAEL ROLOVICH	PT153	Female	74	0	0	No Info	28.12	5	100	0
KEVIN CASHMAN	PT156	Male	50	0	0	former	37.16	9	159	1

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

```
select * from diabetes_prediction_csv order by HbA1c_level Desc limit 1;
select * from diabetes_prediction_csv order by HbA1c_level limit 1;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content: | Fetch rows:

EmployeeName	Patient_id	gender	age	hypertension	heart_disease	smoking_history	bmi	HbA1c_level	blood_glucose_level	diabetes
Miao Ling Huang	PT99613	Male	60	0	0	former	27.32	9	280	1

Result Grid | Filter Rows: | Export: | Wrap Cell Content: | Fetch rows:

EmployeeName	Patient_id	gender	age	hypertension	heart_disease	smoking_history	bmi	HbA1c_level	blood_glucose_level	diabetes
Franz Brustmeyer	PT98827	Female	51	0	0	current	49.11	3.5	100	0

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

```
SELECT Patient_id, age,YEAR(curDATE())-age AS AgeInYearsFROM diabetes_prediction_csv;
```

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

```
select Patient_id,gender,blood_glucose_level, rank() over (partition by  gender order by
blood_glucose_level desc) AS GlucoseRank from diabetes_prediction_csv;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Co

Patient_id	gender	blood_glucose_level	GlucoseRank
PT96902	Female	300	1
PT97622	Female	300	1
PT96346	Female	300	1
PT97955	Female	300	1
PT97671	Female	300	1
PT96328	Female	300	1
PT98911	Female	300	1
PT98538	Female	300	1
PT97570	Female	300	1
PT98454	Female	300	1
PT99638	Female	300	1
PT97141	Female	300	1
PT96371	Female	300	1
PT96815	Female	300	1
PT96814	Female	300	1
PT97830	Female	300	1

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

```
UPDATE diabetes_prediction_csv
SET smoking_history = 'Ex-smoker’ WHERE age > 50;
```

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

```
INSERT INTO diabetes_prediction_csv (EmployeeName, Patient_id, gender, age,
hypertension,heart_disease, smoking_history,
bmi,HbA1c_level,blood_glucose_level,diabetes)VALUES ('Ravi','PT1141', 'Male', 23, 0,0, 'Never',
20, 6.6, 100,0);
```

#14 Delete all patients with heart disease from the database

```
delete from diabetes_prediction_csv where heart_disease = 1;
```

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

```
select Patient_id from diabetes_prediction_csv where hypertension =1  EXCEPT
select Patient_id from diabetes_prediction_csv where diabetes =1;
```

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

```
ALTER TABLE diabetes_prediction_csvADD UNIQUE (patient_id(255));
```

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

```
CREATE VIEW PatientInfo ASSELECT Patient_id, age, BMIFROM diabetes_prediction_csv;

SELECT * FROM PatientInfo;
```