Report (Diabetes)

Diabetes: is a disease that occurs when your blood glucose, also called blood sugar, is too high. Blood glucose is your main sourc e of energy and comes from the food you eat. Insulin, a hormone m ade by the pancreas, helps glucose from food get into your cells to be used for energy.

1 - Data:

	ID	No_Pation	Gender	AGE	Urea	Cr	HbA1c	Chol	TG	HDL	LDL	VLDL	BMI	CLASS
0	502	17975	F	50	4.7	46	4.9	4.2	0.9	2.4	1.4	0.5	24.0	N
1	735	34221	M	26	4.5	62	4.9	3.7	1.4	1.1	2.1	0.6	23.0	N
2	420	47975	F	50	4.7	46	4.9	4.2	0.9	2.4	1.4	0.5	24.0	N
3	680	87656	F	50	4.7	46	4.9	4.2	0.9	2.4	1.4	0.5	24.0	N
4	504	34223	M	33	7.1	46	4.9	4.9	1.0	0.8	2.0	0.4	21.0	N
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
95	171	45434	M	40	22.0	159	5.4	3.1	1.6	1.1	1.3	0.7	24.0	N
96	180	454316	F	54	6.3	106	5.6	4.3	2.0	1.3	2.2	0.9	20.0	N
97	181	454316	F	50	3.3	70	4.9	3.7	0.9	1.2	2.7	0.6	24.0	N
98	182	454316	M	60	7.5	70	0.9	3.4	5.3	1.1	3.6	0.8	24.0	N
99	186	454316	M	77	5.0	106	5.4	0.0	2.8	0.8	1.8	0.7	19.0	N

These are the data I have analyzed..

- Gender:

  M: male.

  F: Female.

- Class:

  N: NO (Not have diabetes)

  P: prediabetes

  Y: Yes (have diabetes)

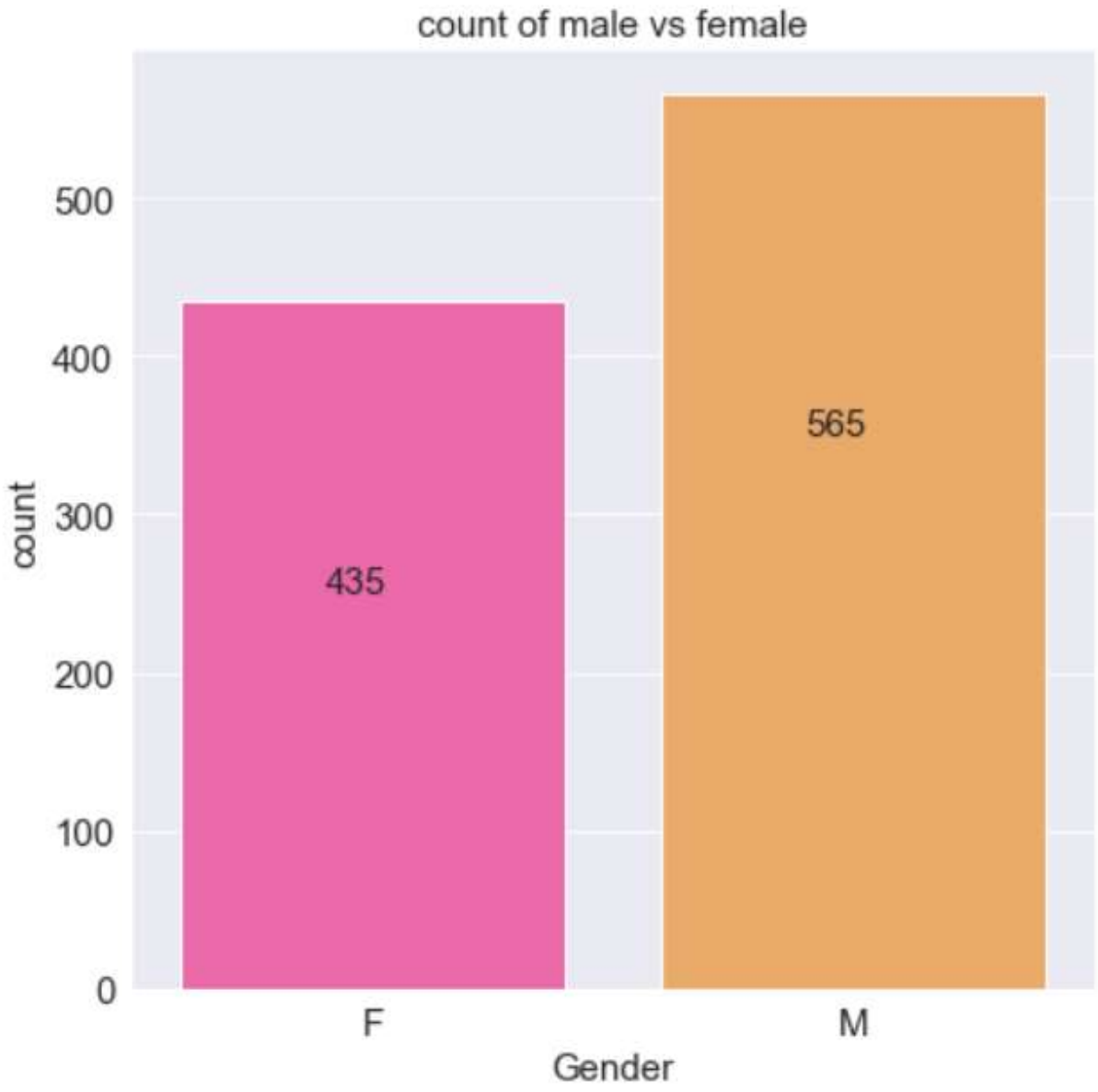
business plan:

Q1: what is the percentage of male and female in this data? and the percentage of disease?

Q2: What is the range of age the disease is common?

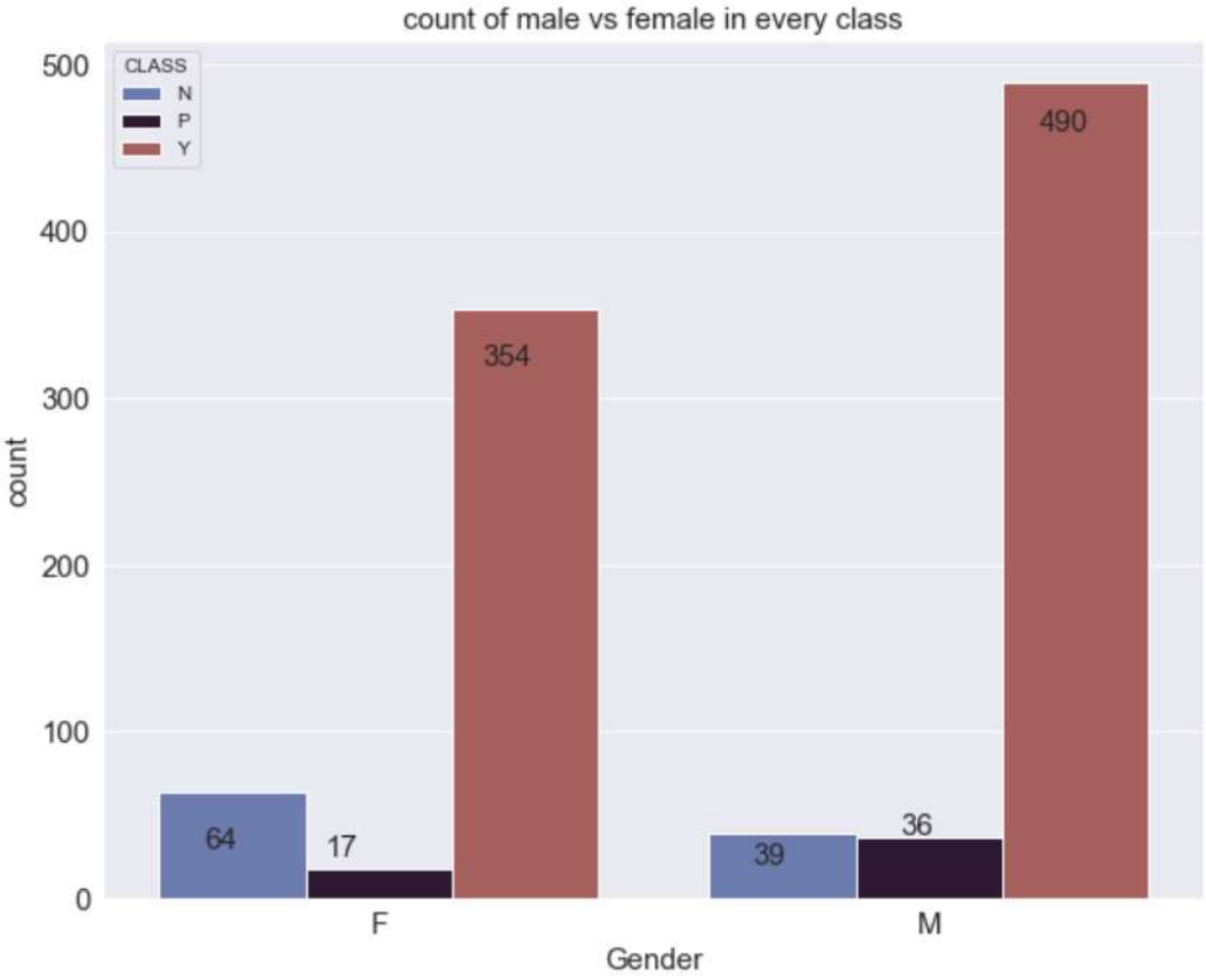
Q3 : what is the features are related with the class ? and how we can classification the disease?

Q1: what is the percentage of male and female in this data? and the percentage of disease?



The count of males =565, and females= 435, the total is 1000.

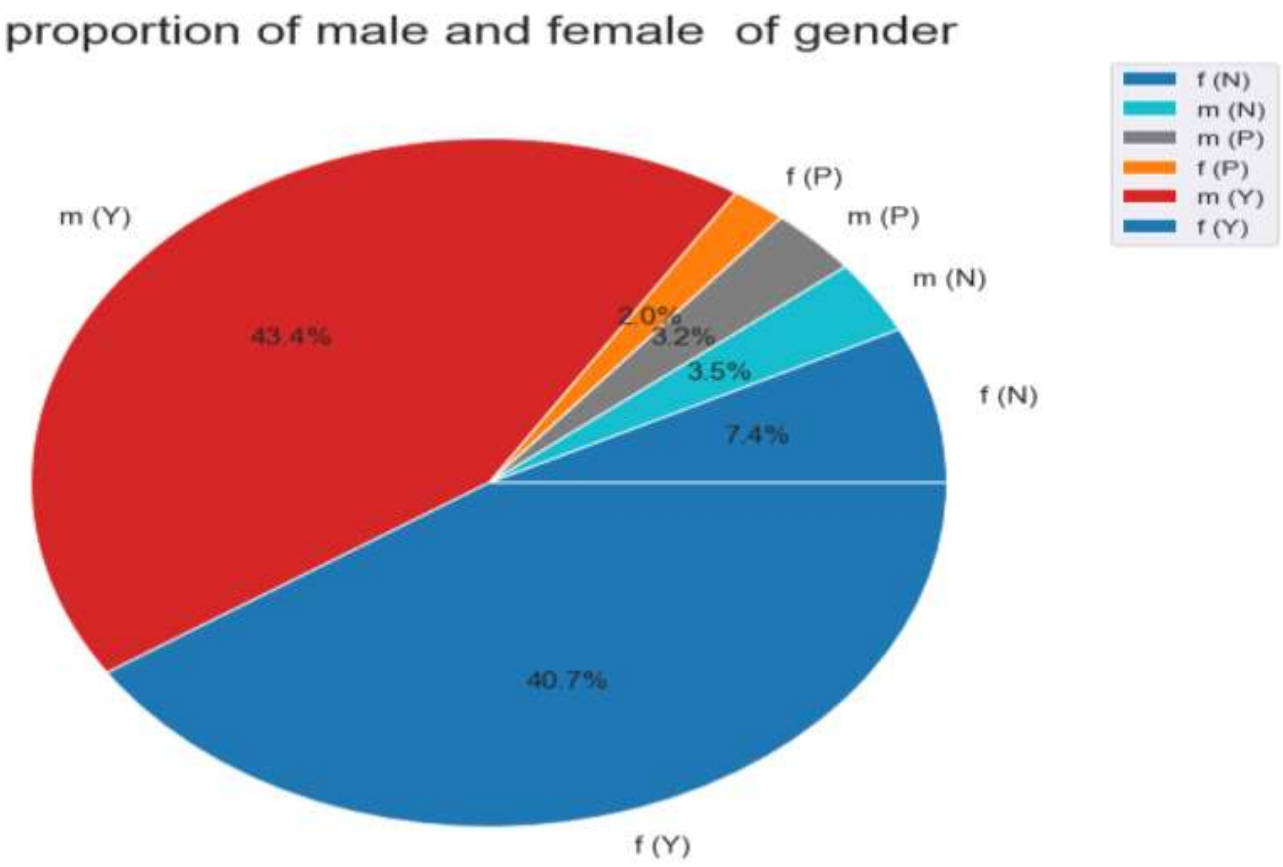
It is mean the count of males is larger than females.

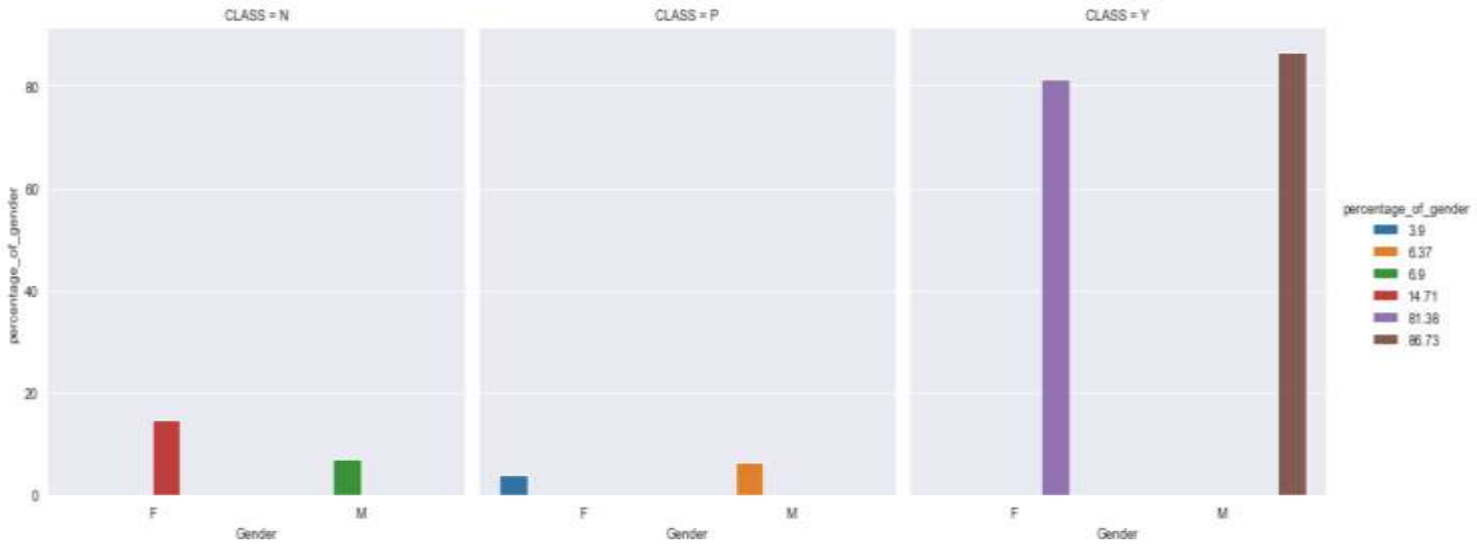


The males who have diabetes are **490**, prediabetes are **36** and **39** don't have diabetes from **565**

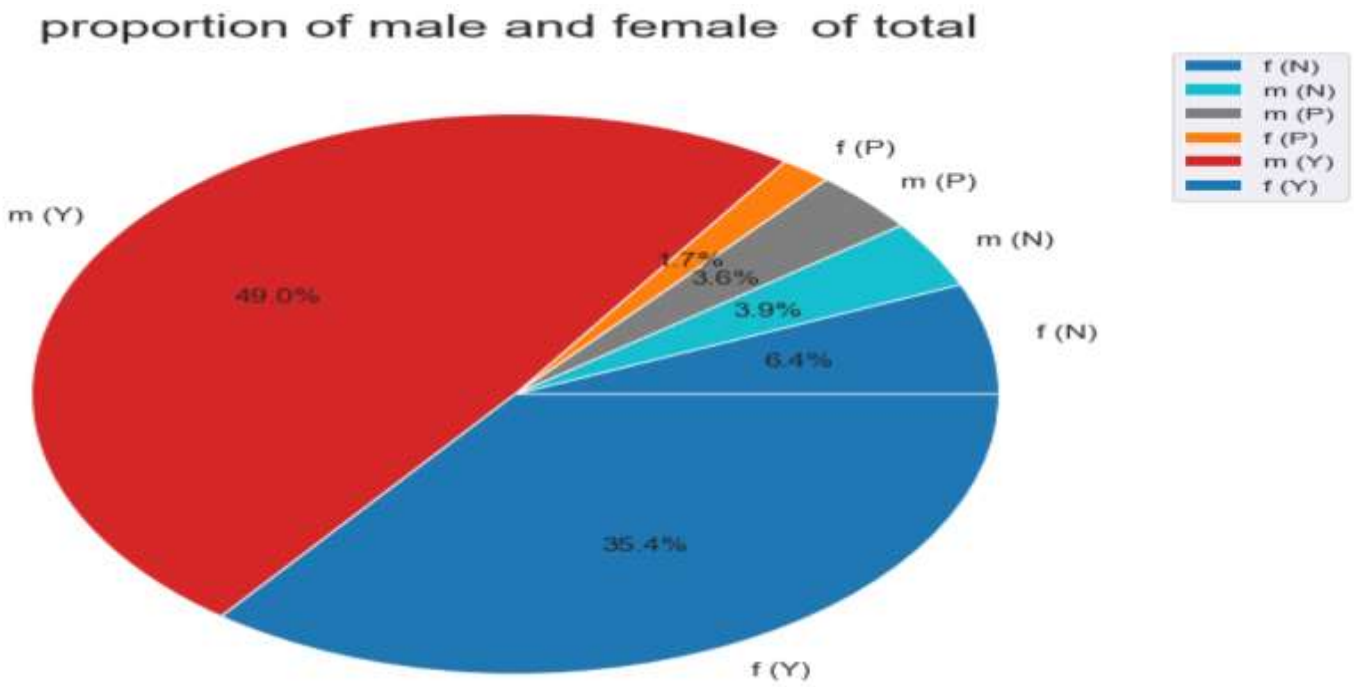
The females who have diabetes are **354,** prediabetes is **16,** and **64** do not have diabetes from **435**.

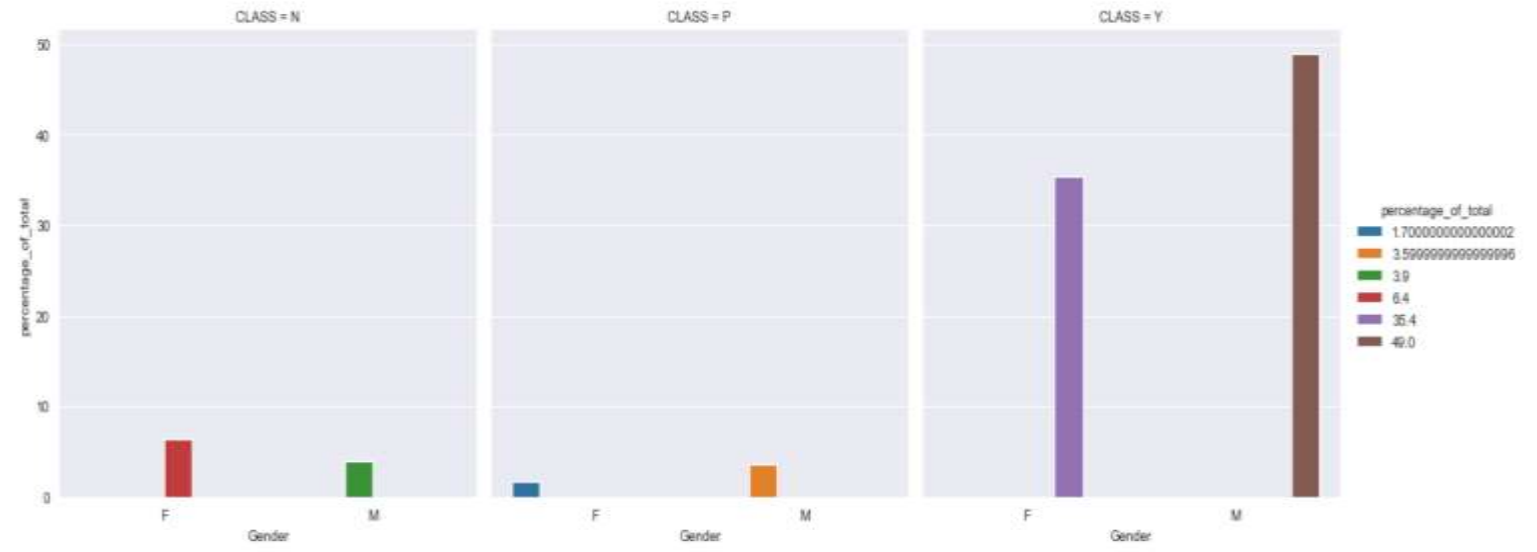
proportion of male and female of gender





proportion of male and female of total

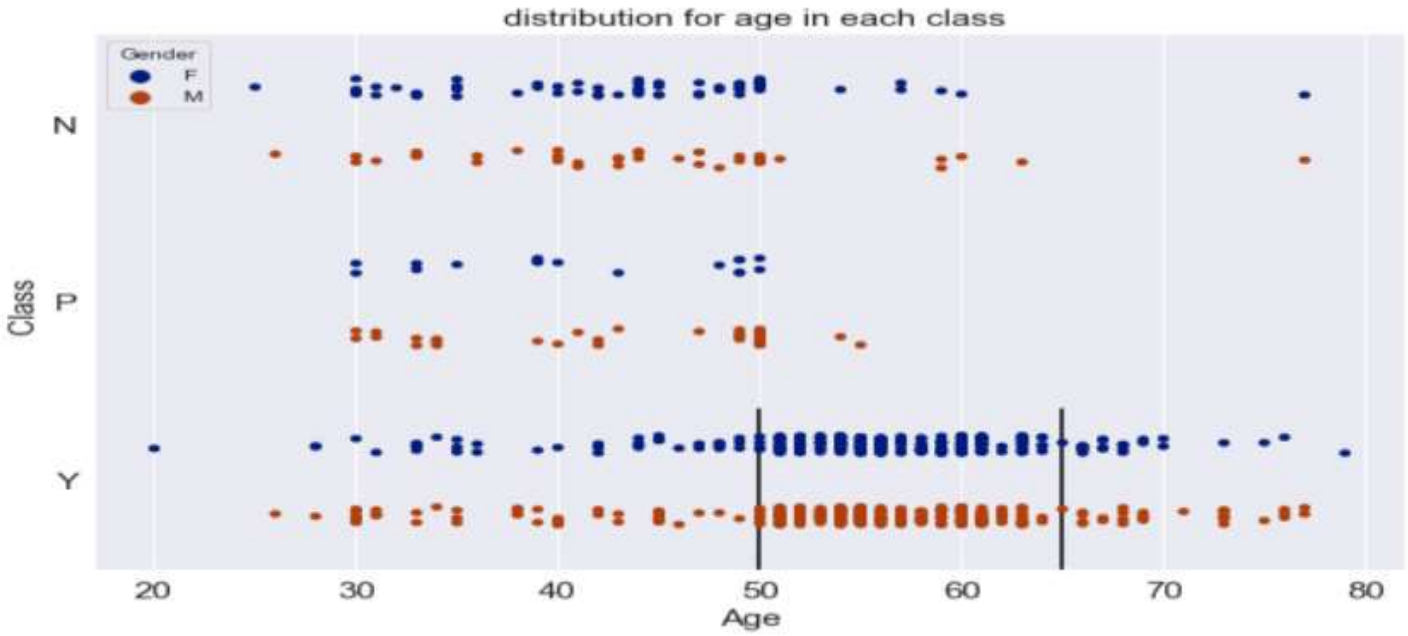


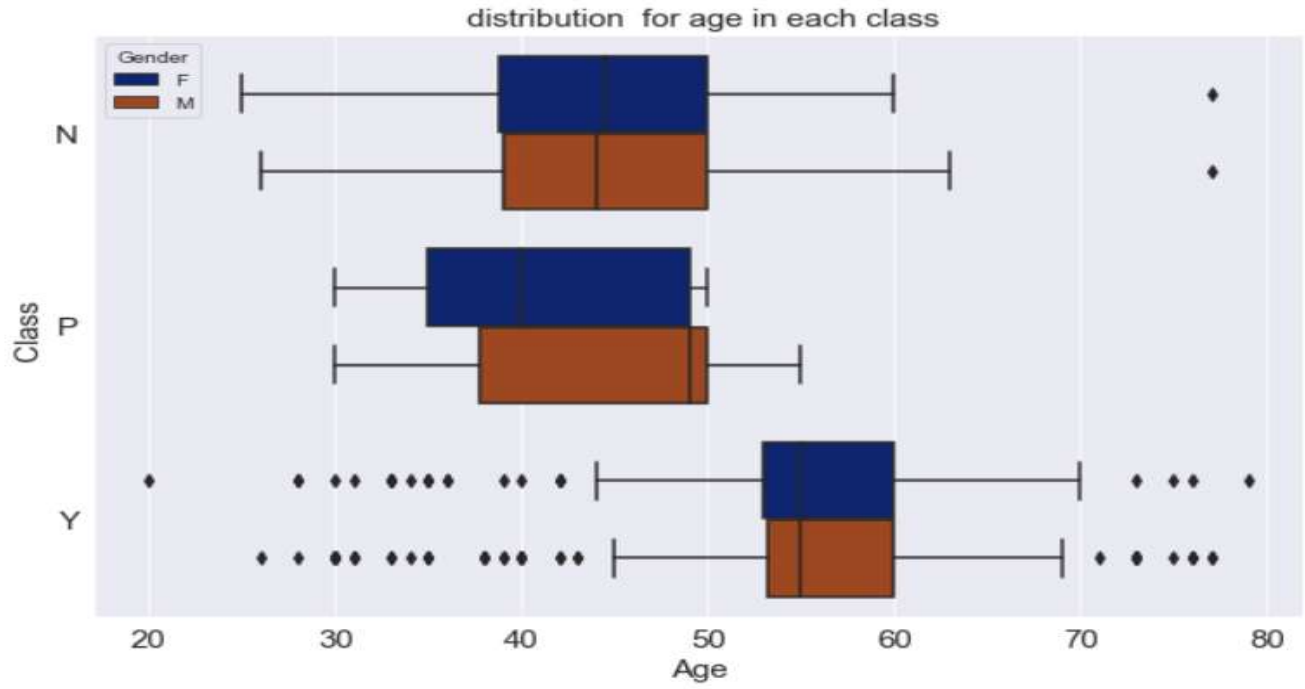


The proportion of males who have diabetes is **86.37 %** from males, and **49%** from total , and prediabetes is **6.37%** from males & **3.6%** from total.

while the proportion of females who have diabetes is **81.38%** from females and **35.4%** from total and prediabetes is **3.9 %** from females & **1.7%** from total.

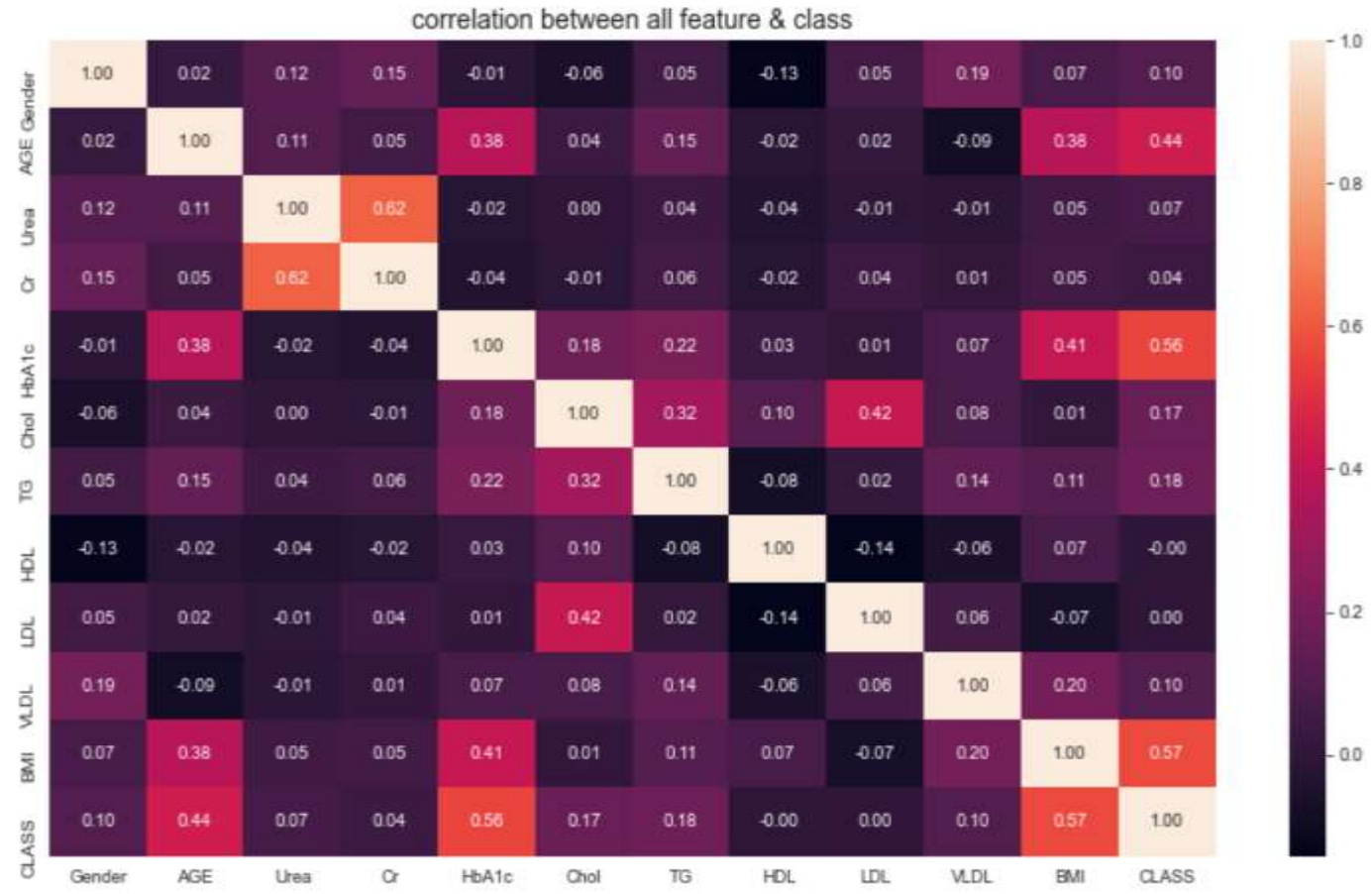
Q2: What is the range of age the disease is common?





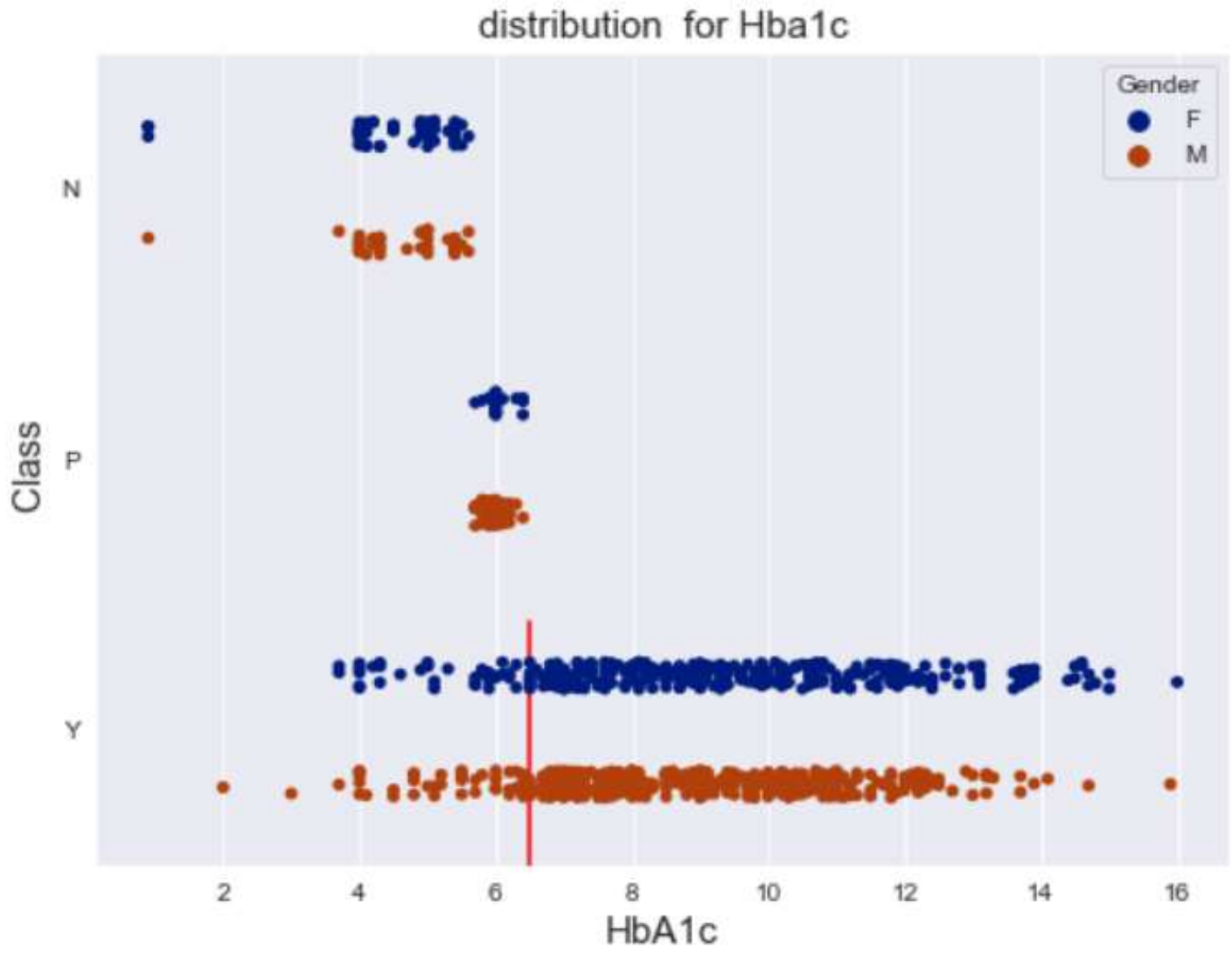
We see an increased risk of diabetes between age (**50-65**)

Q3 : what is the features are related with the class ? and how we can classification the disease ?



The correlation between ['Urea', 'Cr', 'HDL', 'LDL'] and class is very low.

distribution for Hba1c



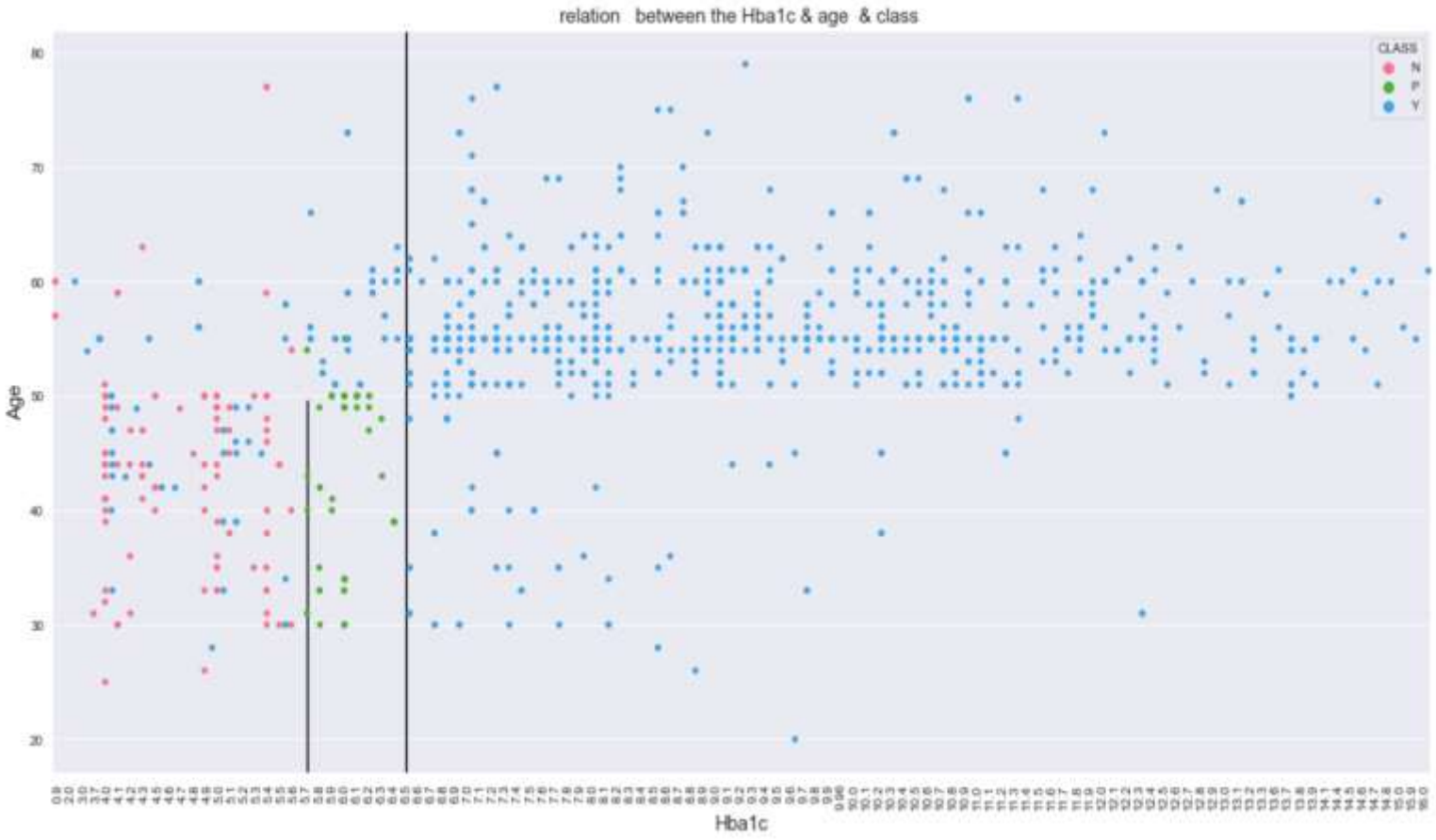
Normal : Below 5.7%

Prediabetes:5.7% to 6.4%

Diabetes :6.5% or above

but we found patients have diabetes and the HbA1c less than 6.5 !! why?

relation between the Hba1c & age & class



HbA1C >**6.5** .... class be Y

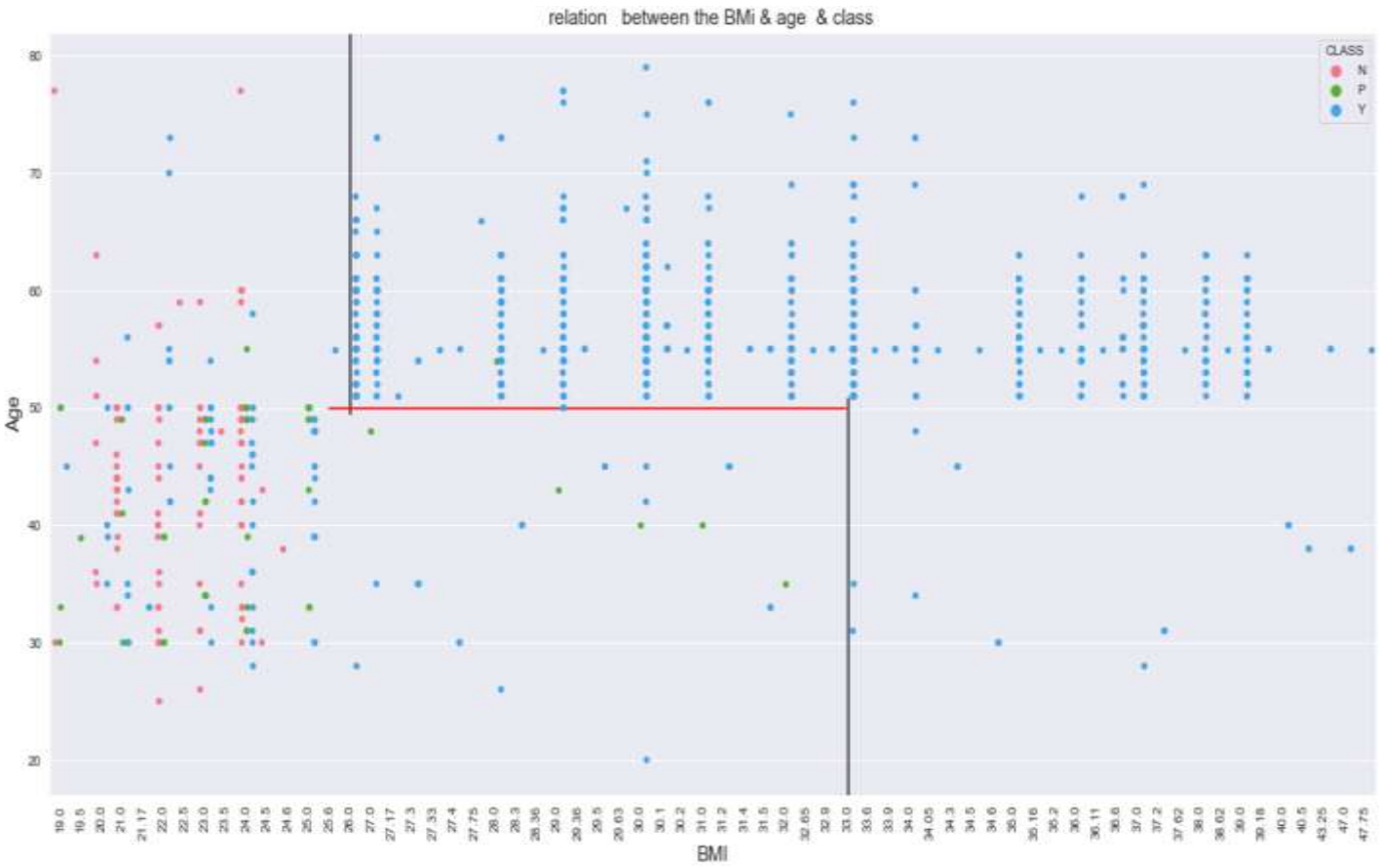
Age between **51-53** & HbA1C **6** or above .....class be Y

age >**53** & HbA1C **2** or above the most ..... class be Y

age between 30-50 & HbA1C 5.7-6.4..... classs be P

age between **20-80** & HbA1C between **3-5.6** ..... class can be Y or N what the feature related with this is case ?

relation between the BMi & age & class



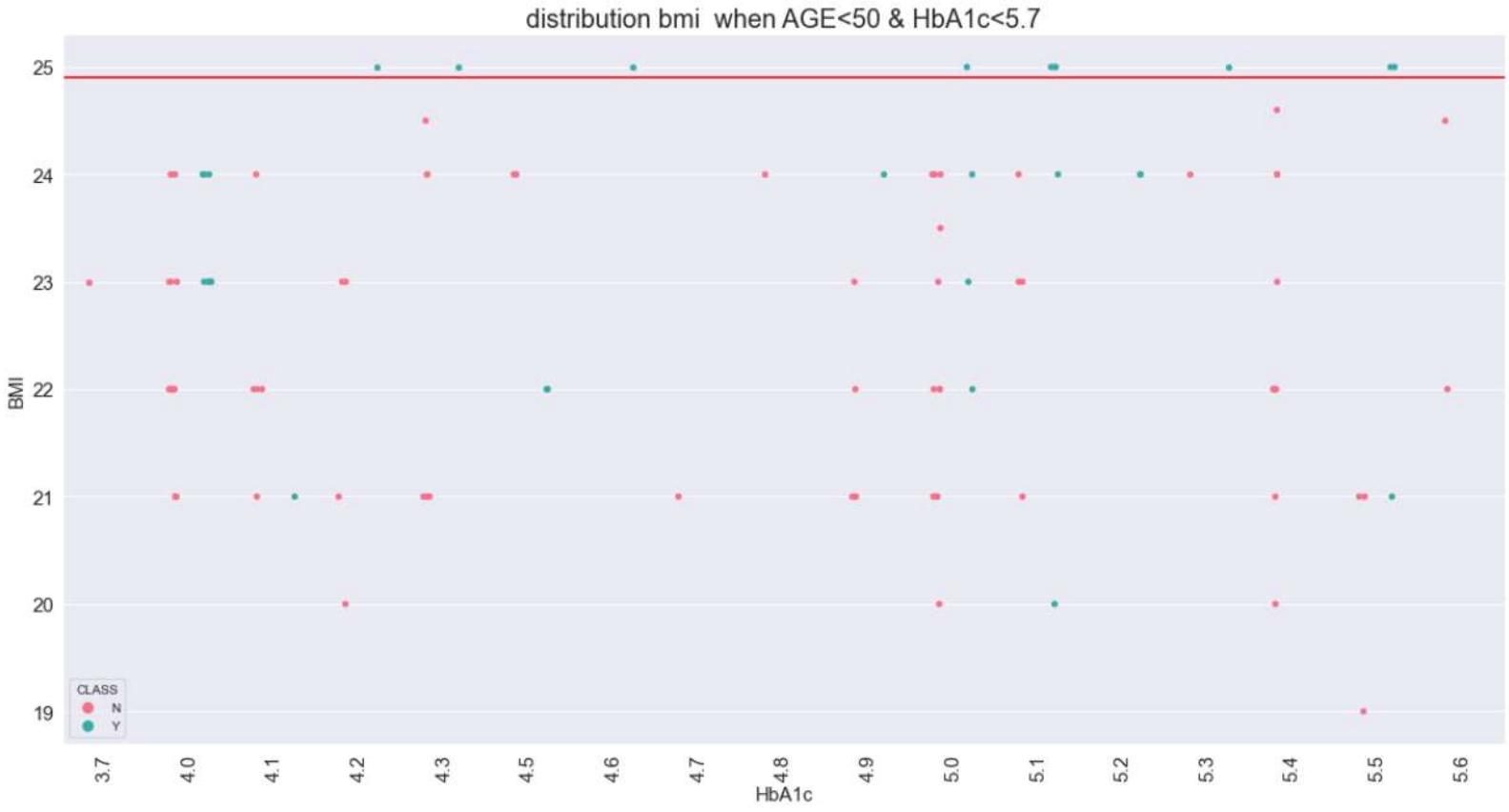
age >50 & bmi >= **26** ..... class be Y

20 < age <30 & bmi >=**24** ..... class be Y

age between 20-50 & bmi >=**33** ..... class be Y

age between 20-50 or above & bmi between **19-24** ..... class can be Y or N

distribution bmi when AGE<50 & HbA1c<5.7

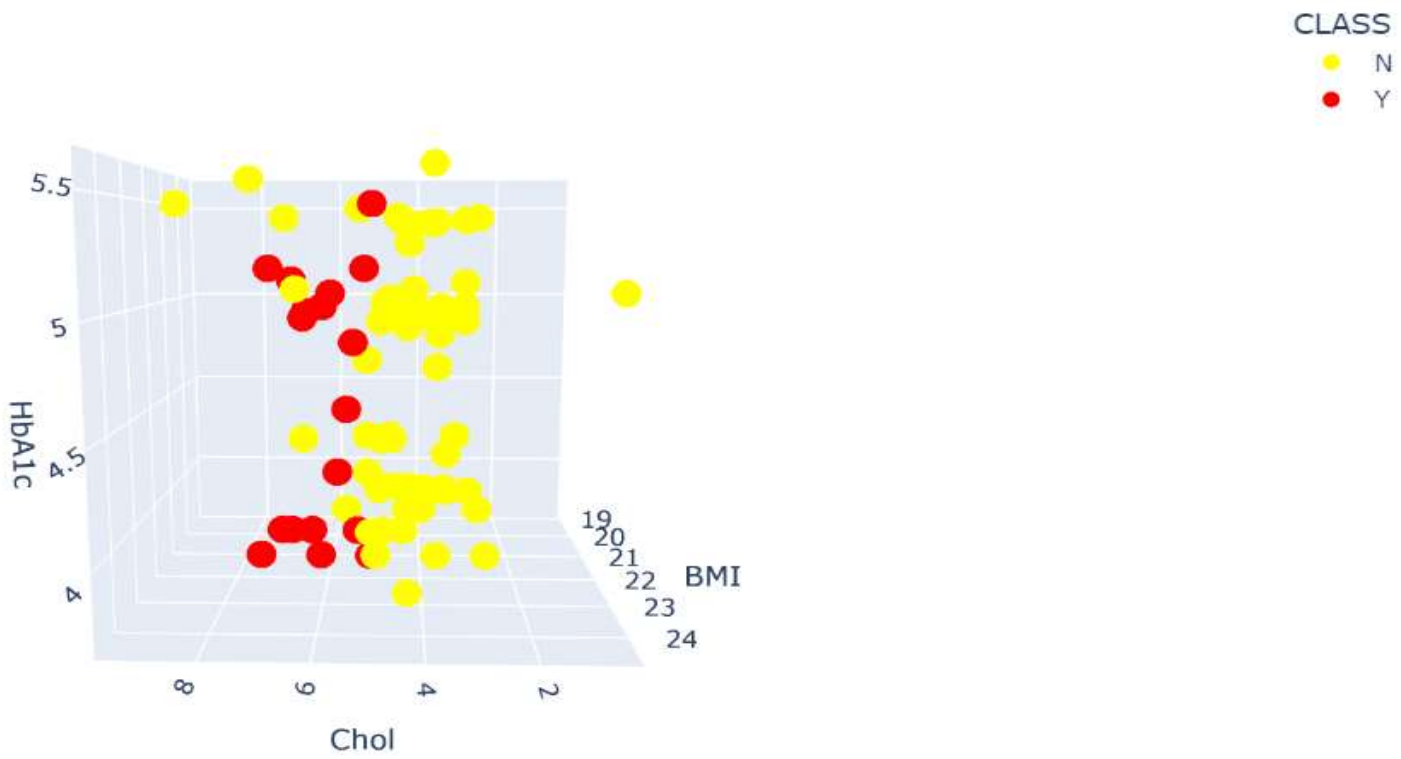


when HbA1c <5.7 & bmi =25 ..... class be Y

discover the feature is most related when AGE<50 & HbA1c<5.7 &BMI!=25

I found it (chol)

Plotting 3d for chol when AGE<50 & HbA1c<5.7



when age <50 & HbA1c<5.7 &bmi<25 & chol >=5 .....class be Y

Summary :

When *:*

Age between **51**-**53** & HbA1C **6** or above .....class be Y

age >**53** & HbA1C **2** or above the most ..... class be Y

age between 30-50 & HbA1C 5.7-6.4..... classs be P

age >50 & bmi >= **26** ..... class be Y

20 < age <30 & bmi >=**24** ..... class be Y

age between 20-50 & bmi >=**33** ..... class be Y

AGE<50 & HbA1c<5.7 & bmi =25 ..... class be Y

age <50 & HbA1c<5.7 &bmi<25 & chol >=5 .....class be Y

Report of model Decision tree

```
Training Model DT
--------------
Training Accuracy: 1.0
Testing Accuracy: 0.995
Testing Confusion Matrix:
[[ 27   0   0]
 [  0   9   0]
 [  1   0 163]]
Testing Recall: 0.9979674796747968
Testing Precesion: 0.9880952380952381
Testing F1 score :0.9929200259475489
------------------------------
```