```
import pandas as pd
import numpy as np
import seaborn as sns
from sklearn.model_selection import train_test_split
```

```
df = pd.read_csv("T3_kidney_disease.csv", na_values=["?"])
```

```
df
```

	@relation Chronic_Kidney_Disease	Unnamed: 1	Unnamed: 2	Unnamed: 3	Unnamed: 4	Unnamed: 5	Unnamed: 6	Unnamed: 7	Unnamed: 8	Unnamed: 9	...
0	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	...
1	@attribute 'age' numeric	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	...
2	@attribute 'bp' numeric	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	...
3	@attribute 'sg' {1.005	1.01	1.015	1.02	1.025}	NaN	NaN	NaN	NaN	NaN	...
4	@attribute 'al' {0	1	2.000	3.00	4	5}	NaN	NaN	NaN	NaN	...
...	...	...	...	...	...	...	...	...	...	...	...
423	55	80	1.020	0.00	0	normal	normal	notpresent	notpresent	140.0	...
424	42	70	1.025	0.00	0	normal	normal	notpresent	notpresent	75.0	...
425	12	80	1.020	0.00	0	normal	normal	notpresent	notpresent	100.0	...
426	17	60	1.025	0.00	0	normal	normal	notpresent	notpresent	114.0	...
427	58	80	1.025	0.00	0	normal	normal	notpresent	notpresent	131.0	...

428 rows × 26 columns

```
df.drop([0,1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,21,22,23,24,25,26,27], axis=0, inplace=True)
```

```
df
```

	@relation Chronic_Kidney_Disease	Unnamed: 1	Unnamed: 2	Unnamed: 3	Unnamed: 4	Unnamed: 5	Unnamed: 6	Unnamed: 7	Unnamed: 8	Unnamed: 9	...
28	48	80	1.020	1.0	0	NaN	normal	notpresent	notpresent	121.0	...
29	7	50	1.020	4.0	0	NaN	normal	notpresent	notpresent	NaN	...
30	62	80	1.010	2.0	3	normal	normal	notpresent	notpresent	423.0	...
31	48	70	1.005	4.0	0	normal	abnormal	present	notpresent	117.0	...
32	51	80	1.010	2.0	0	normal	normal	notpresent	notpresent	106.0	...
...	...	...	...	...	...	...	...	...	...	...	...
423	55	80	1.020	0.0	0	normal	normal	notpresent	notpresent	140.0	...
424	42	70	1.025	0.0	0	normal	normal	notpresent	notpresent	75.0	...
425	12	80	1.020	0.0	0	normal	normal	notpresent	notpresent	100.0	...
426	17	60	1.025	0.0	0	normal	normal	notpresent	notpresent	114.0	...
427	58	80	1.025	0.0	0	normal	normal	notpresent	notpresent	131.0	...

400 rows × 26 columns

```
df.columns = ['age', 'bp', 'sg', 'al', 'su', 'rbc', 'pc', 'pcc', 'ba', 'bgr', 'bu', 'sc', 'sod', 'pot', 'hemo', 'pcv', 'wc
```

```
df
```

	age	bp	sg	al	su	rbc	pc	pcc	ba	bgr	...	wc	rc	htn	dm	cad	appet	pe	ane	clas
28	48	80	1.020	1.0	0	NaN	normal	notpresent	notpresent	121.0	...	7800	5.2	yes	yes	no	good	no	no	cl
29	7	50	1.020	4.0	0	NaN	normal	notpresent	notpresent	NaN	...	6000	NaN	no	no	no	good	no	no	cl
30	62	80	1.010	2.0	3	normal	normal	notpresent	notpresent	423.0	...	7500	NaN	no	yes	no	poor	no	yes	cl
31	48	70	1.005	4.0	0	normal	abnormal	present	notpresent	117.0	...	6700	3.9	yes	no	no	poor	yes	yes	cl
32	51	80	1.010	2.0	0	normal	normal	notpresent	notpresent	106.0	...	7300	4.6	no	no	no	good	no	no	cl
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	
423	55	80	1.020	0.0	0	normal	normal	notpresent	notpresent	140.0	...	6700	4.9	no	no	no	good	no	no	notcl
424	42	70	1.025	0.0	0	normal	normal	notpresent	notpresent	75.0	...	7800	6.2	no	no	no	good	no	no	notcl
425	12	80	1.020	0.0	0	normal	normal	notpresent	notpresent	100.0	...	6600	5.4	no	no	no	good	no	no	notcl
426	17	60	1.025	0.0	0	normal	normal	notpresent	notpresent	114.0	...	7200	5.9	no	no	no	good	no	no	notcl
427	58	80	1.025	0.0	0	normal	normal	notpresent	notpresent	131.0	...	6800	6.1	no	no	no	good	no	no	notcl

400 rows × 26 columns

```
df = df.reset_index(drop=True)
df.index = df.index + 1
```

```
df
```

	age	bp	sg	al	su	rbc	pc	pcc	ba	bgr	...	wc	rc	htn	dm	cad	appet	pe	ane	clas
1	48	80	1.020	1.0	0	NaN	normal	notpresent	notpresent	121.0	...	7800	5.2	yes	yes	no	good	no	no	cl
2	7	50	1.020	4.0	0	NaN	normal	notpresent	notpresent	NaN	...	6000	NaN	no	no	no	good	no	no	cl
3	62	80	1.010	2.0	3	normal	normal	notpresent	notpresent	423.0	...	7500	NaN	no	yes	no	poor	no	yes	cl
4	48	70	1.005	4.0	0	normal	abnormal	present	notpresent	117.0	...	6700	3.9	yes	no	no	poor	yes	yes	cl
5	51	80	1.010	2.0	0	normal	normal	notpresent	notpresent	106.0	...	7300	4.6	no	no	no	good	no	no	cl
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	
396	55	80	1.020	0.0	0	normal	normal	notpresent	notpresent	140.0	...	6700	4.9	no	no	no	good	no	no	notcl
397	42	70	1.025	0.0	0	normal	normal	notpresent	notpresent	75.0	...	7800	6.2	no	no	no	good	no	no	notcl
398	12	80	1.020	0.0	0	normal	normal	notpresent	notpresent	100.0	...	6600	5.4	no	no	no	good	no	no	notcl
399	17	60	1.025	0.0	0	normal	normal	notpresent	notpresent	114.0	...	7200	5.9	no	no	no	good	no	no	notcl
400	58	80	1.025	0.0	0	normal	normal	notpresent	notpresent	131.0	...	6800	6.1	no	no	no	good	no	no	notcl

400 rows × 26 columns

```
df.shape
```

```
(400, 26)
```

```
pd.isna(df)
```

	age	bp	sg	al	su	rbc	pc	pcc	ba	bgr	...	wc	rc	htn	dm	cad	appet	pe	ane
1	False	False	False	False	False	True	False	False	False	False	...	False	False	False	False	False	False	False	False
2	False	False	False	False	False	True	False	False	False	True	...	False	True	False	False	False	False	False	False
3	False	False	False	False	False	False	False	False	False	False	...	False	True	False	False	False	False	False	False
4	False	False	False	False	False	False	False	False	False	False	...	False	False	False	False	False	False	False	False
5	False	False	False	False	False	False	False	False	False	False	...	False	False	False	False	False	False	False	False
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
396	False	False	False	False	False	False	False	False	False	False	...	False	False	False	False	False	False	False	False
397	False	False	False	False	False	False	False	False	False	False	...	False	False	False	False	False	False	False	False
398	False	False	False	False	False	False	False	False	False	False	...	False	False	False	False	False	False	False	False
399	False	False	False	False	False	False	False	False	False	False	...	False	False	False	False	False	False	False	False
400	False	False	False	False	False	False	False	False	False	False	...	False	False	False	False	False	False	False	False

400 rows × 26 columns

```
df.isnull().sum()
```

```
age        9
bp        12
sg        47
al        46
su        49
rbc      152
pc        65
pcc        4
ba         4
bgr       44
bu        19
sc        17
sod       87
pot       88
hemo      52
pcv       70
wc       105
rc       130
htn        2
dm         3
cad        2
appet      1
pe         1
ane        1
class      0
ckd      399
dtype: int64
```

```
num_cols = ['age','bp','sg','al','su','bgr','bu','sc','sod','pot','hemo','pcv','wc','rc']
notnum_cols = [c for c in df.columns if c not in num_cols]
```

```
for col in num_cols:
    df[col] = pd.to_numeric(df[col], errors="coerce")
df.dtypes
```

```
age      float64
bp       float64
sg       float64
al       float64
su       float64
rbc       object
pc        object
pcc       object
ba        object
bgr      float64
bu       float64
sc       float64
sod      float64
pot      float64
hemo     float64
pcv      float64
wc       float64
rc       float64
htn       object
dm        object
cad       object
appet     object
pe        object
ane       object
class     object
ckd       object
dtype: object
```

```
df.describe()
```

	age	bp	sg	al	su	bgr	bu	sc	sod	pot	
count	391.000000	388.000000	353.000000	354.000000	351.000000	356.000000	381.000000	383.000000	313.000000	312.000000	3
mean	51.483376	76.469072	1.017408	1.016949	0.450142	148.036517	57.425722	3.072454	137.528754	4.627244	
std	17.169714	13.683637	0.005717	1.352679	1.099191	79.281714	50.503006	5.741126	10.408752	3.193904	
min	2.000000	50.000000	1.005000	0.000000	0.000000	22.000000	1.500000	0.400000	4.500000	2.500000	
25%	42.000000	70.000000	1.010000	0.000000	0.000000	99.000000	27.000000	0.900000	135.000000	3.800000	
50%	55.000000	80.000000	1.020000	0.000000	0.000000	121.000000	42.000000	1.300000	138.000000	4.400000	
75%	64.500000	80.000000	1.020000	2.000000	0.000000	163.000000	66.000000	2.800000	142.000000	4.900000	
max	90.000000	180.000000	1.025000	5.000000	5.000000	490.000000	391.000000	76.000000	163.000000	47.000000	

```
(df==0).sum()
```

```
age        0
bp         0
sg         0
al       199
su       290
rbc        0
pc         0
pcc        0
ba         0
bgr        0
bu         0
sc         0
sod        0
pot        0
hemo       0
pcv        0
wc         0
rc         0
htn        0
dm         0
cad        0
appet      0
pe         0
ane        0
class      0
ckd        0
dtype: int64
```

```
df.std(numeric_only=True)
```

```
age       17.169714
bp        13.683637
sg         0.005717
al         1.352679
su         1.099191
bgr       79.281714
bu        50.503006
sc         5.741126
sod       10.408752
pot        3.193904
hemo       2.912587
pcv        8.990105
wc      2944.474190
rc         1.025323
dtype: float64
```

```
df.mean(numeric_only=True)
```

```
age       51.483376
bp        76.469072
sg         1.017408
al         1.016949
su         0.450142
bgr      148.036517
bu        57.425722
sc         3.072454
sod      137.528754
pot        4.627244
hemo      12.526437
pcv       38.884498
wc      8406.122449
rc         4.707435
dtype: float64
```

```
df.mode()
```

	age	bp	sg	al	su	rbc	pc	pcc	ba	bgr	...	wc	rc	htn	dm	cad	appet	pe	ane
0	60.0	80.0	1.02	0.0	0.0	normal	normal	notpresent	notpresent	99.0	...	9800.0	5.2	no	no	no	good	no	no
1	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	...	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN

2 rows × 26 columns

```
df.median(numeric_only=True)
```

```
age       55.00
bp        80.00
sg         1.02
```

```
al         0.00
su         0.00
bgr      121.00
bu        42.00
sc         1.30
sod      138.00
pot        4.40
hemo      12.65
pcv       40.00
wc      8000.00
rc         4.80
dtype: float64
```

```
df = df.fillna(df[notnum_cols].mode().iloc[0])
df
```

	age	bp	sg	al	su	rbc	pc	pcc	ba	bgr	...	wc	rc	htn	dm	cad	appet	pe	ane
1	48.0	80.0	1.020	1.0	0.0	normal	normal	notpresent	notpresent	121.0	...	7800.0	5.2	yes	yes	no	good	no	no
2	7.0	50.0	1.020	4.0	0.0	normal	normal	notpresent	notpresent	NaN	...	6000.0	NaN	no	no	no	good	no	no
3	62.0	80.0	1.010	2.0	3.0	normal	normal	notpresent	notpresent	423.0	...	7500.0	NaN	no	yes	no	poor	no	yes
4	48.0	70.0	1.005	4.0	0.0	normal	abnormal	present	notpresent	117.0	...	6700.0	3.9	yes	no	no	poor	yes	yes
5	51.0	80.0	1.010	2.0	0.0	normal	normal	notpresent	notpresent	106.0	...	7300.0	4.6	no	no	no	good	no	no
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
396	55.0	80.0	1.020	0.0	0.0	normal	normal	notpresent	notpresent	140.0	...	6700.0	4.9	no	no	no	good	no	no
397	42.0	70.0	1.025	0.0	0.0	normal	normal	notpresent	notpresent	75.0	...	7800.0	6.2	no	no	no	good	no	no
398	12.0	80.0	1.020	0.0	0.0	normal	normal	notpresent	notpresent	100.0	...	6600.0	5.4	no	no	no	good	no	no
399	17.0	60.0	1.025	0.0	0.0	normal	normal	notpresent	notpresent	114.0	...	7200.0	5.9	no	no	no	good	no	no
400	58.0	80.0	1.025	0.0	0.0	normal	normal	notpresent	notpresent	131.0	...	6800.0	6.1	no	no	no	good	no	no

400 rows × 26 columns

```
df = df.fillna(df[num_cols].mean())
df
```

	age	bp	sg	al	su	rbc	pc	pcc	ba	bgr	...	wc	rc	htn	dm	cad	appet
1	48.0	80.0	1.020	1.0	0.0	normal	normal	notpresent	notpresent	121.000000	...	7800.0	5.200000	yes	yes	no	good
2	7.0	50.0	1.020	4.0	0.0	normal	normal	notpresent	notpresent	148.036517	...	6000.0	4.707435	no	no	no	good
3	62.0	80.0	1.010	2.0	3.0	normal	normal	notpresent	notpresent	423.000000	...	7500.0	4.707435	no	yes	no	poor
4	48.0	70.0	1.005	4.0	0.0	normal	abnormal	present	notpresent	117.000000	...	6700.0	3.900000	yes	no	no	poor
5	51.0	80.0	1.010	2.0	0.0	normal	normal	notpresent	notpresent	106.000000	...	7300.0	4.600000	no	no	no	good
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
396	55.0	80.0	1.020	0.0	0.0	normal	normal	notpresent	notpresent	140.000000	...	6700.0	4.900000	no	no	no	good
397	42.0	70.0	1.025	0.0	0.0	normal	normal	notpresent	notpresent	75.000000	...	7800.0	6.200000	no	no	no	good
398	12.0	80.0	1.020	0.0	0.0	normal	normal	notpresent	notpresent	100.000000	...	6600.0	5.400000	no	no	no	good
399	17.0	60.0	1.025	0.0	0.0	normal	normal	notpresent	notpresent	114.000000	...	7200.0	5.900000	no	no	no	good
400	58.0	80.0	1.025	0.0	0.0	normal	normal	notpresent	notpresent	131.000000	...	6800.0	6.100000	no	no	no	good

400 rows × 26 columns

```
df.isnull().sum()
```

```
age      0
bp       0
sg       0
al       0
su       0
rbc      0
pc       0
pcc      0
ba       0
bgr      0
bu       0
sc       0
sod      0
pot      0
hemo     0
pcv      0
wc       0
```

```
rc       0
htn      0
dm       0
cad      0
appet    0
pe       0
ane      0
class    0
ckd      0
dtype: int64
```

```
from sklearn.model_selection import train_test_split
X= df.drop(columns=['class'])
Y= df['class']
X_train, X_test, Y_train, Y_test = train_test_split(X, Y,test_size=0.2, random_state=42)
```

```
X_train
```

	age	bp	sg	al	su	rbc	pc	pcc	ba	bgr	...	pcv	wc	rc	h
4	48.0	70.0	1.005000	4.000000	0.000000	normal	abnormal	present	notpresent	117.0	...	32.0	6700.000000	3.900000	y
19	60.0	100.0	1.025000	0.000000	3.000000	normal	normal	notpresent	notpresent	263.0	...	37.0	11400.000000	4.300000	y
203	78.0	60.0	1.017408	1.016949	0.450142	normal	normal	notpresent	notpresent	114.0	...	24.0	8406.122449	4.707435	
251	40.0	80.0	1.025000	0.000000	0.000000	normal	normal	notpresent	notpresent	140.0	...	48.0	10400.000000	4.500000	
275	19.0	80.0	1.020000	0.000000	0.000000	normal	normal	notpresent	notpresent	107.0	...	44.0	8406.122449	4.707435	
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	
72	46.0	60.0	1.010000	1.000000	0.000000	normal	normal	notpresent	notpresent	163.0	...	28.0	14600.000000	3.200000	y
107	50.0	90.0	1.017408	1.016949	0.450142	normal	normal	notpresent	notpresent	89.0	...	17.0	6500.000000	4.707435	y
271	23.0	80.0	1.025000	0.000000	0.000000	normal	normal	notpresent	notpresent	111.0	...	41.0	7200.000000	5.000000	
349	38.0	80.0	1.020000	0.000000	0.000000	normal	normal	notpresent	notpresent	99.0	...	44.0	7300.000000	6.400000	
103	17.0	60.0	1.010000	0.000000	0.000000	normal	normal	notpresent	notpresent	92.0	...	52.0	7000.000000	4.707435	

320 rows × 25 columns

```
Y_train
```

```
4         ckd
19        ckd
203       ckd
251    notckd
275    notckd
        ...
72        ckd
107       ckd
271    notckd
349    notckd
103       ckd
Name: class, Length: 320, dtype: object
```

```
X_test
```

	age	bp	sg	al	su	rbc	pc	pcc	ba	bgr	...	pcv	wc
210	19.0	70.0	1.020000	0.000000	0.000000	normal	normal	notpresent	notpresent	148.036517	...	38.884498	6900.000000
281	47.0	80.0	1.017408	1.016949	0.450142	normal	normal	notpresent	notpresent	93.000000	...	52.000000	8100.000000
34	60.0	100.0	1.020000	2.000000	0.000000	abnormal	abnormal	notpresent	notpresent	140.000000	...	29.000000	8406.122449
211	59.0	100.0	1.015000	4.000000	2.000000	normal	normal	notpresent	notpresent	255.000000	...	20.000000	9800.000000
94	73.0	100.0	1.010000	3.000000	2.000000	abnormal	abnormal	present	notpresent	295.000000	...	30.000000	7000.000000
...	...	...	...	...	...	...	...	...	...	...	...	...	...
247	48.0	110.0	1.015000	3.000000	0.000000	abnormal	normal	present	notpresent	106.000000	...	26.000000	5000.000000
228	57.0	80.0	1.015000	0.000000	0.000000	normal	normal	notpresent	notpresent	120.000000	...	36.000000	7200.000000
370	75.0	70.0	1.020000	0.000000	0.000000	normal	normal	notpresent	notpresent	107.000000	...	46.000000	10300.000000
177	21.0	90.0	1.010000	4.000000	0.000000	normal	abnormal	present	present	107.000000	...	23.000000	12400.000000
290	42.0	70.0	1.020000	0.000000	0.000000	normal	normal	notpresent	notpresent	93.000000	...	43.000000	7100.000000

80 rows × 25 columns

```
Y_test
```

```
210       ckd
281    notckd
34        ckd
211       ckd
94        ckd
        ...
247       ckd
228       ckd
370        no
177       ckd
290    notckd
Name: class, Length: 80, dtype: object
```

```
subset1= df[df['age']>58]
subset1
```

	age	bp	sg	al	su	rbc	pc	pcc	ba	bgr	...	wc	rc	htn	dm	cad	appe
3	62.0	80.0	1.010	2.0	3.0	normal	normal	notpresent	notpresent	423.0	...	7500.000000	4.707435	no	yes	no	poo
6	60.0	90.0	1.015	3.0	0.0	normal	normal	notpresent	notpresent	74.0	...	7800.000000	4.400000	yes	yes	no	goo
7	68.0	70.0	1.010	0.0	0.0	normal	normal	notpresent	notpresent	100.0	...	8406.122449	4.707435	no	no	no	goo
12	63.0	70.0	1.010	3.0	0.0	abnormal	abnormal	present	notpresent	380.0	...	4500.000000	3.800000	yes	yes	no	poo
13	68.0	70.0	1.015	3.0	1.0	normal	normal	present	notpresent	208.0	...	12200.000000	3.400000	yes	yes	yes	poo
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	.
380	62.0	80.0	1.025	0.0	0.0	normal	normal	notpresent	notpresent	78.0	...	5400.000000	5.700000	no	no	no	goo
381	59.0	60.0	1.020	0.0	0.0	normal	normal	notpresent	notpresent	113.0	...	6500.000000	4.900000	no	no	no	goo
382	71.0	70.0	1.025	0.0	0.0	normal	normal	notpresent	notpresent	79.0	...	5800.000000	5.900000	no	no	no	goo
384	80.0	80.0	1.025	0.0	0.0	normal	normal	notpresent	notpresent	119.0	...	5100.000000	5.000000	no	no	no	goo
386	63.0	70.0	1.020	0.0	0.0	normal	normal	notpresent	notpresent	113.0	...	8000.000000	5.100000	no	no	no	goo

161 rows × 26 columns

```
subset2 = df.loc[df['pc']=='abnormal', ['age', 'bp','pc']]
subset2
```

	age	bp	pc
4	48.0	70.000000	abnormal
8	24.0	76.469072	abnormal
9	52.0	100.000000	abnormal
10	53.0	90.000000	abnormal
11	50.0	60.000000	abnormal
...	...	...	...
243	69.0	70.000000	abnormal
245	64.0	90.000000	abnormal
248	54.0	90.000000	abnormal
249	59.0	70.000000	abnormal
250	56.0	90.000000	abnormal

76 rows × 3 columns

```
subset3= df.loc[df['wc']>6000, ['pc', 'pcc', 'ba', 'wc']]
subset3
```

	pc	pcc	ba	wc
1	normal	notpresent	notpresent	7800.0
3	normal	notpresent	notpresent	7500.0
4	abnormal	present	notpresent	6700.0
5	normal	notpresent	notpresent	7300.0
6	normal	notpresent	notpresent	7800.0
...	...	...	...	...
396	normal	notpresent	notpresent	6700.0
397	normal	notpresent	notpresent	7800.0
398	normal	notpresent	notpresent	6600.0
399	normal	notpresent	notpresent	7200.0
400	normal	notpresent	notpresent	6800.0

342 rows × 4 columns

```
subset4= df[df['pe']=='yes']
subset4
```

	age	bp	sg	al	su	rbc	pc	pcc	ba	bgr	...	wc	rc	htn	dm	cad
4	48.0	70.000000	1.005	4.0	0.0	normal	abnormal	present	notpresent	117.0	...	6700.000000	3.900000	yes	no	no
6	60.0	90.000000	1.015	3.0	0.0	normal	normal	notpresent	notpresent	74.0	...	7800.000000	4.400000	yes	yes	no
8	24.0	76.469072	1.015	2.0	4.0	normal	abnormal	notpresent	notpresent	410.0	...	6900.000000	5.000000	no	yes	no
12	63.0	70.000000	1.010	3.0	0.0	abnormal	abnormal	present	notpresent	380.0	...	4500.000000	3.800000	yes	yes	no
13	68.0	70.000000	1.015	3.0	1.0	normal	normal	present	notpresent	208.0	...	12200.000000	3.400000	yes	yes	yes
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
238	80.0	70.000000	1.015	2.0	2.0	normal	normal	notpresent	notpresent	141.0	...	9600.000000	4.707435	yes	yes	no
241	65.0	70.000000	1.015	1.0	0.0	normal	normal	notpresent	notpresent	203.0	...	5000.000000	4.100000	yes	yes	no
243	69.0	70.000000	1.010	4.0	3.0	normal	abnormal	present	present	214.0	...	11500.000000	3.300000	yes	yes	yes
248	54.0	90.000000	1.025	1.0	0.0	normal	abnormal	notpresent	notpresent	150.0	...	8406.122449	4.707435	no	no	no
250	56.0	90.000000	1.010	4.0	1.0	normal	abnormal	present	notpresent	176.0	...	5400.000000	2.100000	yes	yes	no

76 rows × 26 columns

```
subset5 = df.iloc[100:200]
subset5
```

	age	bp	sg	al	su	rbc	pc	pcc	ba	bgr	...	wc	rc
101	34.0	70.0	1.015000	4.000000	0.000000	abnormal	abnormal	notpresent	notpresent	153.000000	...	8406.122449	4.707435
102	71.0	90.0	1.015000	2.000000	0.000000	normal	abnormal	present	present	88.000000	...	10700.000000	3.900000
103	17.0	60.0	1.010000	0.000000	0.000000	normal	normal	notpresent	notpresent	92.000000	...	7000.000000	4.707435
104	76.0	70.0	1.015000	2.000000	0.000000	normal	abnormal	present	notpresent	226.000000	...	12700.000000	4.200000
105	55.0	90.0	1.017408	1.016949	0.450142	normal	normal	notpresent	notpresent	143.000000	...	8406.122449	4.707435
...	...	...	...	...	...	...	...	...	...	...	...	...	...
196	70.0	90.0	1.020000	2.000000	1.000000	abnormal	abnormal	notpresent	present	184.000000	...	8406.122449	4.707435
197	49.0	100.0	1.010000	3.000000	0.000000	abnormal	abnormal	notpresent	notpresent	129.000000	...	9600.000000	3.500000
198	57.0	80.0	1.017408	1.016949	0.450142	normal	normal	notpresent	notpresent	148.036517	...	4300.000000	3.000000
199	59.0	100.0	1.020000	4.000000	2.000000	normal	normal	notpresent	notpresent	252.000000	...	26400.000000	3.900000
200	65.0	80.0	1.015000	0.000000	0.000000	normal	normal	notpresent	notpresent	92.000000	...	10700.000000	3.200000

100 rows × 26 columns

```
Start coding or generate with AI.
```

```
sns.pairplot(df)
```

```
<seaborn.axisgrid.PairGrid at 0x20240a08ec0>
```