In [81]:
```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
```

In [184]:
```
df=pd.read_csv(r"C:\USERS\user\Downloads\C6_bmi - C6_bmi.csv")
```

Out[184]:

	Gender	Height	Weight	Index
0	Male	174	96	4
1	Male	189	87	2
2	Female	185	110	4
3	Female	195	104	3
4	Male	149	61	3
...	...	...	...	...
495	Female	150	153	5
496	Female	184	121	4
497	Female	141	136	5
498	Male	150	95	5
499	Male	173	131	5

500 rows × 4 columns

In [185]:
```
df.columns
```

Out[185]:
```
Index(['Gender', 'Height', 'Weight', 'Index'], dtype='object')
```

In [186]:
```
df=df.head(21)
```

Out[186]:

	Gender	Height	Weight	Index
0	Male	174	96	4
1	Male	189	87	2
2	Female	185	110	4
3	Female	195	104	3
4	Male	149	61	3
5	Male	189	104	3
6	Male	147	92	5
7	Male	154	111	5
8	Male	174	90	3
9	Female	169	103	4
10	Male	195	81	2
11	Female	159	80	4
12	Female	192	101	3
13	Male	155	51	2
14	Male	191	79	2
15	Female	153	107	5
16	Female	157	110	5
17	Male	140	129	5
18	Male	144	145	5
19	Male	172	139	5
20	Male	157	110	5

In [187]:
```
a=df[['Gender', 'Height', 'Weight', 'Index']]
```

Out[187]:

	Gender	Height	Weight	Index
0	Male	174	96	4
1	Male	189	87	2
2	Female	185	110	4
3	Female	195	104	3
4	Male	149	61	3
5	Male	189	104	3
6	Male	147	92	5
7	Male	154	111	5
8	Male	174	90	3
9	Female	169	103	4
10	Male	195	81	2
11	Female	159	80	4
12	Female	192	101	3
13	Male	155	51	2
14	Male	191	79	2
15	Female	153	107	5
16	Female	157	110	5
17	Male	140	129	5
18	Male	144	145	5
19	Male	172	139	5
20	Male	157	110	5

In [188]:

Out[188]:
```
Male      14
Female     7
Name: Gender, dtype: int64
```

In [189]:
```
x=a.drop('Gender',axis=1)
```

In [190]:
```
g1={"Gender":{'Male':1,'Female':2}}
a=a.replace(g1)
```

	Gender	Height	Weight	Index
0	1	174	96	4
1	1	189	87	2
2	2	185	110	4
3	2	195	104	3
4	1	149	61	3
5	1	189	104	3
6	1	147	92	5
7	1	154	111	5
8	1	174	90	3
9	2	169	103	4
10	1	195	81	2
11	2	159	80	4
12	2	192	101	3
13	1	155	51	2
14	1	191	79	2
15	2	153	107	5
16	2	157	110	5
17	1	140	129	5
18	1	144	145	5
19	1	172	139	5
20	1	157	110	5

In [191]:
```
from sklearn.model_selection import train_test_split
```

In [192]:
```
from sklearn.ensemble import RandomForestClassifier

rfc=RandomForestClassifier()
```

Out[192]:
```
RandomForestClassifier()
```

In [198]:
```
parameters={'max_depth':[1,2,3,4,5],
            'min_samples_leaf':[5,10,15,20,25],
```

In [199]:
```
from sklearn.model_selection import GridSearchCV

grid_search=GridSearchCV(estimator=rfc,param_grid=parameters,cv=2,scoring="acc
```

Out[199]:
```
GridSearchCV(cv=2, estimator=RandomForestClassifier(),
             param_grid={'max_depth': [1, 2, 3, 4, 5],
                         'min_samples_leaf': [5, 10, 15, 20, 25],
                         'n_estimators': [10, 20, 30, 40, 50]},
             scoring='accuracy')
```

In [200]:

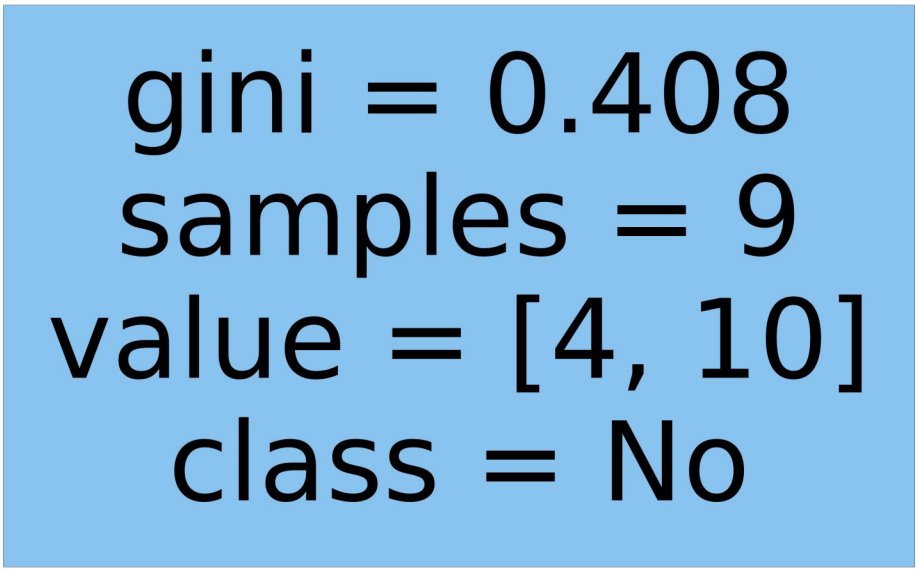
Out[200]:
```
0.6428571428571428
```

In [201]:

In [203]:
```
from sklearn.tree import plot_tree

plt.figure(figsize=(80,40))
```

Out[203]:
```
[Text(2232.0, 1087.2, 'gini = 0.408\nsamples = 9\nvalue = [4, 10]\nclass = No
')]
```



In [ ]: