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

In [204]:
```python
df=pd.read_csv(r"C:\USERS\user\Downloads\C8_loan-test - C8_loan-test.csv")
```

Out[204]:

	Loan_ID	Gender	Married	Dependents	Education	Self_Employed	ApplicantIncome	Coap
0	LP001015	Male	Yes	0	Graduate	No	5720	
1	LP001022	Male	Yes	1	Graduate	No	3076	
2	LP001031	Male	Yes	2	Graduate	No	5000	
3	LP001035	Male	Yes	2	Graduate	No	2340	
4	LP001051	Male	No	0	Not Graduate	No	3276	
...	...	...	...	...	...	...	...	
362	LP002971	Male	Yes	3+	Not Graduate	Yes	4009	
363	LP002975	Male	Yes	0	Graduate	No	4158	
364	LP002980	Male	No	0	Graduate	No	3250	
365	LP002986	Male	Yes	0	Graduate	No	5000	
366	LP002989	Male	No	0	Graduate	Yes	9200	

367 rows × 12 columns

In [205]:

Out[205]:
```
Index(['Loan_ID', 'Gender', 'Married', 'Dependents', 'Education',
       'Self_Employed', 'ApplicantIncome', 'CoapplicantIncome', 'LoanAmount',
       'Loan_Amount_Term', 'Credit_History', 'Property_Area'],
      dtype='object')
```

In [206]: df=df.head(21)

Out[206]:

	Loan_ID	Gender	Married	Dependents	Education	Self_Employed	ApplicantIncome	Coapp
0	LP001015	Male	Yes	0	Graduate	No	5720	
1	LP001022	Male	Yes	1	Graduate	No	3076	
2	LP001031	Male	Yes	2	Graduate	No	5000	
3	LP001035	Male	Yes	2	Graduate	No	2340	
4	LP001051	Male	No	0	Not Graduate	No	3276	
5	LP001054	Male	Yes	0	Not Graduate	Yes	2165	
6	LP001055	Female	No	1	Not Graduate	No	2226	
7	LP001056	Male	Yes	2	Not Graduate	No	3881	
8	LP001059	Male	Yes	2	Graduate	NaN	13633	
9	LP001067	Male	No	0	Not Graduate	No	2400	
10	LP001078	Male	No	0	Not Graduate	No	3091	
11	LP001082	Male	Yes	1	Graduate	NaN	2185	
12	LP001083	Male	No	3+	Graduate	No	4166	
13	LP001094	Male	Yes	2	Graduate	NaN	12173	
14	LP001096	Female	No	0	Graduate	No	4666	
15	LP001099	Male	No	1	Graduate	No	5667	
16	LP001105	Male	Yes	2	Graduate	No	4583	
17	LP001107	Male	Yes	3+	Graduate	No	3786	
18	LP001108	Male	Yes	0	Graduate	No	9226	
19	LP001115	Male	No	0	Graduate	No	1300	
20	LP001121	Male	Yes	1	Not Graduate	No	1888	

In [208]: `a=df[['Gender', 'ApplicantIncome', 'CoapplicantIncome','LoanAmount','Loan_Amou`

Out[208]:

	Gender	ApplicantIncome	CoapplicantIncome	LoanAmount	Loan_Amount_Term
0	Male	5720	0	110.0	360.0
1	Male	3076	1500	126.0	360.0
2	Male	5000	1800	208.0	360.0
3	Male	2340	2546	100.0	360.0
4	Male	3276	0	78.0	360.0
5	Male	2165	3422	152.0	360.0
6	Female	2226	0	59.0	360.0
7	Male	3881	0	147.0	360.0
8	Male	13633	0	280.0	240.0
9	Male	2400	2400	123.0	360.0
10	Male	3091	0	90.0	360.0
11	Male	2185	1516	162.0	360.0
12	Male	4166	0	40.0	180.0
13	Male	12173	0	166.0	360.0
14	Female	4666	0	124.0	360.0
15	Male	5667	0	131.0	360.0
16	Male	4583	2916	200.0	360.0
17	Male	3786	333	126.0	360.0
18	Male	9226	7916	300.0	360.0
19	Male	1300	3470	100.0	180.0
20	Male	1888	1620	48.0	360.0

In [209]:

Out[209]:
```
Male      19
Female     2
Name: Gender, dtype: int64
```

In [210]: `x=a.drop('Gender',axis=1)`

```
In [211]: g1={"Gender":{'Male':1,'Female':2}}
          a=a.replace(g1)
```

	Gender	ApplicantIncome	CoapplicantIncome	LoanAmount	Loan_Amount_Term
0	1	5720	0	110.0	360.0
1	1	3076	1500	126.0	360.0
2	1	5000	1800	208.0	360.0
3	1	2340	2546	100.0	360.0
4	1	3276	0	78.0	360.0
5	1	2165	3422	152.0	360.0
6	2	2226	0	59.0	360.0
7	1	3881	0	147.0	360.0
8	1	13633	0	280.0	240.0
9	1	2400	2400	123.0	360.0
10	1	3091	0	90.0	360.0
11	1	2185	1516	162.0	360.0
12	1	4166	0	40.0	180.0
13	1	12173	0	166.0	360.0
14	2	4666	0	124.0	360.0
15	1	5667	0	131.0	360.0
16	1	4583	2916	200.0	360.0
17	1	3786	333	126.0	360.0
18	1	9226	7916	300.0	360.0
19	1	1300	3470	100.0	180.0
20	1	1888	1620	48.0	360.0

```
In [212]: from sklearn.model_selection import train_test_split
```

```
In [213]: from sklearn.ensemble import RandomForestClassifier

          rfc=RandomForestClassifier()
```

```
Out[213]: RandomForestClassifier()
```

```
In [214]: parameters={'max_depth':[1,2,3,4,5],
                      'min_samples_leaf':[5,10,15,20,25],
```

```
In [215]: from sklearn.model_selection import GridSearchCV

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

```
Out[215]: 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')
```
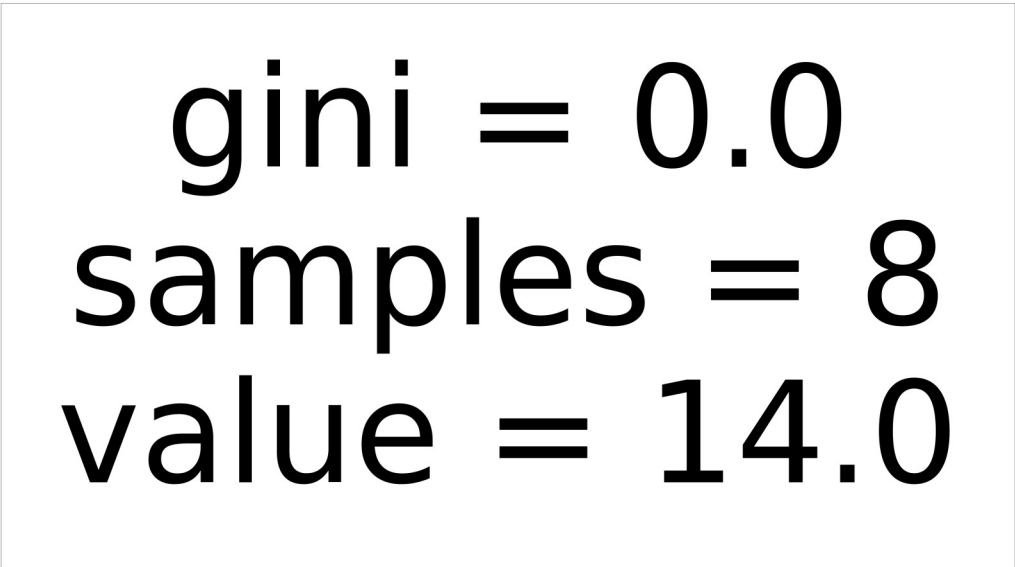
```
In [216]:
```

```
Out[216]: 1.0
```

```
In [217]:
```

In [218]:
```python
from sklearn.tree import plot_tree

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

Out[218]: [Text(2232.0, 1087.2, 'gini = 0.0\nsamples = 8\nvalue = 14.0')]

gini = 0.0
samples = 8
value = 14.0

In [ ]: