```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
from sklearn.model_selection import train_test_split
from sklearn.tree import DecisionTreeClassifier,plot_tree
from sklearn.metrics import accuracy_score, classification_report
```

Insert code cell below (Ctrl+M B)

```
data=sns.load_dataset('titanic')
```

```
data.dropna(inplace=True)
```

```
data['sex']=data['sex'].astype('category').cat.codes
data['class']=data['class'].astype('category').cat.codes
data['embarked']=data['embarked'].astype('category').cat.codes
```

```
x_train,x_test,y_train,y_test=train_test_split(x,y,test_size=0.2,random_state=42)
```

```
x=data.drop(['sex','class','embarked','who','adult_male','deck','embark_town','alive','alone'],axis=1)
y=data['survived']
```

```
clf=DecisionTreeClassifier()
clf.fit(x_train,y_train)
```

▾ DecisionTreeClassifier

```
DecisionTreeClassifier()
```

```
y_pred=clf.predict(x_test)
```

```
data
```

	survived	pclass	sex	age	sibsp	parch	fare	embarked	class	who	adult_male	deck	embark_town	alive	alone
1	1	1	0	38.0	1	0	71.2833	0	0	woman	False	C	Cherbourg	yes	False
3	1	1	0	35.0	1	0	53.1000	2	0	woman	False	C	Southampton	yes	False
6	0	1	1	54.0	0	0	51.8625	2	0	man	True	E	Southampton	no	True
10	1	3	0	4.0	1	1	16.7000	2	2	child	False	G	Southampton	yes	False
11	1	1	0	58.0	0	0	26.5500	2	0	woman	False	C	Southampton	yes	True
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
871	1	1	0	47.0	1	1	52.5542	2	0	woman	False	D	Southampton	yes	False
872	0	1	1	33.0	0	0	5.0000	2	0	man	True	B	Southampton	no	True
879	1	1	0	56.0	0	1	83.1583	0	0	woman	False	C	Cherbourg	yes	False
887	1	1	0	19.0	0	0	30.0000	2	0	woman	False	B	Southampton	yes	True
889	1	1	1	26.0	0	0	30.0000	0	0	man	True	C	Cherbourg	yes	True

182 rows × 15 columns

Next steps: Generate code with data | New interactive sheet

```
plt.figure(figsize=(8,5))
plot_tree(clf,
          filled=True,
          feature_names=x_train.columns,
          class_names=['Not Survived','Survived'],
          rounded=True,
          proportion=True,
          fontsize=8)
plt.tight_layout()
plt.show()
```

Insert code cell below (Ctrl+M B)