```
import pandas as pd
df = pd.read_csv("/content/tested.csv")
df.head()
```

	PassengerId	Survived	Pclass	Name	Sex	Age	SibSp	Parch	Ticket	Fare	Cabin	Embarked
0	892	0	3	Kelly, Mr. James	male	34.5	0	0	330911	7.8292	NaN	Q
1	893	1	3	Wilkes, Mrs. James (Ellen Needs)	female	47.0	1	0	363272	7.0000	NaN	S
2	894	0	2	Myles, Mr. Thomas Francis	male	62.0	0	0	240276	9.6875	NaN	Q
3	895	0	3	Wirz, Mr. Albert	male	27.0	0	0	315154	8.6625	NaN	S
4	896	1	3	Hirvonen, Mrs. Alexander (Helga E Lindqvist)	female	22.0	1	1	3101298	12.2875	NaN	S

```
df.info
```

```
<bound method DataFrame.info of      PassengerId  Survived  Pclass  \
0            892         0       3
1            893         1       3
2            894         0       2
3            895         0       3
4            896         1       3
..           ...       ...     ...
413         1305         0       3
414         1306         1       1
415         1307         0       3
416         1308         0       3
417         1309         0       3

                                          Name     Sex   Age  SibSp  Parch  \
0                              Kelly, Mr. James    male  34.5      0      0
1              Wilkes, Mrs. James (Ellen Needs)  female  47.0      1      0
2                     Myles, Mr. Thomas Francis    male  62.0      0      0
3                              Wirz, Mr. Albert    male  27.0      0      0
4    Hirvonen, Mrs. Alexander (Helga E Lindqvist)  female  22.0      1      1
..                                            ...     ...   ...    ...    ...
413                            Spector, Mr. Woolf    male   NaN      0      0
414                Oliva y Ocana, Dona. Fermina  female  39.0      0      0
415                Saether, Mr. Simon Sivertsen    male  38.5      0      0
416                         Ware, Mr. Frederick    male   NaN      0      0
417                  Peter, Master. Michael J    male   NaN      1      1

                Ticket      Fare Cabin Embarked
0               330911    7.8292   NaN        Q
1               363272    7.0000   NaN        S
2               240276    9.6875   NaN        Q
3               315154    8.6625   NaN        S
4              3101298   12.2875   NaN        S
..                 ...       ...   ...      ...
413           A.5. 3236    8.0500   NaN        S
414            PC 17758  108.9000  C105        C
415  SOTON/O.Q. 3101262    7.2500   NaN        S
416              359309    8.0500   NaN        S
417                2668   22.3583   NaN        C

[418 rows x 12 columns]>
```

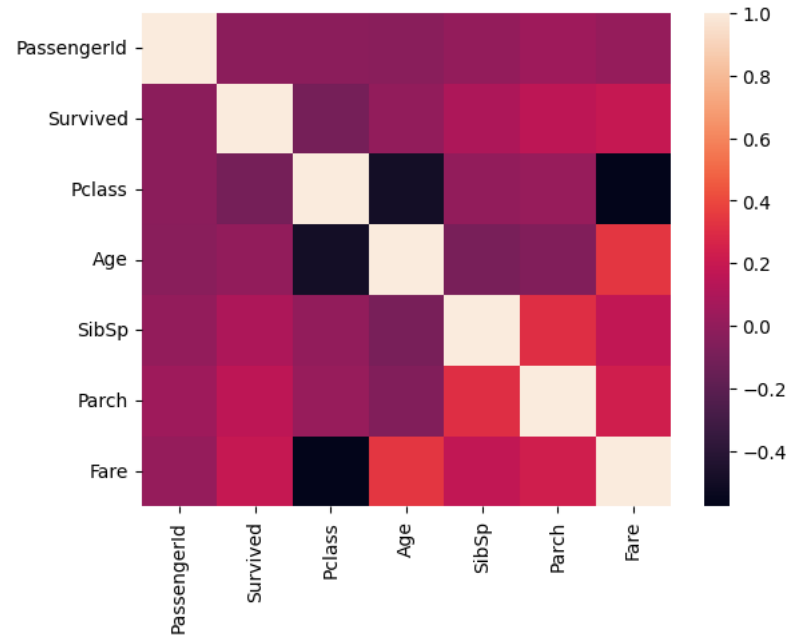
```
df.corr()
```

```
<ipython-input-3-2f6f6606aa2c>:1: FutureWarning: The default value of numeric_only in DataFrame.corr is deprecated. In a future ver
  df.corr()
```

	PassengerId	Survived	Pclass	Age	SibSp	Parch	Fare
PassengerId	1.000000	-0.023245	-0.026751	-0.034102	0.003818	0.043080	0.008211
Survived	-0.023245	1.000000	-0.108615	-0.000013	0.099943	0.159120	0.191514
Pclass	-0.026751	-0.108615	1.000000	-0.492143	0.001087	0.018721	-0.577147
Age	-0.034102	-0.000013	-0.492143	1.000000	-0.091587	-0.061249	0.337932
SibSp	0.003818	0.099943	0.001087	-0.091587	1.000000	0.306895	0.171539
Parch	0.043080	0.159120	0.018721	-0.061249	0.306895	1.000000	0.230046
Fare	0.008211	0.191514	-0.577147	0.337932	0.171539	0.230046	1.000000

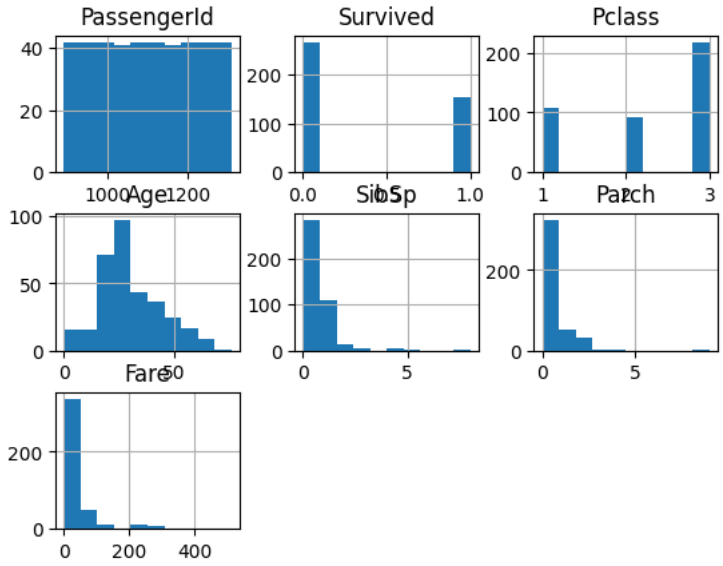
```
import seaborn as sns
sns.heatmap(df.corr())
```

```
<ipython-input-4-534f4f3c80b7>:2: FutureWarning: The default value of numeric_only in DataFrame.corr is deprecated. In a future ver
  sns.heatmap(df.corr())
<Axes: >
```

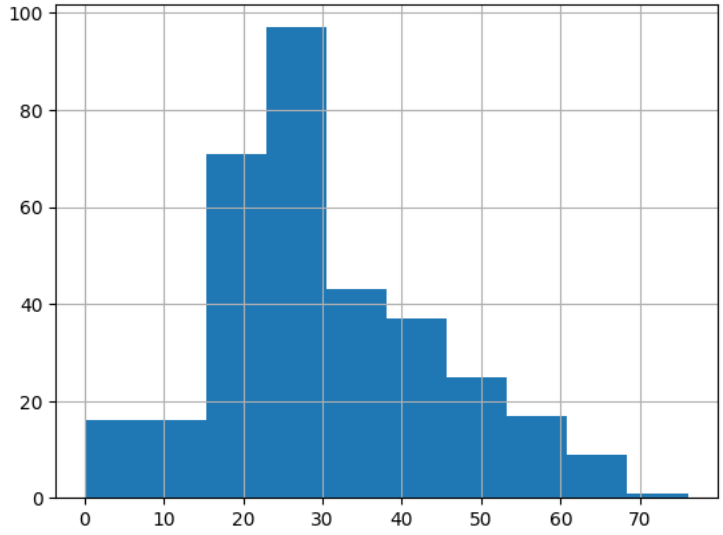



```
df.hist()
```

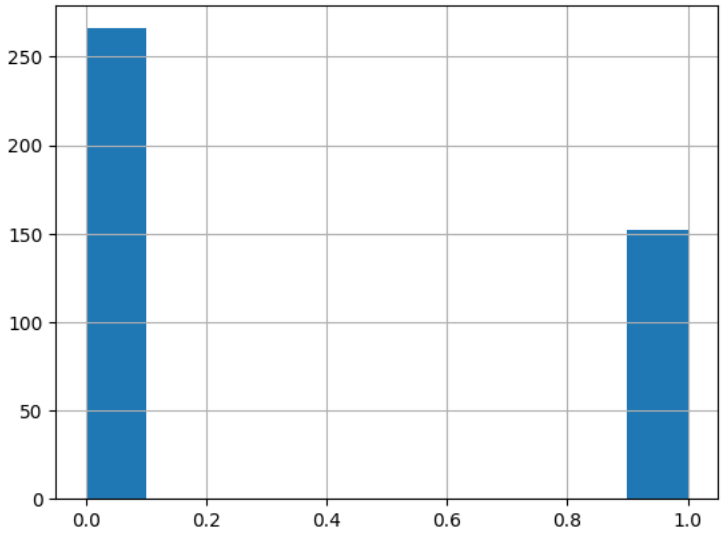
```
array([[<Axes: title={'center': 'PassengerId'}>,
        <Axes: title={'center': 'Survived'}>,
        <Axes: title={'center': 'Pclass'}>],
       [<Axes: title={'center': 'Age'}>,
        <Axes: title={'center': 'SibSp'}>,
        <Axes: title={'center': 'Parch'}>],
       [<Axes: title={'center': 'Fare'}>, <Axes: >, <Axes: >]],
      dtype=object)
```




```
df["Age"].hist();
```




```
df["Survived"].hist();
```




```
import matplotlib.pyplot as plt

sns.countplot(df, x="Survived", hue="Sex")
plt.xlabel('Survived')
plt.ylabel('Count')
```

```
Text(0, 0.5, 'Count')
```