```
import pandas as pd
```

```
data=pd.read_csv("titanic.csv")
```

```
data.head()
```

	PassengerId	Survived	Pclass	Name	Sex	Age	SibSp	Parch	Ticket	Fare
0	1	0	3	Braund, Mr. Owen Harris	male	22.0	1	0	A/5 21171	7.250
1	2	1	1	Cumings, Mrs. John Bradley (Florence Briggs	female	38.0	1	0	PC 17599	71.283

```
print(data)
```

```
     PassengerId  Survived  Pclass  \
0              1         0       3
1              2         1       1
2              3         1       3
3              4         1       1
4              5         0       3
..           ...       ...     ...
886          887         0       2
887          888         1       1
888          889         0       3
889          890         1       1
890          891         0       3

                                                 Name     Sex   Age  SibSp  \
0                             Braund, Mr. Owen Harris    male  22.0      1
1    Cumings, Mrs. John Bradley (Florence Briggs Th...  female  38.0      1
2                              Heikkinen, Miss. Laina  female  26.0      0
3          Futrelle, Mrs. Jacques Heath (Lily May Peel)  female  35.0      1
4                            Allen, Mr. William Henry    male  35.0      0
..                                                ...     ...   ...    ...
886                            Montvila, Rev. Juozas    male  27.0      0
887                     Graham, Miss. Margaret Edith  female  19.0      0
888         Johnston, Miss. Catherine Helen "Carrie"  female   NaN      1
889                                Behr, Mr. Karl Howell    male  26.0      0
890                                  Dooley, Mr. Patrick    male  32.0      0

     Parch            Ticket     Fare Cabin Embarked
0        0         A/5 21171   7.2500   NaN        S
1        0          PC 17599  71.2833   C85        C
2        0  STON/O2. 3101282   7.9250   NaN        S
3        0            113803  53.1000  C123        S
4        0            373450   8.0500   NaN        S
..     ...               ...      ...   ...      ...
886      0            211536  13.0000   NaN        S
887      0            112053  30.0000   B42        S
888      2        W./C. 6607  23.4500   NaN        S
889      0            111369  30.0000  C148        C
890      0            370376   7.7500   NaN        Q

[891 rows x 12 columns]
```

```
data.isnull()
```

	PassengerId	Survived	Pclass	Name	Sex	Age	SibSp	Parch	Ticket	Fare	Cabi
0	False	False	False	False	False	False	False	False	False	False	Tru
1	False	False	False	False	False	False	False	False	False	False	Fals
2	False	False	False	False	False	False	False	False	False	False	Tru
3	False	False	False	False	False	False	False	False	False	False	Fals
4	False	False	False	False	False	False	False	False	False	False	Tru
...	...	...	...	...	...	...	...	...	...	...	...
886	False	False	False	False	False	False	False	False	False	False	Tru
887	False	False	False	False	False	False	False	False	False	False	Fals
888	False	False	False	False	False	True	False	False	False	False	Tru
889	False	False	False	False	False	False	False	False	False	False	Fals
890	False	False	False	False	False	False	False	False	False	False	Tru

891 rows × 12 columns

```
Data_temp=data
```

```
data.info()
```

```
---------------------------------------------------------------------------
AttributeError                            Traceback (most recent call last)
<ipython-input-31-6208d269f320> in <cell line: 1>()
----> 1 data.info()

AttributeError: 'function' object has no attribute 'info'
```

SEARCH STACK OVERFLOW

```
data.dropna()
```

	PassengerId	Survived	Pclass	Name	Sex	Age	SibSp	Parch	Ticket	Far
1	2	1	1	Cumings, Mrs. John Bradley (Florence Briggs Th...	female	38.0	1	0	PC 17599	71.283
3	4	1	1	Futrelle, Mrs. Jacques Heath (Lily May Peel)	female	35.0	1	0	113803	53.100
6	7	0	1	McCarthy, Mr. Timothy J	male	54.0	0	0	17463	51.862
10	11	1	3	Sandstrom, Miss. Marguerite Rut	female	4.0	1	1	PP 9549	16.700
				Bonnell						

```
data.dropna(axis=0)
#if we write axis=1 it will delete the column means a attribute or feature whole column in which if we had atleast 1 null value it
#will deleted  ut if we write axis=0 it will dele the row which are null.
```

	PassengerId	Survived	Pclass	Name	Sex	Age	SibSp	Parch	Ticket	Far
1	2	1	1	Cumings, Mrs. John Bradley (Florence Briggs Th...	female	38.0	1	0	PC 17599	71.283
3	4	1	1	Futrelle, Mrs. Jacques Heath (Lily May Peel)	female	35.0	1	0	113803	53.100
6	7	0	1	McCarthy, Mr. Timothy J	male	54.0	0	0	17463	51.862
10	11	1	3	Sandstrom, Miss. Marguerite Rut	female	4.0	1	1	PP 9549	16.700
				Bonnell						

```
data=data.drop
```

```
data.describe()
```

```
---------------------------------------------------------------------------
AttributeError                            Traceback (most recent call last)
<ipython-input-21-2bb0b18689d4> in <cell line: 1>()
----> 1 data.describe()

AttributeError: 'function' object has no attribute 'describe'
```

SEARCH STACK OVERFLOW

```
Data_temp=Data_temp["Age"].fillna(30)
```

```
print(data)
```

```
<bound method DataFrame.drop of      PassengerId  Survived  Pclass  \
0              1         0       3
1              2         1       1
2              3         1       3
3              4         1       1
4              5         0       3
..           ...       ...     ...
886          887         0       2
887          888         1       1
888          889         0       3
889          890         1       1
890          891         0       3

                                                 Name     Sex   Age  SibSp  \
0                             Braund, Mr. Owen Harris    male  22.0      1
1    Cumings, Mrs. John Bradley (Florence Briggs Th...  female  38.0      1
2                              Heikkinen, Miss. Laina  female  26.0      0
3        Futrelle, Mrs. Jacques Heath (Lily May Peel)  female  35.0      1
4                            Allen, Mr. William Henry    male  35.0      0
..                                                ...     ...   ...    ...
886                            Montvila, Rev. Juozas    male  27.0      0
887                     Graham, Miss. Margaret Edith  female  19.0      0
888         Johnston, Miss. Catherine Helen "Carrie"  female   NaN      1
889                            Behr, Mr. Karl Howell    male  26.0      0
890                              Dooley, Mr. Patrick    male  32.0      0

     Parch            Ticket     Fare Cabin Embarked
0        0         A/5 21171   7.2500   NaN        S
1        0          PC 17599  71.2833   C85        C
2        0  STON/O2. 3101282   7.9250   NaN        S
3        0            113803  53.1000  C123        S
4        0            373450   8.0500   NaN        S
..     ...               ...      ...   ...      ...
886      0            211536  13.0000   NaN        S
887      0            112053  30.0000   B42        S
888      2        W./C. 6607  23.4500   NaN        S
889      0            111369  30.0000  C148        C
```

```
890      0            370376   7.7500   NaN        Q

[891 rows x 12 columns]>
```

```
data=pd.read_csv("\titanic.csv")
```

```
---------------------------------------------------------------------------
FileNotFoundError                         Traceback (most recent call last)
<ipython-input-26-fd5a6c88d20c> in <cell line: 1>()
----> 1 data=pd.read_csv("\titanic.csv")

6 frames
/usr/local/lib/python3.10/dist-packages/pandas/io/common.py in get_handle(path_or_buf,
mode, encoding, compression, memory_map, is_text, errors, storage_options)
    854         if ioargs.encoding and "b" not in ioargs.mode:
    855             # Encoding
--> 856             handle = open(
    857                 handle,
    858                 ioargs.mode,

FileNotFoundError: [Errno 2] No such file or directory: '\titanic.csv'
```

SEARCH STACK OVERFLOW

```
Data_temp["Pclass"]=Data_temp["Pclass"].fillna("temp")
```

```
Data_temp["Age"]=Data_temp["Age"].fillna(30)
```

```
Data_temp["Cabin"]=Data_temp["Cabin"].fillna(5)
```

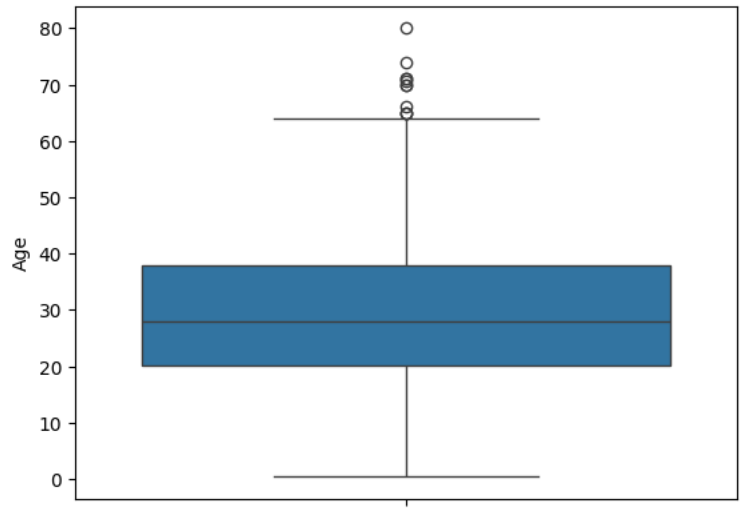
```
#box plot
```

```
data=pd.read_csv("titanic.csv")
```

```
import seaborn as sns
```

```
sns.boxplot(data["Age"])
```

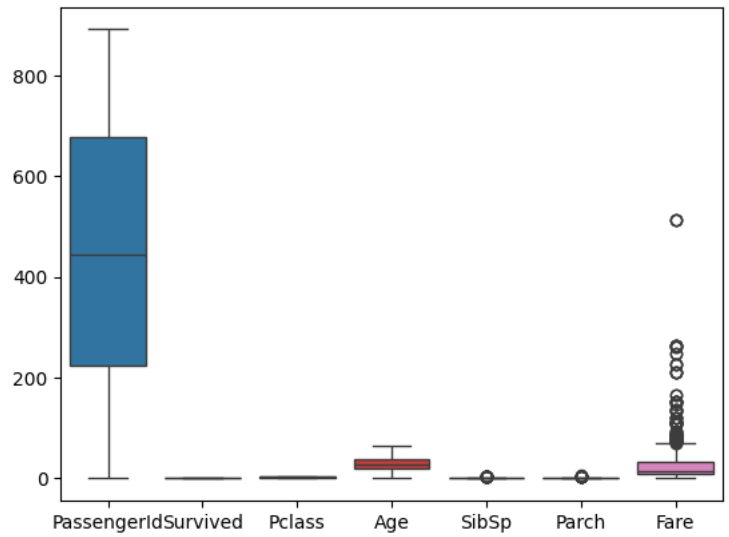
```
<Axes: ylabel='Age'>
```




```
removedoutliers=data[data["Age"]<65]
```

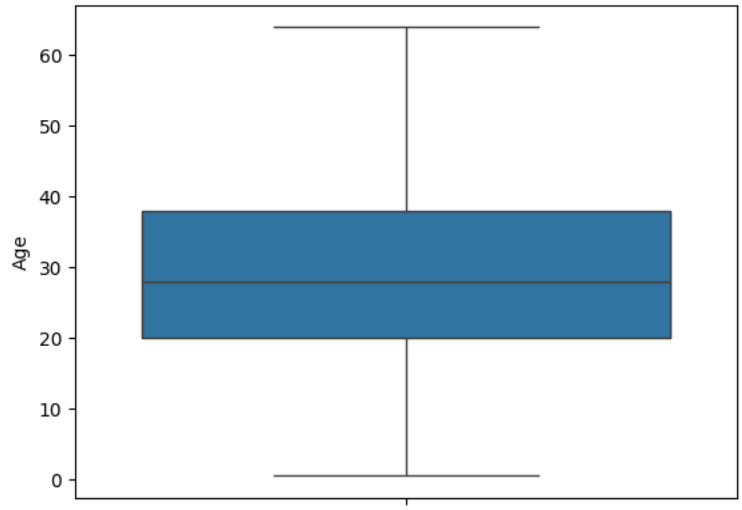
```
sns.boxplot(removedoutliers)
```

```
<Axes: >
```




```
sns.boxplot(removedoutliers["Age"])
```

```
<Axes: ylabel='Age'>
```

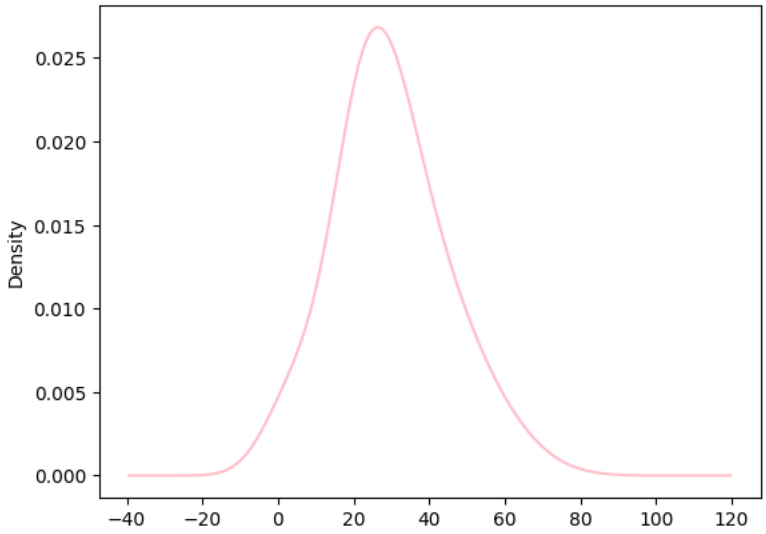



```
new_data=data[data["Age"]>65]


#duplicates
data=data.sort_values("Name")


dup=data["Name"].duplicated()


data[dup]
```

	PassengerId	Survived	Pclass	Name	Sex	Age	SibSp	Parch	Ticket	Fare	Cabin	Emba

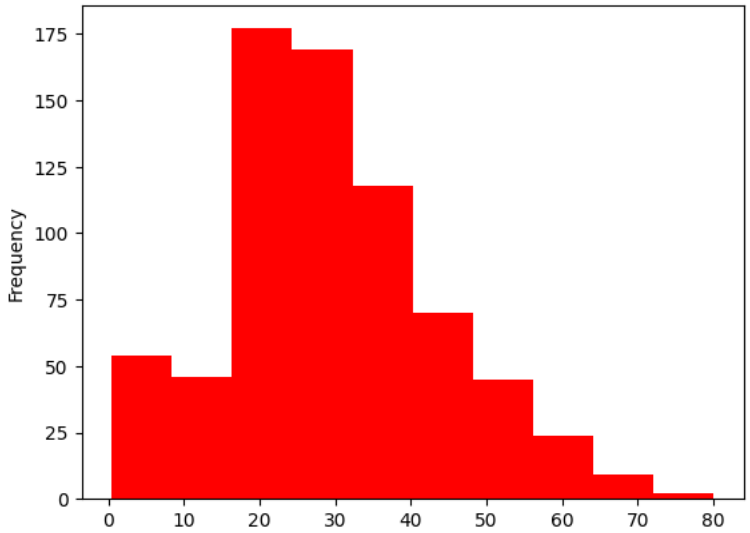
```
data_age=data["Age"]
scaler=MinMax Scaler()



data=pd.read_csv("titanic.csv")


print=(data.isnull().sum())


data["Age"].plot.density(bw_method=0.5,color="pink")
#you can change the color as per your choice
```

```
<Axes: ylabel='Density'>
```




```
data["Age"].plot.hist(color="red")
```

```
<Axes: ylabel='Frequency'>
```




```
data["Age"].plot.pie(color="black")
```

```
---------------------------------------------------------------------------
TypeError                                 Traceback (most recent call last)
<ipython-input-20-8bc78f0aecdf> in <cell line: 1>()
----> 1 data["Age"].plot.pie(color="black")

                                  5 frames
/usr/local/lib/python3.10/dist-packages/matplotlib/__init__.py in inner(ax, data, *args, **kwargs)
```