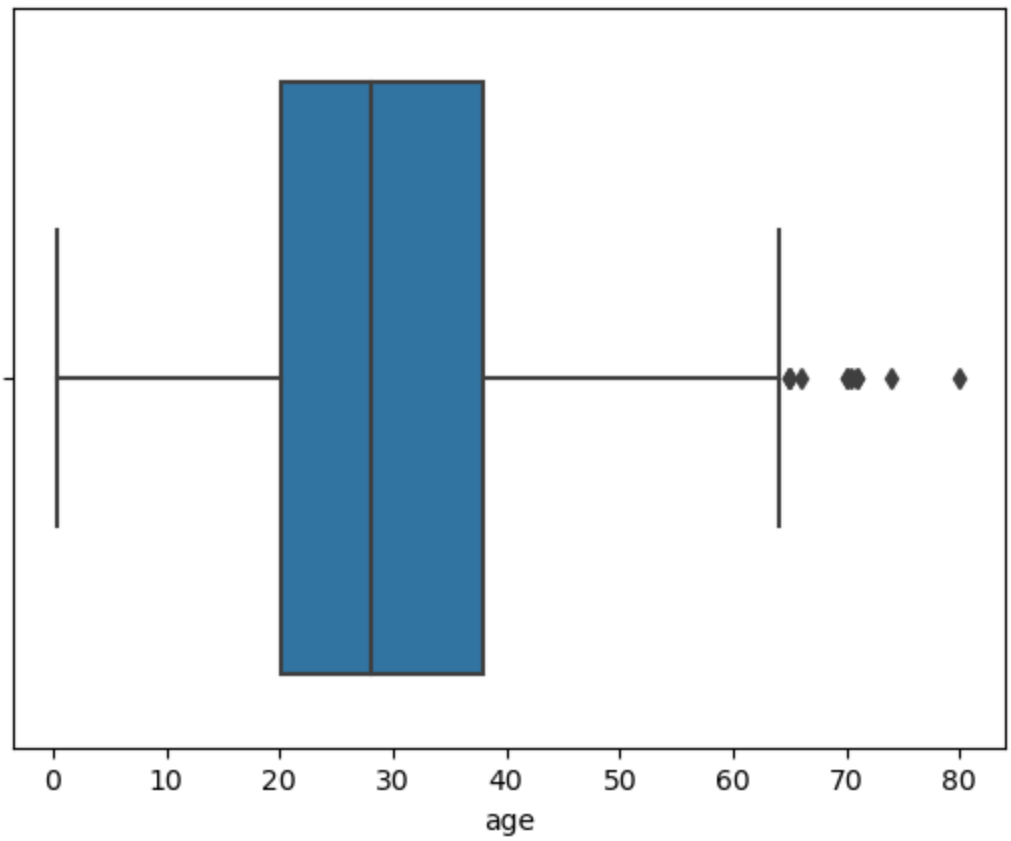
In [ ]:
```
##OUTLIERS
```

In [1]:
```
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt
pd.set_option('display.max_column',None) ##... göstermesin diye
pd.set_option('display.width',700) #sağ tarafa doğru daha fazla sütun(columns)
df=sns.load_dataset("titanic")
df.head()
```

Out[1]:

	survived	pclass	sex	age	sibsp	parch	fare	embarked	class	who	adult_male	d
0	0	3	male	22.0	1	0	7.2500	S	Third	man	True	
1	1	1	female	38.0	1	0	71.2833	C	First	woman	False	
2	1	3	female	26.0	0	0	7.9250	S	Third	woman	False	
3	1	1	female	35.0	1	0	53.1000	S	First	woman	False	
4	0	3	male	35.0	0	0	8.0500	S	Third	man	True	

In [4]:
```
sns.boxplot(x=df["age"])
plt.show()
```



In [5]:
```
q1 = df["age"].quantile(0.25)
```

In [6]:
```
q1
```

Out[6]: 20.125

In [7]:
```
q3= df["age"].quantile(0.75)
```

In [8]:
```
q3
```

Out[8]: 38.0

In [9]:
```
iqr = q3-q1
```

In [10]:
```
up = q3+1.5*iqr
low = q1-1.5*iqr
```

In [11]:
```
up
```

Out[11]: 64.8125

In [14]:
```
df[(df["age"]>up) | (df["age"]<low)]
```

Out[14]:

	survived	pclass	sex	age	sibsp	parch	fare	embarked	class	who	adult_male	d
33	0	2	male	66.0	0	0	10.5000	S	Second	man	True	
54	0	1	male	65.0	0	1	61.9792	C	First	man	True	
96	0	1	male	71.0	0	0	34.6542	C	First	man	True	
116	0	3	male	70.5	0	0	7.7500	Q	Third	man	True	
280	0	3	male	65.0	0	0	7.7500	Q	Third	man	True	
456	0	1	male	65.0	0	0	26.5500	S	First	man	True	
493	0	1	male	71.0	0	0	49.5042	C	First	man	True	
630	1	1	male	80.0	0	0	30.0000	S	First	man	True	
672	0	2	male	70.0	0	0	10.5000	S	Second	man	True	
745	0	1	male	70.0	1	1	71.0000	S	First	man	True	
851	0	3	male	74.0	0	0	7.7750	S	Third	man	True	

In [15]:
```
df[(df["age"]>up) | (df["age"]<low)].index
```

Out[15]: Index([33, 54, 96, 116, 280, 456, 493, 630, 672, 745, 851], dtype='int64')

In [ ]:
```
##SİLME
```

In [18]:
```
df.shape
```

Out[18]: (891, 15)

In [20]:
```
df[~((df["fare"]<low)|(df["fare"]>up))].shape
```

Out[20]: (773, 15)

In [21]:
```
df=df[~((df["fare"]<low)|(df["fare"]>up))]
```

In [22]:
```
df
```

Out[22]:

	survived	pclass	sex	age	sibsp	parch	fare	embarked	class	who	adult_mal
0	0	3	male	22.0	1	0	7.2500	S	Third	man	Tru
2	1	3	female	26.0	0	0	7.9250	S	Third	woman	Fals
3	1	1	female	35.0	1	0	53.1000	S	First	woman	Fals
4	0	3	male	35.0	0	0	8.0500	S	Third	man	Tru
5	0	3	male	NaN	0	0	8.4583	Q	Third	man	Tru
...	...	...	...	...	...	...	...	...	...	...	
886	0	2	male	27.0	0	0	13.0000	S	Second	man	Tru
887	1	1	female	19.0	0	0	30.0000	S	First	woman	Fals
888	0	3	female	NaN	1	2	23.4500	S	Third	woman	Fals
889	1	1	male	26.0	0	0	30.0000	C	First	man	Tru
890	0	3	male	32.0	0	0	7.7500	Q	Third	man	Tru

773 rows × 15 columns

In [23]:
```
df.loc[(df["fare"]>up),"fare"] = up
```

In [24]:
```
df
```

Out[24]:

	survived	pclass	sex	age	sibsp	parch	fare	embarked	class	who	adult_mal
0	0	3	male	22.0	1	0	7.2500	S	Third	man	Tru
2	1	3	female	26.0	0	0	7.9250	S	Third	woman	Fals
3	1	1	female	35.0	1	0	53.1000	S	First	woman	Fals
4	0	3	male	35.0	0	0	8.0500	S	Third	man	Tru
5	0	3	male	NaN	0	0	8.4583	Q	Third	man	Tru
...	...	...	...	...	...	...	...	...	...	...	
886	0	2	male	27.0	0	0	13.0000	S	Second	man	Tru
887	1	1	female	19.0	0	0	30.0000	S	First	woman	Fals
888	0	3	female	NaN	1	2	23.4500	S	Third	woman	Fals
889	1	1	male	26.0	0	0	30.0000	C	First	man	Tru
890	0	3	male	32.0	0	0	7.7500	Q	Third	man	Tru

773 rows × 15 columns

In [25]:
```
df.isnull().values.any()
```

Out[25]:
```
True
```

In [26]:
```
df.isnull().sum()
```

Out[26]:
```
survived         0
pclass           0
sex              0
age            162
sibsp            0
parch            0
fare             0
embarked         0
class            0
who              0
adult_male       0
deck           660
embark_town      0
alive            0
alone            0
dtype: int64
```

In [ ]:

In [ ]: