In [42]:
```
import pandas as pd
```

In [43]:
```
df=pd.read_csv("dataset_3.csv")
```

In [44]:
```
df.shape
```

Out[44]:
```
(32, 5)
```

In [45]:
```
df.columns
```

Out[45]:
```
Index(['Duration', 'Date', 'Pulse', 'Maxpulse', 'Calories'], dtype='object')
```

In [46]:
```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 32 entries, 0 to 31
Data columns (total 5 columns):
 #   Column    Non-Null Count  Dtype  
---  ------    --------------  -----  
 0   Duration  32 non-null     int64  
 1   Date      31 non-null     object 
 2   Pulse     32 non-null     int64  
 3   Maxpulse  32 non-null     int64  
 4   Calories  30 non-null     float64
dtypes: float64(1), int64(3), object(1)
memory usage: 1.4+ KB
```

In [47]:
```
df.head()
```

Out[47]:

	Duration	Date	Pulse	Maxpulse	Calories
0	60	'2020/12/01'	110	130	409.1
1	60	'2020/12/02'	117	145	479.0
2	60	'2020/12/03'	103	135	340.0
3	45	'2020/12/04'	109	175	282.4
4	45	'2020/12/05'	117	148	406.0

In [48]:
```
df.head(7)
```

Out[48]:

	Duration	Date	Pulse	Maxpulse	Calories
0	60	'2020/12/01'	110	130	409.1
1	60	'2020/12/02'	117	145	479.0
2	60	'2020/12/03'	103	135	340.0
3	45	'2020/12/04'	109	175	282.4
4	45	'2020/12/05'	117	148	406.0
5	60	'2020/12/06'	102	127	300.0
6	60	'2020/12/07'	110	136	374.0

In [49]:
```
df.tail(7)
```

Out[49]:

	Duration	Date	Pulse	Maxpulse	Calories
25	60	'2020/12/25'	102	126	334.5
26	60	20201226	100	120	250.0
27	60	'2020/12/27'	92	118	241.0
28	60	'2020/12/28'	103	132	NaN
29	60	'2020/12/29'	100	132	280.0
30	60	'2020/12/30'	102	129	380.3
31	60	'2020/12/31'	92	115	243.0

In [50]:
```
df["Calories"]
```

Out[50]:
```
0     409.1
1     479.0
2     340.0
3     282.4
4     406.0
5     300.0
6     374.0
7     253.3
8     195.1
9     269.0
10    329.3
11    250.7
12    250.7
13    345.3
14    379.3
15    275.0
16    215.2
17    300.0
18      NaN
19    323.0
20    243.0
21    364.2
22    282.0
23    300.0
24    246.0
25    334.5
26    250.0
27    241.0
28      NaN
29    280.0
30    380.3
31    243.0
Name: Calories, dtype: float64
```

In [60]:
```
df[["Calories","Date"]].head()
```

Out[60]:

	Calories	Date
0	409.1	'2020/12/01'
1	479.0	'2020/12/02'
2	340.0	'2020/12/03'
3	282.4	'2020/12/04'
4	406.0	'2020/12/05'

In [52]:
```
df.iloc[1:4]
```

Out[52]:

	Duration	Date	Pulse	Maxpulse	Calories
1	60	'2020/12/02'	117	145	479.0
2	60	'2020/12/03'	103	135	340.0
3	45	'2020/12/04'	109	175	282.4

In [53]:
```
df.iloc[1:4,0:3]
```

Out[53]:

	Duration	Date	Pulse
1	60	'2020/12/02'	117
2	60	'2020/12/03'	103
3	45	'2020/12/04'	109

In [54]:
```
df.iloc[[0,8,10]]
```

Out[54]:

	Duration	Date	Pulse	Maxpulse	Calories
0	60	'2020/12/01'	110	130	409.1
8	30	'2020/12/09'	109	133	195.1
10	60	'2020/12/11'	103	147	329.3

In [55]:
```
df.describe()
```

Out[55]:

	Duration	Pulse	Maxpulse	Calories
count	32.000000	32.000000	32.000000	30.000000
mean	68.437500	103.500000	128.500000	304.680000
std	70.039591	7.832933	12.998759	66.003779
min	30.000000	90.000000	101.000000	195.100000
25%	60.000000	100.000000	120.000000	250.700000
50%	60.000000	102.500000	127.500000	291.200000
75%	60.000000	106.500000	132.250000	343.975000
max	450.000000	130.000000	175.000000	479.000000

In [56]:
```
df2=df.drop("Date",axis=1)
```

In [57]:
```
df2.mean(0)
```

Out[57]:
```
Duration     68.4375
Pulse       103.5000
Maxpulse    128.5000
Calories    304.6800
dtype: float64
```

In [59]:
```
durationfunc=lambda x: x>=60
df2["Duration_mins_above_60"]=df["Duration"].apply(durationfunc)
df2.head()
```

Out[59]:

	Duration	Pulse	Maxpulse	Calories	Duration_mins_above_60
0	60	110	130	409.1	True
1	60	117	145	479.0	True
2	60	103	135	340.0	True
3	45	109	175	282.4	False
4	45	117	148	406.0	False