

```
In [1]: import pandas as pd
import pickle
import warnings
warnings.filterwarnings("ignore")
```

```
In [2]: a=pd.read_csv("calories.csv")
```

```
In [3]: a
```

```
Out[3]:
```

	User_ID	Gender	Age	Height	Weight	Duration	Heart_Rate	Body_Temp	Calories
0	14733363	male	68	190	94	29	105	40.8	231
1	14861698	female	20	166	60	14	94	40.3	66
2	11179863	male	69	179	79	5	88	38.7	26
3	16180408	female	34	179	71	13	100	40.5	71
4	17771927	female	27	154	58	10	81	39.8	35
...	...	...	...	...	...	...	...	...	...
14995	15644082	female	20	193	86	11	92	40.4	45
14996	17212577	female	27	165	65	6	85	39.2	23
14997	17271188	female	43	159	58	16	90	40.1	75
14998	18643037	male	78	193	97	2	84	38.3	11
14999	11751526	male	63	173	79	18	92	40.5	98

15000 rows × 9 columns

```
In [4]: a.head(10)
```

```
Out[4]:
```

	User_ID	Gender	Age	Height	Weight	Duration	Heart_Rate	Body_Temp	Calories
0	14733363	male	68	190	94	29	105	40.8	231
1	14861698	female	20	166	60	14	94	40.3	66
2	11179863	male	69	179	79	5	88	38.7	26
3	16180408	female	34	179	71	13	100	40.5	71
4	17771927	female	27	154	58	10	81	39.8	35
5	15130815	female	36	151	50	23	96	40.7	123
6	19602372	female	33	158	56	22	95	40.5	112
7	11117088	male	41	175	85	25	100	40.7	143
8	12132339	male	60	186	94	21	97	40.4	134
9	17964668	female	26	146	51	16	90	40.2	72

```
In [5]: a.tail(10)
```

Out[5]:

	User_ID	Gender	Age	Height	Weight	Duration	Heart_Rate	Body_Temp	Calories
14990	19715870	female	22	190	79	19	96	40.3	89
14991	10050978	male	51	181	87	9	91	39.6	44
14992	14722670	male	27	170	70	13	92	40.1	46
14993	13584585	male	45	179	78	11	98	39.9	60
14994	18209611	female	48	159	57	10	94	39.8	52
14995	15644082	female	20	193	86	11	92	40.4	45
14996	17212577	female	27	165	65	6	85	39.2	23
14997	17271188	female	43	159	58	16	90	40.1	75
14998	18643037	male	78	193	97	2	84	38.3	11
14999	11751526	male	63	173	79	18	92	40.5	98

In [6]: `a.info()`

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 15000 entries, 0 to 14999
Data columns (total 9 columns):
#   Column          Non-Null Count  Dtype
---  -
0   User_ID         15000 non-null  int64
1   Gender          15000 non-null  object
2   Age             15000 non-null  int64
3   Height          15000 non-null  int64
4   Weight          15000 non-null  int64
5   Duration        15000 non-null  int64
6   Heart_Rate      15000 non-null  int64
7   Body_Temp       15000 non-null  float64
8   Calories        15000 non-null  int64
dtypes: float64(1), int64(7), object(1)
memory usage: 1.0+ MB
```

In [7]: `a.shape`

Out[7]: (15000, 9)

In [8]: `a.isnull().sum()`

Out[8]:

```
User_ID      0
Gender       0
Age          0
Height       0
Weight       0
Duration     0
Heart_Rate   0
Body_Temp    0
Calories     0
dtype: int64
```

In [ ]:

In [9]: `b=a.drop('User_ID',axis=1)`  
**b**

Out[9]:

	Gender	Age	Height	Weight	Duration	Heart_Rate	Body_Temp	Calories
0	male	68	190	94	29	105	40.8	231
1	female	20	166	60	14	94	40.3	66
2	male	69	179	79	5	88	38.7	26
3	female	34	179	71	13	100	40.5	71
4	female	27	154	58	10	81	39.8	35
...	...	...	...	...	...	...	...	...
14995	female	20	193	86	11	92	40.4	45
14996	female	27	165	65	6	85	39.2	23
14997	female	43	159	58	16	90	40.1	75
14998	male	78	193	97	2	84	38.3	11
14999	male	63	173	79	18	92	40.5	98

15000 rows × 8 columns

In [10]:

```
b=pd.get_dummies(b, dtype=int)
```

In [11]:

```
b
```

Out[11]:

	Age	Height	Weight	Duration	Heart_Rate	Body_Temp	Calories	Gender_female	Gender
0	68	190	94	29	105	40.8	231	0	
1	20	166	60	14	94	40.3	66	1	
2	69	179	79	5	88	38.7	26	0	
3	34	179	71	13	100	40.5	71	1	
4	27	154	58	10	81	39.8	35	1	
...	...	...	...	...	...	...	...	...	...
14995	20	193	86	11	92	40.4	45	1	
14996	27	165	65	6	85	39.2	23	1	
14997	43	159	58	16	90	40.1	75	1	
14998	78	193	97	2	84	38.3	11	0	
14999	63	173	79	18	92	40.5	98	0	

15000 rows × 9 columns

◀

▶

In [12]:

```
y=b['Calories']  
x=b.drop('Calories', axis=1)
```

In [13]:

```
y
```

```
Out[13]: 0      231
          1      66
          2      26
          3      71
          4      35
          ...
14995     45
14996     23
14997     75
14998     11
14999     98
Name: Calories, Length: 15000, dtype: int64
```

```
In [14]: x
```

```
Out[14]:
```

	Age	Height	Weight	Duration	Heart_Rate	Body_Temp	Gender_female	Gender_male
<b>0</b>	68	190	94	29	105	40.8	0	1
<b>1</b>	20	166	60	14	94	40.3	1	0
<b>2</b>	69	179	79	5	88	38.7	0	1
<b>3</b>	34	179	71	13	100	40.5	1	0
<b>4</b>	27	154	58	10	81	39.8	1	0
...	...	...	...	...	...	...	...	...
<b>14995</b>	20	193	86	11	92	40.4	1	0
<b>14996</b>	27	165	65	6	85	39.2	1	0
<b>14997</b>	43	159	58	16	90	40.1	1	0
<b>14998</b>	78	193	97	2	84	38.3	0	1
<b>14999</b>	63	173	79	18	92	40.5	0	1

15000 rows × 8 columns

```
In [15]: from sklearn.model_selection import train_test_split
```

```
In [16]: x_train,x_test,y_train,y_test=train_test_split(x,y,test_size=0.22,random_state=49)
```

```
In [17]: x_test.head(10)
```

Out[17]:

	Age	Height	Weight	Duration	Heart_Rate	Body_Temp	Gender_female	Gender_male
<b>9944</b>	55	183	83	24	106	40.7	0	1
<b>12</b>	66	171	79	11	90	40.0	0	1
<b>5052</b>	60	191	95	2	72	38.1	0	1
<b>14845</b>	26	186	81	5	72	39.3	0	1
<b>12637</b>	22	177	78	25	102	40.7	0	1
<b>9721</b>	24	183	81	12	91	40.1	0	1
<b>14141</b>	33	168	63	5	81	38.8	1	0
<b>2372</b>	77	175	79	4	84	38.8	0	1
<b>431</b>	23	173	72	28	110	40.9	1	0
<b>6357</b>	59	189	95	22	110	40.5	0	1

In [18]:

x\_train.head(10)

Out[18]:

	Age	Height	Weight	Duration	Heart_Rate	Body_Temp	Gender_female	Gender_male
<b>11805</b>	28	193	97	13	91	40.0	0	1
<b>10315</b>	22	145	43	9	77	39.9	1	0
<b>12700</b>	63	162	58	19	107	40.5	1	0
<b>1557</b>	78	170	78	14	103	40.1	0	1
<b>7973</b>	29	160	55	5	85	39.3	1	0
<b>2827</b>	38	158	63	21	103	40.6	1	0
<b>1785</b>	39	169	62	29	113	41.0	1	0
<b>13914</b>	34	175	65	5	85	38.7	1	0
<b>11389</b>	26	187	85	25	107	40.5	0	1
<b>2342</b>	25	181	70	21	108	40.2	1	0

In [19]:

y\_train.head(10)

Out[19]:

```

11805    52
10315    28
12700   131
1557     109
7973     20
2827     125
1785     204
13914     20
11389    151
2342     129
Name: Calories, dtype: int64

```

In [20]:

y\_test.head(10)

```
Out[20]: 9944    174
          12      58
          5052    5
          14845   3
          12637  124
          9721   42
          14141  18
          2372   20
          431    176
          6357   183
          Name: Calories, dtype: int64
```

```
In [21]: from sklearn.linear_model import LinearRegression
```

```
In [22]: reg=LinearRegression()
          reg.fit(x_train,y_train)
```

```
Out[22]: LinearRegression()
```

```
In [23]: ypred=reg.predict(x_test)
```

```
In [24]: ypred
```

```
Out[24]: array([161.64451313,  61.70210153, -3.32685259, ...,  64.76631124,
                88.71631251, 157.51493911])
```

```
In [25]: from sklearn.metrics import r2_score
```

```
In [26]: r2_score(y_test,ypred)
```

```
Out[26]: 0.966580087122947
```

```
In [27]: from sklearn.metrics import mean_squared_error
```

```
In [28]: mean_squared_error(ypred,y_test)
```

```
Out[28]: 128.97933287284866
```

```
In [29]: res=pd.DataFrame(columns=['Calories','predicted'])
          res['Calories']=y_test
          res['predicted']=ypred
          res=res.reset_index()
          res['ID']=res.index
```

```
In [30]: res.head(10)
```

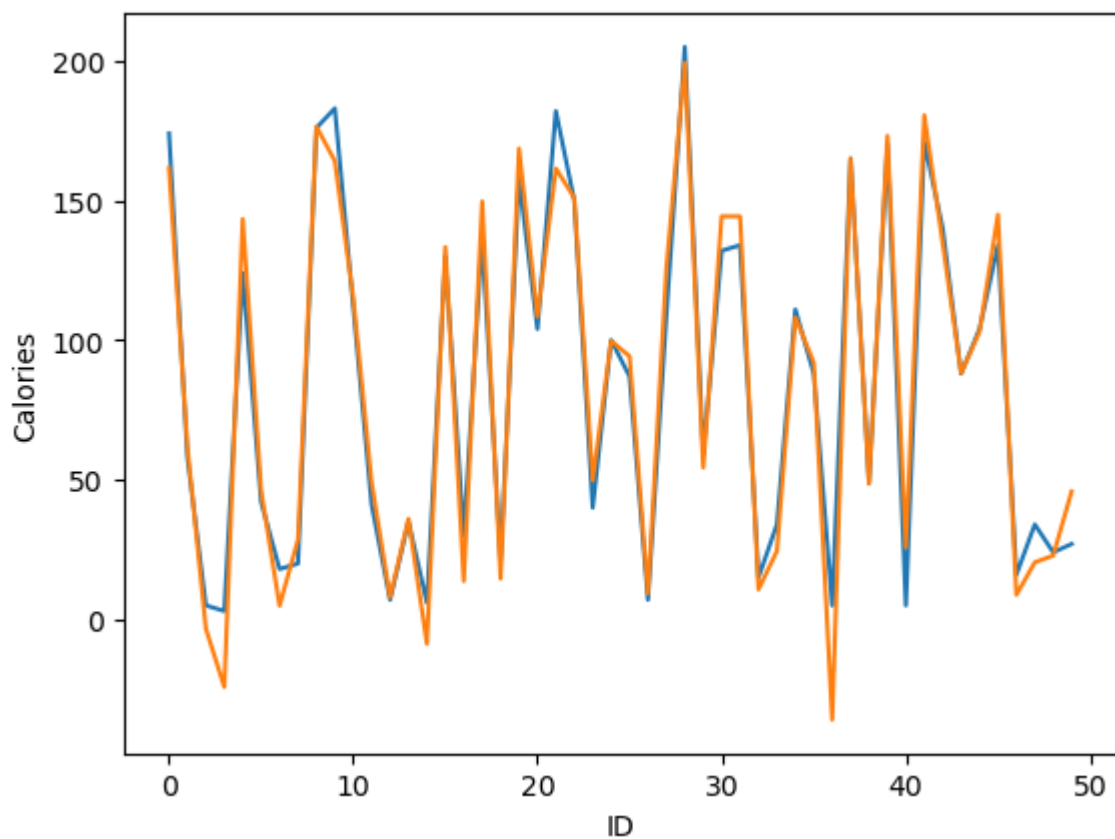
Out[30]:

	index	Calories	predicted	ID
0	9944	174	161.644513	0
1	12	58	61.702102	1
2	5052	5	-3.326853	2
3	14845	3	-24.238001	3
4	12637	124	143.377735	4
5	9721	42	46.007323	5
6	14141	18	4.872376	6
7	2372	20	28.560636	7
8	431	176	176.460214	8
9	6357	183	164.176525	9

In [31]: `import seaborn as sns`  
`import matplotlib.pyplot as plt`

In [32]: `sns.lineplot(x='ID',y='Calories',data=res.head(50))`  
`sns.lineplot(x='ID',y='predicted',data=res.head(50))`  
`plt.plot()`

Out[32]: []



In [33]: `new=[[22,175,80,25,75,43.5,1,0]]`

In [34]: `real=reg.predict(new)`  
`real`

Out[34]: array([43.57603218])

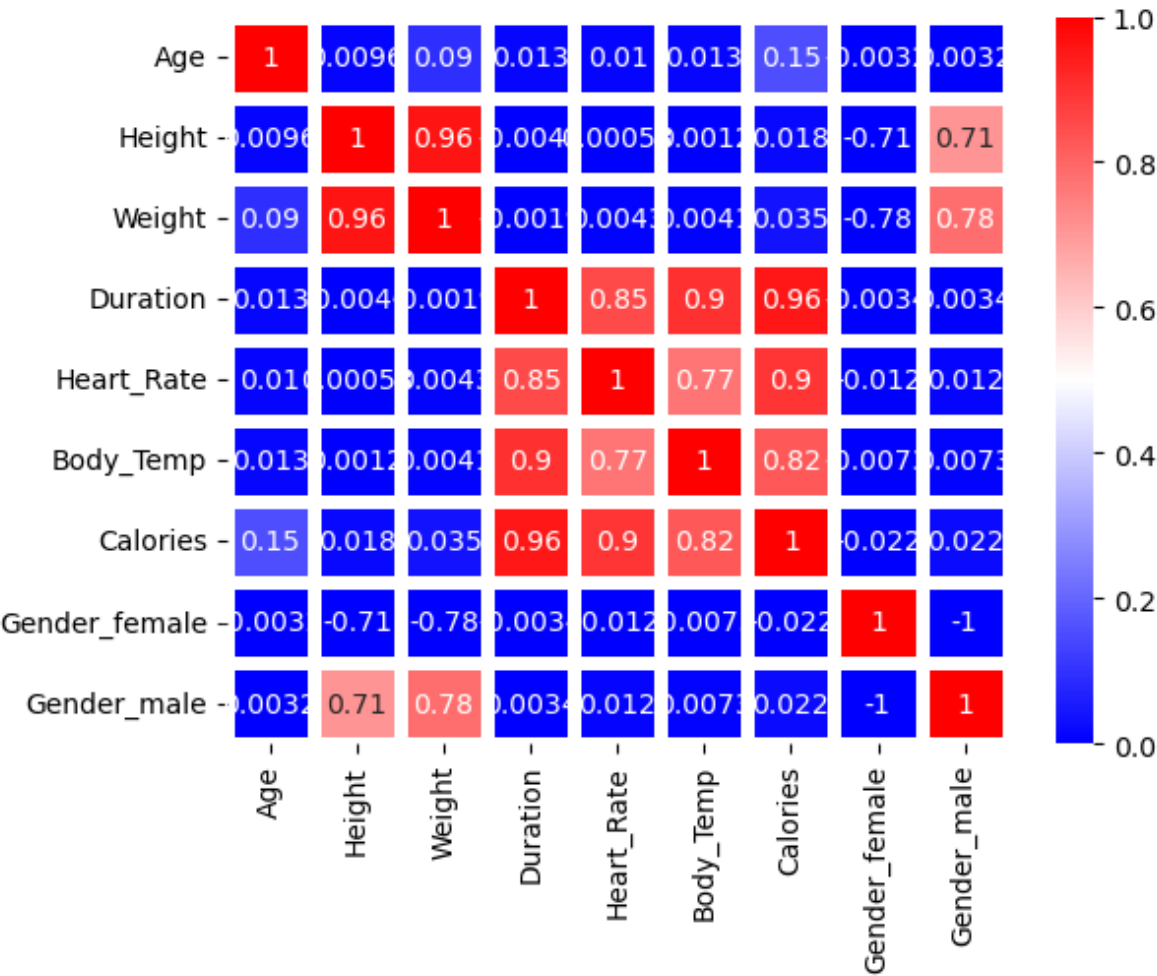
In [35]: cor\_mat=b.corr()  
cor\_mat

Out[35]:

	Age	Height	Weight	Duration	Heart_Rate	Body_Temp	Calories	Gen
Age	1.000000	0.009554	0.090094	0.013247	0.010482	0.013175	0.154395	
Height	0.009554	1.000000	0.958451	-0.004625	0.000528	0.001200	0.017537	
Weight	0.090094	0.958451	1.000000	-0.001884	0.004311	0.004095	0.035481	
Duration	0.013247	-0.004625	-0.001884	1.000000	0.852869	0.903167	0.955421	
Heart_Rate	0.010482	0.000528	0.004311	0.852869	1.000000	0.771529	0.897882	
Body_Temp	0.013175	0.001200	0.004095	0.903167	0.771529	1.000000	0.824558	
Calories	0.154395	0.017537	0.035481	0.955421	0.897882	0.824558	1.000000	
Gender_female	-0.003222	-0.710534	-0.783186	-0.003440	-0.011555	-0.007264	-0.022357	
Gender_male	0.003222	0.710534	0.783186	0.003440	0.011555	0.007264	0.022357	

In [36]: sns.heatmap(cor\_mat,vmax=1,vmin=0,annot=True,linewidth=5,cmap='bwr')

Out[36]: <AxesSubplot:>



In [ ]: