In [1]:

```
import pandas as pd
import seaborn as sn
```

In [2]:

```
df=pd.read_csv("weight-height.csv")
df.head()
```

Out[2]:

	Gender	Height	Weight
0	Male	73.847017	241.893563
1	Male	68.781904	162.310473
2	Male	74.110105	212.740856
3	Male	71.730978	220.042470
4	Male	69.881796	206.349801

In [3]:

```
df.shape
```

Out[3]:

```
(10000, 3)
```

In [6]:

```
df.Height.describe()
```

Out[6]:

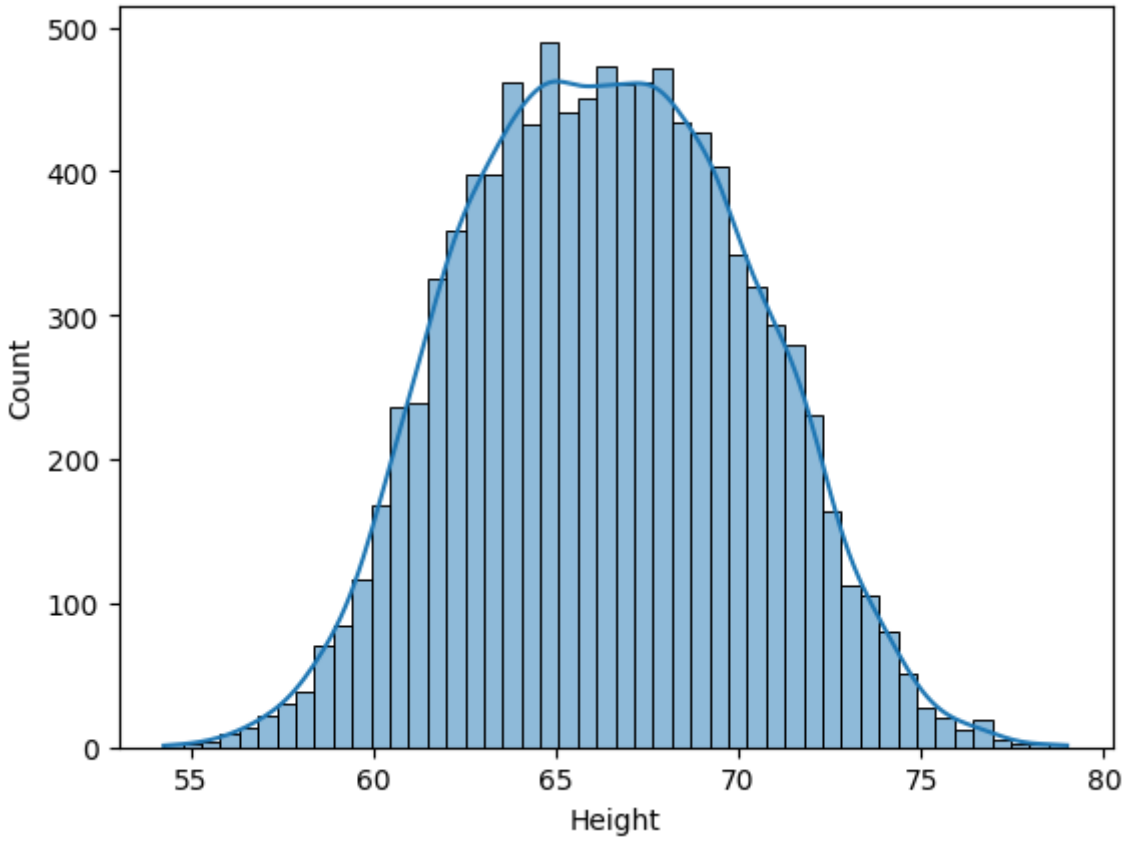
```
count    10000.000000
mean        66.367560
std          3.847528
min         54.263133
25%         63.505620
50%         66.318070
75%         69.174262
max         78.998742
Name: Height, dtype: float64
```

In [7]:

```
sn.histplot(df.Height, kde=True)
```

Out[7]:

```
<AxesSubplot: xlabel='Height', ylabel='Count'>
```



In [9]:

```
mean=df.Height.mean()
mean
```

Out[9]:

```
66.36755975482124
```

In [10]:

```
std=df.Height.std()
std
```

Out[10]:

```
3.8475281207732293
```

In [11]:

```
mean - 3*std
```

Out[11]:

```
54.82497539250156
```

In [12]:

```
mean + 3*std
```

Out[12]:

```
77.91014411714093
```

In [15]:

```
df[(df.Height<54.82497539250156)|(df.Height>77.91014411714093)]
```

Out[15]:

	Gender	Height	Weight
994	Male	78.095867	255.690835
1317	Male	78.462053	227.342565
2014	Male	78.998742	269.989699
3285	Male	78.528210	253.889004
3757	Male	78.621374	245.733783
6624	Female	54.616858	71.393749
9285	Female	54.263133	64.700127

In [20]:

```
df_with_no_outlier=df[(df.Height>54.82497539250156)&(df.Height<77.91014411714093)]
df_with_no_outlier
```

Out[20]:

	Gender	Height	Weight
0	Male	73.847017	241.893563
1	Male	68.781904	162.310473
2	Male	74.110105	212.740856
3	Male	71.730978	220.042470
4	Male	69.881796	206.349801
...	...	...	...
9995	Female	66.172652	136.777454
9996	Female	67.067155	170.867906
9997	Female	63.867992	128.475319
9998	Female	69.034243	163.852461
9999	Female	61.944246	113.649103

9993 rows × 3 columns

In [21]:

```
df_with_no_outlier.shape
```

Out[21]:

```
(9993, 3)
```

In [ ]:

```

```