Heart Diesease

The Gender which is suffering with high rate of HeartDiseses.

In [1]:
```
import numpy as np
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt
```

In [2]:
```
df=pd.read_csv(r"C:\Users\yoshitha lakshmi\OneDrive\Desktop\python\heart_2020_cleaned.csv")
df
```

Out[2]:

	HeartDisease	BMI	Smoking	AlcoholDrinking	Stroke	PhysicalHealth	MentalHealth	DiffWalking	Sex	AgeCategory	Race	Diabe
0	No	16.60	Yes	No	No	3.0	30.0	No	Female	55-59	White	Y
1	No	20.34	No	No	Yes	0.0	0.0	No	Female	80 or older	White	I
2	No	26.58	Yes	No	No	20.0	30.0	No	Male	65-69	White	Y
3	No	24.21	No	No	No	0.0	0.0	No	Female	75-79	White	I
4	No	23.71	No	No	No	28.0	0.0	Yes	Female	40-44	White	I
...	...	...	...	...	...	...	...	...	...	...	...	
319790	Yes	27.41	Yes	No	No	7.0	0.0	Yes	Male	60-64	Hispanic	Y
319791	No	29.84	Yes	No	No	0.0	0.0	No	Male	35-39	Hispanic	I
319792	No	24.24	No	No	No	0.0	0.0	No	Female	45-49	Hispanic	I
319793	No	32.81	No	No	No	0.0	0.0	No	Female	25-29	Hispanic	I
319794	No	46.56	No	No	No	0.0	0.0	No	Female	80 or older	Hispanic	I

319795 rows × 18 columns

In [3]:
```
1 df.head()
```

Out[3]:

	HeartDisease	BMI	Smoking	AlcoholDrinking	Stroke	PhysicalHealth	MentalHealth	DiffWalking	Sex	AgeCategory	Race	Diabetic	Phy
0	No	16.60	Yes	No	No	3.0	30.0	No	Female	55-59	White	Yes	
1	No	20.34	No	No	Yes	0.0	0.0	No	Female	80 or older	White	No	
2	No	26.58	Yes	No	No	20.0	30.0	No	Male	65-69	White	Yes	
3	No	24.21	No	No	No	0.0	0.0	No	Female	75-79	White	No	
4	No	23.71	No	No	No	28.0	0.0	Yes	Female	40-44	White	No	

In [4]:
```
1 df.tail()
```

Out[4]:

	HeartDisease	BMI	Smoking	AlcoholDrinking	Stroke	PhysicalHealth	MentalHealth	DiffWalking	Sex	AgeCategory	Race	Diabe
319790	Yes	27.41	Yes	No	No	7.0	0.0	Yes	Male	60-64	Hispanic	Y
319791	No	29.84	Yes	No	No	0.0	0.0	No	Male	35-39	Hispanic	
319792	No	24.24	No	No	No	0.0	0.0	No	Female	45-49	Hispanic	
319793	No	32.81	No	No	No	0.0	0.0	No	Female	25-29	Hispanic	
319794	No	46.56	No	No	No	0.0	0.0	No	Female	80 or older	Hispanic	

In [5]:
```
1 df.shape
```

Out[5]: (319795, 18)

In [6]:
```
1 df.describe()
```

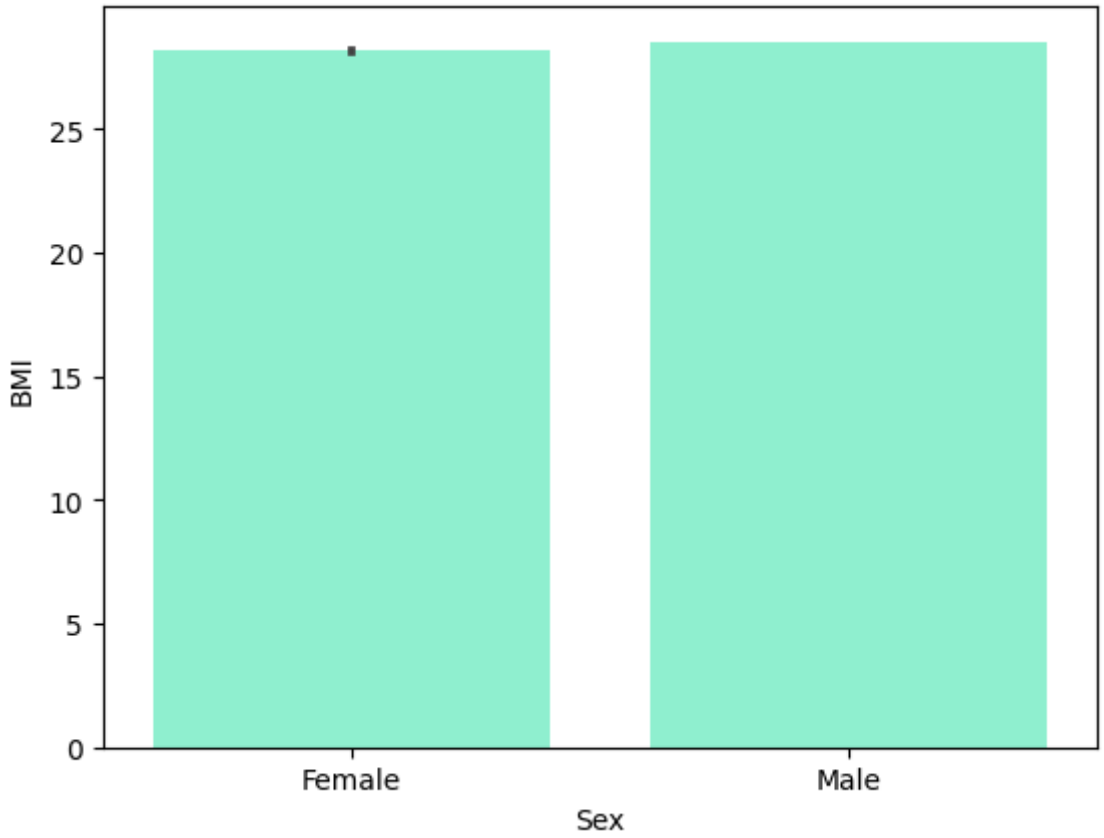
Out[6]:

	BMI	PhysicalHealth	MentalHealth	SleepTime
count	319795.000000	319795.00000	319795.000000	319795.000000
mean	28.325399	3.37171	3.898366	7.097075
std	6.356100	7.95085	7.955235	1.436007
min	12.020000	0.00000	0.000000	1.000000
25%	24.030000	0.00000	0.000000	6.000000
50%	27.340000	0.00000	0.000000	7.000000
75%	31.420000	2.00000	3.000000	8.000000
max	94.850000	30.00000	30.000000	24.000000

In [7]:
```
1 df.info()
```

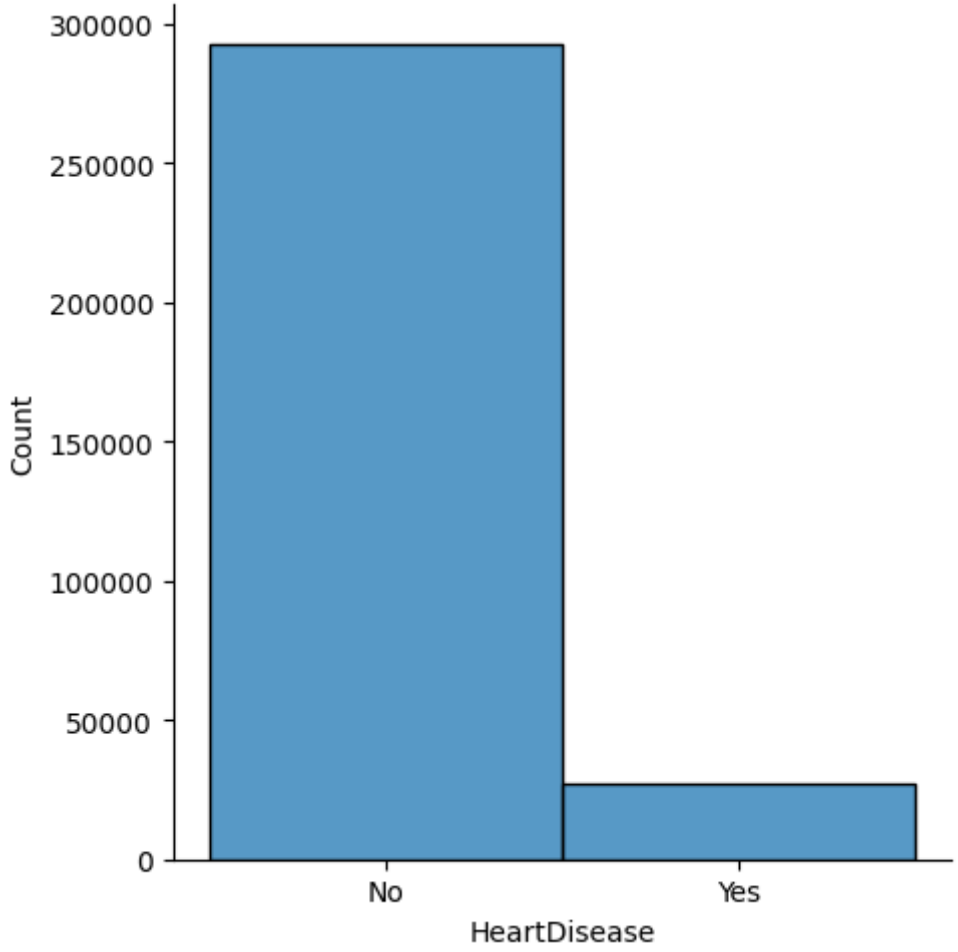
```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 319795 entries, 0 to 319794
Data columns (total 18 columns):
 #   Column            Non-Null Count   Dtype  
---  ------            --------------   -----  
 0   HeartDisease      319795 non-null  object 
 1   BMI               319795 non-null  float64
 2   Smoking           319795 non-null  object 
 3   AlcoholDrinking   319795 non-null  object 
 4   Stroke            319795 non-null  object 
 5   PhysicalHealth    319795 non-null  float64
 6   MentalHealth      319795 non-null  float64
 7   DiffWalking       319795 non-null  object 
 8   Sex               319795 non-null  object 
 9   AgeCategory       319795 non-null  object 
 10  Race              319795 non-null  object 
 11  Diabetic          319795 non-null  object 
 12  PhysicalActivity  319795 non-null  object 
 13  GenHealth         319795 non-null  object 
 14  SleepTime         319795 non-null  float64
 15  Asthma            319795 non-null  object 
 16  KidneyDisease     319795 non-null  object 
 17  SkinCancer        319795 non-null  object 
dtypes: float64(4), object(14)
memory usage: 43.9+ MB
```

In [8]:
```
1 sns.barplot(x='Sex',y='BMI',data=df,color='aquamarine')
2 plt.show()
```



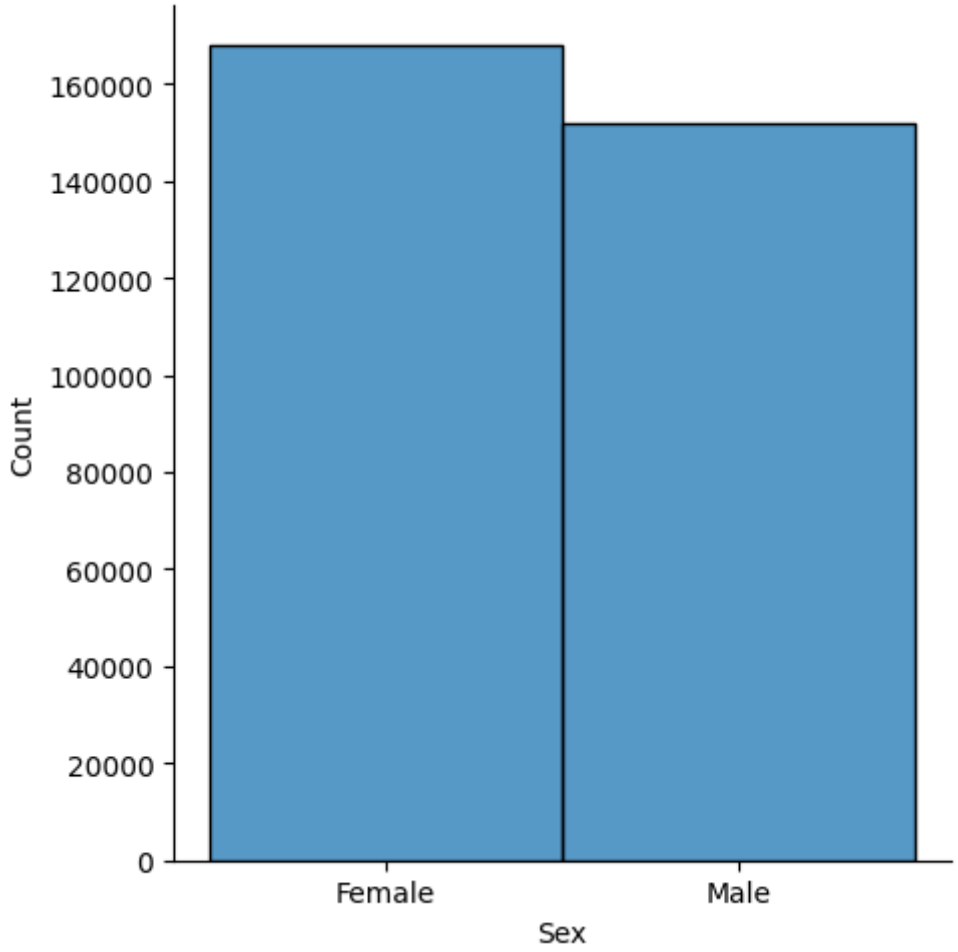
In [9]:
```
1 sns.displot(df['HeartDisease'])
2
```

Out[9]: <seaborn.axisgrid.FacetGrid at 0x158781d6da0>



In [10]:
```
1 sns.displot(df['Sex'])
```

Out[10]: <seaborn.axisgrid.FacetGrid at 0x15877d47d00>

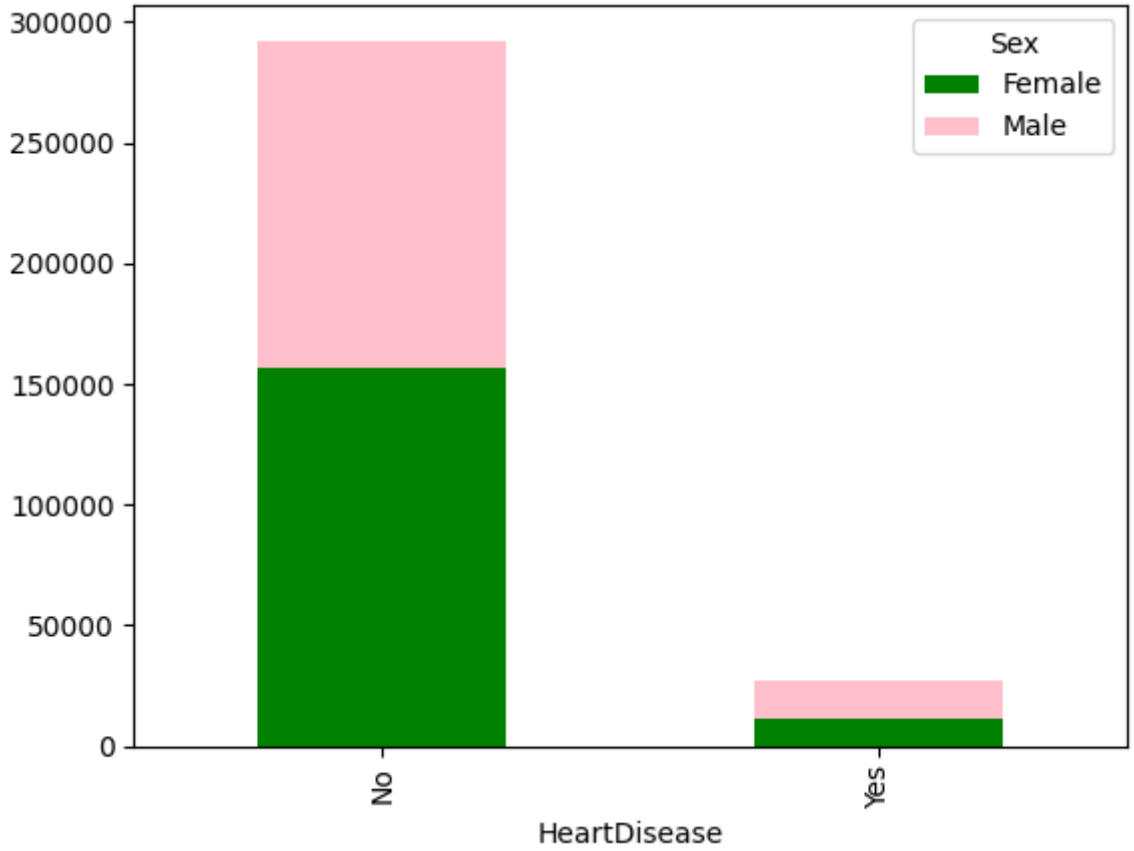


In [11]:
```
x=pd.crosstab(df['HeartDisease'],df['Sex'])
print(x)
```

```
Sex           Female    Male
HeartDisease
No            156571  135851
Yes            11234   16139
```

In [13]:
```
x.plot(kind='bar',stacked=True,color=['green','pink'],grid=False)
```

Out[13]: <Axes: xlabel='HeartDisease'>



conclusion

The males has high chances of getting Heart Diseases.

```
1
```