

In [2]:

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
import warnings
warnings.filterwarnings('ignore')
```

In [4]:

```
df=pd.read_csv('Downloads/Heart_Disease_Prediction.csv')
```

In [5]:

```
df.head()
```

Out[5]:

|   | Age | Sex | Chest<br>pain<br>type | BP  | Cholesterol | FBS<br>over<br>120 | EKG<br>results | Max<br>HR | Exercise<br>angina | ST<br>depression | Slope<br>of ST | Number<br>of<br>vessels<br>fluro |
|---|-----|-----|-----------------------|-----|-------------|--------------------|----------------|-----------|--------------------|------------------|----------------|----------------------------------|
| 0 | 70  | 1   | 4                     | 130 | 322         | 0                  | 2              | 109       | 0                  | 2.4              | 2              |                                  |
| 1 | 67  | 0   | 3                     | 115 | 564         | 0                  | 2              | 160       | 0                  | 1.6              | 2              |                                  |
| 2 | 57  | 1   | 2                     | 124 | 261         | 0                  | 0              | 141       | 0                  | 0.3              | 1              |                                  |
| 3 | 64  | 1   | 4                     | 128 | 263         | 0                  | 0              | 105       | 1                  | 0.2              | 2              |                                  |
| 4 | 74  | 0   | 2                     | 120 | 269         | 0                  | 2              | 121       | 1                  | 0.2              | 1              |                                  |

In [6]:

```
df.isnull().sum()
```

Out[6]:

```
Age          0
Sex          0
Chest pain type  0
BP           0
Cholesterol  0
FBS over 120  0
EKG results  0
Max HR       0
Exercise angina  0
ST depression  0
Slope of ST  0
Number of vessels fluro  0
Thallium     0
Heart Disease  0
dtype: int64
```

In [7]:

```
print(df.info())
```

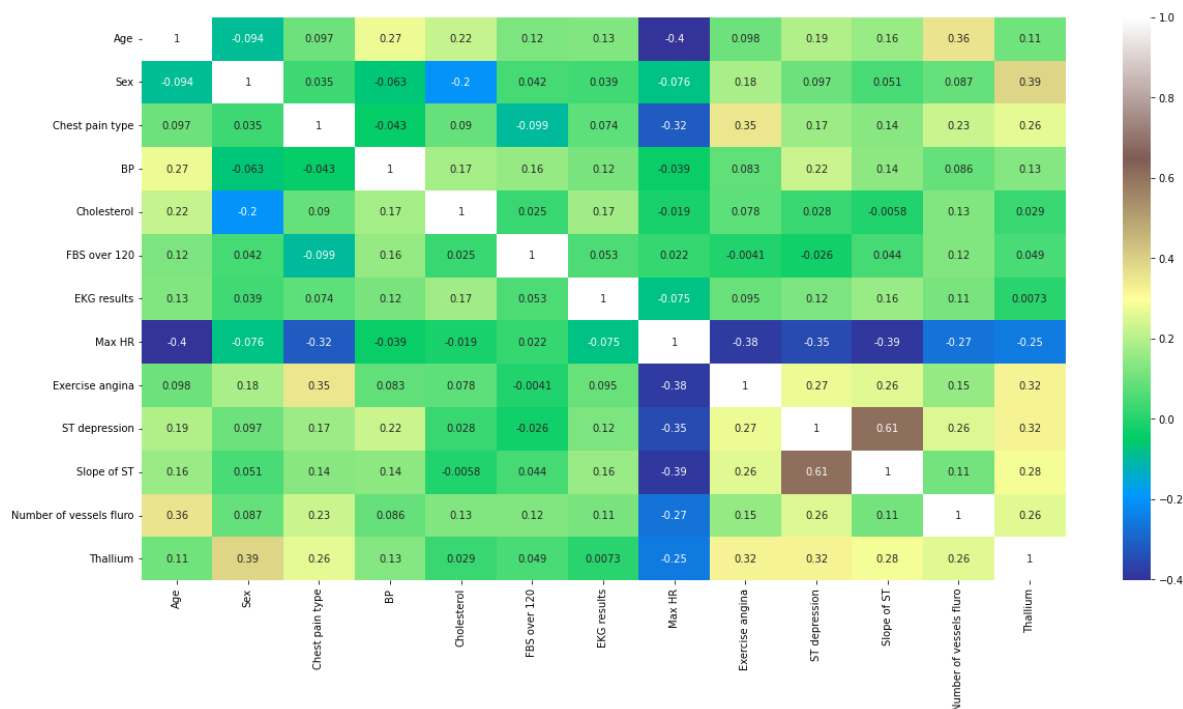
```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 270 entries, 0 to 269
Data columns (total 14 columns):
#   Column                                Non-Null Count  Dtype
---  ---                                -
0   Age                                    270 non-null    int64
1   Sex                                    270 non-null    int64
2   Chest pain type                        270 non-null    int64
3   BP                                     270 non-null    int64
4   Cholesterol                           270 non-null    int64
5   FBS over 120                          270 non-null    int64
6   EKG results                           270 non-null    int64
7   Max HR                                270 non-null    int64
8   Exercise angina                        270 non-null    int64
9   ST depression                          270 non-null    float64
10  Slope of ST                            270 non-null    int64
11  Number of vessels fluro                270 non-null    int64
12  Thallium                               270 non-null    int64
13  Heart Disease                          270 non-null    object
dtypes: float64(1), int64(12), object(1)
memory usage: 29.7+ KB
None
```

In [9]:

```
plt.figure(figsize=(20,10))
sns.heatmap(df.corr(), annot=True, cmap='terrain')
```

Out[9]:

&lt;AxesSubplot:&gt;

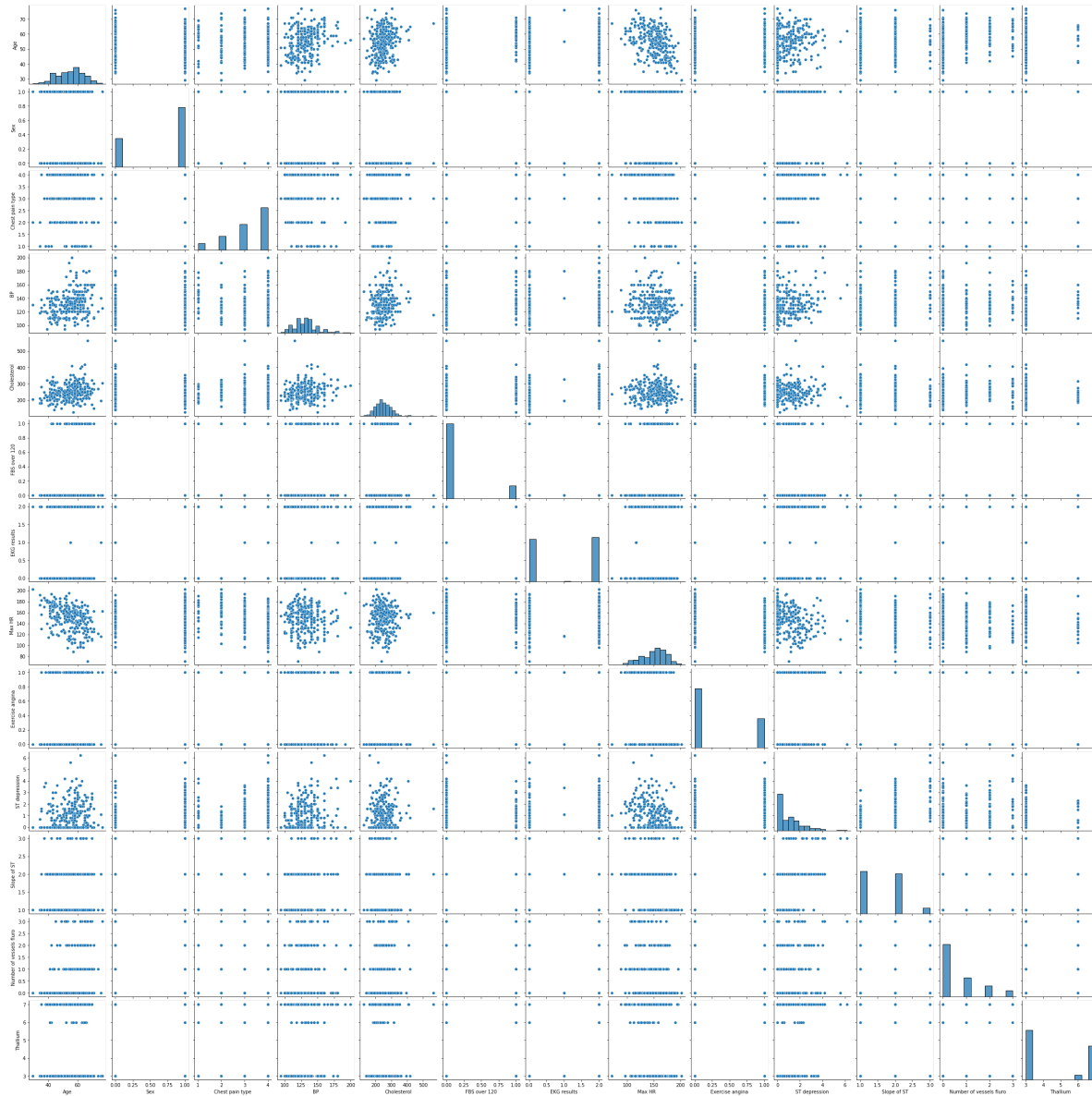


In [10]:

```
sns.pairplot(data=df)
```

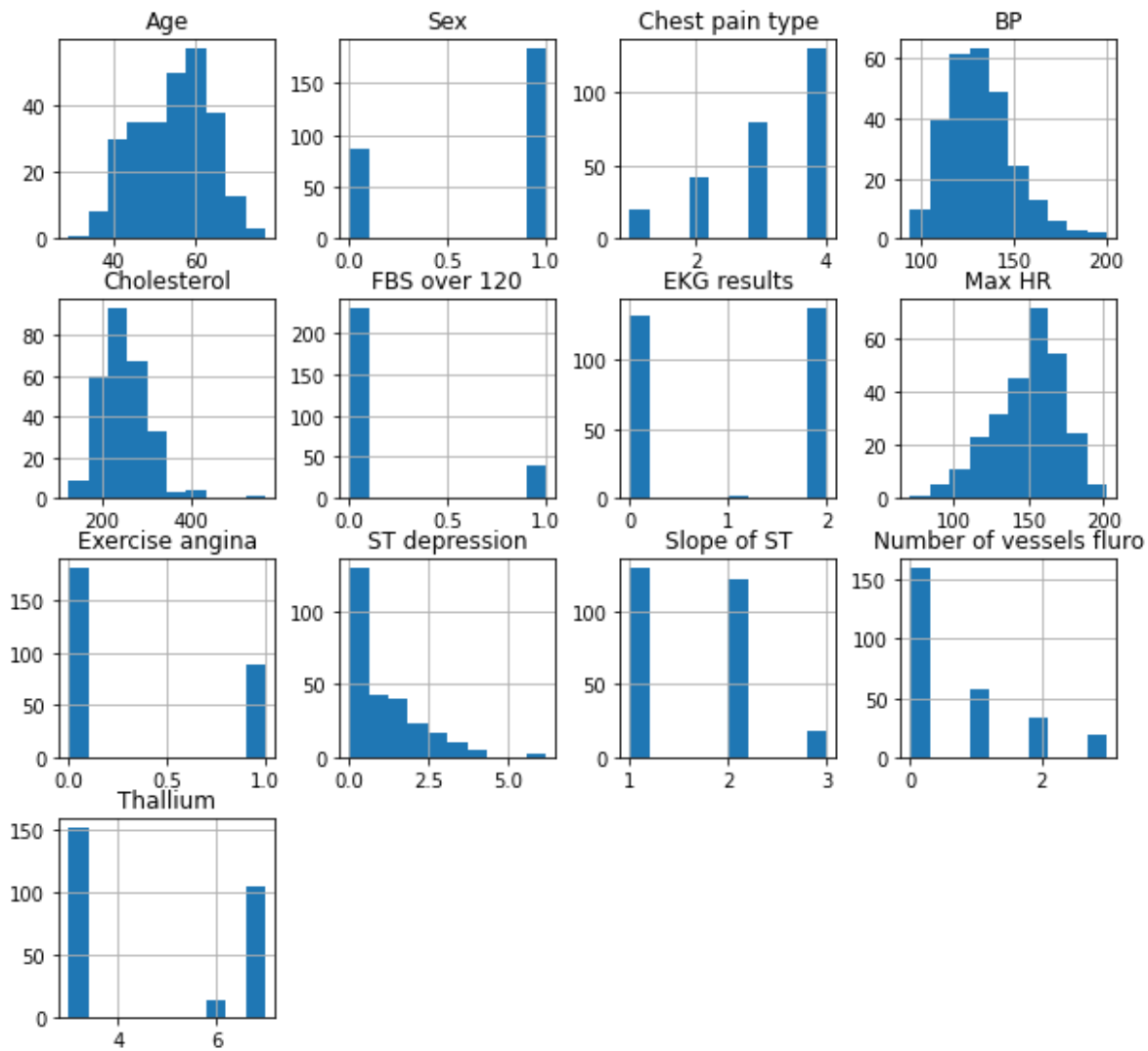
Out[10]:

```
<seaborn.axisgrid.PairGrid at 0x2059aec2448>
```



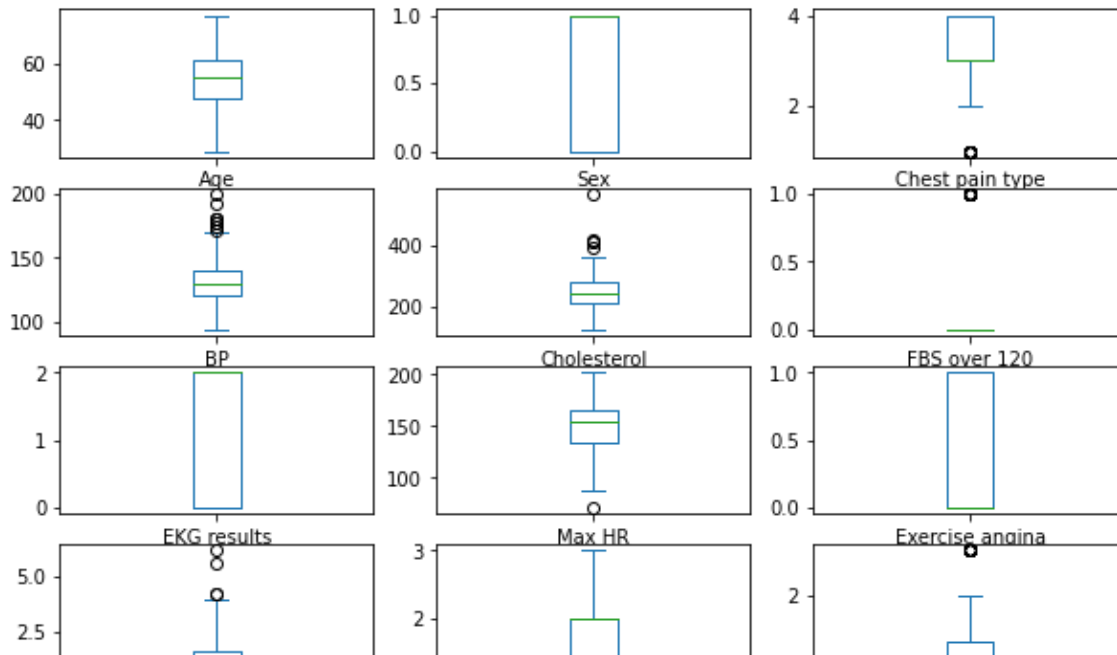
In [11]:

```
df.hist(figsize=(10,12), layout=(5,4));
```



In [13]:

```
df.plot(kind='box', subplots=True, layout=(6,3), figsize=(10,10))
plt.show()
```

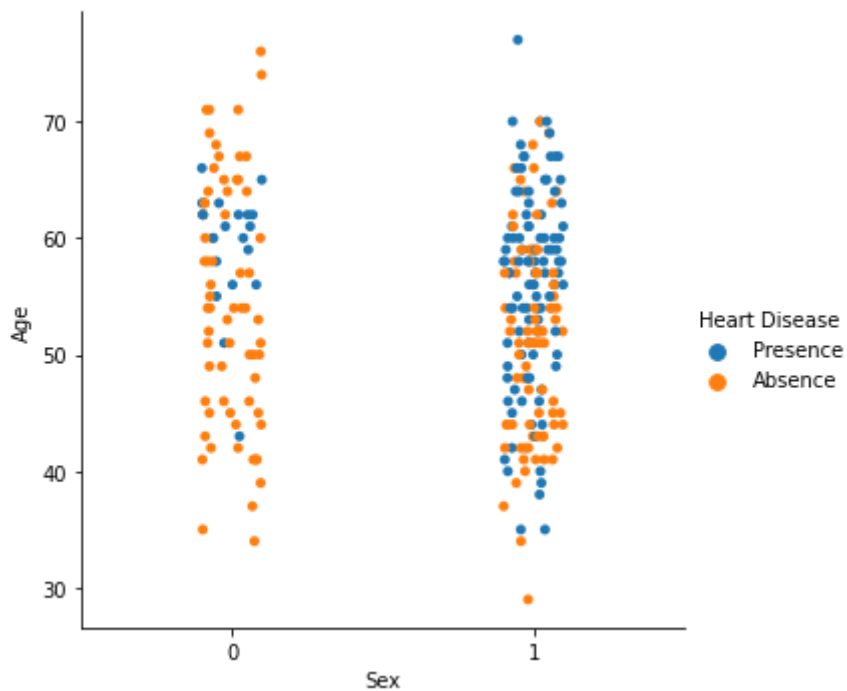


In [19]:

```
sns.catplot(data=df, x='Sex', y='Age', hue='Heart Disease', palette='tab10')
```

Out[19]:

```
<seaborn.axisgrid.FacetGrid at 0x205a367dcc8>
```

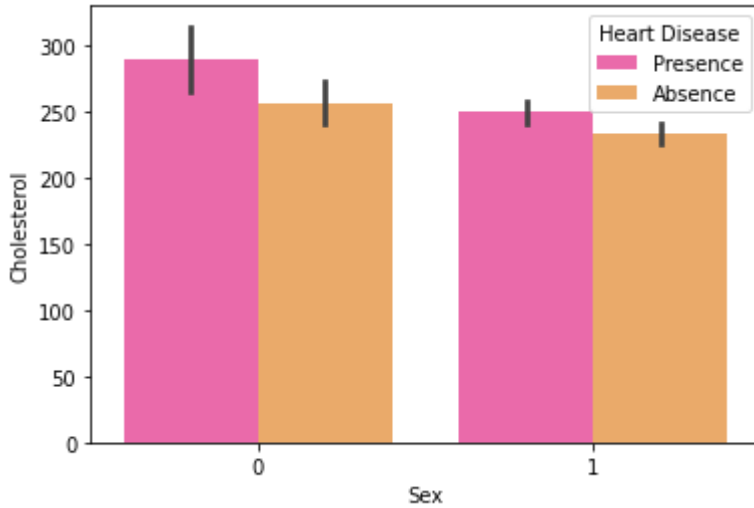


In [20]:

```
sns.barplot(data=df, x='Sex', y='Cholesterol', hue='Heart Disease', palette='spring')
```

Out[20]:

```
<AxesSubplot:xlabel='Sex', ylabel='Cholesterol'>
```



In [21]:

```
df['Sex'].value_counts()
```

Out[21]:

```
1    183
0     87
Name: Sex, dtype: int64
```

In [22]:

```
df['Chest pain type'].value_counts()
```

Out[22]:

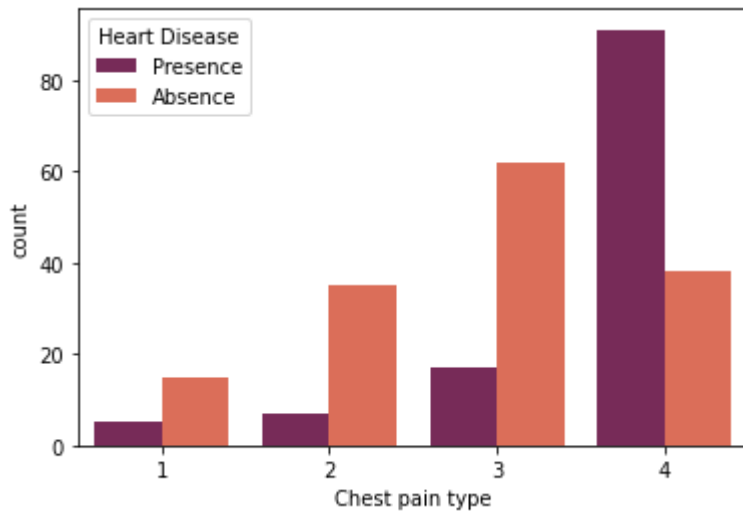
```
4    129
3     79
2     42
1     20
Name: Chest pain type, dtype: int64
```

In [23]:

```
sns.countplot(x='Chest pain type', hue='Heart Disease' , data=df, palette='rocket')
```

Out[23]:

```
<AxesSubplot:xlabel='Chest pain type', ylabel='count'>
```



In [24]:

```
gen = pd.crosstab(df['Sex'], df['Heart Disease'])
print(gen)
```

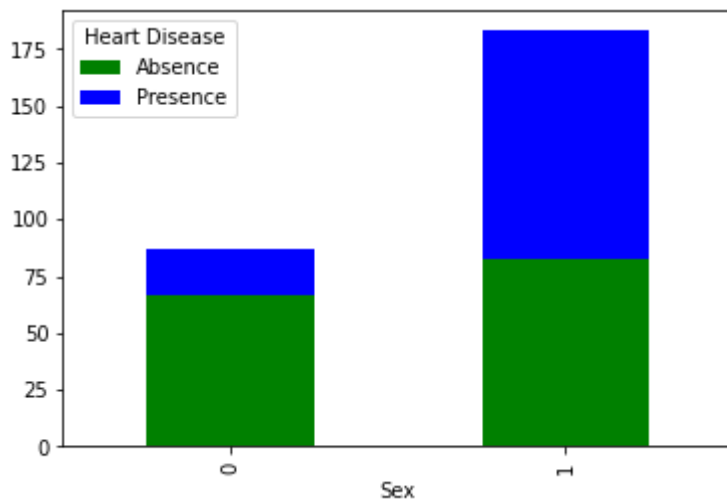
| Heart Disease | Absence | Presence |
|---------------|---------|----------|
| Sex           |         |          |
| 0             | 67      | 20       |
| 1             | 83      | 100      |

In [25]:

```
gen.plot(kind='bar', stacked='True', color=['green','blue'],grid=False)
```

Out[25]:

<AxesSubplot:xlabel='Sex'>



In [ ]:

In [ ]: