

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

```
In [2]: filepath=(r"C:\Users\Jhuvvi Sree\OneDrive\Desktop\Book1.csv")
data=pd.read_csv(filepath)
print(data)
```

	s.no	gender	age	hypertension	heart_disease	obesity	stress	\
0	1	female	58	0	0	1	1	
1	2	male	35	1	0	0	0	
2	3	male	25	0	0	0	1	
3	4	female	78	0	1	1	1	
4	5	male	89	0	0	1	0	
5	6	male	55	1	1	0	1	
6	7	male	77	1	0	1	0	
7	8	female	18	0	0	1	1	
8	9	female	45	0	0	0	0	
9	10	female	75	0	1	0	1	
10	11	male	63	0	1	1	0	
11	12	male	65	1	0	1	1	
12	13	female	29	1	0	0	1	
13	14	male	74	1	1	1	0	
14	15	female	73	0	1	0	1	
15	16	male	71	1	0	1	0	
16	17	male	50	0	0	0	1	
17	18	female	29	1	0	0	0	
18	19	female	45	0	0	1	1	
19	20	female	56	1	0	0	1	

	smoking_history	bmi	b.p level	diabetes
0	never	20.50	140	0
1	current	25.60	125	0
2	current	32.50	160	1
3	no info	28.60	180	1
4	never	46.00	140	0
5	never	26.50	120	0
6	current	46.20	80	1
7	current	15.30	150	0
8	never	28.60	110	1
9	current	14.50	110	1
10	never	16.40	120	0
11	current	25.10	125	1
12	no info	27.35	130	0
13	no info	28.00	140	0
14	current	23.50	180	0
15	never	20.30	170	1
16	never	23.50	140	0
17	no info	22.00	180	1
18	no info	19.20	100	0
19	current	15.20	80	1

```
In [4]: data.shape
```

```
Out[4]: (20, 11)
```

```
In [5]: data.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 20 entries, 0 to 19
Data columns (total 11 columns):
#   Column                Non-Null Count  Dtype
---  -
0   s.no                  20 non-null    int64
1   gender                20 non-null    object
2   age                  20 non-null    int64
3   hypertension          20 non-null    int64
4   heart_disease         20 non-null    int64
5   obesity               20 non-null    int64
6   stress               20 non-null    int64
7   smoking_history       20 non-null    object
8   bmi                  20 non-null    float64
9   b.p level            20 non-null    int64
10  diabetes              20 non-null    int64
dtypes: float64(1), int64(8), object(2)
memory usage: 1.8+ KB
```

```
In [6]: data.head()
```

```
Out[6]:
```

	s.no	gender	age	hypertension	heart_disease	obesity	stress	smoking_history	bmi	b.p level	diabetes	
0	1	female	58		0	0	1	1	never	20.5	140	0
1	2	male	35		1	0	0	0	current	25.6	125	0
2	3	male	25		0	0	0	1	current	32.5	160	1
3	4	female	78		0	1	1	1	no info	28.6	180	1
4	5	male	89		0	0	1	0	never	46.0	140	0

```
In [7]: data.tail()
```

```
Out[7]:
```

	s.no	gender	age	hypertension	heart_disease	obesity	stress	smoking_history	bmi	b.p level	diabetes	
15	16	male	71		1	0	1	0	never	20.3	170	1
16	17	male	50		0	0	0	1	never	23.5	140	0
17	18	female	29		1	0	0	0	no info	22.0	180	1
18	19	female	45		0	0	1	1	no info	19.2	100	0
19	20	female	56		1	0	0	1	current	15.2	80	1

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

```
Out[8]: s.no      0
gender    0
age       0
hypertension  0
heart_disease  0
obesity    0
stress     0
smoking_history  0
bmi        0
b.p level  0
diabetes   0
dtype: int64
```

```
In [9]: data.columns
```

```
Out[9]: Index(['s.no', 'gender', 'age', 'hypertension', 'heart_disease', 'obesity',
              'stress', 'smoking_history', 'bmi', 'b.p level', 'diabetes'],
              dtype='object')
```

```
In [10]: data.size
```

```
Out[10]: 220
```

```
In [11]: data
```

Out[11]:

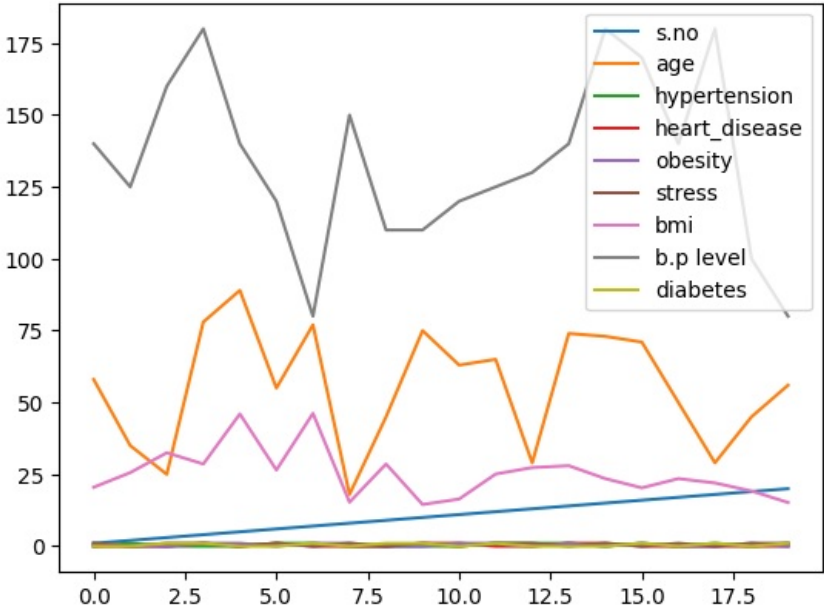
	s.no	gender	age	hypertension	heart_disease	obesity	stress	smoking_history	bmi	b.p level	diabetes
0	1	female	58	0	0	1	1	never	20.50	140	0
1	2	male	35	1	0	0	0	current	25.60	125	0
2	3	male	25	0	0	0	1	current	32.50	160	1
3	4	female	78	0	1	1	1	no info	28.60	180	1
4	5	male	89	0	0	1	0	never	46.00	140	0
5	6	male	55	1	1	0	1	never	26.50	120	0
6	7	male	77	1	0	1	0	current	46.20	80	1
7	8	female	18	0	0	1	1	current	15.30	150	0
8	9	female	45	0	0	0	0	never	28.60	110	1
9	10	female	75	0	1	0	1	current	14.50	110	1
10	11	male	63	0	1	1	0	never	16.40	120	0
11	12	male	65	1	0	1	1	current	25.10	125	1
12	13	female	29	1	0	0	1	no info	27.35	130	0
13	14	male	74	1	1	1	0	no info	28.00	140	0
14	15	female	73	0	1	0	1	current	23.50	180	0
15	16	male	71	1	0	1	0	never	20.30	170	1
16	17	male	50	0	0	0	1	never	23.50	140	0
17	18	female	29	1	0	0	0	no info	22.00	180	1
18	19	female	45	0	0	1	1	no info	19.20	100	0
19	20	female	56	1	0	0	1	current	15.20	80	1

In [12]: data.index

Out[12]: RangeIndex(start=0, stop=20, step=1)

In [13]: data.plot()

Out[13]: <Axes: >



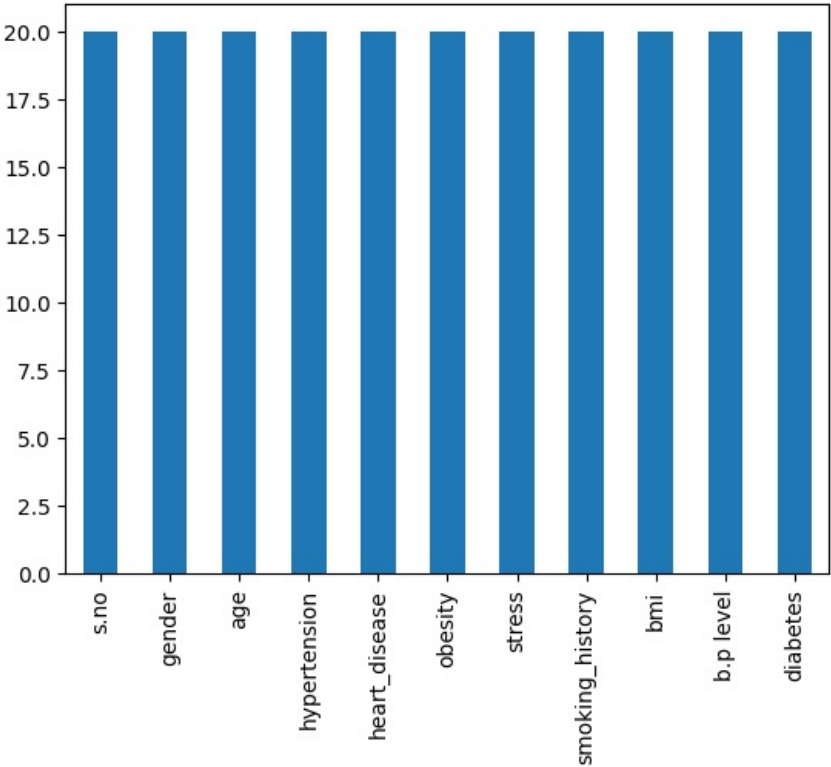
In [48]: data.sample(10)

Out[48]:

	s.no	gender	age	hypertension	heart_disease	obesity	stress	smoking_history	bmi	b.p level	diabetes
15	16	male	71	1	0	1	0	never	20.3	170	1
16	17	male	50	0	0	0	1	never	23.5	140	0
3	4	female	78	0	1	1	1	no info	28.6	180	1
2	3	male	25	0	0	0	1	current	32.5	160	1
1	2	male	35	1	0	0	0	current	25.6	125	0
7	8	female	18	0	0	1	1	current	15.3	150	0
0	1	female	58	0	0	1	1	never	20.5	140	0
10	11	male	63	0	1	1	0	never	16.4	120	0
6	7	male	77	1	0	1	0	current	46.2	80	1
8	9	female	45	0	0	0	0	never	28.6	110	1

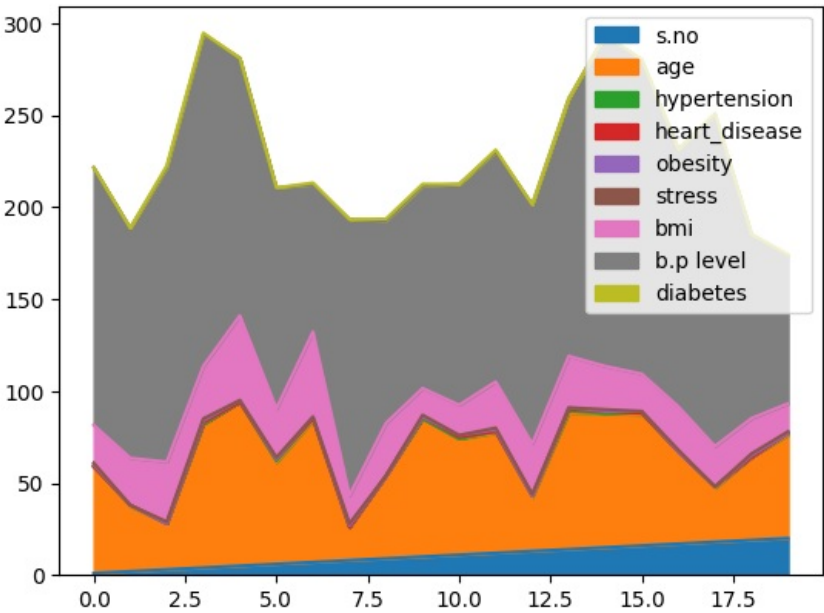
In [14]: data.count().plot.bar()#

Out[14]: <Axes: >



In [15]: data.plot.area()

Out[15]: <Axes: >



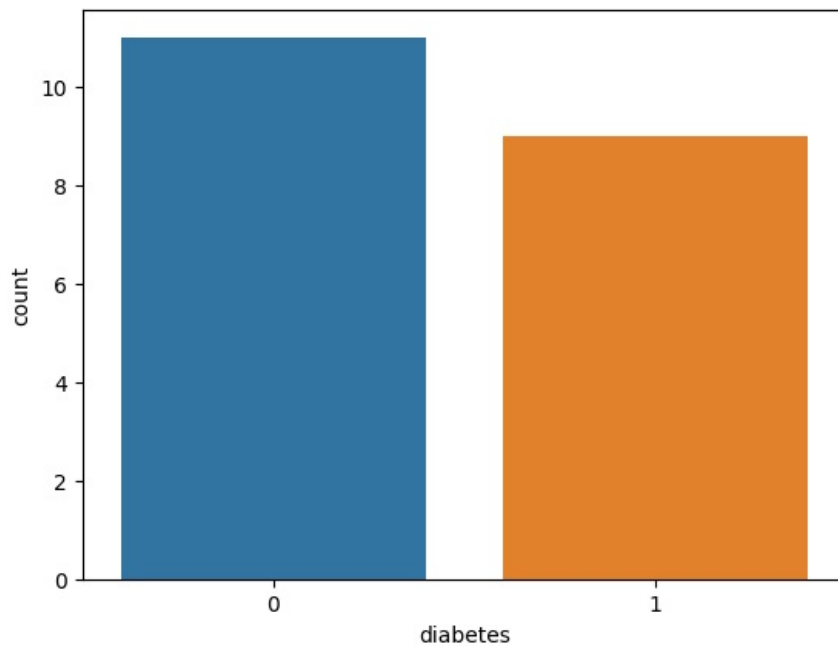
In [40]: data.info()

```
data.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 20 entries, 0 to 19
Data columns (total 11 columns):
#   Column                Non-Null Count  Dtype  
---  -
0   s.no                   20 non-null    int64  
1   gender                 20 non-null    object  
2   age                   20 non-null    int64  
3   hypertension           20 non-null    int64  
4   heart_disease          20 non-null    int64  
5   obesity                20 non-null    int64  
6   stress                 20 non-null    int64  
7   smoking_history        20 non-null    object  
8   bmi                    20 non-null    float64 
9   b.p level              20 non-null    int64  
10  diabetes                20 non-null    int64  
dtypes: float64(1), int64(8), object(2)
memory usage: 1.8+ KB
```

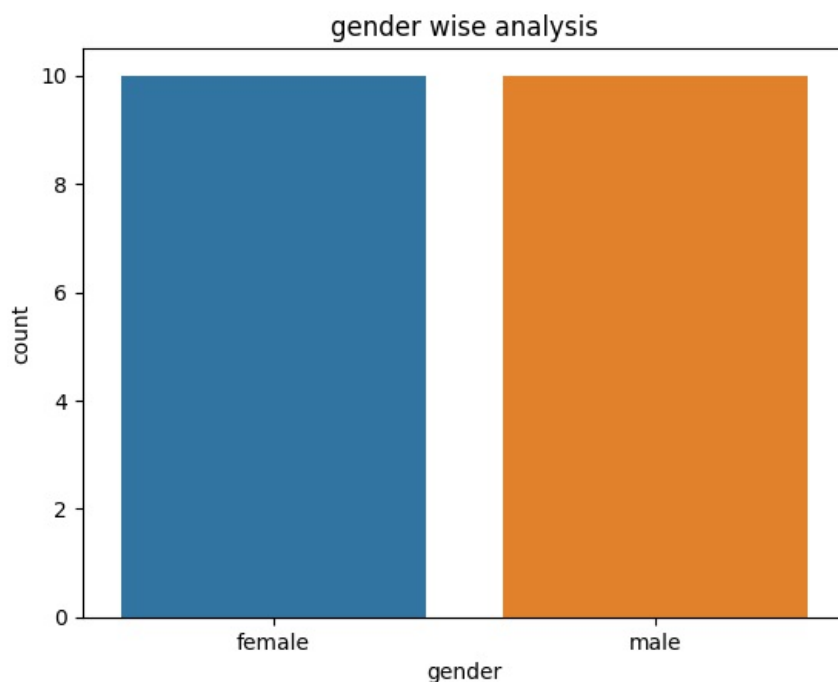
```
In [17]: sns.countplot(x='diabetes',data=data)
```

```
Out[17]: <Axes: xlabel='diabetes', ylabel='count'>
```



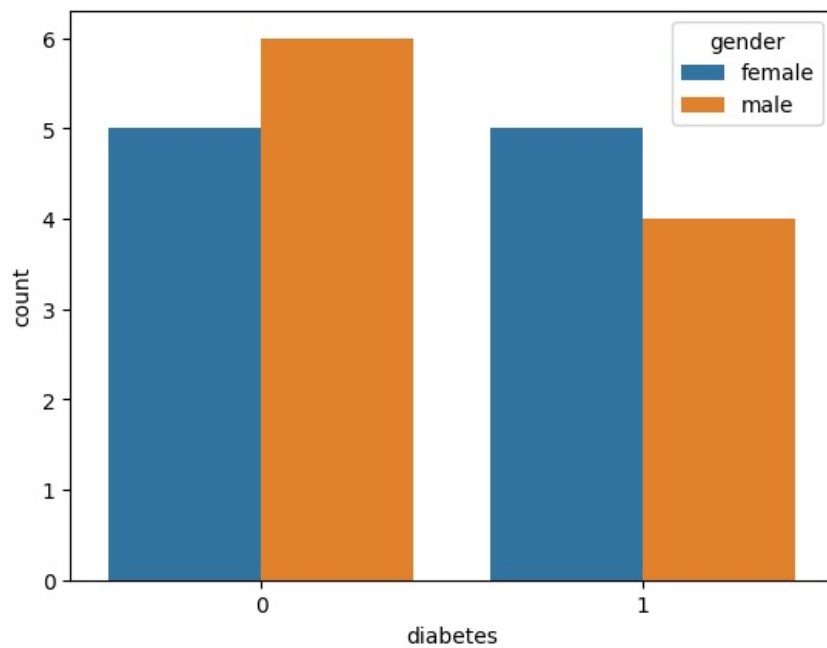
```
In [18]: plt.title('gender wise analysis')
sns.countplot(x='gender',data=data)
```

```
Out[18]: <Axes: title={'center': 'gender wise analysis'}, xlabel='gender', ylabel='count'>
```



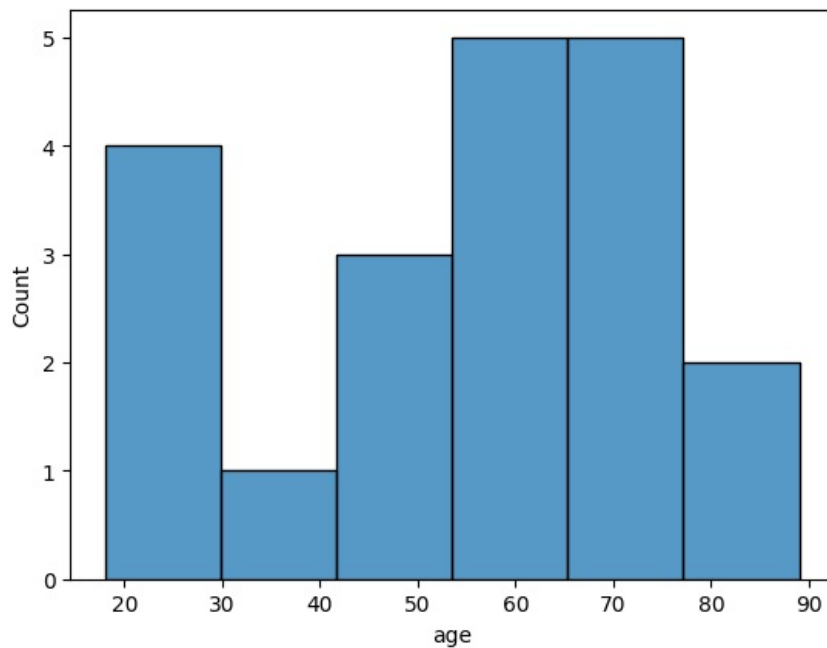
```
In [19]: sns.countplot(x='diabetes',data=data,hue='gender')
```

```
Out[19]: <Axes: xlabel='diabetes', ylabel='count'>
```



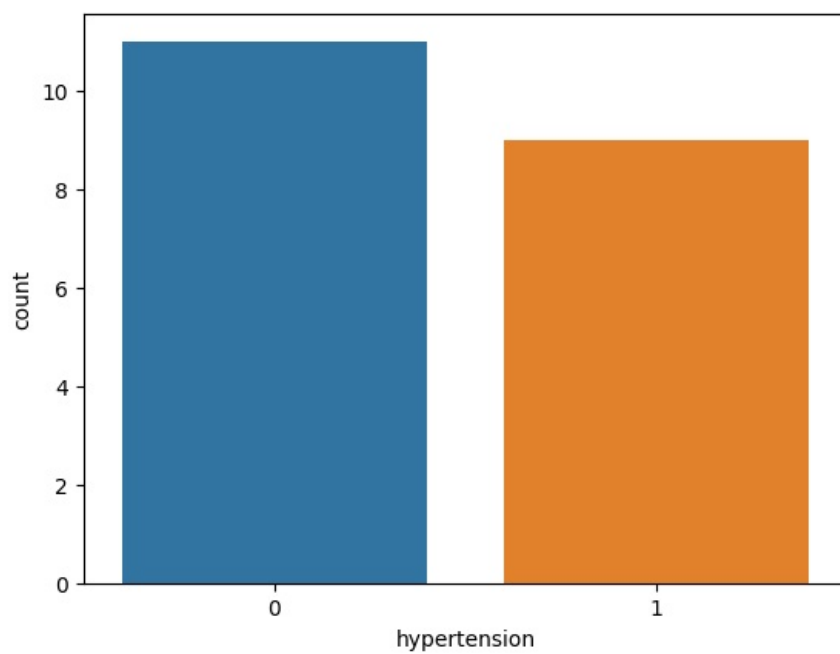
```
In [23]: sns.histplot(x='age',data=data)
```

```
Out[23]: <Axes: xlabel='age', ylabel='Count'>
```



```
In [24]: sns.countplot(x='hypertension',data=data)
```

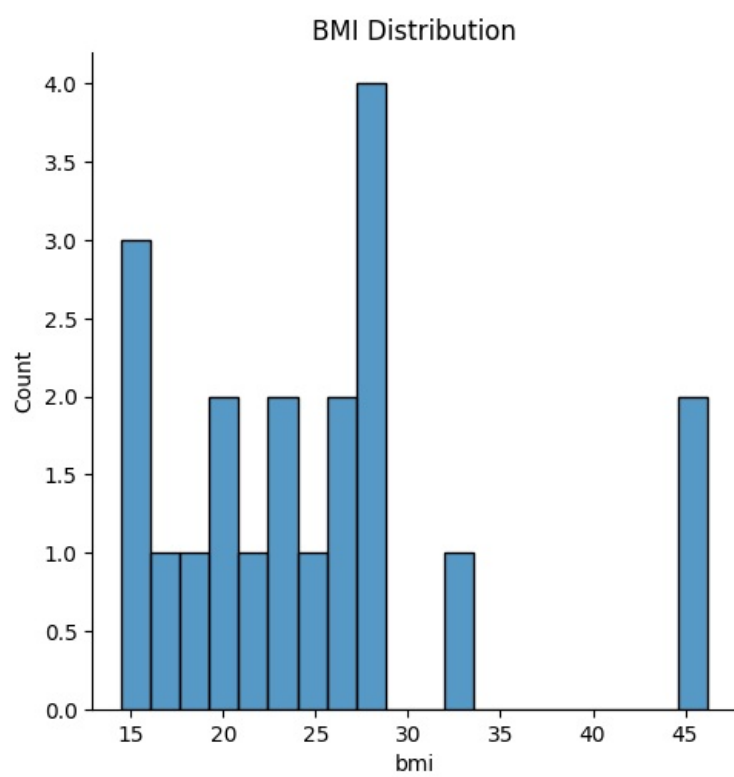
```
Out[24]: <Axes: xlabel='hypertension', ylabel='count'>
```



```
In [26]: sns.displot(data['bmi'],bins=20)
plt.title('BMI Distribution')
```

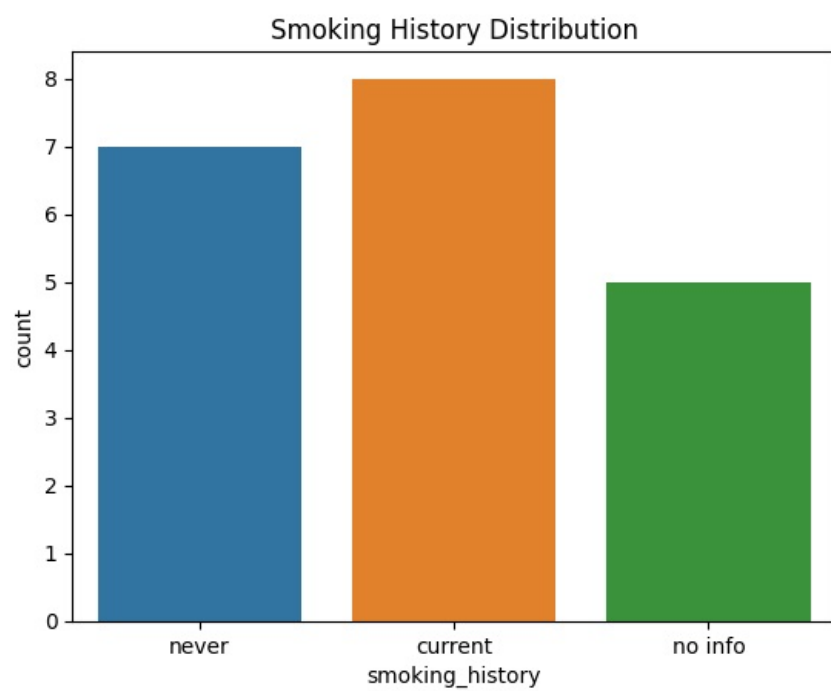
C:\Users\Jhuvvi Sree\AppData\Local\Programs\Python\Python310\lib\site-packages\seaborn\axisgrid.py:118: UserWarning: The figure layout has changed to tight
self._figure.tight_layout(*args, **kwargs)

```
Out[26]: Text(0.5, 1.0, 'BMI Distribution')
```



```
In [28]: sns.countplot(x='smoking_history', data=data)
plt.title('Smoking History Distribution')
```

```
Out[28]: Text(0.5, 1.0, 'Smoking History Distribution')
```

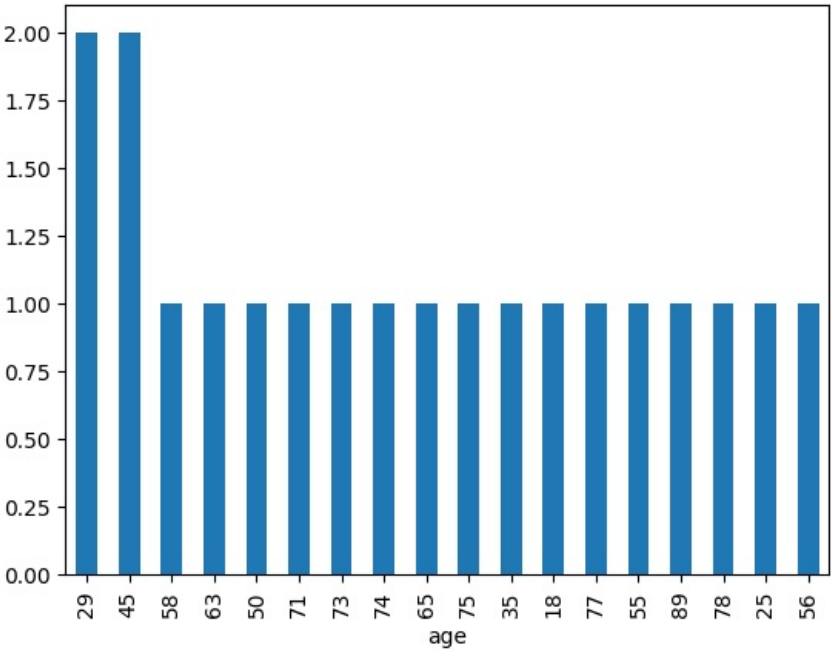
In [29]: data

Out[29]:

	s.no	gender	age	hypertension	heart_disease	obesity	stress	smoking_history	bmi	b.p level	diabetes
0	1	female	58	0	0	1	1	never	20.50	140	0
1	2	male	35	1	0	0	0	current	25.60	125	0
2	3	male	25	0	0	0	1	current	32.50	160	1
3	4	female	78	0	1	1	1	no info	28.60	180	1
4	5	male	89	0	0	1	0	never	46.00	140	0
5	6	male	55	1	1	0	1	never	26.50	120	0
6	7	male	77	1	0	1	0	current	46.20	80	1
7	8	female	18	0	0	1	1	current	15.30	150	0
8	9	female	45	0	0	0	0	never	28.60	110	1
9	10	female	75	0	1	0	1	current	14.50	110	1
10	11	male	63	0	1	1	0	never	16.40	120	0
11	12	male	65	1	0	1	1	current	25.10	125	1
12	13	female	29	1	0	0	1	no info	27.35	130	0
13	14	male	74	1	1	1	0	no info	28.00	140	0
14	15	female	73	0	1	0	1	current	23.50	180	0
15	16	male	71	1	0	1	0	never	20.30	170	1
16	17	male	50	0	0	0	1	never	23.50	140	0
17	18	female	29	1	0	0	0	no info	22.00	180	1
18	19	female	45	0	0	1	1	no info	19.20	100	0
19	20	female	56	1	0	0	1	current	15.20	80	1

In [30]: data.age.value_counts().plot.bar()

Out[30]: <Axes: xlabel='age'>



In [31]: data

Out[31]:

	s.no	gender	age	hypertension	heart_disease	obesity	stress	smoking_history	bmi	b.p level	diabetes
0	1	female	58	0	0	1	1	never	20.50	140	0
1	2	male	35	1	0	0	0	current	25.60	125	0
2	3	male	25	0	0	0	1	current	32.50	160	1
3	4	female	78	0	1	1	1	no info	28.60	180	1
4	5	male	89	0	0	1	0	never	46.00	140	0
5	6	male	55	1	1	0	1	never	26.50	120	0
6	7	male	77	1	0	1	0	current	46.20	80	1
7	8	female	18	0	0	1	1	current	15.30	150	0
8	9	female	45	0	0	0	0	never	28.60	110	1
9	10	female	75	0	1	0	1	current	14.50	110	1
10	11	male	63	0	1	1	0	never	16.40	120	0
11	12	male	65	1	0	1	1	current	25.10	125	1
12	13	female	29	1	0	0	1	no info	27.35	130	0
13	14	male	74	1	1	1	0	no info	28.00	140	0
14	15	female	73	0	1	0	1	current	23.50	180	0
15	16	male	71	1	0	1	0	never	20.30	170	1
16	17	male	50	0	0	0	1	never	23.50	140	0
17	18	female	29	1	0	0	0	no info	22.00	180	1
18	19	female	45	0	0	1	1	no info	19.20	100	0
19	20	female	56	1	0	0	1	current	15.20	80	1

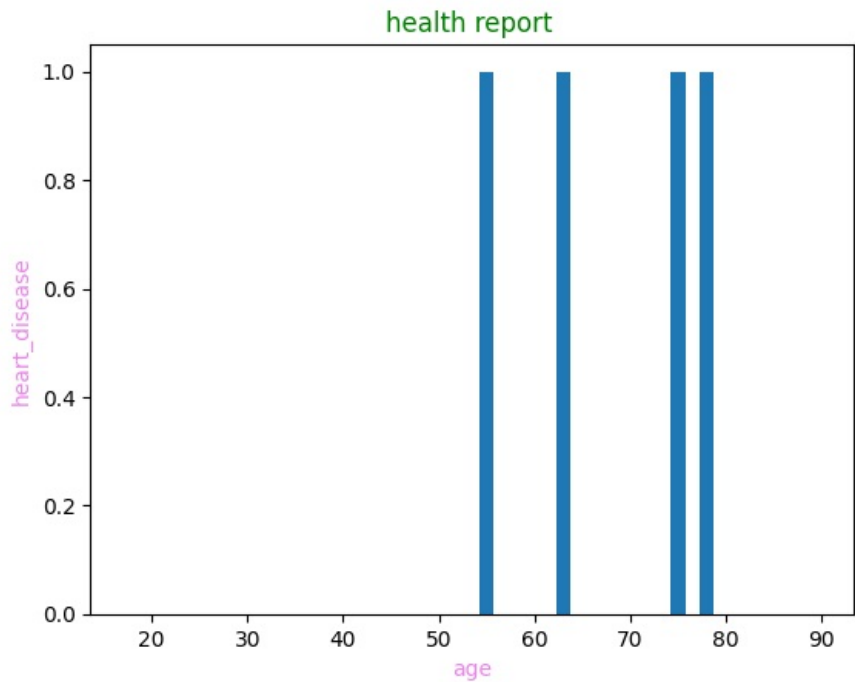
```
In [34]: x = data['heart_disease'].head(19)
y = data['age'].head(19)
```

```
In [35]: fig = plt.figure(figsize =(15,15))
```

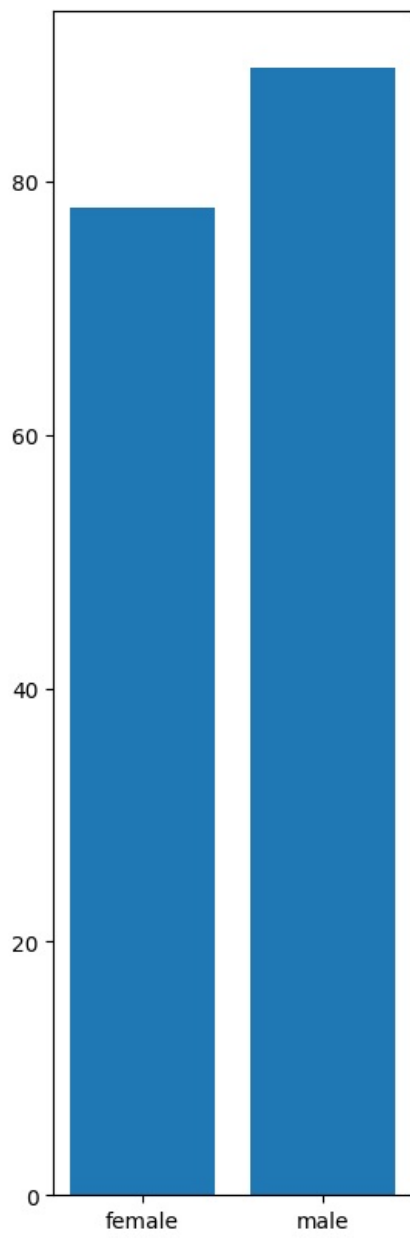
<Figure size 1500x1500 with 0 Axes>

```
In [38]: plt.bar(y[0:12], x[0:12],width=1.5)
plt.ylabel("heart_disease",color="violet")
plt.xlabel("age",color="violet")
plt.title("health report ",color="green")
```

```
Out[38]: Text(0.5, 1.0, 'health report ')
```



```
In [39]: plt.figure(figsize=(3,10))
plt.bar(data.gender,data.age)
plt.show()
```



In [47]: data

Out[47]:

	s.no	gender	age	hypertension	heart_disease	obesity	stress	smoking_history	bmi	b.p level	diabetes
0	1	female	58	0	0	1	1	never	20.50	140	0
1	2	male	35	1	0	0	0	current	25.60	125	0
2	3	male	25	0	0	0	1	current	32.50	160	1
3	4	female	78	0	1	1	1	no info	28.60	180	1
4	5	male	89	0	0	1	0	never	46.00	140	0
5	6	male	55	1	1	0	1	never	26.50	120	0
6	7	male	77	1	0	1	0	current	46.20	80	1
7	8	female	18	0	0	1	1	current	15.30	150	0
8	9	female	45	0	0	0	0	never	28.60	110	1
9	10	female	75	0	1	0	1	current	14.50	110	1
10	11	male	63	0	1	1	0	never	16.40	120	0
11	12	male	65	1	0	1	1	current	25.10	125	1
12	13	female	29	1	0	0	1	no info	27.35	130	0
13	14	male	74	1	1	1	0	no info	28.00	140	0
14	15	female	73	0	1	0	1	current	23.50	180	0
15	16	male	71	1	0	1	0	never	20.30	170	1
16	17	male	50	0	0	0	1	never	23.50	140	0
17	18	female	29	1	0	0	0	no info	22.00	180	1
18	19	female	45	0	0	1	1	no info	19.20	100	0
19	20	female	56	1	0	0	1	current	15.20	80	1

In []:

Loading [MathJax]/jax/output/CommonHTML/fonts/TeX/fontdata.js