In [1]:

```
import pandas as pd

```

In [2]:

```
data=pd.read_csv("E:\Data files\Brain Stroke.csv")
```

In [3]:

```
data
```

Out[3]:

	gender	age	hypertension	heart_disease	ever_married	work_type	Residence_type	avg
0	Male	67.0	0	1	Yes	Private	Urban	
1	Male	80.0	0	1	Yes	Private	Rural	
2	Female	49.0	0	0	Yes	Private	Urban	
3	Female	79.0	1	0	Yes	Self-employed	Rural	
4	Male	81.0	0	0	Yes	Private	Urban	
...	...	...	...	...	...	...	...	
4976	Male	41.0	0	0	No	Private	Rural	
4977	Male	40.0	0	0	Yes	Private	Urban	
4978	Female	45.0	1	0	Yes	Govt_job	Rural	
4979	Male	40.0	0	0	Yes	Private	Rural	
4980	Female	80.0	1	0	Yes	Private	Urban	

4981 rows × 11 columns

In [4]:

```
data.isnull().sum()
```

Out[4]:

```
gender               0
age                  0
hypertension         0
heart_disease        0
ever_married         0
work_type            0
Residence_type       0
avg_glucose_level    0
bmi                  0
smoking_status       0
stroke               0
dtype: int64
```

In [5]:

```
1 data["Residence_type"].unique()
```

Out[5]:

```
array(['Urban', 'Rural'], dtype=object)
```

In [6]:

```
1 data.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 4981 entries, 0 to 4980
Data columns (total 11 columns):
 #   Column             Non-Null Count  Dtype  
---  ------             --------------  -----  
 0   gender             4981 non-null   object 
 1   age                4981 non-null   float64
 2   hypertension       4981 non-null   int64  
 3   heart_disease      4981 non-null   int64  
 4   ever_married       4981 non-null   object 
 5   work_type          4981 non-null   object 
 6   Residence_type     4981 non-null   object 
 7   avg_glucose_level  4981 non-null   float64
 8   bmi                4981 non-null   float64
 9   smoking_status     4981 non-null   object 
 10  stroke             4981 non-null   int64  
dtypes: float64(3), int64(3), object(5)
memory usage: 428.2+ KB
```

In [7]:

```
1 data.describe()
```

Out[7]:

	age	hypertension	heart_disease	avg_glucose_level	bmi	stroke
count	4981.000000	4981.000000	4981.000000	4981.000000	4981.000000	4981.000000
mean	43.419859	0.096165	0.055210	105.943562	28.498173	0.049789
std	22.662755	0.294848	0.228412	45.075373	6.790464	0.217531
min	0.080000	0.000000	0.000000	55.120000	14.000000	0.000000
25%	25.000000	0.000000	0.000000	77.230000	23.700000	0.000000
50%	45.000000	0.000000	0.000000	91.850000	28.100000	0.000000
75%	61.000000	0.000000	0.000000	113.860000	32.600000	0.000000
max	82.000000	1.000000	1.000000	271.740000	48.900000	1.000000

In [8]:

```
1 data.head(10)
```

Out[8]:

	gender	age	hypertension	heart_disease	ever_married	work_type	Residence_type	avg_gl
0	Male	67.0	0	1	Yes	Private	Urban	
1	Male	80.0	0	1	Yes	Private	Rural	
2	Female	49.0	0	0	Yes	Private	Urban	
3	Female	79.0	1	0	Yes	Self-employed	Rural	
4	Male	81.0	0	0	Yes	Private	Urban	
5	Male	74.0	1	1	Yes	Private	Rural	
6	Female	69.0	0	0	No	Private	Urban	
7	Female	78.0	0	0	Yes	Private	Urban	
8	Female	81.0	1	0	Yes	Private	Rural	
9	Female	61.0	0	1	Yes	Govt_job	Rural	

In [9]:

```
1 data.tail(10)
```

Out[9]:

	gender	age	hypertension	heart_disease	ever_married	work_type	Residence_type	av
4971	Male	37.00	0	0	Yes	Private	Rural	
4972	Male	72.00	0	1	Yes	Private	Rural	
4973	Male	1.32	0	0	No	children	Rural	
4974	Male	58.00	0	0	Yes	Govt_job	Urban	
4975	Male	31.00	0	0	No	Private	Urban	
4976	Male	41.00	0	0	No	Private	Rural	
4977	Male	40.00	0	0	Yes	Private	Urban	
4978	Female	45.00	1	0	Yes	Govt_job	Rural	
4979	Male	40.00	0	0	Yes	Private	Rural	
4980	Female	80.00	1	0	Yes	Private	Urban	

In [10]:

```
1 #Slicing
2 data1=data.iloc[121:221,3:10]
```

In [11]:

```
1 data1
```

Out[11]:

	heart_disease	ever_married	work_type	Residence_type	avg_glucose_level	bmi	smoking
121	0	Yes	Private	Urban	151.16	27.5	formerly
122	0	Yes	Private	Urban	67.29	24.6	never
123	0	Yes	Govt_job	Rural	67.41	32.9	never
124	1	Yes	Private	Rural	239.07	26.1	never
125	1	Yes	Private	Urban	223.83	31.9	formerly
...	...	...	...	...	...	...	
216	1	Yes	Private	Urban	57.08	22.0	formerly
217	0	Yes	Private	Rural	162.96	39.4	never
218	0	Yes	Private	Rural	73.50	26.1	formerly
219	0	Yes	Private	Rural	95.04	42.4	never
220	0	No	Private	Rural	85.37	33.0	never

100 rows × 7 columns

In [12]:

```
1 data.shape
```

Out[12]:

```
(4981, 11)
```

In [13]:

```
1 data12=data1.loc[data1["heart_disease"]==1]
```

In [14]:

```
1 data12
```

Out[14]:

	heart_disease	ever_married	work_type	Residence_type	avg_glucose_level	bmi	smoking
124	1	Yes	Private	Rural	239.07	26.1	never
125	1	Yes	Private	Urban	223.83	31.9	formerly
129	1	Yes	Private	Rural	231.56	36.9	never
149	1	Yes	Private	Urban	112.24	37.4	
151	1	Yes	Self-employed	Urban	116.69	34.5	formerly
152	1	Yes	Govt_job	Urban	215.94	27.9	formerly
157	1	Yes	Private	Rural	271.74	31.1	
163	1	Yes	Private	Urban	175.29	31.5	formerly
177	1	No	Private	Urban	205.33	31.0	
179	1	Yes	Private	Urban	210.40	40.0	
180	1	Yes	Govt_job	Rural	219.73	28.6	never
181	1	Yes	Private	Urban	250.89	28.1	
186	1	Yes	Private	Urban	216.58	31.0	l
199	1	Yes	Private	Urban	91.92	35.9	
203	1	Yes	Private	Urban	247.51	40.5	formerly
213	1	Yes	Self-employed	Rural	243.53	27.0	never
216	1	Yes	Private	Urban	57.08	22.0	formerly

In [15]:

```
1 data12["avg_glucose_level"].mean()
```

Out[15]:

```
195.84294117647062
```

In [16]:

```
1 data12["avg_glucose_level"].var()
```

Out[16]:

```
3943.766322058824
```

In [17]:

```
1 data12["avg_glucose_level"].std()
```

Out[17]:

```
62.79941338944834
```

In [18]:

```
1 data12["Residence_type"].value_counts()
```

Out[18]:

```
Urban    12
Rural     5
Name: Residence_type, dtype: int64
```

In [ ]:

```
1
```