introduction This project going to predict the most influncer factor whether patients will attend or agnore thier opointment

my Research quetion is there relationship between age and hyperntion?

In [ ]:
```
we are going to import all libraraies that we going to use in this project
```

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

we gonna read the dataset into pandas

In [13]:
```
ds=pd.read_csv('noshowappointments-kagglev2-may-2016.csv')

ds.head()
```

Out[13]:

	PatientId	AppointmentID	Gender	ScheduledDay	AppointmentDay	Age	Neighbourhood	Scho
0	2.987250e+13	5642903	F	2016-04-29T18:38:08Z	2016-04-29T00:00:00Z	62	JARDIM DA PENHA	
1	5.589978e+14	5642503	M	2016-04-29T16:08:27Z	2016-04-29T00:00:00Z	56	JARDIM DA PENHA	
2	4.262962e+12	5642549	F	2016-04-29T16:19:04Z	2016-04-29T00:00:00Z	62	MATA DA PRAIA	
3	8.679512e+11	5642828	F	2016-04-29T17:29:31Z	2016-04-29T00:00:00Z	8	PONTAL DE CAMBURI	
4	8.841186e+12	5642494	F	2016-04-29T16:07:23Z	2016-04-29T00:00:00Z	56	JARDIM DA PENHA	

we are going to chech whither there is missiing value or nan

In [48]:
```
ds.isna().sum().sum()
```

Out[48]: 0

In [49]:
```
ds.isnull().sum().sum()
```

Out[49]: 0

'great the data is cleaned '

In [24]:
```
show_up=ds[ds['No-show']=='Yes']
```

In [27]:
```
show_up_corr=show_up.corr()
```

In [36]:
```
show_up_corr.head()
```

Out[36]:

	PatientId	AppointmentID	Age	Scholarship	Hipertension	Diabetes	Alcoholism
PatientId	1.000000	-0.006088	-0.000908	0.000109	0.000230	-0.005386	-0.007139
AppointmentID	-0.006088	1.000000	0.006243	0.022385	0.011716	0.032301	0.034371
Age	-0.000908	0.006243	1.000000	-0.079770	0.521557	0.306683	0.099859
Scholarship	0.000109	0.022385	-0.079770	1.000000	-0.023824	-0.024142	0.057870
Hipertension	0.000230	0.011716	0.521557	-0.023824	1.000000	0.439078	0.090336

'there is clear correlation between age and Hipertension'

In [55]:
```
Hipertension=show_up[show_up['Hipertension']==1]
H= Hipertension['Hipertension'].value_counts()
age=Hipertension['Age']
```

In [64]:
```
age
```

Out[64]:
```
44        78
212       62
270       45
272       51
352       62
          ..
110386    48
110399    17
110492    33
110496    37
110515    33
Name: Age, Length: 3772, dtype: int64
```