

## Step 1 : Importing Libraries

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

## Step 2 : Loading Dataset

```
In [5]: data= pd.read_csv("C:/Users/Suhel Gundre/Downloads/California.csv")
```

## Data Discription:

- gender: male or female(Binary) Here (binary: "1", means "Male", "0" means "Female")

- age: Age of the patient;(Continuous - Although the recorded ages have been truncated to whole numbers, the concept of age is continuous)

==> Behavioral

- CurrentSmoker: whether or not the patient is a current smoker.((binary: "1", means "Yes", "0" means "No") )

- CigsPerDay: the number of cigarettes that the person smoked on average in one day. (can be considered continuous as one can have any number of cigarettes, even half a cigarette.)

==> Medical( history)

- BPMeds: whether or not the patient was on blood pressure medication (Binary)

- prevalentStroke: whether or not the patient had previously had a stroke (Binary) .(Stroke is a type of disease occurs when something blocks blood supply to part of the brain.)

- prevalentHyp: whether or not the patient was hypertensive (Binary).(it is a heart problems that occur because of high blood pressure that is present over a long time.)

- diabetes: whether or not the patient had diabetes (Binary)

==> Medical(current)

- totChol: total cholesterol level (Continuous) (normal cholesterol is less than 200 mg/dl.

Blood pressure is measured using two numbers:

- SysBP: systolic blood pressure (Continuous) (it measures the pressure in your arteries when