data acquisition

```
In [3]:  #Aim : To perform operation of data acquisition
```

```
In [ ]:  #Name :siddhi n. sakharkar
         #Roll no: 51
         #Subject : DSS
         #Sec:B
```

```
In [4]:  import pandas as pd
```

```
In [9]:  import os
```

```
In [15]:  os.getcwd()
```

```
Out[15]:  'C:\\Users\\HP\\Desktop'
```

```
In [18]:  os.chdir("C:\\Users\\HP\\Desktop")
```

```
In [2]:  df=pd.read_csv("diabetes.csv")
```

```
---------------------------------------------------------------------------
NameError                                 Traceback (most recent call last)
~\AppData\Local\Temp\ipykernel_5104\1462985696.py in <module>
----> 1 df=pd.read_csv("diabetes.csv")

NameError: name 'pd' is not defined
```

```
In [20]:  df.head()
```

Out[20]:

	Pregnancies	Glucose	BloodPressure	SkinThickness	Insulin	BMI	DiabetesPedigreeFunction	Age	Outcome
0	6	148	72	35	0	33.6	0.627	50	1
1	1	85	66	29	0	26.6	0.351	31	0
2	8	183	64	0	0	23.3	0.672	32	1
3	1	89	66	23	94	28.1	0.167	21	0
4	0	137	40	35	168	43.1	2.288	33	1

```
In [21]:  df.tail()
```

Out[21]:

	Pregnancies	Glucose	BloodPressure	SkinThickness	Insulin	BMI	DiabetesPedigreeFunction	Age	Outcome
763	10	101	76	48	180	32.9	0.171	63	0
764	2	122	70	27	0	36.8	0.340	27	0
765	5	121	72	23	112	26.2	0.245	30	0
766	1	126	60	0	0	30.1	0.349	47	1
767	1	93	70	31	0	30.4	0.315	23	0

```
In [23]:  df.head(10)
```

Out[23]:

	Pregnancies	Glucose	BloodPressure	SkinThickness	Insulin	BMI	DiabetesPedigreeFunction	Age	Outcome
0	6	148	72	35	0	33.6	0.627	50	1
1	1	85	66	29	0	26.6	0.351	31	0
2	8	183	64	0	0	23.3	0.672	32	1
3	1	89	66	23	94	28.1	0.167	21	0
4	0	137	40	35	168	43.1	2.288	33	1
5	5	116	74	0	0	25.6	0.201	30	0
6	3	78	50	32	88	31.0	0.248	26	1
7	10	115	0	0	0	35.3	0.134	29	0
8	2	197	70	45	543	30.5	0.158	53	1
9	8	125	96	0	0	0.0	0.232	54	1

```
In [24]:  df.shape
```

```
Out[24]:  (768, 9)
```

```
In [26]:  df.size
```

```
Out[26]:  6912
```