```
# Experinment no. 1

# Aim : To perform operation of data acquisition

#Name = Agink Shriram Chiple

# Roll no. = 03
#Sec = c
#Subject = ET 1
#Date = 25-07-2025

# importing the basic library
import pandas as pd

import os

os.getcwd()
```

```
'C:\\Users\\shiva'
```

```
os.chdir('C:\\Users\\shiva\\OneDrive\\Desktop\\datasets')

data=pd.read_csv('diabetes.csv')

data.head()
```

```
   Pregnancies  Glucose  BloodPressure  SkinThickness  Insulin
BMI  \
0            6      148             72             35        0  33.6

1            1       85             66             29        0  26.6

2            8      183             64              0        0  23.3

3            1       89             66             23       94  28.1

4            0      137             40             35      168  43.1


   DiabetesPedigreeFunction  Age  Outcome
0                     0.627   50        1
1                     0.351   31        0
2                     0.672   32        1
3                     0.167   21        0
4                     2.288   33        1
```

```
data.shape
```

```
(768, 9)
```

```
data.size
```

```
6912
```

```
data.ndim
```

```
2
```

```
data.columns
```

```
Index(['Pregnancies', 'Glucose', 'BloodPressure', 'SkinThickness',
'Insulin',
       'BMI', 'DiabetesPedigreeFunction', 'Age', 'Outcome'],
      dtype='object')
```

```
data.describe()
```

```
       Pregnancies     Glucose  BloodPressure  SkinThickness
Insulin  \
count   768.000000  768.000000     768.000000     768.000000
768.000000
mean      3.845052  120.894531      69.105469      20.536458
79.799479
std       3.369578   31.972618      19.355807      15.952218
115.244002
min       0.000000    0.000000       0.000000       0.000000
0.000000
25%       1.000000   99.000000      62.000000       0.000000
0.000000
50%       3.000000  117.000000      72.000000      23.000000
30.500000
75%       6.000000  140.250000      80.000000      32.000000
127.250000
max      17.000000  199.000000     122.000000      99.000000
846.000000

               BMI   DiabetesPedigreeFunction          Age      Outcome
count   768.000000                768.000000  768.000000  768.000000
mean     31.992578                  0.471876   33.240885    0.348958
std       7.884160                  0.331329   11.760232    0.476951
min       0.000000                  0.078000   21.000000    0.000000
25%      27.300000                  0.243750   24.000000    0.000000
50%      32.000000                  0.372500   29.000000    0.000000
75%      36.600000                  0.626250   41.000000    1.000000
max      67.100000                  2.420000   81.000000    1.000000
```

```
data.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 768 entries, 0 to 767
Data columns (total 9 columns):
 #   Column                    Non-Null Count  Dtype
---  ------                    --------------  -----
 0   Pregnancies               768 non-null    int64
 1   Glucose                   768 non-null    int64
```

```
 2   BloodPressure              768 non-null    int64
 3   SkinThickness              768 non-null    int64
 4   Insulin                    768 non-null    int64
 5   BMI                        768 non-null    float64
 6   DiabetesPedigreeFunction  768 non-null    float64
 7   Age                        768 non-null    int64
 8   Outcome                    768 non-null    int64
dtypes: float64(2), int64(7)
memory usage: 54.1 KB

data.head(30)
```

	Pregnancies	Glucose	BloodPressure	SkinThickness	Insulin	BMI
0	6	148	72	35	0	33.6
1	1	85	66	29	0	26.6
2	8	183	64	0	0	23.3
3	1	89	66	23	94	28.1
4	0	137	40	35	168	43.1
5	5	116	74	0	0	25.6
6	3	78	50	32	88	31.0
7	10	115	0	0	0	35.3
8	2	197	70	45	543	30.5
9	8	125	96	0	0	0.0
10	4	110	92	0	0	37.6
11	10	168	74	0	0	38.0
12	10	139	80	0	0	27.1
13	1	189	60	23	846	30.1
14	5	166	72	19	175	25.8
15	7	100	0	0	0	30.0
16	0	118	84	47	230	45.8
17	7	107	74	0	0	29.6
18	1	103	30	38	83	43.3

19	1	115	70	30	96	34.6
20	3	126	88	41	235	39.3
21	8	99	84	0	0	35.4
22	7	196	90	0	0	39.8
23	9	119	80	35	0	29.0
24	11	143	94	33	146	36.6
25	10	125	70	26	115	31.1
26	7	147	76	0	0	39.4
27	1	97	66	15	140	23.2
28	13	145	82	19	110	22.2
29	5	117	92	0	0	34.1

	DiabetesPedigreeFunction	Age	Outcome
0	0.627	50	1
1	0.351	31	0
2	0.672	32	1
3	0.167	21	0
4	2.288	33	1
5	0.201	30	0
6	0.248	26	1
7	0.134	29	0
8	0.158	53	1
9	0.232	54	1
10	0.191	30	0
11	0.537	34	1
12	1.441	57	0
13	0.398	59	1
14	0.587	51	1
15	0.484	32	1
16	0.551	31	1
17	0.254	31	1
18	0.183	33	0
19	0.529	32	1
20	0.704	27	0
21	0.388	50	0
22	0.451	41	1
23	0.263	29	1
24	0.254	51	1

```
25                      0.205   41          1
26                      0.257   43          1
27                      0.487   22          0
28                      0.245   57          0
29                      0.337   38          0
```