

```
In [60]: # Name:Khushi Chandrashekhar Satpute
#Aim:To perform and analysis for Normal Distribution in given dataset
#Roll No:43
#Sec:B
#Sub:ET-II
```

```
In [62]: import pandas as pd
```

```
In [64]: import os
```

```
In [66]: os.getcwd()
```

```
Out[66]: 'C:\\Users\\asus\\Downloads'
```

```
In [68]: os.chdir("C:\\Users\\asus\\Downloads")
```

```
In [70]: data=pd.read_csv("diabetes.csv")
```

```
In [72]: data.head()
```

Out[72]:

	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 [74]: data.tail()
```

Out[74]:

	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 [76]: data.info
```

```
Out[76]: <bound method DataFrame.info of
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
..    ...    ...    ...    ...    ...    ...
763    10     101      76      48     180      32.9
764     2     122      70      27      0      36.8
765     5     121      72      23     112      26.2
766     1     126      60      0      0      30.1
767     1      93      70      31      0      30.4

      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
..                ...    ...        ...
763                0.171    63         0
764                0.340    27         0
765                0.245    30         0
766                0.349    47         1
767                0.315    23         0

[768 rows x 9 columns]>
```

```
In [78]: data.describe
```

```
Out[78]: <bound method NDFrame.describe of
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
..    ...    ...    ...    ...    ...    ...
763    10     101      76      48     180      32.9
764     2     122      70      27      0      36.8
765     5     121      72      23     112      26.2
766     1     126      60      0      0      30.1
767     1      93      70      31      0      30.4

      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
..                ...    ...        ...
763                0.171    63         0
764                0.340    27         0
765                0.245    30         0
766                0.349    47         1
767                0.315    23         0

[768 rows x 9 columns]>
```

```
In [80]: data.isna()
```

Out[80]:

	Pregnancies	Glucose	BloodPressure	SkinThickness	Insulin	BMI	DiabetesPedigreeFunction	Age	Outcome
0	False	False	False	False	False	False	False	False	False
1	False	False	False	False	False	False	False	False	False
2	False	False	False	False	False	False	False	False	False
3	False	False	False	False	False	False	False	False	False
4	False	False	False	False	False	False	False	False	False
...	...	...	...	...	...	...	...	...	...
763	False	False	False	False	False	False	False	False	False
764	False	False	False	False	False	False	False	False	False
765	False	False	False	False	False	False	False	False	False
766	False	False	False	False	False	False	False	False	False
767	False	False	False	False	False	False	False	False	False

768 rows x 9 columns

```
In [82]: data.isna().any()
```

```
Out[82]: Pregnancies      False
Glucose      False
BloodPressure      False
SkinThickness      False
Insulin      False
BMI      False
DiabetesPedigreeFunction      False
Age      False
Outcome      False
dtype: bool
```

```
In [84]: data.isna().sum()
```

```
Out[84]: Pregnancies      0
Glucose      0
BloodPressure      0
SkinThickness      0
Insulin      0
BMI      0
DiabetesPedigreeFunction      0
Age      0
Outcome      0
dtype: int64
```

