# **Project Name: Al Based Diabetes Prediction System Project Code:203476**

## **Coding:**

import pandas as pd

import numpy as np

import seaborn as sns

import matplotlib.pyplot as plt

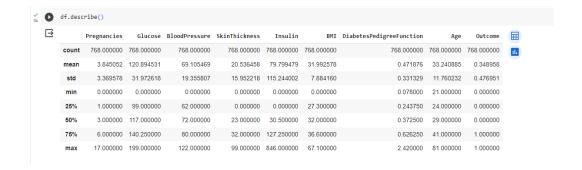
%matplotlib inline

```
import pandas as pd
import numpy as np
import seaborn as sns
import matplotlib.pyplot as plt
%matplotlib inline
```

df=pd.read\_csv('/content/diabetes (1).csv')
df.head()

```
df=pd.read_csv('/content/diabetes (1).csv')
\supseteq
       Pregnancies Glucose BloodPressure SkinThickness Insulin BMI DiabetesPedigreeFunction Age Outcome
                       148
                                       72
                                                              0 33.6
                                                                                         0.627 50
                                                                                                             ıl.
                 1
                                                              0 26.6
                                                                                         0.351 31
                                                                                                         0
                       183
                                       64
                                                              0 23.3
                                                                                         0.672 32
                                       66
                                                             94 28.1
                                                            168 43.1
                                                                                         2.288 33
```

df.describe()



#### df.info()

```
√ [7] df.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
           Glucose
                                     768 non-null
                                                    int64
                                   768 non-null int64
768 non-null int64
           BloodPressure
           SkinThickness
            Insulin
                                     768 non-null
           BMI
                                     768 non-null
                                                    float64
        6 DiabetesPedigreeFunction 768 non-null
                                                  float64
                                                    int64
           Age
                                     768 non-null
        8 Outcome
                                     768 non-null
                                                    int64
       dtypes: float64(2), int64(7)
       memory usage: 54.1 KB
```

## df.isnull().values.any()



print(df.isnull().sum())

```
print(df.isnull().sum())

→ Pregnancies

                                 0
       Glucose
                                 0
       BloodPressure
                                 0
       SkinThickness
                                0
       Insulin
                                0
       BMI
       DiabetesPedigreeFunction 0
       Age
       Outcome
       dtype: int64
```

# df.shape

```
(768, 9)
```