

```
In [2]: import pandas as pd
```

```
In [4]: df = pd.read_csv("diabetes.csv")
```

```
In [10]: df.head
```

```
Out[10]: <bound method NDFrame.head of
ess  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
..      ...    ...    ...    ...    ...
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 [11]: df.tail
```

```
Out[11]: <bound method NDFrame.tail of
ess  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
..      ...    ...    ...    ...    ...
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 [12]: df.info
```

```
Out[12]: <bound method DataFrame.info of
kness  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
..      ...     ...    ...     ...     ...
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 [13]: df.nunique
```

```
Out[13]: <bound method DataFrame.nunique of
hickness  Insulin  BMI  \
0          6    148    72
1          1     85    66
2          8    183    64
3          1     89    66
4          0    137    40
..        ...    ...    ...
763        10    101    76
764         2    122    70
765         5    121    72
766         1    126    60
767         1     93    70
Pregnancies  Glucose  BloodPressure  SkinT
0           35         0  33.6
1           29         0  26.6
2            0         0  23.3
3           23        94  28.1
4           35       168  43.1
..        ...    ...    ...
763          48       180  32.9
764          27         0  36.8
765          23       112  26.2
766           0         0  30.1
767          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 [15]: df[['BMI', 'Age']].describe()
```

```
Out[15]:
```

	BMI	Age
count	768.000000	768.000000
mean	31.992578	33.240885
std	7.884160	11.760232
min	0.000000	21.000000
25%	27.300000	24.000000
50%	32.000000	29.000000
75%	36.600000	41.000000
max	67.100000	81.000000

```
In [16]: media_bp = df["BMI"].mean()
mediana_bp = df["BMI"].median()
desv_bp = df["BMI"].std()
media_skin = df["Age"].mean()
mediana_skin = df["Age"].median()
desv_skin = df["Age"].std()
print("BMI:")
```

```

print(f" Media: {media_bp:.2f}")
print(f" Mediana: {mediana_bp:.2f}")
print(f" Desviación estándar: {desv_bp:.2f}\n")
print("Age:")
print(f" Media: {media_skin:.2f}")
print(f" Mediana: {mediana_skin:.2f}")
print(f" Desviación estándar: {desv_skin:.2f}")

```

BMI:

Media: 31.99

Mediana: 32.00

Desviación estándar: 7.88

Age:

Media: 33.24

Mediana: 29.00

Desviación estándar: 11.76

In [17]: `df.iloc[0]`

```

Out[17]: Pregnancies      6.000
          Glucose        148.000
          BloodPressure   72.000
          SkinThickness   35.000
          Insulin         0.000
          BMI            33.600
          DiabetesPedigreeFunction  0.627
          Age            50.000
          Outcome         1.000
          Name: 0, dtype: float64

```

In [18]: `df.iloc[0:2]`

```

Out[18]:
   Pregnancies  Glucose  BloodPressure  SkinThickness  Insulin  BMI  DiabetesPedigreeFunc
0            6     148             72             35         0  33.6                0.627
1            1      85             66             29         0  26.6                0.349

```



In [19]: `df[['BMI', 'Age']]`

Out[19]:

	BMI	Age
0	33.6	50
1	26.6	31
2	23.3	32
3	28.1	21
4	43.1	33
...
763	32.9	63
764	36.8	27
765	26.2	30
766	30.1	47
767	30.4	23

768 rows × 2 columns