

## Carlos Chinchilla - A00841171 - Ingeniería Industrial

In [1]: `import pandas as pd`

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

In [6]: `df.head()`

Out[6]:

	Pregnancies	Glucose	BloodPressure	SkinThickness	Insulin	BMI	DiabetesPedigreeFunc
--	-------------	---------	---------------	---------------	---------	-----	----------------------

<b>0</b>	6	148	72	35	0	33.6	0.351
<b>1</b>	1	85	66	29	0	26.6	0.671
<b>2</b>	8	183	64	0	0	23.3	0.167
<b>3</b>	1	89	66	23	94	28.1	0.178
<b>4</b>	0	137	40	35	168	43.1	0.334



In [7]: `df.tail()`

Out[7]:

	Pregnancies	Glucose	BloodPressure	SkinThickness	Insulin	BMI	DiabetesPedigreeFunc
--	-------------	---------	---------------	---------------	---------	-----	----------------------

<b>763</b>	10	101	76	48	180	32.9	0.178
<b>764</b>	2	122	70	27	0	36.8	0.344
<b>765</b>	5	121	72	23	112	26.2	0.178
<b>766</b>	1	126	60	0	0	30.1	0.327
<b>767</b>	1	93	70	31	0	30.4	0.344



In [8]: `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
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
```

```
In [9]: df.nunique()
```

```
Out[9]: Pregnancies            17
Glucose                136
BloodPressure          47
SkinThickness          51
Insulin                186
BMI                   248
DiabetesPedigreeFunction 517
Age                   52
Outcome                2
dtype: int64
```

```
In [12]: df[['BloodPressure', 'SkinThickness']].describe()
```

```
Out[12]:
```

	BloodPressure	SkinThickness
<b>count</b>	768.000000	768.000000
<b>mean</b>	69.105469	20.536458
<b>std</b>	19.355807	15.952218
<b>min</b>	0.000000	0.000000
<b>25%</b>	62.000000	0.000000
<b>50%</b>	72.000000	23.000000
<b>75%</b>	80.000000	32.000000
<b>max</b>	122.000000	99.000000

```
In [16]: media_bp = df["BloodPressure"].mean()
mediana_bp = df["BloodPressure"].median()
desv_bp = df["BloodPressure"].std()

media_skin = df["SkinThickness"].mean()
mediana_skin = df["SkinThickness"].median()
desv_skin = df["SkinThickness"].std()
```

```
print("BloodPressure:")
print(f"  Media: {media_bp:.2f}")
print(f"  Mediana: {mediana_bp:.2f}")
print(f"  Desviación estándar: {desv_bp:.2f}\n")

print("SkinThickness:")
print(f"  Media: {media_skin:.2f}")
print(f"  Mediana: {mediana_skin:.2f}")
print(f"  Desviación estándar: {desv_skin:.2f}")
```

BloodPressure:

Media: 69.11

Mediana: 72.00

Desviación estándar: 19.36

SkinThickness:

Media: 20.54

Mediana: 23.00

Desviación estándar: 15.95

Consulta 1:

Las personas con diabetes tienen una presión diastólica un poco más alta (promedio de 70.8 mm Hg) que las que no la tienen (68.1 mm Hg). Aunque la diferencia no es grande, puede indicar que los diabéticos tienden a tener una presión un poco más elevada debido a los cambios que la enfermedad produce en el cuerpo. Sin embargo, esta variable por sí sola no determina si alguien tiene diabetes; solo aporta una pista que debe analizarse junto con otras, como el nivel de glucosa o el índice de masa corporal (BMI).

Consulta 2:

El grosor del pliegue de la piel tiene un promedio de 20.5 mm, pero los valores cambian mucho entre las personas (la desviación estándar es alta, 15.9 mm). Esto significa que hay mucha diferencia en la cantidad de grasa corporal de una persona a otra. En general, las personas con un pliegue cutáneo más grueso podrían tener más grasa corporal y un mayor riesgo de desarrollar diabetes, aunque también se debe considerar que algunos valores en el dataset son ceros, lo que puede afectar el análisis.

Consulta 3: Se encontraron 232 personas sin diabetes con una presión arterial menor a 70 mm Hg. Esto sugiere que una presión baja puede estar asociada con un mejor estado general de salud, o con personas que tienen un peso corporal más bajo. En resumen, los que no tienen diabetes tienden a tener presiones diastólicas más bajas, lo que coincide con la idea de que la diabetes suele acompañarse de una presión un poco más alta.