

# A1.1 Aprendizaje estadístico-automático – Inteligencia Artificial

**Sergio Alejandro Zamora Dávila 616148**

Yo, he realizado esta actividad con integridad académica

1. Importa los datos del archivo "Obesidad.csv" a tu ambiente de trabajo en una Jupyter Notebook y muestra en consola un resumen, imprimiendo las primeras 10 filas de datos. Recuerda que es conveniente que el archivo "csv" y tu libreta estén en el mismo directorio.

```
In [4]: import pandas as pd
data = pd.read_csv("A1.1 Obesidad.csv")
# Lo tengo que guardar con el /, ya que al estar editando esto desde Visual Studio
print(data.head(10))
# Busque como hacer que imprimiera los 10 que se necesitan, solo habia que ponerle
```

	Sexo	Edad	Estatura	Peso	FamiliarConSobrepeso	ComeMuchasCalorias	\
0	Female	21.0	1.62	64.0		yes	no
1	Female	21.0	1.52	56.0		yes	no
2	Male	23.0	1.80	77.0		yes	no
3	Male	27.0	1.80	87.0		no	no
4	Male	22.0	1.78	89.8		no	no
5	Male	29.0	1.62	53.0		no	yes
6	Female	23.0	1.50	55.0		yes	yes
7	Male	22.0	1.64	53.0		no	no
8	Male	24.0	1.78	64.0		yes	yes
9	Male	22.0	1.72	68.0		yes	yes

	ComeVegetales	Fumador	ConsumoDeAgua	NivelDeObesidad	
0		2.0	no	2.0	Normal_Weight
1		3.0	yes	3.0	Normal_Weight
2		2.0	no	2.0	Normal_Weight
3		3.0	no	2.0	Overweight_Level_I
4		2.0	no	2.0	Overweight_Level_II
5		2.0	no	2.0	Normal_Weight
6		3.0	no	2.0	Normal_Weight
7		2.0	no	2.0	Normal_Weight
8		3.0	no	2.0	Normal_Weight
9		2.0	no	2.0	Normal_Weight

2. Imprime en consola o genera un cuadro de descripción que muestre un mensaje donde indiques qué otra variable hubieras decidido medir si formarás parte del grupo de investigación, indicando claramente si la información recopilada se consideraría como cuantitativa o cualitativa.

```
In [15]: print("Si formara parte del grupo de investigación añadiría las siguientes variable
#Este primero me parece muy simple, ya que la cantidad de ejercicio que la persona
print("CantEjer, me gustaria saber la cantidad de horas semanales de ejercicio que
#Este dato podria ayudar a complementar informacion, ya que si una persona va const
print("VisitasAlDoctor: Quisiera saber si la persona suele revisarse con el doctor.
```

Si formara parte del grupo de investigación añadiría las siguientes variables:

CantEjer, me gustaria saber la cantidad de horas semanales de ejercicio que cada persona hace, sería una variable cuantitativa ya que se mediría en horas

VisitasAlDoctor: Quisiera saber si la persona suele revisarse con el doctor. Lo clasificaría en 5 categorías (variable cualitativa): 1 = nunca, 2 = solo cuando es necesario, 3 = cada 3 meses aproximadamente, 4 = al menos una vez al mes, y 5 = casos especiales de revisiones diarias o muy frecuentes.

3. Si tú fueras el líder del proyecto y quisieras realizar un estudio de inferencia, ¿qué variable de la base de datos definirías como la salida o respuesta? Aquella variable para la que te interesa encontrar asociaciones específicas. Imprime el promedio de dicha variable, si es cuantitativa, o la cantidad de personas que pertenecen a cada categoría, si es cualitativa.

- a. Para calcular el promedio puedes usar la función “mean()” de pandas.

```
In [16]: # Utilizaria el valor de Peso como salida o respuesta, con los demas datos de la ba
import numpy as np
print("El peso promedio de las personas en la base de datos es de:")
# Esto lo aprendi de la "Lectura previa"
print(np.mean(data.Peso))
```

El peso promedio de las personas en la base de datos es de:  
86.58605812648035

- b. Para calcular la cantidad de personas que pertenecen a cada categoría puedes usar la función “value\_counts()” de pandas.

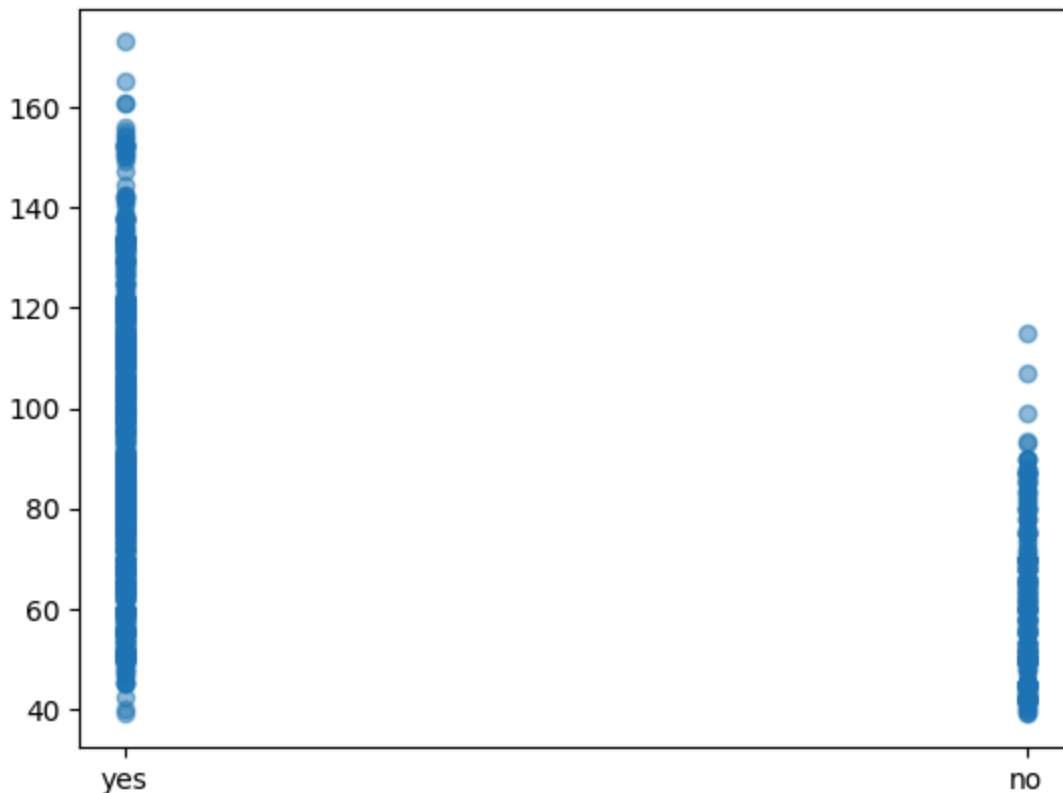
```
In [17]: # Esto es similar al anterior, ya que se podria inferir un peso estimado del indivi
print("Cantidad de personas en cada categoría de NivelDeObesidad:")
print(data['NivelDeObesidad'].value_counts())
```

Cantidad de personas en cada categoría de NivelDeObesidad:

```
NivelDeObesidad
Obesity_Type_I      351
Obesity_Type_III    324
Obesity_Type_II     297
Overweight_Level_I  290
Overweight_Level_II 290
Normal_Weight       287
Insufficient_Weight 272
Name: count, dtype: int64
```

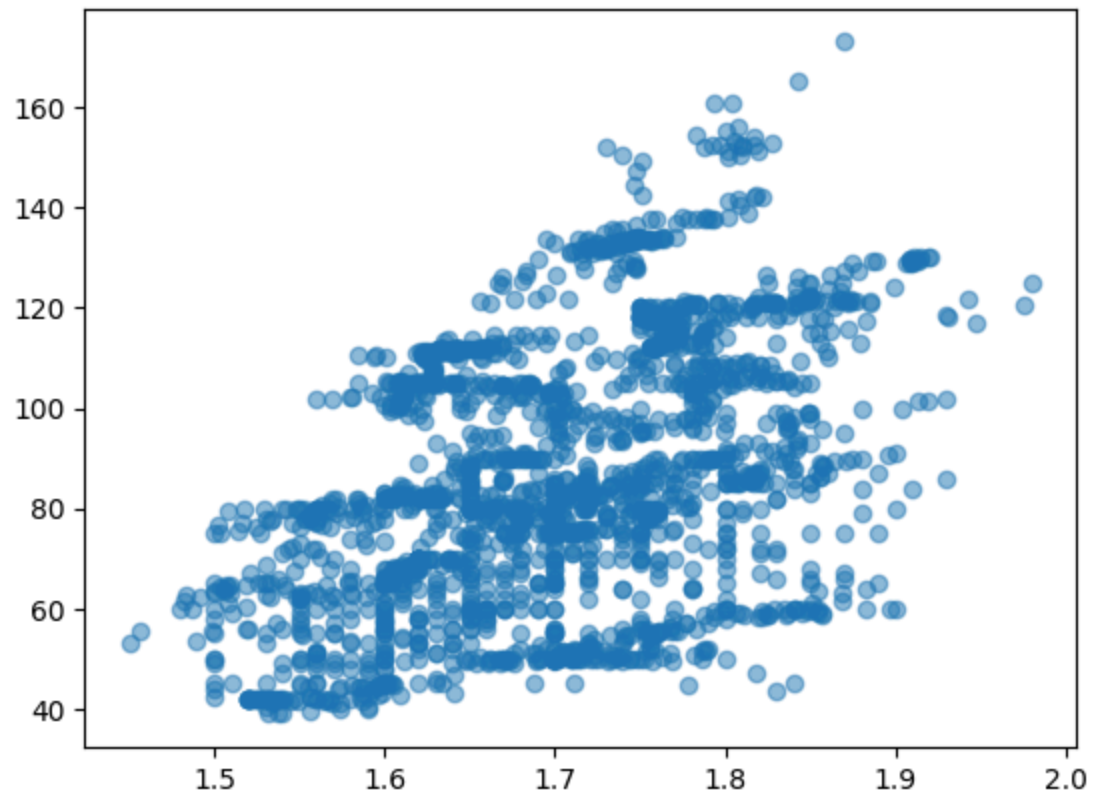
Vi que en la r brica viene "Puedo representar gr ficamente informaci n relevante de una base de datos, modificando par metros b sicos de visualizaci n.", entonces aqu  le va unas gr ficas ac  bien padres

```
In [14]: %matplotlib inline
import matplotlib.pyplot as plt
plt.scatter(x = data.FamiliarConSobrepeso, y = data.Peso, alpha = 0.5)
plt.show()
print("Gracias a la grafica se puede suponer que las personas que tienen un Familia
```



Gracias a la grafica se puede suponer que las personas que tienen un Familiar con Sobrepeso tienen mayores probabilidades de tener un peso mayor

```
In [13]: plt.scatter(x = data.Estatura, y = data.Peso, alpha = 0.5)
plt.show()
print("La grafica muestra una leve de relacion positiva entre las variables de Esta
```



La grafica muestra una leve de relacion positiva entre las variables de Estatura y P  
eso