

# Proyecto - Sección 1.2

Probabilidad y Estadística ejercicios

Daniela Jijón, Juan Francisco Cisneros y Luciana Valdivieso

18 de junio de 2022

---

## Estadísticas Descriptivas

### Cálculos

Los cálculos se realizaron con Excel y Python.

El tamaño de la muestra de los datos recolectados es  $n=202$ .

### Media

La media representa el valor promedio de los datos, se representa como la sumatoria de todos los datos dividido para el numero de datos  $n$ . La sumatoria de los  $n$  valores se realiza iterando por cada valor de la variable y sumándolos.

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n}$$

Edad:

$$\bar{x} = \frac{\sum_{i=1}^{202} x_i}{202} = 19.7, 20 \text{ años}$$

Año de estudio:

$$\bar{x} = \frac{\sum_{i=1}^{202} x_i}{202} = 2.3, \text{segundo año de carrera}$$

Horas de estudio por semana:

$$\bar{x} = \frac{\sum_{i=1}^{202} x_i}{202} = 12.2$$

GPA ponderado:

$$\bar{x} = \frac{\sum_{i=1}^{202} x_i}{202} = 3.47$$

### Mediana

Para calcular el valor de la mediana necesitamos la posición del valor. Los valores deben estar ordenados de menor a mayor, en el código se transfieren los  $n$  valores de cada variable a una lista para poder ser ordenados y acceder a la posición del valor de la media. Dado que  $n$  es par, por lo que se usa la fórmula que es el promedio entre  $(\frac{n}{2})^{n-esima}$  y  $(\frac{n}{2}+1)^{n-esima}$  posición.

$$\left(\frac{202}{2}\right)^{n-esima} + \left(\frac{202}{2} + 1\right)^{n-esima} = 203$$

$$\frac{203}{2} = 101.5^{n-esima} \text{ posición}$$

Edad:  
 $\tilde{x} = 19$

Año de estudio:  
 $\tilde{x} = 2$

Horas de estudio por semana:  
 $\tilde{x} = 13$

GPA ponderado:  
 $\tilde{x} = 3.5$

### Moda

La moda es el valor más frecuente.

Edad:  
 $Mo = 18$

Año de estudio:  
 $Mo = 2$

Horas de estudio por semana:  
 $Mo = 8$

GPA ponderado:  
 $Mo = 4$

### Varianza y Desviación Estandar

$$s^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1}$$

$$s^2 = \frac{\sum_{i=1}^{202} (x_i - \bar{x})^2}{202-1}$$

$$s = \sqrt{s^2}$$

Edad:  
 $s^2 = \frac{\sum_{i=1}^{202} (x_i - 19.7)^2}{202-1} = 3.4$   
 $s = 1.98$

Año de estudio:  
 $s^2 = \frac{\sum_{i=1}^{202} (x_i - 2.3)^2}{202-1} = 1.35$   
 $s = 1.21$

Horas de estudio por semana:  
 $s^2 = \frac{\sum_{i=1}^{202} (x_i - 12.2)^2}{202-1} = 49.28$   
 $s = 7.06$

GPA ponderado:

$$s^2 = \frac{\sum_{i=1}^{202} (x_i - 3.47)^2}{202-1} = 0.16$$

$$s = 0.4$$

### Cuartiles

Para los cuartiles debemos usar la fórmula de percentiles, son los percentiles 25, 50, 75 y 100.

$$i = \frac{25}{100}(n+1) = 50.5$$

$$i = \frac{50}{100}(n+1) = 101$$

$$i = \frac{75}{100}(n+1) = 151.5$$

Para las posiciones no enteras, se deben promediar las posiciones  $i$  e  $i+1$

Edad:

$$Q_1 = 18$$

$$Q_2 = 19$$

$$Q_3 = 20$$

Año de estudio:

$$Q_1 = 1$$

$$Q_2 = 2$$

$$Q_3 = 3$$

Horas de estudio por semana:

$$Q_1 = 8$$

$$Q_2 = 13$$

$$Q_3 = 18$$

GPA ponderado:

$$Q_1 = 3.2$$

$$Q_2 = 3.5$$

$$Q_3 = 3.8$$

### Rango Intercuartil

Usamos los valores obtenidos para los cuartiles y reemplazamos en la fórmula del rango intercuartil (Ric), este valor resultante indica la distancia entre el primer y tercer cuartil.

$$Ric = Q_3 - Q_1$$

Edad:

$$Ric = 20 - 18 = 2$$

Año de estudio:

$$Ric = 3 - 1 = 2$$

Horas de estudio por semana:

$$Ric = 18 - 8 = 10$$

GPA ponderado:

$$Ric = 3.8 - 3.2 = 0.6$$

## Percentil 10

$$i = \frac{k}{100}(n)$$
$$i = \frac{10}{100}(202) = 20,3$$

Ya que el valor de  $i$  no es entero, la posición del percentil es el entero inmediato superior, posición 21 de los ordenados.

Edad:

$$P_{10} = 18$$

Año de estudio:

$$P_{10} = 1$$

Horas de estudio por semana:

$$P_{10} = 2.5$$

GPA ponderado:

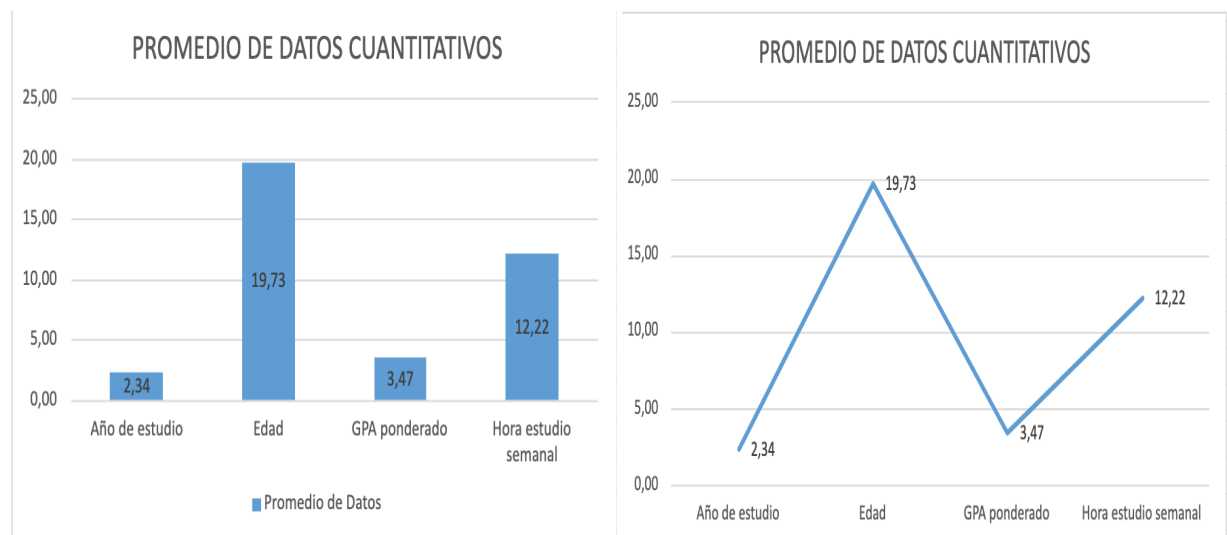
$$P_{10} = 3$$

## Límites superior e inferior

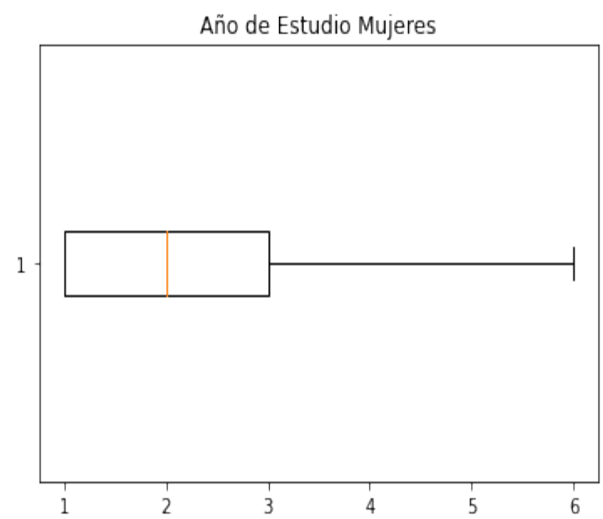
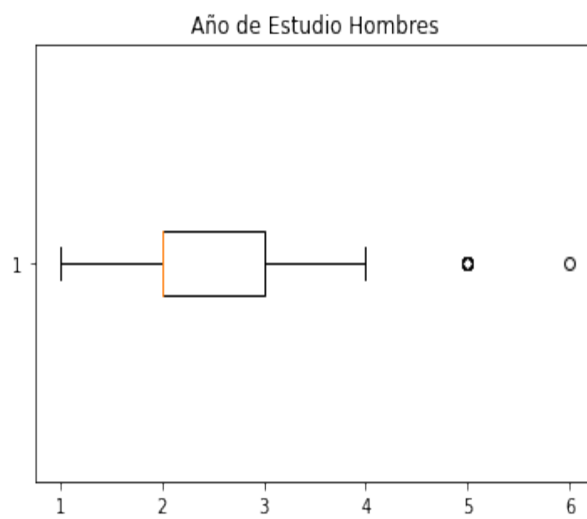
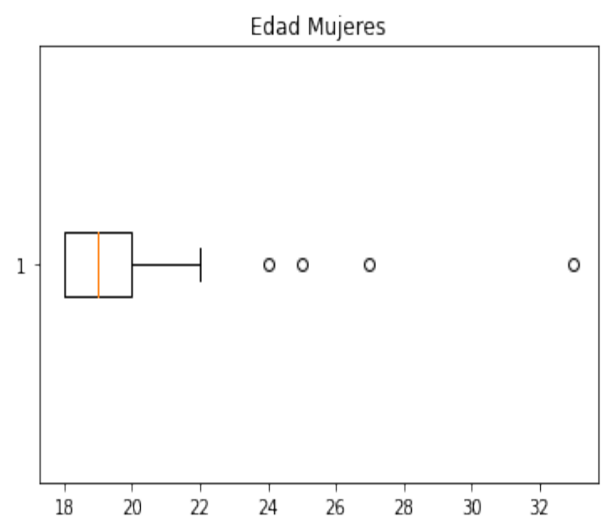
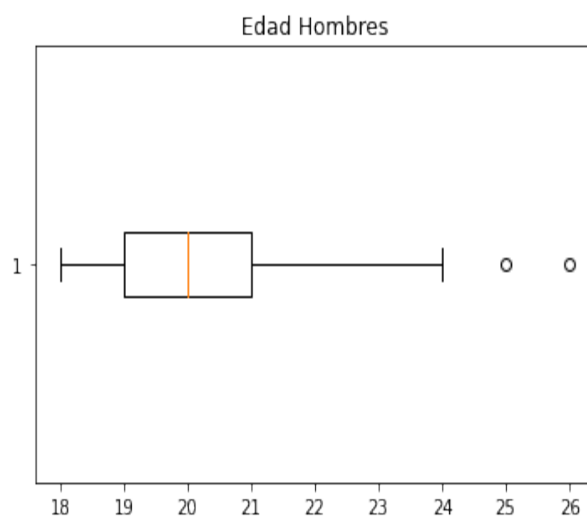
Para calcular los límites superiores e inferiores debemos ordenar nuestros datos de menor a mayor. Una vez ordenados tomamos como límite inferior el valor más pequeño y como superior el valor más alto.

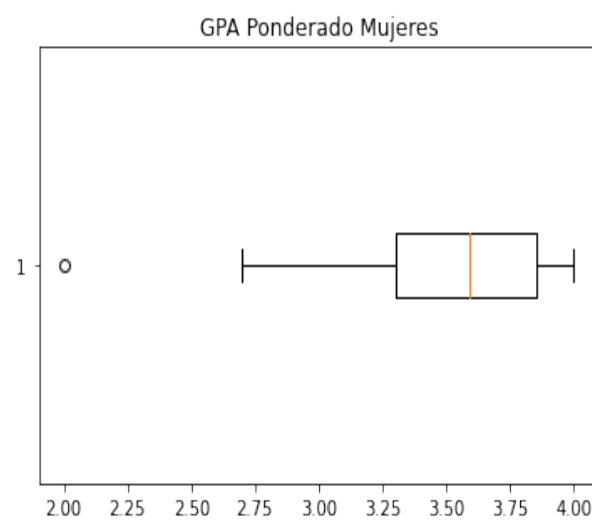
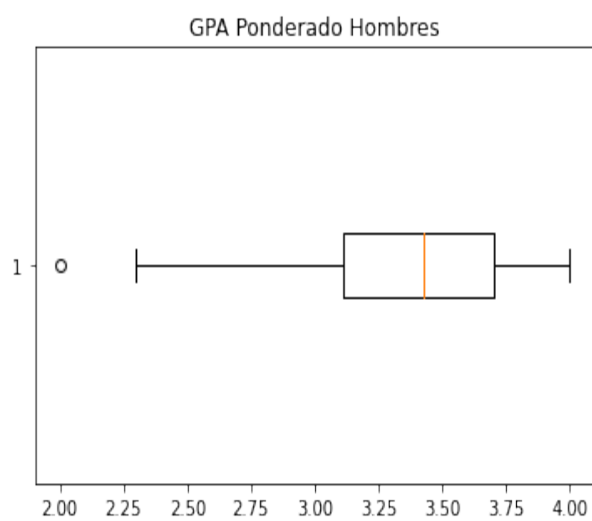
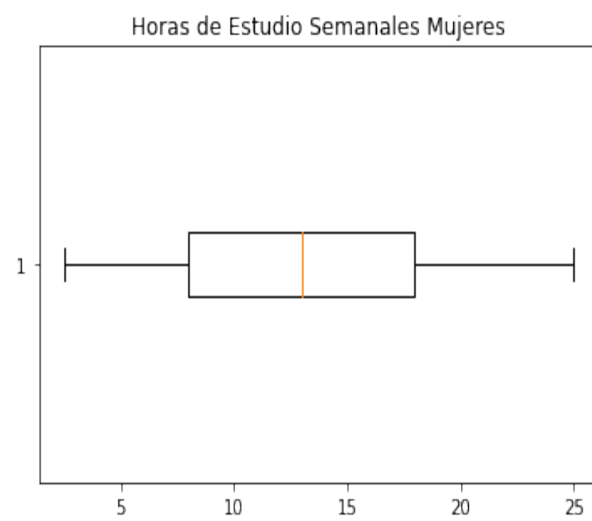
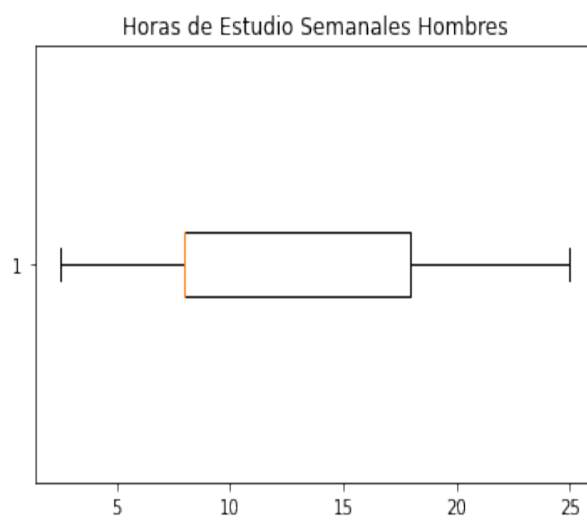
## Gráficos

### Variables Cuantitativas



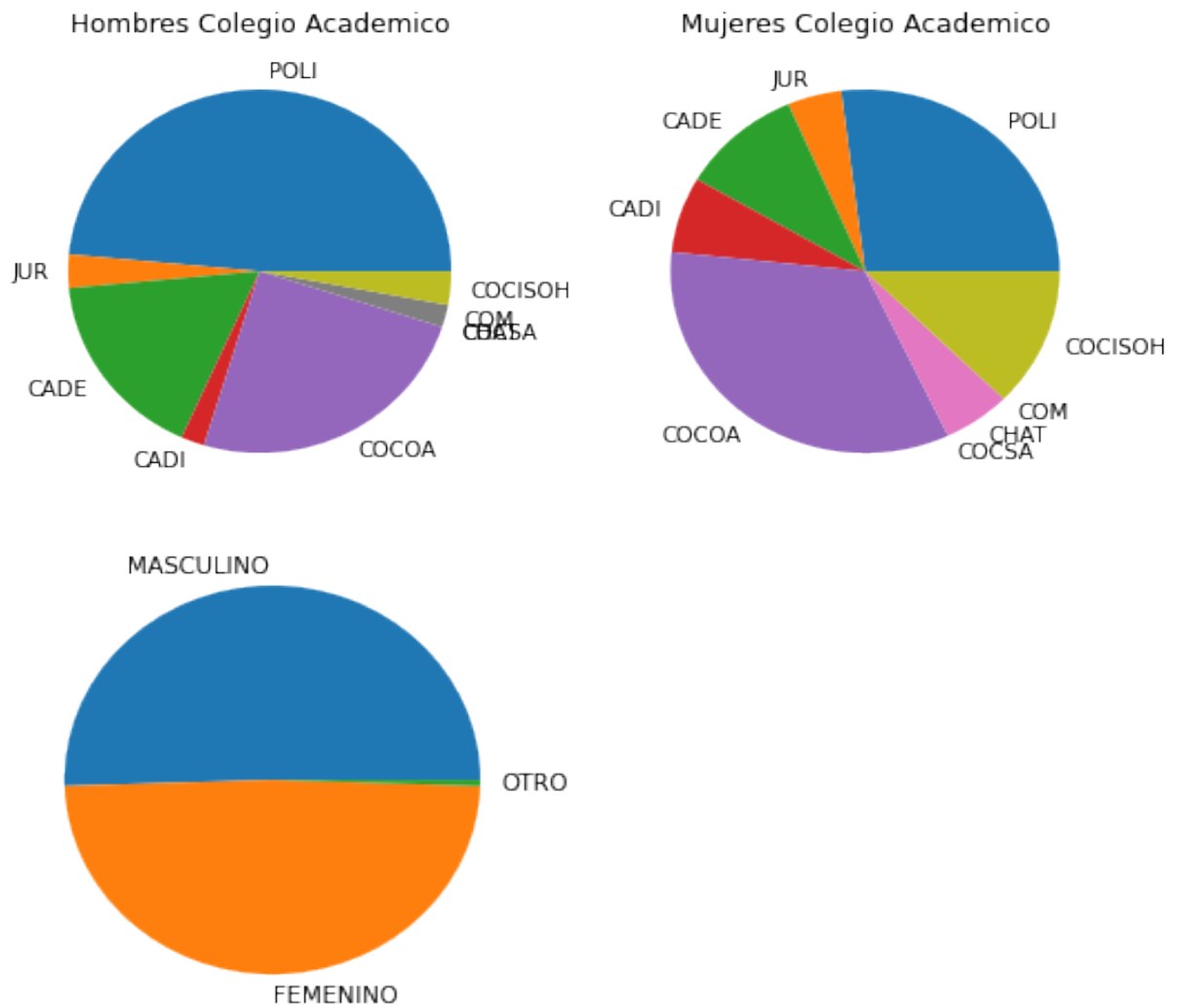
Box Plots comparativos:





## Variables Cualitativas

Pie charts comparativos:



Bar charts comparativos:

