

Complete la tabla y luego responda:

Edad	Marca de clase ( $x_{mc}$ )	Frecuencia absoluta (f)	Desviación ( $x_{mc}-x$ )	$(x_{mc}-x)^2 \cdot f$
[16, 20[		2		
[20, 24[		8		
[24, 28[		8		
[28, 32[		18		
[32, 36[		20		
[36, 40[		18		
[40, 44[				
[44, 48[		8		
[48, 52]		3		
Total		100		

- a) ¿Cuál es la varianza de las edades?
- b) ¿Cuál es la desviación estándar de las edades en que los encuestados contrajeron matrimonio?
- c) ¿Los datos están distribuidos de forma homogénea? Justifique.

# CONCLUSIONES DE LA ÚLTIMA ACTIVIDAD



- ¿Por qué la suma de la 3ra columna no dio 0? D:
- ¿Los datos están distribuidos de forma homogénea?
- Interpretemos los datos.



- **El coeficiente de variación (CV):** permite realizar comparaciones entre conjuntos con respecto a la dispersión de sus datos, e incluso entre variables que se miden con diferentes unidades de medida. Matemáticamente, corresponde al cociente entre la desviación estándar y la media aritmética. Esto es:

$$CV = \frac{S}{\bar{x}}$$

Para expresar el CV en porcentaje, basta con multiplicar el cociente obtenido por 100.

Mientras **menor sea el coeficiente de variación**, el conjunto es **más homogéneo** (los datos son más parecidos entre sí).

Mientras **mayor sea el coeficiente de variación**, el conjunto es **más heterogéneo** (los datos se diferencian más entre sí).

# ¡IMPORTANTE!

## El coeficiente de variación

Se le atribuye más a la estadística inferencial, contenido que abordaremos en la unidad 2 y 4.



# MEDIDAS DE POSICIÓN RELATIVA

(Relativa a un grupo)



Cuartiles



Quintiles

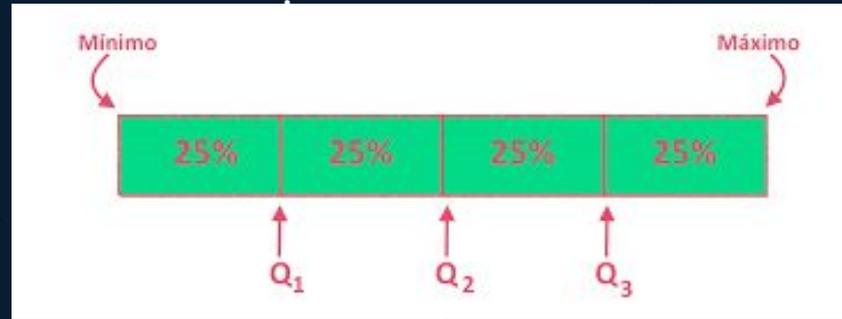


Deciles



Percentiles

- **Cuartiles:** Son los tres valores de una distribución que la dividen en cuatro partes iguales.

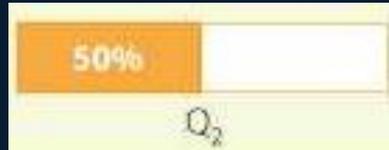


¿A qué valor corresponde el  $Q_2$  (cuartil 2)?

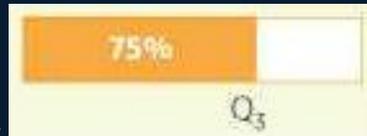
- **Primer cuartil ( $Q_1$ ):** es el valor de la variable que deja a la izquierda el 25% de los datos.



- **Segundo cuartil ( $Q_2$ ):** es el valor de la variable que deja a la izquierda el 50% de los datos. Es equivalente a la mediana.



- **Tercer cuartil ( $Q_3$ ):** es el valor de la variable que deja a la izquierda el 75% de los datos.



- **Quintiles:** son los cuatro valores de una distribución que la dividen en cinco partes iguales.

- **Primer quintil ( $Q_1$ ):** es el valor de la variable que deja a la izquierda el 20% de los datos.



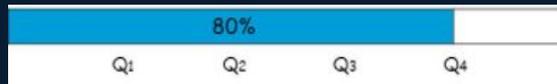
- **Segundo quintil ( $Q_2$ ):** es el valor de la variable que deja a la izquierda el 40% de los datos.



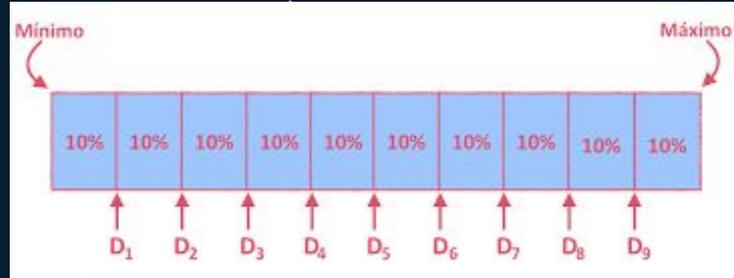
- **Tercer quintil ( $Q_3$ ):** es el valor de la variable que deja a la izquierda el 60% de los datos.



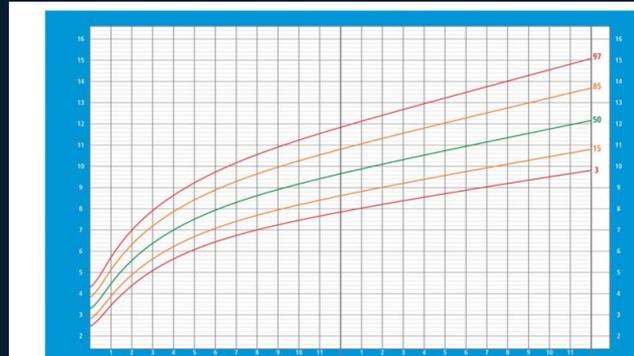
- **Cuarto quintil ( $Q_4$ ):** es el valor de la variable que deja a la izquierda el 80% de los datos.



- **Deciles:** son los nueve valores de una distribución que la dividen en diez partes iguales.



- **Percentiles ( $P_n$ ):** son los 99 valores de una distribución que la dividen en 100 partes iguales.



# IMPORTANTE

## Caso general: Cuantiles

Para cualquier valor  $q$  entre 0 y 100, se denomina cuantil  $q$  a un valor tal que:

- AL menos un  $q\%$  de las observaciones son menores o iguales a él.
- Al menos un  $(100-q)\%$  de las observaciones son mayores o iguales a él.

¿A qué corresponde los cuantiles 25, 50 y 75?



# EN LA COTIDIANIDAD... ¿CÓMO SE OCUPA LAS MEDIDAS DE POSICIÓN?

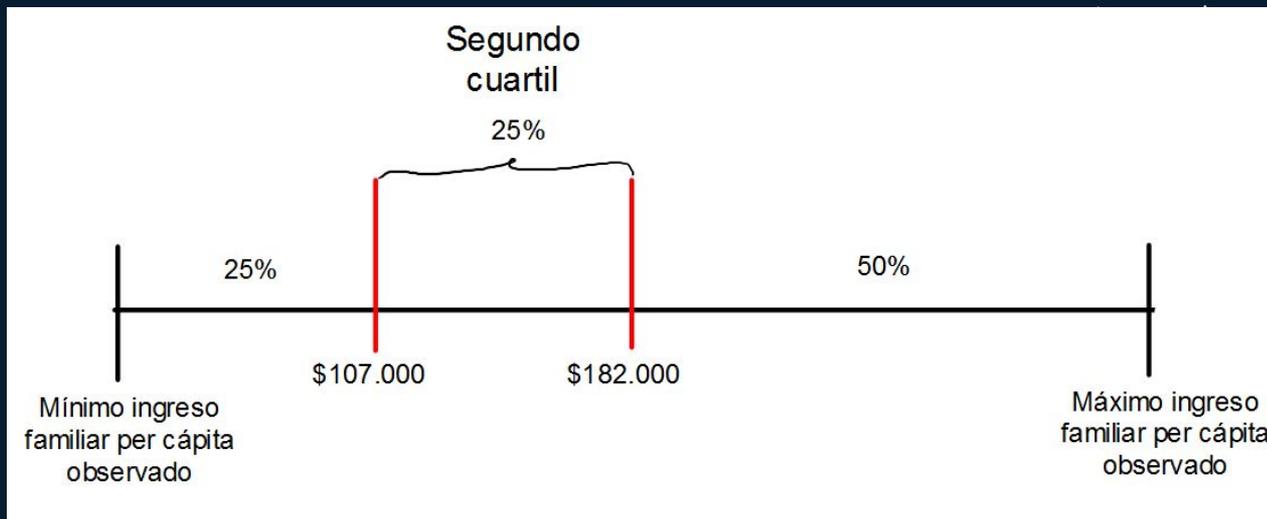
Ejemplo 1:

En el año 2007, se reportó que el primer cuartil de ingreso familiar per cápita en Chile correspondía a alrededor de \$107.000 y la mediana a alrededor de \$182.000

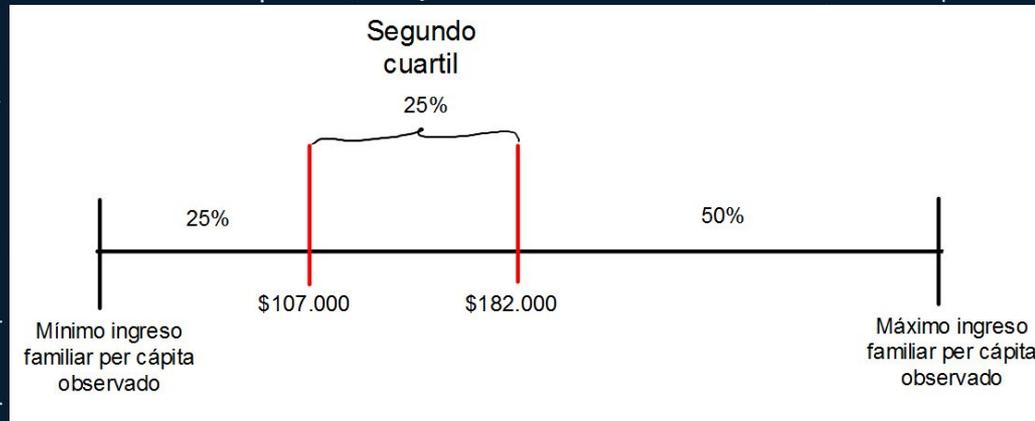


## Ejemplo 2:

En el año 2007, el segundo cuartil de ingreso familiar per cápita en Chile estaba entre \$107.000 y \$182.000

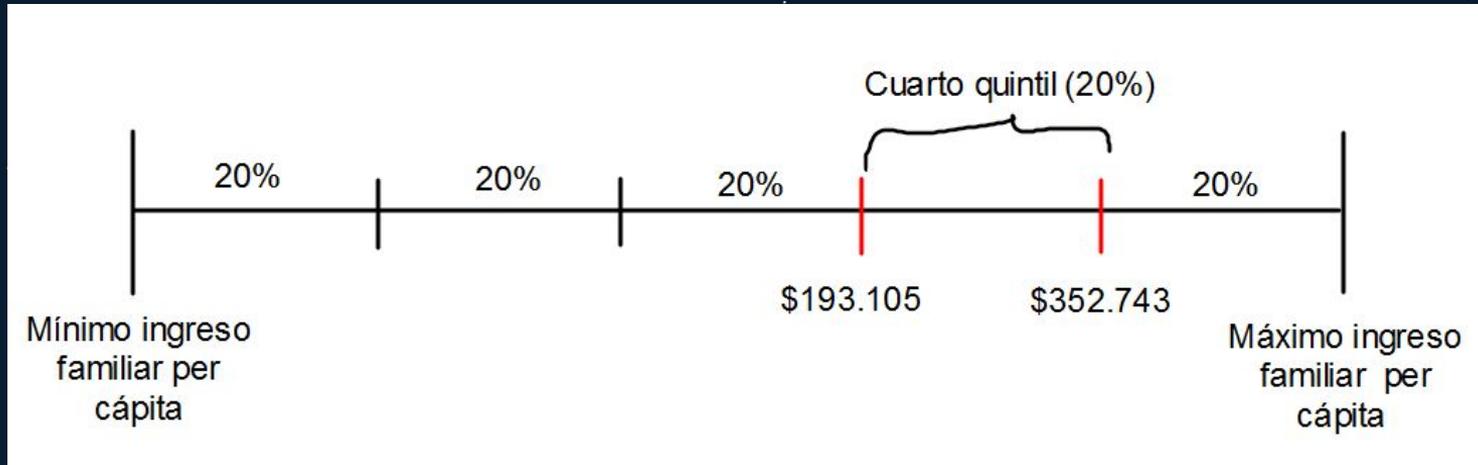


## Dos significados para el término *cuartil*



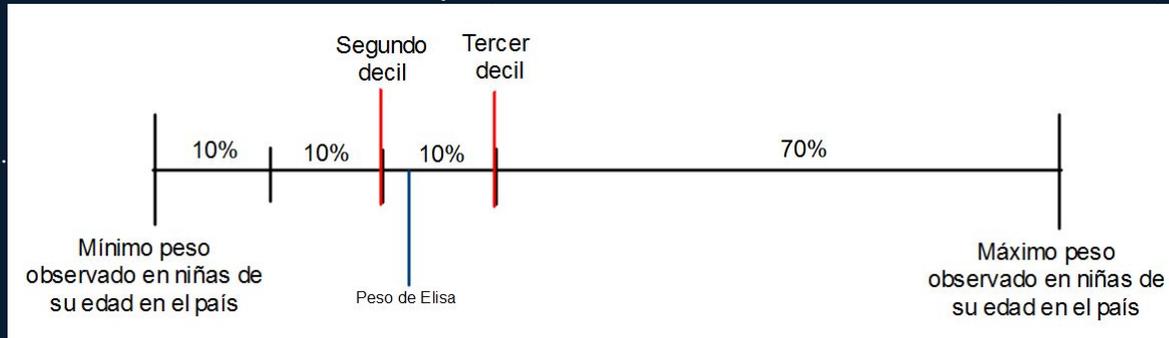
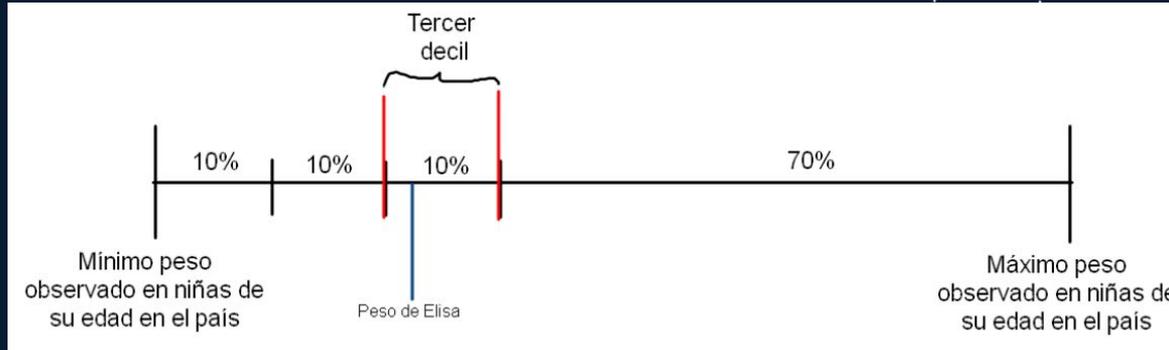
### Ejemplo 3

El cuarto quintil de ingreso familiar per cápita en Chile está entre \$193.105 y \$352.743



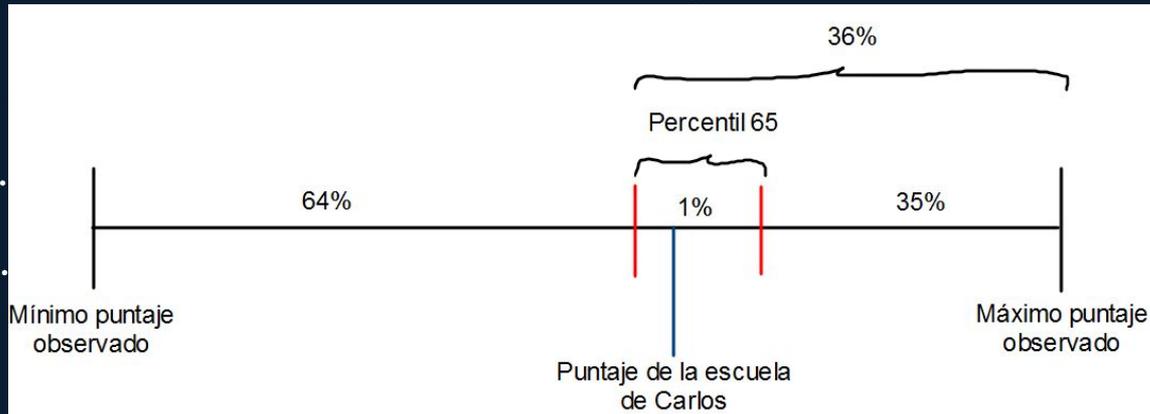
#### Ejemplo 4:

El peso de Elisa se encuentra en el tercer decil de los pesos de las niñas del país de su misma edad.



### Ejemplo 5:

El puntaje de la escuela de Carlos en la prueba SIMCE de Lenguaje de los octavos básicos del año 2012 se encuentra en el percentil 65 de los puntajes de los cursos del mismo nivel del país que rindieron dicha prueba.



### Ejemplo 6:

El médico dijo a Josefa que su estatura se encuentra en el percentil 7 de las estaturas de las niñas de su edad. Cuando Josefa le pidió que le explicara, el médico dijo: "imagina un grupo de 100 niñas de tu edad. Siete de ellas miden como tu, o menos."

# EN RESUMEN:

- **Formalmente**, los cuantiles (cuartiles, quintiles, deciles y percentiles) corresponden a valores puntuales de la variable de interés.
- **Informalmente** (pero frecuentemente, y no es considerado incorrecto) se habla de algunos cuantiles como rango de valores.



# ANALICEMOS

- Percentil PSU año 2010, prueba de matemática, página 129. Fuente: demre

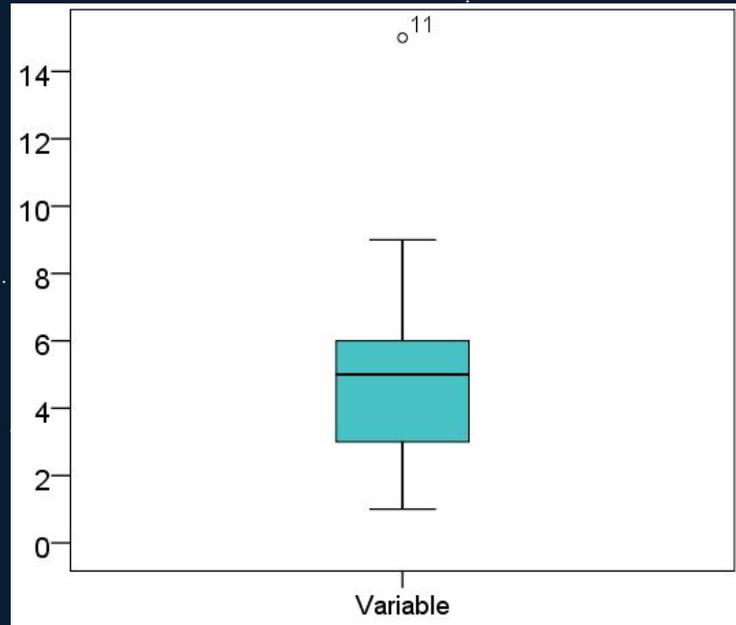
<https://demre.cl/estadisticas/documentos/p2010/2010-compendio-estadistico.pdf>

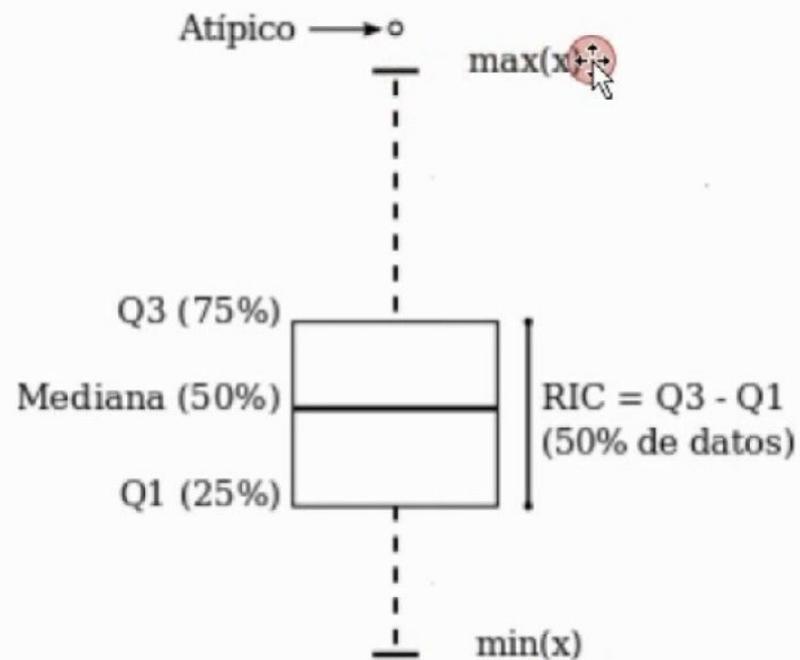
- Deciles socioeconómicos año 2017. Fuente: El Mercurio (23 de Octubre de 2017)

<https://www.emol.com/noticias/Nacional/2017/10/23/880299/Conoce-a-que-decil-pertenece-s-para-postular-a-la-gratuidad-y-becas-de-la-educacion-superior.html>

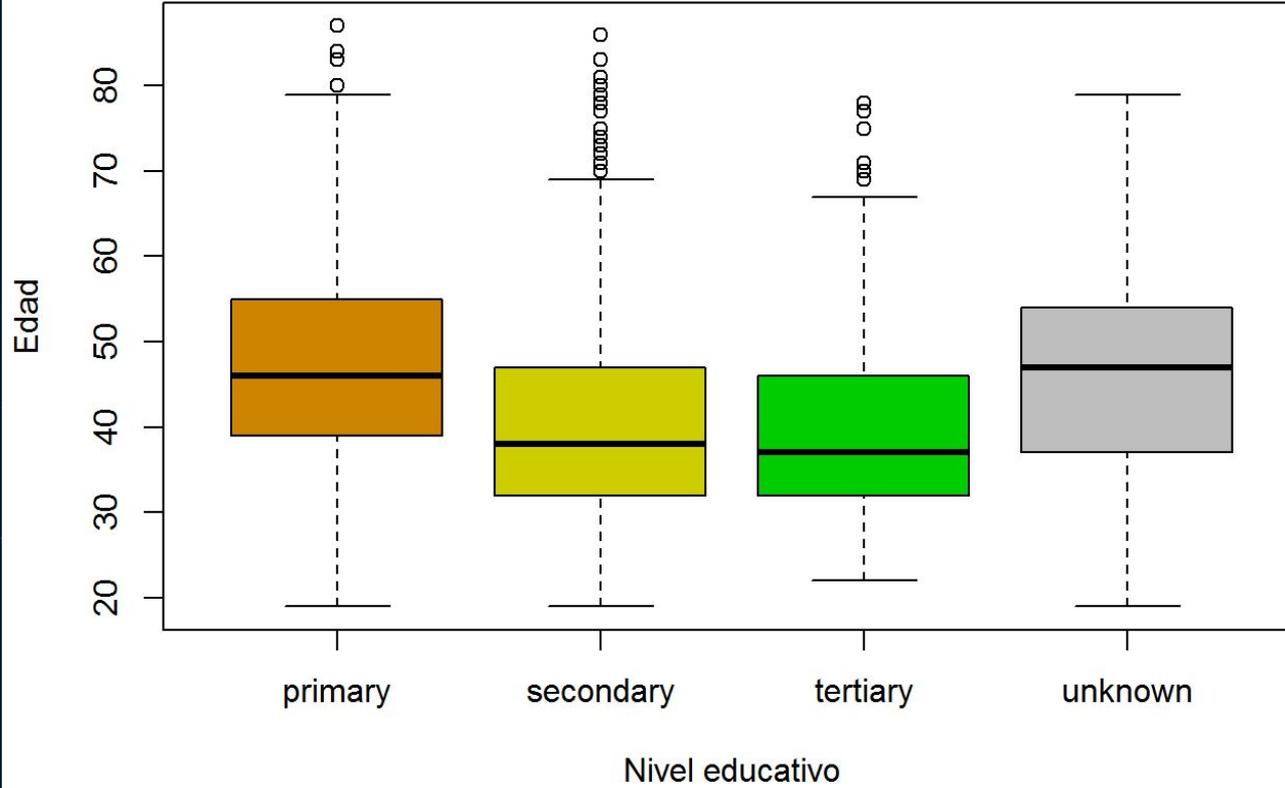
# DIAGRAMA DE CAJA Y BIGOTES

- Representa gráficamente lo que se conoce como **los cinco números de resumen** de un conjunto de datos: mínimo, primer cuartil, mediana, tercer cuartil y máximo.

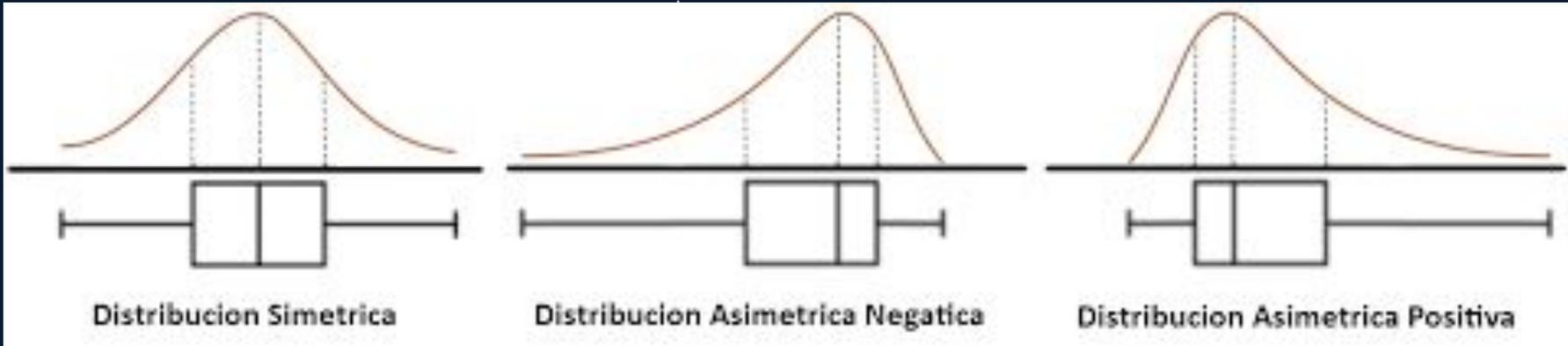




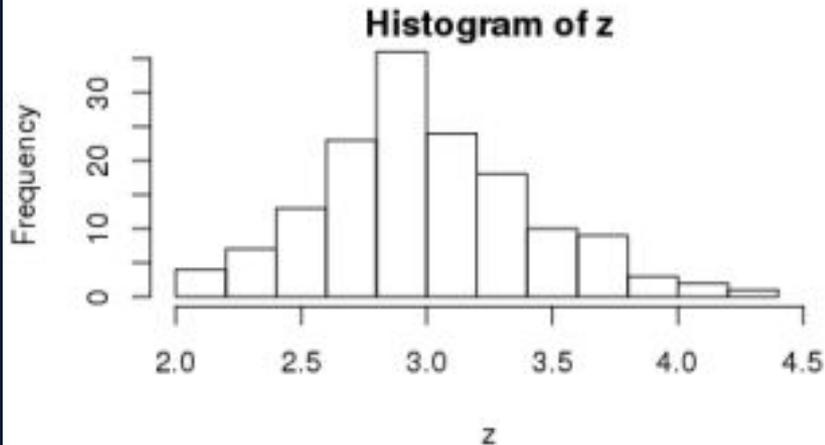
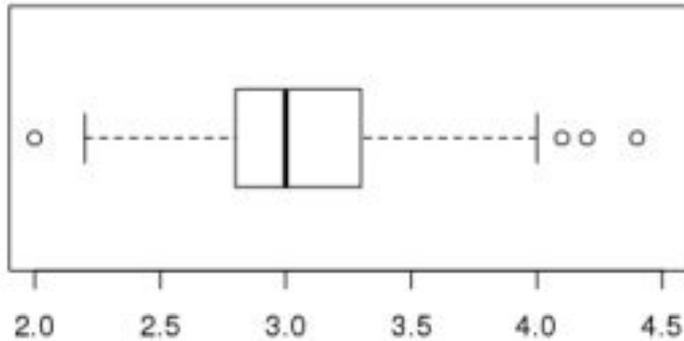
## Edad por nivel educativo



- Si la mediana está alejada del centro de la caja (cajón), esto indica que la distribución de las observaciones es asimétrica.



- Si la mediana está bastante más cercana al tercer cuartil que al primero. Esto dice que los valores de las observaciones se encuentran más concentrados en la zona superior de la caja.



- El diagrama de caja y bigotes para un conjunto de datos es único, a diferencia del histograma, que depende del número o del ancho de los intervalos, definidos por el usuario.
- Existe una relación directa entre la información que entrega un diagrama de caja y bigotes y un histograma acerca de la distribución de los datos, sin embargo, hay características que solo son identificables en uno y no en el otro.

# EN CONCLUSIÓN

## En un diagrama de caja y bigotes

- No se visualiza la media de las observaciones.
- Se puede descartar simetría global y/o simetría central de las observaciones.
- Es posible detectar la presencia de valores extremos en el conjunto de datos.



# ¿CÓMO SE CONSTRUYE UN DIAGRAMA DE CAJA Y BIGOTES?

1. Se ordenan los datos de menor a mayor.
2. Calcular puntos de interés: valor mínimo y máximo, los cuartiles y la mediana.
3. Se sitúan estos valores sobre algún eje y se dibuja la “caja” desde el primer al tercer cuartil (el recorrido intercuartílico).
4. Se dibujan los “bigotes” desde los extremos de la caja hacia los valores mínimo o máximo, según corresponda.

**NOTA:** La longitud de los bigotes no debe exceder una vez y media (1,5 veces) la de la caja, si hay valores extremos que superan esa medida se dibujan como puntos aislados.

# EJERCITEMOS JUNTOS:

La tabla muestra el consumo diario de agua, en ml, de los 20 alumnos de una clase.

Juan	1650
Luis	1300
Alma	2400
Toño	2000
Rosa	2100
Lupe	1700
Paco	1900
Tere	1500
Iris	1900
Pepe	1850
Marco	2000
Lisa	2200
Julio	2300
Maya	1600
Alex	1900
Beto	2500
Rita	2200
Marta	1650
Omar	2100
David	1750

Luis	1300
Tere	1500
Maya	1600
Marta	1650
Juan	1650
Lupe	1700
David	1750
Pepe	1850
Alex	1900
Iris	1900
Paco	1900
Marco	2000
Toño	2000
Omar	2100
Rosa	2100
Lisa	2200
Rita	2200
Julio	2300
Alma	2300
Beto	2500

Mín

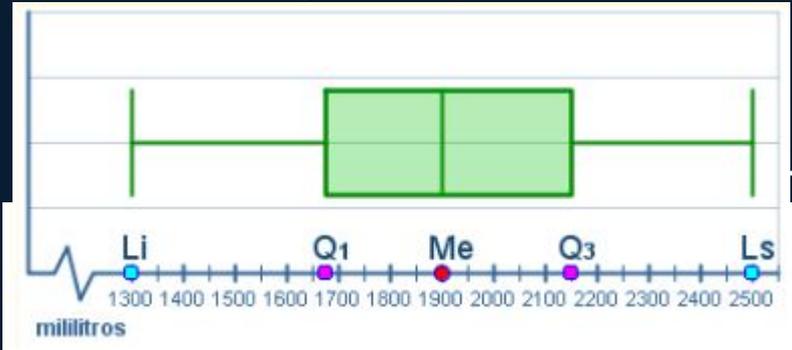
Q<sub>1</sub>

Me

Q<sub>3</sub>

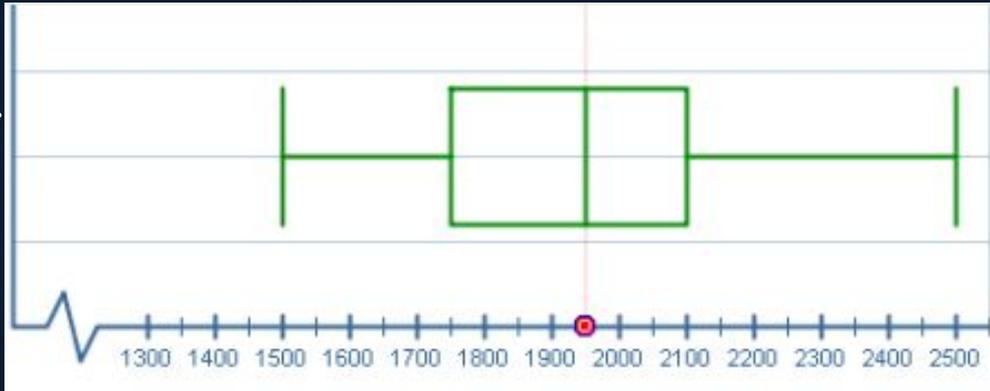
Máx

mín=1300 Q<sub>1</sub>=1675 Me=1900 Q<sub>3</sub>=2150 máx=2500



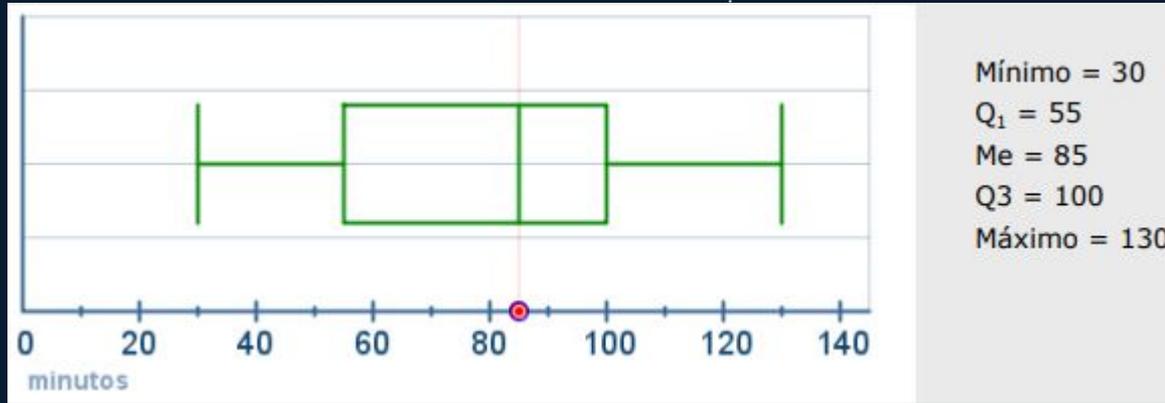
# AHORA TU:

1. Analiza el siguiente diagrama de caja y bigotes y calcula, a partir de él, los valores máximo y mínimo, la mediana y los cuartiles.



Mínimo = 1500  
 $Q_1 = 1750$   
Me = 1950  
 $Q_3 = 2100$   
Máximo = 2500

2. . Analiza el siguiente diagrama de caja y bigotes. Muestra los minutos que tarda en hacer efecto un medicamento en una población. Interpreta la información que presenta y responde a las preguntas.



- ¿A qué porcentaje de la población había hecho efecto al cabo de 30 minutos?
- Al cabo de cuántos minutos había hecho efecto al 50 % de la población?
- ¿Cuántos minutos tardó en hacer efecto al 100% de la población?
- A qué porcentaje había hecho efecto a los 55 minutos?
- ¿Cuánto tardó en hacer efecto a las tres cuartas partes de la población?

Respuestas:

- a. Al 0%, 30 es el valor mínimo.
- b. a los 85 minutos (la mediana)
- c. 130 minutos (valor máximo)
- d. 55 es el primer cuartil, al 25%
- e. 100 minutos,  $\frac{3}{4}$  partes son el 75%

