

Diagrama de Caja y Bigotes

Los diagramas de Caja-Bigotes (boxplots o box and whiskers) son una presentación visual que describe varias características importantes, al mismo tiempo, tales como la dispersión y simetría.

Para su realización se representan los tres cuartiles y los valores mínimo y máximo de los datos, sobre un rectángulo, alineado horizontal o verticalmente.

Una gráfica de este tipo consiste en una caja rectangular, donde los lados más largos muestran el recorrido intercuartílico. Este rectángulo está dividido por un segmento vertical que indica dónde se posiciona la mediana y por lo tanto su relación con los cuartiles primero y tercero (recordemos que el segundo cuartil coincide con la mediana).

Esta caja se ubica a escala sobre un segmento que tiene como extremos los valores mínimo y máximo de la variable. Las líneas que sobresalen de la caja se llaman bigotes. Estos bigotes tienen un límite de prolongación, de modo que cualquier dato o caso que no se encuentre dentro de este rango es marcado e identificado individualmente.

Ejemplo de distribución de edades:

36 25 37 24 39 20 36 45 31 31 39 24 29 23 41 40 33 24 34 40

Ordenar los datos

Para calcular los parámetros estadísticos, lo primero es ordenar la distribución:

Q1	Q2	Q3
20 23 24 24 24	25 29 31 31	33 34 36 36 37 39 39

Diagrama de Caja y Bigotes

Calculo de cuartiles

Q_1 , el cuartil Primero es el valor mayor que el 25% de los valores de la distribución.

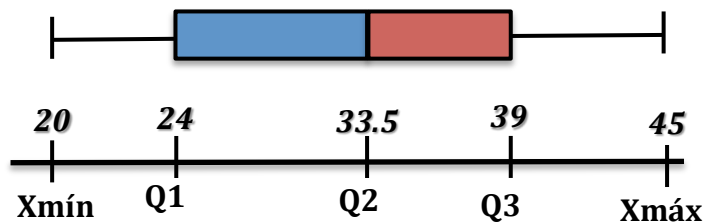
De acuerdo a lo explicado en el video, encontramos que

$$Q_1 = 24$$

$$Q_2 = 33 + 34/2 = 33.5$$

$$Q_3 = 39$$

Dibujar la caja y los bigotes



El bigote de la izquierda representa al colectivo de edades ($X_{\text{mín}}$, Q_1)

La primera parte de la caja a (Q_1 , Q_2),

La segunda parte de la caja a (Q_2 , Q_3)

El bigote de la derecha viene dado por (Q_3 , $X_{\text{máx}}$).

Diagrama de Caja y Bigotes

Información del diagrama

Podemos obtener abundante información de una distribución a partir de estas representaciones. Veamos alguna:

La parte izquierda de la caja es mayor que la de la derecha; ello quiere decir que las edades comprendidas entre el 25% y el 50% de la población están más dispersas que entre el 50% y el 75%.

El bigote de la izquierda ($X_{mín}$, $Q1$) es más corto que el de la derecha; por ello el 25% de los más jóvenes están más concentrados que el 25% de los mayores.

El rango intercuartílico = $Q3 - Q1 = 14,5$; es decir, el 50% de la población está comprendido en 14,5 años

Referencia:

(s/d). (2009). Diagrama de caja y bigotes. Recuperado a partir de: http://inst-mat.otalca.cl/tem/sitiolmde/primeroguias-liceo/recuperacion/Diagrama_de_Caja_y_Bigotes-2.pdf