

**Minnaard, Claudia**

Universidad CAECE, Universidad Nacional de Lomas de Zamora

**Condesse, Viviana**

Universidad de Buenos Aires, Universidad Nacional de Lomas de Zamora

**Minnaard, Vivian**

Instituto Superior de Formación Docente N° 19

**Rabino, Cecilia**

Universidad Nacional de Mar del Plata

Número 35/8  
10-5-05

## 1. Resumen

En 1977 John Tukey (citado por Hildebrand, 1997) publicó un tipo de gráfico estadístico para resumir información utilizando 5 medidas estadísticas: el valor mínimo, el primer cuartil, la mediana, el tercer cuartil y el valor máximo. Este tipo de gráfico recibe el nombre de *gráfico de caja* (boxplot).

En el presente trabajo mostramos la construcción e interpretación de los gráficos de caja. Así como también, las propuestas didácticas que se han presentado en clase

## 2. Introducción

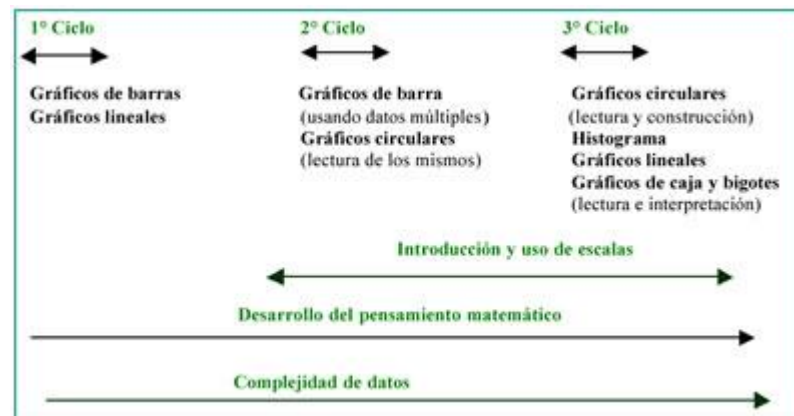
Si bien se ha considerado a la matemática como una poderosa auxiliar de otras ciencias, hoy en día, se han ampliado tanto sus posibilidades, fundamentalmente por el aporte de nuevas tecnologías, que se abren en innumerables e insospechados sentidos los servicios que puede prestar. Tal es el caso de la interpretación de gráficos estadísticos. El gráfico es una de las herramientas más útiles en el estudio de la mayoría de las disciplinas, ya que permite una visión de conjunto del fenómeno sometido a investigación, más rápidamente perceptible que la observación directa de los datos numéricos.

Las propuestas observadas, en general, en los libros de texto se refieren a la construcción y lectura de gráficos sin presentarlos en un contexto determinado. Muchos investigadores consideran importante no solo la construcción, sino también la elección del gráfico adecuado de acuerdo al contexto.

En el primer ciclo de la EGB se sugiere presentar tablas simples y gráficos de barra. Si bien conviene en una primera instancia el uso de objetos reales antes que representaciones más abstractas. En el segundo ciclo se siguen usando los gráficos que se introdujeron en años anteriores pero estos serán más complejos. En el tercer ciclo se aumentará la sofisticación matemática, la cantidad de datos, su complejidad. El alumno será capaz de comparar grupos de datos.

En el siguiente cuadro indicamos los gráficos estadísticos recomendados por los investigadores de acuerdo al nivel educativo:

**Cuadro 1: Progresión sugerida para la introducción de los diferentes gráficos estadísticos (Adaptación de Friel, S. ; Curcio, F. y Bright, G. 2001)**



En el presente trabajo focalizamos en la construcción y análisis de los gráficos de caja, también

llamados gráficos de caja y bigotes, en el nivel polimodal.

### 3. Desarrollo

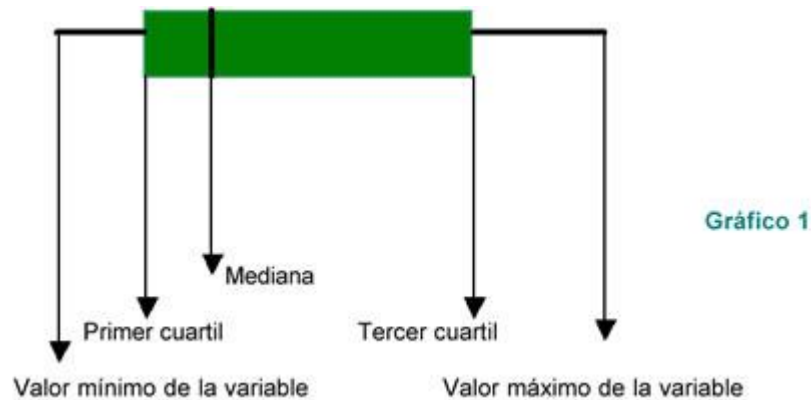
En 1977 John Tukey (citado por Hildebrand, 1997) publicó un tipo de gráfico estadístico para resumir información utilizando 5 medidas estadísticas: el valor mínimo, el primer cuartil, la mediana, el tercer cuartil y el valor máximo. Este tipo de gráfico recibe el nombre de *gráfico de caja* (boxplot).

Un gráfico de este tipo consiste en un rectángulo (caja), donde los lados más largos muestran el recorrido intercuartílico (RIC). Este rectángulo está dividido por un segmento vertical que indica donde 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).

Este rectángulo se ubica a escala sobre un segmento que tiene como extremos los valores mínimo y máximo de la variable. Estos segmentos que quedan a izquierda y a derecha de la caja se llaman bigotes. ( Freund, Williams y Perles, 1992).

Tomemos un ejemplo: (Gráfico 1)

La variable medida en este caso es : tiempo en segundos para recorrer 100 m



Mediana

Primer cuartil Tercer cuartil

Valor mínimo de la variable Valor máximo de la variable

Los bigotes tienen un límite de prolongación, de modo que aquellos valores atípicos que se separan del cuerpo principal de datos se indican individualmente. A diferencia de otros métodos de presentación de datos, los gráficos de caja muestran los valores atípicos de la variable. Llamaremos valores atípicos de la variable a aquellos que están tan apartados del cuerpo principal de los datos que bien pueden representar los efectos de causas extrañas, como algún error de medición o registro. Su eliminación no se justifica, ya que el propósito del gráfico de caja consiste en brindarnos un mayor conocimiento de la forma en que se distribuyen los datos.

Tukey ( citado por Hildebrand, 1997) introduce un criterio para fijar los extremos de los bigotes. Para esto calcula 4 barreras, dos interiores y dos exteriores:

Barrera interior inferior=Primer cuartil – 1,5 . RIC

Barrera interior superior=Tercer cuartil + 1,5 . RIC

Barrera exterior inferior=Primer cuartil – 3 . RIC

Barrera exterior superior=Tercer cuartil + 3 . RIC

Recordemos que RIC (Recorrido Intercuartílico) es igual a la diferencia entre el Tercer cuartil y el Primero.

Si se consideran los valores de la variable comprendidos entre las dos barreras interiores, el valor mínimo de la variable y el valor máximo son los extremos de los bigotes.

Si existen valores de la variable comprendidos entre las barreras interiores y exteriores se consideran valores atípicos y se indican con \*. Si existieren valores fuera de las barreras exteriores se consideran valores todavía más atípicos y se indican con · .

Por otra parte, este tipo de gráfico nos proporciona información con respecto a la simetría o asimetría de la distribución. Se utilizan los siguientes criterios: si la mediana está en el centro de la caja o cerca de él, constituye un indicio de simetría de los datos, si la mediana está considerablemente más cerca del primer cuartil indica que los datos son positivamente asimétricos y si está más cerca del tercer cuartil, señala que los datos son negativamente asimétricos. Asimismo, la longitud relativa de los bigotes se puede emplear como un indicio de su asimetría.

Una vez realizado el gráfico, ¿qué tipo de preguntas debemos formular para una mejor comprensión?

Algunas preguntas podrían ser las siguientes:

- ¿Qué porcentaje de los datos está representado por la caja?
- ¿Qué porcentaje representa cada uno de los bigotes?
- ¿Puede ser un bigote más largo que otro?. ¿Cuál es el significado?
- ¿Se encuentra la mediana siempre en el centro de la caja?

#### 4. Actividades Propuestas

1)Hildebrand (1997) propone el siguiente problema donde se muestra como actúan las barreras interiores y exteriores:

Suponga que el rendimiento de las inversiones de 21 compañías al término de un año en cierto ramo de la industria es:

-24,6	2,6	2,4	2,7	3,8	5,6	5,9
6,7	7,0	7,2	7,5	8,0	8,2	8,5
8,6	8,8	9,0	9,2	9,7	10,0	20,5

Trace un diagrama de caja para estos datos, señalando valores atípicos

#### Solución

En base a los datos obtenemos que:

Mediana: 7,5

Cuartil 1: 5,6

Cuartil 3: 8,8

RIC : 3,2

Las barreras son:

Barrera exterior inferior=Q1 - 3.0 RIC=5,6 - 3.0 (3,2)=-4,0

Barrera exterior superior=Q3 + 3.0 RIC=8,8 + 3.0 (3,2)=18,4

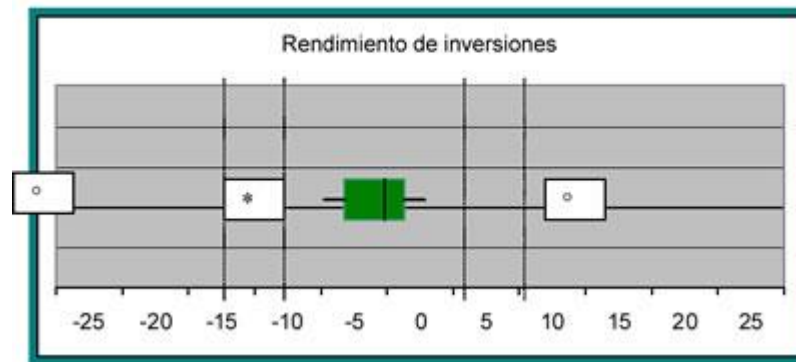
Barrera interior inferior=Q1 - 1.5 RIC=5,6 - 1.5 (3,2)=0,8

Barrera interior superior=Q3 + 1.5 RIC=8,8 + 1.5 (3,2)=13,6

La prueba de las barreras identifica dos valores atípicos importantes, -24,6 y 20,5 y un posible valor atípico, -2,6. ( Una gráfica de los datos indica que los valores atípicos importantes son obviamente valores extremos y que el valor dudoso queda posiblemente excluído).

El gráfico de caja para este problema es el siguiente:

Gráfico 2



2) Los gráficos de caja son muy útiles para hacer comparaciones, es por esto que se presenta la siguiente actividad surgida a partir de las inquietudes de algunos alumnos.

Un grupo de alumnos solicitó realizar, utilizando el gráfico de cajas, el estudio comparativo de las alturas de los jugadores de basket-ball de la N.B.A. en diferentes décadas.

A partir de los datos buscados, se obtuvieron los siguientes resultados:

	1970	1980	1990	2005
<b>Mediana</b>	200,66	205,74	200,66	205,74
<b>Cuartil I</b>	194,95	202,56	198,12	198,12
<b>Cuartil III</b>	207,01	208,28	204,47	211,45
<b>V. Máx</b>	215,9	213,36	218,44	228,6
<b>V. mín</b>	187,96	185,42	187,96	182,88
<b>Promedio</b>	201,29	204,89	201,12	203,83

Es interesante destacar cómo a partir de un ejercicio no visiblemente distinto a cualquier otro, y que por otra parte surgió de los mismos alumnos, implicó un trabajo de investigación no previsto. Como las alturas de los jugadores las buscaron en la página web de la N.B.A., notaron de inmediato que no estaban expresadas en metros si no que lo estaban en pies. Fue menester, en primer lugar, buscar la conversión a centímetros. Pero inmediatamente después hubo que analizar qué sucedía con los valores no enteros. ¿utilizan un sistema centesimal? Fue luego de varias discusiones, de búsquedas y de

cotejar con aquellos que conocían la altura de algunos jugadores, que llegaron a la conversión correcta de pies y pulgadas a centímetros.

Por otra parte, los alumnos habían supuesto a priori que las estaturas de los jugadores tendían a ser mayores a medida que pasaba el tiempo. Aunque Salzano (2002) afirma que las estaturas tienden a aumentar, esto no quedó reflejado en el análisis realizado por los alumnos.

Gráfico 3



## 5. Conclusiones

Los gráficos tienen como finalidad aclarar o facilitar la comprensión del texto que acompañan, por lo cual favorecen el aprendizaje.

Es conveniente, en la medida que avanzan en su escolaridad, ir señalando las ventajas, desventajas y la información que brindan cada tipo de gráfico.

Así, en un gráfico de tallo y hojas queda representado cada uno de los valores sin la pérdida que significa agruparlos. Un pictograma tiene una finalidad "casi decorativa", generalmente dirigido a personas no especializadas con un fin de divulgación o propaganda. Los gráficos de líneas son ideales para representar la tendencia de datos durante un intervalo de tiempo y uno de sectores puede utilizarse de manera efectiva para representar los componentes de un total.

¿Cuál es la importancia entonces del uso de los gráficos de caja?

En particular, los gráficos de caja vinculan los conceptos de mediana, cuartiles, valor mínimo y máximo que los alumnos manejan individualmente pero no en forma global.

## 6. Bibliografía

Alaminos, A. (1993). *Gráficos. Cuadernos metodológicos*. Centro de Investigaciones Sociológicas. Madrid.

Freund, J., Williams, F. y Perles, B. (1992). *Estadística para la Administración con enfoque moderno*. Editorial Prentice Hall Hispanoamericana. pp. 8 – 35, 76 – 77

Friel, S., Curcio, F. y Bright, G. (2001). *Making sense of Graphs: critical factors, influencing comprensión y instructional implication*. Journal of Research in Mathematics Education. 32 (2), pp124 –158.

Hildebrand, D. y Lyman Ott, R. (1997). *Estadística aplicada a la administración y a la economía*. Addison

– Wesley Iberoamericana.pp.36 - 37

Minnaard, V. , Minnaard, C. , Rabino, C. et al (2002). *El uso de las gráficas en la escuela: otro lenguaje de las ciencias*. Revista Iberoamericana de Educación . OEI.

Dirección en Internet: <http://www.campus-oei.org/revista/experiencias34.htm>

Salzano, F.M. y Bortolini, M. C. (2002). *The evolution and genetics of Latin American populations*.Cambridge University Press. Cambridge.