

REPRESENTACIONES TALLO-HOJA

La técnica para compendiar datos numéricos conocida como representación **tallo-hoja** en inglés, (*stem-and-leaf*) ha ganado popularidad recientemente. Estas representaciones son idóneas para aplicarse en el uso de la computadora. Muy fácil de obtener y utilizar, dicha técnica es una combinación de dos procedimientos: uno gráfico y otro de ordenación. (Ordenar datos, es formar una lista de ellos en un orden dado de acuerdo con su valor numérico). Este tipo de representación es similar a un histograma debido a que los valores de los datos se presentan en intervalos y despliegan en barras. Sin embargo, de un diagrama de tallo y hoja se puede recobrar más información de los dígitos de cada uno de los número y también se puede ver si algún valor es identificado como atípico.

Los valores de los datos se utilizan para efectuar tal ordenación. El tallo se forma con el (los) primer(os) dígito(s) del dato, mientras que la hoja se forma con los demás dígitos siguientes. Por ejemplo, el valor numérico 458 se dividiría en 45-8 como se muestra a continuación:

| Dígitos iniciales | Dígitos sucesivos |
|-----------------------------|--------------------------------|
| 45 | 8 |
| Utilizados en la ordenación | Mostrados en la representación |

Ilustración 2-1: Construir una representación tallo-hoja para el siguiente conjunto de 20 calificaciones.

82 74 88 66 58 74 78 84 96 76
62 68 72 92 86 76 52 76 82 78

En una inspección rápida se ve que hay valores en los grupos del 50, 60, 70, 80 y 90. El primer dígito de cada valor debe utilizarse como tallo, y el segundo como hoja. Casi siempre la representación se construye en disposición vertical. Trácese una recta vertical y colóquense los tallos a la izquierda, en columna.

5
6
7
8
9

En seguida colóquese cada hoja junto a su tallo. Esto se hace escribiendo el dígito sucesivo correspondiente a la derecha de la recta vertical. El valor del primer dato es 82; 8 es el tallo y 2 es la hoja. Por tanto, se ubica 2 frente al 8.

$$8 \mid 2$$

El valor del dato siguiente es 74, de manera que se escribe la hoja 4 frente al tallo 7.

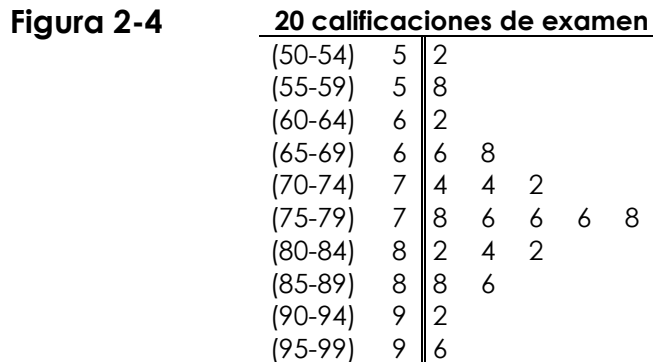
$$\begin{array}{c} 7 \\ 8 \end{array} \mid \begin{array}{c} 4 \\ 2 \end{array}$$

El proceso continúa hasta que se coloquen todas las 18 hojas restantes. La figura 2-3 muestra la representación tallo-hoja que resulta.



Generalmente es sencillo formar una representación con tallos diferentes una vez que los datos se han ingresado en una computadora.

En la Figura 2-3 se han puesto en la misma línea de hojas o rama todas las calificaciones puntajes que tienen las mismas decenas. Suponga que se reconstruye la representación y ahora en vez de agrupar diez valores posibles para cada tallo, se agrupan los valores de manera que sólo puedan estar cinco para cada tallo. ¿Se nota alguna diferencia en el aspecto de la Figura 2-4?. Es muy semejante la forma general, aproximadamente simétrica respecto al 70. Es típico que muchas variables muestren una distribución que se concentra (formando un montículo) alrededor de un valor central y se extiende a un lado y otro.



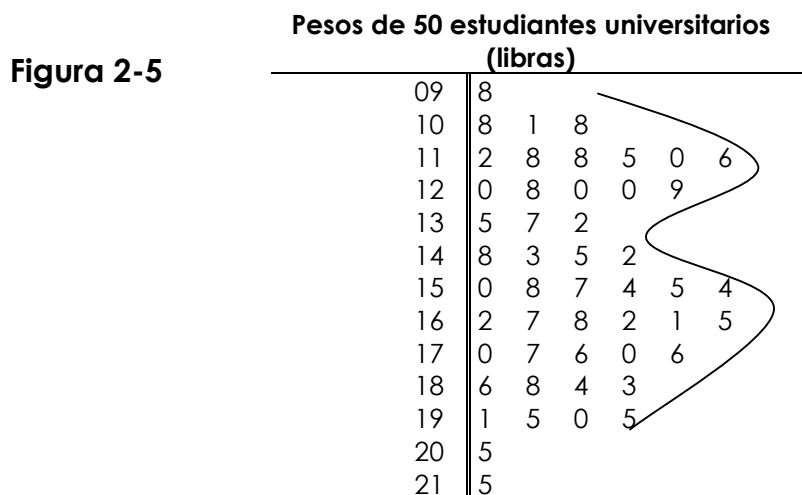
Una representación gráfica revela frecuentemente algo que el analista pudo o no haber previsto. La ilustración 2-2 ejemplifica lo que ocurre generalmente cuando se muestran dos poblaciones en forma conjunta.

Ilustración 2-2. Se seleccionó una muestra al azar de 50 estudiantes de una universidad. De sus registro médicos se obtuvieron sus pesos (en libras). Los datos obtenidos se presentan en la Tabla 2-2.

Tabla 2-2

| | | | | | | | | | | |
|--------------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Estudiante | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| Hombre/Mujer | M | H | M | H | H | M | M | H | H | M |
| Peso | 98 | 150 | 108 | 158 | 162 | 112 | 118 | 167 | 170 | 120 |
| Estudiante | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 |
| Hombre/Mujer | H | H | H | M | M | H | M | H | H | M |
| Peso | 177 | 186 | 191 | 128 | 135 | 195 | 137 | 205 | 190 | 120 |
| Estudiante | 21 | 22 | 23 | 24 | 25 | 26 | 27 | 28 | 29 | 30 |
| Hombre/Mujer | H | H | M | H | M | M | H | H | H | H |
| Peso | 188 | 176 | 118 | 168 | 115 | 115 | 162 | 157 | 154 | 148 |
| Estudiante | 31 | 32 | 33 | 34 | 35 | 36 | 37 | 38 | 39 | 40 |
| Hombre/Mujer | M | H | H | M | H | M | H | M | H | H |
| Peso | 101 | 143 | 145 | 108 | 155 | 110 | 154 | 116 | 161 | 165 |
| Estudiante | 41 | 42 | 43 | 44 | 45 | 46 | 47 | 48 | 49 | 50 |
| Hombre/Mujer | M | H | M | H | H | M | M | H | H | H |
| Peso | 142 | 184 | 120 | 170 | 195 | 132 | 129 | 215 | 176 | 183 |

Se debe advertir que los pesos varían entre 98 y 215 libras. Los valores de peso se agrupan para tallos de 10 unidades utilizando los dígitos de las centenas y decenas como tallos. Y los de las unidades como hojas (Figura 2-5).



Una inspección más detallada de la Figura 2-5 indica que puede haber dos distribuciones traslapadas. Esto es exactamente lo que se tiene: una distribución para los pesos de las alumnas y otra para los pesos de los alumnos. La Figura 2-6, la cual muestra una representación tallo-hoja de tipo adosado para ese conjunto de datos, evidencia que en realidad se tienen dos distribuciones distintas.

Figura 2-6

