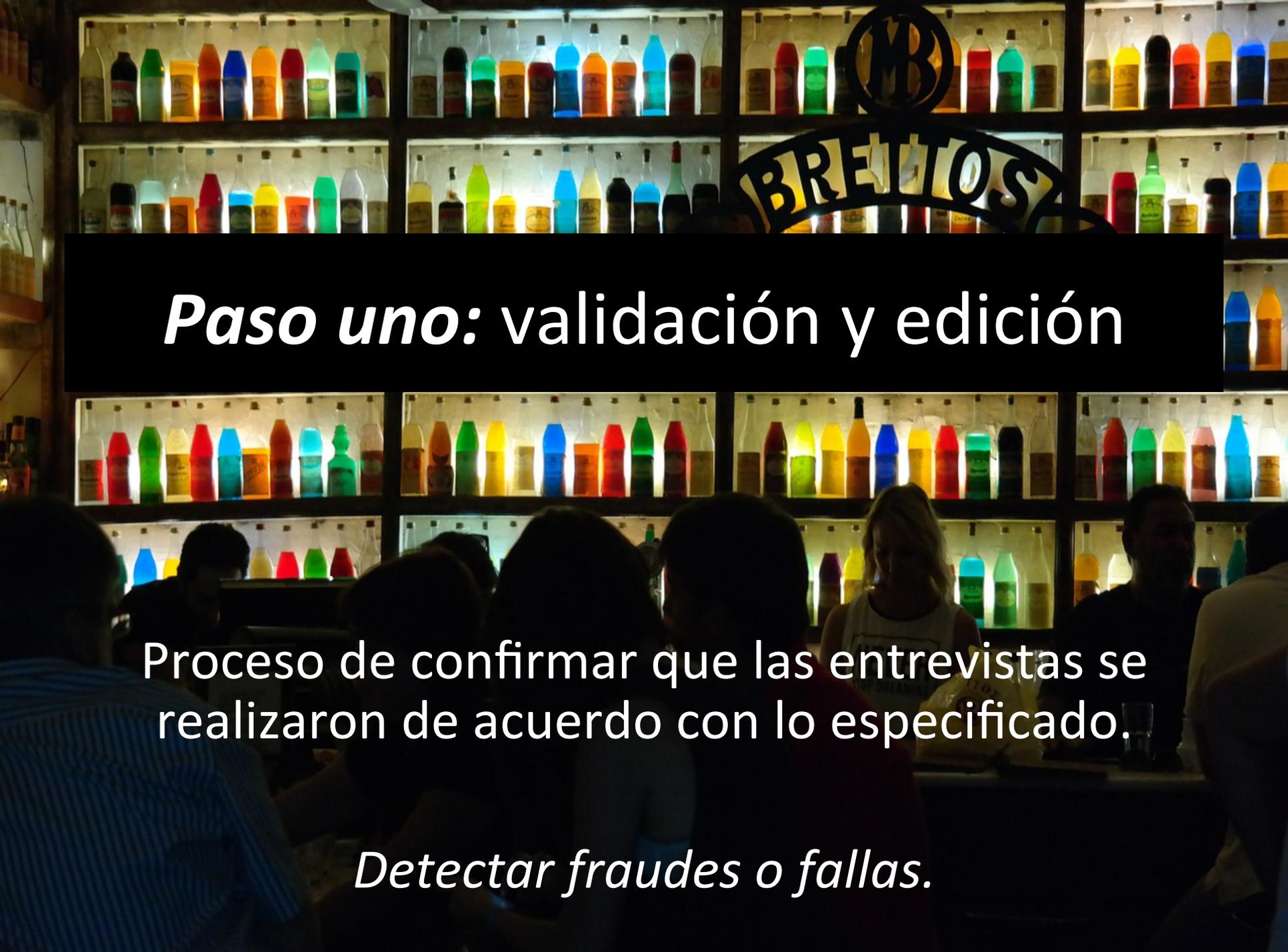




Procesamiento de datos y análisis fundamental de datos

Mtro. Carlos Estrada Zamora

Versión enero 2017

A photograph of a bar with a backlit display of colorful bottles. The bottles are arranged in rows and are illuminated from behind, creating a vibrant, multi-colored effect. A sign with the word "BRETTIOS" is visible in the upper right. The foreground shows the silhouettes of people sitting at the bar.

Paso uno: validación y edición

Proceso de confirmar que las entrevistas se realizaron de acuerdo con lo especificado.

Detectar fraudes o fallas.

La **edición** es el proceso de confirmar que los cuestionarios se llenaron adecuada y completamente.

Si el entrevistador dejó de hacer ciertas preguntas o de registrar las respuestas de ciertas preguntas.

Si se registraron los patrones de saltos.

Si el entrevistador parafraseó respuestas del encuestado a preguntas abiertas.

Paso dos: codificación

Proceso de agrupar y asignar códigos numéricos a las diversas respuestas a una pregunta.

El proceso de codificar respuestas a preguntas abiertas puede ser tedioso y tardado. Además el procedimiento es en cierto grado subjetivo. Por estas razones, los investigadores tienden a evitar preguntas abiertas a menos que sean absolutamente necesarias.

Proceso de codificación

Enlistar las respuestas.



Consolidar respuestas.



Establecer códigos.



Introducir códigos.

“ We predicted the result of Brexit a week before the vote, by combining our analytics and Lexalytics' sentiment analysis.

Nova Spivack, CEO & Co-Founder Bottlenose



Semantria.com ofrece varias herramientas y opciones de procesamiento de texto para procesar resultados de preguntas abiertas y alcanzar grados de síntesis difíciles de obtener.

Paso tres: entrada de datos

Proceso de convertir la información en un formato electrónico (que pueda ser leída por una computadora).

Entrada inteligente de datos

Forma de entrada de datos en la que la información que se introduce en el dispositivo de entrada de datos se verifica en cuanto a su lógica interna.

Proceso de entrada de datos

Introducir los datos a la base de datos de entrada.

Tecnología de escaneo

Forma de entrada de datos en la que respuestas en cuestionarios son leídas automáticamente por el dispositivo de entrada de datos.



Paso cuatro: depuración lógica de
datos

Última corrección computarizada de errores en
los datos.

Paso cinco: tabulación y análisis estadístico

Tablas unidireccionales de frecuencias

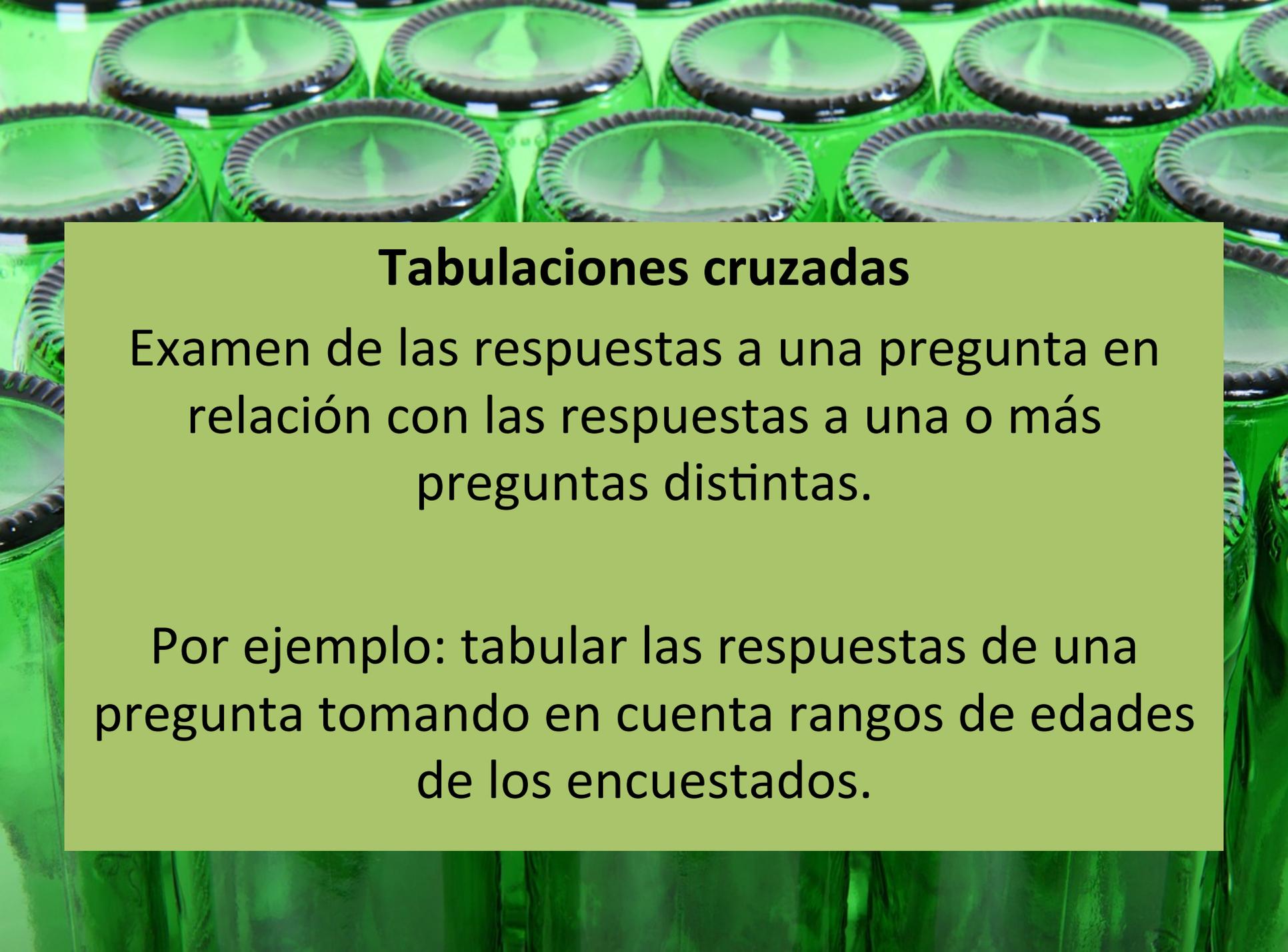
Tabla que muestra el número de encuestados que eligieron cada respuesta a una pregunta de encuesta.

Se pueden tomar como base:

Encuestados
totales.

Número de
personas a las que
se les hizo la
pregunta.

Número de
personas que
contestaron la
pregunta.

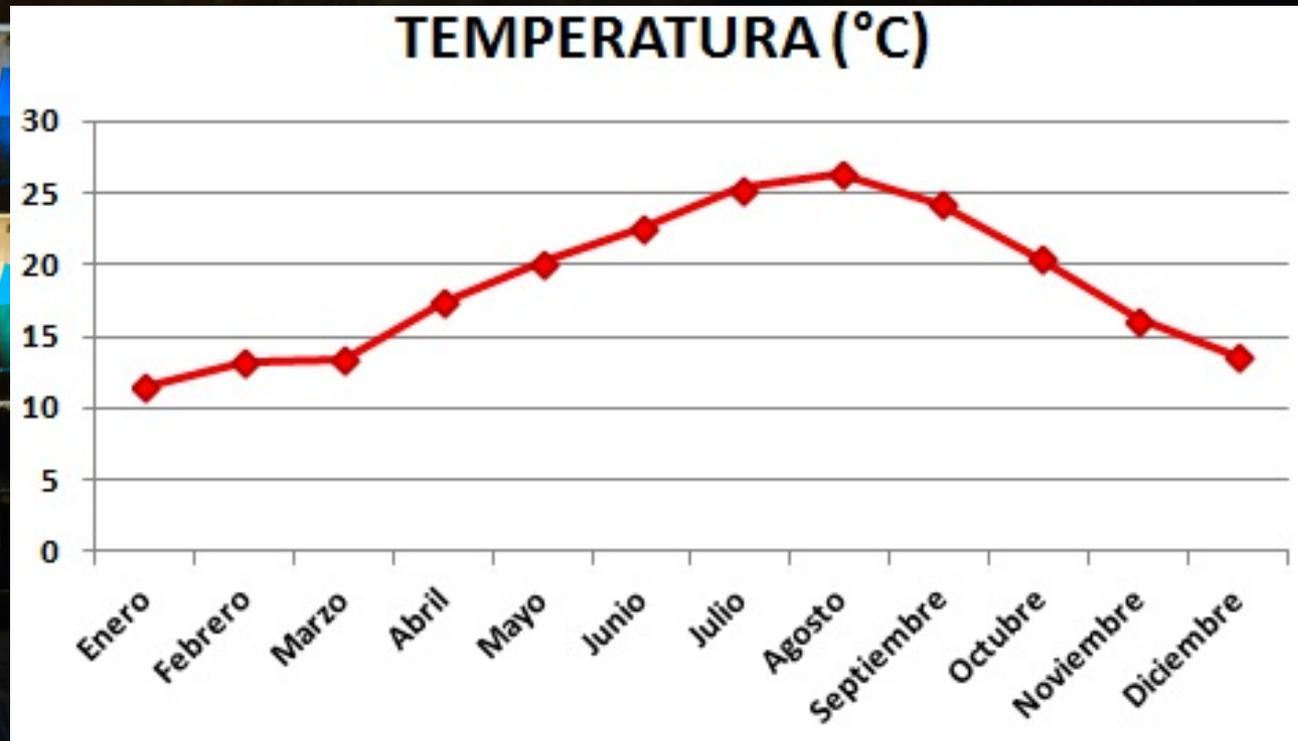


Tabulaciones cruzadas

Examen de las respuestas a una pregunta en relación con las respuestas a una o más preguntas distintas.

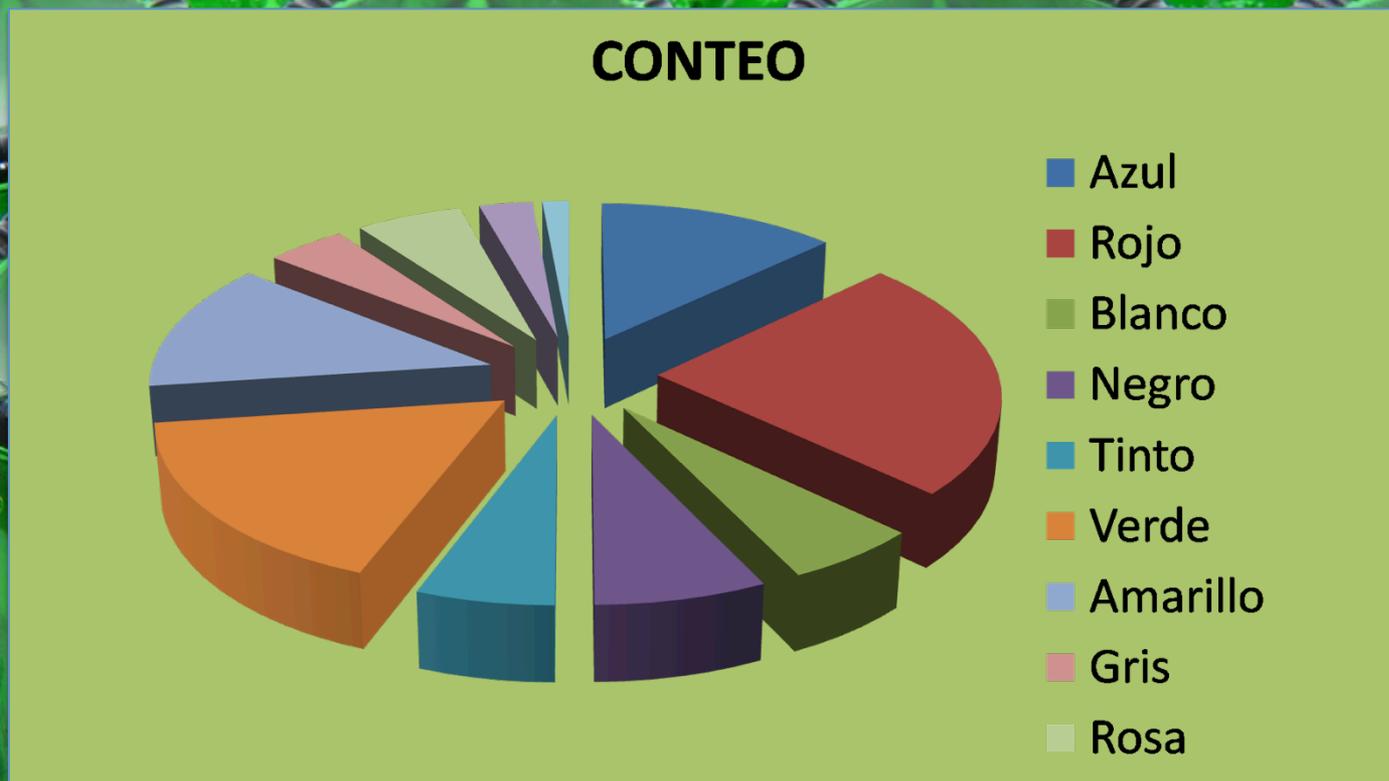
Por ejemplo: tabular las respuestas de una pregunta tomando en cuenta rangos de edades de los encuestados.

Representaciones gráficas de datos



Gráficas lineales

Representaciones gráficas de datos

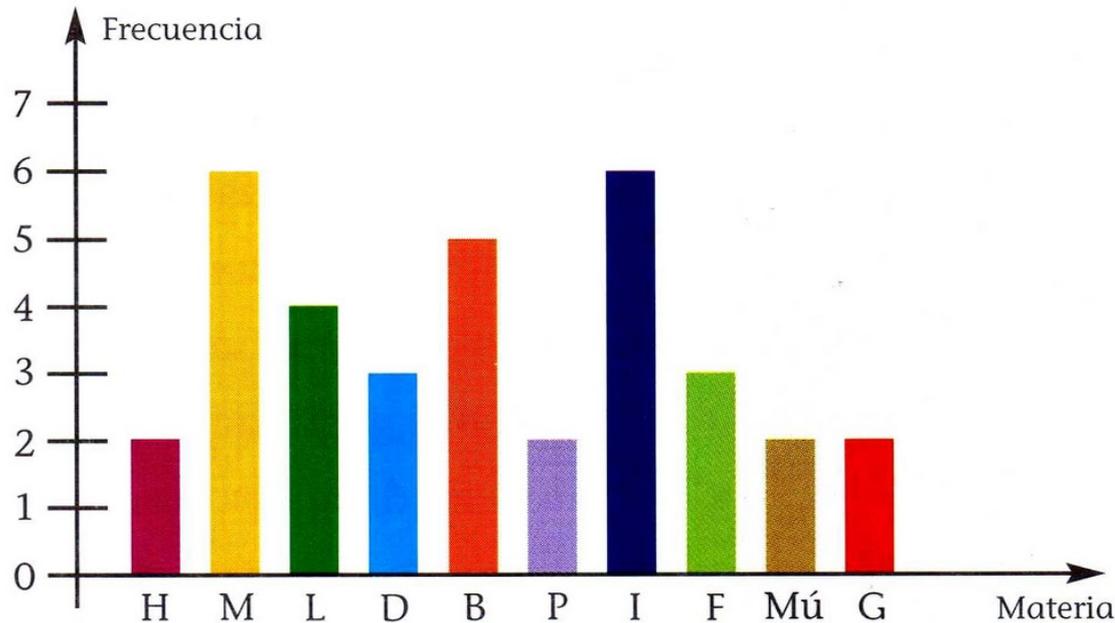


Gráficas circulares

Representaciones gráficas de datos

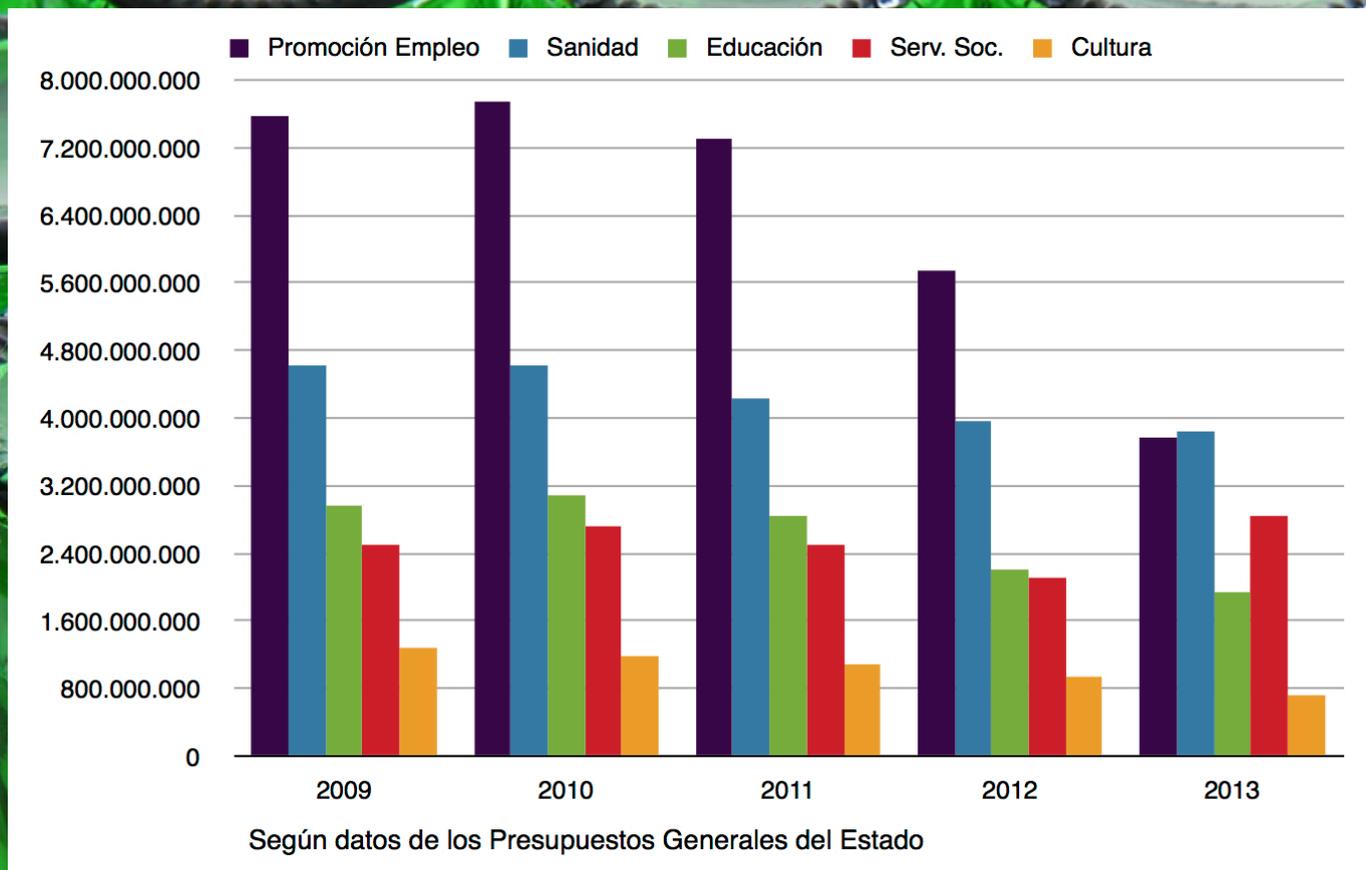


Este gráfico muestra la cantidad de alumnos de un curso que prefieren cada materia escolar.



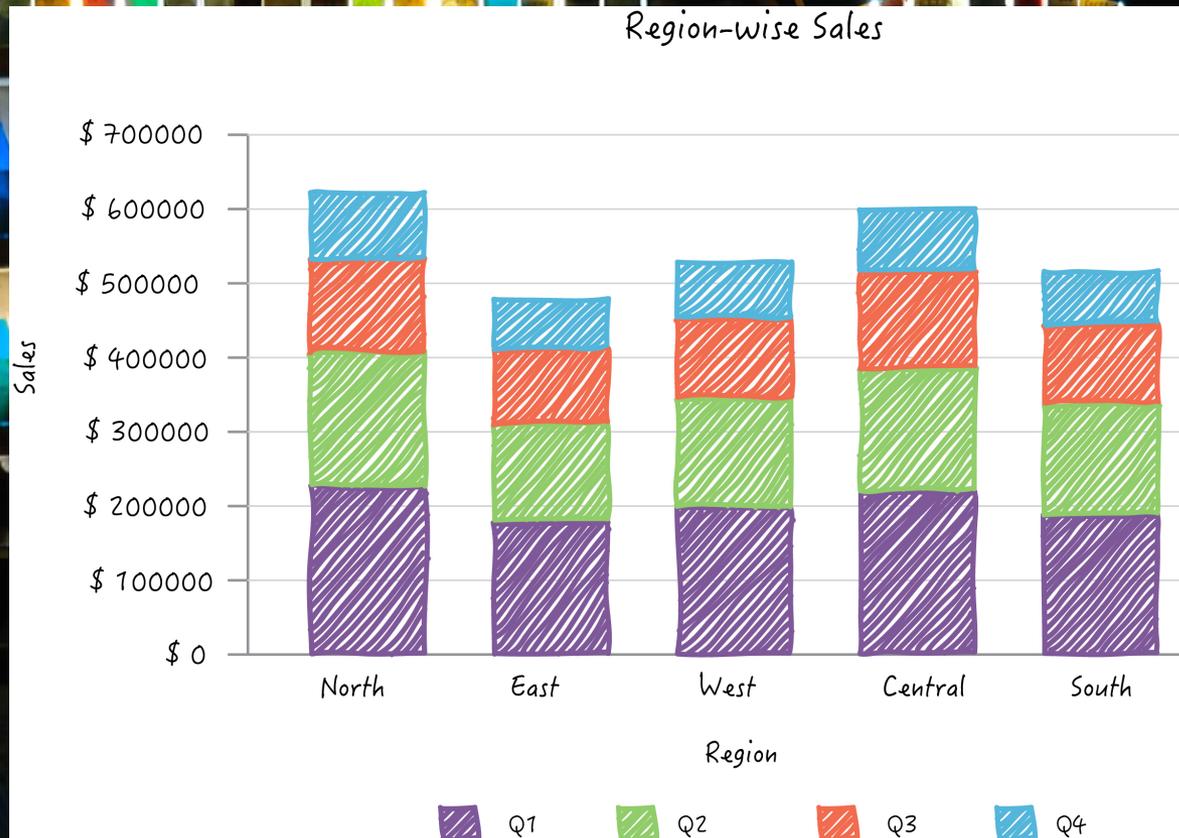
Gráfica de barras simple

Representaciones gráficas de datos



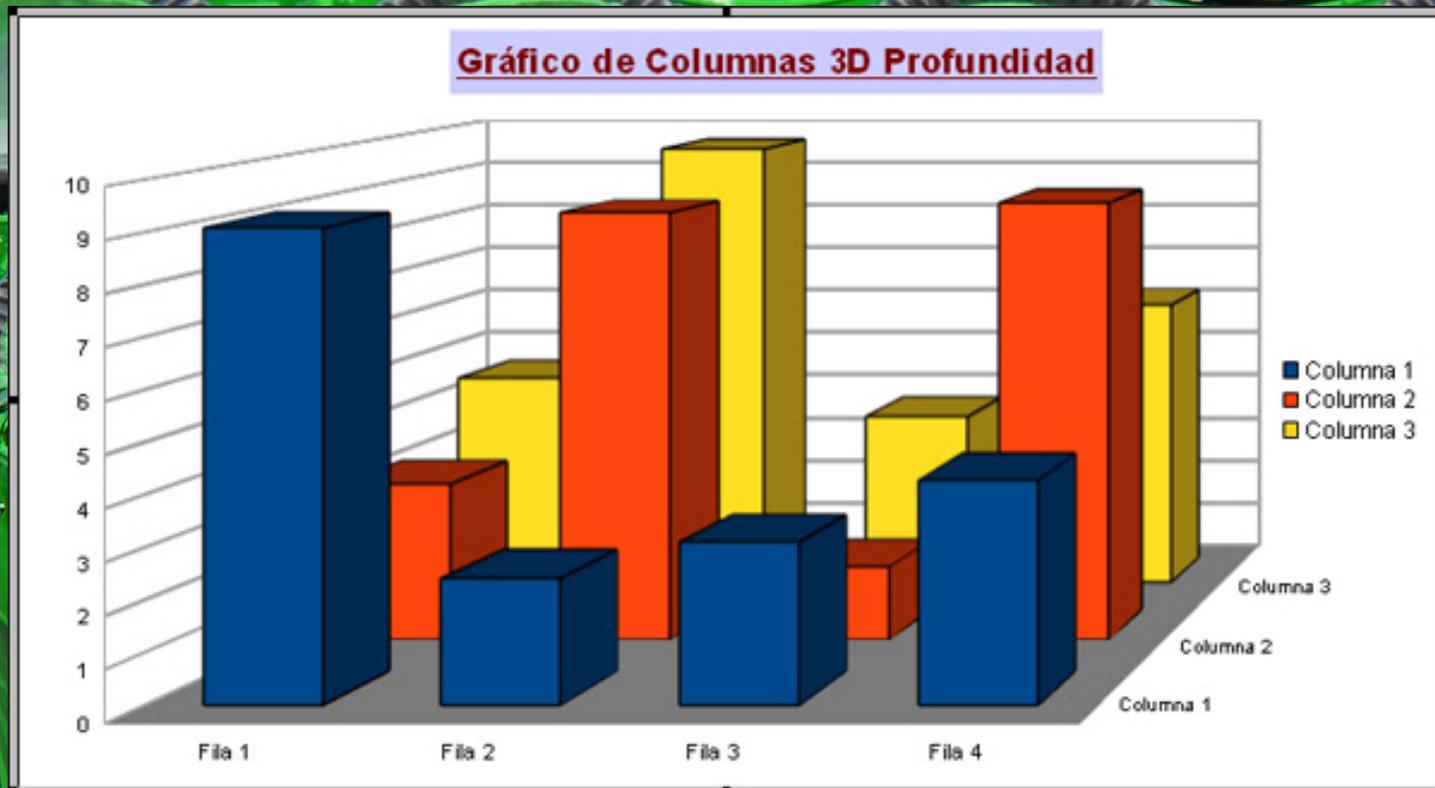
Gráfica de barras agrupada

Representaciones gráficas de datos



Gráfica de barras apilada

Representaciones gráficas de datos



**Gráfica de barras tridimensional
de filas múltiples**



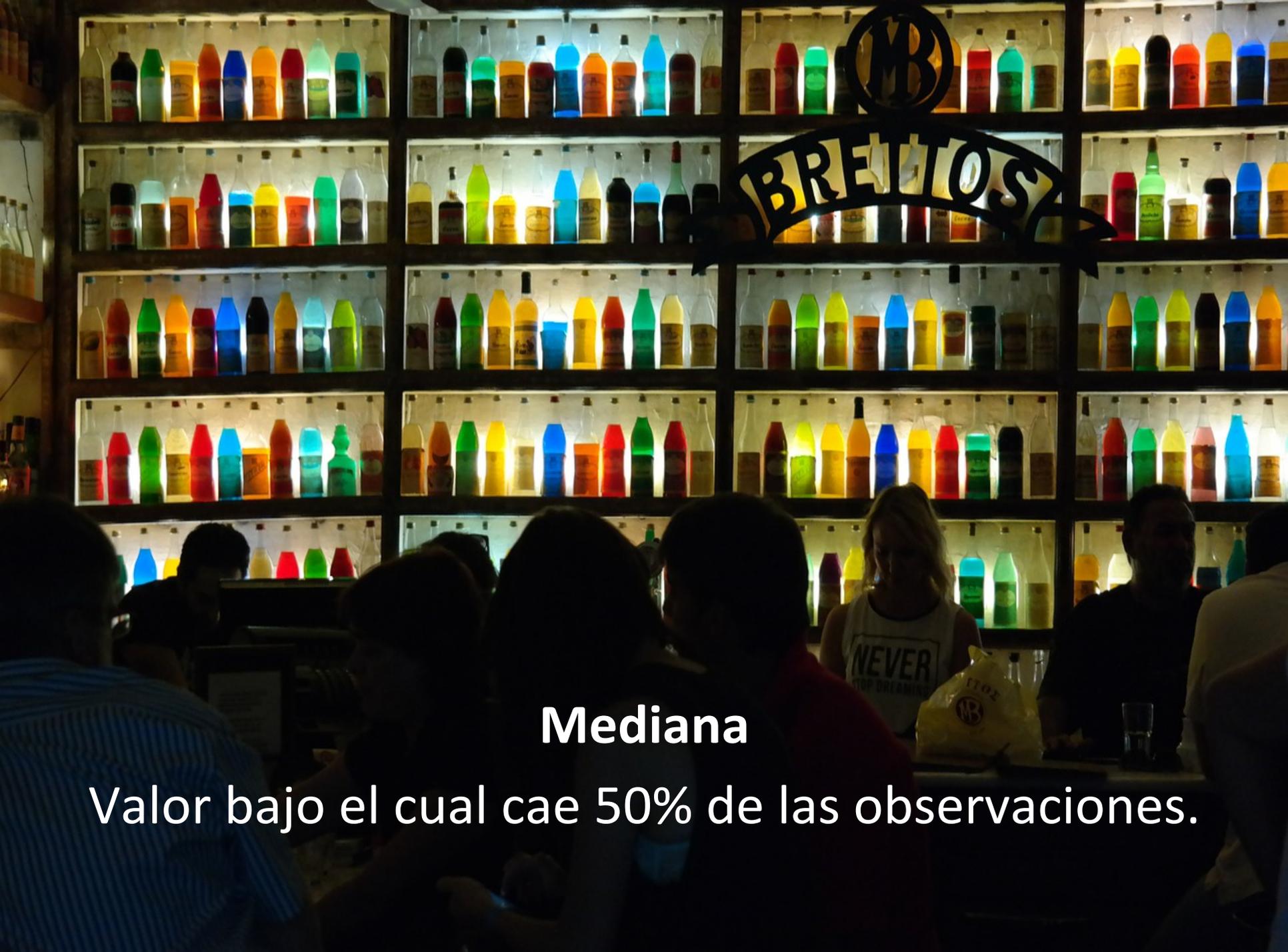
Estadística descriptiva

Medio eficiente de resumir las características básicas de grandes conjuntos de datos.

Medidas de tendencia central

Media

Suma de los valores de todas las observaciones de una variable dividida entre el número de observaciones.



Mediana

Valor bajo el cual cae 50% de las observaciones.



Moda

Valor que ocurre con más frecuencia.

FIGURA 15.20**Media, mediana y moda de datos de consumo de cerveza**

Un total de 10 bebedores de cerveza (que toman una o más latas, botellas o vasos de cerveza al día en promedio) fueron entrevistados en un estudio de intercepción en un centro comercial. Se les preguntó cuántas latas, botellas o vasos de cerveza beben en un día promedio.

Encuestado	Número de latas/botellas/ vasos por encuestado al día
1	2
2	2
3	3
4	2
5	5
6	1
7	2
8	2
9	10
10	1

Moda = 2 latas/botellas/vasos

Mediana = 2 latas/botellas/vasos

Media = 3 latas/botellas/vasos

Medidas de dispersión

Indican qué tan esparcidos están los datos.

Desviación estándar de la muestra

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

FIGURA 15.21**Medidas de dispersión y medidas de tendencia central**

Considérese los datos de consumo de cerveza presentados a continuación acerca de dos mercados.

Encuestados	Número de latas/ botellas/vasos Mercado uno	Número de latas/ botellas/vasos Mercado dos
1	2	1
2	2	1
3	3	1
4	2	1
5	5	1
6	1	1
7	2	1
8	2	3
9	10	10
10	1	10
Media	3	3
Desviación estándar	2.7	3.7

Aunque el consumo medio es igual en ambos mercados, la desviación estándar indica que hay mayor dispersión en el consumo en el mercado dos.

Contenido obtenido de:

INVESTIGACIÓN DE MERCADOS

McDaniel y Gates (2016)

Cengage Learning

Décima edición