Tablas de frecuencias relativas

Encontremos frecuencias relativas de datos categóricos.

2.1 Observa y pregúntate: Nivel de estudios de los profesores

A varios adultos que estaban en un edificio escolar les preguntaron cuál era el nivel de estudios más alto que habían completado y les preguntaron si eran profesores.

¿Qué observas? ¿Qué te preguntas?

	es profesor	no es profesor
técnico	4%	16%
profesional	52%	64%
maestría o superior	44%	20%

2.2 Gato de ciudad, gato de campo

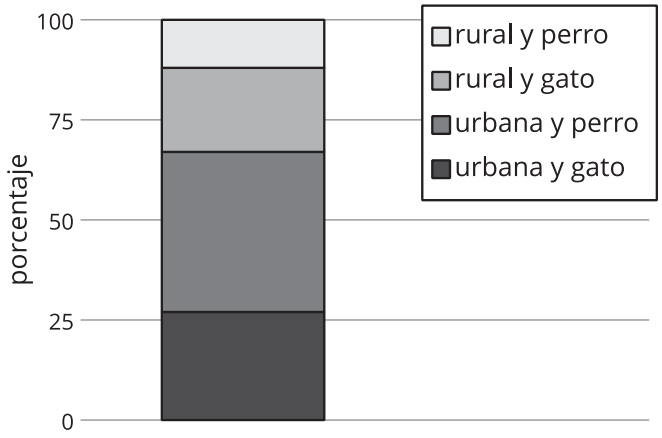
A 200 personas les preguntaron si preferían perros o gatos, y si vivían en un área rural o en un área urbana.

Los valores reales recolectados en la encuesta están en la primera tabla.

	urbana	rural	total
gato	54	42	96
perro	80	24	104
total	134	66	200

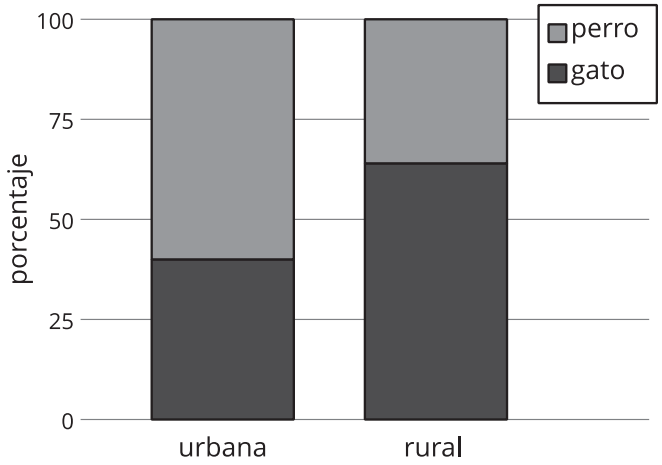
La siguiente tabla muestra qué porcentaje del total de 200 personas están representadas por cada combinación de categorías. El gráfico de barras segmentadas representa la misma información de forma gráfica.

	urbana	rural
gato	27%	21%
perro	40%	12%



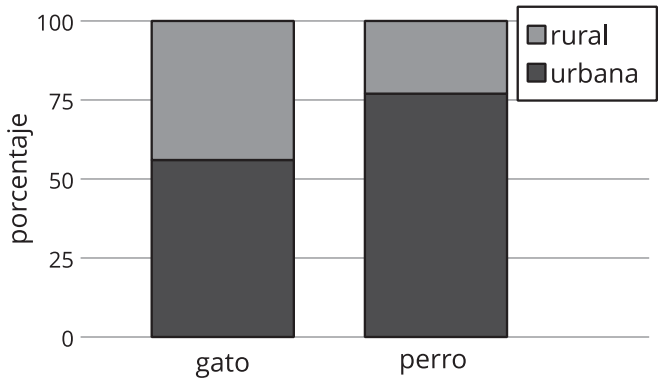
La siguiente es una tabla de frecuencias relativas por columna y muestra el porcentaje de personas representadas por cada columna que tienen cierta preferencia por una mascota. El gráfico de barras segmentadas representa la misma información de forma gráfica.

	urbana	rural
gato	40%	64%
perro	60%	36%



La última es una tabla de frecuencias relativas por fila y muestra el porcentaje de personas representadas por cada fila que viven en cierta área. El gráfico de barras segmentadas representa la misma información de forma gráfica.

	urbana	rural
gato	56%	44%
perro	77%	23%



1. En cada tabla de frecuencias relativas, selecciona un porcentaje y explica cómo se usaron los números de la tabla original para obtener ese porcentaje.

2. ¿Qué porcentaje de las personas encuestadas vive en un área urbana y prefiere perros?

3. De las personas encuestadas que prefieren perros, ¿qué porcentaje vive en un área urbana?
4. De las personas encuestadas que viven en un área urbana, ¿qué porcentaje prefiere perros?

5. ¿Cuántas de las personas respondieron que prefieren perros y que viven en un área urbana?

6. Entre las personas encuestadas, ¿hay más personas que prefieren perros o más personas que prefieren gatos?

7. Tu empresa de comida para mascotas tiene acceso a una valla publicitaria en un área rural. ¿Recomendarías que en esta valla publicitaria se promocionara comida para perros o comida para gatos? ¿Cuál tabla usaste para tomar esta decisión? Explica tu razonamiento.

2.3 Analicemos un estudio con dos tratamientos

Un experimento busca comprobar la eficacia de la vitamina C para reducir la duración de los resfriados. En el experimento, a dos grupos de personas con resfriado se les da una pastilla que deben tomar una vez al día. Un grupo toma una pastilla que contiene 1,000 mg de vitamina C, mientras que el otro grupo toma una pastilla de efecto placebo. Los investigadores registran los resultados en una tabla.

	grupo A	grupo B
el resfriado dura menos de una semana	16	27
el resfriado dura una semana o más	17	53

1. Primero, los investigadores quieren saber qué porcentaje de personas (al redondear al porcentaje entero más cercano) está en cada combinación de categorías. El 14% de todos los participantes tuvo un resfriado de menos de una semana y estaba en el grupo A. ¿Qué porcentaje de todos los participantes tuvo un resfriado de menos de una semana y estaba en el grupo B? Completa el resto de la tabla de frecuencias relativas con los porcentajes correspondientes.

	grupo A	grupo B
el resfriado dura menos de una semana	14% ($\frac{16}{16+27+17+53} \approx 0.14$)	
el resfriado dura una semana o más		

2. Después, los investigadores observaron que de todos los participantes que tuvieron resfriados de menos de una semana, el 37% estaba en el grupo A. De todos los participantes que tuvieron resfriados de una semana o más, ¿qué porcentaje estaba en el grupo B? Completa la tabla con los porcentajes correspondientes.

	grupo A	grupo B
el resfriado dura menos de una semana	37% ($\frac{16}{16+27} \approx 0.37$)	
el resfriado dura una semana o más		

3. Por último, los investigadores observaron que de todos los participantes del grupo A, el 48% tuvo resfriados de menos de una semana. De todos los participantes del grupo B, ¿cuántos tuvieron resfriados de una semana o más? Completa la tabla con los porcentajes correspondientes.

	grupo A	grupo B
el resfriado dura menos de una semana	$48\%\ (\frac{16}{16+17} \approx 0.48)$	
el resfriado dura una semana o más		

4. Para entender los resultados, los investigadores quieren saber lo siguiente: de todas las personas que tuvieron resfriados de menos de una semana, ¿qué porcentaje está en el grupo B? Responde y explica tu razonamiento.

5. Si los investigadores creen que la vitamina C tiene un leve efecto en la duración de un resfriado, ¿cuál grupo es más probable que haya recibido las pastillas que contienen vitamina C? Explica tu razonamiento.

¿Estás listo para más?

Un profesor encuestó a un grupo de 25 estudiantes de octavo grado y a un grupo de 20 estudiantes de doceavo grado que dijeron que conocían algún lenguaje de programación. Python y Scratch son lenguajes de programación. Los resultados de la encuesta de octavo grado se muestran en la siguiente tabla de doble entrada.

	Python es el lenguaje que mejor conozco	Scratch es el lenguaje que mejor conozco	Conozco mejor otro lenguaje de programación
Me enseñaron un lenguaje de programación en la escuela	8	6	1
No me enseñaron un lenguaje de programación en la escuela	1	7	2

Los resultados de la encuesta de doceavo grado se muestran en esta tabla de doble entrada.

	Python es el lenguaje que mejor conozco	Scratch es el lenguaje que mejor conozco	Conozco mejor otro lenguaje de programación
Me enseñaron un lenguaje de programación en la escuela	25%	35%	0%
No me enseñaron un lenguaje de programación en la escuela	30%	5%	5%

1. ¿Cuál lenguaje de programación conoce mejor la mayoría de los estudiantes encuestados de octavo grado?
2. ¿Cuál lenguaje de programación conoce mejor la mayoría de los estudiantes encuestados de doceavo grado?
3. ¿Cuántos de los estudiantes encuestados de doceavo grado dicen que les enseñaron un lenguaje de programación en la escuela?
4. ¿Qué porcentaje de los estudiantes encuestados de octavo grado dice que Python es el lenguaje de programación que mejor conoce y que no le enseñaron un lenguaje de programación en la escuela?
5. Según como está dada la información en las tablas, ¿por qué es difícil decidir si los estudiantes de doceavo grado o los estudiantes de octavo grado usan más Python?

Resumen de la lección 2

Convertir tablas de doble entrada en *tablas de frecuencias relativas* puede ayudarnos a mostrar patrones en variables categóricas que están emparejadas. Para crear una **tabla de frecuencias relativas**, se divide el valor de cada celda de una tabla de doble entrada entre el número total de valores de toda la tabla, o entre el número total de valores de una fila o de una columna. Se pueden usar distintos tipos de tablas de frecuencias relativas dependiendo de qué patrones nos interesen. Si se quiere analizar cómo se relaciona cada combinación de variables categóricas con el grupo entero, se divide cada valor de la tabla de doble entrada entre el número total de valores de toda la tabla. Así se encuentra la frecuencia relativa.

Por ejemplo, esta tabla de doble entrada muestra el estado y el precio de cada libro de los 120 libros que hay en una librería universitaria.

	$10 o menos	más de $10 pero menos de $30	$30 o más
nuevo	3	9	27
usado	33	36	12

Para crear una tabla de doble entrada de frecuencias relativas, dividimos cada número de la tabla de doble entrada entre 120, porque hay 120 valores (3+9+27+33+36+12) en este conjunto de datos. La tabla de doble entrada de frecuencias relativas que resulta se puede representar usando fracciones, decimales o porcentajes.

	$10 o menos	más de $10 pero menos de $30	$30 o más
nuevo	0.025	0.075	0.225
usado	0.275	0.300	0.100

Esta tabla de doble entrada de frecuencias relativas nos permite ver qué proporción del total representa cada número de la tabla de doble entrada. El número 33 en la tabla de doble entrada original representa el número de libros que son usados y que también cuestan $10 o menos, que es el 27.5% de todos los libros del conjunto de datos. Esta tabla de doble entrada de frecuencias relativas nos permite ver que hay muy pocos libros que son nuevos y que también son baratos (2.5%), y que el 10% de los libros que se venden son caros y usados.

En otras situaciones, tiene sentido analizar las proporciones por fila o por columna en una tabla de frecuencias relativas. Por ejemplo, para convertir la tabla de doble entrada original en una tabla de frecuencias relativas por columna usando las proporciones de la columna, cada valor se divide entre la suma de los valores de su columna.

	$10 o menos	**más de $10 pero menos de $30**	**$30 o más**
nuevo	0.083	0.2	0.692
usado	0.917	0.8	0.308

Esto muestra que aproximadamente el 91.7% ($\frac{33}{3+33} \approx 0.917$) de los libros que se venden por $10 o menos son usados. Observa que en esta tabla de frecuencias relativas por columna, cada columna muestra las proporciones de los libros nuevos o usados para el rango de precio correspondiente. Observa además que la suma de las frecuencias relativas de cada columna es igual a 1. En particular, esto muestra que la mayoría de los libros baratos y de precio moderado son usados, y la mayoría de los libros caros son nuevos.