https://doi.org/10.62486/978-628-97230-1-4

Capítulo 5

https://doi.org/10.62486/978-628-97230-1-4.ch05

Ponerlo en práctica: ejercicios resueltos con datos agrupados

Hasta aquí hemos revisado cómo se organiza y analiza la información cuando se trabaja con datos agrupados. Sabemos que este tipo de datos se presenta en forma de intervalos, que usamos marcas de clase en lugar de valores individuales, y que los cálculos deben adaptarse para reflejar esa forma resumida de los datos. Ahora es momento de aplicar esas ideas a situaciones reales.

Vamos a resolver un ejercicio completo, tomado directamente de un contexto real. Este ejemplo te permitirá observar cómo se construye una tabla de frecuencias, cómo se utilizan las fórmulas adecuadas y, sobre todo, cómo se interpretan los resultados.

Ejercicio 1: Emisiones de hidrocarburos en automóviles modelo 2000

En una muestra de 20 automóviles del modelo 2000, se midieron las emisiones de hidrocarburos en reposo, expresadas en partes por millón (ppm). Los valores obtenidos fueron los siguientes:

141, 359, 247, 940, 882, 494, 306, 210, 105, 880,

200, 223, 188, 940, 241, 190, 300, 435, 241, 380

Dado que estos datos están dispersos, el primer paso es agruparlos en intervalos de clase. Para hacerlo, seguimos estos pasos básicos:

a) Identificamos el valor mínimo (105) y el máximo (940).
b) Calculamos el rango: Rango = 940 – 105 = 835
c) Elegimos un número adecuado de clases, por ejemplo 6.

d) Calculamos la amplitud de clase (A): A ≈ R / k = 835 / 6 ≈ 139.2 → redondeamos a 140.

Con esto, proponemos los siguientes intervalos:

- 100 – 239
- 240 – 379
- 380 – 519
- 520 – 659
- 660 – 799
- 800 – 939
- 940 – 1079

Ya que el valor 940 aparece dos veces, agregamos una clase más para incluirlo.

Tabla 1

Construcción de frecuencia

Intervalo	Marca de clase (xi)	Frecuencia (fi)
100 – 239	169.5	6
240 – 379	309.5	6
380 – 519	449.5	3
520 – 659	589.5	0
660 – 799	729.5	0
800 – 939	869.5	2
940 – 1079	1009.5	3

Total de observaciones: ∑fi = 20

Cálculos principales

Media (μ):

Usamos la fórmula:

$$\mu = (\Sigma fi * xi) / \Sigma fi$$

$\mu = (6\times169.5 + 6\times309.5 + 3\times449.5 + 2\times869.5 + 3\times1009.5) / 20$

$\mu \approx (1017 + 1857 + 1348.5 + 1739 + 3028.5) / 20$

$\mu \approx 8989 / 20 = 449.45$

La media indica que, en promedio, los vehículos de esta muestra emiten cerca de 449 ppm de hidrocarburos. Sin embargo, debemos observar que hay valores extremos muy altos (880, 940) que elevan esta media. ¿Está distorsionada? Para saberlo, veamos otras medidas.

Mediana (Me): Para encontrar la mediana, buscamos el intervalo donde se encuentra el valor de posición n/2 = 20 / 2 = 10.

Acumulando las frecuencias:

- Intervalo 1 (100–239): 6
- Intervalo 2 (240–379): 6 + 6 = 12 → contiene la mediana.

Aplicamos la fórmula:

$Me = Li + [(n/2 - F<) / fi] \times A$

$Me = 240 + [(10 - 6) / 6] \times 140$

$Me = 240 + (4 / 6) \times 140 = 240 + 93.33 = 333.33$

La mediana es mucho menor que la media. Esto confirma que la distribución está sesgada positivamente, es decir, hacia la derecha, lo cual ocurre cuando hay valores atípicos altos.

Moda (Mo): La moda se encuentra en el intervalo con mayor frecuencia. En este caso, los dos primeros intervalos tienen frecuencia 6, pero usaremos el segundo (240–379) como clase modal por su posición más central.

$$Mo = Lmod + [(fmod - f_{-1}) / ((fmod - f_{-1}) + (fmod - f_{+1}))] \times A$$

$$Mo = 240 + [(6 - 6) / ((6 - 6) + (6 - 3))] \times 140$$

$$Mo = 240 + [0 / (0 + 3)] \times 140 = 240$$

El cálculo nos da como moda el límite inferior de la clase modal. En este caso, esto indica que la mayoría de los vehículos se concentra en valores cercanos al límite inferior del intervalo.

Rango y dispersión:

- Rango: 940 – 105 = 835 ppm
- Varianza y desviación estándar requieren aplicar la fórmula: $\sigma^2 = (\Sigma fi * xi^2)/n - \mu^2$

Interpretación general

Los datos analizados muestran que, si bien la media está cerca de 450 ppm, tanto la mediana como la moda están por debajo de ese valor, lo que revela una asimetría positiva: unos pocos vehículos con emisiones muy altas (como los que están por encima de 880 ppm) alteran el promedio.

La mayoría de los vehículos se agrupa entre los 240 y 380 ppm, y los valores extremos, aunque importantes, son excepcionales. Este tipo de análisis es útil, por ejemplo, en estudios ambientales, donde una lectura incorrecta podría generar una falsa alarma o, por el contrario, pasar por alto un caso grave.

¿Qué sigue?

En el próximo ejercicio, veremos un nuevo conjunto de datos agrupados, como los tiempos de secado de pintura esmaltada, para reforzar los procedimientos y ampliar tu capacidad de análisis.

Pero por ahora, si has llegado hasta aquí, ya dominas los elementos esenciales para interpretar datos agrupados con confianza.

Ejercicio 2: Tiempos de secado de pintura esmaltada

Imaginemos que una empresa de recubrimientos industriales necesita evaluar el rendimiento de una pintura esmaltada. Para ello, realiza pruebas de laboratorio con distintas muestras del producto, registrando el tiempo de secado en horas.

Se obtienen los siguientes 40 datos:

2.2, 4.1, 3.5, 4.5, 3.2, 3.7, 3.0, 2.6, 3.4, 1.6,

3.1, 3.3, 3.8, 3.1, 4.7, 3.7, 2.5, 4.3, 3.4, 3.6,

2.9, 3.3, 3.9, 3.1, 3.3, 3.1, 3.7, 4.4, 3.2, 4.1,

1.9, 3.4, 4.7, 3.8, 3.2, 2.6, 3.9, 3.0, 4.2, 3.5

Paso 1: Determinar los intervalos

Primero, identificamos el valor mínimo (1.6) y el valor máximo (4.7).

Calculamos el rango:

R = 4.7 – 1.6 = 3.1

Elegimos usar 5 clases → Amplitud ≈ R / k = 3.1 / 5 = 0.62 → Redondeamos a 0.7

Proponemos los siguientes intervalos (sin traslape):

- 1.6 – 2.3
- 2.3 – 3.0
- 3.0 – 3.7
- 3.7 – 4.4
- 4.4 – 5.1

Paso 2: Construcción de la tabla de frecuencia

Contamos cuántos datos caen en cada intervalo:

Intervalo	xi (marca de clase)	fi
1.6 – 2.3	1.95	2
2.3 – 3.0	2.65	5
3.0 – 3.7	3.35	18
3.7 – 4.4	4.05	12
4.4 – 5.1	4.75	3
Total	—	40

Paso 3: Calcular la media

Media = (Σxi * fi) / n

$\mu = (1.95\times2 + 2.65\times5 + 3.35\times18 + 4.05\times12 + 4.75\times3) / 40$

$\mu = (3.9 + 13.25 + 60.3 + 48.6 + 14.25) / 40$

$\mu = 140.3 / 40 = 3.51$ horas

Paso 4: Mediana

n = 40 → n/2 = 20

Sumando frecuencias:

- Intervalo 1: 2
- Intervalo 2: 2 + 5 = 7
- Intervalo 3: 7 + 18 = 25 → contiene la mediana

Usamos la fórmula:

$Me = Li + [(n/2 – F<) / fi] \times A$

$Me = 3.0 + [(20 – 7) / 18] \times 0.7$

$Me = 3.0 + (13 / 18) \times 0.7$

$Me \approx 3.0 + 0.505 = 3.51$ horas

La mediana coincide con la media, lo que sugiere una distribución bastante simétrica.

Paso 5: Moda

La clase con mayor frecuencia es 3.0 – 3.7 (fi = 18) → clase modal

Usamos la fórmula:

$Mo = Li + [(fmod – f–1) / ((fmod – f–1) + (fmod – f+1))] \times A$

$Mo = 3.0 + [(18 – 5) / ((18 – 5) + (18 – 12))] \times 0.7$

$Mo = 3.0 + (13 / 19) \times 0.7$

$Mo \approx 3.0 + 0.479 = 3.48$ horas

Paso 6: Interpretación

Los tres valores clave —media, mediana y moda— son muy similares (entre 3.48 y 3.51). Esto indica que la distribución es simétrica, sin valores extremos o atípicos significativos. La mayoría de las pinturas seca entre 3.0 y 3.7 horas, con muy pocos casos más lentos o rápidos. Esto es un buen indicador de consistencia en el producto.

Este ejercicio muestra cómo los datos agrupados pueden ofrecer una visión clara y rápida del comportamiento de un fenómeno. Con una tabla bien construida y fórmulas aplicadas correctamente, es posible llegar a conclusiones útiles incluso sin tener todos los datos individuales.

Ejercicio 3: Análisis de nidos de tortugas marinas

En una playa del Caribe, un grupo de biólogos está estudiando el comportamiento reproductivo de las tortugas marinas. Como parte de su investigación, se seleccionó una muestra de 25 nidos y se contó el número de huevos depositados en cada uno. Los datos recolectados fueron los siguientes:

40, 37, 60, 10, 30, 45, 55, 27, 40, 70, 30, 50, 35, 40, 60, 80, 50, 60, 65, 50,

55, 40, 35, 48, 50

Paso 1: Determinar el rango y clases

Mínimo: 10

Máximo: 80

Rango: 80 – 10 = 70

Elegimos 6 clases → Amplitud ≈ 70 / 6 = 11.6 → redondeamos a 12

Paso 2: Construcción de intervalos y tabla de frecuencias

Proponemos los siguientes intervalos de clase:

- 10 – 21.9
- 22 – 33.9
- 34 – 45.9
- 46 – 57.9
- 58 – 69.9
- 70 – 81.9

Agrupamos los datos dentro de esos rangos y contamos frecuencias:

Intervalo	xi (marca de clase)	fi
10 – 21.9	16.0	1
22 – 33.9	28.0	2
34 – 45.9	40.0	6
46 – 57.9	52.0	8
58 – 69.9	64.0	5
70 – 81.9	76.0	3
Total	—	25

Paso 3: Calcular la media

Media = (Σxi × fi) / n

$\mu = (16\times1 + 28\times2 + 40\times6 + 52\times8 + 64\times5 + 76\times3) / 25$

$\mu = (16 + 56 + 240 + 416 + 320 + 228) / 25$

$\mu = 1276 / 25 = 51.04$ huevos por nido

https://doi.org/10.62486/978-628-97230-1-4

Paso 4: Calcular la mediana

$n = 25 \rightarrow n/2 = 12.5$

Acumulamos frecuencias:

- 1 (clase 1)
- 1 + 2 = 3 (clase 2)
- 3 + 6 = 9 (clase 3)
- 9 + 8 = 17 → la mediana está en la clase 4 (46–57.9)

Aplicamos la fórmula:

$Me = Li + [(n/2 - F<) / fi] \times A$

$Me = 46 + [(12.5 - 9) / 8] \times 12$

$Me = 46 + (3.5 / 8) \times 12 = 46 + 5.25 = 51.25$ huevos

Paso 5: Calcular la moda

La clase con mayor frecuencia es 46–57.9 (fi = 8), por lo tanto, es la clase modal.

$Mo = Li + [(fmod - f{-1}) / ((fmod - f{-1}) + (fmod - f{+1}))] \times A$

$Mo = 46 + [(8 - 6) / ((8 - 6) + (8 - 5))] \times 12$

$Mo = 46 + (2 / (2 + 3)) \times 12 = 46 + (2/5) \times 12 = 46 + 4.8 = 50.8$ huevos

Paso 6: Análisis

La media (51.04), la mediana (51.25) y la moda (50.8) son prácticamente iguales. Esta coincidencia indica una distribución simétrica, donde no hay sesgo hacia la

https://doi.org/10.62486/978-628-97230-1-4

derecha ni hacia la izquierda. En este caso, la mayoría de las tortugas deposita entre 46 y 58 huevos por nido, lo cual parece ser un comportamiento reproductivo común en esta población.

Además, las clases más altas (60–80) tienen valores frecuentes, lo que indica que algunas tortugas superan con facilidad ese promedio. Es decir, si bien hay uniformidad, también hay un grupo de ejemplares más prolíficos.

Este análisis ayuda a los investigadores a entender los patrones reproductivos de las tortugas marinas. Saber cuántos huevos deposita cada una en promedio es crucial para estimar la cantidad de crías esperadas, planificar acciones de conservación y evaluar la salud de la especie.

La estadística, aplicada con claridad y orden, permite extraer información valiosa de lo que a primera vista parece solo una lista de números.

Ejercicio 4: Análisis del peso de estudiantes universitarios

En una universidad se recolectaron los pesos, redondeados a la libra más cercana, de 40 estudiantes. El objetivo era conocer la distribución general del peso en un grupo aleatorio, para utilizar los resultados como parte de un estudio sobre salud y bienestar estudiantil.

Los datos fueron los siguientes:

138, 164, 150, 132, 144, 125, 149, 157, 146, 158,

140, 147, 136, 148, 152, 144, 168, 126, 138, 176,

163, 119, 154, 165, 146, 173, 142, 147, 135, 153,

140, 135, 161, 145, 135, 142, 150, 156, 145, 128

Paso 1: Determinar rango y clases

- Mínimo: 119
- Máximo: 176
- Rango: 176 – 119 = 57

Elegimos 6 clases → Amplitud ≈ 57 / 6 ≈ 9.5 → redondeamos a 10

Paso 2: Definimos intervalos y frecuencia

Intervalos propuestos:

- 115 – 124
- 125 – 134
- 135 – 144
- 145 – 154
- 155 – 164
- 165 – 174
- 175 – 184 (solo para incluir el 176)

Contamos los datos dentro de cada intervalo:

Intervalo	xi (marca de clase)	fi
115 – 124	119.5	1
125 – 134	129.5	3
135 – 144	139.5	13
145 – 154	149.5	13
155 – 164	159.5	6

165 – 174	169.5	3
175 – 184	179.5	1
Total	—	40

Paso 3: Calcular la media

Media = (Σxi × fi) / n

μ = (119.5×1 + 129.5×3 + 139.5×13 + 149.5×13 + 159.5×6 + 169.5×3 + 179.5×1) / 40

μ = (119.5 + 388.5 + 1813.5 + 1943.5 + 957 + 508.5 + 179.5) / 40

μ = 5910 / 40 = 147.75 libras

Paso 4: Calcular la mediana

n = 40 → n/2 = 20

Acumulando frecuencias:

- 1 (intervalo 1)
- 1 + 3 = 4 (intervalo 2)
- 4 + 13 = 17 (intervalo 3)
- 17 + 13 = 30 → la mediana está en el intervalo 145–154

Aplicamos la fórmula:

Me = Li + [(n/2 – F<) / fi] × A

Me = 145 + [(20 – 17) / 13] × 10

Me = 145 + (3 / 13) × 10 ≈ 145 + 2.31 = 147.31 libras

Paso 5: Calcular la moda

La clase modal es compartida entre dos clases con frecuencia 13: 135–144 y 145–154. Elegimos la segunda (más alta en el rango) por estar más cercana a la media.

$Mo = Li + [(fmod - f{-1}) / ((fmod - f{-1}) + (fmod - f{+1}))] \times A$

$Mo = 145 + [(13 - 13) / ((13 - 13) + (13 - 6))] \times 10$

$Mo = 145 + [0 / (0 + 7)] \times 10 = 145$ libras

Interpretación

Los valores de media (147.75), mediana (147.31) y moda (145) están muy cerca entre sí, lo que indica una distribución bastante simétrica, con un leve sesgo a la izquierda (ligeramente más valores por debajo de la media).

La mayoría de los estudiantes pesa entre 135 y 154 libras, una distribución común en adultos jóvenes. La dispersión es moderada, y los valores extremos (119 y 176) no afectan gravemente el promedio.

Este tipo de análisis puede ayudar a tomar decisiones sobre programas de salud universitaria, nutrición o diseño de mobiliario adecuado, por ejemplo, para instalaciones deportivas.

Con este último ejercicio, queda demostrado que los datos agrupados, aunque pierden algo de precisión individual, siguen siendo una herramienta poderosa para entender fenómenos reales. Has aprendido a construir intervalos, organizar frecuencias, aplicar fórmulas de MTC, dispersión y posición, e interpretar cada resultado con criterio.

Has llegado al final del capítulo más técnico, y ahora estás listo para ver cómo todas estas operaciones pueden realizarse con mayor facilidad utilizando un software especializado. En el siguiente capítulo, te enseñaré cómo usar JASP, una herramienta libre que te permite aplicar todo lo aprendido con solo unos clics.

ISBN: 978-628-97230-1-4

DOI: 10.62486/978-628-97230-1-4.ch05