

Capítulo 4

<https://doi.org/10.62486/978-628-97230-1-4.ch04>

Agrupar para entender: cómo analizar datos cuando vienen por intervalos

Hasta ahora hemos trabajado con datos individuales, no agrupados. Es decir, listas completas donde cada valor representa una observación específica. Este enfoque es útil cuando la cantidad de datos es manejable, pero ¿qué ocurre cuando tenemos decenas, cientos o incluso miles de observaciones? En esos casos, agrupar los datos en clases o intervalos no solo es más práctico, sino necesario. Aquí es donde entra la estadística descriptiva para datos agrupados.

Agrupar datos significa organizarlos en rangos o intervalos de clase, en lugar de listar cada uno por separado. Esto permite ver patrones más claramente y facilita el cálculo de medidas de resumen. Por ejemplo, si tenemos los tiempos de secado de cien muestras de pintura, en lugar de escribir cada valor, podemos organizar los resultados en intervalos como: 1–2 horas, 2–3 horas, 3–4 horas, y así sucesivamente, contando cuántas veces se repiten los valores dentro de cada intervalo.

Esta organización da lugar a una herramienta fundamental: la tabla de frecuencia. Esta tabla muestra cada intervalo, su frecuencia absoluta (cuántas veces ocurre), frecuencia acumulada, frecuencia relativa (porcentaje respecto al total) y en algunos casos, la marca de clase, que representa el punto medio del intervalo. Con esta información, es posible aplicar fórmulas específicas para obtener la media, mediana, moda, desviación estándar y otras medidas, adaptadas a este tipo de organización.

Para ilustrarlo, imagina que registramos la cantidad de hidrocarburos emitidos por una serie de vehículos de modelos distintos. Supongamos que, en lugar de

anotar los valores exactos, los datos ya vienen agrupados en intervalos: de 100 a 200, de 200 a 300, de 300 a 400, y así sucesivamente. Dentro de cada intervalo, contamos cuántos vehículos se encuentran. Esa frecuencia es la base del análisis.

A diferencia de los datos no agrupados, aquí ya no sabemos los valores exactos dentro de cada grupo, por lo que necesitamos hacer una estimación basada en los promedios de cada clase. Por eso, usamos la marca de clase (también llamada punto medio), que se obtiene sumando los límites inferior y superior del intervalo y dividiéndolo entre dos. Por ejemplo, si el intervalo es de 200 a 300, su marca de clase sería $(200 + 300) / 2 = 250$.

La media para datos agrupados se calcula multiplicando cada marca de clase por su frecuencia y luego sumando los productos. Esa suma total se divide entre el número total de datos. Es una media aproximada, pero muy útil para grandes volúmenes de información.

La mediana en datos agrupados se localiza dentro del intervalo que contiene el valor que divide a la distribución en dos mitades. Para hallarla, usamos una fórmula que involucra el límite inferior de la clase mediana, la frecuencia de esa clase, la frecuencia acumulada anterior y la amplitud del intervalo. Aunque la fórmula puede parecer compleja al principio, al aplicarla con calma se vuelve una herramienta poderosa para resumir la posición central de datos agrupados.

La moda, por su parte, se localiza en la clase con mayor frecuencia, conocida como clase modal. Para calcular su valor aproximado, se utiliza una fórmula que considera las frecuencias de las clases anterior y posterior, así como la amplitud del intervalo. Al igual que con la mediana, se obtiene un valor aproximado, pero suficiente para realizar comparaciones o interpretaciones útiles.

Las medidas de dispersión, como la varianza o la desviación estándar, también pueden calcularse en datos agrupados. Para ello, utilizamos nuevamente la

marca de clase y aplicamos fórmulas que integran la frecuencia y el punto medio al cuadrado, comparado con la media obtenida.

Además de estas medidas, los datos agrupados permiten calcular cuartiles, deciles y percentiles con fórmulas similares a la de la mediana, ya que también requieren localizar la clase donde se encuentra el valor de interés. Estas herramientas permiten hacer análisis más finos, como saber qué porcentaje de casos se encuentra por debajo de cierto umbral.

En este punto, es importante recordar que todos estos cálculos son aproximaciones, pero no por ello menos útiles. En estadística, muchas veces se trabaja con muestras y estimaciones, y el objetivo no es encontrar un valor exacto, sino describir el comportamiento general de los datos con precisión suficiente para tomar decisiones.

Otro aspecto clave al trabajar con datos agrupados es la representación gráfica. Aquí aparecen herramientas como el histograma, que es una gráfica de barras donde cada barra representa un intervalo, y su altura indica la frecuencia. También están los polígonos de frecuencia y los diagramas de caja y bigote, que ayudan a visualizar la distribución, simetría y presencia de valores atípicos.

El análisis de datos agrupados es especialmente valioso en contextos donde la información se recolecta en grandes cantidades, como estudios poblacionales, investigaciones de mercado, análisis de rendimiento escolar o experimentos científicos. Su utilidad radica en su capacidad para simplificar la complejidad sin perder de vista lo esencial.

A continuación, en el próximo capítulo, trabajaremos con ejercicios reales de datos agrupados. Verás paso a paso cómo se construyen las tablas, cómo se aplican las fórmulas y, lo más importante, cómo interpretar los resultados de manera clara. La práctica hará que todo esto te resulte natural, y descubrirás que,

con las herramientas adecuadas, la estadística agrupada puede ser tan comprensible como la individual.

Derechos de Autor (Copyright) Guillermo Alejandro Zaragoza Alvarado 2025 ©

Este texto está protegido por una licencia Creative Commons 4.0.



Usted es libre de compartir, copiar y redistribuir el material en cualquier medio o formato, así como de adaptarlo, remezclarlo, transformarlo y crear a partir de él para cualquier propósito, incluso con fines comerciales. Sin embargo, debe cumplir con la condición de atribución, lo que significa que debe otorgar el crédito correspondiente a la obra original de manera adecuada, proporcionar un enlace a la licencia e indicar si se han realizado modificaciones. Puede hacerlo en cualquier formato razonable, pero no de manera que sugiera que cuenta con el respaldo del licenciante o que recibe algún beneficio por el uso de la obra.

[Resumen de licencia](#) – [Texto completo de la licencia](#)

ISBN: 978-628-97230-1-4

DOI: 10.62486/978-628-97230-1-4.ch04