

### **1. ¿Crees que estos centros puedan ser representativos de los datos? ¿Por qué?**

Sí, los centros representan bien a los datos porque resumen el comportamiento promedio de cada grupo.

Por ejemplo, uno agrupa publicaciones con pocas reacciones, otro con actividad media y otro con muchas interacciones.

Así podemos entender cómo se comportan las publicaciones en general, sin revisarlas una por una.

### **2. ¿Cómo obtuviste el valor de k a usar?**

Usé el método del codo, que consiste en probar varios valores de k y ver en qué punto la mejora empieza a ser mínima.

Ese punto se llama “codo” y marca el número de grupos más adecuado.

En este caso, el codo apareció en  $k = 3$ , así que usamos tres grupos.

### **3. ¿Los centros serían más representativos si usaras un valor más alto? ¿Más bajo?**

Depende:

- Si pongo un k más alto, habría más grupos y los resultados serían más detallados, pero también más confusos.
- Si pongo un k más bajo, los grupos serían más generales y se perdería información. Por eso elegimos un valor intermedio ( $k=3$ ) para que sea equilibrado y claro.

### **4. ¿Qué distancia tienen los centros entre sí? ¿Hay alguno que esté muy cercano a otros?**

Los centros están bien separados, lo que indica que los grupos son distintos entre sí.

Puede que algunos estén un poco más cerca, especialmente los de interacción media y alta, porque tienen comportamientos parecidos.

Pero en general, sí se nota una buena diferencia entre cada tipo de grupo.

### **5. ¿Qué pasaría con los centros si tuviéramos muchos outliers en el análisis de cajas y bigotes?**

Si hay muchos outliers, los centros se moverían hacia ellos y podrían perder representatividad. Por eso es importante limpiar esos valores extremos antes de hacer el análisis, ya que pueden distorsionar los resultados.

### **6. ¿Qué puedes decir de los datos basándote en los centros?**

Los centros nos muestran tres tipos de publicaciones:

1. Las que tienen pocas reacciones.
2. Las que tienen actividad media.
3. Y las que destacan mucho, que serían las publicaciones virales.

En resumen, la mayoría de los posts no tienen muchas interacciones, pero hay unos pocos que destacan mucho, justo como vimos en el boxplot.