

## 86. Aplicación de técnicas de agrupación en clústeres

---

[Inteligencia Tips & Datos - Implementar agrupación de clúster en Power BI \(dds.pe\)](#)

La agrupación en Clústeres permite identificar un segmento (clúster) de datos similares entre sí, pero diferentes al resto de los datos. El proceso de agrupación en clústeres es diferente al proceso de agrupación que ha llevado a cabo antes.

La característica de agrupación en clústeres de Power BI permite encontrar rápidamente grupos de puntos de datos similares en un subconjunto de los datos. Analiza el conjunto de datos para identificar similitudes y diferencias en los valores de atributos y, después, separa los que tienen similitudes en un subconjunto de los datos. Estos subconjuntos de datos se denominan **Clústeres**.

---

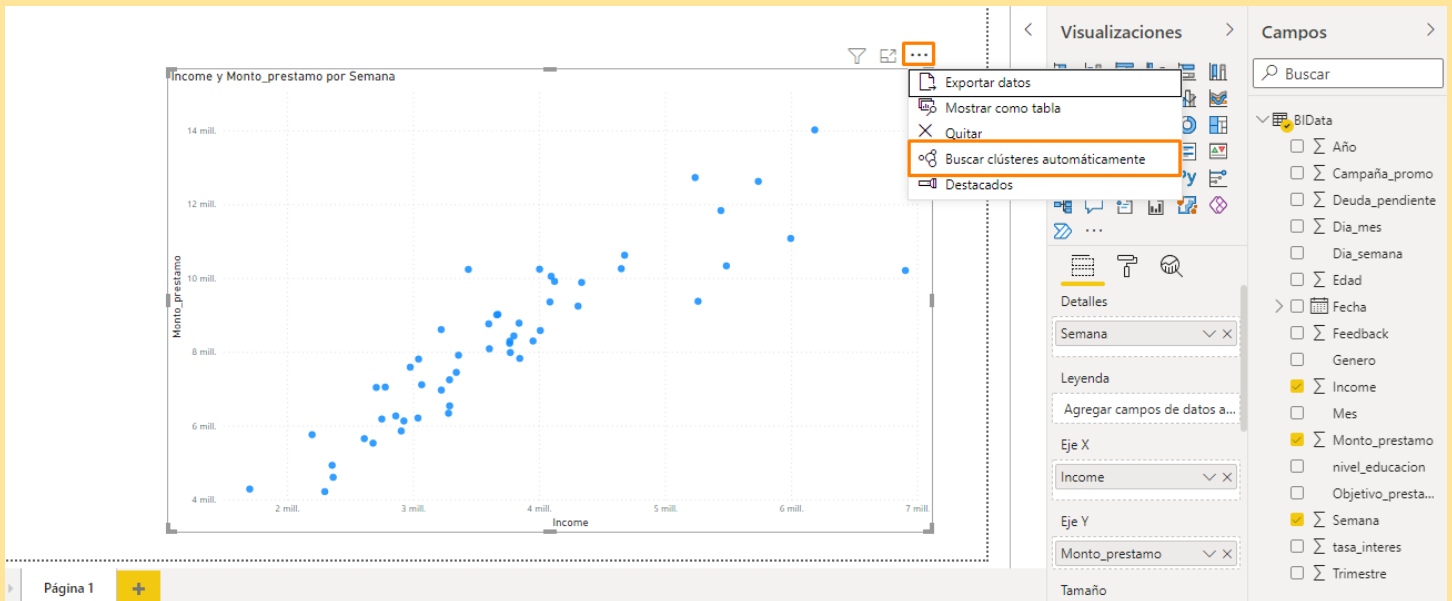
La agrupación en clúster es un algoritmo de aprendizaje automático no supervisado, que busca patrones en los datos dividiéndolos en grupos. Estas agrupaciones se crean de manera que los puntos sean homogéneos dentro del grupo y heterogéneos entre los grupos. La agrupación en clústeres se usa comúnmente en la segmentación de mercado y en varias áreas de análisis de marketing.

En nuestro nuevo “tip”, aprenderá a implementar la agrupación de clústeres en Power BI Desktop, para este caso, dos tipos de agrupación en clústeres: bivariante y multivariante.

### **Ejemplo: Agregar visualización para agrupación bivariante**

En nuestro “tip”, trabajaremos con un conjunto de datos ficticios de desembolsos de préstamos bancarios a lo largo de los años. Los datos contienen 3,000 observaciones y 17 variables. Puede descargar el conjunto de datos aquí:

1. Fecha: Fecha de desembolso del préstamo.
2. Ingresos: Ingresos anuales del solicitante (en dólares estadounidenses).
3. Monto\_prestamo: Monto del préstamo (en dólares estadounidenses) desembolsado por el banco.
4. Edad: Edad del solicitante en años.
5. Genero: Si el solicitante es mujer (F) o hombre (M).
6. Tasa\_interes: Tasa de interés anual, en porcentaje, que se cobra por el préstamo desembolsado.
7. Objetivo\_prestamo: Objeto para el que se tomó el préstamo.
8. Semana: Número de semana del año.
9. Deuda\_pendiente: Deuda pendiente (en dólares) del solicitante antes del desembolso del préstamo.



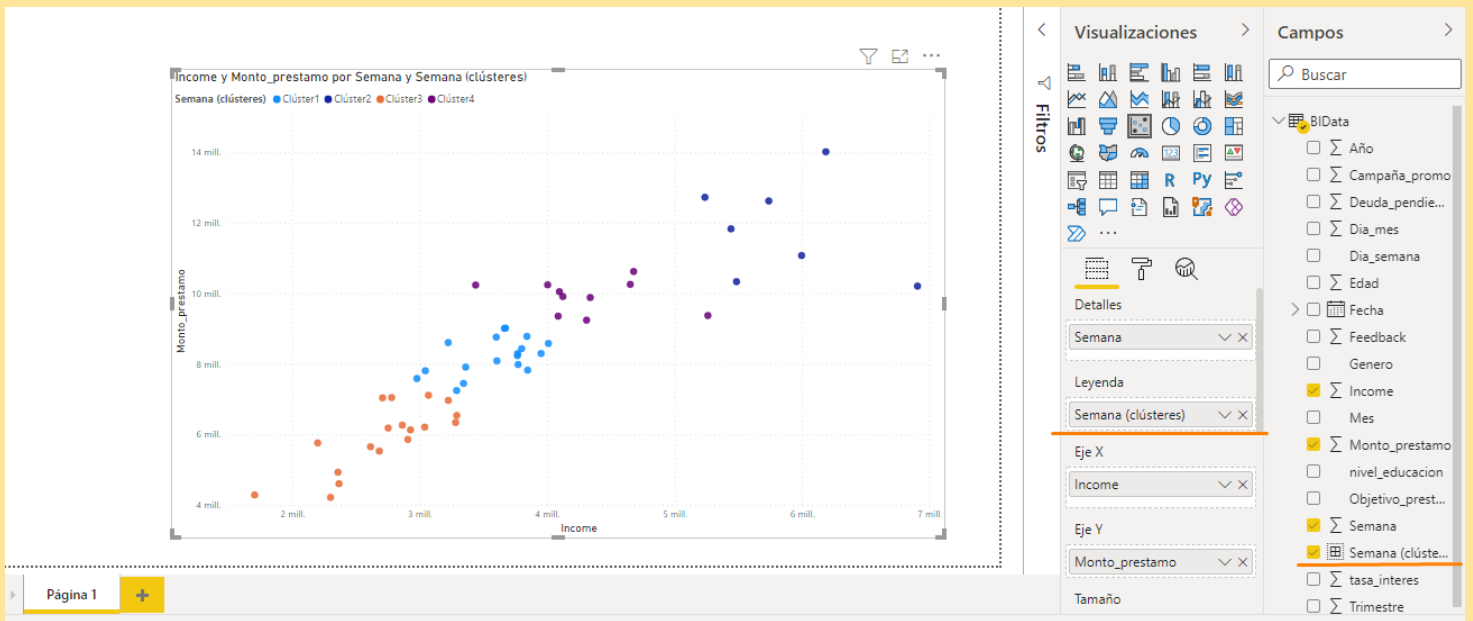
Colocamos el número de clústeres

## Clústeres

Nombre  Campo

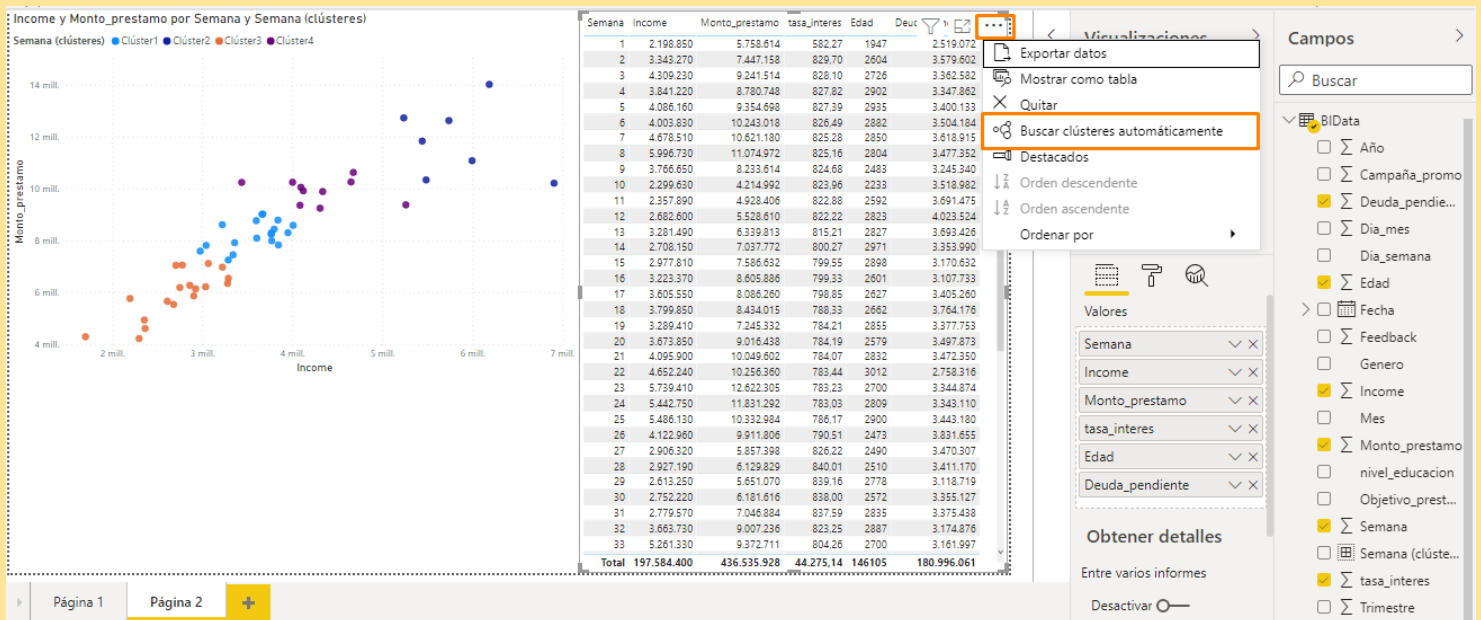
Descripción

Número de clústeres



En el resultado anterior, puede ver que Power BI ha creado una nueva variable Semana (clústeres), y la ha colocado en el campo Leyenda. Esto le brinda la flexibilidad de usar esta nueva variable de la misma manera que cualquier otra variable en el conjunto de datos original.

### **Ejemplo: Agregar visualización para agrupación multivariante**



Clústeres

Nombre

Clúster Multivariante

Campo

Semana

Descripción

Clústeres de Semana

Número de clústeres

Automático

Aceptar

Cancelar

