**Análisis de datos**

Johan Alfonso Felix Fuquen

Ingeniería, Universidad Nacional Abierta y a Distancia UNAD

Análisis de datos

Dr. Breyner Alexander Parra

Noviembre de 2024

**INTERPRETACION DEL K-MEANS**

El dataset contiene las siguientes columnas:

1. CustomerID: Identificador único de cada cliente.
2. Gender: Género del cliente (Male/Female).
3. Age: Edad del cliente.
4. Annual Income (k$): Ingreso anual en miles de dólares.
5. Spending Score (1-100): Puntuación de gasto asignada al cliente (basada en su comportamiento y datos de gasto).

**Análisis Exploratorio de los Datos:**

1. Distribución de las variables numéricas:
   * Edad: La mayoría de los clientes tienen entre 20 y 50 años, con pocos extremos jóvenes o mayores.
   * Ingreso Anual: Oscila entre $15k y $137k, con una distribución relativamente uniforme.
   * Puntuación de Gasto: La distribución es más dispersa, con clientes en un amplio rango de valores (1 a 100).
2. Correlación:
   * No hay fuertes correlaciones entre las variables numéricas, lo que indica que cada una podría aportar información única al modelo.
3. Valores faltantes:
   * No se encontraron valores nulos en el dataset.