

**Diplomatura en Data Science aplicada**  
**APRENDIZAJE NO SUPERVISADO**

# Segmentación de mercado

Ferrero, N - Ghelfi, A - Guzman, C - Manoni, C.



## PUNTO DE PARTIDA

# Situación

Determinar características principales distintivas y definir patrones ocultos en un centro comercial para poder generar segmentos con distintos tipos de clientes.

# Objetivo

Identificar la base de clientes con mayor potencial para la venta de productos.

# Datos



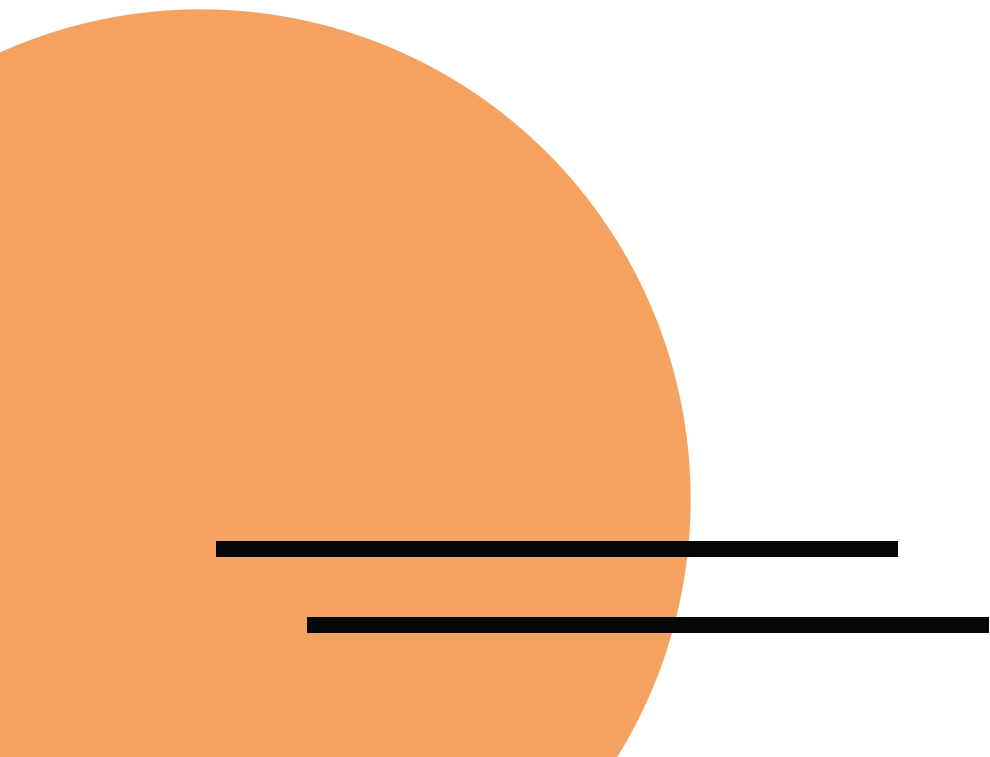
**200 FILAS DE CLIENTES**



**5 COLUMNAS DESCRIPTIVAS**



ID cliente  
Edad  
Género  
Ingresos anuales  
Nivel de gastos



# Datos



200 FILAS DE CLIENTES



5 COLUMNAS DESCRIPTIVAS



ID cliente  
Edad  
Género  
Ingresos anuales  
Nivel de gastos

## Descripción



### EDAD

Mínima: 18

Máxima: 70



Promedio: 38



### GENERO



112 Mujeres

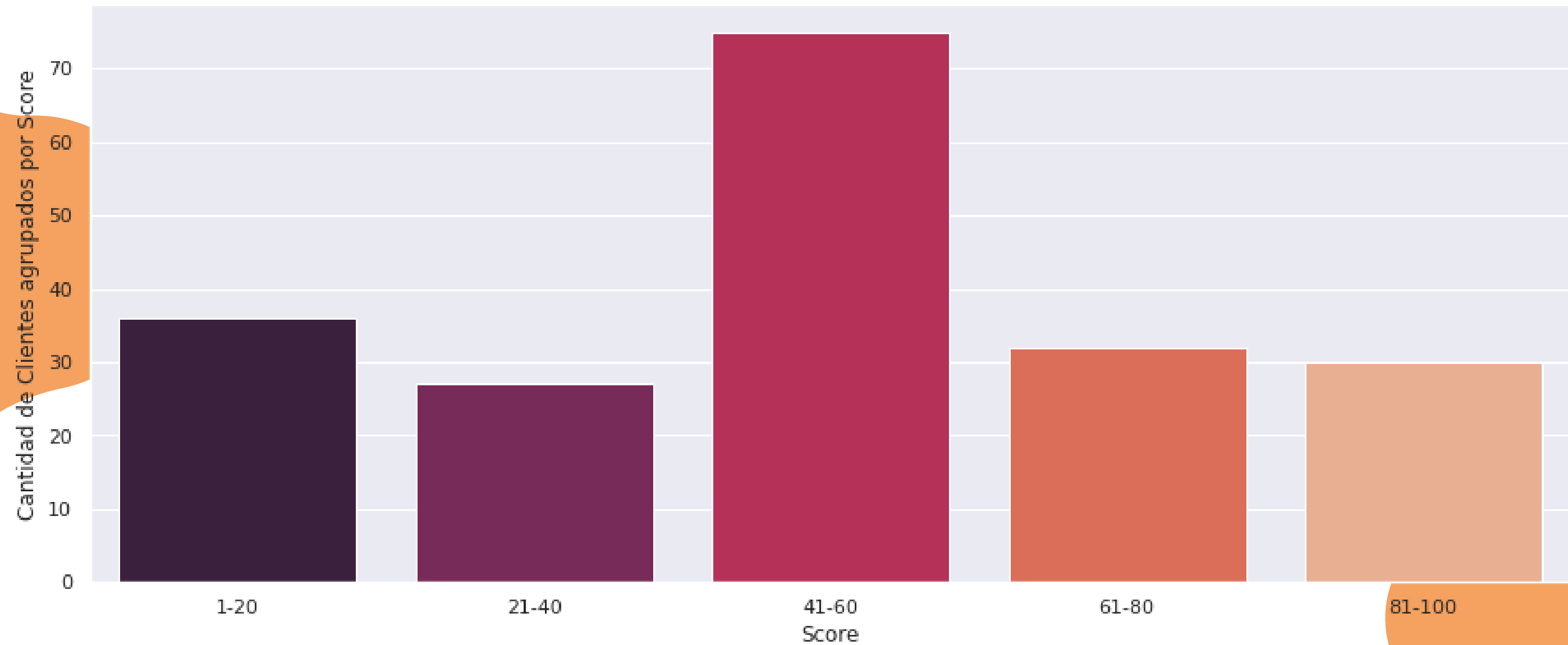


88 Hombres



## NIVEL DE GASTOS

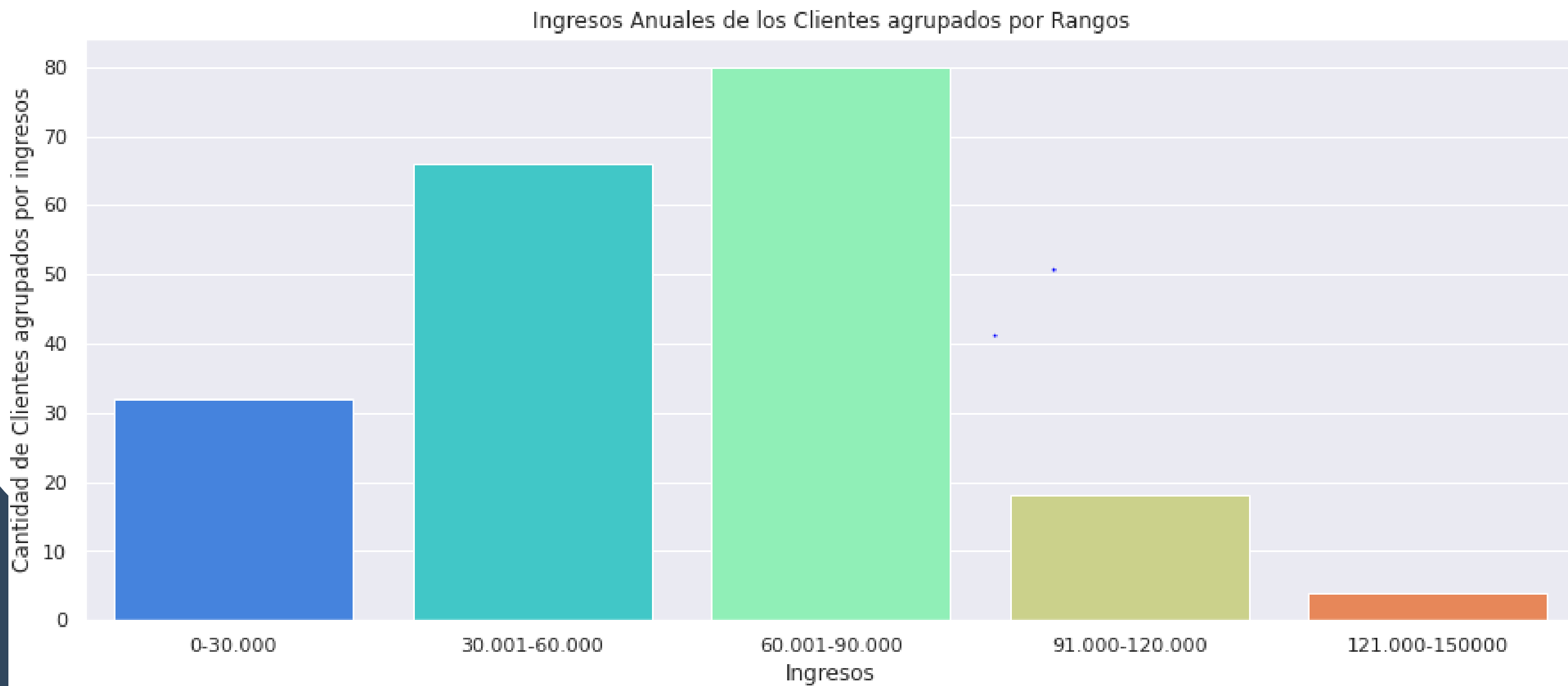
Gasto de los Clientes en funcion del Rango de Score



**Descripción**



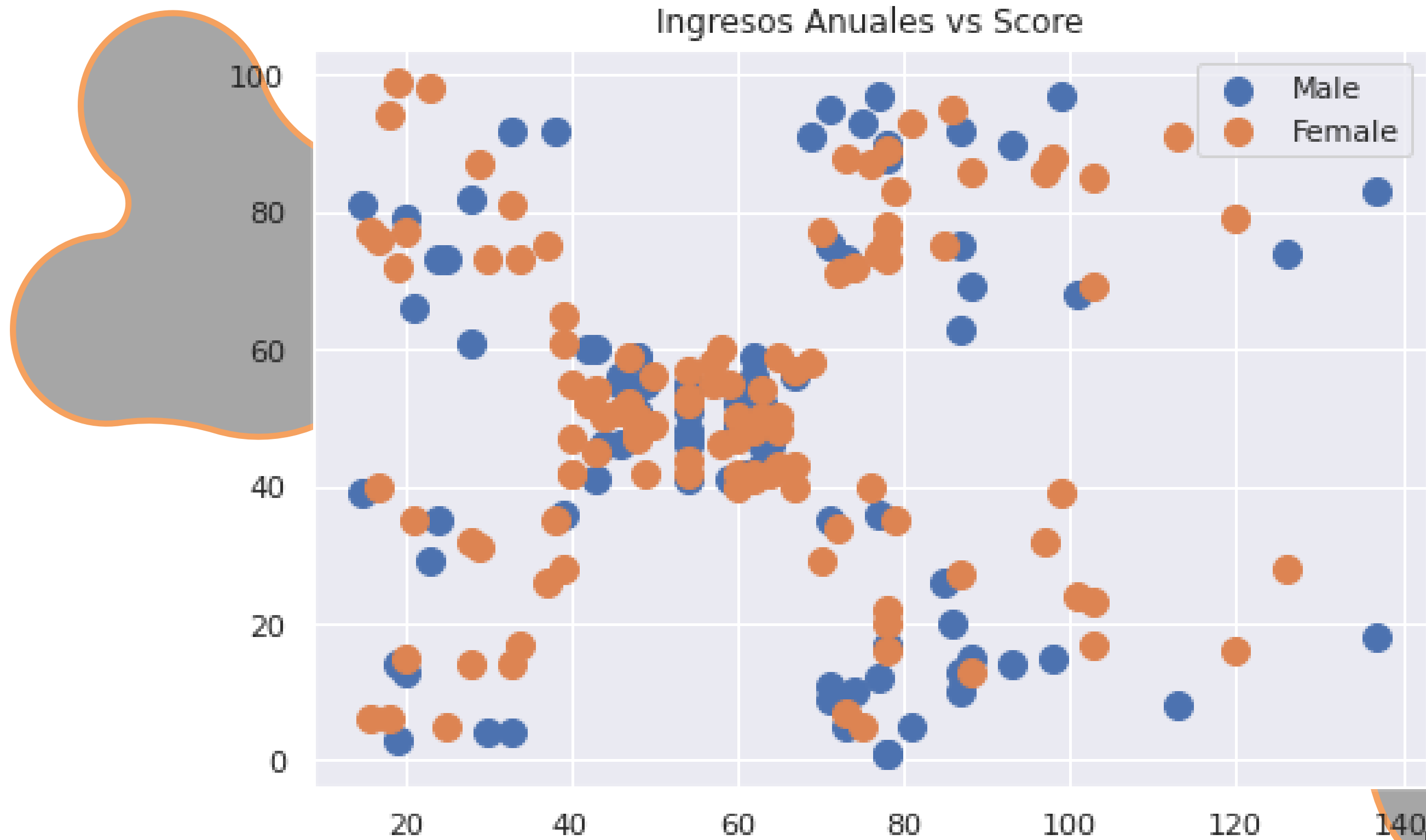
## INGRESOS ANUALES



**Descripción**



## NIVEL DE GASTOS & INGRESOS ANUALES



Descripción

# Segmentación

## Variables consideradas

INGRESOS ANUALES

NIVEL DE GASTOS

Seleccionadas debido a que consideramos que están más relacionadas al poder de compra de los clientes





## K-MEANS

5 grupos de clientes



## DBSCAN

Considerando radio de 0.4 y 5 vecinos  
4 grupos de clientes



## CLUSTERING JERÁRQUICO

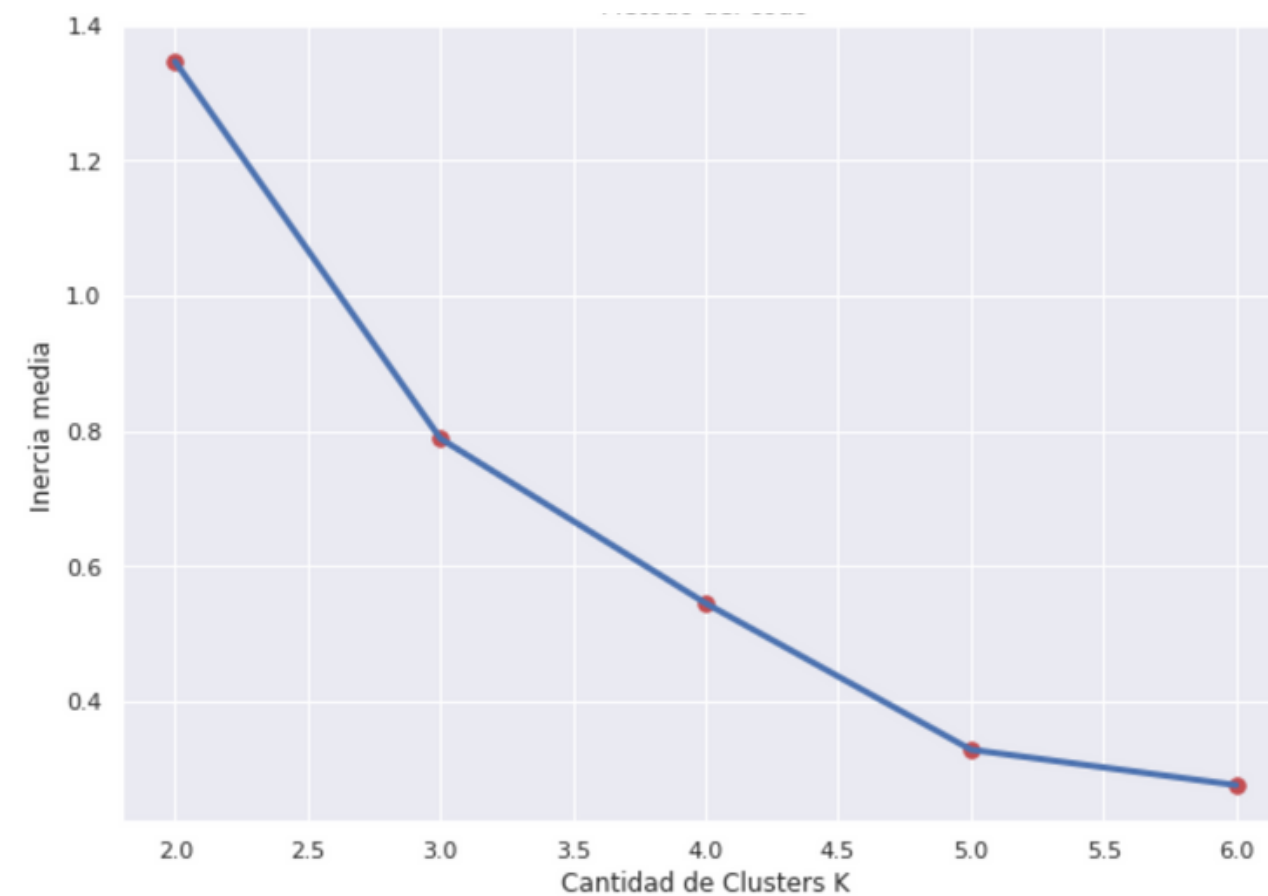
5 grupos de clientes

# Segmentación

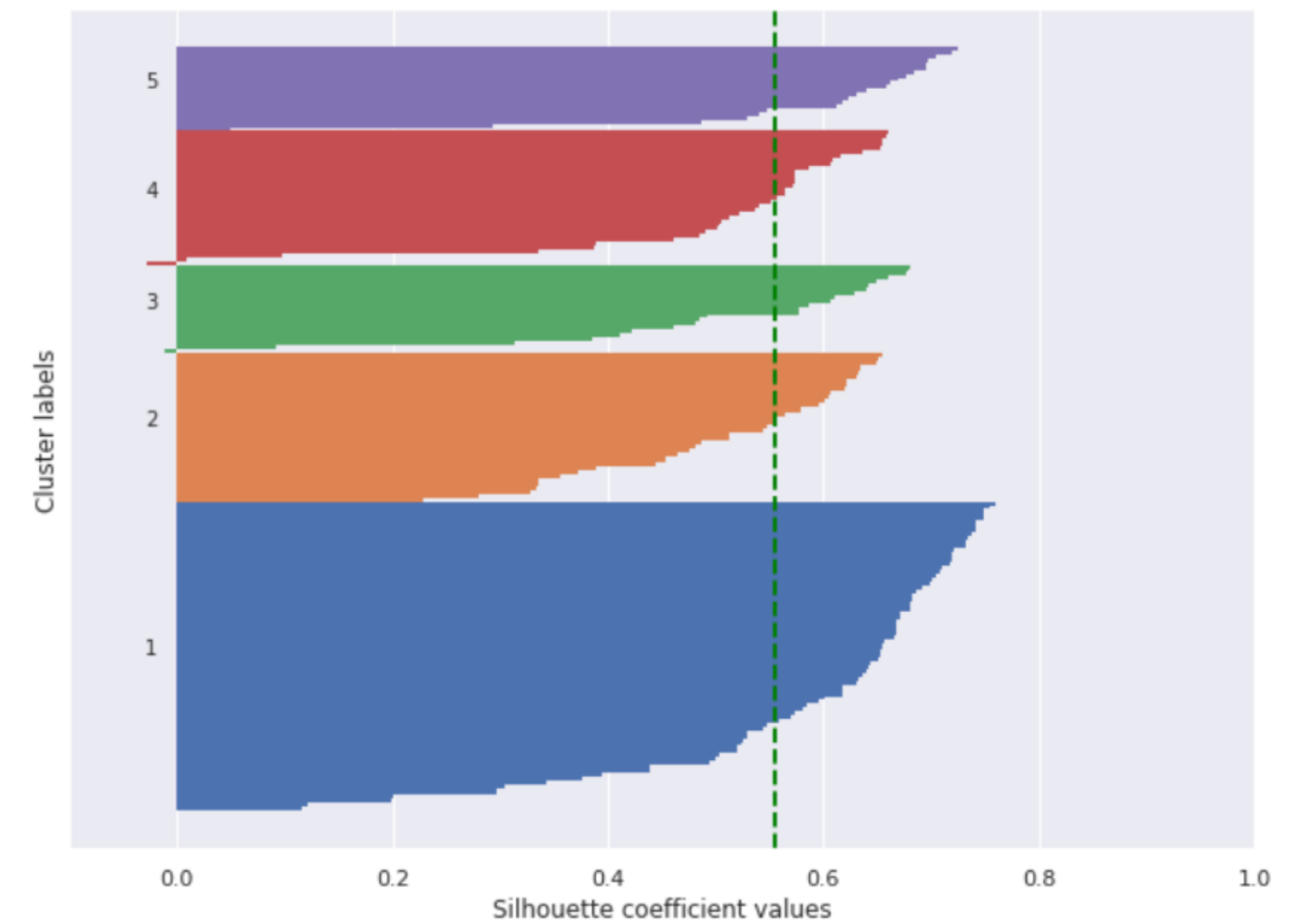
# K-Means

Procedimiento para determinar la cantidad de grupos óptimos

CODO



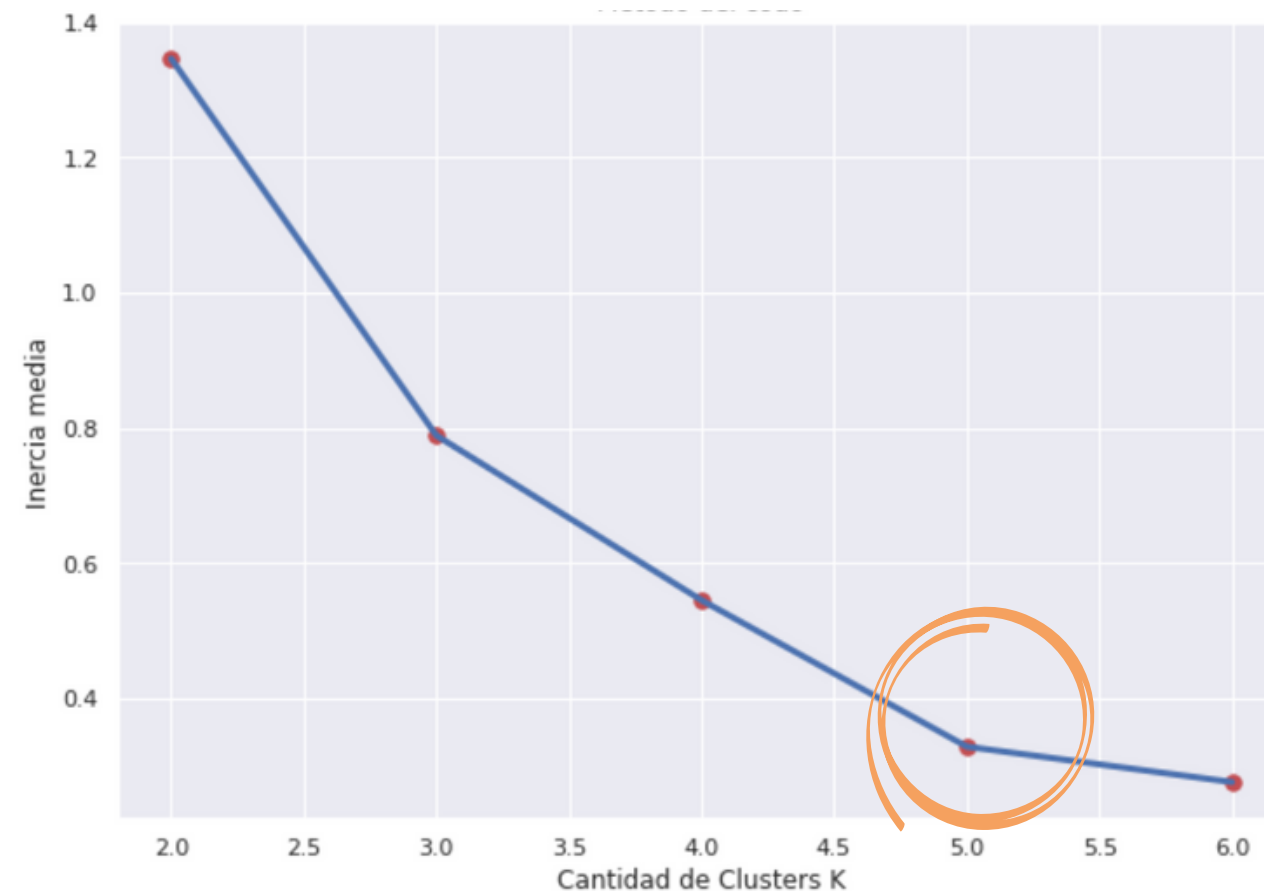
SILUETA



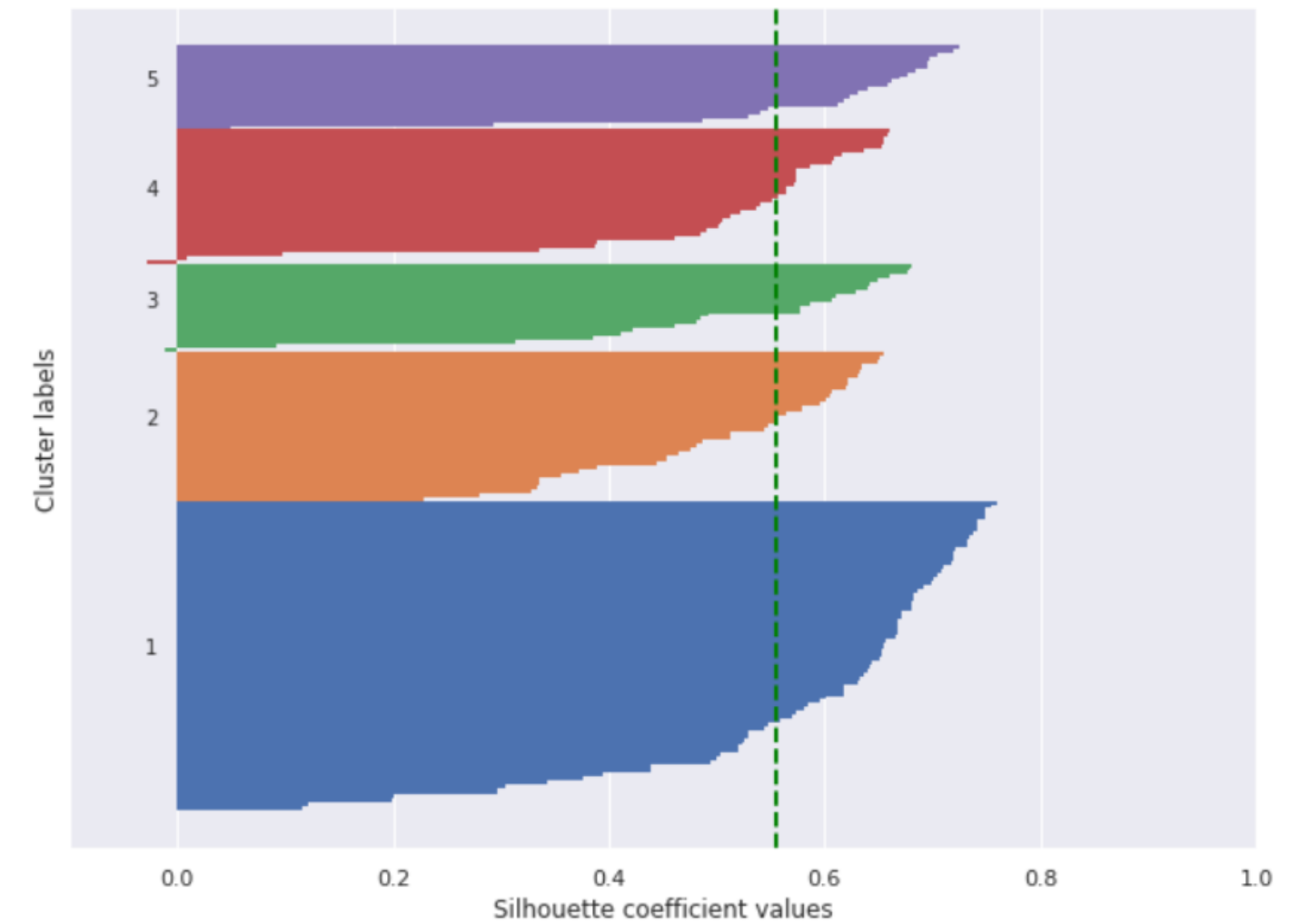
# K-Means

Procedimiento para determinar la cantidad de grupos óptimos

CODO

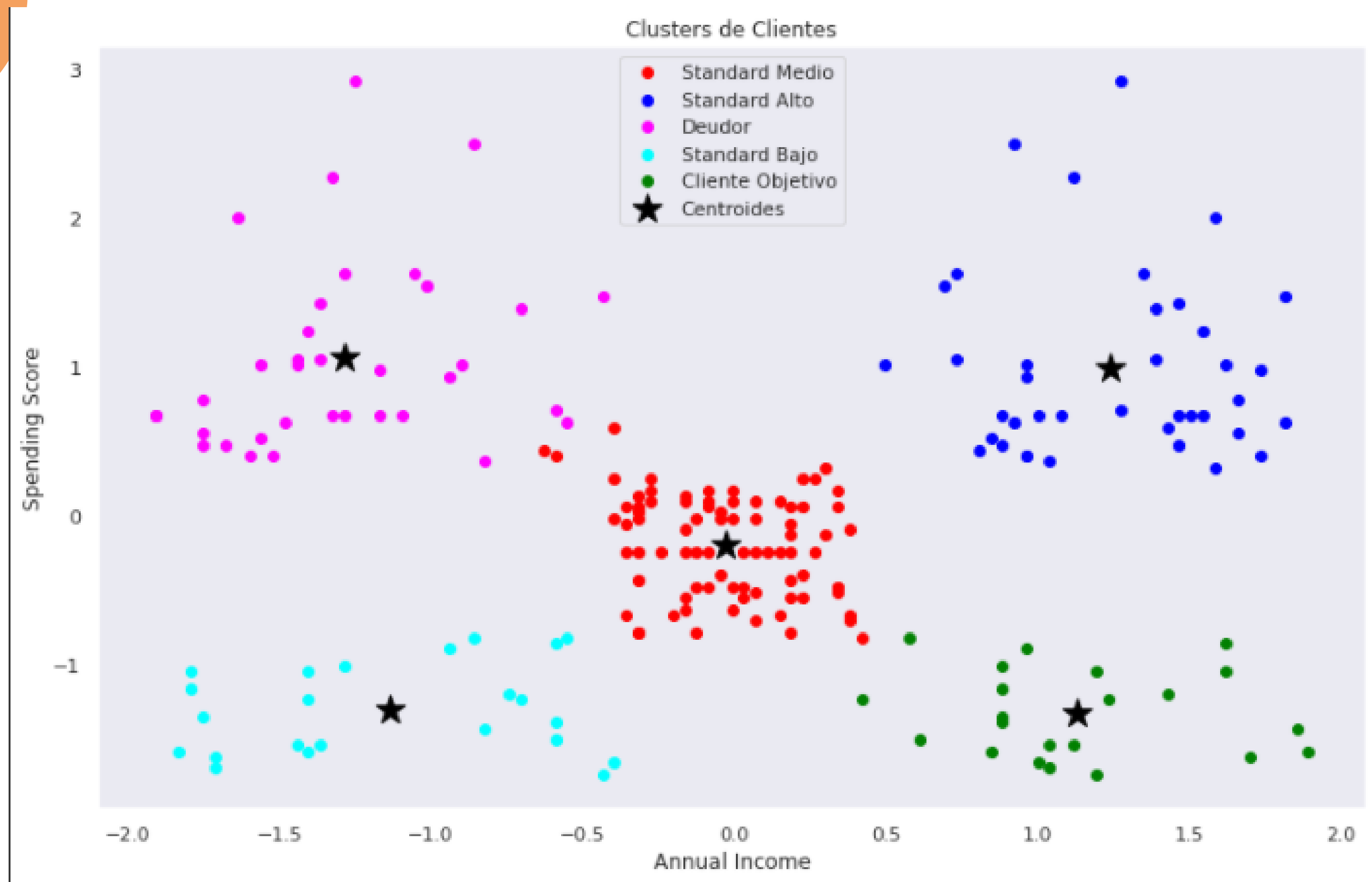
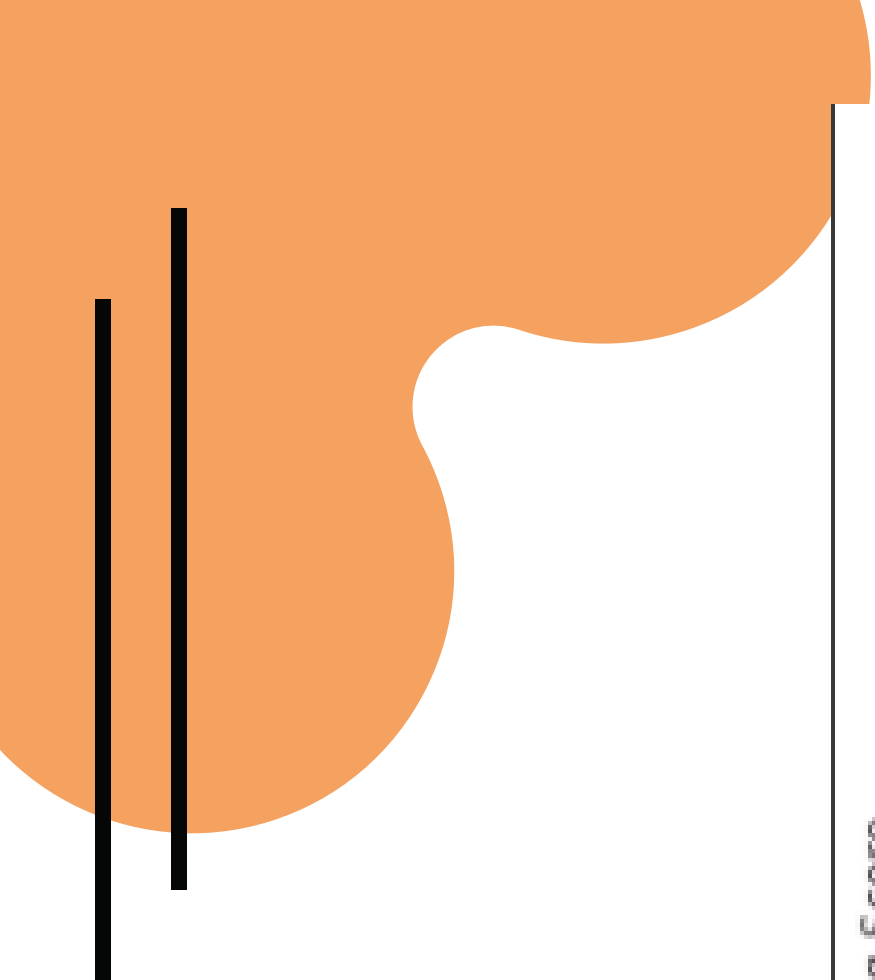


SILUETA

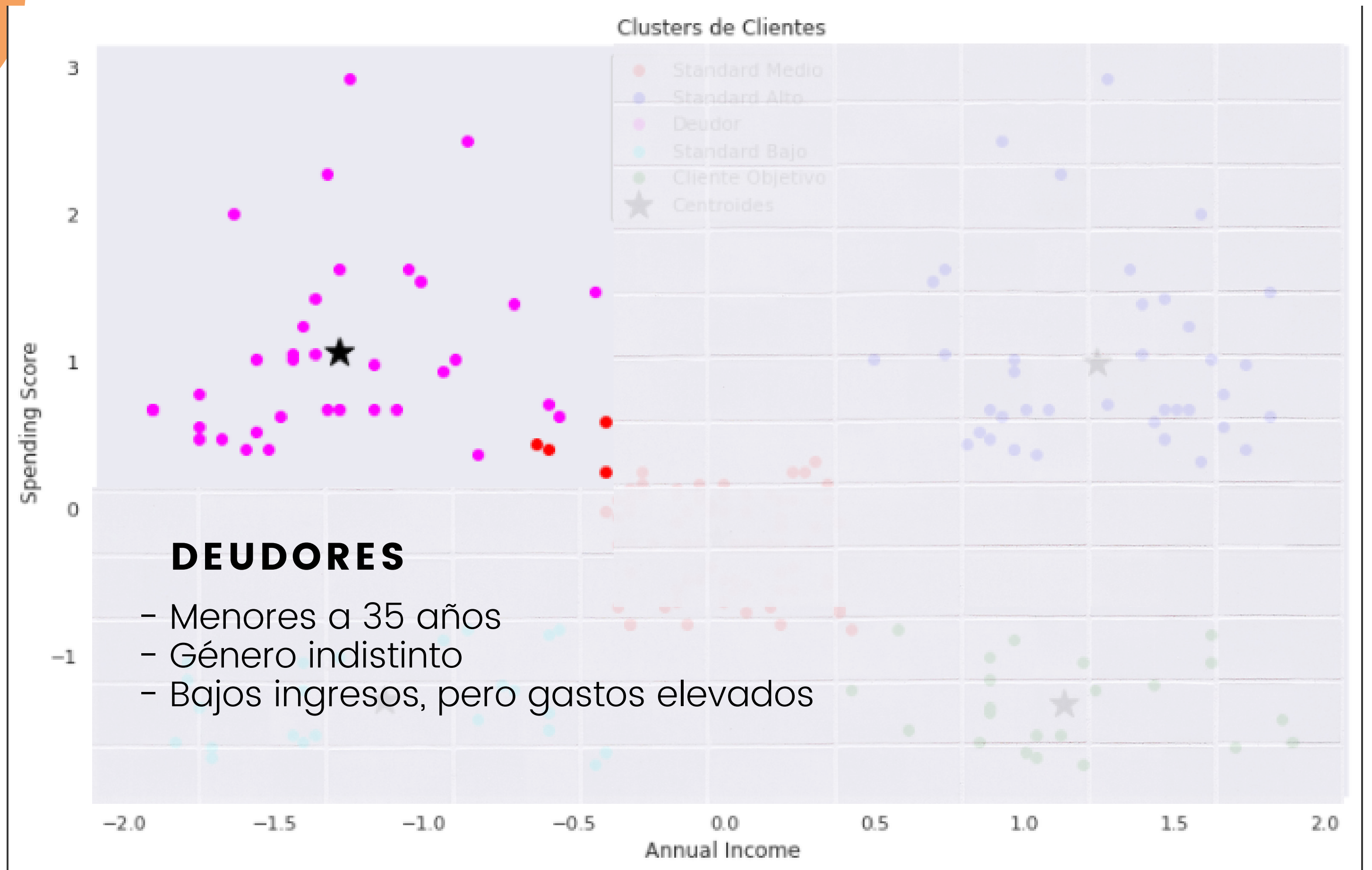
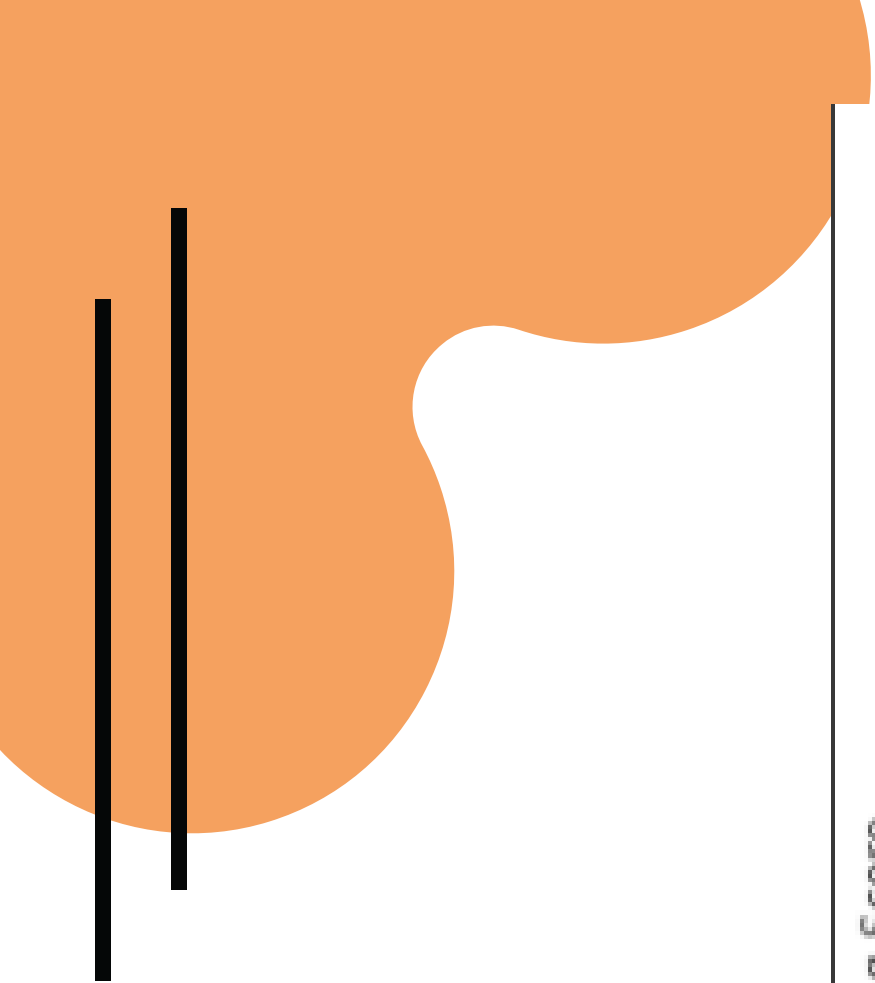


Separación por distancia con

**K=5**



# Resultados

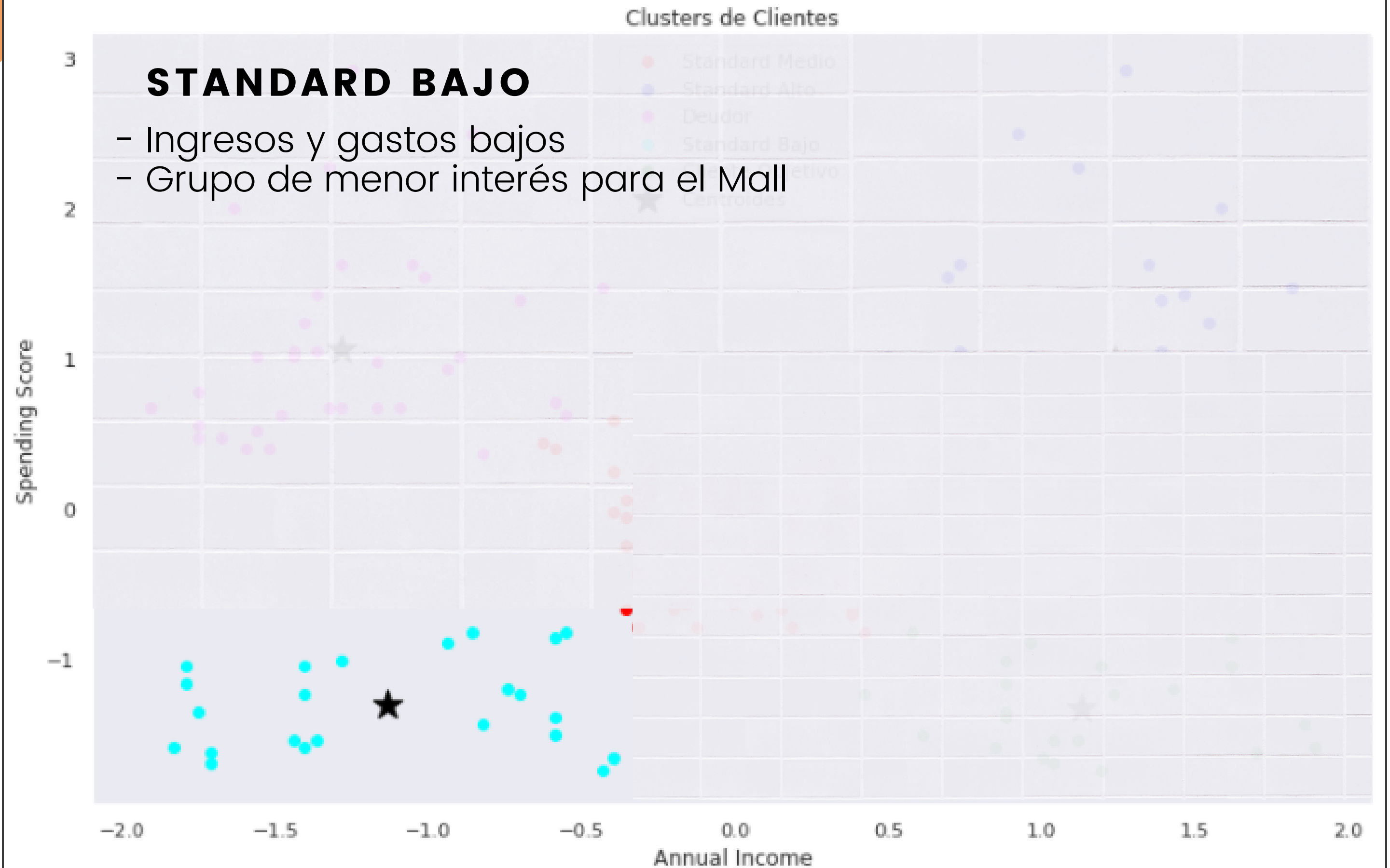


# Resultados



## STANDARD BAJO

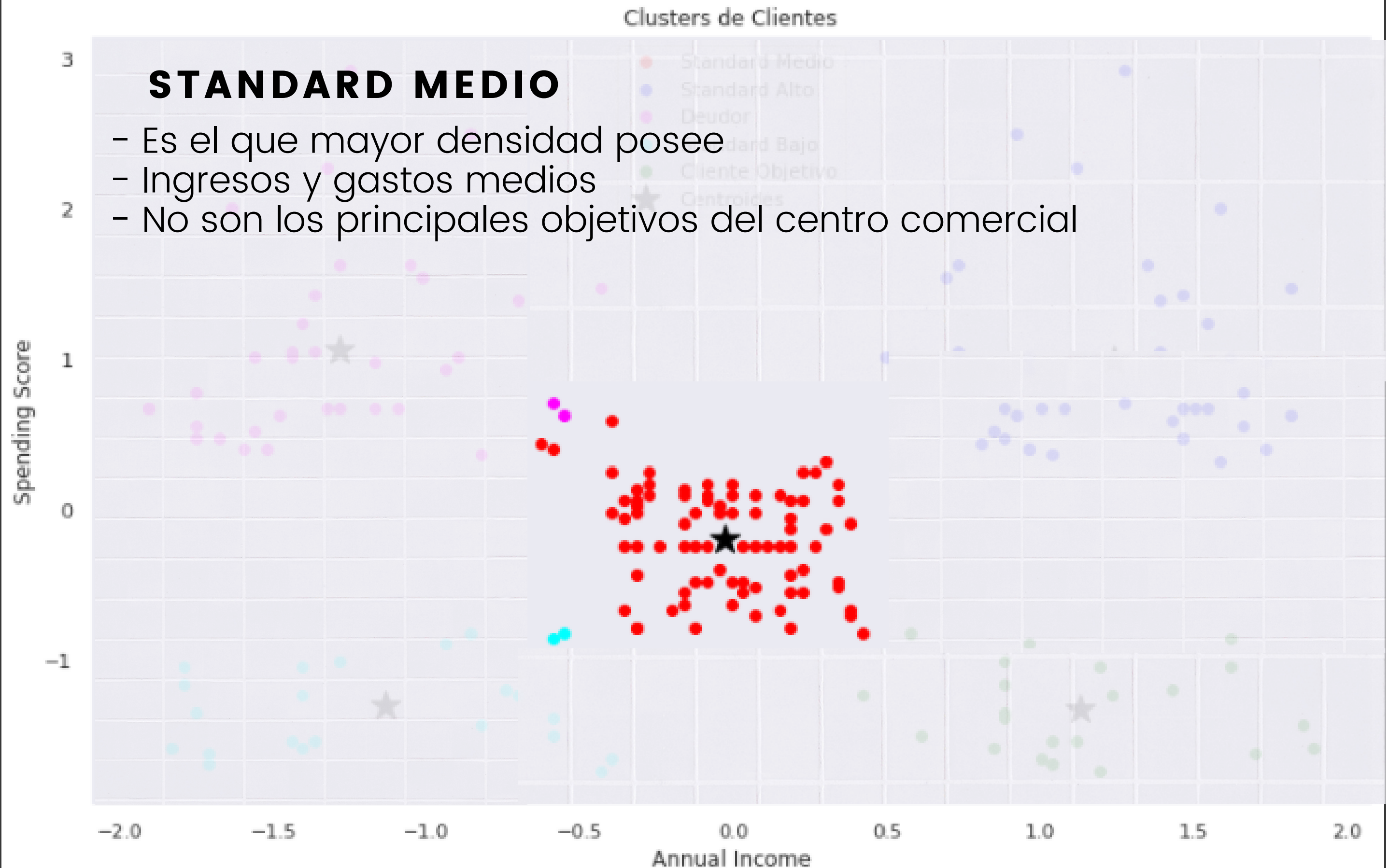
- Ingresos y gastos bajos
- Grupo de menor interés para el Mall



# Resultados

## STANDARD MEDIO

- Es el que mayor densidad posee
- Ingresos y gastos medios
- No son los principales objetivos del centro comercial

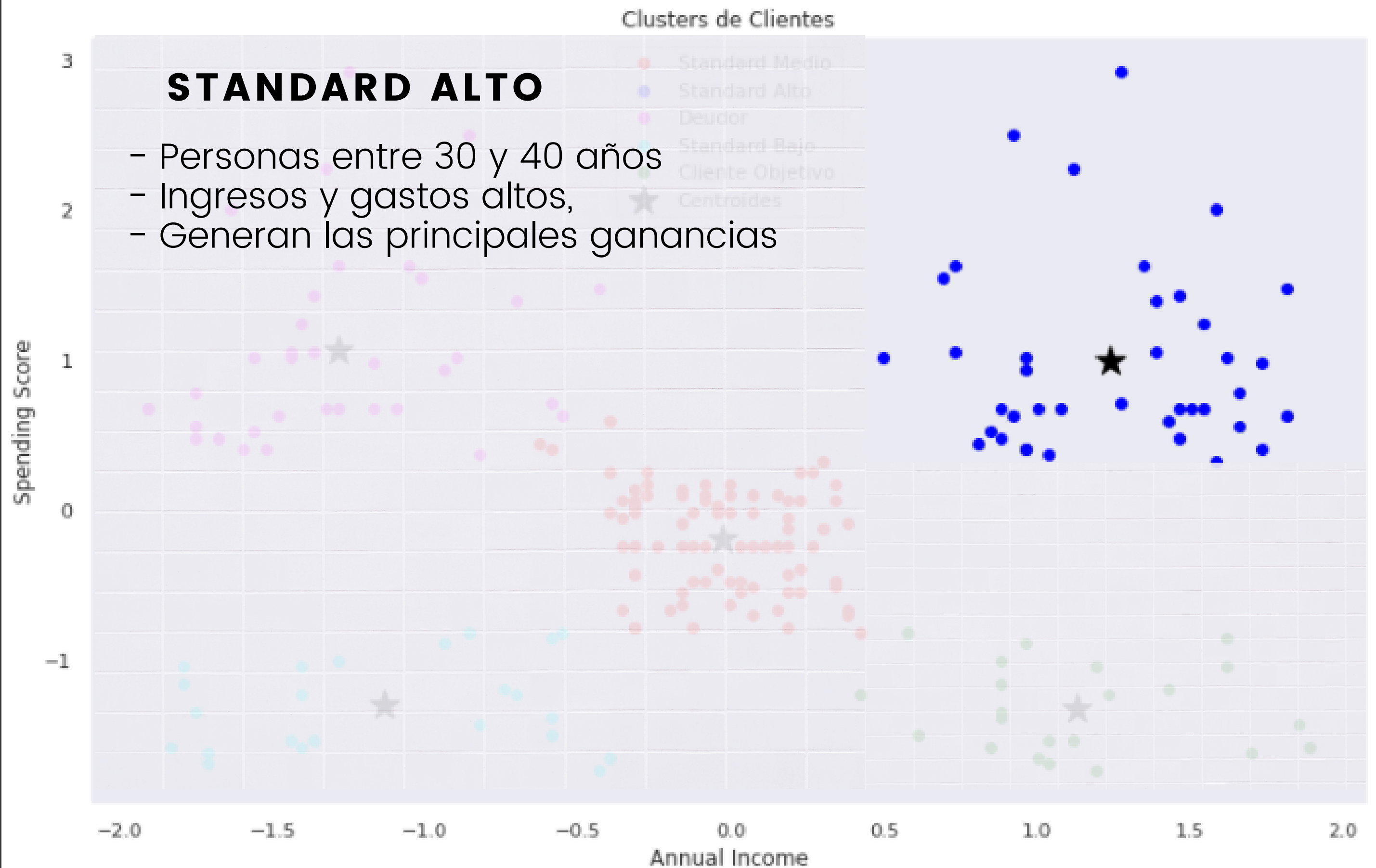


# Resultados



## STANDARD ALTO

- Personas entre 30 y 40 años
- Ingresos y gastos altos,
- Generan las principales ganancias

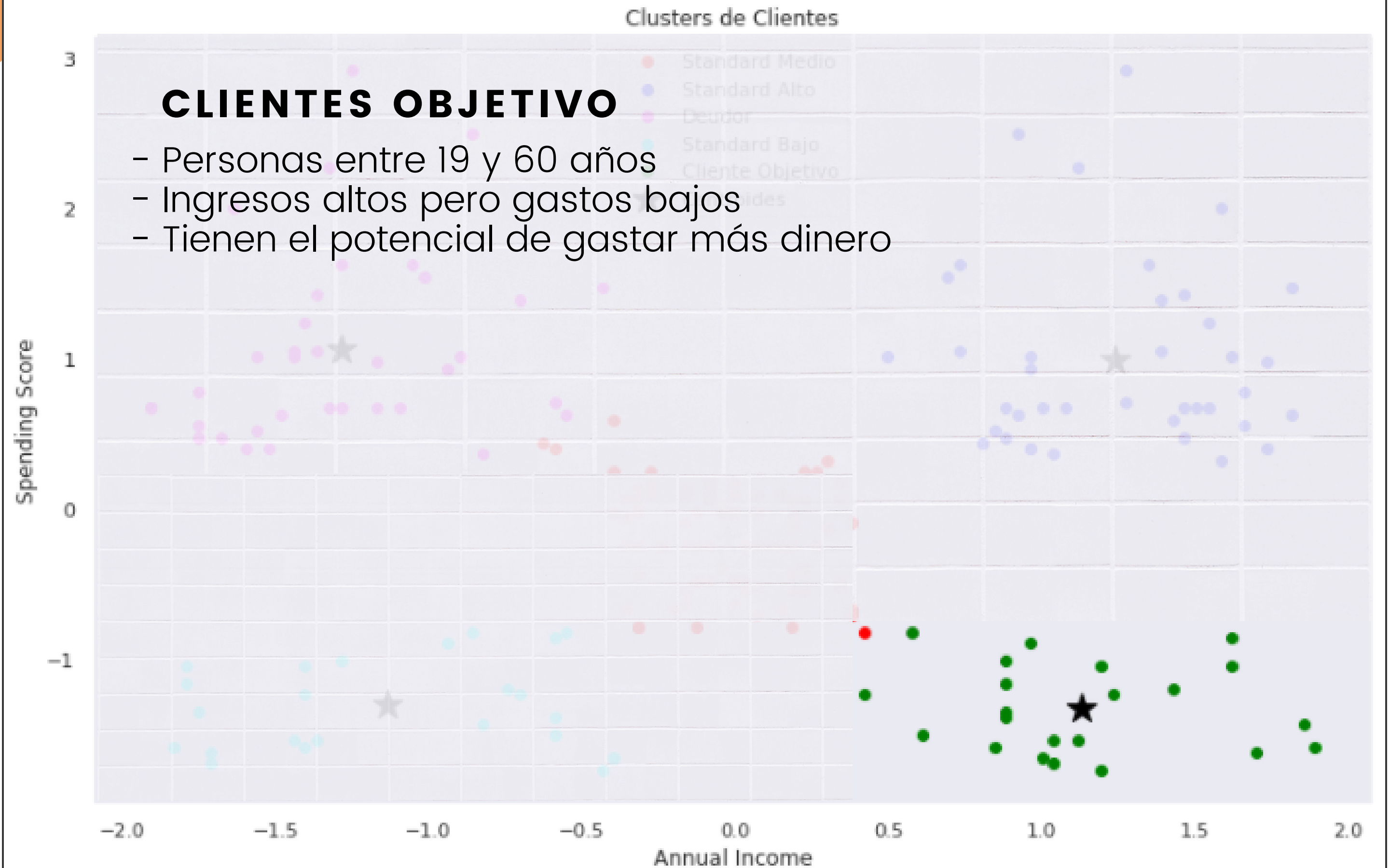


# Resultados



## CLIENTES OBJETIVO

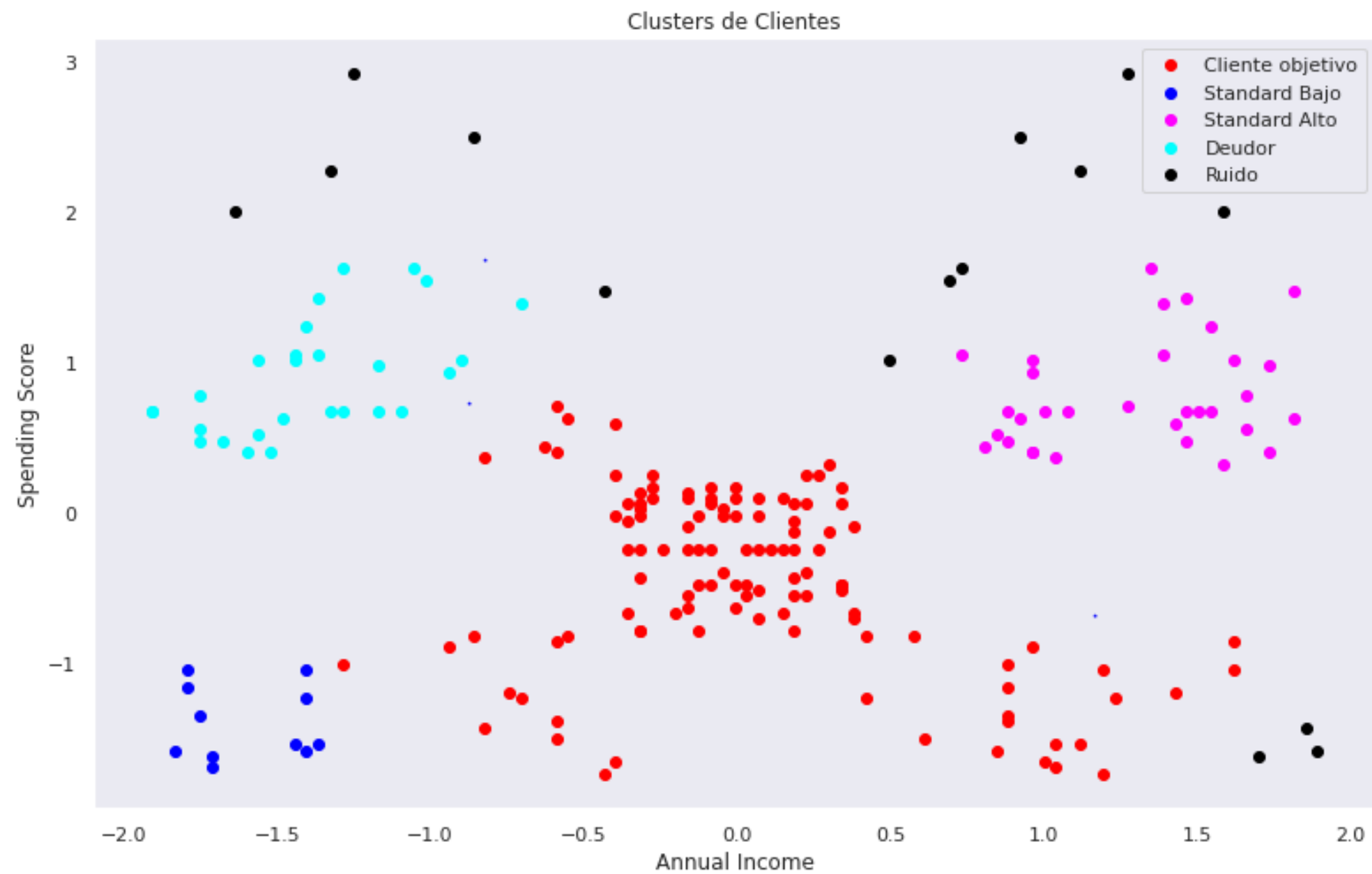
- Personas entre 19 y 60 años
- Ingresos altos pero gastos bajos
- Tienen el potencial de gastar más dinero



# Resultados

# Dbscan

Segmentación óptima de acuerdo a la densidad



# Dbscan

Segmentación óptima de acuerdo a la densidad

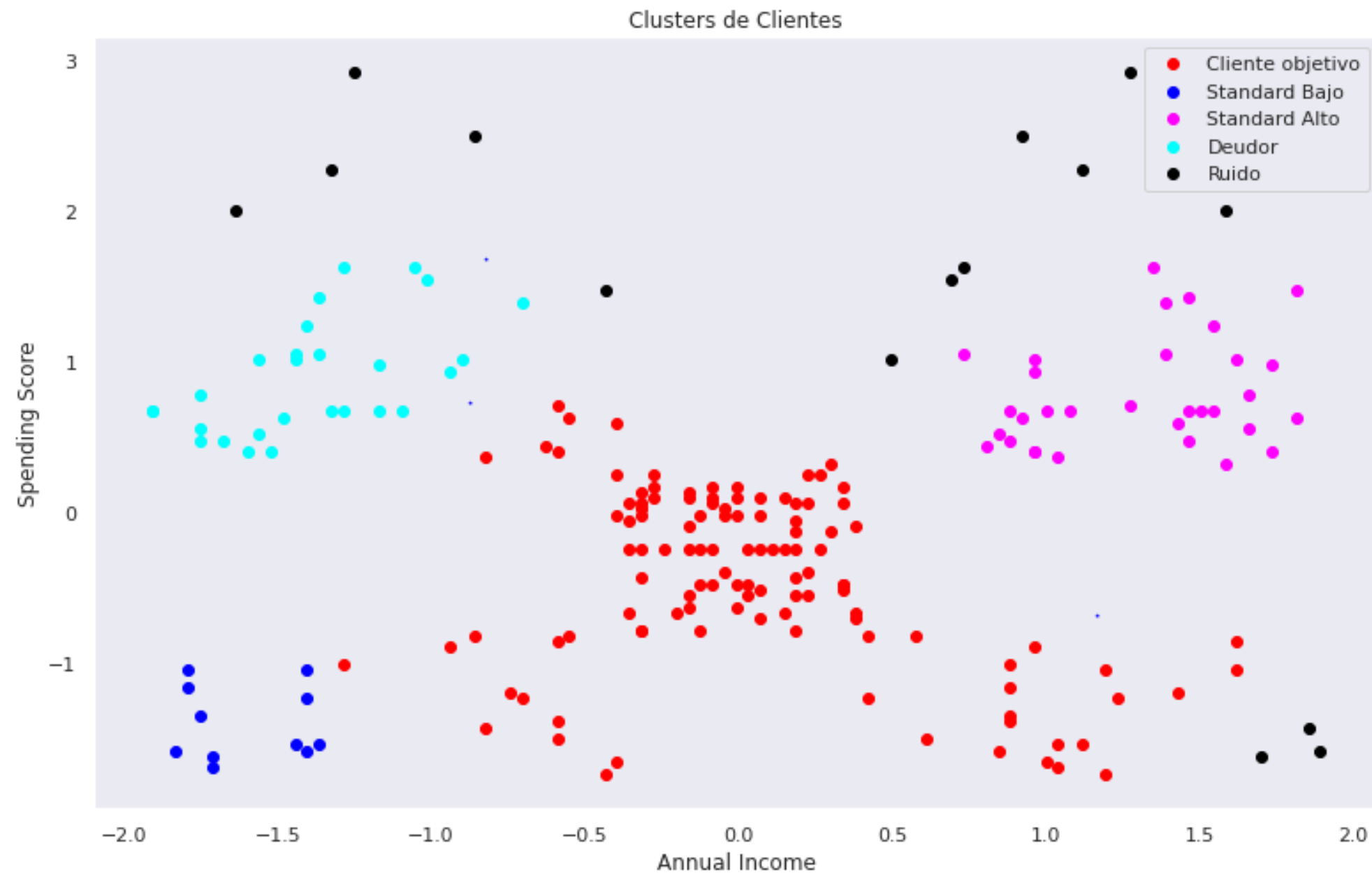


## RUIDO

- 👉 Puntajes de gastos más altos
- 👉 Standar Alto y Deudores de Kmeans

# Dbscan

Segmentación óptima de acuerdo a la densidad



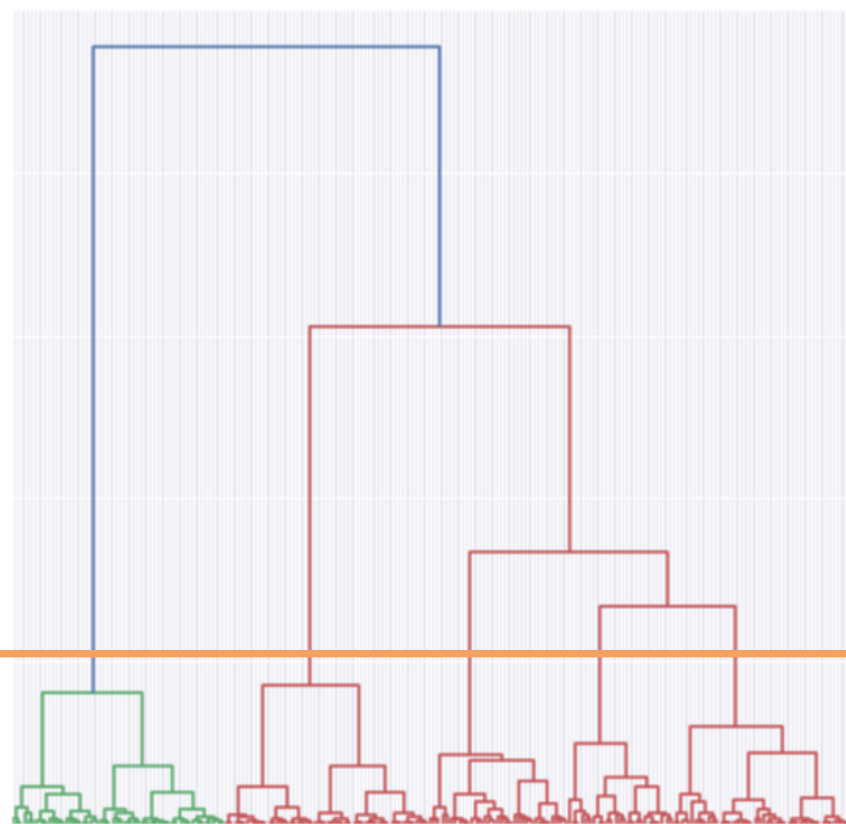
## CLIENTES OBJETIVO

- ➡ Densidad y cantidad muy alta en relación con los demás
- ➡ Absorbió a los Clientes Standard Medio de Kmeans
- ➡ 60% integrantes femeninos
- ➡ Edad indistinta
- ➡ Ingresos medio/altos

# Jerárquico

DENDOGRAMA

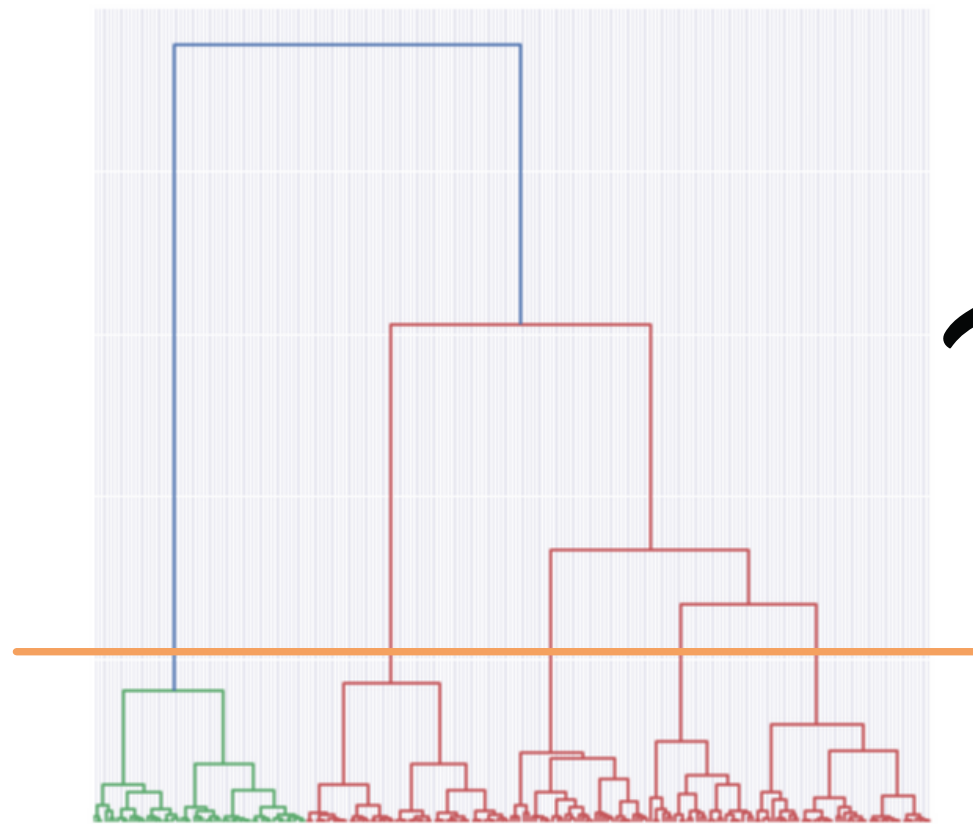
K=5



# Jerárquico

DENDOGRAMA

K=5



RESULTADOS

# Jerárquico

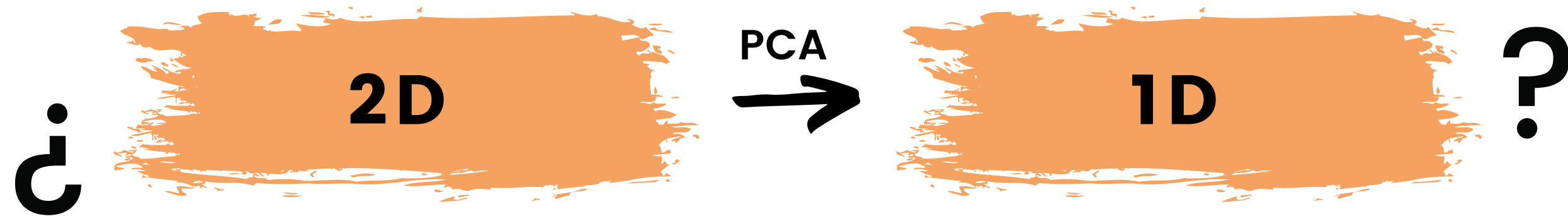
Grupos de clientes similares a los obtenidos con el algoritmo k-means



RESULTADOS



# Reducción de dimensionalidad



Consideramos que no es aplicable el algoritmo de PCA al trabajar originalmente con dos variables.



# Cientes objetivo

## CARACTERÍSTICAS

- Edad indistinta
- Género indistinto
- Ingresos altos
- Gastos bajos

## INTERROGANTES DE NEGOCIO

- ¿Estarán insatisfechas con los servicios del centro comercial?
- ¿Será errónea la estrategia de marketing?

## SUGERENCIAS

Intentar agregar nuevas instalaciones/servicios que puedan atraer y tentar a estos clientes

