



EL FARO

LIBRERÍA

EL FARO

Proyecto de Ciencia de Datos

Caso: Librería “El Faro”

Ximena Macías

Edgar Chacón

Emiliano Salinas







Compras Basadas en Intuición

Exceso de Inventario

Pérdida de Ventas

El objetivo del proyecto es abandonar el modelo de negocio y compra “basado en la intuición” y adoptar una estrategia de reinversión de inventario basada en datos.



Predecir



Predecir



Clasificar



Predecir



Clasificar



Buscar Relaciones



Predecir

Tiempo



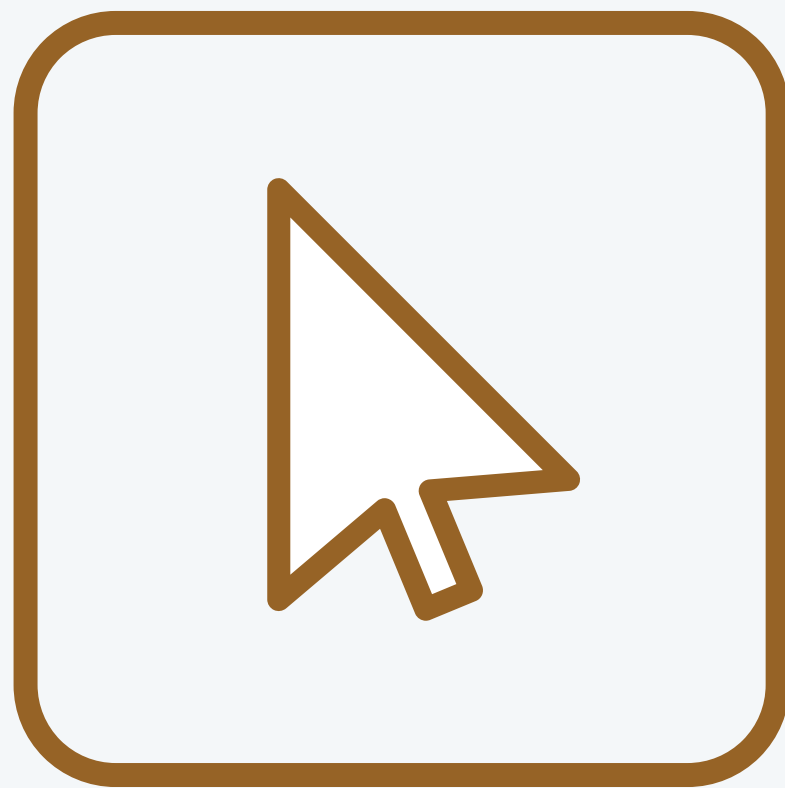
Clasificar

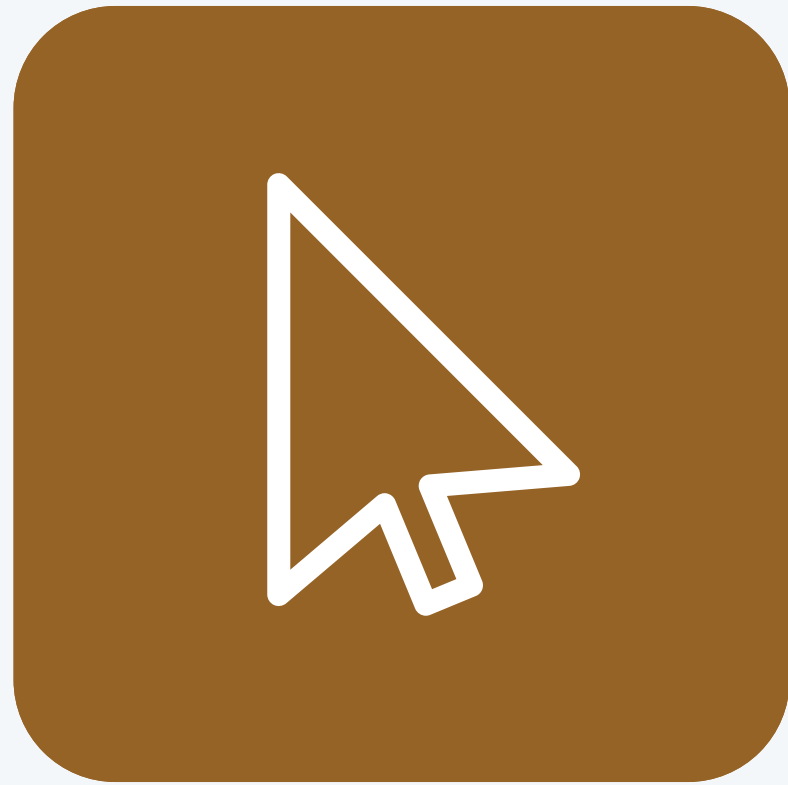
Disponibilidad



Buscar Relaciones

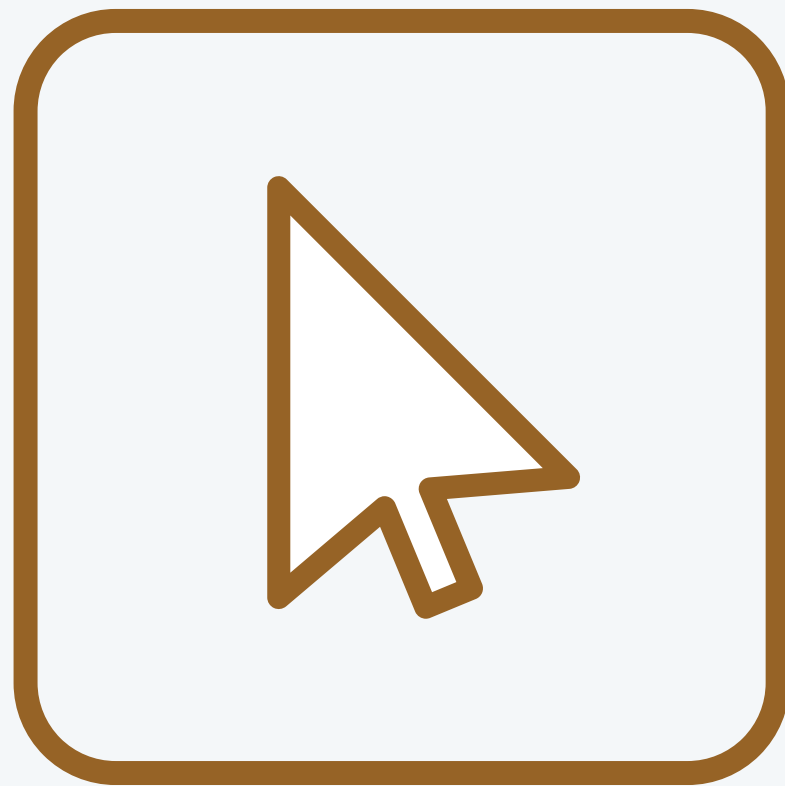
Calidad de los datos





Recolección de URLs





Extracción de Atributos

| Métrica | Price | Availability | Rating |
|------------------------|------------------|--------------|------------------|
| Media | 35.07 | 8.58 | 2.92 |
| Mediana | 35.98 | 7 | 3 |
| Rango | 10.00 - 59.99 | 1 - 22 | 10.00 - 59.99 |
| Desviación Estándar | 14.45 | 5.65 | 1.43 |

Proceso de limpieza

Valores nulos y duplicados

Eliminación de columnas

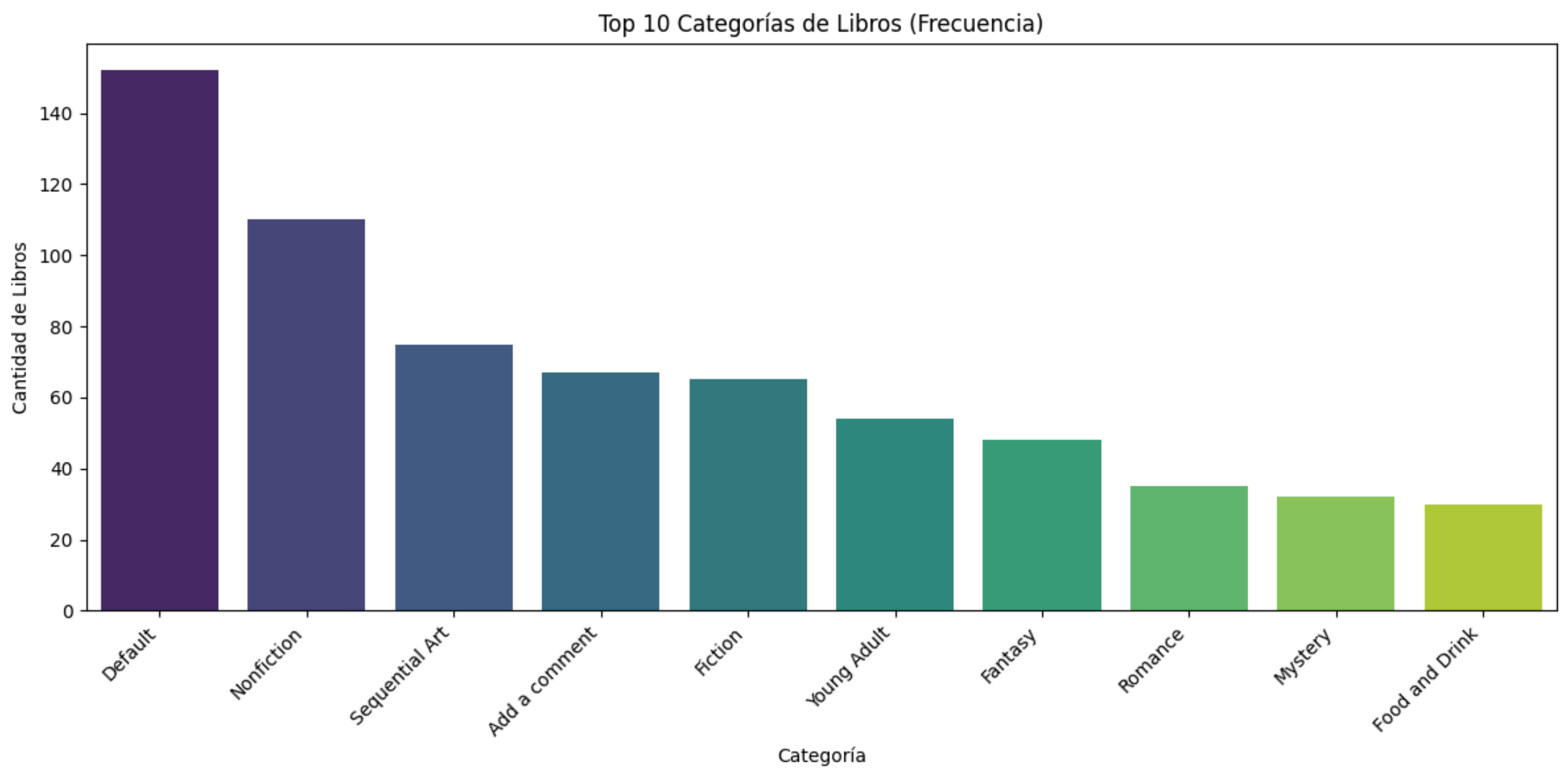
Outliers

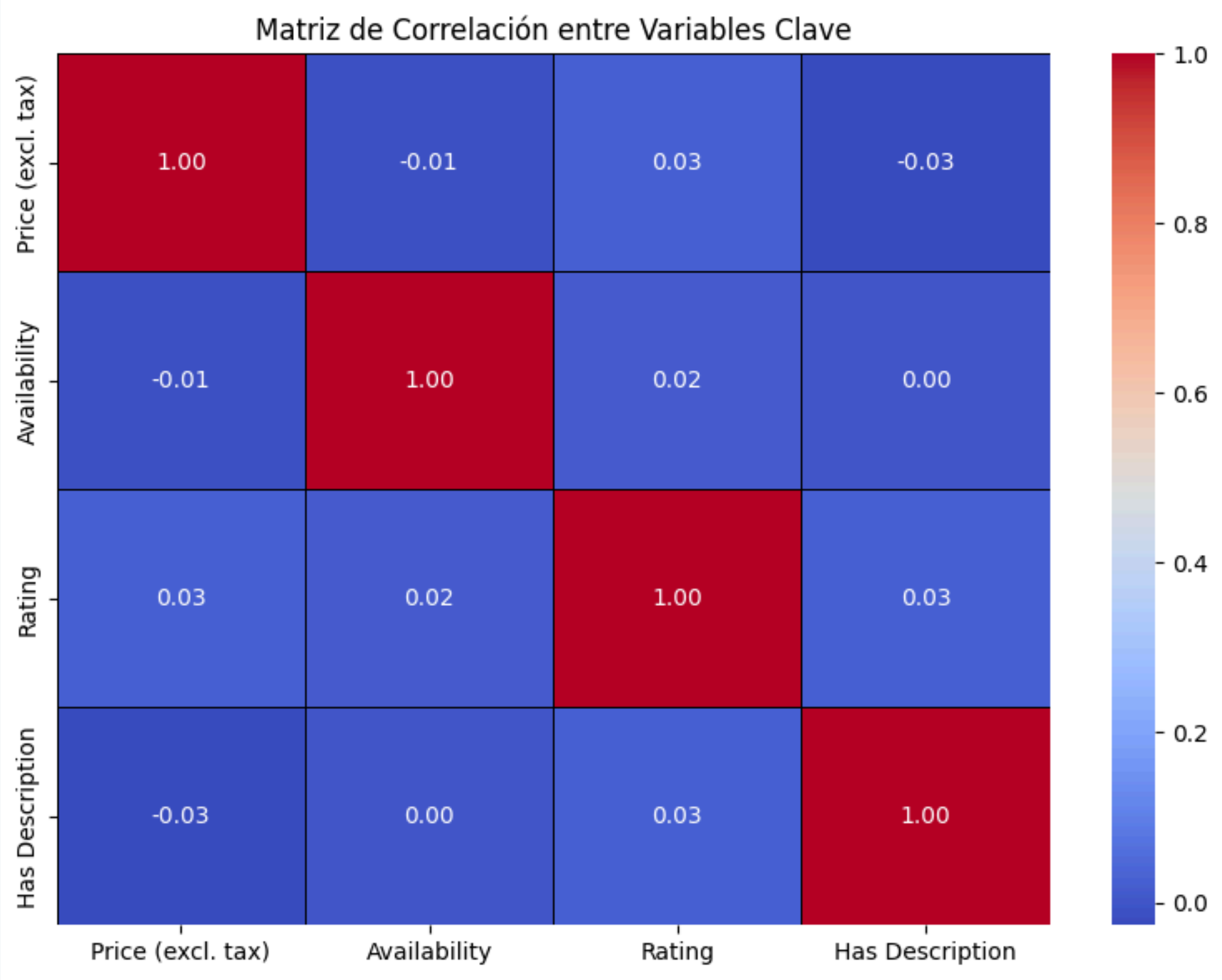
Transformaciones

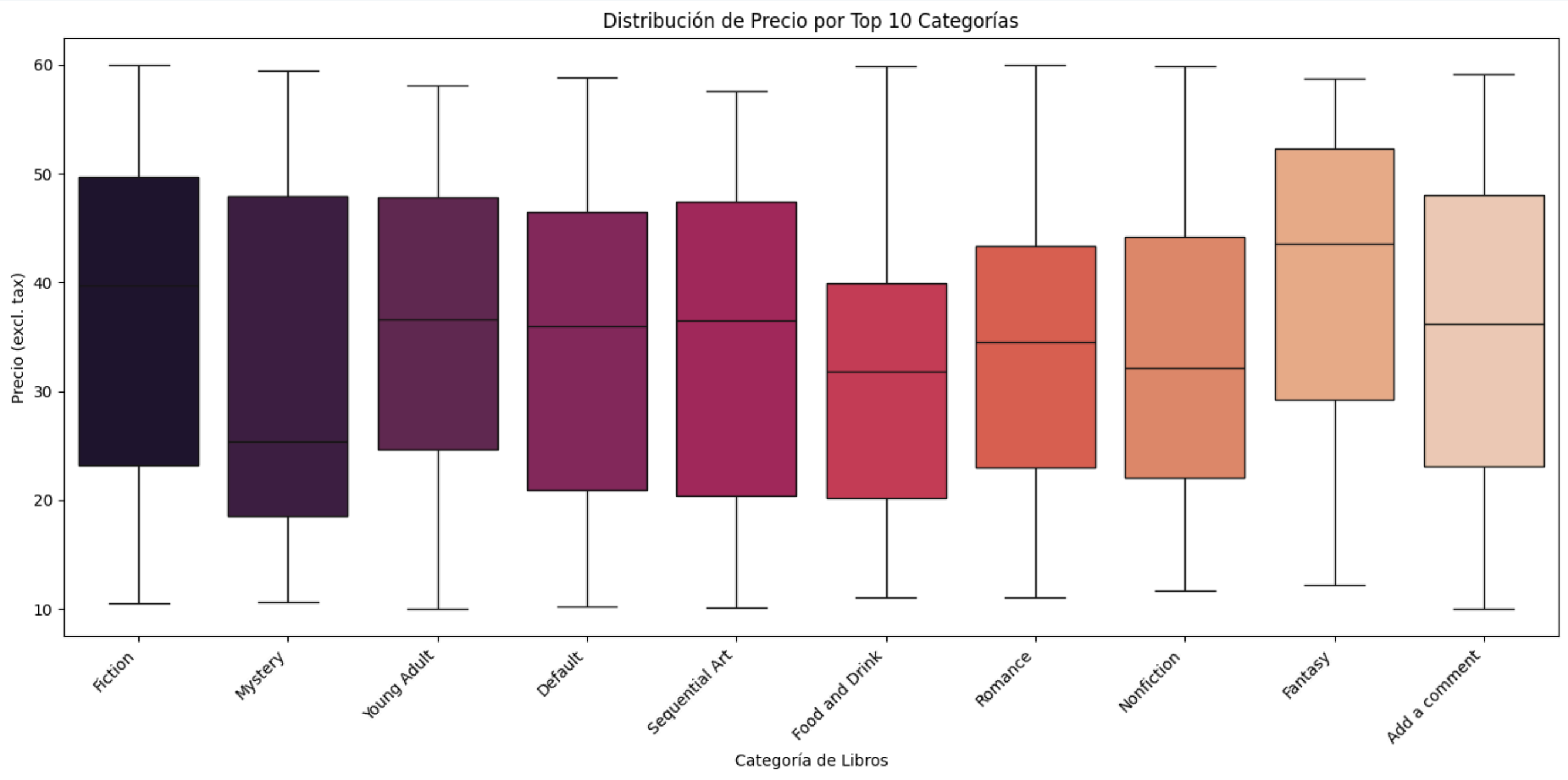
Uniformidad y Tipo de Dato

Extracción de la Disponibilidad

One-Hot Encoding







Hipótesis 1: ¿Los libros con alta calificación (≥ 4) son más caros?

Hipótesis 2: ¿Hay diferencia de precio entre Ficción y No Ficción?

Hipótesis 3: ¿Difieren las calificaciones entre Ciencia Ficción y Misterio?

Hipótesis 4: ¿Los libros Clásicos son más baratos que el promedio?

Hipótesis 5: ¿Hay diferencia de disponibilidad entre Romance y Thriller?

42 Componentes
a retener → **80%** Varianza
total

MODELO DE REGRESIÓN

Objetivo: Estimar el precio exacto de un libro

Modelos:

- Regresión Lineal Múltiple
- Regresión Ridge

Métrica: RMSE (Error Cuadrático Medio)

¿Cuál será el precio (sin impuestos) de un libro en función de su rating, disponibilidad y categoría?

MODELO DE CLASIFICACIÓN

Objetivo: Etiquetar cada libro como "caro" o "no caro" según el precio mediano.

Modelos:

- Regresión Logística
- Máquinas de Vectores de Soporte (SVC)
- K-Vecinos más Cercanos (KNN)


Métrica: Accuracy

¿Un libro será caro
(precio por encima de la
mediana) o será no
caro?

Fase 1: Reporte Ejecutivo

Fase 2: Dashboard Interactivo

Fase 3: Sistema de Soporte a la Decisión



EL FARO | Conclusiones y Retroalimentación