

Presentado por:
Angélica Corrales y Keren López

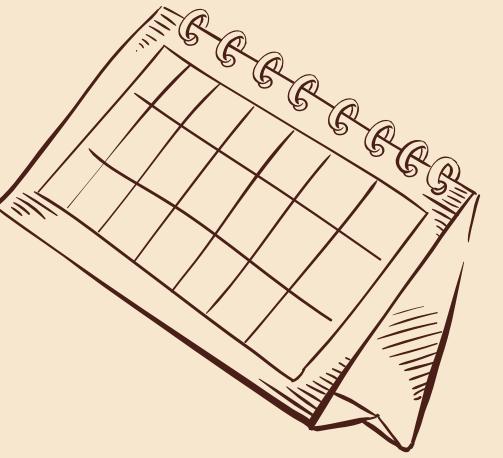


Recomendaciones de prendas de vestir

Inteligencia Artificial, Universidad Icesi



Agenda



Contexto del negocio

¿Qué es H&M?
¿Cuál es el objetivo del negocio?
Indicadores de éxito y beneficios

Antecedentes

Datos disponibles y exploración

Composición del dataset (.csv y variables)
Exploración y limpieza de datos

Modelos implementados

Despliegue

01

02

03

04

05

Contexto del negocio

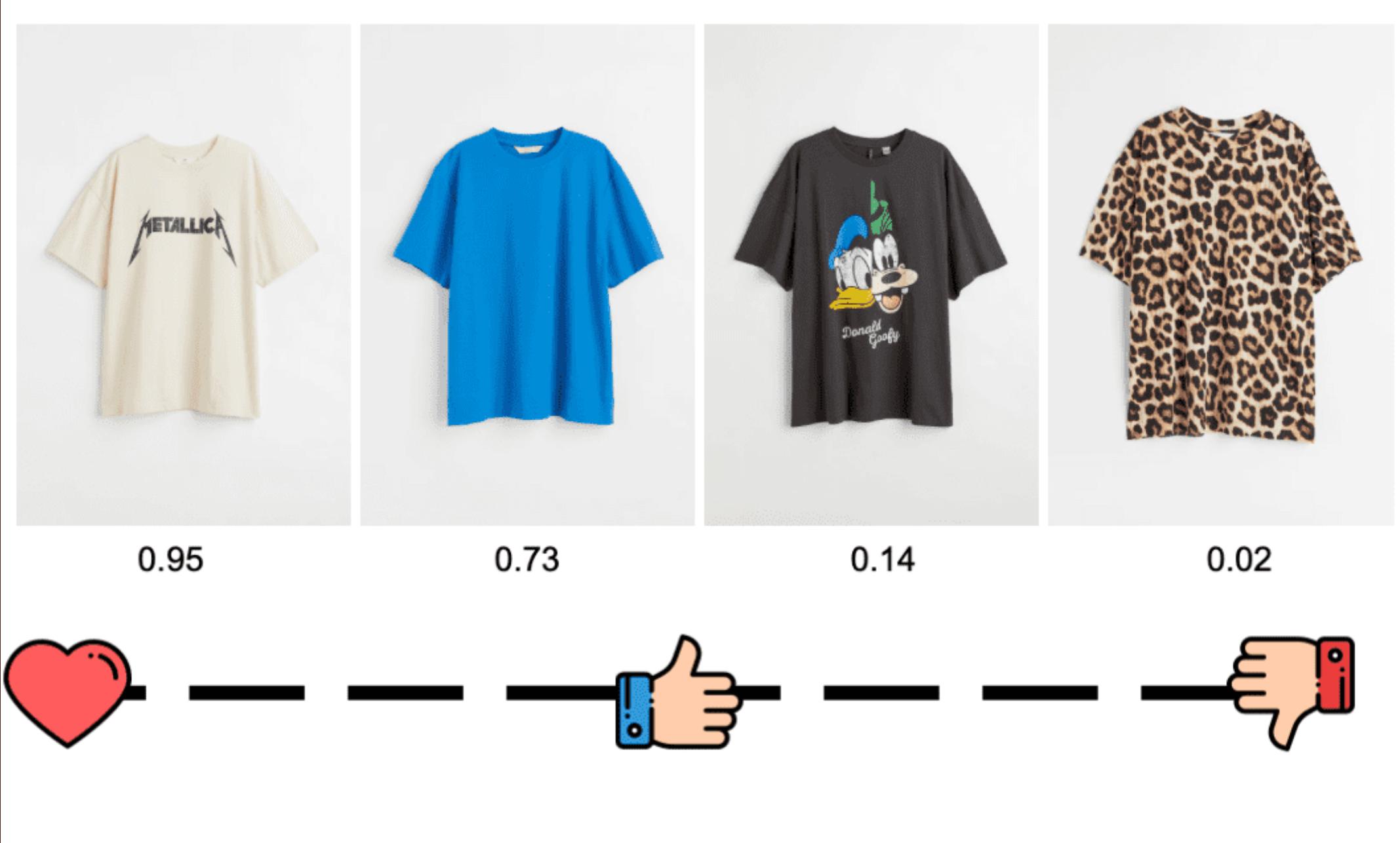
¿Qué es H&M?

Cuando se trata de la industria de la moda es una de las marcas más importantes a nivel mundial, cuenta con 53 mercados en línea y aproximadamente 4850 tiendas





¿Cuál es el objetivo del negocio?



Se busca realizar recomendaciones de artículos de ropa disponibles, de acuerdo con el historial de compras de los clientes. Esto con el fin de mejorar la experiencia de compra de las personas y así, aumentar las ventas.



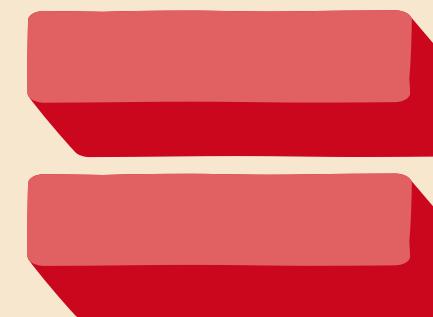


Indicadores de éxito y beneficios

De reducción de las personas que ingresan al sitio web/tiendas y no realizan ninguna compra.

De satisfacción de los clientes con el sistema de recomendación planteado.

De aumento de ventas físicas/línea a partir de la fecha de implementación del sistema de recomendación.



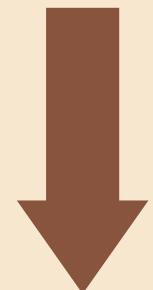
Aumento de ingresos



Mejor experiencia de compra



Mejor control de la producción de productos con baja demanda



Antecedentes

Sistemas recomendadores

Filtrado colaborativo

Identificar perfiles similares y aprender de los datos para recomendar productos de manera individual.

Basado en contenido

Utiliza los atributos de un artículo para recomendar otros similares a las preferencias del usuario.

Popularidad

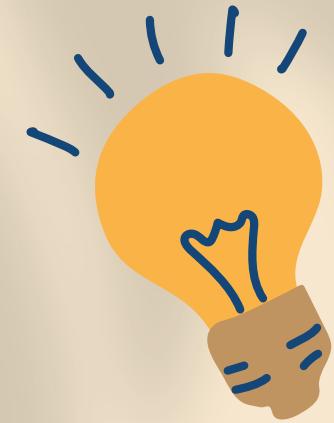
Por ejemplo, los artículos que más se han vendido.

02

Antecedentes



03



Artículos

25 columnas de variables categóricas y 105542 filas



Clientes

7 columnas de variables categóricas y 1.371.980 filas



Transacciones

5 columnas de variables categóricas y 31.788.324 filas

Datos disponibles

SMART Question

¿Cómo podemos recomendar artículos de ropa a los clientes de H&M con la ayuda de un modelo de clasificación no supervisado, teniendo en cuenta sus historiales de compra?



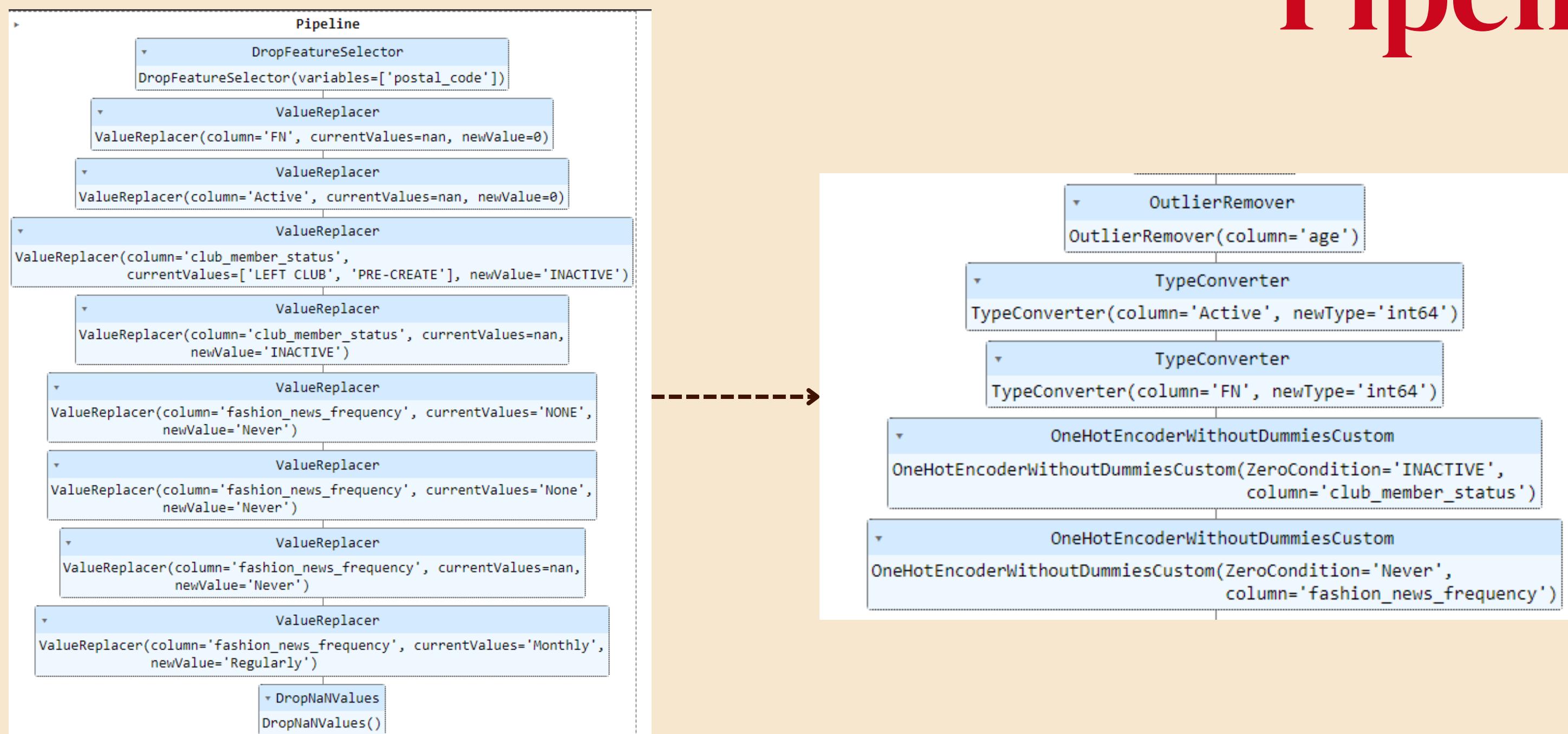
Análisis exploratorio



Google Colab

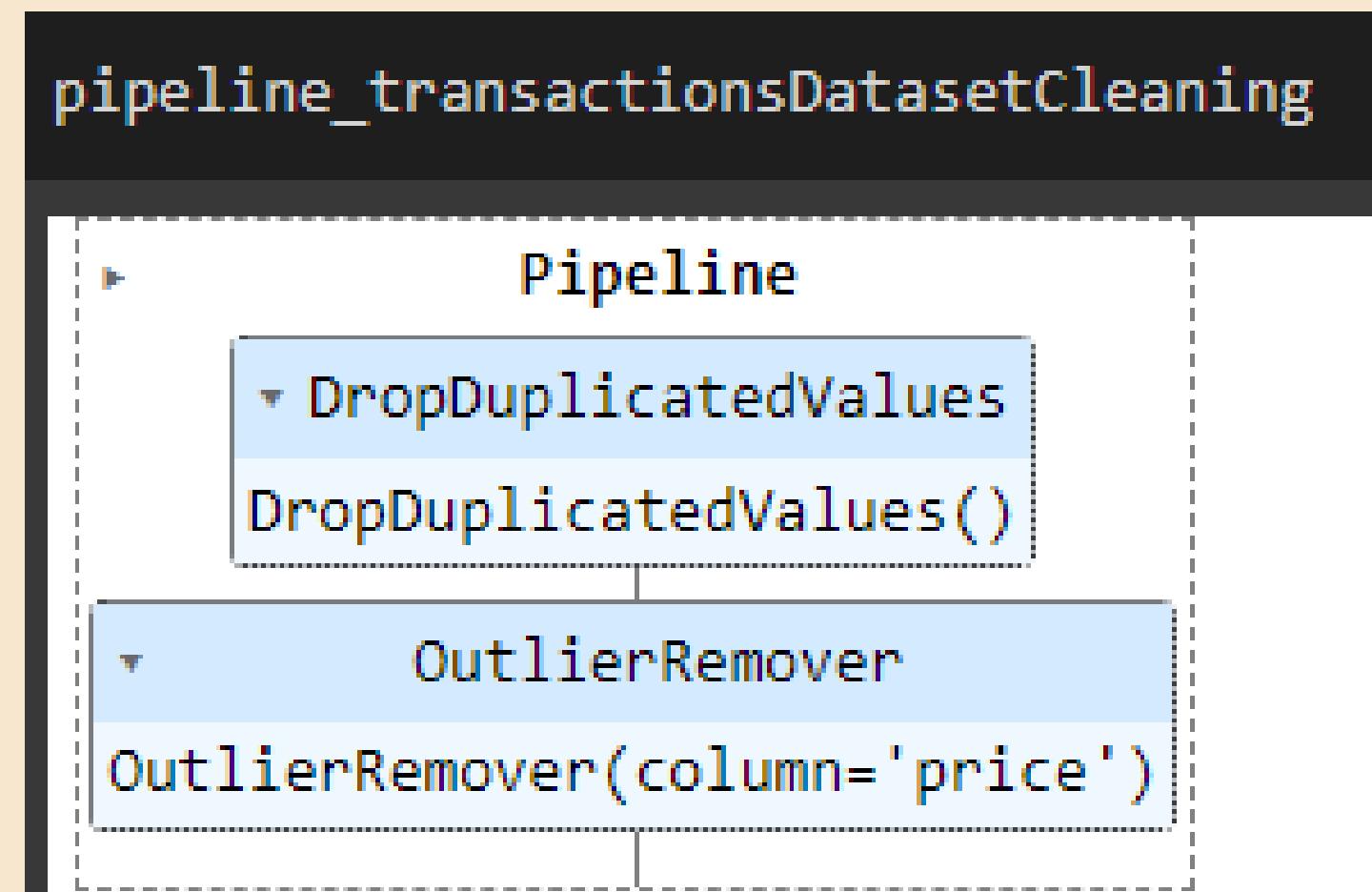
Limpieza del dataset de Clientes

Pipelines



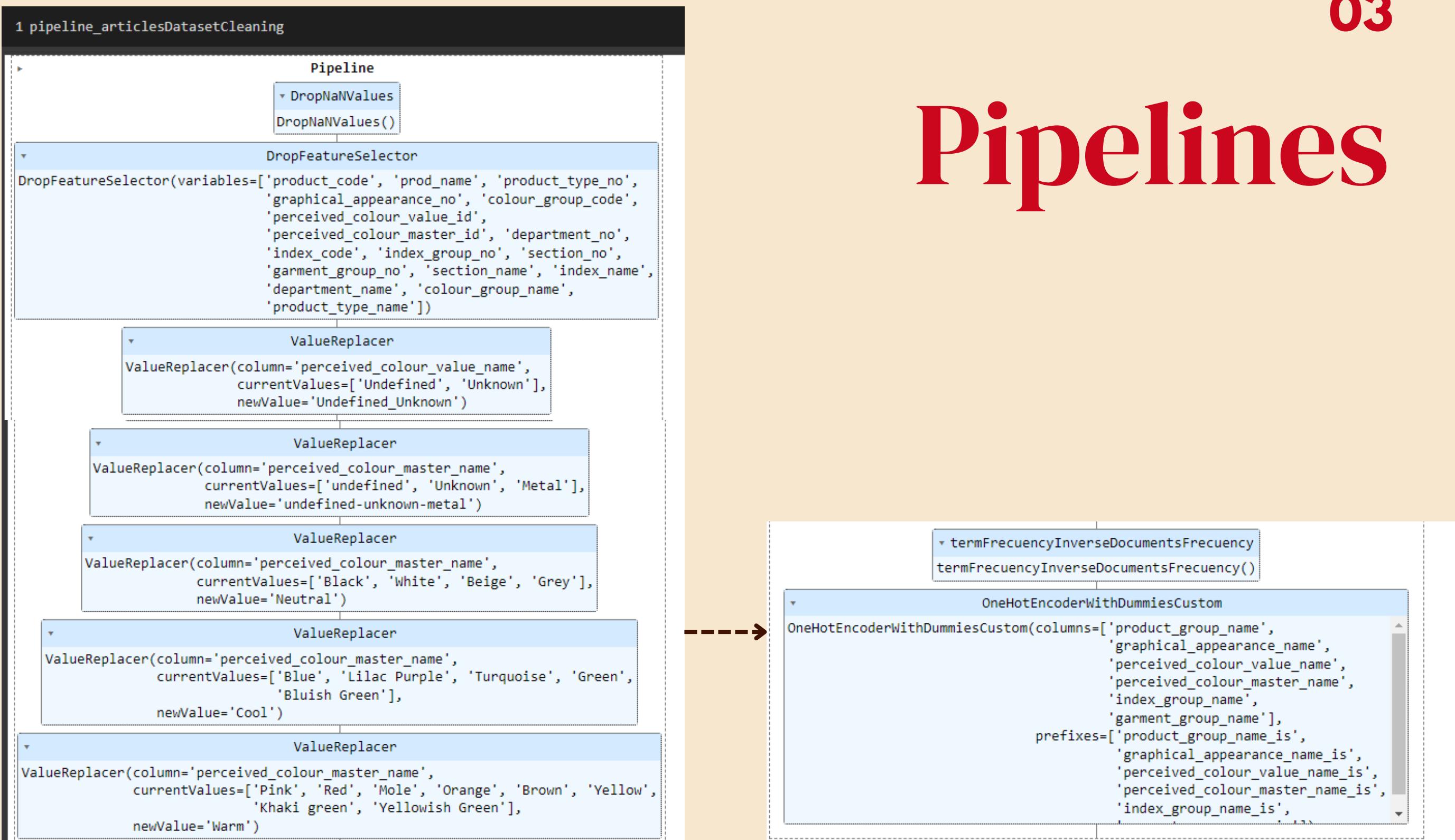
Limpieza del dataset de Transacciones

Pipelines



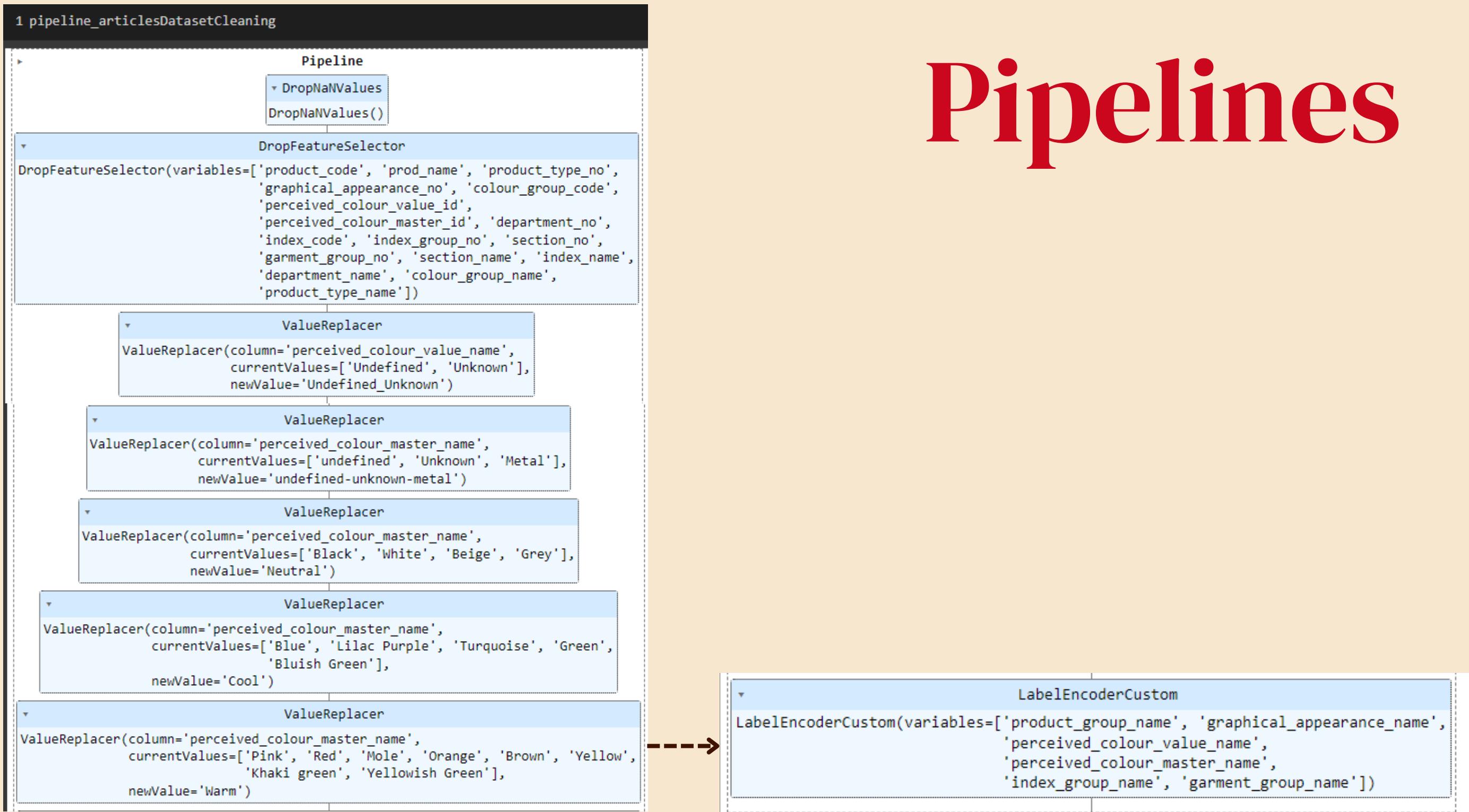
Artículos

K - Means

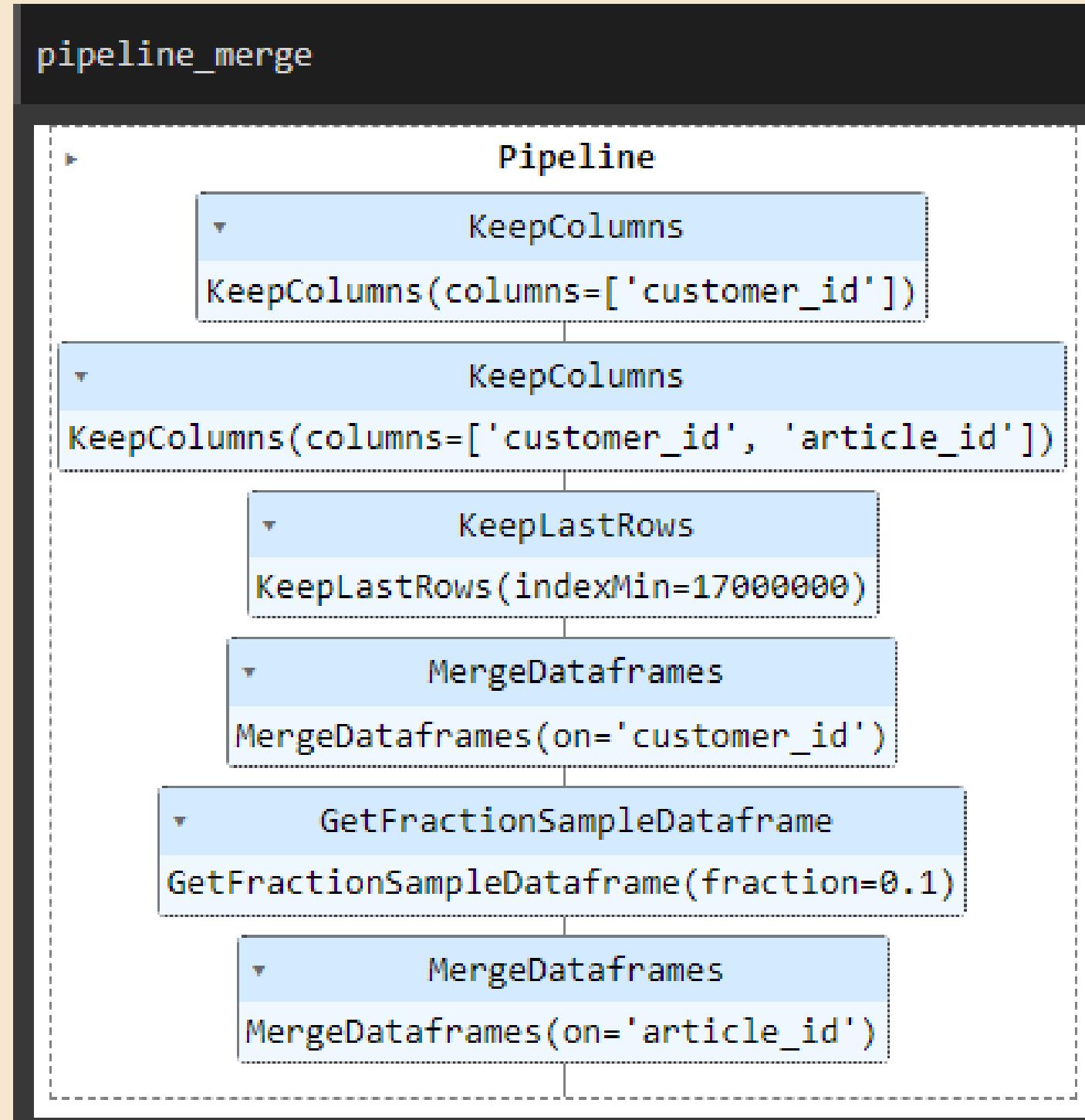


Artículos

K - Modes



Unión de conjuntos de datos



Pipelines

Modelos implementados

K-Means

- Content-Based Filtering.

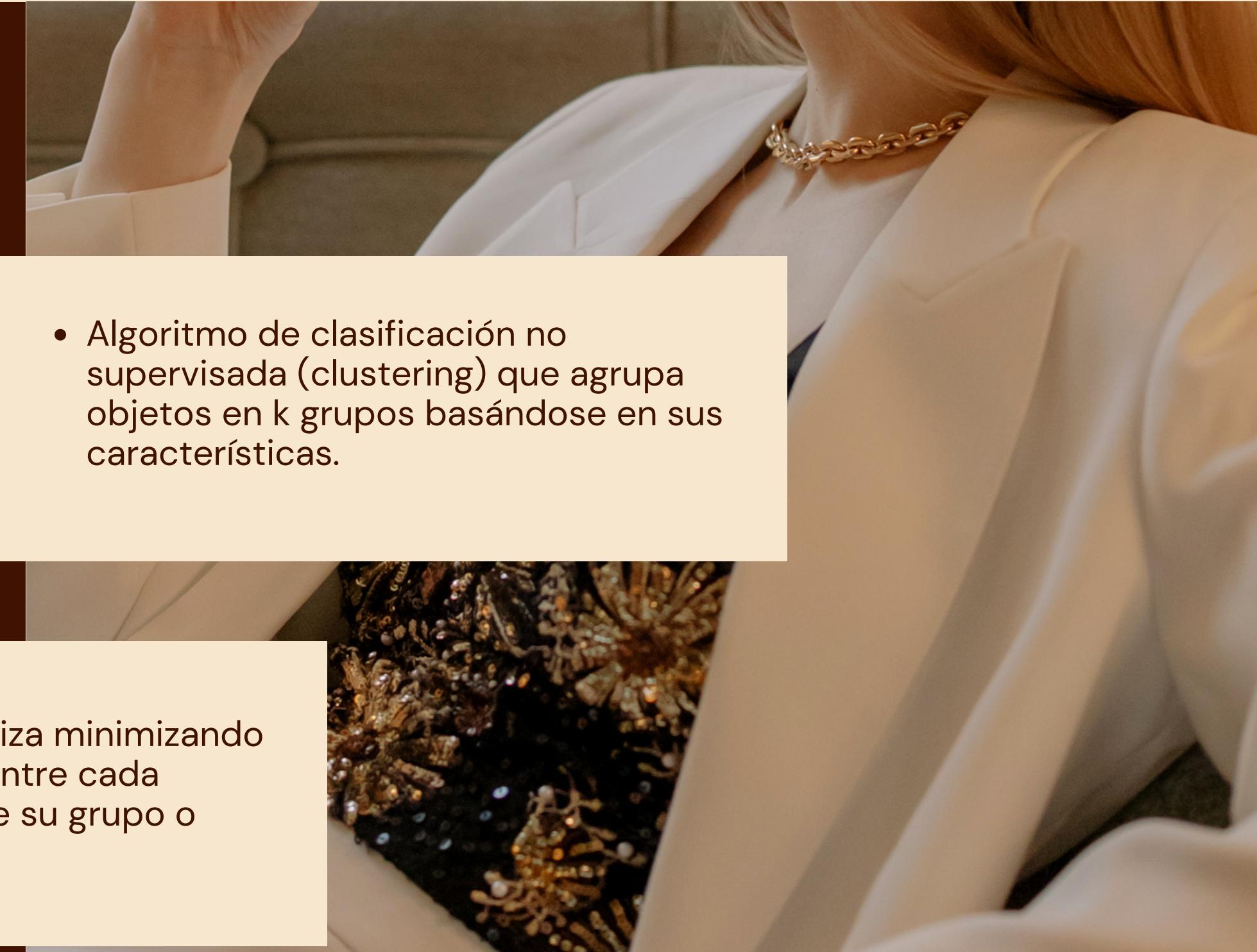


- El agrupamiento se realiza minimizando la suma de distancias entre cada objeto y el centroide de su grupo o cluster.

- Retroalimentación de usuario



- Algoritmo de clasificación no supervisada (clustering) que agrupa objetos en k grupos basándose en sus características.



Term frequency – Inverse document frequency (TFIDF)

- Para procesar el lenguaje natural, el texto debe representarse como una característica numérica

- Convierte el texto en un vector de características contando la aparición de palabras



- La importancia de una palabra en el texto es de gran importancia en la recuperación de información.

PCA



- Técnica de aprendizaje no supervisado para reducir la dimensionalidad de los datos.

- Aumenta la interpretabilidad y, al mismo tiempo, minimiza la pérdida de información.

- Ayuda a encontrar las características más significativas de un conjunto de datos.

Modelos implementados

K-Modes

- Content-Based Filtering.



- Algoritmo de clustering para agrupar datos similares en grupos basados en sus atributos categóricos

- Retroalimentación de usuario



- Identifica las modas o los valores más frecuentes dentro de cada clúster para determinar su centroide.



Despliegue



Github



Referencias

- H&M Personalized Fashion Recommendations | Kaggle. (s. f.). <https://www.kaggle.com/competitions/h-and-m-personalized-fashion-recommendations/data>
- Aprende Machine Learning. (2019). Sistemas de Recomendación. Aprende Machine Learning. <https://www.aprendemachinelearning.com/sistemas-de-recomendacion/>
- Biswal, A. (2023). Principal Component Analysis in Machine Learning: Complete Guide. Simplilearn.com. <https://www.simplilearn.com/tutorials/machine-learning-tutorial/principal-component-analysis>
- Bonthu, H. (2023). KModes Clustering Algorithm for Categorical data. Analytics Vidhya. <https://www.analyticsvidhya.com/blog/2021/06/kmodes-clustering-algorithm-for-categorical-data/>
- Sistemas de recomendación: la clave del éxito en el eCommerce. (2017, 9 abril). IONOS Digital Guide. <https://www.ionos.es/digitalguide/online-marketing/vender-en-internet/los-sistemas-de-recomendacion-en-el-eCommerce/>
- Universidad de Oviedo. (s. f.). El algoritmo k-means aplicado a clasificación y procesamiento de imágenes. https://www.unioviedo.es/comnum/laboratorios_py/kmeans/kmeans.html
- What is a Recommendation System? (s. f.). NVIDIA Data Science Glossary. <https://www.nvidia.com/en-us/glossary/data-science/recommendation-system/>



Gracias