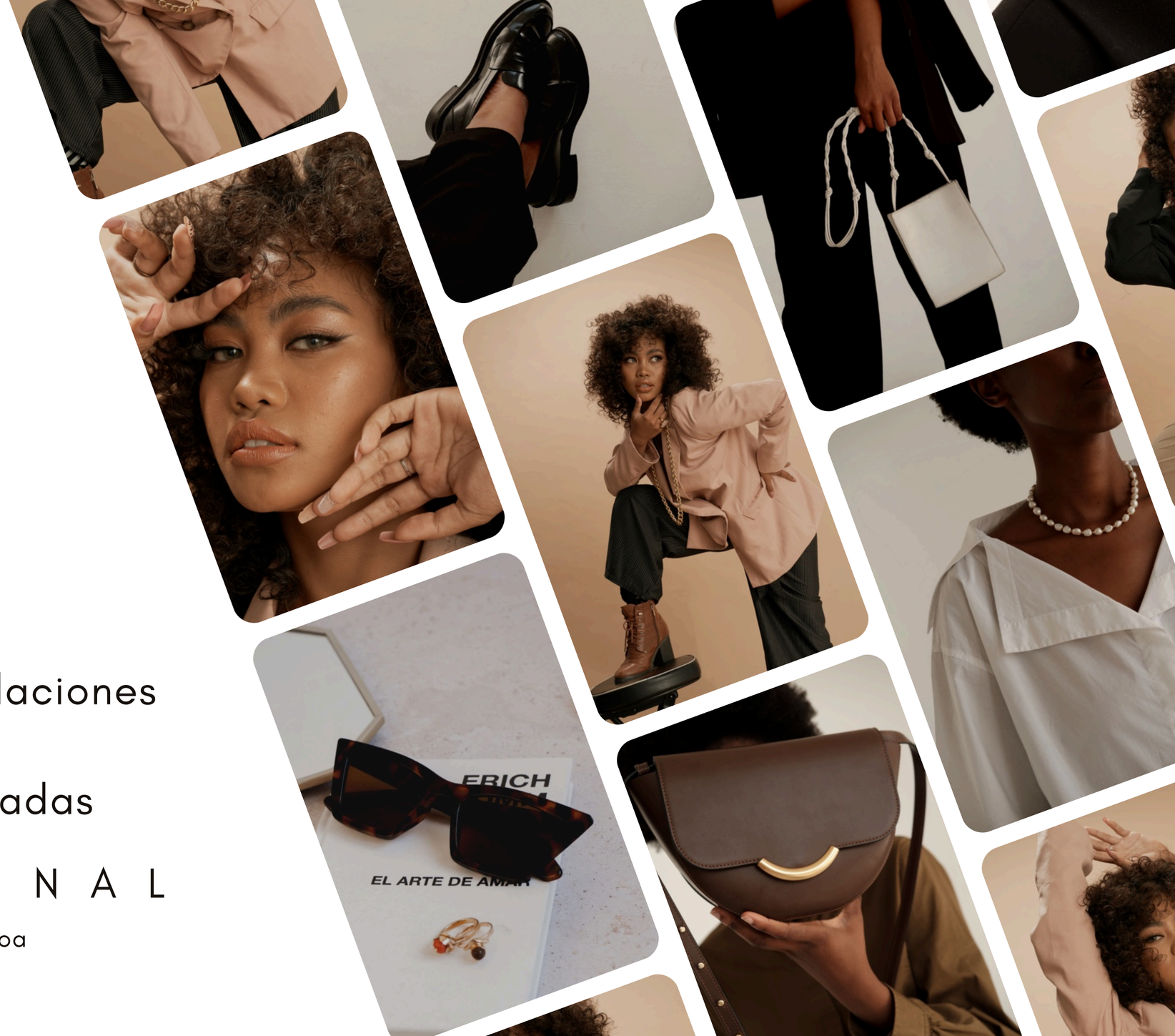
Recomendaciones de Moda Personalizadas

PROYECTO FINAL

- Sofía Escobar
- Elisa Samayoa

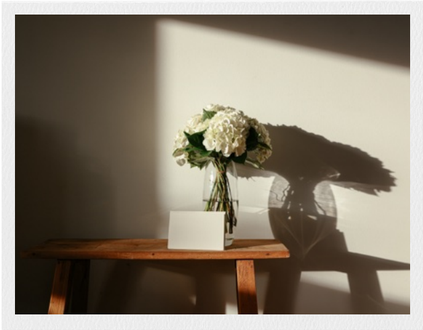


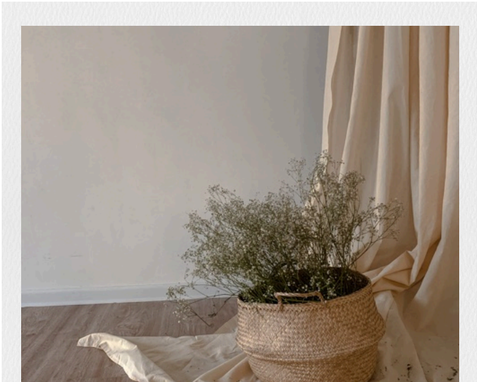
HENNES & MAURITZ

H&M es una **cadena multinacional sueca de tiendas de ropa**, con presencia en Europa, Oriente Próximo, África, Asia y Latinoamérica. Para finales de 2021, H&M contaba con **más de 4800 tiendas físicas y 53 en línea.**

Aunque las tiendas en línea ofrecen una amplia selección de productos para explorar, muchas veces, los clientes no pueden encontrar rápidamente lo que están buscando y pueden no realizar una compra.

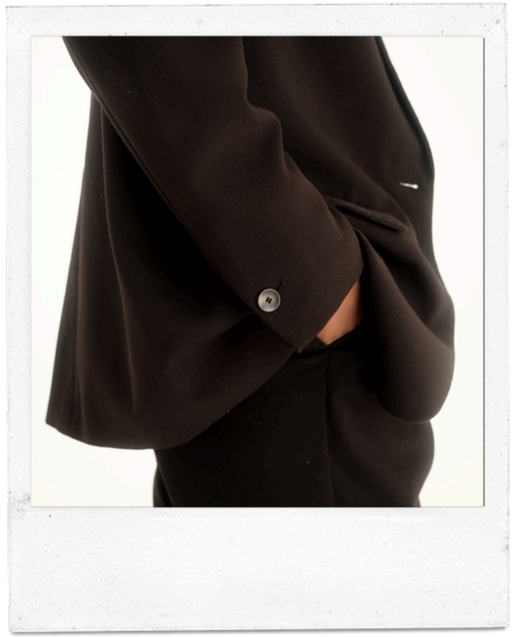
Para mejorar la experiencia de compra, las recomendaciones de productos son clave, sobre todo, para ayudar a los clientes a tomar las decisiones correctas, lo que también tiene implicaciones positivas para la sostenibilidad, ya que reduce las devoluciones y, por lo tanto, minimiza las emisiones del transporte.

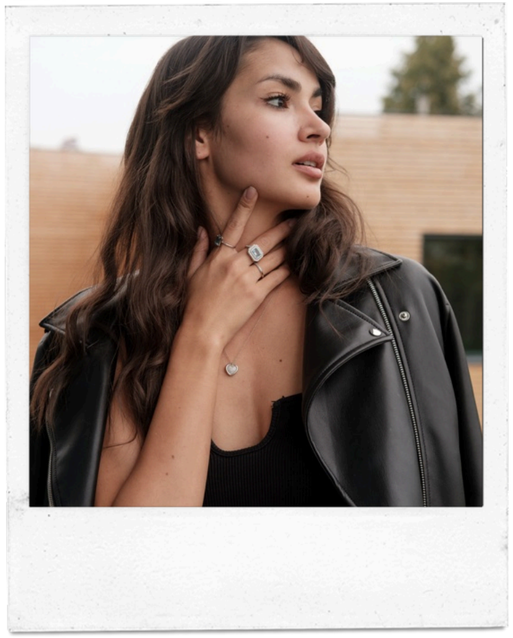




Objetivo

Desarrollar recomendaciones de productos basadas en datos de transacciones anteriores, así como en metadatos de clientes y productos.





DATOS DISPONIBLES

1. imágenes/ - carpeta con imágenes con una carpeta de imágenes correspondiente a cada artículo_id;
2. artículos.csv: metadatos detallados para cada artículo_id disponible para su compra
3. clientes.csv: metadatos para cada id_cliente en el conjunto de datos
4. sample_submission.csv: un archivo de envío de muestra en el formato correcto
5. transacciones_train.csv: los datos de capacitación, que consisten en las compras de cada cliente para cada fecha, así como información adicional.

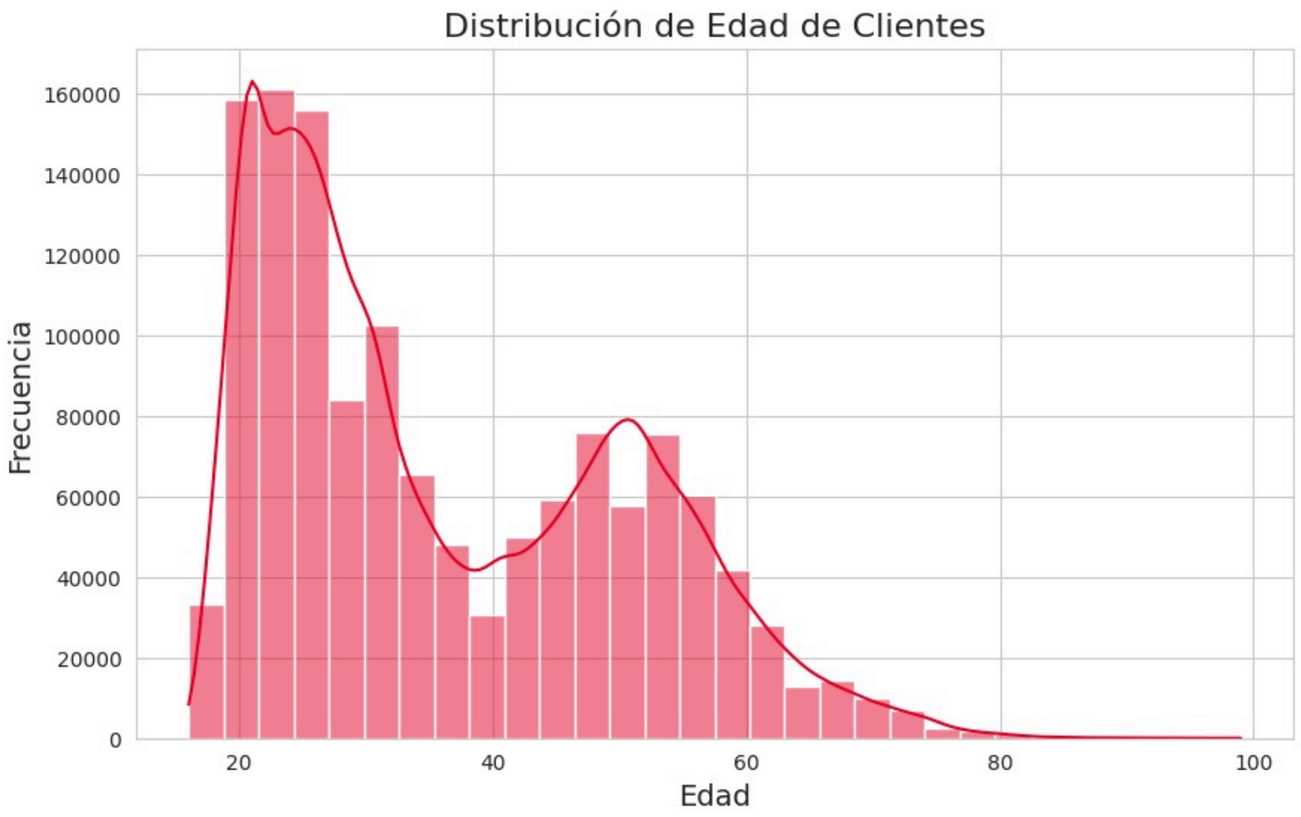
1. ANÁLISIS EXPLORATORIO

Clientes, artículos y transacciones

Clientes

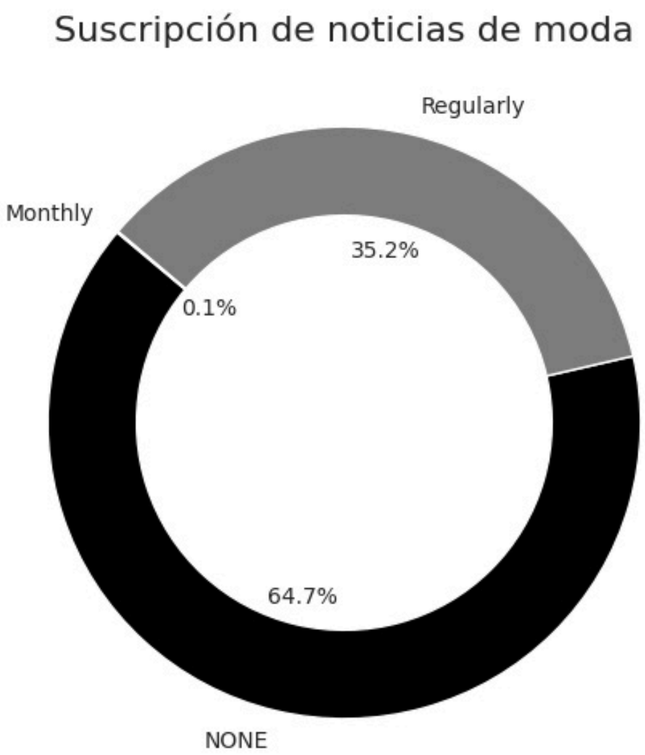
- Customer_id y postal_code cifrados
- FN y Active se excluyeron del análisis
- Eliminación de nulls en el resto de variables

```
RangeIndex: 1371980 entries, 0 to 1371979
Data columns (total 7 columns):
 #   Column                  Non-Null Count    Dtype
---  ------                  --------------    -----
 0   customer_id             1371980 non-null  object
 1   FN                      476930 non-null   float64
 2   Active                  464404 non-null   float64
 3   club_member_status      1365918 non-null  object
 4   fashion_news_frequency  1355969 non-null  object
 5   age                     1356119 non-null  float64
 6   postal_code             1371980 non-null  object
```



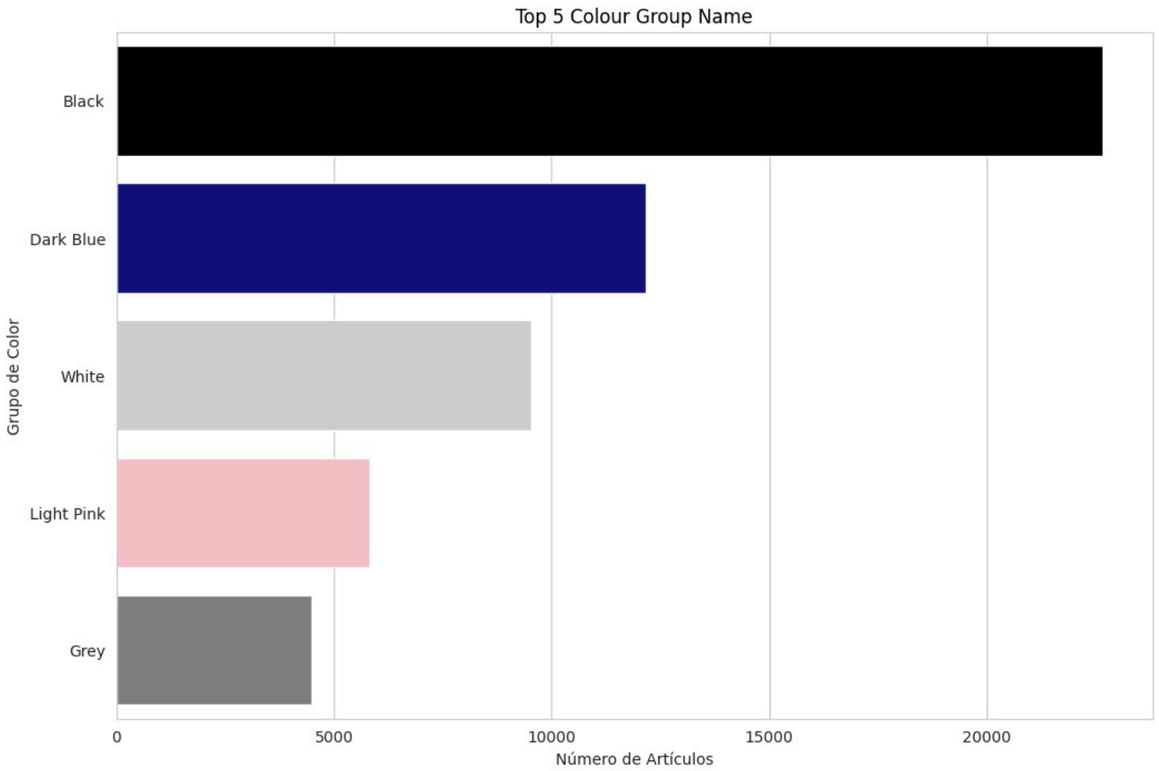
Clientes

Estado de Membresía del Club
PRE-CREATE
LEFT CLUB
6.8%
0.0%
93.2%
ACTIVE


Suscripción de noticias de moda
Regularly
Monthly
35.2%
0.1%
64.7%
NONE

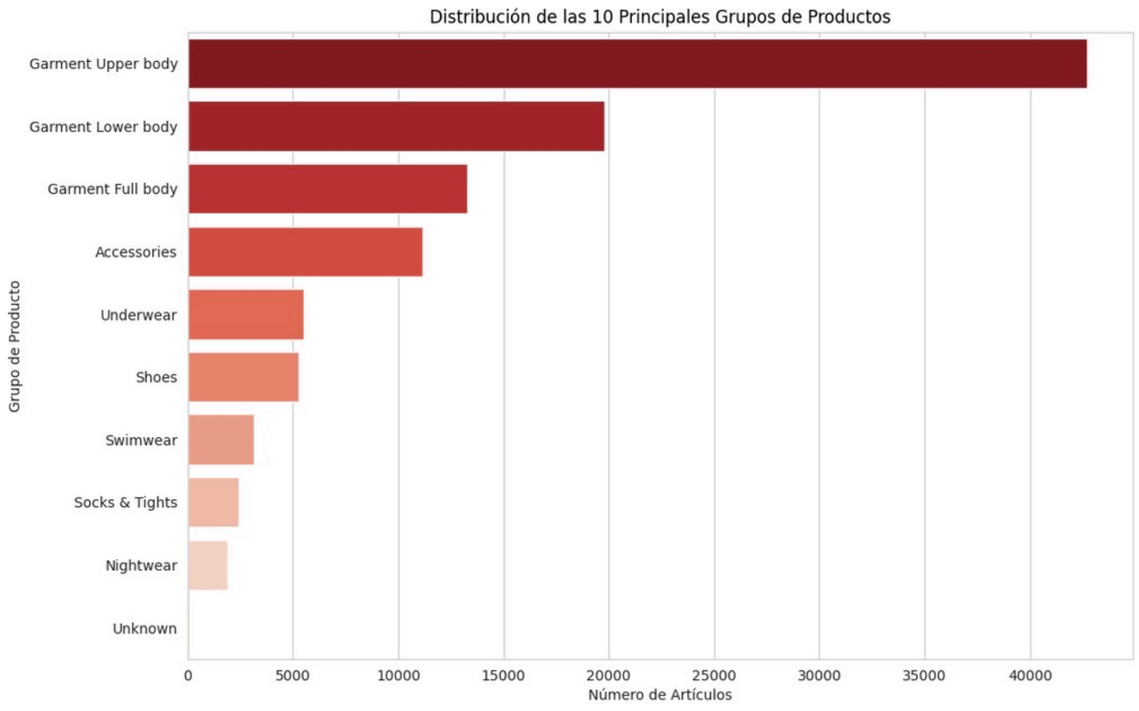
Artículos

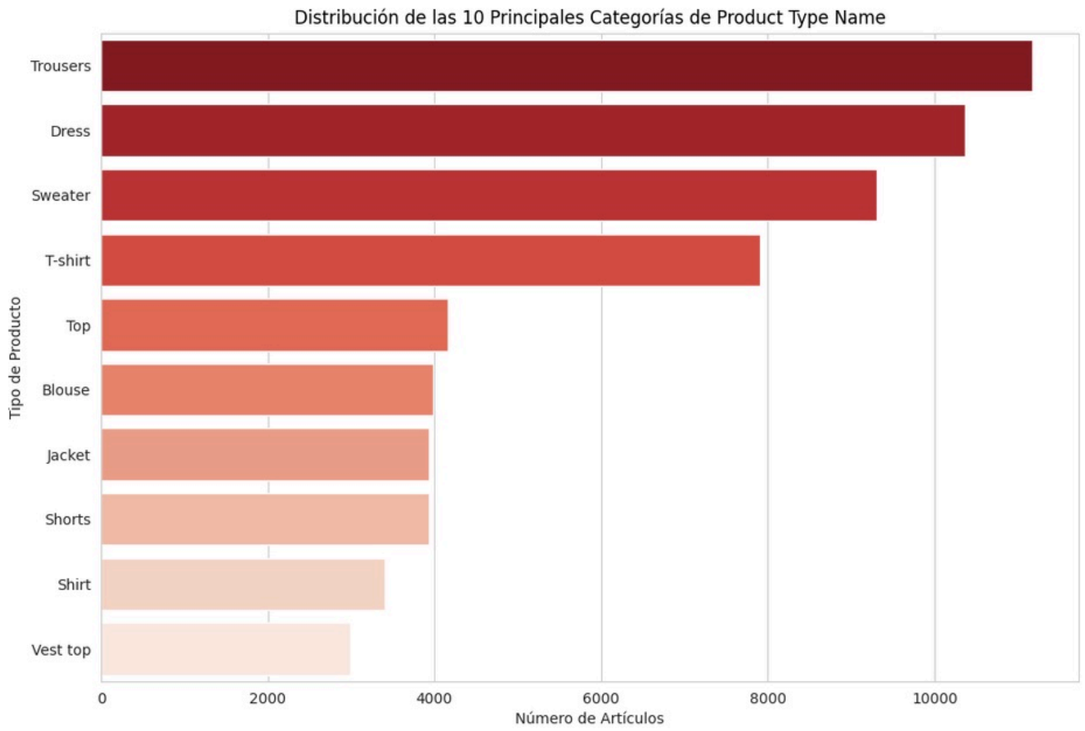
```
RangeIndex: 105542 entries, 0 to 105541
Data columns (total 25 columns):
 #   Column                        Non-Null Count   Dtype
---  ------                        --------------   -----
 0   article_id                    105542 non-null  int64
 1   product_code                  105542 non-null  int64
 2   prod_name                     105542 non-null  object
 3   product_type_no               105542 non-null  int64
 4   product_type_name             105542 non-null  object
 5   product_group_name            105542 non-null  object
 6   graphical_appearance_no       105542 non-null  int64
 7   graphical_appearance_name     105542 non-null  object
 8   colour_group_code             105542 non-null  int64
 9   colour_group_name             105542 non-null  object
 10  perceived_colour_value_id     105542 non-null  int64
 11  perceived_colour_value_name   105542 non-null  object
 12  perceived_colour_master_id    105542 non-null  int64
 13  perceived_colour_master_name  105542 non-null  object
 14  department_no                 105542 non-null  int64
 15  department_name               105542 non-null  object
 16  index_code                    105542 non-null  object
 17  index_name                    105542 non-null  object
 18  index_group_no                105542 non-null  int64
 19  index_group_name              105542 non-null  object
 20  section_no                    105542 non-null  int64
 21  section_name                  105542 non-null  object
 22  garment_group_no              105542 non-null  int64
 23  garment_group_name            105542 non-null  object
 24  detail_desc                   105126 non-null  object
```



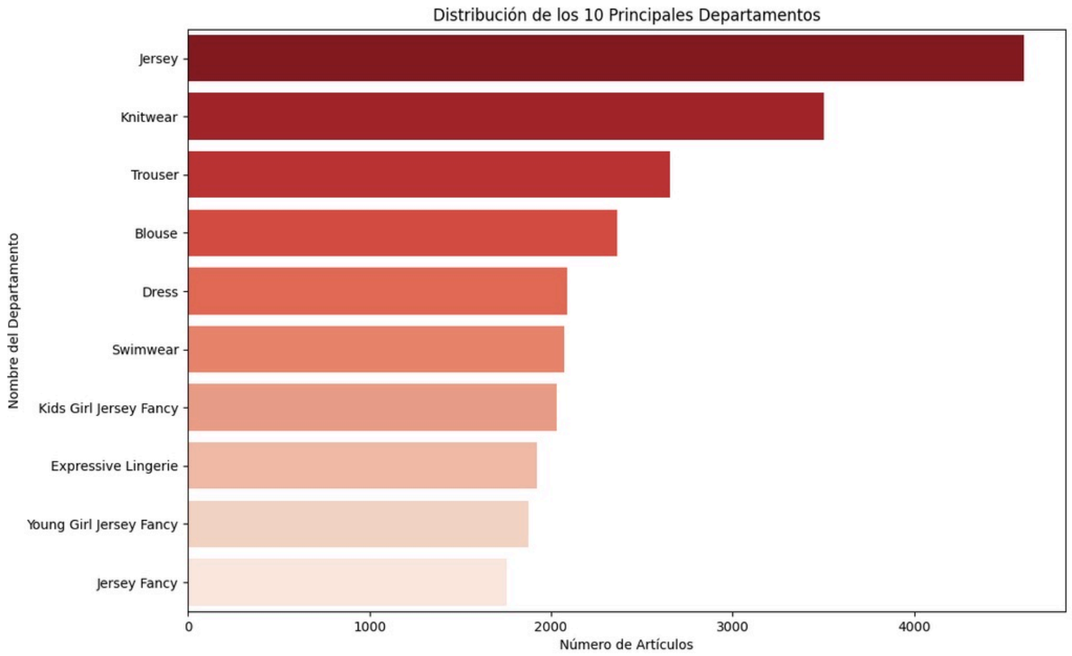
105,542 artículos únicos

Artículos


Distribución de las 10 Principales Grupos de Productos
Garment Upper body
Garment Lower body
Garment Full body
Accessories
Underwear
Shoes
Swimwear
Socks & Tights
Nightwear
Unknown
Grupo de Producto
0
5000
10000
15000
20000
25000
30000
35000
40000
Número de Artículos


Distribución de las 10 Principales Categorías de Product Type Name
Trousers
Dress
Sweater
T-shirt
Top
Blouse
Jacket
Shorts
Shirt
Vest top
Tipo de Producto
0
2000
4000
6000
8000
10000
Número de Artículos

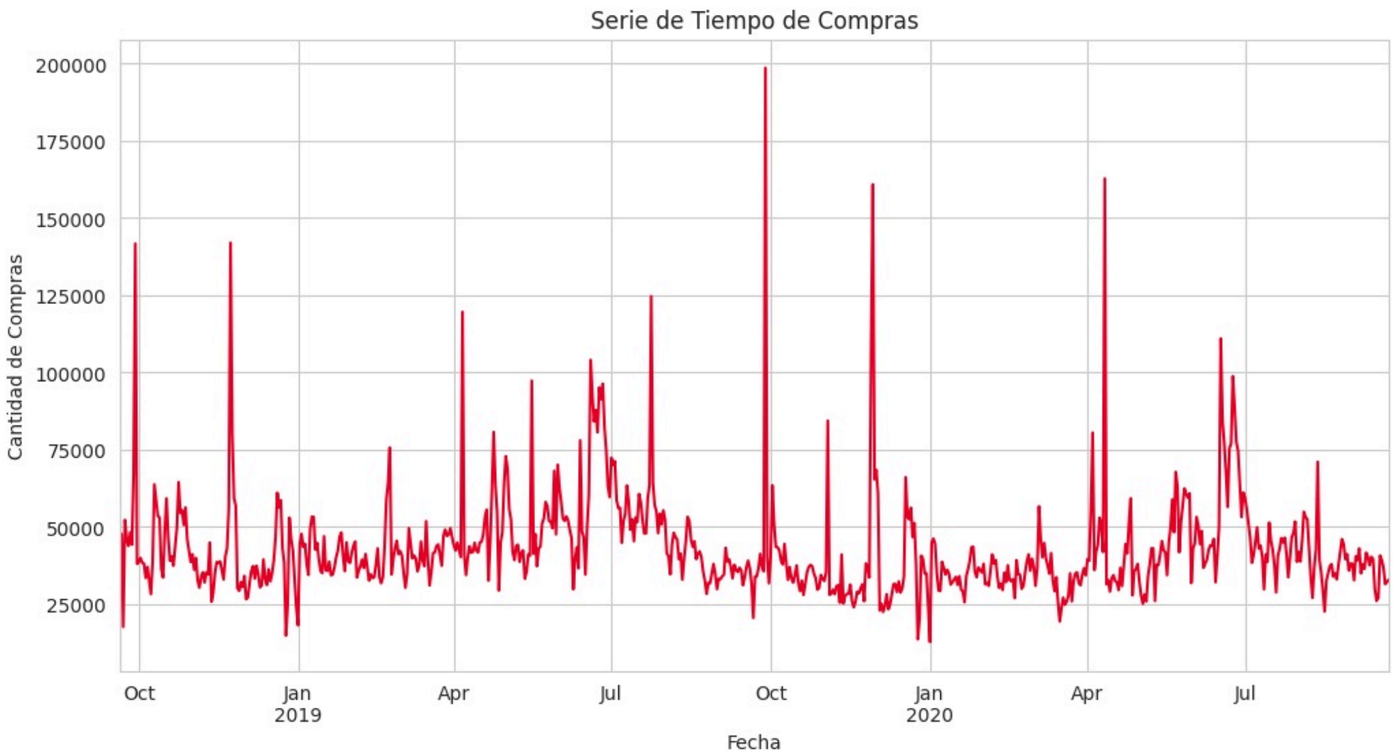
Artículos


Distribución de los 10 Principales Departamentos
Jersey
Knitwear
Trouser
Blouse
Dress
Swimwear
Kids Girl Jersey Fancy
Expressive Lingerie
Young Girl Jersey Fancy
Jersey Fancy
Nombre del Departamento
0
1000
2000
3000
4000
Número de Artículos

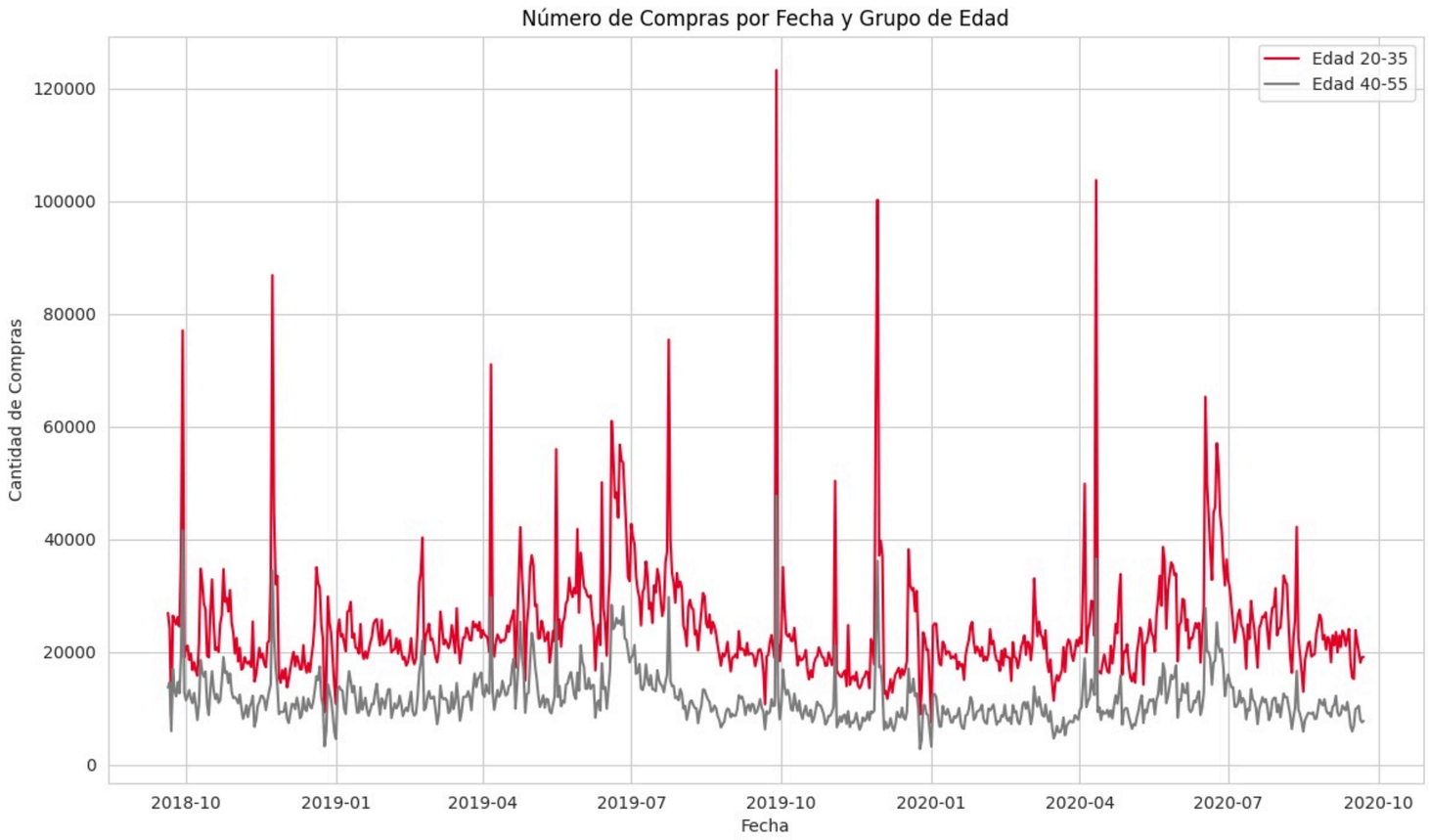
WordCloud de Detail Desc
side
pocket
zip fly
front pocket
shoulder
strap
elasticated
waist
long
sleeve
ribbing
around
back
pocket
cuff
Long sleeved
top
button
sweatshirt
fabric
cotton
jersey
press
stud
rounded
hem
V
neck
fine knit
woven fabric
neckline cuffs
imitation leather

Transacciones

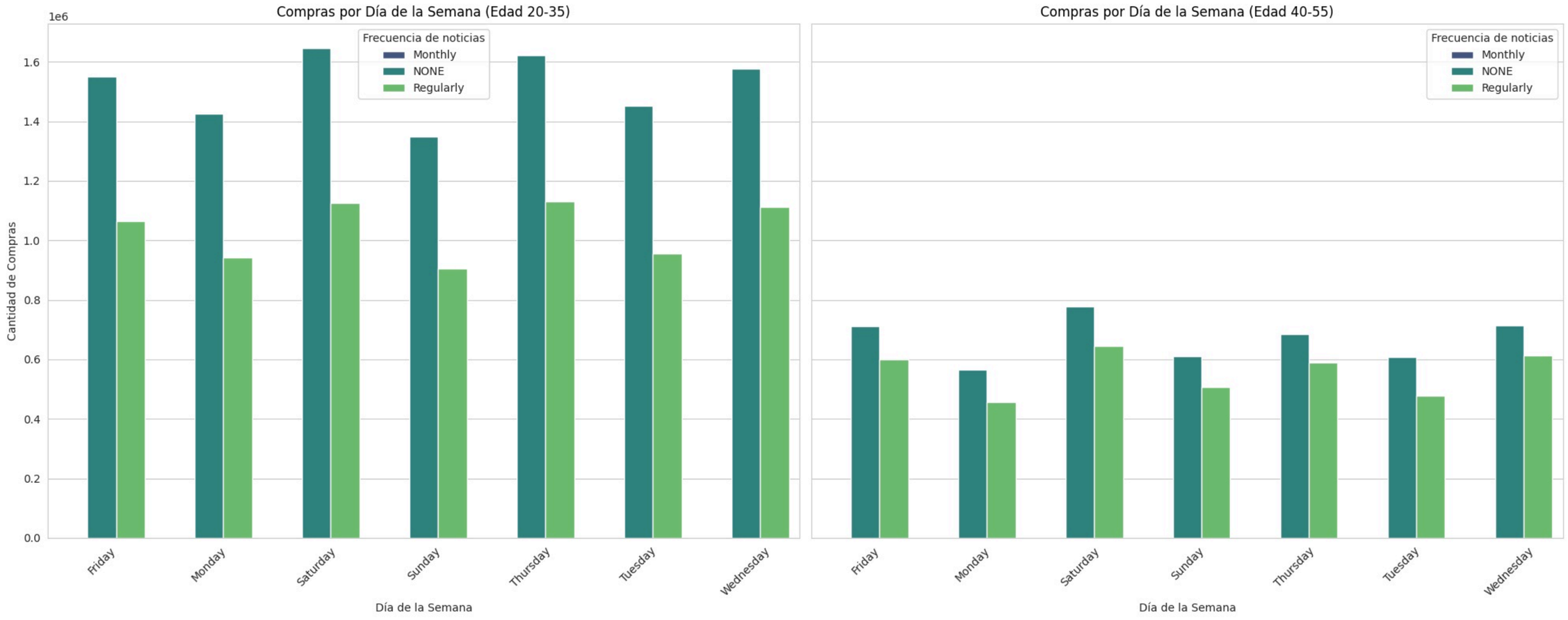
```
RangeIndex: 31788324 entries, 0 to 31788323
Data columns (total 5 columns):
 #   Column           Dtype
---  ------           -----
 0   t_dat            object
 1   customer_id      object
 2   article_id       int64
 3   price            float64
 4   sales_channel_id int64
```



Transacciones



Transacciones



Transacciones

Top 5 Artículos Más Comprados Globalmente:

	article_id	prod_name	product_type_name	section_name	count
0	706016001	Jade HW Skinny Denim TRS	Trousers	Divided Collection	50287
1	706016002	Jade HW Skinny Denim TRS	Trousers	Divided Collection	35043
2	372860001	7p Basic Shaftless	Socks	Womens Nightwear, Socks & Tigh	31718
3	610776002	Tilly (1)	T-shirt	Womens Everyday Basics	30199
4	759871002	Tilda tank	Vest top	Divided Complements Other	26329

Transacciones

Información del artículo más comprado en cada año:

	year	article_id	prod_name	product_type_name	section_name	count
0	2018	399223001	Curvy Jeggings HW Ankle	Trousers	Ladies Denim	7414
1	2019	706016001	Jade HW Skinny Denim TRS	Trousers	Divided Collection	29869
2	2020	706016001	Jade HW Skinny Denim TRS	Trousers	Divided Collection	18208

2018

2019 y 2020

Transacciones

Artículo más comprado cada año para clientes de 20-35 años:

	year	article_id	prod_name	product_type_name	section_name	count
0	2018	399223001	Curvy Jeggings HW Ankle	Trousers	Ladies Denim	4824
1	2019	706016001	Jade HW Skinny Denim TRS	Trousers	Divided Collection	19371
2	2020	706016001	Jade HW Skinny Denim TRS	Trousers	Divided Collection	11750

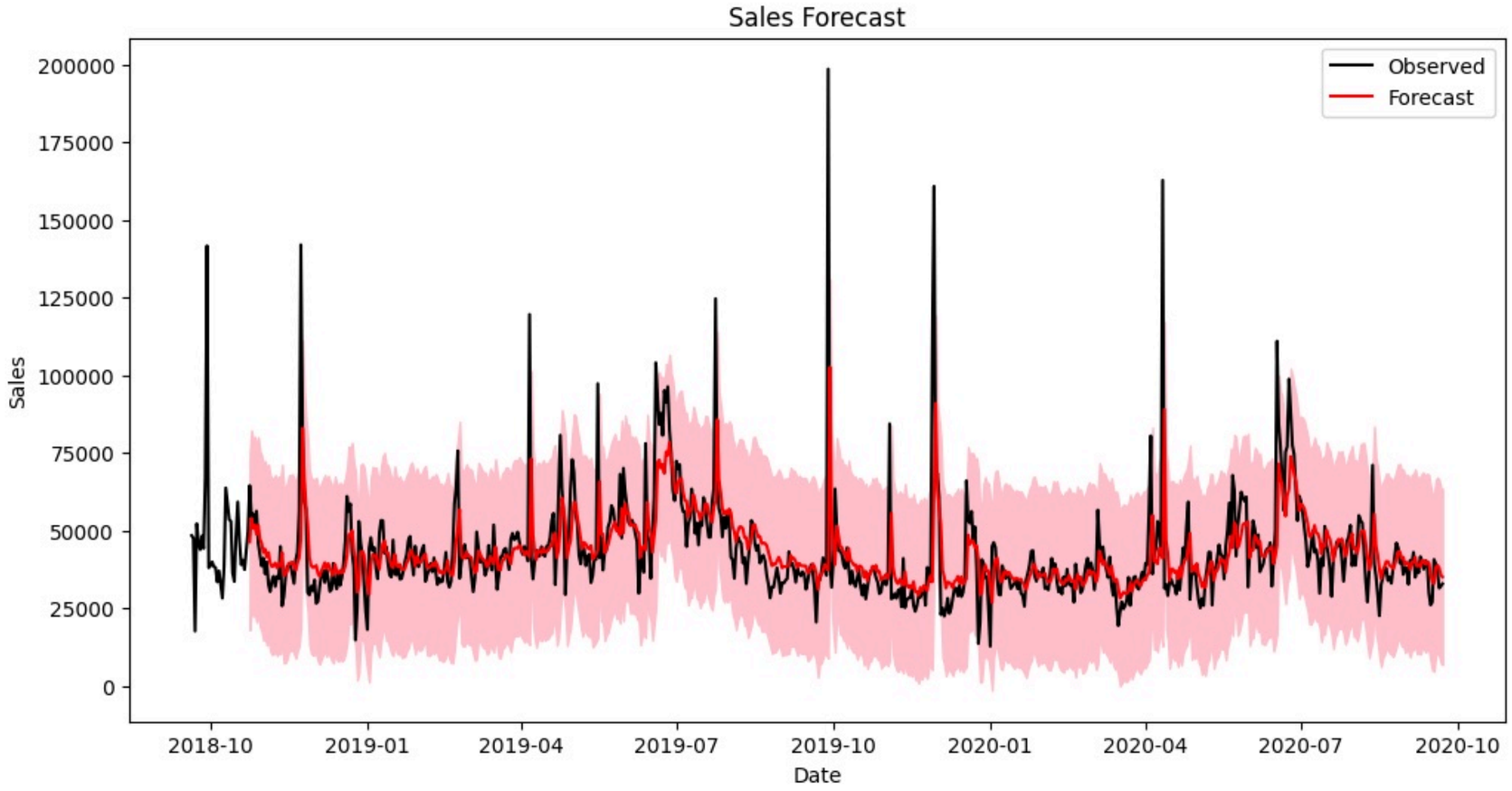
Artículo más comprado cada año para clientes de 40-55 años:

	year	article_id	prod_name	product_type_name	section_name	count
0	2018	539723005	Jade Denim TRS	Trousers	Divided Collection	1859
1	2019	706016001	Jade HW Skinny Denim TRS	Trousers	Divided Collection	6904
2	2020	751471001	Pluto RW slacks (1)	Trousers	Womens Everyday Collection	4255

2. PRONÓSTICO

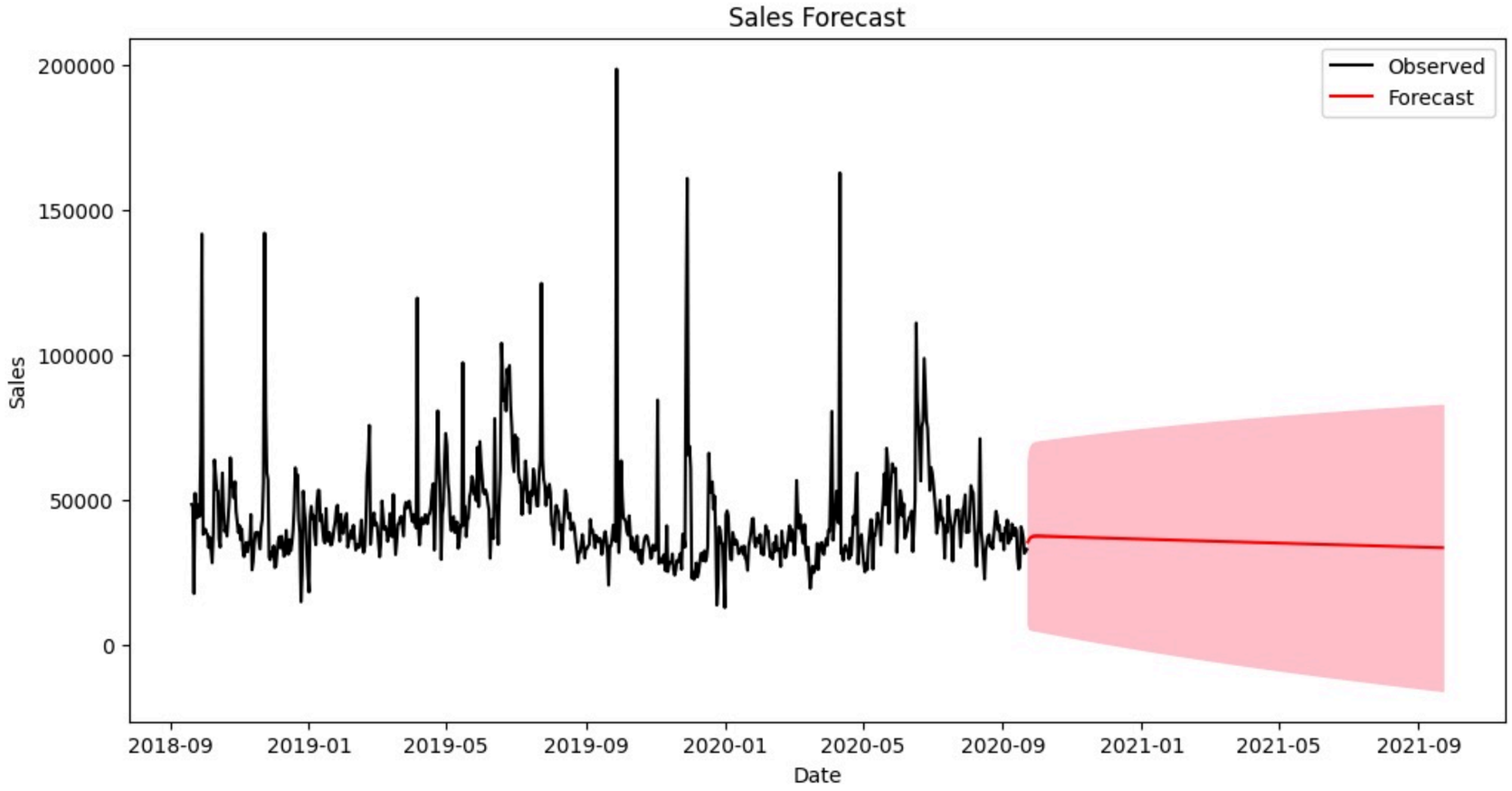
Series de tiempo

Sarima



Mean Absolute Error (MAE): 7382.128803835926

Sarima (pronóstico)



3. NEAREST NEIGHBORS

Para los artículos más comprados en cada año

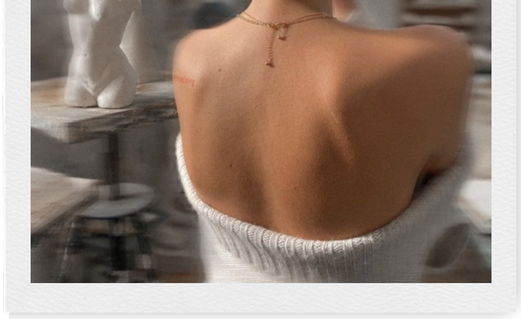
2018

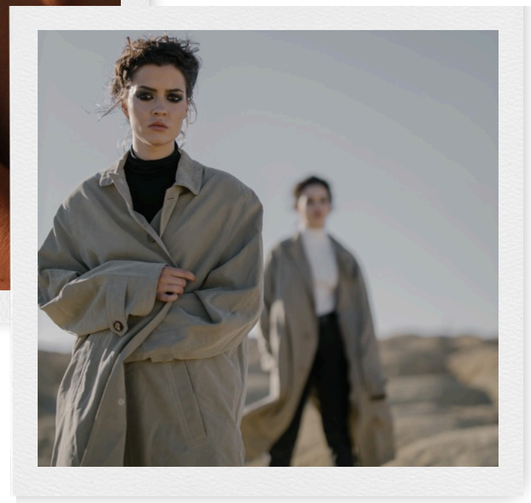
```
Artículos más similares al artículo con article_id 399223001 :
      article_id              prod_name product_type_name  section_name  \
2205   399087047      S.Skinny L.W Epic          Trousers  Ladies Denim
2200   399087028      S.Skinny L.W Epic          Trousers  Ladies Denim
2198   399087019  Super Skinny L.W Epic          Trousers  Ladies Denim
2206   399087051      S.Skinny L.W Epic          Trousers  Ladies Denim
2192   399087001  Super Skinny L.W Epic          Trousers  Ladies Denim

     garment_group_name colour_group_name
2205     Trousers Denim             Black
2200     Trousers Denim             Black
2198     Trousers Denim             Black
2206     Trousers Denim             Black
2192     Trousers Denim             Black
```

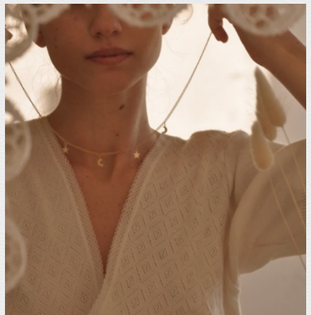
2019-2020

```
Artículos más similares al artículo con article_id 706016001 :
      article_id                prod_name product_type_name  \
2541   412707028                    Julia          Trousers
1070   316242054             Victoria (1)          Trousers
1071   316242055             Victoria (1)          Trousers
2607   416961006    Jade Denim Petite Trs          Trousers
2597   416511001   Julia Denim Petite Trs          Trousers

            section_name garment_group_name colour_group_name
2541  Divided Collection           Trousers             Black
1070  Divided Collection           Trousers             Black
1071  Divided Collection           Trousers             Black
2607  Divided Collection           Trousers             Black
2597  Divided Collection           Trousers             Black
```

4. SISTEMA DE RECOMENDACIÓN







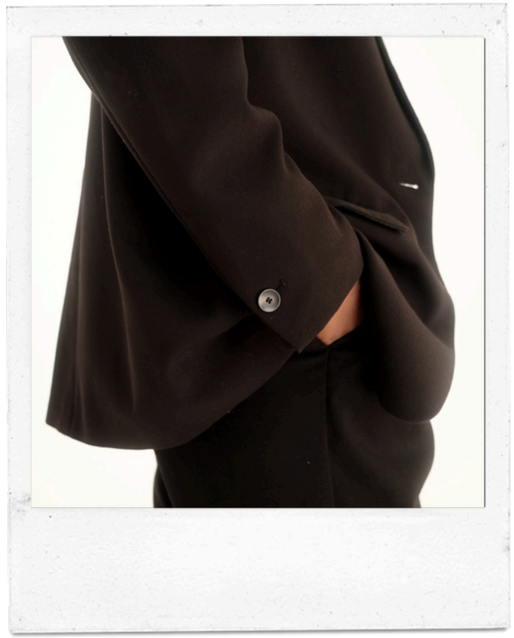
RECOMENDACIONES PARA LOS CLIENTES

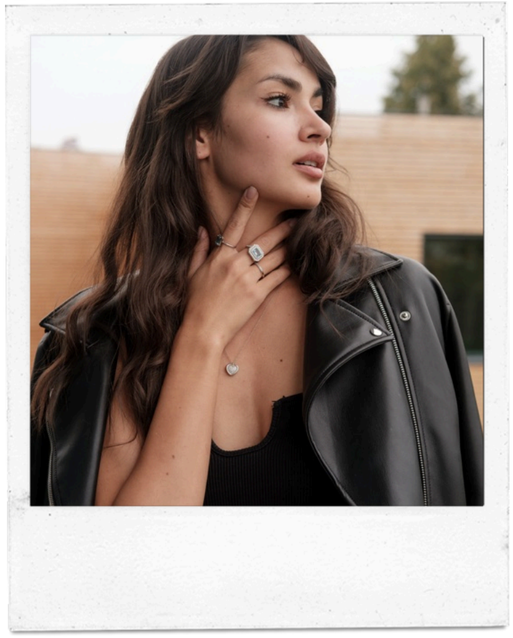
Modelo usado: NFM

- REDUCCIÓN DE DIMENSIONALIDAD
- DETECCIÓN DE PATRONES LATENTES
- RECOMENDACIONES PERSONALIZADAS
- SIMPLICIDAD E INTERPRETABILIDAD

DATOS UTILIZADOS

1. price: Rango de precios específico, y conocer esta información puede mejorar la precisión de las recomendaciones.
2. product_type_name: Ayuda a clasificar los productos en categorías. Es útil para recomendar productos similares dentro de las categorías que el cliente ha mostrado preferencia.
3. article_id y customer_id: Estas variables son esenciales para construir la matriz de interacción entre clientes y productos.

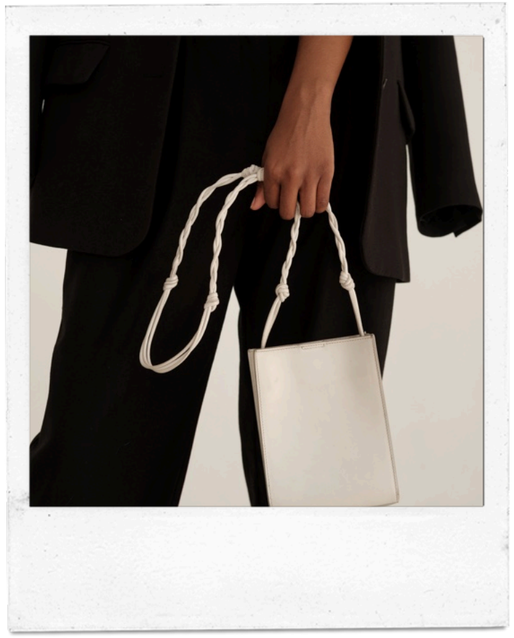




USUARIOS CON MÁS COMPRAS

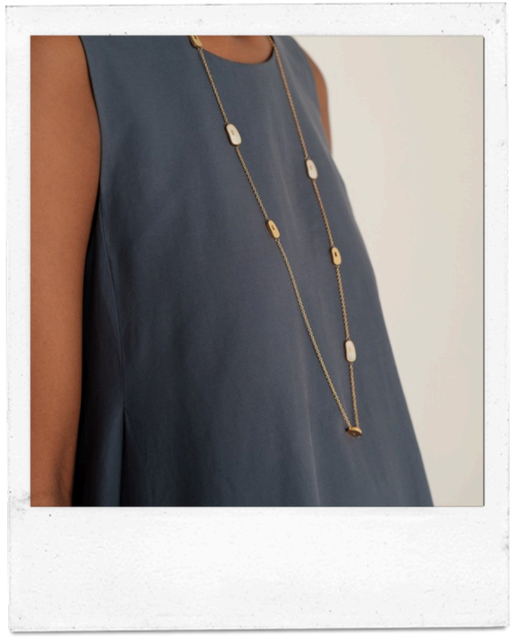
USUARIO A

TRANSACCIONES: 1641
ÚNICOS COMPRADOS: 1346



USUARIO B

TRANSACCIONES: 1321
ÚNICOS COMPRADOS: 1068



USUARIO C

TRANSACCIONES: 1304
ÚNICOS COMPRADOS: 906

USUARIO D

TRANSACCIONES: 1233
ÚNICOS COMPRADOS: 1088

article_id	prod_name	product_type_name
160442007	3p Sneaker Socks	Socks
160442010	3p Sneaker Socks	Socks
351484002	Lazer Razer Brief	Swimwear bottom
448509014	Perrie Slim Mom Denim TRS	Trousers
464297007	Greta Thong Mynta Low 3p	Underwear bottom
590928001	New Girl Push Top	Bikini top
599580017	Timeless Midrise Brief	Swimwear bottom
684209004	Simple as That Triangle Top	Bikini top
688537004	Simple as that Cheeky Tanga	Swimwear bottom
720125001	SUPREME RW tights	Leggings/Tights

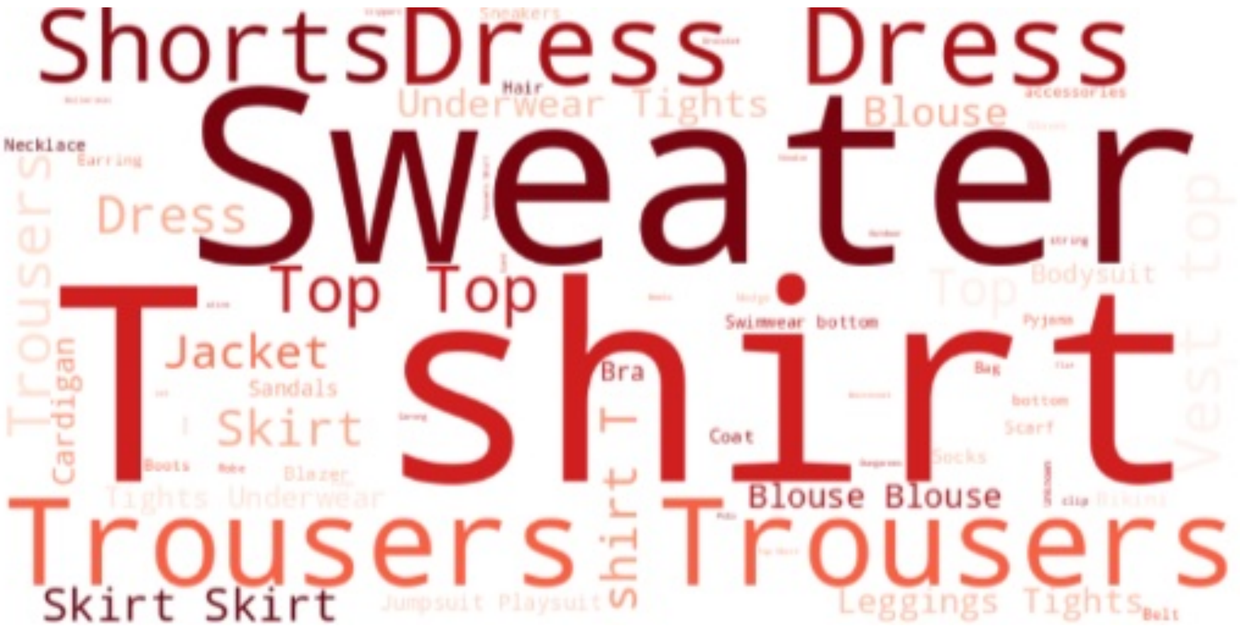
USUARIO A

article_id	prod_name	product_type_name
448509014	Perrie Slim Mom Denim TRS	Trousers
568601006	Mariette Blazer	Blazer
573716012	Kanta slacks RW	Trousers
579541001	Calista cardigan.	Cardigan
673396002	Ringo hipbelt	Belt
673677002	Henry polo. (1)	Sweater
678942001	Harrison short sleeve top CN	Top
751471001	Pluto RW slacks (1)	Trousers
783346001	Primo slacks	Trousers
806388002	Therese tee	T-shirt

USUARIO B

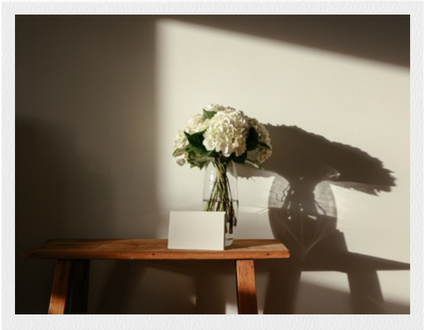
article_id	prod_name	product_type_name
351484002	Lazer Razer Brief	Swimwear bottom
448509014	Perrie Slim Mom Denim TRS	Trousers
464297007	Greta Thong Mynta Low 3p	Underwear bottom
562245001	Luna skinny RW	Trousers
590928001	New Girl Push Top	Bikini top
599580017	Timeless Midrise Brief	Swimwear bottom
684209004	Simple as That Triangle Top	Bikini top
688537004	Simple as that Cheeky Tanga	Swimwear bottom
720125001	SUPREME RW tights	Leggings/Tights
759871002	Tilda tank	Vest top

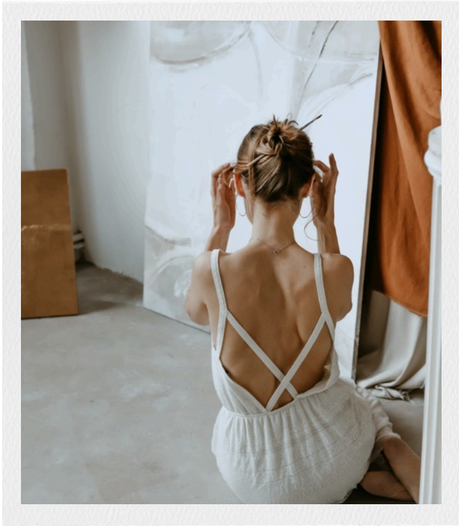
USUARIO C

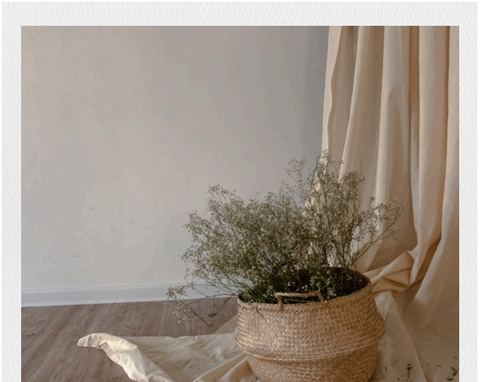


article_id	prod_name	product_type_name
448509014	Perrie Slim Mom Denim TRS	Trousers
568601006	Mariette Blazer	Blazer
573716012	Kanta slacks RW	Trousers
579541001	Calista cardigan.	Cardigan
673396002	Ringo hipbelt	Belt
673677002	Henry polo. (1)	Sweater
678942001	Harrison short sleeve top CN	Top
751471001	Pluto RW slacks (1)	Trousers
783346001	Primo slacks	Trousers
806388002	Therese tee	T-shirt

USUARIO D



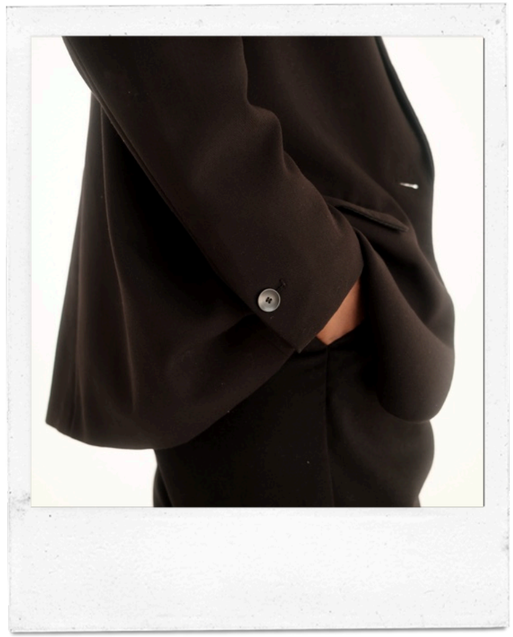


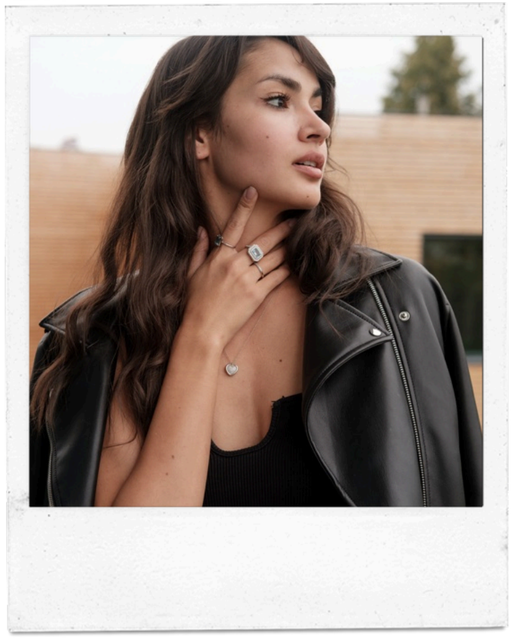


Conclusiones

1. El análisis detallado de las transacciones permite identificar distintos perfiles de clientes basados en su comportamiento de compra, como frecuencia de transacciones, variedad de productos y patrones de gasto.

2. El sistema de recomendación ofrece sugerencias personalizadas y relevantes al analizar las compras previas y las características de los productos. La eficiencia en el manejo de grandes volúmenes de datos asegura su escalabilidad y buen rendimiento.





Recomendaciones

1. Agregar más variables en las bases de datos que permitan identificar de mejor manera a los clientes y los artículos que compran.
2. Seccionar los datos antes de implementar el sistema de recomendación para mejorar la eficiencia.