

E-COMMERCE CON LOS OJOS VENDADOS

Entendiendo una web anónima a través del análisis de datos de recomendación

Esther Veguillas

<https://www.linkedin.com/in/eveguillas/>

OBJETIVO

- A partir de un conjunto de datos sobre valoración y recomendación de producto de una tienda anónima de ropa femenina online, se pretende comprender cómo es el cliente de este ecommerce, el tipo de prendas que se ofrecen y cuáles son las preferencias de los clientes.
- De este análisis “a ciegas” se podrán obtener conclusiones y oportunidades de mejora.

NATURALEZA DE LOS DATOS

- Se ha analizado una web de ecommerce de moda femenina a partir de los datos de recomendación de sus productos.

22.641	Comentarios
18 a 99	Rango de edades
1.206	Prendas únicas
20	Tipos de prendas
3	Niveles de clasificación
si/no	Recomendación
1 a 5	Estrellas de valoración

- Se desconocen otros datos de interés como volumen de ventas, país de origen, catálogo completo o número de referencias para cada categoría de prendas.

MÉTODO DE ANÁLISIS

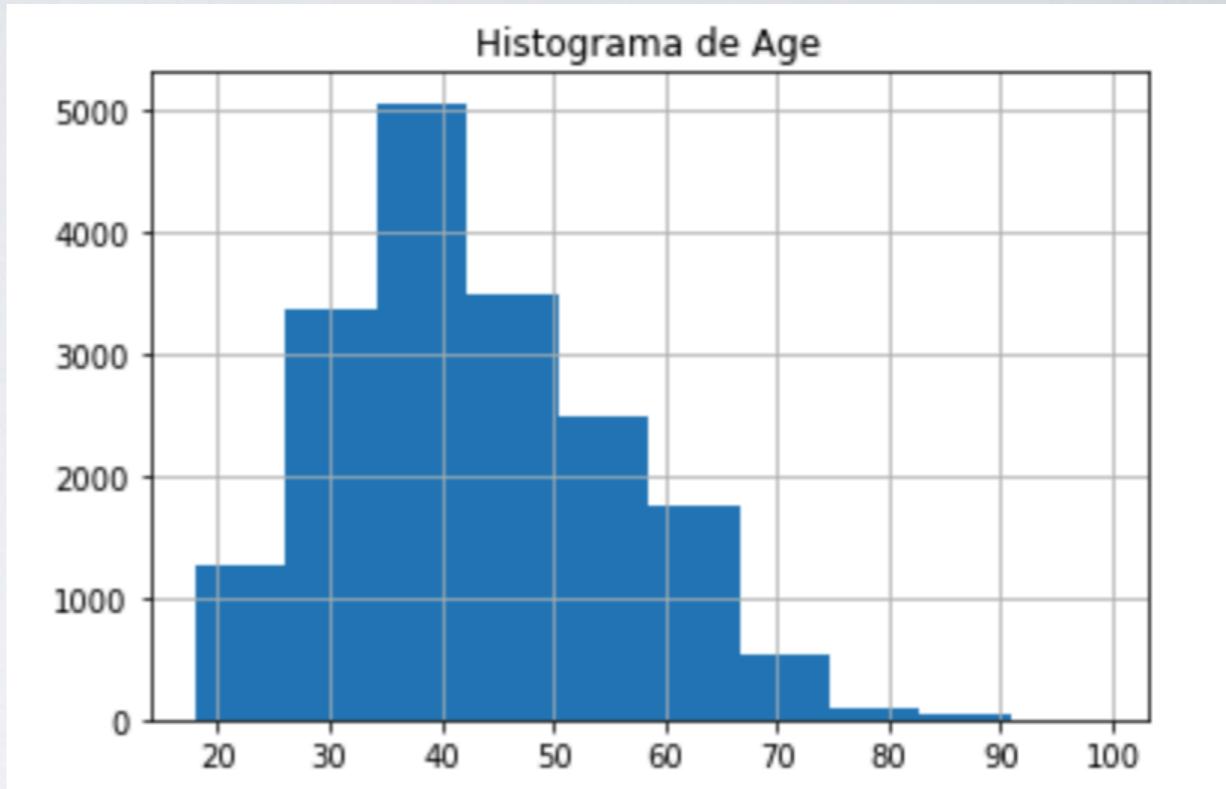
- Los datos, procedentes de un fichero en formato .csv, se han filtrado, limpiado y analizado con Python.
- Se han eliminado registros incompletos, sin información importante para el análisis, como tipo de prenda, valoración o recomendación o con datos presumiblemente falsos como edad por encima de 70 años.
- En casos de pérdida de información no imprescindible, como categoría de la prenda, se han utilizado distintas técnicas de completado del dato faltante.

EDAD

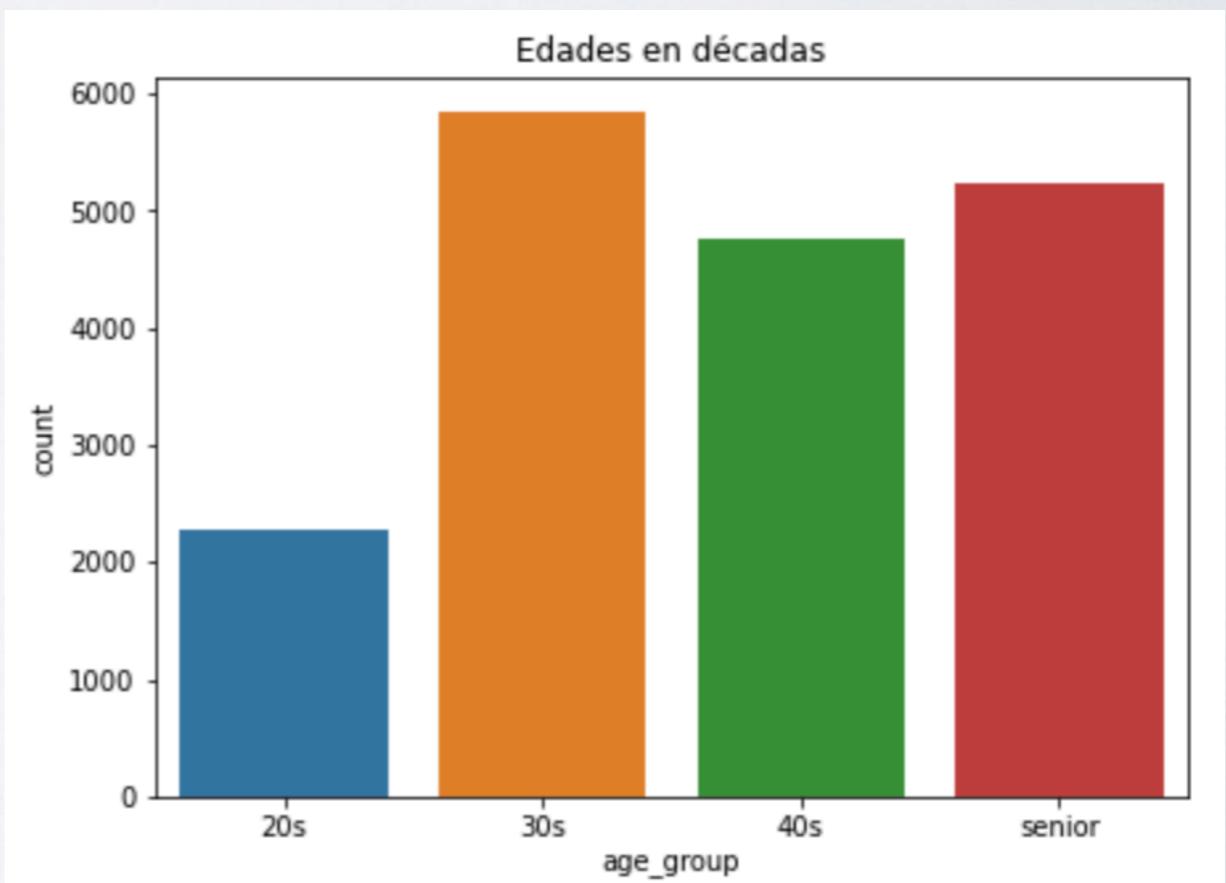
Rangos, grupos y reparto de población



- Aunque la franja de edad va de 18 a 99 años, las edades desde 70 en adelante se pueden considerar residuales por su volumen y de nulo interés para el estudio, ya que resultan poco fiables como clientes online.
- La distribución sigue una curva Log-normal. Entre 30 y 50 años es donde se concentra el público principal, que se va escalonando para acabar en los valores residuales mencionados.



- Se agrupa por tramos de décadas (18-29, 30-39, 40-49) siendo el senior el que recoge los compradores a partir de 50 años.
- Los compradores más jóvenes son también el grupo más reducido. Menos del 50% de la siguiente categoría (40's).
- Los clientes senior son el segundo grupo más importante en cuanto a volumen, aunque sus edades se encuentran escalonadas, teniendo mayor presencia los de menor edad.



EDAD: CONCLUSIONES

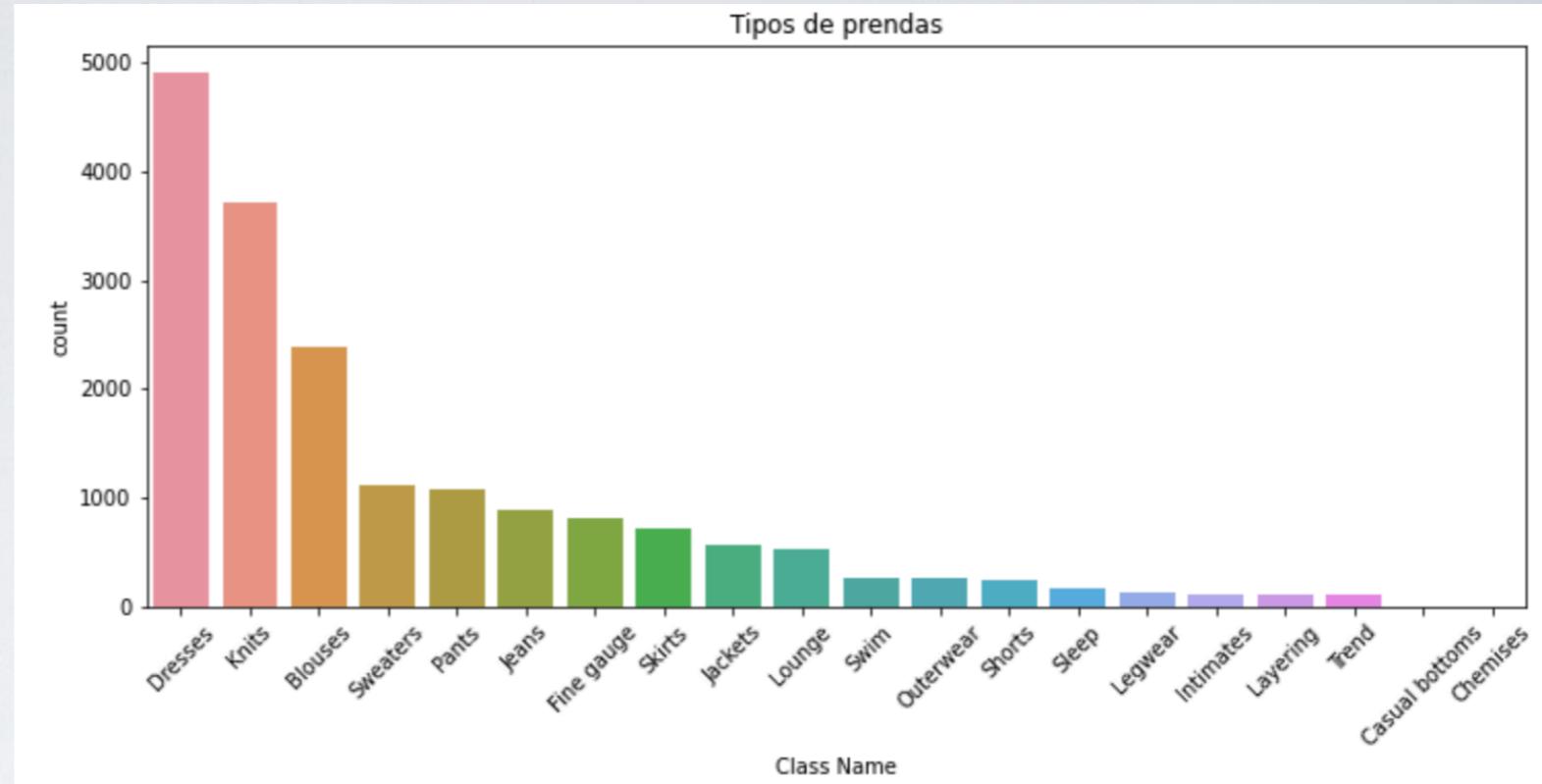
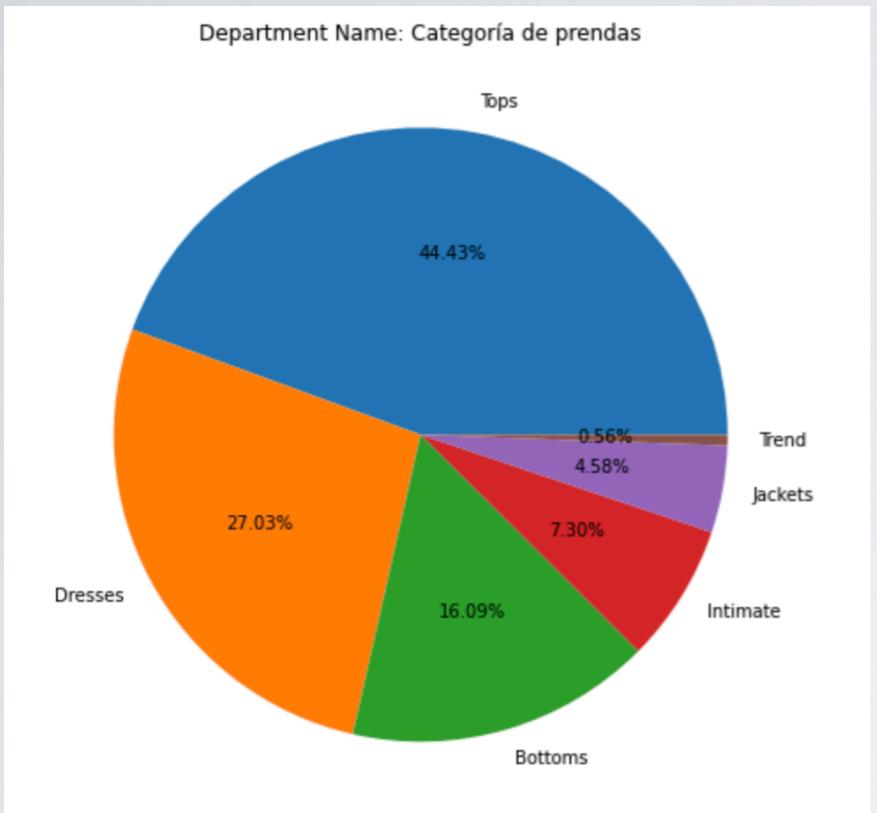
- Los compradores principales se sitúan entre alrededor de los 40 años, aunque los mayores de 50 tienen una aportación importante.
- Los compradores más jóvenes no son el público principal de esta tienda online.



TIPOS DE PRENDAS

Clasificación de prendas y categorías en las que se agrupan





- La web cuenta con 20 tipos de prendas diferentes, entre las que se desmarcan 3: los vestidos (27,1%) , seguidos de jerseys de punto (20,5%) y blusas (13,1%), todas ellas con más de 2.000 menciones cada una. El resto de tipos está alrededor o por debajo de 1.000 comentarios.
- Analizando las prendas tipo “knits”, hay que aclarar que se refiere solamente a jersey confeccionados en punto y entra en la categoría de “tops”. El catálogo cuenta con jerseys confeccionados con otro material (“sweaters”) con la tercera parte de comentarios que los de lana. Si se toman ambas prendas en conjunto, los jerseys estarán a la altura de los vestidos con 4.839 comentarios frente a los 4.905 de la prenda que lidera el ranking.
- Agrupados según las categorías que asigna la tienda, entre “tops” (blusas, jerseys...) y “dresses” se reparten el 71,5% de los comentarios de los clientes.
- Las prendas inferiores “bottoms” (vaqueros, faldas, pantalones...) son la tercera categoría y suponen menos de la mitad de los comentarios que la categoría principal (“tops”). Además, marcan una separación entre las populares y las minoritarias.
- Estas categorías minoritarias, que no llegan al 10% cada una, consideradas en conjunto tienen una cuota de 12,44%. En estas 3 categorías se agrupan 10 tipos de prendas

PRENDAS: CONCLUSIONES

- El vestido es la prenda estrella de la tienda, aunque compartiría su liderazgo con los jerseys si en ellos no se hicieran diferencias por tipo de confección. El vestido tiene su propia categoría y como tal ocupa el segundo puesto.
- Los “tops” son la categoría principal, por delante de vestidos y “bottoms”, ya que recoge todos los tipos de jerseys, blusas y otras prendas superiores.
- Los “bottom” tienen casi la mitad de comentarios que sus prendas complementarias, los “tops”. Esto podría deberse a que con dos piezas superiores y una inferior se consiguen dos atuendos diferentes.
- Las categorías de chaquetas, complementos y ropa interior, se reparten algo más del 12% del interés de los clientes que comentan.

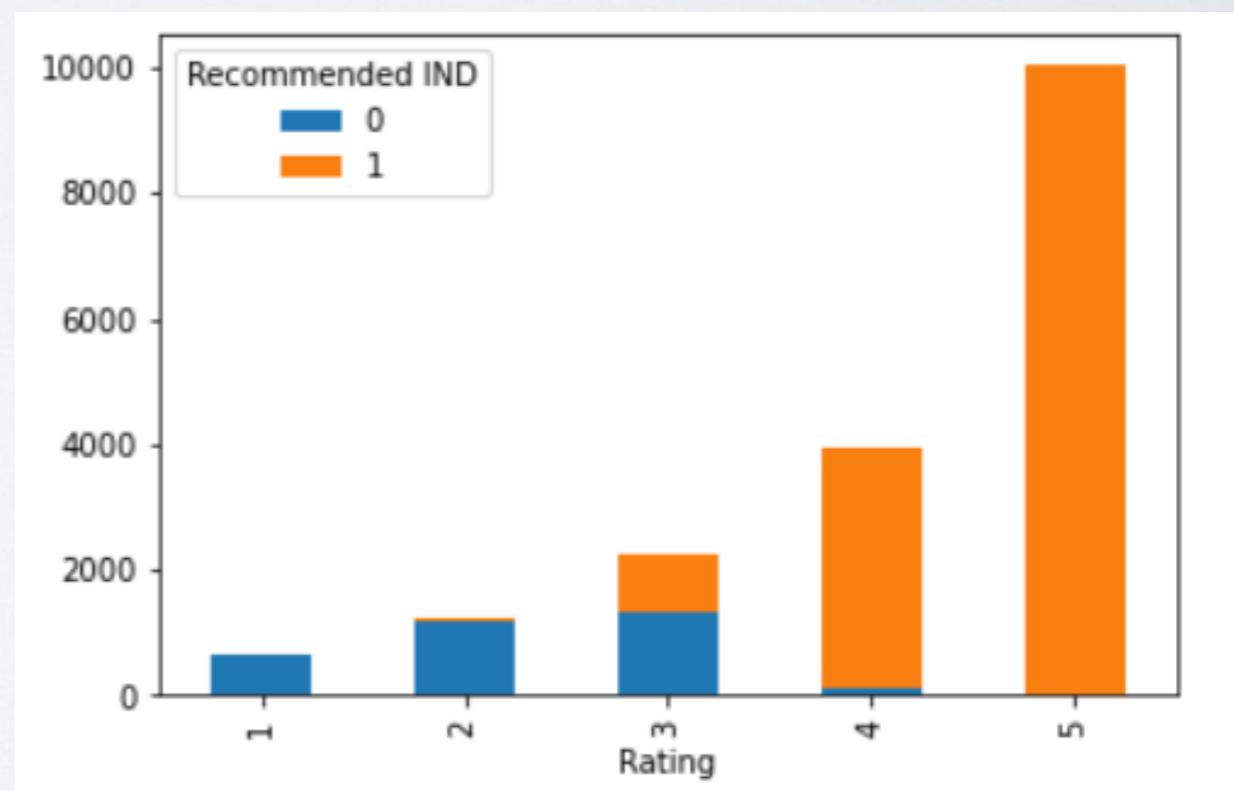
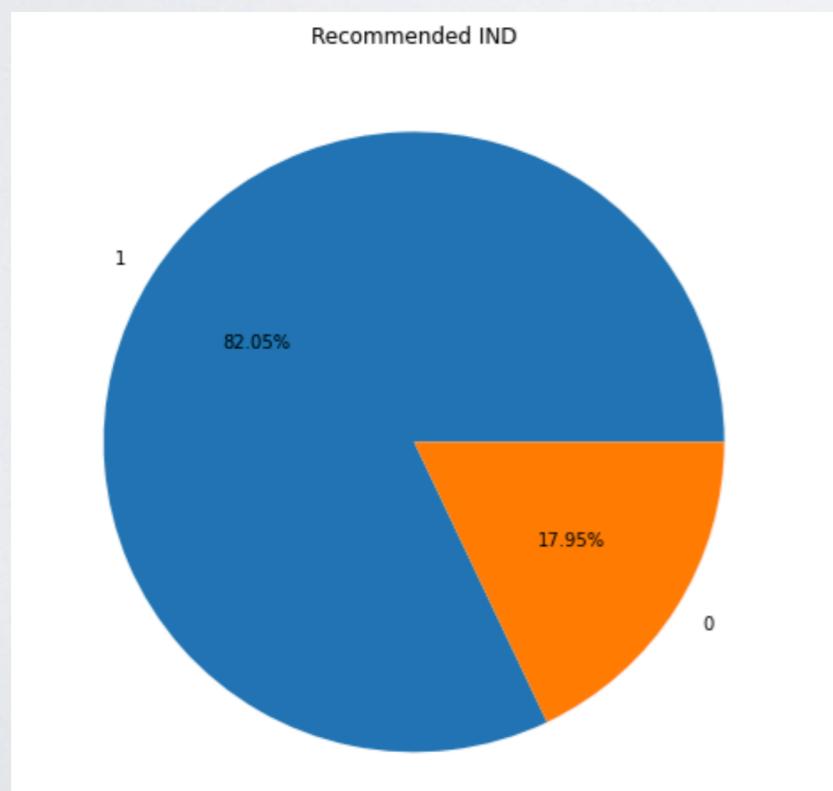
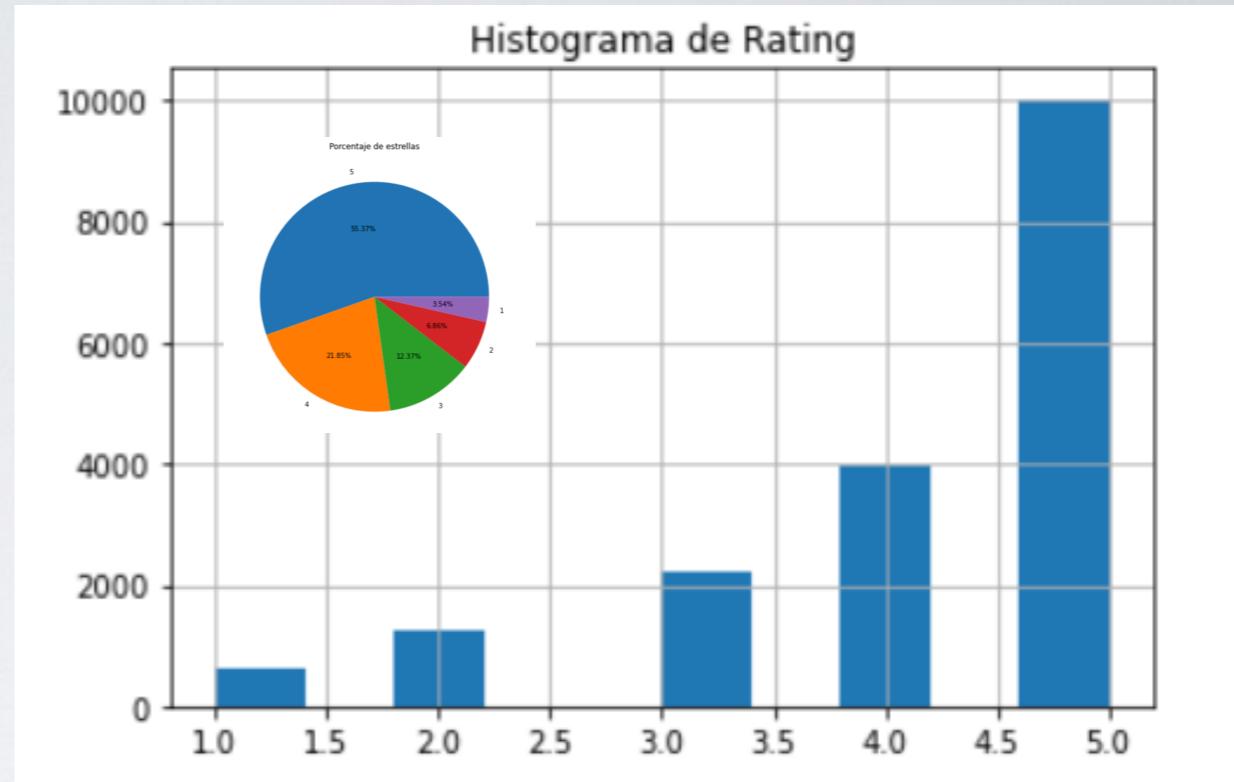


VALORACIONES Y RECOMENDACIONES

Valoración de 1 a 5 estrellas e indicaciones sobre si se recomienda o no la prenda comentada.



- Las valoraciones se realizan con puntuación de 1 a 5 estrellas, siendo 1 la más baja y 5 la más alta.
- La valoración es mayoritariamente positiva y muy alta: 5 estrellas en un 55.37% y 4 en un 21.85% de los casos.
- La recomendación puede ser sí o no, representada en las gráficas como 1 y 0 respectivamente.
- Las recomendaciones también son mayoritarias y multiplican por 5 los casos que no se recomiendan.
- Las puntuaciones de 4 y 5 estrellas llevan asociada la recomendación del producto. Con 3 estrellas predomina la recomendación frente a la no recomendación.



VALORACIONES: CONCLUSIONES

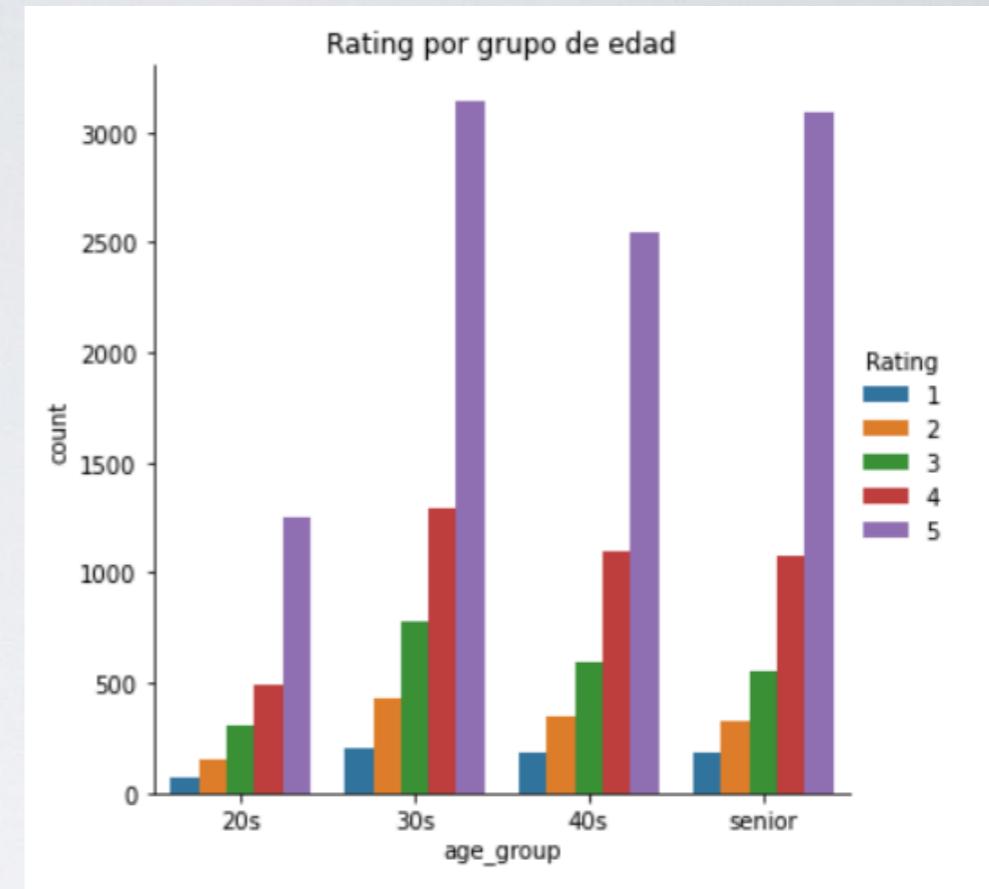
- En su mayoría, los comentarios que los clientes hacen de las prendas son muy positivos y van acompañados de la recomendación del producto.
- Los casos que no recomiendan la prenda corresponden a valoraciones de 1 y 2 estrellas. Los clientes que valoran con 3 recomiendan en más de la mitad de las ocasiones. Los casos de 4 estrellas en los que no se recomienda, aunque existen, soy ínfimos.
- No hay valoración de 5 estrellas que no lleve asociada la recomendación de la prenda comentada.



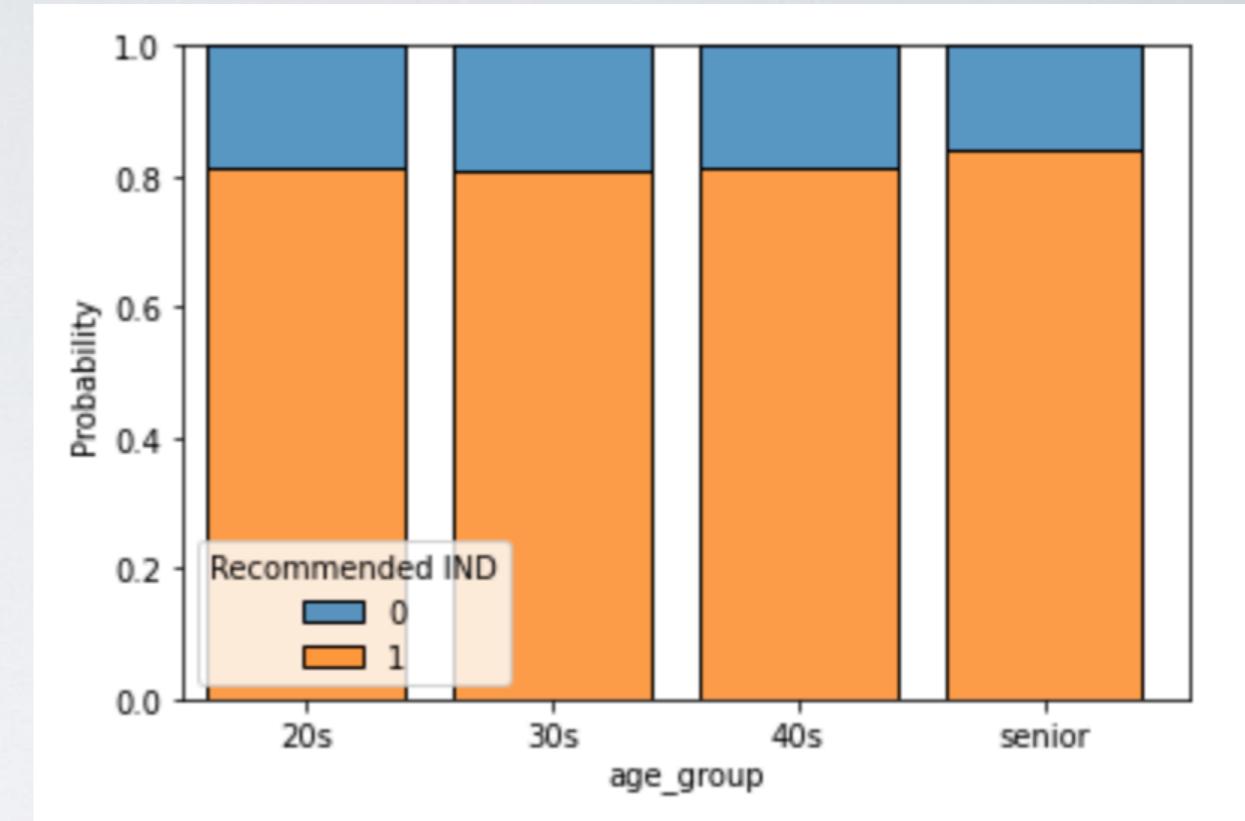
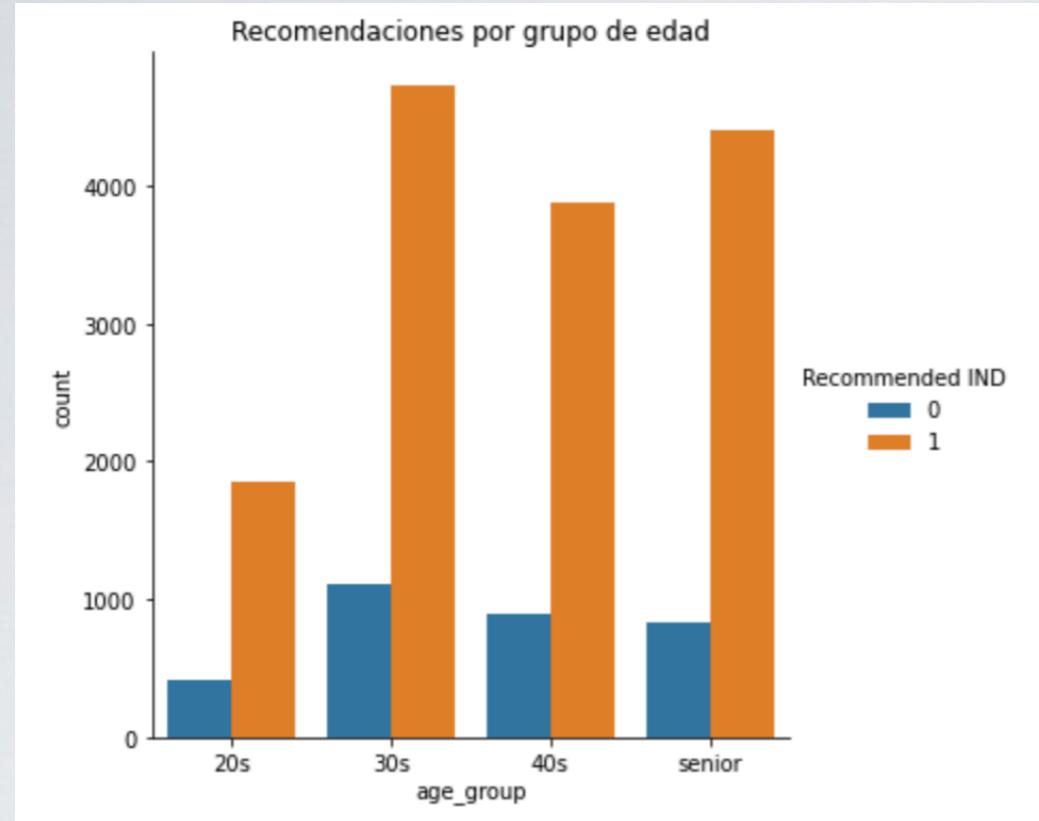
RELACIÓN ENTRE DATOS

Valoraciones, tipos de prenda,
recomendación y edad
combinados entre sí

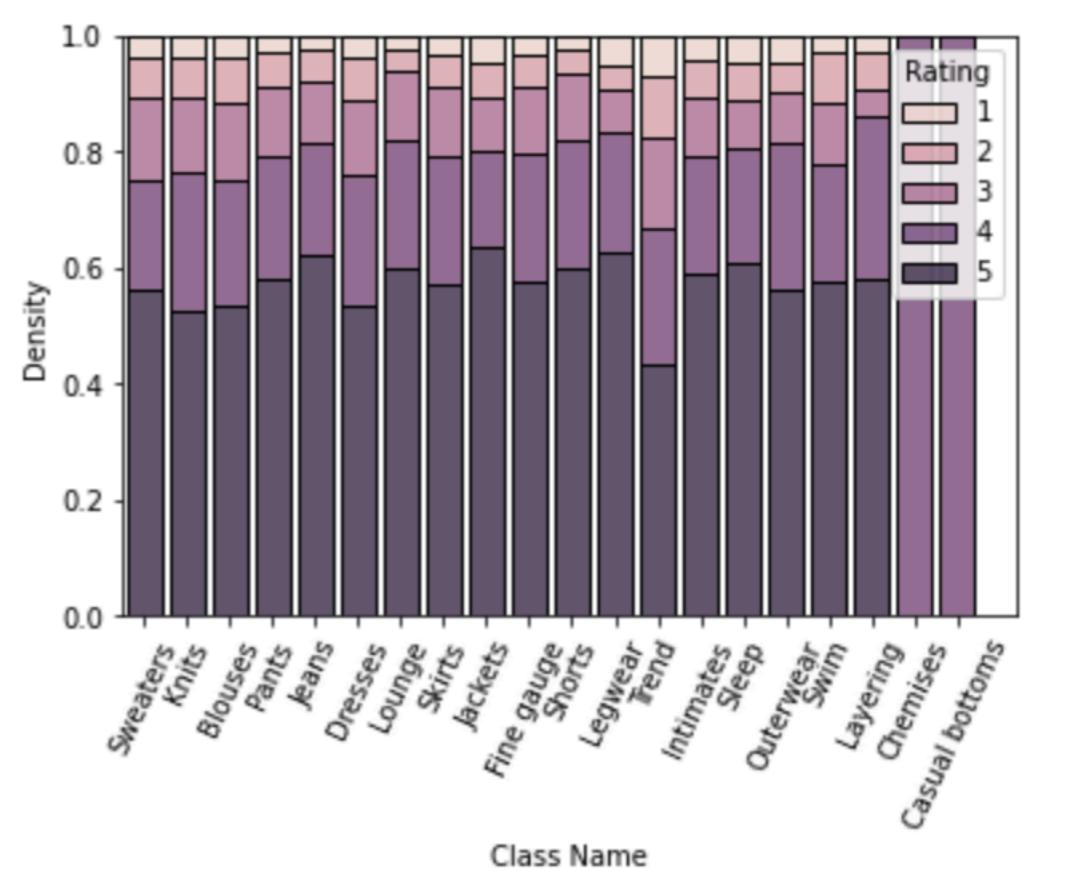
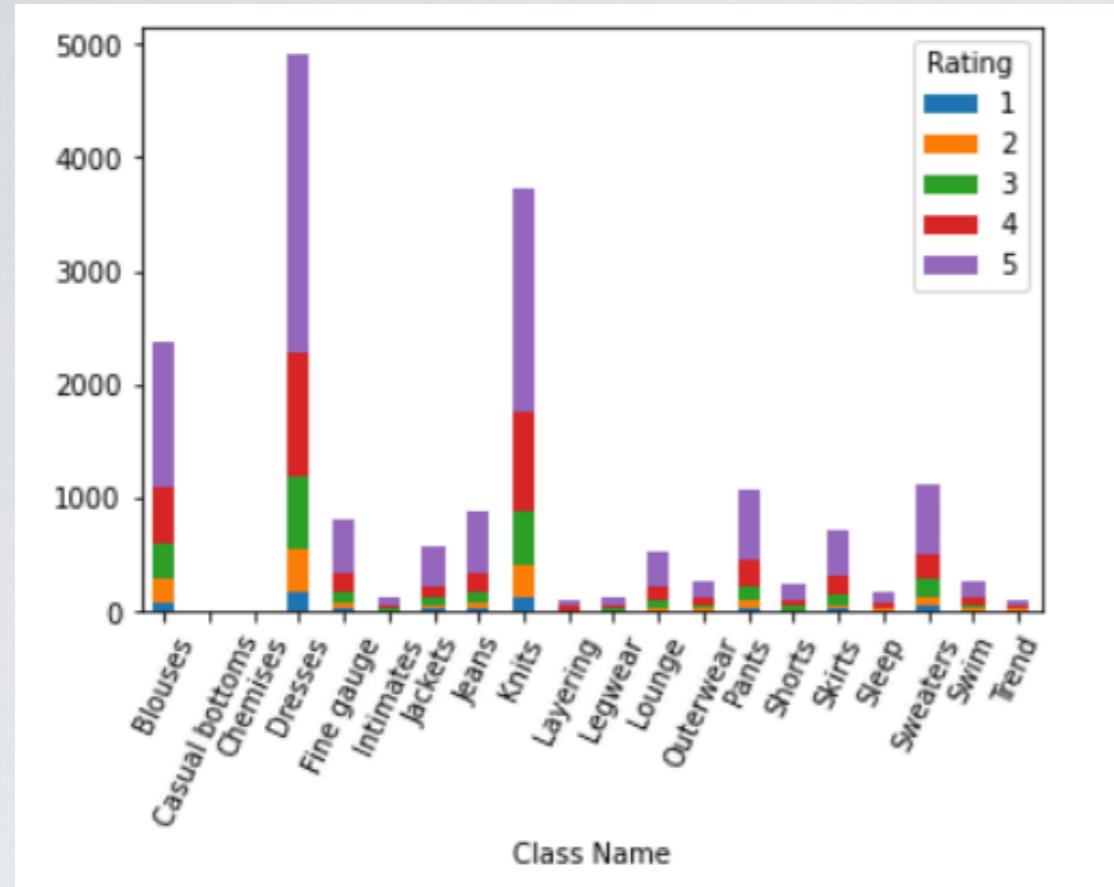




- La valoración de 1 a 5 estrellas se reparte por edades siguiendo un patrón similar en todas ellas.
- Todos los grupos de edad valoran mayoritariamente con 5 estrellas. Las valoraciones de 4 ocupan el segundo puesto y en ningún caso llegan al 50% de la puntuación máxima.
- Las mujeres de 30 y las senior encabezan la participación en número de comentarios de la web y también en la puntuación máxima.
- Las más jóvenes son el grupo con menos presencia en las recomendaciones y por tanto tienen menor número de recomendaciones, sin embargo, porcentualmente son igual de positivas que en el resto de edades.



- La recomendación en función de la edad está muy relacionada con el volumen de comentarios y la valoración de 5 estrellas, siendo mayoritariamente positiva. Como en esos casos, el patrón es una mayor recomendación de los grupos de 30 y 50 o más años, y menor cantidad de recomendaciones en el grupo más joven.
- En la recomendación negativa, destaca que el grupo de 40 años adelanta al senior en volumen de no recomendaciones, situándose en el segundo lugar. Sin embargo, en un análisis porcentual, la población de 20, 30 y 40 años tiene un comportamiento similar (80-20%) y son las mujeres senior las que reducen ligeramente los casos de no recomendación del producto.



- Con respecto al tipo de ropa, todas las prendas, excepto tres que tienen muy pocas valoraciones (complementos, camisas y casual bootoms), tienen más de 50% de sus votos con 5 estrellas. Además, la combinación de 4 y 5 está por encima del 75% de las valoraciones en cada tipo de prenda.
- Las tres prendas con mayor presencia (dresses, knits y blouses) tienen buena valoración general, pero no son, sin embargo, las mejor valoradas en proporción. Con 5 estrellas se sitúan a la cabeza las chaquetas, y con 4+5 las capas (layering). Estas prendas se encuentran en la segunda mitad del ranking por volumen de comentarios.
- En cuanto a las valoraciones negativas, con 1 y 2 estrellas, encontramos a los complementos, seguidos de ropa de baño, blusas y en cuarto lugar vestidos y pijamas.

DATOS RELACIONADOS: CONCLUSIONES

- Las valoraciones por grupos de edad no presentan sorpresas, ya que todos siguen un comportamiento similar de valoración muy positiva o positiva.
- Por edades, la recomendación también es muy alta y similar en todas las franjas, si bien es un poco más alta entre el público más senior.
- Las prendas con más volumen de comentarios están bien valoradas, aunque no son las que destacan en cuanto a proporción de 5 y 4 estrellas.
- Aparecen, sin embargo, dos tipos de prendas que sobresalen en valoración alta/muy alta aunque tienen un peso medio o muy bajo en cuanto a volumen de comentarios.



CONCLUSIONES

- Los datos analizados corresponden a eCommerce de ropa de mujer con un catálogo amplio en cuanto a categorías de productos, ya que cuenta con ropa variada: vestidos, pantalones, abrigos, ropa interior, complementos, pijamas y bañadores.
- Está orientada a un público de edad media entorno a los 40 años, que es consumidor online de prendas de todo tipo, aunque principalmente vestidos, prendas superiores, como jerseys o blusas e inferiores como pantalones y faldas.
- Las mujeres más jóvenes no son el público objetivo de la web, aunque ellas también son clientes de la tienda y tienen un comportamiento similar al resto.

- En general, todos los productos de la tienda están muy bien valorados, con 5 ó 4 estrellas, y se recomiendan en 8 de cada 10 ocasiones.
- Los vestidos y los jerseys son las prendas con más éxito, acaparan los comentarios en la web y sus valoraciones son altas en más del 75% de los casos.
- Las prendas mejor valoradas, sin embargo, son las chaquetas y las capas, dos categorías con un volumen muy bajo de menciones, pero con ratios de valoración altos/muy altos por encima del resto de productos.
- Las prendas con peor valoración y también volumen muy bajo de comentarios son bañadores y complementos.

OPORTUNIDADES: EDAD

- Los productos parecen bien segmentados hacia un público objetivo basado en mujeres que visten a la moda independientemente de su edad, con un foco destacable en mujeres más maduras a las que les gusta vestir como las más jóvenes.
- Oportunidad: Potenciar como diferenciación de la marca la ruptura de barreras de edad en la moda.
 - Aumentar el foco en el público senior, que ya se siente identificado.
 - Atraer al público más joven, ya que parece interesado en el estilo de la marca, con prendas con las que se sientan identificadas y romper también por ese extremo las barreras de edad.

OPORTUNIDADES: PRENDAS

- Las prendas estrella son los vestidos, las prendas de punto y los top en general.
- Oportunidades: Servir de gancho para otras prendas
 - Reforzar el catálogo en otras categorías que puedan complementar estas prendas, para que al realizar la compra se sugieran complementos, pantalones/faldas que resulten igualmente atractivos y bien coordinados con la prenda que se está viendo.

OPORTUNIDADES: PRENDAS

- Las prendas con mejor valoración tienen poca presencia en el volumen total de prendas.
- Oportunidad: Aumentar el volumen en esta categoría
 - Dar más peso a estos productos en el catálogo, con más prendas de estas categorías que mantengan el nivel de calidad de estos productos.
 - Dar más visibilidad en la web a estos productos resaltando sus buenas valoraciones y recomendaciones.

OPORTUNIDADES: VALORACIONES

- Desconocemos si ya se está sacando partido a las valoraciones y recomendaciones de los productos que están haciendo los clientes.
- Oportunidades:
 - Dar visibilidad a los comentarios y recomendaciones de los clientes en la ficha de cada prenda y en distintas zonas de web o RR.SS.
 - Interactuar con los clientes que valoran y recomiendan, agradeciendo, escuchando sugerencias y buscando posibilidades de mejora del servicio.

- Créditos:
 - Análisis de los datos y preparación del informe: Esther Véguillas
<https://www.linkedin.com/in/eveguillas/>
 - Fuente de datos: <https://www.kaggle.com/nicapotato/womens-ecommerce-clothing-reviews>
 - Código Python del análisis realizado: https://github.com/Estherveg/ML_e-commerce_recomendation/blob/main/women_ecommerce.ipynb
 - Bancos de imágenes: Pexels (christian-diokno, ksenia-chernaya, rosana-solis)