



PRODUCT RECOMMENDATION MODEL

DATA ANALYTICS FINAL PROYECT

THE GOAL

El fin de este proyecto es trabajar con sets de data de: productos, consumidores y transacciones de H&M.
A partir de estos datos, creé un modelo de recomendación que, en base a un producto seleccionado, nos recomiende 10 productos similares y así crear más oportunidades de compra.



CONTENIDO

Paso a paso del proceso de desarrollo de mi modelo.

01

DATASETS

03

SEASONAL VOLUMES

05

CORRELATION SEARCH

02

CATEGORICAL EXPLORATION

04

AGE GROUPS

06

MODELING



01 DATASETS

Para este proyecto trabajé con un total de 31 Millones de observaciones distribuidas en 3 Datasets:



ARTICULOS [25 Variables]

	article_id	product_code	prod_name	product_type_no	product_type_name	product_group_name	graphical_appearance_no	graphical_appearance_name
78011	792027002	792027	Mini J Melina Playsuit.	274	Shorts	Garment Lower body	1010001	All over pattern
4037	474461030	474461	Evert Shorts	274	Shorts	Garment Lower body	1010016	Solid
6627	507854003	507854	5pk regular fancy	302	Socks	Socks & Tights	1010012	Mixed solid/pattern

CLIENTES [5 Variables]

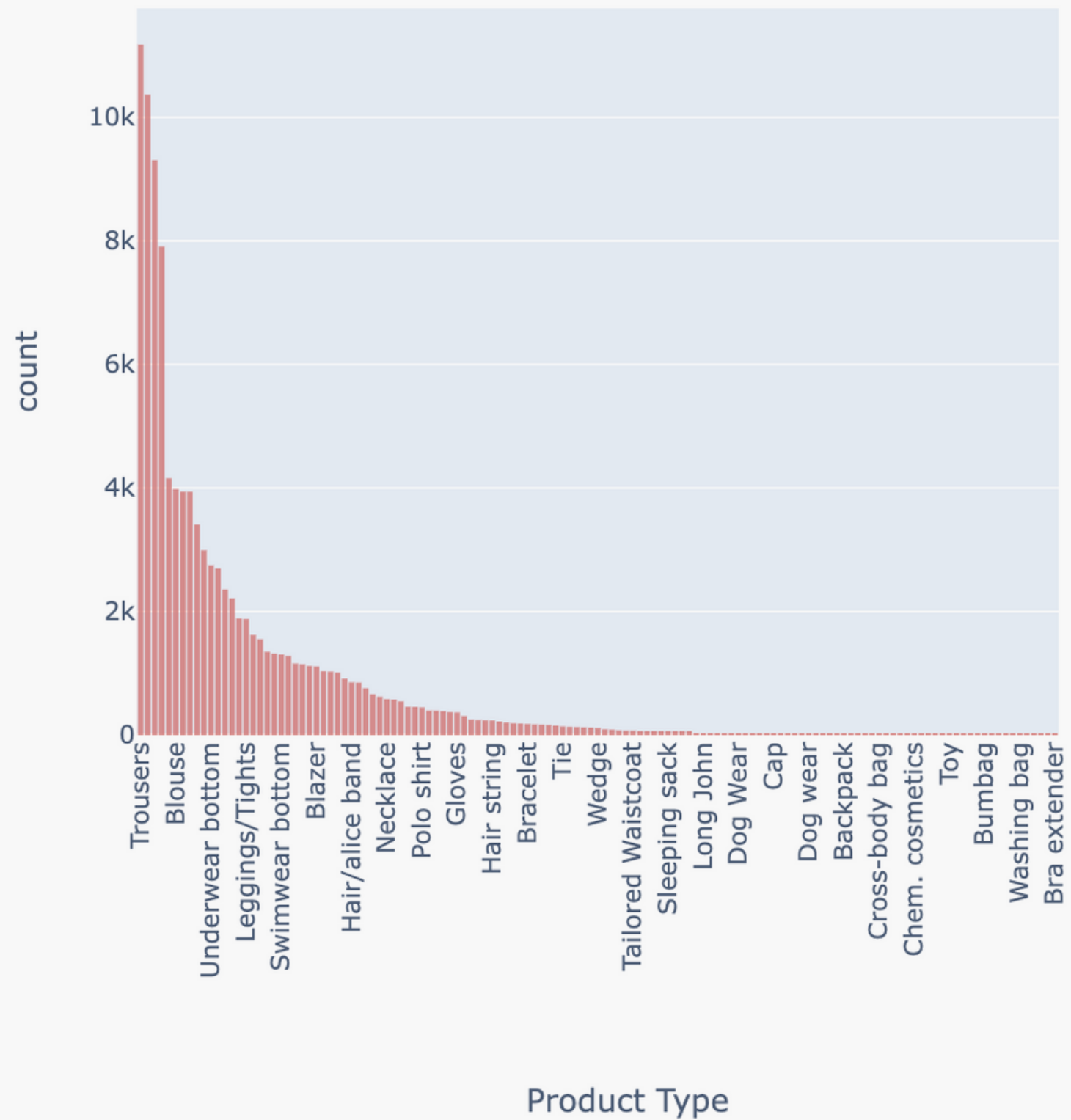
	t_dat	customer_id	article_id	price	sales_channel_id
0	2018-09-20	000058a12d5b43e67d225668fa1f8d618c13dc232df0ca...	663713001	0.050831	2
1	2018-09-20	000058a12d5b43e67d225668fa1f8d618c13dc232df0ca...	541518023	0.030492	2
2	2018-09-20	00007d2de826758b65a93dd24ce629ed66842531df6699...	505221004	0.015237	2
3	2018-09-20	00007d2de826758b65a93dd24ce629ed66842531df6699...	685687003	0.016932	2
4	2018-09-20	00007d2de826758b65a93dd24ce629ed66842531df6699...	685687004	0.016932	2

TRANSACCIONES [6 Variables]

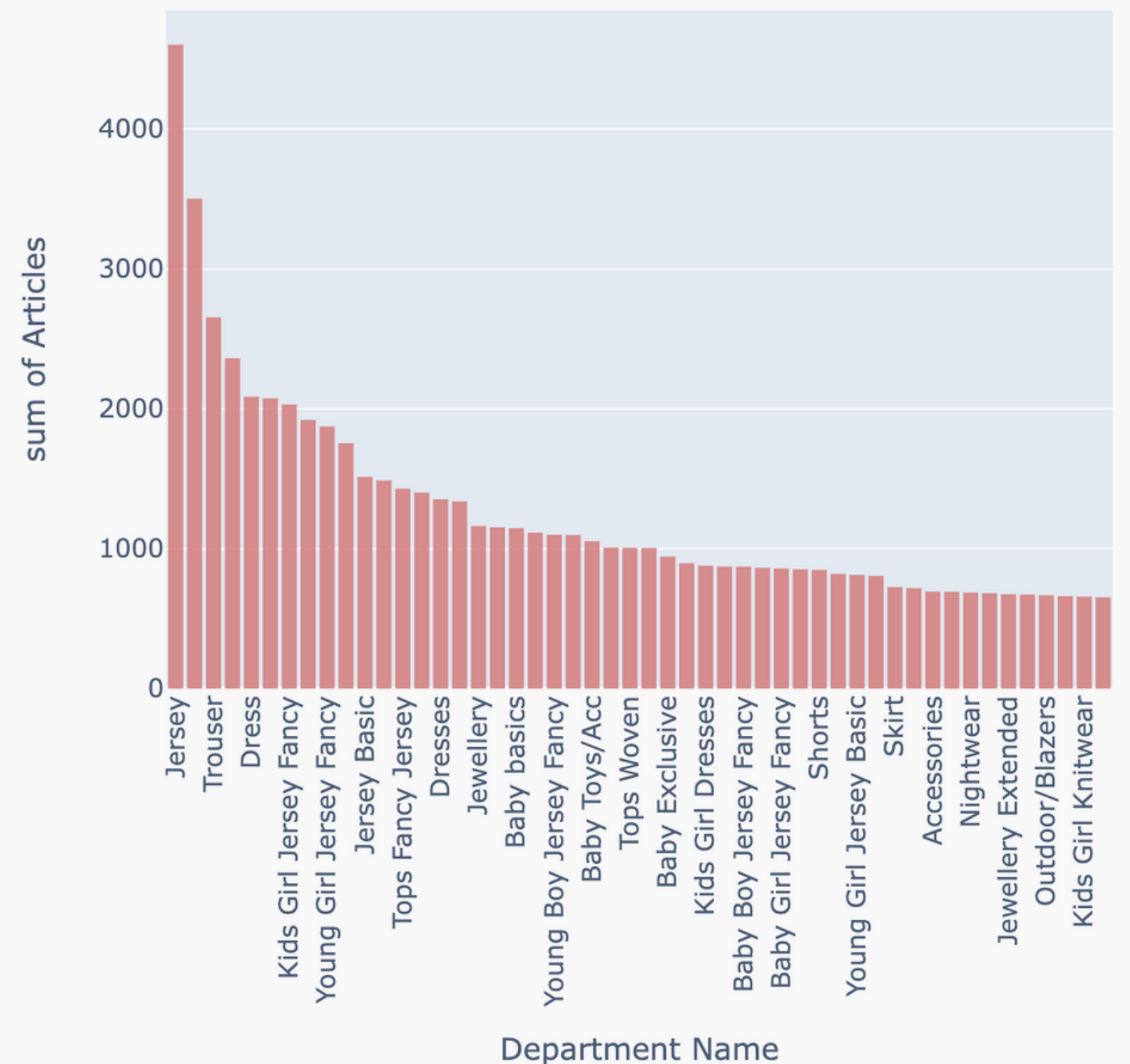
	customer_id	FN	Active	club_member_status	fashion_news_frequency	age
0	00000dbacae5abe5e23885899a1fa44253a17956c6d1c3...	NaN	NaN	ACTIVE	NONE	49.0
1	0000423b00ade91418cceaf3b26c6af3dd342b51fd051e...	NaN	NaN	ACTIVE	NONE	25.0
2	000058a12d5b43e67d225668fa1f8d618c13dc232df0ca...	NaN	NaN	ACTIVE	NONE	24.0
3	00005ca1c9ed5f5146b52ac8639a40ca9d57aeff4d1bd2...	NaN	NaN	ACTIVE	NONE	54.0
4	00006413d8573cd20ed7128e53b7b13819fe5cfc2d801f...	1.0	1.0	ACTIVE	Regularly	52.0

02 CATEGORICAL EXPLORATION

Distribution by Product Type



Number of Articles by Department



03

SEASONAL VOLUMES

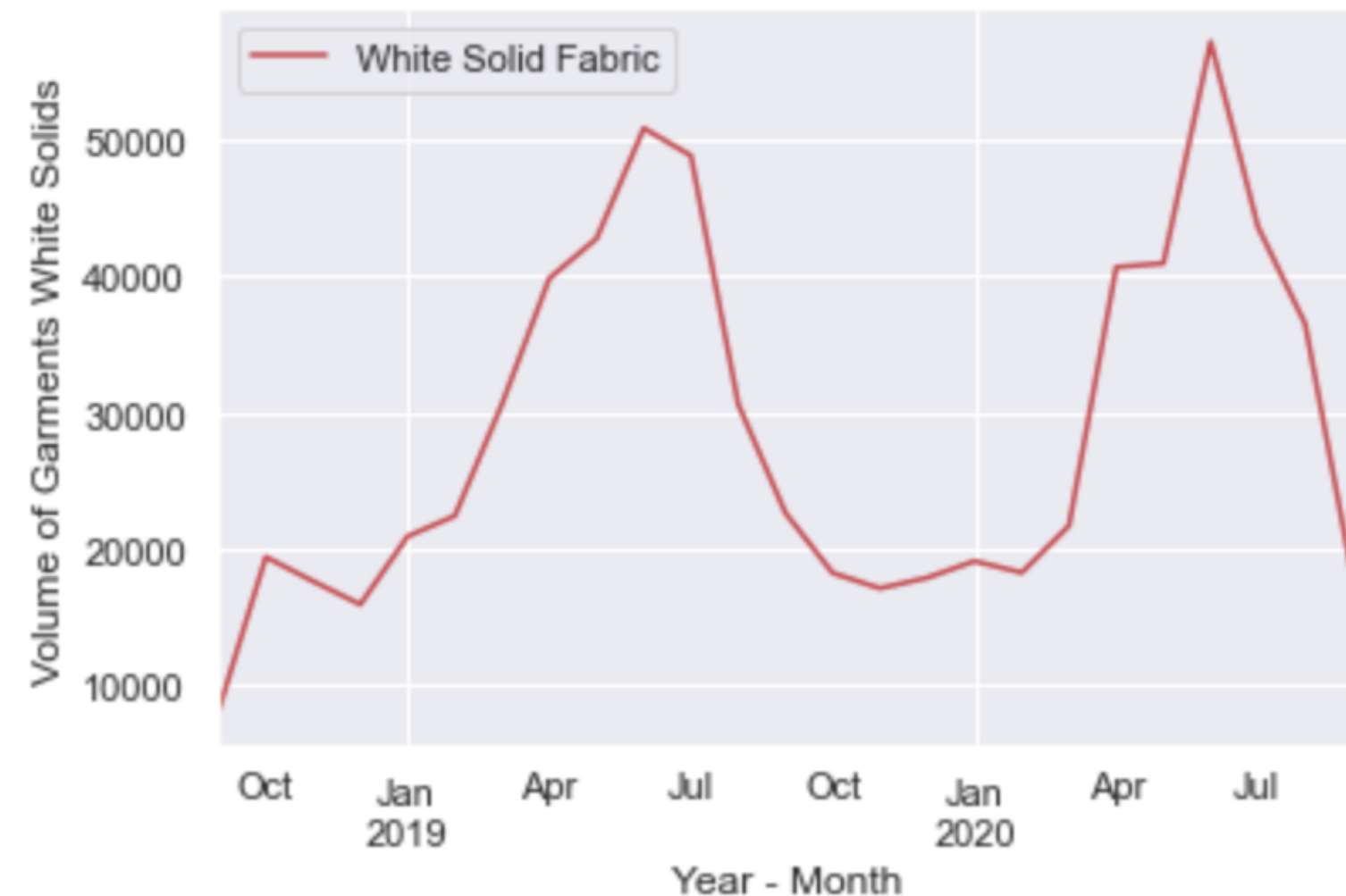
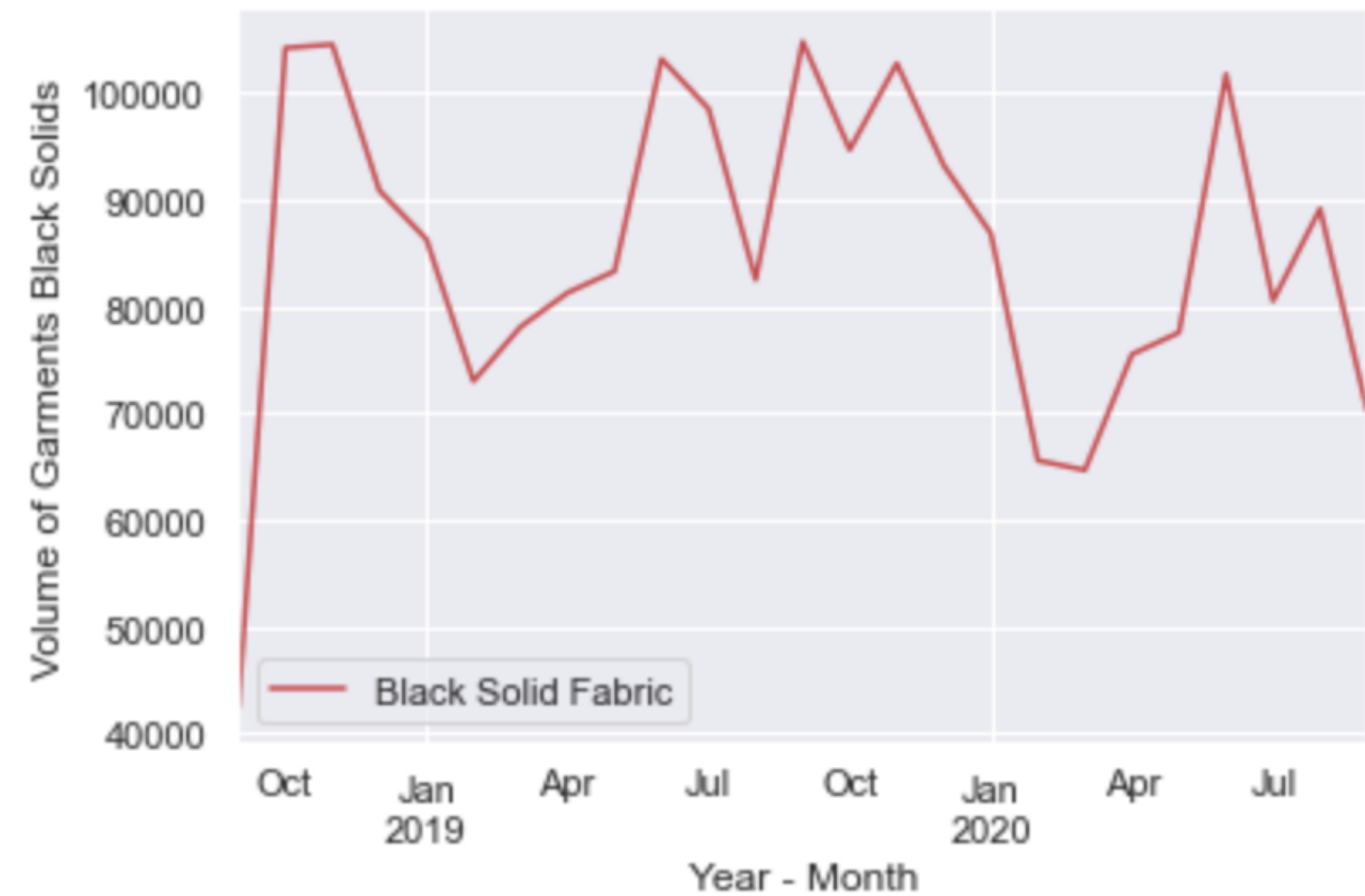
GROUPED VARIABLES

Agrupé las ventas diarias desde Sep 2018 hasta Sep 2019 para poder determinar cuales eran las combinaciones de materiales + colores de las prendas más vendidas a lo largo del tiempo.

Podemos ver las variaciones de venta de estos materiales a lo largo del tiempo dependiendo de su color:

Prendas | Tela Negra Solida 2,136,243

Prendas | Tela Blanca Solida 718,842

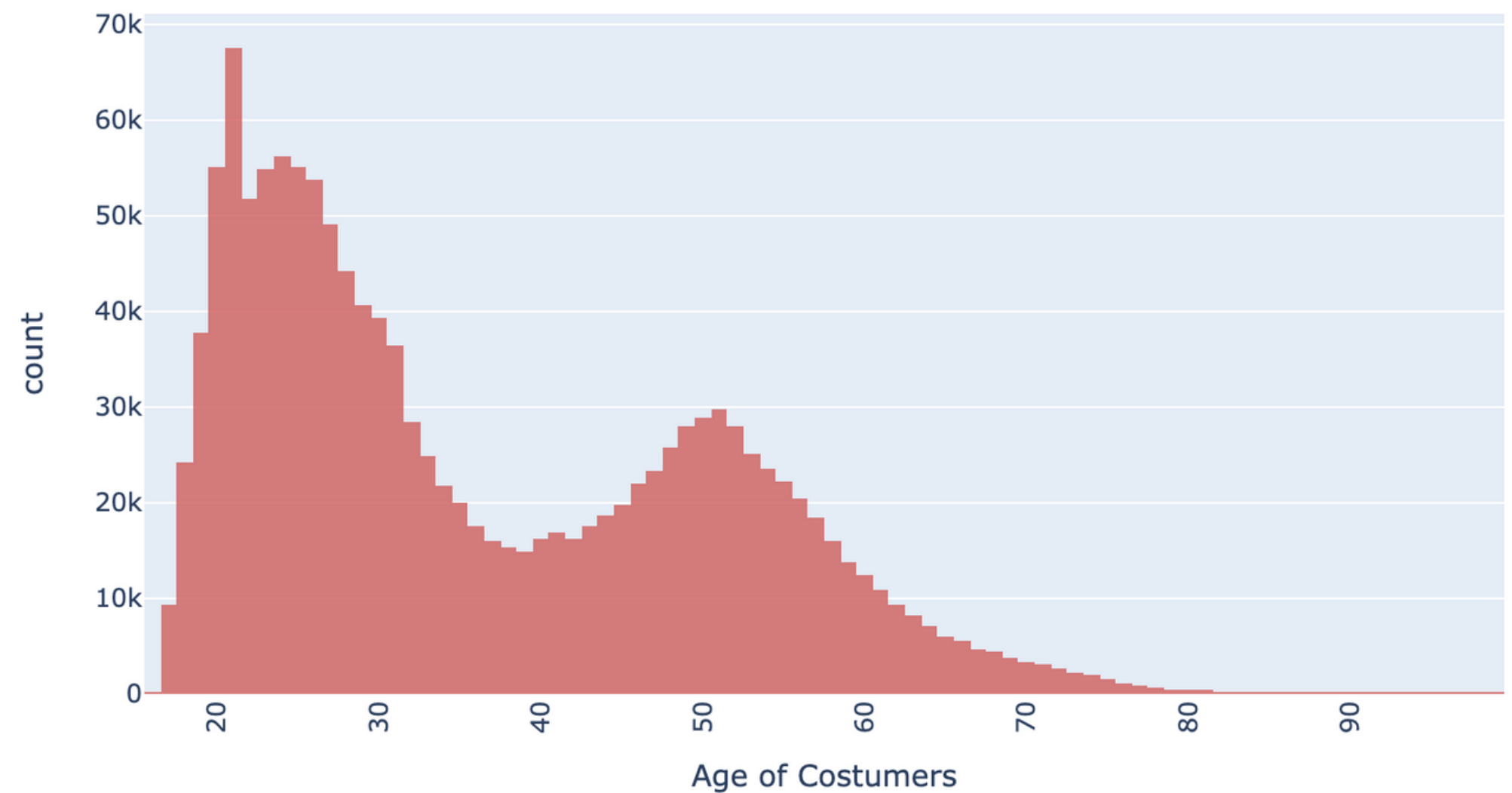


04 AGE GROUPS

La distribución de edades en los clientes esta definida bi-modalmente en dos grupos mayoritarios. Clientes entre 20-30 años dominan las compras, seguidos por clientes entre 45-50 años. La mayor cantidad de clientes tiene 21 años, este grupo supera los 60mil clientes



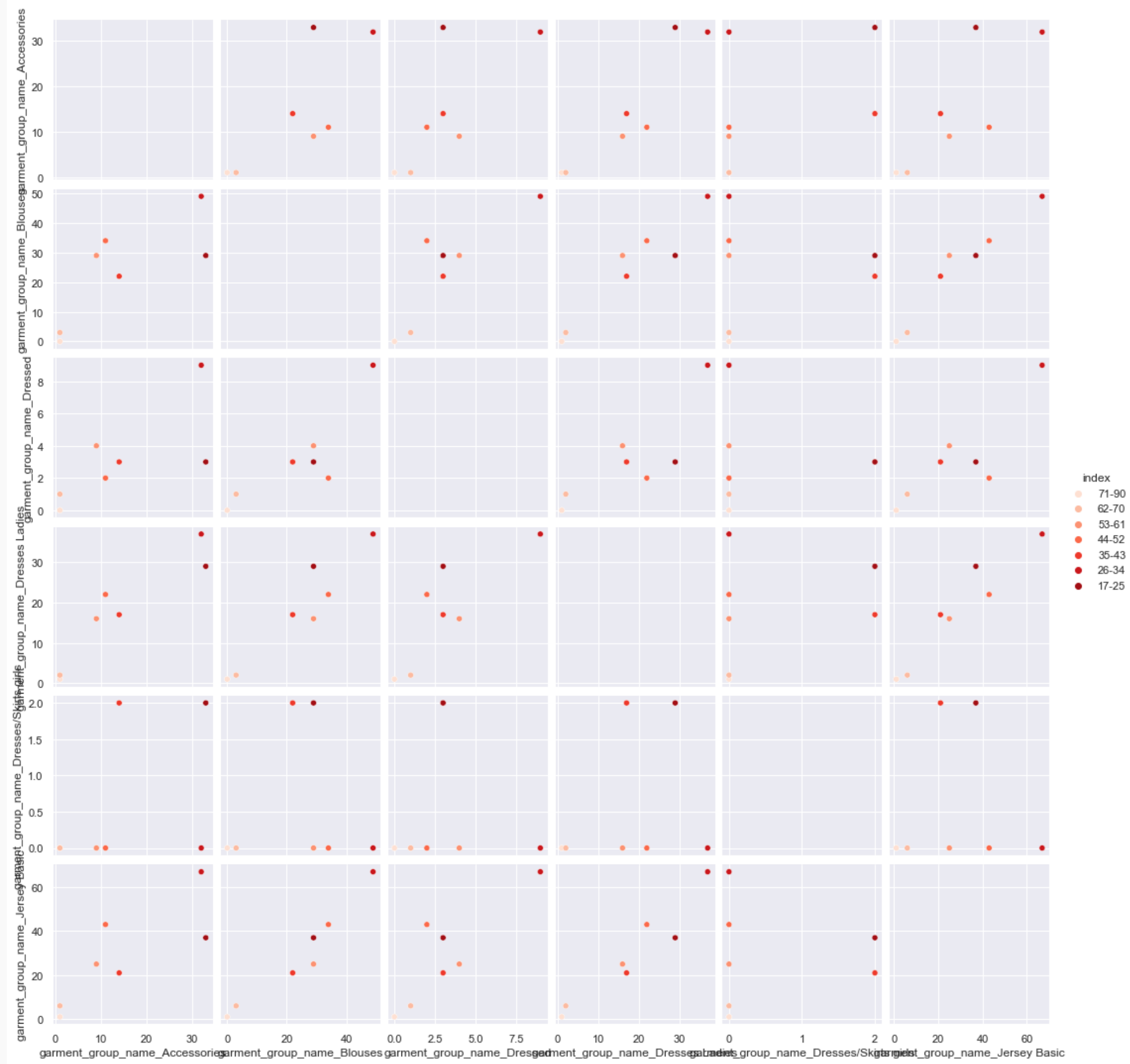
No. of Customers By Age Distribution



05 CORRELATION SEARCH

Luego de limpiar y explorar a detalle mis sets de datos busque correlación entre las variables características de mis productos y la edad de mis clientes para determinar si la edad seria un factor determinante para mi modelo de recomendación.

Determiné que no existe relación alguna entre la edad y la selección de productos por lo que recurrí a otro método.



1 PRODUCTO

Seleccionado Entra al Pipeline

50.000

Puntos Entran al Pipeline

SIMILITUD DE COSENOS

`['product_name']+['product_group_name']+['colour_group_name']+['product_type_name']+ ['index_group_name']`



**MODELO SESGADO POR LOS NOMBRES DEL
PRODUCTO**

1 PRODUCTO

Seleccionado Entra al Pipeline

50.000

Puntos Entran al Pipeline

SIMILITUD DE COSENOS

['product_group_name']+['colour_group_name']+['product_type_name']+ ['index_group_name']+
['garment_group_name']



**TOP TEN RECOMENDACIONES DE PRODUCTOS
SIMILARES**

06

MODEL
OUTPUTS | WOMEN
BOHO BLACK DRESS



de **50.000**
artículos que entran al
pipeline

9%

son

VESTIDOS

16%

son

PATRON

24%

son

MUJER

06

MODEL
OUTPUTS | WOMEN
BOHO BLACK DRESS



Utah dress



Napolitana Dress



Jenny boho dress



Jill l/s dress w



POWER DRESS



SPEED Binx



Fiona dress



SPEED Barbados W



Biker pinafore Dress



Fancy Dress



Hunt dress W



06

MODEL
OUTPUTS | GIRL RAIN BOOTS



de **50.000**
artículos que entran al
pipeline

5%

son

ZAPATOS

47%

son

SOLIDOS

20%

son

NIÑOS/AS

06

MODEL
OUTPUTS | GIRL RAIN BOOTS



Jackie waterproof SG



May boot SG



Blake chelsea boot
SG



West puffer boot
waterproof SB



Olivia boot SG



Jane Fancy BG



Lara high boot BG



Katie playful SG



Katie boot SG.



Reeze Boot Baby



Rainboot SB

06

MODEL
OUTPUTS | MEN JACKET



de **50.000**
artículos que entran al
pipeline

8%

son

CHAQUETAS

47%

son

SOLIDOS

12%

son

HOMBRE

06

MODEL
OUTPUTS | MEN JACKET



Jon varsity



Elliot Quilted Jkt



Paul windbreaker



Bobby (1)



Simz anorak



CAD puffer



Moon racer jacket



Tom coach jkt



Java bomber



Kurt Cord Jacket

06

MODEL
OUTPUTS | KIDS JOGGER



H&M

de **50.000**
artículos que entran al
pipeline

10%

son

PANTALONES

16%

son

PATRON

20%

son

NIÑOS/AS

06

MODEL
OUTPUTS | KIDS JOGGER



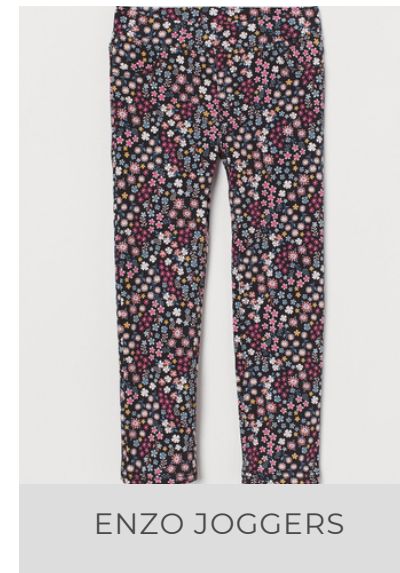
Timothy Jogger



BAT jogger



Meadow trs



ENZO JOGGERS



Naomi



Hazza



DRAKE sweatpants TP



North utility jogger



SMITH fancy sweatpants



Tuxedo junior pants



Kacey pants