



Materials Recommendation System

For Fabric Central

Supriya Sambhus
GA - DSIF1



About Fabric Central*

.....

WHAT IS FABRIC CENTRAL

Fabric Central is a digital sourcing platform that connects apparel brands and fabric suppliers in the most efficient, accessible and transparent marketplace.

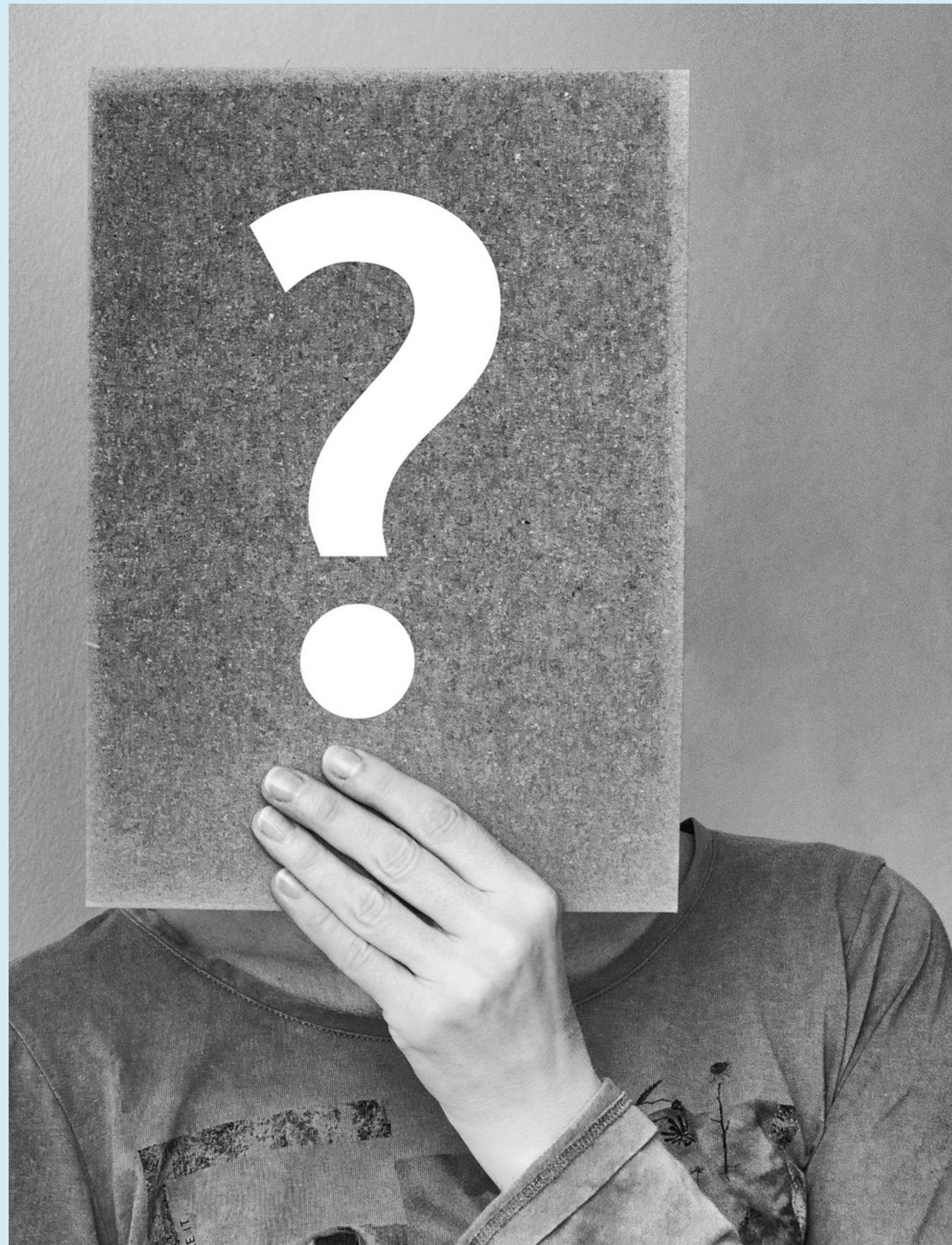
GROWTH SO FAR

- 78 suppliers onboard
- 359 brands registered
- Total 570 users
- Around 2000 materials uploaded on the platform

BUSINESS PRIORITIES

- Onboard more suppliers to improve the variety of materials on the platform
- Increase engagement and conversion rate of the brand users and the sales of materials

* Name changed to for client privacy



Problem Statement

BUILDING A RECOMMENDATION SYSTEM

Help Fabric Central improve their sales by increasing customer engagement, by implementing a recommendation system, which will suggest relevant fabrics at appropriate touchpoints in the web application.



CONTENT BASED FILTERING

The system finds the similarity between products based on its features.

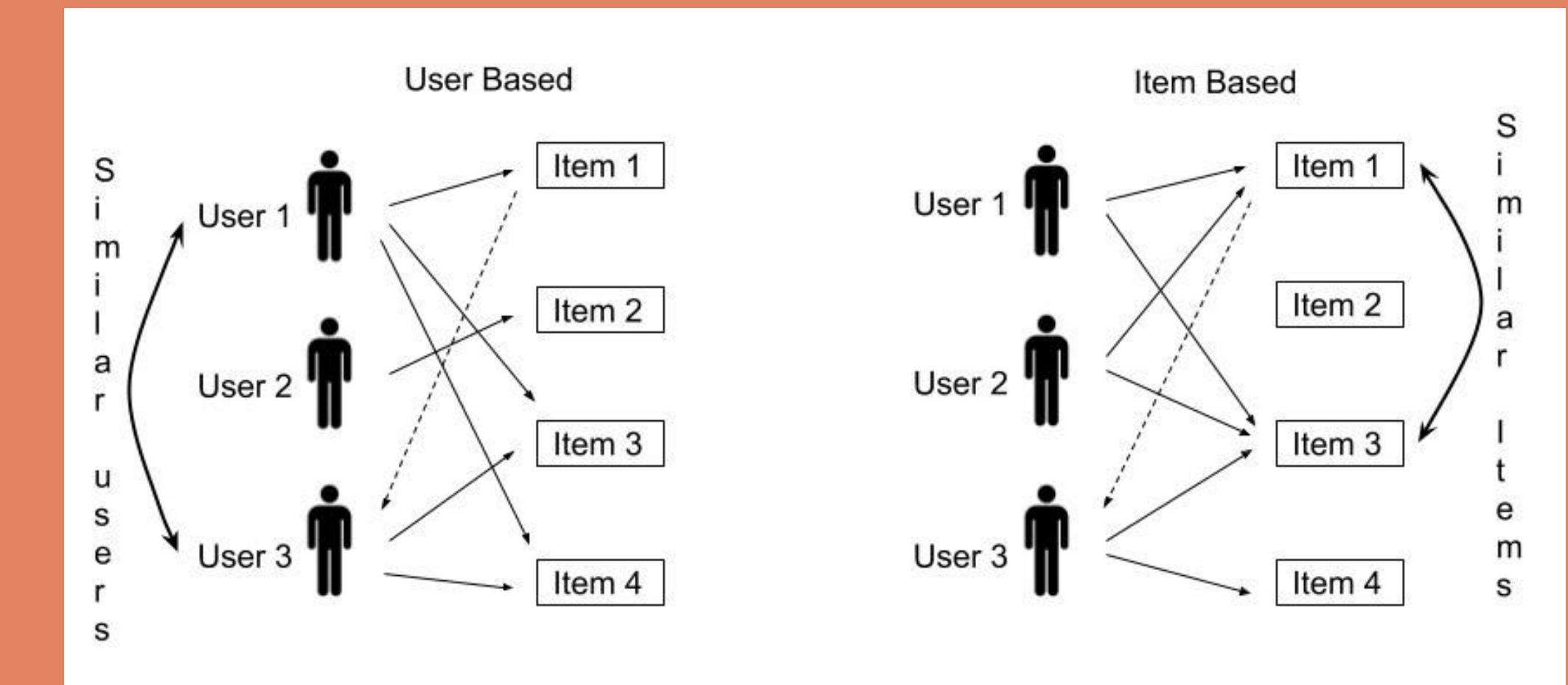
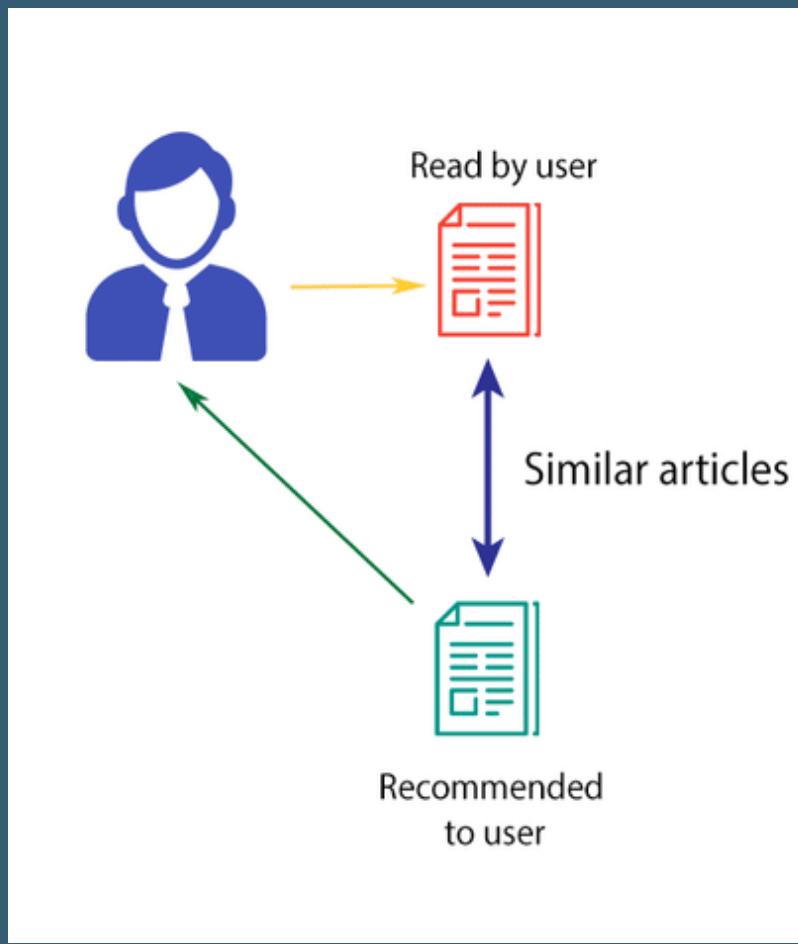
COLLABORATIVE FILTERING

Recommendations are done based on the user's behavior.

Item Based

User Based

CONTENT BASED VS COLLABORATIVE FILTERING



DATASETS

- Materials (1954 x 46)
- Fabric requests (1063 x 23)
- Orders (95 x 18)
- Users (570 x 22)
- Brands (359 x 19)

DATASETS

- ◆ Materials (1954 x 46)
- ◆ Fabric requests (1063 x 23)
- ◆ Orders (95 x 18)
- ◆ Users (570 x 22)
- ◆ Brands (359 x 19)

DATA CLEANING



HANDLING NULL VALUES

- Drop columns with too many null values
- Impute values for fabric weight for leather fabrics as 500 g/m-sq.
- Impute values for other weight values as mean of the fabric blend.

REPLACE NULL VALUES FOR CATEGORICAL VARIABLES

Many materials had missing values for categorical variables such as fabric_weave, fabric_blend etc. Instead of dropping these, the values were replaced with "blank" values.

REMOVE ARCHIVED / UNPUBLISHED MATERIALS

Some materials which were archived by the suppliers or marked with Published = False were removed

DROP FABRIC REQUESTS MADE BY DEMO USERS

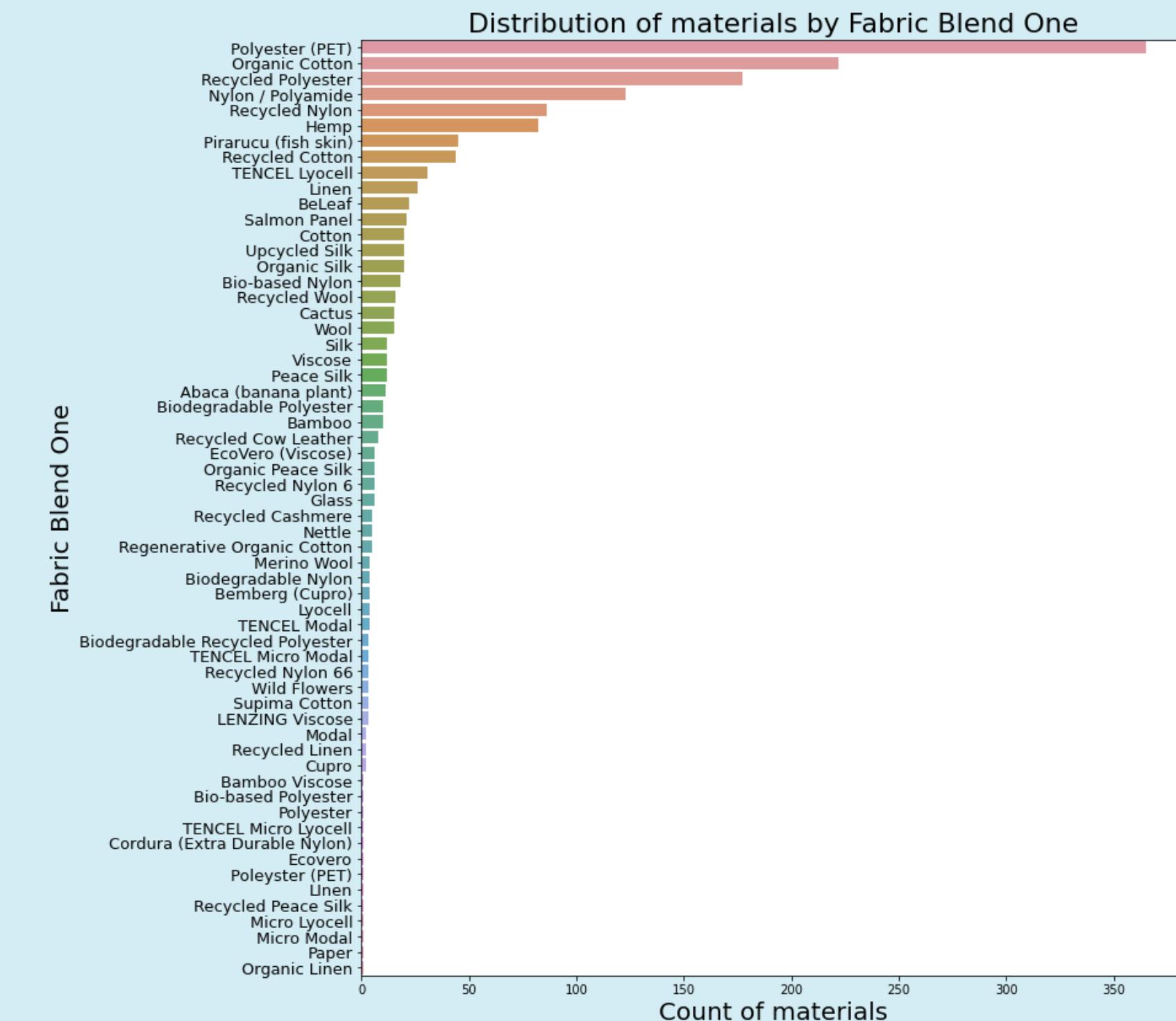
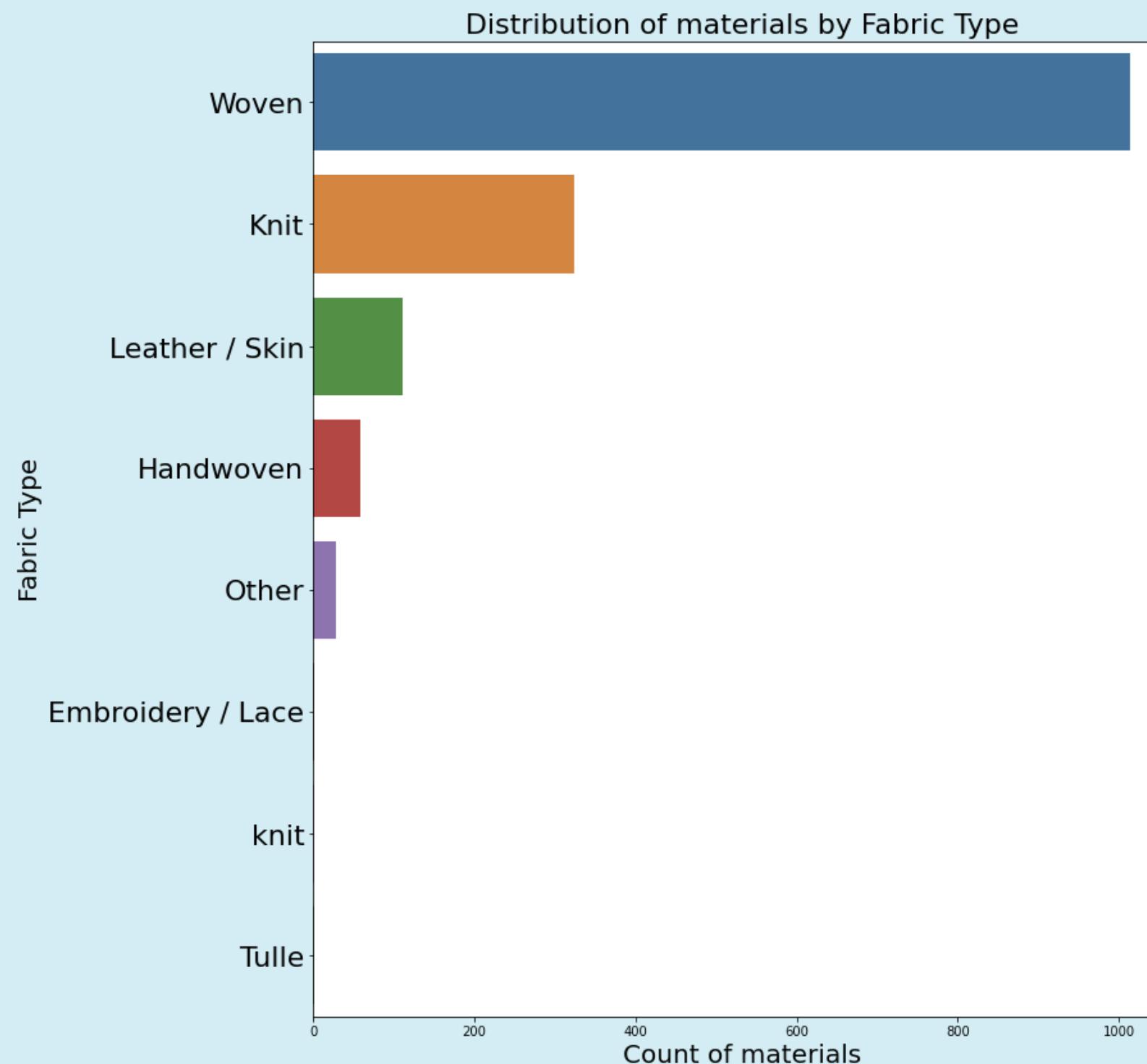
Many fabric requests were initiated by Techstyle staff, these couldn't be used as they would have skewed the recommendations.

MATERIALS
1539 Materials

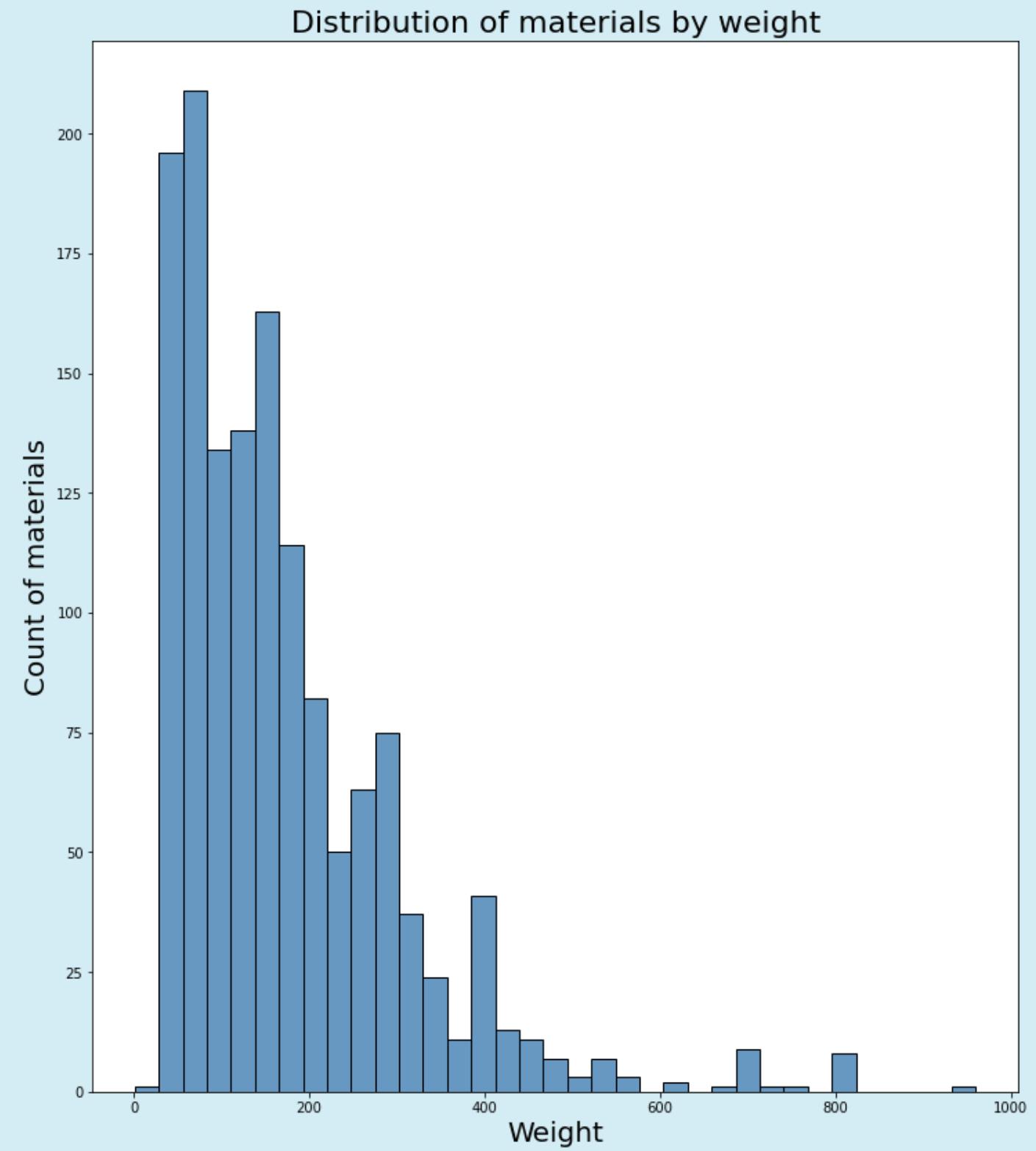
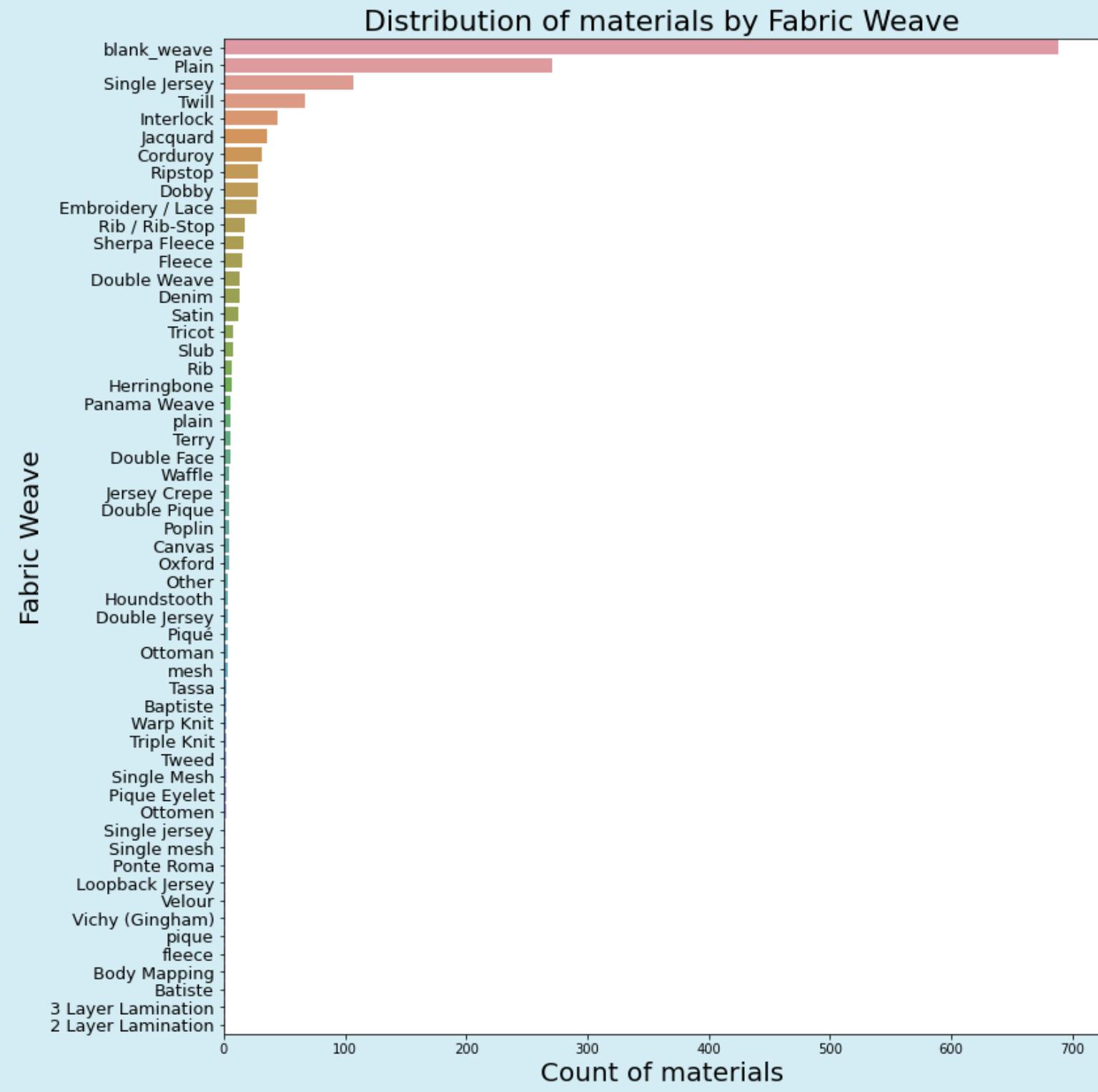
FABRIC REQUESTS

808 Fabric Requests

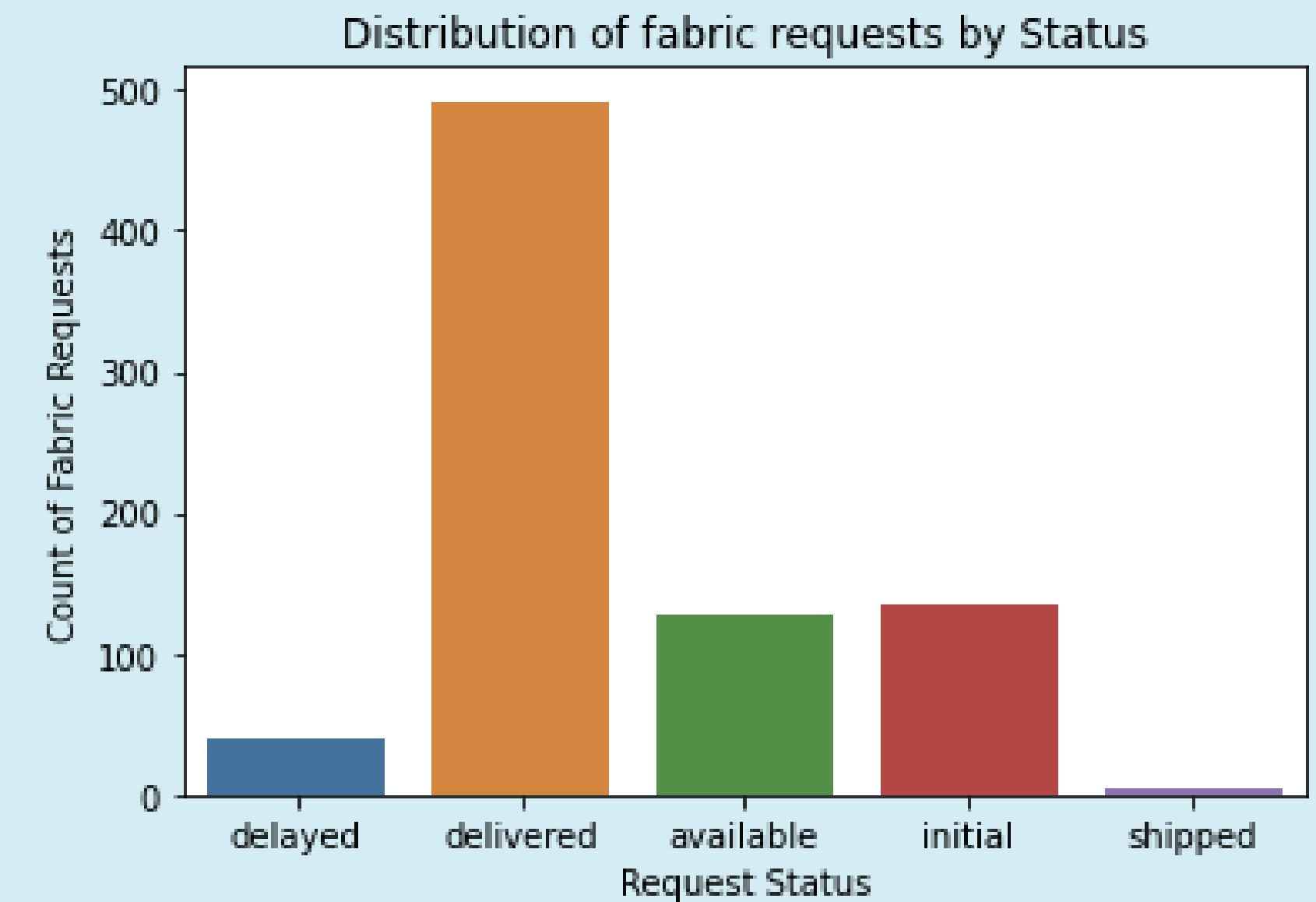
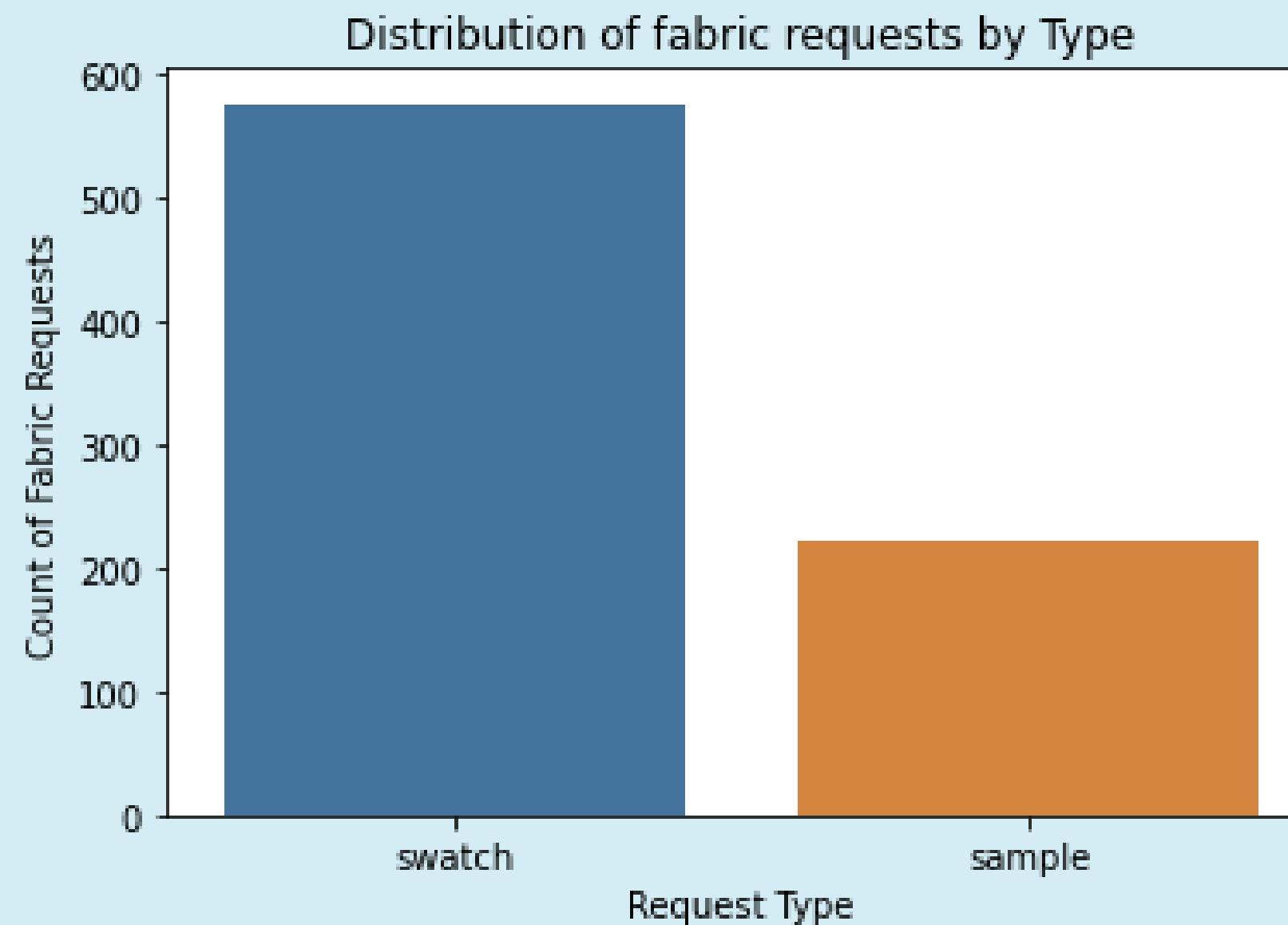
EDA - MATERIALS



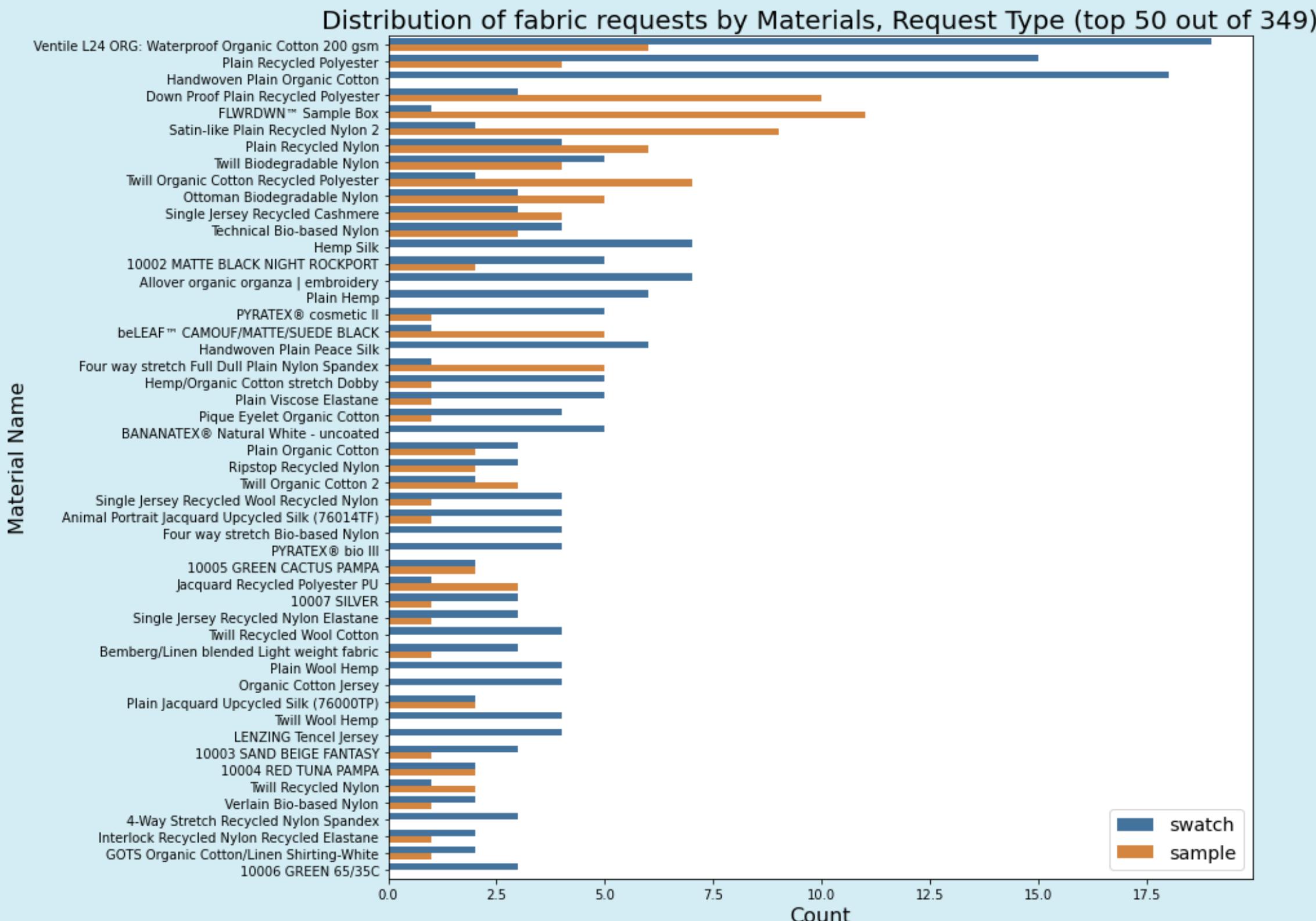
EDA - MATERIALS



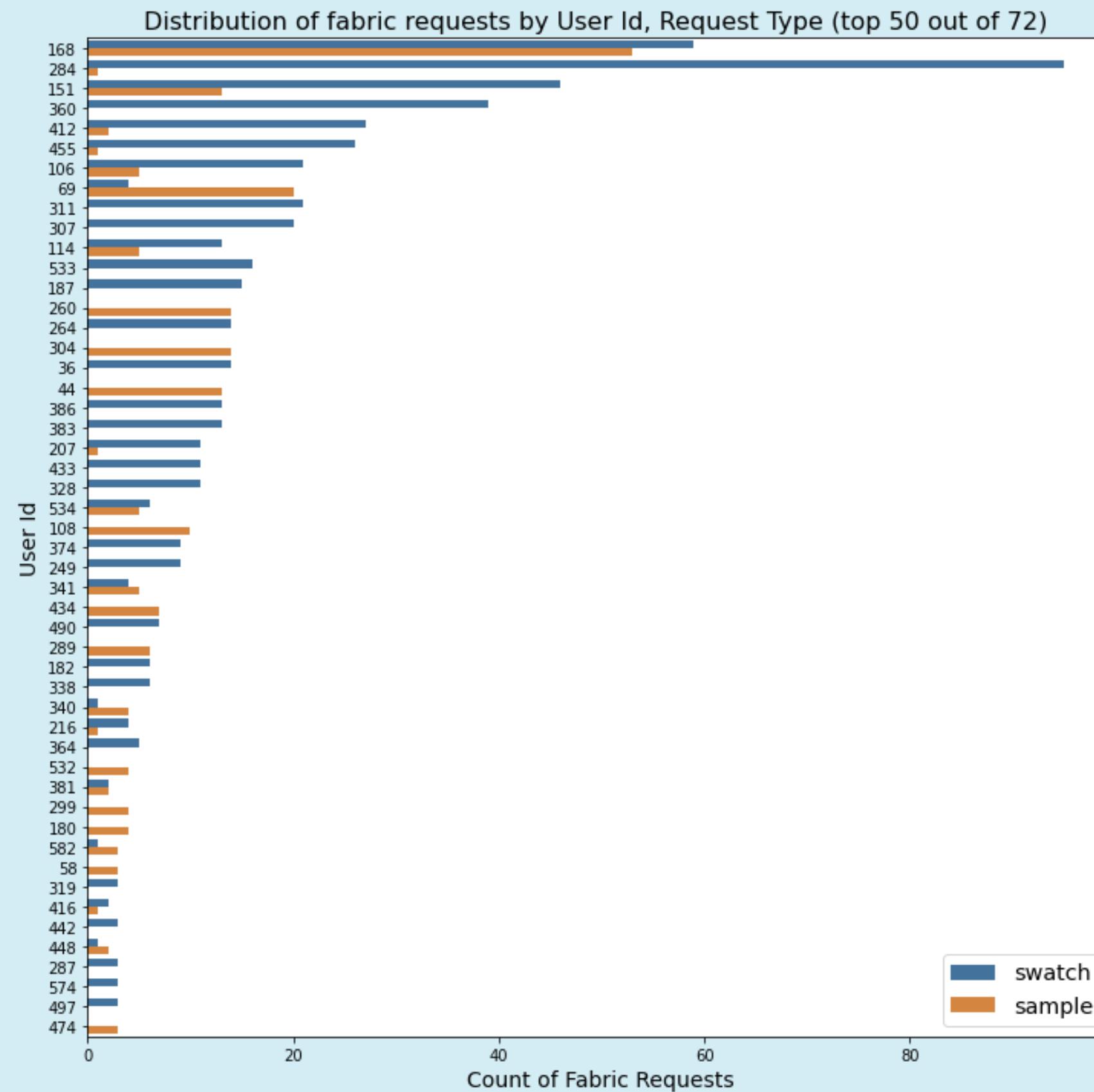
EDA - FABRIC REQUESTS



EDA - FABRIC REQUESTS



EDA - FABRIC REQUESTS



FEATURE ENGINEERING & PRE-PROCESSING



COMBINE PROPERTIES & END-USAGES

Properties and EndUses were split across multiple columns e.g. end_use/0, end_use/1 etc. These were combined into properties and end_use columns

ASSIGN HIGHER WEIGHT FOR SAMPLE FABRIC REQUESTS

Assign weights to Fabric Requests such that Sample Requests are given higher importance than swatch requests. These weights will be used in calculating the user-item matrix

SCALING THE DATA

Standard Scaler was used to standardize the data for Clustering

MinMax Scaler was used to standardize the data for Cosine Similarity

FEATURE SELECTION

MATERIALS

ID
CREATED_AT
UPDATED_AT
CODE
FABRIC_WEAVE
CURRENCY
SUPPLIER_ID
PRICE
NAME
FABRIC_BLEND_ONE
FABRIC_BLEND_ONE_PERCENT
FABRIC_BLEND_TWO
FABRIC_BLEND_TWO_PERCENT
FABRIC_BLEND_THREE
FABRIC_BLEND_THREE_PERCENT
WEIGHT
WEIGHT_UNIT
END_USE
FABRIC_TYPE
PRICE_UNIT
STATUS
DELETED_AT
ONLY_FOR_BRANDS
PUBLISHED
DESCRIPTION
DISPATCH_TIME_SAMPLE
FACTORY_LOCATION
CONSTRUCTION
WIDTH
WIDTH_UNIT
FINISHING
PROPERTIES
DISPATCH_TIME_MASS_QUANTITY_MIN
DISPATCH_TIME_MASS_QUANTITY_MAX
ORIGIN
COLOUR
TSA_STAMPS
MINIMAL_ORDER_QUANTITY
CROWD_SOURCE_AVAILABLE
MOQ_DETAILS
UUID
NOTIFY_FC_TEAM
MOQ
SWATCH_PRICE
SWATCH_QUANTITY
ARCHIVED_AT

FABRIC REQUESTS

ID
STATUS
TYPE
QUANTITY
REQUESTED_AT
CLOSED_AT
DELAY_DAYS
DELAY_NOTE
MATERIAL_ID
USER_ID
PROJECT_ID
CREATED_AT
UPDATED_AT
ORDER_ID
PRICE
PRICE_UNIT
SURCHARGE
ACCEPTED_AT
DELETED_AT
ORIGIN_QUANTITY
VARIANT_ID
CURRENCY
BOUGHT_PRICE



FEATURE SELECTION

MATERIALS

ID
CREATED_AT
UPDATED_AT
CODE
FABRIC_WEAVE
CURRENCY
SUPPLIER_ID
PRICE
NAME
FABRIC_BLEND_ONE
FABRIC_BLEND_ONE_PERCENT
FABRIC_BLEND_TWO
FABRIC_BLEND_TWO_PERCENT
FABRIC_BLEND_THREE
FABRIC_BLEND_THREE_PERCENT
WEIGHT
WEIGHT_UNIT
END_USE
FABRIC_TYPE
PRICE_UNIT
STATUS
DELETED_AT
ONLY_FOR_BRANDS
PUBLISHED
DESCRIPTION
DISPATCH_TIME_SAMPLE
FACTORY_LOCATION
CONSTRUCTION
WIDTH
WIDTH_UNIT
FINISHING
PROPERTIES
DISPATCH_TIME_MASS_QUANTITY_MIN
DISPATCH_TIME_MASS_QUANTITY_MAX
ORIGIN
COLOUR
TSA_STAMPS
MINIMAL_ORDER_QUANTITY
CROWD_SOURCE_AVAILABLE
MOQ_DETAILS
UUID
NOTIFY_FC_TEAM
MOQ
SWATCH_PRICE
SWATCH_QUANTITY
ARCHIVED_AT

FABRIC REQUESTS

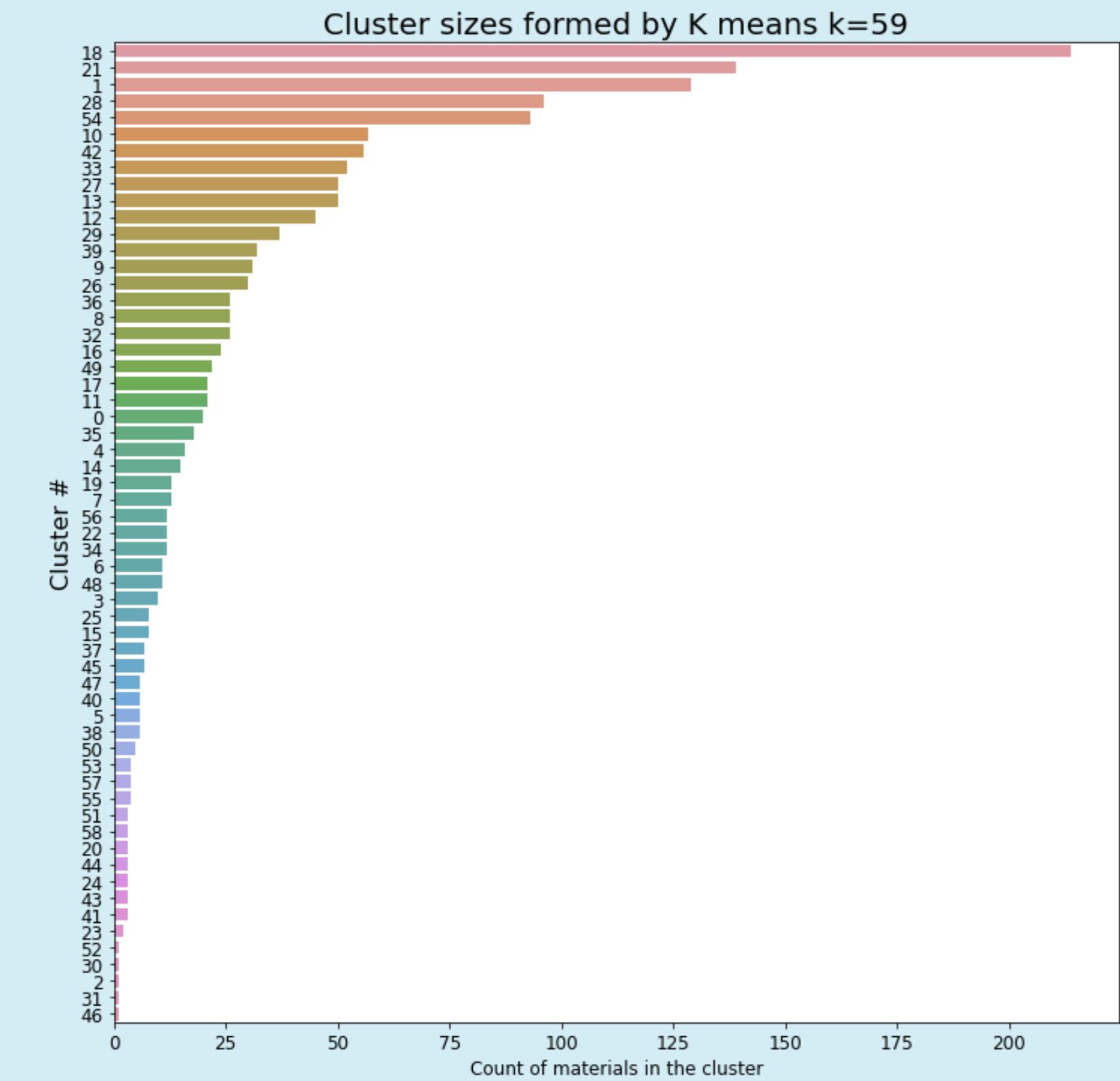
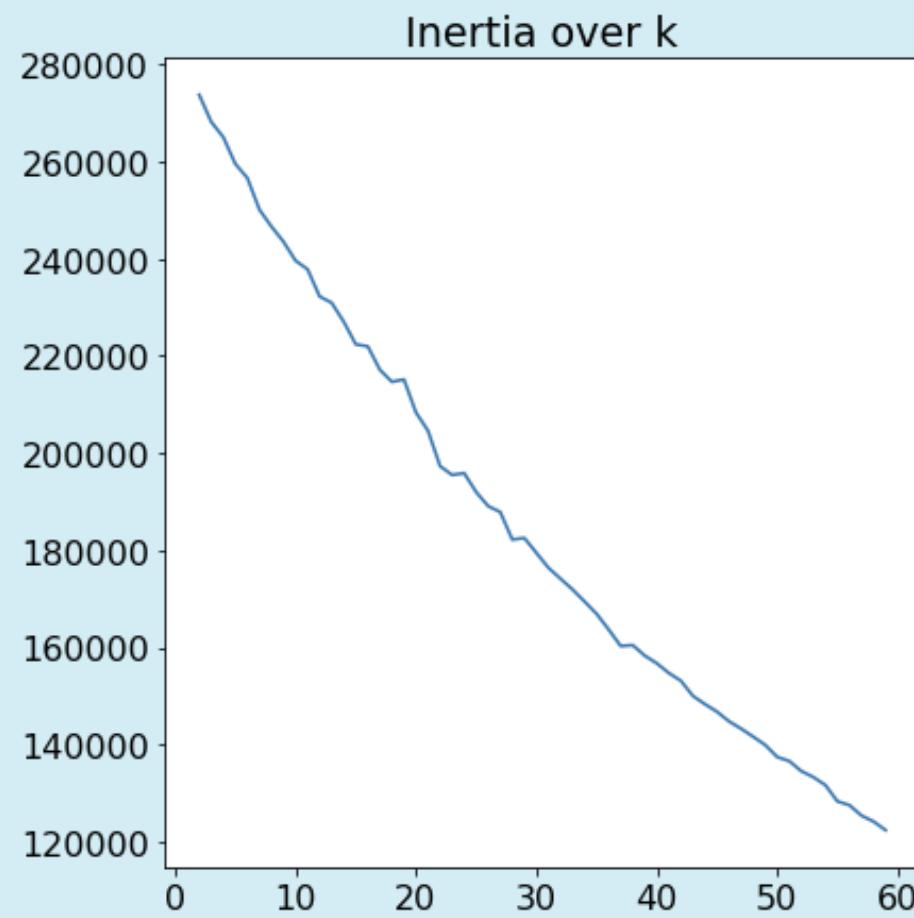
ID
STATUS
TYPE
QUANTITY
REQUESTED_AT
CLOSED_AT
DELAY_DAYS
DELAY_NOTE
MATERIAL_ID
USER_ID
PROJECT_ID
CREATED_AT
UPDATED_AT
ORDER_ID
PRICE
PRICE_UNIT
SURCHARGE
ACCEPTED_AT
DELETED_AT
ORIGIN_QUANTITY
VARIANT_ID
CURRENCY
BOUGHT_PRICE



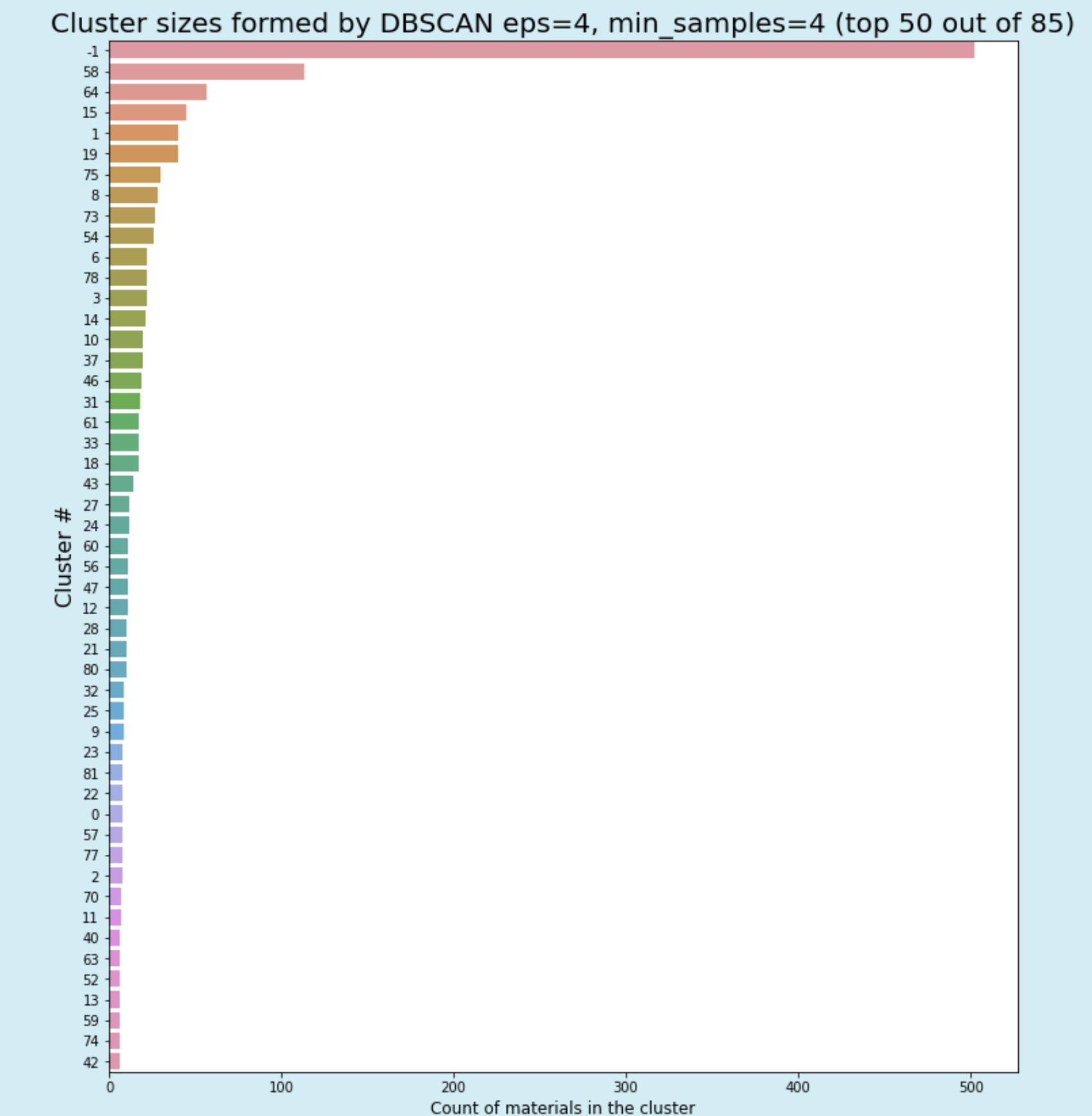
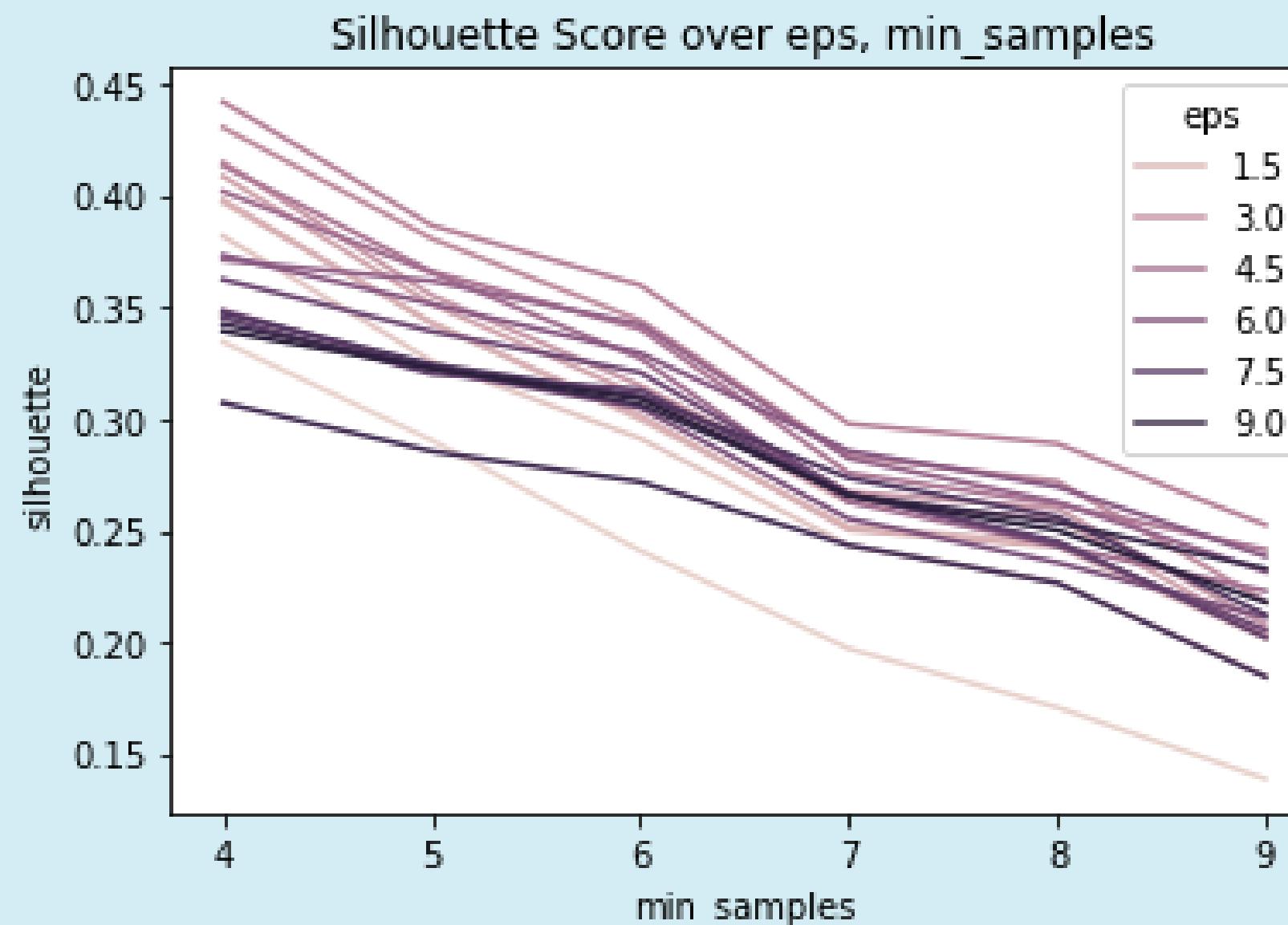
Building the recommendation system

- ◆ Content Based Filtering
("You may also like" or "More like this")
 - ◆ Clustering
 - ◆ Cosine Similarity
- ◆ Collaborative Filtering
("Users who liked this also liked")
 - ◆ Nearest Neighbours

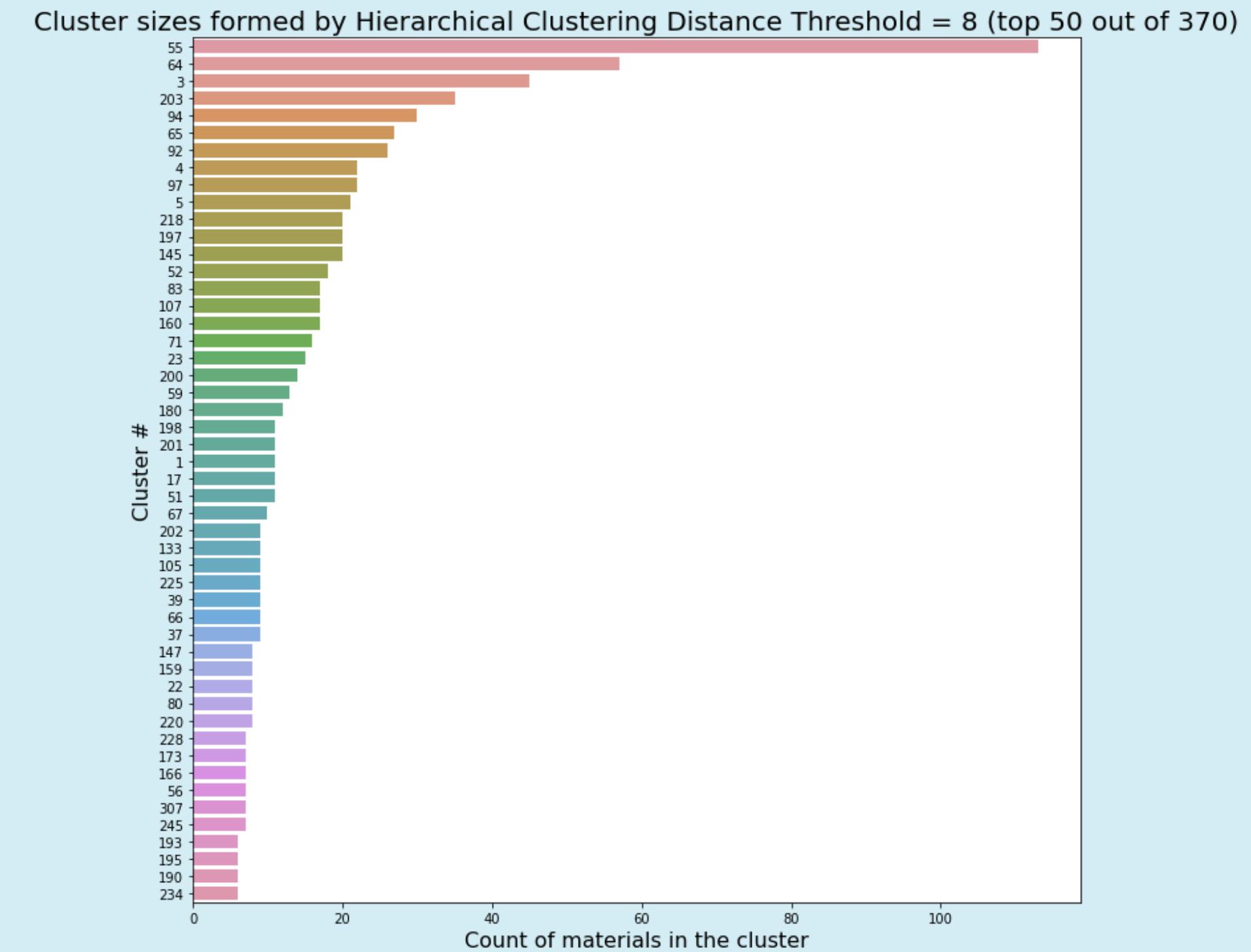
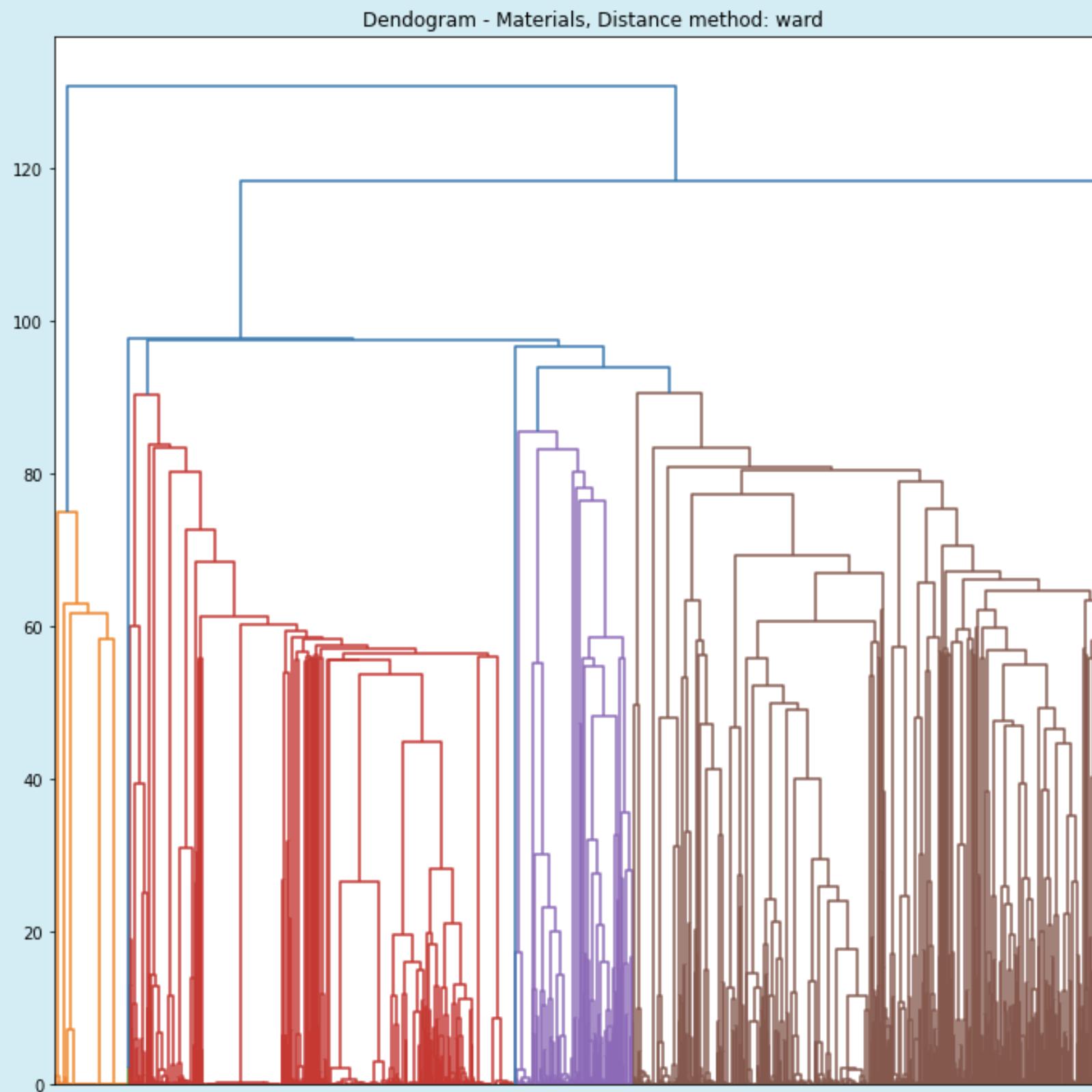
CLUSTERING- K Means



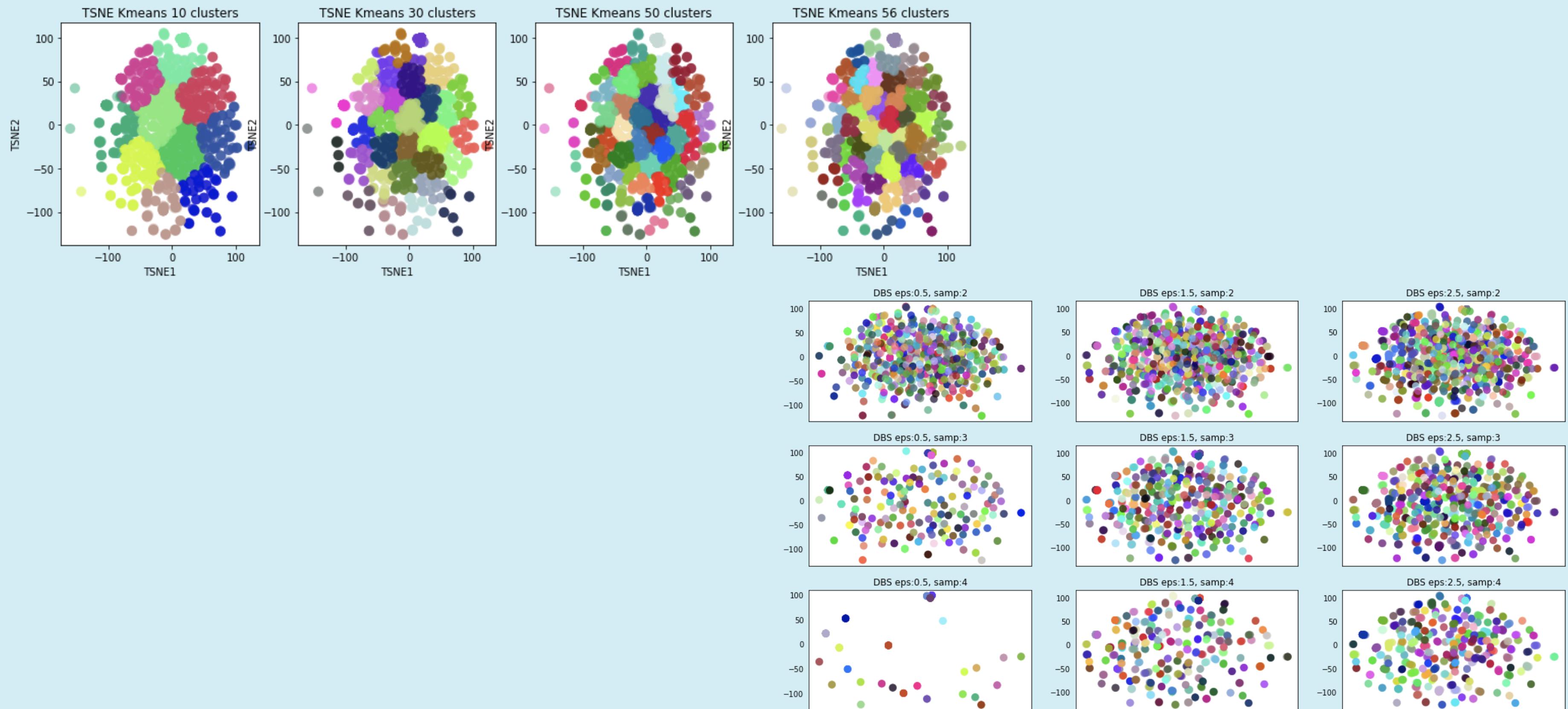
CLUSTERING- DB Scan



CLUSTERING- Hierarchical Agglomerative



CLUSTERING- TSNE with Kmeans and DBSCAN



Clustering Conclusion

- ◆ Clusters were uneven in size, ranging from too big to too small
- ◆ Visualisation shows no natural clusters being formed
- ◆ Not suitable for recommendations here

CONTENT BASED- Cosine Similarity

- Cosine Similarity is a measurement that quantifies the similarity between two or more vectors.
- Steps to get recommendations based on cosine similarity:
 - Build a Cosine Similarity matrix

id	2253	557	947	2276	1840	1814	1924	1255	217	2488
2253	1	0.19830528	0.26619967	0.19044629	0.0560013	0.13138872	0.17310236	0.00731087	0.00843747	0.30036669
557	0.19830528	1	5.63E-05	0.00051157	0.00130634	0.00045321	0.00059303	0.54723284	7.40E-05	0.29814224
947	0.26619967	5.63E-05	1	0.22834334	0.03098632	0.00190828	0.00316041	0.00396482	0.0047149	0.19469331
2276	0.19044629	0.00051157	0.22834334	1	0.06970018	0.0045703	0.0074517	0.16067208	0.01051266	0.00692018
1840	0.0560013	0.00130634	0.03098632	0.06970018	1	0.01002807	0.01629304	0.01971949	0.02268617	0.01503643
1814	0.13138872	0.00045321	0.00190828	0.0045703	0.01002807	1	0.65505059	0.00146545	0.0014208	0.28944526
1924	0.17310236	0.00059303	0.00316041	0.0074517	0.01629304	0.65505059	1	0.00232091	0.00234306	0.37904569
1255	0.00731087	0.54723284	0.00396482	0.16067208	0.01971949	0.00146545	0.00232091	1	0.00291653	0.13783733
217	0.00843747	7.40E-05	0.0047149	0.01051266	0.02268617	0.0014208	0.00234306	0.00291653	1	0.00221959
2488	0.30036669	0.29814224	0.19469331	0.00692018	0.01503643	0.28944526	0.37904569	0.13783733	0.00221959	1

- To get recommendations for a material A, filter by material id e.g. 1814, and get top n materials when ordered by cosine similarity in descending order.

id	2253	557	947	2276	1840	1814	1924	1255	217	2488
2253	1	0.19830528	0.26619967	0.19044629	0.0560013	0.13138872	0.17310236	0.00731087	0.00843747	0.30036669
557	0.19830528	1	5.63E-05	0.00051157	0.00130634	0.00045321	0.00059303	0.54723284	7.40E-05	0.29814224
947	0.26619967	5.63E-05	1	0.22834334	0.03098632	0.00190828	0.00316041	0.00396482	0.0047149	0.19469331
2276	0.19044629	0.00051157	0.22834334	1	0.06970018	0.0045703	0.0074517	0.16067208	0.01051266	0.00692018
1840	0.0560013	0.00130634	0.03098632	0.06970018	1	0.01002807	0.01629304	0.01971949	0.02268617	0.01503643
1814	0.13138872	0.00045321	0.00190828	0.0045703	0.01002807	1	0.65505059	0.00146545	0.0014208	0.28944526
1924	0.17310236	0.00059303	0.00316041	0.0074517	0.01629304	0.65505059	1	0.00232091	0.00234306	0.37904569
1255	0.00731087	0.54723284	0.00396482	0.16067208	0.01971949	0.00146545	0.00232091	1	0.00291653	0.13783733
217	0.00843747	7.40E-05	0.0047149	0.01051266	0.02268617	0.0014208	0.00234306	0.00291653	1	0.00221959
2488	0.30036669	0.29814224	0.19469331	0.00692018	0.01503643	0.28944526	0.37904569	0.13783733	0.00221959	1

CONTENT BASED- Cosine Similarity

Original

	id	name	fabric_blend_one	price	weight	supplier_id	end_use	properties	
0	2253	Ripstop Organic Cotton	Organic Cotton	11.9	310.0	23	['Coat', 'Jacket', 'Pants']		

Recommendations

	id	name	fabric_blend_one	price	weight	supplier_id	end_use	properties	
	139	2271	Organic Cotton Polyester	Organic Cotton	14.95	340.0	23	['Coat', 'Jacket', 'Pants', 'Shorts']	
	229	2254	Twill Organic Cotton	Organic Cotton	12.90	320.0	23	['Coat', 'Jacket', 'Pants', 'Shorts']	
	891	491	Twill Organic Cotton 2	Organic Cotton	17.90	260.0	23	['Coat', 'Jacket', 'Pants', 'Shorts']	
	909	488	Plain Organic Cotton Recycled Polyester	Organic Cotton	9.50	245.0	23	['Coat', 'Jacket', 'Pants', 'Shorts']	
	1145	490	Twill Organic Cotton 1	Organic Cotton	18.90	320.0	23	['Coat', 'Jacket', 'Pants', 'Shorts', 'Trench ...']	
	554	487	Twill Organic Cotton Polyester 1	Organic Cotton	11.95	215.0	23	['Coat', 'Jacket', 'Pants', 'Shorts']	['Water Repellent (W/R)']
	1125	659	Plain Organic Cotton	Organic Cotton	12.50	145.0	23	['Coat', 'Jacket', 'Pants', 'Shorts']	['Water Repellent (W/R)']
	688	2149	Organic Cotton	Organic Cotton	11.88	420.0	57	['Coat', 'Jacket', 'Pants']	
	45	2270	Twill Polyester	Polyester (PET)	7.85	320.0	23	['Coat', 'Jacket', 'Pants']	
	82	489	Twill Organic Cotton Recycled Polyester	Organic Cotton	11.20	205.0	23	['Jacket', 'Coat']	['Water Repellent (W/R)']

Content Based Recommendations results

- ◆ Recommendations generated for all materials and shared with Domain Expert in csv format for evaluation
- ◆ Pro - No cold start problem, Recommendations are available for all materials
- ◆ Con - Does not consider user behaviour

COLLABORATIVE FILTERING (ITEM BASED)

Nearest Neighbours

- Nearest Neighbours is an unsupervised learning algorithm which takes a sparse matrix with m samples and n features and finds nearest samples based on the features
- Steps to get recommendations based Nearest neighbours:
 - Merge Materials and Fabric Request tables
 - Build a pivot table with material id as rows, user id as columns, and intersection as count of fabric requests (sample requests were assigned higher weight)

user_id	0.0	36.0	44.0	58.0	69.0	87.0	101.0	106.0	108.0	114.0	...
id											
1172	0.0	1.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...
1173	0.0	1.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...
1174	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...
1175	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...
1176	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...
...

- Convert Pivot Table to a compressed sparse matrix and pass to Sklearn Nearest Neighbour algorithm, to get k nearest neighbours

COLLABORATIVE FILTERING (ITEM BASED)

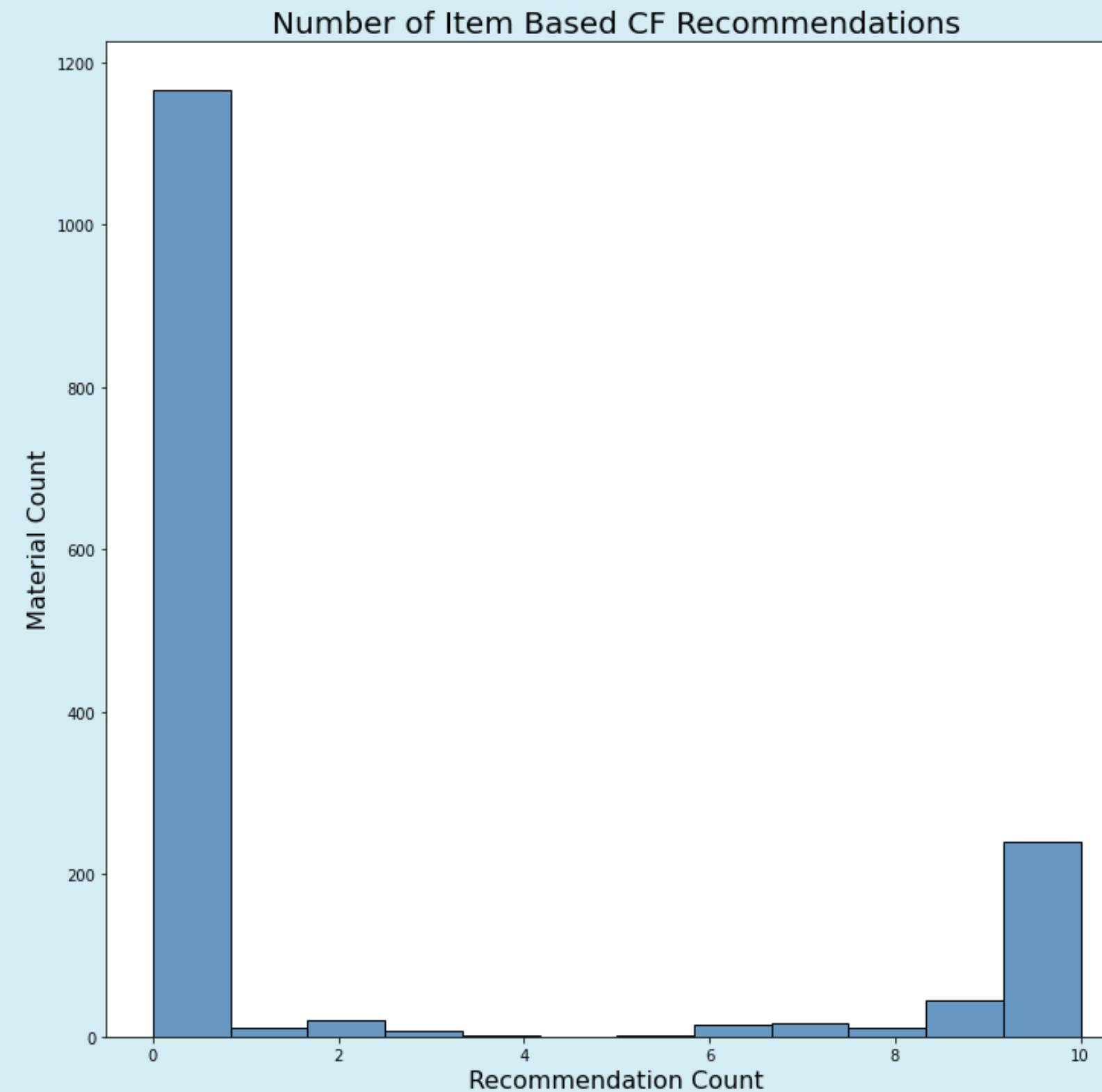
Original

id	name	fabric_blend_one	price	weight	supplier_id	end_use	properties
82 489	Twill Organic Cotton Recycled Polyester	Organic Cotton	11.2	205.0	23	['Jacket', 'Coat']	['Water Repellent (W/R)']

Recommendations

id	name	fabric_blend_one	price	weight	supplier_id	end_use	properties
90 650	Ottoman Biodegradable Nylon	Biodegradable Nylon	17.50	260.0	23	['Coat', 'Softshell Jacket', 'Jacket', 'Trench...']	['Water Repellent (W/R)']
141 2269	Recycled Polyester PU	Recycled Polyester	16.55	120.0	23	['Softshell Jacket', 'Jacket']	
216 381	Gold-Q Membrane Plain Polyester	Polyester (PET)	5.95	106.0	18		['Durable Water Repellent (DWR)']
469 2265	Reflective Glass Polyester	Glass	24.45	420.0	23	['Coat', 'Jacket']	['Reflective']
563 656	Plain Recycled Polyester	Recycled Polyester	9.80	280.0	23	['Coat', 'Jacket', 'Softshell Jacket']	['Water Repellent (W/R)']
649 495	Plain Recycled Nylon 1	Recycled Nylon	13.90	80.0	23	['Jacket', 'Softshell Jacket']	['Water Repellent (W/R)']
1028 496	Plain Recycled Polyester PU	Recycled Polyester	15.90	135.0	23	['Jacket', 'Softshell Jacket']	['Water Repellent (W/R)', 'Waterproof (W/P)', ...]
1290 1236	Summer Jersey	Hemp	23.40	190.0	43	['Base Layer', 'Sweatshirt', 'T-Shirt']	
1472 662	Dry-dyed Plain Nylon	Recycled Nylon	9.95	110.0	23	['Jacket', 'Swimming Shorts']	['Water Repellent (W/R)']
1509 1798	Plain Wool Hemp	Wool	66.00	370.0	57	['Jacket', 'Pants', 'Shirt']	['Breathable']

COLLABORATIVE FILTERING (ITEM BASED)

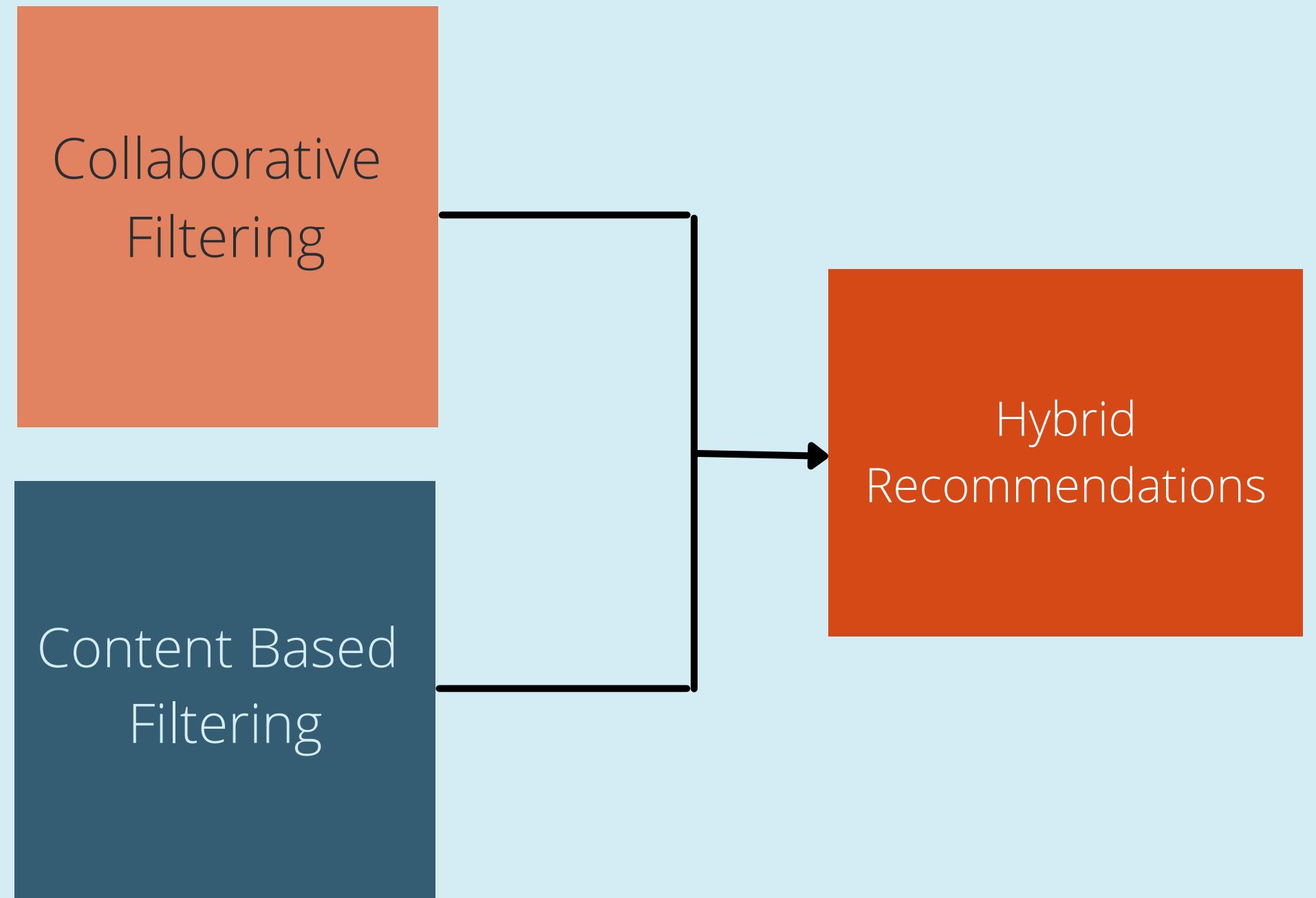


Item Based Collaborative Filtering results

- ❖ Recommendations generated for all materials which have fabric requests and shared with Domain Expert for evaluation.
- ❖ Pro - Takes into consideration user behaviour and can recommend similar as well as complementary materials
- ❖ Con - Insufficient data / cold start problem

Hybrid Recommendations

To avoid the cold start
problem



Evaluation

Post Deployment

EVALUATION USING DOMAIN KNOWLEDGE

Recommendations were shared with domain experts and materials were sampled and recommendations verified

TRACKING USER CLICKS

Tracking user clicks from specific web pages to measure metrics such as Click Through Rate, Conversion Rate etc.

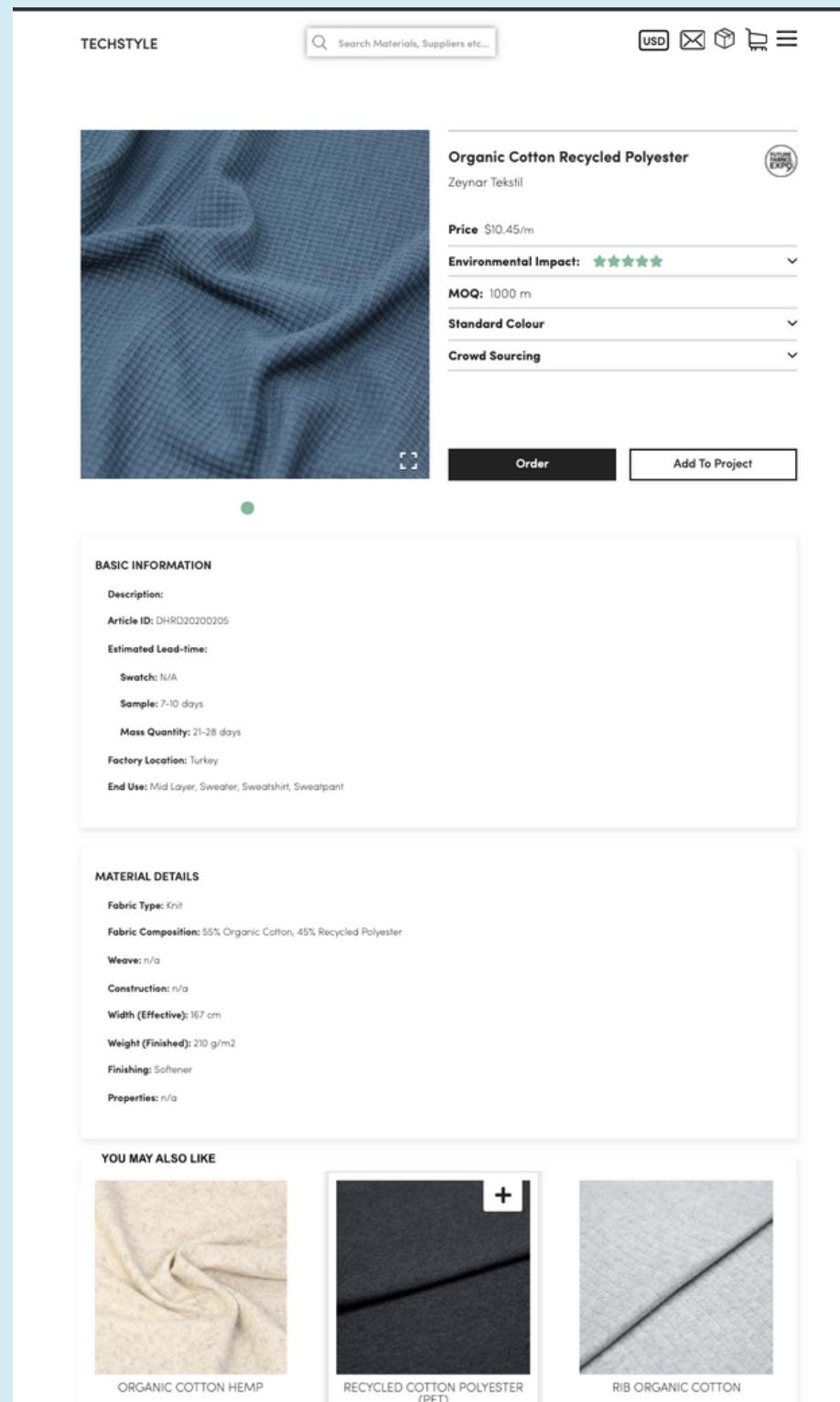
A/B TESTING

Comparing recommendations from various methods by using A/B testing.

Conclusion & Recommendations

- Initial evaluation by domain experts, shows satisfactory recommendations for item based Collaborative Filtering.
- Further fine tuning in progress for assigning higher weights to specific features in content based filtering
- Integration & Deployment planned in October 2021

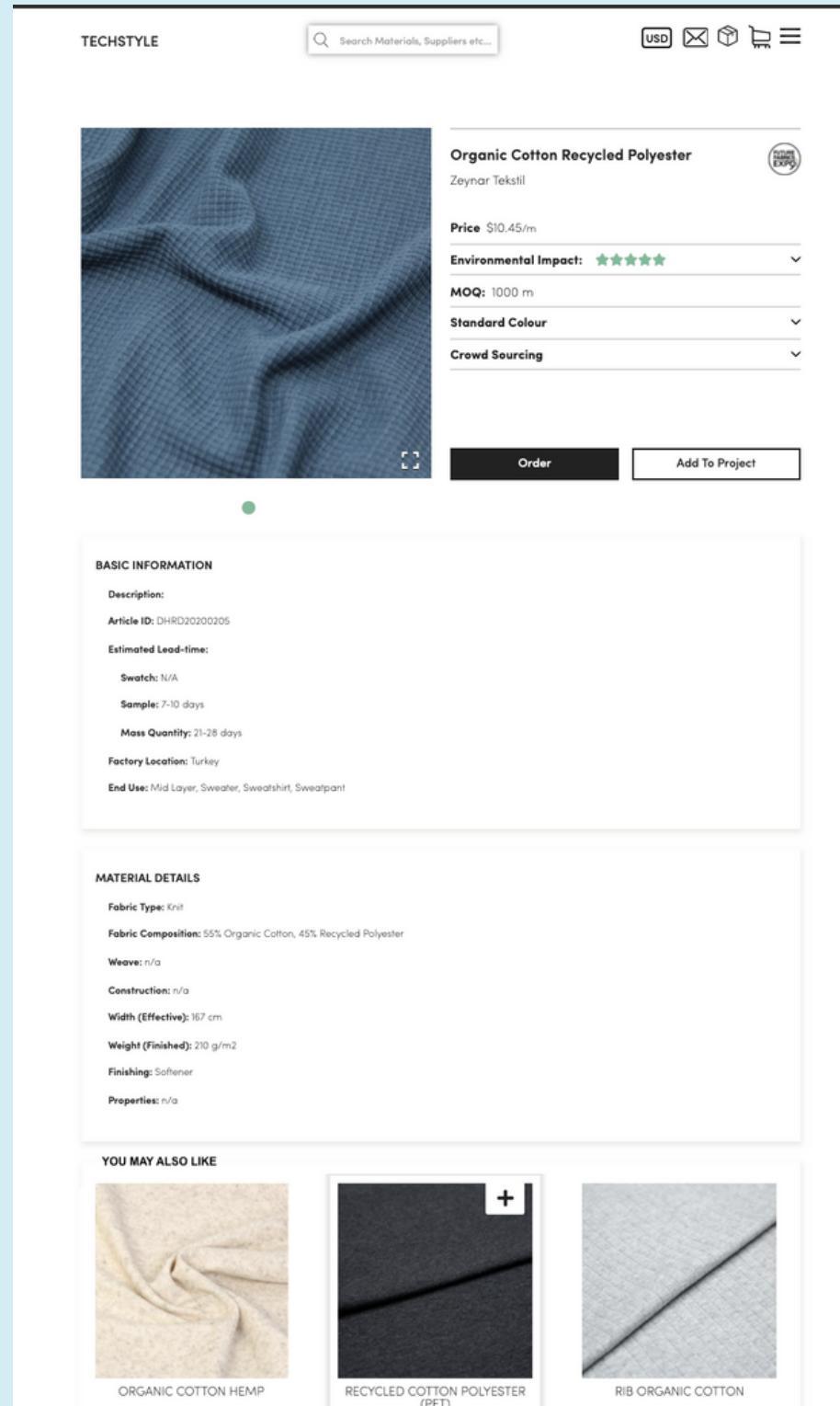
Conclusion & Recommendations



Where to display the Recommendations?

- Content Based Materials can be displayed on Material Search page
- Hybrid Recommendations can be displayed on the materials details page
- Optionally, Collaborative Filtering recommendations should be displayed on Order Completion page
- Targetted email campaigns to recommend materials to users

Conclusion & Recommendations



Other Recommendations:

- User clicks from recommendations, should be tracked upon deployment of the recommendation system, to assess performance
- Implement explicit user feedback such as reviews/ratings for materials to get more accurate recommendations
- Gather more brand attributes to enable better user based collaborative filtering

FUTURE STEPS

- ◆ Integration & Deployment of the recommendation system to integrate with existing web application
- ◆ User Based Collaborative Filtering, based on similarity of user attributes
- ◆ Explore other algorithms for item based collaborative filtering like Matrix Factorization, Deep Learning
- ◆ Implement A/B testing to assess relative performance of various models



**THANK
YOU!**