Which Dove branded products to promote in order to increase overall sales for Dove

EMORY
GOIZUETA BUSINESS SCHOOL
Master of Science in Business Analytics MSBA

MKT_680_4100: Recommender Systems

Group 5

Yi Cai

Sameera Vattikuti

Nicholas Yang

Zeyu Zhu

Table of Contents

Chapter 1: Introduction

1.1 Business Background

Pernalonga, a leading supermarket chain with 421 stores which sells ~10K products in 430 categories with a consumer base of ~7900. Currently, 30% of their sales are through promotions (mostly in-store) executed in partnership with suppliers.

Unilever, one such supplier, wants to find out which Dove branded products to promote in order to increase overall sales for the Dove brand. More specifically, they want a data supported marketing campaign that is scheduled to run for two weeks in April 2019.

1.2 Problem Statement

The business problem could be translated to a data problem statement as 'Build a recommender system to identify consumer-product combo in order to promote brands of Dove'

Chapter 2: Overview

2.1 Data Overview

For the purpose of our analysis, we have transactional data related to ~2.69 Million unique transactions and 29.6 Million line-items. This data corresponds to 421 stores, selling ~10000 unique products belonging to 429 categories. The stores serviced 7920 unique customers in the given time period.

2.1.1 Assumptions & Considerations

- To uniquely identify each transaction, we created a **unique transaction id** that is a combination of customer id, transaction date and store id with the assumption that a customer visits a store only once in a day.
- Since the available data does not convey meaningful definitions of **sub-categories**, we choose to focus on a product or a category level analysis only.
- Since all products (from both Dove vs competitors) in the operating segment of dove only have one **SKU – CT,** we will not address it in further analysis.

2.2 Business Overview

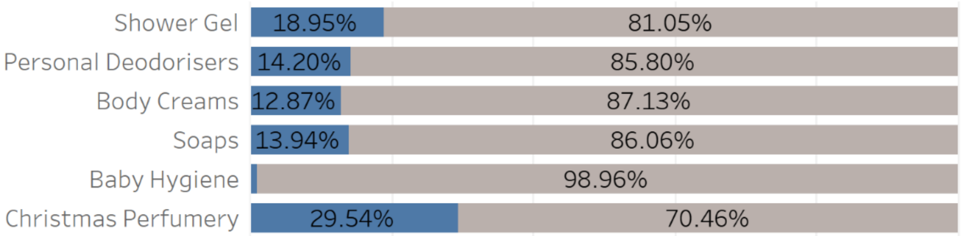
2.2.1 Dove Overview

Sales (by revenue & volume)

Dove Sales				
	Revenue		Volume/Qty	
Row Labels	**Revenue**	**% Contribution**	**Volume/Qty**	**% Contribution**
SHOWER GEL	48864	45%	14316	37%
PERSONAL DEODORISERS	33110	31%	15380	40%
BODY CREAMS	12399	11%	2953	8%
SOAPS	11165	10%	5192	14%
OUT. BABY HYGIENE ITEMS	1569	1%	424	1%
CHRISTMAS PERFUMERY	799	1%	126	0%
Grand Total	**107906**	**100%**	**38391**	**100%**

Dove operates in 6 categories with 52 product units. It derives ~78% of sales by volume and ~76% sales by revenue from 50% of its products. Product 999527671 alone accounts for 12% of its sales revenue. The categories shower gel (45%) and Personal deodorizers (31%) together contribute to 76% of its sales revenue (after discount) closely followed by body creams (11%) and soaps (10%). We focus more on the former 3 categories as they are high-value products when compared to soap and therefore generate more sales revenue per additional purchase when a consumer accepts the marketing promotion.

Market Share

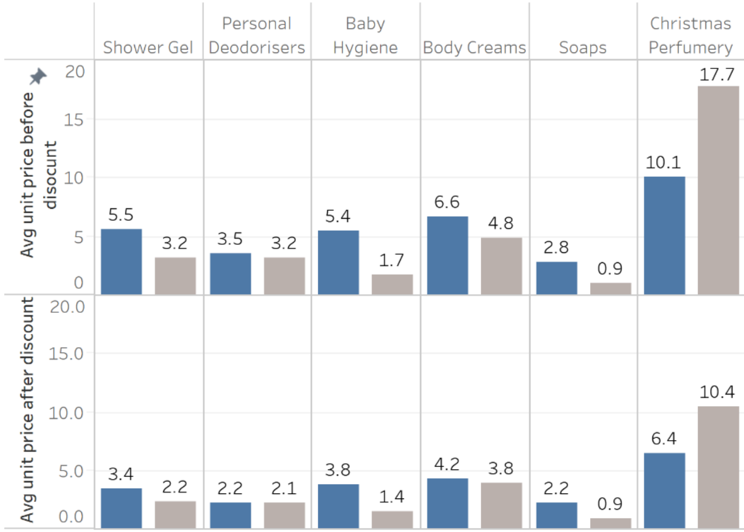


Considering its FMCG, Dove has a high market share in the range of 12 – 18% in most categories especially in the categories that contribute most to its sales revenue. It has extremely low market share in baby hygiene which could be a category to explore promotions on.

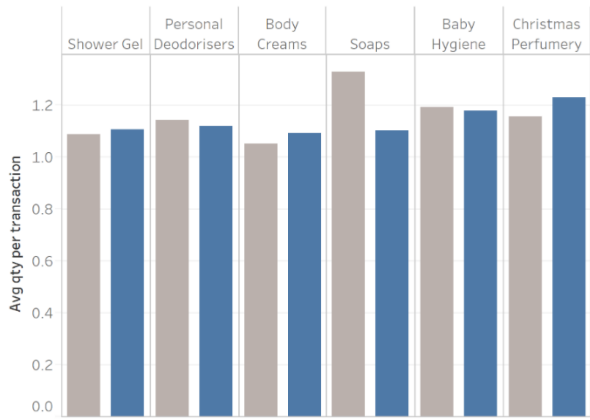
Pricing & Promotion

In general, Dove's products are slightly priced above their competition (both before and after discount) in each segment.

At a product level, most of Dove's products are priced above average apart from the personal deodorizers segment. As the plot shows, almost all categories offer discounts so frequently and in absolute currency value, Dove typically offers higher discounts than its competitors as a function of it being more expensive.



Quantity per Transaction

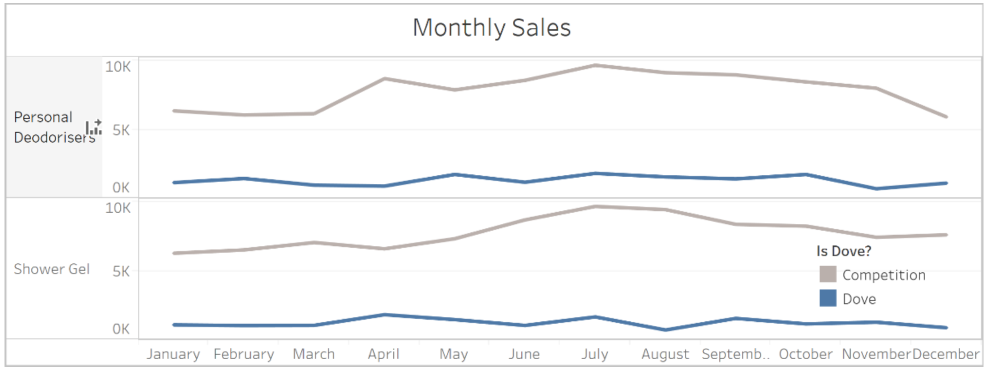


Quantity of each product category purchased per transaction is lower for dove in the soaps category. So, the recommended promotions could be used to increase the quantity purchased in a transaction as well.

2.2.1 Operating Segment

2.2.1 Assumptions & Considerations

Seasonality



Since the marketing promotions are set to take place in April, we had to capture the seasonality effect if any exists. Assuming the data is a good representation of yearly sales patterns, there is a mild effect of seasonality observed in the categories of personal deodorizers and shower gel, with a more relevant difference in the former category. There is a spike in sales (in Vol) for competitors' products while there is a slight dip in Dove's deodorizers. One possible explanation is that since April, peak summers is a popular time for consumers to purchase deodorizers, other brands start promoting around this time. So, this would be a good opportunity to promote Dove's deodorizers.

Private Label

This is a Pernalonga-owned brand which has presence in Dove's operating segment. Despite not targeting consumers of this brand, we still include the brand's transactional data in our analysis as we want to capture all useful purchase patterns. Post analysis, we exclude consumers of this brand from the promotional target list and go with the next best set of consumers.

2.3 Modeling Overview

Following is a brief description of possible consumers' motivations to purchase and the corresponding approaches we considered while building the recommender models.

2.3.1 Brand Preference

Brand preference is when a consumer chooses to purchase a specific brand despite other available options which are closely priced. We believe that the current consumers of Dove have a strong preference for the brand. We hypothesize that if a product of Dove from a different category is recommended to an existing consumer of Dove, there is a chance that consumer will opt to purchase the recommended product. Given that the consumer already has favorable outlook on Dove, this kind of promotion could serve as a recall for the recommended product and an additional incentive to purchase the Dove product from a different category.

For this, we do an item-based collaborative filtering based on product purchases from within Dove's portfolio.

2.3.2 Product Feature Preference

All of Dove's products here are in categories related to personal care. We hypothesize that consumers who are conscious of their needs prefer products with certain characteristics. For instance, a consumer with dry skin would usually purchase products from different brands as long as they meet his skin needs – A consumer could use products from an assortment of brands like Dove soap, Rexona shower gel, Nivea body cream, etc. To this end, we explore similarities among product purchases from the categories within in the personal care operating segment and recommend relevant Dove products to him/her.

For this, we do an item-based collaborative filtering based on product purchases from the operating segment (i.e. the 6 categories Dove has presence in)

2.3.3 Co-purchase among overall categories

Finally, we hypothesize that there are certain categories of products which are purchased together similar to the 'beer & diapers' story. We identify consumers and transactions which show these similarities in purchases at a category level & for all those consumers whose purchase pattern shows a similarity to any of Dove's operating categories, we recommend Dove's brands though a promotion.

For this, we do an item-based collaborative filtering based on overall past purchases.

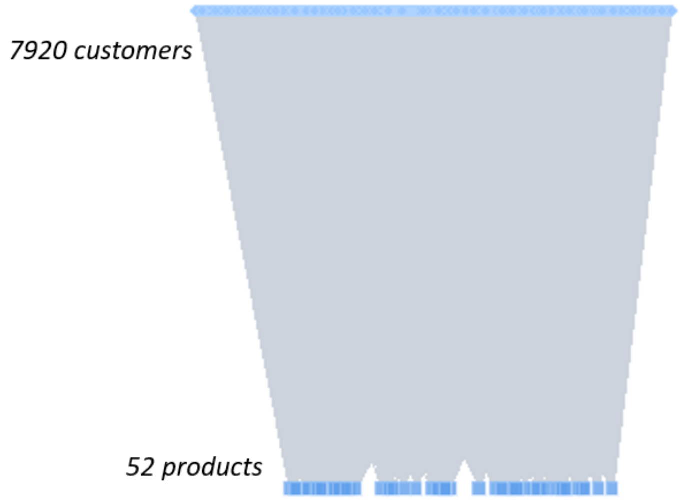
3 different levels of models have been built to recommend products. Whether or not the results of these models converge to an extent informs us about validity of our assumptions, theories and strengthens the recommendations.

Chapter 3: Recommendations & Modeling

3.1 Item-based collaborative filtering within Dove's portfolio

Promote shower gel 999163859 on purchase of 999527671 to consumers from 'frequent customers segment'[1]

Transaction information ***aggregated at a consumer level*** for each consumer who has purchased any of Dove's 52 products was used to create a bipartite network as shown below in graph 1



Graph1 customer-Dove products network viz

[1] Segmentation from project 1: These are the consumers with the highest number of transactions with a reasonably high percent of average discount. since these customers visit stores with a high frequency, they have a higher exposure to the sales being offered in store and so take advantage of that

From the co-purchasing transactional data, we obtained the matrix of products-to-products co-purchased times and their Jaccard similarity.

999233738	999233749	999251576	999333463	999399984	999526011	999186544	999494540	999163869	999177474
0.00000000	0.09448819	0.141472868	0.18270999	0.07665505	0.00000000	0.05297474	0.085585586	0.00000000	0.000000000
0.09448819	0.00000000	0.048037889	0.11380145	0.16829268	0.05310345	0.06534273	0.171166582	0.03894298	0.000000000
0.14147287	0.04803789	0.000000000	0.10467883	0.03806584	0.00000000	0.03278689	0.048192771	0.00000000	0.000000000
0.18270999	0.11380145	0.104678826	0.00000000	0.07317073	0.00000000	0.05680224	0.095603272	0.04612979	0.000000000
0.07665505	0.16829268	0.038065844	0.07317073	0.00000000	0.04223865	0.03981481	0.142207792	0.03808487	0.000000000
0.02846975	0.05310345	0.000000000	0.03549849	0.04223865	0.00000000	0.02657343	0.035525321	0.00000000	0.000000000
0.05297474	0.06534273	0.000000000	0.05680224	0.03981481	0.00000000	0.00000000	0.051675978	0.04160475	0.040247678
0.08558559	0.17116658	0.048192771	0.09560327	0.14220779	0.03552532	0.05167598	0.000000000	0.00000000	0.000000000
0.03981481	0.03894298	0.000000000	0.04612979	0.03808487	0.00000000	0.04160475	0.000000000	0.00000000	0.164351852
0.03696682	0.03460452	0.000000000	0.03971406	0.03467562	0.00000000	0.04024768	0.000000000	0.16435185	0.000000000

Graph3 a hearder of the Jaccard similarity matrix of Dove products

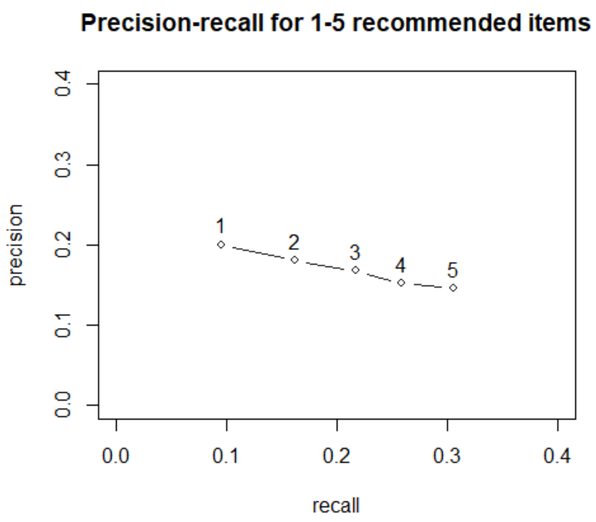
Combined with the Dove's products sales records, we find that Dove's products 999527671, 999399982 and 999233749 are the top selling products from categories deodorizers, soaps and shower gel separately.

Here if a customer has purchased a popular product 999527671 - a deodorizer, there is chance that he/she would likely purchase Dove's products from another category if recommended to him. The Jaccard similarity matrix informs us that the 5 products most likely to be co-purchased with 999527671 include 4 shower gels and 1 deodorizer. Hence, we recommend that it would be good for Dove to promote shower gel 999163859 to customers who have bought deodorizer 999527671.

Product	Category	Jaccard similarity
999163859	*Shower gel*	*0.2235*
999230948	*Shower gel*	*0. 2079*
999233738	*Shower gel*	*0.2028*
999233749	*Dedorisers*	*0.1681*
999246477	*Shower gel*	*0.1342*

Similarly, we also looked at another set of popular products 999399982 and 999233749, along with their Jaccard similarity with other Dove products. And recommendations would be to ***promote shower gel 999527672 to customers who bought Soaps 999399982*** and ***deodorizers 999399984 to customers who bought shower gel 999233749***.

Using this algorithm, we can support a recommendation strategy which recommends Dove's less popular products from one category to a consumer on his/her purchase of one of Dove's popular product from a different category. This strategy tries to leverage on consumer's brand loyalty to encourage product purchase from a different category.



The plot here shows the precision-recall for this recommendation model. Which shows that there is 20% precision when 1 product is recommended, around 18% for 2 & ~16% when 3 products are recommended. Also there is a slightly sharper fall in precision from 3-4 product recommendations. Since this precision is per product, the overall chance of redemption increases when the consumer is exposed to many offers. However, too many options/offers overwhelms the consumers and reduces the efficiency of promotions. So, we suggest that the number of promotions a consumer is exposed should be capped at 3, which brings up the redemption rate to 3*16% from one of the recommended products.

3.2 Item-based collaborative filtering from operating segment

Promote Dove's product 999527671 to customers who's purchase Nivea's 999251780 from the 'skin-care liker segment'[2]

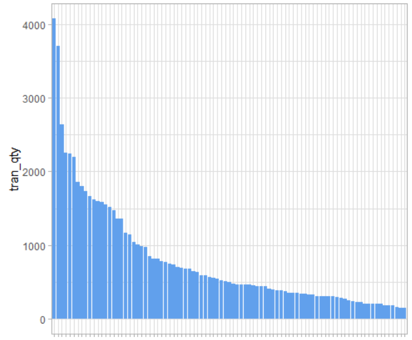
Similarly, we built recommender system in all 6 categories containing products by Dove and other companies as well.

First, we looked at the category 95704(Christmas Perfumery). It's a niche market containing only 4 products, one from Dove and three from Nivea. The similarity matrix suggests that products don't have any apparent co-purchasing relationship, with Jaccard similarity being only around 0.04. So we would not recommend a product in this category based on a customer's buying behavior on same category products from another brand.

Then we have shower gel which contains Dove's most popular product from within its portfolio, contributing to ~12% of Doves sales by revenue. The next 4 of 5 top selling products in this category do not belong to Dove. If there is a high Jaccard similarity between Dove's product 999527671 and popular products from other brands, we could recommend this Dove product to customers who purchased other brand's. The result showed that Nivea's product 999251870 has the highest co-purchasing relationship, with a Jaccard similarity being 0.224.

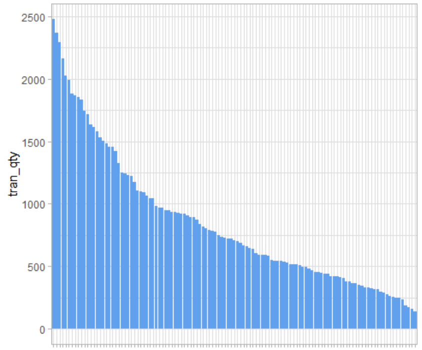
We could promote the most featured Dove's products 999527671 to customers who's buying Nivea's 999251780, to attract potential customers.

prod_id	tran_qty	brand_desc
999527671	4079	DOVE
999251870	3701	NIVEA
999301537	2634	NATURAL HONEY
999226694	2252	PALMOLIVE
999426139	2245	PALMOLIVE



[2] Segmentation from project 1 – Consumers who have high number of transactions involving skin and personal care related products. Consists of consumers who are aware of product features they need or those who try a different assortment of brands. Either way, would serve as good category for current recommendation

Similarly, we looked at deodorizer which also accounts for a big part of Dove's sales. Doves' products sell well in this category, but we still can find a way to increase its market share. We analyzed the Jaccard similarity for the top Rexona product and found that one of Dove's less popular product 999173739 is often purchased with the Rexona product. So, we could promote this product to people who bought product 999152445 and gain customers from other brands.



prod_id	tran_qty	brand_desc
999152445	445	REXONA
999159419	416	DOVE
999160347	421	DOVE
999160407	316	SANEX
999163869	454	DOVE

3.3 item-based collaborative filtering based on overall past purchases

Promote Dove's products to consumers purchasing products from categories that are frequently co-purchased with Dove's categories.

Here we use Jaccard Similarity to identify customers buying which categories are likely to buy the categories that Dove operates in. The network is built on category_id and the customer_id. Higher the number of transactions for a particular customer that involve two categories, higher is the similarity.

For shower gel, categories like hair conditioner, oral elixirs, combing products, make up, face treatment & sunscreen are the most similar as can be seen from the table below:

category_id	95746	category_desc_eng
95730	0.7841220	HAIR CONDITIONERS
95740	0.5600316	ORAL ELIXIRS
95729	0.5263573	PRODUCTS FOR COMBING
95707	0.4503040	MAKE UP
95780	0.4419955	BRIGHTENING DISHWASHING MACHINE DETERGENTS
95710	0.4288176	FACE TREATMENT
95703	0.4235854	SOLAR
95709	0.4133722	FACE CLEANING
95785	0.4088325	MACHINE ADDITIVES CLOTHING
95385	0.3887712	SANDWICHES
95405	0.3871181	UHT SOY DRINKS
95723	0.3819315	PRODUCTS FOR HAIR REMOVAL APPLICATION
95052	0.3756117	LACTOSE UHT MILK
95718	0.3549583	INCONTINENTS
95722	0.3348434	SHEETS AND WAXING MACHINES
95775	0.3247149	DECONTAMINATORS
95784	0.3184350	MACHINE WASH DESCALER
95727	0.3145933	COLORING
95733	0.3131985	BABY SHAMPOO AND CONDITIONER
95736	0.3065032	TOPS
95433	0.3036712	DERMATOLOGICAL TREATMENT
95735	0.3035881	INTIMATE CARE PRODUCTS
95686	0.3027804	KITCHEN PLASTIC ARTICLES

For personal deodorizers, categories like hair conditioners, air fresheners, packaged gums, combing products, make up, sunscreen & face treatment are most similar and relevant as shown below

category_id	95747	category_desc_eng
95730	0.7849462	HAIR CONDITIONERS
95758	0.7175998	AIR FRESHENER N / ELECT
95087	0.6767309	PACKAGED GUMS
95676	0.5400601	PLASTIC DISPOSABLE ITEMS
95729	0.5278068	PRODUCTS FOR COMBING
95873	0.5215743	FROZEN PIZZAS
95872	0.4997384	FROZEN MEALS
95688	0.4695504	CLOTHING ACCESSORIES
95707	0.4517187	MAKE UP
95703	0.4281957	SOLAR
95710	0.4277646	FACE TREATMENT
95709	0.4154373	FACE CLEANING
95692	0.4103236	TEFLON KITCHEN ITEMS
95593	0.4003150	ACCESSORIES FOR ANIMALS
95723	0.3858631	PRODUCTS FOR HAIR REMOVAL APPLICATION
95475	0.3623607	CANNED SEA SPECIALTIES
95808	0.3575095	CONCENTRATED DRINKS
95570	0.3437993	FOOD PREPARATION
95722	0.3353586	SHEETS AND WAXING MACHINES
95775	0.3258383	DECONTAMINATORS
95727	0.3156299	COLORING
95733	0.3130355	BABY SHAMPOO AND CONDITIONER
95736	0.3069202	TOPS
95735	0.3036871	INTIMATE CARE PRODUCTS
95686	0.3023715	KITCHEN PLASTIC ARTICLES
95757	0.2955776	AIR FRESHENERS ELECT
95713	0.2951855	PHARMACY COTTON
95792	0.2942107	BEVERAGE MIXES
95799	0.2726554	BEER WITH ALCOHOL
95802	0.2498018	SPORTS DRINKS
95726	0.2466385	AFTER SHAVE

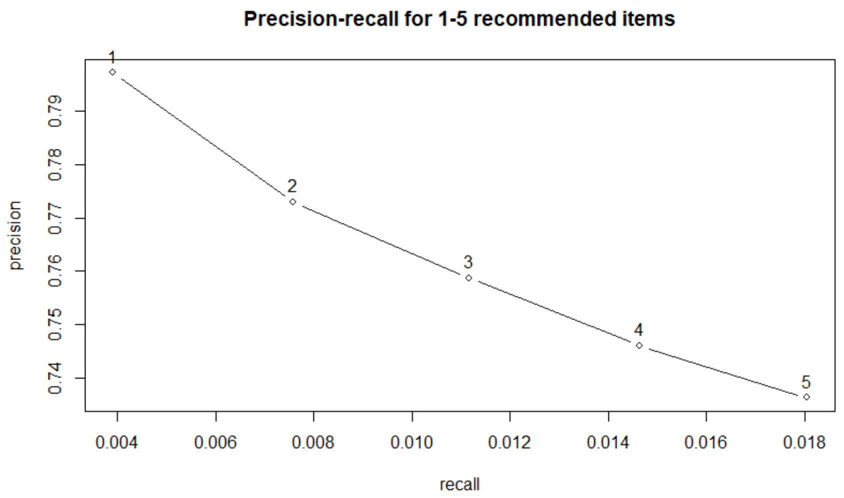
For Baby Hygiene, categories like Diaper, confectionary, Candy, lolliop, party articles share high similarity and are relevant as shown below:

category_id	95732	category_desc_eng
95087	0.6762409	PACKAGED GUMS
95993	0.5936213	CONFECTIONERY
95740	0.5569133	ORAL ELIXIRS
95676	0.5400906	PLASTIC DISPOSABLE ITEMS
95674	0.5121236	PARTY ARTICLES
95705	0.4905029	BEAUTY TREATMENT COTTON
95707	0.4500602	MAKE UP
95088	0.4485569	DROPS/CANDY
95703	0.4239829	SOLAR
95709	0.4102358	FACE CLEANING
95385	0.3881817	SANDWICHES
95723	0.3831160	PRODUCTS FOR HAIR REMOVAL APPLICATION
95828	0.3760397	CHILLED JUICES
95090	0.3738028	LOLLIPOP
95720	0.3568420	DIAPERS
95718	0.3552367	INCONTINENTS
95384	0.3366935	MENUS
95722	0.3363059	SHEETS AND WAXING MACHINES
95775	0.3242481	DECONTAMINATORS
95733	0.3220247	BABY SHAMPOO AND CONDITIONER
95784	0.3178888	MACHINE WASH DESCALER
95727	0.3125840	COLORING
95736	0.3114665	TOPS
95433	0.3071622	DERMATOLOGICAL TREATMENT
95735	0.3037307	INTIMATE CARE PRODUCTS
95832	0.3014994	BARRIER CREAMS

In general, the Face cleaning, face treatment and solar have high similarity to most of the categories that Dove carries.

Although Soap is a major category for DOVE, there are not many other brands carry the same category. As a result, the customer purchasing information related to soaps is very limited, thus the Jaccard similarity is low when comparing with other categories.

The following plot shows the precision – recall according to the number of items recommended to the consumer. Even a single prediction has a high precision as there are based on category co-occurrences. Since there are other factors to consider here like price, discount amount and personal preferences, the actual redemption for this would be less than what the plot suggests. A good estimate to take would be the expected redemption rate of the category multiplied by the market share of Dove in respective category.



Chapter 4: Conclusion

Another approach we considered for recommendation is to work with the consumer segmentation from the earlier project and build a separate recommender system based on data from each segment. However, focusing only on one section of consumers takes away valuable information about overall behavior related to transactions of this category. So, we built the recommender model on overall data and applied it to relevant consumer segments.

Since we are basing our recommendations on co-purchases and not ratings of these products, Jaccard similarity – a binary co-purchasing matrix serves as a good distance/similarity measure.