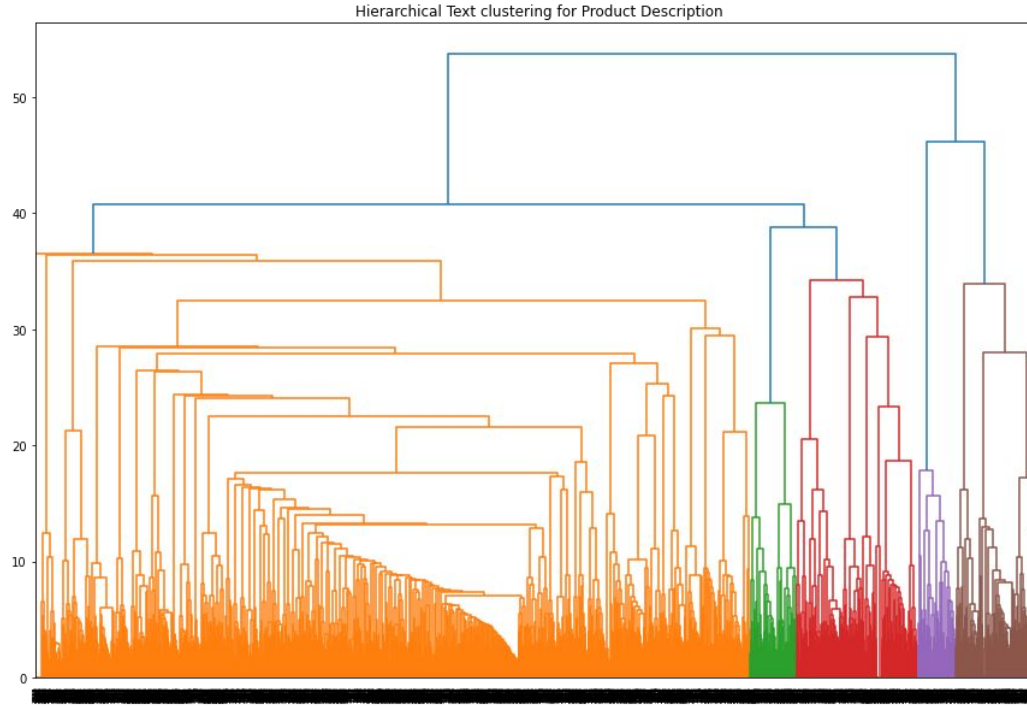


E-Commerce: Product Range Analysis

Gome Steinkoler

Product Description Analysis

Five clusters are identified in the hierarchical text clustering of the Product Description.

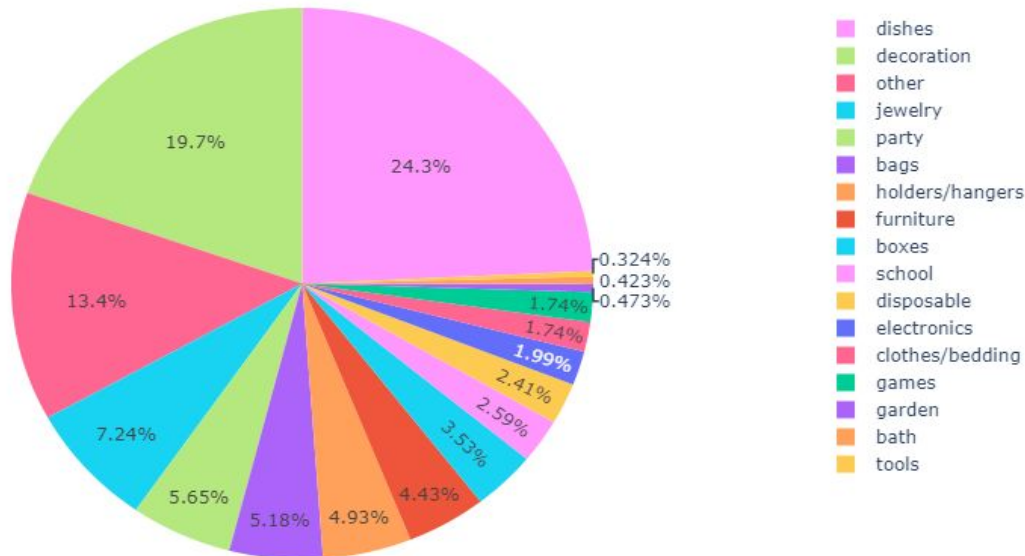


The identified text cluster keywords:

- cluster 0 keywords:
 - glass candle card bag flower necklace christmas
- cluster 1 keywords:
 - vintage design wrap bag christmas mug gift
- cluster 2 keywords:
 - box key ring bling letter trinket gift
- cluster 3 keywords:
 - heart decoration hanging holder love zinc bell
- cluster 4 keywords:
 - decoration egg easter hanging chick christmas rabbit

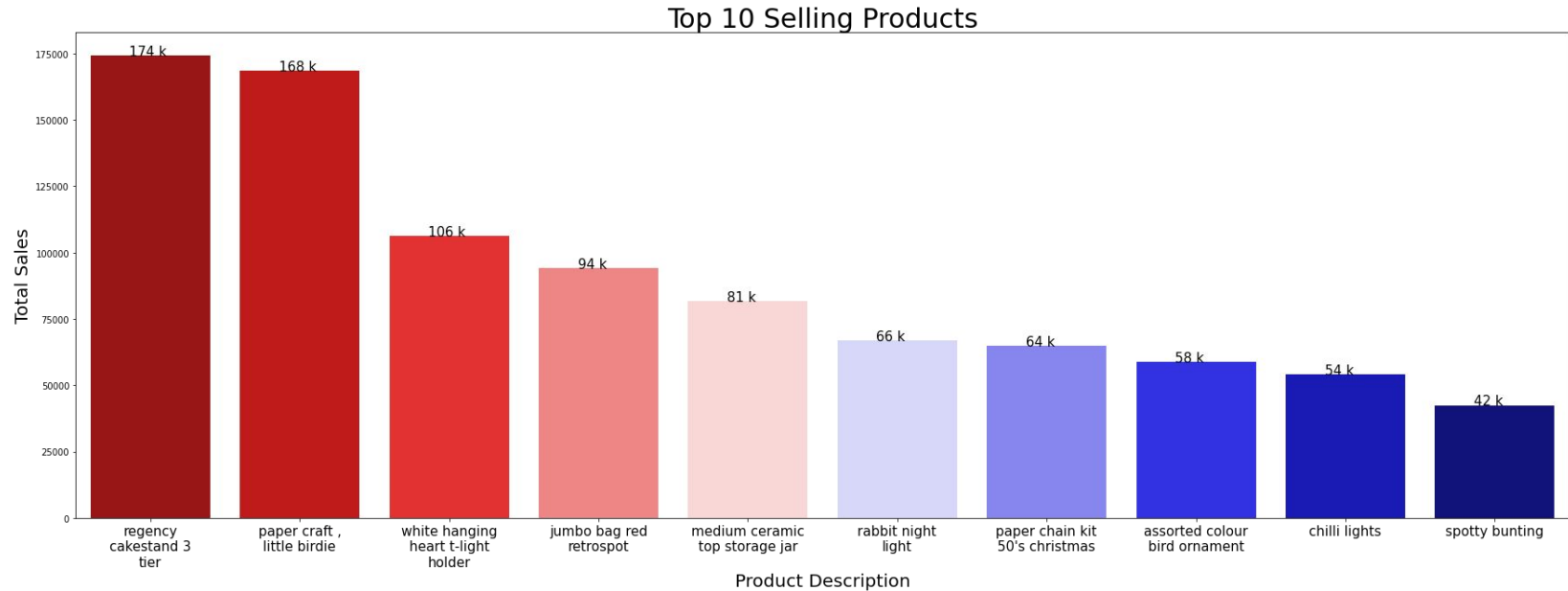
The leading category in number of items is dishes!

Split by categories



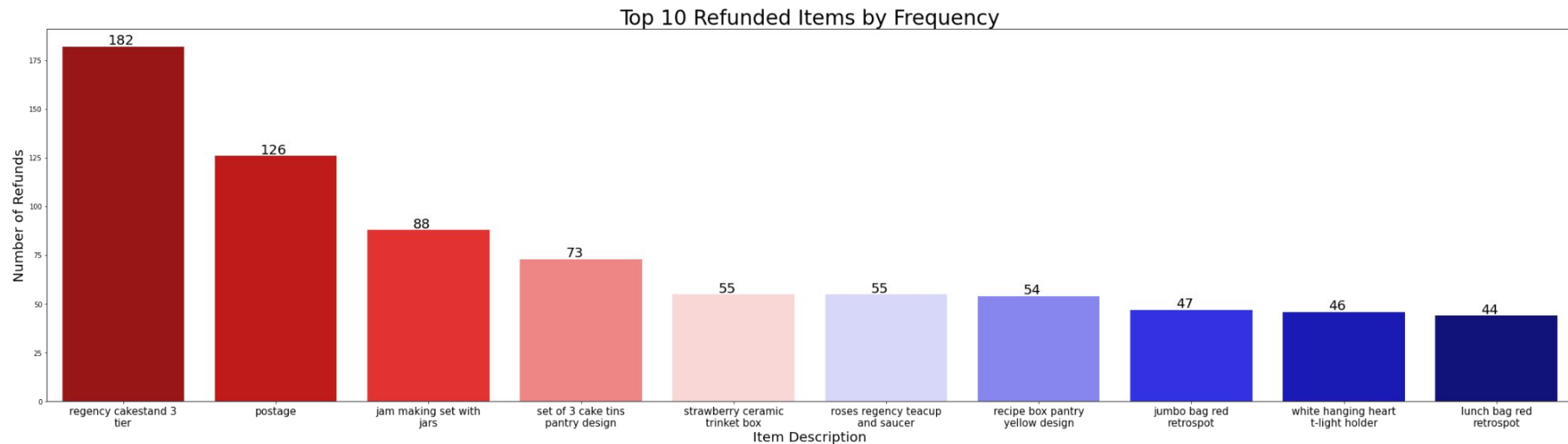
- The leading product category in number of items is dishes (24.3%) with decoration coming second (19.7%).
- The category with the smallest range of product is tools with 13 items only!

The top selling product is cake stand!

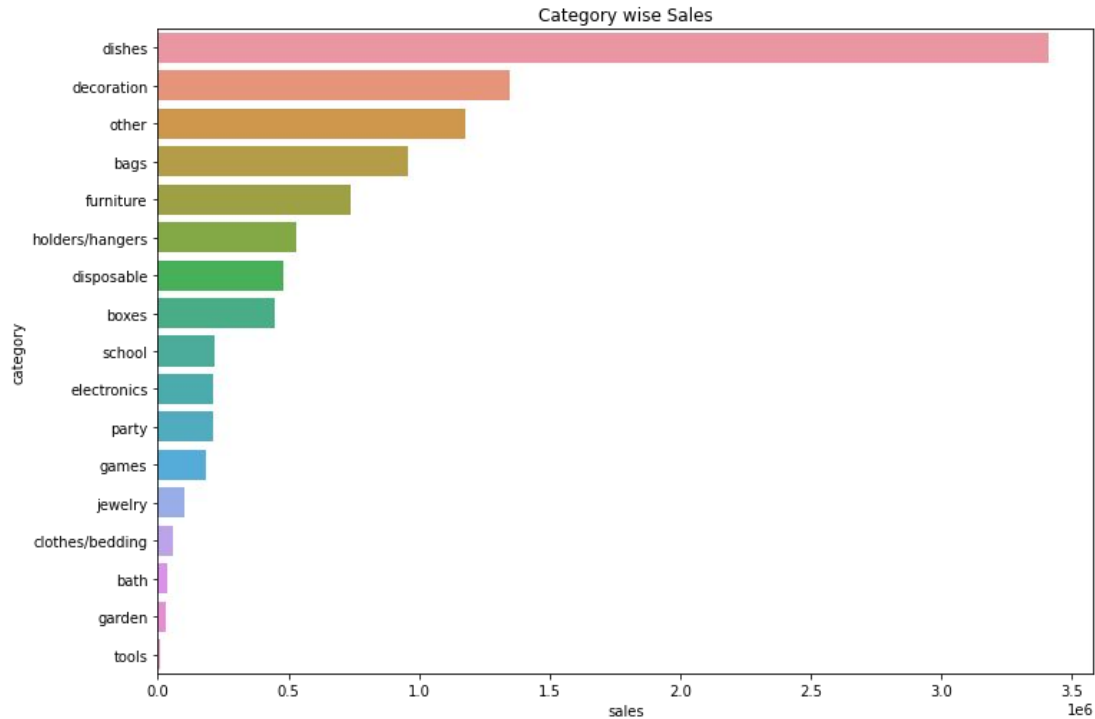


- The leading product in total sales is cake stand with paper little birdie coming second in sales.

The most frequently refunded product is also cake stand!



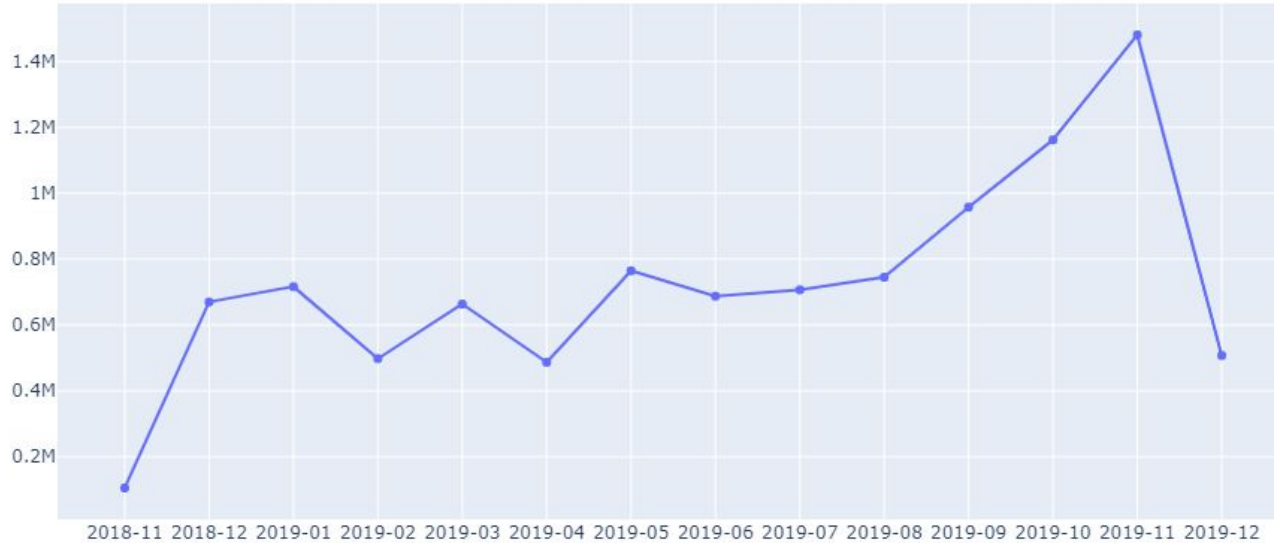
The leading category in sales is also dishes!



- The leading product category sales is dishes with decoration coming second.

The strongest month is November!

Montly Revenue



Monthly revenue generally increases until a peak at November 2019 followed by a sharp decrease at December 2019.

Basket Analysis

Basket Analysis: The highest confidence (77.4%) of the baskets goes to: 'rose regency teacup saucer' and 'regency teacup saucer'.

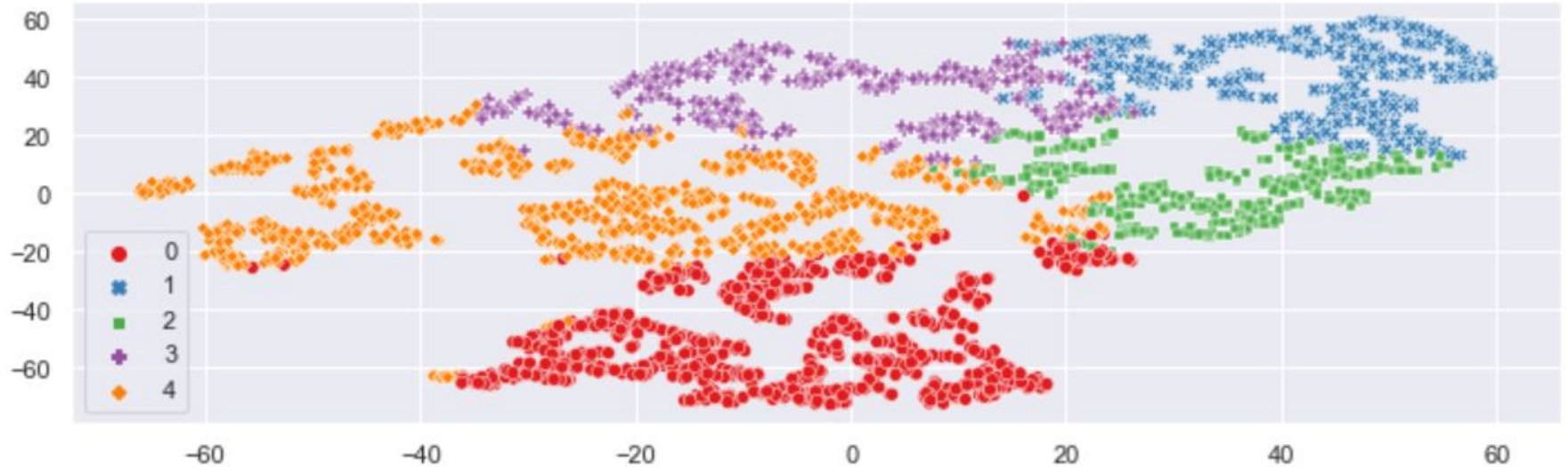
	antecedents	consequents	antecedent support	consequent support	support	confidence	lift	leverage	conviction
14	(rose regency teacup saucer)	(regency teacup saucer)	0.053893	0.058038	0.041709	0.773921	13.334635	0.038581	4.166519
15	(regency teacup saucer)	(rose regency teacup saucer)	0.058038	0.053893	0.041709	0.718641	13.334635	0.038581	3.362635
1	(bag polkadot)	(bag retrospot)	0.061577	0.105763	0.041709	0.677340	6.404294	0.035196	2.771451
5	(storage bag suki)	(bag retrospot)	0.059858	0.105763	0.036603	0.611486	5.781646	0.030272	2.301687
3	(shopper vintage paisley)	(bag retrospot)	0.059403	0.105763	0.034378	0.578723	5.471869	0.028095	2.122683
8	(lunch bag polkadot)	(lunch bag retrospot)	0.055106	0.079070	0.030637	0.555963	7.031301	0.026280	2.073996
13	(lunch bag suki design)	(lunch bag retrospot)	0.064965	0.079070	0.033064	0.508949	6.436713	0.027927	1.875428
10	(lunch bag skull)	(lunch bag retrospot)	0.064358	0.079070	0.032406	0.503535	6.368236	0.027318	1.854975
12	(lunch bag retrospot)	(lunch bag suki design)	0.079070	0.064965	0.033064	0.418159	6.436713	0.027927	1.607028
11	(lunch bag retrospot)	(lunch bag skull)	0.079070	0.064358	0.032406	0.409847	6.368236	0.027318	1.585422

- The highest confidence 77.4% of the baskets goes to: 'rose regency teacup saucer' and 'regency teacup saucer'. Their total support is 4.1% , antecedent support (of the first) is 5.3% ,consequent support (of the last) is 5.8%. This pair has the highest lift (13.3%).
- Another pair with pretty high confidence of 67.7% of the baskets is: 'bag polkadot' and 'storage bag suki'. Their total support is 4.1% , antecedent support (of the first) is 6.1% ,consequent support (of the last) is 10.5%. This pair also has the high lift (6.4%) .

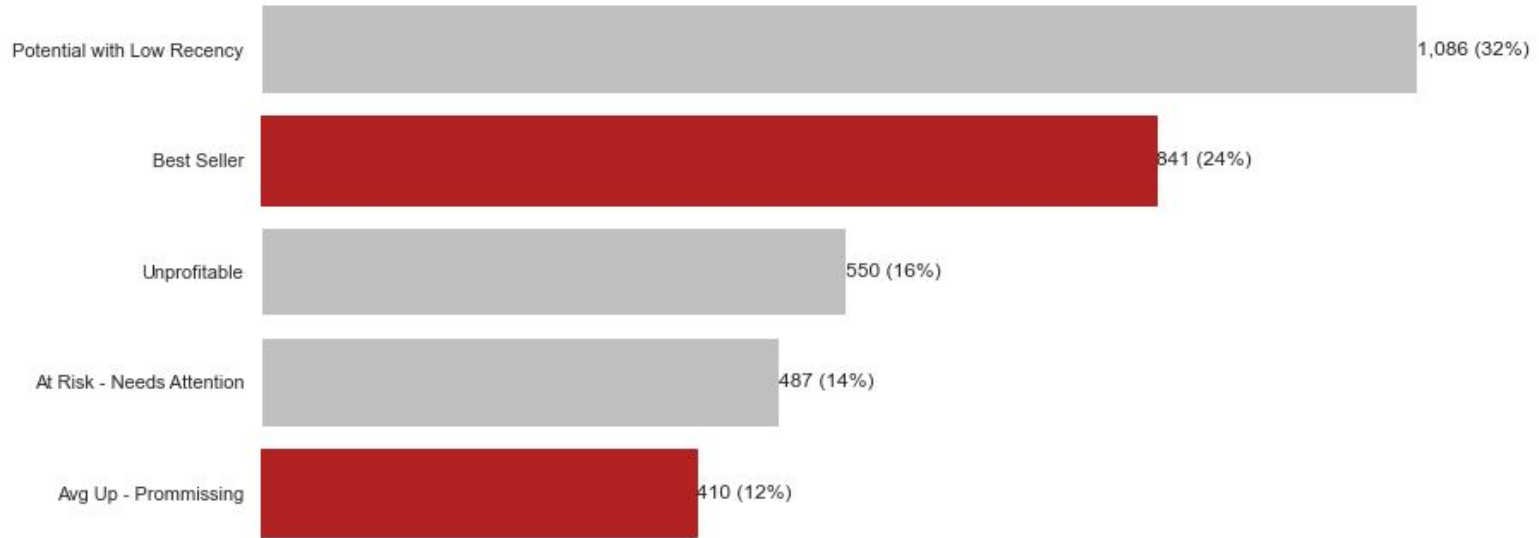
RFM Analysis

The KMEANS of the RFM (log-scaled and fitted) produces 5 well distinguished clusters

Flattened Graph of 5 Clusters



The five sales clusters: 36% are best sellers or Promising products!



- There are a lot of products which are not bought frequently or recently and do not generate a good profit.
- However, 36% of the items are best sellers or promising (red bars) meaning that they are bought frequently, generate good profit and their recency is also high.
- 32% of the products have a potential and with an appropriate strategy will become even more profitable!

- Please see my [dashboard](#) supporting this RFM analysis with more details and interactive tools available.

Summary

- Product Description Analysis
 - The number of orders increases from August 2019 until November 2019 which has the best sales considering units, range of product type and revenue. It is suggested to further investigate why November 2018 is the worst month.
 - The leading product in total sales is cake-stand with paper little birdie coming second and hanging heart t-light holder coming third in sales.
- Basket Analysis
 - The basket pair with the highest lift (13.3%) is 'rose regency teacup saucer' and 'regency teacup saucer'. Another strong basket pair is: 'bag polkadot' and 'storage bag suki'.
 - The online store can use a similar basket tool to find the top 10 products that resemble items which are missing or to imply marketing strategy on them. It is suggested to also investigate the products with the lowest support or the ones we consider unpopular for other reasons.
- RFM
 - The RFM analysis identified 36% of the products as best sellers or promising meaning that they are bought frequently, generate great profit and their recency is also high.
 - Another 32% of the products have a good potential and and it is suggested to find a strategy to make them more profitable.
 - There are also a lot of products (30%) which are not bought frequently or recently and do not generate a good profit, these perhaps should removed from the store or require a drastic change in marketing or pricing.