
KMart NLP Team Presentation

Presented by
Nachammai Palaniappan
Komali Reddy
Uwais Mansuri

Saint Peter's University

Problem Statement

- The purpose of our project is to understand the importance of what people say through reviews about Kmart products, in order to make better decisions and improve customer experience.
- To achieve this, we have analyzed reviews by using text analysis, sentiment analysis topic modeling, and aspect based analysis.

About Kmart



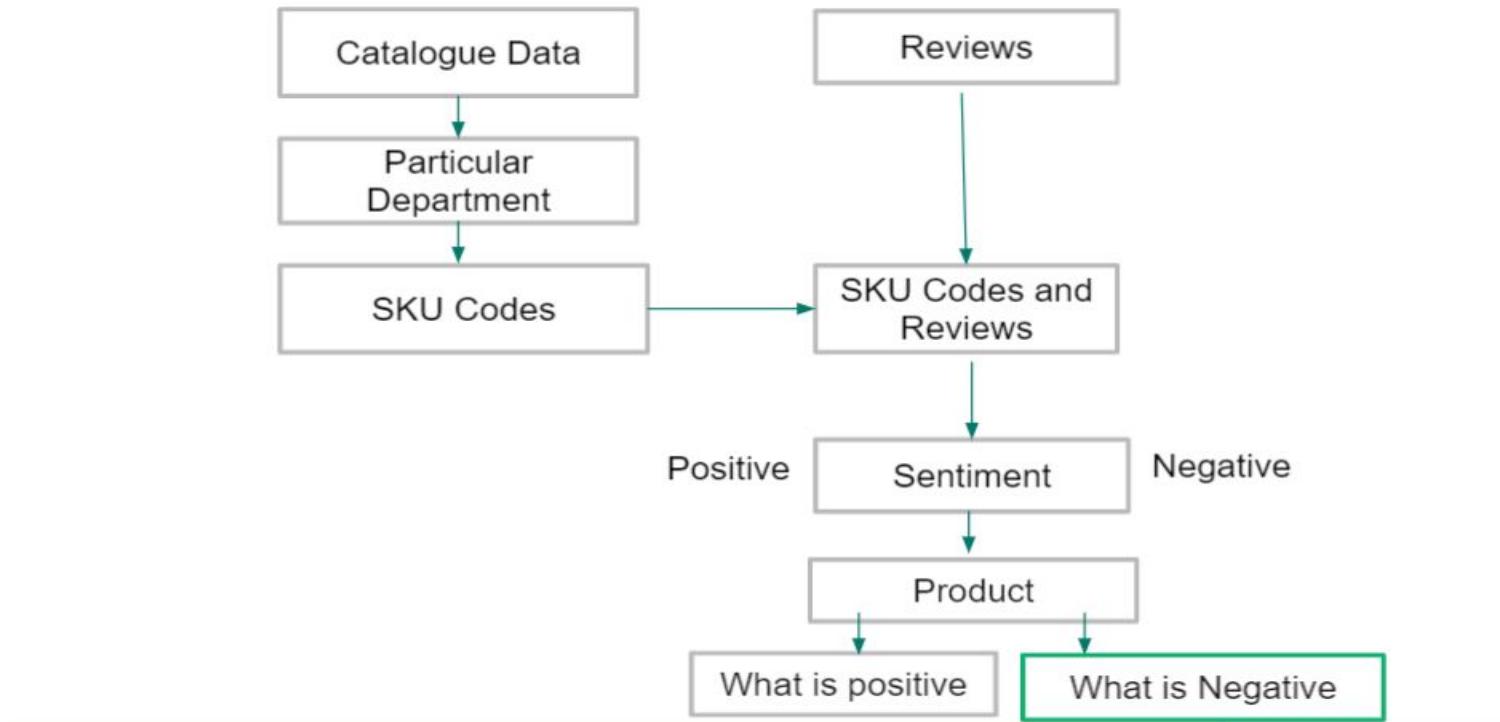
- Kmart Australia Limited is an Australian-based chain of affordable retail stores owned by the Kmart Group division of Wesfarmers.
- The company operates 234 stores across Australia and New Zealand, consisting of 209 stores in Australia and 25 stores in New Zealand, with its head office located in Mulgrave, Melbourne. Kmart is a nationwide retailer of electronics, toys, clothing, bedding, furniture & home decor. Buy online, pick up in the store.

Goal

Compared the KMart reviews for the years 2019 and 2020. Distribution of reviews for the year 2019 and 2020. Determined the two highly reviewed products in Kmart. Product Rating Analysis was done. Collaborative filtering was used for recommendation systems. Sentiment analysis, Aspect Model was done. Finally deployed the output in Flask.



Workflow



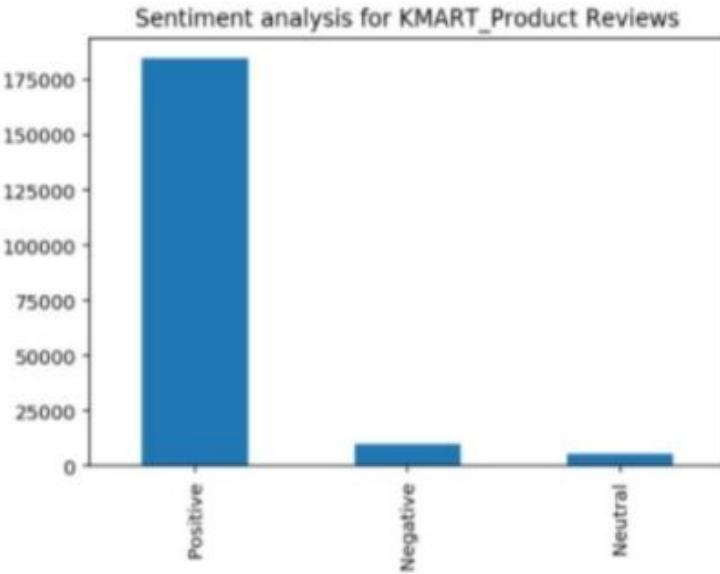
About the Dataset

There is a Catalogue data provided in which they do have Department and sub sub class. Grouped the data into department wise and sub sub class. Time series analysis for 2019 and 2020 data. Determined lots of 1 star reviews for the month 3,4,5,6 of 2020.

Tools and Methods

1. Python
2. IBM NLU
3. IBM Discovery
4. Excel
5. Power BI
6. Tableau

Sentiment analysis for KMAR T product reviews

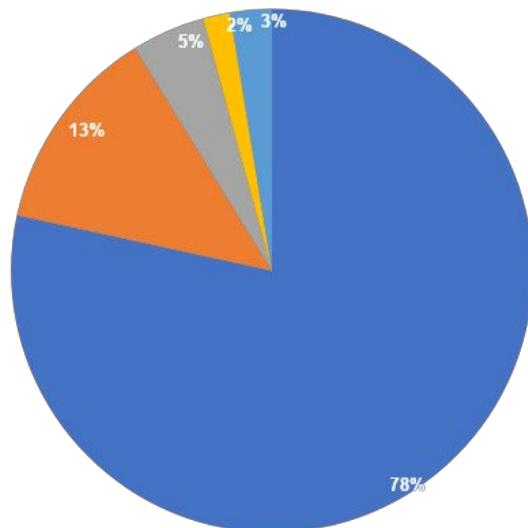


```
1 print("Number of Positive reviews: {}, which is: {}".format(  
2 print("Number of Negative reviews: {}, which is: {}".format(  
3 print("Number of Neutral reviews: {}, which is: {}".format(  
4
```

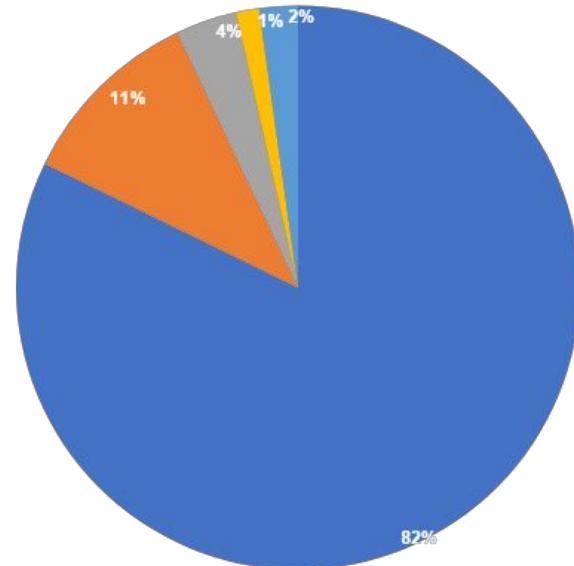
Number of Positive reviews: 184728, which is: 92.20448623880687
Number of Negative reviews: 9943, which is: 4.962914158505785
Number of Neutral reviews: 5675, which is: 2.832599602687351

Python (2019 and 2020 same data)

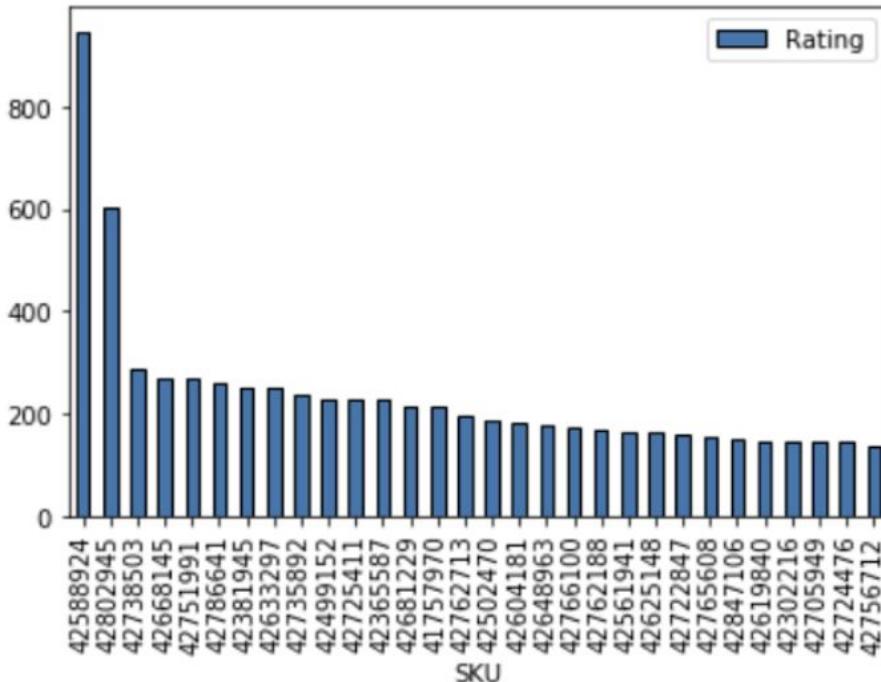
distribution of reviews 2019



distribution of reviews 2020



Products with highest number of Ratings



The products with SKU Code 42588924 has the highest number of ratings followed by 42802945 and other SKU codes.

Let us have a look on the highly reviewed products (our two outliers) from kmart website



3.2 Litre Air Fryer

\$69.00



5.3 Litre Air Fryer

\$89.00

Grouping the data

File Home Insert Page Layout Formulas Data Review View Help

Cut Copy Format Painter

Paste

Font

Alignment

Number

General

Conditional Formatting

S29

	A	B	C	D	E	F	G	H	I	J
1	SKU Code	Department	Sub Sub Class							
2	39333018	KIDS INDOOR PLAY	DRAWING/CASES							
3	39333018	KIDS INDOOR PLAY	DRAWING/CASES							
4	39333018	KIDS INDOOR PLAY	DRAWING/CASES							
5	39333018	KIDS INDOOR PLAY	DRAWING/CASES							
6	42163053	KIDS INDOOR PLAY	DRAWING/CASES							
7	42163053	KIDS INDOOR PLAY	DRAWING/CASES							
8	42163053	KIDS INDOOR PLAY	DRAWING/CASES							
9	42163053	KIDS INDOOR PLAY	DRAWING/CASES							
10	42163053	KIDS INDOOR PLAY	DRAWING/CASES							
11	42163053	KIDS INDOOR PLAY	DRAWING/CASES							
12	42163053	KIDS INDOOR PLAY	DRAWING/CASES							
13	42163053	KIDS INDOOR PLAY	DRAWING/CASES							
14	42163053	KIDS INDOOR PLAY	DRAWING/CASES							
15	42163053	KIDS INDOOR PLAY	DRAWING/CASES							
16	42163053	KIDS INDOOR PLAY	DRAWING/CASES							
17	42163053	KIDS INDOOR PLAY	DRAWING/CASES							
18	42163053	KIDS INDOOR PLAY	DRAWING/CASES							
19	42163053	KIDS INDOOR PLAY	DRAWING/CASES							
20	42163053	KIDS INDOOR PLAY	DRAWING/CASES							
21	41882917	KIDS INDOOR PLAY	DRAWING/CASES							
22	41882917	KIDS INDOOR PLAY	DRAWING/CASES							
23	41882917	KIDS INDOOR PLAY	DRAWING/CASES							
24	41882917	KIDS INDOOR PLAY	DRAWING/CASES							
25	41882917	KIDS INDOOR PLAY	DRAWING/CASES							
26	40891675	KIDS INDOOR PLAY	ACCESSORIES							
27	40891675	KIDS INDOOR PLAY	ACCESSORIES							
28	41882900	KIDS INDOOR PLAY	PADS/SCRAPBOOKS							
29	41882900	KIDS INDOOR PLAY	FADS/SCRAPBOOKS							

Product Hierarchy MASSAGE ELECTRONIC HANGING DECORATION GIRL FASHION EQUIPMENT DECOR

Sub sub classes

```
graph TD; SS[Sub sub classes] --> D[DRAWING/CASES]; SS --> A[ACCESSORIES]; SS --> P[PADS/SCRAPBOOKS]; SS --> F[FADS/SCRAPBOOKS]; D --> M[MASSAGE]; E[ELECTRONIC]; H[HANGING DECORATION]; D --> G[GIRL]; E --> F[FASHION]; H --> E; D --> D[DECOR];
```

Sample sub sub class

File Home Insert Page Layout Formulas Data Review View Help

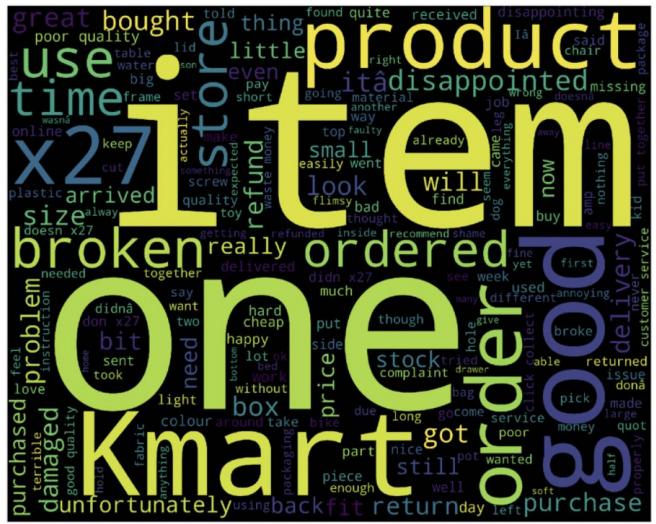
Cut Copy Format Painter

Font Alignment Number

Conditional Formatting

G16	A	B	C	D	E	F	G	H	I
1	SKU Code	Department	Sub Sub Class						
2	42833857	PET CARE	ELECTRONIC						
3	42833840	PET CARE	ELECTRONIC						
4	42833840	PET CARE	ELECTRONIC						
5	42833840	PET CARE	ELECTRONIC						
6	42833840	PET CARE	ELECTRONIC						
7	42833840	PET CARE	ELECTRONIC						
8	42833840	PET CARE	ELECTRONIC						
9	42833840	PET CARE	ELECTRONIC						
10	42833840	PET CARE	ELECTRONIC						
11	42833840	PET CARE	ELECTRONIC						
12	42833840	PET CARE	ELECTRONIC						
13	42833840	PET CARE	ELECTRONIC						
14	42833840	PET CARE	ELECTRONIC						
15	42833840	PET CARE	ELECTRONIC						
16	42833840	PET CARE	ELECTRONIC						
17	42833840	PET CARE	ELECTRONIC						
18	42833840	PET CARE	ELECTRONIC						
19	42833840	PET CARE	ELECTRONIC						
20	42833840	PET CARE	ELECTRONIC						
21	42833840	PET CARE	ELECTRONIC						
22	42833840	PET CARE	ELECTRONIC						
23	42833840	PET CARE	ELECTRONIC						
24	42833840	PET CARE	ELECTRONIC						
25	42833840	PET CARE	ELECTRONIC						
26	42833840	PET CARE	ELECTRONIC						
27	42833840	PET CARE	ELECTRONIC						
28	42719502	PET CARE	ELECTRONIC						
29	42719502	PET CARE	ELECTRONIC						

Negative words



Positive words



Word Cloud image for K-Mart Product reviews 2019/ 2020

	C	D	E	F	G	H	N	O	P	Q	R	S
1	Review Tit	Review Co	Rating	Product Title	Product URL	SKU Code	sku row number		SKU Code	Department	Sub Sub Class	
2	I hated it.	I hated it.	1	Jute Runner	https://www.kmart.com.au/product/jute-runner	42570677	38017		39333018	KIDS INDOOR PLAY	DRAWING/CASES	
3	FIRE HAZA	As soon as	1	3 in 1 Accessory Pack	https://www.kmart.com.au/product/3-in-1-accessory-pack	42551249	82022		39333018	KIDS INDOOR PLAY	DRAWING/CASES	
4	Good.	Good fit.	1	5 Pack Work Socks-Blk/gry,6-	https://www.kmart.com.au/product/5-pac	68635367	139039		39333018	KIDS INDOOR PLAY	DRAWING/CASES	
5	Used twice	The handh	1	2 in 1 Cordless Stick Vacuum	https://www.kmart.com.au/product/2-in-1-cordless-stick-vacuum	42735892	265668		39333018	KIDS INDOOR PLAY	DRAWING/CASES	
6	Product go	Unable to	1	Skate Helmet - Medium Blue	https://www.kmart.com.au/product/skate-helmet-medium-blue	42642329	156420		42163053	KIDS INDOOR PLAY	DRAWING/CASES	
7	Nice top	Bought th	1	Extended Sleeve Top-Blue Pe	https://www.kmart.com.au/product/extended-sleeve-top-blue-pe	69510304	194414		42163053	KIDS INDOOR PLAY	DRAWING/CASES	
8	Broken	This clock	1	Metro Wall Clock	https://www.kmart.com.au/product/metro-wall-clock	41337158	37347		42163053	KIDS INDOOR PLAY	DRAWING/CASES	
9	busted bin	Received t	1	30L Brushed Stainless Steel B	https://www.kmart.com.au/product/30l-brushed-stainless-steel-bin	42747826	40144		42163053	KIDS INDOOR PLAY	DRAWING/CASES	
10	Products v	Products v	1	White Desk Drawers	https://www.kmart.com.au/product/white-desk-drawers	42698494	24418		42163053	KIDS INDOOR PLAY	DRAWING/CASES	
11	This air be	This air be	1	Queen Air Bed with Built-in P	https://www.kmart.com.au/product/queen-air-bed-with-built-in-pump	42666394	218981		42163053	KIDS INDOOR PLAY	DRAWING/CASES	
12	Not compl	3 glazed bo	1	3 Glazed Pot Plants	https://www.kmart.com.au/product/3-glazed-pot-plants	42643944	174358		42163053	KIDS INDOOR PLAY	DRAWING/CASES	
13	Disappoint	Ordered sh	1	Tan Ottoman	https://www.kmart.com.au/product/tan-ottoman	42760351	133482		42163053	KIDS INDOOR PLAY	DRAWING/CASES	
14	Oversized	Very overs	1	Plus Size Long Sleeve Texture	https://www.kmart.com.au/product/plus-size-long-sleeve-texture	69379918	257608		42163053	KIDS INDOOR PLAY	DRAWING/CASES	
15	Shelf	Awsome s	1	4 Tier Shelf with Bamboo Sh	https://www.kmart.com.au/product/4-tier-shelf-with-bamboo-shelves	42187585	211084		42163053	KIDS INDOOR PLAY	DRAWING/CASES	
16	Cheap and Horrible	st	1	Emma Wiggles Costume - To	https://www.kmart.com.au/product/emma-wiggles-costume-toy-story	42598497	41693		42163053	KIDS INDOOR PLAY	DRAWING/CASES	
17	This vacuu	This vacuu	1	1200W Upright Vacuum	https://www.kmart.com.au/product/1200w-upright-vacuum	42701668	221266		42163053	KIDS INDOOR PLAY	DRAWING/CASES	
18	Absolutely	My tablec	1	Coffee Table Oak Look	https://www.kmart.com.au/product/coffee-table-oak-look	42619697	31370		42163053	KIDS INDOOR PLAY	DRAWING/CASES	
19	Flimsy &ar	top was ca	1	Walnut Look Desk	https://www.kmart.com.au/product/walnut-look-desk	42715696	144175		42163053	KIDS INDOOR PLAY	DRAWING/CASES	
20	Item not d	Item not J	1	Fish Bowl Vase	https://www.kmart.com.au/product/fish-bowl-vase	40345437	161431		42163053	KIDS INDOOR PLAY	DRAWING/CASES	
21	Was appal	Was appal	1	Vacuum Food Sealer	https://www.kmart.com.au/product/vacuum-food-sealer	42584896	244868		41882917	KIDS INDOOR PLAY	DRAWING/CASES	
22	I had a ten	I had a ten	1	Large White Embossed Canis	https://www.kmart.com.au/product/large-white-embossed-canis	42576617	171432		41882917	KIDS INDOOR PLAY	DRAWING/CASES	
23	Dissapoint	Does not s	1	VTech Lullaby Lambs Mobile	https://www.kmart.com.au/product/vtech-lullaby-lambs-mobile	42638698	229694		41882917	KIDS INDOOR PLAY	DRAWING/CASES	
24	Came dam	Came dam	1	Industrial 4 Tier Bookshelf	https://www.kmart.com.au/product/industrial-4-tier-bookshelf	42758358	117804		41882917	KIDS INDOOR PLAY	DRAWING/CASES	
25	Delivered i	One out of	1	Hanging Sloth Pot	https://www.kmart.com.au/product/hanging-sloth-pot	42754183	104782		41882917	KIDS INDOOR PLAY	DRAWING/CASES	
26	Great bed	Great bed	1	Ball Fibre Mattress Topper - I	https://www.kmart.com.au/product/ball-fibre-mattress-topper	42127369	260701		40891675	KIDS INDOOR PLAY	ACCESSORIES	
27	Love it	Love it	1	5 Piece Cookware Set	https://www.kmart.com.au/product/5-piece-cookware-set	42689140	209040		40891675	KIDS INDOOR PLAY	ACCESSORIES	

Full list of non-reviewed products

The screenshot shows a Microsoft Excel spreadsheet titled "G1003". The table consists of 20 rows and 7 columns. The columns are labeled A through G. Column A contains row numbers from 1 to 20. Column B contains SKU codes, and Column C contains product descriptions. Column D contains a boolean value (TRUE or FALSE) indicating if the product is not reviewed. Column E is empty. Column F contains the formula =SUM(D:D), which sums up all the TRUE values in column D, resulting in the value 16274. Column G is also empty.

	A	B	C	D	E	F	G
1	SKU Code catalogue	Department		SKU's not reviewed full list			
665	41732564	RECREATION & BIKES	TRUE	69376061			
855	42712565	STATIONERY & CRAFT	TRUE	42757696			
944	42223238	ACTION/VEHICLE TOY	TRUE	69602368			
965	42699637	KIDS INDOOR PLAY	TRUE	42774402			
1003	68140120	KIDS U/WEAR SOCKS	TRUE	69019999			
1038	69229688	WMS WVN TOPS/DRESS	TRUE	42740438			
1138	42691020	STATIONERY & CRAFT	TRUE	69261107			
1191	68967338	WMS WVN TOPS/DRESS	TRUE	42641148			
1297	68406790	MENS U/WEAR/SOCKS	TRUE	42774297			
1401	42460633	FISHING	TRUE	68976477			
1442	42791676	BOOKS	TRUE	42751663			
1464	42719014	FISHING	TRUE	69395468			
1518	42468394	FISHING	TRUE	42784982			
1587	42774907	BOOKS	TRUE	42787990			
1601	69446269	MENS T/S & TOPS	TRUE	42797364			
1614	42803508	BOOKS	TRUE	41991947			
1616	42818465	BOOKS	TRUE	42587774			
1644	42818106	BOOKS	TRUE	42806080			
1740	42722595	BOOKS	TRUE	42800064			
1749	69333835	MENS T/S & TOPS	TRUE	42792239			
1806	69387395	WMS WVN TOPS/DRESS	TRUE	69528279			
1811	67491940	ACTIVE KIDSWEAR	TRUE	69546891			
1935	67826070	WMS BRAS/CO-ORDS	TRUE	69287282			
1937	68583156	WMS BRAS/CO-ORDS	TRUE	68023423			
1960	42494317	LAUNDRY & STORAGE	TRUE	69394447			
1988	69386534	BOYSWEAR 1-8	TRUE	67527601			
2004	69322242	MENS T/S & TOPS	TRUE	42771104			
2014	69155390	GIRLSWEAR 7-16	TRUE	68449353			

Non-reviewed products based on department

The image shows three side-by-side Microsoft Excel spreadsheets. Each spreadsheet has a header row with columns A through E. The first column (SKU Code) contains numerical values, the second column (Department) contains categorical names, and the third column (total) contains numerical values. The fourth column (D) is highlighted in green in all three spreadsheets.

Sheet1 (XMAS TRIM & WRAP):

	A	B	C	D	E
1	SKU Code	Department		total	221
2	42788614	XMAS TRIM & WRAP			
3	42789031	XMAS TRIM & WRAP			
4	42789055	XMAS TRIM & WRAP			
5	42789000	XMAS TRIM & WRAP			
6	42788829	XMAS TRIM & WRAP			
7	42788751	XMAS TRIM & WRAP			
8	42789062	XMAS TRIM & WRAP			
9	42788768	XMAS TRIM & WRAP			
10	42789086	XMAS TRIM & WRAP			
11	42789017	XMAS TRIM & WRAP			
12	42788775	XMAS TRIM & WRAP			
13	42788621	XMAS TRIM & WRAP			
14	42788676	XMAS TRIM & WRAP			
15	42788980	XMAS TRIM & WRAP			
16	42789024	XMAS TRIM & WRAP			
17	42788638	XMAS TRIM & WRAP			
18	42789048	XMAS TRIM & WRAP			
19	42789093	XMAS TRIM & WRAP			
20	42788690	XMAS TRIM & WRAP			
21	42788737	XMAS TRIM & WRAP			
22	42789079	XMAS TRIM & WRAP			
23	42789130	XMAS TRIM & WRAP			
24	42778813	XMAS TRIM & WRAP			
25	42786405	XMAS TRIM & WRAP			
26	42775348	XMAS TRIM & WRAP			
27	42779582	XMAS TRIM & WRAP			
28	42778820	XMAS TRIM & WRAP			
29	42788652	XMAS TRIM & WRAP			

Sheet1 (DINING):

	A	B	C	D	E
1	SKU Code	Department		total	435
2	42692386	DINING			
3	42692409	DINING			
4	41617403	DINING			
5	42799061	DINING			
6	42436164	DINING			
7	42834670	DINING			
8	42104315	DINING			
9	41371244	DINING			
10	42438007	DINING			
11	42758860	DINING			
12	42787525	DINING			
13	42624356	DINING			
14	42436126	DINING			
15	42485377	DINING			
16	42891529	DINING			
17	42799085	DINING			
18	42104018	DINING			
19	42758891	DINING			
20	42758631	DINING			
21	42436188	DINING			
22	42485414	DINING			
23	42436140	DINING			
24	42874867	DINING			
25	42569114	DINING			
26	42875123	DINING			
27	42811909	DINING			
28	42758600	DINING			
29	42710929	DINING			

Sheet1 (ACTION/VEHICLE TOY):

	A	B	C	D	E
1	SKU Code	Department		total	505
2	42223238	ACTION/VEHICLE TOY			
3	42794653	ACTION/VEHICLE TOY			
4	42739302	ACTION/VEHICLE TOY			
5	42620778	ACTION/VEHICLE TOY			
6	42732327	ACTION/VEHICLE TOY			
7	42762539	ACTION/VEHICLE TOY			
8	42732341	ACTION/VEHICLE TOY			
9	42661276	ACTION/VEHICLE TOY			
10	42756101	ACTION/VEHICLE TOY			
11	42475743	ACTION/VEHICLE TOY			
12	42690047	ACTION/VEHICLE TOY			
13	42690016	ACTION/VEHICLE TOY			
14	42756187	ACTION/VEHICLE TOY			
15	42788089	ACTION/VEHICLE TOY			
16	42690023	ACTION/VEHICLE TOY			
17	42689966	ACTION/VEHICLE TOY			
18	42606000	ACTION/VEHICLE TOY			
19	42804895	ACTION/VEHICLE TOY			
20	41779675	ACTION/VEHICLE TOY			
21	42854326	ACTION/VEHICLE TOY			
22	42638155	ACTION/VEHICLE TOY			
23	42850564	ACTION/VEHICLE TOY			
24	42797364	ACTION/VEHICLE TOY			
25	42798323	ACTION/VEHICLE TOY			
26	42798415	ACTION/VEHICLE TOY			
27	42798293	ACTION/VEHICLE TOY			
28	42798378	ACTION/VEHICLE TOY			
29	42325802	ACTION/VEHICLE TOY			

SKU's with less rating

Quality Sco..	SKU Code	value for money score					
		1	2	3	4	5	(blank)
1	41993194		1.000				
	42001416	1.000					
	42001898	1.000					
	42002086	1.000					
	42008460	1.00					
	42010395	1.00					
	42014249	1.00					
	42015932	2.00					
	42018551	1.00					
	42026785	1.000					
	42028406	1.000					
	42029946		1.000				
	42038382	1.000					
	42049203		1.000				
	42051381	1.000	1.000	2.000	1.000		
	42069522			1.000			
	42069560			1.000			
	42074410	1.000					
	42076315			2.000			
	42077459	1.000					
	42080176				1.000		
	42082712					1.000	
	42099666	1.000					
	42099888	1.000					
	42104735			1.000			
	42106692				3.000		

Locating Single star reviews in catalogue data

	C	D	E	F	G	H	N	O	P	P	Q	R
1	Review Tit	Review Co	Rating	Product Title	Product URL	SKU Code	sku row number	SKU Code				
2	I hated it.	I hated it.	1	Jute Runner	https://www.kmart.com.au/product/jute-runner	42570677	38017	39333018	42707028	INDOOR LIVING	WALL ACCESSORIES	
3	FIRE HAZA	As soon as	1	3 in 1 Accessory Pack	https://www.kmart.com.au/product/3-in-1-accessory-pack	42551249	82022	39333018	42707028	INDOOR LIVING	WALL ACCESSORIES	
4	Good.	Good fit.	1	5 Pack Work Socks-Blk/gry/6-7	https://www.kmart.com.au/product/5-pack-work-socks-blk-gry-6-7	68635367	139039	39333018	42707028	INDOOR LIVING	WALL ACCESSORIES	
5	Used twice	The handh	1	2 in 1 Cordless Stick Vacuum	https://www.kmart.com.au/product/2-in-1-cordless-stick-vacuum	42735892	265668	39333018	42707028	INDOOR LIVING	WALL ACCESSORIES	
6	Product g	Unable to	1	Skate Helmet - Medium Blue	https://www.kmart.com.au/product/skate-helmet-medium-blue	42642329	156420	42163053	42707028	INDOOR LIVING	WALL ACCESSORIES	
7	Nice top	Bought thi	1	Extended Sleeve Top-Blue Pe	https://www.kmart.com.au/product/extended-sleeve-top-blue-pe	69510304	194414	42163053	42707028	INDOOR LIVING	WALL ACCESSORIES	
8	Broken	This clock	1	Metro Wall Clock	https://www.kmart.com.au/product/metro-wall-clock	41337158	37347	42163053	42707028	INDOOR LIVING	WALL ACCESSORIES	
9	busted bin	Received t	1	30L Brushed Stainless Steel B	https://www.kmart.com.au/product/30l-brushed-stainless-steel-bin	42747826	40144	42163053	42694786	INDOOR LIVING	DECORATIVE ITEMS	
10	Products v	Products v	1	White Desk Drawers	https://www.kmart.com.au/product/white-desk-drawers	42698494	24418	42163053	42694786	INDOOR LIVING	DECORATIVE ITEMS	
11	This air be	This air be	1	Queen Air Bed with Built-in P	https://www.kmart.com.au/product/queen-air-bed-with-built-in-pump	42666394	218981	42163053	42694786	INDOOR LIVING	DECORATIVE ITEMS	
12	Not compl	3 glazed px	1	3 Glazed Pot Plants	https://www.kmart.com.au/product/3-glazed-pot-plants	42643944	174358	42163053	42694786	INDOOR LIVING	DECORATIVE ITEMS	
13	Disappoint	Ordered th	1	Tan Ottoman	https://www.kmart.com.au/product/tan-ottoman	42760351	133482	42163053	42694786	INDOOR LIVING	DECORATIVE ITEMS	
14	Oversized	Very overs	1	Plus Size Long Sleeve Texture	https://www.kmart.com.au/product/plus-size-long-sleeve-texture	69379918	257608	42163053	41337158	INDOOR LIVING	CLOCKS	
15	Shelf	Awsome s	1	4 Tier Shelf with Bamboo Shelves	https://www.kmart.com.au/product/4-tier-shelf-with-bamboo-shelves	42187585	211084	42163053	41337158	INDOOR LIVING	CLOCKS	
16	Cheap and Horrible ch	1	Emma Wiggles Costume - To	https://www.kmart.com.au/product/emma-wiggles-costume-toy	42598497	41693	42163053	41337158	INDOOR LIVING	CLOCKS		
17	This vacuu	This vacuu	1	1200W Upright Vacuum	https://www.kmart.com.au/product/1200w-upright-vacuum	42701668	221266	42163053	41337158	INDOOR LIVING	CLOCKS	
18	Absolutely My	table c	1	Coffee Table Oak Look	https://www.kmart.com.au/product/coffee-table-oak-look	42619697	31370	42163053	41337158	INDOOR LIVING	CLOCKS	
19	Flimsy & ar	top was da	1	Walnut Look Desk	https://www.kmart.com.au/product/walnut-look-desk	42715696	144175	42163053	41337158	INDOOR LIVING	CLOCKS	
20	Item not d	Item not d	1	Fish Bowl Vase	https://www.kmart.com.au/product/fish-bowl-vase	40345437	161431	42163053	41337158	INDOOR LIVING	CLOCKS	
21	Was appal	Was appal	1	Vacuum Food Sealer	https://www.kmart.com.au/product/vacuum-food-sealer	42584896	244868	41882917	41337158	INDOOR LIVING	CLOCKS	
22	I had a ter	I had a ter	1	Large White Embossed Canis	https://www.kmart.com.au/product/large-white-embossed-canis	42576617	171432	41882917	41337158	INDOOR LIVING	CLOCKS	
23	Dissapoint	Does not s	1	VTech Lullaby Lambs Mobile	https://www.kmart.com.au/product/vtech-lullaby-lambs-mobile	42638698	229694	41882917	41337158	INDOOR LIVING	CLOCKS	
24	Came dam	Came dam	1	Industrial 4 Tier Bookshelf	https://www.kmart.com.au/product/industrial-4-tier-bookshelf	42758358	117804	41882917	41337158	INDOOR LIVING	CLOCKS	
25	Delivered i	One out of	1	Hanging Sloth Pot	https://www.kmart.com.au/product/hanging-sloth-pot	42754183	104782	41882917	41337158	INDOOR LIVING	CLOCKS	
26	Can't find	Can't find	4	Small Mirror	https://www.kmart.com.au/product/small-mirror	42622200	200384	40000000	69816628	GIRLSWEAR 1-8	KNIT TOPS	

Locating Single star reviews in catalogue data

E	F	G	H	N	O	P	Q	R
Rating	Product Title	Product URL	SKU Code	sku row number	SKU Code	P	INDOOR LIVING	SMALL
1	Jute Runner	https://www.kmart.com.au/product/jute-runner	42570677	38017	3933018	42570677	INDOOR LIVING	SMALL
1	3 in 1 Accessory Pack	https://www.kmart.com.au/product/3-in-1-accessory-pack	42551249	82022	3933018	42570677	INDOOR LIVING	SMALL
1	5 Pack Work Socks-Blk/gry,6- https://www.kmart.com.au/product/5-pack-work-socks-blk-gry-6	https://www.kmart.com.au/product/5-pack-work-socks-blk-gry-6	68635367	139039	3933018	42570677	INDOOR LIVING	SMALL
1	2 in 1 Cordless Stick Vacuum	https://www.kmart.com.au/product/2-in-1-cordless-stick-vacuum	42735892	265668	3933018	42570677	INDOOR LIVING	SMALL
1	Skate Helmet - Medium Blue	https://www.kmart.com.au/product/skate-helmet-medium-blue	42642329	156420	42163053	42570677	INDOOR LIVING	SMALL
1	Extended Sleeve Top-Blue Pe	https://www.kmart.com.au/product/extended-sleeve-top-blue-pe	69510304	194414	42163053	42570677	INDOOR LIVING	SMALL
1	Metro Wall Clock	https://www.kmart.com.au/product/metro-wall-clock	41337158	37347	42163053	42570677	INDOOR LIVING	SMALL
1	30L Brushed Stainless Steel B	https://www.kmart.com.au/product/30l-brushed-stainless-steel-bin	42747826	40144	42163053	42570677	INDOOR LIVING	SMALL
1	White Desk Drawers	https://www.kmart.com.au/product/white-desk-drawers	42698494	24418	42163053	42570677	INDOOR LIVING	SMALL
1	Queen Air Bed with Built-in Pump	https://www.kmart.com.au/product/queen-air-bed-with-built-in-pump	42666394	218981	42163053	42570677	INDOOR LIVING	SMALL
1	3 Glazed Pot Plants	https://www.kmart.com.au/product/3-glazed-pot-plants	42643944	174358	42163053	42570677	INDOOR LIVING	SMALL
1	Tan Ottoman	https://www.kmart.com.au/product/tan-ottoman	42760351	133482	42163053	42570677	INDOOR LIVING	SMALL
1	Plus Size Long Sleeve Texture	https://www.kmart.com.au/product/plus-size-long-sleeve-texture	69379918	257608	42163053	42570677	INDOOR LIVING	SMALL
1	4 Tier Shelf with Bamboo Shelves	https://www.kmart.com.au/product/4-tier-shelf-with-bamboo-shelves	42187585	211084	42163053	42570677	INDOOR LIVING	SMALL
1	Emma Wiggles Costume - Toddler	https://www.kmart.com.au/product/emma-wiggles-costume-toddler	42598497	41693	42163053	42570677	INDOOR LIVING	SMALL
1	1200W Upright Vacuum	https://www.kmart.com.au/product/1200w-upright-vacuum	42701668	221266	42163053	42570677	INDOOR LIVING	SMALL
1	Coffee Table Oak Look	https://www.kmart.com.au/product/coffee-table-oak-look	42619697	31370	42163053	42570677	INDOOR LIVING	SMALL
1	Walnut Look Desk	https://www.kmart.com.au/product/walnut-look-desk	42715696	144175	42163053	42570677	INDOOR LIVING	SMALL
1	Fish Bowl Vase	https://www.kmart.com.au/product/fish-bowl-vase	40345437	161431	42163053	42570677	INDOOR LIVING	SMALL

Images of some Single star review products based on department

Indoor living

Free Click & Collect on orders over \$20



Metro Wall Clock

 32 Review

\$10.00

bga

Add to bag



Available on orders \$70 to \$200

[Learn More](#)



Available on orders \$70 to \$100

[Learn More](#)

kitchen

Free Click & Collect on orders over \$20



Search Stores Cart Sign in

Christmas Latest Arrivals Home & Living Technology Toys Women Men Kids & Baby Sports & Outdoor Inspo Catalogue



30L Brushed Stainless Steel Bin

 31 Reviews

\$27.0

 Add to wishlist

1

Add to bag



Available on orders \$70 to \$1000

[Learn More](#)

Insights:

2019 & 2020

There is a common word and topic observed from 2019 & 2020 is x27



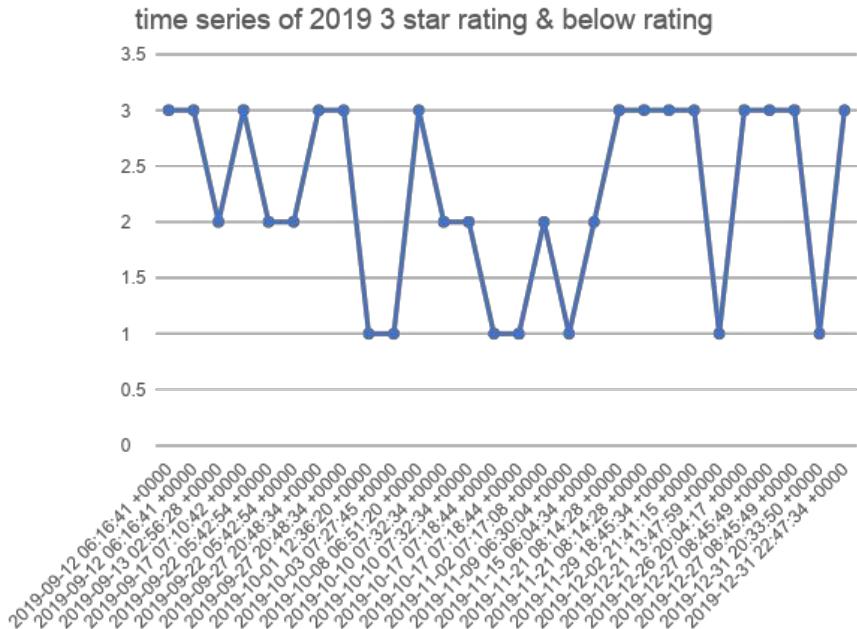
In 2019 customers spoke good about online order but for 2020 it was spoken less about online orders



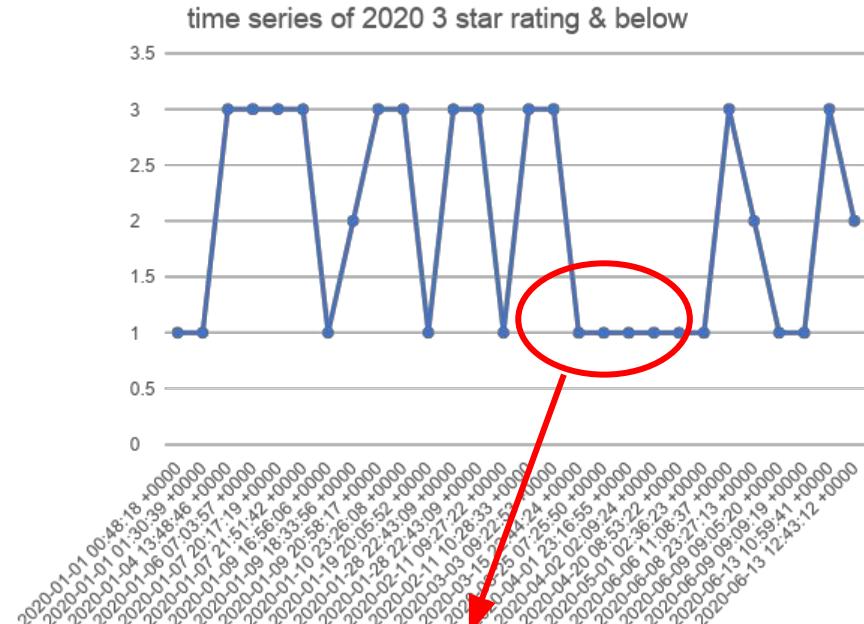
2020

Time series analysis for 2019 & 2020 reviews

2019



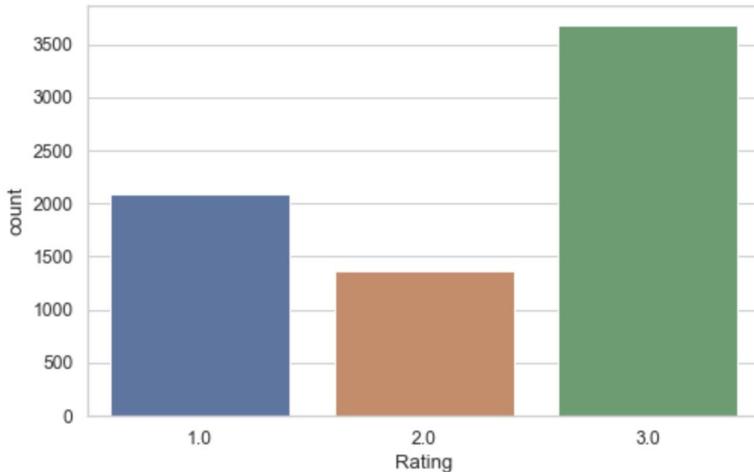
2020



lot of 1 star reviews in months 3,4,5,6 of 2020

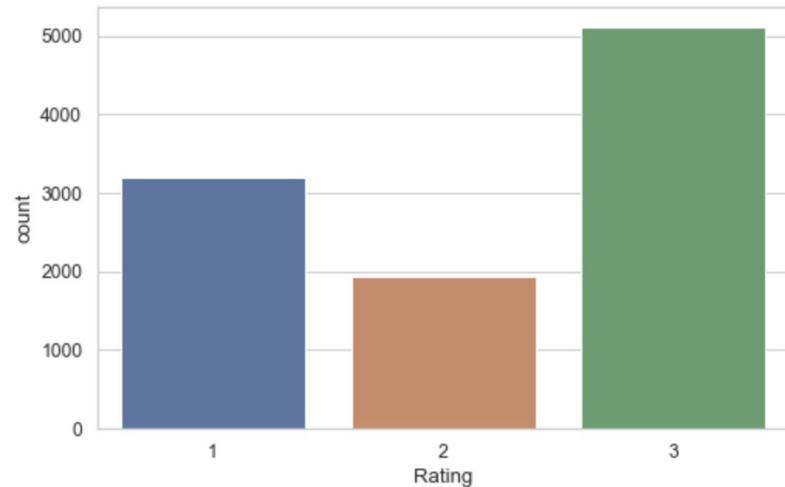
Rating for 1,2 and 3 star

2019



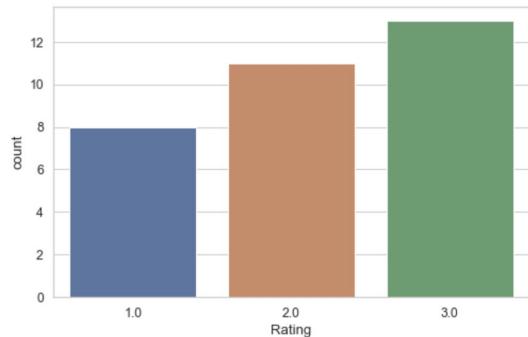
For the 2019 data people more than 3500 have rated the products to be 3 star, More than 2000 people have rated 1 star and Around 1400 people have rated the product to be 2 star.

2020

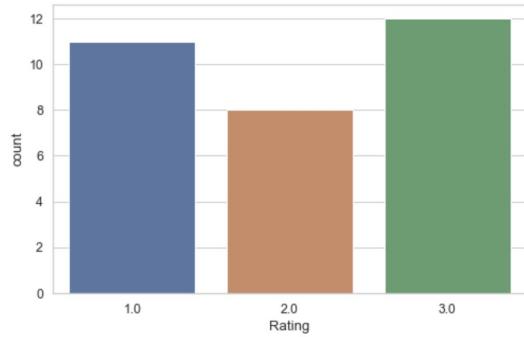


In 2020 people more than 5000 have given rating for 3 star, 1-star rating has been given by more than 3000 people and 2-star rating was given by people around count of 1900.

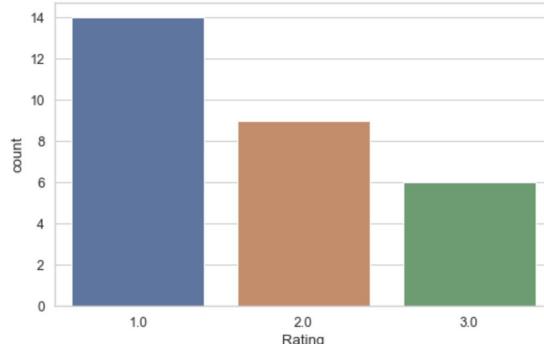
2019 top 10 frequently rated products



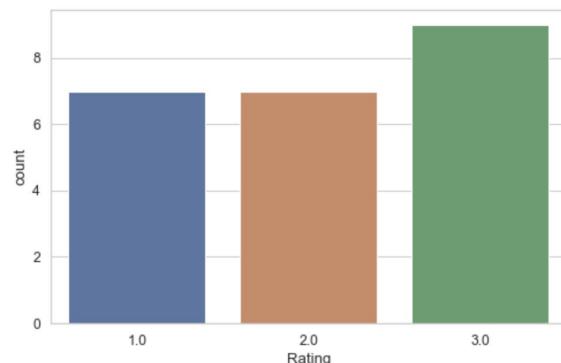
2 in 1 Cordless Stick Vacuum



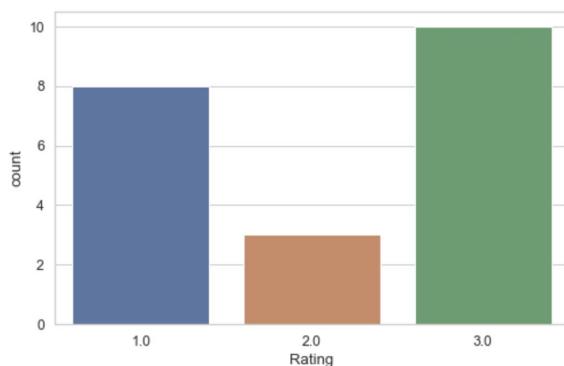
8 Cube Unit White



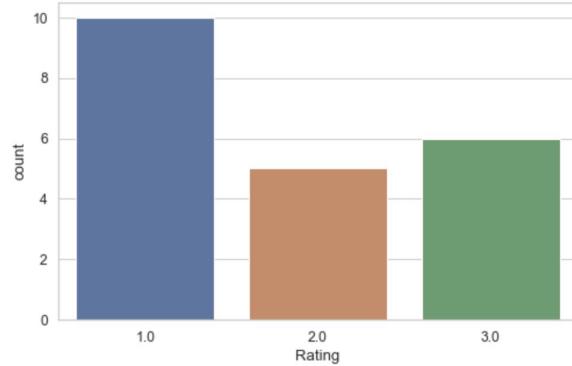
Cordless Stick Vacuum Cleaner



30L Soft Close Bin Black

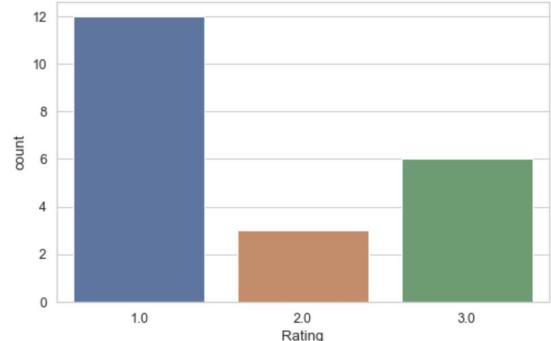


2000W Bagless Vacuum



Magnetic Exercise Bike

2019 top 10 frequently rated products



Industrial 2 Drawer Unit

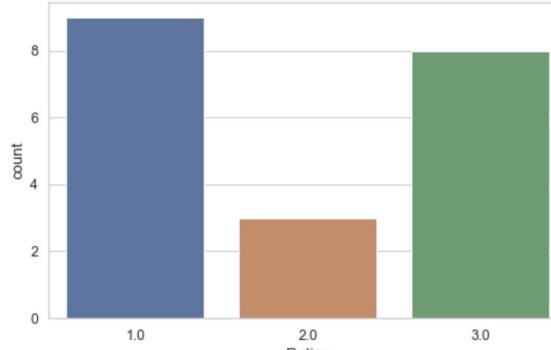
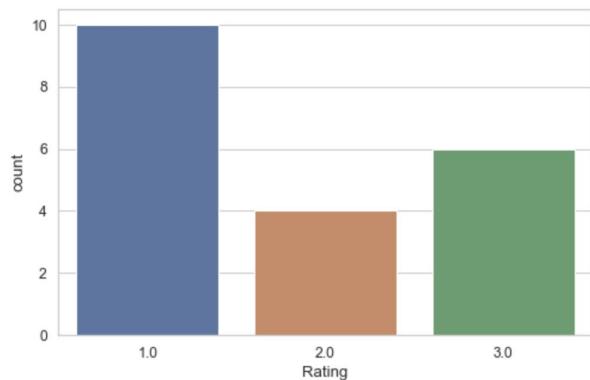
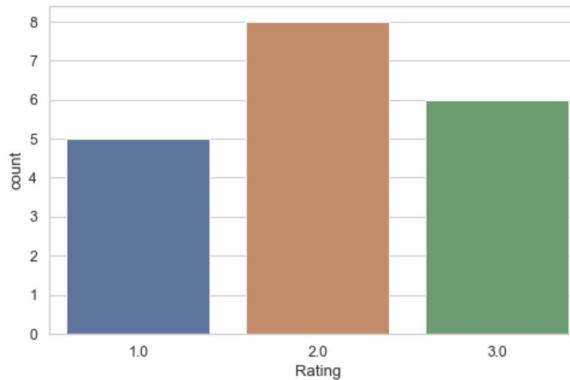


Table Tennis Table

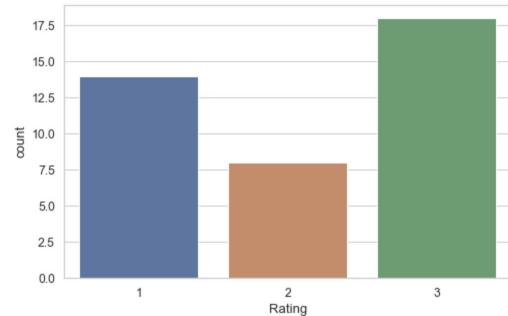


True Wireless Earphones - White

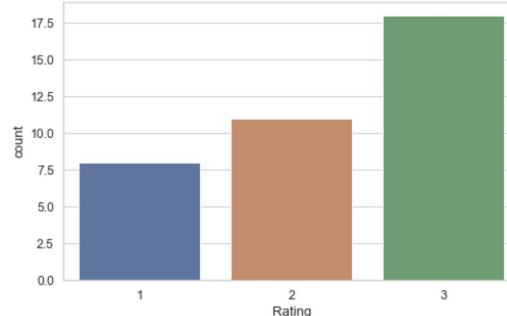


Flokati Style Rug

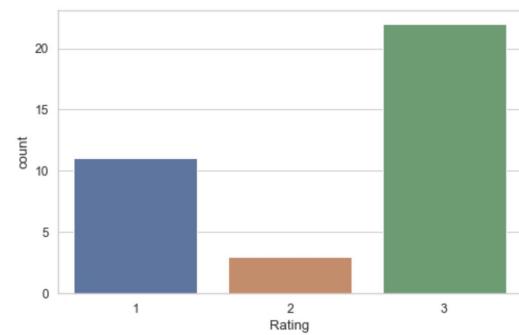
2020 top 10 frequently rated products



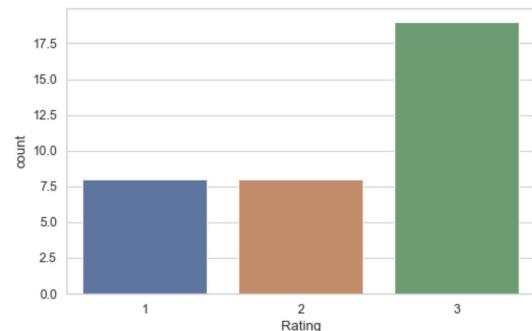
8 Cube Unit White



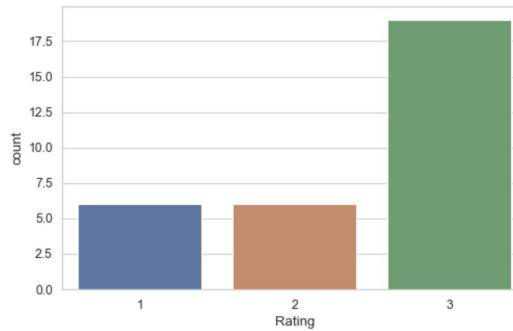
Industrial 2 Drawer Unit



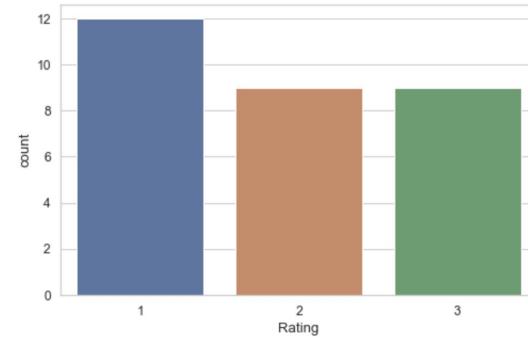
2 in 1 Cordless Stick Vacuum



2000W Bagless Vacuum

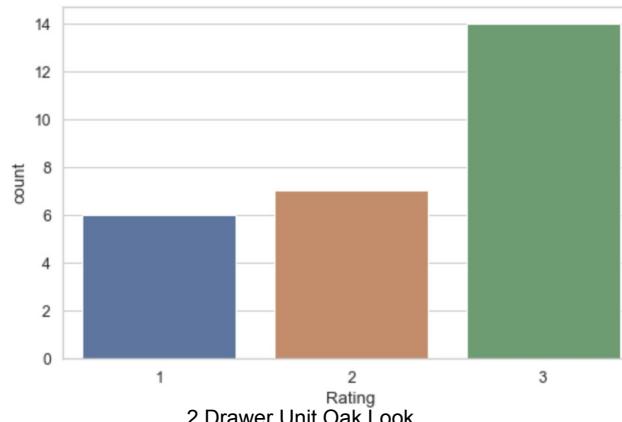
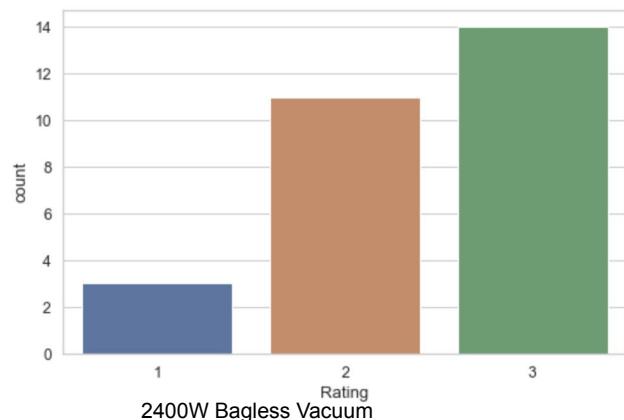
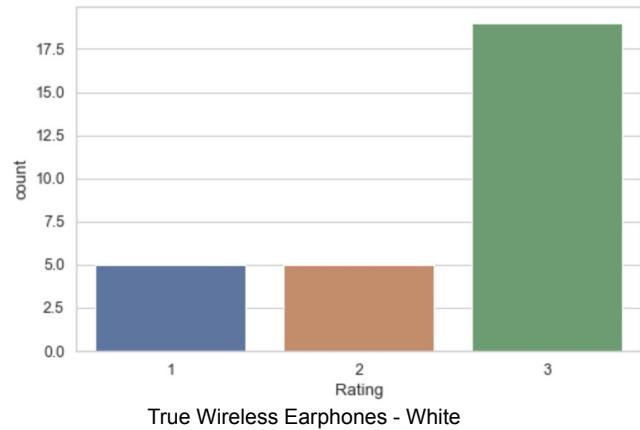


Oak Look Chest of Drawers



LED Strip Light with Remote - 5m Cable Length

2020 top 10 frequently rated products



Product Rating Analysis

```
In [15]: 1 df_2019.groupby('Product Title')['Rating'].count().sort_values(ascending=False)
```

Out[15]: Product Title

3.2 Litre Air Fryer	455
Foldable Beach Trolley	242
Pie Maker	234
2 in 1 Cordless Stick Vacuum	230
5.3 Litre Air Fryer	226
...	
Striped Pinafore-Stripe,7	1
L.O.L. Surprise Lils Makeover Series Doll	1
Shirred Cami-Arty,8	1
L.O.L. Surprise Knit Pyjama Set-L.o.l,6	1
Short Sleeve V-Neck Top-Sage Green,14	1
Name: Rating, Length: 15801, dtype: int64	

```
In [33]: 1 df_2020.groupby('Product Title')['Rating'].count().sort_values(ascending=False)
```

Out[33]: Product Title

3.2 Litre Air Fryer	947
5.3 Litre Air Fryer	602
3-in-1 Air Fryer Oven	287
34L Microwave	270
2000W Bagless Vacuum	268
...	
Full Figure Soft Cup Underwire Bra-Slate Rose,22D	1
Full Length Leggings-Black Marl,8	1
Full Length Leggings-Black,8	1
Full Length Leggings-Sky Captain,8	1
\$50 iTunes Cards	1
Name: Rating, Length: 20871, dtype: int64	

So to avoid misleadings, we have plot the total number of ratings for a particular product.

Product Title, Ratings and Number of Rating Counts

In [18]: 1 ratings_mean_count

Out[18]:

Product Title	Rating	rating_counts
1 Litre Glass Food Storage Container	4.500000	4
1 Litre Glass Measuring Jug	5.000000	10
1 Pack 40cm x 50cm Stretched Canvas	4.777778	9
1 Pack Sports Compression Socks-Black,11-14	4.000000	4
1 Pack Sports Compression Socks-Black,6-10	5.000000	1
...
iPhone XR Silicone Case - Black	5.000000	2
iPhone XR Silicone Case - Blush	4.500000	4
iPhone XS MAX Guard Flip Case - Black	5.000000	2
iPhone XS Max Glass Screen Protector	1.000000	1
smART Sketcher Projector	4.666667	3

15801 rows × 2 columns

In [34]: 1 ratings_mean_count

Out[34]:

Product Title	Rating	rating_counts
\$50 iTunes Cards	1.000000	1
1 Litre Glass Food Storage Container	4.760000	25
1 Litre Glass Measuring Jug	4.666667	6
1 Litre Measuring Jug - Assorted	5.000000	4
1 Pack 40cm x 50cm Stretched Canvas	4.571429	14
...
iPhone XS MAX Guard Flip Case - Rose Gold Look	5.000000	4
iPhone XS Max / 11 Pro Max Privacy Hardened Glass Screen Protector	5.000000	2
iPhone XS Max Glass Screen Protector	5.000000	2
iPhone XS Max/11 Pro Max Hardened Glass Screen Protector	4.000000	4
smART Sketcher Projector	4.285714	7

20871 rows × 2 columns

Sentiment Analysis

- Sentiment analysis is part of the Natural Language Processing (NLP) techniques that consists in extracting emotions related to some raw texts.
- This is usually used on social media posts and customer reviews in order to automatically understand if some users are positive or negative and why.

**CUSTOMER FEEDBACK
SENTIMENT ANALYSIS**



Sentiment Analysis

Sentiment Analysis



My experience so far has been **fantastic!**

POSITIVE



The product is **ok I guess**

NEUTRAL



Your support team is **useless**

NEGATIVE

Lemmatization

A method that switches any kind of a word to its base root mode is called Lemmatization. Lemmatization is a method responsible for grouping different inflected forms of words into the root form, having the same meaning. It is similar to stemming, in turn, it gives the stripped word that has some dictionary meaning.

Stemming

word has one root-base form but having different variations, for example, “play” is a root-base word and playing, played, plays are the different forms of a single word. So, these words get stripped out, they might get the incorrect meanings or some other sort of errors. The process of reducing inflection towards their root forms are called Stemming

Stop words

- A stop word is a commonly used word (such as “the”, “a”, “an”, “in”) that a search engine has been programmed to ignore, both when indexing entries for searching and when retrieving them as the result of a search query.

Sentiment Analysis

```
In [12]: 1 reviewsDf['New Text'] = reviewsDf['Review Content'].apply(preproc)
```

```
In [13]: 1 reviewsDf.head()
```

	Rating	Review Content	New Text
0	5	Excellent great value for money my son uses it...	excellent great value money son us lego table
1	5	Our dogs love them. Easy to break in half to m...	dog love easy break half make last long
2	5	GreAt work k mart ...was delivered quickly and...	great work k mart deliver quickly safely
3	5	Great product	great product
4	5	Very quick service great product,excellent price	quick service great product excellent price

- Created a new text after using STOP WORDS
- STEMMING
- LEMMATIZATION

DEFINED A CLASS POSITIVE OR NEGATIVE FROM REVIEWS

```
In [15]: 1 reviewsDf.head()
```

	Rating	Review Content	New Text	Class
0	5	Excellent great value for money my son uses it...	excellent great value money son us lego table	Positive
1	5	Our dogs love them. Easy to break in half to m...	dog love easy break half make last long	Positive
2	5	GreAt work k mart ...was delivered quickly and...	great work k mart deliver quickly safely	Positive
3	5	Great product	great product	Positive
4	5	Very quick service great product,excellent price	quick service great product excellent price	Positive

Sentiment analysis and level of severity

A	B	C	D	E
1	Review content	Sentiment	Score	level of severity
2	Excellent experience	positive	84%	good
3	Our dogs love it	neutral	53%	could be better
4	Great work!	positive	68%	good
5	Great product	positive	79%	good
6	Very quick delivery	positive	87%	good
7	The hair wash is	negative	25%	very severe
8	Great decoration	positive	82%	good
9	They fade quickly	negative	31%	severe
10	Great quality	positive	90%	good
11	For the price, it's	positive	95%	good
12	Little miss	positive	70%	good
13	The price is reasonable	positive	79%	good
14	The lids are	neutral	46%	could be better
15	It's a deal	positive	79%	good
16	Half of the time	negative	20%	very severe
17	Really nice	positive	81%	good
18	This is a great product	neutral	55%	could be better
19	I hated it.	neutral	60%	could be better
20	Very good	positive	78%	good
21	I love this	positive	89%	good
22	This is the best	positive	75%	good
23	Very nice product	negative	39%	severe
24	Love it	positive	82%	good
25	Perfect fit, I like it	positive	91%	good
26	As soon as I received it	negative	24%	very severe
27	Jumper is very comfortable	negative	26%	severe
28	My son loves it	positive	61%	good
29	Good fit.	positive	74%	good

- 0 to 25% very severe
- 25 to 39% severe
- 40 to 60 could be better
- 60 and above good



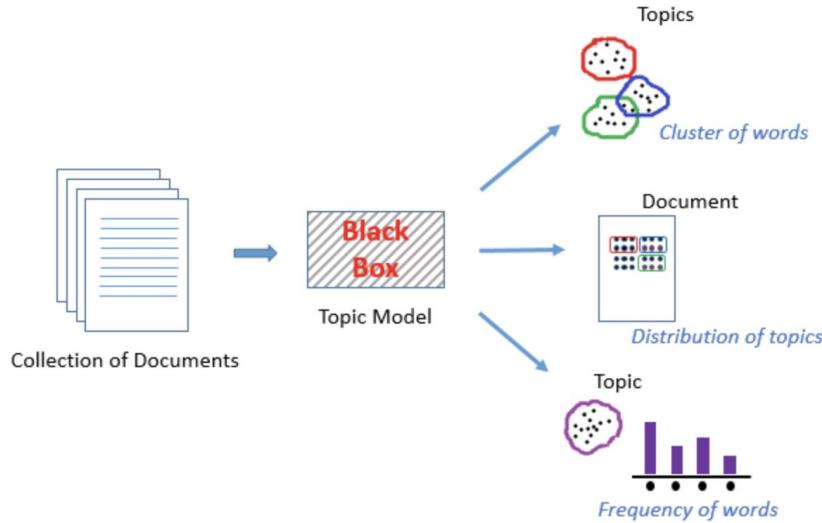
Topic modelling and LDA

In general terms topic modelling helps us to understand the ‘topics’ that best represents the information from them

Used LDA (*Latent Dirichlet Allocation*) algorithm for implementing topic modelling

Topic Modelling

Topic modeling is an unsupervised machine learning technique that's capable of scanning a set of documents, detecting word and phrase patterns within them, and automatically clustering word groups and similar expressions that best characterize a set of documents.



Topics for the year 2019 & 2020

2019

```
***** Topic 1 *****  
['like', 'size', 'small', 'need', 'work', 'make', 'use', 'little', 'bit', 'easy']  
  
***** Topic 2 *****  
['love', 'great', 'fit', 'nice', 'look', 'size', 'comfortable', 'colour', 'perfect', 'little']  
  
***** Topic 3 *****  
['great', 'good', 'quality', 'price', 'value', 'product', 'happy', 'money', 'excellent', 'purchase']
```

2020

```
***** Topic 1 *****  
['great', 'good', 'quality', 'price', 'value', 'size', 'fit', 'look', 'money', 'love']  
  
***** Topic 2 *****  
['love', 'great', 'easy', 'little', 'work', 'make', 'kid', 'perfect', 'use', 'look']  
  
***** Topic 3 *****  
['purchase', 'love', 'product', 'delivery', 'kmart', 'happy', 'order', 'item', 'service', 'recommend']
```

- The topics for the year 2019 and 2020
- For the year 2019 the topics named as
 - TOPIC 1: Home Décor
 - TOPIC 2: Clothing
 - TOPIC 3: Delivery

For the year 2020 the topics named as

- TOPIC 1: Home & Living
- TOPIC 2: Clothing
- TOPIC 3: Service

2019_Topic 3_Delivery

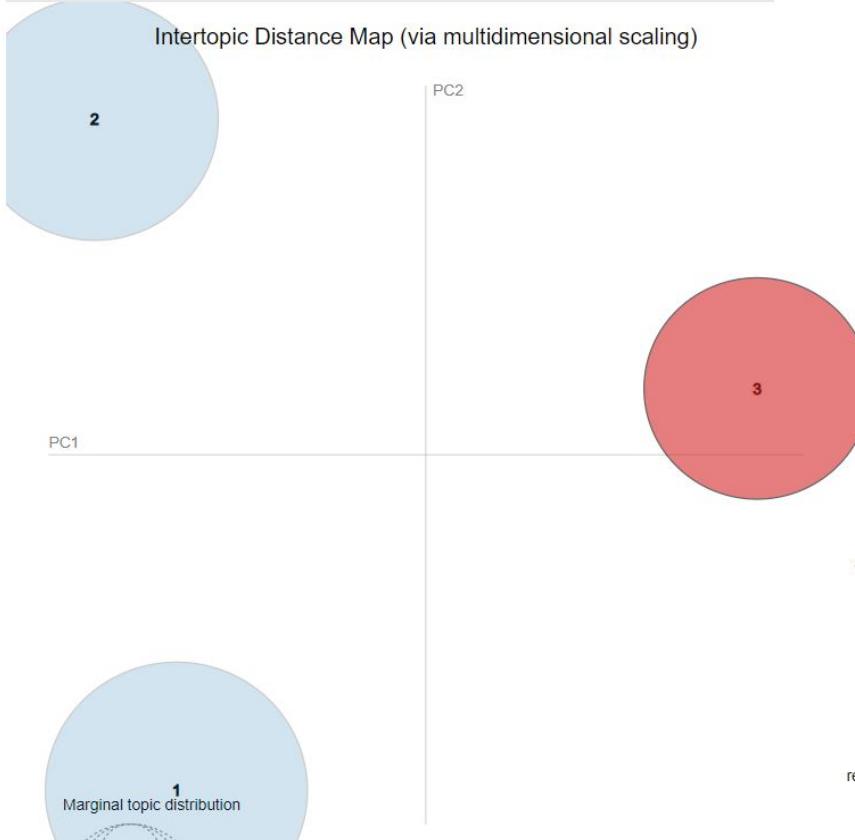
Selected topic: 3 | Previous topic | Next topic | Clear topic

Slide to adjust relevance metric:⁽²⁾

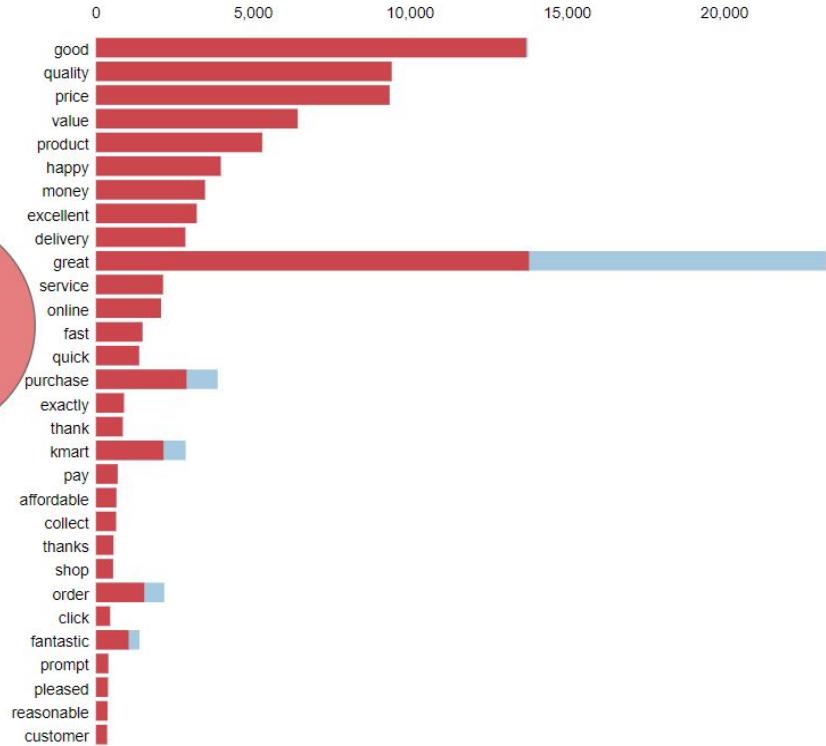
$\lambda = 0.23$

0.0 0.2 0.4 0.6 0.8 1.0

Intertopic Distance Map (via multidimensional scaling)



Top-30 Most Relevant Terms for Topic 3 (28.2% of tokens)



2020_Topic 1_Home and Living



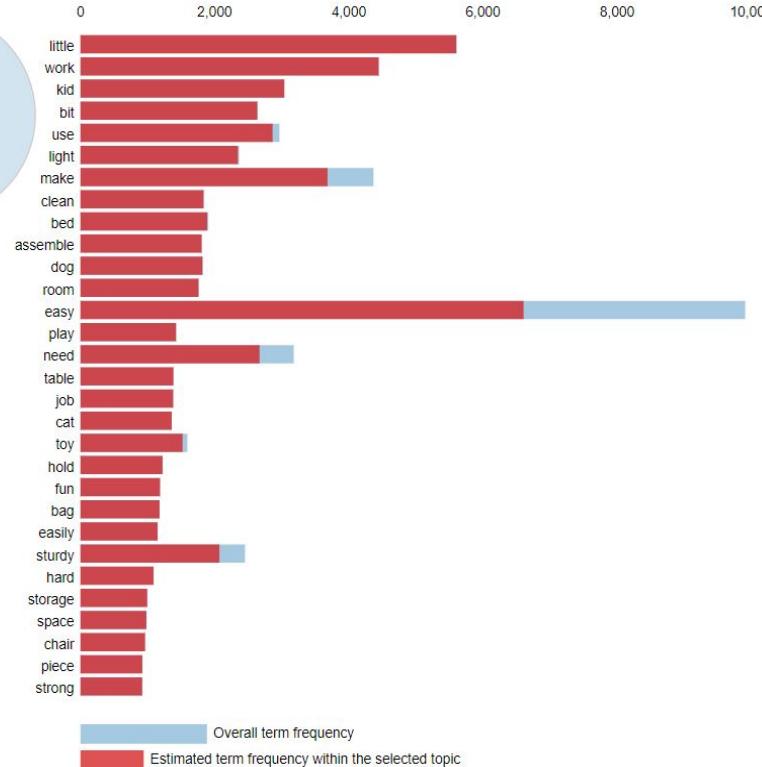
Intertopic Distance Map (via multidimensional scaling)



Marginal topic distribution



Top-30 Most Relevant Terms for Topic 1 (40.3% of tokens)



Overall term frequency

Estimated term frequency within the selected topic

1. saliency(term w) = frequency(w) * $\sum_{t} b(t | w) * \log(b(t | w) / b(t))$ for topics t; see Chuang et. al (2012)

Word Cloud Image for Most Negative Reviews



Insights from Word Cloud Images are:

Quality

Terrible

Service

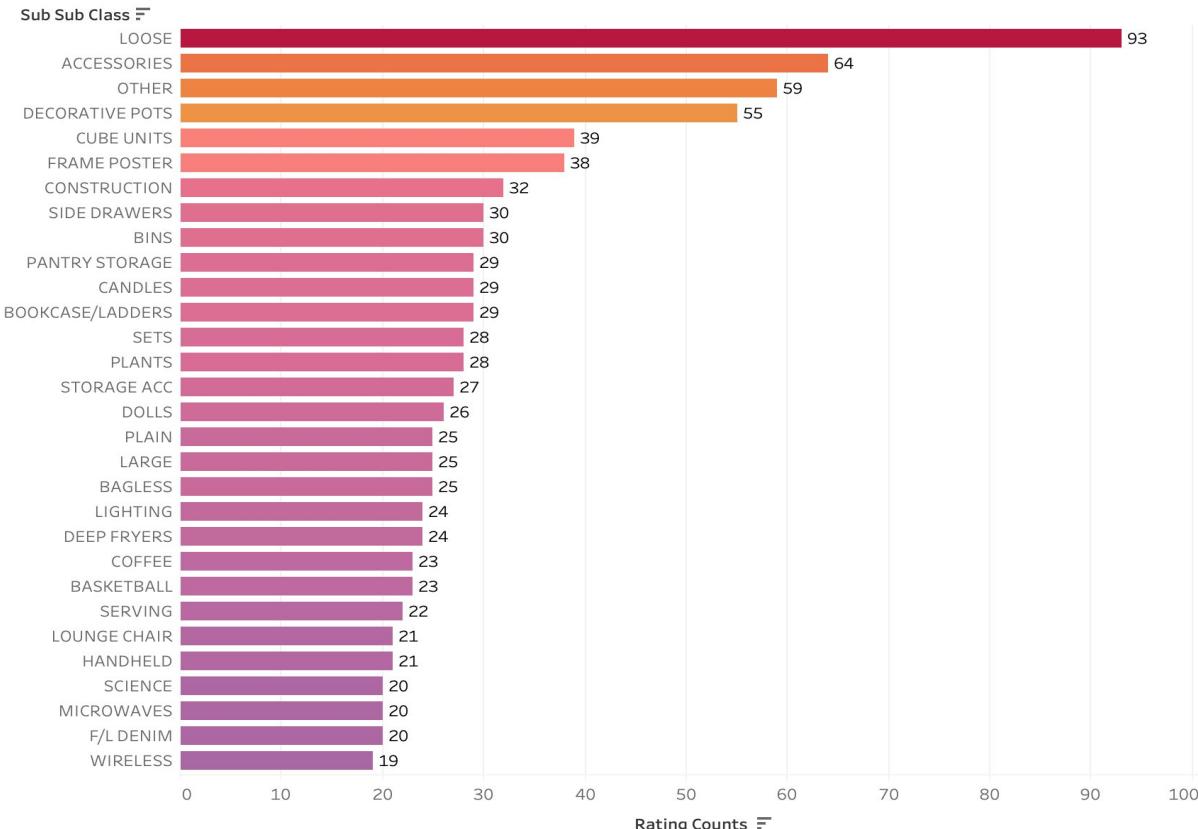
Broke

Delivery

Online

Shipping

Products with number of negative ratings



List of number of products with Negative rating.

“Loose” is for dining and kitchen department that has been rated negative by 93 people followed by Accessories, ”Other” includes pet care, indoor living.

SKU Code with number of negative ratings

- We have considered only the negative reviews. The SKU code 42633297 has total rating counts 23.



8 Cube Unit White

\$35.00

SKU Code	rating_counts
42633297	23
42766049	12
42588924	12
42735892	11
42753193	11
...	...
42832119	1
42832164	1
42832577	1
42832768	1
14323287	1

2415 rows × 1 columns

[35] neg_review_wordcloud

	cleaned_review	SKU	Code	Sub	Sub	Class
8489	false illusion similar quality four cube shelf...	42633297		CUBE	UNITS	
8497	extremely easy put together less 30mins hate f...	42633297		CUBE	UNITS	
8501	terrible hadnx27t pulled taken back like onei ...	42633297		CUBE	UNITS	
8508	ordered 8 jan reveived 22 jan bad experience	42633297		CUBE	UNITS	
8513		poor quality	42633297		CUBE	UNITS
8515	first time haven't able put something together...	42633297		CUBE	UNITS	
8517	strong first delivered broken second broke son...	42633297		CUBE	UNITS	
8519	unpacked box build several wooden screws broke...	42633297		CUBE	UNITS	
8520	don't buy product cheap crap quality slants st...	42633297		CUBE	UNITS	
8534		ok trouble desk purchased	42633297		CUBE	UNITS
8535	looked plan pieces practical person parts inst...	42633297		CUBE	UNITS	
8541	smallish flimsyish regretted spending bit extr...	42633297		CUBE	UNITS	
8549	flatpacks usually mediocre put together many e...	42633297		CUBE	UNITS	
8554		poor quality difficult build due enough pieces	42633297		CUBE	UNITS
8562		absolutely horrible instructions	42633297		CUBE	UNITS
8577		let say hate flat packs instructions vague	42633297		CUBE	UNITS

Negative Review for 8 cube unit white

8 Cube Unit White

\$35.00

Topics for Negative terms for Cube Units

Topic #0: hard nails, going hard, buy hard

Topic #1: hate flat, 30mins hate, flat packs

Topic #2: wrong made, shelf wrong, shelf wrong made

Topic #3: half assembled, box half, box half assembled

Topic #4: cheap nasty, nasty chip, cheap nasty chip

Topic #5: together less, less 30mins, together less 30mins

Topic #6: together unfortunately, unfortunately two, together unfortunately two



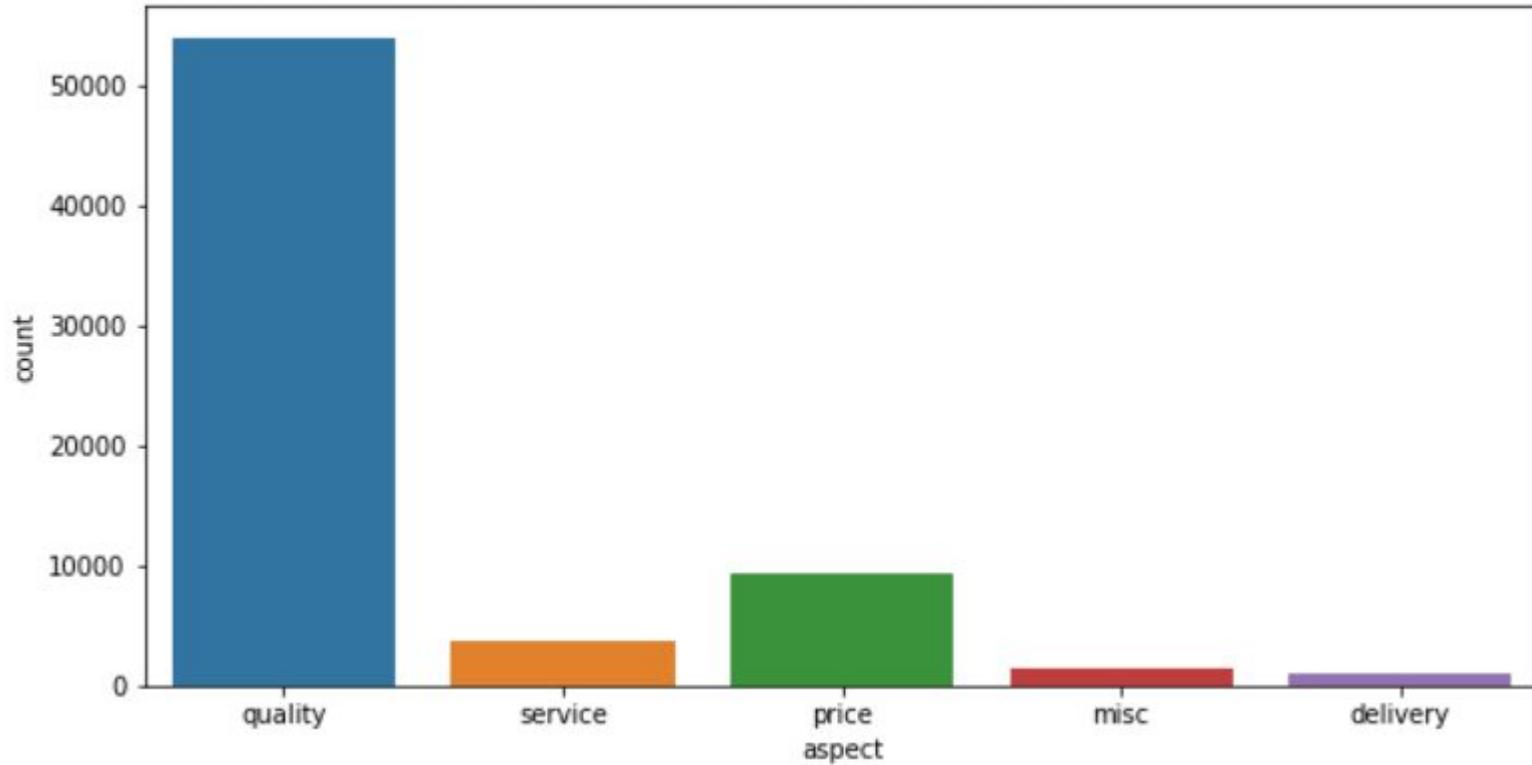
8 Cube Unit White

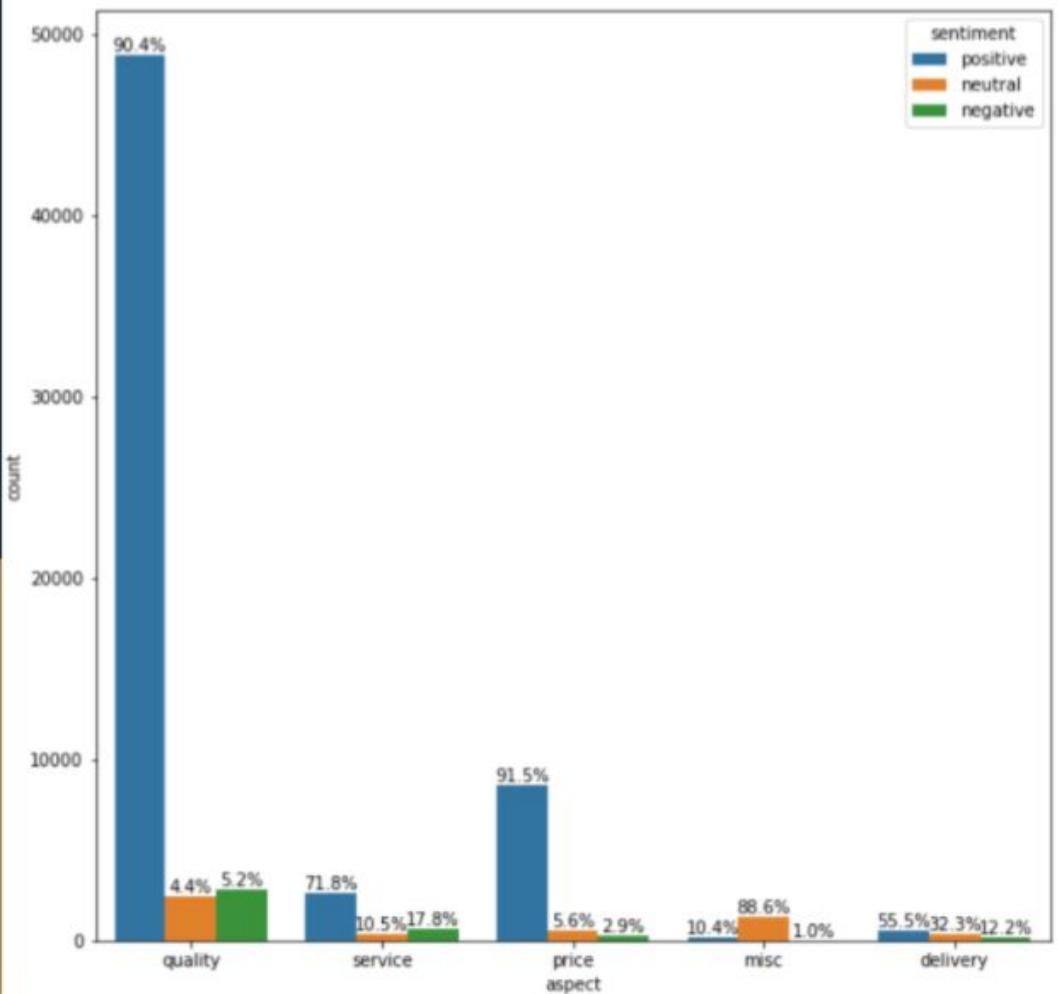
\$35.00

Aspect Based sentiment analysis

Aspect-Based Sentiment Analysis aims to extract both the entity described in the text (attributes or components of a product or service) and the sentiment expressed towards such entities.

<matplotlib.axes._subplots.AxesSubplot at 0x7f18bb97e510>





Kmart Reviews

Aspect Model with
percentage of aspects.



Value Counts for Department

```
[ ] df.value_counts('Department')
```

Department	
APPLIANCES	4975
PET CARE	3732
INDOOR LIVING	3498
BED LINEN	3226
INDOOR FURNITURE	2966
...	
FISHING	29
WMS MATERNITY	21
WMS BRIEFS/FUNCTIO	17
SEASONAL FOOD	8
GIFT CARDS & RCHRG	1
Length: 66, dtype: int64	

Appliances Department has the most number of reviews.

So, Appliances department has been taken for our analysis.

Value counts for SKU Code for Appliances Department

```
[ ] df1.value_counts('SKU Code')
```

```
SKU Code
42588924    454
42802945    290
42738503    153
42751991    141
42668145    137
...
42680437      1
42671190      1
42630166      1
42630029      1
38937620      1
Length: 213, dtype: int64
```

Below SKU Codes has more number of reviews within appliances department

1. 42588924
2. 42802945
3. 42738503
4. 42751991
5. 42668145

Products

SKU Code:
42588924



3.2 Litre Air Fryer
\$69.00

 1694 Reviews

SKU Code:
42802945



5.3 Litre Air Fryer
\$89.00

 1213 Reviews



BACK ONLINE SOON

3-in-1 Air Fryer Oven
\$129.00 SKU Code:
42738503

 428 Reviews

Products

SKU Code:
42751991



2000W Bagless Vacuum

\$47.00



390 Reviews



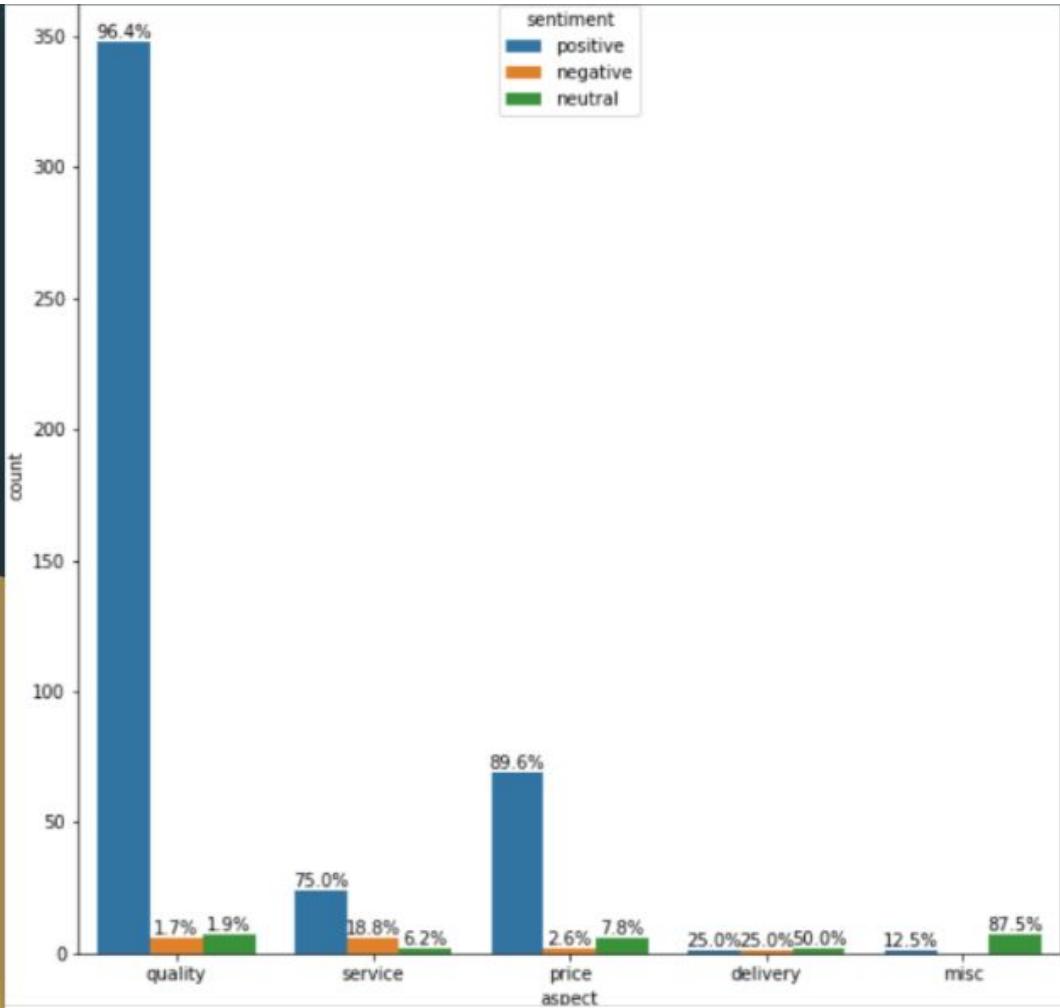
SKU Code: 42668145

34L Microwave

\$99.00



476 Reviews

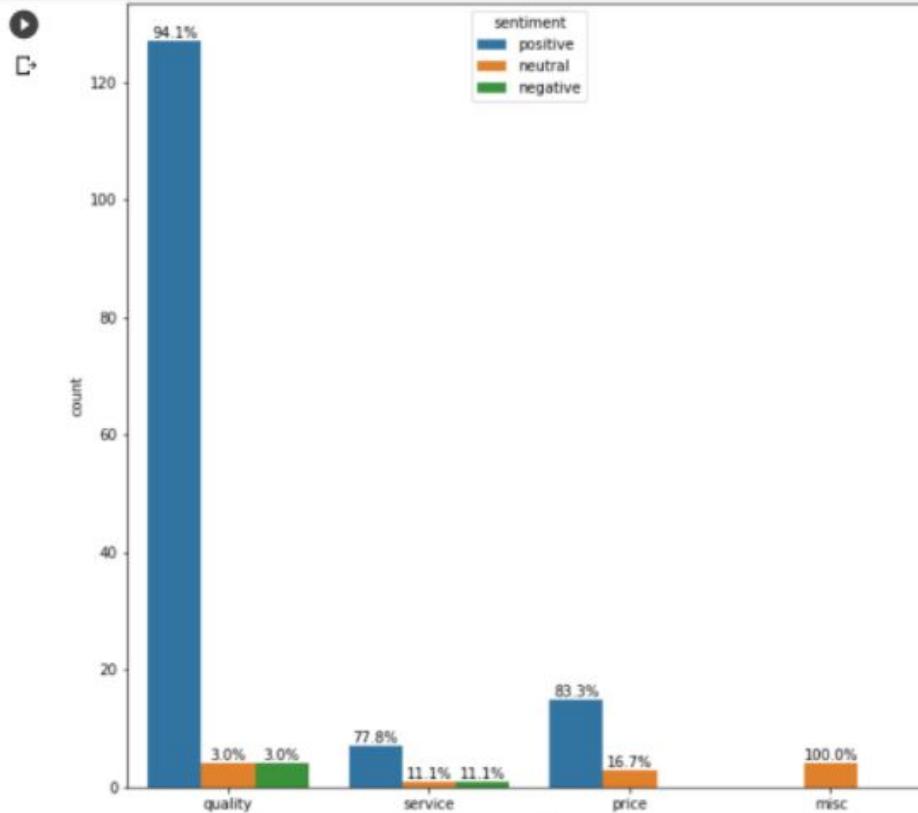


3.2 litre Air Fryer

Aspect Model with
percentage of aspects.

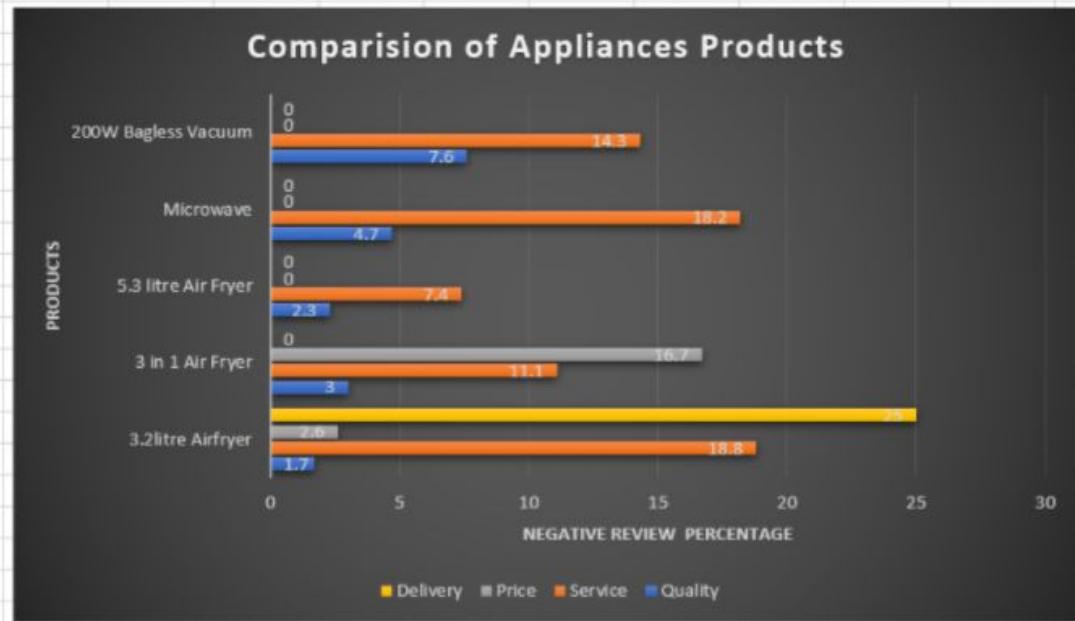


+ Code + Text



Percentage of Aspect Model Output for 3 in 1 Air Fryer

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1	Products	Quality	Service	Price	Delivery											
2	3.2 litre Airfryer	1.7	18.8	2.6	25											
3	3 in 1 Air Fryer	3	11.1	16.7	0											
4	5.3 litre Air Fryer	2.3	7.4	0	0											
5	Microwave	4.7	18.2	0	0											
6	200W Bagless Vacuum	7.6	14.3	0	0											



Introduction to FLASK

What is Flask?

- A Microframework written in Python which is used to create websites and api's to consume the model.
- Build an attractive User Interface for the model where we can provide the inputs for the model and see the model outputs.

Why Flask?

1. Easy to learn
2. Small but can easily handle large apps
3. Awesome documentation
4. Written in Python
5. Import only what you need

Flask Implementation

1. Implementation of a model that predicts the positive percentage of a review.
2. Using Sentiment Intensity Analyzer model, polarity score is determined.
3. Calculating the positive percentage (compound score) using the polarity score.
4. The compound score is the sum of positive, negative & neutral scores which is then normalized between -1(most extreme negative) and +1 (most extreme positive). The more Compound score closer to +1, the higher the positivity of the text.



Sentiment Analyzer

Review	Sentiment	Aspect
excellent product true size best maternity top i have owned	Positive	quality
terrible quality dinnerware seal neat plates bowls unfinished scratching table would not recommend	Negative	quality
good quality right sizes	Positive	quality
really bad jeans fades after wash	Negative	quality
very comfortable	Positive	quality
great price worth money	Positive	price
terrible quality screws and holding brackets worse	Negative	quality
contacting kmart problem solved quickly service	Positive	service
failed delivery	Negative	delivery
perfect pasta love shape	Positive	quality
disgusting service terrible experience disgusting system	Negative	service
amazing quality product	Positive	quality
as stated quick delivery	Negative	delivery
happy service product	Positive	service
mug huge made loved it	Positive	quality
delivery not as stated took time	Negative	delivery
not value for money product	Negative	price
terrible smashed plate	Negative	quality
thank god online purchase surrounding stores never size also pick chosen store perfect good quality good fit	Positive	quality
worst jug ever pouring cup water leaks	Negative	quality
christmas quality terrible embarrassing give gift legs came expecting	Negative	quality
really nice print style sits like sack does not look flattering good slouch around house hoping something could wear stroll	Positive	quality
bit tight but excellent value money worth price	Positive	price

Discussion

In future work, in Flask Deployment,

1. Reviews can be loaded through .csv file
2. Determine the graph for Sentiment analysis
3. Determine the Aspect level percentage

Thank you