

A photograph showing a long, parallel line of red shopping carts. The carts are made of a wireframe design and have black wheels. They are parked on a light-colored concrete or asphalt surface. The perspective is from the side, looking down the length of the row.

RETAIL ANALYTICS

Methodology to Find the Most Optimize Shelf Space Allocation, Shopping Cart, Customer Segmentation & Price Elasticity in Retail Industry

TEAM 3

SHELF SPACE ALLOCATION

Key challenges:

- Constantly growing number of available products.
- Retail space limitations (increasing number of stores).
- Short product life cycles (constant need to re-optimize).

Metrics:



- Width, Monthly_demand, Unit_margin, Profit from Display Products & Binary Indicator

MAXIMUM PROFIT

STEP 2

Technique: Excel Solver

Considering:

- Self width to be max. 3600 mm

Considering:

- Self width to be max. 3600 mm
- At least 3-product per brand

OBJECTIVE FUNCTION	
Max Profit	\$ 1,243.24

CONSTRAINTS	LHS	RHS	Profit
Width	3594.03234	<=	3600
Binary			
Brand 35	12		\$ 602.63
Brand 128	4		\$ 96.37
Brand 348	11		\$ 232.65
Brand 371	5		\$ 195.51
Brand 423	0		\$ -
Brand 424	5		\$ 116.09

OBJECTIVE FUNCTION	
Max Profit	\$ 1,205.51

CONSTRAINTS	LHS	RHS	Profit
Width	3540.47535	<=	3600
Binary			
Brand 35	12	>=	3 \$ 602.63
Brand 128	3	>=	3 \$ 76.97
Brand 348	8	>=	3 \$ 190.55
Brand 371	5	>=	3 \$ 195.51
Brand 423	3	>=	3 \$ 23.76
Brand 424	5	>=	3 \$ 116.09

COMPARISON FOR THE NEW PRODUCT IN RELATION TO PROFIT

STEP 3

Technique: Excel Solver

Considering:

- At least 8-product for new brand as additional constraint

With New Product (Brand 48)				
OBJECTIVE FUNCTION				
Max Profit	\$ 1,165.54			
CONSTRAINTS	LHS		RHS	Profit
Width	3598.61715	<=	3600	
Binary				
Brand 35	10	>=	3	\$ 569.89
Brand 128	3	>=	3	\$ 76.97
Brand 348	6	>=	3	\$ 151.83
Brand 371	4	>=	3	\$ 175.78
Brand 423	3	>=	3	\$ 24.71
Brand 424	3	>=	3	\$ 73.62
Brand 48 (new)	8	>=	8	\$ 92.74

Without				
OBJECTIVE FUNCTION				
Max Profit	\$ 1,205.51			
CONSTRAINTS	LHS		RHS	Profit
Width	3540.47535	<=	3600	
Binary				
Brand 35	12	>=	3	\$ 602.63
Brand 128	3	>=	3	\$ 76.97
Brand 348	8	>=	3	\$ 190.55
Brand 371	5	>=	3	\$ 195.51
Brand 423	3	>=	3	\$ 23.76
Brand 424	5	>=	3	\$ 116.09

CUSTOMER SEGMENTATION

RFM Metrics



RECENCY

The freshness of the customer activity, be it purchases or visits

E.g. Time since last order or last engaged with the product



FREQUENCY

The frequency of the customer transactions or visits

E.g. Total number of transactions or average time between transactions/engaged visits



MONETARY

The intention of customer to spend or purchasing power of customer

E.g. Total or average transactions value

OUR RFM STEPS USING EXCEL:

min_20_freq	20%
	6
min_30_freq	30
	4
min_30_spend	
	1280.991

	Recency	Freq	Monetary
Min	0	1	0
Q1	17	1	299.51
Q2	50	2	641.98
Q3	141	5	1572.235
Max	373	210	259657.3

Segment	Definition
3.67 - 4	LOYAL CUSTOMER (e.g: retiree, elderly)
3 - 3.33	LOCAL CUSTOMER (e.g: neighbourhood)
2.33 - 2.67	MODERATE CUSTOMER
1.67 - 2	TEMPORARY CUSTOMER (e.g: seasonal customers)
1 - 1.33	VISITING CUSTOMER (e.g: tourist)

STEP 1

DATA PRE-PROCESSING:

CLEAN DATA, DETERMINE THE PERCENTILE

STEP 2

DATA MASSAGING:

SET A SCORE FROM QUARTILE (1-4)

STEP 3

DATA INTERPRETATION:

CREATE CUSTOMER SEGMENTATION PERSONA

CUSTOMER SEGMENTATION FINDINGS:

- Average Spending for the past 2-week (top 20% frequency quantile): \$8,326.27
- Average Recency (top 30% quantile in both frequency and spending): 29.65 days
- Top-10 Most Valuable Customer:



CustomerID
17675
13777
12748
15311
18102
13798
13408
17677
14606
15189

Top 10 Most Valuable Customer

Factors that Affecting Price Elasticity:

- Availability of close substitutes
- Nature of goods (necessity, comfort or luxury)
- Distribution of income
- Time elapsed since a change in price
- Possibility of Postponement



Luxuries vs. Necessities



Elastic



Inelastic

PRICE ELASTICITY

Metrics: Excel Regression



STEP 4

$$Y^{\wedge} = a + bX$$

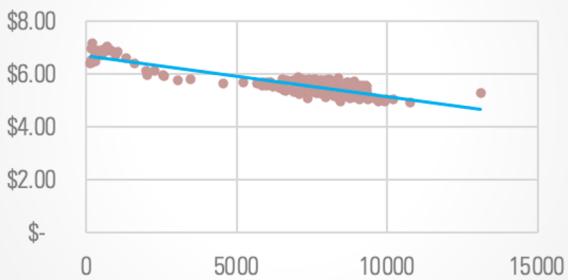
PRICE VS DEMANDS

	Coefficients
Intercept	35624
X Variable 1	-5122

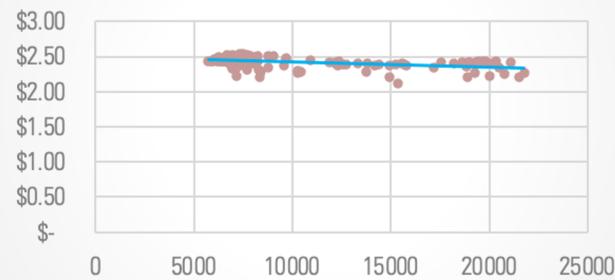
	Coefficients
Intercept	79627
X Variable 1	-28725

	Coefficients
Intercept	18912
X Variable 1	-3034

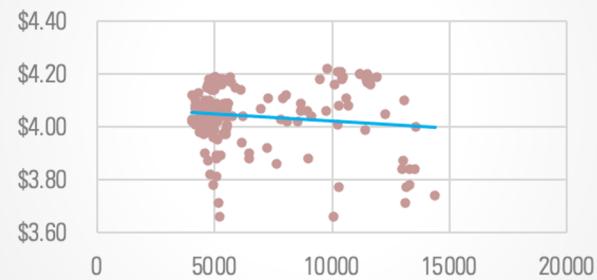
Demand Curve_P2



Demand Curve_P3



Demand Curve_P4



SHOPPING CART ANALYSIS USING PYTHON



Using python to sort the top purchased products



```
[ ]: # top items by purchase frequency: lower support threshold to find more items (longer runtime)
freq_items = apriori(df_trans, min_support=0.001, use_colnames=True)
freq_items.sort_values(by = "support", ascending = False).head(30)
```

STEP 1

	support	itemsets
24	0.255516	(whole milk)
22	0.193493	(other vegetables)
53	0.183935	(rolls/buns)
99	0.174377	(soda)
29	0.139502	(yogurt)
98	0.110524	(bottled water)
19	0.108998	(root vegetables)
14	0.104931	(tropical fruit)
156	0.098526	(shopping bags)
1	0.093950	(sausage)
56	0.088968	(pastry)
13	0.082766	(citrus fruit)
103	0.080529	(bottled beer)
151	0.079817	(newspapers)
104	0.077682	(canned beer)
15	0.075648	(pip fruit)
1298	0.074835	(other vegetables, whole milk)
101	0.072293	(fruit/vegetable juice)
30	0.071683	(whipped/sour cream)
55	0.064870	(brown bread)
52	0.063447	(domestic eggs)
0	0.058973	(frankfurter)
67	0.058566	(margarine)

STEP 2

Shopping Cart Product Association: Rolls/Buns

STEP 3

METRICS:

Lift = 5

Confidence = 0.5

Support = 0.01

ALGORITHM:

- Systematic Data Filtering
- Sort purchases by frequency

RESULT AND INSIGHTS:

The assumption is that rolls/buns are usually purchased with other goods such as frankfurters to make hot dogs, however our data tells us that there's a higher chance of rolls/buns being purchased along with vegetables and milk.

This proves that rolls/buns are a staple food item in many kitchens based on the sample data.

SHOPPING CART PRODUCT ASSOCIATION:

TOP 3 ASSOCIATIONS

	antecedents	consequents	antecedent support	consequent support	support	confidence	lift	leverage	conviction
15262	(tropical fruit, rolls/buns, other vegetables,...)	(root vegetables, whole milk)	0.001729	0.048907	0.001118	0.647059	13.230402	0.001034	2.694764
13642	(herbs, rolls/buns, other vegetables)	(root vegetables, whole milk)	0.001830	0.048907	0.001017	0.555556	11.359436	0.000927	2.139959
15263	(tropical fruit, rolls/buns, whole milk, beef)	(root vegetables, other vegetables)	0.002135	0.047382	0.001118	0.523810	11.055079	0.001017	2.000498

CONCLUSION:

- For shelf space allocation, sometimes less is more.
- The best benefits of custom segmentation will help the sales team to design the sales roadmap, understand most and least engaged customers at granular level and the best way to approach them.
- Price elasticity of demand is perhaps the most well-recognized, measuring how demand changes for an item if its price changes, with some goods' demand more sensitive to price than others.
- Shopping cart association will help businesses improve revenue through up-selling or cross-selling strategies, understanding these associations creates a more pleasant shopping experience for the consumer.

APPENDIX:

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
product_id	category_id	brand_id	width	height	depth	weight	onthly_dema	price	unit_margin	FACING/ NO FACING	PROFIT						
21824	212	35	116.022102	90.8532032	104.057607	1.15744577	7.66666667	14.973333	7.591111	1	58.1985177						
101120	212	35	116.095343	84.7268624	121.180416	0.86335293	6.33333333	14.44898	10.785	1	68.305						
28643	212	35	80.3414269	113.268308	103.57429	1.05172569	14.3333333	14.939252	8.326324	1	119.343977						
24214	212	35	121.956231	98.9439591	89.9113464	0.97959534	4.33333333	15.390004	6.425221	1	27.8426243						
32858	212	35	103.884262	83.3847328	92.34091	0.92907354	10	15.392857	8.829384	1	88.29384						
113790	212	35	88.7958955	118.195918	104.037858	0.96720841	2.33333333	16.524194	7.798387	0	0						
103984	212	35	99.5905005	76.319076	76.6823999	1.07100209	3.66666667	9.235294	9.235294	1	33.8627447						
32857	212	35	123.952506	112.410127	76.6519186	0.82643822	2.66666667	16.25	10.642857	1	28.380952						
30713	212	35	78.2055211	97.8636106	97.8636106	0.97863611	1	33.987045	14.547045	0	0						
101119	212	35	76.9939594	116.584737	123.057061	0.7672574	5.66666667	15.7666667	11.722222	1	66.4259247						
103983	212	35	101.898435	87.0467751	84.7268624	1.21180416	1.66666667	9.235294	8.974265	0	0						
21825	212	35	117.070344	118.7645475	102.134818	0.80476725	7	15.120002	5.557501	1	38.902507						
32855	212	35	102.407891	90.7717594	75.5730123	0.97241096	5.33333333	14.696272	7.5625	1	40.3333333						
113792	212	35	85.3268939	113.532961	106.145085	0.86041981	1.33333333	15.924051	7.639241	0	0						
19492	212	128	116.533539	83.3107218	112.95876	0.90318357	29.3333333	1.077037	0.593704	0	0						

OBJECTIVE FUNCTION
Max Profit \$ 1,165.54

CONSTRAINTS	LHS	RHS	Profit
Width	3598.61715	<=	3600
Binary			
Brand 35	10	>=	\$ 569.89
Brand 128	3	>=	\$ 76.97
Brand 348	6	>=	\$ 151.83
Brand 371	4	>=	\$ 175.78
Brand 423	3	>=	\$ 24.71
Brand 424	3	>=	\$ 73.62
Brand 48 (new)	8	>=	\$ 92.74

Shelf Space including New Product (Constraints Applied)

product_id	category_id	brand_id	width	height	depth	weight	onthly_dema	price	unit_margin	FACING/ NO FACING	PROFIT						
21824	212	35	116.0221	90.853203	104.05761	1.1574458	7.66666667	14.973333	7.591111	1	58.198518						
101120	212	35	116.09534	84.726862	121.18042	0.8633529	6.33333333	14.44898	10.785	1	68.305						
28643	212	35	80.34142	113.26831	103.57429	1.0517257	14.3333333	14.939252	8.326324	1	119.34398						
24214	212	35	121.95623	98.943959	89.911346	0.97959534	4.33333333	15.390004	6.425221	1	27.842624						
32858	212	35	103.88426	83.38473	92.34091	0.9290735	10	15.392857	8.829384	1	88.29384						
113790	212	35	88.795896	118.19592	104.03786	0.9672084	2.33333333	16.524194	7.798387	1	18.196236						
103984	212	35	99.59050	76.319076	76.6824	1.0710021	3.66666667	9.235294	9.235294	1	33.862745						
32857	212	35	123.95251	112.41013	76.651919	0.8264382	2.66666667	16.25	10.642857	1	28.380952						
30713	212	35	78.205521	97.863611	97.863611	0.9786361	1	33.987045	14.547045	1	14.547045						
101119	212	35	76.993959	116.584737	123.05706	0.7672574	5.66666667	15.7666667	11.722222	1	66.425925						
103983	212	35	101.898435	87.0467751	84.726862	1.21180416	1.66666667	9.235294	8.974265	0	0						
21825	212	35	117.070344	118.7645475	102.134818	0.80476725	7	15.120002	5.557501	1	38.902507						
32855	212	35	102.407891	90.7717594	75.573012	0.97241096	5.33333333	14.696272	7.5625	1	40.3333333						
113792	212	35	85.3268939	113.532961	106.145085	0.86041981	1.33333333	15.924051	7.639241	0	0						

OBJECTIVE FUNCTION
Max Profit \$ 1,205.51

CONSTRAINTS	LHS	RHS	Profit
Width	3540.4753	<=	3600
Binary			
Brand 35	12	>=	\$ 602.63
Brand 128	3	>=	\$ 76.97
Brand 348	8	>=	\$ 190.55
Brand 371	5	>=	\$ 195.51
Brand 423	3	>=	\$ 23.76
Brand 424	5	>=	\$ 116.09

Shelf Space without New Product (Constraints Applied)

product_id	category_id	brand_id	width	height	depth	weight	onthly_dema	price	unit_margin	FACING/ NO FACING	PROFIT			
21824	212	35	116.022102	90.8532032	104.057607	1.15744577	7.66666667	14.973333	7.591111	1	58.1985177			
101120	212	35	116.095343	84.7268624	121.180416	0.86335293	6.33333333	14.44898	10.785	1	68.305			
28643	212	35	80.3414269	113.268308	103.57429	1.05172569	14.33333333	14.939252	8.326324	1	119.343977			
24214	212	35	121.956231	98.9439591	89.9113464	0.97959534	4.33333333	15.390004	6.425221	1	27.8426243			
32858	212	35	103.884262	83.3847328	92.34091	0.92907354	10	15.392857	8.829384	1	88.29384			
113790	212	35	88.7958955	118.195918	104.037858	0.96720841	2.33333333	16.524194	7.798387	1	18.1962363			
103984	212	35	99.5905005	76.319076	76.6823999	1.07100209	3.66666667	9.235294	9.235294	1	33.8627447			
32857	212	35	123.952506	112.410127	76.6519186	0.82643822	2.66666667	16.25	10.642857	1	28.380952			
30713	212	35	78.2055211	97.8636106	97.8636106	0.97863611	1	33.987045	14.547045	1	14.547045			
101119	212	35	76.9939594	116.584737	123.057061	0.7672574	5.66666667	15.766667	11.722222	1	66.4259247			
103983	212	35	101.898435	87.0467751	84.7268624	1.21180416	1.66666667	9.235294	8.974265	0	0			
21825	212	35	117.070344	118.764574	102.134818	0.80476725	7	15.120002	5.557501	1	38.902507			
32855	212	35	102.407891	90.7717594	75.5730123	0.97241096	5.33333333	14.696272	7.5625	1	40.3333333			
113792	212	35	85.3268939	113.532961	106.145085	0.86041981	1.33333333	15.924051	7.639241	0	0			

Shelf Space without New Product (Only Width Constraints Applied)

CustomerID	recency	frequency	monetary	min_20_freq	20%
13874	358	1	540.45	6	
17423	46	2	698.41		
16555	35	2	285.89		
15469	50	2	538.55	min_30_freq	30
16477	63	7	2494.46	4	
17935	138	2	145.79		
17094	322	1	302	min_30_spend	
16455	63	7	1841.18	1280.991	
14662	28	4	741.37		

Customer Segmentation (Quantile 20% and 30%)

CustomerID	recency	frequency	monetary	R_SCORE	F_SCORE	M_SCORE	TOTAL_SCORE	SEGMENT_STATUS	Recency	Freq	Monetary	Segment	Definition	
13576	10	18	6717.34	4.00	4.00	4.00	4.00	LOYAL CUSTOMER	Min	0	1	0	3.67 - 4	LOYAL CUSTOMER (e.g.: retiree, elderly)
13985	4	15	7060.13	4.00	4.00	4.00	4.00	LOYAL CUSTOMER	Q1	17	1	299.51	3 - 3.33	LOCAL CUSTOMER (e.g.: neighbourhood)
16326	5	14	3039.22	4.00	4.00	4.00	4.00	LOYAL CUSTOMER	Q2	50	2	641.98	2.33 - 2.67	MODERATE CUSTOMER
17675	1	31	20374.28	4.00	4.00	4.00	4.00	LOYAL CUSTOMER	Q3	141	5	1572.235	1.67 - 2	TEMPORARY CUSTOMER (e.g: seasonal customers)
17243	1	26	8540.05	4.00	4.00	4.00	4.00	LOYAL CUSTOMER	Max	373	210	259657.3	1 - 1.33	VISITING CUSTOMER (e.g: tourist)
13319	5	26	10845.55	4.00	4.00	4.00	4.00	LOYAL CUSTOMER						
16210	1	18	21086.3	4.00	4.00	4.00	4.00	LOYAL CUSTOMER						
15218	10	11	5756.89	4.00	4.00	4.00	4.00	LOYAL CUSTOMER						
13709	3	13	6307.65	4.00	4.00	4.00	4.00	LOYAL CUSTOMER						
15251	9	11	10484.99	4.00	4.00	4.00	4.00	LOYAL CUSTOMER						
15046	9	15	5090.42	4.00	4.00	4.00	4.00	LOYAL CUSTOMER						
17644	1	11	2895.64	4.00	4.00	4.00	4.00	LOYAL CUSTOMER						

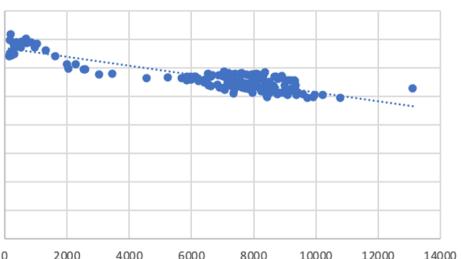
Customer Segmentation (Including Segment Persona for Customers)

CustomerID	recency	frequency	monetary	R_SCORE	F_SCORE	M_SCORE	TOTAL_SCORE	SEGMENT_STATUS
17675	1	31	20374.28	4.00	4.00	4.00	4.00	TOP 10 LOYAL CUSTOMER
13777	0	33	25977.16	4.00	4.00	4.00	4.00	TOP 10 LOYAL CUSTOMER
12748	0	210	33053.19	4.00	4.00	4.00	4.00	TOP 10 LOYAL CUSTOMER
15311	0	91	60632.75	4.00	4.00	4.00	4.00	TOP 10 LOYAL CUSTOMER
18102	0	60	259657.3	4.00	4.00	4.00	4.00	TOP 10 LOYAL CUSTOMER
13798	1	57	37153.85	4.00	4.00	4.00	4.00	TOP 10 LOYAL CUSTOMER
13408	1	62	28117.04	4.00	4.00	4.00	4.00	TOP 10 LOYAL CUSTOMER
17677	1	30	16345.38	4.00	4.00	4.00	4.00	TOP 10 LOYAL CUSTOMER
14606	1	93	12076.15	4.00	4.00	4.00	4.00	TOP 10 LOYAL CUSTOMER
15189	1	41	16225.39	4.00	4.00	4.00	4.00	TOP 10 LOYAL CUSTOMER

Top 10 Most Loyal Customers

Q_P2	price_P2	revenue_P2	"-C2"	"C1"	P_Average	Q_Average	Elasticity	Cost_P1	P_Optimal
8240.77	\$ 5.17	42604.78		-5121.88	35623.56	5.67	6579.87	4.41	3.90
7104.73	\$ 5.33	37868.21							
6554.08	\$ 5.39	35326.49							
6901.22	\$ 5.32	36714.49							
6624.8	\$ 5.37	35575.18							
5855.77	\$ 5.56	32558.08							
5975.24	\$ 5.70	34058.87							
7160.74	\$ 5.52	39527.28							
8001.87	\$ 5.37	42970.04							
7347.32	\$ 5.37	39455.11							
7725.72	\$ 5.28	40791.80							
7634.54	\$ 5.30	40463.06							
7348.96	\$ 5.40	39684.38							
9342.85	\$ 5.38	50264.53							
6218.2	\$ 5.54	34448.83							
6833.65	\$ 5.41	36970.05							
6540.92	\$ 5.44	35582.60							

Demand Curve_P2



Product 2 Price Elasticity and Optimal Price

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.894947
R Square	0.800930133
Adjusted R S	0.799637471
Standard Err:	1325.208996
Observations	156

ANOVA

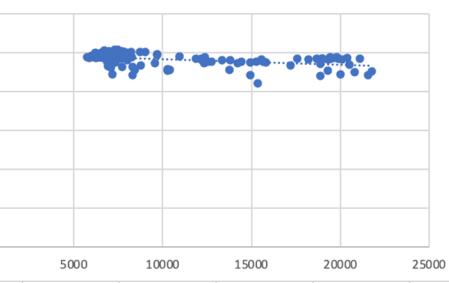
	df	SS	MS	F	Significance F
Regression	1	1088124473	1088124473	619.597744	7.5581E-56
Residual	154	270451547.9	1756178.88		
Total	155	1358576021			

	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%
Intercept	35623.56311	1171.615671	30.4055024	5.6569E-67	33309.0503	37938.0759	33309.0503	37938.0759
X Variable 1	-5121.881816	205.7664853	-24.89172	7.5581E-56	-5528.3711	-4715.3926	-5528.3711	-4715.3926

Regression Analysis for Product 2

Q_P3	price_P3	revenue_P3	"-C2"	"C1"	P_Average	Q_Average	Elasticity	Cost_P1	P_Optimal
7428.31	\$ 2.41	17902.23		-28725.14	79627.37	2.42	10167.77	6.83	1.8
7436.82	\$ 2.43	18071.47							2.29
6967.51	\$ 2.43	16931.05							
6927.91	\$ 2.44	16904.10							
6757.23	\$ 2.45	16555.21							
7163.16	\$ 2.41	17263.22							
7233.72	\$ 2.40	17360.93							
7098.25	\$ 2.41	17106.78							
6822.62	\$ 2.43	16578.97							
6730.52	\$ 2.44	16422.47							
6834.72	\$ 2.43	16608.37							
6803.58	\$ 2.43	16532.70							
7083.52	\$ 2.42	17142.12							
7997.06	\$ 2.41	19272.91							
6686.17	\$ 2.44	16314.25							
6013.81	\$ 2.46	14793.97							
6444.45	\$ 2.44	15724.46							
6210.60	\$ 2.44	14422.00							

Demand Curve_P3



Product 3 Price Elasticity and Optimal Price

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.448633094
R Square	0.201271653
Adjusted R S	0.196085106
Standard Err	4331.600804
Observations	156

ANOVA

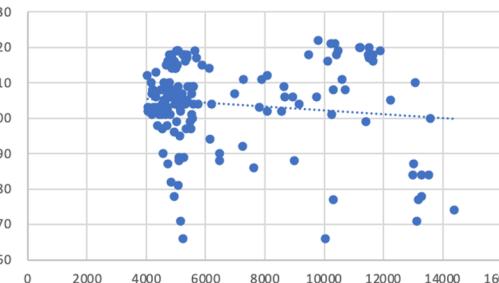
	df	SS	MS	F	Significance F
Regression	1	728116861.1	728116861	38.8064787	4.2647E-09
Residual	154	2889465891	18762765.5		
Total	155	3617582752			

	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%
Intercept	79627.36507	11155.52667	7.13792969	3.4985E-11	57589.7553	101664.975	57589.7553	101664.975
X Variable 1	-28725.13745	4611.157932	-6.2294846	4.2647E-09	-37834.425	-19615.85	-37834.425	-19615.85

Regression Analysis for Product 3

Q_P4	price_P4	revenue_P3	"-C2"	"C1"	P_Average	Q_Average	Elasticity	Cost_P1	P_Optimal
4736.72	\$ 4.03	19088.98	-3033.63	18912.20	4.04	6656.32	1.84	2.9	4.57
4525.64	\$ 4.02	18193.07							
4434.44	\$ 4.01	17782.10							
4477.84	\$ 4.02	18000.92							
4371.58	\$ 3.98	17398.89							
4647.6	\$ 4.01	18636.88							
4539.41	\$ 4.02	18248.43							
4465.96	\$ 4.02	17953.16							
4348.35	\$ 4.03	17523.85							
4277.81	\$ 4.02	17196.80							
4509.82	\$ 4.01	18084.38							
4469.84	\$ 4.02	17968.76							
4690.94	\$ 3.98	18669.94							
5497.87	\$ 3.97	21826.54							
4189.18	\$ 4.02	16840.50							
4824.51	\$ 3.82	18429.63							
4720.50	\$ 3.87	19207.77							

Demand Curve_P4



Product 4 Price Elasticity and Optimal Price

SUMMARY OUTPUT									
Regression Statistics									
Multiple R	0.12744296								
R Square	0.01624171								
Adjusted R S	0.00985367								
Standard Err	2835.33976								
Observations	156								
ANOVA									
	df	SS	MS	F	Significance F				
Regression	1	20439686.3	20439686.3	2.54251785	0.11286821				
Residual	154	1238029340	8039151.56						
Total	155	1258469026							
	Coefficients	standard Error	t Stat	P-value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%	
Intercept	18912.196	7689.56065	2.45946379	0.0150209	3721.56075	34102.8312	3721.56075	34102.8312	
X Variable 1	-3033.6316	1902.527	-1.5945275	0.11286821	-6792.0509	724.787807	-6792.0509	724.787807	

Regression Analysis for Product 4

```
In [ ]: rules = association_rules(freq_items, metric="confidence", min_threshold=0.5)
rules
```

Out[]:

	antecedents	consequents	antecedent support	consequent support	support	confidence	lift	leverage	conviction
0	(specialty cheese)	(other vegetables)	0.008541	0.193493	0.004270	0.500000	2.584078	0.002618	1.613015
1	(rice)	(other vegetables)	0.007626	0.193493	0.003965	0.520000	2.687441	0.002490	1.680224
2	(rice)	(whole milk)	0.007626	0.255516	0.004677	0.613333	2.400371	0.002729	1.925390
3	(honey)	(whole milk)	0.001525	0.255516	0.001118	0.733333	2.870009	0.000729	2.791815
4	(cereals)	(whole milk)	0.005694	0.255516	0.003660	0.642857	2.515917	0.002206	2.084555
...
5824	(domestic eggs, butter, whole milk, yogurt, tr...	(other vegetables)	0.001220	0.193493	0.001017	0.833333	4.306796	0.000781	4.839044
5825	(domestic eggs, butter, other vegetables, yogurt)	(tropical fruit, whole milk)	0.001729	0.042298	0.001017	0.588235	13.906957	0.000944	2.325848
5826	(tropical fruit, domestic eggs, butter, other ...)	(whole milk, yogurt)	0.001729	0.056024	0.001017	0.588235	10.499626	0.000920	2.292512
5827	(tropical fruit, domestic eggs, butter, whole ...)	(other vegetables, yogurt)	0.001830	0.043416	0.001017	0.555556	12.795993	0.000937	2.152313
5828	(tropical fruit, domestic eggs, butter, yogurt)	(other vegetables, whole milk)	0.001322	0.074835	0.001017	0.769231	10.279055	0.000918	4.009049

5829 rows × 9 columns

```
In [ ]: rules = association_rules(freq_items, metric = "lift", min_threshold = 5 )
rules
```

Out[]:

	antecedents	consequents	antecedent support	consequent support	support	confidence	lift	leverage	conviction
0	(ham)	(processed cheese)	0.026029	0.016573	0.003050	0.117188	7.070792	0.002619	1.113970
1	(processed cheese)	(ham)	0.016573	0.026029	0.003050	0.184049	7.070792	0.002619	1.193663
2	(hamburger meat)	(pasta)	0.033249	0.015048	0.002745	0.082569	5.486920	0.002245	1.073597
3	(pasta)	(hamburger meat)	0.015048	0.033249	0.002745	0.182432	5.486920	0.002245	1.182473
4	(Instant food products)	(hamburger meat)	0.008033	0.033249	0.003050	0.379747	11.421438	0.002783	1.558640
...
15787	(tropical fruit, yogurt)	(domestic eggs, butter, other vegetables, whol...	0.029283	0.003050	0.001017	0.034722	11.383102	0.000927	1.032811
15788	(domestic eggs)	(other vegetables, butter, whole milk, yogurt,...	0.063447	0.002339	0.001017	0.016026	6.852703	0.000868	1.013910
15789	(butter)	(domestic eggs, other vegetables, whole milk, ...)	0.055414	0.001627	0.001017	0.018349	11.278670	0.000927	1.017034
15790	(yogurt)	(domestic eggs, other vegetables, butter, whol...	0.139502	0.001322	0.001017	0.007289	5.514129	0.000832	1.006011
15791	(tropical fruit)	(domestic eggs, other vegetables, butter, whol...	0.104931	0.001322	0.001017	0.009690	7.330799	0.000878	1.008450

15792 rows × 9 columns

```
In [ ]: # top items by purchase frequency: lower support threshold to find more items (longer runtime)
freq_items = apriori(df_trans, min_support=0.001, use_colnames=True)
freq_items.sort_values(by = "support", ascending = False).head(30)
```

Out[]:

	support	itemsets
24	0.255516	(whole milk)
22	0.193493	(other vegetables)
53	0.183935	(rolls/buns)
99	0.174377	(soda)
29	0.139502	(yogurt)
98	0.110524	(bottled water)
19	0.108998	(root vegetables)
14	0.104931	(tropical fruit)
156	0.098526	(shopping bags)
1	0.093950	(sausage)
56	0.088968	(pastry)
13	0.082766	(citrus fruit)
103	0.080529	(bottled beer)
151	0.079817	(newspapers)
104	0.077682	(canned beer)
15	0.075648	(pip fruit)
1298	0.074835	(other vegetables, whole milk)
101	0.072293	(fruit/vegetable juice)
30	0.071683	(whipped/sour cream)
55	0.064870	(brown bread)
52	0.063447	(domestic eggs)
0	0.058973	(frankfurter)
67	0.058566	(margarine)

```
In [ ]: #This is to sort top 3:  
rules_sel.sort_values('lift', ascending=False).head(10)
```

Out[]:

	antecedents	consequents	antecedent support	consequent support	support	confidence	lift	leverage	conviction
15262	(tropical fruit, rolls/buns, other vegetables,...	(root vegetables, whole milk)	0.001729	0.048907	0.001118	0.647059	13.230402	0.001034	2.694764
13642	(herbs, rolls/buns, other vegetables)	(root vegetables, whole milk)	0.001830	0.048907	0.001017	0.555556	11.359436	0.000927	2.139959
15263	(tropical fruit, rolls/buns, whole milk, beef)	(root vegetables, other vegetables)	0.002135	0.047382	0.001118	0.523810	11.055079	0.001017	2.000498
9665	(tropical fruit, rolls/buns, beef)	(root vegetables, whole milk)	0.002745	0.048907	0.001423	0.518519	10.602141	0.001289	1.975347
15246	(root vegetables, other vegetables, whole milk...)	(beef)	0.002034	0.052466	0.001118	0.550000	10.483043	0.001012	2.105632
9877	(butter, beef, rolls/buns)	(root vegetables, whole milk)	0.002237	0.048907	0.001118	0.500000	10.223493	0.001009	1.902186
5957	(herbs, rolls/buns)	(root vegetables, whole milk)	0.003050	0.048907	0.001525	0.500000	10.223493	0.001376	1.902186
15258	(tropical fruit, root vegetables, beef, rolls/...)	(other vegetables, whole milk)	0.001729	0.074835	0.001118	0.647059	8.646499	0.000989	2.621301
10891	(pastry, whole milk, rolls/buns, whipped/sour ...	(citrus fruit)	0.001525	0.082766	0.001017	0.666667	8.054873	0.000891	2.751703
13639	(herbs, root vegetables, rolls/buns)	(other vegetables, whole milk)	0.001830	0.074835	0.001017	0.555556	7.423762	0.000880	2.081622

REFERENCES:

- <https://www.intercom.com/blog/customer-segmentation/>
- <https://www.economicsdiscussion.net/elasticity-of-demand/9-factors-that-influence-price-elasticity-of-demand/3493>