MIND OVER DATA

Retail Challenge

Group O

Challenges

DATASET TRANSFORMATION

From 27 Gb to 4Gb

POS QUARTERLY ANALYSIS

- Top Products by units sold
- Market Share (Family, Categories)
- Product co-occurrencies

CLUSTERING

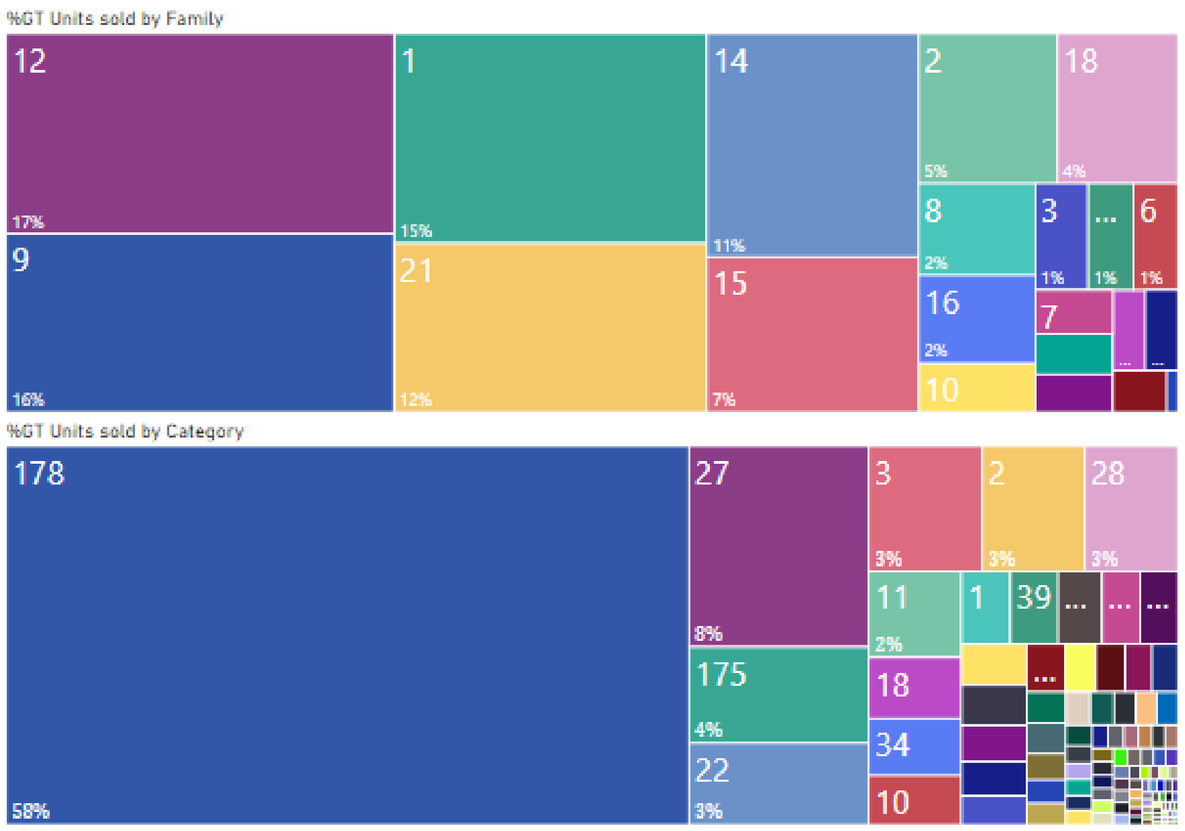
- Stores Clustering
- Products Clustering

FORECAST

- Units sold
- Point of sale

Quarterly Analysis Example — STORE 292 QUARTER 1 2019

MARKET SHARE



TOP 10 PRODUCTS

Top 10 Products Sold

Family	Category	Brand	Name	Units sold
14	28	574	1147	2633
15	2	487	993	2244
14	27	226	481	2040
14	27	618	1234	1851
14	27	1266	2372	1749
9	178	122	226	1536
12	178	1410	2609	1499
14	27	67	130	1497
8	178	1509	2802	1265
9	178	732	1408	1121
Total				**17435**

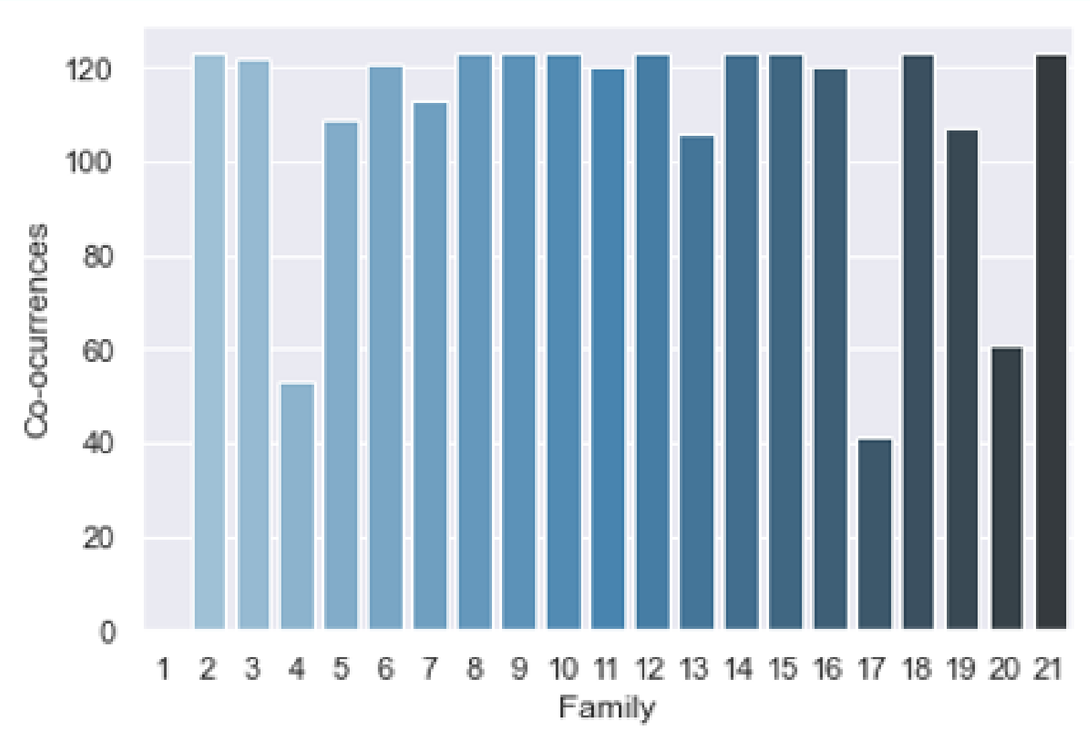
Quarterly Analysis Example

STORE 292
QUARTER 1
2019

PRODUCT CO-OCCURRENCES

Which Families of products were mostly bought with Family 1?

(Family	
1	0.0
2	123.0
3	122.0
4	53.0
5	109.0
6	121.0
7	113.0
8	123.0
9	123.0
10	123.0
11	120.0
12	123.0
13	106.0
14	123.0
15	123.0
16	120.0
17	41.0
18	123.0
19	107.0
20	61.0
21	123.0



Quarterly Analysis Example

PRODUCT CO-OCCURRENCES

Creation of association rules

First-quarter association rules example

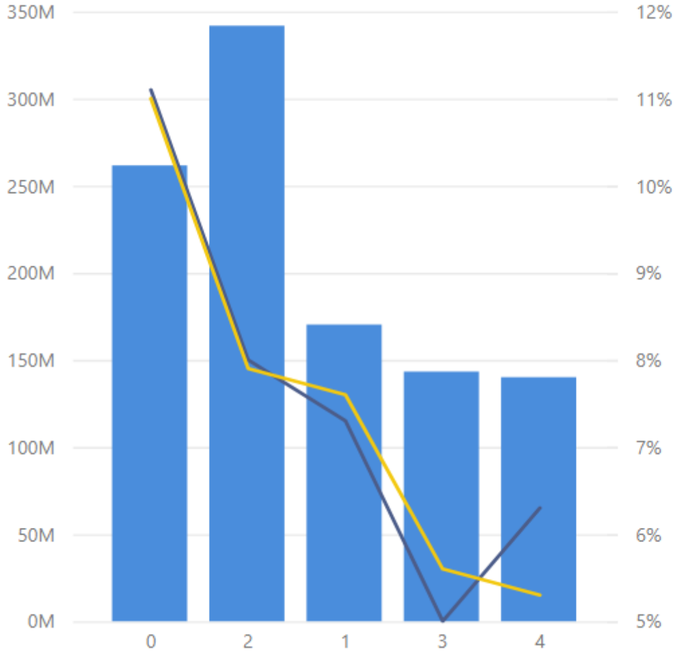
antecedents	consequents	antecedent support	consequent support	support	confidence	lift
(3, 8, 9, 12, 15, 16, 21)	(1, 2, 5, 14, 18)	0.82	0.74	0.67	0.82	1.1
(1, 2, 5, 14, 18)	(3, 8, 9, 12, 15, 16, 21)	0.74	0.82	0.67	0.90	1.1
(3, 8, 9, 12, 16, 21)	(1, 2, 5, 14, 15, 18)	0.82	0.74	0.67	0.82	1.1
(1, 3, 8, 9, 12, 16, 21)	(2, 5, 14, 15, 18)	0.82	0.74	0.67	0.82	1.1
(3, 8, 9, 12, 16, 21)	(2, 5, 14, 15, 18)	0.82	0.74	0.67	0.82	1.1
(2, 5, 14, 15, 18)	(1, 3, 8, 9, 12, 16, 21)	0.74	0.82	0.67	0.90	1.1
(1, 2, 5, 14, 15, 18)	(3, 8, 9, 12, 16, 21)	0.74	0.82	0.67	0.90	1.1
(2, 5, 14, 15, 18)	(3, 8, 9, 12, 16, 21)	0.74	0.82	0.67	0.90	1.1
(1, 3, 8, 9, 12, 15, 16, 21)	(2, 18, 5, 14)	0.82	0.74	0.67	0.82	1.1
(2, 18, 5, 14)	(3, 8, 9, 12, 15, 16, 21)	0.74	0.82	0.67	0.90	1.1

Clustering

ON STORE, CATEGORY AND COMBINED

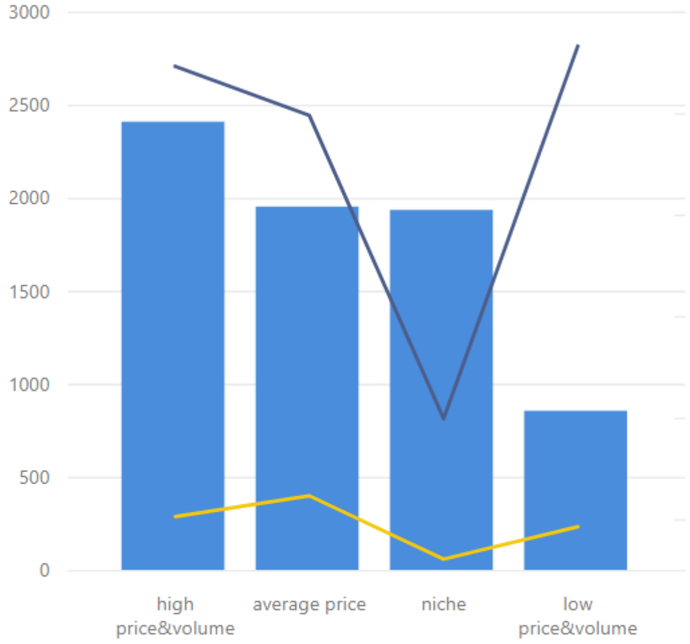
count stores per cluster

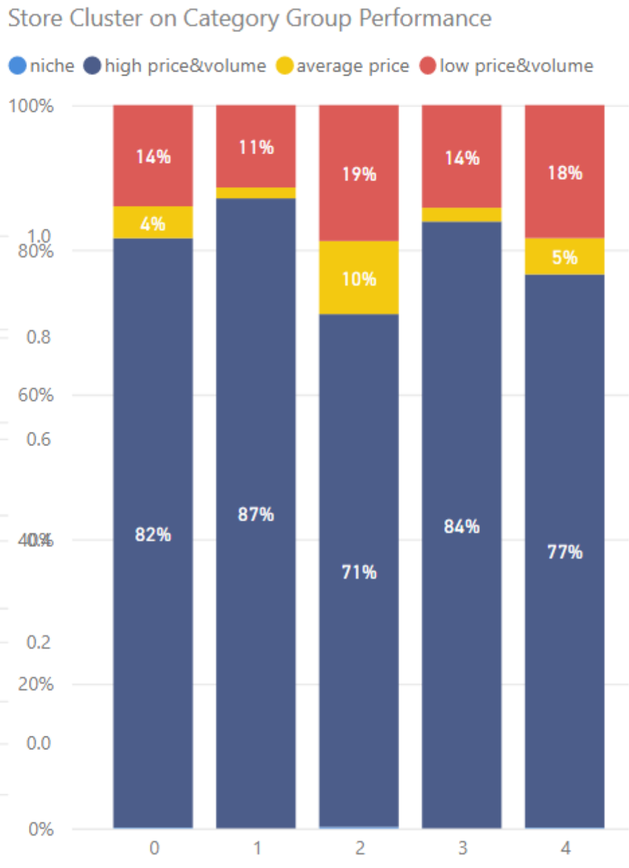
0	1	2	3	4	Total
62	97	45	119	85	**408**



count categories per cluster

0	1	2	3	Total
44	65	31	22	**162**





Forecasting

TRAIN From 2017 to June 2019

TEST July 2019 - October 2019

MODELS Theta or ETS based on R^2

1.FOR EACH PRODUCT NAME

79% couldn't be correctly predicted we used category or family forecast instead

2.FOR EACH STORE

48% Theta Model
36% ETS

3. COMBINED

Combination of product's contribution for each store and the product forecast

=

product forecast at store level

Thank you!

GROUP O