



ΙΟΝΙΟ ΠΑΝΕΠΙΣΤΗΜΙΟ
ΤΜΗΜΑ ΠΛΗΡΟΦΟΡΙΚΗΣ

ΠΤΥΧΙΑΚΗ ΕΡΓΑΣΙΑ

ΘΕΜΑ: «Εξόρυξη προσωποποιημένης πληροφόρησης από περιβάλλοντα super-market με σκοπό την ανάλυση αγορών προϊόντων»

ΣΠΟΥΔΑΣΤΡΙΑ: Παπααρηγορίου Ελισσάβετ

ΕΠΟΠΤΕΥΩΝ ΚΑΘΗΓΗΤΗΣ: Μυλωνάς Φοίβος

ΑΘΗΝΑ 2020

**Στην οικογένεια μου
και στους φίλους μου Μάνο, Έλενα και Νίκο.**

.....

.....

ΔΗΛΩΣΗ ΠΕΡΙ ΜΗ ΛΟΓΟΚΛΟΠΗΣ

Δηλώνω ότι είμαι ο η συγγραφέας της παρούσας εργασίας με τίτλο «Εξόρυξη προσωποποιημένης πληροφόρησης από περιβάλλοντα super-market με σκοπό την ανάλυση αγορών προϊόντων». Η αναφερόμενη εργασία δεν αποτελεί αντιγραφή ούτε προέρχεται από ανάθεση σε τρίτους. Οι πηγές που χρησιμοποιήθηκαν αναφέρονται σαφώς στη βιβλιογραφία ενώ κάθε εξωτερική βοήθεια, αναγνωρίζεται ρητά.

ΕΛΙΣΣΑΒΕΤ ΠΑΠΑΓΡΗΓΟΡΙΟΥ - 27/08/2020

ΕΥΧΑΡΙΣΤΙΕΣ

Θα ήθελα να ευχαριστήσω την οικογένειά μου Πέτρο, Θωμούλα και Μανώλη για όλα όσα μου έχουν προσφέρει στη διάρκεια των μαθητικών και φοιτητικών μου χρόνων και για την στήριξη τους σε κάθε μου επιλογή.

Θα ήθελα επιπλέον να ευχαριστήσω τους φίλους και συμφοιτητές μου Μάνο, Έλενα και Νίκο για την πολύτιμη και ουσιαστική υποστήριξή τους και συμπαράστασή τους τόσο στην εκπόνηση της παρούσας εργασίας αλλά και για όλη τη διάρκεια των φοιτητικών μου χρόνων.

Τέλος, Θα ήθελα να ευχαριστήσω τον επιβλέποντα καθηγητή μου Ανδρέα Καναβό για τον πολύτιμο χρόνο που αφιέρωσε στην παρούσα εργασία καθώς και για την ουσιαστική βοήθεια και καθοδήγηση που μου παρείχε.

ΠΕΡΙΕΧΟΜΕΝΑ

ΚΕΦΑΛΑΙΟ 1^ο - ΕΙΣΑΓΩΓΙΚΑ

1.1. ΕΙΣΑΓΩΓΗ.....σελ.8

ΚΕΦΑΛΑΙΟ 2^ο – ΘΕΩΡΗΤΙΚΟ ΠΛΑΙΣΙΟ

2.1. ΑΓΟΡΕΣ.....σελ.9

2.2. ΜΑΡΚΕΤΙΝΓΚ.....σελ.10

2.3. DATA ANALYSIS.....σελ.11

ΚΕΦΑΛΑΙΟ 3^ο – ΜΕΘΟΔΟΛΟΓΙΑ

3.1. ΜΕΘΟΔΟΛΟΓΙΑ.....σελ.12

ΚΕΦΑΛΑΙΟ 4^ο – ΠΙΝΑΚΕΣ ΠΩΛΗΣΕΩΝ

4.1. ΠΙΝΑΚΑΣ ΣΥΝΟΛΙΚΩΝ ΠΩΛΗΣΕΩΝ.....σελ.20
--

4.2. ΠΙΝΑΚΑΣ ΜΟΝΑΔΙΚΩΝ ΠΩΛΗΣΕΩΝ.....σελ.23
--

ΚΕΦΑΛΑΙΟ 5^ο – ΣΥΝΔΙΑΣΤΙΚΕΣ ΑΓΟΡΕΣ ΠΡΟΙΟΝΤΩΝ

5.1. ΣΥΝΔΙΑΣΤΙΚΕΣ ΑΓΟΡΕΣ – ΠΕΡΙΣΣΟΤΕΡΕΣ ΠΩΛΗΣΕΙΣ σελ.25
--

5.2. ΣΥΝΔΙΑΣΤΙΚΕΣ ΑΓΟΡΕΣ – ΛΙΓΟΤΕΡΕΣ ΠΩΛΗΣΕΙΣ σελ.52

ΚΕΦΑΛΑΙΟ 6^ο – ΕΠΙΛΟΓΟΣ

6.1. ΓΕΝΙΚΕΣ ΠΑΡΑΤΗΡΗΣΕΙΣ σελ.64

6.2. ΠΗΓΕΣ σελ.65

ΕΥΡΕΤΗΡΙΟ ΠΙΝΑΚΩΝ – ΓΡΑΦΗΜΑΤΩΝ – ΕΙΚΟΝΩΝ:

ΠΙΝΑΚΕΣ

Πίνακας 1 – Πίνακας Συνολικών Πωλήσεων	σελ.20
Πίνακας 2 – Πίνακας Μοναδικών Πωλήσεων	σελ.24

ΓΡΑΦΗΜΑΤΑ

Γράφημα 1 – Word Cloud	σελ.22
Γράφημα 2 – Unique Sales	σελ.23

ΕΙΚΟΝΕΣ

Εικόνα 1 – Dataset Prototype	σελ.12
Εικόνα 2 – Dataset Format	σελ.12
Εικόνα 3 – Table Basket	σελ.13
Εικόνα 4 – Table Products	σελ.14
Εικόνα 5 – Table Products With Count	σελ.16
Εικόνα 6 – Total	σελ.16
Εικόνα 7 – Percent Of Sales Transformations	σελ.17
Εικόνα 8 – Percent Of Sales	σελ.17
Εικόνα 9 – Transformed Basket	σελ.18
Εικόνα 10 – Mineral Water Set	σελ.25
Εικόνα 11 – Eggs Set	σελ.27
Εικόνα 12 – Spaghetti Set	σελ.29
Εικόνα 13 – French Fries Set	σελ.30
Εικόνα 14 – Chocolate Set	σελ.31
Εικόνα 15 – Green Tea Set	σελ.32
Εικόνα 16 – Milk Set	σελ.34
Εικόνα 17 – Ground Beef Set	σελ.35
Εικόνα 18 – Frozen Vegetables Set	σελ.37
Εικόνα 19 – Pancakes Set	σελ.38
Εικόνα 20 – Burgers Set	σελ.39
Εικόνα 21 – Cake Set	σελ.41

Εικόνα 22 – Cookies Set	σελ.42
Εικόνα 23 – Escalope Set	σελ.44
Εικόνα 24 – Low Fat Yogurt Set	σελ.46
Εικόνα 25 – Shrimp Set	σελ.47
Εικόνα 26 – Tomatoes Set	σελ.48
Εικόνα 27 – Olive Oil Set	σελ.49
Εικόνα 28 – Frozen Smoothie Set	σελ.50
Εικόνα 29 – Turkey Set	σελ.51
Εικόνα 30 – Oatmeal Set	σελ.52
Εικόνα 31 – Chocolate Bread Set	σελ.54
Εικόνα 32 – Mashed Potato Set	σελ.55
Εικόνα 33 – Chutney Set	σελ.56
Εικόνα 34 – Tea Set	σελ.57
Εικόνα 35 – Bramble Set	σελ.58
Εικόνα 36 – Cream Set	σελ.59
Εικόνα 37 – Napkins Set	σελ.61
Εικόνα 38 – Water Spray Set	σελ.62
Εικόνα 39 – Asparagus Set	σελ.63

ΕΙΣΑΓΩΓΗ

Η παρούσα εργασία με θέμα: Εξόρυξη προσωποποιημένης πληροφόρησης από περιβάλλοντα super-market με σκοπό την ανάλυση αγορών προϊόντων πραγματοποιήθηκε στα πλαίσια της διπλωματικής μου εργασίας. Η εργασία περιλαμβάνει τρία μέρη:

1. Θεωρητικό πλαίσιο: Θα παρουσιαστούν σύντομα κάποιοι όροι ώστε να κατανοηθεί καλύτερα το πεδίο που θα ασχοληθούμε στην εργασία αυτή.
2. Μεθοδολογία: Θα παρουσιαστούν αναλυτικά όλοι οι μέθοδοι που χρησιμοποιήθηκαν στα πλαίσια της εργασίας αυτής.
3. Συμπεράσματα: Θα παρουσιαστούν πίνακες και γραφικές απεικονίσεις μαζί με συμπεράσματα για το dataset.

ΘΕΩΡΗΤΙΚΟ ΠΛΑΙΣΙΟ

ΑΓΟΡΕΣ ΠΡΟΙΟΝΤΩΝ

Γενικά με τον όρο αγορά στην οικονομική επιστήμη χαρακτηρίζεται ένας θεωρητικός χώρος στον οποίο απαντάται η προσφορά και ζήτηση αγαθών ή παροχή υπηρεσιών, ή συγκεκριμένος χώρος που συγκεντρώνονται συναλλασσόμενοι εμπορικά ή ακόμα και το σύνολο αυτών των συναλλασσόμενων ή του είδους αυτών, κατά ακολουθούμενη πρακτική.

Γενικά σήμερα ο όρος αγορά αποδίδεται με τη σημασία κάποιου χώρου όπου γίνονται αγοραπωλησίες αγαθών, π.χ. Δημοτική Αγορά Αθήνας. Στη σύγχρονη όμως Οικονομική και τα οικονομικά συστήματα η έννοια της αγοράς έχει μια ευρύτερη σημασία θεωρητικού χώρου όπου πραγματοποιούνται οι εμπορικές συναλλαγές.

Οι αγορές διακρίνονται κυρίως κατά τόπο δραστηριότητας, κατά μορφή ή τύπο δραστηριότητας και κατά είδος διακινούμενου αγαθού.

- Κατά τόπο δραστηριότητας οι αγορές διακρίνονται σε: "τοπικές", "εθνικές" και "διεθνείς".
- Κατά τύπο ή μορφή, οι αγορές διακρίνονται σε: ελεύθερη αγορά, ή αγορά ελεύθερου συναγωνισμού, αγορά ευτελούς συναγωνισμού και αγορά άνευ συναγωνισμού, και
- Κατά διακινούμενα αγαθά, οι αγορές διακρίνονται σε: "αγορά πρωτογενών προϊόντων," "βιομηχανικών προϊόντων", "αγορά υπηρεσιών" (παροχής), (π.χ. μεταφορές), αγορά εργασίας κλπ.
- Κατά μια συμφωνία αγοράς: π.χ. αγορά ενός οικοπέδου δηλαδή το να πληρώνει κανείς για να αποκτήσει κάτι κλπ.

MARKETING

Το μάρκετινγκ (marketing), συνίσταται στην οργανωμένη προσπάθεια μιας επιχείρησης ή ενός οργανισμού να ικανοποιήσει τις ανάγκες αλλά και τις επιθυμίες των καταναλωτών. Προσπαθεί δηλαδή, με κύριο εργαλείο την έρευνα αγοράς, να αντιστοιχίσει τα προϊόντα ή τις υπηρεσίες που παράγει με τον πελάτη-στόχο που τα χρειάζεται ή τα επιθυμεί, ή ακόμα καλύτερα αφού κατανοήσει τις ανάγκες και τις επιθυμίες του, να κατασκευάσει τα αντίστοιχα προϊόντα / υπηρεσίες με τα χαρακτηριστικά και τις ιδιότητες που ο πελάτης επιθυμεί, να του τα γνωστοποιήσει (διαφήμιση και προώθηση), να τα καταστήσει διαθέσιμα μέσα από τα κανάλια διανομής (τα μαγαζιά και οι τοποθεσίες που αυτά είναι διαθέσιμα) στην τιμή που θα πρέπει αυτά να πωλούνται.

Ο βασικός στόχος του μάρκετινγκ, ειδικότερα όσον αφορά τα καταναλωτικά αγαθά, δηλαδή αυτά που απευθύνονται στους τελικούς καταναλωτές (και όχι σε επιχειρήσεις και οργανισμούς, οπότε γίνεται λόγος για βιομηχανικά προϊόντα/υπηρεσίες) είναι οι επαναλαμβανόμενες πωλήσεις. Ειδικότερα τα στελέχη μάρκετινγκ μιας επιχείρησης ή ενός οργανισμού προσπαθούν να δημιουργήσουν μία μακροχρόνια σχέση με τον πελάτη προσφέροντας υψηλή αξία στα προϊόντα τους, αξία υψηλότερη από τα αντίστοιχα προϊόντα του ανταγωνισμού. Μέσα από την υψηλότερη αυτή αξία επιτυγχάνεται το ανταγωνιστικό πλεονέκτημα της επιχείρησης, δηλαδή η υπεροχή της και των προϊόντων της έναντι του ανταγωνισμού. Όλα τα παραπάνω με τη σειρά τους οδηγούν σε πιστότητα των πελατών και τελικά σε επαναλαμβανόμενες πωλήσεις των προϊόντων / υπηρεσιών της επιχείρησης.

Ειδικότερα τώρα, όσον αφορά τα super markets, υπάρχουν δοκιμασμένες και πλέον ευρέως χρησιμοποιούμενες στρατηγικές για την επίτευξη καλύτερων αποτελεσμάτων στις πωλήσεις που πετυχαίνει το εκάστοτε κατάστημα. Ενδεικτικά, μερικές από αυτές τις στρατηγικές:

- Εποχιακές καμπάνιες
- Διαφήμιση πολυχρηστικών προϊόντων
- Προσφορές προϊόντων
- Συνεργασία με επώνυμες μάρκες προϊόντων
- Δωρεάν δείγματα
- Καμπάνιες μέσω κοινωνικών μέσων

DATA ANALYSIS

Η επιστήμη των Δεδομένων είναι ένα διεπιστημονικό πεδίο του οποίου αντικείμενο είναι η εξαγωγή της γνώσης από αδόμητα ή δομημένα δεδομένα. Αποτελεί τη συνέχεια επιστημών όπως η στατιστική, η ανάλυση προγνωστικών (predictive analytics), η μηχανική μάθηση (machine learning) και η εξόρυξη δεδομένων (data mining). Με τις τεχνολογικές προόδους τις δύο τελευταίες δεκαετίες, σε συνδυασμό, εν μέρη με την έκρηξη του διαδικτύου, έχει προκύψει μια νέα μορφή ανάλυσης δεδομένων.

Ο ορισμός του ακριβή ρόλου ενός επιστήμονα δεδομένων περιλαμβάνει ένα συγκεκριμένο υπόβαθρο σε κάποιους επιστημονικούς τομείς αλλά και ένα σύνολο ικανοτήτων που δεν έχουν προσδιοριστεί απολύτως καθώς η συγκεκριμένη επιστήμη αποτελεί το συνδυασμό περισσότερων από τους συνήθεις τομείς έρευνας. Εν γένει, ένας επιστήμονας δεδομένων χρησιμοποιεί τις τρέχουσες υπολογιστικές τεχνικές για την ανάλυση δεδομένων. Οι εκάστοτε ερευνητές, όπως οι στατιστικοί, οι φυσικοί, οι βιολόγοι και οι οικονομολόγοι από πάντα αναζητούσαν τη πληροφορία στους αντίστοιχους τομείς έρευνάς τους, πλέον, όμως, η σφαίρα της πληροφορίας έχει αλλάξει, ως εκ τούτου, απαιτείται ένα νέο σύνολο δεξιοτήτων για την αποδοτικότερη ανάλυση των δεδομένων που λαμβάνονται. Παράλληλα, παρατηρείται εκθετική αύξηση της υπολογιστικής ισχύος η οποία παρέχει νέες μεθόδους για να επεξεργαστούν οι συνεχώς αυξανόμενες ποσότητες δεδομένων που συλλέγονται κάθε δευτερόλεπτο της ημέρας.

Συνεπώς, κάθε επιστήμονας δεδομένων θα πρέπει να επενδύσει χρόνο να εξοικειωθεί με υπολογιστικές τεχνικές οι οποίες διαχειρίζονται τον όγκο δεδομένων και τη πολυπλοκότητα των δεδομένων σήμερα. Πέρα από τις μαθηματικές και τις στατιστικές γνώσεις, οι υπόλοιπες δεξιότητες ενός επιστήμονα δεδομένων δεν αφορούν έναν μόνο συγκεκριμένο τομέα.

ΜΕΘΟΔΟΛΟΓΙΑ

Στην παρούσα εργασία θα εξετάσουμε ένα [dataset](#) που περιέχει πωλήσεις προϊόντων super market και θα προσπαθήσουμε να βγάλουμε συμπεράσματα για τους καταναλωτές με βάση τις αγορές τους. Η διαδικασία αυτή θα πραγματοποιηθεί με τη δημιουργία μιας βάσης δεδομένων όπου θα αποθηκευτούν και στη συνέχεια θα επεξεργαστούν τα δεδομένα του dataset, με τη χρήση του Microsoft SQL Server.

Το dataset που επιλέξαμε στην παρούσα εργασία αφορά καλάθια αγορών που έγιναν σε κατάστημα σουπερ-μάρκετ. Στην παρακάτω εικόνα βρίσκεται στην πρωταρχική του μορφή όπως το πήραμε από το Kaggle.com σε αρχείο CSV.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T
1	shrimp	almonds	avocado	vegetable	green gra	whole we	yams	cottage ch	energy dri	tomato ju	low fat yo	green tea	honey	salad	mineral w	salmon	antioxyda	frozen sm	spinach	olive oil
2	burgers	meatballs	eggs																	
3	chutney																			
4	turkey	avocado																		
5	mineral w	milk	energy ba	whole wh	green tea															
6	low fat yogurt																			
7	whole wh	french fries																		
8	soup	light crear	shallot																	
9	frozen vej	spaghetti	green tea																	
10	french fries																			
11	eggs	pet food																		
12	cookies																			
13	turkey	burgers	mineral w	eggs	cooking oil															
14	spaghetti	champagn	cookies																	
15	mineral w	salmon																		
16	mineral water																			
17	shrimp	chocolate	chicken	honey	oil	cooking oi	low fat yogurt													
18	turkey	eggs																		
19	turkey	fresh tuna	tomatoes	spaghetti	mineral w	black tea	salmon	eggs	chicken	extra dark chocolate										
20	meatballs	milk	honey	french frie	protein bar															

Εικόνα 1 – Dataset Prototype

Το αρχείο περιλαμβάνει σε κάθε γραμμή ένα καλάθι αγορών με το κάθε προϊόν να βρίσκεται σε ξεχωριστό κελί. Αυτό που κάναμε όπως φαίνεται στην παρακάτω εικόνα είναι να μετρήσουμε το μέγιστο αριθμό στηλών όπου υπάρχει εγγραφή και να τοποθετήσουμε την πρώτη γραμμή που χωρίζει τις στήλες (Product_1, Product_2, ..., Product_20). Αυτό το κάναμε με σκοπό να διευκολυνθούμε στην τοποθέτηση του πίνακα στην βάση.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T
1	Product_1	Product_2	Product_3	Product_4	Product_5	Product_6	Product_7	Product_8	Product_9	Product_1	Product_1	Product_1	Product_1	Product_1	Product_1	Product_1	Product_1	Product_1	Product_1	Product_20
2	shrimp	almonds	avocado	vegetable	green gra	whole we	yams	cottage ch	energy dri	tomato ju	low fat yo	green tea	honey	salad	mineral w	salmon	antioxyda	frozen sm	spinach	olive oil
3	burgers	meatballs	eggs																	
4	chutney																			
5	turkey	avocado																		
6	mineral w	milk	energy ba	whole wh	green tea															
7	low fat yogurt																			
8	whole wh	french fries																		
9	soup	light crear	shallot																	
10	frozen vej	spaghetti	green tea																	
11	french fries																			
12	eggs	pet food																		
13	cookies																			
14	turkey	burgers	mineral w	eggs	cooking oil															
15	spaghetti	champagn	cookies																	
16	mineral w	salmon																		
17	mineral water																			
18	shrimp	chocolate	chicken	honey	oil	cooking oi	low fat yogurt													
19	turkey	eggs																		
20	turkey	fresh tuna	tomatoes	spaghetti	mineral w	black tea	salmon	eggs	chicken	extra dark chocolate										

Εικόνα 2 – Dataset Format

Παρακάτω παρουσιάζεται η μεθοδολογία που ακολούθησα για την διεκπεραίωση της παρούσας εργασίας.

1. Δημιουργία του table Basket (9002x31)

	ID	Product_1	Product_2	Product_3	Product_4	Product_5	Product_6	Product_7	Product_8	Product_9	Product_10	Product_11
1	1	shrimp	almonds	avocado	vegetables mix	green grapes	whole wheat flour	yams	cottage cheese	energy drink	tomato juice	low fat yogurt
2	2	burgers	meatballs	eggs	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL
3	3	chutney	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL
4	4	turkey	avocado	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL
5	5	mineral water	milk	energy bar	whole wheat rice	green tea	NULL	NULL	NULL	NULL	NULL	NULL
6	6	low fat yogurt	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL
7	7	whole wheat pasta	french fries	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL
8	8	soup	light cream	shallot	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL
9	9	frozen vegetables	spaghetti	green tea	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL
10	10	french fries	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL
11	11	eggs	pet food	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL
12	12	cookies	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL
13	13	turkey	burgers	mineral water	eggs	cooking oil	NULL	NULL	NULL	NULL	NULL	NULL
14	14	spaghetti	champagne	cookies	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL
15	15	mineral water	salmon	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL
16	16	mineral water	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL
17	17	shrimp	chocolate	chicken	honey	oil	cooking oil	low fat yogurt	NULL	NULL	NULL	NULL
18	18	turkey	eggs	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL
19	19	turkey	fresh tuna	tomatoes	spaghetti	mineral water	black tea	salmon	eggs	chicken	extra dark chocolate	NULL
20	20	meatballs	milk	honey	french fries	protein bar	NULL	NULL	NULL	NULL	NULL	NULL

Εικόνα 3 – Table Basket

Στην παραπάνω εικόνα φαίνεται ο πρώτος πίνακας «Basket» που δημιουργήσαμε στην βάση μας ο οποίος περιλαμβάνει το dataset ακριβώς όπως ήταν. Η μεταφορά του στην βάση έγινε με την χρήση του εργαλείου import wizard του SQL Server. Επιπλέον προστέθηκε στον πίνακα το column «id» το οποίο είναι ένας μοναδικός αριθμός που δίνεται σε κάθε γραμμή του πίνακα. Οι γραμμές του πίνακα αφορούν την κάθε αγορά που πραγματοποιήθηκε από κάποιον πελάτη, δηλαδή τα προϊόντα που βρέθηκαν μαζί σε ένα καλάθι και οι στήλες αφορούν το κάθε προϊόν του καλαθιού.

Ο παραπάνω πίνακας έγινε ώστε να υπάρχει ένας πίνακας που να περιέχει το dataset όπως ακριβώς είναι.

2. Δημιουργία του table Products (120x4)

	Product_ID	Product_Name	COUNT	%
1	1	shampoo	37	0.12
2	2	extra dark chocolate	90	0.3
3	3	spaghetti	1306	4.44
4	4	hand protein bar	39	0.13
5	5	tomato juice	228	0.77
6	6	chocolate	1229	4.18
7	7	cream	7	0.02
8	8	cookies	603	2.05
9	9	cottage cheese	239	0.81
10	10	barbecue sauce	81	0.27
11	11	ham	199	0.67
12	12	soup	379	1.29
13	13	olive oil	494	1.68
14	14	herb & pepper	371	1.26
15	15	ketchup	33	0.11
16	16	burgers	654	2.22
17	17	cauliflower	36	0.12
18	18	shrimp	536	1.82
19	19	cooking oil	383	1.3

Εικόνα 4 – Table Products

Στην συνέχεια δημιουργήθηκε ο παραπάνω πίνακας «Products», προκειμένου να βρεθούν όλα τα προϊόντα που υπάρχουν στο dataset. Ο πίνακας αυτός δημιουργήθηκε ώστε να καταγραφούν όλα τα προϊόντα που υπάρχουν στο dataset με σκοπό να μπορούν μετέπειτα να συγκριθούν μεταξύ τους. Περιέχει ένα μοναδικό id για κάθε προϊόν (Product_id), όλα τα προϊόντα που βρέθηκαν στο dataset (Product_Name), τις συνολικές πωλήσεις που έκανε το κάθε προϊόν (COUNT), το ποσοστό τις εκατό που έκανε το κάθε προϊόν σε σχέση με όλες τις πωλήσεις (%).

```
select distinct Product_1 into Products
from [Basket-analysis].[dbo].[Basket]
```

Το παραπάνω query βρίσκει όλα τα προϊόντα που υπάρχουν στην πρώτη στήλη του πίνακα «Basket» και τα τοποθετεί στον πίνακα «Products» στην στήλη «Product_Name».

```
insert into products(Product_Name)
select distinct Product_2
from [Basket-analysis].[dbo].[Basket]
where Product_2 not in (select Product_Name from products) and isnull(Product_2, '')<>''
```

Έπειτα συνεχίσαμε με το παραπάνω query στο οποίο περνάμε στο πίνακα «Products» και τα προϊόντα που υπάρχουν στην δεύτερη στήλη τα οποία δεν βρέθηκαν στην πρώτη στήλη. Επιπλέον εδώ υπάρχει και ο

περιορισμός για την κενή εγγραφή. Στην πρώτη περίπτωση δεν το ελέγξαμε επειδή ήταν η πρώτη στήλη όπου δεν υπάρχουν κενές εγγραφές.

Έπειτα συνεχίσαμε με το ίδιο query και για τις υπόλοιπες στήλες με τη λογική του παραπάνω query αλλάζοντας μόνο τον αριθμό της στήλης που ελέγχουμε κάθε φορά “Product_(αριθμός προϊόντος)”. Τελειώνοντας τη διαδικασία αυτή δημιουργήθηκε ένας πίνακας με όλα τα προϊόντα που υπάρχουν στο dataset μας.

Έπειτα με τη χρήση του SQL server τοποθετήθηκε η έξτρα στήλη με τα μοναδικά ID.

```
update Products
set COUNT = (select count(ID)
from [Basket-analysis].[dbo].[Market_Basket_Optimisation]
where Product_1 = 'extra dark chocolate' or Product_2 = 'extra dark chocolate' or Product_3
= 'extra dark chocolate' or Product_4 = 'extra dark chocolate'
or Product_5 = 'extra dark chocolate' or Product_6 = 'extra dark chocolate' or Product_7 =
'extra dark chocolate' or Product_8 = 'extra dark chocolate'
or Product_9 = 'extra dark chocolate' or Product_10 = 'extra dark chocolate' or Product_11
= 'extra dark chocolate' or Product_12 = 'extra dark chocolate'
or Product_13 = 'extra dark chocolate' or Product_14 = 'extra dark chocolate' or Product_15
= 'extra dark chocolate' or Product_16 = 'extra dark chocolate'
or Product_17 = 'extra dark chocolate' or Product_18 = 'extra dark chocolate' or Product_19
= 'extra dark chocolate' or Product_20 = 'extra dark chocolate')
where Product_Name = 'extra dark chocolate'
```

Το παραπάνω query χρησιμοποιήθηκε προκειμένου να βρούμε των αριθμό των πωλήσεων που έκανε το κάθε προϊόν. Στην συγκεκριμένη περίπτωση ψάχνουμε το σύνολο για το προϊόν “extra dark chocolate”. Αυτό το κάνουμε χρησιμοποιώντας τη συνάρτηση count() της SQL η οποία αθροίζει όλες εγγραφές βρίσκει που πληρούν τα κριτήρια. Στο συγκεκριμένο query μετράει όλες τις εγγραφές που βρίσκει το όνομα του προϊόντος που δώσαμε “extra dark chocolate”.

Αντίστοιχα το ίδιο query το τρέξαμε και για τα υπόλοιπα προϊόντα και βγάλαμε τη στήλη COUNT.

	Product_ID	Product_Name	COUNT
1	90	mineral water	1788
2	77	eggs	1348
3	3	spaghetti	1306
4	110	french fries	1282
5	6	chocolate	1229
6	104	green tea	991
7	95	milk	972
8	47	ground beef	737
9	59	frozen vegetables	715
10	107	pancakes	713
11	16	burgers	654
12	96	cake	608
13	8	cookies	603
14	98	escalope	595
15	40	low fat yogurt	574
16	18	shrimp	536
17	34	tomatoes	513
18	13	olive oil	494
19	42	frozen smoothie	475
20	93	turkey	469
21	21	chicken	450
22	102	whole wheat rice	439
23	86	grated cheese	393
24	19	cooking oil	383

Εικόνα 5 – Table Products With Count

Πραγματοποιώντας όλα τα παραπάνω ο πίνακάς μας έχει διαμορφωθεί όπως φαίνεται στην παραπάνω εικόνα.

```
select sum(COUNT) from Products
```

	(No column name)
1	29358

Εικόνα 6 – Total Count

Με το παραπάνω query βρίσκω το σύνολο όλων των πωλήσεων χρησιμοποιώντας την συνάρτηση sum() της SQL.

```
select *, cast(COUNT as float)/29358, (cast(COUNT as float)/sum(COUNT))*100,
left((cast(COUNT as float)/29358)*100, 4) as '(%)'
from Products
```

	Product_ID	Product_Name	COUNT	%	(No column name)	(No column name)	(%)
1	1	shampoo	37	0,12	0,00126030383541113	0,126030383541113	0.12
2	2	extra dark chocolate	90	0,3	0,00306560392397302	0,306560392397302	0.30
3	3	spaghetti	1306	4,44	0,0444853191634308	4,44853191634308	4.44
4	4	hand protein bar	39	0,13	0,00132842836705498	0,132842836705498	0.13
5	5	tomato juice	228	0,77	0,00776619660739832	0,776619660739832	0.77
6	6	chocolate	1229	4,18	0,0418625246951427	4,18625246951427	4.18
7	7	cream	7	0,02	0,000238435860753457	0,0238435860753457	0.02
8	8	cookies	603	2,05	0,0205395462906193	2,05395462906193	2.05
9	9	cottage cheese	239	0,81	0,00814088153143947	0,814088153143947	0.81

Εικόνα 7 – Percent Of Sales Transformations

Εδώ φαίνονται οι μετατροπές που έγιναν προκειμένου να δημιουργηθεί η στήλη με τα ποσοστά από τις πωλήσεις του κάθε προϊόντος. Διαιρέσαμε τις πωλήσεις με το σύνολο που βρήκαμε πριν αφού μετατρέψαμε τους αριθμούς σε δεκαδικούς με τη συνάρτηση cast (as float), ώστε να μπορούμε να δεχθούμε σαν αποτέλεσμα δεκαδικά ψηφία. Έπειτα πολλαπλασιάζουμε με 100 προκειμένου να φτιάξουμε το ποσοστό και κρατάμε τα 4 πρώτα ψηφία του αριθμού με τη συνάρτηση left.

```
update Products set [%] = left((cast(COUNT as float)/sum(COUNT))*100, 4)
```

	Product_ID	Product_Name	COUNT	%
1	1	shampoo	37	0,12
2	2	extra dark chocolate	90	0,3
3	3	spaghetti	1306	4,44
4	4	hand protein bar	39	0,13
5	5	tomato juice	228	0,77
6	6	chocolate	1229	4,18
7	7	cream	7	0,02
8	8	cookies	603	2,05
9	9	cottage cheese	239	0,81

Εικόνα 8 – Percent Of Sales

Εδώ κάνουμε το τελικό update στον πίνακά μας και τον βλέπουμε πως είναι στην τελική του μορφή.

3. Δημιουργία του table Transformed_Basket (150,650x5)

	Row_ID	ProductID	Product_Name	Product_with_it	IDSale
1	1	105	almonds	NULL	254
2	2	105	almonds	eggs	4670
3	3	105	almonds	pancakes	4742
4	4	105	almonds	eggs	1395
5	5	105	almonds	french wine	3589
6	6	105	almonds	cake	731
7	7	105	almonds	eggs	2071
8	8	105	almonds	french wine	5923
9	9	105	almonds	eggs	6878
10	10	105	almonds	cookies	7011
11	11	105	almonds	NULL	7454
12	12	84	antioxydant juice	french fries	6842
13	13	84	antioxydant juice	NULL	5759
14	14	84	antioxydant juice	NULL	6646
15	15	84	antioxydant juice	NULL	5528
16	16	84	antioxydant juice	french fries	5688
17	17	84	antioxydant juice	escalope	349
18	18	84	antioxydant juice	NULL	3193
19	19	84	antioxydant juice	chocolate	3312
20	20	84	antioxydant juice	chocolate	628
21	21	84	antioxydant juice	NULL	2850
22	22	84	antioxydant juice	mint green tea	2873
23	23	84	antioxydant juice	french fries	3390
24	24	84	antioxydant juice	brownies	3702
25	25	84	antioxydant juice	NULL	3706
26	26	84	antioxydant juice	NULL	3707
27	27	84	antioxydant juice	chocolate	5197
28	28	84	antioxydant juice	green tea	5204
29	29	84	antioxydant juice	escalope	4274
30	30	35	asparagus	salad	89

Εικόνα 9 – Transformed Basket

Στην συνέχεια για να μπορέσουν να συγκριθούν τα προϊόντα μεταξύ τους δημιουργήθηκε ο παραπάνω πίνακας «Transformed Basket». Ο πίνακας αυτός προέκυψε από τον αρχικό μας πίνακα «Basket» πολλαπλασιάζοντας ουσιαστικά τις γραμμές με τις στήλες του πίνακα. Δηλαδή, χωρίστηκαν τα προϊόντα του σετ δεδομένων ανά δύο με όλους τους συνδυασμούς που μπορούσαν να έχουν μεταξύ τους στο ίδιο καλάθι αγορών. Αυτό έγινε με τη βοήθεια του πίνακα «Products». Αυτό συνέβη διότι για τις συγκρίσεις που θέλουμε να κάνουμε στην παρούσα εργασία τα προϊόντα χρειάζεται να βρίσκονται σε ζευγάρια.

Ο πίνακας «Transformed Basket» περιέχει ένα μοναδικό id (Row_ID), το id του προϊόντος από τον πίνακα Products (ProductID), το όνομα του προϊόντος (Product_Name), το κάθε προϊόν που υπήρχε στο καλάθι μαζί με το προϊόν που εξετάζουμε (Product_with_it), το id του καλαθιού στο οποίο έγινε η αγορά που παίρνει από τον πίνακα Basket.

```
insert into [dbo].[Transformed_Basket](ProductID, Product_Name, Product_with_it, IDSale)
select product_id, product_name, product_2 from products a join basket b on a.product_name
= b.product_1 order by product_name
```

```
insert into [dbo].[Transformed_Basket](productid, Product_Name, Product_with_it, IDSale)
select a.product_id, a.product_name, b.product_3, b.id from products a join basket b on
a.product_name = b.product_1
where isnull(product_2, '')<>''
order by product_name
```

Εδώ βλέπουμε τη διαδικασία της μεταφοράς των προϊόντων από τον αρχικό πίνακα Basket, στον πίνακα Transformed Basket που είναι χωρισμένα ανά δύο. Στην ουσία κάθε φορά παίρνουμε ένα προϊόν από τον πίνακα με κάθε άλλο προϊόν που βρίσκεται στην ίδια εγγραφή και αυτό το κάνουμε για όλα τα προϊόντα. Παραπάνω φαίνεται το πρώτο μαζί με το δεύτερο και αντίστοιχα το πρώτο σε σχέση με το τρίτο. Ελέγχουμε πάντα για την προηγούμενη εγγραφή να μην είναι κενή γιατί σημαίνει ότι δεν υπάρχουν άλλα προϊόντα στην εγγραφή μας. Ωστόσο δεν περιορίζουμε το δεύτερο προϊόν στο πρώτο query επειδή αφορά την πρώτη στήλη και θέλουμε να καταγράψουμε και τα προϊόντα που βρέθηκαν μόνο τους στα καλάθια.

Ο πίνακας Transformed Basket είναι ο τελικός πίνακας που χρησιμοποιήσαμε για να πάρουμε αποτελέσματα από το datasets μας. Το γεγονός ότι έχουμε πλέον τα προϊόντα σε ζευγάρια μας βοηθάει στο να μπορούμε να τα συγκρίνουμε ευκολότερα.

4. Δημιουργία του Pivot table Transformed Basket.

Έπειτα με τη χρήση του Excel δημιουργήθηκε ένα pivot table το οποίο συγχρονίστηκε με τον πίνακα «Transformed Basket» από τη βάση μας. Με τη χρήση του pivot δημιουργήθηκαν πίνακες και σχεδιαγράμματα από τα οποία έγινε η ανάλυση των δεδομένων του dataset και η εξαγωγή συμπερασμάτων από αυτά.

ΠΙΝΑΚΑΣ ΣΥΝΟΛΙΚΩΝ ΠΩΛΗΣΕΩΝ

Product_Name	Count	Count(%)	Product_Name	Count	Count(%)	Product_Name	Count	Count(%)
1 mineral water	1788	6.09%	41 energy bar	203	0.69%	81 green grapes	68	0.23%
2 eggs	1348	4.59%	42 energy drink	200	0.68%	82 flax seed	68	0.23%
3 spaghetti	1306	4.45%	43 pepper	199	0.68%	83 antioxydant juice	67	0.23%
4 french fries	1282	4.37%	44 ham	199	0.68%	84 bug spray	65	0.22%
5 chocolate	1229	4.19%	45 vegetables mix	193	0.66%	85 bacon	65	0.22%
6 green tea	991	3.38%	46 cereals	193	0.66%	86 green beans	65	0.22%
7 milk	972	3.31%	47 muffins	181	0.62%	87 clothes accessories	63	0.21%
8 ground beef	737	2.51%	48 oil	173	0.59%	88 toothpaste	61	0.21%
9 frozen vegetables	715	2.44%	49 french wine	169	0.58%	89 shallot	58	0.20%
10 pancakes	713	2.43%	50 fresh tuna	167	0.57%	90 strong cheese	58	0.20%
11 burgers	654	2.23%	51 strawberries	160	0.54%	91 spinach	53	0.18%
12 cake	608	2.07%	52 meatballs	157	0.53%	92 gluten free bar	52	0.18%
13 cookies	603	2.05%	53 almonds	153	0.52%	93 pet food	49	0.17%
14 escalope	595	2.03%	54 parmesan cheese	149	0.51%	94 sparkling water	47	0.16%
15 low fat yogurt	574	1.96%	55 mushroom cream sauce	143	0.49%	95 soda	47	0.16%
16 shrimp	536	1.83%	56 rice	141	0.48%	96 mayonnaise	46	0.16%
17 tomatoes	513	1.75%	57 protein bar	139	0.47%	97 chili	46	0.16%
18 olive oil	494	1.68%	58 mint	131	0.45%	98 pickles	45	0.15%
19 frozen smoothie	475	1.62%	59 white wine	124	0.42%	99 burger sauce	44	0.15%
20 turkey	469	1.60%	60 pasta	118	0.40%	100 mint green tea	42	0.14%
21 chicken	450	1.53%	61 light cream	117	0.40%	101 hand protein bar	39	0.13%
22 whole wheat rice	439	1.50%	62 carrots	115	0.39%	102 shampoo	37	0.13%
23 grated cheese	393	1.34%	63 black tea	107	0.36%	103 salad	37	0.13%
24 cooking oil	383	1.30%	64 tomato sauce	106	0.36%	104 corn	36	0.12%
25 soup	379	1.29%	65 fromage blanc	102	0.35%	105 cauliflower	36	0.12%
26 herb & pepper	371	1.26%	66 gums	101	0.34%	106 asparagus	35	0.12%
27 honey	356	1.21%	67 eggplant	99	0.34%	107 sandwich	34	0.12%
28 champagne	351	1.20%	68 extra dark chocolate	90	0.31%	108 babies food	34	0.12%
29 fresh bread	323	1.10%	69 melons	90	0.31%	109 dessert wine	33	0.11%
30 salmon	319	1.09%	70 yams	86	0.29%	110 ketchup	33	0.11%
31 brownies	253	0.86%	71 body spray	86	0.29%	111 oatmeal	33	0.11%
32 avocado	250	0.85%	72 magazines	82	0.28%	112 chocolate bread	32	0.11%
33 hot dogs	243	0.83%	73 barbecue sauce	81	0.28%	113 mashed potato	31	0.11%
34 cottage cheese	239	0.81%	74 cider	79	0.27%	114 chutney	31	0.11%
35 tomato juice	228	0.78%	75 nonfat milk	78	0.27%	115 tea	29	0.10%
36 butter	226	0.77%	76 candy bars	73	0.25%	116 bramble	14	0.05%
37 whole wheat pasta	221	0.75%	77 zucchini	71	0.24%	117 cream	7	0.02%
38 red wine	211	0.72%	78 whole weat flour	70	0.24%	118 napkins	5	0.02%
39 yogurt cake	205	0.70%	79 blueberries	69	0.24%	119 water spray	3	0.01%
40 light mayo	204	0.69%	80 salt	69	0.24%	120 asparagus	1	0.00%
						Total	29358	100.00%

Πίνακας 1 – Πίνακας Συνολικών Πωλήσεων

Στον παραπάνω πίνακα βλέπουμε τις συνολικές πωλήσεις για όλα τα προϊόντα που υπάρχουν στο dataset. Ο πίνακας δημιουργήθηκε στο Excel από τον πίνακα “Products” της βάσης χρησιμοποιώντας το query

```
--Φέρνει όλα τα στοιχεία για τα πεδία Product_Name και COUNT από τον πίνακα Products
ταξινομημένα κατά φθίνουσα σειρά βάση το column COUNT
select Product_Name, [COUNT] from [elly].[dbo].[Products] order by count desc
```

Η πρώτη, πέμπτη και ένατη στήλη () του πίνακα είναι ένας μοναδικός αριθμός που δόθηκε σε κάθε προϊόν σύμφωνα με τη σειρά που αυτό καταλαμβάνει στον πίνακα. Η δεύτερη, έκτη και δέκατη στήλη (Product_Name) είναι το εκάστοτε προϊόν. Η τρίτη, έβδομη και ενδέκατη στήλη (Count) είναι το σύνολο των φορών που βρέθηκε το προϊόν σε ένα καλάθι μόνο του ή και με άλλα προϊόντα. Τέλος, η τέταρτη, όγδοη και δωδέκατη στήλη (Count(%)) είναι το ποσοστό επί τις εκατό (%) των πωλήσεων που κατέχει το κάθε προϊόν σε σχέση με τις πωλήσεις του συνόλου. Η τελευταία γραμμή του πίνακα (Total) περιέχει τον συνολικό αριθμό των πωλήσεων όλων των προϊόντων και το συνολικό ποσοστό τους.

Παρατηρώντας των παραπάνω πίνακα μπορούμε να διακρίνουμε τα πιο δημοφιλή προϊόντα. Βλέπουμε ότι «κυριαρχεί» το προϊόν “Mineral Water” το οποίο θα μπορούσε να είναι λογικό εάν το dataset προέρχεται από κατάσταση σε περιοχή με μη πόσιμο νερό. Έπειτα ακολουθούν προϊόντα όπως τα “Eggs”, “Spaghetti”, “French Fries” κλπ. τα οποία είναι προϊόντα τα οποία είναι απαραίτητα σε κάθε νοικοκυριό αφού αποτελούν βασικά συστατικά για τις περισσότερες παρασκευές γευμάτων που ένα νοικοκυριό χρειάζεται καθημερινά. Σημαντικό είναι να παρατηρήσουμε τα προϊόντα που δεν είναι απαραίτητα «αναγκαία» για ένα νοικοκυριό όπως το προϊόν “Chocolate” το οποίο είναι ίσως το πιο δημοφιλές σνακ και οι πωλήσεις του ξεπερνάνε άλλα προϊόντα που είναι σχεδόν απαραίτητα όπως το λάδι ή τα λαχανικά.

Πολύ σημαντικό επίσης είναι να παρατηρήσουμε τα προϊόντα με τις χαμηλότερες πωλήσεις και να κατανοήσουμε γιατί συμβαίνει αυτό. Στον συγκεκριμένο πίνακα τελευταία στη σειρά βρίσκονται τα προϊόντα “Asparagus”, “Water Spray”, “Napkins” κτλ. Καλό θα ήταν να παρατηρήσουμε την πορεία των προϊόντων αυτών, δηλαδή να κάνουμε την ίδια διαδικασία σε άλλη χρονική περίοδο και να δούμε το σκορ τους ξανά αφού υπάρχουν πολλοί παράγοντες που παίζουν ρόλο στις αγορές των καταναλωτών όπως η εποχή κτλ. Καλό θα ήταν να προσέξουμε τις τιμές που έχουν τα συγκεκριμένα προϊόντα. Για παράδειγμα το προϊόν “Napkins” (χαρτοπετσέτες) είναι ένα προϊόν το οποίο είναι απαραίτητο στα περισσότερα νοικοκυριά, επομένως δεν δικαιολογείτε να είναι τόσο χαμηλά στην κατάταξη. Θα πρέπει λοιπόν να δούμε ποιος είναι ο λόγος που κάνει τις χαρτοπετσέτες να μην πουλιούνται. Ένας παράγοντας θα μπορούσε να είναι η τιμή τους στο συγκεκριμένο κατάσταση οπότε οι πελάτες να στρέφονται σε ανταγωνιστές. Πολύ σημαντικό είναι και η τάση προς την προστασία του περιβάλλοντος που παρατηρείται το τελευταίο διάστημα επομένως θα ήταν σημαντικό να προσέξουμε την παρασκευή του προϊόντος αλλά και των υπόλοιπων προϊόντων μας ή σε περίπτωση που εισάγουμε τα προϊόντα, τους προμηθευτές μας. Να φροντίσουμε λοιπόν οι παρασκευές να γίνονται με τρόπο ώστε να μην βλάπτουν το περιβάλλον και να το αναγράφουμε αυτό στα προϊόντα μας.

Αντίστοιχα, θα πρέπει να ελέγξουμε και την ποιότητα των προϊόντων μας. Το προϊόν “Asparagus” (σπαράγγι) είναι λαχανικό και παρόλο που δεν είναι τόσο διαδεδομένο προϊόν, χρησιμοποιείται από μερικά νοικοκυριά. Το γεγονός ότι σημείωσε μόνο μια πώληση θα έπρεπε να μας προβληματίζει αρκετά. Το πρόβλημα θα μπορούσα να είναι η ποιότητα του προϊόντος μας επομένως θα μπορούσαμε να αντικαταστήσουμε τα σπαράγγια μας με άλλα καλύτερης ποιότητας δοκιμαστικά και να ελέγξουμε πάλι τις

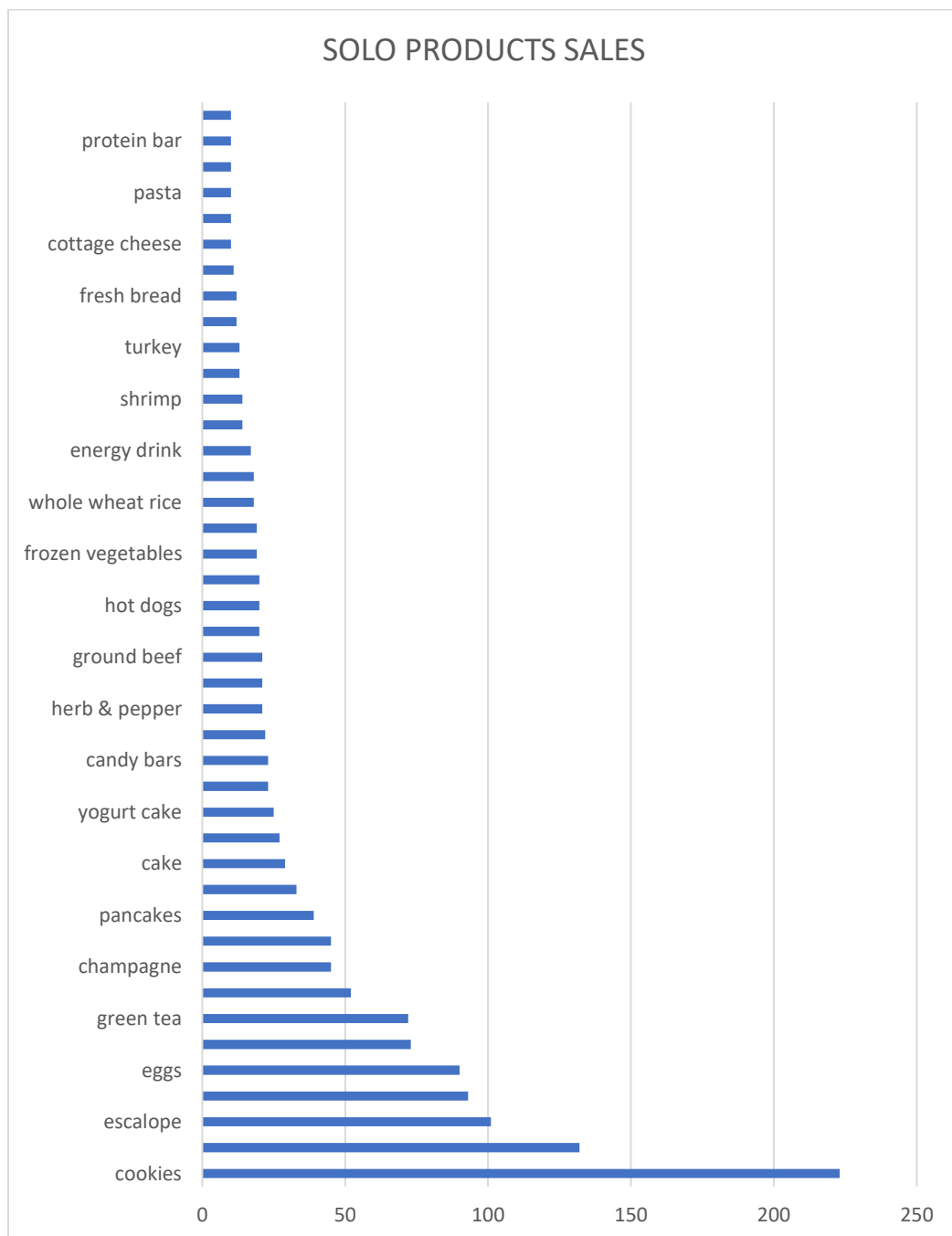
πωλήσεις τους. Αν δούμε ότι πάλι δεν έχουμε πωλήσεις θα πρέπει ίσως να «ρίξουμε» την τιμή για τον ίδιο λόγο με τις χαρτοπετσέτες ή ακόμη και να αποσύρουμε τελείως το προϊόν. Ένα προϊόν που δεν πουλάει είναι ζημία για το κατάστημα. Άλλωστε οι πελάτες του καταστήματος είναι μια μικρή κοινωνία με συγκεκριμένο lifestyle και ανάγκες και το προϊόν μας μπορεί να μην απευθύνεται σε αυτούς.



Γράφημα 1 – Word Cloud

Εδώ παρουσιάζεται ένα Word Cloud για τον παραπάνω πίνακα. Δημιουργήθηκε με τη χρήση του Wordart και περιέχει όλα τα προϊόντα του dataset. Τα προϊόντα που είναι περισσότερο ορατά είναι αυτά που συγκέντρωσαν μεγαλύτερο αριθμό πωλήσεων και τα λιγότερο ορατά αντίστοιχα με τον μικρότερο αριθμό. Δόθηκαν τα προϊόντα (Product_Name) σαν Word και ο αριθμός των πωλήσεων (Count) σαν Frequency. Το συγκεκριμένο σχήμα είναι ένας γρήγορος τρόπος να μελετήσουμε τις πωλήσεις μας χωρίς να υπάρχουν αριθμοί που μερικές φορές «μπερδεύουν» και να επικεντρωθούμε μόνο στα σημαντικά σημεία (στην προκειμένη περίπτωση στα δημοφιλή προϊόντα). Βέβαια, εάν θέλουμε να δούμε περισσότερες λεπτομέρειες θα πρέπει να ανατρέξουμε στον πίνακα.

ΜΟΝΑΔΙΚΕΣ ΠΩΛΗΣΕΙΣ ΠΡΟΙΟΝΤΩΝ



Γράφημα 2 – Unique Sales

Στην εικόνα παρουσιάζεται το διάγραμμα με τις πωλήσεις που έκανε το κάθε προϊόν από το dataset μόνο του. Δηλαδή όλες τις φορές που ένα προϊόν βρισκόταν μόνο του σε ένα καλάθι αγορών. Στην αριστερή πλευρά του διαγράμματος φαίνεται το προϊόν και αντίστοιχα οι πωλήσεις που έγιναν για το κάθε προϊόν. Τα προϊόντα

που παρουσιάζονται στο διάγραμμα είναι αυτά που βρέθηκαν σε ένα καλάθι μόνα τους 10 ή περισσότερες φορές (42 προϊόντα). Παρακάτω παρουσιάζονται αναλυτικά.

Product	Count	Count(%)	Product	Count	Count(%)	Product	Count	Count(%)
cookies	223	12.71%	olive oil	23	1.31%	shrimp	14	0.80%
french fries	132	7.53%	candy bars	23	1.31%	vegetables mix	13	0.74%
escalope	101	5.76%	energy bar	22	1.25%	turkey	13	0.74%
mineral water	93	5.30%	herb & pepper	21	1.20%	oil	12	0.68%
eggs	90	5.13%	grated cheese	21	1.20%	fresh bread	12	0.68%
chocolate	73	4.16%	ground beef	21	1.20%	salmon	11	0.63%
green tea	72	4.10%	burgers	20	1.14%	cottage cheese	10	0.57%
spaghetti	52	2.96%	hot dogs	20	1.14%	red wine	10	0.57%
champagne	45	2.57%	frozen smoothie	20	1.14%	pasta	10	0.57%
low fat yogurt	45	2.57%	frozen vegetables	19	1.08%	tomato juice	10	0.57%
pancakes	39	2.22%	tomatoes	19	1.08%	protein bar	10	0.57%
milk	33	1.88%	whole wheat rice	18	1.03%	chicken	10	0.57%
cake	29	1.65%	brownies	18	1.03%	Other	68	13.91%
muffins	27	1.54%	energy drink	17	0.97%	Total	1578	100.00%
yogurt cake	25	1.43%	strawberries	14	0.80%			

Πίνακας 2 – Πίνακας Μοναδικών Πωλήσεων

Όπως προαναφέραμε, παρουσιάζονται τα προϊόντα που έκαναν 10 ή περισσότερες πωλήσεις μόνα τους σε ένα καλάθι. Στην πρώτη στήλη «Product» βλέπουμε τα προϊόντα. Στην τελευταία τιμή «Other», έχουν συνυπολογιστεί όλα τα υπόλοιπα προϊόντα που βρίσκονται στο dataset τα οποία έχουν βρεθεί τουλάχιστον μία φορά σε καλάθι μόνα τους και έχουν κάτω από 10 πωλήσεις.

Στην δεύτερη στήλη «Count» βλέπουμε τις πωλήσεις που αντιστοιχούν σε κάθε προϊόν. Στην Τρίτη στήλη «Count(%)» βλέπουμε το ποσοστό επί τις εκατό που κατέχει κάθε προϊόν σε σχέση με τις συνολικές πωλήσεις στο σύνολο.

Παρατηρώντας τον πίνακα μπορούμε να δούμε τα πιο «δημοφιλή» προϊόντα, αυτά που έχουν δηλαδή τον μεγαλύτερο αριθμό σε πωλήσεις με το προϊόν cookies να «ισοπεδώνει» τα υπόλοιπα προϊόντα. Στην προκειμένη περίπτωση δημοφιλή χαρακτηρίζονται τα προϊόντα cookies, french fries, escalope, mineral water κλπ.

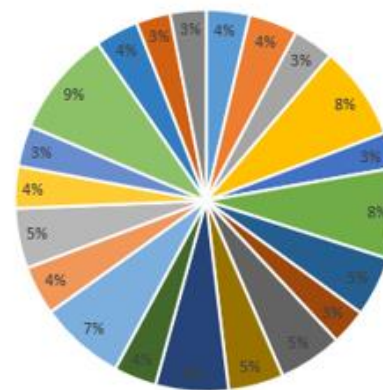
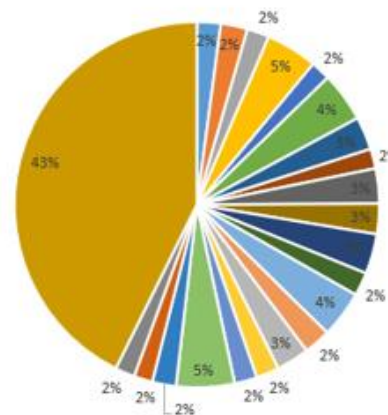
ΣΥΝΔΙΑΣΤΙΚΕΣ ΑΓΟΡΕΣ ΠΡΟΙΟΝΤΩΝ

Σε αυτήν την ενότητα θα παρουσιάσουμε πίνακες με ποσοστά για τα προϊόντα που βρέθηκαν σε ένα καλάθι μαζί με ένα συγκεκριμένο προϊόν που εξετάζουμε κάθε φορά. Θα αναλύσουμε τα πρώτα είκοσι προϊόντα που έκαναν τις περισσότερες πωλήσεις και τα τελευταία δέκα με τις λιγότερες πωλήσεις. Για όλες τις συγκρίσεις χρησιμοποιούμε τον πίνακα “Transformed Basket” της βάσης μας ο οποίος έχει γίνει linked με ένα pivot table στο excel όπου παίρνουμε και τα αποτελέσματα μας. Τα charts παράγονται από τον ίδιο πίνακα στο excel.

ΣΥΝΔΙΑΣΤΙΚΕΣ ΑΓΟΡΕΣ – ΠΕΡΙΣΣΟΤΕΡΕΣ ΠΩΛΗΣΕΙΣ

1. MINERAL WATER

Products_with_it	Count	Count(%)
spaghetti	448	5.16%
chocolate	396	4.56%
eggs	382	4.40%
milk	360	4.15%
ground beef	307	3.54%
frozen vegetables	268	3.09%
pancakes	253	2.91%
french fries	253	2.91%
green tea	233	2.68%
olive oil	207	2.38%
cake	206	2.37%
burgers	183	2.11%
tomatoes	183	2.11%
low fat yogurt	180	2.07%
shrimp	177	2.04%
soup	173	1.99%
chicken	171	1.97%
frozen smoothie	152	1.75%
whole wheat rice	151	1.74%
cooking oil	151	1.74%
Other	3848	44.32%
Grand Total	8682	100.00%



Εικόνα 10 – Mineral Water Set

Η δομή που φαίνεται στην παραπάνω εικόνα (πίνακας, charts (x2), λεζάντα) είναι αυτή που θα κρατήσουμε για να παρουσιαστούν τα προϊόντα που βρέθηκαν μαζί σε καλάθι για κάθε ένα από τα προϊόντα με τις μεγαλύτερες πωλήσεις. Ο τίτλος αφορά το προϊόν που εξετάζουμε κάθε φορά (Στην προκειμένη περίπτωση εξετάζουμε το προϊόν “Mineral Water”). Η πρώτη στήλη “Product_with_it” αφορά όλα τα προϊόντα που βρέθηκαν μαζί με το εκάστοτε προϊόν που εξετάζουμε σε ένα καλάθι αγορών. Η Τρίτη στήλη (Count) αφορά τις φορές που καταγράφηκε το προϊόν από την πρώτη στήλη. Τέλος, η τέταρτη στήλη αφορά το ποσοστό επί τις εκατό που κατέχει το προϊόν αυτό σε σχέση με το σύνολο των αγορών όλων των προϊόντων που αγοράστηκαν μαζί με το εκάστοτε προϊόν (Mineral Water). Ο πίνακας περιέχει τα 20 πρώτα προϊόντα με τις περισσότερες πωλήσεις που ήταν μαζί με το προϊόν που εξετάζουμε, τα οποία είναι ταξινομημένα κατά φθίνουσα σειρά με βάση τις πωλήσεις που κατέχουν. Όλα τα υπόλοιπα προϊόντα συνυπολογίζονται στην τιμή “Other”.

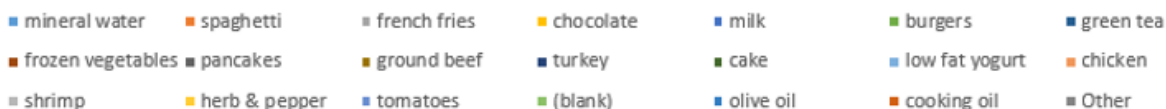
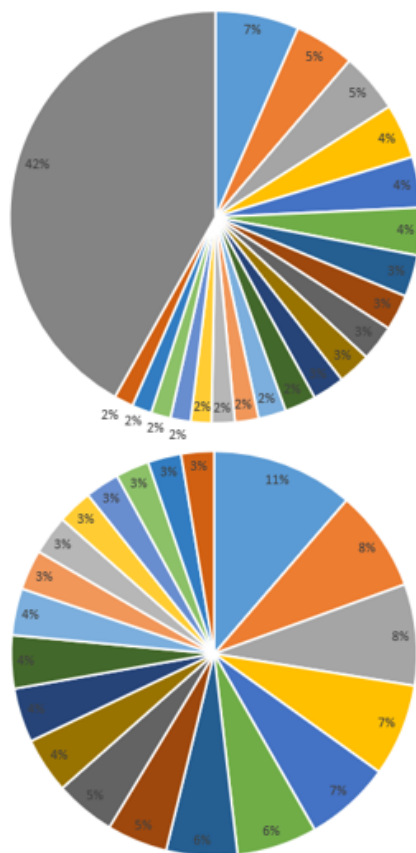
Η τελευταία γραμμή “Grand Total” εκφράζει το σύνολο των αγορών για τα καλάθια με το προϊόν που εξετάζουμε και δεν έχει σχέση με το σύνολο των αγορών καθώς αυτό περιέχει για κάθε προϊόν όλα τα προϊόντα που βρέθηκαν με αυτό, δηλαδή για ένα καλάθι με τρεις εγγραφές: cake, chocolate, tea έχουν δημιουργηθεί οι εξής εγγραφές: cake – chocolate, cake – tea, chocolate – cake, chocolate – tea, tea – cake, tea – chocolate. Αυτό συνέβη για λόγους σύγκρισης, να μπορούμε δηλαδή να έχουμε τα προϊόντα ανά δύο κάθε φορά και να έχουμε συσχέτιση για όλα τα προϊόντα με τα υπόλοιπα. Αντίστοιχα τα ποσοστά αφορούν πάλι το συγκεκριμένο σύνολο (των προϊόντων που υπήρξαν σε καλάθι με το “Mineral Water”).

Όσον αφορά τις πίτες – charts, το κάθε προϊόν από την πρώτη στήλη (Product_with_it) αντιστοιχεί σε ένα χρώμα το οποίο φαίνεται στην λεζάντα κάτω από τον πίνακα και τις πίτες. Επάνω στα charts αναγράφεται και το ποσοστό που κατέχει το κάθε προϊόν σε σχέση με το σύνολο που εξετάζουμε. Το πρώτο chart αφορά τον πίνακα στα αριστερά με τα προϊόντα ενώ το δεύτερο τον ίδιο πίνακα με τη διαφορά ότι η τιμή “Other” έχει εξαιρεθεί από το σύνολο ώστε να μελετήσουμε μόνο τα 20 που μας ενδιαφέρουν και να έχουμε μια πιο ξεκάθαρη εικόνα γι’ αυτά.

Παρατηρώντας τα γραφήματα και τον πίνακα βλέπουμε ότι υπερτερεί το προϊόν “Spaghetti” το οποίο όμως στην κατάταξη με τις συνολικές πωλήσεις που έκαναν όλα τα προϊόντα είχε πιο χαμηλή θέση. Αυτό είναι λογικό επειδή τα μακαρόνια είναι ένα προϊόν το οποίο για τις περισσότερες παρασκευές του, το νερό είναι απαραίτητο συστατικό. Ενώ αντίστοιχα για την σοκολάτα, το γάλα ή τα αυγά που βρίσκονται παρακάτω στην κατάταξη, το νερό δεν είναι απαραίτητο συστατικό. Γενικά μπορούμε να πούμε ότι ο πίνακας περιέχει προϊόντα τα οποία δεν αποκλίνουν ιδιαίτερα από τον πίνακα με τα «δημοφιλή» προϊόντα που είδαμε παραπάνω. Επιπλέον, δεν υπάρχει κάποια ακραία τιμή στα ποσοστά που τα προϊόντα κατέχουν ώστε να έχει στατιστικό ενδιαφέρον.

2. EGGS

Product_with_it	Count	Count(%)
mineral water	382	6.58%
spaghetti	274	4.72%
french fries	273	4.70%
chocolate	249	4.29%
milk	231	3.98%
burgers	216	3.72%
green tea	191	3.29%
frozen vegetables	163	2.81%
pancakes	163	2.81%
ground beef	150	2.58%
turkey	146	2.51%
cake	143	2.46%
low fat yogurt	126	2.17%
chicken	108	1.86%
shrimp	106	1.83%
herb & pepper	94	1.62%
tomatoes	92	1.58%
(blank)	90	1.55%
olive oil	90	1.55%
cooking oil	88	1.52%
<i>Other</i>	<i>2431</i>	<i>41.87%</i>
Grand Total	5806	100.00%



Εικόνα 11 – Eggs Set

Το συγκεκριμένο σετ αφορά το προϊόν “Eggs”. Βλέπουμε ότι ο πίνακας μας στα πρώτα τέσσερα προϊόντα είναι ίδιος με τον πίνακα με τις συνολικές αγορές όσον αφορά την κατάταξή τους και ότι αυτά κατέχουν μεγαλύτερα ποσοστά στο σύνολο τους σε σχέση με το “Mineral Water” που εξετάζαμε προηγουμένως.

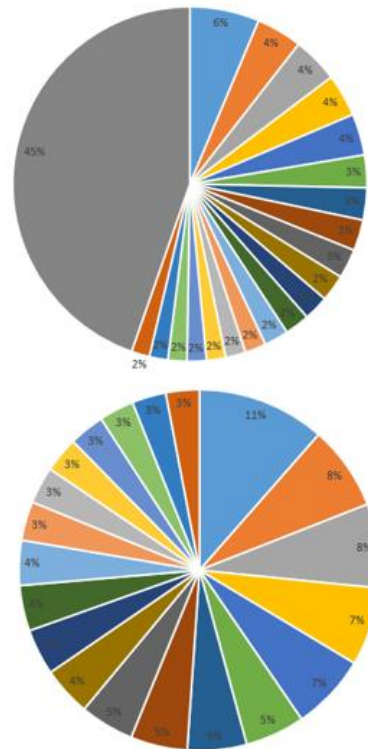
Το προϊόν “Herb & pepper” είναι ένα προϊόν το οποίο ήταν στην 26^η θέση στην κατάταξη με τις συνολικές αγορές όμως στον συγκεκριμένο πίνακα έχει ανέβει αρκετά (16^η θέση), το οποίο είναι λογικό αφού μιλάμε για ένα προϊόν το οποίο «ταιριάζει» με το προϊόν που εξετάζουμε (Eggs). Είναι ωστόσο περίεργο το γεγονός ότι υπάρχει στην λίστα μας αλλά δεν υπάρχει το πιπέρι ή το αλάτι τα οποία είναι απαραίτητα συστατικά και ειδικά όσον αφορά το συγκεκριμένο προϊόν. Θα μπορούσαμε λοιπόν να υποθέσουμε ότι υπάρχει τάση από τους καταναλωτές στα τρόφιμα τα οποία προσφέρουν έτοιμες λύσεις για μια παρασκευή γεύματος, όπως στην συγκεκριμένη περίπτωση το mix μπαχαρικών, και υψηλότερη προτίμηση τους σε σχέση με την πρώτη ύλη όπως το αλάτι ή το πιπέρι κτλ. που

θα χρειαστούν για την ίδια παρασκευή γεύματος. Επομένως, καλό θα ήταν τα προϊόντα που προωθούνται να προσφέρουν έξυπνες λύσεις για «νόστιμες» παρασκευές και να μειώνουν τον χρόνο παρασκευής ενός πιάτου.

Τέλος, παρατηρούμε ότι στον συγκεκριμένο πίνακα βρέθηκε η τιμή “(blank)” ή οποία εκφράζει όλες τις φορές τις οποίες έχει καταγραφεί το προϊόν χωρίς κανένα άλλο προϊόν στο καλάθι. Το προϊόν “Eggs” ήταν πολύ υψηλά στην κατάταξη στον πίνακα “Solo Products” όπως και το προϊόν “Mineral Water” το οποίο ξεπερνούσε σε πωλήσεις το προϊόν μας. Ωστόσο, στο παραπάνω προϊόν δεν ήταν τόσο μεγάλος ο αριθμός που αυτό πουλήθηκε μόνο του σε σχέση με σύνολο του όπως στο προϊόν μας. Μιλάμε δηλαδή για έναν μεγάλο αριθμό καταναλωτών οι οποίοι έρχονται στο κατάστημα παίρνουν το προϊόν “Eggs” και φεύγουν με άδειο καλάθι. Θα μπορούσαμε να περιορίσουμε την τάση αυτήν βάζοντας προσφορές σε προϊόντα που σχετίζονται με το συγκεκριμένο προϊόν όταν αυτά αγοράζονται μαζί, όπως το “Herbs & Pepper” το οποίο από μόνο του δεν «πουλάει» τόσο πολύ. Ακόμη, θα μπορούσαμε να τοποθετούμε τα συγκεκριμένα προϊόντα που θέλουμε να προωθήσουμε και συσχετίζονται με το προϊόν μας στα καλάθια αγορών, σε κοντινά σημεία στο κατάστημα ώστε να είναι ευδιάκριτα στον πελάτη που ετοιμάζεται να αγοράσει το προϊόν “Eggs”.

3. SPAGHETTI

Product_with_it	Count	Count(%)
mineral water	448	6.36%
chocolate	295	4.18%
ground beef	294	4.17%
eggs	274	3.89%
milk	266	3.77%
frozen vegetables	209	2.96%
french fries	207	2.94%
green tea	199	2.82%
pancakes	189	2.68%
olive oil	172	2.44%
burgers	161	2.28%
shrimp	159	2.26%
tomatoes	157	2.23%
cake	136	1.93%
chicken	129	1.83%
grated cheese	124	1.76%
turkey	124	1.76%
herb & pepper	122	1.73%
cooking oil	119	1.69%
frozen smoothie	117	1.66%
Other	3148	44.66%
Grand Total	7049	100.00%

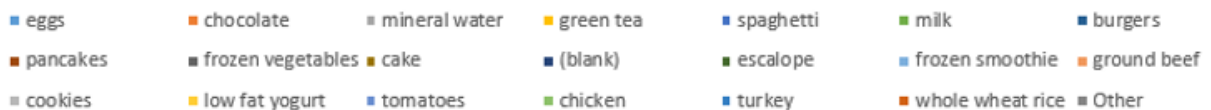
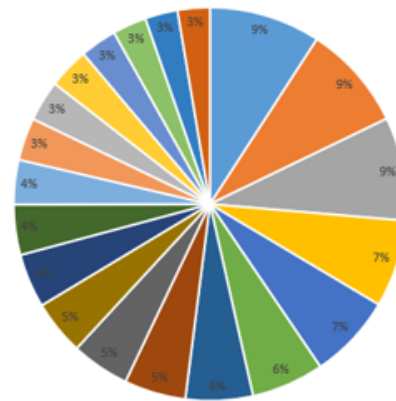
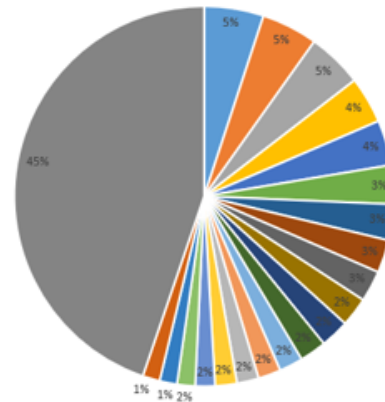


Εικόνα 12 – Spaghetti Set

Το παραπάνω σεν αφορά το προϊόν “Spaghetti”. Τα προϊόντα που βλέπουμε στον πίνακα είναι κορυφαία στην κατάταξη στον πίνακα με τις συνολικές πωλήσεις και τα ποσοστά που καταλαμβάνουν πάλι δεν έχουν κάποιο στατιστικό ενδιαφέρον. Το μόνο που θα μπορούσαμε να παρατηρήσουμε εδώ είναι ότι η τιμή “Other” είναι μεγαλύτερη σε σχέση με τα παραπάνω προϊόντα (45%) το οποίο είναι σημαντικό αφού υπάρχει καλύτερη κατανομή στις αγορές των υπόλοιπων προϊόντων του καταστήματος και όχι περισσότερο των “δημοφιλών” συγκριτικά με τα δύο προηγούμενα προϊόντα.

4. FRENCH FRIES

Product_with_it	Count	Count(%)
eggs	273	5.07%
chocolate	258	4.79%
mineral water	253	4.70%
green tea	214	3.97%
spaghetti	207	3.84%
milk	178	3.30%
burgers	165	3.06%
pancakes	151	2.80%
frozen vegetables	143	2.65%
cake	134	2.49%
(blank)	132	2.45%
escalope	123	2.28%
frozen smoothie	109	2.02%
ground beef	104	1.93%
cookies	100	1.86%
low fat yogurt	100	1.86%
tomatoes	90	1.67%
chicken	83	1.54%
turkey	80	1.49%
whole wheat rice	79	1.47%
Other	2411	44.76%
Grand Total	5387	100.00%



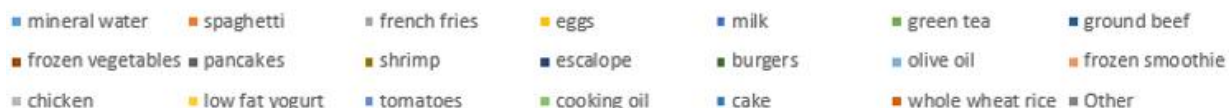
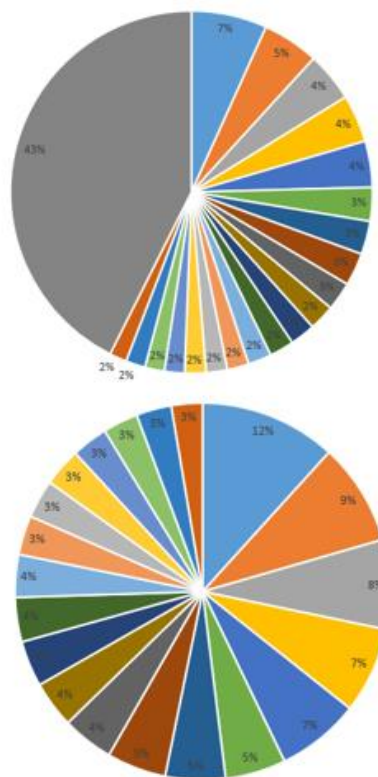
Εικόνα 13 – French Fries Set

Το παραπάνω σεν αφορά το προϊόν “French Fries”. Πάλι δεν έχουμε κάποια μεγάλη απόκλιση από τον πίνακα με τις συνολικές πωλήσεις. Η διαφορά εδώ είναι ότι ποσοστικά δεν υπάρχει ένα προϊόν που ξεχωρίζει αλλά τρία (9%) και τα επόμενα δύο σε πωλήσεις είναι πολύ κοντά (μόλις 2%) σε σχέση με τα παραπάνω προϊόντα που είχαμε το προϊόν “Mineral Water” να «ισοπεδώνει» τα υπόλοιπα προϊόντα. Πάλι μπορούμε να πούμε ότι έχουμε καλύτερη κατανομή πωλήσεων.

Βλέπουμε ότι και στο συγκεκριμένο προϊόν υπάρχει η τιμή “(blank)” και μάλιστα αρκετά ψηλά στην κατάταξη, επομένως θα πρέπει να προσέξουμε κι εδώ το προϊόν μας. Εφόσον και στο προϊόν “Eggs” αλλά και στο προϊόν που εξετάζουμε τώρα παρατηρήθηκε το ίδιο φαινόμενο και εφόσον και τα δύο προϊόντα οι περισσότεροι πελάτες επιλέγουν να τα αγοράζουν μαζί, θα μπορούσαν να γίνουν προσφορές που να αφορούν τα δύο προϊόντα συνδυαστικά.

5. CHOCOLATE

Product_with_it	Count	Count(%)
mineral water	396	6.76%
spaghetti	295	5.04%
french fries	258	4.41%
eggs	249	4.25%
milk	241	4.12%
green tea	176	3.01%
ground beef	174	2.97%
frozen vegetables	172	2.94%
pancakes	149	2.54%
shrimp	135	2.31%
escalope	132	2.25%
burgers	128	2.19%
olive oil	123	2.10%
frozen smoothie	113	1.93%
chicken	111	1.90%
low fat yogurt	111	1.90%
tomatoes	105	1.79%
cooking oil	102	1.74%
cake	102	1.74%
whole wheat rice	90	1.54%
Other	2493	42.58%
Grand Total	5855	100.00%

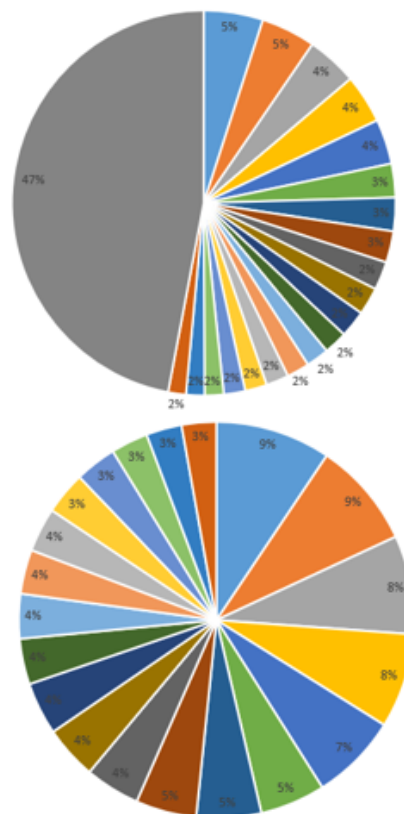


Εικόνα 14 – Chocolate Set

Το παραπάνω σεν αφορά το προϊόν “Chocolate”, ένα αρκετά δημοφιλές προϊόν εφόσον δεν είναι αναγκαίο προϊόν ωστόσο οι πελάτες το προτιμούν. Η παρατήρηση που μπορούμε να κάνουμε εδώ είναι ότι σε σχέση με τα παραπάνω προϊόντα που εξετάσαμε το ποσοστό του προϊόντος “Mineral Water” κατέχει την υψηλότερη τιμή (12%). Στο αντίστοιχο σεν του προϊόντος “Mineral Water” φαίνεται κορυφαίο προϊόν να είναι το προϊόν “Spaghetti” και όχι το προϊόν που εξετάζουμε. Αυτό όμως οφείλεται στο γεγονός ότι το προϊόν “Spaghetti” κατέχει μεγαλύτερες πωλήσεις. Επομένως, μπορούμε να υποθέσουμε ότι οι πελάτες που αγοράζουν το προϊόν “Mineral Water” έχουν μεγαλύτερη πιθανότητα να αγοράσουν το προϊόν “Chocolate” από ότι προϊόν “French Fries”.

6. GREEN TEA

Product_with_it	Count	Count(%)
mineral water	233	5.00%
french fries	214	4.59%
spaghetti	199	4.27%
eggs	191	4.10%
chocolate	176	3.77%
milk	132	2.83%
burgers	131	2.81%
pancakes	123	2.64%
ground beef	111	2.38%
frozen vegetables	108	2.32%
cake	106	2.27%
tomatoes	92	1.97%
cookies	90	1.93%
turkey	89	1.91%
chicken	89	1.91%
shrimp	86	1.84%
frozen smoothie	84	1.80%
olive oil	75	1.61%
(blank)	72	1.54%
honey	70	1.50%
Other	2193	47.02%
Grand Total	4664	100.00%



Εικόνα 15 – Green Tea Set

Το παραπάνω σεν αφορά το προϊόν “Green tea”. Βλέπουμε ότι πάλι πρωταγωνιστές είναι πάλι τα κορυφαία προϊόντα από τον πίνακα με τις συνολικές πωλήσεις και ότι πάλι δεν υπάρχει μεγάλη απόκλιση από αυτόν.

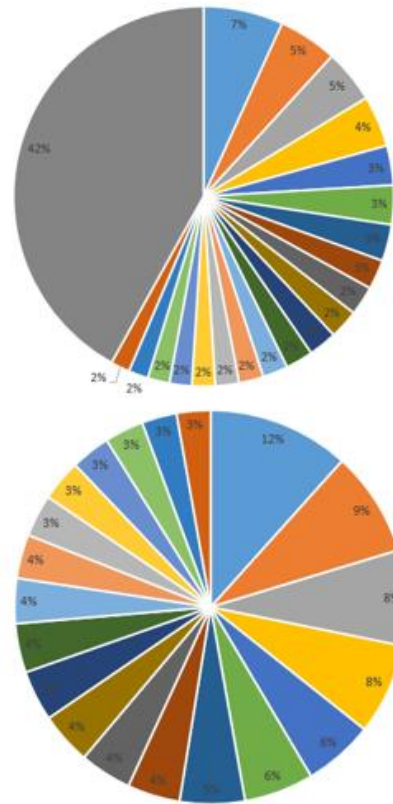
Το περίεργο ίσως είναι ότι αυτήν τη φορά έχουμε και το προϊόν “honey” που βρίσκεται στην τελευταία θέση, το οποίο είναι πιο χαμηλά στην κατάταξη (συγκεκριμένα θέση 27) και εμφανίζεται για πρώτη φορά στα 20 πρώτα προϊόντα. Παρατηρώντας και τις πωλήσεις του προϊόντος “honey” βλέπουμε ότι ένα ποσοστό περίπου 20% από τις συνολικές πωλήσεις του ήταν μαζί με το προϊόν “Green tea”. Άρα μπορούμε να υποθέσουμε ότι τα δύο προϊόντα πωλούνται συχνά μαζί.

Εξίσου παράξενο είναι το γεγονός ότι το προϊόν μας βρίσκεται στην 6^η θέση στον πίνακα με τις πωλήσεις, ενώ αντίστοιχα το προϊόν “tea” βρίσκεται στην 115^η θέση με ελάχιστες πωλήσεις. Θα μπορούσαμε λοιπόν να υποθέσουμε ότι οι καταναλωτές προτιμούν πιο σύνθετες γεύσεις όσον αφορά το συγκεκριμένο ρόφημα, αφού επιλέγουν να αγοράσουν πράσινο τσάι και όχι τσάι.

Τέλος, πάλι βρέθηκε η τιμή “(blank)”, οπότε θα πρέπει να δώσουμε προσοχή. Σε αυτή την περίπτωση καλό θα ήταν να προωθηθεί το προϊόν “honey”, αφού τα δύο προϊόντα έχουν καλή συσχέτιση μεταξύ τους και το μέλι είναι ένα προϊόν που χρειάζεται να «ανέβει» περισσότερο από τα υπόλοιπα.

7. MILK

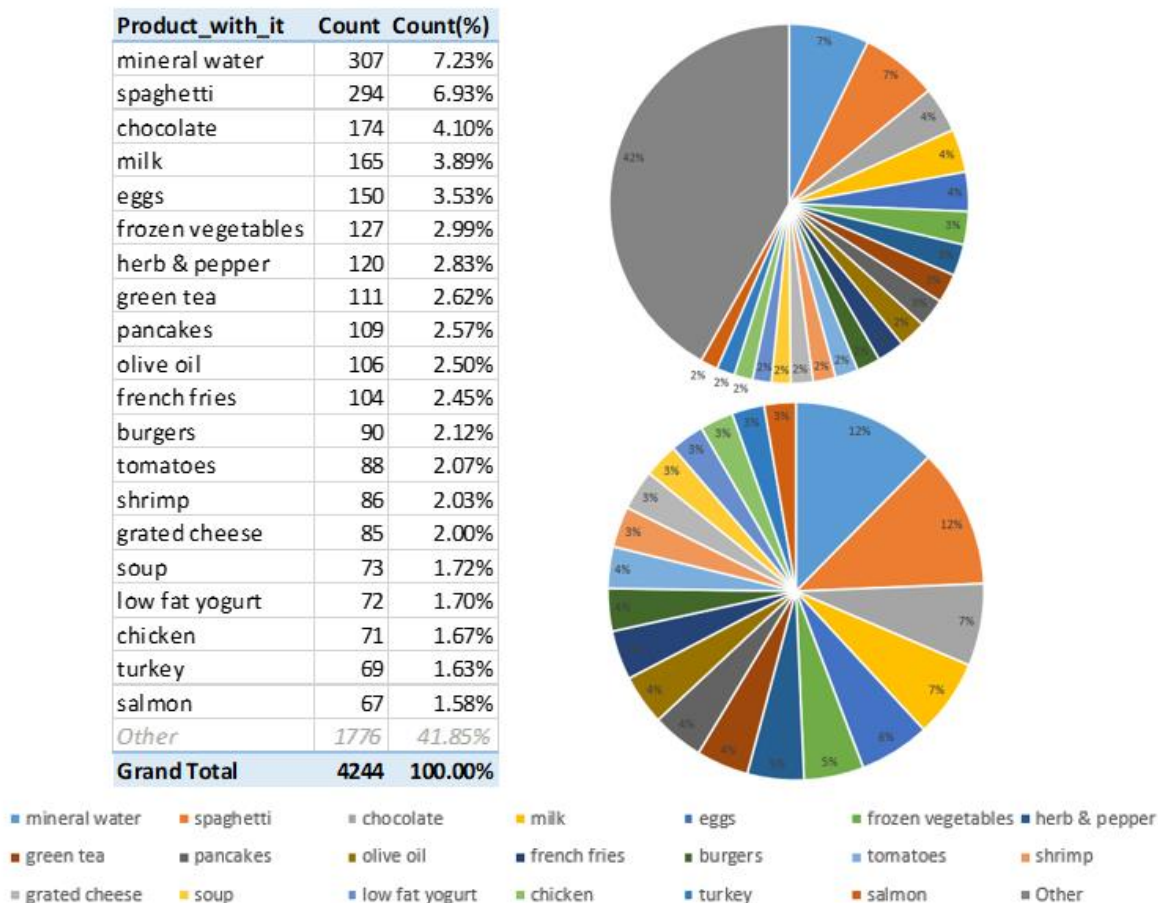
Product_with_it	Count	Count(%)
mineral water	360	6.79%
spaghetti	266	5.02%
chocolate	241	4.54%
eggs	231	4.36%
french fries	178	3.36%
frozen vegetables	177	3.34%
ground beef	165	3.11%
burgers	134	2.53%
green tea	132	2.49%
shrimp	132	2.49%
olive oil	128	2.41%
pancakes	124	2.34%
soup	114	2.15%
chicken	111	2.09%
frozen smoothie	107	2.02%
tomatoes	105	1.98%
cake	100	1.89%
low fat yogurt	99	1.87%
whole wheat rice	89	1.68%
cooking oil	86	1.62%
Other	2225	41.95%
Grand Total	5304	100.00%



Εικόνα 16 – Milk Set

Το παραπάνω σεν αφορά το προϊόν “Milk” και όπως μπορούμε να παρατηρήσουμε είναι παρόμοιο με τους υπόλοιπους πίνακες. Πάλι η ταξινόμηση των προϊόντων είναι παρόμοια με αυτήν του πίνακα συνολικών πωλήσεων, με το προϊόν “Mineral Water” να πρωταγωνιστεί. Δεν παρατηρείται κάποια καινούρια τιμή στον πίνακα ώστε να αξίζει να σχολιαστεί και στα αποτελέσματα δεν υπάρχει κάποιο στατιστικό ενδιαφέρον.

8. GROUND BEEF



Εικόνα 17 – Ground Beef Set

Το παραπάνω σετ αφορά το προϊόν “ground beef”. Παρατηρώντας και τον πίνακα των συνολικών πωλήσεων μπορούμε να πούμε πως το προϊόν μας είναι το κορυφαίο προϊόν σε πωλήσεις ανάμεσα στα προϊόντα κρέατος. Βλέπουμε πως τα προϊόντα “chicken” και “turkey” είναι πολύ χαμηλότερα στην κατάταξη στον πίνακα με τις συνολικές πωλήσεις αλλά ακόμη και στον παραπάνω πίνακα όπου έχουν «πέσει» στην κατάταξη σε σχέση με τους υπόλοιπους πίνακες που αφορούσαν παραπάνω προϊόντα. Επομένως μπορούμε να βγάλουμε δύο συμπεράσματα από αυτό:

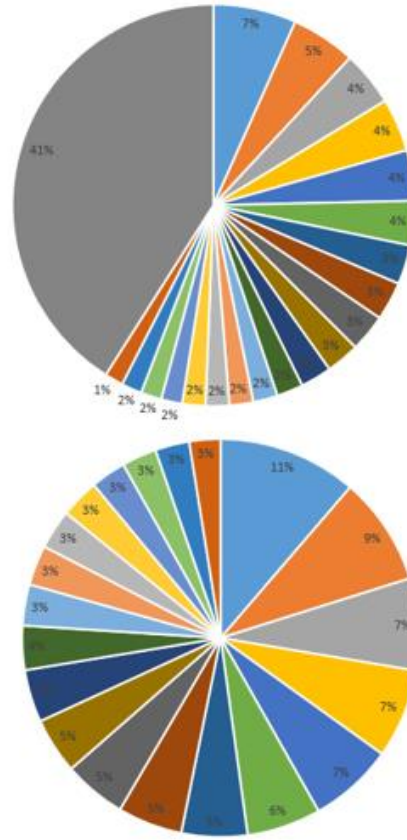
- Το κρέας που προτιμούν να αγοράζουν οι πελάτες είναι ο κιμάς.
- Οι πελάτες που αγοράζουν κιμά, δεν έχουν μεγάλη πιθανότητα να αγοράσουν και άλλο προϊόν του ίδιου είδους (κρέας).

Εξίσου σημαντικό, είναι το γεγονός ότι για πρώτη φορά μέχρι στιγμής έχουμε ένα προϊόν το οποίο φτάνει ποσοστικά το προϊόν “Mineral Water”. Στην συγκεκριμένη περίπτωση έχουμε το προϊόν “spaghetti” το οποίο είναι λογικό να «ανέβει» τόσο πολύ αφού υπάρχει άμεση συσχέτιση με το προϊόν που εξετάζουμε. Τα μακαρόνια με κιμά είναι ίσως το πιο δημοφιλές και αγαπημένο πιάτο και απαιτεί για την παρασκευή του τα δύο προϊόντα επομένως εξηγεί και τις υψηλές πωλήσεις συνδυαστικά των δύο.

Ανεβασμένα είναι και τα προϊόντα “grated cheese” και “herb & pepper” σε σχέση με τις κατατάξεις τους στα προηγούμενα σελ που εξετάσαμε. Αυτό είναι λογικό αφού μιλάμε για προϊόντα τα οποία είναι εξαρτώμενα μεταξύ τους για συγκεκριμένες παρασκευές (πχ. μακαρόνια με κιμά που προαναφέραμε). Άρα μπορούμε να συμπεράνουμε ότι υπάρχει μεγάλη τάση από πελάτες να γεμίζουν το καλάθι τους με προϊόντα τα οποία θα χρησιμοποιηθούν άμεσα για την παρασκευή ενός «γρήγορου» γεύματος ημέρας.

9. FROZEN VEGETABLES

Product_with_it	Count	Count(%)
mineral water	268	6.68%
spaghetti	209	5.21%
milk	177	4.41%
chocolate	172	4.29%
eggs	163	4.06%
french fries	143	3.56%
ground beef	127	3.17%
shrimp	125	3.12%
tomatoes	121	3.02%
green tea	108	2.69%
pancakes	101	2.52%
olive oil	85	2.12%
burgers	79	1.97%
cake	77	1.92%
low fat yogurt	76	1.89%
frozen smoothie	75	1.87%
whole wheat rice	68	1.69%
chicken	67	1.67%
turkey	66	1.65%
soup	60	1.50%
Other	1645	41.00%
Grand Total	4012	100.00%

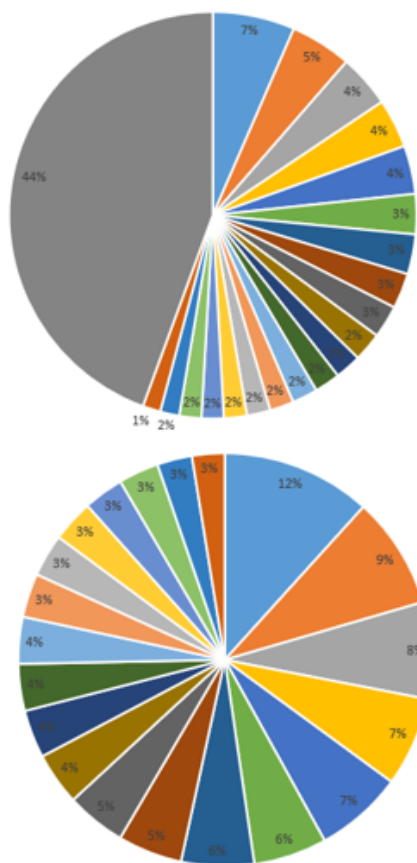


Εικόνα 18 – Frozen Vegetables Set

Το παραπάνω σεν αφορά το προϊόν “frozen vegetables”. Τα αποτελέσματα για το συγκεκριμένο προϊόν είναι παρεμφερή με τα υπόλοιπα προϊόντα. Ο πίνακας μοιάζει με τον πίνακα με τις συνολικές πωλήσεις όσον αφορά τα προϊόντα που βλέπουμε και δεν έχουμε κάποιο καινούριο προϊόν ή κάποια αύξηση ή μείωση στα ποσοστά των προϊόντων σε βαθμό που να αξίζει να σημειωθεί.

10. PANCAKES

Product_with_it	Count	Count(%)
mineral water	253	6.52%
spaghetti	189	4.87%
eggs	163	4.20%
french fries	151	3.89%
chocolate	149	3.84%
milk	124	3.20%
green tea	123	3.17%
ground beef	109	2.81%
frozen vegetables	101	2.60%
cake	89	2.29%
olive oil	81	2.09%
shrimp	79	2.04%
burgers	79	2.04%
tomatoes	74	1.91%
turkey	72	1.86%
low fat yogurt	71	1.83%
chicken	68	1.75%
frozen smoothie	68	1.75%
cooking oil	59	1.52%
herb & pepper	55	1.42%
Other	1724	44.42%
Grand Total	3881	100.00%



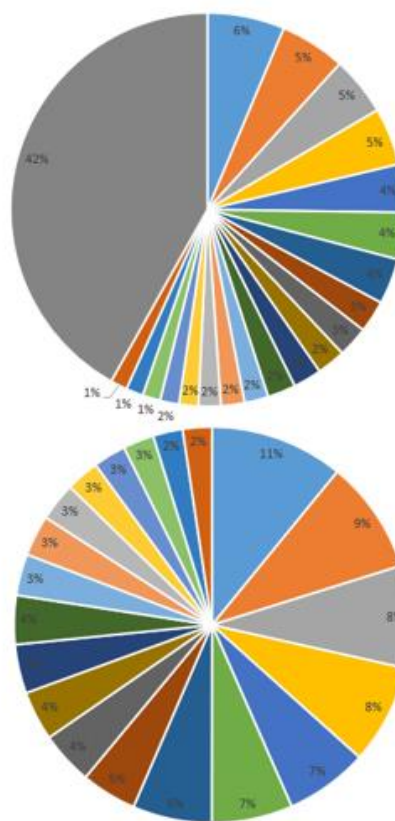
Εικόνα 19 – Pancakes Set

Το παραπάνω σεν αφορά το προϊόν “pancakes”. Όπως και στις προηγούμενες περιπτώσεις πάλι έχουμε ταύτιση με τον πίνακα με τις συνολικές πωλήσεις. Η μόνη διαφορά εδώ είναι ότι εμφανίζεται πάλι το προϊόν “herb & pepper” που βρίσκεται στην τελευταία θέση.

Στα παραπάνω σεν έχει παρατηρηθεί ότι το προϊόν “olive oil” είναι πάντα πολύ ψηλά και το προϊόν “cooking oil” που είναι παρόμοιο προϊόν, εμφανίζεται είτε ανάμεσα στα τελευταία της κατάταξης, είτε καθόλου. Στο συγκεκριμένο παράδειγμα υπάρχει στο σύνολο μας και αυτό είναι λογικό αφού είναι αναγκαίο για την παρασκευή αρκετών γευμάτων που περιέχουν το προϊόν “pancakes”.

11. BURGERS

Product_with_it	Count	Count(%)
eggs	216	6.35%
mineral water	183	5.38%
french fries	165	4.85%
spaghetti	161	4.73%
milk	134	3.94%
green tea	131	3.85%
chocolate	128	3.76%
ground beef	90	2.64%
cake	86	2.53%
turkey	80	2.35%
frozen vegetables	79	2.32%
pancakes	79	2.32%
whole wheat rice	67	1.97%
shrimp	66	1.94%
tomatoes	62	1.82%
frozen smoothie	54	1.59%
escalope	53	1.56%
low fat yogurt	50	1.47%
olive oil	48	1.41%
soup	47	1.38%
Other	1425	41.86%
Grand Total	3404	100.00%



Εικόνα 20 – Burgers Set

Το παραπάνω σεν αφορά το προϊόν “burgers”. Δεν μπορούμε να είμαστε σίγουροι εάν το συγκεκριμένο προϊόν είναι το μπιφτέκι ή αν είναι έτοιμα burger που πουλάει το κατάστημα. Δεδομένου όμως των μεγάλων πωλήσεων, το πιο πιθανό είναι να είναι το μπιφτέκι. Υποθέτοντας λοιπόν ότι το προϊόν μας είναι έτοιμα μπιφτέκια μπορούμε να παρατηρήσουμε ότι ενώ το προϊόν μας υπάρχει πάντα στις λίστες με τα κορυφαία προϊόντα και δεν ξεπερνάει σε πωλήσεις ποτέ τον μοσχαρίσιο κιμά από το οποίο μπορεί να παρασκευαστεί. Ωστόσο, οι πωλήσεις των προϊόντων είναι πολύ κοντά. Συγκεκριμένα:

- Ground beef: 737 πωλήσεις, 2.51% από το σύνολο
- Burgers: 654 πωλήσεις, 2.23% από το σύνολο

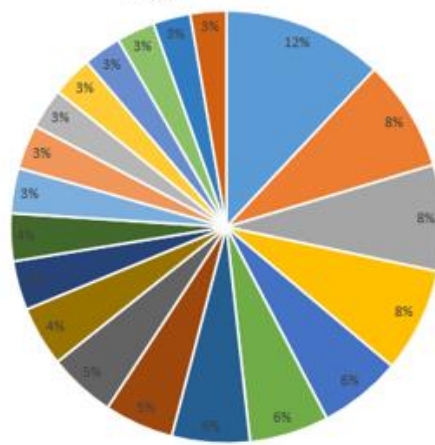
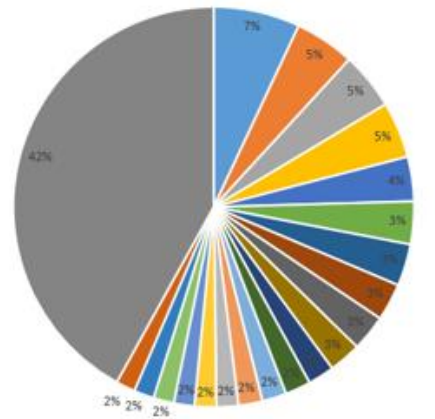
Άρα μπορούμε να υποθέσουμε ότι οι πελάτες προτιμούν να αγοράζουν έτοιμα μπιφτέκια παρά να μπαίνουν στην διαδικασία να τα φτιάξουν μόνοι τους.

Αν υποθέσουμε ότι το προϊόν μας είναι έτοιμα burger που παρέχει το κατάστημά στους πελάτες του τότε μιλάμε για ένα πολύ αποδοτικό προϊόν αφού οι πωλήσεις του βρίσκεται στα ύψη, παράλληλα με αναγκαία οικιακά προϊόντα.

Εξίσου σημαντικό είναι το γεγονός ότι το προϊόν “Eggs” σε αυτό το σετ ξεπέρασε το προϊόν “Mineral water” και πρέπει να σημειωθεί ότι μέχρι στιγμής αυτή είναι η πρώτη φορά που κάποιο από τα προϊόντα το ξεπερνάει. Ωστόσο πρόκειται για το δεύτερο προϊόν στην κατάταξη με τις πωλήσεις το οποίο είναι λογικό. Το γεγονός ότι ξεπέρασε το “Mineral Water” σε πωλήσεις και δεδομένου της γνώσης που αποκτήσαμε ότι στην συγκεκριμένη κοινωνία το εμφιαλωμένο νερό είναι αναγκαίο συστατικό μπορούμε να πούμε ότι υπάρχει τεράστια συσχέτιση μεταξύ των δύο προϊόντων. Μπορούμε λοιπόν να συμπεράνουμε ότι σε πολύ μεγάλο πλήθος καλαθιών που θα υπάρχει το προϊόν “burger”, θα υπάρχει και το προϊόν “eggs”.

12. CAKE

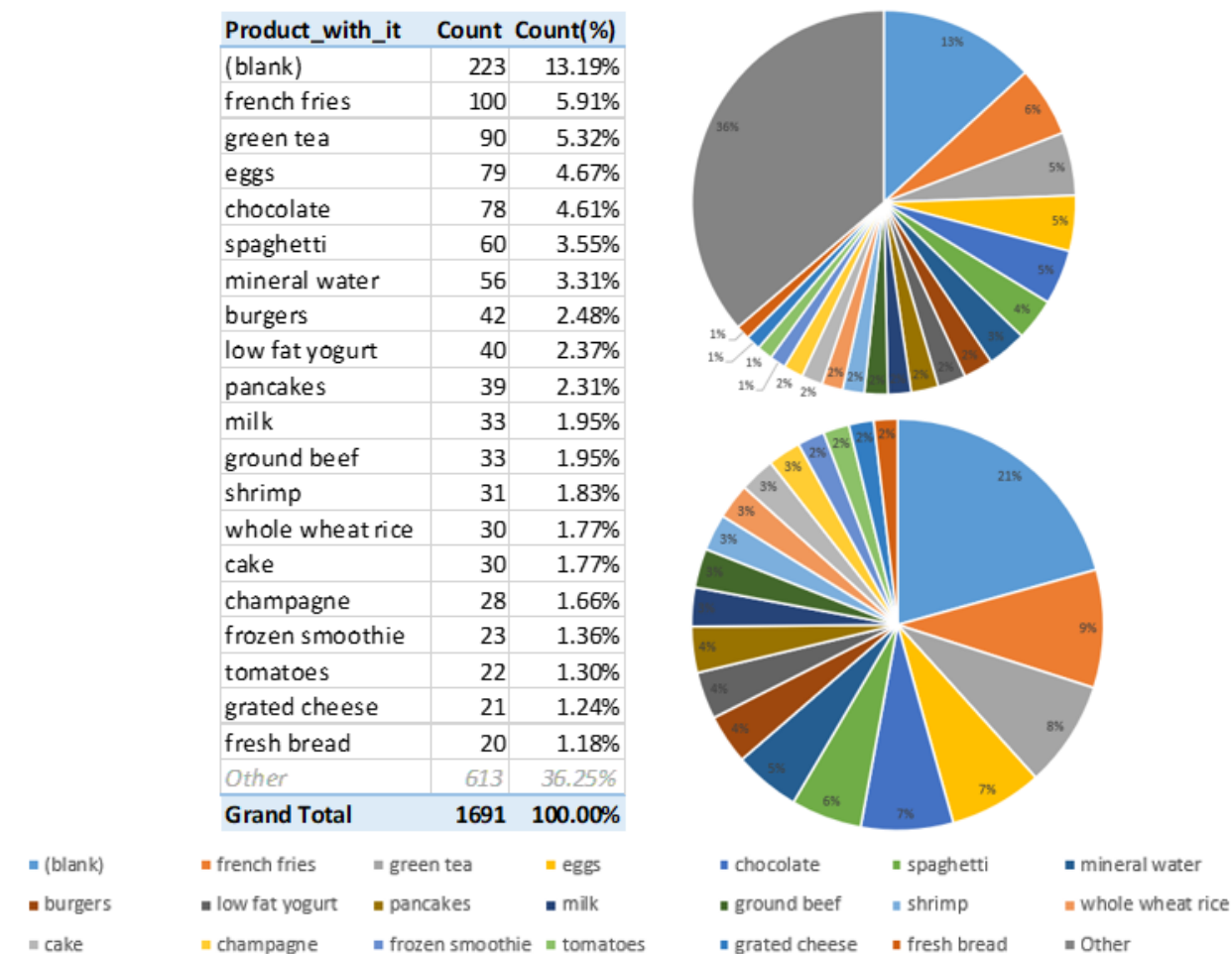
Product_with_it	Count	Count(%)
mineral water	206	6.98%
eggs	143	4.84%
spaghetti	136	4.61%
french fries	134	4.54%
green tea	106	3.59%
chocolate	102	3.45%
milk	100	3.39%
pancakes	89	3.01%
burgers	86	2.91%
frozen vegetables	77	2.61%
tomatoes	63	2.13%
ground beef	61	2.07%
whole wheat rice	58	1.96%
shrimp	57	1.93%
low fat yogurt	52	1.76%
turkey	52	1.76%
chicken	49	1.66%
olive oil	49	1.66%
frozen smoothie	48	1.63%
soup	46	1.56%
Other	1239	41.96%
Grand Total	2953	100.00%



Εικόνα 21 – Cake Set

Το παραπάνω σετ αφορά το προϊόν “cake”. Τα αποτελέσματα που βλέπουμε είναι παρεμφερή με τα αποτελέσματα των παραπάνω προϊόντων που εξετάσαμε (δεν υπάρχει μεγάλη απόκλιση από τον πίνακα με τις συνολικές πωλήσεις ούτε έχουμε κάποια καινούρια τιμή).

13. COOKIES



Εικόνα 22 – Cookies Set

Το παραπάνω σεν αφορά το προϊόν “cookies” και σίγουρα αποτελεί το πιο «παράξενο» σεν που έχουμε εξετάσει μέχρι στιγμής.

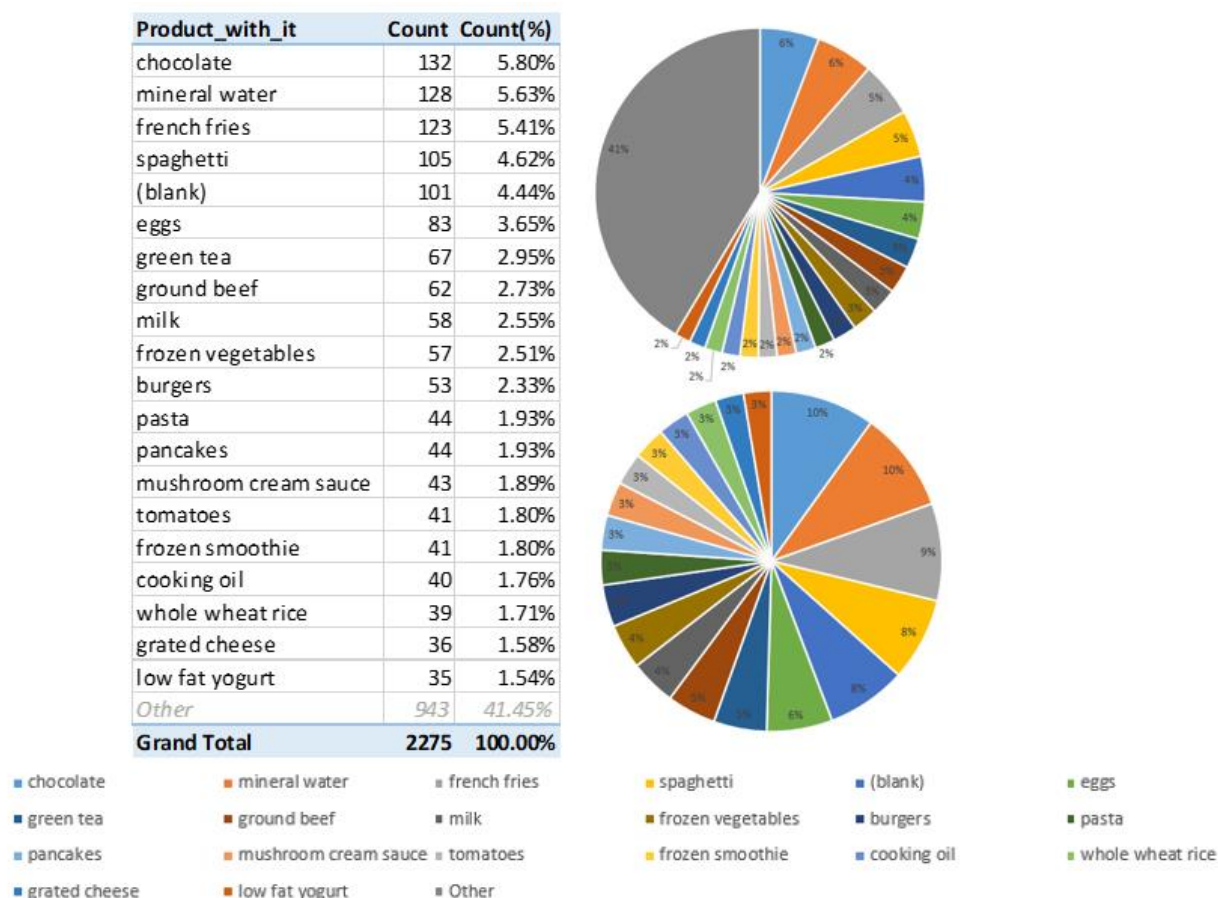
Αρχικά, για πρώτη φορά έχουμε προϊόν που υπερτερούν οι φορές που αγοράστηκε μόνο του σε σχέση με τις συνδυαστικές αγορές με τα υπόλοιπα προϊόντα. Αυτό θα πρέπει να το περιορίσουμε και να εφαρμόσουμε τακτικές όπως αυτές που προαναφέρθηκαν για τα υπόλοιπα προϊόντα που βρέθηκε η τιμή “(blank)” στα αποτελέσματά τους.

Παρατηρούμε επίσης πως το προϊόν “Mineral water” που πάντα «κυριαρχούσε» σε όλα τα υπόλοιπα σεν τώρα έχει κατέβει πολύ. Αυτό εξηγείται από το γεγονός ότι το συγκεκριμένο προϊόν έχει κάνει ένα πολύ μεγάλο πλήθος πωλήσεων μόνο του, όμως δεν δικαιολογεί το γεγονός ότι και πολλά άλλα προϊόντα έχουν ξεπεράσει το εμφιαλωμένο νερό (πατάτες, τσάι, αυγά, σοκολάτα, μακαρόνια). Θα μπορούσαμε ίσως να υποθέσουμε ότι οι πελάτες που αγοράζουν μπισκότα, δεν αγοράζουν τόσο πολύ εμφιαλωμένο νερό σε σχέση

με τους πελάτες που δεν αγοράζουν μπισκότα. Η συγκεκριμένη υπόθεση δεν μπορεί να στηριχθεί κάπου αλλού οπότε θα πρέπει να επανεξεταστεί και άλλες φορές για να αποτελεί συμπέρασμα.

Επιπλέον, στο συγκεκριμένο σετ βλέπουμε δυο προϊόντα για πρώτη φορά στα κορυφαία 20. Αυτά είναι τα προϊόντα “champagne” και “fresh bread” τα οποία είναι βέβαια σχετικά ψηλά στην κατάταξη στον πίνακα με τις συνολικές πωλήσεις. Αφού ως τώρα δεν τα έχουμε ξανασυναντήσει θα μπορούσαμε να υποθέσουμε ότι υπάρχει συσχέτιση με τα δύο προϊόντα και το προϊόν μας.

14. ESCALOPE



Εικόνα 23 – Escalope Set

Το παραπάνω σετ αφορά το προϊόν “escalope”. Παρατηρούμε ότι στο συγκεκριμένο σετ το προϊόν “chocolate” έχει βγει πρώτο σε πωλήσεις και πως έχει ξεπεράσει το προϊόν “mineral water” που είναι το «κορυφαίο» μας προϊόν σχεδόν σε όλα τα σετ. Οι πωλήσεις των δύο προϊόντων δεν έχουν πολύ μεγάλη διαφορά οπότε το γεγονός ότι η σοκολάτα ξεπέρασε το εμφιαλωμένο νερό θα μπορούσε να είναι ένα τυχαίο γεγονός και για να αποτελεί συμπέρασμα θα πρέπει να επανεξεταστούν τα δύο προϊόντα σε σχέση με το προϊόν που εξετάζουμε. Ωστόσο, η σοκολάτα συνήθως βρίσκεται πιο χαμηλά στην κατάταξη στα σετ, πράγμα που δεν συμβαίνει εδώ. Θα μπορούσαμε να μιλήσουμε λοιπόν για μεγάλη συσχέτιση μεταξύ των δύο προϊόντων.

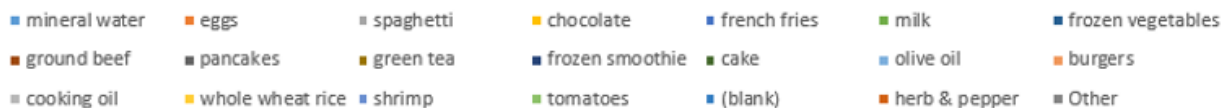
Επιπλέον, παρατηρούμε ότι στο συγκεκριμένο σετ έχουμε τα προϊόντα “pasta” και “mushroom cream sauce” τα οποία όχι μόνο εμφανίζονται για πρώτη φορά, αλλά βρίσκονται και αρκετά χαμηλά στην κατάταξη στον πίνακα με τις συνολικές πωλήσεις πράγμα που σημαίνει ότι το μεγαλύτερο μέρος των πωλήσεων αυτών των προϊόντων οφείλεται στο προϊόν “escalope”.

Εξίσου σημαντικό είναι το γεγονός ότι η τιμή “(blank)” βρέθηκε το οποίο πρέπει να μας προβληματίζει. Από τη μία πλευρά τα γεγονόσ ότι το προϊόν μας πουλάει μόνο του είναι σημαντικό γιατί καταλαβαίνουμε ότι

μιλάμε για ένα δημοφιλές προϊόν. Θα πρέπει όμως να εφαρμόσουμε τακτικές όπως αυτές που αναφέρθηκαν στα παραπάνω σεν για να περιορίσουμε όσο μπορούμε αυτό το γεγονός.

15. LOW FAT YOGURT

Product_with_it	Count	Count(%)
mineral water	180	6.83%
eggs	126	4.78%
spaghetti	114	4.33%
chocolate	111	4.21%
french fries	100	3.80%
milk	99	3.76%
frozen vegetables	76	2.89%
ground beef	72	2.73%
pancakes	71	2.70%
green tea	57	2.16%
frozen smoothie	55	2.09%
cake	52	1.97%
olive oil	51	1.94%
burgers	50	1.90%
cooking oil	49	1.86%
whole wheat rice	49	1.86%
shrimp	47	1.78%
tomatoes	45	1.71%
(blank)	45	1.71%
herb & pepper	44	1.67%
Other	1141	43.32%
Grand Total	2634	100.00%

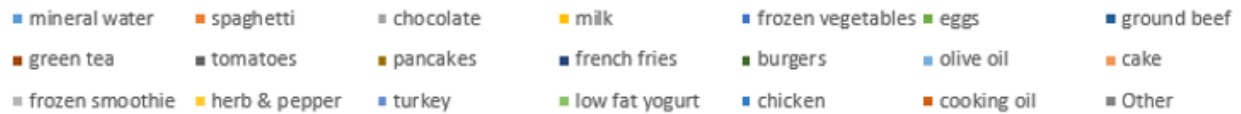
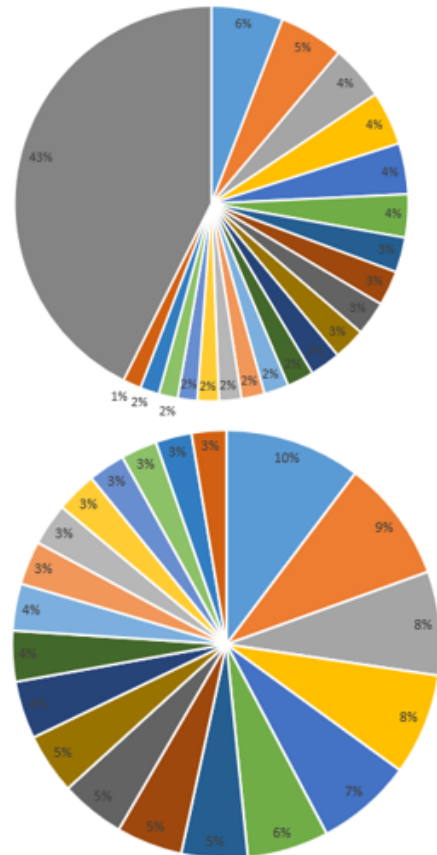


Εικόνα 24 – Low Fat Yogurt Set

Το παραπάνω σετ αφορά το προϊόν “low fat yogurt”. Το μόνο που μπορούμε να παρατηρήσουμε εδώ είναι ότι πάλι βρέθηκε η τιμή “(blank)”. Ανεξάρτητα από αυτό η ταξινόμηση των προϊόντων που έχουμε εδώ είναι «καλή» και παρόμοια με τον πίνακα με τις συνολικές πωλήσεις.

16. SHRIMP

Product_with_it	Count	Count(%)
mineral water	177	5.89%
spaghetti	159	5.29%
chocolate	135	4.50%
milk	132	4.40%
frozen vegetables	125	4.16%
eggs	106	3.53%
ground beef	86	2.86%
green tea	86	2.86%
tomatoes	84	2.80%
pancakes	79	2.63%
french fries	75	2.50%
burgers	66	2.20%
olive oil	61	2.03%
cake	57	1.90%
frozen smoothie	55	1.83%
herb & pepper	53	1.76%
turkey	48	1.60%
low fat yogurt	47	1.57%
chicken	47	1.57%
cooking oil	45	1.50%
Other	1280	42.62%
Grand Total	3003	100.00%

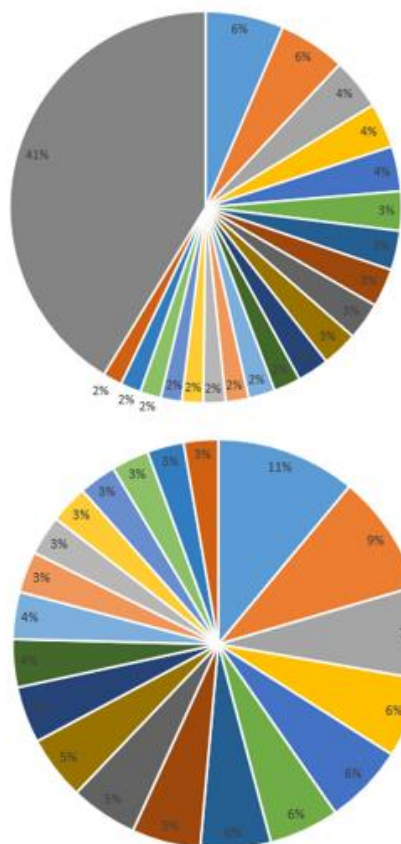


Εικόνα 25 – Shrimp Set

Το παραπάνω σετ αφορά το προϊόν “shrimp”. Τα αποτελέσματα που βλέπουμε εδώ δεν έχουν μεγάλη απόκλιση από τον πίνακα με τις συνολικές πωλήσεις και δεν έχουμε κάποια καινούρια τιμή. Είναι ένα σετ το οποίο μας δίνει μια «κανονική» εικόνα, δηλαδή αυτό που έχουμε συνηθίσει να βλέπουμε μέχρι τώρα.

17. TOMATOES

Product_with_it	Count	Count(%)
mineral water	183	6.46%
spaghetti	157	5.54%
frozen vegetables	121	4.27%
milk	105	3.71%
chocolate	105	3.71%
eggs	92	3.25%
green tea	92	3.25%
french fries	90	3.18%
ground beef	88	3.11%
shrimp	84	2.97%
pancakes	74	2.61%
cake	63	2.22%
burgers	62	2.19%
olive oil	54	1.91%
soup	52	1.84%
cooking oil	50	1.76%
frozen smoothie	49	1.73%
turkey	49	1.73%
chicken	48	1.69%
low fat yogurt	45	1.59%
Other	1170	41.30%
Grand Total	2833	100.00%

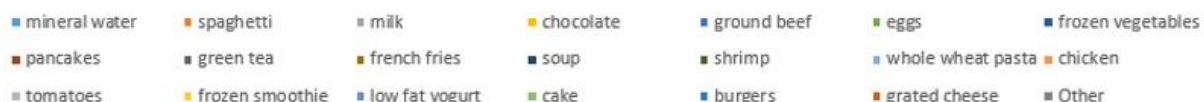
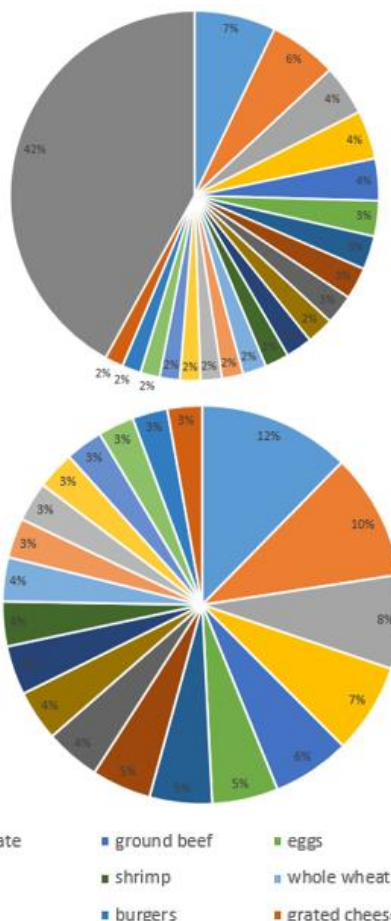


Εικόνα 26 – Tomatoes Set

Το παραπάνω σεν αφορά το προϊόν “tomatoes”. Όσον αφορά τα αποτελέσματα πάλι δεν αποκλίνουμε από τον πίνακα με τις συνολικές πωλήσεις. Ωστόσο, εδώ πρέπει να σημειωθεί ότι το συγκεκριμένο προϊόν (ντομάτες) είναι λαχανικό και συγκριτικά με τα υπόλοιπα λαχανικά που το κατάστημα εμπορεύεται έχει τις περισσότερες πωλήσεις. Άρα μπορούμε να πούμε με σιγουριά ότι στο συγκεκριμένο κατάστημα η ντομάτα είναι το πιο δημοφιλές λαχανικό για τους καταναλωτές. Έχουμε εξαίρεσει το προϊόν “frozen vegetables” διότι δεν είναι ένα λαχανικό όπως η ντομάτα αλλά πιθανός κάποιο σεν με λαχανικά και δεν γνωρίζουμε αν περιλαμβάνει και την ντομάτα μέσα ή τι είδους λαχανικά περιλαμβάνει. Αλλιώς αυτό θα ήταν το «κορυφαίο» προϊόν στον τομέα των λαχανικών, το οποίο μας δείχνει ότι οι πελάτες έχουν προτίμηση για εύκολες γρήγορες και ίσως οικονομικές λύσεις. Δηλαδή, προτιμούν μία έτοιμη σαλάτα από το να μπουν στην διαδικασία να αγοράσουν ένα-ένα τα λαχανικά που χρειάζονται, να κόψουν τα λαχανικά και στη συνέχεια να δημιουργήσουν μόνοι τους τη σαλάτα. Αυτό μπορεί να συμβαίνει για λόγους εξοικονόμησης χρόνου (η δεύτερη διαδικασία είναι χρονοβόρα), ή για λόγους εξοικονόμησης χρημάτων (το να αγοράσουν τα προϊόντα ένα-ένα ίσως κοστίζει περισσότερο από το προϊόν με τη μίξη λαχανικών).

18. OLIVE OIL

Product_with_it	Count	Count(%)
mineral water	207	7.15%
spaghetti	172	5.94%
milk	128	4.42%
chocolate	123	4.25%
ground beef	106	3.66%
eggs	90	3.11%
frozen vegetables	85	2.94%
pancakes	81	2.80%
green tea	75	2.59%
french fries	69	2.38%
soup	67	2.32%
shrimp	61	2.11%
whole wheat pasta	60	2.07%
chicken	54	1.87%
tomatoes	54	1.87%
frozen smoothie	52	1.80%
low fat yogurt	51	1.76%
cake	49	1.69%
burgers	48	1.66%
grated cheese	47	1.62%
Other	1215	41.98%
Grand Total	2894	100.00%



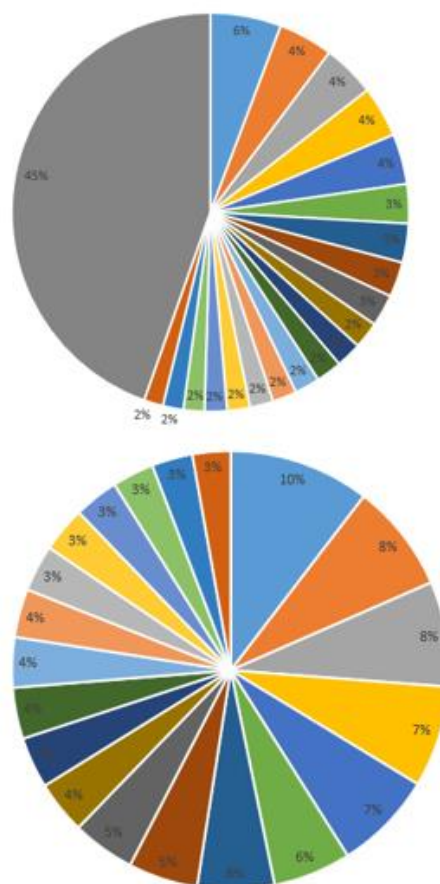
Εικόνα 27 – Olive Oil Set

Το παραπάνω σετ αφορά το προϊόν “olive oil”. Όσον αφορά τα προϊόντα που βλέπουμε στο σετ πάλι μιλάμε για μια «κανονική» εικόνα και δεν αξίζει να αναφερθεί κάτι σχετικά με τα ποσοστά τους ή τις τιμές που έχουμε εδώ.

Ωστόσο, αξίζει να αναφερθεί εδώ, το γεγονός πως στο συγκεκριμένο σετ δεν έχουμε καθόλου το προϊόν “cooking oil” πράγμα λογικό αφού τα δύο προϊόντα ανήκουν στην ίδια κατηγορία προϊόντων (λάδι). Όμως, το λάδι για μαγείρεμα είναι ένα προϊόν απαραίτητο για τα περισσότερα γεύματα και δεδομένου των αποτελεσμάτων που έχουμε στον πίνακα με τις συνολικές πωλήσεις αφού: τα προϊόντα “cooking oil” και “oil” βρίσκονται αρκετά πιο κάτω στην κατάταξη από ότι το προϊόν “olive oil” μπορούμε να πούμε με σιγουριά ότι το συγκεκριμένο προϊόν προτιμάτε στο είδος του. Αλλά, μπορούμε να υποθέσουμε ακόμη ότι οι καταναλωτές του συγκεκριμένου καταστήματος ή προτιμούν το προϊόν “olive oil” για να τηγανίσουν ή οι παρασκευές που απαιτούν τηγάνισμα δεν ταιριάζει στο lifestyle τους. Στην πρώτη περίπτωση θα πρέπει να μελετήσουμε για πιο λόγο συμβαίνει αυτό (πχ. κακή ποιότητα προϊόντος, πολύ ανταγωνιστικό το άλλο προϊόν, κτλ.). Στην δεύτερη περίπτωση θα πρέπει να προτείνουμε άλλα προϊόντα τα οποία είναι πιο συμβατά με τις ανάγκες των καταναλωτών.

19. FROZEN SMOOTHIE

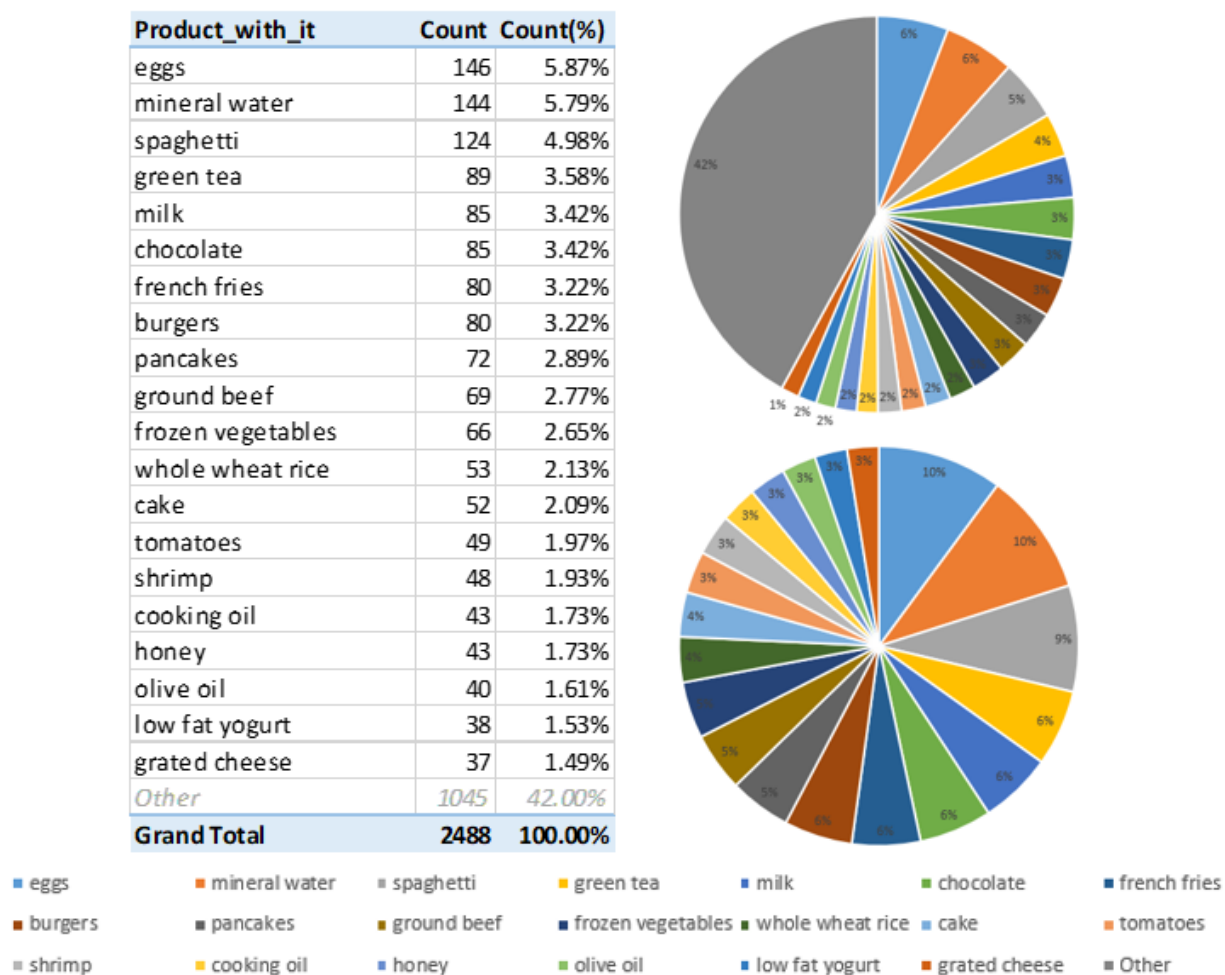
Product_with_it	Count	Count(%)
mineral water	152	5.78%
spaghetti	117	4.45%
chocolate	113	4.29%
french fries	109	4.14%
milk	107	4.07%
green tea	84	3.19%
eggs	83	3.15%
frozen vegetables	75	2.85%
pancakes	68	2.58%
ground beef	57	2.17%
shrimp	55	2.09%
low fat yogurt	55	2.09%
burgers	54	2.05%
olive oil	52	1.98%
chicken	50	1.90%
tomatoes	49	1.86%
cake	48	1.82%
whole wheat rice	45	1.71%
honey	44	1.67%
escalope	41	1.56%
Other	1173	44.58%
Grand Total	2631	100.00%



Εικόνα 28 – Frozen Smoothie Set

Το παραπάνω σεν αφορά το προϊόν “frozen smoothie”. Τα αποτελέσματα που βλέπουμε στο συγκεκριμένο σεν είναι παρόμοια με αυτά του πίνακα με τις συνολικές πωλήσεις. Το μόνο ίσως που αξίζει να σημειωθεί εδώ, είναι το γεγονός ότι έχουμε το προϊόν “honey” ανάμεσα στα κορυφαία 20 προϊόντα, πράγμα το οποίο δεν έχουμε συνηθίσει να βλέπουμε. Δεδομένου αυτού, μπορούμε να πούμε ότι το συγκεκριμένο προϊόν «ταιριάζει» με το προϊόν μας άρα μιλάμε για συσχέτιση μεταξύ των δύο προϊόντων. Οι πελάτες δηλαδή που αγοράζουν smoothies έχουν μεγάλη πιθανότητα να αγοράσουν μέλι.

20. TURKEY



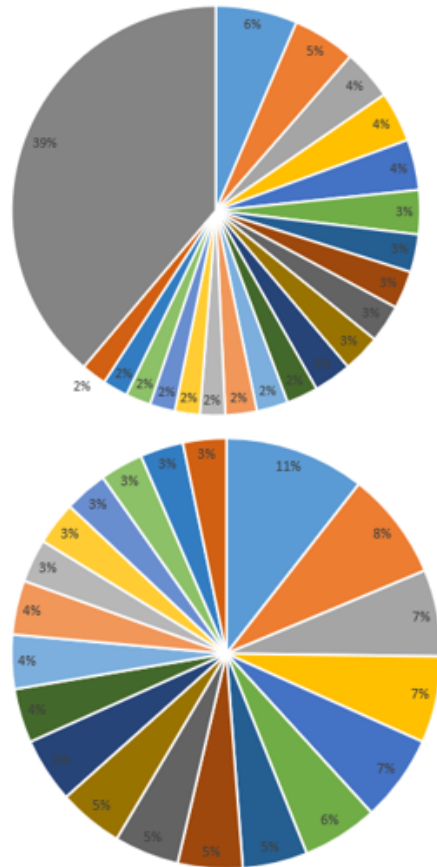
Εικόνα 29 – Turkey Set

Το παραπάνω σεν αφορά το προϊόν “turkey”. Στο συγκεκριμένο σεν μπορούμε να παρατηρήσουμε ότι πάλι έχουμε προϊόν το οποίο ξεπέρασε σε πωλήσεις το προϊόν “mineral water”. Πιο συγκεκριμένα μιλάμε για το προϊόν “eggs” το οποίο έχουμε συνηθίσει να έχει αυξημένες πωλήσεις ωστόσο το γεγονός πως ξεπέρασε το εμφιαλωμένο νερό σε πωλήσεις, παρόλο που η διαφορά των δύο προϊόντων είναι μικρή, μας κάνει να μπορούμε να μιλήσουμε για μεγάλη συσχέτιση μεταξύ των δύο προϊόντων (turkey-eggs). Ακόμη, μπορούμε να υποθέσουμε πως οι πελάτες που αγοράζουν γαλοπούλα έχουν μεγαλύτερη πιθανότητα να αγοράσουν αυγά από ότι εμφιαλωμένο νερό. Αν η διαφορά των δύο προϊόντων δεν ήταν τόσο μικρή θα μπορούσε ο ισχυρισμός αυτός να είναι συμπέρασμα, ωστόσο μπορεί το γεγονός ότι βρέθηκαν τα αυγά στην κορυφή να είναι τυχαίο.

Ανεξάρτητα από αυτό όμως, το εμφιαλωμένο νερό έχουμε συνηθίσει όχι μόνο να ξεπερνάει τα υπόλοιπα προϊόντα αλλά και με αρκετά υψηλή διαφορά, επομένως μπορούμε να ισχυριστούμε πως οι πελάτες που αγοράζουν γαλοπούλα δεν αγοράζουν τόσο πολύ εμφιαλωμένο νερό σε σχέση με τους πελάτες που δεν αγοράζουν γαλοπούλα.

111. OATMEAL

Product_with_it	Count	Count(%)
mineral water	13	6.47%
chocolate	10	4.98%
spaghetti	8	3.98%
green tea	8	3.98%
ground beef	8	3.98%
milk	7	3.48%
olive oil	6	2.99%
pancakes	6	2.99%
french fries	6	2.99%
eggs	6	2.99%
frozen vegetables	6	2.99%
shrimp	5	2.49%
soup	5	2.49%
chicken	5	2.49%
herb & pepper	4	1.99%
tomatoes	4	1.99%
whole wheat rice	4	1.99%
burgers	4	1.99%
cooking oil	4	1.99%
frozen smoothie	4	1.99%
Other	78	38.81%
Grand Total	201	100.00%



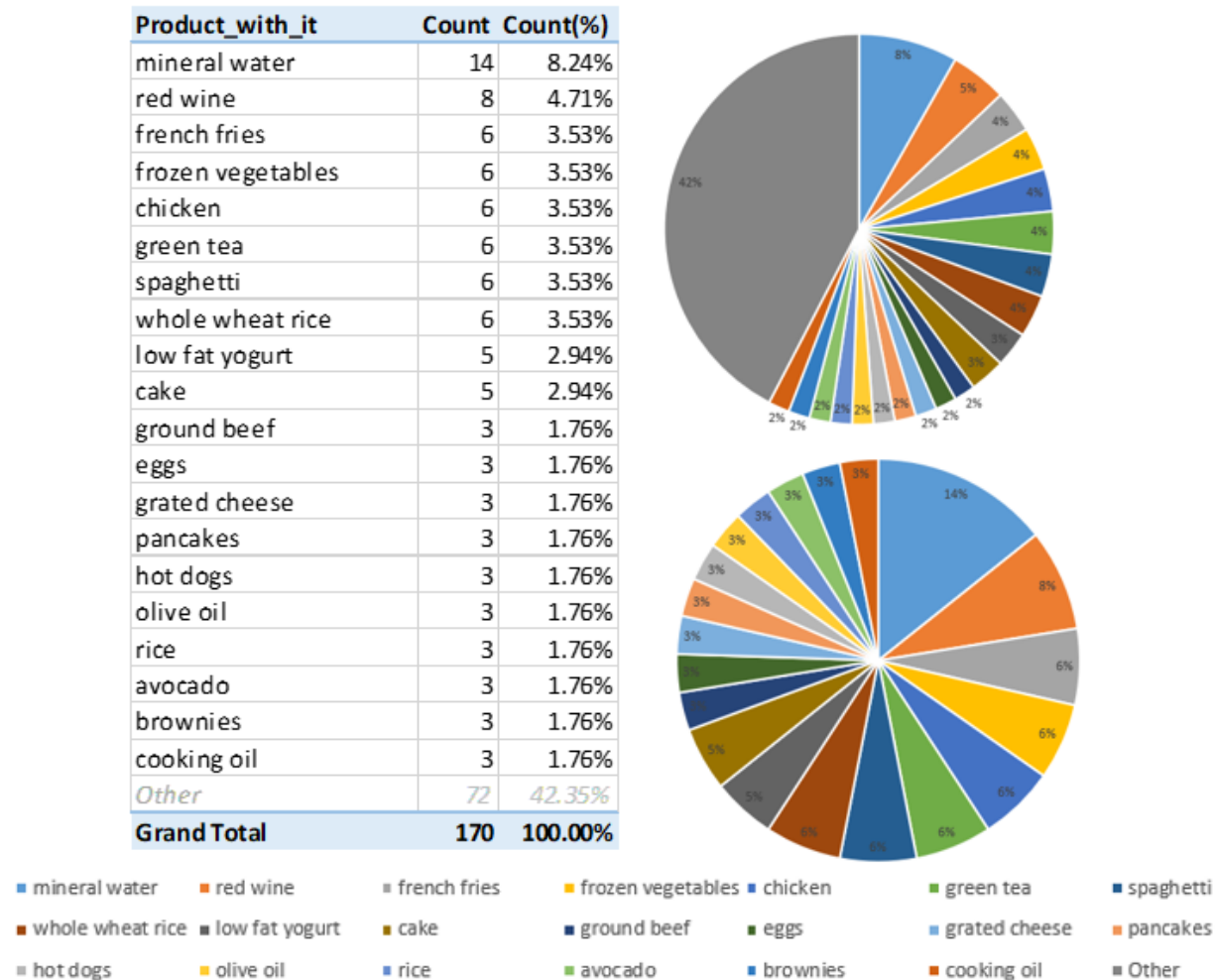
- mineral water
- chocolate
- spaghetti
- green tea
- ground beef
- milk
- olive oil
- pancakes
- french fries
- eggs
- frozen vegetables
- shrimp
- soup
- chicken
- herb & pepper
- tomatoes
- whole wheat rice
- burgers
- cooking oil
- frozen smoothie
- Other

Εικόνα 30 – Oatmeal Set

Όπως είπαμε και παραπάνω, συνεχίζουμε αναλύοντας τα τελευταία 10 προϊόντα σε πωλήσεις από τον πίνακα με τις συνολικές πωλήσεις. Η δομή θα είναι η ίδια που ακολουθήσαμε και στα παραπάνω προϊόντα. Είναι πολύ σημαντικό να εξετάσουμε και αυτά τα προϊόντα ώστε να καταλάβουμε ποιος είναι ο λόγος που τα κάνει να μην πουλιούνται εάν μπορούμε. Ακόμη θα παρατηρήσουμε εάν υπάρχει συσχέτιση με κάποιο άλλο προϊόν, ώστε να μπορέσουμε να τα προωθήσουμε μαζί για να επωφεληθεί από αυτό το «αδύναμο» προϊόν.

Ξεκινάμε με το παραπάνω σετ το οποίο αφορά το προϊόν “oatmeal” το οποίο είναι το προϊόν με τις μεγαλύτερες πωλήσεις ανάμεσα στα τελευταία δέκα προϊόντα. Τα αποτελέσματα που βλέπουμε εδώ, είναι παρόμοια με αυτά του πίνακα με τις συνολικές πωλήσεις και αντίστοιχα με τα περισσότερα σετ. Επομένως, δεν υπάρχει κάποιο προϊόν που θα μπορούσε να ταυτιστεί το προϊόν μας. Το προϊόν “oatmeal” είναι λοιπόν ένα προϊόν που δεν «πουλάει» αρκετά και το καλύτερο που θα μπορούσαμε να κάνουμε είναι να ελέγξουμε την ποιότητά του και αν αυτό ταιριάζει με τις επιλογές των καταναλωτών.

112. CHOCOLATE BREAD

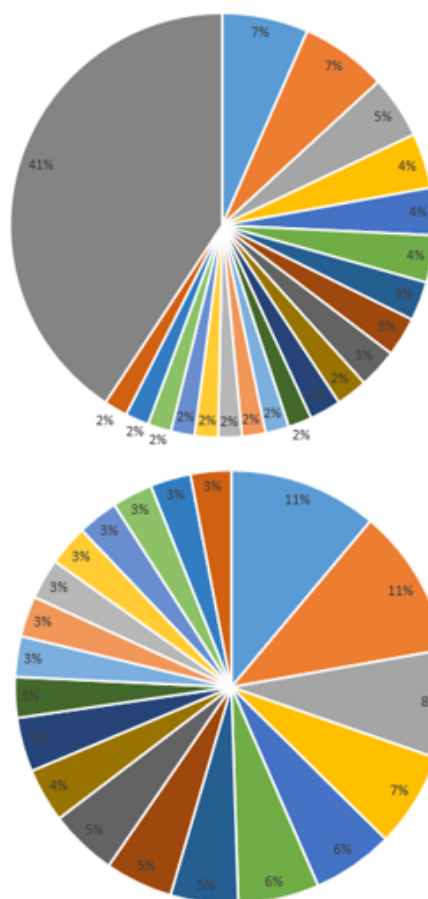


Εικόνα 31 – Chocolate Bread Set

Το παραπάνω σεν αφορά το προϊόν “chocolate bread”. Παρατηρώντας τα αποτελέσματα εδώ, βλέπουμε ότι έχουμε αρκετά προϊόντα (hot dogs, rice, brownies, κτλ.) τα οποία δεν έχουν εμφανιστεί ξανά στα πρώτα 20. Ωστόσο το σημαντικό εδώ είναι το προϊόν “red wine” το οποίο είναι σχετικά δημοφιλές προϊόν (38^η θέση), αλλά κατέχει τις μεγαλύτερες πωλήσεις στο συγκεκριμένο προϊόν μετά το “mineral water”. Πιο συγκεκριμένα το προϊόν μας έχει σύνολο 32 πωλήσεις από τις οποίες οι 8 περιέχουν και το προϊόν “red wine”. Δηλαδή ένα ποσοστό 25%, που είναι ένα καλό ποσοστό, από τις πωλήσεις του “chocolate bread” αντιστοιχούν στο “red wine”. Δεδομένου αυτού, και του γεγονότος πως το προϊόν μας είναι ένα προϊόν τύπου σνακ άρα όχι αναγκαίο για τις διατροφικές συνήθειες των νοικοκυριών, μπορούμε να κάνουμε προσφορές που αφορούν τα δύο προϊόντα μαζί (πχ. κάποια έκπτωση).

113. MASHED POTATO

Product_with_it	Count	Count(%)
spaghetti	11	6.59%
mineral water	11	6.59%
chocolate	8	4.79%
french fries	7	4.19%
eggs	6	3.59%
frozen vegetables	6	3.59%
soup	5	2.99%
escalope	5	2.99%
cake	5	2.99%
cooking oil	4	2.40%
brownies	4	2.40%
burgers	3	1.80%
green tea	3	1.80%
honey	3	1.80%
salmon	3	1.80%
shrimp	3	1.80%
tomato juice	3	1.80%
pancakes	3	1.80%
black tea	3	1.80%
cottage cheese	3	1.80%
Other	68	40.72%
Grand Total	167	100.00%



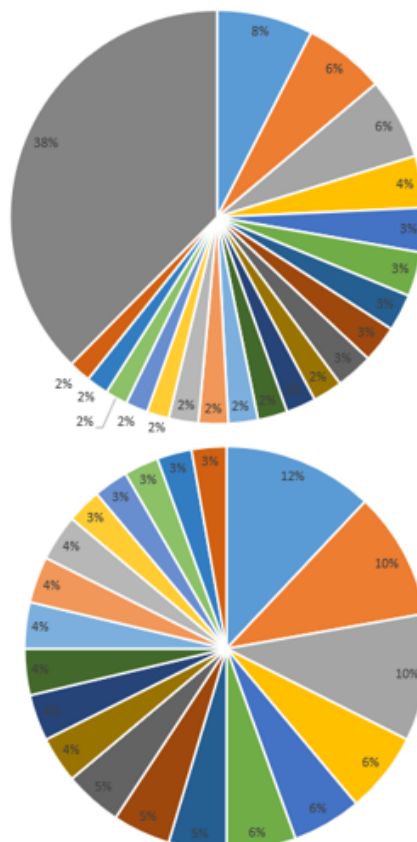
Εικόνα 32 – Mashed Potato Set

Το παραπάνω σεν αφορά το προϊόν “mashed potato”. Παρατηρώντας τα αποτελέσματα που έχουμε στην συγκεκριμένη περίπτωση βλέπουμε ότι έχουμε αρκετά μη «δημοφιλή» προϊόντα όπως “honey”, “black tea”, “tomato juice”, κτλ. Ωστόσο αυτό που είναι ίσως παράξενο εδώ, είναι το γεγονός πως έχουμε το προϊόν “black tea” το οποίο είναι αρκετά χαμηλά στην κατάταξη σε σχέση με το προϊόν “green tea” το οποίο είναι ανάμεσα στα κορυφαία προϊόντα κάθε φορά και έκτο σε πωλήσεις ανάμεσα σε όλα τα προϊόντα. Δεδομένου αυτού και του γεγονότος πως και τα δύο προϊόντα ανήκουν στην ίδια κατηγορία «τσάι», μπορούμε να πούμε πως οι πελάτες που αγοράζουν πουρέ πατάτας προτιμούν το μαύρο τσάι από το πράσινο τσάι και τα υπόλοιπα είδη τσαγιού.

Στο παραπάνω σεν μπορούμε να στοχεύσουμε και για την προώθηση του προϊόντος μας που δεν πουλάει τόσο. Αφού υπάρχει συσχέτιση αν γίνει κάποια έκπτωση στους πελάτες που αγοράζουν και τα δύο προϊόντα μαζί, αυτό θα βοηθήσει και τα δύο προϊόντα. Αφού ούτε το προϊόν “black tea” «πουλάει» τόσο.

114. CHUTNEY

Product_with_it	Count	Count(%)
mineral water	13	7.51%
spaghetti	11	6.36%
eggs	11	6.36%
frozen vegetables	7	4.05%
chocolate	6	3.47%
milk	6	3.47%
cereals	5	2.89%
ground beef	5	2.89%
grated cheese	5	2.89%
cooking oil	4	2.31%
soup	4	2.31%
salmon	4	2.31%
escalope	4	2.31%
french wine	4	2.31%
olive oil	4	2.31%
burgers	3	1.73%
shrimp	3	1.73%
pepper	3	1.73%
fresh tuna	3	1.73%
pancakes	3	1.73%
Other	65	37.57%
Grand Total	173	100.00%

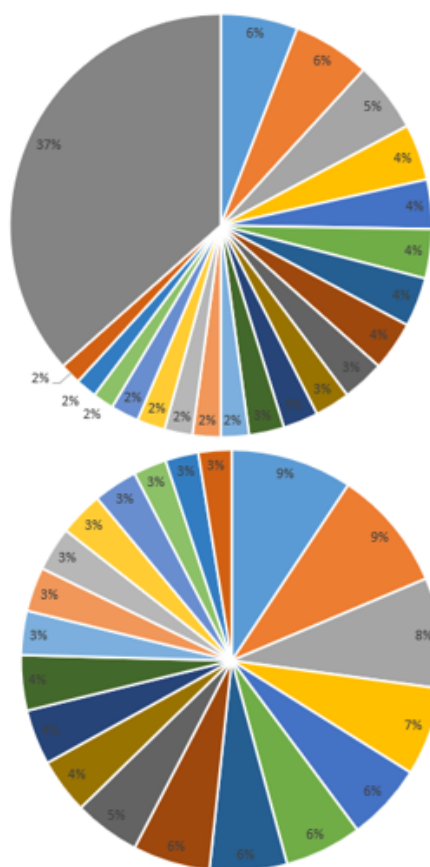


Εικόνα 33 – Chutney Set

Το παραπάνω σετ αφορά το προϊόν “chutney”. Παρατηρώντας τα αποτελέσματα στο σετ μας βλέπουμε ότι οι περισσότερες πωλήσεις του προϊόντος μας είναι με αρκετά δημοφιλή προϊόντα. Επομένως, δεν μπορούμε να μιλήσουμε για κάποια συσχέτιση εδώ. Θα πρέπει λοιπόν να προσπαθήσουμε να προωθήσουμε το προϊόν μας μόνο του.

115. TEA

Product_with_it	Count	Count(%)
mineral water	11	5.91%
spaghetti	11	5.91%
frozen vegetables	10	5.38%
eggs	8	4.30%
chocolate	7	3.76%
milk	7	3.76%
escalope	7	3.76%
pancakes	7	3.76%
soup	6	3.23%
tomatoes	5	2.69%
frozen smoothie	5	2.69%
ground beef	5	2.69%
chicken	4	2.15%
green tea	4	2.15%
magazines	4	2.15%
turkey	4	2.15%
pepper	4	2.15%
tomato juice	3	1.61%
shrimp	3	1.61%
salmon	3	1.61%
Other	68	36.56%
Grand Total	186	100.00%



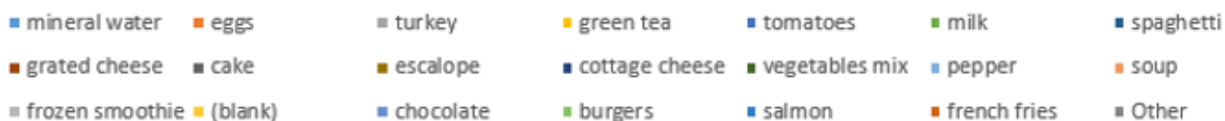
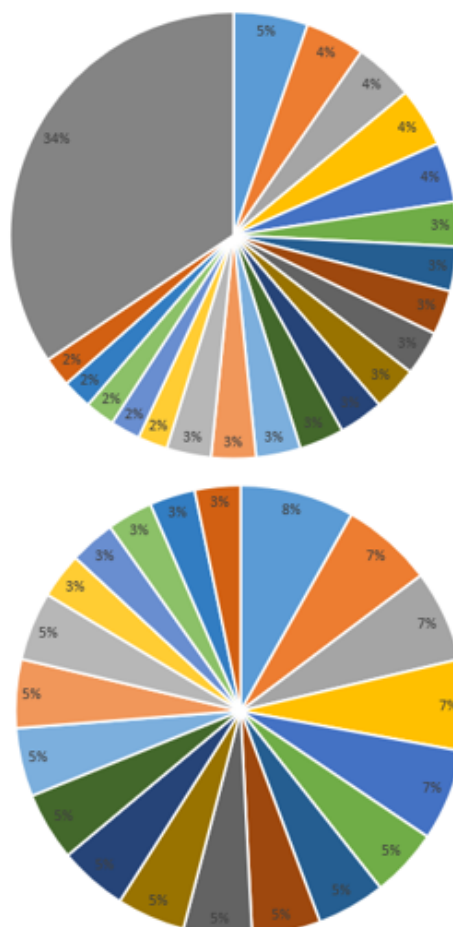
Εικόνα 34 – Tea Set

Το παραπάνω σεν αφορά το προϊόν “tea”. Παρατηρώντας τα αποτελέσματα το μόνο που αξίζει να σχολιάσουμε εδώ είναι πως έχουμε τέσσερις πωλήσεις του προϊόντος μας με το προϊόν “green tea” το οποίο ανήκει στην ίδια κατηγορία προϊόντος (τσάι) και θα μπορούσαμε να πούμε ότι είναι ανταγωνιστής του.

Πέρα από αυτό, το προϊόν μας έχει ελάχιστες πωλήσεις σε σχέση με τα υπόλοιπα προϊόντα τσαγιού (green tea, black tea, mint green tea). Επομένως μπορούμε να καταλάβουμε ότι οι καταναλωτές όσον αφορά τα συγκεκριμένα προϊόντα, προτιμούν αυτά που έχουν κάποια γεύση από ότι απλό τσάι, με ιδιαίτερη προτίμηση στο πράσινο τσάι.

116. BRAMBLE

Product_with_it	Count	Count(%)
mineral water	5	5.38%
eggs	4	4.30%
turkey	4	4.30%
green tea	4	4.30%
tomatoes	4	4.30%
milk	3	3.23%
spaghetti	3	3.23%
grated cheese	3	3.23%
cake	3	3.23%
escalope	3	3.23%
cottage cheese	3	3.23%
vegetables mix	3	3.23%
pepper	3	3.23%
soup	3	3.23%
frozen smoothie	3	3.23%
(blank)	2	2.15%
chocolate	2	2.15%
burgers	2	2.15%
salmon	2	2.15%
french fries	2	2.15%
Other	32	34.41%
Grand Total	93	100.00%

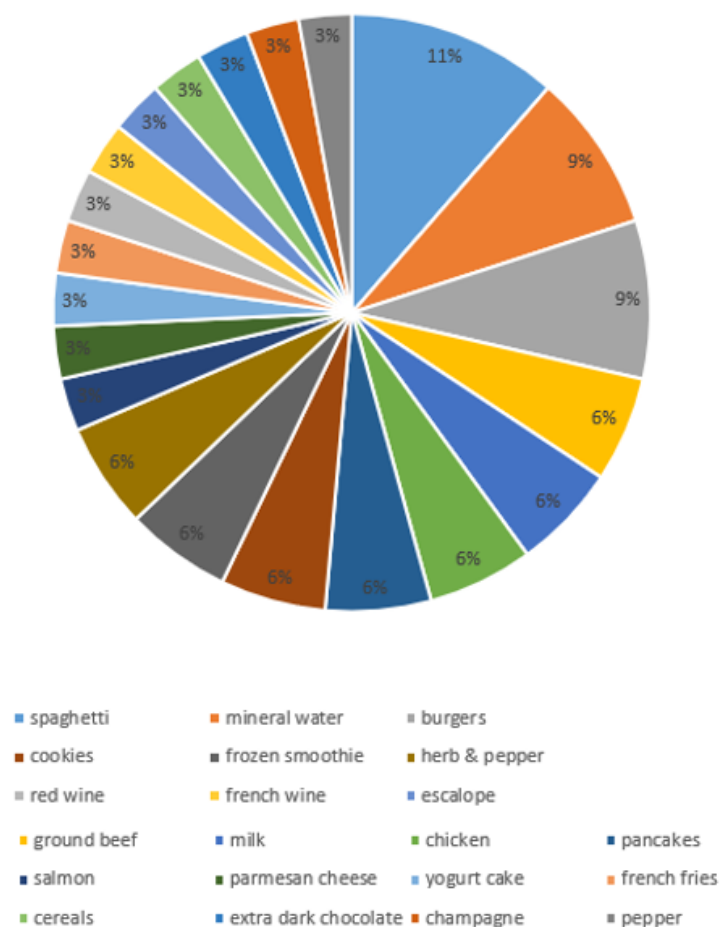


Εικόνα 35 – Bramble Set

Το παραπάνω σεν αφορά το προϊόν “bramble”. Όσον αφορά τα αποτελέσματα του πίνακα δεν παρατηρούμε κάτι περίεργο. Το μόνο που αξίζει να σημειωθεί είναι ότι έχουμε την τιμή (blank). Αυτό είναι αρκετά παράξενο για προϊόν που είναι από τα τελευταία σε πωλήσεις. Επομένως εάν σκεφτούμε και τη φύση του προϊόντος μας, ότι το bramble είναι κάποιο έτοιμο κοκτέιλ και αποτελεί είδος πολυτελείας (δηλαδή δεν είναι απαραίτητο προϊόν για τις διατροφικές συνήθειες ενός νοικοκυριού), μπορούμε να καταλάβουμε ότι οι λίγες πωλήσεις που κατέχει δεν σημαίνουν ότι το προϊόν μας δεν είναι ποιοτικό. Απλά, δεν αγοράζεται τόσο συχνά γιατί δεν απευθύνεται σε όλους τους πελάτες.

117. CREAM

Product_with_it	Count	Count(%)
spaghetti	4	8.89%
mineral water	3	6.67%
burgers	3	6.67%
ground beef	2	4.44%
milk	2	4.44%
chicken	2	4.44%
pancakes	2	4.44%
cookies	2	4.44%
frozen smoothie	2	4.44%
herb & pepper	2	4.44%
salmon	1	2.22%
parmesan cheese	1	2.22%
yogurt cake	1	2.22%
french fries	1	2.22%
red wine	1	2.22%
french wine	1	2.22%
escalope	1	2.22%
cereals	1	2.22%
extra dark chocolate	1	2.22%
champagne	1	2.22%
pepper	1	2.22%
almonds	1	2.22%
rice	1	2.22%
hot dogs	1	2.22%
soup	1	2.22%
blueberries	1	2.22%
spinach	1	2.22%
black tea	1	2.22%
(blank)	1	2.22%
oil	1	2.22%
eggs	1	2.22%
Grand Total	45	100.00%



Εικόνα 36 – Cream Set

Το παραπάνω σεν αφορά το προϊόν “cream”. Η διαφορά του με τα υπόλοιπα σεν είναι ότι έχουν μπει όλα τα αποτελέσματα που έχουμε για το συγκεκριμένο προϊόν και δεν υπάρχει η τιμή “other”. Αντίστοιχα το chart που βλέπουμε είναι αυτό με όλα τα αποτελέσματα που έχουμε για τον πίνακα μας.

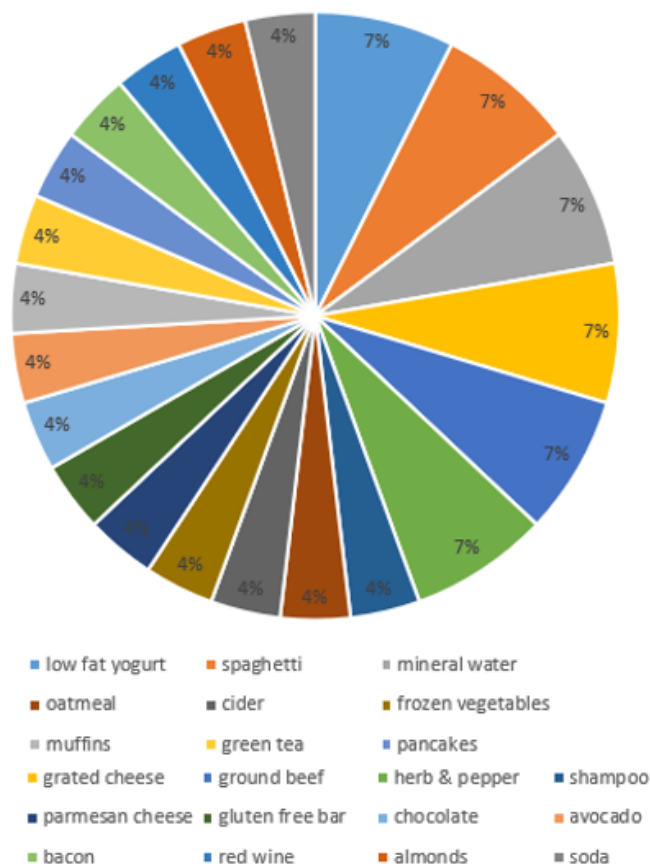
Παρατηρώντας τα αποτελέσματα, βλέπουμε ότι τις μεγαλύτερες πωλήσεις τις έχουν προϊόντα τα οποία είναι δημοφιλή, οπότε δεν μπορούμε να μιλήσουμε για κάποια συσχέτιση ούτε εδώ. Τα προϊόντα που έχουν μόνο μία πώληση δεν μας ενδιαφέρουν γιατί το προϊόν μας έχει συνολικά 7 πωλήσεις, άρα πιθανόν πολλά

από αυτά τα προϊόντα να βρίσκονται στο ίδιο καλάθι αγοράς (4 καλάθια με εμφιαλωμένο νερό, 1 (blank), άρα μένουν 2 καλάθια).

Η τιμή (blank) σημαίνει ότι το προϊόν αγοράστηκε μία φορά μόνο του. Αυτό θα μπορούσε να σημαίνει ότι υπάρχει κάποια προτίμηση στο προϊόν μας, αφού κάποιος πελάτης μπήκε στο κατάστημα μόνο γι' αυτό. Όμως, επειδή έχουμε μόνο ένα τέτοιο παράδειγμα, δεν μπορούμε να είμαστε σίγουροι.

118. NAPKINS

Product_with_it	Count	Count(%)
low fat yogurt	2	6.45%
spaghetti	2	6.45%
mineral water	2	6.45%
grated cheese	2	6.45%
ground beef	2	6.45%
herb & pepper	2	6.45%
shampoo	1	3.23%
oatmeal	1	3.23%
cider	1	3.23%
frozen vegetables	1	3.23%
parmesan cheese	1	3.23%
gluten free bar	1	3.23%
chocolate	1	3.23%
avocado	1	3.23%
muffins	1	3.23%
green tea	1	3.23%
pancakes	1	3.23%
bacon	1	3.23%
red wine	1	3.23%
almonds	1	3.23%
soda	1	3.23%
light mayo	1	3.23%
tomato sauce	1	3.23%
butter	1	3.23%
carrots	1	3.23%
Grand Total	31	100.00%

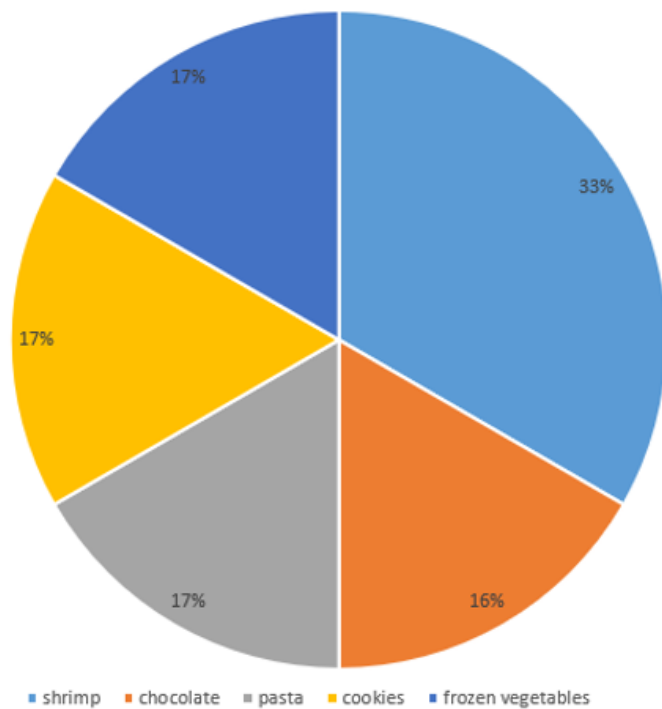


Εικόνα 37 – Napkins Set

Το παραπάνω σεν αφορά το προϊόν “napkins”. Ακόμη μια φορά δεν μπορούμε να σχολιάσουμε κάτι αφού τα προϊόντα που έχουν πάνω από μια πώληση, είναι αρκετά δημοφιλή προϊόντα. Αυτό που μπορούμε να παρατηρήσουμε σίγουρα είναι πως, όσο πιο λίγες πωλήσεις έχει ένα προϊόν, τόσο πιο πολλά μη δημοφιλή προϊόντα βλέπουμε. Αν προσέξουμε εδώ, όπως και στα παραπάνω σεν με τα λιγότερο δημοφιλή προϊόντα παρατηρούμε ότι συνέχεια βρίσκουμε καινούρια προϊόντα. Ενώ θα περίμενε κανείς τα καλάθια των πελατών να αποτελούνται από προϊόντα δημοφιλή και ίσως ένα ή δύο προϊόντα μη δημοφιλή. Αυτό ίσως σημαίνει ότι οι πελάτες που αγοράζουν ένα προϊόν που δεν πουλάει, τότε θα αγοράσουν και άλλα προϊόντα που δεν πουλάνε.

119. WATER SPRAY

Product_with_it	Count	Count(%)
shrimp	2	33.33%
chocolate	1	16.67%
pasta	1	16.67%
cookies	1	16.67%
frozen vegetables	1	16.67%
Grand Total	6	100.00%

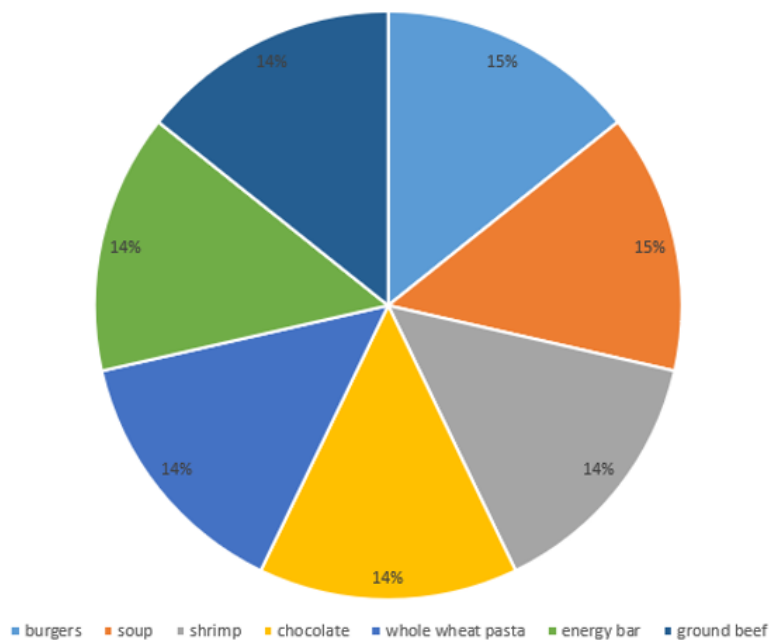


Εικόνα 38 – Water Spray Set

Το παραπάνω σετ αφορά το προϊόν “water spray”. Οι συνολικές του πωλήσεις είναι 3 γι’ αυτό και ο πίνακας του είναι τόσο μικρός. Όσον αφορά τα αποτελέσματα, μπορούμε να πούμε ότι είναι αναμενόμενα. Δηλαδή ότι οι αγορές του προϊόντος μας έχουν μοιραστεί σε προϊόντα αρκετά δημοφιλή με πρώτο το προϊόν “shrimp”.

120. ASPARAGUS

Product_with_it	Count	Count(%)
burgers	1	14.29%
soup	1	14.29%
shrimp	1	14.29%
chocolate	1	14.29%
whole wheat pasta	1	14.29%
energy bar	1	14.29%
ground beef	1	14.29%
Grand Total	7	100.00%



Εικόνα 39 – Asparagus Set

Το παραπάνω σετ αφορά το προϊόν “asparagus”. Επειδή το συγκεκριμένο προϊόν έχει μόνο μία πώληση στον πίνακα και στο γράφημα βλέπουμε ολόκληρο το καλάθι της αγοράς αυτής. Δεν μπορούμε να σχολιάσουμε κάτι για τα προϊόντα διότι πρόκειται για μία πώληση.

ΓΕΝΙΚΕΣ ΠΑΡΑΤΗΡΗΣΕΙΣ

Έχοντας υπόψιν όλους τους πίνακες που μελετήσαμε μπορούμε να βγάλουμε τα εξής συμπεράσματα για το κατάστημα από το οποίο προέρχεται το συγκεκριμένο dataset και για τους πελάτες του:

1. Στην περιοχή που βρίσκεται το κατάστημα το νερό δεν είναι πόσιμο ή δεν θεωρείται από τους τοπικούς κατοίκους «καλής ποιότητας». Αυτό συνεπάγεται από το γεγονός ότι έχουμε αυξημένες πωλήσεις στο εμφιαλωμένο νερό.
2. Οι πελάτες του καταστήματος ανήκουν στην μεσαία ή ανώτερη τάξη. Αυτό εξηγείται από το γεγονός ότι υπάρχει μεγάλος αριθμός αγορών σε γλυκά και σνακ, τα οποία είναι προϊόντα που δεν αποτελούν ανάγκη για ένα νοικοκυριό.
3. Υπάρχει μεγάλο πλήθος νεαρών ηλικιακά πελατών. Αυτό μπορούμε να το συμπεράνουμε από τον υπερβολικό αριθμό πωλήσεων σε μπισκότα χωρίς να υπάρχει συνοδεία άλλων προϊόντων στα καλάθια. Τα περισσότερα νοικοκυριά όταν κάνουν τα ψώνια τους γεμίζουν τα καλάθια τους με όλα τα αναγκαία υλικά που θα χρειαστούν για ολόκληρη την εβδομάδα ή και περισσότερο. Εάν έχουν παραλείψει κάποιο προϊόν, ναι θα το αγοράσουν μόνο του (πάλι όμως το καλάθι συνοδεύεται και από κάτι ακόμη που έχουν παραλείψει). Τα μπισκότα δεν είναι ένα προϊόν που θα τους κάνει όμως να βγουν από το σπίτι για να το προμηθευτούν. Επομένως εδώ πιθανώς μιλάμε για άτομα νεαρής ηλικίας τα οποία συχνάζουν κοντά στο κατάστημα ή το σπίτι τους βρίσκεται κοντά στο κατάστημα και «λαχταρούν» ένα σνακ.
4. Περίπου το 5% των πελατών φεύγει με καλάθι που περιέχει μόνο ένα προϊόν. (1578 -μοναδικές πωλήσεις/29358 – συνολικές πωλήσεις)

ΠΗΓΕΣ

1. <https://www.kaggle.com/dragonheir/basket-optimisation>
2. <https://docs.microsoft.com/en-us/sql/sql-server/sql-server-technical-documentation?view=sql-server-2017>
3. <https://wordart.com/create>
4. https://el.wikipedia.org/wiki/%CE%92%CE%AC%CF%83%CE%B7_%CE%B4%CE%B5%CE%B4%CE%BF%CE%BC%CE%AD%CE%BD%CF%89%CE%BD
5. <https://libguides.cut.ac.cy/etds/structure>
6. https://www.dept.aueb.gr/sites/default/files/mbc/useful/Marketing_defined.pdf
7. https://en.wikipedia.org/wiki/Data_analysis
8. https://www.albionresearch.com/data_mining/market_basket.php
9. <https://www.w3schools.com/sql/default.asp>