

Συστήματα Επιχειρηματικής Ευφυΐας

Η αξία τους στις σύγχρονες
επιχειρήσεις

&

εφαρμογή σε περιβάλλον της
Microsoft Power BI

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ
ΤΟΥ
ΔΗΜΗΤΡΙΟΥ ΚΟΜΝΗΝΟΥ

Επιβλέπων: Νικόλαος Παναγιώτου
Καθηγητής Ε.Μ.Π.

Σχολή Μηχανολόγων Μηχανικών
Ε.Μ.Π.

Τομέας Βιομηχανικής Διοίκησης και
Επιχειρησιακής Έρευνας

Αθήνα , Οκτώβριος 2023



Πίνακας Περιεχομένων

Πίνακας Περιεχομένων	3
Πίνακες	5
Διαγράμματα	5
Σχήματα	5
Εικόνες	6
Εξισώσεις	8
Διάρθρωση διπλωματικής	9
Περίληψη	12
Abstract	13
Εισαγωγή	15
1. Η επιλογή του θέματος της Διπλωματικής	17
2.Συστήματα Επιχειρηματικής Ευφυΐας	19
2.1 Ορισμός Επιχειρηματικής Ευφυΐας	19
2.1.1 Ιστορικά στοιχεία για τα συστήματα Επιχειρηματικής Ευφυΐας.....	20
2.2 Πληροφοριακά Συστήματα	22
2.2.1 Λειτουργικά Συστήματα	23
2.2.2 Αναλυτικά Συστήματα	24
2.3 Αρχιτεκτονική αναλυτικών συστημάτων DW/BI	26
2.3.1 Οι αρχιτεκτονικές συστημάτων DW/BI των R.Kimball και W.H.Inmon	27
2.3.2 Stand-alone Marts	29
2.3.3 Σύγκριση των Αρχιτεκτονικών του DW/BI	30
2.4 Η κουλτούρα των εταιριών γύρω από το BI & τα Analytics	31
2.5 Παραδοσιακά και self-service BI συστήματα	34
2.6 Self-service BI λογισμικά της αγοράς	37
3. Η πλατφόρμα Επιχειρηματικής Ευφυΐας της Microsoft Power BI	41
3.1 Το οικοσύστημα του Power BI	41
3.1.1 Συνιστώσες του Power BI	43
3.2 Power BI Desktop	44
3.3 Power BI Desktop - Ανακάλυψη Δεδομένων	48
3.3.1 Tabular processing [10]	48
3.3.2 Σύνδεση Δεδομένων.....	53

3.3.3 Διαμόρφωση Δεδομένων	56
3.4 Power BI Desktop - Μοντελοποίηση Δεδομένων	58
3.4.1 Μοντέλα Δεδομένων	58
3.4.2 Σχήμα Αστέρι – Star Schema	62
3.4.3 Σχέσεις του Μοντέλου Δεδομένων - Model Relationships	68
3.4.4 Μετρήσεις - Measures.....	77
3.5 Power BI Desktop - Οπτικοποίηση Δεδομένων	78
3.6 Power BI Service - Δημοσίευση και Συνεργασία.....	80
3.6.1 Άδειες Χρήσης και δυνατότητες	81
3.6.2 Δημοσίευση στο Power BI Service	83
3.6.3 Ολοκληρωμένο Self-service σύστημα στο Power BI	91
4. Αξιολόγηση Εταιρείας Ηλεκτρονικού Εμπορίου	97
4.1 Συλλογή επιχειρηματικών απαιτήσεων – Καταγραφή των Αναγκών του οργανισμού... 100	
4.2 Εισαγωγή Δεδομένων και κατασκευή Μοντέλου Δεδομένων	102
4.3 Πρόταση Λύσης στην Υπόθεση Εργασίας.....	120
4.3.1 Παρουσίαση Καρτελών Αξιολόγησης	121
4.3.2 Παραδείγματα επίδειξης των δυνατοτήτων του εργαλείου Power BI.....	132
4.3.3 Ανάλυση Προφίλ Πελατών – Customers Analysis	141
4.3.4 Ανάλυση Πωλήσεων και Αξιολόγηση των Πωλητών-Συνεργατών	152
4.3.5 Ανάλυση Προσέλκυσης Νέων Συνεργατών στο δίκτυο της εταιρείας – Τμήμα Μάρκετινγκ.....	159
4.3.6 Ολοκλήρωση του self-service συστήματος στο Power BI Service	167
5. Συμπεράσματα.....	174
5.1 Αποτίμηση εταιρείας	174
5.2 Αποτίμηση self-service συστημάτων επιχειρηματικής ευφυΐας και το μέλλον τους	175
6. Ευχαριστίες.....	178
Βιβλιογραφία	179
Παράρτημα.....	181
Συνομογραφίες	182

Πίνακες

Πίνακας 1 : Λειτουργικά vs Αναλυτικά Συστήματα	26
Πίνακας 2 : Σύγκριση των Αρχιτεκτονικών Σχεδίασης των Αποθηκών Δεδομένων	31
Πίνακας 3 : Σύγκριση παραδοσιακού και self-service BI [6]	37
Πίνακας 4 : Οι Leaders στον χώρο των self-service BI εργαλείων [8].....	39
Πίνακας 5: Σύγκριση των μεθόδων σύνδεσης δεδομένων στο Power BI	51
Πίνακας 6 : Σύγκριση Μοντέλων Δεδομένων	68
Πίνακας 7 :Σχέσεις μοντέλου δεδομένων.....	77
Πίνακας 8 : Περιγραφή διαδικασιών λειτουργίας ενός συστήματος με κεντρικοποιημένο μοντέλο – Managed self-service BI.....	93
Πίνακας 9 : Περιγραφή διαδικασιών λειτουργίας ενός self-service συστήματος σε εταιρικό επίπεδο – Enterprise BI.....	96
Πίνακας 10 : Πίνακας επιχειρηματικών δραστηριοτήτων οργανισμού – Olist.....	101
Πίνακας 11 : Τα δεδομένα που εισάγονται στο Power BI Desktop	103

Διαγράμματα

Διάγραμμα 1 : Πληροφοριακά Συστήματα	22
Διάγραμμα 2 : Παγκόσμιος όγκος αναζητήσεων των self-service εργαλείων σύμφωνα με το Google Trends	40
Διάγραμμα 3 : Storage mode.....	53
Διάγραμμα 4: Tabular Μοντέλο Δεδομένων στον γραφικό περιβάλλον του Power BI Desktop.....	59
Διάγραμμα 5 : Κανονικοποιημένος πίνακας Sales.....	61
Διάγραμμα 6 : Μοντέλο Δεδομένων στο Power BI – Model View.....	65
Διάγραμμα 7 : Διαστάσεις Χιονονιφάδας.....	66
Διάγραμμα 8 : Many-to-one σχέση	71
Διάγραμμα 9 : One-to-one σχέση σε εκφυλισμένη διάσταση.....	72
Διάγραμμα 10 : One-to-one σχέση όπου τα δεδομένα έρχονται από διαφορετικά πηγαία συστήματα. Πηγή Εικόνας: Πηγή Εικόνας	73
Διάγραμμα 11 : Τα ετήσια έσοδα της Olist στην πορεία των χρόνων	97

Σχήματα

Σχήμα 1 : Κύκλος ζωής ενός συστήματος BI	20
Σχήμα 2 : Η αρχιτεκτονική της αποθήκης δεδομένων του W.H.Inmon : the Corporate Information Factory	27
Σχήμα 3 : Η αρχιτεκτονική της αποθήκης δεδομένων του Ralph Kimball : Dimensional Data Warehouse	28
Σχήμα 4 : Πολλαπλά Stand-alone Data Marts	30
Σχήμα 5 : BI και το μοντέλο ωρίμανσης των Analytics στην εταιρεία	32
Σχήμα 6 : Συνοπτική πορεία εργασιών σε σύστημα BI	36
Σχήμα 7 : Οι φάσεις ενός BI project στο Power BI Desktop	45
Σχήμα 8 : Η αρχιτεκτονική στο SSAS για Tabular models [11].....	50
Σχήμα 9 : Σύνθετο μοντέλο δεδομένων (composite model) [13].....	52
Σχήμα 10 : Μοντέλα Δεδομένων	60
Σχήμα 11: Σχηματική αναπαράσταση του μοντέλου δεδομένων	64
Σχήμα 12 : Παράδειγμα επίλυσης προβλήματος σύνδεσης πινάκων Διαστάσεων.....	74

Σχήμα 13 : Σύνδεση πινάκων με διαφορετικό επίπεδο αποθήκευσης πληροφορίας	75
Σχήμα 14 : Δεύτερο σετ δεδομένων Μάρκετινγκ – Marketing Dataset	99
Σχήμα 15 : Πρώτο σετ δεδομένων με τις παραγγελίες της εταιρείας – Brazilian E-Commerce Dataset	99
Σχήμα 16 : Στοιχεία ανάλυσης προφίλ πελατών	141

Εικόνες

Εικόνα 1 : Πίνακας ελέγχου (Dashboard) ιστορικών συστημάτων BI – IBM – Cognos Powerplay (1997)	21
Εικόνα 2 : Το μαγικό τεταρτοκύκλιο του Gartner	38
Εικόνα 3 : Τα μέρη της Power Platform	41
Εικόνα 4 : Οι συνιστώσες του Power BI	43
Εικόνα 5 : Το γραφικό περιβάλλον του Power BI Desktop – Report View	46
Εικόνα 6 : Το γραφικό περιβάλλον του Power BI Desktop – Data View	46
Εικόνα 7 : Το γραφικό περιβάλλον του Power BI Desktop – Model View	47
Εικόνα 8 : Παρασκηνιακή λειτουργία του Microsoft SSAS στο Task Manager	48
Εικόνα 9: Το Get Data κουμπί.....	55
Εικόνα 10: Σύνδεση κοινόχρηστου μοντέλου δεδομένων	55
Εικόνα 11 Κορδέλα επεξεργασίας Transform – Power Query Editor	57
Εικόνα 12 : Κορδέλα επεξεργασίας Home – Power Query Editor	57
Εικόνα 13 : Στιγμιότυπο από τον Επεξεργαστή – Power Query Editor	57
Εικόνα 14 : Ιεραρχίες επιχειρηματικών οντοτήτων στο Power BI – Model View	67
Εικόνα 15 : Εκφυλισμένη Διάσταση.....	72
Εικόνα 16 : Inner Join – Διάγραμμα Venn Πηγή Εικόνας	73
Εικόνα 17 : Παράδειγμα σύνδεσης πινάκων με σχέση τύπου Many-to-many – Model View Πηγή Εικόνας	75
Εικόνα 18 : Παράδειγμα εφαρμογής Μέτρησης – Report View	78
Εικόνα 19 : Ενδεικτικές οπτικοποιήσεις του Power BI	79
Εικόνα 20 : Το πάνελ οπτικοποίησης - Report View	80
Εικόνα 21 : Παράδειγμα δημιουργίας αναφοράς - Report View	80
Εικόνα 22 : Αρχική σελίδα Power BI Service στον φυλλομετρητή – Power BI Service	81
Εικόνα 23 : Δημοσίευση στο Power BI Service – Power BI Desktop	83
Εικόνα 24 : Παράδειγμα Χώρου Εργασίας – Power BI Service	84
Εικόνα 25 : Ρύθμιση Προσβασιμότητας στον χώρο εργασίας – Power BI Service	84
Εικόνα 26 : Ανάθεση Ρόλων στους χώρους εργασίας – Power BI Service	84
Εικόνα 27 : Παράδειγμα Πίνακα ελέγχου (Dashboard) – Power BI Service	86
Εικόνα 28 : Εισαγωγή ροής ζωντανών δεδομένων – Power BI Service	87
Εικόνα 29 : Καρτέλα κατεβασμένων εφαρμογών – Power BI Service	88
Εικόνα 30 : Δημιουργία πολλαπλών αναφορών από κοινό μοντέλο δεδομένων ,Data Linage – Power BI Service	89
Εικόνα 31 : Diagram View, Dataflow - Power BI Service Πηγή Εικόνας	90
Εικόνα 32 : Ρύθμιση Λειτουργίας Row Level Security – Power BI Service	91
Εικόνα 33 : Κεντρικοποιημένο μοντέλο χρήσης – Managed self-service BI Πηγή εικόνας	92
Εικόνα 34 : Ολοκληρωμένο Σύστημα Επιχειρηματικής Ευφυΐας σε επίπεδο εταιρείας Πηγή Εικόνας	94
Εικόνα 35: Εισαγωγή .csv αρχείου	104
Εικόνα 36 : Μετάβαση στο Power Query	104
Εικόνα 37 : Επιλογή λειτουργίας Column Quality – Power Query Editor	104
Εικόνα 38 : Ποιότητα δεδομένων στήλης – Power Query Editor	104
Εικόνα 39 : Επιλογή Custom Column – Power Query Editor	105
Εικόνα 40 : Μετατροπή από Date/Time σε τύπο Text	106
Εικόνα 41 : Δημιουργία στήλης Date_Key – Power Query Editor	106
Εικόνα 42 : Αλγοριθμική στήλη – Conditional Formatting – Power Query Editor	106
Εικόνα 43 : Διαχωρισμός στήλης – Power Query Editor	106

Εικόνα 44 : Ο αλγόριθμος δημιουργίας της στήλης [Time of the Day].....	107
Εικόνα 45 : Ο αλγόριθμος δημιουργίας της στήλης [Time of the day Sorting] – Power Query Editor.....	108
Εικόνα 46 : Δημιουργία στήλης τύπου Index - Power Query Editor	108
Εικόνα 47 : Αφαίρεση διπλότυπων σειρών πίνακα Reviews – Power Query Editor	110
Εικόνα 48 : Σύνδεση πίνακα Reviews με την τεχνική του πίνακα «Γέφυρα» - Model View	111
Εικόνα 49 : Σύνδεση πίνακα Order_Payments στο μοντέλο - Model View.....	112
Εικόνα 50 : Ομαδοποίηση των παραγγελιών για την δημιουργία του πίνακα Order_Items_Aggregations – Power Query Editor	112
Εικόνα 51 : Ομαδοποίηση των παραγγελιών για την δημιουργία του πίνακα Order_Payments_Agg – Power Query Editor	113
Εικόνα 52 : Προσθήκη πινάκων Products, Sellers, Category Name – Model View	113
Εικόνα 53 : Δημιουργία πίνακα Date – Data View	114
Εικόνα 54 : Ορίζουμε τον πίνακα Date του συστήματος – Model View	115
Εικόνα 55 : Σύνδεση πινάκων Geolocation, Customers, Customers Aggregation, Customer Data - Model View	116
Εικόνα 56 : Συγχώνευση Πινάκων – Power Query Editor	118
Εικόνα 57 : Σύνδεση του σετ δεδομένων Μάρκετινγκ - Model View	119
Εικόνα 58 : Το τελικό μοντέλο δεδομένων – Model View	120
Εικόνα 59 : Περιεχόμενα Αναφοράς – Report View	121
Εικόνα 60 : Καρτέλα Αξιολόγησης των Κριτικών	122
Εικόνα 61 : Πίνακας “Average Review by Geolocation by Order Status “– Reviews Καρτέλα	123
Εικόνα 62 : Πίνακας “Average Review by Geolocation by Order Status “ - Reviews Καρτέλα	123
Εικόνα 63 : Κάρτες αξιολογούμενων δεδομένων – Reviews Καρτέλα.....	124
Εικόνα 64 : Οπτικοποίηση Γεωγραφικών Δεδομένων – Reviews Καρτέλα.....	124
Εικόνα 65 : Cluster Columns – Reviews Καρτέλα	124
Εικόνα 66 : Οπτικοποίηση τύπου «σμήνους λέξεων» - Reviews Καρτέλα 2.....	125
Εικόνα 67 : Καρτέλα Αξιολόγησης Εφοδιαστικής Αλυσίδας - Logistics Evaluation	126
Εικόνα 68 : Καρτέλα 2 Αξιολόγησης Εφοδιαστικής Αλυσίδας- Logistics Evaluation 2	127
Εικόνα 69 : Καρτέλα Αξιολόγησης Πελατών – Customers Analysis	128
Εικόνα 70 : Δημογραφικά Στοιχεία Πελατών.....	129
Εικόνα 71 : Καρτέλα Αξιολόγησης Παραγγελιών – Orders Evaluation	130
Εικόνα 72 : Ανάλυση Πωλητών στην πλατφόρμα της εταιρείας	130
Εικόνα 73 : Αξιολόγηση Πληρωμών – Payments Analysis	131
Εικόνα 74 : Ανάλυση Ευκαιριών Ανάπτυξης – Payments Analysis	132
Εικόνα 75 : Μέση βαθμολογία κριτικής ανά κατηγορία προϊόντων – Reviews Καρτέλα	133
Εικόνα 76 : Πολιτείες με περισσότερες παραγγελίες στην κατηγορία «Home & Office Furniture» - Reviews Καρτέλα	134
Εικόνα 77 : Η πολιτεία RJ αναδεικνύεται ως κύριος παράγοντας στην μειωμένη κριτική – Reviews Καρτέλα 2	134
Εικόνα 78 : Φιλτράρισμα καρτέλας Αξιολογήσεων για TK 20271-25041 – Reviews Καρτέλα	135
Εικόνα 79 : Λειτουργία Drill-down σε επίπεδο πόλης – Reviews Καρτέλα.....	135
Εικόνα 80 : Η πολιτεία RJ αναδεικνύεται ως κύριος παράγοντας στο μειωμένο ποσοστό έγκαιρης παράδοσης – Reviews Καρτέλα 2	136
Εικόνα 81 : Φιλτράρισμα Πωλητών που προμηθεύουν την περιοχή RJ – Logistics Καρτέλα 2.....	137
Εικόνα 82 : Τα ποσοστά έγκαιρης παράδοσης στο RJ ανά πολιτεία Πωλητή.....	137
Εικόνα 83 :Χαμηλότερες Κριτικές στην Κατηγορία «Home & Office” - Reviews Καρτέλα	138
Εικόνα 84: Σύγκριση χρόνων παραγγελιών στις περιοχές AL,PA,PI,MA,AC – Logistics Καρτέλα	138
Εικόνα 85 :Φιλτράρισμα Πωλητών που προμηθεύουν τις περιοχές AL,PA,PI,MA,AC – Logistics Καρτέλα 2.....	139
Εικόνα 86 : Λειτουργία Drill Down σε επίπεδο TK – Logistics Καρτέλα 2	139
Εικόνα 87 : Καρτέλα Αξιολόγησης των Πελατών	142
Εικόνα 88 : Τα ηλικιακά γκρουπ των πελατών της εταιρείας.....	142
Εικόνα 89 : Δείκτης HDI και ΑΕΠ στις πολιτείες της Βραζιλίας.....	143
Εικόνα 90 : Πολυπαραμετρική ανάλυση πελατών	144
Εικόνα 91 : Το διάστημα της μέρας που προτιμούν οι πελάτες για τις αγορές τους	145
Εικόνα 92 : Οι μέρες της εβδομάδας που κάνουν παραγγελίες οι πελάτες.....	146
Εικόνα 93 : Πίνακας με τις κατηγορίες προϊόντων που προτιμούν οι πελάτες	146
Εικόνα 94: Πολυπαραμετρική ανάλυση Πελατών – Διάγραμμα τύπου Radar.....	147

Εικόνα 95 : Αξιολογήσεις για το 80% των παραγγελιών	148
Εικόνα 96 : Οι χειρότερες αξιολογήσεις των παραγγελιών	148
Εικόνα 97 : Η πορεία του αριθμού των παραγγελιών ανά μήνα.....	149
Εικόνα 98 : Οι παραγγελίες της εταιρείας με ανά μήνα και σύγκριση ετών	150
Εικόνα 99 : Τα έσοδα της εταιρείας στην πορεία του χρόνου.....	150
Εικόνα 100 : Ο τρόπος πληρωμής που επιλέγουν οι πελάτες	151
Εικόνα 101 : Στοιχεία Εσόδων για τις εγχώριες πωλήσεις	153
Εικόνα 102 : Ρυθμός Ανάπτυξης Εσόδων ανά έτος 2017-18.....	155
Εικόνα 103 : Πίνακας Περιθωρίου Κερδών Εγχώριων πωλήσεων	155
Εικόνα 104 : Καρτέλα Αξιολόγησης Πωλητών – Power BI Service	157
Εικόνα 105 : Ποσοστό πρόσφατης δραστηριότητας ανά πολιτεία Πωλητή	158
Εικόνα 106 : Οι διαπιστευμένες καινούργιες επαφές του μάρκετινγκ μέσα στο χρόνο	159
Εικόνα 107 : Οι πηγές των διαπιστευμένων νέων επαφών του Μάρκετινγκ.....	160
Εικόνα 108 Το ποσοστό των συμφωνιών ανά πηγή	161
Εικόνα 109 : Τα έσοδα από τις ενσωματωμένες νέες συμφωνίες – Power BI Service	161
Εικόνα 110: Σημαντικά στοιχεία για το κλείσιμο νέων συμφωνιών.....	162
Εικόνα 111 : Η μετατροπή των διαπιστευμένων επαφών σε συμφωνίες με πωλητές – Power BI Service	162
Εικόνα 112 : Πίνακας νέων συμφωνιών ανά επιχειρηματικό κλάδο	163
Εικόνα 113 : Καρτέλα Αξιολόγησης των διαπιστευμένων επαφών - Power BI Service	165
Εικόνα 114 : Τα έσοδα από τις κατηγορίες Online Big και Online Medium – Power BI Service.....	165
Εικόνα 115 : Καρτέλα Αξιολόγησης των Αντιπροσώπων Πωλήσεων	166
Εικόνα 116 : Χώρος Εργασίας αφιερωμένος για την αξιολόγηση της εταιρείας – Power BI Service	167
Εικόνα 117 : Πίνακας ελέγχου αξιολόγησης της εταιρείας – Power BI Service	168
Εικόνα 118 : Ρόλοι στον χώρο εργασίας Olist.....	169
Εικόνα 119 : Διαφορετικές Άδειες θέασης – Power BI Service.....	169
Εικόνα 120 : Λειτουργία Endorsement – Power BI Service	170
Εικόνα 121 : Χώρος εμφάνισης των δημοσιοποιημένων μοντέλων δεδομένων – Power BI Service	170
Εικόνα 122 Παρακολούθηση σημαντικών KPIs μέσω των καρτελών Scorecards.....	171
Εικόνα 123 : Η εφαρμογή Olist στο Power BI Service	172

Εξισώσεις

Εξίσωση 1 : Ακαθάριστα έσοδα εταιρείας	152
Εξίσωση 2 : Μεικτό Περιθώριο Κέρδους	152
Εξίσωση 3 : Ρυθμός Ανάπτυξης Εσόδων ανά έτος.....	153

Διάρθρωση διπλωματικής

2.Συστήματα Επιχειρηματικής Ευφυΐας

Εδώ γίνεται ανάλυση του θεωρητικού υπόβαθρου των συστημάτων επιχειρηματικής ευφυΐας (**Business Intelligence - BI**). Αρχικά, γίνεται η κατάταξη των συστημάτων BI στην ευρύτερη οικογένεια, αναλυτικών συστημάτων και γίνεται διαχωρισμός αυτών από τα λειτουργικά συστήματα. Στη συνέχεια, αναλύονται σε βάθος οι Αποθήκες Δεδομένων ή συστήματα **DW/BI (Data Warehouse/ Business Intelligence)**, τα οποία περιλαμβάνουν και τον τρόπο αποθήκευσης και επεξεργασίας της πληροφορίας την οποία συλλέγει η εταιρεία, σε ειδικά διαμορφωμένες αποθήκες δεδομένων για σκοπούς του BI. Έπειτα, γίνεται αναφορά στην κουλτούρα των εταιρειών γύρω από συστήματα BI και κατάταξη αυτών των εταιρειών σε επίπεδα ωριμότητας αναφορικά με την ικανότητα τους να λαμβάνουν αποφάσεις με βάση δεδομένα.

Μετά γίνεται ένας βασικός διαχωρισμός μεταξύ των παραδοσιακών συστημάτων BI και των self-service συστημάτων BI, όπως επί παραδείγματι είναι η πλατφόρμα της Microsoft, **Power BI**. Στο τέλος της ενότητας εξερευνάται η αγορά των **self-service** πλατφορμών και γίνεται η σύγκριση μεταξύ των κορυφαίων εξ' αυτών.

3. Η υποδομή του self-service εργαλείου Power BI

Στην ενότητα αυτή αναλύεται η λειτουργικότητα της πλατφόρμας της Microsoft, Power BI. Αρχικά, γίνεται η επεξήγηση του ευρύτερου

οικοσυστήματος στο οποίο ανήκει το Power BI, το οποίο ονομάζεται Power Platform και αποτελεί μια συλλογή εφαρμογών και λογισμικών. Στην συνέχεια γίνεται αναφορά της τεχνολογίας, την οποία αξιοποιεί η πλατφόρμα. Το υπόλοιπο κεφάλαιο αναπτύσσεται γύρω από τα τέσσερα βασικά βήματα που απαιτούνται για κάθε πρότζεκτ στο Power BI. Αυτά τα τέσσερα βήματα είναι : «**1. Ανακάλυψη των Δεδομένων**», «**2. Μοντελοποίηση των Δεδομένων**», «**3. Οπτικοποίηση των Δεδομένων**» και «**4. Δημοσίευση των Δεδομένων.**» Παράλληλα με την ανάλυση αυτών βημάτων, γίνεται η επεξήγηση των επιμέρους συνιστωσών του Power BI, δηλαδή του **Power BI Desktop**, **Power BI Service** και **Power BI Mobile App**. Στο τέλος της ενότητας, απεικονίζεται πως είναι ένα ολοκληρωμένο σύστημα Power BI ,από μια high-level ματιά, σε επίπεδο ολόκληρης της εταιρείας.

4.Αξιολόγηση Εταιρείας Ηλεκτρονικού Εμπορίου

Σε αυτό το κεφάλαιο γίνεται η αξιολόγηση μιας εταιρείας, με βάση στην Βραζιλία, που δραστηριοποιείται στο **Ηλεκτρονικό Εμπόριο**. Γίνεται αξιοποίηση του εργαλείου Power BI για την δημιουργία αναφορών αξιολόγησης στον τομέα της εφοδιαστικής αλυσίδας της εταιρείας, των πωλήσεων ,των παραγγελιών, των πελατών, των πληρωμών και τέλος για το τμήμα μάρκετινγκ της εταιρείας. Έγιναν δύο κύριες αναλύσεις και μια συμπληρωματική. Η πρώτη ανάλυση σχετίζεται με την αξιολόγηση των συνηθειών των πελατών της εταιρείας και η δεύτερη με την αξιολόγηση των πωλήσεων, ενώ η συμπληρωματική ανάλυση σχετίζεται με την αξιολόγηση της δραστηριότητας των Πωλητών-Συνεργατών που είναι εγγεγραμμένοι στην πλατφόρμα της εταιρείας . Τέλος ,δημιουργήθηκε μια αναφορά για την αξιολόγηση των νέων

συμφωνιών που κλείνουν οι αντιπρόσωποι της εταιρείας με υποψήφιους Πωλητές-Συνεργάτες.

Συμπεράσματα

Στην τελευταία ενότητα της διπλωματικής, γίνεται η συγκέντρωση των σημαντικότερων ευρημάτων από την ανάλυση που έγινε στην εταιρεία και παράλληλα, με βάση αυτά γίνεται εκτίμηση για το μέλλον της εταιρείας. Στην συνέχεια ακολουθεί η γενική αποτίμηση των συστημάτων επιχειρηματικής ευφυΐας και οι λόγοι για τους οποίους είναι ζωτικής σημασίας η ύπαρξής τους στον σημερινό κόσμο των επιχειρήσεων.

Περίληψη

Η κύρια προσπάθεια της παρούσας διπλωματικής εργασίας είναι να αναδείξει τη σημασία των συστημάτων επιχειρηματικής ευφυΐας. Στο πρώτο μέρος της εργασίας, τίθενται τα θεωρητικά θεμέλια αυτών των συστημάτων. Αρχικά, τα Συστήματα Επιχειρηματικής Ευφυΐας κατατάσσονται στην ευρύτερη οικογένεια των αναλυτικών συστημάτων. Στη συνέχεια, τα διακρίνουμε σε παραδοσιακά και self-service συστήματα επιχειρηματικής ευφυΐας, όπου κάθε είδος διαθέτει μοναδικά πλεονεκτήματα και μειονεκτήματά. Έπειτα, συζητείται εκτενώς το υποκείμενο μοντέλο δεδομένων τέτοιων συστημάτων και ορίζεται ως πιο αποτελεσματικός τρόπος σχεδίασης, ο σχεδιασμός διαστάσεων. Όπως φανερώνεται, η χρήση κατάλληλων πρακτικών κατά τη δημιουργία ενός μοντέλου δεδομένων αποδεικνύεται ιδιαίτερα επωφελής στην πορεία δημιουργίας ενός συστήματος επιχειρηματικής ευφυΐας. Το μεγαλύτερο τμήμα της εργασίας σχετίζεται με την επίδειξη ενός self-service συστήματος επιχειρηματικής ευφυΐας. Αξιοποιώντας την πλατφόρμα της Microsoft, Power BI, δημιουργούμε μια επιχειρηματική έκθεση που αναλύει ένα πραγματικό επιχειρηματικό σενάριο. Τα επιχειρηματικά δεδομένα προέρχονται από μια σχετικά νέα επιχείρηση ηλεκτρονικού εμπορίου με έδρα τη Βραζιλία, την Olist. Γίνεται μια αξιολόγηση σε επίπεδο επιχειρηματικών διαδικασιών των τμημάτων της επιχείρησης, όπως το τμήμα Logistics, το τμήμα Marketing και το οικονομικό τμήμα. Στο τελευταίο μέρος, με βάση την ανάλυση που προηγήθηκε, γίνονται παρατηρήσεις και συστάσεις σχετικά με τη στρατηγική επιχειρησιακής ανάπτυξης της εν λόγω εταιρείας.

Abstract

The main strive of this thesis is to demonstrate the importance of Business Intelligence Systems. In the first part of the thesis, the theoretical foundation of such systems is being laid. Business Intelligence Systems are classified as a member of Analytical Systems. Later, we distinguish them into Traditional and Self-service Business Intelligence Systems, each type with its unique advantages and disadvantages. Afterwards, the underlying data model of such systems is extensively discussed, and dimensional design is appointed as the most efficient. Using proper practices when building a data model is proved greatly beneficial along the way of creating a BI System. Such practices are discussed in this thesis. The largest section of the thesis is the demonstration of a self-service Business Intelligence System. Exploiting the self-service Platform of Microsoft, Power BI, we create a BI report analysing a real business case scenario. The business data comes from a relatively new e-Commerce Enterprise based in Brazil, Olist. An evaluation on business-process level for departments of the company such as the Logistics, Marketing and Financial department is being made. In the last part, based on the analysis made with the Power BI tool, observations and recommendations are also made concerning the strategic business development of this enterprise.

Εισαγωγή

Η *Επιχειρησιακή Ευφυΐα ή Γνώση (Business Intelligence)* δεν είναι κάτι καινούργιο, αλλά υπάρχει εδώ και αιώνες σε διαφορετικές μορφές. Ο Sun Tzu, μεγάλη Κινέζικη στρατιωτική ιδιοφυΐα και φιλόσοφος, που έζησε κατά την περίοδο των εμπόλεμων κρατών πριν από περισσότερα από 2500 χρόνια στην Κίνα, στο βιβλίο του Η «Τέχνη του Πολέμου» ισχυρίστηκε ότι για να πετύχεις στον πόλεμο, είναι κρίσιμο να έχεις πλήρη γνώση της δικής σου δύναμης καθώς και των αδυναμιών σου. Η έλλειψη ενός από τα δύο μπορεί να οδηγήσει σε ήττα. Η δραστηριότητα των επιχειρήσεων σε ένα ανταγωνιστικό περιβάλλον είναι σαν να δίνεις μάχη. Στην επιχείρηση, οι εταιρείες είναι στρατοί. Οι ανταγωνιστές είναι οι εχθροί και το αποτέλεσμα του ανταγωνισμού μπορεί να είναι μοιραίο για τη μία ή την άλλη πλευρά. Όπως και στον πόλεμο, το κλειδί για την επιβίωση στο «εχθρικό έδαφος» είναι οι πληροφορίες. «Αν γνωρίζεις τον εαυτό σου και τον εχθρό σου, σε εκατό μάχες δεν θα φοβηθείς ποτέ το αποτέλεσμα. Όταν γνωρίζεις τον εαυτό σου αλλά όχι τον εχθρό, οι πιθανότητές σου να κερδίσεις ή να χάσεις είναι ίσες. Αν δεν γνωρίζεις ούτε τον εαυτό σου ούτε τον εχθρό σου, είναι βέβαιο ότι θα κινδυνεύεις σε κάθε μάχη», Sun Tzu, The Art of War. [1]

Ο «πατέρας» της Επιχειρηματικής Ευφυΐας θεωρείται ο Stefan Dedijer. Ο ίδιος υπηρέτησε ως αξιωματικός στον αμερικανικό στρατό και είχε τόνισε την ιστορική μετάβαση από τον πόλεμο στις επιχειρήσεις. Σύμφωνα με τον ίδιο, οι πληροφορίες στις επιχειρήσεις έχουν την ίδια βαρύτητα όσο και οι πληροφορίες στον πόλεμο [2]. Οι πληροφορίες, όπως συμπέρανε ο Dedijer, ήταν αναπόσπαστο και απαραίτητο μέρος κάθε συστήματος αυτομεγιστοποίησης, σε συνθήκες συναγωνισμού. Τα συστήματα αυτά είναι στρατοί, επιχειρήσεις, έθνη, ανθρώπινα όντα, ζώα, οικοσυστήματα, ακόμη και ολόκληρος ο κόσμος. Του άρεσε να αναφέρει τον Ηράκλειτο: «Ο πόλεμος είναι ο πατέρας των πάντων». **Όπου υπάρχει συναγωνισμός, πρέπει να εμπλέκεται η πληροφορία.** [1] Ως εκ τούτου, για να συμβαδίσουν με τον σημερινό ανταγωνιστικό κόσμο, και σε ορισμένες περιπτώσεις, για να «επιβιώσουν», πολλές εταιρείες σαρώνουν το περιβάλλον τους, παρακολουθούν στενά τον ανταγωνισμό τους και προσπαθούν να κατανοήσουν τους πελάτες τους. Οι εταιρείες δεν έχουν την πολυτέλεια να μείνουν αδιάφορες, ακόμα κι αν είναι ηγέτες στον κλάδο. Συγκεντρώνουν και χρησιμοποιούν πληροφορίες για να λάβουν τεκμηριωμένες αποφάσεις. [3]

Οι τρεις κλασικοί πόροι που είναι απαραίτητοι για μια βιομηχανική επιχείρηση, όπως περιγράφονται συχνά από οικονομολόγους είναι η γη, η εργασία και τα οικονομικά. [4] Όμως, τις τελευταίες δύο δεκαετίες, το επιχειρηματικό περιβάλλον έχει σταδιακά μετατραπεί σε μια λεγόμενη οικονομία βασισμένη στη πληροφορία. Συχνά λέγεται στον κόσμο των επιχειρήσεων, ότι τα δεδομένα είναι το νέο «πετρέλαιο» και αυτό σημαίνει ότι απαιτείται ένα σωστό σύστημα για να αξιοποιούνται αυτές τις πληροφορίες. Αυτές τις ανάγκες εξυπηρετούν συστήματα BI, επίγονος συστημάτων Υποστήριξης Λήψης Αποφάσεων (Decision Support System – DSS). Τα συστήματα αυτά υπάρχουν εδώ και δεκαετίες στην αγορά είτε συστηματικά κατασκευασμένα είτε ad-hoc εφαρμοζόμενα ανά περίπτωση.

Ενώ δεν υπάρχει κάποιο πρόβλημα με την αξιοποίηση παραδοσιακών συστημάτων BI ,η κατασκευή και συντήρηση τους απαιτεί πολυπληθές εξειδικευμένο προσωπικό και είναι φυσικά αρκετά χρονοβόρο και κοστοβόρο στις λύσεις που παρέχει. Με την ωρίμανση της τεχνολογίας πλέον, έχουν προκύψει **self-service BI συστήματα**, στα οποία οι διαδικασίες ενός παραδοσιακού BI συστήματος έχουν απλοποιηθεί και «εκδημοκρατιστεί». Γραμμένα σε low-level κώδικα, τα συστήματα αυτά βοηθούν τους end-users και business users ,που δεν είναι εξειδικευμένοι επαγγελματίες, να αξιοποιούν τα δεδομένα που συλλέγει ή αγοράζει η επιχείρηση, να τα μετατρέψουν σε πραγματικά ανταγωνιστικό πλεονέκτημα, με λιγότερο κόστος και σε λιγότερο χρόνο. Αξίζει, να σημειωθεί πως τα παραδοσιακά συστήματα και τα self-service δεν είναι ανταγωνιστικά συστήματα, αλλά συμπληρωματικά ,καθώς πάντα υπάρχουν ζητήματα τα οποία απαιτούν πιο μεγάλη εξειδίκευση.

1. Η επιλογή του θέματος της Διπλωματικής

Η αφετηρία αυτής της διπλωματικής φαινομενικά είναι λίγο παράδοξη και κάπως έκκεντρη σε σχέση με τον πυρήνα των σπουδών μου στο τμήμα των Μηχανολόγων Μηχανικών στο Πολυτεχνείο. Η αρχική μου επιθυμία ήταν η διπλωματική μου να σχετίζεται με το μάρκετινγκ.

Η επιθυμία μου αυτή δημιουργήθηκε , κατά την παρακολούθηση του κατ'εκλογήν υποχρεωτικού μαθήματος «Εισαγωγή στο Μάρκετινγκ» του 7^{ου} εξαμήνου , που το διδάσκει ο επιβλέπων της διπλωματικής μου, καθηγητής κ. Νικόλαος Παναγιώτου. Ο οποίος, ξεκίνησε τα μαθήματα με φράση κλειδί για εμένα , «Θα σας κάνω μάρκετινγκ για το μάρκετινγκ». Όπως , έμαθα το μάρκετινγκ έχει επιτελικό ρόλο και πολλές φορές ο διευθυντής του μάρκετινγκ είναι το «δεξί χέρι» του διευθύνοντα συμβούλου. Για παράδειγμα ,σε μια βιομηχανική εταιρεία ο ρόλος του μάρκετινγκ είναι **να συγκεράσει να αντιφατικούς στόχους** μεταξύ του τομέα της Παραγωγής και των Πωλήσεων. Ο τομέας της Παραγωγής επιθυμεί τα προϊόντα να είναι όσον το δυνατόν πιο απλά, λιγότερο εξειδικευμένα και προσωποποιημένα, με σκοπό το κόστος παραγωγής να είναι χαμηλό. Από την άλλη ,ο τομέας Πωλήσεων επιθυμεί να κάνει όσον τον δυνατόν μεγαλύτερο όγκο πωλήσεων και για να το πετύχει αυτό , απαιτεί εξειδίκευση και προσωποποίηση στον πελάτη. Σε αυτό το σημείο, δουλειά του στελέχους του μάρκετινγκ είναι να βρει την σωστή ισορροπία και να προτείνει στα μέλη της διοίκησης την πιο συμφέρουσα λύση.

Μια ακόμα λειτουργία που έχει το μάρκετινγκ, πέρα από το τομέα της διαφήμισης, είναι να μπορεί να διακρίνει τις ευκαιρίες και τις απειλές στο ρευστό οικοσύστημα που ανήκει η εταιρεία, να αξιολογεί τις δυνατότητες και αδυναμίες της εταιρείας και να προτείνει λύσεις.

Με λίγα λόγια , **το μάρκετινγκ** μπορεί να **έχει στρατηγικό χαρακτήρα**, κάτι το οποίο μου κεντρίζει το ενδιαφέρον. Παλαιότερα οι επιχειρήσεις λαμβάνανε αποφάσεις αποκλειστικά με το ένστικτο, την εμπειρία των στελεχών της και τις φημολογίες που υπήρχαν στην αγορά. Πλέον, όμως στην καθημερινή λειτουργία των εταιρειών έχουν μπει τα δεδομένα . Ένας τεράστιος όγκος δεδομένων καταναλώνεται καθημερινά από τις εταιρείες, δεδομένα που είτε δημιουργούν οι ίδιες είτε αγοράζουν, οδηγώντας πλέον την διαδικασία λήψης αποφάσεων όλο και περισσότερο βασιζόμενη στα δεδομένα. Στο σημείο αυτό ανακύπτει η ανάγκη των **συστημάτων BI** και αναδεικνύεται η εξαιρετική χρησιμότητα τους.

Τα συστήματα αυτά και συγκεκριμένα τα **self-service BI** συστήματα, στα οποία εστιάζει και η διπλωματική μου, επιτρέπουν στους χρήστες της επιχείρησης να έρθουν πιο κοντά στα δεδομένα και να τα «καταναλώνουν» με μεγαλύτερη ευχέρεια. Τα συστήματα αυτά απλοποιούν τις διαδικασίες συγκέντρωσης και επεξεργασίας δεδομένων, βοηθούν τους

χρήστες να δημιουργούν αναφορές και συγκεντρωτικούς πίνακες ελέγχου (dashboards) ,ενώ παράλληλα επιτρέπει την συνεργασία μεταξύ μελών της ομάδας, τόσο κατά την φάση κατασκευής των συστημάτων όσο και κατά την φάση συντήρησης ή αναβάθμισης αυτών, με μεγαλύτερη ασφάλεια και ταχύτητα. Συνολικά ,τα συστήματα αυτά δίνουν ένα ανταγωνιστικό πλεονέκτημα, παρέχοντας έναν ουσιαστικό τρόπο αξιοποίησης των δεδομένων που συλλέγονται από τις εταιρείες.

2.Συστήματα Επιχειρηματικής Ευφυΐας

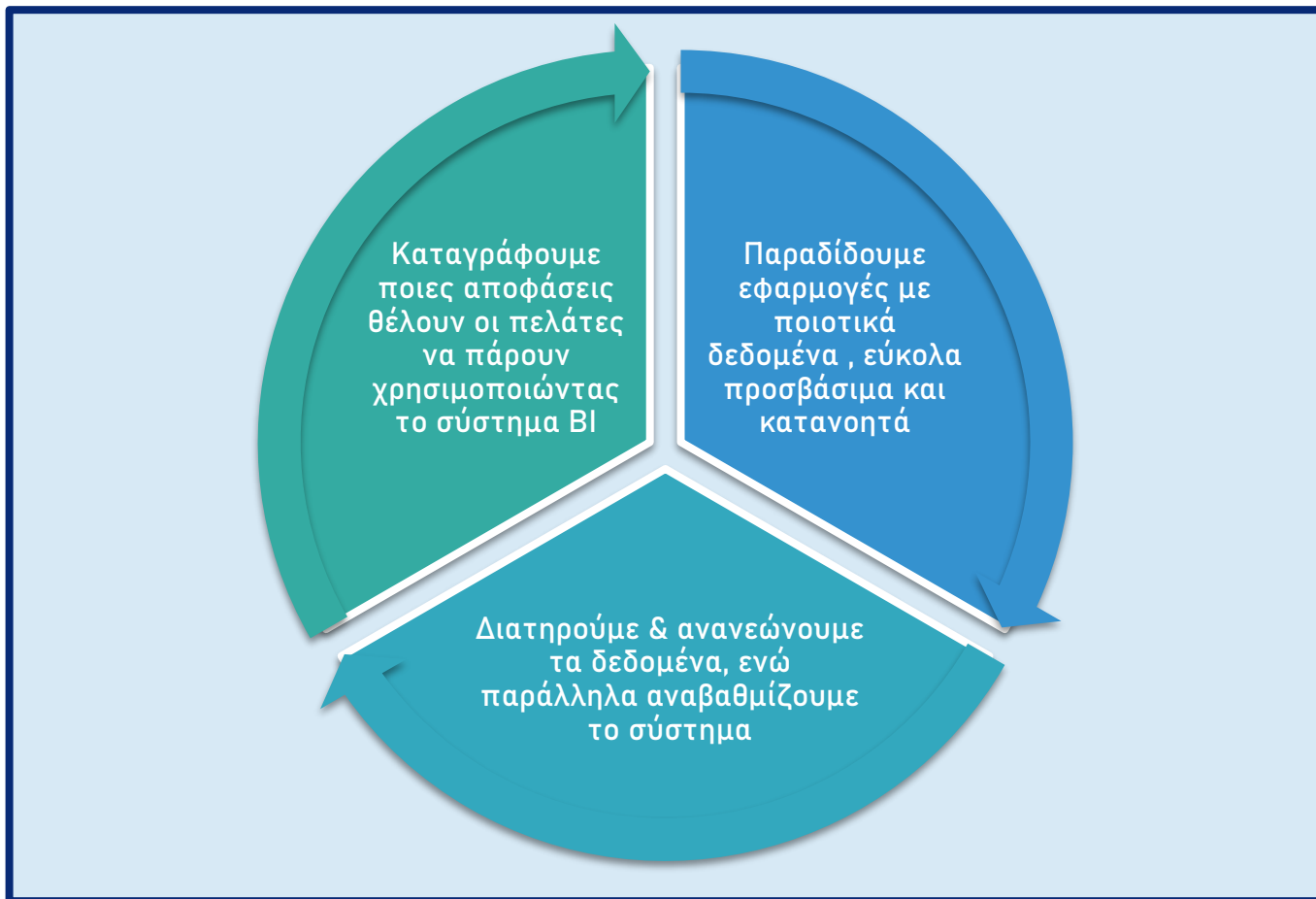
2.1 Ορισμός Επιχειρηματικής Ευφυΐας

Η **Επιχειρηματική Ευφυΐα** ή **Business Intelligence** ή **BI** είναι ένας τεχνικός όρος ο οποίος αναφέρεται στις «μεθόδους, τεχνολογίες και συστήματα που αφορούν τη *συλλογή, αποθήκευση, επεξεργασία, ανάλυση και εν τέλει οπτικοποίηση δεδομένων* για επιχειρησιακούς σκοπούς». Οι περισσότεροι επαγγελματίες BI θα όριζαν ως σκοπό ενός BI συστήματος «την λήψη πρωτογενών δεδομένων και η μετατροπή τους σε γνώση, η οποία ενημερώνει και βοηθάει τα στελέχη μια επιχείρησης να πάρουν αποφάσεις, σημαντικές για τον οργανισμό τους.» Το BI μπορεί να μας δώσει τα δεδομένα που χρειαζόμαστε για να λάβουμε μια τεκμηριωμένη απόφαση για το *τι συνέβη, πού και πότε συνέβη, γιατί συνέβη, ακόμα και τι μπορεί να συμβεί μετά*. Θα μπορούσε να θεωρηθεί ο απόγονος συστημάτων υποστήριξης για λήψη αποφάσεων (*Decision Support System - DSS*), (*Executive Information System - EIS*) και (*Manager Information Systems - MIS*).

Οι στόχοι του BI μπορούν να εντοπιστούν εύκολα, αν απλώς κανείς βρεθεί στις αίθουσες οποιουδήποτε οργανισμού και ακούγοντας την διοίκηση. Αυτά τα επαναλαμβανόμενα θέματα υπάρχουν για περισσότερα από τρεις δεκαετίες:

- «Συλλέγουμε τόνους δεδομένων, αλλά δεν μπορούμε να έχουμε πρόσβαση σε αυτά».
- «Πρέπει να μπορούμε να μεταχειριζόμαστε τα δεδομένα μας με κάθε τρόπο».
- «Οι επιχειρηματίες πρέπει να κατανοούν εύκολα τα δεδομένα».
- «Απλώς δείξε μου τι είναι σημαντικό».
- «Περνάμε ολόκληρες συναντήσεις μαλώνοντας για το ποιος έχει τα σωστά νούμερα».
- «Θέλουμε οι άνθρωποι να χρησιμοποιούν περισσότερο δεδομένα για να υποστηρίξουν αποφάσεις και όχι απλά το ένστικτό τους.» [5]

Συστήματα BI χρησιμοποιούνται οπουδήποτε χρειάζεται να αξιολογηθεί μια διαδικασία ενός οργανισμού. Αυτός οργανισμός ενδέχεται να είναι ένα σχολείο ,ένα νοσοκομείο ,μια επιχείρηση στο λιανεμπόριο, κάποιο τραπεζικό σύστημα, μια εταιρεία τηλεπικοινωνιών ,μια ασφαλιστική εταιρεία ή μια μεταφορική εταιρεία και ούτω καθεξής. Μερικά παραδείγματα λύσεων που θα μπορούσε να δώσει ένα BI σύστημα είναι οι αναφορές μπάντζετ (**Budget Reporting**) για το τμήμα Μάρκετινγκ, προβλέψεις πωλήσεων (**Sales Forecasting**) για το τμήμα Πωλήσεων, καταμέτρηση εργαζομένων (**Employee Headcount**) για λογαριασμό του τμήματος Διαχείρισης Προσωπικού, να αξιολογήσει ένα σύστημα CRM, να βοηθήσει στην αξιολόγηση των πωλητών ή του προσωπικού, να εντοπίσει τους πιο πιστούς πελάτες ,να βοηθήσει στην ανάπτυξη ενός εκλεπτυσμένου συστήματος μανάτζμεντ ή να αναδείξει κυριάρχους παράγοντες που επηρεάζουν μια μέτρηση που μας ενδιαφέρει (**Key Influencers**).

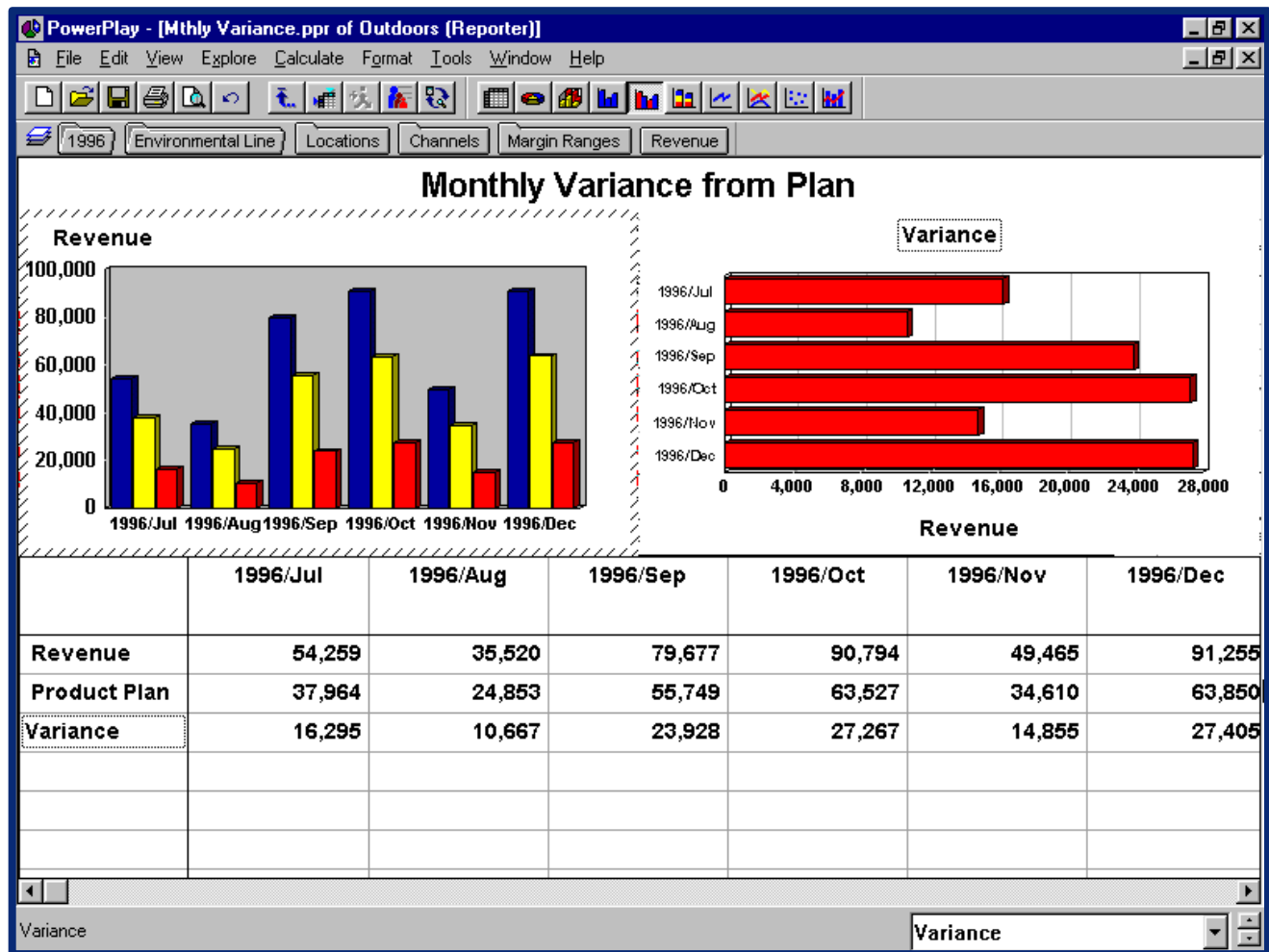


Σχήμα 1 : Κύκλος ζωής ενός συστήματος BI

2.1.1 Ιστορικά στοιχεία για τα συστήματα Επιχειρηματικής Ευφυΐας

Τα πρώτα ολοκληρωμένα συστήματα BI αναπτύχθηκαν από την IBM και τη Siebel Systems (που αποκτήθηκαν από την Oracle το 2006) μεταξύ 1970 και 1990. Οι αρχικές εκδόσεις των συστημάτων αυτών ήταν αποθήκες δεδομένων που χρησίμευαν ως κεντρικές αποθήκες ενσωματωμένων δεδομένων από μία ή περισσότερες διαφορετικές πηγές. Αυτές οι αποθήκες δεδομένων ήταν τεχνικά πολύπλοκες και απαιτούσαν εκτεταμένο προσωπικό πληροφορικής (IT) για τη συντήρηση και τη διαχείρισή τους. Χρειάζονταν προσωπικό εξειδικευμένο στο BI για την εξαγωγή πληροφοριών από τα δεδομένα και τη δημιουργία αναλυτικών αναφορών για ξεχωριστά τμήματα και επιχειρηματικές λειτουργίες. Με άλλα λόγια, η κατοχή και η λειτουργία του BI ήταν μια δαπανηρή υπόθεση, βιώσιμη αποκλειστικά για επιχειρήσεις πλούσιες σε πόρους.

Με την εξέλιξη της τεχνολογίας, οι προμηθευτές εργαλείων BI άρχισαν να αναπτύσσουν εργαλεία που δεν ήταν τόσο δαπανηρά ή τεχνικά περίπλοκα ,ενώ παράλληλα εμφανίστηκε ο πίνακας ελέγχου (**dashboards**). Οι πίνακες ελέγχου εξακολουθούσαν να εξαρτώνται από την εμπλοκή του προσωπικού IT για τη λήψη των δεδομένων, αλλά έκαναν το BI προσβάσιμο στους χρήστες ,επιτρέποντας τους να δημιουργούν αναφορές και να θέτουν ερωτήματα στην βάση δεδομένων (**querying**) οι ίδιοι.



Εικόνα 1 : Πίνακας ελέγχου (Dashboard) ιστορικών συστημάτων BI – IBM – Cognos Powerplay (1997)

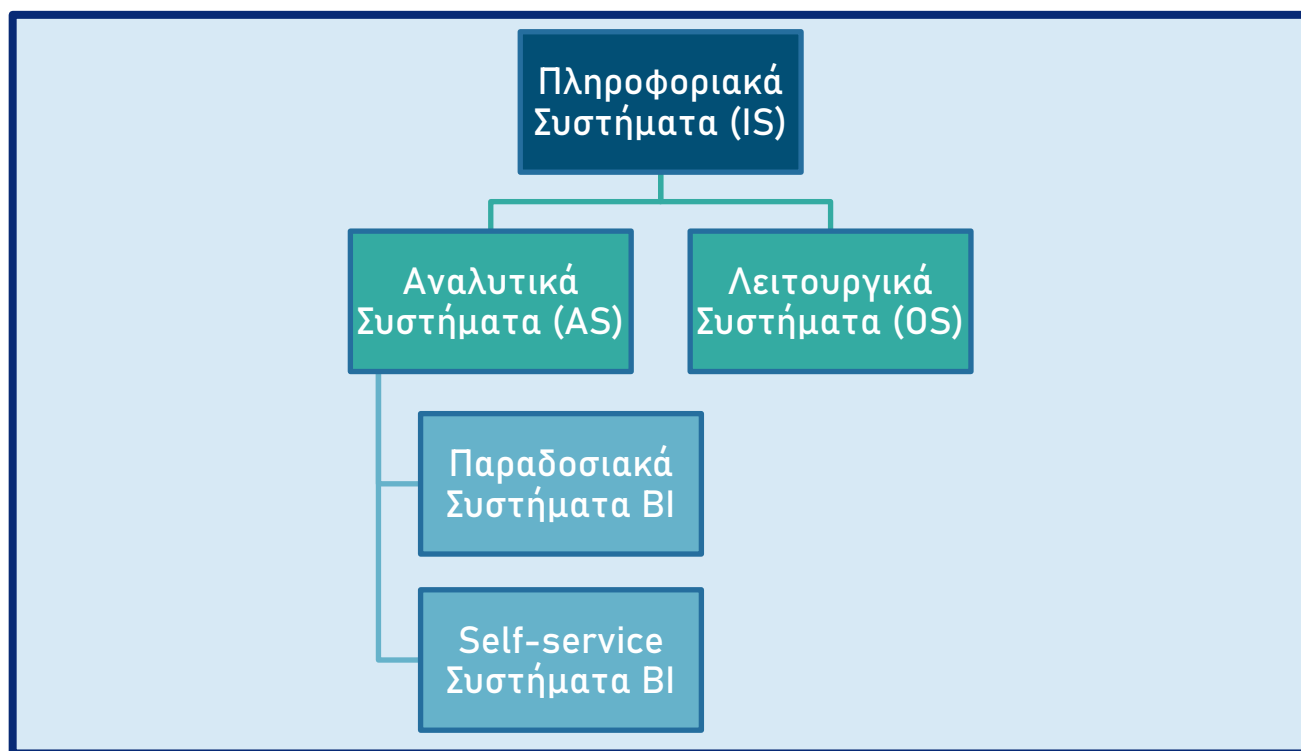
Μεγάλες εταιρείες τεχνολογίας όπως η *Microsoft*, η *IBM*, η *Oracle*, η *SAP* και η *SAS* ήταν οι πρωτοπόροι αυτής της εποχής BI και κυριάρχησαν στον χώρο των πλατφορμών για BI and Analytics. Σήμερα, οι εταιρίες αυτές έχουν φέρει στο προσκήνιο εδώ και αρκετά χρόνια ,self-service BI εργαλεία όπως το **Power BI** της **Microsoft** ή το **Tableau** της **Salesforce** ,τα

οποία είναι cloud-based SaaS πλατφόρμες ,όπου επιτρέπουν στους χρήστες να εκτελούν πρότζεκτ του BI και να τα μοιράζονται με τους συνεργάτες τους στο cloud. [6]

2.2 Πληροφοριακά Συστήματα

Τα συστήματα BI, με τα οποία ασχολείται η παρούσα διπλωματική ,αποτελούν μια υποκατηγορία των *Πληροφοριακών συστημάτων (Information Systems – IS)*. Τα συστήματα αυτά που χρησιμοποιούνται από τις σημερινές επιχειρήσεις, επιτρέπουν την εκτέλεση των λειτουργικών διαδικασιών των επιχειρήσεων ,τη συλλογή και διαχείριση των δεδομένων τους ,καθώς και αξιολόγηση των διαδικασιών αυτών.

Τα πληροφοριακά συστήματα χωρίζονται σε δύο μεγάλες κατηγορίες: αυτά που υποστηρίζουν την *εκτέλεση επιχειρηματικών διαδικασιών* και εκείνα που υποστηρίζουν την *ανάλυση των επιχειρηματικών διαδικασιών*. Τα πρώτα ονομάζονται *Λειτουργικά συστήματα (Operational Systems – OS)* και τα δεύτερα είναι τα *Αναλυτικά συστήματα (Analytical Systems – AS)* .



Διάγραμμα 1 : Πληροφοριακά Συστήματα

Τα αναλυτικά συστήματα και τα λειτουργικά συστήματα εξυπηρετούν θεμελιωδώς διαφορετικούς σκοπούς, το οποίο με τη σειρά του υποδηλώνει ότι διαφορετικές αρχές καθοδηγούν το σχεδιασμό τους ,εννοώντας μεταξύ άλλων, το βαθμό καθετοποίησης του μοντέλου δεδομένων (**Degree of Normalization**). Συνήθως στα αναλυτικά συστήματα ακολουθείται ο Διαστασιακός Σχεδιασμός (**Dimensional Design**). Τα BI συστήματα αποτελούν μέλος της οικογένειας των αναλυτικών συστημάτων. Αναλόγως το πώς είναι τεχνικά στημένα, τα συστήματα BI διαιρούνται περαιτέρω σε δύο κατηγορίες. Αυτά που ονομάζουμε «**Παραδοσιακά**» BI συστήματα και τα «**Self-service**». Αυτά τα συστήματα μπορούν να έχουν κάποια κοινά μέρη ή να συνεργάζονται μεταξύ τους. Η βασική τους διαφοροποίηση έγκειται στην πρόσβαση των χρηστών ή «καταναλωτών» στα δεδομένα του συστήματος αυτού. Στην περίπτωση των «παραδοσιακών» συστημάτων ,η εγκατάσταση τους, ανάπτυξη τους ,η συντήρηση και η αξιοποίηση τους αφορά κατά αποκλειστικότητα επαγγελματίες του τμήματος IT της εταιρείας. Στην περίπτωση των «self-service» συστημάτων, η χρήση τους μπορεί να γίνεται από έναν αρκετά πιο διευρυμένο αριθμό χρηστών που δεν ανήκουν στο τμήμα των πληροφορικής, καθώς είναι ουσιαστικά απλούστερα στην χρήση και συνήθως δίνουν την δυνατότητα στον χρήστη υποβολής ερωτημάτων στην βάση δεδομένων σχεδόν σε φυσική γλώσσα (*Natural Language Querying*). [5]

2.2.1 Λειτουργικά Συστήματα

Ένα λειτουργικό σύστημα (*Operational System – OS*) υποστηρίζει την *άμεση εκτέλεση μιας επιχειρηματικής διαδικασίας* με την αποτύπωση λεπτομερειών σχετικά με σημαντικά γεγονότα ή συναλλαγές, και δημιουργεί ένα **αρχείο δραστηριοτήτων**. Ένα σύστημα πωλήσεων, για παράδειγμα, καταγράφει πληροφορίες σχετικά με παραγγελίες, αποστολές και επιστροφές. Ένα σύστημα διαχείρισης ανθρωπίνων πόρων συλλέγει πληροφορίες σχετικά με την πρόσληψη ,την προαγωγή εργαζομένων και διαχειρίζεται τις πληρωμές του προσωπικού. Ένα λογιστικό σύστημα συλλέγει πληροφορίες σχετικά με τη διαχείριση των χρηματοοικονομικών περιουσιακών στοιχείων και υποχρεώσεων της επιχείρησης. Οι δραστηριότητες που καταγράφονται από αυτά τα συστήματα είναι μερικές φορές γνωστές ως «*συναλλαγές*». Τα ίδια συστήματα ονομάζονται μερικές φορές «*online transaction processing systems – OLTP*» ή «**συστήματα συναλλαγών**» για συντομία.

Τα λειτουργικά συστήματα πρέπει να επιτρέπουν πολλές τύπους αλληλεπίδρασης με τις βάσεις δεδομένων, συμπεριλαμβανομένων των εισαγωγών (*insert*), ενημερώσεων (*update*) και διαγραφών (*delete*). Οι αλληλεπιδράσεις αυτές είναι σχεδόν πάντα **ατομικές** ,μια συγκεκριμένη παραγγελία, μια αποστολή, μια επιστροφή χρημάτων. Επίσης, *είναι* εξαιρετικά προβλέψιμες στη φύση τους. Για παράδειγμα, ένα σύστημα εισαγωγής παραγγελιών πρέπει να παρέχει τη λίστα προϊόντων, πελατών και πωλητών ,την εισαγωγή παραγγελιών, την εκτυπωμένη περίληψη παραγγελιών, τιμολογίων και λιστών συσκευασίας ,καθώς και την κατάσταση παρακολούθησης της παραγγελίας.

Επειδή αυτά τα συστήματα επικεντρώνονται στην εκτέλεση της διαδικασίας, το λειτουργικό σύστημα είναι πιθανό να *ενημερώνει* τα δεδομένα καθώς προκύπτουν αλλαγές, να *εκκαθαρίσει* ή να *αρχειοθετήσει* δεδομένα μόλις τελειώσει η λειτουργική τους χρησιμότητα. Όταν ένας πελάτης μετακομίζει, για παράδειγμα, η παλιά του/της διεύθυνση δεν είναι πλέον χρήσιμη για την αποστολή προϊόντων ή αποστολή τιμολογίων, έτσι απλά αντικαθίσταται. Δηλαδή, **το λειτουργικό σύστημα δεν αποθηκεύει ιστορικά δεδομένα** .

Στα λειτουργικά συστήματα, βρίσκουν εφαρμογή σχεσιακές βάσεις δεδομένων (**Relational Database**) και ο βέλτιστος σχεδιασμός τους (*Optimal schema*), θεωρείται ότι είναι όταν το σύστημα βρίσκεται σε κάποια κανονικοποιημένη μορφή (*Normalized format*) ,συνήθως είτε στην 3^η είτε στην 5^η (*Third Normal Form - 3NF* ή *Fifth Normal Form - 5NF*). Η σχεδίαση μπορεί να απεικονιστεί ως ένα μοντέλο σχέσης οντότητας ή μοντέλο **ER** (*Entity Relationship*) .

Τα διαγράμματα σχέσεων οντοτήτων (*διαγράμματα ER* ή *ERD*) είναι σχέδια που επικοινωνούν τις σχέσεις μεταξύ των πινάκων. Και τα μοντέλα 3NF και διαστάσεων μπορούν να αντιπροσωπεύονται σε ERD επειδή και οι δύο αποτελούνται από ενωμένους σχεσιακούς πίνακες. Η διαφορά μεταξύ μοντέλων 3NF και μοντέλων διαστάσεων είναι ο βαθμός κανονικοποίησης. [5] *

* Αξίζει να σημειωθεί ότι, όπως θα δείχτεί στο κεφάλαιο του «Διαστασιακού Σχεδιασμού», οι πίνακες γεγονότων (**Fact Tables**) οργανώνονται συνήθως σε 3NF μορφή.

2.2.2 Αναλυτικά Συστήματα

Ενώ το επίκεντρο του λειτουργικού συστήματος είναι η *εκτέλεση μιας επιχειρηματικής διαδικασίας*, το **αναλυτικό σύστημα υποστηρίζει την αξιολόγηση της διαδικασίας**. Τα λειτουργικά συστήματα εξυπηρετούν τις καθημερινά απαιτούμενες διαδικασίες της επιχείρησης (*day to day operations*). Τα αναλυτικά συστήματα, όπως υποδεικνύει και το όνομα τους, έπονται χρονικά και συνήθως σχετίζονται με την αξιοποίηση «ιστορικών» δεδομένων, όπου ενδέχεται να έχουν προκύψει από τα λειτουργικά συστήματα. Τα συστήματα Επιχειρηματικής Ευφυΐας, φυσικά είναι μέλος της οικογένειας των αναλυτικών συστημάτων.

Τα αναλυτικά συστήματα απαντάνε σε ερωτήματα, όπως:

- Πώς κινούνται οι παραγγελίες αυτόν τον μήνα σε σχέση με τον τελευταίο;
- Πού μας βάζει αυτό σε σύγκριση με τους στόχους πωλήσεών μας για το τρίμηνο;
- Η προώθηση του μάρκετινγκ είχε αντίκτυπο στις πωλήσεις;
- Ποιοι είναι οι καλύτεροι πελάτες μας;

- Ποιοι είναι οι πιο αποδοτικοί ή συνεπείς εργαζόμενοι;
- Ποιοι είναι οι καλύτεροι πωλητές μας;

Αν γίνεται αξιολόγηση των παραγγελιών μιας επιχείρησης για παράδειγμα ,οι ερωτήσεις μπορεί να αφορούν τη μέτρηση των *συνολικών παραγγελιών* κάποιας συγκεκριμένης χρονικής περιόδου, αντί να ρωτούν για κάποια μεμονωμένη παραγγελία ,όπως συμβαίνει στα OS. Η αλληλεπίδραση με ένα αναλυτικό σύστημα πραγματοποιείται αποκλειστικά μέσω «ερωτημάτων» (*queries*) που ανακτούν δεδομένα από τις διάφορες επιχειρηματικές διαδικασίες. **Τα δεδομένα αυτά δεν δημιουργούνται, ούτε τροποποιούνται** αλλά φιλτράρονται για να απαντήσουν σε ερωτήματα που ενδιαφέρουν την επιχείρηση. Αυτά τα ερωτήματα μπορούν να περιλαμβάνουν μεγάλο αριθμό συναλλαγών, σε αντίθεση με την τυπική εστίαση του λειτουργικού συστήματος σε μεμονωμένες συναλλαγές. Οι συγκεκριμένες ερωτήσεις που τίθενται στην βάση δεδομένων είναι λιγότερο προβλέψιμες και πιο πιθανό αλλάζουν με το χρόνο. **Τα ιστορικά δεδομένα που συλλέγονται παραμένουν σημαντικά** για το αναλυτικό σύστημα, για πολύ καιρό μετά την ολοκλήρωση της λειτουργικής του χρήσης.

Για τα αναλυτικά συστήματα, η μοντελοποίηση διαστάσεων (*Dimensional Modelling*) είναι η ενδεδειγμένη ως προς το μοντέλο δεδομένων (*Data Model*). Αναλόγως, σε τι είδους βάση δεδομένων γίνεται η εφαρμογή, έχουμε δύο είδη μοντελοποίησης διαστάσεων: την μοντελοποίηση σε *Σχήμα Αστέρι (Star Schema)* ,αν εφαρμόζεται σε Σχεσιακή Βάση Δεδομένων (*Relational Database - RDB*) και τον *Κύβο (Cube)* ,αν εφαρμόζεται σε πολυδιάστατη βάση δεδομένων (*Multi-Dimensional Database - MDB*).**

** Μόνο το μοντέλο **Σχήμα Αστέρι** παρουσιάζεται στην παρούσα διπλωματική εργασία.

Στα OS συναντάμε μεγάλη καθετοποίηση των δεδομένων (*nth Nominal Form*) ,όπου οι απαιτήσεις του συστήματος είναι πάντα τα δεδομένα να είναι σε **ατομική μορφή**. Ατομικά δεδομένα είναι τα δεδομένα τα οποία έχουν συλλεχθεί στο επίπεδο εστίασης με την μεγαλύτερη δυνατή λεπτομέρεια ,δηλαδή δεν γίνεται δηλαδή να αναλυθούν σε χαμηλότερο επίπεδο πληροφορία. Η καταγραφή *μιας συναλλαγής με ένα πελάτη*. Ένα επίπεδο υψηλότερο Στα AS ,μπορούμε να έχουμε ατομικά δεδομένα ,χωρίς όμως να είναι απαραίτητο. Αυτό, εξαρτάται από την επιλογή του *επιπέδου αποθήκευσης των δεδομένων* ,που έχει αποφασίσει ο σχεδιαστής του συστήματος στην αρχή.

Η σχεδίαση διαστάσεων θεωρείται ευρέως ως, η ιδανική για αναλυτικά συστήματα ,παρά ταύτα εξακολουθούμε να συναντάμε υποτιθέμενα συστήματα BI όπου, η σχεδίαση τους είναι μια αντιγραφή των λειτουργικών συστημάτων, με τα αρχεία τους απλά να είναι αποθηκευμένα σε μια ξεχωριστή πλατφόρμα. [5]

	Λειτουργικά Συστήματα (OS)	Αναλυτικά Συστήματα (AS)
Σκοπός	Εκτέλεση διαδικασιών επιχείρησης	Μέτρηση/Αξιολόγηση μιας διαδικασίας της επιχείρησης
Κυρίαρχη αλληλεπίδραση με δεδομένα	Εισαγωγή, Ανανέωση, Query, Διαγραφή	Ερώτηση στην βάση (Query)
Μοτίβο querying	Προβλεπόμενο και σταθερό	Απρόβλεπτο και εναλλασσόμενο
Σκοπιά εστίασης	Παρόντα δεδομένα	Παρόντα & Ιστορικά δεδομένα
Αρχή Σχεδίασης	Entity - Relationship (ER) design in third normal form (3NF)	Dimensional design (Star Schema ή Cube)
Άλλες ονομασίες	Σύστημα συναλλαγών Online Transaction Processing System (OLTP) Source System	Data Warehouse System Data Mart

Πίνακας 1 : Λειτουργικά vs Αναλυτικά Συστήματα

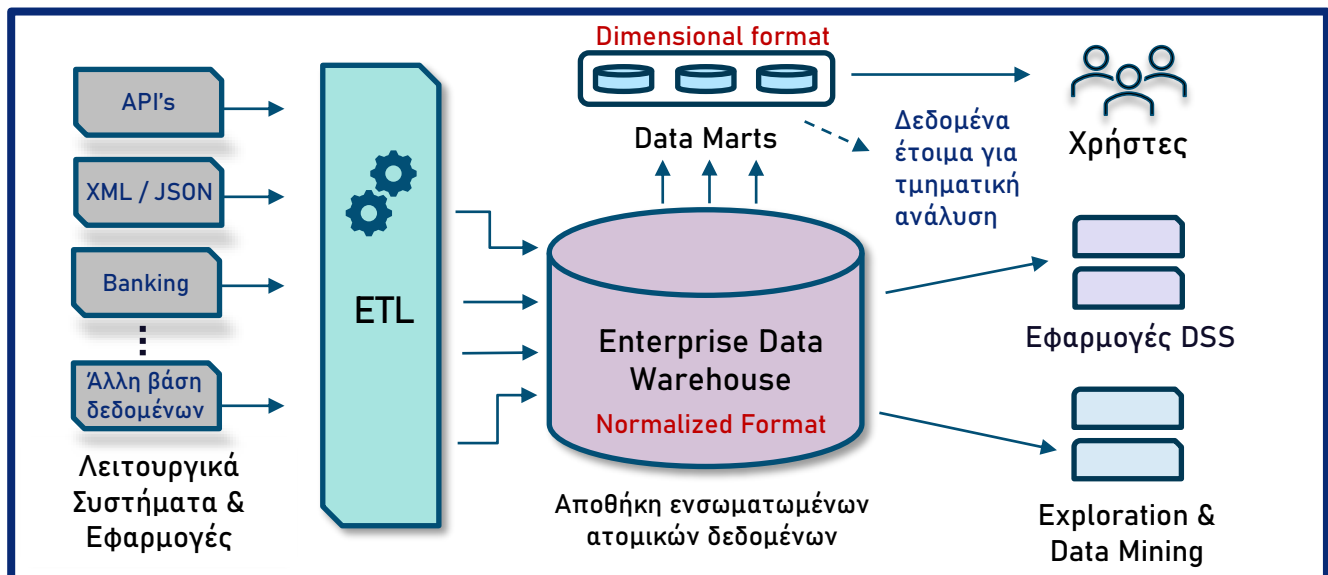
2.3 Αρχιτεκτονική αναλυτικών συστημάτων DW/BI

Στην κλάδο των αναλυτικών συστημάτων συνήθως τα συστήματα BI συχνά αναφέρονται και ως **DW/BI** ή **DWBI**. Αυτή η ονομασία προκύπτει από τις λέξεις **Data Warehouse / Business Intelligence**, καθώς το Data Warehousing θεωρείται βασικό κομμάτι για την ορθολογική ανάπτυξη ενός ολοκληρωμένου συστήματος *Επιχειρηματικής Ευφυΐας*. Η «Αποθήκευση Δεδομένων» (DW) είναι μια διαδικασία συλλογής, ενσωμάτωσης και διαχείρισης δεδομένων από διάφορες πηγές για την παροχή ουσιαστικών επιχειρηματικών πληροφοριών. Τα DW διαθέτουν κεντρικές αποθήκες ενσωματωμένων δεδομένων από μία ή περισσότερες διαφορετικές πηγές. Αποθηκεύουν τρέχοντα και ιστορικά δεδομένα σε ένα μόνο μέρος είτε φυσικό είτε λογικό, όταν έχουμε πολλαπλές βάσεις δεδομένων. [5]

2.3.1 Οι αρχιτεκτονικές συστημάτων DW/BI των R.Kimball και W.H.Inmon

Στα περισσότερα παραδοσιακά συστήματα BI η αρχιτεκτονική τους προσομοιάζει κάποια από αυτές που πρότειναν δύο αρχιτέκτονες αποθηκών δεδομένων για αναλυτικά συστήματα (**Data Warehouses**). Αυτές δεν είναι άλλες από του **R.Kimball** και του **W.H.Inmon**.

Οι δύο αυτές αρχιτεκτονικές αξιοποιούν αρχές του διαστασιακού μοντελισμού (**Star Schema**) σε κάποια σημεία τους αλλά όχι αναγκαστικά στα ίδια, το οποίο θα αναλυθεί σε επόμενο κεφάλαιο.



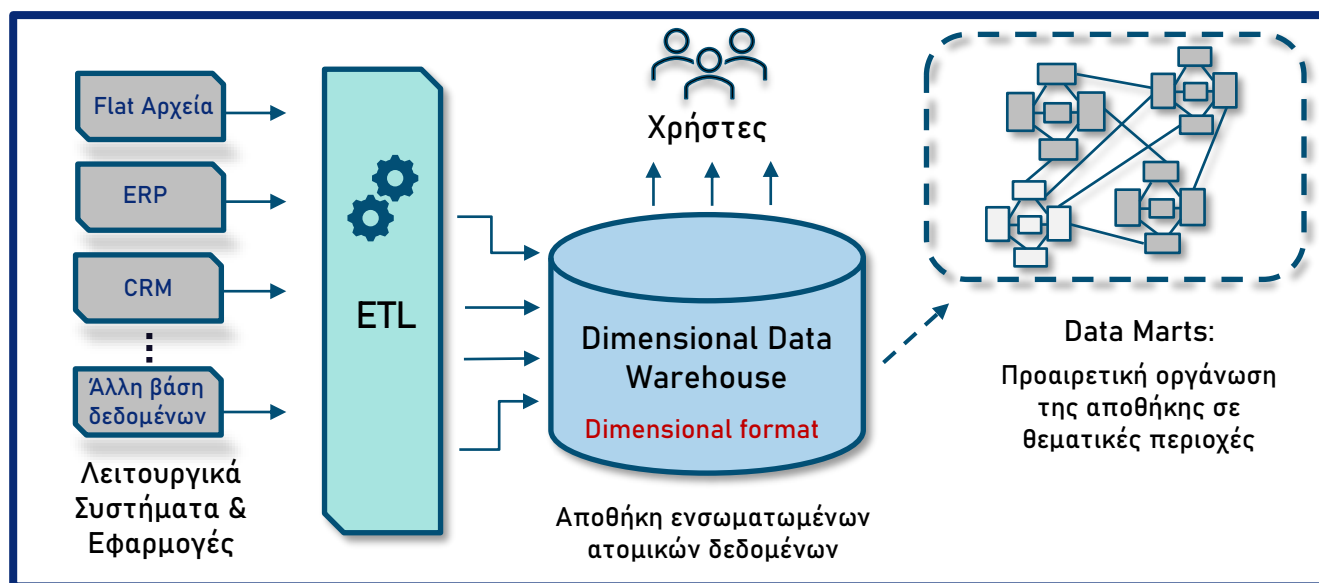
Σχήμα 2 : Η αρχιτεκτονική της αποθήκης δεδομένων του W.H.Inmon : the Corporate Information Factory

Και οι δύο αρχιτεκτονικές ξεκινούν με την υπόθεση ενός **διαχωρισμού του λειτουργικού και του αναλυτικού συστήματος**. Τα Λειτουργικά Συστήματα εμφανίζονται στο αριστερό άκρο του **Σχήματος 2** & **Σχήματος 3**. Αυτά τα συστήματα μπορεί ενσωματώνουν αποθήκες δεδομένων που είναι σχεσιακές ή όχι, ενώ ενδέχεται να είναι πολυάριθμες. Η πιο συνήθης μορφή τους είναι να βρίσκονται σε κανονικοποιημένη μορφή (**Normalized Format**).

Προχωρώντας προς τα δεξιά του **Σχήματος 2** & **Σχήματος 3**, ακολουθεί μια διαδικασία με την ονομασία **ETL**, όπου ενοποιεί πληροφορίες από τα διάφορα λειτουργικά συστήματα, τα ενσωματώνει και τα φορτώνει σε ένα ενιαίο χώρο αποθήκευσης. Τα αρχικά ETL προκύπτουν από τις αγγλικές λέξεις «Extract», «Transform» και «Load», δηλαδή «Εξαγωγή», «Μορφοποίηση» και «Φόρτωση» των δεδομένων. Σκοπός αυτού του συστήματος ή καλύτερα σύνολο συστημάτων είναι να εισάγει, να καθαρίσει και να

προετοιμάσει τα δεδομένα τα οποία λαμβάνει από τα διάφορα πηγαία συστήματα (**source systems**) όπως βάσεις δεδομένων, υπολογιστικά φύλλα, αρχεία κειμένου, σχεσιακά δεδομένα, ιεραρχικά δεδομένα και ούτω καθεξής και εν τέλει να φορτώσει τα δεδομένα στην ειδικά διαμορφωμένη αποθήκη δεδομένων.

Στη συνέχεια, γίνεται η ενσωμάτωση όλων των δεδομένων μαζί, σε ένα κεντροποιημένο μέρος όπως μια Αποθήκη Δεδομένων (*Data Warehouse*) που είναι ένα ειδικό είδος βάσης δεδομένων που χρησιμοποιείται για σκοπούς της Επιχειρηματικής Ευφυΐας. Και στις δύο περιπτώσεις αποθήκης δεδομένων, είτε του Inmon είτε του Kimball πρόκειται για ένα **ολοκληρωμένο αποθετήριο ατομικών δεδομένων**.



Σχήμα 3 : Η αρχιτεκτονική της αποθήκης δεδομένων του Ralph Kimball : Dimensional Data Warehouse

Έχοντας ενσωματώσει τα διάφορα λειτουργικά συστήματα, η **Αποθήκη Δεδομένων περιέχει μια οριστική και συνεπή αναπαράσταση των επιχειρηματικών δραστηριοτήτων** σε έναν ενιαίο χώρο. Τα δεδομένα σε αυτό το αποθετήριο καταγράφονται στο χαμηλότερο δυνατό επίπεδο λεπτομέρειας, γι' αυτό ονομάζονται και ατομικά. Η αρχιτεκτονική δομείται έτσι, διότι είναι κατάλληλα σχεδιασμένη για σκοπούς του BI, όπου οι **χρήστες στέλνουν ερωτήματα στην βάση δεδομένων (queries) που συνήθως αφορούν μεγάλους όγκους δεδομένων**, σε αντίθεση με τα λειτουργικά συστήματα που συνήθως στέλνονται ερωτήματα που απαιτούν μια απάντηση, ένα δεδομένο, δηλαδή κάποια είδους συναλλαγή. [5]

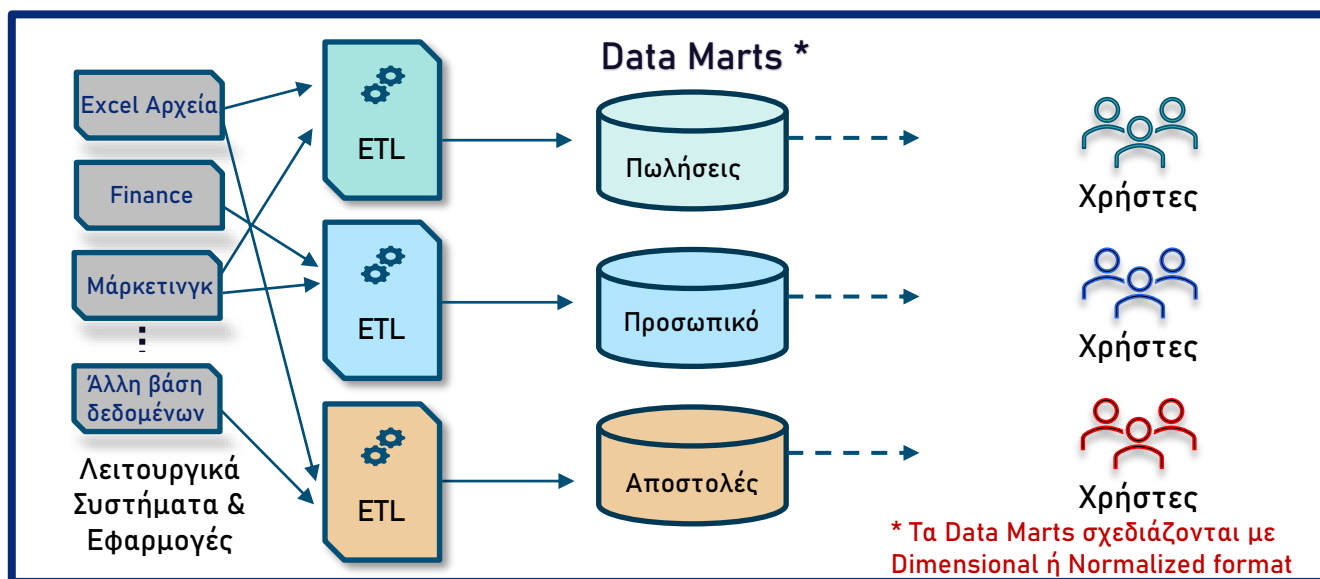
Η διαφοροποίηση των δύο αρχιτεκτονικών έγκειται **στα σημεία του αναλυτικού συστήματος, όπου έχουν πρόσβαση οι χρήστες για να ανακτήσουν πληροφορίες**. Στην περίπτωση του «*The Corporate Information Factory*», η κεντρική αποθήκη είναι σε

κανονικοποιημένη μορφή (3NF) και είναι βασισμένη σε αρχές του ER modelling και οι χρήστες δεν έχουν άμεση πρόσβαση σε αυτή, αλλά πρώτα παρεμβάλλονται κάποια **Data Marts** . Αυτά είναι σε διαστασιακή μορφή και υποστηρίζουν την «τμηματική σκοπιά της πληροφορίας» (departmental view of information). Στην περίπτωση του «*Dimensional Data Warehouse*» , τα «*Data Marts*» βρίσκονται εντός της Αποθήκης Δεδομένων και βρίσκονται ήδη σε διαστασιακή μορφή. Ουσιαστικά πρόκειται για ένα λογικό διαχωρισμό. Σε αυτή την αρχιτεκτονική οι χρήστες των αναλυτικών συστημάτων μπορούν να έχουν άμεση πρόσβαση στην πληροφορία.

Αξίζει να σημειωθεί ότι παρόλο που υποδεικνύεται ότι έχουμε *μια* κεντρικοποιημένη αποθήκη στο **Σχήμα 2 & Σχήμα 3**, στην πραγματικότητα αυτό είναι μια λογική απεικόνιση και άρα, μπορούμε να έχουμε πολλαπλές φυσικές βάσεις δεδομένων που να την αποτελούν.

2.3.2 Stand-alone Marts

Για λόγους πληρότητας θα ήταν να χρήσιμο να αναφέρουμε τις αρχιτεκτονικές **Stand-alone Marts**, δηλαδή αυτόνομα αποθηκευτήρια τα οποία, ενώ συναντώνται συχνά στην αγορά, δεν αποτελούν καλή πρακτική και προορίζονται κυρίως για τμηματική ανάλυση (**departmental analysis**). Βασικό μειονεκτήματα τους είναι ότι δεν παρέχουν δυνατότητα για **cross-functional** ανάλυση, δηλαδή για διασταύρωση της πληροφορίας ανάμεσα σε διαφορετικά τμήματα της επιχείρησης, καθώς ίδιες οντότητες σε κάθε ξεχωριστό mart μπορεί να ορίζονται με διαφορετικό τρόπο και να κάνουν αδύνατη την σύγκριση της πληροφορίας. Επίσης, σημαντικό μειονέκτημα είναι ότι αποθηκεύουν την πληροφορία όχι απαραίτητα στο επιθυμητό επίπεδο πληροφορίας, δηλαδή σε ατομικό επίπεδο (*atomic data – high granularity*) και **άρα δεν είναι ευλύγιστα συστήματα ως προς διαφορετικές ερωτήσεις που μπορεί να θέσει ένας χρήστης στο μέλλον**. Στα πλεονεκτήματα αυτών των συστημάτων συγκαταλέγουμε την ευκολία εγκατάστασης τους. Αυτό τα καθιστά καλή επιλογή για πρότζεκτ με βραχύ χρονικό ορίζοντα ολοκλήρωσης ή για ένα περιορισμένο φάσμα εργασιών μια επιχείρησης, δηλαδή για τμηματική ανάλυση. Παραδείγματα λύσεων στα οποία αξιοποιούνται τέτοια συστήματα αποτελούν το τμήμα πωλήσεων, τμήμα μάρκετινγκ ή το τμήμα προσωπικού.



Σχήμα 4 : Πολλαπλά Stand-alone Data Marts

2.3.3 Σύγκριση των Αρχιτεκτονικών του DW/BI

Το **Corporate Information Factory (CIF)** του Inmon και η **Διαστασιακή Αποθήκη Δεδομένων** του Kimball διαθέτουν μια «*επιχειρηματική εστίαση*» που αντιμετωπίζουν τα δεδομένα. Ως αποτέλεσμα του κοινού πεδίου εστίασης τους, οι δύο εταιρικές αρχιτεκτονικές μοιράζονται ένα κοινό αρχιτεκτονικό χαρακτηριστικό: έχουν ένα ενιαίο ενσωματωμένο χώρο αποθήκευσης ατομικών δεδομένων. Η μορφή αποθήκευσης των δεδομένων στις αποθήκες του CIF είναι σε καθιερωμένη μορφή, ενώ η διαστασιακή αποθήκη του Kimball είναι, όπως υποδεικνύει και το όνομα, σε διαστασιακή μορφή. Στο CIF η πρόσβαση δεν γίνεται απευθείας αλλά μέσω Data Marts, ενώ στην αποθήκη Kimball η πρόσβαση γίνεται απευθείας αλλά και εμμέσως μέσω Data Marts, αναλόγως την σχεδίαση.

Τα Stand-alone Data Marts στερούνται «*επιχειρηματικής εστίασης*». Η ανάπτυξή τους λαμβάνει υπόψη τις απαιτήσεις μιας ομάδας ή τμήματος και όχι ολόκληρης της επιχείρησης και δεν διαθέτουν αναγκαστικά ούτε ενσωματωμένα, ούτε ατομικά δεδομένα. Ενώ όσον αφορά τον τρόπο μοντελοποίησης των Stand-alone Data Marts, συχνά ακολουθείται ο διαστασιακός σχεδιασμός.

Συγκεντρωτικά γίνεται η σύγκριση των διαφορετικών αρχιτεκτονικών στον **Πίνακα 2**. [5]

Αρχιτεκτονική	Περιγραφή	Data Marts	Format	Άμεση Πρόσβαση	Σκοπιά Εστίασης
Inmon's Corporate Information Factory	<ul style="list-style-type: none"> Κεντριοποιημένο απηθηκευτήριο ατομικών δεδομένων Τα Data Marts της οργανώνουν τα δεδομένα για τμηματική ανάλυση 	Φυσικά	3NF	ΟΧΙ	Εταιρική
Kimball's Dimensional Datawarehouse	<ul style="list-style-type: none"> Κεντριοποιημένο απηθηκευτήριο ατομικών δεδομένων Τα Data Marts δεν απαιτούν ξεχωριστή βάση δεδομένων 	Λογικά*	Διαστασιακό	NAI*	Εταιρική
Stand-alone Data Marts	<ul style="list-style-type: none"> Εφαρμογή ανά θεματική/ τμηματική περιοχή 	Φυσικά	n/a	n/a	Τμηματική

*Προαιρετικό

Πίνακας 2 : Σύγκριση των Αρχιτεκτονικών Σχεδίασης των Αποθηκών Δεδομένων

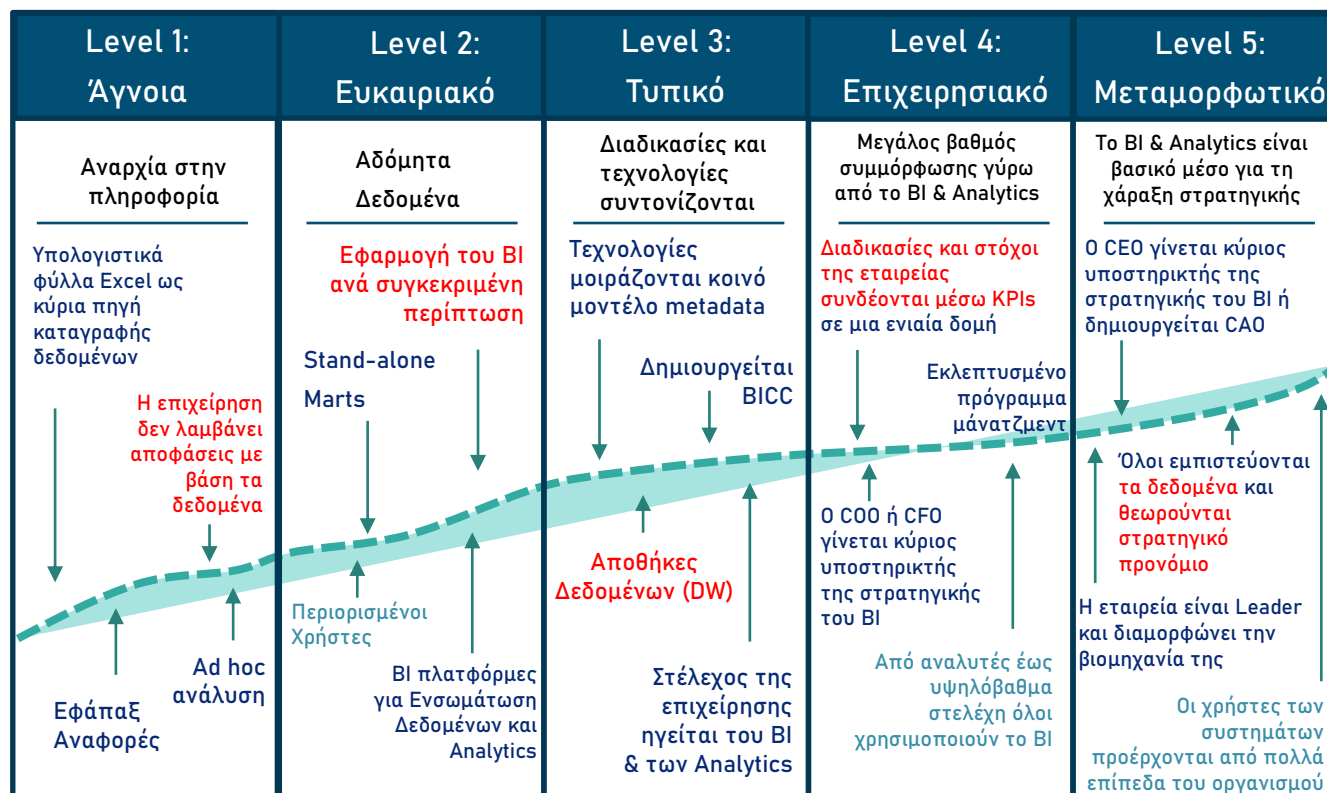
2.4 Η κουλτούρα των εταιριών γύρω από το BI & τα Analytics

Τα συστήματα αποθήκευσης της πληροφορίας σε ένα κεντριοποιημένο μέρος (DW) αποτελούν μέρος ενός *ολοκληρωμένου συστήματος BI*. Όμως, η ύπαρξη τους ή όχι εν τέλει συναρτάται ,κυρίως από το επίπεδο «ωρίμανσης» στο οποίο βρίσκεται η εταιρεία όσον αφορά την κουλτούρα της γύρω από την λήψη αποφάσεων. Λαμβάνει η εταιρεία τις αποφάσεις της με βάση δεδομένα; Εμπιστεύεται τα δεδομένα που συλλέγει; Θεωρούνται τα δεδομένα κρίσιμο στοιχείο για την χάραξη στρατηγικής ή στηρίζεται στην εμπειρία και το ένστικτο των στελεχών της; Οι απαντήσεις σε αυτά τα ερωτήματα υποδεικνύουν σε τι επίπεδο ωρίμανσης βρίσκεται η εταιρεία.

Κάποιες επιχειρήσεις επιλέγουν να αναπτύξουν στρατηγική BI μαζί με τη δημιουργία Αποθηκών Δεδομένων (DW). Άλλες αξιοποιούν υπάρχοντα συστήματα τύπου Stand-alone Mart, που διαθέτουν περιορισμένη σκοπιά σε σχέση με τον τρόπο που αξιοποιούν την πληροφορία και άρα είναι απομονωμένα ,ενώ άλλες δεν αναπτύσσουν συστήματα Data Warehousing χρησιμοποιώντας απλές «στατικές» πηγές δεδομένων ,όπως αρχεία Excel και

με βάση αυτά χαράσσουν στρατηγική BI.

Ανεξάρτητα από το πόσο «ώριμη» είναι η εταιρεία σε σχέση με την αξιοποίηση των δεδομένων της, οι εταιρείες μπορούν να αξιοποιήσουν **self-service BI tools**, όπως το **Power BI**, για να τις βοηθήσουν στην διαδικασία λήψης αποφάσεων. Οποσδήποτε όμως, τα αποτελέσματα βελτιώνονται και γίνονται πιο αξιόπιστα, όσο πιο «ώριμη» είναι μια εταιρεία. Στο **Σχήμα 5** γίνεται μια συνοπτική απεικόνιση των επιπέδων «ωριμότητας» μιας επιχείρησης και της κουλτούρα της σε σχέση με τον τρόπο που παίρνει αποφάσεις, αν είναι **data-driven decision making** ή όχι.



Σχήμα 5 : BI και το μοντέλο ωρίμανσης των Analytics στην εταιρεία

Level 1: Άγνοια

Στο πρώτο επίπεδο ωριμότητας μια εταιρείας δεν υπάρχει συγκεκριμένη στρατηγική γύρω από την αξιοποίηση συστημάτων BI. Δεδομένα συλλέγονται χωρίς κάποια είδους συστηματική προσπάθεια και δεν υπάρχει κάποια υποδομή για να το υποστηρίξει. Συνήθως είναι σε μορφή υπολογιστικών φύλλων Excel. Διευθυντές και στελέχη ζητάνε πληροφορίες και οι χρήστες κάνουν **ad-hoc ανάλυση** και querying από κάποια λειτουργικά συστήματα. Δεν υπάρχει κάποια κουλτούρα για λήψη αποφάσεων με βάση δεδομένα. Η εταιρεία για να λάβει αποφάσεις στηρίζεται κυρίως, στην εμπειρία και το ένστικτο των στελεχών της.

Level 2: Ευκαιριακό

Σε αυτό το επίπεδο, οι επιχειρηματικές μονάδες αναλαμβάνουν κάθε πρότζεκτ BI ξεχωριστά για να βελτιστοποιήσουν μια διαδικασία της εταιρείας. Κάθε διαδικασία δεν αντιμετωπίζεται ως μέρος μιας ευρύτερης δομής αλλά, ως μεμονωμένη. Αυτό αποτυπώνεται στην ασυνέπεια των δεδομένων που συλλέγονται και στο πως αποθηκεύονται. Συνήθως χρησιμοποιούνται αποθήκες δεδομένων τύπου Stand-alone Data Marts, τα οποία είναι απομονωμένα και περιέχουν δεδομένα κατάλληλα για τμηματική ανάλυση. Σταδιακά, η εταιρεία αρχίζει να χρησιμοποιεί λογισμικά και εφαρμογές BI για την ενσωμάτωση δεδομένων και ανάλυση αυτών. Σε αυτό το στάδιο, τα άτομα που ασχολούνται με το BI & τα Analytics είναι ακόμη μικρές ομάδες IT. Οι χρήστες των εφαρμογών της επιχείρησης και οι επιχειρησιακοί διευθυντές είναι ακόμη ένας μικρός αριθμός ατόμων στην επιχείρηση.

Level 3: Τυπικό

Στο τρίτο επίπεδο οι διαδικασίες, οι τεχνολογίες και οι άνθρωποι αρχίζουν να συντονίζονται σε όλη την επιχείρηση. Ένα ανώτερο στέλεχος της εταιρείας γίνεται βασικός υποστηρικτής (**sponsor**) και ηγέτης της πρωτοβουλίας του BI & των Analytics. Δημιουργείται ένα **BICC** (Business Intelligence Competence Centre), το οποίο είναι ένα κέντρο αριστείας Analytics, απαρτιζόμενο από ομάδες IT, χρήστες της επιχείρησης και αναλυτές, οι οποίοι βελτιώνουν την συνέπεια και την ενσωμάτωση των δεδομένων, των εφαρμογών BI και Data Warehousing. Ολοκληρωμένα αποθηκευτήρια δεδομένων (DW) δημιουργούνται και διαχειρίζονται από το BICC. Το πολύ 1-2 διαδικασίες της εταιρείας μοιράζονται κοινό master - μοντέλο δεδομένων και τα μοντέλα μεταδεδομένων (**metadata**) είναι κοινά για κάθε τεχνολογία. Για παράδειγμα, εργαλεία ενσωμάτωσης δεδομένων μοιράζονται ένα μοντέλο μεταδεδομένων, ενώ BI πλατφόρμες χρησιμοποιούν ένα άλλο.

Level 4: Επιχειρησιακό

Στο «Επιχειρησιακό» επίπεδο ωριμότητας, πλέον είτε ο CEO σε μικρότερη εταιρεία, είτε ο CFO είτε ο COO γίνονται υποστηρικτές και έχουν την πρωτοβουλία των BI & Analytics. Στην εταιρεία έχει δημιουργηθεί μια ενιαία δομή-δίκτυο (**framework**) με την οποία συνδέονται στόχοι της εταιρείας και διαδικασίες μέσω μετρήσιμων και άμεσα συγκρίσιμων KPIs (Key Performance Indicators). Αυτά τα KPI καθορίζουν την στρατηγική της εταιρείας. Πλέον οι εφαρμογές και πλατφόρμες BI υποστηρίζουν αποφάσεις που σχετίζονται με την συνολική λειτουργία της εταιρείας και όχι μόνο την τμηματική. Τα στελέχη μπορούν να εντοπίσουν ξεκάθαρα την σχέση μεταξύ αίτιου και αιτιατού σε κρίσιμες δραστηριότητες του οργανισμού. Πλέον, όλοι από απλούς αναλυτές, διευθυντές και ανώτερα στελέχη αξιοποιούν εφαρμογές των BI συστημάτων. Τα συστήματα διαχείρισης της πληροφορίας (EIM) και διαμοιρασμού

αυτής έχουν ωριμάσει αρκετά και λαμβάνουν μεγάλη χρηματοδότηση από τον προϋπολογισμό τη επιχείρησης. Ο οργανισμός δείχνει μεγάλη πειθαρχία στα BI & Analytics πρότζεκτ της. Κοινά μοντέλα δεδομένων και κανόνες μοντελοποίησης μεταξύ των διαδικασιών του οργανισμού επιβάλλουν μικρότερο αριθμό εκδόσεων της ίδιας πληροφορίας, επιτρέποντας πιο εκλεπτυσμένα προγράμματα μάνατζμεντ και γρήγορης ανάπτυξης τεχνολογιών και διαδικασιών.

Level 5: Μεταμορφωτικό

Στο ανώτερο στάδιο της ωριμότητας των εταιρειών, η πληροφορία και τα δεδομένα πραγματικά θεωρούνται **περιουσιακό στοιχείο** και αποτελούν κλειδί για τον ανταγωνισμό. Εμπλεκόμενοι από διάφορα επίπεδα του οργανισμού χρησιμοποιούν εφαρμογές του BI & Analytics, **εμπιστεύονται τα δεδομένα** και λαμβάνουν αποφάσεις με βάση αυτά. Τα ανώτερα στελέχη γνωρίζουν ότι τα δεδομένα μπορούν να παράξουν επιπλέον εισόδημα. Στον οργανισμό ο CEO γίνεται βασικός υποστηρικτής (**sponsor**) της BI στρατηγικής είτε δημιουργείται η θέση του CAO (Chief Analytics Officer) ή CDO (Chief Data Officer). Έχει ολοκληρωθεί η δομή η οποία αξιολογεί τις διαδικασίες της εταιρείας και μπορεί να έχει **επεκταθεί και σε εξωτερικούς συνεργάτες** και πελάτες. Επί παραδείγματι, μπορεί η εταιρεία να θέλει να αξιολογήσει την αλυσίδα τροφοδοσίας της. Πλέον όλα τα πρότζεκτ του οργανισμού χρησιμοποιούν σπάντα διαδικασίες και μοντέλα, με μικρές διορθώσεις προσαρμογής ανά περιοχή ή πρότζεκτ. Ενώ, σε χαμηλότερα επίπεδα ωρίμανσης τα συστήματα αυτά επιστρατεύονται για να μοντελοποιήσουν και να βελτιστοποιήσουν εσωτερικές διαδικασίες και μετρήσεις, σε αυτό το επίπεδο η εστίαση γίνεται στο πως προσθέτουν αξία στην εταιρεία. Γενικά, επιχειρήσεις και οργανισμοί που βρίσκονται σε τόσο μεγάλη ωριμότητα, πρωτοτυπούν και παράγουν νέες τεχνολογίες, ενώ συνήθως είναι **ηγέτες στην βιομηχανία τους** και την διαμορφώνουν. [7]

2.5 Παραδοσιακά και self-service BI συστήματα

Όπως, γίνεται φανερό από τις ενότητες που προηγήθηκαν ένα ολοκληρωμένο σύστημα BI είναι ενταγμένο σε ένα ευρύτερο οικοσύστημα εφαρμογών και συστημάτων. Σε αυτό λαμβάνουν χώρα διάφορες διαδικασίες και εμπλέκονται πολλαπλές ομάδες επαγγελματιών από διάφορα επίπεδα ενός οργανισμού, σε όλα τα στάδια ανάπτυξης του κύκλου ζωής ενός συστήματος BI.

Το *παραδοσιακό BI* είναι ο «παλιός τρόπος», χωρίς να είναι ξεπερασμένος, εφαρμογής εργαλείων ανάλυσης δεδομένων. Απαιτεί συνήθως ένα πολύπλοκο περιβάλλον πληροφορικής, χώρο για αποθήκευση δεδομένων (data warehousing) και σχεδόν συνεχή συμμετοχή από το προσωπικό πληροφορικής. Έχει χαρακτήρα τοπικό, καθώς δημιουργείται

εντός της εταιρείας (**on-premises**) αξιοποιώντας συστήματα και πόρους της ίδιας της εταιρείας, ενώ συντηρείται και αναπτύσσεται από την ίδια.

Στα παραδοσιακά συστήματα BI, αυτοί που τα οργανώνουν ανήκουν σε κάποια ομάδα επαγγελματιών από το IT. Τα άτομα αυτά διαθέτουν αρκετές δεξιότητες συμπεριλαμβανομένης της γνώσης συστημάτων επεξεργασίας και αποθήκευσης δεδομένων (**data processing & storing systems**), έννοιες επιχειρηματικής μοντελοποίησης (**business modelling**), και τουλάχιστον κάποιες βασικές δεξιότητες μαθηματικών και προγραμματισμού.

Το *self-service BI*, από την άλλη, είναι η εφαρμογή του Data Analytics με τρόπο που επιτρέπει στους χρήστες να έχουν πρόσβαση και να χρησιμοποιούν δεδομένα χωρίς να έχουν εμπειρογνωμοσύνη σε χρήση στατιστικής, ανάλυσης και διαχείρισης δεδομένων. Αυτή η προσέγγιση βασίζεται σε εργαλεία BI που *επιτρέπουν στους χρήστες να φιλτράρουν, να ταξινομούν, να αναλύουν και να οπτικοποιούν δεδομένα για την εξαγωγή πληροφοριών χωρίς την εξάρτηση από προγραμματιστές ή ειδικούς δεδομένων ή πληροφορικής.* [6]

Η πρόοδος της τεχνολογίας στον τομέα της **μηχανικής μάθησης** και της **τεχνητής νοημοσύνης** απλοποιεί και συγκεντρώνει διαδικασίες προετοιμασίας δεδομένων, μοντελοποίησης δεδομένων και δημιουργίας πληροφοριών. Ο προγραμματισμός σε **φυσική γλώσσα** παρέχει στους χρήστες μια εμπειρία που μοιάζει με αναζήτηση, αντί για μια διαδικασία «δημιουργίας ερωτημάτων» (query-building) ,όπως στην περίπτωση των παραδοσιακών συστημάτων. Τέλος, οι καινοτομίες στην τεχνολογία **cloud** επιτρέπουν την κατηγοριοποίηση και την επεξεργασία δεδομένων σε όγκους που δεν ήταν ποτέ δυνατοί στο παρελθόν.

Ένα ολοκληρωμένο παραδοσιακό **on-premises** σύστημα απαιτεί πολλαπλά συστήματα και ομάδες επαγγελματιών για να πραγματοποιήσει τις εργασίες του συστήματος. Οι σύγχρονες *self-service* πλατφόρμες συγκεντρώνουν και απλοποιούν αυτές τις διαδικασίες ,όπου ακολουθούν στο **Σχήμα 6**.



Σχήμα 6 : Συνοπτική πορεία εργασιών σε σύστημα BI

Αξίζει να σημειωθεί ότι, όπως φανερώνεται και από την ενότητα **2.4 «Η κουλτούρα των εταιριών γύρω από το BI & τα Analytics»**, ανεξαρτήτως του επίπεδου ωρίμανσης ή της κουλτούρας που έχει η εταιρεία για να λαμβάνει αποφάσεις, εργαλεία και πλατφόρμες BI μπορούν να εφαρμοστούν και να έχουν σημαντικό αντίκτυπο στην απόδοσή της, χωρίς βέβαια να παρέχουν την ίδια ποιότητα αποτελεσμάτων και ούτε την ίδια ευκινησία (robustness) σε ερωτήματα που τίθενται από τους χρήστες, συγκριτικά με έναν ώριμο οργανισμό.

	Παραδοσιακό BI (on-premises)	Self-service BI
Χρόνος Εγκατάστασης	Μεγάλος	Σημαντικά μικρότερος
Κόστος Εγκατάστασης	Μεγάλο	Σημαντικά μικρότερο
Ανάμειξη ατόμων από το τμήμα του IT	Κατά την εγκατάσταση ,συντήρηση & ανάπτυξη του συστήματος	Μετά την εγκατάσταση τα άτομα του IT έχουν μικρή ανάμειξη
Προσαρμοστικότητα	Συνήθως μικρή Εξαρτάται από το επίπεδο ωριμότητας της εταιρείας	Αυξημένη λόγω του πλήθους των αναλυτών
Πρόσβαση σε δεδομένα	Συνήθως τα άτομα που έρχονται σε επαφή με τα δεδομένα είναι κυρίως από το IT	Ο αριθμός των αναλυτών και των χρηστών είναι αυξημένος, καθώς οι διαδικασίες είναι απλοποιημένες
Χρόνος εκτέλεσης	Συνήθως μεγάλος, λόγω του ότι οι αναλύσεις αποκλειστικά προέρχονται από το IT	Σημαντικά μικρότερος
Διαχείριση των δεδομένων	Οι διαδικασίες σύνδεσης, ενσωμάτωσης και επεξεργασίας γίνονται από το IT ,όποτε είναι ξεκάθαρες και οικουμενικές	Κίνδυνος για μεγάλη αταξία στον τρόπο που συλλέγεται και διαμορφώνεται η πληροφορία
Ασφάλεια στα Δεδομένα	Μεγαλύτερη	Ενδεχομένως μειωμένη

Πίνακας 3 : Σύγκριση παραδοσιακού και self-service BI [6]

2.6 Self-service BI λογισμικά της αγοράς

Στην αγορά μετά το 2010 έχουν εμφανιστεί διάφορες πλατφόρμες και εργαλεία BI , τόσο από μεγάλες πολυεθνικές εταιρείες όσο και μικρότερες. Η πρόοδος της μηχανικής μάθησης και της τεχνικής νοημοσύνης καθιστούν τα self-service εργαλεία ,σε σχέση με τα παραδοσιακά εργαλεία που είναι βασισμένα σε ομάδες του IT ,όλο και πιο προσιτά στις επιχειρήσεις που θέλουν να αναμορφώσουν τον τρόπο λήψης αποφάσεων τους.






















Πρωταγωνιστές ,σαφώς, στην αγορά (2022) είναι η **Microsoft** (Power BI) , η **Salesforce** (Tableau), **SAP** (Business Objects) ,**Qlik** (Qlik Sense), **Google** (Google Analytics), **Oracle** (Oracle Analytics Cloud) και **SAS**.

Η συμβουλευτική εταιρεία Gartner, όπου συγκρίνει προϊόντα και υπηρεσίες τεχνολογίας, στο γνωστό της “Μαγικό” τεταρτοκύκλιο (**Magic Quadrant for Analytics and Business Intelligence Platforms**) ,κατατάσσει τις εταιρείες στο πεδίο των Analytics και BI πλατφορμών σύμφωνα με την ικανότητα τους να εκτελούν τις διάφορες απαιτούμενες εργασίες αλλά και το συνολικό όραμα τους ,όπως απεικονίζεται στην **Εικόνα 2**. [6]



Εικόνα 2 : Το μαγικό τεταρτοκύκλιο του Gartner

Τρεις εταιρείες ανήκουν στο τεταρτημόριο *Leaders* : η Microsoft ,η Salesforce και η Qlik. Στον **Πίνακα 4** γίνεται μια εποπτική ανάλυση και σύγκριση των τριών κορυφαίων προϊόντων στον χώρο του BI & Analytics.

	Power BI - Microsoft	Tableau - Salesforce	Qlik Sense - Qlik
Περιβάλλον Ανάπτυξης	Desktop & Cloud	Desktop & Cloud	Cloud /Online
Επεξεργασία Δεδομένων			
Μοντελοποίηση Δεδομένων			
Προτιμώμενο Μοντέλο Δεδομένων	Διαστατικός Σχεδιασμός	Πίνακες Flat	Τύπου Snowflake
Υποστηρίζουν R & Python			
Ενσωματωμένα AI - Analytics			
Εφαρμογή στο Κινητό			
Πρόσβαση σε Μοντέλο Δεδομένων τρίτων χρηστών			
Πρόσβαση σε open-source Οπτικοποιήσεις			
Κόστος	Χαμηλότερο στην αγορά	Συγκριτικά Αυξημένο	Συγκριτικά Αυξημένο

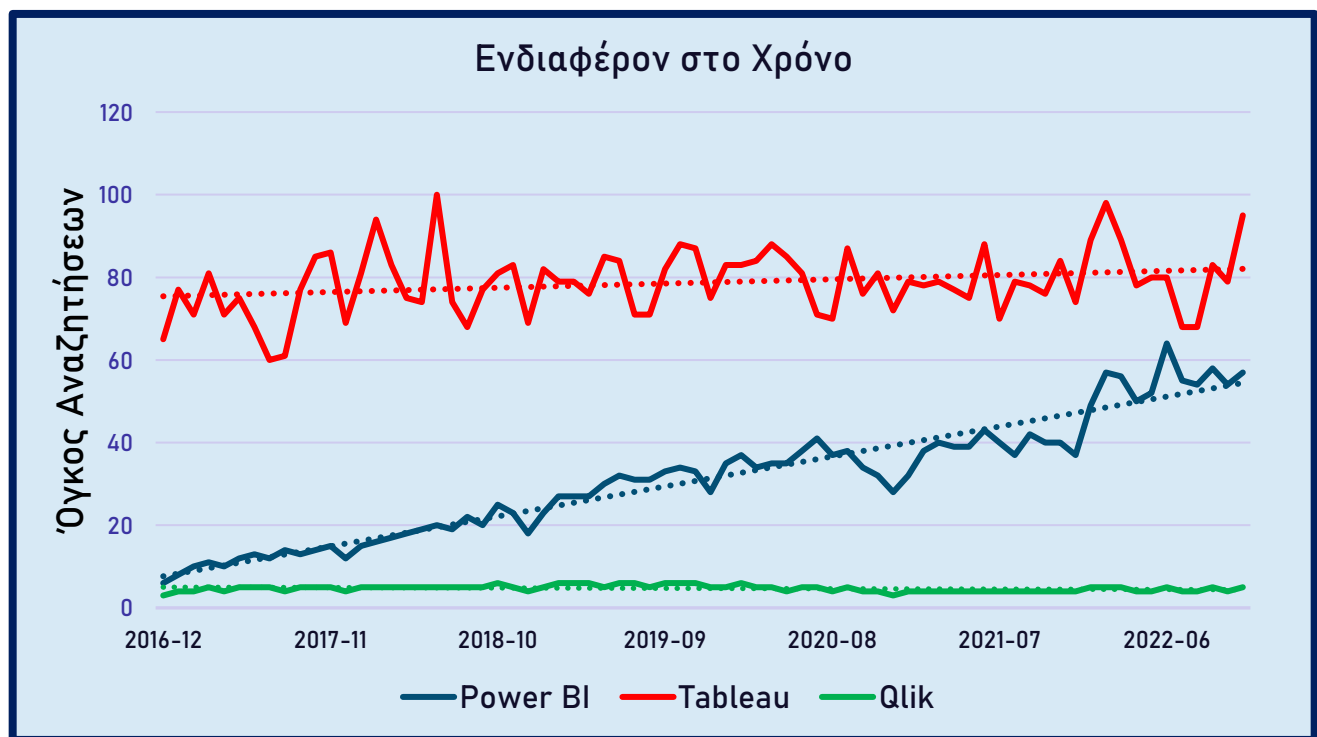
Πίνακας 4 : Οι Leaders στον χώρο των self-service BI εργαλείων [8]

Στην αγορά οι περισσότεροι επαγγελματίες συμφωνούν ότι κυρίαρχος στον χώρο της οπτικοποίησης είναι το **Tableau**. Το σημείο στο οποίο φαίνεται ότι υπερτερεί το **Power BI**,

είναι στο χώρο της προετοιμασίας και επεξεργασίας των δεδομένων, καθώς και στα εργαλεία που παρέχει για την κατασκευή του Μοντέλου Δεδομένων. Οι περισσότερες πλατφόρμες δεν διαθέτουν ενσωματωμένα εργαλεία για τις αναφερθείσες διαδικασίες και έτσι χρειάζονται ξεχωριστά εργαλεία με επιπλέον κόστος.

Το Power BI, έχει έναν διαρκώς αυξανόμενο όγκο αναζητήσεων παγκοσμίως και έχει φτάσει αυτόν του Tableau ,όπως φαίνεται από το **Google Trends**. Αυτό σημαίνει ότι διαθέτει μια τεράστια κοινότητα μελών για την αλληλοϋποστήριξη τους.

Τέλος, ένα στοιχείο το οποίο ενισχύει το Power BI της Microsoft είναι το κόστος, καθώς συγκριτικά με τα υπόλοιπα πακέτα είναι πιο οικονομικό τόσο για εξατομικευμένους χρήστες όσο και για μεγαλύτερους οργανισμούς, όπου το κόστος μειώνεται ακόμα περισσότερο ανά χρήστη.



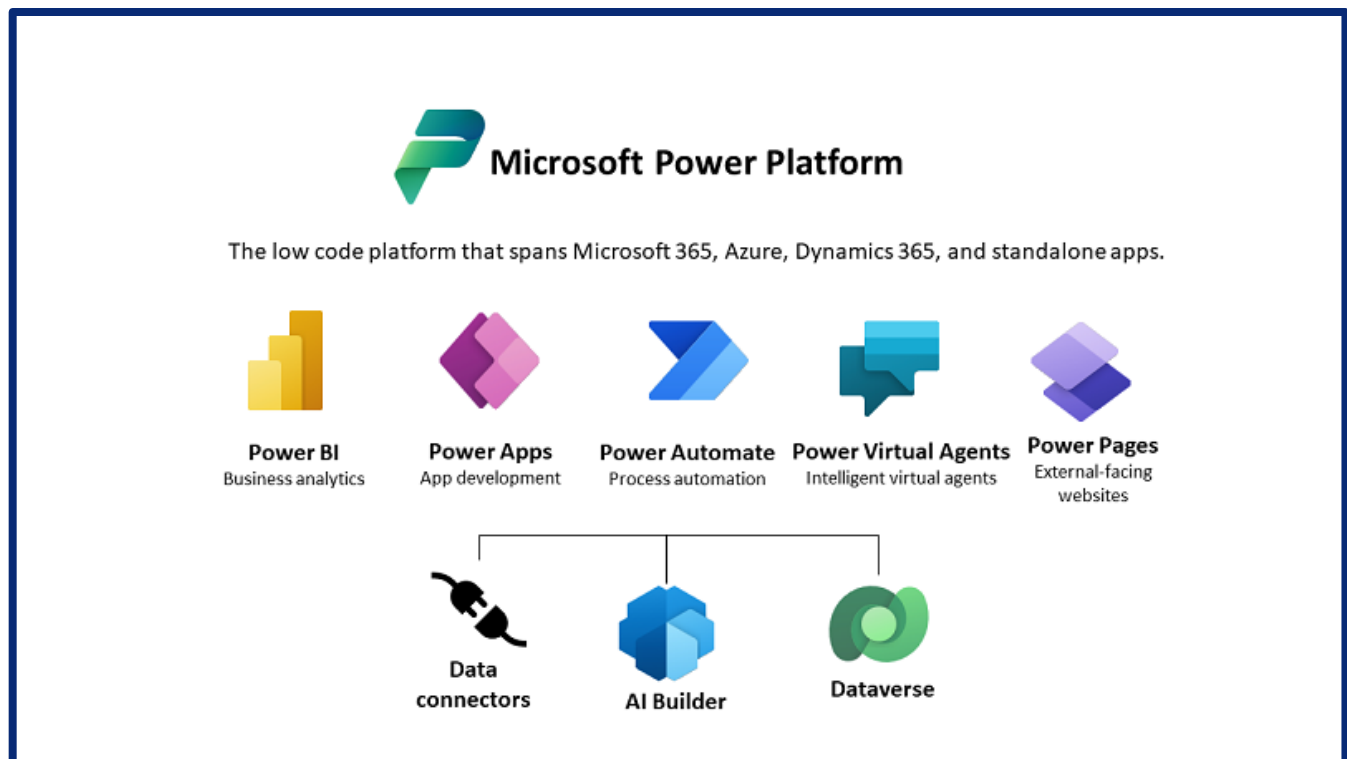
Διάγραμμα 2 : Παγκόσμιος όγκος αναζητήσεων των self-service εργαλείων σύμφωνα με το Google Trends

3. Η πλατφόρμα Επιχειρηματικής Ευφυΐας της Microsoft Power BI

3.1 Το οικοσύστημα του Power BI

Το αναλυτικό εργαλείο **Power BI** ανήκει σε μια συλλογή υπηρεσιών της **Microsoft** που ονομάζεται **Power Platform**. Η πλατφόρμα απευθύνεται κυρίως σε **citizen developers**, δηλαδή εργαζόμενους οι οποίοι παράγουν εφαρμογές για τους ίδιους ή τις επιχειρήσεις που ανήκουν και δεν είναι προγραμματιστές κατά βάση .

Η πλατφόρμα συνδυάζει ανάπτυξη εφαρμογών **κώδικα χαμηλού επιπέδου**, **αυτοματοποίηση ροής εργασιών**, ανάπτυξη bot AI και ανάλυση δεδομένων με ευρεία συνδεσιμότητα μέσω του Microsoft Dataverse - όλα σχεδιασμένα για να λειτουργούν με τις υπηρεσίες cloud της Microsoft των Azure, Dynamics 365 και Microsoft 365. Παρακάτω, στην **Εικόνα 3** φαίνεται το οικοσύστημα της πλατφόρμας. [9]



Εικόνα 3 : Τα μέρη της Power Platform

Το **Power BI** είναι μια συλλογή υπηρεσιών λογισμικού (**software services**), εφαρμογών (**apps**) και συνδέσεων (**connectors**) που συνεργάζονται για να μετατρέψουν τις μη σχετιζόμενες πηγές δεδομένων μας σε συνεκτικές, οπτικά πλούσιες, επεξηγηματικές και διαδραστικές απαντήσεις για κρίσιμα επιχειρησιακά ερωτήματα. Τα δεδομένα μας μπορεί να είναι από ένα απλό υπολογιστικό φύλλο **Excel**, έως και μια συλλογή υβριδικών αποθηκών δεδομένων (**hybrid data warehouses**).

Οι υβριδικές αυτές αποθήκες δεδομένων απαρτίζονται από **cloud-based** αποθήκες και από εσωτερικές - τοπικές αποθήκες (**on premises**). Το Power BI μας επιτρέπει να συνδεόμαστε εύκολα με τις πηγές δεδομένων μας, να οπτικοποιήσουμε και να ανακαλύψουμε, τι είναι σημαντικό και να το μοιραστούμε με οποιονδήποτε επιθυμούμε.

Συνήθως, όταν γίνεται η πρώτη γνωριμία με το **Power BI**, οι περισσότεροι χρήστες που δεν έχουν προηγούμενη επαφή με self-service λογισμικά για Business Analysis & Analytics και θεωρούν ότι πρόκειται για ένα εργαλείο οπτικοποίησης. Βέβαια, αυτό απέχει από την πραγματικότητα, καθώς αν και κρίσιμο μέρος του εργαλείου, η οπτικοποίηση αποτελεί το 25% των δυνατοτήτων του λογισμικού.

Αρχικά και κυριότερα, το **Power BI** αποτελεί ένα εργαλείο για την κατασκευή του μοντέλου δεδομένων μας. Το μοντέλο δεδομένων αποτελείται από τουλάχιστον έναν πίνακα δεδομένων αλλά, συνήθως αποτελείται από πολλαπλούς πίνακες. Αυτοί οι πίνακες μπορούν είτε να υπάρχουν ανεξάρτητα ο ένας από τον άλλο, ή μπορούμε να δημιουργήσουμε συνδέσεις μεταξύ τους χρησιμοποιώντας δεδομένα που μοιράζονται. Οι συνδέσεις μεταξύ πινάκων στο μοντέλο δεδομένων μας ονομάζονται σχέσεις (**relationships**).

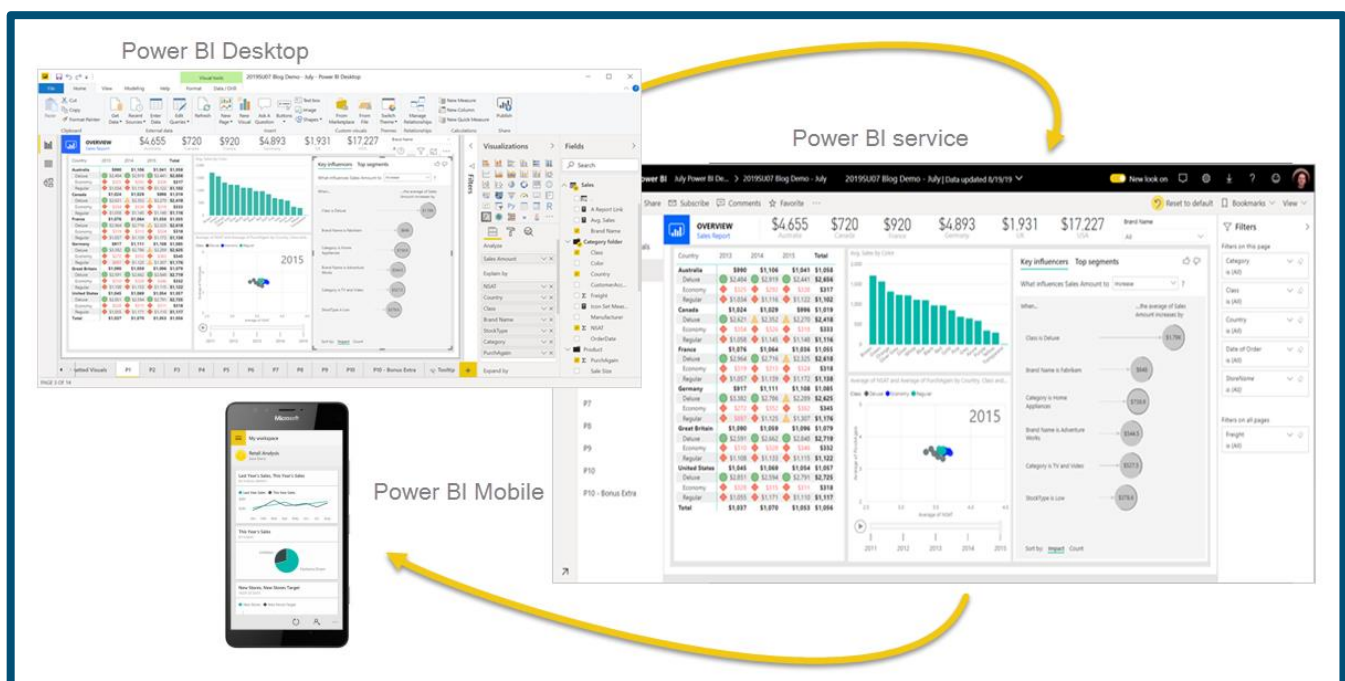
Η παραδοσιακή μοντελοποίηση BI απαιτεί πολύ χρόνο, καθώς χρειάζεται η γνώση κατασκευής και διαχείρισης ενός πολύπλοκου συστήματος BI με Αποθήκες Δεδομένων, όπως του **Ralph Kimball** (Dimensional Data Warehouse) είτε του **W.H. Inmon** (Corporate Information Factory). Τέτοια συστήματα απαιτούν μεγάλο αριθμό ατόμων και ομάδων από το IT σε ρόλους, όπως: Διαχειριστής Βάσης Δεδομένων, Προγραμματιστής για BI εφαρμογές, Αρχιτέκτονας Δεδομένων, Διαχειριστής Μεταδεδομένων και αρκετές άλλες κατηγορίες.

Τα self-service BI εργαλεία συντομεύουν αυτές τις διαδικασίες άρα και τον χρόνο παράδοσης. Ένας από τους λόγους που το Power BI είναι τόσο δημοφιλές και εύκολο στη χρήση είναι ότι συντομεύει τη διαδικασία μοντελοποίησης δεδομένων και δίνει αυτές τις δυνατότητες στα χέρια των ανθρώπων που δεν είναι προγραμματιστές λογισμικού (**citizen developers**). Με τη σωστή εκπαίδευση στο Power BI, οι αναλυτές δεν χρειάζεται να αναμένουν από τους προγραμματιστές να παραδώσουν αξιόπιστες BI λύσεις, αλλά τις δημιουργούν οι ίδιοι. Έτσι, οι ανθρώπινοι πόροι της εταιρείας αξιοποιούνται πιο ισορροπημένα. Οι χρήστες της επιχείρησης πλέον, μπορούν να εισάγουν τα δεδομένα, να κάνουν τους μετασχηματισμούς και να δημιουργούν μόνοι τους ένα ισχυρό μοντέλο δεδομένων.

3.1.1 Συνιστώσες του Power BI

Το Power BI αποτελείται από διάφορα μέρη, τα οποία είτε λειτουργούν συνεργατικά είτε και ανεξάρτητα. Τα βασικά μέρη που απαρτίζεται το **Power BI** είναι τα εξής:

- Μια εφαρμογή για Windows λογισμικό ,που ονομάζεται **Power BI Desktop** .
- Μια online υπηρεσία τύπου **SaaS** (Software as a Service) ,με όνομα **Power BI Service**.
- Μια εφαρμογή για κινητές συσκευές ,**Power BI Mobile**, κατάλληλη για Windows , IOS και Android λογισμικά.



Εικόνα 4 : Οι συνιστώσες του Power BI

Πέρα από αυτά τα βασικά μέρη υπάρχουν και δύο ακόμα:

- **Power BI Report Server:** Λειτουργεί ως ένας τοπικός (on-premises) διακομιστής για την δημοσίευση αναφορών, και όχι cloud-based διακομιστής, όπως στην περίπτωση του Power BI Service. Βέβαια, αυτός ο διακομιστής έχει περιορισμένη λειτουργικότητα και υποστήριξη από την Microsoft, αλλά σε κάποιους οργανισμούς είναι χρήσιμος.

- **Power BI Report Builder:** Πρόκειται για ένα λογισμικό, το οποίο εξυπηρετεί στην δημιουργία εκτυπώσιμων αναφορών, έτοιμες για να δημοσιευτούν στο Power BI Service, είτε να χρησιμοποιηθούν όπως, οι συνηθισμένες αναφορές. [9]

3.2 Power BI Desktop

Το **Power BI Desktop** είναι η βασική εφαρμογή του οικοσυστήματος του **Power BI** και παρέχει εξαιρετικές δυνατότητες όπως:

- Συλλογή των δεδομένων από τα διάφορα πηγαία συστήματα.
- Ενσωμάτωση των δεδομένων σε ενιαία και εύκολα διαχειρίσιμη μορφή.
- Καθαρισμός και η Επεξεργασία/Προετοιμασία των δεδομένων.
- Δημιουργία του πινακοποιημένου Μοντέλου Δεδομένων.
- Καθορισμός των σχέσεων μεταξύ των πινάκων στο Μοντέλο Δεδομένων.
- Δημιουργία Αναφορών (**Reports**) και Πινάκων (**Dashboards**).
- Δημοσίευση των Αναφορών και Πινάκων.
- Ανανέωση των Δεδομένων.

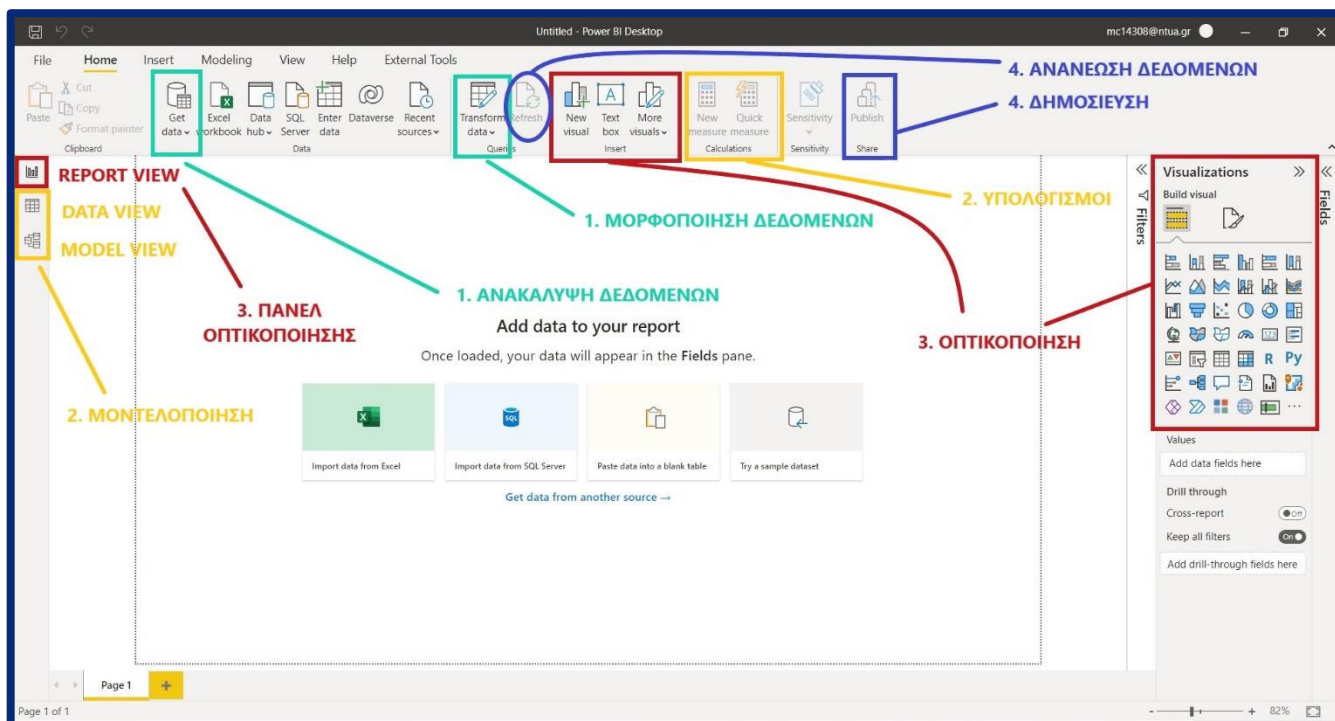
Οι κύριες εργασίες σε ένα πρότζεκτ BI μπορούν να χωρισθούν σε τέσσερις μεγάλες κατηγορίες. Συγχρόνως, κάθε κατηγορία αποτελεί και ένα βήμα στην ροή εργασιών, όπου κάθε βήμα είναι απαραίτητο για να προχωρήσουμε στο επόμενο. Επί παραδείγματι, για να γίνει η δημιουργία του Μοντέλου Δεδομένων θα πρέπει νωρίτερα να έχει γίνει το βήμα «Προετοιμασίας των Δεδομένων».

Στο **Σχήμα 7** οργανώνονται οι εργασίες αυτές σε 4 επίπεδα. Ασφαλώς, πάνω στην πορεία εργασιών μπορεί να πρέπει να αναθεωρηθούν ορισμένα βήματα ,είτε προσθέτοντας δεδομένα, είτε αλλάζοντας το μοντέλο είτε κάποια οπτικοποίηση. Έτσι, η **πορεία εργασιών δεν είναι απαραίτητα γραμμική αλλά κυκλική**, όμως αξίζει να γίνει εμφανής η τυπική πορεία εργασιών.

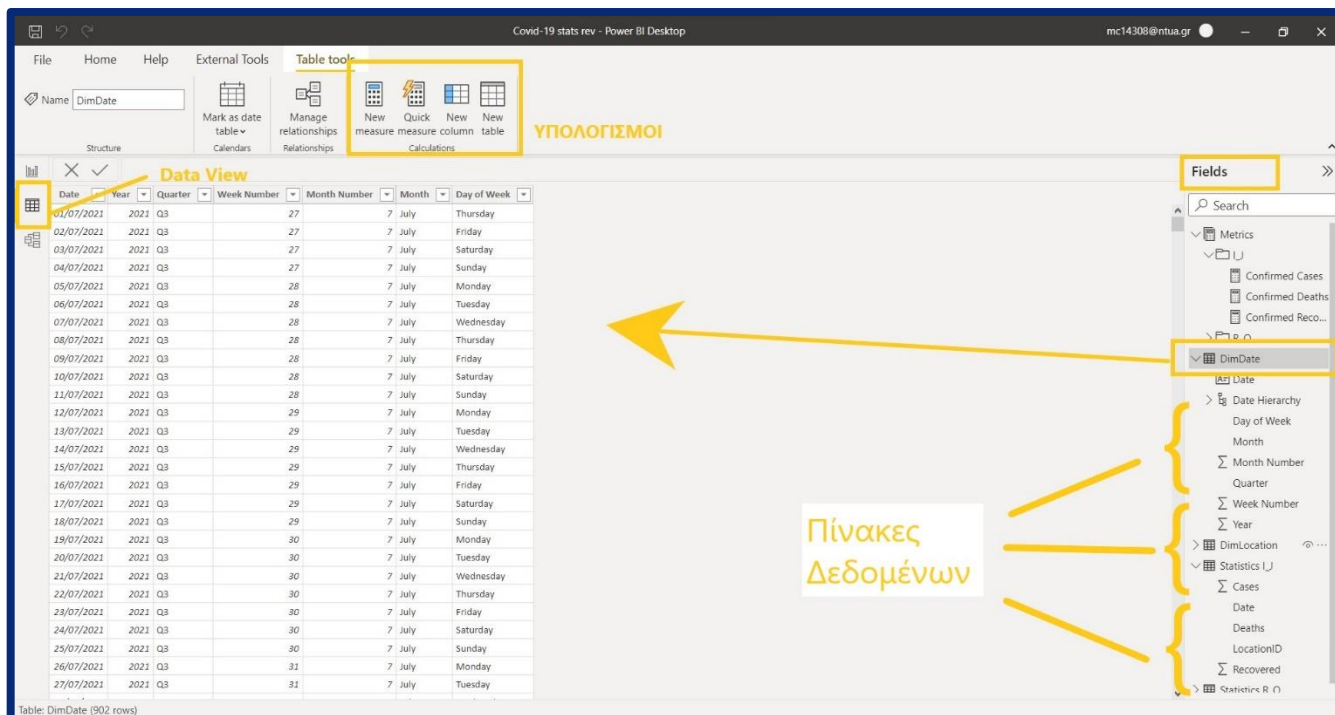


Σχήμα 7 : Οι φάσεις ενός BI project στο Power BI Desktop

Ας γνωρίσουμε το γραφικό περιβάλλον της εφαρμογής. Ανοίγοντας την εφαρμογή οδηγούμαστε κατευθείαν στο Πάνελ Οπτικοποίησης (**Report View**), η οποία είναι η καρτέλα στην οποία θα εμφανίζονται όλες οι Οπτικοποιήσεις (**Visualizations**). Η επόμενη καρτέλα ονομάζεται καρτέλα Δεδομένων (**Data View**), όπου εκεί εμφανίζονται όλα μας τα δεδομένα σε μορφή Πινάκων.



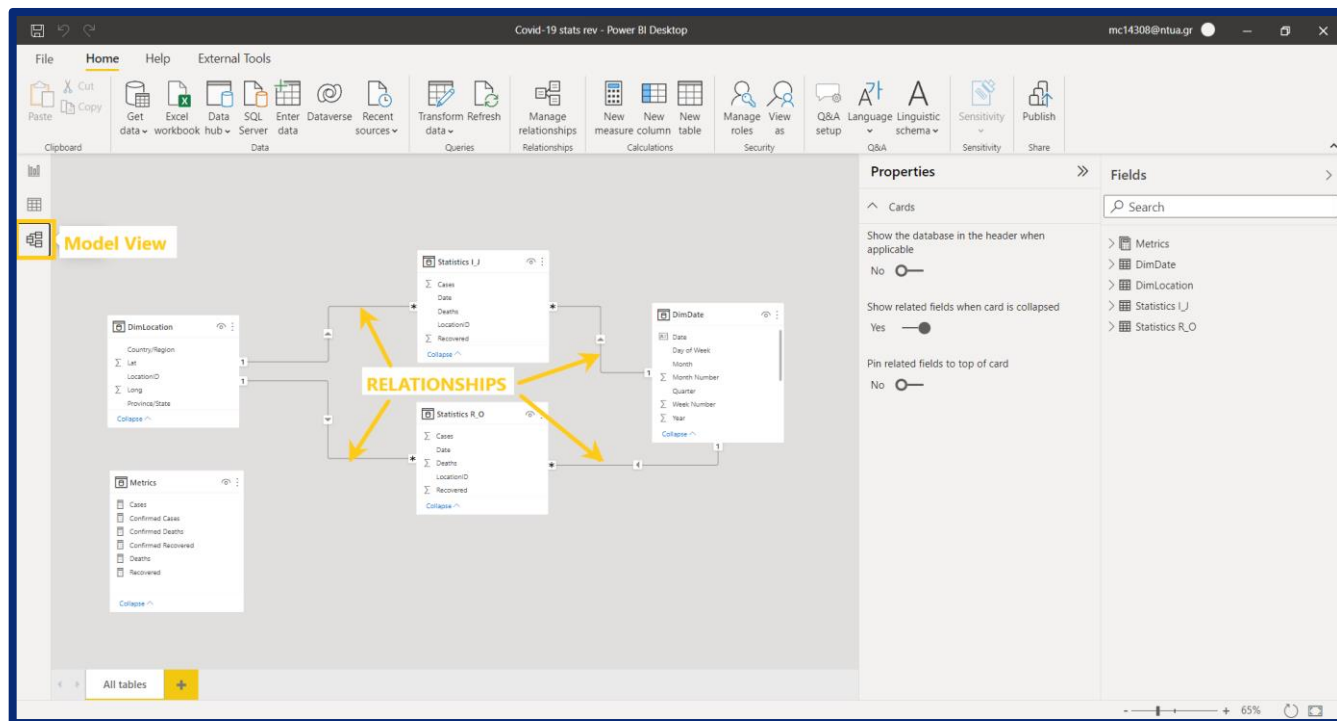
Εικόνα 5 : Το γραφικό περιβάλλον του Power BI Desktop – Report View



Εικόνα 6 : Το γραφικό περιβάλλον του Power BI Desktop – Data View

Η τρίτη καρτέλα ονομάζεται Μοντέλο Δεδομένων (**Model View**), όπου εκεί απεικονίζεται το μοντέλο δεδομένων μας, δηλαδή η πινακοποιημένη μορφή των δεδομένων μας, οι σχέσεις

(relationships) μεταξύ αυτών, καθώς και υπολογισμοί που προκύπτουν από την σύνοψη στηλών και αποθηκεύονται ως νέες στήλες στους πίνακες (measures).



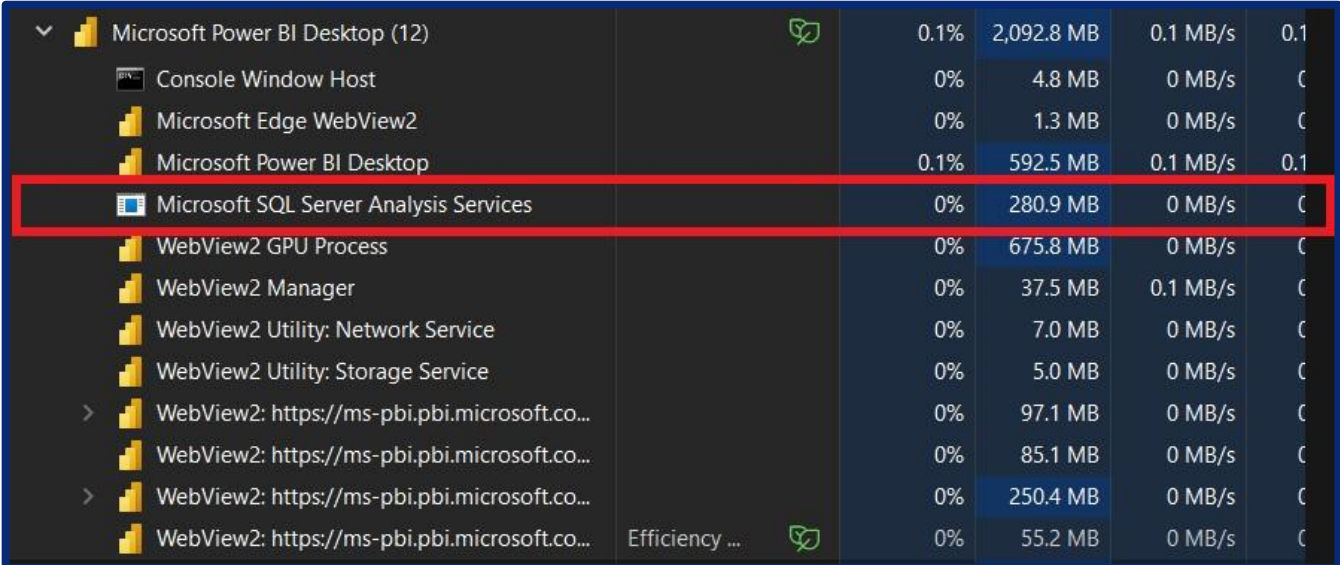
Εικόνα 7 : Το γραφικό περιβάλλον του Power BI Desktop – Model View

3.3 Power BI Desktop - Ανακάλυψη Δεδομένων

3.3.1 Tabular processing [10]

Το Power BI αξιοποιεί την αναλυτική μηχανή της Microsoft, **SQL Server Analysis Services (SSAS) Tabular**, κάνοντας χρήση μοντέλων **Tabular** (πινακοποιημένη μορφή), τα οποία μοντέλα είναι βάσεις δεδομένων κατάλληλα σχεδιασμένες για σκοπούς του BI. Τα Tabular μοντέλα παρέχουν ένα σχεδιαστικό περιβάλλον για την δημιουργία σημασιολογικών αντικείμενων ενός μοντέλου (**semantic model objects**), όπως πίνακες, σχέσεις, ιεραρχίες, μέτρα (measures) και KPIs.

Το **SSAS είναι εγκατεστημένο στο τοπικό σύστημα** και αξιοποιείται από το Power BI Desktop, αλλά και από το Power BI service, με δύο διαφορετικούς τρόπους όταν προσπαθούμε να θέσουμε ερώτημα (query) στην πηγή δεδομένων μας. Η διαφοροποίηση έγκειται στο αν τα δεδομένα αποθηκεύονται στην RAM του τοπικού συστήματος ή όχι. Εξετάζουμε και τις δύο διαφορετικές πορείες εργασιών παρακάτω. Το SSAS τρέχει παρασκηνιακά όταν χρησιμοποιούμε το Power BI Desktop, όπως φαίνεται και από το Task Manager στην **Εικόνα 8**.



Process Name	CPU	Private Bytes	Working Set	Private Bytes (KB)	Private Bytes (MB)
Microsoft Power BI Desktop (12)	0.1%	2,092.8 MB	0.1 MB/s	0.1	0.1
Console Window Host	0%	4.8 MB	0 MB/s	0	0
Microsoft Edge WebView2	0%	1.3 MB	0 MB/s	0	0
Microsoft Power BI Desktop	0.1%	592.5 MB	0.1 MB/s	0.1	0.1
Microsoft SQL Server Analysis Services	0%	280.9 MB	0 MB/s	0	0
WebView2 GPU Process	0%	675.8 MB	0 MB/s	0	0
WebView2 Manager	0%	37.5 MB	0.1 MB/s	0	0
WebView2 Utility: Network Service	0%	7.0 MB	0 MB/s	0	0
WebView2 Utility: Storage Service	0%	5.0 MB	0 MB/s	0	0
WebView2: https://ms-pbi.pbi.microsoft.co...	0%	97.1 MB	0 MB/s	0	0
WebView2: https://ms-pbi.pbi.microsoft.co...	0%	85.1 MB	0 MB/s	0	0
WebView2: https://ms-pbi.pbi.microsoft.co...	0%	250.4 MB	0 MB/s	0	0
WebView2: https://ms-pbi.pbi.microsoft.co...	0%	55.2 MB	0 MB/s	0	0

Εικόνα 8 : Παρασκηνιακή λειτουργία του Microsoft SSAS στο Task Manager

Αρχικά δημιουργείται ένα ερώτημα (query) στο Power BI γραμμένο στην γλώσσα που χρησιμοποιεί η εφαρμογή, την **DAX (Data Analysis Expressions)**. Η σύνταξη κώδικα στην

γλώσσα DAX μοιάζει με την σύνταξη της γλώσσας του Excel ,μάλιστα έχουν πολλές κοινές συναρτήσεις. Όμως, η σύνταξη στο Excel είναι σχεδιασμένη να ασχολείται με υπολογιστικά κελιά, ενώ η DAX είναι βελτιστοποιημένη στο να δουλεύει με πίνακες και στήλες του μοντέλου δεδομένων. Η DAX παρέχει διάφορους τρόπους για να συνοψίζει και να αναλύει δεδομένα και είναι εξαιρετικά γρήγορη στους υπολογισμούς.

Στη συνέχεια το ερώτημα τίθεται στο SSAS για αναλυθεί και να δοκιμαστεί. Το SSAS για να πραγματοποιήσει απαιτούμενους υπολογισμούς στέλνει το ερώτημα στην μηχανή σύνταξης (**Formulation Engine – FE**) και αυτή με τη σειρά της θέτει το ερώτημα στην μηχανή αποθήκευσης (**Storage Engine – SE**).

Η μηχανή αποθήκευσης **SE** είναι υπεύθυνη για την ανάκτηση πληροφοριών από την βάση δεδομένων ή πιο γενικά την πηγή δεδομένων, ενώ η μηχανή σύνταξης **FE** χρησιμοποιεί τα αποτελέσματα της SE για να κάνει ακόμη πιο πολύπλοκους υπολογισμούς ,όπως το σορτάρισμα (sorting) των στηλών και η επιλογή συγκεκριμένων τιμών μίας στήλης. Το σημείο αυτό είναι εκεί που ξεκινάει η διαφοροποίηση στην πορεία εργασιών, αναλόγως αν επιθυμούμε να γίνεται η αξιοποίηση του τοπικού εσωτερικού συστήματος ή αν χρησιμοποιείται εξωτερική βάση δεδομένων από αυτή του συστήματος. [10]

In-memory model (Vertipaq)

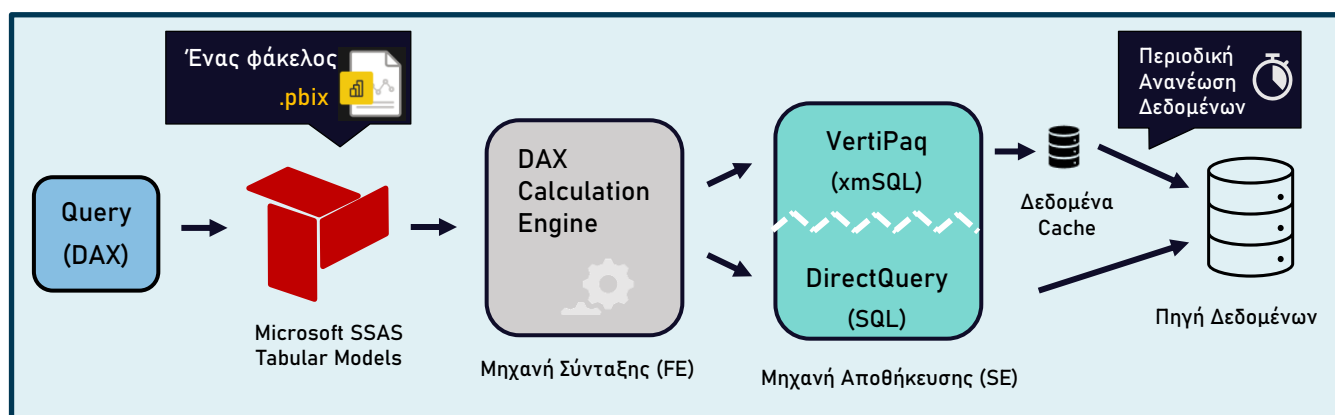
Στην πρώτη περίπτωση, το SSAS Tabular αποθηκεύει ένα αντίγραφο των δεδομένων σε εσωτερική βάση δεδομένων σε μορφή στήλης (**in-memory columnar database**) η οποία ονομάζεται **Vertipaq**. Εδώ τον ρόλο της **SE** έχει η Vertipaq. Στο μοντέλο αυτό δημιουργείται αντίγραφο των δεδομένων στην κρυφή μνήμη (**cached data**). Η Vertipaq χρησιμοποιεί υπερσύγχρονους αλγόριθμους συμπίεσης και επεξεργαστή ερωτημάτων (querying) πολλαπλών νημάτων, και παρέχει γρήγορη πρόσβαση σε αντικείμενα και δεδομένα μοντέλων που είναι σε πινακοειδή μορφή. Ουσιαστικά τα ερωτήματα στην βάση (queries) τίθενται στο αντίγραφο των δεδομένων στην εσωτερική μνήμη, που έχει αποθηκευτεί μέχρι εκείνη την στιγμή και μπορεί να ανανεωθεί αν αλλάξουν τα δεδομένα. Η γλώσσα που χρησιμοποιεί στα queries της SE ονομάζεται xMSQL, και έχει την μορφή ανθρώπινης γλώσσας.

Υπάρχουν δύο τυπικές πηγές δεδομένων :

- Μια αποθήκη δεδομένων που περιέχει όλα τα δεδομένα που σχετίζονται με την εταιρεία, τα οποία είναι έτοιμα για ανάλυση. Σε αυτή την περίπτωση έχουμε μόνο μια πηγή δεδομένων η οποία είναι σχεσιακή βάση δεδομένων.
- Η δεύτερη περίπτωση είναι όταν πολλαπλές πηγές δεδομένων τροφοδοτούν το SSAS , πιθανόν με δεδομένα διαφόρων τύπων .Το SSAS ενσωματώνει τα δεδομένα αυτά σε μία ενιαία μορφή. Εδώ συναντάμε σχεσιακές βάσεις δεδομένων και άλλες πηγές δεδομένων όπως αρχεία text, αρχεία Excel και άλλες πηγές.

DirectQuery model

Στην δεύτερη περίπτωση, δεν γίνεται εισαγωγή των δεδομένων σε εσωτερική μνήμη αλλά στέλνονται τα δεδομένα εξωτερικά του τοπικού συστήματος. Αυτό γίνεται μέσω της τεχνολογίας του **DirectQuery** που εισήχθη από το SSAS Tabular models το 2012. Το DirectQuery μετατρέπει το μοντέλο Tabular σε μια λωρίδα μεταδεδομένων πάνω από μια εξωτερική βάση δεδομένων, όπως ένας SQL server. Στην πραγματικότητα, χρησιμοποιώντας το DirectQuery, ο μηχανισμός Tabular δεν επεξεργάζεται πλέον δεδομένα στη δική του εσωτερική βάση δεδομένων. Αντίθετα, το ερώτημα μετατρέπεται σε ένα σύνολο ερωτημάτων SQL που αποστέλλονται στην υποκείμενη σχεσιακή βάση δεδομένων. Μόλις η βάση δεδομένων SQL επιστρέψει τα ζητούμενα δεδομένα, η μηχανή Tabular επεξεργάζεται περαιτέρω αυτά τα δεδομένα. Εδώ η μηχανή αποθήκευσης **SE** είναι το DirectQuery. Το ότι δεν αποθηκεύονται δεδομένα στο τοπικό σύστημα επιτρέπει να δημιουργείτε μοντέλα δεδομένων σε πραγματικό χρόνο (**real-time data**), επειδή δεν χρειάζεται πλέον η τακτική ενημέρωση των εσωτερικών δομών δεδομένων. Έτσι, τα δεδομένα είναι πάντα σε πραγματικό χρόνο, δεν υπάρχει χρόνος επεξεργασίας, δεν υπάρχουν περιορισμοί στη μνήμη και τα δεδομένα δεν χρειάζεται να μετακινηθούν από τη βάση δεδομένων πηγής στη βάση δεδομένων SSAS. Από την άλλη, χάνεται η τεράστια ταχύτητα της μηχανής VertiPaq. Επιπλέον, **δεν είναι δυνατόν να ενσωματωθούν δεδομένα που προέρχονται από διαφορετικές πηγές δεδομένων** και η ταχύτητα επεξεργασίας των ερωτημάτων εξαρτάται από την υπολογιστική δύναμη της μηχανής της εξωτερικής βάσης δεδομένων.



Σχήμα 8 : Η αρχιτεκτονική στο SSAS για Tabular models [11]

Η αξιοποίηση του ενός τρόπου σύνδεσης δεδομένων είτε του άλλου, παρέχει πλεονεκτήματα και μειονεκτήματα. Κάποιες φορές, είναι ασύμφορο ή πρακτικά αδύνατο να εισάγουμε τα δεδομένα μέσω VertiPaq, λόγω του μεγέθους του μοντέλου δεδομένων, όπως στην περίπτωση το μοντέλο δεδομένων να υπερβαίνει τη μνήμη RAM του συστήματος. Λύση

θα μπορούσε να ήταν να αυξηθεί η μνήμη του συστήματος είτε να γίνει βελτιστοποίηση του μοντέλου για να μικρίνει το μέγεθος του. Σε αυτή την περίπτωση είναι χρήσιμη η ύπαρξη του DirectQuery η οποία βεβαίως έχει το μειονέκτημα ότι δεν κάνει χρήση της υπερσύγχρονης μηχανής VertiPaq του Power BI. Γενικώς, η συμβουλή από τους σχεδιαστές μοντέλων δεδομένων είναι να αποφεύγεται η χρήση του DirectQuery ,όπου είναι εφικτό και τα δεδομένα πάντα να εισάγονται στο σύστημα. [10]

Στην περίπτωση όπου τα δεδομένα στις πηγές δεδομένων μπορεί να αλλάζουν τακτικά ,οπότε η χρήση του DirectQuery πλεονεκτεί σε σχέση με την κλασσική εισαγωγή δεδομένων με την VertiPaq.

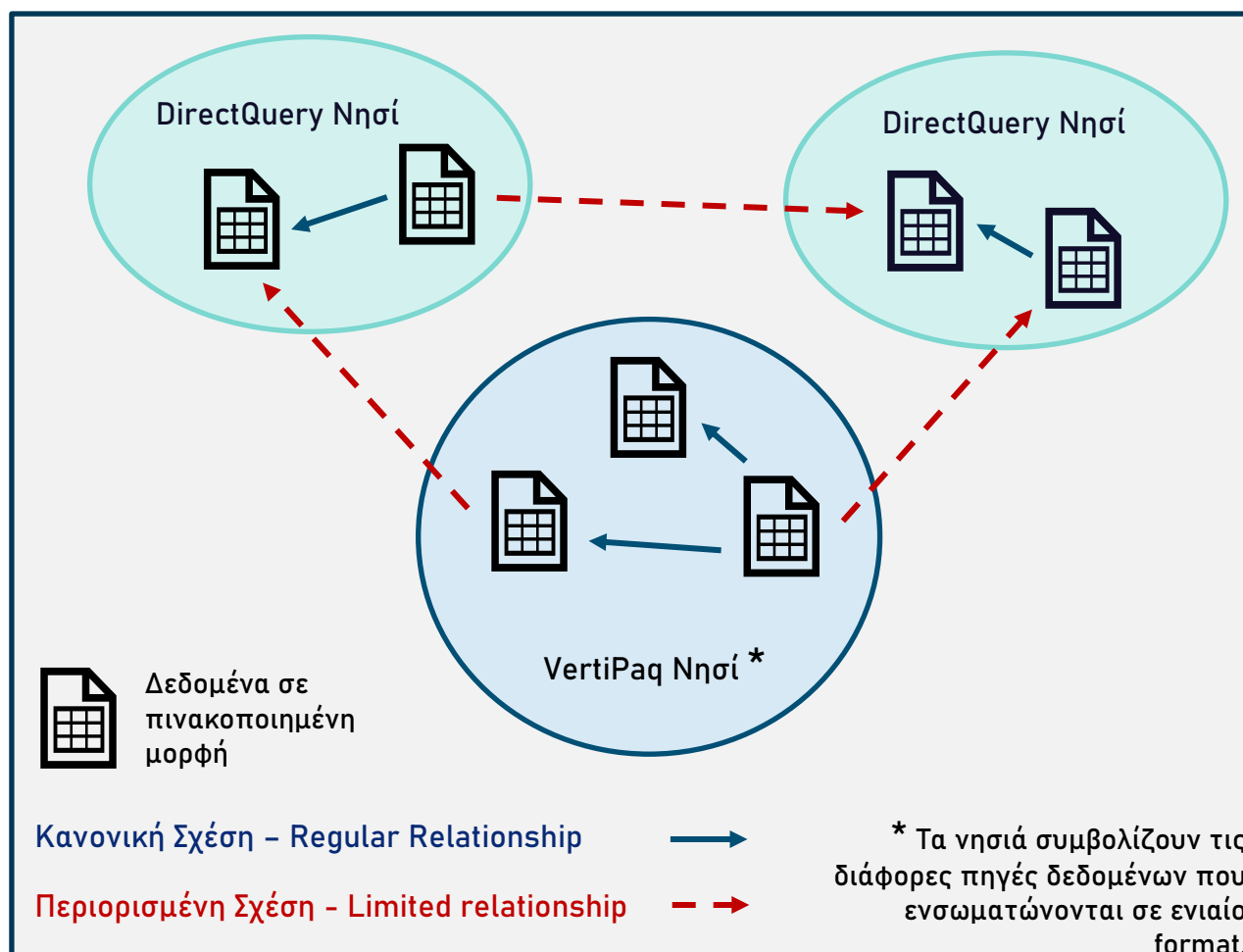
	VertiPaq	DirectQuery
Αποθήκευση Δεδομένων	Γίνεται εισαγωγή στην μνήμη RAM ,σε μορφή στήλης	Δεν γίνεται αποθήκευση
Υπολογιστική Δύναμη CPU	Διαθέτει μεγάλη υπολογιστική δύναμη ,με αντίστοιχα μεγάλο κόστος στους πόρους	Δεν απαιτεί ,αλλά εξαρτάται πλήρως από εξωτερική βάση δεδομένων
Περιορισμοί	Έχει περιορισμό στο μέγεθος του μοντέλου δεδομένου ,λόγω της πεπερασμένης μνήμης RAM του συστήματος	Δεν υπάρχει περιορισμός για το μέγεθος του μοντέλου
Χρόνος απάντησης στα queries	Βέλτιστος Χρόνος	Αυξημένος ,εξαρτάται από την εξωτερική βάση δεδομένων
Real-time queries	Δεν είναι δυνατόν	Πάντα είναι
Ενσωμάτωση διαφόρων πηγών δεδομένων	Τα δεδομένα μπορούν να ενσωματωθούν σε ένα μοναδικό format	Μέχρι το 2018 δεν ήταν εφικτό, όπου συστήθηκαν τα σύνθετα μοντέλα (composite models)

Πίνακας 5: Σύγκριση των μεθόδων σύνδεσης δεδομένων στο Power BI

Composite model [12]

Μέχρι το 2018 δεν ήταν εφικτή η ενσωμάτωση περισσότερων από μία πηγή δεδομένων στο μοντέλο δεδομένων, όταν γινόταν χρήση του DirectQuery ,ασχέτως αν αυτή η σύνδεση γινόταν μέσω εισαγωγής δεδομένων (**import mode**) ,δηλαδή της μηχανής VertiPaq είτε εκ νέου μέσω του DirectQuery. Πλέον, το Power BI δίνει τη δυνατότητα να

κατασκευαστούν σύνθετα μοντέλα (**composite models**) δεδομένων. Δηλαδή σε ένα μοντέλο δεδομένων μπορούν να ενσωματωθούν ,πηγές δεδομένων είτε προέρχονται μέσω DirectQuery είτε εισάγονται απευθείας στο σύστημα.



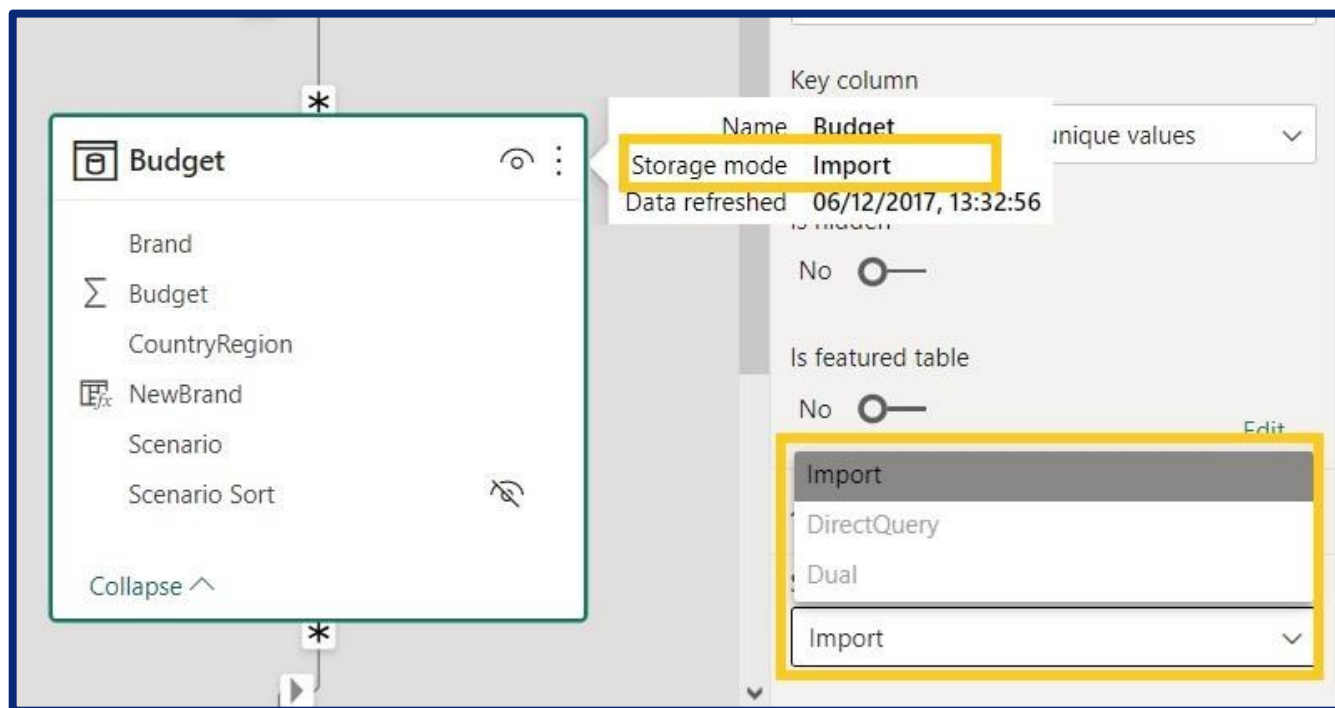
Σχήμα 9 : Σύνθετο μοντέλο δεδομένων (composite model) [13]

Μαζί με τα σύνθετα μοντέλα (composite models) στο Power BI εγκαινιάσθηκαν και οι σχέσεις **Many-to-many (*:*)** μεταξύ πινάκων του μοντέλου δεδομένων. Αυτές οι σχέσεις ,οι οποίες θα αναλυθούν εκτενέστερα στην υποενότητα **3.5.3 Σχέσεις του Μοντέλου Δεδομένων – Model Relationships**, μαζί με τις σχέσεις μεταξύ διαφορετικών νησιών – πηγών, όπως φαίνεται στο **Σχήμα 9**, αποτελούν ένα ειδικό είδος σχέσεων οι οποίες ονομάζονται περιορισμένες (**limited relationships**). Οι σχέσεις αυτές έχουν το χαρακτηριστικό ότι το Power BI δεν μπορεί να επιβεβαιώσει ότι ο ένας πίνακας από τους δύο ,στις σχέσεις αυτές, περιέχει μοναδικές τιμές, όπως συμβαίνει στις συνήθεις σχέσεις (**regular relationships**) .

Επιπλέον, εγκαινιάσθηκε η ιδιότητα **Storage Mode** στο *Model View* του Power BI Desktop. Αυτή η ιδιότητα επιτρέπει τον εκ των υστέρων ορισμό της μορφής που συνδέονται οι πίνακες στο μοντέλο δεδομένων και διαθέτει τρεις ρυθμίσεις:

- Import
- DirectQuery
- και Dual

Στην ρύθμιση **Import** οι πίνακες πάνε στην cache μνήμη. Άπαξ και επιλεχθεί αυτή η ρύθμιση δεν μπορεί να αντιστραφεί. Στην ρύθμιση **DirectQuery**, τα δεδομένα δεν εισάγονται στην μνήμη cache, ενώ στην **Dual** οι πίνακες ανάλογα με το περιεχόμενο του ερωτήματος που τίθεται στο μοντέλο μπορούν να μπουν ή όχι στη μνήμη cache. Η ρύθμιση μπορούμε να καθορίσουμε ποιες οπτικοποιήσεις υποβάλλουν queries σε back-end πηγές δεδομένων. [12]



Διάγραμμα 3 : Storage mode

3.3.2 Σύνδεση Δεδομένων

Το πρώτο βήμα σε κάθε αναλυτικό σύστημα ασφαλώς είναι η διασύνδεση των

δεδομένων με το σύστημα, στην συνέχεια ακολουθεί η ενσωμάτωση των δεδομένων που προέρχονται από διάφορες πηγές, βάσεις δεδομένων, λειτουργικά συστήματα είτε από κάποια Αποθήκη Δεδομένων (DW) που περιέχει όλα τα δεδομένα της εταιρείας. Έπειτα ακολουθεί ο καθαρισμός και επεξεργασία των στοιχείων και τελικά η παράδοση τους προς αξιοποίηση. Με το Power BI Desktop, υπάρχουν τρεις τρόποι με τους οποίους μπορούμε να συνδέσουμε δεδομένα:

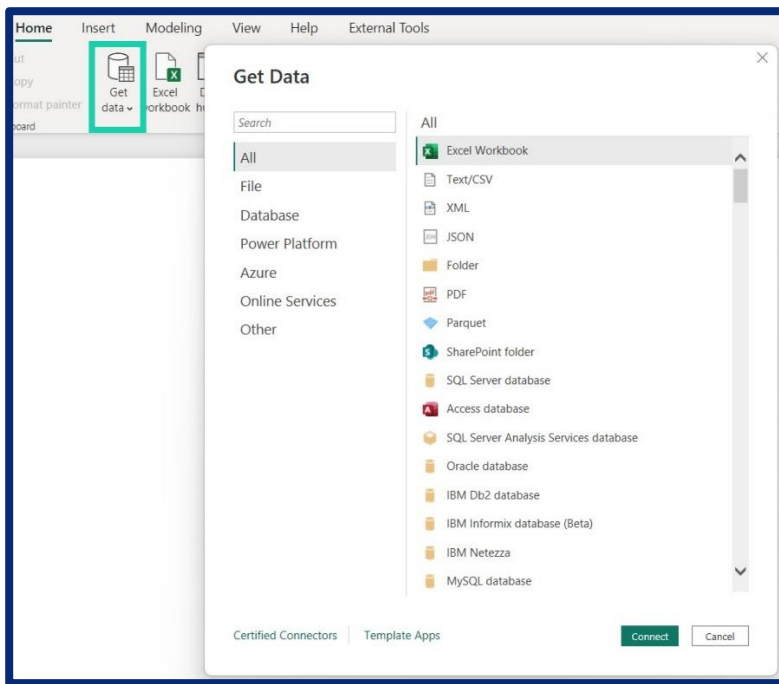
- Εισαγωγή (Import)
- Μέσω του *DirectQuery*
- Και *live connection*

Οι δύο πρώτοι τρόποι αναλύθηκαν αναλυτικά στην υποενότητα **3.4.1 Tabular Processing**. Η σύνδεση μέσω *live connection* μοιάζει πολύ με το *DirectQuery* στον τρόπο που σχετίζεται με την πηγή δεδομένων. Δεν αποθηκεύει δεδομένα στο Power BI και στέλνει ερωτήματα στην πηγή δεδομένων κάθε φορά. Ωστόσο, είναι διαφορετικό από το *DirectQuery*. Η Ζωντανή Σύνδεση υποστηρίζεται μόνο για αυτές τις πηγές δεδομένων.

- Azure Analysis Services
- SQL Server Analysis Services (SSAS) Tabular
- SQL Server Analysis Services (SSAS) Multi-Dimensional
- **Power BI Dataset** από το Power BI Service

Επειδή αυτές οι πηγές δεδομένων είναι οι ίδιες οι μηχανές μοντελοποίησης, **το Power BI** συνδέεται με αυτές και **ανακτά μόνο τα μεταδεδομένα του μοντέλου**. Με αυτήν τη μέθοδο, η μοντελοποίηση γίνεται αποκλειστικά στην ίδια την πηγή και το Power BI ,απλώς εμφανίζει αυτά τα δεδομένα μέσω της Οπτικοποίησης.

Ο πιο συνήθης τρόπος που συνδέουμε δεδομένα είναι από το κουμπί **Get Data** στην κορδέλα, στο πάνω μέρος του interface.

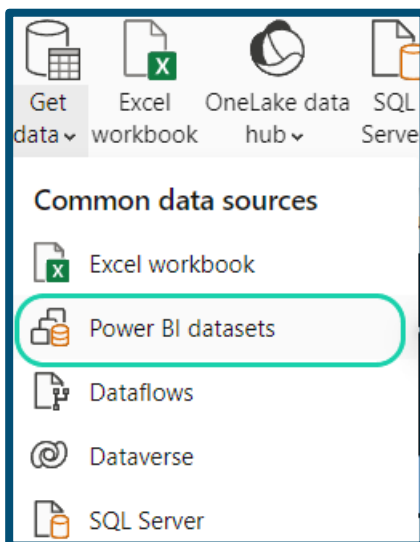


Εικόνα 9: Το Get Data κουμπί

Από εκεί μπορούν να συνδεθούν είτε να φορτωθούν τύποι αρχείων όπως *.xlsx*, *.xlsm*, *.csv*, *.json*, *.pdf*, να γίνει σύνδεση με τις βάσεις δεδομένων πιστοποιημένων υπηρεσιών, όπως *Google Analytics*, *GitHub*, *Salesforce Objects*, *Dynamics 365*, *SharePoint*, *Azure*. Η συνδεσιμότητα είναι σχεδόν καθολική.

Η σύνδεση με τα μέσω **Power BI Dataset**, έχει αξία να αναφερθεί καθώς χρησιμοποιείται ευρέως από ομάδες σχεδιαστών και είναι εξαιρετικά χρήσιμη. Μέσω αυτής

της επιλογής, πραγματοποιείται σύνδεση ολόκληρου του μοντέλου δεδομένων (Πίνακες, σχέσεις, υπολογισμοί) ως αυτούσιο. Ο χρήστης μπορεί να δημιουργήσει αναφορές και οπτικοποιήσεις βασισόμενος πάνω σε αυτό το μοντέλο. Η επιλογή αυτή αφορά κοινόχρηστα μοντέλα δεδομένων ή **Shared Datasets**. Τα κοινόχρηστα μοντέλα δεδομένων είναι διαθέσιμα στον διακομιστή του **Power BI Service** και μπορούν να χρησιμοποιηθούν από όποιον έχει άδεια σε αυτά. Το τοπικό σύστημα πραγματοποιεί σύνδεση με το **Power BI Service** μέσω *live connection*. Όταν γίνει σύνδεση ενός **Shared Dataset** με το τοπικό σύστημα, πλέον δεν μπορούν να προστεθούν και άλλα δεδομένα, για λόγους που σχετίζονται με την κοινόχρηστη φύση του μοντέλου και από



Εικόνα 10: Σύνδεση κοινόχρηστου μοντέλου δεδομένων

άλλους χρήστες. Θα ακολουθήσει περισσότερη ανάλυση τους στο αντίστοιχο κεφάλαιο για το **3.6 Power BI Service – Δημοσίευση και Συνεργασία**.

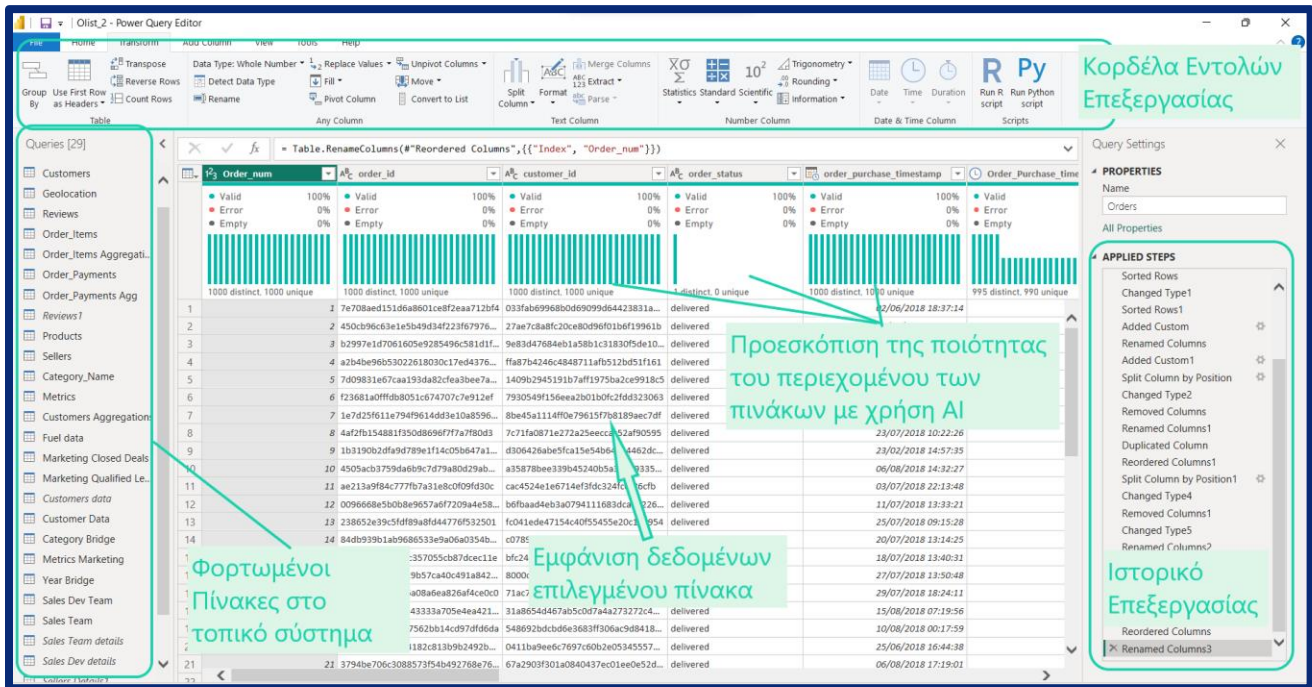
Ακόμη ένας σημαντικός τρόπος σύνδεσης δεδομένων στο **Power BI Desktop** είναι να μέσω των **Dataflows**. Αυτό είναι ένα self-service εργαλείο επεξεργασίας, όπως το **Power Query Editor** στο **Power BI Desktop**, αλλά το οποίο βρίσκεται αποκλειστικά στο **Power BI Service**. Ουσιαστικά, τα δεδομένα που εισάγουμε μέσω αυτής της επιλογής έχουν διαμορφωθεί στο Cloud, δηλαδή το **Power BI Service**. Παρέχει παρόμοια εμπειρία επεξεργασίας δεδομένων στο χρήστη, παρέχοντας κάποια πλεονεκτήματα που δεν έχει η επεξεργασία στο τοπικό σύστημα, όπως η άμεση ενημέρωση των μοντέλων και των αναφορών που βρίσκονται στο **Power BI Service**, μετά από κάποια επεξεργασία που πραγματοποιήθηκε σε αυτά. Ένα χαρακτηριστικό εξαιρετικά χρήσιμο για ομάδες επαγγελματιών που χρησιμοποιούν το Power BI.

3.3.3 Διαμόρφωση Δεδομένων

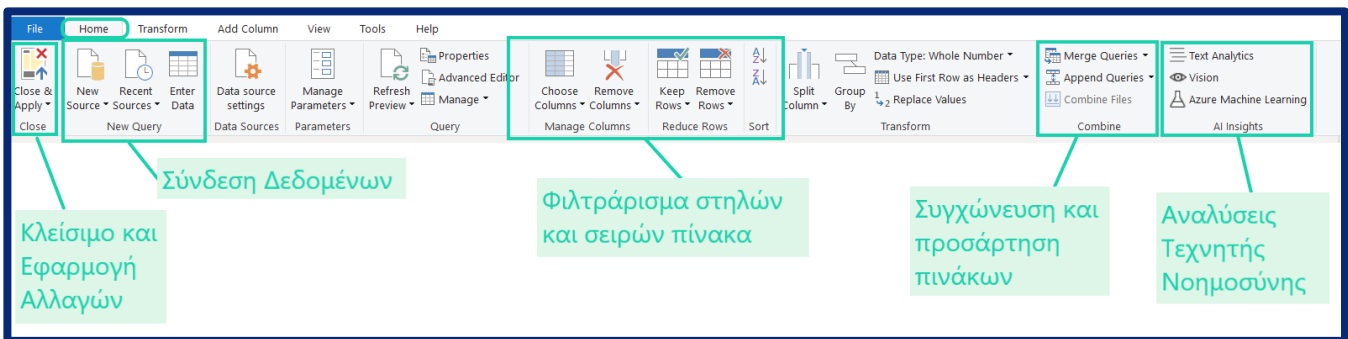
Αφού γίνει η σύνδεση των δεδομένων στο τοπικό σύστημα, ακολουθεί η διαμόρφωση, ο καθαρισμός και επεξεργασία των δεδομένων. Για να το κάνει αυτό το **Power BI Desktop** χρησιμοποιεί τον δικό του Editor, το **Power Query Editor**. Αυτός ο επεξεργαστής έχει πανίσχυρες δυνατότητες επεξεργασίας σε απλοποιημένη γλώσσα, την *Power Query M*.

Οι περισσότερες τροποποιήσεις και λειτουργίες μπορούν να γίνουν απλά με το πάτημα ενός κουμπιού. Είναι ένας Editor μπολιασμένος με αλγόριθμους τεχνητής νοημοσύνης για να μπορεί να κάνει έξυπνες λειτουργίες όπως η προεπισκόπηση των δεδομένων και η παροχή χρήσιμων παρατηρήσεων στον χρήστη για τα δεδομένα που μεταχειρίζεται. Παρέχει λειτουργίες όπως *Προσθήκη/Αφαίρεση Στηλών Αφαίρεση Διπλότυπων Τιμών*, *φιλτράρισμα συγκεκριμένων σειρών*, *Συγχώνευση Πινάκων*, *Ταξινόμηση Στηλών*, *Ομαδοποίηση Τιμών*.

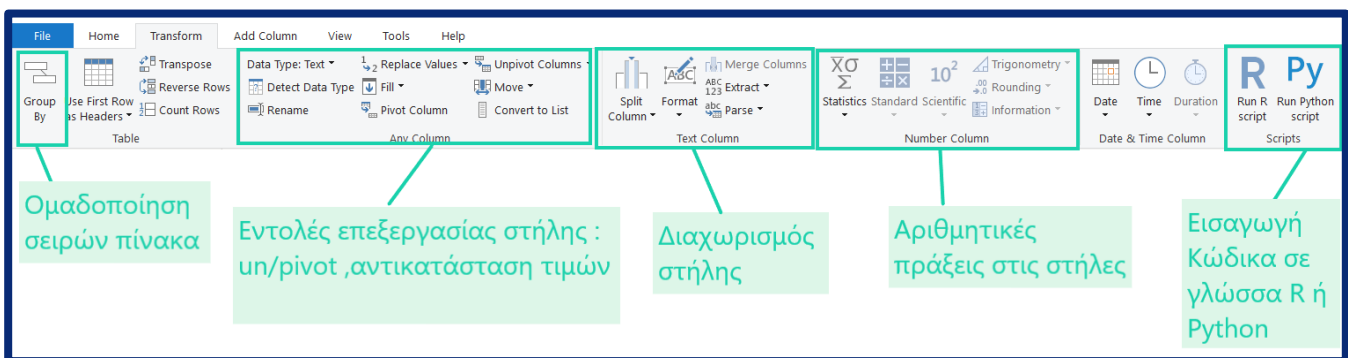
Ας γνωρίσουμε την διεπιφάνεια του Επεξεργαστή.



Εικόνα 13 : Στιγμιότυπο από τον Επεξεργαστή – Power Query Editor



Εικόνα 12 : Κορδέλα επεξεργασίας Home – Power Query Editor



Εικόνα 11 Κορδέλα επεξεργασίας Transform – Power Query Editor

Ορισμένες από τις δυνατότητες του επεξεργαστή φαίνονται στην παραπάνω εικόνα .

Μετά την επεξεργασία των δεδομένων ακολουθεί η σύνδεση τους στο βασικό μοντέλο δεδομένων.

* Παραδείγματα αξιοποίησης του επεξεργαστή θα ακολουθήσουν στην υποενότητα **4.2 Εισαγωγή Δεδομένων και κατασκευή Μοντέλου Δεδομένων**

3.4 Power BI Desktop - Μοντελοποίηση Δεδομένων

Μια βασική έννοια που πρέπει να οριστεί πριν προχωρήσουμε στην λειτουργικότητα της πλατφόρμας του Power BI είναι αυτή του **Μοντέλου Δεδομένων**.

«Το μοντέλο δεδομένων είναι μια αναπαράσταση πραγμάτων που έχουν σημασία για έναν οργανισμό ,καθώς και η σχέση μεταξύ αυτών. Απεικονίζει την υποκείμενη δομή των δεδομένων του οργανισμού και αντανακλά την δομή των βάσεων δεδομένων που το υποστηρίζουν» [15]

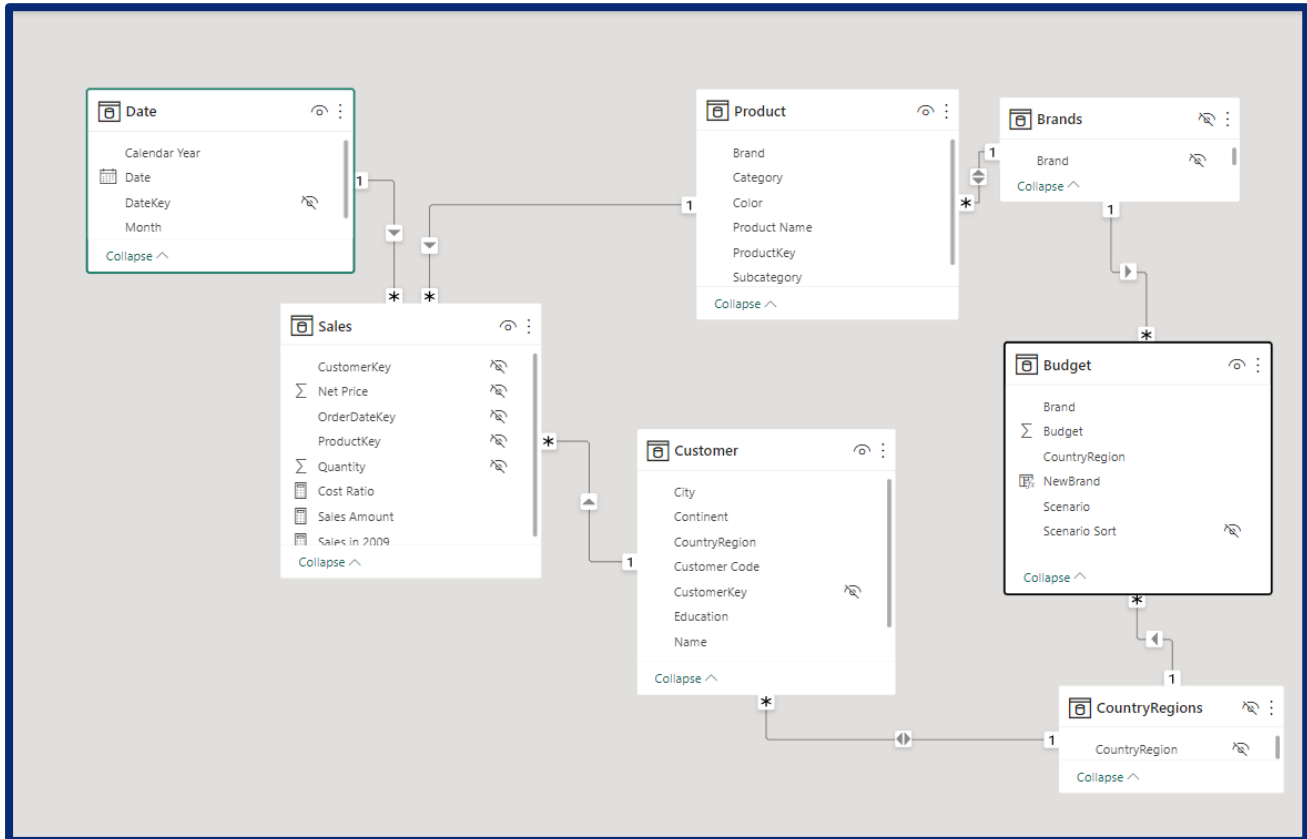
Εξυπηρετεί δύο σκοπούς στα αναλυτικά συστήματα. Αρχικά, ως μια τεχνική γραφικής απεικόνισης, βοηθάει την επικοινωνία μεταξύ αναλυτών και των τελικών χρηστών , μεταξύ των ίδιων των αναλυτών ,όσο και με αυτούς που σχεδιάζουν το σύστημα. Επίσης, θέτει ένα αυστηρό πλαίσιο ,στο οποίο είναι ξεκάθαρες σε όλους οι λογικές συνέπειες των σχέσεων των δεδομένων.

3.4.1 Μοντέλα Δεδομένων

Το μοντέλο δεδομένων ως αναπαράσταση είναι σε πινακοειδή μορφή και είναι μέρος ενός πινακοειδούς μοντέλου - **Tabular Model**. Ένα Tabular Μοντέλο είναι ένα σύνολο μεταδεδομένων όπως πίνακες, σχέσεις, στήλες, KPIs, ομάδες υπολογισμού, ιεραρχίες, ρόλους ασφαλείας και πολλά άλλα στοιχεία που αποτελούν το "σημασιολογικό μοντέλο" (**semantic model**) που χρησιμοποιείται ως σύστημα πλοήγησης σε εργαλεία όπως το Power BI και το Excel. Σε ένα μοντέλο Tabular, μια φυσική σχέση συνδέει δύο πίνακες μέσω μιας ισοδυναμίας σε μια στήλη. Ο σκοπός μιας σχέσης σε ένα μοντέλο Tabular είναι να διαδίδει ένα φίλτρο όταν γίνεται ερώτηση στο μοντέλο δεδομένων [10].

Το Power BI κάνει χρήση της μηχανής των **Analysis Services** για analytics , την **VertiPaq**. Δηλαδή, το μοντέλο δεδομένων στο Power BI είναι σε πινακοποιημένη μορφή. Συγκεκριμένα, το Power BI κάνει χρήση του TOM (Tabular Object Model) ,το οποίο είναι επέκταση της βιβλιοθήκης – client **AOM** (Analysis Management Object). Το TOM παρέχει έναν

προγραμματιστικό τρόπο χειρισμού διαχειριστικών λειτουργιών όπως η δημιουργία μοντέλων, η εισαγωγή και η ανανέωση δεδομένων και η εκχώρηση ρόλων και δικαιωμάτων [16]. Το Power BI είναι σχεδιασμένο ώστε να λειτουργεί ταχύτερα αν δημιουργείται Διαστασιακό Μοντέλο Δεδομένων (**Dimensional Data Model**), δηλαδή σε σχήμα Αστέρι (**Star Schema**) ή σε Κύβο (**Cube**) αν εφαρμόζεται σε πολύ-διαστατικές βάσεις δεδομένων.



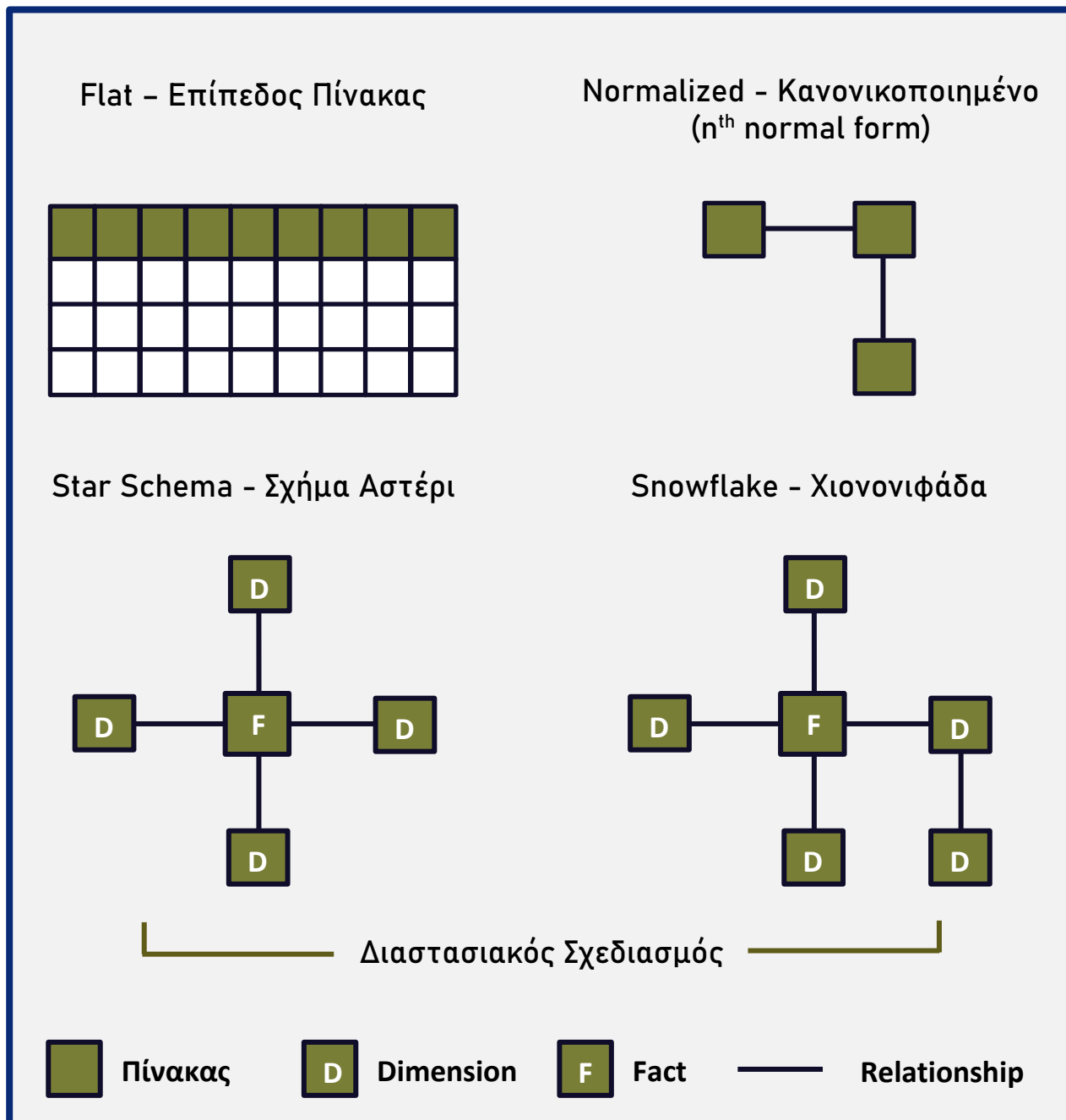
Διάγραμμα 4: Tabular Μοντέλο Δεδομένων στον γραφικό περιβάλλον του Power BI Desktop

Κάθε πίνακας στο μοντέλο περιέχει στήλες τα οποία στην βιβλιογραφία συχνά αναφέρονται ,ως πεδία (Fields), όπως ο πίνακας Product , στο παράδειγμα του Διαγράμματος 5 , διαθέτει τα πεδία Product[Brand], Product[Category], Product[Color], Product[Product Name], Product[ProductKey] και Product[Subcategory] ,ενώ κάθε σειρά από την κάθε στήλη ονομάζεται καταγραφή (record).

Στην πράξη **συναντάμε τέσσερα μοντέλα** δεδομένων :

- Επίπεδου πίνακα (**Flat**)
- Κανονικοποιημένου πίνακα (**Normalized**)

- Σχήματος Αστεριού (**Star Schema**)
- και της Χιονονιφάδας (**Snowflake**)

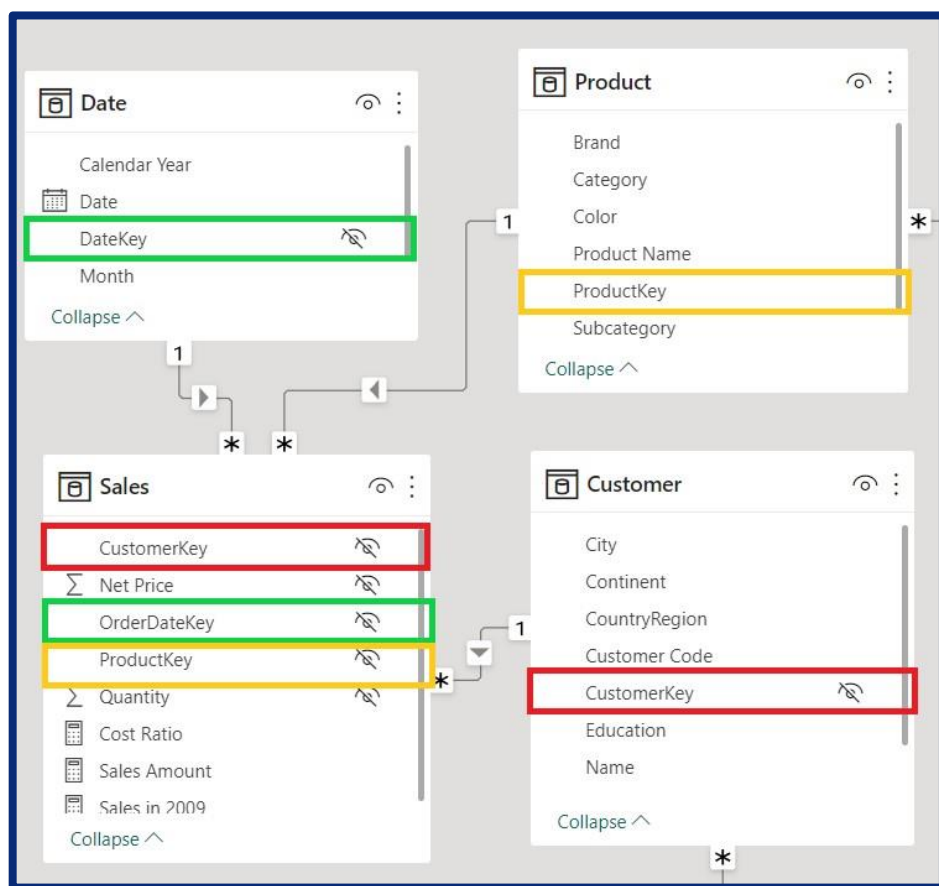


Σχήμα 10 : Μοντέλα Δεδομένων

Επίπεδος Πίνακας (Flat): Αποτελεί το πιο απλοϊκό είδος μοντέλου δεδομένων και είναι αυτό που χρησιμοποιούν οι περισσότερες επιχειρήσεις. Είναι συνήθως ένας επιμήκης πίνακας πολλών στηλών και θεωρείται αποκαθετοποιημένος, δηλαδή έχει επαναλαμβανόμενες τιμές.

Τέτοιοι πίνακες προέρχονται από υπολογιστικά φύλλα Excel και είναι κατασκευασμένες από προσωπικό που δεν είναι οικείο με την κατασκευή μοντέλο δεδομένων.

Κανονικοποιημένος (Normalized) : Κανονικοποίηση είναι ο όρος που χρησιμοποιείται για να περιγράψει δεδομένα που αποθηκεύονται με τρόπο που μειώνει τα επαναλαμβανόμενα δεδομένα. Καθετοποιημένος πίνακας θεωρείται αυτός στον οποίο αποθηκεύονται στήλες με μοναδικές τιμές κλειδιά (**unique key value column**). Για παράδειγμα ,ο πίνακας *Sales* στο **Διάγραμμα 6**, θεωρείται κανονικοποιημένος, καθώς διαθέτει στήλες με μοναδικά κλειδιά ,όπως την *Sales[ProductKey]* και *Sales[CostumerKey]* και *Sales[OrderDateKey]* , τα οποία χρησιμοποιεί για να συνδεθεί με τους γειτονικούς πίνακες *Product*, *Costumer* και *Date* αντίστοιχα ,χωρίς να υπάρχει η ανάγκη να υπάρχουν οι υπόλοιπες στήλες από τους πίνακες αυτούς ,στον πίνακα *Sales*. Γενικώς η κανονικοποίηση ή αλλιώς καθετοποίηση αυξάνει δραστικά την απόδοση της μια αναλυτικής μηχανής ,όπως αυτής του Power BI , στην απόδοση ερωτημάτων (querying). Η μορφή του κανονικοποίησης ,δηλαδή αν θα πρόκειται για 2ND Normal Form - 2NF ,3NF ,4NF ή 5NF, εξαρτάται από την πόσο μεγάλη καθετοποίηση μας ενδιαφέρει να έχει η εκάστοτε εφαρμογή.



Διάγραμμα 5 : Κανονικοποιημένος πίνακας Sales

Οι έννοιες του Σχήματος Αστεριού (**Star Shema**) και η παραλλαγή αυτού, Χιονονιφάδα (**Snowflake**), εξηγούνται με λεπτομέρεια στην επόμενη ενότητα ,καθώς το μοντέλο δεδομένων που αξιοποιεί σε μέγιστο βαθμό την λειτουργικότητα του Power BI είναι αυτό του σχήματος αστεριού και γι' αυτό συνήθως χρησιμοποιείται στα αναλυτικά συστήματα.

3.4.2 Σχήμα Αστέρι – Star Schema

Στη δεκαετία του 1990, ο R.Kimball ήταν αυτός που ευθύνεται σε μεγάλο βαθμό για τη διάδοση του Σχήματος Αστέρι. Μέσα από τα γραπτά του, ο R.Kimball συνέθεσε και συστηματοποίησε μια σειρά από τεχνικές που είχαν χρησιμοποιηθεί ήδη από τη δεκαετία του 1960. Η μοντελοποίηση με το σχήμα Αστέρι ή διαφορετικά η **μοντελοποίηση διαστάσεων** είναι ευρέως αποδεκτή ως η τεχνική για την παρουσίαση αναλυτικών δεδομένων επειδή καλύπτει δύο ταυτόχρονες απαιτήσεις:

- Παρέχει δεδομένα που είναι κατανοητά στους χρήστες της επιχείρησης.
- Παρέχει καλή απόδοση ερωτήματος (fast query performance). [16]

«Ο σχεδιασμός διαστάσεων είναι σχεδιασμένος ώστε να υποστηρίξει την ανάλυση μιας επιχειρηματικής διαδικασίας μοντελοποιώντας τον τρόπο μέτρησής της.» [7]

Είναι εύκολο να διακρίνει κάποιος ποιες είναι οι μετρήσεις που αξιολογούνται σε έναν οργανισμό.

Για παράδειγμα ,ερωτήσεις που ακούει κάποιος σε έναν οργανισμό είναι οι εξής :

- Ποια είναι τα μικτά περιθώρια κέρδους ανά κατηγορία προϊόντων για τον Φεβρουάριο;
- Πόσες ημέρες ασθενείας πήραν οι υπάλληλοι του μάρκετινγκ πέρυσι;
- Ποιες είναι οι εκκρεμείς υποχρεώσεις ανά προμηθευτή;
- Ποιο είναι το ποσοστό επιστροφής ανά προμηθευτή;

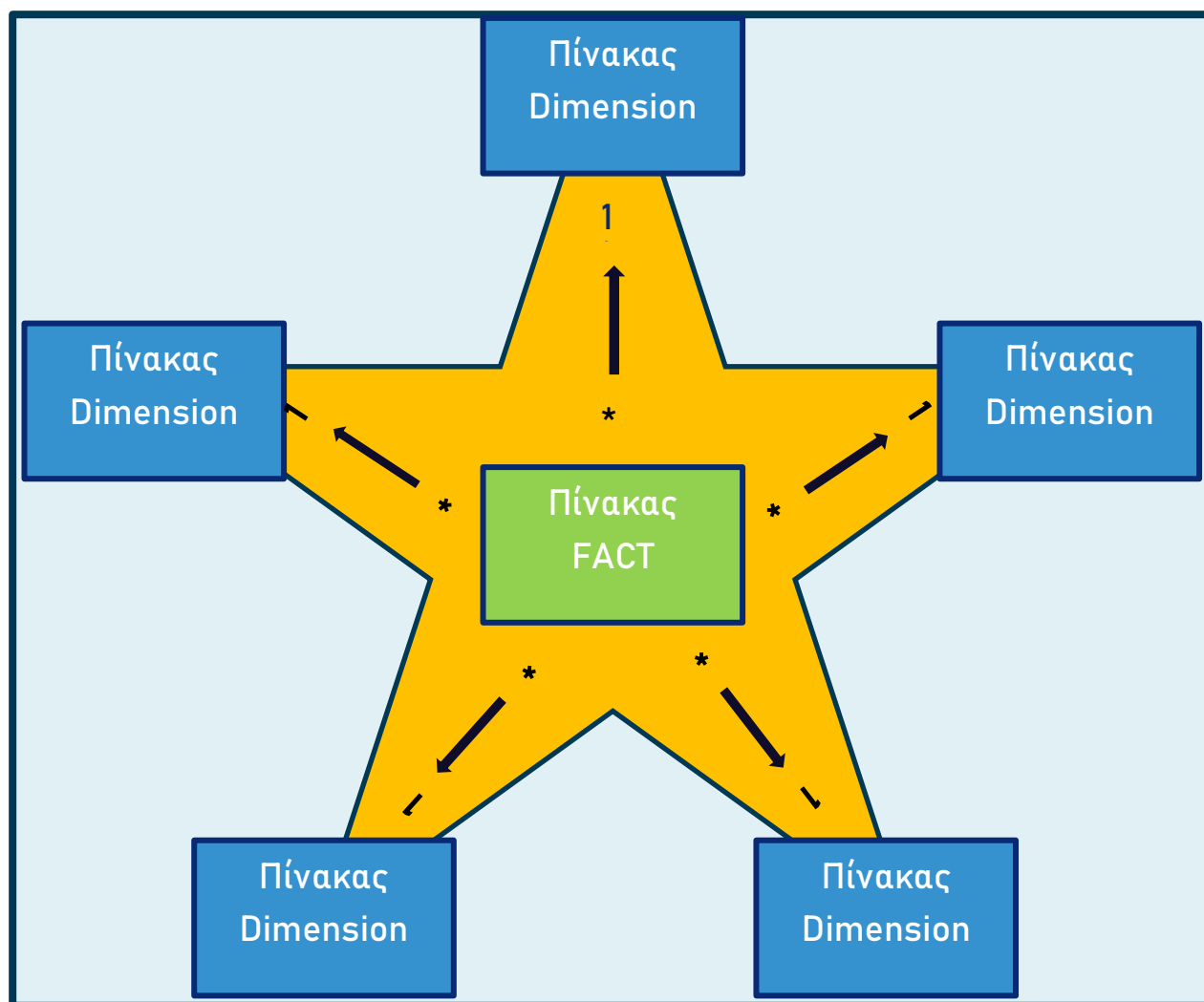
Αυτές είναι ερωτήσεις που ενέχουν **μετρήσεις** για συγκεκριμένες διαδικασίες της εταιρείας όπως τμήμα πωλήσεων θέλει να μετρήσει το μικτό κέρδος (Gross Margin),το τμήμα του HR μετράει τις απουσίες, το οικονομικό τμήμα θέλει να αξιολογήσει τις οφειλές και οι διευθυντές αγορών τις επιστροφές. Ενώ , οι μετρήσεις είναι αυτό που μας ενδιαφέρει να καταγραφούν , απαιτείται και περιεχόμενο γύρω από την μέτρηση .Το **περιεχόμενο** περιορίζει ή κάνει συγκεκριμένη την μέτρηση. Δηλαδή στις ερωτήσεις που προηγήθηκαν μας ενδιέφερε το μικτό κέρδος ανά κατηγορία προϊόντων ,σε συγκεκριμένο χρονικό διάστημα. Μας ενδιαφέρουν οι μέρες απουσίας αλλά μόνο για τα άτομα του μάρκετινγκ ,τον περσινό χρόνο. Μας ενδιαφέρει το ποσοστό επιστροφής ανά προμηθευτή. Η λέξη κλειδί είναι το ανά και η λέξη σε. Αυτές οι λέξεις μας συγκεκριμενοποιούν την μέτρηση.

Οι λέξεις **μετρήσεις** και **περιεχόμενο** είναι οι θεμελιώδεις έννοιες που εισήγαγε ο R.Kimball στο μοντέλο δεδομένων, μέσω του διαχωρισμού των πινάκων σε δύο κατηγορίες.

Στους πίνακες με τα γεγονότα - **Fact Table** και οι περιφερικοί πίνακες ,που ονομάζονται Διαστάσεις – **Dimension Table**.Οι μετρήσεις αντιστοιχούν στα γεγονότα και οι διαστάσεις στο περιεχόμενο.

Fact table: Περιέχει παρατηρήσεις ή συμβάντα και μπορεί να είναι παραγγελίες πωλήσεων, υπόλοιπα μετοχών, συναλλαγματικές ισοτιμίες, θερμοκρασίες ,ποσότητες μονάδων στην αποθήκη και πολλά άλλα είδη. Ένας πίνακας γεγονότων περιέχει δύο ειδών στήλες: στήλες με τιμές από τα κλειδιά διαστάσεων (**dimension key columns**) που σχετίζονται με τους αντίστοιχους πίνακες διαστάσεων και στήλες αριθμητικών μετρήσεων. Οι στήλες των κλειδιών διαστάσεων καθορίζουν τη πόσες διαστάσεις έχει το σχήμα αστέρι, ενώ οι τιμές των κλειδιών διαστάσεων καθορίζουν την βαθμό λεπτομέρειας (**granularity**) ενός πίνακα γεγονότων. Ο βαθμός λεπτομέρειας ενός πίνακα γεγονότων καθορίζει σε ποιο επίπεδο αποθηκεύουμε την πληροφορία. Για παράδειγμα, ο τυπικός πίνακας-διάσταση *Date* (δηλαδή οι ημερομηνίες που καταγράφονται τα γεγονότα) που υπάρχει σε κάθε μοντέλο δεδομένων, αν αποθηκεύει τα γεγονότα ανά εβδομάδα θεωρούμε ότι το επίπεδο λεπτομέρειας του είναι η εβδομάδα. Το επίπεδο λεπτομέρειας ονομάζεται στην βιβλιογραφία ως κόκκος (**grain**) και ουσιαστικά , σε ένα πίνακα γεγονότων καθορίζει τι εκπροσωπεί η κάθε σειρά του. Ατομικός κόκκος (**atomic grain**) θεωρείται το χαμηλότερο επίπεδο στο οποίο μπορεί να αποθηκευτεί η πληροφορία. *Κάθε πίνακας γεγονότων θα πρέπει να έχει τα δεδομένα τους στο ίδιο επίπεδο πληροφορίας ή να έχουν το ίδιο granularity* [13]. Στο **Διάγραμμα 3** ο πίνακας *Sales* αποτελεί παράδειγμα ενός πίνακα γεγονότων , ενώ όπως προτείνει ο R.Kimball είναι κανονικοποιημένος στην 3NF μορφή.

Dimension table: Οι πίνακες διαστάσεις μας παρέχουν το «Ποιος;», το «Πότε;», το «Γιατί;», το «Που;» και το «Πώς;». Περιγράφουν επιχειρηματικές οντότητες—τα πράγματα που μοντελοποιούμε. Οι οντότητες μπορούν να περιλαμβάνουν προϊόντα, άτομα, μέρη και έννοιες όπως του χρόνου. Ο πιο συνεπής πίνακας που βρίσκει κάποιος σε ένα σχήμα αστέρι είναι ένας πίνακας διαστάσεων ημερομηνίας (*Date*). Ένας πίνακας διαστάσεων περιέχει δύο ειδών στήλες: μια βασική στήλη (ή στήλες) που λειτουργεί ως μοναδικό αναγνωριστικό και στήλες με περιγραφικό περιεχόμενο των γεγονότων που καταγράφουμε. Όπως ,εξηγείται και στην ενότητα με τις Σχέσεις του Μοντέλου Δεδομένων , η ύπαρξη του μοναδικού αναγνωριστικού , εξασφαλίζει ότι οι σχέση που συνάπτεται με τον πίνακα γεγονότων είναι **One-to-many (1:*)**. Δηλαδή μια τιμή στον πίνακα διάστασης μπορεί να εντοπιστεί πολλαπλές φορές στον πίνακα γεγονότων.



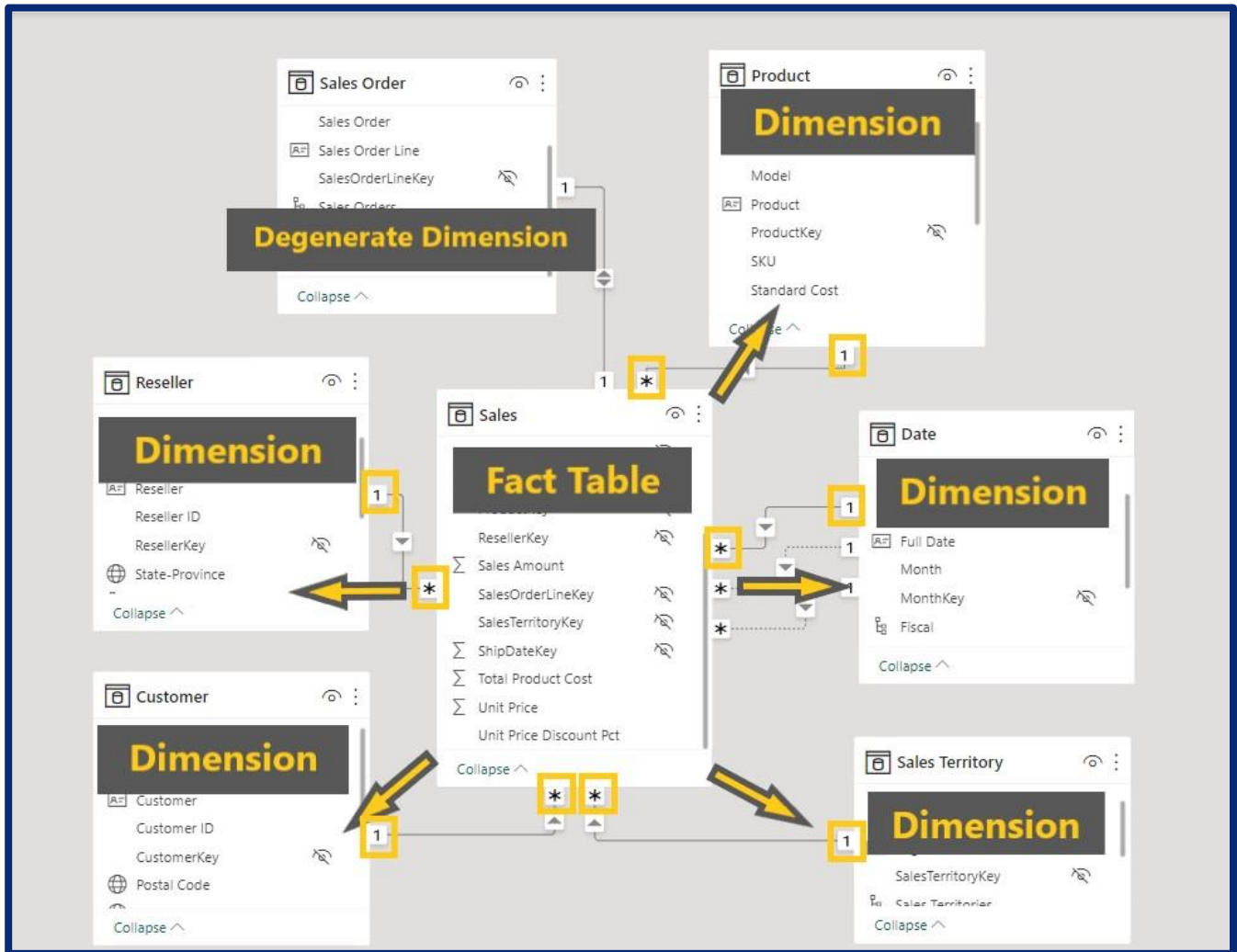
Σχήμα 11: Σχηματική αναπαράσταση του μοντέλου δεδομένων

Επίσης από την λειτουργική πλευρά τους οι πίνακες του σχήματος αστέρι επιτελούν τα εξής:

- **Dimension table** :Υποστηρίζουν το φιλτράρισμα και την ομαδοποίηση δεδομένων στα ερωτήματα (queries) που τίθενται από τον χρήστη.
- **Fact table** : Υποστηρίζουν την άθροιση/σύνοψη των δεδομένων στη μορφή που μας ενδιαφέρει.

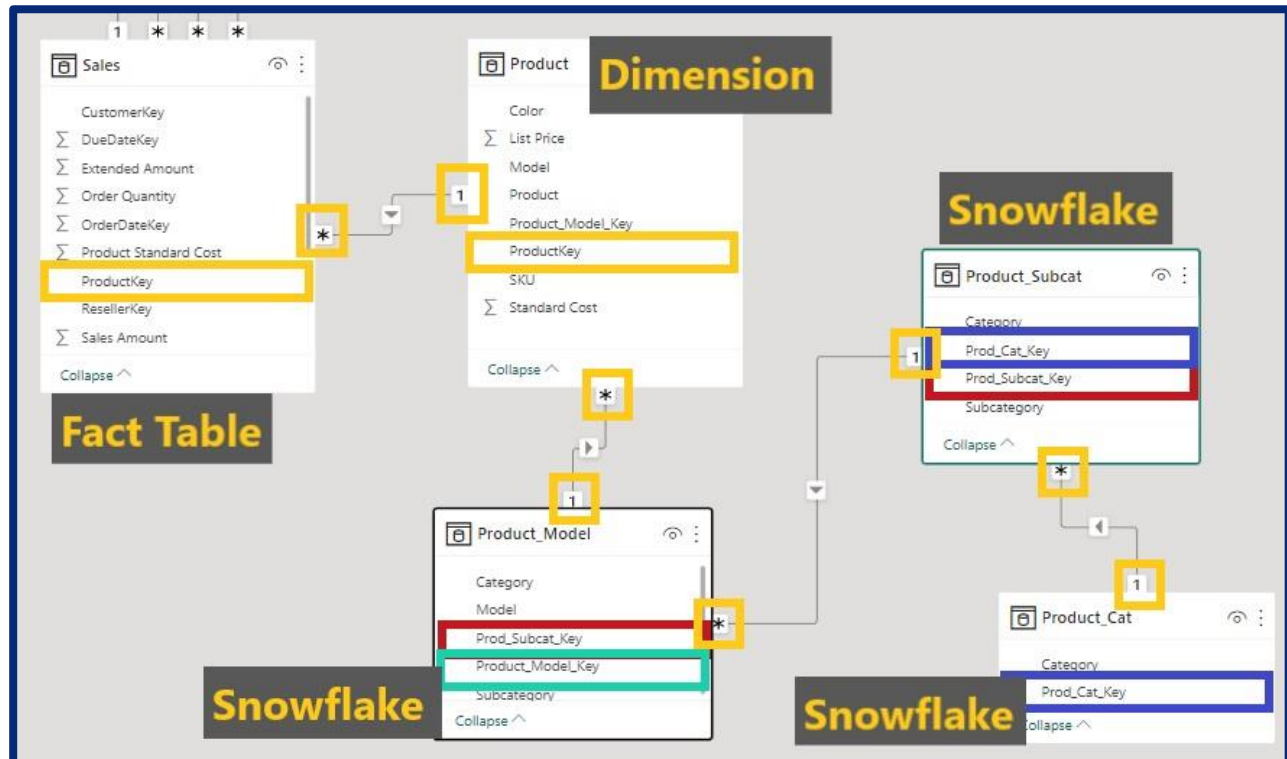
Ακολουθεί ένα παράδειγμα εφαρμογής της Σχεδίασης Αστέρι στο Power BI Desktop. Τα δεδομένα είναι από δείγμα της Microsoft για της πωλήσεις μια φανταστικής εταιρείας ,που ονομάζεται Adventure Works. Αποτελεί μια τυπική σχεδίαση ενός μοντέλου σε σχήμα Αστέρι ,όπου έχουμε έναν πίνακα γεγονότων στο κέντρο του μοντέλου [*Sales*] και περιμετρικά βρίσκονται οι πίνακες διαστάσεις [*Reseller*] , [*Costumer*] , [*Sales Territory*] , [*Date*] και [*Product*]. Υπάρχει και ένας επιπλέον πίνακας [*Sales Order*], ο οποίος δεν αποτελεί την

τυπική διάσταση αλλά μια εκφυλισμένη διάσταση , η οποία εξηγείται αναλυτικά σε επόμενη υποενότητα .



Διάγραμμα 6 : Μοντέλο Δεδομένων στο Power BI – Model View

Snowflake Schema - Σχήμα Χιονονιφάδας

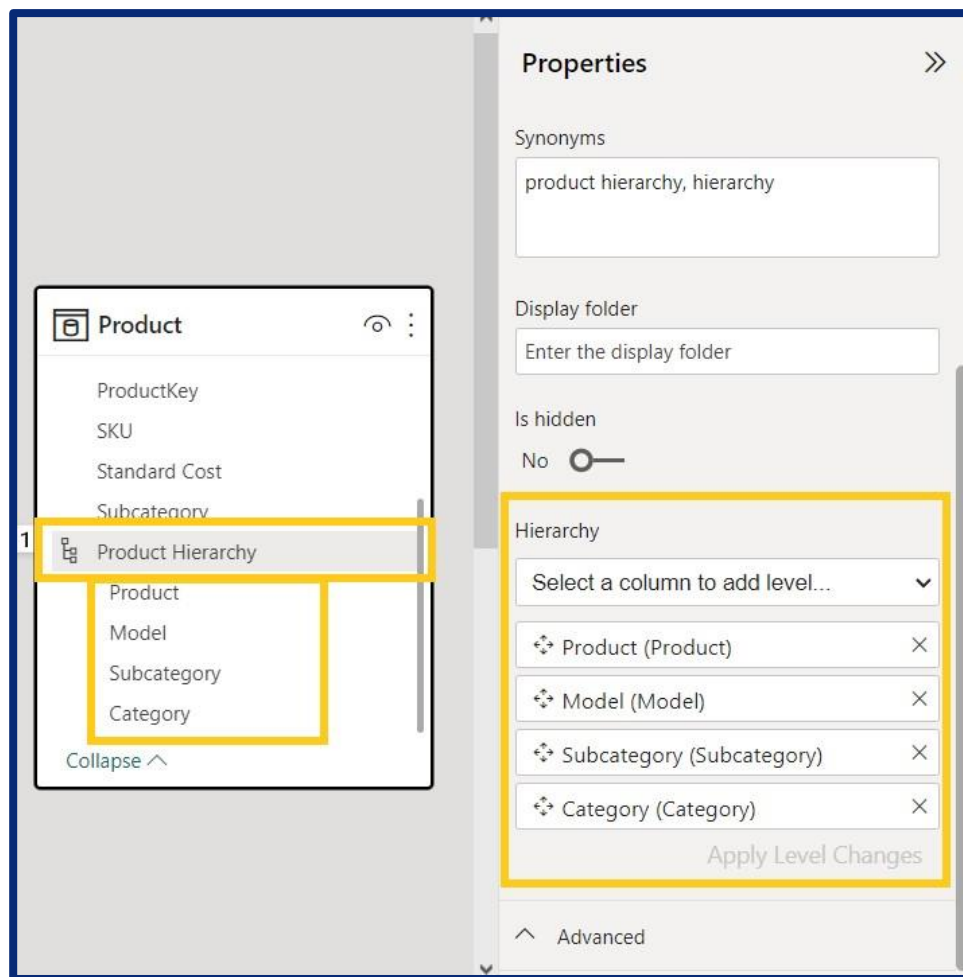


Διάγραμμα 7 : Διαστάσεις Χιονονιφάδας

Η κανονικοποίηση των πινάκων Διαστάσεων ενός σχήματος Αστερί δημιουργεί Χιονονιφάδες (**Snowflakes**). Η Διάσταση Χιονονιφάδας είναι ένα σύνολο κανονικοποιημένων πινάκων για μια μεμονωμένη επιχειρηματική οντότητα. Για παράδειγμα, στο **Διάγραμμα 7**, το οποίο έχει προέλθει από το βασικό παράδειγμα με του Σχήματος Αστεριού του **Διαγράμματος 6**, έχει γίνει καθετοποίηση του πίνακα – διάσταση [*Product*], ο οποίος κατηγοριοποιείται ανά [*Model*], στη συνέχεια ανά [*Subcategory*] και τέλος ανά [*Category*]. Αν αυτή η καθετοποίηση γινόταν σε όλες τις Διαστάσεις θα προέκυπτα η συμβολικά λεγόμενη «Χιονονιφάδα» λόγω του σχήματος του οποίου θα έπαιρνε το σχήμα. Παρατηρούμε ότι οι σχέσεις μεταξύ πινάκων είναι **Many-to-one (*:1)**. Κάτι το οποίο είναι λογικό καθώς, όσο κινούμαστε μακριά από τον πίνακα Διάσταση, κατηγοριοποιείται σε όλο και ευρύτερες υπερκατηγορίες.

Ένας από τους βασικούς λόγους που οι αναλυτές επιλέγουν σχήματα Χιονονιφάδες, είναι για να μειώσουν την επαναληψιμότητα (**Redundancy**) των τιμών στους πίνακες διαστάσεις. Αυτό το επιχείρημα, ενώ γενικά ισχύει, δεν δημιουργεί κάποια ουσιαστική διαφορά στο μέγεθος του μοντέλου δεδομένων, καθώς οι πίνακες Γεγονότων είναι αυτοί που έχουν τις περισσότερες επαναλαμβανόμενες τιμές, λόγω της φύσης τους και άρα επιβαρύνουν το μέγεθος του μοντέλου [13]. Επίσης, τα σχήματα Χιονονιφάδας επιτρέπουν στον χρήστη να δει

πιο εύκολα τις διάφορα επίπεδα κατηγοριοποίησης της κάθε διάστασης. Κάτι το οποίο , μπορεί να συμβεί αρκετά εύκολα στο Power BI μέσω της δημιουργίας Ιεραρχιών (**Hierarchies**) των επιχειρησιακών οντοτήτων. Μια λειτουργία εξαιρετικά χρήσιμη για οποιαδήποτε αναφορά.



Εικόνα 14 : Ιεραρχίες επιχειρηματικών οντοτήτων στο Power BI – Model View

Γενικά η χρήση των σχημάτων Χιονονιφάδων αντιπροτείνεται , καθώς αυξάνει την πολυπλοκότητα ανάγνωσης από τους χρήστες , το οποίο είναι βασικό στον διαστασιακό σχεδιασμό. Επίσης λόγω της μεγάλης καθετοποίησης απαιτείται να γραφούν πιο πολύπλοκα ερωτήματα (queries) κάτι το οποίο ζημιώνει την απόδοση.

Συγκεντρωτικά , μπορούμε να κάνουμε μια σύγκριση των τεσσάρων μεγάλων μοντέλων δεδομένων. Τα μοντέλο επίπεδου πίνακα δεν θα έχει καλή απόδοση σε πρότζεκτ με αναλυτικό χαρακτήρα , λόγω της μεγάλης επαναληψιμότητας των δεδομένων. Η μηχανή του Power BI όταν κάνει τις προσπελάσεις δεν μπορεί να βελτιστοποιήσει την διαδικασία , καθώς πρόκειται μόνο για έναν πίνακα. Ο μεγάλος αριθμός των πινάκων στην περίπτωση του

καθετοποιημένου μοντέλου επίσης επιβαρύνει την απόδοση για σκοπούς αναλυτικής φύσης ,αλλά η απόδοση είναι βελτιωμένη. Στον διαστασιακό σχεδιασμό έχουμε μέγιστη απόδοση, καθώς είναι η ενδεδειγμένη σχεδίαση για αναλυτικά συστήματα. Επίσης στον διαστασιακό σχεδιασμό έχουμε μικρό αριθμό (όγκο) ερωτημάτων αλλά μεγάλη πολυπλοκότητα στον τρόπο γραφής τους. Αντίθετα στα συστήματα συναλλαγών (transactional systems) όπου χρησιμοποιούνται κανονικοποιημένα σχήματα έχουμε μεγάλο αριθμό ερωτημάτων και μικρή πολυπλοκότητα στα ερωτήματα. Τέλος, όσον αφορά την συμπίεση η οποία γίνεται για την βελτιστοποίηση της απόδοσης του μοντέλου , στην περίπτωση του επίπεδου πίνακα και του κανονικοποιημένου μοντέλου ,η συμπίεση συμβαίνει στη σειρά (row) , για τον διαστασιακό σχεδιασμό γίνεται ανά στήλη .

	Flat Table	Normalized (3NF)	Star Schema	Snowflake
Απόδοση στα Analytics	Χαμηλή	Μεσαία	Υψηλή	Υψηλή
Φόρτος εργασίας για την δημιουργία	Χαμηλός	Υψηλός	Υψηλός	Υψηλός
Όγκος ερωτημάτων (queries)	Χαμηλός	Υψηλός	Χαμηλός	Χαμηλός
Πολυπλοκότητα ερωτημάτων (queries)	Χαμηλή	Χαμηλή	Υψηλή	Υψηλή
Εφαρμογή	Δεν υπάρχει	Λειτουργικά Συστήματα/ Εφαρμογές CRM / ERP	Αναλυτικά Συστήματα/ Αποθήκες Δεδομένων (Data Warehouses)	Αναλυτικά Συστήματα/ Αποθήκες Δεδομένων (Data Warehouses)
Μέθοδος συμπίεσης	Row	Row	Column	Column

Πίνακας 6 : Σύγκριση Μοντέλων Δεδομένων

3.4.3 Σχέσεις του Μοντέλου Δεδομένων - Model Relationships

Το επόμενο βασικό μέρος ενός μοντέλου δεδομένων αποτελούν οι σχέσεις του μοντέλου (**model relationships**). Όπως, αναφέρθηκε στην παράγραφο «**3.4.1 Tabular Processing**», η λειτουργία του **Power BI** στηρίζεται σε ένα σημασιολογικό μοντέλο, το οποίο

έχει την βάση του στο Tabular Object Model (**TOM**). Σε ένα μοντέλο Tabular, μια **φυσική σχέση** συνδέει δύο πίνακες μέσω μιας ισοδυναμίας μεταξύ δύο στηλών των πινάκων*. Ο σκοπός μιας σχέσης σε ένα μοντέλο Tabular είναι να **μεταφέρει ένα φίλτρο** όσο πραγματοποιεί ερωτήσεις (querying) στο μοντέλο. Τα φίλτρα θα διαδίδονται εφόσον υπάρχει μια διαδρομή σχέσης που μπορούν να ακολουθήσουν. Οι διαδρομές σχέσεων είναι ντετερμινιστικές, που σημαίνει ότι τα φίλτρα διαδίδονται πάντα με τον ίδιο τρόπο και χωρίς τυχαίες διακυμάνσεις, μέσω των κατευθύνσεων που έχουν οριστεί. Οι κατευθύνσεις που μεταδίδεται ένα φίλτρο, αναλύονται εκτενώς παρακάτω.

* Κάτω από ορισμένες συνθήκες που θα περιγράψουμε αργότερα, μια σχέση μπορεί να *θέσει έναν περιορισμό στο περιεχόμενο ενός πίνακα*.

Πέραν της μεταφοράς φίλτρου μέσω των σχέσεων μοντέλου, η μεταφορά ενός φίλτρου είναι δυνατή και με τη σύνταξη κώδικα DAX χωρίς να προϋποθέτει φυσική σχέση στο μοντέλο. Έτσι ορίζουμε δύο κατηγορίες σχέσεων:

- **Εικονικές σχέσεις** (Virtual Relationships) : πρόκειται για σχέσεις που δεν εμφανίζονται στο μοντέλο δεδομένων αλλά μπορούν να περιγράψουν στο λογικό μοντέλο. Οι εικονικές σχέσεις υλοποιούνται μέσω εκφράσεων DAX που μεταφέρουν το περιεχόμενο του φίλτρου (**filter context**) από τον έναν πίνακα στον άλλο. Βασική διαφοροποίηση της κατηγορίας είναι ότι η εικονική σχέση μπορεί να συμπεριλαμβάνει περισσότερο από ένα ζευγάρι στηλών.
- **Φυσικές σχέσεις** (Physical Relationships) : αυτές είναι οι σχέσεις που ορίζονται σε ένα μοντέλο Tabular και είναι αντιληπτές στον χρήστη μέσω της οπτικής απεικόνισης στην καρτέλα **Model View**. Οι φυσικές σχέσεις μπορούν να είναι είτε **ενεργές** είτε **ανενεργές**, αναλόγως ποιο φίλτρο μας ενδιαφέρει να μεταφέρουμε. Οπτικά, ο χρήστης διακρίνει την ανενεργή σχέση από την ενεργή, καθώς η πρώτη αποτυπώνεται στο οπτικό μοντέλο δεδομένων μέσω διακοπτόμενης γραμμής, ενώ η δεύτερη συναντάται με συνεχή γραμμή. Η μηχανή διαδίδει αυτόματα το περιεχόμενο του φίλτρου μέσω ερωτήματος, σύμφωνα με την κατεύθυνση φίλτρου της ενεργής σχέσης. Οι ανενεργές σχέσεις αγνοούνται, όμως υπάρχουν συναρτήσεις DAX που μπορούν να αλλάξουν την κατάσταση (ενεργή ή ανενεργή) και την κατεύθυνση του φίλτρου των φυσικών σχέσεων. Τυπικές συναρτήσεις της DAX που πραγματοποιούν αυτή την λειτουργία είναι η **USERELATIONSHIP()** και **CROSSFILTER()**.

Οι σχέσεις στο μοντέλο δεδομένων επίσης πέραν του χαρακτηριστικού που αναλύθηκε, δηλαδή αν θα είναι εικονικές ή φυσικές, μπορούν να διακρίνονται ανάλογα με την τοποθεσία των πηγαίων συστημάτων των οποίων συνδέουν.

Όπως είδαμε και στην υποενότητα **3.4.1 Tabular Processing** σε:

- Κανονικές (**Regular Relationships**)
- Περιορισμένες (**Limited Relationships**)

Φυσικές σχέσεις μοντέλου δεδομένων

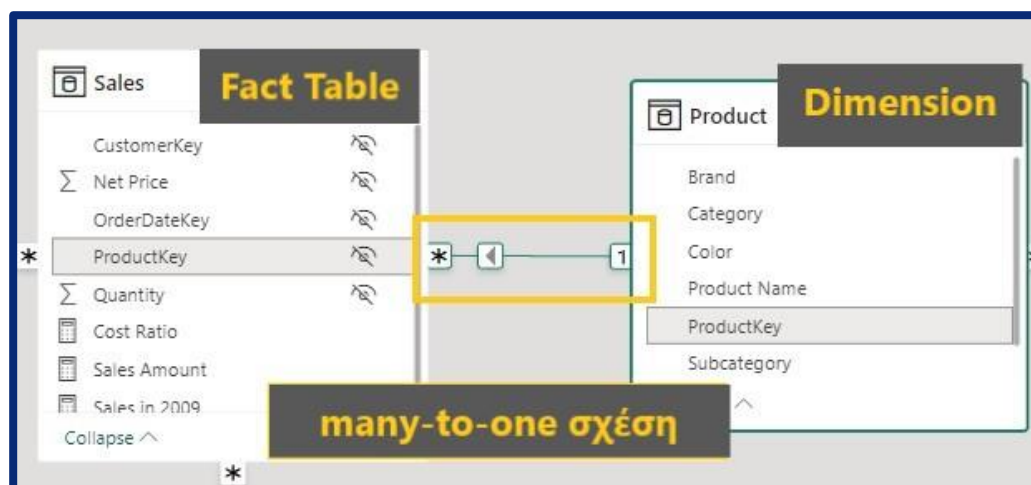
Στο Power BI υπάρχουν 3 Φυσικών σχέσεων:

- **One-to-many (1 : *)** ή **Many-to-one relationships (* : 1)** ,
- **One-to-one relationship (1 : 1)** και η
- **Many-to-Many relationship (Many : Many)**

One-to-many (1 : *) ή **Many-to-one relationships (* : 1):**

Αυτός είναι ο πιο κοινός τύπος σχέσεων. Στη μία πλευρά της σχέσης (one), η στήλη πρέπει να έχει μοναδικές τιμές , ενώ στην άλλη πλευρά (many) μπορεί ,και συνήθως έτσι συμβαίνει, να περιέχει διπλότυπες τιμές. Ουσιαστικά είτε One-to-many είτε Many-to-one σχέση , πρόκειται για την ίδια σχέση.

Για παράδειγμα στο **Διάγραμμα 7** η σχέση one-to-many μεταξύ των πινάκων Product και Sales είναι η ίδια με τη σχέση many-to-one μεταξύ των Sales και Product. Στο παράδειγμα αυτό οι δύο πίνακες έχουν κοινή στήλη την $Sales[ProductKey] = Product[ProductKey]$ και μέσω αυτής δημιουργείται η σχέση. Ο πίνακας *Product* έχει μοναδικές τιμές ,ενώ ο πίνακας Sales έχει διπλότυπες τιμές.

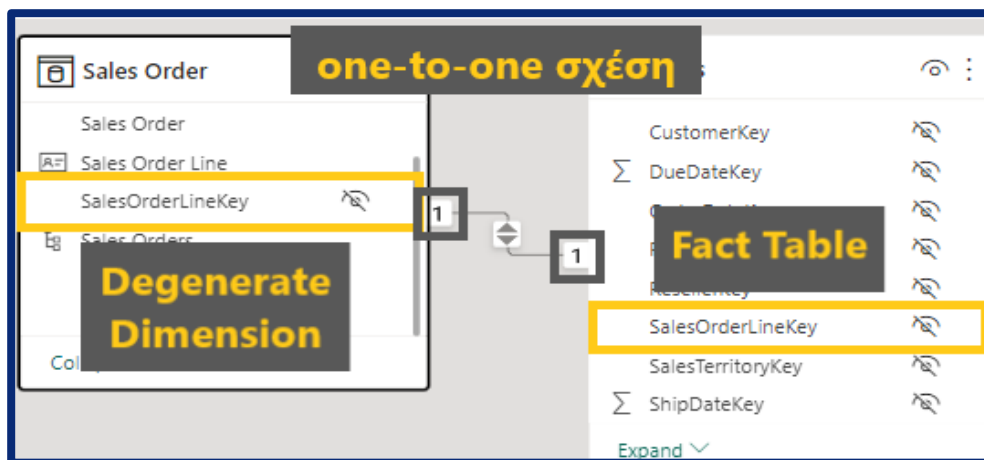


Διάγραμμα 8 : Many-to-one σχέση

One-to-one relationship (1 : 1)

Αυτός είναι ένας μάλλον ασυνήθιστος τύπος σχέσεων. Και στις δύο πλευρές της σχέσης οι στήλες πρέπει να έχουν μοναδικές τιμές, όμως οι τιμές του ενός πίνακα δεν βρίσκονται αναγκαστικά και στον άλλον πίνακα.

Αν και η γενική οδηγία από τους ειδικούς είναι να αποφεύγονται στον διαστασιακό σχεδιασμό. Στην πρώτη περίπτωση, δημιουργούνται όταν στήλες από τον βασικό πίνακα γεγονότων (**Fact Table**), μπορούν να ομαδοποιηθούν σε ξεχωριστό πίνακα-διάσταση ώστε, να επιτύχουμε μείωση του απασχολούμενου αποθηκευτικού χώρου, να αυξήσουμε την επίδοση σε ερωτήματα προς την βάση και για να απλοποιήσουμε τους υπολογισμούς. Αυτές οι στήλες ουσιαστικά εκφυλίζονται σε ένα δευτερογενή πίνακα που έχει προκύψει από τον βασικό πίνακα γεγονότων. Αυτοί οι πίνακες ονομάζονται εκφυλισμένες διαστάσεις (**Degenerate Dimensions**). Στο **Διάγραμμα 9** οι στήλες *Sales Order*[Sales Order] και *Sales Order* [Sales Order Line] δεν μπήκαν στον Πίνακα Sales. Μεταφέρθηκαν σε ξεχωριστή διάσταση (Sales Order) και ενώθηκαν με τον πίνακα γεγονότων με μέσω της στήλης κλειδί *Sales Order*[SalesOrderLineKey] = *Sales*[SalesOrderLineKey]. Σε αυτή την τεχνική, δημιουργούμε αυτό που ονομάζουμε παρένθετο κλειδί (**surrogate key**).

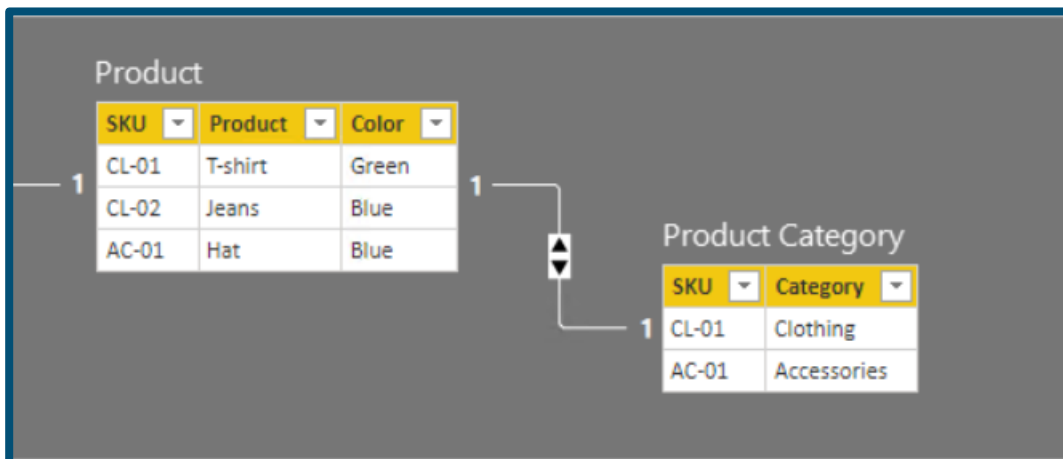


Διάγραμμα 9 : One-to-one σχέση σε εκφυλισμένη διάσταση

SalesOrderLineKey	Sales Order	Sales Order Line	Channel
53465001	SO53465	SO53465 - 01	Reseller
53465002	SO53465	SO53465 - 02	Reseller
53465003	SO53465	SO53465 - 03	Reseller
53465004	SO53465	SO53465 - 04	Reseller
53465005	SO53465	SO53465 - 05	Reseller
53465006	SO53465	SO53465 - 06	Reseller
53465007	SO53465	SO53465 - 07	Reseller
53465008	SO53465	SO53465 - 08	Reseller
53465009	SO53465	SO53465 - 09	Reseller
53465010	SO53465	SO53465 - 10	Reseller
53465011	SO53465	SO53465 - 11	Reseller
53465012	SO53465	SO53465 - 12	Reseller
53465013	SO53465	SO53465 - 13	Reseller
53465014	SO53465	SO53465 - 14	Reseller
53465015	SO53465	SO53465 - 15	Reseller

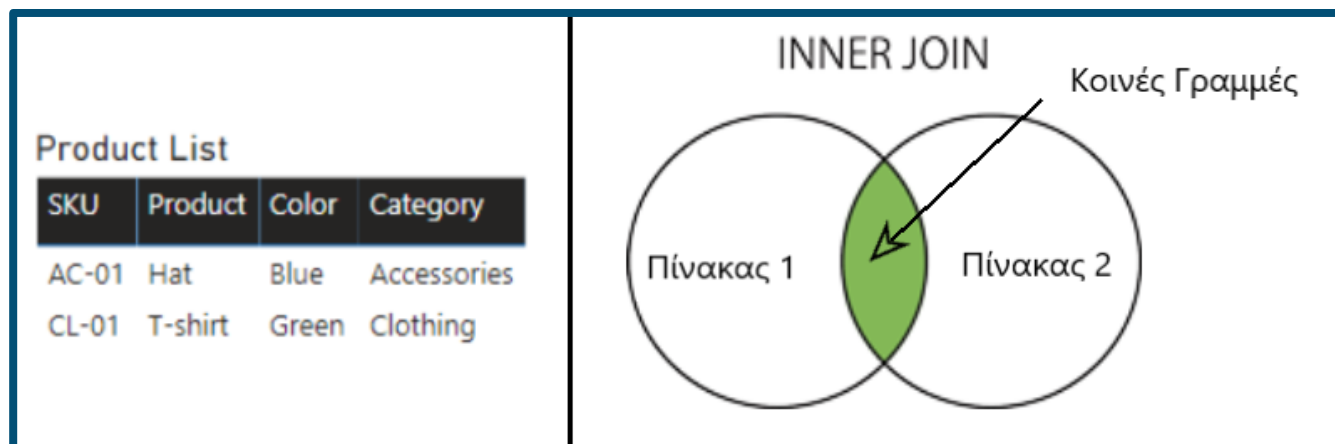
Εικόνα 15 : Εκφυλισμένη Διάσταση

Η δεύτερη περίπτωση που συναντάμε σχέσεις τύπου *one-to-one*, αν και γενικώς αντιπροτείνεται, συναντάται όταν μια μεμονωμένη επιχειρηματική οντότητα φορτώνεται ως δύο (ή περισσότεροι) πίνακες, πιθανώς επειδή **τα δεδομένα τους προέρχονται από διαφορετικά πηγή δεδομένων**. Αυτή η περίπτωση στην βιβλιογραφία ,ονομάζεται «Επέκταση των δεδομένων σε πολλαπλούς πίνακες» (“**Row data span across tables**”). Για παράδειγμα, θα μπορούσε δεδομένα του κύριου προϊόντος αποθηκεύονται σε ένα λειτουργικό σύστημα πωλήσεων και οι συμπληρωματικά δεδομένα του προϊόντος αποθηκεύονται σε διαφορετική πηγή.



Διάγραμμα 10 : One-to-one σχέση όπου τα δεδομένα έρχονται από διαφορετικά πηγαία συστήματα. Πηγή Εικόνας: [Πηγή Εικόνας](#)

Στο **Διάγραμμα 10** συναντάμε ένα τέτοιο παράδειγμα όπου, δύο επιχειρηματικές οντότητες έχουν εισαχθεί ξεχωριστά και ενώνονται μέσω της στήλης-κλειδί $Product[SKU] = Product\ Category[SKU]$. Αξίζει να σημειωθεί ότι είναι πιθανόν οι στήλες να μην περιέχουν όλες τις ίδιες μοναδικές τιμές. Οι πίνακες έχουν προέλθει από διαφορετικά πηγαία συστήματα και έχει δημιουργηθεί περιορισμένη σχέση (**limited relationship**) μεταξύ τους, γι' αυτό σε ένα πιθανό φιλτράρισμα μπορεί να εμφανίζονται τιμές ,που είναι αποκλειστικά κοινές και για τους δύο πίνακες. Η διαδικασία που καθορίζει αυτή την ιδιότητα ονομάζεται ,από την θεωρία Βάσεων Δεδομένων, **Συγχώνευση Πινάκων (Merging Data Tables)** και συγκεκριμένα συγχώνευση τύπου **Inner Join** .



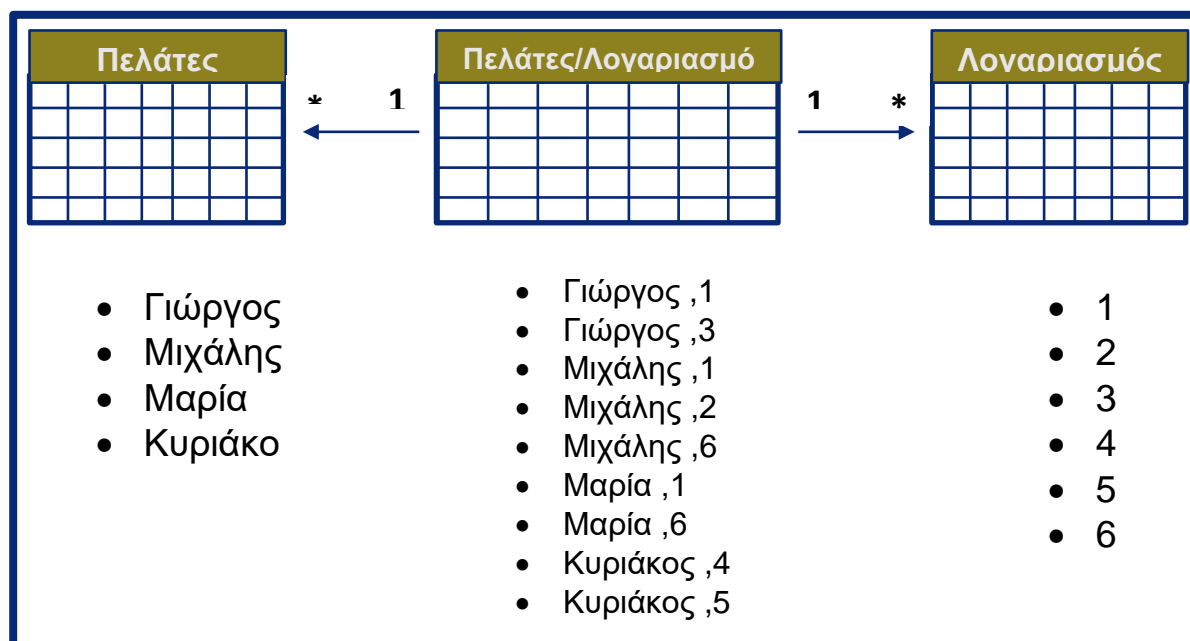
Εικόνα 16 : Inner Join – Διάγραμμα Venn [Πηγή Εικόνας](#)

Many-to-many relationship (* : *)

Και στις δύο πλευρές της σχέσης οι στήλες μπορεί να έχουν διπλότυπες τιμές. Αυτό το χαρακτηριστικό εισήχθη το 2018, πριν δεν γινόταν να δημιουργηθεί μια σχέση μεταξύ πινάκων που τουλάχιστον ο ένας από τους δύο

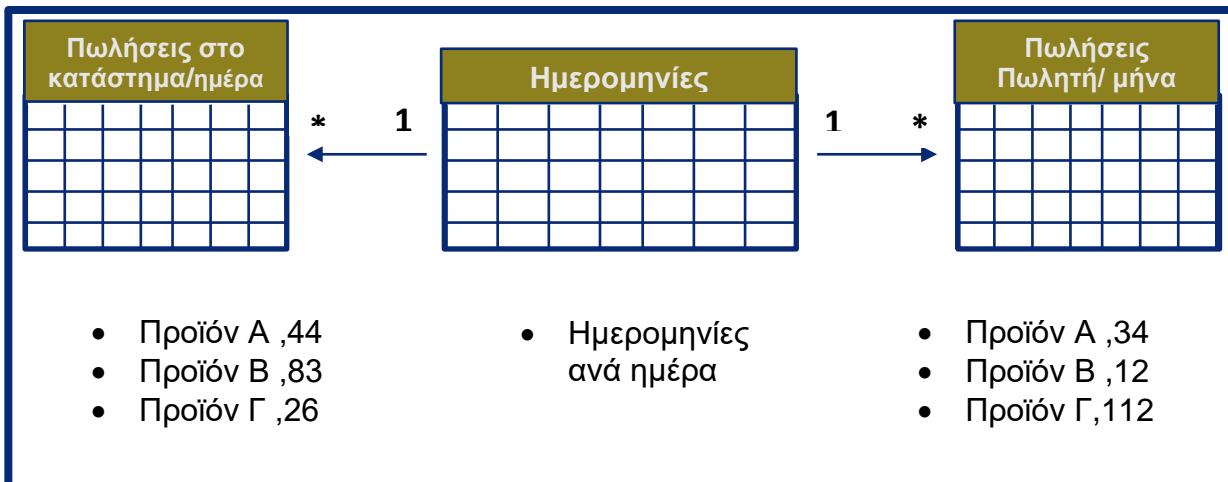
πίνακες είχε μοναδικές τιμές. Βέβαια, αυτό δεν συμβαίνει πάντα στην πράξη. Με αυτή την προσθήκη μπορούν να μην έχουν διπλότυπες τιμές. Δύο είναι οι συνήθεις περιπτώσεις που συναντάμε διπλότυπες τιμές και στους δύο συνδεόμενους πίνακες :

- **Συνδέοντας Πίνακες - Διαστάσεις** : Ένα κλασσικό παράδειγμα που συναντάμε είναι η περίπτωση των τραπεζικών λογαριασμών. Ένας πελάτης μπορεί να έχει περισσότερους από έναν λογαριασμούς στο όνομα του, ενώ την ίδια στιγμή, οι τραπεζικοί λογαριασμοί μπορούν να έχουν περισσότερο από έναν διαχειριστή, όπως συμβαίνει στην περίπτωση των κοινών λογαριασμών. Αυτό το πρόβλημα για να επιλυθεί παλαιότερα οι δημιουργοί των μοντέλων έβαζαν ένα πίνακα ενδιάμεσο με μοναδικές τιμές, τον οποίο ονόμαζαν Πίνακα Γέφυρα (**Bridge Table**), και συνέδεαν σε αυτόν τους δύο πίνακες με τις διπλότυπες τιμές με δύο σχέσεις τύπου Many-to-One (* - 1).

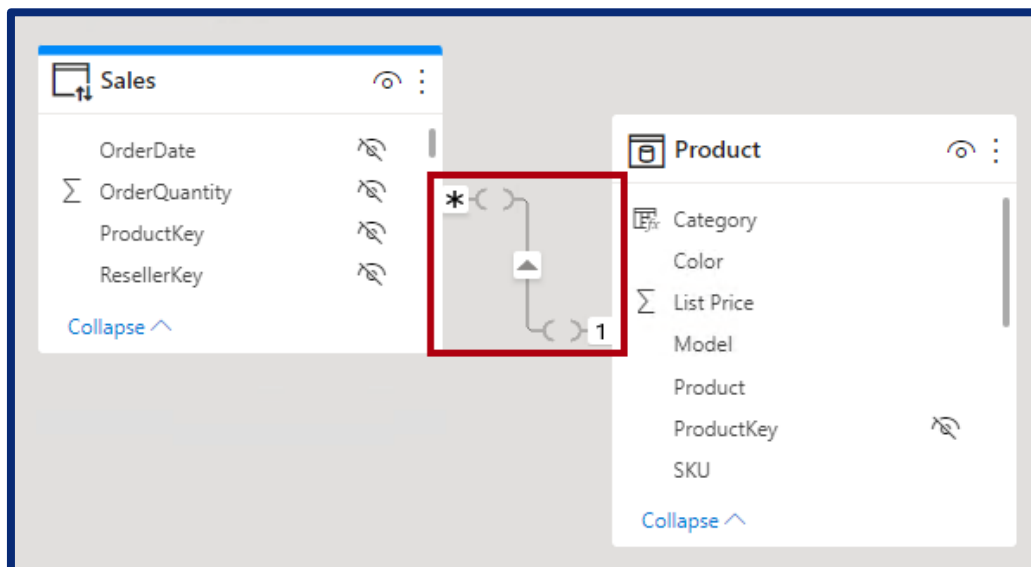


Σχήμα 12 : Παράδειγμα επίλυσης προβλήματος σύνδεσης πινάκων Διαστάσεων

- **Συνδέοντας πίνακες με διαφορετικό επίπεδο αποθήκευσης πληροφορίας (granularity):** Συχνά βλέπουμε να υπάρχουν δεδομένα σε πίνακες στους οποίους είναι αποθηκευμένες πληροφορίες σε διαφορετικό επίπεδο αποθήκευσης της πληροφορίας. Για παράδειγμα, στον ένα πίνακα του μοντέλου να αποθηκεύονται οι **πωλήσεις των προϊόντων στο φυσικό κατάστημα να γίνονται ανά ημέρα** και σε ένα άλλο πίνακα να αποθηκεύονται, **οι πωλήσεις που κάνει ένας πωλητής σε επίπεδο μήνα**. Αυτοί δύο πίνακες παλαιότερα θα συνδέονταν μέσω ενός πίνακα που περιέχει ημερομηνίες στο χαμηλότερο επιθυμητό επίπεδο.



Σχήμα 13 : Σύνδεση πινάκων με διαφορετικό επίπεδο αποθήκευσης πληροφορίας



Εικόνα 17 : Παράδειγμα σύνδεσης πινάκων με σχέση τύπου Many-to-many – Model View [Πηγή Εικόνας](#)

Πλέον η σύνδεση των πινάκων ,μπορεί να γίνει απευθείας με την σχέση **Many-to-Many (*-*)**. Οι σχέσεις αυτές παρ' όλ' αυτά καλό είναι να αποφεύγονται από τους χρήστες καθώς, μπορεί να δημιουργηθούν ασάφειες στο μοντέλο και να παράγει λάθος δεδομένα , αν δεν γίνουν με σωστό τρόπο. Αυτές οι σχέσεις είναι πάντα **περιορισμένες** ή **limited**. Αυτό σημαίνει ότι η μηχανή του Power BI δεν μπορεί να επιβεβαιώσει ότι υπάρχουν μοναδικές τιμές και στους δύο πίνακες που συνδέονται και γι'αυτό όταν γίνεται κάποιο query ,αν δεν υπάρχει και στις δύο πλευρές κάποια τιμή ,απλά δεν προσμετράται στο αποτέλεσμα.

Διαφορετικά ,όπως αναφέρεται στην βιβλιογραφία δεν μπορεί να διαπιστωθεί η αναφορική ακεραιότητα των πινάκων (**referential integrity**).

Συμπερασματικά , ορισμένες τιμές μπορεί να χαθούν στο τελικό αποτέλεσμα και άρα να έχουμε λανθασμένο αποτέλεσμα.

Κατευθύνσεις των σχέσεων

Ένα εξαιρετικό σημαντικό στοιχείο ενός μοντέλου δεδομένων είναι οι κατευθύνσεις των σχέσεων .Αυτές δείχνουν την κατεύθυνση της προς την οποία μεταδίδεται η εφαρμογή ενός φίλτρου σε ένα μοντέλο δεδομένων ,όταν γίνεται μια ερώτηση στην βάση δεδομένων, το querying δηλαδή. Υπάρχουν δύο τύποι κατευθύνσεων :

- **Μονή κατεύθυνση (Single)** : Το φίλτρο διαδίδεται πάντα προς τη μία κατεύθυνση της σχέσης και όχι το αντίστροφο. Σε μια σχέση ένα One-to-Many, η κατεύθυνση είναι πάντα από τη One πλευρά της σχέσης προς την Many πλευρά. Αυτή είναι η τυπική και πιο επιθυμητή συμπεριφορά.[13]
- **Αμφίπλευρη κατεύθυνση (Both)**: Το φίλτρο διαδίδεται και στις δύο κατευθύνσεις της σχέσης. Αυτό ονομάζεται επίσης αμφίδρομο διασταυρούμενο φίλτρο (**bidirectional cross-filter**). Σε μια σχέση ένα προς την πλευρά Many, το φίλτρο εξακολουθεί να διατηρεί το χαρακτηριστικό του να διαδίδεται από τη μία πλευρά στην πολύπλευρη, αλλά διαδίδεται επίσης από την Many πλευρά στην One πλευρά. [13]

Συγκεντρώνουμε όλες τις περιπτώσεις των σχέσεων σε ένα μοντέλο δεδομένων.

Τύπος σχέσης	Κατεύθυνση σχέσης	Μετάδοση Φίλτρου	Περιορισμένη ή Κανονική
One-to-One	Μονή	Από την πλευρά One προς την πλευρά Many	Περιορισμένη αν είναι σε composite μοντέλο ,αλλιώς κανονική

One-to-One	Αμφίπλευρη	Αμφίπλευρη	Περιορισμένη αν είναι σε composite μοντέλο ,αλλιώς κανονική
One-to-One	Αμφίπλευρη	Αμφίπλευρη	Περιορισμένη αν είναι σε composite μοντέλο ,αλλιώς κανονική
Many-to-One	Μονή	Επιλέγουμε πίνακα πηγή	Πάντα περιορισμένη
Many-to-One	Αμφίπλευρη	Αμφίπλευρη	Πάντα Περιορισμένη

Πίνακας 7 :Σχέσεις μοντέλου δεδομένων

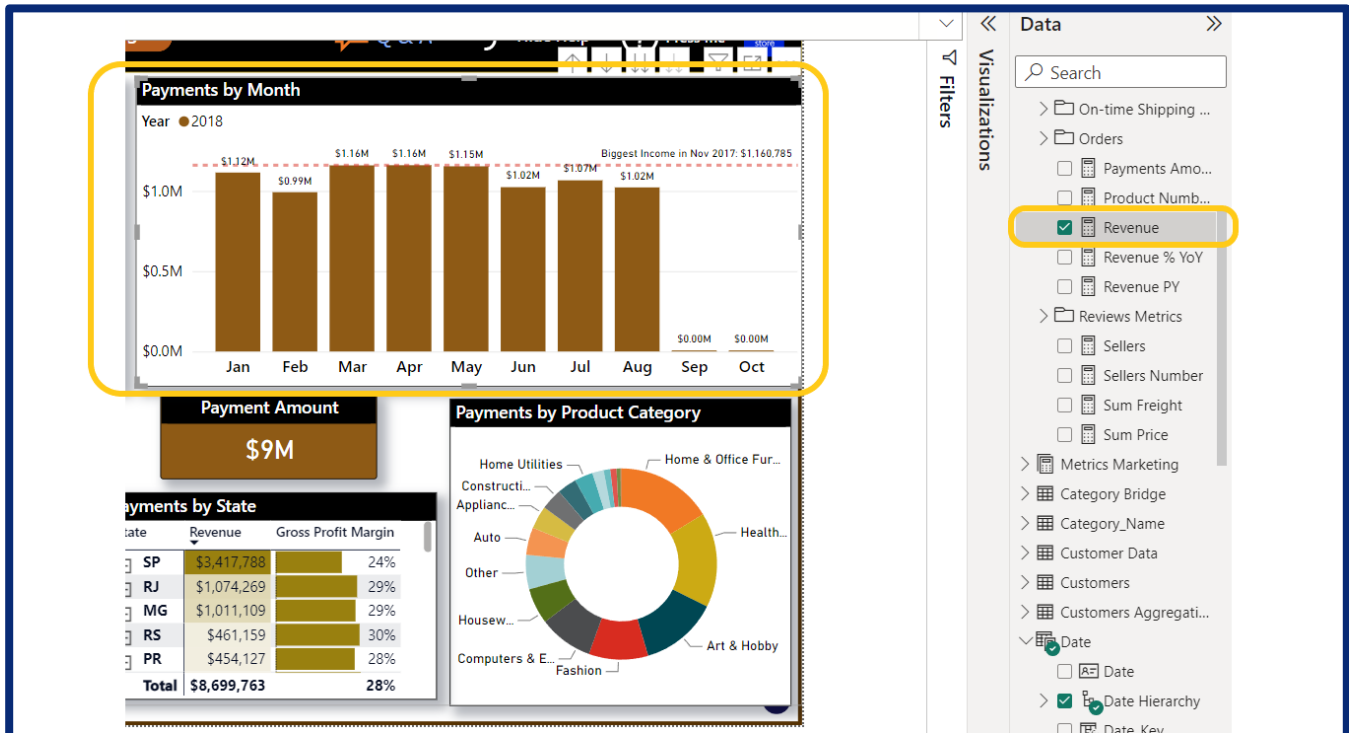
3.4.4 Μετρήσεις - Measures

Οι **Μετρήσεις** ή **Μέτρα** (στα αγγλικά **Measures**) , αποτελούν το τελευταίο κομμάτι στην μοντελοποίηση .Αποτελούν ένα είδος Υπολογισμών βασιζόμενοι στις στήλες και τους πίνακες του μοντέλου. Τα μέτρα χρησιμοποιούνται σε μερικές από τις πιο κοινές αναλύσεις δεδομένων. Απλές συνοψίσεις, όπως αθροίσματα, μέσοι όροι, ελάχιστο, μέγιστο και μετρήσεις μπορούν να οριστούν μέσα από τα πεδία των πινάκων. **Τα υπολογισμένα αποτελέσματα των μέτρων αλλάζουν πάντα ανάλογα με την αλληλεπίδρασή με τις αναφορές**, επιτρέποντας γρήγορη και δυναμική ad-hoc εξερεύνηση δεδομένων. Είναι ο πιο εύκολος τρόπος για να πραγματοποιηθούν φιλτραρίσματα στον πίνακα γεγονότων του μοντέλου.

Το περιεχόμενο των μέτρων συντάσσεται με την γλώσσα του Power BI – **DAX** (Data Analysis Expressions). Η DAX περιλαμβάνει μια βιβλιοθήκη με περισσότερες από 200 συναρτήσεις, τελεστές και περιοριστές. Ακολουθεί ένα παράδειγμα υπολογισμού των συνολικών εσόδων μιας επιχείρησης.Συντάχθηκε ο εξής κώδικας για τον υπολογισμό του **μέτρου** :

Revenue = `SUM(Order_Payments[payment_value])`

Χρησιμοποιήσαμε από τον πίνακα Order_Payments, το πεδίο Order_Payments[payment_value] , όπου αθροίσαμε το συνολικό ποσό όλων των κελιών της στήλης .



Εικόνα 18 : Παράδειγμα εφαρμογής Μέτρησης – Report View

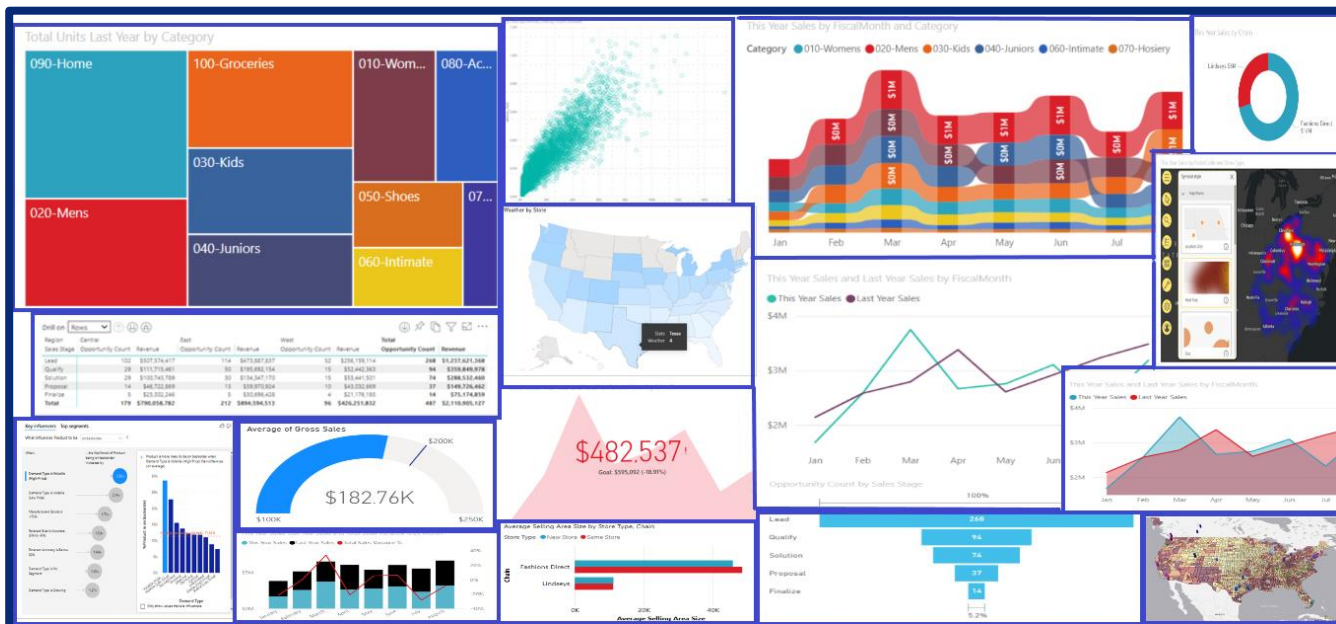
Με βάση αυτό το μέτρο, υπολογίζονται και εμφανίζονται στις οπτικοποιήσεις το ποσό των συνολικών εσόδων, ανά μήνα του έτους αλλά και το συνολικό. Το αποτέλεσμα έχει φιλταριστεί, ώστε να φαίνονται οι τιμές για μόνο το έτος 2018. Αν επιλέξουμε και άλλη χρονιά, το συνολικό ποσό που υπολογίζεται από το μέτρο αλλάζει.

Εδώ φαίνεται και η χρησιμότητα των μέτρων καθώς αλλάζουν τον υπολογισμό που πραγματοποιούν, ανάλογα το επιβαλλόμενο φίλτρο, ενώ επίσης μπορούν να επαναχρησιμοποιηθούν και μέσα στην σύνταξη κάποιου άλλου μέτρου, για τον υπολογισμό κάποιου πολύπλοκου μέτρου.

3.5 Power BI Desktop - Οπτικοποίηση Δεδομένων

Οι οπτικοποιήσεις στο Power BI, αποτελούν την λωρίδα του μοντέλου που μπορεί ο χρήστης να αλληλεπιδράσει και να αντλήσει συμπεράσματα για την ανάλυση που πραγματοποιεί. Είναι το μέσο που απογειώνει την εμπειρία του χρήστη, καθώς μέσω αυτών γίνεται η απεικόνιση των δεδομένων του μοντέλου. Το Power BI διαθέτει μια ευρεία γκάμα οπτικοποιήσεων όπως: Διαγράμματα γραμμών, Χάρτη ArcGIS, Azure Map, Shape Map, Πίνακες, διαγράμματα Πίτας και τύπου Ντόνατ, διαγράμματα Κορδέλας, απλά διαγράμματα

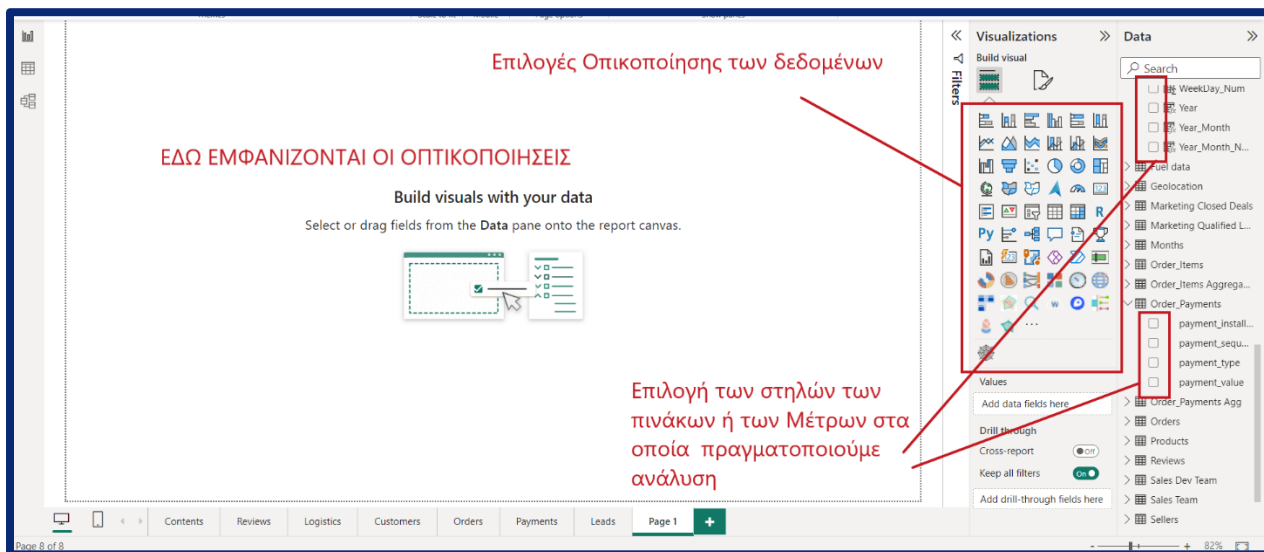
Σηλών, διαγράμματα Διασποράς, τύπου Ιστού , για την μέτρηση KPIs, τύπου Χωνιού, Καρτέλες και πολλά άλλα που μπορεί ο χρήστης να κατεβάσει από μια πλατφόρμα όπου είναι ανεβασμένα από χρήστες και εταιρείες , τα περισσότερα από τα οποία.



Εικόνα 19 : Ενδεικτικές οπτικοποιήσεις του Power BI

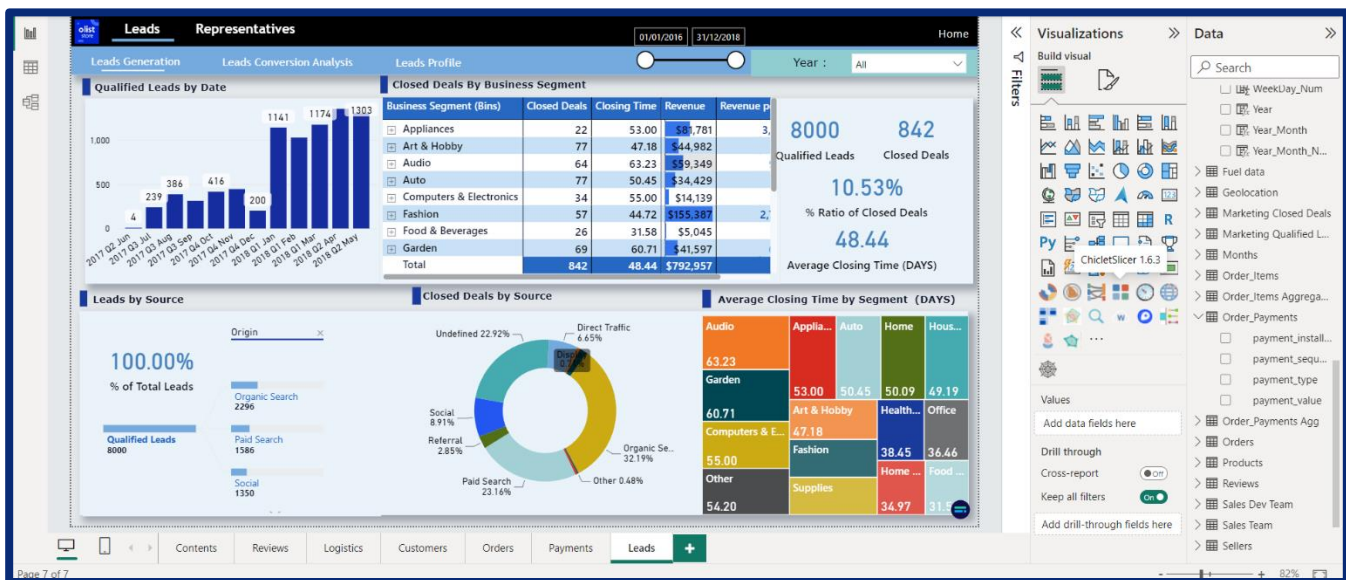
Το πιο σημαντικό χαρακτηριστικό το οποίο έχουν, εξαιτίας του μοντέλου δεδομένων που τρέχει στο παρασκήνιο, είναι η δυνατότητα που έχουν να αλλάζουν το περιεχόμενο της οπτικοποίησης , αναλόγως το εφαρμοζόμενο φίλτρο σε αυτά ή στην σελίδα ή σε ολόκληρη την αναφορά. Αυτό το χαρακτηριστικό επιτρέπει την άμεση απεικόνιση των ερωτημάτων στην βάση ,με τον πιο καταληπτό τρόπο.

Οι οπτικοποιήσεις προστίθενται στις αναφορές από την καρτέλα του **Power BI Desktop , Report View**.



Εικόνα 20 : Το πάνελ οπτικοποίησης - Report View

Στην επόμενη Εικόνα φαίνεται παράδειγμα αναφοράς στο Power BI Desktop κατά την διάρκεια δημιουργίας της αναφοράς.

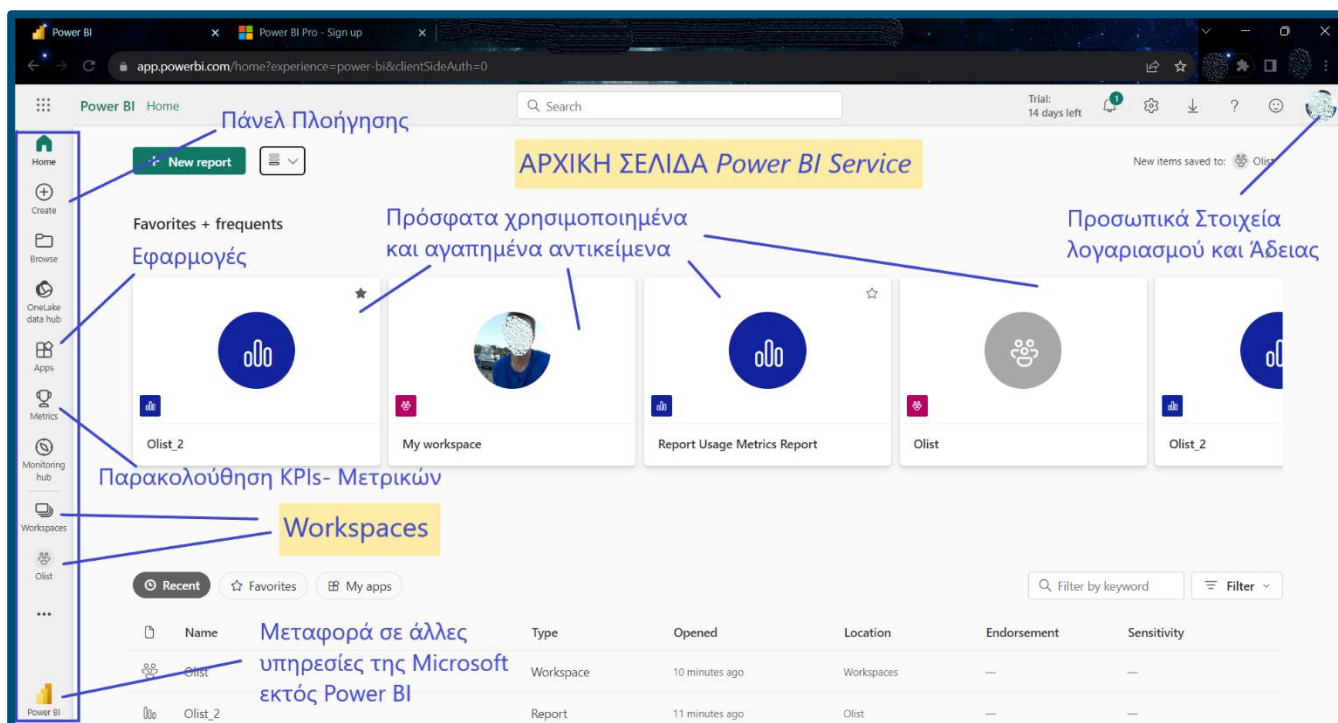


Εικόνα 21 : Παράδειγμα δημιουργίας αναφοράς - Report View

3.6 Power BI Service - Δημοσίευση και Συνεργασία

Το **Power BI Service** είναι η κύρια πλατφόρμα της *Microsoft* για φιλοξενία και κοινή χρήση αναφορών Power BI. Είναι ένα υπερσύγχρονο *cloud-based* λογισμικό που επιτρέπει να μοιραζόμαστε τις οπτικοποιήσεις των δεδομένων μας, τις αναφορές και άλλα στοιχεία με

την εταιρεία μας, το σχολείο ή ολόκληρο τον κόσμο. Το **Power BI Service** εκτελείται σε υποδομή της Microsoft. Αυτό σημαίνει ότι όλα συμβαίνουν σε μια τεράστια γκάμα διακομιστών που κατέχει η Microsoft και λειτουργεί κάπου απομακρυσμένα στον κόσμο. Η υπηρεσία του **PBI Service** είναι και αυτή που ολοκληρώνει ένα σύστημα επιχειρηματικής ευφυΐας. **Χωρίς μια τέτοια υποδομή, ένα σύστημα self-service BI δεν θα μπορούσε να έχει την ολοκληρωμένη μορφή που επιθυμούμε από τέτοια συστήματα**, όπου περιλαμβάνουν την συνεργασία χρηστών ,καταναλωτών του οργανισμού ή ομάδων εργαζομένων, τον άμεσο συγχρονισμό των δεδομένων ,την ζωντανή παρακολούθηση KPIs , την χρήση κοινών μοντέλων δεδομένων και άλλες λειτουργίες. Θα επρόκειτο για μια στατική διαδικασία ,μακριά από την ζωντανή και συνεχή αναβάθμιση και αλληλεπίδραση .Το **PBI Service** εκτός των δυνατοτήτων για δημοσίευση και συνεργασία ,παρέχει την δυνατότητα να δημιουργίας αναφορών και μοντέλων δεδομένων (**Datasets**) ,χωρίς την χρήση του **Power BI Desktop**. Ασφαλώς δεν διαθέτει όλες τις δυνατότητες του μοντελοποίησης ,οπτικοποιήσεων και λειτουργιών ,αλλά οπωσδήποτε είναι αρκετό για απλές λύσεις που μπορεί να χρειάζεται κάποιος σχεδιαστής.



Εικόνα 22 : Αρχική σελίδα Power BI Service στον φυλλομετρητή – Power BI Service

3.6.1 Άδειες Χρήσης και δυνατότητες

Για την χρήση του **PBI Service**, υπάρχουν διάφορες άδειες χρήσης που μπορεί κάποιο άτομο ή οργανισμός να αγοράσει ,αναλόγως τις λειτουργίες που επιθυμεί να έχει.

Οι άδειες χρήσης είναι οι παρακάτω :

- **Power BI Free**
- **Power BI Pro**
- **Power BI Premium ανά χρήστη**
- **Power BI Premium ανά χωρητικότητα** (για οργανισμούς)

Ο δωρεάν λογαριασμός (**Power BI Free**) δεν απαιτεί κάποια χρέωση ,όπως προκύπτει και από το όνομά του, αλλά έχει περιορισμένες δυνατότητες. Το μόνο που χρειάζεται είναι μόνο το e-mail του οργανισμού στο οποίο ανήκει ο χρήστης. Σε αυτόν τον τύπο άδειας, ο σχεδιαστής δεν μπορεί να μοιραστεί αντικείμενα σε κοινόχρηστους χώρους εργασίας (**Shared Workspaces**) και έχει μόνο πρόσβαση στο προσωπικό χώρο εργασίας (**Personal Workspace**).

Οι άδειες χρήσης **Power BI Pro** είναι ατομικές άδειες χρήσης. Οι άδειες Pro προορίζονται για τους ανθρώπους στον οργανισμό που δημιουργούν αναφορές και άλλο περιεχόμενο εντός του **Power BI Service**. Επιτρέπουν στους χρήστες να έχουν πρόσβαση σε κοινόχρηστους χώρους εργασίας ,όπου εκεί μπορούν να συνεργαστούν σε πρότζεκτ του Power BI με άλλα άτομα στον οργανισμό . Σε αυτή την άδεια οι σχεδιαστές μπορούν να δημοσιεύουν εφαρμογές Power BI (**Apps**) , οι οποίες χρησιμεύουν στο να επικοινωνεί ο σχεδιαστής το περιεχόμενο που δημιουργεί σε μεγαλύτερα κοινά . Οι άδειες αυτές είναι κατάλληλες για προσωπική χρήση και για μικρές ομάδες αναλυτών.

Οι άδειες τύπου **Premium ανά χωρητικότητα**, είναι αυτές που αναδεικνύουν πραγματικά τις δυνατότητες του **Power BI Service**. Πρόκειται βασικά για εταιρικές άδειες , τις οποίες αγοράζει η εταιρεία ή ο οργανισμός και καλύπτει όλη την εταιρεία. Η Microsoft δίνει στον οργανισμό αποκλειστική χωρητικότητα στην υποδομή του Power BI. Αυτό σημαίνει ότι όλοι οι χρήστες στον οργανισμό σας μπορούν να έχουν πρόσβαση σε αναφορές και πίνακες ελέγχου (**Dashboards**) στο **Power BI Service** ακόμα κι αν δεν έχουν άδεια **Pro**. Αυτό λοιπόν θα κάλυπτε τους εργαζόμενους που δεν δημιουργούν αναφορές στο Power BI, αλλά πρέπει να μπορούν να βλέπουν και αλληλεπιδρούν με τις αναφορές. Οι δυνατότητες , που πλέον είναι διαθέσιμες στην συγκεκριμένη άδεια , είναι εμπλουτισμένες και περιλαμβάνουν πιο συχνές ανανεώσεις των δεδομένων (**Scheduled Refreshes**), περισσότερο cloud- based χώρο αποθήκευσης , υπηρεσίες μηχανικής μάθησης και εργαλεία τεχνητής νοημοσύνης μέσα στο **Microsoft Azure**.

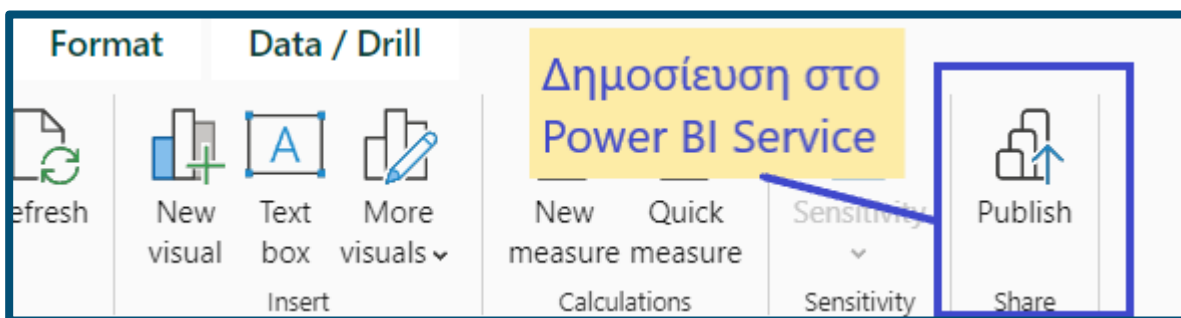
Τέλος, οι άδειες **Premium ανά χρήστη**, είναι προσωπικές άδειες , όπως και οι **Pro** , αλλά διαθέτουν δυνατότητες **Premium**. Οι άδειες αυτές έχουν ένα σημαντικό περιορισμό , καθώς το περιεχόμενο που δημιουργούν οι σχεδιαστές με αυτή την άδεια, μπορεί να προβληθεί μόνο από χρήστες που επίσης έχουν **Premium** άδεια.

3.6.2 Δημοσίευση στο *Power BI Service*

Όταν έχουμε ολοκληρώσει την αναφορά που έχουμε δημιουργήσει στο **Power BI Desktop**, είναι η ώρα να το δημοσιεύσουμε στο **Power BI Service**, ώστε να το δουν οι συνεργάτες μας ή το κοινό που επιθυμούμε. Ακολουθούν ορισμένοι βασικοί ορισμοί για αντικείμενα και λειτουργίες του **Power BI Service**.

Χώροι Εργασίας – Workspaces

Όταν προσπαθήσουμε να κάνουμε την δημοσίευση στο **Power BI** θα μας ζητήσει να επιλέξουμε το χώρο εργασίας (**Workspace**) όπου θέλουμε να το δημοσιεύσουμε. Οι χώροι εργασίας είναι το πρωταρχικό σημείο όπου φορτώνονται οι αναφορές που δημιουργούμε.



Εικόνα 23 : Δημοσίευση στο Power BI Service – Power BI Desktop

Υπάρχουν δύο ειδών χώροι εργασίας :

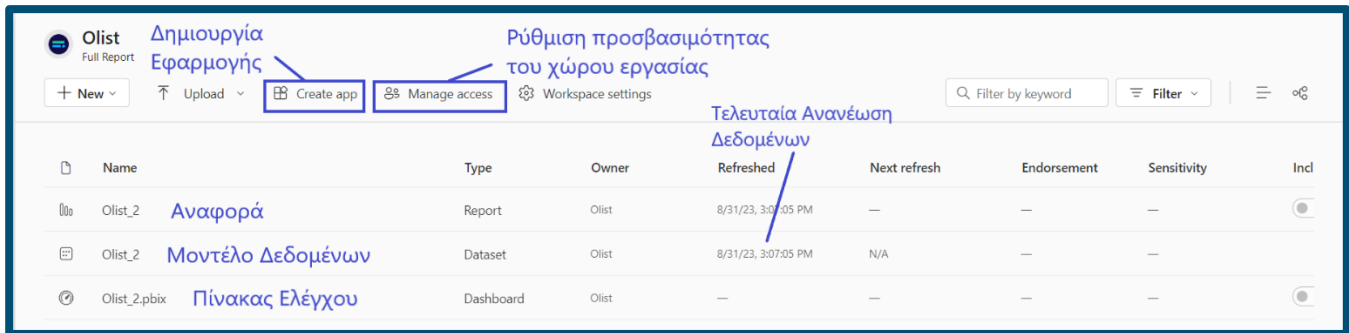
- Ο προσωπικός χώρος εργασίας (**Personal Workspace**)
- Οι κοινόχρηστοι χώροι εργασίας (**Shared Workspaces**)

Στον προσωπικό χώρο εργασίας έχει πρόσβαση μόνο ο εκάστοτε σχεδιαστής ή χρήστης, ενώ στους κοινόχρηστους μπορεί να έχει πρόσβαση, όποιος ορίσει ο σχεδιαστής.

Οι χώρους εργασίας χρησιμεύουν :

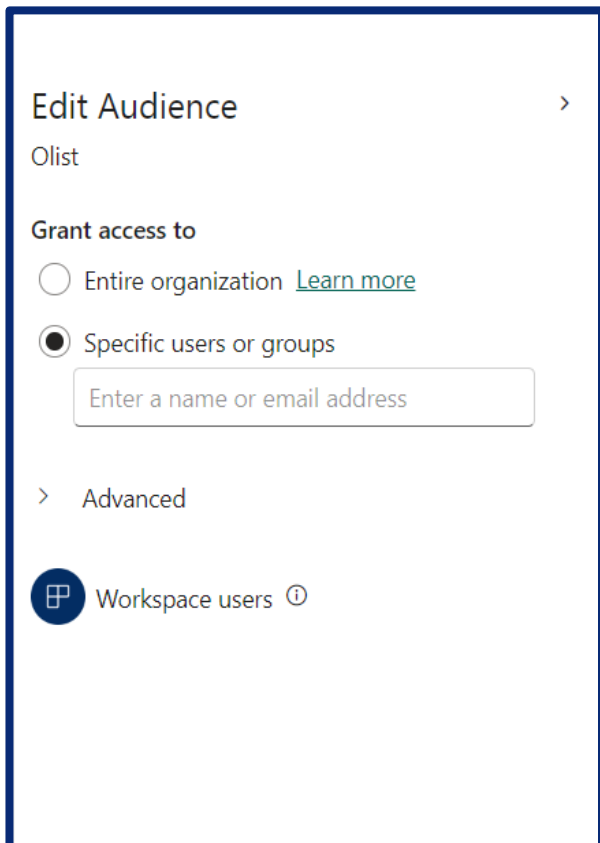
- Στον ανέβασμα μοντέλων δεδομένων (Datasets) και αναφορών (Reports)
- δημιουργία Πινάκων Ελέγχου (*Dashboards*)
- Στην συνεργασία με συναδέλφους
- Για την δημιουργία **Dataflows** και

- Για την δημιουργία **Εφαρμογών Power BI**

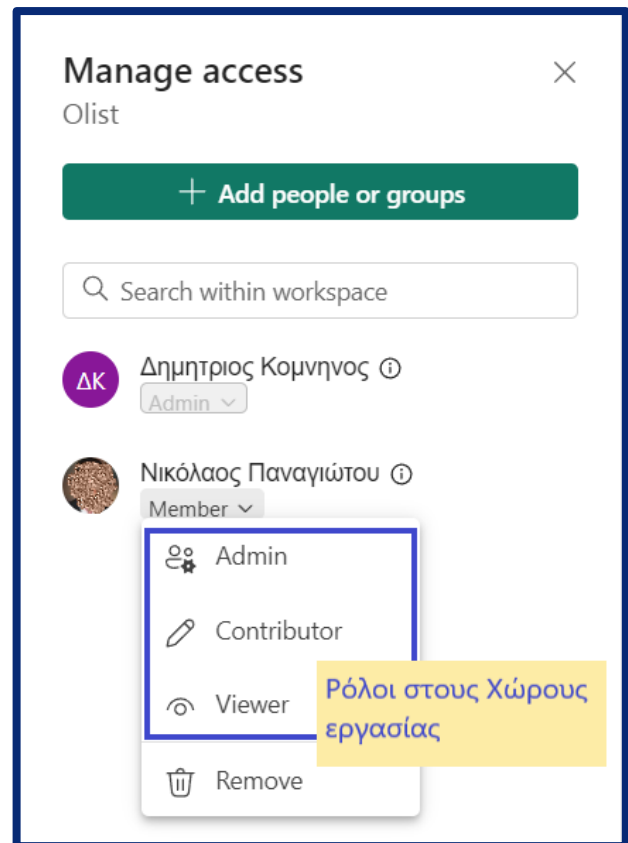


Εικόνα 24 : Παράδειγμα Χώρου Εργασίας – Power BI Service

Με ευκολία μπορούμε να ρυθμίσουμε το κοινό που επιτρέπουμε να δει ή και να επεξεργαστεί τον χώρο εργασίας που δημιουργούμε .Μπορούμε να δώσουμε πρόσβαση σε ολόκληρο τον οργανισμό ή κάποια ομάδες ή μεμονωμένα άτομα του οργανισμού.



Εικόνα 25 : Ρύθμιση Προσβασιμότητας στον χώρο εργασίας – Power BI Service



Εικόνα 26 : Ανάθεση Ρόλων στους χώρους εργασίας – Power BI Service

Ρόλοι Χώρων Εργασίας – Workspaces Roles

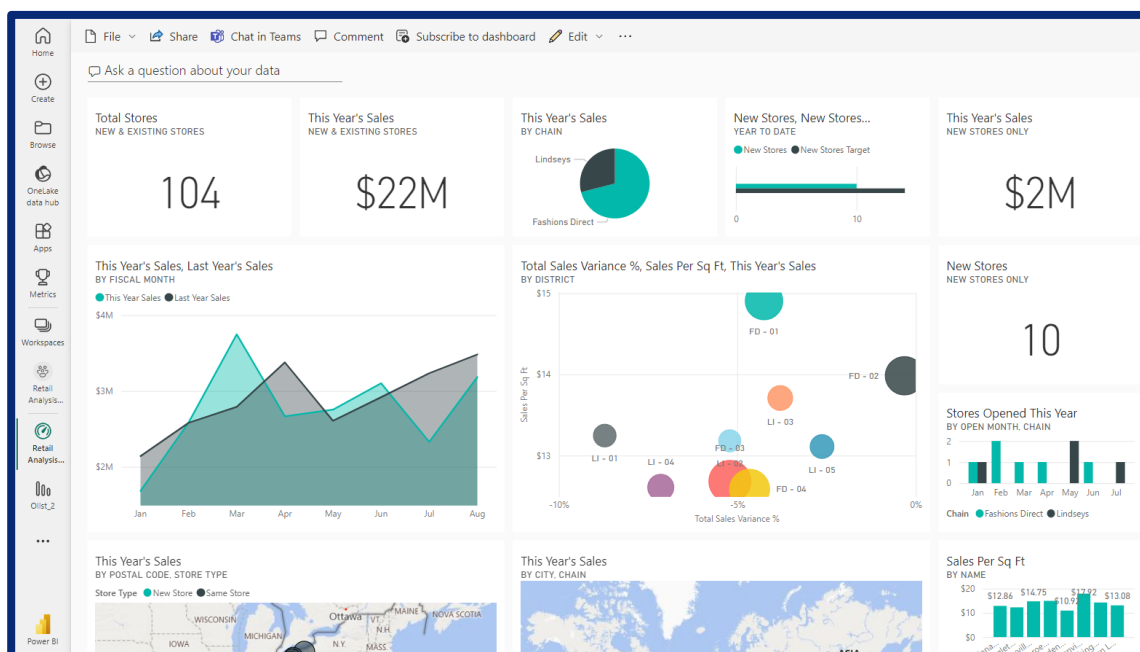
Στους χώρους εργασίας μπορούν να ανατεθούν ρόλοι από τον δημιουργό του χώρου εργασίας για τους χρήστες τους οποίους επιθυμεί να αλληλεπιδρούν με το περιεχόμενο του. Μια λειτουργία πολύ χρήσιμη **για να διατηρείται το περιεχόμενο οργανωμένο και αξιοποιήσιμο**, για παράδειγμα από χρήστες κάποιας υποομάδας ενός τμήματος μιας επιχείρησης. Υπάρχουν 4 βασικοί ρόλοι που μπορούν να έχουν οι χρήστες και είναι οι εξής:

- **Viewer** (Θεατής) :
 - Μπορεί να δει και να αλληλεπιδράσει με αναφορές και πίνακες ελέγχου
- **Contributor** (Συνεισφέρων) :
 - Μπορεί να δει και να αλληλεπιδράσει με αναφορές και πίνακες ελέγχου
 - Μπορεί να δημιουργήσει και επεξεργαστεί περιεχόμενο μέσα στους χώρους εργασίας και στα μοντέλα δεδομένων (**Datasets**)
 - Μπορεί να χρησιμοποιήσει κοινά μοντέλα δεδομένων (**shared datasets**)
- **Member** (Μέλος) :
 - Μπορεί να κάνει ότι και ο **Contributor**
 - Μπορεί να επεξεργαστεί την προσβασιμότητα στο χώρο εργασίας και την ασφάλεια
 - Μπορεί να δημοσιεύσει **Εφαρμογή Power BI**
- **Admin** (Διαχειριστής) :
 - Μπορεί να κάνει τα πάντα
 - Μπορεί να προσθέσει & να επεξεργασθεί περιεχόμενο
 - Μπορεί να προσθέσει ή να αφαιρέσει άλλους Διαχειριστές
 - Μπορεί να διαγράψει το Χώρο Εργασίας

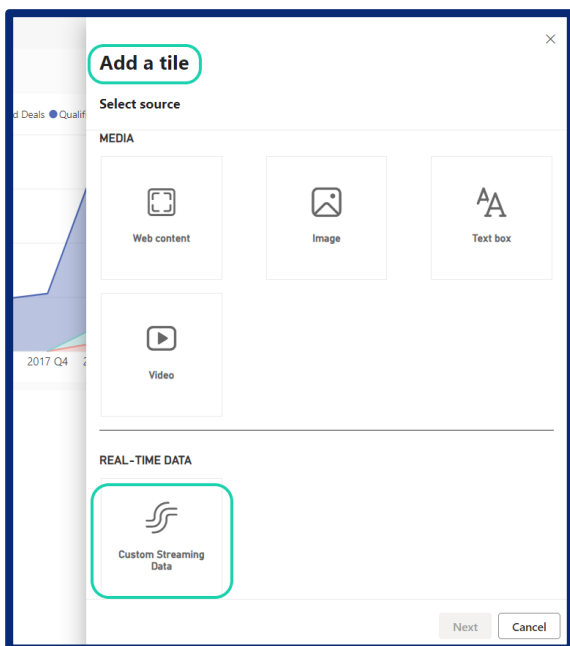
Πίνακες ελέγχου - Dashboards

Οι πίνακες ελέγχου ή **Dashboards**, αποτελούν ένα ακόμα τρόπο απεικόνισης των δεδομένων μας πέραν των αναφορών στο *PBI Service*. Μπορούν να δημιουργηθούν αποκλειστικά στο *PBI Service* και περιέχουν συγκεκριμένες οπτικοποιήσεις από αναφορές ή ολόκληρες αναφορές που έχουμε φτιάξει και έχουμε ανεβάσει σε κάποιο Χώρο Εργασίας. Αυτό το οποίο προσφέρουν στην εμπειρία του χρήστη είναι ότι, **μπορούν να έχουν περιεχόμενο από διάφορες αναφορές** με σκοπό να έχει ο χρήστης μια πιο γενική εικόνα. Επίσης, μόνο στους Πίνακες ελέγχου υπάρχει η δυνατότητα να προστεθούν Έκτακτες Ειδοποιήσεις (**Alert Notifications**) για τυχόν κρίσιμα στοιχεία που άλλαξαν στα δεδομένα και επιθυμούμε το *PBI Service* να μας ειδοποιήσει για αυτό. Ένα παράδειγμα, θα μπορούσε να είναι, να αλλάξει η τιμή είναι ένα σημαντικός δείκτης KPI, ο οποίος άλλαξε μετά από κάποια ανανέωση του μοντέλου δεδομένων. Ακόμη, στον πίνακα ελέγχου μπορούν να προστεθούν εκτός από οπτικοποιήσεις των αναφορών, περιεχόμενο που από το διαδίκτυο είτε κάποια **ζωντανά δεδομένα (Real - Time Data)** για την μετάδοσή τους.

Στην εικόνα που ακολουθεί, παρατηρούμε διάφορες οπτικοποιήσεις, οι οποίες ανήκουν σε διαφορετικές αναφορές και έχουν συγκεντρωθεί στον εμφανιζόμενο πίνακα ελέγχου. Αν πατήσει ο χρήστης πάνω σε κάποια οπτικοποίηση, μεταφέρεται αυτόματα σε αυτή την αναφορά για να δει περισσότερες λεπτομέρειες.



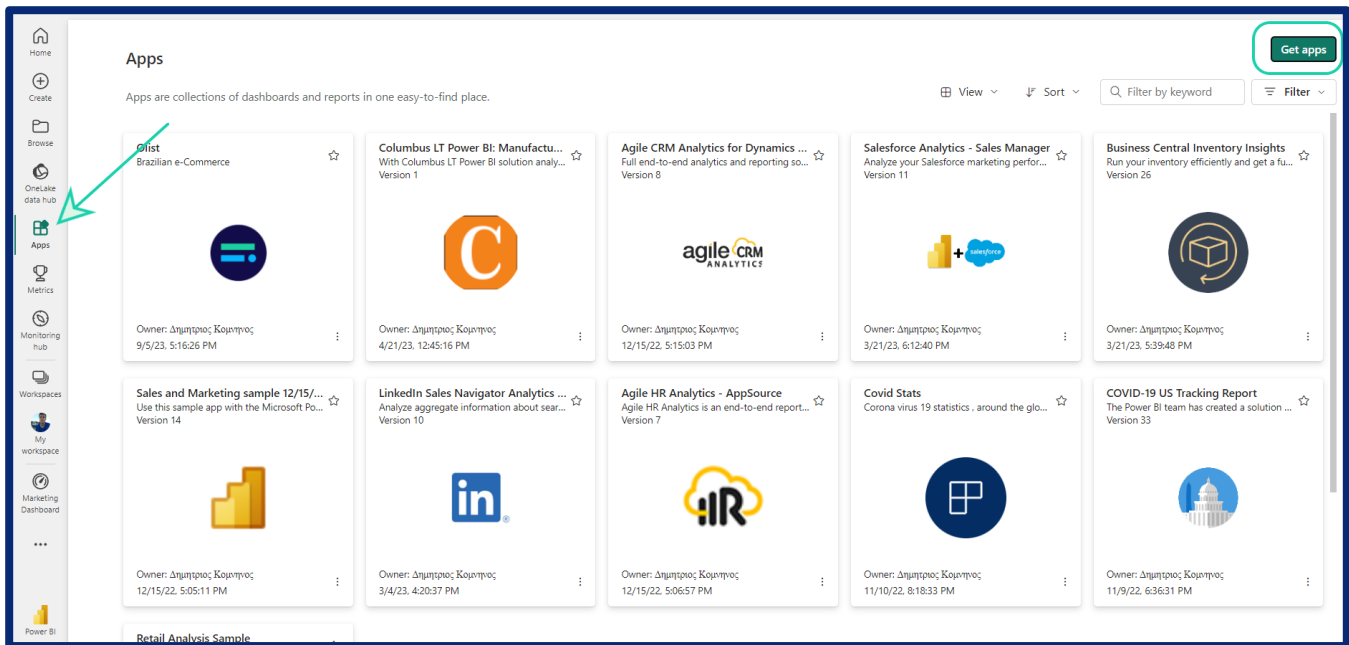
Εικόνα 27 : Παράδειγμα Πίνακα ελέγχου (Dashboard) – Power BI Service



Εικόνα 28 : Εισαγωγή ροής ζωντανών δεδομένων – Power BI

Εφαρμογές Power BI

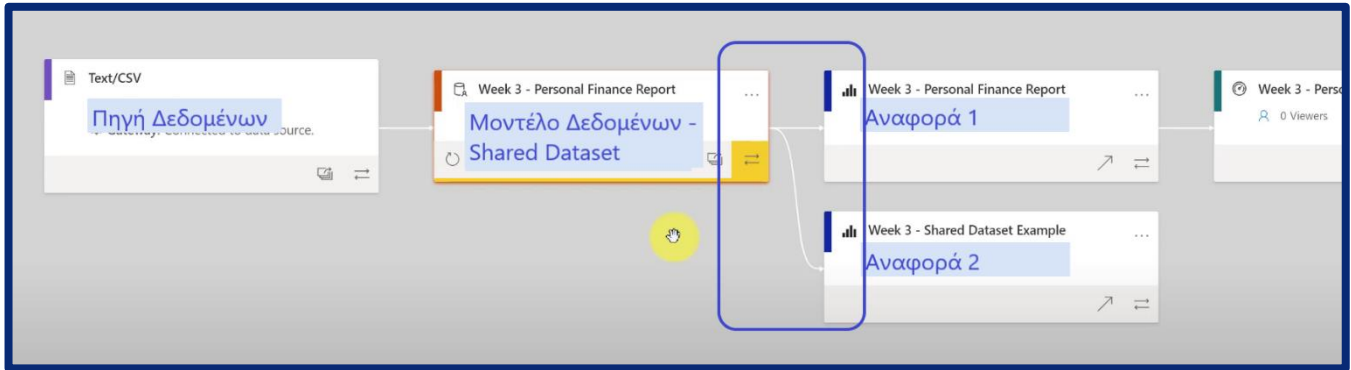
Η εφαρμογή Power BI είναι σαν το συμπυκνωμένο περιεχόμενο από έναν χώρο εργασίας. Σε αυτές μπορούμε να προσθέσουμε τις καλύτερες αναφορές και τους πίνακες ελέγχου μας από τον χώρο εργασίας ώστε να προσεγγίζουν ένα ευρύτερο κοινό. Αυτό **επιτρέπει σε άτομα εκτός του οργανισμού** μας, άτομα που δεν είναι προγραμματιστές Power BI, **να προβάλουν και να αλληλεπιδρούν με τις αναφορές** μας χωρίς να είναι συνεισφέρων μέλος του χώρου εργασίας. Ένα καλό παράδειγμα θα ήταν ο διευθυντής μιας άλλης εταιρείας, στον οποίο μπορούμε να στείλουμε ένα σύνδεσμο για να κατεβάσει και να δει την εφαρμογή που δημιουργήσαμε. Όπως και στους χώρους εργασίας, έτσι και στις **εφαρμογές Power BI** μπορούμε να ορίσουμε το κοινό ή τα άτομα τα οποία, μπορούν να κατεβάσουν ή να αλληλεπιδράσουν με το περιεχόμενο της.



Εικόνα 29 : Καρτέλα κατεβασμένων εφαρμογών – Power BI Service

Κοινόχρηστα Μοντέλα Δεδομένων - Shared Datasets

Ένα από τα στοιχεία που ξεχωρίζουν τα συστήματα επιχειρηματικής ευφυΐας από την απλή δημιουργία και την παρουσίαση αναφορών, είναι το στοιχείο της συνεργασίας των ατόμων σε ένα χώρο οργανωμένο και έτοιμο προς αξιοποίηση. Όταν κάποιος σχεδιαστής δημιουργεί ένα μοντέλο δεδομένων, με δεδομένα από το τμήμα των πωλήσεων για παράδειγμα, μπορεί να κάνει αυτό το μοντέλο ως την βάση για την δημιουργία και άλλων αναφορών και υπολογισμών από τους συναδέλφους του ή για άλλες αναφορές δικές του. Έτσι δεν χρειάζεται κάθε σχεδιαστής να φτιάχνει το δικό του μοντέλο, με τους δικούς του υπολογισμούς και παραμέτρους. Έτσι αποφεύγονται πιθανή σύγχυση για το ποιο μοντέλο είναι σωστό ή ποια αποτελέσματα είναι τελικά σωστά. Όλοι επικοινωνούν με «κοινή γλώσσα». Αυτό που μόλις περιγράψαμε είναι η λειτουργία ενός κοινόχρηστου μοντέλου δεδομένων (**Shared Dataset**).



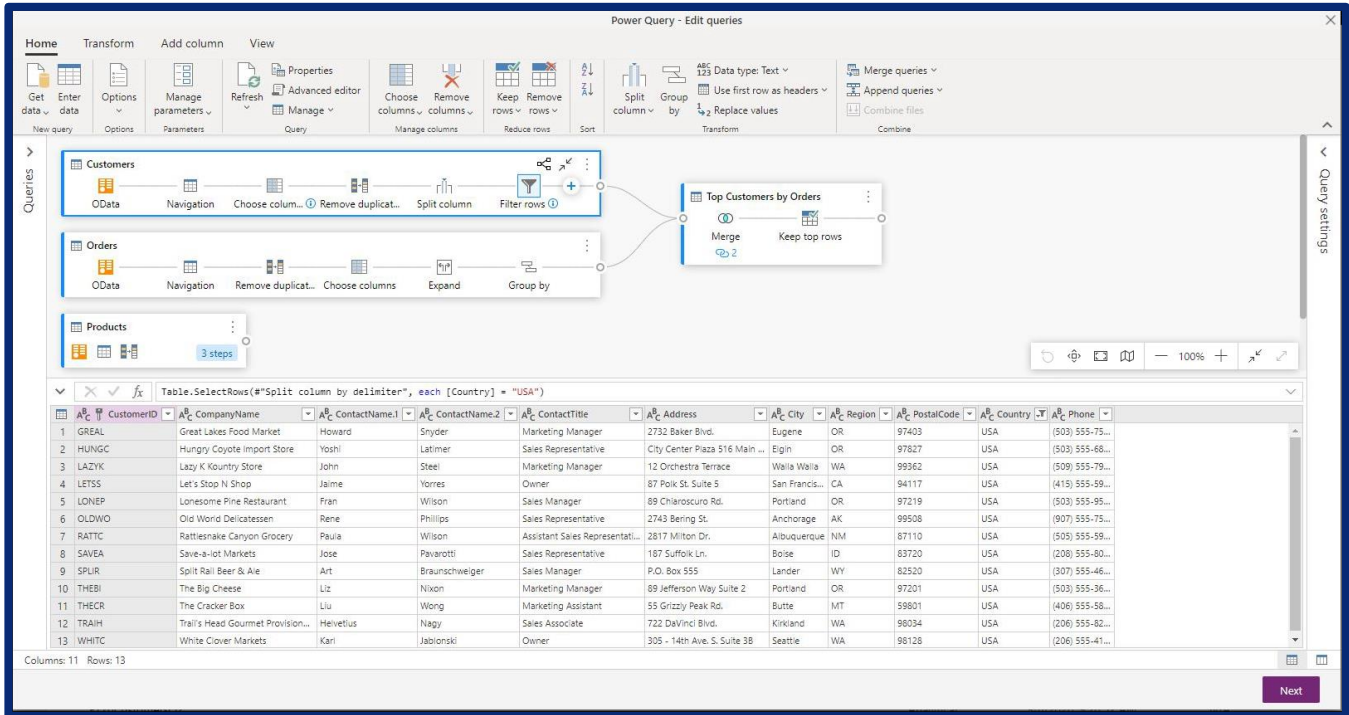
Εικόνα 30 : Δημιουργία πολλαπλών αναφορών από κοινό μοντέλο δεδομένων ,Data Lineage – Power BI Service

Χάριν σε μια λειτουργία του **Power BI Service**, που ονομάζεται **Data Lineage** μπορούμε να δούμε τα βήματα της πορείας δημιουργίας των αναφορών και των πινάκων ελέγχου. Στην παραπάνω εικόνα, φαίνεται παράδειγμα από την δημιουργία δύο αναφορών από ένα κοινό μοντέλο δεδομένων.

Dataflows

Το μέλλον του Power BI είναι στο **Power BI Service**. Όλο και περισσότερες δυνατότητες προστίθενται εκεί με σκοπό, όλες οι εργασίες να μπορούν να πραγματοποιούνται online. Αυτό προσφέρει μεγαλύτερη ευελιξία στο σύστημα επιχειρηματικής ευφυΐας ,εφόσον όλες οι αλλαγές συμβαίνουν στο cloud και όλες οι αλλαγές μεταφέρονται αυτόματα στις αναφορές και στις εφαρμογές. Με γνώμονα αυτό, έχει προστεθεί στο **Power BI Service** και η δυνατότητα για φόρτωση και επεξεργασία των δεδομένων σε αυτό ,μέσω των **Dataflows**. Μέσω αυτής της λειτουργίας τα δεδομένα πλέον συνδέονται με τον διακομιστή του PBI Service. Αυτός ο online Editor, έχει παρόμοια μορφή και λειτουργία με το **Power Query Editor** ,που βρίσκεται στο **Power BI Desktop**. Μπορούμε να καθαρίσουμε και να επεξεργαστούμε τα δεδομένα αποκλειστικά στο cloud.

Ένα χρήσιμο χαρακτηριστικό στα **Dataflow**, είναι το 'Diagram View' ,όπου φαίνεται το ιστορικό των μετατροπών που έχουμε κάνει στα δεδομένα .

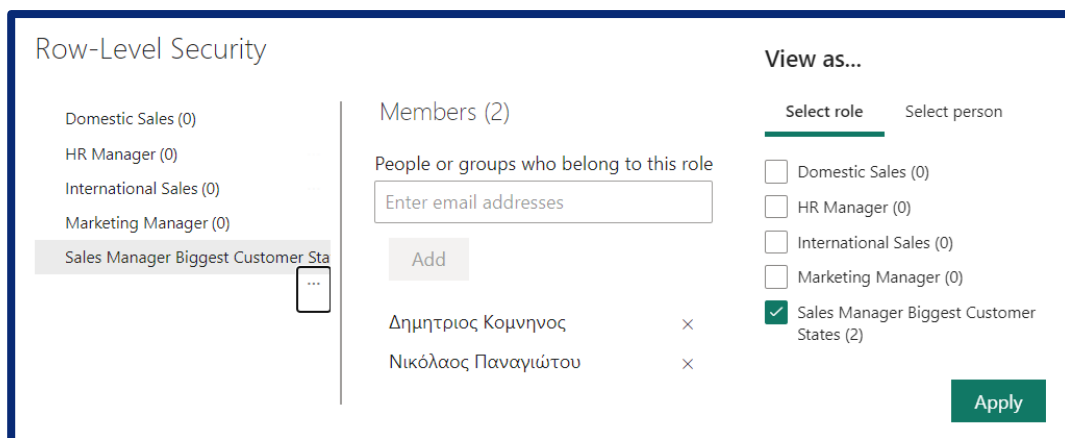


Εικόνα 31 : Diagram View, Dataflow - Power BI Service [Πηγή Εικόνας](#)

Row Level Security

Ένα ακόμα στοιχείο το οποίο επιτελεί στην ομαλή και οργανωμένη λειτουργία ενός συστήματος είναι δημιουργία ρόλων στο επίπεδο μοντελοποίησης ,στο **Power BI Desktop**. Για να γίνουμε πιο ακριβής υπάρχει η δυνατότητα να αποθηκευτεί ένα συγκεκριμένο φιλτράρισμα των δεδομένων (**row level security**) και ανατεθεί σε έναν συγκεκριμένο ρόλο. Ένα κλασικό παράδειγμα αποτελεί την δημιουργία ρόλων για μάνατζερ που ενδιαφέρονται για πωλήσεις σε διαφορετικές χώρες. Αυτοί οι ρόλοι μπορούν να ανατεθούν σε πραγματικά πρόσωπα και αυτοί να βλέπουν τα δεδομένα που πραγματικά ενδιαφέρονται ,αξιοποιώντας το ίδιο μοντέλο δεδομένων.

Έχουν προστεθεί στο τελικό παράδειγμα της ενότητας **4. Αξιολόγηση Εταιρείας Ηλεκτρονικού Εμπορίου** , δύο ρόλοι .Ο πρώτος για τις συνολικές πωλήσεις εντός Βραζιλίας και ο δεύτερος για πωλήσεις για τις περιοχές της Βραζιλίας με τον μεγαλύτερο όγκο πωλήσεων.



Εικόνα 32 : Ρύθμιση Λειτουργίας Row Level Security – Power BI Service

3.6.3 Ολοκληρωμένο Self-service σύστημα στο Power BI

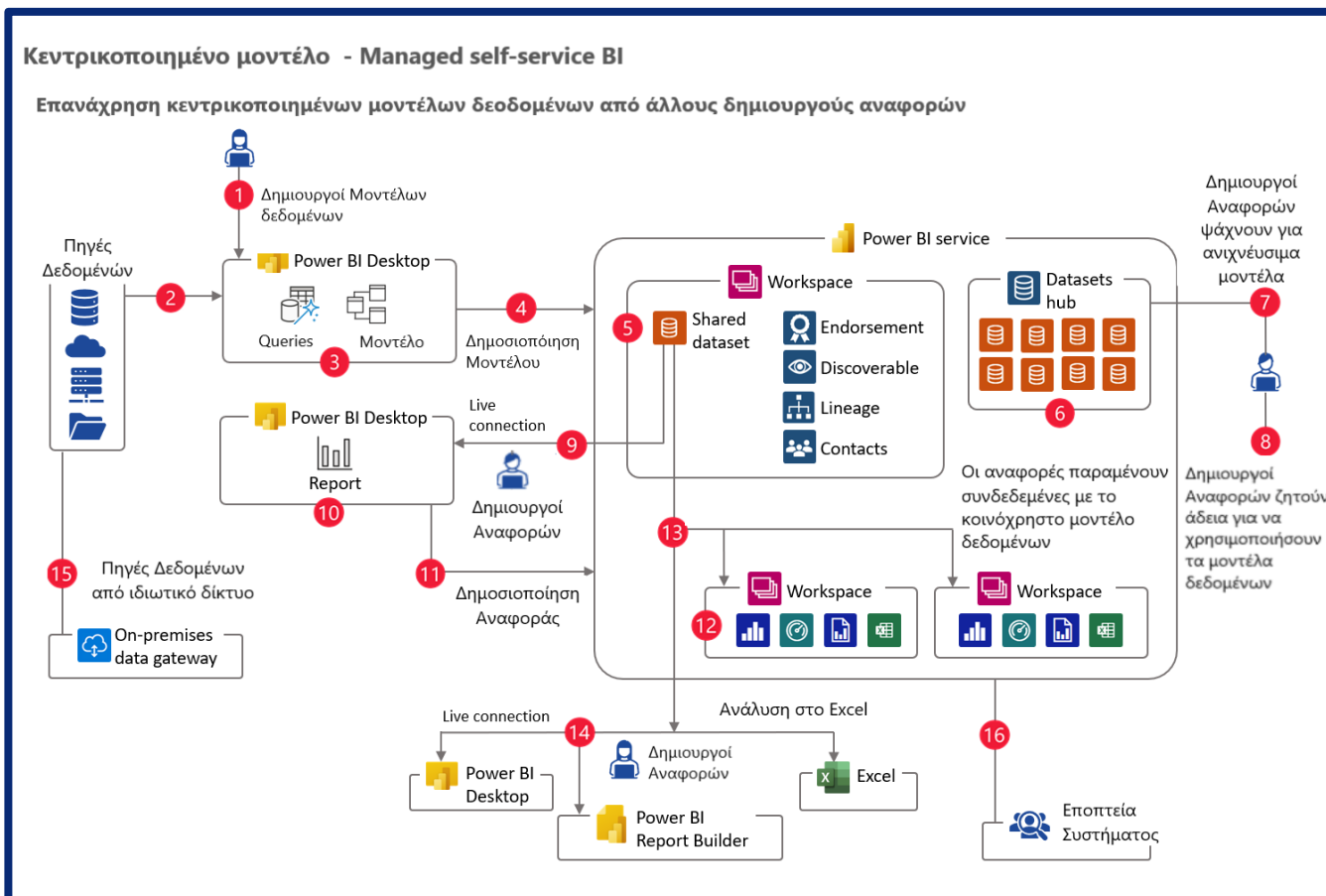
Σε ένα σύστημα επιχειρηματικής ευφυΐας για να μπορέσει να αναδειχθεί ολόκληρη η λειτουργικότητά του, ο οργανισμός ή ομάδα επαγγελματιών που το χειρίζεται οφείλει να έχει ένα κοινό προσανατολισμό και τρόπο προσέγγισης. Αυτό προκύπτει και από το επίπεδο ωριμότητας της εταιρείας για τον τρόπο που διαχειρίζεται τα δεδομένα της, όπως αναλύθηκε στην υποενότητα 2.4 Η κουλτούρα των εταιριών γύρω από το BI και τα Analytics.

Υπάρχουν διάφορα σενάρια αρχιτεκτονικής ενός συστήματος επιχειρηματικής ευφυΐας. Για κάθε οργανισμό ταιριάζει κάτι διαφορετικό, πάντα εξειδικευμένο στις ανάγκες της εταιρείας. Συνήθως σε επίπεδο οργανισμού, η βασική διαφοροποίηση για την δομή του συστήματος σχετίζεται αν θα υπάρχει **κεντροποιημένη δομή για την δημιουργία μοντέλων δεδομένων**. Η **Microsoft** έχει ορίσει ορισμένα τέτοια τυπικά σενάρια χρήσης σε επίπεδο οργανισμού.

Κεντροποιημένο μοντέλο χρήσης - Managed self-service BI [18]

Η αρχιτεκτονική ενός τέτοιου συστήματος διατηρείται συνήθως **από μια κεντροποιημένη ομάδα ειδικών BI, που ασχολούνται αποκλειστικά με την δημιουργία των μοντέλων όλου του οργανισμού**. Τα μοντέλα που δημιουργούνται αποθηκεύονται σε ένα κεντροποιημένο χώρο αποθήκευσης στο **Power BI Service**, στον οποίο δεν έχουν πρόσβαση όλοι, παρά μόνον οι όσοι έχουν άδεια. Η ευθύνη για την δημιουργία αναφορών βρίσκεται σε άλλα άτομα στον οργανισμό. Συνήθως, υπάρχουν πολύ περισσότεροι δημιουργοί αναφορών από δημιουργοί μοντέλων δεδομένων. Αυτοί οι δημιουργοί αναφορών μπορούν να υπάρχουν σε οποιαδήποτε περιοχή του οργανισμού.

Επειδή οι δημιουργοί αναφορών χρειάζεται συχνά να παράγουν γρήγορα περιεχόμενο, μια τέτοια δομή τους επιτρέπει να επικεντρώνονται αποκλειστικά σε αυτό το κομμάτι. Αυτός ο τύπος δόμησης του self-service συστήματος επιχειρηματικής ευφυΐας έχει αυστηρούς κανόνες στον πυρήνα λειτουργείας του και προσφέρει μεγάλη ευελιξία στους χρήστες του οργανισμού.



Εικόνα 33 : Κεντριοποιημένο μοντέλο χρήσης – Managed self-service BI [Πηγή εικόνας](#)

Αρ.	Περιγραφή
1	Οι δημιουργοί μοντέλων χρησιμοποιούν το Power BI Desktop για να δημιουργήσουν τα μοντέλα. Αυτοί οι δημιουργοί ανήκουν συνήθως σε κάποια κεντριοποιημένη ομάδα ειδικών όπως το IT.
2	Το Power BI Desktop συνδέεται με κάποια πηγή δεδομένων
3	Οι δημιουργοί των μοντέλων δεδομένων προσπαθούν να δημιουργήσουν ακόμα πιο φιλική προς τον χρήστη μορφή των μοντέλων ,για τους δημιουργούς αναφορών.
4	Όταν το μοντέλο δεδομένων είναι έτοιμο ,δημοσιοποιείται αρχείο .pbix στο Power BI Service.

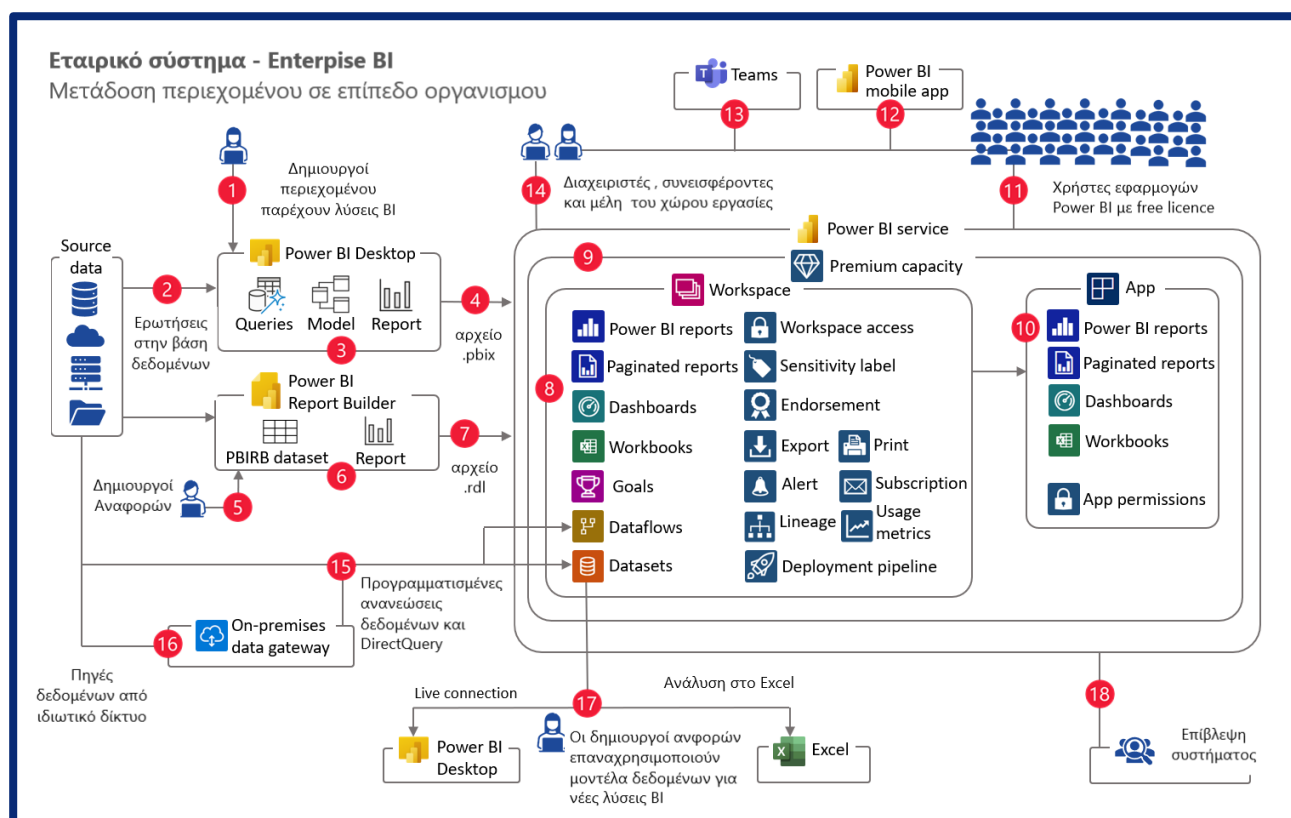
5	Το μοντέλο δημοσιοποιείται σε ένα κεντρικοποιημένο χώρο εργασίας αφιερωμένο για στα κοινόχρηστα μοντέλα δεδομένων. Τα μοντέλα δεδομένων προορίζονται για επανάχρηση και μαρκάρονται με ως πιστοποιημένα για χρήση.
6	Υπάρχει και μια άλλη κεντρικοποιημένη δομή αποθήκευσης (Data Hub) όλων των μοντέλων δεδομένων, η οποία ενεργοποιείται αφού , τα δεδομένα έχουν μαρκαριστεί ως Discoverable . Δηλαδή μοντέλα δεδομένων τα οποία κανονικά δεν ήταν διαθέσιμα σε όλους , πλέον έχουν αρθεί οι περιορισμοί.
7	Οι δημιουργοί αναφορών μπορούν να ψάξουν και να χρησιμοποιήσουν μοντέλα δεδομένων που έχουν μαρκαριστεί ως Discoverable , για την δημιουργία αναφορών.
8	Αν οι δημιουργοί των αναφορών δεν έχουν άδεια να χρησιμοποιήσουν το μοντέλο που τους ενδιαφέρει , ζητούν άδεια σε αυτά.
9	Οι δημιουργοί για να χρησιμοποιήσουν τα κοινόχρηστα μοντέλα δεδομένων χρησιμοποιούν τον τύπο σύνδεσης δεδομένων, <i>live connection</i> .
10	Οι δημιουργοί αναφορών ,φτιάχνουν αναφορές στο Power BI Desktop.
11	Όταν ετοιμαστούν οι αναφορές δημοσιοποιούνται στο Power BI Service.
12	Οι αναφορές αποθηκεύονται σε συγκεκριμένους χώρους εργασίας, αφιερωμένοι αποκλειστικά για αναφορές και πίνακες ελέγχου.
13	Οι αναφορές παραμένουν συνδεδεμένες με τα μοντέλα δεδομένων, από τα οποία έχουν προκύψει και ανανεώνονται αυτόματα αν προκύψει κάποια αλλαγή σε αυτά.
14	Άλλοι δημιουργοί αναφορών χρησιμοποιούν τα κοινόχρηστα μοντέλα δεδομένων για περισσότερη ανάλυση ή την σύνταξη διαφορετικών αναφορών στο Power BI Desktop ,Power BI Report Builder, το Power BI Service ή το Excel.
15	Όταν τα δεδομένα προέρχονται από κάποιον τοπικό διακομιστή ή δίκτυο, πρέπει να δημιουργηθεί μια πύλη δεδομένων (On-premises Data Gateway) για να ανανεώνονται τα δεδομένα.
16	Οι διαχειριστές του Power BI συστήματος στον οργανισμό, επιβλέπουν και ελέγχουν την δραστηριότητα στο Power BI Service.

Πίνακας 8 : Περιγραφή διαδικασιών λειτουργίας ενός συστήματος με κεντρικοποιημένο μοντέλο – Managed self-service BI

Εταιρικό self-service σύστημα – Enterprise BI [19]

Το εταιρικό BI έχει παρόμοιες βασικές αρχές με το κεντρικοποιημένο μοντέλο συστήματος, αλλά διαφοροποιείται στον αριθμό των ατόμων που έχουν πρόσβαση στο περιεχόμενο του συστήματος. Συγκεκριμένα το εταιρικό BI απευθύνεται σε έναν πολύ μεγαλύτερο αριθμό καταναλωτών περιεχομένου, σε σύγκριση με έναν πολύ μικρότερο αριθμό συντακτών που δημιουργούν και δημοσιεύουν περιεχόμενο.

Στο εταιρικό BI γίνεται χρήση των δυνατοτήτων της άδειας χρήσης **Power BI Premium**, η οποία επιτρέπει την ευρεία διανομή περιεχομένου σε καταναλωτές που έχουν δωρεάν άδεια χρήσης (**Free licence**) Power BI. **Οι καταναλωτές μπορούν να περιλαμβάνουν χρήστες εντός του οργανισμού, καθώς και χρήστες-επισκέπτες που είναι εξωτερικοί του οργανισμού.**



Εικόνα 34 : Ολοκληρωμένο Σύστημα Επιχειρηματικής Ευφύϊας σε επίπεδο εταιρείας
[Πηγή Εικόνας](#)

Αρ.	Περιγραφή
1	Οι δημιουργοί περιεχομένου Power BI αναπτύσσουν λύσεις BI με χρήση του Power BI Desktop. Σε ένα εταιρικό σενάριο BI, είναι σύνηθες οι δημιουργοί να είναι μέλη μιας κεντρικής ομάδας όπως το IT.
2	Το Power BI Desktop συνδέεται σε δεδομένα από μία ή περισσότερες πηγές δεδομένων και τα ενσωματώνει σε ενιαία μορφή στο Power Query.

3	Η ανάπτυξη μοντέλων δεδομένων και η δημιουργία αναφορών γίνονται στο Power BI Desktop. Το περιεχόμενο αυτό πρέπει να είναι σε απλή και φιλική προς τον χρήστη μορφή.
4	Όταν το μοντέλο δεδομένων είναι έτοιμο ,δημοσιοποιείται αρχείο <i>.pbix</i> στο Power BI Service.
5	Οι δημιουργοί αναφορών αναπτύσσουν σελιδοποιημένες αναφορές χρησιμοποιώντας το Power BI Report Builder.
6	Το Power BI Report Builder αναφορών υποβάλλει ερωτήματα στα δεδομένα. Εδώ δημιουργούνται σελιδοποιημένες αναφορές (Paginated Report) μπορούν να εκτυπωθούν.
7	Όταν είναι έτοιμο, οι δημιουργοί αναφορών δημοσιεύουν το αρχείο <i>.rdl</i> που δημιούργησαν στο Power BI Report Builder στο Power BI Service.
8	Πολλοί τύποι στοιχείων Power BI μπορούν να δημοσιευτούν σε έναν χώρο εργασίας Premium.
9	Σε αυτό το τύπο συστήματος γίνεται χρήση της άδειας χρήσης Power BI Premium, ώστε να γίνεται πλήρης αξιοποίηση των δυνατοτήτων του Power BI Service.
10	Ορισμένες ή όλες οι αναφορές και οι πίνακες ελέγχου δημοσιεύονται ως εφαρμογή Power BI, με σκοπό την δημιουργία ενός πιο φιλικού προς τον χρήστη περιβάλλοντος.
11	Μέσω της άδειας Power BI Premium, χρήστες που έχουν δωρεάν άδεια μπορούν να καταναλώσουν το περιεχόμενο των αναφορών και των εφαρμογών του οργανισμού.
12	Το περιεχόμενο των εφαρμογών και των χώρων εργασίας που δημιουργεί ο οργανισμός είναι διαθέσιμο και σε εφαρμογή smartphone, την Power BI Mobile app.
13	Οι χρήστες που εργάζονται συχνά στο Microsoft Teams , μπορούν να προβάλουν το περιεχόμενο που δημιουργεί ο οργανισμός απευθείας εκεί.
14	Διαχειριστές , συνεισφέροντα και απλά μέλη μπορούν να διαχειρίζονται το περιεχόμενο των χώρων εργασίας
15	Ρυθμίζεται η προγραμματισμένη ανανέωση δεδομένων (Scheduled refresh data) στο Power BI Service για να διατηρούνται ενημερωμένα τα δεδομένα που έχουν εισαχθεί σε αυτό.
16	Όταν τα δεδομένα προέρχονται από κάποιον τοπικό διακομιστή ή δίκτυο, πρέπει να δημιουργηθεί μια πύλη δεδομένων (On-premises Data Gateway) για να ανανεώνονται τα δεδομένα.

17	Άλλοι δημιουργοί αναφορών χρησιμοποιούν τα κοινόχρηστα μοντέλα δεδομένων για περισσότερη ανάλυση ή την σύνταξη διαφορετικών αναφορών στο Power BI Desktop ,Power BI Report Builder, το Power BI Service ή το Excel.
18	Οι διαχειριστές του Power BI συστήματος στον οργανισμό, επιβλέπουν και ελέγχουν την δραστηριότητα στο Power BI Service.

Πίνακας 9 : Περιγραφή διαδικασιών λειτουργίας ενός self-service συστήματος σε εταιρικό επίπεδο – Enterprise BI

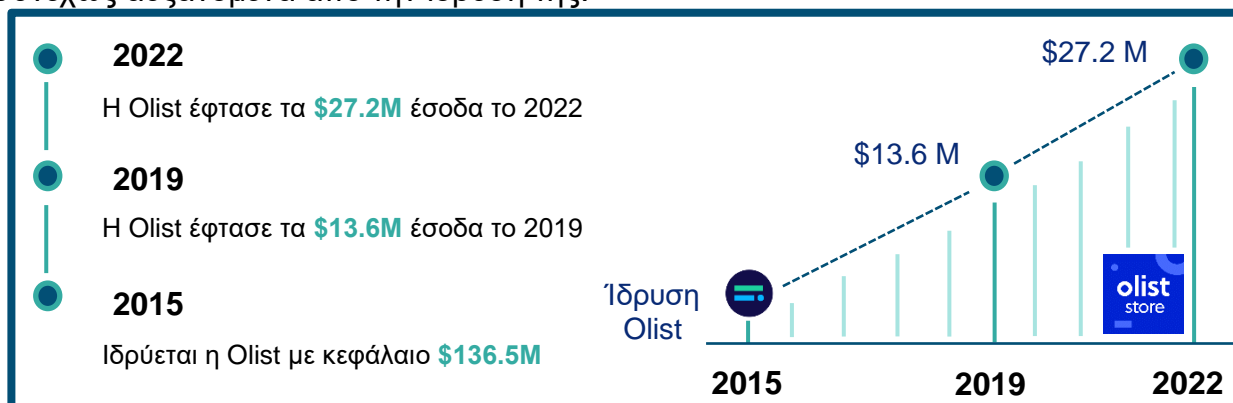
Η επιλογή της δομής για τον κάθε οργανισμό είναι κάτι ξεχωριστό σε κάθε περίπτωση ,καθώς ο κάθε οργανισμός έχει τις ιδιαιτερότητές του και τις ειδικές ανάγκες του. Η κοινή αρχή σχεδίασης και ανάπτυξη τους όμως, για όλα τα συστήματα, είναι τα δεδομένα και το περιεχόμενο που δημιουργούν οι χρήστες στο Power BI Service, να ακολουθεί κοινούς κανόνες και πρακτικές ώστε να φτάνουν όλοι στα ίδια συμπεράσματα. Αυτό επιτυγχάνεται αν υπάρχει εκπαίδευση στους χρήστες από τον οργανισμό ,για να ακολουθούν τις οδηγίες που έχουν εκδοθεί από τους σχεδιαστές του συστήματος. Ενώ, ταυτόχρονη πρέπει να είναι και η παρακολούθηση του συστήματος από τους «επόπτες» ,οι οποίοι είναι σε θέση να αξιολογούν τις καταστάσεις και να παρεμβαίνουν για να διατηρείται αυτό το κοινό σύστημα επικοινωνίας.

4. Αξιολόγηση Εταιρείας Ηλεκτρονικού Εμπορίου

Αφού προηγήθηκε το θεωρητικό υπόβαθρο για τα συστήματα επιχειρηματικής ευφυΐας , στην παρούσα ενότητα θα προσπαθήσουμε να αναδείξουμε μέσω ενός ρεαλιστικού παραδείγματος, την χρησιμότητα τέτοιων αναλυτικών συστημάτων.

Εν μέσω επιβράδυνσης της οικονομικής δραστηριότητας, ο COVID-19 οδήγησε σε άνοδο του ηλεκτρονικού εμπορίου και επιτάχυνε τον ψηφιακό μετασχηματισμό. Καθώς τα lockdown έγιναν το νέο φυσιολογικό, οι επιχειρήσεις και οι καταναλωτές «πέρασαν ψηφιακά», παρέχοντας και αγοράζοντας περισσότερα αγαθά και υπηρεσίες στο διαδίκτυο, αυξάνοντας το μερίδιο του ηλεκτρονικού εμπορίου στο παγκόσμιο λιανικό εμπόριο από 14% το 2019 σε περίπου 17% το 2020 [20] .

Δεδομένα για την δημιουργία αυτού του παραδείγματος συλλέχθηκαν από το διαδίκτυο και συγκεκριμένα μέσω της ιστοσελίδας - απηθηκευτήριο *Kaggle.com* [21] ,[22]. Τα δεδομένα αυτά έχουν μεταφορτωθεί στην ιστοσελίδα από εταιρία **ηλεκτρονικού εμπορίου** με βάση την Βραζιλία, την **Olist**. Η Olist έχει αναπτύξει το ταχύτερα αναπτυσσόμενο εργαλείο ενδυνάμωσης εμπόρων λιανικής στη Βραζιλία που δημιουργεί λύσεις για να βοηθήσει τους πωλητές να βρουν αγοραστές, επηρεάζοντας θετικά χιλιάδες εταιρείες από πολύ μικρές έως μεγάλες μάρκες στην αγορά. Στο **επιχειρηματικό μοντέλο** (business model) της εταιρείας ,οι επιχειρήσεις - συνεργάτες εγγράφονται στην διαδικτυακή πλατφόρμα της εταιρείας για να επικοινωνήσουν τα προϊόντα τους στους πελάτες ,ενώ συγχρόνως αξιοποιούν το διευρυμένο δίκτυο Logistics της εταιρείας για να μεταφέρουν τα προϊόντα που επιλέγουν οι πελάτες ,σε αυτούς. Από την τιμή των προϊόντων η εταιρεία κρατάει το **21% από την τιμή του προϊόντος**. από την Η Olist ιδρύθηκε αρχικά στην Βραζιλία το 2015 με κεφάλαιο \$136.5 εκατομμυρίων , ενώ στην συνέχεια επεκτάθηκε στην αγορά της Νιγηρίας .Απασχολεί συνολικά 593 υπαλλήλους και 93 αντιπροσώπους πωλήσεων. Τα ετήσια έσοδα της βαίνουν συνεχώς αυξανόμενα από την ίδρυση της.



Διάγραμμα 11 : Τα ετήσια έσοδα της Olist στην πορεία των χρόνων

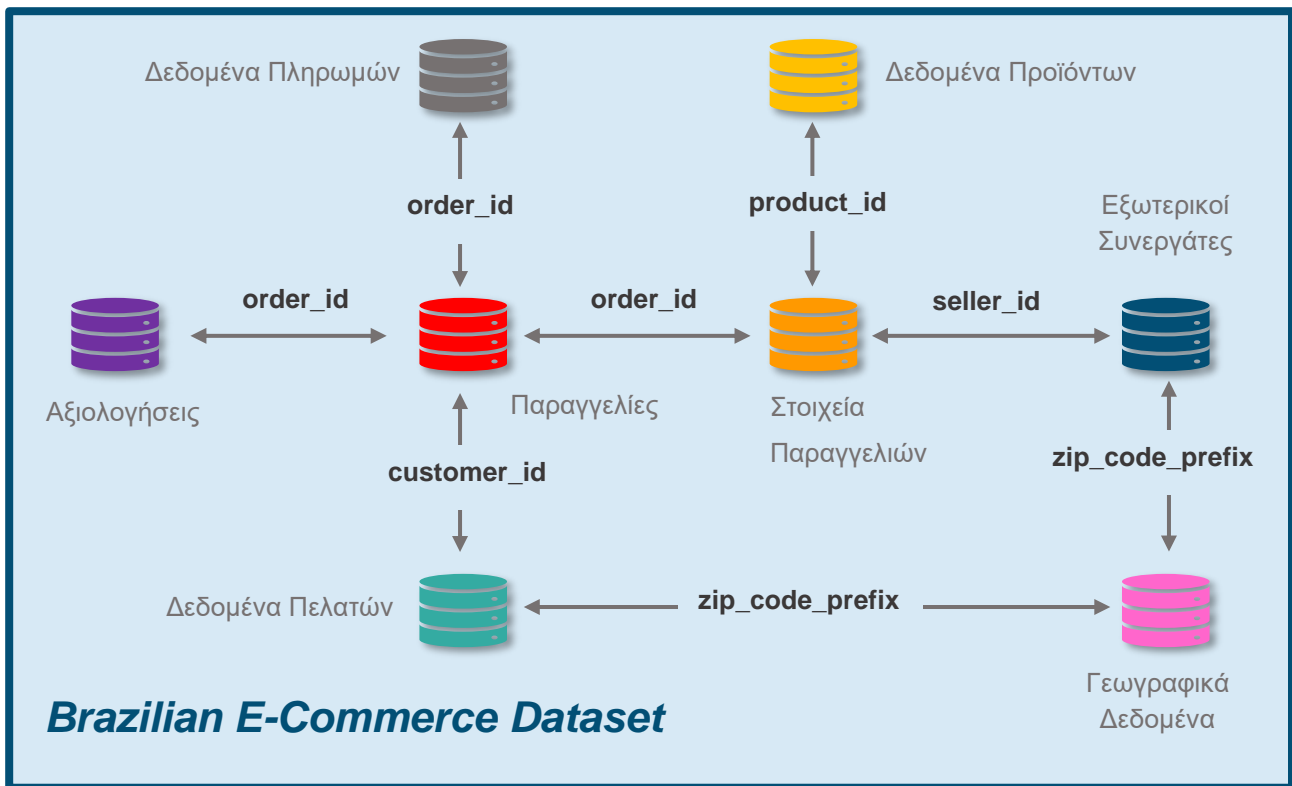
Το σύνολο των δεδομένων περιέχει πληροφορίες για 100.000 παραγγελίες από το 2016 έως το 2018 που έγιναν σε ηλεκτρονικές αγορές (**marketplaces**) στη Βραζιλία και κάποιες σε άλλες χώρες. Τα χαρακτηριστικά του σετ δεδομένων επιτρέπουν την ανάλυση μιας παραγγελίας σε πολλές διαστάσεις: από την κατάσταση της παραγγελίας, την τιμή, την απόδοση πληρωμής και μεταφοράς έως την τοποθεσία του πελάτη, τα χαρακτηριστικά του προϊόντος και, τέλος, τις κριτικές που γράφτηκαν από πελάτες. Εμπεριέχονται επίσης, ένα σύνολο δεδομένων γεωγραφικής τοποθεσίας που συσχετίζει τους ταχυδρομικούς κώδικες της Βραζιλίας με τις συντεταγμένες.

Το παραπάνω σετ δεδομένων (**dataset**) συνδυάστηκε και με ένα άλλο σετ δεδομένων που επίσης δημοσιεύτηκε από την εταιρεία και αφορά ένα σύνολο δεδομένων μάρκετινγκ από επιχειρήσεις - πωλητές που πραγματοποίησαν αιτήματα επικοινωνίας με την εταιρεία, με οποιονδήποτε τρόπο, για να πουλήσουν τα προϊόντα τους στο Olist Store. Το σύνολο δεδομένων περιέχει πληροφορίες 8 χιλιάδων Πιστοποιημένων εν Δυνάμει Πελατών (**Marketing Qualified Leads - MQL**) που ήρθαν σε επαφή με την εταιρεία, μεταξύ της 1ης Ιουνίου 2017 και της 1ης Ιουνίου 2018. Η δειγματοληψία τους έγινε τυχαία από το σύνολο των MQL. Τα χαρακτηριστικά των δεδομένων επιτρέπουν, όπως και πριν, την προβολή μιας διαδικασίας πώλησης από πολλαπλές διαστάσεις όπως, κατηγορία δυνητικού πελάτη, μέγεθος καταλόγου εταιρείας, προφίλ συμπεριφοράς κ.λπ. Αυτά είναι πραγματικά δεδομένα, τα οποία έχουν ανωνυμοποιηθεί.

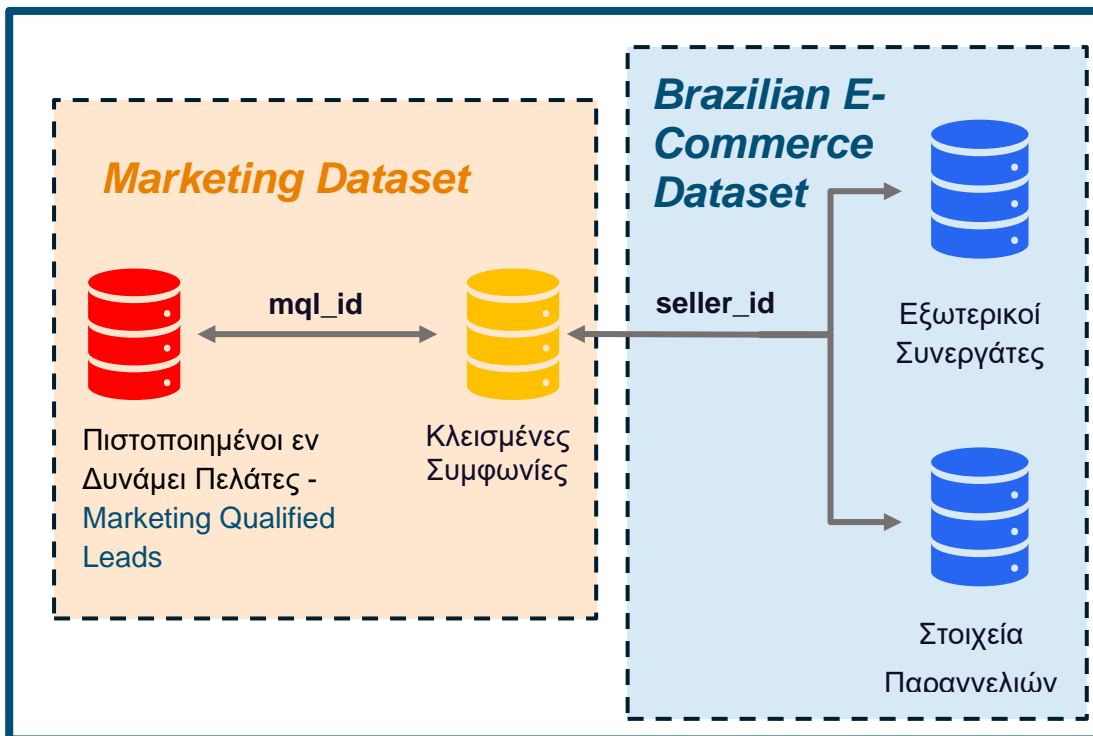
Συνδυαστικά αυτός ο πλούσιος όγκος δεδομένων, μπορεί να βοηθήσει να γίνει μια πολύπλευρη διαστατική ανάλυση και αξιολόγηση της απόδοσης διαφόρων τομέων της εταιρείας αλλά και του προσωπικού της εταιρείας. Αυτές οι αναλύσεις είναι χρήσιμες για τομείς της επιχείρησης όπως: οι Πωλήσεις, το Μάρκετινγκ, την Οικονομική Διεύθυνση, την Γενική Διεύθυνση και την Διεύθυνση Διαχείρισης Εμπορευμάτων (Logistics).

Αξίζει να σημειωθεί ότι, για το στήσιμο της προσομοίωσης μας, πραγματοποιήσαμε **Αντίστροφη Μηχανική (Reverse Engineering)** του προβλήματος. Αυτό σημαίνει ότι τα δεδομένα, που αξιοποιήθηκαν ήταν ήδη έτοιμα και δεν χρειάστηκε να προηγηθούν συνεντεύξεις με άτομα της εταιρείας για να συλλεχθούν. Η ίδια η εταιρεία έδωσε τα δεδομένα έτοιμα προς ανάλυση και εμείς προσπαθήσαμε να ανακατασκευάσουμε το υπόβαθρο της επιλογής αυτών, δηλαδή τις πιθανές ανάγκες της εταιρείας και στην συνέχεια με βάση αυτές, έγινε η δημιουργία του αναλυτικού συστήματος.

Στα επόμενα δύο σχήματα που ακολουθούν φαίνονται οι πίνακες δεδομένων που δόθηκαν από την εταιρεία. Στην **Σχήμα 13** φαίνεται το πρώτο σετ δεδομένων, όπου αποτελεί τις παραγγελίες τις οποίες έχει καταγράψει το σύστημα της εταιρείας. Στην **Σχήμα 14** απεικονίζεται το δεύτερο σετ δεδομένων, τα οποία έχουν τα δεδομένα του Μάρκετινγκ, δηλαδή δεδομένα από την αλληλεπίδραση της εταιρείας με πιθανούς συνεργάτες-επιχειρήσεις, ενώ παράλληλα φαίνεται ο τρόπος σύνδεσης των δύο σετ δεδομένων με κάποιες στήλες- κλειδιά που έχουν κοινές τιμές μεταξύ τους.



Σχήμα 15 : Πρώτο σετ δεδομένων με τις παραγγελίες της εταιρείας – Brazilian E-Commerce Dataset



Σχήμα 14 : Δεύτερο σετ δεδομένων Μάρκετινγκ – Marketing Dataset

4.1 Συλλογή επιχειρηματικών απαιτήσεων – Καταγραφή των Αναγκών του οργανισμού

Το πρώτο βήμα για την υλοποίηση ενός αναλυτικού συστήματος , τύπου επιχειρηματικής ευφυΐας, είναι να γίνει η εννοιολογική σύλληψη του μοντέλου. Σε αυτή την προκαταρκτική φάση γίνεται συζήτηση με τον πελάτη , με σκοπό :

- την καταγραφή ήδη γνωστών αναγκών που επιθυμεί ο πελάτης να καλύψει με το σχεδιαζόμενο σύστημα, όπως KPIs (Key Performance Indicators).
- την πιθανή ανακάλυψη άγνωστων μέχρι εκείνη την στιγμή αναγκών.
- τον εντοπισμό *ευκαιριών* και *προβλημάτων* στον τρόπο λειτουργίας του οργανισμού.
- να μάθουμε τον τρόπο λειτουργίας της εταιρείας, τις διάφορες επιχειρηματικές διαδικασίες (**business processes**) που λαμβάνουν χώρα.
- την κουλτούρα του οργανισμού γύρω από την διαχείριση των δεδομένων της.
- το επίπεδο ωριμότητας του οργανισμού ,όπως αναφέρθηκε στην **Ενότητα 2.4**
- να εντοπίσουμε τις ομάδες ατόμων και τους επαγγελματίες που ασχολούνται με την σύνταξη των δεδομένων, τον καθαρισμό τους , την συντήρησή τους ,την αποθήκευση και την διαχείριση τους (Αρχιτέκτονας Δεδομένων, Σχεδιαστής Βάσης Δεδομένων, ETL developer κ) . Αυτοί είναι τα άτομα που θα μας βοηθήσουμε να καταλάβουμε καλύτερα όλα τα παραπάνω και επιπλέον τι συστήματα αξιοποιεί η εταιρεία στον παρόντα χρόνο αλλά και τι μπορεί να υλοποιηθεί με βάση αυτά στο μέλλον .

Στην πορεία των τέτοιων συζητήσεων ,χωρίς να γίνεται άμεση ερώτηση προς τους χρήστες της επιχείρησης , ο σχεδιαστής αναλυτικών συστημάτων μπορεί να αντιληφθεί πως γίνεται η συλλογή δεδομένων στην εταιρεία , που και σε ποια μορφή γίνεται η αποθήκευση της πληροφορίας , πως γίνεται η καταγραφή και η αξιολόγηση της προόδου του οργανισμού, σε ποιο επίπεδο γίνεται αυτή η αξιολόγηση (**grain - granularity of data**) όπως εννοείται στην υποενότητα **3.4.2**. Με βάση αυτές τις πληροφορίες μπορεί ο σχεδιαστής ευφυών συστημάτων να ξεκινήσει να σχεδιάζει το σύστημα με βάση τις αρχές διαστατικού σχεδιασμού ,όπως αναλύθηκε στην αντίστοιχη ενότητα **3.4.2**.

Το πιο βασικό κομμάτι στον αρχικό σχεδιασμό του ενός τέτοιου συστήματος είναι η ανακάλυψη των επιχειρηματικών διαδικασιών του οργανισμού. Στην καταγραφή και την διασύνδεση αυτών των διαδικασιών σε ένα ενιαίο μοντέλο ,βοηθάει ένα εργαλείο από τον μεγάλο σχεδιαστή αναλυτικών συστημάτων τον R.Kimball στην δεκαετία του 1990, όπως αναλύεται και στο σχετικό βιβλίο **[17]**. Αυτό το εργαλείο ονομάζεται μεταξύ άλλων, πίνακας συμβάντων (**enterprise data warehouse bus matrix**). Στις γραμμές του πίνακα βρίσκουμε τις υποψήφια επιχειρηματικές δραστηριότητες και στις στήλες βρίσκουμε συνήθεις διαστάσεις, όπως ορίζονται στον διαστατικό σχεδιασμό στην υποενότητα **3.4.2**. Στη συνέχεια,

σημειώνουμε με ένα X κελία του πίνακα για να δηλώσουμε ποια διαδικασίες της εταιρείας συνδέονται με ποιες διαστάσεις. Φυσικά ένα σύστημα επιχειρηματικής ευφυΐας εξελίσσεται και να ανανεώνεται διαρκώς ,όπως και η ίδια η εταιρεία ,ως εκ τούτου δεν είναι εφικτή η εξαρχής καταγραφή όλων των δραστηριοτήτων και αυτό είναι θεμιτό. Επίσης είναι σημαντικό να γίνει διαχωρισμός μεταξύ των επιχειρηματικών δραστηριοτήτων και των τμημάτων ενός οργανισμού, αυτά τα δύο δεν ταυτίζονται. Μια επιχειρηματική διαδικασία ,όπως η καταγραφή μιας παραγγελίας στο σύστημα του οργανισμού ,μπορεί να αποτελεί μέρος πολλαπλών τμημάτων ,μεταξύ αυτών την Διεύθυνση Πωλήσεων , την Διεύθυνση Διαχείρισης Αποθεμάτων ή το Μάρκετινγκ. Στον παρακάτω πίνακα φαίνονται οι επιχειρηματικές δραστηριότητες και οι διαστάσεις της εταιρείας.

Επιχειρηματικές Δραστηριότητες	Διαστάσεις							
	Ημ/νια	Γεωγραφικά δεδομένα	Προϊόν	Συεργάτες – Επιχειρ.	Πελάτης	Αξιολο γήσεις	Αντιπρ. Πωλήσεων	Marketing Leads
Κατάθεση Παραγγελίας	X	X	X	X	X	X		
Μεταφορά Προϊόντων (Logistics)	X	X	X	X	X	X		
Συγκέντρωση Αξιολογήσεων Πελατών	X	X	X	X		X		
Λήψη Πληρωμών	X	X	X		X			
Αξιολόγηση εργαζομένων	X	X	X			X	X	X
CRM υπηρεσίες*	X	X	X	X	X			X

*(Απόκτηση Δυνητικού Πελάτη, Μετατροπή σε Πελάτη ,Διατήρηση Πελάτη)
= (Lead acquisition ,conversion, retention, customer loyalty)

Πίνακας 10 : Πίνακας επιχειρηματικών δραστηριοτήτων οργανισμού – Olist

Ο παραπάνω πίνακας που έχουν καταγραφεί δραστηριότητες της εταιρείας , ασφαλώς δεν είναι πλήρης αλλά εκφράζει τις διαδικασίες οι οποίες έχουν προκύψει από τα ελεύθερα προσβάσιμα δεδομένα, δηλαδή η εταιρεία έχει και άλλες πολλές δραστηριότητες οι οποίες δεν συμπεριλήφθηκαν σ' αυτό το αρχικό σύστημα που δημιουργήσαμε.

Το σύστημα καταγράφει τις επιχειρηματικές διαδικασίες ,όπως η κατάθεση της παραγγελίας **σε ατομικό επίπεδο πληροφορίας (granularity)** ,δηλαδή κάθε γραμμή στον πίνακα αντιστοιχεί σε μία παραγγελία , για ένα πελάτη , από ένα πωλητή. Το επίπεδο αυτό πληροφορίας ,αν είναι διαθέσιμο , είναι το πιο χρήσιμο διότι με βάση αυτό μπορούμε να

συνοψίσουμε γεγονότα προς τα άνω (roll-up), πχ σε μηνιαίες παραγγελίες ή σε επίπεδο γεωγραφικού νομού. Το μέγεθος του μοντέλου αυξάνει αλλά, παρέχει την μεγαλύτερη ευελιξία στον σχεδιαστή και τον χρήστη του συστήματος.

Έχοντας δημιουργήσει τον πίνακα επιχειρηματικών δραστηριοτήτων , είμαστε σε θέση να κατανοήσουμε με μεγαλύτερη ευκολία τις ανάγκες του πελάτη αλλά και τον τρόπο σχεδιασμού του μοντέλου μας.

Ο πελάτης θέλει να έχει στην διάθεση του ένα σύστημα για να :

- Αξιολογήσει την απόδοση των συστημάτων της στην μεταφορά προϊόντων από τους συνεργάτες - πωλητές μέχρι και τους πελάτες (**Logistics**)
- Να ενημερώνεται για τις **αξιολογήσεις** που γίνονται από τους πελάτες , από τις διάφορες γεωγραφικές περιοχές και να βγάζει χρήσιμα συμπεράσματα
- Να αντιληφθεί ποιες **κατηγορίες προϊόντων είναι πιο επικερδής** , από ποια γεωγραφικά σημεία
- Να έχει στην διάθεση του σε «εύπεπτη» μορφή το **προφίλ των πελατών**, να γνωρίζει ποιοι είναι καλύτεροι πελάτες του ,τον τρόπο που επιλέγουν να πληρώνουν
- Να πραγματοποιήσει **αξιολόγηση στο προσωπικό της εταιρείας** (Αντιπρόσωπους Πωλήσεων ,Αντιπροσώπους Ανάπτυξης Πωλήσεων)
- Να αξιολογήσει την **καμπάνια του μάρκετινγκ**
- Να μπορεί να έχει **οργανωμένα** και από **παντού προσβάσιμη την πληροφορία** που τον ενδιαφέρει

4.2 Εισαγωγή Δεδομένων και κατασκευή Μοντέλου Δεδομένων

Ξεκινάμε την κατασκευή του συστήματος μας με την εισαγωγή των δεδομένων στο *Power BI Desktop*. Τα αρχεία βρίσκονται σε μορφή .csv ,οπότε η φόρτωση τους θα γίνει απευθείας στο τοπικό σύστημα του υπολογιστή , ώστε να γίνει χρήση της πανίσχυρης μηχανής *Vertipaq* της Microsoft. Δεν θα χρειαστεί να γίνει κάποια σύνδεση με απομακρυσμένο σέρβερ μέσω *DirectQuery* ή *live connection* (βλ. υποενότητα **3.3.1**). Η φόρτωση των αρχείων .csv γίνεται από την κορδέλα στο πάνω μέρος ,όπως φαίνεται στην **Εικόνα 35**.

Το τελικό μοντέλο δεδομένων περιέχει τα αρχεία που δόθηκαν από την εταιρεία (***Brazilian E-commerce Dataset , Marketing Dataset***),ενώ προστέθηκαν κάποια επιπλέον είτε για τον εμπλουτισμό του μοντέλου είτε για σκοπούς της παρουσίασης. Για παράδειγμα , δημιουργήθηκαν μέσω module (**faker**) βιβλιοθήκης της γλώσσας *προγραμματισμού Python 3.0*, «πλασματικά» στοιχεία εργαζομένων της εταιρείας αλλά και των επιχειρήσεων που είναι

εγγεγραμμένοι στην πλατφόρμα της εταιρείας. Αυτό έγινε καθώς, τα δεδομένα της εταιρείας ήταν ανώνυμα και περιείχαν απλώς κωδικούς-κλειδιά ,αντί ονομάτων ,τα οποία δεν είναι ευπαρουσίαστα. Επίσης , προστέθηκαν ο πίνακας-διάσταση των Ημερομηνιών (*[Date]*) και ένας πίνακας που δείχνει την διακύμανση των τιμών των καυσίμων για την Βραζιλία στην πάροδο των χρόνων (Fuel_data.csv).

Brazilian E-commerce Dataset

- olist_customers_dataset.csv
- olist_geolocation_dataset.csv
- olist_order_items_dataset.csv
- olist_order_payments_dataset.csv
- olist_order_reviews_dataset.csv
- olist_orders_dataset.csv
- olist_products_dataset.csv
- olist_sellers_dataset.csv
- product_category_name_translation.csv

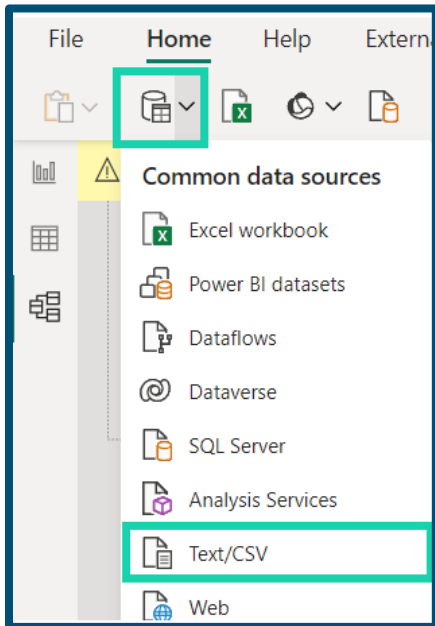
Marketing Dataset

- olist_closed_deals_dataset.csv
- olist_marketing_qualified_leads.csv

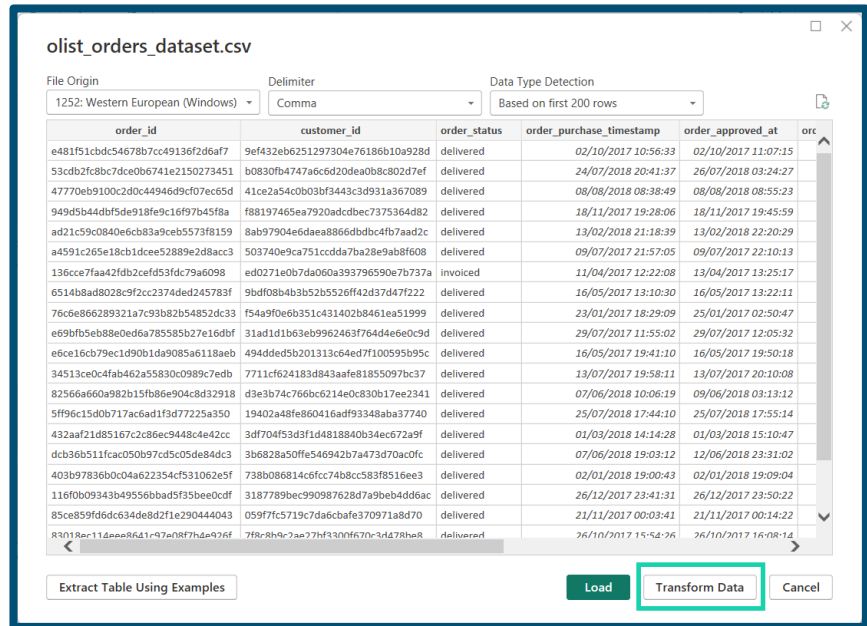
Βοηθητικοί Πίνακες

- Customer_data.csv
- Sales_Team.csv
- Sales_Dev_Team.csv
- Sellers Details.csv
- Fuel_data.csv
- Date (Blank Query)

Πίνακας 11 : Τα δεδομένα που εισάγονται στο Power BI Desktop

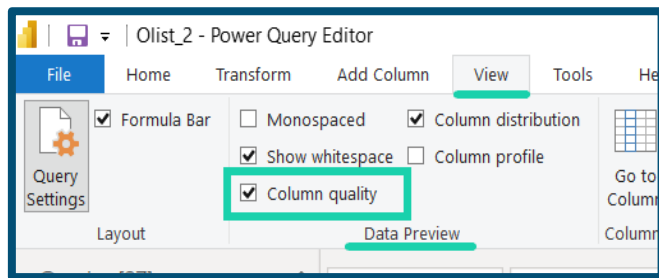


Εικόνα 35: Εισαγωγή .csv αρχείου

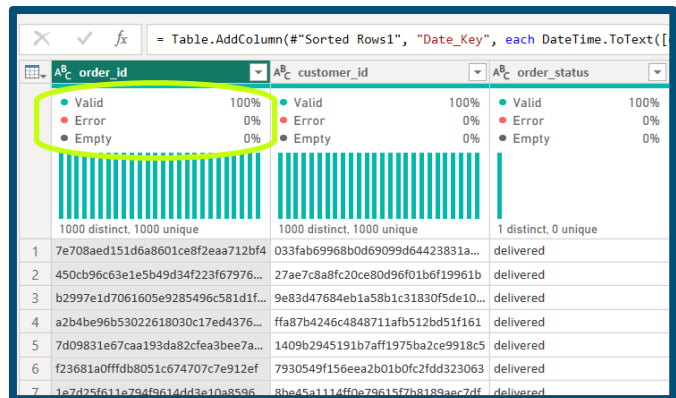


Εικόνα 36 : Μετάβαση στο Power Query

Το πρώτο αρχείο που κάνουμε εισαγωγή στο **Power Query Editor** είναι αυτό με τις παραγγελίες *olist_orders_dataset.csv*. Παρατηρούμε ότι μπορούμε να αξιολογήσουμε την ποιότητα των δεδομένων μας, μέσω της προ-απεικόνισης δεδομένων που προσφέρει το **Power Query Editor**. Το εντοπίζει κανείς από την καρτέλα *View/Data Preview/Column Quality*.



Εικόνα 37 : Επιλογή λειτουργίας Column Quality – Power Query Editor



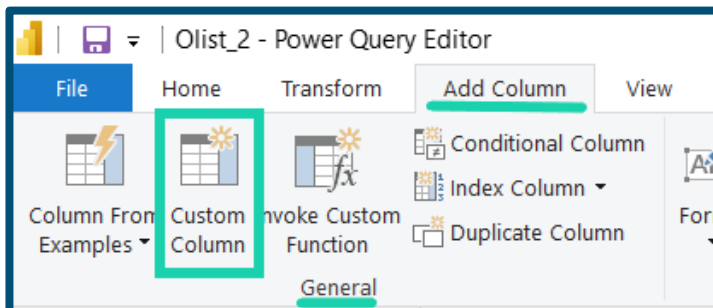
Εικόνα 38 : Ποιότητα δεδομένων στήλης – Power Query Editor

Τα δεδομένα μας δεν περιέχουν κενά (*Empty 0%*) και είναι χωρίς λάθη (*Error 0%*), πρόκειται άρα για «καθαρά» ,πλήρη και έτοιμα προς ανάλυση δεδομένα.

Ο πίνακας αυτός περιέχει :

- στήλες-κλειδιά
`Orders[order_id]`,
`Orders[customer_id]`
- και στήλες που σχετίζονται με τα logistics κάθε παραγγελίας
`Orders[order_status]`
`Orders[order_purchase_stamp]`,
`Orders[order_approved_at]`,
`Orders [order_delivered_carrier_date]`,
`Orders[order_delivered_customer_date]`,
`Orders[order_estimated_delivery_date]`

Μετά την εισαγωγή του αρχείου παραγγελιών ,είναι η ώρα να το επεξεργαστούμε. Θα προσθέσουμε στο αρχείο αυτό μια στήλη-κλειδί κατάλληλη για να συνδέσουμε στο μοντέλο μας ,σε επόμενο βήμα, την βασική διάσταση της Ημερομηνίας (**Date Dimension**). Στην κατάλληλη καρτέλα βρίσκεται η εντολή Add Colum/General/**Custom Column**.

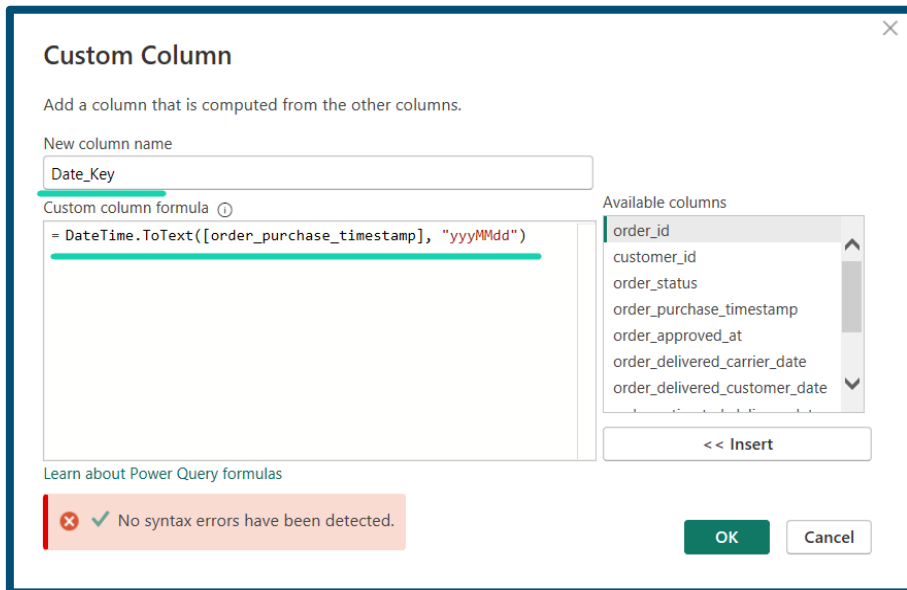


Εικόνα 39 : Επιλογή Custom Column – Power Query Editor

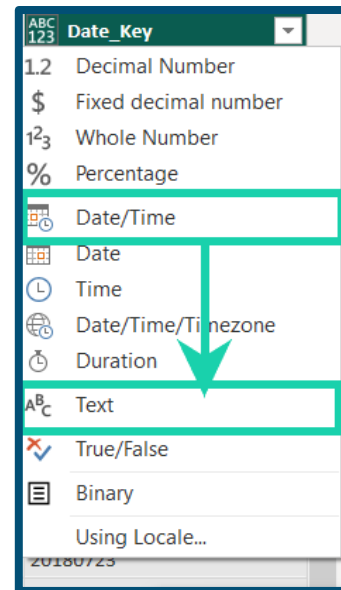
Για να δημιουργήσουμε την νέα στήλη απαιτείται να γράψουμε στην γλώσσα του **Power Query M** , στην απλοποιημένη μορφή της. Ουσιαστικά πρόκειται για μια γλώσσα , σχεδιασμένη για να φιλτράρει , να συνδυάζει και να ενσωματώνει τα δεδομένα του **Power BI Desktop**. Εν προκειμένω, για την δημιουργία της στήλης που θα περιέχει το κλειδί ημερομηνίας (*Date_Key*) ,

χρησιμοποιήσαμε μια άλλη στήλη του αρχείου ,την `[order_purchase_timestamp]`, με βάση την οποία μετατρέψαμε την ημερομηνία σε κείμενο. Αυτή η στήλη είναι μια στήλη με δεδομένα τύπου *ημερομηνίας/ώρας* και δηλώνει την ημερομηνία και ώρα καταχώρισης της παραγγελίας του πελάτη στο σύστημα της εταιρείας. Η εντολή που χρησιμοποιήσαμε ήταν :

```
Date_Key = DateTime.ToText([order_purchase_timestamp], "yyyMMdd")
```

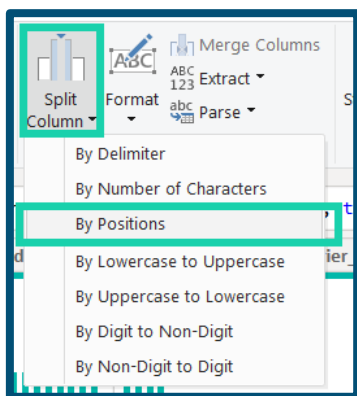


Εικόνα 41 : Δημιουργία στήλης Date_Key – Power Query

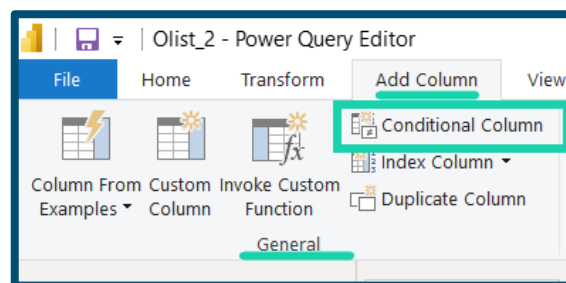


Εικόνα 40 : Μετατροπή από Date/Time σε τύποText

Στη συνέχεια, η εταιρεία στα πλαίσια χαρτογράφησης του προφίλ των πελατών ,επιθυμεί να γνωρίζει τις ώρες προτιμούν οι πελάτες να κάνουν τις αγορές τους. Εκ νέου με βάση , την στήλη *[order_purchase_timestamp]*, θα δημιουργήσουμε μια νέα στήλη που θα υποδηλώνει το χρονικό διάστημα αγοράς (πρώι, μεσημέρι ,απόγευμα ,βράδυ). Δημιουργούμε ένα αντίγραφο της στήλης *[order_purchase_timestamp]* και στην συνέχεια χωρίζουμε το αντίγραφο σε 2 στήλες με την κατάλληλη εντολή στην καρτέλα Transform/Text Column/**Split Column by Positions** ,η πρώτη περιέχει την ημερομηνία αγοράς και η δεύτερη την ώρα.



Εικόνα 43 : Διαχωρισμός στήλης – Power Query Editor



Εικόνα 42 : Αλγοριθμική στήλη – Conditional Formatting – Power Query Editor

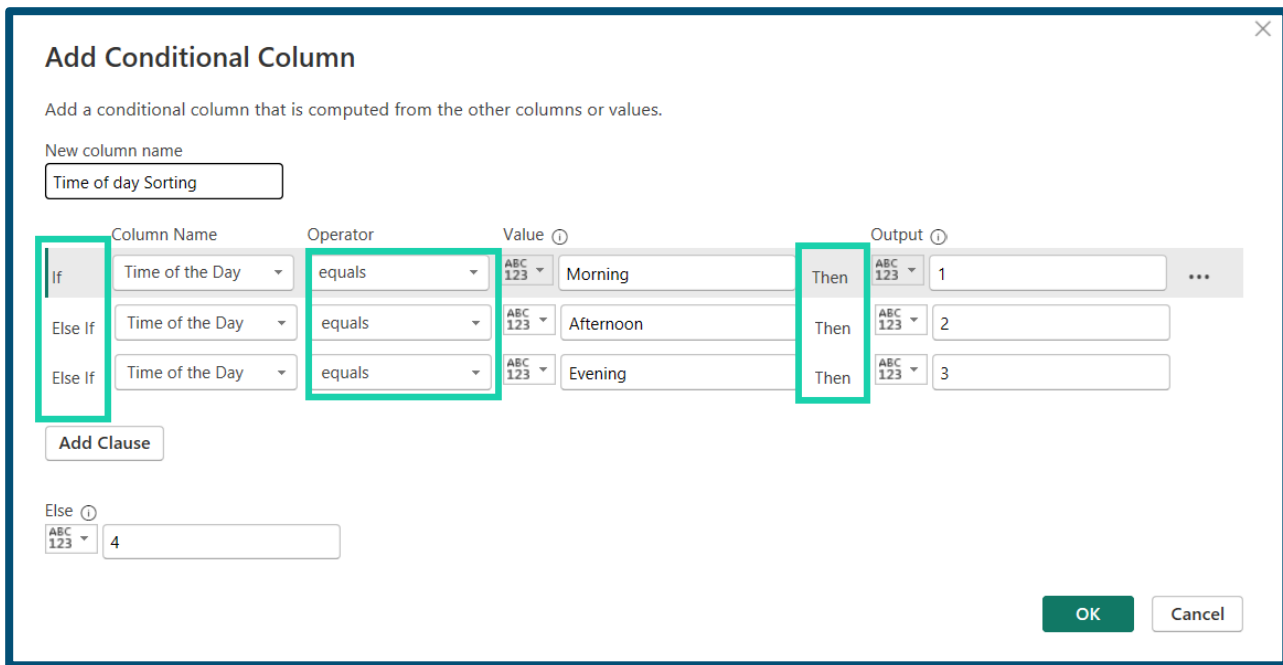
Η στήλη της ημερομηνίας είναι περιττή και διαγράφεται. Μετονομάζουμε το αντίγραφο που δημιουργήσαμε σε *[Time of the Day]* και αλλάζουμε τον τύπο του σε στήλη με τύπο Ώρας.

Στην συνέχεια , με βάση την στήλη αυτή χρησιμοποιούμε την εντολή του Power Query ,Add Column/General/ **Conditional column**. Με βάση αυτή την εντολή δημιουργούμε την νέα αλγοριθμική στήλη με τα δηλωμένα χρονικά διαστήματα της ημέρας ,όπου γίνεται αντιστοίχιση στις ώρες παραγγελίας. Ο αλγόριθμος εμφανίζεται στην παρακάτω **Εικόνα 22**.

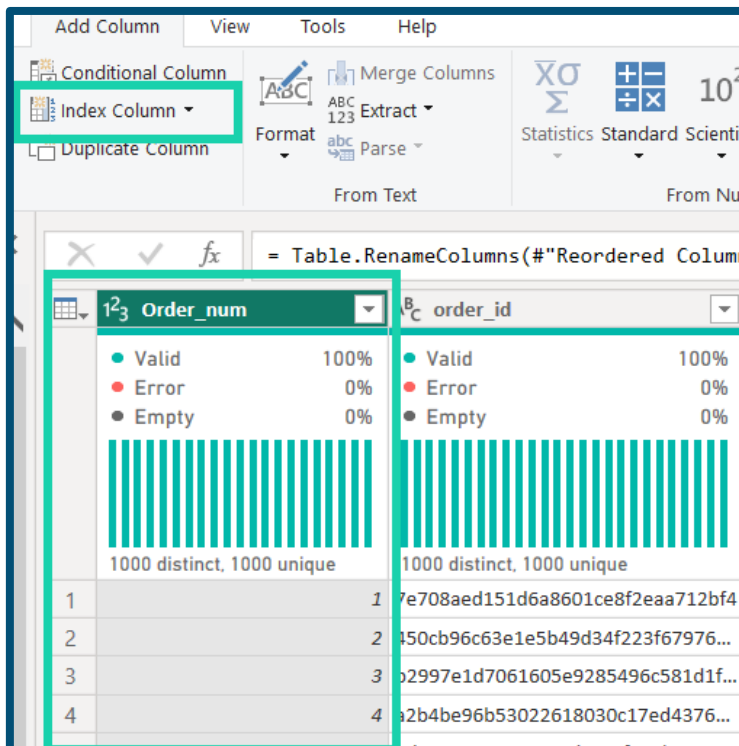
	Column Name	Operator	Value	Output
If	Order_Purchase_t...	is greater than	18:00:00	Evening
Else If	Order_Purchase_t...	is greater than	12:00:00	Afternoon
Else If	Order_Purchase_t...	is greater than	06:00:00	Morning
Else If	Order_Purchase_t...	is greater than	00:00:00	Night

Εικόνα 44 : Ο αλγόριθμος δημιουργίας της στήλης [Time of the Day]

Για τεχνικούς λόγους στην πορεία χρειάστηκε να δημιουργήσουμε ακόμη 2 στήλες . Η πρώτη στήλη αφορά λόγους οπτικοποίησης ,οι οποίοι θα φανούν όταν φτάσουμε στην παρουσίαση, όπου η ταξινόμηση των περιόδων της ημέρας γινόταν αλφαβητικά (πρώτα Afternoon , έπειτα Evening κλπ).Οπότε δημιουργήθηκε ακόμα μια αλγοριθμική στήλη με ονομασία [Time of day Sorting] ,όπως φαίνεται παρακάτω.



Εικόνα 45 : Ο αλγόριθμος δημιουργίας της στήλης [Time of the day Sorting] – Power Query Editor

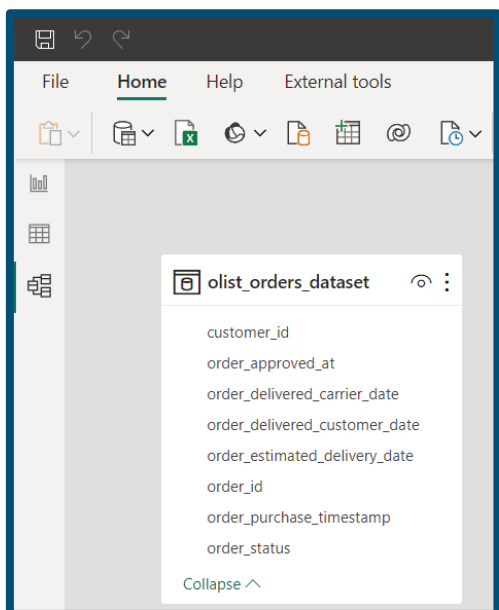


Εικόνα 46 : Δημιουργία στήλης τύπου Index - Power Query Editor

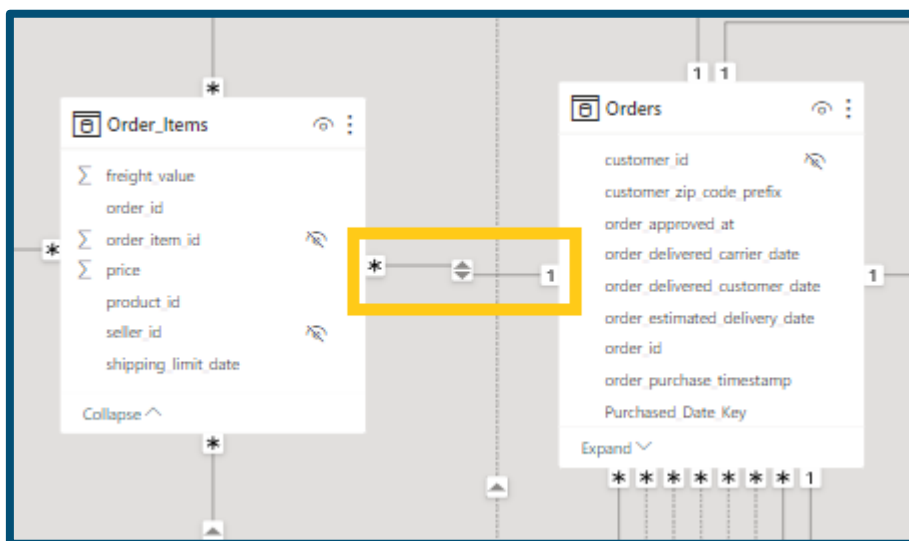
Η δεύτερη στήλη που δημιουργήσαμε είναι μια απλή αρίθμηση για να γίνεται άμεσα αντιληπτός ο αριθμός των σειρών του πίνακα. Αυτό έγινε μέσω της τυποποιημένης εντολής στο Power Query, **Index Column\General\Add Column**.

Πλέον έχουν διαμορφωθεί τα δεδομένα του πίνακα *Orders* οπότε, τα φορτώνουμε στο μοντέλο δεδομένων πατώντας την εντολή *Close & Apply*. Αυτομάτως ενημερώνεται το *Model View* (βλ. ενότητα 3.2) και ο πίνακας

πλέον είναι εμφανής. Στην συνέχεια προσθέτουμε το αρχείο *olist_items_dataset.csv* , το οποίο θα είναι και ο πίνακας Γεγονότων (**Fact table**) του μοντέλου δεδομένων. Αυτός ο πίνακας είναι και ο πυρήνας της σχεδίασης συστημάτων σε σχήμα Αστέρι (*Star Schema* - βλ. υποενότητα **3.4.2**). Και αυτός ο πίνακας δεν χρειάζεται να «καθαριστεί» ,καθώς περιέχει τα δεδομένα ούτε εμπεριέχουν σφάλματα ούτε κενά. Οπότε, απλά το φορτώνουμε στο σύστημα μας και εμφανίζεται στο *Model View*. Το Power BI αυτομάτως δημιουργεί μια σχέση μεταξύ των πινάκων. Η σχέση που επιλέγεται είναι Many-to-one (* - 1) σχέση με την κατεύθυνση του φίλτρου να είναι μονή και να δείχνει προς την πλευρά του πίνακα Γεγονότων (πλευρά Many *). Η σχέση που δημιουργήθηκε αν και σωστή , στην πορεία δημιουργίας του μοντέλου απαιτήθηκε να μετατραπεί η κατεύθυνση μεταφοράς του φίλτρου και προς τις δύο κατευθύνσεις.

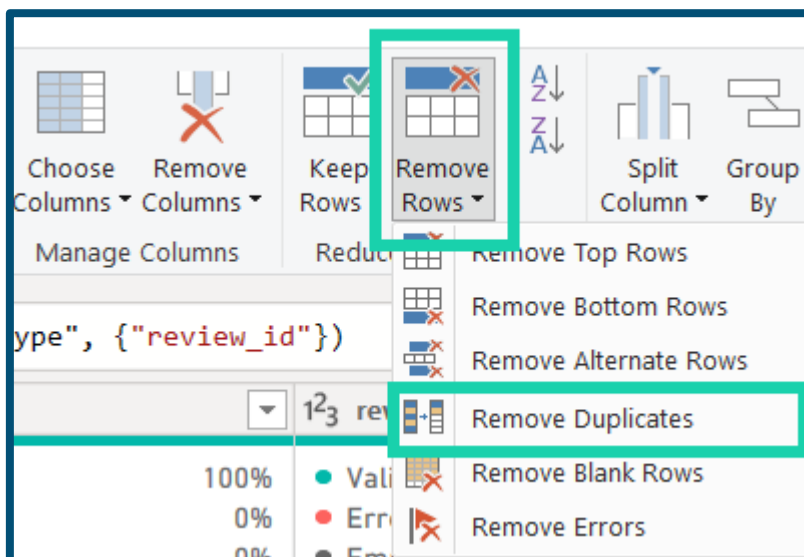


Εικόνα 47 : Φόρτωση πίνακα Orders - Model View



Εικόνα 48 : Σύνδεση πίνακα Order_Items - Model View

Έπειτα συνδέουμε τον πίνακα με τις αξιολογήσεις *olist_reviews_dataset.csv*. Αυτός ο πίνακας χρειάστηκε να αφαιρεθούν ορισμένες διπλότυπες σειρές ,οι οποίες περιείχαν κριτικές που είχαν περαστεί προφανώς περισσότερες από μια φορά.

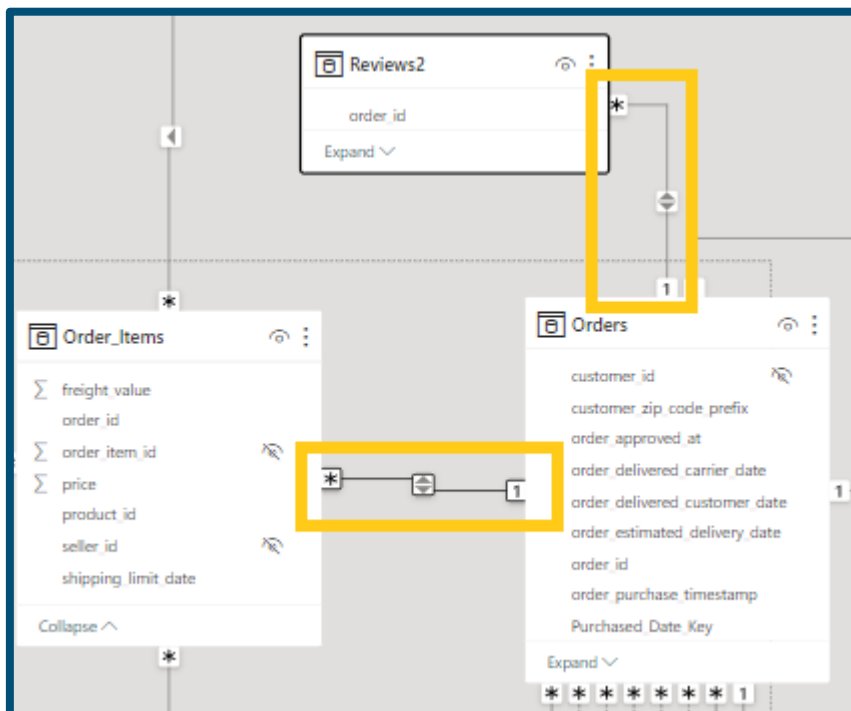


Εικόνα 47 : Αφαίρεση διπλότυπων σειρών πίνακα Reviews – Power Query Editor

Ο τρόπος σύνδεσης των πινάκων γίνεται με διαφορετικό τρόπο απ' ό τι πριν. Ο πίνακας **Reviews** ,όπως μετονομάστηκε, δεν συνδέεται απευθείας με τον πίνακα Γεγονότων αλλά πρώτα με τον πίνακα *Orders*. Στον πίνακα **Reviews** ,κάθε αξιολόγηση δεν αντιστοιχεί σε διαφορετική παραγγελία αναγκαστικά, καθώς κάποιος πελάτης μπορεί να έχει βάλει περισσότερες από μια αξιολογήσεις για την ίδια

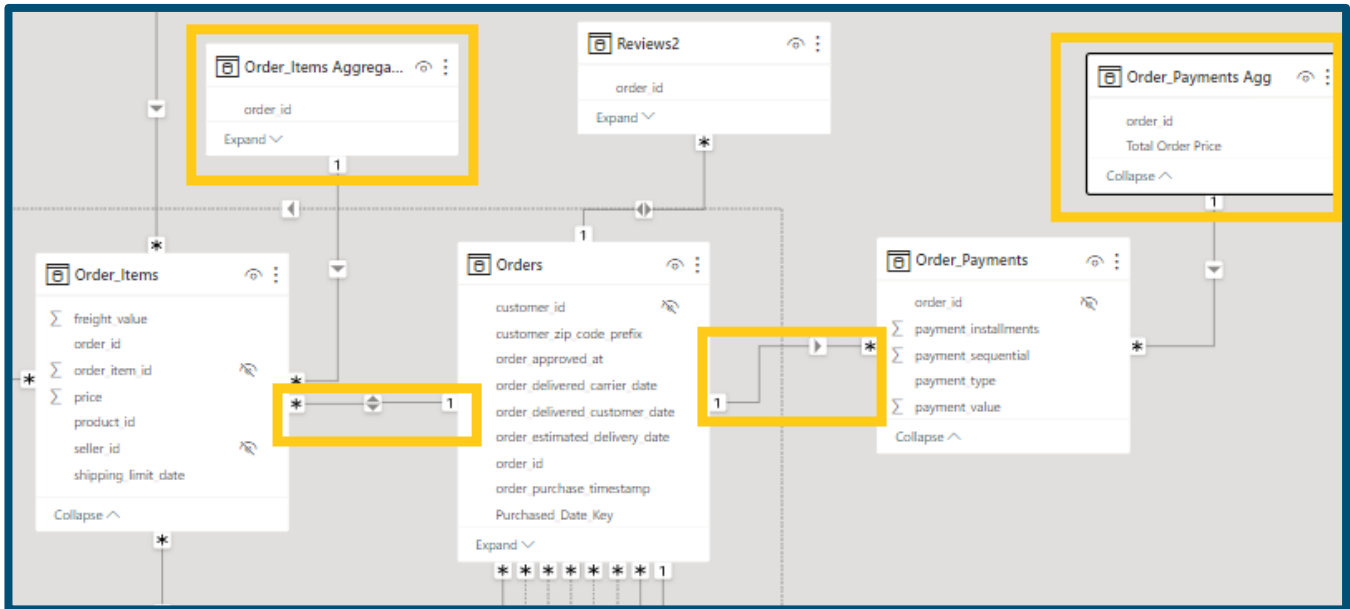
παραγγελία. Το λειτουργικό σύστημα της εταιρείας καταγράφει όλες αξιολογήσεις για το ίδιο προϊόν και απλά αναθέτει ένα διαφορετικό κωδικό για την κάθε αξιολόγηση. Έτσι, αν συνδέαμε τον πίνακα **Reviews** ,στον πίνακα γεγονότων θα δημιουργούσαμε σχέση Many-to-Many (* - *), η οποία σχέση ,αν και πλέον είναι διαδεδομένη στα μοντέλα δεδομένων , δεν αποτελεί την ιδανική σχεδίαση και μπορεί να προκαλέσει λάθος αποτελέσματα στις απεικονίσεις που δημιουργούμε. Έτσι χρησιμοποιούμε την τεχνική του πίνακα «γέφυρα»

(*Bridge Table*), όπου εδώ τον ρόλο της γέφυρας μεταξύ των δύο πινάκων έχει ο πίνακας **Orders**. Η σύνδεση είναι εκ νέου δι-κατευθυντήρια .

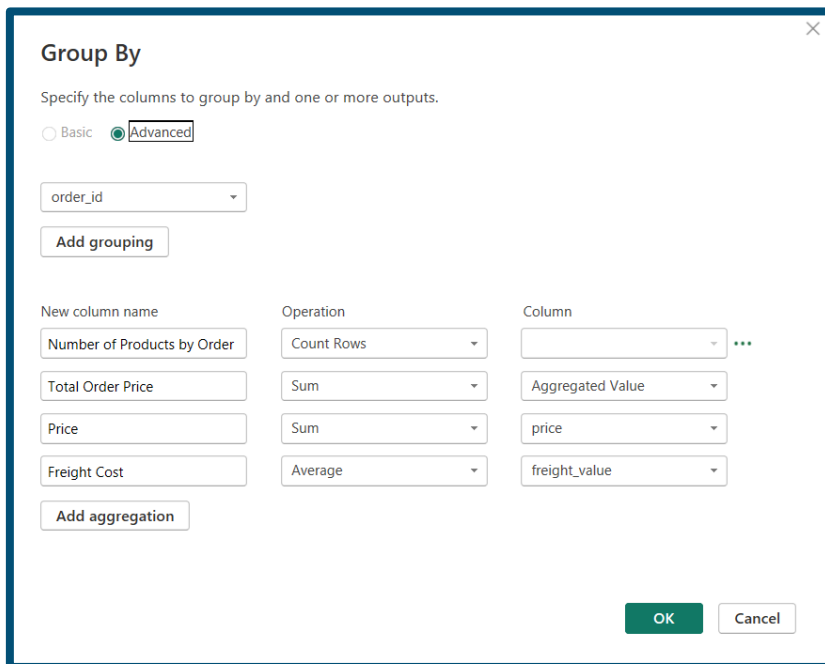


Εικόνα 48 : Σύνδεση πίνακα Reviews με την τεχνική του πίνακα «Γέφυρα» - Model View

Επόμενος πίνακας που συνδέουμε είναι από το αρχείο *olist_order_payments_dataset.csv*. Ξανά ο πίνακας αυτός θα συνδεθεί με ενδιάμεσο πίνακα και όχι απευθείας με τον πίνακα γεγονότων. Ο πίνακας με τις πληρωμές (**Order_Payments**) περιέχει διπλότυπες τιμές του κλειδιού *order_id* (μέσω του οποίου γίνεται η σύνδεση των πινάκων) , καθώς μια παραγγελία μπορεί να πληρωθεί σε δόσεις και έτσι μια παραγγελία εμφανίζεται πολλές φορές στον πίνακα. Για να αποφύγουμε την δημιουργία σχέσης τύπου Many-to-Many (* - *), επαναλαμβάνουμε την ίδια διαδικασία με πριν και δημιουργούμε 2 σχέσεις Many-to-One και One-to-Many.



Εικόνα 49 : Σύνδεση πίνακα Order_Payments στο μοντέλο - Model View

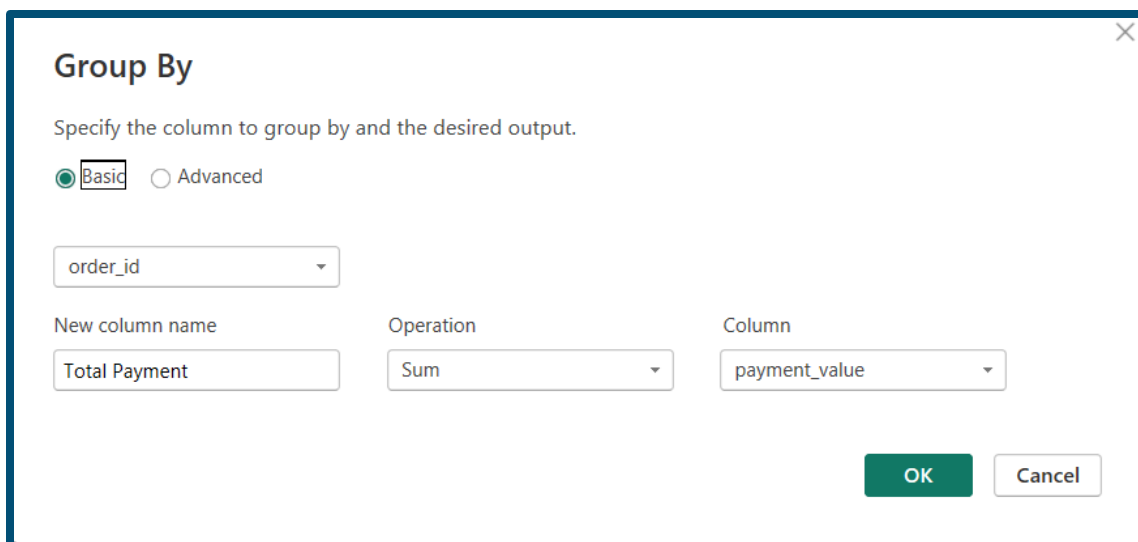


Εικόνα 50 : Ομαδοποίηση των παραγγελιών για την δημιουργία του πίνακα Order_Items_Aggregations – Power Query Editor

παραγγελία ,σε αντίθεση με πριν όπου ,η πληρωμή καταγραφόταν στο σύστημα ανά δόση πληρωμής και όχι την συνολική.

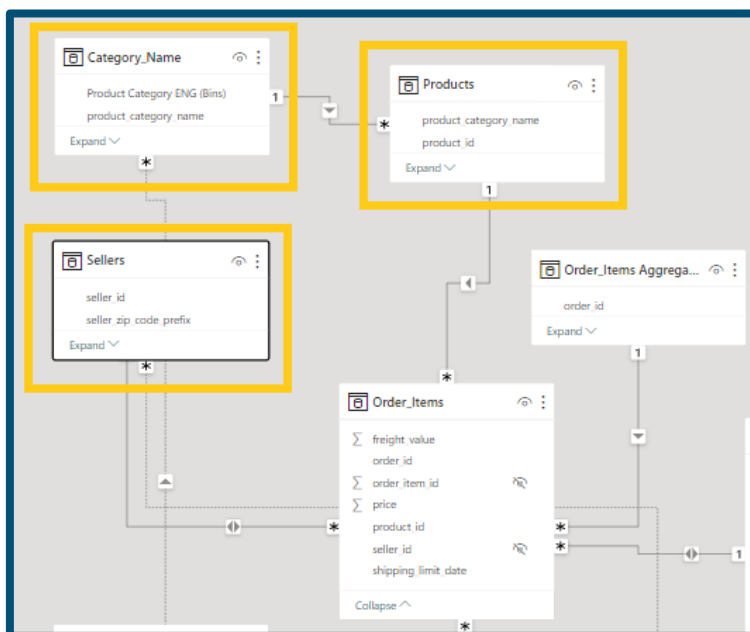
Κάθε παραγγελία διαθέτει διάφορα αντικείμενα και όπως είναι οργανωμένος ο πίνακας **Order_Items** , εμφανίζει τις πληροφορίες ανά προϊόν παραγγελίας και όχι ανά παραγγελία . Έτσι, δημιουργήσαμε τον πίνακα **Order_Items_Aggregations** που εμφανίζεται στην παραπάνω εικόνα και πλέον μπορούμε να βλέπουμε πληροφορίες όπως, το συνολικό κόστος και το κόστος μεταφοράς, ανά παραγγελία και όχι ανά προϊόν παραγγελίας.

Με παρόμοια λογική δημιουργήσαμε και τον πίνακα **Order_Payments_Agg**, δηλαδή αθροίσαμε τις πληρωμές ανά



Εικόνα 51 : Ομαδοποίηση των παραγγελιών για την δημιουργία του πίνακα Order_Payments_Agg – Power Query Editor

Στην συνέχεια, προσθέτουμε τον πίνακα με την περιγραφή των προϊόντων (**Products**), τον πίνακα με τους συνεργάτες – πωλητές της εταιρείας (**Sellers**) και τον πίνακα με την κατηγοριοποίηση των προϊόντων σε ευρείες κατηγορίες (**Category_Name**). Οι πρώτοι δύο πίνακες δεν χρειάστηκαν κάποια επιπρόσθετη επεξεργασία και συνδέθηκαν απευθείας με τον πίνακα γεγονότων **Order_Items** με τις κλασικές σχέσεις Many-to-One, με δι-κατευθυντήριες κατεύθυνση φιλτραρίσματος για τον πίνακα **Sellers**. Ο πίνακας **Category_Name** συνδέθηκε μόνο με τον πίνακα **Products**, σε σχεδίαση που μοιάζει με την σχεδίαση «Χιονονιφάδα» (βλ. υποενότητα 3.4.1). Αυτός ο πίνακας απλά έχει χρησιμότητα μόνο ως προς την κατηγοριοποίηση των προϊόντων, οπότε δεν μπορεί να συνδεθεί με τον πίνακα γεγονότων.



Εικόνα 52 : Προσθήκη πινάκων Products, Sellers, Category Name – Model View

Ο επόμενος πίνακας που θα προστεθεί είναι ένας από τους πλέον βασικούς πίνακες διαστάσεις, ο πίνακας Ημερομηνιών (**Date**). Αυτός πίνακας δημιουργήθηκε στον **Data View**, με την μέθοδο των **Υπολογιζόμενων Στηλών (Calculated Columns)**, γράφοντας σε κώδικα DAX για να δημιουργήσουμε της στήλης του πίνακα.

The screenshot shows the Power BI Desktop interface. The DAX formula bar contains the formula: `Date = CALENDAR(DATE(2016,01,01),DATE(2018,12,31))`. Below the formula bar, a table is displayed with the following columns and data:

Date	Year	Month	Month_Num	Day	Day_Num	Date_Key	Year_Month	Quarter	Year_Month_Num	WeekDay_Num
01/01/2016 00:00:00	2016	Jan	01	Fri	01	20160101	Jan 2016	Q1	01 2016	5
02/01/2016 00:00:00	2016	Jan	01	Sat	02	20160102	Jan 2016	Q1	01 2016	6
03/01/2016 00:00:00	2016	Jan	01	Sun	03	20160103	Jan 2016	Q1	01 2016	7
04/01/2016 00:00:00	2016	Jan	01	Mon	04	20160104	Jan 2016	Q1	01 2016	1
05/01/2016 00:00:00	2016	Jan	01	Tue	05	20160105	Jan 2016	Q1	01 2016	2
06/01/2016 00:00:00	2016	Jan	01	Wed	06	20160106	Jan 2016	Q1	01 2016	3
07/01/2016 00:00:00	2016	Jan	01	Thu	07	20160107	Jan 2016	Q1	01 2016	4
08/01/2016 00:00:00	2016	Jan	01	Fri	08	20160108	Jan 2016	Q1	01 2016	5

Εικόνα 53 : Δημιουργία πίνακα Date – Data View

Οι «υπολογισμοί» των στηλών που γράφτηκαν σε γλώσσα DAX είναι οι εξής:

```

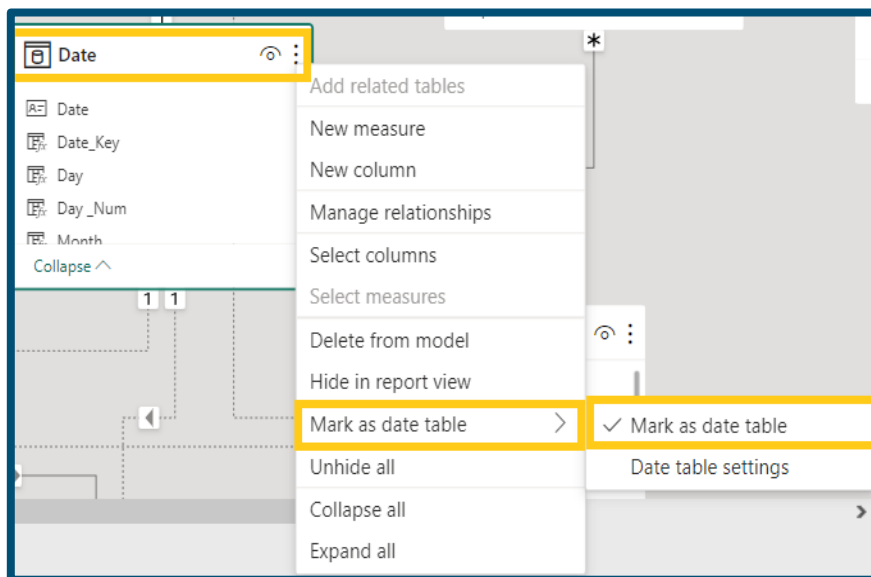
Date = CALENDAR(DATE(2016,01,01),DATE(2018,12,31))
Year = FORMAT('Date'[Date],"YYYY")
Month = FORMAT('Date'[Date],"mmm")
Month_Num = FORMAT('Date'[Date],"MM")
Day = FORMAT('Date'[Date],"ddd")
Day_Num = FORMAT('Date'[Date],"dd")
Date_Key = FORMAT([Date],"yyyyMMdd")
Year_Month = FORMAT('Date'[Date],"MMm yyyy")
Quarter = "Q" & FORMAT(QUARTER('Date'[Date]),"")
Year_Month_Num = FORMAT('Date'[Date],"MM yyyy")
WeekDay_Num = WEEKDAY('Date'[Date],2)

```

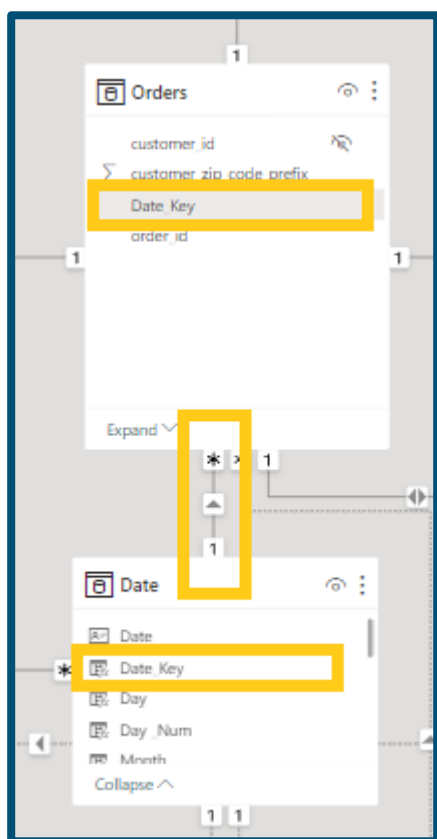
Συνδέσαμε τον πίνακα αυτόν με τον πίνακα **Orders**, και όχι με τον πίνακα γεγονότων ,καθώς σε αυτόν τον πίνακα υπάρχουν δεδομένα χρονικά για τις παραγγελίες .Αυτό είναι θεμιτό ,καθώς και με αυτόν τρόπο τα δεδομένα μπορούν να φιλτραριστούν σωστά. Ο τρόπος σύνδεσης των πινάκων γίνεται μέσω μιας στήλης κλειδί **Orders[Date_Key] = Date[Date_Key]**,όπως αναλύθηκε παραπάνω όταν επεξεργαζόμασταν τον πίνακα **Orders**. Επιλέγουμε να δημιουργήσουμε την νέα στήλη κλειδί **Orders[Date_Key]** ,με βάση την στήλη **Orders[order_purchase_timestamp]**, καθώς αυτή η στήλη περιέχει οπωσδήποτε περισσότερες τιμές από τις υπόλοιπες στήλες ,όπως **[order_approved_at]** ή

[*order_delivered_customer_date*] που εμπεριέχουν ημερομηνίες, εφόσον κάποια παραγγελία μπορεί να μην ολοκληρώσει το δρομολόγιο της και άρα δεν θα περαστεί τιμή στο σύστημα.

Διαφορετικά θα δημιουργούταν πρόβλημα ακεραιότητας των δεδομένων (**referential integrity**). Η διάσταση **Date** αντιμετωπίζεται διαφορετικά από το Power BI και γι' αυτό καλό είναι να επιλέγουμε μια σχετική ρύθμιση ,αν δεν έχει γίνει αυτόματα , όπου επιτρέπει στο σύστημα να αναγνωρίσει ότι υπάρχει τέτοιος πίνακας μέσα στο μοντέλο.



Εικόνα 54 : Ορίζουμε τον πίνακα Date του συστήματος – Model View



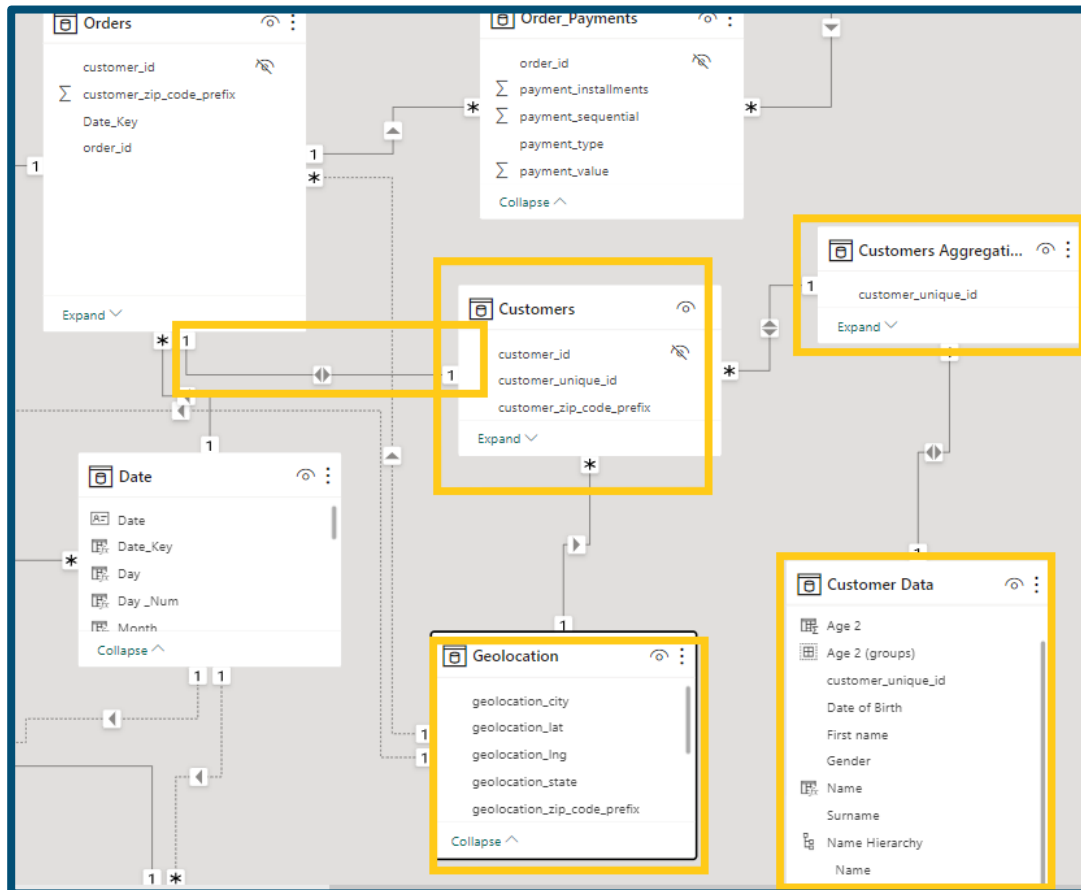
Στην συνέχεια, συνδέουμε τον πίνακα **Customers** , ο οποίος δεν χρειάστηκε κάποια επιπλέον επεξεργασία, με πίνακα **Orders** με σχέση *One-to-One*. Αν και γενικώς αντισυντασσεται να συνδέονται πίνακες με αυτή την σχέση (γιατί μπορεί να ενσωματωθεί με τον συνδεόμενο πίνακα απευθείας), εδώ επιλέγουμε να διατηρήσουμε τη σχέση αυτή , καθώς θα μεγάλωνε αρκετά το πλήθος των στηλών στον πίνακα **Orders** .

Στον πίνακα **Customers** , θα συνδέσουμε τον πίνακα με τα γεωγραφικά δεδομένα **Geolocation** μέσω της στήλης που περιέχει τον ταχυδρομικό κώδικα (**Customers[geolocation_zip_code_prefix] = Geolocation[geolocation_zip_code_prefix]**) με σχέση *on* . Σε αυτόν τον πίνακα χρειάστηκε να κάνουμε μια μικρή επεξεργασία. Ο πίνακας **Geolocation** περιέχει τις γεωγραφικές πληροφορίες για κάθε πελάτη και πωλητή σε «επίπεδο πληροφορίας» (**granularity** ,βλ. υποενότητα **3.4.2**)

Εικόνα 36 : Σύνδεση πίνακα Date στο μοντέλο – Model View

,γεωγραφικών συντεταγμένων. Αυτό το επίπεδο πληροφορίας είναι περισσότερο απ' όσο χρειαζόμαστε ,οπότε ανεβαίνουμε «επίπεδο πληροφορίας» σε επίπεδο ταχυδρομικού κώδικα. Γι' αυτό αποκόπτουμε ορισμένες σειρές του πίνακα **Geolocation**, αφαιρώντας διπλότυπες σειρές που διαθέτουν ίδιο ταχυδρομικό κώδικα στο *Power Query Editor*.

Επίσης για να έχουμε μια πιο ευπαρουσίαστη μορφή, των δεδομένων των πελατών ,όπως αναφέρθηκε και προηγουμένως δημιουργήσαμε κάποια «πλαστά» στοιχεία σε γλώσσα Python τα οποία συνδέθηκαν με τον πίνακα **Customers**, μέσω ενδιάμεσου πίνακα που περιέχει τον μοναδικό κωδικό κάθε πελάτη (**Customers[customer_unique_id]**).



Εικόνα 55 : Σύνδεση πινάκων **Geolocation**, **Customers**, **Customers Aggregation**, **Customer Data** - Model View

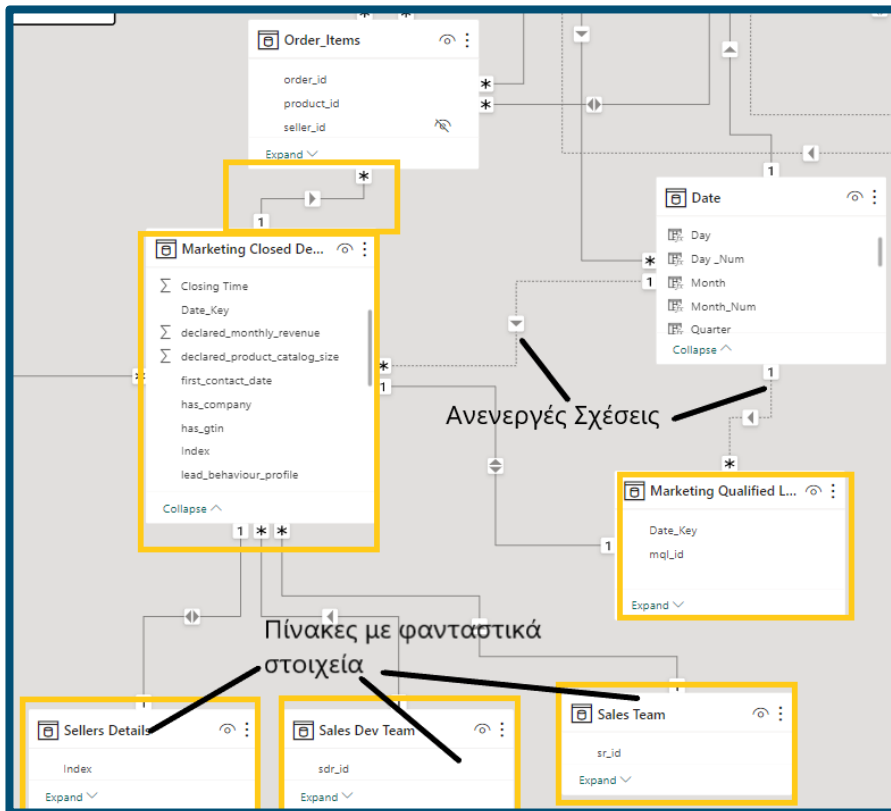
Αξίζει να σημειώσουμε ότι συνδέουμε μέσω ανενεργής σχέσης τον πίνακα **Geolocation** με τον πίνακα **Sellers** μέσω της στήλης **Sellers[geolocation_zip_code_prefix]** = **Geolocation[geolocation_zip_code_prefix]**, καθώς και οι πωλητές έχουν γεωγραφικά δεδομένα, διαφορετικά από αυτά των πελατών. Αυτή η σχέση είναι χρήσιμη παρακάτω όταν

```
Sellers Number = CALCULATE(  
    DISTINCTCOUNT(Sellers[seller_id]),  
    USERELATIONSHIP(Sellers[seller_zip_code_prefix],  
        Geolocation[geolocation_zip_code_prefix])  
)
```

δημιουργήσαμε το *μετρικό (Measure)* για τον υπολογισμό του αριθμού των Πωλητών, όπως φαίνεται παρακάτω.

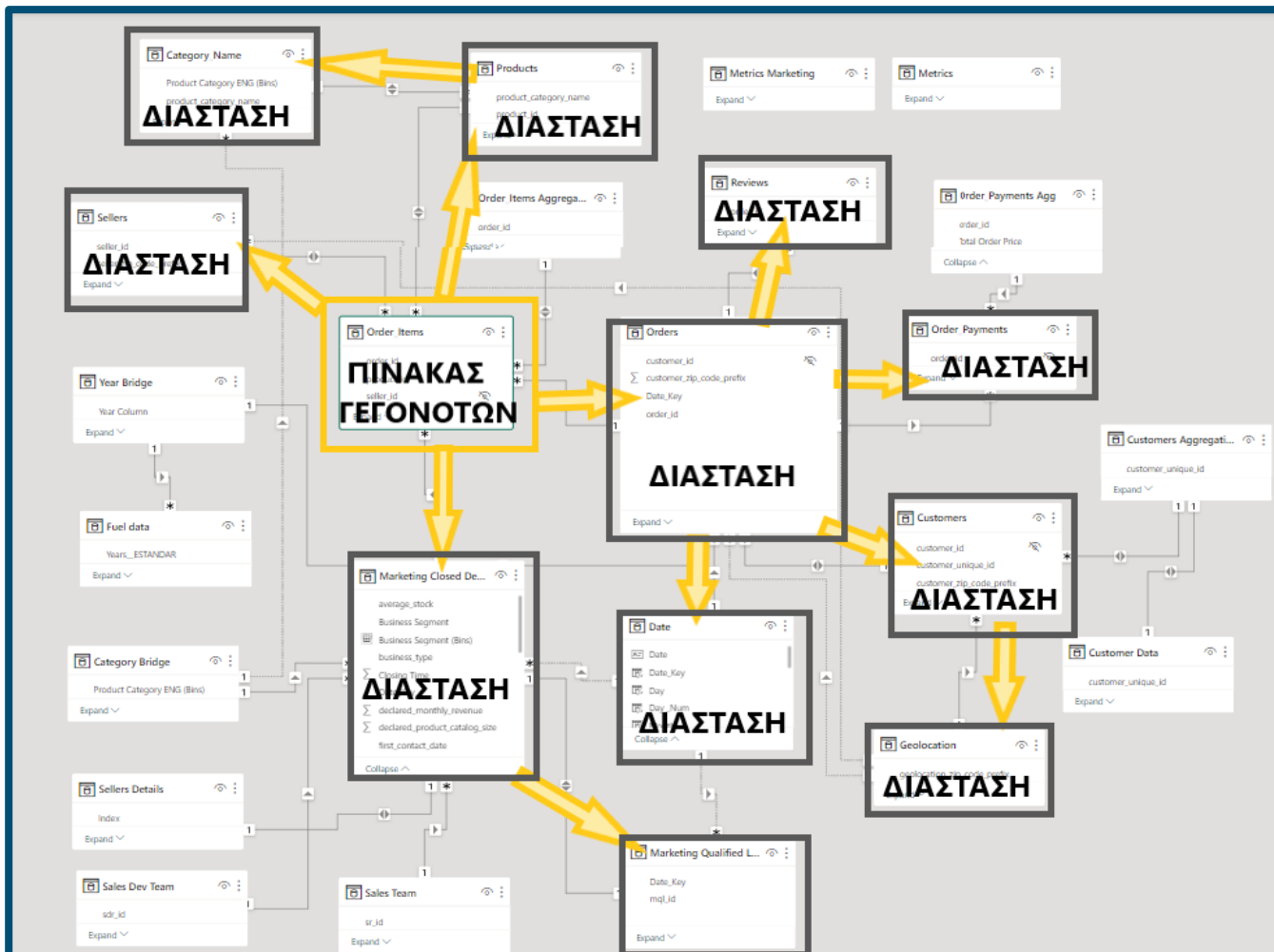
Παρόμοιες ανενεργές σχέσεις έχουν δημιουργηθεί και για άλλους πίνακες ,απλά αναφέρεται εδώ μόνο αυτή η περίπτωση για λόγους μη επαναληψιμότητας. Ο λόγος που δημιουργούμε ανενεργές σχέσεις είναι διότι , όταν συνδέουμε πίνακες φροντίζουμε να μην δημιουργούνται κλειστές διαδρομές σχέσεων στο μοντέλο μας , καθώς αν υπάρχουν κλειστές διαδρομές μετάδοσης φίλτρων ,τα αποτελέσματα που λαμβάνουμε δεν είναι ντετερμινιστικά αλλά τυχαία και άρα ενδεχομένως λάθος.

Στη συνέχεια είναι η στιγμή να προσθέσουμε του πίνακες του Μάρκετινγκ. Ο πίνακας **Marketing Closed Deals** συνδέεται με τον πίνακα γεγονότων , **Order_Items** ,με απλή σχέση Many-to-One με μονή κατεύθυνση μετάδοσης φίλτρου μέσω των στηλών **Marketing Closed Deals[sellers_id] = Order_Items[sellers_id]**. Ο πίνακας **Marketing Qualified Leads** συνδέεται με σχέση One-to-One ,μέσω του κλειδιού *[mql_id]* που δηλώνει την κάθε ξεχωριστή αλληλεπίδραση του πιθανού πελάτη με την εταιρεία.



Εικόνα 57 : Σύνδεση του σετ δεδομένων Μάρκετινγκ - Model View

Στο τελικό μοντέλο προστίθενται και οι βοηθητικοί πίνακες **Year Bridge** , **Category Bridge** και ο πίνακας **Fuel Data**, που θα βοηθήσει σε παραμετρική ανάλυση για τα έξοδα καυσίμων. Το τελικό μοντέλο δεδομένων διαμορφώθηκε σύμφωνα με τους κανόνες τους διαστατικού σχεδιασμού, δηλαδή είναι δομημένο με πυρήνα το πίνακα γεγονότων και περιμετρικά βρίσκονται οι περιγραφικές διαστάσεις. Λόγω της φύσης των στοιχείων που δόθηκαν ο τελικός σχεδιασμός πλησιάζει την σχεδίαση σε σχήμα «Χιονοφιλάδας» (**Snowflake Schema**) και όχι σε σχήμα «Αστεριού». Αυτός ο σχεδιασμός είναι ο δεύτερος βέλτιστος από την άποψη απόδοσης σε ερωτήματα στην βάση (querying). Στο τελικό μοντέλο που φαίνεται στην επόμενη εικόνα , έχουν σημειωθεί μόνο οι βασικές διαστάσεις και ο πίνακας γεγονότων.



Εικόνα 58 : Το τελικό μοντέλο δεδομένων – Model View

4.3 Πρόταση Λύσης στην Υπόθεση Εργασίας

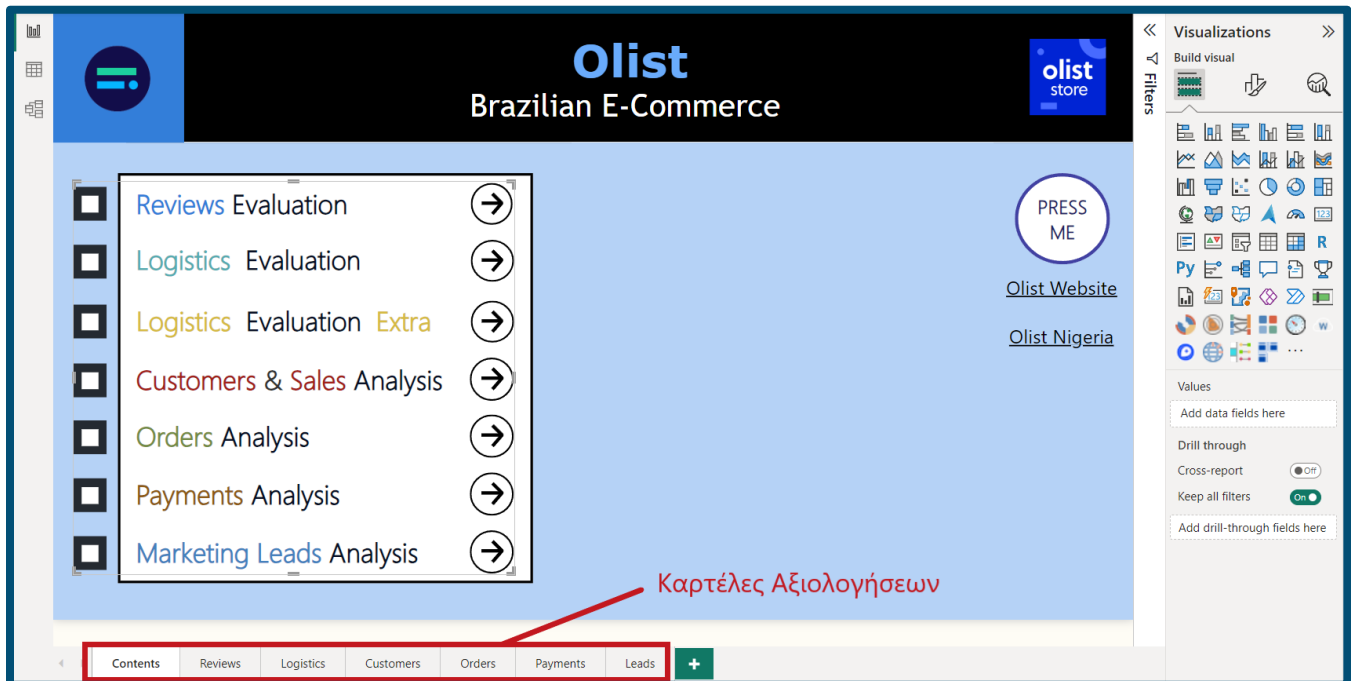
Η αξιολόγηση των διαφόρων δραστηριοτήτων της εταιρείας απαιτεί να γίνει μια πολυεπίπεδη ανάλυση των δεδομένων και απευθύνεται σε χρήστες από διάφορους τομείς της επιχείρησης. Διαφορετικές ανάγκες πρέπει να καλυφθούν για τον Διευθυντή Μάρκετινγκ , διαφορετικές για την Διεύθυνση Εφοδιαστικής Αλυσίδας και επίσης διαφορετικές για την γενική διεύθυνση.

Ένα ολοκληρωμένο σύστημα Επιχειρηματικής Ευφυΐας περιλαμβάνει Αναφορές (*Reports*) , Πίνακες Ελέγχου (*Dashboards*) αλλά και κοινούς διαδικτυακούς χώρους εργασίας (*Workspaces*) ,όπου οι χρήστες της επιχείρησης μπορούν να μοιράζονται τις αναφορές τους και να βλέπουν άλλων μελών του οργανισμού. Το σύστημα πρέπει να δίνει την δυνατότητα οι

χρήστες του να **χρησιμοποιούν κοινές υποδομές και στοιχεία** του συστήματος αλλά και να **θέτει περιορισμούς**, ώστε η πληροφορία να φτάνει στο σωστό κοινό και να είναι οργανωμένη.

Το οπτικό μέρος της λύσης που δημιουργήσαμε περιλαμβάνει διαδραστικούς πίνακες, διαγράμματα και άλλες οπτικοποιήσεις οργανωμένα σε καρτέλες μιας συγκεντρωτικής αναφοράς .Σε αυτήν την αναφορά υπάρχουν οι εξής καρτέλες :

- Αξιολόγηση των Κριτικών
- Αξιολόγηση Εφοδιαστικής Αλυσίδας (Logistics)
- Αξιολόγηση Πελατών
- Αξιολόγηση Παραγγελιών και Πωλητών της πλατφόρμας
- Αξιολόγηση Πληρωμών
- Αξιολόγηση Μάρκετινγκ



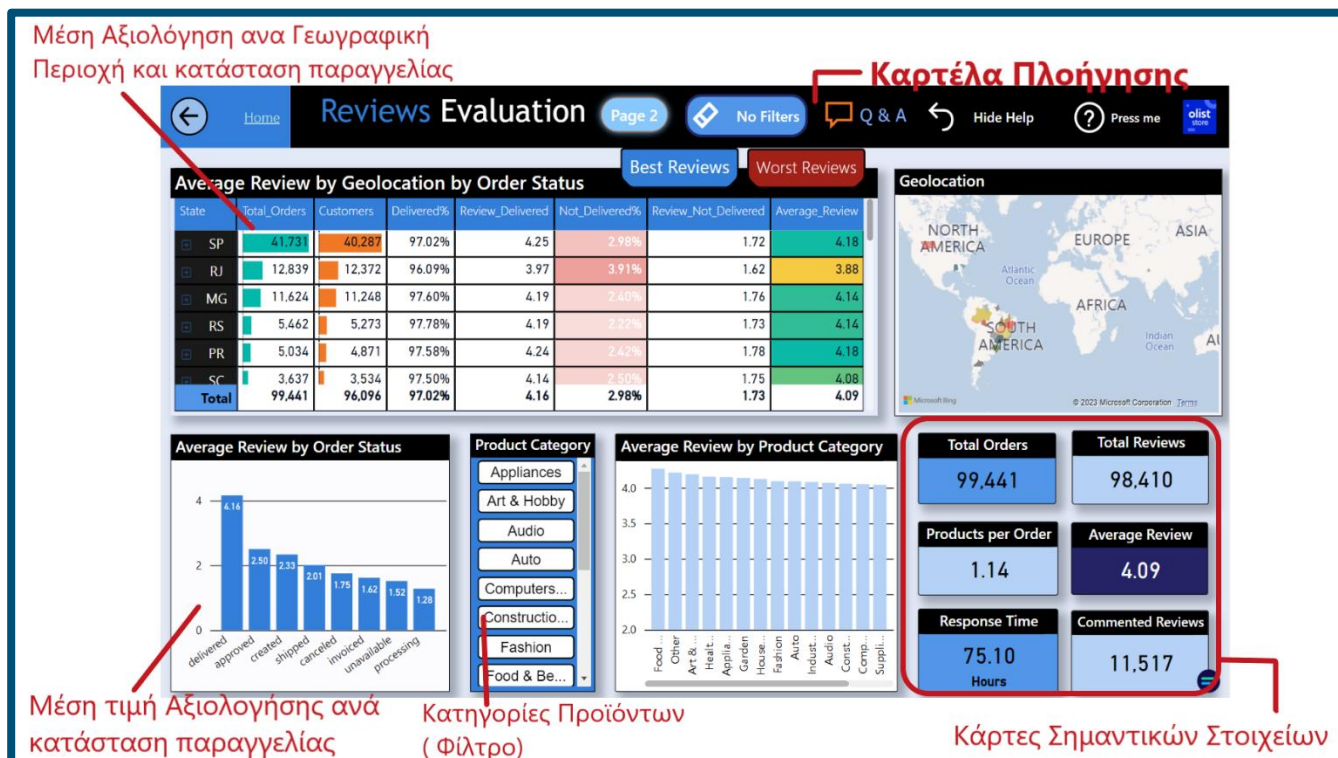
Εικόνα 59 : Περιεχόμενα Αναφοράς – Report View

4.3.1 Παρουσίαση Καρτελών Αξιολόγησης

Αξιολόγηση Κριτικών - Reviews Evaluation

Ένα βασικό σημείο ενδιαφέροντος που αφορά την Διεύθυνση του Μάρκετινγκ και την

Διοίκηση του οργανισμού είναι οι Αξιολογήσεις που αφήνουν οι πελάτες για τις παραγγελίες που πραγματοποιούν.



Εικόνα 60 : Καρτέλα Αξιολόγησης των Κριτικών

Στην καρτέλα αξιολόγησης των κριτικών ,όπως και σε κάθε καρτέλα, έχει προστεθεί και μια μπάρα-καρτέλα πλοήγησης για να μπορεί να κάνει την περιήγηση του χρήστη πιο απλή και κατανοητή. Στην καρτέλα αυτή υπάρχουν διαδραστικά κουμπιά όπου, πληροφορούν το χρήστη για το περιεχόμενο της καρτέλας ,αφαιρούν οποιοδήποτε φίλτρο έχει προστεθεί από τον χρήστη ,τον οδηγούν σε άλλες καρτέλες. Επί πλέον ,αν ο χρήστης θέλει να δει επιπλέον διαγράμματα μπορεί να αξιοποιήσει το κουμπί Q&A , το οποίο υποστηρίζεται από τεχνητή νοημοσύνη και απαντάει σε ερωτήσεις οι οποίες δεν καλύπτονται από την συγκεκριμένη καρτέλα

Η βασική συλλογιστική πίσω από τον σχεδιασμό αυτής της καρτέλας είναι να μπορεί ο χρήστης να δει την **μέση τιμή αξιολογήσεων** , ανά κάποιο ή κάποια πεδία .Για παράδειγμα , ο πίνακας με τίτλο **“Average Review by Geolocation by Order Status”** παρέχει στον χρήστη την εμποτική εικόνα της μέσης αξιολόγησης ανά γεωγραφικό διαμέρισμα ,αλλά και ανά την κατάσταση της παραγγελίας (παράδοση , όχι παράδοση ,ακυρωμένη , εγκεκριμένη κλπ.). Τα κελιά του πίνακα χρωματίζονται αναλόγως με τον αριθμό των παραγγελιών ,το ποσοστό των παραδομένων παραγγελιών και την μέση τιμή αξιολόγησης. Πατώντας το κάθε κελί με τον κέρσορα ,μπορεί να φιλτράρει το περιεχόμενο του πίνακα και της καρτέλας με αντίστοιχο τρόπο.

State	Total_Orders	Customers	Delivered%	Review_Delivered	Not_Delivered%	Review_Not_Delivered	Average_Review
SP	41,731	40,287	97.02%	4.25	2.98%	1.72	4.18
RJ	12,839	12,372	96.09%	3.97	3.91%	1.62	3.88
MG	11,624	11,248	97.60%	4.19	2.40%	1.76	4.14
RS	5,462	5,273	97.78%	4.19	2.22%	1.73	4.14
PR	5,034	4,871	97.58%	4.24	2.42%	1.78	4.18
SC	3,637	3,534	97.50%	4.14	2.50%	1.75	4.08
BA	3,371	3,268	96.35%	3.93	3.65%	1.84	3.86

Εικόνα 61 : Πίνακας “Average Review by Geolocation by Order Status “– Reviews Καρτέλα

Πατώντας πάνω για παράδειγμα πάνω στην στήλη *Total Orders* , τα δεδομένα ταξινομούνται ανά αριθμό παραγγελιών ανά γεωγραφική περιοχή. Ο χρήστης πληροφορείται άμεσα για το ποιες γεωγραφικές περιοχές έχουν το μεγαλύτερο όγκο παραγγελιών, ποιες είναι οι αξιολογήσεις των πελατών σε αυτές τις περιοχές και ποιο ποσοστό των παραγγελιών έχουν φτάσει στους πελάτες. Αν επιθυμεί ο χρήστης να μάθει ποιες περιοχές έχουν την χειρότερη αξιολόγηση ,μπορεί να πατήσει πάνω στον τίτλο της στήλης *Average_Review* και τα αποτελέσματα ταξινομούνται.

State	Total_Orders	Customers	Delivered%	Review_Delivered	Not_Delivered%	Review_Not_Delivered	Average_Review
RR	46	45	89.13%	3.90	10.87%	1.20	3.61
AL	412	400	96.12%	3.85	3.88%	1.50	3.76
MA	743	722	96.10%	3.84	3.90%	1.93	3.77
SE	349	341	95.70%	3.90	4.30%	1.60	3.80
PA	972	946	97.12%	3.91	2.88%	1.79	3.85
CE	1,332	1,310	95.72%	3.95	4.28%	1.75	3.86
BA	3,371	3,268	96.35%	3.93	3.65%	1.84	3.86

Εικόνα 62 : Πίνακας “Average Review by Geolocation by Order Status “ - Reviews Καρτέλα

Αξιοσημείωτα στοιχεία της αναφοράς έχουν τοποθετεί σε κάρτες στο κάτω δεξιά μέρος της καρτέλας. Αυτά τα στοιχεία αφορούν τον **συνολικό Αριθμό Παραγγελιών** , τις **συνολικό αριθμό Αξιολογήσεων** , τον **μέσο όρο Προϊόντων ανά Παραγγελία**, την **Μέση Αξιολόγηση** ,ο χρόνος που χρειάζεται η εταιρεία για να απαντήσει σε σχόλια των πελατών και τέλος το ποσοστό των αξιολογήσεων που περιέχουν σχόλια από τους πελάτες .Ενώ σε

διαφορετική οπτικοποίηση χρωματίζεται ο παγκόσμιος χάρτης αναλόγως το φιλτράρισμα των γεωγραφικών δεδομένων.

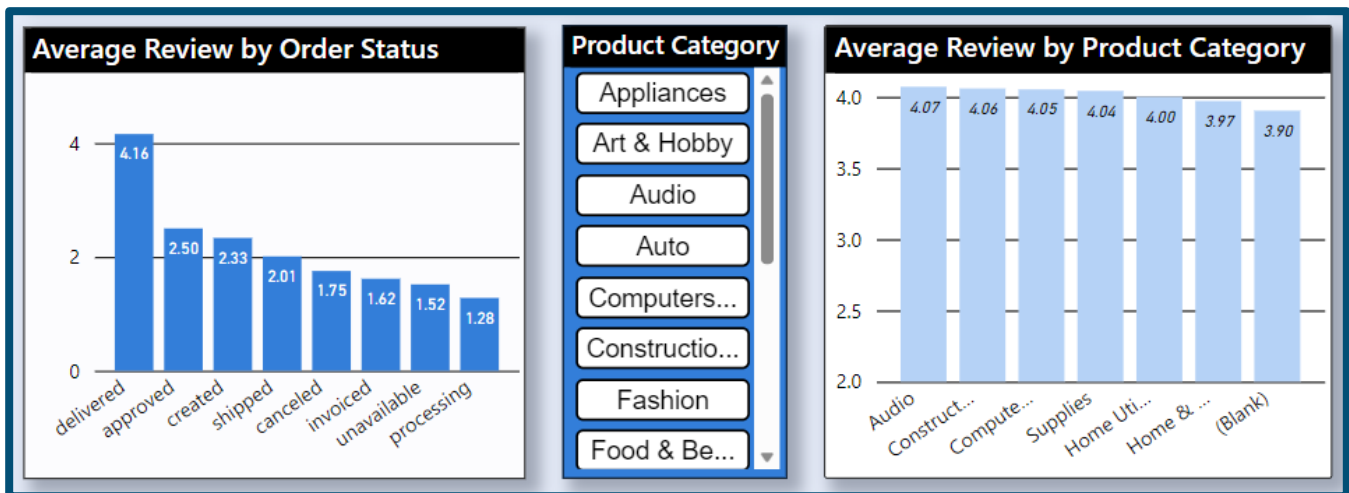


Εικόνα 64 : Οπτικοποίηση Γεωγραφικών Δεδομένων – Reviews Καρτέλα



Εικόνα 63 : Κάρτες αξιοσημείωτων δεδομένων – Reviews Καρτέλα

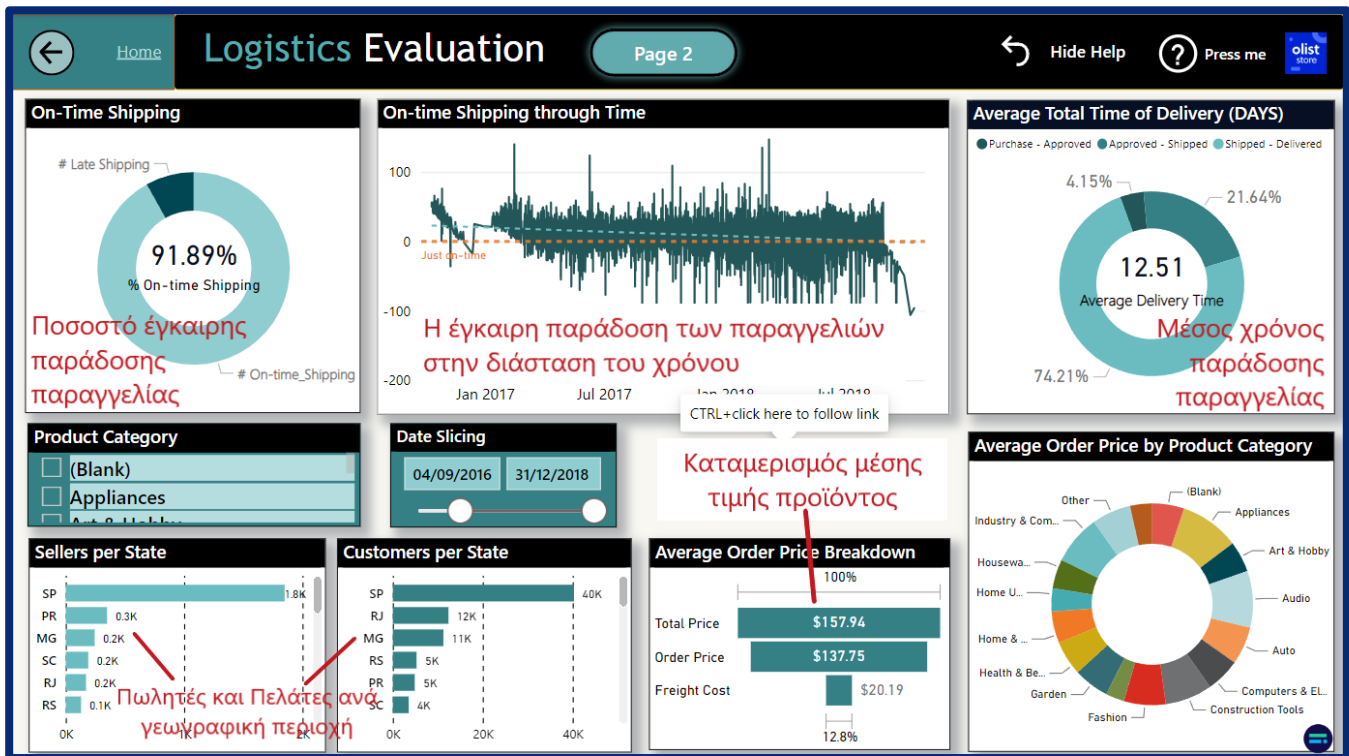
Επίσης έχουν προστεθεί και διαγράμματα τύπου *Cluster columns* για την εύκολη οπτικοποίηση των κατηγοριών στην κατάσταση παραγγελίας και στην κατηγορία προϊόντος.



Εικόνα 65 : Cluster Columns – Reviews Καρτέλα

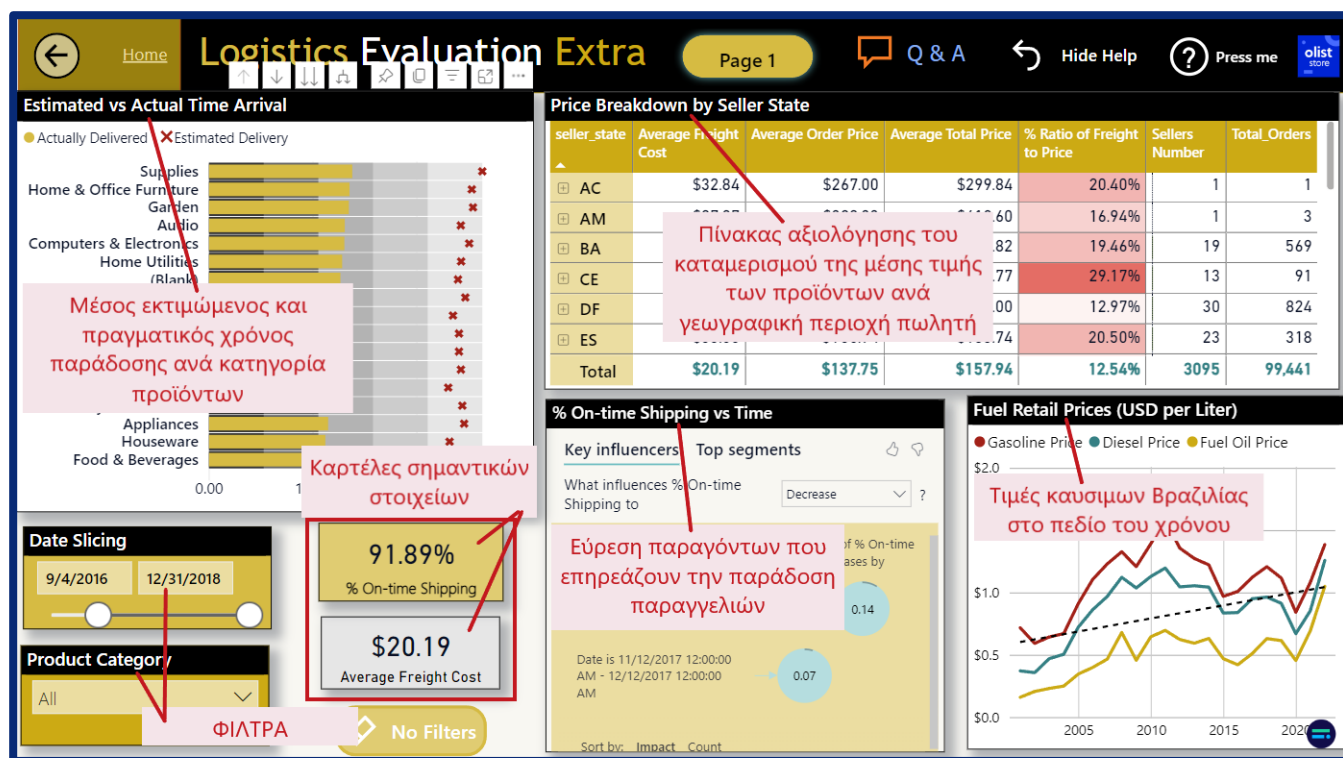
Έχει επίσης, δημιουργηθεί και μια δεύτερη καρτέλα σχετιζόμενη με αξιολογήσεις και ο χρήστης μπορεί να την επισκεφθεί πατώντας πάνω στο σχετικό κουμπί στην καρτέλα περιήγησης. Σε αυτή την καρτέλα έχει προστεθεί μια οπτικοποίηση, όπου με βάση τις γραπτές αξιολογήσεις των πελατών, εμφανίζει ένα σμήνος λέξεων με τις πιο συχνά

εξαιρετικά σημαντική και είναι απαραίτητη η παρακολούθηση του από τα στελέχη της εταιρείας. Αυτή η αξιολόγηση γίνεται φυσικά τόσο στο πεδίο του χρόνου όσο και ανά γεωγραφική περιοχή.



Εικόνα 67 : Καρτέλα Αξιολόγησης Εφοδιαστικής Αλυσίδας - Logistics Evaluation

Επίσης, δεύτερη καρτέλα όπου εστιάζει στο κόστος μεταφορικών τελών των προϊόντων ,μπορούμε να συσχετίσουμε το ποσοστό έγκαιρης παράδοσης, το μέσο μεταφορικό κόστος της παραγγελίας και το ποσοτή των μεταφορικών σε σχέση με την τιμή του προϊόντος, μεταξύ την πολιτεία του πωλητή και την πολιτεία του πελάτη και να βγάλουμε χρήσιμα συμπεράσματα.



Εικόνα 68 : Καρτέλα 2 Αξιολόγησης Εφοδιαστικής Αλυσίδας- Logistics Evaluation 2

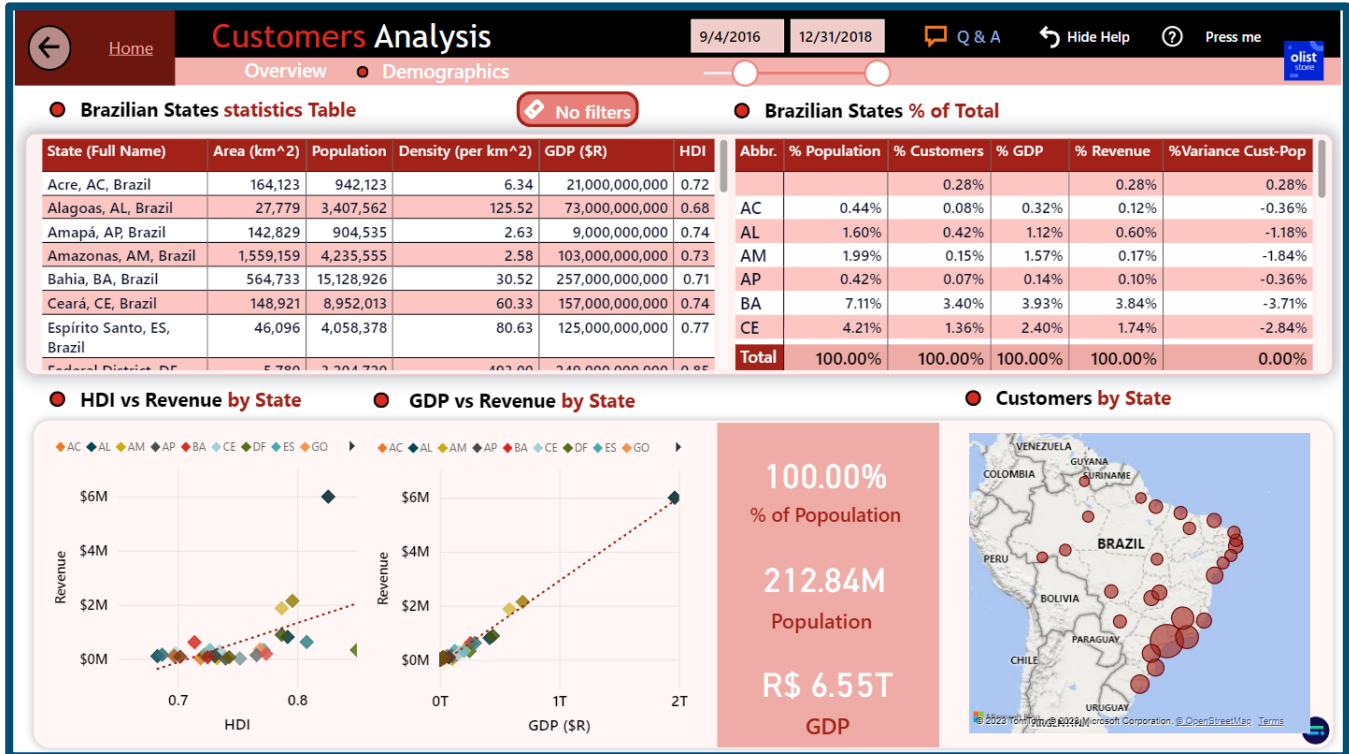
Ανάλυση Πελατών – Customers Analysis

Η αξιολόγηση του προφίλ των πελατών είναι ένα από τα βασικά αντικείμενα αναζήτησης και επένδυσης σε κάθε εταιρεία, η οποία επιθυμεί να αυξήσει τα έσοδα της επενδύοντας τους πόρους της με σωστό τρόπο. Σε όρους Μάρκετινγκ συνήθως αποκαλείται χαρτογράφηση του προφίλ των πελατών και βοηθάει την Τμηματοποίηση της Αγοράς (**Market Segmentation**). Στοιχεία για το προφίλ των πελατών βρίσκονται σε αυτή την καρτέλα, όπως οι κατηγορίες των προϊόντων που επιλέγουν οι πελάτες αναλόγως, το φύλο τους, την ηλικία τους, την πολιτεία που μένουν, τον αριθμό των παραγγελιών που πραγματοποιούν στην πλατφόρμα της εταιρείας και την μέση τιμή των παραγγελιών αυτών. Εξαιρετικά χρήσιμη είναι η **πολυπαραμετρική ανάλυση** που συσχετίζει την μέση τιμή της παραγγελίας (AOV - *Average Order Value*), με τον αριθμό των παραγγελιών και τα έσοδα της εταιρείας ανά κατηγορία προϊόντων. Η ανάλυση αυτή γίνεται τόσο υπό την μορφή πίνακα όσο και διαγράμματος τύπου **Radar** στο κάτω αριστερό μέρος της αξιολόγησης.



Εικόνα 69 : Καρτέλα Αξιολόγησης Πελατών – Customers Analysis

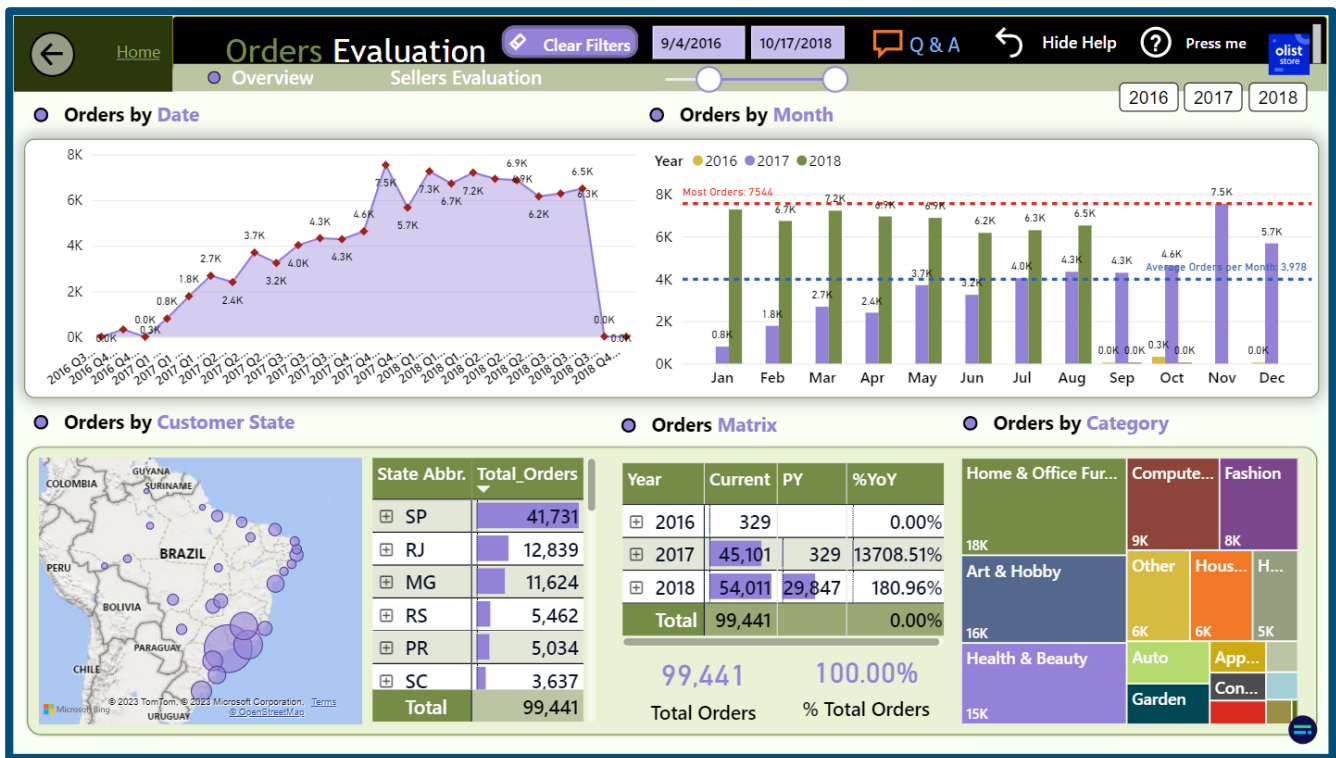
Επίσης έχει δημιουργηθεί καρτέλα που περιέχει επιπλέον **συγκεντρωτικά δημογραφικά στοιχεία για όλη την Βραζιλία**, όπως το βιοτικό επίπεδο, το κατά κεφαλήν εισόδημα, το επίπεδο μόρφωσης των πελατών ανά γεωγραφική περιοχή για την πιο ενδελεχή αξιολόγηση του προφίλ τους. Αυτές οι πληροφορίες είναι κρίσιμες για την **χάραξη στρατηγικής επενδύσεων** και επέκτασης του δικτύου της εταιρείας.



Εικόνα 70 : Δημογραφικά Στοιχεία Πελατών

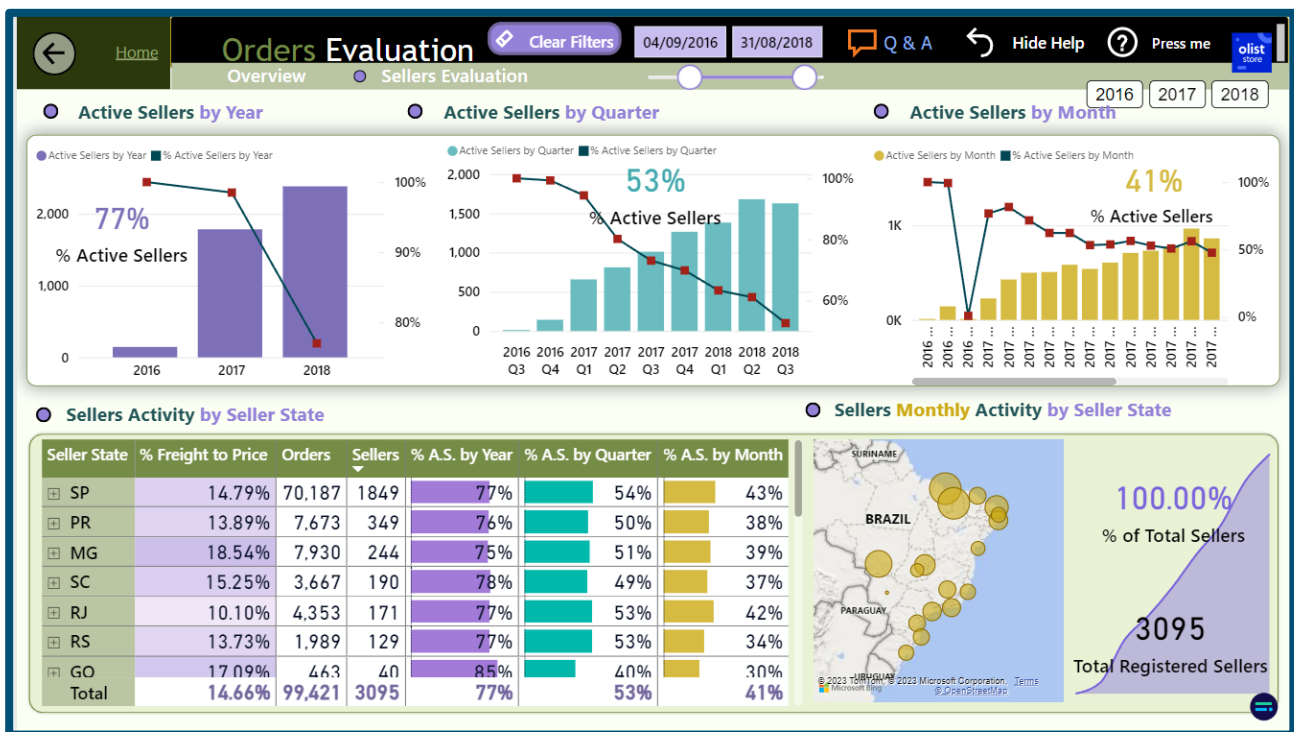
Ανάλυση Παραγγελιών – Orders Evaluation

Στην καρτέλα αξιολόγησης των παραγγελιών, βρίσκουμε στοιχεία για το πως κινείται ο αριθμός των παραγγελιών σε διάφορα επίπεδα χρόνου (ημέρας, εβδομάδας, μήνα, τετράμηνο, έτους στον χρόνο). Μπορούμε να αξιολογήσουμε ποιες κατηγορίες προϊόντων είναι πιο δημοφιλείς από άποψη αριθμό παραγγελιών και ποιος είναι ο μέσος αριθμός παραγγελιών ανά μήνα. Επίσης, υπάρχει πίνακας που βοηθάει τον χρήστη της αναφοράς να διακρίνει ακραίες τιμές και αλλαγές στον αριθμό των παραγγελιών σε σχέση με κάποια προηγούμενη χρονική περίοδο. Ακόμη, να μάθει ποιες πολιτείες λαμβάνουν τις περισσότερες παραγγελίες και ποιες τις λιγότερες. Όπως είναι στημένο το μοντέλο, δίνεται η δυνατότητα μέσω αυτών των καρτελών στον χρήστη, να κάνει **ad-hoc ερωτήσεις** στην βάση για να απαντήσει ερωτήματα που τον ενδιαφέρουν την δεδομένη στιγμή αλλά και άλλα ερωτήματα που είναι απρόβλεπτα από τον σχεδιαστή της αναφοράς.



Εικόνα 71 : Καρτέλα Αξιολόγησης Παραγγελιών – Orders Evaluation

Ανάλυση Πωλητών – Sellers Evaluation

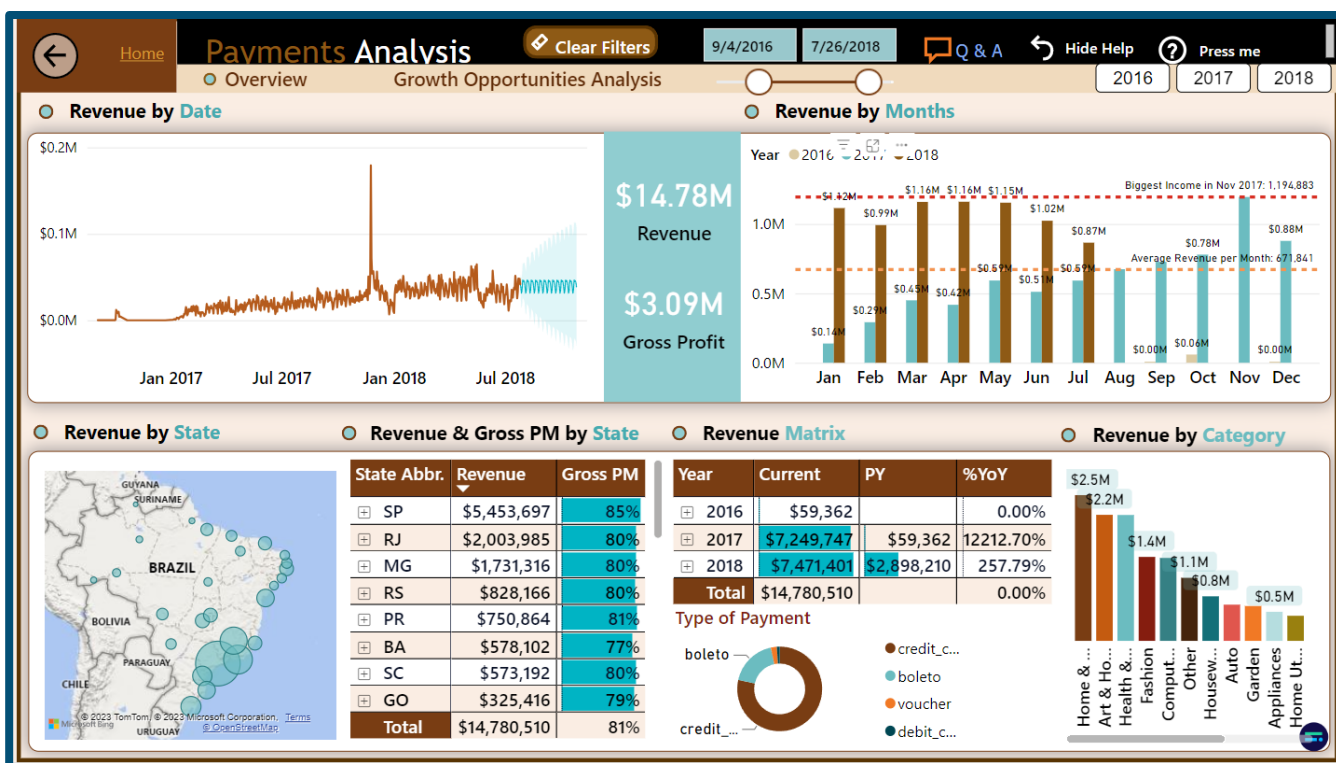


Εικόνα 72 : Ανάλυση Πωλητών στην πλατφόρμα της εταιρείας

Επίσης εφόσον το επιχειρηματικό μοντέλο της εταιρείας στηρίζεται στους πωλητές που εγγράφονται στην πλατφόρμα της και διαφημίζουν τα προϊόντα τους είναι κρίσιμο να γνωρίζει η εταιρεία τον **ρυθμό με τον οποίο προστίθενται καινούργιοι** αλλά και **πόσοι παραμένουν ενεργοί**. Ενεργοί θεωρούνται οι πωλητές οι οποίες έχουν πουλήσει πρόσφατα κάποιο προϊόν τους μέσω της πλατφόρμας της εταιρείας. Το ποσοστό των πωλητών από τους συνολικά εγγεγραμμένους στην πλατφόρμα που παραμένουν ενεργοί, αξιολογείται σε επίπεδο έτους, τετραμήνου και μήνα, συνολικά και ανά γεωγραφική περιοχή.

Ανάλυση Πληρωμών – Payments Analysis

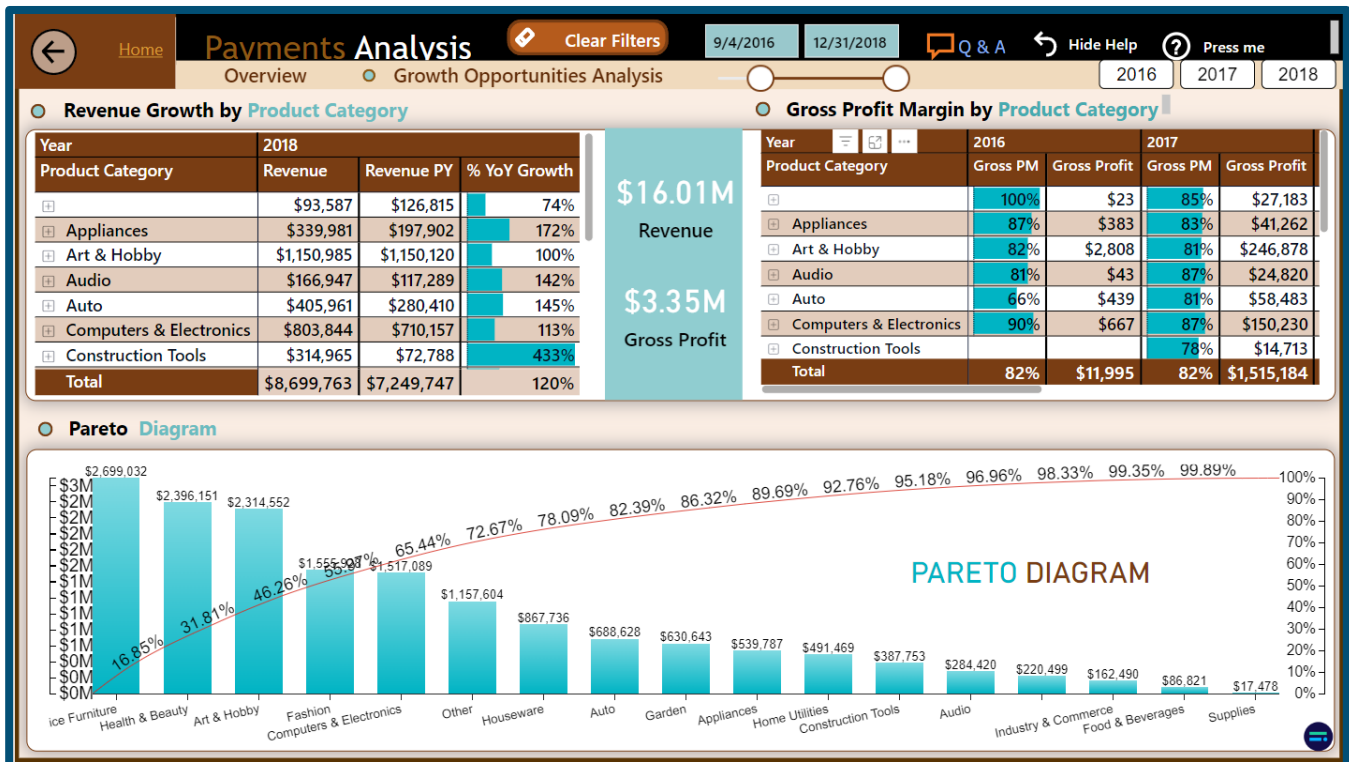
Η ανάλυση πληρωμών γίνεται σε δύο καρτέλες. Σε αυτές μπορούμε παρατηρήσουμε τα γενικά έσοδα της εταιρείας στον χρόνο, **να κάνουμε σύγκριση των εσόδων της εταιρείας ανά μήνα** για κάθε έτος. Να δούμε την διακύμανση των εσόδων ανά κατηγορία προϊόντων και ανά γεωγραφική περιοχή ώστε, να εντοπίσουμε ποιες είναι οι προσοδοφόρες σε επίπεδο εσόδων και μεικτού περιθωρίου κέρδους. Επίσης, υπάρχουν στοιχεία για τον τρόπο που επιλέγουν οι πελάτες να καταθέσουν μια πληρωμή (τραπεζική κάρτα, μετρητά, voucher).



Εικόνα 73 : Αξιολόγηση Πληρωμών – Payments Analysis

Στους πίνακες της δεύτερης καρτέλας μπορούμε να **ελέγξουμε** τον ρυθμό ανάπτυξης των

εσόδων αλλά και το περιθωρίου μεικτούς κέρδους ανά κατηγορία προϊόντων που εμπορεύεται η εταιρεία. Επίσης, εκεί βρίσκεται το χρήσιμο διάγραμμα **Pareto** για τα έσοδα της εταιρείας ανά κατηγορία προϊόντων. Η καρτέλα αυτή επιχειρεί να βοηθήσει την διοίκηση και την διεύθυνση Μάρκετινγκ να κεφαλαιοποιήσει τις ευκαιρίες της αγοράς χαράσσοντας σωστή επενδυτική κατεύθυνση, ελέγχοντας δύο βασικά **KPI**, του **ρυθμού ανάπτυξης** των κατηγοριών προϊόντων αλλά και τα **περιθώρια κερδών** από αυτές.



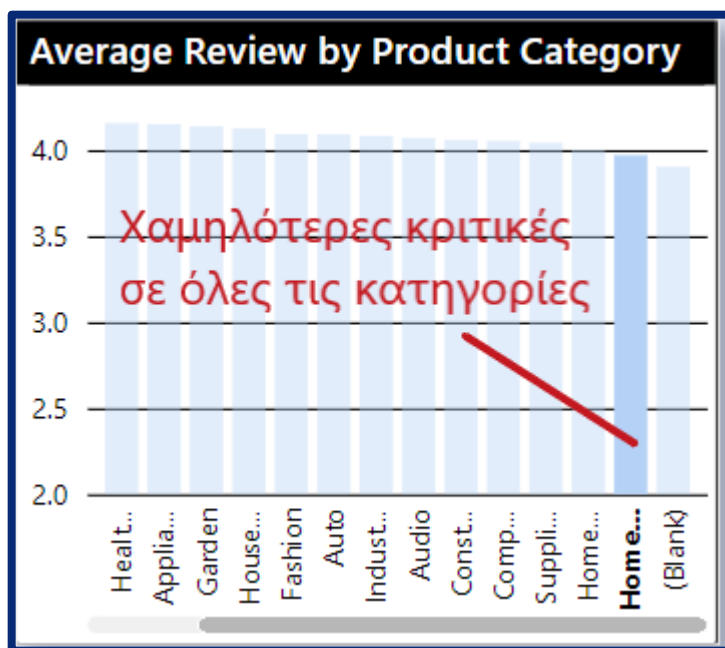
Εικόνα 74 : Ανάλυση Ευκαιριών Ανάπτυξης – Payments Analysis

4.3.2 Παραδείγματα επίδειξης των δυνατοτήτων του εργαλείου Power BI

Έχοντας μια πρώτη γνωριμία με τις καρτέλες, θα ήταν χρήσιμο να γίνει μια επίδειξη των απαντήσεων σε ερωτήματα που μπορεί να θέσει ένας χρήστης. Ένα σύστημα επιχειρηματικής ευφυΐας, χάρη στο μοντέλο δεδομένων στο οποίο στηρίζεται στο παρασκήνιο, καταφέρνει να «αντέξει» σε σχεδόν σε οποιοδήποτε μοτίβο απρόβλεπτων ερωτημάτων στην βάση δεδομένων του (**query building**). Ένας αναλυτής δεν μπορεί να γνωρίζει από πριν όλα τα ερωτήματα θέλει να θέσει στην βάση, αλλά ένα καλό μοντέλο θα πρέπει να είναι ευέλικτο και να δίνει την δυνατότητα για πληθώρα ερωτημάτων προβλεπόμενων και μη.

Με μια αρχική ματιά στην καρτέλα μας δείχνει ότι η κατηγορία προϊόντων με τις χαμηλότερες κριτικές είναι η κατηγορία «**Home & Office Furniture**». Η Γενική διεύθυνση και η διεύθυνση του Μάρκετινγκ επιθυμούν να γνωρίζουν:

- I. Ποιος είναι ο σημαντικότερος παράγοντας που διαμορφώνει τις χαμηλές κριτικές για την συγκεκριμένη κατηγορία και επίσης
- II. Ποιες περιοχές συγκεντρώνουν τις χαμηλότερες βαθμολογίες σε αυτή την κατηγορία προϊόντων και να βρεθούν αίτια για αυτό.



Εικόνα 75 : Μέση βαθμολογία κριτικής ανά κατηγορία προϊόντων – Reviews Καρτέλα

I.

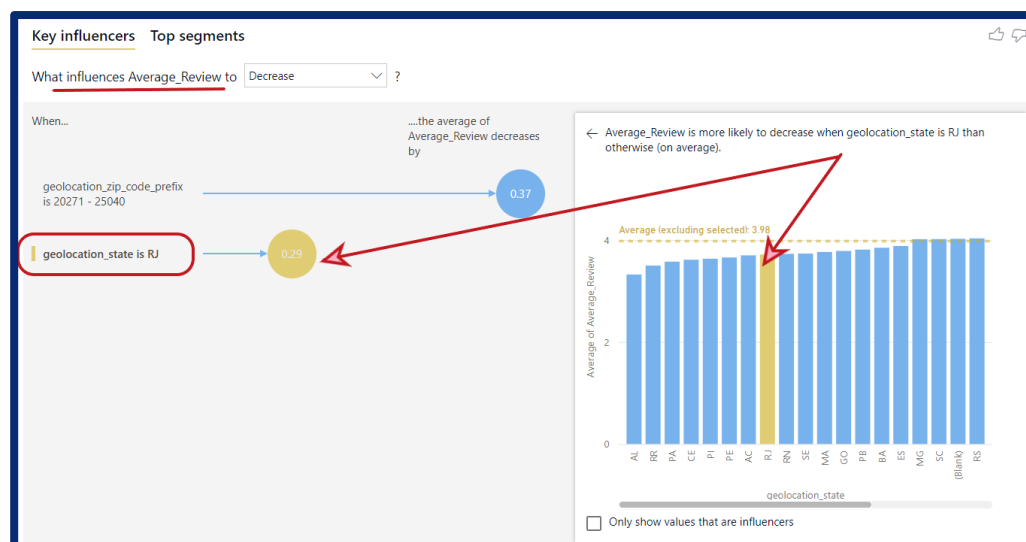
Έχοντας επιλέξει ως φίλτρο την κατηγορία προϊόντος «**Home & Office Furniture**», θα παρατηρούμε τις μέσες βαθμολογίες κριτικών. Οι περιοχές οι οποίες έχουν περισσότερες παραγγελίες θα επηρεάζουν ασφαλώς περισσότερο την μέση βαθμολογία των κριτικών. Οι πολιτείες οι οποίες έχουν τις περισσότερες παραγγελίες είναι, όπως διακρίνουμε από τον πίνακα είναι οι **SP**, **RJ**, και **MG**.

Ερωτήματα σαν και αυτά είναι τα πιο συχνά στην καθημερινότητα των εταιριών, και εδώ φαίνεται η «μαγεία» των αναλυτικών συστημάτων, καθώς τα ερωτήματα αυτά αναδεικνύουν την δυνατότητα των συστημάτων αυτών να φιλτράρουν ταυτόχρονα πολλούς πίνακες του μοντέλου και άρα να πραγματοποιείται πολυδιάστατη ανάλυση. Αν έχει χτιστεί ένα καλό μοντέλο δεδομένων με τις κατάλληλες σχέσεις, μπορούν να απαντηθούν σχεδόν όλες οι πιθανές ερωτήσεις.

Average Review by Geolocation by Order Status							
State	Total_Orders	Customers	Delivered%	Review_Delivered	Not_Delivered%	Review_Not_Delivered	Average_Review
SP	7,966	7,741	98.20%	4.12	1.80%	1.82	4.08
RJ	2,610	2,519	97.13%	3.78	2.87%	1.94	3.73
MG	2,060	1,992	98.88%	4.00	1.12%	2.39	3.99
RS	1,072	1,041	98.97%	4.01	1.03%	2.36	4.00
PR	879	856	98.63%	4.04	1.37%	2.25	4.02
SC	610	588	98.20%	4.00	1.80%	1.55	3.96
Total	17,861	17,315	98.17%	4.01	1.83%	1.92	3.97

Εικόνα 76 : Πολιτείες με περισσότερες παραγγελίες στην κατηγορία «Home & Office Furniture» - Reviews Καρτέλα

Η πολιτεία η οποία μειώνει έντονα την μέση κριτική φαίνεται ότι είναι η **RJ**. Αυτή η διαπίστωση μπορεί γίνει και από τις σπικτοποιήσεις **Key Influencers**, όπου αναδεικνύουν τον μεγαλύτερο παράγοντα που επηρεάζει την μέση βαθμολογία κριτικής. Συγκεκριμένα η πολιτεία **RJ** ρίχνει την μέση κριτική στην επιλεγμένη κατηγορία κατά **0.29 μονάδες**.



Εικόνα 77 : Η πολιτεία RJ αναδεικνύεται ως κύριος παράγοντας στην μειωμένη κριτική – Reviews Καρτέλα 2

Επίσης το εργαλείο τεχνητής νοημοσύνης μας πληροφορεί ότι όταν η διεύθυνση του πελάτη έχει ΤΚ 20271 έως 25041, τότε η μέση βαθμολογία μειώνεται κατά 0,37 μονάδες. Οι Ταχυδρομικοί Κώδικες αυτοί όπως μπορούμε να διαπιστώσουμε ανήκουν όλοι στην πολιτεία **RJ**.



Εικόνα 78 : Φιλτράρισμα καρτέλας Αξιολογήσεων για TK 20271-25041 – Reviews Καρτέλα

Μάλιστα μπορούμε να μάθουμε και ποια πόλη σε αυτή την πολιτεία επηρεάζει περισσότερο. Πατώντας την επιλογή **drill down**, εφόσον έχουμε δημιουργήσει *Ιεραρχία Δεδομένων* στα γεωγραφικά δεδομένα του μοντέλου μας (Πολιτεία > Πόλη > Ταχυδρομικός Κώδικας), αναπτύσσεται ο πίνακας των πόλεων στην πολιτεία **RJ**.

	RJ	2,610	2,519	97.13%	3.78	2.87%	1.94	3.73
rio de janeiro	1,438	1,384	96.52%	3.75	3.48%	2.00	3.69	
niteroi	106	101	100.00%	3.90			3.90	
nova iguacu	76	74	97.37%	3.62	2.63%	1.00	3.58	
sao goncalo	68	66	97.06%	3.68	2.94%	3.00	3.67	
duque de caxias	57	56	94.74%	3.89	5.26%	2.00	3.82	
niterói	55	53	100.00%	3.80			3.80	
volta redonda	54	52	100.00%	4.06			4.06	
campos dos goytacazes	47	47	97.87%	4.00	2.13%	1.00	3.93	

Εικόνα 79 : Λειτουργία Drill-down σε επίπεδο πόλης – Reviews Καρτέλα

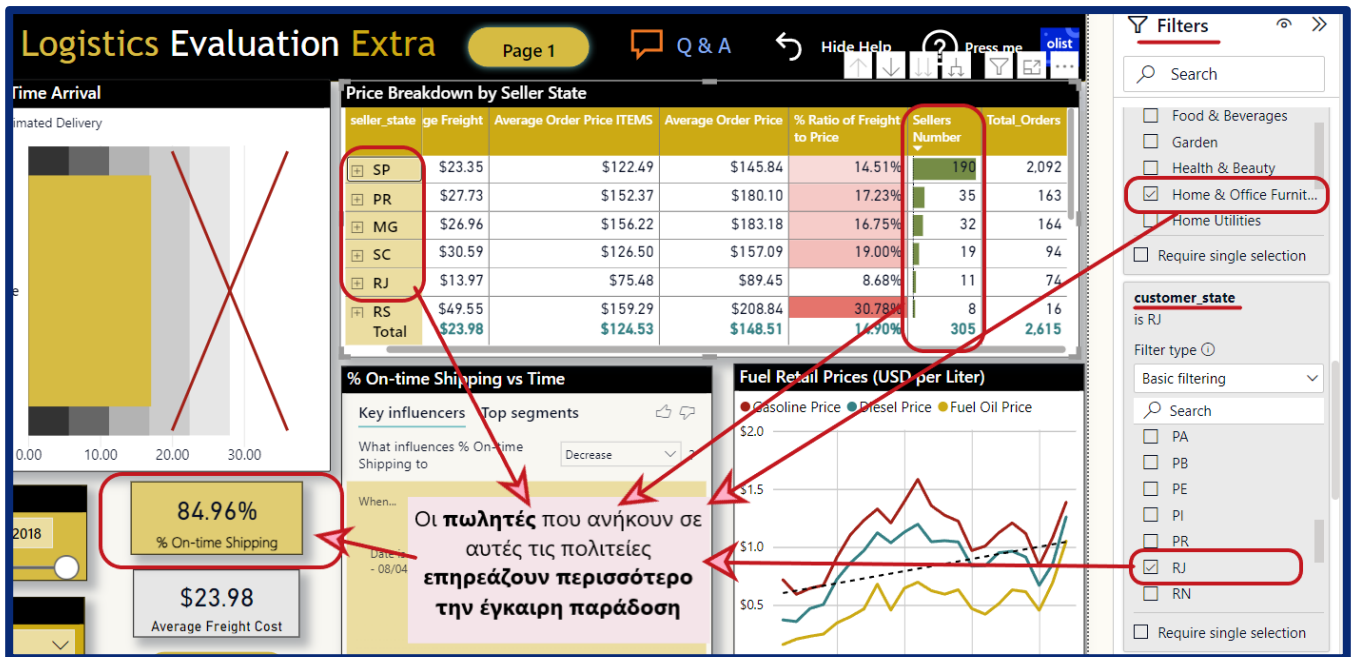
Μια από τις πιο συνήθεις αιτίες ,για κακές βαθμολογήσεις είναι ,εκτός από την κακή κατάσταση των προϊόντων είναι και η **καθυστερημένη παράδοση των προϊόντων**. Γι' αυτό θα αξιοποιήσουμε ξανά την τεχνητή νοημοσύνη ,για να διαπιστώσουμε αν το ποσοστό μη έγκαιρης παράδοσης είναι αυξημένο για την περίπτωση που εξετάζουμε ,δηλαδή η διεύθυνση πελάτη να είναι η πολιτεία **RJ** και να έχει παραγγελθεί προϊόν από την κατηγορία «**Home & Office Furniture**».



Εικόνα 80 : Η πολιτεία RJ αναδεικνύεται ως κύριος παράγοντας στο μειωμένο ποσοστό έγκαιρης παράδοσης – Reviews Καρτέλα 2

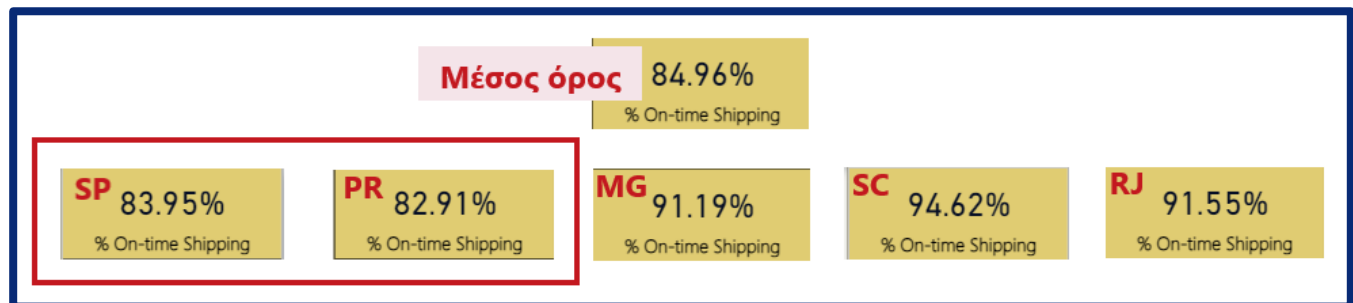
Πράγματι , από την οπτικοποίηση **Key Influencers**, διακρίνουμε εύκολα ότι όταν η πολιτεία παράδοσης είναι η **RJ** ,τότε το ποσοστό έγκαιρης παράδοσης επηρεάζεται περισσότερο από άλλες πολιτείες.

Αν θέλουμε να ελέγξουμε ποιοι είναι οι Πωλητές που σχετίζονται με αυτές τις παραγγελίες **φιλτράρουμε την καρτέλα 2 των Logistics** ,έχοντας επιλέξει παράδοση μόνο για **RJ** και αυτόματα βλέπουμε που βρίσκονται οι Πωλητές που προμηθεύουν την περιοχή αυτή. Έχοντας κάνει το φιλτράρισμα ,εντοπίζουμε τους πιο σημαντικούς σε αριθμό. Οι περισσότεροι βρίσκονται στις περιοχές **SP ,PR, MG, SC και RJ**. Το μέσο ποσοστό έγκαιρης παράδοσης είναι **84.96%**. Τώρα είμαστε σε θέση να ελέγξουμε που ανήκουν γεωγραφικά οι πωλητές με το χαμηλότερο ποσοστό.



Εικόνα 81 : Φιλτράρισμα Πωλητών που προμηθεύουν την περιοχή RJ – Logistics Καρτέλα 2

Το χαμηλότερο ποσοστό έγκαιρης παράδοσης έχουν οι Πωλητές στις πολιτείες του **SP** και **PR**. Τώρα γνωρίζοντας σε ποιες πολιτείες δημιουργούνται προβλήματα , είναι αρκετά πιο εύκολο να εντοπιστούν προβλήματα στην εφοδιαστική αλυσίδα σε αυτές τις περιοχές και όχι σε όλο το δίκτυο.

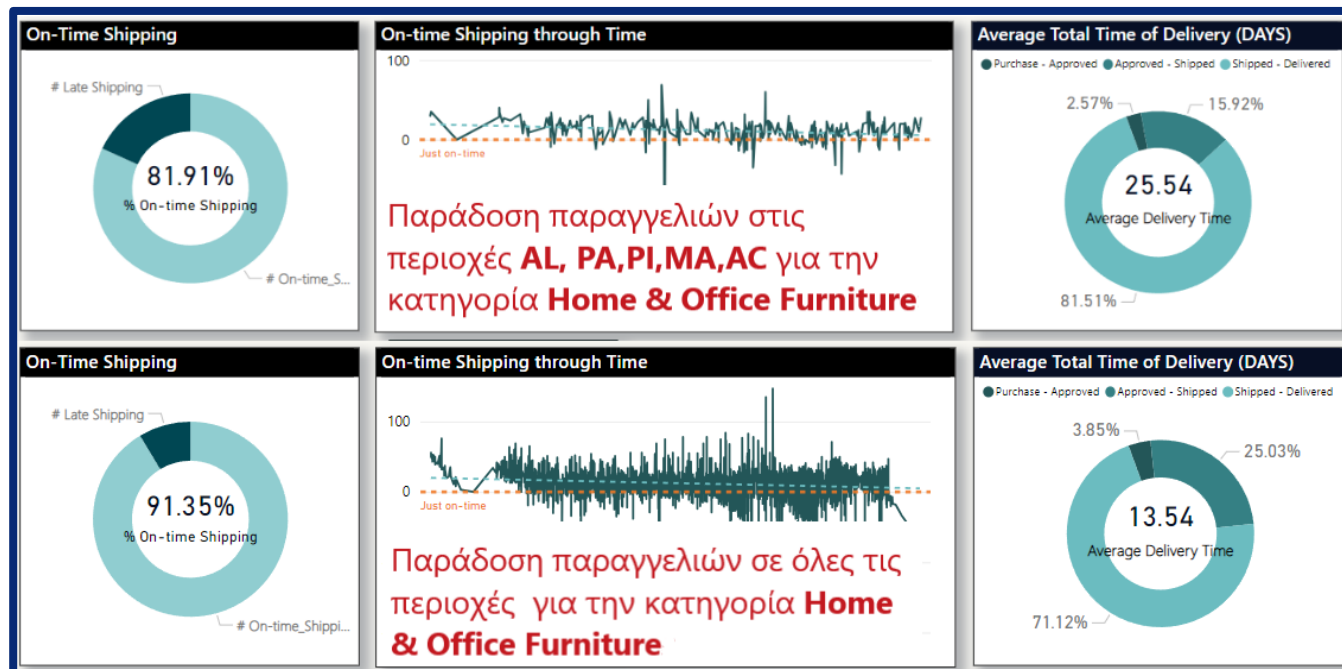


Εικόνα 82 : Τα ποσοστά έγκαιρης παράδοσης στο RJ ανά πολιτεία Πωλητή

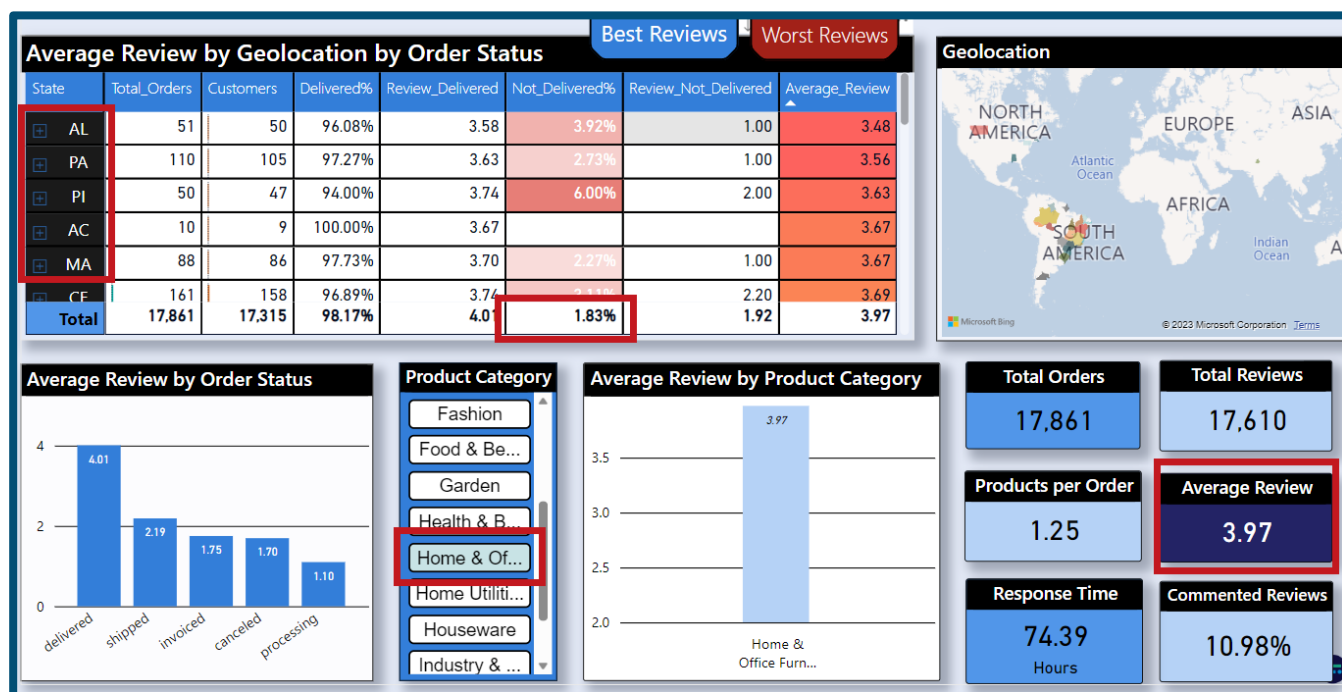
II.

Η δεύτερη ερώτηση αφορά τις πολιτείες με την χαμηλότερη βαθμολογία για την ίδια κατηγορία προϊόντων .Εκ νέου από τον πίνακα της καρτέλας Αξιολογήσεων μπορούμε να τις ταξινομήσουμε ανά βαθμολογία κριτικής. Στην περίπτωση αυτή οι περιοχές που έχουν την χαμηλότερη βαθμολογία είναι οι : **AL** (Αλαγκόας), **PA** (Παρά) ,**PI** (Πιάουσι) ,**AC** (Άκρε) και **MA** (Μαρανάο).

Επίσης παρατηρούμε ότι το ποσοστό των παραγγελιών που δεν έφτασαν ποτέ στον πελάτη είναι αυξημένο σε σχέση με τον μέσο όρο που διαμορφώνεται στο **1.83%**. Αυτό μπορούμε να το διαπιστώσουμε και από επόμενη καρτέλα που αφορά την εφοδιαστική αλυσίδα, όπου οι



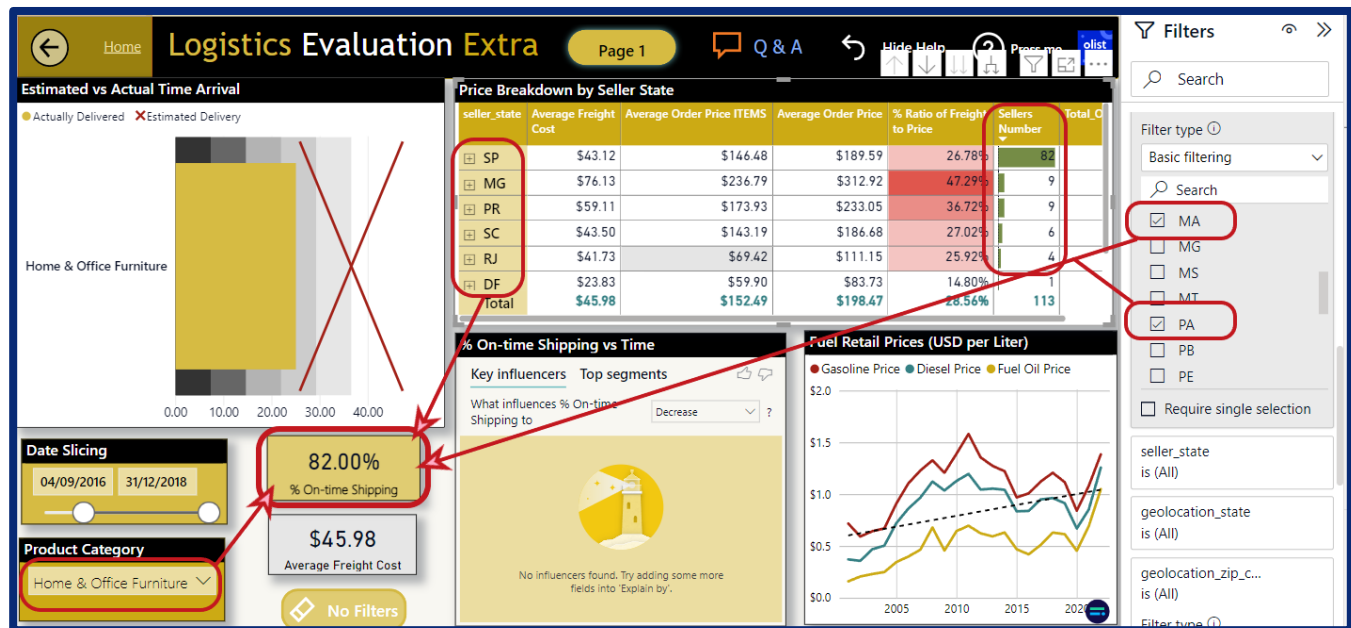
Εικόνα 84: Σύγκριση χρόνων παραγγελιών στις περιοχές AL,PA,PI,MA,AC – Logistics Καρτέλα



Εικόνα 83 :Χαμηλότερες Κριτικές στην Κατηγορία «Home & Office» - Reviews Καρτέλα

μέσοι χρόνοι άφιξης των παραγγελιών αλλά και το ποσοστό έγκαιρων παραδόσεων διαφέρουν αρκετά σε σχέση με τον μέσο όρο.

Όπως και στο προηγούμενο παράδειγμα I. μπορούμε να εντοπίσουμε ποιοι Πωλητές σχετίζονται με τις κακά ποσοστά έγκαιρης παράδοσης.



Εικόνα 85 :Φιλτράρισμα Πωλητών που προμηθεύουν τις περιοχές AL,PA,PI,MA,AC – Logistics Καρτέλα 2

Πράγματι οι περισσότεροι Πωλητές που προμηθεύουν τις πολιτείες **AL, PA, PI, MA, AC** βρίσκονται στις ίδιες πολιτείες με πριν ,δηλαδή τις **SP, PR, MG, SC και RJ**. Συγκεκριμένα η περιοχή **SP** που επηρεάζει περισσότερο το ποσοστό έγκαιρων παραδόσεων ,έχει ποσοστό **82.11%** και οι περιοχές **PR, MG** ακολουθούν με ποσοστό **78.95%** και **75%** αντίστοιχα. Χάριν της λειτουργίας **Drill Down** μπορούμε να κάνουμε ακόμα πιο συγκεκριμένο τον εντοπισμό των Πωλητών , σε επίπεδο ταχυδρομικού κώδικα.

seller_state	Drill Down	Average Freight Cost	Average Order Price ITEMS	Average Order Price	% Ratio of Freight to Price	Sellers Number	Total Orders
<input checked="" type="checkbox"/>	SP	\$43.12	\$146.48	\$189.59	26.78%	82	255
<input type="checkbox"/>	sao paulo	\$37.24	\$171.12	\$208.37	23.13%	24	78
<input type="checkbox"/>	ibitinga	\$38.33	\$113.46	\$151.79	23.81%	20	81
<input checked="" type="checkbox"/>	santo andre	\$34.96	\$104.46	\$139.42	21.72%	4	11
	9171	\$38.12	\$93.50	\$131.62	23.68%	1	2
	9190	\$19.04	\$49.90	\$68.94	11.83%	1	1
	9220	\$24.62	\$76.13	\$100.75	15.29%	1	4
	9270	\$47.71	\$151.92	\$199.62	29.63%	1	4

Εικόνα 86 : Λειτουργία Drill Down σε επίπεδο TK – Logistics Καρτέλα 2

Συγκεντρωτικά μπορούμε να συμπεράνουμε ότι:

- Ο κύριος παράγοντας για τις χαμηλές κριτικές στην κατηγορία «**Home and Office Furniture**» είναι οι παραδόσεις των παραγγελιών στην περιοχή **RJ** (Ρίο ντε ζανέιρο) καθώς ρίχνουν την μέση κριτική κατά **0.29 μονάδες**. Ενώ πιθανότατα οι Πωλητές που σχετίζονται με το χαμηλό ποσοστό βρίσκονται κυρίως στις περιοχές του **SP** (Σάο Πάολο) και **PR** (Παρανά).
- Οι πολιτείες που συγκεντρώνουν τις χαμηλότερες βαθμολογίες είναι οι **AL,PA,PI,MA** και **AC**. Σε αυτό οφείλονται κατά πάσα πιθανότητα η μειωμένη απόδοση της έγκαιρης παράδοσης των παραγγελιών αλλά και τους μεγαλύτερους σε διάρκεια χρόνους παράδοσης. Οι Πωλητές του δικτύου της εταιρείας που επηρεάζουν το ποσοστό βρίσκονται στις περιοχές **SP ,PR** και **MG**.

4.3.3 Ανάλυση Προφίλ Πελατών – Customers Analysis

Η αξιολόγηση του προφίλ των πελατών είναι ένα από τα βασικά αντικείμενα αναζήτησης και επένδυσης σε κάθε εταιρεία, η οποία επιθυμεί να αυξήσει τα έσοδα της επενδύοντας τους πόρους της με σωστό τρόπο. Σε όρους Μάρκετινγκ συνήθως αποκαλείται χαρτογράφηση των πελατών και βοηθάει την Τμηματοποίηση της Αγοράς (**Market Segmentation**), ώστε να γίνεται πιο εύκολο το **Target Positioning**.

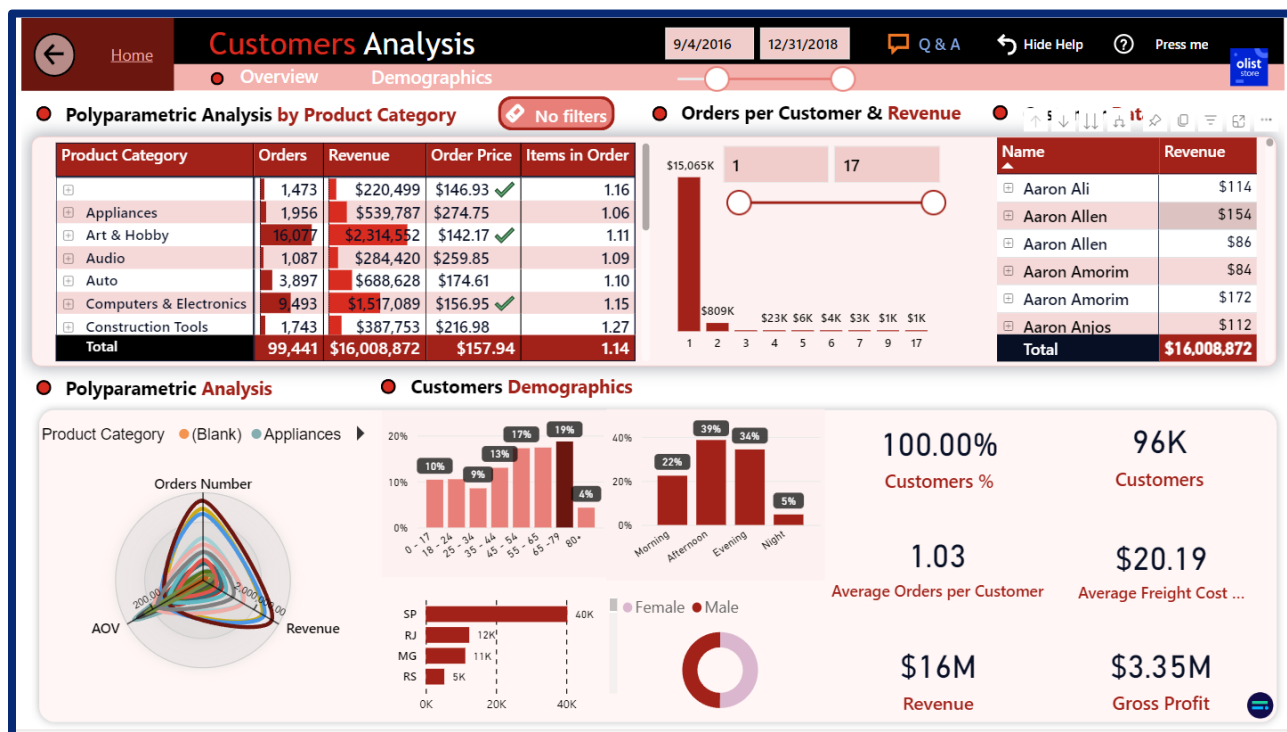
Υπάρχουν 4 κατηγορίες που γίνεται συνήθως η τμηματοποίηση της αγοράς :

- I. **Δημογραφικά στοιχεία**
- II. **Γεωγραφικά στοιχεία**
- III. **Ψυχογραφικά στοιχεία**
- IV. **Συμπεριφορικά στοιχεία**

Από την καρτέλα της αναφοράς «**Customers Analysis**», μπορούν να εξαχθούν σημαντικά συμπεράσματα για την συμπεριφορά και το προφίλ των πελατών της εταιρείας.

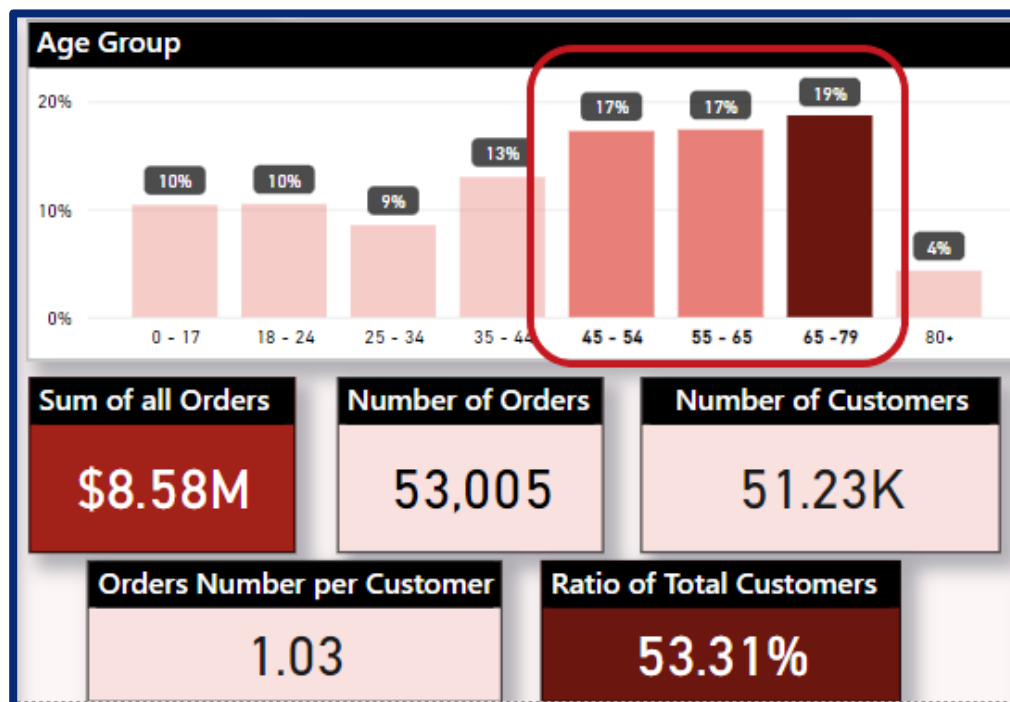


Σχήμα 16 : Στοιχεία ανάλυσης προφίλ πελατών



Εικόνα 87 : Καρτέλα Αξιολόγησης των Πελατών

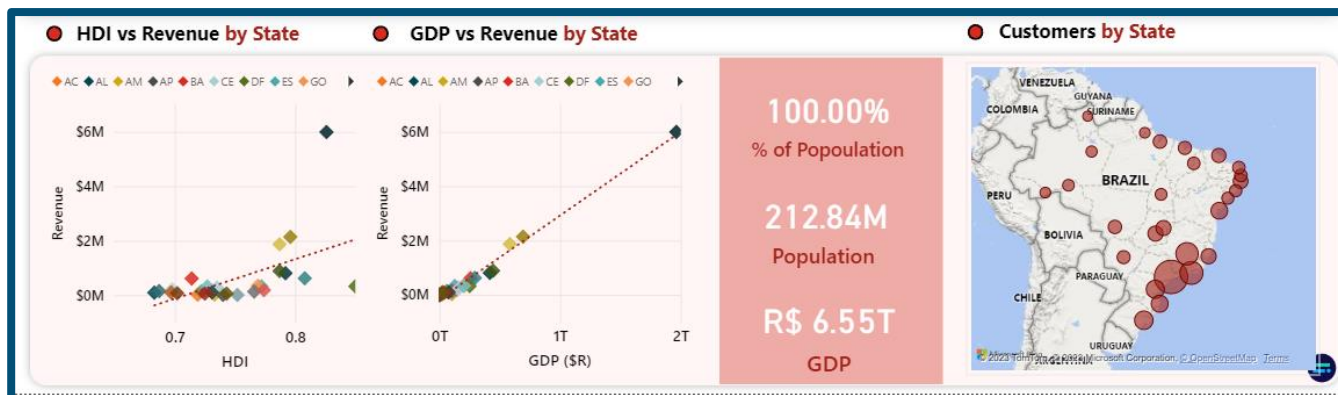
Τα πρώτα συμπεράσματα που βγάζουμε από την εποπτική εικόνα των καρτελών **Customers** και **Orders Analysis**.



Εικόνα 88 : Τα ηλικιακά γκρουπ των πελατών της εταιρείας

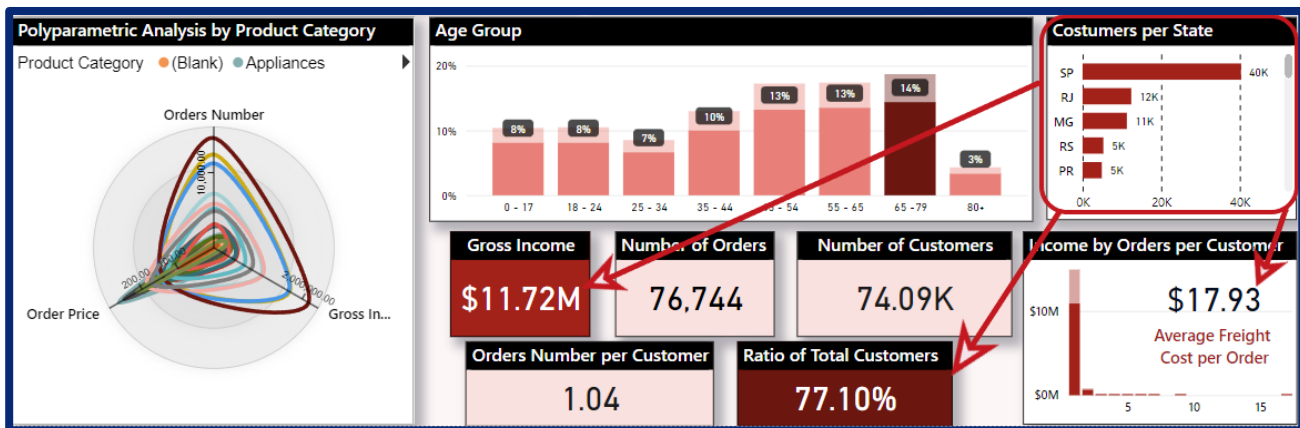
- Οι μοναδικοί πελάτες της εταιρείας αυτή την στιγμή πελάτες είναι περίπου **96χιλ.**
- Τα ηλικιακά γκρουπ που χρησιμοποιούν περισσότερο την εταιρεία για τις αγορές τους είναι **45-79 έτη** με ποσοστό **53%** των συνολικών πελατών και πραγματοποιώντας τζίρο ανάλογο του ποσοστού αυτού (**8.53εκ. \$** από τα **16.1εκ \$**).
- Το **σεξουαλικό φύλο** των πελατών είναι **μοιρασμένα σχεδόν εξίσου**.
- Συνολικά οι **5 πρώτες πολιτείες αποτελούν το 77,1% όλου του πελατολογίου** με τζίρο **11,72εκ. \$** , το οποίο αντιστοιχεί σε **73% του συνολικού τζίρου** της εταιρείας. Η περιοχή **SP** ,δηλαδή του Σάο Πάολο, διαθέτει τους περισσότερους πελάτες με διαφορές από τις υπόλοιπες περιοχές **40,3χιλ** πελατών. Η κατάταξη των πρώτων 5 περιοχών είναι η ακόλουθη:
 - **SP** με **41,94%**
 - **RJ** με **12,89%**,
 - **MG** με **11,72%**,
 - **RS** με **5,49%** και
 - η **PR** με **5,08%**.

Οι πολιτείες αυτές αποτελούν μεγάλα αστικά κέντρα , με πυκνή κατοίκηση. Οι κάτοικοι των περιοχών αυτών έχουν μεγαλύτερο κατά κεφαλήν εισόδημα και ο δείκτης Ανθρώπινης Ανάπτυξης (Human Development Index - **HDI**) είναι υψηλότερος συγκριτικά με τις επαρχιακές περιοχές.



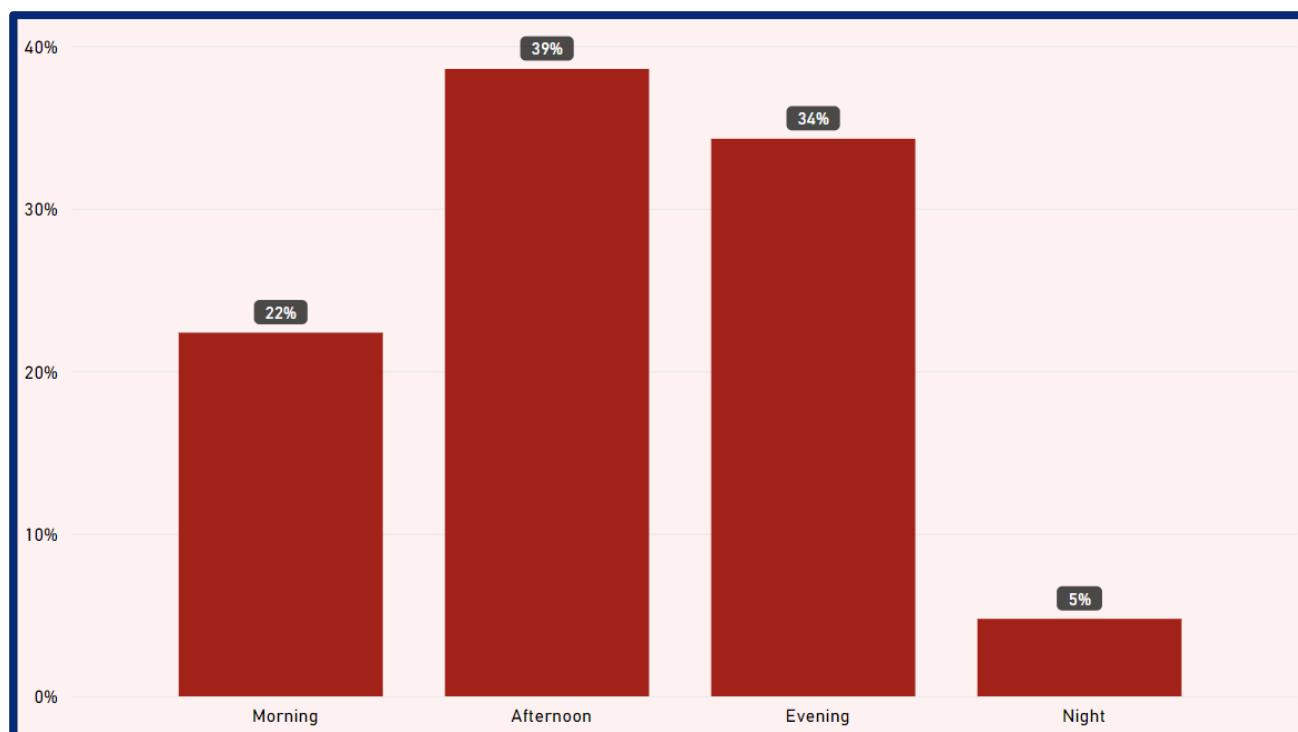
Εικόνα 89 : Δείκτης HDI και ΑΕΠ στις πολιτείες της Βραζιλίας

Επίσης το γενικό μέσο κόστος μεταφορικών μέσων από τα **20,19\$** , σε αυτές τις πολιτείες διαμορφώνεται στα **17,93\$** , μια μείωση **12%**.



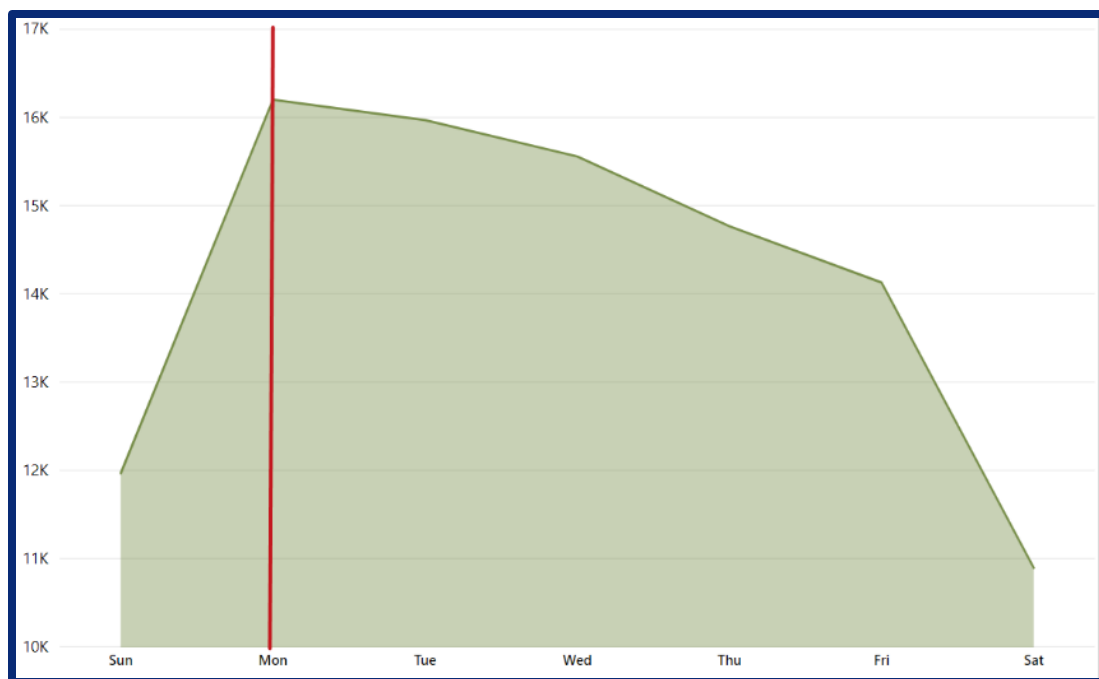
Εικόνα 90 : Πολυπαραμετρική ανάλυση πελατών

- Ο μέσος όρος προϊόντων ανά παραγγελία είναι πολύ κοντά στην μονάδα ,συγκεκριμένα στα **1,03 προϊόντα ανά παραγγελία**.
- Οι περισσότερες παραγγελίες που δέχεται οι εταιρεία είναι από διαφορετικούς πελάτες. Συγκεκριμένα :
 - **1 παραγγελία** έχει πραγματοποιήσει το **96,88%**
 - **2 παραγγελίες** έχουν πραγματοποιήσει το **2,86%**
 - **3 παραγγελίες** έχουν πραγματοποιήσει το **0,21%**
 - Ενώ, περισσότερες από **3 και μέχρι 17 παραγγελίες** ,το **0,26%**
- Επίσης, φαίνεται ότι πελάτες της εταιρείας πραγματοποιούν τις αγορές τους, κυρίως τις μεσημεριανές και απογευματινές ώρες .Συγκεκριμένα :
 - Οι **πρωινές αγορές** αποτελούν το **22%** (06:00 -12:00)
 - Οι **απογευματινές αγορές** αποτελούν το **39%** (12:00 -18:00)
 - Οι **βραδινές αγορές** αποτελούν το **34%** (18:00 -00:00)
 - και οι **νυχτερινές αγορές** αποτελούν το **5%** (12:00 -18:00)



Εικόνα 91 : Το διάστημα της μέρας που προτιμούν οι πελάτες για τις αγορές τους

- Οι μέρες που προτιμούν να αγοράζουν προϊόντα οι πελάτες είναι κυρίως οι καθημερινές ημέρες της εβδομάδας , με πρώτη την **Δευτέρα**.
 - Την **Δευτέρα** το ποσοστό αγορών είναι **16,3%**
 - Την **Τρίτη** το ποσοστό αγορών είναι **16,1%**
 - Την **Τετάρτη** το ποσοστό αγορών είναι **15,6%**
 - Την **Πέμπτη** το ποσοστό αγορών είναι **14,8%**
 - Την **Παρασκευή** το ποσοστό αγορών είναι **14,2%**
 - Την **Σάββατο** το ποσοστό αγορών είναι **11,0%**
 - και την **Κυριακή** το ποσοστό αγορών είναι **12,0%**.



Εικόνα 92 : Οι μέρες της εβδομάδας που κάνουν παραγγελίες οι πελάτες

Μια από τις κορυφαίες αναλύσεις που ενδιαφέρουν την εταιρεία είναι , **να μάθει ποιες κατηγορίες προϊόντων είναι πιο επικερδής**. Τα αποτελέσματα είναι ενδιαφέροντα και ιδιαίτερως χρήσιμα. Ο παρακάτω πίνακας δείχνει τις πιο επικερδής κατηγορίες

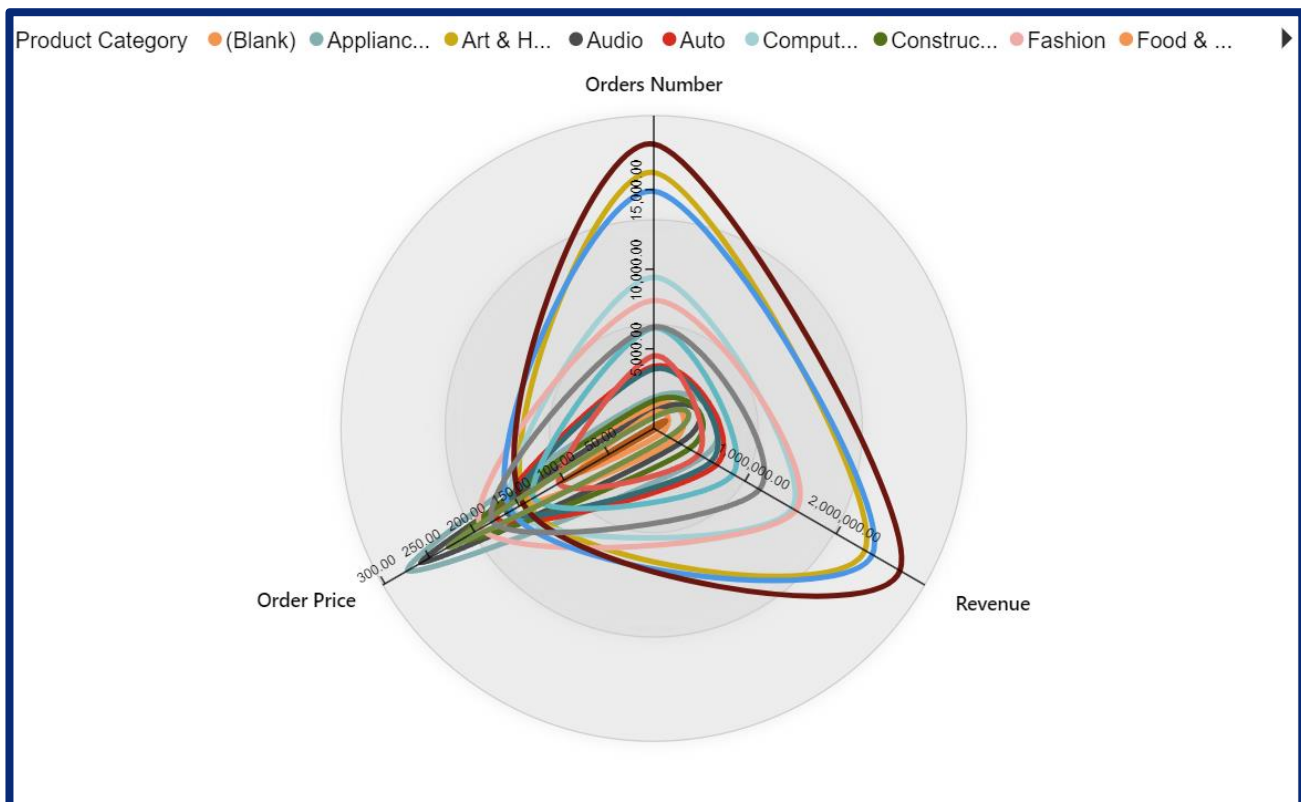
Product Category	Orders Number	Gross Income	Order Price	Items per Order
Appliances	1,956	\$539,787.44	\$274.75	1.06
Art & Hobby	16,077	\$2,314,551.75	\$142.17	1.11
Audio	1,087	\$284,419.59	\$259.85	1.09
Auto	3,897	\$688,628.01	\$174.61	1.10
Computers & Electronics	9,493	\$1,517,088.78	\$156.95	1.15
Construction Tools	1,743	\$387,752.94	\$216.98	1.27
Fashion	8,031	\$1,555,927.86	\$192.49	1.08
Food & Beverages	974	\$86,820.81	\$85.76	1.23
Garden	3,739	\$630,642.97	\$163.61	1.27
Health & Beauty	14,897	\$2,396,151.30	\$159.24	1.10
Home & Office Furniture	17,861	\$2,699,032.48	\$145.84	1.25
Home Utilities	4,557	\$491,468.93	\$105.69	1.11
Houseware	6,304	\$867,736.49	\$133.77	1.21
Industry & Commerce	697	\$162,490.18	\$229.24	1.17
Other	6,376	\$1,157,603.70	\$179.82	1.09
Supplies	167	\$17,477.82	\$101.42	1.19
Total	97,256	\$15,638,698.58	\$158.12	1.14

Εικόνα 93 : Πίνακας με τις κατηγορίες προϊόντων που προτιμούν οι πελάτες

προϊόντων ,μαζί με τον όγκο πωλήσεων τους , την τιμή AOV και τον μέσο αριθμό προϊόντων ανά παραγγελία. Συγκεκριμένα οι πιο επικερδής κατηγορίες είναι οι ακόλουθες.

- Έπιπλα Σπιτιού και Γραφείου με 2,7εκ. \$,
- Προϊόντα Υγείας και Ομορφιάς με 2,4εκ. \$ και
- Προϊόντα Τέχνης και Χόμπι με 2,3εκ. \$

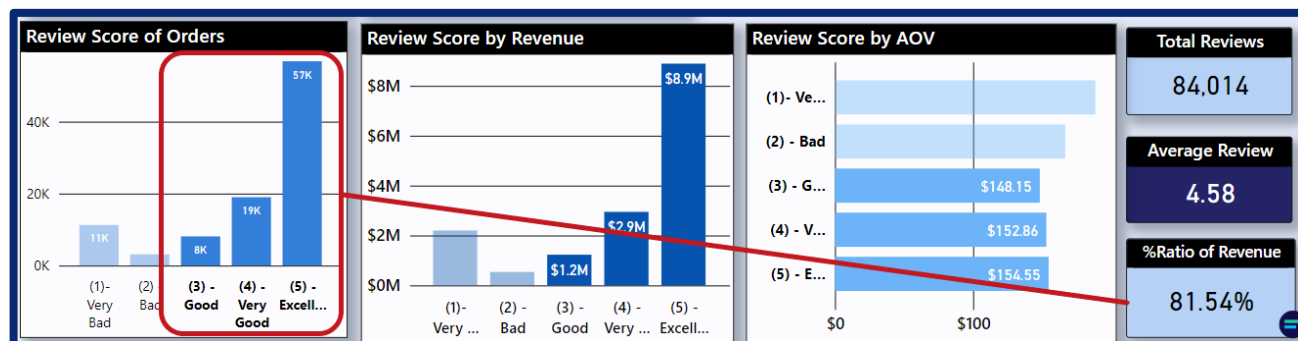
Μια διαφορετική οπτικοποίηση, με πολυπαραμετρική δυνατότητα ανάλυσης, των παραπάνω δεδομένων ,μας αποκαλύπτει κάτι πολύ ενδιαφέρον για τον τις συνήθειες των πελατών της εταιρείας. Συγκεκριμένα ,αποκαλύπτεται η συσχέτιση του αριθμού παραγγελιών και των εσόδων ,που ήταν αναμενόμενη ,αλλά και **η συσχέτιση των εσόδων με την μέση τιμή της παραγγελίας**. Υπάρχει ένα εύρος τιμών της μέσης τιμής της παραγγελίας η οποία αποδεικνύεται ότι προσφέρει τα περισσότερα έσοδα για την εταιρεία. Συγκεκριμένα οι κατηγορίες των προϊόντων που κατέκτησαν την κορυφή ,έχουν **ένα εύρος μέσης τιμής μεταξύ 140.0\$ και 160.0\$ ανά παραγγελία**. Οι κατηγορίες με χαμηλότερη μέση τιμή (AOV) τείνουν να έχουν μέση απόδοση εσόδων και οι κατηγορίες με μεγάλη μέση τιμή παραγγελίας (AOV) έχουν τα μικρότερα έσοδα σε σχέση με τις άλλες κατηγορίες. Όσον αφορά τον **μέσο αριθμό αντικειμένων που περιέχει η παραγγελία δεν φαίνεται να υπάρχει κάποια ισχυρή συσχέτιση με τα υπόλοιπα μεγέθη**, καθώς οι περισσότερες παραγγελίες περιέχουν 1 προϊόν συνήθως.



Εικόνα 94: Πολυπαραμετρική ανάλυση Πελατών – Διάγραμμα τύπου Radar

Η ανάλυση μέσα από την σκοπιά των **αξιολογήσεων** που αφήνουν οι πελάτες ,δείχνει παρόμοια αποτελέσματα. Συγκεκριμένα έχει γίνει παρόμοια ανάλυση για τον συσχετισμό μέσης τιμής παραγγελίας (AOV - Average Order Value) και του τζίρου της εταιρείας.

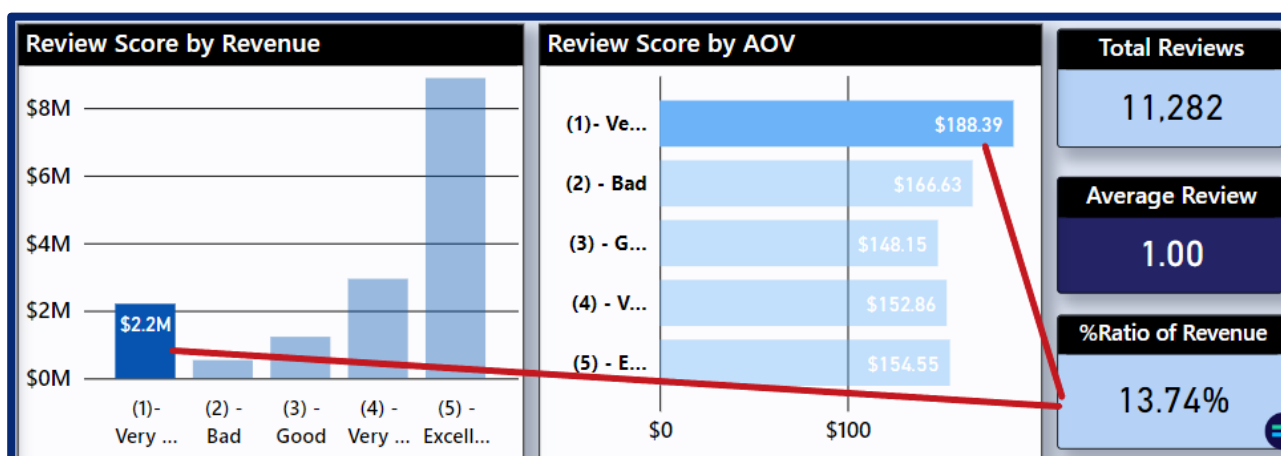
Οι παραγγελίες που έχουν λάβει **ικανοποιητική βαθμολογία (3/5 έως 5/5)** αποτελούν **ποσοστό μεγαλύτερο του 80%** των συνολικών. Αυτό δείχνει ότι η εταιρεία έχει καλή απόδοση στις πωλήσεις της.



Εικόνα 95 : Αξιολογήσεις για το 80% των παραγγελιών

Παράλληλα, οι αξιολογήσεις με χαμηλότερο σκορ αξιολόγησης φαίνεται ότι έχουν και **μεγαλύτερη μέση τιμή παραγγελίας (AOV)** και φέρνουν το **13.74%** του συνολικού τζίρου της εταιρείας. Αυτό υποδηλώνει ότι **τα προϊόντα με πολύ χαμηλή βαθμολογία είναι τα πιο ακριβά είδη**, όπου οι πελάτες δεν έχουν καλές εμπειρίες παρά το γεγονός ότι η εταιρεία φτιάχνει καλύτερο AOV.

Παρόλα αυτά, συγκριτικά οι παραγγελίες πολύ χαμηλής βαθμολογίας (1/5) έχουν αυξημένο όγκο πωλήσεων και εσόδων σε σχέση με τις παραγγελίες που αξιολογήθηκαν ως κακές και καλές (2/5 και 3/5) .

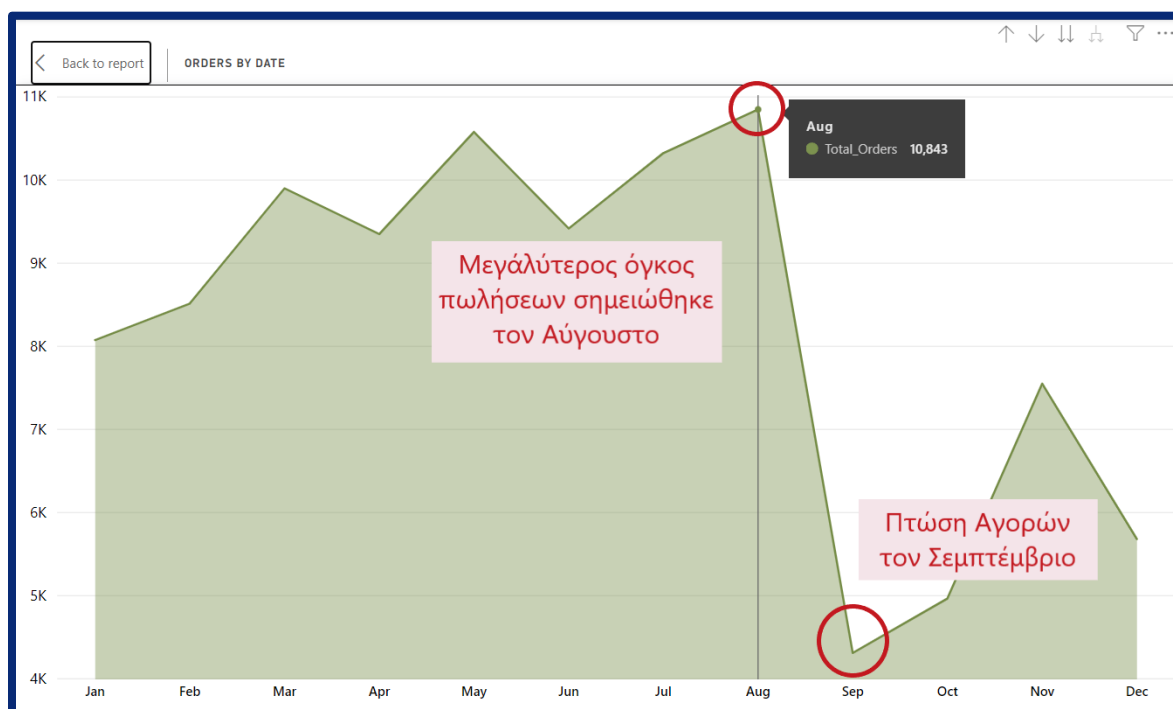


Εικόνα 96 : Οι χειρότερες αξιολογήσεις των παραγγελιών

Μια εναλλακτική εξήγηση για το γεγονός ότι τα προϊόντα με μεγαλύτερο όγκο πωλήσεων και βαθμολογίας έχουν χαμηλό AOV ,θα μπορούσε να είναι ότι τα προϊόντα αυτά βρίσκονται σε μια πιο ανταγωνιστική αγορά, γεγονός που θα οδηγεί σε χαμηλότερες τιμές άρα. Ταυτόχρονα αυτό σημαίνει χαμηλότερα περιθώρια κέρδους, παρά τον υψηλότερο όγκο πωλήσεών τους.

Από την **ανάλυση Παραγγελιών και Πληρωμών** μπορούμε να βγάλουμε και άλλα χρήσιμα συμπεράσματα για την συμπεριφορά των πελατών στις αγορές τους. Συγκεκριμένα, κάνοντας **Drill Down σε επίπεδο μήνα** στο διάγραμμα γραμμής ,μεταξύ παραγγελιών και χρόνου, μπορούμε να εξάγουμε συμπέρασμα για τους μήνες που επιλέγουν οι πελάτες να κάνουν τις αγορές.

- Ο μήνας του έτους με τον **μεγαλύτερο όγκο πωλήσεων** και για τα τρία χρόνια δεδομένων **είναι ο Αύγουστος**
- και ο **Σεπτέμβριος** μήνας με τον **μικρότερο όγκο πωλήσεων**.

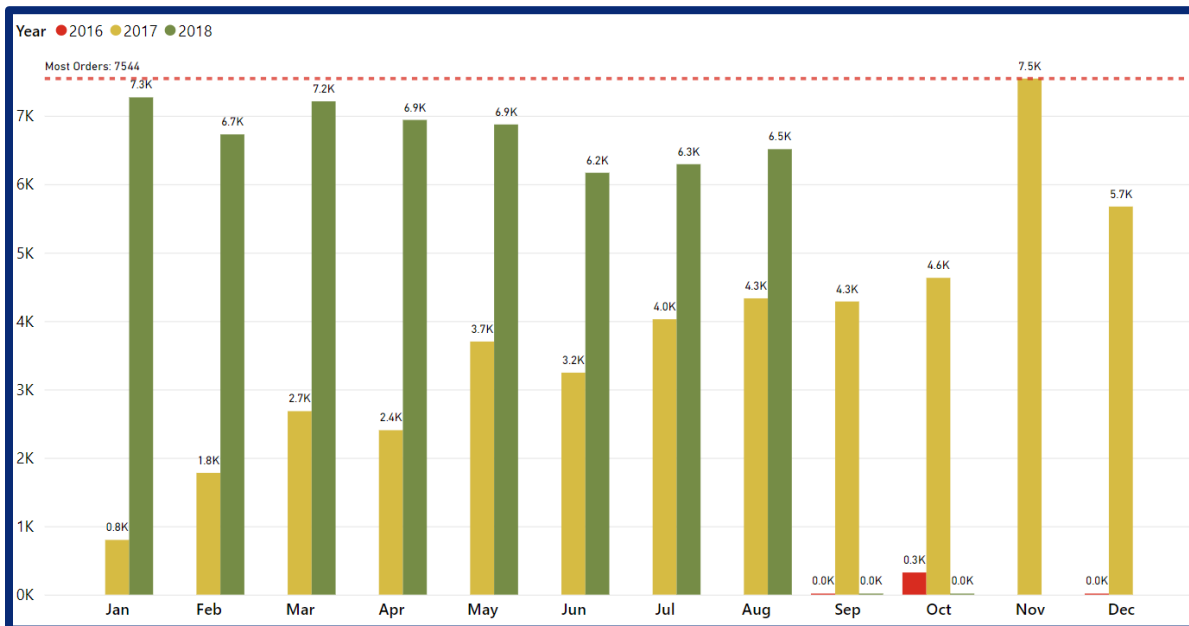


Εικόνα 97 : Η πορεία του αριθμού των παραγγελιών ανά μήνα

Αν δούμε όμως το διάγραμμα στηλών με την σύγκριση των μηνών ανά έτος η παρατήρηση, που γίνεται είναι διαφορετική. Οι μήνες που επιλέγουν οι πελάτες τις αγορές τους ,έχουν διαφορά αναλόγως το έτος, πιθανόν λόγω ελλείπων δεδομένων προς το κλείσιμο του έτους 2018. Αν θεωρήσουμε τα δεδομένα ακέραια , θα παρατηρούσαμε τεράστια μεταβολή των παραγγελιών μεταξύ των δύο ετών , ιδιαίτερα τους πρώτους μήνες του έτους.

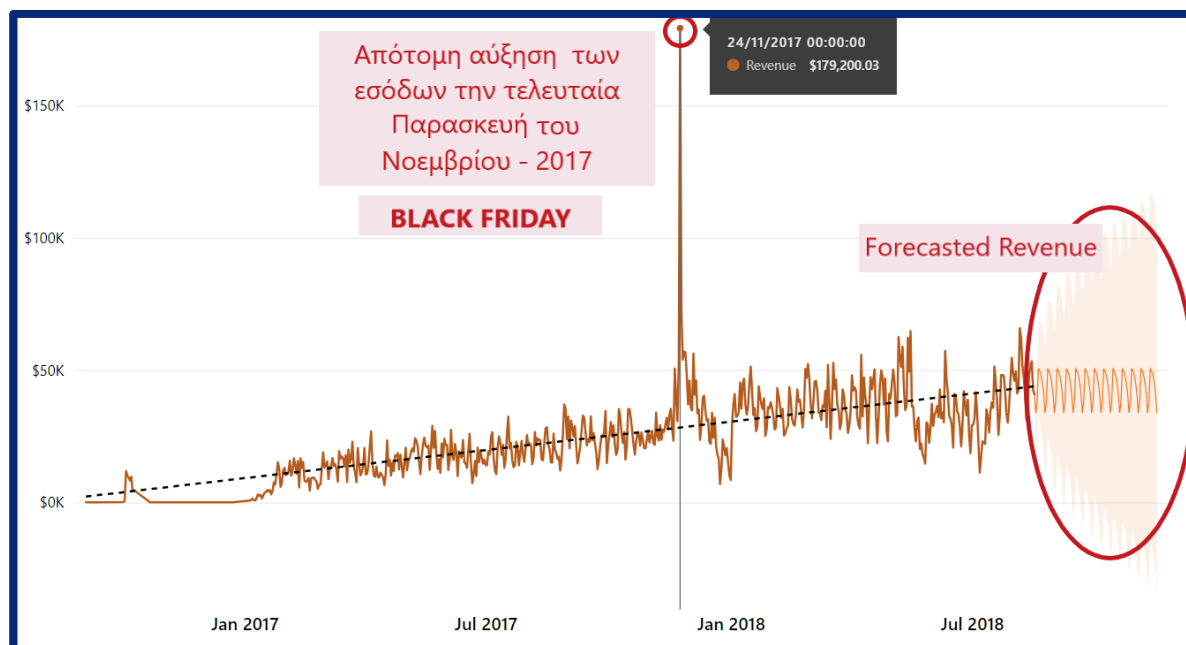
- Το **2017** οι περισσότερες αγορές έγιναν τον **Νοέμβριο**,

- ΕΝΩ ΤΟ 2018 ΕΓΙΝΑΝ ΤΟΝ Ιανουάριο και τον Μάρτιο.



Εικόνα 98 : Οι παραγγελίες της εταιρείας με ανά μήνα και σύγκριση ετών

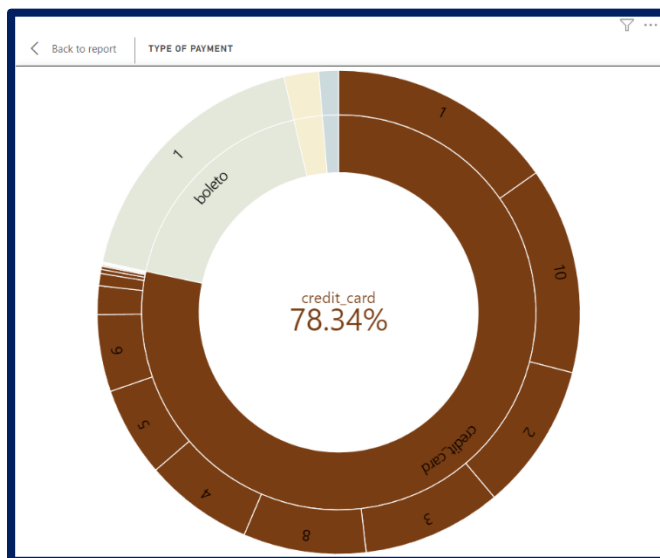
Ακόμη, αν δούμε το διάγραμμα γραμμής που δείχνει τα έσοδα της εταιρείας μέσα στο έτος, παρατηρούμε **μια μεγάλη αύξηση των εσόδων τον μήνα Νοέμβριο για το έτος 2017.** Αυτή η τιμή συμπίπτει με την τελευταία εβδομάδα του Νοεμβρίου, όπου γνωρίζουμε ότι συμπίπτει με την μαζική εμπορική συνήθεια των καταναλωτών, όπου εκμεταλλεύονται τις εκπτώσεις για να πραγματοποιήσουν τις αγορές τους. Αυτό ωφέλησε την εταιρεία αρκετά



Εικόνα 99 : Τα έσοδα της εταιρείας στην πορεία του χρόνου

,αφού οι πελάτες επέλεξαν την πλατφόρμα της και δημιούργησαν μεγάλο όγκο πωλήσεων και εσόδων.

Επίσης ένα ενδιαφέρον στοιχείο που σχετίζεται με την συμπεριφορά των πελατών είναι **ο τρόπος πληρωμής των παραγγελιών**. Ο κυρίαρχος τρόπος πληρωμής είναι η πιστωτική κάρτα. Αυτό είναι και λογικό καθώς , πλέον αυτός είναι και ο κυρίαρχος τρόπος πληρωμών παγκοσμίως ,λόγω της ασφάλειας που προσφέρει και ίσως εξαιτίας ορισμένων



Εικόνα 100 : Ο τρόπος πληρωμής που επιλέγουν οι πελάτες

ανταποδοτικών προγραμμάτων των τραπεζών που υπάρχουν. Στην Βραζιλία υπάρχει και ένας άλλος τρόπος πληρωμής που δεν είναι γνωστός στις υπόλοιπες χώρες που ονομάζεται **Boleto**. Τα Boleto είναι ένας άλλος δημοφιλής τύπος πληρωμής στη Βραζιλία, ειδικά για πελάτες που δεν έχουν πρόσβαση σε πιστωτικές κάρτες ή δεν αισθάνονται άνετα να τις χρησιμοποιούν για διαδικτυακές συναλλαγές. Τα Boleto είναι ουσιαστικά τραπεζικά κουπόνια που οι πελάτες μπορούν να πληρώσουν σε οποιαδήποτε τράπεζα, ταχυδρομείο ή παντοπωλείο.

Οι τρόποι πληρωμής σε ποσοστιαία απεικόνιση είναι :

1. **Πιστωτική κάρτα** με ποσοστό **78,34%**,
2. **Boleto** με ποσοστό **17,92%**,
3. **Voucher** με ποσοστό **2,37%**,
4. **Χρεωστική κάρτα** με ποσοστό **1,36%**.

4.3.4 Ανάλυση Πωλήσεων και Αξιολόγηση των Πωλητών-Συνεργατών

Τα έσοδα της εταιρείας είναι συνεχώς αυξανόμενα από το 2016 μέχρι και τον Σεπτέμβριο του 2018. Όλο και περισσότεροι πελάτες και Πωλητές χρησιμοποιούν την πλατφόρμα της Olist. Στα τρία χρόνια που υπάρχουν διαθέσιμα δεδομένα η εταιρεία έχει συγκεντρώσει περισσότερα από **16 εκατομμύρια \$** σε εγχώριες πωλήσεις.

Στο επιχειρησιακό μοντέλο της εταιρείας, εφαρμόζεται **μια πολιτική ποσοστιαίας τιμολόγησης των παραγγελιών** που δέχονται οι εγγεγραμμένοι Συνεργάτες – Πωλητές της πλατφόρμας. Από την τιμολόγηση του προϊόντος, το 21% της τιμής και ένα επιπλέον **σταθερό τέλος 5 \$R**, αποτελούν τα ακαθάριστα έσοδα της εταιρείας από την κάθε παραγγελία. Το υπόλοιπο ποσοστό της τιμής της παραγγελίας είναι τα έσοδα του Πωλητή-Συνεργάτη. Τα μεταφορικά κόστη για την παραγγελία, τα επωμίζεται ο πελάτης αν η παραγγελία συνολικά κοστίζει 79 R\$ και κάτω, ενώ αν κοστίζει περισσότερα από 79 R\$ η χρώση περνάει εξ' ημισείας στον εταιρεία και τον Συνεργάτη. Αυτό γίνεται στα πλαίσια μίας προωθητικής ενέργειας, ώστε να δίνεται κίνητρο στους καταναλωτές, για μεγαλύτερες αγορές.

Ορίζουμε την εξίσωση μικτών εσόδων της εταιρείας ανά παραγγελία :

$$\text{Ακαθάριστα Έσοδα} = 0.21 \times \text{Τιμή προϊόντος} + \text{Σταθερό τέλος (= \$ 5)} \\ - \text{μεταφορικά τέλη (> 79 R\$)}$$

Εξίσωση 1 : Ακαθάριστα έσοδα εταιρείας

Επιπλέον, ορίζουμε το **Μεικτό Περιθώριο Κέρδους**:

$$\text{Μ.Π.Κ.} = \frac{\text{Ακαθάριστα έσοδα} - \text{Τέλος μεταφορικών}}{\text{Ακαθάριστα έσοδα}} \times 100\%$$

Εξίσωση 2 : Μεικτό Περιθώριο Κέρδους

Πέρα από τον δείκτη του Μεικτό Περιθώριο Κέρδους εξαιρετική σημασία έχει για κάθε εταιρεία και ο Ρυθμός Ανάπτυξης Εσόδων ανά Έτος (**Revenue YoY% Growth**)

Ορίζουμε το Ρυθμός Ανάπτυξης Εσόδων (P.A.E. ανά Έτος :

$$P. A. E. \text{ ανά Έτος} = \frac{\text{Ακαθαρ. έσοδα τωρινού έτους}}{\text{Ακαθαρ. έσοδα περσινού έτους}} \times 100\%$$

Εξίσωση 3 : Ρυθμός Ανάπτυξης Εσόδων ανά έτος

Με βάση αυτούς τους δύο δείκτες και τα συνολικά έσοδα έγινε η ανάλυση ανά κατηγορία προϊόντων που εμπορεύεται η εταιρεία, ώστε να βρεθούν οι πιο επικερδής και οι πιο υποσχόμενες κατηγορίες προϊόντων. Ακολουθούν τα σημαντικότερα ευρήματα από την ανάλυση.

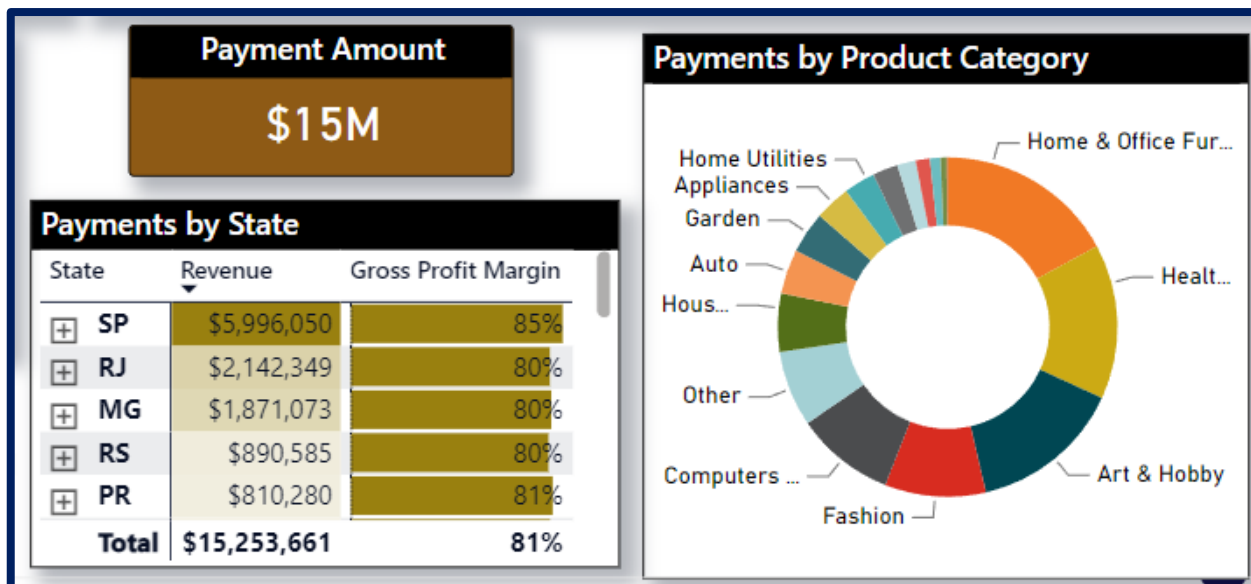
Συγκεντρωτικά από τις καρτέλες αξιολόγησης των πληρωμών που δημιουργήθηκαν, έχουμε τα επόμενα στοιχεία.

Για τα **ακαθάριστα έσοδα** συμπεραίνουμε ότι :

- Οι **εγχώριες πωλήσεις** παράγουν τζίρο **16,01εκ. R\$**.

Οι οικονομικά **πιο επικερδής περιοχές** είναι οι :

- **SP, RJ, MG, RS** και **PR**



Εικόνα 101 : Στοιχεία Εσόδων για τις εγχώριες πωλήσεις

Οι **περιοχές με το μεγαλύτερο μεικτό περιθώριο κέρδους** είναι :

- **SP** με **85%**
- **PR** με **81%**
- **DF** με **80%**

- **MG** με **80%**
- **RJ** με **80%**
- και γενικός μέσος όρος **81%**.

Οι πιο επικερδής κατηγορίες προϊόντων σε **απόλυτο αριθμό εσόδων** για το **2018** είναι :

- **Έπιπλα Σπιτιού και Γραφείου** με έσοδα **1,43εκ. \$**
 - **Υγεία και Ομορφιά** με έσοδα **1,15εκ. \$**
 - **Τέχνη και Χόμπι** με έσοδα **1,15εκ. \$**
-

Οι πιο **ταχέως αναπτυσσόμενες κατηγορίες προϊόντων** για τα έτη **2017-2018** είναι:

- **Κατασκευαστικά Εργαλεία** με ρυθμό ανάπτυξης **433%**
 - **Τρόφιμα και Ποτά** ρυθμό ανάπτυξης **217%**
 - **Προμήθειες** με ρυθμό ανάπτυξης **191%**.
-

Οι **κατηγορίες προϊόντων** με το μεγαλύτερο **μεικτό περιθώριο κέρδους** συνολικά είναι :

- **Οι Παροχές Σπιτιού (τηλεφωνία, ασφάλεια)** με ποσοστό **90%**
- **Τρόφιμα και Ποτά** με ποσοστό **87%**
- **Είδη Μόδας** με **87%**

Year		2017		2018		Total	
Product Category	Profit	Gross PM	Gross Profit	Gross PM	Gross Profit	Gross PM	Gross Profit
⊕ Home Utilities	\$339	90%	\$48,382	90%	\$59,549	90%	\$108,270
⊕ Food & Beverages	\$22	85%	\$5,997	89%	\$13,292	87%	\$19,311
⊕ Fashion	\$920	89%	\$145,957	86%	\$189,523	87%	\$336,400
⊕ Audio	\$43	87%	\$24,820	85%	\$34,670	85%	\$59,533
⊕ Computers & Electronics	\$667	87%	\$150,230	82%	\$171,547	84%	\$322,444
⊕ Appliances	\$383	83%	\$41,262	85%	\$70,676	84%	\$112,321
⊕	\$23	85%	\$27,183	79%	\$20,138	83%	\$47,344
⊕ Health & Beauty	\$977	83%	\$210,890	81%	\$298,460	82%	\$512,327
⊕ Industry & Commerce	\$399	85%	\$13,455	79%	\$19,600	82%	\$33,454
⊕ Art & Hobby	\$808	81%	\$246,878	80%	\$245,906	81%	\$495,592
⊕ Auto	\$439	81%	\$58,483	81%	\$85,595	81%	\$144,517
⊕ Construction Tools		78%	\$14,713	80%	\$64,848	80%	\$79,561
⊕ Garden	\$311	78%	\$69,649	77%	\$58,502	78%	\$128,462
⊕ Houseware	\$315	81%	\$71,797	75%	\$108,075	77%	\$180,187
⊕ Other	\$440	77%	\$135,024	76%	\$106,407	76%	\$241,871
⊕ Home & Office Furniture	\$909	76%	\$262,037	74%	\$289,289	75%	\$553,235
⊕ Supplies		85%	\$1,295	69%	\$2,361	75%	\$3,656
Total	995	82%	\$1,515,184	80%	\$1,818,565	81%	\$3,345,744

Εικόνα 103 : Πίνακας Περιθωρίου Κερδών Εγχώριων πωλήσεων

Year	2018		
Product Category	Revenue	Revenue PY	% YoY Growth
⊕ Construction Tools	\$314,965	\$72,788	433%
⊕ Food & Beverages	\$59,344	\$27,381	217%
⊕ Supplies	\$11,470	\$6,008	191%
⊕ Appliances	\$339,981	\$197,902	172%
⊕ Houseware	\$527,452	\$338,742	156%
⊕ Industry & Commerce	\$96,155	\$64,323	149%
⊕ Auto	\$405,961	\$280,410	145%
⊕ Health & Beauty	\$1,403,041	\$979,475	143%
⊕ Audio	\$166,947	\$117,289	142%
⊕ Fashion	\$878,287	\$673,350	130%
⊕ Home Utilities	\$273,437	\$216,548	126%
⊕ Computers & Electronics	\$803,844	\$710,157	113%
⊕ Home & Office Furniture	\$1,425,266	\$1,264,265	113%
⊕ Art & Hobby	\$1,150,985	\$1,150,120	100%
⊕ Garden	\$289,318	\$339,860	85%
⊕ Other	\$510,085	\$645,407	79%
⊕	\$93,587	\$126,815	74%
Total	\$8,699,763	\$7,249,747	120%

Εικόνα 102 : Ρυθμός Ανάπτυξης Εσόδων ανά έτος 2017-18

Η εταιρεία από τα συνολικά έσοδα 16εκ. \$ απορροφά μεικτό κέρδος 3,3εκ. R\$, με μεικτό περιθώριο κερδών, 81%.

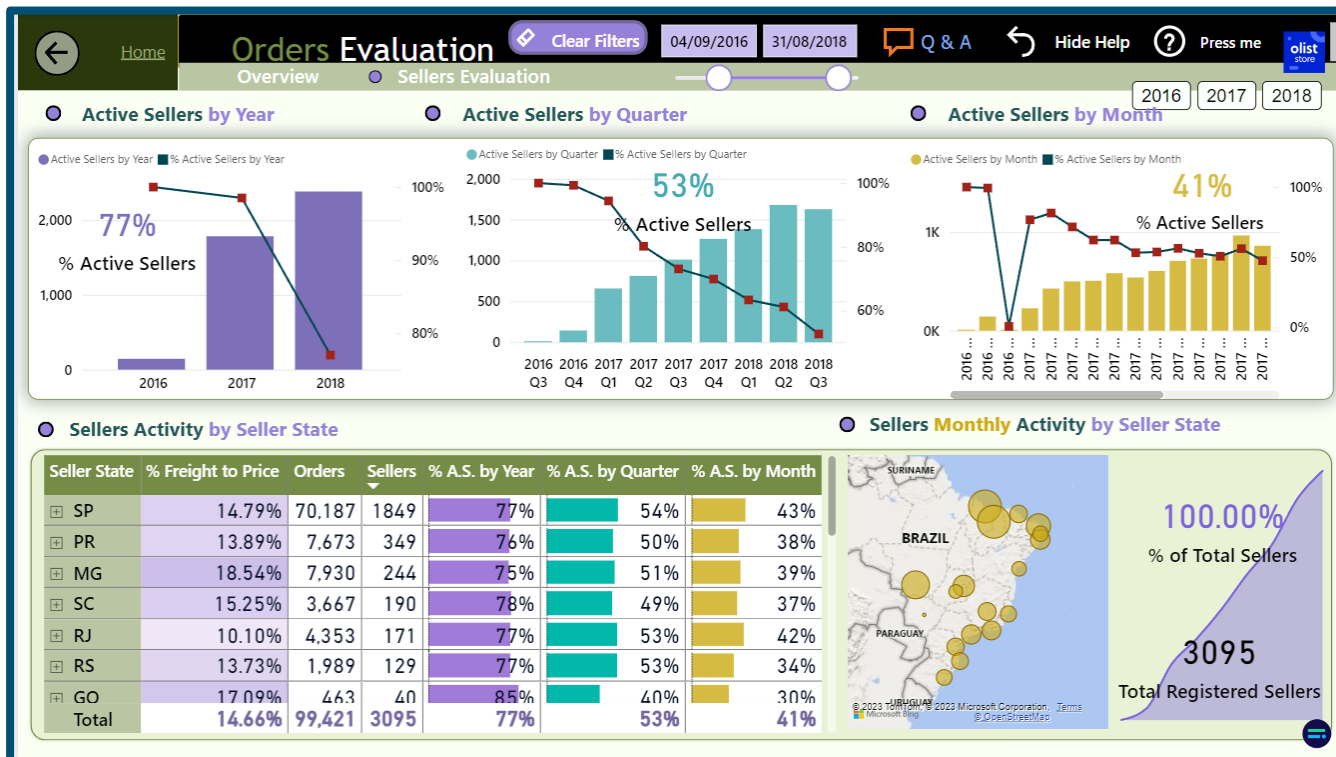
Οι κατηγορίες **Κήπος** και **Διάφορα** φαίνεται να έχουν αρνητικό ρυθμό ανάπτυξης ,ίσως άλλων ανταγωνιστικών αγορών στις περιοχές αυτές, και γι' αυτό η εταιρεία καλό ήταν να επικεντρωθεί σε συνεργασίες με πωλητές που εμπορεύονται άλλες κατηγορίες προϊόντων. Οι πιο υποσχόμενες κατηγορίες και από άποψη ρυθμού ανάπτυξης αλλά και περιθωρίου κέρδους είναι οι κατηγορίες **Τρόφιμα και Ποτά**, **Κατασκευαστικά Εργαλεία** ,**Προμήθειες** ,**Συσκευές** και **Είδη Μόδας**. Ασφαλώς, οι κατηγορίες προϊόντων που φέρνουν αυτήν περίοδο τα περισσότερα έσοδα είναι οι κατηγορίες **Έπιπλα Γραφείου και Σπιτιού** και **Υγεία και Ομορφιά**. Η ενίσχυση των εμπορικών σχέσεων τόσο με τις περιοχές που προσφέρουν τα περισσότερα έσοδα της εταιρείας και η διευκόλυνση των Πωλητών που εμπορεύονται τις πιο επικερδής κατηγορίες προϊόντων σε αυτές τις περιοχές, θα επιτρέψουν την δημιουργία πιστών πελατών ,οι οποίοι θα προσφέρουν σταθερά κέρδη στην εταιρεία.

Τέλος, από την ανάλυση που έγινε για το προφίλ των πελατών της εταιρείας προέκυψε το συμπέρασμα ότι **όσο περισσότερα προϊόντα έχει μια παραγγελία , τόσο φθηνότερα μεταφορικά κόστη** έχει άρα τόσο μεγαλύτερο περιθώριο κέρδους.

Αξιολόγηση ενεργής δραστηριότητας των Πωλητών στην πλατφόρμα της εταιρείας

Το επιχειρηματικό μοντέλο της επιχείρησης βασίζεται στην παρακράτηση ενός ποσοστού από την συνολική τιμή της παραγγελίας προϊόντων που εμπορεύονται οι συνεργάτες – πωλητές της εταιρείας . Η προσέλκυση νέων συνεργατών στην πλατφόρμα της εταιρείας, αυξάνει την ποικιλία των προϊόντων ,τα έσοδα και αναπόφευκτα τα κέρδη της εταιρείας. Άρα η οικονομική αξιολόγηση για το μέλλον της εταιρείας , είναι ζωτικής σημασίας να συμπεριλαμβάνει και την αξιολόγηση της δραστηριότητας των συνεργατών της εταιρείας.

Ένα μετρικό που βοηθάει στην αξιολόγηση αυτή, είναι το ποσοστό που δείχνει πόσοι πωλητές έχουν πραγματοποιήσει έστω και μία πώληση τον τελευταίο μήνα, τρίμηνο και έτος ή με άλλα λόγια έχουν παραμείνει ενεργοί στην πλατφόρμα της εταιρείας. Στην αναφορά έχουν προστεθεί τρία μετρικά, για διάφορα χρονικά διαστήματα δραστηριότητας ,καθώς κάποιοι Πωλητές έχουν εποχικά είδη και αν το διάστημα δραστηριότητας ήταν λιγότερο από χρόνο, πιθανότατα τα αποτελέσματα να μην ήταν έγκυρα.



Εικόνα 104 : Καρτέλα Αξιολόγησης Πωλητών – Power BI Service

Η αρχική αξιολόγηση μας δείχνει ότι ο **απόλυτος αριθμός των ενεργών πωλητών βαίνει διαρκώς αυξανόμενος** για οποιοδήποτε διάστημα δραστηριότητας, κάτι το οποίο οπωσδήποτε προσμετράται ως θετικό ποιοτικό χαρακτηριστικό. Το δεύτερο σημαντικό ποιοτικό χαρακτηριστικό το οποίο παρατηρούμε, είναι ότι το **σχετικό ποσοστό των ενεργών Πωλητών βαίνει διαρκώς μειούμενο για οποιοδήποτε διάστημα δραστηριότητας**. Αυτό δηλώνει ότι το ποσοστό των ενεργών πωλητών μειώνεται. Συγκεκριμένα, τον τελευταίο μήνα μόλις **το 41% του συνόλου των εγγεγραμμένων Πωλητών, δέχτηκαν παραγγελία** από πελάτη μέσω της πλατφόρμας της εταιρείας. Το αντίστοιχο ποσοστό **για το τελευταίο τρίμηνο**, όπου έχουμε δεδομένα, είναι το **53%** και το αντίστοιχο ποσοστό ενεργών Πωλητών **το τελευταίο έτος είναι 77%**

Κάτι τέτοιο δικαιολογείται μερικώς από το ότι η επιχείρηση έχει λίγα χρόνια δραστηριότητας και η εισροή πολλών νέων συνεργατών αυξάνει τον παρονομαστή του κλάσματος. Αυτό που έχει την βαρύνουσα σημασία στην ένδειξη αυτή βέβαια, είναι το γεγονός πως οι πελάτες δεν πραγματοποιούν επαναληπτικές παραγγελίες από την πλατφόρμα της εταιρείας και άρα δεν επιλέγουν σαν αποτέλεσμα τους ίδιους πωλητές. Αν η εταιρεία συνεχίσει να μην διατηρεί την ενεργή δραστηριότητα των πωλητών υψηλή, **κινδυνεύει να χάσει τους συνεργάτες της από άλλες ανταγωνιστικές πλατφόρμες**, αν αυτό δεν έχει συμβεί ήδη.

Όσο αφορά το ποσοστό δραστηριότητας των πωλητών ανά πολιτεία, παρατηρούμε ότι οι πολιτείες με τους περισσότερους πωλητές (**SP, PR, MG, SC, RJ**) δεν έχουν το μεγαλύτερο ποσοστό ενεργών πωλήσεων. Αυτό είναι σχετικά λογικό, καθώς όσο περισσότεροι πωλητές από την ίδια πολιτεία τόσο μεγαλύτερη και η πιθανότητα ο ανταγωνισμός να είναι αυξημένος και να διαμοιράζονται οι παραγγελίες σε περισσότερους Πωλητές.

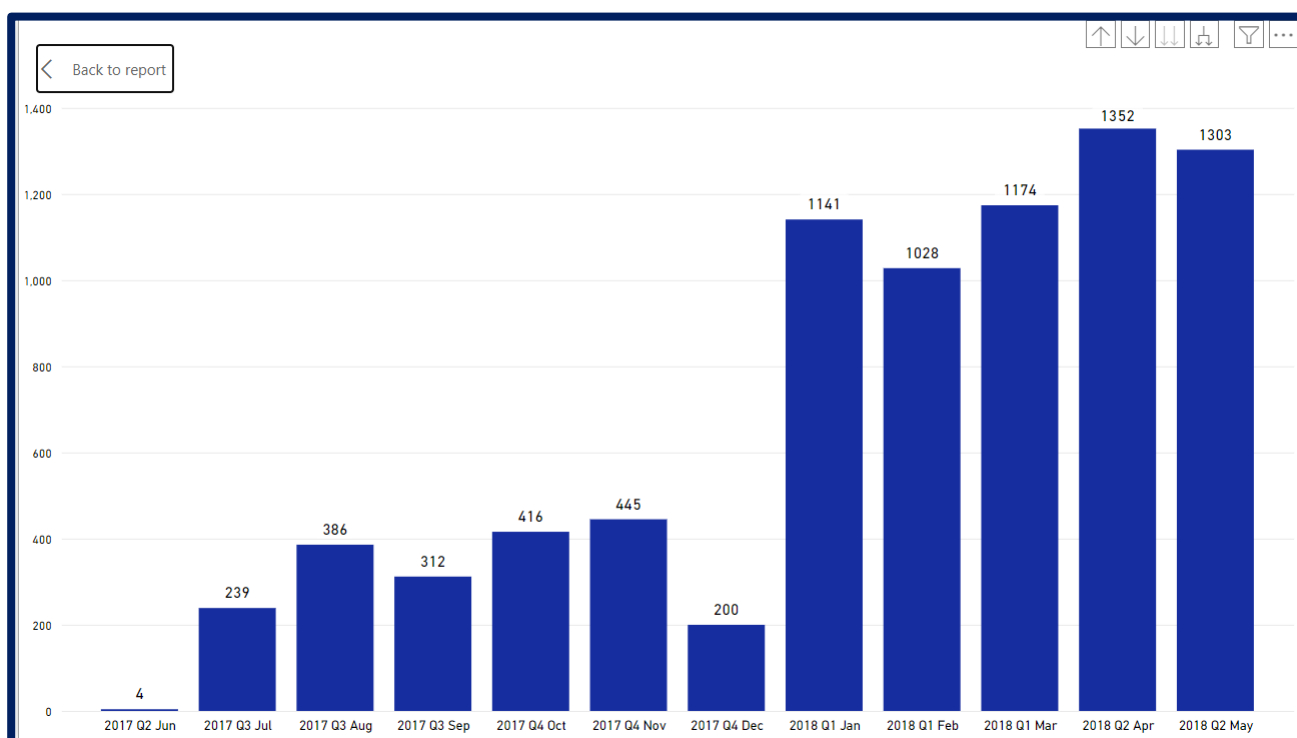
Seller State	% Freight to Price	Orders	Sellers	Orders per Seller	% A.S. by Year	% A.S. by Quarter	% A.S. by Month
MA	32.26%	392	1	392.0	100%	100%	100%
PI	16.41%	12	1	12.0	100%	100%	100%
MT	25.70%	137	4	34.3	100%	75%	75%
RN	11.06%	51	5	10.2	100%	80%	60%
DF	17.53%	824	30	27.5	83%	63%	50%
PE	12.43%	406	9	45.1	89%	56%	44%
SP	14.79%	70,187	1849	38.0	77%	54%	43%
RJ	10.10%	4,353	171	25.5	77%	53%	42%
MG	18.54%	7,930	244	32.5	75%	51%	39%
CE	20.70%	91	13	7.0	85%	38%	38%
PR	13.89%	7,673	349	22.0	76%	50%	38%
SC	15.25%	3,667	190	19.3	78%	49%	37%
ES	21.90%	318	23	13.8	87%	52%	35%
RS	13.73%	1,989	129	15.4	77%	53%	34%
PB	7.88%	36	6	6.0	83%	67%	33%
BA	6.21%	569	19	29.9	89%	42%	32%
GO	17.08%	442	40	11.1	85%	40%	30%
Total	14.66%	99,421	3095	32.1	77%	53%	41%

Εικόνα 105 : Ποσοστό πρόσφατης δραστηριότητας ανά πολιτεία Πωλητή

Πρωταθλητές στην σχετική κατηγορία αναδεικνύονται οι περιοχές **MA, PI, MT, RN** και **DF**. Οι πολιτείες αυτές δεν έχουν πολλούς πωλητές εγγεγραμμένους στην πλατφόρμα της εταιρείας, αυτό σημαίνει ότι το δίκτυο της εταιρείας σε αυτές τις περιοχές δεν έχει επεκταθεί αρκετά.

4.3.5 Ανάλυση Προσέλκυσης Νέων Συνεργατών στο δίκτυο της εταιρείας – Τμήμα Μάρκετινγκ

Το επιχειρηματικό μοντέλο της εταιρείας στηρίζεται στην συνεργασία της εταιρείας με Πωλητές ανά την Βραζιλία. Οι Πωλητές βρίσκουν χώρο να προωθήσουν τα προϊόντα τους στην πλατφόρμα της εταιρείας και χρησιμοποιούν το δίκτυο συνεργατών που έχει η Olist για να μεταφέρουν τα προϊόντα τους. Οπότε η ανάλυση δυνητικών συνεργασιών με Πωλητές είναι κρίσιμη για το μέλλον της εταιρείας για να συνεχίσει να δημιουργεί έσοδα. Οι επιχειρήσεις που ενδιαφέρονται να εγγραφούν στην πλατφόρμα της εταιρείας έρχονται σε επαφή με την εταιρεία με κάποιον τρόπο (αναζήτηση στο διαδίκτυο, μείλ, μέσω διαφημίσεων). Όταν αυτό το ενδιαφέρον είναι διαπιστευμένο τότε, καταγράφεται ως δυνητικός συνεργάτης από το σύστημα της εταιρείας. Από εκεί και πέρα αναλαμβάνουν οι αντιπρόσωποι πωλήσεων και ανάπτυξης πωλήσεων της εταιρείας για την γνωριμία με τον υποψήφιο συνεργάτη.



Εικόνα 106 : Οι διαπιστευμένες καινούργιες επαφές του μάρκετινγκ μέσα στο χρόνο

Συνολικά τα δεδομένα που διαθέτουμε δείχνουν ότι η εταιρεία έχει συνολικά 8000 διαπιστευμένες επαφές τα έτη 2017 και 2018. Από το διάγραμμα στηλών, παρατηρούμε ότι οι περισσότερες διαπιστευμένες επαφές για συνεργασία με την εταιρεία έγιναν το έτος 2018 , με κατά μέσο όρο 1200 καινούργιες επαφές ανά μήνα.

Αξιολόγηση απόκτησης Νέων Συνεργατών - Partner Acquisition

Δεν γνωρίζουμε για όλες τις επαφές σε ποια κατηγορία προϊόντων δραστηριοποιούνται, παρά μόνο για όσες τελικά προχώρησαν σε συμφωνία με την εταιρεία. Γνωρίζουμε όμως πως ήρθαν σε επαφή με την εταιρεία, δηλαδή αν έψαξαν οι ίδιοι οι Πωλητές για την Olist, αν ήρθαν



Εικόνα 107 : Οι πηγές των διαπιστευμένων νέων επαφών του Μάρκετινγκ

στην ιστοσελίδα της εταιρείας μέσω κάποιας άλλης ή από συγκεκριμένο τύπο διαφήμισης. Το διάγραμμα δέντρου μας διευκολύνει να αντιληφθούμε ποιοι είναι οι καλοί τρόποι να έρθουν σε επαφή οι Πωλητές με την εταιρεία.

Η οργανική αναζήτηση (**Organic Search**) είναι ο πιο συχνός τρόπος που επιλέγουν να έρθουν σε επαφή με την εταιρεία οι Πωλητές, σε ποσοστό **28,7%**. Τα οργανικά αποτελέσματα αναζήτησης είναι οι **καταχωρίσεις χωρίς πληρωμή**, που εμφανίζονται σε μια σελίδα αποτελεσμάτων μηχανής αναζήτησης. Αυτά τα αποτελέσματα βασίζονται σε παράγοντες όπως, η συνάφεια με το ερώτημα αναζήτησης του χρήστη.

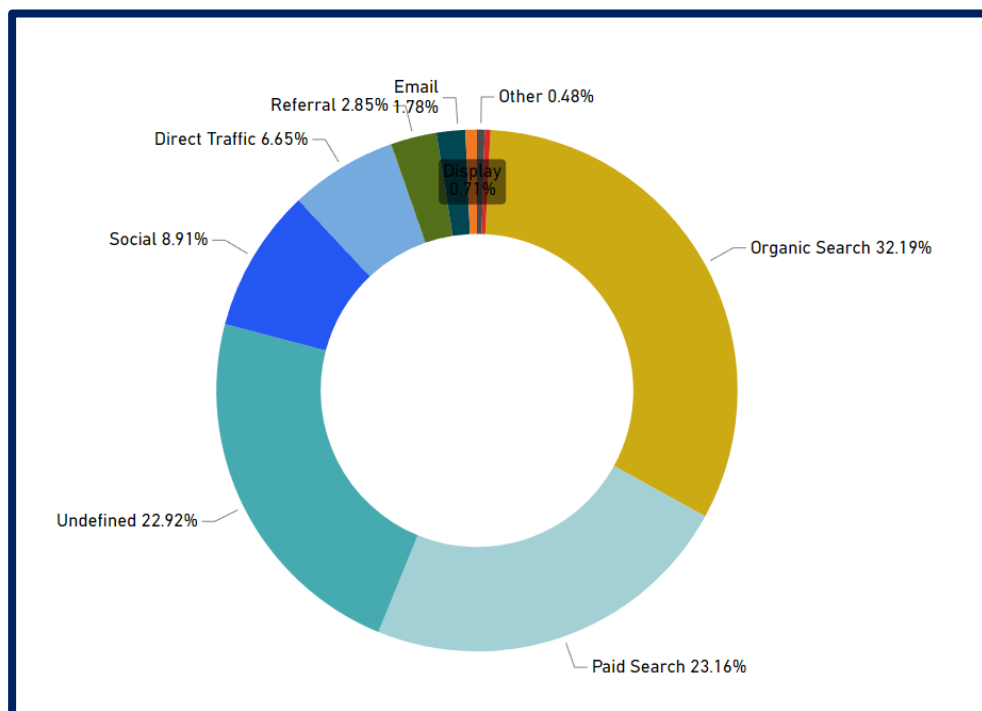
Στη δεύτερη θέση έρχονται τα πληρωμένα αποτελέσματα αναζήτησης (**Paid Search**) με ποσοστό σχεδόν **20%**. Οι πληρωμένες αναζητήσεις βρίσκονται στα πρώτα αποτελέσματα της αναζήτησης. Οι κοινωνικές αναζητήσεις έρχονται τρίτες σε σειρά με ποσοστό **17%**. Τα αποτελέσματα των αναζητήσεων αυτών φιλτράρονται με βάση κριτήρια που σχετίζονται με τον ίδιο τον χρήστη, το «κοινωνικό» του περιβάλλον και την χρήση που κάνει στο διαδίκτυο. Τέλος **14,5%** έρχεται σε επαφή με την εταιρεία με άγνωστο τρόπο (**Undefined**).

Αν αξιολογήσουμε **το ποσοστό των συμφωνιών που έγιναν με βάση τον αρχικό τρόπο επικοινωνίας** με τον υποψήφιο συνεργάτη, τα αποτελέσματα είναι αρκετά παρόμοια.

Οι πιο αποδοτικοί τρόποι επαφής είναι σε σειρά οι εξής :

1. **Οργανική Αναζήτηση** με ποσοστό συμφωνιών **32%**,

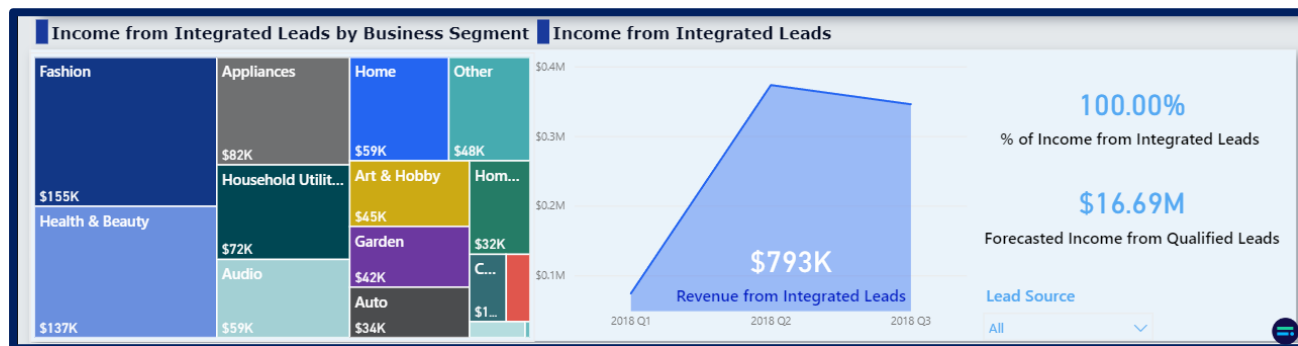
2. Πληρωμένη Αναζήτηση με ποσοστό συμφωνιών **23%**,
3. Ακαθόριστος Τρόπος με ποσοστό συμφωνιών **23%** και
4. Κοινωνική Αναζήτηση με ποσοστό συμφωνιών **9%**.



Εικόνα 108 Το ποσοστό των συμφωνιών ανά πηγή

Οι υπόλοιποι τρόποι δίνουν μικρή απόδοση των επενδεδυμένων κεφαλαίων (**ROI – Return of Investment**) στην προώθηση της εταιρείας.

Τα έσοδα που είχε η εταιρεία από τις πωλήσεις, από τις συμφωνίες που προέκυψαν, από τους 4 αυτούς τρόπους πρώτης επαφής (**Οργανική Αναζήτηση, Πληρωμένη Αναζήτηση, Ακαθόριστος Τρόπος και Κοινωνική Αναζήτηση**) αποτελούν σχεδόν το **92%** εσόδων της και σε απόλυτο αριθμό **726χιλ. R\$**.



Εικόνα 109 : Τα έσοδα από τις ενσωματωμένες νέες συμφωνίες – Power BI Service

Συνολικά, οι αναζητήσεις σε μηχανή αναζήτησης είναι ο τρόπος που επιλέγουν οι Πωλητές να έρθουν σε επαφή με την εταιρεία και τελικά είναι πιο κερδοφόρες για την εταιρεία. Οπότε η εταιρεία πρέπει να αξιοποιήσει τους πόρους της Διαφήμισης, αποκλειστικά στις **Πληρωμένες Αναζητήσεις**. Ακόμη θα ήταν χρήσιμο για την εταιρεία να ξεκαθαρίσουν ποια είναι η μορφή της επικοινωνίας, που ανήκει στην κατηγορία **Ακαθόριστος Τρόπος**, καθώς αποτελεί μεγάλο «κομμάτι της πίτας» του τρόπου που έρχονται σε επαφή οι Πωλητές.



Εικόνα 110: Σημαντικά στοιχεία για το κλείσιμο νέων συμφωνιών

Επόμενο σημαντικό κομμάτι της ανάλυση των Δυνητικών Συνεργατών - Πωλητών, είναι να γνωρίζει η εταιρεία το ποσοστό των επαφών που είχε η εταιρεία που κατέληξαν στο κλείσιμο συμφωνίας, αλλά και το χρόνο που χρειάστηκε για να γίνει αυτό.

Στην ανάλυση που έγινε βρέθηκαν τα επόμενα συγκεντρωτικά στοιχεία. Από το σύνολο των **8000 πιθανών συνεργατών**, έχουν κλειστεί **842 συμφωνίες**, ποσοστό **10,5%**. Ενώ ο μέσος χρόνος κλεισίματος της συμφωνίας, από την πρώτη επαφή μέχρι την ολοκλήρωση της συμφωνίας, είναι σχεδόν **49 ημέρες** κατά μέσο όρο.

Το ποσοστό 10% είναι αρκετά ικανοποιητικό και δηλώνει ότι αφενός οι πωλητές εμπιστεύονται την εταιρεία και

αφετέρου οι **Εκπρόσωποι Πωλήσεων** (Sales Representatives) και οι **Εκπρόσωποι Ανάπτυξης Πωλήσεων** (Sales Development Representatives) κάνουν ικανοποιητική δουλειά στην ανάπτυξη του δικτύου της εταιρείας.

Μια εποπτική εικόνα, για την μετατροπή των Διαπιστευμένων Επαφών (**Qualified Leads**), σε Κλεισμένες Συμφωνίες (**Closed Deals**) και η μετατροπή αυτών σε Ενσωματωμένες Συμφωνίες (**Integrated Leads**) μέσα στον χρόνο, αποτυπώνεται στην σχετική καρτέλα Αξιολόγησης που δημιουργήθηκε.



Εικόνα 111 : Η μετατροπή των διαπιστευμένων επαφών σε συμφωνίες με πωλητές – Power Bi Service

Αξιολόγηση εσόδων από τις ενσωματωμένες συμφωνίες

Ένα ποσοστό των συμφωνιών που έχουν κλειστεί με τους Πωλητές , ήδη έχουν ενσωματωθεί (**Revenue from Integrated Leads**) στην πλατφόρμα της εταιρείας και έχουν ξεκινήσει να αποδίδουν κέρδη στην εταιρεία. Ασφαλώς ,είναι χρήσιμο η εταιρεία να γνωρίζει ποιες κατηγορίες προϊόντων έχουν προσφέρει τα περισσότερα έσοδα και πόσο γρήγορα κλείστηκαν οι συμφωνίες για την κάθε κατηγορία προϊόντων.

Business Segment (Bins)	Closed Deals	Closing Time	Revenue	Revenue per Deal
+ Fashion	57	44.72	\$155,387	2,726.09
+ Health & Beauty	133	38.45	\$136,795	1,028.53
+ Appliances	22	53.00	\$81,781	3,717.33
+ Household Utilities	84	49.19	\$71,858	855.46
+ Audio	64	63.23	\$59,349	927.33
+ Home	105	50.09	\$58,602	558.11
+ Other	40	54.20	\$47,970	1,199.24
+ Art & Hobby	77	47.18	\$44,982	584.18
+ Garden	69	60.71	\$41,597	602.85
+ Auto	77	50.45	\$34,429	447.13
+ Home & Office Furniture	36	34.97	\$32,293	897.02
+ Computers & Electronics	34	55.00	\$14,139	415.85
+ Office	13	36.46	\$9,105	700.41
+ Food & Beverages	26	31.58	\$5,045	194.04
+ Supplies	5	44.20	\$430	85.94
Total	842	48.44	\$792,957	941.75

Εικόνα 112 : Πίνακας νέων συμφωνιών ανά επιχειρηματικό κλάδο

Όπως φαίνεται, η πιο επικερδής κατηγορίες προϊόντων που έχουν ενσωματωθεί στην πλατφόρμα της Olist είναι η **Ένδυση** και **Υγεία & Ομορφιά**, ενώ οι συγκεκριμένες κατηγορίες έχουν επίσης μικρό χρόνο κλεισίματος συμφωνίας και προσφέρουν περισσότερα έσοδα ανά συμφωνία. Ξεχωρίζει επίσης και η κατηγορία των **Συσκευών**, οι οποία αν και έχει μεγάλο μέσο χρόνο κλεισίματος μια συμφωνίας, προσδίδει **έσοδα 3,7χιλ. \$** ανά συμφωνία , που σημαίνει ότι είναι επικεφαλής στον δείκτη **Έσοδα ανά Συμφωνία**. Οι εκπρόσωποι και οι πόροι της εταιρείας πρέπει να επικεντρωθούν σε νέες συμφωνίες με επιχειρηματίες που εμπορεύονται **Συσκευές**.

Μια εποπτική εικόνα ,για την μετατροπή των Διαπιστευμένων Επαφών (**Qualified Leads**), σε Κλεισμένες Συμφωνίες (**Closed Deals**) και η μετατροπή αυτών σε Ενσωματωμένες Συμφωνίες (**Integrated Leads**) μέσα στον χρόνο, αποτυπώνεται στην σχετική καρτέλα Αξιολόγησης που δημιουργήθηκε.

Αξιολόγηση Προφίλ Διαπιστευμένων Επαφών

Για τους μάνατζερ και τους αντιπροσώπους στο τμήμα Πωλήσεων και Μάρκετινγκ , είναι χρήσιμο **να έχουν συνοπτικά όλες τις κρίσιμες πληροφορίες** για τους πιθανούς συνεργάτες της εταιρείας, **οργανωμένες σε ένα σημείο** ώστε να μπορούν να προετοιμάζονται καλύτερα για τις συναντήσεις μαζί τους και να κάνουν αξιολόγηση του συνολικού προφίλ τους. Γι' αυτούς τους σκοπούς , ένα σύστημα Επιχειρηματικής Ευφυΐας μπορεί να αποδειχτεί εξαιρετικά χρήσιμο.

Σε αυτή την καρτέλα βρίσκονται όλες οι διαθέσιμες πληροφορίες για τις επιχειρήσεις που ενδιαφέρονται να μπουν στην πλατφόρμα της εταιρείας. Συγκεκριμένα ,οι χρήστες τους συστήματος μπορούν να αντλήσουν πληροφορίες όπως:

- Τα βασικά στοιχεία του ιδιοκτήτη
- Τα στοιχεία επικοινωνίας
- Την ψυχογραφική περιγραφή του επιχειρηματία ,σύμφωνα με την κατηγοριοποίηση κατά **DiSC Personality Types**
- Ο τύπος της επιχείρησης ,δηλαδή αν είναι Online επιχείρηση ή βιομηχανία
- Αν είναι κατασκευαστής ή όχι
- Τον επιχειρηματικό κλάδο που δραστηριοποιούνται
- Το μέγεθός της επιχείρησης ,σε οικονομικό επίπεδο και επίπεδο καταλόγου προϊόντων
- Αν έχει τα σωστά νομικά έγγραφα
- Αν έχουν τα προϊόντα barcode (GTIN)

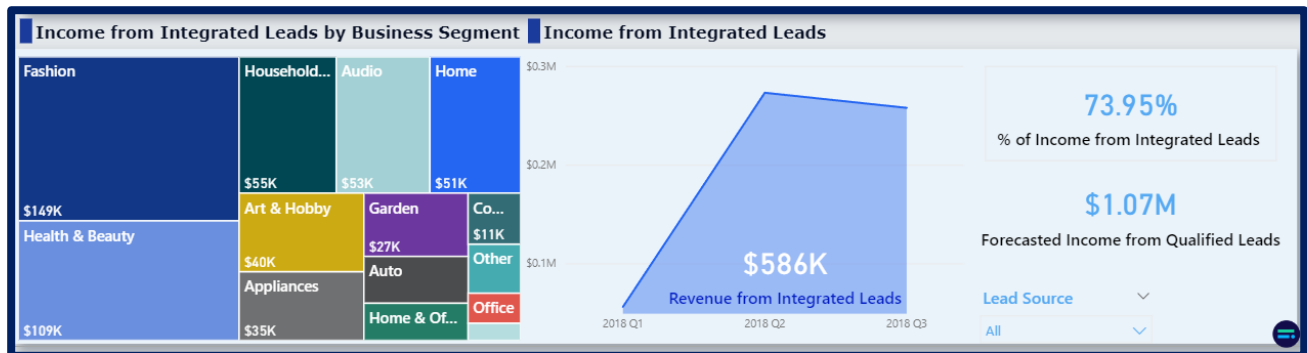


Εικόνα 113 : Καρτέλα Αξιολόγησης των διαπιστευμένων επαφών - Power BI Service

Παρατηρούμε ότι οι περισσότερες επιχειρήσεις που έχουν κλείσει συμφωνία με την εταιρεία ανήκουν στην κατηγορία **Online Medium**, δηλαδή μεσαία επιχείρηση που δεν έχει φυσικό κατάστημα, με ποσοστό **39%** επί των συνολικών συμφωνιών. Τα μέσα δηλωμένα έσοδα για αυτές τις επιχειρήσεις ανέρχεται στις **34χιλ. \$/μήνα**.

Δεύτερη σε αριθμό συμφωνιών είναι η κατηγορία **Online Big** όπου αποτελούν το **15%** των συμφωνιών και έχουν δηλώσει μέσο μηνιαίο εισόδημα **1,13εκ. \$**.

Συνολικά οι **αυτές κατηγορίες έχουν αποδώσει το 74% των εσόδων** της εταιρείας, από τις πρόσφατα ενταγμένες εταιρείες στην πλατφόρμα της εταιρείας με το ποσό των **586χιλ. \$**.

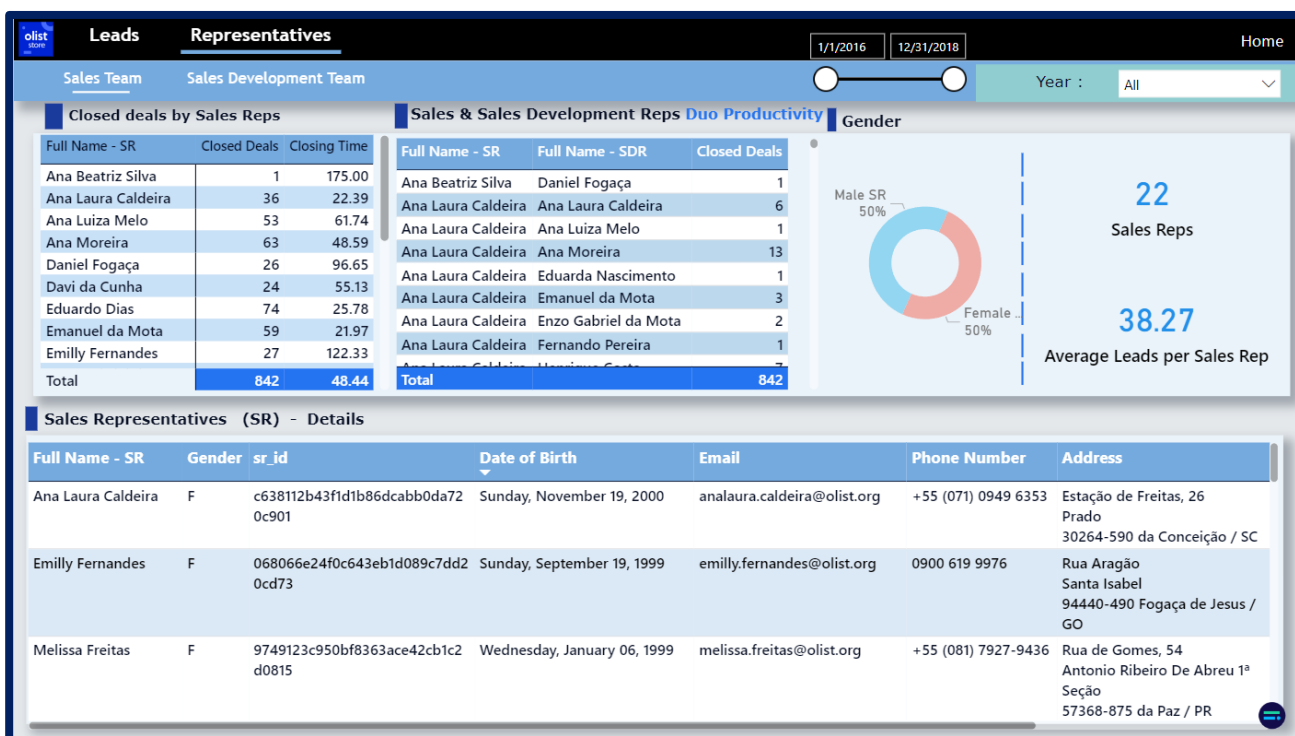


Εικόνα 114 : Τα έσοδα από τις κατηγορίες Online Big και Online Medium – Power Bi Service

Ενώ, με βάση αυτά τα νούμερα προβλέπεται να προσφέρουν συνολικά οι κλεισμένες συμφωνίες **1,07εκ. \$**.

Αξιολόγηση Αντιπροσώπων

Η αξιολόγηση των αντιπροσώπων των τμημάτων Πωλήσεων και Μάρκετινγκ, ενδιαφέρει τις αντίστοιχες Διευθύνσεις, ώστε να μπορούν να βλέπουν την απόδοσή τους. Σε κάθε καινούργια συμφωνία που κλείνεται, αρχικά έρχονται σε πρώτη επαφή με τον πιθανό συνεργάτη ένας **Εκπρόσωπος Ανάπτυξης Πωλήσεων (Sales Development Rep)** και στην συνέχεια, σε ποιο προχωρημένο στάδιο των συνομιλιών εμπλέκονται οι **Εκπρόσωποι Πωλήσεων (Sales Rep)** για να κλείσουν την συμφωνία. Κάθε εκπρόσωπος έχει μια εταιρική ταυτότητα και με βάση αυτή δημιουργήσαμε ένα σύστημα συγκέντρωσης των κλεισμένων συμφωνιών τους.



Εικόνα 115 : Καρτέλα Αξιολόγησης των Αντιπροσώπων Πωλήσεων

Από αυτή την καρτέλα ο μάνατζερ του τμήματος είναι σε θέση βλέπει όλα τα στοιχεία των πωλητών του και να αξιολογεί την επίδοση των πωλητών και πωλητών ανάπτυξης .

Στους πίνακες της καρτέλας φαίνονται στοιχεία για τον **Αντιπρόσωπο Πωλήσεων**:

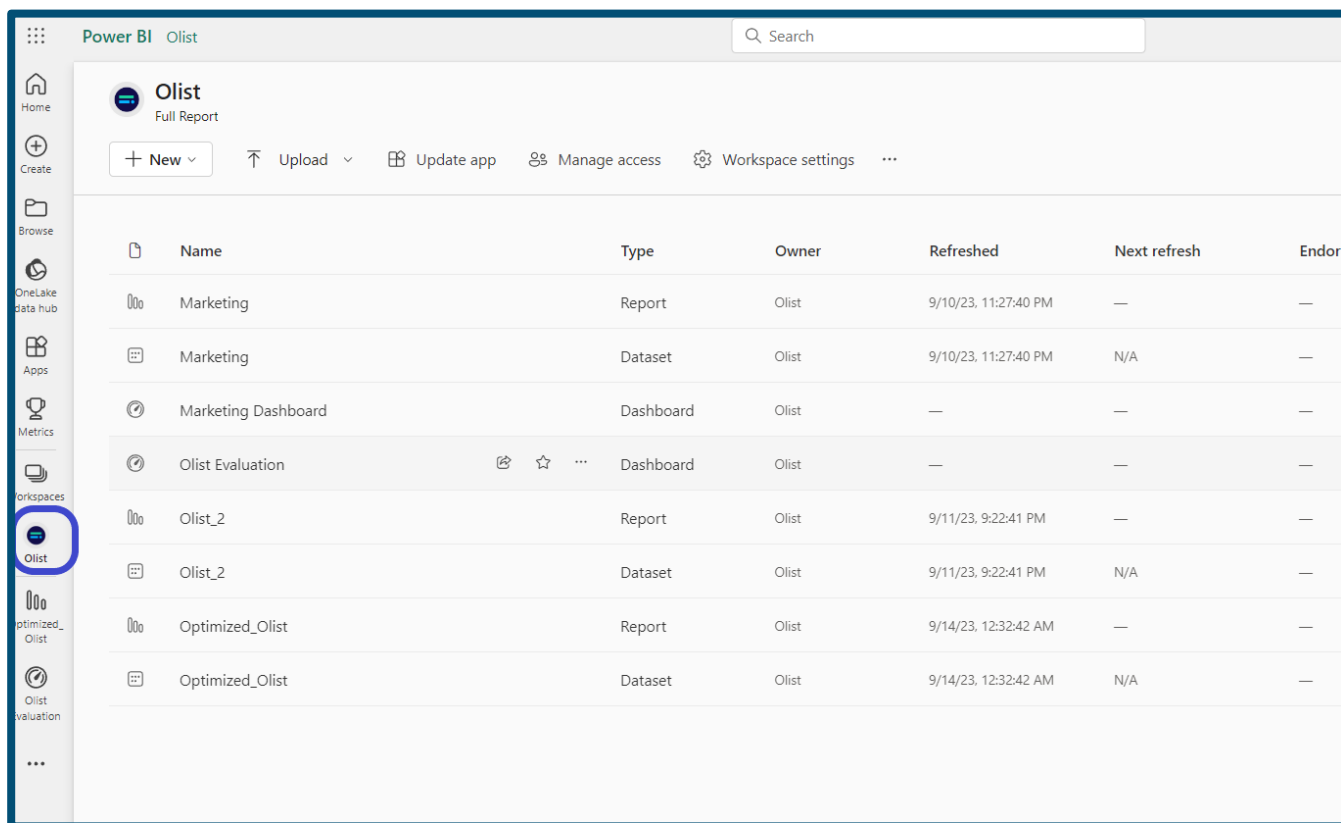
- Για τον **αριθμό των συμφωνιών** που έκλεισε
- τον **χρόνο που χρειάστηκε να κλείσει την συμφωνία**
- και με ποιον Αντιπρόσωπο Ανάπτυξης Πωλήσεων **συνεργάστηκε πιο αποδοτικά**.

Αντίστοιχη καρτέλα δημιουργήθηκε με τα στοιχεία των **Αντιπροσώπων Ανάπτυξης Πωλήσεων**.

4.3.6 Ολοκλήρωση του self-service συστήματος στο Power BI Service

Μετά την ολοκλήρωση των αναφορών, είναι η ώρα να δημοσιευτούν σε ένα κοινόχρηστο χώρο, ώστε να μπορούν να χρησιμοποιηθούν και να «καταναλωθούν» από τους διάφορους χρήστες της επιχείρησης. Ο χώρος αυτός παρέχεται στο cloud μέρος της εφαρμογής, το **Power BI Service**.

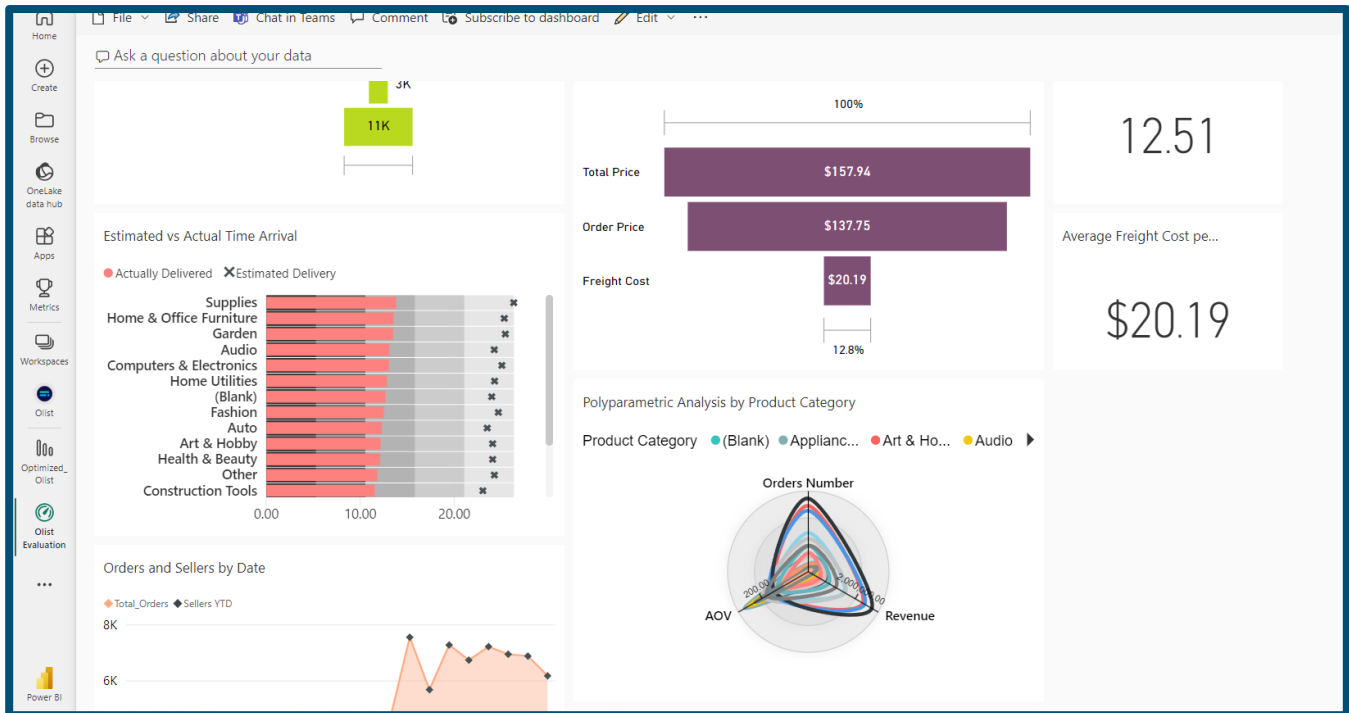
Εκεί δημιουργούμε ένα χώρο εργασίας, θέτοντας ορισμένους περιορισμούς για το ποιοι επιθυμούμε να έχουν πρόσβαση. Σε αυτόν τον χώρο έχουμε δημιουργήσει και προσθέσει



Εικόνα 116 : Χώρος Εργασίας αφιερωμένος για την αξιολόγηση της εταιρείας – Power BI Service

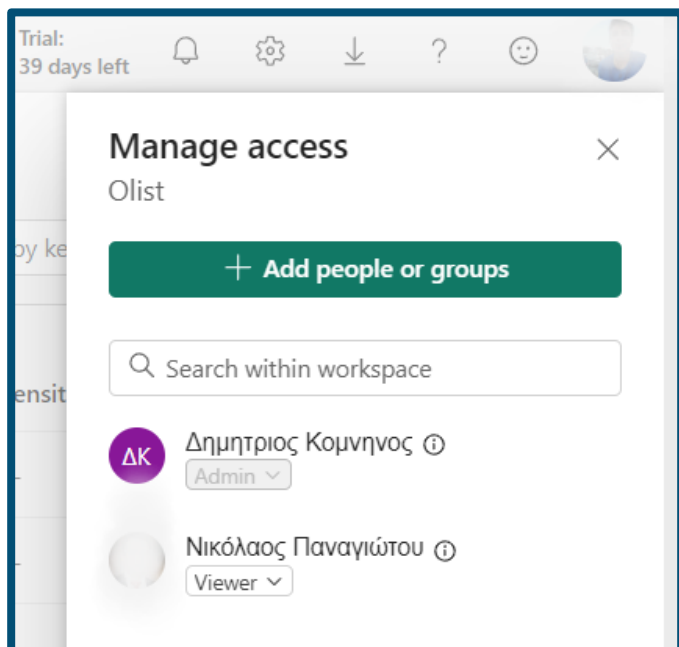
Πίνακες Ελέγχου ,ενώ επίσης έχουν προστεθεί τα μοντέλα δεδομένων και οι αναφορές που δημιουργήσαμε.

Αν επιλέξουμε τον πίνακα ελέγχου που δημιουργήσαμε, μπορούμε να δούμε τα σημαντικότερα στοιχεία της αναφοράς , οργανωμένα σε αυτόν τον πίνακα. Πατώντας κάποια σπτικοποίηση από αυτές οδηγούμαστε στην αντίστοιχη αναφορά ,από την οποία προήλθε. Ο πίνακας ελέγχου βοηθάει σε μια πρώτη γνωριμία με τα δεδομένα και δίνει μια εμποπτική εικόνα αυτών στον χρήστη.



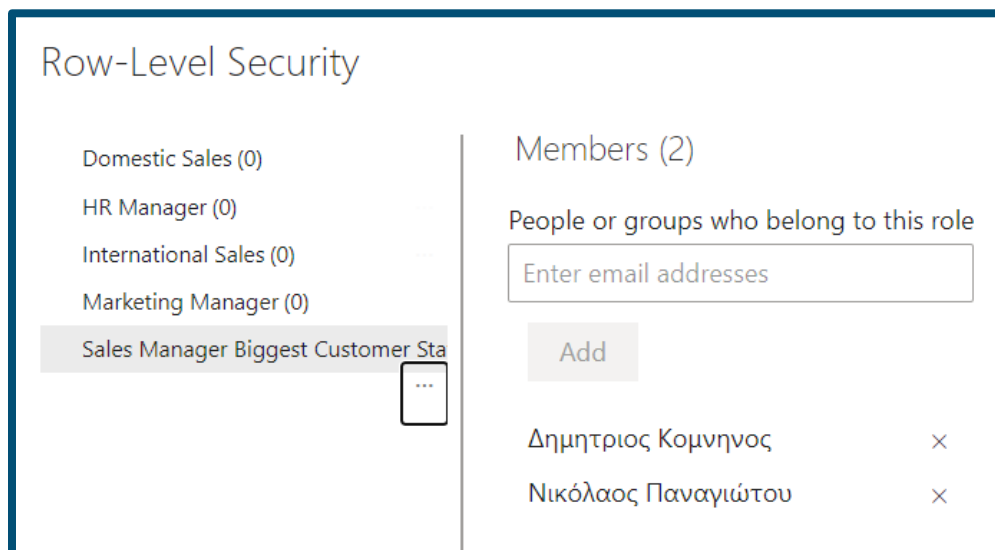
Εικόνα 117 : Πίνακας ελέγχου αξιολόγησης της εταιρείας – Power BI Service

Για την κατάλληλη οργάνωση του περιεχομένου ,ελέγχουμε τα άτομα που θα έχουν πρόσβαση σε αυτό το χώρο εργασίας ,ώστε να βλέπουν το περιεχόμενο και το διαχειρίζονται.



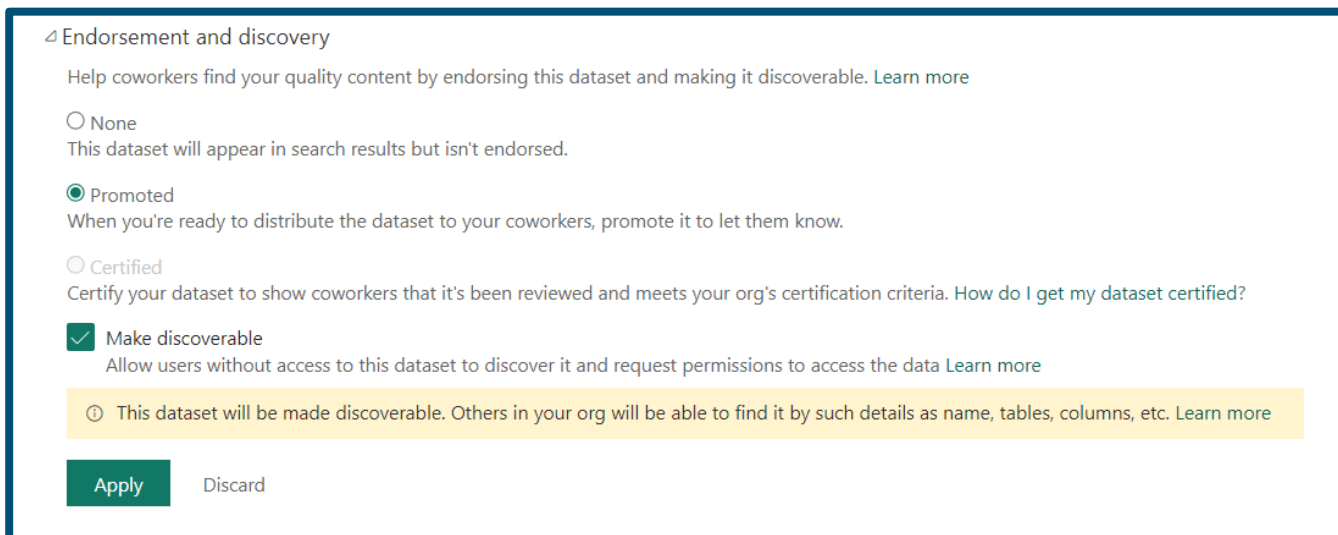
Εικόνα 118 : Ρόλοι στον χώρο εργασίας Olist

Επίσης, νωρίτερα είχαμε οργανώσει ρόλους για τον διευθυντή πωλήσεων για τις περιοχές της Βραζιλίας με τον μεγαλύτερο όγκο πωλήσεων ,δηλαδή τις περιοχές του Σάο Πάολο (SP) ,Ρίο ντε Ζανέιρο (RJ) και Μίνας Γκεράις (MG), για να γίνεται καλύτερη παρακολούθηση των δεδομένων. Με αυτόν τρόπο ,**οι αναφορές μένουν ίδιες σε δομή αλλά διαφορετικές σε περιεχόμενο μετά το φιλτράρισμα.**



Εικόνα 119 : Διαφορετικές Άδειες θέασης – Power BI Service

Επί πλέον θέτουμε το μοντέλο δεδομένων που δημιουργήσαμε ως **Endorsed** για να ενημερώσουμε τους συναδέλφους μας ότι, αυτό το μοντέλο έχει δημιουργηθεί και μπορεί να αξιοποιηθεί από αυτούς.



Endorsement and discovery

Help coworkers find your quality content by endorsing this dataset and making it discoverable. [Learn more](#)

None
This dataset will appear in search results but isn't endorsed.

Promoted
When you're ready to distribute the dataset to your coworkers, promote it to let them know.

Certified
Certify your dataset to show coworkers that it's been reviewed and meets your org's certification criteria. [How do I get my dataset certified?](#)

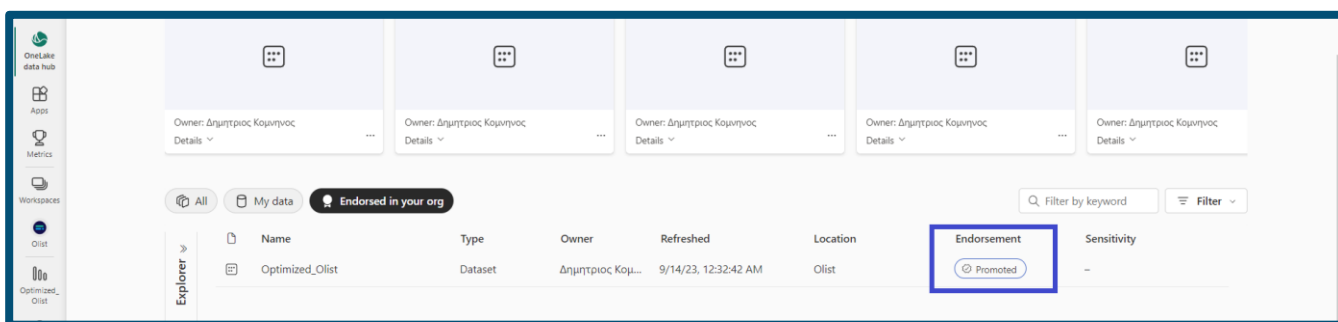
Make discoverable
Allow users without access to this dataset to discover it and request permissions to access the data [Learn more](#)

ⓘ This dataset will be made discoverable. Others in your org will be able to find it by such details as name, tables, columns, etc. [Learn more](#)

[Apply](#) [Discard](#)

Εικόνα 120 : Λειτουργία Endorsement – Power BI Service

Πλέον το μοντέλο δεδομένων εμφανίζεται στο Datasets Hub και βρίσκεται στην κατηγορία **Endorsed** και μπορεί να βρεθεί εύκολα από τους υπόλοιπους ενδιαφερόμενους χρήστες. Αυτή η ευκολία στην επικοινωνία της δραστηριότητας στο Service, είναι ένα από τα μεγάλα οφέλη των συστημάτων που εξετάζουμε. Πλέον όλοι οι χρήστες έχουν ένα κοινό σημείο αναφοράς, το κοινό μοντέλο (**Shared Dataset**), με βάση το οποίο μπορούν να δημιουργήσουν το δικό τους αξιόπιστο περιεχόμενο.

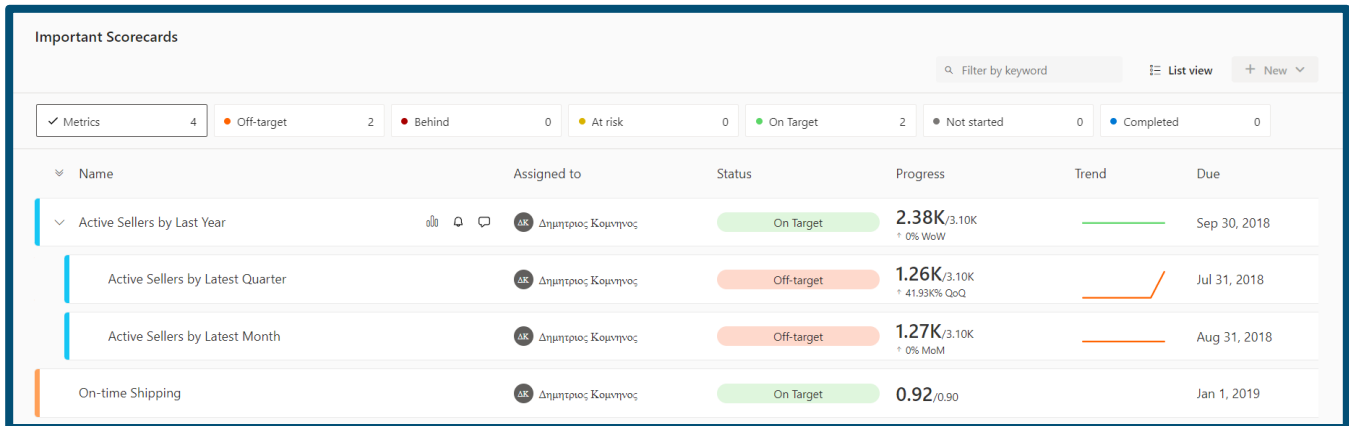


Name	Type	Owner	Refreshed	Location	Endorsement	Sensitivity
Optimized_Olist	Dataset	Δημητριος Κομ...	9/14/23, 12:32:42 AM	Olist	Promoted	-

Εικόνα 121 : Χώρος εμφάνισης των δημοσιοποιημένων μοντέλων δεδομένων – Power BI Service

Ακόμη, το Power BI Service δίνει την δυνατότητα για την παρακολούθηση σημαντικών KPIs μέσω των «Καρτελών Σκορ» (“**Metrics Scorecards**”). Στο παράδειγμα που δημιουργήσαμε κρίνεται ως σημαντικοί από την διοίκηση της εταιρείας να γνωρίζουν δύο

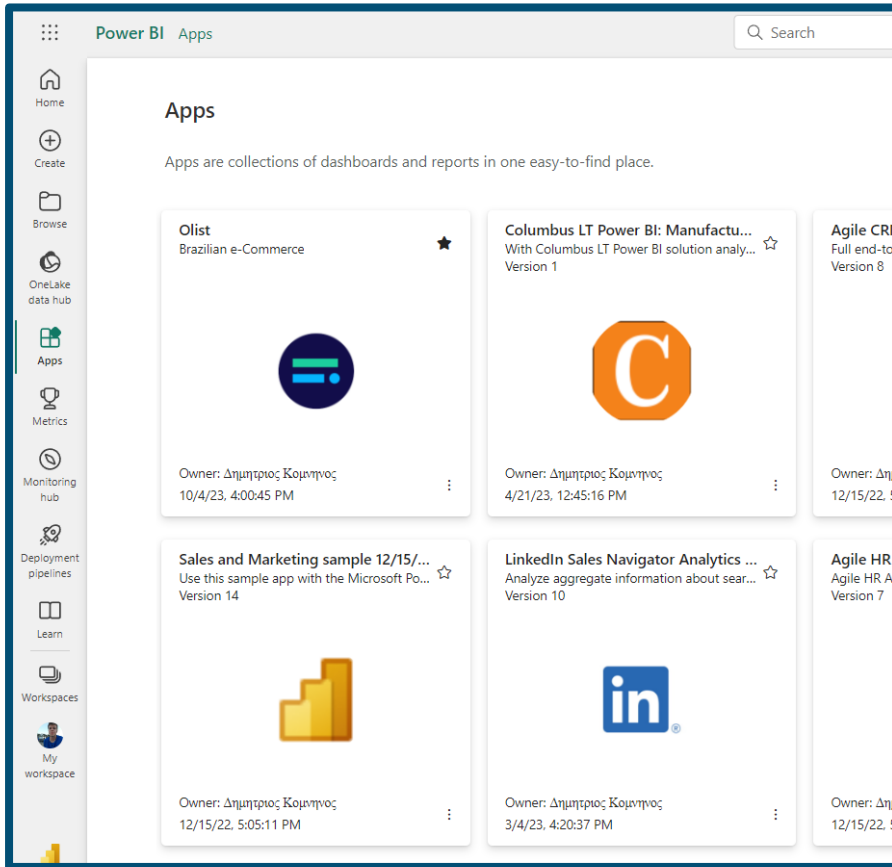
συγκεκριμένα μετρικά. Αυτά είναι το ποσοστό της έγκαιρης παράδοσης και το ποσοστό των ενεργών πωλητών στην πλατφόρμα σε επίπεδο έτους, τετραμήνου και μήνα.



Εικόνα 122 Παρακολούθηση σημαντικών KPIs μέσω των καρτελών Scorecards

Ένα από τα πλεονεκτήματα αυτών των καρτελών είναι ότι μπορούν να στέλνουν **αυτοματοποιημένη ειδοποίηση** μέσω της εφαρμογής **Power BI Mobile** ή μέσω email και να ενημερώνουν τους χρήστες για κάποια αλλαγή των σημαντικών μετρικών για την επιχείρηση, όταν τα συνδεδεμένα δεδομένα στο μοντέλο ανανεωθούν.

Αφού έχουμε δημιουργήσει τις αναφορές, τους πίνακες ελέγχου και την καρτέλα με σημαντικότερα Μετρικά, είναι χρήσιμο όλα αυτά να οργανωθούν σε ένα χώρο πέρα από τον χώρο εργασίας που ήδη βρίσκονται, στην μορφή μιας **εφαρμογής (App)**. Αυτή η λειτουργία του **Power BI Service**, εξυπηρετεί το να κοινοποιούμε σε συνεργάτες, που ενδεχομένως να μην έχουν πρόσβαση στον χώρο εργασίας, το περιεχόμενο τους. Είναι ένας απλός τρόπος να δείξουμε ότι το περιεχόμενο είναι έτοιμο προς παρουσίαση και δεν είναι υπό επεξεργασία. Στο παράδειγμα μας, το τμήμα του μάρκετινγκ έχει ολοκληρώσει την αξιολόγηση των πελατών και επιθυμεί να την κοινοποιήσει στην διοίκηση της εταιρείας. **Η εφαρμογή είναι το πιο κατάλληλο μέσο για να το πράξει αυτό.**



Εικόνα 123 : Η εφαρμογή Olist στο Power BI Service

Στοιχεία σαν και όλα τα παραπάνω συνθέτουν ένα απλό σύστημα επιχειρηματικής ευφυΐας, μέσω του οποίου οι χρήστες της επιχείρησης όπως ο διευθυντής πωλήσεων ,αναλυτές από την διεύθυνση Marketing ή μέλη της διοίκησης, μπορούν να βλέπουν το περιεχόμενο των αναφορών και αν τους επιτρέπεται να έχουν πρόσβαση στο μοντέλο, να μπορούν να δημιουργήσουν καινούργιες αναφορές και εφαρμογές. Τα δεδομένα έτσι παραμένουν ακέραια και η πληροφορία φτάνει στο σωστό ακροατήριο.

5. Συμπεράσματα

5.1 Αποτίμηση εταιρείας

Το ηλεκτρονικό εμπόριο τα τελευταία έτη έχει ανθίσει λόγω και των έκτακτων υγειονομικών καταστάσεων που ζήσαμε όλοι, λόγω της Covid-19. Αυτό αναμφισβήτητα είναι μια έξοχη ευκαιρία της εταιρείας για αύξηση του όγκου των πωλήσεων της. Την ίδια στιγμή λόγω της αύξησης των ηλεκτρονικών αγορών , αφενός όλο και περισσότεροι ανταγωνιστές εμφανίζονται και αφετέρου **το δίκτυο των συνεργατών της εφοδιαστικής αλυσίδας της εταιρείας πρέπει να αναπτύσσεται με τον ίδιο ρυθμό** , ώστε οι πελάτες να λαμβάνουν το ίδιο επίπεδο εξυπηρέτησης.

Οι αξιολογήσεις των πελατών στα προϊόντα που βρίσκονται στην πλατφόρμα της εταιρείας, είναι ένας από τους κύριους μηχανισμούς θετικής ανατροφοδότησης για τις μελλοντικές πωλήσεις της εταιρείας. Γενικώς, δεν υπάρχει κάποια κατηγορία προϊόντων που να εμπορεύεται ,η οποία έχει μεγάλη απόκλιση στην μέση βαθμολογία κριτικής. Όμως, η κατηγορία προϊόντων «**Έπιπλα Σπιτιού και Γραφείου**» που εμπορεύεται η εταιρεία ,η οποία έχει τον μεγαλύτερο όγκο πωλήσεων και παρέχει το μεγαλύτερο ποσοστό των εσόδων της εταιρείας ,είναι κατηγορία με το χαμηλότερο σκορ κριτικών. Δύο είναι οι παράγοντες που επηρεάζουν την βαθμολογία των κριτικών των πελατών. Πρώτον, η κατάσταση του προϊόντος όταν φτάνει στον πελάτη ,αν δηλαδή το προϊόν ανταποκρίνεται στις απαιτήσεις που έχουν δημιουργηθεί στον τελευταίο και αν έχει γίνει **έγκαιρη παράδοση των προϊόντων**. Σύμφωνα, με την ανάλυση που έχει γίνει **το ποσοστό έγκαιρων παραδόσεων των παραγγελιών της εταιρείας** είναι γενικώς αρκετά ικανοποιητικό ,καθώς αγγίζει το **92%**. Στην περίπτωση της κατηγορίας «**Έπιπλα Σπιτιού και Γραφείου**» , εξετάζοντας φαίνεται από τα αποτελέσματα της ανάλυσης ότι η περιοχή του «**Ρίο ντε Ζανέιρο**» σχετίζεται με τις χαμηλότερες βαθμολογίες. Το ποσοστό των έγκαιρων παραγγελιών είναι σχετικά χαμηλότερο από τις υπόλοιπες περιοχές που εξυπηρετεί η εταιρεία.

Από την αξιολόγηση του προφίλ των πελατών , τρία είναι τα σημαντικότερα ευρήματα της ανάλυσης. Πρώτον ,**το ποσοστό των πελατών που έχουν πραγματοποιήσει περισσότερες από μια παραγγελίες** από την πλατφόρμα της εταιρείας είναι εξαιρετικά μικρό , κοντά στο **3%**. Αυτό σημαίνει ότι η εταιρεία δεν έχει εγκαταστήσει ένα σταθερό πελατολόγιο ,όπου θα προσφέρει σταθερά έσοδα. Λόγω της μικρής «πιστότητας των πελατών» (**Customer Loyalty**), είναι σίγουρο ότι η εταιρεία δαπανάει πολύ περισσότερους πόρους στο να αποκτήσει νέους πελάτες, καθώς είναι γνωστό στον κόσμο των επιχειρήσεων ότι η διατήρηση ενός πελάτη κοστίζει υποπολλαπλάσιο προϋπολογισμό , απ' ότι να αποκτηθεί ένας νέος. Δεύτερο κρίσιμο συμπέρασμα που προέκυψε από την ανάλυση ,είναι ότι οι πελάτες προτιμούν προϊόντα με μια **μέτρια μέση τιμή παραγγελίας (AOV)**, κοντά στα **140,0\$ ανά παραγγελία**. Η εταιρεία αξίζει να επενδύσει στην συνεργασία με περισσότερους

εμπόρους που έχουν κατάλογο προϊόντων κοντά σε αυτήν τιμή. Τρίτο σημαντικό πόρισμα είναι το γεγονός ότι, οι περισσότεροι πελάτες που χρησιμοποιούν την πλατφόρμα της εταιρείας, κατοικούν στον αστικό ιστό της Βραζιλίας (Σάο Πάολο, Ρίο ντε Ζανέιρο, Μίνας Γκεράις) ,όπου το κατά κεφαλήν εισόδημα και ο δείκτης βιοτικού επιπέδου είναι μεγαλύτερος συγκριτικά με επαρχιακές περιοχές.

Από την οικονομική ανάλυση των πωλήσεων της εταιρείας, προέκυψαν σημαντικά συμπεράσματα. Την δεδομένη στιγμή ,**οι κατηγορίες προϊόντων** που έχουν την **μεγαλύτερη ζήτηση** είναι οι κατηγορίες «**Έπιπλα Σπιτιού και Γραφείου**» και «**Υγεία και Ομορφιά**». Ενώ, οι πιο ταχέως αναπτυσσόμενες κατηγορίες είναι «**Κατασκευαστικά Εργαλεία**» , «**Φαγητό και Ποτά**» και «**Προμήθειες**». Επίσης, από την αξιολόγηση των νέων συνεργατών που εγγράφονται στην πλατφόρμα της εταιρείας, καταλήγουμε ότι ,ενώ ο αριθμός των νέων συνεργατών βαίνει αυξανόμενος το ποσοστό που παραμένει ενεργό ,δηλαδή όπου συνεχίσει και δέχεται παραγγελίες, βαίνει σταθερά μειούμενο. Συγκριμένα το **ποσοστό των Πωλητών που έχουν δείχτει παραγγελία μέσω της πλατφόρμα** της εταιρείας **τον τελευταίο μήνα** των δεδομένων, **ήταν 41% των συνολικών εγγεγραμμένων.**

Τέλος, από την ανάλυση στους Δυνητικούς Συνεργάτες-Πωλητές της εταιρείας , παρατηρούμε ότι το **ποσοστό κλεισίματος νέων συμφωνιών είναι πάνω από 10%**, κάτι που δηλώνει ότι οι αντιπρόσωποι είναι αρκετά αποδοτικοί στην προσέλκυση νέων Πωλητών, ενώ συνολικά αποτυπώνει το ευοίωνο μέλλον της εταιρείας.

Στα τρωτά σημεία της εταιρείας οπωσδήποτε συγκαταλέγονται η **χαμηλή «συγκράτηση» των πελατών της**, και προτείνεται να αξιοποιηθούν ελκυστικά προγράμματα διατήρησης πελατών, όπως αυτό που εφαρμόζει ήδη, όπου τα κόστη μεταφορικών για μια τιμή παραγγελίας και άνω, καλύπτονται από την εταιρεία και τους πωλητές. Συνολικά, η αξιολόγηση της εταιρείας δείχνει έναν οργανισμό με μεγάλες προοπτικές, αν κεφαλαιοποιήσει τις ευκαιρίες της αγοράς και αξιοποιήσει τους πόρους της προς την σωστή κατεύθυνση, δηλαδή στο κλείσιμο νέων συμφωνιών με πωλητές που εμπορεύονται κατηγορίες προϊόντων με μεγάλο περιθώριο μεικτού κέρδους. **Η περεταίρω ανάπτυξη του δικτύου** πελατών και συνεργατών σε κυρίως **σε αστικές περιοχές** και με μεγάλο δείκτη HDI (Human Development Index) , όπως οι περιοχές του **Parana – PR** και **Bahia – BA** ,όπου γνωρίζουμε ότι αποτελούν πιο εύκολα προσεγγίσιμο κοινό θα προσφέρει περισσότερα και θα επιτρέψει την οικονομική της ανάπτυξη της εταιρείας.

5.2 Αποτίμηση self-service συστημάτων επιχειρηματικής ευφυΐας και το μέλλον τους

Η ιδιότητα των συστημάτων Επιχειρηματικής Ευφυΐας να συγκεντρώνουν την πληθώρα δεδομένων σε ένα οργανωμένο περιβάλλον, με καταχωρημένους ρόλους και περιορισμούς, επιτρέπουν στους οργανισμούς να διαχειρίζονται την πληροφορία που έχουν

με απλό και καταληπτό τρόπο και να φτάνουν σε χρήσιμα συμπεράσματα και τελικά αποφάσεις. Οποίο μοντέλο συστήματος και αν ακολουθεί κάποια εταιρεία που χρησιμοποιεί τέτοια συστήματα καταφέρνει να αξιοποιεί τους πόρους της με βέλτιστο τρόπο. Απαλλάσσει το προσωπικό του IT από την δημιουργία μοντέλων και αναφορών που είναι κοστοβόρες ,αναδιανέμοντας τους προς εργασίες που έχουν την πραγματική ανάγκη τους. Έτσι, δίνει την ευκαιρία σε αναλυτές και άλλες ομάδες ενός οργανισμού να έχουν τον πραγματικό έλεγχο των δεδομένων τους.

Επίσης, μέσα από την self-service πλατφόρμας της Microsoft , Power BI , είδαμε τις ικανότητες τέτοιων συστημάτων τόσο στην οργάνωση όσο και στην οπτικοποίηση των δεδομένων. Κορυφαία χρησιμότητα έχει το **δυναμικό φιλτράρισμα** που γίνεται σε όλα δεδομένα, όταν ο χρήστης στέλνει ad-hoc και απρόβλεπτα ερωτήματα στην βάση, επιτρέποντας την πολυπαραγοντική ανάλυση σε οποιοδήποτε αντικείμενο μελέτης. Οι πανίσχυρες και εντυπωσιακές οπτικοποιήσεις του εργαλείου, του προσθέτουν ακόμα μεγαλύτερη αξία.

Σαφώς, το σημείο το οποίο αλλάζει το παιχνίδι για την χρησιμότητα τέτοιων συστημάτων είναι η ευκολία που οργανώνεται το περιεχόμενο της εταιρείας στο cloud μέρος της πλατφόρμας, το **Power BI Service**. Αν θεσπίζονται σωστοί κανόνες από τους επιτηρητές του συστήματος, οι χρήστες μπορούν να συνεργαστούν με μεγάλη ευκολία, χωρίς να υπάρχει σύγχυση για το ποιο μοντέλο δεδομένων είναι σωστό ,αν τα δεδομένα είναι σωστά ή αν ο χρήστης έχει δικαιοδοσία στο κάνει τροποποιήσεις σε κάποιο μοντέλο. Η πληροφορία φτάνει πάντα στους σωστούς ακροατές και μένει πάντα σύγχρονη και ενημερωμένη.

Είναι σημαντικό να αναφέρουμε ότι τα self-service συστήματα Επιχειρηματικής Ευφυΐας δεν υποκαθιστούν τα κλασικά. Και τα δύο μπορούν να συνεργάζονται σε σημεία μεταξύ τους και να παρέχουν τις ίδιες δυνατότητες στους χρήστες τους. Κάθε κατηγορία έχει τα δικά της πλεονεκτήματα και μειονεκτήματα. Η επιλογή μεταξύ των, γίνεται με βάση την υποδομή την οποία διαθέτει ο οργανισμός, τον απαιτούμενο χρόνο για να στήσει ένα τέτοιο σύστημα και φυσικά τους πόρους που διαθέτει ο οργανισμός.

Αναφορικά με το **μέλλον των συστημάτων μπορούμε να πούμε με σιγουριά ,ότι αυτό είναι στο cloud**. Οι δεσπόζουσες πλατφόρμες επιχειρηματικής ευφυΐας, προωθούν όλο και περισσότερο την ανάπτυξη των συστημάτων διαδικτυακά, εκεί όπου η πληροφορία είναι πάντα πρόσφατη, ενημερωμένη και άμεσα προσβάσιμη από παντού στον κόσμο. Στην Ελλάδα, τα συστήματα Επιχειρηματικής Ευφυΐας κερδίζουν διαρκώς έδαφος, χωρίς ακόμα η πλειονότητα των επιχειρηματιών να είναι οικεία με την αξία τους. Ακόμα οι περισσότερες μικρομεσαίες επιχειρήσεις στην χώρα μας ,οι οποίες είναι και η πλειονότητα των επιχειρήσεων, στηρίζονται σε εμπειρικά δεδομένα.Όπως έγινε γνωστό, το μέγεθος μιας επιχείρησης συναρτάται άμεσα με την πιθανότητα να συλλέγει και να διαχειρίζεται σωστά τα δεδομένα της.

Ολοκληρώνοντας, η βελτίωση λειτουργίας ενός οργανισμού ή μιας επιχείρησης και η επίλυση των προβλημάτων τους, όπως και στην ζωή, ξεκινάει πάντα με την αναγνώριση

αυτών των προβλημάτων. Αυτό ακριβώς είναι και το σημείο το οποίο γεννάει την ανάγκη ύπαρξης των Συστημάτων Επιχειρηματικής Ευφυΐας.

6. Ευχαριστίες

Θα ήθελα να ευχαριστώ συνολικά όσους με βοήθησαν στο να δημιουργήσω αυτήν την διπλωματική εργασία. Αρχικά, θα ήθελα να ευχαριστήσω τον καθηγητή μου για την εμπιστοσύνη που μου έδειξε και για την καθοδήγηση του στο επίτιμο αλλά και συναρπαστικό έργο της συγγραφής του πρώτου ολοκληρωμένου «επαγγελματικού τέκνου» μου. Επιπλέον, θα ήθελα να ευχαριστήσω την οικογένεια μου για την στήριξη της όλο αυτόν καιρό. Τέλος, η ομάδα κοινότητας του Ψυχοδράματος και η ομάδα Πλοήγησης της εταιρείας HAGAP (*Hellenic Association of Group Analysis and Psychotherapy*) και η Αναλυτική μου ομάδα υπήρξαν τα στέρεα θεμέλια στα οποία στήριξα τα πόδια μου, για να ολοκληρώσω με επιτυχία την ακαδημαϊκή πορεία μου στο Πολυτεχνείο. Τους ευχαριστώ!

Βιβλιογραφία

- [1] **Ahmad, A., 2015.** Business intelligence for sustainable competitive advantage. In *Sustaining competitive advantage via business intelligence, knowledge management, and system dynamics*. Emerald Group Publishing Limited.
- [2] **Marren, P., 2004.** *The father of business intelligence*. The Journal of Business Strategy, 25(6), 5–7.
- [3] **Liautaud, B., & Hammond, M., 2001.** E-business intelligence: Turning information into knowledge into profit. New York, NY: McGraw-Hill.
- [4] **Sullivan, P. H., 2000.** Profiting from intellectual capital: Extracting value from innovation. New York, NY: Wiley.
- [5] **Adamson, C., 2010.** *Star schema the complete reference*. McGraw Hill Professional.
- [6] **Rupal Bhandari, 2020.** ‘Traditional vs. Self-Service BI: Analytics Alternatives Explained’ June 12, 2020. Διαθέσιμο εδώ: <https://www.softwareadvice.com/resources/traditional-bi-vs-self-service/> (Προσπελάστηκε στις 07/12/2022).
- [7] **Howson, C. & Duncan, A. D., 2015.** ‘ITScore Overview for BI and Analytics’, Gartner Research, 24 September 2015. Διαθέσιμο εδώ: <https://www.gartner.com/en/documents/3136418> (Προσπελάστηκε στις 07/12/2022).
- [8] **Petrossian ,G. ,2021** ‘Business Intelligence Tools: A Pros and Cons Comparison Chart’ Διαθέσιμο εδώ : <https://skypointcloud.com/blog/business-intelligence-tools-comparison-chart/> (Προσπελάστηκε στις 10/12/2022).
- [9] **Microsoft Power BI – Documentation** <https://learn.microsoft.com/en-us/power-bi/fundamentals/power-bi-overview> (Προσπελάστηκε στις 10/12/2022).
- [10] **Russo, M. & Ferrari, A. ,2017.** Technical Whitepaper: ‘DirectQuery in SQL Server 2016 Analysis Services.
- [11] **Mone, N. 2021.** ‘VertiPaaS Engine Power BI: An Introduction’ Διαθέσιμο εδώ: <https://blog.enterprisedna.co/vertipaas-engine-power-bi-an-introduction/> (Προσπελάστηκε στις 18/12/2022)

[12] **Microsoft Power BI – Documentation**, Διαθέσιμο εδώ : <https://learn.microsoft.com/en-us/power-bi/transform-model/desktop-composite-models> (Προσπελάστηκε στις 10/12/2022).

[13] **Russo, M. 2021.** ‘Relationships in Power BI and Tabular models’ Διαθέσιμο εδώ : <https://www.sqlbi.com/articles/relationships-in-power-bi-and-tabular-models/> (Προσπελάστηκε στις 11/12/2022).

[14] **Microsoft Power BI – Documentation** : <https://learn.microsoft.com/en-us/power-bi/connect-data/desktop-quickstart-connect-to-data> (Προσπελάστηκε στις 15/12/2022).

[15] **Hay, D.C., 2013.** Data model patterns: conventions of thought. Pearson Education.

[16] **Microsoft Power BI – Documentation** : <https://learn.microsoft.com/en-us/analysis-services/tom/introduction-to-the-tabular-object-model-tom-in-analysis-services-amo?view=asallproducts-allversions> (Προσπελάστηκε στις 15/12/2022)

[17] **Ross, M. and Kimball, R., 2013.** The data warehouse toolkit: *the definitive guide to dimensional modeling*. John Wiley & Sons.

[18] **Microsoft Power BI – Documentation** , Διαθέσιμο εδώ : <https://learn.microsoft.com/el-gr/power-bi/guidance/powerbi-implementation-planning-usage-scenario-managed-self-service-bi> (Προσπελάστηκε στις 10/09/2023)

[19] **Microsoft Power BI – Documentation** , Διαθέσιμο εδώ : <https://learn.microsoft.com/el-gr/power-bi/guidance/powerbi-implementation-planning-usage-scenario-enterprise-bi> (Προσπελάστηκε στις 10/09/2023)

[20] **Anonymous, UNCTAD** ‘How COVID-19 triggered the digital and e-commerce turning point’ Διαθέσιμο εδώ : <https://unctad.org/news/how-covid-19-triggered-digital-and-e-commerce-turning-point> (Προσπελάστηκε στις : 13/09/23)

[21] **Sionek, A. and Olist, 2018** *Brazilian E-Commerce Public Dataset by Olist* Διαθέσιμο εδώ: <https://www.kaggle.com/datasets/olistbr/brazilian-ecommerce> ,(Προσπελάστηκε στις 24/03/2023)

[22] **Sionek, A. and Olist, 2018** *Marketing Funnel by Olist* Διαθέσιμο εδώ: [Marketing Funnel Olist](#), (Προσπελάστηκε στις 24/03/2023)

Παράρτημα

Οι συντομογραφίες των πολιτειών που εξυπηρετεί η εταιρεία.

Συντομογραφία Πολιτείας	Πλήρες Όνομα
AC	ΑΚΡΕ
AL	ΑΛΑΜΠΑΜΑ
AM	ΑΜΑΖΟΝΑΣ
AP	ΑΜΑΠΑ
BA	ΜΠΑΧΙΑ
CE	ΣΕΑΡΑ
DF	ΟΜΟΣΠΟΝΔΙΑΚΗ ΠΕΡΙΟΧΗ (ΜΠΡΑΖΙΛΙΑ)
ES	ΕΠΣΙΡΙΤΟ ΣΑΝΤΟ
GO	ΓΚΟΙΑΣ
MA	ΜΑΣΑΧΟΥΣΕΤΗ
MG	ΜΙΝΑΣ ΓΚΕΡΑΙΣ
MS	ΜΑΤΟ ΓΚΡΟΣΟ ΝΤΟ ΣΟΥΛ
MT	ΜΟΝΤΑΝΑ
PA	ΠΕΝΣΥΛΒΑΝΙΑ
PB	ΠΑΡΑΙΜΠΑ
PE	ΠΕΡΝΑΜΠΟΥΚΟ
PI	ΠΙΑΟΥΙ
PR	ΠΑΡΑΝΑ
RJ	ΡΙΟ ΝΤΕ ΖΑΝΕΙΡΟ
RN	ΡΙΟ ΓΚΡΑΝΤΕ ΝΤΟ ΝΟΡΤΕ
RO	ΡΟΝΔΟΝΙΑ
RR	ΡΟΡΑΙΜΑ
RS	ΡΙΟ ΓΚΡΑΝΤΕ ΝΤΟ ΣΟΥΛ
SC	ΣΑΝΤΑ ΚΑΤΑΡΙΝΑ
SE	ΣΕΡΓΚΙΠΕ
SP	ΣΑΟ ΠΑΟΛΟ
TO	ΤΟΚΑΝΤΙΣ

Συντομογραφίες

3NF: 3rd Normal Form
AS: Analytical System
BI: Business Intelligence
BICC: Business Intelligence Competence Centre
CAO: Chief Analytics Officer
CDO: Chief Data Officer
CEO: Chief Executive Officer
CFO: Chief Financial Officer
COO: Chief Operating Officer
CRM: Customer Relationship Management
DSS: Decision Support System
DW: Data Warehouse
EIM: Enterprise Information Management
EIS: Executive Information System
ER: Entity Relationship
ERP: Enterprise Resource Planning
ETL : Extract Transform Load
IS: Information System
KPIs: Key Performance Indicators
MDB: Multi-Dimensional Database
MIS: Manager Information System
OS: Operational System
RDB: Relational Database
SaaS: Software as a Service