



ΤΕΧΝΟΛΟΓΙΚΟ ΕΚΠΑΙΔΕΥΤΙΚΟ
ΙΔΡΥΜΑ ΚΑΛΑΜΑΤΑΣ
ΣΧΟΛΗ ΔΙΟΙΚΗΣΗΣ ΟΙΚΟΝΟΜΙΑΣ
ΤΜΗΜΑ ΔΙΟΙΚΗΣΗΣ ΜΟΝΑΔΩΝ ΥΓΕΙΑΣ
ΠΡΟΝΟΙΑΣ

ΠΤΥΧΙΑΚΗ ΕΡΓΑΣΙΑ

«ΕΓΚΑΤΑΣΤΑΣΗ ΚΑΙ ΜΕΛΕΤΗ ΕΡΓΑΛΕΙΩΝ
ΔΙΑΧΕΙΡΙΣΗΣ ΔΕΔΟΜΕΝΩΝ OLAP»

ΓΚΑΡΝΑΡΑ ΜΑΡΙΝΑ
ΚΑΚΙΑΣ ΓΡΗΓΟΡΗΣ
ΚΟΜΜΑΤΑΣ ΣΤΑΥΡΟΣ

ΕΠΙΒΛΕΠΩΝ : Δρ. ΚΟΤΣΙΛΙΕΡΗΣ ΘΕΟΔΩΡΟΣ

Καλαμάτα 2008

ΠΕΡΙΛΗΨΗ

Σκοπός της παρούσας πτυχιακής εργασίας είναι η μελέτη εργαλείων OLAP καθώς και η διαπίστωση αν αυτά τα εργαλεία μπορούν να διαχειριστούν επιχειρηματικές βάσεις δεδομένων μεγάλου όγκου.

Μέσω του διαδικτύου μελετήθηκαν διάφορα εργαλεία OLAP, ανοιχτού κώδικα και μη, από τα οποία επιλέχτηκε το εργαλείο Sisense Prism Desktop, πληρώντας όλα τα κριτήρια επιλογής εργαλείου όπως π.χ. να παρέχεται δωρεάν, να είναι εύχρηστο και να είναι ολοκληρωμένο. Επίσης, στην εργασία παρουσιάζεται η διαδικασία εγκατάστασης του εργαλείου καθώς και η διαδικασία εφαρμογής του με κάποια παραδείγματα για την καλύτερη κατανόησή του.

Από την εφαρμογή και την μελέτη του διαπιστώθηκε ότι είναι εύχρηστο, δεν χρειάζεται ειδικές γνώσεις στον προγραμματισμό για να θέσει κάποιος ερωτήματα καθώς και ότι μπορεί να διαχειριστεί μεγάλες επιχειρηματικές βάσεις δεδομένων και μη, χωρίς να χρειάζεται μεγάλη υπολογιστική ισχύ.

Τέλος, η εργασία απέδειξε ότι η OLAP τεχνολογία είναι ο ποιο σύγχρονος και αποδοτικός τρόπος για την επεξεργασία μεγάλου όγκου δεδομένων. Παρόλα αυτά, στην Ελλάδα βρίσκεται ακόμη σε πρώιμο στάδιο.

ΕΥΧΑΡΙΣΤΙΕΣ

Σε αυτό το σημείο, θα θέλαμε να ευχαριστήσουμε τον καθηγητή μας κ. Κοτσιλιέρη Θεόδωρο για την πολύτιμη βοήθεια του στην εκπόνηση της πτυχιακής μας εργασίας και τους γονείς μας για όλη την υποστήριξη που μας παρείχαν κατά την διάρκεια των σπουδών μας.

ΠΕΡΙΛΗΨΗ.....	i
ΕΥΧΑΡΙΣΤΙΕΣ.....	ii
ΠΕΡΙΕΧΟΜΕΝΑ.....	iii
ΠΕΡΙΕΧΟΜΕΝΑ ΠΙΝΑΚΩΝ.....	iv
ΠΕΡΙΕΧΟΜΕΝΑ ΣΧΗΜΑΤΩΝ.....	v
ΣΥΝΤΟΜΟΓΡΑΦΙΕΣ.....	viii
ΕΙΣΑΓΩΓΗ.....	1
ΚΕΦΑΛΑΙΟ 1: ΑΜΕΣΗ ΑΝΑΛΥΤΙΚΗ ΕΠΕΞΕΡΓΑΣΙΑ ΤΩΝ ΔΕΔΟΜΕΝΩΝ ON-LINE ANALYTICAL PROCESSING (OLAP).....	3
1.1 Γενικά.....	3
1.2 Κατηγορίες OLAP.....	6
1.3 Χαρακτηριστικά OLAP.....	8
1.4 Εφαρμογές OLAP.....	9
1.5 Γενική Αρχιτεκτονική Συστημάτων Διαχείρισης Βάσεων Δεδομένων για OLAP.....	11
1.6 Βασικές Λειτουργίες του Πολυδιάστατου Μοντέλου για OLAP.....	17
ΚΕΦΑΛΑΙΟ 2: ΑΠΟΘΗΚΕΣ ΔΕΔΟΜΕΝΩΝ.....	21
2.1 Γενικά.....	21
2.2 Κατηγορίες.....	22
2.3 Χαρακτηριστικά.....	23
2.4 Δομή και λειτουργία.....	24
2.5 Ροή και Μετασχηματισμός Δεδομένων.....	27
2.6 Σχεδίαση και Εφαρμογές Αποθηκών Δεδομένων	29
2.7 Συνεργασία Αποθηκών Δεδομένων και OLAP.....	33
ΚΕΦΑΛΑΙΟ 3: ΑΝΟΙΧΤΟΣ ΚΩΔΙΚΑΣ (OPEN SOURCE).....	38
3.1 Η έννοια του ανοιχτού κώδικα.....	38
3.2 Πλεονεκτήματα και Μειονεκτήματα	39
3.3 Χρήση Ανοιχτού Κώδικα και Ελεύθερου Λογισμικού.....	40
3.4 Διαφορά Ανοιχτού Κώδικα και Ελεύθερου Λογισμικού.....	42
ΚΕΦΑΛΑΙΟ 4: ΕΡΓΑΛΕΙΑ OLAP.....	44
4.1 Λειτουργία εργαλείων OLAP.....	44
4.2 Ανάλυση εργαλείων.....	45
ΚΕΦΑΛΑΙΟ 5: ΑΝΑΛΥΣΗ ΕΡΓΑΛΕΙΟΥ SISENCE PRISM DESKTOP ΚΑΙ ΔΙΑΔΙΚΑΣΙΑ ΕΓΚΑΤΑΣΤΑΣΗΣ ΤΟΥ.....	54
5.1 Γενικά.....	54
5.2 Ορισμοί	55
5.3 Εργασιακό περιβάλλον του Prism	61
5.4 Σύνδεση με μια πηγή δεδομένων Excel.....	65
5.5 Λειτουργία Άξονα (Pivot) και Δεικτών	68
5.6 Συνεργασία αριθμητικών τύπων με το Prism	73
5.7 Φίλτρα	76
5.8 Διαδικασία εγκατάστασης.....	93
ΚΕΦΑΛΑΙΟ 6: ΕΦΑΡΜΟΓΗ ΕΡΓΑΛΕΙΟΥ PRISM.....	94
6.1 Επεξεργασία δεδομένων εταιρειών κινητής τηλεφωνίας	94
6.2 Επεξεργασία δεδομένων εταιριών σταθερής τηλεφωνίας.....	101
6.3 Επεξεργασία δεδομένων καταστημάτων πληροφορικής	110
ΣΥΜΠΕΡΑΣΜΑΤΑ.....	121
ΒΙΒΛΙΟΓΡΑΦΙΑ.....	122
ΠΑΡΑΡΤΗΜΑ.....	124

ΠΕΡΙΕΧΟΜΕΝΑ ΠΙΝΑΚΩΝ

Πίνακας Εργαλείων OLAP.....	45
Πίνακας 5-1: Πολυδιάστατος πίνακας.....	55
Πίνακας 5-2: Κατηγορίες Συναθροίσεων.....	59
Πίνακας 5-3: Θέσεις Widgets.....	63

ΠΕΡΙΕΧΟΜΕΝΑ ΣΧΗΜΑΤΩΝ

Σχήμα 1-1: Ο ρόλος ενός OLAP Server.....	5
Σχήμα 1-2: Σχήμα αστέρα.....	8
Σχήμα 1-3: Τα στοιχεία μιας OLAP εφαρμογής.....	9
Σχήμα 1-4: Εφαρμογή OLAP τριών επιπέδων.....	10
Σχήμα 1-5: Αρχιτεκτονική ενός ΣΔΒΔ που χρησιμοποιείται για OLAP.....	12
Σχήμα 1-6: Παράδειγμα Πολυδιάστατων Δεδομένων.....	14
Σχήμα 1-7: Παραδείγματα Ιεραρχιών Διαστάσεων	16
Σχήμα 1-8: Παράδειγμα λειτουργίας pivoting.....	18
Σχήμα 1-9: Παράδειγμα λειτουργίας roll-up.....	19
Σχήμα 1-10: Παράδειγμα λειτουργίας περιστροφής(rotation).....	20
Σχήμα 2-1: Γενική αρχιτεκτονική ΣΔΒΔ.....	21
Σχήμα 2-2: Η δομή μιας αποθήκης δεδομένων.....	24
Σχήμα 2-3: Οι λειτουργίες στις αποθήκες δεδομένων.....	27
Σχήμα 2-4: Η ροή των δεδομένων μέσα σε μια αποθήκη δεδομένων.....	28
Σχήμα 2-5: Τα δεδομένα και η χρήση τους.....	29
Σχήμα 2-6: Το σύνολο των εφαρμογών μιας αποθήκης δεδομένων.....	32
Σχήμα 4-1: Αρχιτεκτονική Mondrian.....	48
Σχήμα 4-2: Απεικόνιση Εργαλείου Palo	48
Σχήμα 4-3: Cell-Based Palo	49
Σχήμα 4-4: Multidimensional.....	50
Σχήμα 4-5: Hierarchies.....	50
Σχήμα 4-6: OLAP X Application.....	52
Σχήμα 4-7: OLAP Cube.....	53
Σχήμα 5-1: Δέντρο διαστάσεων πολυδιάστατου πίνακα.....	56
Σχήμα 5-2: Date (Calendar).....	57
Σχήμα 5-3: Date (Months).....	57
Σχήμα 5-4: Date (Week Days).....	57
Σχήμα 5-5: Γονικά και Θυγατρικά Μέλη.....	58
Σχήμα 5-6: Drill-Down και Drill-Up.....	58
Σχήμα 5-7: Συμβολισμός Μέτρων.....	59
Σχήμα 5-8: Αποθήκη Φίλτρων.....	60
Σχήμα 5-9: Περιβάλλον Prism.....	61
Σχήμα 5-10: Sheets.....	62
Σχήμα 5-11: Δημιουργία Formula.....	62
Σχήμα 5-12: Επεξεργασία Formula.....	62
Σχήμα 5-13: Δημιουργία νέας Formulas.....	62
Σχήμα 5-14: Γραμμή Εργαλείων.....	63
Σχήμα 5-15: Widget's Data Layout (Pivot-Chart).....	63
Σχήμα 5-16: Widget's Data Layout (Indicator-Calendar)	64
Σχήμα 5-17: Widgets.....	65
Σχήμα 5-18: Create New Data Source.....	65
Σχήμα 5-19: Σύνδεση με Excel	66
Σχήμα 5-20: Εισαγωγή Δεδομένων.....	66

Σχήμα 5-21: Μετονομασία Πηγής Δεδομένων.....	67
Σχήμα 5-22: Επιλογή Διαστάσεων.....	67
Σχήμα 5-23: Δέντρο Διαστάσεων.....	68
Σχήμα 5-24: Εισαγωγή νέου Άξονα.....	69
Σχήμα 5-25: Εισαγωγή δεδομένων σε Widget.....	70
Σχήμα 5-26: Μορφοποίηση Άξονα.....	70
Σχήμα 5-27: Span Selection Feature.....	71
Σχήμα 5-28: Μορφοποίηση χρώματος.....	71
Σχήμα 5-29: Εισαγωγή Χρόνου.....	72
Σχήμα 5-30: Ταξινόμηση Μελών.....	73
Σχήμα 5-31: Κέρδος ανά Μήνα.....	74
Σχήμα 5-32: Κέρδος ανά Πόλη.....	74
Σχήμα 5-33: Κέρδος ανά Μήνα και Πόλη.....	75
Σχήμα 5-34: Κέρδος ανά Μήνα.....	76
Σχήμα 5-35: Κέρδος ανά Πόλη.....	76
Σχήμα 5-36: Θέση φίλτρου σε διάσταση Customers.....	77
Σχήμα 5-37: Δημιουργία νέου Φίλτρου.....	77
Σχήμα 5-38: Είδη Φίλτρων.....	78
Σχήμα 5-39: Top Ranking.....	78
Σχήμα 5-40: Πλαίσιο By.....	79
Σχήμα 5-41: Περιορισμός Φίλτρου.....	79
Σχήμα 5-42: Νέο Φίλτρο.....	80
Σχήμα 5-43: Δημιουργία νέου Άξονα.....	80
Σχήμα 5-44: Εισαγωγή Φίλτρου.....	80
Σχήμα 5-45: Αποτελέσματα Φίλτρου.....	81
Σχήμα 5-46: Εισαγωγή διάστασης Χρόνου.....	81
Σχήμα 5-47: Αποτελέσματα.....	82
Σχήμα 5-48: Φιλτράρισμα με Μέτρο.....	83
Σχήμα 5-49: Επιλογή Μέτρου / Τύπου.....	83
Σχήμα 5-50: Πεδίο εφαρμογής / τύπου.....	83
Σχήμα 5-51: Κριτήριο Box.....	84
Σχήμα 5-52: Δεύτερος τελεστής.....	84
Σχήμα 5-53: Πρόσθεση νέου κριτηρίου.....	84
Σχήμα 5-54: Αλλαγή δεύτερου τελεστή.....	85
Σχήμα 5-55: User Interface.....	86
Σχήμα 5-56: Εισαγωγή κριτηρίου.....	86
Σχήμα 5-57: Label.....	86
Σχήμα 5-58: Πρόσθεση νέου κριτηρίου.....	87
Σχήμα 5-59: Αποτελέσματα κριτηρίων.....	87
Σχήμα 5-60: User Interface.....	88
Σχήμα 5-61: Κριτήριο.....	88
Σχήμα 5-62: Box Value.....	89
Σχήμα 5-63: Φιλτράρισμα με βάση την αξία.....	89
Σχήμα 5-64: PartNumber.....	90
Σχήμα 5-65: Define a Scope.....	90
Σχήμα 5-66: Επιλογή Μηνών.....	91
Σχήμα 5-67: Εφαρμογή διαδικασίας.....	92
Σχήμα 5-68: Τελικό Φίλτρο.....	92

Σχήμα 6-1: Εισαγωγή δεδομένων εταιρειών κινητής τηλεφωνίας.....	94
Σχήμα 6-2: Ονομασία πηγής δεδομένων εταιρειών κινητής τηλεφωνίας.....	95
Σχήμα 6-3: Επιλογές διαστάσεων εταιρειών κινητής τηλεφωνίας.....	95
Σχήμα 6-4: Δέντρο διαστάσεων εταιρειών κινητής τηλεφωνίας.....	96
Σχήμα 6-5: Widgets Data Layout.....	97
Σχήμα 6-6: Αποτελέσματα πρώτου παραδείγματος εταιρειών κινητής τηλεφωνίας.....	98
Σχήμα 6-7: Δημιουργία φίλτρου με μέτρο.....	99
Σχήμα 6-8: Εμφάνιση φίλτρου με μέτρο.....	99
Σχήμα 6-9: Αποτέλεσμα δεύτερου παραδείγματος εταιρειών κινητής τηλεφωνίας με χρήση φίλτρου.....	100
Σχήμα 6-10: Αποτέλεσμα τρίτου παραδείγματος εταιρειών κινητής τηλεφωνίας.....	101
Σχήμα 6-11: Εισαγωγή δεδομένων εταιρειών σταθερής τηλεφωνίας.....	102
Σχήμα 6-12: Ονομασία πηγής δεδομένων εταιρειών σταθερής τηλεφωνίας.....	102
Σχήμα 6-13: Επιλογές διαστάσεων εταιρειών σταθερής τηλεφωνίας.....	103
Σχήμα 6-14: Δέντρο διαστάσεων εταιρειών σταθερής τηλεφωνίας.....	104
Σχήμα 6-15: Γραμμή εργαλείων.....	104
Σχήμα 6-16: Widgets Data Layout.....	105
Σχήμα 6-17: Αποτέλεσμα πρώτου παραδείγματος εταιρειών σταθερής τηλεφωνίας.....	106
Σχήμα 6-18: Επιλογή φίλτρου.....	107
Σχήμα 6-19: Φίλτρο με αξία.....	107
Σχήμα 6-20: Εμφάνιση φίλτρου με αξία.....	108
Σχήμα 6-21: Αποτέλεσμα δεύτερου παραδείγματος εταιρειών σταθερής τηλεφωνίας.....	109
Σχήμα 6-22: Αποτέλεσμα τρίτου παραδείγματος εταιρειών σταθερής τηλεφωνίας.....	110
Σχήμα 6-23: Εισαγωγή δεδομένων καταστημάτων πληροφορικής.....	111
Σχήμα 6-24: Ονομασία πηγής δεδομένων καταστημάτων πληροφορικής.....	111
Σχήμα 6-25: Επιλογές διαστάσεων καταστημάτων πληροφορικής.....	112
Σχήμα 6-26: Δέντρο διαστάσεων καταστημάτων πληροφορικής.....	113
Σχήμα 6-27: Γραμμή εργαλείων.....	113
Σχήμα 6-28: Widgets Data Layout.....	114
Σχήμα 6-29: Αποτέλεσμα πρώτου παραδείγματος καταστημάτων πληροφορικής.....	115
Σχήμα 6-30: Αποτέλεσμα δεύτερου παραδείγματος καταστημάτων πληροφορικής.....	116
Σχήμα 6-31: Αποτέλεσμα τρίτου παραδείγματος καταστημάτων πληροφορικής.....	117
Σχήμα 6-32: Πλαίσιο εισαγωγής τύπου.....	118
Σχήμα 6-33: Ονομασία τύπου.....	118
Σχήμα 6-34: Εμφάνιση τύπου.....	119
Σχήμα 6-35: Αποτέλεσμα τρίτου παραδείγματος καταστημάτων πληροφορικής με εφαρμογή τύπου.....	120

ΣΥΝΤΟΜΟΓΡΑΦΙΕΣ

BI	BUSINESS INTELLIGENCE
DBMS	DATA BASE MANAGEMENT SYSTEM
LAN	LOCAL AREA NETWORK
MDBMS	MULTIDIMENSIONAL DATA BASE MANAGEMENT SYSTEM
MDX	MULTIDIMENSIONAL EXPRESSIONS
MOLAP	MULTIDIMENSIONAL ON-LINE ANALYTICAL PROCESSING
OLAP	ON-LINE ANALYTICAL PROCESSING
OLTP	ON-LINE TRANSACTION PROCESSING
RDBMS	RELATIONAL DATA BASE MANAGEMENT SYSTEM
ROLAP	RELATIONAL ON-LINE ANALYTICAL PROCESSING
SQL	STRUCTURED QUERY LANGUAGE
ΒΔ	ΒΑΣΗ ΔΕΔΟΜΕΝΩΝ
ΕΛ	ΕΛΕΥΘΕΡΟ ΛΟΓΙΣΜΙΚΟ
ΛΑΚ	ΛΟΓΙΣΜΙΚΟ ΑΝΟΙΧΤΟΥ ΚΩΔΙΚΑ
ΣΔΒΔ	ΣΥΣΤΗΜΑΤΑ ΔΙΑΧΕΙΡΙΣΗΣ ΒΑΣΕΩΝ ΔΕΔΟΜΕΝΩΝ
ΣΥΑ	ΣΥΣΤΗΜΑΤΑ ΥΠΟΣΤΗΡΙΞΗΣ ΑΠΟΦΑΣΕΩΝ

ΕΙΣΑΓΩΓΗ

Η παρούσα μελέτη εκπονήθηκε στα πλαίσια των πτυχιακών εργασιών του Τμήματος Διοίκηση Μονάδων Υγείας και Πρόνοιας, της σχολής Διοίκησης και Οικονομίας του Τ.Ε.Ι Καλαμάτας.

Αντικείμενο της εργασίας είναι η εγκατάσταση και η μελέτη εργαλείων διαχείρισης δεδομένων OLAP.

Η υπόθεση εργασίας είναι ότι η τεχνολογία OLAP βοηθάει στην ταχεία πρόσβαση και επεξεργασία μεγάλων επιχειρηματικών βάσεων δεδομένων με την πολυδιάστατη ανάλυσή τους.

Σκοπός της εργασίας είναι η μελέτη εργαλείων OLAP καθώς και η διαδικασία εφαρμογής τους με κάποια παραδείγματα για την καλύτερη κατανόησή τους.

Για την προσέγγιση του σκοπού της εργασίας τέθηκαν δύο στόχοι. Ο πρώτος στόχος ήταν η εύρεση και η εγκατάσταση ενός εργαλείου OLAP ανοιχτού κώδικα και η περαιτέρω μελέτη και ανάλυσή του. Ο δεύτερος στόχος ήταν η εφαρμογή του εργαλείου σε διάφορες βάσεις δεδομένων για την καλύτερη κατανόησή του.

Για την επίτευξη του πρώτου στόχου μελετήσαμε διάφορων ειδών εργαλεία μέσω του διαδικτύου, ανοιχτού κώδικα και μη, με αποτέλεσμα να καταλήξουμε στο εργαλείο Sisense Prism Desktop. Για την επίτευξη του δεύτερου στόχου δημιουργήσαμε, με ίδια έρευνα, διάφορες βάσεις δεδομένων με στοιχεία που βρήκαμε στο διαδίκτυο.

Η εργασία, εκτός της παρούσας Εισαγωγής, είναι δομημένη σε έξι κεφάλαια και ένα συμπερασματικό τμήμα. Στο Πρώτο Κεφάλαιο αναλύεται η Άμεση Αναλυτική Επεξεργασία Δεδομένων (OLAP). Ο όρος OLAP επινοήθηκε το 1993 από τον Ted Codd, ο οποίος έχει χαρακτηριστεί ως “ο πατέρας της σχεσιακής βάσης δεδομένων”.

Στο Δεύτερο Κεφάλαιο παρουσιάζονται οι Αποθήκες Δεδομένων, οι οποίες είναι ειδικές βάσεις δεδομένων που περιέχουν κατάλληλα επεξεργασμένα δεδομένα. Ως βάση δεδομένων εννοείται η οργάνωση και καταχώρηση της πληροφορίας, ώστε να ενημερώνεται και να ανακαλείται με ευέλικτο τρόπο.

Το Τρίτο Κεφάλαιο αναφέρεται στη χρήση του Ανοιχτού Κώδικα καθώς και στην διαφοροποίησή του από το ελεύθερο λογισμικό.

Στο Τέταρτο Κεφάλαιο παρουσιάζονται ορισμένα εργαλεία, ανοιχτού κώδικα και μη, καθώς και τα κριτήρια επιλογής για την εύρεση ενός εργαλείου OLAP.

Στο Πέμπτο Κεφάλαιο αναλύεται το εργαλείο Sisense Prism Desktop καθώς και η διαδικασία εγκατάστασής του.

Τέλος, στο Έκτο Κεφάλαιο παρουσιάζονται κάποια παραδείγματα εφαρμογής του εργαλείου Sisense Prism Desktop, προς διευκόλυνση του αναγνώστη.

Ως βασική βιβλιογραφία της εργασίας χρησιμοποιήθηκαν μελέτες και έρευνες ελληνικής και ξένης βιβλιογραφίας αλλά και πηγές από το διαδίκτυο.

ΚΕΦΑΛΑΙΟ 1

ΑΜΕΣΗ ΑΝΑΛΥΤΙΚΗ ΕΠΕΞΕΡΓΑΣΙΑ ΤΩΝ ΔΕΔΟΜΕΝΩΝ ON-LINE ANALYTICAL PROCESSING (OLAP)

1.1 Γενικά

Η άμεση αναλυτική επεξεργασία των δεδομένων (OLAP) είναι μια τεχνολογία που βοηθά στην ταχεία πρόσβαση και πολυδιάστατη επεξεργασία μεγάλου όγκου δεδομένων με σκοπό την παρουσίαση και τη λύση των προβλημάτων της επιχείρησης στις πραγματικές τους διαστάσεις και την υποστήριξη της επιχειρηματικής ευφυΐας. Οι βάσεις δεδομένων OLAP χωρίζονται σε έναν ή περισσότερους κύβους και κάθε κύβος είναι οργανωμένος και σχεδιασμένος από ένα διαχειριστή κύβου.

Ο κύβος είναι μια ομάδα από κελιά δεδομένων (data cells). Κάθε κελί χαρακτηρίζεται μονοσήμαντα από τις αντίστοιχες τιμές των διαστάσεων του κύβου. Τα περιεχόμενα του κελιού ονομάζονται μέτρα (measures).

Η διάσταση είναι ένα σύνολο μιας ή περισσοτέρων οργανωμένων ιεραρχικών επιπέδων σε έναν κύβο, της οποίας ένας χρήστης κατανοεί και χρησιμοποιεί ως βάση για την ανάλυση δεδομένων. Για παράδειγμα, μια γεωγραφική διάσταση ενδέχεται να περιλαμβάνει επίπεδα για τη Χώρα/Περιοχή, το Νομό/Επαρχία και την Πόλη.

Η ιεραρχία είναι μια λογική δομή δέντρου, η οποία οργανώνει τα μέλη μιας διάστασης κατά τέτοιον τρόπο, ώστε κάθε μέλος να έχει ένα γονικό μέλος και ένα ή περισσότερα θυγατρικά μέλη. Το θυγατρικό μέλος είναι το επόμενο χαμηλότερο επίπεδο σε μια ιεραρχία, το οποίο συνδέεται απευθείας με το τρέχων μέλος. Για παράδειγμα, σε μια χρονική ιεραρχία που περιέχει τα επίπεδα “Τρίμηνο”, “Μήνας” και “Ημέρα”, το στοιχείο “Ιανουάριος” είναι θυγατρικό στοιχείο του Πρώτου Τριμήνου. Το γονικό είναι ένα μέλος του επόμενου υψηλότερου επιπέδου μιας ιεραρχίας, το οποίο συνδέεται απευθείας με το τρέχων μέλος. Η γονική τιμή είναι συνήθως μια συνένωση των τιμών όλων των θυγατρικών τιμών της. Για παράδειγμα, σε μια χρονική ιεραρχία που περιέχει τα επίπεδα “Τρίμηνο”, “Μήνας” και “Ημέρα”, το στοιχείο Πρώτο Τρίμηνο είναι το γονικό μέλος του στοιχείου “Ιανουάριος”.¹

¹ Επισκόπηση της ηλεκτρονικής διαδικασίας ανάλυσης (On-Line Analytical Processing – OLAP), 5 Αυγούστου 2008, 17:21, <http://office.microsoft.com/el-gr/excel/HP101774371032.aspx>

Η επιχειρηματική ευφυΐα είναι η διαδικασία εξαγωγής δεδομένων από μια βάση δεδομένων OLAP και η ανάλυσή τους, με σκοπό την εξαγωγή πληροφοριών που μπορούμε να χρησιμοποιήσουμε για τη λήψη εμπειριστατωμένων επιχειρηματικών αποφάσεων καθώς και για τη λήψη μέτρων. Για παράδειγμα, η διαδικασία OLAP και η επιχειρηματική ευφυΐα βοηθούν να δοθούν απαντήσεις στους ακόλουθους τύπους ερωτήσεων που αφορούν επιχειρησιακά δεδομένα:

- Με ποιόν τρόπο είναι δυνατό να συγκριθούν οι συνολικές πωλήσεις όλων των προϊόντων για το 2007 με τις συνολικές πωλήσεις τους 2006;
- Πως μπορεί να συγκριθεί η τρέχουσα αποδοτικότητά μας με αυτήν του ίδιου χρονικού διαστήματος κατά τη διάρκεια των προηγούμενων 5 ετών;
- Πόσα χρήματα ξόδεψαν πέρυσι οι πελάτες με ηλικία μεγαλύτερη των 35 ετών και πως έχει αλλάξει αυτή στη διάρκεια του χρόνου;
- Πόσα προϊόντα πωλήθηκαν σε δύο συγκεκριμένες χώρες/περιοχές αυτόν το μήνα σε σύγκριση με τον ίδιο μήνα πέρυσι;
- Για κάθε ηλικιακή ομάδα πελατών, ποια είναι η ανάλυση της αποδοτικότητας (τόσο το ποσοστιαίο όσο και το συνολικό περιθώριο) ανά κατηγορία προϊόντων;
- Βρείτε τους κορυφαίους και τους κατώτατους πωλητές, διανομείς, προμηθευτές, πελάτες, συνεργάτες ή καταναλωτές.²

Ο Codd (1993), πρωτοέθεσε και τις απαιτήσεις που πρέπει να τηρεί ένα OLAP σύστημα.

- Πολυδιάστατη αναπαράσταση πληροφορίας
- Προσπελάσιμη
- Αρχιτεκτονική πελάτη/εξυπηρετητή
- Πολυχρηστική υποστήριξη
- Ευέλικτη αναφορά και απόκριση
- Ενιαίος χρόνος απόκρισης
- Απεριόριστος αριθμός διαστάσεων και επιπέδων ιεραρχίας

Μια πολυδιάστατη βάση δεδομένων είναι σε θέση να εκτελέσει πολύπλοκους μαθηματικούς υπολογισμούς πολύ εύκολα. Τα δεδομένα αναφέρονται και ορίζονται μαθηματικά. Η δομή αναπαράστασης του μαθηματικού μοντέλου είναι τέτοια ώστε επιτρέπει στο χρήστη να εξερευνά ελεύθερα τα δεδομένα και να λαμβάνει τον τύπο και τη μορφή των δεδομένων που επιθυμεί το δυνατόν συντομότερα. Οι OLAP εξυπηρετητές χρησιμοποιούν πολυδιάστατες δομές για την αποθήκευση των δεδομένων και των συσχετίσεών τους. Αυτές

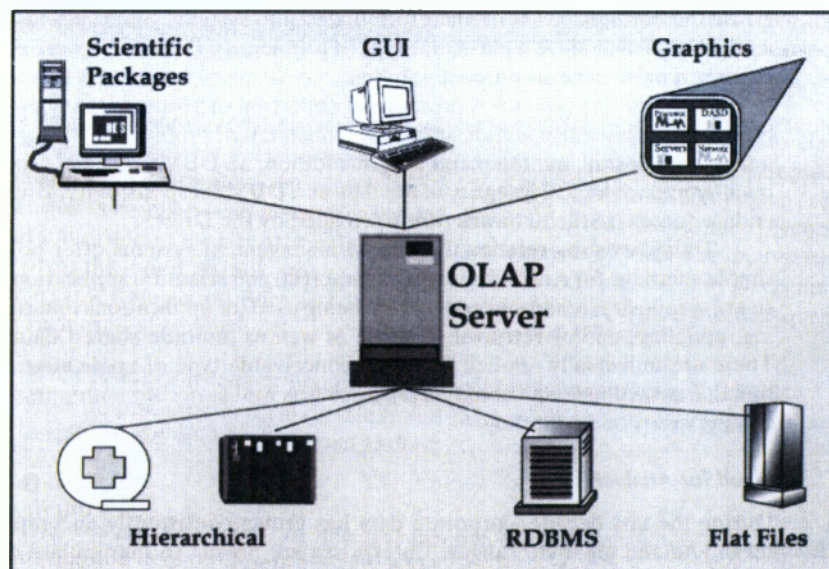
² Στο ίδιο, <http://office.microsoft.com/el-gr/excel/HP101774371032.aspx>

οι δομές αναπαριστούνται πιο εύκολα με υπερκύβους κάθε ακμή των οποίων θεωρείται και μια διάσταση.

Οι εφαρμογές OLAP χαρακτηρίζονται από δυναμική πολυδιάστατη ανάλυση των δεδομένων της επιχείρησης, παρέχοντας επιπλέον στο χρήστη δυνατότητες μοντελοποίησης των προβλημάτων, άντλησης των κατώτερων και λεπτομερέστερων δεδομένων και υπολογισμών.

Γενικά, οι εφαρμογές OLAP βοηθούν τον χρήστη να δημιουργεί αναλύσεις μέσα από πολλαπλές ερωτήσεις του τύπου “what-if” και έτσι να μοντελοποιεί το σενάριό του. Οι εφαρμογές OLAP έχουν γίνει συνώνυμες με την πολυδιάστατη παρουσία των δεδομένων. Αυτή η πολυδιάστατη παρουσίαση ενισχύεται και υποστηρίζεται από πολυδιάστατες βάσεις δεδομένων παρέχοντας έτσι στις OLAP εφαρμογές τη βάση για τον υπολογισμό και την ανάλυση των δεδομένων.

Η ανάπτυξη και η εξέλιξη της απευθείας αναλυτικής διαδικασίας (OLAP) οφείλεται κυρίως σε δύο λόγους: Στη ραγδαία αύξηση των ποσοτήτων των δεδομένων και την ταυτόχρονη ανάγκη για ταχεία ανάλυσή τους. Τα δεδομένα προέρχονται από στοιχεία και καταχωρήσεις προερχόμενα μέσα ή έξω από την επιχείρηση. Στο σχήμα 1-1 παρουσιάζεται ο server υπολογιστής OLAP εφαρμογής και ο ρόλος του ανάμεσα στους διάφορους τύπους βάσεων δεδομένων και αρχείων. Φαίνεται χαρακτηριστικά ο μεσολαβητικός ρόλος της OLAP εφαρμογής ανάμεσα στις βάσεις δεδομένων και στα δεδομένα που χρειάζονται οι χρήστες – αναλυτές. Η εφαρμογή βρίσκεται στο κέντρο του διαγράμματος.



Σχήμα 1-1: Ο ρόλος ενός OLAP Server

Πηγή: *Συστήματα Υποστήριξης Αποφάσεων*, Κεφ.11, Data Warehouse and OLAP, σελ 243, <http://www.ergasva.tuc.gr/Users/Matsatsinis/Courses/SYA/sya.htm>

Η πολυδιάστατη ανάλυση που παρέχουν οι OLAP εφαρμογές δίνουν τη δυνατότητα της άμεσης πρόσβασης σε μεγάλα ποσά δεδομένων και πληροφοριών από διαφορετικές πτυχές της επιχείρησης ή και έξω από αυτή. Πολλές φορές ακόμα, η εξαγωγή ενός αποτελέσματος δεν βγαίνει από την απλή παράθεση των δεδομένων αλλά απαιτείτε σύγκριση, ανάλυση και αλληλοσυσχέτισή τους.

Η συσχέτιση των ιστορικών δεδομένων, των ήδη επεξεργασμένων και συγκεντρωτικών στοιχείων και των έτοιμων πληροφοριών δημιουργεί το πολυδιάστατο περιβάλλον που χρειάζεται για να εφαρμοστεί μια OLAP εφαρμογή. Για να επιτευχθεί αυτό στο μέγιστο βαθμό απαιτούνται μηχανήματα μεγάλης υπολογιστικής ισχύος.

Ο αναλυτής πρέπει να ακολουθήσει συγκεκριμένα βήματα για την επεξεργασία των δεδομένων του:

- Ανάπτυξη σύνθετων αναλύσεων. Ένα OLAP σύστημα πρέπει να είναι ικανό να προσαρμόζεται στις απαιτήσεις του χρήστη. Ο βαθμός ευελιξίας του συστήματος αντικατοπτρίζεται ακόμα από τη δυνατότητα που παρέχει να αναπτύσσει σύνθετα και πολύπλοκα αναλυτικά μοντέλα

- Ο χρήστης OLAP ορίζει τις διαστάσεις σε ένα πολυδιάστατο πρόβλημα αποτελούμενο από διαφορετικά δεδομένα. Μπορεί να δημιουργήσει μοντέλα μερικών έως και εκατοντάδων διαστάσεων. Σαν παράδειγμα τεσσάρων διαστάσεων μπορεί να αναφερθεί τα εβδομαδιαία έσοδα διαφορετικών προϊόντων, ανά τμήμα και ανά κατάσταση αλυσίδας καταστημάτων

- Η χρήση διαφορετικών συνόλων του διαμορφωμένου χώρου του προβλήματος. Ο χρήστης αναλύει τα δεδομένα σύμφωνα με τις απαιτήσεις του προβλήματος, παίρνοντας υπόψη διαφορετικά τμήματα του πολυδιάστατου χώρου.³

1.2 Κατηγορίες OLAP

Υπάρχουν τέσσερις κύριες κατηγορίες OLAP:

1) Πολυδιάστατο (Multidimensional): Το πολυδιάστατο σύστημα OLAP (το οποίο μερικές φορές αναφέρεται και ως MOLAP) αποτελεί την παραδοσιακή μορφή του OLAP. Αυτή η μορφή OLAP δουλεύει καλά με τις βάσεις δεδομένων που περιλαμβάνουν χαρακτηριστικά όπως κώδικες, τοποθεσίες, προϊόντα ή χρονικά διαστήματα.

³ Συστήματα Υποστήριξης Αποφάσεων, Κεφ.11, Data Warehouse and OLAP, 7 Αυγούστου 2008, 14:30, σελ 242-244, <http://www.ergasya.tuc.gr/Users/Matsatsinis/Courses/SYA/sva.htm>

2) **Σχεσιακό (Relational):** Το σχεσιακό σύστημα OLAP παρέχει λειτουργικότητα OLAP από τη χρήση σχεσιακών βάσεων δεδομένων και οικεία σχεσιακά εργαλεία για να αποθηκεύσει και να αναλύσει πολυδιάστατα δεδομένα. Αυτή η προσέγγιση στηρίζεται σε υπάρχουσα σχεσιακή τεχνολογία και αναπαριστά μια φυσική προέκταση όλων των εταιρειών που ήδη χρησιμοποιούν σχεσιακά συστήματα διαχείρισης βάσεων δεδομένων στους οργανισμούς τους.⁴

3) **Υβριδικό (Hybrid):** Το υβριδικό OLAP μπορεί να λάβει διάφορες μορφές και το κύριο χαρακτηριστικό του είναι η χρήση εξειδικευμένων συσχετισμών και αποθήκευσης.

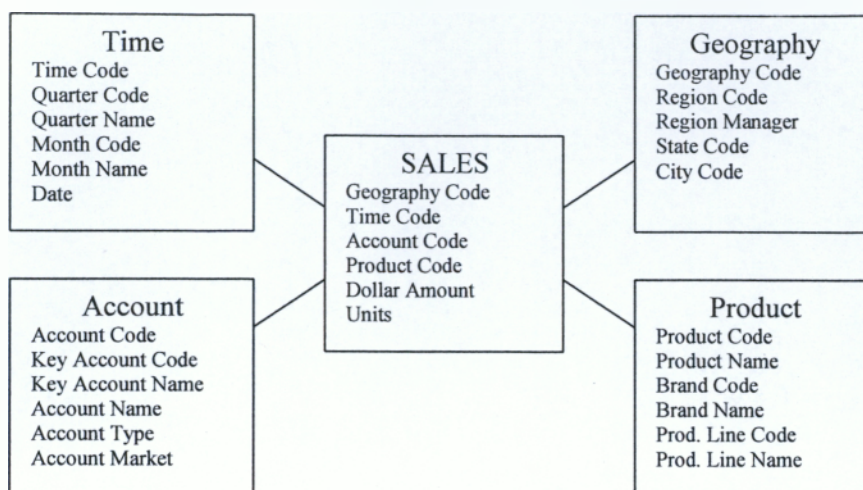
4) **Άλλα:** Τα συστήματα που δεν εμπίπτουν στις παραπάνω κατηγορίες είναι τα Desktop (DOLAP), Real-Time (RTOLAP) και βασίζονται στο Web (WOLAP) OLAPs.⁵

Οι πιο βασικές κατηγορίες γύρω από τις οποίες συγκεντρώνονται τα εργαλεία OLAP είναι η αρχιτεκτονική Multidimensional OLAP και η αρχιτεκτονική Relational OLAP, οι οποίες έχουν να κάνουν με την φυσική αποθήκευση δεδομένων. Το πλεονέκτημα της MOLAP είναι ότι παρέχει μια άμεση πολυδιάστατη όψη των δεδομένων ενώ τα πλεονεκτήματα της ROLAP είναι ότι μπορεί να ενσωματωθεί εύκολα σε υπάρχοντα σχεσιακά συστήματα και τα σχεσιακά δεδομένα μπορούν να αποθηκευθούν πιο αποδοτικά από τα πολυδιάστατα δεδομένα.

Σε μια αρχιτεκτονική ROLAP, τα δεδομένα οργανώνονται σε σχήμα αστέρα (star) ή νιφάδας (snowflake). Ένα σχήμα αστέρα αποτελείται από ένα κεντρικό πίνακα πληροφοριών (fact table) και διάφορους αποκανονικοποιημένους πίνακες διάστασης (dimension tables). Τα μέτρα (measures) της πληροφορίας αποθηκεύονται στο πίνακα πληροφοριών (πχ. Dollar Amount, Units στο πίνακα SALES). Για κάθε διάσταση του πολυδιάστατου μοντέλου υπάρχει και ένας πίνακας διάστασης (πχ. Geography, Product, Time, Account) με όλα τα επίπεδα συνάθροισης και τις επιπλέον ιδιότητες των επιπέδων αυτών. Η κανονικοποιημένη εκδοχή του σχήματος αστέρα είναι το σχήμα νιφάδας όπου κάθε επίπεδο μιας διάστασης έχει το δικό του πίνακα. (Για πιο ολοκληρωμένη άποψη δείτε το σχήμα 1-2 της επόμενης σελίδας).

⁴ P. Rob & C. Coronel, *Database Systems: Design, Implementation, & Management*, Course Technology, Cambridge χ.χ., σελ. 591, 593

⁵ Τι είναι OLAP, 8 Αυγούστου 2008, 15:32, <http://www.tech-faq.com/lang/el/olap.shtml>



Σχήμα 1-2: Σχήμα αστέρα

Πηγή: Π. Βασιλειάδης, Σ. Σκιαδόπουλος και Τ. Σελλής, *Θέματα Μοντελοποίησης και Επερώτησης Πολυδιάστατων Βάσεων Δεδομένων*, Εθνικό Μετσόβιο Πολυτεχνείο, Αθήνα χ.χ., σελ. 4

Τα πολυδιάστατα συστήματα βάσεων δεδομένων (Multidimensional Data Base Management System – MDBMS) αποθηκεύουν τα δεδομένα σε πολυδιάστατους πίνακες. Κάθε διάσταση του πίνακα αναπαριστά μια διάσταση του κύβου. Τα περιεχόμενα του πίνακα είναι τα μέτρα του κύβου. Τα συστήματα MDBMS απαιτούν τον προϋπολογισμό όλων των πιθανών συναθροίσεων, γι' αυτό και είναι πιο γρήγορα από τα τυπικά σχεσιακά συστήματα, αλλά πιο απαιτητικά στην ανανέωση της πληροφορίας και την διαχείρισή της.⁶

1.3 Χαρακτηριστικά OLAP

Μία OLAP εφαρμογή πρέπει να έχει τα παρακάτω χαρακτηριστικά:

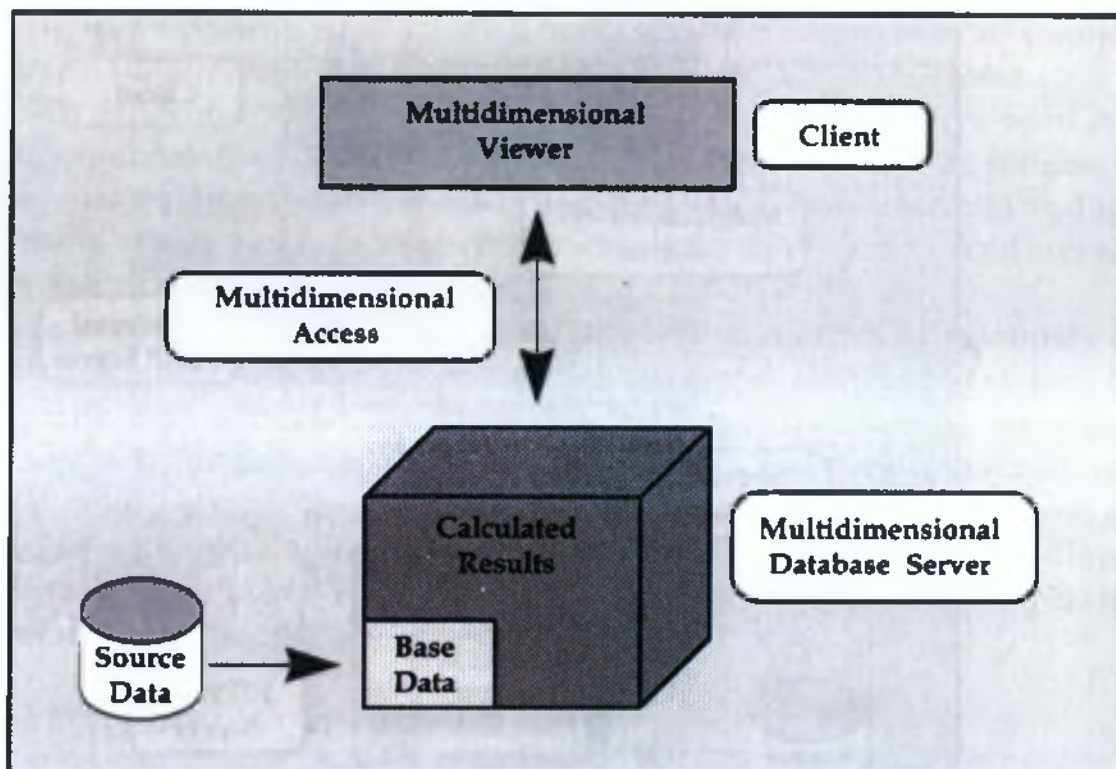
- Να περιέχει μαθηματικά, στατιστικά και αναλυτικά μοντέλα όπως ποσοστά, κατανομές, μεθόδους παλινδρόμησης κ.α.
- Να συνενώνει και να συσχετίζει τα διάφορα δεδομένα (αρχικά και εξαγόμενα)
- Να υποστηρίζει αναλύσεις και σχεδιασμούς του τύπου “what-if”
- Να έχει άμεση απόκριση και να παρέχει ταχεία εξαγωγή αποτελεσμάτων
- Να μπορεί να εξελίσσεται εύκολα και γρήγορα
- Να διαχειρίζεται μεγάλα ποσά δεδομένων

⁶ Π. Βασιλειάδης, Σ. Σκιαδόπουλος και Τ. Σελλής, *Θέματα Μοντελοποίησης και Επερώτησης Πολυδιάστατων Βάσεων Δεδομένων*, Εθνικό Μετσόβιο Πολυτεχνείο, Αθήνα χ.χ., σελ. 3-4

Τα παραπάνω χαρακτηριστικά αποτελούν ταυτόχρονα και τις απαιτήσεις για τον σχεδιασμό και την ανάπτυξη ενός OLAP περιβάλλοντος. Επιπλέον η μελέτη του υπάρχοντος συστήματος στην επιχείρηση κρίνεται αναγκαία για την ευκολότερη και γρηγορότερη εφαρμογή μιας OLAP εφαρμογής (οι περισσότεροι σχεδιαστές OLAP εφαρμογών αναφέρουν ως πλεονέκτημα των εφαρμογών τους την άμεση προσαρμογή στα υπάρχοντα πληροφοριακά συστήματα).⁷

1.4 Εφαρμογές OLAP

Μια απευθείας αναλυτική διαδικασία (OLAP) αποτελείται από ένα αριθμό εφαρμογών και δεδομένων τα οποία εμπεριέχονται σε μία συγκεκριμένη πολυδιάστατη βάση δεδομένων. Είναι απαραίτητο να κατανοηθεί ο ρόλος του κάθε τμήματος, ώστε να μπορεί ο χρήστης να απολαμβάνει την μέγιστη απόδοση από την εφαρμογή.



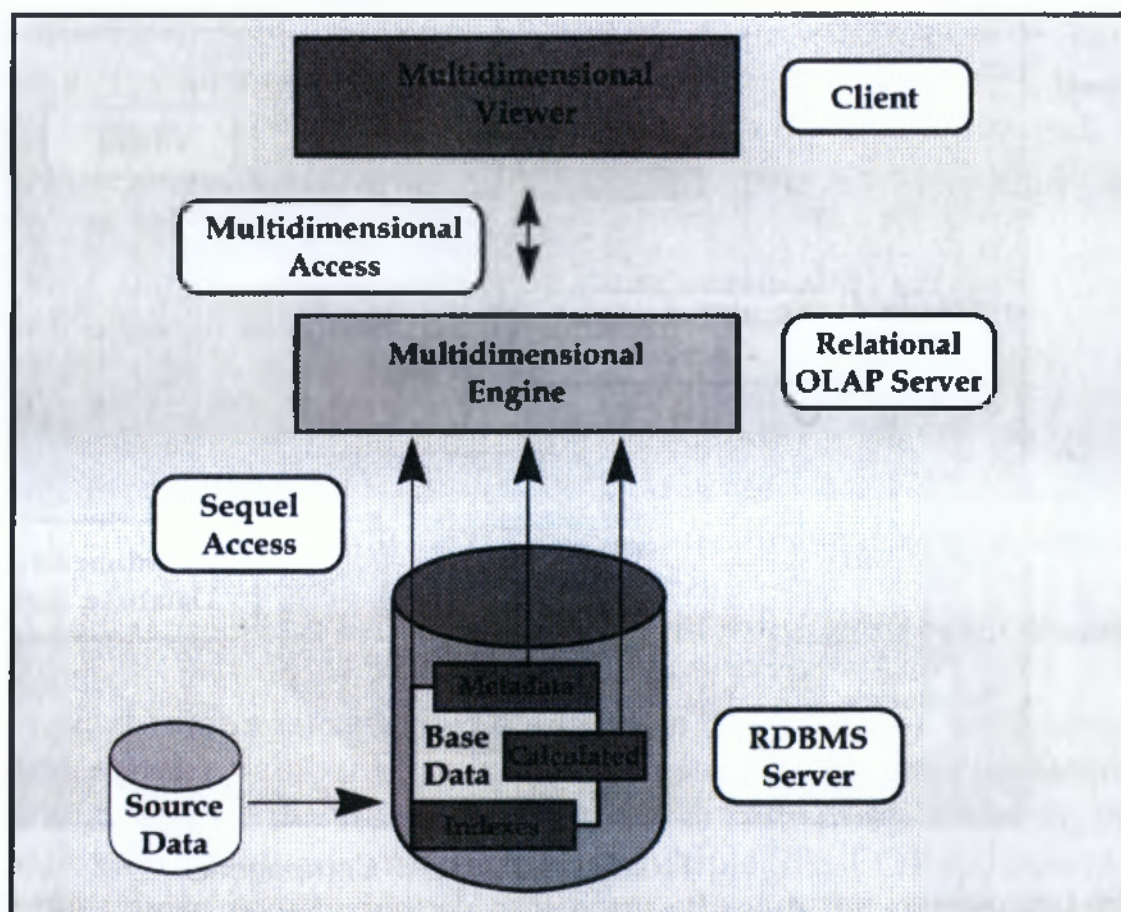
Σχήμα 1-3: Τα στοιχεία μιας OLAP εφαρμογής

Πηγή: *Συστήματα Υποστήριξης Αποφάσεων*, Κεφ.11, Data Warehouse and OLAP, σελ. 245, <http://www.ergasya.tuc.gr/Users/Matsatsinis/Courses/SYA/sya.htm>

⁷ *Συστήματα Υποστήριξης Αποφάσεων*, Κεφ.11, Data Warehouse and OLAP, 7 Αυγούστου 2008, 14:30, σελ. 247 <http://www.ergasya.tuc.gr/Users/Matsatsinis/Courses/SYA/sva.htm>

Τα συγκεντρωτικά και συνοπτικά δεδομένα καθώς και τα μερικώς επεξεργασμένα αποτελέσματα είναι οι αρχικές πηγές μιας OLAP εφαρμογής. Δεν αποτελούν τα στοιχεία τα οποία χρησιμοποιεί άμεσα η απευθείας αναλυτική διαδικασία, αλλά είναι η βάση (data base) των δεδομένων που θα χρησιμοποιηθούν. Το κύριο πλεονέκτημα των βασικών δεδομένων είναι η επιτάχυνση των επεξεργασιών και η άμεση πρόσβαση τους.

Η προετοιμασία των δεδομένων πριν τη χρήση της απευθείας αναλυτικής διαδικασίας γίνεται με εργαλεία, όπως SQL (Structured Query Language) και μηχανές δημιουργίας πολυδιάστατων χώρων. Υπάρχουν, πλέον και OLAP εφαρμογές που δίνουν την δυνατότητα στον χρήστη να επεξεργαστεί τα δεδομένα μέσα από αυτές και να δημιουργήσει, για παράδειγμα, τον πολυδιάστατο χώρο που εκείνος χρειάζεται. Στο σχήμα παρουσιάζεται μία τέτοια εφαρμογή OLAP τριών επιπέδων.



Σχήμα 1-4: Εφαρμογή OLAP τριών επιπέδων

Πηγή: *Συστήματα Υποστήριξης Αποφάσεων*, Κεφ.11, Data Warehouse and OLAP, σελ. 246, <http://www.ergasya.tuc.gr/Users/Matsatsinis/Courses/SYA/sva.htm>

Σε αυτό το σημείο θα πρέπει να αναφερθεί ο ρόλος των metadata δεδομένων. Όπως και στις εφαρμογές αποθήκευσης δεδομένων, έτσι και στις OLAP εφαρμογές, το metadata είναι ένα απαραίτητο στοιχείο για τη λειτουργία τους. Χρησιμοποιείται για να περιγράψει πολλές εφαρμογές όπως ιεραρχικές συσχετίσεις, αποθηκευτικές φόρμουλες, σημαντικές πληροφορίες για τους μετασχηματισμούς των δεδομένων (χρονικές, ποιοτικές, τοπολογικές), στοιχεία που αφορούν τις αναλυτικές διαδικασίες κ.α. Από τα παραπάνω διαφαίνεται ο ουσιαστικός ρόλος των δεδομένων metadata και στις OLAP εφαρμογές, αφού περικλείει όλες τις απαραίτητες πληροφορίες για τη σύνδεση των στοιχείων της και τη λειτουργία της.

Η πολυδιάστατη προβολή (multidimensional viewing) των δεδομένων είναι ο τρόπος επικοινωνίας του χρήστη με το σύστημα. Προσαρμόζεται ανάλογα με τις ανάγκες και τις απαιτήσεις του χρήστη. Έτσι, λογιστές χρησιμοποιούν πολυδιάστατες προβολές μέσα από λογιστικά φύλλα, διοικητικά στελέχη απαιτούν απλές και κατανοητές φόρμες προβολής, αναλυτές μάρκετινγκ χρησιμοποιούν γραφικές απεικονίσεις, ενώ άλλοι χρησιμοποιούν τα δεδομένα των αναλύσεων μέσα από άλλα προγράμματα ή από γλώσσες προγραμματισμού όπως η Visual Basic.

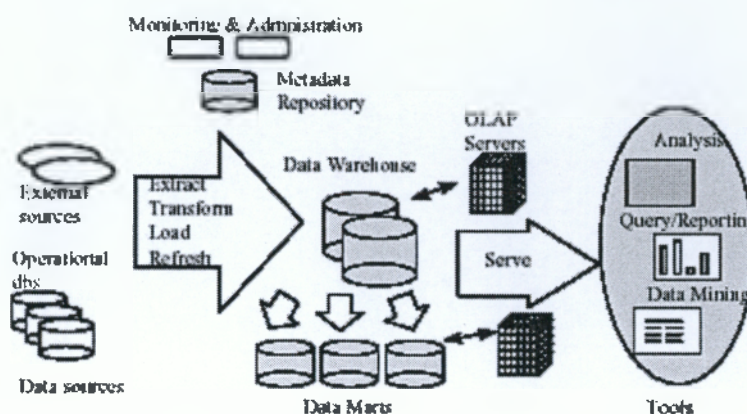
Τα εργαλεία επεξεργασίας των δεδομένων καθορίζονται ανάλογα με τις απαιτήσεις των χρηστών και τις δυνατότητες του συστήματος. Το τελευταίο είναι ένας περιοριστικός παράγοντας για τη δημιουργία OLAP εφαρμογών, αφού απαιτείται μεγάλη υπολογιστική ισχύς για ολοκληρωμένα συστήματα που περιέχουν πολλά αναλυτικά μοντέλα επεξεργασίας των δεδομένων και προβολής των πληροφοριών.⁸

1.5 Γενική Αρχιτεκτονική Συστημάτων Διαχείρισης Βάσεων Δεδομένων για OLAP

Ένα ΣΔΒΔ που χρησιμοποιείται για OLAP περιέχει συνήθως, εκτός από την Αποθήκη Δεδομένων, και εργαλεία για εξαγωγή (extract) των δεδομένων από πολλαπλές επιχειρησιακές Βάσεις Δεδομένων και γενικότερα εξωτερικές πηγές, καθώς και για καθαρισμό (cleaning), μετατροπή και ενσωμάτωση των δεδομένων αυτών στην Αποθήκη Δεδομένων. Επίσης, υποστηρίζονται εργαλεία για ανανέωση (refresh) ανά περιοδικά χρονικά Διαστήματα της Αποθήκης Δεδομένων με νέα δεδομένα, τα οποία αντιστοιχούν στις αλλαγές των δεδομένων στις πηγές, καθώς και για εξαγωγή παλαιότερων δεδομένων από την Αποθήκη Δεδομένων, τοποθετώντας τα ίσως σε πιο αργά αποθηκευτικά μέσα για αρχειακούς λόγους.

⁸ Στο ίδιο , σελ 244-247

Μια τυπική αρχιτεκτονική ενός ΣΔΒΔ που χρησιμοποιείται για OLAP φαίνεται στο σχήμα 1-5.



Σχήμα 1-5: Αρχιτεκτονική ενός ΣΔΒΔ που χρησιμοποιείται για OLAP

Πηγή: Ι. Ρούσσοσ , *Εργαστήριο Συστημάτων Βάσεων Γνώσεων και Δεδομένων* , Εθνικό Μετσόβιο Πολυτεχνείο , Αθήνα 2003, σελ 24

Καθώς μια Αποθήκη Δεδομένων χρησιμοποιείται για προχωρημένη στήριξη αποφάσεων, είναι ιδιαίτερα σημαντικό τα δεδομένα που είναι αποθηκευμένα σε αυτή να είναι σωστά. Εφόσον όμως, στην πράξη, υπάρχουν μεγάλες ποσότητες δεδομένων που προέρχονται από πολλές διαφορετικές πηγές, η πιθανότητα να εμφανιστούν λάθη ή ανωμαλίες στα εισερχόμενα δεδομένα είναι αρκετά αυξημένη. Γι' αυτό τον λόγο, εργαλεία που βοηθάνε στον εντοπισμό και τη διόρθωση τυχόν ανωμαλιών είναι πολύτιμα. Παραδείγματα, όπου ο καθαρισμός των δεδομένων είναι απαραίτητος, είναι ασυνεπή μήκη πεδίων, ασυνεπείς περιγραφές, ασυνεπείς αναθέσεις τιμών, έλλειψη κάποιων πεδίων, εγγραφές που έρχονται σε αντίθεση με τους περιορισμούς ακεραιότητας, διπλοεγγραφές και αναληθή στοιχεία (προερχόμενα συνήθως από φόρμες προς συμπλήρωση, όπως ερωτηματολόγια). Τα εργαλεία για τον καθαρισμό των δεδομένων κυμαίνονται από εργαλεία που κάνουν χρήση απλών κανόνων μετατροπής έως εργαλεία που χρησιμοποιούν ασαφής (fuzzy) τεχνικές και προχωρημένες τεχνικές ανάλυσης των δεδομένων, με στόχο την ανακάλυψη κανόνων και συσχετίσεων σε αυτά.

Ύστερα από την εξαγωγή των δεδομένων από τις πηγές, τον καθαρισμό και τη μετατροπή τους, υπάρχει περίπτωση να απαιτείται επιπλέον επεξεργασία τους πριν εισαχθούν στην Αποθήκη Δεδομένων. Στα επιπλέον αυτά στάδια συγκαταλέγονται: ο έλεγχος των τελικών δεδομένων εάν ικανοποιούν τους περιορισμούς ακεραιότητας, η ταξινόμησή τους, η σύνοψη (summarization) ή η εφαρμογή κάποιας συναθροιστικής συνάρτησης (aggregation) σε ένα τμήμα των δεδομένων για την κατασκευή των παραγόμενων πινάκων της Αποθήκης

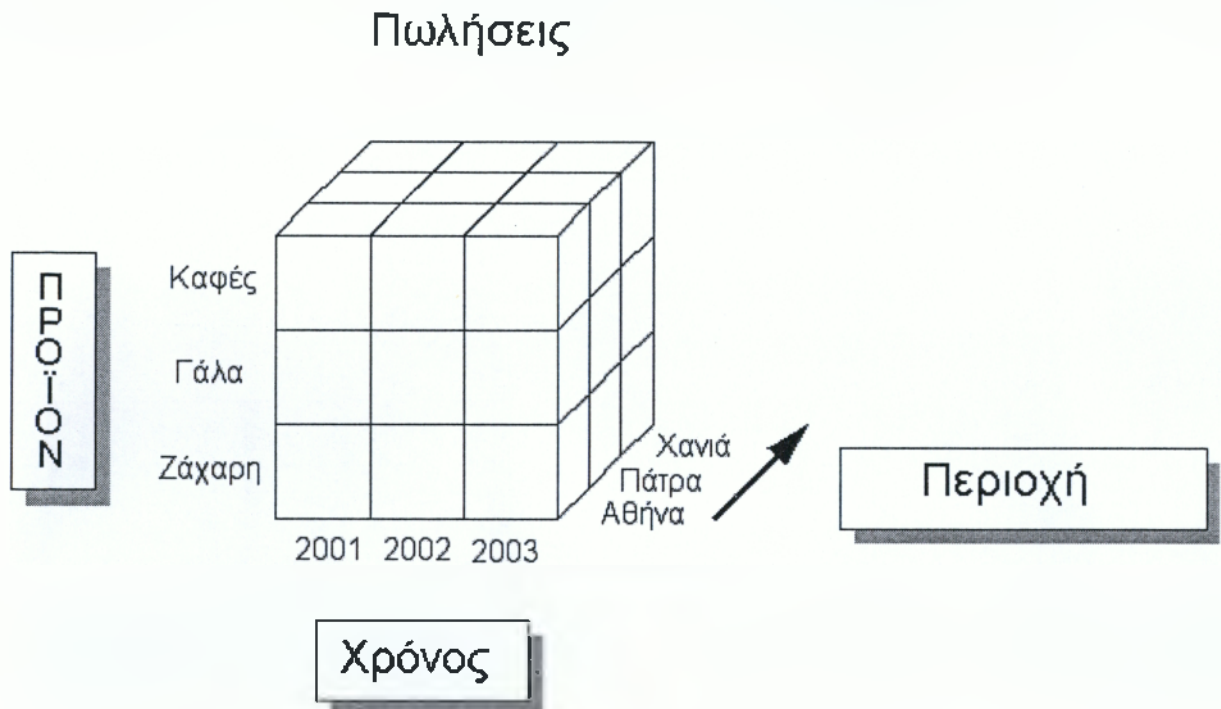
Δεδομένων. Συνήθως, η εισαγωγή των δεδομένων στην Αποθήκη Δεδομένων γίνεται με μαζική φόρτωσή τους (batch load), κάνοντας χρήση μεθόδων επεξεργασίας pipelining και διαμερισμένου (partitioned) παραλληλισμού των εργασιών, καθώς το μέγεθος των δεδομένων είναι ιδιαίτερα μεγάλο και υπάρχουν μικρά χρονικά περιθώρια (συνήθως βραδινές ώρες), κατά τα οποία η Αποθήκη Δεδομένων μπορεί να τεθεί εκτός λειτουργίας για να πραγματοποιηθούν οι εργασίες αυτές. Η χρήση σειριακών μεθόδων θα χρειαζόταν πολύ μεγάλα χρονικά Διαστήματα, καθώς η φόρτωση ενός Terabyte δεδομένων μπορεί να διαρκέσει εβδομάδες έως και μήνες. Όμως, ακόμα και με τη χρήση παραλληλισμού, η πλήρης φόρτωση (full load) μιας Αποθήκης Δεδομένων εξακολουθεί να χρειάζεται πολύ χρόνο, οπότε συνήθως υπάρχει και η επιλογή της αυξητικής φόρτωσης (incremental load), η οποία ελαττώνει το μέγεθος των δεδομένων που εισάγονται κάθε φορά στην Αποθήκη Δεδομένων. Χρησιμοποιείται κυρίως κατά την ανανέωση (refresh) της Αποθήκης Δεδομένων, όπου μόνο οι καινούργιες ή οι αλλαγμένες εγγραφές εισάγονται.

Επιπρόσθετα της βασικής Αποθήκης Δεδομένων, είναι δυνατόν να υπάρχει και ένας αριθμός από data marts. Τα data marts είναι ακριβώς ίδια (ως σχεδιασμός και γενικότερη αρχιτεκτονική) με μία Αποθήκη Δεδομένων, μόνο που είναι πολύ μικρότερα σε μέγεθος και εξυπηρετούν τις απαιτήσεις ορισμένων, συνήθως λίγων, συγκεκριμένων εφαρμογών. Αποτελούν μια όψη της κεντρικής Αποθήκης Δεδομένων, διατηρώντας μόνο τις μετρήσιμες τιμές και τις Διαστάσεις που χρειάζονται στις εφαρμογές που τα χρησιμοποιούν. Παραδείγματος χάριν, μία αλυσίδα πολυκαταστημάτων μπορεί να έχει, εκτός από την κεντρική Αποθήκη Δεδομένων, ένα data mart με τις πωλήσεις μόνο ως προς το χρόνο και το προϊόν, για να εξυπηρετεί εφαρμογές ανάλυσης που δεν ενδιαφέρονται για την περιοχή πώλησης ή τα στοιχεία του πωλητή. Τα data marts αποτελούν αποκεντρωμένο (decentralized) τμήμα του συστήματος και έχουν το ιδιαίτερο χαρακτηριστικό ότι τα δεδομένα τους δεν προέρχονται από τις εξωτερικές πηγές αλλά από την κεντρική Αποθήκη Δεδομένων. Ο στόχος των data marts είναι να προσφέρουν γρηγορότερες απαντήσεις, λόγω του μικρότερου και πιο ευέλικτου σχήματός τους, σε συχνά χρησιμοποιούμενες εφαρμογές ανάλυσης, οι οποίες δεν χρειάζονται όλο το εύρος των δεδομένων που υπάρχουν στην Αποθήκη Δεδομένων και συγχρόνως να ελαφρύνουν το φόρτο εργασίας της κεντρικής Αποθήκης Δεδομένων.

Η διαχείριση των δεδομένων της Αποθήκης Δεδομένων και των data marts πραγματοποιείται από έναν ή περισσότερους OLAP servers, οι οποίοι προσφέρουν πολυδιάστατη όψη των αποθηκευμένων δεδομένων σε ένα σύνολο από εφαρμογές, όπως εργαλεία επερωτήσεων (query tools), ανάλυσης, αυτόματης εγγραφής report και εξόρυξης

δεδομένων (data mining). Επίσης, υπάρχει ένας αποθηκευτικός χώρος (repository) για την αποθήκευση και διαχείριση των μεταδεδομένων (metadata), στα οποία συγκαταλέγονται πληροφορίες για τους χρήστες, τις διαθέσιμες βάσεις, τις Διαστάσεις, τα data marts, τις πηγές από τις οποίες προέρχονται τα δεδομένα και άλλες διαχειριστικής φύσης πληροφορίες. Τέλος, διαθέσιμα είναι εργαλεία για επίβλεψη (monitoring) και διαχείριση (administration) του συστήματος.

Το λογικό μοντέλο που χρησιμοποιείται από τις εφαρμογές OLAP είναι η πολυδιάστατη περιγραφή των δεδομένων μίας Αποθήκης Δεδομένων.



Σχήμα 1-6: Παράδειγμα Πολυδιάστατων δεδομένων

Πηγή: Ι. Ρούσσος, *Εργαστήριο Συστημάτων Βάσεων Γνώσεων και Δεδομένων*, Εθνικό Μετσόβιο Πολυτεχνείο, Αθήνα 2003, σελ 26

Στην πολυδιάστατη μοντελοποίηση των δεδομένων, υπάρχει ένα σύνολο από μετρήσιμες τιμές ή μέτρα (measures) ή γεγονότα (facts), οι οποίες αποτελούν τα αντικείμενα της ανάλυσης. Παραδείγματα τέτοιων μετρήσιμων τιμών είναι οι πωλήσεις, οι αποδοχές, ο αριθμός των επισκεπτών σε μία έκθεση ή ο αριθμός των συνδρομητών ενός περιοδικού. Οι μετρήσιμες τιμές εξαρτώνται από ένα σύνολο από Διαστάσεις, οι οποίες χαρακτηρίζουν μοναδικά τις μετρήσιμες τιμές και τις τοποθετούν σε μία περιοχή του πολυδιάστατου χώρου, όπως φαίνεται και από το απλοποιημένο παράδειγμα τριών Διαστάσεων του σχήματος 1-6.

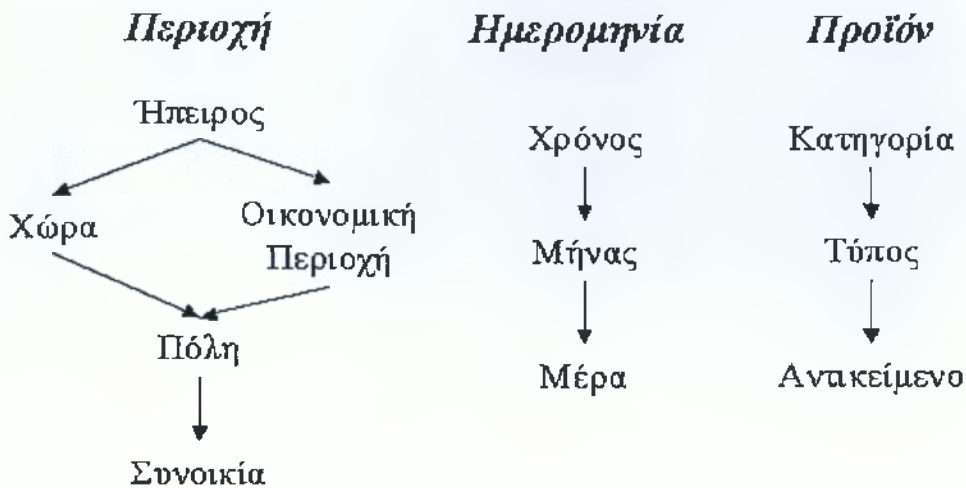
Για παράδειγμα, οι Διαστάσεις που θα μπορούσαν να χαρακτηρίζουν τη μετρήσιμη τιμή «πωλήσεις» είναι η περιοχή, το όνομα προϊόντος και η ημερομηνία αγοραπωλησίας.

Τα μέτρα είναι κυρίως αριθμητικές τιμές, οι οποίες αντιστοιχούν στις μετρήσεις κάποιας παραμέτρου σχετιζόμενης με ένα συμβάν σε συγκεκριμένα χρονικά σημεία (π.χ. χρηματικό ποσό που εμφανίζεται σε μια γραμμή ενός τιμολογίου σε μια συγκεκριμένη ημέρα ή το υπόλοιπο ενός λογαριασμού στο τέλος κάθε ημέρας κλπ.) και αναμένεται να αλλάζουν συνεχώς και γρήγορα, καθώς νέα συμβάντα προκύπτουν (π.χ. νέες γραμμές τιμολογίου εμφανίζονται κ.ο.κ.). Τα δεδομένα Διαστάσεων (ή απλά οι Διαστάσεις) θεωρούνται σχεδόν στατικά (ή αργά μεταβαλλόμενα) στο χρόνο. Οι τιμές διάστασης χαρακτηρίζουν μια συγκεκριμένη τιμή ενός μέτρου με τον ίδιο τρόπο που οι συντεταγμένες προσδιορίζουν ένα συγκεκριμένο σημείο σε ένα πολυδιάστατο χώρο.

Αυτή η πολυδιάστατη αναπαράσταση των δεδομένων εμπεριέχει έναν μεγαλύτερο βαθμό οργάνωσης των δεδομένων σε σχέση με έναν σχεσιακό πίνακα, καθώς οι συσχετίσεις ανάμεσα στα δεδομένα είναι ενσωματωμένες στην πολυδιάστατη δομή. Αυτό γίνεται εύκολα αντιληπτό, εάν σκεφτούμε ότι μια επερώτηση ανάλυσης των δεδομένων ή οποιαδήποτε άλλη αίτηση επεξεργασίας τους είναι συνήθως μία αίτηση επιλογής και οργάνωσής τους με έναν συγκεκριμένο τρόπο. Σε αντίθεση με τη δόμηση των δεδομένων στο σχεσιακό μοντέλο, όπου απαιτείται σάρωση όλων των εγγραφών για να ανακαλυφθούν αυτές που πιθανώς είναι οι κατάλληλες, στην πολυδιάστατη αναπαράσταση τα δεδομένα είναι ήδη οργανωμένα και απλά ζητείται ένα τμήμα του συνολικού χώρου.

Οι Διαστάσεις οργανώνονται σε διάφορα συναθροιστικά επίπεδα καθορίζοντας μία ιεραρχία, δηλαδή ένα μονοπάτι συνάθροισης. Παραδείγματα ιεραρχιών Διαστάσεων φαίνονται στο σχήμα 1-7, όπου το πιο λεπτομερές επίπεδο βρίσκεται στο κάτω μέρος και το επίπεδο με τη μικρότερη διακριτότητα στο επάνω.

Διαστάσεις:



Σχήμα 1-7: Παραδείγματα ιεραρχιών Διαστάσεων

Πηγή: Ι. Ρούσσος, *Εργαστήριο Συστημάτων Βάσεων Γνώσεων και Δεδομένων*, Εθνικό Μετσόβιο Πολυτεχνείο, Αθήνα 2003, σελ 27

Καλούμε το πιο λεπτομερές επίπεδο, επίπεδο ίχνους (grain level) της διάστασης. Μια συγκεκριμένη τιμή σε ένα επίπεδο L μιας διάστασης D καλείται μέλος (member) του L , π.χ., η τιμή "Αθήνα" είναι μέλος του επιπέδου Πόλη της διάστασης Περιοχή.

Τα πεδία (attributes) των Διαστάσεων που ορίζουν τα επίπεδα των ιεραρχιών ονομάζονται ιεραρχικά πεδία (hierarchical attributes) και σχετίζονται μεταξύ τους με τον τρόπο που σχετίζονται και τα αντίστοιχα συναθροιστικά επίπεδα. Παραδείγματος χάριν, μια συνοικία ανήκει σε μια συγκεκριμένη πόλη και αυτή σε μία συγκεκριμένη χώρα και αντίστροφα μια χώρα έχει πολλές πόλεις και αυτές πολλές συνοικίες (σχέση 1:N). Είναι σύνηθες για μία διάσταση να αποτελείται από περισσότερες της μιας ιεραρχίες. Παράδειγμα τέτοιας διάστασης αποτελεί η διάσταση 'Περιοχή' του σχήματος 1-7, της οποίας οι ιεραρχίες είναι: «συνοικία, πόλη, χώρα, ήπειρος» και «συνοικία, πόλη, οικονομική περιοχή, ήπειρος». Τα συναθροιστικά επίπεδα που ορίζουν τις διαφορετικές ιεραρχίες μια διάστασης θα μπορούσαν να ήταν εντελώς διαφορετικά, συνήθως όμως τουλάχιστον το επίπεδο ίχνους (grain level) είναι κοινό.

Εκτός από τα ιεραρχικά πεδία, σε μία διάσταση συνήθως υπάρχουν και ορισμένα πεδία τα οποία χαρακτηρίζουν κάποια ιεραρχικά επίπεδα και τα οποία ονομάζονται πεδία γνωρίσματος (feature attributes). Τα πεδία γνωρίσματος δεν συμμετέχουν στη δημιουργία των ιεραρχιών, καθώς ο στόχος τους είναι η αποθήκευση κάποιων επιπλέον πληροφοριών για ένα

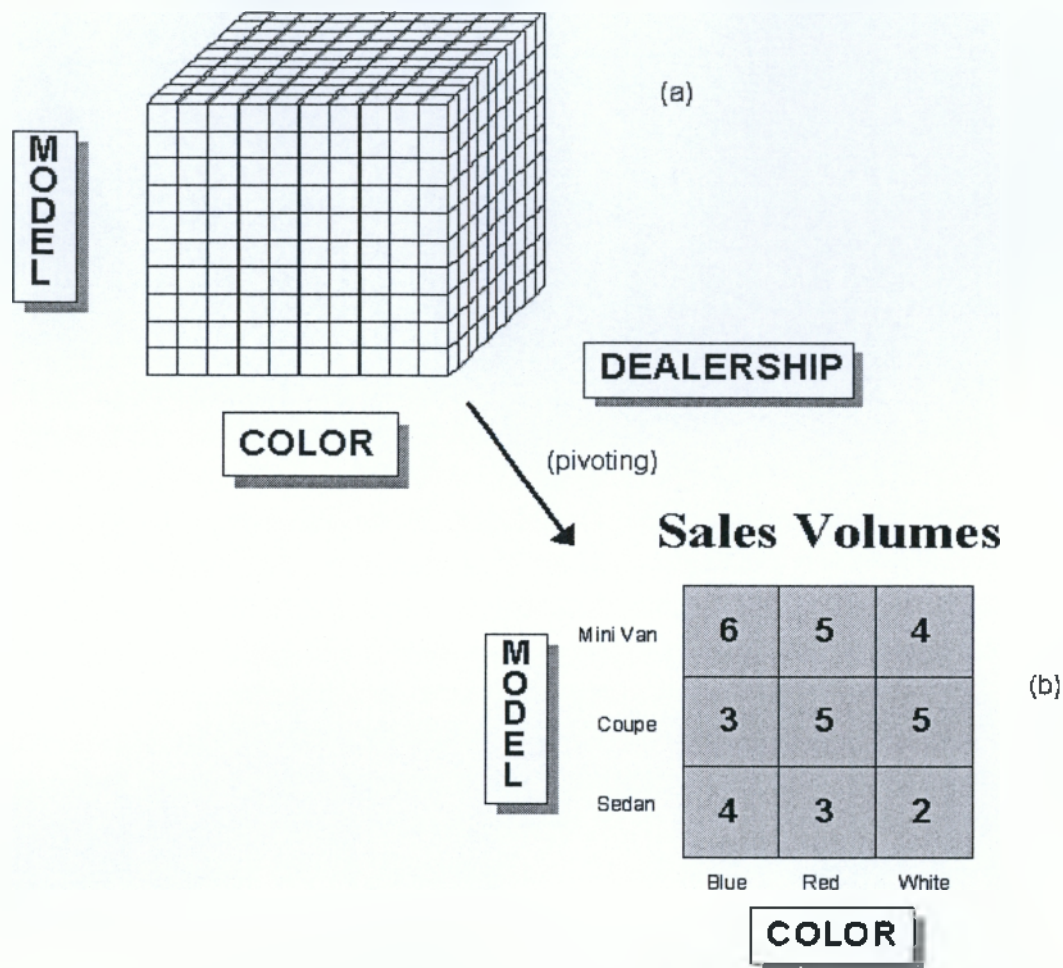
ιεραρχικό επίπεδο. Παραδείγματα πεδίων γνωρίσματος είναι ο αριθμός των κατοίκων μίας πόλης, το κατά κεφαλήν εισόδημα μιας οικονομικής περιοχής ή το βάρος ενός προϊόντος.

Ο χρόνος είναι μία ιδιαίτερα σημαντική διάσταση για το πεδίο της προχωρημένης στήριξης αποφάσεων, καθώς οι περισσότερες επερωτήσεις απαιτούν ανάλυση των δεδομένων κατά τη διάρκεια χρονικών Διαστημάτων (π.χ. η ανάλυση και εύρεση τάσεων στα δεδομένα). Σε πολλές πολυδιάστατες Βάσεις Δεδομένων, ο χρόνος αποτελεί ενσωματωμένη στο σύστημα διάσταση με ορισμένες εκ των προτέρων ιεραρχίες και ειδικές λειτουργίες. Αυτή η ιδιαίτερη μεταχείριση του χρόνου προσφέρει δύο σημαντικά πλεονεκτήματα. Αρχικά, ελαχιστοποιεί την προσπάθεια που απαιτείται για τον ορισμό εξεζητημένων ιεραρχιών κάθε φορά που κατασκευάζεται μία αποθήκη δεδομένων. Επιπλέον, καθώς η διάσταση του χρόνου είναι ενσωματωμένη στο σύστημα, βελτιώνεται η απόδοση των επερωτήσεων κατά τις συχνές αναλύσεις των δεδομένων ως προς τον χρόνο, την ανάλυση χρονοσειρών ή την εύρεση τάσεων.

1.6 Βασικές Λειτουργίες του Πολυδιάστατου Μοντέλου για OLAP

Λειτουργία κλειδί του λογικού μοντέλου για OLAP είναι η συνάθροιση (aggregation) των μέτρων ως προς μία ή περισσότερες Διαστάσεις. Παράδειγμα τέτοιας λειτουργίας είναι ο υπολογισμός των συνολικών πωλήσεων για κάθε χώρα (ή και για κάθε χρόνο). Χαρακτηριστικό παράδειγμα αποτελεί η λειτουργία του pivoting. Ας θεωρήσουμε το παρακάτω πολυδιάστατο σχήμα (schema), όπου αποθηκεύονται οι πωλήσεις αυτοκινήτων ως προς το μοντέλο του αυτοκινήτου, το χρώμα του και τον έμπορο (Σχήμα 1-8.a):

Sales Volumes

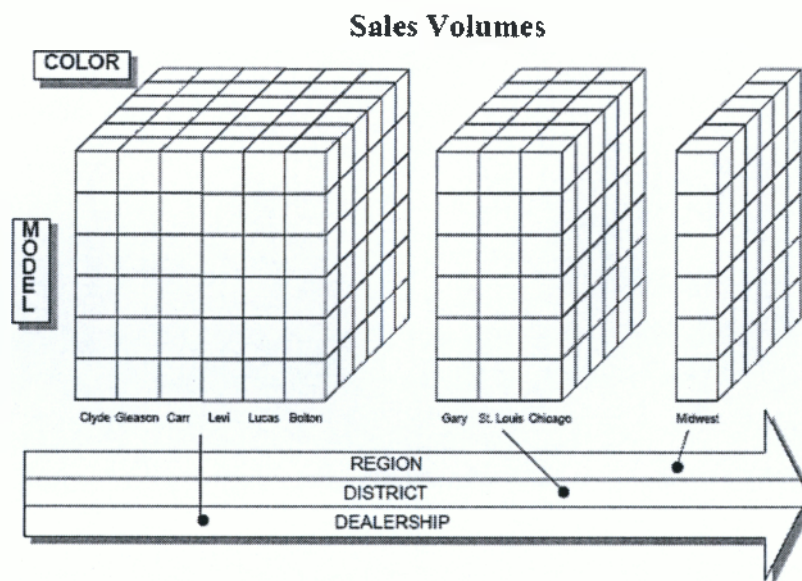


Σχήμα 1-8: Παράδειγμα λειτουργίας pivoting

Πηγή: Ι. Ρούσσοσ , *Εργαστήριο Συστημάτων Βάσεων Γνώσεων και Δεδομένων* , Εθνικό Μετσόβιο Πολυτεχνείο, Αθήνα 2003, σελ. 29

Η πιο απλή χρήση του pivoting είναι η επιλογή δύο Διαστάσεων, ως προς τις οποίες προβάλλονται τα συναθροιστικά αποτελέσματα ενός μέτρου, στην περίπτωση μας των πωλήσεων αυτοκινήτων. Τα συναθροιστικά αποτελέσματα προβάλλονται συνήθως σε ένα πλέγμα (grid), όπου κάθε τιμή στις (x,y) συντεταγμένες αντιστοιχεί στην τιμή που προκύπτει από τη εφαρμογή της συναθροιστικής συνάρτησης σε όλες τις υπόλοιπες Διαστάσεις, όταν η πρώτη διάσταση (στο παράδειγμα του σχήματος 1-8.b το χρώμα) έχει την x κατά σειρά τιμή και η δεύτερη διάσταση (στο παράδειγμα του σχήματος 1-8.b το μοντέλο του αυτοκινήτου) έχει την y κατά σειρά τιμή. Παραδείγματος χάριν, εάν οι επιλεγμένες Διαστάσεις είναι το χρώμα και το μοντέλο του αυτοκινήτου, το πάνω αριστερά κελί του σχήματος 1-8.b περιέχει τις συνολικές πωλήσεις μπλε Mini Van για όλους τους εμπόρους.

Άλλες δημοφιλείς λειτουργίες είναι η συναθροιστική άνοδος (roll-up) και η αναλυτική κάθοδος (drill-down), οι οποίες έχουν ως αποτέλεσμα την κίνηση κατά μήκος ενός ή περισσότερων μονοπατιών συνάθροισης. Η λειτουργία roll-up σε κάποια δεδομένα έχει ως αποτέλεσμα τη μετακίνηση σε περισσότερα συναθροισμένες (aggregated) όψεις των δεδομένων, και αντιστοιχεί στο group-by του σχεσιακού σχήματος. Παραδείγματος χάριν, ας θεωρήσουμε στο προηγούμενο παράδειγμα ότι ο έμπορος (dealership) αποτελεί το επίπεδο ίχνους (grain level) μιας διάστασης με την ακόλουθη ιεραρχία: «έμπορος, περιφέρεια, περιοχή». Η λειτουργία roll-up στη συγκεκριμένη διάσταση έχει ως αποτέλεσμα να βλέπουμε τα δεδομένα ανά περιφέρεια (district) και στη συνέχεια ανά περιοχή (region), όπου πλέον κάθε κελί περιέχει τις συνολικές πωλήσεις όλων των εμπόρων της συγκεκριμένης περιφέρειας ή περιοχής (Σχήμα 1-9). Αντίστοιχα, η λειτουργία drill-down έχει το ακριβώς αντίθετο αποτέλεσμα από την roll-up, δηλαδή τη μετακίνηση σε πιο λεπτομερή δεδομένα.

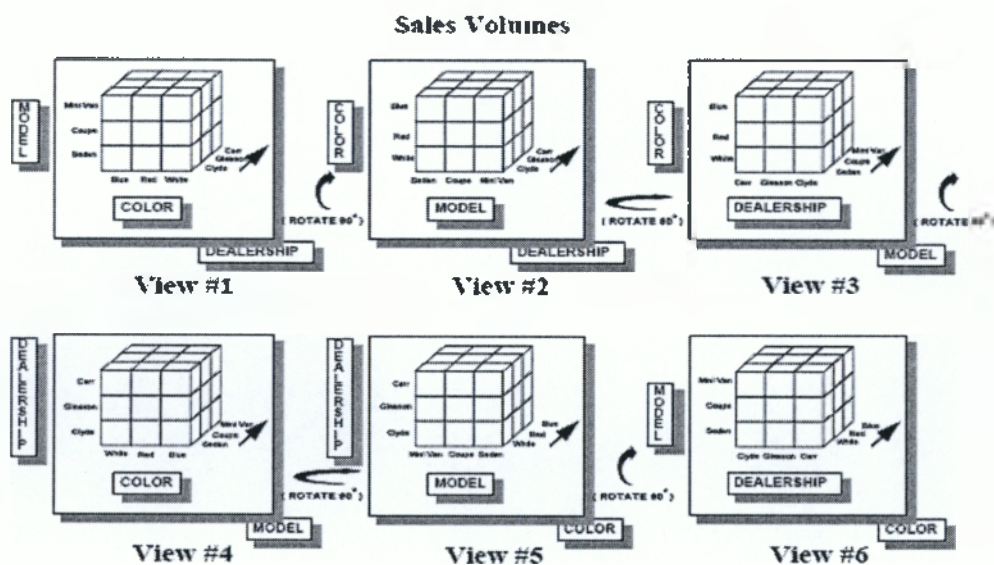


Σχήμα 1-9: Παράδειγμα λειτουργίας roll-up

Πηγή: Ι. Ρούσσος, *Εργαστήριο Συστημάτων Βάσεων Γνώσεων και Δεδομένων*, Εθνικό Μετσόβιο Πολυτεχνείο, Αθήνα 2003, σελ. 30

Τέλος, χρήσιμη λειτουργία, κυρίως κατά την οπτικοποίηση των δεδομένων, είναι αυτή της περιστροφής (rotation). Χρησιμοποιώντας ως παράδειγμα το ίδιο πολυδιάστατο σχήμα με τις τρεις Διαστάσεις, ένας χρήστης (αναλυτής) μπορεί να αποφασίσει ότι δεν τον ικανοποιεί να βλέπει τα δεδομένα ως προς το μοντέλο και το χρώμα του αυτοκινήτου, με τη διάσταση των εμπόρων να βρίσκεται κάθετα στο background, αλλά ότι θα επιθυμούσε να έχει τα δεδομένα ως προς το χρώμα και τους εμπόρους, με τη διάσταση των μοντέλων να βρίσκεται

κάθετα στο background. Σε ένα σχεσιακό περιβάλλον αυτή η αλλαγή όψης θα απαιτούσε την αποτίμηση μιας νέας πολύπλοκης επερώτησης ή την ταξινόμηση των δεδομένων. Αντίθετα, στο πολυδιάστατο μοντέλο, η νέα όψη προκύπτει με απλή περιστροφή στον n-διάστατο χώρο (3-διάστατο στην περίπτωσή μας) της παλιάς όψης. Για την ακρίβεια, εάν χρησιμοποιείται και για την αποθήκευση των δεδομένων στο φυσικό μέσο μια πολυδιάστατη δομή, η νέα όψη προκύπτει χωρίς νέους υπολογισμούς ή επερωτήσεις στα δεδομένα, απλά με την εφαρμογή στην παλιά όψη δύο περιστροφών 90° (μετάβαση από τη view #1 στη view #3 του σχήματος 1-10). Τα παραπάνω πραγματοποιούνται ταχύτατα σε σχέση με το σχεσιακό μοντέλο, καθώς είναι δυνατή η υποστήριξη από το σύστημα της περιστροφής 90° ως τελεστή σε φυσικό επίπεδο στην πολυδιάστατη δομή.⁹ Στο σχήμα 1-10 παρουσιάζονται ορισμένα παραδείγματα χρήσης της λειτουργίας περιστροφής:



Σχήμα 1-10: Παράδειγμα λειτουργίας περιστροφής (rotation)

Πηγή: Ι. Ρούσσος, *Εργαστήριο Συστημάτων Βάσεων Γνώσεων και Δεδομένων*, Εθνικό Μετσόβιο Πολυτεχνείο, Αθήνα 2003, σελ. 31

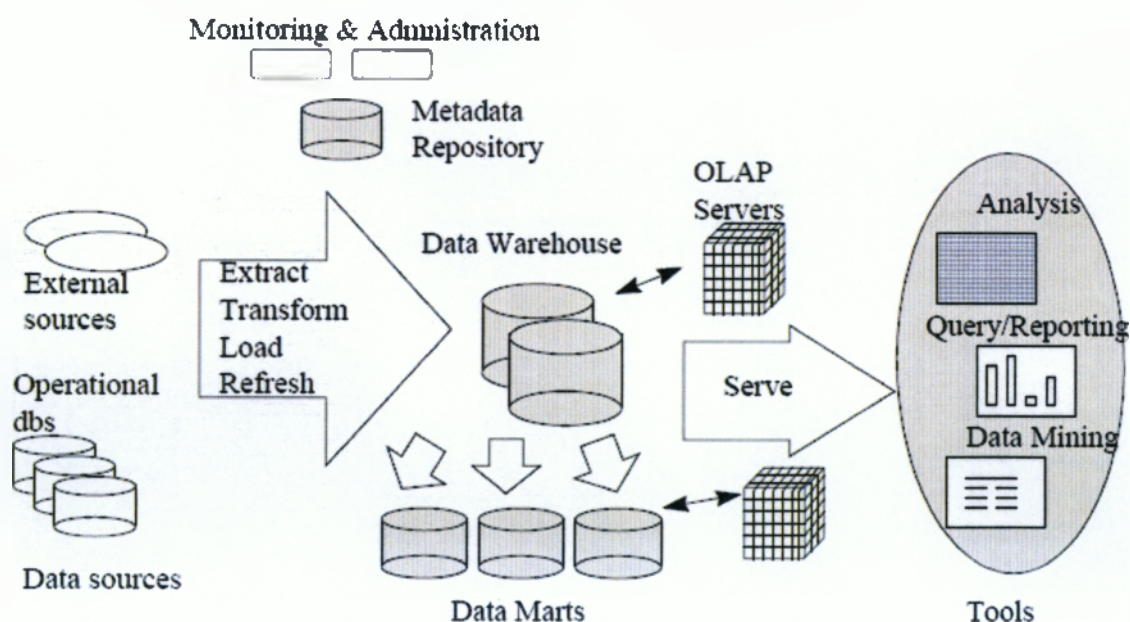
⁹ Ι. Ρούσσος, *Εργαστήριο Συστημάτων Βάσεων Γνώσεων και Δεδομένων*, Εθνικό Μετσόβιο Πολυτεχνείο, Αθήνα 2003, σελ. 23-31

ΚΕΦΑΛΑΙΟ 2

ΑΠΟΘΗΚΕΣ ΔΕΔΟΜΕΝΩΝ

2.1 Γενικά

Σημαντικά στοιχεία της υποστήριξης αποφάσεων αποτελούν τα data warehousing και το on-line analytical processing (OLAP). Οι απαιτήσεις για υποστήριξη αποφάσεων όσον αφορά τις ΒΔ είναι διαφορετικές από τις παραδοσιακές on-line transaction processing applications (OLTPA).



Σχήμα 2-1: Γενική αρχιτεκτονική ΣΔΒΔ

Πηγή: *Συστήματα Υποστήριξης Αποφάσεων*, Κεφ.11, Data Warehouse and OLAP, σελ. 229, <http://www.ergasva.tuc.gr/Users/Matsatsinis/Courses/SYA/sya.htm>

Οι αποθήκες δεδομένων (data warehouses) είναι ειδικές βάσεις δεδομένων που περιέχουν κατάλληλα επεξεργασμένα δεδομένα, προερχόμενα από τις επιμέρους βάσεις δεδομένων μιας επιχείρησης, που σκοπό έχουν να υποβοηθήσουν τον αποφασίζοντα ή να υποστηρίξουν ένα σύστημα υποστήριξης αποφάσεων.

Ως αποθήκευση δεδομένων (data warehousing) εννοούμε την διαδικασία ενσωμάτωσης δεδομένων που σχετίζονται μέσα σε απλές αποθήκες δεδομένων (data warehouses) από όπου ο χρήστης μπορεί εύκολα να διατυπώσει ειδικές ερωτήσεις, να εξάγει

λίστες-αναφορές και να παρουσιάζει αναλύσεις. Ένας άλλος ορισμός για τις αποθήκες δεδομένων αναφέρει ότι data warehouse ονομάζεται η αποθήκη των μόνο προς ανάγνωση αρχείων (read only data records), που σκοπό έχει να υποβοηθήσει τους αποφασίζοντες.

Υπάρχουν εμφανείς διακρίσεις μεταξύ των κλασικών βάσεων δεδομένων (databases) και των αποθηκών δεδομένων. Από τους ορισμούς διαφαίνεται ότι οι αποθήκες δεδομένων περιέχουν μόνο προς ανάγνωση δεδομένα. Ακόμα, στην αποθήκη δεδομένων καταλήγουν κατάλληλα επεξεργασμένα δεδομένα των επιμέρους βάσεων δεδομένων, διαφοροποιώντας την έτσι ως προς το περιεχόμενο των πληροφοριών. Πολλές φορές μάλιστα αυτά τα δεδομένα αποτελούν δομημένες πληροφορίες και όχι απλά μία καταγραφή απλών στοιχείων -πράγματα που εμφανίζονται στις απλές βάσεις δεδομένων.¹

2.2 Κατηγορίες

Υπάρχουν πολλές κατηγορίες αποθηκών δεδομένων, γεγονός που δικαιολογείται από την πολυπλοκότητα και την πολυμορφία των συστημάτων, καθώς από τη δυνατότητα που παρέχεται στον σχεδιαστή μιας data warehouse να εφαρμόσει διαφορετικές τεχνικές σύμφωνα με την κρίση του. Έτσι διακρίνουμε τις παρακάτω κατηγορίες:

- Αποθήκες δεδομένων μεγάλης υπολογιστικής ισχύος (host-based mainframe data warehouse), οι οποίες ανήκουν σε συστήματα υψηλής χωρητικότητας και μεγάλης υπολογιστικής ισχύος, όπως για παράδειγμα τα συστήματα IBM System/390, UNISYS και Data General Sequent Systems υποστηριζόμενα από βάσεις δεδομένων όπως Sybase, Oracle, Informix και DB2.

- Αποθήκες δεδομένων δικτύων LAN (LAN data warehouse), όπου τα δεδομένα λαμβάνονται είτε από κεντρικό υπολογιστή server, είτε μέσα από δίκτυα-workgroups. Το μέγεθος της αποθήκης εξαρτάται από τον τρόπο σχεδίασης του συστήματος.

- Πολυεπίπεδες αποθήκες δεδομένων (multistage data warehouses).
- Μόνιμες αποθήκες δεδομένων (stationary data warehouses)
- Αποθήκες διανομής δεδομένων (distributed data warehouses)
- Εικονικές αποθήκες δεδομένων (virtual data warehouses)²

¹ Συστήματα Υποστήριξης Αποφάσεων, Κεφ.11, Data Warehouse and OLAP, 7 Αυγούστου 2008, 14:30, σελ 229-230, <http://www.ergasya.tuc.gr/Users/Matsatsinis/Courses/SYA/sya.htm>

² Στο ίδιο, σελ 240-241

2.3 Χαρακτηριστικά

Χαρακτηριστικά που περιγράφουν ένα Data Warehouse:

Θεματοστρεφής: Τα δεδομένα οργανώνονται σύμφωνα με το θέμα προς εξέταση αντί της εφαρμογής π.χ. μια ασφαλιστική εταιρία που χρησιμοποιεί ένα Data Warehouse θα οργάνωνε τα δεδομένα της ανά πελάτη (υποχρεώσεις και απαιτήσεις) αντί ανά κατηγορία προϊόντος (ασφάλειες ζωής, αυτοκινήτου κτλ).

Ενιαία: Όταν τα δεδομένα μοιράζονται σε πολλές διαφορετικές εφαρμογές στο λειτουργικό τους περιβάλλον η κωδικοποίησή τους μπορεί να είναι διαφορετική και ασύμβατη. Για παράδειγμα σε μία εφαρμογή μπορεί το γένος να κωδικοποιείται με m και f (male, female) και σε μία άλλη με 1 και 0. Όταν δεδομένα εισάγονται στο data warehouse είναι δεδομένο ότι χρησιμοποιείται μία μορφή κωδικοποίησης.

Χρονολογικά δεδομένα: Το Data Warehouse δίνει την δυνατότητα να αποθηκευτούν σε αυτό δεδομένα ηλικίας έως 5 με 10 χρόνια προκειμένου να χρησιμοποιηθούν για συγκρίσεις και προβλέψεις.

Αμεταβλητότητα: Τα δεδομένα που εισάγονται στο Data Warehouse δεν αλλάζουν από την στιγμή που θα μπουν σε αυτό. Μόνο φορτώνονται και προσπελούνται.

Το Data Warehouse ανακτά δεδομένα λοιπόν από πολλά ετερογενή περιβάλλοντα τα μετασχηματίζει και τα αποθηκεύει βάσει ενός συλλεκτικού μοντέλου. Ο μετασχηματισμός και η τροποποίηση πραγματοποιείται κάθε φορά που απαιτείται μία ενημέρωση του Data Warehouse. Η πληροφορία που περιγράφει το μοντέλο μετασχηματισμού και αποθήκευσης της πρωτογενούς πληροφορίας καλείται metadata. Το metadata είναι ο τρόπος με τον οποίο ο τελικός χρήστης αντιλαμβάνεται τα δεδομένα στο Data Warehouse.

Τα metadata πρέπει να περιλαμβάνουν τουλάχιστον τα παρακάτω:

- Την δομή της πληροφορίας
- Τον αλγόριθμο επιλογής των δεδομένων
- Την συνάρτηση μετασχηματισμού των δεδομένων από το λειτουργικό περιβάλλον στο

Data Warehouse

Η αποθήκευση δεδομένων αποτελεί το πρώτο βήμα για την διαχείριση του τεράστιου όγκου των δεδομένων. Έχει γίνει αναπόσπαστο κομμάτι πολλών έτοιμων πληροφοριακών συστημάτων διοίκησης και ακόμα περισσότερο σε συστήματα υποστήριξης αποφάσεων αφού διαχειρίζεται δισεκατομμύρια bytes δεδομένων ή ακόμα και terabytes.

Ακόμα, η αποθήκη δεδομένων είναι αναγκαία λόγω της ανάγκης για ταχεία ανάλυση και εξαγωγή πληροφοριών για την επιχείρηση. Τα υπάρχοντα πληροφοριακά συστήματα

παρέχουν ως ένα βαθμό τη δυνατότητα στον χρήστη να εξαγει τέτοιου είδους πληροφορίες, από την άλλη όμως πολλές φορές τον παγιδεύουν λόγω της αδυναμίας τους να έχουν:

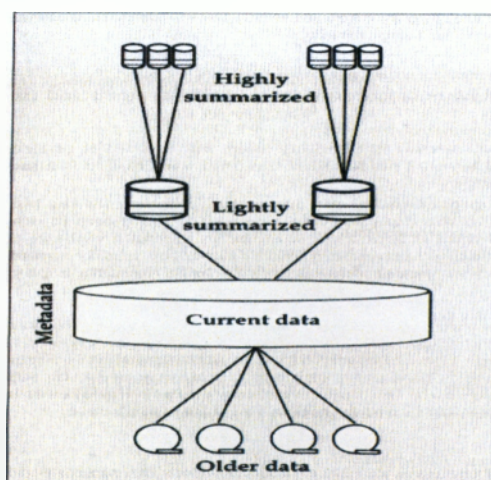
- Ιστορικά δεδομένα
- Πρόσβαση σε δεδομένα από διαφορετικές διαχειριστικές βάσεις δεδομένων (operational databases)
- Δυνατότητα άντλησης πληροφοριών από κατάλληλα διατυπωμένες ερωτήσεις
- Δυνατότητα πλήρους υποστήριξης για την λήψη απόφασης (υποστήριξη σε συστήματα υποστήριξης αποφάσεων).

Έτσι, λόγω των μεγάλων ωφελειών που παρέχει εφαρμόζεται σε:

Τηλε-επικοινωνίες	Εκπαίδευση
Τραπεζικές λειτουργίες	Ασφάλειες
Παραγωγή	Πωλήσεις καταναλωτικών αγαθών
Υγεία	Στρατηγικές μάρκετινγκ
Διαχείριση κινδύνου	Στατιστικές αναλύσεις
Χρηματοοικονομικές αναλύσεις	Ανάλυση στόχων
Μάρκετινγκ	Τάσεις αγοράς

2.4 Δομή και Λειτουργία

Μία αποθήκη δεδομένων έχει συγκεκριμένη δομή. Υπάρχουν διαφορετικά επίπεδα και στοιχεία που την αποτελούν. Στο σχήμα 2-2 παρουσιάζονται τα διαφορετικά συστατικά μιας τυπικής αποθήκης δεδομένων.



Σχήμα 2-2: Η δομή μίας αποθήκης δεδομένων

Πηγή: Συστήματα Υποστήριξης Αποφάσεων, Κεφ.11, Data Warehouse and OLAP, σελ. 232, <http://www.ergasya.tuc.gr/Users/Matsatsinis/Courses/SYA/sya.htm>

Από το σχήμα φαίνεται ότι τα κύρια στοιχεία μιας αποθήκης δεδομένων είναι:

- 1) Τα σύγχρονα δεδομένα (current data)
- 2) Τα παλιά δεδομένα (older data)
- 3) Τα συγκεντρωτικά δεδομένα (summarized data)
- 4) Τα metadata

Τα σύγχρονα δεδομένα αποτελούν και τα κύρια συστατικά μιας data warehouse. Τα δεδομένα αυτά: 1) παρουσιάζουν τα πιο πρόσφατα γεγονότα, γεγονότα όμως που εμφανίζουν κάποιο ενδιαφέρον, 2) είναι μεγάλα σε όγκο και αποθηκεύονται στα χαμηλότερα επίπεδα αποθήκευσης, 3) αποθηκεύονται σε δίσκους αποθήκευσης ώστε να είναι εύκολα προσπελάσιμα, κάνοντας όμως την διαχείριση τους πολύπλοκη (αυτό οφείλεται και στο μεγάλο όγκο τους).

Τα παλιά δεδομένα αποτελούνται από δεδομένα που δεν χρησιμοποιούνται συχνά. Λόγω της μη ευρείας χρήσης τους δεν είναι υποχρεωτική η αποθήκευσή τους σε ταχείς δίσκους ή ακόμα πολλές φορές μετά από την πάροδο του χρόνου διαγράφονται.

Τα συνοπτικά δεδομένα διακρίνονται σε δύο κατηγορίες, ανάλογα με την ανάγκη για επεξεργασία τους και την δυνατότητα αποθήκευσής τους.

1) Τα ελαφρώς συνοπτικά δεδομένα (lightly summarized data) τα οποία προέκυψαν από την επεξεργασία δεδομένων χωρίς να παρέχουν μεγάλες ποσότητες πληροφοριών. Αυτού του είδους τα δεδομένα αποθηκεύονται συνήθως σε αποθηκευτικούς δίσκους.

2) Τα πολύ συνοπτικά δεδομένα (highly summarized data) τα οποία είναι συμπαγή και πολύ εύκολα στην προσέγγισή τους. Αποθηκεύονται στις κύριες μονάδες των αποθηκών δεδομένων.

Πρέπει να σημειωθεί ότι η αποθήκευση συνοπτικών δεδομένων γίνεται μόνο στις περιπτώσεις όπου αυτά μπορούν να χρησιμοποιηθούν συχνά. Με άλλα λόγια, εάν παράγονται συγκεντρωτικά δεδομένα που η πιθανότητα να ξαναχρησιμοποιηθούν είναι μικρή, τότε δε γίνεται η αποθήκευσή τους.

Τα metadata είναι ένα πολύ σημαντικό στοιχείο μίας αποθήκης δεδομένων. Τα metadata είναι το “αρχείο” για το δεδομένο ή αλλιώς τα “δεδομένα” για το δεδομένο. Στις περισσότερες των περιπτώσεων τα metadata αποτίθενται σε διαφορετικά μέρη από τα υπόλοιπα δεδομένα μιας data warehouse. Και αυτό οφείλεται στο ότι δεν περιέχει απευθείας δεδομένα από τις διαχειριστικές βάσεις δεδομένων. Παρόλα αυτά, είναι επιθυμητή η γνώση από που προήλθε τα δεδομένα, πότε και από ποιον δημιουργήθηκαν, ποιο σύστημα χρησιμοποιήθηκε για την δημιουργία τους και διάφορες άλλες πληροφορίες που θα

βοηθήσουν στην τοποθέτησή τους μέσα στην αποθήκη. Συγκεντρωτικά, τα metadata χρησιμεύουν:

- 1) Στην τοποθέτηση των περιεχομένων μιας αποθήκης
- 2) Στην ταυτοποίηση των δεδομένων που προέρχονται από τις διαχειριστικές βάσεις
- 3) Κατά την επεξεργασία των δεδομένων των διαχειριστικών βάσεων
- 4) Σαν οδηγός στους αλγόριθμους που χρησιμοποιούνται για την συγχώνευση των σύγχρονων και των συγκεντρωτικών δεδομένων

Τα metadata περιέχουν πληροφορίες τουλάχιστον για:

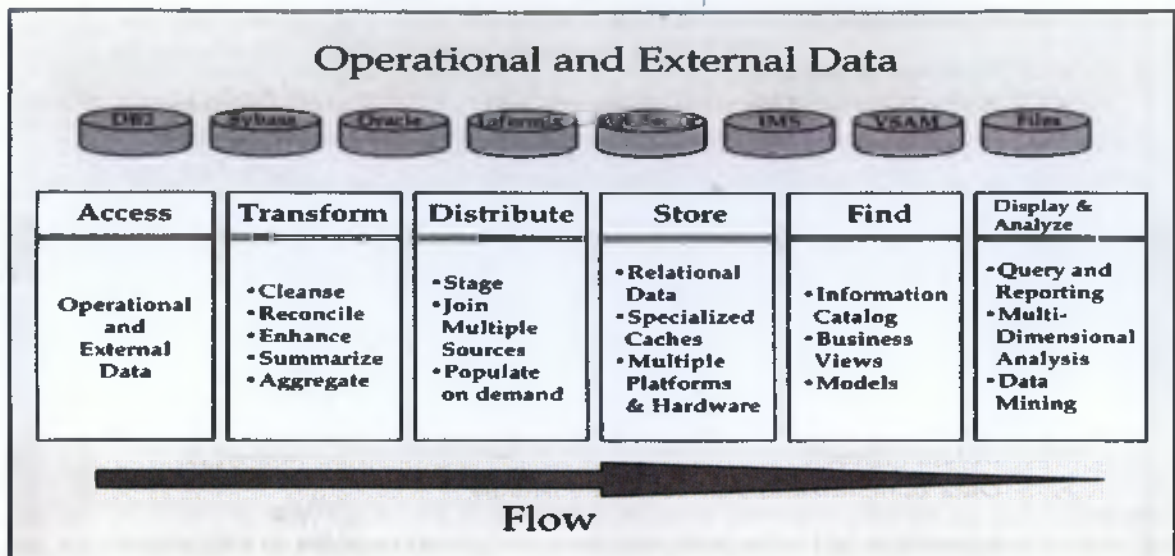
- Τη δομή του δεδομένου
- Τους αλγόριθμους που είναι αναγκαίους για την ταξινόμησή του, και
- Την χαρτογράφησή του από την διαχειριστική βάση δεδομένων στην αποθήκη

ενώ μπορεί ακόμα να περιέχει πληροφορίες για:

- Την καταγωγή και τις μετατροπές του
- Χρήσιμα στατιστικά στοιχεία που αφορούν τα δεδομένα

Έτσι, το metadata χαρακτηρίζεται σαν το κύριο συστατικό για την σωστή δημιουργία ενός δεδομένου που θα τοποθετηθεί μετέπειτα σε μία αποθήκη δεδομένων. Σε μία τυπική εφαρμογή, η αποθήκη δεδομένων είναι άμεσα συνδεδεμένη με το λειτουργικό της σύστημα μέσω των metadata. Κάθε αλλαγή που υπεισέρχεται στην αποθήκη, παρουσιάζεται άμεσα στο τερματικό του χρήστη της αποθήκης δεδομένων. Για παράδειγμα μπορεί να αναφερθεί η απαλοιφή ενός επιπέδου στην ιεραρχική δομή της επιχείρησης. Η εφαρμογή πρέπει να επανασχεδιάσει, χρησιμοποιώντας τα ανάλογα metadata, την ιεράρχηση όταν αυτό ζητηθεί.

Όσον αφορά την λειτουργία των αποθηκών δεδομένων στο παρακάτω σχήμα παρουσιάζεται η ροή των δεδομένων από την αρχική πηγή τους μέχρι τον χρήστη, συμπεριλαμβανομένων και των εφαρμογών που παρεμβάλλονται. Λόγω των ετερογενών πηγών των δεδομένων (διαφορετικές διαχειριστικές βάσεις δεδομένων), και λοιπόν προβλημάτων που προκύπτουν σε τέτοια ανομοιογενή συστήματα, παρεμβάλλονται μηχανισμοί που τα αποκαθιστούν, μετατρέποντάς τα και παραδίδοντάς τα στις αποθηκευτικές βάσεις δεδομένων. Αυτός ο μηχανισμός βασίζεται σε προεπιλεγμένο μοντέλο (σε αυτό βασίζεται ουσιαστικά και ο σχεδιασμός των data warehouses). Κύριο συστατικό -περιγράφει το μοντέλο ενώ παράλληλα διευκρινίζει τα στοιχεία των δεδομένων- είναι τα metadata.



Σχήμα 2-3: Οι λειτουργίες στις αποθήκες δεδομένων

Πηγή: Συστήματα Υποστήριξης Αποφάσεων, Κεφ.11, Data Warehouse and OLAP, σελ. 234, <http://www.ergasya.tuc.gr/Users/Matsatsinis/Courses/SYA/sya.htm>

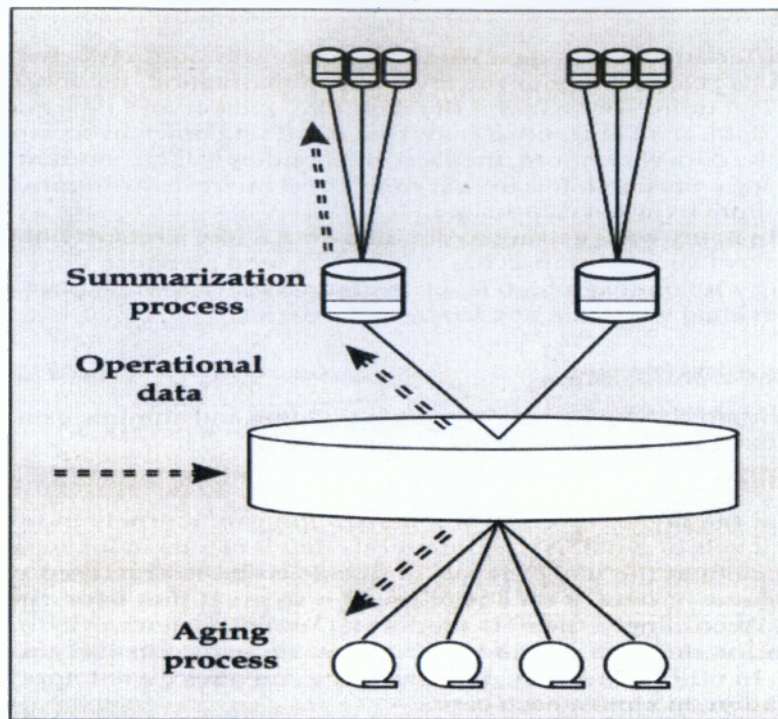
2.5 Ροή και Μετασχηματισμός Δεδομένων

Τα περισσότερα δεδομένα εισέρχονται μέσα στην αποθήκη δεδομένων απ' ευθείας από τις διαχειριστικές βάσεις δεδομένων. Στο σχήμα 2-4 παρουσιάζεται η ροή των δεδομένων μέσα σε μία αποθήκη δεδομένων, ανάλογα με τη χρήση που υπόκεινται. Έτσι, τα δεδομένα διακρίνονται σε:

- Διαγραφέντα (τα δεδομένα προς διαγραφή, μικρή χρησιμότητα)
- Συνοπτικά (είναι δεδομένα που χρησιμοποιούνται σε μεγαλύτερη συχνότητα)
- Δεδομένα-αρχεία

Η διαδικασία “παλαιώσης” (aging process) σε μία βάση δεδομένων μετατρέπει τα δεδομένα ανάλογα με την ηλικία τους από σύγχρονα δεδομένα (current data) σε παλαιά (older data) και για αυτό τον λόγο, τις περισσότερες φορές τα αποθέτει σε αποθηκευτικούς δίσκους. Η διαδικασία σύνοψης (summarized process) χρησιμοποιεί προσεκτικά επιλεγμένα δεδομένα με σκοπό να τα μετατρέψει σε συνοπτικά δεδομένα (summarized data), είτε ελαφρώς συνοπτικά δεδομένα (lightly summarized data), είτε πολύ συνοπτικά δεδομένα (highly summarized data).

Όσο πιο συγκεντρωτικά και συνοπτικά είναι τα δεδομένα σε μία data warehouse, τόσο ταχύτερη και αποτελεσματικότερη είναι η πρόσβαση για τη χρήση τους.



Σχήμα 2-4: Η ροή των δεδομένων μέσα σε μία αποθήκη δεδομένων

Πηγή: Συστήματα Υποστήριξης Αποφάσεων, Κεφ.11, Data Warehouse and OLAP, σελ. 235, <http://www.ergasva.tuc.gr/Users/Matsatsinis/Courses/SYA/sya.htm>

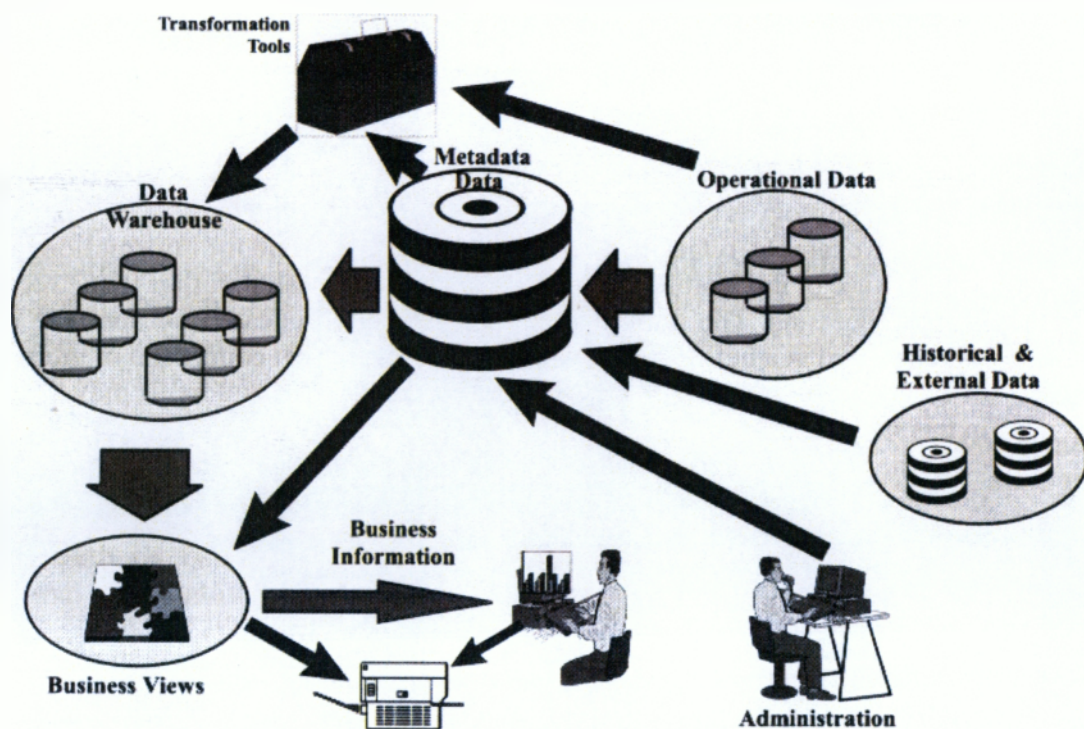
Οι περισσότερες επιχειρήσεις χρειάζονται δύο διαφορετικά είδη δεδομένων. Το ένα αφορά της διαχειριστικές εφαρμογές της (operational applications), ενώ το άλλο χρησιμοποιείται άμεσα από τα πληροφοριακά συστήματα. Στο σχήμα 2-5 παρουσιάζονται τα δύο αυτά συστήματα και τα δεδομένα που χρησιμοποιούν. Τα δεδομένα, όπως έχει αναφερθεί, διαφοροποιούνται ανά εφαρμογή. Για παράδειγμα, οι εφαρμογές, οι βάσεις δεδομένων και τα λειτουργικά συστήματα που χρησιμοποιούνται στο κατώτερα επίπεδα της επιχείρησης είναι σχεδιασμένα να παρέχουν ταχύτατη πρόσβαση, χωρίς όμως να είναι ικανά να παρέχουν πληροφορίες κατάλληλες για την λήψη μιας απόφασης -όπως δηλαδή κάνουν τα συστήματα αποφάσεων υποστηριζόμενα από data warehouses.

Τα δεδομένα μιας data warehouse (επομένως και ενός συστήματος υποστήριξης αποφάσεων) προκύπτουν από κατάλληλο μετασχηματισμό των διαχειριστικών δεδομένων. Το ποσοστό του μετασχηματισμού που απαιτείται, εξαρτάται από την ποιότητα και την δομή του αρχικού δεδομένου.

Μερικοί από τους μετασχηματισμούς που υπόκεινται τα δεδομένα προκειμένου να ενταχθούν σε μία αποθήκη δεδομένων είναι ο καθαρισμός (cleansing), η προσαρμογή-τυποποίηση (adjusting) και ο εμπλουτισμός (enhance procedure). Κατά τον καθαρισμό, ο

τεράστιος όγκος των δεδομένων των διαχειριστικών βάσεων ξεδιαλύνεται και καθαρίζεται από λανθασμένες καταχωρίσεις και διπλοεγγραφές. Κατά την προσαρμογή-τυποποίηση τα δεδομένα εντάσσονται σε ευρύτερες κατηγορίες ανάλογα με την διαμόρφωσή τους, καθώς και την πληροφορία που παρέχουν. Κατά τον εμπλουτισμό γίνεται η σύνθεση και η συμπίκνωση των δεδομένων, πολλές φορές δε προκύπτουν νέα δεδομένα εμπλουτισμένα από στοιχεία πολλών εκατοντάδων διαχειριστικών δεδομένων.

Θα πρέπει να αναφερθεί εδώ η σημαντικότητα των δεδομένων metadata, τα οποία παρέχουν την δυνατότητα, με τις πληροφορίες που περιέχουν, του ευκολότερου μετασχηματισμού και εν κατακλείδι, του καταλληλότερου σχεδιασμού της αποθήκης δεδομένων.



Σχήμα 2-5: Τα δεδομένα και η χρήση τους

Πηγή: Συστήματα Υποστήριξης Αποφάσεων, Κεφ.11, Data Warehouse and OLAP, σελ. 237, <http://www.ergasva.tuc.gr/Users/Matsatsinis/Courses/SYA/sya.htm>

2.6 Σχεδίαση και Εφαρμογές Αποθηκών Δεδομένων

Το μεγαλύτερο ποσοστό των επιχειρήσεων επιλέγουν τον σχεδιασμό και την εφαρμογή κεντρικών και ανεξάρτητων αποθηκών δεδομένων. Οι σχεδιαστές των συστημάτων αποσκοπούν σε:

- Μία απλή και ανεξάρτητη εφαρμογή που να διαχειρίζεται εύκολα μεγάλο όγκο δεδομένων
- Συγκεντρωτική παρουσίαση των δεδομένων της επιχείρησης στους αποφασίζοντες, συλλέγοντας τα διάσπαρτα δεδομένα από όλες τις πτυχές της
- Διαχωρισμό ανά κατηγορία και θέμα των δεδομένων (πολλές φορές σε διαφορετικές αποθήκες δεδομένων), αφού πολλές φορές η συγκέντρωση σε μία και μόνη αποθήκη δυσχεραίνει την πρόσβαση

Είναι αυτονόητο ότι μία εφαρμογή αποθήκης δεδομένων -η ακόμα περισσότερο μία έτοιμη εμπορική εφαρμογή- δεν ταιριάζει σε μία επιχείρηση. Για αυτό τον λόγο και τις περισσότερες φορές ο σχεδιασμός από την αρχή είναι επιβεβλημένος. Πολλές εταιρίες παρέχουν την δυνατότητα προσαρμογής ήδη υπάρχοντων εφαρμογών, ανάλογα με τις απαιτήσεις της κάθε επιχείρησης.

Σε κάθε περίπτωση όμως, οι παράγοντες που θα πρέπει να λαμβάνονται υπόψη είναι:

1) Πολλές επιχειρήσεις γνωρίζουν την αναγκαιότητα μιας αποθήκης δεδομένων, αλλά δεν είναι σε θέση να θέσουν τις προτεραιότητες και τις προδιαγραφές, πράγματα που μόνο οι εταιρίες εφαρμογών μπορούν να ξέρουν και να θέσουν. Σαν τέτοιες χαρακτηρίζονται το μέγεθος, η εγκατάσταση, η συχνότητα χρήσης και η συντήρηση. Για πολλές εταιρίες, όμως, μία ολοκληρωμένη μελέτη και εφαρμογή κρίνεται αντιοικονομική.

2) Η κατανόηση του ήδη υπάρχοντος διαχειριστικού συστήματος είναι ένας σημαντικός παράγοντας. Πάνω σε αυτό το σύστημα θα στηριχτεί η όποια εφαρμογή αποθήκης δεδομένων. Έτσι, πρώτος στόχος είναι να διασαφηνιστεί η ακριβής προέλευση των δεδομένων και της δομής τους, ώστε να επιλεγθούν τα κατάλληλα εργαλεία για την εφαρμογή.

3) Η ανάγκη για μετακίνηση των δεδομένων είναι ένας άλλος παράγοντας που θα πρέπει να εξεταστεί. Πολλές φορές κρίνεται συμφέρουσα και οικονομική η παραμονή των δεδομένων στις αρχικές βάσεις δεδομένων. Τα κριτήρια για την μετακίνηση ή όχι των δεδομένων είναι:

- Η ποιότητα των δεδομένων
- Το μέγεθος και η χρηστικότητα τους
- Η δομή τους
- Η δυνατότητα να εφαρμοστούν στη νέα εφαρμογή
- Η ευκολία πρόσβασης
- Η αναγκαιότητά τους στο ήδη υπάρχον σύστημα

4) Ένα πολύ σημαντικό σημείο στη μελέτη και τον σχεδιασμό είναι και ο τελικός αποδέκτης των μετακινούμενων δεδομένων. Πριν καν μελετηθεί η μετακίνηση των δεδομένων, πρέπει να γίνει ουσιαστικά ο σχεδιασμός της δομής των δεδομένων που θα συμπεριληφθούν στην αποθήκη δεδομένων.

5) Για να γίνει περισσότερο κατανοητή η χρήση της αποθήκης δεδομένων πρέπει να εντάσσεται σε ένα πληροφοριακό σύστημα διοίκησης ή σε ένα σύστημα υποστήριξης αποφάσεων. Τα ήδη υπάρχοντα συστήματα είναι ένας παράγοντας που πρέπει να λαμβάνεται υπόψη κατά τον σχεδιασμό μιας data warehouse.

6) Ο βαθμός μετατροπής των δεδομένων, καθώς και τα εργαλεία που απαιτούνται για αυτό.

7) Η επιλογή των κατάλληλων εργαλείων (στην περίπτωση όπου δεν υπάρχει εφαρμογή ΣΥΑ) για την επεξεργασία των δεδομένων της αποθήκης δεδομένων. Τις περισσότερες των περιπτώσεων απαιτούνται διαφορετικά εργαλεία, ανάλογα σε τι είδους χρήστη απευθύνεται (για παράδειγμα σε έμπειρους χρήστες αποθηκών δεδομένων, βελτιωτές εφαρμογών, υψηλόβαθμα στελέχη της επιχείρησης κ.α.).

8) Καθορισμός των κανόνων χρήσης και λειτουργίας της αποθήκης δεδομένων. Καθορισμός των χρηστών με πρόσβαση μερική ή ολική σε αυτή.

Οι παραπάνω παράγοντες δεν είναι οι μοναδικοί για τον σχεδιασμό μιας data warehouse. Η ευελιξία που δίνεται στον σχεδιαστή μέσω των διαφόρων τεχνολογικών εφαρμογών που ποικίλουν πλέον στην αγορά, καθώς και η πληθώρα των λογισμικών που είναι διαθέσιμα για τον σχεδιασμό βάσεων δεδομένων, έχουν ως αποτέλεσμα την εμφάνιση πολλών και διαφορετικών αποθηκών δεδομένων.

Με βάση τους παραπάνω παράγοντες, διαπιστώνει κανείς ότι ο υπεύθυνος σχεδίασης μιας εφαρμογής αποθήκης δεδομένων έχει την διακριτική ικανότητα να επιλέξει και να διαμορφώσει τα δικά του χαρακτηριστικά για την εφαρμογή, αλλά πάντα με βάση τις απαιτήσεις της επιχείρησης. Έτσι, οι εφαρμογές των αποθηκών δεδομένων διακρίνονται σε δύο βασικές κατηγορίες:

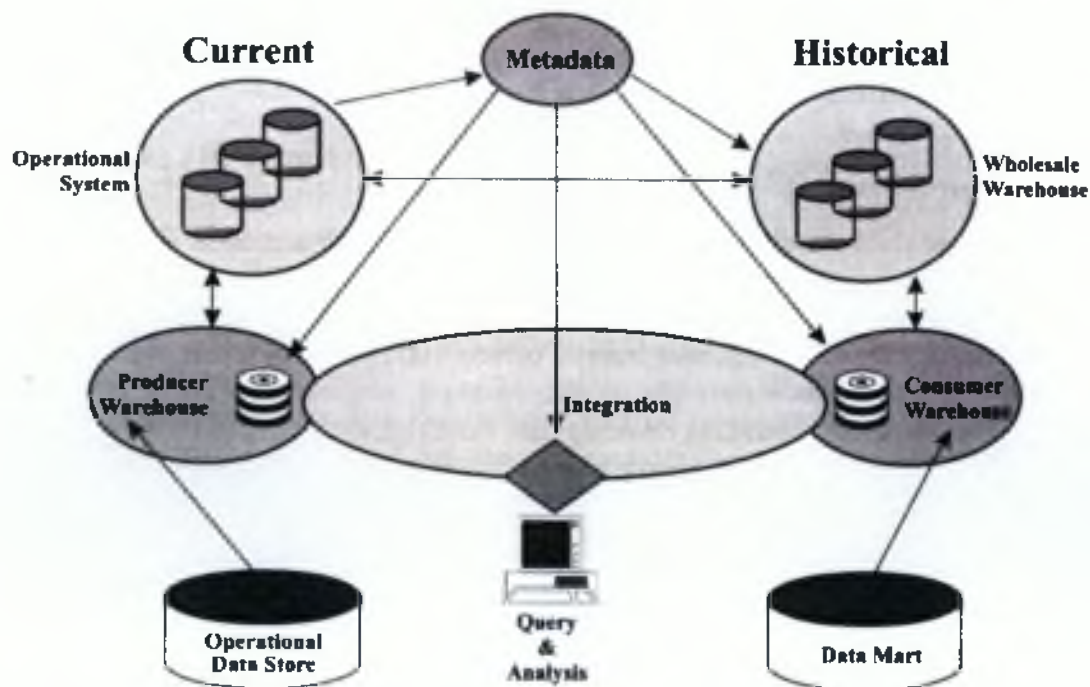
- **Host-based warehouses:** Η διαχείρισή τους γίνεται μέσω πληροφοριακών συστημάτων (έλεγχος μέσω IT-Information Technology) ή παραδοσιακών συστημάτων.

- **LAN-based Warehouses:** Συνεργάζονται με τοπικό δίκτυο, στηρίζουν δηλαδή λειτουργίες ενός μικρού συνόλου χρηστών και η διαχείρισή γίνεται μέσω αυτού ή μέσω του κεντρικού υπολογιστή.

Ένας ακόμα διαχωρισμός που αναφέρεται σε αποθήκες δεδομένων στηρίζεται στη τοποθεσία και την πηγή των δεδομένων της. Έτσι έχουμε:

- Διαχειριστικές αποθήκες δεδομένων (Operational data warehouses)
- Αποθήκες δεδομένων LAN δικτύου (LAN-based workgroup data warehouses)
- Αποθήκες πολλαπλών βαθμίδων (Multistage data warehouses)
- Μόνιμες αποθήκες δεδομένων (Stationary data warehouses)
- Κατανεμημένες αποθήκες δεδομένων (Distributed data warehouses)

Στο σχήμα 2-6 παρουσιάζεται το εύρος των εφαρμογών μιας αποθήκης δεδομένων και πως συνεργάζεται με τα διάφορα τμήματα μιας επιχείρησης. Θα πρέπει να παρατηρηθεί η ιδιαίτερη σημασία των αρχείων-δεδομένων metadata και ο ρόλος τους στο σύνολο της εφαρμογής.³



Σχήμα 2-6: Το σύνολο των εφαρμογών μιας αποθήκης δεδομένων

Πηγή: Συστήματα Υποστήριξης Αποφάσεων, Κεφ.11, Data Warehouse and OLAP, σελ. 240, <http://www.crgasya.tuc.gr/Users/Matsatsinis/Courses/SYA/sya.htm>

³ Στο ίδιο, σελ 230-240

2.7 Συνεργασία Αποθηκών Δεδομένων και OLAP

Οι πηγές των δεδομένων και των πληροφοριών μιας OLAP εφαρμογής μπορεί να είναι διαχειριστικές βάσεις δεδομένων, αποθήκες δεδομένων καθώς και πηγές έξω από την επιχείρηση. Η ανάγκη όμως για πολυδιάστατη ανάλυση αναδεικνύει τις αποθήκες δεδομένων (data warehouses) σαν τη κύρια πηγή άντλησης των πληροφοριών.

Οι αποθήκες δεδομένων περιέχουν συγκεντρωτικά και συνοπτικά, όλα τα απαραίτητα δεδομένα για τη δημιουργία των πολυδιάστατων χώρων που χρησιμοποιεί η απευθείας αναλυτική διαδικασία. Οι επιχειρήσεις με την εφαρμογή και εγκατάσταση μιας αποθήκης δεδομένων, ταυτόχρονα εφαρμόζουν τεχνικές OLAP οι οποίες συνεργάζονται άμεσα με αυτές. Άλλωστε, η μελέτη και εφαρμογή μιας αποθήκης δεδομένων (ή τις περισσότερες φορές μιας data mart) γίνεται για να υποστηρίξει μια απευθείας αναλυτική μέθοδο, η οποία δεν είναι δυνατό να συνεργαστεί απόλυτα με τις διαχειριστικές βάσεις δεδομένων.

Ο σχεδιασμός των αποθηκών δεδομένων (καθώς και των data marts) γίνεται τότε ώστε να επιτυγχάνεται η ταχύτερη και άμεση προσπέλαση στα δεδομένα από την OLAP εφαρμογή, ή ακόμα σχεδιάζεται για να μειώσει κατά κάποιο βαθμό την επεξεργασία των δεδομένων που απαιτούνται να έχουν για μια OLAP εφαρμογή.⁴

Οι Αποθήκες Δεδομένων αποτελούν μια συλλογή από τεχνολογίες στήριξης αποφάσεων (decision support) και μαζί με την Άμεση Αναλυτική Επεξεργασία των Δεδομένων παίζουν ουσιαστικά το ρόλο της κινητήριας δύναμης πίσω από την επιχειρησιακή νοημοσύνη (business intelligence). Ο όρος επιχειρησιακή νοημοσύνη περιγράφει το σύνολο των τεχνολογιών που παρέχουν στον τελικό χρήστη την δυνατότητα της αποκτήσεως πολύτιμης πληροφορίας από δεδομένα, τα οποία συγκεντρώνονται συνήθως από την καθημερινή λειτουργία μιας επιχείρησης και η οποία προάγει τη στρατηγική λήψη αποφάσεων.

Πρακτικά, μια Αποθήκη Δεδομένων αποτελεί ένα αντίγραφο των δεδομένων που προκύπτουν από τις συναλλαγές ενός οργανισμού και είναι δομημένη με τέτοιο τρόπο ώστε να εξυπηρετούνται πολύπλοκες ερωτήσεις (queries) και η ανάλυση των δεδομένων. Συνήθως, η Αποθήκη Δεδομένων αποτελεί ξεχωριστό τμήμα από τις επιχειρησιακές βάσεις δεδομένων (operational databases) ενός οργανισμού, καθώς ο στόχος της είναι η υποστήριξη Άμεσης Αναλυτικής Επεξεργασίας των Δεδομένων (OLAP). Το παραπάνω επιβάλλεται καθώς η υποστήριξη OLAP έχει εντελώς διαφορετικές λειτουργικές απαιτήσεις και απαιτήσεις απόδοσης σε σχέση με την Άμεση Επεξεργασία Δοσοληπιών (On-Line

⁴ Στο ίδιο, σελ 248

Transaction Processing - OLTP) που υποστηρίζουν οι επιχειρησιακές βάσεις δεδομένων.⁵ Τα OLTP συστήματα εστιάζουν σε συγκεκριμένες συχνές συναλλαγές με τη βάση δεδομένων, καλύπτοντας κυρίως διαδικασίες εισαγωγής νέων δεδομένων, μεταβολής/ενημέρωσης, διαγραφής και αναζήτησης των αποθηκευμένων δεδομένων. Τα συστήματα αυτά ικανοποιούν τις βασικές ανάγκες καθημερινής λειτουργίας του οργανισμού (σύστημα πωλήσεων, σύστημα παραγγελιών, σύστημα αποθήκης κ.ά.) αλλά δεν προσφέρονται για εύκολη ανάλυση των αποθηκευμένων πληροφοριών. Τα OLAP συστήματα σχεδιάστηκαν με σκοπό να καλύψουν το κενό αυτό και να προσφέρουν στα αρμόδια διοικητικά κλιμάκια τη δυνατότητα πρόσβασης, παρουσίασης και ανάλυσης των πληροφοριών του οργανισμού από πολλές διαφορετικές διαστάσεις, ανεξάρτητα από την πολυπλοκότητα της ερώτησης (query) προς τη βάση δεδομένων.⁶

Οι Αποθήκες Δεδομένων, σε αντίθεση, απευθύνονται στο πεδίο της προχωρημένης στήριξης αποφάσεων (decision support). Στις Αποθήκες Δεδομένων τα ιστορικά, αθροιστικά και ενοποιημένα δεδομένα είναι πιο σημαντικά από λεπτομερείς, ανεξάρτητες εγγραφές. Λέγοντας ιστορικά δεδομένα εννοούμε τη διατήρηση όλων των αλλαγών μιας εγγραφής και όχι μόνο της τελευταίας (έγκυρης) έκδοσής της, π.χ. η ποσότητα ενός προϊόντος που πωλήθηκε ανά μέρα για τις τελευταίες j μέρες και όχι μόνο για σήμερα. Η ανακάλυψη τάσεων στα δεδομένα ή η δυνατότητα πρόβλεψης απαιτούν την ύπαρξη ιστορικών δεδομένων, εφαρμογές στις οποίες οι επιχειρησιακές βάσεις δεδομένων αδυνατούν να αντεπεξέρθουν, καθώς αποθηκεύουν μόνο τα πιο πρόσφατα δεδομένα. Ενοποιημένα δεδομένα είναι τα δεδομένα που προκύπτουν από πολλές διαφορετικές πηγές, π.χ. από i διαφορετικούς πίνακες της ίδιας βάσης ή ακόμα και από τις βάσεις δεδομένων k διαφορετικών υποκαταστημάτων μίας επιχείρησης. Οι διαφορετικές πηγές είναι δυνατόν να περιέχουν δεδομένα διαφορετικής ποιότητας ή να χρησιμοποιούν διαφορετικούς συμβολισμούς, κωδικοποίηση και διάταξη (formats), οπότε τα εισερχόμενα δεδομένα καθαρίζονται και μετατρέπονται σε μια κοινή τελική μορφή. Τέλος, αθροιστικά δεδομένα προκύπτουν όταν παρατηρούμε τα δεδομένα σε διαφορετικά επίπεδα διακριτότητας σε σχέση με το πλέον λεπτομερές επίπεδο, δηλαδή τα δεδομένα που πρακτικά προκύπτουν από την

⁵ Ι. Ρούσσος, *Εργαστήριο Συστημάτων Βάσεων Γνώσεων και Δεδομένων*, Εθνικό Μετσόβιο Πολυτεχνείο, Αθήνα 2003, σελ 19-20

⁶ *Προς ένα Ολοκληρωμένο Πληροφοριακό Σύστημα Συλλογής, Αποθήκευσης και Επεξεργασίας Εκπαιδευτικών Δεδομένων: Από το Όραμα στο Σύστημα Edubase*, 25 Ιουλίου 2008, 20:10, σελ 617, <http://www.etpe.gr/files/proceedings/uploads/p113.pdf>

εφαρμογή αθροιστικών τελεστών στα δεδομένα, π.χ. ο αριθμός κατοίκων ανά πόλη ή ανά χώρα ή ανά ήπειρο.

Καθώς οι Αποθήκες Δεδομένων περιέχουν ενοποιημένα δεδομένα, πιθανώς από αρκετές επιχειρησιακές βάσεις δεδομένων και συχνά καλύπτοντας διάρκεια μεγάλων χρονικών περιόδων, τείνουν να είναι κάποιες τάξεις μεγέθους μεγαλύτερες από τις επιχειρησιακές βάσεις δεδομένων, φτάνοντας σε μέγεθος εκατοντάδων gigabytes έως και μερικών terabytes. Επίσης, οι απαιτήσεις και τα χαρακτηριστικά ενός Συστήματος Διαχείρισης Βάσεων Δεδομένων (ΣΔΒΔ) (Database Management System – DBMS) που χρησιμοποιείται για OLAP, και το οποίο βασίζεται κυρίως σε Αποθήκες Δεδομένων, είναι σημαντικά διαφορετικές απ’ αυτές ενός συστήματος για εφαρμογές OLTP. Το πρώτο χρησιμοποιείται σε ένα περιβάλλον «κυρίως ανάγνωσης» (read-mostly), όπου οι ενημερώσεις λαμβάνουν χώρα περιοδικά και με μαζικό τρόπο (batch mode), αντί για αυθαιρέτως και με μεγάλη συχνότητα. Το γεγονός αυτό αλλάζει ριζικά την εστίαση της οργάνωσης των δεδομένων, επιτρέποντας την χρήση τεχνικών, όπως ο πλεονασμός των δεδομένων μέσω αποθήκευσης μη κανονικοποιημένων πινάκων ή η φυσική συγκέντρωση (clustering) των δεδομένων, ώστε να μειωθεί ο χρόνος απόκρισης των επερωτήσεων και οι οποίες δεν θα χρησιμοποιούνταν ποτέ σε ένα σύστημα OLTP, αφού θα επιδείωναν τον χρόνο εκτέλεσης των ενημερώσεων. Επιπλέον, ο έλεγχος συγχρονισμού (concurrency control) δεν είναι τόσο σημαντικός στα συστήματα OLAP, όσο είναι στα συστήματα OLTP, αφού τα πρώτα στοχεύουν πρωτίστως σε μια μικρή μόνο ομάδα στρατηγικών (για την επιχείρηση) χρηστών. Στα συστήματα OLAP, η απόδοση (throughput) και ο χρόνος απόκρισης των queries είναι πιο σημαντικά από το throughput των δοσοληψιών.

Συγχρόνως, η ανάγκη για μια προχωρημένη ανάλυση των δεδομένων επέβαλλε στις επερωτήσεις OLAP να είναι πιο σύνθετες και να συνεπάγονται την επεξεργασία μεγάλου όγκου δεδομένων. Έτσι, το μεγαλύτερο ποσοστό του φόρτου εργασίας ενός DBMS που χρησιμοποιείται για OLAP οφείλεται κυρίως σε ad hoc (κάθε αυθόρμητη ή μη προσχεδιασμένη ερώτηση) σύνθετες επερωτήσεις στα δεδομένα, οι οποίες είναι δυνατόν να χρειάζεται να προσπελάσουν εκατομμύρια εγγραφές και να κάνουν μεγάλο αριθμό από συζεύξεις (joins), σαρώσεις (scans) στους πίνακες και συναθροίσεις. Αυτή η μορφή των ad hoc OLAP επερωτήσεων, για τις οποίες δεν υπάρχει καμία εκ των προτέρων γνώση (a priori), είναι απαραίτητη για την προχωρημένη στήριξη αποφάσεων, καθώς η ανάλυση των δεδομένων και η εύρεση νέων συσχετίσεων σε αυτά δεν είναι δυνατόν να ακολουθεί προκαθορισμένους κανόνες ούτε οι επερωτήσεις να προέρχονται από ένα προκαθορισμένο πεπερασμένο σύνολο, όπως γίνεται στις OLTP εφαρμογές.

Τα παραπάνω καθιστούν τις *ad hoc* επερωτήσεις τις πλέον σημαντικές για ένα σύστημα OLAP, με αποτέλεσμα η απόδοση μιας Αποθήκης Δεδομένων ως προς αυτές να είναι κρίσιμη για τη συνολική απόδοση όλου του συστήματος. Αυτού του τύπου οι επερωτήσεις συνδυάζουν πολλαπλά χαρακτηριστικά μιας επιχείρησης (ή του όποιου άλλου συστήματος περιγράφεται μέσω της Αποθήκης Δεδομένων) με στόχο την παροχή ενός ενοποιημένου αποτελέσματος ως προς κάποιες μετρήσιμες τιμές. Αυτό οδηγεί σε μια πολυδιάστατη (multidimensional) μοντελοποίηση των δεδομένων, όπου ένας αριθμός από Διαστάσεις (dimensions) χαρακτηρίζει ορισμένες μετρήσιμες τιμές (measures). Παραδείγματος χάριν, στην Αποθήκη Δεδομένων με τις πωλήσεις μιας αλυσίδας πολυκαταστημάτων, ο χρόνος πώλησης, η διεύθυνση του συγκεκριμένου καταστήματος, το όνομα του πωλητή ή ο αριθμός του ταμείου θα μπορούσαν να είναι ορισμένες από τις Διαστάσεις μέσω των οποίων χαρακτηρίζεται μοναδικά η συγκεκριμένη καταχώρηση, της οποίας μετρήσιμη τιμή αποτελεί το χρηματικό ποσό της συγκεκριμένης αγοραπωλησίας. Σημαντική λεπτομέρεια αποτελεί το γεγονός ότι το ποιά μεγέθη είναι Διαστάσεις και το ποιά είναι μετρήσιμες τιμές εξαρτάται μόνο από τις ιδιαιτερότητες του κάθε οργανισμού και της μορφής της ανάλυσης που θέλει να πραγματοποιήσει. Οι μετρήσιμες τιμές αποτελούν την “καθαρή” πληροφορία του συστήματος, αλλά οι συσχετίσεις ή ομαδοποιήσεις πραγματοποιούνται μέσω των Διαστάσεων, μέσω των οποίων γίνεται η τοποθέτηση της πληροφορίας σε κάποια περιοχή του πολυδιάστατου χώρου.

Η ανάγκη για αποτελέσματα σε διαφορετική διακριτότητα (granularity) οδήγησε στην οργάνωση των Διαστάσεων σε ιεραρχίες διαφόρων συναθροιστικών επιπέδων (aggregation levels). Παραδείγματος χάριν, για τη διάσταση τοποθεσία πώλησης μια τέτοια ιεραρχία επιπέδων θα μπορούσε να ήταν η ακόλουθη: {κατάστημα, περιοχή, πόλη, χώρα, ήπειρος}, με το πιο λεπτομερές επίπεδο να είναι το ‘κατάστημα’ και το επίπεδο με τη μικρότερη διακριτότητα το ‘ήπειρος’, και ένα συγκεκριμένο κατάστημα να είναι το: {κατάστημα17, Άνω Πατήσια, Αθήνα, Ελλάδα, Ευρώπη}. Έτσι, οι ιεραρχίες επιτρέπουν ερωτήσεις σε διαφορετικά επίπεδα διακριτότητας όπως: “δείξε το σύνολο των πωλήσεων ανά χρόνο ανά κατάστημα ” ή “ δείξε το σύνολο των πωλήσεων ανά μήνα ανά χώρα”. Οι ιεραρχίες των Διαστάσεων παίζουν κυρίαρχο ρόλο στα φορτία επερωτήσεων OLAP, εφ’ όσον οι πιο τυπικές επερωτήσεις περιέχουν περιορισμούς πάνω στις ιεραρχίες σε συνδυασμό με ομαδοποίηση (grouping) και συνάθροιση (aggregation) πάνω στα διάφορα επίπεδα των ιεραρχιών. Τυπικές λειτουργίες (operations) OLAP ως προς μία ή περισσότερες ιεραρχίες Διαστάσεων είναι οι rollup (ελάττωση της διακριτότητας ή επιπλέον εφαρμογή συναθροιστικών τελεστών στα δεδομένα) και drill-down (αύξηση της διακριτότητας).

Τέλος, χρησιμοποιώντας πάντα την προαναφερθείσα πολυδιάστατη μοντελοποίηση των δεδομένων στο λογικό επίπεδο, οι Αποθήκες Δεδομένων μπορούν να υλοποιηθούν είτε χρησιμοποιώντας κλασικά ή τροποποιημένα σχεσιακά Συστήματα Διαχείρισης Βάσεων Δεδομένων, οπότε ονομάζονται σχεσιακά συστήματα OLAP (Relational OLAP - ROLAP), είτε χρησιμοποιώντας ειδικά κατασκευασμένα πολυδιάστατα συστήματα OLAP (Multidimensional OLAP - MOLAP). Τα συστήματα ROLAP θεωρούν ότι τα δεδομένα αποθηκεύονται σε σχεσιακές Βάσεις Δεδομένων και υποστηρίζουν επεκτάσεις της SQL και ειδικές μεθόδους για την πρόσβαση στις λειτουργίες του πολυδιάστατου μοντέλου δεδομένων. Από την άλλη, τα συστήματα MOLAP αποθηκεύουν απ' ευθείας τα πολυδιάστατα δεδομένα σε ειδικές πολυδιάστατες δομές (όπως πολυδιάστατοι πίνακες) και υποστηρίζουν τις λειτουργίες του πολυδιάστατου μοντέλου δεδομένων άμεσα, μέσω πράξεων σε φυσικό επίπεδο στις δομές αυτές.⁷

⁷ Ι. Ρούσσοις, *Εργαστήριο Συστημάτων Βάσεων Γνώσεων και Δεδομένων*, Εθνικό Μετσόβιο Πολυτεχνείο, Αθήνα 2003, σελ 20-23

ΚΕΦΑΛΑΙΟ 3

ΑΝΟΙΧΤΟΣ ΚΩΔΙΚΑΣ (OPEN SOURCE)

3.1 Η έννοια του ανοιχτού κώδικα

Ως Ανοιχτός Κώδικας (Open Source) αναφερόμαστε στο μοντέλο διάθεσης λογισμικού, όπου ο πηγαίος κώδικας είναι διαθέσιμος σε όποιον ενδιαφέρεται για αυτόν. Οι όροι διανομής του λογισμικού αυτού πρέπει να πληρούν συγκεκριμένα κριτήρια, τα βασικότερα από τα οποία είναι η ελεύθερη χρήση, αντιγραφή/αναδιανομή και μεταβολή/βελτίωσή του (ώστε να εξυπηρετεί τις ανάγκες του ατόμου/συνόλου που τον μεταβάλλει). Οι όροι αυτοί περιγράφονται στην άδεια χρήσης του λογισμικού. Υπάρχουν διάφορες άδειες χρήσης, με πιο διαδεδομένες τις BSD License και MIT License. Η άδεια χρήσης BSD, επιτρέπει τις αλλαγές χωρίς επαναδιανομή του πηγαίου κώδικα.¹

Γύρω από αυτή τη λογική δημιουργήθηκε μια τεράστια κοινότητα χρηστών και προγραμματιστών, με βάση το Διαδίκτυο, οι οποίοι συμβάλλουν από κοινού στη συνεχή βελτίωση του λογισμικού, παρέχοντας δωρεάν τις γνώσεις και τη δουλειά τους σε όλους. Ένα ανοικτό δίκτυο "εθελοντών" προγραμματιστών και εταιρειών του κλάδου, οι οποίοι αναπτύσσουν και διορθώνουν τον κώδικα των προγραμμάτων παράλληλα, κυκλοφορώντας ταχύτατα και σε μεγάλη συχνότητα τις νέες εκδόσεις Ελεύθερου Λογισμικού/Λογισμικού Ανοιχτού Κώδικα. Το Διαδίκτυο αποτελεί τη βασική πρόσβαση στο διαθέσιμο ΕΛ/ΛΑΚ. Η σχέση τους, εξ αρχής, υπήρξε στενή και αμφίδρομη. Η εξάπλωση του ΕΛ/ΛΑΚ έχει στηριχθεί στην ευρεία χρήση του Διαδικτύου και η διαδικασία ανάπτυξης του Διαδικτύου βασίζεται, κατά κύριο λόγο, στο ΕΛ/ΛΑΚ. Το Λογισμικό Ανοιχτού Κώδικα, με συνεχείς βελτιώσεις και αυξημένη πλέον φιλικότητα προς το χρήστη, κερδίζει διαρκώς νέους φίλους παγκοσμίως. Στην εκπαίδευση, στη δημόσια διοίκηση και στις επιχειρήσεις, ενδιαφέρονται, ενημερώνονται και αποκτούν ένα ιδιαίτερα ελκυστικό εργαλείο, αξιόπιστο, σταθερό στη λειτουργία και απαλλαγμένο από σημαντικά κόστη, απόκτησης και αναβαθμίσεων.

¹ *Ελεύθερο Λογισμικό/Λογισμικό Ανοιχτού Κώδικα*, 10 Αυγούστου, 14:22, <http://www.opensource.gr/faq.php?language=el>

Οι βασικότεροι στόχοι του ΕΛ/ΛΑΚ είναι:

- Η παραγωγή καλύτερης ποιότητας λογισμικού
- Η συλλογική βελτίωση της τεχνογνωσίας με τη διάθεση του πηγαίου κώδικα προς μελέτη
- Η ανεξαρτησία του χρήστη από τις εταιρείες παραγωγής λογισμικού, καθώς η κατοχή του πηγαίου κώδικα διευκολύνει τη μετάβαση από μία μορφή διατήρησης δεδομένων σε μία άλλη

Για την επίτευξή τους:

- Προβάλλονται πρωτοβουλίες σχετικές με το ΕΛ/ΛΑΚ και παράλληλα αναλαμβάνονται καινούριες
- Γίνεται προσπάθεια για την ενημέρωση όσων ενδιαφέρονται, με νέα από την Ελλάδα και τον κόσμο, εκδηλώσεις και πληροφορίες σχετικά με το ΕΛ/ΛΑΚ²

3.2 Πλεονεκτήματα και Μειονεκτήματα

Τα κυριότερα πλεονεκτήματα του ΕΛ/ΛΑΚ είναι η ποιότητα και η αξιοπιστία, επειδή η φιλοσοφία του επιτρέπει αφενός μεν τον ποιοτικό έλεγχο του από πολλούς ανθρώπους που δεν έχουν κάποια σχέση μεταξύ τους, αφετέρου δε την άμεση εξέλιξη του πηγαίου κώδικα. Ωστόσο, αυτό δεν ισχύει στον ίδιο βαθμό σε όλες τις περιπτώσεις.

Για παράδειγμα, το ΕΛ/ΛΑΚ υποδομής (δηλαδή αυτό που υλοποιεί λειτουργικά συστήματα, όπως GNU/Linux, και διαφόρων τύπων δικτυακές υπηρεσίες, όπως τον web server Apache) έχει αποδειχθεί πάρα πολύ σταθερό. Αντίθετα, ΕΛ/ΛΑΚ εφαρμογές που απευθύνονται στον τελικό χρήστη είναι πολλές φορές πιο ασταθείς από τις αντίστοιχες proprietary. Ένα άλλο πλεονέκτημα είναι το γεγονός ότι η χρήση ΕΛ/ΛΑΚ δεν δημιουργεί εξαρτήσεις από κάποια συγκεκριμένη εταιρεία. Τέλος, όσοι έχουν τις τεχνικές γνώσεις και την διάθεση, έχουν τη δυνατότητα να μελετήσουν τον πηγαίο κώδικα, να μάθουν πως λειτουργεί και ενδεχομένως να τον τροποποιήσουν ώστε να ταιριάζει περισσότερο στις ανάγκες τους.

Τα κυριότερα μειονεκτήματα του ΕΛ/ΛΑΚ είναι ότι το επίπεδο της τεκμηρίωσης και της υποστήριξης δεν είναι πάντοτε αποδεκτό. Σε αντιστοιχία με τα προαναφερθέντα πλεονεκτήματα της ποιότητας και αξιοπιστίας, τα open-source προϊόντα λογισμικού που προσφέρουν υψηλής ποιότητας τεκμηρίωση και υποστήριξη είναι, κατά κανόνα, μόνο εκείνα

² *Ελεύθερο Λογισμικό/Λογισμικό Ανοικτού Κώδικα*, 10 Αυγούστου 2008, 14:22, <http://www.open-source.gr/whatis.php?language=el>

που υλοποιούν λειτουργικά συστήματα και διάφορες δικτυακές υπηρεσίες (GNU/Linux, Apache κτλ.). Επιπλέον, ένα πρόβλημα που συναντιέται συχνά είναι ότι κάποιο λογισμικό αν και μπορεί να στηρίζεται σε μία πολύ καλή ιδέα, συνήθως δεν είναι αρκετά καλό στις πρώτες εκδόσεις του και δεν βελτιώνεται αισθητά έως ότου γίνει αρκετά δημοφιλές ώστε να διατεθούν αρκετοί πόροι για την ανάπτυξή του.³

3.3 Χρήση Ανοιχτού Κώδικα και Ελεύθερου Λογισμικού

Όταν λέμε το λογισμικό “ελεύθερο”, εννοούμε ότι σέβεται τις βασικές ελευθερίες του χρήστη που αναφέραμε πιο πάνω. Αυτές οι ελευθερίες είναι ζωτικής σημασίας. Είναι απαραίτητες, όχι μόνο για το καλό του χρήστη, αλλά επειδή προωθούν κοινωνική αλληλεγγύη που είναι ο διαμοιρασμός και η συνεργασία. Γίνονται ακόμα πιο σημαντικές, όσο και πιο πολύ από την κουλτούρα και τις ζωτικές μας ενέργειες, όσο αυτές ψηφιοποιούνται. Σε έναν κόσμο ψηφιακών ήχων, εικόνων και λέξεων, το ελεύθερο λογισμικό αυξάνεται ώστε να εξισοροποιηθεί με την ελευθερία γενικότερα.

Δεκάδες εκατομμυρίων ανθρώπων σε όλο τον κόσμο πλέον χρησιμοποιούν ελεύθερο λογισμικό. Τα σχολεία σε περιοχές στην Ινδία και την Ισπανία πλέον διδάσκουν όλους τους μαθητές να χρησιμοποιούν το ελεύθερο λειτουργικό σύστημα GNU/Linux. Αλλά οι περισσότεροι από αυτούς τους χρήστες δεν έχουν ακούσει ποτέ τους ηθικούς λόγους για τους οποίους αναπτύχθηκε αυτό το σύστημα και χτίστηκε η κοινότητα του ελεύθερου λογισμικού, επειδή σήμερα αυτό το σύστημα και αυτή η κοινότητα πιο συχνά περιγράφονται σαν “ανοιχτός κώδικας” και απευθύνονται σε μία διαφορετική φιλοσοφία στην οποία οι ελευθερίες δύσκολα αναφέρονται.

Το κίνημα του ελεύθερου λογισμικού δραστηριοποιείται για την ελευθερία των χρηστών-υπολογιστών από το 1983. Το 1984 ξεκίνησε την ανάπτυξη του ελεύθερου λογισμικού GNU, ώστε να αποφευχθούν τα μη ελεύθερα λειτουργικά συστήματα τα οποία αρνούνται την ελευθερία στους χρήστες τους. Κατά την διάρκεια της δεκαετίας του 80, αναπτύχθηκαν τα περισσότερα από τα απαραίτητα περιεχόμενα ενός τέτοιου συστήματος, καθώς επίσης και η Γενική Άδεια Δημόσιας Χρήσης GNU, μία άδεια σχεδιασμένη ειδικά για να προστατεύσει την ελευθερία για όλους τους χρήστες ενός προγράμματος.

Παρ’ όλα αυτά, δεν συμφώνησαν όλοι οι χρήστες και οι δημιουργοί του ελεύθερου λογισμικού με αυτούς τους στόχους του κινήματος του ελεύθερου λογισμικού. Το 1998, ένα μέρος της κοινότητας του ελεύθερου λογισμικού διασπάστηκε και ξεκίνησε να

³ *ΕΛΛΑΚ*, 10 Αυγούστου 2008, 14:26, <http://www.open-source.gr/faq.php?ctn=1&pos=10&language=e>

δραστηριοποιείται με το όνομα “ανοιχτός κώδικας.” Αυτός ο όρος, αρχικά, προτάθηκε για να αποφευχθούν πιθανές παρεξηγήσεις του όρου “ελεύθερο λογισμικό”, αλλά σύντομα συσχετίστηκε με φιλοσοφικές απόψεις αρκετά διαφορετικές από εκείνους του κινήματος του ελεύθερου λογισμικού.

Κάποιοι από τους υποστηρικτές του “ανοιχτού κώδικα” το θεώρησαν σαν μία “εμπορική δραστηριοποίηση για το ελεύθερο λογισμικό”, το οποίο θα άρесе σε επιχειρηματικά στελέχη αναφέροντας πρακτικά οφέλη, ενώ παράλληλα θα απέφευγαν ιδέες σωστού και λάθους τις οποίες δεν θα ήθελαν να ακούσουν. Άλλοι υποστηρικτές, απέρριψαν κατηγορηματικά τις ηθικές και κοινωνικές αξίες του κινήματος του ελεύθερου λογισμικού. Όποιες κι αν ήταν οι απόψεις τους, όταν δραστηριοποιούνταν για τον “ανοιχτό κώδικα”, δεν ανέφεραν ή συνηγόρησαν υπέρ εκείνων των αξιών. Ο όρος “ανοιχτός κώδικας” γρήγορα συσχετίστηκε με την πράξη τους να αναφέρουν μόνο πρακτικές αξίες, όπως το να κάνουν ισχυρό κι αξιόπιστο λογισμικό. Οι περισσότεροι υποστηρικτές του “ανοιχτού κώδικα” έχουν εισαχθεί σε αυτό από τότε και αυτή την πρακτική την δέχονται όπως είναι.

Σχεδόν όλο το λογισμικό ανοιχτού κώδικα είναι ελεύθερο λογισμικό. Οι δύο όροι περιγράφουν σχεδόν την ίδια κατηγορία λογισμικού, αλλά αντιπροσωπεύουν κάποιες απόψεις βασισμένες σε βασικές διαφορετικές αξίες. Ο ανοιχτός κώδικας είναι μία μεθοδολογία ανάπτυξης. Το ελεύθερο λογισμικό είναι ένα κοινωνικό κίνημα. Για το κίνημα ελεύθερου λογισμικού, το ελεύθερο λογισμικό είναι μία ηθική προσταγή, επειδή μόνο το ελεύθερο λογισμικό σέβεται την ελευθερία του χρήστη. Σε αντίθεση, η φιλοσοφία του ανοιχτού κώδικα σχετίζεται μόνο με θέματα για το πως να κάνει το λογισμικό “καλύτερο” με καθαρά πρακτικό τρόπο. Λέγεται πως το μη ελεύθερο λογισμικό είναι μία χειριστή λύση.

Έχει μεγάλη σημασία για το πιο όνομα θα χρησιμοποιήσουμε, αν θα το πούμε δηλαδή Ελεύθερο λογισμικό ή Ανοιχτός κώδικας. Αυτό συμβαίνει, επειδή διαφορετικές λέξεις εκφράζουν διαφορετικές ιδέες. Ένα ελεύθερο πρόγραμμα με οποιοδήποτε άλλο όνομα θα έδινε την ίδια ελευθερία σήμερα, καθιδρύνοντας την ελευθερία με έναν τρόπο που θα κρατήσει και πάνω από όλα να διδάξει στους ανθρώπους την αξία την ελευθερίας.⁴

⁴ Γιατί “Ο Ανοιχτός Κώδικας” χάνει το νόημα του ελεύθερου λογισμικού, 10 Αυγούστου 2008, 15:00, <http://www.gnu.org/philosophy/open-source-misses-the-point.el.html>

3.4 Διαφορά Ανοιχτού Κώδικα και Ελεύθερου Λογισμικού

Καταρχήν πρέπει να διευκρινιστεί ότι οι δύο όροι αναφέρονται σε αντίστοιχες ιδεολογίες, κατ' ουσίαν στο ίδιο πράγμα. Τόσο η μία ιδεολογία όσο και η άλλη, προσπαθούν να προωθήσουν την κοινή γνώση και να δημιουργήσουν τελικά καλύτερο λογισμικό. Η κύρια διαφορά τους είναι ότι ενώ το Ελεύθερο Λογισμικό δίνει έμφαση στην ελευθερία (για μάθηση, προσαρμογή, βελτίωση), το Λογισμικό Ανοιχτού Κώδικα δίνει έμφαση στη δημιουργία καλύτερου λογισμικού, μέσω της συνεργασίας των προγραμματιστών και των μηχανικών όλου του κόσμου. Συνεπώς, το ΛΑΚ είναι λιγότερο αυστηρό και πιο φιλικό προς τις εταιρίες που επιθυμούν να αξιοποιήσουν αλγορίθμους που υπάρχουν σε έργα τύπου ΛΑΚ.

Οι όροι του Α/Κ και Ε/Λ είναι σαφώς συσχετισμένοι, αν και μπορούν να γίνουν πολλές φιλοσοφικού τύπου συζητήσεις γύρω από το θέμα. Δεν πρέπει κάποιος να επηρεάζεται από τις προφανείς (προερχόμενες από απλή μετάφραση των αγγλικών όρων) και κοινώς διαδεδομένες ερμηνείες των παραπάνω όρων, καθώς αυτές είναι σε μεγάλο βαθμό λανθασμένες. Συγκεκριμένα, δεν είναι σωστό το να ονομάσει κάποιος ως open source κάθε λογισμικό στο οποίο είναι διαθέσιμος ο πηγαίος κώδικας (αν και όταν το λογισμικό είναι open source ο πηγαίος κώδικας είναι οπωσδήποτε διαθέσιμος), καθώς υπάρχουν περιπτώσεις λογισμικού semi-free ή και proprietary στο οποίο είναι επίσης διαθέσιμος ο πηγαίος κώδικας. Αντίστοιχα, free δεν ονομάζεται απαραίτητα το δωρεάν λογισμικό (αν και στις περισσότερες περιπτώσεις το free λογισμικό διατίθεται δωρεάν), αλλά το λογισμικό εκείνο για το οποίο υπάρχει ελευθερία χρήσης, μεταβολής και επαναδιανομής του, που παρέχει η άδεια χρήσης του λογισμικού και όχι στο κόστος απόκτησης του. Τυπικό παράδειγμα είναι το open-source λειτουργικό σύστημα GNU/Linux, που μπορεί κάποιος να κατεβάσει και δωρεάν από το internet αλλά και να αγοράσει μέσω κάποιας διανομής.

Η ανάπτυξη λογισμικού με τη φιλοσοφία του open source δεν αποσκοπεί στο κέρδος (χωρίς αυτό να σημαίνει πως οι δύο έννοιες συγκρούονται). Σκοπός της είναι αφενός μεν η δημιουργία όσο το δυνατόν καλύτερου λογισμικού μέσα από τη συνεργασία της κοινότητας για την βελτίωση του κώδικα, αφετέρου δε η βελτίωση της κοινής διαθέσιμης γνώσης μέσα από τη διαδικασία μελέτης του πηγαίου κώδικα. Το κέρδος, εφόσον, είναι επιθυμητό μπορεί να προκύψει τόσο από την πώληση του λογισμικού σε binary form (έτοιμου προς χρήση δηλαδή) μαζί με τον πηγαίο κώδικα, όσο και από την πώληση υπηρεσιών που σχετίζονται με το λογισμικό αυτό.

Η επιτυχία ενός προϊόντος λογισμικού, εάν αυτό δοθεί στην κοινότητα με άδεια τύπου ΕΛ/ΛΑΚ, δεν είναι πάντα σίγουρη. Σημαντικό ρόλο παίζει το είδος του λογισμικού, καθώς αυτό καθορίζει αφενός μεν το μέγεθος του κοινού στο οποίο απευθύνεται και επομένως τον

αριθμό των προγραμματιστών που επιθυμούν να ασχοληθούν με την ανάπτυξή του, αφετέρου δε το ποσοστό αυτών που είναι ικανοί να φέρουν σε πέρας τέτοιο έργο. Για παράδειγμα, ένα λειτουργικό σύστημα ή μία ευρέως χρησιμοποιούμενη δικτυακή υπηρεσία απευθύνονται σε μεγάλη μερίδα χρηστών και προσελκύουν τους πόρους που είναι απαραίτητοι για την ανάπτυξη τους. Αντίθετα, μία πολύ εξειδικευμένη εφαρμογή (π.χ. πολύ μικρού μεγέθους εφαρμογή επεξεργασίας ήχου και εικόνας) απευθύνεται σε περιορισμένο κοινό και υπάρχουν πολλοί λίγοι προγραμματιστές στον κόσμο που να ενδιαφέρονται (και ακόμη λιγότεροι που να μπορούν) να ασχοληθούν με την ανάπτυξή της. Επιπλέον, μία εταιρεία που αναπτύσσει τέτοιου είδους εφαρμογές συνήθως επενδύει χρόνο και προσωπικό για την ανάπτυξη νέων αλγορίθμων που να επιλύουν το εκάστοτε πρόβλημα και αναμένει κάποια απόδοση στην επένδυση που έκανε. Το είδος του λογισμικού, όμως, από μόνο του δεν αρκεί. Η επιτυχία και η διάδοση μιας τέτοιας εργασίας εξαρτάται από την ανάγκη για το εν λόγω προϊόν λογισμικού, την ποιότητα αυτού όταν παραδίδεται στην κοινότητα η πρώτη έκδοση, από την οργάνωση της διαδικασίας ανάπτυξής του ώστε να υποστηρίξει με οργανωτικές δομές τους προγραμματιστές που ενδιαφέρονται, καθώς και από παράγοντες όπως η «διαφήμιση» σε σχετικά sites, η πρωτοπορία της χρησιμοποιούμενης τεχνολογίας κτλ.

Στη συντριπτική πλειοψηφία των περιπτώσεων, ένα open-source έργο ξεκινάει επειδή ο ίδιος ο developer είχε τη ανάγκη για να δημιουργήσει κάποιο εργαλείο. Η λογική είναι ότι εφόσον δεν υπάρχει χρηματικό κέρδος, η ανάπτυξη γίνεται με βάση τις δικές του ανάγκες και προτιμήσεις. Δεν υπάρχει νόημα στην επιλογή ενός έργου. Στην ερώτηση γιατί κατασκευάστηκε, συνήθως η απάντηση είναι είτε γιατί υπήρχε η ανάγκη, είτε γιατί απλώς μπορούσε να γίνει.

Τέλος, ο αριθμός των χρηστών που ενδέχεται να χρησιμοποιήσουν το λογισμικό εξαρτάται καθαρά από το έργο. Για παράδειγμα, ένα μεγάλο έργο περιλαμβάνει εκατοντάδες χιλιάδες χρήστες, με αποτέλεσμα η δουλειά να είναι πολύ μεγάλου μεγέθους με εκατοντάδες ενεργούς developers. Παρ' όλα αυτά πολλά μεγάλα έργα έχουν αποτύχει, ενώ άλλα πολύ μικρότερα έχουν ιδιαίτερα μεγάλη απήχηση. Συνήθως οι μικρές εργασίες που χρησιμοποιούνται ευρέως είναι σε επίπεδο συστήματος (π.χ. βιβλιοθήκες κρυπτογράφησης, επεξεργασίας εικόνας κλπ) και αξιοποιούνται από μεγάλες εφαρμογές. Σε γενικές γραμμές το κοινό χρησιμοποιεί λογισμικό το οποίο έχει ανάγκη και απαιτεί από αυτό σταθερότητα στη λειτουργία.⁵

⁵ *Ελεύθερο Λογισμικό/Λογισμικό Ανοικτού Κώδικα*, 10 Αυγούστου 2008, 14:22, <http://www.open-source.gr/faq.php?language=el>

ΚΕΦΑΛΑΙΟ 4

ΕΡΓΑΛΕΙΑ OLAP

4.1 Λειτουργία εργαλείων OLAP

Τα εργαλεία των OLAP εφαρμογών υποστηρίζουν τρεις εργασίες των χρηστών: 1) την αναζήτηση και την παρουσίαση γνωστών δεδομένων-πληροφοριών, 2) την ανάλυση των γνωστών δεδομένων και πληροφοριών και 3) την ανακάλυψη αγνώστων πληροφοριών.

Οι εργασίες αυτές αποσκοπούν στην οπτική προβολή των στοιχείων που είναι καταχωρημένα ή εξάγονται από το σύστημα μέσα από γραφικές απεικονίσεις (Graphical User Interfaces GUIs) και λίστες αναφορών. Η διαδικασία γίνεται απευθείας (on-line) και η απόκριση του συστήματος είναι γρήγορη και αποτελεσματική.

1) Κατά την αναζήτηση, την παρουσίαση και την ανάλυση των γνωστών δεδομένων και πληροφοριών, ο χρήστης γνωρίζει το είδος και την ποιότητα των εξαγόμενων συμπερασμάτων και αποτελεσμάτων. Η διαδικασία είναι σύντομη και αποτελεσματική αφού στηρίζεται σε γνωστά στοιχεία και δεδομένα.

2) Η εξερεύνηση δεδομένων αφορά την εύρεση αγνώστων έως τώρα δεδομένων και πληροφοριών που περιέχονται όμως μέσα στον όγκο των καταχωρήσεων. Πολλές OLAP εφαρμογές περικλείουν τεχνικές data mining κάνοντας την εύρεση τέτοιων δεδομένων πιο εύκολη.¹ Ανακάλυψη ή εξόρυξη γνώσης (data mining) είναι η διαδικασία οργανωμένης διερεύνησης των διαθέσιμων στον οργανισμό δεδομένων, για ανεύρεση νέων πληροφοριών (όπως συσχετισμούς στοιχείων, κανονικότητες, ομαδοποιήσεις, πρότυπα (patterns) δεδομένων, πιθανούς ενγενείς κανόνες, μοντέλα περιγραφής και πρόβλεψης κ.ά.) που ήταν κρυμμένες στον μεγάλο όγκο δεδομένων. Οι πληροφορίες αυτές, αφού περάσουν από το στάδιο της πιστοποίησης της ακρίβειας και της αξιοπιστίας τους (με κατάλληλη ανάλυση στο ίδιο ενιαίο περιβάλλον του εργαλείου) και εφόσον τα δεδομένα στα οποία στηρίχθηκαν αντικατοπτρίζουν επακριβώς στοιχεία της πραγματικής λειτουργίας, μπορούν να αξιοποιηθούν στις διαδικασίες λήψης αποφάσεων, τόσο για την επιλογή άμεσων ενεργειών όσο και για το μεσοπρόθεσμο και μακροπρόθεσμο σχεδιασμό του οργανισμού.²

¹ *Συστήματα Υποστήριξης Αποφάσεων*, Κεφ.11, Data Warehouse and OLAP, 7 Αυγούστου 2008, 14:30, σελ 247
<http://www.ergasva.tuc.gr/Users/Matsatsinis/Courses/SYA/sva.htm>

² *Προς ένα Ολοκληρωμένο Πληροφοριακό Σύστημα Συλλογής, Αποθήκευσης και Επεξεργασίας Εκπαιδευτικών Δεδομένων: Από το Όραμα στο Σύστημα Edubase*, 25 Ιουλίου 2008, 20:10, σελ. 618,
<http://www.etpe.gr/files/proceedings/uploads/p113.pdf>

3) Η επεξεργασία και η αναλυτική διεργασία που χρησιμοποιεί ο χρήστης για την εξαγωγή των επιθυμούμενων αποτελεσμάτων, οδηγούν στη εύρεση αγνώστων αλλά χρήσιμων έως τώρα πληροφοριών. Αποτελεί μία διαδικασία πολύπλοκη, η οποία απαιτεί μεγάλη υπολογιστική ισχύ και αρκετό χρόνο για την εξαγωγή των αποτελεσμάτων. Οι OLAP εφαρμογές συνεργάζονται με αποθήκες δεδομένων που περιέχουν ιστορικά και συγκεντρωτικά δεδομένα.³

4.2 Ανάλυση εργαλείων

Ορισμένα εργαλεία OLAP, τα οποία υποστηρίζουν τις παραπάνω εργασίες, παρουσιάζονται και αναλύονται παρακάτω:

Πίνακας Εργαλείων OLAP

ΚΑΤΗΓΟΡΙΕΣ ΕΡΓΑΛΕΙΑ	ΕΓΚΑΤΑΣΤΑΣΗ	ΚΟΣΤΟΣ	ΣΥΝΕΡΓΑΣΙΑ	OPEN SOURCE	ΤΥΠΟΣ
Sisense Prism Desktop	Free Download	0,00 \$	Microsoft Access Excel	Yes	Full Setup.exe
Mondrian	Free Download	0,00 \$	Microsoft Access	Yes	Full Zip File
Palo	Free Download	0,00 \$	Microsoft Excel	Yes	Full Zip File
Olap X Application	Registration	99 \$	Microsoft Excel	No	Shareware
Olap Cube	Registration	29 \$	Microsoft Excel	No	Shareware

Πηγή: Ιδία Έρευνα

1) **Sisense Prism Desktop:** Είναι ένα εργαλείο ανοιχτού κώδικα (open source), το οποίο μπορεί να αποκτηθεί δωρεάν μέσω του διαδικτύου σε ολοκληρωμένη μορφή. Είναι ένα σύνολο εργαλείων λογισμικού που δίνει στους χρήστες τη δυνατότητα, να διαχειρίζονται και να μοιράζονται εφαρμογές επιχειρηματικής ευφυΐας, χωρίς να χρειάζεται να μάθουν πώς να

³ Συστήματα Υποστήριξης Αποφάσεων, Κεφ.11, Data Warehouse and OLAP, 7 Αυγούστου 2008, 14:30, σελ 247 <http://www.ergasya.tuc.gr/Users/Matsatsinis/Courses/SYA/sya.htm>

προγραμματίσουν, πώς να θέτουν ερωτήματα και πώς να εκτελούν βελτιστοποιήσεις βάσεων δεδομένων που υποστηρίζουν ερωτήματα εντατικής εργασίας.⁴

2) **Mondrian**: Είναι ένα εργαλείο ανοιχτού κώδικα, όπου επιτρέπει την ανάλυση μεγάλων ποσοτήτων δεδομένων τα οποία είναι αποθηκευμένα σε SQL βάσεις δεδομένων χωρίς να χρησιμοποιείται η SQL. Παρέχει την δυνατότητα απόκτησής του δωρεάν, μέσω του διαδικτύου σε ολοκληρωμένη μορφή. Επίσης, η εισαγωγή των δεδομένων του γίνεται μέσω της Microsoft Access.

Ο στόχος του Mondrian είναι να παρουσιαστεί η πολυδιάστατη ανάλυση στις μάζες. Για να το κάνει αυτό πρέπει να είναι:

- Ελεύθερο (free)
- Φορητό (portable)
- Εύκολο στην εγκατάσταση (easy to install)
- Εύκολο στην ενσωμάτωση (easy to integrate)
- Εύκολο στην κατανόηση (easy to understand)

Ως ένας OLAP Server ανοιχτού κώδικα, που γράφεται σε Java, ικανοποιεί τους παραπάνω στόχους και το γεγονός ότι είναι ανοιχτού κώδικα και συνδυάζεται με την Java κάνει το Mondrian αρκετά εύκαμπτο, έτσι ώστε να ενσωματώνεται σε ποικίλα περιβάλλοντα.

Το Mondrian είναι μέρος του Pentaho Open Source BI Suite. Το Pentaho στοχεύει να παραδώσει την καλύτερη δυνατή εμπειρία στους χρήστες, με την ενσωμάτωση του Mondrian σε αυτό.

Το Mondrian δεν μπορεί να τα κάνει όλα. Εάν τα έκανε όλα, θα ήταν δύσκολο για κάποιον να το κατεβάσει από το διαδίκτυο, δύσκολο να το εγκαταστήσει και ακόμα πιο δύσκολο να το ενσωματώσει με άλλο λογισμικό. Το πλεονέκτημα είναι ότι είναι ανοιχτού κώδικα. Εάν ένα χαρακτηριστικό γνώρισμα λείπει, είναι συχνά εύκολο να προστεθεί το χαρακτηριστικό γνώρισμα σε Mondrian ή να ενσωματωθεί με ένα άλλο προϊόν ανοιχτού κώδικα που παρέχει το χαρακτηριστικό γνώρισμα.

Το Mondrian χρησιμοποιείται για:

- Υψηλή επίδοση, διαλογική ανάλυση των μεγάλων ή μικρών όγκων των πληροφοριών
- «Διαστατική» εξερεύνηση των στοιχείων, παραδείγματος χάριν αναλύει τις πωλήσεις ανά παραγωγή, ανά περιοχή και ανά χρονικό διάστημα

⁴ Prism FAQ, 30 Αυγούστου 2008, 16:10, <http://sisense.com/help/learning-center/prism-faq/>

- Ανάλυση της πολυδιάστατης γλώσσας έκφρασης (MDX) σε δομημένη γλώσσα διατύπωσης ερωτήσεων (SQL) για να ανακτήσει τις απαντήσεις στις διαστατικές ερωτήσεις
- Μεγάλες ερωτήσεις μέσω της χρήσης συνολικών πινάκων στο RDBMS
- Προηγμένους υπολογισμούς που χρησιμοποιούν τις εκφράσεις υπολογισμού της γλώσσας MDX

Ένα σύστημα Mondrian OLAP αποτελείται από τέσσερα στρώματα: το στρώμα παρουσίασης, το διαστατικό στρώμα, το στρώμα αστεριών και το στρώμα αποθήκευσης.

Το στρώμα παρουσίασης καθορίζει τι βλέπει ο τελικός χρήστης στο όργανο ελέγχου και πώς μπορεί να αλληλεπιδράσει για να υποβάλει τις νέες ερωτήσεις. Υπάρχουν πολλοί τρόποι να παρουσιαστούν τα πολυδιάστατα σύνολα δεδομένων, συμπεριλαμβανομένων των συγκεντρωτικών πινάκων – αξόνων (Pivot), της πίτας, της γραμμής, των ιστογραμμάτων και των προηγμένων εργαλείων απεικόνισης, όπως οι clickable χάρτες και η δυναμική γραφική παράσταση.

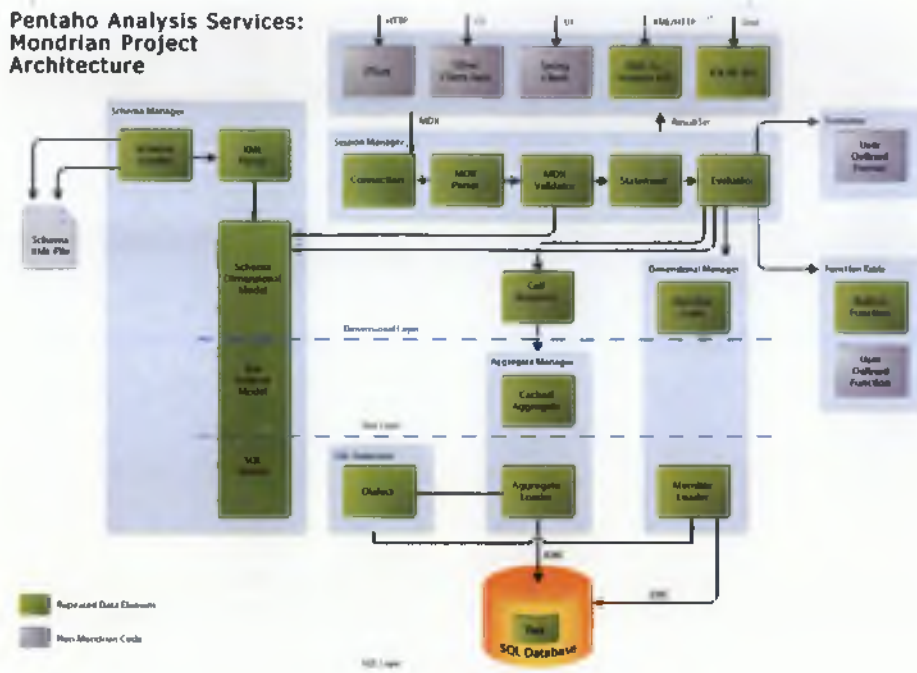
Το δεύτερο στρώμα είναι το διαστατικό στρώμα. Το διαστατικό στρώμα αναλύει, επικυρώνει και εκτελεί τις ερωτήσεις MDX. Πρώτα, υπολογίζονται οι άξονες και έπειτα οι τιμές των κελιών μέσα στους άξονες.

Το τρίτο στρώμα είναι το στρώμα αστεριών και είναι αρμόδιο για τη διατήρηση μιας συνολικής κρύπτης. Το διαστατικό στρώμα στέλνει τα αιτήματα στο σύνολο των κελιών. Εάν τα ζητούμενα κελιά δεν είναι στην κρύπτη, ο διευθυντής συνάθροισης στέλνει ένα αίτημα στο στρώμα αποθήκευσης.

Το στρώμα αποθήκευσης είναι ένα RDBMS. Είναι αρμόδιο για την παροχή των αθροισμένων στοιχείων κελιών και για τα μέλη από τους πίνακες διάστασης. Αυτά τα συστατικά μπορούν όλα να υπάρξουν στην ίδια μηχανή ή μπορούν να διανεμηθούν μεταξύ των μηχανών. Τα στρώματα 2 και 3, που περιλαμβάνουν τον κεντρικό υπολογιστή Mondrian, πρέπει να είναι στην ίδια μηχανή. Το στρώμα αποθήκευσης θα μπορούσε να είναι σε μια άλλη μηχανή. Σε ένα σύστημα πολλών χρηστών, το στρώμα παρουσίασης θα υπήρχε στη μηχανή κάθε τελικού χρήστη.⁵

⁵ Pentaho Analysis Services: Mondrian Project, 22 Αυγούστου 2008, 17:40, <http://mondrian.pentaho.org/>

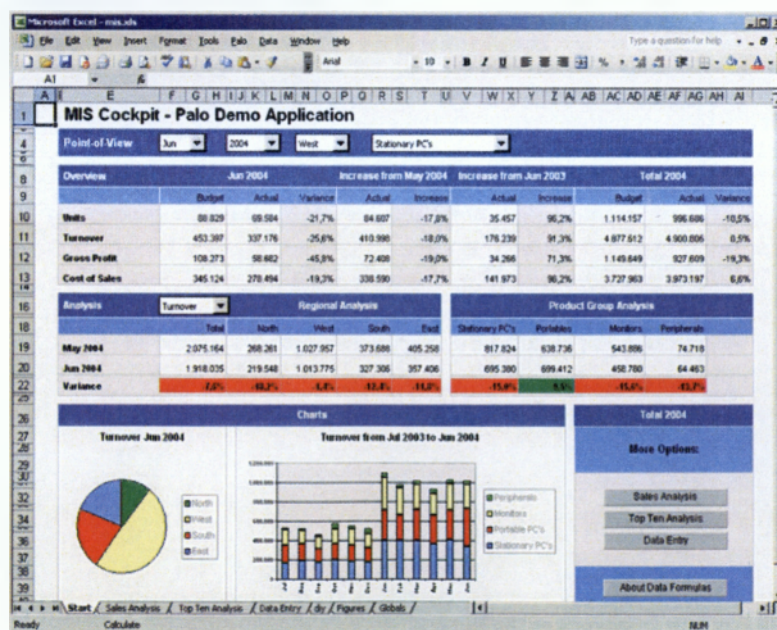
Pentaho Analysis Services: Mondrian Project Architecture



Σχήμα 4-1: Αρχιτεκτονική Mondrian

Πηγή: Pentaho Analysis Services: Mondrian Project, <http://mondrian.pentaho.org/>

3) **Palo**: Είναι ένα εργαλείο ανοιχτού κώδικα (open source), το οποίο παρέχεται δωρεάν μέσω του διαδικτύου. Η εισαγωγή των δεδομένων του γίνεται μέσω του Microsoft Excel και παρέχεται ολοκληρωμένο. Μια απεικόνιση του Palo φαίνεται στο παρακάτω σχήμα.



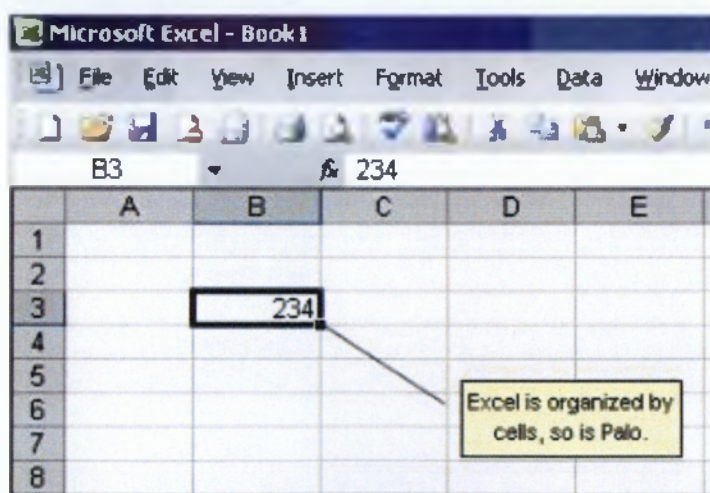
Σχήμα 4-2: Απεικόνιση Εργαλείου Palo

Πηγή: Jedox Palo OLAP Server-Open Source Business Intelligence for Planning, Analysis, Reporting and Consolidation, <http://www.jedox.com/en/enterprise-spreadsheet-server/excel-olap-server/palo-server.html>

Το Palo είναι ένα εξαιρετικά αποδοτικό εργαλείο ανοιχτού κώδικα, το οποίο διατίθεται ελεύθερο με πλήρη άδεια και επαγγελματική υποστήριξη. Είναι ένας OLAP Server, αλλά πιο συγκεκριμένα είναι ένας πολυδιάστατος MOLAP Server που δεν μπορεί να χρησιμοποιηθεί εάν αυτό που αναζητεί κανείς είναι ένας ROLAP Server.

Τα πιο σημαντικά χαρακτηριστικά του Palo και του MOLAP είναι τα εξής:

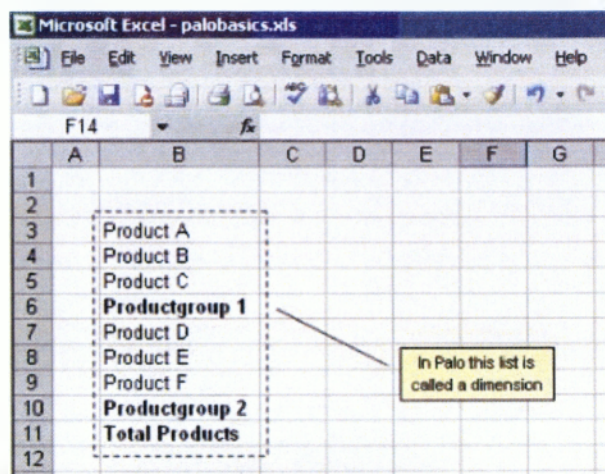
- **Write-Back:** Εκτός από την παροχή της πολυδιάστατης ανάλυσης στοιχείων των υπάρχοντων στοιχείων, το Palo επιτρέπει επίσης την εισαγωγή των νέων στοιχείων άμεσα στον κύβο και διευκολύνει και τον από επάνω προς τα κάτω (top-down), καθώς επίσης και από κάτω προς τα επάνω (bottom-up) προγραμματισμό. Αυτό το σημαντικό χαρακτηριστικό, διαφοροποιεί το Palo από τους συγγενικούς κεντρικούς υπολογιστές OLAP.
- **User Rights:** Η λειτουργία αυτή είναι ένα νέο χαρακτηριστικό γνώρισμα του Palo 1.5. Επιτρέπει στο χρήστη να διαχειριστεί τα δικαιώματα πρόσβασης για τις βάσεις δεδομένων, τους κύβους και τα στοιχεία της κάθε διάστασης.
- **Cell-Based:** Εάν κάνετε μια ερώτηση σε μια σχεσιακή βάση δεδομένων, το μικρότερο αποτέλεσμα που θα πάρετε είναι ένα αρχείο (συνήθως που αποτελείται από διάφορους τομείς). Εάν κάνετε μια ερώτηση σε Palo, το αποτέλεσμά σας είναι μια συγκεκριμένη τιμή πχ. την τιμή ενός συγκεκριμένου κελιού. Γι' αυτό το λόγο, το Palo συνεργάζεται άψογα με το Excel, αφού και αυτό αποτελείται από κελιά.



Σχήμα 4-3: Cell-Based Palo

Πηγή: *Jedox Palo OLAP Server-Open Source Business Intelligence for Planning, Analysis, Reporting and Consolidation*, <http://www.jedox.com/en/enterprise-spreadsheet-server/excel-olap-server/palo-server.html>

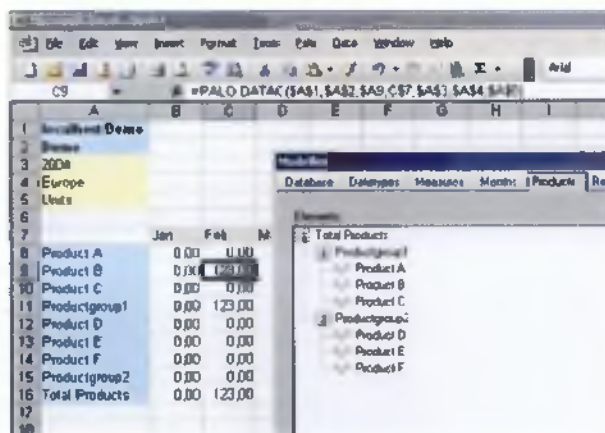
- **Multidimensional:** Ο πολυδιάστατος χαρακτήρας είναι ένας εναλλακτικός τρόπος να οργανωθούν τα στοιχεία σε μια βάση δεδομένων. Σε μια σχεσιακή βάση δεδομένων χρησιμοποιούνται οι πίνακες, τα αρχεία και οι τομείς. Σε μια πολυδιάστατη βάση δεδομένων υπάρχουν οι κύβοι, οι διαστάσεις και τα κελιά.



Σχήμα 4-4: Multidimensional

Πηγή: *Jedox Palo OLAP Server-Open Source Business Intelligence for Planning, Analysis, Reporting and Consolidation*, <http://www.iedox.com/en/enterprise-spreadsheet-server/excel-olap-server/palo-server.html>

- **Hierarchies:** Η χρησιμοποίηση των ιεραρχιών κατά τον καθορισμό μιας διάστασης προσθέτει πολλή δύναμη στους κύβους στοιχείων Palo. Με τις ιεραρχικές διαστάσεις, ο κύβος είναι σε θέση να κάνει τις πολυδιάστατες συναθροίσεις. Έτσι δεν είναι απαραίτητο να γίνουν αυτοί οι υπολογισμοί με τους τύπους του Excel, αλλά διαμορφώνονται κατάλληλα στον κύβο.



Σχήμα 4-5: Hierarchies

Πηγή: *Jedox Palo OLAP Server-Open Source Business Intelligence for Planning, Analysis, Reporting and Consolidation*, <http://www.iedox.com/en/enterprise-spreadsheet-server/excel-olap-server/palo-server.html>

- **Memory-based & Real-time:** Το Palo βασίζεται στη μνήμη. Αυτό σημαίνει ταχύτητα. Η ανάκτηση στοιχείων από τα κελιά και επίσης από τα αθροισμένα κελιά γίνονται σε πραγματικό χρόνο. Υπάρχουν άλλες πολυδιάστατες βάσεις δεδομένων στην αγορά που είναι βασισμένες σε μια συγγενική τεχνολογία (βάσεις δεδομένων ROLAP). Έχουν τα πλεονεκτήματά τους από την άποψη της ικανότητας, αλλά είναι συνήθως πιο αργοί και δεν υποστηρίζουν την εισαγωγή νέων στοιχείων στους κύβους.
- **Server-Based:** Το Palo μπορεί να εγκατασταθεί τοπικά ή σε ένα δίκτυο επιχείρησης. Στον τόπο δικτύων που όλοι οι χρήστες εργάζονται με τα ίδια στοιχεία, οι αλλαγές σε έναν υπολογισμό με λογιστικό φύλλο (spreadsheet) είναι αμέσως ορατές στους άλλους τερματικούς σταθμούς στο ενδοδίκτυο.
- **Subsets:** Ένα υποσύνολο είναι ένα σύνολο μέσα σε μια διάσταση. Με ένα υποσύνολο κάποιος μπορεί να επιλέξει τα στοιχεία ή τα υποσύνολα των στοιχείων και να τα ταξινομήσει σύμφωνα με ορισμένα κριτήρια. Παράδειγμα: Αν επιθυμούσε κάποιος να πάρει τα 10 κορυφαία στοιχεία ενός καταλόγου προϊόντων, θα επιθυμούσε στα αποτελέσματα να είναι μέσα στις καθορισμένες οριακές τιμές.
- **Rules:** Ο παράγοντας σταθεροποίησης των στοιχείων χρησιμοποιείτε για να πραγματοποιήσει τους απλούς υπολογισμούς. Οι επιχειρηματικοί κανόνες Palo ενισχύουν αυτήν την λειτουργία και επιτρέπουν την εκτέλεση σύνθετων υπολογισμών (πολλαπλασιασμού, λογικών και διάφορων λειτουργίες τμήματος).⁶

4) **OLAP X Application:** Είναι ένα εργαλείο, το οποίο δεν παρέχεται δωρεάν. Χρειάζεται εγγραφή στην επίσημη σελίδα του προγράμματος και συνεργάζεται με το Microsoft Excel. Δεν είναι εργαλείο ανοιχτού κώδικα και δεν παρέχεται ολοκληρωμένο.

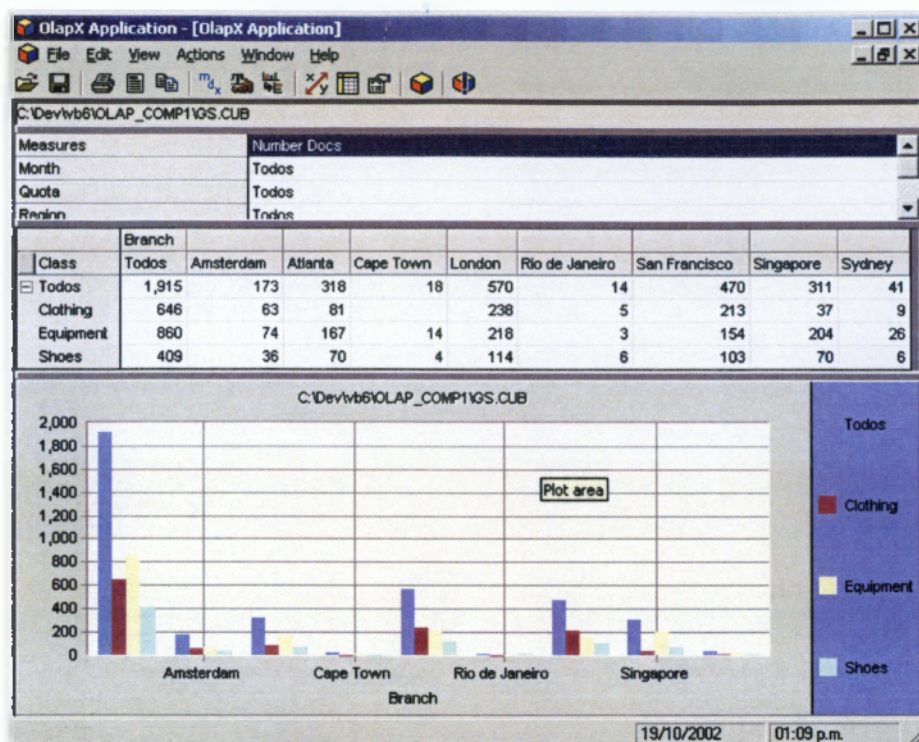
Η εφαρμογή OlapX είναι ένα περίπλοκο πολυδιάστατο εργαλείο ξεφυλλίσματος βάσεων δεδομένων. Επιτρέπει τη διαλογική ανάλυση, την υποβολή έκθεσης και την παρουσίαση δημιουργιών των βάσεων δεδομένων υπηρεσιών ανάλυσης της Microsoft. Η εφαρμογή OlapX σχεδιάζεται για τους χρήστες σε οποιοδήποτε επιχείρηση ή τεχνικό επίπεδο ικανότητας σε μια επιχείρηση για να εκτελέσει την πολυδιάστατη ανάλυσή τους και να δημιουργήσει τις εκθέσεις με τις οποίες θα πάρουν την καλύτερη απόφαση.⁷

⁶ *Jedox Palo OLAP Server-Open Source Business Intelligence for Planning, Analysis, Reporting and Consolidation*, 18 Ιουλίου 2008, 20:10,

<http://www.jedox.com/en/enterprise-spreadsheet-server/excel-olap-server/palo-server.html>

⁷ *Olap X Application 3.0*, 19 Ιουλίου 2008, 21:34,

<http://www.download25.com/olap-x-application-download.html>



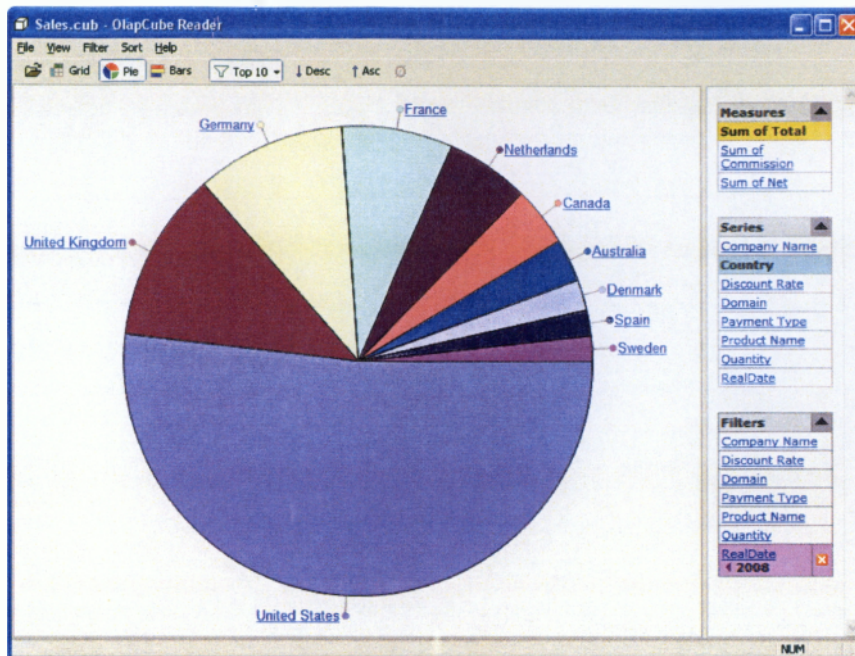
Σχήμα 4-6: OLAP X Application

Πηγή: *Olap X Application 3.0*, <http://www.download25.com/olap-x-application-download.html>

5) **OLAP Cube:** Είναι ένα εργαλείο, το οποίο παρέχεται δωρεάν, αλλά δεν είναι ανοιχτού κώδικα. Συνεργάζεται με το Microsoft Excel και δεν παρέχεται ολοκληρωμένο.

Το OlapCube είναι ένα απλό, όμως ισχυρό εργαλείο ανάλυσης στοιχείων. Το OlapCube δημιουργεί τοπικούς κύβους από τα στοιχεία που αποθηκεύονται σε οποιαδήποτε σχεσιακή βάση δεδομένων (συμπεριλαμβανομένου MySQL, PostgreSQL, της Microsoft Access, SQL aServer, κεντρικός υπολογιστής SQL σαφής, Oracle, Oracle σαφής). Ο κύβος OLAP αποτελείται από τα αριθμητικά γεγονότα που καλούνται μέτρα και ταξινομούνται σε διαστάσεις.⁸

⁸ *OlapCube 1.1*, 19 Ιουλίου 2008, 22:20, <http://www.download25.com/olapcube-download.html>



Σχήμα 4-7: OLAP Cube

Πηγή: *OlapCube 1.1*, <http://www.download25.com/olapcube-download.html>

Σύμφωνα με τα παρακάτω κριτήρια δηλαδή: α) το εργαλείο να είναι ανοιχτού κώδικα, β) να παρέχεται δωρεάν μέσω διαδικτύου, γ) να παρέχεται ολοκληρωμένο και δ) να είναι εύχρηστο, επιλέχτηκε για την ανάλυση πολυδιάστατων δεδομένων το εργαλείο Sisense Prism Desktop, το οποίο θα αναλυθεί στο επόμενο κεφάλαιο.

ΚΕΦΑΛΑΙΟ 5

ΑΝΑΛΥΣΗ ΕΡΓΑΛΕΙΟΥ SISENSE PRISM DESKTOP ΚΑΙ ΔΙΑΔΙΚΑΣΙΑ ΕΓΚΑΤΑΣΤΑΣΗΣ ΤΟΥ

5.1 Γενικά

Όπως ήδη αναφέρθηκε στο προηγούμενο κεφάλαιο, το Sisense Prism Desktop είναι ένα σύνολο εργαλείων λογισμικού που δίνει στους χρήστες τη δυνατότητα, να διαχειρίζονται και να μοιράζονται εφαρμογές επιχειρηματικής ευφυΐας, χωρίς να χρειάζεται να μάθουν πώς να προγραμματίσουν, πώς να θέτουν ερωτήματα και πώς να εκτελούν βελτιστοποιήσεις βάσεων δεδομένων που υποστηρίζουν ερωτήματα εντατικής εργασίας.

Με το Prism, οι χρήστες δημιουργούν πίνακες, εκθέσεις και καθοδηγούμενες αναλυτικές αιτήσεις για επιχειρηματικές λειτουργίες, όπως πωλήσεις, δραστηριότητες, marketing, χρηματοδότηση, ανθρώπινους πόρους ή οποιαδήποτε βάση δεδομένων υποστηρίζεται από το σύστημα. Είναι εύκολο στην εγκατάσταση και δεν απαιτεί ειδικές υποδομές. Μόλις εγκατασταθεί, μπορεί να συνδεθεί με τις διαθέσιμες πηγές δεδομένων και εφαρμογών BI και να παραδώσει μέσα σε λίγα λεπτά αποτελέσματα.

Κάθε Prism BI αίτηση είναι αποθηκευμένη σε ένα Prism έγγραφο, το οποίο είναι ένας χώρος με πηγές δεδομένων και μικροσυσκευές (Widgets). Οι χρήστες μπορούν να χρησιμοποιήσουν BI εφαρμογές ανοίγοντας ένα Prism έγγραφο, το οποίο μπορεί να έχει ληφθεί από άλλους χρήστες ή να έχει δημιουργηθεί από το μηδέν για προσωπική και ομαδική χρήση. Οι Prism BI εφαρμογές μπορούν να παρουσιάσουν πίνακες οργάνων (dashboard), εκθέσεις, επιχειρηματικές παρουσιάσεις, ad-hoc εργαλεία και οτιδήποτε άλλο αφορά την παρουσίαση και την ανάλυση πληροφοριών.

Κάθε Prism BI εφαρμογή έχει την εξής δομή:

- **Πηγές δεδομένων (Data Sources):** Μια πηγή δεδομένων αντιπροσωπεύει ένα δεσμό, που κρατά τα χαρακτηριστικά συνδεσιμότητας σε φυσικό χώρο αποθήκευσης δεδομένων, όπως για παράδειγμα ένα υπολογιστικό φύλλο του Excel, ένας OLAP κύβος ή μια βάση δεδομένων πίνακα.

- **Δελτία (Sheets):** Είναι οπτικά δοχεία των Widgets. Αποτελούν την κύρια περιοχή, στην οποία ο κάθε χρήστης μπορεί να δημιουργήσει την δική του BI εφαρμογή με drag & drop λειτουργίες.

- **Μικροσυσκευές (Widgets):** Οπτικοποιούν τις πληροφορίες και επιτρέπουν στους χρήστες να αλληλεπιδρούν με αυτές. Για παράδειγμα, περιλαμβάνουν άξονες (Pivots), γραφήματα, δείκτες, εικόνες, πλούσια πλαίσια κειμένου και υπερσυνδέσμους.
- **Ερωτήσεις (Questions):** Αποτελούν εικονικές αναζητήσεις που αντιπροσωπεύουν περισσότερο πολύπλοκες ροές πληροφοριών (χωρισμός σε τμήματα, διασταυρώσεις κ.α.).¹

5.2 Ορισμοί

Πολυδιάστατη παρουσίαση μιας βάσης δεδομένων είναι μια παρουσίαση, στην οποία διάφορα Κλειδιά (Στήλες) διαφόρων βάσεων δεδομένων, παρουσιάζονται ως Διαστάσεις και όχι ως σχετικές στήλες. Η πολυδιάστατη παρουσίαση επιτρέπει την καλύτερη προσβασιμότητα και ανάλυση των δεδομένων που αποθηκεύονται στη βάση δεδομένων. Στην παρουσίαση ενός πίνακα, κάθε Κλειδί (Ημερομηνία, Ηλικία, ZIP κ.λπ.) παρουσιάζεται ως μια στήλη και κάθε σειρά υποβάλει τα δεδομένα που σχετίζονται με αυτό το κλειδί.

Πίνακας 5-1: Πολυδιάστατος πίνακας

Customers ID	RegDate	Date	Discount Level	Age	ZIP	Region	Town	Shop	Gender	Product	Sales	Quantity
000522	06.09.2006	01.01.2007	1	69	220035	220	T 78	TA 1	F	Accessory 1	41.860	251
000522	06.09.2006	01.08.2007	1	69	220035	220	T 78	TA 1	F	Accessory 1	42.202	271
000522	06.09.2006	01.07.2007	1	69	220035	220	T 78	TA 1	F	Accessory 1	38.596	161
000522	06.09.2006	01.03.2007	1	69	220035	220	T 78	TA 1	F	Accessory 1	14.972	75
000522	06.09.2006	01.02.2007	1	69	220035	220	T 78	TA 1	F	Accessory 1	7.558	38
000522	06.09.2006	01.05.2007	1	69	220035	220	T 78	TA 1	F	Accessory 1	22.817	111
000522	06.09.2006	01.12.2007	1	69	220035	220	T 78	TA 1	F	Accessory 1	17.714	86

Πηγή: *Definitions*, <http://sisense.com/help/learning-center/definitions/>

¹ *Prism Frequently Asked Questions (FAQ)*, 1 Σεπτεμβρίου 2008, 17:01, <http://sisense.com/help/learning-center/prism-faq/>

Η πολυδιάστατη παρουσίαση του ίδιου πίνακα είναι ως εξής:



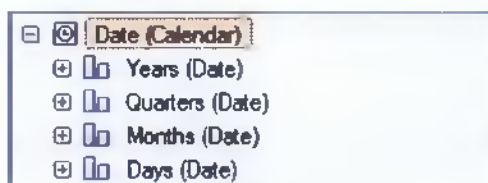
Σχήμα 5-1: Δέντρο διαστάσεων πολυδιάστατου πίνακα

Πηγή: *Definitions*, <http://sisense.com/help/learning-center/definitions/>

Στην παρουσίαση του πολυδιάστατου πίνακα, δεν υπάρχουν ούτε στήλες ούτε σειρές μόνο Διαστάσεις, όπου κάθε διάσταση αντιπροσωπεύει ένα Κλειδί στην πηγή δεδομένων. Οι ποσοτικές τιμές του πίνακα (μέτρα) είναι συγκεντρωμένες για κάθε μία από τις Διαστάσεις. Η παρουσίαση αυτή ονομάζεται **Δέντρο Διαστάσεων** και οι διερευνήσεις γίνονται κυριολεκτικά από την επιλογή συνδυασμού μιας ή περισσότερων Διαστάσεων και ενός ή περισσότερων μέτρων. Μια διάσταση είναι ένα περιγραφικό κλειδί με το οποίο ένας πίνακας μπορεί να αναλυθεί π.χ. σε ημερομηνία, CustomerID, Φύλο, Product, μέτρα κλπ. Οι Διαστάσεις θα περιλαμβάνουν κανονικά τα περιγραφικά στοιχεία (Πόλη, Όνομα προϊόντος, ημερομηνία κτλ), ενώ τα μέτρα θα περιλαμβάνουν αριθμητικές τιμές.

Το Prism χειρίζεται την Διάσταση Χρόνος (Time Dimension), με ιδιαίτερο τρόπο. Μετατρέπει αυτόματα κάθε στήλη που περιέχει χρονικές αναφορές (Timestamps) σε τρεις διαστάσεις ως εξής:

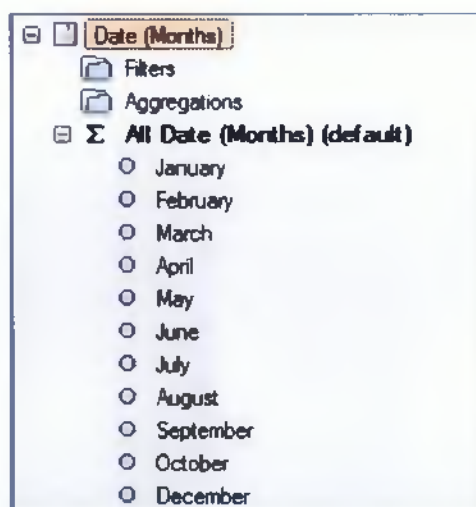
1. Date (Calendar): Αυτή είναι η κύρια Διάσταση Χρόνος που επιτρέπει την διερεύνηση σε Χρόνια, Τεταρτημόρια, Μήνες και Ημέρες.



Σχήμα 5-2: Date (Calendar)

Πηγή: *Definitions*, <http://sisense.com/help/learning-center/definitions/>

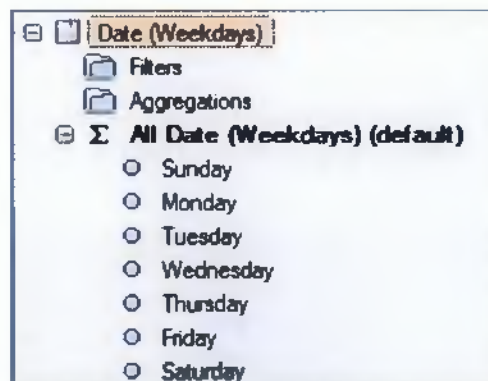
2. Date (Month): Η Διάσταση αυτή επιτρέπει την ανάλυση πολλών ετών, μέσω της διερεύνησης των μηνών του έτους.



Σχήμα 5-3: Date (Months)

Πηγή: *Definitions*, <http://sisense.com/help/learning-center/definitions/>

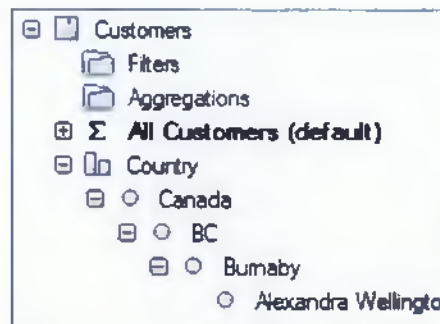
3. Date (Weekdays): Η Διάσταση αυτή επιτρέπει την ανάλυση πολλών μηνών και χρόνων, μέσω της διερεύνησης των ημερών του έτους.



Σχήμα 5-4: Date (Weekdays)

Πηγή: *Definitions*, <http://sisense.com/help/learning-center/definitions/>

Μια διάσταση αποτελείται από μία ή περισσότερες ιεραρχίες. Κάθε κόμβος σε αυτό το Δέντρο Διαστάσεων λέγεται **Μέλος** και μπορεί να χρησιμοποιηθεί σαν τελεστής στην ανάλυση. Κάθε μέλος αποτελείται από ένα γονικό και από ένα ή περισσότερα θυγατρικά μέλη. Για παράδειγμα, ο Καναδάς (Canada) είναι το θυγατρικό μέλος της Χώρας (Country) και το γονικό μέλος της BC.





Σχήμα 5-5: Γονικά και Θυγατρικά Μέλη

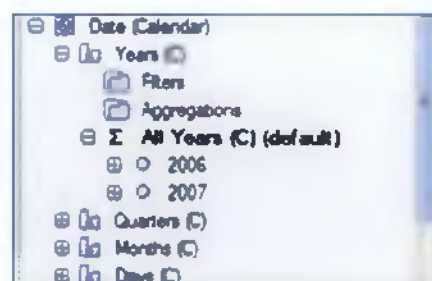
Πηγή: *Definitions*, <http://sisense.com/help/learning-center/definitions/>

Στο παραπάνω παράδειγμα, η Διάσταση Πελάτες (Customers) μπορεί να είναι μόνο γονικό μέλος, επειδή είναι το υψηλότερο επίπεδο, ενώ η Διάσταση Alexandra Wellington μπορεί να είναι μόνο θυγατρικό μέλος, επειδή είναι το χαμηλότερο επίπεδο.

Το **Drill Down** είναι η ικανότητα να παρατηρούνται τα μέλη από το υψηλότερο προς το χαμηλότερο επίπεδο.

Το **Drill Up** είναι η ικανότητα να παρατηρούνται τα μέλη από το χαμηλότερο προς το υψηλότερο επίπεδο.

Τα σύμβολα  και  στην αριστερή πλευρά της Διάστασης / Επίπεδο υποδεικνύουν ότι το drill-down και το drill-up είναι δυνατόν να πραγματοποιηθούν. Για παράδειγμα, στο σχήμα 5-6, το **Date (Calendar)** δείχνει ότι έχει τέσσερα θυγατρικά μέλη το **Years**, **Quarters**, **Months** και **Days**. Το μέλος **Years**, μπορεί κάποιος να παρατηρήσει ότι, έχει τρία θυγατρικά μέλη το **Filters**, **Aggregations** και **All Years**. Το μέλος **All Years** υποδεικνύει ότι έχει δύο θυγατρικά μέλη το **2006** και το **2007**.









Σχήμα 5-6: Drill-Down και Drill-Up

Πηγή: *Definitions*, <http://sisense.com/help/learning-center/definitions/>

Μέτρο (Measure) είναι ένας ποσοτικός δείκτης που αντιπροσωπεύει τη συνολική τιμή του μέλους π.χ. Sales (μέτρο) από CustomerID (διάσταση), Profit (μέτρο) από Region (διάστασης) κ.λπ. Τα μέτρα μιας ομάδας ονομάζονται μέτρα διάστασης και κάθε μέτρο στην ομάδα ονομάζεται μέλος.

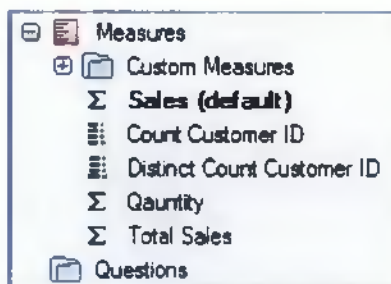
Συνυπολογισμός μέτρων (Measures Aggregation) είναι μια πράξη, που συνυπολογίζει πολλά από τα ίδια μέλη σε ένα αποτέλεσμα. Όλα τα μέτρα συνυπολογίζονται και κάποιος μπορεί να εφαρμόσει πολλαπλές συναθροίσεις στο ίδιο μέτρο. Το Prism υποστηρίζει τις ακόλουθες συναθροίσεις:

Πίνακας 5-2: Κατηγορίες Συναθροίσεων

ΟΜΑΔΟΠΟΙΗΣΗ	ΣΥΜΒΟΛΟ	ΠΕΡΙΓΡΑΦΗ
SUM		Προσθέτει ενός αξίες ενός μέτρου
AVERAGE		Υπολογίζει το μέσο όρο των αξιών ενός μέτρου
MIN		Υπολογίζει την μικρότερη τιμή ενός μέτρου ή μιας διάστασης
MAX		Υπολογίζει την μεγαλύτερη τιμή ενός μέτρου ή μιας διάστασης
COUNT		Μετράει τον αριθμό των γραμμών σε ένα μέτρο ή σε μια διάσταση
DISTINCT COUNT		Μετράει τον αριθμό των μοναδικών αξιών σε ένα μέτρο ή σε μια διάσταση

Πηγή: *Definitions*, <http://sisense.com/help/learning-center/definitions/>

Στο Δέντρο Μέτρων, κάθε μέτρο παρουσιάζεται με το όνομα και το σύμβολό του, που παρουσιάζει τον τύπο της συνάθροισης. Το Prism επιτρέπει πολλαπλές συναθροίσεις του ίδιου μέτρου π.χ. Sum [Sales], Average [Sales], Max [Sales] κ.λπ.





Σχήμα 5-7: Συμβολισμός Μέτρων

Πηγή: *Definitions*, <http://sisense.com/help/learning-center/definitions/>

Συναθροίσεις Διαστάσεων (Dimensional Aggregations) είναι οι υπολογισμοί, που συναθροίζουν ένα μέτρο με μια διάσταση. Ένας υπολογισμός δημιουργεί νέα μέλη, τα οποία αρχικά δεν ήταν διαθέσιμα στην πηγή δεδομένων.

Μέτρα Repository περιέχουν τους τύπους, που δημιουργήθηκαν από τον ίδιο τον χρήστη. Η εξοικονόμηση αυτών των υπολογισμών επιτρέπει την περαιτέρω χρήση των τύπων σε διάφορες αναλύσεις και εφαρμογές.

Φίλτρο είναι μια προϋπόθεση, που καθορίζει το υποσύνολο ενός συγκεκριμένου μέλους. Ένα φίλτρο “θα επιστρέψει” μόνο τα μέλη που πληρούν τους όρους, που καθορίζει το φίλτρο. Η προϋπόθεση δεν περιορίζει τον αριθμό των μελών, αλλά μπορεί να εξαρτηθεί από όσα μέλη είναι απαραίτητα. Ένα φίλτρο είναι αναγνωρισμένο από το ακόλουθο σύμβολο: 

Μια **ομάδα** είναι ένα φίλτρο, στο οποίο τα μέλη έχουν επιλεγθεί χειροκίνητα από τον χρήστη, αντί να φιλτραριστούν από μια σειρά προϋποθέσεων. Μια ομάδα είναι αναγνωρισμένη από το ακόλουθο σύμβολο: 

Κάθε διάσταση έχει ένα **Filter Member Repository**, όπου αποθηκεύονται όλα τα φίλτρα και οι ομάδες που έχουν οριστεί για μια συγκεκριμένη διάσταση.



Σχήμα 5-8: Αποθήκη Φίλτρων

Πηγή: *Definitions*, <http://sisense.com/help/learning-center/definitions/>

Μικροσυσκευές (Widgets) είναι αντικείμενα, τα οποία παρουσιάζουν τις πληροφορίες με οπτικό-γραφικό τρόπο. Μια Widget μπορεί να είναι ένας πίνακας, ένα γράφημα, ένας δείκτης ενός ημερολογίου κλπ.

Φόντο (Background) είναι μια προκαθορισμένη κατάσταση, με την οποία το Widget εμφανίζει πληροφορίες. Η προϋπόθεση αυτή μπορεί να είναι ένα φίλτρο, μία διάσταση, ένα επίπεδο, ένα μέτρο κλπ.

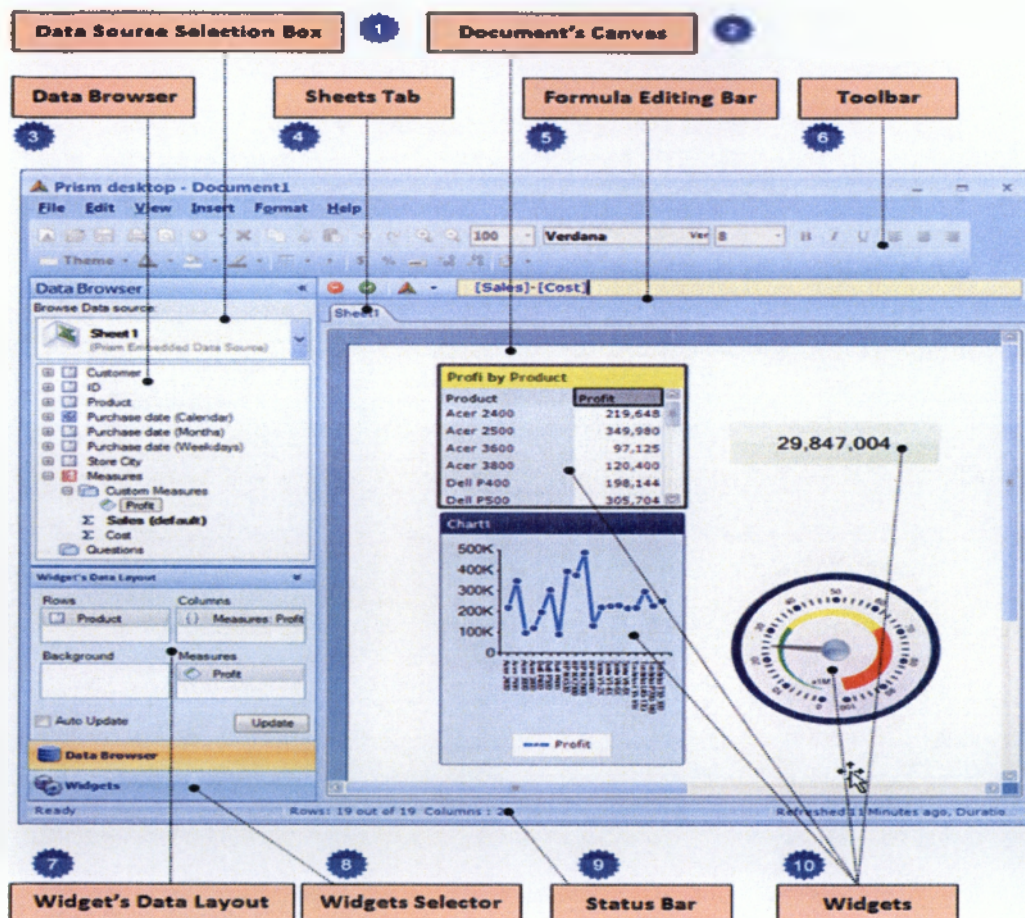
Πηγή Δεδομένων (Data Source) είναι μια βάση δεδομένων ή ένα αρχείο, που περιέχει τα δεδομένα που ο χρήστης θέλει να αξιολογήσει. Η πηγή δεδομένων μπορεί να είναι ένα αρχείο Excel, ένας κύβος OLAP ή αρχείο κειμένου.²

² *Definitions*, 1 Σεπτεμβρίου 2008, 19:45, <http://sisense.com/help/learning-center/definitions/>

5.3 Εργασιακό περιβάλλον του Prism

Το εργασιακό περιβάλλον του Prism αποτελείται από έναν Καμβά (Canvas), στον οποίο οι μικροσυσκευές επιλέγονται είτε από τον Επιλογέα Widget (Widget Selector), είτε κάνοντας δεξί κλικ οπουδήποτε στον καμβά.

Ακολουθεί μια περιγραφή των εργασιών Prism:

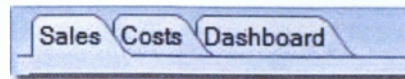


Σχήμα 5-9: Περιβάλλον Prism

Πηγή: *The Prism Environment*, <http://sisense.com/help/learning-center/the-prism-environment/>

1. **Επιλογή πηγής δεδομένων (Data Source Selection Box)** είναι ο χώρος, που μπορεί κάποιος να συνδεθεί με μια πηγή δεδομένων ή να δημιουργήσει μια καινούργια.
2. **Έγγραφο του Καμβά (Document's Canvas)** είναι ο χώρος, που μπορεί κάποιος να δημιουργήσει το Prism και να εφαρμόσει διάφορες γραφικές μικροσυσκευές που θα θέλει να δει στο έγγραφό του. Δεν υπάρχει περιορισμός στον αριθμό των συνδυασμών και των μικροσυσκευών που μπορεί να φιλοξενήσει το έγγραφο του καμβά.
3. Το **Data Browser** χρησιμοποιείτε για την επιλογή των μελών από την πηγή δεδομένων που θέλει να δει κάποιος σε μία από τις διαθέσιμες μικροσυσκευές.

4. Το **Φύλλο Καρτελών (Sheet Tabs)** χρησιμοποιείτε για την πολλαπλή προβολή των δεδομένων σε ένα ή πολλαπλά φύλλα.



Σχήμα 5-10: Sheets

Πηγή: *The Prism Environment*, <http://sisense.com/help/learning-center/the-prism-environment/>

5. **Γραμμή Δημιουργίας Formula (Formula Editing Bar)** χρησιμοποιείτε για την σύνθεση νέων τρόπων επεξεργασίας και έχει τρεις μορφές:

- Δημιουργία Formula σε πίνακα Pivot

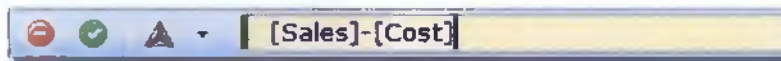


Σχήμα 5-11: Δημιουργία Formula

Πηγή: *The Prism Environment*, <http://sisense.com/help/learning-center/the-prism-environment/>

Με αυτήν την επιλογή κάποιος μπορεί να επεξεργαστεί μια Formula που υπάρχει ήδη στο Pivot.

- Επεξεργασία της Formula με το Repository

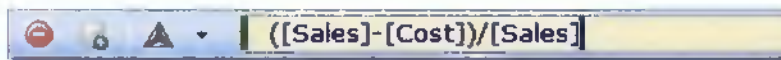


Σχήμα 5-12: Επεξεργασία Formula

Πηγή: *The Prism Environment*, <http://sisense.com/help/learning-center/the-prism-environment/>

Με αυτήν την επιλογή κάποιος μπορεί να δημιουργήσει μια Formula που υπάρχει ήδη στο Repository.

- Δημιουργία νέας Formula







Σχήμα 5-13: Δημιουργίας νέας Formulas

Πηγή: *The Prism Environment*, <http://sisense.com/help/learning-center/the-prism-environment/>

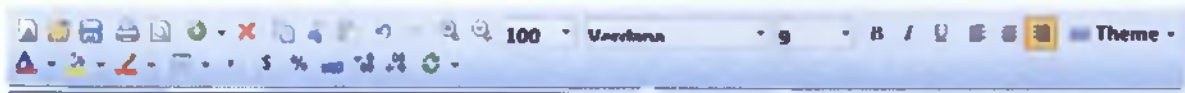
Με αυτήν την επιλογή κάποιος μπορεί να δημιουργήσει νέα Formula και να την αποθηκεύσει στο Repository για περαιτέρω χρήση.

Τα εικονίδια που χρησιμοποιούνται στην Formula Editing Bar είναι τα εξής:

-  Ακύρωση της λειτουργίας
-  Πρόσθεση νέας Formula στο Repository

-  Επιλογή στοιχείων για την νέα Formula
-  Ενημέρωση της Formula στο Repository

6. Η Γραμμή Εργαλείων (Toolbar) περιέχει διάφορα εικονίδια όπως τα εξής:



Σχήμα 5-14: Γραμμή Εργαλείων

Πηγή: *The Prism Environment*, <http://sisense.com/help/learning-center/the-prism-environment/>

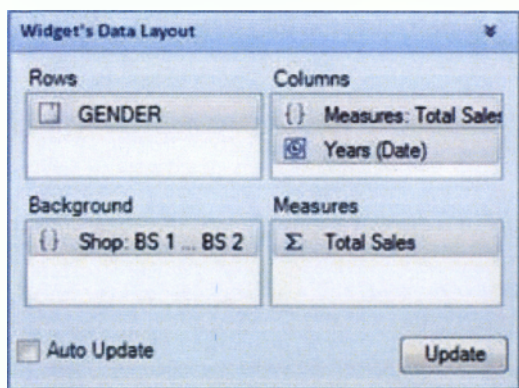
7. Το **Widget's Data Layout** είναι ένα δίκτυο, το οποίο καθορίζει την δομή των μικροσυσκευών. Όταν κάποιος έχει πολλαπλές μικροσυσκευές σε ένα έγγραφο, κάθε μικροσυσκευή θα έχει μια μοναδική Widget Data Layout. Οι μικροσυσκευές Pivot και Chart περιέχουν τέσσερις θέσεις που είναι οι εξής:

Πίνακας 5-3: Θέσεις Widgets

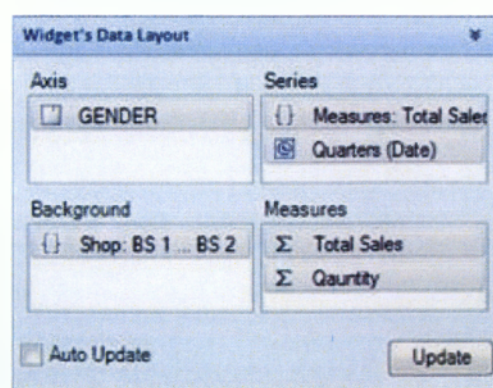
Σειρές (Rows)	Επιλέγει κάποιος από το Data Browser τα μέλη που θέλει να δει στις σειρές του Pivot, σύροντας και αφήνοντας τα μέλη μέσα στο κουτί
Στήλες (Columns)	Επιλέγει κάποιος από το Data Browser τα μέλη που θέλει να δει στις στήλες του Pivot, σύροντας και αφήνοντας τα μέλη μέσα στο κουτί
Μέτρα (Measures)	Επιλέγει κάποιος από το Data Browser τα μέτρα που θέλει να δει στο Pivot, σύροντας και αφήνοντας τα μέτρα μέσα στο κουτί
Φόντο (Background)	Επιλέγει κάποιος από το Data Browser τα μέλη, ανάλογα με το ποιες αξίες θα υπολογίσει το Pivot

Πηγή: *The Prism Environment*, <http://sisense.com/help/learning-center/the-prism-environment/>

Pivot's Data Layout



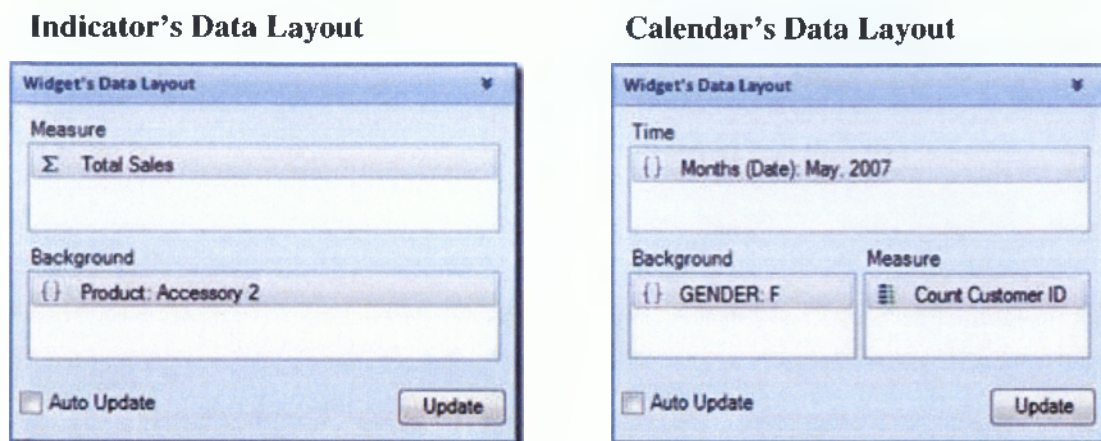
Chart's Data Layout



Σχήμα 5-15: Widget's Data Layout (Pivot – Chart)

Πηγή: *The Prism Environment*, <http://sisense.com/help/learning-center/the-prism-environment/>

Οι μικροσυσκευές **Ημερολόγιο (Calendar)** και **Δείκτης (Indicator)** περιέχουν δύο και τρεις θέσεις όπως οι εξής:



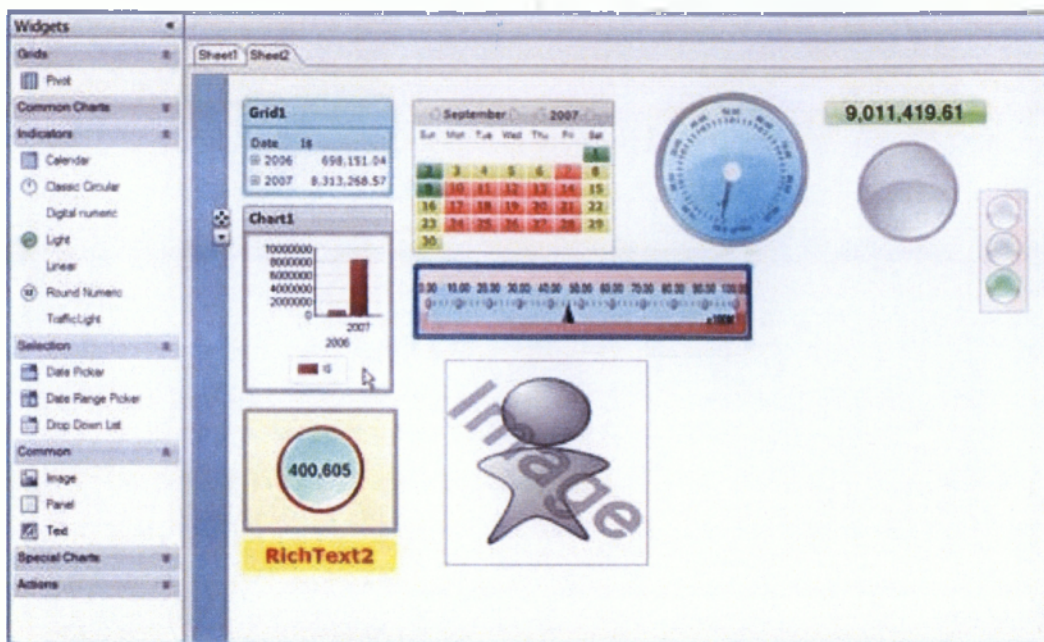
Σχήμα 5-16: Widget's Data Layout (Indicator – Calendar)

Πηγή: *The Prism Environment*, <http://sisense.com/help/learning-center/the-prism-environment/>

Τα μέλη μπορούν να προστεθούν στο Widget's Data Layout ή να αφαιρεθούν από αυτό ανά πάσα στιγμή. Τσεκάροντας την επιλογή **Auto Update** θα υπάρχει αυτόματη ενημέρωση της σχετικής μικροσυσκευής, όταν ένα νέο μέλος εισέρχεται στο Widget's Data Layout.

8. Ο **Επιλογέας μικροσυσκευών (Widgets Selector)** περιέχει μια λίστα των διαθέσιμων μικροσυσκευών.
9. Η **Γραμμή Κατάστασης (Status Bar)** δείχνει τον αριθμό των γραμμών (rows) και των στηλών (columns), καθώς και τον χρόνο που χρειάστηκε για να ολοκληρωθεί το ερώτημα.
10. **Μικροσυσκευές (Widgets)**. Ένα έγγραφο μπορεί να περιέχει οποιονδήποτε αριθμό, συνδυασμό καθώς και ένα μείγμα μικροσυσκευών, οι οποίες μπορούν να συνδεθούν με διαφορετικές πηγές δεδομένων και άλλες μικροσυσκευές.³

³ *The Prism Environment*, 2 Σεπτεμβρίου 2008, 15:30, <http://sisense.com/help/learning-center/the-prism-environment/>



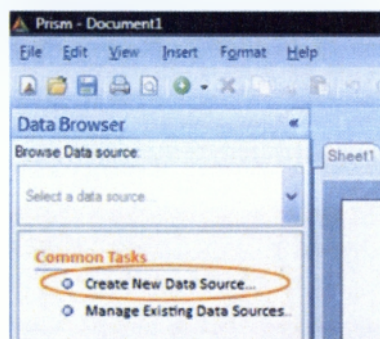
Σχήμα 5-17: Widgets

Πηγή: *The Prism Environment*, <http://sisense.com/help/learning-center/the-prism-environment/>

5.4 Σύνδεση με μια πηγή δεδομένων Excel

Τα βήματα για την σύνδεση με μια πηγή δεδομένων είναι τα εξής:

- Το άνοιγμα του οδηγού της πηγής δεδομένων, γίνεται κάνοντας κλικ στην επιλογή “Create New Data Source” ή από το κεντρικό μενού πατώντας **File**→ **Create New Data Source**.



Σχήμα 5-18: Create New Data Source

Πηγή: *Connecting to Excel*, http://www.sisense.com/Support/connect_to_data.aspx

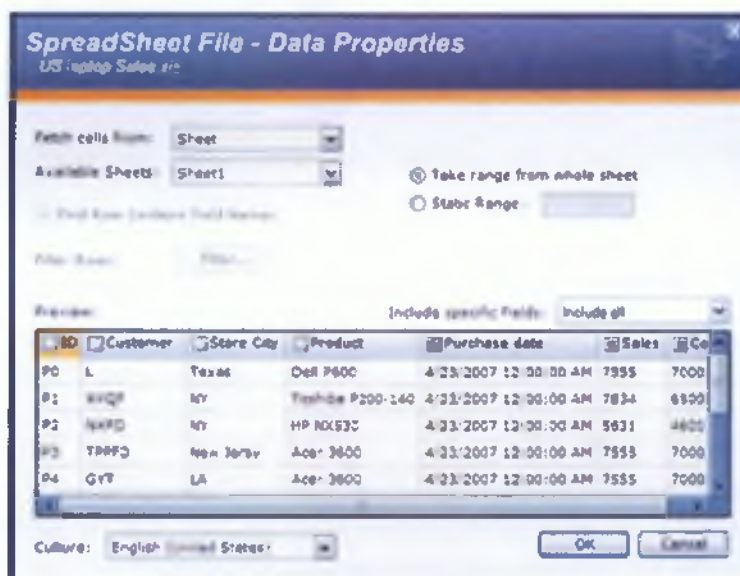
- Η επιλογή του προμηθευτή δεδομένων (Data Provider) γίνεται από τον κατάλογο φορέων παροχής δεδομένων. Στην προκειμένη περίπτωση συνδεόμαστε με μια πηγή δεδομένων Excel.



Σχήμα 5-19: Σύνδεση με Excel

Πηγή: *Connecting to Excel*, http://www.sisense.com/Support/connect_to_data.aspx

- Επιλογές σύνδεσης. Αφού γίνει η επιλογή της πηγής δεδομένων, τότε γίνεται η εισαγωγή των στοιχείων από το φύλλο του Excel. Κάποιος μπορεί να επιλέξει ανάμεσα από ένα συγκεκριμένο δελτίο, μια σειρά και ένα στατικό φάσμα και μετά να κάνει κλικ στο κουμπί OK.



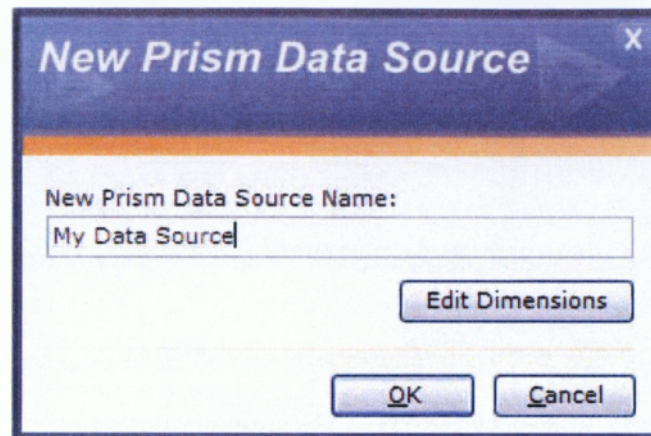
Σχήμα 5-20: Εισαγωγή Δεδομένων

Πηγή: *Connecting to Excel*, http://www.sisense.com/Support/connect_to_data.aspx

Προκειμένου το Prism να αναγνωρίσει σωστά τις αξίες, ο χρήστης πρέπει να έχει βεβαιωθεί σε ποια γλώσσα έχει δημιουργηθεί το αρχείο και να επιλέξει την κατάλληλη από

την επιλογή “Culture” στο κάτω αριστερό μέρος του παραθύρου και να κάνει κλικ στο κουμπί OK.

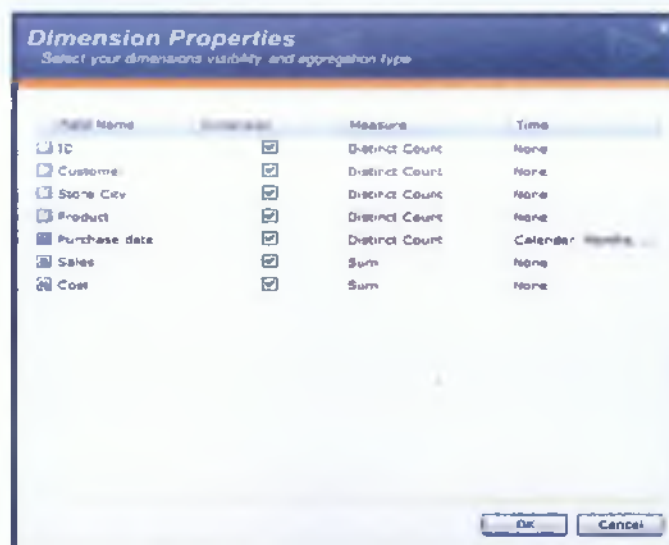
- Το επόμενο παράθυρο που θα δει ο χρήστης είναι το εξής:



Σχήμα 5-21: Μετονομασία Πηγής Δεδομένων

Πηγή: *Connecting to Excel*, http://www.sisense.com/Support/connect_to_data.aspx

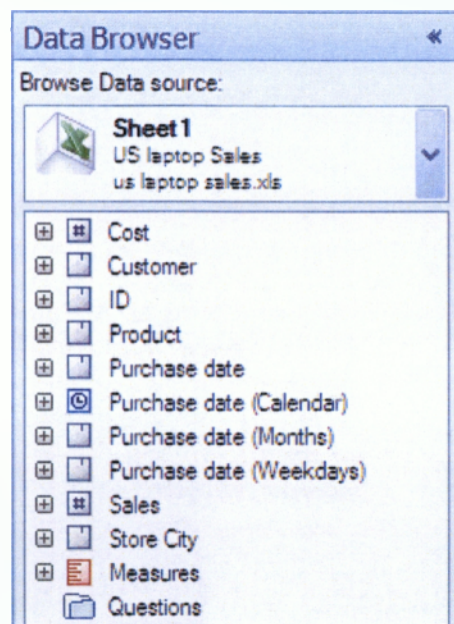
Εδώ ο χρήστης έχει την δυνατότητα να ονομάσει την πηγή δεδομένων όπως θέλει και αν επιθυμεί μπορεί να αλλάξει και να καθορίσει ο ίδιος τις διαστάσεις της πηγής δεδομένων, επιλέγοντας το “Edit Dimensions”. Αφού το επιλέξει θα εμφανιστεί το παρακάτω παράθυρο που φαίνεται στο σχήμα 5-21.



Σχήμα 5-22: Επιλογή Διαστάσεων

Πηγή: *Connecting to Excel*, http://www.sisense.com/Support/connect_to_data.aspx

- Αφού ο χρήστης κάνει τις αλλαγές που θέλει στις διαστάσεις, στα μέτρα και στον χρόνο της πηγής δεδομένων, πατώντας το κουμπί OK επιστρέφει στο παράθυρο του σχήματος 5-20 (New Prism.....) και πατώντας ξανά το κουμπί OK συνδέεται τελικά με την πηγή δεδομένων, που ο ίδιος έχει επιλέξει.⁴



Σχήμα 5-23: Δέντρο Διαστάσεων

Πηγή: *Connecting to Excel*, http://www.sisense.com/Support/connect_to_data.aspx

5.5 Λειτουργία Άξονα (Pivot) και Δεικτών

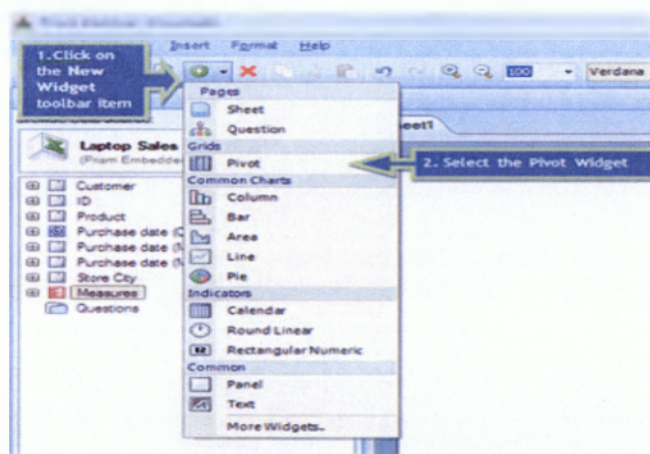
Ένας άξονας Pivot δίνει τη δυνατότητα της εύκολης ανάλυσης πολλαπλών διαστάσεων, χρησιμοποιώντας τα Widgets για ανάλυση, διερεύνηση, φιλτράρισμα, ταξινόμηση και παρουσίαση συγκεντρωτικών δεδομένων. Οι λειτουργίες του Pivot είναι οι εξής:

1. Εισαγωγή νέου Άξονα (Pivot)
2. Εισαγωγή δεδομένων σε μικροσυσκευές
3. Προσθήκη υποσυνόλων
4. Μορφοποίηση Άξονα
5. Προσθήκη χρόνου σε μια διάσταση
6. Φίλτρο διάστασης
7. Μετονομασία πεδίων και μελών
8. Ταξινόμηση μελών

⁴ *Connecting to Excel*, 2 Σεπτεμβρίου 2008, 16:08, http://www.sisense.com/Support/connect_to_data.aspx

1) Εισαγωγή νέου Άξονα (Pivot)

Για να γίνει αυτό, θα πρέπει σε πρώτη φάση ο χρήστης να είναι συνδεδεμένος με μια πηγή δεδομένων, έτσι ώστε να τοποθετήσει ένα νέο άξονα (Pivot) σε ένα φύλλο πατώντας την επιλογή “New Widget” από την μπάρα εργαλείων και την επιλογή “Pivot Widget” από το επιμέρους μενού, όπως φαίνεται στο σχήμα 5-24.

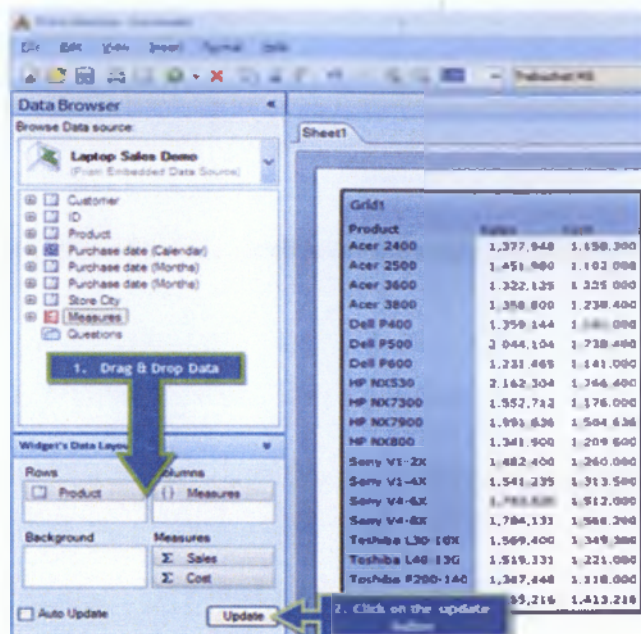


Σχήμα 5-24: Εισαγωγή νέου Άξονα

Πηγή: *Working with Pivot tables*, http://www.sisense.com/Support/working_with_pivots.aspx

2) Εισαγωγή δεδομένων σε μικροσυσκευές

Από τη στιγμή που θα γίνει η εισαγωγή του άξονα, ο χρήστης θα πρέπει να εισάγει στη μικροσυσκευή δεδομένα και αργότερα να τα προβάλει όπως αυτός θέλει. Για να γίνει αυτό, ο χρήστης σύρει την **Διάσταση Προϊόν (Product Dimension)** από τα στοιχεία του Data Browser και την αφήνει στον χώρο των Γραμμών (Rows) εντός του **Widgets Data Layout Panel**. Μετά ο χρήστης επιλέγει την **Διάσταση Μέτρα (Measures Dimension)**, την σύρει και την αφήνει στον χώρο των **Στηλών (Columns)** εντός του **Widgets Data Layout Panel** και αργότερα κάνοντας κλικ στην επιλογή “Update” γίνεται η προβολή των δεδομένων που έχει επιλέξει ο χρήστης όπως αυτός τα θέλει.



Σχήμα 5-25: Εισαγωγή δεδομένων σε Widget

Πηγή: : *Working with Pivot tables*, http://www.sisense.com/Support/working_with_pivots.aspx

3) Προσθήκη υποσυνόλων

Ακόμα, ο χρήστης χρησιμοποιώντας το Pivot έχει μια πληθώρα επιλογών όπως π.χ. την πρόσθεση υποσυνόλων για την διάσταση προϊόντα, κάνοντας δεξί κλικ πάνω της, μέσα στον χώρο του Pivot και επιλέγοντας στη συνέχεια την επιλογή “Sup Totals”.

4) Μορφοποίηση Άξονα

Εύκολα μπορεί να γίνει και η μορφοποίηση των κεφαλίδων του Pivot, επιλέγοντας σε πρώτη φάση όλες τις κεφαλίδες που περιέχονται στα κελιά όπως ακριβώς και σε ένα λογιστικό φύλλο. Η μορφοποίηση γίνεται χρησιμοποιώντας τα κατάλληλα κουμπιά μορφοποίησης από την γραμμή εργαλείων.

Product	Sales	Cost
Acer 2400	1,377,948	1,158,300
Acer 2500	1,451,980	1,102,000
Acer 3600	1,322,125	1,225,000
Acer 3800	1,358,800	1,238,400

Σχήμα 5-26: Μορφοποίηση Άξονα

Πηγή: *Working with Pivot tables*, http://www.sisense.com/Support/working_with_pivots.aspx

Μερικές φορές όμως ο χρήστης μπορεί να θέλει να κάνει μορφοποίηση της λίστας των μελών, χωρίς όμως να θέλει να επιλέξει όλα τα κελιά για να το κάνει. Για να το κάνει

αυτό, μπορεί να χρησιμοποιήσει την επιλογή “**span selection feature**”, όπου του δίνει τη δυνατότητα να επιλέξει γραμμές, στήλες και ομάδες, κάνοντας απλώς ένα κλικ. Για παράδειγμα, αν ο χρήστης θέλει να επιλέξει όλα τα μέλη της **Διάστασης Προϊόν (Product Dimension)** χωρίς όμως την κεφαλίδα, τοποθετεί τον κέρσορα μεταξύ των μελών (π.χ. Acer 2400, Acer 2500) της **Διάστασης Προϊόν**. Μερικές επιλογές **span** φαίνονται στο σχήμα 5-26.

Επιλογή μελών

Product	Sales	Cost
Acer 2400	1,377,948	1,158,300
Acer 2500	1,451,980	1,102,000
Acer 3600	1,322,125	1,225,000
Acer 3800	1,358,800	1,238,400

Επιλογή στήλης και τιμών

Product	Sales	Cost
Acer 2400	1,377,948	1,158,300
Acer 2500	1,451,980	1,102,000
Acer 3600	1,322,125	1,225,000
Acer 3800	1,358,800	1,238,400

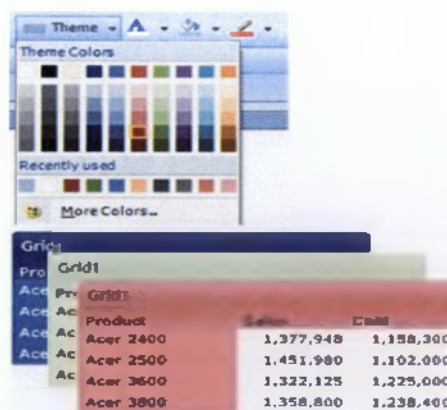
Επιλογή πλήρη σειρών και μερικών σειρών

Product	Sales	Cost
Acer 2400	1,377,948	1,158,300
Acer 2500	1,451,980	1,102,000
Acer 3600	1,322,125	1,225,000
Acer 3800	1,358,800	1,238,400

Σχήμα 5-27: Span Selection Feature

Πηγή : *Working with Pivot tables*, http://www.sisense.com/Support/working_with_pivots.aspx

Η μορφοποίηση δεν περιορίζεται μόνο στην επιλογή στηλών και σειρών, αλλά και στην επιλογή χρώματος, όπως απεικονίζεται παρακάτω.

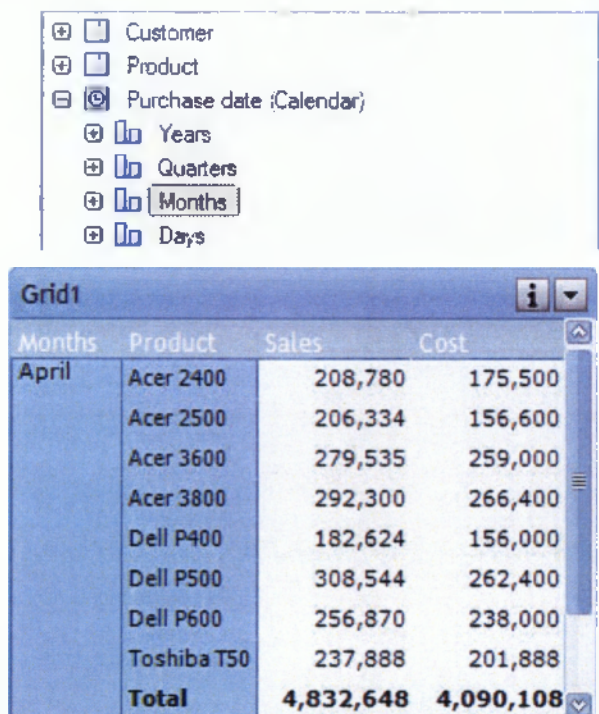


Σχήμα 5-28: Μορφοποίηση χρώματος

Πηγή: *Working with Pivot tables*, http://www.sisense.com/Support/working_with_pivots.aspx

5) Προσθήκη χρόνου σε μια διάσταση

Το Pivot έχει πολλές ενδιαφέρουσες επιλογές, όπως π.χ. η προσθήκη χρόνου στα μέλη της Διάστασης Ημερολόγιο (Calendar Dimension). Από το Data Browser ο χρήστης επιλέγει από την Διάσταση Purchase Date (Calendar) τους Μήνες (Months) και τους τοποθετεί, σύροντας τους πριν το στοιχείο Product, μέσα στην περιοχή των σειρών.



The screenshot shows the Data Browser interface with the 'Purchase date (Calendar)' dimension expanded to 'Months'. Below it, a PivotTable grid displays the following data:

Months	Product	Sales	Cost
April	Acer 2400	208,780	175,500
	Acer 2500	206,334	156,600
	Acer 3600	279,535	259,000
	Acer 3800	292,300	266,400
	Dell P400	182,624	156,000
	Dell P500	308,544	262,400
	Dell P600	256,870	238,000
	Toshiba T50	237,888	201,888
	Total		4,832,648

Σχήμα 5-29: Εισαγωγή Χρόνου

Πηγή: *Working with Pivot tables*, http://www.sisense.com/Support/working_with_pivots.aspx

6) Φίλτρο διάστασης

Κάνοντας δεξί κλικ στην Διάσταση Προϊόν (Product Dimension), επιλέγουμε από το υποσύνολο την επιλογή "Filter→ Apply New Filter". Με αυτή τη λειτουργία, μπορεί κάποιος να παρατηρήσει αλλαγές στην περιεκτικότητα του άξονα, ανάλογα με το τι ζητάει κανείς από το φίλτρο. Για παράδειγμα, από τη στιγμή που δίπλα στο Προϊόν (Product) έχουμε Μήνες (Months), το φίλτρο θα διενεργεί εβδομαδιαίους υπολογισμούς κάθε μήνα. Αν είχαμε τεταρτημόρια ή χρόνια πριν το Προϊόν (Product), τότε το φίλτρο θα το λάμβανε υπόψη και θα διενεργούσε εβδομαδιαίους υπολογισμούς ανά τρίμηνο ή ανά έτος.

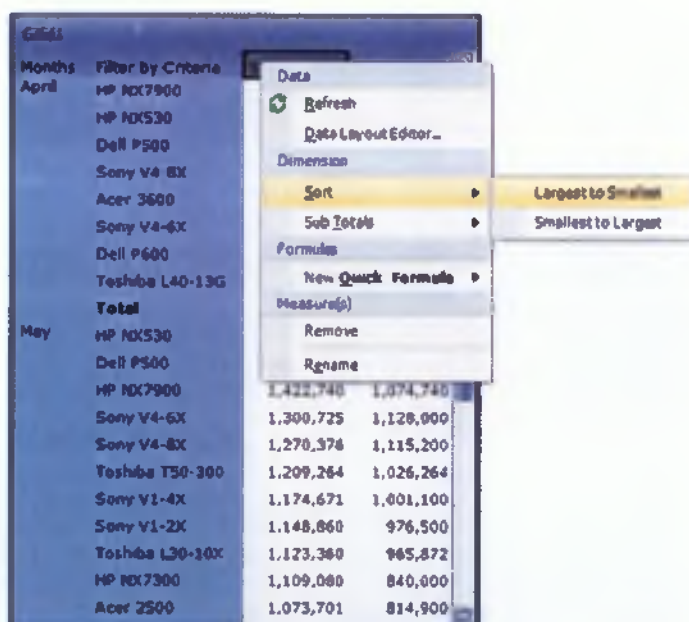
Αυτού του είδους το φιλτράρισμα ονομάζεται ευαίσθητο και βοηθά στην εξοικονόμηση χρόνου, όταν δημιουργούνται και χρησιμοποιούνται φίλτρα.

7) Μετονομασία πεδίων και μελών

Αυτή η λειτουργία γίνεται κάνοντας διπλό κλικ στο πεδίο με το όνομα **Filter by Criteria**, αλλάζοντας την λεζάντα σε **Product Sales > 200K** και πατώντας ENTER.

8) Ταξινόμηση μελών

Η επιλογή των προϊόντων, ανάλογα με τις πωλήσεις τους, γίνεται κάνοντας δεξί κλικ στο μέτρο **Πωλήσεις (Sales)** στο εσωτερικό του άξονα και από το αναδυόμενο παράθυρο γίνεται επιλογή της ταξινόμησης (Sort)→ Από μεγαλύτερη σε μικρότερη (Largest to Smallest).⁵



Months	Filter by Criteria			
April	HP NX7900			
	HP NX530			
	Dell P500			
	Sony V4-6X			
	Acer 3600			
	Sony V4-6X			
	Dell P600			
	Toshiba L40-13G			
	Total			
	May	HP NX530		
Dell P500				
HP NX7900		1,422,740	1,074,740	
Sony V4-6X		1,300,725	1,128,000	
Sony V4-6X		1,270,376	1,115,200	
Toshiba T50-300		1,209,264	1,026,264	
Sony V1-4X		1,174,671	1,001,100	
Sony V1-2X		1,148,860	976,500	
Toshiba L30-10X		1,123,360	965,872	
HP NX7300		1,109,080	840,000	
Acer 2500	1,073,701	814,900		

Σχήμα 5-30: Ταξινόμηση Μελών

Πηγή: *Working with Pivot tables*, http://www.sisense.com/Support/working_with_pivots.aspx

5.6 Συνεργασία αριθμητικών τύπων με το Prism

Οι τύποι (Formulas) είναι πολυδιάστατες εξισώσεις, που εκτελούν υπολογισμούς σε τιμές που βρίσκονται σε πηγές δεδομένων. Ένας τύπος μπορεί να περιέχει ένα ή περισσότερα από τα ακόλουθα:

- Λειτουργίες (Functions)
- Διαχειριστές (Operators)
- Σταθερούς Τελεστές (Constant Operands) π.χ. μέλη, επίπεδα, διαστάσεις, ομάδες
- Δυναμικούς Τελεστές (Dynamic Operands) π.χ. μετρούμενες τιμές, φίλτρα και άλλοι τύποι

⁵ *Working with Pivot tables*, 3 Σεπτεμβρίου 2008, 19:45, http://www.sisense.com/Support/working_with_pivots.aspx

Το Prism υποστηρίζει τις εξής κατηγορίες λειτουργιών:

- Αριθμητικές λειτουργίες
- Μετρούμενες τιμές
- Χρονικές λειτουργίες
- Λειτουργίες μελών⁶

Οι αριθμητικοί τύποι υπολογίζουν νέα μέτρα που δεν υπάρχουν στις πηγές δεδομένων, αλλά απαιτούνται για την ανάλυσή τους. Οι τιμές εξαρτώνται μόνο από το πλαίσιο των μικροσυσκευών (Widgets), στο οποίο είναι ενσωματωμένες και όχι από τον ίδιο τύπο. Ο ίδιος ακριβώς τύπος θα αποφέρει διαφορετικά αποτελέσματα, όταν αναπτυχθεί σε διαφορετικό περιβάλλον. Μερικοί απλοί τύποι μοιάζουν με τους παρακάτω:

$$[\text{Όριο}] = 1 - [\text{Κόστος}] / [\text{Πωλήσεις}]$$

$$[\text{ΦΠΑ}] = [\text{Πωλήσεις}] / 1,2 * 0,2$$

$$[\text{Κέρδη}] = [\text{Πωλήσεις}] * [\text{Όριο}]$$

Παράδειγμα 1:

Ορίζουμε $[\text{Κέρδος}] = [\text{Πωλήσεις}] - [\text{Κόστος}]$, όπου $[\text{Πωλήσεις}]$ και $[\text{Κόστος}]$ είναι μέτρα στην πηγή δεδομένων. Όταν το **Κέρδος** τοποθετηθεί στις στήλες (Columns) και οι **Μήνες** στις γραμμές (Rows), τότε θα εμφανιστεί στον άξονα (Pivot) το Κέρδος ανά Μήνα.

Profit by Month	
Months	Profit
April	742,540
May	3,347,987
June	582,445

Σχήμα 5-31: Κέρδος ανά Μήνα

Πηγή: *Numeric Formulas*, <http://sisense.com/help/learning-center/numeric-formulas/>

Αν ο ίδιος τύπος του κέρδους τοποθετηθεί στον άξονα, του οποίου οι γραμμές δείχνουν τα καταστήματα ανά πόλη, τότε ο άξονας θα εμφανίσει το κέρδος ανά πόλη.

Profit by City	
Store City	Profit
Boston	860,719
Houston	953,237
London	954,964
Los Angeles	934,953
New York	969,099

Σχήμα 5-32: Κέρδος ανά Πόλη

Πηγή: *Numeric Formulas*, <http://sisense.com/help/learning-center/numeric-formulas/>

⁶ *Introduction to Formulas*, 4 Σεπτεμβρίου 2008, 16:20, <http://sisense.com/help/learning-center/introduction-to-formulas/>

Αν ο ίδιος τύπος του κέρδους τοποθετηθεί σε διαφορετικούς άξονες (στους οποίους έχουν τεθεί διαφορετικά ερωτήματα), τότε θα υπάρξουν διαφορετικά αποτελέσματα.

Profit by City		Profit by Month and City			
Store City	Profit	Store City	Profit		
			April	May	June
Boston	860,719	Boston	142,423	606,639	111,657
Houston	953,237	Houston	143,700	686,232	123,305
London	954,964	London	161,235	682,172	111,557
Los Angeles	934,953	Los Angeles	151,743	671,200	112,010
New York	969,099	New York	143,439	701,744	123,916

Σχήμα 5-33: Κέρδος ανά Μήνα και Πόλη

Πηγή: *Numeric Formulas*, <http://sisense.com/help/learning-center/numeric-formulas/>

Το Prism υποστηρίζει πολλαπλούς τύπους, που περιέχουν πολλαπλούς τελεστές και μπορούν να ενσωματωθούν στο εσωτερικό του ίδιου τύπου. Αυτοί οι υπολογισμοί αποτελούνται από μέτρα και διαστάσεις. Αντίθετα με τους απλούς τύπους, οι πιο πολύπλοκοι τύποι εξαρτώνται από το φόντο των μικροσυσκευών και από το περιεχόμενο του τύπου. Οι πολύπλοκοι τύποι έχουν πολλά μέτρα που δεν είναι τελεστές και ένα μόνο μέτρο που είναι τελεστής.

Σύνταξη τύπου:

$([Dimension1Member]), \{ [Dimension2Member1], [Dimension2Member2], \dots \}, [Measure]$

Τελεστής μπορεί να είναι ένα μέλος ή περισσότερα μέλη από την ίδια διάσταση. Κάθε μέλος θα πρέπει να περικλείεται από τετράγωνες αγκύλες ([]). Τα πολλαπλά μέλη θα πρέπει να περικλείονται από αγκύλες ({}), ενώ κάθε μέλος εντός παρενθέσεως περικλείεται από τετράγωνες αγκύλες.

Παράδειγμα 2:

$[\text{Κέρδη στο Λονδίνο}] = ([\text{Κέρδη}], [\text{Λονδίνο}])$

Σε αυτή την έκφραση, ορίζεται ένα μέτρο που ονομάζεται **[Κέρδη στο Λονδίνο]**, το οποίο είναι συνάρτηση του μέτρου **[Κέρδη]** και του μέλους **[Λονδίνο]** της διάστασης **[Store City]**. Στον παρακάτω άξονα, παρουσιάζεται το **[Κέρδη στο Λονδίνο]** για κάθε μήνα.

Profit by Month		
Months	Profit	Profit in London
April	742,540	161,235
May	3,347,987	682,172
June	582,445	111,557

Σχήμα 5-34: Κέρδος ανά Μήνα

Πηγή: *Numeric Formulas*, <http://sisense.com/help/learning-center/numeric-formulas/>

Στον παρακάτω άξονα, ο ίδιος τύπος θα αποφέρει την ίδια αξία για όλες τις πόλεις που ορίζονται στις γραμμές, επειδή εξαρτάται από το [Λονδίνο], και όχι από μια πόλη γενικά.

Profit by City		
Store City	Profit	Profit in London
Boston	860,719	954,964
Houston	953,237	954,964
London	954,964	954,964
Los Angeles	934,953	954,964
New York	969,099	954,964

Σχήμα 5-35: Κέρδος ανά Πόλη

Πηγή: *Numeric Formulas*, <http://sisense.com/help/learning-center/numeric-formulas/>

[Κέρδη στο Λονδίνο το 200704] = ([Κέρδη], [Λονδίνο], [Απριλίου, 2007])

Σε αυτή την έκφραση, θα επικεντρωθούμε περισσότερο στο κέρδος από το προηγούμενο παράδειγμα για έναν συγκεκριμένο μήνα.

[Κέρδη στο LA & New York NY] = ([[New York], [Los Angeles]], [Κέρδη], [Απριλίου], [Μάιος])

Σε αυτή την έκφραση, υπολογίζουμε το κέρδος των δύο συγκεκριμένων πόλεων σε δύο συγκεκριμένους μήνες.⁷

5.7 Φίλτρα

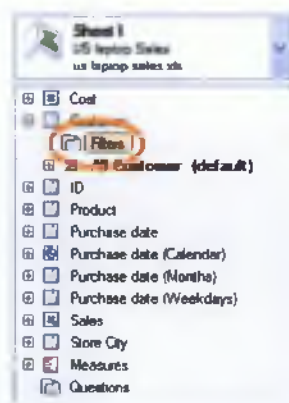
Τα φίλτρα είναι πολύ ισχυρά, αλλά απλά δομικά στοιχεία του Prism. Χρησιμοποιώντας τα, εξαπολύονται ατελείωτες δυνατότητες για την ενίσχυση των πινάκων οργάνων και μια βαθύτερη αναζήτηση στα δεδομένα. Τα φίλτρα επικεντρώνονται στους αριθμούς και στις στήλες, που κάποιος βλέπει και αναλύει. Μερικές φορές αν όχι πάντα, κάποιος μπορεί να μην θέλει να δει όλα τα στοιχεία μιας διάστασης. Για παράδειγμα, εάν

⁷ *Numeric Formulas*, 4 Σεπτεμβρίου 2008, 18:30, <http://sisense.com/help/learning-center/numeric-formulas/>

υπάρχουν δεδομένα πωλήσεων για χιλιάδες προϊόντα, το να τοποθετηθούν όλα μαζί σε έναν άξονα δεν θα ήταν πολύ χρήσιμο.

Ωστόσο, το φιλτράρισμα τμημάτων των δεδομένων είναι αρκετά χρήσιμο. Για παράδειγμα, μπορεί κάποιος να δει τις δέκα καλύτερες πωλήσεις προϊόντων, όλα τα προϊόντα που πουλήθηκαν αυτή την εβδομάδα, όλα τα προϊόντα που πουλήθηκαν παγκοσμίως, όλα τα προϊόντα τύπου “Laptop”, κλπ. Για να αποκτήσει κάποιος αυτά τα τμήματα πληροφορίας από όλη τη διάσταση, χρησιμοποιεί τα φίλτρα.

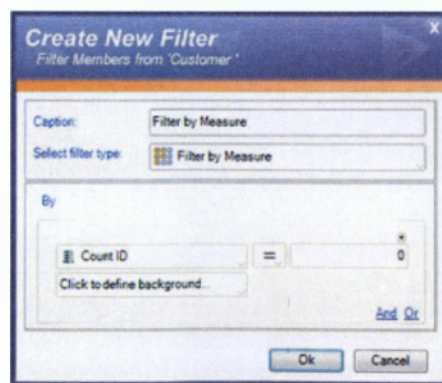
Η δημιουργία ενός νέου φίλτρου, γίνεται εάν αναπτυχθεί οτιδήποτε από το δέντρο διαστάσεων. Διαλέγοντας την διάσταση Customers και αναπτύσσοντάς την, βλέπει κάποιος ότι σε κάθε κόμβο υπάρχει η επιλογή Φίλτρο (Filter), όπως π.χ.



Σχήμα 5-36: Θέση φίλτρου σε διάσταση Customers

Πηγή: *Introduction to Filter*, <http://sisense.com/help/learning-center/introduction-to-filters/>

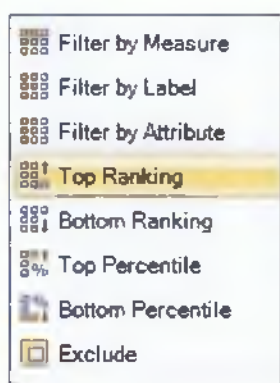
Κάνοντας δεξί κλικ πάνω σε αυτή την επιλογή, για να δημιουργηθεί το νέο φίλτρο, πρέπει να επιλεγθεί το “New Filter”, ανοίγοντας αυτόματα το παράθυρο διαλόγου “Create New Filter”.



Σχήμα 5-37: Δημιουργία νέου Φίλτρου

Πηγή: *Introduction to Filter*, <http://sisense.com/help/learning-center/introduction-to-filters/>

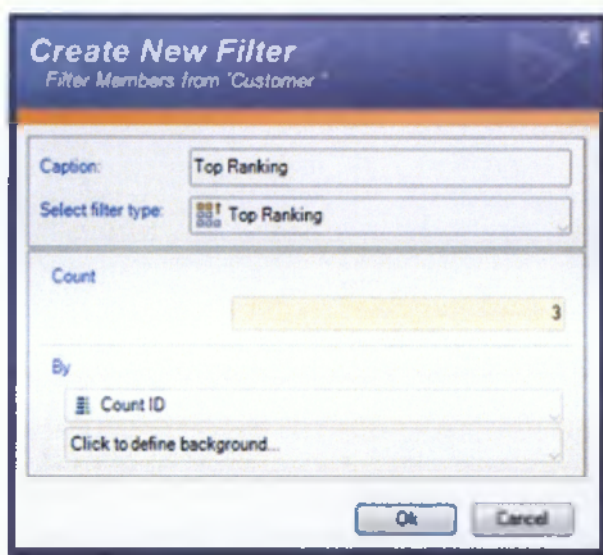
Η επιλογή “Select filter type”, εμφανίζει όλα τα διαθέσιμα είδη φίλτρων όπως π.χ.



Σχήμα 5-38: Είδη Φίλτρων

Πηγή: *Introduction to Filter*, <http://sisense.com/help/learning-center/introduction-to-filters/>

Όταν επιλεγθεί ένα φίλτρο, το παράθυρο διαλόγου αλλάζει ανάλογα με τον τύπο του φίλτρου που έχει επιλεγθεί. Στην συγκεκριμένη περίπτωση, ο χρήστης μπορεί να επιλέξει την επιλογή “Top Ranking”, με αποτέλεσμα το παράθυρο διαλόγου να μοιάζει κάπως έτσι:



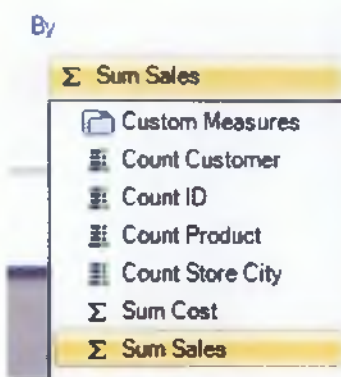
Σχήμα 5-39: Top Ranking

Πηγή: *Introduction to Filter*, <http://sisense.com/help/learning-center/introduction-to-filters/>

Το πεδίο “Count” επιτρέπει στον χρήστη να επιλέξει τον αριθμό των κορυφαίων πελατών της κατάταξης. Για παράδειγμα, Top 10, Top 5, Top 3, κλπ. Ας υποθέσουμε ότι ο χρήστης επιλέγει να δει τους τρεις κορυφαίους πελάτες (Top 3).

Στο πλαίσιο “By”, ο χρήστης μπορεί να επιλέξει το στυλ της κατάταξης όπως π.χ. οι δέκα κορυφαίοι πελάτες από τις πωλήσεις (Top 10 By Sales) ή οι δέκα κορυφαίες δαπάνες

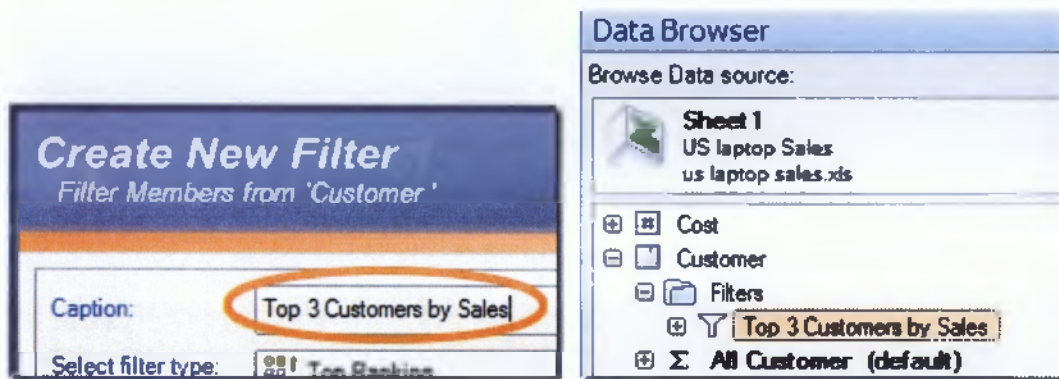
(Top 10 By Costs). Ας υποθέσουμε ότι ο χρήστης επιλέγει να δει τις συνολικές πωλήσεις (Sum Sales):



Σχήμα 5-40: Πλαίσιο By

Πηγή: *Introduction to Filter*, <http://sisense.com/help/learning-center/introduction-to-filters/>

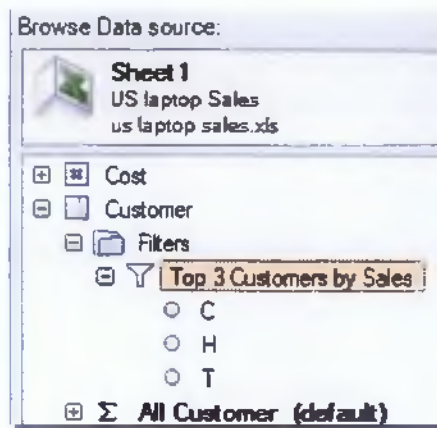
Στο πλαίσιο “Click to define background.....” ο χρήστης έχει τη δυνατότητα να περιορίσει το φίλτρο υπό ορισμένες προϋποθέσεις, ανάλογα με το τι θέλει να κάνει ο ίδιος. Εφόσον τελειώσει με τις προϋποθέσεις, ο χρήστης μπορεί να δώσει ότι όνομα θέλει στο νέο φίλτρο που δημιούργησε, όπως φαίνεται στο παρακάτω σχήμα:



Σχήμα 5-41: Περιορισμός Φίλτρου

Πηγή: *Introduction to Filter*, <http://sisense.com/help/learning-center/introduction-to-filters/>

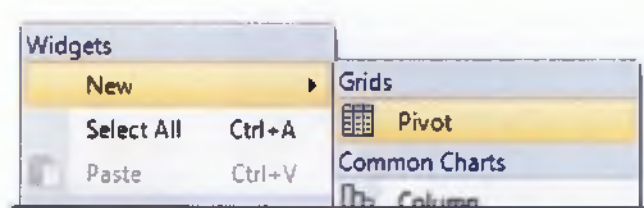
Το νέο φίλτρο, όπως φαίνεται στο παραπάνω σχήμα, έχει προστεθεί κάτω από τον κόμβο “Filters” και αναπτύσσοντας το νέο φίλτρο ο χρήστης μπορεί να δει τα αποτελέσματα:



Σχήμα 5-42: Νέο Φίλτρο

Πηγή: *Introduction to Filter*, <http://sisense.com/help/learning-center/introduction-to-filters/>

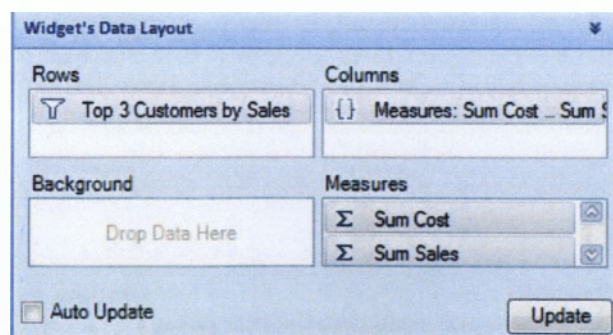
Η χρήση του φίλτρου στον πίνακα οργάνων γίνεται με την δημιουργία ενός Συγκεντρωτικού Πίνακα, κάνοντας δεξί κλικ στον καμβά και επιλέγοντας “New”→ “Pivot”:



Σχήμα 5-43: Δημιουργία νέου Άξονα

Πηγή: *Introduction to Filter*, <http://sisense.com/help/learning-center/introduction-to-filters/>

Για να δει ο χρήστης τους τρεις κορυφαίους πελάτες, σύρει το νέο φίλτρο που δημιούργησε στις γραμμές της διάταξης των δεδομένων και τα δύο μέτρα “Sum Cost” και “Sum Sales” στο πεδίο “Measures” της διάταξης των δεδομένων, όπως φαίνεται στο σχήμα 5-44.



Σχήμα 5-44: Εισαγωγή Φίλτρου

Πηγή: *Introduction to Filter*, <http://sisense.com/help/learning-center/introduction-to-filters/>

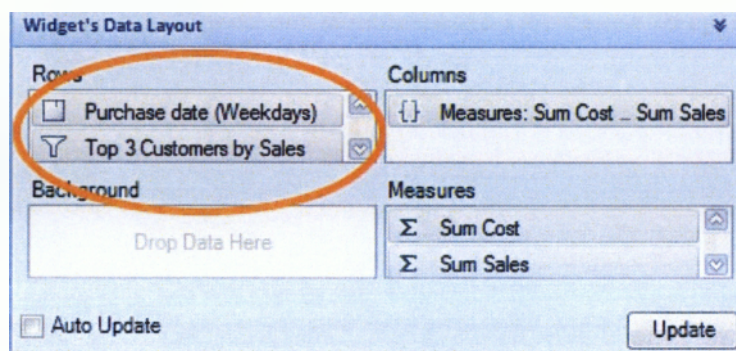
Πατώντας ο χρήστης το κουμπί Update, το αποτέλεσμα που θα δει θα μοιάζει με το εξής:

Pivot1		
Top 3 Customers by Sales	Sum Cost	Sum Sales
C	243,664	289,062
H	246,152	292,916
T	261,654	310,716

Σχήμα 5-45: Αποτελέσματα Φίλτρου

Πηγή: *Introduction to Filter*, <http://sisense.com/help/learning-center/introduction-to-filters/>

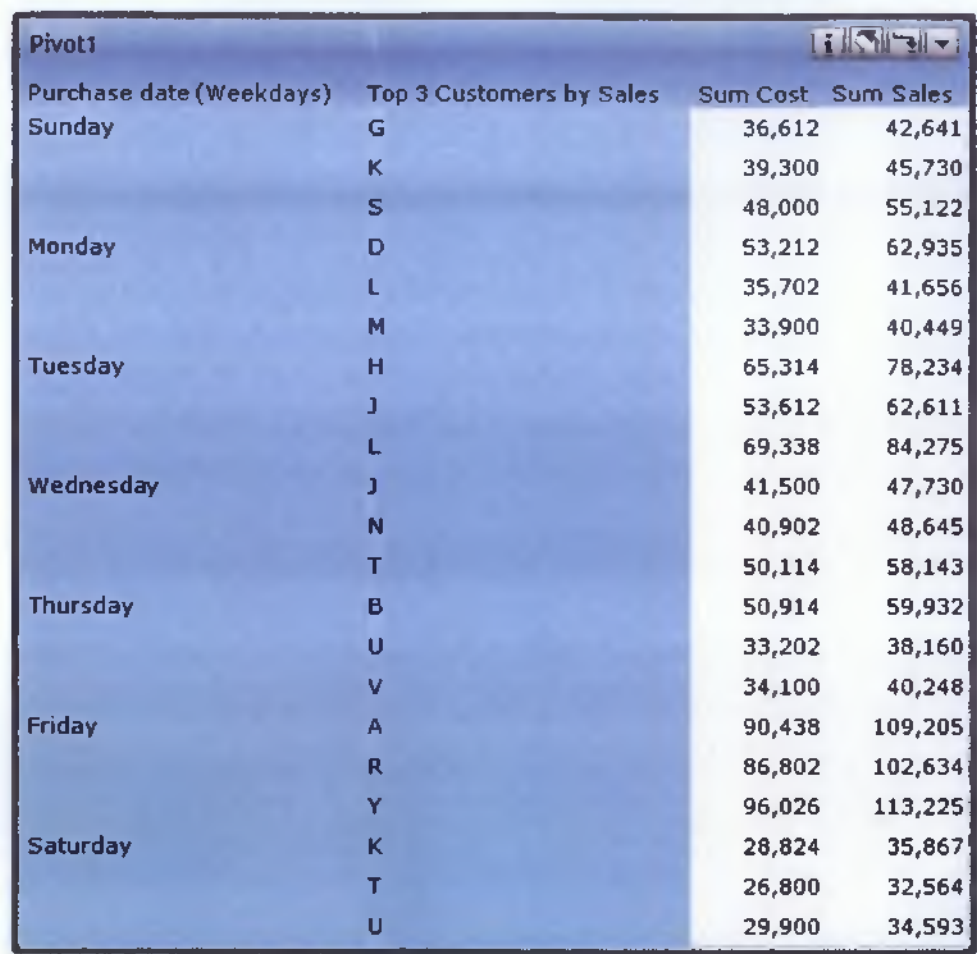
Ας υποθέσουμε ότι ο χρήστης δεν θέλει να δει τους τρεις κορυφαίους πελάτες των συνολικών πωλήσεων, αλλά θέλει να δει τους τρεις κορυφαίους πελάτες για κάθε ημέρα της εβδομάδας. Από το δέντρο διαστάσεων, σύρει το “Purchase Date (Weekdays)”, αφήνοντάς το μέσα στις γραμμές (Rows) του Widgets Data Layout, πάνω από το φίλτρο με τους τρεις κορυφαίους πελάτες (Top 3 Customers By Sales) που τοποθέτησε ο χρήστης εκεί νωρίτερα, όπως φαίνεται στο σχήμα 5-46.



Σχήμα 5-46: Εισαγωγή διάστασης Χρόνου

Πηγή: *Introduction to Filter*, <http://sisense.com/help/learning-center/introduction-to-filters/>

Κάνοντας κλικ στο κουμπί Update, το αποτέλεσμα θα είναι το εξής:



Purchase date (Weekdays)	Top 3 Customers by Sales	Sum Cost	Sum Sales
Sunday	G	36,612	42,641
	K	39,300	45,730
	S	48,000	55,122
Monday	D	53,212	62,935
	L	35,702	41,656
	M	33,900	40,449
Tuesday	H	65,314	78,234
	J	53,612	62,611
	L	69,338	84,275
Wednesday	J	41,500	47,730
	N	40,902	48,645
	T	50,114	58,143
Thursday	B	50,914	59,932
	U	33,202	38,160
	V	34,100	40,248
Friday	A	90,438	109,205
	R	86,802	102,634
	Y	96,026	113,225
Saturday	K	28,824	35,867
	T	26,800	32,564
	U	29,900	34,593

Σχήμα 5-47: Αποτελέσματα

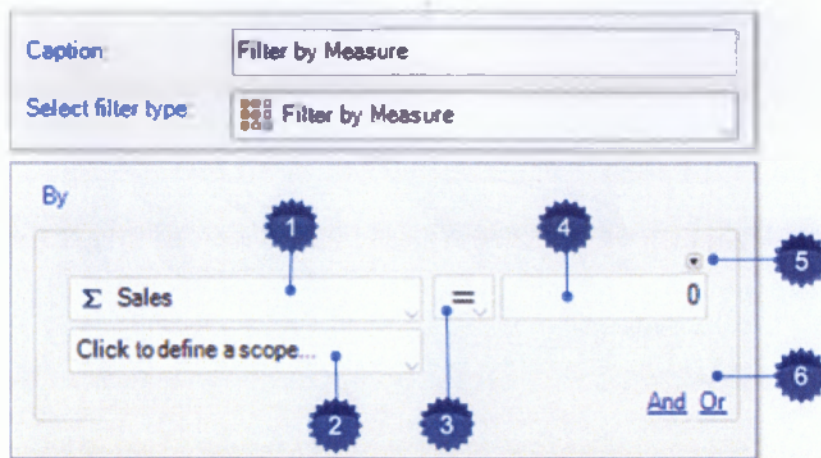
Πηγή: *Introduction to Filter*, <http://sisense.com/help/learning-center/introduction-to-filters/>

Βλέποντας το παραπάνω σχήμα, παρατηρούμε ότι οι τρεις κορυφαιοί πελάτες αλλάζουν από μέρα σε μέρα, χωρίς ο χρήστης να έχει αλλάξει κάτι στο φίλτρο.⁸

Τα είδη των φίλτρων χωρίζονται στα παρακάτω:

- Το **Φιλτράρισμα με Μέτρο (Filter By Measure)** δίνει σαν αποτέλεσμα τα μέλη που πληρούν μια αριθμητική προϋπόθεση. Η προϋπόθεση του φίλτρου έχει δύο τελεστές. Ο ένας είναι πάντα ένα Μέτρο (Measure) και ο δεύτερος είναι μία Αριθμητική Τιμή ή ένα Μέτρο.

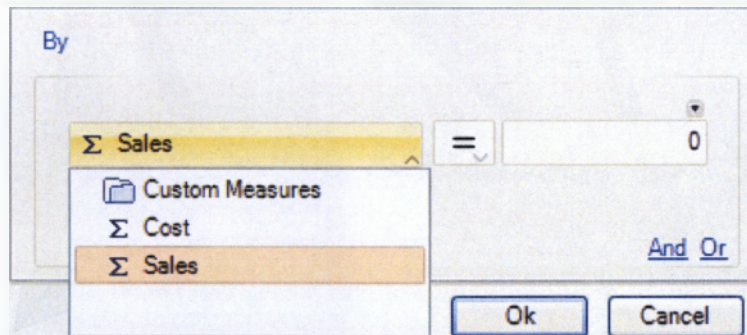
⁸ *Introduction to Filter*, 5 Σεπτεμβρίου 2008, 15:20, <http://sisense.com/help/learning-center/introduction-to-filters/>



Σχήμα 5-48: Φιλτράρισμα με Μέτρο

Πηγή: *Filter by Measure*, <http://sisense.com/help/learning-center/filter-by-measure/>

1. Επιλογή Μέτρου / Τύπου (Selection Measure / Formula)

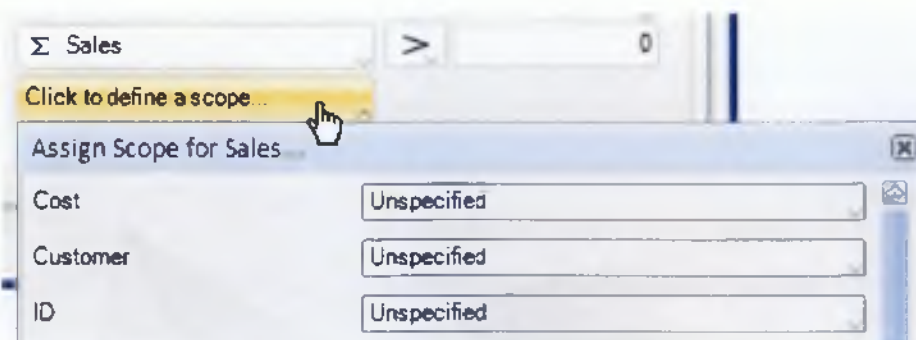


Σχήμα 5-49: Επιλογή Μέτρου / Τύπου

Πηγή: *Filter by Measure*, <http://sisense.com/help/learning-center/filter-by-measure/>

2. Πεδίο εφαρμογής Μέτρου / Τύπου (Scope on Measure / Formula)

Επιτρέπει στον χρήστη να εφαρμόσει ένα μέτρο στο πεδίο εφαρμογής από άλλες διαστάσεις, οι οποίες περιλαμβάνονται στο φίλτρο.

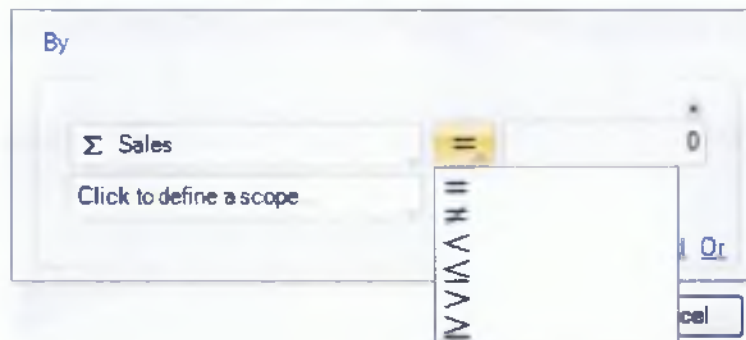


Σχήμα 5-50: Πεδίο εφαρμογής / τύπου

Πηγή: *Filter by Measure*, <http://sisense.com/help/learning-center/filter-by-measure/>

3. Κριτήριο Box (Box Criterion)

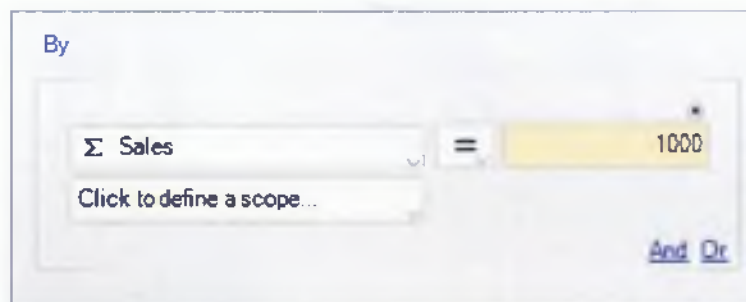
Δεύτερος Τελεστής (Μέτρο ή Αριθμητική Τιμή)



Σχήμα 5-51: Κριτήριο Box

Πηγή: *Filter by Measure*, <http://sisense.com/help/learning-center/filter-by-measure/>

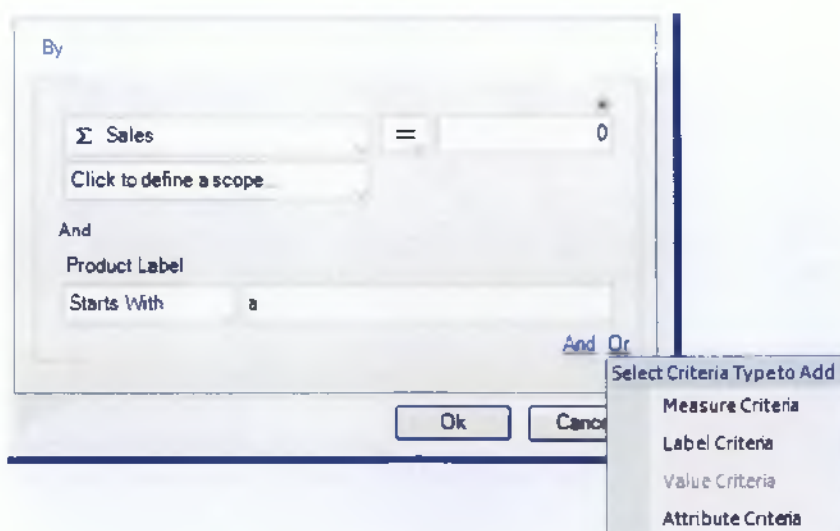
4. Δεύτερος Τελεστής (Μέτρο ή Αριθμητική Τιμή)



Σχήμα 5-52: Δεύτερος τελεστής

Πηγή: *Filter by Measure*, <http://sisense.com/help/learning-center/filter-by-measure/>

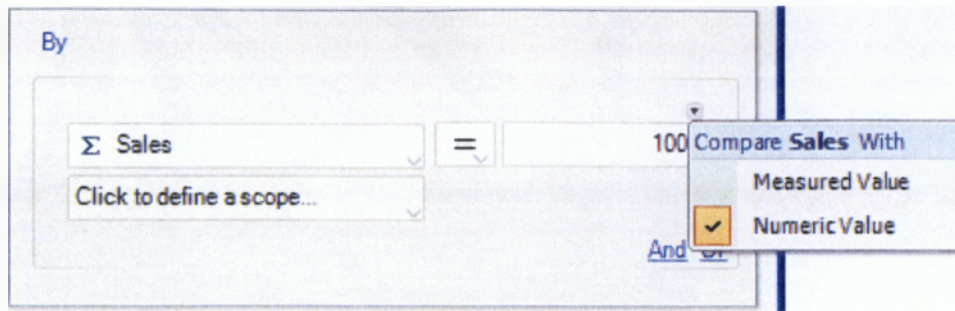
5. Πρόσθεση νέου κριτηρίου



Σχήμα 5-53: Πρόσθεση νέου κριτηρίου

Πηγή: *Filter by Measure*, <http://sisense.com/help/learning-center/filter-by-measure/>

6. Αλλαγή δεύτερου τελεστή⁹



Σχήμα 5-54: Αλλαγή δεύτερου τελεστή

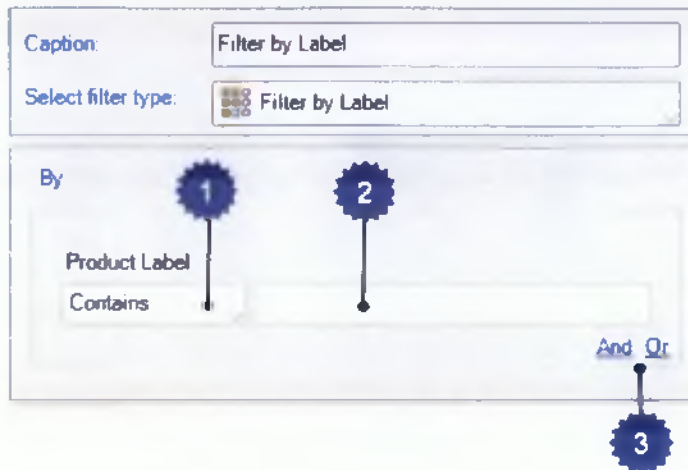
Πηγή: *Filter by Measure*, <http://sisense.com/help/learning-center/filter-by-measure/>

- Το **Φιλτράρισμα με Ετικέτα (Filter By Label)** δίνει τη δυνατότητα της ανάκτησης των μελών, τα οποία πληρούν ένα συγκεκριμένο Κριτήριο Κειμένου (Textual Criterion) στο όνομά τους. Το φίλτρο υποστηρίζει τα ακόλουθα κριτήρια:

- Contains “ABC” – Ανάκτηση μελών που περιέχουν “ABC”
- Doesn’t Contain “ABC” – Ανάκτηση μελών που δεν περιέχουν “ABC” και αποκλεισμός εκείνων που περιέχουν
- Start’s with “A” – Ανάκτηση μελών που ξεκινούν με “A”
- Doesn’t start with “A” – Ανάκτηση μελών που δεν ξεκινούν με “A” και αποκλεισμός εκείνων που ξεκινούν
- End’s with “Q” – Ανάκτηση μελών που τελειώνουν σε “Q”
- Doesn’t end with “Q” – Ανάκτηση μελών που δεν τελειώνουν σε “Q” και αποκλεισμός εκείνων που τελειώνουν
- Equals with “Frank” – Ανάκτηση μελών που ισούνται με “Frank”
- Doesn’t equal with “Frank” – Ανάκτηση μελών που δεν ισούνται με “Frank” και αποκλεισμός εκείνων που ισούνται

⁹ *Filter by Measure*, 6 Σεπτεμβρίου 2008, 11:45, <http://sisense.com/help/learning-center/filter-by-measure/>

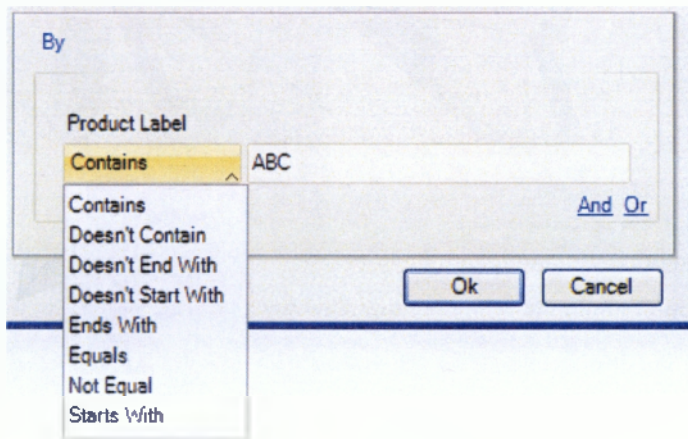
User Interface



Σχήμα 5-55: User Interface

Πηγή: *Filter By Label*, <http://sisense.com/help/learning-center/filter-by-label/>

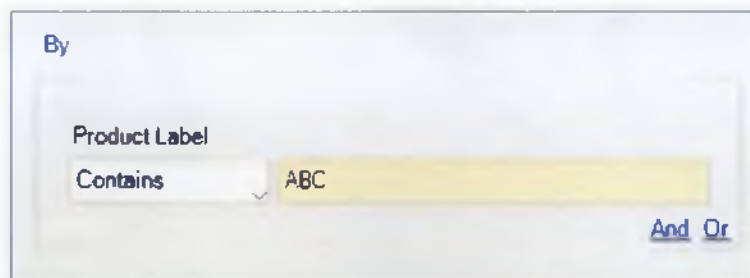
1. Κριτήριο



Σχήμα 5-56: Εισαγωγή κριτηρίου

Πηγή: *Filter By Label*, <http://sisense.com/help/learning-center/filter-by-label/>

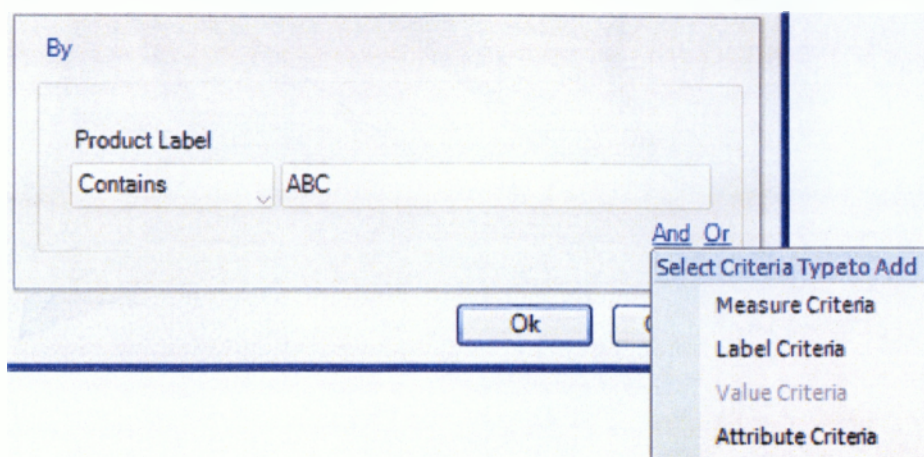
2. Ετικέτα



Σχήμα 5-57: Label

Πηγή: *Filter By Label*, <http://sisense.com/help/learning-center/filter-by-label/>

3. Πρόσθεση νέου κριτηρίου

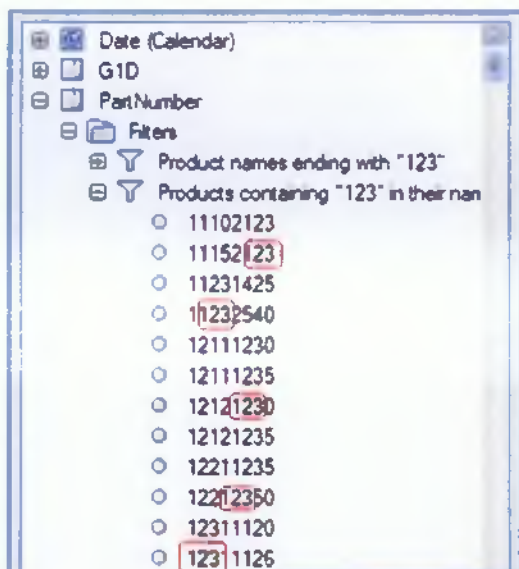


Σχήμα 5-58: Πρόσθεση νέου κριτηρίου

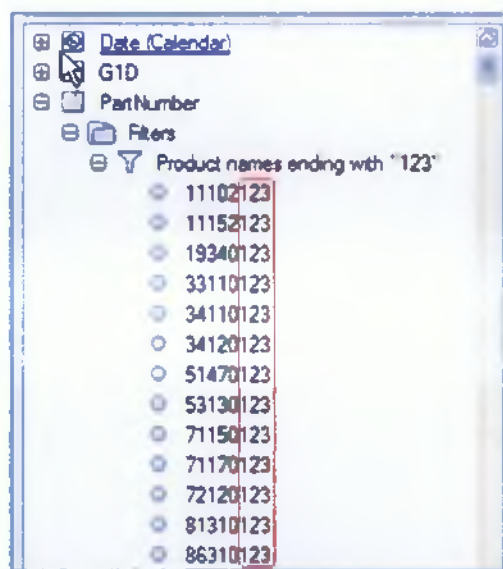
Πηγή: *Filter By Label*, <http://sisense.com/help/learning-center/filter-by-label/>

Το Prism δίνει τη δυνατότητα στο χρήστη να συνδυάσει διάφορους τύπους Κριτηρίων Φίλτρων (Filter Criteria) μέσα σε ένα Φίλτρο, όπως φαίνεται στο σχήμα 5-58.¹⁰

Μέλη που περιέχουν "123"



Μέλη που τελειώνουν σε "123"



Σχήμα 5-59: Αποτελέσματα κριτηρίων

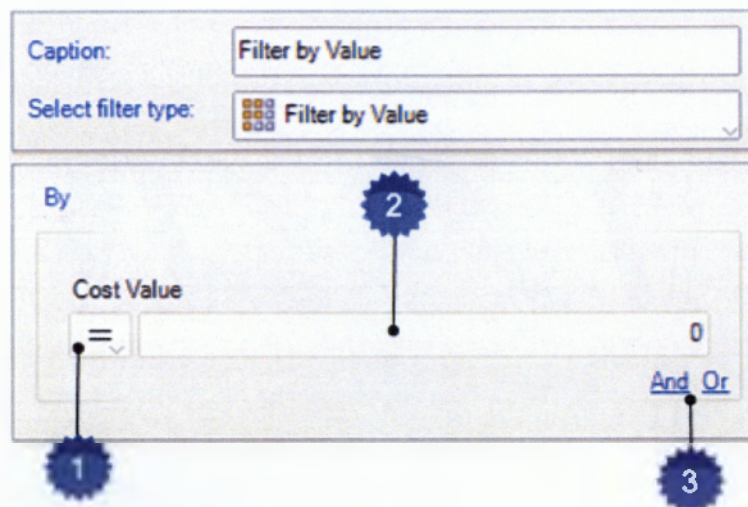
Πηγή: *Filter By Label*, <http://sisense.com/help/learning-center/filter-by-label/>

- Το **Φιλτράρισμα με βάση την Αξία (Filter By Value)** δίνει τη δυνατότητα της ανάκτησης των μελών, τα οποία πληρούν ένα συγκεκριμένο αριθμητικό κριτήριο στο όνομά τους. Το φίλτρο υποστηρίζει τα ακόλουθα κριτήρια:

¹⁰ *Filter By Label*, 8 Σεπτεμβρίου 2008, 15:48, <http://sisense.com/help/learning-center/filter-by-label/>

- Equals to (=) – Ισούται με (=)
- Doesn't equal to (≠) – Δεν ισούται με (≠)
- Greater than (>) – Μεγαλύτερο από (>)
- Less than (<) – Μικρότερο από (<)
- Greater than or equal to (>=) – Μεγαλύτερο ή ίσο από
- Less than or equal to (<=) – Μικρότερο ή ίσο από

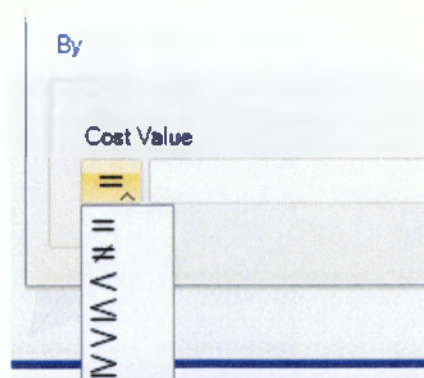
User Interface



Σχήμα 5-60: User Interface

Πηγή: *Filter By Value*, <http://sisense.com/help/learning-center/filter-by-value/>

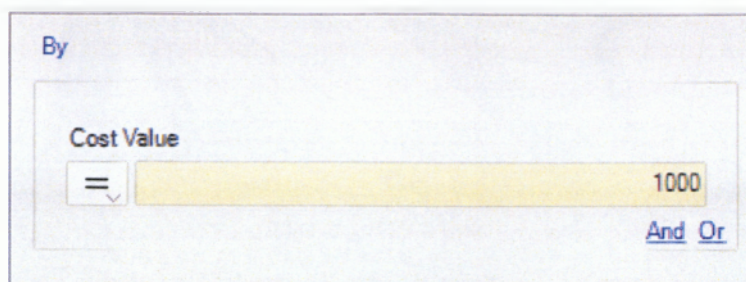
1. Κριτήριο



Σχήμα 5-61: Κριτήριο

Πηγή: *Filter By Value*, <http://sisense.com/help/learning-center/filter-by-value/>

2. Αξία Box

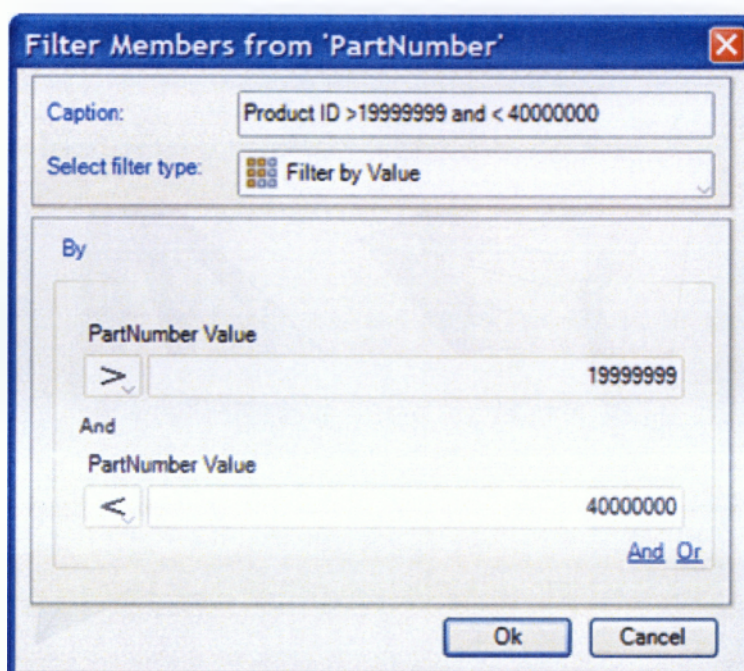


Σχήμα 5-62: Box Value

Πηγή: *Filter By Value*, <http://sisense.com/help/learning-center/filter-by-value/>

Παράδειγμα:

Το φιλτράρισμα με βάση την αξία θα ανακτήσει τα προϊόντα, τα οποία έχουν ProductID μεταξύ (και συμπεριλαμβανομένων) 20000000 και 39999999.¹¹



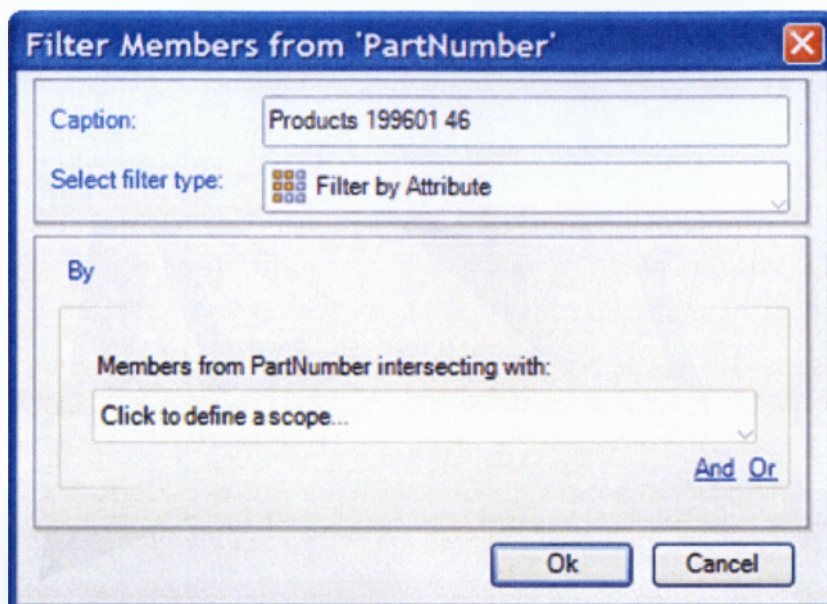
Σχήμα 5-63: Φιλτράρισμα με βάση την αξία

Πηγή: *Filter By Value*, <http://sisense.com/help/learning-center/filter-by-value/>

- Το **Φιλτράρισμα με Ιδιότητα (Filter by Attribute)** δίνει τη δυνατότητα της ανάκτησης των μελών, που έχουν ένα υπάρχον χαρακτηριστικό (πραγματική αξία σε μια άλλη διάσταση) στην πηγή δεδομένων. Από το Data Browser, ο χρήστης επιλέγει μια διάσταση που θέλει να φιλτράρει, ανοίγει το παράθυρο του φίλτρου (Filter's Window) και ορίζει ένα όνομα για το

¹¹ *Filter By Value*, 9 Σεπτεμβρίου 2008, 18:20, <http://sisense.com/help/learning-center/filter-by-value/>

φίλτρο. Κάνοντας κλικ στο πλαίσιο “Select filter type”, εμφανίζεται μια αναπτυσσόμενη λίστα από την οποία ο χρήστης επιλέγει το “Filter by Attribute”, όπως φαίνεται στο σχήμα 5-64.



Σχήμα 5-64: PartNumber

Πηγή: *Filter By Attribute*, <http://sisense.com/help/learning-center/filter-by-attribute/>

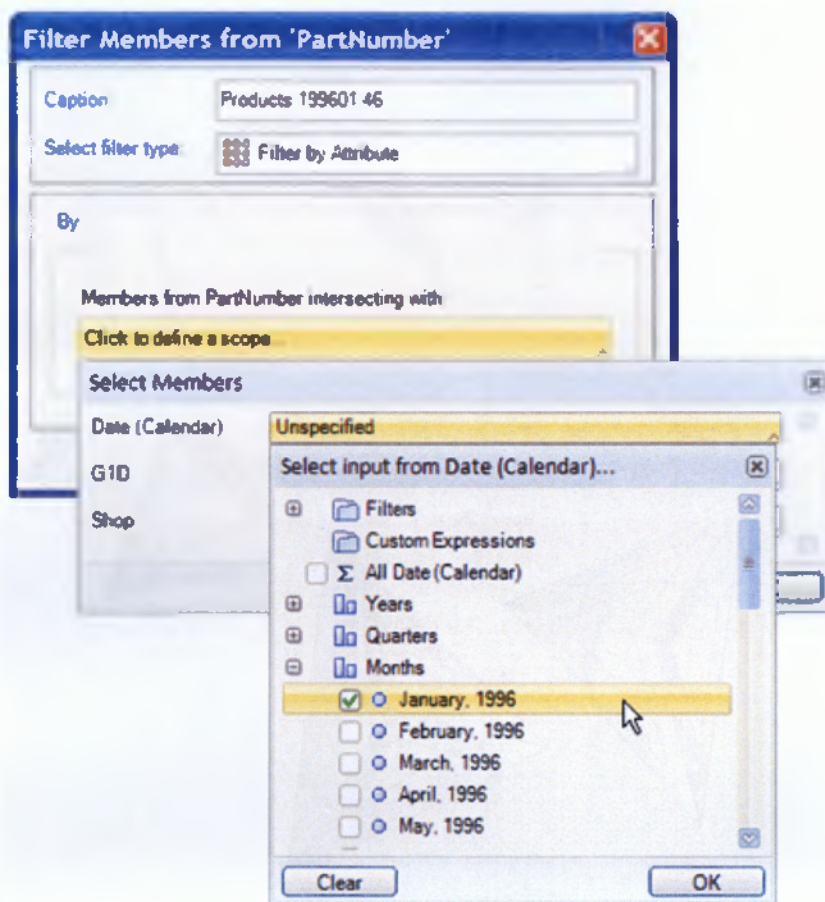
Κάνοντας κλικ στο “Define a Scope”, το παράθυρο θα αλλάξει με το ακόλουθο:



Σχήμα 5-65: Define a Scope

Πηγή: *Filter By Attribute*, <http://sisense.com/help/learning-center/filter-by-attribute/>

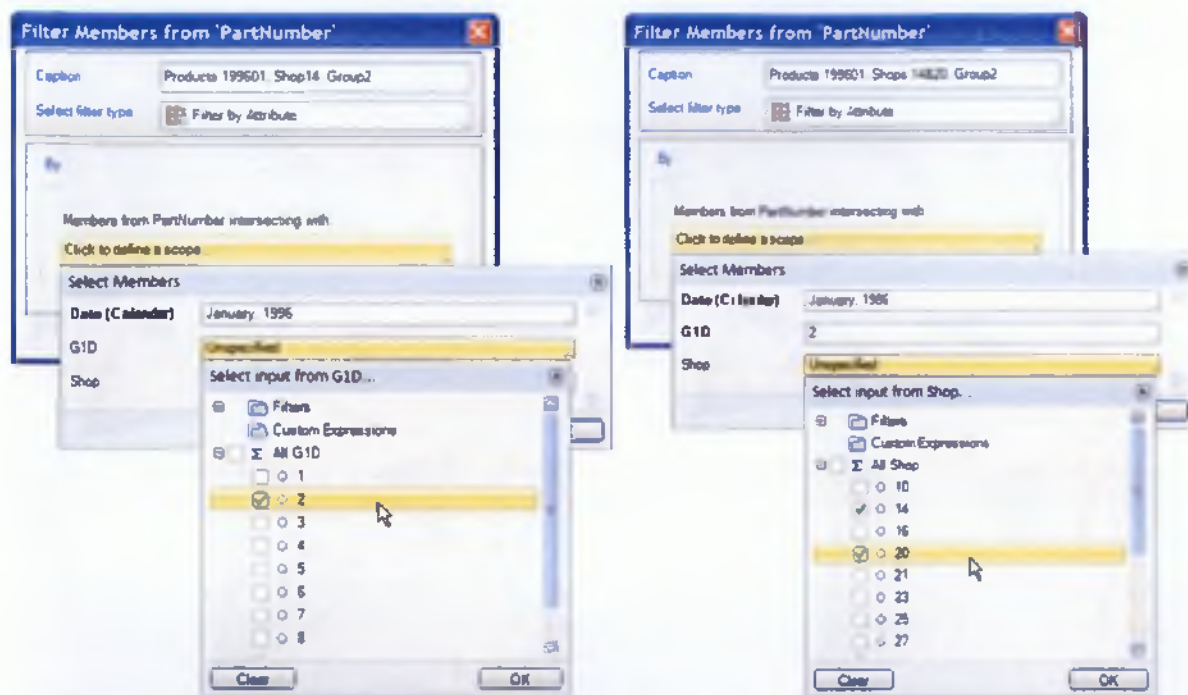
Θα ανοίξει μια αναπτυσσόμενη λίστα, η οποία περιλαμβάνει όλα τα δεδομένα των διαστάσεων του Data Browser. Από αυτή τη λίστα γίνεται η επιλογή του Πεδίου Εφαρμογής (Scope). Κάνοντας κλικ στο πλαίσιο Date (Calendar), εφαρμόζεται το drill-down στο επίπεδο Μήνας (Month) και γίνεται η επιλογή των απαιτούμενων μηνών, όπως φαίνεται παρακάτω:



Σχήμα 5-66: Επιλογή Μηνών

Πηγή: *Filter By Attribute*, <http://sisense.com/help/learning-center/filter-by-attribute/>

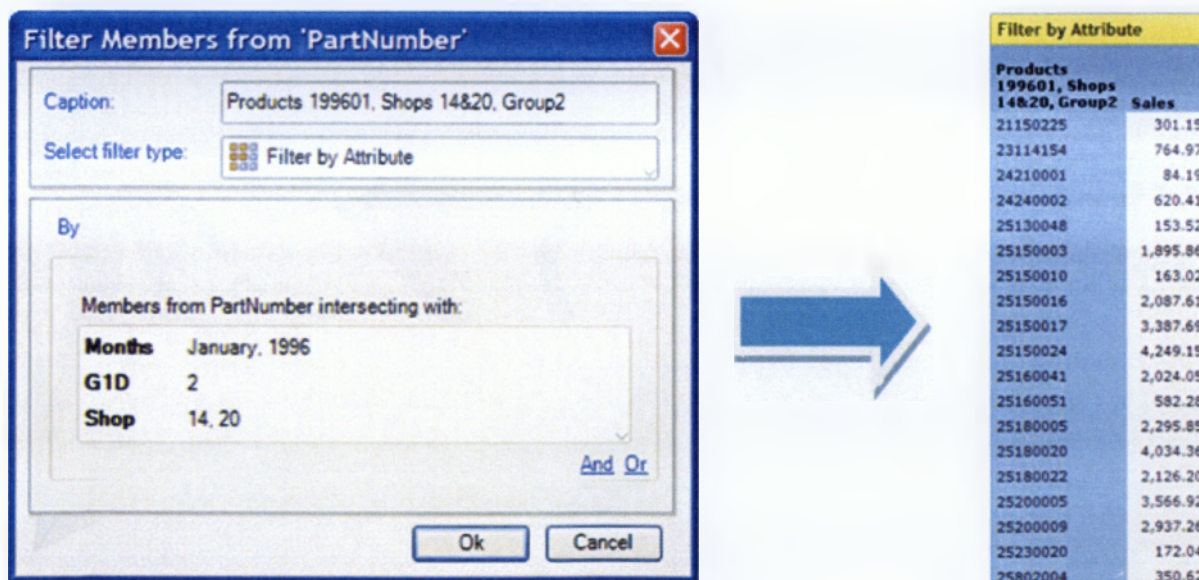
Ακολουθείται η ίδια διαδικασία για κάθε ένα μέλος στην λίστα.



Σχήμα 5-67: Εφαρμογή διαδικασίας

Πηγή: *Filter By Attribute*, <http://sisense.com/help/learning-center/filter-by-attribute/>

Το τελικό φίλτρο θα μοιάζει με το εξής:¹²



Σχήμα 5-68: Τελικό Φίλτρο

Πηγή: *Filter By Attribute*, <http://sisense.com/help/learning-center/filter-by-attribute/>

¹² *Filter By Attribute*, 9 Σεπτεμβρίου 2008, 19:45, <http://sisense.com/help/learning-center/filter-by-attribute/>

5.8 Διαδικασία εγκατάστασης

Το πρώτο βήμα για την εγκατάσταση του εργαλείου Prism είναι η εγγραφή του χρήστη στο επίσημο site του Prism, χωρίς καμία χρέωση. Στην επίσημη σελίδα, υπάρχει η επιλογή “Download Now”, όπου πατώντας την θα οδηγήσει τον χρήστη στην σχετική φόρμα εγγραφής. Εφόσον ο χρήστης συμπληρώσει τα στοιχεία του στη φόρμα και αποδεχτεί τους όρους εγγραφής, πρέπει να κάνει κλικ πάνω στην επιλογή “Submit”.

Το επόμενο βήμα είναι να συνδεθεί με το λογαριασμό του προσωπικού του e-mail και να ανοίξει το μήνυμα με την ένδειξη “Sisense Team”. Στο μήνυμα που εμφανίζεται, πρέπει να κάνει κλικ στην επιλογή “Activate”. Εφόσον γίνει η ενεργοποίηση του λογαριασμού, θα ανοιχτεί αυτόματα μια σελίδα, από την οποία ο χρήστης θα μπορεί να “κατεβάσει” (Download) το Prism. Αφού “κατεβάσει” το αρχείο “PrismDesktopSetup”, κάνοντας διπλό κλικ “τρέχει” το αρχείο και ακολουθεί τις οδηγίες εγκατάστασης που εμφανίζονται στην οθόνη.

ΚΕΦΑΛΑΙΟ 6

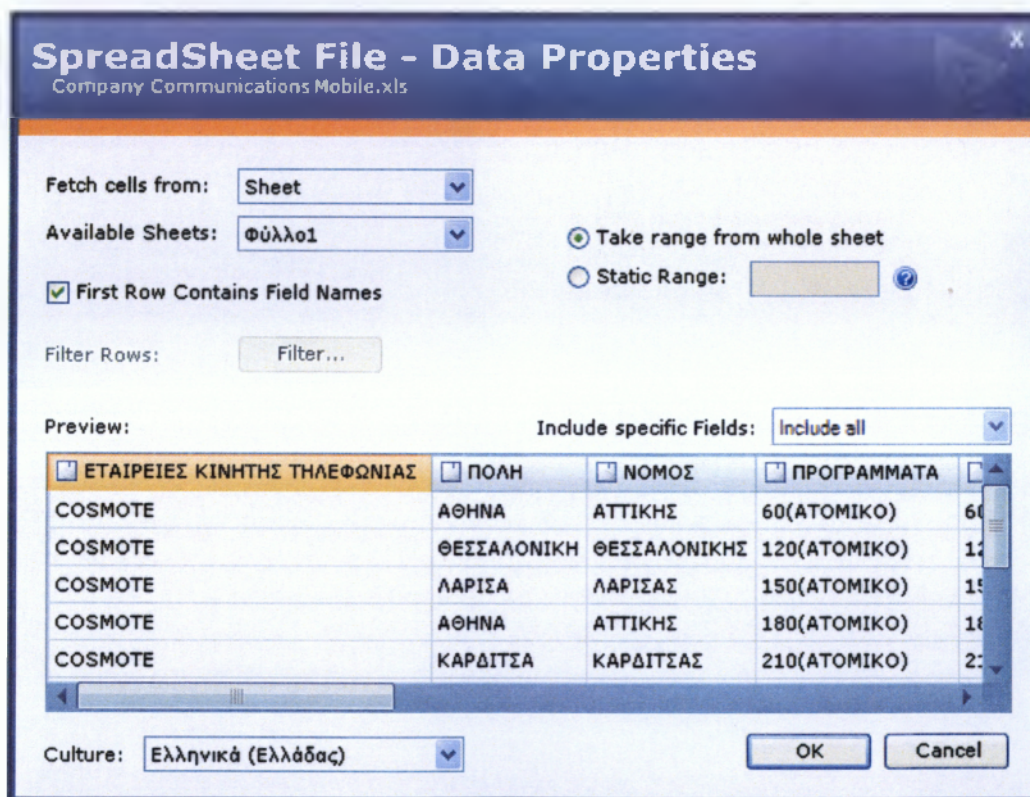
ΕΦΑΡΜΟΓΗ ΕΡΓΑΛΕΙΟΥ PRISM

6.1 Επεξεργασία δεδομένων εταιρειών κινητής τηλεφωνίας

Σε αυτό το κεφάλαιο, ακολουθούν κάποια παραδείγματα για την καλύτερη κατανόηση της λειτουργίας και της εφαρμογής του εργαλείου Prism. Όλα τα παραδείγματα περιέχουν δεδομένα, που περιέχονται σε βάση δεδομένων Excel.

Παράδειγμα 1:

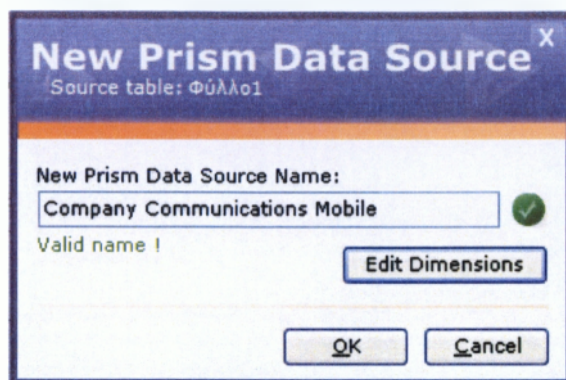
Στο συγκεκριμένο παράδειγμα, υπάρχουν δεδομένα εταιρειών κινητής τηλεφωνίας. Αρχικά, εισάγουμε την πηγή δεδομένων στο Prism (η διαδικασία αναφέρεται αναλυτικά στο Κεφάλαιο 5) και εμφανίζεται στην οθόνη το παράθυρο διαλόγου “Spreadsheet File – Data Properties”, από το οποίο έχουμε μια πρώτη όψη των δεδομένων.



Σχήμα 6-1: Εισαγωγή δεδομένων εταιρειών κινητής τηλεφωνίας

Πηγή: Ιδία Έρευνα

Πατώντας το κουμπί OK, εμφανίζεται το παράθυρο διαλόγου “New Prism Data Source”, από το οποίο μπορούμε να ονομάσουμε την πηγή δεδομένων, όπως επιθυμούμε. Στο συγκεκριμένο παράδειγμα ονομάζουμε την πηγή δεδομένων “Company Communications Mobile”, όπως φαίνεται στο σχήμα 6-2 παρακάτω.



Σχήμα 6-2: Ονομασία πηγής δεδομένων εταιρειών κινητής τηλεφωνίας

Πηγή: Ιδία Έρευνα

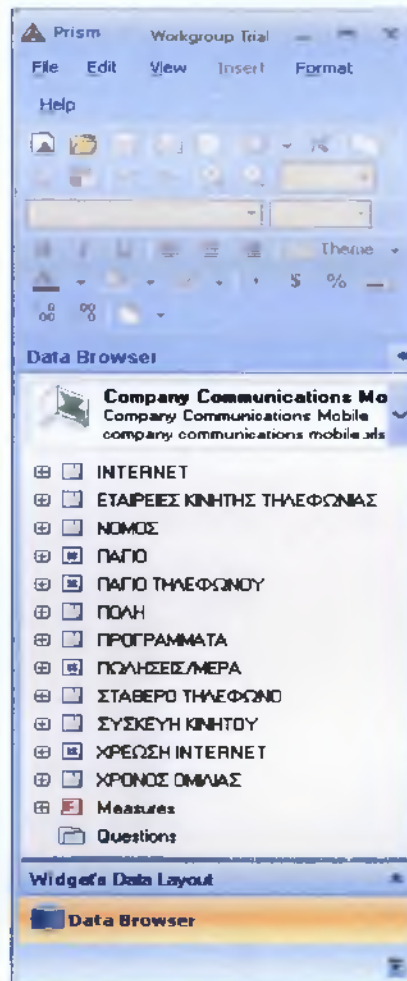
Πατώντας την επιλογή “Edit Dimensions”, εμφανίζεται στην οθόνη ένα παράθυρο, όπου βλέπουμε τις διαστάσεις της βάσης μας. Εκεί, μας δίνεται η δυνατότητα να **τροποποιήσουμε** τις διαστάσεις, ανάλογα με τις απαιτήσεις μας. Στο συγκεκριμένο παράδειγμα, αφήνουμε τις διαστάσεις όπως έχουν (αυτή η επιλογή είναι προαιρετική).



Σχήμα 6-3: Επιλογές διαστάσεων εταιρειών κινητής τηλεφωνίας

Πηγή: Ιδία Έρευνα

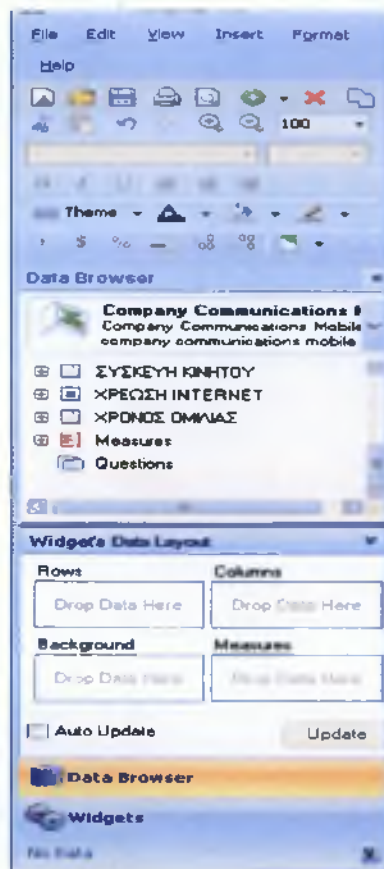
Πατώντας την επιλογή OK, επανερχόμαστε στο προηγούμενο παράθυρο διαλόγου στο σχήμα 6-2 και πατώντας και εκεί OK γίνεται η εισαγωγή των δεδομένων από την πηγή δεδομένων στο Prism, εμφανίζοντας το δέντρο διαστάσεων στα αριστερά του προγράμματος.



Σχήμα 6-4: Δέντρο διαστάσεων εταιρειών κινητής τηλεφωνίας

Πηγή: Ιδία Έρευνα

Ανοίγοντας το πρόγραμμα, το Prism δημιουργεί αυτόματα ένα φύλλο (sheet) και κάνοντας δεξιά κλικ στην περιοχή του καμβά, επιλέγουμε από την αναπτυσσόμενη λίστα την επιλογή “New→ Pivot”. Αυτό έχει σαν αποτέλεσμα, την δημιουργία άξονα (Pivot) χωρίς δεδομένα. Κάνοντας ένα απλό κλικ στον άξονα, αυτόματα εμφανίζεται κάτω από το δέντρο διαστάσεων το Widgets Data Layout, στο οποίο περιέχονται τα πλαίσια των Γραμμών (Rows), των Στηλών (Columns), των Μέτρων (Measures) και του Φόντου (Background), όπως φαίνεται στο παρακάτω σχήμα.



Σχήμα 6-5: Widgets Data Layout

Πηγή: Ίδια Έρευνα

Εφόσον δημιουργηθεί ο άξονας, το εργαλείο είναι έτοιμο να δώσει απαντήσεις στα ερωτήματα που θέλουμε να θέσουμε. Ας υποθέσουμε ότι θέλουμε να εξετάσουμε, πόσες εταιρείες κινητής τηλεφωνίας υπάρχουν, σε ποιες πόλεις, με ατομικά προγράμματα 300 λεπτών και ποιο το πάγιο τους.

Από το δέντρο διαστάσεων, επιλέγουμε την διάσταση “ΕΤΑΙΡΕΙΕΣ ΚΙΝΗΤΗΣ ΤΗΛΕΦΩΝΙΑΣ” και με την απλή διαδικασία Drag and Drop σύρουμε την διάσταση στο Widgets Data Layout και την αφήνουμε στο πλαίσιο των γραμμών. Με τον ίδιο τρόπο, σύρουμε την διάσταση “ΠΟΛΗ” και την τοποθετούμε στο πλαίσιο των γραμμών κάτω από την διάσταση “ΕΤΑΙΡΕΙΕΣ ΚΙΝΗΤΗΣ ΤΗΛΕΦΩΝΙΑΣ”. Όσον αφορά το τρίτο σκέλος του ερωτήματος, αναπτύσσουμε τη διάσταση “ΠΡΟΓΡΑΜΜΑΤΑ” και ακολούθως το μέλος “All ΠΡΟΓΡΑΜΜΑΤΑ” και επιλέγουμε το θυγατρικό μέλος “300(ΑΤΟΜΙΚΟ)”, σύροντάς το στο πλαίσιο των γραμμών κάτω από την διάσταση και το μέλος που τοποθετήσαμε πριν. Τέλος, αναπτύσσουμε την διάσταση μέτρων, επιλέγουμε το μέτρο “Sum ΠΑΓΙΟ” και το σύρουμε – αφήνουμε στο πλαίσιο των στηλών, συμπληρώνοντας το Prism αυτόματα το πλαίσιο των μέτρων. Πατώντας την επιλογή “Update” το αποτέλεσμα θα είναι το εξής:

ΕΤΑΙΡΕΙΕΣ ΚΙΝΗΤΗΣ ΤΗΛΕΦΩΝΙΑΣ	ΠΟΛΗ	Members from ΠΡΟΓΡΑΜΜΑΤΑ	Sum ΠΑΓΙΟ
COSMOTE	ΤΡΙΚΑΛΑ	300(ΑΤΟΜΙΚΟ)	44.50
VODAFONE	ΑΘΗΝΑ	300(ΑΤΟΜΙΚΟ)	44.50
WIND	ΣΕΡΡΕΣ	300(ΑΤΟΜΙΚΟ)	40

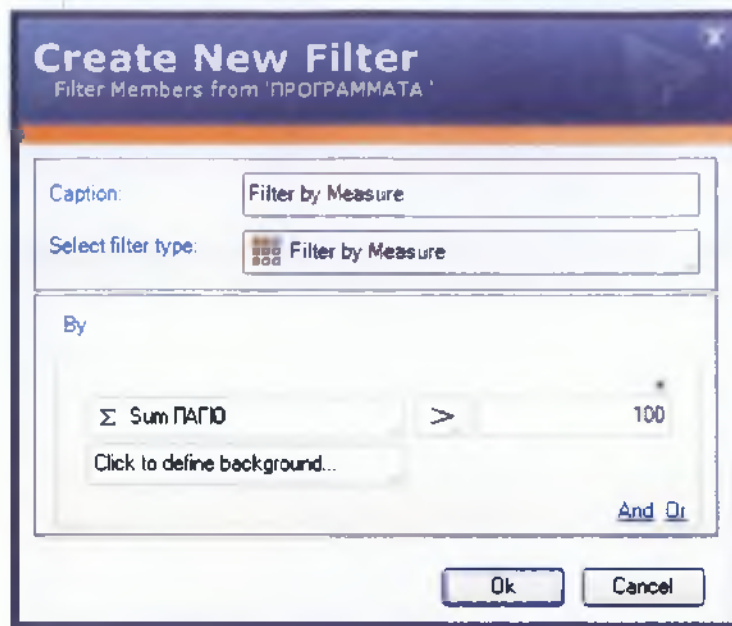
Σχήμα 6-6: Αποτέλεσμα πρώτου παραδείγματος εταιρειών κινητής τηλεφωνίας

Πηγή: Ίδια Έρευνα

Όπως βλέπουμε το αποτέλεσμα είναι αυτό, το οποίο επιθυμούσαμε. Το εργαλείο έδωσε συγκεκριμένη απάντηση, στο ερώτημα που του τέθηκε. Δηλαδή, ότι υπάρχουν τρεις εταιρείες κινητής τηλεφωνίας (COSMOTE, VODAFONE, WIND), στις πόλεις (Τρίκαλα, Αθήνα, Σέρρες) οι οποίες παρέχουν το πακέτο των 300 λεπτών (300 ΑΤΟΜΙΚΟ) με πάγιο (44,50 €, 44,50 €, 40,00 €) αντίστοιχα.

Παράδειγμα 2:

Ας υποθέσουμε ότι θέλουμε να εξετάσουμε, πόσες εταιρείες κινητής τηλεφωνίας έχουν προγράμματα με πάγιο μεγαλύτερο των 100 €. Αυτό το ερώτημα απαιτεί τη δημιουργία ενός φίλτρου, που θα μας δώσει τα επιθυμητά αποτελέσματα. Από το δέντρο διαστάσεων με Drag and Drop, αφήνουμε την διάσταση “ΕΤΑΙΡΕΙΕΣ ΚΙΝΗΤΗΣ ΤΗΛΕΦΩΝΙΑΣ” στο πλαίσιο των γραμμών. Στη συνέχεια, αναπτύσσουμε την διάσταση “ΠΡΟΓΡΑΜΜΑΤΑ”. Κάνουμε δεξί κλικ στην επιλογή “Filter” και από την αναπτυσσόμενη λίστα επιλέγουμε “New”. Το παράθυρο διαλόγου που εμφανίζεται, μας δίνει τη δυνατότητα να επιλέξουμε το είδος του φίλτρου που θέλουμε. Στο συγκεκριμένο παράδειγμα, επιλέξαμε το Φίλτρο με Μέτρο (Filter by Measure). Στο πλαίσιο “By”, επιλέγουμε το “Σ Sum ΠΑΓΙΟ”, σαν τελεστή το “>” και στο πλαίσιο δίπλα από τον τελεστή γράφουμε τον αριθμό 100 (σχήμα 6-7). Πατώντας το κουμπί OK, βλέπουμε στην διάσταση “ΠΡΟΓΡΑΜΜΑΤΑ”, στο δέντρο διαστάσεων, ότι έχει εμφανιστεί το νέο φίλτρο που δημιουργήσαμε “Filter By Measure” (σχήμα 6-8).



Σχήμα 6-7: Δημιουργία φίλτρου με μέτρο
 Πηγή: Ιδία Έρευνα



Σχήμα 6-8: Εμφάνιση φίλτρου με μέτρο
 Πηγή: Ιδία Έρευνα

Με Drag and Drop τοποθετούμε το νέο φίλτρο στο πλαίσιο των γραμμών, κάτω από την διάσταση “ΕΤΑΙΡΕΙΕΣ ΚΙΝΗΤΗΣ ΤΗΛΕΦΩΝΙΑΣ” και από την διάσταση μέτρα επιλέγουμε το “Σ Sum ΠΑΓΙΟ” και με Drag and Drop το τοποθετούμε στο πλαίσιο των στηλών, συμπληρώνοντας αυτόματα το Prism το πλαίσιο των μέτρων. Πατώντας το κουμπί “Update”, το αποτέλεσμα είναι το εξής:

ΕΤΑΙΡΕΙΕΣ ΚΙΝΗΤΗΣ ΤΗΛΕΦΩΝΙΑΣ	Filter by Measure	Sum ΠΑΓΙΟ
COSMOTE	1000(ΑΤΟΜΙΚΟ)	125
COSMOTE	1000(ΕΠΑΓΓΕΛΜΑΤΙΚΟ)	130
VODAFONE	1000(ΑΤΟΜΙΚΟ)	113
VODAFONE	1000(ΕΠΑΓΓΕΛΜΑΤΙΚΟ)	117
WIND	1000(ΑΤΟΜΙΚΟ)	125
WIND	1000(ΕΠΑΓΓΕΛΜΑΤΙΚΟ)	114

Σχήμα 6-9: Αποτέλεσμα δεύτερου παραδείγματος εταιρειών κινητής τηλεφωνίας με χρήση φίλτρου

Πηγή: Ίδια Έρευνα

Το αποτέλεσμα δείχνει, ότι οι εταιρείες (COSMOTE, VODAFONE, WIND) προσφέρουν ατομικά και επαγγελματικά πακέτα των 1000 λεπτών χρόνου ομιλίας, με πάγιο μεγαλύτερο των 100€.

Παράδειγμα 3:

Ας υποθέσουμε ότι θέλουμε να εξετάσουμε, ποιες εταιρείες κινητής τηλεφωνίας προσφέρουν προγράμματα παροχής internet, ποιο το πάγιο των προγραμμάτων και ποια η χρέωση του internet. Από το δέντρο διαστάσεων, επιλέγουμε την διάσταση “ΕΤΑΙΡΕΙΕΣ ΚΙΝΗΤΗΣ ΤΗΛΕΦΩΝΙΑΣ” και με την διαδικασία Drag and Drop την σύρουμε και την αφήνουμε στο πλαίσιο των γραμμών. Με την ίδια διαδικασία, σύρουμε και αφήνουμε την διάσταση “ΠΡΟΓΡΑΜΜΑΤΑ” στο πλαίσιο των γραμμών κάτω από την προηγούμενη διάσταση. Από το δέντρο διαστάσεων, αναπτύσσουμε την διάσταση “INTERNET” και ακολούθως το μέλος “All INTERNET”, από το οποίο επιλέγουμε το θυγατρικό μέλος “YES”

και με Drag and Drop το αφήνουμε στο πλαίσιο των γραμμών, κάτω από τις προηγούμενες διαστάσεις. Στη συνέχεια, από το δέντρο διαστάσεων επιλέγουμε την διάσταση “ΠΑΓΙΟ” και την σύρουμε – αφήνουμε στο πλαίσιο των γραμμών. Τέλος, από την διάσταση μέτρων επιλέγουμε το “Σ Sum ΧΡΕΩΣΗ INTERNET” και το σύρουμε – αφήνουμε στο πλαίσιο των στηλών. Πατώντας το κουμπί “Update”, το αποτέλεσμα θα είναι το εξής:

ΕΤΑΙΡΕΙΕΣ ΚΙΝΗΤΗΣ ΤΗΛΕΦΩΝΙΑΣ	ΠΡΟΓΡΑΜΜΑΤΑ	Members from INTERNET	ΠΑΓΙΟ	Sum ΧΡΕΩΣΗ INTERNET
COSMOTE	120(ΑΤΟΜΙΚΟ)	YES	22	0
	120(ΕΠΑΓΓΕΛΜΑΤΙΚΟ)	YES	36	0
	180(ΦΟΙΤΗΤΙΚΟ)	YES	25	0
	300(ΑΤΟΜΙΚΟ)	YES	44,5	29,50
	600(ΑΤΟΜΙΚΟ)	YES	82	25
	(ΟΙΚΟΓΕΝΕΙΑΚΟ-4 ΑΤΟΜΩΝ)	YES	29	19,90
VODAFONE	120(ΑΤΟΜΙΚΟ)	YES	20	0
	360(ΕΠΑΓΓΕΛΜΑΤΙΚΟ)	YES	73,5	22,80
	600(ΑΤΟΜΙΚΟ)	YES	69,5	16,50
	800(ΦΟΙΤΗΤΙΚΟ)	YES	65	10
	(ΟΙΚΟΓΕΝΕΙΑΚΟ-3 ΑΤΟΜΩΝ)	YES	23,99	15,50
	180(ΦΟΙΤΗΤΙΚΟ)	YES	30	0
	360(ΕΠΑΓΓΕΛΜΑΤΙΚΟ)	YES	75	19,50
	600(ΦΟΙΤΗΤΙΚΟ)	YES	60	0

Σχήμα 6-10: Αποτέλεσμα τρίτου παραδείγματος εταιρειών κινητής τηλεφωνίας

Πηγή: Ίδια Έρευνα

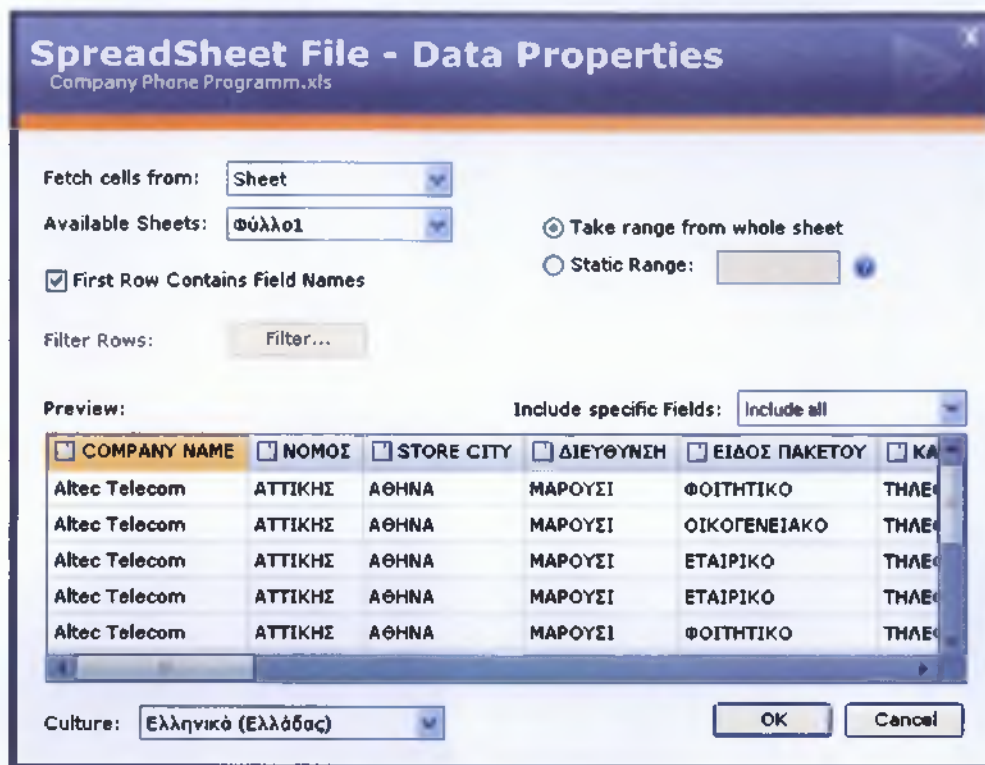
Το αποτέλεσμα δείχνει ότι οι εταιρείες (COSMOTE, VODAFONE, WIND) παρέχουν internet στα παραπάνω προγράμματα. Επίσης, φαίνεται το πάγιο των πακέτων και η χρέωση του internet για το κάθε πρόγραμμα. Οι μηδενικές τιμές δείχνουν ότι σε αυτή την περίπτωση, ο πελάτης πληρώνει μόνο το πάγιο του πακέτου και δεν έχει επιπλέον χρέωση για την παροχή του internet.

6.2 Επεξεργασία δεδομένων εταιρειών σταθερής τηλεφωνίας

Παράδειγμα 1:

Σε αυτό το παράδειγμα, έχουμε δεδομένα εταιρειών σταθερής τηλεφωνίας. Ακολουθούμε τα ίδια βήματα που εφαρμόσαμε στο προηγούμενο παράδειγμα. Αρχικά, εισάγουμε την πηγή δεδομένων στο εργαλείο Prism και εμφανίζεται στην οθόνη το παράθυρο

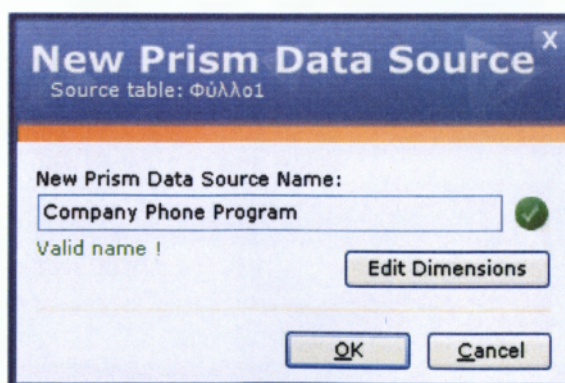
διαλόγου “Spreadsheet File – Data Properties”, από το οποίο έχουμε μια πρώτη όψη των δεδομένων που θέλουμε να εισάγουμε.



Σχήμα 6-11: Εισαγωγή δεδομένων εταιρειών σταθερής τηλεφωνίας

Πηγή: Ιδία Έρευνα

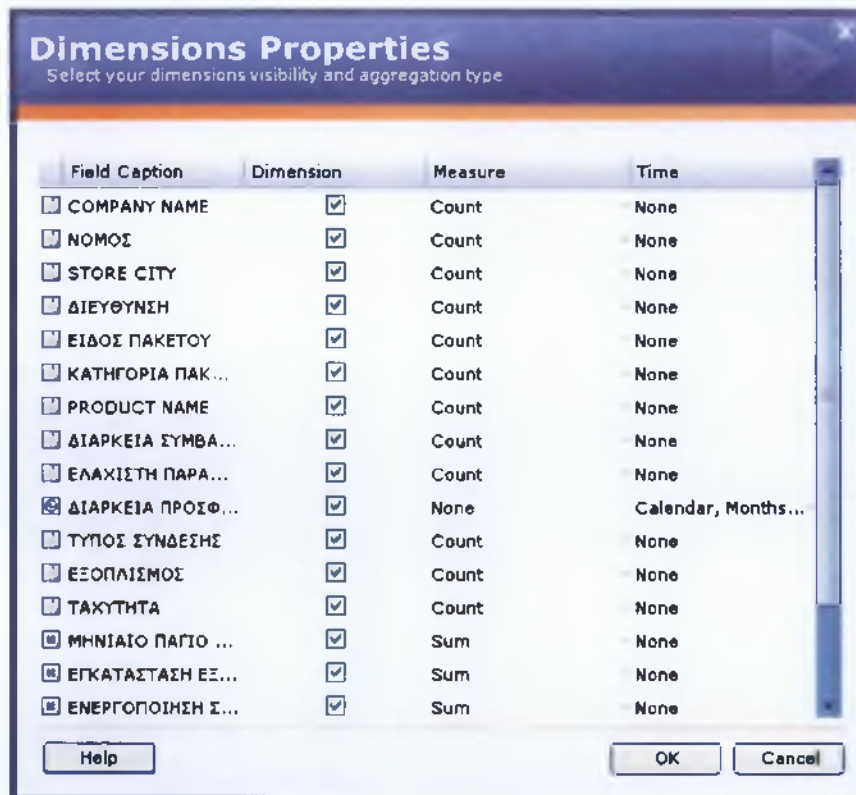
Πατώντας το κουμπί OK, εμφανίζεται το παράθυρο διαλόγου “New Prism Data Source”, από το οποίο μπορούμε να ονομάσουμε την πηγή δεδομένων, όπως επιθυμούμε. Στο συγκεκριμένο παράδειγμα ονομάζουμε την πηγή δεδομένων “Company Phone Program”, όπως φαίνεται στο σχήμα 6-12 παρακάτω.



Σχήμα 6-12: Ονομασία πηγής δεδομένων εταιρειών σταθερής τηλεφωνίας

Πηγή: Ιδία Έρευνα

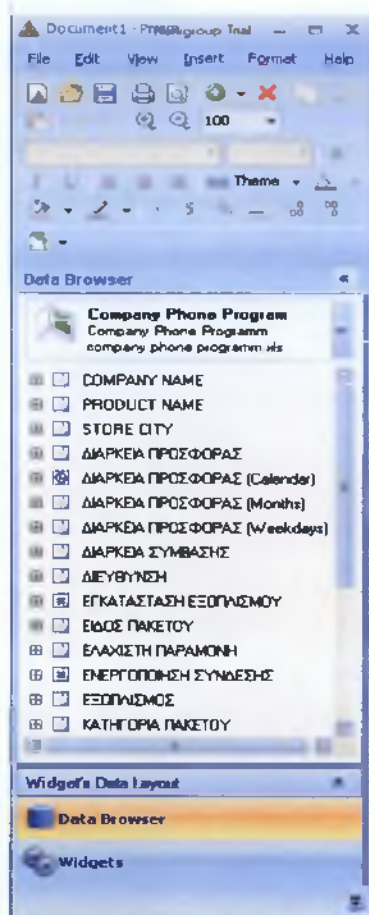
Πατώντας την επιλογή “Edit Dimensions”, όπως εφαρμόσαμε και στο προηγούμενο παράδειγμα, εμφανίζεται στην οθόνη ένα παράθυρο διαλόγου, στο οποίο βλέπουμε τις διαστάσεις της βάσης μας. Εκεί, μας δίνεται η δυνατότητα να τροποποιήσουμε τις διαστάσεις, ανάλογα με τις απαιτήσεις μας. Στο συγκεκριμένο παράδειγμα, αφήνουμε τις διαστάσεις όπως έχουν, όπως και στο προηγούμενο παράδειγμα.



Σχήμα 6-13: Επιλογές διαστάσεων εταιρειών σταθερής τηλεφωνίας

Πηγή: Ιδία Έρευνα

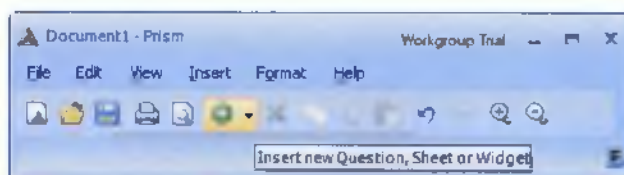
Πατώντας την επιλογή OK, επανερχόμαστε στο προηγούμενο παράθυρο διαλόγου του σχήματος 6-12 και πατώντας και εκεί OK γίνεται η εισαγωγή των δεδομένων από την πηγή δεδομένων στο Prism, εμφανίζοντας το δέντρο διαστάσεων στα αριστερά του προγράμματος.



Σχήμα 6-14: Δέντρο διαστάσεων εταιρειών σταθερής τηλεφωνίας

Πηγή: Ιδία Έρευνα

Ένας άλλος τρόπος για την εισαγωγή άξονα στον καμβά είναι μέσω της γραμμής εργαλείων (toolbar), πατώντας το πράσινο κουμπάκι (σχήμα 6-15) και επιλέγοντας από την αναπτυσσόμενη λίστα την επιλογή “Ρίνοι”.

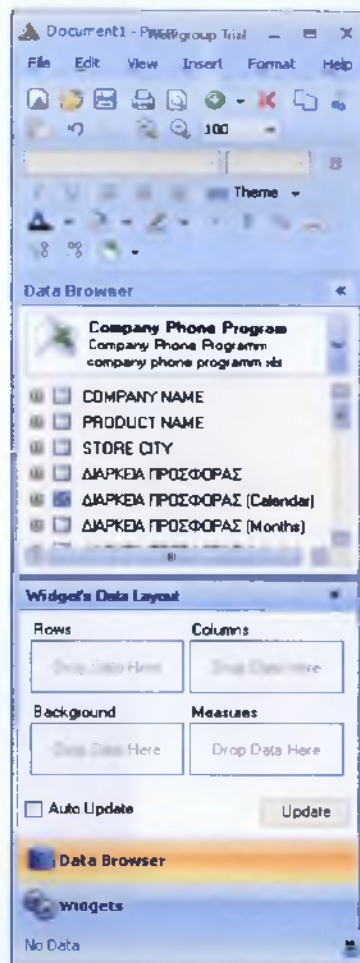


Σχήμα 6-15: Γραμμή εργαλείων

Πηγή: Ιδία Έρευνα

Έχοντας δημιουργήσει τον άξονα και κάνοντας ένα απλό κλικ πάνω του, είμαστε έτοιμοι να θέσουμε τα ερωτήματά μας, τοποθετώντας τις διαστάσεις και τα μέτρα που είναι απαραίτητα, στα πλαίσια των γραμμών (Rows), των στηλών (Columns), των μέτρων

(Measures) και του φόντου (Background), που βρίσκονται μέσα στο Widgets Data Layout κάτω από το δέντρο διαστάσεων.



Σχήμα 6-16: Widgets Data Layout

Πηγή: Ιδία Έρευνα

Ας υποθέσουμε ότι θέλουμε να δούμε, πόσες εταιρείες προσφέρουν οικογενειακό πακέτο με τηλέφωνο και internet, ποιο είναι το όνομα των πακέτων και ποιο είναι το κόστος ανά μήνα που θα πρέπει να πληρώνει ο κάθε πελάτης. Από το δέντρο διαστάσεων, επιλέγουμε την διάσταση “COMPANY NAME” και με Drag and Drop, την σύρουμε – αφήνουμε στο πλαίσιο των γραμμών. Στη συνέχεια, αναπτύσσουμε την διάσταση “ΕΙΔΟΣ ΠΑΚΕΤΟΥ” και ακολούθως το μέλος “All ΕΙΔΟΣ ΠΑΚΕΤΟΥ”. Επιλέγουμε το θυγατρικό μέλος “ΟΙΚΟΓΕΝΕΙΑΚΟ” και το σύρουμε – αφήνουμε στο πλαίσιο των γραμμών, κάτω από την διάσταση που τοποθετήσαμε πριν. Με τον ίδιο ακριβώς τρόπο, αναπτύσσουμε την διάσταση “ΚΑΤΗΓΟΡΙΑ ΠΑΚΕΤΟΥ” και ακολούθως το μέλος “All ΚΑΤΗΓΟΡΙΑ ΠΑΚΕΤΟΥ” και επιλέγουμε, σύρουμε και αφήνουμε το θυγατρικό μέλος “ΤΗΛΕΦΩΝΟ ΚΑΙ INTERNET” στο πλαίσιο των γραμμών κάτω από την διάσταση και τα δύο μέλη, που τοποθετήσαμε πριν.

Επιλέγουμε την διάσταση “PRODUCT NAME” και την τοποθετούμε και αυτή στο πλαίσιο των γραμμών. Τέλος, αναπτύσσουμε την διάσταση των μέτρων και επιλέγουμε, σύρουμε – αφήνουμε το “Sum ΚΟΣΤΟΣ ΑΝΑ ΜΗΝΑ” στο πλαίσιο των στηλών και πατάμε το κουμπι “Update”. Το αποτέλεσμα είναι το εξής:

COMPANY NAME	Members from ΕΙΔΟΣ ΠΑΚΕΤΟΥ	Members from ΚΑΤΗΓΟΡΙΑ ΠΑΚΕΤΟΥ	PRODUCT NAME	Sum ΚΟΣΤΟΣ ΑΝΑ ΜΗΝΑ
Algonet Telecom	ΟΙΚΟΓΕΝΕΙΑΚΟ	ΤΗΛΕΦΩΝΟ ΚΑΙ INTERNET	Max Value	15.35
Altet Telecom	ΟΙΚΟΓΕΝΕΙΑΚΟ	ΤΗΛΕΦΩΝΟ ΚΑΙ INTERNET	Altetnet ADSL 24 Mbit	19.99
Forthnet	ΟΙΚΟΓΕΝΕΙΑΚΟ	ΤΗΛΕΦΩΝΟ ΚΑΙ INTERNET	Forthnet 2play	39.90
Hellas OnLine	ΟΙΚΟΓΕΝΕΙΑΚΟ	ΤΗΛΕΦΩΝΟ ΚΑΙ INTERNET	hol double-play	39.90

Σχήμα 6-17: Αποτέλεσμα πρώτου παραδείγματος εταιρειών σταθερής τηλεφωνίας

Πηγή: Ιδία Έρευνα

Όπως βλέπουμε, το Prism μας έδωσε εκείνες τις εταιρείες που παρέχουν οικογενειακό πακέτο με τηλέφωνο και internet, αποκλείοντας παράλληλα εκείνες τις εταιρείες που δεν παρέχουν οικογενειακό πακέτο. Επίσης, μας έδωσε την ονομασία των πακέτων και την τιμή που καλείται να πληρώσει ο κάθε πελάτης κάθε μήνα.

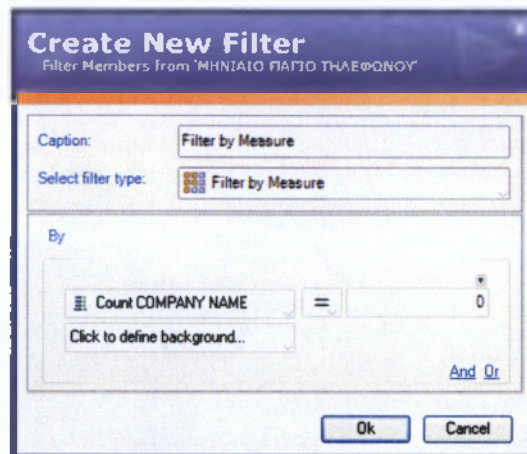
Παράδειγμα 2:

Ας υποθέσουμε ότι θέλουμε να εξετάσουμε, ποιες εταιρείες σταθερής τηλεφωνίας προσφέρουν εταιρικό πακέτο με μηνιαίο πάγιο άνω των 20 € με σταθερό τηλέφωνο, ποιο είναι το όνομα του πακέτου και ποιο το κόστος ενεργοποίησης της σύνδεσης.

Από το δέντρο διαστάσεων, επιλέγουμε την διάσταση “COMPANY NAME” και με Drag and Drop σύρουμε – αφήνουμε την διάσταση στο πλαίσιο των γραμμών. Αναπτύσσουμε την διάσταση “ΕΙΔΟΣ ΠΑΚΕΤΟΥ” και ακολούθως το μέλος “All ΕΙΔΟΣ ΠΑΚΕΤΟΥ” και επιλέγουμε το θυγατρικό μέλος “ΕΤΑΙΡΙΚΟ” και το τοποθετούμε στο πλαίσιο των γραμμών κάτω από την διάσταση “COMPANY NAME”. Με το ίδιο τρόπο, αναπτύσσουμε την διάσταση “ΚΑΤΗΓΟΡΙΑ ΠΑΚΕΤΟΥ” και από το μέλος “All ΚΑΤΗΓΟΡΙΑ ΠΑΚΕΤΟΥ”

επιλέγουμε το θυγατρικό μέλος “ΤΗΛΕΦΩΝΟ” και το τοποθετούμε στο πλαίσιο των γραμμών.

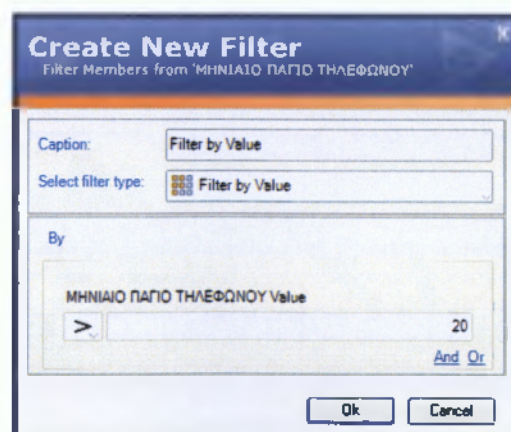
Όσον αφορά το μηνιαίο πακέτο των 20 €, από το δέντρο διαστάσεων αναπτύσσουμε την διάσταση “ΜΗΝΙΑΙΟ ΠΑΓΙΟ ΤΗΛΕΦΩΝΟΥ”. Κάνουμε δεξί κλικ πάνω στο “filter” και από την αναπτυσσόμενη λίστα πατάμε την επιλογή “New”. Στο παράθυρο διαλόγου που άνοιξε (σχήμα 6-18), μας δίνεται η δυνατότητα να επιλέξουμε το είδος του φίλτρου που θα χρησιμοποιήσουμε, πατώντας απλώς στην επιλογή “Select filter type”.



Σχήμα 6-18: Επιλογή φίλτρου

Πηγή: Ιδία Έρευνα

Από την αναπτυσσόμενη λίστα, επιλέγουμε το φίλτρο με αξία (Filter By Value) και το παράθυρο διαλόγου αλλάζει μορφή. Στο πλαίσιο “By”, επιλέγουμε τον τελεστή “>” και στο διπλανό πλαίσιο γράφουμε τον αριθμό 20 (σχήμα 6-19), αφού θέλουμε το μηνιαίο πάγιο να είναι μεγαλύτερο των 20 €.



Σχήμα 6-19: Φίλτρο με αξία

Πηγή: Ιδία Έρευνα

Πατώντας την επιλογή OK, το φίλτρο εμφανίζεται στην διάσταση “ΜΗΝΙΑΙΟ ΠΑΓΙΟ ΤΗΛΕΦΩΝΟΥ” και με Drag and Drop το τοποθετούμε στο πλαίσιο των γραμμών κάτω από τα προηγούμενα, όπως φαίνεται στο επόμενο σχήμα.



Σχήμα 6-20: Εμφάνιση φίλτρου με αξία

Πηγή: Ιδία Έρευνα

Στη συνέχεια, επιλέγουμε από το δέντρο διαστάσεων την διάσταση “PRODUCT NAME” και με Drag and Drop την τοποθετούμε στο πλαίσιο των γραμμών. Τέλος, αναπτύσσουμε την διάσταση των μέτρων και επιλέγουμε το “Σ Sum ΕΝΕΡΓΟΠΟΙΗΣΗ ΣΥΝΔΕΣΗΣ” και το σύρουμε – αφήνουμε στο πλαίσιο των στηλών. Πατώντας το κουμπί “Update”, το αποτέλεσμα θα είναι το εξής:

COMPANY NAME	Members from ΕΙΔΟΣ ΠΑΚΕΤΟΥ	Members from ΚΑΤΗΓΟΡΙΑ ΠΑΚΕΤΟΥ	Filter by Value	PRODUCT NAME	Sum ΕΝΕΡΓΟΠΟΙΗΣΗ ΣΥΝΔΕΣΗΣ
Hellas Online	ΕΤΑΙΡΙΚΟ	ΤΗΛΕΦΩΝΟ	21,43	hol business telephony CPS	34,52
			25,5	hol business telephony PRI	34,52
OTE	ΕΤΑΙΡΙΚΟ	ΤΗΛΕΦΩΝΟ	20,94	ISDN BRA	0
Telfas	ΕΤΑΙΡΙΚΟ	ΤΗΛΕΦΩΝΟ	39,9	Telfas Corporate Telephony	0
Vivodi	ΕΤΑΙΡΙΚΟ	ΤΗΛΕΦΩΝΟ	24,32	Business Voice	0

Σχήμα 6-21: Αποτέλεσμα δεύτερου παραδείγματος εταιρειών σταθερής τηλεφωνίας
 Πηγή: Ιδία έρευνα

Το αποτέλεσμα δείχνει τις εταιρείες που προσφέρουν εταιρικό πακέτο τηλεφώνου, με μηνιαίο πάγιο άνω των 20 €, αποκλείοντας ταυτόχρονα τις εταιρείες που δεν προσφέρουν εταιρικό πακέτο ή προσφέρουν κάτω των 20 €. Επίσης, στην εικόνα φαίνεται το όνομα του κάθε πακέτου των εταιρειών καθώς και το κόστος ενεργοποίησης της σύνδεσης.

Παράδειγμα 3:

Ας υποθέσουμε ότι θέλουμε να εξετάσουμε, ποιες εταιρείες παρέχουν πακέτου τηλεφώνου και internet, ποιο είναι το όνομα και η ταχύτητα του κάθε πακέτου, πόσο θα πληρώνει ο κάθε πελάτης κάθε μήνα, πόσο θα στοιχίσει η εγκατάσταση του εξοπλισμού και ποιο το κόστος απεγκατάστασής του.

Από το δέντρο διαστάσεων, επιλέγουμε την διάσταση “COMPANY NAME” και με Drag and Drop την τοποθετούμε στο πλαίσιο των γραμμών. Στη συνέχεια, αναπτύσσουμε την διάσταση “ΚΑΤΗΓΟΡΙΑ ΠΑΚΕΤΟΥ” και ακολούθως το μέλος “All ΚΑΤΗΓΟΡΙΑ ΠΑΚΕΤΟΥ”, από το οποίο επιλέγουμε το θυγατρικό μέλος “ΤΗΛΕΦΩΝΟ ΚΑΙ INTERNET”. Με Drag and Drop το σύρουμε – αφήνουμε στο πλαίσιο των γραμμών. Επίσης, επιλέγουμε τις διαστάσεις “PRODUCT NAME”, “ΤΑΧΥΤΗΤΑ” και τις σύρουμε στο πλαίσιο των γραμμών. Τέλος, αναπτύσσουμε την διάσταση των μέτρων και επιλέγουμε τα μέτρα “Sum ΚΟΣΤΟΣ ΑΝΑ ΜΗΝΑ”, “Sum ΕΓΚΑΤΑΣΤΑΣΗ ΕΞΟΠΛΙΣΜΟΥ” και “Sum ΚΟΣΤΟΣ ΑΠΕΓΚΑΤΑΣΤΑΣΗΣ” αφήνοντάς τα στο πλαίσιο των στηλών. Πατώντας το κουμπί “Update”, το αποτέλεσμα θα είναι το εξής:

The screenshot shows a Prism application window with a PivotTable. The PivotTable is filtered by 'Members from ΚΑΤΗΓΟΡΙΑ ΠΑΚΕΤΟΥ' and displays data for various companies and their products. The columns are: COMPANY NAME, MEMBERS FROM ΚΑΤΗΓΟΡΙΑ ΠΑΚΕΤΟΥ, PRODUCT NAME, ΤΑΧΥΤΗΤΑ, Sum ΚΟΣΤΟΣ ΑΝΑ ΜΗΝΑ, Sum ΕΓΚΑΤΑΣΤΑΣΗ ΕΞΟΦΛΙΣΜΟΥ, and Sum ΚΟΣΤΟΣ ΑΠΕΓΚΑΤΑΣΤΑΣΗΣ.

COMPANY NAME	MEMBERS FROM ΚΑΤΗΓΟΡΙΑ ΠΑΚΕΤΟΥ	PRODUCT NAME	ΤΑΧΥΤΗΤΑ	Sum ΚΟΣΤΟΣ ΑΝΑ ΜΗΝΑ	Sum ΕΓΚΑΤΑΣΤΑΣΗ ΕΞΟΦΛΙΣΜΟΥ	Sum ΚΟΣΤΟΣ ΑΠΕΓΚΑΤΑΣΤΑΣΗΣ
Algonet Telecom	ΤΗΛΕΦΩΝΟ ΚΑΙ INTERNET	Max 24	24Kbps	14.50	0	10.38
		Max Value	24Kbps	15.35	25.40	10.38
		Plus Web Office	2048Kbps	44.40	0	20.76
Altec Telecom	ΤΗΛΕΦΩΝΟ ΚΑΙ INTERNET	Altecnet ADSL 2048	24Mbps	43.57	107.08	0
		Altecnet ADSL 24 Mbit	24Mbps	57.43	107.08	0
		Altecnet Student ADSL 1024	24Mbps	9.50	0	0
		Forthnet 2play	24Mbps	36.90	0	0
Forthnet	ΤΗΛΕΦΩΝΟ ΚΑΙ INTERNET	Forthnet ADSL College	24Mbps	24.90	0	10.12
		Forthnet ADSL Economy	24Mbps	24.90	0	0
		Forthnet ADSL Student	8Mbps	14.90	0	0
		hol business double-play	12Mbps	66.90	0	0
Hellas OnLine	ΤΗΛΕΦΩΝΟ ΚΑΙ INTERNET	hol double-play	24Mbps	79.90	0	0
		hol double-play lite	24Mbps	39.90	0	0
		hol double-play lite	6Mbps	21.90	0	0
Lannet	ΤΗΛΕΦΩΝΟ ΚΑΙ INTERNET	Livebox	8Mbps	25.15	50	55

Σχήμα 6-22: Αποτέλεσμα τρίτου παραδείγματος εταιρειών σταθερής τηλεφωνίας

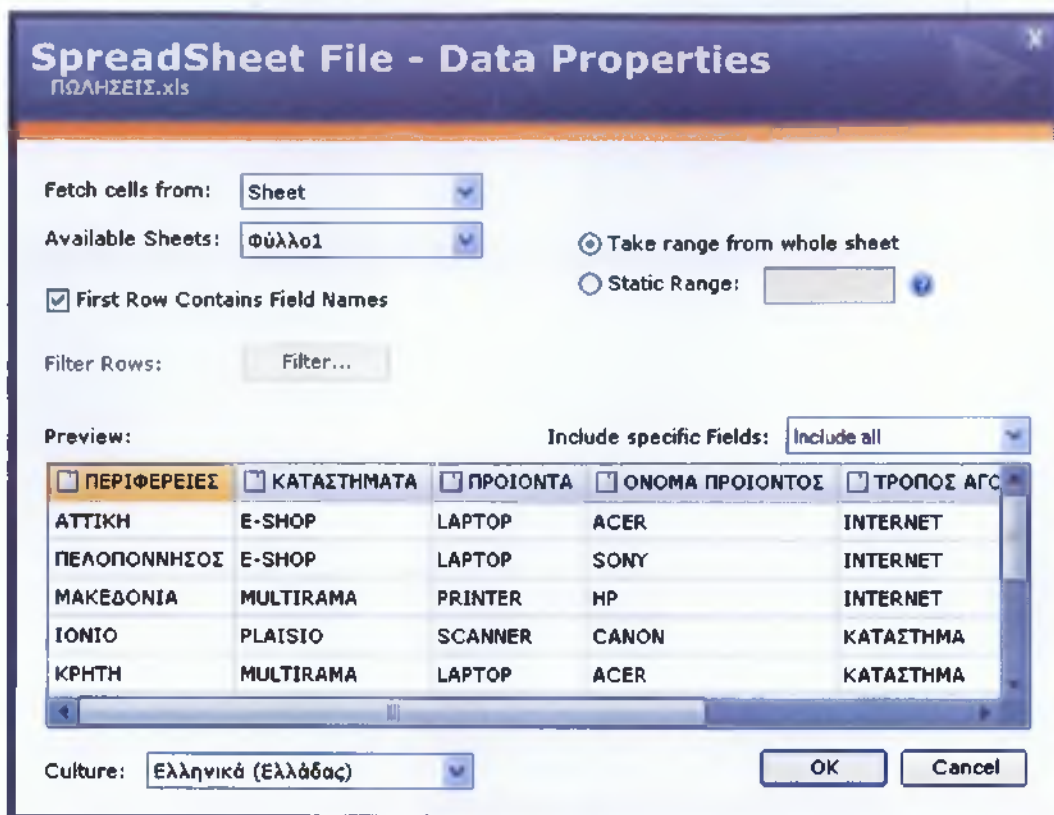
Πηγή: Ίδια έρευνα

Όπως βλέπουμε, το αποτέλεσμα είναι για ακόμη μια φορά το αναμενόμενο. Το Prism μας έδωσε τις εταιρείες που παρέχουν τηλέφωνο και internet, αποκλείοντας όσες δεν προσφέρουν. Ακόμα μας έδωσε το όνομα του κάθε πακέτου, σε τι ταχύτητες κινούνται, πόσο θα πληρώνει ο πελάτης κάθε μήνα, καθώς και πόσο θα κοστίσει στον πελάτη η εγκατάσταση και η απεγκατάσταση του πακέτου.

6.3 Επεξεργασία δεδομένων καταστημάτων πληροφορικής

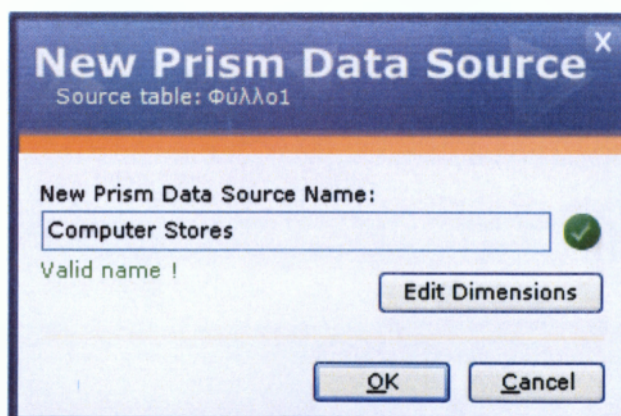
Παράδειγμα 1:

Σε αυτό το παράδειγμα, έχουμε δεδομένα καταστημάτων πληροφορικής. Ακολουθούμε τα ίδια βήματα που εφαρμόσαμε στα προηγούμενα παραδείγματα. Αρχικά, εισάγουμε την πηγή δεδομένων στο εργαλείο Prism και εμφανίζεται στην οθόνη το παράθυρο διαλόγου “Spreadsheet File – Data Properties”, από το οποίο έχουμε μια πρώτη όψη των δεδομένων που θέλουμε να εισάγουμε.



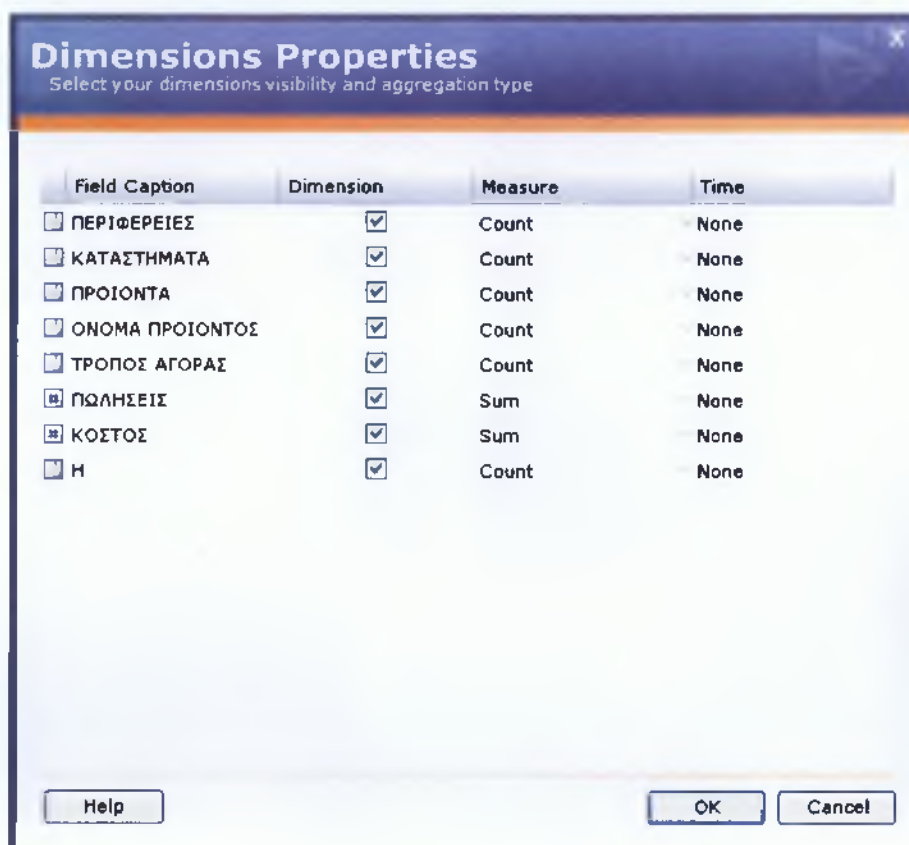
Σχήμα 6-23: Εισαγωγή δεδομένων καταστημάτων πληροφορικής
Πηγή: Ιδία Έρευνα

Πατώντας το κουμπί OK, εμφανίζεται το παράθυρο διαλόγου “New Prism Data Source”, από το οποίο μπορούμε να ονομάσουμε την πηγή δεδομένων, όπως επιθυμούμε. Στο συγκεκριμένο παράδειγμα ονομάζουμε την πηγή δεδομένων “Computer Stores”, όπως φαίνεται στο σχήμα 6-24 παρακάτω.



Σχήμα 6-24: Ονομασία πηγής δεδομένων καταστημάτων πληροφορικής
Πηγή: Ιδία έρευνα

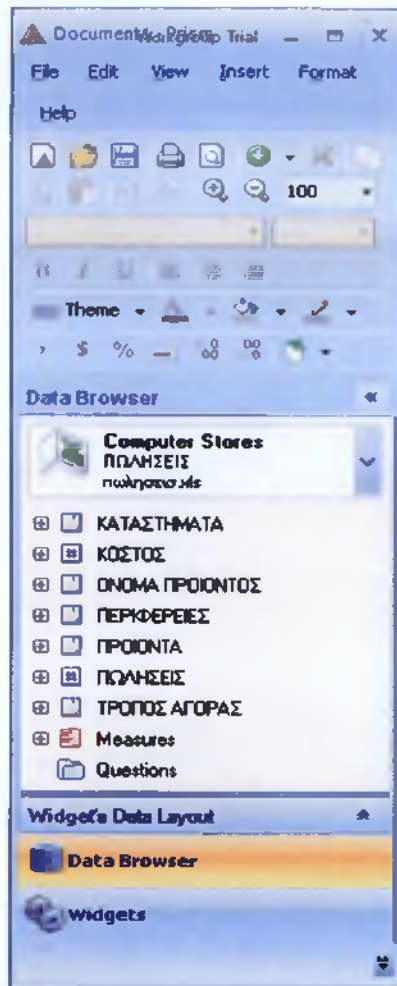
Πατώντας την επιλογή “Edit Dimensions”, όπως εφαρμόσαμε και στα προηγούμενα παραδείγματα, εμφανίζεται στην οθόνη ένα παράθυρο διαλόγου, στο οποίο βλέπουμε τις διαστάσεις της βάσης μας. Εκεί, μας δίνεται η δυνατότητα να τροποποιήσουμε τις διαστάσεις, ανάλογα με τις απαιτήσεις μας. Στο συγκεκριμένο παράδειγμα, αφήνουμε τις διαστάσεις όπως έχουν, όπως και στα προηγούμενα παραδείγματα.



Σχήμα 6-25: Επιλογές διαστάσεων καταστημάτων πληροφορικής

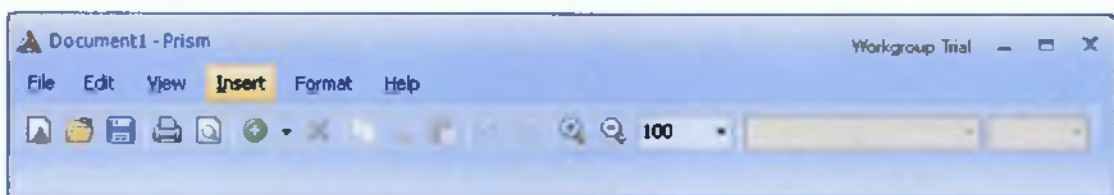
Πηγή: Ιδία έρευνα

Πατώντας την επιλογή OK, επανερχόμαστε στο προηγούμενο παράθυρο διαλόγου του σχήματος 6-24 και πατώντας και εκεί OK γίνεται η εισαγωγή των δεδομένων από την πηγή δεδομένων στο Prism, εμφανίζοντας το δέντρο διαστάσεων στα αριστερά του προγράμματος.



Σχήμα 6-26: Δέντρο διαστάσεων καταστημάτων πληροφορικής
 Πηγή: Ιδία έρευνα

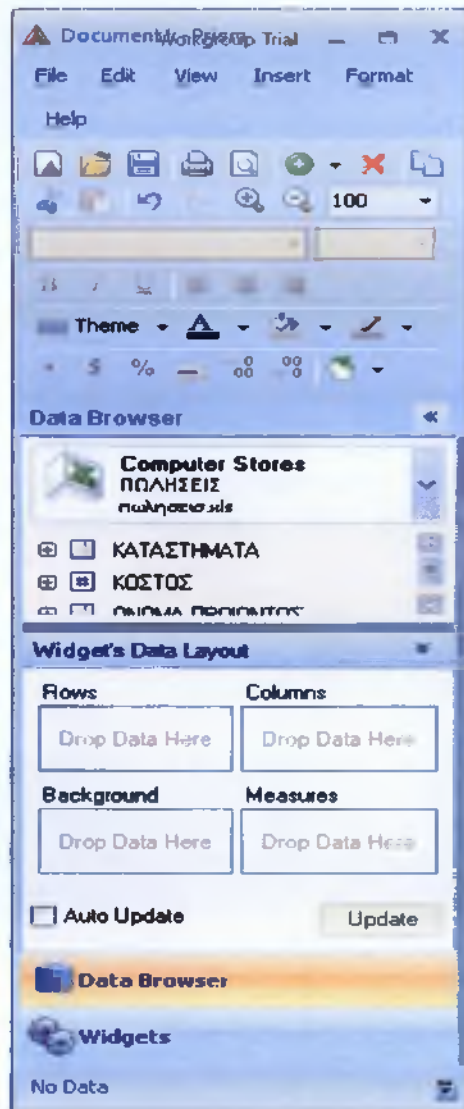
Ένας τρίτος τρόπος για την εισαγωγή άξονα στον καμβά είναι μέσω της επιλογής “Insert” πάνω από την γραμμή εργαλείων (toolbar), κάνοντας κλικ εκεί και επιλέγοντας από την αναπτυσσόμενη λίστα την επιλογή “Pinot”, όπως φαίνεται παρακάτω.



Σχήμα 6-27: Γραμμή εργαλείων
 Πηγή: Ιδία έρευνα

Έχοντας δημιουργήσει τον άξονα και κάνοντας ένα απλό κλικ πάνω του, είμαστε έτοιμοι να θέσουμε τα ερωτήματά μας, τοποθετώντας τις διαστάσεις και τα μέτρα που είναι

απαραίτητα, στα πλαίσια των γραμμών (Rows), των στηλών (Columns), των μέτρων (Measures) και του φόντου (Background), που βρίσκονται μέσα στο Widgets Data Layout κάτω από το δέντρο διαστάσεων.



Σχήμα 6-28: Widgets Data Layout

Πηγή: Ιδία έρευνα

Ας υποθέσουμε ότι θέσουμε να εξετάσουμε, ποια καταστήματα, σε ποιες περιφέρειες πωλούν υπολογιστές τύπου Laptop και ποιες είναι οι πωλήσεις αυτών.

Από το δέντρο διαστάσεων, επιλέγουμε τις διαστάσεις “ΚΑΤΑΣΤΗΜΑΤΑ”, “ΠΕΡΙΦΕΡΕΙΕΣ” και με Drag and Drop τις τοποθετούμε στο πλαίσιο των γραμμών. Στη συνέχεια, αναπτύσσοντας την διάσταση “ΠΡΟΪΟΝΤΑ” και ακολούθως το μέλος “All ΠΡΟΪΟΝΤΑ” επιλέγουμε το θυγατρικό μέλος “Laptop” και το σύρουμε – αφήνουμε στο πλαίσιο των γραμμών. Τέλος, αναπτύσσουμε την διάσταση των μέτρων και επιλέγουμε το

μέτρο “Sum ΠΩΛΗΣΕΙΣ” και το σύρουμε – αφήνουμε στο πλαίσιο των στηλών. Πατώντας το κουμπί “Update”, το αποτέλεσμα θα είναι το εξής:

ΚΑΤΑΣΤΗΜΑΤΑ	ΠΕΡΙΦΕΡΕΙΕΣ	Μέλη από ΠΡΟΪΟΝΤΑ	Sum ΠΩΛΗΣΕΙΣ
Ε-SHOP	ΑΤΤΙΚΗ	ΛΑΡΤΟΡ	30.824
	ΗΠΕΙΡΟΣ	ΛΑΡΤΟΡ	7.890
	ΚΡΗΤΗ	ΛΑΡΤΟΡ	24.505
	ΜΑΚΕΔΟΝΙΑ	ΛΑΡΤΟΡ	16.415
	ΠΕΛΟΠΟΝΝΗΣΟΣ	ΛΑΡΤΟΡ	24.445
MULTIRAMA	ΑΙΓΑΙΟ	ΛΑΡΤΟΡ	8.700
	ΑΤΤΙΚΗ	ΛΑΡΤΟΡ	17.590
	ΘΡΑΚΗ	ΛΑΡΤΟΡ	8.650
PLAISIO	ΚΡΗΤΗ	ΛΑΡΤΟΡ	7.543
	ΠΕΛΟΠΟΝΝΗΣΟΣ	ΛΑΡΤΟΡ	8.590
E-SHOP	ΣΤΕΡΕΑ ΕΛΛΑΔΑ	ΛΑΡΤΟΡ	8.790
	ΣΤΕΡΕΑ ΕΛΛΑΔΑ	ΛΑΡΤΟΡ	7.980

Σχήμα 6-29: Αποτέλεσμα πρώτου παραδείγματος καταστημάτων πληροφορικής

Πηγή: Ιδία έρευνα

Στο αποτέλεσμα φαίνονται ποιες εταιρείες πωλούν υπολογιστές τύπου Laptop, σε ποιες περιφέρειες και ποιες είναι οι πωλήσεις των καταστημάτων ανά περιφέρειες το τρέχον έτος.

Παράδειγμα 2:

Ας υποθέσουμε ότι θέλουμε να εξετάσουμε, ποια καταστήματα πουλάνε προϊόντα μέσω internet και ποιο είναι το συνολικό κόστος των πωλήσεων.

Από το δέντρο διαστάσεων, επιλέγουμε τις διαστάσεις “ΚΑΤΑΣΤΗΜΑΤΑ”, “ΠΡΟΪΟΝΤΑ” και με Drag and Drop τις τοποθετούμε στο πλαίσιο των γραμμών. Με το ίδιο τρόπο, αναπτύσσουμε την διάσταση “ΤΡΟΠΟΣ ΑΓΟΡΑΣ” και από το μέλος “All ΤΡΟΠΟΣ ΑΓΟΡΑΣ ” επιλέγουμε το θυγατρικό μέλος “INTERNET ” και το τοποθετούμε στο πλαίσιο των γραμμών. Τέλος, αναπτύσσουμε την διάσταση των μέτρων και επιλέγουμε το μέτρο “Sum ΚΟΣΤΟΣ ”, αφήνοντάς το στο πλαίσιο των στηλών. Πατώντας το κουμπί “Update”, το αποτέλεσμα θα είναι το εξής:

ΚΑΤΑΣΤΗΜΑΤΑ	Members from ΤΡΟΠΟΣ ΑΓΟΡΑΣ	ΠΡΟΪΟΝΤΑ	Sum ΚΟΣΤΟΣ
E-SHOP	INTERNET	LAPTOP	49.700
		PRINTER	4.300
		SCANNER	4.500
		USB	3.300
MULTIRAMA	INTERNET	LAPTOP	22.800
		PRINTER	4.300
		SCANNER	19.100
PLASIO	INTERNET	USB	600
		PRINTER	4.900
		SCANNER	14.100

Σχήμα 6-30: Αποτέλεσμα δεύτερου παραδείγματος καταστημάτων πληροφορικής

Πηγή: Ιδία έρευνα

Το αποτέλεσμα δείχνει, ποια καταστήματα πουλάνε τα προϊόντα τους μέσω internet και ποιο είναι το συνολικό κόστος των πωλήσεων.

Παράδειγμα 3:

Ας υποθέσουμε ότι θέλουμε να εξετάσουμε, ποιες είναι οι πωλήσεις και το κόστος των προϊόντων στα καταστήματα πληροφορικής και ποιο είναι το ετήσιο κέρδος.

Από το δέντρο διαστάσεων, επιλέγουμε τις διαστάσεις “ΚΑΤΑΣΤΗΜΑΤΑ”, “ΠΡΟΪΟΝΤΑ” και τις σύρουμε, με Drag and Drop, στο πλαίσιο των γραμμών. Από την διάσταση των μέτρων, επιλέγουμε τα μέτρα “Sum ΠΩΛΗΣΕΙΣ”, “Sum ΚΟΣΤΟΣ” και τις τοποθετούμε στο πλαίσιο των στηλών. Πατώντας το κουμπί “Update”, το αποτέλεσμα θα είναι το εξής:

ΚΑΤΑΣΤΗΜΑΤΑ	ΠΡΟΙΟΝΤΑ	Sum ΠΩΛΗΣΕΙΣ	Sum ΚΟΣΤΟΣ
E-SHOP	LAPTOP	104.079	91.900
	PRINTER	9.903	8.200
	SCANNER	5.687	4.500
	USB	7.440	5.385
MULTIRAMA	LAPTOP	59.863	52.400
	PRINTER	25.276	21.500
	SCANNER	28.663	24.100
	USB	8.794	1.500
PLAISIO	LAPTOP	7.980	7.200
	PRINTER	11.480	9.200
	SCANNER	26.657	22.700
	USB	2.457	1.900

Σχήμα 6-31: Αποτέλεσμα τρίτου παραδείγματος καταστημάτων πληροφορικής
 Πηγή: Ίδια έρευνα

Όπως βλέπουμε το Prism μας έδωσε τα καταστήματα, τα προϊόντα που πωλούν αυτά τα καταστήματα καθώς και τις ετήσιες πωλήσεις και κόστος αυτών.

Στα προηγούμενα παραδείγματα χρησιμοποιήσαμε δύο είδη φίλτρων, το φίλτρο με μέτρο και το φίλτρο με αξία. Στο παράδειγμα αυτό θα δείξουμε πως μπορούμε να χρησιμοποιήσουμε τύπους.

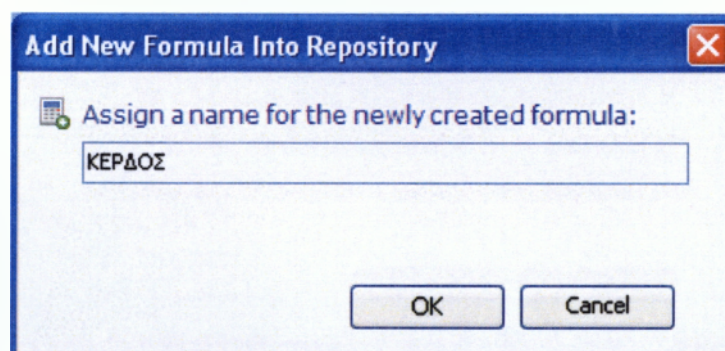
Ας υποθέσουμε ότι θέλουμε να δούμε, το κέρδος των καταστημάτων του παραπάνω παραδείγματος. Από το δέντρο διαστάσεων, αναπτύσσουμε την διάσταση των μέτρων, κάνουμε δεξί κλικ στο “Custom Measures” και επιλέγουμε “New”. Κάτω από την γραμμή εργαλείων θα εμφανιστεί το πλαίσιο, στο οποίο μπορούμε να εισάγουμε τον τύπο. Στο συγκεκριμένο παράδειγμα, θέλουμε να δούμε το κέρδος των καταστημάτων. Αυτό σημαίνει ότι θα πρέπει να γίνει μια αφαίρεση μεταξύ των μέτρων “Sum ΠΩΛΗΣΕΙΣ” και “Sum ΚΟΣΤΟΣ”. Στη συνέχεια, από την διάσταση των μέτρων επιλέγουμε το μέτρο “Sum ΠΩΛΗΣΕΙΣ” και με Drag and Drop το σύρουμε και το αφήνουμε στο πλαίσιο των τύπων. Το ίδιο κάνουμε και με το μέτρο “Sum ΚΟΣΤΟΣ”, τοποθετώντας το δίπλα στο προηγούμενο μέτρο, συμπληρώνοντας ανάμεσά τους το σύμβολο της αφαίρεσης (-), όπως φαίνεται στο σχήμα 6-32 παρακάτω.



Σχήμα 6-32: Πλαίσιο εισαγωγής τύπου

Πηγή: Ιδία έρευνα

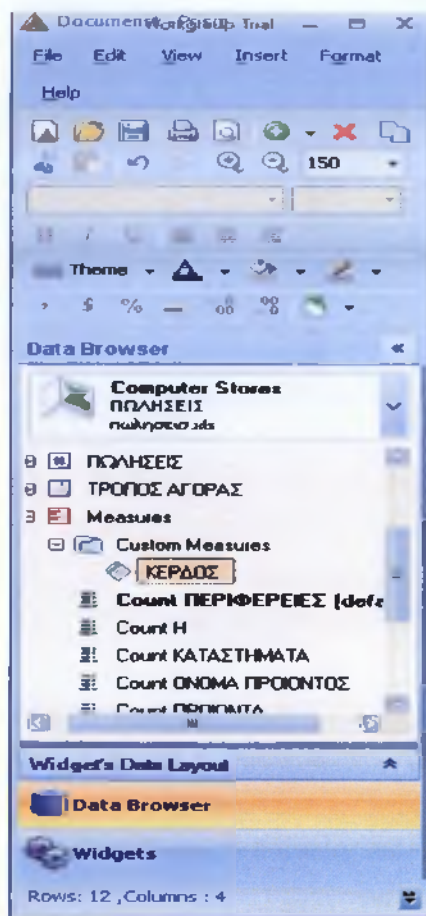
Η πράσινη ένδειξη, στα αριστερά του μέτρου “Sum ΠΩΛΗΣΕΙΣ”, δείχνει ότι ο τύπος έχει γραφτεί σωστά. Πατώντας το κουμπί enter, εμφανίζεται ένα παράθυρο διαλόγου, από το οποίο μας δίνεται η δυνατότητα να ονομάσουμε τον νέο τύπο όπως θέλουμε και με αυτό το όνομα να τοποθετηθεί στο repository. Σε αυτό το παράδειγμα, ονομάζουμε τον νέο τύπο “ΚΕΡΔΟΣ”.



Σχήμα 6-33: Ονομασία τύπου

Πηγή: Ιδία έρευνα

Πατώντας το κουμπί OK, αυτόματα εμφανίζεται ο τύπος στο δέντρο διαστάσεων και συγκεκριμένα στην διάσταση μέτρα με την ονομασία που του είχαμε δώσει, όπως φαίνεται στο σχήμα 6.34 παρακάτω.



Σχήμα 6-34: Εμφάνιση τύπου

Πηγή: Ιδία έρευνα

Κάνοντας αριστερό κλικ πάνω στον τύπο και κρατώντας πατημένο το κουμπί, σύρουμε τον τύπο στον άξονα και τον τοποθετούμε δίπλα από τα μέτρα “Sum ΠΩΛΗΣΕΙΣ” και “Sum ΚΟΣΤΟΣ”. Αφήνοντας το κουμπί, δημιουργείτε μια νέα στήλη δίπλα από τα μέτρα, δείχνοντας το κέρδος των καταστημάτων ανά προϊόν. Το αποτέλεσμα είναι αυτό που φαίνεται στο σχήμα 6-35 παρακάτω.

The screenshot shows an Excel PivotTable titled 'Pivot' with the following data:

ΚΑΤΑΣΤΗΜΑΤΑ	ΠΡΟΪΟΝΤΑ	Sum ΠΩΛΗΣΕΙΣ	Sum ΚΟΣΤΟΣ	ΚΕΡΔΟΣ
E SHOP	LAPTOP	104 079	91 900	12 179
	PRINTER	9 903	8 200	1 703
	SCANNER	5 687	4 500	1 187
	USB	7 440	5 385	2 055
MULTIRAMA	LAPTOP	59 863	52 400	7 463
	PRINTER	25 276	21 500	3 776
	SCANNER	28 663	24 100	4 563
PLAISIO	USB	8 794	1 500	7 294
	LAPTOP	7 980	7 200	780
	PRINTER	11 480	9 200	2 280
	SCANNER	26 657	22 700	3 957
	USB	2 457	1 900	557

Σχήμα 6-35: Αποτέλεσμα τρίτου παραδείγματος καταστημάτων πληροφορικής με εφαρμογή τύπου
 Πηγή: Ιδία έρευνα

Στο παράδειγμα, εκτός από τις πωλήσεις και το κόστος, εμφανίστηκε και το κέρδος των καταστημάτων.

ΣΥΜΠΕΡΑΣΜΑΤΑ

Στην παρούσα πτυχιακή εργασία μελετήθηκε η τεχνολογία OLAP. Η εξέλιξη της τεχνολογίας OLAP οφείλεται στην ραγδαία αύξηση των ποσοτήτων των δεδομένων και στην ανάγκη για ταχεία ανάλυσή τους. Αυτή η ανάλυση γίνεται με πολυδιάστατο τρόπο. Η ανάγκη για πολυδιάστατη ανάλυση προϋποθέτει τη συνεργασία Αποθηκών Δεδομένων και της τεχνολογίας OLAP. Για αυτό το λόγο, οι Αποθήκες Δεδομένων αναδείχθηκαν ως η κύρια πηγή άντλησης των πληροφοριών.

Οι επιχειρήσεις με την εφαρμογή και εγκατάσταση μιας αποθήκης δεδομένων, ταυτόχρονα εφαρμόζουν τεχνικές OLAP οι οποίες συνεργάζονται άμεσα με αυτές. Άλλωστε, η μελέτη και εφαρμογή μιας αποθήκης δεδομένων, γίνεται για να υποστηρίξει την τεχνολογία OLAP, η οποία δεν είναι δυνατό να συνεργαστεί απόλυτα με τις διαχειριστικές βάσεις δεδομένων.

Ο σχεδιασμός των αποθηκών δεδομένων γίνεται έτσι, ώστε να επιτυγχάνεται η ταχύτερη και άμεση προσπέλαση στα δεδομένα από την OLAP εφαρμογή, ή ακόμα σχεδιάζεται για να μειώσει κατά κάποιο βαθμό την επεξεργασία των δεδομένων, που απαιτείται να έχει για μια OLAP εφαρμογή.

Επίσης, μελετήθηκαν ορισμένα εργαλεία OLAP. Από αυτά τα εργαλεία επιλέχθηκε το εργαλείο Sisense Prism Desktop, το οποίο είναι εύχρηστο, μπορεί να συνεργαστεί με διάφορα είδη βάσεων δεδομένων, είναι εύκολο στην εγκατάσταση, προσαρμόζεται εύκολα στις απαιτήσεις του χρήστη και δεν υπάρχει περιορισμός επεξεργασίας διαστάσεων.

Η μελέτη αυτού απέδειξε ότι η τεχνολογία OLAP βοηθάει στην ανάλυση και επεξεργασία μεγάλων επιχειρηματικών βάσεων, ώστε να υπάρχουν άμεσα αποτελέσματα. Επίσης, διαπιστώθηκε ότι η ταχύτητα ανάδειξης των αποτελεσμάτων εξαρτάται από την υπολογιστική ισχύ και το μέγεθος της βάσης δεδομένων που χρησιμοποιεί μια επιχείρηση ή κάποιος χρήστης.

Τέλος, η εργασία απέδειξε ότι η τεχνολογία OLAP είναι ο πιο σύγχρονος και αποδοτικός τρόπος για τη διαχείριση μεγάλου όγκου δεδομένων. Η τεχνολογία OLAP εφαρμόζεται στις περισσότερες χώρες, κυρίως από το 1993, αλλά στην Ελλάδα βρίσκεται ακόμη σε πρώιμο στάδιο.

ΒΙΒΛΙΟΓΡΑΦΙΑ

1. Βασιλειάδης, Π., Σκιαδόπουλος, Σ., και Σελλής, Τ., *Θέματα Μοντελοποίησης και Επερώτησης Πολυδιάστατων Βάσεων Δεδομένων*, Εθνικό Μετσόβιο Πολυτεχνείο, Αθήνα χ.χ.
2. Ρούσσο, Ι., *Εργαστήριο Συστημάτων Βάσεων Γνώσεων και Δεδομένων*, Εθνικό Μετσόβιο Πολυτεχνείο, Αθήνα 2003
3. Rob, P., & Coronel, C., *Database Systems: Design, Implementation, & Management*, Course Technology, Cambridge χ.χ.

ΠΗΓΕΣ ΔΙΑΔΙΚΤΥΟΥ

Ελληνικές

1. Γιατί “Ο Ανοιχτός Κώδικας” χάνει το νόημα του ελεύθερου λογισμικού, <http://www.gnu.org/philosophy/open-source-misses-the-point.el.html>
2. Ελεύθερο Λογισμικό/Λογισμικό Ανοικτού Κώδικα, <http://www.opensource.gr/faq.php?language=el>,
<http://www.open-source.gr/whatis.php?language=el>,
<http://www.open-source.gr/faq.php?ctn=1&pos=10&language=el>
1. Επισκόπηση της ηλεκτρονικής διαδικασίας ανάλυσης (On-Line Analytical Processing – OLAP), <http://office.microsoft.com/el-gr/excel/HP101774371032.aspx>
2. Προς ένα Ολοκληρωμένο Πληροφοριακό Σύστημα Συλλογής, Αποθήκευσης και Επεξεργασίας Εκπαιδευτικών Δεδομένων: Από το Όραμα στο Σύστημα Edubase, <http://www.etpe.gr/files/proceedings/uploads/p113.pdf>
3. Συστήματα Υποστήριξης Αποφάσεων, Κεφ.11, Data Warehouse and OLAP, <http://www.ergasya.tuc.gr/Users/Matsatsinis/Courses/SYA/sya.htm>
4. Τι είναι OLAP, <http://www.tech-faq.com/lang/el/olap.shtml>

Ξένες

1. Connecting to Excel, http://www.sisense.com/Support/connect_to_data.aspx
2. Definitions, <http://sisense.com/help/learning-center/definitions/>
3. Filter By Attribute, <http://sisense.com/help/learning-center/filter-by-attribute/>
4. Filter By Label, <http://sisense.com/help/learning-center/filter-by-label/>
5. Filter by Measure, <http://sisense.com/help/learning-center/filter-by-measure/>

6. *Filter By Value*, <http://sisense.com/help/learning-center/filter-by-value/>
7. *Introduction to Filter*, <http://sisense.com/help/learning-center/introduction-to-filters/>
8. *Introduction to Formulas*, <http://sisense.com/help/learning-center/introduction-to-formulas/>
9. Jedox Palo OLAP Server-Open Source Business Intelligence for Planning, Analysis, Reporting and Consolidation, <http://www.jedox.com/en/enterprise-spreadsheet-server/excel-olap-server/palo-server.html>
10. *Numeric Formulas*, <http://sisense.com/help/learning-center/numeric-formulas/>
11. *Olap X Application 3.0*, <http://www.download25.com/olap-x-application-download.html>
12. *OlapCube 1.1*, <http://www.download25.com/olapcube-download.html>
13. *Pentaho Analysis Services: Mondrian Project*, <http://mondrian.pentaho.org/>
14. *Prism FAQ*, <http://sisense.com/help/learning-center/prism-faq/>
15. *The Prism Environment*, <http://sisense.com/help/learning-center/the-prism-environment/>
16. *Working with Pivot tables*, http://www.sisense.com/Support/working_with_pivots.aspx

ΠΑΡΑΡΤΗΜΑ

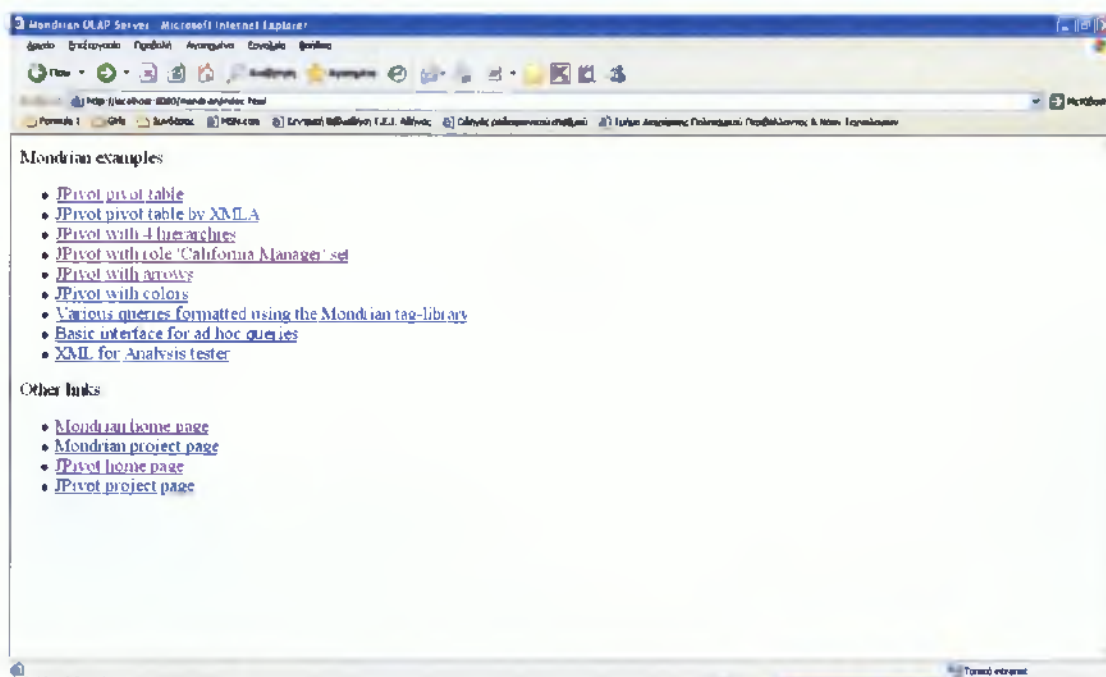
ΕΦΑΡΜΟΓΗ ΕΡΓΑΛΕΙΟΥ MONDRIAN

ΠΕΡΙΕΧΟΜΕΝΑ ΣΧΗΜΑΤΩΝ ΠΑΡΑΡΤΗΜΑΤΟΣ

Σχήμα I: Mondrian examples.....	1
Σχήμα II: Test Query.....	2
Σχήμα III: Access Data Source.....	2
Σχήμα IV: Συγκεντρωτικός πίνακας.....	3
Σχήμα V: Query Buttons.....	3
Σχήμα VI: Olap Navigator.....	4
Σχήμα VII: Επιλογή διαστάσεων.....	4
Σχήμα VIII: Επιλογή μέτρου.....	5
Σχήμα IX: Επιλογή θυγατρικού μέλους USA.....	6
Σχήμα X: Επιλογή θυγατρικού μέλους Supermarket.....	6
Σχήμα XI: Επιλογή θυγατρικού μέλους Customers USA.....	7
Σχήμα XII: Επιλογή θυγατρικού μέλους Marital.....	8
Σχήμα XIII: Επιλογή θυγατρικού μέλους \$150K+.....	8
Σχήμα XIV: Επιλογή θυγατρικού μέλους Drink.....	9
Σχήμα XV: Αποτέλεσμα ερωτήματος.....	9
Σχήμα XVI: Ανάλυση διαστάσεων.....	10

Όπως έχουμε αναφέρει στο Κεφάλαιο 4, το Mondrian είναι ένα εργαλείο ανοιχτού κώδικα, όπου επιτρέπει την ανάλυση μεγάλων ποσοτήτων δεδομένων τα οποία είναι αποθηκευμένα σε SQL βάσεις δεδομένων, χωρίς να χρησιμοποιείται η SQL. Παρέχει την δυνατότητα απόκτησής του δωρεάν, μέσω του διαδικτύου σε ολοκληρωμένη μορφή. Επίσης, η εισαγωγή των δεδομένων του γίνεται μέσω της Microsoft Access.

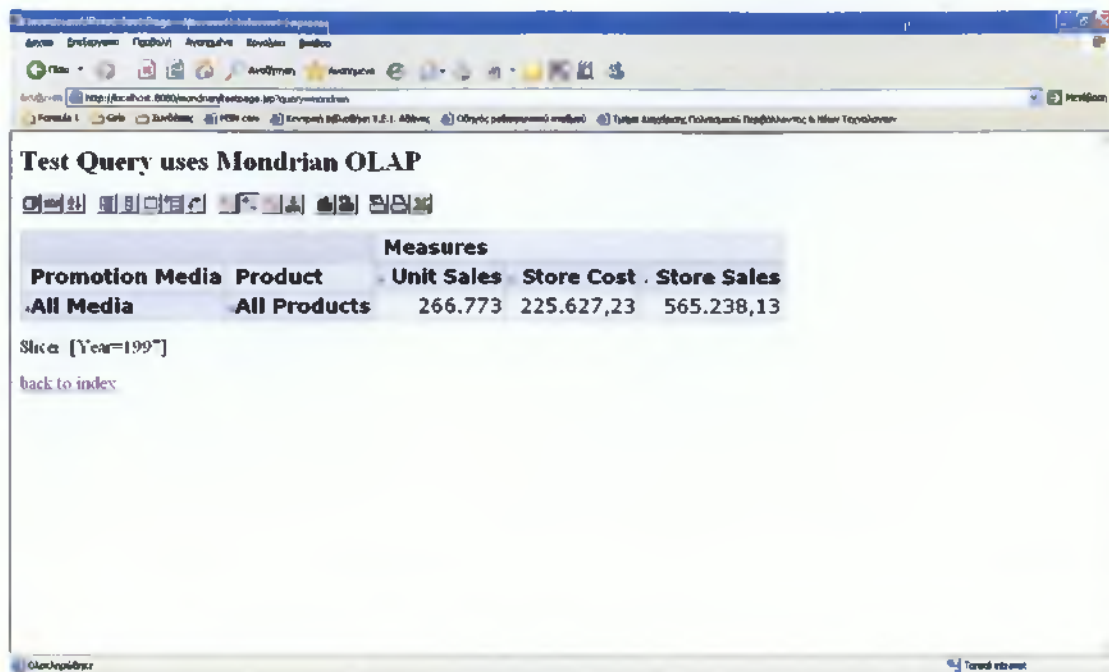
Έχοντας εγκαταστήσει το Mondrian, ανοίγουμε τον Internet Explorer και πληκτρολογούμε την διεύθυνση <http://localhost:8080/mondrian/index.html>. Το περιεχόμενο αυτής της σελίδας φαίνεται στο σχήμα I.



Σχήμα I: Mondrian examples

Πηγή: Ιδία επεξεργασία

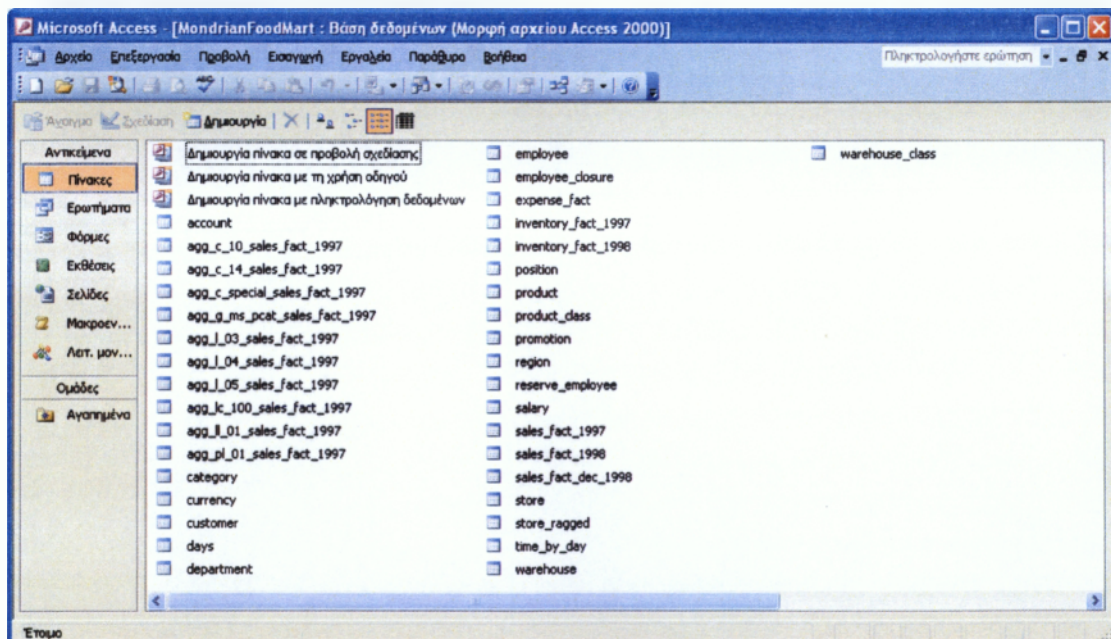
Από τις επιλογές που φαίνονται παραπάνω, επιλέγουμε την “JPivot pivot table”, κάνοντας ένα απλό κλικ πάνω της και εισερχόμαστε στον άξονα του Mondrian. Εκεί, μπορούμε να δούμε το αποτέλεσμα, του ερωτήματος που θέλουμε να θέσουμε.



Σχήμα II: Test Query

Πηγή: Ιδία επεξεργασία

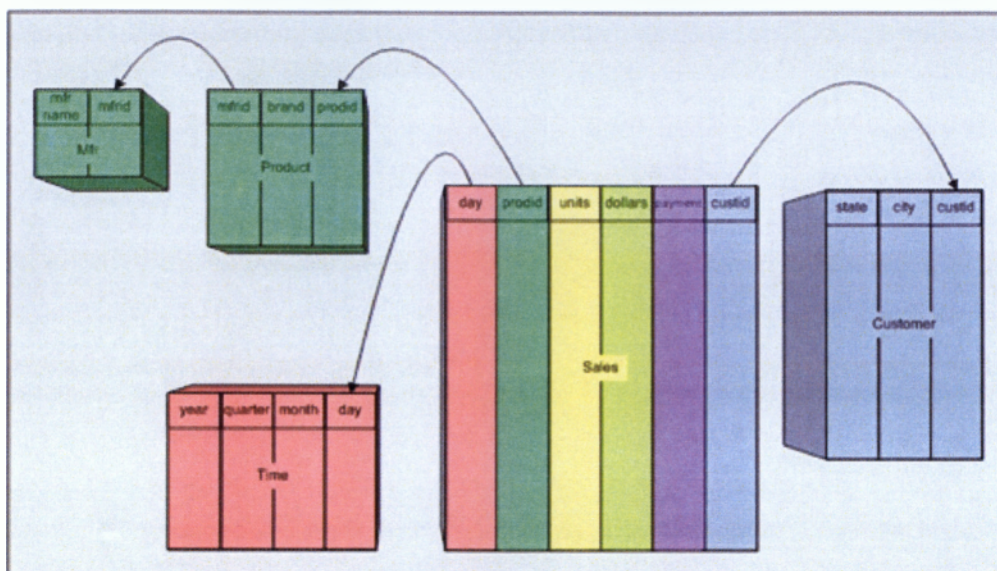
Το Mondrian περιέχει μια πηγή δεδομένων τύπου Microsoft Access, στην οποία περιέχεται ένας μεγάλος αριθμός πινάκων, όπως φαίνεται στο παρακάτω σχήμα.



Σχήμα III: Access Data Source

Πηγή: Πηγή: Ιδία επεξεργασία

Όταν εισάγουμε ένα ερώτημα, το Mondrian “παίρνει” τα δεδομένα και τις στήλες που χρειάζεται από τους πίνακες και τα παρουσιάζει σε έναν ενιαίο πίνακα, όπως φαίνεται χαρακτηριστικά στο παρακάτω σχήμα.



Σχήμα IV: Συγκεντρωτικός πίνακας

Πηγή: *Pentaho Analysis Services: Mondrian Project*, <http://mondrian.pentaho.org/>

Ας δούμε όμως ένα παράδειγμα εφαρμογής του Mondrian, ώστε να γίνει πιο κατανοητό.

Παράδειγμα:

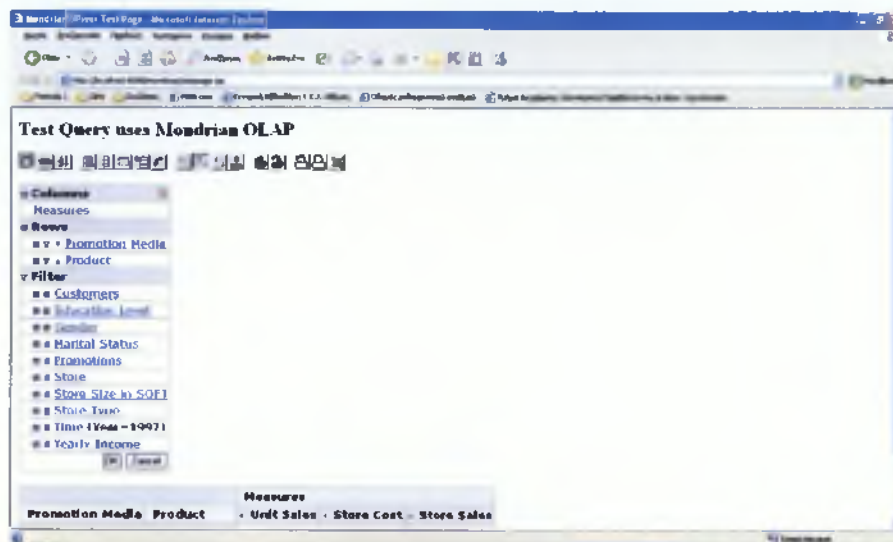
Ας υποθέσουμε ότι θέλουμε να εξετάσουμε, ποιες είναι οι πωλήσεις των σουπερμάρκετ της Αμερικής (USA) σε ποτά, με πελάτες που έχουν ετήσιο εισόδημα πάνω από 150.000 \$ και είναι έγγαμοι.



Σχήμα V: Query Buttons

Πηγή: Ίδια επεξεργασία

Από τα Query Buttons που φαίνονται στο σχήμα V, κάνουμε κλικ στο αριστερό κουμπί που έχει την ονομασία “Open OLAP Navigator”. Αμέσως στην οθόνη μας, θα εμφανιστεί το πεδίο των στηλών (Columns), των γραμμών (Rows) και του φίλτρου (Filter), όπως φαίνεται στο σχήμα VI.



Σχήμα VI: Olap Navigator

Πηγή: Ιδία επεξεργασία

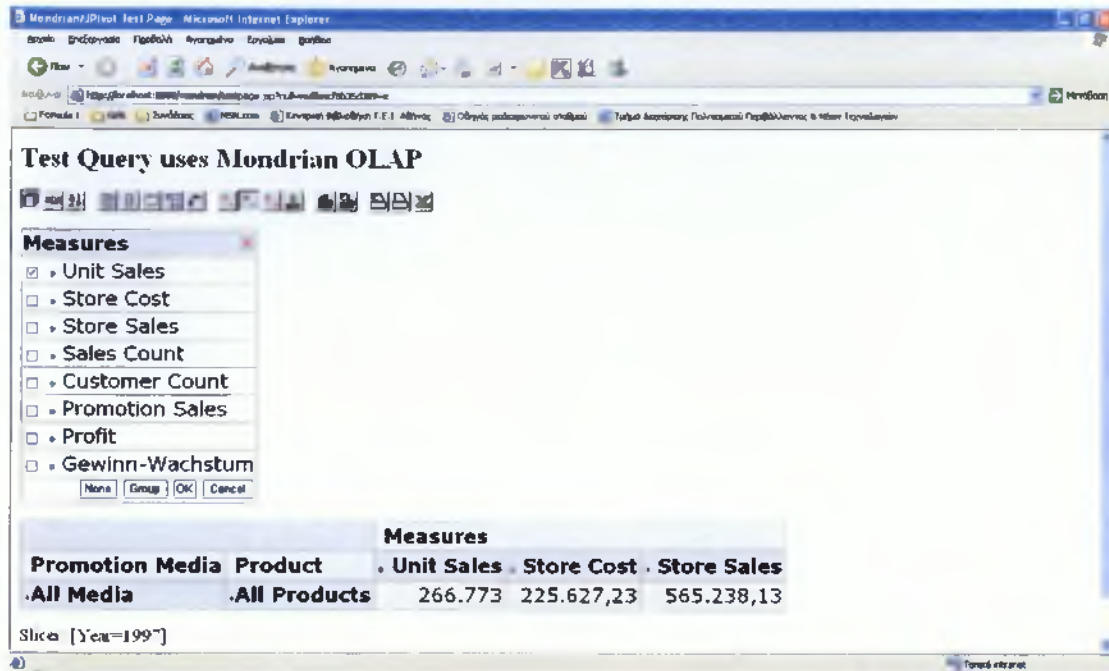
Με βάση τα πεδία αυτά τοποθετούμε τις διαστάσεις στα πλαίσια των γραμμών και των στηλών, πατώντας τις επιλογές που υπάρχουν αριστερά των διαστάσεων, ανάλογα με το ερώτημα που θέλουμε να εξετάσουμε, όπως φαίνεται παρακάτω.



Σχήμα VII: Επιλογή διαστάσεων

Πηγή: Ιδία επεξεργασία

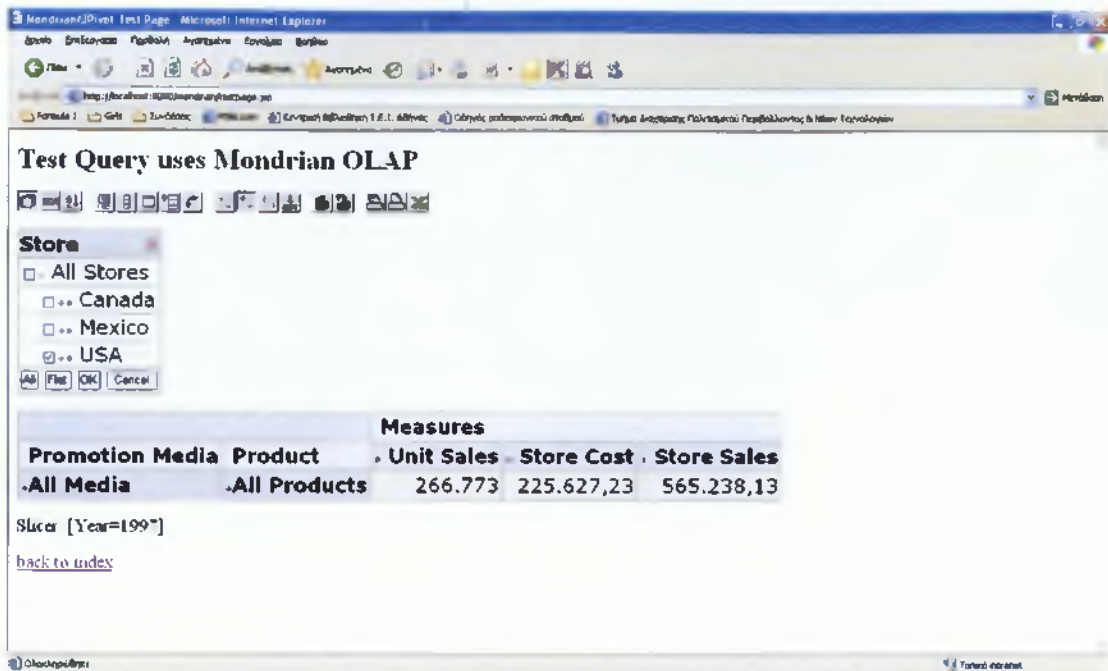
Πηγαίνουμε στο πλαίσιο των στηλών και κάνουμε κλικ πάνω στα Measures. Το παράθυρο που ανοίγει είναι το εξής:



Σχήμα VIII: Επιλογή μέτρου

Πηγή: Ιδία επεξεργασία

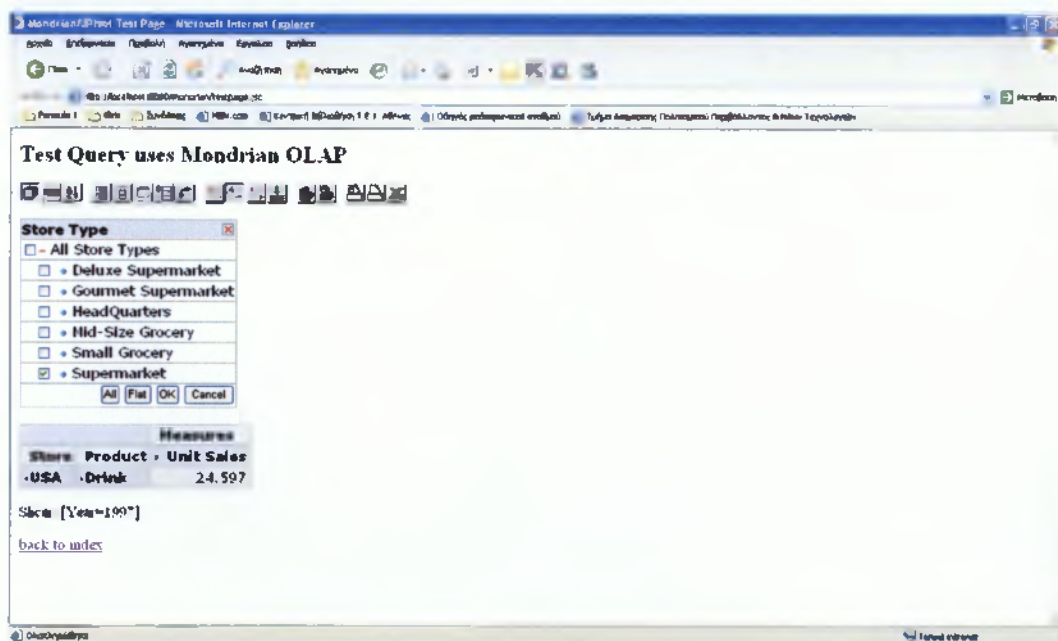
Από τα μέτρα που εμφανίζονται επιλέγουμε μόνο το “Unit Sales”, όπως φαίνεται παραπάνω και πατάμε OK. Επιστρέφουμε στο μενού του σχήματος VII και κάνουμε κλικ στην διάσταση “Store” του πλαισίου των γραμμών. Αναπτύσσουμε το μέλος “All Stores” και επιλέγουμε μόνο το θυγατρικό μέλος “USA”, όπως φαίνεται στο σχήμα IX.



Σχήμα IX: Επιλογή θυγατρικού μέλους USA

Πηγή: Ίδια επεξεργασία

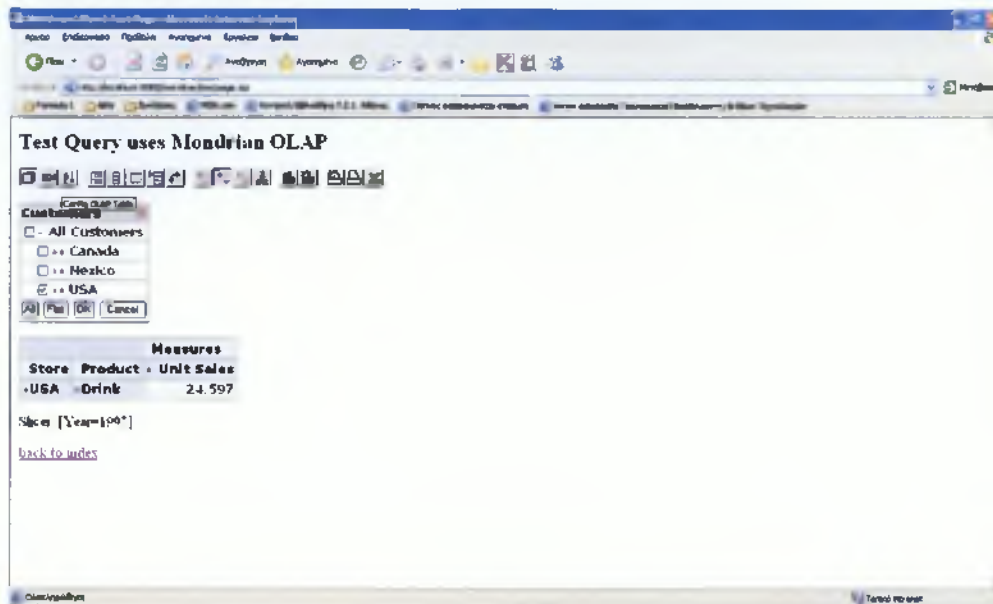
Κάνοντας κλικ στο κουμπί OK, επιστρέφουμε στο μενού του σχήματος VII και κάνουμε κλικ στην διάσταση "Store Type" του πλαισίου των γραμμών. Αναπτύσσουμε το μέλος "All Store Types" και επιλέγουμε μόνο το θυγατρικό μέλος "Supermarket", όπως φαίνεται στο σχήμα X.



Σχήμα X: Επιλογή θυγατρικού μέλους Supermarket

Πηγή: Ίδια επεξεργασία

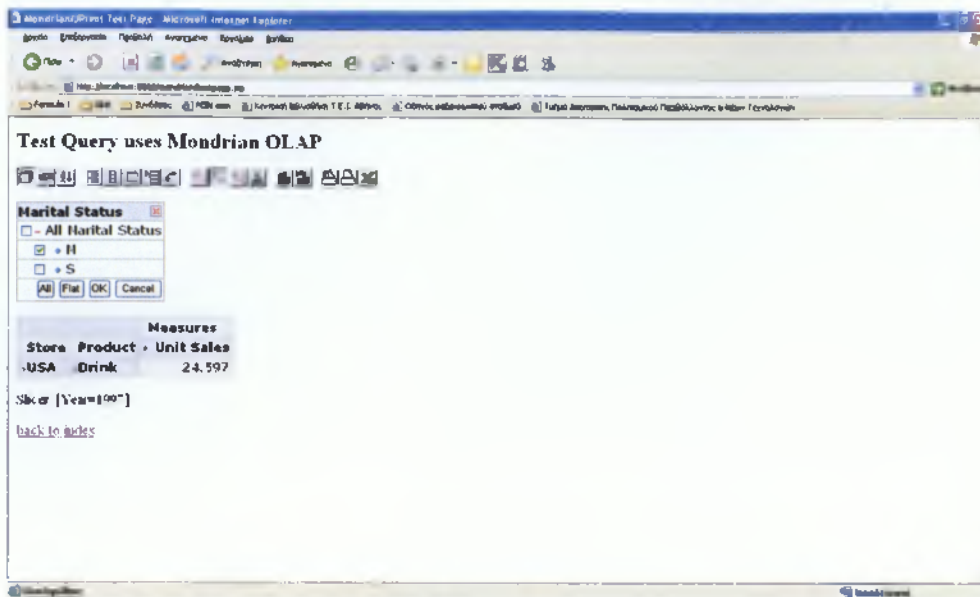
Κάνοντας κλικ στο κουμπί OK, επιστρέφουμε στο μενού του σχήματος VII και κάνουμε κλικ στην διάσταση “Customers” του πλαισίου των γραμμών. Αναπτύσσουμε το μέλος “All Customers” και επιλέγουμε μόνο το θυγατρικό μέλος “USA”, όπως φαίνεται στο σχήμα XI.



Σχήμα XI: Επιλογή θυγατρικού μέλους Customers USA

Πηγή: Ιδία επεξεργασία

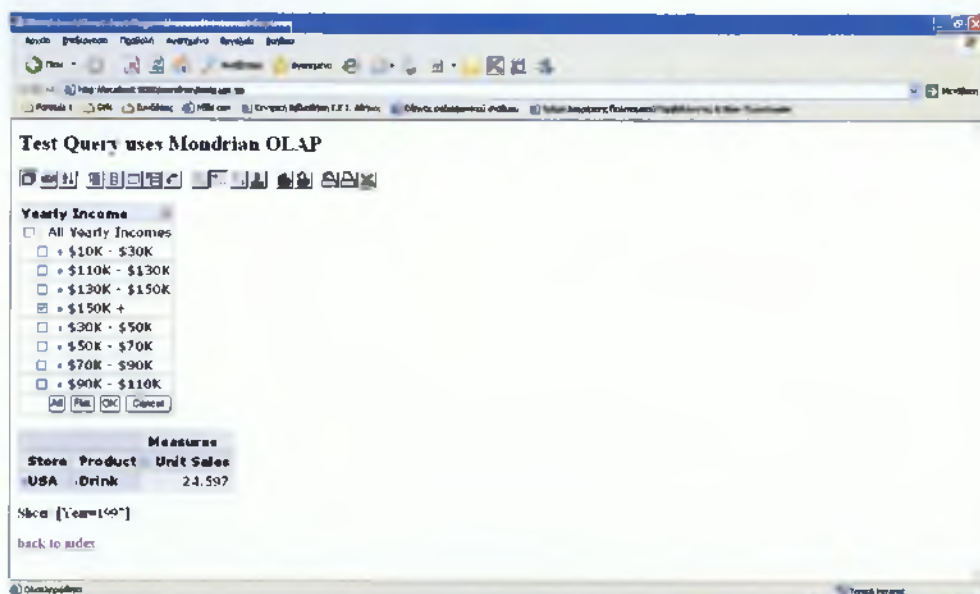
Κάνοντας κλικ στο κουμπί OK, επιστρέφουμε στο μενού του σχήματος VII και κάνουμε κλικ στην διάσταση “Marital Status” του πλαισίου των γραμμών. Αναπτύσσουμε το μέλος “All Marital Status” και επιλέγουμε μόνο το θυγατρικό μέλος “Marital”, όπως φαίνεται στο σχήμα XII.



Σχήμα XII: Επιλογή θυγατρικού μέλους Marital

Πηγή: Ίδια επεξεργασία

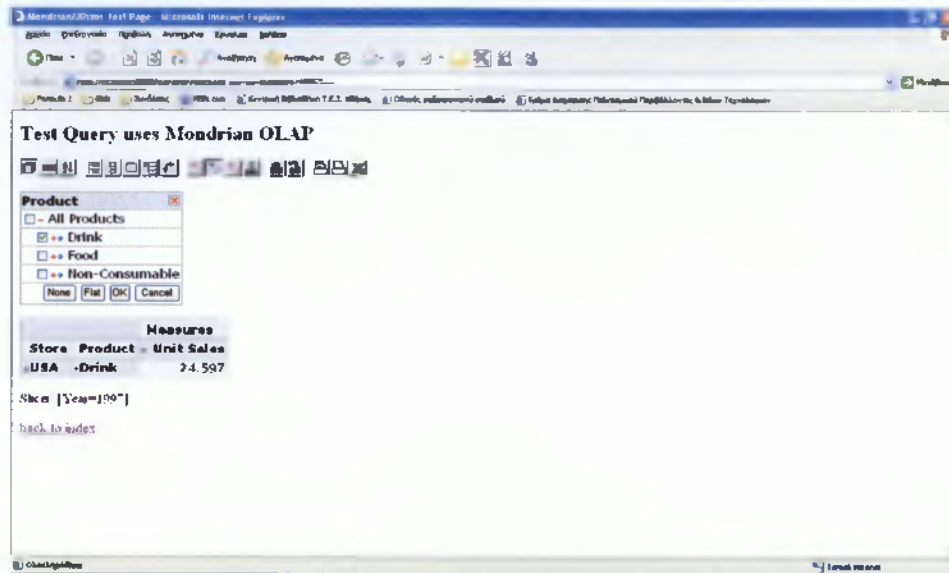
Κάνοντας κλικ στο κουμπί OK, επιστρέφουμε στο μενού του σχήματος VII και κάνουμε κλικ στην διάσταση “Yearly Income” του πλαισίου των γραμμών. Αναπτύσσουμε το μέλος “All Yearly Incomes ” και επιλέγουμε μόνο το θυγατρικό μέλος “\$150K +”, όπως φαίνεται στο σχήμα XIII.



Σχήμα XIII: Επιλογή θυγατρικού μέλους \$150K +

Πηγή: Ίδια επεξεργασία

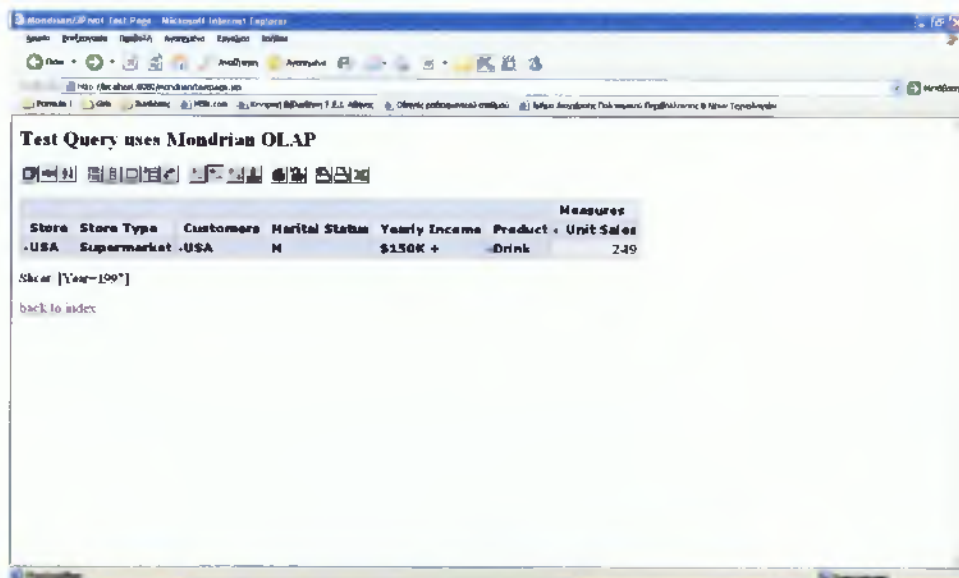
Κάνοντας κλικ στο κουμπί OK, επιστρέφουμε στο μενού του σχήματος VII και κάνουμε κλικ στην διάσταση “Product” του πλαισίου των γραμμών. Αναπτύσσουμε το μέλος “All Products” και επιλέγουμε μόνο το θυγατρικό μέλος “Drink”, όπως φαίνεται στο σχήμα XIV.



Σχήμα XIV: Επιλογή θυγατρικού μέλους Drink

Πηγή: Ίδια επεξεργασία

Κάνοντας κλικ στο κουμπί OK, επιστρέφουμε στο μενού του σχήματος VII και πατώντας ξανά OK, έχουμε το εξής αποτέλεσμα:



Σχήμα XV: Αποτέλεσμα ερωτήματος

Πηγή: Ίδια επεξεργασία

Στο σχήμα XV, παρουσιάζονται όλες οι διαστάσεις που ζητήθηκαν στο ερώτημα. Πατώντας το σύμβολο (+) αριστερά των διαστάσεων, οι διαστάσεις αναλύονται περισσότερο (σχήμα XVI). Το ίδιο ισχύει και για τα μέλη των διαστάσεων.

Test Query uses Mondrian OLAP

						Measures
Store	Store Type	Customers	Marital Status	Yearly Income	Product	Unit Sales
USA	Supermarket	USA	M	\$150K +	Drink	249
		CA	M	\$150K +	Drink	87
		OR	M	\$150K +	Drink	29
		WA	M	\$150K +	Drink	133
CA	Supermarket	USA	M	\$150K +	Drink	87
OR	Supermarket	USA	M	\$150K +	Drink	29
WA	Supermarket	USA	M	\$150K +	Drink	133

Slice [Year=1997]
[back to index](#)

Σχήμα XVI: Ανάλυση διαστάσεων

Πηγή: Ιδία επεξεργασία