



## **ΔΙΑΤΜΗΜΑΤΙΚΟ ΠΡΟΓΡΑΜΜΑ ΜΕΤΑΠΤΥΧΙΑΚΩΝ ΣΠΟΥΔΩΝ ΣΤΑ ΠΛΗΡΟΦΟΡΙΑΚΑ ΣΥΣΤΗΜΑΤΑ**

Διπλωματική Εργασία

### **Η ΕΠΙΧΕΙΡΗΜΑΤΙΚΗ ΕΥΦΥΪΑ ΣΤΗ ΣΥΓΧΡΟΝΗ ΕΠΙΧΕΙΡΗΣΗ ΚΑΙ Η ΑΝΑΛΥΤΙΚΗ ΕΠΕΞΕΡΓΑΣΙΑ ΔΕΔΟΜΕΝΩΝ (OLAP) – ΠΡΑΚΤΙΚΗ ΕΦΑΡΜΟΓΗ**

ΥΦΑΝΤΙΔΗΣ Λ. ΔΗΜΗΤΡΙΟΣ  
ΕΠΙΒΛΕΠΩΝ ΚΑΘΗΓΗΤΗΣ: ΕΥΑΓΓΕΛΙΔΗΣ ΓΕΩΡΓΙΟΣ

Υποβλήθηκε ως απαιτούμενο για την απόκτηση του Μεταπτυχιακού  
Διπλώματος Ειδίκευσης στα Πληροφορικά Συστήματα

Θεσσαλονίκη, Σεπτέμβριος 2018

## Αφιερώσεις

Στην κόρη μου Φαίδρα, που πριν τρεις μήνες ήρθε στον κόσμο και μου άλλαξε τη ζωή.

## Ευχαριστίες

Η παρακολούθηση και ολοκλήρωση του Διατμηματικού Προγράμματος Μεταπτυχιακών Σπουδών στα Πληροφοριακά Συστήματα απαιτεί προσήλωση, διαρκή προσπάθεια και χρόνο. Σε όλα τα παραπάνω είχα την αμέριστη συμπαράσταση της συζύγου μου την οποία και ευχαριστώ.

Επίσης, ευχαριστώ θερμά τον επιβλέποντα καθηγητή μου κ. Ευαγγελίδη Γεώργιο για τη στήριξή του και την υπομονή που υπέδειξε μέχρι την ολοκλήρωση της διπλωματικής μου εργασίας, καθώς και για το απαραίτητο υλικό που μου παρείχε.

Τέλος, θα ήθελα να ευχαριστήσω την κα Μιχαηλίδου Ναταλία Δήμητρα για την πολύτιμη τεχνική υποστήριξη που μου προσέφερε.

## Περίληψη

Η παρούσα μελέτη αναδεικνύει την αξία που προσδίδει στη σύγχρονη επιχείρηση η επιχειρηματική ευφυΐα και ειδικότερα η ανάλυση δεδομένων μέσω των εργαλείων OLAP. Αρχικά, γίνεται μια εκτενής θεωρητική ανασκόπηση γύρω από το εξεταζόμενο αντικείμενο, προκειμένου να γίνει κατανοητό στον αναγνώστη, και στη συνέχεια μελετάται η περίπτωση επιχείρησης του τριτογενούς τομέα παραγωγής, εντάσεως οικονομίας, όπου δημιουργείται μια βάση δεδομένων στον MS SQL Server με τα στοιχεία των πωλήσεών της, τα οποία έχουν εξαχθεί από τα συστήματα συναλλαγών της επιχείρησης σε λογιστικά φύλλα MS Excel, στη συνέχεια δημιουργείται ένα πολυδιάστατο σχήμα (κύβος OLAP) με τα MS SQL Server Data Tools και το MS Visual Studio και τέλος εκτελούνται ορισμένα ερωτήματα και δημιουργούνται αναφορές με το TARGIT Decision Suite από τα οποία και προκύπτει η ουσιαστική και πολύτιμη πληροφορία που σχετίζεται με τα καταστήματα, τους υπαλλήλους, τους πελάτες και τα προϊόντα της επιχείρησης και βοηθάει στον εντοπισμό δυνατών σημείων και αδυναμιών της, καθώς επίσης προσφέρει χρήσιμη πληροφορία για τη λήψη σημαντικών αποφάσεων. Βάσει της προκύπτουσας γνώσης προτείνεται μια άμεσα εφαρμοστέα λύση στην επιχείρηση αλλά και μελλοντικές προσθήκες προκειμένου να επιτευχθεί ακόμη μεγαλύτερη παραγωγή πολύτιμης πληροφορίας.

## **Abstract**

This study highlights the value that business intelligence and in particular data analysis through OLAP tools attributes to the modern company. Initially, it provides a thorough review regarding the theoretical framework of the object being discussed, in order to provide better insight to the reader. Subsequently, it describes the case of a heavy economically related company of the tertiary sector of production, where a database is created in MS SQL Server with the provided sales data, which have been extracted in MS Excel spreadsheets from the company's transaction systems, subsequently a multi-dimensional schema (OLAP cube) is created with the MS SQL Server Data Tools and MS Visual Studio and finally specific queries are being executed and reports are being produced with TARGIT Decision Suite to provide meaningful and valuable information, related to the company's stores, employees, customers and products, that helps identifying potential strengths and weaknesses, as well as providing useful information for decision support. Based on the resulting knowledge, a directly applicable solution is proposed to the company as well as future additions to achieve even greater production of valuable information.

## Πίνακας Περιεχομένων

Αφιέρώσεις.....	ii
Ευχαριστίες.....	iii
Περίληψη.....	iv
Abstract.....	v
Πίνακας Περιεχομένων.....	vi
Πίνακας Εικόνων.....	viii
1. Εισαγωγή.....	1
2. Ανασκόπηση Βιβλιογραφίας.....	2
2.1. Επιχειρηματική Ευφυΐα.....	2
2.1.1. Ορισμός.....	2
2.1.2. Η αξία της Επιχειρηματικής Ευφυΐας.....	2
2.1.3. Οι περιορισμοί της Επιχειρηματικής Ευφυΐας.....	5
2.1.4. Ανάλυση συστήματος Επιχειρηματικής Ευφυΐας.....	6
2.1.4.1. Επίπεδο Πηγών Δεδομένων.....	7
2.1.4.2. Επίπεδο διεργασιών ETL.....	8
2.1.4.3. Επίπεδο Αποθήκης Δεδομένων.....	10
2.1.4.4. Επίπεδο Τελικού Χρήστη.....	12
2.1.4.5. Επίπεδο Μεταδεδομένων.....	12
2.2. Οι Αποθήκες Δεδομένων αναλυτικότερα.....	14
2.2.1. Δομή Αποθήκης Δεδομένων.....	14
2.2.2. Σχήματα Αποθήκης Δεδομένων.....	16
2.3. OLAP.....	20
2.3.1. Χρησιμότητα της OLAP.....	20
2.3.2. Βασικοί τύποι OLAP.....	21
2.3.2.1. ROLAP.....	21
2.3.2.2. MOLAP.....	22
2.3.2.3. HOLAP.....	23
2.3.3. Χαρακτηριστικά εργαλείων OLAP.....	24
2.3.4. Κύβος.....	27
2.3.5. Ιεραρχίες.....	28
2.3.6. Βασικές λειτουργίες των κύβων OLAP.....	30
2.3.6.1. Περιστροφή (Pivot).....	30

2.3.6.2.	Συναθροιστική Άνοδος & Αναλυτική Κάθοδος (Roll-up & Drill-down)	31
2.3.6.3.	Οριζόντιος και Κάθετος Τεμαχισμός (Slice & Dice)	32
3.	Μεθοδολογία	34
3.1.	Πρόλογος	34
3.2.	Παρουσίαση Επιχείρησης	34
3.3.	Παρουσίαση Λογισμικού	37
3.3.1.	MS Excel	37
3.3.2.	Visual Paradigm	38
3.3.3.	MS SQL Server	38
3.3.4.	MS Visual Studio και SQL Server Data Tools	39
3.3.5.	TARGIT Decision Suite	39
3.4.	Σχεδιασμός και εκτέλεση	40
4.	Ανάλυση Δεδομένων και Προτάσεις	64
	Κατάλογος Αναφορών και Παραπομπών	67

## Πίνακας Εικόνων

Εικόνα 1: Οι προτεραιότητες των CEO's μεγάλων επιχειρήσεων (Πηγή: Gartner 2017)	3
Εικόνα 2: Τα πέντε επίπεδα ενός συστήματος επιχειρηματικής ευφυΐας (Lih Ong, Hwa Siew, & Wong, 2011)	7
Εικόνα 3: Οι διεργασίες ETL πλαισιωμένες από τις διεργασίες Profiling, Analyzing, Monitoring και Cleansing (element61 N.V. – Moore Stephens, 2018)	8
Εικόνα 4: Η δομή της Αποθήκης Δεδομένων (Inmon, 2005)	15
Εικόνα 5: Σχήμα Αστέρα (Moody & Kortink, 2000)	17
Εικόνα 6: Σχήμα Αστερισμού (Moody & Kortink, 2000)	18
Εικόνα 7: Σχήμα Χιονονιφάδας (Moody & Kortink, 2000)	19
Εικόνα 8: Η αρχιτεκτονική της σχεσιακής OLAP (Saranya, 2013)	21
Εικόνα 9: Η αρχιτεκτονική της πολυδιάστατης OLAP (Saranya, 2013)	22
Εικόνα 10: Η αρχιτεκτονική της υβριδικής OLAP (Saranya, 2013)	23
Εικόνα 11: Ο κύβος OLAP (Loudcher, Jakawat, Pavel Soriano Morales, & Favre, 2015)	28
Εικόνα 12: Γραμμική ιεραρχία (Πηγή: πρωτότυπο)	29
Εικόνα 13: Ιεραρχία πλέγματος (Πηγή: πρωτότυπο)	29
Εικόνα 14: Η λειτουργία Pivot (California State University, 2018)	30
Εικόνα 15: Οι λειτουργίες Roll-up και Drill-down (California State University, 2018)	31
Εικόνα 16: Οι λειτουργίες Slice και Dice (California State University, 2018)	32
Εικόνα 17: Τα βήματα της μελέτης κατά την παρουσίαση της πρακτικής εφαρμογής (Πηγή: πρωτότυπο)	34
Εικόνα 18: Αρχική μορφή δεδομένων εξαγμένων από τα συστήματα συναλλαγών της επιχείρησης	35
Εικόνα 19: Λειτουργία "Εύρεσης και Αντικατάστασης" στο Excel	36
Εικόνα 20: Αρχικό μενού Visual Paradigm	40
Εικόνα 21: Επιλογές στοιχείων για τη σχεδίαση του σχήματος της αποθήκης δεδομένων	41
Εικόνα 22: Σχήμα χιονονιφάδας της αποθήκης δεδομένων	42
Εικόνα 23: Εισαγωγή φίλτρου στα δεδομένα του αρχείου υπολογιστικών φύλλων	44
Εικόνα 24: Καταχωρήσεις πίνακα fact_sales	45
Εικόνα 25: Καταχωρήσεις πίνακα dim_time	45
Εικόνα 26: Δημιουργία της βάσης δεδομένων στον MS SQL Server	46
Εικόνα 27: Εισαγωγή δεδομένων στη βάση	47
Εικόνα 28: Δημιουργία έργου στο Analysis Services	48
Εικόνα 29: Σύνδεση κύβου με βάση δεδομένων	49
Εικόνα 30: Επιλογή πινάκων για χρήση στον κύβο	50
Εικόνα 31: Περιγραφή σύνδεσης μεταξύ των πινάκων	51
Εικόνα 32: Επιλογή διαστάσεων	52
Εικόνα 33: Εισαγωγή διαστάσεων και μέτρων στον κύβο	53
Εικόνα 34: Τελική διαμόρφωση κύβου	54
Εικόνα 35: Έργο για την ανανέωση των δεδομένων της βάσης και του κύβου	55



Εικόνα 36: Ανάθεση έργου σε εργασία του SQL Server Agent .....	56
Εικόνα 37: Επαναληψιμότητα έργου.....	56
Εικόνα 38: Σύνδεση κύβου με TARGIT Decision Suite.....	57
Εικόνα 39: Ημερήσια αναφορά .....	58
Εικόνα 40: Εβδομαδιαία αναφορά .....	59
Εικόνα 41: Μηνιαία αναφορά .....	60
Εικόνα 42: Ετήσια αναφορά.....	61
Εικόνα 43: Πωλήσεις σε σχέση με το φύλο του πωλητή (0: Γυναίκα, 1: Άντρας).....	62
Εικόνα 44: Πωλήσεις ανάλογα με το ωράριο του πωλητή.....	63
Εικόνα 45: Πωλήσεις σε ιδιαίτερες περιόδους.....	63



## 1. Εισαγωγή

Οι επιχειρήσεις της σύγχρονης εποχής αντιμετωπίζουν αστάθμητες προκλήσεις λόγω του έντονα ασταθούς οικονομικού περιβάλλοντος όσο και της συνεχούς και με αλματώδεις ρυθμούς εξέλιξης της τεχνολογίας. Για να επιβιώσει αλλά και να ξεχωρίσει μια επιχείρηση πρέπει να είναι έτοιμη ανά πάσα στιγμή να αφομοιώνει τις αλλαγές και να προσαρμόζει αναλόγως τις πρακτικές της σε θέματα διοίκησης και τρόπου λήψης αποφάσεων. Προκειμένου να μπορεί να επιτύχει το ζητούμενο πρέπει να είναι σε θέση να αντιλαμβάνεται πλήρως τόσο το εξωτερικό όσο και το εσωτερικό της περιβάλλον και να έχει τη δυνατότητα να παράγει τη σωστή πληροφορία στην κατάλληλη χρονική στιγμή. Η επιχειρηματική ευφυΐα, ως ένας συνεχώς αναπτυσσόμενος επιστημονικός τομέας, έρχεται να καλύψει αυτήν την ανάγκη.

Η επιχειρηματική ευφυΐα επικεντρώνεται κυρίως στο να εντοπίζει, να εξάγει και να αναλύει σημαντικά επιχειρησιακά δεδομένα. Σκοπός της μελέτης είναι να εξηγήσει το θεωρητικό πλαίσιο γύρω από την επιχειρηματική ευφυΐα και των συστημάτων αυτής, τις έννοιες των αποθηκών δεδομένων και της OLAP, ως εργαλείου ανάλυσης των δεδομένων, καθώς και να παρουσιάσει μια πρακτική εφαρμογή αυτών, προκειμένου να αναδείξει στο δυνητικό χρήστη-στέλεχος επιχείρησης τη σημαντικότητα και τη χρησιμότητα αυτού του τομέα.

Στο πλαίσιο συγγραφής της μελέτης θα γίνει ανασκόπηση της θεωρίας που περιβάλλει την επιχειρηματική ευφυΐα, τις αποθήκες δεδομένων και τα εργαλεία OLAP. Για την πρακτική εφαρμογή θα δημιουργηθεί μια βάση δεδομένων με στοιχεία επιχείρησης του οικονομικού κλάδου, στη συνέχεια θα αναπτυχθεί ένα πολυδιάστατο μοντέλο (κύβος OLAP), θα τεθούν στοχευμένα ερωτήματα και αναφορές στα δεδομένα και τέλος θα αξιολογηθεί το αποτέλεσμα και θα γίνουν προτάσεις προς την επιχείρηση.

## **2. Ανασκόπηση Βιβλιογραφίας**

### **2.1. Επιχειρηματική Ευφυΐα**

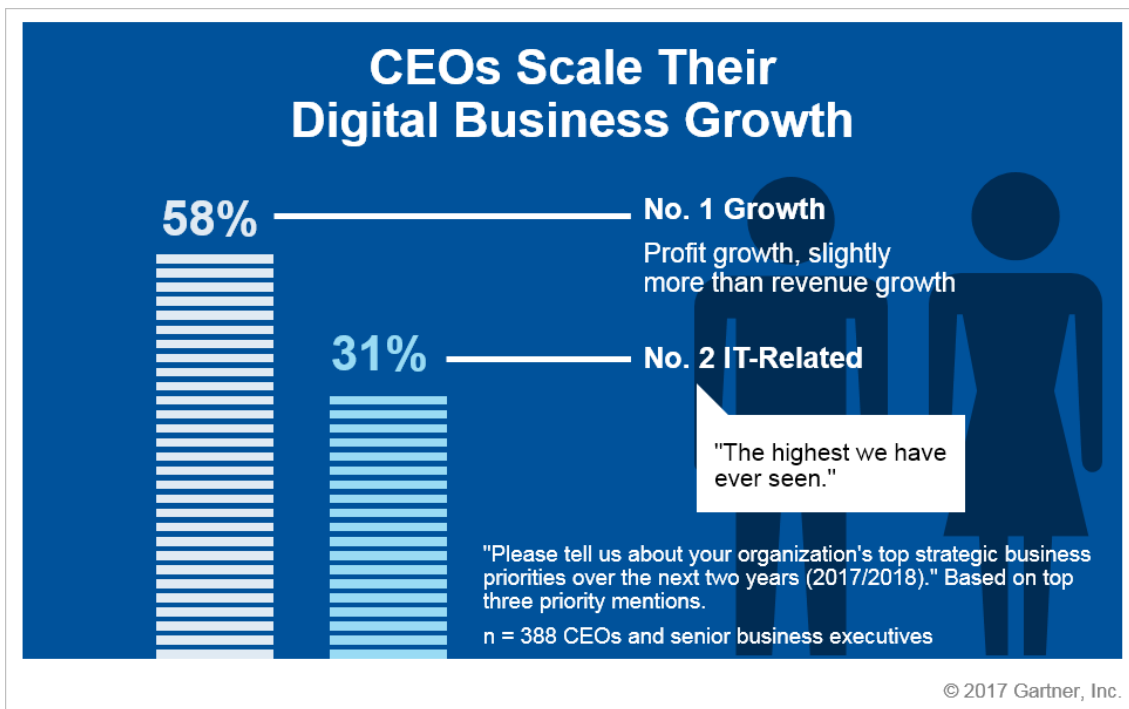
#### **2.1.1. Ορισμός**

Ο όρος επιχειρηματική ευφυΐα (Business Intelligence) αναφέρεται σε τεχνολογίες, εφαρμογές και πρακτικές για τη συλλογή, ενσωμάτωση, ανάλυση και παρουσίαση επιχειρηματικών πληροφοριών. Ο σκοπός της επιχειρηματικής ευφυΐας είναι να υποστηρίξει τη διαδικασία λήψης επιχειρηματικών αποφάσεων. Ουσιαστικά, τα συστήματα επιχειρηματικής ευφυΐας είναι συστήματα υποστήριξης αποφάσεων (Decision Support Systems) οδηγούμενα από δεδομένα (data-driven). Η επιχειρηματική ευφυΐα χρησιμοποιείται μερικές φορές εναλλακτικά των γραφικών συστημάτων λήψης επιχειρηματικών αποφάσεων (briefing books), των εργαλείων αναφορών και ερωτημάτων (report and query tools) και των πληροφοριακών συστημάτων ανώτατης διοίκησης (Executive Information Systems) (OLAP.com, 2018).

Ένας εναλλακτικός ορισμός αναφέρει ότι «η επιχειρηματική ευφυΐα ορίζεται ως ένα σύνολο μεθοδολογιών, διαδικασιών, αρχιτεκτονικών και τεχνολογιών που ενισχύει το προϊόν των διαδικασιών των πληροφοριακών συστημάτων διοίκησης, το οποίο χρησιμοποιείται σε αναλύσεις, αναφορές, στη διοίκηση απόδοσης και στην παροχή πληροφοριών. Το πεδίο εφαρμογής περιλαμβάνει τους πίνακες ανώτατης διοίκησης (executive dashboards) καθώς και εργαλεία αναφορών και ερωτημάτων» (Forrester, 2018).

#### **2.1.2. Η αξία της Επιχειρηματικής Ευφυΐας**

Η σύγχρονη επιχείρηση οφείλει να επικεντρωθεί στον ψηφιακό μετασχηματισμό προκειμένου να ανταπεξέλθει στις προκύπτουσες ανάγκες. Στην Εικόνα 1 εμφανίζεται το αποτέλεσμα σχετικής έρευνας του οίκου Gartner (Raskino, 2017), η οποία αναδεικνύει τη σημαντικότητα που προσδίδουν CEO's μεγάλων επιχειρήσεων ανά τον κόσμο σε αυτόν τον τομέα.



Εικόνα 1: Οι προτεραιότητες των CEO's μεγάλων επιχειρήσεων (Πηγή: Gartner 2017)

Η εκμετάλλευση της επιχειρηματικής ευφυΐας ως πολύτιμου πόρου για την επιχείρηση καθίσταται ως ένας από τους σημαντικότερους στόχους των διοικήσεων. Σύμφωνα με την πρόσφατη βιβλιογραφία (Κύρκος, 2015), η επιχειρηματική ευφυΐα κρίνεται απαραίτητη για ποικίλους λόγους. Αρχικά, βοηθάει στη λήψη επιχειρηματικών αποφάσεων μέσα στις συνθήκες αβεβαιότητας της σύγχρονης εποχής, κυρίως σε στρατηγικό επίπεδο όπου η αβεβαιότητα και η διασπορά των δεδομένων καθιστά αυτές τις αποφάσεις ιδιαίτερα περίπλοκες. Εδώ έρχεται να προστεθεί το σύγχρονο, απαιτητικό και κυρίως ασταθές επιχειρηματικό περιβάλλον, αυξάνοντας ακόμη περισσότερο το βαθμό πολυπλοκότητας και δυσκολίας. Ενδεικτικά στοιχεία αυτού του περιβάλλοντος είναι η μεγάλη αύξηση του ανταγωνισμού, η έντονη χρονική πίεση λόγω της εντατικοποίησης της λειτουργίας, η διεθνοποίηση των επιχειρήσεων, η πληθώρα των υπάρχουσών επιλογών και ο τεράστιος όγκος πληροφοριών που είναι πλέον διαθέσιμος σε όλους. Προκειμένου λοιπόν να ανταποκριθεί η διοίκηση της επιχείρησης στην ιδιαίτερα βεβαρυσμένη διαδικασία λήψης των σωστών αποφάσεων, ώστε να αποκτήσει συγκριτικό πλεονέκτημα έναντι των ανταγωνιστών της, πρέπει είναι σε θέση να λαμβάνει την κατάλληλη πληροφόρηση στη σωστή χρονική στιγμή. Αυτόν το ρόλο λαμβάνουν τα συστήματα επιχειρηματικής ευφυΐας προσφέροντας πολύτιμη πληροφορία και ελαχιστοποιώντας το επίπεδο της αβεβαιότητας στη λήψη αποφάσεων (Ferrari, 2011).

Επιπλέον, ενισχύει την επιχείρηση στο μέτωπο της αντιμετώπισης των προκλήσεων που προκύπτουν από το φαινόμενο της παγκοσμιοποίησης. Οι επιχειρήσεις δραστηριοποιούνται πλέον σε ένα περιβάλλον που αλλάζει συνεχώς με ραγδαίους ρυθμούς και απαιτεί συνεχή προσπάθεια για μείωση του κόστους και επίτευξη οικονομικών κλίμακας. Επίσης, καλούνται να αντιμετωπίσουν ανταγωνισμό σε διεθνή κλίμακα λόγω των ελαχιστοποιήσεων των περιοριστικών μέτρων και την ελάφρυνση των κανονισμών του διεθνούς εμπορίου. Πολλές φορές δραστηριοποιούνται σε μεγάλες γεωγραφικές εκτάσεις, σε διαφορετικές χώρες, όπου πρέπει να διαχειριστούν και διαφορετικές κουλτούρες, θρησκείες και εργασιακές τακτικές, κυρίως με την είσοδο των υπό ανάπτυξη χωρών που πλέον αντιμετωπίζονται ως αναπτυγμένες λόγω συγκρίσιμων οικονομικών μεγεθών. Τέλος, η αλματώδης εξέλιξη της τεχνολογίας έχει δημιουργήσει νέα κανάλια επικοινωνίας και πιο απαιτητικούς καταναλωτές. Γίνεται λοιπόν αντιληπτό ότι η σύνθεση του σύγχρονου περιβάλλοντος, στο οποίο καλείται να δραστηριοποιηθεί η επιχείρηση, είναι πολύπλοκη και συχνά αβέβαιη, οπότε είναι απαραίτητη η βελτίωση των πρακτικών λήψης αποφάσεων μέσω της άντλησης των ορθών και έγκαιρων πληροφοριών.

Ένας ακόμη λόγος που προσδίδει σημαντική αξία στη χρήση της επιχειρηματικής ευφυΐας είναι η αξιοποίηση της πληθώρας των διαθέσιμων δεδομένων. Πλέον, οι επιχειρήσεις έχουν πρόσβαση σε ένα τεράστιο όγκο δεδομένων, που προέρχονται είτε από το εσωτερικό τους περιβάλλον μέσω των πληροφοριακών συστημάτων τους (ERP, CRM, SRM), της ιστοσελίδας τους κ.α., είτε από το εξωτερικό περιβάλλον μέσω του κρατικού μηχανισμού, των μέσων κοινωνικής δικτύωσης και της εν γένει ελεύθερης πληροφορίας που υφίσταται σήμερα στο διαδίκτυο. Τα συστήματα επιχειρηματικής ευφυΐας έχουν τη δυνατότητα να συγκεντρώσουν, να ομογενοποιήσουν και να επεξεργαστούν αυτά τα δεδομένα προκειμένου να παραγάγουν πολύτιμη πληροφορία για τις επιχειρήσεις.

Ένας τελευταίος αλλά πολύ σημαντικός παράγοντας, που προσέδωσε αξία στην επιχειρηματική ευφυΐα και την ανέδειξε ως έναν από τους πολυτιμότερους πόρους μιας επιχείρησης, είναι το γεγονός ότι με την αλματώδη ανάπτυξη των νέων τεχνολογιών προέκυψαν και νέες μέθοδοι επεξεργασίας των δεδομένων, οι οποίες προσφέρουν άμεση, στοχευμένη και πολύτιμη πληροφόρηση. Το σύνολο των προαναφερθέντων

παραγόντων σε συνδυασμό με την ανάγκη αφομοίωσης και αποδοχής των νέων μεθόδων είναι αυτό που έχει καταστήσει την επιχειρηματική ευφυΐα ως έναν από τους σημαντικότερους τομείς που οφείλει να εκμεταλλευτεί μια επιχείρηση.

### **2.1.3. Οι περιορισμοί της Επιχειρηματικής Ευφυΐας**

Το μεγαλύτερο ίσως μειονέκτημα της επιχειρηματικής ευφυΐας είναι το κόστος απόκτησης, συντήρησης και λειτουργίας του συστήματος που την υποστηρίζει. Αρχικά μια επιχείρηση πρέπει να εξετάσει το κόστος απόκτησης των λογισμικών που θα χρησιμοποιηθούν στο σύστημα και βάσει αυτών να προμηθευτεί και τον κατάλληλο υλικοτεχνικό εξοπλισμό. Στις περισσότερες περιπτώσεις απαιτείται και παραμετροποίηση του λογισμικού στις ανάγκες και τα δεδομένα της επιχείρησης, διαδικασία που είναι ιδιαίτερα δαπανηρή. Τέλος, δεδομένης της πολυπλοκότητας των συστημάτων αυτών, θα πρέπει να προβεί και στην εκπαίδευση του προσωπικού της για τη χρήση του συστήματος ή ακόμα και να αναγκαστεί να προσλάβει νέο και εξειδικευμένο προσωπικό, κόστος το οποίο έρχεται να προστεθεί στα προαναφερθέντα.

Όπως αναφέρθηκε και προηγουμένως, τα συστήματα επιχειρηματικής ευφυΐας είναι ιδιαίτερα πολύπλοκα. Αυτό συνεπάγεται πολλές φορές με δυσκολία και λάθη κατά το χειρισμό τους αλλά και μεγάλο απαιτούμενο διάστημα εξοικείωσης στη χρήση τους, γεγονότα που έχουν αρνητική επίδραση στην ποιότητα του αποτελέσματος που λαμβάνει η επιχείρηση από την επένδυση που έχει υλοποιήσει.

Μια ακόμη δυσκολία που πρέπει να αντιμετωπίσουν τα συστήματα επιχειρηματικής ευφυΐας είναι η μεγάλη εξάρτησή τους από την ποιότητα των δεδομένων που χρησιμοποιούν. Με γνώμονα ότι λαμβάνουν δεδομένα από διάφορες πηγές της επιχείρησης, σχεδόν πάντα ανομοιογενή και «μη καθαρά», είναι πολύ πιθανό κατά την εισαγωγή τους στο σύστημα να μη φιλτραριστούν με τον ενδεδειγμένο τρόπο και να καταλήξουν λανθασμένα στη βάση του συστήματος, με αποτέλεσμα οι όποιες πληροφορίες παραχθούν να είναι χαμηλής αξίας.

Ένας ακόμη παράγοντας που θα πρέπει να εξεταστεί είναι το γεγονός ότι λόγω της ανάγκης πολυεπίπεδης και ταυτόχρονης χρήσης των συστημάτων αυτών από πολλούς

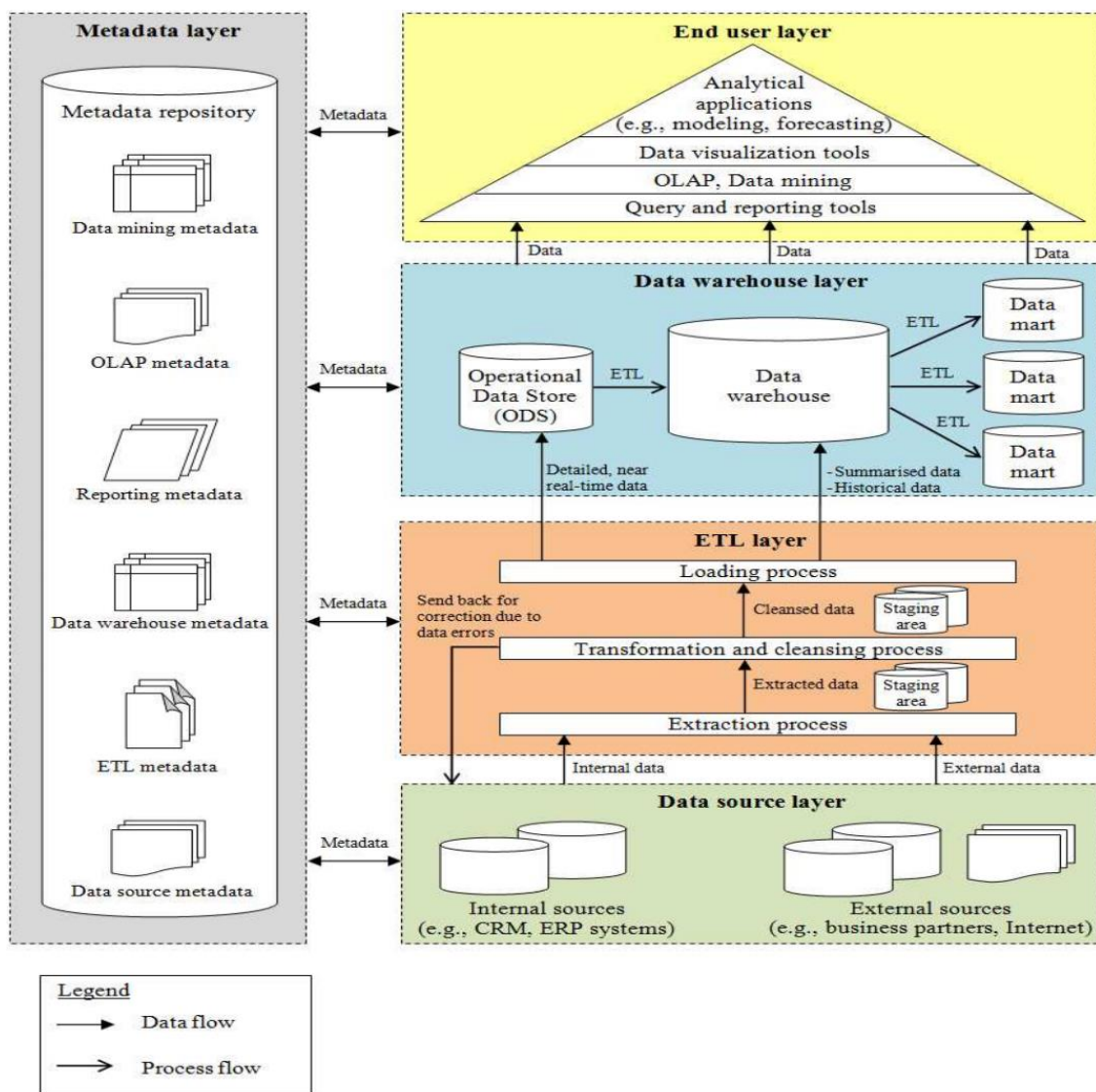
χρήστες, δημιουργούνται πολλές φορές προβλήματα με την ακεραιότητα των δεδομένων αλλά και την ασφάλεια και διαθεσιμότητα του ίδιου του συστήματος.

Τέλος, είναι πολύ σημαντικό να ληφθεί υπόψη ότι δεδομένου ότι ο χρήστης είναι αυτός που θα έχει την τελική ερμηνεία της πληροφορίας που θα παραχθεί από το σύστημα, πολλές φορές παρατηρείται το γεγονός ότι διαφορετικοί χρήστες καταλήγουν σε διαφορετικά συμπεράσματα. Γι' αυτό το λόγο, πρέπει πάντα η επιχείρηση να λαμβάνει υπόψιν της ότι δεν μπορεί να επιδεικνύει υπερβολική εμπιστοσύνη στο ίδιο το σύστημα και να στηρίζεται τυφλά στο όποιο αποτέλεσμα παράγεται και παρουσιάζεται από αυτό αλλά να αναγνωρίζει ότι ο ανθρώπινος παράγοντας είναι ο πολυτιμότερος πόρος ακόμη και στα συστήματα επιχειρηματικής ευφυΐας.

#### **2.1.4. Ανάλυση συστήματος Επιχειρηματικής Ευφυΐας**

Ένα σύστημα επιχειρηματικής ευφυΐας θα μπορούσε να αναλυθεί ως μια αρχιτεκτονική πέντε επιπέδων, όπως αυτά παρουσιάζονται και στην Εικόνα 2, στην οποία το κατώτερο επίπεδο είναι οι πηγές δεδομένων (Data Sources), ακολουθεί η διεργασία εξαγωγής, μετασχηματισμού και φόρτωσης των δεδομένων (ETL - Extract/Transform/Load), μετά συναντάται το επίπεδο των αποθηκών δεδομένων (Data Warehouses) και καταλήγει στο επίπεδο του τελικού χρήστη (End User). Παράλληλα με τα τέσσερα αυτά επίπεδα βρίσκεται το πέμπτο επίπεδο που αφορά τα μεταδεδομένα (Metadata) (Lih Ong, Hwa Siew, & Wong, 2011).





Εικόνα 2: Τα πέντε επίπεδα ενός συστήματος επιχειρηματικής ευφυΐας (Lih Ong, Hwa Siew, & Wong, 2011)

### 2.1.4.1. Επίπεδο Πηγών Δεδομένων

Οι πηγές δεδομένων είναι το κατώτερο επίπεδο ενός συστήματος επιχειρηματικής ευφυΐας και το τροφοδοτούν με τα ακατέργαστα δεδομένα που εν συνεχεία θα υποστούν επεξεργασία για να παραγάγουν την επιθυμητή πληροφορία.

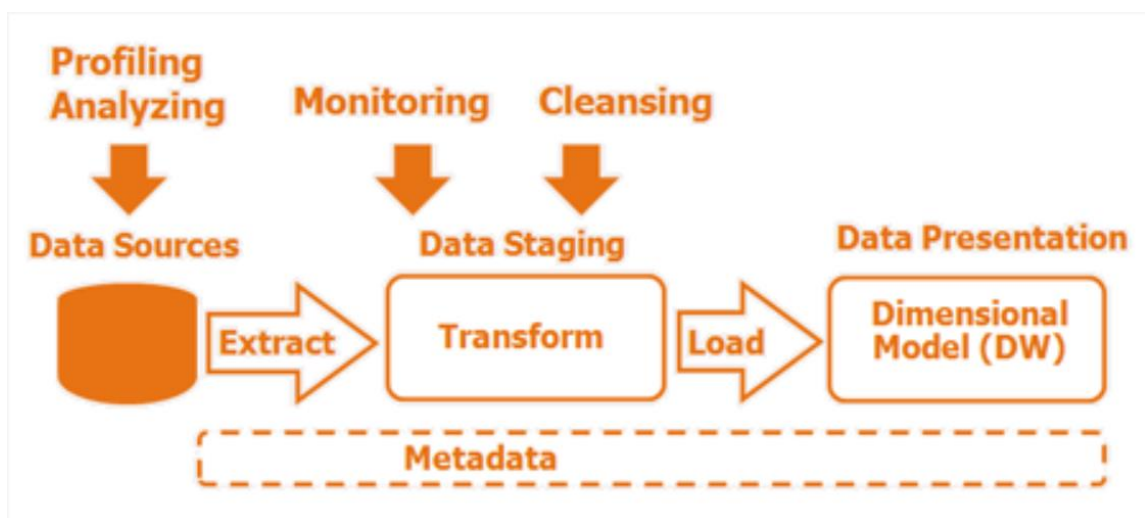
Τα δεδομένα αυτά σε μια επιχείρηση συγκεντρώνονται, κυρίως αλλά όχι αποκλειστικά, από τα συστήματα επεξεργασίας συναλλαγών (OLTP systems – Online Transaction Processing) και αφορούν σε όλες τις λεπτομερείς εγγραφές κάθε στοιχειώδους πράξης που εκτελούν αυτά τα συστήματα (εσωτερικές πηγές). Στη μορφή τους αυτή τα δεδομένα είναι απαραίτητα για την καθημερινή λειτουργία της επιχείρησης αλλά

ακατάλληλα για χρήση σε στρατηγικό επίπεδο και λήψη αποφάσεων. Χαρακτηριστικά αυτών των δεδομένων είναι η μεγάλη διασπορά σε διακριτά συστήματα της επιχείρησης, ο υπερβολικά μεγάλος όγκος, η μεταξύ τους ανομοιογένεια και πολλές φορές οι ελλείψεις.

Επιπροσθέτως, ανάλογα δεδομένα μπορούν να συγκεντρωθούν και από το εξωτερικό περιβάλλον της επιχείρησης, όπως οι συνεργάτες, το πολιτικό περιβάλλον, το διαδίκτυο κ.α. και αφορούν κυρίως στοιχεία του ανταγωνισμού, της κατάστασης της αγοράς, καθώς και της πορείας της οικονομίας.

### 2.1.4.2. Επίπεδο διεργασιών ETL

Προκειμένου τα δεδομένα να εξαχθούν από τις πηγές, να μεταφερθούν στην αποθήκη δεδομένων, να «καθαριστούν» από λάθη και κενά, να μετασχηματιστούν ώστε να είναι ομοιογενή και να φορτωθούν εν τέλει σε ένα κοινό σχήμα, εκτελούνται κάποιες διεργασίες οι οποίες ονομάζονται ETL, ακρωνύμιο για το Extract, Transform & Load (εξαγωγή, μετασχηματισμός και φόρτωση) (Vassiliadis, 2009). Οι σύγχρονες αρχιτεκτονικές των συστημάτων αυτών προσθέτουν και κάποιες επιπλέον έννοιες στις πραγματοποιούμενες διεργασίες αυτού του επιπέδου. Όπως μπορούμε να δούμε και στην Εικόνα 3, αυτές είναι η δημιουργία προφίλ (Profiling), η ανάλυση (Analyzing), η παρακολούθηση (Monitoring) και ο καθαρισμός (Cleansing). Παρακάτω θα αναλυθούν με τη σειρά που πραγματοποιούνται όλες αυτές οι διεργασίες.



Εικόνα 3: Οι διεργασίες ETL πλαισιωμένες από τις διεργασίες Profiling, Analyzing, Monitoring και Cleansing (element61 N.V. – Moore Stephens, 2018)

## **Προφίλ και Ανάλυση**

Σκοπός αυτών των διεργασιών είναι να βοηθήσουν στην κατανόηση της πηγής δεδομένων. Κατά τη δημιουργία του προφίλ της πηγής χρησιμοποιούνται τεχνικές επικύρωσης μοτίβων και μορφών δεδομένων αλλά και εντοπισμού πλεοναζόντων δεδομένων για να διερευνηθεί το περιεχόμενο, η δομή και η ποιότητα των δεδομένων. Στη συνέχεια η ανάλυση θα χρησιμοποιήσει αυτό το αποτέλεσμα για να εντοπίσει ποιοτικά προβλήματα, όπως ελλιπή, ασυνεπή ή λάθος δεδομένα. Η ορθότητα και η ποιότητα της ανάλυσης αυτής είναι πολύ σημαντική δεδομένου ότι θα βοηθήσει τη διεργασία του μετασχηματισμού που θα ακολουθήσει.

## **Εξαγωγή**

Η διεργασία της εξαγωγής έχει ως στόχο να εντοπίσει και να αντλήσει το ζητούμενο υποσύνολο δεδομένων από την πηγή. Σημαντικά σημεία σε αυτή τη διεργασία είναι:

- η προσπάθεια για την ελάχιστη δυνατή επιβάρυνση του συστήματος που φιλοξενεί την πηγή δεδομένων
- η ταχύτητα στην άντληση αυτών των δεδομένων
- η ελάχιστη (έως μηδενική) παρέμβαση στο λογισμικό αυτού του συστήματος
- το γεγονός ότι κάθε φορά πρέπει να επιλεγθεί η πιο κατάλληλη μέθοδος ανάλογα με το συνεργαζόμενο σύστημα
- η ανάγκη για συμπίεση και κρυπτογράφηση των δεδομένων κατά την εξαγωγή τους για λόγους απόδοσης και ασφάλειας

## **Μετασχηματισμός και Καθαρισμός**

Στη διεργασία του μετασχηματισμού εφαρμόζονται ορισμένοι κανόνες στα δεδομένα που έχουν προέλθει από τη διεργασία της εξαγωγής προκειμένου να χρησιμοποιηθούν στη διεργασία της φόρτωσης. Ένας πολύ σημαντικός παράγοντας του σταδίου αυτού είναι ο καθαρισμός των δεδομένων αυτών, ο οποίος στοχεύει στο να περάσουν μόνο τα κατάλληλα δεδομένα στην επόμενη διεργασία. Το φάσμα των κανόνων που εφαρμόζονται είναι ιδιαίτερα ευρύ και περιλαμβάνει, μεταξύ άλλων, κανόνες επιλογής συγκεκριμένων τμημάτων της βάσης προέλευσης, τεχνικές τακτοποίησης, συσχέτισης και συνάθροισης δεδομένων, τη μετάφραση, κωδικοποίηση και δημιουργία νέων τιμών, κανόνες επικύρωσης δεδομένων κ.α.

## **Φόρτωση**

Η φόρτωση αποτελεί το τελευταίο στάδιο της συνολικής διαδικασίας ETL και στην ουσία φορτώνει τα δεδομένα στην αποθήκη δεδομένων. Αυτή η διεργασία έχει πολλές εκφάνσεις, οι οποίες εξαρτώνται από επιχειρηματικά κριτήρια, που απορρέουν από τις στρατηγικές αποφάσεις της κάθε επιχείρησης. Σε ορισμένες περιπτώσεις εφαρμόζεται ανανέωση των δεδομένων σε τακτά χρονικά διαστήματα, προσθήκη ιστορικών δεδομένων ή αντικατάσταση υπάρχουσας πληροφορίας με σωρευτική. Σε αυτό το σημείο ενεργοποιούνται και οι περιορισμοί που προέρχονται από το σχήμα της αποθήκης δεδομένων ή και από την ίδια τη διεργασία και οι οποίοι είναι πολύ σημαντικοί για την ποιοτική απόδοση του όλου συστήματος.

## **Παρακολούθηση**

Η παρακολούθηση είναι μια διεργασία που συμμετέχει σχεδόν σε όλο το σύστημα ETL και στην ουσία παρακολουθεί και επικυρώνει τα δεδομένα σε κάθε στάδιο. Ο έλεγχος των δεδομένων πρέπει να πραγματοποιείται σε τέτοιο βαθμό που να μην επηρεάζει τη συνολική απόδοση του συστήματος, η οποία με τη σειρά της καταγράφεται σε αυτό το στάδιο μέσω άντλησης μεταδεδομένων που αφορούν στη λειτουργία του συστήματος και τις πραγματοποιηθείσες εγγραφές του. Τα μεταδεδομένα αυτά θα βοηθήσουν στη γενικότερη επισκόπηση της απόδοσης της λειτουργίας του συστήματος καθώς και στην πληρότητα των δεδομένων που εν τέλει εισήχθησαν στην αποθήκη δεδομένων.

### **2.1.4.3. Επίπεδο Αποθήκης Δεδομένων**

Μετά το στάδιο των διεργασιών ETL ακολουθεί η αποθήκευση των καθαρών και ομογενοποιημένων δεδομένων στην αποθήκη δεδομένων. Το επίπεδο αυτό αποτελεί «μια ενιαία, ολοκληρωμένη και σταθερή αποθήκη δεδομένων, τα οποία εξάγονται από πολλαπλές πηγές και καθίστανται διαθέσιμα στους τελικούς χρήστες, με τρόπο κατανοητό προς αυτούς ώστε να χρησιμοποιηθούν για επιχειρησιακή ανάλυση» (Kour, 2015). Συναντώνται διάφοροι τρόποι υλοποίησης αυτού του επιπέδου. Η δομή που θα εξετασθεί περιλαμβάνει την αποθήκη λειτουργικών δεδομένων (ODS - Operational Data Store), την αποθήκη δεδομένων (Data Warehouse) και τα υποσύνολα της αποθήκης δεδομένων (Data Marts).

## **Operational Data Store**

Αποτελεί ένα υποεπίπεδο ενοποίησης μεταξύ των διεργασιών ETL και της αποθήκης δεδομένων. Εδώ ενσωματώνονται και αποθηκεύονται όλα τα δεδομένα που προέρχονται από το στάδιο των διεργασιών ETL και στη συνέχεια φορτώνονται στην αποθήκη δεδομένων. Τα χαρακτηριστικά των αποθηκών λειτουργικών δεδομένων συνοψίζονται στα εξής (Imhoff, Galemno, & Geiger, 2003):

- Είναι θεματικά προσανατολισμένες, δηλαδή οργανώνονται γύρω από συγκεκριμένα στοιχεία, όπως προϊόντα και πωλήσεις. Δεν συμπεριλαμβάνουν δεδομένα που δεν συνεισφέρουν στη διαδικασία λήψης αποφάσεων.
- Περιλαμβάνουν δεδομένα σχεδόν πραγματικού χρόνου.
- Είναι ευμετάβλητες υπό την έννοια ότι δέχονται συνεχώς νέα δεδομένα, τα οποία αντικαθιστούν τα παλαιότερα. Ως εκ τούτου, δεν μπορούν να διατηρήσουν ιστορικά δεδομένα.
- Υποστηρίζουν διαδικασίες του λειτουργικού επιπέδου μιας επιχείρησης.

## **Data Warehouse**

Στην αποθήκη δεδομένων συγκεντρώνονται και αποθηκεύονται όλα τα δεδομένα που έχουν προέλθει από την αλυσίδα της επεξεργασίας που περιγράφηκε και χρησιμοποιούνται για ανάλυση και λήψη αποφάσεων στο στρατηγικό επίπεδο μιας επιχείρησης. Εδώ αποθηκεύεται πολύ μεγάλος όγκος συγκεντρωτικών δεδομένων, χρησιμοποιώντας μια πολυδιάστατη δομή. Τα δεδομένα επικαιροποιούνται σε τακτά χρονικά διαστήματα και διατηρούνται για πολύ μεγάλο χρονικό διάστημα (έως και δέκα χρόνια), ενώ τηρείται και μεγάλος όγκος ιστορικών δεδομένων. Στην παράγραφο 2.2 θα εξετασθεί αναλυτικότερα η αποθήκη δεδομένων διότι αποτελεί το σημαντικότερο ίσως στοιχείο του συστήματος της επιχειρηματικής ευφυΐας.

## **Data Marts**

Προς εξυπηρέτηση συγκεκριμένων τμημάτων μιας επιχείρησης ή διακριτών λειτουργιών της, δημιουργούνται υποσύνολα της αποθήκης δεδομένων (Data Marts), τα οποία είναι προσαρμοσμένα στις ανάγκες του σκοπού που εξυπηρετούν. Τα υποσύνολα αυτά μπορεί να έχουν πολλά διαφορετικά χαρακτηριστικά ανάλογα με το σκοπό για τον οποίο δημιουργήθηκαν αλλά έχουν και ορισμένα κοινά, όπως το ότι πάντα περιλαμβάνουν ένα υποσύνολο των δεδομένων της αποθήκης δεδομένων, ότι μπορεί να

φιλοξενούνται στο ίδιο φυσικό μέσο με την αποθήκη δεδομένων ή σε ξεχωριστό και ότι χαρακτηρίζονται από μεγάλο εύρος στον όγκο των δεδομένων.

#### 2.1.4.4. Επίπεδο Τελικού Χρήστη

Σε αυτό το στάδιο συναντώνται όλα τα εργαλεία και οι μέθοδοι που χρησιμοποιούνται για την εξαγωγή και οπτικοποίηση της πληροφορίας. Ανάλογα με το επίπεδο της διοίκησης που αυτά αναφέρονται (άρα και το σκοπό χρήσης τους), μεταβάλλεται και το είδος τους. Αρχικά, συναντώνται τα **εργαλεία ερωτημάτων και αναφορών** (query and reporting tools), τα οποία προσφέρουν ταχεία πρόσβαση του χρήστη στα δεδομένα και τη δημιουργία αναφορών για την υποβοήθηση της λήψης αποφάσεων σε λειτουργικό επίπεδο. Στη συνέχεια ακολουθούν τα **εργαλεία OLAP** (Online Analytical Processing tools), τα οποία χρησιμοποιούνται για την πολυδιάστατη μοντελοποίηση και ανάλυση δεδομένων. Στην παράγραφο 2.3 θα αναπτυχθεί εκτενώς η θεματολογία γύρω από τα εργαλεία OLAP. Συνδυάζοντας τις λειτουργίες των εργαλείων OLAP και των αποθηκών δεδομένων, συναντώνται στο επόμενο σκαλοπάτι τα **εργαλεία εξόρυξης δεδομένων** (Data Mining tools), τα οποία, μέσω στατιστικών τεχνικών, προσφέρουν τον αυτόματο προσδιορισμό πολύτιμης πληροφορίας, όπως τάσεις, κρυφές σχέσεις και ασυνήθιστα μοτίβα που είναι κρυμμένα σε τεράστιους όγκους δεδομένων. Τα **εργαλεία οπτικοποίησης των δεδομένων** (Data Visualization tools), τα οποία αφορούν στο επόμενο επίπεδο, είναι εργαλεία που χρησιμοποιούνται από την ανώτατη διοίκηση για καθημερινή, σύντομη και συνοπτική, οπτικοποιημένη πληροφόρηση σχετικά με την απόδοση και την πορεία της επιχείρησης. Τα εργαλεία αυτά συνήθως αφορούν σε ψηφιακά ταμπλό (dashboards), τα οποία περιλαμβάνουν πίνακες, διαγράμματα κ.α. Τέλος, στο ανώτατο επίπεδο υφίστανται οι **εφαρμογές ανάλυσης** (Analytical Applications), οι οποίες προσφέρουν στη διοίκηση της επιχείρησης τη δυνατότητα να κατανοήσει σε βάθος και να βελτιώσει την απόδοση των επιχειρηματικών λειτουργιών της μέσω τεχνικών μοντελοποίησης, προβλέψεων κ.α.

#### 2.1.4.5. Επίπεδο Μεταδεδομένων

Τα μεταδεδομένα (Metadata) αποτελούν πληροφορίες που συλλέγονται για τα ίδια τα δεδομένα, βάσει των οποίων περιγράφονται και ορίζονται η δομή και όλες οι

λειτουργίες και τα περιεχόμενα του συστήματος επιχειρηματικής ευφυΐας (Kimball & Ross, 2013). Χρησιμοποιούνται τρεις τύποι μεταδεδομένων:

Τα **τεχνικά μεταδεδομένα** που περιλαμβάνουν πληροφορίες για τη φυσική δομή του συστήματος, καθώς και για τις λειτουργίες που μεταφέρουν και μετασχηματίζουν δεδομένα μέσα σε αυτό.

Τα **επιχειρηματικά μεταδεδομένα** που περιγράφουν τη δομή και τα στοιχεία των δεδομένων, όπως επίσης και την επιχειρηματική λογική και λειτουργία της επιχείρησης.

Τα **μεταδεδομένα διαδικασιών** που αποδίδουν πληροφορίες, όπως μέτρηση απόδοσης, επίπεδο ποιότητας δεδομένων, έλεγχος κ.α., για όλες τις λειτουργίες του συστήματος.

## 2.2. Οι Αποθήκες Δεδομένων αναλυτικότερα

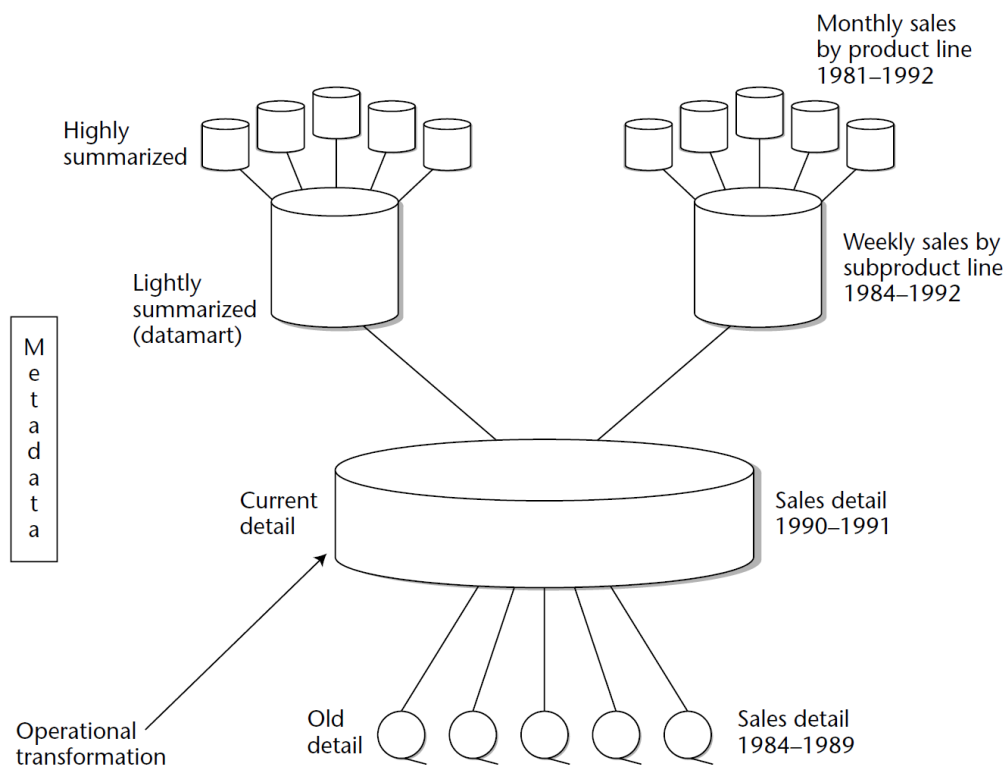
Η αποθήκη δεδομένων είναι «μια θεματικά προσανατολισμένη, ολοκληρωμένη, χρονικά διαφοροποιούμενη και μη ευμετάβλητη συλλογή δεδομένων, που χρησιμοποιείται για την υποστήριξη της ανάλυσης και της λήψης αποφάσεων σε στρατηγικό επίπεδο» (Inmon, 2005). Αναλυτικότερα, τα τέσσερα αυτά χαρακτηριστικά έχουν ως εξής (Κύρκος, 2015):

- **Θεματικά προσανατολισμένη (Subject-oriented):** Τα δεδομένα που συγκεντρώνονται από διάφορες πηγές, οργανώνονται και ομαδοποιούνται βάσει των εννοιών στις οποίες θέλει να επικεντρωθεί μια επιχείρηση, όπως οι πελάτες, οι πωλήσεις, τα προϊόντα κ.α.
- **Ολοκληρωμένη (Integrated):** Όπως έχει ήδη αναφερθεί, στην αποθήκη δεδομένων συγκεντρώνονται δεδομένα από διάφορες πηγές. Τα δεδομένα αυτά είναι ανομοιογενή, με διαφορετική μορφή, ενώ συνήθως έχουν και προβλήματα. Προκειμένου να μεταφερθούν και να αποθηκευτούν στην αποθήκη δεδομένων, απαλλάσσονται από τα προβλήματα και ομογενοποιούνται.
- **Χρονικά διαφοροποιούμενη (Time-variant):** Τα δεδομένα που αποθηκεύονται στην αποθήκη δεδομένων έχουν πάντα τη διάσταση του χρόνου. Με αυτόν τον τρόπο καταγράφεται η ιστορικότητα των δεδομένων και μπορούν να εντοπισθούν αλλαγές και τάσεις.
- **Μη ευμετάβλητη (Non-volatile):** Τα νέα δεδομένα που αποθηκεύονται τακτικά στην αποθήκη δεδομένων δεν μεταβάλουν και δεν διαγράφουν τα ήδη υπάρχοντα, τα οποία παραμένουν ανεπηρέαστα.

### 2.2.1. Δομή Αποθήκης Δεδομένων

Η δομή της αποθήκης δεδομένων, η οποία παρουσιάζεται στη σχετική Εικόνα 4, χαρακτηρίζεται από πέντε στοιχεία. Τα στοιχεία αυτά είναι τα **σύγχρονα δεδομένα** (current detail), τα **παλαιά δεδομένα** (old detail), τα **ελαφρώς συγκεντρωτικά δεδομένα** (lightly summarized), τα **έντονα συγκεντρωτικά δεδομένα** (highly summarized) και τα **μεταδεδομένα** (metadata). Η αποθήκευση των στοιχείων αυτών συνήθως δεν πραγματοποιείται στο ίδιο φυσικό μέσο (Gray & Watson, 1998).





Εικόνα 4: Η δομή της Αποθήκης Δεδομένων (Inmon, 2005)

### Σύγχρονα δεδομένα

Αυτά είναι τα πιο πρόσφατα δεδομένα που αποθηκεύονται στην αποθήκη δεδομένων, αφού καθαριστούν και ομογενοποιηθούν. Αποτελούν αντίγραφα των δεδομένων συναλλαγών (συνήθως όχι πλήρη) και αφορούν στη χρονική στιγμή που αυτά μεταφέρθηκαν στην αποθήκη δεδομένων. Ο όγκος τους είναι πάρα πολύ μεγάλος, γιατί αποθηκεύονται κυρίως στο χαμηλότερο επίπεδο λεπτομέρειάς τους. Ο λόγος που μεταφέρονται στην αποθήκη δεδομένων είναι για να μην επιβαρύνεται η λειτουργία των συστημάτων συναλλαγών κάθε φορά που υποβάλλονται ερωτήματα στο σύστημα.

### Παλαιά δεδομένα

Όπως αναφέρθηκε ήδη, οι αποθήκες δεδομένων συνήθως τηρούν ιστορικότητα μέχρι και για δέκα χρόνια. Τα ιστορικά δεδομένα μεταφέρονται σε άλλο φυσικό μέσο από τα σύγχρονα δεδομένα, συνήθως χαμηλότερων δυνατοτήτων σε ταχύτητα, μετά από συγκεκριμένο χρονικό διάστημα, που ορίζεται από τους κανόνες του συστήματος, για το λόγο ότι χρησιμοποιούνται σε λιγότερα ερωτήματα και αναφορές, οπότε δεν υπάρχει και η ανάγκη για τη μέγιστη δυνατή ταχύτητα στην ανάκτησή τους.

### **Ελαφρώς συγκεντρωτικά δεδομένα**

Η σύνοψη των στοιχείων βοηθάει στην ταχεία ανταπόκριση και βελτιώνει τη γενικότερη χρήση του συστήματος. Όπως περιγράφηκε και στην παράγραφο 2.1.4.3, αυτό επιτυγχάνεται με τη χρήση υποσυνόλων της αποθήκης δεδομένων (Data Marts). Πέραν των χαρακτηριστικών που αναφέρθηκαν ήδη, εξετάζονται δύο σημαντικά στοιχεία που αφορούν στην επιλογή: i) των στοιχείων που θα συνοψισθούν και ii) της μονάδας του χρόνου που θα χρησιμοποιηθεί για τη σύνοψη. Αυτά τα δεδομένα είναι χρήσιμα για συγκεκριμένους σκοπούς τμημάτων της επιχείρησης (όπως για παράδειγμα το τμήμα πωλήσεων, το τμήμα μάρκετινγκ κ.α.).

### **Εντόνως συγκεντρωτικά δεδομένα**

Τα δεδομένα αυτά απευθύνονται στην ανώτατη διοίκηση μιας επιχείρησης. Εδώ παρατηρείται ακόμη υψηλότερο επίπεδο σύνοψης και πολύ μεγάλη ιστορικότητα. Τα δεδομένα είναι πολύ συνοπτικά και εύκολα προσβάσιμα, με τέτοιο τρόπο που βελτιώνουν κατά πολύ την ταχύτητα απόκρισης και βοηθούν στον εντοπισμό τάσεων και αδυναμιών της επιχείρησης.

### **Μεταδεδομένα**

Όπως παρουσιάστηκε και προηγουμένως, τα μεταδεδομένα αφορούν σε πληροφορίες για τα ίδια τα δεδομένα. Για την αποθήκη δεδομένων περιλαμβάνουν:

- Ένα ευρετήριο για το τι περιλαμβάνεται και που αυτό αποθηκεύεται.
- Χαρτογράφηση των δεδομένων για τη μετατροπή της μορφής των δεδομένων από αυτή των συστημάτων πηγής στην κατάλληλη για την αποθήκη δεδομένων.
- Τους κανόνες που χρησιμοποιούνται κατά τη σύνοψη των στοιχείων.

## **2.2.2. Σχήματα Αποθήκης Δεδομένων**

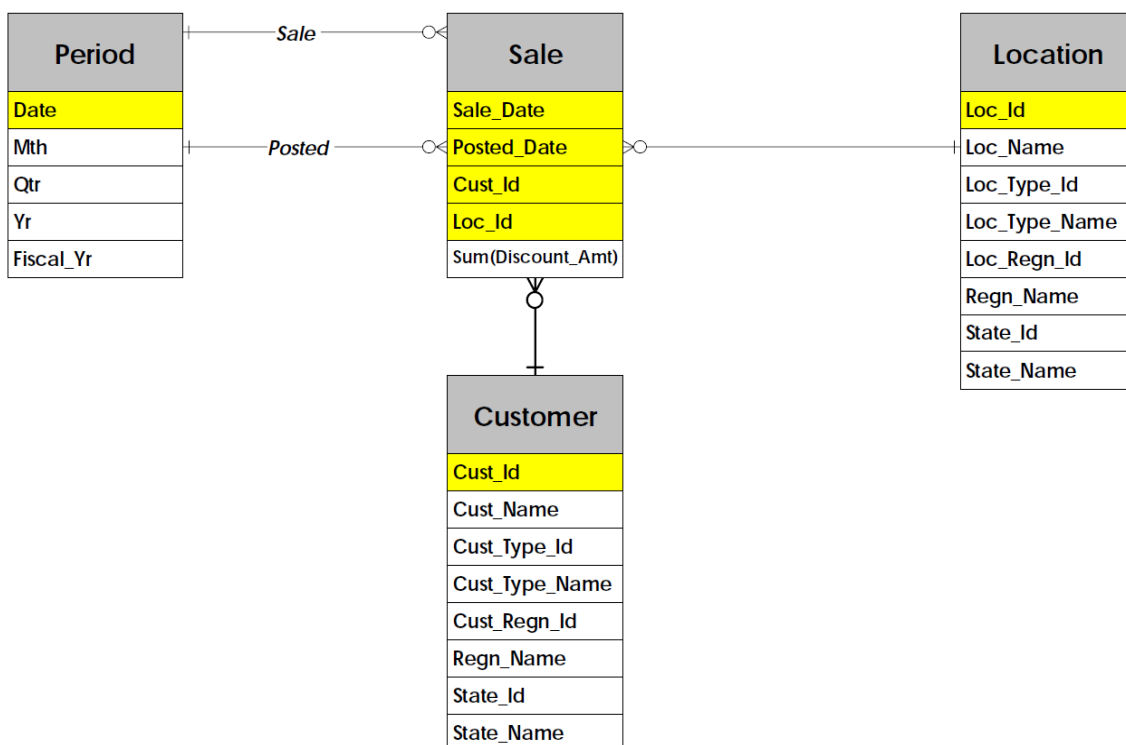
Για τις αποθήκες δεδομένων η αρχιτεκτονική των σχεσιακών πινάκων μετασχηματίζεται σε σχήματα που επικεντρώνονται σε αντικείμενα ενδιαφέροντος, ώστε να διευκολυνθεί η πολυδιάστατη ανάλυση. Αυτά είναι το **σχήμα αστέρα** (Star Schema), το **σχήμα αστερισμού** (Constellation Schema) και το **σχήμα χιονονιφάδας** (Snowflake Schema) (Theodoratos & Sellis, 1998).

## Σχήμα Αστέρα

Το σχήμα αστέρα απορρέει από το σχεσιακό μοντέλο οντοτήτων. Η δομή του περιλαμβάνει τα εξής:

- Πίνακας γεγονότων (Fact Table): Αποτελεί τον κεντρικό πίνακα και αναφέρεται στο αντικείμενο που πρέπει να αναλυθεί. Περιέχει τιμές-κλειδιά από τους πίνακες διαστάσεων και ένα ή περισσότερα μέτρα (measures).
- Πίνακες διαστάσεων (Dimension Tables): Αυτοί οι πίνακες αναφέρονται στις ιδιότητες των γεγονότων.

Στην Εικόνα 5 παρουσιάζεται ένα παράδειγμα σχήματος αστέρα που αφορά στις πωλήσεις, εξ ου και ο κεντρικός πίνακας γεγονότων. Περιμετρικά παρατηρούνται τρεις πίνακες διαστάσεων που αφορούν σε ιδιότητες του κεντρικού πίνακα: ημερομηνία, τόπος και πελάτης. Και οι τρεις διαστάσεις περιέχουν ενσωματωμένες ιεραρχίες. Οι τιμές του κεντρικού πίνακα αποτελούνται από τις τιμές-κλειδιά των πινάκων διαστάσεων εκτός από μια τιμή, αυτήν του αθροίσματος των πωλήσεων, που αποτελεί το μέτρο.

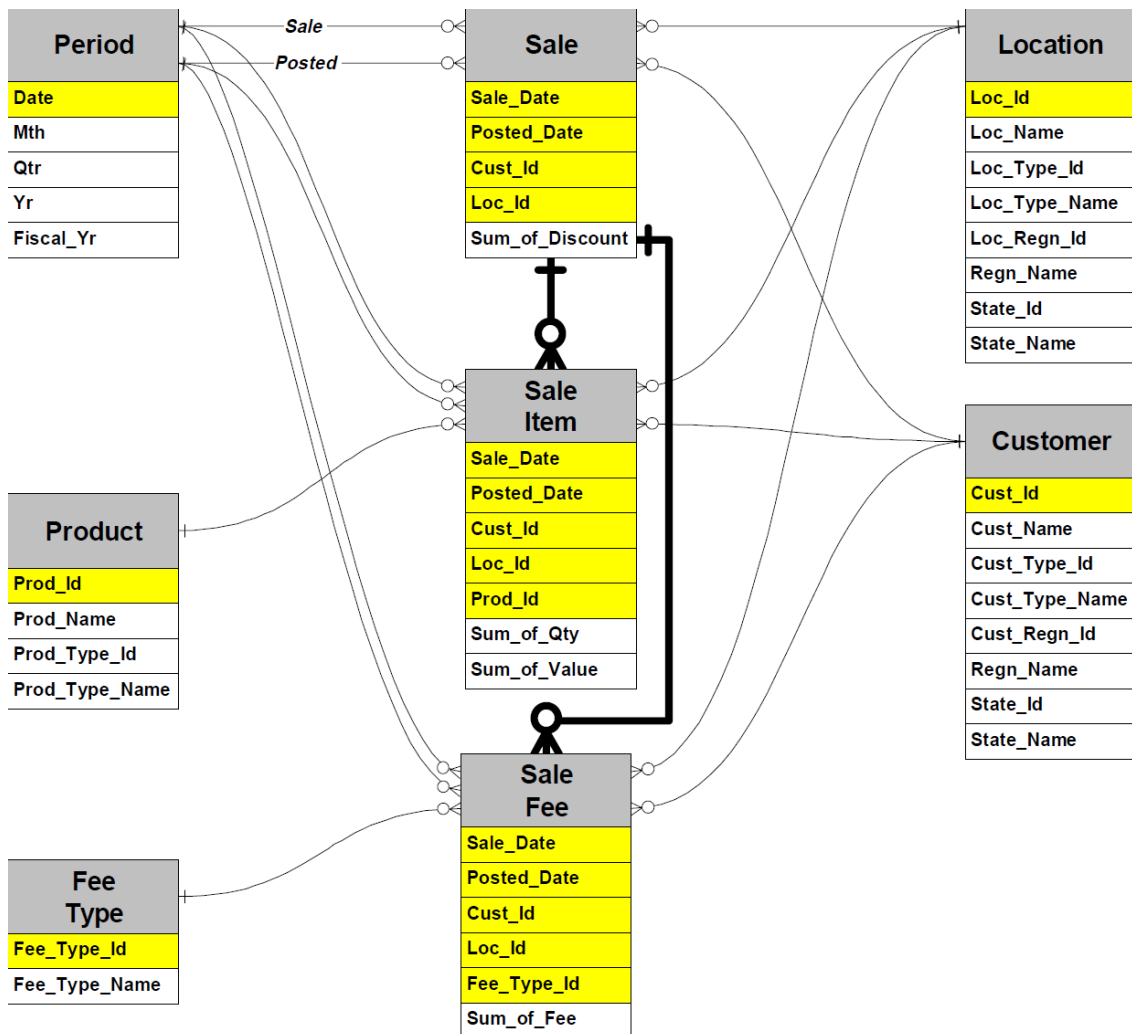


Εικόνα 5: Σχήμα Αστέρα (Moody & Kortink, 2000)

## Σχήμα Αστερισμού

Το σχήμα αστερισμού είναι μια παράλλαξη του σχήματος αστέρα, υπό την έννοια ότι αποτελείται από ένα σύνολο (δύο ή περισσότερα) σχημάτων αστέρα. Αυτό το σχήμα διασυνδέει τους πίνακες γεγονότων και δίνει τη δυνατότητα εξειδίκευσης στη λεπτομέρεια των περιγραφόμενων γεγονότων.

Στο παράδειγμα της Εικόνα 6 γίνεται αντιληπτό ότι προστέθηκαν δύο πίνακες γεγονότων στο προηγούμενο σχήμα αστέρα (Εικόνα 5), ένας με τα πωλούμενα αντικείμενα κι ένας με το κόστος της κάθε πώλησης. Προστέθηκαν επίσης οι πίνακες του προϊόντος και του τύπου κόστους που αποτελούν διαστάσεις για τους ανωτέρω πίνακες γεγονότων αντίστοιχα. Πρέπει να σημειωθεί ότι η σχέση του αρχικού πίνακα γεγονότων με τους δύο νέους είναι ένα-προς-πολλά.

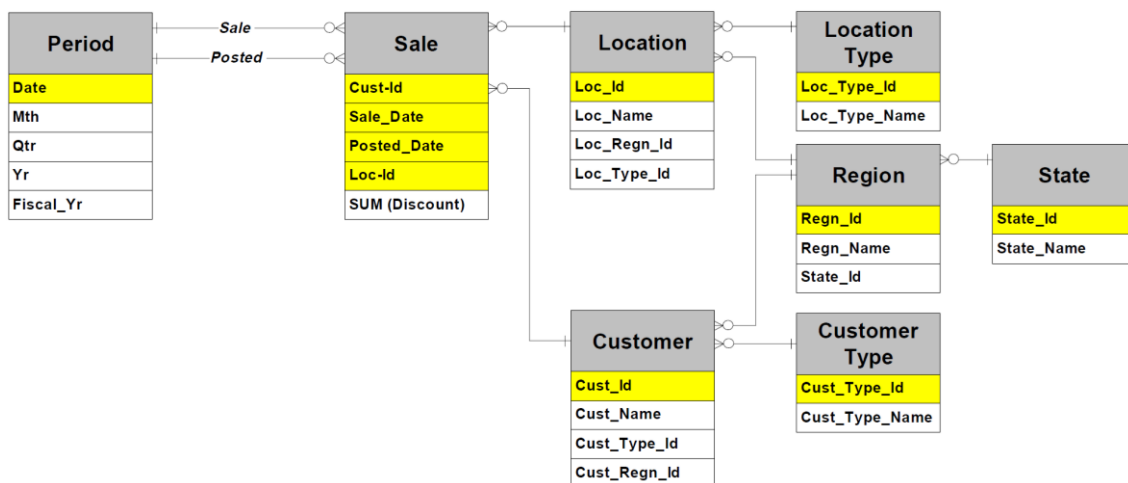


Εικόνα 6: Σχήμα Αστερισμού (Moody & Kortink, 2000)

## Σχήμα Χιονονιφάδας

Το σχήμα χιονονιφάδας διαφοροποιείται από το σχήμα αστέρα στο ότι οι πίνακες διαστάσεων μπορούν να αναλυθούν περισσότερο χρησιμοποιώντας βοηθητικούς πίνακες. Η λειτουργία αυτή προσομοιάζει στη λειτουργία της κανονικοποίησης στις σχεσιακές βάσεις.

Στην Εικόνα 7 παρουσιάζεται το προηγούμενο σχήμα αστέρα (Εικόνα 5), στο οποίο όμως έχουν προστεθεί βοηθητικοί πίνακες (τύπος τοποθεσίας, περιοχή, πολιτεία, είδος πελάτη) που συνδέονται με τις διαστάσεις του τόπου και του πελάτη. Με αυτόν τον τρόπο επιτυγχάνεται ελαχιστοποίηση των πλεοναζόντων δεδομένων.



Εικόνα 7: Σχήμα Χιονονιφάδας (Moody & Kortink, 2000)

## 2.3. OLAP

Η αναλυτική επεξεργασία άμεσης σύνδεσης (OLAP - Online Analytical Processing) είναι μια τεχνολογία που χρησιμοποιείται για την υποστήριξη της επιχειρηματικής ευφυΐας, ενώ αποτελεί ένα ισχυρό εργαλείο για την παραγωγή πολύτιμης πληροφορίας μέσω δυνατοτήτων όπως η παραγωγή μεγάλου αριθμού αναφορών, οι σύνθετοι υπολογισμοί και ο προγνωστικός σχεδιασμός, που είναι απαραίτητος σε στρατηγικό επίπεδο. Τα συστήματα OLAP σχετίζονται με τη διάταξη των δεδομένων σε ένα πολυδιάστατο και πολυεπίπεδο μοντέλο. Η έννοια OLAP θα μπορούσε να εκφραστεί και ως η θέαση του υλικού αλλά και του κοινωνικού κόσμου ως ένα σύνολο από πολλές διαστάσεις, κάθε μια από τις οποίες έχει πολλές ιεραρχίες, στάση η οποία επιτρέπει την καλύτερη κατανόησή του (Thomsen, 2002).

### 2.3.1. Χρησιμότητα της OLAP

Όπως έχει ήδη γίνει σαφές, η σύγχρονη επιχείρηση έχει ανάγκη από έγκαιρη και ορθή πληροφορία. Τα εργαλεία OLAP έχουν τη δυνατότητα να αναδεικνύουν αναντιστοιχίες και τάσεις στα δεδομένα, πληροφορίες που μέχρι πρότινος δεν ήταν αντιληπτές από την επιχείρηση, μέσω της σε βάθος αναζήτησης σε ενοποιημένα και συναθροισμένα δεδομένα, τα οποία προσφέρονται από τη δομή τους. Τα εργαλεία OLAP προσφέρουν μια ευρεία γκάμα εφαρμογών όπως η πρόβλεψη, η ανάλυση κεφαλαίων, η οργάνωση πόρων, οι προϋπολογισμοί, η διαχείριση κινδύνου κ.α. Συγκεκριμένα τα εργαλεία OLAP έχουν τις εξής δυνατότητες:

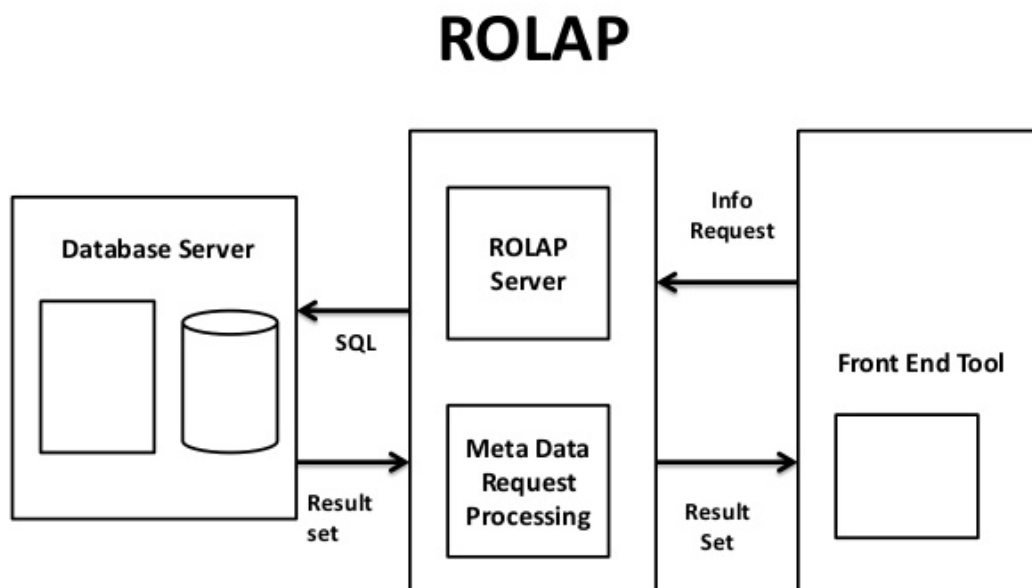
- Διαχειρίζονται δεδομένα με πολλές διαστάσεις και ιεραρχίες, οι οποίες μπορούν να αυξηθούν όσο μεγεθύνεται η επιχείρηση
- Ενοποιούν και συναθροίζουν εκ των προτέρων δεδομένα για ταχύτερα ερωτήματα και αναφορές
- Υποστηρίζουν την πρόσβαση πολλαπλών χρηστών και την εκχώρηση πολλαπλών ερωτημάτων
- Προσφέρουν ταχεία πρόσβαση, υπολογισμό και σύνοψη των επιχειρηματικών δεδομένων ενός οργανισμού

## 2.3.2. Βασικοί τύποι OLAP

Υπάρχουν τρεις βασικοί τύποι συστημάτων OLAP αναφορικά με τον τρόπο φυσικής αποθήκευσης των δεδομένων, οι οποίοι είναι: i) τα συστήματα **σχεσιακού OLAP** (ROLAP - Relational OLAP), ii) τα συστήματα **πολυδιάστατου OLAP** (MOLAP - Multidimensional OLAP) και iii) τα **υβριδικά OLAP** (HOLAP - Hybrid OLAP). Κάθε ένα από αυτά τα συστήματα έχει συγκεκριμένα πλεονεκτήματα και μειονεκτήματα, τα οποία πρέπει να εξετάσει κάθε επιχείρηση πριν υιοθετήσει ένα εξ αυτών (Συμεωνίδης & Γούναρης, 2015).

### 2.3.2.1. ROLAP

Τα συστήματα ROLAP αποθηκεύουν όλα τα δεδομένα, συμπεριλαμβανομένων των συναθροίσεων, στη σχεσιακή βάση δεδομένων, όπως γίνεται αντιληπτό και στην Εικόνα 8, στην οποία παρουσιάζεται η αρχιτεκτονική του συστήματος. Αυτός ο τύπος αποθήκευσης είναι κατάλληλος για επιχειρήσεις που χειρίζονται πολύ μεγάλο όγκο δεδομένων. Τα ROLAP χρησιμοποιούν εργαλεία αναφοράς SQL για την άντληση δεδομένων απευθείας από την αποθήκη δεδομένων.



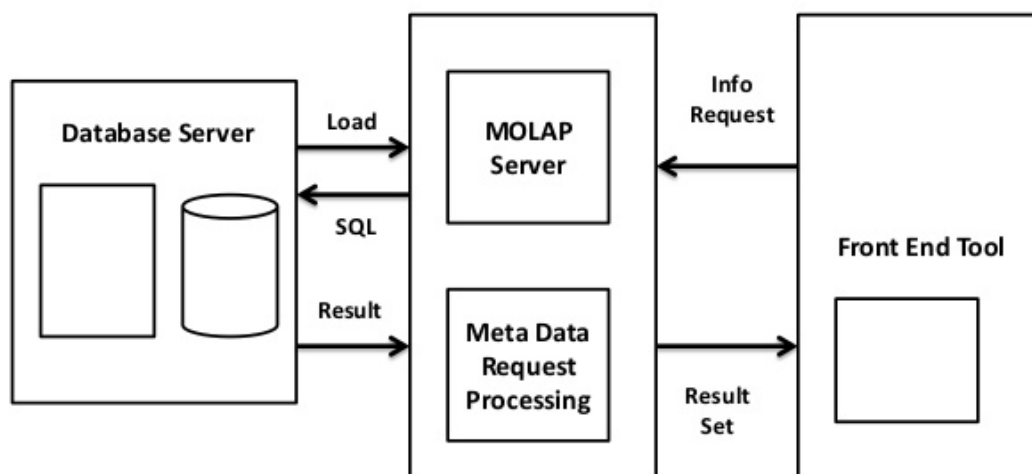
Εικόνα 8: Η αρχιτεκτονική της σχεσιακής OLAP (Saranya, 2013)

Τα κύρια πλεονεκτήματα των ROLAP αφορούν στην ικανότητα διαχείρισης πολύ μεγάλου όγκου δεδομένων και στη δυνατότητα αποτελεσματικής διαχείρισης τόσο αριθμητικών δεδομένων όσο και δεδομένων κειμένου. Επίσης, μπορούν να εκμεταλλευτούν λειτουργικότητες που υπάρχουν στη σχεσιακή βάση δεδομένων. Ωστόσο η ταχύτητα εκτέλεσης των ερωτημάτων είναι μειωμένη σε σχέση με τους υπόλοιπους βασικούς τύπους OLAP, ενώ περιορίζονται και από τις λειτουργικότητες της SQL.

### 2.3.2.2. MOLAP

Τα συστήματα MOLAP είναι ο πιο παραδοσιακός τύπος OLAP. Στην Εικόνα 9 παρουσιάζεται η αρχιτεκτονική τους και ο τρόπος ροής της λειτουργίας τους. Στα MOLAP τόσο τα στοιχεία βάσης όσο και οι υπολογισμοί συνάθροισης αποθηκεύονται σε ένα πολυδιάστατο μοντέλο πινάκων. Αυτός ο τύπος OLAP αποτελεί τη βέλτιστη επιλογή όσον αφορά την ταχύτητα ανάκτησης δεδομένων αλλά απαιτεί επίσης το μεγαλύτερο χώρο αποθήκευσης.

## MOLAP Architecture



Εικόνα 9: Η αρχιτεκτονική της πολυδιάστατης OLAP (Saranya, 2013)

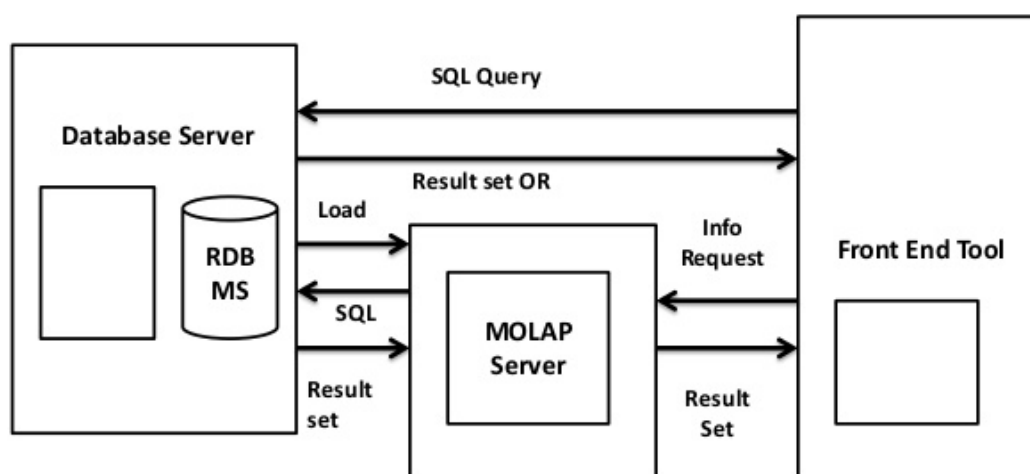


Τα συστήματα MOLAP είναι βελτιστοποιημένα για την καλύτερη απόδοση από άποψη ταχύτητας στα ερωτήματα και την ανάκτηση συνοπτικών δεδομένων. Ωστόσο οι πίνακες είναι αραιοί, αυξάνοντας τον όγκο αποθήκευσης, ενώ ο κύριος περιορισμός τους είναι ότι δεν μπορούν να χειριστούν μεγάλο αριθμό δεδομένων για το λόγο ότι απαιτούν τον εκ των προτέρων υπολογισμό των απαραίτητων συναθροίσεων.

### 2.3.2.3. HOLAP

Τα συστήματα HOLAP προσπαθούν να συνδυάσουν τα πλεονεκτήματα των συστημάτων ROLAP και MOLAP σε ένα ενιαίο σύστημα. Όπως παρουσιάζεται στην Εικόνα 10, τα HOLAP αποθηκεύουν μεγάλο όγκο δεδομένων σε σχεσιακούς πίνακες, ενώ οι υπολογισμοί συναθροίσεων αποθηκεύονται σε κύβους που έχουν εκ των προτέρων υπολογισθεί. Το γεγονός αυτό προσδίδει ευελιξία στην πρόσβαση σε διάφορες πηγές δεδομένων, ταχύτητα στην επεξεργασία των δεδομένων και μεγαλύτερη δυνατότητα κλιμάκωσης.

## HOLAP/MQE/Hybrid architecture



Εικόνα 10: Η αρχιτεκτονική της υβριδικής OLAP (Saranya, 2013)

### 2.3.3. Χαρακτηριστικά εργαλείων OLAP

Ο όρος OLAP παρουσιάστηκε για πρώτη φορά το 1993 από τον Dr. Edgar F. Codd σε σχετικό του άρθρο (Codd, Codd, & Salley, 1993). Στο ίδιο άρθρο ανέφερε δώδεκα κανόνες τους οποίους πρέπει να ακολουθούν τα εργαλεία OLAP.

- **Πολυδιάστατη οπτικοποίηση**

Ο χρήστης πρέπει να έχει τη δυνατότητα να προβάλει τα στοιχεία μιας επιχείρησης σε διάφορες διαστάσεις, όπως για παράδειγμα να εμφανίσει τις πωλήσεις ανά περιοχή, ανά κατάσταση ή ανά πωλητή. Τα μοντέλα πολυδιάστατων δεδομένων επιτρέπουν στο χρήστη την πλήρη διαχείριση αυτών μέσω λειτουργιών όπως η συναθροιστική άνοδος και η αναλυτική κάθοδος.

- **Διαφάνεια**

Ο χρήστης πρέπει να λαμβάνει το αποτέλεσμα του εργαλείου OLAP χωρίς να χρειάζεται να γνωρίζει την αρχιτεκτονική του συστήματος και τη λειτουργία αυτού ή την προέλευση των δεδομένων.

- **Προσβασιμότητα**

Το εργαλείο OLAP (και όχι ο χρήστης) πρέπει να είναι ικανό να ορίζει τον τρόπο πρόσβασης στα διάφορα δεδομένα και να προβαίνει αυτόματα στην απαραίτητη επεξεργασία τους.

- **Σταθερή απόδοση συστήματος αναφορών**

Η απόδοση του εργαλείου OLAP δεν πρέπει να επηρεάζεται σημαντικά όσο αυξάνεται ο αριθμός των διαστάσεων.

- **Αρχιτεκτονική Πελάτη/Διακομιστή**

Ο διακομιστής (Server) του εργαλείου OLAP πρέπει να είναι επαρκώς ικανός να συνδέεται με τους διάφορους πελάτες (Clients) με ελάχιστη προσπάθεια και να καθορίζει και να ενοποιεί δεδομένα από ανόμιες βάσεις δεδομένων.

- **Γενική μορφή διαστάσεων**

Οι διαστάσεις πρέπει να είναι όμοιες μεταξύ τους ως προς τη δομή και τις λειτουργικές δυνατότητες.

- **Δυναμική διαχείριση αραιών πινάκων**

Ο διακομιστής του εργαλείου OLAP πρέπει να έχει την κατάλληλη φυσική δομή ώστε να μπορεί να διαχειρίζεται στο βέλτιστο βαθμό αραιούς πίνακες.

- **Υποστήριξη πολλών χρηστών**

Το εργαλείο OLAP πρέπει να προσφέρει ταυτόχρονη πρόσβαση για ανάκτηση και μεταβολή, ασφάλεια και ακεραιότητα.

- **Άνευ περιορισμών λειτουργίες μεταξύ των διαστάσεων**

Ο υπολογιστικός εξοπλισμός πρέπει να επιτρέπει τους υπολογισμούς και τη διαχείριση των δεδομένων μεταξύ οποιουδήποτε αριθμού διαστάσεων και να μην περιορίζει την εφαρμογή σχέσεων μεταξύ των δεδομένων.

- **Εύληπτη διαχείριση δεδομένων**

Οι εγγενείς λειτουργίες που σχετίζονται με τη διαχείριση των δεδομένων, όπως η συναθροιστική άνοδος και η αναλυτική κάθοδος, πρέπει να επιτυγχάνονται μέσω άμεσων ενεργειών στο μοντέλο και να μην απαιτούν τη χρήση επιπλέον επιλογών ή πολλαπλών διαδρομών στο περιβάλλον του χρήστη.

- **Ευέλικτες αναφορές**

Οι διαδικασίες αναφορών πρέπει να μπορούν να παρουσιάζουν την πληροφορία με οποιοδήποτε τρόπο επιθυμεί ο χρήστης.

- **Απεριόριστες διαστάσεις και επίπεδα συνάθροισης**

Ο αριθμός των υποστηριζόμενων διαστάσεων δεδομένων πρέπει να είναι δυνητικά απεριόριστος. Κάθε διάσταση πρέπει να επιτρέπει στο χρήστη να ορίζει απεριόριστα επίπεδα συνάθροισης μέσω οποιασδήποτε διαδρομής ενοποίησης.

Σε μια εναλλακτική οπτική, η οποία προσπαθεί να απλουστεύσει και να καταστήσει ανεξάρτητο προϊόντων λογισμικού το χαρακτηρισμό ενός εργαλείου ως εργαλείου OLAP, ο Nigel Pendse σε σχετικό του άρθρο προκρίνει το FASMI test (Fast Analysis of Shared Multidimensional Information test) προκειμένου να παρουσιάσει τα βασικά χαρακτηριστικά ενός εργαλείου OLAP, χωρίς αυτά να σχετίζονται και να περιορίζονται από τη διαδικασία υλοποίησής του (Pendse, 2002). Αναλυτικότερα οι κατηγορίες που εξετάζονται είναι οι εξής:

- **Ταχύτητα (Fast)**

Το σύστημα πρέπει να είναι ικανό να ανταποκριθεί στην πλειονότητα των ερωτημάτων του χρήστη μέσα σε πέντε δευτερόλεπτα, με τα πιο απλά ερωτήματα να απαιτούν μόλις ένα δευτερόλεπτο ενώ τα πιο σύνθετα να απαιτούν στο μέγιστο τα 20 δευτερόλεπτα. Η απαίτηση αυτή απορρέει από το

γεγονός ότι, σύμφωνα με μελέτες που έχουν γίνει, οι χρήστες υποθέτουν ότι ένα σύστημα έχει καταρρεύσει εάν δεν ανταποκριθεί μέσα σε τριάντα δευτερόλεπτα και προσπαθούν να τερματίσουν απότομα τη διαδικασία εκτός και εάν το σύστημα τους έχει προειδοποιήσει ότι θα χρειαστεί πολύ χρόνο για το συγκεκριμένο ερώτημα. Ακόμη και σε αυτήν την περίπτωση, η πολύ μεγάλη αναμονή οδηγεί σε απώλεια της συγκέντρωσης του χρήστη με δυσμενή αποτελέσματα για την ποιότητα της ανάλυσης. Φυσικά, αυτή η ταχύτητα δεν είναι εύκολο να επιτευχθεί λόγω του μεγάλου όγκου των δεδομένων, καθώς και της σε πραγματικό χρόνο απαίτησης νέων υπολογισμών. Προς αυτήν την κατεύθυνση, έχουν αναπτυχθεί τεχνικές για τη βελτίωση της ταχύτητας, όπως ειδικές φόρμες καταχώρησης δεδομένων, εκτενείς εκ των προτέρων υπολογισμοί και συγκεκριμένες απαιτήσεις σε εξοπλισμό αλλά υπάρχει ακόμη μεγάλο πεδίο για βελτίωση.

- **Ανάλυση (Analysis)**

Το σύστημα πρέπει να μπορεί να επεξεργαστεί οποιαδήποτε επιχειρησιακή λογική και στατιστική ανάλυση σχετική με το πεδίο εφαρμογής και με το χρήστη και να τη διατηρήσει απλή γι' αυτόν. Είναι απαραίτητο να επιτρέπει στο χρήστη να προβαίνει σε υπολογισμούς για συγκεκριμένους σκοπούς ως μέρος της αναλυτικής διαδικασίας και να παράγει αναφορές με οποιονδήποτε τρόπο επιθυμεί, χωρίς την ανάγκη επαναπρογραμματισμού του συστήματος και ανεξαρτήτως εάν αυτή η ανάλυση γίνει από το ίδιο το σύστημα ή από κάποιο τρίτο πρόσθετο πρόγραμμα.

- **Κοινή χρήση (Shared)**

Το σύστημα πρέπει να ενσωματώνει όλες τις απαραίτητες τεχνικές προδιαγραφές ασφαλείας προκειμένου να είναι δυνατή η ταυτόχρονη πρόσβαση για μεταβολή δεδομένων από διάφορους χρήστες διασφαλίζοντας την ακεραιότητα του συστήματος.

- **Πολυδιάστατο (Multidimensional)**

Αποτελεί το κύριο προαπαιτούμενο δεδομένου ότι είναι το βασικό χαρακτηριστικό ενός εργαλείου OLAP. Το σύστημα πρέπει να προσφέρει τη δυνατότητα για πολυδιάστατη οπτική στα δεδομένα, συμπεριλαμβάνοντας πολλαπλές ιεραρχίες, υπό το πρίσμα ότι αυτός είναι ο πιο δόκιμος τρόπος ανάλυσης επιχειρήσεων και οργανισμών.

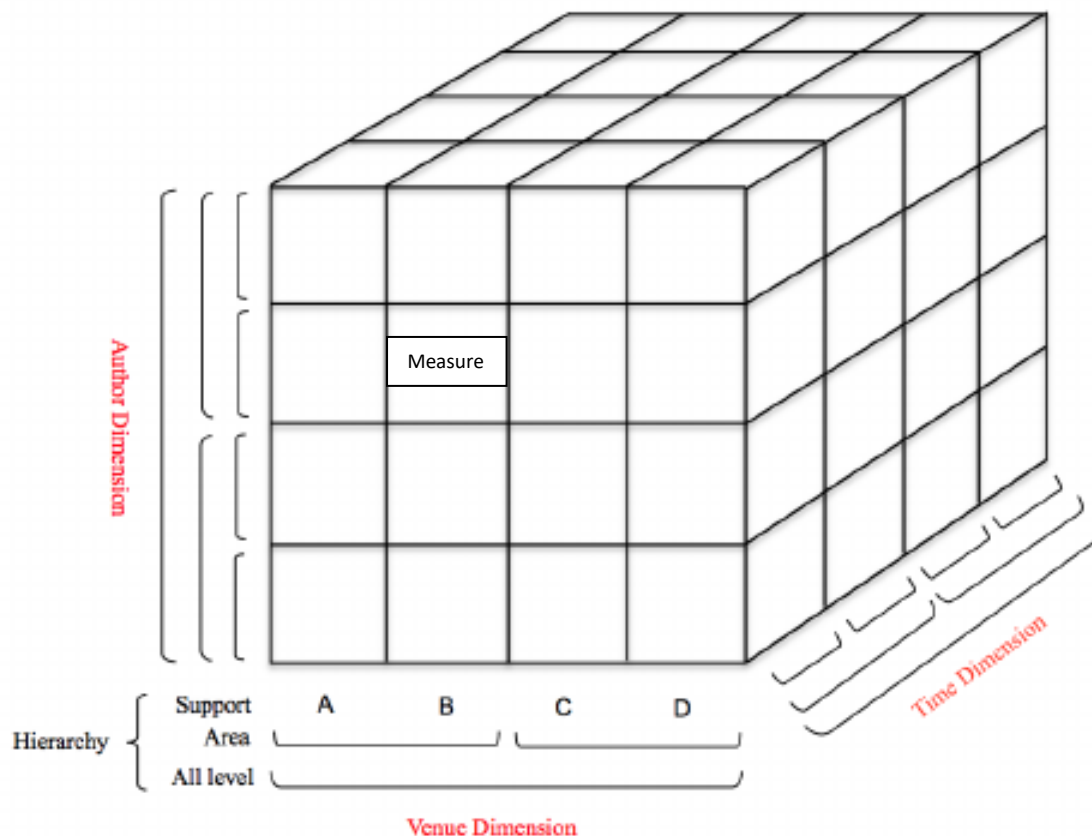
- **Πληροφορία (Information)**

Η έννοια αυτή εμπεριέχει το σύνολο των δεδομένων και των συμπληρωματικών πληροφοριών που χρειάζονται, από όπου και αν προέρχονται αυτά και ανεξαρτήτως του εάν είναι σχετικά με την εφαρμογή. Η μέτρηση της χωρητικότητας ενός συστήματος αναφέρεται στο πόσα δεδομένα μπορεί να διαχειριστεί ως εισροή και όχι πόση χωρητικότητα απαιτείται για να αποθηκευτούν.

### 2.3.4. Κύβος

Ο κύβος OLAP είναι μια πολυδιάστατη συστοιχία δεδομένων και αποτελεί μια βελτιστοποιημένη βάση δεδομένων για χρήση με εφαρμογές OLAP και αποθήκες δεδομένων. Είναι ένας τρόπος αποθήκευσης σχεσιακών δεδομένων σε ένα πολυδιάστατο μοντέλο ώστε να διευκολύνει την αποτελεσματική και ποιοτική ανάλυση αλλά και ώστε το αποτέλεσμα να είναι πιο λογικό και κατανοητό όταν χρησιμοποιούνται τα δεδομένα αυτά για τη δημιουργία αναφορών.

Στον κύβο, όπως παρατηρείται και στην Εικόνα 11, τα δεδομένα αποθηκεύονται και κατηγοριοποιούνται σε **διαστάσεις** (dimensions), οι οποίες χαρακτηρίζονται από **ιεραρχίες** (hierarchies) και από **μέτρα** (measures). Ουσιαστικά οι διαστάσεις ορίζουν τη δομή των δεδομένων, ενώ οι ιεραρχίες αντιπροσωπεύουν το επίπεδο γενίκευσης των διαστάσεων και τα μέτρα είναι αυτά που αποδίδουν τις αριθμητικές τιμές στα ερωτήματα του χρήστη. Το μέτρο κάθε «υποκύβου» έχει υπολογισθεί εκ των προτέρων, γεγονός που βελτιώνει κατά πολύ την ταχύτητα των ερωτημάτων στις σχεσιακές βάσεις δεδομένων.



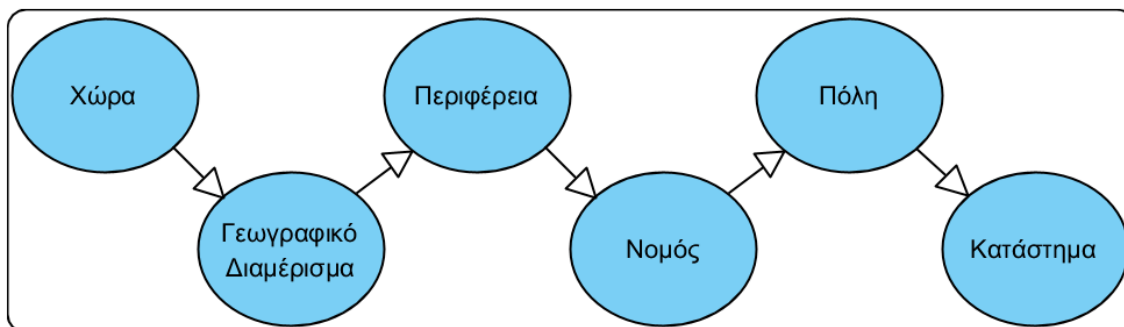
Εικόνα 11: Ο κύβος OLAP (Loudcher, Jakawat, Pavel Soriano Morales, & Favre, 2015)

Ο σχεδιασμός του κύβου γίνεται με επιχειρησιακή λογική ώστε να εξυπηρετήσει την επιχειρηματική ευφυΐα και δεν περιορίζεται από συγκεκριμένες απαιτήσεις συστήματος, οπότε είναι βελτιστοποιημένος για την αναλυτική επεξεργασία. Σε αυτή τη λογική, ο κύβος μπορεί να έχει πάνω από τρεις διαστάσεις, δυνητικά άπειρες, με αποτέλεσμα να μετατρέπεται σε υπερκύβο, που επιτρέπει την εκ βάθους ανάλυση πολύπλοκων και μεγάλου όγκου δεδομένων, ώστε να είναι σε θέση να εξυπηρετήσει πολύπλοκα επιχειρηματικά ερωτήματα.

### 2.3.5. Ιεραρχίες

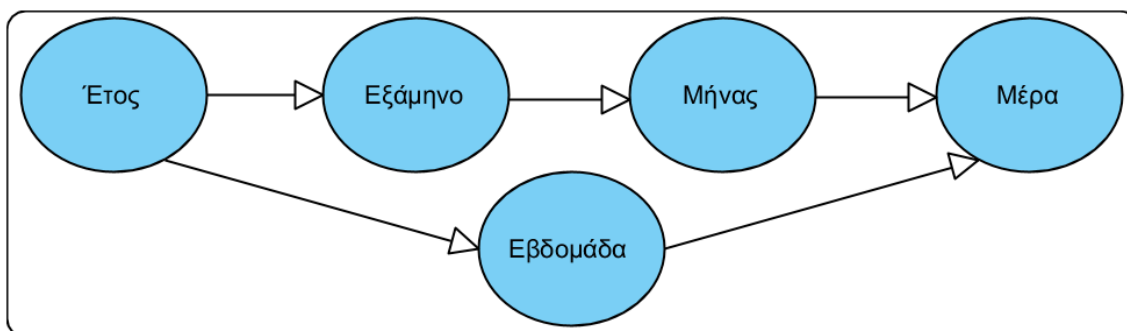
Όπως αναφέρθηκε ήδη, βασικό χαρακτηριστικό των συστημάτων OLAP είναι η ιεραρχιοποίηση των διαστάσεων. Η ιεραρχία είναι μια διαβαθμισμένη κλίμακα απεικόνισης της γενίκευσης, στην οποία ψηλά συναντάται το γενικό και κατεβαίνοντας εφαρμόζεται εξειδίκευση (Κύρκος, 2015). Χαρακτηριστικό παράδειγμα είναι αυτό του τύπου σε ένα πολυδιάστατο μοντέλο πωλήσεων μιας επιχείρησης με πολλά

καταστήματα. Αν θεωρηθεί ότι η επιχείρηση έχει καταστήματα μόνο στην Ελλάδα, η ιεραρχία θα ξεκινούσε από την έννοια «Χώρα» και κατεβαίνοντας θα παρατηρούνταν κατά σειρά το «Γεωγραφικό Διαμέρισμα», η «Περιφέρεια», ο «Νομός», η «Πόλη» και τέλος το «Κατάστημα», όπως φαίνεται και στην Εικόνα 12. Αυτό είναι ένα παράδειγμα γραμμικής ιεραρχίας, δηλαδή παρατηρείται μόνο μια διαδρομή από το γενικότερο στο ειδικότερο.



Εικόνα 12: Γραμμική ιεραρχία (Πηγή: πρωτότυπο)

Στον αντίποδα, υπάρχει η ιεραρχία πλέγματος, στην οποία συναντώνται τουλάχιστον δύο διαδρομές από το γενικότερο στο ειδικότερο. Στο ίδιο μοντέλο πωλήσεων που χρησιμοποιήθηκε παραπάνω, χαρακτηριστικό παράδειγμα είναι η ιεραρχία της διάστασης του χρόνου. Ξεκινώντας από το γενικό που είναι το «Έτος», εξειδικεύεται κατεβαίνοντας στο «Εξάμηνο», στο «Μήνα» και τέλος στη «Μέρα». Μια άλλη διαδρομή που θα μπορούσε να ακολουθηθεί όμως είναι μετά το «Έτος», να ακολουθήσει η «Εβδομάδα» και τέλος η «Μέρα». Το παράδειγμα αυτό παρουσιάζεται στην Εικόνα 13.



Εικόνα 13: Ιεραρχία πλέγματος (Πηγή: πρωτότυπο)

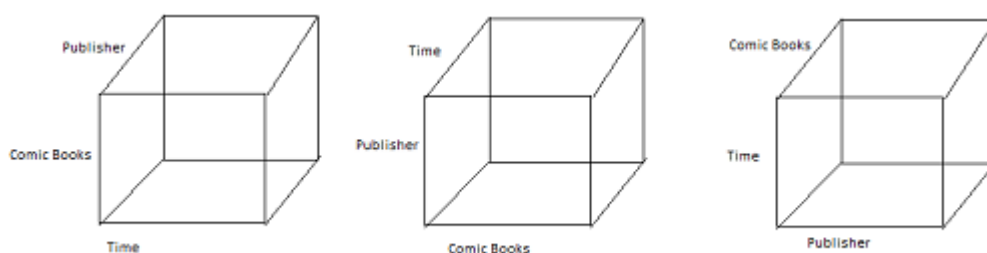
Τα δύο ανωτέρω παραδείγματα αποτελούν τις πιο συχνές και ταυτόχρονα απλές περιπτώσεις. Οι ιεραρχίες πολλές φορές όμως μπορεί να είναι αρκετά πολύπλοκες. Στο παράδειγμα που χρησιμοποιήθηκε προηγουμένως μια τέτοια περίπτωση θα μπορούσε να είναι η ιεραρχία του προϊόντος, η οποία θα μπορούσε να περιέχει πολλές διακριτές διαδρομές ανάλογα με το αν στη διαδρομή χρησιμοποιηθεί η έννοια του «Προμηθευτή», της «Κατηγορίας Προϊόντος», της «Εταιρείας Παραγωγής» κ.α.

### 2.3.6. Βασικές λειτουργίες των κύβων OLAP

Μετά τη δημιουργία του κύβου OLAP και τον ορισμό των ιεραρχιών, το σύνολο των δεδομένων είναι πλέον διαθέσιμο για πολυδιάστατη θεώρηση με τη χρήση των λειτουργιών που προσφέρει το σύστημα. Οι πιο βασικές από αυτές είναι ο **οριζόντιος** και **κάθετος τεμαχισμός** (Slice & Dice), η **συναθροιστική άνοδος** (Roll-up), η **αναλυτική κάθοδος** (Drill-down) και η **περιστροφή** (Pivot) (van der Aalst, 2013).

#### 2.3.6.1. Περιστροφή (Pivot)

Με τη λειτουργία της περιστροφής προσφέρεται η δυνατότητα προβολής του κύβου από άλλη οπτική γωνία, αλλάζοντας τη διάταξη των αξόνων του, χωρίς να μεταβληθούν τα δεδομένα ή να πραγματοποιείται κάποιος εκ νέου υπολογισμός. Στην Εικόνα 14 παρουσιάζεται ένα παράδειγμα περιστροφής κύβου, ο οποίος έχει τρεις διαστάσεις, «χρόνος», «βιβλίο», «εκδότης».



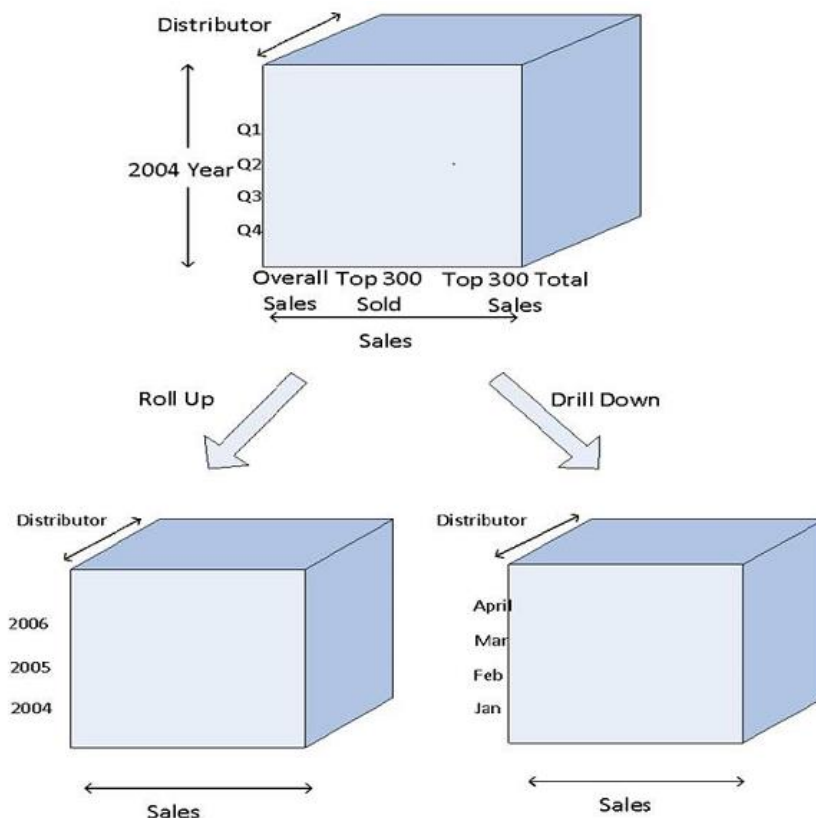
Εικόνα 14: Η λειτουργία Pivot (California State University, 2018)



### 2.3.6.2. Συναθροιστική Άνοδος & Αναλυτική Κάθοδος (Roll-up & Drill-down)

#### Συναθροιστική άνοδος

Η λειτουργία αυτή παρουσιάζει τα δεδομένα του κύβου από μια πιο γενικευμένη οπτική. Αυτό επιτυγχάνεται είτε με την άνοδο στην ιεραρχία μιας διάστασης, όπως φαίνεται στο αριστερό κομμάτι της Εικόνα 15, στην οποία η διάσταση του χρόνου ανεβαίνει επίπεδο από το έτος 2014 στην τριετία 2014-2016, είτε με μείωση των διαστάσεων, όπου αφαιρείται εντελώς μια διάσταση, για παράδειγμα ο χρόνος, και τα δεδομένα εμφανίζονται συνολικά γι' αυτήν τη διάσταση.



Εικόνα 15: Οι λειτουργίες Roll-up και Drill-down (California State University, 2018)

#### Αναλυτική Κάθοδος

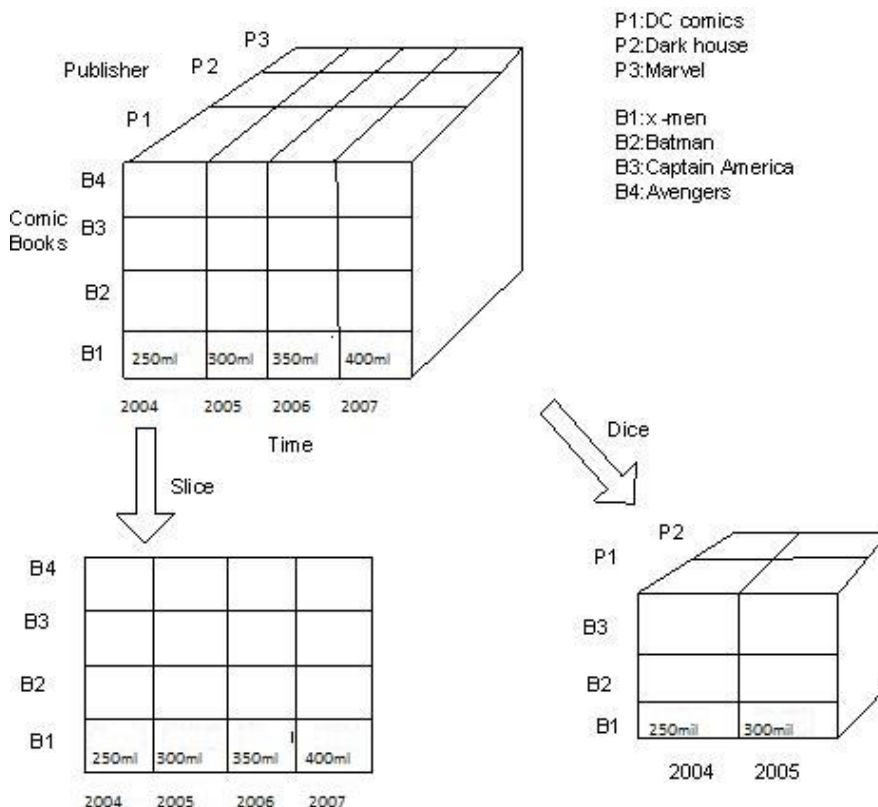
Η λειτουργία αυτή ενεργεί αντιστρόφως της συναθροιστικής ανόδου, μεταβαίνει δηλαδή από τη γενικευμένη οπτική σε επίπεδο μεγαλύτερης λεπτομέρειας. Οι τιμές του κύβου επιμερίζονται σε μικρότερα σύνολα ανάλογα με το εύρος της λεπτομέρειας που έχει επιλέξει ο χρήστης. Αυτό επιτυγχάνεται είτε με την κάθοδο στην ιεραρχία μιας διάστασης, όπως φαίνεται στο δεξιό κομμάτι της Εικόνα 15, στην οποία η διάσταση του

χρόνου κατεβαίνει επίπεδο από το έτος 2014 σε συγκεκριμένους μήνες του ίδιου έτους, είτε με την αύξηση των διαστάσεων του κύβου, για παράδειγμα σε ένα κύβο με δύο διαστάσεις χρόνου και προϊόντος, ο χρήστης μπορεί να προσθέσει τη διάσταση του τόπου και να επιμερίσει τα δεδομένα σε μικρότερα υποσύνολα ανάλογα και με τον τόπο.

### 2.3.6.3. Οριζόντιος και Κάθετος Τεμαχισμός (Slice & Dice)

#### Οριζόντιος τεμαχισμός

Με τη λειτουργία του οριζόντιου τεμαχισμού προκύπτει ένας νέος κύβος (δισδιάστατος ή πολυδιάστατος) κατόπιν επιλογής από το χρήστη μίας ή περισσότερων τιμών από μία μόνο διάσταση. Στο αριστερό κομμάτι της Εικόνα 16 παρατηρείται ότι από τον αρχικό μας κύβο που έχει τις διαστάσεις «χρόνος», «βιβλίο» και «εκδότης», έχει επιλεγθεί μόνο ένας εκδότης και έχει προκύψει ένας νέος δισδιάστατος κύβος που περιέχει όλα τα βιβλία που έχει εκδώσει αυτός ο εκδότης σε όλο το χρονικό διάστημα που αναλύεται στον αρχικό μας κύβο.



Εικόνα 16: Οι λειτουργίες Slice και Dice (California State University, 2018)

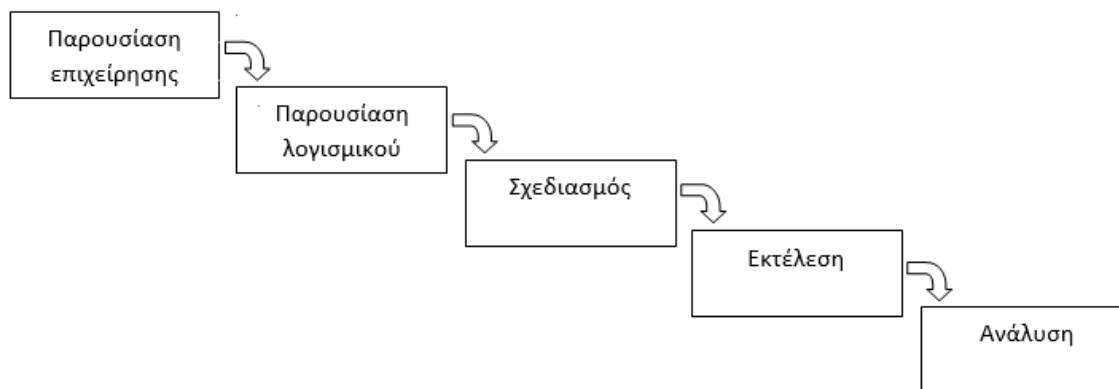
### **Κάθετος τεμαχισμός**

Με τη λειτουργία του κάθετου τεμαχισμού προκύπτει ένας νέος πολυδιάστατος κύβος κατόπιν επιλογής από το χρήστη μίας ή περισσότερων τιμών από δύο ή περισσότερες διαστάσεις. Στο δεξιό κομμάτι της Εικόνα 16 παρατηρείται ότι από τον αρχικό κύβο που έχει τις διαστάσεις «χρόνος», «βιβλίο» και «εκδότης», έχουν επιλεγθεί δύο από τα τέσσερα έτη, τρία από τα τέσσερα βιβλία και δύο από τους τρεις εκδότες, με αποτέλεσμα να προκύψει ένας νέος πολυδιάστατος κύβος που προβάλλει μόνο τα ζητούμενα δεδομένα.

## 3. Μεθοδολογία

### 3.1. Πρόλογος

Η παρούσα μελέτη χωρίστηκε σε δύο μέρη. Στο πρώτο μέρος έγινε ανασκόπηση της βιβλιογραφίας γύρω από τα εξεταζόμενα αντικείμενα, ήτοι την επιχειρηματική ευφυΐα, τις αποθήκες δεδομένων και τα εργαλεία OLAP. Στόχος ήταν να γίνει αντιληπτή η δυναμική χρησιμότητα και η ουσιαστική αναγκαιότητα των συστημάτων επιχειρηματικής ευφυΐας και να γίνει πιο κατανοητή η τεχνολογία πίσω από αυτά τα συστήματα. Στο δεύτερο μέρος θα γίνει παρουσίαση μιας πρακτικής εφαρμογής σε πραγματική επιχείρηση. Αρχικά, όπως φαίνεται και στην Εικόνα 17, θα γίνει μια παρουσίαση όλου του λογισμικού που θα χρησιμοποιηθεί κατά τη διάρκεια της μελέτης. Στη συνέχεια, θα δημιουργηθούν οι απαραίτητες δομές του συστήματος, θα γίνει η κατάλληλη επεξεργασία των πηγών δεδομένων της επιχείρησης, ώστε αυτά να τροφοδοτήσουν το σύστημα, με σκοπό την ανάλυση των δεδομένων και την παρουσίαση αναφορών. Τέλος, θα αξιολογηθεί η όλη διαδικασία.



Εικόνα 17: Τα βήματα της μελέτης κατά την παρουσίαση της πρακτικής εφαρμογής (Πηγή: πρωτότυπο)

### 3.2. Παρουσίαση Επιχείρησης

Στην παρούσα μελέτη θα εξετασθεί μια επιχείρηση του τριτογενούς τομέα. Η επιχείρηση αυτή δραστηριοποιείται στην Ελλάδα και κατά την εξεταζόμενη περίοδο (2016-2017) είχε έξι καταστήματα σε τρεις πόλεις, ενώ είχε εξαήμερη λειτουργία ακολουθώντας το ωράριο εμπορικών καταστημάτων.

Τα στοιχεία που παρουσιάζονται προέρχονται από τα συστήματα συναλλαγών της επιχείρησης, έχουν εξαχθεί σε αρχείο υπολογιστικών φύλλων και παρατίθενται σε πίνακα με πέντε στήλες: i) ημερομηνία, ii) κατάστημα, iii) προϊόν, iv) πελάτης και v) πωλητής, όπως φαίνεται και στην Εικόνα 18. Οι συνολικές εγγραφές φτάνουν τις 330.668.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1	Date	Store	Product	Customer	Salesman											
2	4/1/2016	Store A	Product A	89574124726	Salesman A											
3	4/1/2016	Store A	Product A	89574124726	Salesman A											
4	4/1/2016	Store A	Product A	89574124726	Salesman A											
5	4/1/2016	Store A	Product A	89574124726	Salesman A											
6	4/1/2016	Store A	Product A	89574124726	Salesman A											
7	4/1/2016	Store A	Product A	89574124726	Salesman A											
8	4/1/2016	Store A	Product A	89574124726	Salesman A											
9	4/1/2016	Store A	Product A	89574124726	Salesman A											
10	4/1/2016	Store A	Product A	89574124726	Salesman A											
11	4/1/2016	Store A	Product A	89574124726	Salesman A											
12	4/1/2016	Store A	Product A	89574124726	Salesman A											
13	4/1/2016	Store A	Product A	89574124726	Salesman A											
14	4/1/2016	Store A	Product A	89574124726	Salesman A											
15	4/1/2016	Store A	Product A	89574124726	Salesman A											
16	4/1/2016	Store A	Product A	89574124726	Salesman A											
17	4/1/2016	Store A	Product A	89574124726	Salesman A											
18	4/1/2016	Store A	Product A	89574124726	Salesman A											
19	4/1/2016	Store A	Product A	89574124726	Salesman A											
20	4/1/2016	Store A	Product A	89574124726	Salesman A											
21	4/1/2016	Store A	Product A	89574124726	Salesman B											
22	4/1/2016	Store A	Product A	89574124726	Salesman B											
23	4/1/2016	Store A	Product A	89574124726	Salesman B											
24	4/1/2016	Store A	Product A	89574124726	Salesman A											
25	4/1/2016	Store D	Product A	89574141353	Salesman C											
26	4/1/2016	Store D	Product A	89574141353	Salesman C											
27	4/1/2016	Store D	Product A	89574141353	Salesman C											
28	4/1/2016	Store D	Product A	89574141353	Salesman C											
29	4/1/2016	Store D	Product A	89574141353	Salesman C											
30	4/1/2016	Store D	Product A	89574141353	Salesman C											
31	4/1/2016	Store D	Product A	89574141353	Salesman C											
32	4/1/2016	Store D	Product A	89574141353	Salesman C											
33	4/1/2016	Store D	Product A	89574141353	Salesman C											
34	4/1/2016	Store D	Product A	89574141353	Salesman C											
35	4/1/2016	Store D	Product A	89574141353	Salesman C											
36	4/1/2016	Store D	Product A	89574141353	Salesman C											
37	4/1/2016	Store D	Product A	89574141353	Salesman C											
38	4/1/2016	Store D	Product A	89574141353	Salesman C											
39	4/1/2016	Store D	Product A	89574141353	Salesman C											
40	4/1/2016	Store D	Product A	89574141353	Salesman C											
41	4/1/2016	Store D	Product A	89574141353	Salesman C											
42	4/1/2016	Store D	Product A	89574141353	Salesman C											
43	4/1/2016	Store D	Product A	89574141353	Salesman C											
44	4/1/2016	Store D	Product A	89574141353	Salesman C											
45	4/1/2016	Store D	Product A	89574141353	Salesman C											
46	4/1/2016	Store D	Product A	89574141353	Salesman C											
47	4/1/2016	Store D	Product A	89574141353	Salesman C											

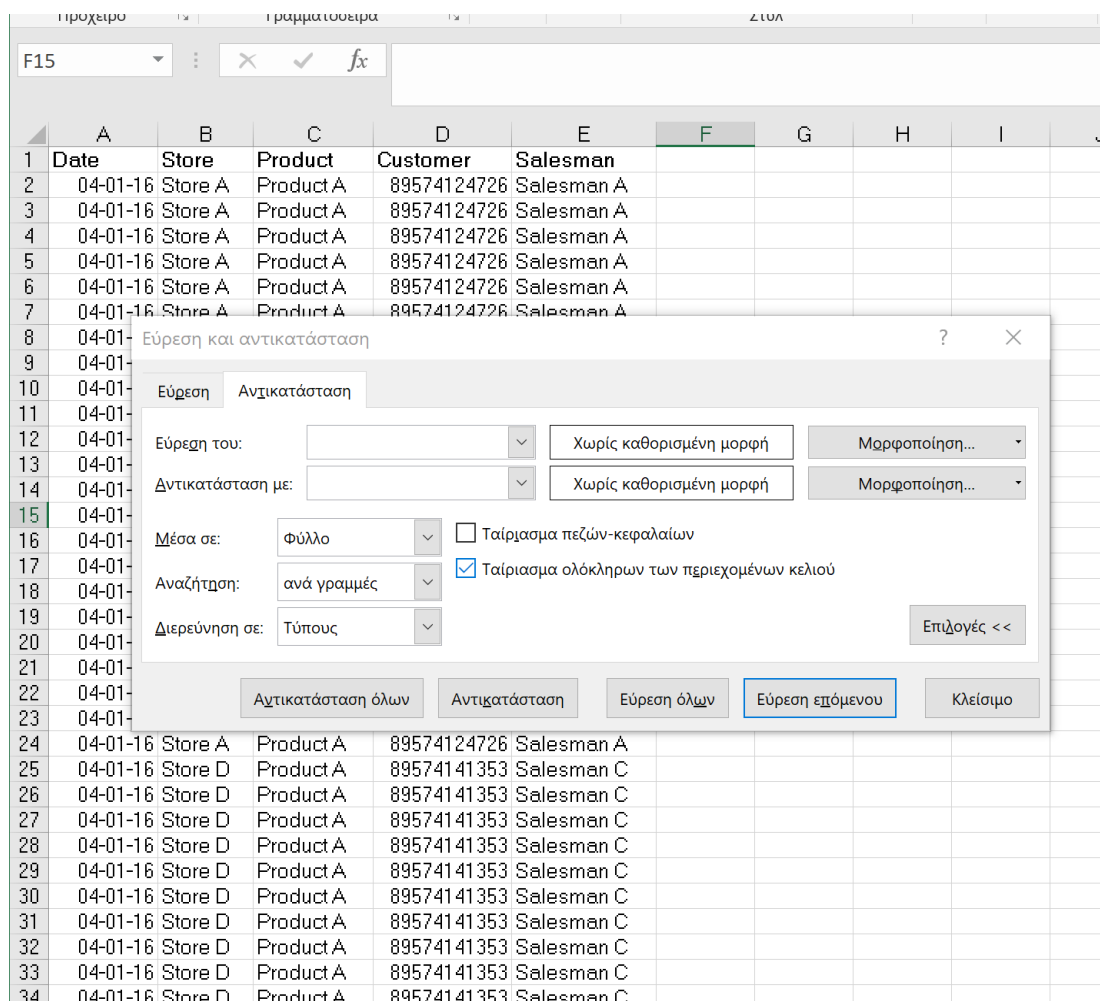
Εικόνα 18: Αρχική μορφή δεδομένων εξαγμένων από τα συστήματα συναλλαγών της επιχείρησης

Για λόγους εμπιστευτικότητας, όλες οι ονομασίες έχουν αντικατασταθεί με γενικά ονόματα. Αυτό επιτυγχάνεται ακολουθώντας τα παρακάτω βήματα:

- Προκειμένου να αντληθούν οι μοναδικές τιμές από κάθε στήλη και να αντικατασταθούν με γενικές ονομασίες, χρησιμοποιείται η φόρμουλα «=IFERROR(INDEX(list;MATCH(0;COUNTIF(unique;list);0));"")», στην οποία ως list αναφέρεται η περιοχή με όλες τις καταχωρήσεις από τις οποίες

πρέπει να αντληθούν οι μοναδικές τιμές, ενώ unique είναι το κελί κάτω από το οποίο θα αρχίσει να συμπληρώνεται η λίστα. Η αναφορά στο κελί θα πρέπει να είναι της μορφής “\$X\$#:X#”, ώστε με τη λειτουργία της αντιγραφής να λειτουργήσει ορθά η φόρμουλα.

- Δημιουργείται ένα νέο αρχείο Excel, στο οποίο μεταφέρονται αυτές οι τιμές και σε κάθε μια αντιστοιχίζεται μια γενική ονομασία. Αυτός ο πίνακας αντιστοίχισης είναι απαραίτητος προκειμένου να είναι δυνατή η ερμηνεία του τελικού αποτελέσματος.
- Τέλος, στο αρχικό αρχείο με τα δεδομένα αντικαθίστανται οι πραγματικές τιμές με τις αντίστοιχες γενικές τους με τη λειτουργία «Εύρεση και Αντικατάσταση» (Εικόνα 19). Πρέπει να δοθεί προσοχή στο ότι στο παράθυρο πρέπει να επιλεγθεί το «Ταίριασμα ολόκληρων των περιεχομένων κελιού».



Εικόνα 19: Λειτουργία "Εύρεσης και Αντικατάστασης" στο Excel

Σε αυτήν τη μορφή, τα δεδομένα είναι δύσκολα επεξεργάσιμα και δεν μπορούν να προσφέρουν πολύτιμη πληροφορία στην επιχείρηση για τη λειτουργία των καταστημάτων της, την απόδοση των υπαλλήλων της και την πορεία πωλήσεων των προϊόντων της. Μέσα από τα βήματα που θα ακολουθήσει αυτή η μελέτη, θα παρουσιασθεί πώς αυτό το σύνολο δεδομένων μπορεί να μετασχηματιστεί σε ένα πολυδιάστατο σύστημα, το οποίο θα είναι ικανό να προσφέρει στην επιχείρηση γνώση που «κρύβεται» μέσα σε αυτό, προκειμένου να εντοπίσει τα δυνατά της σημεία και τις αδυναμίες της.

Προς αυτόν το σκοπό, θα χρησιμοποιηθεί μια σειρά λογισμικού που θα παρουσιαστεί στην επόμενη ενότητα.

### **3.3. Παρουσίαση Λογισμικού**

#### **3.3.1. MS Excel**

Το Microsoft Excel είναι ένα ευρέως διαδεδομένο πρόγραμμα υπολογιστικών φύλλων, το οποίο ανήκει στη σουίτα εφαρμογών Microsoft Office και έχει αναπτυχθεί και κυκλοφορήσει από την εταιρεία Microsoft, με την πρώτη έκδοση να έχει δημοσιευθεί το 1987. Το πρόγραμμα δίνει τη δυνατότητα στο χρήστη να πραγματοποιήσει μια πλειάδα υπολογισμών, βάσει συναρτήσεων ή χρησιμοποιώντας προσαρμοσμένες φόρμουλες, να χρησιμοποιήσει γραφικά εργαλεία, να δημιουργήσει μακροεντολές και συγκεντρωτικούς πίνακες. Το Excel χρησιμοποιεί ένα σύνολο από κελιά διατεταγμένα σε σειρές και στήλες που δίνουν τη δυνατότητα στο χρήστη να οργανώσει και να διαχειριστεί δεδομένα. Ο χρήστης μπορεί επίσης να απεικονίσει αυτά τα δεδομένα σε σειριακά γραφήματα, ιστογράμματα και πίνακες.

Στην παρούσα μελέτη χρησιμοποιείται το MS Excel 2016 μέσω της σουίτας MS Office Professional Plus 2016. Μέσω του προγράμματος έγινε η αρχική συγκέντρωση των δεδομένων, τα οποία επεξεργάστηκαν προκειμένου να καθαριστούν και να ομογενοποιηθούν με διάφορα εργαλεία του προγράμματος. Στη συνέχεια δημιουργήθηκαν οι πίνακες και οι καταχωρήσεις αυτών, σε ξεχωριστά λογιστικά

φύλλα, που θα χρησιμοποιηθούν στην αποθήκη δεδομένων που θα σχεδιασθεί και θα υλοποιηθεί παρακάτω.

### **3.3.2. Visual Paradigm**

Το Visual Paradigm είναι ένα ισχυρό εργαλείο για το σχεδιασμό πληροφοριακών συστημάτων. Περιέχει όλα τα απαραίτητα στοιχεία που μπορούν να βοηθήσουν το χρήστη σε όλη τη διαδικασία της σχεδίασης, της ανάπτυξης του κώδικα και της υλοποίησης ενός πληροφοριακού συστήματος, όπως εργαλεία για μοντελοποίηση με UML, ευέλικτη ανάπτυξη λογισμικού (agile development), σχεδιασμό βάσεων δεδομένων, μοντελοποίηση επιχειρηματικών διαδικασιών κ.α.

Στην παρούσα μελέτη θα χρησιμοποιηθεί η έκδοση 15.0 μέσω ακαδημαϊκής άδειας και ειδικότερα το εργαλείο για τη σχεδίαση βάσεων δεδομένων (Database Modeling), για τη σχεδίαση του σχήματος της βάσης που θα χρησιμοποιηθεί στην αποθήκη δεδομένων.

### **3.3.3. MS SQL Server**

Ο Microsoft SQL Server είναι ένα σύστημα διαχείρισης σχεσιακών βάσεων δεδομένων (RDBMS - Relational Database Management System), το οποίο έχει αναπτυχθεί και κυκλοφορήσει από την εταιρεία Microsoft, με πρώτη έκδοση αυτή του 1989. Προσφέρει ένα μεγάλο εύρος δυνατοτήτων επεξεργασίας συναλλαγών, καθώς και εφαρμογές επιχειρηματικής ευφυΐας και ανάλυσης. Υποστηρίζει την ANSI SQL, τη βασική γλώσσα SQL, καθώς και την T-SQL, η οποία αποτελεί υλοποίηση της εταιρείας. Το κύριο εργαλείο διεπαφής της πλατφόρμας είναι ο SQL Server Management Studio (SSMS).

Στην παρούσα μελέτη θα χρησιμοποιηθεί ο MS SQL Server 2014 προκειμένου να δημιουργηθεί η αποθήκη δεδομένων.



### **3.3.4. MS Visual Studio και SQL Server Data Tools**

Το MS Visual Studio είναι ένα ολοκληρωμένο περιβάλλον ανάπτυξης (IDE - Integrated Development Environment), το οποίο έχει αναπτυχθεί και κυκλοφορήσει από την εταιρεία Microsoft, με πρώτη έκδοση αυτή του 1997, και χρησιμοποιείται κυρίως στη ανάπτυξη εφαρμογών υπολογιστών, ιστοτόπων και φορητών συσκευών. Υποστηρίζει εγγενώς ένα μεγάλο εύρος από γλώσσες προγραμματισμού (πάνω από 30) και μέσω προσθέτων δύναται να υποστηρίζει ακόμη περισσότερες.

Το SQL Server Data Tools, το οποίο εφεξής θα αναφέρεται ως SSDT, είναι ένα πακέτο εργαλείων για τη δημιουργία σχεσιακών βάσεων δεδομένων και υποστηρίζει διαδικασίες ανάλυσης, αναφοράς και ενσωμάτωσης δεδομένων.

Στην παρούσα μελέτη θα χρησιμοποιηθεί το MS Visual Studio Enterprise 2015, με προεγκατεστημένο το πρόσθετο SSDT, ως πλατφόρμα για την εκτέλεση του προτύπου «Analysis Services Multidimensional and Data Mining», προκειμένου να δημιουργηθεί το πολυδιάστατο μοντέλο (κύβος OLAP). Επίσης, με το πρότυπο «Integration Services» θα δημιουργηθεί μια διαδικασία (data migration) για την ανανέωση της βάσης και του κύβου. Τέλος, θα προστεθεί μια εργασία στον SQL Server Agent, η οποία θα εκτελεί τη συγκεκριμένη διαδικασία βάσει ενός χρονοδιαγράμματος.

### **3.3.5. TARGIT Decision Suite**

Το TARGIT Decision Suite είναι μια πλατφόρμα λογισμικού της εταιρείας TARGIT A/S, που κυκλοφόρησε για πρώτη φορά το 1996. Η πλατφόρμα επιτρέπει τη διασύνδεση σε πολλαπλές πηγές δεδομένων και συστημάτων προκειμένου να προσφέρει μια ενιαία λύση επιχειρηματικής ευφυΐας και ανάλυσης. Μεταξύ άλλων παρέχει ψηφιακά ταμπλό (dashboards) με πληροφορία πραγματικού χρόνου, λεπτομερείς αναλύσεις και δυνατότητα δημιουργίας μοτίβων εκ του μηδενός, διεπαφή με ανάγκη ελάχιστης συμμετοχής από το χρήστη, φορητότητα, εικονογραφημένα σενάρια (storyboards) κ.α.

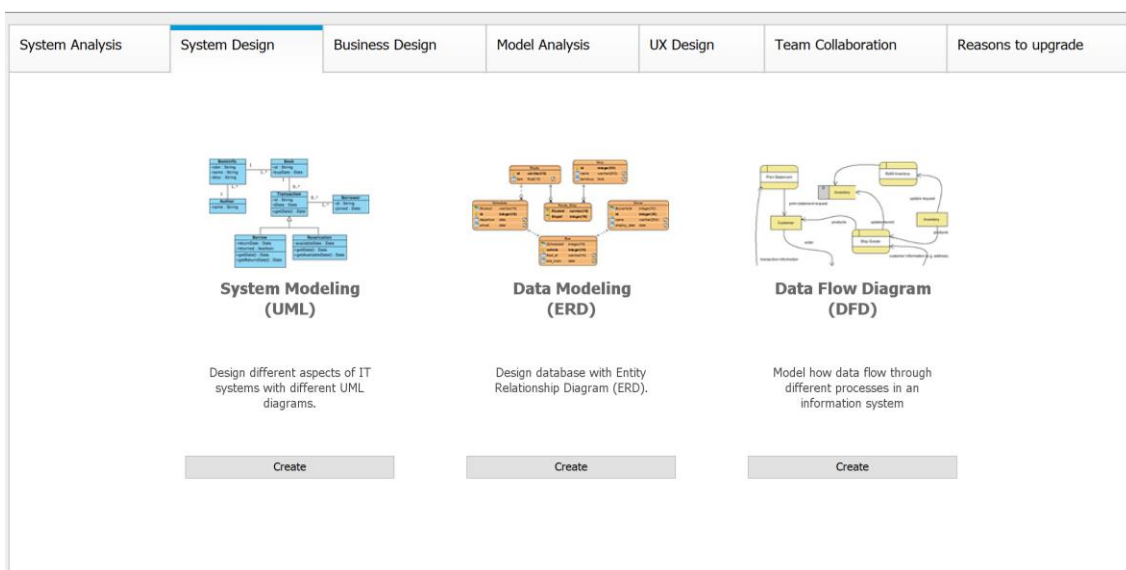
Στην παρούσα εργασία θα χρησιμοποιηθεί το TARGIT Decision Suite 2017 για την οπτικοποίηση της πληροφορίας που θα προκύψει από το σύστημα που έχει σχεδιαστεί

και υλοποιηθεί και η οποία θα προσφέρει ουσιαστικά τη νέα πληροφορία στην επιχείρηση.

### 3.4. Σχεδιασμός και εκτέλεση

Ακολουθώντας το σχεδιασμό ενός συστήματος επιχειρηματικής ευφυΐας, όπως αυτό παρουσιάστηκε στην ενότητα 2.1.4, συναντώνται στο πρώτο επίπεδο οι πηγές δεδομένων, που στην περίπτωση της επιχείρησης που εξετάζεται είναι το σύστημα συναλλαγών της. Δεδομένου ότι στα πλαίσια της μελέτης δεν μπορεί να υπάρξει διασύνδεση με αυτό το σύστημα, τα δεδομένα των πηγών έχουν εξαχθεί εκ των προτέρων σε αρχείο υπολογιστικών φύλλων (Εικόνα 18). Βάσει των στοιχείων αυτών, θα σχεδιασθεί στη συνέχεια το σχήμα της αποθήκης δεδομένων μέσω του Visual Paradigm.

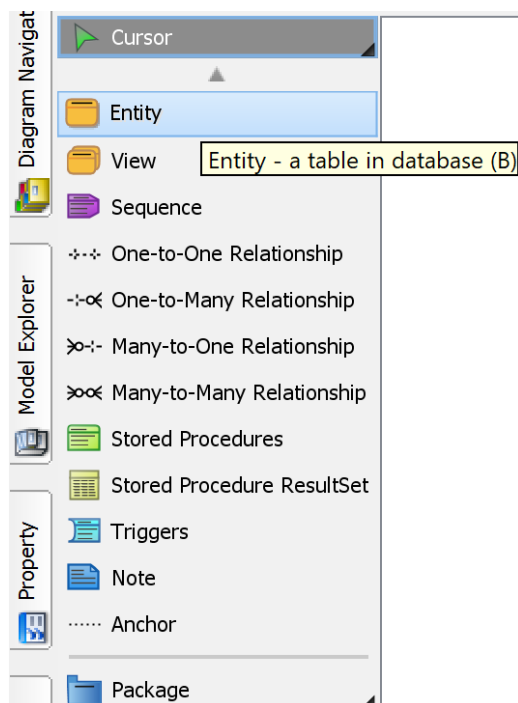
Αφού εκτελεσθεί το πρόγραμμα, παρουσιάζεται το αρχικό μενού επιλογών (Εικόνα 20).



Εικόνα 20: Αρχικό μενού Visual Paradigm

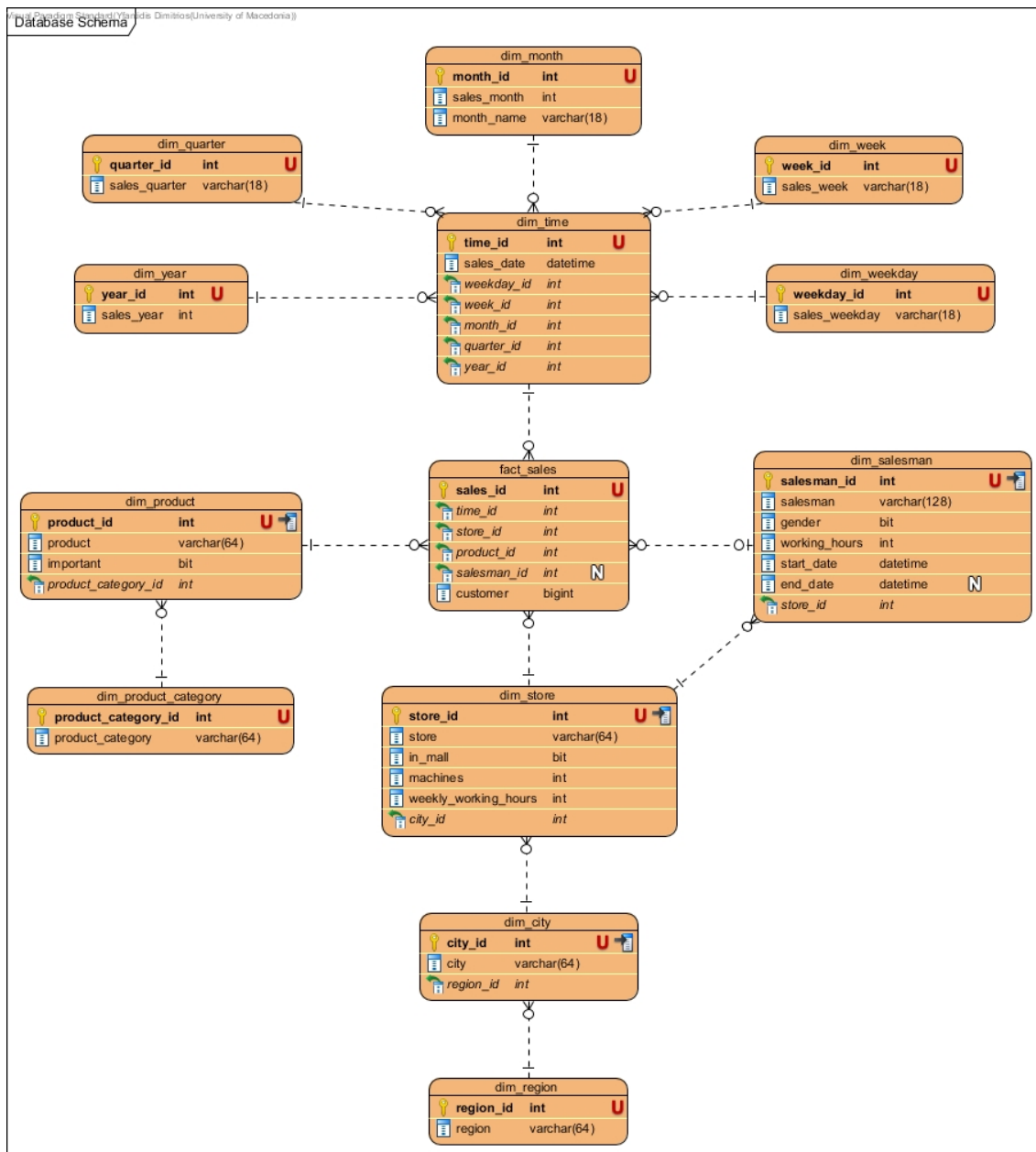
Από εκεί επιλέγεται η καρτέλα «System Design» και στη συνέχεια «Create» στο υπόδειγμα «Data Modeling (ERD)», οπότε και ανοίγει το παράθυρο όπου μπορεί να αρχίσει ο σχεδιασμός του σχήματος της βάσης. Στην αριστερή στήλη υπάρχουν τα διάφορα στοιχεία που μπορούν να προστεθούν με «drag & drop» στο παράθυρο σχεδιασμού (Εικόνα 21). Τα βασικά στοιχεία που θα χρησιμοποιηθούν είναι η

«Οντότητα» (Entity), η οποία θα αναπαριστά τους πίνακες γεγονότων και διαστάσεων, καθώς και η «Σχέση Ένα-προς-Πολλά» (One-to-Many Relationship), που θα αντικατοπτρίζει τη διασύνδεση των πινάκων.



Εικόνα 21: Επιλογές στοιχείων για τη σχεδίαση του σχήματος της αποθήκης δεδομένων

Για τη βάση θα χρησιμοποιηθεί το σχήμα χιονονιφάδας, στο οποίο θα δημιουργηθεί ένας κεντρικός πίνακας γεγονότων (fact\_sales), τέσσερις πίνακες βασικών διαστάσεων (dim\_time, dim\_product, dim\_store και dim\_salesman) και οκτώ πίνακες βοηθητικών διαστάσεων (dim\_weekday, dim\_week, dim\_month, dim\_quarter, dim\_year, dim\_product\_category, dim\_city και dim\_region). Οι σχέσεις είναι όλες Ένα-προς-Πολλά με κατεύθυνση από το γενικό προς το ειδικό. Το τελικό αποτέλεσμα είναι αυτό της Εικόνα 22.



Εικόνα 22: Σχήμα χιονονιφάδας της αποθήκης δεδομένων

Αναλυτικότερα τα στοιχεία των πινάκων έχουν ως εξής:

- fact\_sales: sales\_id – int<sup>1</sup> (PK)<sup>2</sup>, time\_id – int (FK)<sup>3</sup>, store\_id – int (FK), product\_id – int (FK), salesman\_id – int – nullable (FK), customer – bigint<sup>4</sup>

<sup>1</sup> int: integer

<sup>2</sup> PK: Primary Key

<sup>3</sup> FK: Foreign Key

<sup>4</sup> bigint: big integer

- dim\_salesman: salesman\_id – int (PK), salesman – varchar<sup>5</sup>(128), gender – bit<sup>6</sup>, working\_hours – int, start\_date – datetime, end\_date – datetime – nullable, store\_id – int (FK)
- dim\_store: store\_id – int (PK), store – varchar(64), in\_mall – bit, machines – int, weekly\_working\_hours – int, city\_id – int (FK)
- dim\_city: city\_id – int (PK), city – varchar(64), region\_id – int (FK)
- dim\_region: region\_id – int (PK), region – varchar(64)
- dim\_product: product\_id – int (PK), product – varchar(64), important – bit, product\_category\_id – int (FK)
- dim\_product\_category: product\_category\_id – int (PK), product\_category – varchar(64)
- dim\_time: time\_id – int (PK), sales\_date – datetime, weekday\_id – int (FK), week\_id – int (FK), month\_id – int (FK), quarter\_id – int (FK), year\_id – int (FK)
- dim\_weekday: weekday\_id – int (PK), sales\_weekday – varchar(18)
- dim\_week: week\_id – int (PK), sales\_week – varchar(18)
- dim\_month: month\_id – int (PK), sales\_month – int, month\_name – varchar(18)
- dim\_quarter: quarter\_id – int (PK), sales\_quarter – varchar(18)
- dim\_year: year\_id – int (PK), sales\_year – int

Πρέπει να σημειωθεί ότι όλα τα primary keys είναι «unique» και δηλώθηκαν ως «always generated», προκειμένου να δημιουργούνται αυτόματα από το σύστημα σε κάθε νέα καταχώρηση.

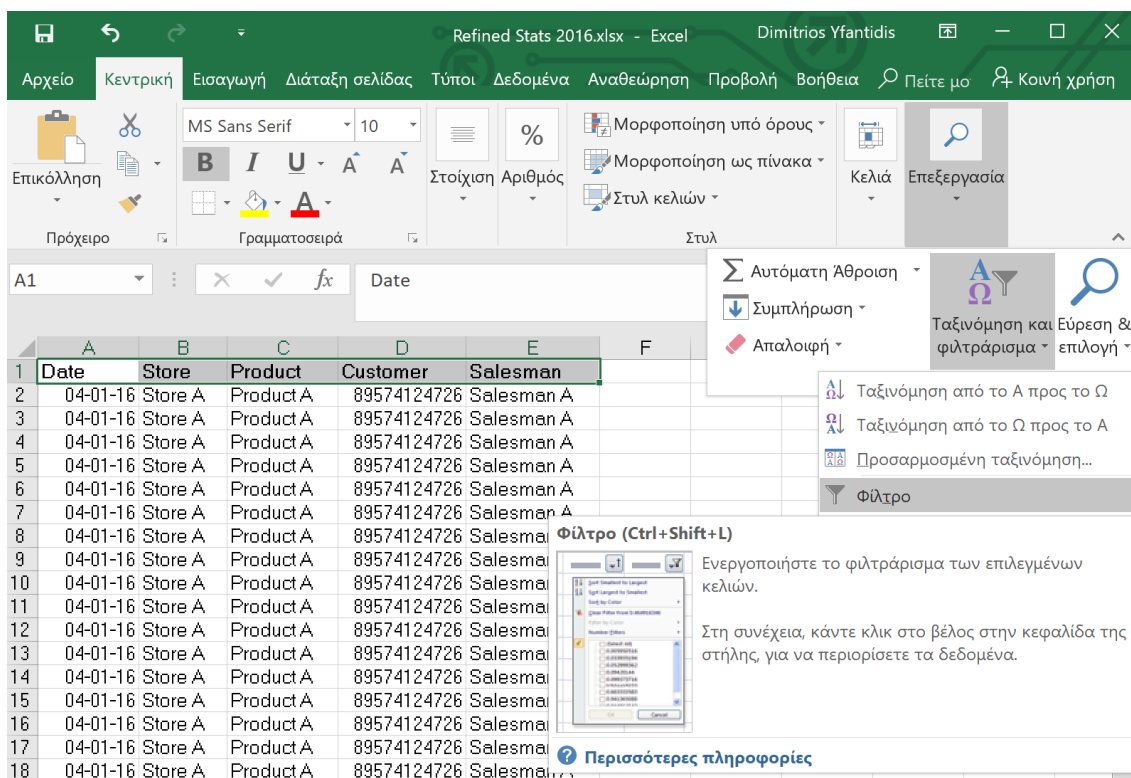
Στη συνέχεια και βάσει αυτού του σχήματος, θα γίνει η επεξεργασία των εξαχθέντων στοιχείων που υπάρχουν στο αρχείο υπολογιστικών φύλλων, προκειμένου να είναι έτοιμα για την εισαγωγή στην αποθήκη δεδομένων. Αυτή η διαδικασία αφορά στο επίπεδο διεργασιών ETL, που στα πλαίσια της παρούσας μελέτης και με δεδομένη την αδυναμία ουσιαστικής διασύνδεσης με το πραγματικό σύστημα συναλλαγών της επιχείρησης, θα γίνει με τη βοήθεια του Excel από το μελετητή.

---

<sup>5</sup> varchar: variable character

<sup>6</sup> bit: binary digit

Ανοίγοντας λοιπόν το αρχείο με τα δεδομένα της επιχείρησης, αρχικά γίνεται προσπάθεια να απομακρυνθούν λανθασμένες εγγραφές, που στην εξεταζόμενη περίπτωση είναι εγγραφές με κενές τιμές σε βασικά πεδία (κελιά). Αυτό επιτυγχάνεται εφαρμόζοντας φίλτρα αναζήτησης στα δεδομένα και επιλέγοντας να εμφανίζονται εγγραφές με κενές τιμές, εγγραφές οι οποίες απομακρύνονται (Εικόνα 23). Στο χρησιμοποιηθέν δείγμα ο αριθμός αυτός ήταν πολύ μικρός (μόλις τέσσερις εγγραφές). Με τα φίλτρα μπορεί επίσης να ελεγχθεί εάν υπάρχουν μη αναμενόμενες ή και λάθος εγγραφές.



Εικόνα 23: Εισαγωγή φίλτρου στα δεδομένα του αρχείου υπολογιστικών φύλλων

Ακολουθώντας, δημιουργείται από ένα νέο φύλλο εργασίας για κάθε πίνακα που έχει σχεδιαστεί στο σχήμα μας. Ξεκινώντας από τους βοηθητικούς πίνακες και καταλήγοντας στον κεντρικό πίνακα γεγονότων, αντικαθίστανται οι τιμές με δείκτες, όπου αυτό προβλέπεται βάσει του σχήματος,. Το τελικό αποτέλεσμα παρουσιάζεται στην Εικόνα 24 και την Εικόνα 25.

sales_id	time_id	store_id	product_id	salesman_id	customer
1	4	1	1	1001	89574124726
2	4	1	1	1001	89574124726
3	4	1	1	1001	89574124726
4	4	1	1	1001	89574124726
5	4	1	1	1001	89574124726
6	4	1	1	1001	89574124726
7	4	1	1	1001	89574124726
8	4	1	1	1001	89574124726
9	4	1	1	1001	89574124726
10	4	1	1	1001	89574124726
11	4	1	1	1001	89574124726
12	4	1	1	1001	89574124726
13	4	1	1	1001	89574124726
14	4	1	1	1001	89574124726
15	4	1	1	1001	89574124726
16	4	1	1	1001	89574124726
17	4	1	1	1001	89574124726
18	4	1	1	1001	89574124726
19	4	1	1	1001	89574124726
20	4	1	1	1001	89574124726
21	4	1	1	1002	89574124726
22	4	1	1	1002	89574124726
23	4	1	1	1002	89574124726
24	4	1	1	1001	89574124726
25	4	4	1	1003	89574141353
26	4	4	1	1003	89574141353
27	4	4	1	1003	89574141353
28	4	4	1	1003	89574141353
29	4	4	1	1003	89574141353
30	4	4	1	1003	89574141353
31	4	4	1	1003	89574141353
32	4	4	1	1003	89574141353
33	4	4	1	1003	89574141353
34	4	4	1	1003	89574141353

Εικόνα 24: Καταχωρήσεις πίνακα fact\_sales

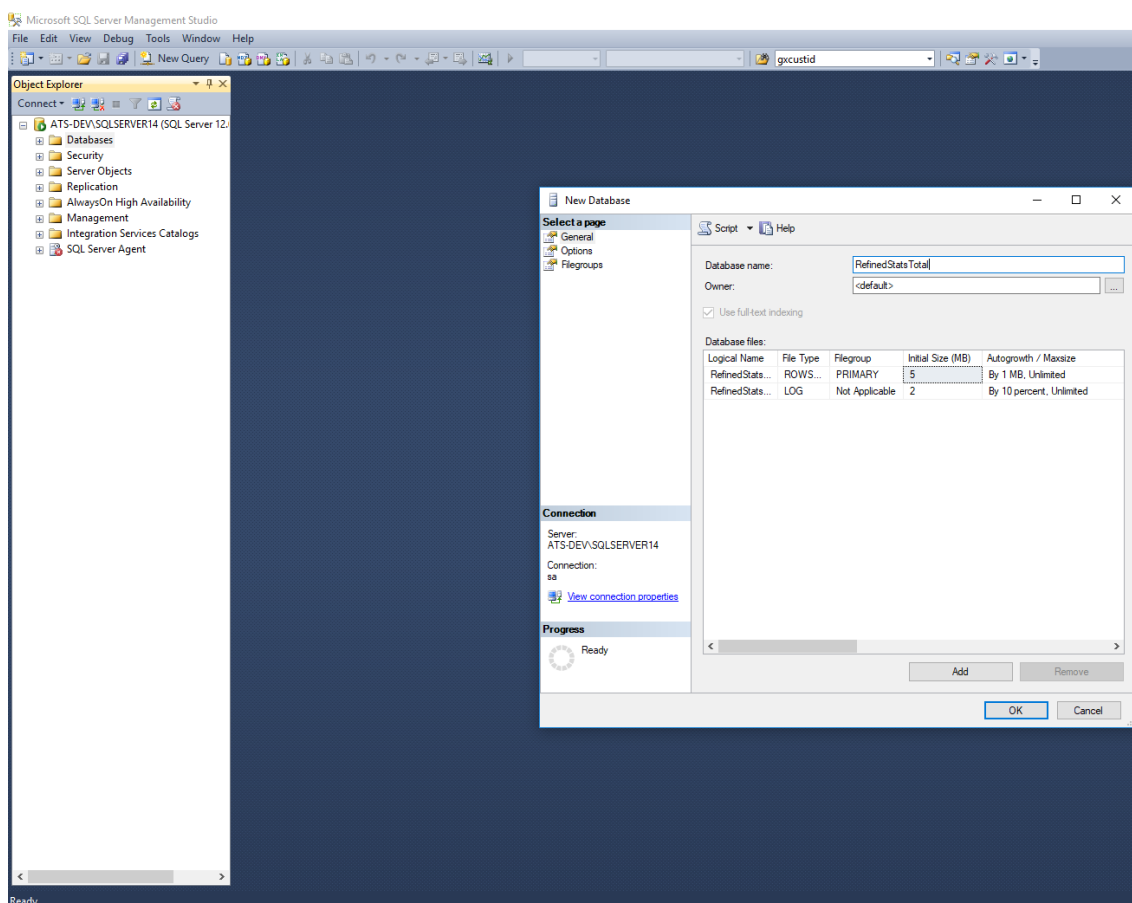
time_id	sales_date	weekday_id	week_id	month_id	quarter_id	year_id
1	01-01-16	5	53	1	1	1
2	02-01-16	6	53	1	1	1
3	03-01-16	7	53	1	1	1
4	04-01-16	1	1	1	1	1
5	05-01-16	2	1	1	1	1
6	06-01-16	3	1	1	1	1
7	07-01-16	4	1	1	1	1
8	08-01-16	5	1	1	1	1
9	09-01-16	6	1	1	1	1
10	10-01-16	7	1	1	1	1
11	11-01-16	1	2	1	1	1
12	12-01-16	2	2	1	1	1
13	13-01-16	3	2	1	1	1
14	14-01-16	4	2	1	1	1
15	15-01-16	5	2	1	1	1
16	16-01-16	6	2	1	1	1
17	17-01-16	7	2	1	1	1
18	18-01-16	1	3	1	1	1
19	19-01-16	2	3	1	1	1
20	20-01-16	3	3	1	1	1
21	21-01-16	4	3	1	1	1
22	22-01-16	5	3	1	1	1
23	23-01-16	6	3	1	1	1
24	24-01-16	7	3	1	1	1
25	25-01-16	1	4	1	1	1
26	26-01-16	2	4	1	1	1
27	27-01-16	3	4	1	1	1
28	28-01-16	4	4	1	1	1
29	29-01-16	5	4	1	1	1
30	30-01-16	6	4	1	1	1
31	31-01-16	7	4	1	1	1
32	01-02-16	1	5	2	1	1
33	02-02-16	2	5	2	1	1

Εικόνα 25: Καταχωρήσεις πίνακα dim\_time

Να σημειωθεί ότι όλα τα κελιά έχουν μορφοποιηθεί ανάλογα με το περιεχόμενό τους (αριθμός, κείμενο, ημερομηνία) μέσω της «Μορφοποίησης κελιών».

Πλέον τα δεδομένα είναι έτοιμα να εισαχθούν στην αποθήκη δεδομένων, που θα υλοποιηθεί στη συνέχεια στον MS SQL Server, διαδικασία που ανήκει στο επόμενο επίπεδο ενός συστήματος επιχειρηματικής ευφυΐας, όπως αυτό παρουσιάστηκε στην ενότητα 2.1.4.

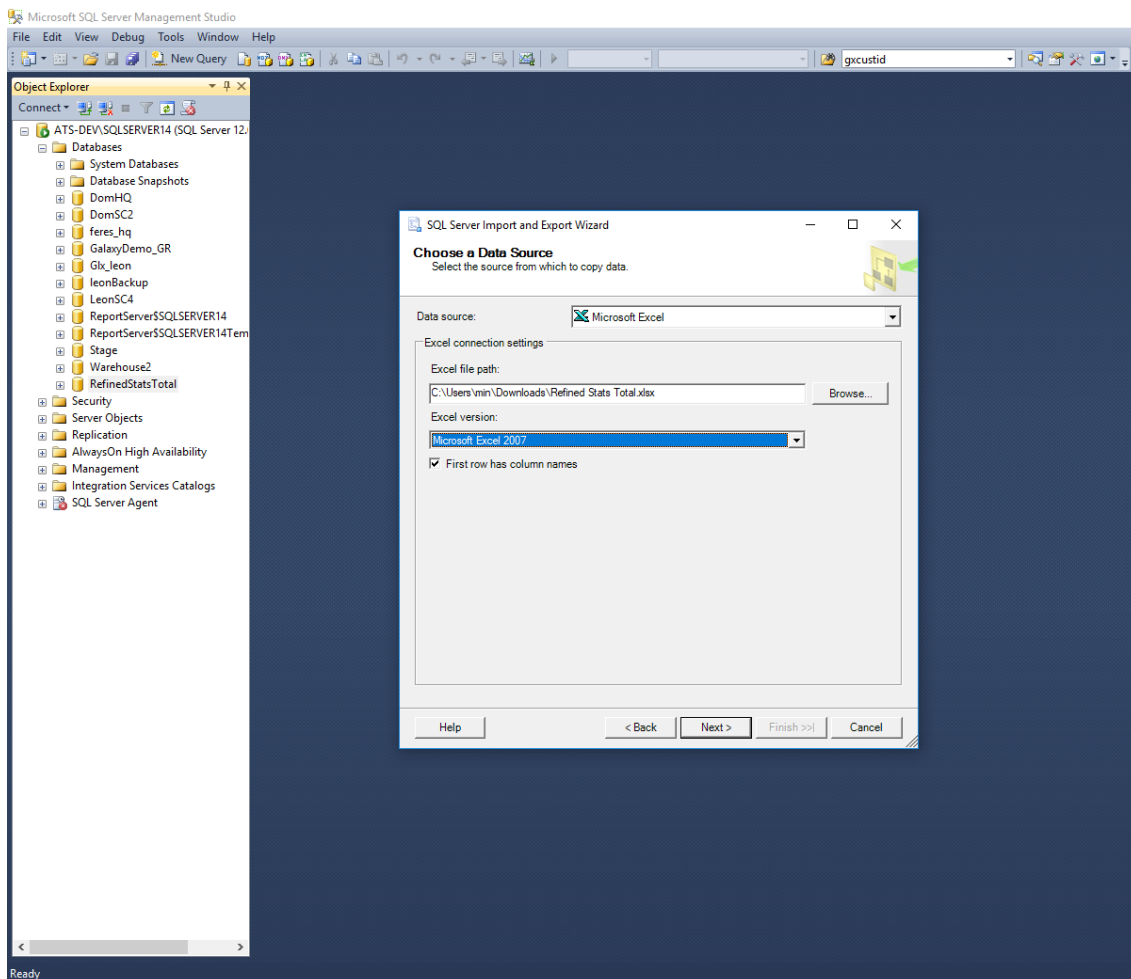
Η διαχείριση νέων και υφιστάμενων βάσεων δεδομένων στον MS SQL Server γίνεται μέσω του MS SQL Server Management Studio. Μέσω αυτού λοιπόν, επιχειρείται η δημιουργία μιας νέας βάσης δεδομένων με όνομα RefinedStatsTotal (Εικόνα 26).



Εικόνα 26: Δημιουργία της βάσης δεδομένων στον MS SQL Server

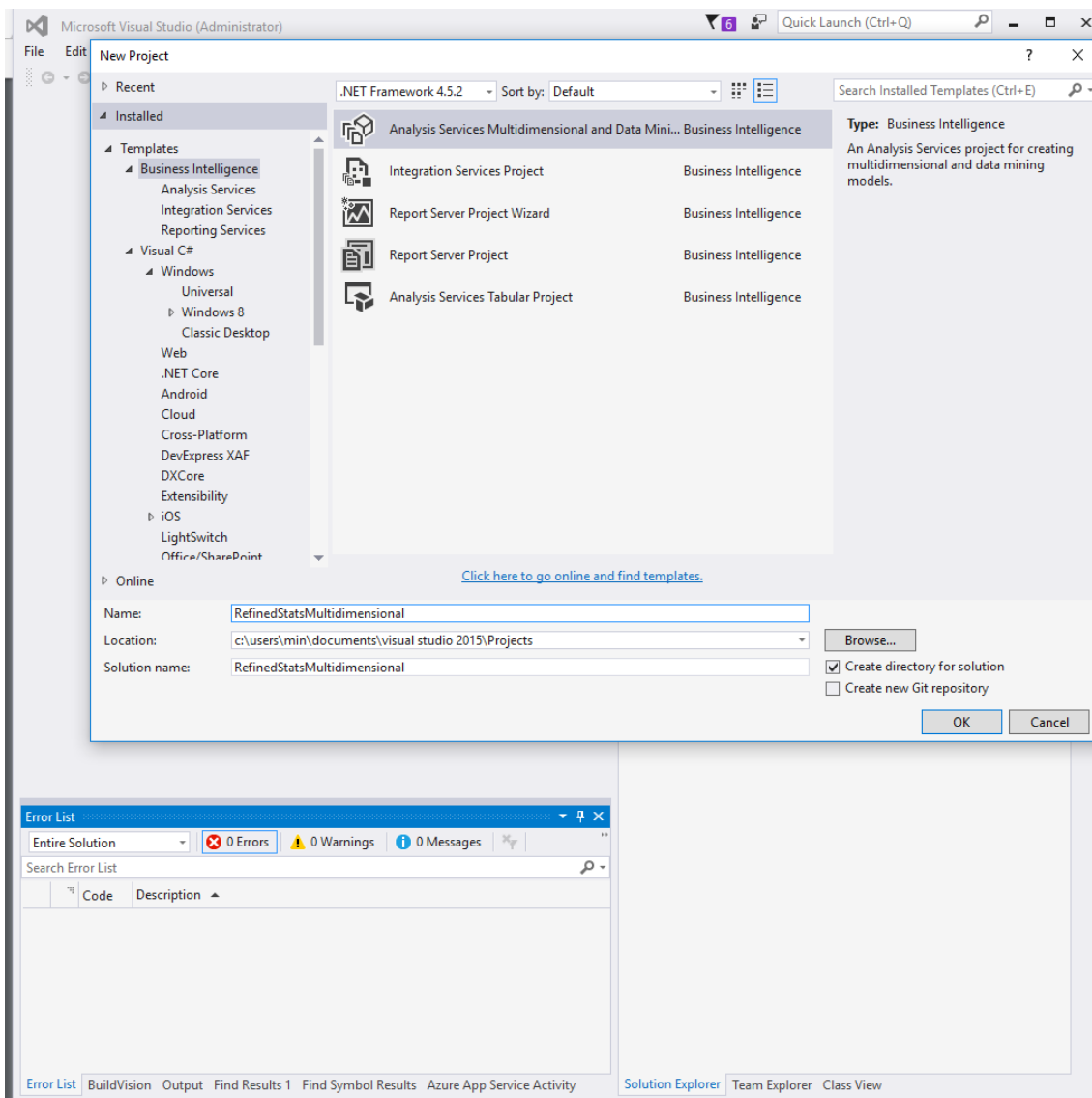


Η βάση υλοποιείται σύμφωνα με το σχήμα που είχε σχεδιαστεί προηγουμένως (Εικόνα 22). Αφού δημιουργηθεί η βάση, ακολουθεί η εισαγωγή των δεδομένων σε αυτή. Αυτό το βήμα πραγματοποιείται με τη βοήθεια του εργαλείου SQL Server Import and Export Wizard (Εικόνα 27). Πηγή των δεδομένων αποτελεί το αρχείο υπολογιστικών φύλλων που δημιουργήθηκε προηγουμένως (Εικόνα 24).



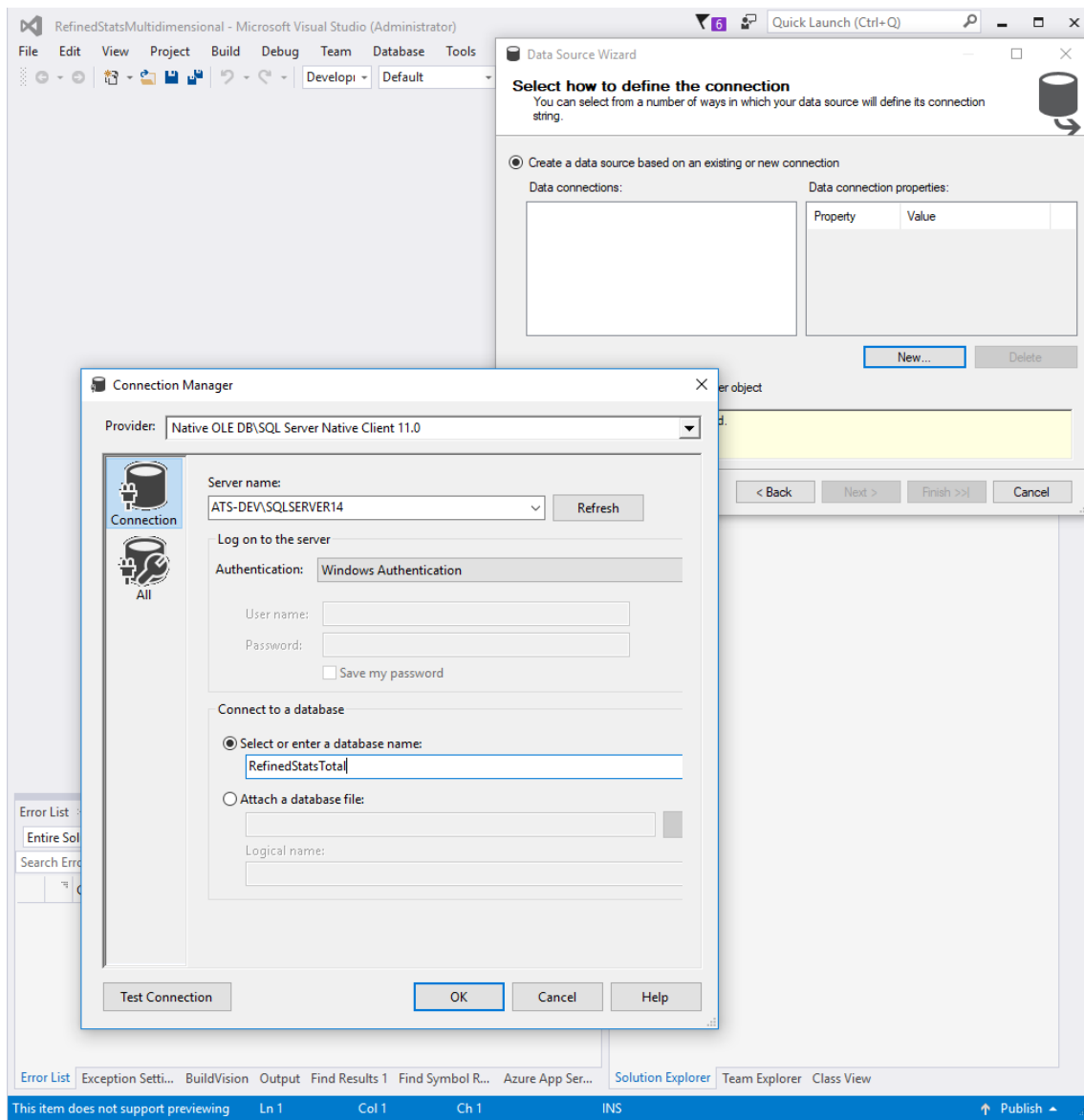
Εικόνα 27: Εισαγωγή δεδομένων στη βάση

Στη συνέχεια δημιουργείται ένα νέο έργο (κύβος OLAP) βάσει του προτύπου Analysis Services Multidimensional and Data Mining Project του MS Visual Studio (Εικόνα 28). Σε αυτό το έργο θα επιλεγούν οι διαστάσεις και τα μέτρα που απαιτούνται για τη ζητούμενη ανάλυση, καθώς και οι ιεραρχίες των ιδιοτήτων των διαστάσεων. Η διαδικασία έχει φτάσει πλέον στο επόμενο επίπεδο ενός συστήματος επιχειρηματικής ευφυΐας, αυτό του τελικού χρήστη.



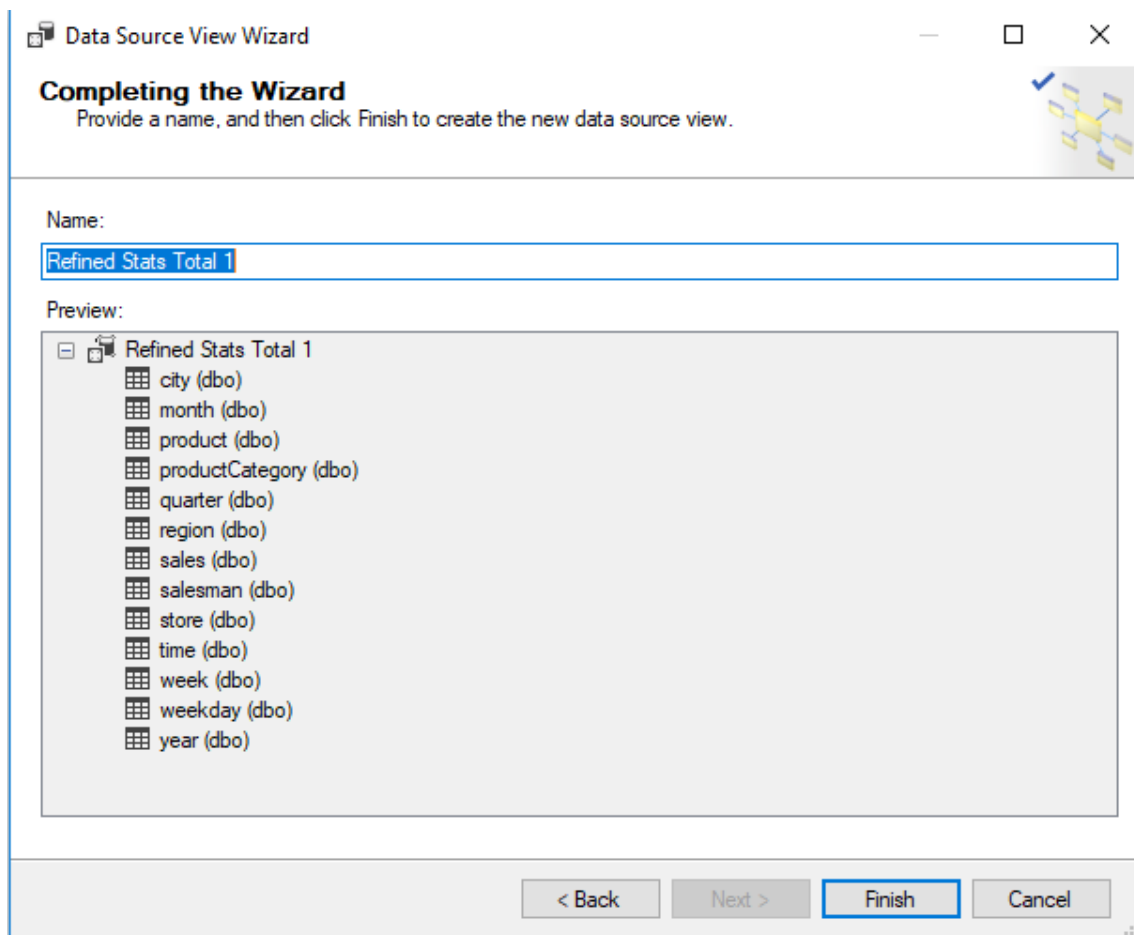
Εικόνα 28: Δημιουργία έργου στο Analysis Services

Αρχικά, ορίζεται η σύνδεση με τη βάση δεδομένων από την οποία θα αντλεί τα δεδομένα ο κύβος (Εικόνα 29).



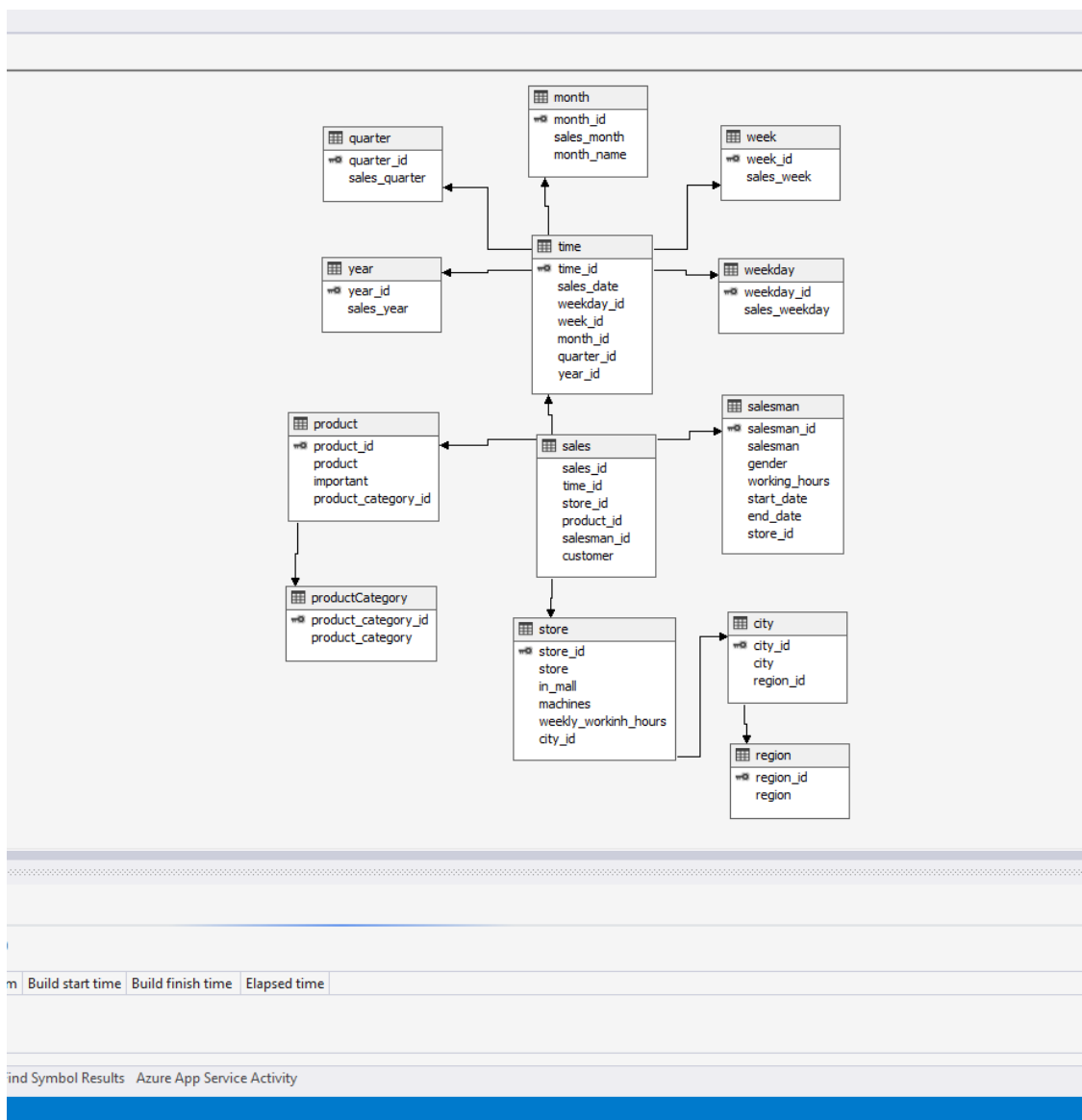
Εικόνα 29: Σύνδεση κύβου με βάση δεδομένων

Έπειτα, με τη χρήση του εργαλείου Data Source View Wizard επιλέγονται οι πίνακες της βάσης, που θα χρησιμοποιηθούν στο συγκεκριμένο έργο (Εικόνα 30).



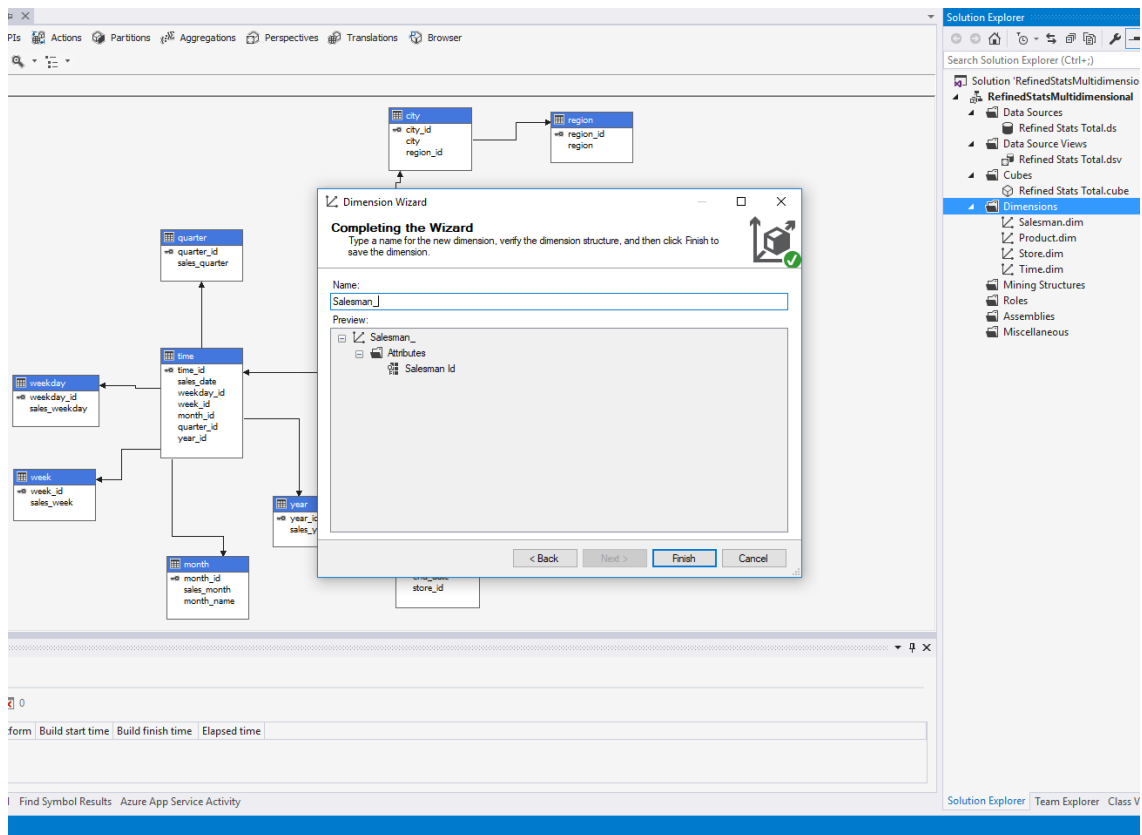
Εικόνα 30: Επιλογή πινάκων για χρήση στον κύβο

Μετά την επιλογή των επιθυμητών πινάκων, θα πρέπει να περιγραφεί η σχέση μεταξύ τους. Η σύνδεση μεταξύ των πινάκων παρουσιάζεται στην Εικόνα 31, όπου το βέλος ξεκινάει από τον πίνακα που περιέχει το ξένο κλειδί και καταλήγει στον πίνακα με το πρωτεύον κλειδί.



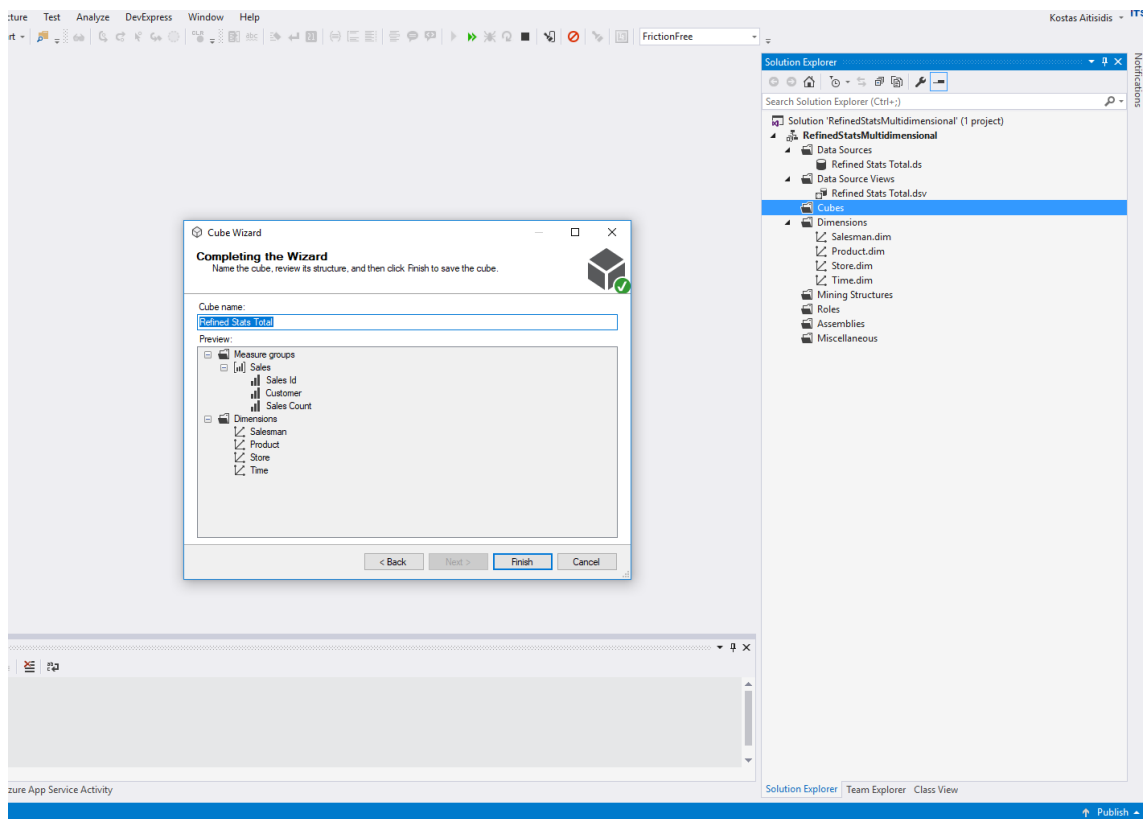
Εικόνα 31: Περιγραφή σύνδεσης μεταξύ των πινάκων

Η επιλογή των διαστάσεων γίνεται με την επιλογή του επιθυμητού πίνακα και των στηλών του, όπως φαίνεται στην Εικόνα 32.



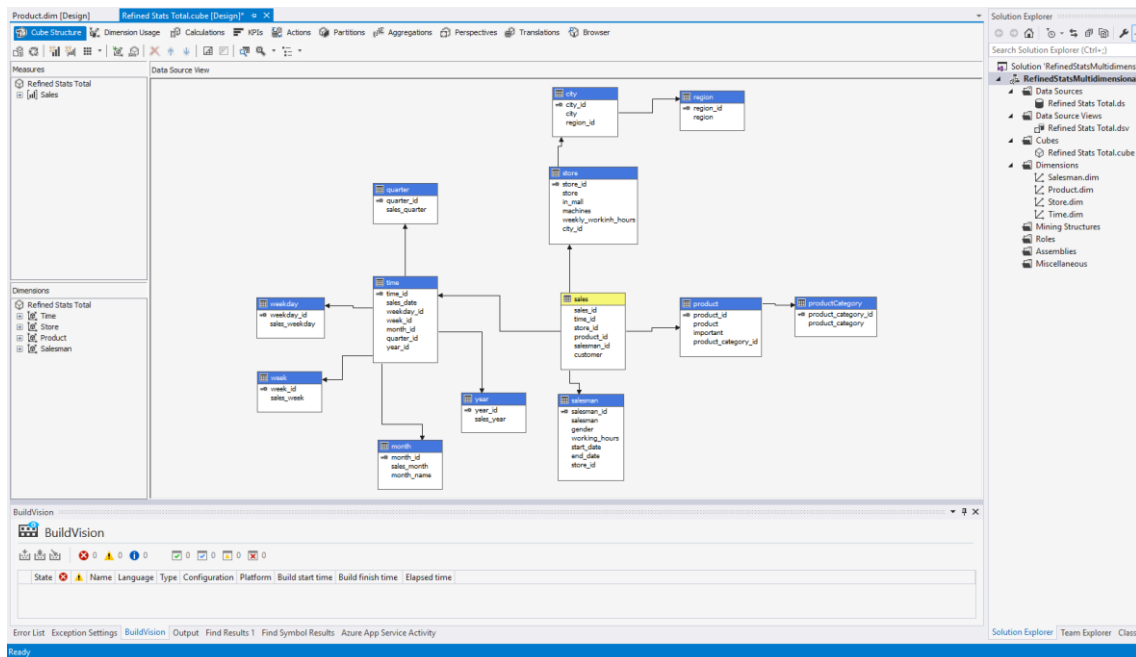
Εικόνα 32: Επιλογή διαστάσεων

Τέλος, σχεδιάζεται ο κύβος με την εισαγωγή των ζητούμενων διαστάσεων και μέτρων (Εικόνα 33).



Εικόνα 33: Εισαγωγή διαστάσεων και μέτρων στον κύβο

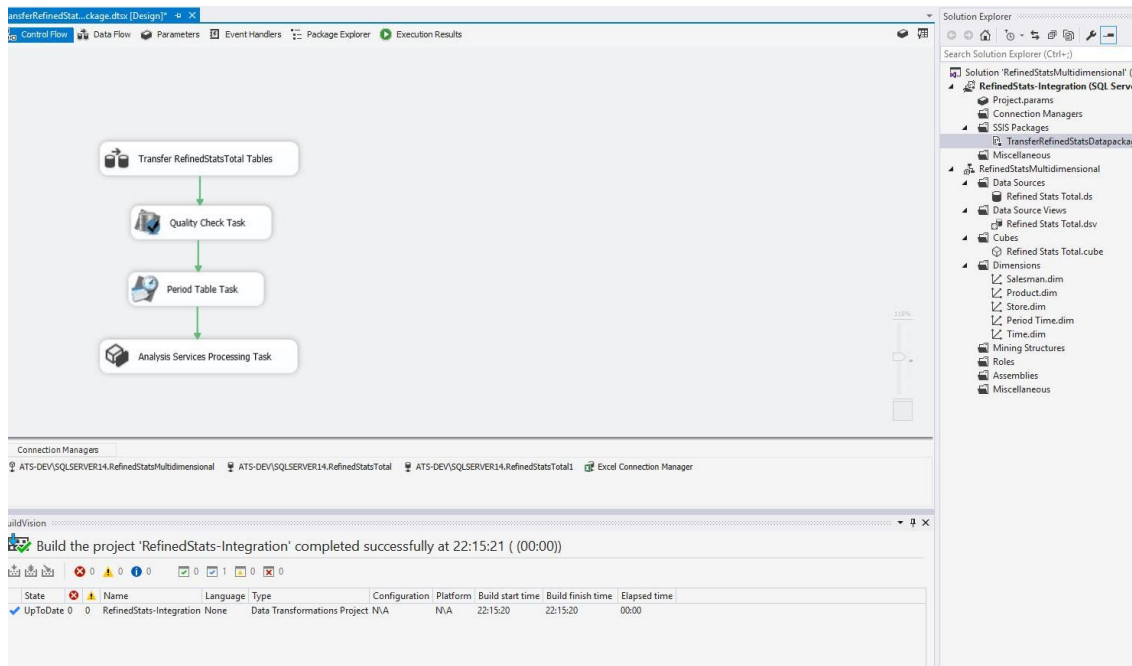
Το τελικώς διαμορφωμένο έργο παρουσιάζεται παρακάτω στην Εικόνα 34.



Εικόνα 34: Τελική διαμόρφωση κύβου

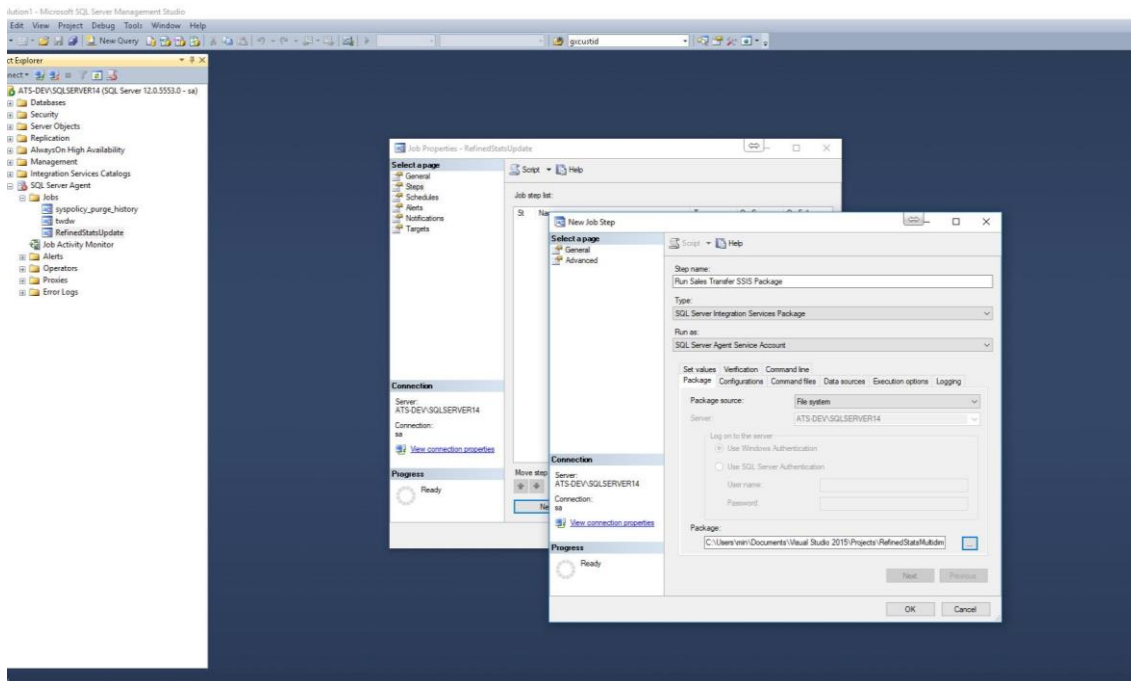


Για την αυτοματοποιημένη προσθήκη δεδομένων στη βάση και την ανανέωση του κύβου προστέθηκε το έργο RefinedStas-Integration (Εικόνα 35), όπου εισάγονται νέα δεδομένα από το αρχείο Excel, δεδομένα που έχουν εξαχθεί από το σύστημα συναλλαγών της επιχείρησης της προηγούμενης ημέρας. Στη συνέχεια πραγματοποιείται έλεγχος για την ακεραιότητα των δεδομένων, εισάγεται ο επεξεργασμένος πίνακας χρόνου και τέλος ανανεώνεται ο κύβος.

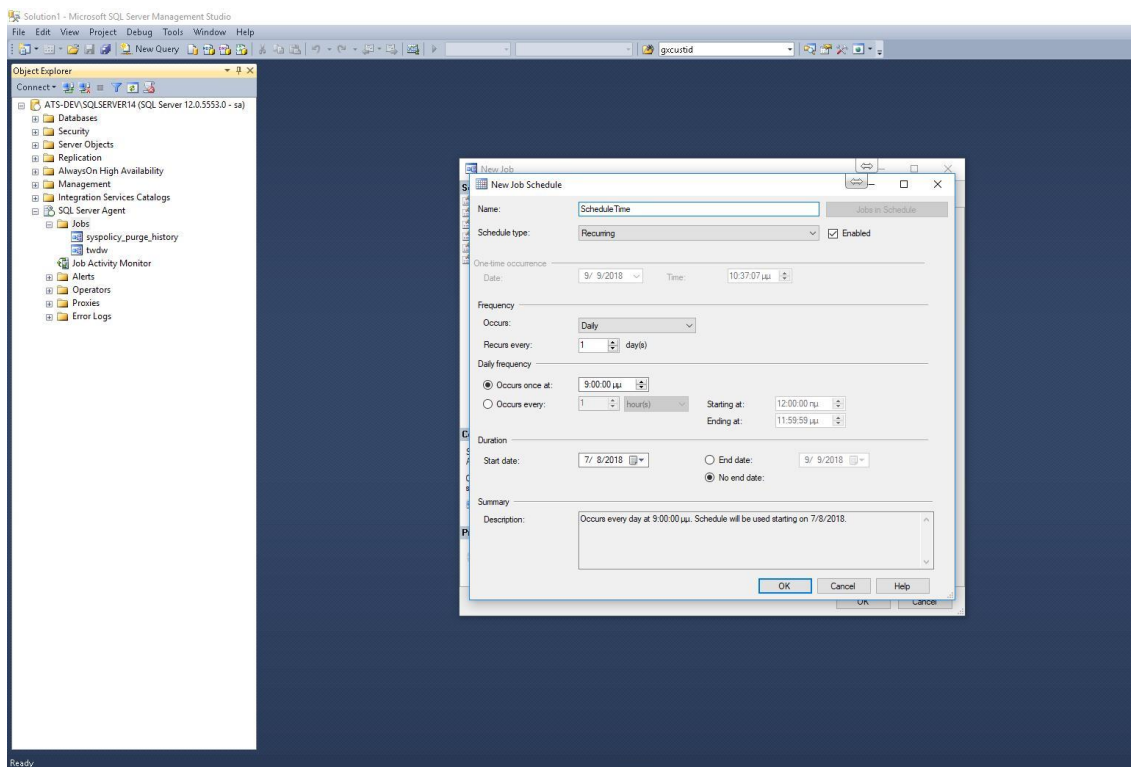


Εικόνα 35: Έργο για την ανανέωση των δεδομένων της βάσης και του κύβου

Στη συνέχεια το συγκεκριμένο έργο ανατίθεται σε μία εργασία του SQL Server Agent (Εικόνα 36) και ρυθμίζεται ώστε να επαναλαμβάνεται η συγκεκριμένη διεργασία κάθε μέρα στις 21:00 (Εικόνα 37).

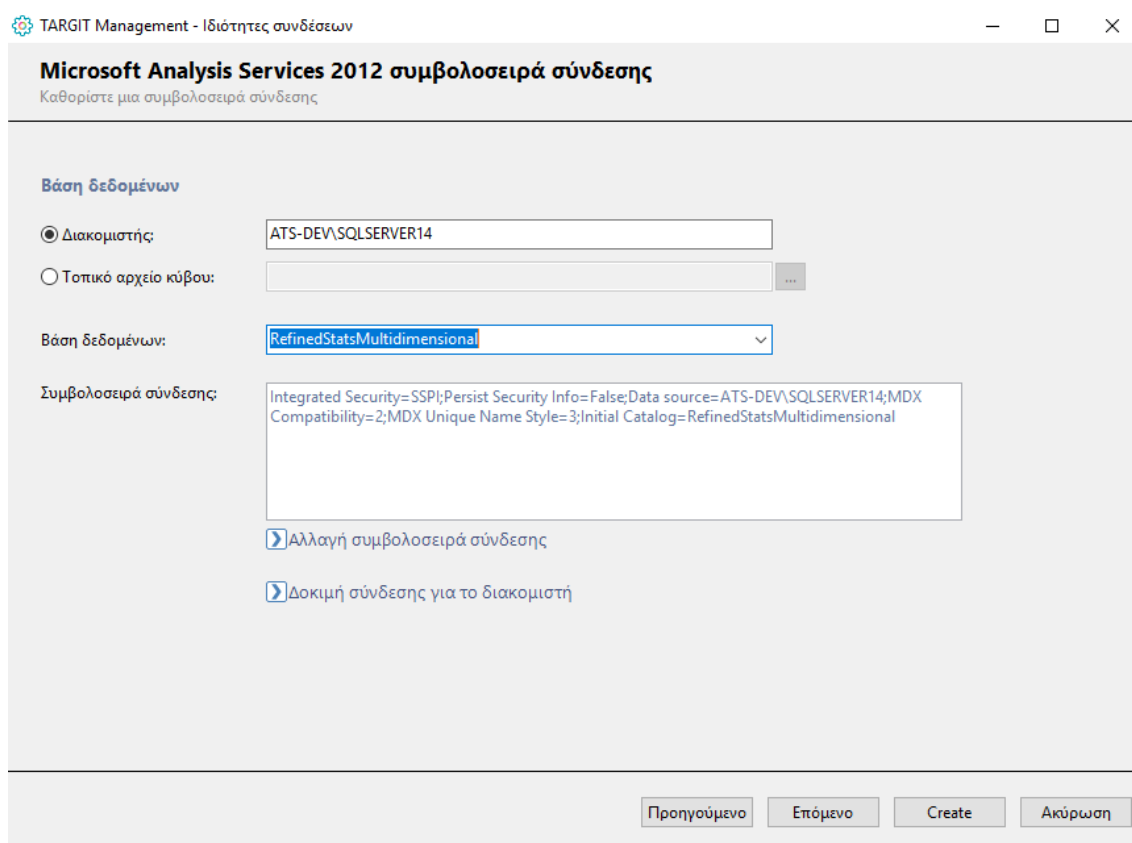


Εικόνα 36: Ανάθεση έργου σε εργασία του SQL Server Agent



Εικόνα 37: Επαναληψιμότητα έργου

Για την δημιουργία των επιθυμητών ερωτημάτων προς τη βάση και την οπτικοποίηση των απαντήσεων θα απαιτηθεί η δημιουργία μιας σύνδεσης μεταξύ της εφαρμογής TARGIT Decision Suite και του κύβου (Εικόνα 38). Με την επιτυχή δημιουργία της σύνδεσης η εφαρμογή είναι έτοιμη για την αναλυτική επεξεργασία δεδομένων και την παρουσίαση τους στην κατάλληλη και επιθυμητή μορφή.



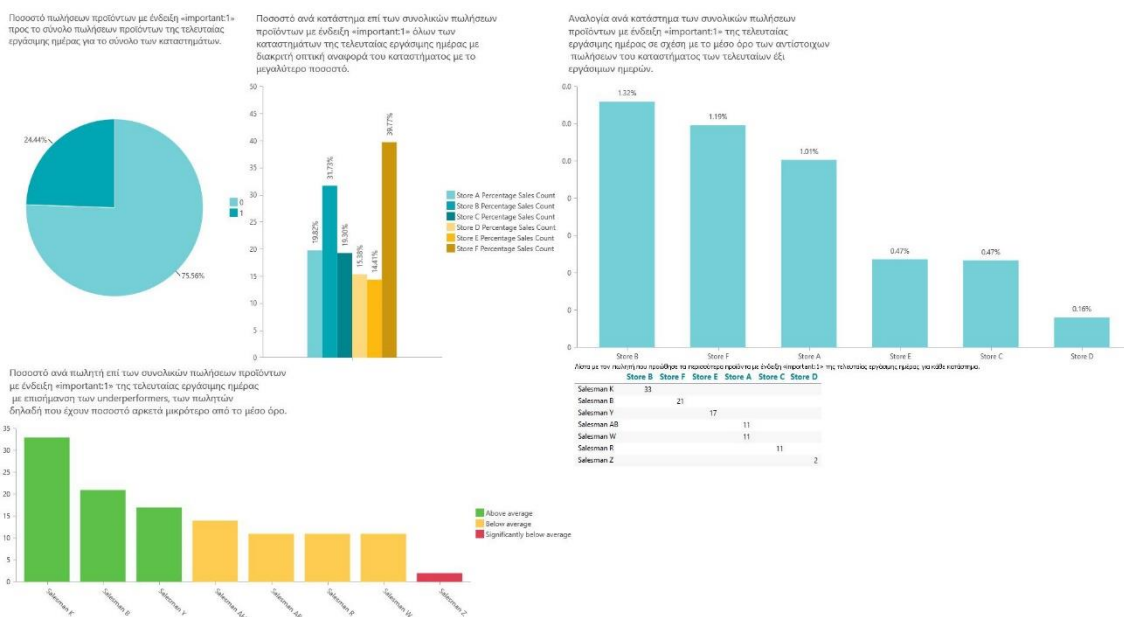
Εικόνα 38: Σύνδεση κύβου με TARGIT Decision Suite

Μέσω της εφαρμογής θα δημιουργηθούν τακτικές αναφορές και παραδείγματα έκτακτων ερωτημάτων. Στις τακτικές αναφορές περιλαμβάνεται ένα καθημερινό ψηφιακό ταμπλό (dashboard), μια εβδομαδιαία, μία μηνιαία και μια ετήσια αναφορά, ενώ οι δείκτες αυτών προέκυψαν κατόπιν συνεντεύξεων με στελέχη της εξεταζόμενης επιχείρησης. Οι έκτακτες αναφορές υλοποιούνται κατά ζήτηση και αφορούν κάθε φορά διαφορετικό θέμα. Στην παρούσα μελέτη θα εξετασθούν ορισμένα παραδείγματα.

## Ημερήσια αναφορά – Ψηφιακό Ταμπλό (Dashboard)

Στο Dashboard (Εικόνα 39) θα παρουσιασθούν οι εξής δείκτες:

- Ποσοστό πωλήσεων προϊόντων με ένδειξη «important:1» προς το σύνολο πωλήσεων προϊόντων της τελευταίας εργάσιμης ημέρας για το σύνολο των καταστημάτων.
- Ποσοστό ανά κατάστημα επί των συνολικών πωλήσεων προϊόντων με ένδειξη «important:1» όλων των καταστημάτων της τελευταίας εργάσιμης ημέρας με διακριτή οπτική αναφορά του καταστήματος με το μεγαλύτερο ποσοστό.
- Αναλογία ανά κατάστημα των συνολικών πωλήσεων προϊόντων με ένδειξη «important:1» της τελευταίας εργάσιμης ημέρας σε σχέση με το μέσο όρο των αντίστοιχων πωλήσεων του καταστήματος των τελευταίων έξι εργάσιμων ημερών.
- Ποσοστό ανά πωλητή επί των συνολικών πωλήσεων προϊόντων με ένδειξη «important:1» της τελευταίας εργάσιμης ημέρας με επισήμανση των underperformers, των πωλητών δηλαδή που έχουν ποσοστό αρκετά μικρότερο από το μέσο όρο.
- Λίστα με τον πωλητή που προώθησε τα περισσότερα προϊόντα με ένδειξη «important:1» της τελευταίας εργάσιμης ημέρας για κάθε κατάστημα.

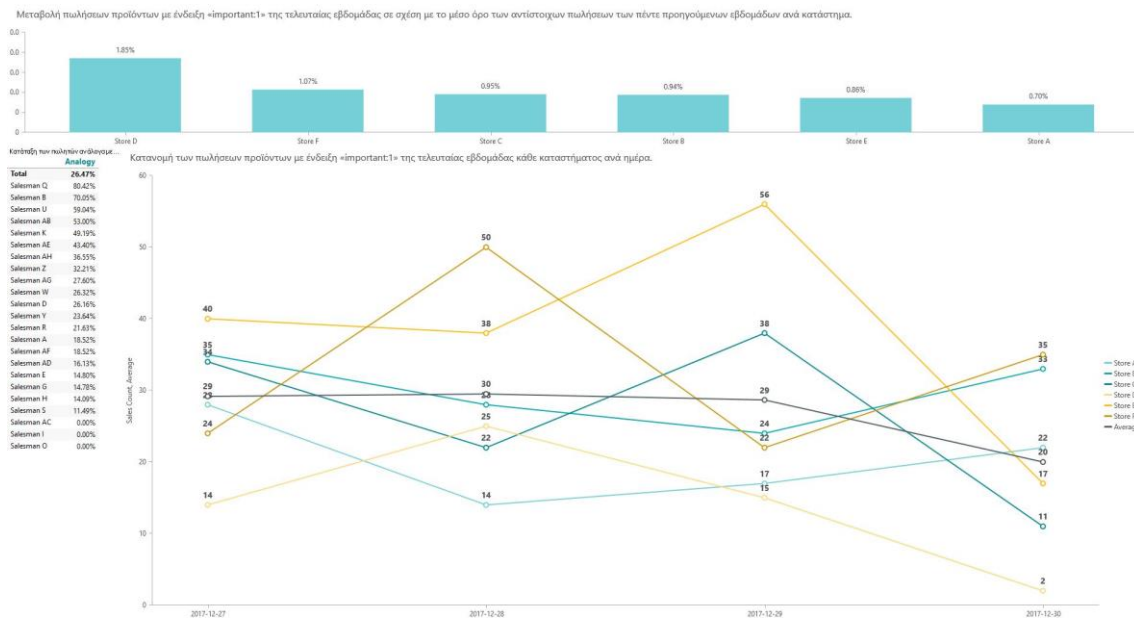


Εικόνα 39: Ημερήσια αναφορά

## Εβδομαδιαία αναφορά

Στην εβδομαδιαία αναφορά (Εικόνα 40) θα παρουσιασθούν οι εξής δείκτες:

- Μεταβολή πωλήσεων προϊόντων με ένδειξη «important:1» της τελευταίας εβδομάδας σε σχέση με το μέσο όρο των αντίστοιχων πωλήσεων των πέντε προηγούμενων εβδομάδων ανά κατάστημα.
- Κατανομή των πωλήσεων προϊόντων με ένδειξη «important:1» της τελευταίας εβδομάδας κάθε καταστήματος ανά ημέρα.
- Κατάταξη των πωλητών ανάλογα με τις πωλήσεις προϊόντων με ένδειξη «important:1» της τελευταίας εβδομάδας σε σχέση με τον εβδομαδιαίο μέσο όρο του καθενός των τελευταίων πέντε εβδομάδων.

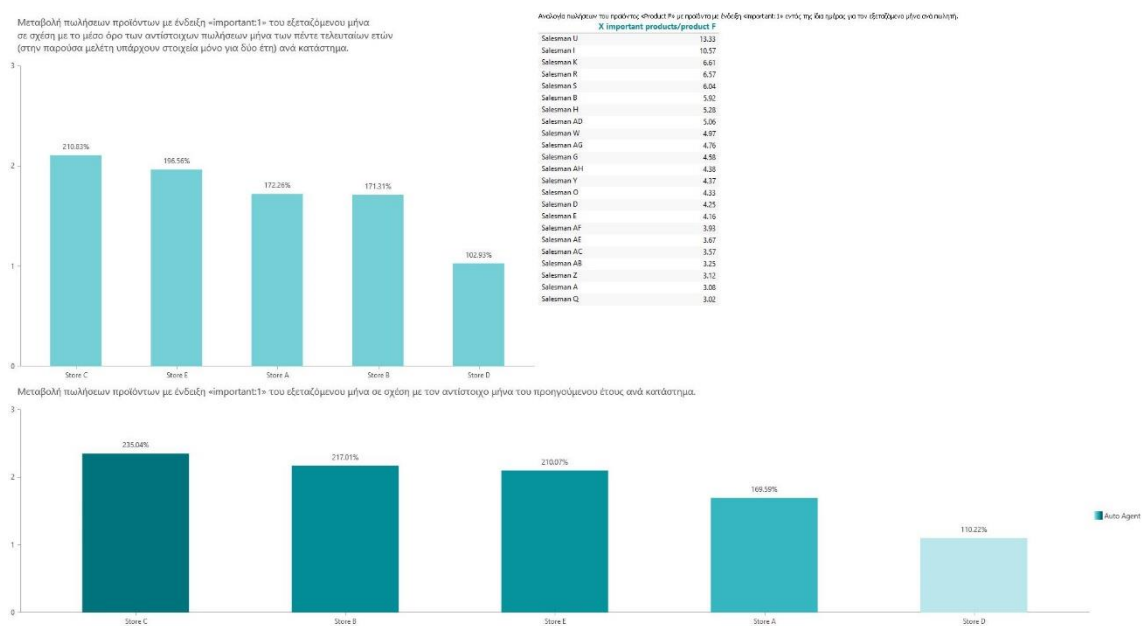


Εικόνα 40: Εβδομαδιαία αναφορά

## Μηνιαία αναφορά

Στη μηνιαία αναφορά () θα παρουσιασθούν οι εξής δείκτες:

- Μεταβολή πωλήσεων προϊόντων με ένδειξη «important:1» του εξεταζόμενου μήνα σε σχέση με το μέσο όρο των αντίστοιχων πωλήσεων μήνα των πέντε τελευταίων ετών (στην παρούσα μελέτη υπάρχουν στοιχεία μόνο για δύο έτη) ανά κατάστημα.
- Μεταβολή πωλήσεων προϊόντων με ένδειξη «important:1» του εξεταζόμενου μήνα σε σχέση με τον αντίστοιχο μήνα του προηγούμενου έτους ανά κατάστημα.
- Αναλογία πωλήσεων του προϊόντος «Product F» με προϊόντα με ένδειξη «important:1» εντός της ίδια ημέρας για τον εξεταζόμενο μήνα ανά πωλητή.

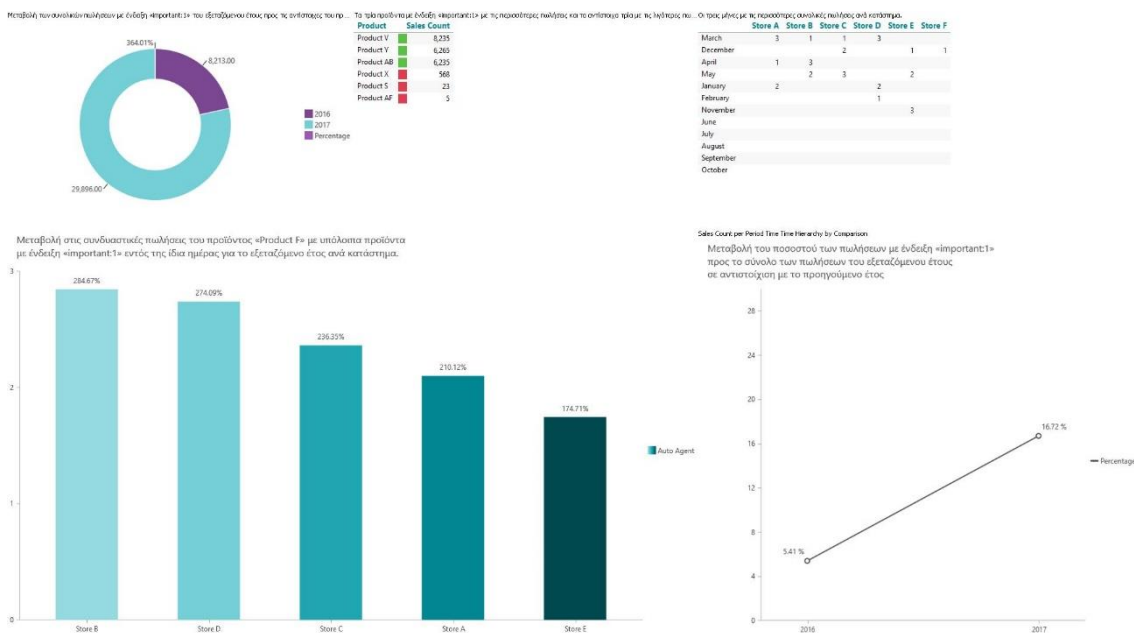


Εικόνα 41: Μηνιαία αναφορά

## Ετήσια αναφορά

Στην ετήσια αναφορά () θα παρουσιασθούν οι εξής δείκτες:

- Μεταβολή των συνολικών πωλήσεων με ένδειξη «important:1» του εξεταζόμενου έτους προς τις αντίστοιχες του προηγούμενου έτους.
- Μεταβολή του ποσοστού των πωλήσεων με ένδειξη «important:1» προς το σύνολο των πωλήσεων του εξεταζόμενου έτους σε αντιστοίχιση με το προηγούμενο έτος.
- Οι τρεις μήνες με τις περισσότερες συνολικές πωλήσεις ανά κατάσταση.
- Τα τρία προϊόντα με ένδειξη «important:1» με τις περισσότερες πωλήσεις και τα αντίστοιχα τρία με τις λιγότερες πωλήσεις.
- Μεταβολή στην αναλογία πωλήσεων του προϊόντος «Product F» με προϊόντα με ένδειξη «important:1» εντός της ίδια ημέρας για το εξεταζόμενο έτος ανά πωλητή σε αντιστοίχιση με το προηγούμενο έτος.



Εικόνα 42: Ετήσια αναφορά

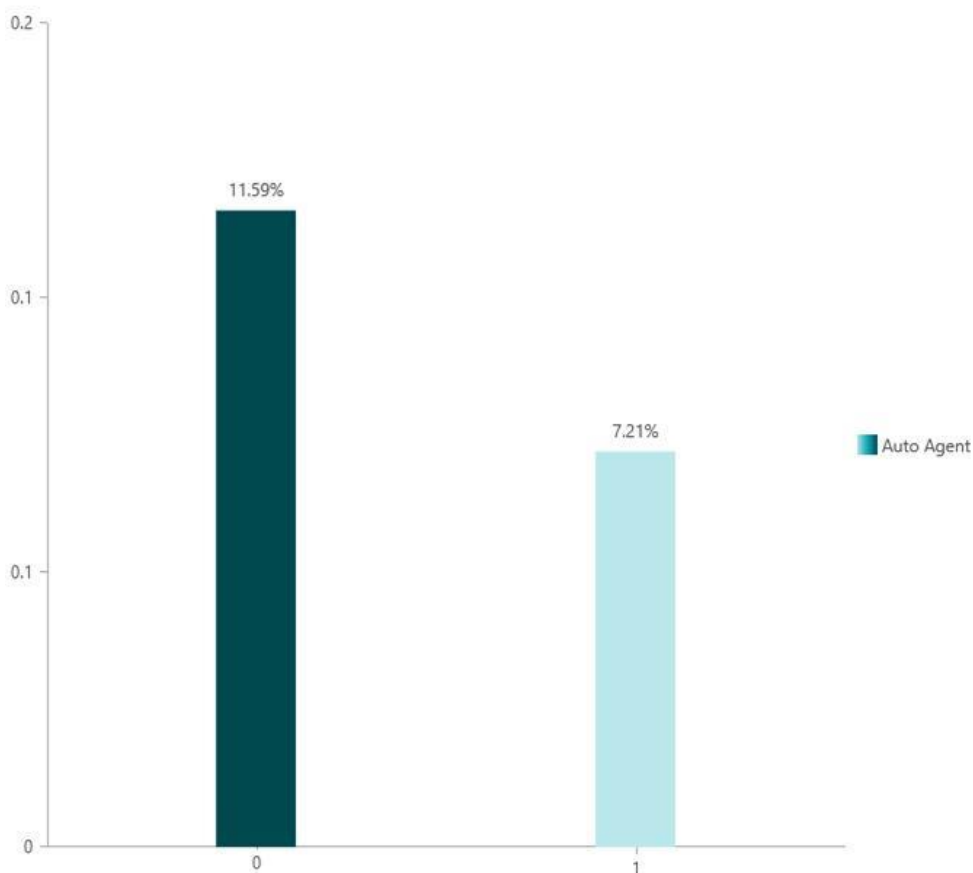
## Παραδείγματα ad-hoc αναφορών

Σε αρκετές περιπτώσεις, προκειμένου η επιχείρηση να λάβει μια απόφαση, θα χρειαστεί να αντλήσει πληροφόρηση για συγκεκριμένα θέματα και σημεία. Σε αυτές τις περιπτώσεις λοιπόν, ζητούνται από το σύστημα έκτακτες αναφορές, χωρίς προκαθορισμένη μορφή αλλά σύμφωνα με τις προκύπτουσες ανάγκες. Ενδεικτικά, αυτές οι αναφορές θα μπορούσαν να περιλαμβάνουν τους εξής δείκτες:

- Ποσοστό πωλήσεων προϊόντων με ένδειξη «important:1» προς το σύνολο των πωλήσεων ανάλογα με το φύλο του πωλητή (Εικόνα 43).
- Ποσοστό πωλήσεων προϊόντων με ένδειξη «important:1» προς το σύνολο των πωλήσεων ανάλογα με το συνολικό ωράριο του πωλητή (Εικόνα 44).
- Μεταβολή πωλήσεων προϊόντων με ένδειξη «important:1» σε σχέση με το διαχρονικό μέσο όρο για ιδιαίτερες περιόδους (Πάσχα, Χριστούγεννα, αρχή ακαδημαϊκού έτους κ.α.) (Εικόνα 45).

Sales Count per Salesman Gender by Comparison

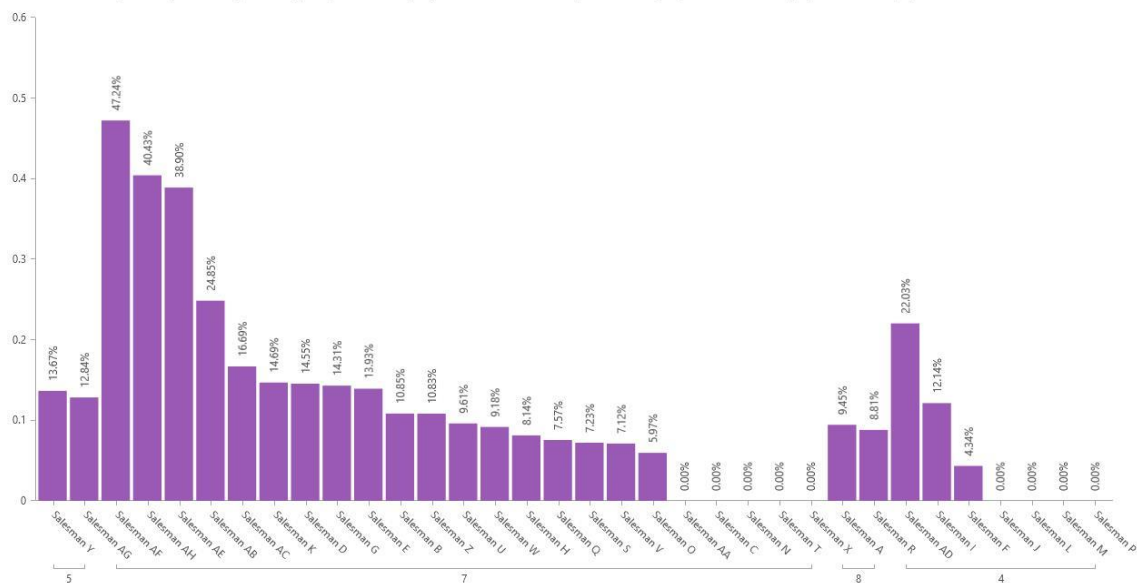
Ποσοστό πωλήσεων προϊόντων με ένδειξη «important:1» προς το σύνολο των πωλήσεων ανάλογα με το φύλο του πωλητή.



Εικόνα 43: Πωλήσεις σε σχέση με το φύλο του πωλητή (0: Γυναίκα, 1: Άντρας)

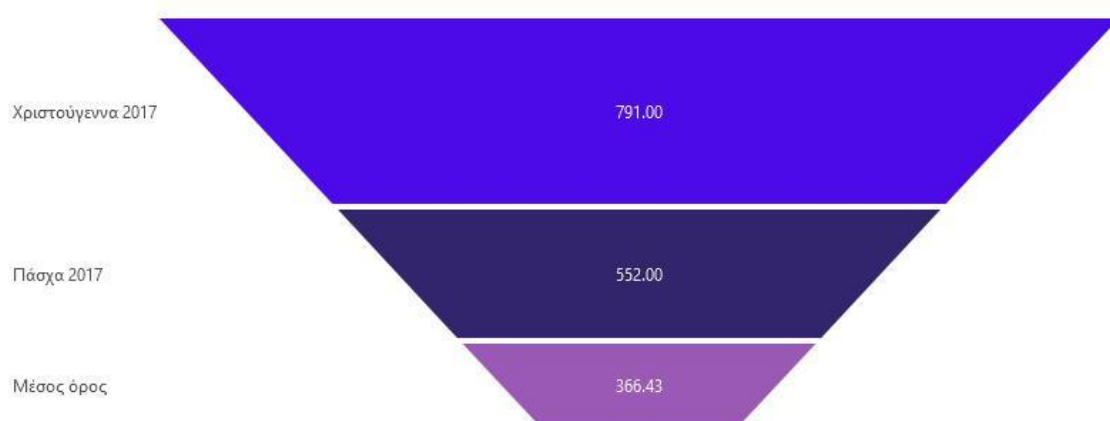


Ποσοστό πωλήσεων προϊόντων με ένδειξη «important:1» προς το σύνολο των πωλήσεων ανάλογα με το συνολικό ωράριο του πωλητή.



Εικόνα 44: Πωλήσεις ανάλογα με το ωράριο του πωλητή

Μεταβολή πωλήσεων προϊόντων με ένδειξη «important:1» σε σχέση με το διαχρονικό μέσο όρο για ιδιαίτερες περιόδους (Πάσχα, Χριστούγεννα, αρχή ακαδημαϊκού έτους κ.α.).



Εικόνα 45: Πωλήσεις σε ιδιαίτερες περιόδους

## 4. Ανάλυση Δεδομένων και Προτάσεις

Συγκρίνοντας την πληροφορία στη μορφή που ήταν αρχικά διαθέσιμη στην επιχείρηση με αυτή που προέκυψε μετά από όλες τις διεργασίες που ακολουθήθηκαν αλλά και από την αναλυτική διαδικασία, είναι κατανοητό ότι έχει προκύψει πολύτιμη γνώση, από την ερμηνεία των συγκεντρωτικών πινάκων και αναφορών. Ενώ οι προβλέψεις στις διακυμάνσεις των μεγεθών της επιχείρησης ήταν εν πολλοίς αυθαίρετες και βασισμένες σε εμπειρικά συστήματα, με μεγάλες πιθανότητες αποκλίσεων, πλέον μέσω του συγκεκριμένου μοντέλου μπορούν να κατηγοριοποιηθούν, να μετρηθούν και να αξιολογηθούν με πολύ μεγάλη ακρίβεια, έχοντας θετικό αντίκτυπο στους εξής τομείς:

- την αναλυτική πορεία των καταστημάτων
- την απόδοση των εργαζομένων
- την αποδοχή ή μη των προϊόντων από την πελατεία
- τις συνήθειες και προτιμήσεις των πελατών

Η γνώση αυτή μπορεί να χρησιμοποιηθεί στη λήψη κρίσιμων αποφάσεων σχετικά με τους προαναφερόμενους τομείς. Παρακάτω αναφέρονται ενδεικτικά ορισμένα θέματα, ανάλογα με τον τομέα ενδιαφέροντος :

### Προσωπικό

- εντοπισμός αδυναμιών υπαλλήλων σε συγκεκριμένο τομέα και προσπάθεια ενδυνάμωσής τους
- την αντικατάσταση ή πρόσληψη επιπλέον προσωπικού
- τη χρήση συστήματος επιβράβευσης του προσωπικού
- την επίδραση του χρόνου παραμονής ενός υπαλλήλου στην απόδοσή του (ωρίμανση)
- την έγκαιρη ανάγνωση προβλημάτων που μπορεί να παρουσιάζονται με έναν υπάλληλο, προκειμένου να υπάρξει άμεση αντιμετώπιση, προτού το πρόβλημα είναι μη αναστρέψιμο
- τη βελτίωση του προγραμματισμού αδειών του προσωπικού

## **Καταστήματα**

- την ενίσχυση καταστημάτων με εποχικό προσωπικό σε περιόδους με δυσανάλογα αυξημένη κίνηση
- την κυκλική μετακίνηση προσωπικού σε καταστήματα της ίδια περιοχής ανάλογα με την κίνηση που παρουσιάζεται
- τη δημιουργία ενός νέου καταστήματος σε περιοχή που ήδη υφίσταται κατάστημα ανάλογα με τις πωλήσεις στην εν λόγω περιοχή
- τη δημιουργία ή μη νέων καταστημάτων σε ανάλογες ή όμοιες γεωγραφικές περιοχές
- την έγκαιρη ανάγνωση προβλημάτων που μπορεί να παρουσιάζονται σε ένα κατάστημα, προκειμένου να υπάρξει άμεση αντιμετώπιση, προτού το πρόβλημα είναι μη αναστρέψιμο

## **Προϊόντα**

- την κατάργηση προϊόντων
- την ομαδοποίηση προϊόντων σε πακέτα για συνδυαστική χορήγηση

Επιπλέον αυτών, το σύστημα θα μπορούσε να απαντήσει και σε ιδιαίτερα σύνθετα ερωτήματα όπως το κατά πόσο επικερδής θα ήταν η εργασία την Κυριακή, εξεταζόμενη υπό το πρίσμα του συγκεκριμένου συστήματος, αν αξίζει δηλαδή η εργασία την Κυριακή ή απλά επιμερίζει τον όγκο εργασιών της εβδομάδας. Εκτός από τον κύκλο εργασιών, θα είναι δυνατόν να εξεταστεί η αποδοτικότητα και η αποτελεσματικότητα των καταστημάτων και των υπαλλήλων και πως επηρεάζεται η απόδοση των προηγούμενων και των επόμενων ημερών, η πελατειακή κίνηση και το τελικό αποτέλεσμα.

Η πρόταση προς την επιχείρηση προς άμεση υλοποίηση θα περιλαμβάνει την αγορά λογισμικού και μηχανολογικού εξοπλισμού προκειμένου να μπορέσει να εφαρμόσει το περιγραφόμενο σύστημα επιχειρηματικής ευφυΐας σε πλήρη διάσταση και για όγκο δεδομένων ανάλογο των εργασιών της. Φυσικά, η διεργασίες ETL θα πρέπει να εκτελούνται αυτοματοποιημένα ώστε να εμπλουτίζουν την αποθήκη δεδομένων χωρίς την απαίτηση ενεργειών από χρήστες. Επιπλέον, θα προταθεί η εκπαίδευση των

χρηστών του συστήματος αλλά και η μελλοντική υποστήριξη για ενδεχόμενη αναπροσαρμογή στις εκάστοτε ανάγκες της επιχείρησης, αναβαθμίσεις του συστήματος, εγγύηση διατήρησης ιδιωτικότητας απορρήτου κ.α.

Τέλος, θα προταθεί στη επιχείρηση να προβεί στις απαραίτητες ενέργειες ώστε να εμπλουτιστεί η βάση με περαιτέρω δεδομένα, τα οποία θα βοηθήσουν στη συνδυαστική ανάλυση. Ενδεικτικά, θα μπορούσε να γίνει:

- εμπλουτισμός της βάσης με στοιχεία για τις βάρδιες (πρωί/απόγευμα) ώστε να μπορέσει να επιτευχθεί η ανάδειξη καλύτερης απόδοσης ομάδας (συνδυασμός δύο ή περισσότερων ατόμων που έχουν καλύτερη απόδοση όταν εργάζονται μαζί) αλλά και ο ορθότερος διαμοιρασμός των υπαλλήλων σε αυτές
- προσθήκη στοιχείου επισκεψιμότητας σε ένα νέο πίνακα γεγονότων με καταχωρήσεις ανά ημέρα και ανά κατάσταση (μετατροπή του σχήματος της βάσης σε συνδυασμό χιονοφιλάδας και αστερισμού) ώστε να μπορούν να κανονικοποιούνται οι δείκτες πωλήσεων αλλά και ως ένα επιπλέον στοιχείο για το κάθε κατάσταση

## Κατάλογος Αναφορών και Παραπομπών

- California State University. (2018). *A case study for data warehousing courseware*. Retrieved from California State University's Web Location: <http://athena.ecs.csus.edu/~casestudy/report4.php> (Accessed on 22 August 2018)
- Codd, E. F., Codd, S. B., & Salley, C. T. (1993). *Providing OLAP (On-Line Analytical Processing) to User-Analysts: An IT Mandate*. Hyperion white papers.
- element61 N.V. – Moore Stephens . (2018). *Best Practice ETL Architecture*. Retrieved from element61: Moore Stephens Belgium Web Location: <https://www.element61.be/en/competence/best-practice-etl-architecture> (Accessed on 24/08/2018)
- Ferrari, A. (2011). Business Intelligence Systems, Uncertainty in Decision-Making and Effectiveness of Organizational Coordination. In A. Carugati, & C. Rossignoli, *Emerging Themes in Information Systems and Organization Studies*. Physica-Verlag HD.
- Forrester. (2018). *Business Intelligence*. Ανάκτηση από Forrester's Web Location: <https://www.forrester.com/Business-Intelligence> (Accessed on 20 August 2018)
- Gray, P., & Watson, H. (1998, June). Present and future directions in data warehousing. *ACM SIGMIS Database*. 29, pp. 83-90. 10.1145/313310.313345.
- Imhoff, C., Gallemmo, N., & Geiger, J. G. (2003). *Mastering Data Warehouse Design: Relational and Dimensional Techniques*. Indianapolis: Wiley Publishing, Inc.
- Inmon, W. H. (2005). *Building the Data Warehouse, Fourth Edition*. Indianapolis: Wiley Publishing, Inc.
- Kimball, R., & Ross, M. (2013). *The Data Warehouse Toolkit: The Definitive Guide to Dimensional Modeling, 3rd Edition*. Indianapolis: John Wiley & Sons, Inc.
- Kour, A. (2015). Data Warehousing, Data Mining, OLAP and OLTP Technologies Are Indispensable Elements to Support Decision-Making Process in Industrial World. *International Journal of Scientific and Research Publications, Volume 5, Issue 5*.
- Lih Ong, I., Hwa Siew, P., & Wong, S. (2011). A Five-Layered Business Intelligence Architecture. *Communications of the IBIMA*, p. DOI: 10.5171/2011.695619.
- Loudcher, S., Jakawat, W., Pavel Soriano Morales, E., & Favre, C. (2015). Combining OLAP and Information Networks for Bibliographic Data Analysis: A Survey. *Scientometrics 103*, DOI: 10.1007/s11192-015-1539-0.
- Moody, D. L., & Kortink, M. A. (2000). From enterprise models to dimensional models: a methodology for data warehouse and data mart design. *Proceedings of the Second Intl. Workshop on Design and Management of Data Warehouses*, (p. 5). Stockholm.

- OLAP.com. (2018). *What is Business Intelligence (BI)?* Retrieved from OLAP's Web Location: <http://olap.com/learn-bi-olap/olap-bi-definitions/business-intelligence/> (Accessed on 18 August 2018)
- Pendse, N. (2002, April 7). *What is OLAP?* Retrieved from DSSResources Web Location: <http://dssresources.com/subscriber/password/papers/features/pendse04072002.htm> (Accessed on 20 August 2018)
- Raskino, M. (2017, April 13). *Highlights of the 2017 CEO Survey: CIOs Must Scale Up Digital Business.* Retrieved from Gartner's Web Location: <https://www.gartner.com/doc/3678617?refval=&pcp=mpe#611978406> (Accessed on 19 August 2018)
- Saranya, V. (2013). *OLAP.* Retrieved from LinkedIn SlideShare: <https://www.slideshare.net/ersaranya/> (Accessed on 22 August 2010)
- Theodoratos, D., & Sellis, T. (1998). Data Warehouse Schema and Instance Design. *Conceptual Modeling – ER '98. ER 1998. Lecture Notes in Computer Science, vol 1507* (pp. 363–376). Springer, Berlin, Heidelberg: T.W. Ling, S. Ram, and M.L. Lee.
- Thomsen, E. (2002). *OLAP Solutions: Building Multidimensional Information Systems, SE.* John Wiley & Sons, Inc.
- van der Aalst, W. (2013). Process Cubes: Slicing, Dicing, Rolling Up and Drilling Down Event Data for Process Mining. In M. Song, M. Wynn, & J. Liu, *Asia Pacific Business Process Management: First Asia Pacific Conference, AP-BPM 2013, Lecture Notes in Business Information Processing.* Springer, Cham.
- Vassiliadis, P. (2009). A Survey of Extract–Transform–Load Technology. *International Journal of Data Warehousing and Mining (IJDWM)*, σ. DOI: 10.4018/jdwm.2009070101.
- Κύρκος, Ε. Γ. (2015). *Επιχειρηματική Ευφυΐα και Εξόρυξη Δεδομένων.* Αθήνα: Σύνδεσμος Ελληνικών Ακαδημαϊκών Βιβλιοθηκών ([www.kallipos.gr](http://www.kallipos.gr)).
- Συμεωνίδης, Π., & Γούναρης, Α. (2015). *Βάσεις, Αποθήκες και Εξόρυξη Δεδομένων με τον SQL Server: Εργαστηριακός Οδηγός.* Αθήνα: Σύνδεσμος Ελληνικών Ακαδημαϊκών Βιβλιοθηκών ([www.kallipos.gr](http://www.kallipos.gr)).