

ΤΕΧΝΟΛΟΓΙΚΟ ΕΚΠΑΙΔΕΥΤΙΚΟ ΙΔΡΥΜΑ ΔΥΤΙΚΗΣ ΕΛΛΑΔΑΣ
ΣΧΟΛΗ ΔΙΟΙΚΗΣΗΣ ΚΑΙ ΟΙΚΟΝΟΜΙΑΣ
ΤΜΗΜΑ ΔΙΟΙΚΗΣΗΣ ΕΠΙΧΕΙΡΗΣΕΩΝ



ΤΙΤΛΟΣ ΠΤΥΧΙΑΚΗΣ ΕΡΓΑΣΙΑΣ
**ΑΝΑΛΥΣΗ ΔΕΔΟΜΕΝΩΝ ΤΗΣ ΕΚΠΑΙΔΕΥΣΗΣ
ΜΕ ΕΡΓΑΛΕΙΑ OLAP**

Πτυχιακή Εργασία του : Μητρούλιας Χρήστος

Αρχικός Επιβλέπων : Παπαϊωάννου Βάϊος

Τρέχον Επιβλέπων : Καρούσος Νικόλαος

ΠΑΤΡΑ,2014

Περίληψη

Στη παρούσα πτυχιακή εργασία γίνεται προσπάθεια να ελεγχθεί η αποτελεσματικότητα των εργαλείων αναλυτικής επεξεργασίας OLAP στο πλαίσιο της εκπαιδευτικής διαδικασίας. Εξετάζονται δεδομένα παρμένα από το χώρο της εκπαίδευσης, με σκοπό να συλλεχθούν και να επεξεργαστούν τα δεδομένα αποδοτικότερα, δίνοντας ποιοτικά αποτελέσματα για τον οργανισμό της εκπαίδευσης και χρήσιμα αποτελέσματα μέσω των εργαλείων της πολυδιάστατης ανάλυσης ώστε να αξιολογηθούν και να αξιοποιηθούν κατάλληλα.

Αρχικά αναλύονται βασικοί όροι που αφορούν την πληροφορία παγκοσμίως, τα συστήματα διαχείρισης της αλλά και τις βάσεις δεδομένων οι οποίες αποτελούν βασικό αντικείμενο εργασίας. Τα ιδιαίτερα χαρακτηριστικά των μοντέλων βάσεων δεδομένων και ο τρόπος λειτουργίας τους γίνονται σαφή στο πρώτο κεφάλαιο της εργασίας. Επίσης γίνεται αναφορά στα συστήματα διαχείρισης βάσεων δεδομένων (ΣΔΒΔ) και στα χαρακτηριστικά τους. Αυτή η ανάλυση θα βοηθήσει στην καλύτερη κατανόηση του τρόπου λειτουργίας των βάσεων αυτών τόσο στο πρακτικό μέρος όσο και στη δημιουργία υπερκύβων.

Στο δεύτερο κεφάλαιο της εργασίας γίνεται λόγος για τα συστήματα αναλυτικής επεξεργασίας (OLAP) και του τρόπου λειτουργίας τους. Πιο συγκεκριμένα γίνεται λόγος για τα πλεονεκτήματα της OLAP σε σχέση με παλαιότερου τύπου βάσεων δεδομένων, για τις Αποθήκες Δεδομένων πάνω στις οποίες στηρίζεται η OLAP και πως μπορεί να δημιουργηθεί μια πολυδιάστατη βάση δεδομένων. Στη συνέχεια γίνεται λόγος για την Αρχιτεκτονική των κύβων OLAP αλλά και εκτενής ανάλυση για τις λειτουργίες του και τις πράξεις που γίνονται σε αυτούς.

Στο τρίτο και τελευταίο μέρος της εργασίας αναφέρεται το πρακτικό μέρος όπου εξετάζεται στη πράξη η εφαρμογή της τεχνολογίας OLAP με δεδομένα από το χώρο της εκπαίδευσης ενώ εξάγονται συμπεράσματα χρήσιμα τα οποία προέκυψαν από την ανάλυση που έγινε σε διάφορους τομείς της.

Περιεχόμενα

Περίληψη	2
Εισαγωγή	4
1. Συστήματα Διαχείρισης Δεδομένων	5
1.1 Γενικά Εισαγωγικά.....	5
1.1.1 Ανάπτυξη Δεδομένων σε Παγκόσμια Κλίμακα.....	5
1.1.2 Big Data	7
1.2 Συστήματα Βάσεων Δεδομένων	8
1.2.1 Ιεραρχικό Μοντέλο Δεδομένων (Hierarchical Data Model)	9
1.2.2 Δικτυωτό Μοντέλο Δεδομένων (Network Data Model)	11
1.2.3 Σχισιακό Μοντέλο Δεδομένων (Relational Data Model)	12
1.2.4 Αντικειμενοστραφές Μοντέλο Δεδομένων (Object Oriented Data Model)	14
1.3 Στόχοι του Σχισιακού Μοντέλου και Δυσκολίες που Αντιμετωπίζονται.....	15
2. Συστήματα Αναλυτικής Επεξεργασίας (OLAP).....	17
2.1 Εισαγωγή.....	17
2.2 Πλεονεκτήματα OLAP.....	18
2.2.1 Πλεονεκτήματα της OLAP Έναντι των Σχισιακών Βάσεων	19
2.3 Αποθήκες Δεδομένων	20
2.3.1 Αρχιτεκτονική Αποθήκης Δεδομένων	20
2.4 Εισαγωγή στην OLAP.....	23
2.5 Βασικές Έννοιες OLAP	24
2.5.1 Αρχιτεκτονική των Συστημάτων OLAP	27
2.6 Μοντέλα Συστημάτων Αναλυτικής Επεξεργασίας	28
2.6.1 Πράξεις στους Υπερκύβους.....	29
2.6.2 Είδη Εργαλείων OLAP	31
3. Κατασκευή Κύβων OLAP για την Εκπαίδευση	33
3.1 Σχεδιασμός και Ανάπτυξη Μοντέλου OLAP.....	33
3.1.1 Μελέτη Περίπτωσης	33
3.1.2 Η Βάση Δεδομένων	34
3.1.3 Μοντέλο OLAP	34
3.2 Σκοπός του Λογισμικού	35
3.3 Τρόπος Λειτουργίας του Λογισμικού	37
3.3.1 Εισαγωγή Δεδομένων της Εκπαίδευσης στο OLAP Access Manager ...	37
3.3.2 Επεξεργασία Δεδομένων και Κατασκευή Κύβων	37
3.4 Παραδείγματα Κύβων OLAP	41
3.4.1 Δημιουργία Ερωτημάτων.....	41
3.4.2 Συμπεράσματα Ερωτημάτων	64
4. Συμπεράσματα	68
Βιβλιογραφία	70

Εισαγωγή

Στη σημερινή εποχή η ανάγκη για την χρήση της πληροφορίας είναι μεγαλύτερη από ποτέ. Η διατήρηση βάσεων δεδομένων όχι μόνο άλλαξε τον τρόπο διαχείρισης και εκτέλεσης πολλών εργασιών αλλά βοήθησε στο να γίνουν αυτές πιο εύκολες. Έτσι γίνεται εύκολα αντιληπτό ότι οι βάσεις δεδομένων είναι φυσιολογικό να χρησιμοποιούνται τόσο για την καταγραφή όσο και για την εκμαίευση πληροφορίας σε κάθε περίπτωση όπου αυτές οι δυο περιπτώσεις συναντώνται. Αυτό έχει ως συνέπεια να εξελιχτούν τα εργαλεία ανάλυσης των βάσεων δεδομένων ώστε να αυξηθούν οι δυνατότητες του χρήστη και να γίνει πιο εύκολο το έργο του. Μια βάση δεδομένων δίνει τη δυνατότητα στον χρήστη με την χρήση των κατάλληλων εργαλείων ανάλυσης να εξάγει συμπεράσματα τα οποία θα τον βοηθήσουν στη λήψη καλύτερων αποφάσεων.

Τα διαθέσιμα εργαλεία για τον χρήστη είναι ποικίλα και αυτό σημαίνει πως είναι δυνατή η μετατροπή και γενικότερα η τροποποίηση των δεδομένων με τέτοιο τρόπο ώστε να μπορούν πέρα από την προφανή πληροφόρηση να δίνουν και πληροφορίες με ένα πιο πολυσύνθετο τρόπο, όπως για παράδειγμα με την πολυδιάστατη ανάλυση δεδομένων.

Στη Σύγχρονη Αναλυτική Επεξεργασία Δεδομένων (OLAP) πιο συγκεκριμένα, συναντάται αυτή η τάση πιο έντονα καθώς στηρίζεται στην ανάγνωση της πληροφορίας με πολυδιάστατο τρόπο. Αυτού του είδους η επεξεργασία είναι προσανατολισμένη στα δεδομένα. Σημειώνεται ότι οι κύβι δεδομένων δεν είναι σωστό να αντιμετωπίζονται σαν απλά σύνολα δεδομένων αλλά σαν ερωτήσεις που παράγονται πάνω στο σύνολο των δεδομένων (Βασιλειάδης Π., Σκιαδόπουλος Σ., Σέλλης Τ., 2002).

Μέσα από την εργασία αυτή δημιουργούνται κύβι OLAP χρησιμοποιώντας τα διαθέσιμα δεδομένα εκπαίδευσης και εξάγονται χρήσιμα αποτελέσματα και συμπεράσματα που αφορούν την εκπαιδευτική διαδικασία. Εφαρμόζοντας την ανάλυση αυτή στο σύνολο της εκπαιδευτικής διαδικασίας, μπορεί μελλοντικά να αποδώσει πολλαπλά οφέλη τόσο στο επίπεδο της περιφερειακής εκπαίδευσης όσο και σε πανελλαδικό επίπεδο. Επίσης μπορεί να φανερώσει τα σημεία που χρήζουν βελτίωσης αλλά και να ελεγχθεί η αποτελεσματικότητα των μέτρων που εφαρμόζονται απλά με την διατήρηση και εκμετάλλευση μιας βάσης δεδομένων για το εκπαιδευτικό σύστημα.

1. Συστήματα Διαχείρισης Δεδομένων

Στο κεφάλαιο αυτό παρουσιάζονται τα Συστήματα Διαχείρισης Δεδομένων και ο λόγος για τον οποίο προέκυψαν αυτά. Πιο συγκεκριμένα γίνεται αναφορά στα δεδομένα σε παγκόσμια κλίμακα, στα Συστήματα Διαχείρισης Βάσεων Δεδομένων και στα σημαντικότερα Μοντέλα Δεδομένων με κυρίαρχο το Σχεσιακό Μοντέλο.

1.1 Γενικά Εισαγωγικά

Κάθε σύστημα διαχείρισης δεδομένων προέκυψε από την ανάγκη επίλυσης ενός συγκεκριμένου τύπου προβλήματος διαχείρισης δεδομένων. Όταν η σχεσιακή βάση δεδομένων εισήχθη στην αγορά, χρειάστηκε ένα σύνολο από εργαλεία ώστε να επιτρέψουν στους διευθυντές να μελετήσουν την σχέση μεταξύ στοιχείων δεδομένων. Όταν οι επιχειρήσεις άρχισαν να αποθηκεύουν αδόμητα δεδομένα, οι αναλυτές χρειάζονταν νέα εργαλεία, όπως φυσικές γλώσσες βασισμένες σε εργαλεία ανάλυσης για να αποκτήσουν γνώσεις που θα ήταν χρήσιμες στις επιχειρήσεις.

Για παράδειγμα, μια εταιρία μηχανών αναζήτησης έχει τεράστιες ποσότητες δεδομένων που μπορεί να αξιοποιήσει έτσι ώστε να έχει κέρδος. Για να αποκτήσουν αξία τα δεδομένα απαιτούνται νέες προσεγγίσεις και κατάλληλα εργαλεία για την αξιοποίηση αυτών.

Τα συστήματα διαχείρισης δεδομένων αναπτύχθηκαν τις τελευταίες πέντε δεκαετίες. Η τελευταία δεκαετία έχει χαρακτηριστεί ως η έναρξη της εποχής των δεδομένων πολύ μεγάλης κλάσης (Big Data).

1.1.1 Ανάπτυξη Δεδομένων σε Παγκόσμια Κλίμακα

Όταν η πληροφορική συνδέθηκε με την εμπορική αγορά στα τέλη του 1960, τα στοιχεία ήταν αποθηκευμένα σε flat αρχεία που δεν επέβαλαν κάποια δομή. Όταν οι εταιρείες απαιτούσαν να φτάσουν σε ένα άλλο πιο λεπτομερή επίπεδο δεδομένων για την κατανόηση των πελατών, έπρεπε να εφαρμοστούν διάφοροι μέθοδοι, οι οποίες συμπεριλάμβαναν πολύ λεπτομερή μοντελοποίηση των δεδομένων ώστε να μπορούν να είναι πιο αξιοποιήσιμα.

Στην δεκαετία του 1970, εφευρέθηκαν τα σχεσιακά μοντέλα δεδομένων και το σύστημα διαχείρισης σχεσιακών βάσεων δεδομένων (relational database management system (RDBMS))(Bakuya T., Matsui M.,1997), που επιβάλλει την

δομή και μια μέθοδο για την αύξηση της απόδοσης. Στο σχεσιακό μοντέλο προστίθεται ένα επίπεδο αφαίρεσης, μια δομημένη γλώσσα ερωτοαπαντήσεων γνωστή ως SQL (Stonebraker M., 2010), στην οποία μπορούν να παραχθούν αναφορές και άλλα εργαλεία διαχείρισης δεδομένων. Με αυτό τον τρόπο γίνεται ευκολότερο για τους προγραμματιστές να ικανοποιήσουν τις αυξανόμενες απαιτήσεις των επιχειρήσεων με την εξαγωγή χρήσιμων δεδομένων και συμπερασμάτων.

Το σχεσιακό μοντέλο προσφέρει ένα μεγάλο αριθμό εργαλείων όπου θα βοηθήσει τις εταιρίες στην καλύτερη οργάνωση των δεδομένων τους και θα είναι σε θέση να συγκρίνουν τις συναλλαγές από μια γεωγραφική θέση σε μία άλλη. Όμως προέκυψε το πρόβλημα αποθήκευσης του μεγάλου όγκου δεδομένου όπου ανέβαζε το κόστος και η πρόσβαση στα δεδομένα ήταν αργή.

Για να επιλυθεί αυτό, βρέθηκε ένα νέο σύνολο τεχνολογιών για να υποστηρίξει το σχεσιακό μοντέλο, το μοντέλο οντοτήτων- συσχετίσεων (Entity-Relationship (ER)), (Pin P., Chen S., 1976). Αυτό το μοντέλο είναι αφαιρετικό έτσι ώστε να αυξάνεται η χρηστικότητα των δεδομένων. Σε αυτό το μοντέλο, κάθε στοιχείο ορίστηκε ανεξάρτητα από την χρήση του. Παρόλα αυτά, οι προγραμματιστές θα μπορούσαν να δημιουργήσουν νέες σχέσεις μεταξύ των πηγών δεδομένων χωρίς πολύπλοκους προγραμματισμούς. Είναι σημαντικό για την διαχείριση των δεδομένων να υπάρχει συναλλαγή υψηλά δομημένων δεδομένων.

Όταν ο όγκος των δεδομένων που οι οργανισμοί χρειάζονται για την διαχείριση αυξήθηκε πάρα πολύ, οι αποθήκες δεδομένων έδωσαν την δυνατότητα στην οργάνωση (information technology-IT) να επιλέξει ένα υποσύνολο των δεδομένων που είναι αποθηκευμένο. Με αυτό τον τρόπο είναι ευκολότερο για τις επιχειρήσεις να ανακτήσουν γνώση από τα δεδομένα, περισσότερο εστιασμένη σε συγκεκριμένες περιοχές. Επιπλέον, οι αποθήκες συχνά αποθηκεύουν δεδομένα από προηγούμενα έτη για τη κατανόηση, την οργανωτική απόδοση, τον εντοπισμό των τάσεων και βοηθούν να εκθέσουν τα πρότυπα της συμπεριφοράς. Επίσης, παρέχει μια ολοκληρωμένη πηγή πληροφόρησης, σε διάφορες πηγές δεδομένων που θα μπορούσε να χρησιμοποιηθεί για ανάλυση.

Από την δεκαετία του 1990 μέχρι και σήμερα, το περιεχόμενο των συστημάτων διαχείρισης δεδομένων και οι αποθήκες δεδομένων είναι σε θέση να επωφεληθούν από τις βελτιώσεις στην επεκτασιμότητα του λογισμικού και του υλικού εξοπλισμού.

Πριν από λίγα χρόνια, οι οργανισμοί βρισκόντουσαν σε κίνδυνο, αποθηκεύοντας στιγμιότυπα ή μικρά σύνολα από σημαντικές πληροφορίες, γιατί το κόστος να αποθηκευτούν και να επεξεργαστούν ήταν απαγορευτικό, ώστε να αναλύσουν ότι είχε αποθηκευτεί.

Τα τελευταία χρόνια, με την ανάπτυξη νέων τεχνολογιών (κοινωνικά δίκτυα, smartphones κ.α.), ο όγκος και η μορφή των δεδομένων μεγάλωσε πολύ, ενώ οι δυνατότητες ανάλυσης και επεξεργασίας αυτών είναι εντυπωσιακές. Δημοσιεύτηκε στο Wired Magazine (2008) ότι « Η αναζήτηση της γνώσης άρχισε με μεγάλες θεωρίες. Πλέον ξεκινά με τεράστιους όγκους δεδομένων».

1.1.2 Big Data

Με τα δεδομένα πολύ μεγάλης κλάσης (Big Data) (Manyika J. et al., 2011), δίνεται η δυνατότητα επεξεργασίας και οπτικοποίησης των δεδομένων, έτσι ώστε να μπορούν να αποθηκεύονται αποτελεσματικά, χρησιμοποιώντας την αποθήκη η οποία είναι βασισμένη κυρίως σε cloud (cloud-based storage) (C. Wang et al., 2009) πιο αποδοτικά. Επιπλέον, με τις βελτιώσεις στην ταχύτητα και την αξιοπιστία του δικτύου είναι σε θέση να διαχειρίζονται τεράστιες ποσότητες δεδομένων σε ένα αποδεκτό ρυθμό.

Τα μεγάλα δεδομένα σχετίζονται με ένα σύνολο διαδικασιών που απαιτούν συνδυαστικές γνώσεις των πεδίων της Πληροφορικής, Στατιστικής, Μηχανικής εκμάθησης και Διοίκησης επιχειρήσεων. Τα μεγάλα δεδομένα χαρακτηρίζονται από τα 3 V, όπου είναι ο όγκος (Volume), η ποικιλία (Variety), και η ταχύτητα (Velocity) (Madden S., 2012).

Ο όγκος δεδομένων χαρακτηρίζει τις μοντέρνες εφαρμογές, γιατί τα δεδομένα που εμφανίζονται είναι ποικίλα (σχεσιακά δεδομένα, εικόνα, ήχος, video), ενώ η παραγωγή των δεδομένων γίνεται σε πολύ μεγάλους ρυθμούς ποικίλων δεδομένων και η εξαγωγή συμπερασμάτων πρέπει να γίνει σε πραγματικό χρόνο που συνδυάζει την ποικιλία με την ταχύτητα των δεδομένων.

Για την ανάγκη των Μεγάλων Δεδομένων αναπτύχθηκε μια γενιά συστημάτων, μοντέλων και προγραμματιστικών εργαλείων, όπως τα Map Reduce (Dean J., Ghemawat S., 2008), Hadoop (White T., 2009), Big Table (Dean J. et al., 2008) και NoSQL (Cattell R., 2010) κ.α. Τα παραπάνω επιτρέπουν την παράλληλη επεξεργασία δεδομένων σε μεγάλη κλίμακα, ενώ ταυτόχρονα αξιοποιούν τα

δεδομένα για την παραγωγή αναλύσεων (analytics) χρησιμοποιώντας τον συνδυασμό της Τεχνητής Νοημοσύνης και Τεχνικών Μηχανικής μάθησης, καθώς και την Επιχειρησιακή Έρευνα και την Διοίκηση Επιχειρήσεων.

1.2 Συστήματα Βάσεων Δεδομένων

Βασικό ρόλο για τα Μεγάλα δεδομένα έπαιξε η σχεσιακή βάση δεδομένων η οποία χρησιμοποιείται έως τώρα, για όλους τους τύπους δεδομένων. Ένα Σύστημα Βάσης Δεδομένων αναλαμβάνει αποκλειστικά την αποθήκευση, προστασία και επεξεργασία των δεδομένων, προσφέροντας σύγχρονους μηχανισμούς πρόσβασης, δικαιωμάτων και ενημέρωσης των δεδομένων. Οι Βάσεις Δεδομένων (ΒΔ) και τα συστήματα που τις διαχειρίζονται (ΣΔΒΔ) δημιουργήθηκαν με σκοπό να εξυπηρετήσουν τις συνεχώς αυξανόμενες ανάγκες αποθήκευσης και επεξεργασίας μεγάλου όγκου δεδομένων.

Η Βάση Δεδομένων αποτελεί μία συλλογή στοιχείων σχετικών μεταξύ τους, τα οποία είναι δομημένα και καταχωρημένα με κατάλληλο τρόπο. Για τη διαχείριση της όπως αναφέραμε χρησιμοποιούνται ειδικά συστήματα τα οποία ονομάζονται Συστήματα Διαχείρισης Βάσεων Δεδομένων (ΣΔΒΔ). Το Σύστημα Διαχείρισης Βάσεων Δεδομένων είναι ένα λογισμικό σύστημα το οποίο πραγματοποιεί όλες τις λειτουργίες που πρέπει να υποστηριχθούν. Αυτές οι λειτουργίες είναι η αναζήτηση, η εισαγωγή, η διαγραφή, ο συγχρονισμός προσπελάσεων, η προστασία και πολλές ακόμη. Το Σύστημα Διαχείρισης Βάσεων Δεδομένων είναι σε θέση να διαχειρίζεται πολλές Βάσεις Δεδομένων ταυτόχρονα, αναλόγως με τις απαιτήσεις των χρηστών.

Τα Συστήματα Βάσεων Δεδομένων υποστηρίζουν εφαρμογές που απαιτούν τη διαχείριση πολλών τμημάτων πληροφορίας και την επεξεργασία πολλών ερωτημάτων ή ενημερώσεων (Μανωλόπουλος Ι., Παπαδόπουλος Α., 2001). Το βασικό τους χαρακτηριστικό είναι η άμεση ταυτόχρονη διαχείριση πολλών συναλλαγών. Για το λόγο αυτό, οι εφαρμογές αυτές καλούνται εφαρμογές Άμεσης Επεξεργασίας Συναλλαγών (On-line Transaction Processing, OLTP). Χαρακτηριστικό παράδειγμα τέτοιων εφαρμογών είναι η διαχείριση δεδομένων σε μεγάλες εταιρείες και οργανισμούς.

Η καλή οργάνωση των δεδομένων μιας μεγάλης εταιρείας ή ενός οργανισμού παίζουν μεγάλο ρόλο στην αποτελεσματική λειτουργία τους, διότι διευκολύνονται και επιταχύνονται οι εσωτερικές τους διεργασίες, οι οποίες πολλές

φορές είναι πολυσύνθετες και χρονοβόρες. Η μεθοδική οργάνωση των δεδομένων έχει επίδραση σε όλα τα τμήματα της εταιρείας, του εκπαιδευτικού οργανισμού ή του εκάστοτε οργανισμού. Για παράδειγμα, τα δεδομένα ενός σχολείου συνήθως σχετίζονται με τα προσωπικά στοιχεία των μαθητών και των καθηγητών, όπως ονοματεπώνυμο, διεύθυνση, θέση, βαθμοί. Η δομημένη οργάνωση της πληροφορίας επιτρέπει τη γρήγορη αναζήτηση στοιχείων που σχετίζονται με οτιδήποτε αναζητεί ο χρήστης.

Οι εντολές της γλώσσας ορισμού δεδομένων μπορούν να περιγράψουν πλήρως το σχήμα μιας βάσης δεδομένων. Ωστόσο, για τη περιγραφή του σχήματος της βάσης δεδομένων χρειάζονται εναλλακτικοί τρόποι, διότι η γλώσσα ορισμού δεδομένων βρίσκεται σε αρκετά χαμηλό επίπεδο και περιέχει αρκετές λεπτομέρειες. Για το λόγο αυτό χρησιμοποιούνται μοντέλα δεδομένων τα οποία περιγράφουν το σχήμα της βάσης δεδομένων σε αρκετά υψηλό επίπεδο (δηλαδή χωρίς λεπτομέρειες υλοποίησης). Ένα μοντέλο δεδομένων αποτελεί μια αναπαράσταση του πραγματικού κόσμου και είναι εστιασμένο στις οντότητες και τις μεταξύ τους σχέσεις που ενδιαφέρουν στις εφαρμογές. Τονίζεται ότι λάθη ή παραλήψεις στο μοντέλο δεδομένων έχουν άμεσο αντίκτυπο τόσο στα αποθηκευμένα δεδομένα όσο και στις λειτουργίες επεξεργασίας δεδομένων. Η αλλαγή του μοντέλου δεδομένων συνεπάγεται αλλαγές σε πολλά τμήματα της βάσης δεδομένων. Υπάρχουν διάφορα είδη μοντέλων δεδομένων, όπως το σχεσιακό μοντέλο, το ιεραρχικό μοντέλο, το δικτυωτό μοντέλο και το αντικειμενοστραφή μοντέλο που θα αναλυθούν στις παρακάτω ενότητες.

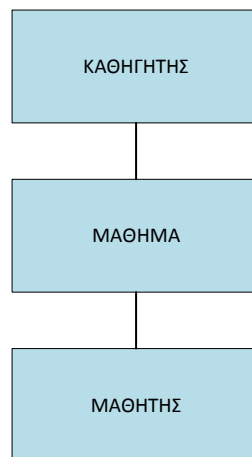
1.2.1 Ιεραρχικό Μοντέλο Δεδομένων (Hierarchical Data Model)

Στις Ιεραρχικές (Hierarchical) βάσεις δεδομένων (Tsichritzis D.C., Lochovsky F.H., 1976) τα δεδομένα αναπαρίστανται με δενδρικής μορφής δομές δεδομένων και συνδέονται μεταξύ τους με συνδέσμους (links). Η κάθε εγγραφή μπορεί να συνδέεται προς τα πάνω μόνο με μία άλλη εγγραφή (γονέας), ενώ μπορεί να έχει έως δύο εγγραφές που να εξαρτώνται απ' αυτήν (παιδιά). Υπάρχει μία μόνο εγγραφή ρίζα (root), απ' την οποία εξαρτώνται όλες οι άλλες εγγραφές της βάσης δεδομένων.

Οι κανόνες ενός ιεραρχικού μοντέλου είναι οι εξής: (i) έχουν ένα και μοναδικό τμήμα, την ρίζα, (ii) η ρίζα μπορεί να έχει πολλούς κόμβους-παιδιά, (iii)

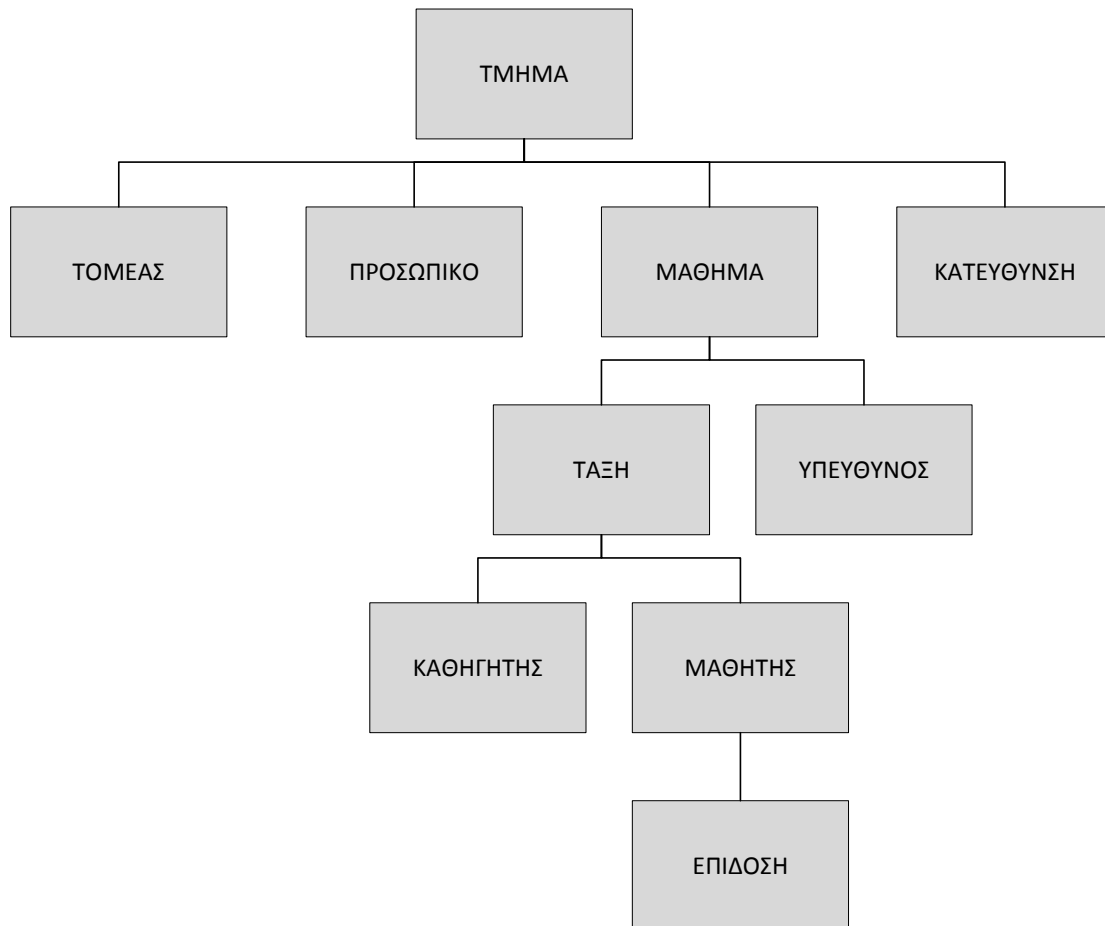
κάθε κόμβος παιδί της ρίζας μπορεί να έχει πολλά τμήματα παιδιά, (iv) δεν επιτρέπεται να έχουμε ένα στιγμιότυπο για κάποιο τμήμα-παιδί, χωρίς να το έχουμε αντιστοιχήσει σε κάποιο στιγμιότυπο ενός τμήματος-πατέρα.

Το ιεραρχικό μοντέλο έχει το μειονέκτημα ότι είναι πολύπλοκο στην επεξεργασία των εγγραφών του (προσθήκη, διαγραφή, τροποποίηση). Επίσης δεν υπάρχει κάποια κατεύθυνση στα βέλη στα παρακάτω σχήματα, αφού η επεξεργασία κινείται από το πατέρα προς το παιδί και δεν μπορεί να πάει αντίστροφα. Έτσι το μοντέλο δεν μπορεί να απαντήσει σε ερωτήσεις όπως «ποιος είναι ο διδακτικός φόρτος του προσωπικού (εύρεση για κάποιον καθηγητή των μαθημάτων, πόσα μαθήματα διδάσκει ο κάθε καθηγητής είτε σαν εκπαιδευτικός είτε σαν υπεύθυνος του μαθήματος), ποιο το ωρολόγιο πρόγραμμα κάθε μαθητή.



Σχήμα 1.1 : Ιεραρχικό μοντέλο για ένα εκπαιδευτικό οργανισμό.

Κάθε κόμβος σχηματίζει μια οντότητα, που είναι μια αναπαράσταση κάποιας αυτόνομης ύπαρξης με υλική ή θεωρητική υπόσταση. Για παράδειγμα, οντότητα μπορεί να είναι ο μαθητής (ένας άνθρωπος με ονοματεπώνυμο, χαρακτηριστικά κ.α.), αλλά και το μάθημα στο σχολείο (Σχήμα 1.1).



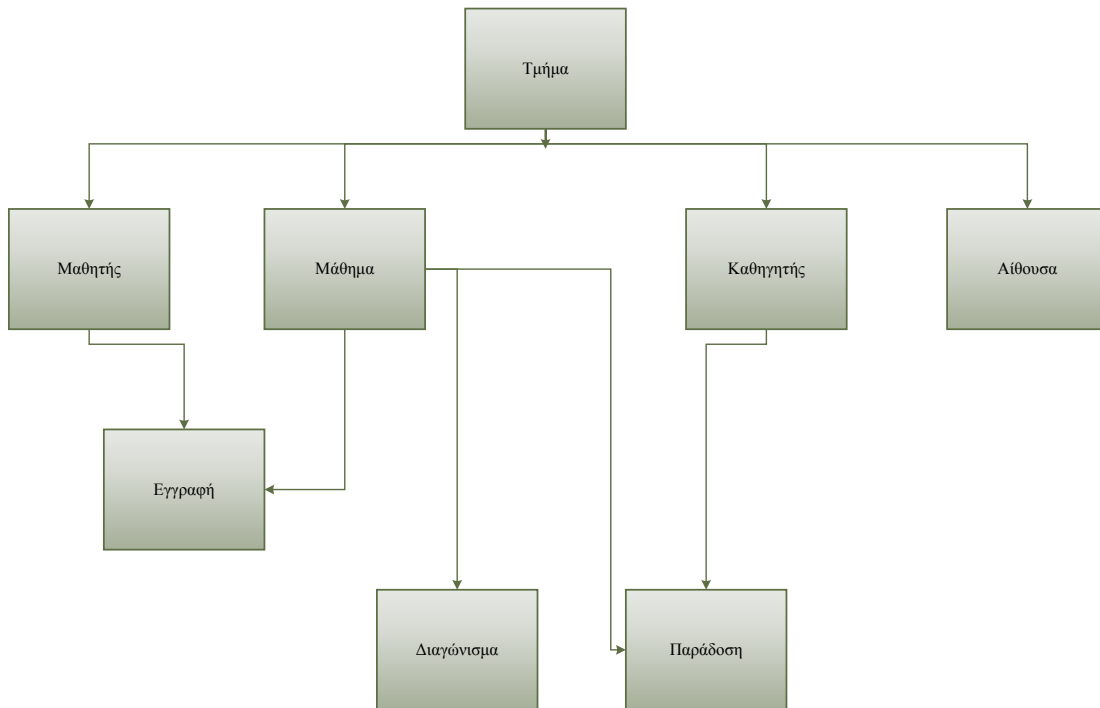
Σχήμα 1.2 : Αναπαράσταση ενός μεγάλου εκπαιδευτικού οργανισμού.

Το πρόβλημα που παρουσιάζεται στο ερώτημα «Ποιο είναι το ωρολόγιο πρόγραμμα ενός μαθητή» στο σχήμα 1.2 εμφανίζεται η απάντηση του ερωτήματος, όπου η απάντηση βρίσκεται μέσα στην βάση της οντότητας μαθητή. Επομένως αν το όνομα του μαθητή εμφανιστεί σε 8 μαθήματα δεν μπορούμε να διακρίνουμε με κάποιο τρόπο αν ο μαθητής είναι ο ίδιος ή κάποιος που έχει το ίδιο όνομα.

1.2.2 Δικτυωτό Μοντέλο Δεδομένων (Network Data Model)

Οι Δικτυωτές βάσεις είναι μια παραλλαγή του ιεραρχικού μοντέλου, όπου αναπαριστά δεδομένα σε μορφή δικτύου. Στις Δικτυωτές (Network) βάσεις δεδομένων (Johnson H.R., Larson J.A. and Lawrence J.D., 1979) τα δεδομένα αναπαριστούνται με μορφές από δομές δένδρων και συνδέονται μεταξύ τους με συνδέσμους (links), όπως ακριβώς και στις ιεραρχικές βάσεις δεδομένων, με τη διαφορά ότι μια εγγραφή μπορεί να συνδέεται προς τα πάνω με περισσότερες από μία πατρικές εγγραφές (parent records). Είναι πιο δύσχρηστες αλλά πιο γρήγορες από τις ιεραρχικές βάσεις δεδομένων. Πλεονέκτημα αυτού του μοντέλου είναι ότι

μειώνει το πλεονασμό της πληροφορίας, όμως υπάρχει δυσκολία κατά την συντήρηση και στην αλλαγή δεδομένων, για παράδειγμα σε ποιο τμήμα και τι μαθήματα έγιναν (Σχήμα 1.3).



Σχήμα 1.3 : Αναπαράσταση Δικτυωτού Μοντέλου Βάσεων Δεδομένων

1.2.3 Σχεσιακό Μοντέλο Δεδομένων (Relational Data Model)

Τα Σχεσιακά Συστήματα Διαχείρισης Βάσεων Δεδομένων αποτελούν την πλειοψηφία των σύγχρονων συστημάτων διαχείρισης βάσεων δεδομένων (ΣΔΒΔ). Τα συστήματα αυτά στηρίζονται στο Σχεσιακό Μοντέλο Δεδομένων (Relational Data Model) (Codd E. F., 1970) το οποίο είναι γνωστό για την απλότητα του και την ευκολία του στη προσαρμογή των εκάστοτε αναγκών του χρήστη. Αυτά τα δύο στοιχεία, καθώς επίσης και η δυνατότητα της συνολικής θεωρητικής περιγραφής του μοντέλου αποτέλεσαν τα βασικά αίτια της μεγάλης απήχησης του. Επίσης, η γλώσσα SQL (Structured Query Language) είναι η πλέον διαδεδομένη γλώσσα Βάσεων Δεδομένων και υλοποιείται από όλα τα σχεσιακά Συστήματα Διαχείρισης Βάσεων Δεδομένων.

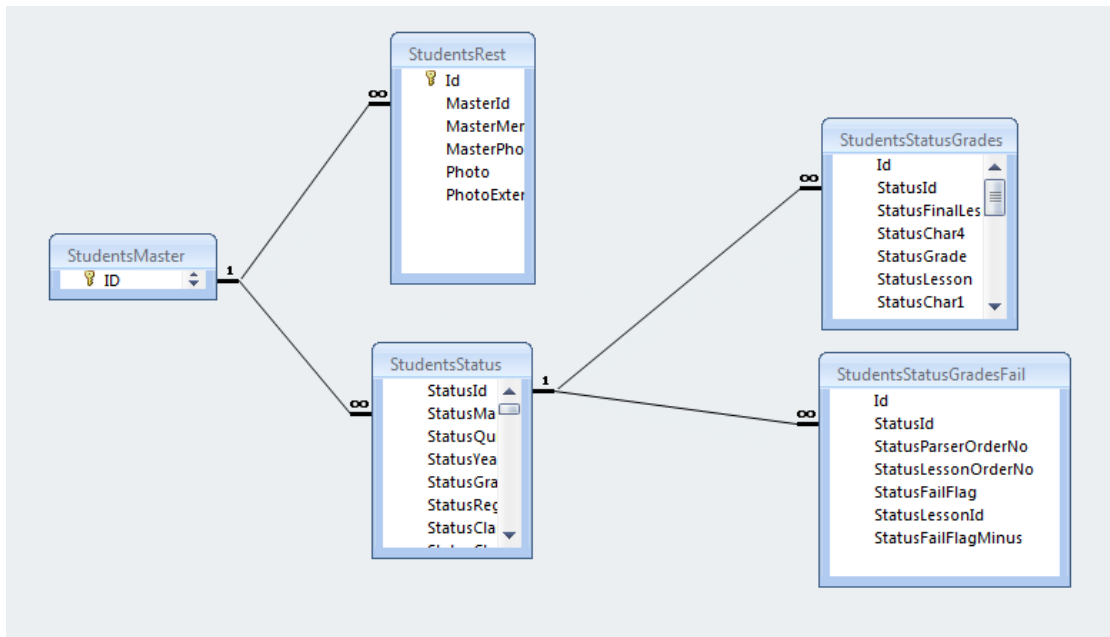
Στις Σχεσιακές (Relational) βάσεις δεδομένων, τα δεδομένα συνδέονται μεταξύ τους με σχέσεις (relations), οι οποίες προκύπτουν από τα κοινά πεδία που υπάρχουν σε διαφορετικά αρχεία. Τα αρχεία αποκαλούνται πίνακες (tables), οι εγγραφές γραμμές (rows) και τα πεδία στήλες (columns). Η ύπαρξη μιας κοινής τιμής στα πεδία δύο αρχείων καθορίζει και μια σχέση μεταξύ των γραμμών διαφορετικών πινάκων.

Το Σχεσιακό Μοντέλο Δεδομένων (relational data model) αναπτύχθηκε από τον Codd το 1970. Σαν μοντέλο έχει μεγάλη απήχηση κάτι το οποίο οφείλεται στην απλή και κατανοητή δομή του και αυτό συμβαίνει τόσο σε κλασικές όσο και σε σύγχρονες εφαρμογές υψηλών απαιτήσεων. Το μεγάλο πλεονέκτημα του μοντέλου είναι ότι μπορεί να περιγράψει με μαθηματικό τρόπο με τη βοήθεια της Θεωρίας Συνόλων ή της Κατηγορηματικής Λογικής. Τα περισσότερα Συστήματα Διαχείρισης Βάσεων Δεδομένων υποστηρίζουν το σχεσιακό μοντέλο δεδομένων. Επίσης το μοντέλο υποστηρίζεται και από πολλά εργαλεία σχεδιασμού (ER Studio, Developer 2000, Visio Enterprise κ.α.) διευκολύνοντας το σχεδιαστή της βάσης δεδομένων.

Τα πρώτα σχεσιακά Συστήματα Διαχείρισης Βάσεων Δεδομένων (ΣΔΒΔ) έθεσαν τα θεμέλια για την περαιτέρω διάδοση του μοντέλου, με αποτέλεσμα να υπάρχουν σήμερα αρκετά σχεσιακά συστήματα με πολλές δυνατότητες και ευκολίες προς τους χρήστες. Σημειώνεται, ότι για την κάλυψη των σύγχρονων εφαρμογών τα περισσότερα συστήματα υλοποιούν επεκτάσεις του καθαρού σχεσιακού μοντέλου. Μερικά από τα σύγχρονα διαθέσιμα ΣΔΒΔ τα οποία υποστηρίζουν το σχεσιακό μοντέλο δεδομένων είναι τα εξής :

- SQL Server της εταιρείας Microsoft
- Oracle 10g της εταιρείας Oracle
- DB2 και Informix της εταιρείας IBM

Στο Σχήμα 1.4 βλέπουμε την απεικόνιση μίας Σχεσιακής Βάσης Δεδομένων.



Σχήμα 1.4 : Απεικόνιση Σχεσιακής Βάσης Δεδομένων

Οι κανόνες που πρέπει να ακολουθούνται είναι οι εξής : (i) κάθε οντότητα πρέπει να παριστάνεται ως ένας ξεχωριστός πίνακας, (ii) κάθε στήλη του πίνακα αντιστοιχεί σε μια ιδιότητα της οντότητας, (iii) κάθε γραμμή πρέπει να είναι μοναδική, (iv) κάθε στήλη έχει μια δική της μοναδική ονομασία, (v) η στήλη που αποτελεί το πρωτεύον κλειδί μιας οντότητας δεν πρέπει να είναι ποτέ κενή, και τέλος (vi) αποκλείεται να υπάρχουν δυο ή και περισσότερες γραμμές που περιέχουν την ίδια τιμή στο πρωτεύον κλειδί.

1.2.4 Αντικειμενοστραφές Μοντέλο Δεδομένων (Object Oriented Data Model)

Το αντικειμενοστραφές Σύστημα Διαχείρισης Βάσεων Δεδομένων (Kim W., 2008) αποθηκεύει δεδομένα και διαδικασίες ως αντικείμενα που μπορούν να ανακτηθούν και να μοιράζονται αυτόματα.

Μια αντικειμενοστραφής βάση δεδομένων μπορεί να επεκτείνει την ύπαρξη των αντικειμένων με τέτοιο τρόπο ώστε να αποθηκεύονται μόνιμα και επομένως τα αντικείμενα μένουν σταθερά μετά τον τερματισμό του προγράμματος που τα χρησιμοποιεί. Τα αντικείμενα μένοντας σταθερά μετά τον τερματισμό του προγράμματος μπορούν να ανακτηθούν αργότερα και να είναι κοινά σε άλλα προγράμματα. Δηλαδή, οι αντικειμενοστραφείς βάσεις δεδομένων αποθηκεύουν διαρκή αντικείμενα (persistent objects) μόνιμα σε δευτερεύουσες μονάδες

αποθήκευσης και επιτρέπουν το διαμοιρασμό των αντικειμένων αυτών μεταξύ πολλαπλών προγραμμάτων και εφαρμογών.

Ένας στόχος των αντικειμενοστραφών βάσεων δεδομένων είναι να διατηρούν άμεση αντιστοιχία μεταξύ των αντικειμένων του πραγματικού κόσμου και των αντικειμένων της βάσης δεδομένων. Αυτό γίνεται ώστε τα αντικείμενα να μην χάνουν την ακεραιότητα και τη ταυτότητα τους και να είναι εύκολη η αναγνώριση τους και η εφαρμογή πράξεων σε αυτά. Επομένως, οι αντικειμενοστραφής βάσεις δεδομένων παρέχουν ένα μοναδικό αναγνωριστικό (object identifier) το οποίο δίνεται από το σύστημα σε κάθε αντικείμενο.

Ένα άλλο χαρακτηριστικό των αντικειμενοστραφών βάσεων δεδομένων είναι ότι επιτρέπουν στο αντικείμενο να έχει δομή οποιασδήποτε πολυπλοκότητας, προκειμένου να περιέχει όλες τις σημαντικές πληροφορίες που περιγράφουν το αντικείμενο.

1.3 Στόχοι του Σχεσιακού Μοντέλου και Δυσκολίες που Αντιμετωπίζονται

Στη σημερινή εποχή όπου ο όγκος των δεδομένων που διακινούνται είναι τεράστιος, υπάρχει η ανάγκη για γρήγορη και αξιόπιστη ανάλυση των δεδομένων προκειμένου οι μεγάλες επιχειρήσεις και οι οργανισμοί να παίρνουν τις πληροφορίες που χρειάζονται γρήγορα και με ασφάλεια.

Το σχεσιακό μοντέλο δεδομένων έχει κάποιους βασικούς στόχους οι οποίοι είναι οι εξής:

- Η υποστήριξη της ανεξαρτησίας δεδομένων, έτσι ώστε αλλαγές στη φυσική δομή και οργάνωση της βάσης δεδομένων να μην απαιτούν αλλαγές στα προγράμματα εφαρμογής.
- Η αποφυγή του πλεονασμού (redundancy) που εμφανίζεται όταν τα ίδια δεδομένα αποθηκεύονται πολλές φορές σε διαφορετικές περιοχές της βάσης δεδομένων.
 - Η διατήρηση της ακεραιότητας και της συνέπειας των δεδομένων.
 - Η υποστήριξη της ανάπτυξης συνολοθεωρητικών γλωσσών χειρισμού δεδομένων, οι οποίες στηρίζονται στη θεωρία συνόλων και διευκολύνουν τη διατύπωση ερωτημάτων προς το ΣΔΒΔ.
 - Ο καταμερισμός (sharing) των ίδιων στοιχείων σε όλους τους χρήστες.

- Η ομοιομορφία (uniformity) στον χειρισμό και την αναπαράσταση των δεδομένων.
- Η επιβολή κανόνων ασφαλείας (security).

Το σχεσιακό μοντέλο δεδομένων το οποίο στηρίζεται στην SQL αντιμετωπίζει κάποιες δυσκολίες. Η SQL είναι αρκετά δύσκολη γλώσσα όσον αφορά τη δυναμική ανάλυση και τις πολύπλοκες εφαρμογές. Επίσης είναι αρκετά ασύμμετρη για αυτό μπορούν να πραγματοποιηθούν πολύπλοκα προβλήματα και υπολογισμοί γενικά μόνο εντός των αρχείων και όχι πέρα από αυτά. Τέλος παρά τις επιδόσεις των δυνατοτήτων ρύθμισης που περιλαμβάνει, εξακολουθεί να είναι πολύ εύκολο να χαθεί ο έλεγχος της απόδοσης (Kimball R., 2007).

2. Συστήματα Αναλυτικής Επεξεργασίας (OLAP)

Μετά από τη περιγραφή των Συστημάτων Διαχείρισης Δεδομένων και μερικών από τα σημαντικότερα μοντέλα βάσεων δεδομένων παρουσιάζεται ένα πιο σύγχρονο μοντέλο διαχείρισης βάσεων δεδομένων το οποίο στηρίζεται σε Αποθήκες Δεδομένων και την Αναλυτική Επεξεργασία και το οποίο έχει πολυδιάστατη μορφή.

2.1 Εισαγωγή

Οι απαιτήσεις των σύγχρονων εφαρμογών οδήγησαν στην ανάγκη μελέτης εναλλακτικών προσεγγίσεων στο θέμα της αποθήκευσης και της επεξεργασίας δεδομένων. Για το λόγο αυτό περιγράφεται ένα μοντέλο, δίνοντας έμφαση στις απαιτήσεις του και στις λειτουργίες που πρέπει να υποστηριχθούν. Σε πολλές περιπτώσεις η χρήση ενός ΣΔΒΔ με επιπλέον δυνατότητες μπορεί να καλύψει τις υπάρχουσες ανάγκες, ενώ σε άλλες περιπτώσεις απαιτούνται νέα συστήματα με διαφορετικές δυνατότητες από αυτές ενός ΣΔΒΔ. Το μοντέλο που θα παρουσιαστεί έχει σαν βάση τις Αποθήκες Δεδομένων και την Αναλυτική επεξεργασία.

Με τη συνεχή συγκέντρωση δεδομένων παρουσιάζεται το πρόβλημα της ανεπαρκούς ανάλυσης τους, λόγω του τεράστιου όγκου της πληροφορίας. Τέτοιο παράδειγμα μεγάλου όγκου δεδομένων είναι τα δεδομένα που λαμβάνονται από δορυφόρους ή μετεωρολογικούς σταθμούς. Για το λόγο αυτό τα δεδομένα συγκεντρώνονται αρχικά σε Αποθήκες Δεδομένων (data warehouses) και στη συνέχεια εκτελούνται εργασίες Αναλυτικής Επεξεργασίας Δεδομένων (on-line analytical processing, OLAP) (Chaudhuri, 1997). Τα δεδομένα μετατρέπονται σε κατάλληλη μορφή ώστε να μπορούν να αποθηκευτούν και να διαχειρίζονται από την αποθήκη δεδομένων. Οι τεχνικές OLAP που εφαρμόζονται επιτρέπουν την ανάλυση των δεδομένων ανάλογα με τις διαστάσεις που έχουν ορισθεί (π.χ. σχολική χρονιά, μαθητές).

2.2 Πλεονεκτήματα OLAP

Τα συστήματα αναλυτικής επεξεργασίας λόγω της εφαρμογής πολυδιάστατης ανάλυσης έχουν αρκετά πλεονεκτήματα σε σχέση με άλλα μοντέλα βάσεων δεδομένων.

Οι τεχνικές OLAP αυξάνουν την παραγωγικότητα των στελεχών των επιχειρήσεων, των προγραμματιστών και ολόκληρων οργανισμών. Η ευελιξία των συστημάτων OLAP βοηθάει ώστε οι χρήστες τους να μπορούν να γίνουν πιο αυτόνομοι. Οι διευθυντές πλέον δεν εξαρτώνται από την πληροφορική για να κάνουν αλλαγές σχήματος ή να δημιουργήσουν συνδέσεις. Ίσως το πιο σημαντικό, είναι ότι η OLAP επιτρέπει στους διευθυντές να μοντελοποιήσουν προβλήματα που θα ήταν αδύνατον χρησιμοποιώντας λιγότερο ευέλικτα συστήματα με μεγάλους και ασυνεπείς χρόνους απόκρισης. Επίσης τα συστήματα OLAP παρέχουν περισσότερο έλεγχο και έγκαιρη πρόσβαση σε στρατηγικές πληροφορίες το οποίο ισούται με πιο αποτελεσματική λήψη αποφάσεων.

Γενικότερα οι προγραμματιστές μπορούν να επωφεληθούν σε μεγάλο ποσοστό από τη χρήση του σωστού λογισμικού OLAP. Αν και είναι δυνατόν να οικοδομηθεί ένα σύστημα OLAP χρησιμοποιώντας λογισμικό που έχει σχεδιαστεί για την επεξεργασία των συναλλαγών ή στοιχείων συλλογής, σίγουρα δεν είναι μια πολύ αποτελεσματική χρήση του χρόνου των προγραμματιστών. Με τη χρήση λογισμικού ειδικά σχεδιασμένο για συστήματα OLAP, οι προγραμματιστές μπορούν να παραδώσουν εφαρμογές στους χρήστες ταχύτερα, προσφέροντας καλύτερες υπηρεσίες. Η ταχύτερη παράδοση των εφαρμογών επίσης μειώνει την καθυστέρηση στη λειτουργία των εφαρμογών.

Η διαδικασία OLAP μειώνει την καθυστέρηση των εφαρμογών ακόμα περισσότερο κάνοντας τους χρήστες αρκετά αυτόνομοι ώστε να δημιουργήσουν δικά τους μοντέλα. Ωστόσο σε αντίθεση με αυτόνομες τμηματικές εφαρμογές οι οποίες τρέχουν σε δίκτυα υπολογιστών, οι εφαρμογές OLAP εξαρτώνται από τις αποθήκες δεδομένων και από τα συστήματα επεξεργασίας συναλλαγών για να ανανεώσουν το επίπεδο πηγαιών δεδομένων. Ως αποτέλεσμα, κερδίζονται περισσότερο αυτόνομοι χρήστες χωρίς να χάνεται ο έλεγχος της ακεραιότητας των δεδομένων. Μέσω της OLAP παρέχεται πιο αποτελεσματική λειτουργία. Με τη χρήση του λογισμικού που έχει σχεδιαστεί για OLAP, μειώνεται η ροή

ερωτημάτων και η κίνηση στο δίκτυο στα συστήματα συναλλαγών ή των Αποθηκών δεδομένων.

Τέλος παρέχει τη δυνατότητα να μοντελοποιηθούν πραγματικά προβλήματα των επιχειρήσεων και των διάφορων οργανισμών με μια πιο αποτελεσματική χρήση των ανθρώπινων πόρων. Η OLAP επιτρέπει στους οργανισμούς να ανταποκρίνονται πιο γρήγορα στις απαιτήσεις της αγοράς και η ανταπόκριση της αγοράς, με τη σειρά της, δίνει συχνά τη βελτίωση των εσόδων και κερδοφορία.

2.2.1 Πλεονεκτήματα της OLAP Έναντι των Σχεσιακών Βάσεων

Το μοντέλο OLAP χρησιμοποιεί σαν βάση το παλιό σχεσιακό μοντέλο δεδομένων το οποίο στην ουσία έρχεται να αντικαταστήσει. Για το λόγο αυτό θα αναφερθούν τα πλεονεκτήματα του έναντι του σχεσιακού μοντέλου δεδομένων. Τα πλεονεκτήματα αυτά είναι τα εξής:

1. Το μοντέλο OLAP γενικά έχει καλύτερη απόδοση από τις σχεσιακές βάσεις όταν οι κύβοι που το απαρτίζουν είναι σωστά σχεδιασμένοι, έχοντας μικρότερη ανάγκη να παίξει περίπλοκα παιχνίδια ρύθμισης.
2. Παρέχει πολύ υψηλότερες δυνατότητες ανάλυσης από τις σχεσιακές βάσεις. Για παράδειγμα η MDX (Γλώσσα Ερωτημάτων για Πολυδιάστατα Δεδομένα- Multidimensional Expressions)(Pasumansky M., Whitehorn M, Zare R., 2007) έχει πιο ισχυρή σημασιολογία για το πέρασμα ανόμοιων και ασύμμετρων ιεραρχιών, όπως τα γραφήματα οργανισμών απ' ότι έχουν οι βάσεις δεδομένων στο σχεσιακό μοντέλο.
3. Τα εργαλεία OLAP για στατιστικές αναφορές και ερωτήματα έχουν γίνει ιστορικά ανώτερα από τα σχεσιακά εργαλεία.
4. Η OLAP δε πάσχει από το πρόβλημα συμμετρίας που περιορίζει τις σχεσιακές βάσεις δεδομένων.
5. Ορισμένα σενάρια φόρτωσης μπορεί να είναι πολύ γρήγορα.
6. Η διαδικασία OLAP επιτρέπει πιο περίπλοκα σενάρια ασφαλείας. Συγκρίνοντας, είναι δύσκολο να ρυθμιστή μια σχεσιακή βάση δεδομένων ώστε να προστατεύει λεπτομερή δεδομένα.

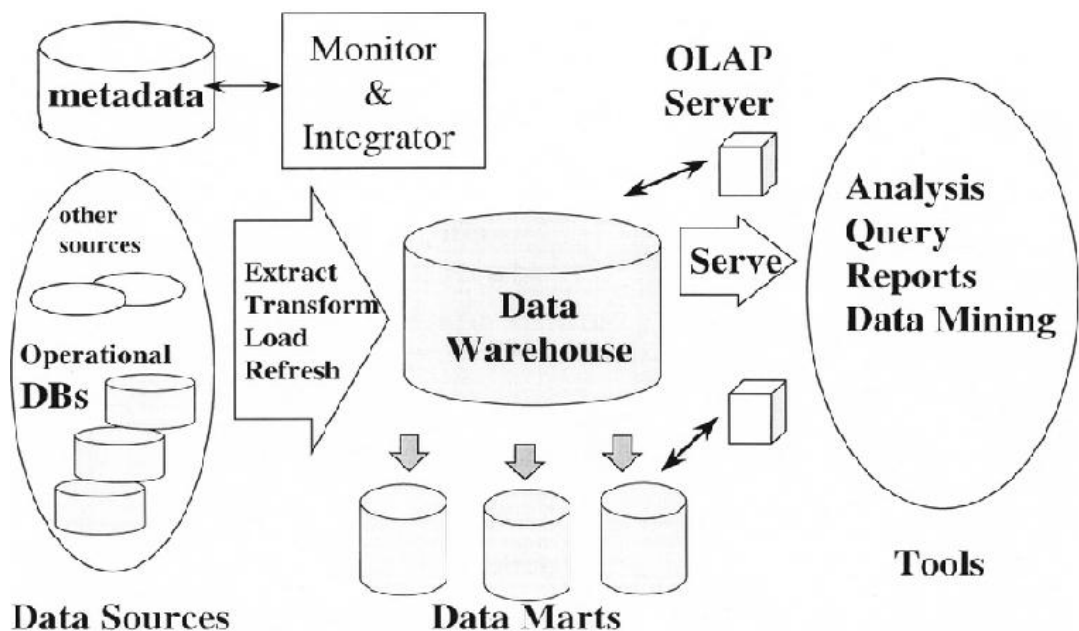
2.3 Αποθήκες Δεδομένων

Με τον όρο Αποθήκες Δεδομένων χαρακτηρίζουμε ένα σύνολο τεχνολογιών στήριξης αποφάσεων, που μαζί με την Αναλυτική Επεξεργασία Δεδομένων παρέχουν στον τελικό χρήστη τη δυνατότητα της απόκτησης πολύτιμης πληροφορίας από δεδομένα, που συγκεντρώνεται συνήθως από τις βάσεις δεδομένων μιας επιχείρησης ή ενός οργανισμού, η οποία προάγει τη στρατηγική λήψη αποφάσεων (Devlin B., Cote L.D., 1996).

Μια αποθήκη δεδομένων αποτελεί ένα αντίγραφο των δεδομένων που συγκεντρώνεται από τις συναλλαγές ενός οργανισμού και είναι δομημένη με τέτοιο τρόπο ώστε να εξυπηρετούνται πολύπλοκες ερωτήσεις και να γίνεται η ανάλυση των δεδομένων. Σύμφωνα με τους Chaudhuri και Dayal, η αποθήκη δεδομένων είναι μια βάση δεδομένων υποστήριξης αποφάσεων, που συντηρείται χωριστά από τη βάση δεδομένων παραγωγής ενός οργανισμού.

2.3.1 Αρχιτεκτονική Αποθήκης Δεδομένων

Η επιλογή της αρχιτεκτονικής μιας αποθήκης δεδομένων πρέπει να ικανοποιεί τις συγκεκριμένες ανάγκες του οργανισμού για τον οποίο δημιουργήθηκε και να εξασφαλίζει τη διαθεσιμότητα και την αποδοτικότητα του συστήματος. Το σχήμα 2.1 παρουσιάζει μια γενική αρχιτεκτονική ενός συστήματος Αποθήκης Δεδομένων.



Σχήμα 2.1: Αρχιτεκτονική Συστήματος Αποθήκης Δεδομένων (http://www.emeraldinsight.com/content_images/file/0670330804028.png)

Το σχήμα 2.1 περιλαμβάνει εργαλεία για την εξαγωγή δεδομένων από πολλαπλές λειτουργικές βάσεις και άλλες πηγές για τον καθορισμό, την μετατροπή και την ολοκλήρωση των δεδομένων, για την φόρτωση δεδομένων στην αποθήκη δεδομένων και για την περιοδική ανανέωση της αποθήκης ώστε να παρουσιάζονται οι ανανεώσεις των πηγών και να καθαριστούν τα δεδομένα από την αποθήκη. Επιπλέον στην βασική αποθήκη μπορεί να υπάρχουν αρκετές τμηματικές συλλογές δεδομένων (Data Marts). Τα δεδομένα στην αποθήκη δεδομένων και στις συλλογές δεδομένων αποθηκεύονται και διαχειρίζονται από ένα ή περισσότερους εξυπηρετητές αποθήκης (warehouse servers), οι οποίοι παρουσιάζουν πολυδιάστατες όψεις δεδομένων σε μια ποικιλία εργαλείων: εργαλεία ερωτημάτων (query tools), συγγραφείς αναφορών (reports writers), εργαλεία ανάλυσης (analysis tools), εργαλεία εξόρυξης δεδομένων (data mining tools). Τελικά, υπάρχει ένας χώρος ο οποίος χρησιμοποιείται για αποθήκευση και διαχείριση μεταδεδομένων, όπως επίσης υπάρχουν και εργαλεία παρακολούθησης και διαχείρισης του συστήματος αποθήκευσης.

Πιο αναλυτικά, τα δομικά μέρη της αρχιτεκτονικής ενός συστήματος Αποθήκης Δεδομένων είναι:

- Εξωτερικές πηγές: Είναι οι πηγές από τις οποίες η αποθήκη αντλεί δεδομένα.

Μπορεί να είναι είτε βάσεις δεδομένων των συστημάτων του οργανισμού είτε διάφορα αρχεία εφαρμογών και αρχεία κειμένου είτε ακόμα εξωτερικές πηγές πληροφοριών, στις οποίες έχει πρόσβαση ο οργανισμός.

- Μεταφορείς-Μετατροπείς: Είναι εφαρμογές που εκτελούν τις διαδικασίες μεταφοράς από τις πηγές στην Αποθήκη. Για κάθε μία διαφορετική πηγή ή κατηγορία πηγής, ένας διαφορετικός μεταφορέας αναλαμβάνει την άντληση των δεδομένα της. Είναι υπεύθυνοι για την εξαγωγή, τη μεταφορά, τον καθαρισμό και τον έλεγχο πληρότητας των δεδομένων.
- Αποθήκη δεδομένων: Είναι τα συστήματα που αποθηκεύονται τα δεδομένα που παρέχονται προς τους χρήστες. Τα δεδομένα αποθηκεύονται σε σχεσιακές ΒΔ και η πρόσβαση σε αυτά τα δεδομένα παρέχεται από μια γλώσσα διαχείρισης δεδομένων που είναι επέκταση της SQL. Εναλλακτικά μπορούν να αποθηκευθούν με πολυδιάστατο τρόπο και να χρησιμοποιηθούν Πολυδιάστατα Συστήματα Αναλυτικής Επεξεργασίας.

- Συλλογές Δεδομένων: Οι συλλογές δεδομένων περιέχουν τμήματα των δεδομένων της Αποθήκης Δεδομένων. Ο κατακερματισμός του περιεχομένου της αποθήκης σε επιμέρους συλλογές γίνεται με οργανωτικά κριτήρια και έχει στόχο την πιο άμεση και αποδοτική πρόσβαση των εφαρμογών ανάλυσης στα δεδομένα της Αποθήκης καθώς και για να ελαφρωθεί το φόρτο εργασίας της κεντρικής Αποθήκης δεδομένων. Η ύπαρξη τους ή όχι είναι επιλογή του διαχειριστή του συστήματος.
- Βάση μεταδεδομένων: Είναι ένα σύστημα αποθήκευσης πληροφορίας σχετικό με τη δομή και τη λειτουργία του συστήματος. Συνήθως περιέχει ένα λεξικό δεδομένων, δηλαδή περιέχει τον ορισμό, τη περιγραφή των δεδομένων και τις μεταξύ τους συσχετίσεις που αποθηκεύονται στην Αποθήκη. Περιέχει επίσης περιγραφή της ροής των δεδομένων μέσα στο σύστημα, δεδομένα ελέγχου των διάφορων εκδοχών των δεδομένων, στατιστικά χρήσης των δεδομένων κ.ά.
- Διαχειριστής: Είναι εφαρμογή που παρέχει τη δυνατότητα διαχείρισης του συστήματος.
- Εφαρμογές ανάλυσης: Είναι εφαρμογές που έχουν πρόσβαση στην Αποθήκη Δεδομένων. Συνήθως είναι συστήματα στήριξης αποφάσεων και αποτελούνται από εργαλεία επερωτήσεων, ανάλυσης και εξόρυξης δεδομένων.

Ο σχεδιασμός μίας αποθήκης δεδομένων είναι μια πολύπλοκη διαδικασία της οποίας τα βήματα είναι τα εξής :

Βήμα 1: Ορίζεται η αρχιτεκτονική, γίνεται ο σχεδιασμός της δυναμικότητας και επιλέγονται οι εξυπηρετητές αποθήκευσης, οι βάσεις δεδομένων, οι εξυπηρετητές OLAP και τα εργαλεία.

Βήμα 2: Ενσωμάτωση των εξυπηρετητών, της αποθήκευσης και των εργαλείων του πελάτη.

Βήμα 3: Σχεδίαση του σχήματος της αποθήκης.

Βήμα 4: Ορισμός της οργάνωσης της φυσικής αποθήκης, της τοποθέτησης των δεδομένων, του διαμερισμού αυτών και των μεθόδων πρόσβασης.

Βήμα 5: Ένωση των πηγών με τη χρήση πυλών, οδηγών ODBC (Chaudhuri S., Dayal U., 1997) και άλλων συσκευαστών.

Βήμα 6: Σχεδιασμός και υλοποίηση σεναρίων για την άντληση στοιχείων, καθαρισμού, μετατροπής, φόρτωσης και ανανέωσης.

Βήμα 7: Συμπλήρωση της αποθήκης με το σχήμα και τη προβολή ορισμών, σεναρίων, και άλλων μεταδεδομένων.

Βήμα 8: Σχεδιασμός και χρήση εφαρμογών τελικού χρήστη.

Βήμα 9: Άνοιγμα της αποθήκης και των εφαρμογών.

2.4 Εισαγωγή στην OLAP

Ένας κύβος OLAP είναι μια τεχνολογία που αποθηκεύει δεδομένα με βελτιστοποιημένο τρόπο παρουσιάζοντας μια γρήγορη απάντηση σε διάφορους τύπους σύνθετων ερωτημάτων χρησιμοποιώντας μεγέθη, όπως διαστάσεις και μέτρα. Οι περισσότεροι κύβοι αποθηκεύουν από πριν τα μέτρα με ειδική δομή αποθήκευσης για να παρέχουν γρήγορη απάντηση σε ερωτήματα.

Ένας κύβος OLAP είναι ένα διάνυσμα δεδομένων κατανοητό από την άποψη μίας ή και περισσότερων διαστάσεων. Το όνομα του κύβου OLAP προήλθε από την διασυνδεδεμένη αναλυτική επεξεργασία (online analytical processing). Ο κύβος OLAP είναι μια τεχνική βασισμένη σε υπολογιστή για την ανάλυση επιχειρησιακών δεδομένων (business data) για την έρευνα της επιχειρηματική νοημοσύνης.

Τα εργαλεία τα οποία χρησιμοποιούνται στην επιχειρησιακή νοημοσύνη αναφορικά είναι:

- Υπολογιστικά φύλλα (Spreadsheets)
- Υποβολή εκθέσεων και επερώτηση του λογισμικού (Reporting and querying software): εργαλεία που εξάγουν, ταξινομούν, συνοψίζουν και παρουσιάζουν επιλεγμένα δεδομένα
- OLAP: Online analytical processing
- Ψηφιακό ταμπλό (Digital dashboards)
- Εξόρυξη δεδομένων (Data mining)
- Αποθήκευση δεδομένων (Data warehousing)
- Απόφαση της μηχανικής (Decision engineering)
- Διαδικασία εξόρυξης (Process mining)
- Διαχείριση επιδόσεων των Επιχειρήσεων (Business performance management)
- Τοπικά συστήματα πληροφοριών (Local information systems)

Προγράμματα ανοιχτού λογισμικού τα οποία προσφέρονται δωρεάν είναι :

- Jaspersoft, όπου αναφέρεται στην υποβολή εκθέσεων, τον πινάκων εργαλείων, στην Ανάλυση Δεδομένων και στην Ενσωμάτωση δεδομένων

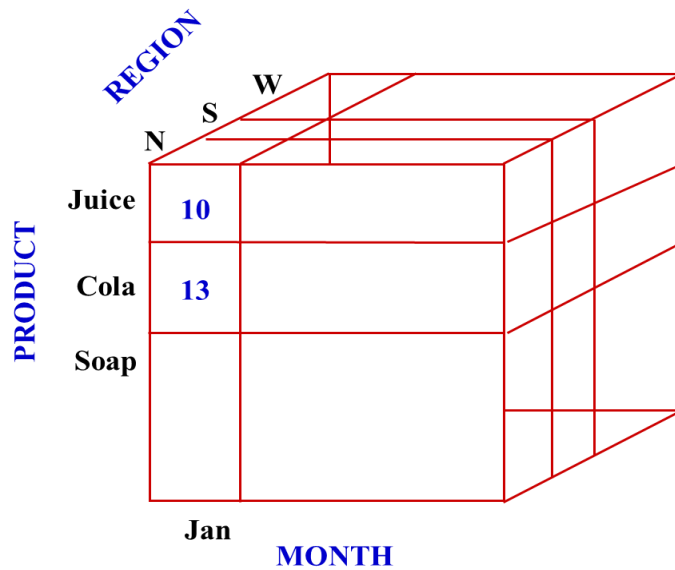
- Palo (OLAP database), όπου αναφέρεται στον OLAP Server, στον Worksheet Server and στον ETL Server
- Pentaho: όπου αναφέρεται στην αναφορά, στην ανάλυση, στους πίνακες, στην εξόρυξη δεδομένων και στη ροή εργασίας των δυνατοτήτων
- TACTIC: όπου αναφέρεται στην αναφορά, στην ανάλυση, στους πίνακες (dashboard), στην εξόρυξη δεδομένων (data mining), στην ενοποίηση (intergration) και στις δυνατότητες ροής εργασίας
- Actuate Corporation: το οποίο παρέχει ανάλυση πολλών δεδομένων (big data), εξόρυξη δεδομένων, οπτικοποίηση δεδομένων, εργαλεία, αναφορές και διαδραστικό ταμπλό (interactive dashboards).

2.5 Βασικές Έννοιες OLAP

Η Σύγχρονη Αναλυτική Επεξεργασία Δεδομένων (OLAP) είναι μια κατηγορία λογισμικού που επιτρέπει σε αναλυτές και διοικητικά στελέχη να αποκτήσουν γνώση των δεδομένων. Αυτό γίνεται μέσω μιας γρήγορης, συνεπούς και αξιόπιστης πρόσβασης σε μια μεγάλη ποικιλία όψεων της πληροφορίας που έχει μετασχηματιστεί από απλά δεδομένα, ώστε να αναπαριστά τη πολυδιάστατη θεώρηση ενός οργανισμού, όπως γίνεται αντιληπτή από τον χρήστη. Η λειτουργικότητα του OLAP χαρακτηρίζεται από τη δυναμική πολυδιάστατη ανάλυση συναθροισμένων δεδομένων οργανισμών και επιχειρήσεων. Αυτά υποστηρίζουν τον τελικό χρήστη στις αναλυτικές πλοηγήσεις του, συμπεριλαμβάνοντας υπολογισμούς και μοντελοποιήσεις που βασίζονται σε διαστάσεις, ιεραρχίες και επίπεδα, ανάλυση τάσεων σε συνεχείς χρονικές περιόδους, τεμαχισμό υποσυνόλων για προβολές στην οθόνη, ανάλυση σε χαμηλότερα επίπεδα συνάθροισης, περιστροφή για νέες συγκρίσεις διαστάσεων κλπ.

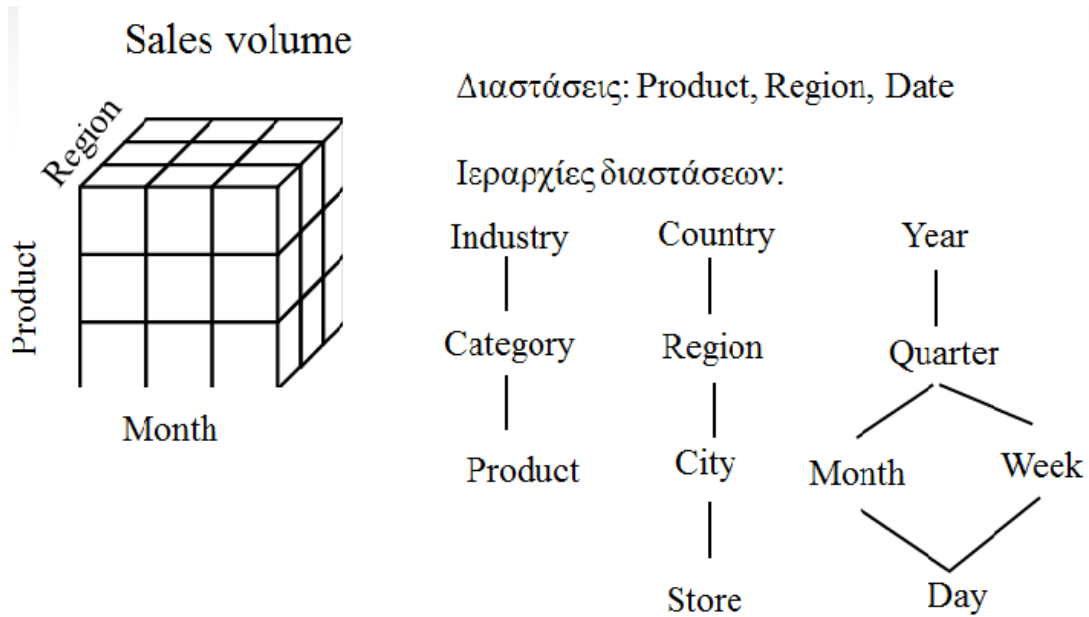
Στόχος των OLAP εργαλείων είναι να περιγράψουν την κάθε πληροφορία με πολυδιάστατο τρόπο. Για να το επιτύχουν αυτό, χρησιμοποιούν πολυδιάστατη μοντελοποίηση και αποθήκευση των δεδομένων. Τα δεδομένα οργανώνονται σε κύβους ή αλλιώς υπερκύβους (cubes και hyper cubes αντίστοιχα), που ορίζονται σε πολυδιάστατους χώρους, αποτελούμενους από πολλές διαστάσεις.

Ο υπερκύβος δεδομένων (data hyper cube) είναι ένας πολυδιάστατος πίνακας (multidimensional array): $DH[d_1, \dots, d_n]$ από λίστες $l=(e_1, \dots, e_m)$.



Σχήμα 2.2 : Κύβος OLAP(διάλεξη Κύριου Βασιλειάδη Π., http://www.cs.uoi.gr/~pvassil/courses/db_III/)

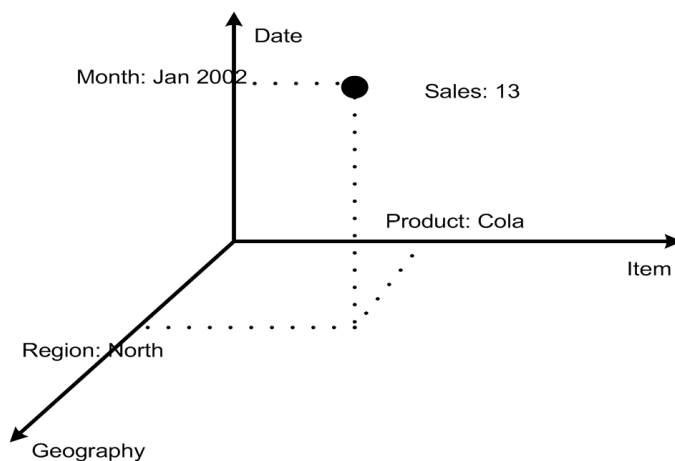
Κάθε στοιχείο μιας λίστας είναι ένα διατεταγμένο ζευγάρι $e_i = \{label_i, value_i\}$ που σημαίνει ότι κάθε στοιχείο καθορίζεται από δύο χαρακτηριστικά: την ετικέτα του, η οποία είναι διακριτή, και τη τιμή της. Όλες οι λίστες του υπερκύβου έχουν τον ίδιο αριθμό στοιχείων και το κάθε στοιχείο κρατά την ετικέτα του σε όλες τις λίστες. Σε κάθε διάσταση d_i , ορίζεται επίσης, μια ετικέτα ($label_i$), που είναι μοναδική σε κάθε διάσταση. Κάθε διάσταση έχει μια σχετική ιεραρχία επιπέδων συνάθροισης των δεδομένων (hierarchy of levels). Αυτό σημαίνει ότι η διάσταση μπορεί να θεωρηθεί από πολλά άλλα επίπεδα όπως για παράδειγμα μέρα, εβδομάδα, μήνας, χρόνος. Κάθε διάσταση οργανώνεται από διαφορετικά επίπεδα λεπτομέρειας και ο χρήστης μπορεί να πλοηγηθεί από το ένα επίπεδο στο άλλο δημιουργώντας νέους κύβους κάθε φορά.



Σχήμα 2.3: κύβος OLAP και Ιεραρχίες Διαστάσεων

(Διαλέξεις. Βασιλειάδη Π., http://www.cs.uoi.gr/~pvassil/courses/db_III/)

Ορίζουμε τις συναρτήσεις, $label(i)$ που επιστρέφει την ετικέτα ενός στοιχείου ή μιας διάστασης, $value(i)$ που επιστρέφει την τιμή ενός στοιχείου και $dom(d_i)$ που επιστρέφει το πεδίο ορισμού των τιμών διάστασης μιας διάστασης d_i .



Σχήμα 2.4 : Διάγραμμα μέτρων

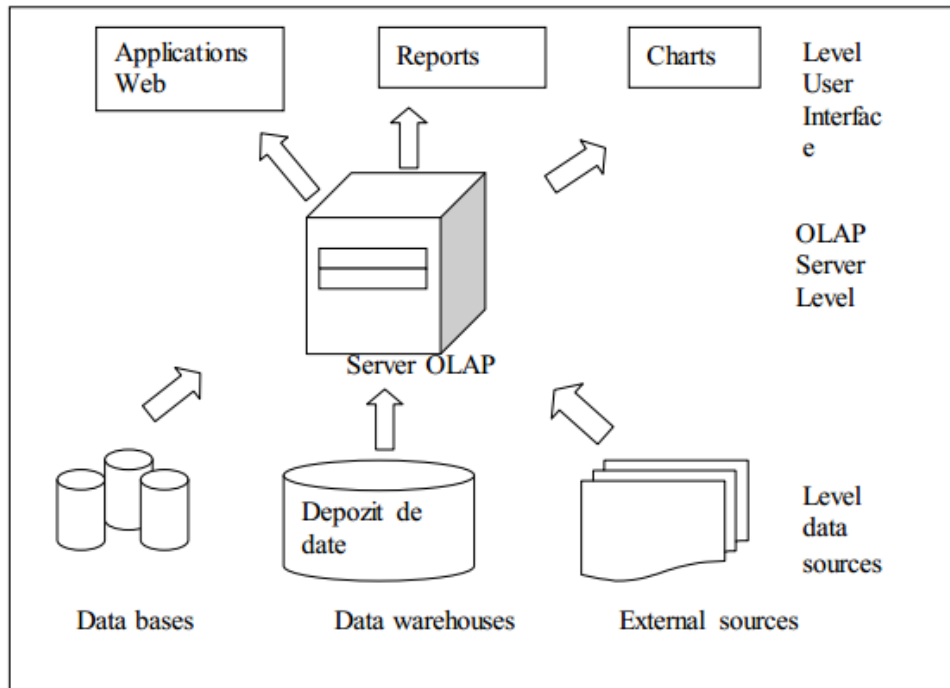
(Διάλεξη Βασιλειάδη Π., http://www.cs.uoi.gr/~pvassil/courses/db_III/)

Τα στοιχεία, που ονομάζονται και «μέτρα» (measures) αναπαριστούν τα δεδομένα του πραγματικού κόσμου και μοντελοποιούν όλους τους τρόπους με τους οποίους τα δεδομένα μπορούν να συναθροιστούν σε σχέση με μια συγκεκριμένη παράμετρο του περιεχομένου τους. Παραδείγματα τέτοιων εννοιών είναι: οι πωλήσεις, οι αγορές, η πελατειακή βάση κ.ά. Η ετικέτα ενός στοιχείου

είναι μια μέτρηση του μεγέθους αυτής της έννοιας. Υπάρχουν δύο ειδών στοιχεία: τα αρχικά (primitive), που είναι μέτρα αυτοκαθοριζόμενων εννοιών (π.χ. κόστος ανά τεμάχιο), και τα παράγωγα (compound), που ορίζονται ως συναρτήσεις άλλων αρχικών ή παράγωγων (π.χ. αριθμός τεμαχίων που πουλήθηκαν με X κόστος ανά τεμάχιο). Κάθε στοιχείο μιας λίστας είτε είναι αρχικό, είτε προήλθε από την εφαρμογή κάποιων συναρτήσεων πάνω σε άλλα στοιχεία της λίστας. Τέτοιες συναρτήσεις είναι οι αριθμητικοί τελεστές, οι στατιστικοί τελεστές (μέγιστο, ελάχιστο, μέσος, απόκλιση κ.ά.), τελεστές που εφαρμόζονται σε μια ακολουθία (π.χ. προηγούμενη περίοδος), κ.ά. Για να προσδιορίσουμε μια λίστα πρέπει πρώτα να προσδιορίσουμε τις διαστάσεις της n-διάστατης λίστας που υλοποιεί τον υπερκύβο. Αυτές οι διαστάσεις αντιπροσωπεύουν τις διαφορετικές όψεις, μέσω των οποίων τα δεδομένα μπορούν να προσπελαστούν. Κατά συνέπεια, ένας υπερκύβος δεδομένων παρέχει μια πολυδιάστατη εννοιολογική εικόνα των δεδομένων.

2.5.1 Αρχιτεκτονική των Συστημάτων OLAP

Τα εργαλεία OLAP συνήθως αποτελούνται από τον διακομιστή OLAP και τον OLAP client. Ο διακομιστής OLAP βρίσκεται ανάμεσα στο χρήστη και την αποθήκη δεδομένων και αποθηκεύει δεδομένα με τη μορφή των πολυδιάστατων πινάκων ή κύβων. Ο OLAP client βρίσκεται στον υπολογιστή του χρήστη και επιτρέπει στο χρήστη να ζητά δεδομένα από πίνακες και κύβους και να παίρνει εκθέσεις με τη μορφή γραφημάτων και πινάκων. Ένα παράδειγμα της αρχιτεκτονικής των συστημάτων OLAP δείχνεται στο Σχήμα 2.5.



Σχήμα 2.5 : Αρχιτεκτονική των συστημάτων OLAP(<http://fse.tibiscus.ro/anale/Lucrari2011/125.pdf>)

2.6 Μοντέλα Συστημάτων Αναλυτικής Επεξεργασίας

Τα αρχιτεκτονήματα δεδομένων (Data Warehouse) και η On-line αναλυτική επεξεργασία δεδομένων (OLAP), αποτελούν κομμάτι των εφαρμογών στήριξης αποφάσεων και συνθέτουν μια σύγχρονη προσέγγιση στο πρόβλημα της υποστήριξης αποφάσεων των εκάστοτε ληπτών αποφάσεων. Η On-line αναλυτική επεξεργασία δεδομένων (OLAP) είναι δυναμική και παρέχει πολυδιάστατη ανάλυση διαφόρων ειδών δεδομένων π.χ. επιχειρησιακών, της εκπαίδευσης κτλ . Ο βασικός στόχος είναι να παρέχει τη δυνατότητα στο χρήστη να παρακολουθεί τα εκάστοτε δεδομένα που του έχουν δοθεί σε διαφορετικά επίπεδα ανάλυσης, από το πιο αναλυτικό στο πιο συγκεντρωτικό και αντίστροφα, από διαφορετικές οπτικές γωνίες και με μια μεθοδολογία η οποία έχει σαν βάση τον άνθρωπο και είναι απαλλαγμένη από τεχνικά θέματα (Βουτσινάς Β., 2003) .

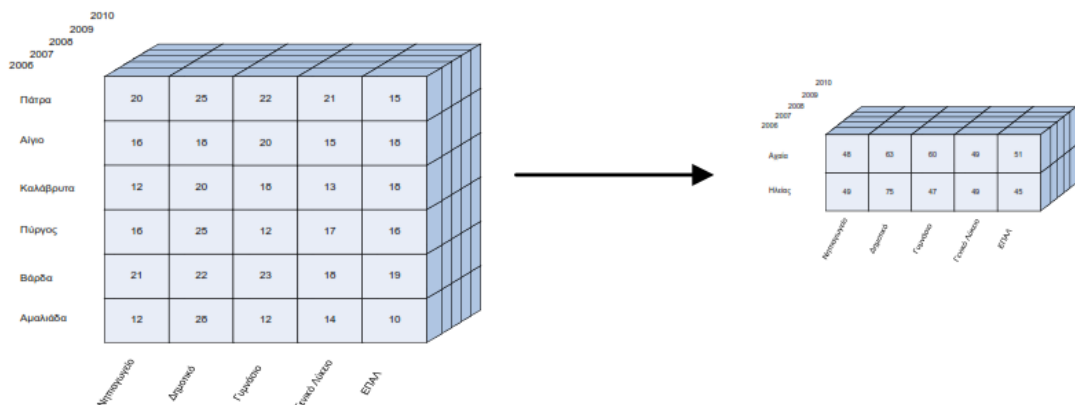
Ο βασικός συντελεστής που καθορίζει την επιτυχία τέτοιων εφαρμογών είναι η μεθοδολογία ανάπτυξης και το OLAP μοντέλο που θα υιοθετηθεί. Τα εργαλεία OLAP γενικότερα προσπαθούν να καλύψουν τις απαιτήσεις για υψηλού επιπέδου πληροφορίες των υψηλόβαθμων διοικητικών στελεχών, όσο και τις ανάγκες για άμεση εξειδικευμένη ανάλυση των πιο χαμηλόβαθμων αναλυτών.

Τα εργαλεία OLAP είναι σχεδιασμένα για να υποστηρίξουν την διαδικασία υποστήριξης λήψης αποφάσεων, λαμβάνοντας υπ' όψη δεδομένα από τυπικές διαδικασίες των επιχειρήσεων και οργανισμών έως και εξειδικευμένες. Αυτό επιτυγχάνεται παρέχοντας στους χρήστες διάφορες OLAP συναρτήσεις, όπως οι δημοφιλείς roll up, drill down, slice, και pivot. Συνήθως, κάθε OLAP συνάρτηση είναι περιεκτική και εύκολα καταλαβαίνει κανείς τη λειτουργία της. Η εξεζητημένη όμως χρήση των OLAP εργαλείων απαιτεί πολύπλοκους συνδυασμούς διαφορετικών OLAP συναρτήσεων, που δεν είναι εύκολα άμεσα χρησιμοποιήσιμες στους χρήστες. Ο ορισμός ενός OLAP μοντέλου το οποίο λαμβάνει υπ' όψιν το αρχιτεκτόνημα δεδομένων είναι το κλειδί της επιτυχίας μίας OLAP εφαρμογής (Roiger R., Geatz M., 2008).

2.6.1 Πράξεις στους Υπερκύβους

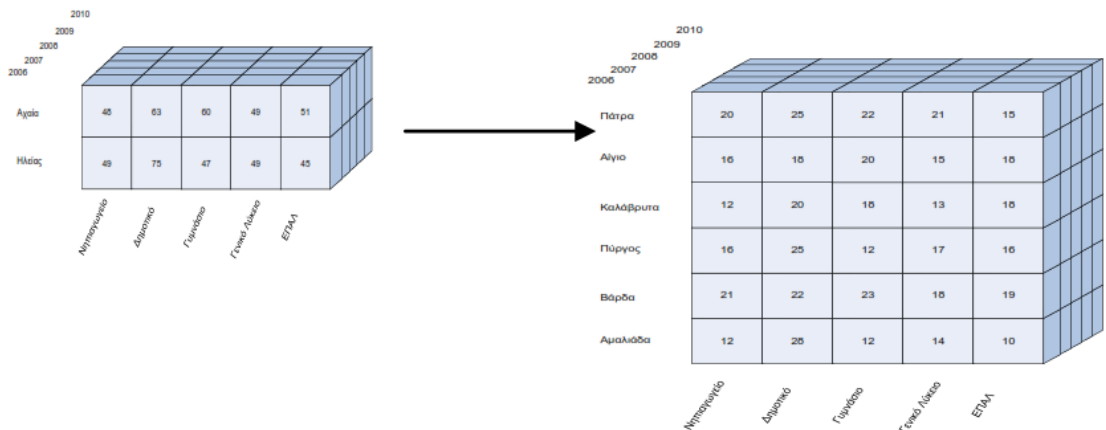
Οι πιο συχνές πράξεις που γίνονται στους κύβους είναι οι Συναθροίσεις (πχ total absents), Συγκρίσεις (Αγόρια vs Κορίτσια), Ταξινόμηση - κατάταξη (top 10), πρόσβαση σε πιο αναλυτική πληροφορία και οπτικοποίηση με διαφορετικό τρόπο για καθετί που απαιτείται. Αυτές οι πράξεις που αναφέραμε είναι οι εξής:

- **Roll-up:** είναι η πράξη με την οποία πραγματοποιούμε ένα βήμα προς τα πάνω στην ιεραρχία μιας διάστασης. Ο κύβος που προκύπτει από την πράξη περιέχει πιο ομαδοποιημένα δεδομένα, με βάση τη διάσταση στην οποία έγινε η ομαδοποίηση. Αν θέλουμε να ανέβουμε περισσότερο στην ιεραρχία εκτελούμε ξανά τη πράξη. (Παράδειγμα: στην διάσταση Τόπος ανεβαίνουμε από το επίπεδο Πόλη στο επίπεδο Νομός, σχήμα 2.6).



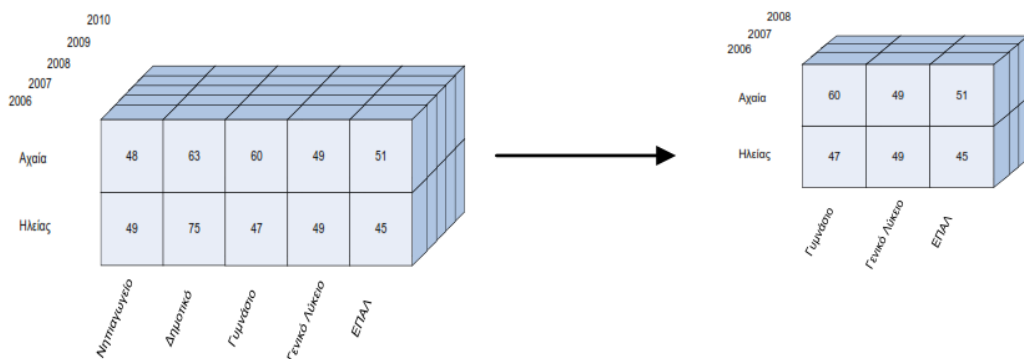
Σχήμα 2.6 : Η Πράξη Roll-up (Ravasopoulos G. et al, 2009)

• **Drill-down:** Είναι η αντίστροφη πράξη του roll-up, όπου εκτελείται ένα βήμα καθόδου από ένα υψηλότερο επίπεδο της ιεραρχίας μιας διάστασης σε ένα χαμηλότερο. Με την εφαρμογή της συγκεκριμένης λειτουργίας, ο κύβος επιστρέφει αποτελέσματα με μεγάλο βαθμό λεπτομέρειας. Επίσης, η λειτουργία αυτή παρέχει τη δυνατότητα στον αναλυτή να διατρέξει ακόμη και ολόκληρη την ιεραρχία μιας κλάσης δεδομένων και να φτάσει στο χαμηλότερο επίπεδο λεπτομέρειας. (Παράδειγμα: στην διάσταση Τόπος κατεβαίνουμε από το επίπεδο Νομός στο επίπεδο Πόλη, σχήμα 2.7).



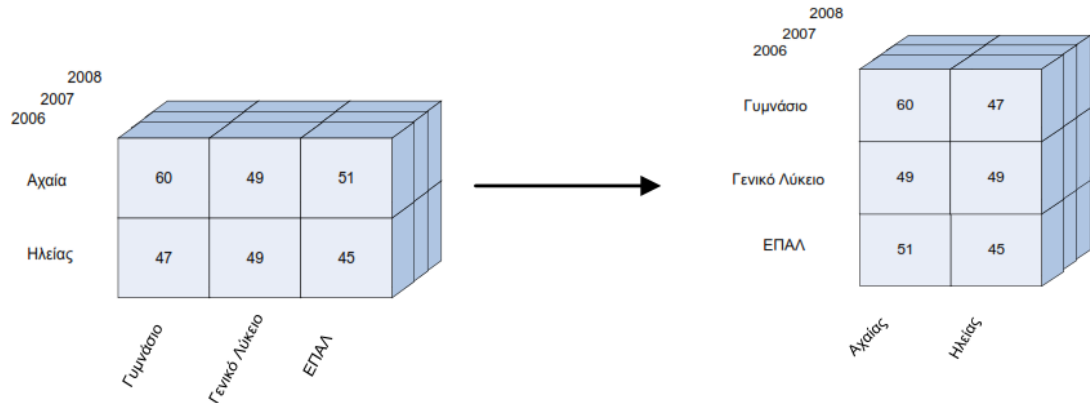
Σχήμα 2.7 : Η πράξη Drill-down (Ravasopoulos G. et al, 2009)

• **Slice:** Πρόκειται για λειτουργία επιλογής δεδομένων σε μία συγκεκριμένη διάσταση. Ένα επίπεδο (slice) είναι ένα υποσύνολο ενός υπερκύβου σύμφωνα με μία περιοχή τιμών ή μια συγκεκριμένη τιμή ενός επιπέδου διάστασης (οριζόντιος τεμαχισμός). Ουσιαστικά, η συγκεκριμένη λειτουργία φιλτράρει τα αποτελέσματα που δίνει ο κύβος ως προς μια διάστασή του. (σχήμα 2.8).



Σχήμα 2.8 : Η πράξη Slice (Ravasopoulos G. et al, 2009)

• **Pivot:** Είναι μια πράξη με την οποία αλλάζει η διάταξη (περιστροφή) των διαστάσεων ώστε να διευκολυνθεί η ανάλυση. Κατά το pivoting δεν μεταβάλλονται ούτε αυξομειώνονται τα δεδομένα του υπερκύβου. Αυτό που αλλάζει είναι ο τρόπος παρουσίασης τους στην εφαρμογή ανάλυσης (σχήμα 2.9).



Σχήμα 2.9 : Η πράξη Pivot (Ravasopoulos G. et al, 2009)

2.6.2 Είδη Εργαλείων OLAP

Υπάρχουν αρκετά είδη εργαλείων OLAP διαθέσιμα στην αγορά, όπως:

- **ROLAP** (MicroStrategy, Mondrian)
- **MOLAP** (Hyperion Essbase, Palo)
- **DOLAP** (Brio Technologies, Business Objects, Corvu)
- **HOLAP** (Microsoft SQL Server Analysis Services)

Τα εργαλεία ROLAP (Relational OLAP tools-Σχεσιακά εργαλεία OLAP) (Chhanda R., 2009) χρησιμοποιούν σαν βάση τους το σχεσιακό μοντέλο. Εξαιτίας αυτού μπορούν να επηρεάσουν το πραγματικό κόσμο με μεγάλη ακρίβεια. Ένα πρόβλημα στη πρακτική χρήση αυτών των εργαλείων μπορεί να προκύψει εάν οι κατασκευαστές των εργαλείων αυτών αποφασίσουν να εφαρμόσουν μια εξειδίκευση στο προϊόν τους. Αυτό μπορεί να οδηγήσει στη δυσκολία πραγματοποίησης τυποποιημένων γλωσσών ερωτημάτων οι οποίες είναι ήδη γνωστές στους χρήστες και άλλες μορφές αλληλεπίδρασης με τη βάση δεδομένων.

Η πολυδιάστατη OLAP ή αλλιώς MOLAP (Chaudhuri S., Dayal U., Ganti V., 2001) είναι ένα σύνολο εργαλείων τα οποία χρησιμοποιούν τη δική τους βάση δεδομένων με μία πολυδιάστατη δομή. Ένα από τα χαρακτηριστικά τέτοιου είδους βάσεων δεδομένων είναι ότι στηρίζονται σε ένα κοινό πρότυπο όπως το σχεσιακό μοντέλο. Εξαιτίας αυτού τα εργαλεία MOLAP είναι λιγότερο προσαρμόσιμα στις αλλαγές του μεγέθους ενός επιχειρησιακού συστήματος. Επίσης όλα τα στοιχεία

του συστήματος θεωρούνται ισότιμα και αντιμετωπίζονται με τον ίδιο τρόπο(δεν είναι εφικτό να αναγνωριστεί καμία συγκεκριμένη ιεραρχία μεταξύ τους). Το πλεονέκτημα των εργαλείων MOLAP είναι η ταχύτητα τους. Το μειονέκτημά τους έγκειται στην ανάγκη τους για ένα μεγάλο μέγεθος του χώρου της μνήμης (λόγω του μεγάλου αριθμού των διαστάσεων).

Τα εργαλεία Dolar (Desktop OLAP) (Rizzi S., 2003) είναι εργαλεία OLAP για επιτραπέζιους υπολογιστές. Αυτά είναι απλά εργαλεία OLAP, με χαμηλότερες τιμές, τα οποία μπορούν να προσαρμοστούν για τη χρήση σε επιτραπέζιους υπολογιστές με μικρή χωρητικότητα μνήμης και τις επιλογές απόδοσης. Τα εργαλεία Dolar τυπικά παρέχουν στήριξη μόνο σε μικρές εφαρμογές που έχουν σχεδιαστεί για να ανταποκρίνονται στις ανάγκες των διάφορων χρηστών. Υπάρχουν επίσης εκδόσεις αυτών των εργαλείων και για φορητούς υπολογιστές.

Τα εργαλεία HOLAP (Hybrid OLAP) (Cardenas N., 2011) είναι υβριδικό προϊόν με το οποίο είναι δυνατόν η ταυτόχρονη διενέργεια πολυδιάστατης ανάλυσης των δεδομένων από την αποθήκη δεδομένων και δεδομένων από το σύστημα διαχείρισης της σχεσιακής βάσης δεδομένων. Τα εργαλεία HOLAP συνδυάζουν τα πλεονεκτήματα των εργαλείων MOLAP και ROLAP. Κάποιες από τις καλύτερες κατασκευάστριες εταιρείες εργαλείων OLAP είναι η Microsoft, η Oracle και η SAS.

3. Κατασκευή Κύβων OLAP για την Εκπαίδευση

Θέλοντας να ελεγχθεί κατά πόσο είναι δυνατόν να επεξεργασθεί ένας μεγάλος όγκος δεδομένων σύντομα και συνοπτικά, χρησιμοποιούνται ερωτήματα που θέτονται σε μια πολυδιάστατη βάση δεδομένων. Για να απαντηθούν τα ερωτήματα αυτά, κατασκευάζεται κύβος OLAP, ο οποίος είναι μια τεχνολογία επερωτήσεων σε βάσεις δεδομένων, που στηρίζεται στη θεώρηση της πληροφορίας με πολυδιάστατο τρόπο, ώστε να εξοικονομηθεί χρόνος επεξεργασίας των δεδομένων. Γι' αυτό τον λόγο, χρησιμοποιείται μια μοντελοποίηση της πολυδιάστατης πληροφορίας με σκοπό να επιτευχθεί υψηλή επίδοση στην ποιότητα της επεξεργασίας ερωτημάτων και εξαγωγής των αντίστοιχων αποτελεσμάτων.

Στο κεφάλαιο αυτό γίνεται μια εμπειρική προσέγγιση της πολυδιάστατης ανάλυσης μέσω του λογισμικού στατιστικών και αναφορών OLAP (OLAP Statistics & Reporting for Microsoft Access) και δημιουργούνται κύβοι OLAP οι οποίοι βασίζονται σε βάση δεδομένων της εκπαίδευσης. Επιλέχθηκε η χρήση του συγκεκριμένου λογισμικού μιας και η βάση στην οποία στηρίζεται η ανάλυση μας είναι σε μορφή Microsoft Access και το λογισμικό αυτό είναι κατασκευασμένο ώστε να πραγματοποιεί την ανάλυση τέτοιου είδους βάσεων δεδομένων. Με αυτό τον τρόπο γίνεται πρακτικά δυνατό να ελεγχθεί κατά πόσο είναι δυνατόν να λάβουμε τις επιθυμητές απαντήσεις στα ερωτήματα που τίθενται σε αυτή τη βάση δεδομένων. Επίσης εξετάζουμε αν αυτά τα ερωτήματα ανταποκρίνονται στις ανάγκες και στις επιθυμίες ενός χρήστη που ενδιαφέρεται να κάνει μια τέτοιου είδους ανάλυση.

3.1 Σχεδιασμός και Ανάπτυξη Μοντέλου OLAP

3.1.1 Μελέτη Περίπτωσης

Σε αυτή την εργασία γίνεται εφαρμογή ενός OLAP μοντέλου προκειμένου να παρουσιασθεί πως αυτό γίνεται χρήσιμο στο τομέα της εκπαίδευσης. Το μοντέλο OLAP είναι ενδεικτικό καθώς η παραμετροποίηση που μπορεί να δεχθεί είναι ιδιαίτερα μεγάλη ανάλογα και με τις εκάστοτε απαιτήσεις. Επίσης είναι ένα δείγμα του πόσο χρήσιμα είναι τέτοιου είδους μοντέλα για την εξαγωγή

συμπερασμάτων και πληροφοριών που αφορούν στην εκπαίδευση και έχουν σαν σκοπό την βελτίωση του εκπαιδευτικού συστήματος.

3.1.2 Η Βάση Δεδομένων

Η βάση που χρησιμοποιήθηκε προκειμένου να γίνει το πρακτικό κομμάτι της εργασίας είναι της εταιρείας «Έπαφος» η οποία περιέχει αρκετά δεδομένα ώστε να δημιουργηθεί ένα ολοκληρωμένο περιβάλλον διαχείρισης για τα σχολεία της Δευτεροβάθμιας Εκπαίδευσης.

Όλο το σύστημα αποτελείται από διαφορετικά συστήματα εφαρμογών που αναφέρονται σε όλες τις δραστηριότητες του σχολείου:

- Διαχείριση των καθημερινών λειτουργιών: Στοιχεία μαθητών, Βαθμολογία, Απουσίες, Υπολογισμός και Έκδοση Αποτελεσμάτων, Εκτύπωση καταστάσεων, Έλεγχοι προόδου, Πιστοποιητικά, Απολυτήρια
- Στατιστικά στοιχεία
- Ακαδημαϊκή κατάρτιση των καθηγητών

Όλο το σύστημα είναι πλήρως παραμετρικό, αρκετά εύκολο στη χρήση του και σχετικά πλήρες όσον αφορά τη βάση δεδομένων που διαχειρίζεται. Το δ-ΒΑΣΗ χρησιμοποιείται σε περισσότερα από 2000 σχολεία στην Ελλάδα και λόγω του γεγονότος ότι είναι σταθερό και προσιτό στη χρήση του, επεκτείνεται και σε άλλα σχολεία σε όλη την χώρα. Τα στοιχεία που επιλέχθηκαν για την κατασκευή του OLAP μοντέλου είναι ενδεικτικά για τη λειτουργία του και τη χρησιμότητα του.

3.1.3 Μοντέλο OLAP

Το επόμενο βήμα είναι ο σχεδιασμός και η ανάπτυξη μιας Αποθήκης δεδομένων, μιας πολυδιάστατης βάσης δεδομένων. Για την ανάγνωση της διαθέσιμης βάσης δεδομένων χρησιμοποιήθηκε το λογισμικό Microsoft Access καθώς πάνω σε αυτό έχει δημιουργηθεί..

Η υλοποίηση του OLAP μοντέλου, η ανάλυση των δεδομένων και η παρουσίαση των αποτελεσμάτων γίνεται με τη χρήση του προγράμματος **OLAP Statistics & Reporting for Microsoft Access** το οποίο αποτελείται από τον OLAP Access Manager και τον OLAP Statistics and Reporting Client . Όπως

σε αυτό πρόγραμμα έτσι και σε άλλα προγράμματα η απαιτούμενη διαδικασία για την υλοποίηση διαφέρει αλλά τα αποτελέσματα παραμένουν ίδια.

Οι προσεγγίσεις των δυο προγραμμάτων έδωσαν τη δυνατότητα να εξαχθούν τόσο οπτικά όσο και σε επίπεδο δεδομένων όλα τα ζητούμενα αποτελέσματα τα οποία θα παρουσιαστούν στη συνέχεια. Στις δυο αυτές περιπτώσεις οι απαιτήσεις είναι ιδιαίτερες, τόσο όσον αφορά το χειρισμό αλλά και τη σωστή διατύπωση των ερωτημάτων που ζητούν απάντηση προκειμένου τα αποτελέσματα να ευσταθούν και να είναι χρήσιμα.

3.2 Σκοπός του Λογισμικού

Το συγκεκριμένο λογισμικό είναι μια εφαρμογή OLAP που κάνει πολύ εύκολη και χωρίς μεγάλη προσπάθεια την ανάλυση της βάσης δεδομένων Access σε οπτικές πολλαπλών διαστάσεων και μας δίνει χρήσιμες πληροφορίες και εικόνα των σχεσιακών δεδομένων.

Ο σκοπός είναι μέσω του λογισμικού να φτιαχτεί ένα μοντέλο OLAP το οποίο να στηριχθεί στην υπάρχουσα βάση δεδομένων η οποία περιέχει εκπαιδευτικά δεδομένα και να παρέχει πληροφορίες σχετικές με: την απόδοση των μαθητών με βάση γεωγραφικά στοιχεία, την παρακολούθηση των μαθητών, την επίδοση των μαθητών, τα δημογραφικά των μαθητών, το επίπεδο σπουδών και το είδος της θέσης των καθηγητών. Ποιο συγκεκριμένα το λογισμικό αναμένεται να απαντά στα παρακάτω ερωτήματα:

- Ποιος είναι ο μέσος όρος βαθμολογίας των μαθητών σε σχέση με την γεωγραφική κατανομή και την τάξη τους
- Ποια είναι η βαθμολογία των Ελλήνων μαθητών ανά φύλο και σχολικό έτος
- Ποια είναι η βαθμολογία των αλλοδαπών μαθητών που βρίσκονται στην Ελλάδα, ανά φύλο, εθνικότητα και σχολικό έτος
- Ποιο είναι το πλήθος των μαθητών που πήραν βαθμό 19 ανά μάθημα και ανά περιοχή
- Πόσες αδικαιολόγητες απουσίες έχουν οι μαθητές ανά τάξη και ανά φύλο
- Ποιο είναι το ποσοστό των δικαιολογημένων απουσιών ανά φύλο και περιοχή των μαθημάτων
- Μαθητές(ονομαστικά) με αυξημένες δικαιολογημένες απουσίες ανά έτος

- Μαθητές (ονομαστικά) που δεν προβιβάστηκαν λόγω απουσιών ανά σχολικό έτος και ανά τάξη
- Ποιος είναι ο μέσος όρος των βαθμολογιών των μαθητών ανά τάξη
- Αριστούχοι μαθητές ανά φύλο και σχολική χρονιά, ονομαστικά
- Ποιο είναι το πλήθος των μαθητών που πήραν βαθμολογίες του 10 ή κάτω από την βάση ανά τάξη και σχολική χρονιά
- Ποιος είναι ο γενικός μέσος όρος των γενικών μαθημάτων ανά τάξη ανά σχολικό έτος
- Ποιος είναι ο γενικός μέσος όρος των μαθημάτων ξένων γλωσσών ανά τάξη και ανά σχολικό έτος
- Διαχωρισμός μαθητών ανά περιοχή και τάξη
- Ποιος είναι ο αριθμός των καθηγητών με μεταπτυχιακές σπουδές ανά φύλο και ειδικότητα
- Πόσοι ήταν οι καθηγητές ανά σχολική χρονιά με βάση το είδος της θέσης τους
- Πόσες είναι οι ώρες μαθήματος ανά είδος θέσης και ανά φύλο

Το πρόγραμμα στατιστικών και αναφορών OLAP (OLAP Statistics & Reporting) είναι κατάλληλο για το σκοπό αυτό. Επιτρέπει να ορισθεί οποιοδήποτε πεδίο ως μέτρο με διαφορετικές συναρτήσεις, όπως άθροισμα, καταμέτρηση, διακριτές μετρήσεις, μέγιστο και ελάχιστο ,αντίθετα με τις στατιστικές διαδικασίες που εκτελούνται. Με αυτό τον τρόπο δίνεται η δυνατότητα ευκολότερης διερεύνησης νέων τάσεων και ανακάλυψης άγνωστων προβλημάτων που προκύπτουν από την ροή των δεδομένων.

3.3 Τρόπος Λειτουργίας του Λογισμικού

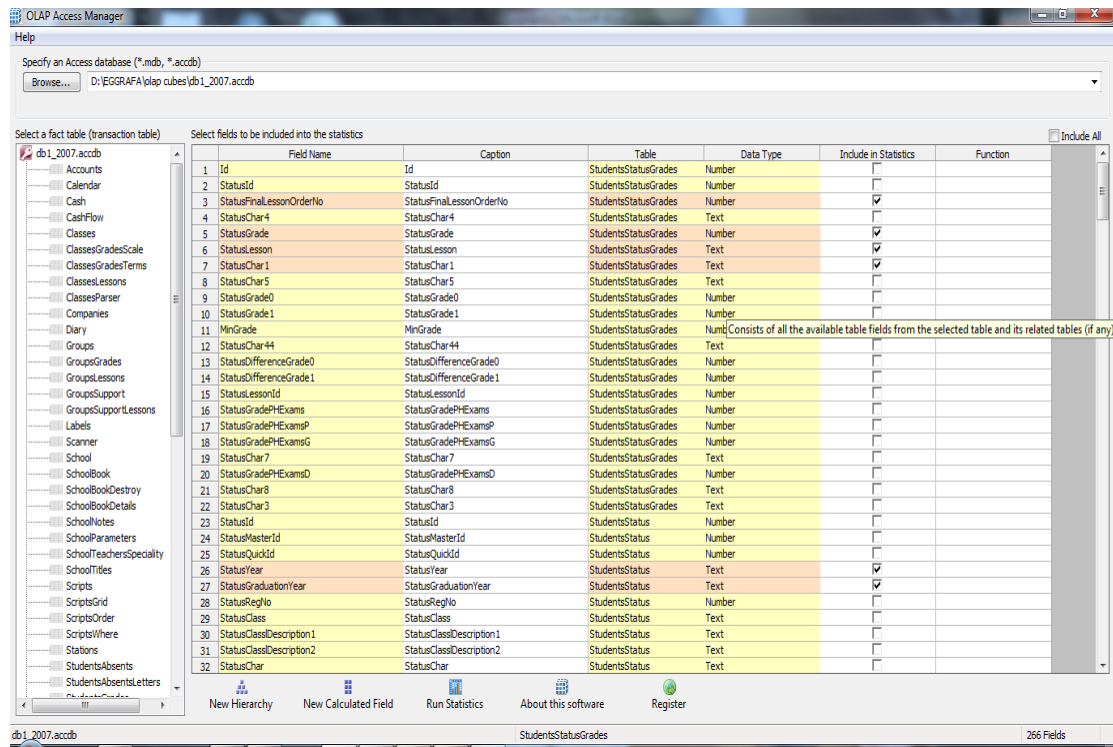
3.3.1 Εισαγωγή Δεδομένων της Εκπαίδευσης στο OLAP Access Manager

Από το πρόγραμμα OLAP Access Manager και πατώντας την επιλογή Browse εντοπίζεται και συνδέεται αρχικά η βάση δεδομένων Access db2007.accdb. Μόλις πραγματοποιηθεί η σύνδεση της βάσης δεδομένων εμφανίζονται στην αριστερή στήλη του περιβάλλοντος του προγράμματος οι περιεχόμενοι πίνακες οι οποίοι είναι διαθέσιμοι για την ανάλυση μας (Εικόνα 3.1).

3.3.2 Επεξεργασία Δεδομένων και Κατασκευή Κύβων

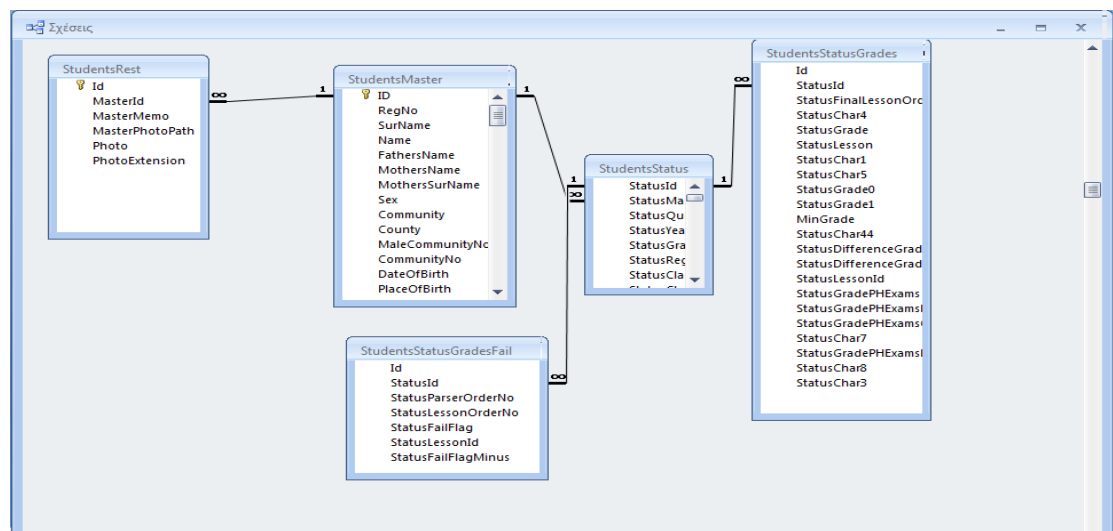
Στη συνέχεια, επιλέγεται ένας πίνακας γεγονότων ή συναλλαγών, έτσι ώστε να εμφανιστούν όλα τα διαθέσιμα πεδία που ορίζονται για αυτόν τον πίνακα (και τους σχετικούς πίνακες πηγής (source table) μέσω του ξένου κλειδιού (foreign key)).

Με αυτόν τον τρόπο επιλέγεται ένας συνδυασμός συνόλων δεδομένων, έτσι ώστε να προκύψουν χρήσιμα δεδομένα και συμπεράσματα για την εκπαίδευση. Στα αριστερά γίνεται η επιλογή ενός πίνακα γεγονότων ή πίνακα συναλλαγών και στα δεξιά εμφανίζονται τα πεδία του επιλεγμένου πίνακα και τα πεδία των σχετικών πινάκων μέσω του ξένου κλειδιού. Κατόπιν, μπορούμε να επιλέξουμε τα πεδία που μπορούν να συμπεριληφθούν στη στατιστική μας ανάλυση, όπως βλέπουμε στην εικόνα 3.1.



Εικόνα 3.1: Επιλογή βάσης δεδομένων και εμφάνιση περιεχομένου της.

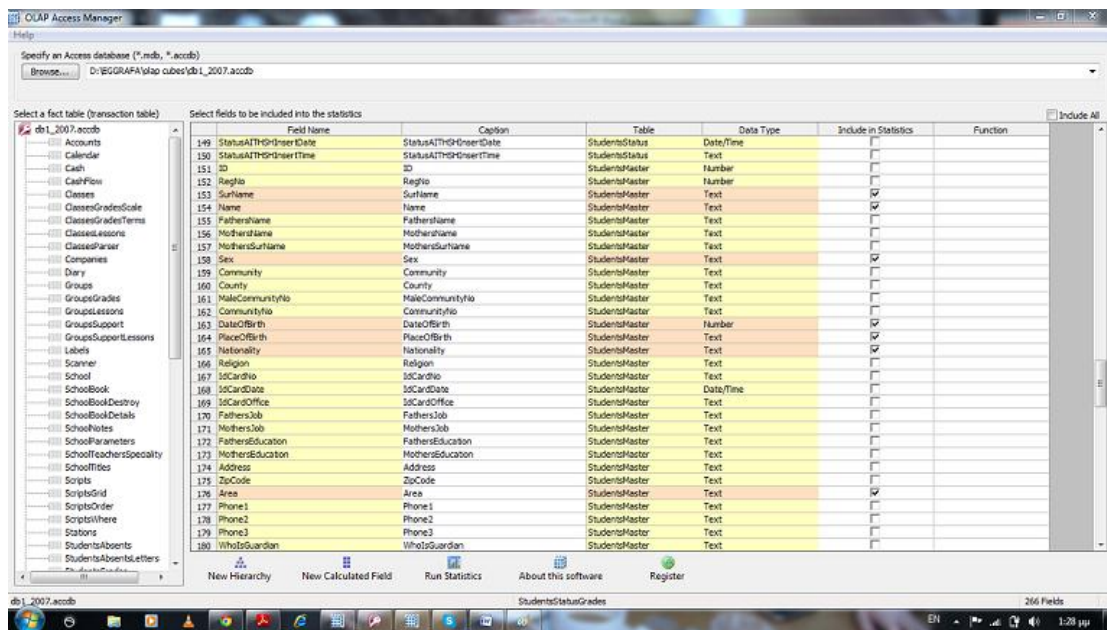
Συνδέουμε τον πίνακα συναλλαγών «StudentMaster» της βάσης δεδομένων μας, του οποίου το σχήμα συσχετίσεων φαίνεται στην εικόνα 3.2.



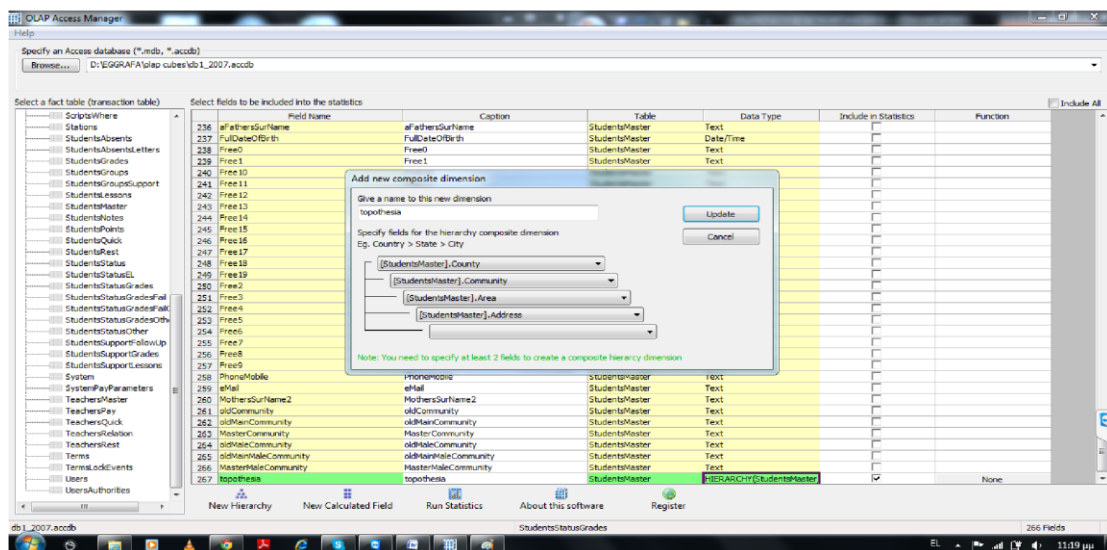
Εικόνα 3.2 : Διάγραμμα οντοτήτων συσχετίσεων (ER diagram)

Στο λογισμικό OLAP Access Manager, παρατηρείται ότι όλα τα πεδία που συσχετίζονται με τους άλλους πίνακες, συνδέονται μέσω ξένων κλειδιών, όπως StudentMaster, StudentRest,(εικόνα 3.2) κλπ. και εμφανίζονται αυτόματα, ώστε να χρησιμοποιηθούν στην κατασκευή του κύβου.

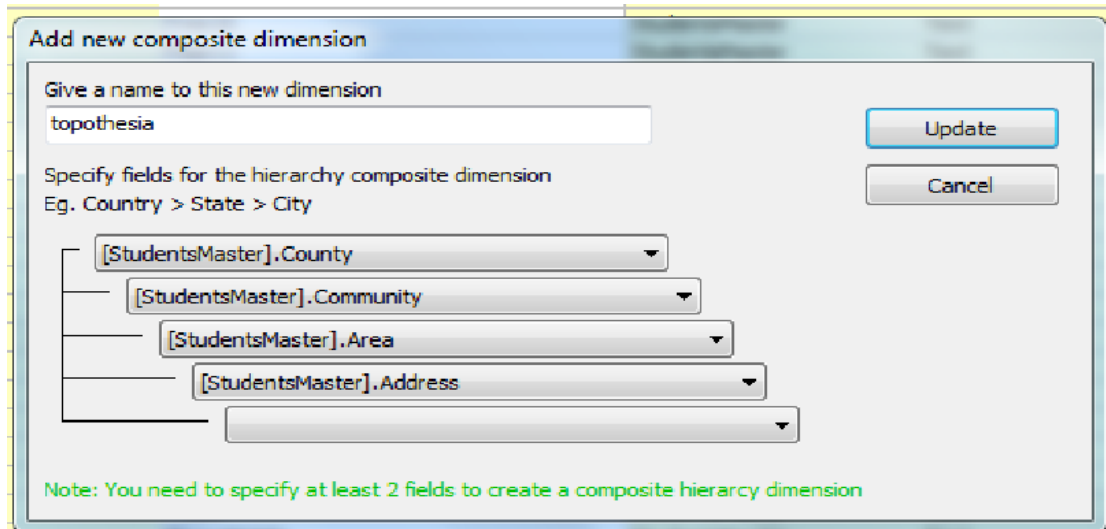
Αφού επιλεγούν ποία πεδία ή διαστάσεις πρέπει να συμπεριληφθούν για τις στατιστικές αναλύσεις (εικόνα 3.3), υπάρχει η επιλογή λειτουργιών για πεδία που λειτουργούν σαν μέτρα στο κύβο OLAP, όπως η στατιστική ανάλυση η οποία μπορεί να αναπαράγεται από άλλα πεδία, βασισμένη στην τιμή των πεδίων «μέτρου» (measure fields).



Εικόνα 3.3 : Επιλογή πεδίων από τους συσχετιζόμενους πίνακες για δημιουργία κύβου



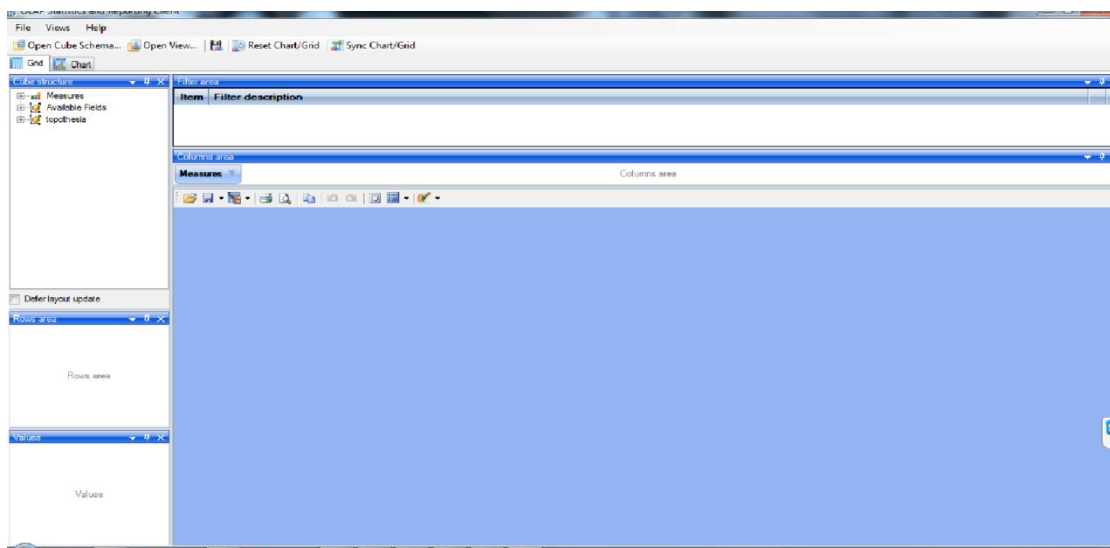
Εικόνα 3.4 : Κατασκευή καινούργιας Ιεραρχίας



Εικόνα 3.5 : Κατασκευή γεωγραφικής Ιεραρχίας

Δημιουργούμε μια ιεραρχία με όνομα «topothesia», όπου απεικονίζεται ο κλάδος των μεταβλητών «Country», «Community», «Area», και «Address» (εικόνες 3.4 και 3.5), θέλοντας να συμβολίσουμε τη τοποθεσία που κάποιος από τους μαθητές βρίσκεται. Αν προβούμε σε «ενημέρωση» (update) της ιεραρχίας, προσθέτεται ένα καινούργιο πεδίο, που βρίσκεται στο τέλος του πίνακα γεγονότος

Αφού κάνουμε τη παραπάνω διαδικασία διαλέγουμε την επιλογή «Run Statistics» και ένα αρχείο σχήματος κύβου (cube schema file) δημιουργείται και τροφοδοτεί τον εξυπηρετητή OLAP (OLAP Statistics and Reporting client) για την επεξεργασία και την εξαγωγή του κύβου από τη βάση δεδομένων (εικόνα 3.6).



Εικόνα 3.6 : Περιβάλλον του «Olap Statistics and Reporting Client»

3.4 Παραδείγματα Κύβων OLAP

3.4.1 Δημιουργία Ερωτημάτων

Σύμφωνα με τη διαδικασία που αναφέρθηκε παραπάνω κατασκευάστηκαν κύβοι με δεδομένα της εκπαίδευσης από τη δοθείσα βάση δεδομένων. Δημιουργήθηκαν ερωτήματα που βοηθούν στην ενδοσκόπηση της λειτουργίας του εκπαιδευτικού συστήματος, στις κατηγορίες που αναφέρονται στους μαθητές, στους καθηγητές και στα μαθήματα τα οποία διδάσκονται. Αναλύθηκαν αυτές οι κύριες κατηγορίες σε μικρότερες, όπου ασχολούνται με τις επιδόσεις των μαθητών στα μαθήματα τους, τον αριθμό των απουσιών τους, τη χώρα καταγωγής τους, το τόπο διαμονής τους κ.α. Επίσης κατασκευάστηκαν ερωτήματα δεδομένων που αφορούν τους καθηγητές διερευνώντας το επίπεδο σπουδών τους και θέσης που κατέχουν. Από την κατασκευή των κύβων, κατασκευάστηκαν ερωτήματα που περιγράφονται παρακάτω και οδήγησαν σε χρήσιμα συμπεράσματα για το χώρο της εκπαίδευσης.

ΕΡΩΤΗΜΑΤΑ

ΜΑΘΗΤΕΣ:

ΑΠΟΔΟΣΗ ΜΑΘΗΤΩΝ ΜΕ ΒΑΣΗ ΓΕΩΓΡΑΦΙΚΑ ΣΤΟΙΧΕΙΑ

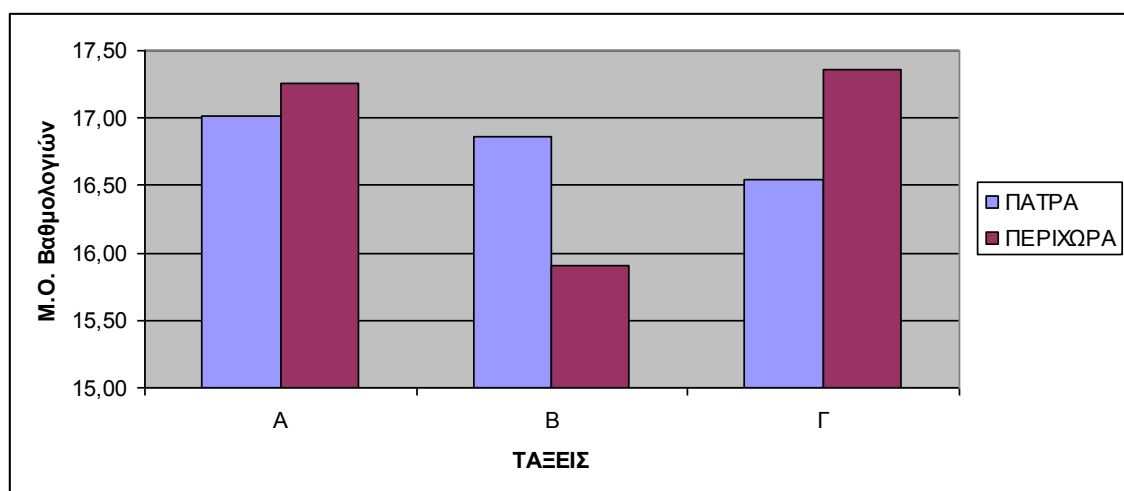
1. Ποιος είναι ο μέσος όρος βαθμολογίας των μαθητών σε σχέση με την γεωγραφική κατανομή και την τάξη τους

Σκοπός της διερεύνησης αυτού του ερωτήματος είναι να αναδειχθούν οι επιπτώσεις της γεωγραφικής περιοχής -εάν υπάρχουν- πάνω στις βαθμολογικές επιδόσεις των μαθητών για κάθε τάξη. Είναι σύνηθες φαινόμενο πιο απομακρυσμένες περιοχές από τα αστικά κέντρα ή περιοχές στις οποίες τα σχολεία που λειτουργούν είναι δυσπρόσιτα από τους μαθητές και το διδακτικό προσωπικό να παρουσιάζουν αποκλίσεις στην επίδοσή τους.

Αυτό μπορεί να ερμηνευθεί με διάφορους τρόπους. Η έλλειψη ενδιαφέροντος μπορεί να πηγάζει από το γεγονός ότι η δύσκολη πρόσβαση στο

σχολικό συγκρότημα κουράζει τους μαθητές με αποτέλεσμα πολλοί από αυτούς να ολοκληρώνουν μόνο την υποχρεωτική εκπαίδευση. Τα ερεθίσματα για τους μαθητές ποικίλουν και στη πλειοψηφία τους οι μαθητές έχουν διάφορες εξωσχολικές δραστηριότητες, οι οποίες τους απασχολούν αρκετές ώρες και αυτό αντικατοπτρίζεται και στις βαθμολογικές επιδόσεις όλων των σχολικών τους χρόνων. Επίσης, για πολλά άτομα του διδακτικού προσωπικού συνήθως οι μετακινήσεις σε τέτοια σχολεία είναι προσωρινές και στην πλειοψηφία τους αποσπάσεις ή νέοι διορισμοί προσωπικού. Αυτό σημαίνει ότι ένας κατά τεκμήριο νέος εκπαιδευτικός καλείται να αντιμετωπίσει μια κατάσταση για την οποία δεν ευθύνεται ούτε ο ίδιος αλλά ούτε και οι μαθητές. Ιδιαίτερα σε ορεινές περιοχές δεν θα ήταν υπερβολή να πει κανείς ότι η απλή μετάβαση και επιστροφή από το σχολείο την περίοδο του χειμώνα είναι από μόνο του ως γεγονός μια πρόκληση. Το ίδιο βέβαια ισχύει και για τον καθηγητή. Αυτό έχει ως αποτέλεσμα τη σωματική και ψυχολογική καταπόνηση όλων των εμπλεκόμενων μερών στην εκπαιδευτική διαδικασία και στη μείωση του βαθμού απόδοσης τόσο του εκπαιδευτικού όσο και των μαθητών σε μακροχρόνια περίοδο.

Επιπλέον στα αστικά κέντρα η συντριπτική πλειοψηφία των μαθητών παραμένει αναπόσπαστη στο ρόλο του μαθητή καθώς δεν τους ζητείται από την οικογένεια κανενός είδους εργασία ή συνεισφορά στις επαγγελματικές υποχρεώσεις της σε αντίθεση με πιο απομακρυσμένες περιοχές όπου οι μαθητές βοηθούν στις αγροτικές αλλά και οποιαδήποτε άλλη εργασία των γονέων τους κάτι που μειώνει το χρόνο αφοσίωσης τους στα σχολικά πράγματα και κατά συνέπεια μειώνει και την απόδοση τους.

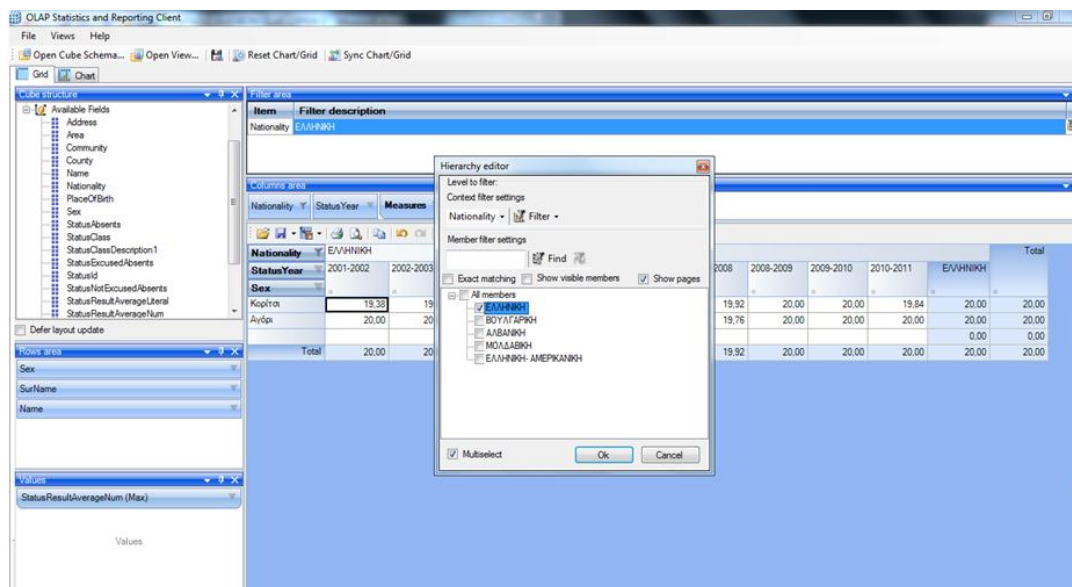


Εικόνα 3.7 : Μέσος όρος Βαθμολογιών ανά τάξη και περιοχή

Σε αντίθεση με τα αναμενόμενα η ανάλυση της βάσης των δεδομένων μας (εικόνα 3.7) δείχνει ότι οι μαθητές στα περιχώρα έχουν υψηλότερες βαθμολογίες σε 2 από τις 3 τάξεις του γυμνασίου. Επίσης οι βαθμολογίες κυμαίνονται σε ιδιαίτερα υψηλά επίπεδα κάτι που δείχνει ότι το επίπεδο των εκπαιδευτικών ιδρυμάτων είναι ιδιαίτερα ικανοποιητικό.

2. Ποιά είναι η βαθμολογία των Ελλήνων μαθητών ανά φύλο και σχολικό έτος.

Χρησιμοποιήθηκε φίλτρο για να απομονώσουμε τους Έλληνες μαθητές και την επίδοσή τους ανά σχολική χρονιά (εικόνα 3.8).



Εικόνα 3.8 : Επιλογή μαθητών με Ελληνική υπηκοότητα

Filter area														
Item	Filter description													
Nationality	ΕΛΛΗΝΙΚΗ													
Columns area														
Nationality	ΕΛΛΗΝΙΚΗ													
StatusYear	2001-2002	2002-2003	2003-2004	2004-2005	2005-2006	2006-2007	2007-2008	2008-2009	2009-2010	2010-2011	ΕΛΛΗΝΙΚΗ	Total		
Sex	Κορίτσια	19,38	19,84	19,92	19,92	19,84	19,76	19,92	20,00	20,00	19,84	20,00	20,00	
	Αγόρια	20,00	20,00	19,92	20,00	20,00	20,00	19,76	20,00	20,00	20,00	20,00	20,00	
	Total	20,00	20,00	19,92	20,00	20,00	20,00	19,92	20,00	20,00	20,00	20,00	20,00	20,00

Εικόνα 3.9 : Διαχωρισμός μαθητών με Ελληνική υπηκοότητα με τον υψηλότερο μέσο όρο ανά έτος με βάση το φύλο

Παρουσιάστηκε πίνακας με τις καλύτερες βαθμολογίες μέσου όρου ανά έτος σε σχέση με το φύλο (εικόνα 3.9) στον φαίνεται ότι τα αγόρια πετυχαίνουν καλύτερη βαθμολογία από ότι τα κορίτσια.

Nationality	A		B		Γ		ΕΛΛΗΝΙΚΗ		StatusResultAverageNum (Max)
	StatusResultAverageNum (Max)	Count	StatusResultAverageNum (Max)	Count	StatusResultAverageNum (Max)	Count	StatusResultAverageNum (Max)	Count	
Κορίτσια	19.92	286	20.00	269	20.00	285	20.00	840	20.00
Αγόρια	20.00	286	20.00	276	20.00	290	20.00	854	20.00
Total	0.00	1	0.00	1	0.00	1	0.00	1	0.00
Total	20.00	573	20.00	547	20.00	575	20.00	1.695	20.00

Εικόνα 3.10: Αριθμός Ελλήνων μαθητών ανά φύλο και τάξη που πέτυχαν τους υψηλότερους μέσους όρους

Τα παιδιά που έχουν ελληνική καταγωγή εμφανίζονται ανά φύλλο σε σχέση με τις επιδόσεις τους σε κάθε τάξη του γυμνασίου (εικόνες 3.10 και 3.11). Στην Α γυμνασίου παρατηρείται μια μικρή πτώση σε σχέση με τις άλλες τάξεις σε κλίμακα καλύτερης βαθμολογίας μέσου όρου. Επίσης παρατηρούμε ότι οι μαθητές στη Β Γυμνασίου ανεξαρτήτου φύλλου που άριστευσαν είναι λιγότεροι σε σχέση με τις άλλες δύο τάξεις του Γυμνασίου.

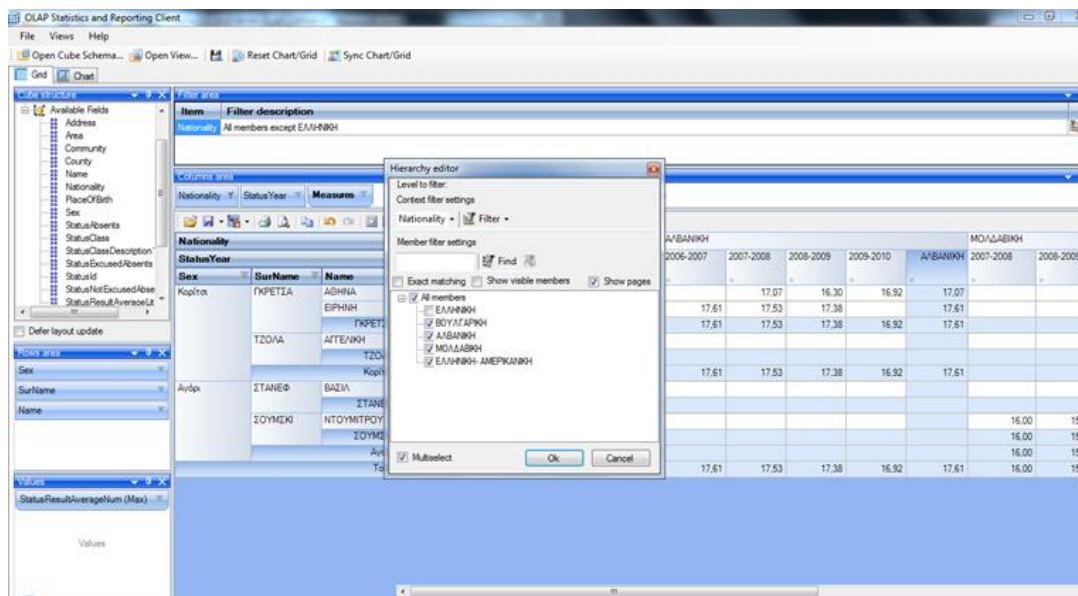
Nationality	2005-2006		2006-2007		2007-2008		2008-2009		2009-2010	
	StatusResultAverageNum (Max)	Count	StatusResultAverageNum (Max)	Count	StatusResultAverageNum (Max)	Count	StatusResultAverageNum (Max)	Count	StatusResultAverageNum (Max)	Count
Κορίτσια	19.69	28	19.76	31	19.69	29	20.00	29	19.61	
Αγόρια	20.00	32	19.61	31	19.76	34	19.92	29	20.00	
Total	20.00	60	19.76	62	19.76	63	20.00	58	20.00	

Εικόνα 3.11 : Αριθμός Ελλήνων μαθητών ανά φύλο για τη Β Γυμνασίου που πέτυχαν τους υψηλότερους μέσους όρους

Από το 2002 έως το 2011 κατά μέσο όρο έχουμε 60 αριστούχους ανά σχολικό έτος με ελαφρά κυριαρχία των αγοριών να παίρνουν καλύτερους βαθμούς. Στα έτη 2005-2006, 2008-2009 και 2009-2010 γίνεται εμφάνιση του βαθμού 20 ως καλύτερος βαθμός του μέσου όρου.

3. Ποια είναι η βαθμολογία αλλοδαπών μαθητών που βρίσκονται στην Ελλάδα, ανά φύλο, εθνικότητα και σχολικό έτος.

Επιλέγονται ως δεδομένα γραμμής (row area), το φύλλο (κορίτσι ή αγόρι), το επίθετο και το όνομα κάθε μαθητή. Ως τιμή (value) επιλέγεται το πεδίο «Status Result Average Num», όπου περιέχει τον ετήσιο μέσο όρο των βαθμολογιών των μαθημάτων του κάθε μαθητή και χρησιμοποιείται επιπλέον η συνάρτηση Max, όπου εμφανίζει το μέγιστο κάθε μέσο όρου. Ως στήλη (Column area), δηλαδή κατακόρυφα, επιλέχθηκαν η Εθνικότητα και η χρονιά ως εργαλεία μέτρησης.



Εικόνα 3.12 : Κύβος με τα στοιχεία των μαθητών, τους καλύτερους μέσους όρους τους και επιλογή μαθητών με ξένη υπηκοότητα

Τοποθετούμε φίλτρο, ώστε να εξαιρέσουμε όλα τα παιδιά που έχουν καταγωγή από την Ελλάδα και με αυτό το τρόπο παρατηρούμε τις επιδόσεις των παιδιών που έχουν έρθει στην Ελλάδα από άλλες χώρες (εικόνα 3.12). Παρατηρούμε ότι στο συγκεκριμένο παράδειγμα στην κατηγορία των κοριτσιών υπάρχουν δυο κορίτσια με το ίδιο επίθετο «Γκρέτσα» και είναι η «Αθηνά» και η «Ειρήνη», οι οποίες κατάγονται από την Αλβανία. Βλέπουμε ότι η «Γκρέτσα Ειρήνη» ήταν στο γυμνάσιο τα έτη 2006-2007,2007-2008,2008-2009 με μέσους όρους βαθμολογίας 17.61, 17.53 και 17.38 αντίστοιχα και τη χρονιά 2006-2007 πέτυχε την καλύτερη βαθμολογία της. Η αδερφή της «Γκρέτσα Αθηνά» ήταν στο γυμνάσιο τα έτη 2007-2008,2008-2009, 2009-2010 με μέσο όρο κατά έτος 17.07,

16.3, 16.9. Τη καλύτερη βαθμολογία φοίτησης την πέτυχε στην Α γυμνασίου (εικόνα 3.13).

Nationality	StatusYear	ΑΛΒΑΝΙΚΗ				ΜΟΛΔΑΒΙΚΗ			
		2006-2007	2007-2008	2008-2009	2009-2010	2006-2007	2008-2009	2009-2010	ΜΟΛΔΑΒΙΚΗ
Κορίτσια	ΓΚΡΕΤΣΑ	ΑΘΗΝΑ	17.07	16.30	16.92	17.07			
		ΕΙΡΗΝΗ	17.61	17.53	17.38	17.61			
	ΤΖΟΛΑ	ΑΓΓΕΛΙΚΗ	17.61	17.53	17.38	16.92	17.61		
		ΤΖΟΛΑ							
Αγόρια	ΣΤΑΝΕΦ	ΒΑΣΙΛ	17.61	17.53	17.38	16.92	17.61		
		ΣΤΑΝΕΦ							
	ΣΟΥΜΣΚΙ	ΝΤΟΥΜΤΡΟΥ					15.00	15.00	15.38
		ΣΟΥΜΣΚΙ					16.00	15.00	15.38
Αγόρια						16.00	15.00	15.38	
Total		17.61	17.53	17.38	16.92	17.61	16.00	15.00	15.38

Εικόνα 3.13 : Οι αλλοδαποί μαθητές με τον υψηλότερο μέσο όρο βαθμολογίας ανά φύλο, εθνικότητα και χρονολογία

Επίσης γίνεται και μια επισκόπηση για την επίδοση των μαθητών από άλλες χώρες εκτός Ελλάδας. Παρατηρείτε ότι υπάρχουν παιδιά από τη Βουλγαρία, Αλβανία, Μολδαβία και Αμερική και υπάρχει μια διαφορετικότητα στις βαθμολογίες τους. Ο μαθητής που κατάγεται από την Βουλγαρία έχει τις χαμηλότερες επιδόσεις (12,92) (εικόνα 3.14), ενώ η μαθήτριά που κατάγεται από την Αμερική έχει τη καλύτερη επίδοση (19,84) από τα άλλα παιδιά η οποία είναι συγκρίσιμη με τους αντίστοιχες των Ελλήνων μαθητών. Παρατηρείται ότι ο μαθητής από την Βουλγαρία ο οποίος δεν είχε καλές επιδόσεις, δεν κατάφερε να περάσει στη τάξη της Γ γυμνασίου.

Nationality	StatusClass	StatusYear	Μέσος όρος			
			Α	Β	Γ	Μέσος όρος
Κορίτσια	Α	2006-2007	12.92	12.53	0.00	19.84
			Count: 1	Count: 1	Count: 1	Count: 3
Αγόρια	Α	2006-2007	12.92	12.53	0.00	19.84
			Count: 1	Count: 1	Count: 1	Count: 3
Total		2006-2007	12.92	12.53	0.00	19.84

Εικόνα 3.14 : Μέσος όρος ανά τάξη του μαθητή που κατάγεται από τη Βουλγαρία

4. Ποιό είναι το πλήθος των μαθητών που πήραν βαθμό 19 ανά μάθημα και ανά περιοχή

Εδώ παρουσιάζεται ο αριθμός των μαθητών που πήραν βαθμό 19 στα μαθήματα τους ανά νομό, δήμο και περιοχή αντίστοιχα (εικόνα 3.15). Παρατηρούμε ότι οι περισσότερες υψηλές βαθμολογίες έχουν επιτευχθεί στη Φυσική αγωγή και στην Αισθητική αγωγή ενώ υψηλά επίπεδα επιδόσεων φαίνεται

να έχουν οι μαθητές στα Θρησκευτικά και στην Οικιακή Οικονομία με φθίνουσα σειρά κατάταξης. Αντιθέτως στο μάθημα της Βιολογίας υπάρχει πολύ μικρός αριθμός μαθητών με υψηλή βαθμολογία, πράγμα που φανερώνει τη δυσκολία του μαθήματος. Επίσης βλέπουμε ότι οι πιο πολλοί μαθητές προέρχονται από το δήμο Μεσσήτιδος και το δήμο Πατρών και οι υπόλοιποι από γειτονικές περιοχές περιμετρικά της πόλης της Πάτρας.

StudentsMaster_County	StudentsMaster_Community	StudentsMaster_Area	Φυσική Αγωγή	Αισθητική Αγωγή	Βιολογία I	Γεωγραφία	Πληροφορική - Τεχνολογία	Οικιακή Οικονομία	Αγγλικά Αρκαρίου	Θρησκευτικά	Αριθμός Μαθητών
ΑΧΑΪΑΣ	ΜΕΣΣΑΤΩΣ		42	35	7	15	16	22	1		24
	ΠΑΤΡΩΝ	ΠΑΤΡΑ	39	17	4	10	8	8	1		20
		ΠΑΤΡΩΝ	39	17	4	10	8	8	1		20
	ΦΑΡΩΝ	ΠΑΤΡΑ	13	8	2	2	3	6			8
		ΜΠΕΓΟΥΛΑΚΙ	3	1							
		ΦΑΡΩΝ	16	9	2	2	3	6			8
	ΔΙΑΚΟΠΤΟΥ	ΠΑΤΡΑ		2			2	2			1
		ΔΙΑΚΟΠΤΟΥ		2			2	2			1
	ΠΑΡΑΛΙΑΣ	ΠΑΤΡΑ	11	9	2	3	3	1	1		4
		ΠΑΡΑΛΙΑΣ	11	9	2	3	3	1	1		4
	ΜΟΒΡΗΣ	ΠΑΤΡΑ	1	3							
		ΜΟΒΡΗΣ	1	3							
	ΑΓ.ΒΑΡΒΑΡΑΣ	ΠΑΤΡΑ	2	2		2					3
		ΑΓ.ΒΑΡΒΑΡΑΣ	2	2		2					3
	ΑΡΟΛΙΑΣ			1							
	ΒΡΑΧΝΑΪΚΩΝ					1	1	2			2
	ΚΑΛΕΝΤΣΙΟΥ										
	ΤΡΙΤΑΙΑΣ		10	3	1	2	3	5			4
	ΣΤΑΥΡΟΔΕΒΡΟΚΟΥ										

Εικόνα 3.15 : Αριθμός μαθητών ανά μάθημα με βαθμό 19 ανά νομό, δήμο και περιοχή

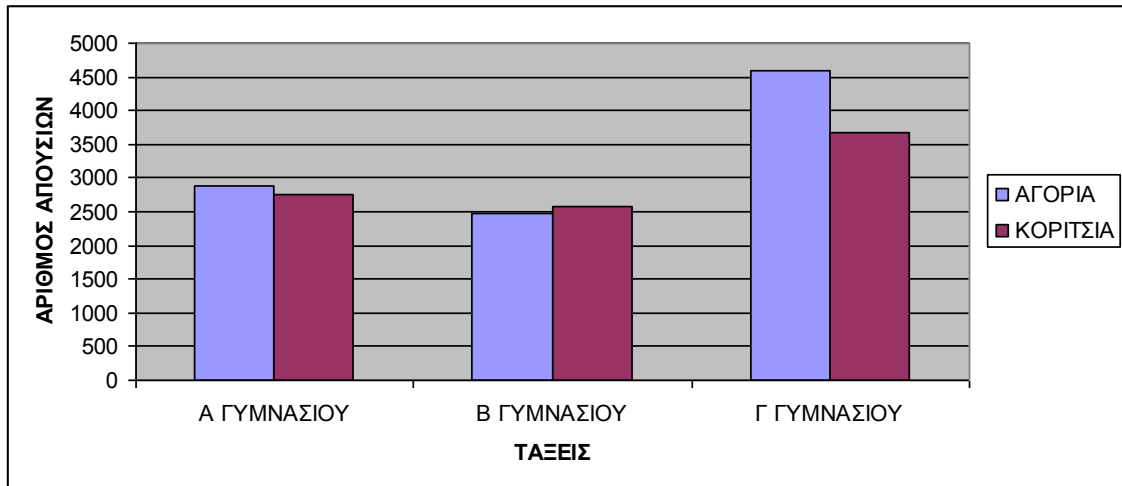
ΠΑΡΑΚΟΛΟΥΘΗΣΗ ΜΑΘΗΤΩΝ

5. Πόσες αδικαιολόγητες απουσίες έχουν οι μαθητές ανά τάξη και ανά φύλο

Σε αυτό το ερώτημα εξετάζεται η εξέλιξη του αριθμού των απουσιών ανά τάξη και φύλο (εικόνα 3.16), με σκοπό να διερευνηθεί το πλήθος τους καθώς οι μαθητές προάγονται σε μεγαλύτερες τάξεις. Επίσης εξετάζεται αν παρουσιάζονται αποκλίσεις στις απουσίες λόγω του φύλου τους. Μιας και στα δεδομένα υπάρχουν μαθητές Γυμνασίου είναι ιδιαίτερα ενδιαφέρουσα η εξέταση της αντιμετώπισης της εκπαιδευτικής διαδικασίας κατ' αρχήν σε όσο μεγαλύτερη τάξη φοιτούν. Συνηθίζεται σε μεγαλύτερες τάξεις οι μαθητές να κάνουν περισσότερες αδικαιολόγητες απουσίες διότι τα ενδιαφέροντα τους επεκτείνονται και ξεπερνούν τα όρια του σχολικού χώρου.

Εδώ διερευνείται τόσο η εξέλιξη του φαινομένου χρονικά όσο και η διαφοροποίηση του ανάμεσα σε αγόρια και κορίτσια. Θα μπορούσε κανείς να

υποθέσει ότι τα αγόρια μαθητές είναι περισσότερο επιρρεπείς στη πραγματοποίηση αδικαιολόγητων απουσιών καθώς παραδοσιακά τα αγόρια θεωρούνται πιο ατίθασα και ευάλωτα σε ομαδικές συμπεριφορές της "παρέας" και σε επιρροές από άλλους συνομηλίκους τους. Αντίθετα είναι κοινή πεποίθηση πως τα κορίτσια υιοθετούν το χαρακτήρα και τη συμπεριφορά της επιμελούς μαθήτριας από μικρότερη ηλικία και μέχρι την αποφοίτησή τους. Η επεξεργασία των δεδομένων της υπάρχουσας βάσης μπορεί να επιβεβαιώσει τους παραπάνω ισχυρισμούς.

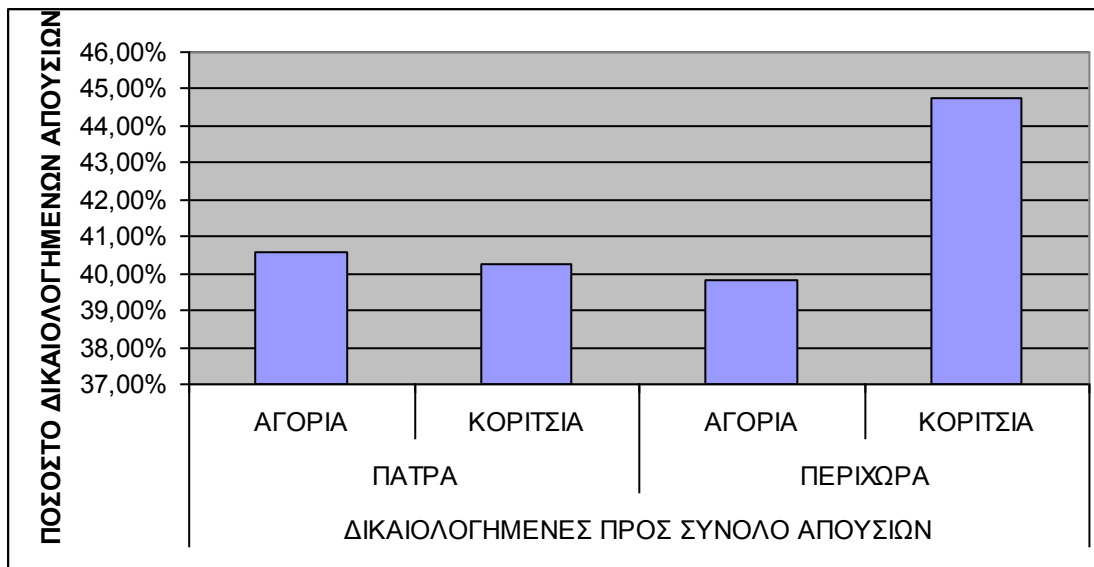


Εικόνα 3.16 : Αριθμός Απουσιών ανά τάξη και φύλο

Σύμφωνα με τα αποτελέσματα της ανάλυσης παρουσιάζεται σημαντική αύξηση του αριθμού των αδικαιολόγητων απουσιών στην τρίτη τάξη του Γυμνασίου κάτι που όπως αναφέρθηκε και στο πρόλογο του ερωτήματος είναι αναμενόμενο καθώς οι συμπεριφορές των μαθητών σε αυτή την ηλικία οδηγούν σε τέτοια αποτελέσματα.

6. Ποιο είναι το ποσοστό των δικαιολογημένων απουσιών ανά φύλο και περιοχή των μαθητών

Σκοπός του ερωτήματος αυτού, είναι να βοηθήσει με τα αποτελέσματα του την ανάλυση των στατιστικών μεγεθών που αφορούν το ποσοστό των δικαιολογημένων απουσιών. Αυτό γίνεται προκειμένου, να έχει η εκπαιδευτική αρχή στη κατοχή της, μια καλύτερη εικόνα για τα εκπαιδευτικά ιδρύματα που έχει υπό την εποπτεία της και να αξιολογεί τις διακυμάνσεις –αν αυτές υπάρχουν– των απουσιών των μαθητών που είναι όμως δικαιολογημένες με βάση μια συγκεκριμένη γεωγραφική κατανομή.



Εικόνα 3.17 : Ποσοστό Δικαιολογημένων Απουσιών ως προς το σύνολο τους ανά περιοχή και φύλο

Στο ερώτημα αυτό γίνεται διερεύνηση για το ποσοστό των απουσιών που δικαιολογούνται από μαθητές της περιοχής των Πατρών και των Περιχώρων. Γενικότερα παρατηρείται πως στο σύνολο των απουσιών δικαιολογείται ένα σταθερό περίπου ποσοστό της τάξης του 40% με εξαίρεση τα κορίτσια των περιφερειακών σχολείων τα οποία δικαιολογούν τις απουσίες τους κατά 45% (εικόνα 3.17).

7. Μαθητές(ονομαστικά) με αυξημένες δικαιολογημένες απουσίες ανά έτος

Στη συνέχεια γίνεται εμφανές ότι κυρίως τα αγόρια έχουν αυξημένο αριθμό δικαιολογημένων απουσιών, αν και γενικά το πλήθος των μαθητών με δικαιολογημένες απουσίες άνω των 65 δεν είναι πολύ μεγάλο. Παρόλα αυτά άξιο αναφοράς είναι ότι υπάρχουν μαθητές με πάνω από 90 δικαιολογημένες απουσίες, αριθμός πολύ υψηλός πάνω από τα επιτρεπόμενα όρια. Αποκορύφωμα σύμφωνα με τα δεδομένα μας, ήταν ο μαθητής ΓΙΑΝΝΑΚΟΠΟΥΛΟΣ ΝΙΚΟΛΑΟΣ ο οποίος τη σχολική χρονιά 2005-2006 ενώ φοιτούσε στην Α Γυμνασίου πραγματοποίησε 130 δικαιολογημένες απουσίες (εικόνα 3.18).

StatusYear	2002-2003	2003-2004	2004-2005	2005-2006	2006-2007	2007-2008	2008-2009	2009-2010
Αγόρι								
ΠΑΥΛΟΠΟΥΛΟΣ						93		
ΜΑΝΙΚΑΣ							71	
ΞΕΝΟΣ			71					
ΣΟΦΙΟΣ			72					
ΑΡΜΕΝΗΣ	80							
ΠΑΠΑΓΕΩΡΓΙΟΥ		72						
ΠΑΝΝΑΚΟΠΟΥΛΟΣ				130	130			
ΝΙΚΟΛΑΟΣ				130	130			
ΠΑΝΝΑΚΟΠΟΥΛΟΣ				130	130			
ΖΕΡΒΑΣ		66	66					
ΠΕΛΕΚΟΥΔΑΣ-ΟΙΚΟΝΟΜΟΥ							73	
ΠΑΡΚΙΑΣ						110		
ΚΑΡΑΠΑΝΑΠΩΤΗΣ							76	
ΡΟΔΗΣ							75	
ΠΑΠΑΔΟΠΟΥΛΟΣ								100
ΑΣΤΕΡΗΣ							79	
ΠΑΠΑΒΕΡΚΙΟΥ								67
ΒΕΡΓΟΣ								69
Κορίτσι								
Αγόρι	80	72	66	130	130	110	76	100
Κορίτσι		70			83	83		67
Total								

Εικόνα 3.18 : Μαθητές με αυξημένες δικαιολογημένες απουσίες ανά έτος

8. Μαθητές (ονομαστικά) που δεν προβιβάστηκαν λόγω απουσιών ανά σχολικό έτος και ανά τάξη

Εξετάσαμε τους μαθητές ως προς τις αδικαιολόγητες απουσίες τους και είδαμε ότι τα τελευταία έτη και συγκεκριμένα το σχολικό έτος 2009-2010 έχει αποτύχει να προαχθεί στην επόμενη τάξη από απουσίες η μαθήτρια ΟΡΤΕΝΤΣΙΟΥ ΕΛΙΣΑΒΕΤ, δηλαδή μόνο ένα άτομο, κάνοντας 66 αδικαιολόγητες απουσίες ενώ το ανώτατο επιτρεπτό όριο ήταν 64.

Εξαιτίας αυτού του γεγονότος παρατηρείται ότι τα παιδιά του Γυμνασίου δεν πραγματοποιούν άσκοπα πολλές αδικαιολόγητες απουσίες και η συντριπτική πλειοψηφία τους γνωρίζει ότι σε αντίθετη περίπτωση θα χάσουν το δικαίωμα Προαγωγής στην επόμενη τάξη (εικόνα 3.19).

Filter area	
Item	Filter description
StatusNotExcusedAbsents (Max)	Not less "64"

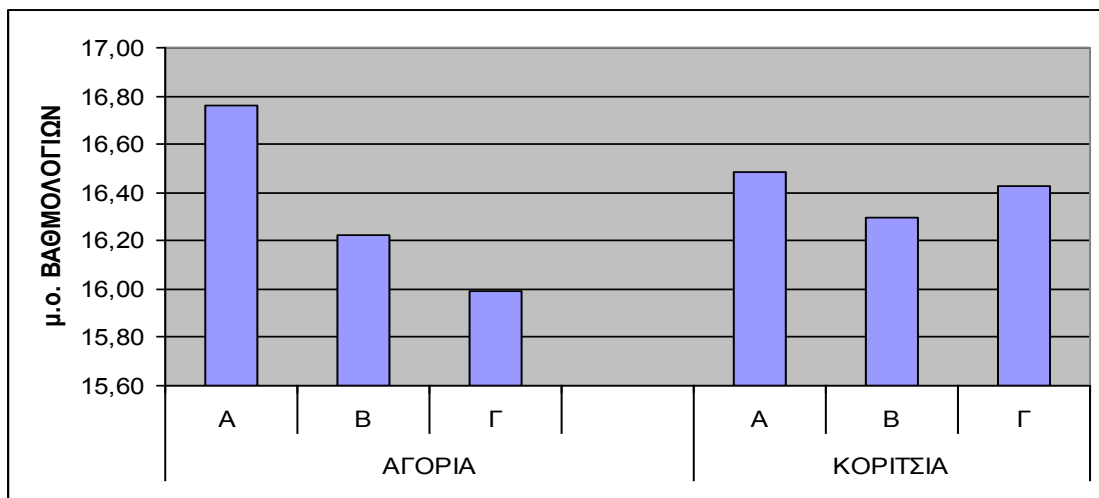
Columns area		
StatusYear	StatusClass	Measures
2002-2003	2002-2003	2008-2009
2008-2009	2009-2010	Total
Sex	SurName	Name
Αγόρι	ΓΡΑΒΑΝΗΣ	ΠΑΝΑΓΙΩΤΗΣ
		ΓΡΑΒΑΝΗΣ
	ΚΩΝΣΤΑΝΤΑΤΟΣ	ΝΙΚΟΛΑΟΣ
		ΚΩΝΣΤΑΝΤΑΤΟΣ
Κορίτσι	ΟΡΤΕΝΤΣΙΟΥ	ΕΛΙΣΑΒΕΤ
		ΟΡΤΕΝΤΣΙΟΥ
	ΜΠΟΥΡΤΖΙΛΑ	ΚΑΛΛΙΟΠΗ
		ΜΠΟΥΡΤΖΙΛΑ
		Κορίτσι
	Total	

Εικόνα 3.19 : Ονοματεπώνυμο μαθητών με αριθμό αδικαιολόγητων απουσιών από 64 και πάνω ανά φύλο, έτος και τάξη

ΕΠΙΔΟΣΗ ΜΑΘΗΤΩΝ

9. Ποιος είναι ο μέσος όρος των βαθμολογιών των μαθητών ανά τάξη

Αυτό το ερώτημα έχει σκοπό να διερευνήσει την εξέλιξη του μέσου όρου της βαθμολογίας των μαθητών ανά τάξη (εικόνα 3.20), προκειμένου να δοθεί μια στατιστική εικόνα ανάμεσα σε αγόρια και κορίτσια και να γίνει η απαραίτητη ανάλυση για το αν υπάρχει λόγος να γίνουν διορθωτικές κινήσεις ανάλογα με τα αποτελέσματα της ανάλυσης αυτής. Ο μέσος όρος της απόδοσης των μαθητών, αποτελεί χαρακτηριστικό δείγμα της γενικότερης απόδοσης των σχολικών συγκροτημάτων, με αποτέλεσμα να είναι σημαντικό στοιχείο ανάλυσης και εξαγωγής συμπερασμάτων.



Εικόνα 3.20 : Μέσος όρος Βαθμολογιών ανά τάξη και φύλο

Σε αυτό το ερώτημα δίνεται μια ξεκάθαρη εικόνα υπεροχής των μαθητριών έναντι των αγοριών μαθητών. Εκτός από αυτό τα αγόρια μαθητές έχουν πτωτική πορεία στην απόδοση τους ενώ τα κορίτσια είναι σταθερά και εν μέρει με αυξητική τάση.

10.Αριστούχοι μαθητές ανά φύλο και σχολική χρονιά, ονομαστικά

Αναζητήσαμε τους μαθητές που πήραν βαθμό «20» όλες τις χρονιές, έτσι ώστε να μπορεί εύκολα το σχολείο να βρει ποίοι είχαν οριστεί σηματοφόροι, απουσιολόγοι και είχαν αποκτήσει αριστεία. Μέσα από αυτό το πίνακα (εικόνα 3.21) επίσης είναι φανερό ότι οι αριστούχοι μαθητές είναι στη πλειοψηφία τους Αγόρια ενώ ανάμεσα τους υπάρχει μόλις ένα Κορίτσι.

Filter area			Columns area									
Item	Filter description											
Status/Result/Average/Num. (Max)	Equal "20"											
Column area												
Status/Year	Status/Class	Nationality	Measures								Total	
Sex	SurName	Name	2001-2002	2002-2003	2004-2005	2005-2006	2006-2007	2008-2009	2009-2010	2010-2011	Total	
Κορίτσια	ΛΙΑΠΗ	ΜΑΡΙΑ						20,00	20,00		20,00	
		Κορίτσια						20,00	20,00		20,00	
Αγόρια	ΚΟΡΕΝΤΖΕΛΟΣ	ΔΗΜΗΤΡΙΟΣ	20,00								20,00	
		ΚΟΡΕΝΤΖΕΛΟΣ	20,00								20,00	
	ΑΡΜΕΝΗΣ	ΓΕΩΡΓΙΟΣ		20,00							20,00	
		ΑΡΜΕΝΗΣ		20,00							20,00	
	ΔΕΛΗΓΙΑΝΝΗΣ	ΠΑΡΑΣΚΕΥΑΣ			20,00		20,00				20,00	
		ΔΕΛΗΓΙΑΝΝΗΣ			20,00		20,00				20,00	
	ΜΠΑΤΣΙΚΑΣ	ΣΠΥΡΙΔΩΝ				20,00	20,00				20,00	
		ΣΠΥΡΟΣ						20,00	20,00	20,00	20,00	20,00
	ΡΟΔΗΣ	ΑΝΔΡΕΑΣ								20,00		20,00
		ΡΟΔΗΣ								20,00		20,00
		Αγόρια	20,00	20,00	20,00	20,00	20,00	20,00	20,00	20,00	20,00	
		Total	20,00	20,00	20,00	20,00	20,00	20,00	20,00	20,00	20,00	

Εικόνα 3.21 : Ονοματεπώνυμο μαθητών με γενικό μέσο όρο «20» ανά έτος και ανά φύλο

Ποιο αναλυτικά μπορούμε να δούμε ποιοι ήταν οι αριστούχοι μαθητές (ονοματεπώνυμο) και σε ποιά τάξη του γυμνασίου ήταν (εικόνα 3.22). Ακόμα μπορεί να προσδιορισθεί και η εθνικότητα του καθενός από τους αριστούχους μαθητές και με μια πρώτη ματιά γίνεται εμφανές ότι δεν υπάρχουν αριστούχοι μαθητές ξένης εθνικότητας. Αυτό εν μέρει είναι φυσιολογικό μιας και η γλώσσα παίζει μεγάλο ρόλο καθώς η Ελληνική γλώσσα θεωρείται από τις δυσκολότερες παγκοσμίως.

StatusYear	2001-2002		2001-2002		2002-2003		2002-2003		2004-2005		2005-2006	
StatusClass	A		A		Γ		Γ		A		B	
Nationality	ΕΛΛΗΝΙΚΗ		ΕΛΛΗΝΙΚΗ		ΕΛΛΗΝΙΚΗ		ΕΛΛΗΝΙΚΗ		ΕΛΛΗΝΙΚΗ		ΕΛΛΗΝΙΚΗ	
Καρίτα	ΚΑΡΙΤΑ	ΚΑΡΙΤΑ										
Ανδρ	ΚΟΡΕΝΤΖΕΛΟΣ	ΔΗΜΗΤΡΙΟΣ	20,00	20,00	20,00							
	ΑΡΜΕΝΗΣ	ΓΕΩΡΓΙΟΣ	20,00	20,00	20,00							
	ΔΕΛΗΓΙΑΝΝΗΣ	ΑΡΜΕΝΗΣ			20,00	20,00	20,00					
	ΜΠΑΤΣΙΚΑΣ	ΔΕΛΗΓΙΑΝΝΗΣ						20,00	20,00	20,00		
	ΡΟΔΗΣ	ΣΠΥΡΙΔΩΝ									20,00	20,00
		ΜΠΑΤΣΙΚΑΣ										20,00
		ΑΝΔΡΕΑΣ										20,00
		ΡΟΔΗΣ	20,00	20,00	20,00	20,00	20,00	20,00	20,00	20,00	20,00	20,00
		Ανδρ	20,00	20,00	20,00	20,00	20,00	20,00	20,00	20,00	20,00	20,00
		Total	20,00	20,00	20,00	20,00	20,00	20,00	20,00	20,00	20,00	20,00

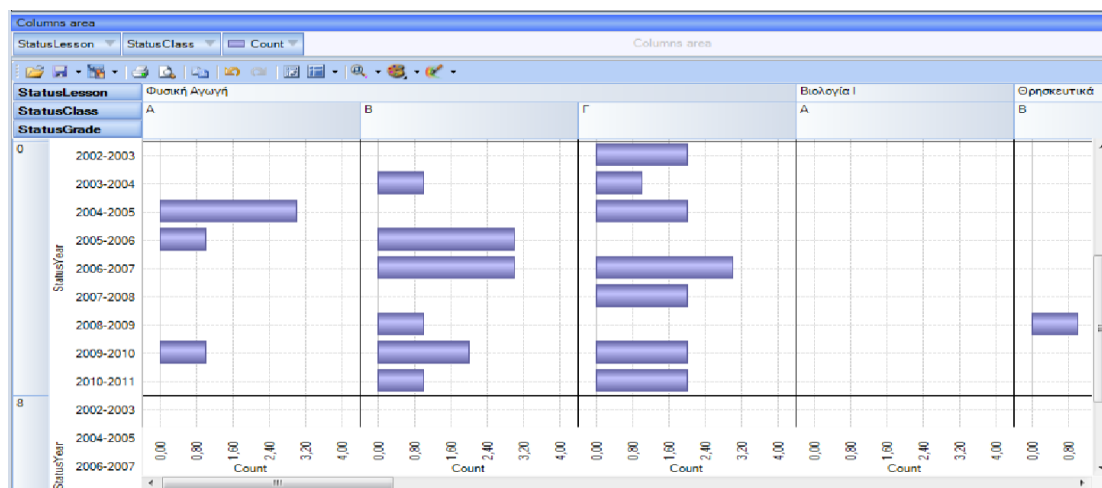
Εικόνα 3.22 : Ονοματεπώνυμο μαθητών με γενικό μέσο όρο «20» ανά έτος, τάξη, και εθνικότητα

11. Ποιο είναι το πλήθος των μαθητών που πήραν βαθμολογίες του 10 ή κάτω από την βάση, ανά τάξη και σχολική χρονιά

Υπολογίστηκαν οι βαθμολογίες που πήραν οι μαθητές ανά σχολικό έτος σε κάθε τάξη, όπου είναι ίσες ή μικρότερες του 10 (εικόνα 3.23). Οι μαθητές που έβγαλαν μέσο όρο 10 σε κάποιο μάθημα ήταν συνολικά 326 μέσα σε 10 χρόνια. Το 2002-2003 είχαμε τους περισσότερους μαθητές με αυτή την βαθμολογία, που ήταν στο σύνολο τους 96. Οι μαθητές που δεν πέρασαν τα μαθήματα τους, δηλαδή είχαν βαθμούς 9, 8 και 0 είναι συνολικά 55 σε αυτά τα δέκα χρόνια. Στο παρακάτω ραβδόγραμμα (εικόνα 3.24), απεικονίζεται το πλήθος των μαθητών στο μάθημα της Φυσικής Αγωγής, όπου υπάρχει μηδενική βαθμολογία κάποιων μαθητών. Τα μεγαλύτερα ποσοστά μαθητών που πήραν βαθμό 0 στο μάθημα αυτό είναι στα έτη 2004-2005 στην Α Γυμνασίου, 2005-2006 και 2006-2007 στην Β Γυμνασίου, και 2006-2007 στην Γ Γυμνασίου.

StatusGrade	10				9				8				0				Total							
StatusClass	A		B		Γ		Γ		A		B		Γ		A		B		Γ		Γ			
StatusYear	=		=		=		=		=		=		=		=		=		=		=			
2001-2002																							1	
2007-2008			4		5																		1	13
2010-2011		3			5																		3	19
2008-2009		2			12																		2	25
2003-2004		4			11																		2	32
2009-2010		12			5																		1	7
2004-2006		3			15																		1	5
2006-2007		8			15																		2	6
2005-2006		15			3																		4	56
2002-2003		14			19																		2	101
Total		62			89																		6	361

Εικόνα 3.23 : Αριθμός μαθητών ανά χρονολογία και τάξη με μέσο όρο κάποιου μαθήματος 10,9,8 ή 0



Εικόνα 3.24 : Γραφική αναπαράσταση αριθμού μαθητών ανά χρονολογία και τάξη με μέσο όρο κάποιου μαθήματος 0

12. Ποιος είναι ο γενικός μέσος όρος γενικών μαθημάτων ανά τάξη ανά σχολικό έτος

Απεικονίζονται οι μέσοι όροι των βαθμολογιών των μαθητών ανά μάθημα και ανά σχολικό έτος στην Α Γυμνασίου (εικόνα 3.25). Παρατηρείται ότι το μάθημα με το μεγαλύτερο μέσο όρο είναι το μάθημα της Αισθητικής Αγωγής, ενώ οι μαθητές δυσκολεύονται περισσότερο στο μάθημα των Μαθηματικών. Οι καλύτερες επιδόσεις των μαθητών εμφανίζονται τη χρονιά 2008-2009 και οι χειρότερες τις χρονιές 2002-2003 και 2003-2004.

StatusYear	Φυσική Αγωγή	Αισθητική Αγωγή	Βιολογία Ι	Γεωγραφία	Πληροφορική - Τεχνολογία	Οικιακή Οικονομία	Θρησκευτικά	Αρχαία Ελληνική Γλώσσα και Γραμματεία	Νεοελληνική Γλώσσα και Γραμματεία	Ιστορία	Μαθηματικά	A
2001-2002	19,63	19,28	16,24	16,59	17,34	17,66	17,69	15,83	16,28	16,17	15,00	16,98
2002-2003	18,42	18,44	15,56	16,58	16,85	17,24	17,31	16,45	16,19	15,31	15,27	16,69
2003-2004	18,42	18,68	17,62	17,45	16,98	16,30	17,38	14,85	14,98	15,78	15,17	16,69
2004-2005	17,53	19,03	16,42	16,03	16,70	17,77	17,93	16,72	16,58	16,03	16,32	17,01
2005-2006	18,80	19,02	15,35	15,27	17,92	18,75	17,85	16,75	16,47	17,02	15,20	17,13
2006-2007	18,63	19,33	16,10	16,70	17,73	18,60	17,75	17,58	17,00	16,80	15,15	17,40
2007-2008	18,40	19,25	17,12	17,63	18,87	18,07	18,07	17,10	17,15	17,12	16,78	17,78
2008-2009	19,07	19,29	18,92	17,90	18,92	18,93	18,29	17,49	17,58	17,61	16,71	18,24
2009-2010	18,59	19,12	16,59	16,95	17,97	18,64	18,47	17,40	17,67	16,45	17,53	17,76
2010-2011	18,92	19,08	16,45	17,20	18,02	18,68	17,82	17,75	17,87	17,47	16,05	17,75
Total	18,59	18,99	16,65	16,84	17,74	18,08	17,86	16,84	16,80	16,59	15,96	17,36

Εικόνα 3.25 : Μέσος όρος των βαθμολογιών των μαθητών της Α Γυμνασίου ανά μάθημα και ανά σχολική χρονιά

StatusClass	Φυσική Αγωγή	Αισθητική Αγωγή	Γεωγραφία	Πληροφορική - Τεχνολογία	Οικιακή Οικονομία	Θρησκευτικά	Αρχαία Ελληνική Γλώσσα και Γραμματεία	Νεοελληνική Γλώσσα και Γραμματεία	Ιστορία	Μαθηματικά	Φυσική - Χημεία	B
2002-2003	18,33	18,50	15,58	18,03	16,48	17,42	15,47	15,98	16,58	15,38	15,72	16,70
2003-2004	18,28	18,64	16,67	16,74	16,52	17,66	16,51	16,79	16,34	14,84	15,87	16,80
2004-2005	18,47	18,53	16,57	16,43	16,60	19,13	14,72	15,42	15,88	16,37	15,98	16,74
2005-2006	17,70	19,10	16,87	18,22	18,75	18,78	17,12	17,17	16,63	15,72	16,27	17,48
2006-2007	17,72	18,95	16,35	17,60	18,05	17,35	16,25	16,47	16,97	15,33	16,70	17,07
2007-2008	18,49	19,29	17,22	18,42	18,59	17,88	16,31	16,96	16,34	15,93	16,61	17,42
2008-2009	18,48	18,90	16,25	18,08	17,77	17,85	16,58	16,88	16,32	15,97	16,12	17,20
2009-2010	17,90	19,32	16,80	18,03	19,17	18,63	17,55	18,03	17,42	16,30	16,58	17,79
2010-2011	18,53	19,03	17,93	18,22	18,52	18,38	17,43	17,63	18,97	16,52	17,23	17,85
Total	18,24	18,92	16,88	17,75	17,82	18,12	16,44	16,77	16,61	15,81	16,34	17,23

Εικόνα 3.26 : Μέσος όρος των βαθμολογιών των μαθητών της Β Γυμνασίου ανά μάθημα και ανά σχολική χρονιά

Ομοίως βλέπουμε τα στοιχεία για τη Β Γυμνασίου(εικόνα 3.26). Παρατηρείται ότι το μάθημα με το μεγαλύτερο μέσο όρο είναι ξανά το μάθημα της Αισθητικής Αγωγής όπως επίσης το μάθημα που τους δυσκόλεψε περισσότερο είναι ξανά τα Μαθηματικά. Οι καλύτερες επιδόσεις των μαθητών παρατηρούνται τη σχολική χρονιά 2010-2011 όπου ο γενικός μέσος όρος των μαθητών έφτασε το 17,85 και οι χειρότερες επιδόσεις τους ήταν τη σχολική χρονιά 2002-2003 όπου ο γενικός μέσος όρος ήταν 16,70.

StatusClass	Φυσική Αγωγή	Αισθητική Αγωγή	Θρησκευτικά	Αρχαία Ελληνική Γλώσσα και Γραμματεία	Νεοελληνική Γλώσσα και Γραμματεία	Ιστορία	Μαθηματικά	Φυσική - Χημεία	Βιολογία II	Πληροφορική	Κοινωνική Αγωγή	Γ	Total
2002-2003	18,49	18,73	17,16	15,85	15,77	15,74	15,44	15,51	15,11	14,75	16,90	16,31	16,31
2003-2004	18,37	18,97	16,77	15,87	16,57	16,15	15,35	16,23	16,88	15,30	17,22	16,71	16,71
2004-2005	18,16	18,95	16,95	16,64	16,84	16,44	15,39	15,85	15,44	14,87	16,97	16,59	16,59
2005-2006	18,63	18,97	17,33	15,75	15,73	14,80	15,82	15,63	14,88	15,98	16,97	16,41	16,41
2006-2007	17,82	19,23	17,48	16,60	16,92	15,43	15,57	16,72	16,32	15,37	16,88	16,76	16,76
2007-2008	17,80	19,32	17,49	16,64	16,81	16,00	16,96	16,49	16,69	18,00	18,31	17,27	17,27
2008-2009	18,75	19,36	17,76	16,71	16,75	16,47	16,69	16,02	17,90	18,61	18,71	17,60	17,60
2009-2010	18,10	19,00	17,28	16,98	17,42	17,12	15,92	16,00	16,28	17,52	18,48	17,28	17,28
2010-2011	18,07	19,52	17,88	17,33	17,95	17,28	16,45	17,32	16,65	18,15	19,00	17,78	17,78
Total	18,26	19,09	17,34	16,45	16,70	16,14	15,87	16,16	16,17	16,40	17,67	16,93	16,93

Εικόνα 3.27 : Μέσος όρος των βαθμολογιών των μαθητών της Γ Γυμνασίου ανά μάθημα και ανά σχολική χρονιά

Τέλος βλέπουμε τους μέσους όρους των βαθμολογιών των μαθητών ανά μάθημα και ανά σχολικό έτος στην Γ Γυμνασίου (εικόνα 3.27). Όπως και στις άλλες δύο τάξεις του Γυμνασίου παρατηρούμε ότι η καλύτερη βαθμολογία των μαθητών ήταν στην Αισθητική Αγωγή και η χειρότερη στα Μαθηματικά. Η αποδοτικότερη σχολική χρονιά για τους μαθητές της Γ Γυμνασίου ήταν το 2010-2011 και η λιγότερο αποδοτική ήταν το 2002-2003.

StatusClass	A	B	Γ	Total
StatusYear				
2001-2002	16,98			16,98
2002-2003	16,69	16,70	16,31	16,53
2003-2004	16,69	16,80	16,71	16,73
2004-2005	17,01	16,74	16,59	16,78
2005-2006	17,13	17,48	16,41	17,01
2006-2007	17,40	17,07	16,76	17,07
2007-2008	17,78	17,42	17,27	17,49
2008-2009	18,24	17,20	17,60	17,68
2009-2010	17,76	17,79	17,28	17,61
2010-2011	17,75	17,85	17,78	17,79
Total	17,36	17,23	16,93	17,17

Εικόνα 3.28 : Μεικτός μέσος όρος των τελικών βαθμολογιών των μαθητών των τριών τάξεων του Γυμνασίου ανά τάξη και ανά σχολική χρονιά

Όπως παρουσιάζεται από τον συνοπτικότερο πίνακα (εικόνα 3.28), οι καλύτερες βαθμολογίες μέσων όρων ήταν στην τάξη της Α Γυμνασίου. Η δυσκολότερη χρονιά φαίνεται να είναι και η τελευταία τάξη του Γυμνασίου αφού ο μέσος όρος έχει πέσει κατά μισό βαθμό περίπου από την καλύτερη επίδοση (17,36). Ακόμα μια σημαντική παρατήρηση που μπορεί να σημειωθεί, είναι ότι όσο περνάνε τα σχολικά έτη, ο γενικός μέσος όρος έχει αυξητική πορεία και υπάρχει μια αύξηση σχεδόν κατά ένα βαθμό μετά το πέρας των δέκα χρόνων.

13. Ποιος είναι ο γενικός μέσος όρος των μαθημάτων ξένων γλωσσών, ανά τάξη και ανά σχολικό έτος

Τις σχολικές χρονιές από το 2001-2002 μέχρι το 2004-2005, παρατηρούμε ότι στο μάθημα των αγγλικών υπήρχε διαχωρισμός επιπέδου σε αρχαρίους και προχωρημένους, ενώ μετά, το μάθημα έγινε ενιαίο για όλους τους μαθητές. Σχετικά με τις ξένες γλώσσες που παρακολουθούσαν οι μαθητές, οι βαθμολογίες τους (εικόνες 3.29 και 3.30) είναι παρόμοιες στα Αγγλικά και τα Γαλλικά, ενώ είναι καλύτερες σε αυτούς που παρακολουθούσαν την Γερμανική γλώσσα στην Α Γυμνασίου. Αυτό όμως αλλάζει στην Β και Γ Γυμνασίου με μια άνοδο στην Αγγλική γλώσσα και πτώση στις άλλες δυο.

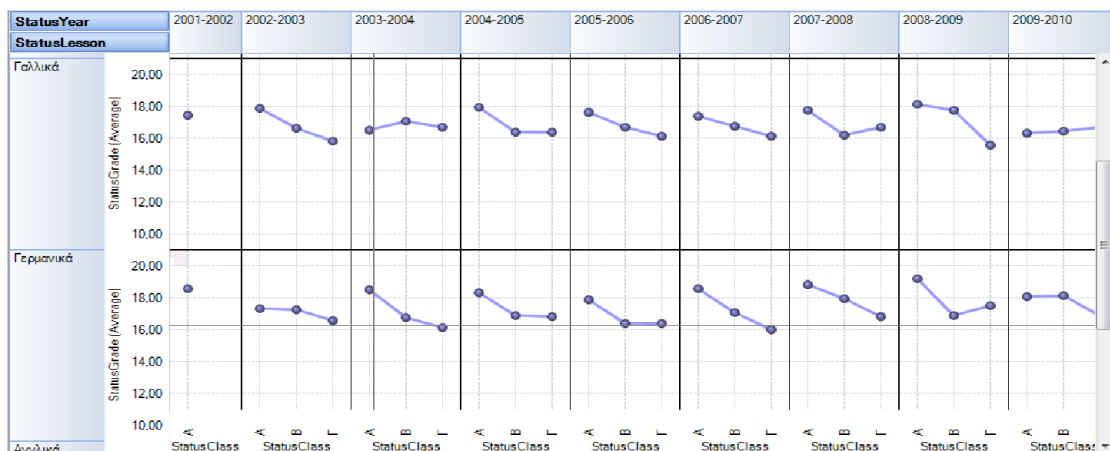
StatusClass	A				B				Γ				Total	
StatusLesson	Αγγλικά Αρχάριοι	Γαλλικά	Γερμανικά	Αγγλικά Προχωρημένοι	Αγγλικά	Γαλλικά	Γερμανικά	Αγγλικά Προχωρημένοι	Αγγλικά	Αγγλικά	Αγγλικά	Αγγλικά	Αγγλικά	
StatusYear														
2001-2002	18,07	17,41	18,58		17,98									17,98
2002-2003		17,85	17,31	16,68		17,14	16,59	17,26	15,96		16,40		16,56	16,68
2003-2004		16,49	18,52	17,50		17,38	17,09	16,75	17,57		17,25		17,28	17,31
2004-2005		17,94	18,32	17,47		17,78	16,35	16,87	15,97		16,26		16,67	16,97
2005-2006		17,62	17,85		16,87	17,29	16,69	16,40		17,57	17,07		16,81	17,05
2006-2007		17,38	18,54		17,08	17,48	16,76	17,08		17,30	17,10		16,85	17,14
2007-2008		17,73	18,82		17,03	17,69	16,21	17,96		16,02	16,48		17,28	17,15
2008-2009		18,10	19,18		17,49	18,05	17,77	16,88		18,07	17,67		17,38	17,70
2009-2010		16,30	18,08		17,57	17,32	16,47	18,13		17,03	17,17		17,75	17,41
2010-2011		17,65	18,72		18,28	18,23	16,00	17,19		17,20	16,87		17,73	17,61
Total	18,07	17,43	18,39	17,21	17,39	17,61	16,63	17,17	16,50	17,20	16,92	17,13	17,23	

Εικόνα 3.29 : Συνολικός μέσος όρος των βαθμολογιών των μαθητών του Γυμνασίου ανά τάξη, ανά διδασκόμενη ξένη γλώσσα και ανά σχολική χρονιά

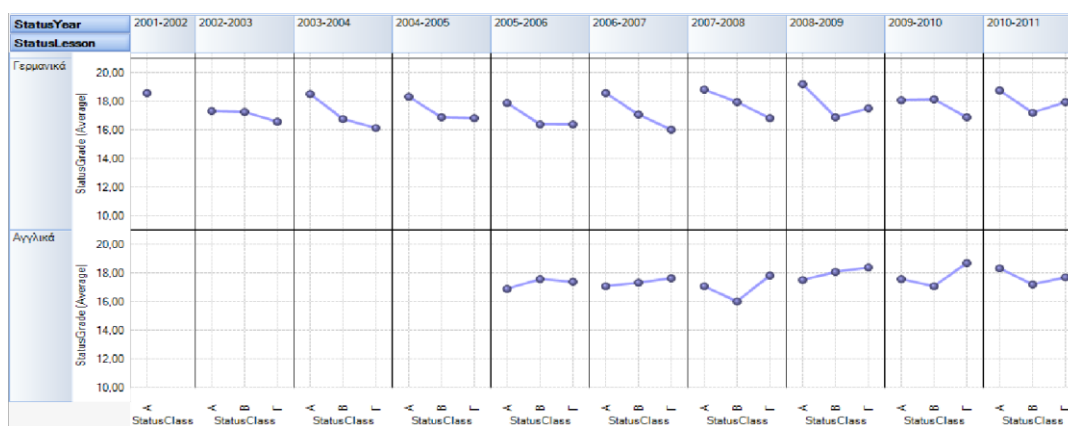
StatusLesson	Αγγλικά Αρχάριοι	Γαλλικά	Γερμανικά	Αγγλικά Προχωρημένοι	Αγγλικά	Total
StatusYear						
2001-2002	18,07	17,41	18,58			17,98
2002-2003		16,50	17,07	16,66		16,68
2003-2004		16,74	17,12	17,72		17,31
2004-2005		16,90	17,32	16,87		16,97
2005-2006		16,79	16,89		17,28	17,06
2006-2007		16,76	17,22		17,33	17,14
2007-2008		16,79	17,99		16,96	17,15
2008-2009		17,07	17,80		17,97	17,70
2009-2010		16,47	17,64		17,77	17,41
2010-2011		17,02	17,97		17,73	17,61
Total	18,07	16,78	17,49	17,06	17,51	17,23

Εικόνα 3.30 : Συνολικός μέσος όρος των βαθμολογιών των μαθητών του Γυμνασίου ανά διδασκόμενη ξένη γλώσσα και ανά σχολική χρονιά

Συνοπτικά υπάρχει μια σταθερή εικόνα στις ξένες γλώσσες σε όλα τα χρόνια με ένα γενικό μέσο όρο της τάξης του 17,23 τον οποίο βγάζουν οι μαθητές στα Αγγλικά, Γερμανικά και Γαλλικά (εικόνες 3.31 και 3.32).



Εικόνα 3.31 : Γραφική απεικόνιση μέσου όρου βαθμολογιών των μαθητών του Γυμνασίου στα Γαλλικά και στα Γερμανικά ανά τάξη και σχολική χρονιά



Εικόνα 3.32 : Γραφική απεικόνιση μέσου όρου βαθμολογιών των μαθητών του Γυμνασίου στα Αγγλικά και στα Γερμανικά ανά τάξη και σχολική χρονιά

StatusLesson	Αγγλικά Αρχάριοι	Αγγλικά Προχωρημένοι	Γαλλικά	Γερμανικά	Αγγλικά	Total	
Αγόρια	2001-2002	12		2	10	24	
	2002-2003		107	60	47	214	
	2003-2004		90	51	39	180	
	2004-2005		92	51	41	184	
	2005-2006			43	49	92	184
	2006-2007			40	53	93	186
	2007-2008			34	55	89	178
	2008-2009			36	52	88	176
	2009-2010			32	55	87	174
	2010-2011			38	51	89	178
Αγόρια	12	289	387	452	538	1.678	
Κορίτσια	2001-2002	17		15	2	34	
	2002-2003		106	77	29	212	
	2003-2004		91	57	34	182	
	2004-2005		89	54	35	178	
	2005-2006			63	25	88	176
	2006-2007			63	24	87	174
	2007-2008			61	28	89	178
	2008-2009			55	35	90	180
	2009-2010			57	34	91	182
	2010-2011			55	36	91	182
Κορίτσια	17	286	557	282	536	1.678	
Total	29	575	944	734	1.074	3.356	

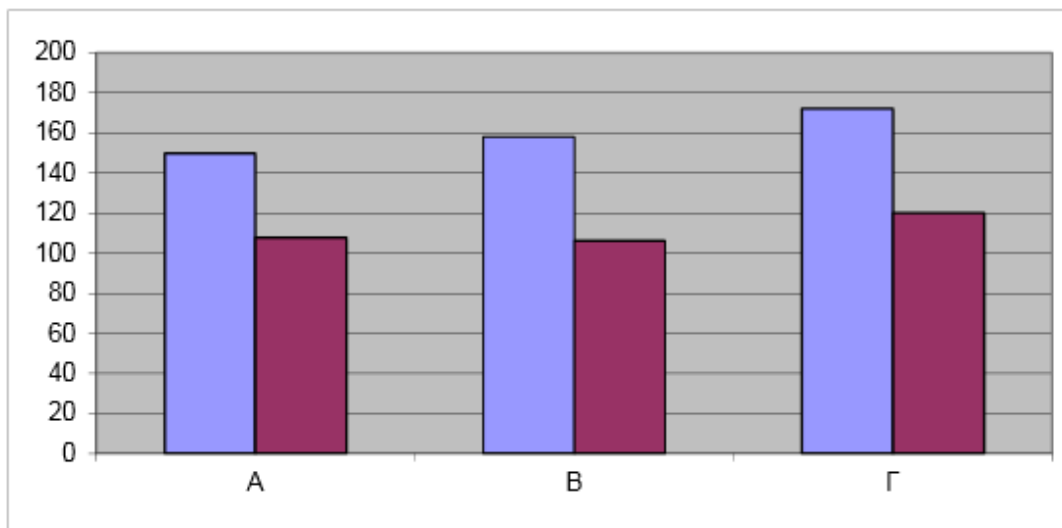
Εικόνα 3.33 : Αριθμός μαθητών ανά διδασκόμενη ξένη γλώσσα και σχολική χρονιά με βάση το φύλο

Βλέπουμε ότι τα Αγγλικά ως μάθημα είναι υποχρεωτικό για όλους τους μαθητές και στις τρεις τάξεις του Γυμνασίου ενώ τα Γαλλικά και τα Γερμανικά είναι μαθήματα επιλογής και οι μαθητές σε κάθε τάξη καλούνται να επιλέξουν ένα από τα δύο. Συνολικά παρατηρούμε ότι οι μαθητές προτιμούν τα Γαλλικά ως δεύτερη ξένη γλώσσα. Ενώ αν εξετάσουμε την επιλογή της δεύτερης ξένης γλώσσας ανά φύλο βλέπουμε ότι τα κορίτσια προτιμούν τα Γαλλικά και τα αγόρια τα Γερμανικά (εικόνα 3.33).

ΔΗΜΟΓΡΑΦΙΚΑ ΜΑΘΗΤΩΝ

14. Διαχωρισμός μαθητών ανά περιοχή και τάξη

Αυτό το ερώτημα στην ουσία αποτελεί ένα από τα πρωταρχικά στοιχεία μιας εκπαιδευτικής αρχής, καθώς αποτελεί το χάρτη της κατανομής των μαθητών. Αυτού του είδους τα αποτελέσματα βοηθούν το σχεδιασμό σε κάθε επίπεδο της εκπαιδευτικής δομής, αλλά αποτελούν και κύριο στοιχείο για την καλύτερη κατανομή των εκπαιδευτικών σε κάθε νέα σχολική χρονιά.



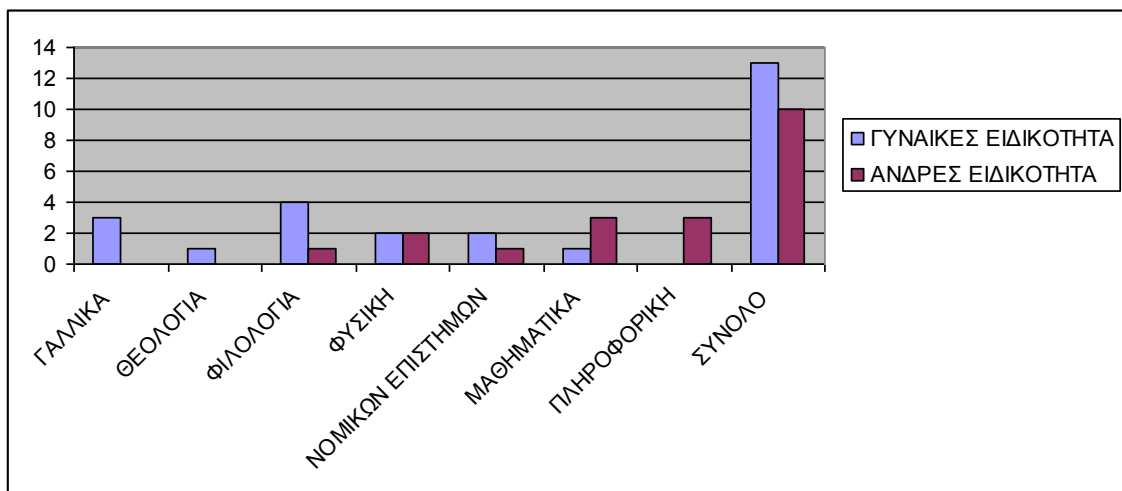
Εικόνα 3.34 : Πλήθος μαθητών ανά περιοχή και τάξη

Σε αυτό το ερώτημα το οποίο αφορά την κατανομή των μαθητών παρατηρούμε ότι όπως είναι αναμενόμενο οι μαθητές των κεντρικών σχολείων της Πάτρας είναι περισσότεροι (γαλάζιες στήλες) ενώ οι μαθητές των σχολείων των περιχώρων (κόκκινες στήλες είναι φυσιολογικά λιγότεροι (εικόνα 3.34). Σε αυτού του είδους τα ερωτήματα όσο μεγαλύτερο είναι το πλήθος των δεδομένων που χρησιμοποιούνται τόσο καλύτερα και πιο διαφωτιστικά είναι τα αποτελέσματα που λαμβάνονται.

ΚΑΘΗΓΗΤΕΣ:**ΕΠΙΠΕΔΟ ΣΠΟΥΔΩΝ ΚΑΘΗΓΗΤΩΝ****15. Ποιος είναι ο αριθμός των καθηγητών με μεταπτυχιακές σπουδές ανά φύλο και ειδικότητα**

Απαντώντας αυτό το ερώτημα γίνεται σαφές ότι είναι δυνατή η ποσοτική ανάλυση των δεδομένων προκειμένου να εξαχθούν και τα αντίστοιχα συμπεράσματα. Κάτι τέτοιο θα βοηθούσε την αξιολόγηση όχι μόνο των εκπαιδευτικών αλλά και την ανάλυση των προσόντων τους με βάση την ειδικότητα αλλά και το φύλο και οποιαδήποτε άλλη παράμετρο έχει ο αναλυτής διαθέσιμη.

Η ποσοτική ανάλυση πέρα από τα προφανή συμπεράσματα για τις δεξιότητες του εκπαιδευτικού προσωπικού θα βοηθούσε τον εκπαιδευτικό οργανισμό γενικότερα προκειμένου να αναδείξει τα πλεονεκτήματα του ή και τις αδυναμίες του ώστε με κάποιον ανάλογο σχεδιασμό να καλυφθούν αυτές ή οι αντίστοιχοι εκπαιδευτικοί να αξιοποιηθούν ακόμα καλύτερα.

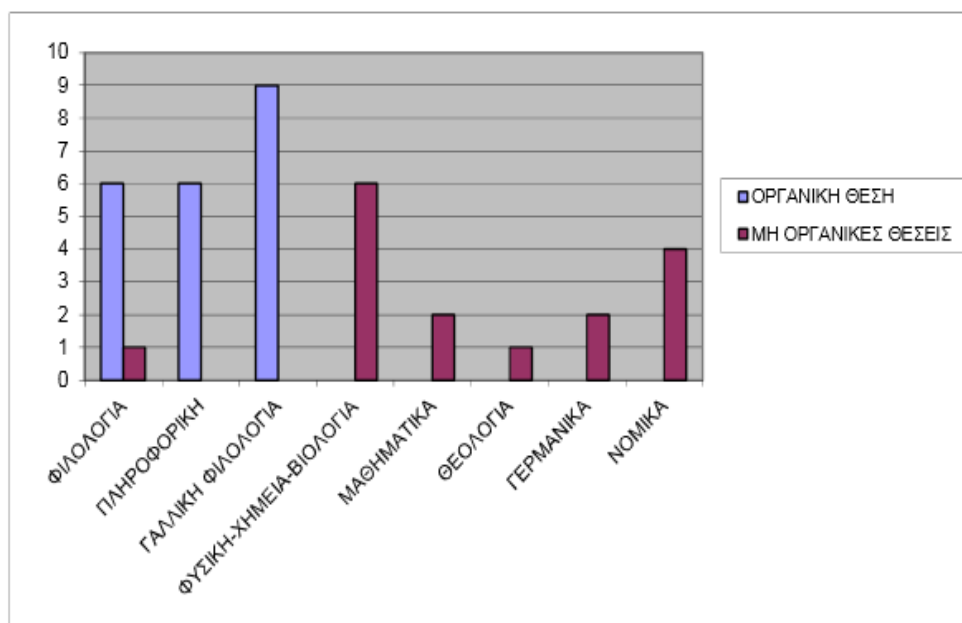


Εικόνα 3.35 : αριθμός καθηγητών με μεταπτυχιακές σπουδές ανά φύλο και ειδικότητα

Σε αυτό το γράφημα (εικόνα 3.35) παρουσιάζεται μια στατιστική ανάλυση των καθηγητών με μεταπτυχιακές σπουδές ανά φύλο και ειδικότητα. Τέτοιους είδους αναλύσεις βοηθούν στην αξιολόγηση του υπηρετούντος προσωπικού έτσι ώστε η εκπαιδευτική αρχή να έχει μια καλύτερη εικόνα για το προσωπικό που απασχολεί.

Ακόμα ένα στοιχείο ποσοτικής ανάλυσης εξάγεται από το παραπάνω ερώτημα και σκοπό έχει να δείξει τον αριθμό των μεταπτυχιακών τίτλων σπουδών σε κάθε εκπαιδευτική ειδικότητα. Εδώ λαμβάνεται υπ' όψιν και το είδος της θέσης για μια πληρέστερη εικόνα της κατάρτισης των καθηγητών.

Άλλο ένα ερώτημα που θέτει ποιοτική διάσταση στη μελέτη του προσωπικού του εκπαιδευτικού οργανισμού είναι το παραπάνω. Όπως φαίνεται από το διάγραμμα (εικόνα 3.36) οι καθηγητές με μεταπτυχιακούς τίτλους είναι στην πλειοψηφία τους καθηγητές με μη οργανικές θέσεις. Αυτό θα μπορούσε να ερμηνευθεί ως ότι αυτές οι θέσεις καλύπτονται από καθηγητές νεαρούς σε ηλικία οι οποίοι μόλις ολοκλήρωσαν τις σπουδές τους, από τους οποίους λόγω του υψηλού ανταγωνισμού απαιτούνται μεταπτυχιακοί τίτλοι σπουδών προκειμένου να προσληφθούν.



Εικόνα 3.36 : αριθμός καθηγητών με μεταπτυχιακές σπουδές ανά ειδικότητα και είδος θέσης

ΕΙΔΟΣ ΘΕΣΗΣ ΚΑΘΗΓΗΤΩΝ

16. Πόσοι ήταν οι καθηγητές ανά σχολική χρονιά με βάση το είδος της θέσης τους

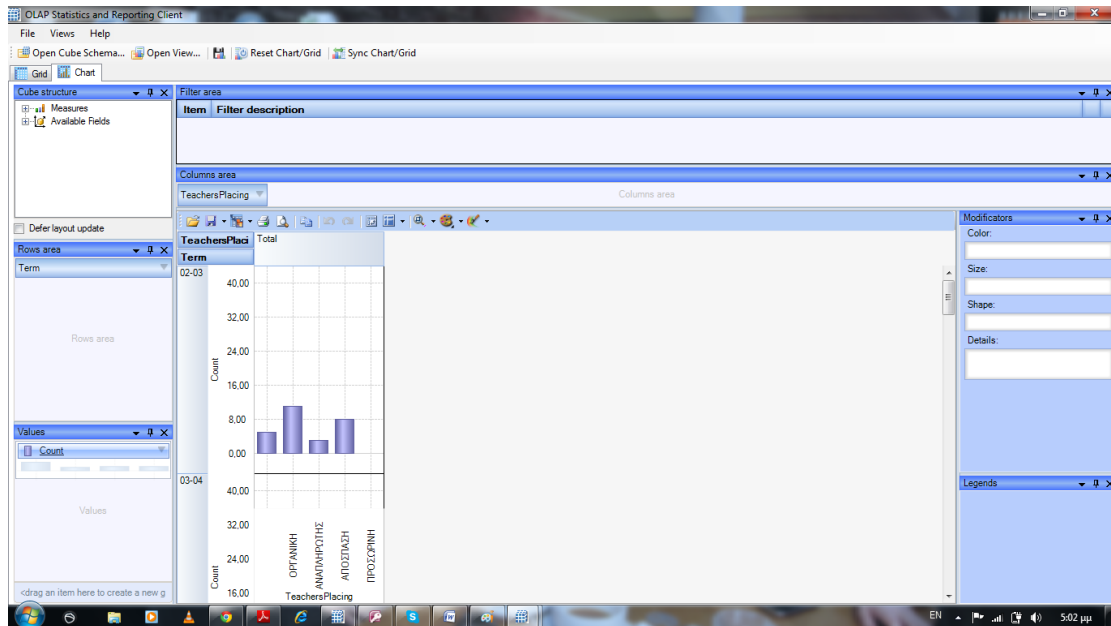
Παρουσιάζεται το πλήθος των καθηγητών τη κάθε σχολική χρονιά σύμφωνα με τη θέση τους (οργανική, προσωρινή, αναπληρωτής, απόσπαση) (εικόνα 3.37). Παρατηρούμε ότι το μεγαλύτερο ποσοστό των καθηγητών που διδάσκουν στο Γυμνάσιο έχει οργανική θέση ενώ το ποσοστό αυτών που έχουν προσωρινή θέση είναι πολύ μικρό.

The screenshot shows the OLAP Statistics and Reporting Client interface. The main area displays a pivot table with 'Term' on the rows and 'TeachersPlacing' on the columns. The data is summarized in the following table:

Term	02-03	03-04	04-05	05-06	06-07	07-08	08-09	09-10	10-11	11-12	Total
TeachersPlacing											
5				1	1	4		2	4	1	18
ΟΡΓΑΝΙΚΗ	11	20	20	18	22	21	24	28		11	175
ΑΝΑΠΛΗΡΩΤΗΣ	3										3
ΑΠΟΣΠΑΣΗ	8	7	8	9	7	6	9	5		1	60
ΠΡΟΣΩΡΙΝΗ		2	2			3	2	2		1	12
Total	27	29	30	28	30	34	35	37	4	14	268

Εικόνα 3.37 : Πλήθος καθηγητών ανά σχολική χρονιά και είδος θέσης

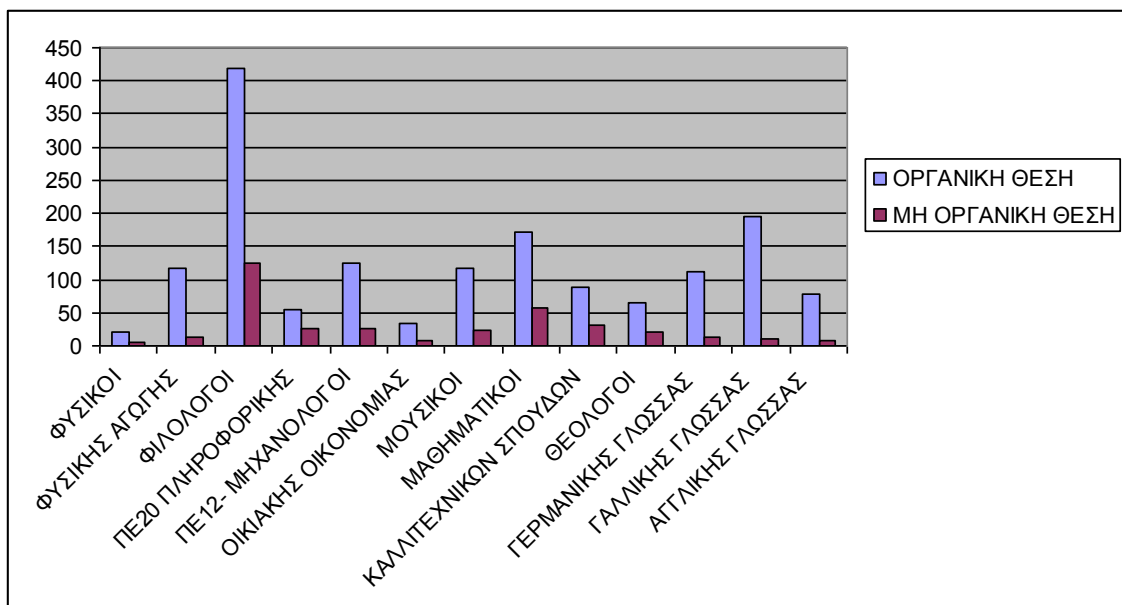
Τέλος υπάρχει και ένα σεβαστό ποσοστό (εικόνα 3.38) καθηγητών που βρίσκονται στο σχολείο με απόσπαση. Με βάση αυτών των δεδομένων καταλήγουμε στο συμπέρασμα ότι τα Γυμνάσια έχουν σταθερό και μόνιμο εκπαιδευτικό προσωπικό, πράγμα που είναι καλό τόσο για τα σχολεία όσο και για τους μαθητές.



Εικόνα 3.38 : Γράφημα με το πλήθος καθηγητών ανά σχολική χρονιά και είδος θέσης

17. Πόσες είναι οι ώρες μαθήματος ανά είδος θέσης και ανά φύλο

Εδώ επιχειρείται να δοθεί άλλη μια ποσοτική διάσταση στην ανάλυση της εκπαιδευτικής διαδικασίας, με σκοπό να πάρει η εκπαιδευτική αρχή τα ανάλογα αποτελέσματα, ώστε να βοηθηθεί στη διαμόρφωση του εκπαιδευτικού προγράμματος με βάση τις πραγματικές ανάγκες των εκπαιδευτικών ιδρυμάτων οι οποίες αποτυπώνονται μέσω των παραπάνω στοιχείων.



Εικόνα 3.39 : Ώρες μαθήματος ανά είδος θέσης και ειδικότητα

Σε αυτό το ερώτημα αποτυπώνεται (εικόνα 3.39) ο φόρτος εργασίας των καθηγητών της εκπαιδευτικής περιφέρειας που εξετάζεται προκειμένου να δοθεί μια εικόνα για τον αριθμό των ωρών που κάθε ειδικότητα απασχολείται. Αν και δεν υπάρχουν στην βάση δεδομένων επαρκή στοιχεία για όλες τις ειδικότητες παρατηρείται ότι σε όλους τους κλάδους οι καθηγητές με οργανικές θέσεις έχουν περισσότερες ώρες όπως είναι αναμενόμενο.

3.4.2 Συμπεράσματα Ερωτημάτων

Εφαρμόζοντας το πρόγραμμα OLAP Statistics & Reporting for Microsoft Access μπορέσαμε να συγκεντρώσουμε τις πληροφορίες που χρειάζονταν στα άτομα που λαμβάνουν αποφάσεις, αλληλεπιδρώντας με τα δεδομένα και ανακαλύπτοντας σχέσεις μέσα από αυτά με απλή πλοήγηση στα δεδομένα μας. Θέσαμε ερωτήματα στη βάση δεδομένων μας και ανακτήσαμε πληροφορίες και αποτελέσματα σε αρκετά ικανοποιητικό χρονικό διάστημα. Ακόμα τέθηκαν πολύπλοκοι υπολογισμοί στους κύβους OLAP παρέχοντας συγκεντρωτικά στοιχεία σε κάποια ερωτήματα. Στη συνέχεια βλέπουμε τα συμπεράσματα κατηγοριοποιημένα ανά ομάδα ερωτημάτων:

Συμπεράσματα όσων αφορά τους μαθητές και τις επιδόσεις τους με βάση τη γεωγραφική τους κατανομή

Από τα αποτελέσματα των ερωτημάτων που αφορούν τους μαθητές και τις επιδόσεις τους με βάση τη γεωγραφική τους κατανομή βγαίνουν αρκετά και χρήσιμα συμπεράσματα. Αρχικά βλέπουμε ότι οι μαθητές που κατοικούν στα περίχωρα της πόλης παρουσιάζουν μεγαλύτερο μ.ο. βαθμολογιών στις δύο από τις τρεις τάξεις του Γυμνασίου. Αυτό και μόνο έρχεται σε αντίθεση με αυτά που περιμέναμε. Οι μαθητές που ζουν μακριά από το σχολείο και σε περιοχές εκτός πόλης γενικά παρουσιάζουν αποκλίσεις στις επιδόσεις τους, μέσα όμως από την έρευνα μας βλέπουμε ότι στη πλειοψηφία τους ο μ.ο. των βαθμολογιών τους είναι καλύτερος από αυτόν των παιδιών που ζουν στη πόλη. Αυτό δείχνει ότι η απόδοση των μαθητών δεν καθορίζεται τόσο από την περιοχή που κατοικούν αλλά από τους ίδιους τους μαθητές και τη θέληση τους.

Στη συνέχεια παρατηρούμε ότι γενικά οι επιδόσεις των ελλήνων μαθητών είναι αρκετά υψηλές και μάλιστα υπάρχουν αρκετοί μαθητές των οποίων ο γενικός μέσος όρος της βαθμολογίας τους σε κάποια από τις τάξεις του Γυμνασίου άγγιξε το «20». Επίσης παρατηρούμε πάντα μέσα από τον μ.ο. των βαθμολογιών των μαθητών ότι η τάξη που τους δυσκολεύει περισσότερο είναι η Β Γυμνασίου και ότι οι βαθμολογίες των αγοριών είναι ελαφρώς υψηλότερες έναντι αυτών των κοριτσιών. Αυτό το τελευταίο ίσως να οφείλεται στο ότι η ηλικία που βρίσκονται τα παιδιά του Γυμνασίου είναι μεταβατική για αυτά και κυρίως για τα κορίτσια.

Στα σχολεία εκτός των ελλήνων μαθητών υπάρχουν και μαθητές που κατάγονται από άλλες χώρες εκτός Ελλάδας. Παρατηρούμε ότι οι επιδόσεις τους στα μαθήματα είναι αρκετά χαμηλές εκτός μερικών ελάχιστων εξαιρέσεων. Αυτό βασικά οφείλεται στο ότι η ελληνική γλώσσα δεν είναι η μητρική τους γλώσσα οπότε αυξάνεται η δυσκολία τους και επίσης στο ότι ίσως από αυτή τη μικρή ηλικία να αναγκάζονται να ψάχνουν κάποια εργασία ώστε να βοηθήσουν την οικογένεια τους αν το εισόδημα τους είναι χαμηλό. Σε αυτή τη περίπτωση θα έπρεπε το κράτος να μεριμνήσει παρέχοντας βοήθεια σε αυτούς τους μαθητές μέσω ενισχυτικής διδασκαλίας στοχευμένη στην αδυναμία κατανόησης της γλώσσας.

Συμπεράσματα για τους μαθητές και το φαινόμενο των απουσιών

Μέσα από την ανάλυση μας διαπιστώσαμε ότι κάθε χρόνο γίνονται αρκετές απουσίες από τους μαθητές αλλά κυρίως μέσα στα ανώτατα επιτρεπτά όρια. Βλέπουμε ότι στις δύο πρώτες τάξεις του Γυμνασίου οι απουσίες είναι περίπου ίδιες στα αγόρια και στα κορίτσια, ενώ στη Γ Γυμνασίου έχουμε αύξηση των απουσιών και κυρίως των αδικαιολόγητων, από τη πλευρά των αγοριών. Από αυτό συμπεραίνουμε ότι τα αγόρια όσο μεγαλώνουν γίνονται πιο ατίθασα και ότι επηρεάζονται από παρέες και εξωσχολικούς ίσως παράγοντες ενώ αντιθέτως τα κορίτσια όσο μεγαλώνουν γίνονται πιο συνεπείς. Το σημαντικότερο όμως είναι ότι σε αυτά τα δέκα χρόνια που εξετάσαμε μόνο ένας μαθητής ξεπέρασε το ανώτατο όριο αδικαιολόγητων απουσιών με αποτέλεσμα να μην προαχθεί στην επόμενη τάξη.

Γενικά συμπεράσματα για την επίδοση των μαθητών

Μέσα από την ανάλυση μας παρατηρούμαι ότι ο μ.ο. των επιδόσεων των αγοριών μειώνεται όσο μεγαλώνει η τάξη, ενώ τα κορίτσια παρουσιάζουν μια σταθερότητα στις επιδόσεις τους στη διάρκεια των τάξεων του Γυμνασίου.

Αναζητώντας τους αριστούχους μαθητές διαπιστώσαμε ότι στη πλειοψηφία τους είναι αγόρια το οποίο έρχεται σε αντίθεση με τη πρωτική πορεία των επιδόσεων τους όσο μεγαλώνει η τάξη. Απ' ότι φαίνεται όμως υπάρχει αρκετή απόκλιση μεταξύ των αριστούχων και των υπολοίπων μαθητών.

Όσον αφορά τους χαμηλούς βαθμούς σε κάποιο από τα μαθήματα, δηλαδή από 10 και κάτω, παρατηρούμε ότι σε αυτά τα δέκα χρόνια δεν υπάρχει μεγάλος αριθμός μαθητών που να πήρε τέτοιες βαθμολογίες σε κάποιο μάθημα. Αυτό μας δείχνει ότι υπάρχει ένα καλό επίπεδο στην εκπαιδευτική διαδικασία.

Τώρα όπως είναι φυσικό υψηλές βαθμολογίες παρουσιάζουν οι μαθητές στα μαθήματα της Φυσικής Αγωγής και τις Αισθητικής Αγωγής μιας και από τη φύση τους είναι τα πιο εύκολα μαθήματα. Στα υπόλοιπα μαθήματα οι βαθμολογίες των μαθητών είναι σαφώς πιο χαμηλές, με το μάθημα των Μαθηματικών να φαίνεται ότι δυσκολεύει πιο πολύ τους μαθητές και στις τρεις τάξεις του Γυμνασίου.

Στις ξένες γλώσσες οι μαθητές φαίνεται να τα πηγαίνουν καλά, κυρίως στα Αγγλικά και κατά δεύτερο λόγο στα Γαλλικά και τα Γερμανικά. Φαίνεται ότι υπάρχει μια προτίμηση στη γλώσσα των Γαλλικών έναντι των Γερμανικών μιας και το μάθημα της δεύτερης ξένης γλώσσας είναι επιλογής. Πάντως παρόλη την προτίμηση των μαθητών προς τα Γαλλικά, παρατηρήθηκαν υψηλότερες βαθμολογίες στους μαθητές που είχαν επιλέξει τα Γερμανικά ως δεύτερη ξένη γλώσσα.

Συμπεράσματα από τα δημογραφικά των μαθητών

Από την ανάλυση των διαθέσιμων δεδομένων παρατηρήσαμε ότι οι μαθητές που κατοικούν στη πόλη και πηγαίνουν σε πλήρη και οργανωμένα σχολεία είναι σαφώς περισσότεροι από τους μαθητές που κατοικούν και πηγαίνουν σε σχολεία εκτός αυτής. Έτσι λοιπόν συμπεραίνουμε ότι τη δεκαετία που

εξετάσαμε υπήρχε μια συγκέντρωση των πολιτών στις πόλεις και αναδείχθηκαν τα προβλήματα όσων κατοικούν στην επαρχία.

Συμπεράσματα όσων αφορά τους καθηγητές

Η ανάλυση της διαθέσιμης βάσης δεδομένων οδήγησε σε αρκετά σημαντικά συμπεράσματα όσων αφορά τους εκπαιδευτικούς που διδάσκουν στη δευτεροβάθμια εκπαίδευση της χώρας. Βλέπουμε ότι στη πλειοψηφία τους οι καθηγητές έχουν οργανική θέση στα σχολεία όπου ανήκουν και διδάσκουν και αυτό τους προσφέρει τη σιγουριά και την ασφάλεια που χρειάζονται ώστε να είναι αποδοτικοί στην εργασία τους.

Επίσης παρατηρούμε ότι υπάρχουν και καθηγητές με προσωρινή θέση στα σχολεία όπου είναι αναγκαίο. Οι περισσότεροι από τους καθηγητές που διαθέτουν μεταπτυχιακό τίτλο σπουδών ανήκουν σε αυτούς που η θέση τους είναι προσωρινή άρα προφανώς είναι μικρότερης ηλικίας και αυτό μας οδηγεί στο συμπέρασμα ότι οι νεότερες γενιές εκπαιδευτικών θα είναι πιο άρτια καταρτισμένες από τις παλαιότερες. Αυτό βέβαια οφείλεται στο ότι στη σημερινή εποχή έχουν αυξηθεί τα κριτήρια πρόσληψης και ο ανταγωνισμός και αυτό ωθεί τους υποψήφιους εκπαιδευτικούς στην απόκτηση περισσότερων τίτλων σπουδών.

4. Συμπεράσματα

Από την ανάλυση που πραγματοποιήθηκε προέκυψαν αρκετά και χρήσιμα συμπεράσματα όσων αφορά την προσέγγιση δεδομένων της εκπαίδευσης με εργαλεία OLAP. Στη πορεία αυτής της ανάλυσης έγινε φανερό ότι η προσέγγιση με εργαλεία OLAP ήταν η πιο κατάλληλη λόγω του όγκου και της φύσης της διαθέσιμης βάσης δεδομένων.

Ο τομέας της εκπαίδευσης είναι πολύ σημαντικός στη λειτουργία ενός κράτους, ο οποίος περιλαμβάνει ένα τεράστιο όγκο δεδομένων και πληροφοριών, ποσοτικών και ποιοτικών. Εξαιτίας αυτών των γεγονότων η ανάγκη της χρήσης OLAP είναι αρκετά μεγάλη μιας και τα εργαλεία OLAP ενδείκνυνται για την ανάλυση και επεξεργασία μεγάλων βάσεων δεδομένων.

Με τη χρήση OLAP η ανάλυση τέτοιων βάσεων δεδομένων γίνεται σε πολύ μικρότερο χρόνο απ' ό,τι με τα εργαλεία αναφορών των σχεσιακών βάσεων και με καλύτερη οπτικοποίηση των αποτελεσμάτων η οποία προσφέρεται από τη πολυδιάστατη μορφή των κύβων OLAP. Από τα παραπάνω γίνεται εμφανές ότι οι απαντήσεις που δίνονται σε ερωτήματα που τίθενται στη βάση δεδομένων εμφανίζονται πιο αναλυτικά με τα εργαλεία αναφορών OLAP και σε μικρότερο χρόνο απ' ό,τι με τα σχεσιακά εργαλεία αναφορών. Δηλαδή η απαντήσεις που δίνονται παρέχουν μεγαλύτερη ανάλυση και ακρίβεια.

Οι όποιες δυσκολίες εμφανίστηκαν στην συγκεκριμένη ανάλυση οφείλονταν στο ότι κάποια δεδομένα της διαθέσιμης βάσης δεδομένων δεν ήταν σωστά ομαδοποιημένα με αποτέλεσμα να μη μπορούν να δημιουργηθούν αρκετές ιεραρχίες κατά την ανάλυση. Επίσης τα δεδομένα στη πλειοψηφία τους ήταν ποιοτικά και όχι ποσοτικά με αποτέλεσμα να μην μπορούν να εξαχθούν όσα συμπεράσματα θα θέλαμε.

Όσων αφορά την εκπαίδευση από την ανάλυση των δεδομένων των μαθητών μέχρι και την ανάλυση ολόκληρης της βάσης δεδομένων στο τέλος κάθε εκπαιδευτικού έτους προκύπτουν χρήσιμα συμπεράσματα όσον αφορά στη διδασκαλία. Επίσης σε επίπεδο οργάνωσης της εκπαίδευσης μπορούν να γίνουν συντονισμένες ενέργειες προκειμένου να βελτιωθεί η λειτουργικότητα του εκπαιδευτικού συστήματος και όχι απλά μόνο η ποιότητα εκπαίδευσης που αυτό παρέχει.

Βέβαια για να γίνουν όλα τα παραπάνω πραγματικότητα πρέπει να έχει κανείς υπ' όψιν ότι απαιτείται η δημιουργία και συνεχής ενημέρωση μιας όσο το δυνατόν περισσότερο λεπτομερούς βάσης δεδομένων. Αυτό θα επιτρέψει στην ανάλυση να είναι πιο εμπειριστατωμένη και να έχει την «πολυδιάστατη» μορφή η οποία εφαρμόστηκε εδώ και φαίνεται να είναι η πιο ενδεδειγμένη για αυτές τις περιπτώσεις.

Οι βάσεις αυτές μπορούν να τηρούνται είτε σε επίπεδο εκπαιδευτικού ιδρύματος είτε σε επίπεδο κεντρικής οργάνωσης προκειμένου να γίνεται από ενδεχομένως πιο εξειδικευμένα άτομα πάνω στην ανάλυση δεδομένων.

Μελλοντικά η παρούσα έρευνα θα μπορούσε μέσω της χρήσης μιας νέας βάσης δεδομένων με περισσότερα στοιχεία όσον αφορά για παράδειγμα τις ώρες διδασκαλίας και τον αριθμό των ατόμων ανά τάξη να εξετάσει ερωτήματα όπως το πώς επηρεάζει ο αριθμός των μαθητών την ποιότητα του μαθήματος ή ακόμα αν ο αριθμός των ωρών διδασκαλίας επηρεάζει το μέσο όρο της βαθμολογίας των μαθητών στο μάθημα αυτό. Ακόμα στοιχεία από περισσότερες γεωγραφικές περιοχές θα επέτρεπαν την εξέταση των σχολείων σε πιο απομακρυσμένες περιοχές και την σύγκριση τους με τα μεγάλα αστικά κέντρα προκειμένου να λάβουμε μια εικόνα για τους παράγοντες που επηρεάζουν την απόδοση τους και αν η απόσταση αποτελεί έναν από αυτούς.

Τέλος πρέπει να αναφερθεί ότι η τεχνολογία OLAP γενικά είναι αρκετά εύχρηστη καθώς υπάρχουν αρκετά λογισμικά τα οποία την υποστηρίζουν και είναι διαθέσιμα στους χρήστες. Αυτά τα λογισμικά ποικίλουν σύμφωνα με τις υπηρεσίες που προσφέρουν, το επίπεδο ανάλυσης και το επίπεδο πολυπλοκότητας τους. Το λογισμικό που χρησιμοποιήθηκε στη συγκεκριμένη ανάλυση είναι αρκετά εύκολο στη χρήση του και παρότι είναι ιδιαίτερα αναλυτικό η εκμάθησή του και η ανάπτυξη του μοντέλου OLAP έγινε σε σχετικά μικρό χρονικό διάστημα.

Βιβλιογραφία

- ABDULLAH A., (2009): «Analysis of mealybug incidence on the cotton crop using ADSS-OLAP (Online Analytical Processing) tool , Volume 69, Issue 1», *Computers and Electronics in Agriculture* **69** (1): 59–72. doi:10.1016/j.compag.2009.07.003
- APOSTOLOS B., (2010): «Business Process Management: A Data Cube To Analyze Business Process Simulation Data For Decision Making», VDM Verlag Dr. Müller e.K. pp. 204 pp. ISBN 978-3-639-22216-6
- AVISON D., FITZGERALD G., (2007): «Προηγμένα Πληροφοριακά Συστήματα», Έκδοση 1^η, Εκδόσεις Νέων Τεχνολογιών
- BAKUYA T., MATSUI M., (1997): «Relational Database Management System», US5680614 A
- BACH P., TORBEN S., (2001): «Multidimensional Database Technology», *Distributed Systems Online (IEEE)*: 40–46. ISSN 0018-9162
- CARDENAS N., (2011): «ROLAP vs MOLAP Systems», *Academic Area: Computational Systems*
- CATTELL R., (2010): «Scalable SQL and NoSQL data stores», Published in: *Newsletter ACM SIGMOD Record archive*, Volume 39 Issue 4, December 2010, Pages 12-27, ACM New York, NY, USA
- CHAUDHURI S., DAYAL U., (1997): «An overview of data», Published in: *Newsletter ACM SIGMOD Record archive*, Volume 39 Issue 4, December 2010, Pages 12-27 ACM New York, NY, USA warehousing and OLAP technology», *SIGMOD Rec. (ACM)* 26 (1): 65. doi :10.1145/248603.248616. Retrieved 2008-03-20
- CHAUDHURI S., DAYAL U. (1997): «Data Warehousing and OLAP for Decision Support», ACM 0-89791-991
- CHAUDHURI S., DAYAL U., GANTI V., (2001): «Database Technology for Decision Support Systems», Published in: *Computer (Volume: 34, Issue 12)*
- CHHANDA R., (2009): «Distributed Database Systems», Published by Dorling Kindersley (India) Pvt. Ltd., licensees of Pearson Education in South Asia
- CODD E.F., (1970): «A relational model of data for large shared data banks», *Communications of the ACM*, Vol. 14, 377-381
- CODD E.F., CODD S.B., AND SALLEY C.T. (1993): «Providing OLAP (On-line Analytical Processing) to User-Analysts: An IT Mandate», Codd & Date, Inc. Retrieved 2008-03-05
- DEAN J. ET AL., (2008): «Bigtable: A Distributed Storage System for Structured Data», *Journal ACM Transactions on Computer Systems (TOCS) TOCS Homepage archive*, Volume 26 Issue 2, June 2008, Article No. 4, ACM New York, NY, USA
- DEAN J., GHEMAWAT S., (2008): «MapReduce: simplified data processing on large clusters», *Magazine Communications of the ACM - 50th anniversary issue: 1958 - 2008 CACM Homepage archive*, Volume 51, Issue 1
- DEEPAK P., (2007): «Business Intelligence for Telecommunications», CRC Press. pp. 294 pp. ISBN 0-8493-8792-2, Retrieved 2008-03-18
- DEVLIN B., COTE L.D., (1996): «Data Warehouse: From Architecture to Implementation», Addison-Wesley Longman Publishing Co., Inc. Boston, MA, USA

- ELNASRI R., NAVATHES S., (1996): «Θεμελιώδεις Αρχές Συστημάτων Βάσεων Δεδομένων», Διάυλος, Αθήνα
- EVERARD B., MORIS G., (1999): «Αποτελεσματική Εκπαιδευτική Διοίκηση», ΕΑΠ, Πάτρα
- GRAY J., CHAUDHURI S., LAYMAN A., REICHART H., VENKATRAO P., PIRAHESH, (1997): «Data Cube: {A} Relational Aggregation Operator Generalizing Group-By, Cross-Tab, and Sub-Totals», J. Data Mining and Knowledge Discovery **1** (1): 29–53. Retrieved 2008-03-20
- HARI M., (2007): «Introduction to OLAP - Slice, Dice and Drill», Data Warehousing Review. Retrieved 2008-03-18
- JOHNSON H. R., LARSON J. A., LAWRENCE J. D., (1979): «Network and Relational Modeling in a Common Data Base Architecture Environment»
- KIMBALL R., (2007): «Dimensional Relational vs. OLAP: The Final Deployment Conundrum»
- KIM W., (2008): «Introduction to Object-Oriented Databases», From Computer Systems Series, ISBN: 9780262512169
- LEMIRE D., (2007): «Data Warehousing and OLAP-A Research-Oriented Bibliography»
- LIU L., ÖZSU T.M., (2009): «Encyclopedia of Database Systems», 4100 p. 60 illus. ISBN 978-0-387-49616-0
- MADDEN S., (2012): «From Databases to Big Data», Massachusetts Institute of Technology
- MANYIKA J. ET AL., (2011): «Big data: The next frontier for innovation, competition, and productivity»,
- MICROSTRATEGY INCORPORATED, (1995): «The Case for Relational OLAP» (PDF). Retrieved 2008-03-20
- O'BRIEN J. A., MARAKAS G. M., (2009): «Management information systems (9th ed.)» Boston, MA: McGraw-Hill/Irwin
- OLAP COUNCIL, (1997): «OLAP Council White Paper», Retrieved 2008-03-18.
- PASUMANSKY M., WHITEHORN M, ZARE R.,(2007): «Fast Track to MDX», Second edition, ISBN 1-84628-174-1
- PIN P., CHEN S., (1976): «Transactions on Database Systems (TODS) - Special issue: papers from the international conference on very large data bases: September 22–24, 1975», Framingham, MA TODS Homepage archive
- RAMAKRISHNAN R. & GEHRKE J. (2002): «Συστήματα Διαχείρισης Βάσεων Δεδομένων», Έκδοση 2η, Εκδόσεις Τζιόλα/McGraw Hill.
- RIZZI S., (2003): «Association for Computing Machinery», ACM Press
- ROIGER R., GEATZ M., (2008): «Εξόρυξη Πληροφορίας – Ένας Εισαγωγικός Οδηγός με Παραδείγματα», Κλειδάριθμος, Αθήνα.
- STONEBRAKER M., (2010): «SQL databases v. NoSQL databases», Magazine Communications of the ACM, Volume 53, Issue 4, pages 10-11
- THIERAUF R., (1994): «Συστήματα Υποστήριξης Αποφάσεων Με Προσανατολισμό στο Χρήστη», Παπαζήση, Αθήνα.
- THOMSEN E., (1997): OLAP Solutions: Building Multidimensional Information Systems, 2nd Edition, John Wiley & Sons, ISBN 978-0-471-14931-6
- TSICHRITZIS D. C., LOCHOVSKY F. H., (1976): « Hierarchical Data-Base Management: A Survey», Journal ACM Computing Surveys (CSUR) Surveys Homepage archive, Volume 8 Issue 1, March 1976, Pages 105-123, ACM New York, NY, USA

- ULLMAN J.D., WIDOM J., (2008): «Βασικές αρχές για τα συστήματα βάσεων δεδομένων», Κλειδάριθμος
- VIESCAS J., (1997): «Ο Οδηγός της Microsoft για την Access», Κλειδάριθμος, Αθήνα.
- WANG C., WANG Q., REN K., LOU W., (2009): «Ensuring Data Storage Security in Cloud Computing», in Proc. of IWQoS'09
- WHITE T., (2009): «HADOOP: THE DEFINITIVE GUIDE», COPYRIGHT 2009, 978-0-596-52197-4
- WILLIAMS C., GARZA V.R., TUCKER S., MARCUS A.M., (1994, January 24): Multidimensional models boost viewing options. InfoWorld, 16(4)
- ΑΘΑΝΑΣΟΥΛΑ-ΡΕΠΠΑ Α., ΚΟΥΤΟΥΖΗΣ Μ., ΜΑΥΡΟΓΙΩΡΓΟΣ Γ., ΧΑΛΚΙΩΤΗΣ Δ., (1999): «Εκπαιδευτική Διοίκηση και Πολιτική», ΕΑΠ, Πάτρα.
- ΒΑΣΙΛΕΙΑΔΗΣ Π., (2009): «Προχωρημένα Θέματα Τεχνολογίας και Εφαρμογών Βάσεων Δεδομένων Αποθήκες Δεδομένων (Data Warehouses)», Πανεπιστήμιο Ιωαννίνων
- ΒΑΣΙΛΕΙΑΔΗΣ Π., ΣΚΙΑΔΟΠΟΥΛΟΣ Σ., ΣΕΛΛΗΣ Τ., (2002): «Θέματα Μοντελοποίησης και Επερώτησης Πολυδιάστατων Βάσεων Δεδομένων»
- ΒΟΥΤΣΙΝΑΣ Β., (2003): «Θέματα Επιχειρηματικής Νοημοσύνης – Θεωρητική Θεμελίωση & Εφαρμογές», Κωσταράκη, Αθήνα.
- ΔΕΡΒΟΣ Δ., (1995): «Μαθήματα Βάσεων Δεδομένων», Τόμος Α, Τσιόλας, Θεσσαλονίκη.
- ΚΑΜΠΟΥΡΙΔΗΣ Γ., (2002): «Οργάνωση και Διοίκηση Σχολικών Μονάδων», Κλειδάριθμος, Αθήνα.
- ΚΟΥΤΟΥΖΗΣ Μ., (1999): «Γενικές Αρχές Μάνατζμεντ», ΕΑΠ, Πάτρα.
- ΛΙΓΟΥΔΙΑΣΤΙΑΝΟΣ Σ., (2008): «Αποθήκες Δεδομένων», Σχολή Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών, Ε.Μ.Π.
- ΜΑΝΩΛΟΠΟΥΛΟΣ Ι., ΠΑΠΑΔΟΠΟΥΛΟΣ Α., (2001): «Συστήματα Βάσεων Δεδομένων, Θεωρία και Πρακτική Εφαρμογή», Αθήνα
- ΜΑΤΣΑΤΣΙΝΗΣ Ν., (2010) : « Συστήματα Υποστήριξης Αποφάσεων », Έκδοση 1^η, Εκδόσεις Νέων Τεχνολογιών.
- ΜΙΧΑΛΛΑΡΙΑΣ Η., (2004): «Μη Παραδοσιακές Τεχνικές Οπτικοποίησης Πολυδιάστατων Δεδομένων», Ε.Μ.Π. Σχολή Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών
- ΞΗΡΟΤΥΡΗΣ Ν., (2012): «Σχεδιασμός και Ανάπτυξη Προσαρμοζόμενου Συστήματος για την Διδασκαλία του Μαθήματος Βάσεις Δεδομένων», Τμήμα Ηλεκτρολόγων Μηχανικών και Τεχνολογίας, Πανεπιστήμιο Πατρών
- ΡΑΒΑΣΟΠΟΥΛΟΣ Γ., ΠΑΠΑΙΩΑΝΝΟΥ Ι., ΒΟΥΤΣΙΝΑΣ Β., (2009): «Η Αναλυτική Επεξεργασία Δεδομένων (On Line Analytical Processing) στην Υποστήριξη Αποφάσεων των Υπευθύνων Περιβαλλοντικής Εκπαίδευσης των Διευθύνσεων Εκπαίδευσης»
- ΣΚΟΥΡΛΑΣ Χ., (2000): «Σχεσιακές Βάσεις Δεδομένων», Έκδοση 2η, Εκδόσεις νέων τεχνολογιών μόν. ΕΠΕ.
- ΣΤΑΥΡΟΠΟΥΛΟΣ Π., (2009): «PhytoKaryon: Μία κυτταρολογική βάση δεδομένων των φυτών της Ευρώπης και της Μεσογείου – Αξιοποίηση και παρουσίαση δεδομένων», Πανεπιστήμιο Πατρών Τμήμα Φυσικής
- ΣΤΡΑΓΓΑΛΙΝΟΣ Ε., (2010): «Επέκταση του OLAP Μοντέλου σημαντικά δίκτυα», Τμήμα Μαθηματικών Πανεπιστημίου Πατρών

Υπερσύνδεσμοι

- http://dspace.lib.uom.gr/bitstream/2159/13821/1/Iosifidis_Msc2010.pdf,
Ιωσιφίδης Ελευθέριος, Διαχείριση Πολυδιάστατων Δεδομένων: Πειραματική
και Συγκριτική Αξιολόγηση της Απόδοσης Εμπορικών και Ανοικτού Κώδικα
DBMS, Θεσσαλονίκη 2010.
- [http://kpekastor.kas.sch.gr/peekpe/proceedings/synedria_9_ereunes/Ravasopoulos
et al.pdf](http://kpekastor.kas.sch.gr/peekpe/proceedings/synedria_9_ereunes/Ravasopoulos_et_al.pdf), Η Αναλυτική Επεξεργασία Δεδομένων (On Line Analytical
Processing) στην Υποστήριξη Αποφάσεων των Υπευθύνων Περιβαλλοντικής
Εκπαίδευσης των Διευθύνσεων Εκπαίδευσης, Βουτσινάς, 2009.
- <http://digilib.lib.unipi.gr/dspace/bitstream/unipi/4835/1/Staugianoudakis.pdf>,
Σταυγιανουδάκης Πέτρος, Συλλογή, Διασταύρωση, Διαχείριση και
Επιχειρησιακή Αξιοποίηση Μεταναστευτικών Δεδομένων Ασφαλιστικών
Οργανισμών, 2012.
- <http://dide.flo.sch.gr/Plinet/Tutorials/Tutorials-DataBasesTheory.html>, Στυλιάδης
Κωνσταντίνος, Η Θεωρία των Βάσεων Δεδομένων.
"OLAP and OLAP Server Definitions". The OLAP Council. 1995. Retrieved 2008-
03-18.
- "Computer Encyclopedia: multidimensional views". Answers.com. Retrieved
2008-03-05.
- "Just What Are Cubes Anyway? (A Painless Introduction to OLAP Technology)".
Msdn.microsoft.com. Retrieved 25-07-2012.
- <http://nemertes.lis.upatras.gr/jspui/bitstream/10889/4343/1/Thesis.pdf>,
Στραγγαλινός Ευάγγελος, Επέκταση του OLAP Μοντέλου με σημαντικά
δίκτυα, Πάτρα, 2010, ecoea@dipe.ach.sch.gr
- http://www.cs.uoi.gr/~pvassil/courses/db_III/, Βασιλειάδης Πάνος, Προχωρημένα
Θέματα Τεχνολογίας και Εφαρμογών Βάσεων Δεδομένων, 2009.
- <http://fse.tibiscus.ro/anale/Lucrari2011/125.pdf>, Codreanu Diana Elena et al,
Architecture on-line analytical processing. envir-edu@dide.ach.sch.gr