



**ΑΝΩΤΑΤΟ ΕΚΠΑΙΔΕΥΤΙΚΟ ΙΔΡΥΜΑ ΠΕΙΡΑΙΑ
ΤΕΧΝΟΛΟΓΙΚΟΥ ΤΟΜΕΑ**

**ΣΧΟΛΗ ΔΙΟΙΚΗΣΗΣ ΚΑΙ ΟΙΚΟΝΟΜΙΑΣ
ΤΜΗΜΑ ΔΙΟΙΚΗΣΗΣ ΕΠΙΧΕΙΡΗΣΕΩΝ
ΠΡΟΓΡΑΜΜΑ ΜΕΤΑΠΤΥΧΙΑΚΩΝ ΣΠΟΥΔΩΝ
«ΔΙΟΙΚΗΣΗ ΕΚΠΑΙΔΕΥΤΙΚΩΝ ΜΟΝΑΔΩΝ»**

ΕΡΕΥΝΗΤΙΚΗ ΜΕΤΑΠΤΥΧΙΑΚΗ ΕΡΓΑΣΙΑ

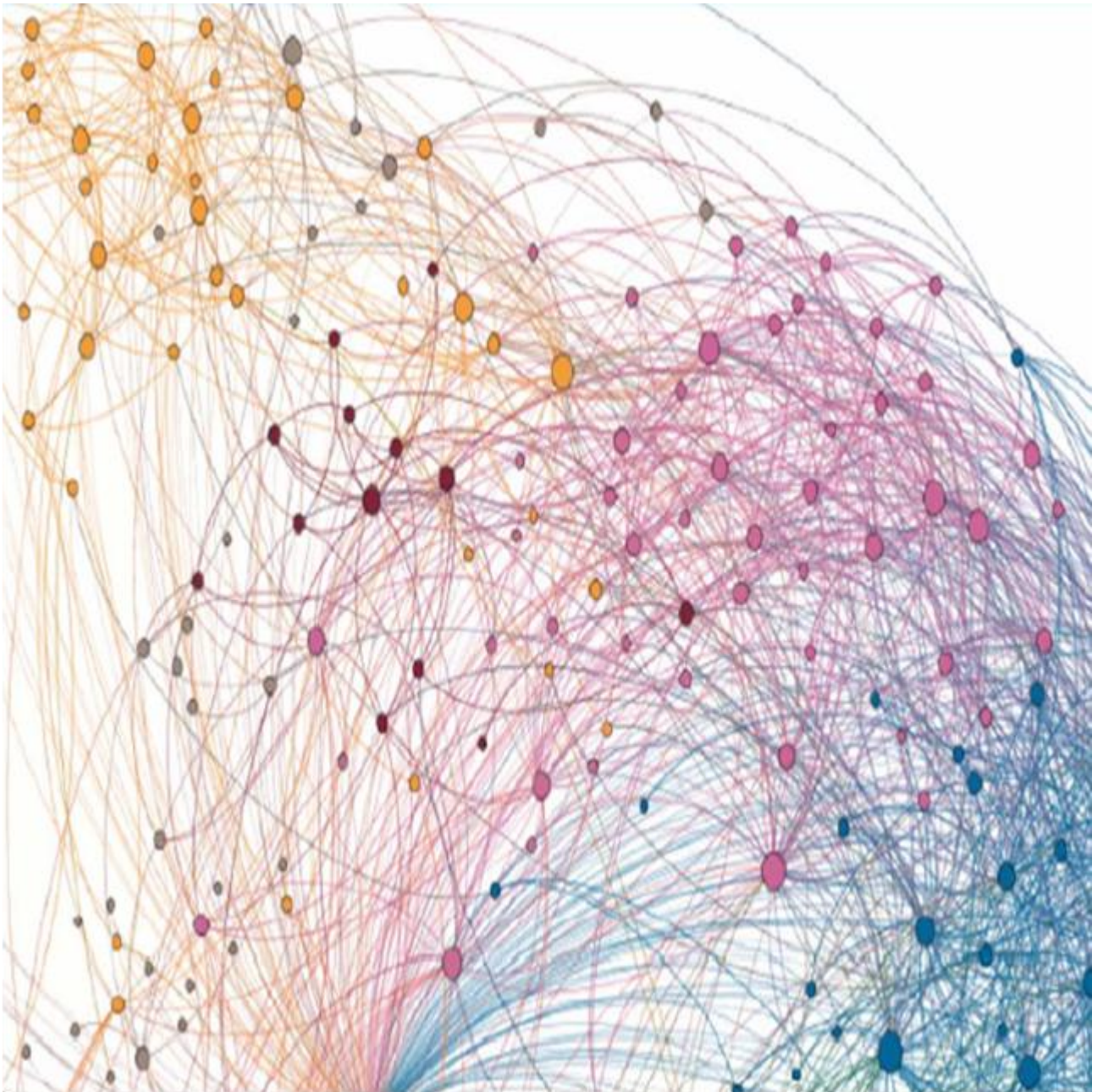
ΤΙΤΛΟΣ: «EDUCATIONAL DATA MINING

**Μια θεωρητική προσέγγιση πάνω στο Educational Data Mining
σε διαδικτυακά περιβάλλοντα και εφαρμογή σε test data»**

**ΝΑΤΑΛΙΑ ΚΩΣΤΟΠΟΥΛΟΥ
ΔΗΜΗΤΡΙΟΣ ΧΑΝΤΖΗΣ**

Επιβλέπων Καθηγητής: Αθανάσιος Σπυριδάκος

Ακαδημαϊκό Έτος 2017-2018



Εικόνα 1: Data Mining (classroom-aid.com)

Ευχαριστίες

Η διπλωματική αυτή εργασία αποτελεί έργο προσωπικής μας προσπάθειας και για να ολοκληρωθεί και να φτάσει στο επιθυμητό αυτό σημείο απαιτήθηκαν ώρες συγκέντρωσης και μελέτης.

Ευχαριστούμε θερμά τον επιβλέποντα καθηγητή μας για τις πολύτιμες συμβουλές και τη βοήθειά του.

Ευχαριστούμε επίσης όλους όσους μας βοήθησαν καθ' όλη την περίοδο εκπόνησης και συγγραφής δίνοντάς μας κουράγιο.

«Ευχαριστώ πολύ τους γονείς μου για τη στήριξή τους.»

Ναταλία Κωστοπούλου

«Ευχαριστώ πολύ τη σύζυγό μου, τον υιό μου και τους γονείς μου για τη βοήθειά τους»

Δημήτρης Χαντζής

Πρόλογος

Η επιστήμη και η τεχνολογία έχουν αλλάξει τη ζωή και την καθημερινότητα του ανθρώπου, από τις πιο απλές καθημερινές διαδικασίες μέχρι τις πιο περίπλοκες ιατρικές επεμβάσεις. Ακόμα και ο τρόπος που μαθαίνουμε έχει πλέον επηρεαστεί και ενισχύεται με τη βοήθεια τεχνολογιών.

Η ανάπτυξη εκπαιδευτικού λογισμικού όσο και ηλεκτρονικών βάσεων δεδομένων με πληροφορίες μαθητών, δημιούργησαν μεγάλες αποθήκες δεδομένων για το πώς μαθαίνουν οι μαθητές. Η χρήση του διαδικτύου στην εκπαίδευση δημιούργησε ένα νέο περιεχόμενο, το e-learning ή web based education όπου και εκεί αποθηκεύονται μεγάλες ποσότητες πληροφοριών σχετικά με την αλληλεπίδραση δασκάλου μαθητή.

Η επιστήμη του Educational Data Mining, συνδυάζοντας μεθόδους και τεχνικές από πολλά επιστημονικά πεδία, προσπάθησε να απαντήσει και προσεγγίσει τα ερωτήματα και τα προβλήματα που προέκυψαν από αυτές τις νέες διαστάσεις της εκπαίδευσης.

Παρακάτω θα δούμε μια παρουσίαση των τεχνικών διεργασιών του Educational Data Mining και της θεωρίας που το περιβάλλει. Θα δούμε τα βήματα που ακολουθούνται και θα προσπαθήσουμε να τα προσεγγίσουμε με ένα παράδειγμα (case).

Abstract

Science and technology have changed life and everyday life of humans, from the simplest daily processes to more complex medical interventions. Even the way we learn is more influenced and enhanced with the help of technologies.

The development of educational software as well as electronic databases with student information have created large data warehouses on how students learn. The use of Internet in education has created a new content, like e-learning or web based education, where large amounts of information about student- teacher interaction is stored.

Educational Data Mining, combining methods and techniques from many disciplines, has tried to answer and address the issues and problems arising from these new dimensions of education.

Below we will see a presentation of the educational processes of Educational Data Mining and the theory that surrounds it. We will look at the steps followed with an example.

ΠΕΡΙΕΧΟΜΕΝΑ

ΕΙΣΑΓΩΓΗ.....	7
1. E-LEARNING.....	9
1.1 ΜΑΘΗΣΗ ΚΑΙ E-LEARNING.....	9
1.2 ΚΑΤΗΓΟΡΙΕΣ E-LEARNING.....	11
1.3 MOODLE.....	13
2. EDUCATIONAL DATA MINING (EDM)	14
2.1 ΤΙ ΕΙΝΑΙ EDM.....	14
2.2 DATA MINING.....	15
2.2.1 Το DM , ή αλλιώς KDD.....	16
2.2.2 ΜΕΘΟΔΟΙ EDM.....	17
2.3 Στάδια εφαρμογής EDM.....	19
2.4 Βασικές Τεχνικές Κατηγοριοποίησης.....	20
3. ΚΥΡΙΕΣ ΕΦΑΡΜΟΓΕΣ ΜΕΘΟΔΩΝ EDM.....	21
4. ΑΝΑΛΥΣΗ ΔΕΔΟΜΕΝΩΝ ΑΠΟ ΕΚΠΑΙΔΕΥΤΙΚΗ ΠΛΑΤΦΟΡΜΑ.....	22
4.1 Κατανόηση του αντικειμένου της Ανάλυσης.....	22
4.2 Προετοιμασία βάσης δεδομένων.....	25
4.3 Εισαγωγή Δεδομένων (Import Data)	26
4.4 Χτίσιμο του Κύβου (Build the Cube)	33
4.5 Εξόρυξη Δεδομένων (Data Mining)	51
Επίλογος Συμπεράσματα.....	71
Πίνακας ακρωνύμων.....	72
Κατάλογος Εικόνων.....	73
ΒΙΒΛΙΟΓΡΑΦΙΑ.....	77



Εικόνα 2: Big Data (laceproject.eu)

ΕΙΣΑΓΩΓΗ

Η επιστήμη και η τεχνολογία έχουν αλλάξει τη ζωή και την καθημερινότητα του ανθρώπου, από τις πιο απλές καθημερινές διαδικασίες μέχρι τις πιο περίπλοκες ιατρικές επεμβάσεις. Ακόμα και ο τρόπος που μαθαίνουμε έχει πλέον επηρεαστεί και ενισχύεται με τη βοήθεια τεχνολογιών.

Η ανάπτυξη εκπαιδευτικού λογισμικού όσο και ηλεκτρονικών βάσεων δεδομένων με πληροφορίες μαθητών, δημιούργησαν μεγάλες αποθήκες δεδομένων για το πώς μαθαίνουν οι μαθητές. Η χρήση του διαδικτύου στην εκπαίδευση δημιούργησε ένα νέο περιεχόμενο, το e-learning ή web based education όπου και εκεί αποθηκεύονται μεγάλες ποσότητες πληροφοριών σχετικά με την αλληλεπίδραση δασκάλου μαθητή.

Η χρήση δεδομένων για τη λήψη αποφάσεων δεν είναι καινούρια ιδέα. Οι επιχειρήσεις χρησιμοποιούν ανάλυση δεδομένων των πελατών τους για τη λήψη αποφάσεων, η ιατρική χρησιμοποιεί δεδομένα και ανάλυση για τη διαγνωστική διαδικασία. Σε πολλούς επιστημονικούς τομείς τα δεδομένα πλέον παίζουν τον πρωταρχικό ρόλο της έρευνας.

Έτσι και στην εκπαίδευση, με τη πάροδο του χρόνου και την ανάπτυξη ηλεκτρονικών συστημάτων ειδικά για την εκπαίδευση άρχισε να διαφαίνεται η ανάγκη της χρησιμοποίησης των δεδομένων αυτών που συγκεντρώνονταν στην βελτίωση λήψης αποφάσεων, διαχείρισης των εκπαιδευτικών ιδρυμάτων και της παιδαγωγικής προσέγγισης στις τάξεις.

Η επιστήμη του Educational Data Mining, συνδυάζοντας μεθόδους και τεχνικές από πολλά επιστημονικά πεδία, προσπάθησε να απαντήσει και προσεγγίσει τα ερωτήματα και τα προβλήματα που προέκυψαν από αυτές τις νέες διαστάσεις της εκπαίδευσης.

Ουσιαστικά αυτό που ξεκίνησε να ψάχνει η EDM είναι απλό. Την εφαρμογή τεχνικών, μεθόδων, και προσεγγίσεων από Data mining (DM) και Business Intelligence (B.I) μετά από τον πειραματισμό που έχει ήδη γίνει στον χώρο των επιχειρήσεων, της ιατρικής κ.τ.λ. Υπό αυτήν την έννοια, η EDM θέλει, χρησιμοποιώντας τα αποθηκευμένα δεδομένα, να φτάσει σε μια καλύτερη κατανόηση των μαθητών και της μάθησης γενικότερα, και να αναπτύξει υπολογιστικές προσεγγίσεις που συνδυάζουν δεδομένα και θεωρία για τη μετατροπή της πρακτικής προς όφελος των μαθητών. (Ventura).

Ο πρώτος στόχος ενός αναλυτή που θέλει να χρησιμοποιήσει Data Mining (DM) στην εκπαίδευση είναι να συλλέξει δεδομένα για τις εκπαιδευτικές διαδικασίες, που στην πρώτη ανάγνωσή τους δεν έχουν κάποιο συγκεκριμένο νόημα, χωρίς την εφαρμογή κάποιας μεθόδου ανάλυσης.

Αντικείμενο της διπλωματικής εργασίας

Στόχος αυτής της εργασίας είναι η παρουσίαση του EDM (Educational Data Mining) μέσω από εικονικά δεδομένα (Test Data) ενός μαθήματος εξ' αποστάσεως εκπαίδευσης έτσι ώστε να γίνει κατανοητή η χρησιμότητα του αλλά και οι δυνατότητες που μπορεί να δώσει στην εκπαιδευτική κοινότητα για την καλυτέρευση της ποιότητας αλλά και της αποτελεσματικότητας των εξ' αποστάσεως μαθημάτων και όχι μόνο.

Η EDUCATIONAL DATA MINING ΣΗΜΕΡΑ

Τα πράγματα αλλάζουν δραματικά τα τελευταία χρόνια στο πως αποθηκεύονται τα εκπαιδευτικά δεδομένα. Έχουμε πλέον τεράστια σύνολα δεδομένων που είναι άμεσα διαθέσιμα. Αυτό προκάλεσε το ενδιαφέρον ερευνητών με αποτέλεσμα τα τελευταία χρόνια να έχουμε μια πληθώρα άρθρων, επιστημονικών εγγράφων, πρακτικών εφαρμογών.

Αρχικά ο στόχος ήταν να βρεθούν σημαντικές και συχνές αναφορές και μοτίβα στα διαθέσιμα δεδομένα. Ωστόσο οι ευκαιρίες που προσφέρονται σήμερα από τη στατιστική λειτουργική έρευνα και τις πληροφοριακές και επικοινωνιακές τεχνολογίες (ICT) ενδυναμώνουν περισσότερο τις δυνατότητες της ανάλυσης δεδομένων προς την κατεύθυνση του data analytics. (chen 2012).

Είναι μια επιστήμη που αναπτύσσεται διαρκώς, προσπαθώντας να απαντήσει σε εκπαιδευτικά ερωτήματα.



Εικόνα 3: E-Learning (tibau.org)

1.E-LEARNING

Υπάρχει στις μέρες μας μια πληθώρα εκπαιδευτικών συστημάτων/ περιβάλλοντα, όπως η παραδοσιακή τάξη, ερωτηματολόγια, τεστ, αλλά διαδικτυακά περιβάλλοντα.

Όλα αυτά μας δίνουν πληροφορίες και δεδομένα διαφορετικά μεταξύ τους. Δημιουργούν όμως και ερωτήματα , προβλήματα και ανάγκες διαφορετικές μεταξύ τους.

Με την ανάπτυξη εκπαιδευτικών συστημάτων μάθησης διαδικτυακά αυξήθηκαν οι ευκαιρίες να εκμεταλευτούμε την τεχνολογία για την διαμόρφωση της αξιολόγησης. Η ίδια αυτή τεχνολογία που υποστηρίζει την μάθηση μέσω υπολογιστών , υποστηρίζει και τη συλογή δεδομένων μέσα από αυτά τα περιβάλλοντα μάθησης.

Τα ηλεκτρονικά περιβάλλοντα μάθησης αφορούν την αλληλεπίδραση των χρηστών τους. Αυτή η αλληλεπίδραση παρέχει και ανατροφοδότηση. Έτσι, τα οφέλη τέτοιων δράσεων αφορούν όχι μόνο τους μαθητές αλλά και τους καθηγητές και διαχειριστές της εκπαίδευσης , και κυρίως την ίδια την εκπαιδευτική διαδικασία. Βασική λειτουργία ενός περιβάλλοντος μάθησης και ανάδρασης είναι η χρήση των καταγεγραμμένων δεδομένων για την αξιολόγηση και τη λήψη αποφάσεων με απώτερο σκοπό τη βελτίωση του.

1.1 ΜΑΘΗΣΗ ΚΑΙ E-LEARNING

Ο στόχος της εκπαίδευσης είναι η μάθηση. Μάθηση ορίζεται η διαδικασία κατά την οποία ο μαθητής αποκτά γνώσεις μέσω εκπαιδευτικού υλικού. Είναι η διαδικασία, που μετά από το αρχικό ερέθισμα, ενεργοποιεί την σταδιακή επεξεργασία πληροφοριών και τη αλληλουχία εσωτερικών λειτουργιών μαθαίνουμε μέσα από εμπειρίες και αντιλήψεις και αυτό έχει αντίκτυπο στη συμπεριφορά και στην προσωπικότητά μας.

Σε μια εποχή που κατακλύζεται από τεχνολογίες, η εκπαίδευση δεν ήταν δυνατό να μείνει ανεπηρέαστη. Η χρήση ηλεκτρονικών υπολογιστών και του διαδικτύου οδήγησε σε νέα μαθησιακά περιβάλλοντα, όπως η εξ αποστάσεως εκπαίδευση και το e-learning.

Στην εξ αποστάσεως εκπαίδευση ορισμένες μαθησιακές μέθοδοι θεωρούνται καταλληλότερες, ειδικά με την αξιοποίηση των νέων τεχνολογιών. Οι κυριότερες διδακτικές λειτουργίες ή στρατηγικές που φαίνεται να είναι αποτελεσματικές σε αυτό το πεδίο είναι η συνεργατική, η ενεργητική, η εποικοδομητική και η εξατομικευμένη μάθηση.

Το e-learning είναι η εκμάθηση που υποστηρίζεται μέσω της χρήσης πληροφοριών και της τεχνολογίας επικοινωνιών. Περιλαμβάνει τη χρήση ενός υπολογιστή ή μιας ηλεκτρονικής συσκευής για να παρέχει το υλικό εκπαίδευσης ή εκμάθησης, και λογισμικό που δημιουργείται για να διδάξει στο χρήστη τις νέες δεξιότητες και παραδίδεται χρησιμοποιώντας την τεχνολογία .

Τα βασικά στοιχεία ενός συστήματος e-learning αναφέρονται στη διασύνδεση των μαθητών, στη διασύνδεση των διαχειριστών, στα εργαλεία του διαχειριστή, στη διασύνδεση του περιεχομένου που περιέχονται σε όλες τις υπηρεσίες, στην αποθήκευση του περιεχομένου που βασίζεται σε συστήματα διαχείρισης σχεσιακών βάσεων δεδομένων και στο περιεχόμενο.

Τα τέσσερα βασικά δομικά στοιχεία μιας πλατφόρμας e-learning.

- Δημιουργία/Authoring
- Παράδοση/Delivery
- Διαχείριση/Management
- Μέτρηση/Measurement

Το e-learning εκμεταλεύεται το πλήθος εργαλείων που έχει αναπτύξει η πληροφορική για να μετατρέψει την τάξη και τα βιβλία σε ένα διαδραστικό εκπαιδευτικό περιβάλλον.

Οι ηλεκτρονικές αυτές πλατφόρμες εκπαιδευτικών εφαρμογών υπολογιστών μπορούν να χαρακτηριστούν με πολλά ονόματα όπως Learning Management Systems (LMS), Course Management Systems (CMS), Virtual Learning Environment (VLE). Και όλες δίνουν την δυνατότητα στους μαθητές να έχουν πρόσβαση σε μαθήματα υπό διάφορες μορφές (κείμενο, βίντεο) , να αλληλεπιδρούν με τους καθηγητές αλλά και μεταξύ τους .

Σύμφωνα με τον Piotrowsky, μια πλατφόρμα e-learning αντιπροσωπεύει ένα σύστημα το οποίο παρέχει ενσωματωμένη υποστήριξη σε έξι σημεία: την δημιουργία, την οργάνωση, την παράδοση, την επικοινωνία, τη συνεργασία και την αξιολόγηση.

Από τεχνική άποψη, υπάρχουν διάφοροι τύποι LMS, ορισμένοι από τους οποίους αντιπροσωπεύουν εμπορικές λύσεις (όπως το Blackboard/ WebCT) και άλλες λύσεις ανοιχτού κώδικα (όπως το Moodle). (Carolina Costa et al).

1.2 ΚΑΤΗΓΟΡΙΕΣ E-LEARNING

- **Συστήματα προσομοίωσης αίθουσας διδασκαλίας(virtual Classrooms)**
- **Απλά συγγραφικά εργαλεία (authoring tools)**
- **Learning Management Systems (LMS)**

Τα τελευταία 20 χρόνια, το ισχυρό λογισμικό για τη διαχείριση σύνθετων βάσεων δεδομένων συνδυάστηκε με ψηφιακά πλαίσια για τη διαχείριση εκπαιδευτικών προγραμμάτων και υλικού και εργαλείων αξιολόγησης. Το LMS επιτρέπει σε οποιονδήποτε να δημιουργεί, να παρακολουθεί, να διαχειρίζεται και να διανέμει εκπαιδευτικό υλικό. Τα προϊόντα και το λογισμικό LMS επιτρέπουν σε οργανισμούς να αναπτύξουν ηλεκτρονικά μαθήματα, να τα παραδώσουν με καινοτόμους τρόπους και ευελιξία και να διαχειριστούν τη χρήση τους στη διάρκεια των εκπαιδευτικών προγραμμάτων .

Οι τεχνολογίες πληροφοριών και επικοινωνίας (ICT) είναι πλέον άρρηκτα συνδεδεμένες με την εκπαίδευση. Τα συστήματα ηλεκτρονικής διαχείρισης ή αλλιώς Learning Management Systems (LMS) που εμπλέκονται σε όλες τις διαδικασίες συλλογής δεδομένων , επεξεργασίας πληροφοριών και δόμησης γνώσης , έχουν σημαντικό ρόλο στη διαδικασία της διδασκαλίας και της μάθησης. Είναι συνυφασμένα με τη διδακτική διδασκαλία που υποστηρίζουν τα LMS όπως το Moodle.

Το Moodle αποτελεί ένα είδος ηλεκτρονικής πλατφόρμας συστήματος LMS . Κύριο χαρακτηριστικό του είναι η αλληλεπίδραση μεταξύ των συμμετεχόντων, η ανατροφοδότηση, η δικτύωση και η συζήτηση. Εμπεριέχει ενότητες και δυνατότητες που επιτρέπουν την δημιουργία , την οργάνωση, την επικοινωνία αλλά και την αξιολόγηση.

Οι ηλεκτρονικές αυτές πλατφόρμες εκπαιδευτικών εφαρμογών υπολογιστών μπορούν να χαρακτηριστούν με πολλά ονόματα όπως Learning Management Systems (LMS), Course Management Systems (CMS), Virtual Learning Environment(VLE). Και όλες δίνουν την δυνατότητα στους μαθητές να έχουν πρόσβαση σε μαθήματα υπό διάφορες μορφές (κείμενο, βίντεο) , να αλληλεπιδρούν με τους καθηγητές αλλά και μεταξύ τους .

• **Κοινά χαρακτηριστικά**

Αυτόματη εγγραφή

- Εργαλεία δημιουργίας πληροφοριακού περιεχομένου (content creation tools)
- Εργαλεία επικοινωνίας και συνεργασίας μεταξύ των χρηστών (communication and collaboration tools)
- Εφαρμογές διαχείρισης (administrative applications)
- Εργαλεία σχεδίασης και διανομής διδακτικών ενοτήτων (course design and delivery tools)
- Εργαλεία ασφάλειας του συστήματος (security tools)
- Εργαλεία αλληλεπίδρασης με άλλα συστήματα (interface tools)
- Πολυγλωσσική υποστήριξη

- Εύκολη προσαρμογή και επανάχρηση υλικών με την πάροδο του χρόνου.
- Περισσότερες επιλογές για τους δημιουργούς του προγράμματος σπουδών, όπως η μέθοδος παράδοσης, ο σχεδιασμός των υλικών και οι τεχνικές αξιολόγησης.
- Οικονομική κλίμακας που καθιστά λιγότερο δαπανηρό για τους οργανισμούς να αναπτύσσουν και να διατηρούν περιεχόμενο για το οποίο βασίζονταν σε τρίτους.
- Βελτιώσεις στην επαγγελματική ανάπτυξη και αξιολόγηση,

Πλατφόρμες λοιπόν όπως το Moodle παρέχουν πολλές δυνατότητες και εργαλεία στους διαχειριστές για τη δημιουργία και διαμόρφωση διαδικτυακών μαθημάτων, ομάδων εργασίας, και γενικότερα για τη δημιουργία μιας μικρής κοινωνίας μάθησης. Η παιδαγωγική διάσταση που έχει ένα εργαλείο σαν αυτό είναι σημαντική και αξιοπρόσεκτη τόσο από την πλευρά του εκπαιδευόμενου, όσο και του εκπαιδευτή. Αποτελεί μια καινούρια και σύγχρονη προσέγγιση της εκπαίδευσης που εκμεταλλεύεται τα σύγχρονα εργαλεία που η τεχνολογία μπορεί να προσφέρει και τα χρησιμοποιεί προς όφελος τόσο του εκπαιδευόμενου όσο και τού μαθητή.

Δημοφιλέστερα συστήματα LMS



Εικόνα 4: Δημοφιλέστερα συστήματα LMS (Capterra, Inc)

1.3 MOODLE

Modular Object-oriented Dynamic Learning Environment.

Το Moodle σαν πλατφόρμα εμπεριέχει εγκατεστημένες λειτουργίες (εργαλεία) που δίνουν τη δυνατότητα αλληλεπίδρασης μεταξύ μαθητών / καθηγητών όπως chats, forums, εργαστήρια, αλλά και τη δυνατότητα ανάπτυξης λειτουργιών/ εργαλείων που μπορούν να συμπληρωθούν στην πλατφόρμα, όπως ερωματολογία , blogs, κ.α. Είναι ένα δωρεάν ηλεκτρονικό σύστημα διαχείρισης μάθησης, που παρέχει σε εκπαιδευτικούς σε όλο τον κόσμο μια λύση ανοιχτού κώδικα για την ηλεκτρονική μάθηση που είναι κλιμακωτή, προσαρμόσιμη και ασφαλής με τη μεγαλύτερη επιλογή διαθέσιμων δραστηριοτήτων. Υποστηρίζεται από ένα ενεργό δίκτυο πιστοποιημένων συνεργατών για βοήθεια με υποστήριξη και μια ενεργή κοινότητα προγραμματιστών, χρηστών και υποστηρικτών. (Capterra Inc.)

Η φιλοσοφία του Moodle περιλαμβάνει μια επικοινωνιακή και κοινωνική κατασκευαστική προσέγγιση στην εκπαίδευση, τονίζοντας ότι οι εκπαιδευόμενοι (και όχι μόνο οι εκπαιδευτικοί) μπορούν να συμβάλλουν στην εκπαιδευτική εμπειρία. Χρησιμοποιώντας αυτές τις παιδαγωγικές αρχές, το Moodle παρέχει ένα περιβάλλον για κοινότητες μάθησης

Χαρακτηριστικές λειτουργίες

- Κινητή εκμάθηση.
- Παρακολούθηση δεξιοτήτων.
- Κοινωνική μάθηση.
- Φοιτητική Πύλη.
- Σύγχρονη μάθηση.
- Δοκιμές / Αξιολογήσεις.
- Εκπαιδευτικές Εταιρείες.
- Βιντεοδιασκέψεις .

Εκτός όμως από την παροχή υπηρεσιών το Moodle ως ηλεκτρονικό σύστημα έχει τη δυνατότητα καταγραφής δεδομένων. Και τα δεδομένα που καταγράφει έχουν επιστημονική αξία. Το Moodle έχει τη δυνατότητα καταγραφής τόσο γενικευμένων όσο και πιο συγκεκριμένων δεδομένων. Παρέχει γενικά δεδομένα τόσο πάνω σε συγκεκριμένες διαδικασίες όσο και στη γενική παρουσία του χρήστη στην πλατφόρμα . Δεδομένα του χρήστη και της συμπεριφοράς του στη πλατφόρμα που είναι ξεχωριστά για κάθε μαθητή. Για παράδειγμα, μπορεί να καταγράψει το αν ένας μαθητής μπήκε στο σύστημα αλλά και πόσο χρόνο αφιέρωσε σε μια συγκεκριμένη λειτουργία του προγράμματος. Μπορεί να καταγράψει το χρόνο ασχολίας ενός μαθητή με ένα συγκεκριμένο έγγραφο ενός μαθήματος και αυτό μπορεί να χρησιμοποιηθεί από τους διαχειριστές για το σχεδιασμό της παρουσίασης του περιεχομένου των μαθημάτων.

Συγκεκριμένα μπορούμε να δημιουργήσουμε πολλών τύπων αναφορές όπως αρχεία καταγραφής για επιλεγμένες δραστηριότητες, καταγραφές σπουδαστές, αντικείμενα και χρονικά διαστήματα. Επίσης ζωντανούς κορμούς, οι οποία περιλαμβάνουν την πρόσφατη δραστηριότητα, εκθέσεις δραστηριότητας, που παρουσιάζουν τον αριθμό των απόψεων κάθε δραστηριότητας σε ένα μάθημα. Συμμετοχή σε μαθήματα, ανάλυση των ενεργειών επιλεγμένων φοιτητών για μια δεδομένη περίοδο και δραστηριότητα και στοιχεία σχετικά με την ολοκλήρωση της δραστηριότητας.

Τα δεδομένα από το Moodle μπορούν να δείξουν τις προτιμήσεις των χρηστών σε έγγραφα και λειτουργίες και αυτό να αποτελέσει χρήσιμο οδηγό για τους καθηγητές στην οργάνωση και δημιουργία των μαθημάτων και του περιεχομένου τους.

Η δυνατότητα αξιολόγησης είναι ένα ακόμα σημαντικό χαρακτηριστικό που παρέχει το Moodle καθώς και η παρακολούθηση των ακαδημαϊκών επιδόσεων των μαθητών. Τα δεδομένα σχετικά με τις επιδόσεις ενός μαθητή δίνουν πολύτιμες πληροφορίες που μπορεί να οδηγήσουν σε προβλέψεις σχετικά με τις επιδόσεις αυτές. Η αξιολόγηση μπορεί επίσης να δώσει καινούριες κατευθύνσεις στους καθηγητές για τη βελτίωση του περιεχομένου, της προσέγγισης και του τελικού στόχου των μαθημάτων .

Οι καταγραφές λοιπόν του Moodle αποτελούν πολύτιμη πηγή πληροφοριών για τους ερευνητές και για τους σχεδιαστές εκπαιδευτικών προγραμμάτων. Πληροφορίες για τα μαθήματα, τους μαθητές, τα ίδια τα εκπαιδευτικά προγράμματα.

Τα δεδομένα αυτά μπορούν να έχουν πολλές χρήσεις. Το δεδομένα που μπορεί να προσφέρει το Moodle μπορεί να βοηθήσουν τους ερευνητές να έχουν καλύτερη αντίληψη των μαθημάτων, της συμμετοχής των μαθητών, των προγραμμάτων, της ίδιας τη εκπαιδευτικής διαδικασίας. Γενικότερα μπορούν να υπάρξουν σημαντικά για τους ερευνητές ευρήματα από μια τέτοια ευρέως χρησιμοποιούμενη πλατφόρμα LMS.

Μπορούμε να εξαγάγουμε μοντέλα βασισμένα στη συμπεριφορά των χρηστών e-learning., μοντέλα που δίνουν τη δυνατότητα στους προγραμματιστές/ κατασκευαστές νέων προγραμμάτων και συστημάτων LMS να βρουν χρήσιμα μονοπάτια στη διαδικασία κατασκευής ενός συστήματος Moodle και να χτίσουν μαθήματα , προγράμματα με τον καλύτερο δυνατό τρόπο. Αλλά και για τους ίδιους τους χρήστες , μαθητές και καθηγητές , δίνουν την ευκαιρία για αυτοαξιολόγηση και βελτίωση. Τα δεδομένα αυτά και η ερμηνεία τους, μπορούν να οδηγήσουν σε βελτίωση της μάθησης και σε διορθώσεις των μαθημάτων ακόμα και κατά τη διάρκεια ενός προγράμματος.

2. EDUCATIONAL DATA MINING (EDM)

2.1 ΤΙ ΕΙΝΑΙ EDM

Σύμφωνα με τον ιστότοπο της κοινότητας για την εξόρυξη δεδομένων σχετικά με την εκπαίδευση, www.educationaldatamining.org, η εξόρυξη εκπαιδευτικών δεδομένων ορίζεται ως εξής: "Η εξόρυξη εκπαιδευτικών δεδομένων είναι μια αναδυόμενη επιστήμη η οποία ασχολείται με την ανάπτυξη μεθόδων για την εξερεύνηση των μοναδικών τύπων δεδομένων που προέρχονται από εκπαιδευτικά περιβάλλοντα και τη χρήση αυτών των μεθόδων για την καλύτερη κατανόηση των μαθητών και των πλαισίων στα οποία μαθαίνουν. Χρησιμοποιεί υπολογιστικές προσεγγίσεις για να αναλύσει εκπαιδευτικά δεδομένα με σκοπό την μελέτη ερωτημάτων που αφορούν την εκπαίδευση (Romero & Ventura 2007)

Έχει προταθεί ότι οι μέθοδοι εξόρυξης εκπαιδευτικών δεδομένων είναι συχνά διαφορετικές από τις συνήθεις μεθόδους εξόρυξης δεδομένων, λόγω της ανάγκης ρητής καταγραφής (και ευκαιριών αξιοποίησης) της ιεραρχίας και της μη ανεξαρτησίας σε εκπαιδευτικά δεδομένα [Baker] . Για το λόγο αυτό, γίνεται ολοένα και πιο συνηθισμένη η χρήση μοντέλων που προέρχονται από την ψυχομετρική βιβλιογραφία σε εκδόσεις εξόρυξης εκπαιδευτικών δεδομένων [Barnes 2005; Desmarais και Pu 2005; Pavlik et al. 2008].

Μιλώντας για EDM λοιπόν, αναφερόμαστε στις τεχνικές για την εύρεση και περιγραφή διαρθρωτικών μοντέλων από δεδομένα, σαν ένα εργαλείο που θα μας βοηθήσει να εξηγήσουμε αυτά τα δεδομένα και να κάνουμε προβλέψεις από αυτά. Να κατακτήσουμε γνώση και την

ικανότητα να τη χρησιμοποιήσουμε.

Το EDM μετατρέπει ωμά/ ακατέργαστα δεδομένα που προέρχονται από την εκπαιδευτικά συστήματα, σε χρήσιμες πληροφορίες που μπορούν να χρησιμοποιηθούν από εκπαιδευτικά προγράμματα, εκπαιδευτικό λογισμικό, καθηγητές, ερευνητές της εκπαίδευσης, κ.τλ.π.

Δεν διαφέρει στην εφαρμογή της από το Data Mining (DM), που χρησιμοποιείται σε επιχειρήσεις, στην ιατρική και σε άλλους τομείς , γιατί χρησιμοποιεί τα ίδια βήματα.

2.2 DATA MINING

Το Data Mining αφορά την επίλυση προβλημάτων μέσω της ανάλυσης δεδομένων που ήδη υπάρχουν σε βάσεις δεδομένων. Συγκεκριμένα το DM χαρακτηρίζεται ως η διαδικασία της ανακάλυψης μοτίβων από δεδομένα. Η διαδικασία αυτή πρέπει να είναι αυτόματη ή ημιαυτόματη και τα μοτίβα να έχουν νόημα, κυρίως στο ότι μπορούν αν οδηγήσουν σε κάποια εξέλιξη/πλεονέκτημα , συνήθως οικονομικό.

Το DM αφορά τεχνικές στην εύρεση και στην ερμηνεία κατασκευαστικών μοντέλων σε δεδομένα, σαν εργαλείο για να μας βοηθήσουν να εξηγήσουμε αυτά τα δεδομένα . Είναι η διαδικασία που εμπλέκεται στην εξαγωγή ενδιαφέρον, ερμηνεύσιμων, χρήσιμων και καινοτόμων πληροφοριών από δεδομένα.

Στο DM τα δεδομένα αποθηκεύονται ηλεκτρονικά και η αναζήτησή τους γίνεται αυτόματα με την ενίσχυση από υπολογιστές. Αυτό δεν είναι κάτι νέο. Οι επιστήμες της οικονομίας και της στατιστικής εργαζόνται εδώ και χρόνια πάνω στην ιδέα μοτίβων από δεδομένα που προήλθαν από αυτόματες διαδικασίες και μπορούν να χρησιμοποιηθούν για να γίνουν προβλέψεις.

Η ραγδαία ανάπτυξη των βάσεων δεδομένων και των συστημάτων αποθήκευσης, φέρνει την επιστήμη του DM στην πρώτη γραμμή των τεχνολογιών των επιχειρήσεων (BI).

Η Business Intelligence (BI) είναι μια τεχνολογικά καθοδηγούμενη διαδικασία για την ανάλυση δεδομένων και την παρουσίαση πληροφοριών που μπορούν να χρησιμοποιηθούν για να βοηθήσουν στελέχη, διαχειριστές και άλλους εταιρικούς τελικούς χρήστες να λαμβάνουν ενημερωμένες επιχειρηματικές αποφάσεις. (searchbusinessanalytics.techtarget.com)

“Έχει υπολογιστεί ότι ο όγκος των δεδομένων που αποθηκεύονται στις παγκόσμιες βάσεις δεδομένων διπλασιάζεται κάθε 20 μήνες, και παρόλο που θα ήταν δύσκολο να δικαιολογήσουμε αυτό το ποσό σε οποιαδήποτε ποσοτική έννοια, μπορούμε όλοι να κατανοήσουμε ποιοτικά τον ρυθμό ανάπτυξης.”(Witten & Frank 2005)

Η αύξηση των δυνατοτήτων μονάδων αποθήκευσης δεδομένων και η εξέλιξη στους υπολογιστές, οδήγησε σε προσβάσιμες μηχανές που είναι ικανές να αναλάβουν την αναζήτηση και ανάλυση των δεδομένων , αυξάνοντας έτσι τις ευκαιρίες για DM. Το DM μπορεί να μας βοηθήσει στην αποσαφήνιση της πολυπλοκότητας των δεδομένων που πλέον έχουμε στα χέρια μας και να οδηγήσει σε νέες ιδέες, νέα επιχειρηματικά μονοπάτια και σε ανταγωνιστικά πλεονεκτήματα.

“Η εξόρυξη δεδομένων ορίζεται ως η διαδικασία ανίχνευσης προτύπων στα δεδομένα. Η διαδικασία πρέπει να είναι αυτόματη ή (πιο συνηθισμένη) ημιαυτόματη. Τα πρότυπα που

ανακαλύφθηκαν πρέπει να έχουν νόημα στο ότι οδηγούν σε κάποιο πλεονέκτημα, συνήθως ένα οικονομικό πλεονέκτημα. Τα δεδομένα είναι πάντοτε παρόντα σε σημαντικές ποσότητες” (Witten & Frank 2005)

2.2.1 Το DM , ή αλλιώς KDD

Η συσσώρευση μεγάλου όγκου δεδομένων κάνει την ανάγκη για ανάλυσή τους ιδιαίτερα σημαντική και η επιστήμη του DM μπορεί αν ικανοποιήσει αυτήν την ανάγκη. Η αύξηση αυτή σε δεδομένα χρειάζεται ισχυρά και πολύπλευρα εργαλεία για την ανακάλυψη πληροφορίας και την μετατροπή της σε γνώση. Έτσι γεννήθηκε η DM. Μπορούμε να τη δούμε σαν φυσική εξέλιξη της τεχνολογίας της πληροφορίας.

KDD

Η KDD αφορά τη γενική διαδικασία της ανακάλυψης χρήσιμης γνώσης από δεδομένα. Περιλαμβάνει πολλά βήματα, όπως την προετοιμασία των δεδομένων, την αναζήτηση μοτίβων την αξιολόγηση των δεδομένων κ. α.

Σύμφωνα με τον Fayyad η KDD αποτελεί “τη διαδικασία της ανακάλυψης χρήσιμων γνώσεων από τα δεδομένα. Επικεντρώνεται στη συνολική διαδικασία της ανακάλυψης γνώσης από τα δεδομένα, συμπεριλαμβανομένου του τρόπου αποθήκευσης και πρόσβασης των δεδομένων, του τρόπου με τον οποίο οι αλγόριθμοι μπορούν να κλιμακωθούν σε μαζικά σύνολα δεδομένων και εξακολουθούν να λειτουργούν αποτελεσματικά, του τρόπου με τον οποίο μπορούν να ερμηνευτούν και να απεικονιστούν τα αποτελέσματα και του τρόπου με τον οποίο η συνολική αλληλεπίδραση ανθρώπου-υπολογιστών μοντελοποιείται και υποστηρίζεται.” (Fayyad et.al. 1996). Περιλαμβάνει πολλά βήματα , όπως την προετοιμασία των δεδομένων, την αναζήτηση μοτίβωντην αξιολόγηση των δεδομένων κ. α.

Το DM χρησιμοποιεί πολλές τεχνολογίες , χαρακτηριστικά μερικές από αυτές:

- στατιστική
- machine learning
- database systems data warehouses
- information retrieval
- buisness intelignce
- web searching engines

MACHINE LEARNING

Το DM χρησιμοποιεί αλγόριθμους Machine learning(ML) για την ανακάλυψη γνώσης μέσα από μεγάλες αποθήκες δεδομένων.

Η μηχανική μάθηση παρέχει την τεχνική βάση της εξόρυξης δεδομένων. Χρησιμοποιείται για την εξαγωγή πληροφοριών από τα ακατέργαστα δεδομένα σε βάσεις δεδομένων που εκφράζονται σε κατανοητή μορφή και μπορούν να χρησιμοποιηθούν για διάφορους σκοπούς. (Witten & Frank 2005)

2.2.2 ΜΕΘΟΔΟΙ EDM

Οι μέθοδοι που χρησιμοποιεί το EDM βασίζονται στις μεθόδους της εξόρυξης δεδομένων DM. Όμως επηρεάζεται και από ποικίλες άλλες επιστήμες όπως η ψυχομετρία, η στατιστική, η απεικόνιση πληροφοριών και computational modeling και άλλα πεδία επιστημών.

Οι Romero και Ventura [2007] κατηγοριοποιούν τις διαδικασίες στην εξόρυξη εκπαιδευτικών δεδομένων στις ακόλουθες κατηγορίες:

- Statistics and visualization
- Web mining
- Clustering, classification, and outlier detection
- Association rule mining and sequential pattern mining
- Text mining

Μια δεύτερη άποψη για την εξόρυξη εκπαιδευτικών δεδομένων δίνεται από τον Baker ο οποίος ταξινομεί τις διαδικασίες στην εξόρυξη εκπαιδευτικών δεδομένων ως εξής:

- Prediction
 - Classification
 - Regression
 - Density estimation
- Clustering
- Relationship mining
 - Association rule mining
 - Correlation mining

-Sequential pattern mining

-Causal data mining

- Distillation of data for human judgment
- Discovery with models

Οι πρώτες τρεις κατηγορίες είναι γνωστές στους ερευνητές καθώς προέρχονται από το DM. Η τέταρτη κατηγορία, αν και δεν είναι απαραίτητα παγκοσμίως θεωρούμενη ως εξόρυξη δεδομένων, συμφωνεί με την κατηγορία στατιστικών και οπτικοποίησης του Romero και Ventura και έχει εξέχουσα θέση τόσο στην δημοσιευμένη έρευνα EDM [Kay et al. 2006] και σε θεωρητικές συζητήσεις για την εξόρυξη εκπαιδευτικών δεδομένων [Tanimoto 2007]. (Baker& Yacef 2009)

Η πέμπτη κατηγορία της ταξινόμησης EDM του Baker δεν αποτελεί κλασική μέθοδο, παρά προκύπτει από την ανακάλυψη ενός μοντέλου, που επικυρώνεται επιστημονικά π.χ. prediction ή knowledge engineering, και μπορεί να χρησιμοποιηθεί σαν στοιχείο σε περαιτέρω αναλύσεις, όπως Relationship mining . Η ανακάλυψη με τα μοντέλα αποτελεί μια δημοφιλή μέθοδο στην έρευνα EDM, υποστηρίζοντας εξελιγμένες αναλύσεις (Baker& Yacef 2009).

Μέθοδος Πρόβλεψης(Prediction)

Η πρόβλεψη επιχειρεί να διαμορφώσει μοτίβα που επιτρέπουν την πρόβλεψη με τα διαθέσιμα δεδομένα εισόδου, που μπορεί να προέρχονται από διάφορες πηγες.

Συσταδοποίηση(Clustering)

Η συσταδοποίηση είναι μια διαδικασία διαίρεσης ενός συνόλου δεδομένων σε ένα σύνολο σημαντικών υποκατηγοριών, που ονομάζονται συμπλέγματα. Μας βοηθά τους να κατανοήσουμε τη φυσική ομαδοποίηση ή τη δομή σε ένα σύνολο δεδομένων.

Σχεσιακή Εξόρυξη (Relationship mining)

Είναι η τεχνική εξόρυξης δεδομένων για σχεσιακές βάσεις δεδομένων. Ενώ οι παραδοσιακοί αλγόριθμοι εξόρυξης δεδομένων αναζητούν μοτίβα σε ένα μόνο πίνακα , οι αλγόριθμοι σχεσιακής εξόρυξης δεδομένων αναζητούν πρότυπα μεταξύ πολλών πινάκων.

Απόσταξη δεδομένων για Ανθρώπινη Κρίση (Distillation of data for human judgment)

Η μέθοδος αυτή συνοψίζει και παρουσιάζει πληροφορίες με χρήσιμο, διαδραστικό και οπτικά ελκυστικό τρόπο, με σκοπό την κατανόηση μεγάλου όγκου εκπαιδευτικών δεδομένων και την καλύτερη υποστήριξη της λήψης αποφάσεων . Βοηθά τους εκπαιδευτικούς στην κατανόηση των πληροφοριών από την χρήση ή αποτελεσματικότητα των μαθητικών δραστηριοτήτων . Οι βασικές εφαρμογές της απόσταξης δεδομένων για την ανθρώπινη κρίση περιλαμβάνουν τον εντοπισμό των προτύπων στη μάθηση, τη συμπεριφορά, τις δυνατότητες συνεργασίας και τα δεδομένα των μαθητών , για μελλοντικές χρήσεις σε μοτίβα πρόβλεψης.

Ανακάλυψη με Μοντέλα(Discovery with models)

Στη αυτή τη μέθοδο, ένα μοντέλο αναπτύσσεται μέσω μηχανικής πρόβλεψης, ομαδοποίησης ή ανθρώπινης συλλογιστικής και στη συνέχεια χρησιμοποιείται ως συστατικό σε μια άλλη ανάλυση. Το μοντέλο που δημιουργείται επιτρέπει την περαιτέρω ανάλυση μεταξύ νέων προβλέψεων και μεταβλητών στη μελέτη.

Ανίχνευση Ακραίων Τιμών(Outlier Detection)

Η Ανίχνευση Ακραίων Τιμών είναι η διαδικασία ανίχνευσης και στη συνέχεια εξαίρεσης από ένα σύνολο δεδομένων. Ένα κομμάτι δεδομένων ή μια παρατήρηση που αποκλίνει από τον κανόνα ή το μέσο όρο του συνόλου δεδομένων μπορεί να είναι τυχαία, αλλά μπορεί επίσης να υποδεικνύει σφάλμα μέτρησης ή ότι το δεδομένο σύνολο δεδομένων έχει μια βαρύτατη κατανομή.

Ανάλυση Κοινωνικών Δικτύων (SNA)ή ανάλυση δόμησης δικτύων(Social Network Analysis)

Η ανάλυση κοινωνικού δικτύου (SNA) είναι μια διαδικασία ποσοτικής και ποιοτικής ανάλυσης ενός κοινωνικού δικτύου. Το SNA μετρά και χαρτογραφεί τη ροή των σχέσεων και των αλλαγών σχέσεων μεταξύ οντοτήτων που συμμετέχουν στη γνώση. Αυτές οι οντότητες περιλαμβάνουν ιστότοπους, υπολογιστές, ανθρώπους, ομάδες. Η δομή της αποτελείται από κόμβους, όπως ρόλοι ή άνθρωποι, και δεσμούς, όπως οι σχέσεις.

Διαδικασία Εξόρυξης(Process Mining)

Η Διαδικασία Εξόρυξης (PM) χρησιμοποιήσει αρχεία καταγραφής, διαδρομές ελέγχου, ή ίχνη για να ανακαλύψει, να παρακολουθήσει και να βελτιώσει τις εκπαιδευτικές διαδικασίες. Παρέχει μια γέφυρα μεταξύ της εξόρυξης δεδομένων (DM) και της μοντελοποίησης και ανάλυσης.

Εξόρυξη από δεδομένα Κειμένου(Text Mining)

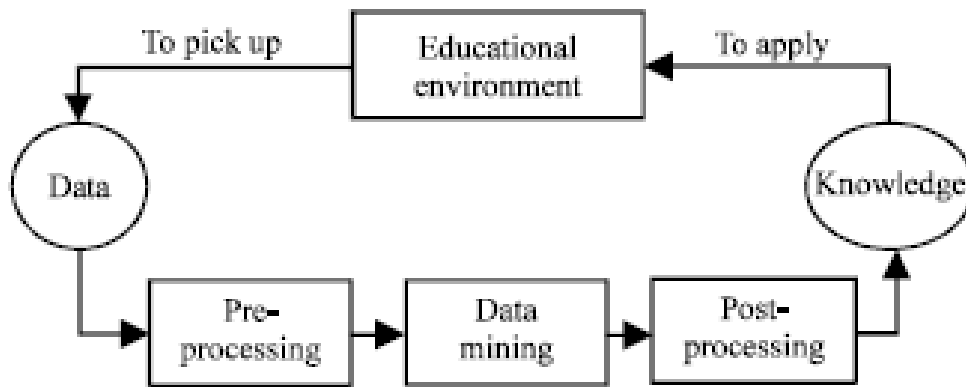
Το Text Analytics, γνωστό και ως εξόρυξη κειμένου, είναι η διαδικασία εξέτασης μεγάλων συλλογών γραπτών πόρων για τη δημιουργία νέων πληροφοριών και τη μετατροπή του κειμένου σε δομημένα δεδομένα για περαιτέρω ανάλυση. Η εξόρυξη κειμένου εντοπίζει γεγονότα και σχέσεις που διαφορετικά θα παρέμεναν θαμμένα στη μάζα μεγάλων δεδομένων κειμένου. Αυτά τα γεγονότα εξάγονται και μετατρέπονται σε δομημένα δεδομένα, για ανάλυση, οπτικοποίηση (π.χ. μέσω πινάκων, χαρτών μυαλού, γραφημάτων), ενσωμάτωσης σε βάσεις δεδομένων ή αποθήκες και περαιτέρω βελτίωσης χρησιμοποιώντας συστήματα μηχανικής μάθησης (ML). (www.linguamatics.com)

Ιχνηλάτηση της Γνώσης(Knowledge Tracking)

Παραγοντοποίηση Μητρώων(Matrix Factorization)

2.3 Στάδια εφαρμογής EDM

- Επιλογή Δεδομένων
- Προ επεξεργασία Δεδομένων
- Μετασχηματισμός Δεδομένων
- Εξόρυξη Γνώσης
- Ερμηνεία –Αξιολόγηση



Εικόνα 5: Educational data mining process (Garcia et al., 2011)

2.4 Βασικές Τεχνικές Κατηγοριοποίησης

Δέντρα Απόφασης (Decision trees)

Ένα δέντρο αποφάσεων είναι μια τεχνική μοντελοποίησης μηχανικής μάθησης για προβλήματα παλινδρόμησης και ταξινόμησης. Τα δέντρα αποφάσεων κάνουν διαδοχικές, ιεραρχικές αποφάσεις σχετικά με τα αποτελέσματα των μεταβλητών με βάση προγνωστικά δεδομένα, για να βρει λύσεις.

Bayesian Classifier

Ένας Bayesian Classifier είναι ένας αλγόριθμος που χρησιμοποιεί το θεώρημα του Bayes για την ταξινόμηση αντικειμένων. Οι κατηγοριοποιητές του Naive Bayes υποθέτουν ισχυρή ή αδύναμη ανεξαρτησία μεταξύ των χαρακτηριστικών των σημείων δεδομένων. Αυτοί οι ταξινομητές χρησιμοποιούνται ευρέως στη μηχανική μάθηση επειδή είναι απλοί στην εφαρμογή τους. Ένας Naive Bayes classifier δεν είναι ένας μόνο αλγόριθμος, αλλά μια οικογένεια αλγορίθμων μηχανικής μάθησης που κάνουν χρήση της στατιστικής ανεξαρτησίας.

Νευρωνικά Δίκτυα

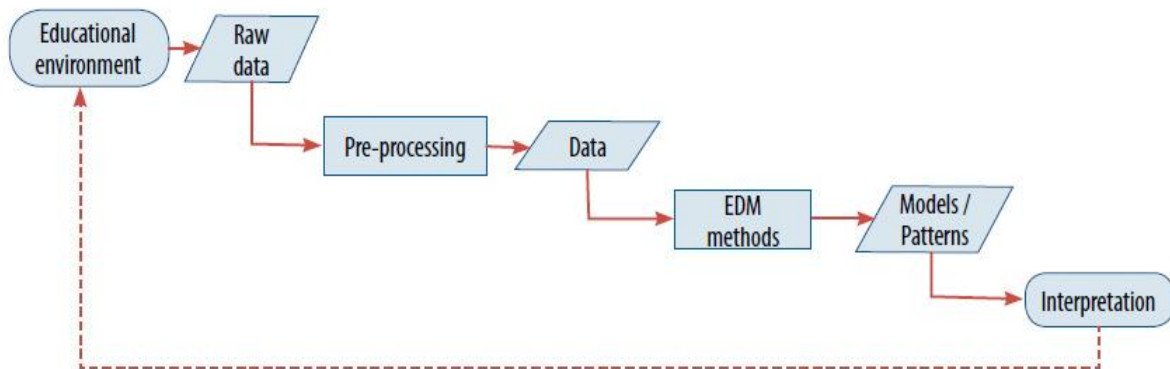
Τα τεχνητά Νευρωνικά δίκτυα βασίζονται στη λειτουργία και τη δομή του ανθρώπινου εγκεφάλου. Επεξεργάζονται και "μαθαίνουν" συγκρίνοντας την πρόβλεψή τους με τη γνωστή πραγματική τιμή της μεταβλητής απόκρισης. Τα σφάλματα από την αρχική πρόβλεψη των πρώτων αρχείων τροφοδοτούνται πίσω στο δίκτυο και χρησιμοποιούνται για να τροποποιήσουν τον αλγόριθμο των δικτύων για δεύτερη φορά. Αυτό συνεχίζεται για πολλές, πολλές επαναλήψεις.

K-Nearest Neighbor Classifiers

Ένας αλγόριθμος K-Nearest Neighbor Classifiers (K-NN), είναι μια προσέγγιση στην ταξινόμηση δεδομένων που υπολογίζει πόσο πιθανό είναι ένα σημείο δεδομένων να είναι μέλος μιας ομάδας ή άλλης. ανάλογα με την ομάδα στην οποία βρίσκονται τα πλησιέστερα σημεία δεδομένων. (www.techopedia.com)

Support Vector Machines

Πρόκειται για ένα τύπο αλγορίθμου που εκτελεί εποπτευόμενη μάθηση για ταξινόμηση ή παλινδρόμηση ομάδων δεδομένων. Χρησιμοποιούνται για την ταξινόμηση δύο ομάδων δεδομένων με παρόμοια ταξινόμηση. Οι αλγόριθμοι τραβούν γραμμές (υπερπληθή) για να διαχωρίσουν τις ομάδες σύμφωνα με τα πρότυπα.



Εικόνα 6: Επισκόπηση του τρόπου με τον οποίο εφαρμόζονται οι μέθοδοι EDM

3 ΚΥΡΙΕΣ ΕΦΑΡΜΟΓΕΣ ΜΕΘΟΔΩΝ EDM

Οι ερευνητές που ασχολούνται με το EDM μελετούν ποικίλους τομείς, μεταξύ των οποίων η ατομική μάθηση από εκπαιδευτικό λογισμικό, η συνεργατική μάθηση που υποστηρίζεται από ηλεκτρονικούς υπολογιστές, οι δοκιμές προσαρμογής στον υπολογιστή (και η δοκιμή γενικότερα) καθώς και τους παράγοντες που σχετίζονται με την αποτυχία.

Κύρια εφαρμογή της EDM αποτελεί η βελτίωση των μαθητικών μοντέλων. Τα μαθητικά μοντέλα αντιπροσωπεύουν πληροφορίες σχετικά με τα χαρακτηριστικά σπουδαστών, όπως οι γνώσεις τους ή οι μαθησιακές συμπεριφορές τους.

Η μοντελοποίηση των ατομικών διαφορών των σπουδαστών επιτρέπει στο λογισμικό να ανταποκρίνεται στις ατομικές αυτές διαφορές, βελτιώνοντας σημαντικά τη μάθηση των μαθητών [Corbett 2001]. Οι μέθοδοι εξόρυξης δεδομένων για εκπαιδευτικούς σκοπούς επιτρέπουν στους ερευνητές να διαμορφώνουν καλύτερα μοντέλα για την παρατήρηση συμπεριφορών και διερεύνηση ερωτημάτων. Για παράδειγμα, τα τελευταία χρόνια, οι ερευνητές χρησιμοποίησαν EDM μεθόδους για να διερευνήσουν το ερώτημα αν ένας φοιτητής εξαπατά το σύστημα [Baker et al. 2004].

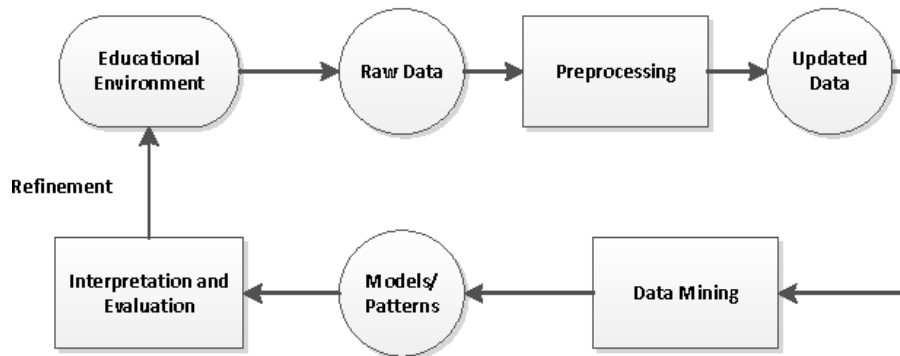
Ένας δεύτερος βασικός τομέας εφαρμογής των μεθόδων EDM ήταν η ανακάλυψη ή η βελτίωση μοντέλων της δομής της γνώσης ενός τομέα. Μέσω του συνδυασμού ψυχομετρικών μοντέλων και αλγόριθμων μηχανικής μάθησης, αρκετοί ερευνητές μπόρεσαν να αναπτύξουν αυτοματοποιημένες προσεγγίσεις που μπορούν να ανακαλύψουν ακριβή μοντέλα δομής της γνώσης.

Ένα τρίτο βασικό πεδίο εφαρμογής των μεθόδων EDM αφορά την παιδαγωγική υποστήριξη σε περιβάλλοντα μάθησης, αλλά και σε λογισμικό εκμάθησης ή μαθησιακές συμπεριφορές, στην ανακάλυψη των τύπων παιδαγωγικής υποστήριξης, δηλαδή που μπορεί είναι πιο αποτελεσματική.

Ένας τέταρτος βασικός τομέας εφαρμογής των μεθόδων EDM είναι η αναζήτηση στοιχείων για την τελειοποίηση και επέκταση εκπαιδευτικών θεωριών και γνωστών εκπαιδευτικών φαινομένων,

και της βαθύτερης κατανόησης των βασικών παραγόντων που επηρεάζουν τη μάθηση, συχνά με σκοπό το σχεδιασμό καλύτερων συστημάτων μάθησης. [Baker et al. 2004].

Για παράδειγμα, οι Gong, Rai και Heffernan [2009] διερεύνησαν τις επιπτώσεις της αυτοπειθαρχίας στην εκμάθηση και διαπίστωσαν ότι, ενώ συσχετίζονταν με την υψηλότερη εισερχόμενη γνώση και λιγότερα λάθη, ο πραγματικός αντίκτυπος στη μάθηση ήταν οριακός.



Εικόνα 7: Διαδικασία ανακάλυψης γνώσης στα Εκπαιδευτικά Ιδρύματα

4 ΑΝΑΛΥΣΗ ΔΕΔΟΜΕΝΩΝ ΑΠΟ ΕΚΠΑΙΔΕΥΤΙΚΗ ΠΛΑΤΦΟΡΜΑ

Στις παραπάνω ενότητες αναφέρθηκαν στην μάθηση μέσω elearning και στην εξόρυξη γνώσης από δεδομένα που σχετίζονται με την εκπαίδευση.

Τα δεδομένα που μπορούν να βγουν από κει μπορούν να χρησιμοποιηθούν για την πρόβλεψη της απόδοσης των εκπαιδευομένων από την πλατφόρμα εξ' αποστάσεως εκπαίδευσης και κατ' επέκταση την τροποποίηση του προγράμματος εκπαίδευσης με βάση τις ανάγκες των εκπαιδευομένων έτσι ώστε να επιτευχθεί ο εκπαιδευτικός στόχος.

4.1 Κατανόηση του αντικειμένου της Ανάλυσης

Το γεγονός ότι στην περίπτωση που θα εξεταστεί, τα δεδομένα μας είναι Test Data και από αυτό το εικονικό διαδικτυακό μάθημα έχουν δημιουργηθεί δεδομένα που σχετίζονται με τους εκπαιδευόμενους και τον τρόπο παρακολούθησης του μαθήματος. Όλα τα δεδομένα που έχουν δημιουργηθεί χωρίζονται σε 3 κατηγορίες και είναι αυτές που ακολουθούν.

Δημογραφικά χαρακτηριστικά του εκπαιδευόμενου

ΦΥΛΟ

ΗΛΙΚΙΑ

ΟΙΚΟΓΕΝΕΙΑΚΗ ΚΑΤΑΣΤΑΣΗ

Δραστηριότητα του εκπαιδευόμενου στην ιστοσελίδα του διαδικτυακού μαθήματος

ΕΠΙΣΚΕΨΕΙΣ

Μ.Ο. ΠΑΡΑΜΟΝΗΣ

ΣΥΝΟΛΙΚΟΣ ΧΡΟΝΟΣ ΠΑΡΑΜΟΝΗΣ

ΚΑΤΕΒΑΣΕ

ΕΙΔΕ url ΣΥΝΟΛΙΚΑ

ΕΙΔΕ url Μ.Ο. ΑΝΑ ΕΠΙΣΚΕΨΗ

ΑΠΟΥΣΙΕΣ

Επίδοση του εκπαιδευόμενου στο διαδικτυακό μάθημα

ΒΑΘΜΟΣ ΕΡΓΑΣΙΑΣ

ΒΑΘΜΟΣ 1 ΕΞΕΤΑΣΤΗ

ΒΑΘΜΟΣ 2 ΕΞΕΤΑΣΤΗ

ΤΕΛΙΚΟΣ ΒΑΘΜΟΣ

ΕΠΙΤΥΧΙΑ / ΑΠΟΤΥΧΙΑ

Περιγραφή των δεδομένων

Το σύνολο των δεδομένων όπως προαναφέρθηκε χωρίζονται σε 3 κατηγορίες, οι οποίες αποτελούνται από 15 μεταβλητές. Πιο συγκεκριμένα είναι οι παρακάτω.

ΦΥΛΟ

Το ΦΥΛΟ είναι το φύλο που έχει ο κάθε εκπαιδευόμενος.

ΗΛΙΚΙΑ

Η ΗΛΙΚΙΑ είναι η ηλικία που έχει ο κάθε ο εκπαιδευόμενος.

ΟΙΚΟΓΕΝΕΙΑΚΗ ΚΑΤΑΣΤΑΣΗ

Η ΟΙΚΟΓΕΝΕΙΑΚΗ ΚΑΤΑΣΤΑΣΗ είναι η οικογενειακή κατάσταση που έχει ο κάθε ο εκπαιδευόμενος (αν είναι παντρεμένος ή όχι).

ΕΠΙΣΚΕΨΕΙΣ

Οι ΕΠΙΣΚΕΨΕΙΣ είναι ο αριθμός επισκέψεων στην ιστοσελίδα του μαθήματος ο κάθε εκπαιδευόμενος.

Μ.Ο. ΠΑΡΑΜΟΝΗΣ

Ο Μ.Ο. ΠΑΡΑΜΟΝΗΣ είναι ο μέσος όρος του χρόνου παραμονής του εκπαιδευόμενου στην ιστοσελίδα του μαθήματος.

ΣΥΝΟΛΙΚΟΣ ΧΡΟΝΟΣ ΠΑΡΑΜΟΝΗΣ

Ο ΣΥΝΟΛΙΚΟΣ ΧΡΟΝΟΣ ΠΑΡΑΜΟΝΗΣ είναι ο συνολικός χρόνος παραμονής όλων των επισκέψεων του εκπαιδευόμενου στην ιστοσελίδα του μαθήματος.

ΚΑΤΕΒΑΣΕ

ΚΑΤΕΒΑΣΕ είναι ο αριθμός των αρχείων που έχει κατεβάσει από τα αρχεία που έχουν αναρτηθεί στην ιστοσελίδα του μαθήματος.

ΕΙΔΕ url ΣΥΝΟΛΙΚΑ

ΕΙΔΕ url ΣΥΝΟΛΙΚΑ είναι ο συνολικός αριθμός του υλικού (βίντεο, αρχεία κτλ) που είδε ο εκπαιδευόμενος στην ιστοσελίδα του μαθήματος.

ΕΙΔΕ url Μ.Ο. ΑΝΑ ΕΠΙΣΚΕΨΗ

ΕΙΔΕ url Μ.Ο. ΑΝΑ ΕΠΙΣΚΕΨΗ είναι ο μέσος όρος ανά επίσκεψη του υλικού (βίντεο, αρχεία κτλ) που είδε ο εκπαιδευόμενος στην ιστοσελίδα του μαθήματος.

ΒΑΘΜΟΣ ΕΡΓΑΣΙΑΣ

ΒΑΘΜΟΣ ΕΡΓΑΣΙΑΣ είναι ο βαθμός της εξαμηνιαίας εργασίας που παραδίδει μέσω της ιστοσελίδας του μαθήματος ο κάθε εκπαιδευόμενος λίγο πριν την λήξη του μαθήματος.

ΒΑΘΜΟΣ 1 ΕΞΕΤΑΣΤΗ

ΒΑΘΜΟΣ 1 ΕΞΕΤΑΣΤΗ είναι η βαθμολογία που έβαλε ο πρώτος εξεταστής στην τελική εξέταση του εκπαιδευόμενου.

ΒΑΘΜΟΣ 2 ΕΞΕΤΑΣΤΗ

ΒΑΘΜΟΣ 2 ΕΞΕΤΑΣΤΗ είναι η βαθμολογία που έβαλε ο δεύτερος εξεταστής στην τελική εξέταση του εκπαιδευόμενου.

ΤΕΛΙΚΟΣ ΒΑΘΜΟΣ

ΤΕΛΙΚΟΣ ΒΑΘΜΟΣ είναι
 $(\text{ΒΑΘΜΟΣ ΕΡΓΑΣΙΑΣ} * 0,40) + (((\text{ΒΑΘΜΟΣ 1 ΕΞΕΤΑΣΤΗ} + \text{ΒΑΘΜΟΣ 2 ΕΞΕΤΑΣΤΗ}) / 2) * 0,6)$

ΑΠΟΥΣΙΕΣ

ΑΠΟΥΣΙΕΣ είναι ο αριθμός των απουσιών στις προγραμματισμένες διαδικτυακές συναντήσεις.

ΕΠΙΤΥΧΙΑ / ΑΠΟΤΥΧΙΑ

ΕΠΙΤΥΧΙΑ / ΑΠΟΤΥΧΙΑ είναι η επιτυχία ή αποτυχία στο να Αριστεύσει ο εκπαιδευόμενος.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	
	ΚΛΕΙΔΙ	ΦΥΛΟ	ΗΛΙΚΙΑ	ΟΙΚΟΓΕΝΕΙΑΚΗ ΚΑΤΑΣΤΑΣΗ	ΕΠΙΣΚΕΦΕΙΣ	Μ.Ο. ΠΑΡΑΜΟΝΗΣ	ΣΥΝΟΛΙΚΟΣ ΧΡΟΝΟΣ ΠΑΡΑΜΟΝΗΣ	ΚΑΤΕΒΑΣΕ	ΕΙΔΕ url	ΕΙΔΕ url	Μ.Ο. ΑΝΑ ΕΡΓΑΣΙΑΣ	ΒΑΘΜΟΣ 1 ΕΞΕΤΑΣΤΗ	ΒΑΘΜΟΣ 2 ΕΞΕΤΑΣΤΗ	ΒΑΘΜΟΣ 3 ΕΞΕΤΑΣΤΗ	ΤΕΛΙΚΟΣ ΒΑΘΜΟΣ	ΑΠΟΥΣΙΕΣ	ΕΠΙΤΥΧΙΑ / ΑΠΟΤΥΧΙΑ
1	1	A	28	A	17	0,20	3,40	12	20	1,18	8	10	10	9,2	1	E	
2	2	Γ	24	A	27	0,03	0,68	7	0	0,00	8	9	10	8,9	1	A	
3	3	A	37	Π	28	0,09	2,52	9	30	1,07	9	9,3	8,5	8,94	0	A	
4	4	Γ	42	Π	32	0,18	5,76	8	29	0,91	8,5	9,3	9	8,89	2	A	
5	5	A	28	A	27	0,24	6,48	4	22	0,81	6,7	9	9	8,08	0	A	
6	6	A	33	Π	32	0,30	9,60	7	26	0,81	8	9	9	8,6	0	A	
7	7	A	34	A	21	0,70	14,70	7	29	1,38	7,8	9	8,5	8,37	0	A	
8	8	Γ	41	A	14	0,12	1,68	6	30	2,14	9	9	9	9	2	A	
9	9	A	29	A	22	0,15	3,30	8	30	1,36	8	10	8,5	8,75	2	A	
10	10	A	30	A	14	0,09	1,26	8	27	1,93	7,8	9,5	9	8,67	1	A	
11	11	Γ	41	Π	18	0,42	7,56	2	28	1,56	10	9,5	9	9,55	0	E	
12	12	A	39	Π	17	0,51	8,67	8	30	1,76	7	9,5	9	8,35	2	A	
13	13	Γ	28	A	15	0,29	4,35	10	36	2,40	8	9,5	10	9,05	1	E	
14	14	A	27	A	29	0,21	6,09	7	15	0,52	9,5	10	10	9,8	0	E	
15	15	Γ	30	A	32	0,09	2,88	6	7	0,22	9	10	10	9,6	0	E	
16	16	Γ	37	A	20	0,45	9,00	9	30	1,50	10	10	10	10	2	E	
17	17	A	32	A	29	0,37	10,73	10	26	0,90	6,5	9,5	10	8,45	1	A	
18	18	Γ	45	A	17	0,28	4,76	6	27	1,59	9	9,5	10	9,45	2	E	
19	19	A	50	Π	22	0,65	14,30	9	30	1,36	9,5	9	9	9,2	0	E	
20	20	A	29	Π	30	0,45	13,50	12	27	0,90	7	9	10	8,5	2	A	
21	21	Γ	45	Π	19	0,32	6,08	9	24	1,26	9	9,5	9	9,15	1	E	
22	22	Γ	44	A	15	0,74	11,10	7	30	2,00	9	9,5	9	9,15	2	E	
23	23	A	35	A	11	0,65	7,15	8	25	2,27	8	9	8,5	8,45	1	A	
24	24	A	34	Π	28	0,87	24,36	9	32	1,14	10	10	10	10	0	E	
25	25	A	28	A	17	0,85	14,45	7	30	1,76	8	10	8,5	8,75	2	A	
26	26	A	36	Π	27	0,21	5,67	6	24	0,89	10	10	10	10	2	E	
27	27	A	42	Π	19	0,23	4,37	7	25	1,32	8	9	8,5	8,45	2	A	

Εικόνα 8: Test Data

Τα δεδομένα μας ουσιαστικά είναι οι παρατηρήσεις από 207 εικονικοί εκπαιδευόμενοι που έχουν παρακολουθήσει το διαδικτυακό μάθημα μέσω της πλατφόρμας εκπαίδευσης.

Στο σύνολο των δεδομένων βλέπουμε ότι υπάρχουν μεταβλητές που αφορούν στοιχεία που δεν μπορούν να επηρεαστούν από την από την αλλαγή της ποιότητας του μαθήματος.

Είναι αυτά που σχετίζονται με δημογραφικά χαρακτηριστικά του εκπαιδευόμενου όπως το φύλο του, η ηλικία του και οικογενειακή του κατάσταση.

Σημαντική κατηγορία μεταβλητών είναι αυτές που σχετίζονται με την δραστηριότητα του εκπαιδευόμενου στην ιστοσελίδα του διαδικτυακού μαθήματος που μπορούν να επηρεαστούν από την μεθοδολογία της διδασκαλίας, το εκπαιδευτικό υλικό και τον τρόπο τοποθέτησης του εκπαιδευτικού υλικού στην στην διαδικτυακή πλατφόρμα μάθησης, διότι ουσιαστικά αυτά είναι τα στοιχεία ποιότητας του μαθήματος και κατ' επέκταση αυτά τα στοιχεία που θα κινητοποιήσουν τον εκπαιδευόμενο.

Τέλος ο Βαθμός της εργασίας θα μπορεί να χρησιμοποιηθεί ως δείκτης για την ανακεφαλαίωση και το υποστηρικτικό υλικό πριν τις τελικές εξετάσεις.

4.2 Προετοιμασία βάσης δεδομένων

Για την προετοιμασία των δεδομένων, την δημιουργία της βάσης δεδομένων και για να καταλήξουμε στην δημιουργία των μοντέλων εξόρυξης χρησιμοποιήθηκαν 3 διαφορετικά λογισμικά που συνεργάζονται μαζί. Αυτά είναι: το Microsoft Excel, το Microsoft SQL Server 2012 και το Microsoft Visual Studio.

Τα δεδομένα μας προέρχονται θεωρητικά από το LMS Moodle σε μορφή αρχείου Microsoft Excel. Έτσι τα δεδομένα αυτά το Microsoft Excel τα τροφοδοτεί στο Microsoft SQL Server και για την αρχιτεκτονική τους παρεμβαίνει το Microsoft Visual Studio.

Θεωρητικά σε αυτό το αρχείο του Excel που βγήκε από το Moodle αποτελείται από 16 στήλες και 207 σειρές. Από αυτές τις στήλες θα προστεθούν άλλες 2, η μια είναι με το πρωτεύον κλειδί.

Απο αυτές τις στήλες θα αφαιρεθεί η στήλη με το όνομα του χρήστη και θα μπει στην θέση αυτής της στήλης μια στήλη με αύξον αριθμό ως πρωτεύον κλειδί για να αποπροσωποποιηθούν τα δεδομένα και να προστατευτούν τα προσωπικά δεδομένα.

Ακόμα θα προστεθεί μια ακόμα στη που με την συνάρτηση `if` θα ελέγξει εάν ο βαθμός είναι άριστα ή όχι. Όταν θα είναι άριστα τότε θα βγει το γράμμα Ε (Επιτυχία), το Α (Αποτυχία).

Ύστερα θα πρέπει να γίνει ο έλεγχος για τυχόν αστοχίες, αυτές οι αστοχίες είναι:

- Κενά κελιά εγγράφων (Null cells)

Σε αυτό το θεωρητικό παράδειγμα δεν βρέθηκαν κενά κλειδιά, αν υπήρχαν αναλόγως την περίπτωση ή θα βάζουμε την τιμή 0 ή θα βάζουμε τον μέσο όρο των τιμών.

- Αφαίρεση όμοιων ή διπλότυπων στηλών

Στο παράδειγμά μας η άντληση των δεδομένων είναι από μια εξωτερική πηγή με αποτέλεσμα να μην γίνει αυτό το σφάλμα. Όμως πάντα πρέπει να ελεγχθεί και αυτή η περίπτωση.

- Στρογγυλοποίηση τιμών

Επειδή στα δεδομένα μας, υπάρχουν μέσοι όροι, πιθανόν να υπάρχουν τιμές στα κελιά του Excel με αρκετά δεκαδικά. Γιαυτό θα γίνονταν στρογγυλοποίηση μέσω αντίστοιχων queries στον SQL Server για να μην υπάρχουν προβλήματα στα πλαίσια της πολυδιάστατης ανάλυσης τους.

- Μορφοποίηση

Οι ονομασίες των μεταβλητών θα πρέπει να είναι τέτοιες ώστε να αναγνωρίζονται από το σύστημα αλλά και να διευκολύνεται η διαδικασία της ανάλυσης.

4.3 Εισαγωγή Δεδομένων (Import Data)

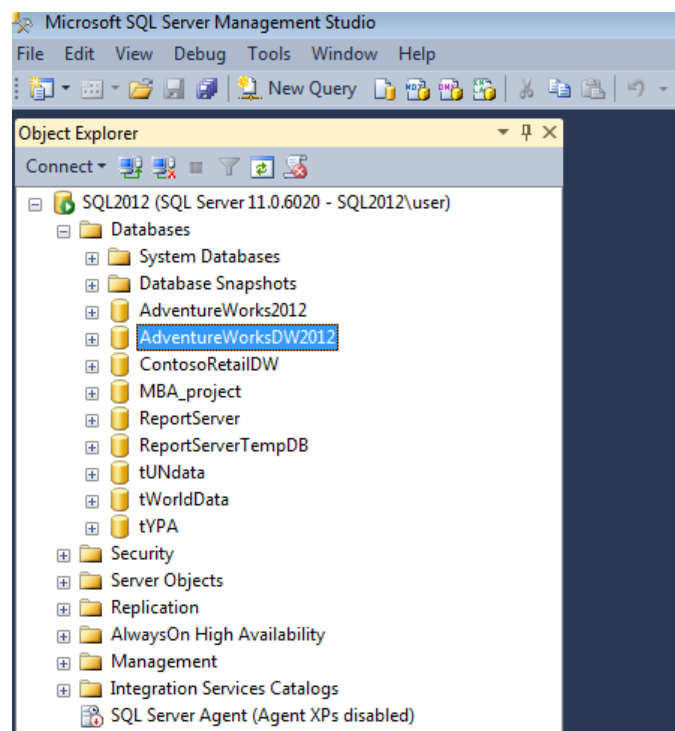
Θα ξεκινήσουμε εισάγοντας τα δεδομένα στην βάση στον SQL Server τα δεδομένα ακολουθώντας τα παρακάτω βήματα.

Ανοίγουμε το Management Studio του sql Server για να πάμε στις βάσεις δεδομένων και συνδεόμαστε με τον server.



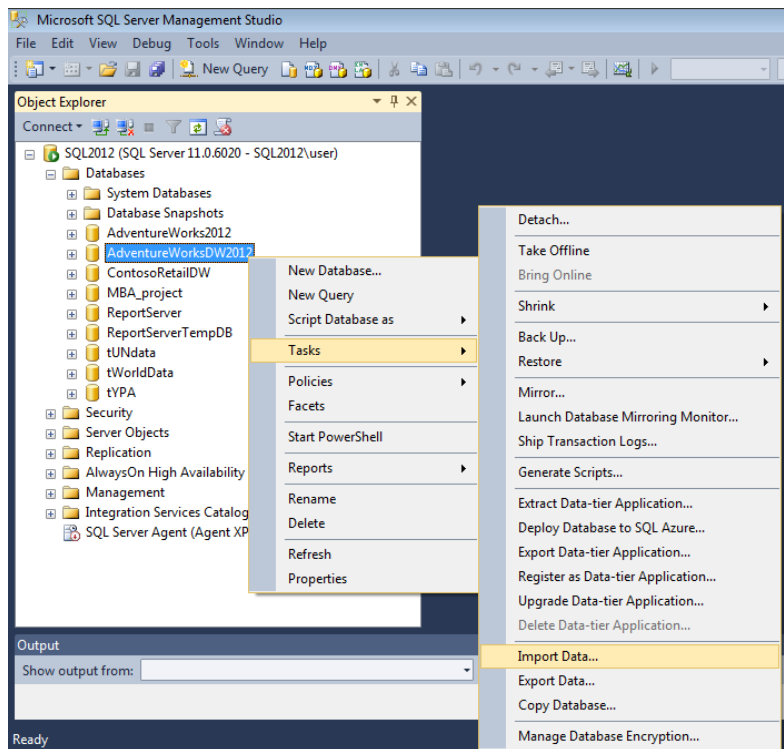
Εικόνα 9: SQL Server

Πατάμε Connect.



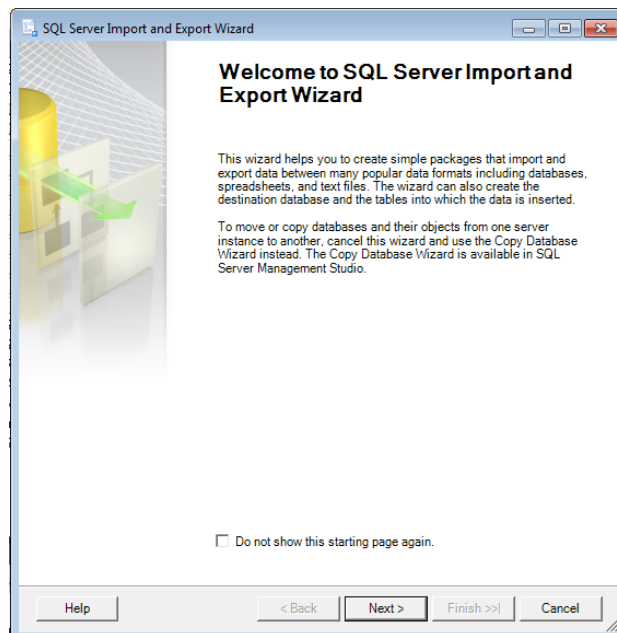
Εικόνα 10: Επιλογή Databases

Επιλέγουμε την βάση που θέλουμε και κάνουμε δεξί κλικ.



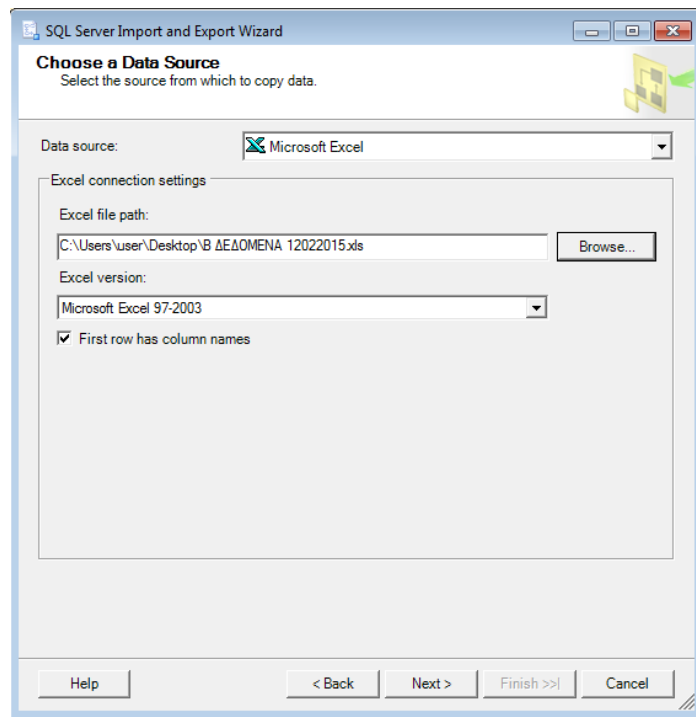
Εικόνα 11: Εισαγωγή Δεδομένων

Μετά Task και μετά Import Data.



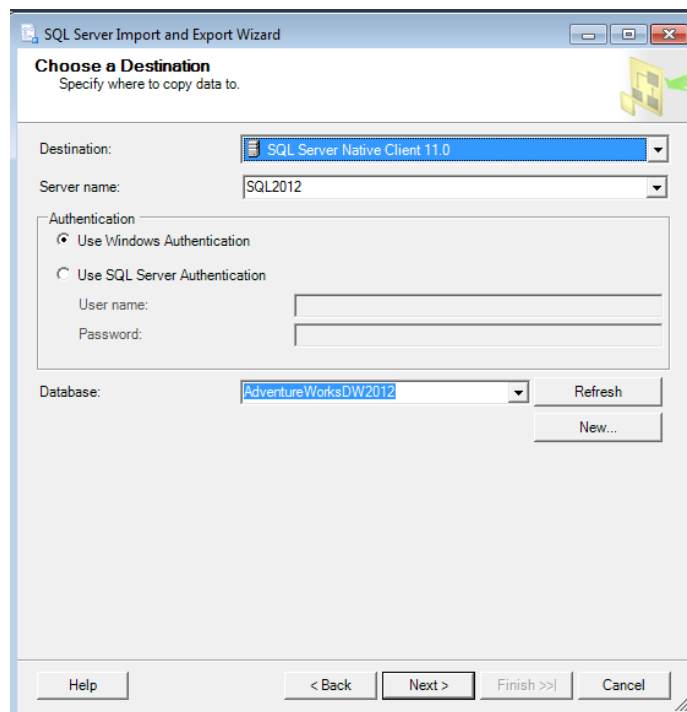
Εικόνα 12: Import Wizard

Στο Wizard πατάμε Next.



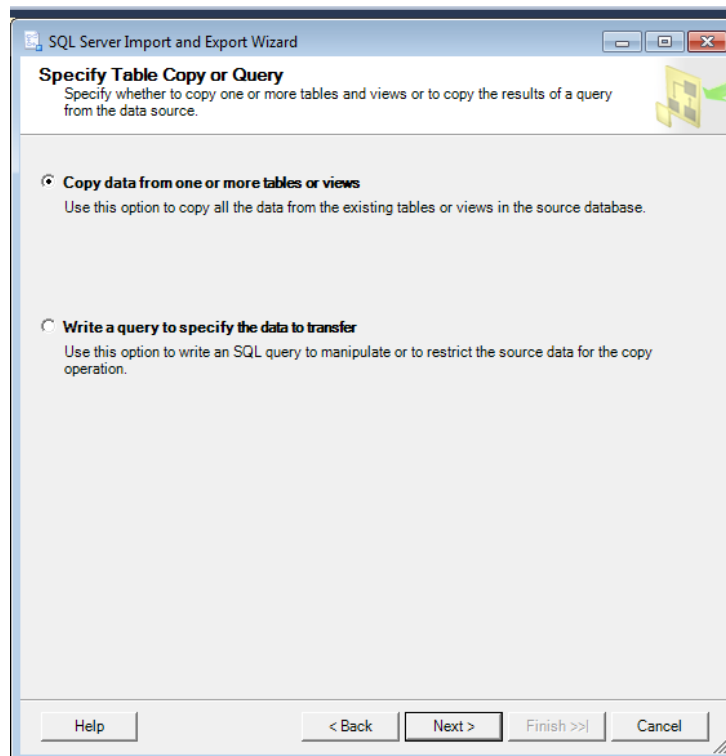
Εικόνα 13: Επιλογή των Δεδομένων

Επιλέγουμε την επιλογή Excel γιατί τα δεδομένα μας είναι σε μορφή Excel. Ύστερα ψάχνουμε με το browser που ακριβώς είναι το αρχείο του Excel και το επιλέγουμε. Επιλέγουμε το First row has column names, δηλαδή ότι στην πρώτη γραμμή στο Excel είναι οι ονομασίες των μεταβλητών. Πατάμε Next.



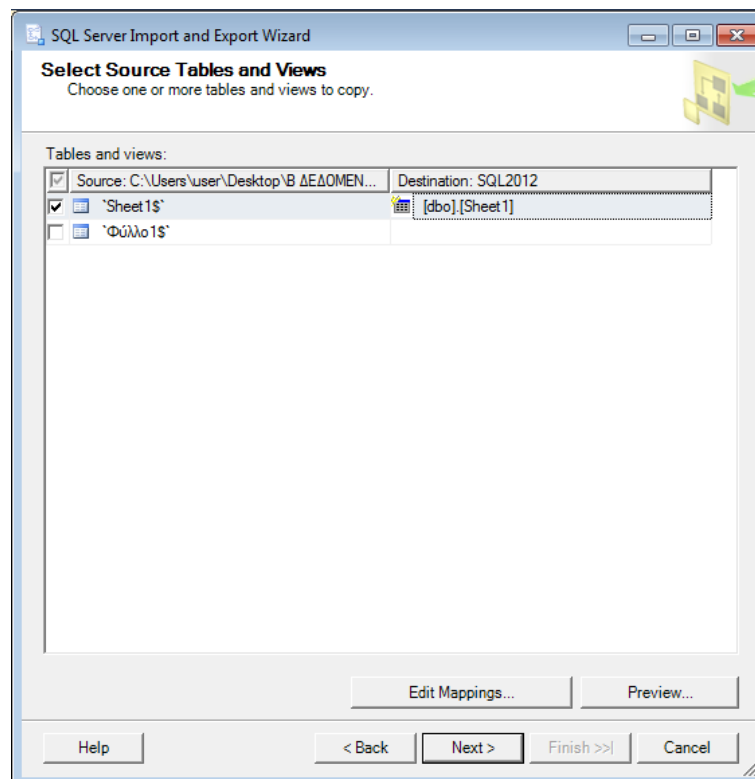
Εικόνα 14: Επιλογή προορισμού

Επιλέγουμε τον προορισμό και τη βάση που θα δημιουργήσουμε τον πίνακα και πατάμε Next.



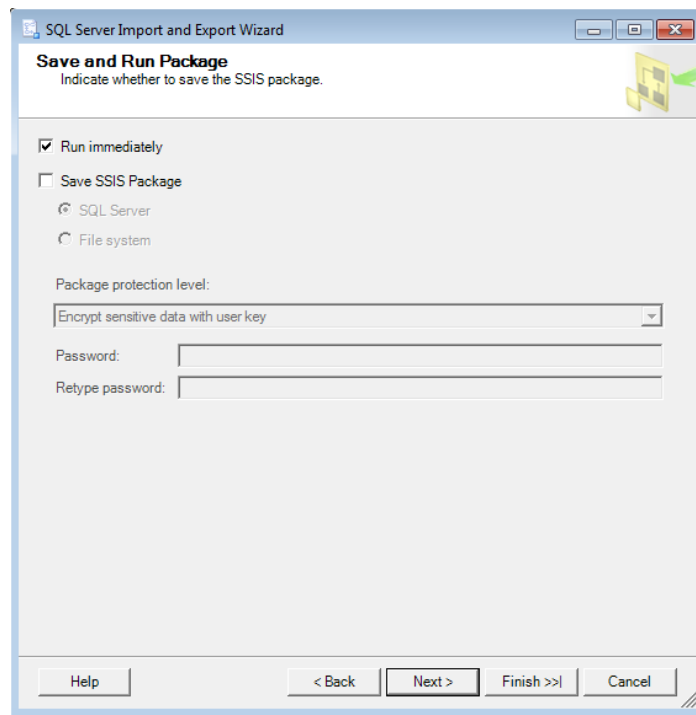
Εικόνα 15: Αντιγραφή Πίνακα

Επιλέγουμε Copy data from one or more tables or views και πατάμε Next.



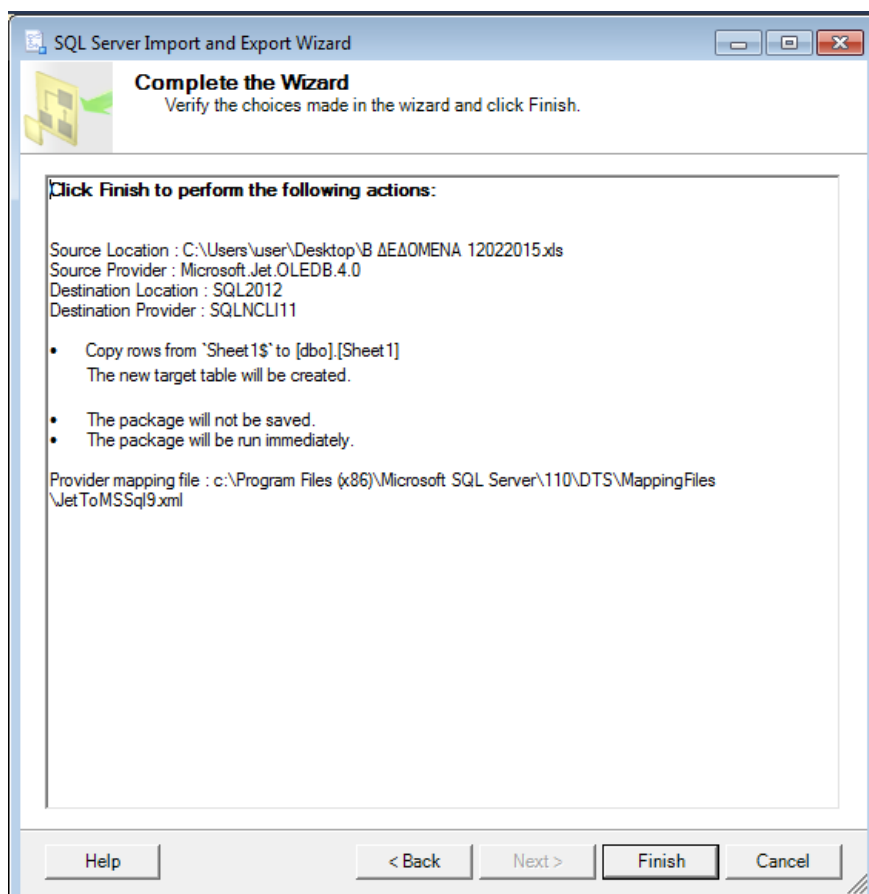
Εικόνα 16: Επιλογή Φύλλου Δεδομένων

Επιλέγουμε το φύλλο του Excel που είναι τα δεδομένα μας και πατάμε Next.



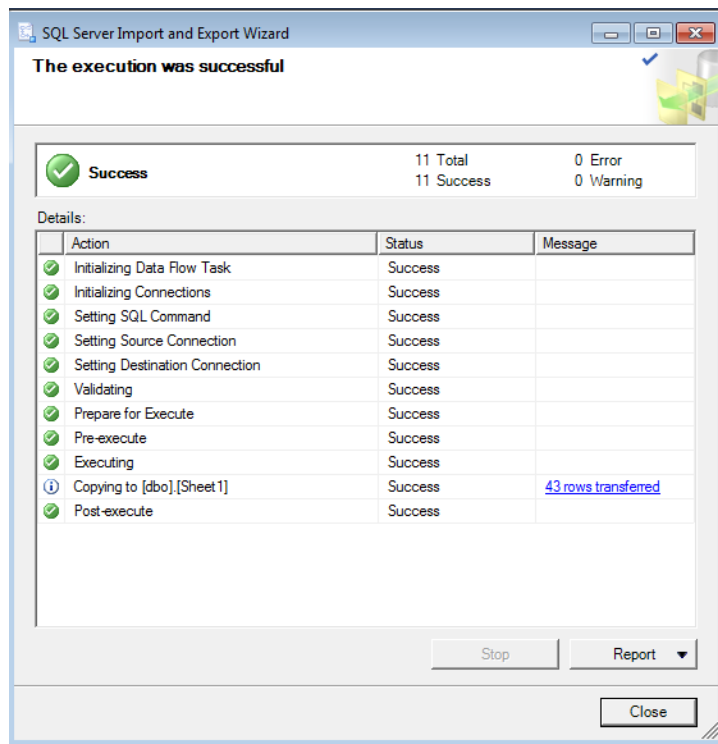
Εικόνα 17: Επιλογή Run Immediately

Επιλέγουμε Run immediately και πατάμε Next.



Εικόνα 18: Complete the Wizard

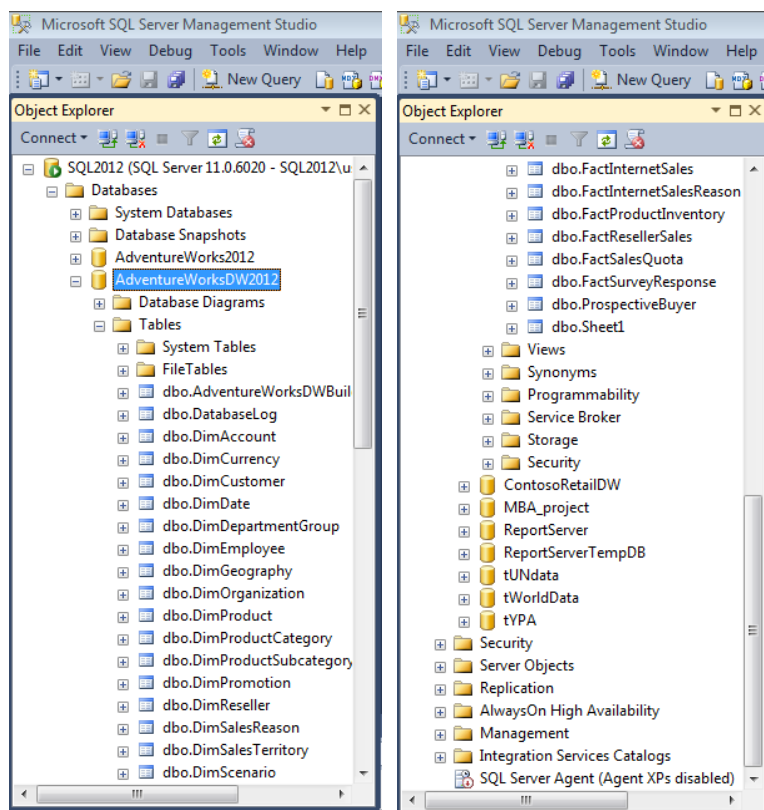
Πατάμε Finish.



Εικόνα 19: Ενημέρωση Ολοκλήρωσης

Πατάμε Close.

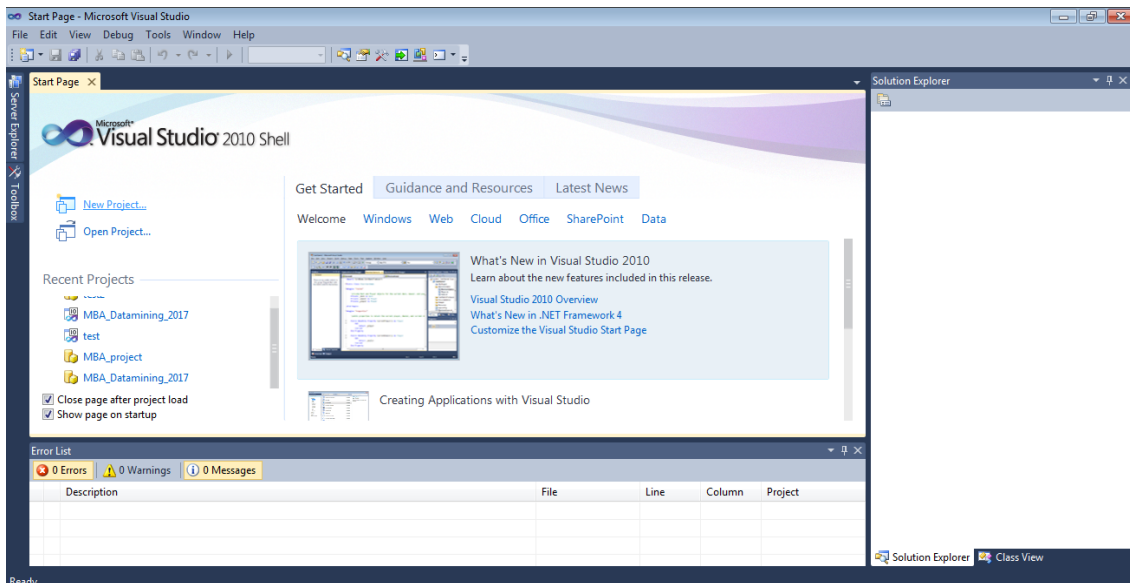
Πλέον έχουμε δημιουργήσει ένα πίνακα με ονομασία dbo.Sheet1



Εικόνα 20: Θέση Πίνακα Sheet1

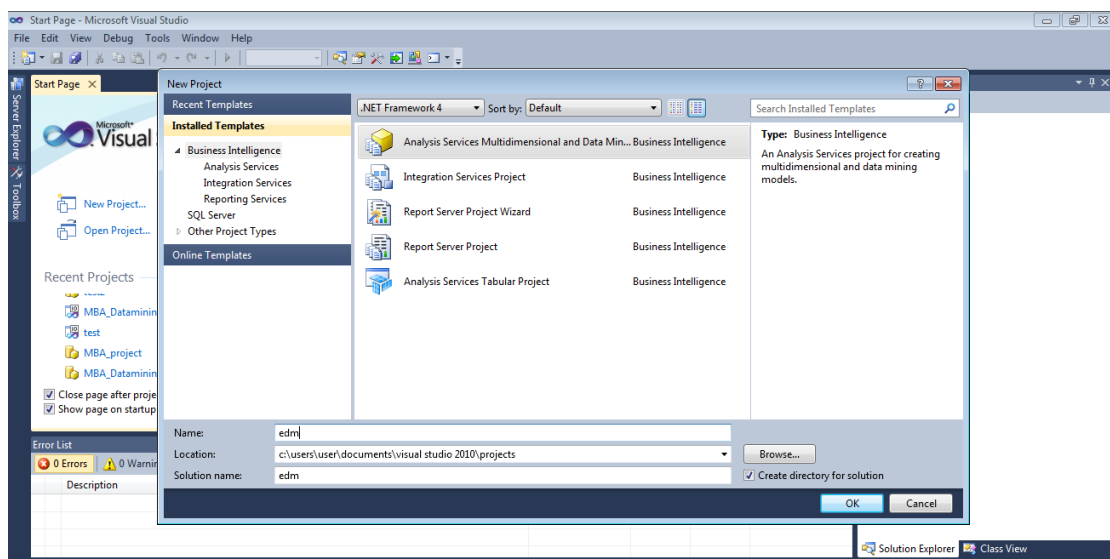
4.4 Χτίσιμο του Κύβου (Build the Cube)

Για το χτίσιμο του κύβου θα ανοίξουμε το Microsoft Visual Studio και θα ακολουθήσουμε τα παρακάτω βήματα.



Εικόνα 21: Επιλογή New Project

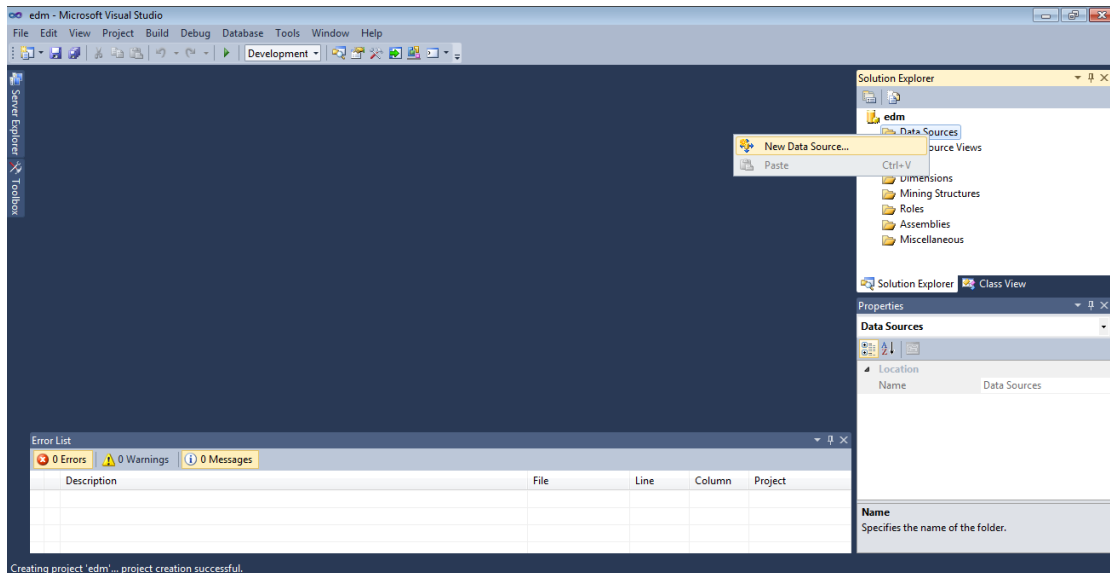
Επιλέγουμε New Project.



Εικόνα 22: Θέση New Project

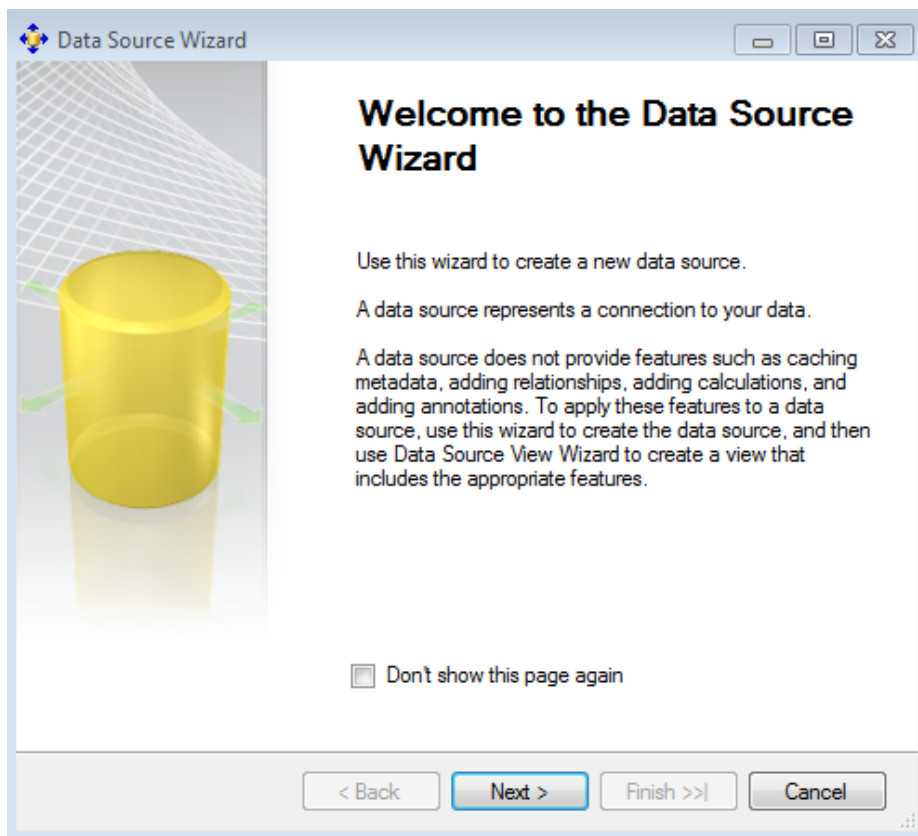
Επιλέγουμε Business Intelligence μετά Analysis Services Multidimensional..

Γράφουμε στο Name το όνομα και πατάμε OK.



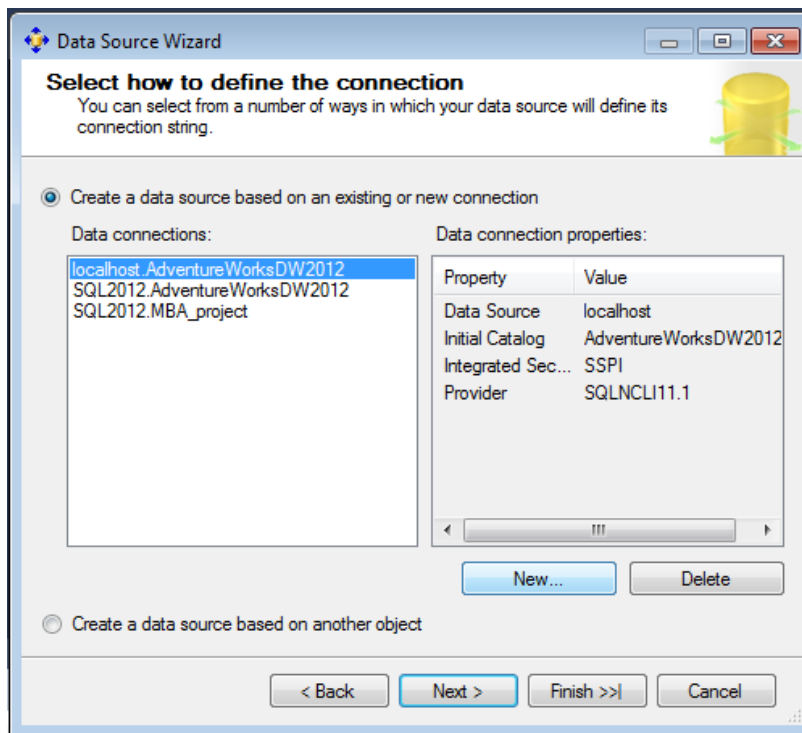
Εικόνα 23: Δημιουργία Data Source

Δημιουργούμε ένα Data Source, Δεξί κλικ στο Data Sources και επιλέγουμε New Data Source.



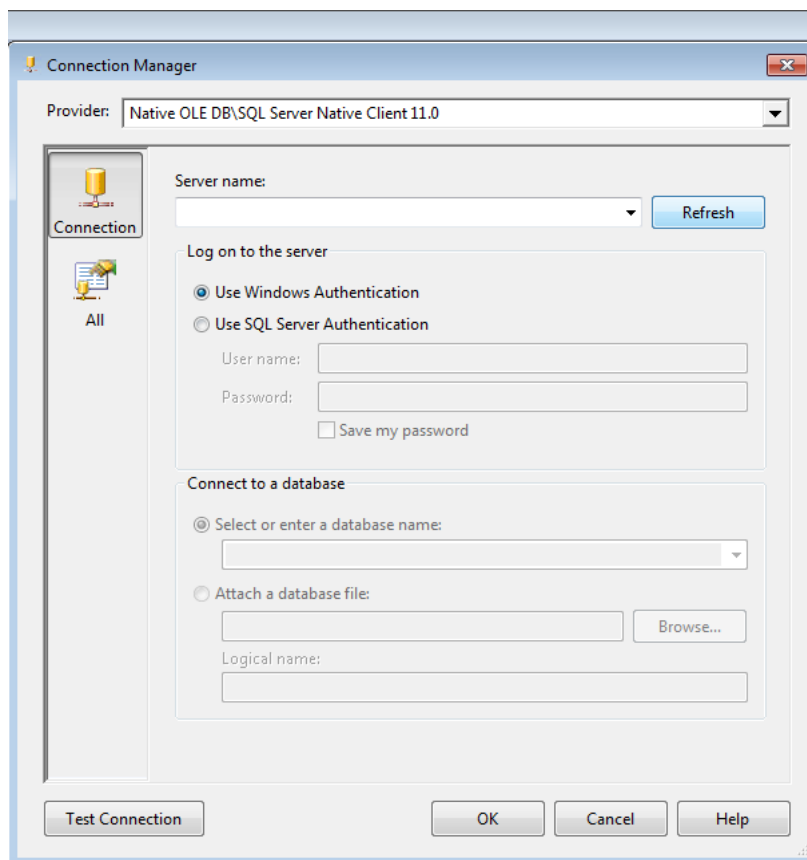
Εικόνα 24: Data Source Wizard

Πατάμε Next στον Wizard



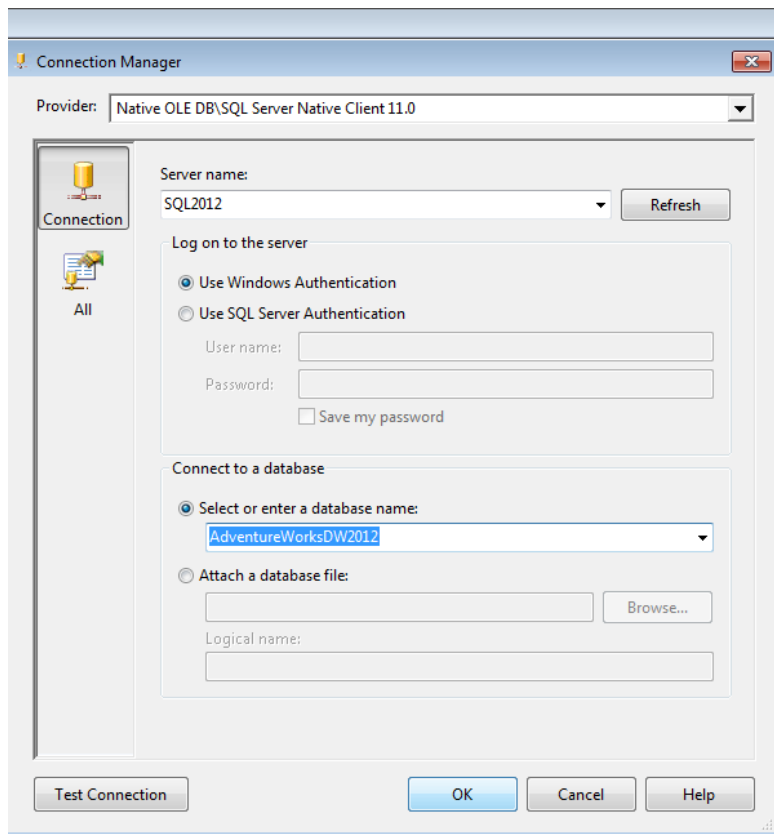
Εικόνα 25:How to define the connection

Πατάμε New



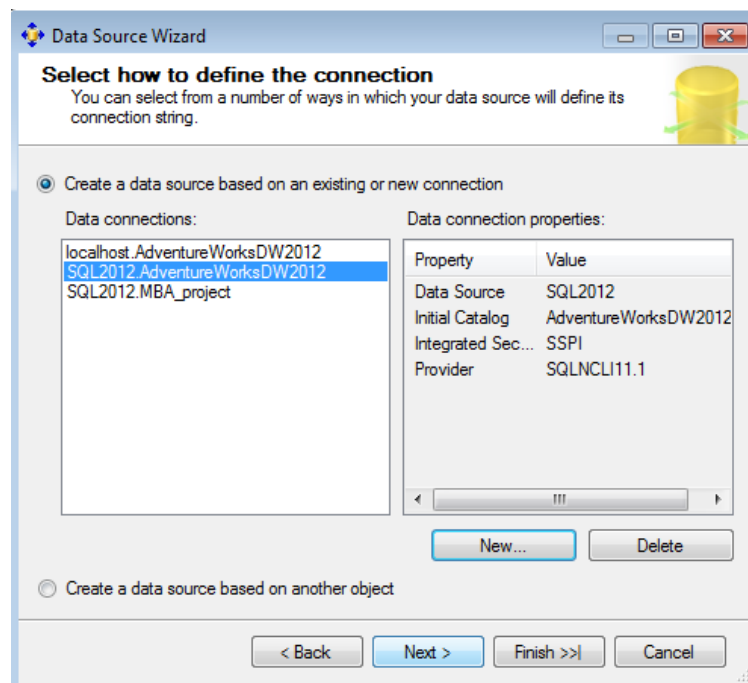
Εικόνα 26: Επιλογή Server

Βρίσκουμε τον Server



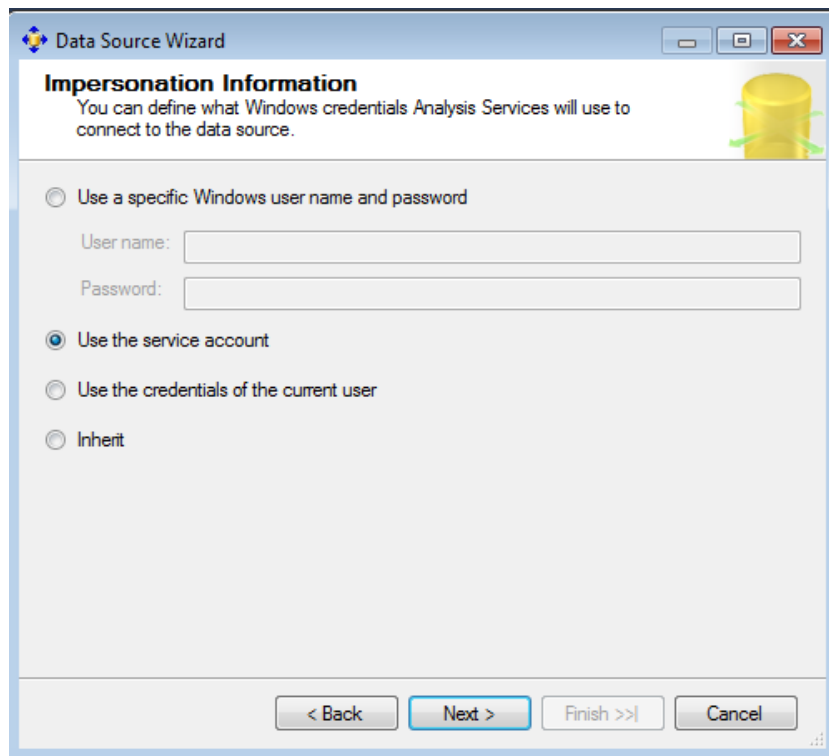
Εικόνα 27: Επιλογή Database

Βρίσκουμε βάση μας και OK



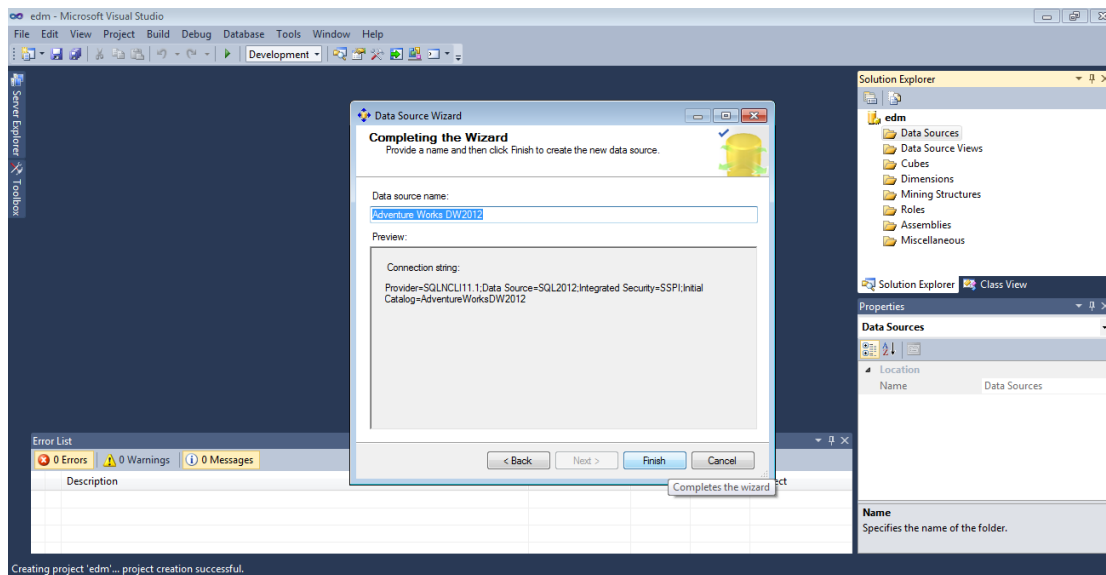
Εικόνα 28:How to define the connection

Πατάμε Next



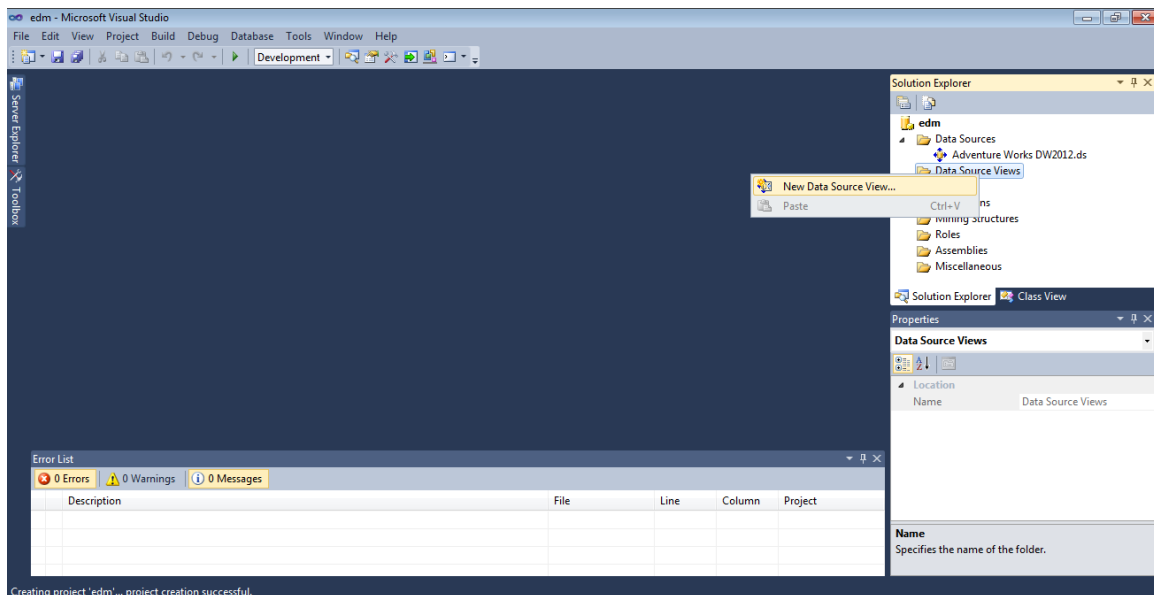
Εικόνα 29: Impersonation Information

Επιλέγουμε Use the service account και μετά Next.



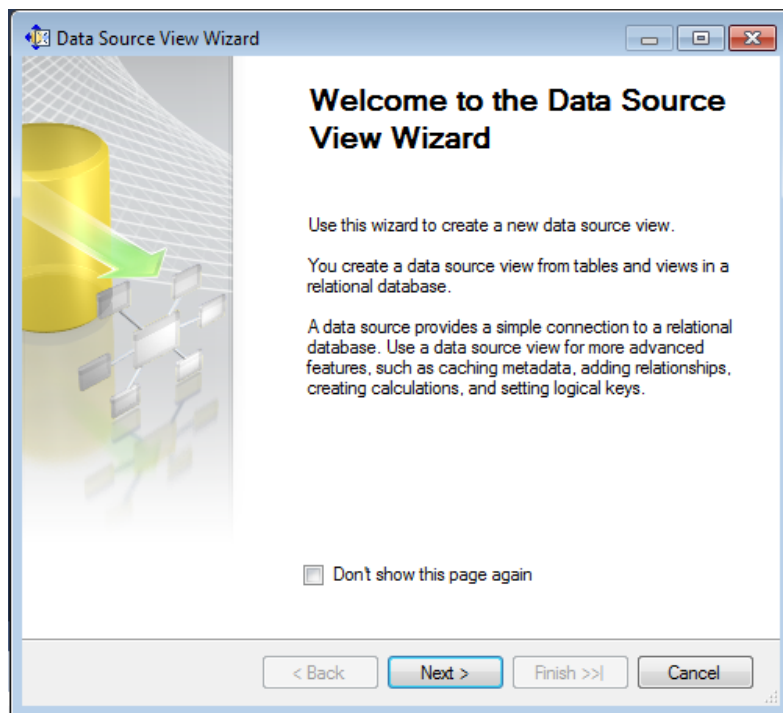
Εικόνα 30: Ολοκλήρωση Data Source

Μετά Finish.



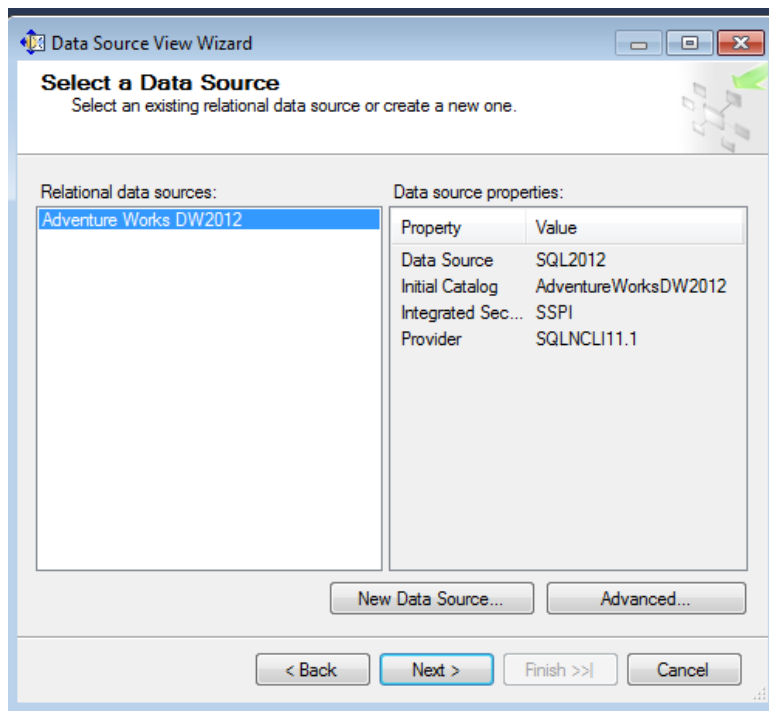
Εικόνα 31: Δημιουργία Data Source View

Δημιουργούμε ένα Data View, Δεξί κλικ στο Data Views και επιλέγουμε New Data View.



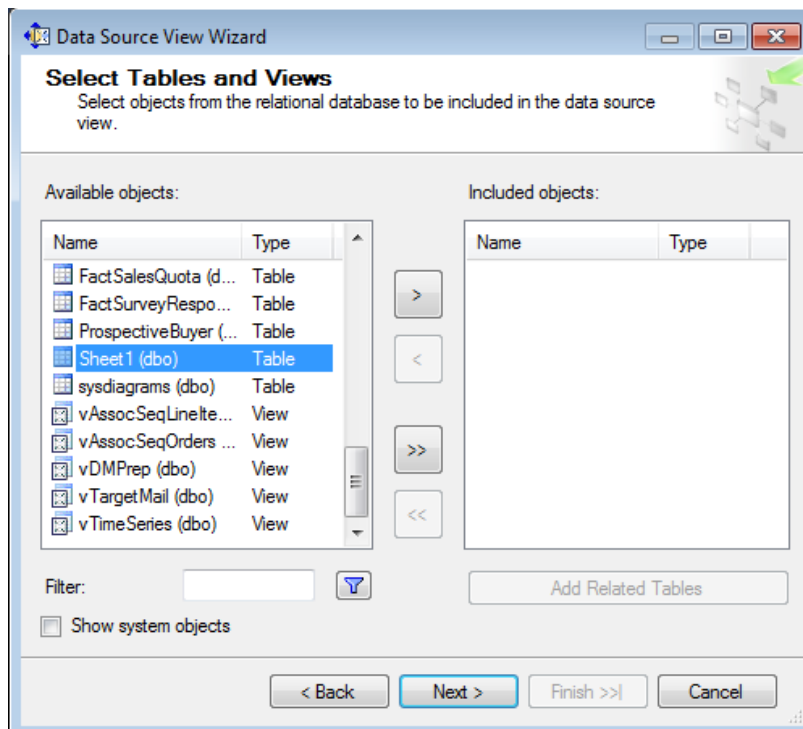
Εικόνα 32: Data Source View Wizard

Πατάμε Next στον Wizard.



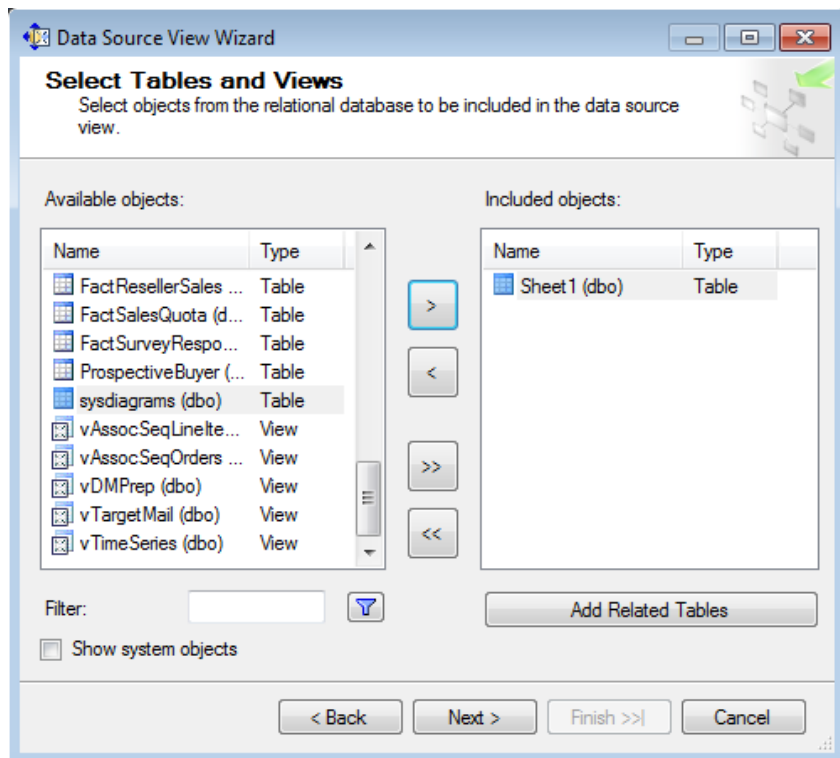
Εικόνα 33: Select a Data Source

Πατάμε Next.



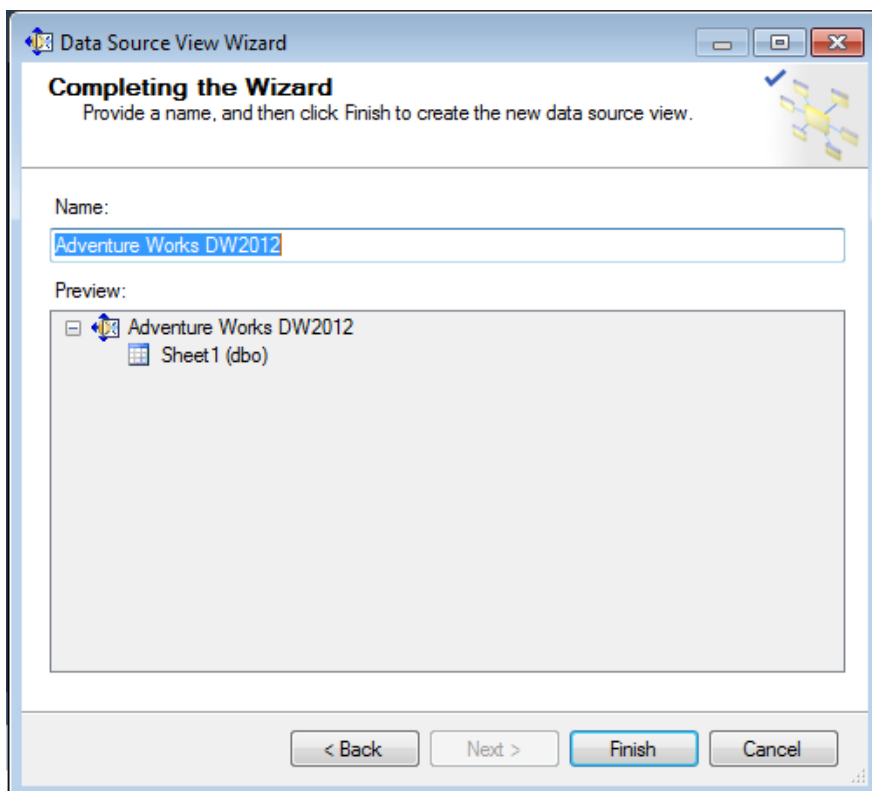
Εικόνα 34: Select Tables and Views

Βρίσκουμε τον πίνακα (Shet1.dbo) που φτιάξαμε και τον επιλέγουμε.



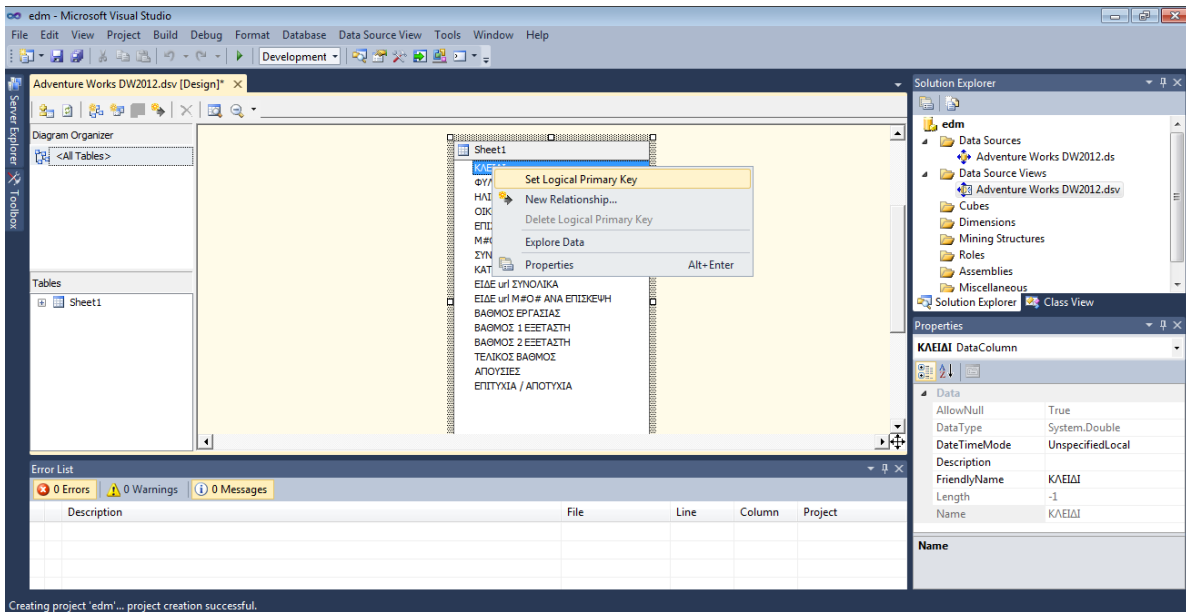
Εικόνα 35: Επιλογή Sheet1

Ύστερα τον μεταφέρουμε στα Included objects και μετά Next.



Εικόνα 36: Ολοκλήρωση Data Source View

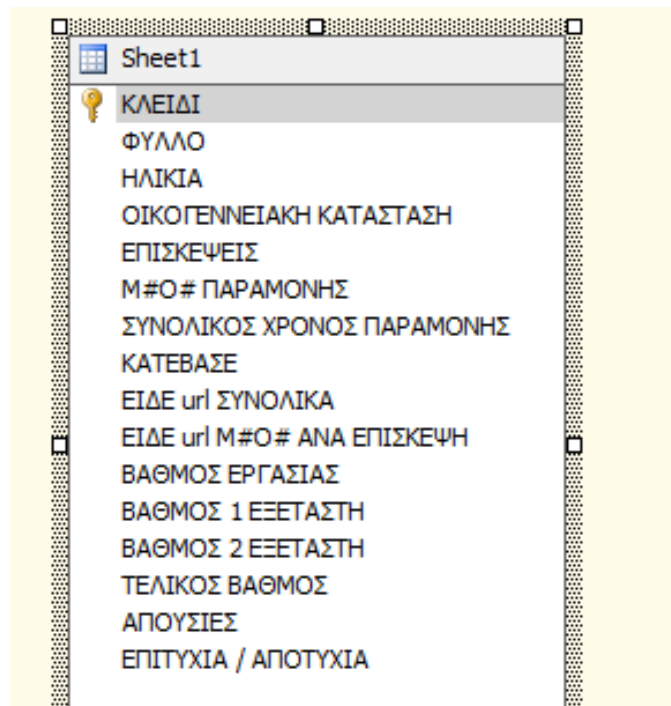
Πατάμε Finish



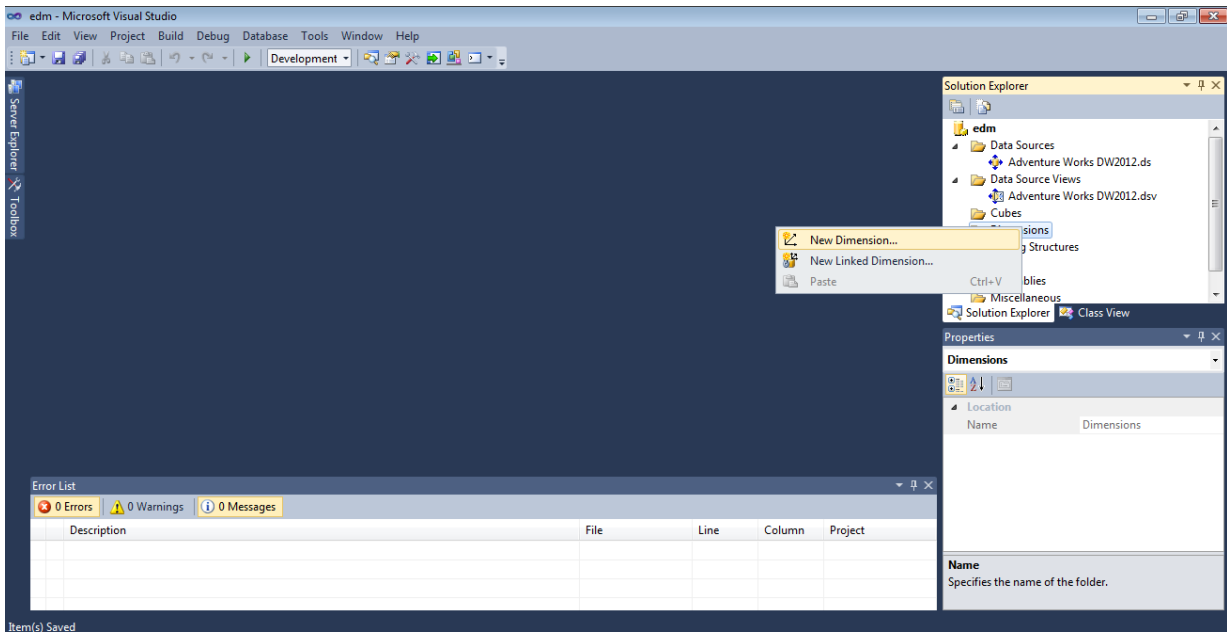
Εικόνα 37: Δημιουργία Πρωτεύον Κλειδί

Πατάμε στο ΚΛΕΙΔΙ και του κάνουμε Δεξί κλικ

Επιλέγουμε Set Logical Primary Key και έτσι έχουμε δημιουργήσει το πρωτεύον κλειδί.

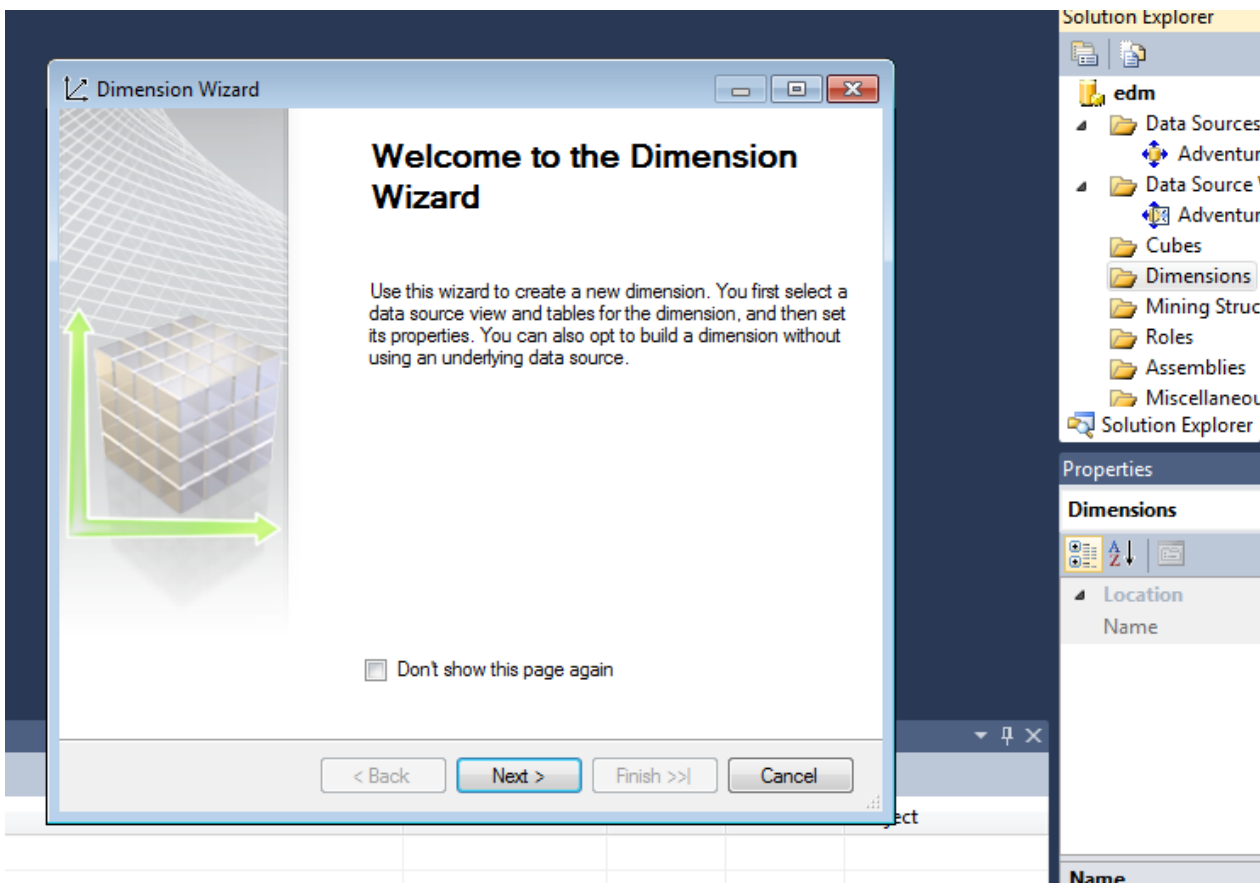


Εικόνα 38: Πρωτεύον Κλειδί



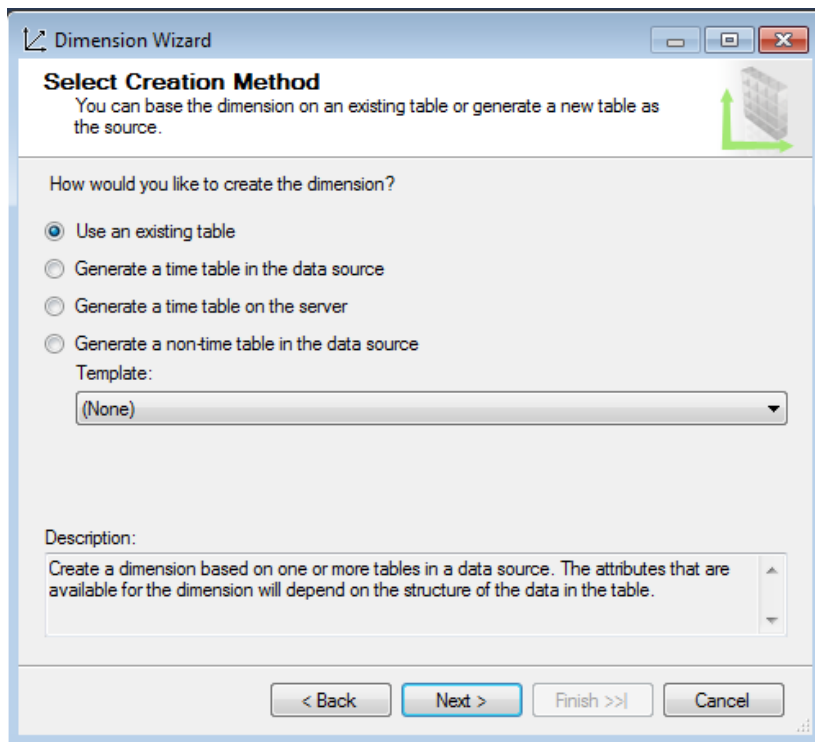
Εικόνα 39: Δημιουργία New Dimension

Δημιουργούμε ένα Dimension, Δεξί κλικ στο Dimensions και επιλέγουμε New Dimension.



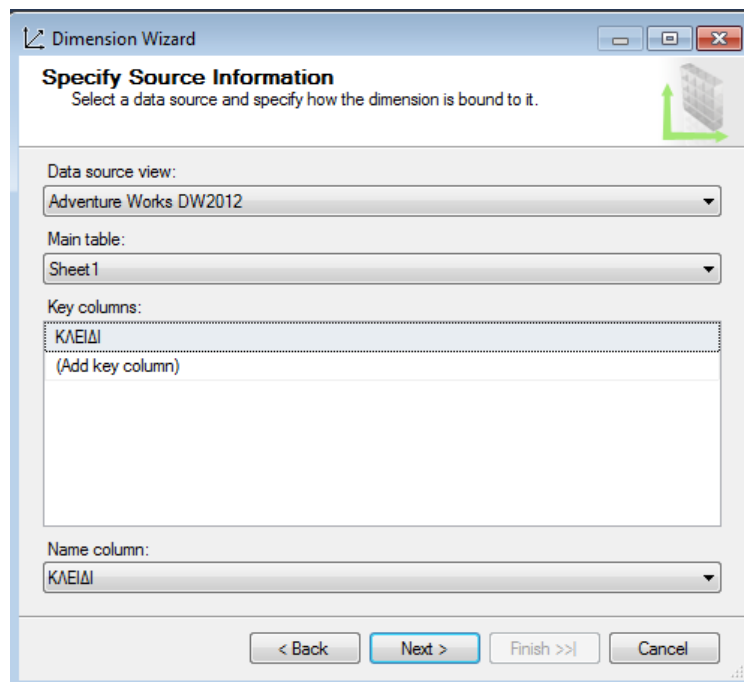
Εικόνα 40: Dimension Wizard

Πατάμε Next στον Wizard



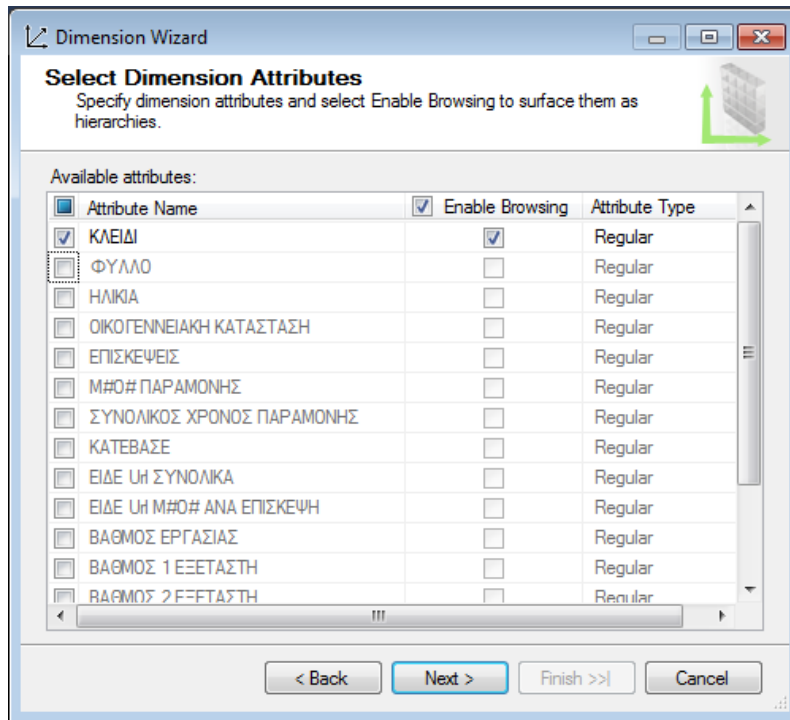
Εικόνα 41:Select Creation Method

Επιλέγουμε Use an existing table και πατάμε Next.



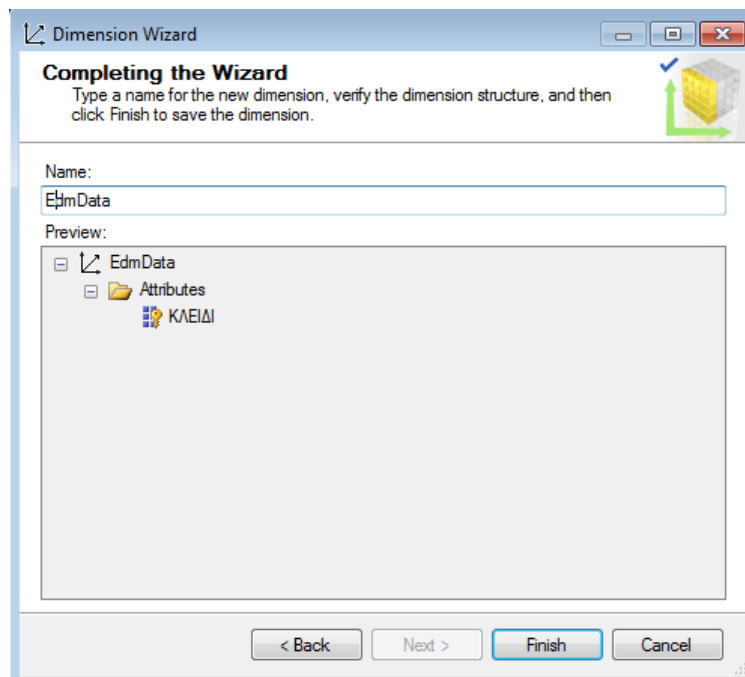
Εικόνα 42:Specify Source Information

Πατάμε Next



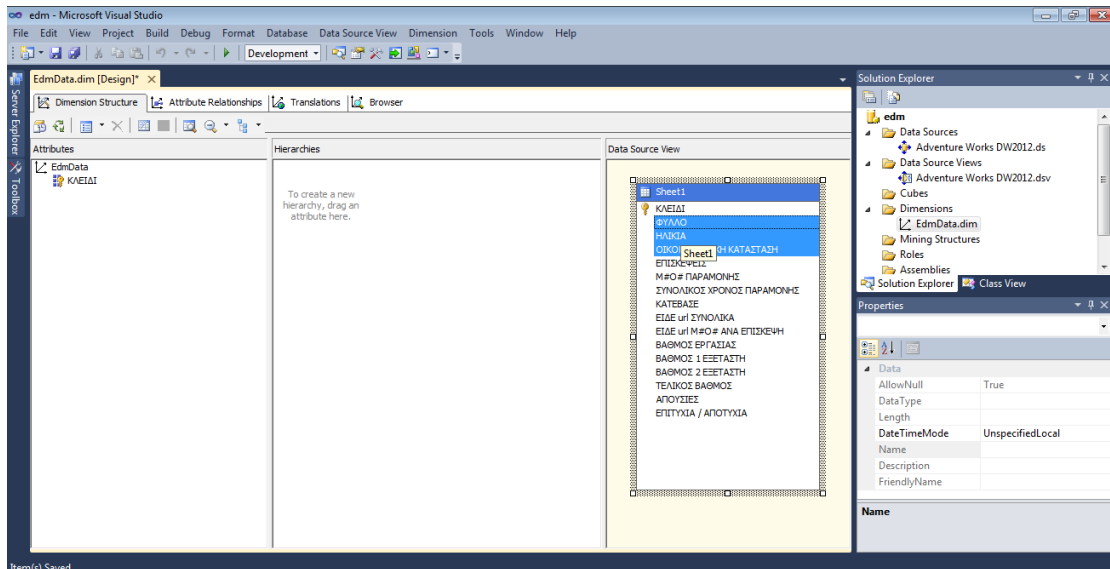
Εικόνα 43:Select Dimension Attributes

Επιλέγουμε το κλειδί και πατάμε Next.



Εικόνα 44: Ολοκλήρωση Dimension

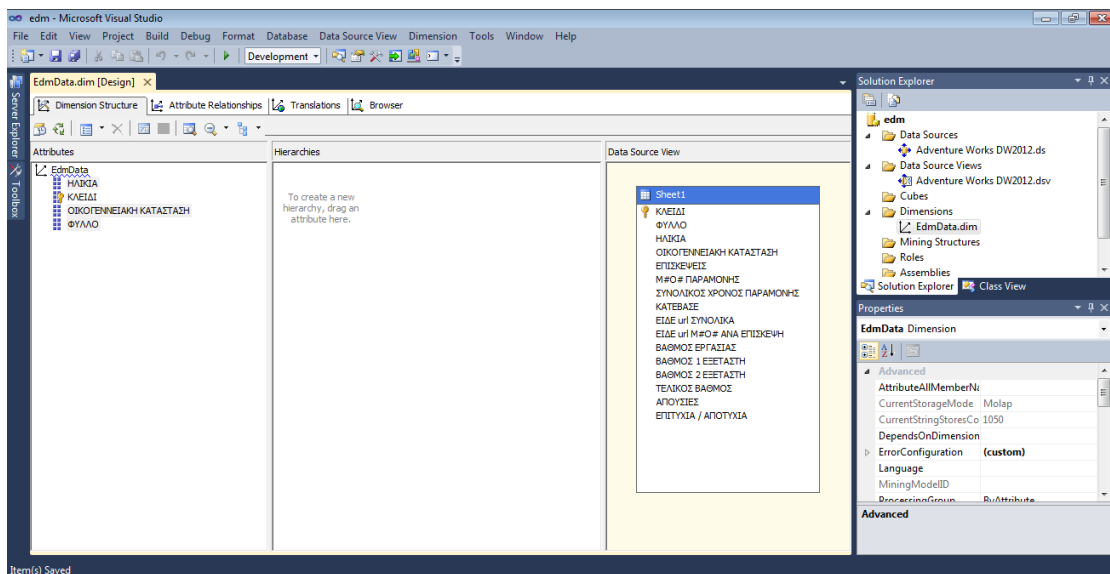
Γράφουμε στο Name StudentInfo και μετά Finish.



Εικόνα 45: Θέση Dimension

Πλέον έχει δημιουργηθεί μια νέα Dimension με όνομα StudentInfo.dim

Επιλέγουμε τα στοιχεία που αφορούν τα χαρακτηριστικά των εκπαιδευόμενων στον πίνακα Sheet1

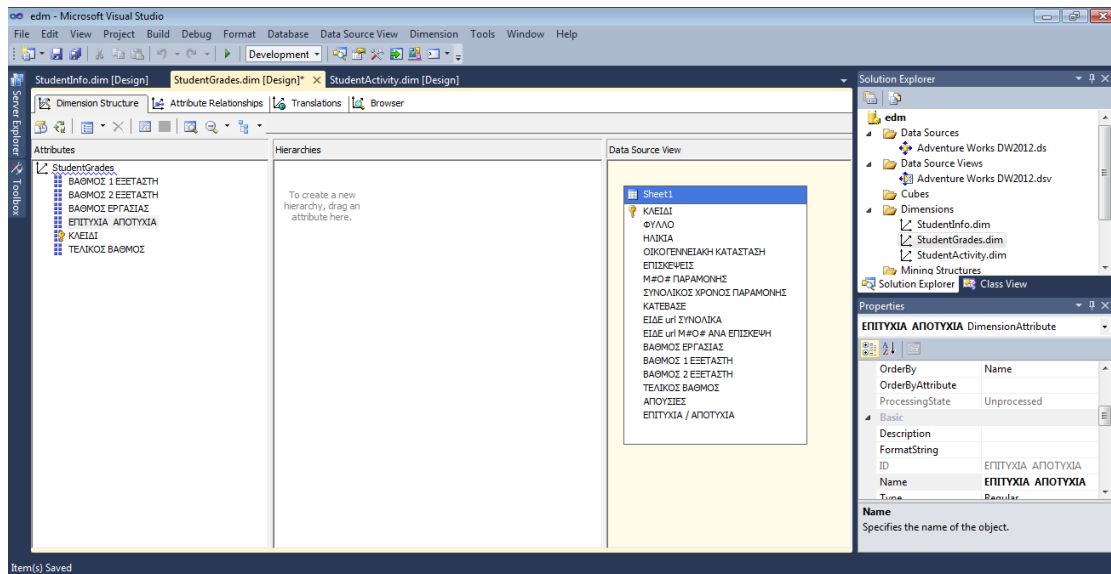


Εικόνα 46:Επιλογή στοιχείων πίνακα Sheet1

Τα επιλεγμένα στοιχεία τα βάζουμε στο Attributes και έτσι έχουμε δημιουργήσει μια διάσταση με τα χαρακτηριστικά των εκπαιδευόμενων (ΦΥΛΟ, ΗΛΙΚΙΑ, ΟΙΚΟΓΕΝΕΙΑΚΗ ΚΑΤΑΣΤΑΣΗ).

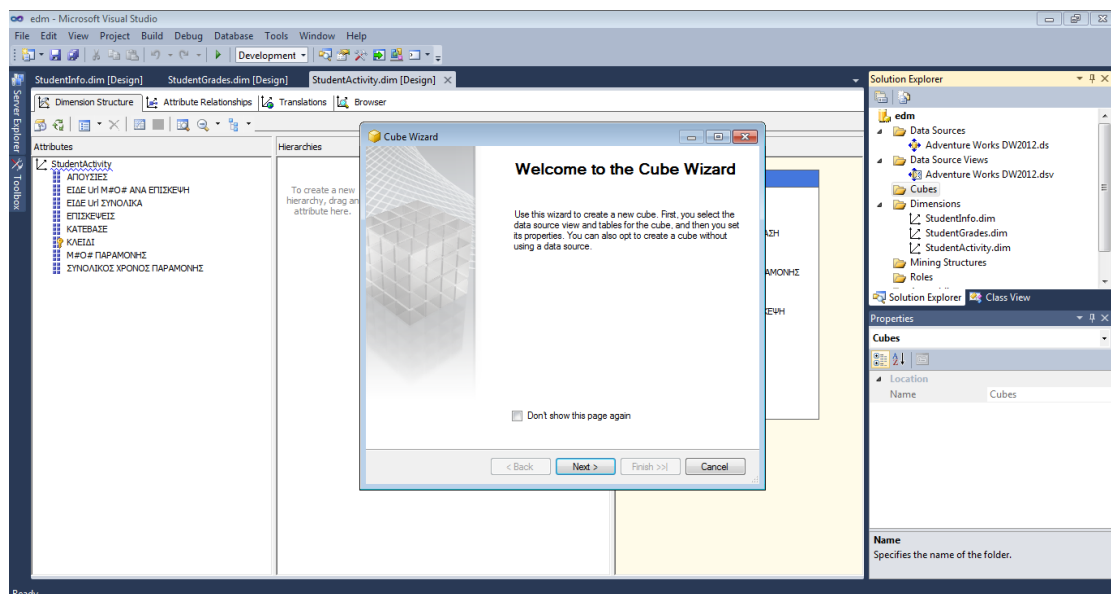
Την ίδια διαδικασία θα την ακολουθήσουμε άλλες τρεις φορές και θα δημιουργήσουμε άλλες δυο διαστάσεις. Η μια διάσταση θα έχει την ονομασία StudentGrandes.dim και θα έχει τις επιδόσεις των εκπαιδευόμενων (ΒΑΘΜΟΣ ΕΡΓΑΣΙΑΣ, ΒΑΘΜΟΣ 1 ΕΞΕΤΑΣΤΗ, ΒΑΘΜΟΣ 2 ΕΞΕΤΑΣΤΗ, ΤΕΛΙΚΟΣ ΒΑΘΜΟΣ). Και η άλλη διάσταση θα έχει την ονομασία StudentActivity.dim και θα έχει τις δραστηριότητες των εκπαιδευόμενων (ΕΠΙΣΚΕΨΕΙΣ , Μ.Ο. ΠΑΡΑΜΟΝΗΣ, ΣΥΝΟΛΙΚΟΣ ΧΡΟΝΟΣ ΠΑΡΑΜΟΝΗΣ, ΚΑΤΕΒΑΣΕ, ΕΙΔΕ url ΣΥΝΟΛΙΚΑ, ΕΙΔΕ url Μ.Ο. ΑΝΑ ΕΠΙΣΚΕΨΗ, ΑΠΟΥΣΙΕΣ).

Η τελευταία διάσταση θα έχει την ονομασία StudentSuccess.dim και θα έχει την Επιτυχία των εκπαιδευόμενων(ΕΠΙΤΥΧΙΑ / ΑΠΟΤΥΧΙΑ)



Εικόνα 47: Διαστάσεις Dimension

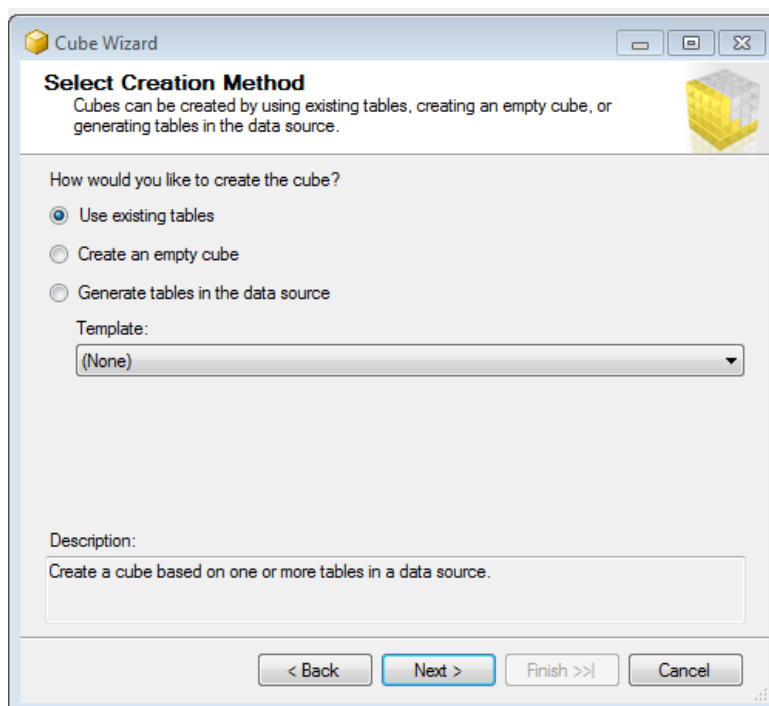
Στο Dimensions βλέπουμε τις τρεις νέες διαστάσεις.



Εικόνα 48: Δημιουργία Cube

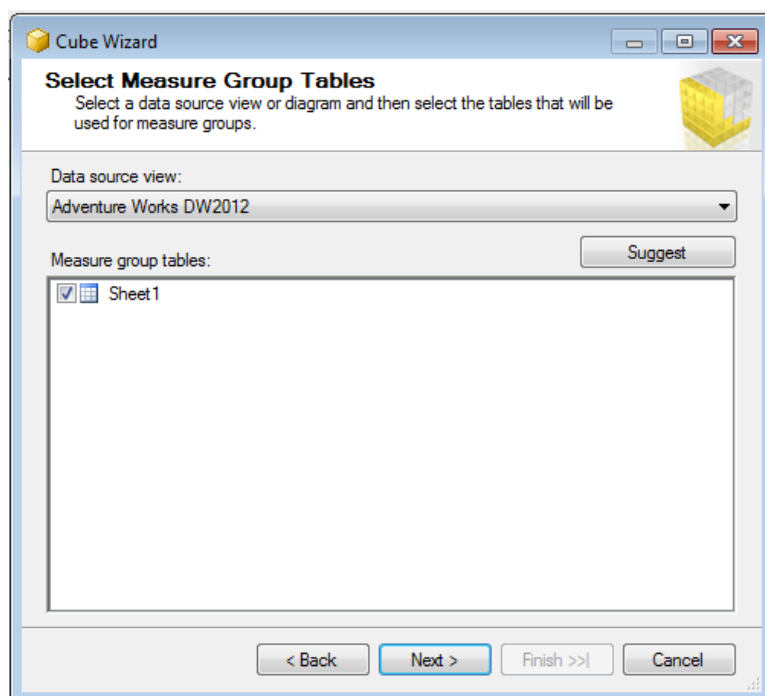
Το επόμενο βήμα είναι να δημιουργήσουμε έναν κύβο. Πατάμε Δεξί κλικ στο Cube και επιλέγουμε New Cube.

Πατάμε Next στον Wizard.



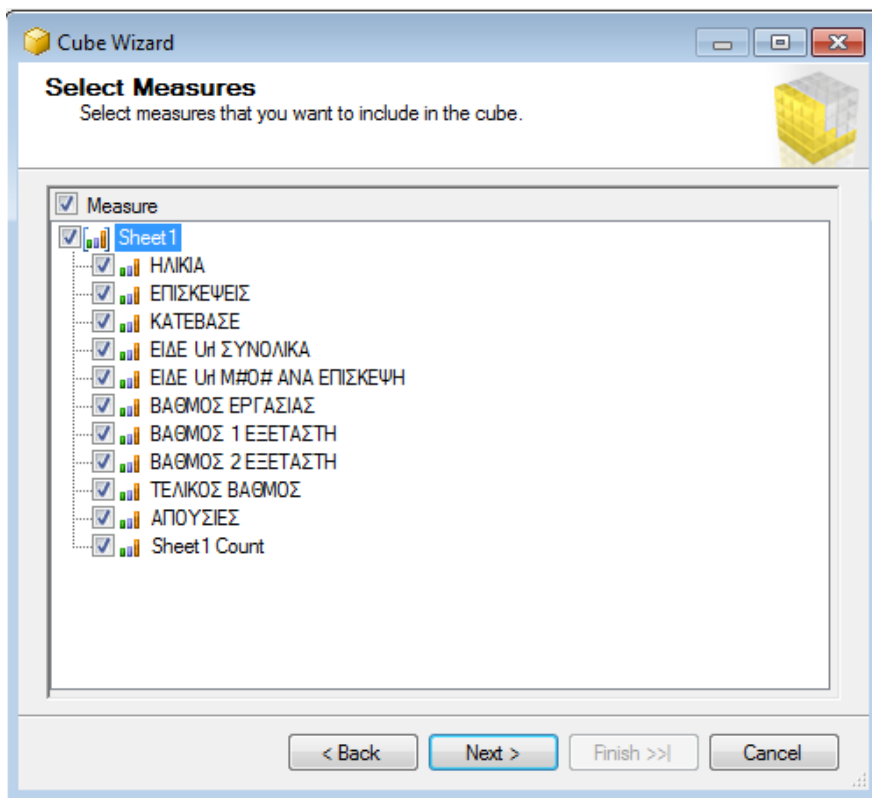
Εικόνα 49: Select Creation Method

Επιλέγουμε Use existing tables και μετά Next.



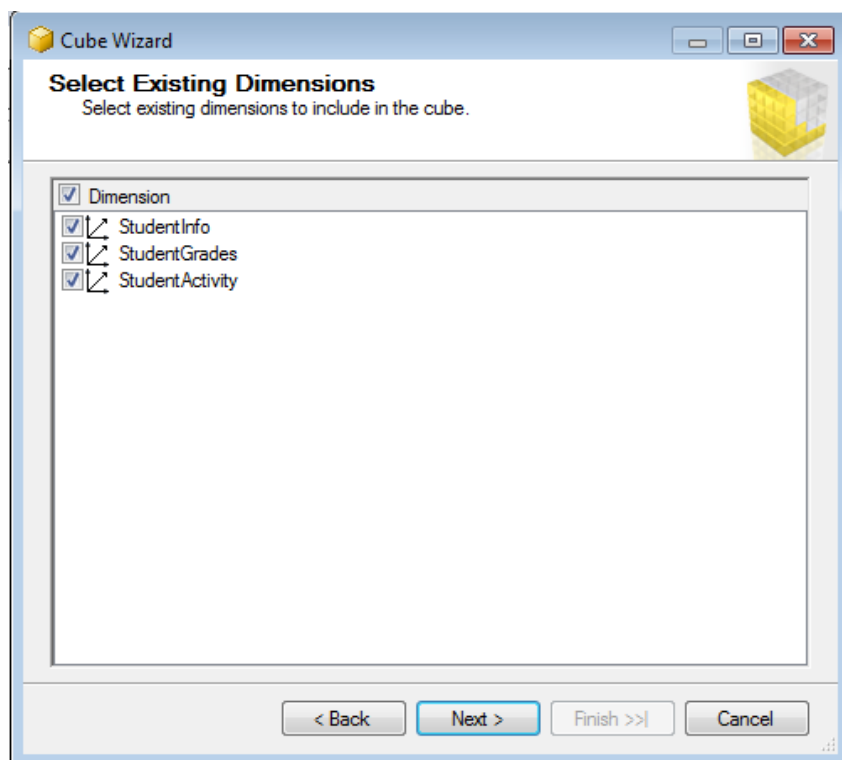
Εικόνα 50: Select Measure Group Tables

Επιλέγουμε το Sheet1 και μετά Next.



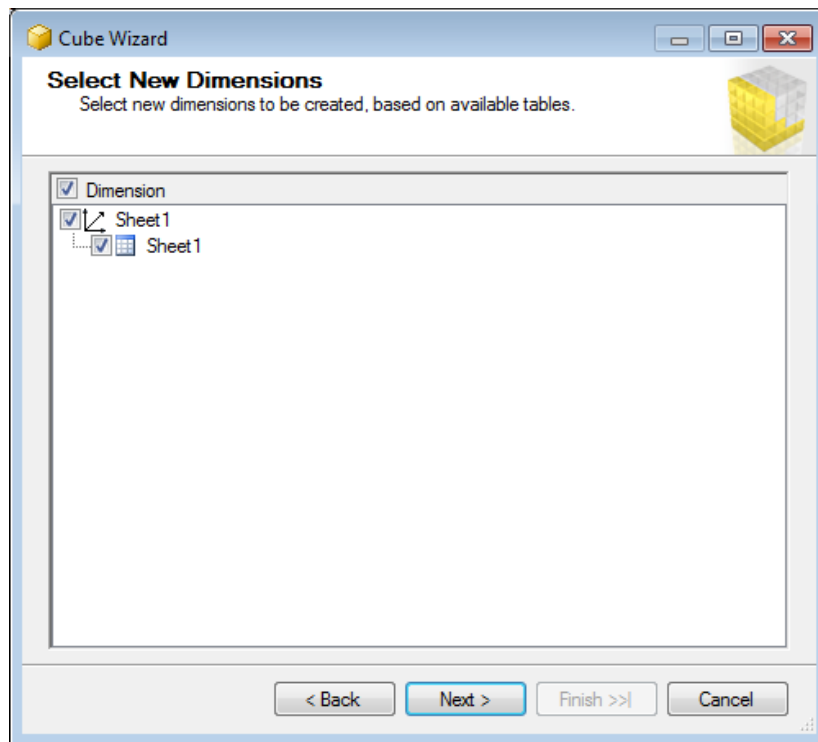
Εικόνα 51: Select Measures

Επιλέγουμε όλα τα Measure και Next.



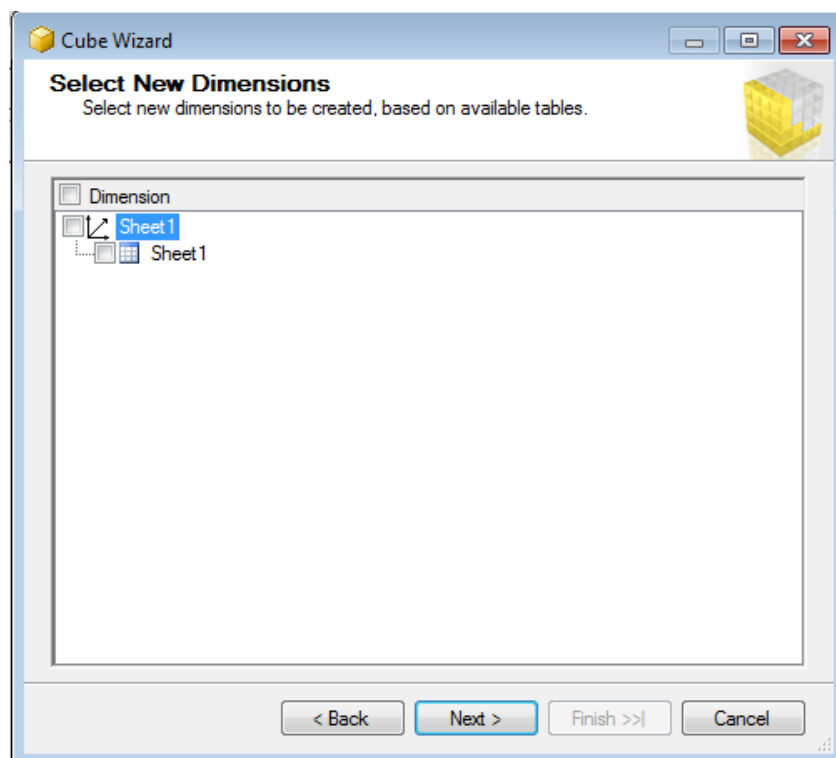
Εικόνα 52: Select Existing Dimensions

Επιλέγουμε όλα τις Dimensions και Next.



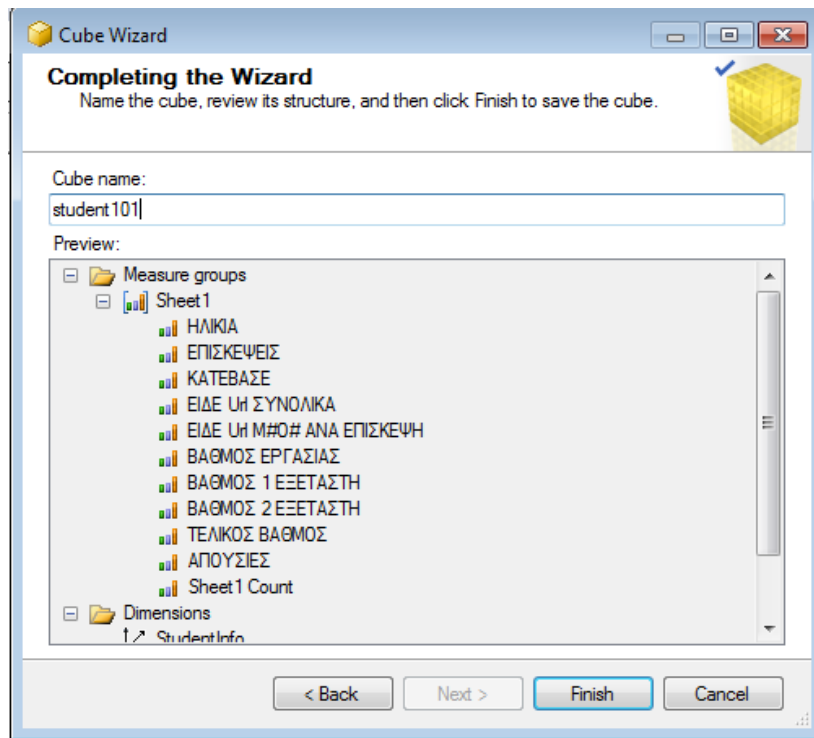
Εικόνα 53: Select New Dimensions

Μας προτίνει να δημιουργήσει μια νέα Dimension.



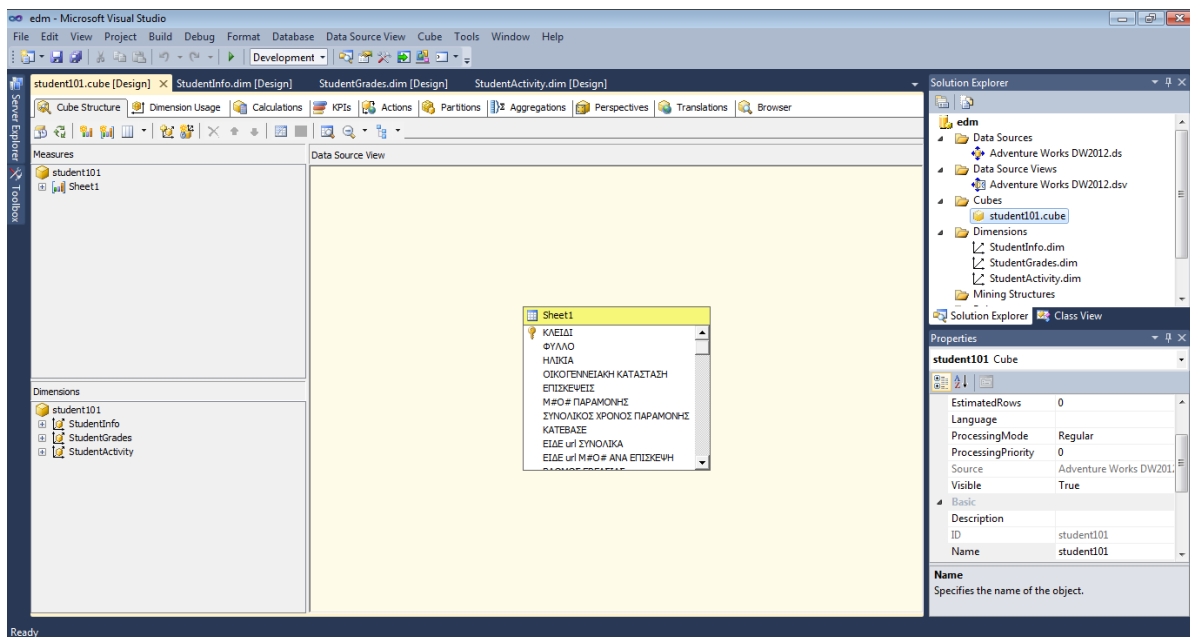
Εικόνα 54: Unselect Dimension

Την ξεκλικάρουμε και μετά πατάμε Next.



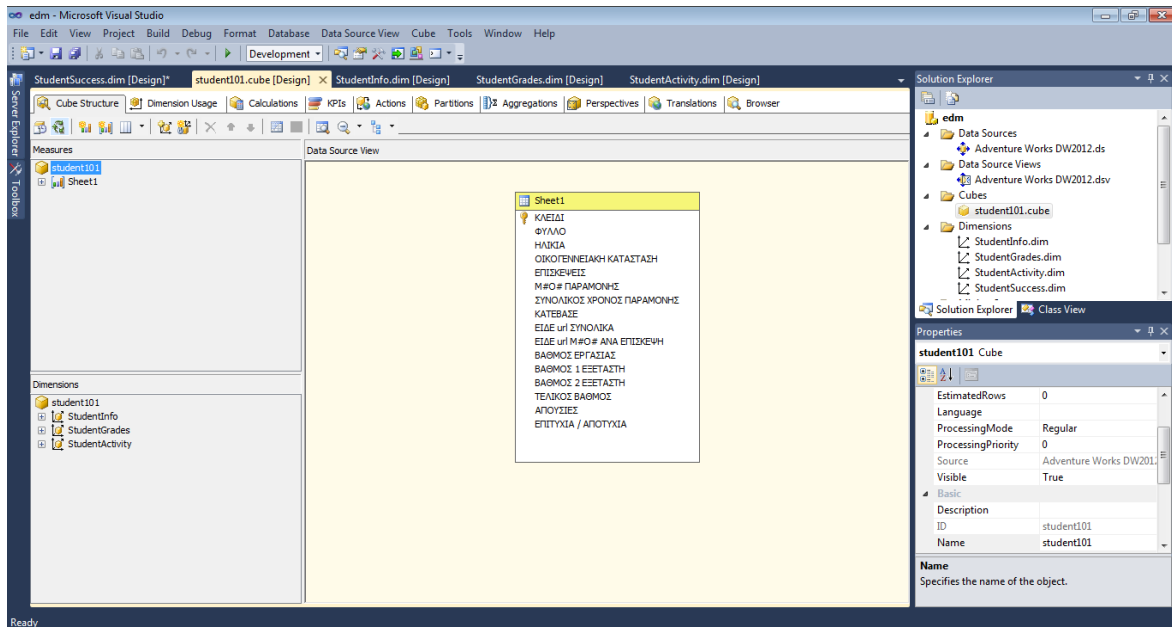
Εικόνα 55: Ολοκλήρωση Cube

Γραφούμε στο Cube name το όνομα που θέλουμε να βάλουμε στον κύβο μας, student101 και μετά Next.



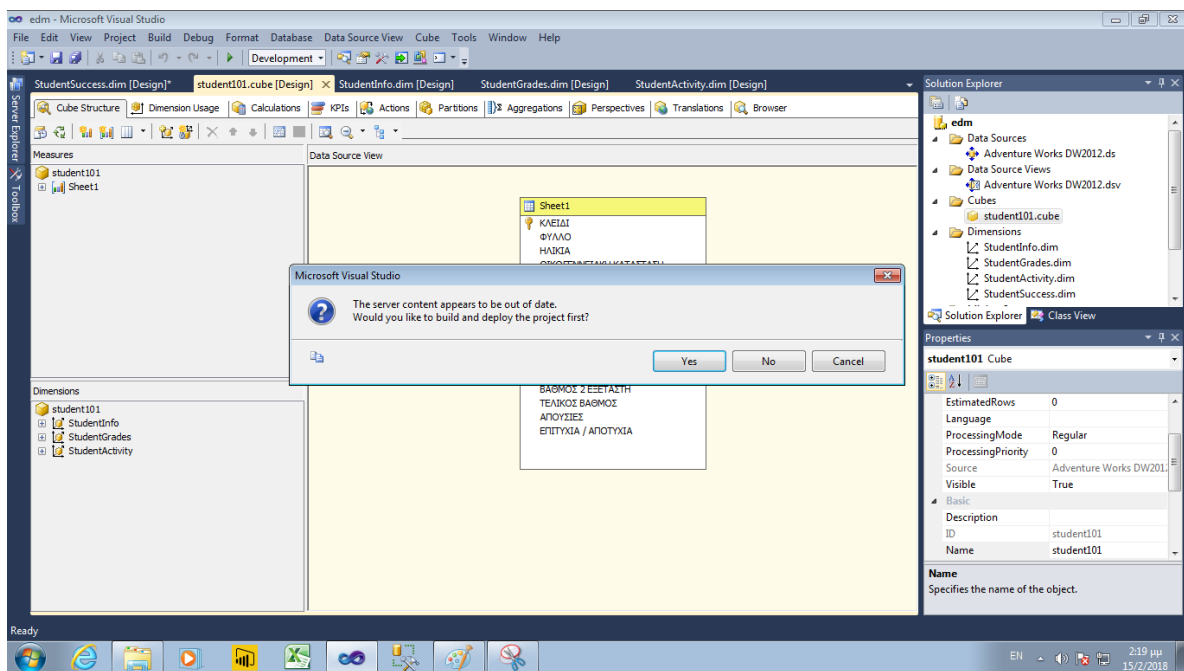
Εικόνα 56: Θέση Cube

Και έτσι δημιουργήθηκε ο κύβος.



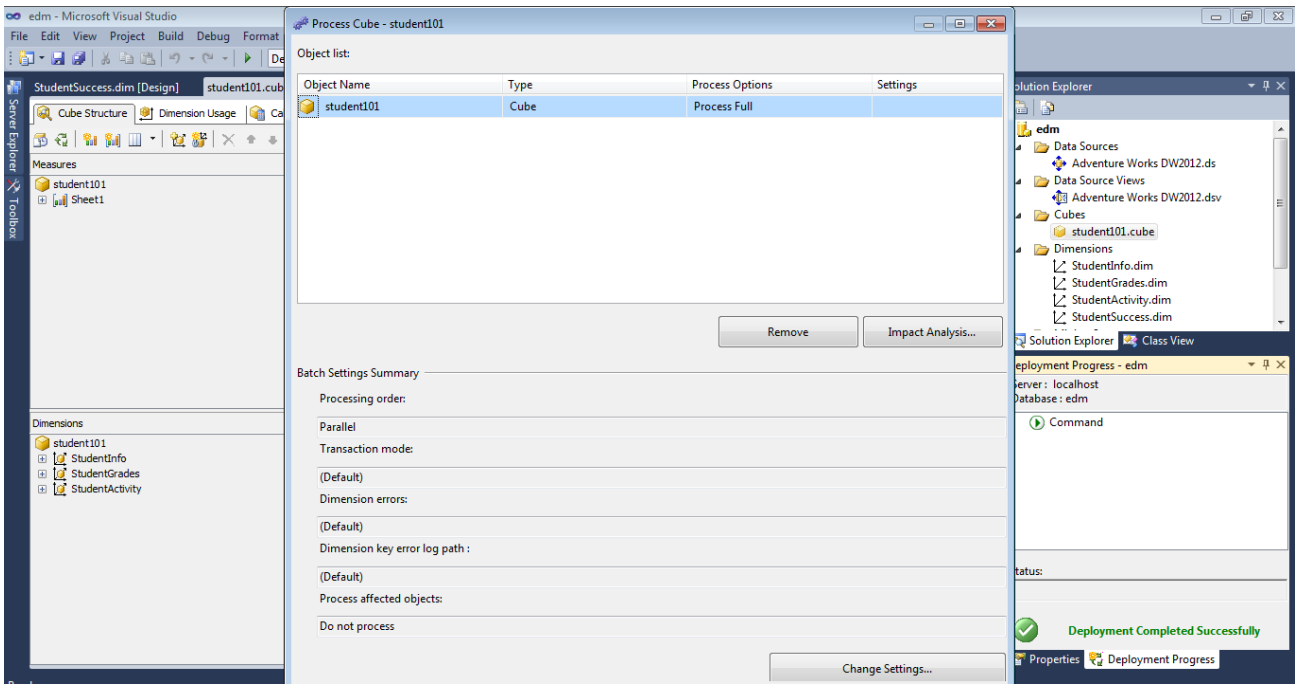
Εικόνα 57: Process

Θα πατήσουμε το εικονίδιο Process για να στείλουμε ότι δημιουργήσαμε στον Server.



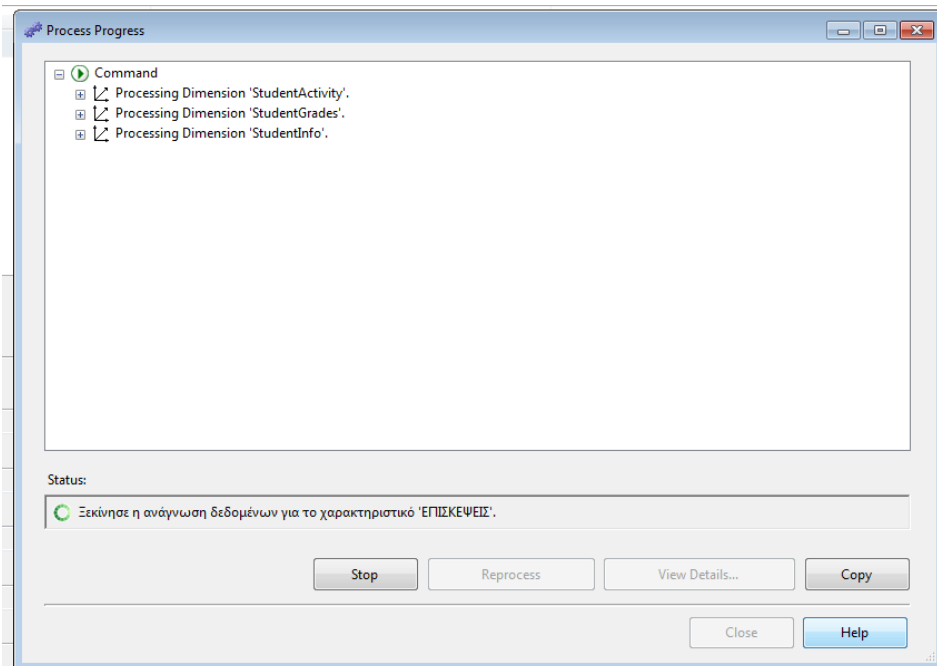
Εικόνα 58:Server Content

Θα πατήσουμε Yes.



Εικόνα 59: Process Cube

Μετά πατάμε Run.

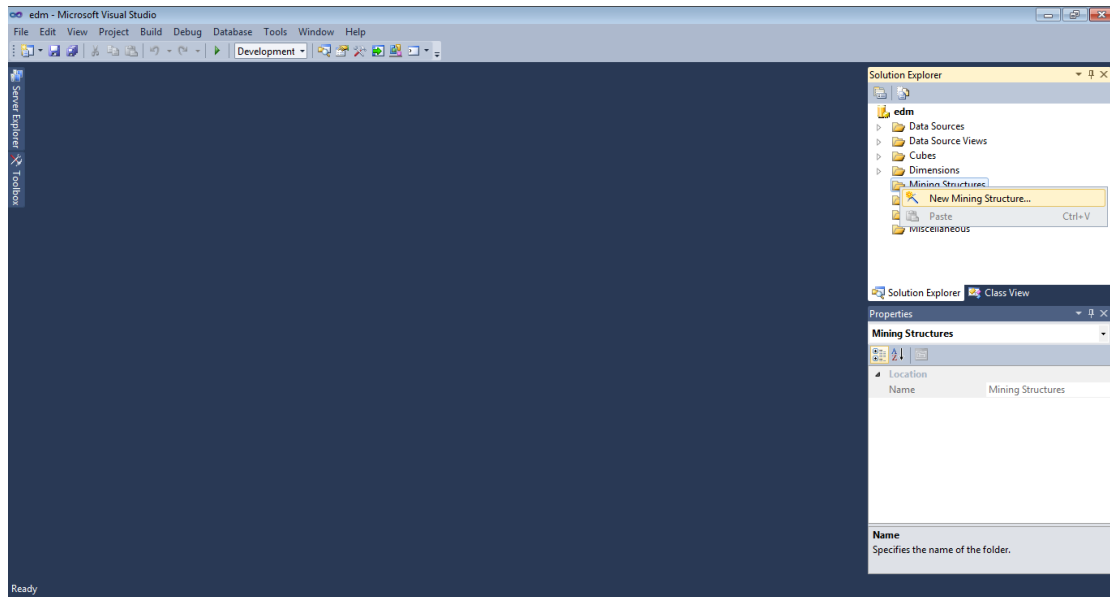


Εικόνα 60: Ολοκλήρωση Process

Όταν τελειώσει το Process πατάμε Close.

4.5 Εξόρυξη Δεδομένων (Data Mining)

Για να κάνουμε την Εξόρυξη Δεδομένων θα ακολουθήσουμε τα παρακάτω βήματα.



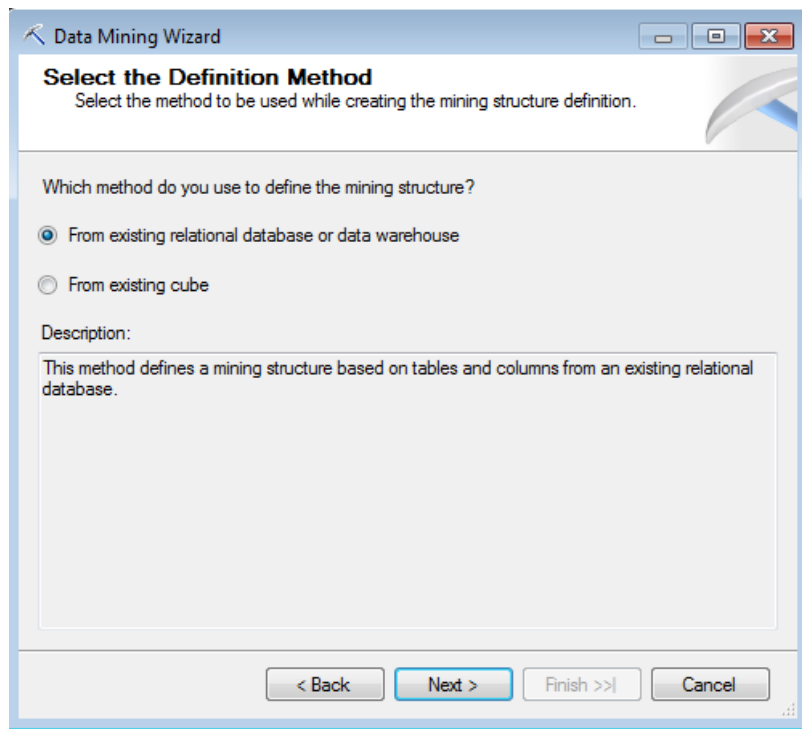
Εικόνα 61: New Mining Structure

Δεξί κλικ στο Mining Structures και θα επιλέξουμε New Mining Structure.



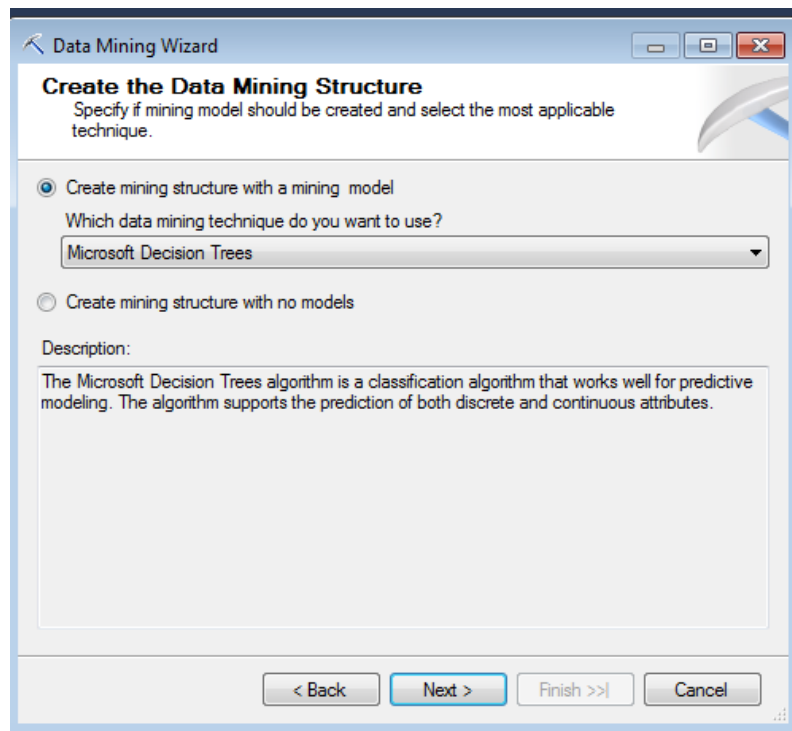
Εικόνα 62: Data Mining Wizard

Πατάμε Next στον Wizard.



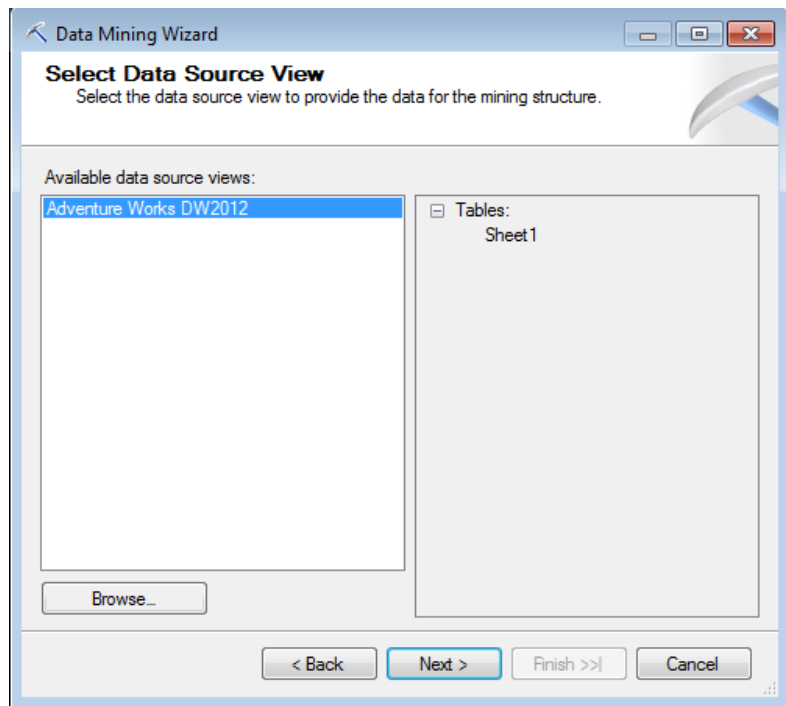
Εικόνα 63: Select the Definition Method

Επιλέγουμε From existing relational database or data warehouse και μετά Next.



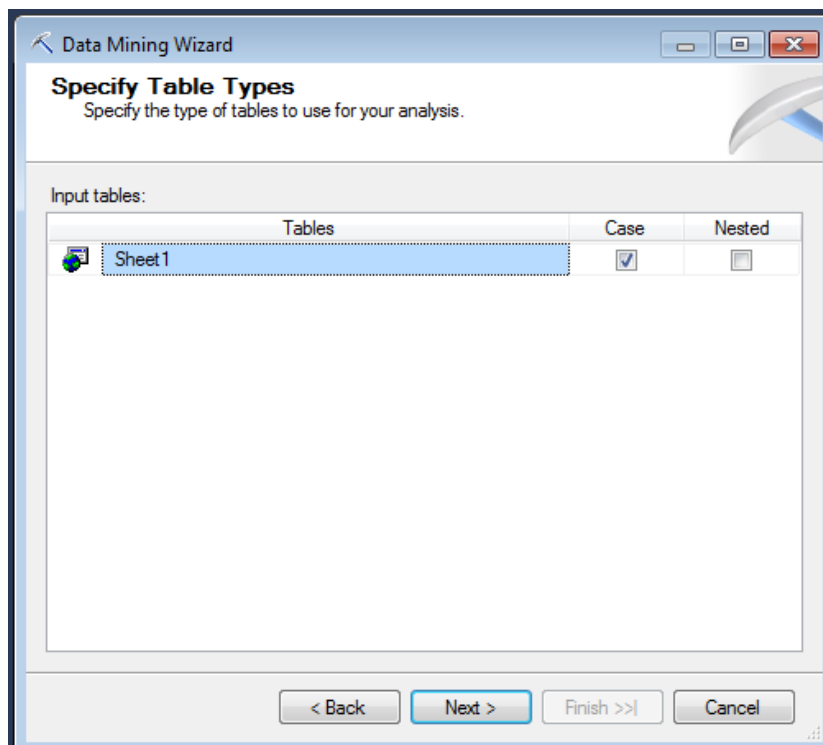
Εικόνα 64: Create the Data Mining Structure

Επιλέγουμε το Microsoft Decision Trees και μετά Next.



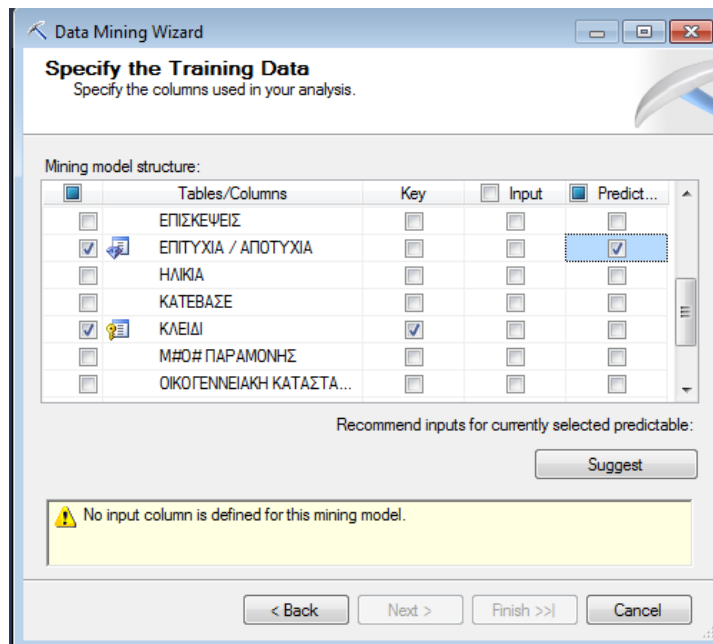
Εικόνα 65: Select Data Source View

Έχει επιλέγει το Data Source Views που θα πάρουμε δεδομένα και πατάμε Next.



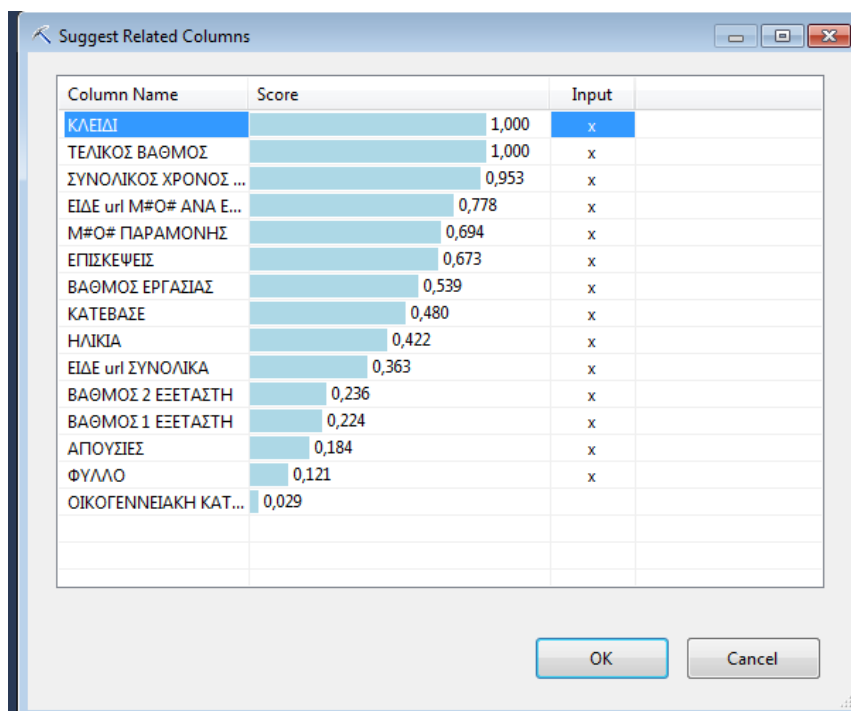
Εικόνα 66: Specify Table Types

Έχει επιλέγει ο πίνακας που έχει τις περιπτώσεις μας και πατάμε Next



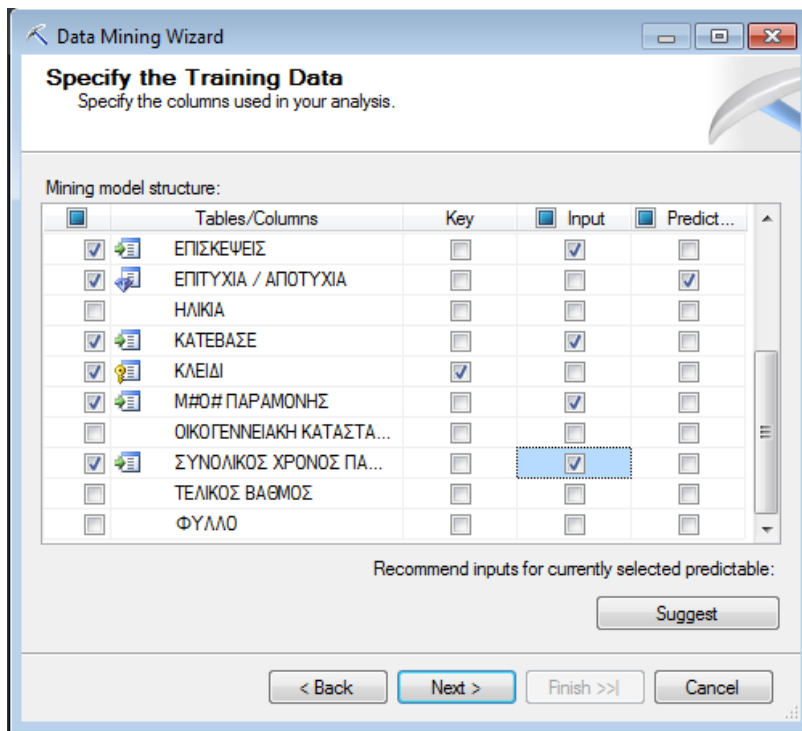
Εικόνα 67: Specify the Training Data

Θα κάνουμε την επιλογή των πεδίων και των μοντέλων. Επιλέγουμε το ΚΛΕΙΔΙ ως Key, το ΕΠΙΤΥΧΙΑ/ΑΠΟΤΥΧΙΑ ως Predictable και θα πατήσουμε Suggest.



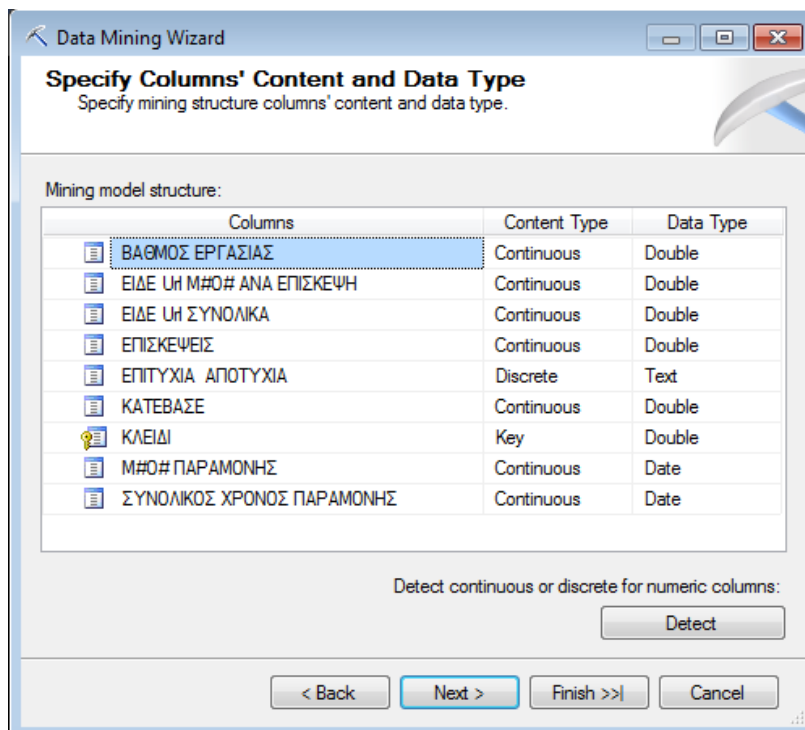
Εικόνα 68: Suggest Related Columns

Μου βγάζει με σειρά τι μου προτίνει να βάλω ως Input και πατάω OK.



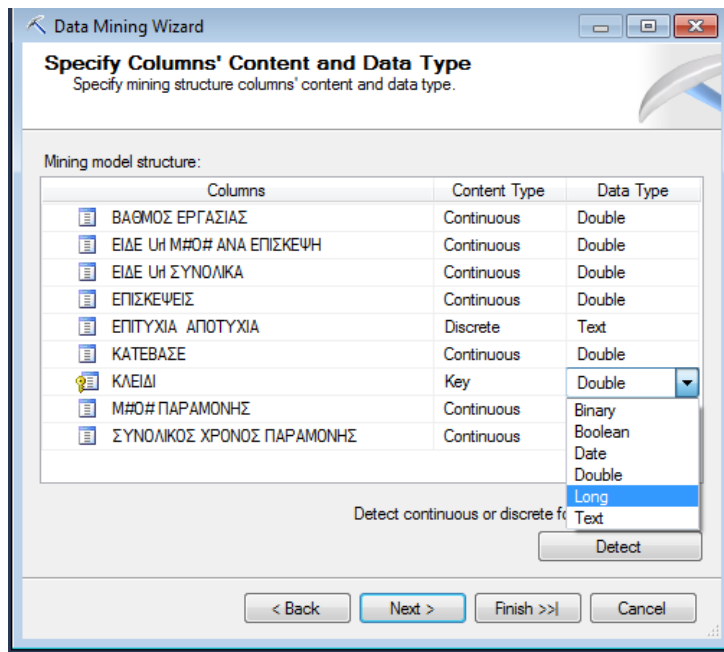
Εικόνα 69: Specify the Training Data

Θα επιλέξουμε ως Input τις ΕΠΙΣΚΕΨΕΙΣ, τον Μ.Ο. ΠΑΡΑΜΟΝΗΣ, το ΣΥΝΟΛΙΚΟΣ ΧΡΟΝΟΣ ΠΑΡΑΜΟΝΗΣ, το ΚΑΤΕΒΑΣΕ, το ΕΙΔΕ url ΣΥΝΟΛΙΚΑ, το ΕΙΔΕ url Μ.Ο. ΑΝΑ ΕΠΙΣΚΕΨΗ και το ΒΑΘΜΟΣ ΕΡΓΑΣΙΑΣ και πατάμε Next.



Εικόνα 70: Specify Columns Content and Data Type

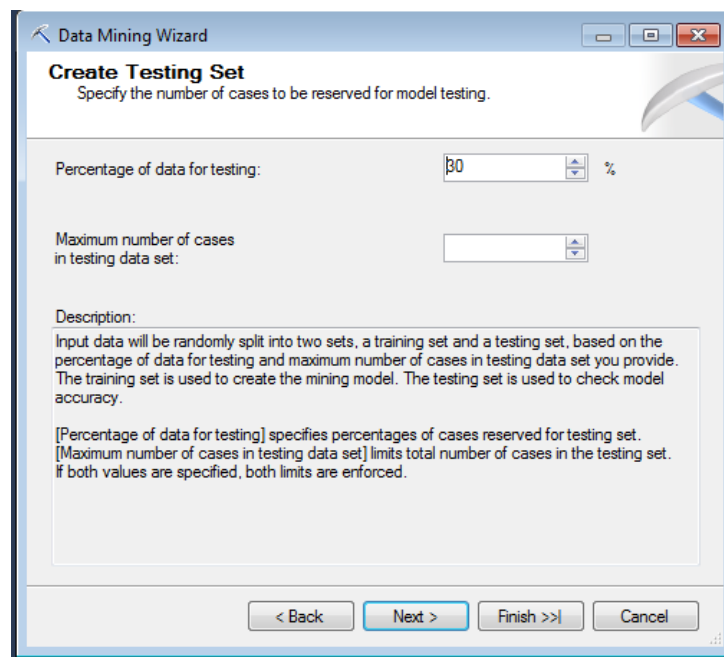
Εδώ θα δηλώσουμε το κλειδί και το ΕΠΙΤΥΧΙΑ ΑΠΟΤΥΧΙΑ ως Key και ως Discrete αντίστοιχα.



Εικόνα 71: Specify Columns Content and Data Type

Ακόμα το Κλειδί θα το δηλώσουμε και ως Long.

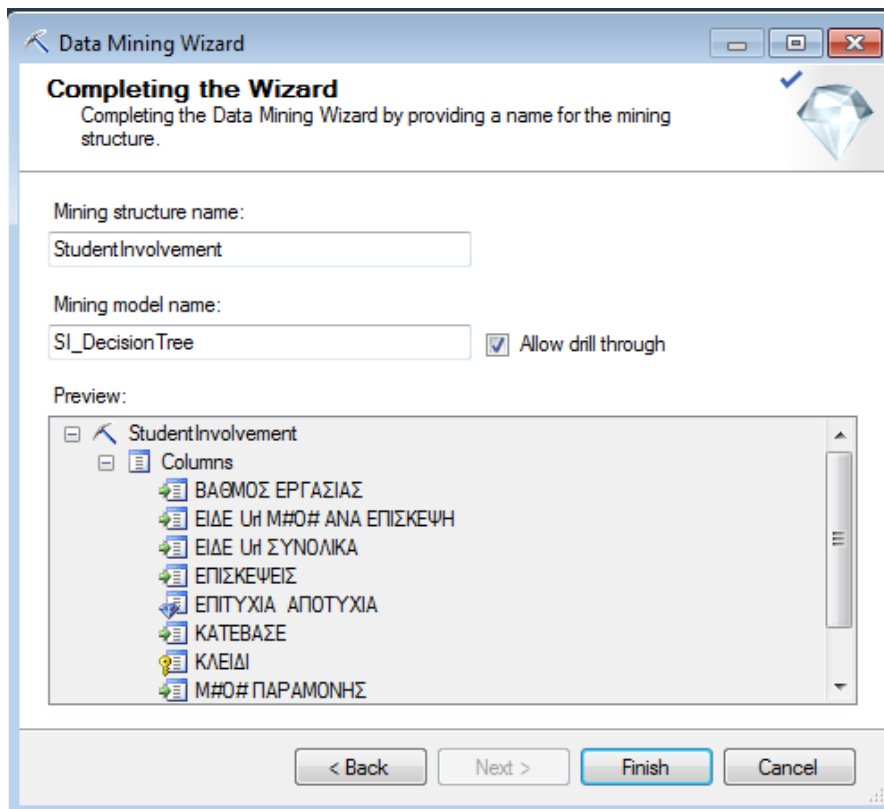
Μετά πατάμε Finish.



Εικόνα 72: Create Testing Set

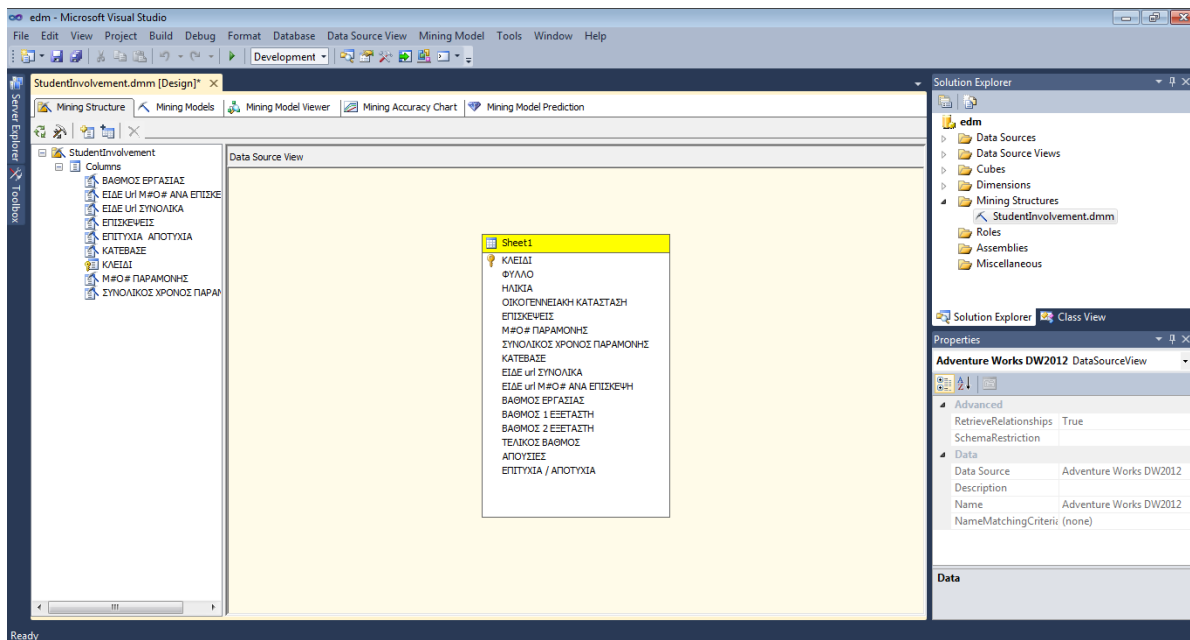
Θα επιλέξουμε το 30% στο Percentage of data for testing και έτσι θα κρατήσει το 30% για να κάνει επαλήθευση του μοντέλου.

Μετά πατάμε Next.



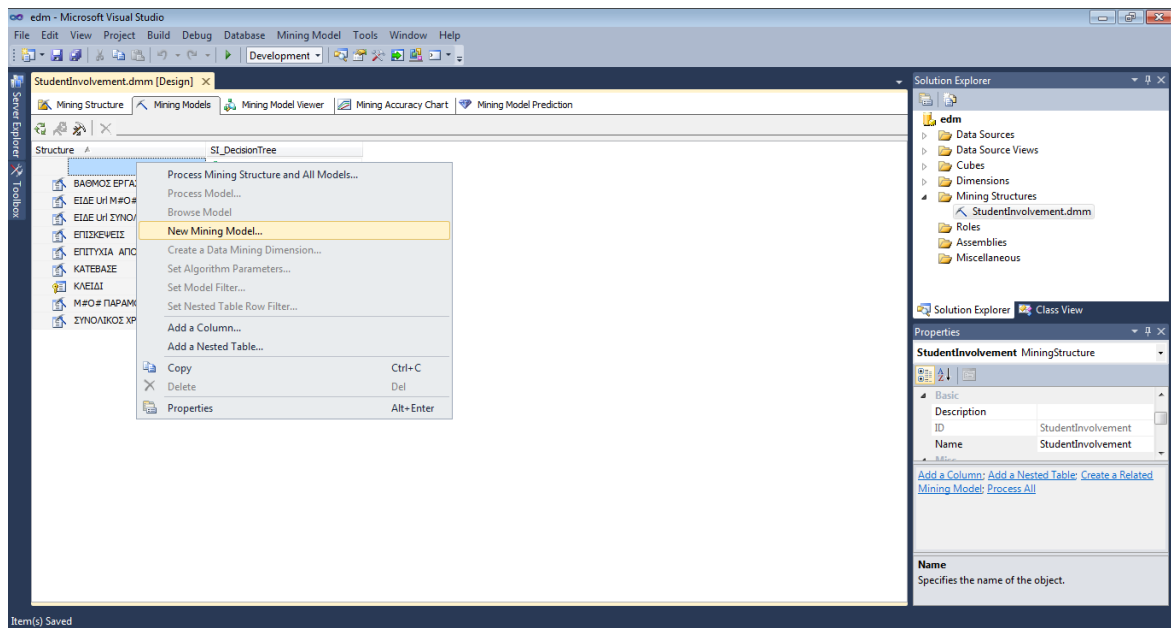
Εικόνα 73: Ολοκλήρωση Mining

Στο Mining structure name θα βάλουμε StudentInvolvement και μετά πατάμε Finish.



Εικόνα 74: Θέση StudentInvolvement.dmm

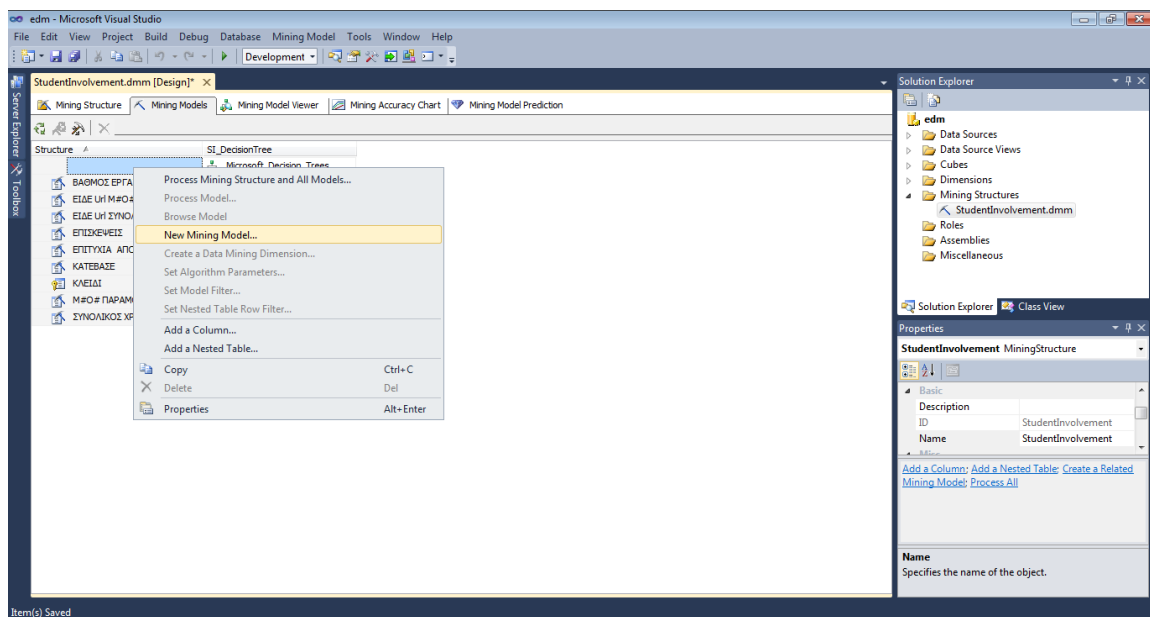
Ανοίξε στο Mining Structures το StudentInvolvement.dmm



Εικόνα 75: New Mining Model

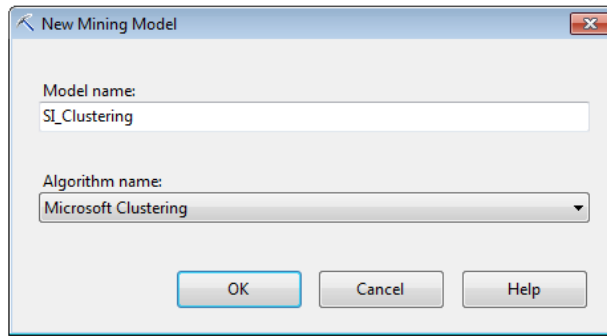
Πάμε στην καρτέλα Mining Models.

Κάνουμε Δεξί κλικ δίπλα στο Structure.



Εικόνα 76: New Mining Model

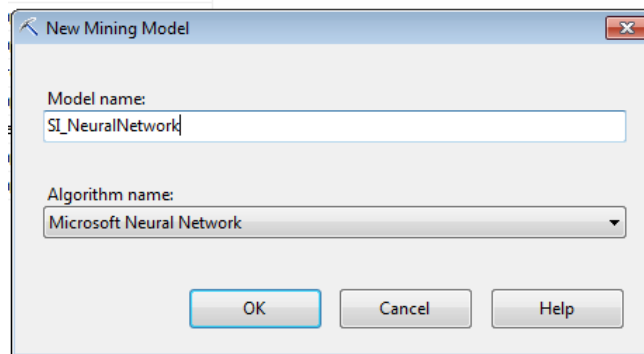
Επιλέγουμε το New Mining Model.



Εικόνα 77: Microsoft Clustering

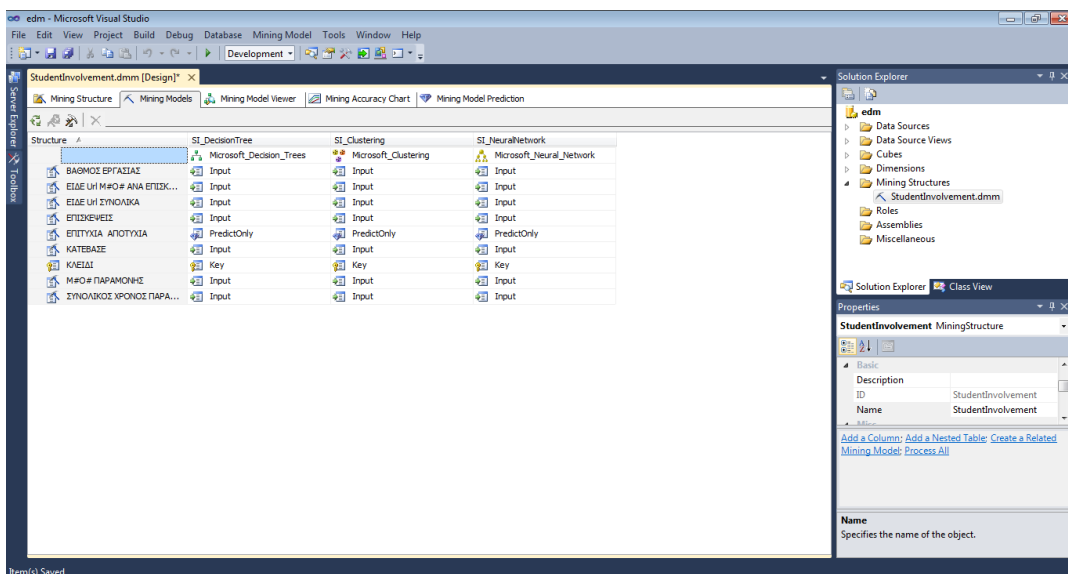
Επιλέγουμε στο Algorithm name το Microsoft Clustering και γράφουμε στο Model name SI_Clustering.

Πατώντας το OK δημιουργήθηκε το Model Clustering.



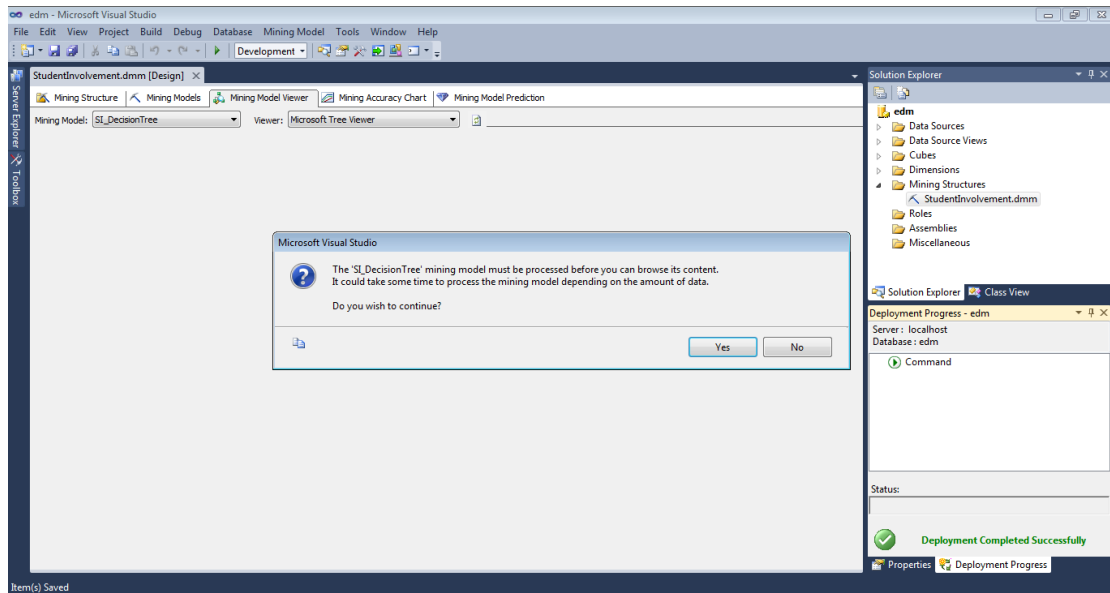
Εικόνα 78: Microsoft Neural Network

Όλη την παραπάνω διαδικασία την ακολουθούμε για να φτιάξουμε και το Model Neural Network με όνομα SI_NeuralNetwork.



Εικόνα 79: Mining Models

Πατώντας το OK βλέπουμε και τα τρία Μοντέλα

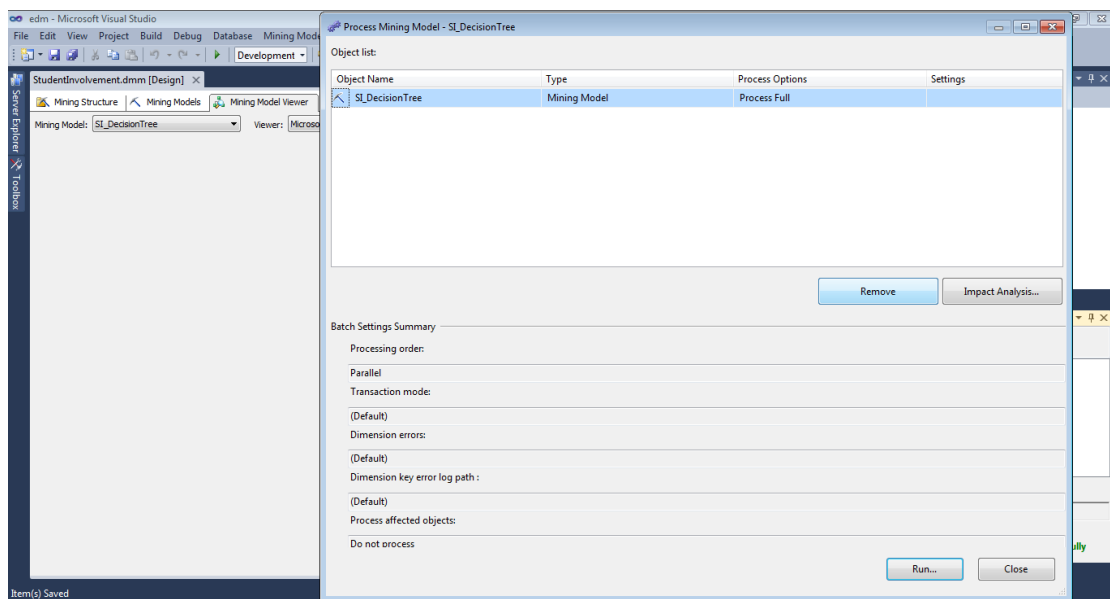


Εικόνα 80: Mining Model Viewer

Επιλέγοντας την καρτέλα Mining Model Viewer για να δούμε τα χαρακτηριστικά των Αλγορίθμων.

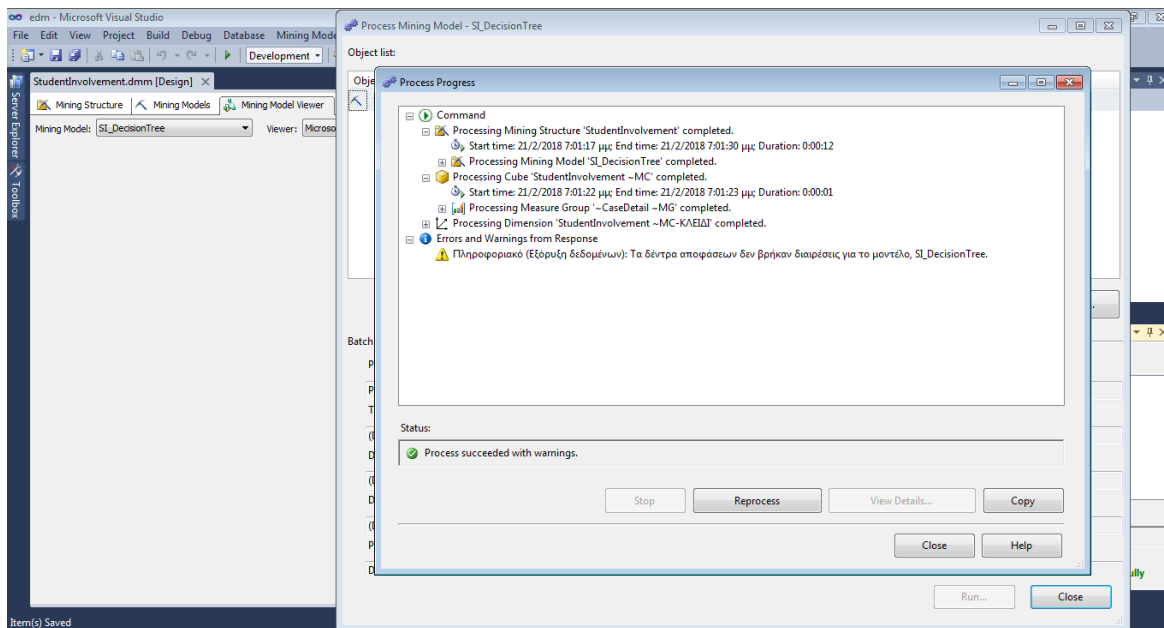
Μετά εμφανίζεται ένα παράθυρο να τα στείλουμε στον server.

Πατάμε Yes για να τα στείλει.



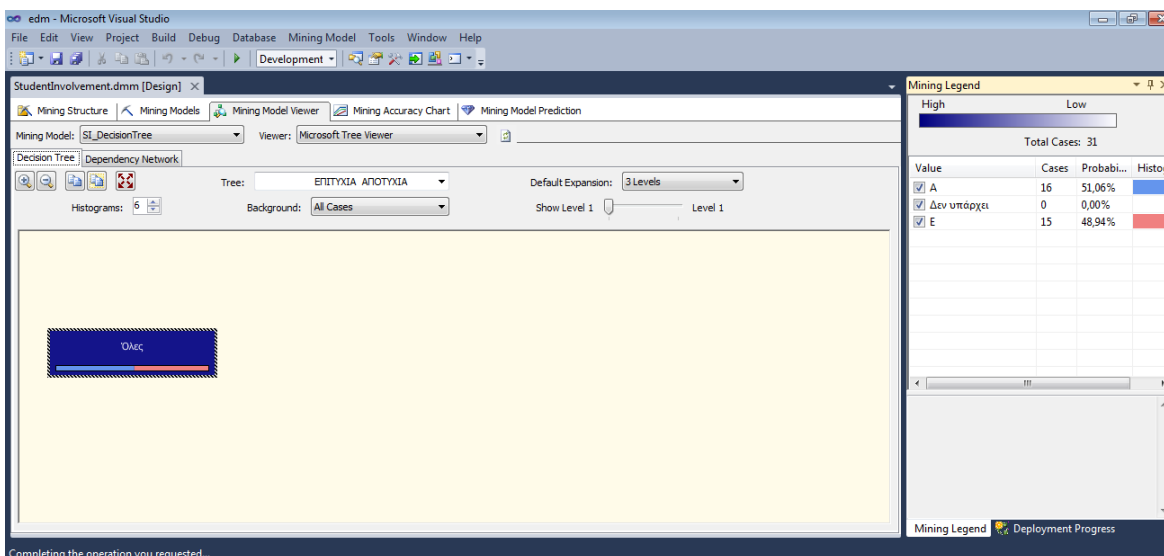
Εικόνα 81: Process Mining Model

Μετά πατάμε Run για να γίνει Process.



Εικόνα 82: Process Progress

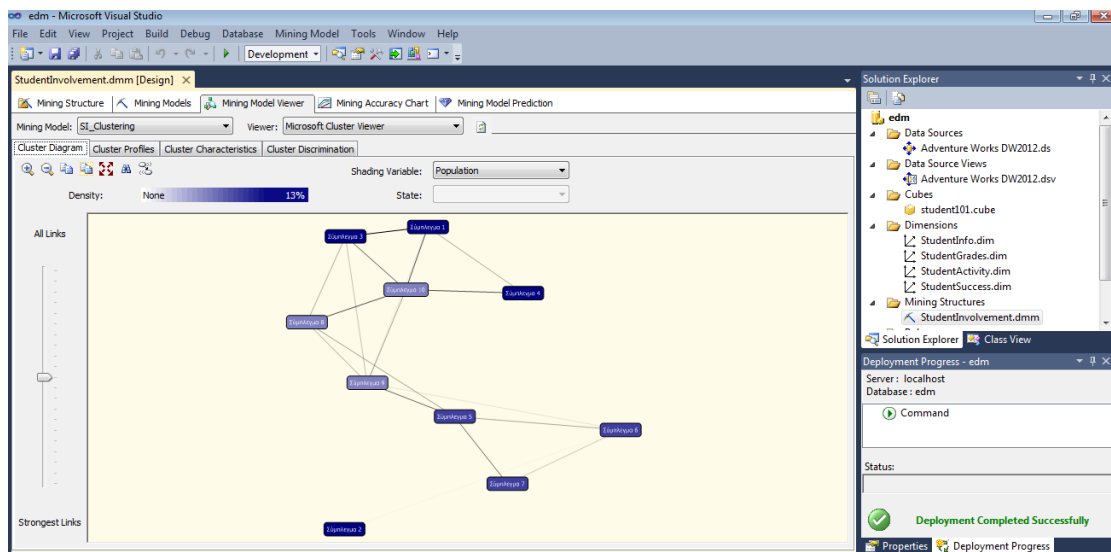
Το Process έγινε και πατάμε Close για να κλείσει η καρτέλα.



Εικόνα 83: Decision Tree

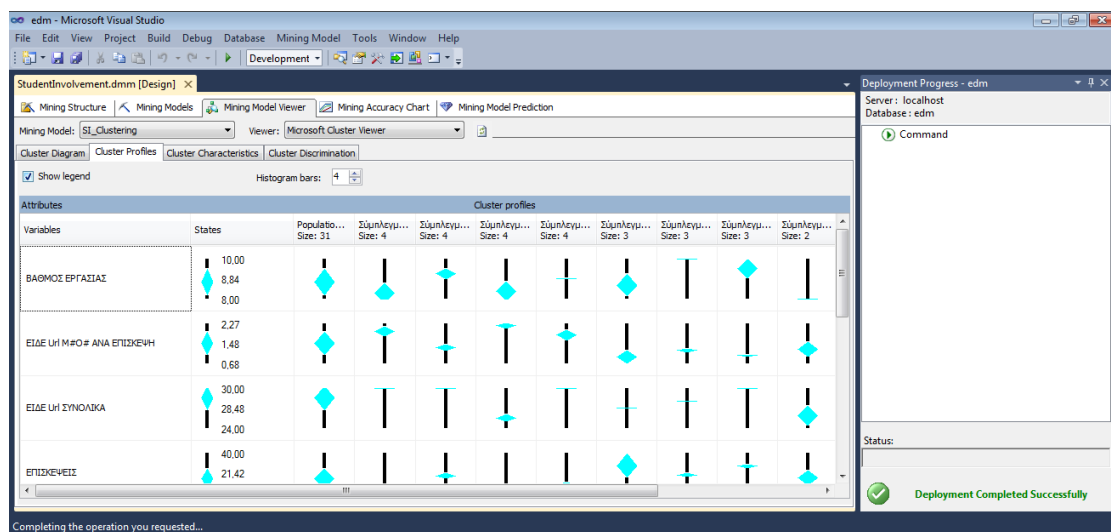
Στην καρτέλα Decision Tree μας βγάζει το διάγραμμα του Αλγόριθμου. Ακόμα παρατηρούμε στην καρτέλα Mining Legend ότι στις 31 περιπτώσεις, οι 16 περιπτώσεις με 51,06% πιθανότητα είναι Αποτυχία και οι 15 περιπτώσεις με 48,94% πιθανότητα είναι Επιτυχία.

Υστερα κάνουμε την ίδια δια διαδικασία για να τα στείλουμε στον Server και του Process για το Μοντέλο Clustering και για το Μοντέλο NeuralNetwork.



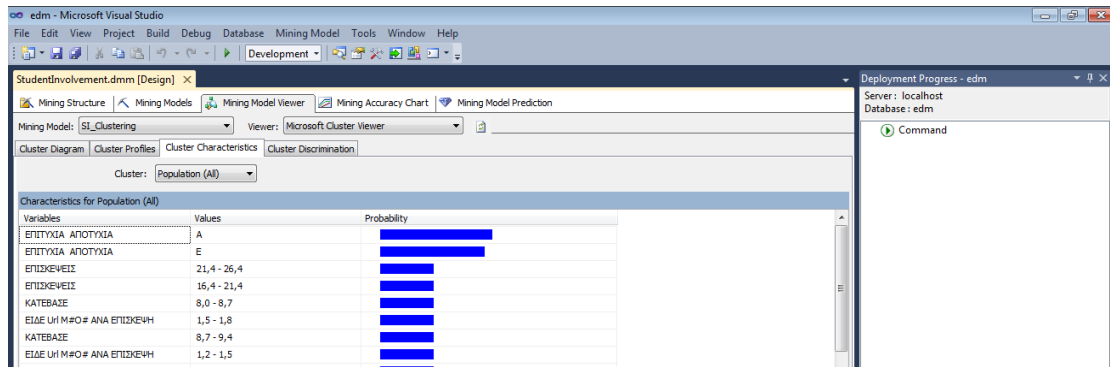
Εικόνα 84: Cluster Diagram

Στην καρτέλα Cluster Diagram μας βγάζει το διάγραμμα του Αλγόριθμου.



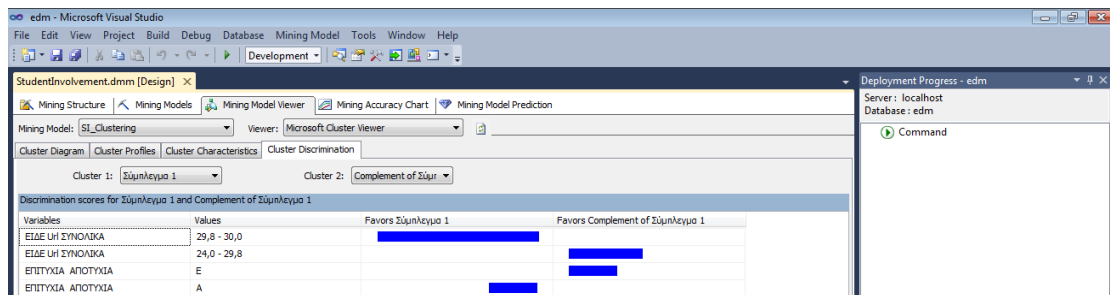
Εικόνα 85: Cluster Profiles

Στην καρτέλα Cluster Profiles παρατηρούμε για κάθε μεταβλητή το συμπλεγμα που είναι και σε τι τιμές κινούνται.



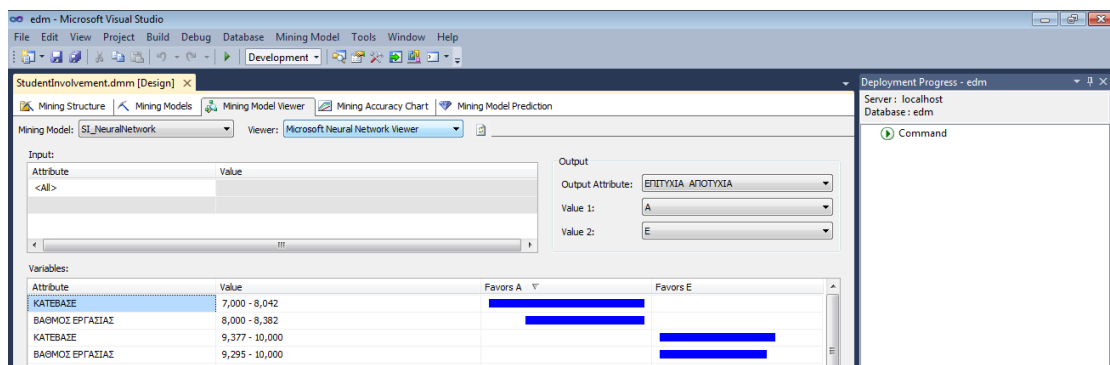
Εικόνα 86: Cluster Characteristics

Στην καρτέλα Cluster Characteristics παρατηρούμε για κάθε σύμπλεγμα την πιθανότητα που μπορεί να πάρει αυτές τις τιμές.



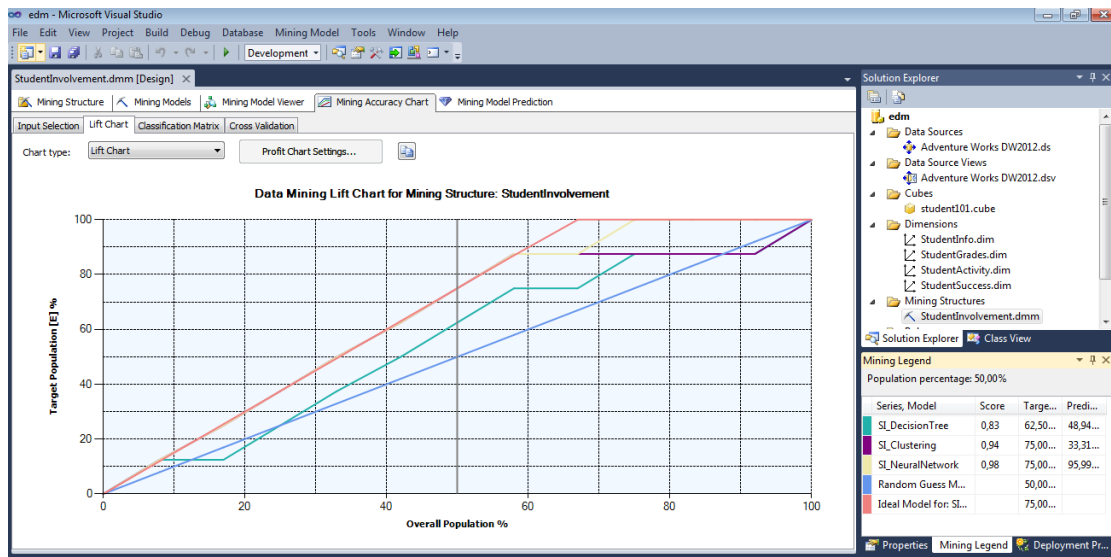
Εικόνα 87: Cluster Discrimination

Στην καρτέλα Cluster Discrimination μπορούμε να συγκρίνουμε τα συμπλέγματα μεταξύ τους.



Εικόνα 88: Mining Model Viewer

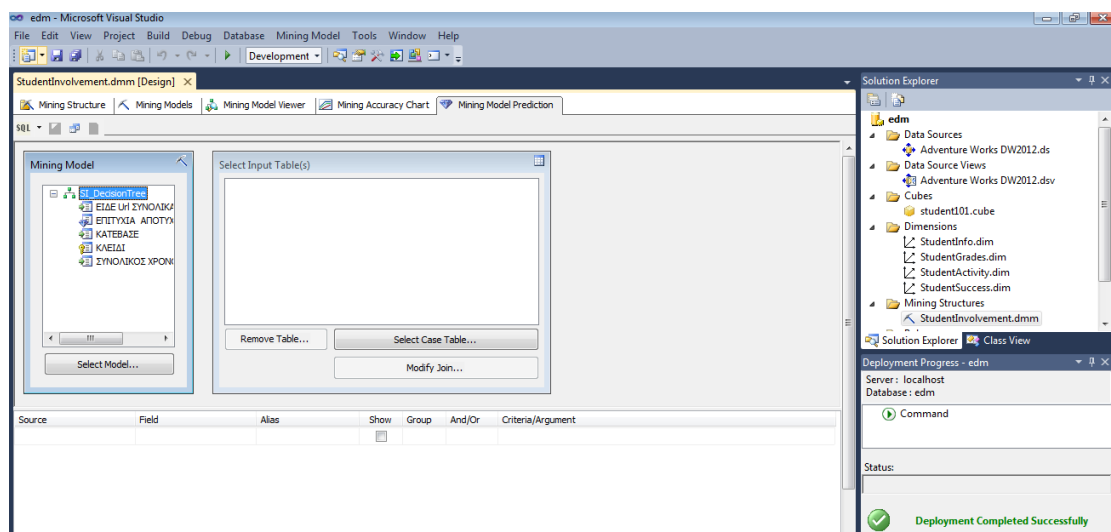
Στην καρτέλα Mining Model Viewer για τον Αλγόριθμο Neural Network μπορούμε να συγκρίνουμε κομβους μεταξύ τους.



Εικόνα 89: Mining Accuracy Chart

Στην καρτέλα Mining Accuracy Chart, στην υποκαρτέλα βλέπουμε το διάγραμμα με όλα τα μοντέλα μας. Αυτό το διάγραμμα μας δείχνει πόσο τα μοντέλα μας πλησιάζουν το ιδανικό μοντέλο. Στην περίπτωση μας το Μοντέλο Neural Network έχει το μεγαλύτερο Score.

Στην καρτέλα Classification Matrix βλέπουμε στο κάθε μοντέλο το αποτελεσματά μας είναι πετυχημένα



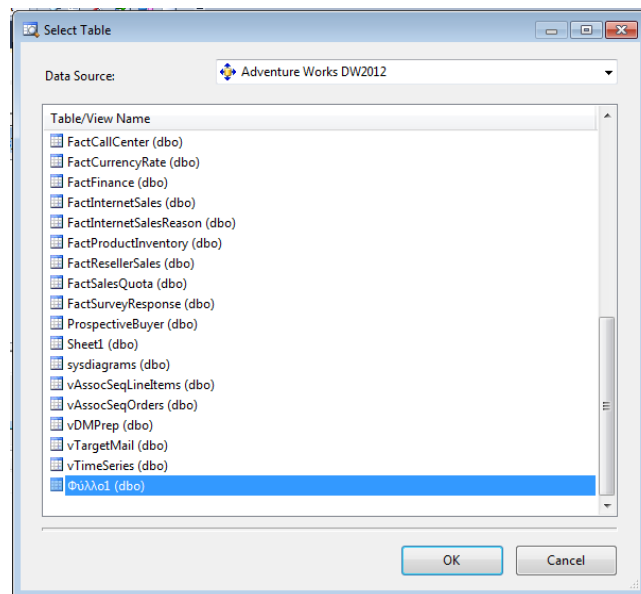
Εικόνα 90: Mining Model Prediction

Στην καρτέλα Mining Model Prediction μπορούμε να κάνουμε πρόβλεψη επιλεγοντας το μοντέλο που επιθυμούμε και βάζοντας εμείς στοιχεία σε κάθε μεταβλητή. Μας δίνει όμως και την δυνατότητα να κάνουμε πολλές προβλέψεις ταυτοχρόνα. Για να γίνει αυτό πρέπει να έχουμε κάνει έναν πίνακα που να έχει για κάθε μεταβλητή τα στοιχεία που επιθυμούμε. Για τον λόγο αυτό δημιουργήθηκε με Test Data το παρακάτω αρχείο Excel.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	
	ΚΛΕΙΔΙ	ΦΥΛΛΟ	ΗΛΙΚΙΑ	ΟΙΚΟΓΕΝΕΙΑΚΗ ΚΑΤΑΣΤΑΣΗ	ΕΠΙΣΚΕΨΕΙΣ	Μ.Ο. ΠΑΡΑΜΟΝΗΣ	ΣΥΝΟΛΙΚΟΣ ΧΡΟΝΟΣ ΠΑΡΑΜΟΝΗΣ	ΚΑΤΕΒΑΣΕ	ΕΙΔΕ url	Μ.Ο. ΑΝΑ ΕΠΙΣΚΕΨΗ	ΒΑΘΜΟΣ ΕΡΓΑΣΙΑΣ	ΒΑΘΜΟΣ 1 ΕΞΕΤΑΣΤΗ	ΒΑΘΜΟΣ 2 ΕΞΕΤΑΣΤΗ	ΒΑΘΜΟΣ 3 ΕΞΕΤΑΣΤΗ	ΤΕΛΙΚΟΣ ΒΑΘΜΟΣ	ΑΠΟΥΣΙΕΣ	ΕΠΙΤΥΧΙΑ / ΑΠΟΤΥΧΙΑ
1	1	A	28	A	18,7	0,22	3,74	10,8	18	1,06	6,4	9	9	8,0	1	E	
2	2	Γ	24	A	29,7	0,03	0,74	6,3	0	0,00	6,4	8,1	9	7,7	1	A	
3	3	A	37	Π	30,8	0,10	2,77	8,1	27	0,96	7,2	8,37	7,65	7,7	0	A	
4	4	Γ	42	Π	35,2	0,20	6,34	7,2	26,1	0,82	6,8	8,37	8,1	7,7	2	A	
5	5	A	28	A	29,7	0,26	7,13	3,6	19,8	0,73	5,36	8,1	8,1	7,0	0	A	
6	6	A	33	Π	35,2	0,33	10,56	6,3	23,4	0,73	6,4	8,1	8,1	7,4	0	A	
7	7	A	34	A	23,1	0,77	16,17	6,3	26,1	1,24	6,24	8,1	7,65	7,2	0	A	
8	8	Γ	41	A	15,4	0,13	1,85	5,4	27	1,93	7,2	8,1	8,1	7,7	2	A	
9	9	A	29	A	24,2	0,17	3,63	7,2	27	1,23	6,4	9	7,65	7,6	2	A	
10	10	A	30	A	15,4	0,10	1,39	7,2	24,3	1,74	6,24	8,55	8,1	7,5	1	A	
11	11	Γ	41	Π	19,8	0,46	8,32	1,8	25,2	1,40	8	8,55	8,1	8,2	0	E	
12	12	A	39	Π	18,7	0,56	9,54	7,2	27	1,59	5,6	8,55	8,1	7,2	2	A	
13	13	Γ	28	A	16,5	0,32	4,79	9	32,4	2,16	6,4	8,55	9	7,8	1	E	
14	14	A	27	A	31,9	0,23	6,70	6,3	13,5	0,47	7,6	9	9	8,4	0	E	
15	15	Γ	30	A	35,2	0,10	3,17	5,4	6,3	0,20	7,2	9	9	8,3	0	E	
16	16	Γ	37	A	18	0,41	8,10	9,9	33	1,65	10	10	10	10,0	2	E	
17	17	A	32	A	29	0,37	10,73	10	26	0,90	6,5	9,5	10	8,5	1	A	
18	18	Γ	45	A	17	0,28	4,76	6	27	1,59	9	9,5	10	9,5	2	E	
19	19	A	50	Π	22	0,65	14,30	9	30	1,36	9,5	9	9	9,2	0	E	
20	20	A	29	Π	30	0,45	13,50	12	27	0,90	7	9	10	8,5	2	A	
21	21	Γ	45	Π	19	0,32	6,08	9	24	1,26	9	9,5	9	9,2	1	E	
22	22	Γ	44	A	15	0,74	11,10	7	30	2,00	9	9,5	9	9,2	2	E	

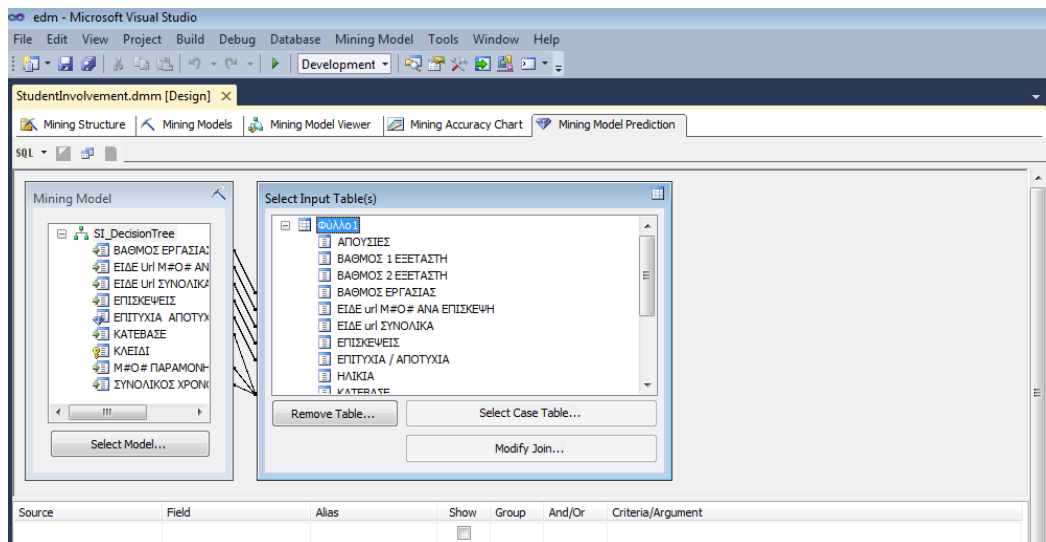
Εικόνα 91: Test Data

Με τα βήματα που έγινε ο πίνακας από το αρχείο του Excel με τα δεδομένα μας, τα ίδια θα ακολουθηθούν για το παραπάνω αρχείο του Excel ώστε να γίνει η πρόβλεψη. Έτσι πατώντας την επιλογή Select Case Table βγαίνει ένα παράθυρο που μας δίνει την δυνατότητα να επιλέξουμε πίνακα.



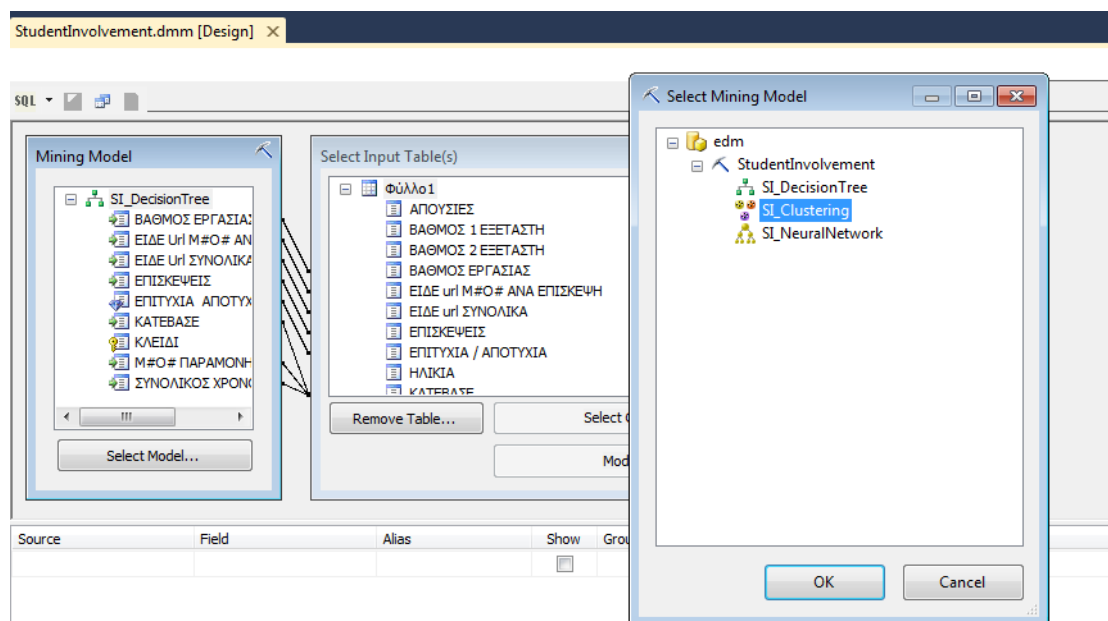
Εικόνα 92: Select Table

Μετά επιλέγουμε τον πίνακα Φύλλο1.dbo και πατάμε OK.



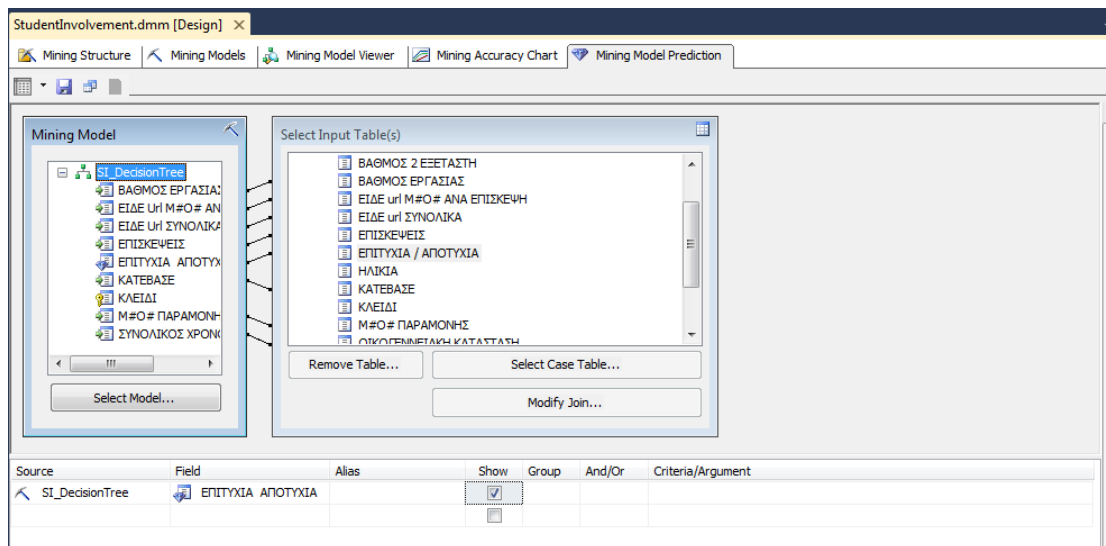
Εικόνα 93: Φύλλο1

Ύστερα πατάμε την επιλογή Select Model



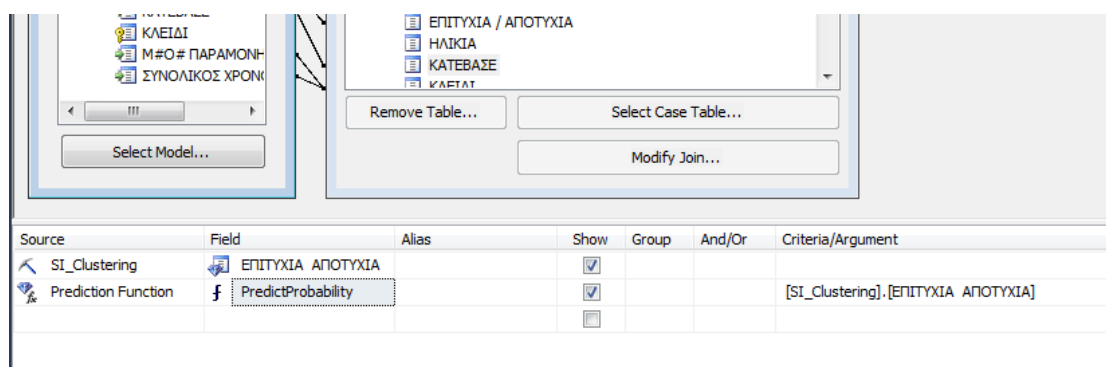
Εικόνα 94: Select Mining Model

Επιλέγουμε όποιο μοντέλο θέλουμε και μετά πατάμε OK.



Εικόνα 95: Συμπλήρωση Field

Επιλέγουμε την μεταβλητή ΕΠΙΤΥΧΙΑ ΑΠΟΤΥΧΙΑ και την τοποθετούμε στην πρώτη σειρά στην στήλη Field.



Εικόνα 96: Συμπλήρωση Criteria/Argument

Μετά στην δεύτερη σειρά στήλη Source επιλέγουμε το Prediction Function, στην στήλη Field επιλέγουμε το PredictProbability και στην στήλη Criteria/Argument τοποθετούμε την μεταβλητή ΕΠΙΤΥΧΙΑ ΑΠΟΤΥΧΙΑ.

Ύστερα πατάμε το κουμπί result view και μας βγαίνει το παρακάτω παράθυρο.

ΕΠΙΤΥΧΙΑ ΑΠΟ...	Expression
A	0,6669322177...
A	1
A	0,6669322177...
A	0,6669322177...
A	0,6669322177...
A	0,6669322177...
A	1
A	0,6669322177...
A	0,6669322177...
E	1
A	1
A	1
E	1
A	0,6669322177...
A	0,6669322177...
E	1
A	0,6669322177...
A	0,6669322177...
A	0,5000396065...
A	0,6669322177...

Εικόνα 97: Mining Model Prediction

Στην αριστερή στήλη βγαίνουν οι προβλέψεις για την κάθε test παρατήρηση που έχουμε βάλει και στην δεξιά στήλη βγαίνει η πιθανότητα αυτή η πρόβλεψη να είναι σωστή.

Επίλογος Συμπεράσματα

Καταλήγοντας, στα συμπεράσματα της μελέτης αυτής είναι ότι επιβεβαιώνονται οι δυνατότητες που δίνει το EDM (Educational Data Mining), για την βελτιστοποίηση της ποιότητας της μάθησης.

Οι δυνατότητες αυτές στηρίζονται πάνω στην πρόβλεψη των αποτελεσμάτων μάθησης και φυσικά οι ευκαιρίες που δίνονται στην εκπαιδευτική κοινότητα από αυτή τη νέα επιστήμη της εξόρυξης γνώσης μέσα από δεδομένα και πιο συγκεκριμένα από εκπαιδευτικά δεδομένα.

Η επιστημονική κοινότητα δηλαδή μπορεί μέσα από αυτή τη διαδικασία να βοηθηθεί στο να αντιληφθεί ποιες είναι αυτές οι μεταβλητές που βοηθούν περισσότερο τους εκπαιδευόμενους στο να κατανοούν καλύτερα, να κινητοποιηθούν και να καλύψουν την ανάγκη τους για την γνώση.

Ποτέ άλλοτε δεν είχε την δυνατότητα κάποιος να έχει στα χέρια του εργαλεία που θα μπορούσε να αναλύσει τόσο όγκο δεδομένων και με τέτοια ακρίβεια για να φέρει το αποτέλεσμα που θέλει.

Έτσι μέσα από αυτή την εργασία παρουσιάστηκε με απλό τρόπο πως κάποιος μπορεί να αντλήσει δεδομένα, να μετρήσει κάποιες μεταβλητές από ένα διαδικτυακό μάθημα και να τις αναλύσει.

Έτσι με ένα θεωρητικό παράδειγμα, δεδομένα επεξεργάστηκαν, αναλύθηκαν και φάνηκε σε ποιες μεταβλητές πρέπει να αλλάξουν και πόσο για να έρθει το κατάλληλο αποτέλεσμα, μέσα από μια συγκεκριμένη μεθοδολογία που είναι το EDM.

Αυτή η εξόρυξη γνώσης είναι πολύ σημαντική για όποιον ασχολείται με την εκπαίδευση αλλά πολύ περισσότερο για όλη την ανθρωπότητα διότι βελτιώνοντας τα μαθησιακά αποτελέσματα αυξάνεται η μόρφωση στον άνθρωπο.

«Επιστήμη ποιητική ευδαιμονίας», (Η γνώση δημιουργεί ευημερία) Πλάτων 427-347 π.Χ.

Πίνακας ακρωνύμιων

Course Management System	CMS
Data Mining	DM
Educational Data Mining	EDM
Information and Communication Technologies	ICT
Knowledge Discovery in Data Bases	KDD
Knowledge Management Systems	KMS
Learning Management Systems	LMS
Machine Learning	ML
Virtual Learning Environments	VLE

Κατάλογος Εικόνων

Εικόνα 1: Data Mining (classroom-aid.com).....	2
Εικόνα 2: Big Data (laceproject.eu).....	7
Εικόνα 3: E-Learning (tibau.org).....	9
Εικόνα 4: Δημοφιλέστερα συστήματα LMS (Capterra,Inc).....	12
Εικόνα 5: Educational data mining process (Garcia et al., 2011).....	20
Εικόνα 6: Επισκόπηση του τρόπου με τον οποίο εφαρμόζονται οι μέθοδοι EDM.....	21
Εικόνα 7: Διαδικασία ανακάλυψης γνώσης στα Εκπαιδευτικά Ιδρύματα.....	22
Εικόνα 8: Test Data.....	25
Εικόνα 9: SQL Server.....	27
Εικόνα 10: Επιλογή Databases.....	27
Εικόνα 11: Εισαγωγή Δεδομένων.....	28
Εικόνα 12: Import Wizard.....	28
Εικόνα 13: Επιλογή των Δεδομένων.....	29
Εικόνα 14: Επιλογή ποορισμού.....	29
Εικόνα 15: Αντιγραφή Πίνακα.....	30
Εικόνα 16: Επιλογή Φύλλου Δεδομένων.....	30
Εικόνα 17: Επιλογή Run Immediately.....	31
Εικόνα 18: Complete the Wizard.....	31
Εικόνα 19: Ενημέρωση Ολοκλήρωσης.....	32
Εικόνα 20: Θέση Πίνακα Sheet1.....	32
Εικόνα 21: Επιλογή New Project.....	33
Εικόνα 22: Θέση New Project.....	33
Εικόνα 23: Δημιουργία Data Source.....	34
Εικόνα 24: Data Source Wizard.....	34
Εικόνα 25: How to define the connection.....	35
Εικόνα 26: Επιλογή Server.....	35
Εικόνα 27: Επιλογή Database.....	36
Εικόνα 28: How to define the connection.....	36
Εικόνα 29: Impersonation Information.....	37

Εικόνα 30: Ολοκλήρωση Data Source.....	37
Εικόνα 31: Δημιουργία Data Source View.....	38
Εικόνα 32: Data Source View Wizard.....	38
Εικόνα 33: Select a Data Source.....	39
Εικόνα 34: Select Tables and Views.....	39
Εικόνα 35: Επιλογή Sheet1.....	40
Εικόνα 36: Ολοκλήρωση Data Source View.....	40
Εικόνα 37: Δημιουργία Πρωτεύον Κλειδί.....	41
Εικόνα 38: Πρωτεύον Κλειδί.....	41
Εικόνα 39: Δημιουργία New Dimension.....	42
Εικόνα 40: Dimension Wizard.....	42
Εικόνα 41: Select Creation Method.....	43
Εικόνα 42: Specify Source Information.....	43
Εικόνα 43: Select Dimension Attributes.....	44
Εικόνα 44: Ολοκλήρωση Dimension.....	44
Εικόνα 45: Θέση Dimension.....	45
Εικόνα 46: Επιλογή στοιχείων πίνακα Sheet1.....	45
Εικόνα 47: Διαστάσεις Dimension.....	46
Εικόνα 48: Δημιουργία Cube.....	46
Εικόνα 49: Select Creation Method.....	47
Εικόνα 50: Select Measure Group Tables.....	47
Εικόνα 51: Select Measures.....	48
Εικόνα 52: Select Existing Dimensions.....	48
Εικόνα 53: Select New Dimensions.....	49
Εικόνα 54: Unselect Dimension.....	49
Εικόνα 55: Ολοκλήρωση Cube.....	50
Εικόνα 56: Θέση Cube.....	50
Εικόνα 57: Process.....	51
Εικόνα 58: Server Content.....	51
Εικόνα 59: Process Cube.....	52

Εικόνα 60: Ολοκλήρωση Process.....	52
Εικόνα 61: New Mining Structure.....	53
Εικόνα 62: Data Mining Wizard.....	53
Εικόνα 63: Select the Definition Method.....	54
Εικόνα 64: Create the Data Mining Structure.....	54
Εικόνα 65: Select Data Source View.....	55
Εικόνα 66: Specify Table Types.....	55
Εικόνα 67: Specify the Training Data.....	56
Εικόνα 68: Suggest Related Columns.....	56
Εικόνα 69: Specify the Training Data.....	57
Εικόνα 70: Specify Columns Content and Data Type.....	57
Εικόνα 71: Specify Columns Content and Data Type.....	58
Εικόνα 72: Create Testing Set.....	58
Εικόνα 73: Ολοκλήρωση Mining.....	59
Εικόνα 74: Θέση StudentInvolvement.dmm.....	59
Εικόνα 75: New Mining Model.....	60
Εικόνα 76: New Mining Model.....	60
Εικόνα 77: Microsoft Clusterin.....	61
Εικόνα 78: Microsoft Neural Network.....	61
Εικόνα 79: Mining Models.....	61
Εικόνα 80: Mining Model Viewer.....	62
Εικόνα 81: Process Mining Model.....	62
Εικόνα 82: Process Progress.....	63
Εικόνα 83: Decision Tree.....	63
Εικόνα 84: Cluster Diagram.....	64
Εικόνα 85: Cluster Profiles.....	64
Εικόνα 86: Cluster Characteristics.....	65
Εικόνα 87: Cluster Discrimination.....	65
Εικόνα 88: Mining Model Viewer.....	65
Εικόνα 89: Mining Accuracy Chart.....	66

Εικόνα 90: Mining Model Prediction.....	66
Εικόνα 91: Test Data.....	67
Εικόνα 92: Select Table.....	67
Εικόνα 93: Φύλλο1.....	68
Εικόνα 94: Select Mining Model.....	68
Εικόνα 95: Συμπλήρωση Field.....	69
Εικόνα 96: Συμπλήρωση Criteria/Argument.....	69
Εικόνα 97: Mining Model Prediction.....	70

ΒΙΒΛΙΟΓΡΑΦΙΑ

- Agasisti, T. (2013). 'The efficiency of Italian secondary schools and the potential role of competition: a data envelopment analysis using OECD-PISA2006 data'. *Education Economics*, 21(5), 520-544
- Andres, J. M. L., Baker, R. S., Siemens, G., Gašević, D., & Spann, C. A. (in press). *Replicating 21 Findings on Student Success in Online Learning*. Technology, Instruction, Cognition, and Learning
- Bach, C. (2010). 'Learning analytics: Targeting instruction, curricula and student support'. Office of the Provost, Drexel University.
- Baker, R. S. (2013). 'Learning, Schooling, and Data Analytics'. In M. Murphy, S. Redding & J. Twyman (Eds.), *Handbook on innovations in learning* Philadelphia, PA: Center on Innovations in Learning, Temple University; Charlotte, NC: Information Age Publishing.
- Baker, R. S., & Inventado, P. S. (2014). 'Educational data mining and learning analytics'. In J. A. Larusson & B. White (Eds.), *Learning Analytics* (pp. 61-75). New York: Springer.
- Baker, R. S., & Yacef, K. (2009). 'The State of Educational Data Mining in 2009: A Review and Future Visions'. *Journal of Educational Data Mining*, 1(1), 3-16
- Bowers, A. J. (2009). 'Reconsidering grades as data for decision making: More than just academic knowledge'. *Journal of Educational Administration*, 47(5), 609-629.
- Bowers, A. J. (2017). 'Quantitative Research Methods Training in Education Leadership and Business Intelligence for Dummies του Scheps (2007).
- Business Intelligence των Sabherwal and Beccera – Fernandez (2010).
- Premise of Learning Analytics for Educational Context: Through Concept to Practice Yasemin GÜLBAHAR, Hale ILGAZ Department of Informatics, Ankara University, Ankara, Turkey Distance Education Center, Ankara University, Ankara, Turkey
(Geliş/Received: 17.03.2014 ; Kabul/Accepted: 12.06.2014)
-
- Carolina Costa et al. / *Procedia Technology* 5 (2012)
- The use of Moodle e-learning platform: a study in a Portuguese University Carolina Costa, Helena Alvelos, Leonor Teixeira
2012 Published by Elsevier Ltd.
- Agasisti, T., Bowers, A.J. (2017) *Data Analytics and Decision-Making in Education: Towards the Educational Data Scientist as a Key Actor in Schools and Higher Education Institutions*. In Johnes, G., Johnes, J., Agasisti, T., López-Torres, L. (Eds.) *Handbook of Contemporary Education Economics* (p.184-210). Cheltenham, UK: Edward Elgar Publishing. ISBN: 978-1-78536-906-3
- Atif, A., Richards. D., Bilgin, A. and Marrone, M. (2013). *A Panorama of Learning Analytics Featuring the Technologies for the Learning and Teaching Domain*. In H. Carter, M. Gosper and J. Hedberg (Eds.), *Electric Dreams. Proceedings ascilite 2013 Sydney*. (pp.68-72)
- Administration Preparation Programs as Disciplined Inquiry for Building School Improvement Capacity. *Journal of Research on Leadership Education*, 12(1), p.72-96. <http://doi.org/10.1177/1942775116659462>.
- Bowers, A. J., Krumm, A. E., Feng, M., & Podkul, T. (2016). 'Building a Data Analytics Partnership to Inform School Leadership Evidence-Based Improvement Cycles'. Paper presented at the Annual meeting of the American Educational Research Association, Washington, DC

- Bowers, A. J., Shoho, A. R., & Barnett, B. G. (2014). 'Considering the Use of Data by School Leaders for Decision Making'. In A. J. Bowers, A. R. Shoho & B. G. Barnett (Eds.), *Using Data in Schools to Inform Leadership and Decision Making* (pp. 1-16). Charlotte, NC: Information Age Publishing.
 - Bowers, A. J., Sprott, R., & Taff, S. (2013). 'Do we know who will drop out? A review of the predictors of dropping out of high school: Precision, sensitivity and specificity'. *The High School Journal*, 96(2), 77-100.
 - Chen, H., Chiang, R. H., & Storey, V. C. (2012). 'Business Intelligence and Analytics: From Big Data to Big Impact'. *MIS Quarterly*, 36(4), 1165-1188.
 - Dawson, S., Gašević, D., Siemens, G., & Joksimovic, S. (2014). 'Current state and future trends: a citation network analysis of the learning analytics field'. Paper presented at the Proceedings of the Fourth International Conference on Learning Analytics And Knowledge, Indianapolis, Indiana, USA.
 - Chung, G. K. W. K. (2014). Toward the Relational Management of Educational Measurement Data. *Teachers College Record*, 116(11), 1-16.
 - Kirkup, C., Sizmur, J., Sturman, L., & Lewis, K. (2005). 'Schools' Use of Data in Teaching and Learning'. Research Report RR671. National Foundation for Educational Research.
 - Romero, C., & Ventura, S. (2007). 'Educational data mining: A survey from 1995 to 2005'. *Expert systems with applications*, 33(1), 135-146.
 - Romero, C., & Ventura, S. (2010). 'Educational data mining: a review of the state of the art'. *Systems, Man, and Cybernetics, Part C: Applications and Reviews*, IEEE Transactions on, 40(6), 601-618.
 - Siemens, G. (2010). 'What are learning analytics'. *ELEARNSPACE: Learning, networks, knowledge, technology, community*.
 - Siemens, G. (2013). 'Learning Analytics: the emergence of a Discipline'. *American Behavioral Scientist*, 57(10), 1380-1400
 - Siemens, G., & Long, P. (2011). 'Penetrating the Fog: Analytics in Learning and Education'. *EDUCAUSE review*, 46(5), 30.
 - Does Self-Discipline impact students' knowledge and learning?
 - Yue Gong, Dovan Rai, Joseph E. Beck and Neil T. Heffernan Computer Science Department, Worcester Polytechnic Institute Educational Data Mining 2009
 - Academic Analytics: A New Tool for a New Era© 2007 John P. Campbell, Peter B. DeBlois, and Diana G. Oblinger *EDUCAUSE Review*, vol. 42, no. 4 (July/August 2007): 40–57
 - Zañane, O. (2001). Web usage mining for a better web-based learning environment. *Proceedings of Conference on Advanced Technology for Education*, 60–64.
 - Romero, C., Ventura, S., de Bra, P., & Castro, C. (2003). Discovering prediction rules in aha! courses. *Proceedings of the International Conference on User Modeling* 25–34.
 - The KDD Process for Extracting Useful Knowledge from Volumes of Data Usama Fayyad, Gregory Piatetsky-Shapiro, and Padhraic Smyth November 1996/Vol. 39, No. 11 COMMUNICATIONS OF THE ACM
 - Handbook of Educational Data Mining Cristobal Romero, Sebastian Ventura, Mykola Pechenizkiy, Ryan S.J.d. Baker CRC Press, 25 Okt 2010
- Enhancing Teaching and Learning Through Educational Data Mining and
- LearningAnalytics: An Issue Brief Marie Bienkowski Mingyu Feng, Barbara Means Center for Technology in Learning SRI International October 2012 U.S. Department of Education Office of Educational Technology

Ιστοθέσεις

- www.information-management.com, ιστοθέση του περιοδικού Information Management | IT Business News (“Information Management | IT Business News,” n.d.).
- www.informationweek.com, ιστοθέση του περιοδικού Information Week (“Information Week,” n.d.).
- tdwi.org, ιστοθέση του Data Warehousing Institute (“TDWI | Advancing all things data,” n.d.). Η ιστοθέση περιέχει πολλά ηλεκτρονικά έγγραφα σχετικά με τη σημασία και την αξία της E.E.
- www.tdan.com, ιστοθέση της ηλεκτρονικής έκδοσης The Data Administration Newsletter (“TDAN. com,” n.d.), που αναφέρεται σε διάφορα θέματα διαχείρισης δεδομένων.
- <http://jedm.educationaldatamining.org/index.php/JEDM>
- <https://moodle.org/>