

Πολυδιάστατη Ανάλυση Δεδομένων

Αντληση δεδομένων από τη βάση

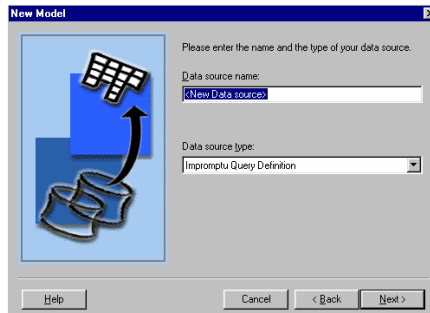
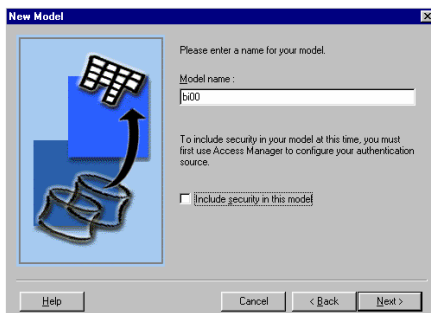
Το πρώτο βήμα είναι η δημιουργία της πολυδιάστατης βάσης δεδομένων (OLAP – On Line Analytical Processing) η οποία απευθύνεται στους καταναλωτές πληροφοριών για πολυδιάστατη ανάλυση δεδομένων.

Η δημιουργία της OLAP βάσης βασίζεται στα υπάρχοντα δεδομένα μιας βάσης δεδομένων (OLTP – On Line Transaction Processing) που προορίζεται για αποθήκευση αναλυτικών δεδομένων λειτουργίας ενός οργανισμού.

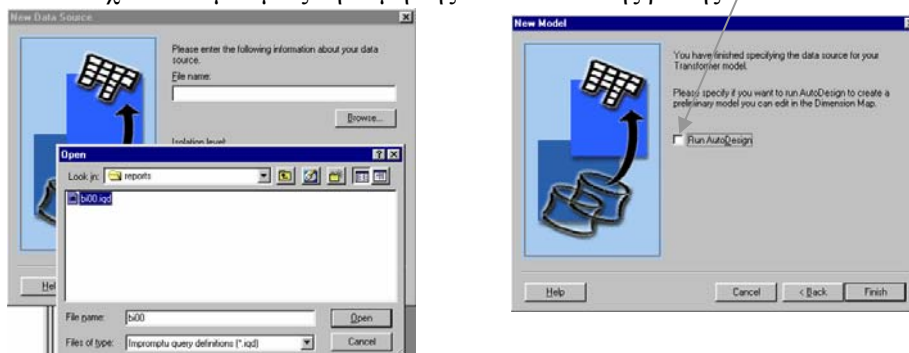
Επιλέγουμε το εικονίδιο PowerPlay Transformer για τη δημιουργία της πολυδιάστατης βάσης (multidimensional cube).



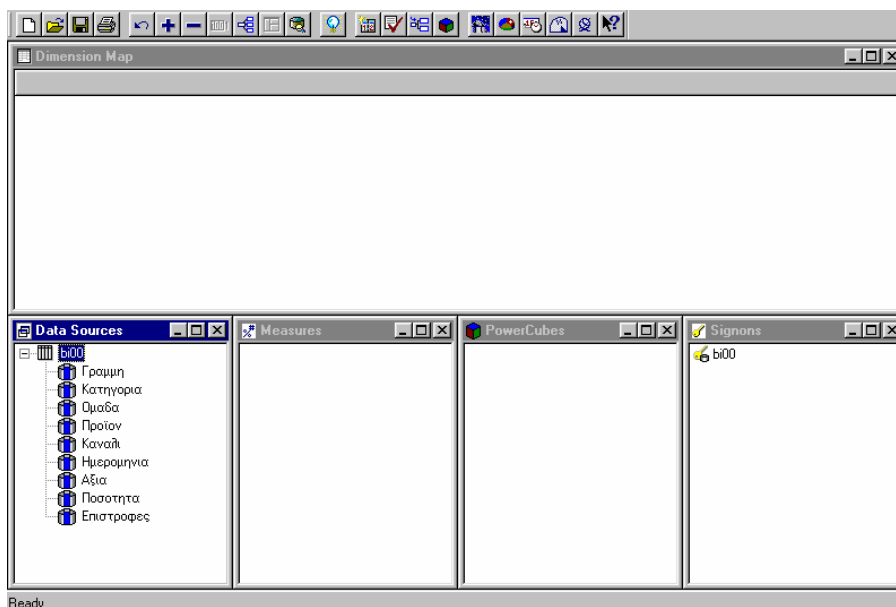
Εμφανίζεται η πρώτη οθόνη και επιλέγουμε File – New για το σχεδιασμό μιας νέας πολυδιάστατης βάσης. Στο πρώτο παράθυρο δίνουμε το όνομα της νέας βάσης biXX και Next. Στο δεύτερο Next.



Στο επόμενο παράθυρο επιλέγουμε Browse, ώστε να αναζητήσουμε το αρχείο υποδομής biXX που δημιουργήσαμε με το Impromptu και Next. Στο επόμενο παράθυρο απενεργοποιούμε την επιλογή Run Autodesign, ώστε να σχεδιάσουμε εμείς τη δομή της πολυδιάστατης βάσης.

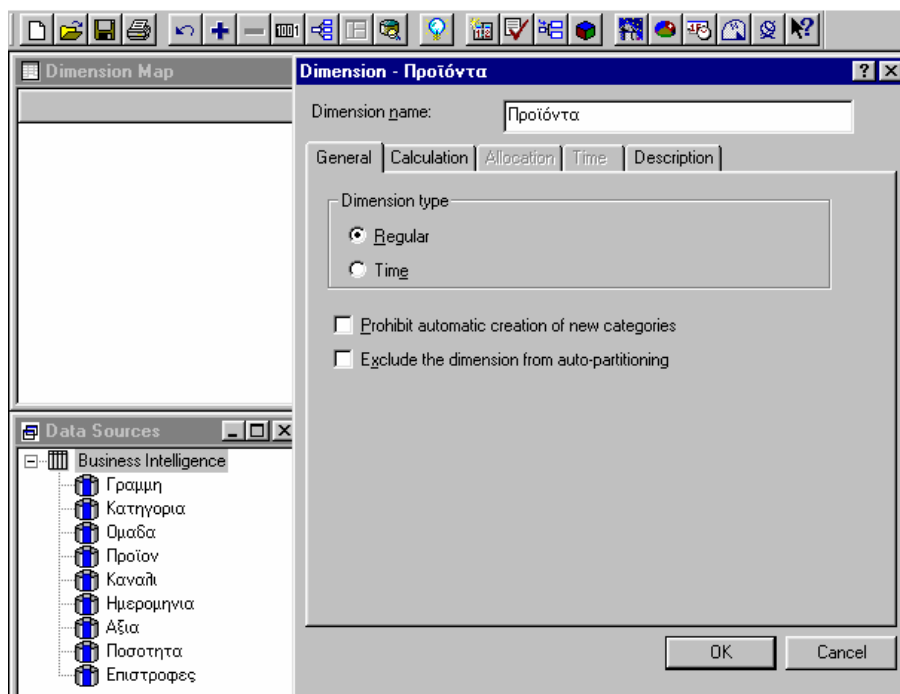


Παρατηρούμε ότι εμφανίζονται τέσσερα παράθυρα (α) των διαστάσεων (dimensions) (β) των πηγών άντλησης (data sources) όπου εμφανίζονται όλα τα πεδία που είχαμε επιλέξει στο biXX, (γ) των τιμών (measures), (δ) των κύβων που θα δημιουργήσουμε (cubes) και (ε) των βάσεων δεδομένων με τις οποίες έχουμε συνδεθεί (signons).



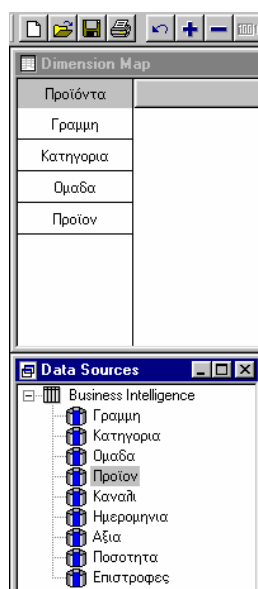
Δημιουργία των διαστάσεων

Τα τέσσερα πρώτα πεδία είναι μια ιεραρχία που καθορίζει τη διάσταση Προϊόντα. Επιλέγουμε το παράθυρο Dimension Map και το κουμπί '+', ώστε να ορίσουμε την πρώτη διάσταση 'Προϊόντα'.



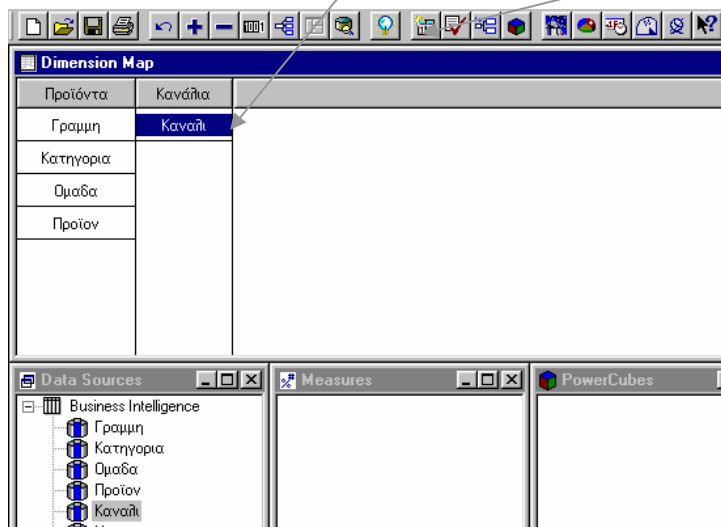
Όταν πατήσουμε OK εμφανίζεται η νέα διάσταση στο παράθυρο Dimension Map. Κατόπιν μεταφέρουμε από το παράθυρο Data Sources ένα - ένα τα πεδία που δομούν την ιεραρχία των προϊόντων, από την κορυφή της ιεραρχίας που είναι οι Γραμμές προϊόντων, μέχρι την χαμηλότερη επιθυμητή ανάλυση που είναι το Προϊόν, με Drag (από το παράθυρο Data Sources) και Drop (στο παράθυρο Dimension Map).

Όπως ορίσαμε τη διάσταση προϊόντα, ορίζουμε τη νέα διάσταση Καναάλια (επιλογή επικεφαλίδας παραθύρου Dimension Map, κουμπί '+' και ορισμός περιγραφής νέας διάστασης).



Μεταφέρουμε το μοναδικό πεδίο ‘Κανάλι’ που ορίζει την ιεραρχία των καναλιών από τα Data Sources στη διάσταση ‘Κανάλια’.

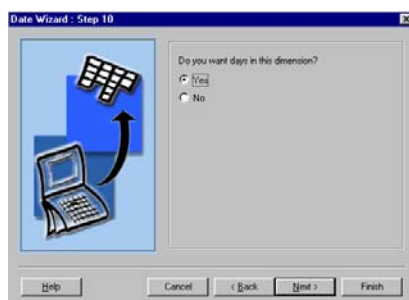
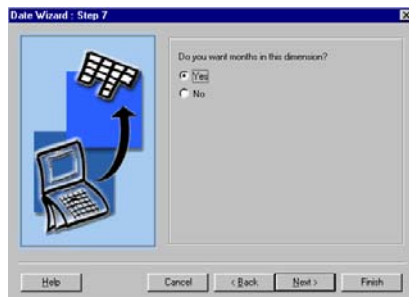
Ορίζουμε την τελευταία διάσταση ‘Ημερομηνία’, επιλέγοντας την επικεφαλίδα του παραθύρου Dimension Map και το κουμπί ‘Date Wizard’.



Στο πρώτο παράθυρο δίνουμε την περιγραφή την νέας διάστασης, στο δεύτερο ορίζουμε το πεδίο, στο τρίτο ότι επιθυμούμε να έχει έτη, στο τέταρτο ότι η βάση θα είναι οι μήνες και όχι οι εβδομάδες.

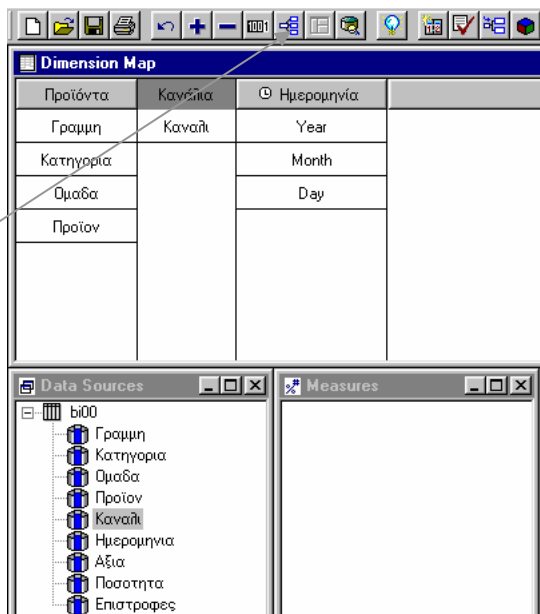


Στο πέμπτο ότι δεν επιθυμούμε να έχει τρίμηνα η διάσταση, στο έκτο ότι θα έχει μήνες, στο έβδομο ότι δεν θα έχει εβδομάδες και στο όγδοο ότι θα έχει ημέρες και Finish.

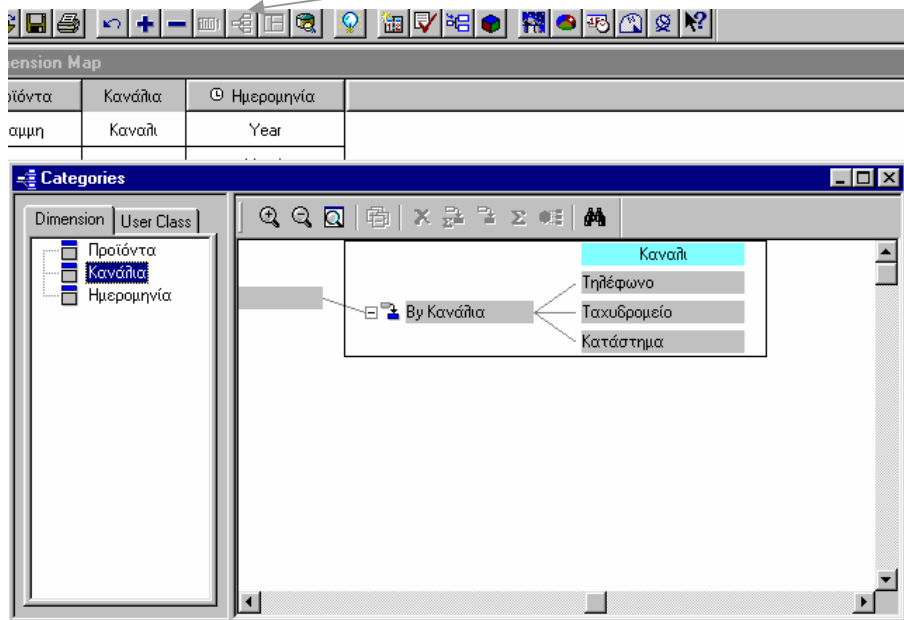


Μετά την ολοκλήρωση εμφανίζεται η διάσταση Ημερομηνία με τα επίπεδα που επιλέξαμε (έτος, μήνας, ημέρα).

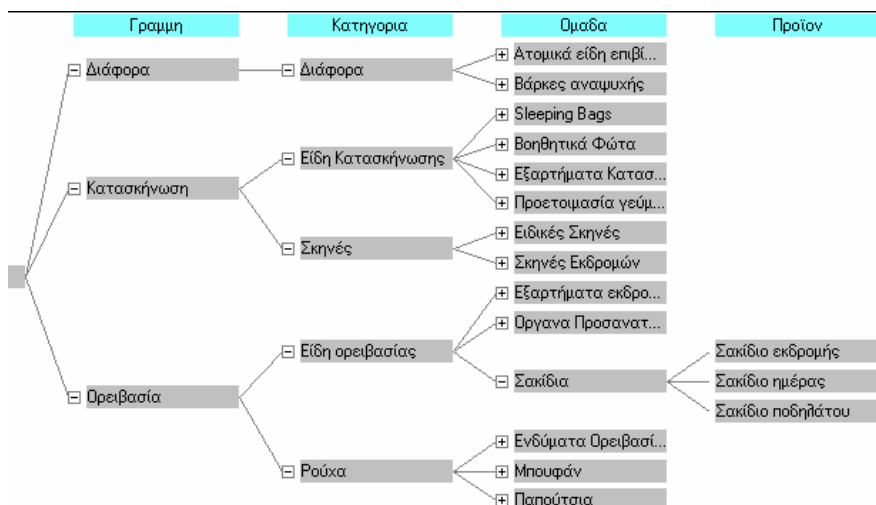
Το τελευταίο βήμα για την ολοκλήρωση των διαστάσεων είναι η δημιουργία των δένδρων κάθε διάστασης που επιτυγχάνεται με το κουμπί Generate Categories.



Μπορούμε να παρατηρήσουμε τη δομή του δένδρου κάθε διάστασης, επιλέγοντας διάσταση (π.χ. Κανάλια) και το κουμπί Show-Diagram. Το αποτέλεσμα είναι η εμφάνιση του δένδρου της διάστασης.

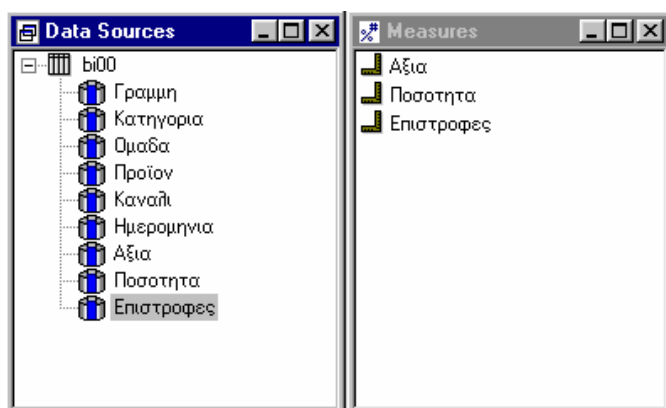


Αντίστοιχα μπορούμε να αναλύσουμε το δένδρο των προϊόντων ή οποιασδήποτε άλλης διάστασης.

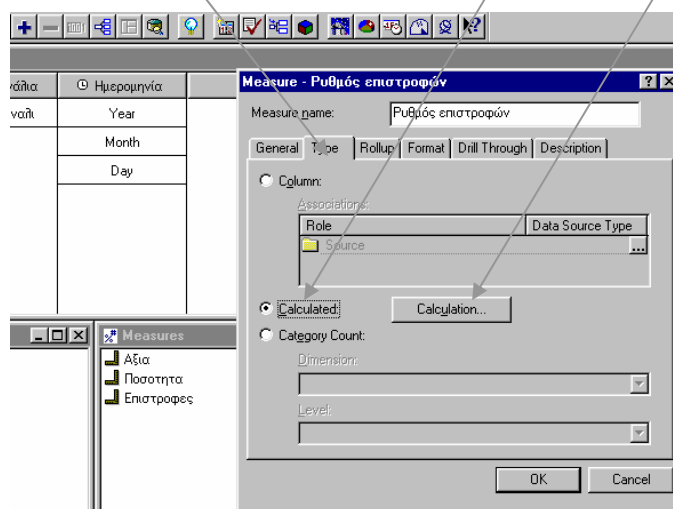


Ορισμός οικονομικών μεγεθών

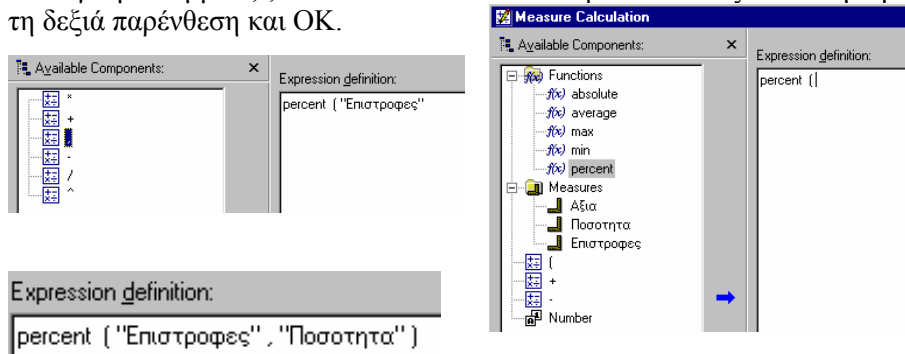
Το τελευταίο βήμα πριν την δημιουργία της πολυδιάστατης βάσης είναι ο ορισμός των οικονομικών μεγεθών. Με drag & drop μεταφέρουμε τα πεδία Αξία, Ποσότητα και Επιστροφές στο παράθυρο Measures.



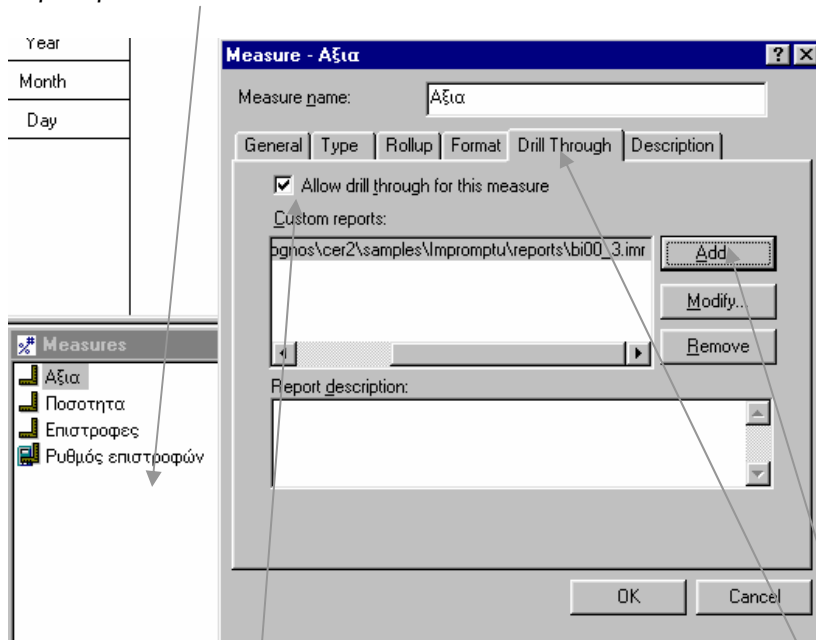
Ένα από τα ζητούμενα του μοντέλου είναι η αξιολόγηση των επιστροφών σε σχέση με την ποσότητα πώλησης. Για το σκοπό αυτό θα ορίσουμε ένα νέο μέγεθος, τον 'ρυθμό επιστροφών'. Επιλέγουμε την επικεφαλίδα του παραθύρου measures και το κουμπί '+', ώστε να προσθέσουμε ένα νέο οικονομικό μέγεθος. Καταχωρούμε το όνομα και στον καρτελοδείκτη Type επιλέγουμε το κουμπί Calculated και Calculation, ώστε να ορίσουμε τρόπο υπολογισμού.



Επιλέγουμε τη συνάρτηση percent και κατόπιν το πεδίο Επιστροφές. Επιλέγουμε κόμμα ',', κατόπιν το πεδίο Ποσότητα και τέλος κλείνουμε με τη δεξιά παρένθεση και OK.

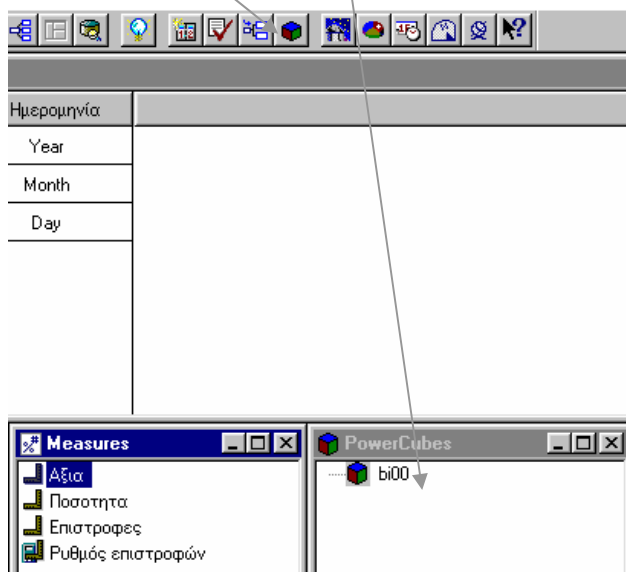


Το αποτέλεσμα είναι η προσθήκη του νέου οικονομικού μεγέθους στο παράθυρο measures.



Η τελευταία απαίτηση είναι η δυνατότητα διερεύνησης των πρωτογενών δεδομένων από την πολυδιάστατη βάση, σαν το τελευταίο επίπεδο ανάλυσης για το μέγεθος αξία. Επιλέγουμε το μέγεθος Αξία στο παράθυρο Measures (2 κλικ) και στον καρτελοδείκτη επιλέγουμε Drill Through, Allow drill through for this measure και με το κουμπί Add αναζητούμε και προσθέτουμε την εκτύπωση biXX_3 που έχουμε ήδη δημιουργήσει. Τέλος επιλέγουμε OK.

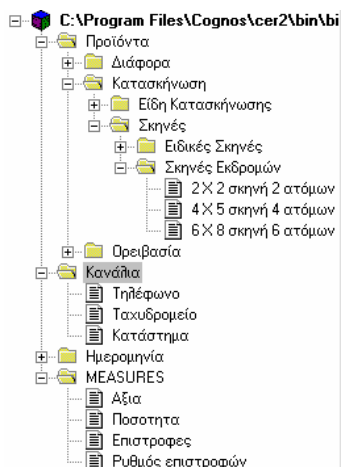
Επόμενο βήμα είναι η δημιουργία της πολυδιάστατης βάσης. Επιλέγουμε το κουμπί Create PowerCubes και μετά από λίγο εμφανίζεται ο νέος κύβος biXX στο παράθυρο PowerCubes.



Ο κύβος έχει ήδη δημιουργηθεί, επομένως δεν χρειάζεται αποθήκευση κατά την έξοδο από το πρόγραμμα.

Η πλοήγηση γίνεται όπως στον MS Explorer και μπορούν να αναλυθούν είτε οι διαστάσεις, είτε τα οικονομικά μεγέθη (measures). Η πλοήγηση αυτή είναι καθαρά πληροφοριακή.

Η πλοήγηση στις διαστάσεις είναι δυναμική. Παρατηρούμε κατ' αρχάς ότι στη γραμμή των διαστάσεων εμφανίζεται το επίπεδο ανάλυσης κάθε διάστασης (στην περίπτωση αυτή το κορυφαίο επίπεδο όλων των διαστάσεων, ήτοι όλα τα προϊόντα, όλα τα κανάλια και όλες οι ημερομηνίες), καθώς και το μέγεθος των υπολογισμών (στην περίπτωση αυτή η αξία).



	Τηλέφωνο	Ταχυδρομείο	Κατάστημα	Κανάλια
Διάφορα	401.739	254.651	183.471	839.861
Κατασκήνωση	5.678.864	4.121.358	14.441.452	24.241.674
Ορειβασία	2.811.827	2.056.795	1.211.534	6.080.156
Προϊόντα	8.892.430	6.432.804	15.836.457	31.161.691

Ένα σημαντικό χαρακτηριστικό των βάσεων πολυδιάστατης ανάλυσης (ΠΑ) είναι η δυνατότητα ανάλυσης (drill-down) ή συγκέντρωσης (drill-up) σε οποιαδήποτε διάσταση και σε οποιοδήποτε βάθος της ιεραρχίας ανά διάσταση. Έτσι μετακινώντας το δρομέα επάνω από τις γραμμές προϊόντων της διάστασης προϊόντα, παρατηρούμε ότι αλλάζει σχήμα και γίνεται συν. Εάν επιλεγεί (δυο κλικ) μια γραμμή π.χ. Κατασκήνωση, τότε η γραμμή αναλύεται στις κατηγορίες από τις οποίες αποτελείται, ενώ η γραμμή προϊόντων (Κατασκήνωση) πηγαίνει στη γραμμή των συνόλων. Παρατηρούμε ότι αλλάζει αυτόματα και η γραμμή των διαστάσεων, ενώ αναπροσαρμόζονται κάθε φορά και οι αξίες.

	Τηλέφωνο	Ταχυδρομείο	Κατάστημα	Κανάλια
Είδη Κατασκήνωσης	2.275.890	1.588.195	1.011.973	4.876.058
Σκηές	3.402.974	2.533.163	13.429.479	19.365.616
Κατασκήνωση	5.678.864	4.121.358	14.441.452	24.241.674

Αυτή η διαδικασία (drill-down) μπορεί να συνεχιστεί μέχρι το χαμηλότερο επίπεδο κάθε διάστασης. Επιστροφή στο προηγούμενο επίπεδο (drill-up) γίνεται αν επιλεγεί (δυο κλικ) η γραμμή των συνόλων, ή στο ανώτερο επίπεδο όλων των διαστάσεων με το κουμπί Reset dimensions.

	Τηλέφωνο	Ταχυδρομείο	Κατάστημα	Κανάλια
Ομαδική σκηνή 20 ατόμων	279.294	194.034	126.420	599.748
Ομαδική σκηνή 50 ατόμων	1.043.088	832.491	509.394	2.384.973
3 X 5 Ελαφριά	51.866	40.229	24.274	116.369
7 X 8 3 Ατόμων Ελαφριά	190.065	141.949	70.974	402.988
8 X 9 4 Ατόμων Ελαφριά	150.139	101.652	48.617	300.408
8 X 9 5 Ατόμων Ελαφριά	147.671	110.712	66.716	325.099
9 X 12 6 Ατόμων Ελαφριά	204.264	150.805	87.117	442.186
9 X 9 5 Ατόμων Ελαφριά	267.019	175.942	95.446	538.407
Ειδικές Σκηνές	2.333.406	1.747.814	1.028.958	5.110.178

Οποιαδήποτε διάσταση μπορεί να αντικατασταθεί με μια άλλη από τη γραμμή των διαστάσεων με τη διαδικασία drag & drop. Έτσι αν σύρουμε από τη γραμμή των διαστάσεων την ημερομηνία και την αφήσουμε κάπου στις επικεφαλίδες των καναλιών, αντικαθίστανται οι στήλες με ημερομηνίες αντί για κανάλια. Αν αναλύσουμε τη χρονιά στο χαμηλότερο επίπεδο (μήνες), παρατηρούμε ότι υπάρχουν πολλοί μήνες με μηδενικές πωλήσεις. Για να εμφανιστούν μόνο οι συνδυασμοί με μη μηδενικές αξίες, επιλέγουμε Explore - Suppress - Zeros - Rows and Columns.

	2004	Ημερομηνία
Ομαδική σκηνή 20 ατόμων	599.748	599.748
Ομαδική σκηνή 50 ατόμων	2.384.973	2.384.973
3 X 5 Ελαφριά	116.369	116.369
7 X 8 3 Ατόμων Ελαφριά	402.988	402.988
8 X 9 4 Ατόμων Ελαφριά	300.408	300.408
8 X 9 5 Ατόμων Ελαφριά	325.099	325.099
9 X 12 6 Ατόμων Ελαφριά	442.186	442.186
9 X 9 5 Ατόμων Ελαφριά	538.407	538.407
Ειδικές Σκηνές	5.110.178	5.110.178

Κατεβαίνουμε στο χαμηλότερο επίπεδο της διάστασης Ημερομηνία (ημέρες). Η οθόνη δείχνει τις πωλήσεις ανά ημέρα για κάθε προϊόν της ομάδας ειδικές σκηνές (ή οποιουδήποτε επιπέδου της διάστασης Προϊόντα).

	2004/01/01	2004/01/04	2004/01/05	2004/01/06	2004/Jan
Ομαδική σκηνή 20 ατόμων	283.704	100.938	105.348	109.758	599.748
Ομαδική σκηνή 50 ατόμων	1.133.988	398.695	414.895	437.395	2.384.973
3 X 5 Ελαφριά	0	31.991	40.829	43.549	116.369
7 X 8 3 Ατόμων Ελαφριά	0	108.951	142.729	151.308	402.988
8 X 9 4 Ατόμων Ελαφριά	0	79.424	107.242	113.742	300.408
8 X 9 5 Ατόμων Ελαφριά	0	83.035	117.113	124.951	325.099
9 X 12 6 Ατόμων Ελαφριά	0	108.566	161.035	172.585	442.186
9 X 9 5 Ατόμων Ελαφριά	0	143.054	190.892	204.461	538.407
Ειδικές Σκηνές	1.417.692	1.054.654	1.280.083	1.357.749	5.110.178

Εάν αντικατασταθεί η διάσταση Προϊόντα με τη διάσταση Κανάλια, η πληροφόρηση είναι διαφορετική, εφ' όσον παρουσιάζονται οι πωλήσεις ανά κανάλι και ημερομηνία.

	2004/01/01	2004/01/04	2004/01/05	2004/01/06	2004/Jan
Τηλέφωνο	626.391	483.969	592.964	630.082	2.333.406
Ταχυδρομείο	486.804	358.684	437.608	464.718	1.747.814
Κατάστημα	304.497	212.001	249.511	262.949	1.028.958
Κανάλια	1.417.692	1.054.654	1.280.083	1.357.749	5.110.178

Παρατηρούμε ότι οι πωλήσεις που εμφανίζονται αφορούν την ομάδα Ειδικές Σκηνές (γραμμή διαστάσεων). Εάν πατηθεί η περιγραφή (Ειδικές Σκηνές) στη γραμμή διαστάσεων, εμφανίζονται τα διάφορα επίπεδα της διάστασης και να επιλεγεί το επιθυμητό (π.χ. Προϊόντα), ώστε να αναλυθούν οι πωλήσεις ανά κανάλι και ημερομηνία για το σύνολο των προϊόντων.

	2004/01/01	2004/01/04	2004/01/05	2004/01/06	2004/Jan
Τηλέφωνο	3.662.716	1.615.273	1.767.440	1.847.001	8.892.430
Ταχυδρομείο	2.663.235	1.168.129	1.276.282	1.325.158	6.432.804
Κατάστημα	1.616.333	698.711	12.742.557	778.856	15.836.457
Κανάλια	7.942.284	3.482.113	15.786.279	3.951.015	31.161.691

Επιλέγοντας Explore – Show values as - % of Column Total, η εικόνα αλλάζει και αντί αξίας, έχουμε κατανομή των πωλήσεων ανά ημέρα ανά κανάλι.

	2004/01/01	2004/01/04	2004/01/05	2004/01/06	2004/Jan
Τηλέφωνο	46,12%	46,39%	11,20%	46,75%	28,54%
Ταχυδρομείο	33,53%	33,55%	8,08%	33,54%	20,64%
Κατάστημα	20,35%	20,07%	80,72%	19,71%	50,82%
Κανάλια	100,00%	100,00%	100,00%	100,00%	100,00%

Επαναφέρουμε τις αξίες αντί ποσοστών (Explore – Show values as – Values) και σύρουμε τη διάσταση Προϊόντα στο αριστερότερο σημείο της γραμμής διαστάσεων. Παρατηρούμε ότι οι διαστάσεις που αναλύονται έγιναν τρεις, η δε διάσταση Προϊόντα μπήκε στον πίνακα σαν σελίδα, οπότε η εναλλαγή των σελίδων παρουσιάζει την ανάλυση του επιπέδου της διάστασης Προϊόντα στο οποίο βρισκόμαστε (στην περίπτωση αυτή γραμμές προϊόντων).

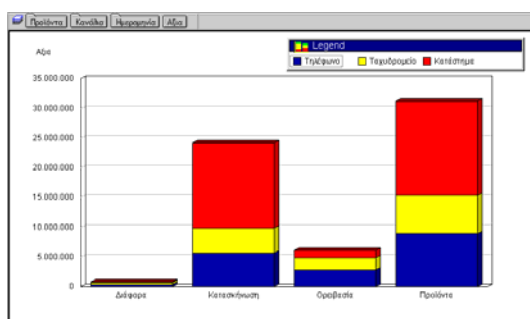
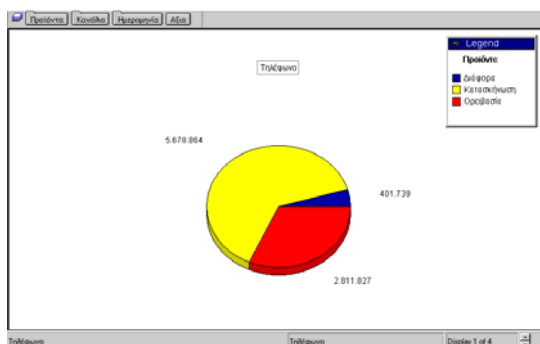
	2004/01/01	2004/01/04	2004/01/05	2004/01/06	2004/Jan
Τηλέφωνο	3.662.716	1.615.273	1.767.440	1.847.001	8.892.430
Ταχυδρομείο	2.663.235	1.168.129	1.276.282	1.325.158	6.432.804
Κατάστημα	1.616.333	698.711	12.742.557	778.856	15.836.457
Κανάλια	7.942.284	3.482.113	15.786.279	3.951.015	31.161.691

Η μέχρι τώρα πλοήγηση έγινε σε πίνακες για καλύτερη κατανόηση της λογικής της πολυδιάστατης ανάλυσης. Μπορεί όμως να γίνει και σε οποιαδήποτε άλλη γραφική μορφή.

	Διάφορα	Κατασκήνωση	Ορειβασία	Προϊόντα
Τηλέφωνο	401.739	5.678.864	2.811.827	8.892.430
Ταχυδρομείο	254.651	4.121.358	2.056.795	6.432.804
Κατάστημα	183.471	14.441.452	1.211.534	15.836.457
Κανάλια	839.861	24.241.674	6.080.156	31.161.691

Επαναφέρουμε την οθόνη στην ανωτέρω μορφή και επιλέγουμε την πίτα (pie chart).

Παρατηρούμε ότι εμφανίζονται οι πωλήσεις για όλα τα προϊόντα με ανάλυση ανά γραμμή και για το κανάλι τηλεφωνο. Η εναλλαγή των καναλιών μπορεί να γίνει κάτω δεξιά, ενώ η διαδικασία ανάλυσης και συγκέντρωσης μπορεί να γίνει και επάνω στο διάγραμμα (δύο κλικ σε οποιοδήποτε σημείο της πίτας, αναλύει στο αμέσως χαμηλότερο επίπεδο).

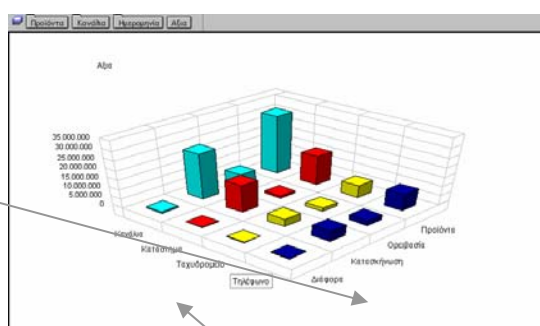


Μπορούμε να επιλέξουμε όποια μορφή διαγράμματος επιθυμούμε και να κάνουμε την πλοήγηση επάνω στα διαγράμματα.

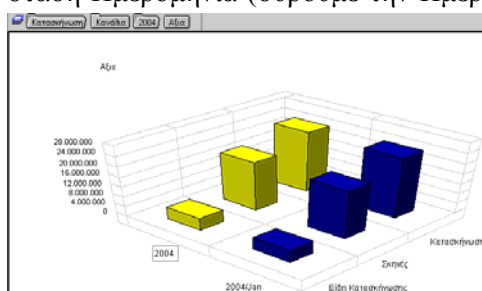
Στο διάγραμμα αυτό παρατηρούμε ότι οι πωλήσεις της γραμμής Κατασκευή

νωση στο κανάλι Κατάστημα είναι ιδιαίτερα υψηλές.

Επιλέγουμε (δύο κλικ) τη γραμμή προϊόντων Κατασκευών, με αποτέλεσμα να κατέβουμε στο χαμηλότερο επίπεδο στη διάσταση Προϊόντα.

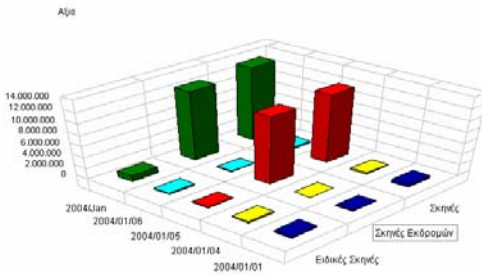


Αλλάζουμε τη διάσταση Κανάλια με τη διάσταση Ημερομηνία (σύρουμε την Ημερομηνία από τη γραμμή διαστάσεων



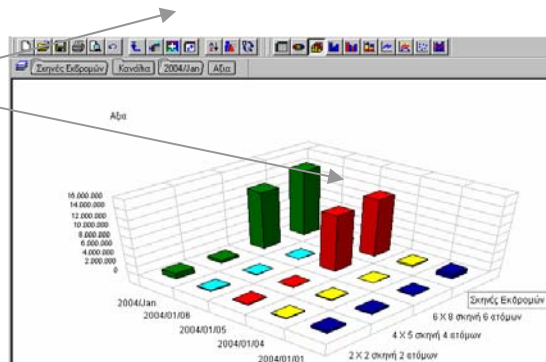
και αφήνουμε επάνω σε οποιοδήποτε σημείο στη διάσταση Κανάλια), ώστε να εξετάσουμε αν η αυξημένη πώληση αφορά κάποια αιχμή. Κατεβαίνουμε ένα επίπεδο στη διάσταση ημερομηνία ώστε να δούμε τις πωλήσεις ανά μήνα.

Κατεβαίνουμε ένα επίπεδο ακόμη και στις δυο διαστάσεις (δυο κλικ στο ιστόγραμμα).



Παρατηρούμε ότι οι υψηλές πωλήσεις αφορούν την ομάδα Σκηνές Εκδρομών, στην ημερομηνία 5/1/2004. Κατεβαίνουμε ένα επίπεδο ακόμη στη διάσταση Προϊόντα (Σκηνές Εκδρομών) ώστε να εντοπίσουμε το συγκεκριμένο προϊόν.

Επιλέγουμε το ιστόγραμμα (κλικ) και κατόπιν το κουμπί Drill Through, ώστε να δούμε τις πρωτογενείς εγγραφές από τη βάση δεδομένων (όπως ορίσαμε για το μέγεθος Αξία στην παράγραφο 'Ορισμός Οικονομικών Μεγεθών').



Παρατηρούμε ότι η κίνηση αφορά μια συγκριτικά πολύ μεγάλη αξία, με χαμηλή ποσότητα (μόλις 4 τεμάχια), συγκριτικά με τις πωλήσεις την ίδια ημέρα στα άλλα κανάλια, επομένως η συγκεκριμένη κίνηση απαιτεί περαιτέρω διερεύνηση. Κλείνουμε το Impromptu και συνεχίζουμε την ανάλυση.

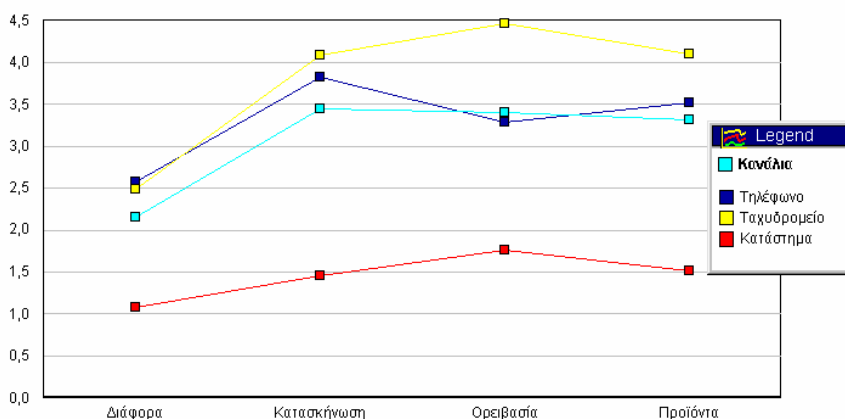
Ημερομηνία	Προϊόν	Κανάλι	Αξία	Ποσότητα
5/1/2004 12:00πμ	6 X 8 σκηνή 6 ατόμων	Κατάστημα	11.999.150,00	4,00
		Ταχυδρομείο	31.516,00	394,00
		Τηλέφωνο	40.955,00	512,00
			12.071.621,00	
			12.071.621,00	
			12.071.621,00	

Πέραν των αξιακών πωλήσεων έχουμε προβλέψει και άλλα οικονομικά μεγέθη, όπως ποσοτικές πωλήσεις, επιστροφές και ρυθμός επιστροφών (επιστροφές / πωλήσεις %).

	Διάφορα	Κατασκήνωση	Ορειβάσια	Προϊόντα
Τηλέφωνο	3	4	3	4
Ταχυδρομείο	2	4	4	4
Κατάστημα	1	1	2	2
Κανάλια	2	3	3	3

Δημιουργούμε τον ανωτέρω πίνακα και επιλέγουμε το γραμμικό διάγραμμα για να αναλύσουμε το ρυθμό επιστροφών ανά κανάλι για όλα τα προϊόντα και όλες τις ημερομηνίες.

Ρυθμός επιστροφών



Παρατηρούμε ότι ο ρυθμός επιστροφών για τις πωλήσεις από το κατάστημα είναι λιγότερο από μισός σε σχέση με τις τηλεφωνικές και τις ταχυδρομικές παραγγελίες, κάτι που είναι αναμενόμενο, εφ' όσον ο πελάτης έχει τη δυνατότητα να δει το προϊόν στο κατάστημα, σε σχέση με τις πωλήσεις μέσω τηλεφώνου και ταχυδρομείου, όπου η επιλογή γίνεται με βάση κάποιον κατάλογο προϊόντων, ή την περιγραφή του προϊόντος στην ιστοσελίδα της επιχείρησης.