

Big Data και Analytics

1η εβδομάδα: Εισαγωγή στα Big Data και στα Analytics

Τμήμα Λογιστικής και Χρηματοοικονομικής

Πλαίσιο της σημερινής συνάντησης

- Πρώτο μέρος: τι είναι τα Big Data και γιατί έχουν στρατηγική σημασία
- Δεύτερο μέρος: τι είναι τα Analytics και πώς στηρίζουν λογιστικές και χρηματοοικονομικές αποφάσεις
- Τρίτο μέρος: παραδείγματα εφαρμογών, εργαλεία του μαθήματος και σύντομη δραστηριότητα με R και Excel

Στόχος

Να τεθεί το εννοιολογικό υπόβαθρο του μαθήματος και να γίνει σαφές γιατί η ανάλυση δεδομένων αποτελεί πλέον βασική δεξιότητα για λογιστές, ελεγκτές και χρηματοοικονομικά στελέχη.

Μαθησιακά αποτελέσματα της 1ης εβδομάδας

Με το τέλος της σημερινής διάλεξης οι φοιτητές αναμένεται να μπορούν:

- να ορίζουν τις έννοιες Big Data και Analytics
- να διακρίνουν τις κύριες μορφές αναλυτικής
- να εξηγούν γιατί τα δεδομένα είναι κρίσιμος πόρος για λογιστική, έλεγχο και χρηματοοικονομική διοίκηση
- να αναγνωρίζουν ενδεικτικές πηγές δεδομένων στις επιχειρήσεις και στις αγορές
- να κατανοούν τον ρόλο της R και του Excel στο μάθημα

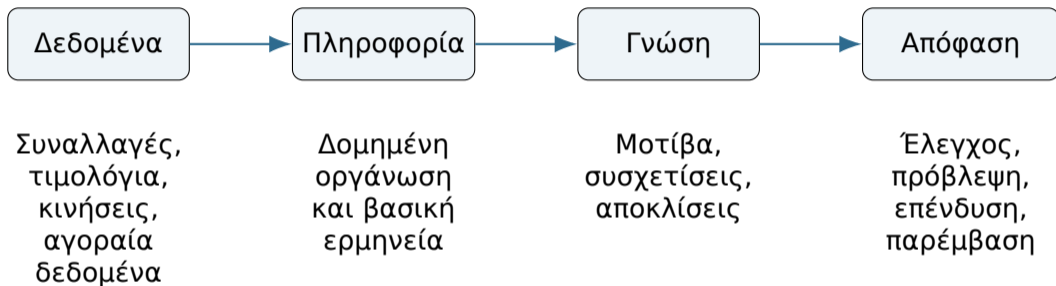
Γιατί αυτό το μάθημα σε τμήμα λογιστικής και χρηματοοικονομικής;

- Οι επιχειρήσεις παράγουν τεράστιο όγκο οικονομικών και λειτουργικών δεδομένων.
- Ο έλεγχος μετακινείται από τη δειγματοληψία προς πιο συστηματική ανάλυση πλήθους συναλλαγών.
- Η χρηματοοικονομική διοίκηση απαιτεί ταχύτερες και τεκμηριωμένες αποφάσεις.
- Η ανίχνευση απάτης, η συμμόρφωση και η διαχείριση κινδύνου εξαρτώνται όλο και περισσότερο από δεδομένα.

Κεντρική ιδέα

Δεν αρκεί πλέον να διαβάζουμε οικονομικές αναφορές. Πρέπει να μπορούμε να εξάγουμε γνώση από δεδομένα μεγάλης κλίμακας και υψηλής πολυπλοκότητας.

Από τα δεδομένα στην απόφαση



Η αναλυτική αξία δεν προκύπτει από την απλή κατοχή δεδομένων αλλά από τη συστηματική μετατροπή τους σε χρήσιμη γνώση.

Λειτουργικός ορισμός

Ως Big Data νοούνται σύνολα δεδομένων τόσο μεγάλα, τόσο γρήγορα παραγόμενα ή τόσο ποικίλα, ώστε οι παραδοσιακές μέθοδοι αποθήκευσης, επεξεργασίας και ανάλυσης να μην επαρκούν.

- Δεν είναι μόνο θέμα όγκου.
- Είναι επίσης θέμα ταχύτητας, μορφής, ποιότητας και αξίας.
- Το κρίσιμο ερώτημα δεν είναι «πόσα δεδομένα έχουμε;» αλλά «τι μπορούμε να εξαγάγουμε από αυτά;»

Τα βασικά χαρακτηριστικά των Big Data

Τα 5V

- Volume: όγκος
- Velocity: ταχύτητα παραγωγής και ροής
- Variety: ποικιλία μορφών
- Veracity: αξιοπιστία και ποιότητα
- Value: επιχειρησιακή αξία

Ερμηνεία για τον χώρο μας

- εκατομμύρια εγγραφές συναλλαγών
- ροές πληρωμών σε πραγματικό χρόνο
- πίνακες, αρχεία, κείμενα, emails, reports
- σφάλματα, ελλείψεις και ασυνέπειες
- ανάγκη μετατροπής σε απόφαση

Τι δεν είναι Big Data

- Δεν είναι απλώς ένα μεγάλο αρχείο Excel.
- Δεν είναι μόνο ένα τεχνολογικό θέμα πληροφορικής.
- Δεν είναι αυτομάτως χρήσιμο επειδή είναι ογκώδες.
- Δεν αντικαθιστά την επαγγελματική κρίση του λογιστή ή του αναλυτή.

Κρίσιμο σημείο

Χωρίς σωστή διατύπωση ερωτήματος, καθαρά δεδομένα και ορθή ερμηνεία, ακόμη και το πιο εντυπωσιακό σύνολο δεδομένων δεν παράγει αξία.

Παραδείγματα Big Data σε επιχειρησιακό περιβάλλον

- δεδομένα πωλήσεων από χιλιάδες συναλλαγές ανά ημέρα
- κινήσεις τραπεζικών λογαριασμών και ηλεκτρονικές πληρωμές
- εγγραφές λογιστικών ημερολογίων και εγκεκριμένες δαπάνες
- στοιχεία πελατών, προμηθευτών και συμβολαίων
- ροές αγοράς, τιμές, ειδήσεις και δεδομένα κοινωνικών δικτύων
- αρχεία ελέγχου από πληροφοριακά συστήματα

Συμπέρασμα

Στον λογιστικό και χρηματοοικονομικό χώρο τα δεδομένα είναι πολλαπλά, διασυνδεδεμένα και συχνά διαχέονται σε διαφορετικά συστήματα.

Τι είναι τα Analytics;

Λειτουργικός ορισμός

Analytics είναι το σύνολο των μεθόδων με τις οποίες συλλέγουμε, οργανώνουμε, αναλύουμε και ερμηνεύουμε δεδομένα ώστε να απαντήσουμε σε επιχειρησιακά ερωτήματα και να στηρίξουμε αποφάσεις.

- συνδυάζει στατιστική, επιχειρησιακή λογική και τεχνολογία
- ξεκινά από ερώτημα και όχι από εργαλείο
- οδηγεί από την περιγραφή στην πρόβλεψη και στη δράση

Οι βασικές κατηγορίες Analytics

Descriptive

Τι συνέβη;

Diagnostic

Γιατί συνέβη;

Predictive

Τι είναι πιθανό
να συμβεί;

Prescriptive

Τι πρέπει να
κάνουμε;

Η αλυσίδα αυτή αποτελεί τον βασικό άξονα του μαθήματος για όλο το εξάμηνο.

Παράδειγμα: ανάλυση καθυστερημένων απαιτήσεων

Τύπος αναλυτικής	Ερώτημα	Ενδεικτικό αποτέλεσμα
Descriptive	Πόσες απαιτήσεις καθυστέρησαν;	Ποσοστό ληξιπρόθεσμων ανά μήνα
Diagnostic	Γιατί αυξήθηκαν;	Συγκέντρωση σε συγκεκριμένους πελάτες ή κλάδους
Predictive	Ποιες απαιτήσεις κινδυνεύουν να μη εισπραχθούν;	Πιθανότητα καθυστέρησης ή αθέτησης
Prescriptive	Ποια ενέργεια είναι καταλληλότερη;	Ιεράρχηση είσπραξης, αναθεώρηση πιστωτικής πολιτικής

Πού χρησιμοποιούνται τα Analytics στη λογιστική;

- ανάλυση εγγραφών ημερολογίου και εντοπισμός εξαιρέσεων
- συμφωνίες και εντοπισμός αποκλίσεων
- κοστολόγηση και ανάλυση περιθωρίων
- ανάλυση απαιτήσεων, υποχρεώσεων και ταμειακών ροών
- υποστήριξη διοικητικής λογιστικής και προϋπολογισμών

Κεντρικό μήνυμα

Η σύγχρονη λογιστική δεν περιορίζεται στην καταγραφή. Προχωρά στην ανάλυση, στην έγκαιρη προειδοποίηση και στη στήριξη της απόφασης.

Πού χρησιμοποιούνται τα Analytics στη χρηματοοικονομική;

- πρόβλεψη πωλήσεων και ταμειακών ροών
- αξιολόγηση πιστωτικού κινδύνου
- αποτίμηση πελατών και χαρτοφυλακίων
- ανάλυση απόδοσης και αποκλίσεων
- υποστήριξη επενδυτικών αποφάσεων

Ενδεικτικά ερωτήματα

Ποιος πελάτης ενδέχεται να καθυστερήσει;
Ποιο τμήμα υπερβαίνει συστηματικά τον προϋπολογισμό;
Πού υπάρχει αυξημένος λειτουργικός κίνδυνος;

Πού χρησιμοποιούνται τα Analytics στον έλεγχο;

- ανάλυση πλήρους πληθυσμού συναλλαγών και όχι μόνο δείγματος
- εντοπισμός ύποπτων ή μη συνήθων λογιστικών εγγραφών
- έλεγχος διπλών πληρωμών, παρακάμψεων έγκρισης και ασυνεπειών
- συνεχής παρακολούθηση βασικών δεικτών και εξαιρέσεων

Γιατί είναι σημαντικό;

Τα analytics ενισχύουν την εμπέλεια, την ταχύτητα και την ποιότητα του ελέγχου, αλλά δεν υποκαθιστούν την τεκμηρίωση και την επαγγελματική κρίση.

Από πού προέρχονται τα δεδομένα;

Εσωτερικές πηγές

- ERP
- λογιστικά συστήματα
- μισθοδοσία
- CRM
- αγορές και αποθήκη
- αρχεία καταγραφής συστημάτων

Εξωτερικές πηγές

- τραπεζικά δεδομένα
- δεδομένα αγοράς
- επιτόκια και συναλλαγματικές ισοτιμίες
- ειδησεογραφία και αναφορές αναλυτών
- δημόσια μητρώα και κανονιστικά δεδομένα

ERP - Enterprise Resource Planning

Ένα **ERP** (Enterprise Resource Planning) σύστημα, στα ελληνικά **Σύστημα Επιχειρησιακού Σχεδιασμού**, είναι ένα λογισμικό που χρησιμοποιούν οι επιχειρήσεις για να διαχειρίζονται και να ενοποιούν τα καθημερινά τους λειτουργικά τμήματα.

Ένα ολοκληρωμένο ERP συνήθως περιλαμβάνει "modules" (ενότητες) για:

- 1 **Οικονομική Διαχείριση:** Λογιστήριο, καθολικό, πληρωτέες/απαιτητές λογαριασμοί.
- 2 **Διαχείριση Αποθήκης:** Αποθέματα, εφοδιαστική αλυσίδα, παραγγελίες.
- 3 **Πωλήσεις & Marketing:** Διαχείριση πελατών (CRM), παραστατικά πώλησης.
- 4 **Αγορές:** Διαχείριση προμηθευτών και κοστολόγηση.
- 5 **Ανθρώπινο Δυναμικό:** Μισθοδοσία, παρουσίες, προσλήψεις.
- 6 **Παραγωγή:** Σχεδιασμός παραγωγής, κοστολόγηση προϊόντων.

CRM - Customer Relationship Management,

CRM σημαίνει Customer Relationship Management, και στα ελληνικά **Διαχείριση Πελατειακών Σχέσεων**.

Ενώ το ERP κοιτάζει όλη την επιχείρηση (λογιστήριο, αποθήκη, προσωπικό), το CRM επικεντρώνεται αποκλειστικά στις πωλήσεις και τους πελάτες.

Είναι ένα λογισμικό που αποθηκεύει όλες τις πληροφορίες για τους πελάτες και τις επαφές σου σε ένα μέρος, ώστε να μην χαθούν ποτέ ξανά σε email, χαρτάκια post-it ή στο μυαλό ενός πωλητή.

Η βασική διαφορά: ERP vs CRM

Το ERP είναι ο "εγκέφαλος της επιχείρησης" για τα οικονομικά και τις λειτουργίες. Το CRM είναι η "μνήμη της ομάδας πωλήσεων".

Το ERP αφορά:

- 1 Τι αγόρασε ο πελάτης,
- 2 πόσο πλήρωσε,
- 3 τι χρωστάει,
- 4 τι έχει η αποθήκη, κ.λπ.

Το CRM αφορά:

- 1 Ποιος μίλησε τελευταία φορά με τον πελάτη,
- 2 τι είπαν,
- 3 πότε πρέπει να τον ξαναπάρουν τηλέφωνο,
- 4 αν του αρέσει το ένα προϊόν ή το άλλο, κ.λπ.

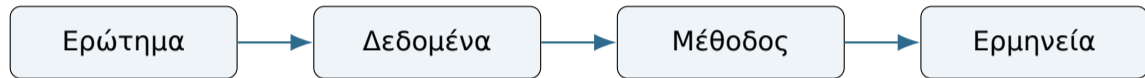
Τα δεδομένα δεν είναι πάντα καθαρά

- ελλιπείς τιμές
- λάθη καταχώρισης
- διπλοεγγραφές
- διαφορετικές ονομασίες για το ίδιο αντικείμενο
- χρονικές ασυνέπειες

Άμεση συνέπεια

Η ποιότητα των αποτελεσμάτων εξαρτάται άμεσα από την ποιότητα των δεδομένων. Το μάθημα δεν θα περιοριστεί σε τεχνικές ανάλυσης αλλά θα καλύψει και την προετοιμασία και αξιολόγηση των δεδομένων.

Το αναλυτικό ερώτημα προηγείται του εργαλείου



- Πρώτα ορίζουμε τι θέλουμε να απαντήσουμε.
- Μετά εντοπίζουμε τα κατάλληλα δεδομένα.
- Στη συνέχεια επιλέγουμε μέθοδο και εργαλείο.
- Τέλος ερμηνεύουμε τα αποτελέσματα σε επιχειρησιακό πλαίσιο.

Ερώτημα

Γιατί μειώθηκε το περιθώριο κέρδους το τελευταίο τρίμηνο;

- Τι δεδομένα χρειαζόμαστε; Πωλήσεις, κόστος, εκπτώσεις, επιστροφές, παραγωγική απόδοση.
- Τι αναζητούμε; Μεταβολές σε τιμές, μίγμα προϊόντων, κόστος πρώτων υλών, πελατειακή σύνθεση.
- Τι θα κάναμε με analytics; Ανάλυση αποκλίσεων, συσχετίσεις, τμηματοποίηση, οπτικοποίηση τάσεων.

Μικρή μελέτη περίπτωσης: ανίχνευση απάτης σε πληρωμές

- διπλές πληρωμές προς τον ίδιο προμηθευτή
- πληρωμές εκτός συνήθους ωραρίου
- ασυνήθιστα ποσά κάτω από όριο έγκρισης
- ασυμφωνία μεταξύ τιμολογίου και παραγγελίας

Αναλυτική λογική

Τα δεδομένα βοηθούν να εντοπίζουμε ενδείξεις. Η επιβεβαίωση απαιτεί έλεγχο, τεκμηρίωση και κρίση.

Ο ρόλος της R και του Excel στο μάθημα

Εργαλείο	Ρόλος στο μάθημα	Πλεονέκτημα
Excel	πρώτη επαφή με δεδομένα, βασική οργάνωση, φίλτρα, απλές συναρτήσεις, αρχική οπτικοποίηση	οικείο και άμεσο για επιχειρησιακή χρήση
R	αναπαραγωγίμη ανάλυση, στατιστική επεξεργασία, οπτικοποίηση, μοντελοποίηση	ισχυρό, ευέλικτο και κατάλληλο για μεγαλύτερη κλίμακα

Το Excel αρκεί για απλές εργασίες. Η R απαιτεί μεγαλύτερη εξοικείωση αλλά προσφέρει πολύ μεγαλύτερη αναλυτική δύναμη.

Μια πρώτη γεύση από R

```
# Ανάγνωση δεδομένων
sales <- read.csv("sales.csv")

# Γενική εικόνα του dataset
summary(sales)

# Συνολικές πωλήσεις ανά μήνα
aggregate(amount ~ month, data = sales, sum)

# Απλό γράφημα
plot(sales$month, sales$amount)
```

Τι δείχνει αυτό το παράδειγμα;

Ότι από την πρώτη στιγμή μπορούμε να μετακινηθούμε από ένα αρχείο δεδομένων σε μια βασική περιγραφική ανάλυση και μια πρώτη οπτική αποτύπωση.

Πού τελειώνει το Excel και πού αρχίζει η R;

Το Excel είναι χρήσιμο όταν

- τα δεδομένα είναι σχετικά περιορισμένα
- χρειάζεται γρήγορος έλεγχος ή βασική αναφορά
- θέλουμε άμεση οπτικοποίηση

Η R είναι προτιμότερη όταν

- ο όγκος και η πολυπλοκότητα αυξάνονται
- απαιτείται επαναληψιμότητα και αυτοματοποίηση
- χρειαζόμαστε στατιστική ή προβλεπτική ανάλυση

Δομή του μαθήματος σε επίπεδο εξαμήνου

- εβδομάδες 1-4: έννοιες, πηγές δεδομένων, ποιότητα και υποδομές
- εβδομάδες 5-6: καθαρισμός, προεπεξεργασία και οπτικοποίηση
- εβδομάδες 7-9: descriptive, diagnostic, predictive analytics και εισαγωγή στο machine learning
- εβδομάδες 10-12: εφαρμογές σε έλεγχο, απάτη, κίνδυνο και χρηματοοικονομική διοίκηση
- εβδομάδα 13: ηθική, περιορισμοί και παρουσιάσεις

Πρακτική κατεύθυνση

Το μάθημα είναι εφαρμοσμένο. Η θεωρία θα συνοδεύεται σταθερά από δεδομένα, ασκήσεις και μελέτες περίπτωσης.

Ενδεικτική κατανομή των 3 ωρών

Χρόνος	Θεματική ενότητα	Διδακτική λογική
1η ώρα	Big Data: έννοιες, χαρακτηριστικά, πηγές	θεωρία και συζήτηση παραδειγμάτων
2η ώρα	Analytics: κατηγορίες, ερωτήματα, εφαρμογές	σύνδεση με λογιστική, έλεγχο και χρηματοοικονομική
3η ώρα	Εργαλεία, μικρή μελέτη περίπτωσης, δραστηριότητα και ανακεφαλαίωση	πρώτη επαφή με R/Excel και εφαρμογή

Σενάριο

Μια επιχείρηση εμφανίζει μείωση λειτουργικής κερδοφορίας και αύξηση καθυστερημένων απαιτήσεων μέσα σε δύο τρίμηνα.

- Ποια δεδομένα θα ζητούσατε πρώτα;
- Ποια ερωτήματα θα θέτατε στη διοίκηση;
- Ποιες μορφές analytics θα ήταν χρήσιμες;
- Τι θα μπορούσε να αναλυθεί αρχικά σε Excel και τι αργότερα σε R;

Στόχος της δραστηριότητας είναι να συνδεθεί αμέσως η θεωρία με πραγματικό πρόβλημα από τον χώρο της επιχείρησης.

Τρία συχνά λάθη στην αρχή της ανάλυσης

- 1 Ξεκινάμε από το εργαλείο αντί από το ερώτημα.
- 2 Θεωρούμε τα δεδομένα αξιόπιστα χωρίς έλεγχο ποιότητας.
- 3 Ερμηνεύουμε συσχέτιση ως αιτιότητα χωρίς προσοχή.

Για το υπόλοιπο εξάμηνο

Η αναλυτική ωριμότητα δεν σημαίνει μόνο τεχνική δεξιότητα. Σημαίνει σωστή διατύπωση ερωτημάτων, πειθαρχημένη επεξεργασία και υπεύθυνη ερμηνεία.

Τι πρέπει να κρατήσουμε από τη σημερινή συνάντηση

- Τα Big Data είναι ευρύτερη έννοια από τον απλό μεγάλο όγκο δεδομένων.
- Τα Analytics μετατρέπουν τα δεδομένα σε γνώση και την γνώση σε απόφαση.
- Η λογιστική, ο έλεγχος και η χρηματοοικονομική μετασχηματίζονται από την αναλυτική.
- Η R θα είναι το βασικό εκπαιδευτικό εργαλείο του μαθήματος, με συμπληρωματική χρήση του Excel.
- Το κρίσιμο ζητούμενο δεν είναι μόνο η τεχνική, αλλά η ορθή επιχειρησιακή ερμηνεία.

Προετοιμασία για την επόμενη εβδομάδα

- Να εγκατασταθεί η R και, αν είναι δυνατόν, το RStudio.
- Να υπάρχει βασική εξοικείωση με αρχεία csv και φύλλα Excel.
- Να εντοπιστούν παραδείγματα πηγών δεδομένων από λογιστικό ή χρηματοοικονομικό περιβάλλον.

Επόμενο μάθημα

Πηγές και κατηγορίες δεδομένων στη λογιστική και στη χρηματοοικονομική ανάλυση.

- Provost, F. and Fawcett, T., Data Science for Business.
- Shmueli, G. et al., Data Mining for Business Analytics.
- Appelbaum, D., Kogan, A. and Vasarhelyi, M., αρθρογραφία για accounting data analytics και audit analytics.

Ευχαριστώ