

ΠΑΝΕΠΙΣΤΗΜΙΟ
ΔΥΤΙΚΗΣ ΜΑΚΕΔΟΝΙΑΣ

ΠΜΣ Ηλεκτρονικό Επιχειρείν και Ψηφιακό Μάρκετινγκ

Τμήμα Διοικητικής Επιστήμης και Τεχνολογίας



M.Sc. in e-Business and
Digital Marketing
Π.Μ.Σ. στο Ηλεκτρονικό Επιχειρείν
και Ψηφιακό Μάρκετινγκ



ΤΜΗΜΑ
ΔΙΟΙΚΗΤΙΚΗΣ
ΕΠΙΣΤΗΜΗΣ &
ΤΕΧΝΟΛΟΓΙΑΣ
ΠΑΝΕΠΙΣΤΗΜΙΟ ΔΥΤΙΚΗΣ ΜΑΚΕΔΟΝΙΑΣ

Έρευνα Αγοράς και Επιστήμη Δεδομένων Ι

ΔΙΔΑΣΚΩΝ

ΔΡ. ΚΩΝΣΤΑΝΤΙΝΟΣ ΠΑΝΥΤΣΙΔΗΣ

Who ?



Δρ. Πανυτσίδης Κωνσταντίνος
Dr. Panitsidis Konstantinos

```
In [ ]: Panitsidis = 'Πανυτσίδης'  
Panytsidis = 'Πανυτσίδης'  
Panitsidis == Panytsidis  
True
```



ΠΑΝΕΠΙΣΤΗΜΙΟ
ΔΥΤΙΚΗΣ ΜΑΚΕΔΟΝΙΑΣ

Panytsidis

Where from ?



Population (2011)^[1]

• Municipality	71,388
• Municipality density	67/km ² (170/sq mi)
• Municipal unit	53,880
• Municipal unit density	150/km ² (380/sq mi)



```
In [ ]: Greece = ['Athens', 'Thessaloniki', 'Kozani']
        'Kozani' in Greece
        True
```

Πανεπιστήμιο Δυτικής Μακεδονίας



ΠΑΝΕΠΙΣΤΗΜΙΟ
ΔΥΤΙΚΗΣ ΜΑΚΕΔΟΝΙΑΣ

Το Πανεπιστήμιο Εκπαίδευση Διοίκηση Έρευνα Επικαιρότητα Επικοινωνία

- Στρατηγικές ρύθμισης των συναισθημάτων για παιδιά και εφήβους
- Πρακτικές συμβουλές για τη διαχείριση του άγχους για τον Covid-19
- Ερωτήσεις-Απαντήσεις για τη νόσο COVID-19 στην Ελληνική Νοηματική Γλώσσα, σε BRAILLE READY και σε Ηχητική μορφή



```
In [ ]: UoWM = {'est':2004, 'dept':7}
        UoWm ['est']
        2004
```

Εισαγωγή

- › Το πλήθος των διαθέσιμων δεδομένων είναι τεράστιο και αυξάνεται εκθετικά κάθε μέρα.
- › Ή μείωση στο κόστος συλλογής και της δυσκολίας στη συλλογή και αποθήκευση των δεδομένων συνετέλεσε σημαντικά στην ανάπτυξή του.

Η Επιστήμη των Δεδομένων

- › Η Επιστήμη των Δεδομένων (**Data Science**), είναι ένας καινούριος όρος, ο οποίος ήρθε να αντικαταστήσει προγενέστερους όρους:
- › Ανακάλυψη Γνώσης από Βάσεις Δεδομένων (**Knowledge Discovery in Data- base**) ή Εξόρυξη Δεδομένων (**Data Mining**).

Η Επιστήμη των Δεδομένων

- › Ένα από τα σημαντικότερα θέματα από πλευράς ουσίας στην Επιστήμη των Δεδομένων αφορά τόσο την εύρεση, όσο και τον χαρακτηρισμό αυτών των **προτύπων**, αποτελούν δομές ή περιστατικά, που εμφανίζονται στα δεδομένα και έχουν κάποια ιδιαίτερη σημασία από **στατιστικής πλευράς**.

Πρότυπο

- › Αναφέρεται και ως μία **διευθέτηση** ή αλλιώς **διάταξη**, στην οποία θεωρείται ότι υπάρχει κάποιου είδους οργάνωση της υποκείμενης δομής.
- › Τα πρότυπα είναι γνωστά και ως **μοντέλα**, ανιχνεύονται κατά κύριο λόγο με τη χρήση μετρήσιμων χαρακτηριστικών γνωρισμάτων ή ιδιοτήτων.

OLAP (OnLine Analytical Processing)

- › Εργαλεία εξερεύνησης δεδομένων, οδηγούμενα από την οπτικοποίηση των αποτελεσμάτων
- › Ή αυτοματοποίηση της διαδικασίας ανακάλυψης προτύπων έλαβε χώρα μέσω τεχνικών μεθοδολογιών και εργαλείων, τα οποία αναπτύχθηκαν στο πλαίσιο της **Επιστήμης των Δεδομένων**

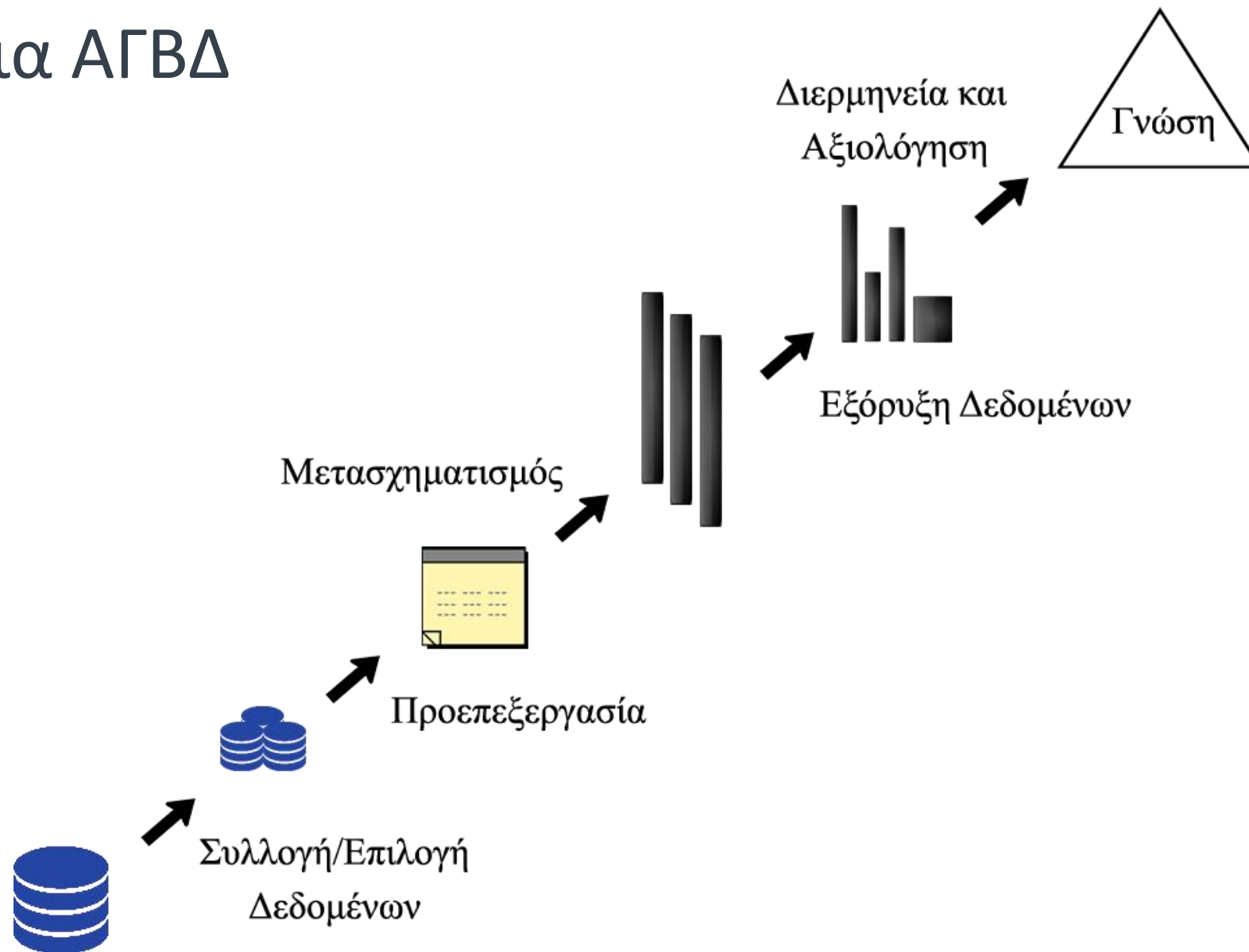
Στόχοι στην πρακτική της Επιστήμης των Δεδομένων

- › η δημιουργία μοντέλων, τα οποία να μπορούν να χρησιμοποιηθούν τόσο για την **πρόβλεψη**, όσο και για την **περιγραφή των δεδομένων**.

Στάδια ΑΓΒΔ

1. Συλλογή Δεδομένων (Data Collection)
2. Προεπεξεργασία Δεδομένων (Preprocessing)
3. Μετασχηματισμός Δεδομένων (Transformation)
4. Εξόρυξη Δεδομένων (Data Mining)
5. Διερμηνεία και Αξιολόγηση (Interpretation/Evaluation)

Στάδια ΑΓΒΔ



Εξόρυξη Δεδομένων

- › Ο βασικός στόχος της **Εξόρυξης Δεδομένων (ΕΔ)** είναι η εξαγωγή μη τετριμμένης, προηγούμενα άγνωστης και πιθανά χρήσιμης πληροφορίας ή προτύπων από το σύνολο των δεδομένων

Συλλογή Δεδομένων

- › Η συλλογή των δεδομένων γίνεται είτε αυτόματα, π.χ. με χρήση αισθητήρων, είτε μη αυτόματα, π.χ. με χρήση ερωτηματολογίων.



Συλλογή/Επιλογή
Δεδομένων

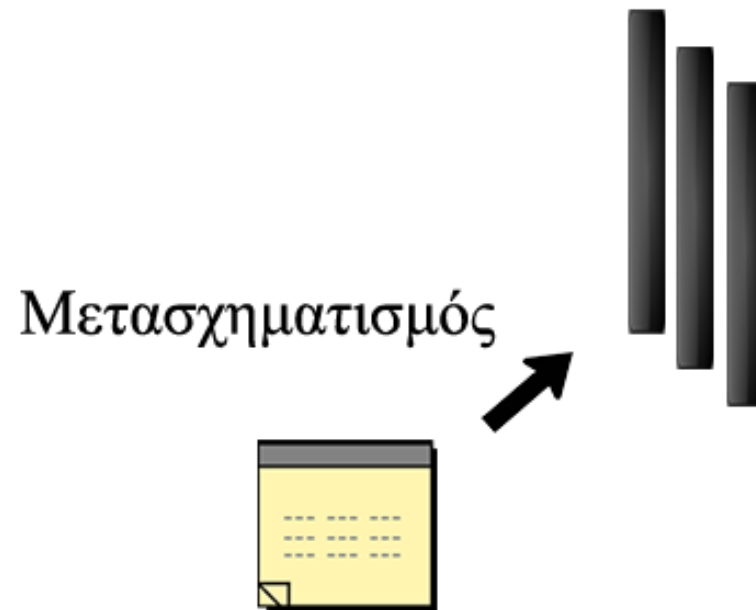
Προ επεξεργασία Δεδομένων

- › Η προ επεξεργασία του συνόλου δεδομένων, γίνεται με στόχο τον καθαρισμό τους, δηλαδή την τακτοποίηση εσφαλμένων, προβληματικών ή ελλειπόντων δεδομένων



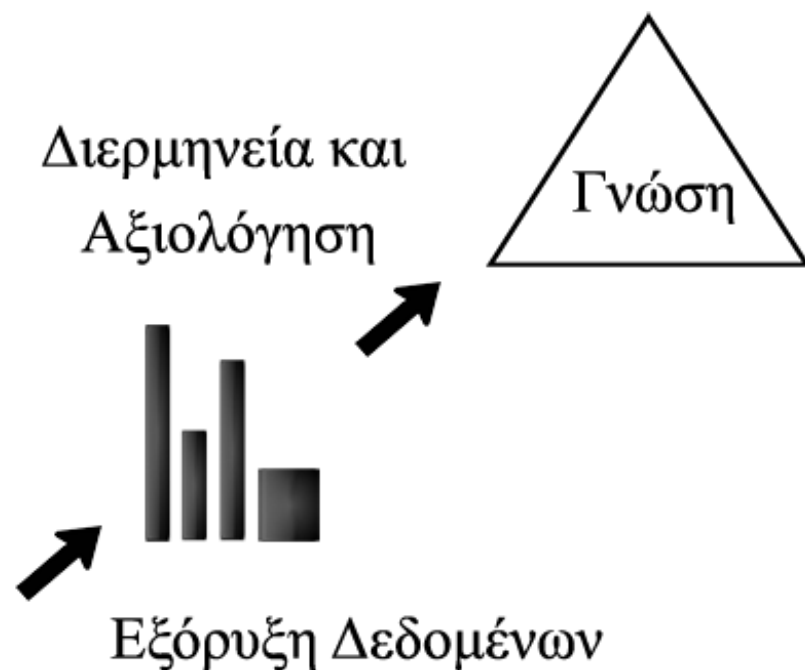
Μετασχηματισμός Δεδομένων

- › Μετατροπή των δεδομένων κάτω από ένα κοινό πλαίσιο, για επεξεργασία.



Εξόρυξη Δεδομένων – Διερμηνεία & Αξιολόγηση

- › Χρήση αλγόριθμου για την παραγωγή μοντέλου.



Τύποι Μοντέλων

- › Μοντέλα πρόβλεψης
- › Περιγραφικά Μοντέλα

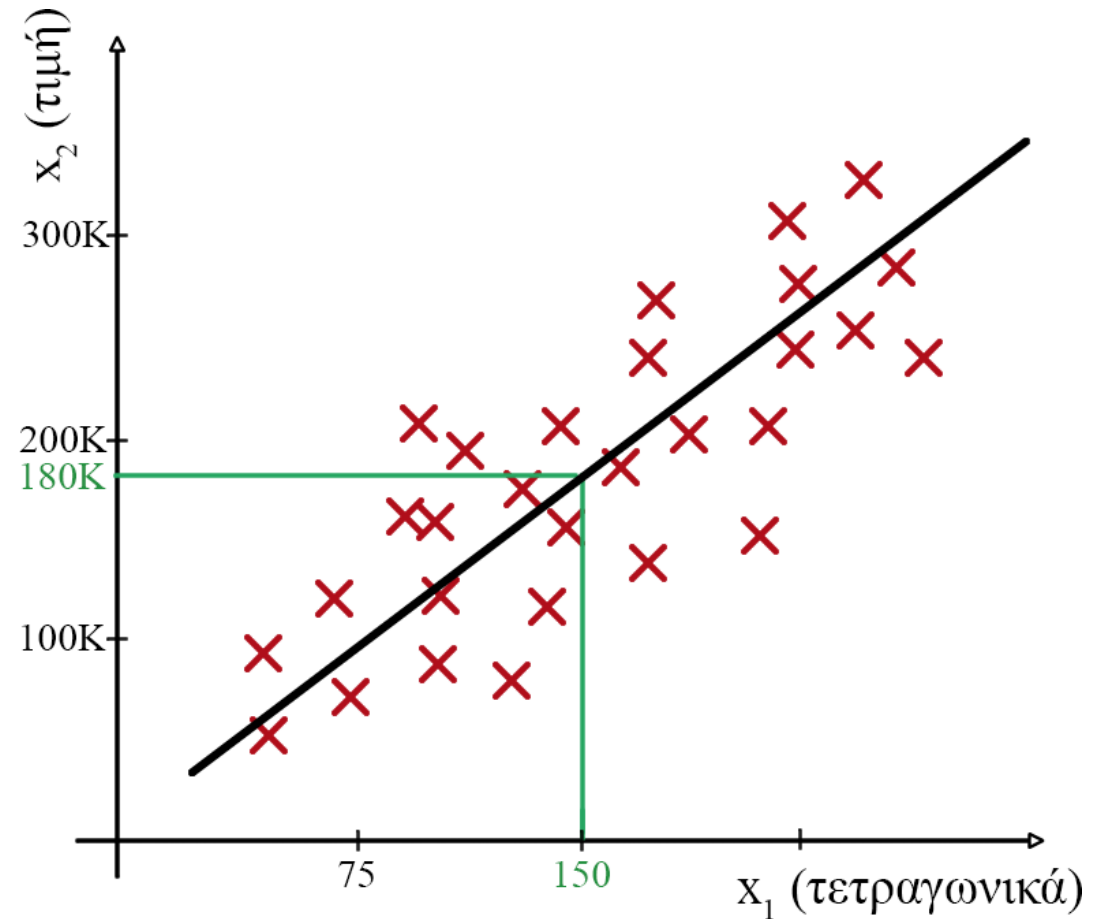
Παραδείγματα ΑΓΒΔ & ΕΔ

Κατηγοριοποίηση Μεθόδων Εξόρυξης

- › Μοντέλα πρόβλεψης
- › Περιγραφικά Μοντέλα

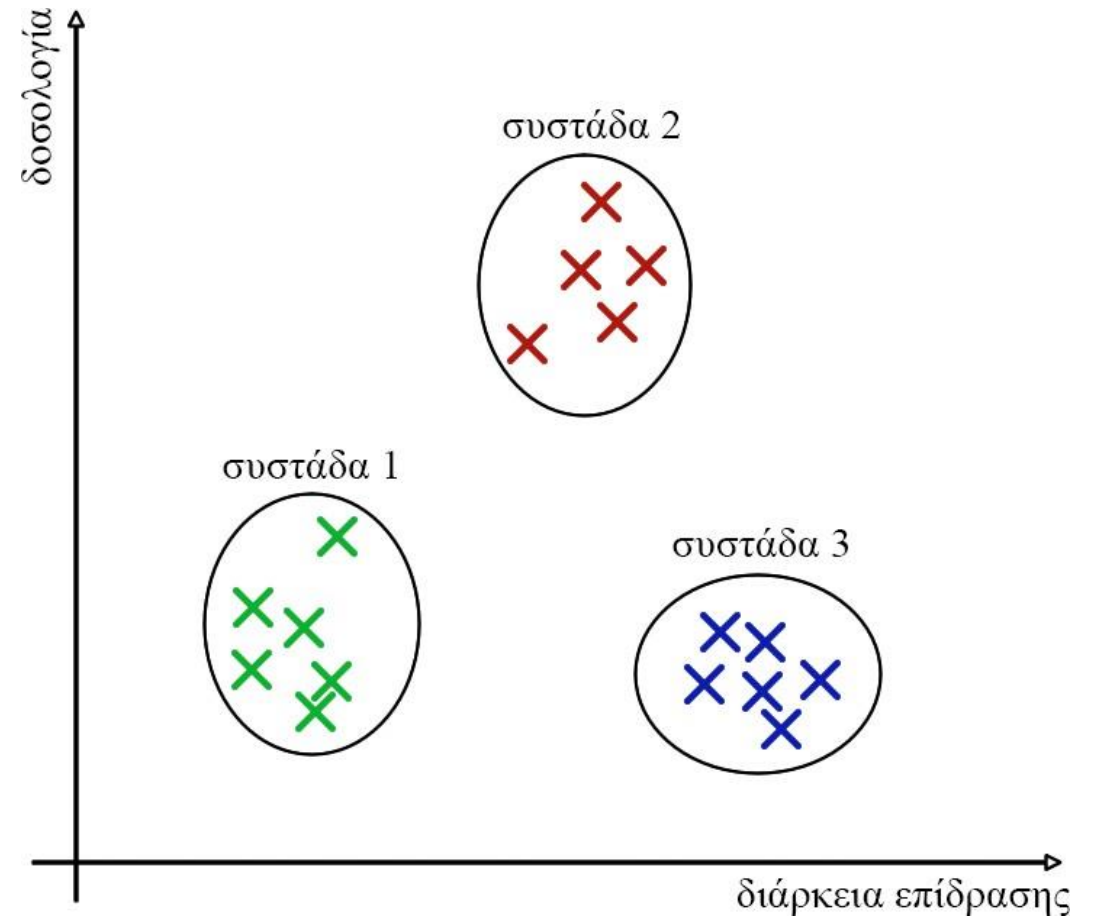
Παλινδρόμηση (regression)

- › Η **εκπαίδευση** (training) μιας συνάρτησης, η οποία απεικονίζει ένα αντικείμενο σε μία πραγματική μεταβλητή. Πρόκειται για μια, προγνωστική μέθοδο.
- › Στόχος είναι με βάση κάποιες **ανεξάρτητες μεταβλητές** (independent variables) να προβλεφθούν οι τιμές μιας εξαρτημένης μεταβλητής (dependent variable)



Συσταδοποίηση (clustering)

- › Η **συσταδοποίηση** (clustering) είναι μια περιγραφική μέθοδος.
- › Έχοντας ένα σύνολο δεδομένων, στόχος της **συσταδοποίησης** είναι η δημιουργία συστάδων (clusters), δηλαδή ομάδων, οι οποίες θα περιέχουν όμοια ή παρεμφερή δείγματα.



Εξαγωγή Κανόνων Συσχέτισης (Mining Association Rules)

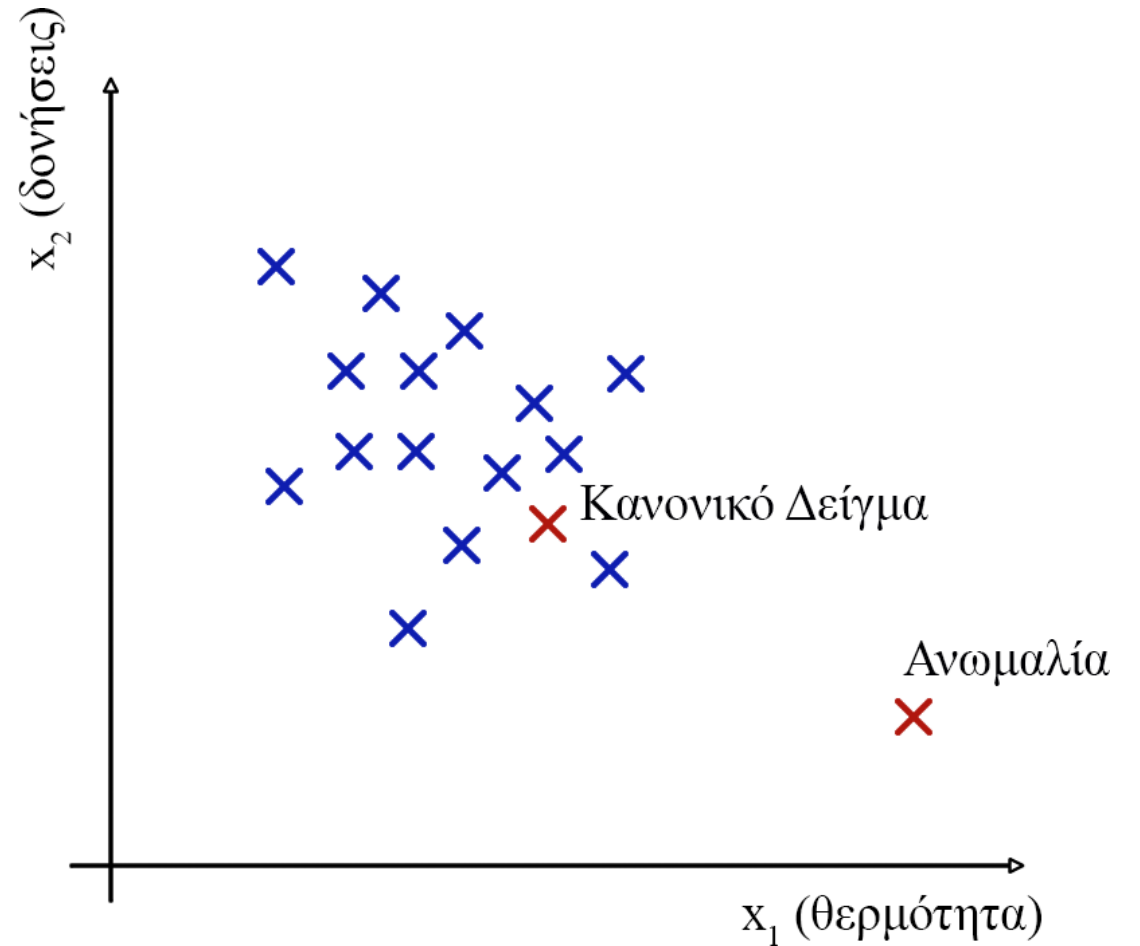
- › Οι κανόνες συσχέτισης ανακαλύπτουν κρυμμένες «συσχετίσεις» μεταξύ των γνωρισμάτων ενός συνόλου των δεδομένων.
- › Αυτοί οι συσχετισμοί παρουσιάζονται στη μορφή **A -> B**, όπου τα A και B αποτελούν σύνολα που αναφέρονται στα χαρακτηριστικά του συνόλου δεδομένων που αναλύουμε.
- › Ένας κανόνας συσχέτισης **A -> B** προβλέπει την εμφάνιση των χαρακτηριστικών του συνόλου **B** δεδομένης της εμφάνισης των χαρακτηριστικών του συνόλου A

Οπτικοποίηση

- › Η οπτικοποίηση των δεδομένων συχνά βοηθάει στην καλύτερη κατανόηση όχι μόνο των ίδιων των δεδομένων, αλλά και των συσχετίσεων που μπορεί να υπάρχουν μεταξύ τους.

Ανίχνευση Ανωμαλιών

- › Η ανίχνευση ανωμαλιών εστιάζει στην ανακάλυψη αποκλίσεων στα δεδομένα σε σχέση με αντίστοιχα δεδομένα, τα οποία έχουν συλλεχθεί στο παρελθόν ή με τυπικές τιμές των δεδομένων αυτών.



π

Επιστήμη Δεδομένων



Data Sources



Data Sources

Evolution of
Technology

IOT

Social Media

Other factors



1,736,111 pictures



347,222 tweets



204,000,000
emails



4,166,667 likes &
200,000 photos

YouTube

300 hours of video
uploaded

Data Sources

Evolution of Technology

IOT

Social Media

Other factors



Data Analysis At Walmart

Halloween and cookie sales



Data scientist at Walmart found a connection between Halloween and the sales of cookies.

Επιστήμη Δεδομένων

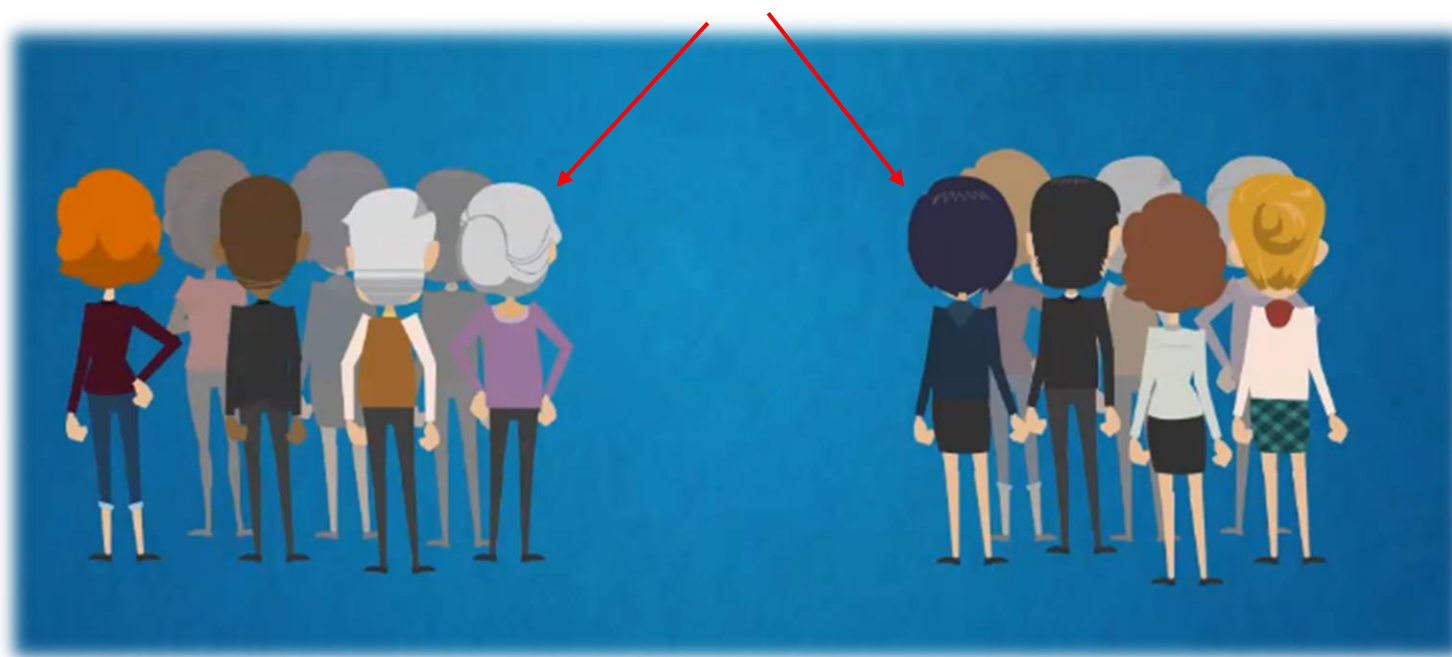
› Δεδομένα -> Αλγόριθμοι -> Οπτικοποίηση



› 30 Παρατηρήσεις

Επιστήμη Δεδομένων

Πως να εξυπηρετήσουμε καλύτερα τους πελάτες



› 30 Παρατηρήσεις

Επιστήμη Δεδομένων

```
In [1]: import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
sns.set()
from sklearn.cluster import KMeans
from sklearn import preprocessing
```

```
In [2]: data = pd.read_csv('Example.csv')
```

```
In [3]: plt.scatter(data['Satisfaction'],data['Loyalty'])
plt.xlabel('Satisfaction')
plt.ylabel('Loyalty')
```

```
Out[3]: Text(0,0.5,'Loyalty')
```



Επιστήμη Δεδομένων

Satisfaction

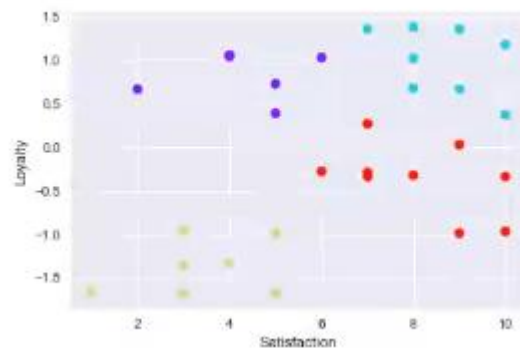
```
In [4]: x = preprocessing.scale(data)
kmeans = KMeans(4)
kmeans.fit(x)
```

```
Out[4]: KMeans(algorithm='auto', copy_x=True, init='k-means++', max_iter=300,
n_clusters=4, n_init=10, n_jobs=1, precompute_distances='auto',
random_state=None, tol=0.0001, verbose=0)
```

```
In [5]: plt.scatter(data['Satisfaction'], data['Loyalty'], c=kmeans.fit_predict(x), cmap='rainbow')
plt.xlabel('Satisfaction')
plt.ylabel('Loyalty')
```

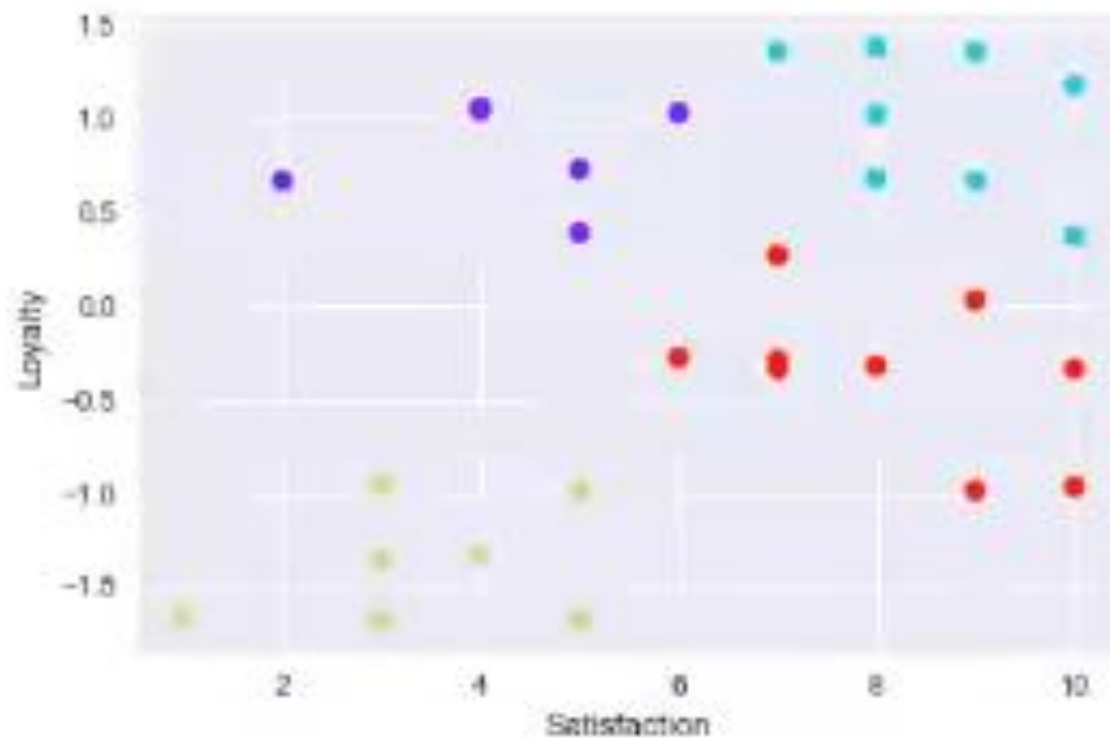
```
Out[5]: Text(0,0.5,'Loyalty')
```

```
Out[5]: Text(0,0.5,'Loyalty')
```



Επιστήμη Δεδομένων

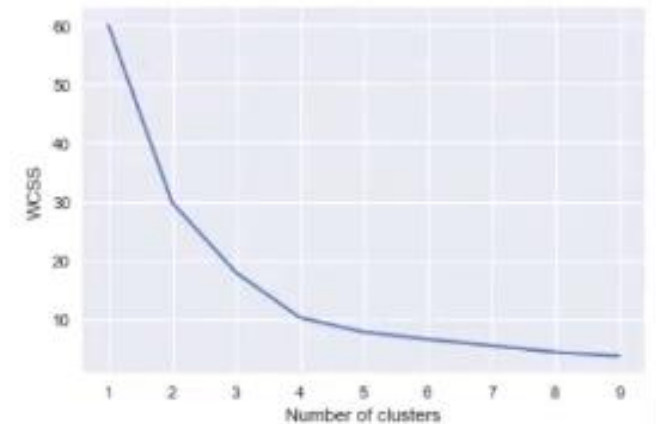
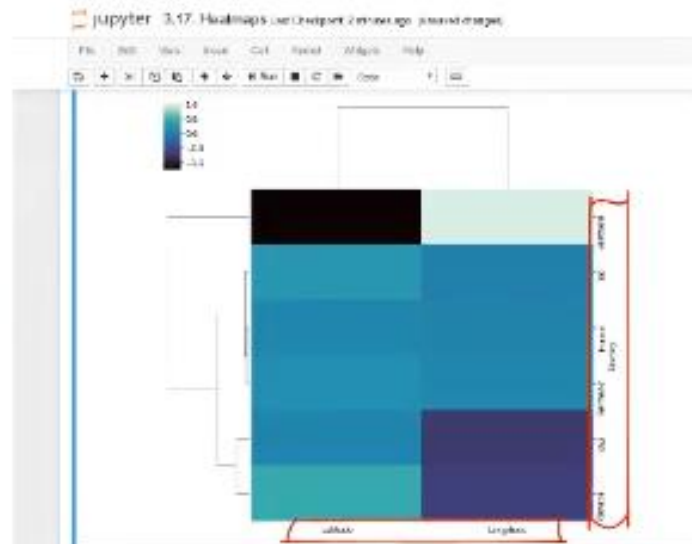
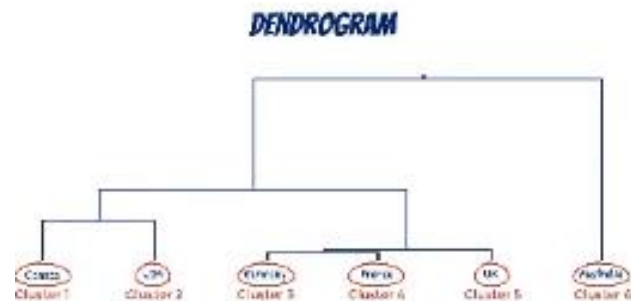
```
Out[5]: Text(0,0.5,'Loyalty')
```



Επιστήμη Δεδομένων



The screenshot shows a Jupyter Notebook window titled 'Market Segmentation example'. It displays a 'Select the features' section with a list of features, a 'Clustering' section with parameters like 'kmeans' and 'n_clusters', and a 'Clustering results' section showing the output of a clustering algorithm. A blue text box at the bottom states: 'The column 'cluster_pred' of 'clusters' will contain the predictions'.



Επιστήμη Δεδομένων

Νέες λέξεις.....



Επιστήμη Δεδομένων



- › Εξαγωγή συνόλων δεδομένων και την εφαρμογή διαφόρων στατιστικών μεθόδων.

Επιστήμη Δεδομένων



- › Εξαγωγή νέων προτύπων δεδομένων
- › ακριβείς προβλέψεις.

Επιστήμη Δεδομένων

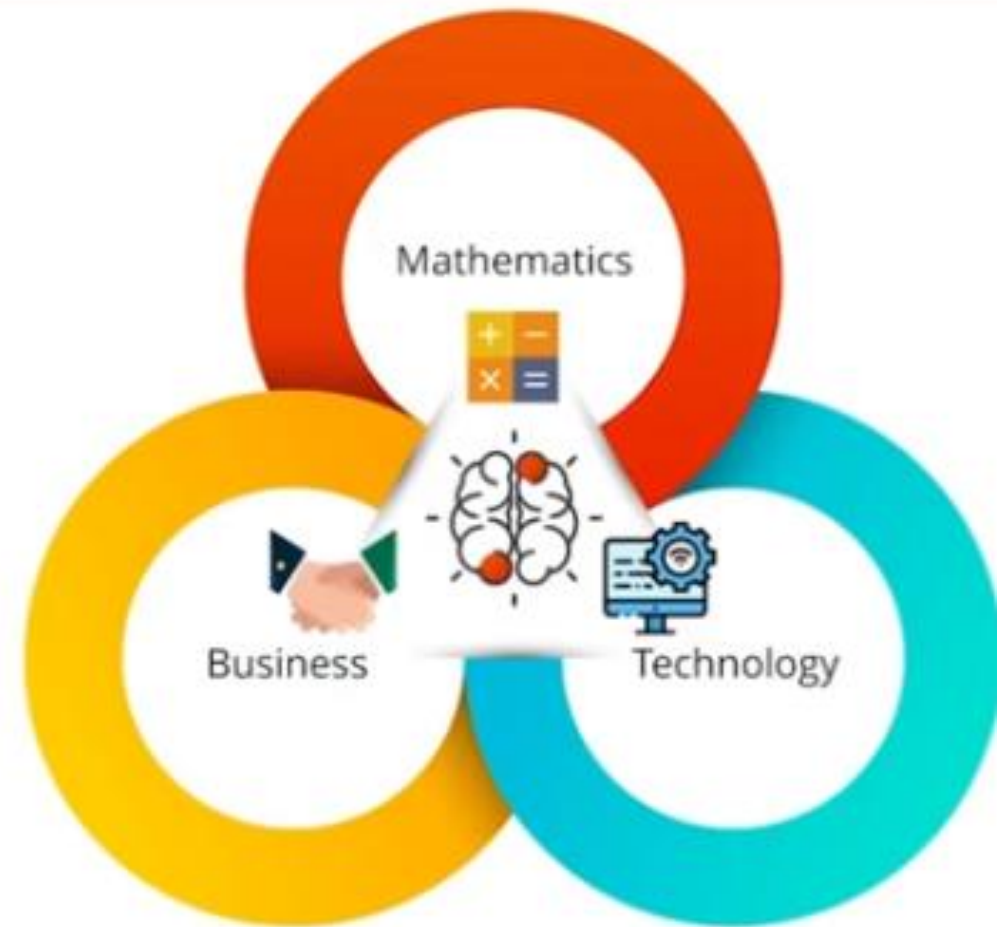


> Η προγνωστική ανάλυση

Επιστήμη Δεδομένων



Who Is A Data Scientist?



Data Science – Skill Set



Statistics



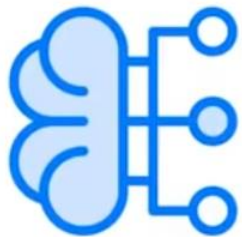
Programming
languages



Data extraction &
processing



Data wrangling &
exploration



Machine Learning



Big Data processing
frameworks



Data visualisation

Data Science Job Roles



Data Scientist



Data Analyst



Data Architect



Data Engineer



Statistician



Database Administrator



Business Analyst



Data & Analytics Manager

Επιστήμη Δεδομένων



π

***ΣΑΣ ΕΥΧΑΡΙΣΤΩ
ΓΙΑ ΤΗΝ ΠΡΟΣΟΧΗ ΣΑΣ***