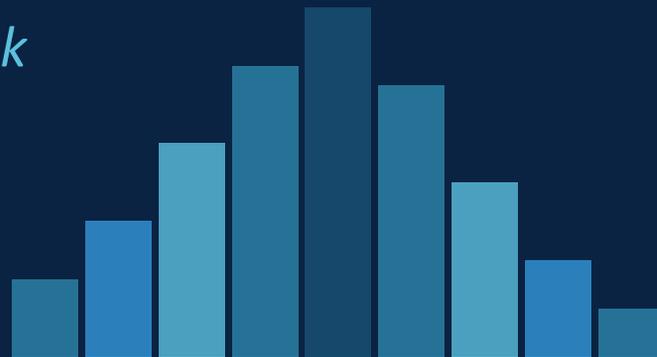


ΙΣΤΟΓΡΑΜΜΑ

Θεωρία, Εφαρμογή & Excel Data Analysis ToolPak



ΠΕΡΙΕΧΟΜΕΝΑ ΠΑΡΟΥΣΙΑΣΗΣ

01

Τι είναι το Ιστόγραμμα;

Ορισμός & θεωρητικό υπόβαθρο

02

Στοιχεία & Ερμηνεία

Κλάσεις, συχνότητες, κατανομές

03

Παράδειγμα Ναυτιλιακών Δεδομένων

Ταχύτητες πλοίων (πρακτική εφαρμογή)

04

Excel – Data Analysis ToolPak

Ενεργοποίηση & βήμα-βήμα οδηγίες

05

Ανάλυση Αποτελεσμάτων

Ερμηνεία εξόδου & βελτιώσεις

06

Ασκήσεις & Συμπεράσματα

Εξάσκηση φοιτητών

01 | ΤΙ ΕΙΝΑΙ ΤΟ ΙΣΤΟΓΡΑΜΜΑ;

Ορισμός

Το ιστόγραμμα είναι γραφική αναπαράσταση της κατανομής συχνοτήτων μιας συνεχούς ποσοτικής μεταβλητής. Αποτελείται από παρακείμενα ορθογώνια (ράβδους) όπου το εμβαδόν κάθε ράβδου αντιστοιχεί στη σχετική συχνότητα της κλάσης.

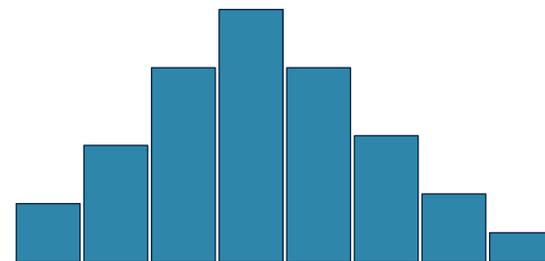
Χρησιμοποιείται για ΣΥΝΕΧΕΙΣ ποσοτικές μεταβλητές (π.χ. ταχύτητα, θερμοκρασία, ύψος)

Ανακαλύφθηκε από τον Karl Pearson το 1895 — η λέξη προέρχεται από το ελληνικό «ιστός» (ιστός, κατάρτι)

Η βάση κάθε ράβδου = εύρος κλάσης· το ύψος = απόλυτη ή σχετική συχνότητα

Δεν υπάρχουν κενά μεταξύ ράβδων (σε αντίθεση με ραβδόγραμμα ποιοτικών δεδομένων)

Ιστόγραμμα vs Ραβδόγραμμα



Ιστόγραμμα (χωρίς κενά)



Ραβδόγραμμα (με κενά)

02 | ΚΥΡΙΑ ΣΤΟΙΧΕΙΑ & ΟΡΟΛΟΓΙΑ

Κλάση (Class)

Διάστημα τιμών που ομαδοποιεί τα δεδομένα, π.χ. [10-15) kt

Απόλυτη Συχνότητα (f)

Αριθμός παρατηρήσεων που ανήκουν σε κάθε κλάση

Σχετική Συχνότητα (rf)

$rf = f / n \rightarrow$ ποσοστό παρατηρήσεων ανά κλάση

Εύρος Κλάσης (h)

Διαφορά άνω και κάτω ορίου:

$$h = \frac{x_{\max} - x_{\min}}{k}$$

Αριθμός Κλάσεων (k)

Κανόνας Sturges: $k \approx 1 + 3.322 \times \log_{10}(n)$

Αθροιστική Συχνότητα (cf)

Άθροισμα συχνοτήτων έως και της τρέχουσας κλάσης

03 | ΠΑΡΑΔΕΙΓΜΑ ΝΑΥΤΙΛΙΑΚΩΝ ΔΕΔΟΜΕΝΩΝ

Σενάριο:

Καταγράφηκαν οι ταχύτητες (σε κόμβους, knots) 30 φορτηγών πλοίων που εισήλθαν στον λιμένα του Πειραιά κατά τη διάρκεια ενός μήνα. Ζητείται η κατασκευή ιστογράμματος.

Δεδομένα: Ταχύτητες Πλοίων (knots) — n = 30 παρατηρήσεις

12.3 kt	14.7 kt	11.5 kt	16.2 kt	13.8 kt	15.0 kt	12.8 kt	17.1 kt	13.2 kt	14.5 kt
16.8 kt	11.9 kt	15.5 kt	13.0 kt	14.2 kt	12.5 kt	16.0 kt	13.7 kt	15.8 kt	14.0 kt
11.2 kt	17.5 kt	12.0 kt	15.2 kt	13.5 kt	16.5 kt	14.8 kt	12.7 kt	15.3 kt	13.9 kt

Βασικά Στατιστικά

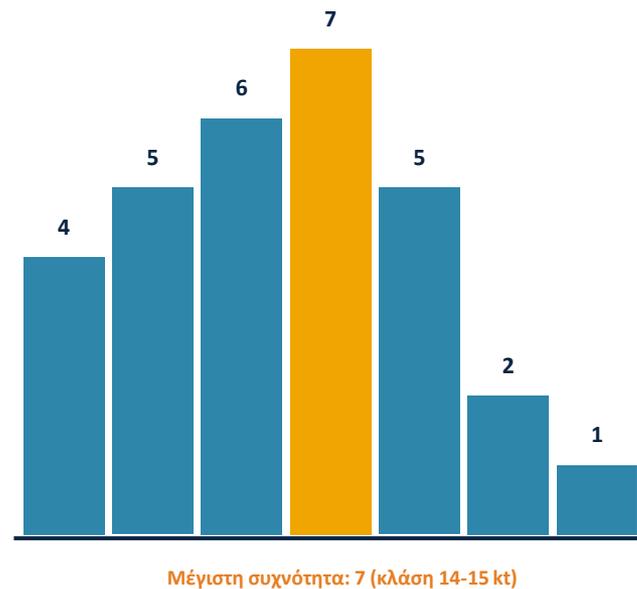
n (πλήθος)	Min	Max	Εύρος	k (Sturges)	h (εύρος κλάσης)
30	11.2 kt	17.5 kt	6.3 kt	≈ 6	≈ 1.0 kt

 Εφαρμογή Sturges: $k = 1 + 3.322 \times \log_{10}(30) = 1 + 3.322 \times 1.477 \approx 5.9 \approx 6$ κλάσεις → εύρος κλάσης $h = 6.3 / 6 \approx 1.05$ kt (στρογγυλοποιούμε σε 1.0 kt)

03 | ΠΙΝΑΚΑΣ ΚΑΤΑΝΟΜΗΣ ΣΥΧΝΟΤΗΤΩΝ

Κλάση (kt)	Κατ. Μέσο	Απόλ. Συχν. (f)
[11.0 – 12.0)	11.5	4
[12.0 – 13.0)	12.5	5
[13.0 – 14.0)	13.5	6
[14.0 – 15.0)	14.5	7
[15.0 – 16.0)	15.5	5
[16.0 – 17.0)	16.5	2
[17.0 – 18.0)	17.5	1
ΣΥΝΟΛΟ	—	30

Οπτική Αναπαράσταση



04 | EXCEL – ΕΝΕΡΓΟΠΟΙΗΣΗ DATA ANALYSIS TOOLPAK

1

Ανοίξτε το Excel

Εκκινήστε το Microsoft Excel και ανοίξτε νέο βιβλίο εργασίας (.xlsx)

2

Μενού Αρχείο → Επιλογές

Κλικ στο "Αρχείο" (File) → "Επιλογές" (Options) → Ανοίγει το παράθυρο Excel Options

3

Πρόσθετα (Add-ins)

Αριστερά: επιλέξτε "Πρόσθετα" (Add-ins). Κάτω στο "Διαχείριση": επιλέξτε "Excel Add-ins" → Κλικ "Μετάβαση"

4

Ενεργοποίηση ToolPak

Στο παράθυρο Add-ins: τσεκάρτε "Analysis ToolPak" → OK. Στη συνέχεια τσεκάρτε και "Analysis ToolPak - VBA" (προαιρετικά)

5

Πρόσβαση στο εργαλείο

Καρτέλα "Δεδομένα" (Data) → Ομάδα "Ανάλυση" → Κλικ "Ανάλυση Δεδομένων" (Data Analysis)

04 | ΒΗΜΑΤΑ ΚΑΤΑΣΚΕΥΗΣ ΙΣΤΟΓΡΑΜΜΑΤΟΣ ΣΤΟ EXCEL

ΒΗΜΑ Α: Εισαγωγή Δεδομένων

1. Στηλη Α (A1): τίτλος «Ταχύτητα (kt)»

2. Κελιά A2:A31: εισάγουμε τις 30 τιμές

3. Στηλη Β (B1): τίτλος «Όρια Κλάσεων»

4. Κελιά B2:B7: εισάγουμε τα άνω όρια:
12.0, 13.0, 14.0, 15.0, 16.0, 17.0, 18.0

(Αυτά ονομάζονται «Bin Range» στο Excel)

ΒΗΜΑ Β: Εκτέλεση Histogram Tool

1. Δεδομένα → Ανάλυση Δεδομένων

2. Επιλογή: «Histogram» → OK

3. Input Range: \$A\$1:\$A\$31 (συμπ. τίτλο)

4. Bin Range: \$B\$1:\$B\$8 (άνω όρια)

5. Labels: ✓ (εάν έχετε τίτλους)

6. Output: New Worksheet Ply ή επιλέξτε κελί

7. ✓ Chart Output — ✓ Cumulative Percentage

8. Κλικ OK → εμφανίζεται ο πίνακας & γράφημα

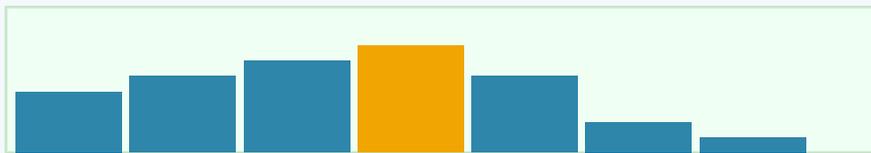


Συμβουλή: Εάν δεν ορίσετε Bin Range, το Excel δημιουργεί αυτόματα κλάσεις — αλλά συνήθως δεν είναι βέλτιστες! Πάντα ορίζετε χειροκίνητα τα όρια με βάση τον κανόνα Sturges

05 | ΕΡΜΗΝΕΙΑ ΕΞΟΔΟΥ ΤΟΥ EXCEL

Έξοδος Excel — Histogram Output

Bin	Frequency	Cumul. %
12	4	13.33%
13	5	30.00%
14	6	50.00%
15	7	73.33%
16	5	90.00%
17	2	96.67%
18	1	100.00%
More	0	100.00%



Πλήθος: 4 5 6 7 5 2 1

Ερμηνεία Αποτελεσμάτων

Bin = Άνω Όριο Κλάσης

Π.χ. «13» σημαίνει κλάση [12–13). Η τιμή 13.0 ΔΕΝ ανήκει σε αυτή την κλάση.

Frequency = Απόλυτη Συχνότητα

Αριθμός τιμών που ανήκουν στην εκάστοτε κλάση (π.χ. 7 πλοία ταξίδεψαν 14–15 kt)

Cumulative %

Αθροιστικό ποσοστό: π.χ. 73.33% των πλοίων είχαν ταχύτητα ≤ 15 knots

«More» (Περισσότερα)

Τιμές εκτός του ορισμένου Bin Range. Αν = 0, τα bin όρια καλύπτουν όλα τα δεδομένα.

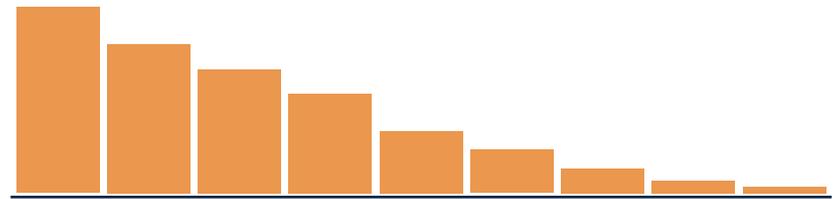
05 | ΜΟΡΦΕΣ ΚΑΤΑΝΟΜΗΣ & ΕΡΜΗΝΕΙΑ

Κανονική (Bell)



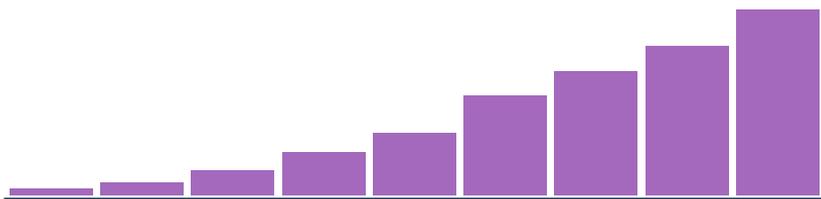
Συμμετρική, bell-shaped. Μέση=Διάμεσος=Επικρατούσα

Δεξιά Λοξή



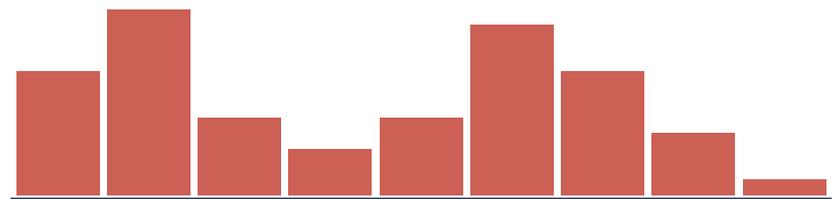
Η ουρά εκτείνεται δεξιά. Μέση > Διάμεσος

Αριστερά Λοξή



Η ουρά εκτείνεται αριστερά. Μέση < Διάμεσος

Διτροπική (Bimodal)



Δύο κορυφές — πιθανές δύο υπο-ομάδες

05 | ΒΕΛΤΙΩΣΗ ΓΡΑΦΗΜΑΤΟΣ ΣΤΟ EXCEL

Το προεπιλεγμένο γράφημα του Excel χρειάζεται βελτίωση. Ακολουθήστε τα παρακάτω βήματα:

Τίτλος & Ετικέτες

- Κλικ στο γράφημα → «Τίτλος Γραφήματος» → πληκτρολογήστε: «Κατανομή Ταχυτήτων Πλοίων (knots)»
- Άξονας Χ: δεξί κλικ → «Τίτλος Άξονα» → «Ταχύτητα (knots)»
- Άξονας Υ: «Απόλυτη Συχνότητα (n)»

Αφαίρεση Κενών Μεταξύ Ράβδων

- Κλικ σε οποιαδήποτε ράβδο → Δεξί κλικ → «Μορφοποίηση Σειράς Δεδομένων»
- «Πλάτος Κενού» (Gap Width): αλλάξτε από 219% σε 0%
- Αυτό μετατρέπει το ραβδόγραμμα σε πραγματικό ιστόγραμμα!

Χρώμα & Περίγραμμα

- Επιλέξτε ράβδους → «Γέμισμα» → επιλέξτε κατάλληλο χρώμα
- Προσθέστε περίγραμμα: «Γραμμή» → σκούρο χρώμα, πάχος 1pt
- Αφαιρέστε γραμμές πλέγματος που δεν χρειάζονται

06 | ΑΣΚΗΣΗ ΕΡΓΑΣΤΗΡΙΟΥ

Εκφώνηση Άσκησης

Σε λιμένα της Μεσογείου καταγράφηκαν οι χρόνοι παραμονής (σε ώρες) 25 φορτηγών πλοίων για φορτοεκφόρτωση. Ζητείται η κατασκευή ιστογράμματος με χρήση του Excel Data Analysis ToolPak (μέσω της λειτουργίας **Histogram**).

Δεδομένα (ώρες παραμονής):

18 24 31 27 19 35 22 28 33 21 26 30 17 25 32 20 29 36 23 27 24 31 19 28 22

Ζητούμενα:

- α) Υπολογίστε χειροκίνητα τον αριθμό κλάσεων k (Sturges) και το εύρος κλάσης h
- β) Δημιουργήστε τον πίνακα κατανομής συχνοτήτων (απόλυτες & σχετικές συχνότητες)
- γ) Κατασκευάστε ιστόγραμμα με το Excel Data Analysis ToolPak
- δ) Βελτιώστε το γράφημα (τίτλοι αξόνων, αφαίρεση κενών μεταξύ ράβδων)
- ε) Περιγράψτε την κατανομή: συμμετρία, λοξότητα, κύρια χαρακτηριστικά

Σύνοψη των βημάτων στο Excel (για τον φοιτητή):

1. **Ενεργοποίηση:** *Data -> Data Analysis -> Histogram.*

2. **Εισαγωγή:**

Input Range: Οι 25 ώρες παραμονής.

Bin Range: Τα άνω όρια που υπολογίστηκαν με Sturges.

3. **Επιλογές:** Τσεκάρετε τα «Labels» (αν υπάρχουν τίτλοι) και οπωσδήποτε το «*Chart Output*».

4. **Μορφοποίηση:** Δεξί κλικ στις ράβδους -> Format Data Series -> Gap Width = 0%

ΣΥΜΠΕΡΑΣΜΑΤΑ & ΑΝΑΚΕΦΑΛΑΙΩΣΗ

 Το ιστόγραμμα αποτελεί βασικό εργαλείο εξερευνητικής ανάλυσης δεδομένων (EDA)

 Η σωστή επιλογή κλάσεων (Sturges) είναι κρίσιμη για αξιόπιστα αποτελέσματα

 Το Excel Data Analysis ToolPak απλοποιεί σημαντικά την κατασκευή ιστογράμματος

 Εφαρμόζεται ευρέως σε ναυτιλιακά δεδομένα: ταχύτητες, χρόνους, φορτία, καιρικές συνθήκες

 Η μορφή κατανομής αποκαλύπτει σημαντικές πληροφορίες: λοξότητα, ακραίες τιμές, ομοιογένεια