

Διάλεξη 1

Περιγράφοντας τα Δεδομένα:

Πίνακες Συχνοτήτων, Κατανομές Συχνοτήτων και Γραφική Παρουσίαση.



ΣΤΟΧΟΙ

1. Οργάνωση ποιοτικών δεδομένων σε *πίνακες συχνοτήτων*.
2. Παρουσίαση του πίνακα συχνοτήτων σε *ραβδόγραμμα ή γράφημα πίτας*.
3. Οργάνωση ποσοτικών δεδομένων σε *κατανομή συχνοτήτων (frequency distribution)*.
4. Παρουσίαση κατανομής συχνοτήτων ποσοτικών δεδομένων με *ιστόγραμμα, πολύγωνα συχνοτήτων και πολύγωνα αθροιστικών συχνοτήτων*.

Πίνακας Συχνοτήτων και Κατανομή Συχνοτήτων

ΠΙΝΑΚΑΣ ΣΥΧΝΟΤΗΤΩΝ: Ομαδοποίηση ποιοτικών δεδομένων σε αμοιβαία αποκλειόμενες κλάσεις οι οποίες δείχνουν τον αριθμό των παρατηρήσεων που ανήκουν σε κάθε κλάση.

Car Type	Number of Cars
Domestic	50
Foreign	30

ΚΑΤΑΝΟΜΗ ΣΥΧΝΟΤΗΤΩΝ: Ομαδοποίηση δεδομένων σε αμοιβαία αποκλειόμενες κλάσεις οι οποίες δείχνουν τον αριθμό των παρατηρήσεων σε κάθε κλάση.

Εύρος κλάσης: Το εύρος μιας κλάσης προκύπτει από την αφαίρεση του κατώτερου ορίου μιας κλάσης από το κατώτερο της επόμενης.

Συχνότητα κλάσης: Ο αριθμός των παρατηρήσεων σε κάθε κλάση.

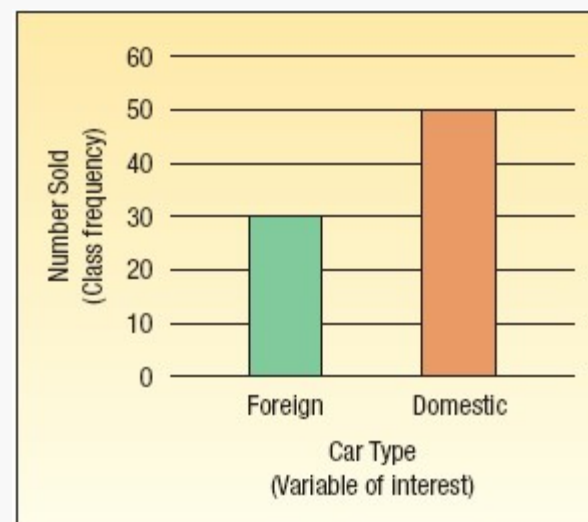
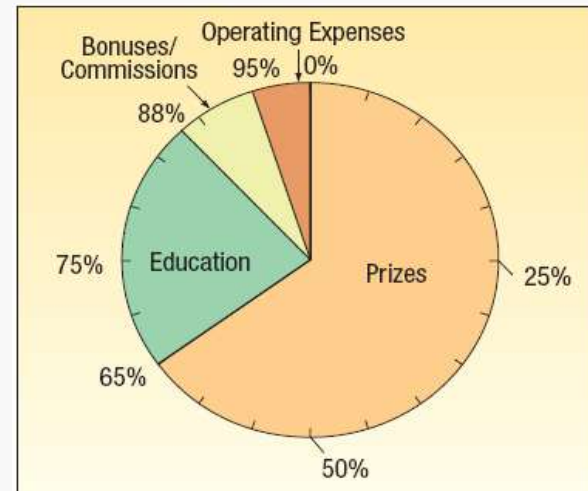
Μέση τιμή κλάσης: Το σημείο στο οποίο χωρίζεται η κλάση σε δύο ίσα μέρη. Αυτός είναι ο μέσος όρος του κατώτερου και του ανώτερου ορίου μιας κλάσης.

Selling Prices (\$ thousands)	Frequency
15 up to 18	8
18 up to 21	23
21 up to 24	17
24 up to 27	18
27 up to 30	8
30 up to 33	4
33 up to 36	2
Total	80

Πίτες και Ραβδογράμματα

ΓΡΑΦΗΜΑ ΠΙΤΑΣ: Ένα γράφημα το οποίο παρουσιάζει την αναλογία ή το ποσοστό των παρατηρήσεων που αντιπροσωπεύει η κάθε κλάση σε σχέση με το συνολικό αριθμό των παρατηρήσεων.

ΡΑΒΔΟΓΡΑΜΜΑ: Γράφημα στο οποίο οι κλάσεις παρουσιάζονται στον οριζόντιο άξονα και οι συχνότητες της κλάσης στον κάθετο άξονα. Οι συχνότητες των κλάσεων είναι ανάλογες του ύψους της κάθε ράβδου.



Σχετική Συχνότητα Κλάσης

- Οι συχνότητες μίας κλάσης μπορούν να μετατραπούν σε **σχετικές συχνότητες** ώστε να δείχνουν πλέον την αναλογία των παρατηρήσεων που περιέχει η κλάση σε σχέση με το συνολικό αριθμό των παρατηρήσεων.
- Η σχετική συχνότητα αποτυπώνει τη σχέση μεταξύ των παρατηρήσεων μίας κλάσης με τον συνολικό αριθμό των παρατηρήσεων.

TABLE 2-2 Relative Frequency Table of Vehicles Sold By Type At Whitner Autoplex Last Month

Vehicle Type	Number Sold	Relative Frequency
Domestic	50	0.625
Foreign	30	0.375
Total	<u>80</u>	<u>1.000</u>

ΠΑΡΑΔΕΙΓΜΑ-Δημιουργώντας Πίνακα Κατανομής Συχνοτήτων

Η Ms. Kathryn Ball της AutoUSA θέλει να κατασκευάσει πίνακες, διαγράμματα και γραφήματα για να παρουσιάσει την τυπική τιμή πώλησης διάφορων αντιπροσώπων. Ο πίνακας δεξιά αποτυπώνει την τιμή μόνο 80 οχημάτων που πουλήθηκαν τον περασμένο μήνα στην Whitner Autoplex.

TABLE 2-4 Prices of Vehicles Sold Last Month at Whitner Autoplex

\$23,197	\$23,372	\$20,454	\$23,591	\$26,651	\$27,453	\$17,266
18,021	28,683	30,872	19,587	23,169	35,851	19,251
20,047	24,285	24,324	24,609	28,670	15,546	15,935
19,873	25,251	25,277	28,034	24,533	27,443	19,889
20,004	17,357	20,155	19,688	23,657	26,613	20,895
20,203	23,765	25,783	26,661	32,277	20,642	21,981
24,052	25,799	15,794	18,263	35,925	17,399	17,968
20,356	21,442	21,722	19,331	22,817	19,766	20,633
20,962	22,845	26,285	27,896	29,076	32,492	18,890
21,740	22,374	24,571	25,449	28,337	20,642	23,613
24,220	30,655	22,442	17,891	20,818	26,237	20,445
21,556	21,639	24,296				

Lowest

Highest

Δημιουργώντας Πίνακα Συχνοτήτων – Παράδειγμα

- **Βήμα 1: Καθορίστε τον αριθμό των κλάσεων.**

Μια χρήσιμη συμβουλή για να προσδιοριστεί ο αριθμός των τάξεων (k) είναι ο “κανόνας 2 εις την k ” έτσι ώστε $2^k > n$. (n ο συνολικός αριθμός των παρατηρήσεων)

$n = 80$. Αν δοκιμάσουμε $k = 6$, δηλαδή 6 κλάσεις, τότε $2^6 = 64 < 80$. Συνεπώς, το 6 δεν είναι ικανοποιητικός αριθμός κλάσεων. Αν $k = 7$, τότε $2^7 = 128 > 80$. Άρα ο απαιτούμενος αριθμός κλάσεων είναι 7.

- **Βήμα 2: Καθορίστε το εύρος της κλάσης.**

Ο τύπος είναι ο εξής: $i \geq (H-L)/k$ όπου i είναι το εύρος της κλάσης, H είναι η παρατήρηση με την υψηλότερη τιμή, L είναι η παρατήρηση με τη χαμηλότερη τιμή και k είναι ο αριθμός των κλάσεων.

$$(\$35,925 - \$15,546)/7 = \$2,911$$

Στρογγυλοποιώντας προς τα πάνω, όπως ένα πολλαπλάσιο του 10 ή του 100.

Χρησιμοποιείτε κλάση εύρους \$3,000

Δημιουργώντας Πίνακα Συχνοτήτων – Παράδειγμα

- **Βήμα 3:** Όρισε τα επιμέρους όρια της κάθε κλάσης

\$15,000 up to 18,000
18,000 up to 21,000
21,000 up to 24,000
24,000 up to 27,000
27,000 up to 30,000
30,000 up to 33,000
33,000 up to 36,000

- **Βήμα 4:** Αντιστοιχίστε τις τιμές των οχημάτων σε κάθε κλάση.

- **Βήμα 5:** Αθροίστε τον αριθμό των παρατηρήσεων κάθε κλάσης

Class	Tallies
\$15,000 up to \$18,000	III
\$18,000 up to \$21,000	III
\$21,000 up to \$24,000	II
\$24,000 up to \$27,000	III
\$27,000 up to \$30,000	III
\$30,000 up to \$33,000	
\$33,000 up to \$36,000	

Selling Prices (\$ thousands)	Frequency
15 up to 18	8
18 up to 21	23
21 up to 24	17
24 up to 27	18
27 up to 30	8
30 up to 33	4
33 up to 36	2
Total	<u>80</u>

Σχετική Κατανομή Συχνοτήτων

Για να μετατρέψουμε μία κατανομή συχνοτήτων σε *σχετική* κατανομή συχνοτήτων, οι συχνότητες κάθε κλάσης διαιρούνται με το συνολικό αριθμό των παρατηρήσεων.

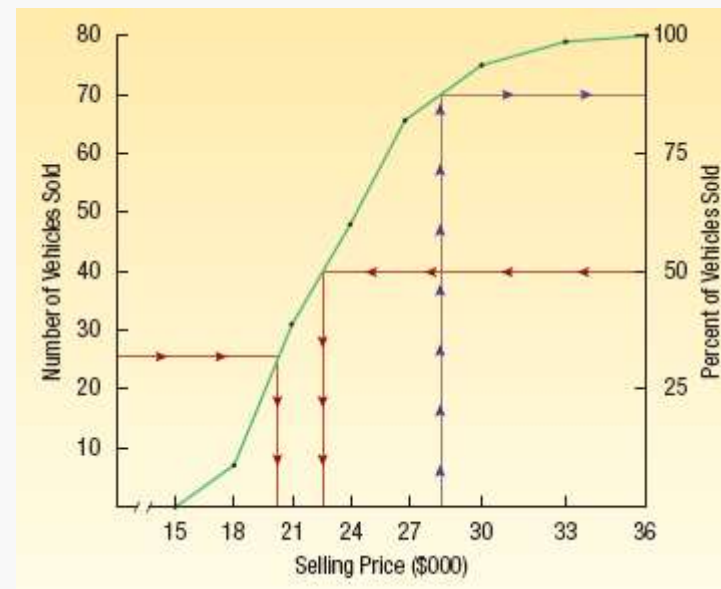
TABLE 2–8 Relative Frequency Distribution of the Prices of Vehicles Sold Last Month at Whitner Autoplex

Selling Price (\$ thousands)	Frequency	Relative Frequency	Found by
15 up to 18	8	0.1000	← 8/80
18 up to 21	23	0.2875	23/80
21 up to 24	17	0.2125	17/80
24 up to 27	18	0.2250	18/80
27 up to 30	8	0.1000	8/80
30 up to 33	4	0.0500	4/80
33 up to 36	2	0.0250	2/80
Total	80	1.0000	

Γραφική Παρουσίαση Κατανομής Συχνοτήτων

Οι τρεις πιο συχνές γραφικές παρουσιάσεις δεδομένων είναι:

- Ιστογράμματα
- Πολύγωνα Συχνοτήτων
- Κατανομές Αθροιστικών Συχνοτήτων



Ιστόγραμμα

ΙΣΤΟΓΡΑΜΜΑ: γράφημα στο οποίο οι κλάσεις τοποθετούνται στον οριζόντιο άξονα και οι συχνότητες στον κάθετο άξονα. Οι συχνότητες των κλάσεων αντιπροσωπεύονται από το ύψος των ράβδων οι οποίες αλληλοδιαδέχονται η μία την άλλη χωρίς κενά.

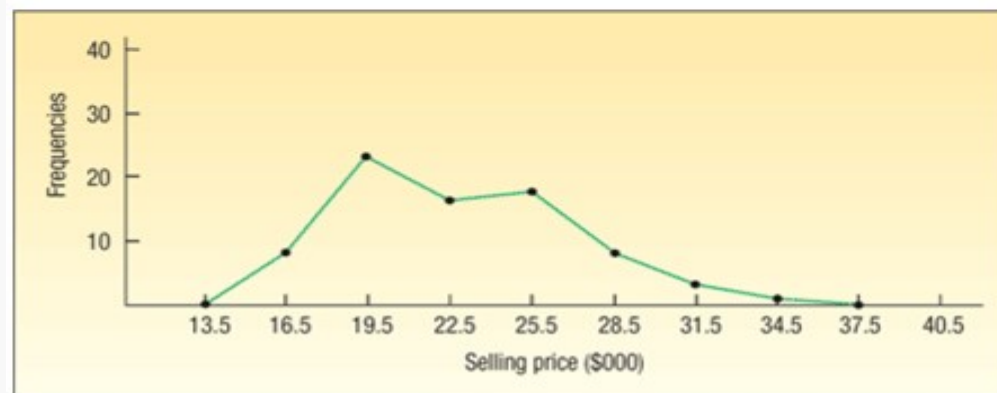


CHART 2-4 Histogram of the Selling Prices of 80 Vehicles at Whitner Autoplex

Πολύγωνο Συχνοτήτων

- Ένα πολύγωνο συχνοτήτων δείχνει το σχήμα της κατανομής και προσομοιάζει αρκετά με το ιστόγραμμα.
- Συγκροτείται από γραμμικά τμήματα τα οποία ορίζονται από τη μέση του εύρους της κάθε κλάσης.

Selling Price (\$ thousands)	Midpoint	Frequency
15 up to 18	16.5	8
18 up to 21	19.5	23
21 up to 24	22.5	17
24 up to 27	25.5	18
27 up to 30	28.5	8
30 up to 33	31.5	4
33 up to 36	34.5	2
Total		80



Ιστόγραμμα VS. Πολύγωνο Συχνοτήτων

- Και τα δύο προσφέρουν μία γρήγορη εικόνα των κύριων χαρακτηριστικών που παρουσιάζουν τα δεδομένα (αύξουσα/φθίνουσα πορεία κλπ.)
- Το ιστόγραμμα έχει το πλεονέκτημα να απεικονίζει κάθε κλάση ως ένα ορθογώνιο, με το ύψος του κάθε ορθογώνιου να αντιπροσωπεύει τον αριθμό των παρατηρήσεων σε κάθε κλάση.
- Το πολύγωνο συχνότητας έχει ένα πλεονέκτημα έναντι του ιστογράμματος. Μας επιτρέπει να συγκρίνουμε άμεσα δύο ή περισσότερες κατανομές συχνοτήτων.

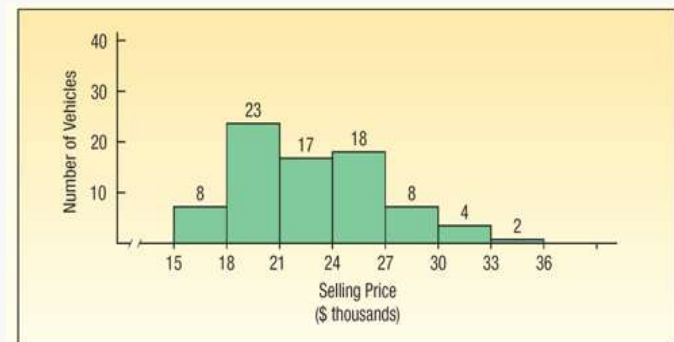
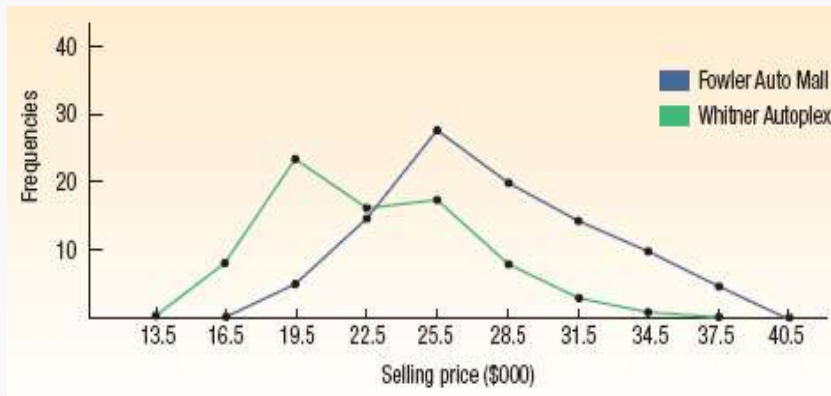


CHART 2-4 Histogram of the Selling Prices of 80 Vehicles at Whitner Autoplex

Αθροιστική Κατανομή Συχνοτήτων

Selling Price (\$ thousands)	Frequency	Cumulative Frequency	Found by
15 up to 18	8	8	
18 up to 21	23	31	← 8 + 23
21 up to 24	17	48	8 + 23 + 17
24 up to 27	18	66	8 + 23 + 17 + 18
27 up to 30	8	74	⋮
30 up to 33	4	78	
33 up to 36	2	80	
Total	<u>80</u>		

