

# ΑΝΑΛΥΣΗ ΔΕΔΟΜΕΝΩΝ

---

**1. ΕΙΔΗ ΔΕΔΟΜΕΝΩΝ, ΣΥΛΛΟΓΗ,  
ΚΩΔΙΚΟΠΟΙΗΣΗ ΚΑΙ ΕΙΣΑΓΩΓΗ**

# Βασικές μορφές Ερωτήσεων - απαντήσεων

---

- Ανοιχτές
- Κλειστές
- Κλίμακας

# Ανοιχτές ερωτήσεις

---

- **Ανοιχτές** είναι οι ερωτήσεις που δε δίνουν δυνατότητα επιλογής απάντησης στον ερωτώμενο

# Ανοιχτές ερωτήσεις

---

## □ Πλεονεκτήματα:

- Επιτρέπουν στους ερωτώμενους να χρησιμοποιήσουν δικές τους εκφράσεις

- Καταγραφή απόψεων

## □ Μειονεκτήματα:

- Δύσκολα κωδικοποιούνται και ερμηνεύονται

- Ατελείς απαντήσεις

# Κλειστές ερωτήσεις

---

- Οι **κλειστές ερωτήσεις** δίνουν επιλογές στο ερωτηματολόγιο που μπορούν να απαντηθούν γρήγορα και εύκολα

# Κλειστές - δίτιμες (Dichotomous)

---

- **Δίτιμες:** Οι ερωτήσεις που έχουν μόνο δύο επιλογές απάντησης όπως «ΝΑΙ» ή «ΟΧΙ»
  - Πλεονέκτημα:
    - Εύκολη διαχείριση και κωδικοποίηση
  - Μειονέκτημα:
    - Υπεραπλούστευση των επιλογών (π.χ. σε ερωτήσεις προτίμησης)

# Κλειστές – πολλαπλών κατηγοριών

---

- **Πολλαπλών κατηγοριών:** έχουν περισσότερες από δύο επιλογές για απάντηση
  - Πλεονεκτήματα:
    - Επιτρέπουν ευρύ φάσμα δυνατών απαντήσεων
    - Απλές στη διαχείριση και την κωδικοποίηση
  - Μειονεκτήματα:
    - Πρέπει να αποσαφηνιστεί το «επιλέγω μία» και όχι «επιλέγω όλες όσες ισχύουν»
    - Ωθούν τους ερωτώμενους να απαντήσουν σε επιλογές που αγνοούν

# Ερωτήσεις κλίμακας

---

- Οι ερωτήσεις κλίμακας χρησιμοποιούν μια κλίμακα που αναπτύχθηκε ή υιοθετήθηκε από τον ερευνητή για μέτρηση – καταγραφή βαθμού ενός χαρακτηριστικού



# Κλίμακας χωρίς ονομασία

---

- **Χωρίς ονομασία:** χρησιμοποιείται καθαρά αριθμητική κλίμακα (πιθανώς μέσα σε κάποια καθορισμένα όρια)
  - Πλεονεκτήματα:
    - Επιτρέπει την έκφραση του βαθμού αφηρημένων εννοιών (συναισθήματα)
    - Απλή στη διαχείριση και κωδικοποίηση
  - Μειονέκτημα:
    - Οι ερωτώμενοι μπορεί να μην προσαρμόζονται καλά στην κλίμακα

# Κλίμακας με ονομασία

---

- **Με ονομασία:** χρησιμοποιείται μια κλίμακα στην οποία όλες οι τιμές ορίζονται από μια περιγραφή
  - Πλεονεκτήματα:
    - Επιτρέπει την έκφραση βαθμονόμησης
    - Απλή στη διαχείριση και κωδικοποίηση
    - Οι ερωτώμενοι προσαρμόζονται στην κλίμακα
  - Μειονέκτημα:
    - Η κλίμακα μπορεί να είναι υπερβολικά λεπτομερής

# Επιλογή μορφής ερωτήσεων και απαντήσεων

---

- Η φύση της ιδιότητας
  - Φύλο=δίτιμη
  - Προτίμηση προϊόντος=κλίμακας
- Προηγούμενες σχετικές έρευνες
  - Συνιστάται η χρήση μορφών που έχουν ήδη χρησιμοποιηθεί σε παρόμοιες έρευνες ειδικά όταν θέλουμε να συγκρίνουμε

# Επιλογή μορφής ερωτήσεων και απαντήσεων

---

- Ο τύπος συλλογής δεδομένων
  - Κάποιες κλίμακες δύσκολα χρησιμοποιούνται σε τηλεφωνικές συνεντεύξεις
- Ικανότητα του ερωτώμενου
  - Τα παιδιά δεν μπορούν να προσαρμοστούν σε ερωτήσεις κλίμακας

# Είδη μεταβλητών

---

- **Ονομαστικές (Nominal)**: αυτές που οι τιμές τους έχουν απλά διαφορετική περιγραφή
- **Διάταξης (Ordinal)**: αυτές που οι τιμές τους μπορούν να διαταχθούν σε μια σειρά
- **Διαστήματος (Interval)**: αυτές στις οποίες μπορεί να οριστεί μόνο η έννοια της απόστασης
- **Αναλογίας (Ratio)**: αυτές στις οποίες υπάρχει η έννοια του μηδενός

# Είδη μεταβλητών – επίπεδα μετρήσεων

Measurement Scales Differ by What Scale Characteristics They Possess

Level of Measurement		Scale Characteristic			
		Description	Order	Distance	Origin
Categorical scales	Nominal	Yes	No	No	No
	Ordinal	Yes	Yes	No	No
Metric scales	Interval	Yes	Yes	Yes	No
	Ratio	Yes	Yes	Yes	Yes

# Σημασία επιπέδου μετρήσεων

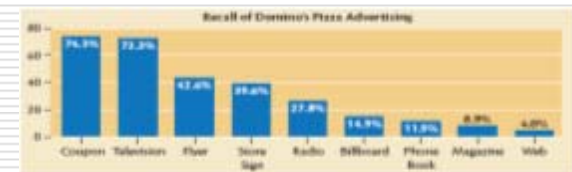
□ Το είδος της μεταβλητής επηρεάζει την ερμηνεία των αποτελεσμάτων του χαρακτηριστικού που μετράμε

## ■ Παραδείγματα:

- Υπολογισμός μέσης τιμής μόνο σε διαστήματος ή αναλογίας
- Η περιγραφή ονομαστικής μεταβλητής ή διάταξης γίνεται με κατανομή συχνοτήτων

How much do you spend on pizza in a typical month?

Average
\$21.80



# Παράδειγμα διαφορετικών επιπέδων μετρήσεων

## A. Nominal-Scaled Questions

1. Please indicate your gender.      \_\_\_ Male \_\_\_ Female
2. Check all the brands you would consider purchasing.  
\_\_\_ Sony  
\_\_\_ Zenith  
\_\_\_ RCA  
\_\_\_ Curtis Mathes
3. Do you recall seeing a Delta Airlines advertisement for “carefree vacations” in the past week?  
\_\_\_ Yes      \_\_\_ No

## B. Ordinal-Scaled Questions

1. Please rank each brand in terms of your preference. Place a “1” by your first choice, a “2” by your second choice, and so on.  
\_\_\_ Arrid  
\_\_\_ Right Guard  
\_\_\_ Mennen
2. For each pair of grocery stores, circle the one you would be more likely to patronize.  
Kroger versus First National  
First National versus A&P  
A&P versus Kroger
3. In your opinion, would you say the prices at Wal-Mart are  
\_\_\_ Higher than Sears,  
\_\_\_ About the same as Sears, or  
\_\_\_ Lower than Sears?

## C. Interval-Scaled Questions

1. Please rate each brand in terms of its overall performance.

Brand	Rating (Circle One)									
	Very Poor					Very Good				
Mont Blanc	1	2	3	4	5	6	7	8	9	10
Parker	1	2	3	4	5	6	7	8	9	10
Cross	1	2	3	4	5	6	7	8	9	10

2. Indicate your degree of agreement with the following statements by circling the appropriate number.

Statement	Strongly Disagree				Strongly Agree				
	a. I always look for bargains.	1	2	3	4	5			
b. I enjoy being outdoors.	1	2	3	4	5				
c. I love to cook.	1	2	3	4	5				

3. Please rate the Pontiac Firebird by checking the line that best corresponds to your evaluation of each item listed.

Slow pickup \_\_\_\_\_ Fast pickup  
 Good design \_\_\_\_\_ Bad design  
 Low price \_\_\_\_\_ High price

## D. Ratio-Scaled Questions

1. Please indicate your age.  
\_\_\_ Years
2. Approximately how many times in the past month have you purchased anything over \$5 in value at a 7-11 store?  
0    1    2    3    4    5    More (specify: \_\_\_)
3. How much do you think a typical purchaser of a \$100,000 term life insurance policy pays per year for that policy?  
\$ \_\_\_\_\_
4. What is the probability that you will use a lawyer’s services when you are ready to make a will?  
\_\_\_ percent



# Μέτρηση αντικειμενικών χαρακτηριστικών

---

- Φυσικά αναγνωρίσιμα χαρακτηριστικά  
π.χ. Ηλικία, φύλο, έσοδα κλπ.

# Μέτρηση υποκειμενικών χαρακτηριστικών

---

- Δεν είναι δυνατή η άμεση παρατήρησή τους γιατί είναι πνευματικά χαρακτηριστικά, π.χ. Προτιμήσεις, γνώμες, συναισθήματα, προθέσεις
- Για τα υποκειμενικά χαρακτηριστικά οι ερευνητές πρέπει να αναπαραστήσουν τις αφηρημένες έννοιες με τιμές
  - (π.χ. Likert Scale)

# Σχεδιασμός ερωτηματολογίου

---

- ❑ Εξαρτάται από τους σκοπούς της έρευνας και τον τρόπο συλλογής των δεδομένων
- ❑ Υπάρχουν προγράμματα λογισμικού που βοηθούν στην ανάπτυξη και διανομή ερωτηματολογίων με προφανή πλεονεκτήματα (ευκολία, ταχύτητα, φιλικότητα, ευελιξία)

WebSurveyor Desktop 4.1 - [Papa John's Qnr.wsx]

File Edit View Notifications Tools Help

Survey Editing

- Question List
- Details
- Languages
- Presentation
- Preview

Survey Management

- Account Details
- Public
- Send
- Manag

Survey R

- Analysis
- Actual Data
- Report

Question List

Add... Edit... Delete...

1) Have you eaten a Papa John's Pizza in the past 30 days?  
 Yes  No  Maybe  
 Yes, No (Go to Question 4), Not Sure (Go to Question 4)

2) About how many Papa John's Pizzas did you eat last month?  
 Numeric: 478  
 Minimum value = 0, Maximum value = 100

3) About how much do you spend on Papa John's Pizza when you have one?  
 Numeric: 478  
 Minimum value = 0, Maximum value = 100

4) Select where you recall seeing or hearing a Papa John's Pizza advertisement.  
 Radio  Billboard  Internet  Coupon  Store sign  Telephone book ad

5) How much do you spend on pizza's (all types) per month?  
 Numeric: 478  
 Minimum value = 0, Maximum value = 100

6) Indicate your opinion of Papa John's Pizza separately on each factor.  
 Categories: Distinctive taste, Freshness of product, Selection of pizza sizes, Competitive price, Choice of toppings  
 Response options: Poor, Fair, Good, Very Good, Excellent

7) When you are selecting a pizza delivery service, how important is each factor?  
 Categories: Distinctive taste, Freshness of product, Selection of pizza sizes, Competitive price, Choice of toppings  
 Response options: Not Important, Slightly Important, Somewhat Important, Important, Very Important

8) Overall, how satisfied are you with Papa John's Pizza?  
 Yes  No  Maybe  
 Extremely dissatisfied, Very dissatisfied, Somewhat dissatisfied, Neither satisfied nor dissatisfied, So

9) Indicate your response to each of the following statements.  
 Categories: I enjoy going out., I have the time to do all the things I want to do., I like doing things that are fun.  
 Response options: Strongly Disagree, Disagree, Neutral, Agree, Strongly Agree

10) Indicate your gender.  
 Yes  No  Maybe  
 Male, Female

11) Indicate your age.  
 Numeric: 478  
 Minimum value = 1, Maximum value = 100

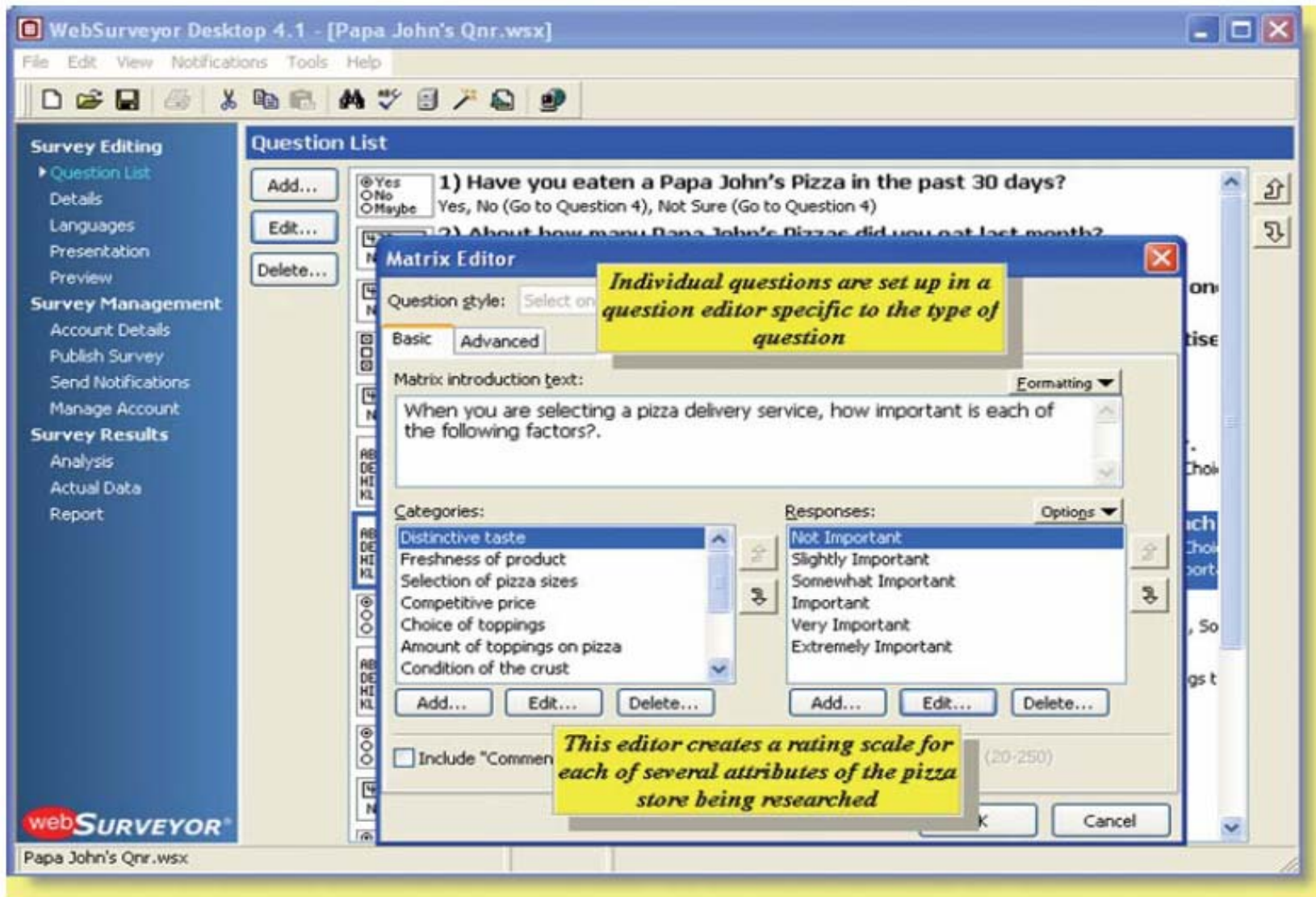
12) Indicate your student classification.

*WebSurveyor provides easy development of questions in a simple navigation window*

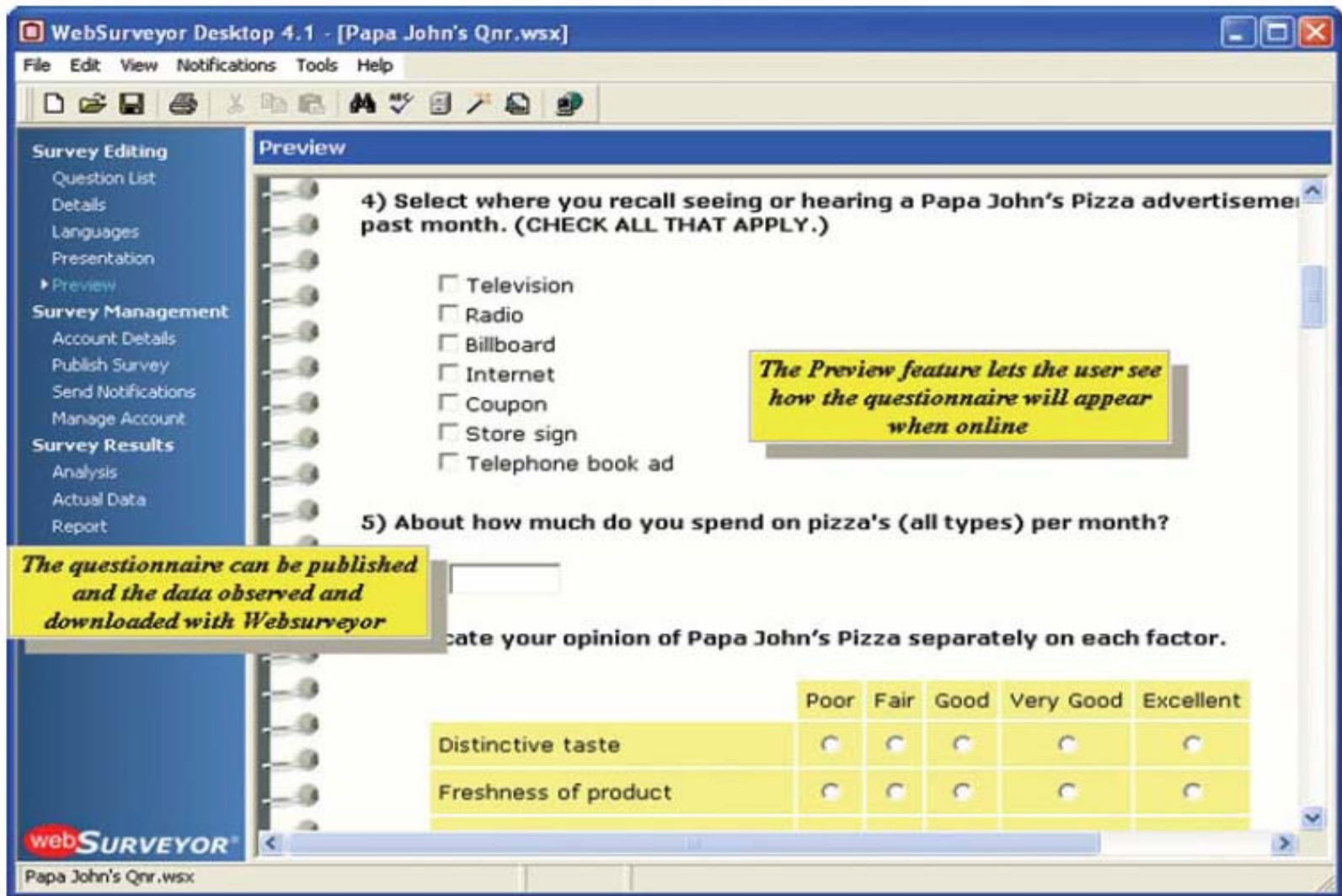
*Questions types are identified by icons and may be added, deleted, moved or edited quickly and easily*

webSURVEYOR®

Papa John's Qnr.wsx







# Συλλογή δεδομένων – δειγματοληψία (sampling)

---

- **Πληθυσμός (population)**: το σύνολο στο οποίο αναφέρεται η έρευνα όπως ορίζεται από τους ερευνητικούς σκοπούς

# Βασικές έννοιες δειγματοληψίας

---

- **Δείγμα (Sample):** υποσύνολο του πληθυσμού που πρέπει να αντιπροσωπεύει όλο τον πληθυσμό
- **Δειγματική μονάδα (Sample unit):** η βασική μονάδα της έρευνας
- **Απογραφή (Census):** η καταγραφή ολόκληρου του πληθυσμού



# Λόγοι για λήψη δείγματος

---

- ❑ Πρακτικοί λόγοι όπως κόστος και μέγεθος πληθυσμού
- ❑ Αδυναμία του ερευνητή να αναλύσει τεράστια ποσά δεδομένων που παράγονται από απογραφές
- ❑ Τα δείγματα μπορούν να δώσουν ακριβή αποτελέσματα με μικρό κόστος

# Μέθοδοι δειγματοληψίας

---

- Μεγάλο πλήθος μεθόδων ανάλογα με το είδος της έρευνας
  - **Probability samples:** αυτά στα οποία τα μέλη ενός πληθυσμού έχουν μια γνωστή πιθανότητα να επιλεγούν στο δείγμα
  - **Non-probability samples:** περιπτώσεις όπου οι πιθανότητες επιλογής μελών του πληθυσμού στο δείγμα είναι άγνωστες
- Πολύ σημαντικός ο καθορισμός **μεγέθους του δείγματος** (απαιτούνται στατιστικές μέθοδοι – διαστήματα εμπιστοσύνης)

# Κωδικοποίηση δεδομένων

---

- **Εισαγωγή δεδομένων (Data entry):**  
Δημιουργία αρχείου που καταγράφονται τα δεδομένα κατάλληλο για ανάλυση
- **Κωδικοποίηση δεδομένων (Data coding):**  
καθορισμός κωδικών που παριστούν τις απαντήσεις – παρατηρήσεις – μετρήσεις που έχουμε συλλέξει
- **Κατάλογος κωδικών (Data code book):**  
καθορισμός όλων των ονομάτων των μεταβλητών και των κωδικών αριθμών που αντιστοιχούν σε κάθε τιμή

# Εισαγωγή δεδομένων

The screenshot displays the SPSS Data Editor interface for a file named 'HobbitData.sav'. The main window shows a list of variables in Variable View. The 'eatrest' variable is highlighted in yellow. A yellow callout box points to the Variable View table with the text: "Variable View shows the variable names, variable labels, value codes, value labels, etc."

Name	Type	Label	Values	Missing
eatrest	N20	Do you eat at this type of restaurant at least once every two weeks	{1, Yes}...	None
totspent	D82	How many total dollars do you spend per month i	None	None
likely	N20	How likely would it be for you to patronize t		
avprice	D82	What would you expect an average evenin		
lisradio	N20	Would you describe yourself as one who lis		
radprogr	N20	To which type of radio programming do you most	{1, Country&Western}	None
vietvne				None
timene			{s}...	None
reapap				None
secpap				None
cityma				None
water				Not None
drive				Not None
waitsta				Not None
desser				Not None
variety				Not None
unusal				Not None
simple		Prefer Simple Decor	{1, Very Strongly	Not None
elegant	N20	Prefer Elegant Decor	{1, Very Strongly Not	None
string	N20	Prefer String Quartet	{1, Very Strongly Not	None

A 'Variables' dialog box is open, showing a list of variables on the left and 'Variable Information' for 'eatrest' on the right. The 'Variable Information' window shows: Label: Do you eat at this type of restaurant at least once every two weeks; Type: F2; Missing Values: none; Measurement Level: Scale; Value Labels: 1 Yes, 2 No. A yellow callout box points to this window with the text: "Variable Information window gives specifics on each variable".

The 'Data View' tab is selected, and the 'Variable View' tab is circled in red at the bottom of the window.

# Εισαγωγή δεδομένων

---

- Τα δεδομένα στο SPSS καταχωρούνται στον **επεξεργαστή δεδομένων (Data Editor)** σε μορφή πίνακα
- Οι γραμμές αντιστοιχούν στις μονάδες ή στις **περιπτώσεις** του δείγματος (**cases**)
- Οι στήλες αντιστοιχούν στις **μεταβλητές** (**variables**)
- Κάθε περίπτωση περιγράφεται με τον ίδιο αριθμό μεταβλητών οι οποίες εκφράζουν διάφορα χαρακτηριστικά που έχουν καταγραφεί κατά τη διάρκεια μιας έρευνας

# Κανόνες κωδικοποίησης των δεδομένων

---

- ❑ Όλα τα δεδομένα πρέπει να είναι αριθμητικά
- ❑ Τα δεδομένα από κάθε περίπτωση πρέπει να καταλαμβάνουν μόνο μια γραμμή (*row*)
- ❑ Κάθε στήλη (*column*) πρέπει να περιέχει δεδομένα της ίδιας μεταβλητής για όλες τις περιπτώσεις
- ❑ Όταν η ίδια μεταβλητή μετριέται περισσότερες από μία φορές για κάθε περίπτωση (π.χ. αποτελέσματα εξέτασης πριν και μετά) τότε πρέπει να εισάγουμε διαφορετικές στήλες με διαφορετικά ονόματα

# Κανόνες κωδικοποίησης των δεδομένων

---

- Όλες οι τιμές (κωδικοί) για μια μεταβλητή πρέπει να είναι αμοιβαία αποκλειόμενες
  - Μόνο μια τιμή μπορεί να καταγραφεί για κάθε μεταβλητή
  - Σε ερωτήσεις (π.χ. προτιμήσεις προϊόντων) όπου οι ερωτώμενοι αφήνονται να συμπληρώσουν περισσότερες από μια απαντήσεις κάθε επιλογή πρέπει να αποτελεί ξεχωριστή μεταβλητή με τιμές 1 που θα αντιστοιχεί στο «ΝΑΙ» (επιλογή) και 0 που θα αντιστοιχεί στο «ΟΧΙ» (δεν το επέλεξε)

# Κανόνες κωδικοποίησης των δεδομένων

---

- Κάθε μεταβλητή πρέπει να κωδικοποιείται ώστε να πετυχαίνουμε μέγιστη πληροφορία
  - Δεν είναι σωστό να ενώνουμε κατηγορίες ή τιμές σε μία όταν κάνουμε την κωδικοποίηση ( μπορούμε να το κάνουμε αργότερα με το SPSS)
  - Γενικά είναι επιθυμητό να κωδικοποιούμε και να εισάγουμε δεδομένα με, όσο το δυνατό, πιο λεπτομερή τρόπο
  - Η συμπίεση και η αυθαίρετη ομαδοποίηση κατά την κωδικοποίηση των δεδομένων προκαλεί την απώλεια της πληροφορίας



# Κανόνες κωδικοποίησης των δεδομένων

---

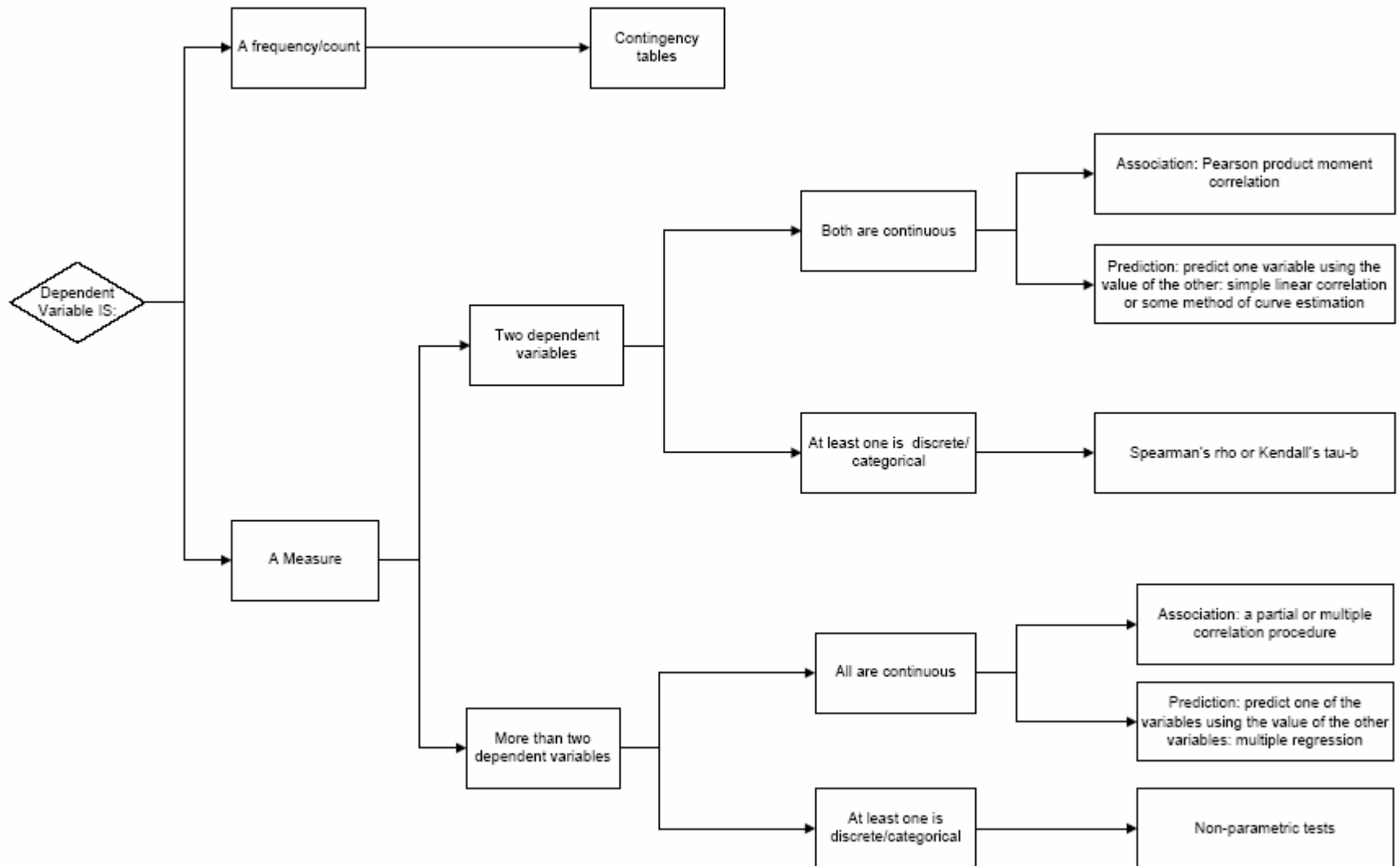
- Για κάθε περίπτωση (case), πρέπει να υπάρχει μια τιμή για κάθε μεταβλητή και αυτή πρέπει να είναι αριθμητική
- Εξαίρεση: οι μεταβλητές στις οποίες έχουμε **χαμένες τιμές (missing values)**
- Συνιστάται οι χαμένες τιμές να αφήνονται κενές (το SPSS είναι σχεδιασμένο να τις αναγνωρίζει)
- Διαφορετικού είδους χαμένες τιμές (π.χ. ερωτήσεις που αφέθηκαν κενές και ερωτήσεις που απαντήθηκαν με μη αποδεκτό τρόπο):
  - Αντιστοιχούμε στις χαμένες τιμές αριθμούς που ξεχωρίζουν (π.χ. 99, 98).
  - Είναι απαραίτητο να το δηλώσουμε στο SPSS

# Κανόνες κωδικοποίησης των δεδομένων

---

- ❑ Οι κανόνες κωδικοποίησης πρέπει να εφαρμόζονται με συνέπεια για όλες τις περιπτώσεις
- ❑ Για μεταβλητές διάταξης συνήθως χρησιμοποιούνται μεγάλοι αριθμοί (κωδικοί) για απαντήσεις όπως «Συμφωνώ», «Καλό» ή «Θετικό»
- ❑ Τα χαμένα δεδομένα προκαλούν προβλήματα στις στατιστικές αναλύσεις. Σε περιπτώσεις όπου κάποια άτομα δεν έχουν απαντήσει σε παραπάνω από τις μισές ερωτήσεις, μπορούμε να απορρίψουμε τα ερωτηματολόγιά τους ως άκυρα
- ❑ Είναι σημαντικό να κατανοούμε τα δεδομένα, δηλαδή να καθορίζουμε τη φύση των μεταβλητών μας

# Μέτρηση σχέσης



# Μέτρηση διαφοράς

