**Chapter 12: Data Processing and Fundamental Data Analysis**

**Multiple Choice**

1. The last step in the data analysis procedure is:

a. coding.

b. machine cleaning of data.

c. tabulation and statistical analysis.

d. interpretation of the results.

e. none of these

Answer: c

1. The process of ascertaining that interviews were actually conducted is known as:

a. validation.

b. coding.

c. machine cleaning

d. data entry

e. checks and balances

Answer: a

1. Checking to determine whether the proper screener questions were invoked would be part of which step in the data analysis procedure?

a. validation

b. coding

c. machine cleaning

d. data entry

e. checks and balances

Answer: a

1. The process of converting information into a form that can be read by a computer is known as \_\_\_\_\_\_\_.

a. scanning

b. coding

c. machine cleaning

d. data entry

e. validation

Answer: d

1. Which type of data entry ensures that data is entered properly?

a. scanning

b. coding

c. intelligent

d. marginal

e. none of these

Answer: c

1. This analytical procedure shows the responses to one question relative to the responses of another question.

a. one-way frequencies

b. cross tabulation

c. t-test

d. z-test

e. none of these

Answer: b

1. Which of the following is the most flexible type of graph?

a. line chart

b. bar chart

c. pie chart

d. scatter plot graph

e. All are equally flexible.

Answer: b

1. Which of the following is not a measure of dispersion?

a. mean

b. standard deviation

c. variance

d. range

e. all of these are measures of dispersion

Answer: a

1. This type of chart is the simplest form of chart and is useful for presenting a measurement taken at several points over time.

a. line chart

b. bar chart

c. pie chart

d. bar chart and pie chart are equally useful

e. All are equally useful.

Answer: a

1. This measure of central tendency can be computed only from interval or ratio data.

a. mode

b. median

c. mean

d. All of these must be computed from interval or ratio data.

e. None of these can be computed from interval or ratio data.

Answer: c

1. Which type of analysis cannot be computed with nominal data?

a. mean

b. mode

c. frequency

d. percentage

e. None can be computed with nominal data

Answer: a

1. If we conclude that the average weekly amount of soft drink consumption by males and females is about the same, but that the standard deviation among males is greater than that for females, what would this mean?

a. Males and females pretty much agree as to how many soft drinks they consume per week.

b. Males consume more soft drinks per week than females.

c. Even though males and females consume, on average, about the same number of soft drinks, there is more diversity of consumption among males.

d. Females are more health conscious than males.

Answer: c

1. Which of the following is not a measure of dispersion?

a. standard deviation

b. median

c. variance

d. range

e. none of these

Answer: b

1. With one-way frequency tables, the base size for a question would *not* take into consideration:

a. total respondents.

b. the number of people asked a particular question.

c. the number of people who answered the question.

d. population size.

e. It would take into account all of these

Answer: d

1. Out of the following, which would typically have the largest base size?

a. total respondents

b. total asked

c. total answering

d. total skipping

Answer: a

1. The quality control aspect of the data analysis procedure refers to:

a. statistical analysis of data.

b. properly training interviewers.

c. validation and editing.

d. coding questionnaires.

e. none of these

Answer: c

1. If an interviewer did not properly follow skip patterns specified in a questionnaire, the \_\_\_\_\_\_\_\_\_\_\_\_\_\_ step in data analysis would reveal the problem.

a. validation

b. editing

c. coding

d. analyzing

e. data cleaning

Answer: a

1. Which of the following does validation typically address?

a. validity of measures used

b. measures of central tendency

c. whether respondents were actually interviewed

d. none of these

Answer: c

1. Finding the value that occurs most frequently determines the\_\_\_\_\_\_\_\_\_\_.

a. standard deviation

b. mode

c. variance

d. mean

e. median

Answer: b

1. Validation of data is an attempt to check the accuracy of the work done by \_\_\_\_\_\_\_.

a. analysts

b. project managers

c. interviewers

d. panel recruiters

e. data entry

Answer: c

1. Which of the following is not part of the coding process regarding open-ended questions?

a. list responses

b. consolidate responses

c. set codes

d. enter codes

e. All are part of the coding process.

Answer: e

1. Probing would typically done by \_\_\_\_\_\_\_\_\_\_\_\_.

a. the client

b. the respondent

c. the interviewer

d. none of these

Answer: c

1. Error checking routines and marginal reports are part of which of the following processes?

a. determination of proper graphical aides in data analysis

b. determination of proper statistical techniques in data analysis

c. machine cleaning of the database

d. validation of questionnaires

e. coding of open-ended questions on questionnaires.

Answer: c

1. The idea behind this type of statistical analysis is to look at the responses to one question in relation to the responses to one or more other questions.

a. one-way frequencies

b. cross tabulations

c. analysis using measures of central tendency

d. analysis using measures of dispersion

e. analysis using line and/or bar charts

Answer: b

1. Which of the following graphic representation is the most visually pleasing for presenting cross tabulation information?

a. multiple row—three dimensional bar chart

b. three dimensional pie chart

c. clustered bar chart

d. stacked bar chart

e. clustered line chart

Answer: a

1. \_\_\_\_\_\_\_\_\_\_\_\_ is the process of ascertaining that questionnaires were filled out properly and completely.

a. Editing

b. Revision

c. Reviewing

d. Checks and Balances

e. None of these

Answer: a

1. \_\_\_\_\_\_\_\_\_\_\_\_ is the process of grouping and assigning numeric codes to the various responses to a question.

a. Alignment

b. Coding

c. Decoding

d. Arrangement

Answer: b

1. \_\_\_\_\_\_\_\_\_\_\_\_\_ is a form of data entry in which responses on questionnaires are read in automatically by the data entry device.

a. Automatic entry

b. Automatic delivery

c. Scanning technology

d. None of these

Answer: c

1. There are some indications that the historically high interest by marketing researchers and their clients is beginning to wane in which of the following statistical analysis methods?

a. one-way frequency tables

b. cross-tabs

c. graphic representations

d. descriptive statistics

e. none of these

Answer: b

1. Which of the following is a text-processing tool for open-ended questions that shows greater prominence to words that appear more frequently in the source text?

a. Word Cloud

b. Semantria

c. TextSmart

d. CATI

e. None of these

Answer: a

**True/False**

1. The coding step pertains to only closed-ended responses on questionnaires.

Answer: False

1. The median can be computed for all types of data except nominal data.

Answer: True

1. The mode of a distribution of responses is the middle response in the distribution, meaning half of the responses are larger and half of the responses are smaller than the mode.

Answer: False

1. Means can be calculated using nominal, ordinal, or interval data.

Answer: False

1. A variable with a large standard deviation would have a large variance as well.

Answer: True

1. Cross tabulations look at differences in mean values.

Answer: False

1. Once a respondent is interviewed in a phone survey, they should never be contacted again regarding that particular study.

Answer: False

1. The first step in the data analysis process is coding.

Answer: False

1. Logical or machine cleaning of data is done with the assistance of computer programs.

Answer: True

1. Error checking routines and marginal reports are two ways of machine cleaning the data.

Answer: True

1. The most complex tabulation is the one-way frequency table.

Answer: False

1. Pie charts are useful for depicting changes over time.

Answer: False

1. Anything that can be shown in a line graph or pie chart also can be shown in a bar chart.

Answer: True

1. The mode is always equal to the mean.

Answer: False

1. Categorical data is typically analyzed with percentages rather than means.

Answer: True

1. Checking to ensure that skip patterns were followed occurs during the data editing stage.

Answer: True

1. Interviewers are often asked to paraphrase the responses to open-ended questions.

Answer: False

1. Data entry must always be done manually.

Answer: False

1. Means are not included in a cross tabulation.

Answer: True

1. The purpose of the first edit, or field edit, is to identify problems when there is still time to contact the respondent again and determine the appropriate answer to questions that were not asked.

Answer: True

1. The TextSmart module of SPSS is not an example of the new breed of automated coding systems.

Answer: False

1. A common way of setting up cross-tabulation is to use columns to represent factors such as demographics and lifestyle characteristics, which are expected to be predictors of the state of mind, behavior, or intentions data, shown as rows of the table.

Answer: True

1. Stacked bar charts are not helpful in graphically representing cross-tabulation results.

Answer: False

1. The process of “validating” online surveys is quite different.

Answer: True

1. A large percentage of online surveys are done using online panels such as Dynata.

Answer: True

1. More and more data are entered manually.

Answer: False

1. Tables are boring and inefficient.

Answer: True

**Essay Questions**

1. Describe the five steps in the analysis process.

Answer:

1. validation and editing,

2. coding,

3. data entry,

4. machine cleaning of data, and

5. tabulation and statistical analysis

1. Using the data below compare and contrast male and female soft drink consumption. You will need to compute the proper measure of central tendency and dispersion to support your answer.

|  |  |
| --- | --- |
| Gender | Weekly Consumption (12 oz. Cans) |
| Male | 5 |
| Female | 2 |
| Female | 3 |
| Male | 8 |
| Female | 4 |
| Male | 9 |
| Male | 7 |
| Female | 3 |
| Female | 2 |
| Female | 1 |
| Male | 7 |
| Male | 4 |

Answer:

|  |  |
| --- | --- |
| Average male consumption | = 6.67 cans per week, |
| Average female consumption | = 2.5 cans per week, |
| Male standard deviation | = 1.86, |
| Female standard deviation | = 1.05, |

Males consume over twice as many soft drinks as females. In addition, there is more diversity of consumption among male drinkers than female drinkers.

1. Illustrate a condition in which a “SKIP PATTERN” is needed.

Answer: Skip patterns are conditional statements that direct a respondent to a particular question sequence given some conditional status. Such as, “If MALE, go to question 3; if FEMALE, go to Question 4.

1. Open-ended questions are often necessary even in surveys whose major objective is to produce quantitative data. Given the open-ended responses below, how would you summarize them for analytical purposes?

QUESTION: How do you think the taste of Pepsi compares with Coca Cola?

Responses: The taste of Pepsi was

1) Better

2) Sweeter

3) Too sweet

4) Not as refreshing

5) More refreshing

6) Sweeter is better

7) Bolder taste

8) They taste about the same

9) I don’t know

10) I don’t have any idea

11) I don’t drink soft drinks

12) Soft drinks are not good for you

13) Sweeter is better

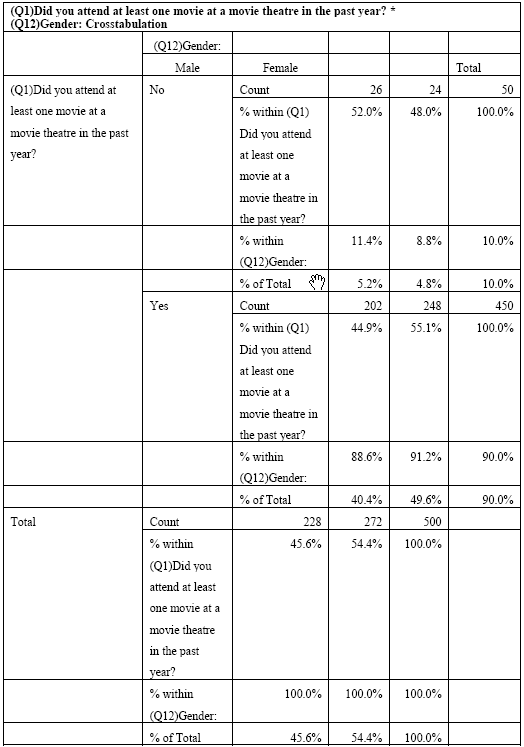
14) Not a refreshing as Coca Cola

15) Coca Cola has always been my favorite brand

Answer:

|  |  |
| --- | --- |
| Consolidated Response Categories | Codes |
| Taste – sweetness | 2, 3, 6 |
| Refreshing | 4, 5, 14 |
| Preferred | 1, 15 |
| Direct Comparison | 7, 8 |
| Don’t know/don’t drink soft drinks | 9, 10, 11 |
| Health statement | 12 |

1. Given the following cross tabulation output, answer the questions below:



**Questions**

1. What percentage of males did not attend at least one movie at a movie theatre in the last year?
2. What percent of all respondents were female and attended at least one movie at a movie theatre?
3. What percent of all respondents that did not go to at least one movie at a movie theatre in the last year were female?

Answer:

1. 11.4%
2. 49.6%
3. 52.0%
4. What is the difference between measurement validity and interview validation?

Answer: Measurement validity refers to the degree to which a questionnaire measures the construct it intends to measure. Interview validation is the process of ensuring that the data recorded on the questionnaire was correctly and accurately collected and recorded.