

Chapter 2 Summarizing & Graphing Data



- A) Overview
- B) Frequency Distributions
- C) Visualizing Data

Overview



❖ Descriptive Statistics

summarize or **describe** the important characteristics of a known set of population data

❖ Inferential Statistics

use sample data to make **inferences (or generalizations)** about a population

Important Characteristics of Data



1. **Center:** A representative or average value that indicates where the middle of the data set is located
2. **Variation:** A measure of the amount that the values vary among themselves
3. **Distribution:** The nature or shape of the distribution of data (such as bell-shaped, uniform, or skewed)
4. **Outliers:** Sample values that lie very far away from the vast majority of other sample values
5. **Time:** Changing characteristics of the data over time

Frequency Distributions



❖ Frequency Distribution

lists data values (either individually or by groups of intervals), along with their corresponding frequencies or counts

Table 2-1 Measured Cotinine Levels in Three Groups

Smoker: The subjects report tobacco use.

ETS: (Environmental Tobacco Smoke) Subjects are nonsmokers who are exposed to environmental tobacco smoke ("secondhand smoke") at home or work.

NOETS: (No Environmental Tobacco Smoke) Subjects are nonsmokers who are not exposed to environmental tobacco smoke at home or work. That is, the subjects do not smoke and are not exposed to secondhand smoke.

Smoker:	1	0	131	173	265	210	44	277	32	3
	35	112	477	289	227	101	222	149	113	401
	130	234	164	198	17	253	87	121	266	290
	123	167	250	245	48	86	284	1	208	173
ETS:	384	0	69	19	1	0	178	2	13	1
	4	0	543	17	1	0	51	0	197	3
	0	3	1	45	13	3	1	1	1	0
	0	551	2	1	1	1	0	74	1	241
NOETS:	0	0	0	0	0	0	0	0	0	0
	0	9	0	0	0	0	0	0	244	0
	1	0	0	0	90	1	0	309	0	0
	0	0	0	0	0	0	0	0	0	0

Table 2-2
Frequency Distribution
of Cotinine Levels
of Smokers

Cotinine	Frequency
0-99	11
100-199	12
200-299	14
300-399	1
400-499	2

Frequency Distributions



Definitions

Copyright © 2004 Pearson Education, Inc.

Lower Class Limits



are the smallest numbers that can actually belong to different classes

Cotinine	Frequency
0-99	11
100-199	12
200-299	14
300-399	1
400-499	2

Copyright © 2004 Pearson Education, Inc.

Lower Class Limits



are the smallest numbers that can actually belong to different classes

Lower Class Limits

Cotinine	Frequency
0-99	11
100-199	12
200-299	14
300-399	1
400-499	2

Copyright © 2004 Pearson Education, Inc.

Upper Class Limits



are the largest numbers that can actually belong to different classes

Upper Class Limits

Cotinine	Frequency
0-99	11
100-199	12
200-299	14
300-399	1
400-499	2

Copyright © 2004 Pearson Education, Inc.

Class Boundaries



are the numbers used to separate classes, but without the gaps created by class limits

Copyright © 2004 Pearson Education, Inc.

Class Boundaries



number separating classes

Cotinine	Frequency	
- 0.5	0-99	11
99.5	100-199	12
199.5	200-299	14
299.5	300-399	1
399.5	400-499	2
499.5		

Copyright © 2004 Pearson Education, Inc.

Class Boundaries

number separating classes

	Cotinine	Frequency
-0.5	0-99	11
99.5	100-199	12
199.5	200-299	14
299.5	300-399	1
399.5	400-499	2
499.5		

Class Boundaries (indicated by arrows pointing to the values in the table)

Copyright © 2004 Pearson Education, Inc.

Class Midpoints

midpoints of the classes

Class midpoints can be found by adding the lower class limit to the upper class limit and dividing the sum by two.

Copyright © 2004 Pearson Education, Inc.

Class Midpoints

midpoints of the classes

	Cotinine	Frequency
0	99	11
100	199	12
200	299	14
300	399	1
400	499	2

Class Midpoints (indicated by arrows pointing to the values in the table)

Copyright © 2004 Pearson Education, Inc.

Class Width

is the difference between two consecutive lower class limits or two consecutive lower class boundaries

	Cotinine	Frequency
100	0-99	11
100	100-199	12
100	200-299	14
100	300-399	1
100	400-499	2

Class Width (indicated by arrows pointing to the values in the table)

Copyright © 2004 Pearson Education, Inc.

Reasons for Constructing Frequency Distributions

1. Large data sets can be summarized.
2. Can gain some insight into the nature of data.
3. Have a basis for constructing graphs.

Copyright © 2004 Pearson Education, Inc.

Constructing A Frequency Table

1. Decide on the number of classes (should be between 5 and 20) .
2. Calculate (round up).

$$\text{class width} \approx \frac{(\text{highest value}) - (\text{lowest value})}{\text{number of classes}}$$
3. Starting point: Begin by choosing a lower limit of the first class.
4. Using the lower limit of the first class and class width, proceed to list the lower class limits.
5. List the lower class limits in a vertical column and proceed to enter the upper class limits.
6. Go through the data set putting a tally in the appropriate class for each data value.

Copyright © 2004 Pearson Education, Inc.

Relative Frequency Distribution

Slide 19

$$\text{relative frequency} = \frac{\text{class frequency}}{\text{sum of all frequencies}}$$

Copyright © 2004 Pearson Education, Inc.

Relative Frequency Distribution

Slide 20

Cotinine	Frequency
0-99	11
100-199	12
200-299	14
300-399	1
400-499	2

Total Frequency = 40

Table 2-3
Relative Frequency Distribution of Cotinine Levels in Smokers

Cotinine	Relative Frequency
0-99	28%
100-199	30%
200-299	35%
300-399	3%
400-499	5%

11/40 = 28%

12/40 = 40%

etc.

Copyright © 2004 Pearson Education, Inc.

Cumulative Frequency Distribution

Slide 21

Cotinine	Frequency
0-99	11
100-199	12
200-299	14
300-399	1
400-499	2

Table 2-4
Cumulative Frequency Distribution of Cotinine Levels in Smokers

Cotinine	Cumulative Frequency
Less than 100	11
Less than 200	23
Less than 300	37
Less than 400	38
Less than 500	40

Cumulative Frequencies

Copyright © 2004 Pearson Education, Inc.

Frequency Tables

Slide 22

Table 2-2
Frequency Distribution of Cotinine Levels of Smokers

Cotinine	Frequency
0-99	11
100-199	12
200-299	14
300-399	1
400-499	2

Table 2-3
Relative Frequency Distribution of Cotinine Levels in Smokers

Cotinine	Relative Frequency
0-99	28%
100-199	30%
200-299	35%
300-399	3%
400-499	5%

Table 2-4
Cumulative Frequency Distribution of Cotinine Levels in Smokers

Cotinine	Cumulative Frequency
Less than 100	11
Less than 200	23
Less than 300	37
Less than 400	38
Less than 500	40

Copyright © 2004 Pearson Education, Inc.

Visualizing Data

Slide 23

Depict the nature of shape or shape of the data distribution

Copyright © 2004 Pearson Education, Inc.

Histogram

Slide 24

A bar graph in which the horizontal scale represents the classes of data values and the vertical scale represents the frequencies.

Cotinine	Frequency
0-99	11
100-199	12
200-299	14
300-399	1
400-499	2

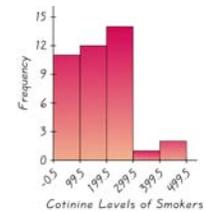


Figure 2-1

Copyright © 2004 Pearson Education, Inc.

Relative Frequency Histogram



Has the same shape and horizontal scale as a histogram, but the vertical scale is marked with relative frequencies.

Cotinine	Relative Frequency
0-99	28%
100-199	30%
200-299	35%
300-399	3%
400-499	5%

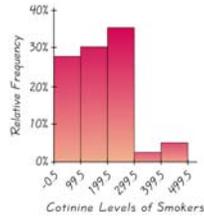


Figure 2-2

Copyright © 2004 Pearson Education, Inc.

Histogram and Relative Frequency Histogram

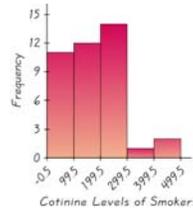


Figure 2-1

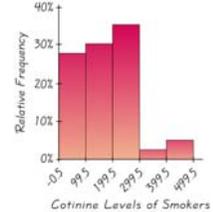


Figure 2-2

Copyright © 2004 Pearson Education, Inc.

Frequency Polygon



Uses line segments connected to points directly above class midpoint values

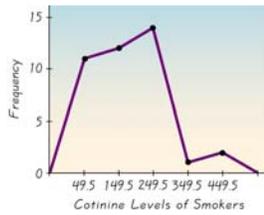


Figure 2-3

Copyright © 2004 Pearson Education, Inc.

Ogive



A line graph that depicts cumulative frequencies

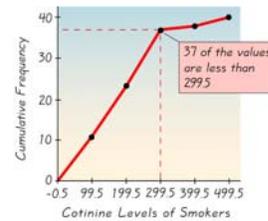


Figure 2-4

Copyright © 2004 Pearson Education, Inc.

Stem-and-Leaf Plot



Represents data by separating each value into two parts: the stem (such as the leftmost digit) and the leaf (such as the rightmost digit)

Stem-and-Leaf Plot

Stem (tens)	Leaves (units)	
6	449	← Values are 64, 64, 69.
7	0111233444455555666778899	
8	0011122233346899	
9	0024	
10		
11		
12	0	← Value is 120.

Copyright © 2004 Pearson Education, Inc.

Pareto Chart



A bar graph for qualitative data, with the bars arranged in order according to frequencies

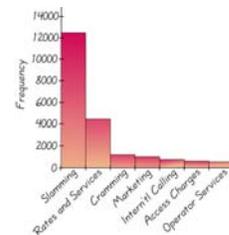


Figure 2-6

Copyright © 2004 Pearson Education, Inc.

Pie Chart

Slide 31

A graph depicting qualitative data as slices of a pie

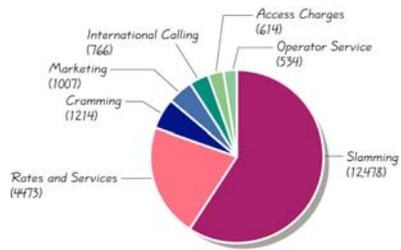


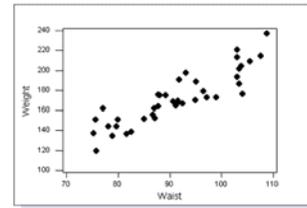
Figure 2-7

Copyright © 2004 Pearson Education, Inc.

Scatter Diagram

Slide 32

A plot of paired (x,y) data with a horizontal x-axis and a vertical y-axis



Copyright © 2004 Pearson Education, Inc.