


Chapter 2 Frequency Distributions and Graphs

Frequency Distribution & Graphs



Descriptive Statistics

Chapter 2 - 1

Descriptive Statistics

- Grouping tables and graphical summary.
- Numerical Summary.

Exploratory Data Analysis

Chapter 2 - 2

Display Distributions (Understand the variability)

Chapter 2 - 3

What type of statistical technique is appropriate?

- Categorical variable?
- Quantitative variable?

Chapter 2 - 4

Display Distributions for Continuous or Quantitative Variables: *Histograms*

Chapter 2 - 5

(A complete list)

ID	Height	Weight	BirthMonth	Exp.	Gender
1	6	135	4		H F
2	63	119	9		H F
3	72	175	11		T M
4	60	106	9		H F
5	65	135	8		T F
6	72	170	10		H M
7	64	180	8		H F
8	71	205	10		H M
9	75	195	6		T M
10	71	185	8		H M
11	71	182	6		T M
12	65	108	8		T F
13	73	150	4		H M
14	67	128	6		T F
15	74	175	6		H M
16	66	160	9		H F
17	65	143	9		T F
18	72	190	11		T M
19	64	180	2		H M
20	61	195	5		T M
21	72	220	7		H M
22	69	285	7		H M

Chapter 2 - 6

Chapter 2 Frequency Distributions and Graphs

Frequency Distribution Table

(From data sheet)

Class	Frequency	Relative Freq.	Cumulative R.F.
90 - <110	2	2/22 = .091	2/22
110 - <130	2	2/22 = .091	4/22
130 - <150	3	3/22 = .136	7/22
150 - <170	2	2/22 = .091	9/22
170 - <190	7	7/22 = .318	16/22
190 - <210	4	4/22 = .182	20/22
210 - <230	1	1/22 = .045	21/22
230 - <250	0	0/22 = .000	21/22
250 - <270	0	0/22 = .000	21/22
270 - <290	1	1/22 = .045	22/22
Total	22	1.000	

Chapter 2 - 7

Classes: Categories for grouping data.

Frequency (class frequency): The number of data values in a class.

Relative frequency: The ratio of the frequency of a class to the total number of pieces of data.

Frequency distribution: A listing of classes and their frequencies.

Relative Frequency distribution: A listing of classes and their relative frequencies.

Upper class limit: The largest value that can go in a class.

Lower class limit: The smallest value that can go in a class.

Class width: The difference between the lower class limit of the given class and the lower class limit of the next higher class.

Class midpoint (class mark): The midpoint of a class.

Chapter 2 - 8

Guidelines for grouping data: (for quantitative variable)

- There should be between five and twenty classes.
- Each piece of data must belong to one, and only one, class. (Mutually Exclusive)
- Whenever feasible, all classes should have the same width.

Chapter 2 - 9

To build a Frequency Table:

- Find the range of the data:
Range = Largest value – smallest value
- Use the **range** and **try different class width** to determine how many classes you need to make frequency table or histogram.

Student data example:

$$\text{Range} = 285 - 106 = 179/20 \approx 9$$

If using a class width of 20, there'll be about 9 classes which is good.

Chapter 2 - 10

Frequency Distribution Table

Data: ~~135~~, ~~119~~, ~~175~~, ~~106~~, ~~135~~, ~~170~~, ~~180~~, ~~205~~, ~~195~~, ...

Class	Tally	Frequency
90 - <110		
110 - <130		
130 - <150		
150 - <170		
170 - <190		
190 - <210		
210 - <230		
230 - <250		
250 - <270		
270 - <290		
Total		

270 to less than 290

Chapter 2 - 11

Frequency Distribution Table

Data: 135, 119, 175, 106, 135, 170, 180, 205, 195, ...

Class	Tally	Frequency
90 - <110		2
110 - <130		2
130 - <150		3
150 - <170		2
170 - <190		7
190 - <210		4
210 - <230		1
230 - <250		0
250 - <270		0
270 - <290		1
Total		22

Chapter 2 - 12

Chapter 2 Frequency Distributions and Graphs

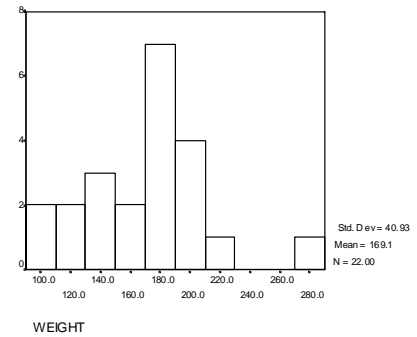
Frequency Distribution Table

(From data sheet)

Class	Frequency	Relative Freq.	Cumulative R.F.
90 - <110	2	2/22 = .091	2/22
110 - <130	2	2/22 = .091	4/22
130 - <150	3	3/22 = .136	7/22
150 - <170	2	2/22 = .091	9/22
170 - <190	7	7/22 = .318	16/22
190 - <210	4	4/22 = .182	20/22
210 - <230	1	1/22 = .045	21/22
230 - <250	0	0/22 = .000	21/22
250 - <270	0	0/22 = .000	21/22
270 - <290	1	1/22 = .045	22/22
Total	22	1.000	

Chapter 2 - 13

Histogram (SPSS)



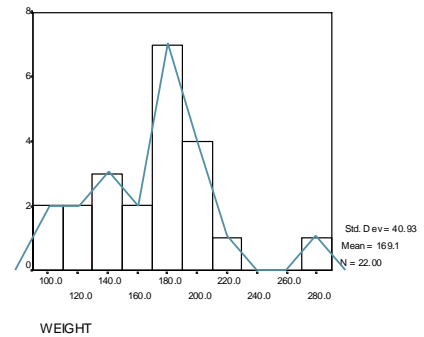
4

What to observe in Histograms?

- **Outliers:** observations that stand out from the rest for some reason.
- **Center:** the “middle” of the data.
- **Spread:** the range; the extent of the data; how far the values are from each other.
- **Shape:** distribution pattern. [Skewness, symmetry, uniform, ...]

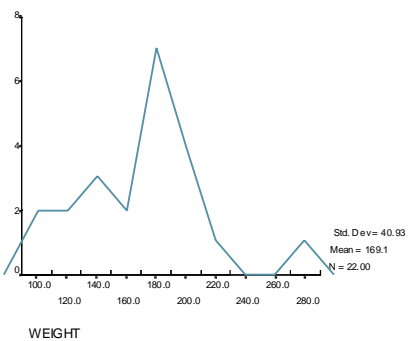
Chapter 2 - 15

Polygon (SPSS)



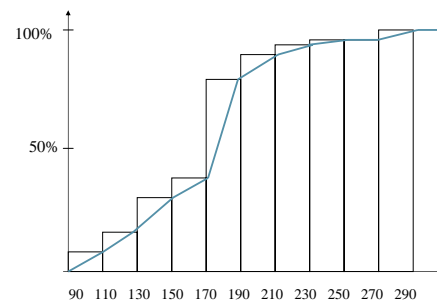
6

Polygon (SPSS)



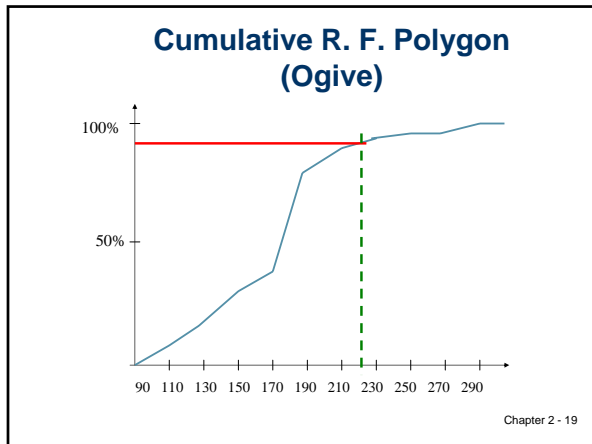
7

Cumulative R. F. Histogram



Chapter 2 - 18

Chapter 2 Frequency Distributions and Graphs

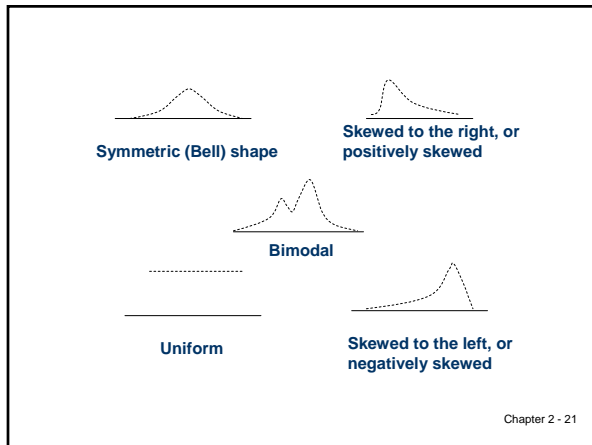


Frequency Distribution Table

(From data sheet with different boundaries)

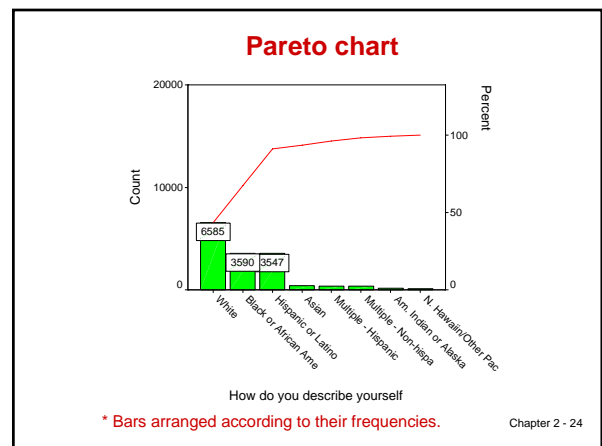
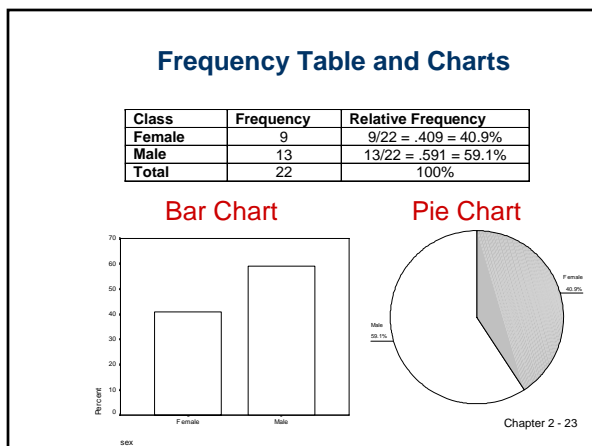
Class	Frequency	Relative Freq.	Cumulative R.F.
90< - 110	2	2/22 = .091	2/22
110< - 130	2	2/22 = .091	4/22
130< - 150	4	4/22 = .182	8/22
150< - 170	2	2/22 = .091	10/22
170< - 190	7	7/22 = .318	17/22
190< - 210	3	3/22 = .136	20/22
210< - 230	1	1/22 = .045	21/22
230< - 250	0	0/22 = .000	21/22
250< - 270	0	0/22 = .000	21/22
270< - 290	1	1/22 = .045	22/22
Total	22	1.000	

Chapter 2 - 20



Grouping and Displaying Categorical Data

Chapter 2 - 22



Chapter 2 Frequency Distributions and Graphs

Examine Bivariate Data

Chapter 2 - 25

Two Categorical Variables

Contingency Table

	Cancer	No cancer	Row Total
Smoker	20	30	50
Non- Smoker	5	45	50
Column Total	25	75	100

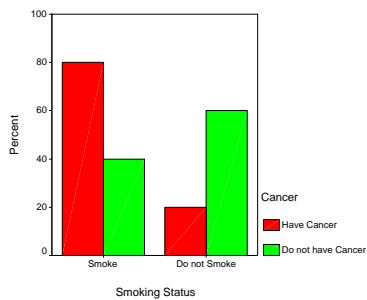
Odds of smoker to have cancer: $20/30 = 6/9$

Odds of nonsmoker to have cancer: $5/45 = 1/9$

Odds Ratio = $(6/9)/(1/9) = 6$

Chapter 2 - 26

Cluster bar chart



Chapter 2 - 27

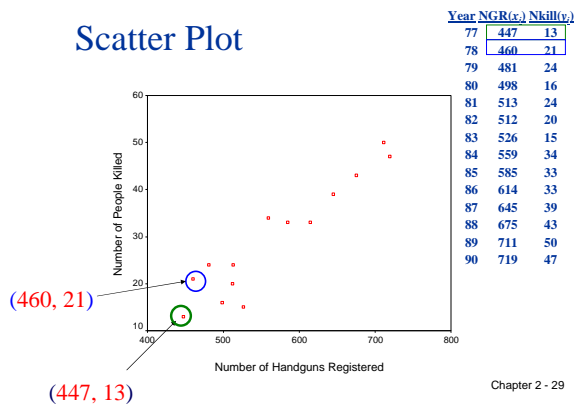
Relation Between Two Quantitative Variables

Is there relation between “number of handguns registered” in the area and “number of people killed by guns”?

Year	NGR(x)	Nkill(y)
77	447	13
78	460	21
79	481	24
80	498	16
81	513	24
82	512	20
83	526	15
84	559	34
85	585	33
86	614	33
87	645	39
88	675	43
89	711	50
90	719	47

Chapter 2 - 28

Scatter Plot



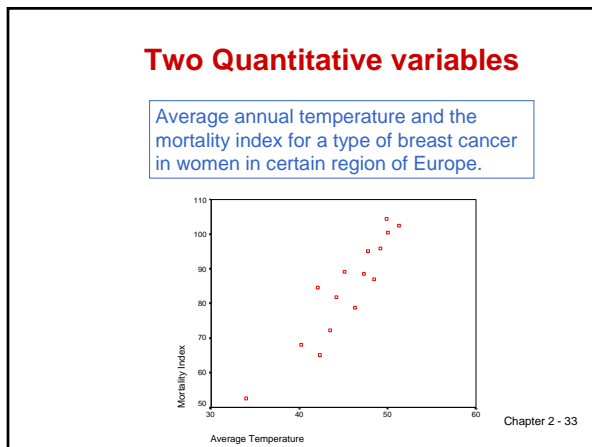
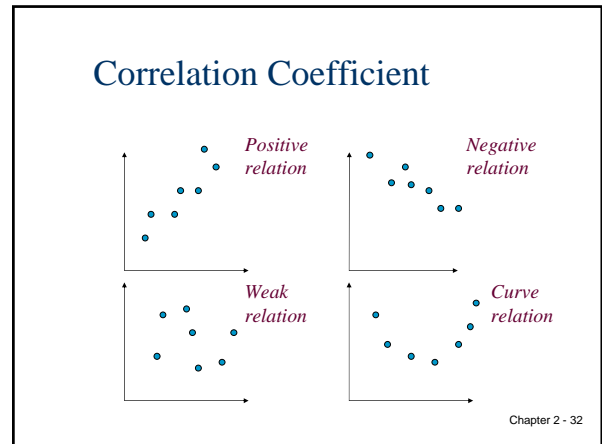
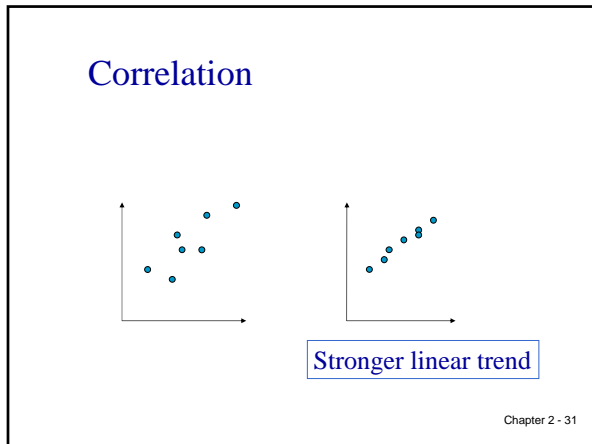
Chapter 2 - 29

Information in Scatter Plot

- Form (Straight line or curve relation)
- Direction (Positive or negative relation)
- Strength (Strong or weak relation)
- Outliers

Chapter 2 - 30

Chapter 2 Frequency Distributions and Graphs

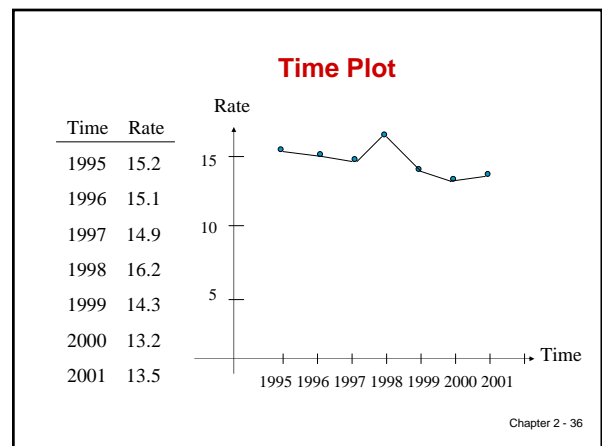
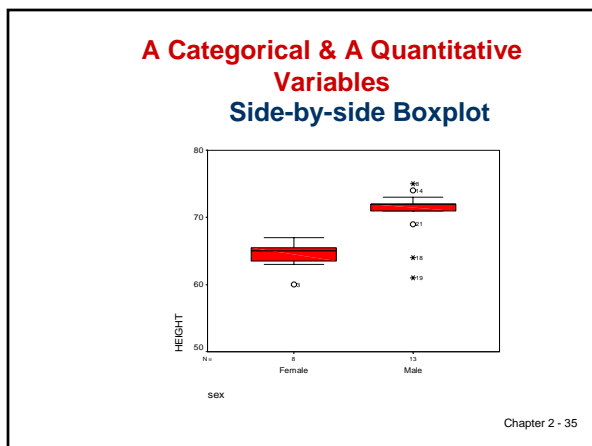


Two Quantitative variables

Data:

Temperature	Mortality Index
34	52
40	68
42	63
42	83
43	72
44	81
45	89
46	77
47	88
48	94
49	86
50	95
51	105
51	100
52	102

Chapter 2 - 34



Chapter 2 Frequency Distributions and Graphs

