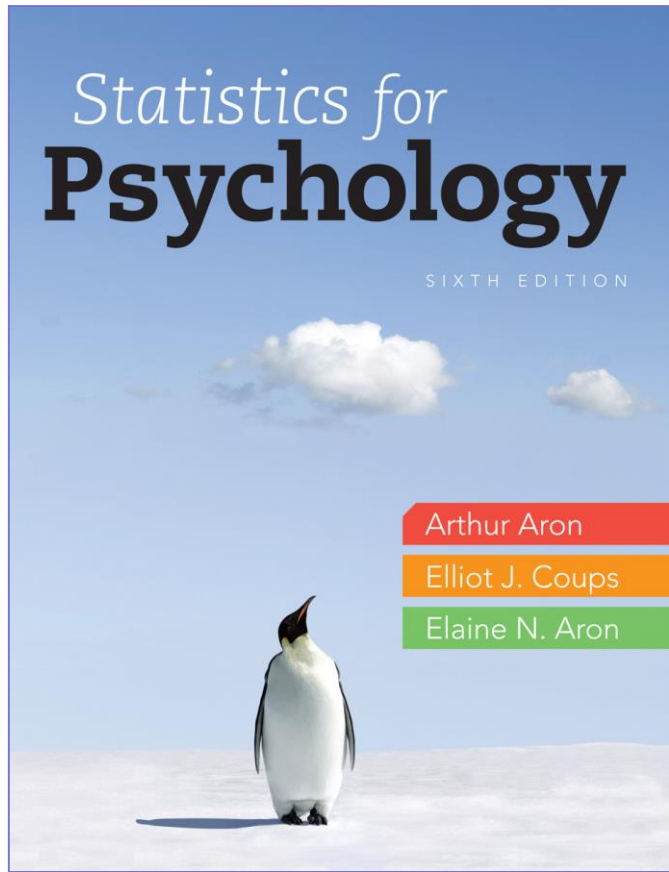# 1

Displaying the Order in a Group of Numbers Using Tables and Graphs

Statistics (stats)

- is a branch of mathematics that focuses on the organization, analysis, and interpretation of a group of numbers.
- is a method of pursuing the truth ☺
- Psychologists use statistical methods to help them make sense of the numbers they collect when conducting research
- The best way of understanding stats is to do the procedures by hand for a while

Two Branches of Statistical Methods

- Descriptive statistics
  - Summarize/organize scores from a research study

- Inferential statistics
  - Draw conclusions/make inferences that go beyond the scores from a research study

Basic Concepts

- Variable
  - A characteristic that can have different values; a characteristics that can vary

- Value
  - A possible number or category that a score can have

- Score
  - A particular person's value on a variable

Basic Concepts

- Research: Suppose that I am interested in your stress level during the last 2,5 weeks

- On a scale of 0 to 10

- Variable?

- Values?

- Score?

Example

- Variable: Age
- Value: Any positive number
- A 20 year old has a score of 20
- Thus, this person's score of 20 is one of many possible values of age

- Your examples?

Levels of measurement

- The scores are numbers that tell you how much there is of what is being measured
- 0-10 rating for stress
- What do the numbers correspond?
- 0 = never; 10 = always
  - The higher the number, the more stress there is
- 0 = always; 10 = never
  - The higher the number, the less stress

Levels of Measurement

A) Numeric (quantitative) variable

1) Equal-interval variable:

- A variable in which the numbers stand for approximately equal amounts of what is being measured
- Eg. GPA
  - Difference b/w 2.5-2.8 means about as much as the difference b/w 3.0-3.3
- Eg. 0-10 stress ratings
  - Difference b/w 4 and 6 means about as much as the difference b/w 7 and 9

Levels of Measurement

1) Equal-interval variable (Ex.: GPA)

1a) Ratio scale

- Some equal-interval variables are measured on what is called a ratio scale
- An equal-interval variable is measured on a ratio scale **if it has an absolute zero point.**
- Most counts or accumulations of things use a ratio scale
- Eg. Number of children
- For these variables, one can make statements a/ the difference in magnitude b/w values
- Eg. A person w/ 4 siblings has twice as many siblings as a person w/ 2 siblings.
- Other examples: distance, time and weight

Levels of Measurement

A) Numeric (quantitative) variable

2) Rank-order (ordinal) variables

- A variable in which the numbers stand only for relative ranking
- Other name: Ordinal (sıra gösteren)
- E.g., A student's standing in his/her graduating class
- A rank-order variable provides less info than an equal-interval variable
- The difference from one rank to the next doesn't tell the exact difference in amount of what is being measured
- ADV.: When rating how much you like your friends from 0-10, you can give same scores for several friends. However if you rank order them; this is avoided

Levels of Measurement

B) Nominal (categorical) variables

- Not a numeric variable at all.
- A variable in which the values are names or categories
- Values are categories; so its other name is categorical variable
- Eg. Gender
- Values: male and female
- A person's score on the gender variable is one of these values
- Eg. Psychiatric diagnosis
- A person's score on the Psychiatric diagnosis variable could be any diagnosis such as major depression, PTSD, schizophrenia, OCD

The Characteristics of the Scores

Types of Data

- Broadly speaking, there are two types of data
  - 1) Nominal and 2) Numerical or score
- Nominal variables are those in which a case falls into one of two or more categories
  - Gender, eye colour, socio-economic status, nationality etc.
  - Nominal variables are also referred to as categorical or qualitative variables

Types of Data

- Numerical variables are those in which a case is assigned a numerical value
  - Age, height, weight, IQ, test scores, income, distance or temperature
  - Numerical variables are also referred to as score or quantitative variables

Nominal or Numerical?

- Which type of data are each of the following?

- Facebook user (i.e. yes or no)

- Use of Facebook (i.e. hours per day)

- Network (i.e. number of friends on Facebook)

- How do you access Facebook (i.e. mobile phone, PC or laptop)

- Use of Facebook (i.e. Less than five hours per day or at least five hours per day)

Measurement Scales

– A *nominal scale* does not indicate an amount; rather, it is used for identification, as a name.

– An *ordinal scale* indicates rank order. There is not an equal unit of measurement separating each score.

Measurement Scales (cont'd)

– An *interval scale* indicates an actual quantity and there is an equal unit of measurement separating adjacent scores. Interval scales do not have a "true" 0.

– A *ratio scale* reflects the true amount of the variable that is present because the scores measure an actual amount, there is an equal unit of measurement, and 0 truly means that zero amount of the variable is present.

Summary of Measurement Scales

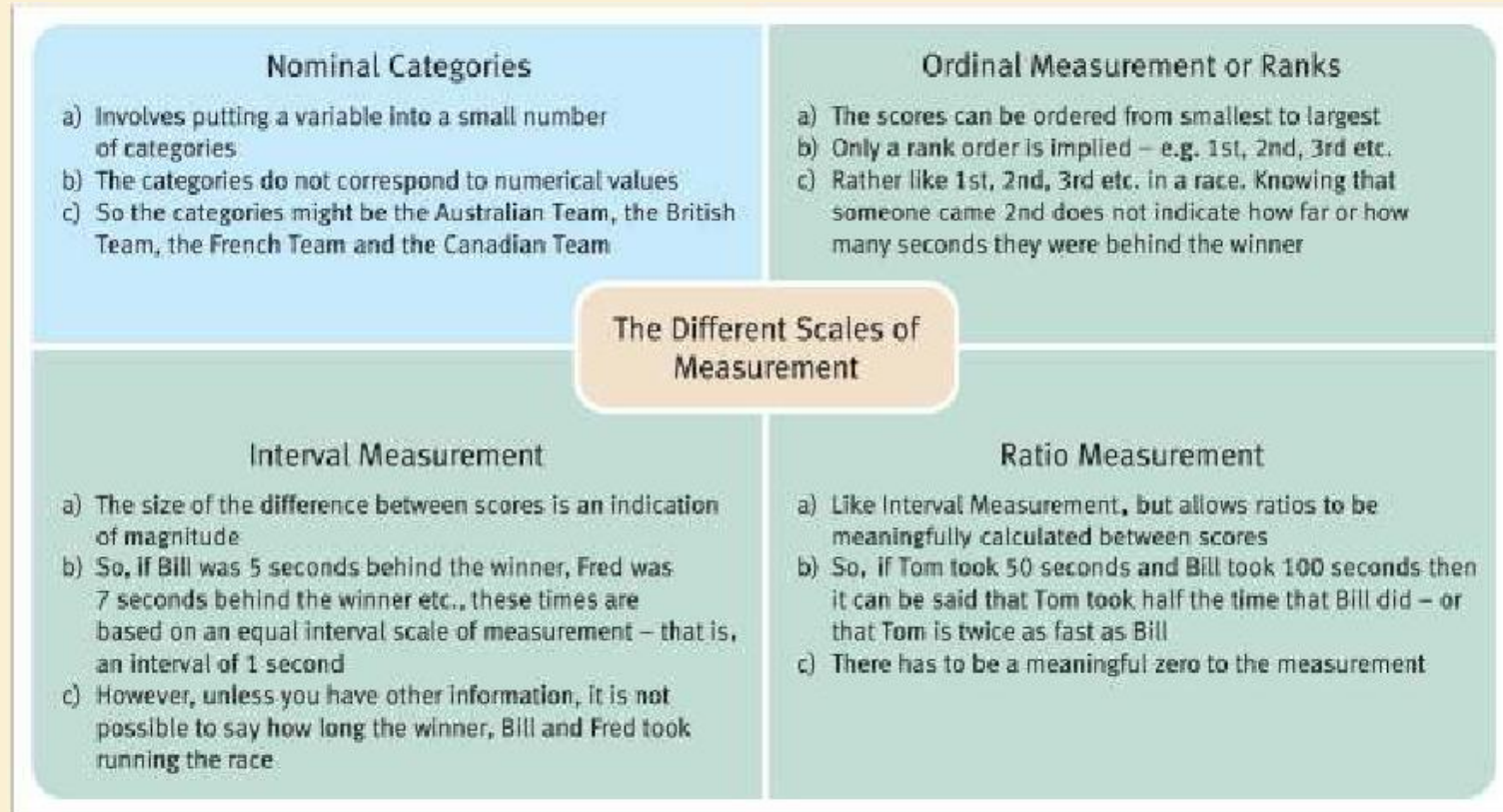|	Type of Measurement Scale			
	Nominal	*Ordinal*	*Interval*	*Ratio*
What Does the Scale Indicate?	Quality	Relative quantity	Quantity	Quantity
Is There an Equal Unit of Measurement?	No	No	Yes	Yes
Is There a True Zero?	No	No	No	Yes
How Might the Scale be Used in Research?	To identify males and females as 1 and 2	To judge who is 1st, 2nd, etc., in aggressiveness	To convey the results of intelligence and personality tests	To count the number of correct answers on a test
Additional Examples	Telephone numbers Social Security numbers	Letter grades Elementary school grade	Checkbook balance Individual's standing relative to class average	Weight Distance traveled

Measurement Characteristics

The Different Scales of Measurement

Nominal Categories

a) Involves putting a variable into a small number of categories

b) The categories do not correspond to numerical values

c) So the categories might be the Australian Team, the British Team, the French Team and the Canadian Team

Ordinal Measurement or Ranks

a) The scores can be ordered from smallest to largest

b) Only a rank order is implied – e.g. 1st, 2nd, 3rd etc.

c) Rather like 1st, 2nd, 3rd etc. in a race. Knowing that someone came 2nd does not indicate how far or how many seconds they were behind the winner

Interval Measurement

a) The size of the difference between scores is an indication of magnitude

b) So, if Bill was 5 seconds behind the winner, Fred was 7 seconds behind the winner etc., these times are based on an equal interval scale of measurement – that is, an interval of 1 second

c) However, unless you have other information, it is not possible to say how long the winner, Bill and Fred took running the race

Ratio Measurement

a) Like Interval Measurement, but allows ratios to be meaningfully calculated between scores

b) So, if Tom took 50 seconds and Bill took 100 seconds then it can be said that Tom took half the time that Bill did – or that Tom is twice as fast as Bill

c) There has to be a meaningful zero to the measurement

Figure 1.1   The different scales of measurement and their main characteristics

Classify the variables in the study as either a score variable or a nominal (category) variable.

- A forensic psychologist is interested in the relationship between geographical factors and criminal victimization. Information is collected from a substantial sample of participants on the following variables.
- Gender
  1 = female; 2 = male
- Age
  recorded as years
- Marital status
  1 = married; 2 = divorced; 3 = separated; 4 = other
- Social class
  1 = working class; 2 = middle class; 3 = upper middle class; 4 = unemployed
- Number of adults (18+) in the household
  number of adults
- Number of children in the household
  number of children

Classify the variables in the study as either a score variable or a nominal (category) variable.

- Geographical location

  1 = inner city council accommodation; 2 = inner city private rented accommodation; 3 = inner city owner occupied; 4 = suburban council accommodation; 5 = suburban private rented accommodation; 6 = suburban owner occupied; 7 = rural council accommodation; 8 = rural private rented accommodation; 9 = rural owner occupied

- Number of times victimized in last year
  number of times

- Number of times victimized by burglary in last year
  number of times

- Number of times victimized by sexual offences in last year
  number of times

- Number of victimizations reported to the police in last year
  number of times

- Whether respondent charged with any offences in last year
  Yes = 1; No = 2

Classify the variables in the study as either a score variable or a nominal (category) variable.

- Type of prison
    - 1=secure; 2=open
- Offence type
    - 1 = violent; 2 = sexual; 3 = property; 4 = mixed
- Offender's age on leaving prison in years
- Length of current sentence before release in years
- Whether on any educational programme
    - 1 = yes; 2 = no
- Whether on any treatment programme
    - 1 = yes; 2 = no
- Offender expresses regret at offending
    - 1 = yes; 2 = no
- Offender was a lone offender
    - 1 = lone offender; 2 = with others
- Whether or not re-convicted for same type of offence (sexual, violent or property) within five years of release
    - 1 = yes; 2 = no

Classify the variables in the study as either a score variable or a nominal (category) variable.

- Number of years before being re-convicted during 5 years following release for a violent crime
  5 years indicates that offender was not re-convicted during study period
- Number of times imprisoned prior to current sentence
  number reported
- Total amount of time imprisoned prior to current sentence
  in years
- Total number of times ever arrested
  number of times
- Offender has job on leaving prison
  1 = yes; 2 = no
- Anxiety score on psychological test on entering prison
  the higher the score the more anxious
- Anxiety score on psychological test shortly before release
  the higher the score the more anxious
- Intelligence in IQ points
  the higher the number the more intelligent the offender

Levels of Measurement

- Why to care about these types of variables?
- The level of measurement selected affects the type of statistics that can be used w/ a variable

Discrete Variables vs. Continuous Variables

- Discrete variables have specific values and cannot have values between these specific values
- Eg. The number of times you went to dentist in the last 12 months
- Answer could be 0,1,2,3 but cannot be 1.65 times
- Nominal variables can also be considered as discrete variables
  - Eg. Gender, religious affiliation, college major (department)

Discrete Variables vs. Continuous Variables

- Continuous variables can, in theory, have an infinite number of values between any two values

- Eg. Age
  - 19 years old
  - 19,25 years old

- Other examples: height, weight and time.

FREQUENCY DISTRIBUTIONS AND PERCENTILES

Frequency Tables

- Provide a listing of individuals having each of the different values for a particular variable.

- Frequency table gives the number of people who used each particular rating.

- Eg.
  - Stress ratings of 30 students: 8,7,4,10,8,6,8,9,9,7,3,7,6,5,0,9,10,7,7,3,6,7,5,2,1,6,7,10,8,8

Steps for Making a Frequency Table

1. Make a list down the page of each possible value, from lowest to highest

2. Go one by one through the scores, making a mark for each next to its value on the list

3. Make a table showing how many times each value on the list is used

4. Figure the percentage of scores for each value

A Frequency Table

Table 1-3 Frequency Table of Number of Students Rating Each Value of the Stress Scale

Stress Rating	Frequency	Percent
0	1	3.3
1	1	3.3
2	1	3.3
3	2	6.7
4	1	3.3
5	2	6.7
6	4	13.3
7	7	23.3
8	5	16.7
9	3	10.0
10	3	10.0

Source: Data based on Aron et al. (1995).

Uses of Frequency Tables

- Describes the data
- Makes the pattern of the data clear
  - Which numbers are given more and which are given less
- Shows how many scores there were for each value on the scale

Grouped Frequency Tables

- Sometimes there are so many possible values that an ordinary frequency table is too awkward to give a simple picture of the scores

- Solution? Make combined categories

- Combined category = interval (aralık)

- A grouped frequency table is a frequency table using intervals based on values.

Grouped Frequency Tables

- Eg. Lets make Table 1-3 (on previous slides) as a grouped frequency table
- The intervals will be like 0-1, 2-3, 4-5, 6-7, 8-9, 10-11

Stress rating interval	Frequency	Percent
0-1	2	6.7
2-3	3	10.0
4-5	3	10.0
6-7	11	36.7
8-9	8	26.7
10-11	3	10.0

Grouped Distributions

Grouped distribution: scores are combined to form small groups

– we report the total *f*, *rel. f*, or *cf* of each group

A Grouped Distribution

Score	f	rel. f	cf	Percentile
40–44	2	.08	25	
35–39	2	.08	23	
30–34	0	.00	21	
25–29	3	.12	21	
20–24	2	.08	18	
15–19	4	.16	16	
10–14	1	.04	12	
5–9	4	.16	11	
0–4	7	.28	7	

An Example:
Grouped Frequency Distribution

- Record the limits of all class intervals, placing the interval containing the highest score value at the top.

- Count up the number of scores in each interval.

Hotel Rates	Frequency
800-899	1
700-799	4
600-699	2
500-599	0
400-499	6
300-399	8
200-299	8
100-199	4
0-99	2

Las Vegas Hotel Rates

52	205	282	325	417	732
76	250	283	373	422	749
100	257	303	384	472	750
136	264	313	384	480	791
186	264	317	400	643	891
196	280	317	402	693	

Frequency Table Guidelines

- Intervals should not overlap, so no score can belong to more than one interval.

- Make all intervals the same width.

- Make the intervals continuous throughout the distribution (even if an interval is empty).

- Choose a convenient interval width.

Hotel Rates	Frequency
800-899	1
700-799	4
600-699	2
500-599	0
400-499	6
300-399	8
200-299	8
100-199	4
0-99	2

An Example: Grouped Frequency Distribution

- Proportion (Relative Frequency)
  - Divide frequency of each class by total frequency.

Hotel Rates	Frequency	Proportion
800-899	1	.03
700-799	4	.11
600-699	2	.06
500-599	0	0
400-499	6	.17
300-399	8	.23
200-299	8	.23
100-199	4	.11
0-99	2	.06

N = 35

An Example: Grouped Frequency Distribution

- Percentage
  - Proportion *100

Hotel Rates	Frequency	Proportion	Percent
800-899	1	.03	3
700-799	4	.11	11
600-699	2	.06	6
500-599	0	0	0
400-499	6	.17	17
300-399	8	.23	23
200-299	8	.23	23
100-199	4	.11	11
0-99	2	.06	6

An Example: Grouped Frequency Distribution

- Cumulative Frequency
  - Shows total number of observations in each class and all lower classes.

Hotel Rates	Frequency	Proportion	Percent	Cumulative Frequency
800-899	1	.03	3	35
700-799	4	.11	11	34
600-699	2	.06	6	30
500-599	0	0	0	28
400-499	6	.17	17	28
300-399	8	.23	23	22
200-299	8	.23	23	14
100-199	4	.11	11	6
0-99	2	.06	6	2

An Example: Grouped Frequency Distribution

- Cumulative Proportion (Cumulative Relative Frequency):
  - Divide Cumulative Frequency by Total Frequency
- Percentile Rank
  - Cumulative Proportion * 100

Hotel Rates	Frequency	Proportion	Percent	Cumulative Frequency	Cumulative Proportion	Percentile
800-899	1	.03	3	35	1	100
700-799	4	.11	11	34	.97	97
600-699	2	.06	6	30	.86	86
500-599	0	0	0	28	.80	80
400-499	6	.17	17	28	.80	80
300-399	8	.23	23	22	.63	63
200-299	8	.23	23	14	.40	40
100-199	4	.11	11	6	.17	17
0-99	2	.06	6	2	.06	6

SUMMARY OF THE FORMULAS

- **Proportion/Relative Frequency:** f/Total Frequency
- **Percentage:** Proportion/Relative Frequency*100
- **Cumulative Proportion/Cumulative Relative Frequency:** Cumulative Frequency/Total Frequency
- **Percentile Rank:** Cumulative Proportion/Cumulative Relative Frequency*100

Histogram

- A graph is a good way to understand a large group of numbers

- Histogram is a barlike graph of a frequency distribution in which the values are plotted along the horizontal axis

- and the **height** of each bar is the **frequency** of that value

- The bars are usually placed next to each other w/out spaces, giving the appearance of a city skyline

Histogram

(a) Frequency Table

❶

STRESS RATING	FREQUENCY
0	1
1	1
2	1
3	2
4	1
5	2
6	4
7	7
8	5
9	3
10	3

❸ ❹ ❷

(b) Grouped Frequency Table

❶

STRESS RATING INTERVAL	FREQUENCY
0 – 1	2
2 – 3	3
4 – 5	3
6 – 7	11
8 – 9	8
10 – 11	3

❸ ❹ ❷
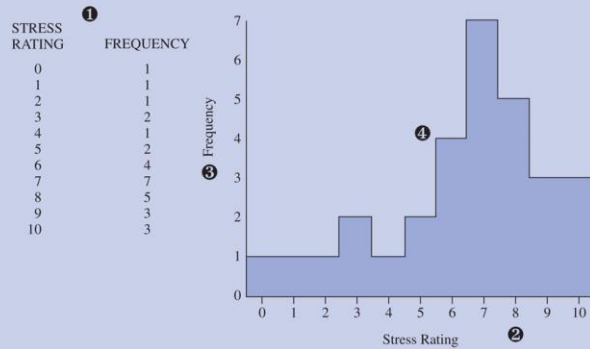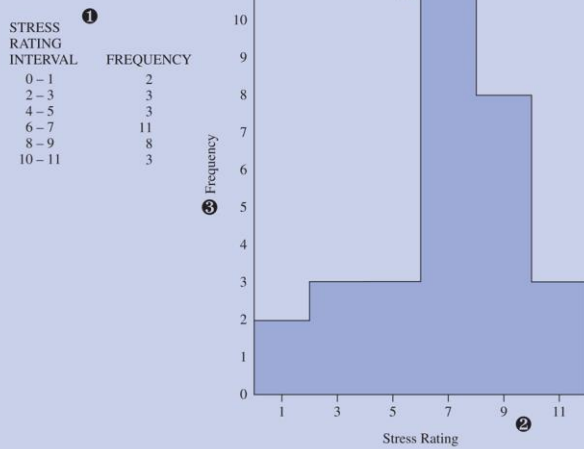
Histogram

- Steps to make a histogram
  - Make a frequency table
  - Put the values along the bottom of the page, from left to right, from lowest to highest
  - Make a scale of frequencies along the left edge of the page that goes from 0 at the bottom to the highest frequency for any value
  - Make a bar above each value with a height for the frequency of that value

Histogram

- Instead of histogram, **a bar chart** would be used to display **nominal data**
- Since the values of a nominal variable are not in any particular order, space is left between bars.

Example of a Pie Chart
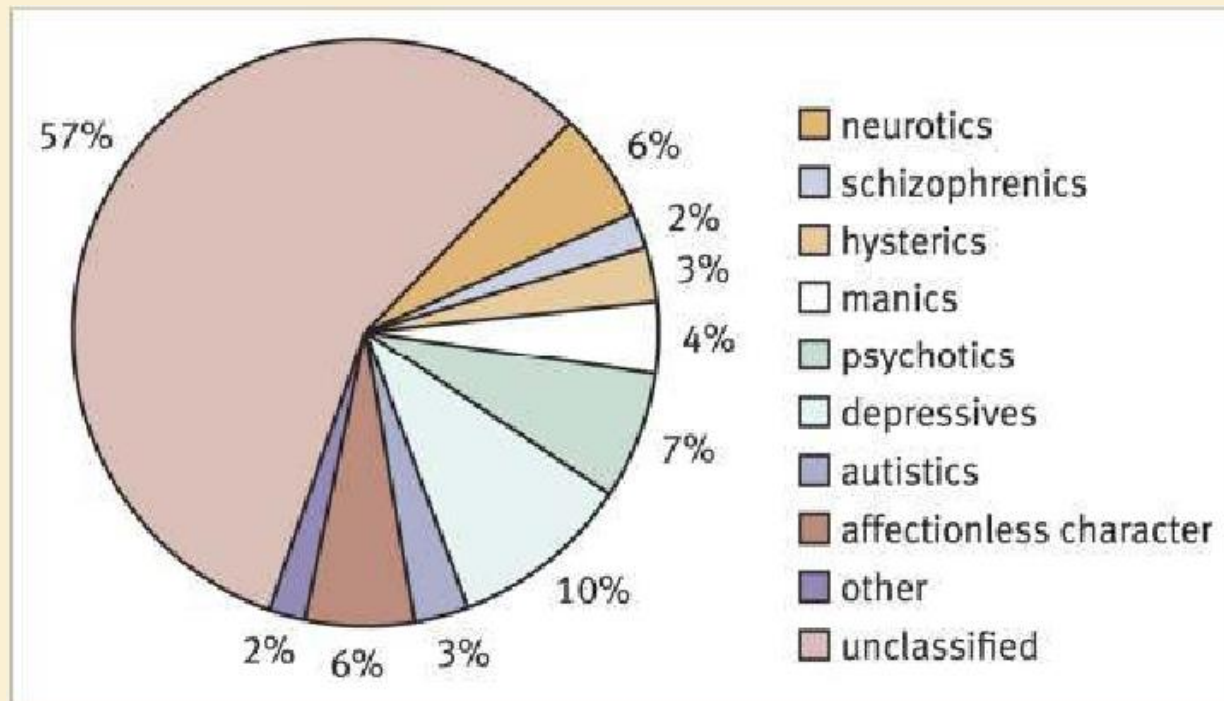
Figure 2.3   A poor pie diagram

Example of a Bar Chart

What this variable might be?
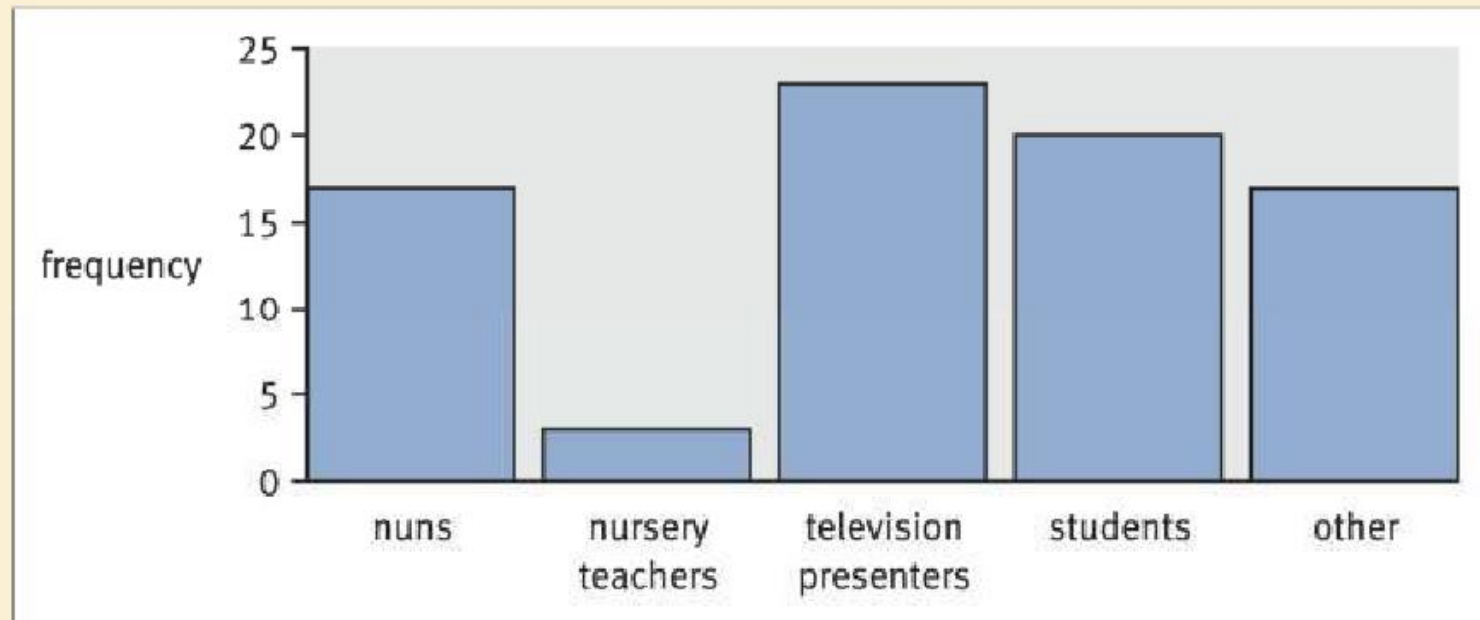
Figure 2.4   Bar chart showing occupational categories in Table 2.1

Describe the below data with a frequency table.
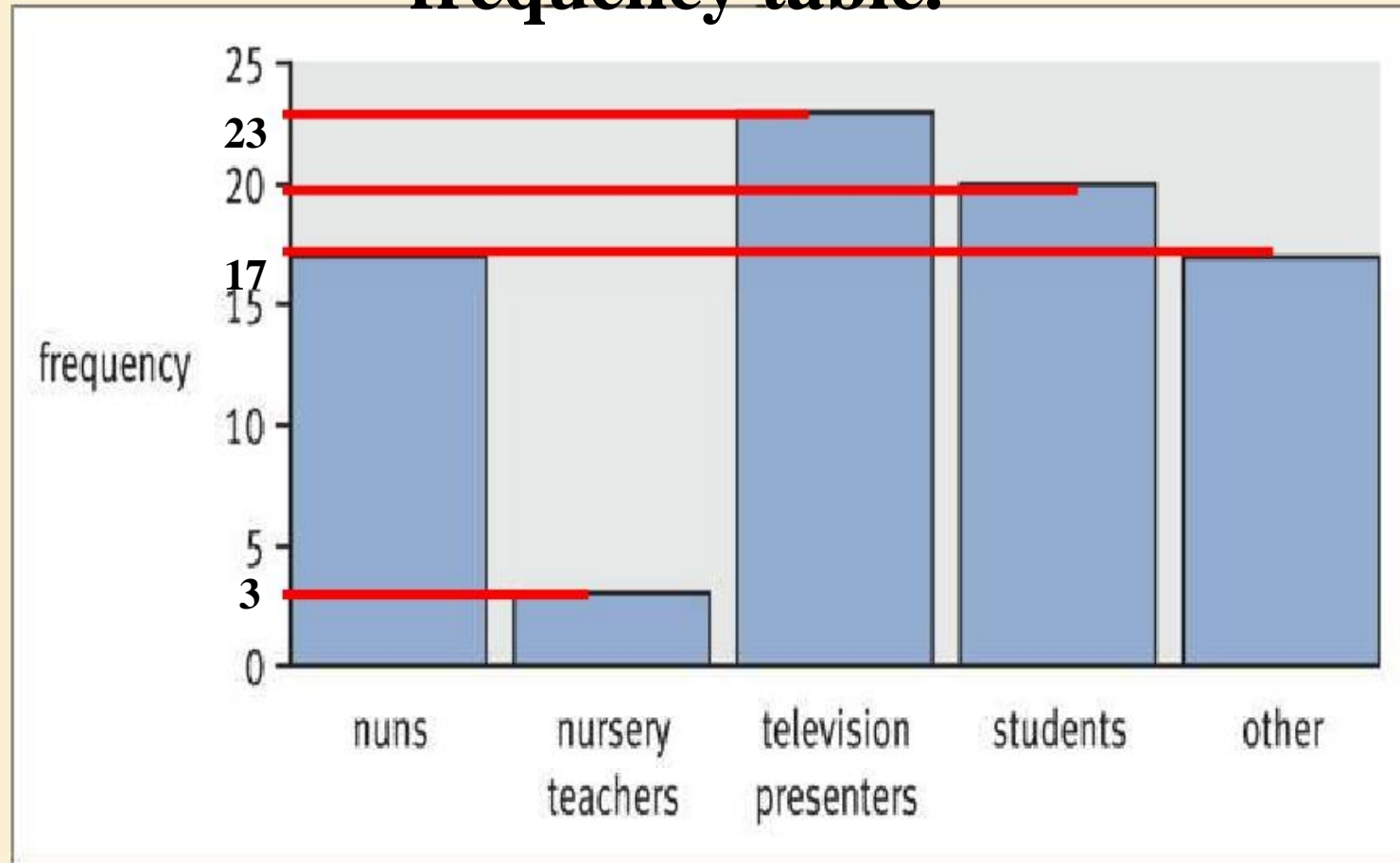
A Forced Choice Variable

- I like statistics.

  - 1) Strongly agree.
  - 2) Agree.
  - 3) Not sure.
  - 4) Disagree.
  - 5) Strongly disagree.

Frequency Table: I like statistics...

| Response	Value	Frequency	Percentage (%)
Strongly agree	1	17	42.5
Agree	2	14	35
Notsure	3	6	15
Disagree	4	2	5
Strongly disagree	5	1	2.5

Howitt and Cramer, *Introduction to Statistics in Psychology*, 5th Edition and *Introduction to SPSS Statistics in Psychology*, 5th Edition © Pearson Education Limited 2011

Example of a Histogram
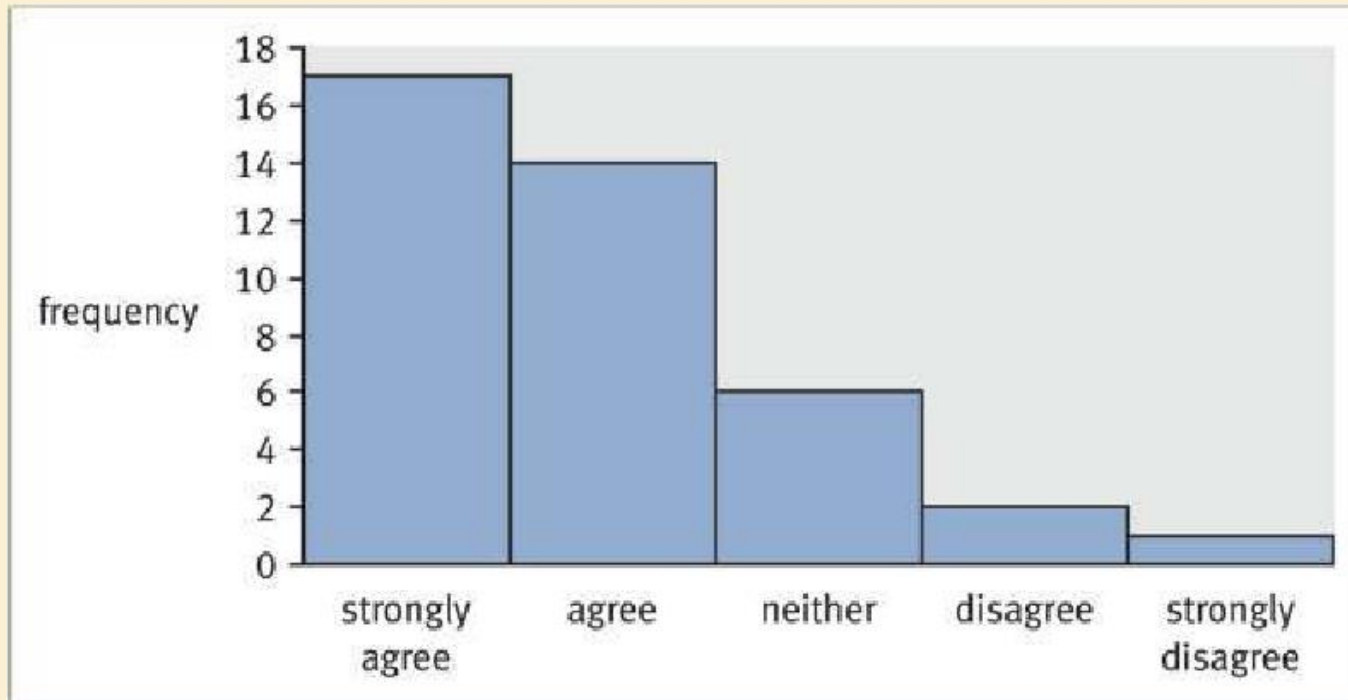
Figure 2.5 Histogram of students' attitudes towards statistics

Example of a Bar Chart
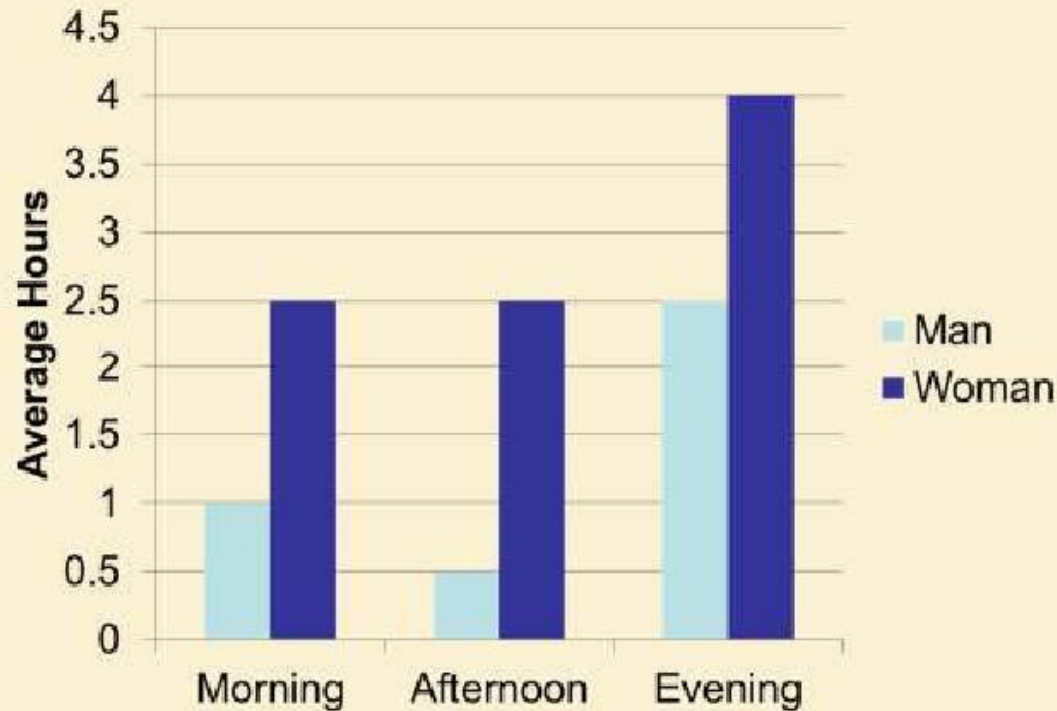
Figure 2.5 Bar chart of number of hours spend with children as a function of time of the day and gender

Example of a Line Diagram

Bar chart of number of hours spend with children in a family as a function of time of the day and gender

Shapes of Frequency Distributions -1

- A frequency distribution shows the pattern of frequencies over the various values

- A frequency table or histogram shows how the frequencies are spread out or "distributed"

Shapes of Frequency Distributions -1

- Unimodal, bimodal, and rectangular
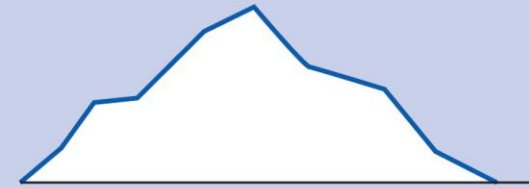
(a) Unimodal

(b) Approximately Bimodal

(c) Approximately Rectangular

Shapes of Frequency Distributions -1

- If a distribution's shape has only one main high point: one height tower on the histogram
  - This is **uni**modal distribution
- If a distribution has two fairly equal high points
  - It is a **bi**modal distribution
- Any distribution which has two or more high points is **multi**modal distribution
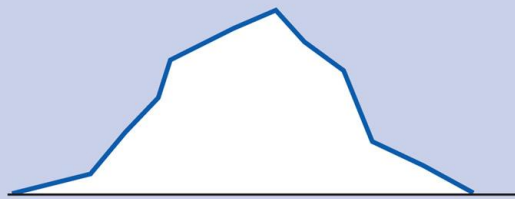- A distribution w/ values of all about the same frequency is **rectangular** distribution

Shapes of Frequency Distributions -1

- Frequency polygons
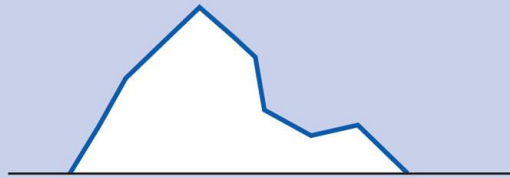- The height of each point shows the number of scores w/ that value
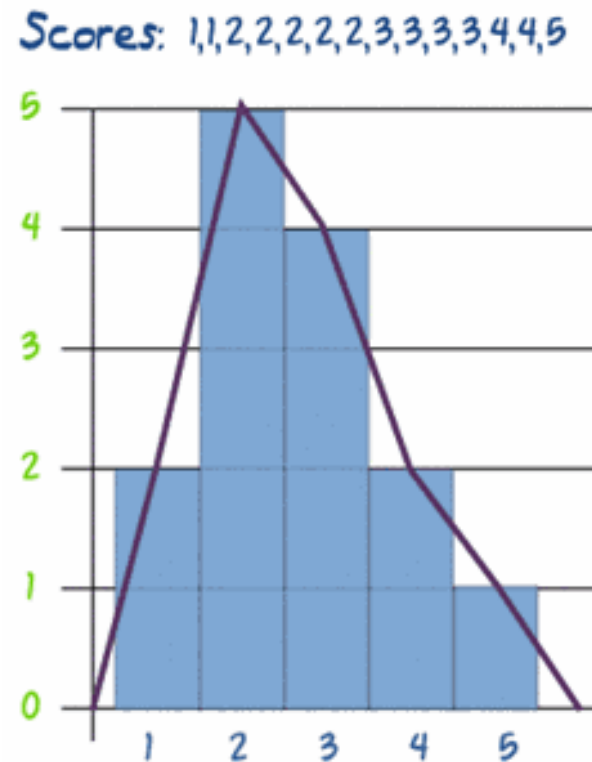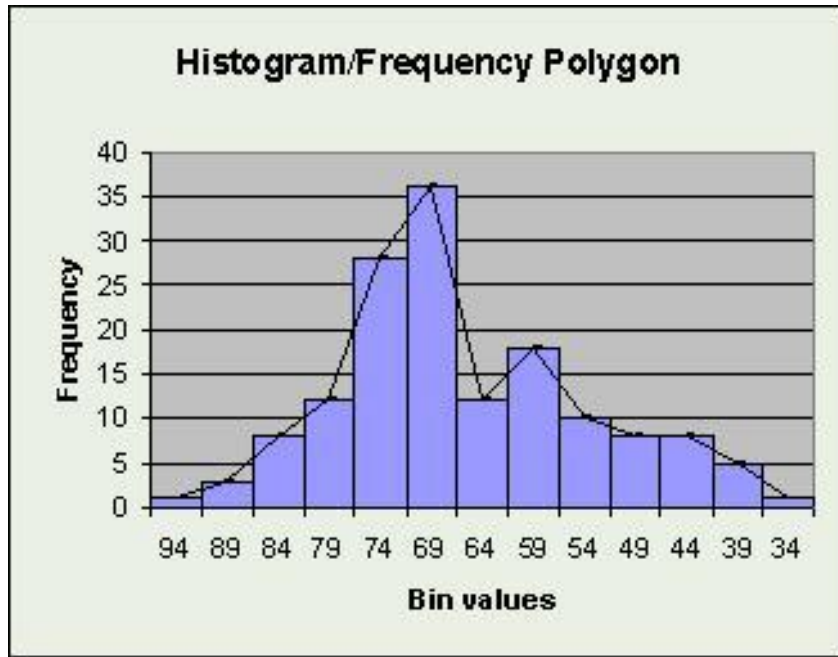
Shapes of Frequency Distributions -2

- Symmetrical and skewed distributions

(a)   (b)   (c)

Shapes of Frequency Distributions -2

- Symmetrical and skewed distributions

- The distributions can be lopsided (orantısız, bir yana eğik). More scores can be near the high or low end => SKEWED DISTRIBUTION

- If equal numbers are placed about equal numbers on both sides of the middle => SYMMETRICAL DISTRIBUTION

Shapes of Frequency Distributions -2

- Skewed distributions are not symmetrical

- Tail shows the direction of the skew

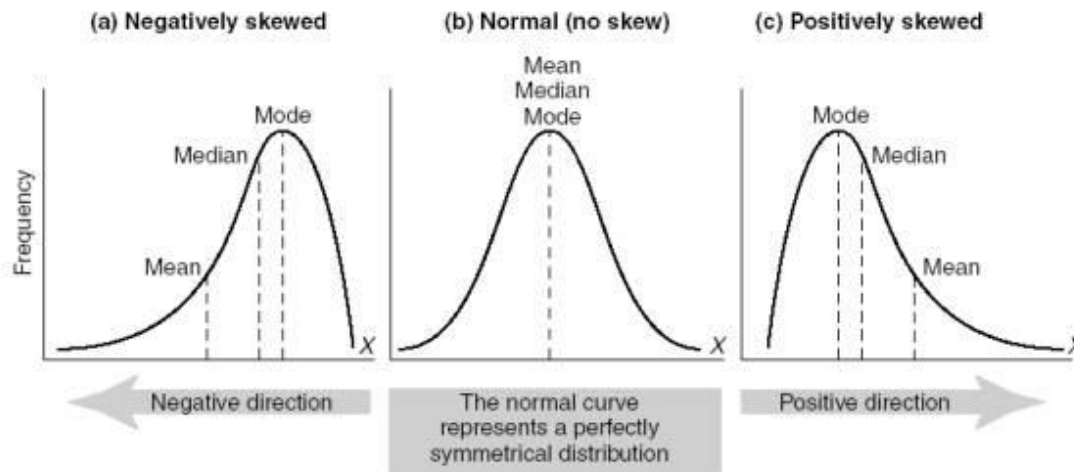- Tail is the side of the distribution w/ fewer scores

(a) Negatively skewed
Mode
Median
Mean
Frequency
Negative direction

(b) Normal (no skew)
Mean
Median
Mode
The normal curve represents a perfectly symmetrical distribution

(c) Positively skewed
Mode
Median
Mean
Positive direction

■ FIGURE 15.6   Examples of normal and skewed distributions

Shapes of Frequency Distributions -2

- A skewed distribution caused by a lower limit => Floor effect

- A skewed distribution caused by an upper limit => Ceiling effect

Shapes of Frequency Distributions -3

• Normal and kurtotic distributions


(a)   (b)   (c)
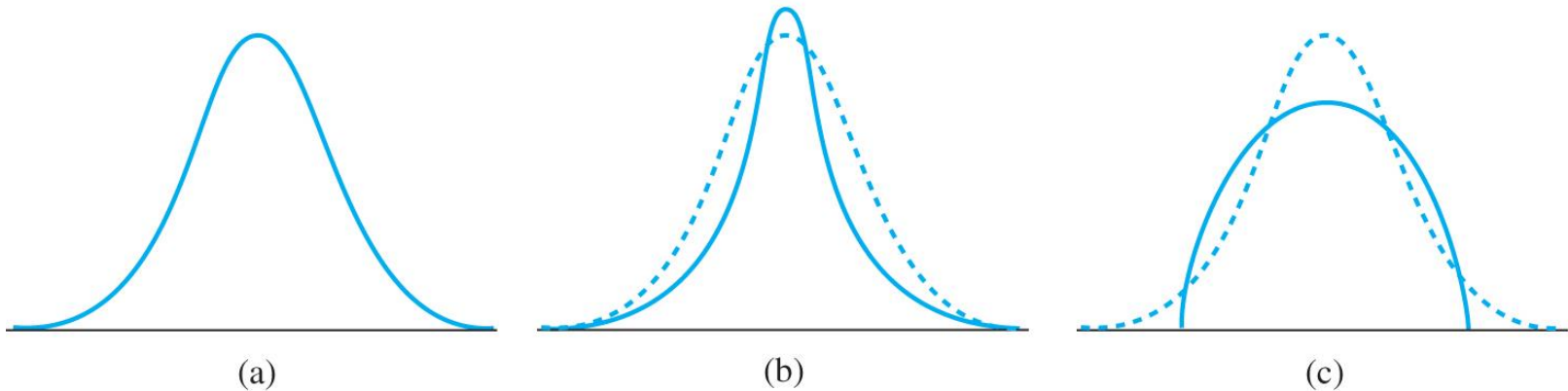
Shapes of Frequency Distributions -3

- Normal and kurtotic distributions
- We describe a distribution in terms of whether the middle of the distribution is particularly peaked or flat
- The standard of comparison is a bell-shaped curve → NORMAL CURVE
- Normal curve is unimodal, symmetrical curve w/ an average peak- the sort of bell-shaped

Shapes of Frequency Distributions -3

- Kurtosis is how much a shape of a distribution differs from a normal curve in terms of whether its curve in the middle is more peaked or flat than the normal curve (DeCarlo, 1997).

Shapes of Frequency Distributions -3

(a)   (b)   (c)
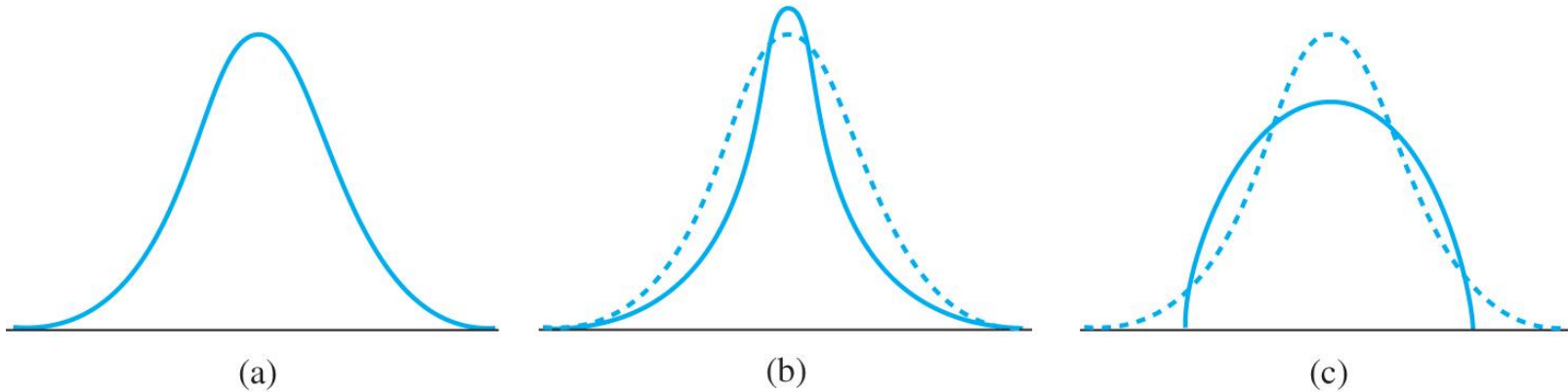
- A) normal curve
- B) heavy-tailed distribution
- C) light-tailed distribution

Exercise 1

The following data are representing answers of 20 participants to the question "I feel I am being paid a fair amount for the work I do." Responses were coded as following

1- Completely agree. 2-Somewhat agree. 3-Neither agree nor disagree. 4-Somewhat disagree 5-Completely disagree

Draw a frequency table for each response category.

- 2, 1, 2, 1, 1, 5, 5, 4, 5, 1, 5, 3, 4, 4, 4, 1, 2, 5, 2, 4,

Exercise 2

- A particular block in a suburban neigborhood has 20 households

- The number of children in these households are as follows:

- 2, 4, 2, 1, 0, 3, 6, 0, 1, 1, 2, 3, 2, 0, 1, 2, 1, 0, 2, 2

- Make a f table including relative frequency, percentage, cumulative frequency, cumulative proportion, percentile

- A histogram

- Describe the general shape of the distribution

Exercise 3

- Following are the speeds of 40 carsclocked by radar on a particular road in a 35-mph zone on Sunday afternoon:

- 30,36,42,36,30,52,36,34,36,33,30,32,35,32,37,34,36,31,35,20,24,46,23,31,32,45,34,37,28,40,34,38,40,52,31,33,15,27,36,40

- Make a f table including relative frequency, percentage, cumulative frequency, cumulative proportion, percentile

- A histogram

- Describe the general shape of the distribution

Exercise 4

- These are the scores on a measure of sensitivity to smell taken by 25 chefs attending a national conference:

- 96,83,59,64,73,74,80,68,87,67,64,92,76,71,68,50,85,75,81,70,76,91,69,83,75

- Make a f table including relative frequency, percentage, cumulative frequency, cumulative proportion, percentile

- A histogram

- Describe the general shape of the distribution

Exercise 5

- The following data are the number of minutes it took each of 34 10-year-olds to do a series of abstract puzzles:

- 24,83,36,22,81,39,60,62,38,66,38,36,45,20,20,67,41,87,41,82,35,82,28,80,80,68,40,27,43,80,31,89,83,24

- Make a f table including relative frequency, percentage, cumulative frequency, cumulative proportion, percentile

- A histogram

- Describe the general shape of the distribution

Overview

- You have collected data. Conducted statistical analyses. Now what?

    – You a need to effectively communicate the key characteristics of your data

    – You have to communicate results in the simplest and the most effective way

- Graphically description of data

    – Tables, graphs or charts

Review of Key Concepts

- What is variable
  - Variable is something which varies
- What is nominal variable
  - Nominal variables are those where cases are placed into groups (e.g. gender, eye colour)
- What is score variable
  - Score variables are those in which cases are given a numerical value (e.g. age, income)
- What are the types of score variable
  - Ordinal, interval and ratio

Howitt and Cramer, *Introduction to Statistics in Psychology*, 5th Edition and *Introduction to SPSS Statistics in Psychology*, 5th Edition © Pearson Education Limited 2011

Review of Key Concepts

- What is raw data

  – Raw data are the actual measures taken from the sample

- What is descriptive statistics

  – Descriptive statistics are visual and numerical techniques for presenting the major features of one's raw data

How to Present Nominal Variables?

- The aim is to display the number of cases which fall into each category

- Frequency table
  - different groups are presented with the number and percentage in each category

- Pie chart
  - It shows the relative size of each group based on percentages

- Bar chart
  - A bar chart represents the size of groups based on the height of bars on a graph

A Forced Choice Variable

- Where do you live? Please choose one of the options below.

    – 1) I live in public dormitory.

    – 2) I live in private dormitory.

    – 3) I rented a house.

    – 4) I am leaving with my family.

Frequency Table: Student Accommodation

Category	Frequency	Percentage (%)
PublicDormitory	600	25
PrivateDormitory	360	15
Homerental	240	10
Family	1200	50
Total	2400	100

An open-ended variable

- What is the diagnosis of the patient?

Errors to Avoid

- There are some key points that you must keep in mind when presenting your data using tables, charts or graphs:

1. Make sure that you provide clear headings that succinctly describe the contents
2. Clearly label all the key elements
3. Make the diagrams as simple as possible
4. Check with other people that they can understand the diagrams and that they convey the key messages as intended

End of Chapter 1