



SHERIDAN
INSTITUTE OF HIGHER EDUCATION

RS292 Research Methods

*Analysing and Presenting Quantitative
Data*

Lecturer: Dr Maya Krayneva

Textbook: Textbook: Gray, D. (2018). *Doing research in the real world* (4th ed.). Los Angeles: Sage. (Chapter 15)



Quantification of Data

- **Quantification Analysis** – the numerical representation and manipulation of observations for the purpose of describing and explaining the phenomena that those observations reflect.

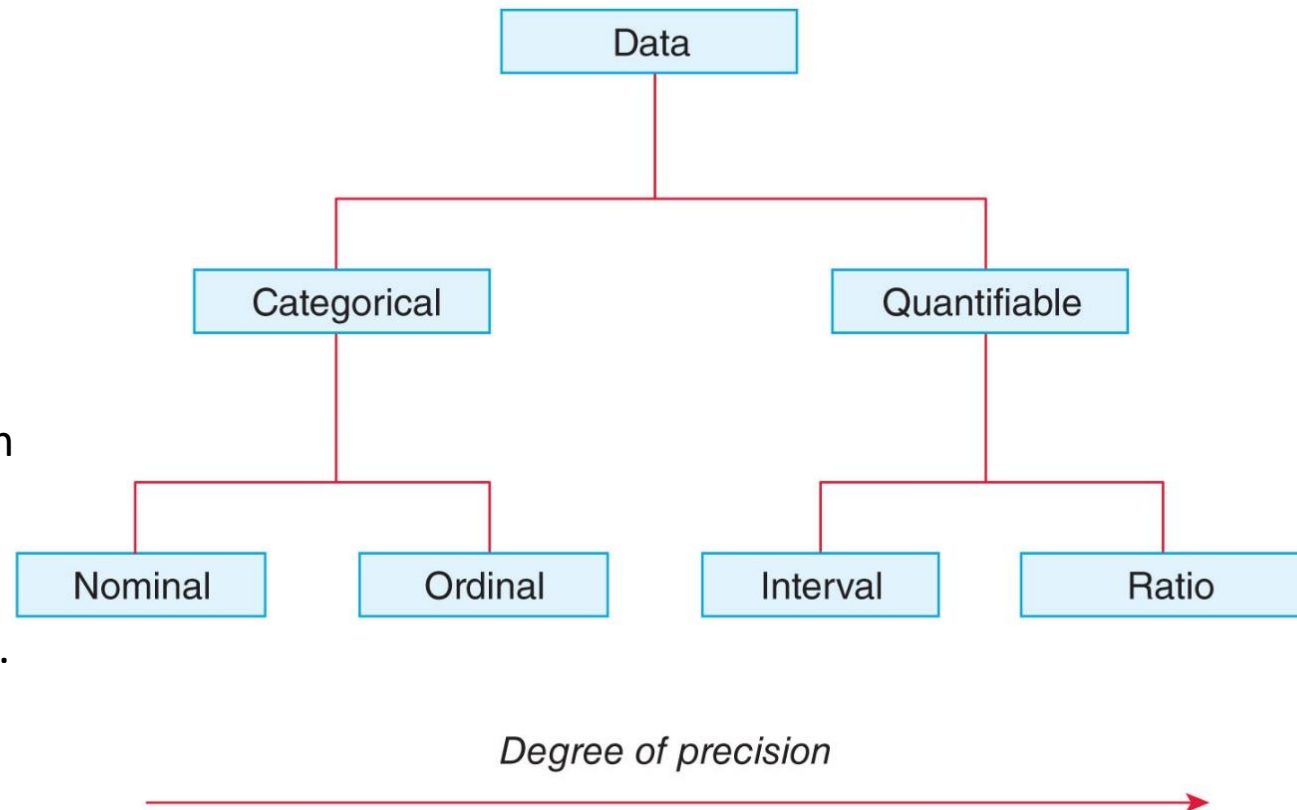


Rule Number 1:

- If you have collected your data and now you are thinking of how to analyse it – you are too late!
- Should select tests at planning stage.



Categorical and Quantifiable Data



- Nominal: a name value or **category** with no order or ranking implied, e.g. type of firm
- Ordinal: an **ordering or ranking** of values, although the intervals between the ranks are not intended to be equal, e.g. strongly agree, agree, disagree, etc.

- Interval: numerical values are assigned along an **interval scale** with equal intervals, but there is no zero point where the trait being measured does not exist, e.g. IQ score or blood pressure
- Ratio: a subset of interval data, but there is an **absolute zero** that represents some meaning, e.g. sales, customer numbers.

Data Entry, Layout, and Quality

- Data analysis is only reliable if data that have been entered into computer accurately
- *Data coding*: allocating an identification number (Id) to data
- Can use a data matrix:
 - column = single variable
 - row = case/profile
- Note: numerical values have different meanings for different variables

Case	Id	Department	Length of service	Seniority
Case 1	1	5	3	2
Case 2	2	2	1	3
Case 3	3	3	12	2

Data coding examples

Age

1 = 1

2 = 2

3 = 3

4 = 4

5 = 5

Sex

Male = 1

Female = 2

Political Affiliation

Democrat = 1

Republican = 2

Independent = 3

Region of Country

West = 1

Midwest = 2

South = 3

Northeast = 4

Handling Missing Data

Non-response bias: has the respondent omitted those questions they feel uneasy about or hostile to answering?

Distinguish between four types of missing values:

'Not applicable' (NA)

'Refused' (RF)

'Did not know' (DK)

'Forgot to answer' (FA)



61.6%: 99.19

104.19

86.72

Analysing
quantitative data



Univariate Analysis

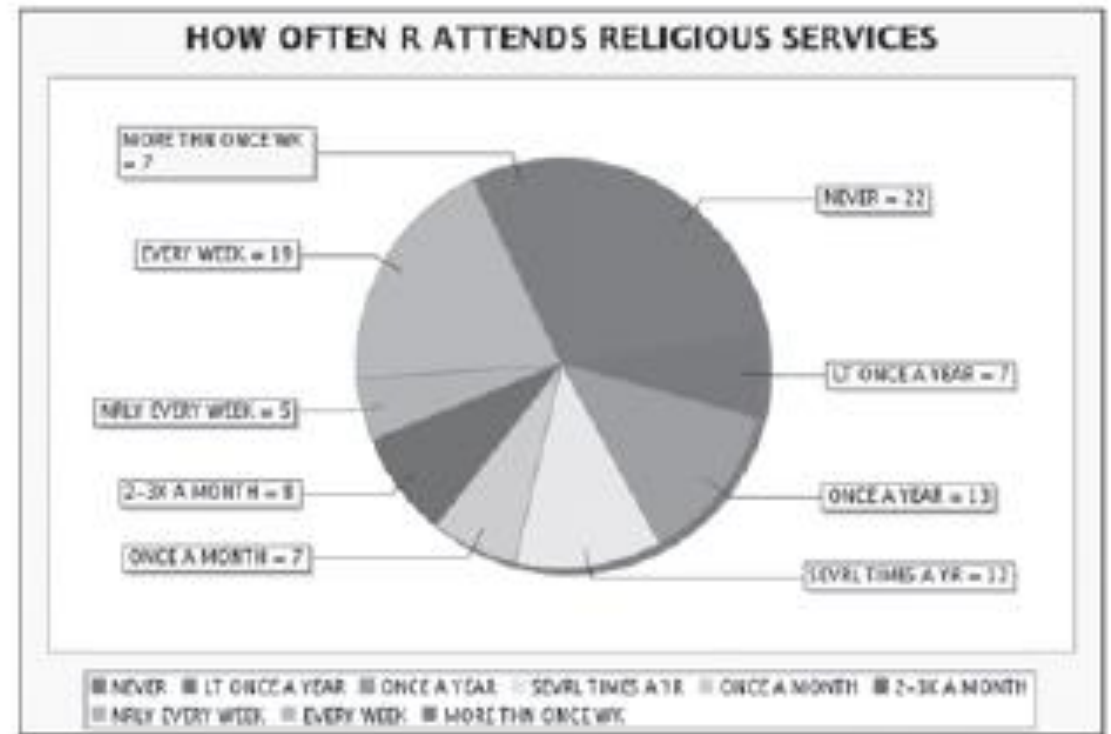
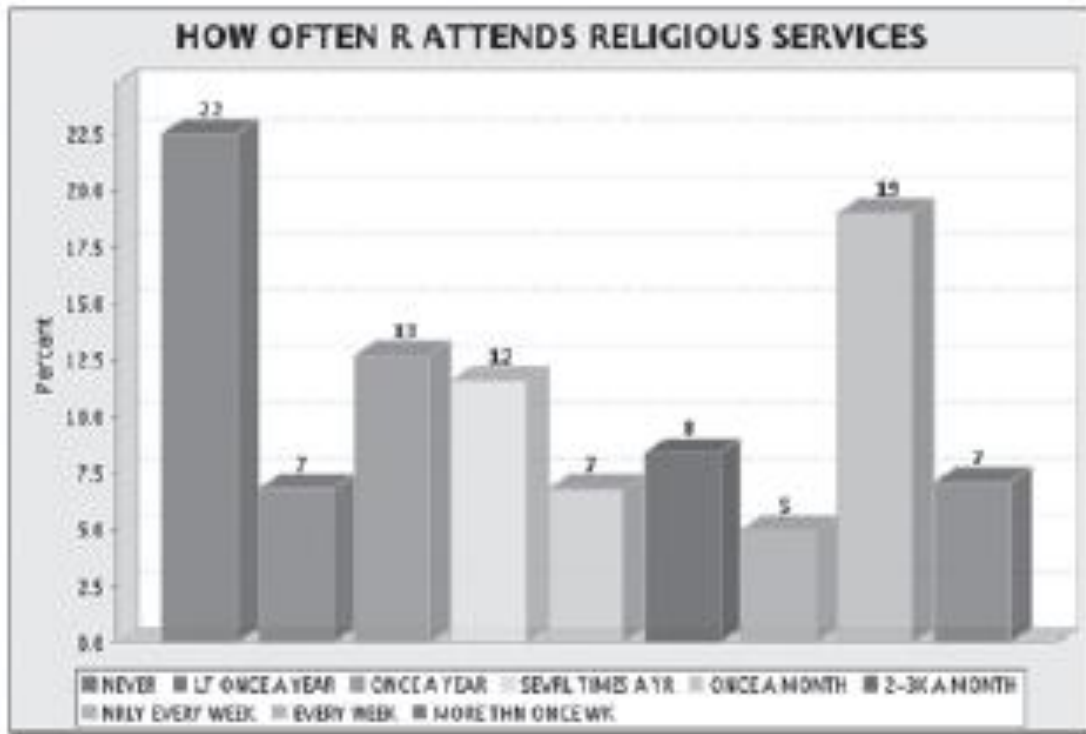
Univariate Analysis – the analysis of a single variable, for purposes of description

Example: Gender

- The number of men in a sample/population and the number of women in a sample/population.

Frequency Distribution

Frequency Distributions – a description of the number of times the various attributes of a variable are observed in a sample i.e. number of instances in a class.



Central Tendency



Measuring Central Tendency by:

- **Mean** – an average computed by summing the values of several observations and dividing by the number of observations
- **Mode** – the most frequently occurring value
- **Median** – the central value when all the scores are arranged in order



Activity 1

The following list represents the scores on a mid-term exam:

100, 94, 88, 91, 75, 61, 93, 82, 70, 88, 71, 88

- Determine the mean.
 - Determine the mode.
 - Determine the median.
-



Activity 1

Determine the mean:

Add up the values and divide by the number of cases. Summing the values yields 1001. Divide this value by 12, the number of cases. The mean is 83.4.



Activity 1

Determine the mode:

One value, 88, occurs more often than any other value. The mode is 88.



Activity 1

Determine the median:

Reordering the cases from low to high yields

61, 70, 71, 75, 82, 88, 88, 88, 91, 93, 94, 100

The two middle values of these 12 cases are both 88; therefore the median is 88.

Activity 2

	Strongly Agree	Agree	Disagree	Strongly Disagree	Total
2012	14	40	32	14	100
2014	21	33	26	20	100

Calculating mean score for Likert scale

Strongly Agree	Agree	Disagree	Strongly Disagree
4	3	2	1

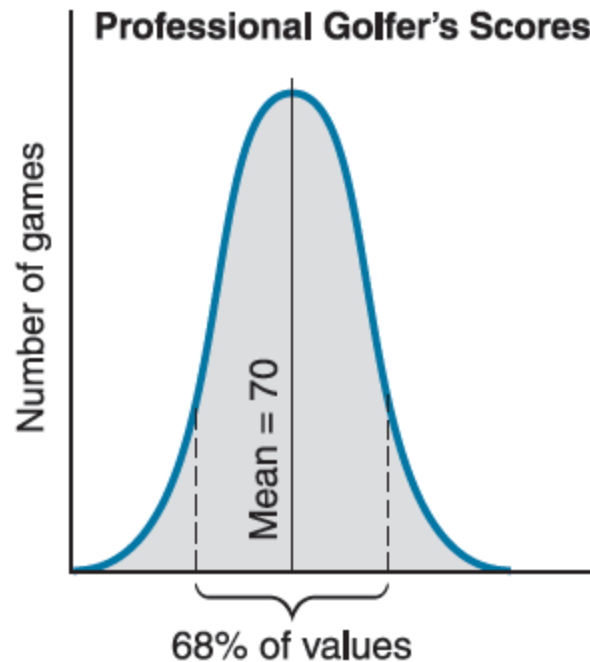
	Strongly Agree	Agree	Disagree	Strongly Disagree	Total
2012	56	120	64	14	254
2014	84	99	52	20	255

Dispersion

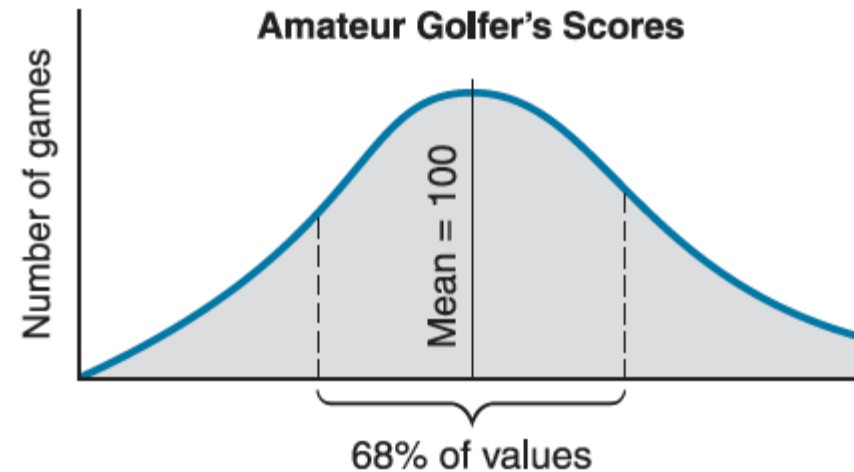
Dispersion – the distribution of values around some central value, such as an average.

- **Range:** difference between highest and lowest scores
- **Standard deviation:** the extent to which responses vary from the mean

b. Low standard deviation = tightly clustered values



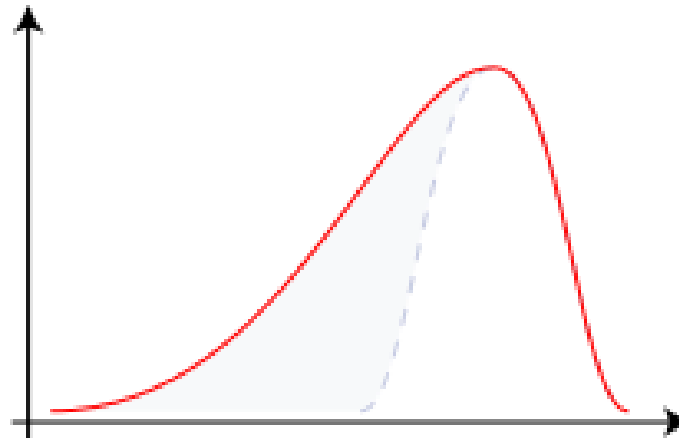
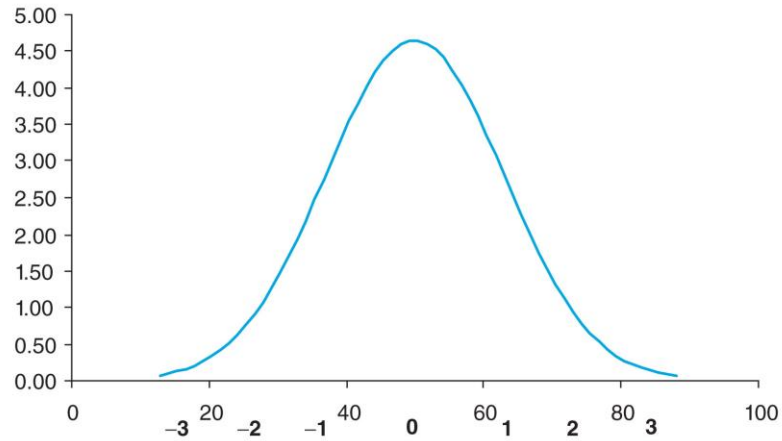
a. High standard deviation = spread-out values



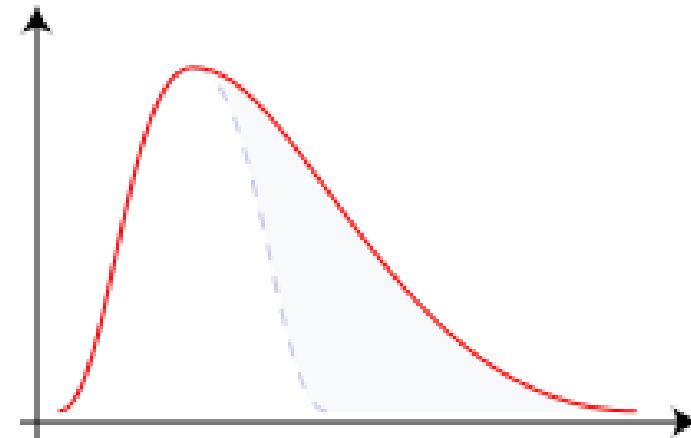
Normal vs Skewed Distributions

Normal distribution: bell-shaped (symmetrical around the mean)

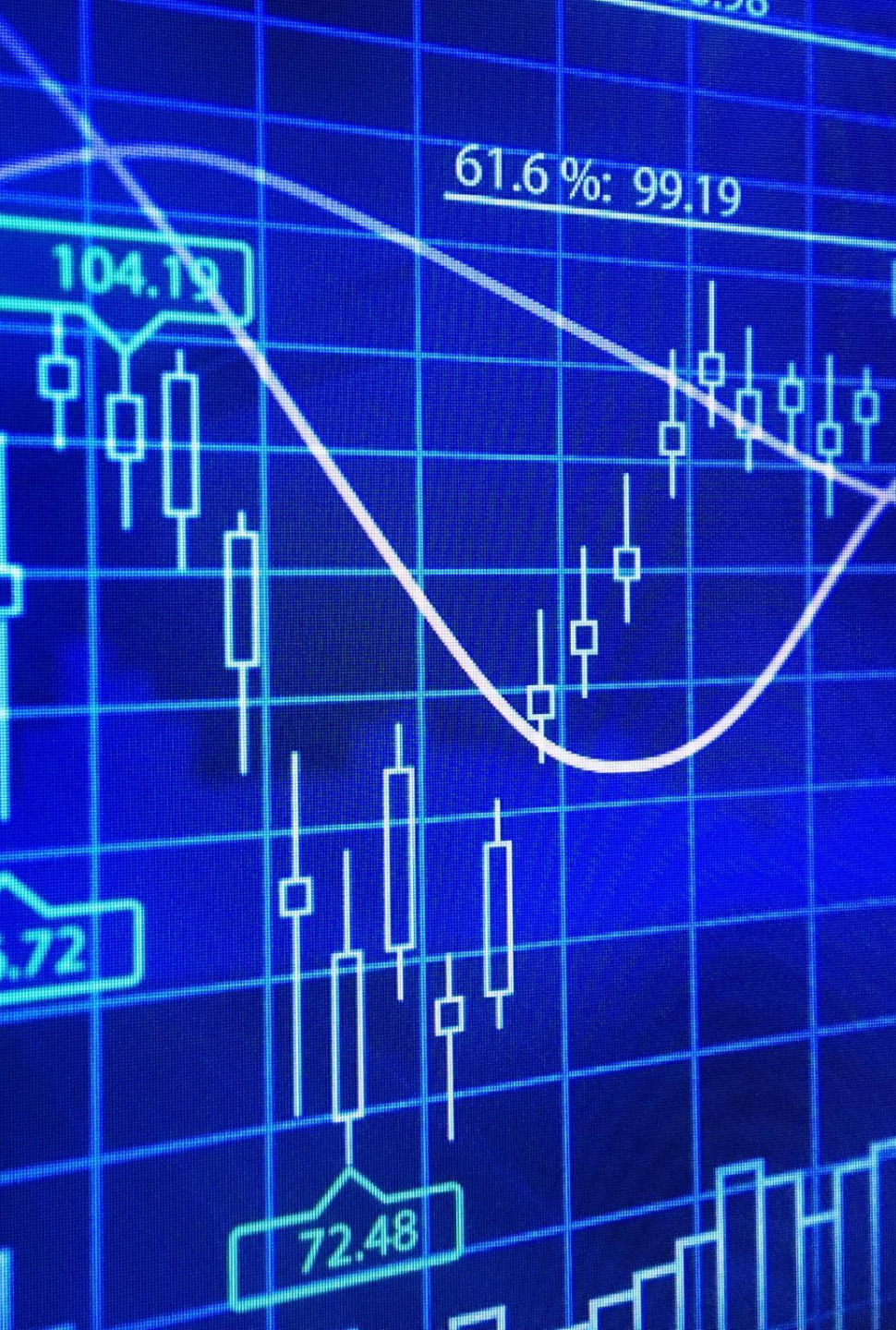
Skewed: not normal



Negative Skew



Positive Skew



Bivariate Analysis

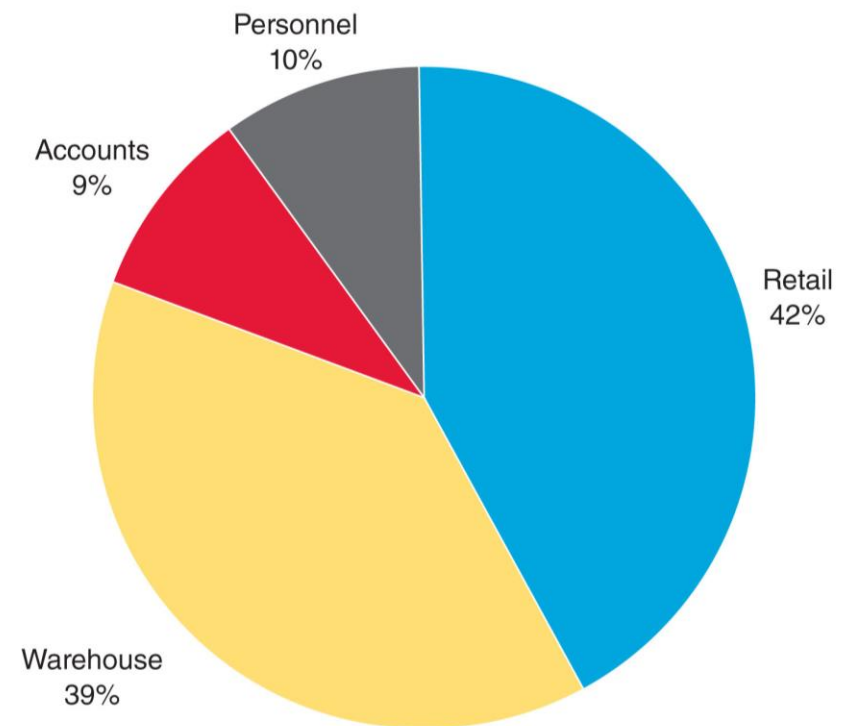
- Bivariate Analysis – the analysis of two variables simultaneously, for the purpose of determining the empirical relationship between them.



Presenting
quantitative data



Pie chart of the nominal data



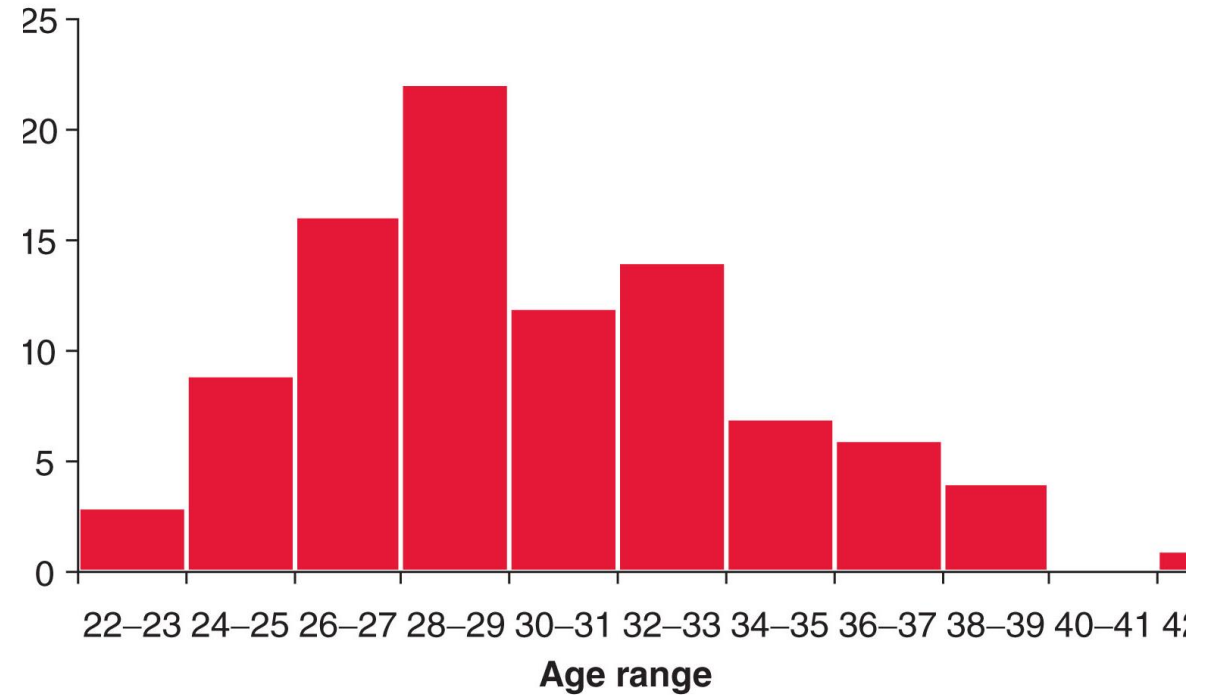
Bar chart of the nominal data

– Give example of using a
bar chart for ordinal data

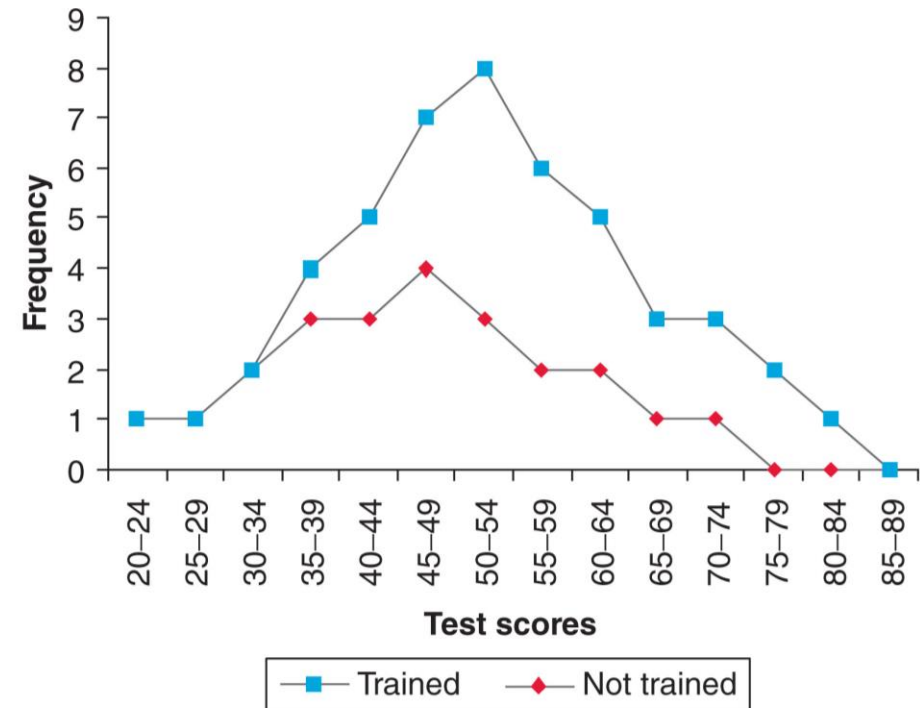




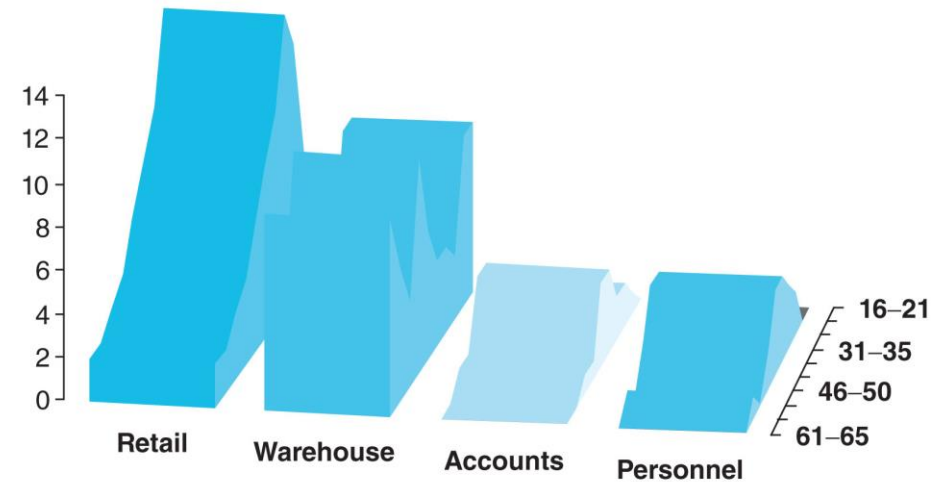
Histogram of the interval data



Frequency polygons
to compare two
groups for continuous
data



Solid polygon
for two variables
(department and age)





Activity 3

20 students in Yrs 2,4,6,8 and 10 were asked to answer the following statement: I get my most creative ideas when I am ?

The researcher was interested in seeing whether there was a difference between older and younger students.

First, she made a tally chart of their answers

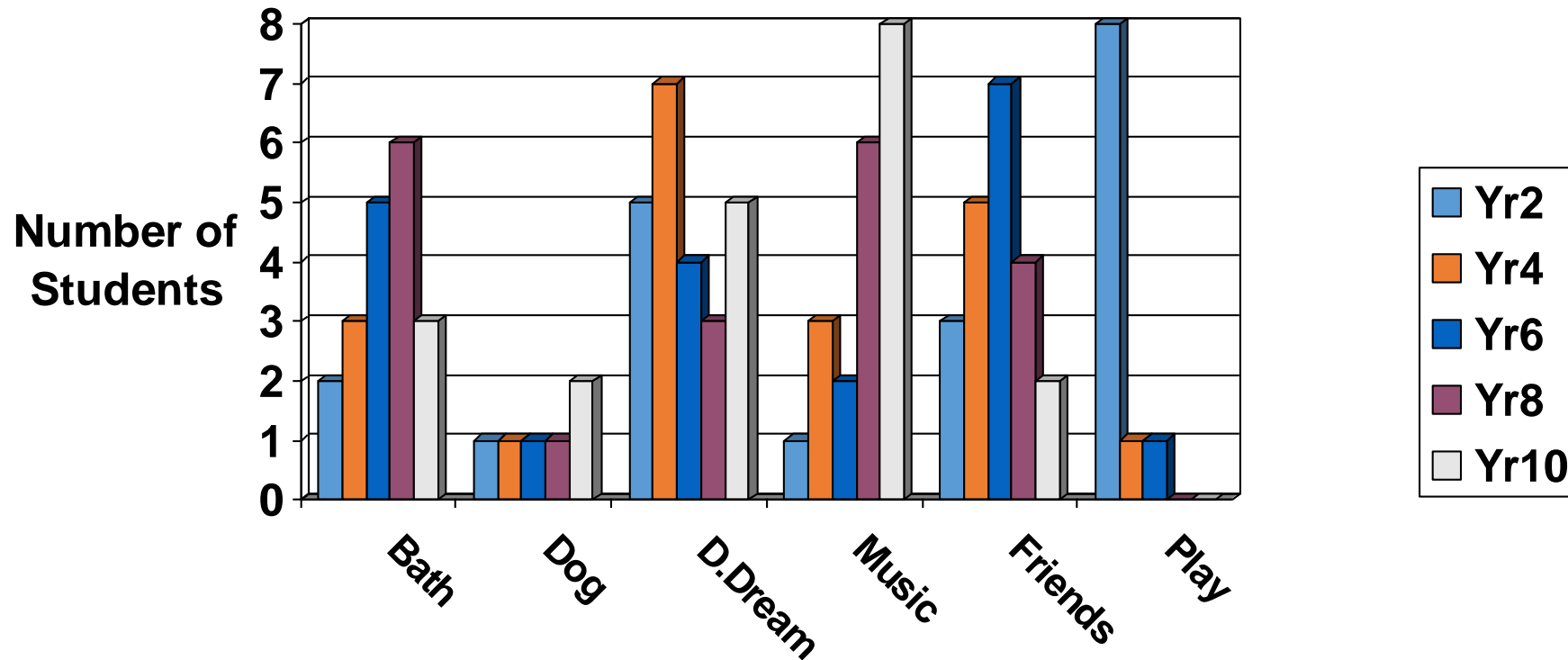
Category	Yr 2	Yr 4	Yr 6	Yr 8	Yr 10
In the bath or shower	xx	xxx	xxxxx	xxxxxxx	xxx
Walking the dog	x	x	x	x	xx
Day-dreaming	xxxxx	xxxxxxxx	xxxx	xxx	xxxxxxx
Listening to music	x	xxx	xx	xxxxxxx	xxxxxxxx
Chatting with friends	xxx	xxxxx	xxxxxxxx	xxxx	xx
Playing	xxxxxxxx	x	x		
	20	20	20	20	20

Next the researcher turned the tally chart into a table which showed the number of students in each category

CATEGORY	YR 2	YR 4	YR 6	YR 8	YR 10
In the bath or shower	2	3	5	6	3
Walking the dog	1	1	1	1	2
Day-dreaming	5	7	4	3	5
Listening to music	1	3	2	6	8
Chatting with friends	3	5	7	4	2
Playing	8	1	1	0	0
	20	20	20	20	20

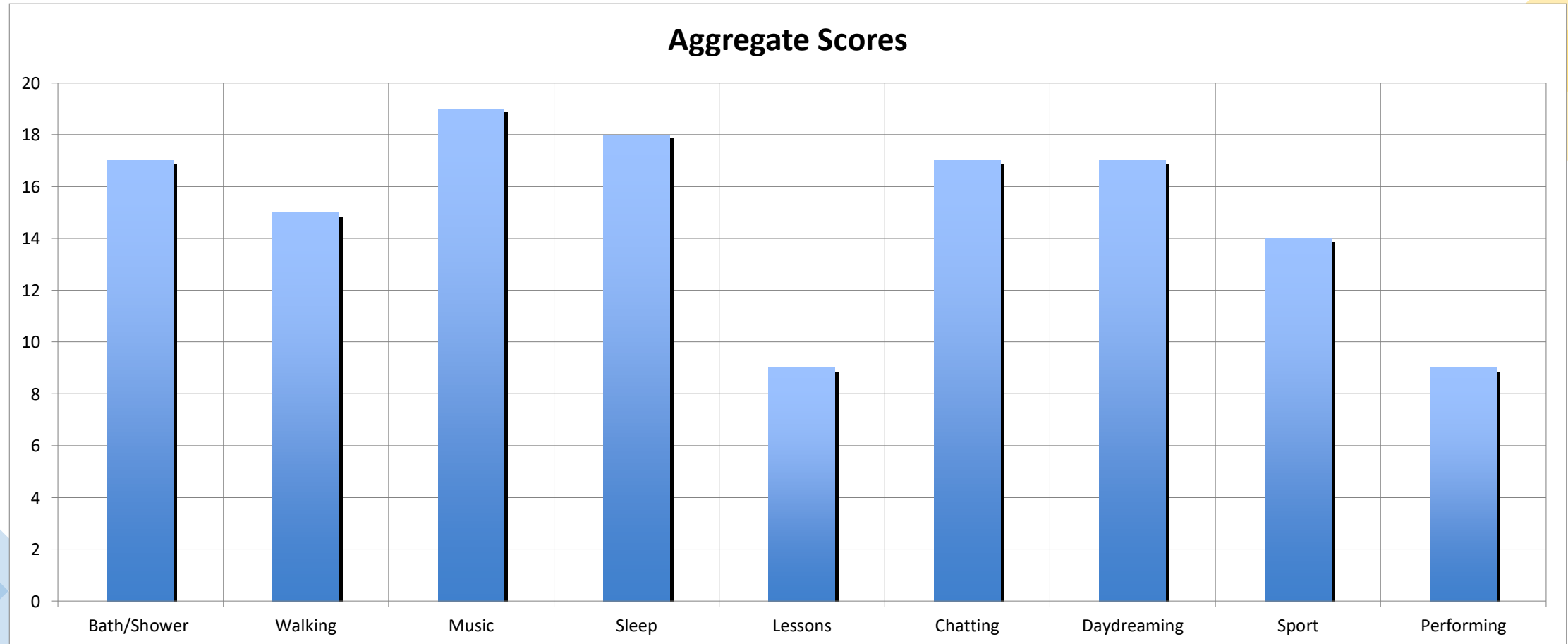
Then the researcher turned the table into a bar chart using the chart tool in PowerPoint. You can also do this using the chart tool in Excel. What does this chart tell you?

Where people get their best ideas



Then the researcher calculated the aggregate scores by multiplying each score point with the numeric value assigned to see which option was most popular.

What does this chart tell you?



Homework

READ CHAPTER 26 (GRAY)

LINKS:

https://www.youtube.com/watch?v=_ihkyrewbds

<https://www.youtube.com/watch?v=pfYhwy3wN14>

<https://www.youtube.com/watch?v=TfkNkrKMF5c>

START COLLECTING DATA

TOPIC WEEK 9:

QUALITATIVE DATA ANALYSIS