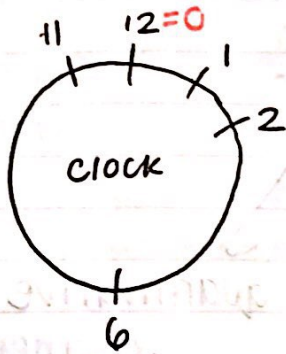


VARIABLE TYPES; RAW FREQUENCY HISTOGRAMS (10/3/19)



what time?

quant.
cont.
interval

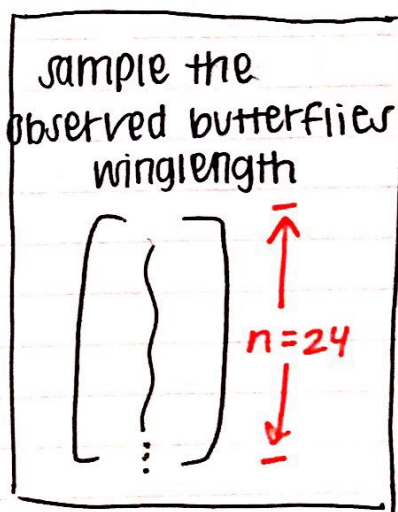
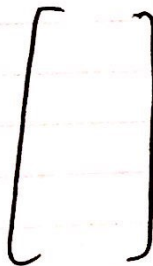
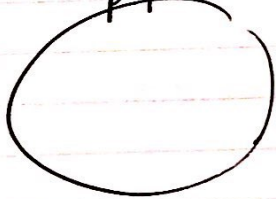
litter size

6
3
2.27
0
0

quant.
discrete
ratio

one row for each
litter

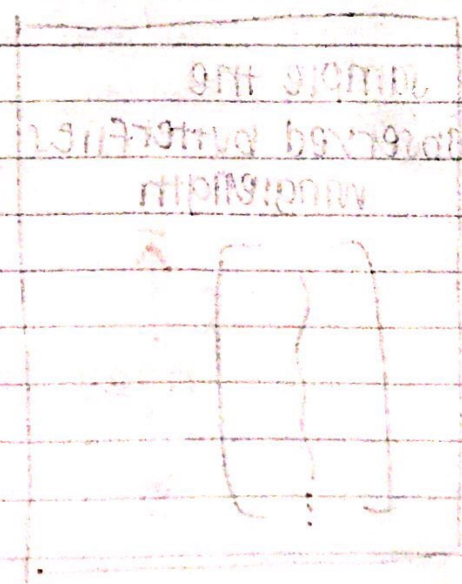
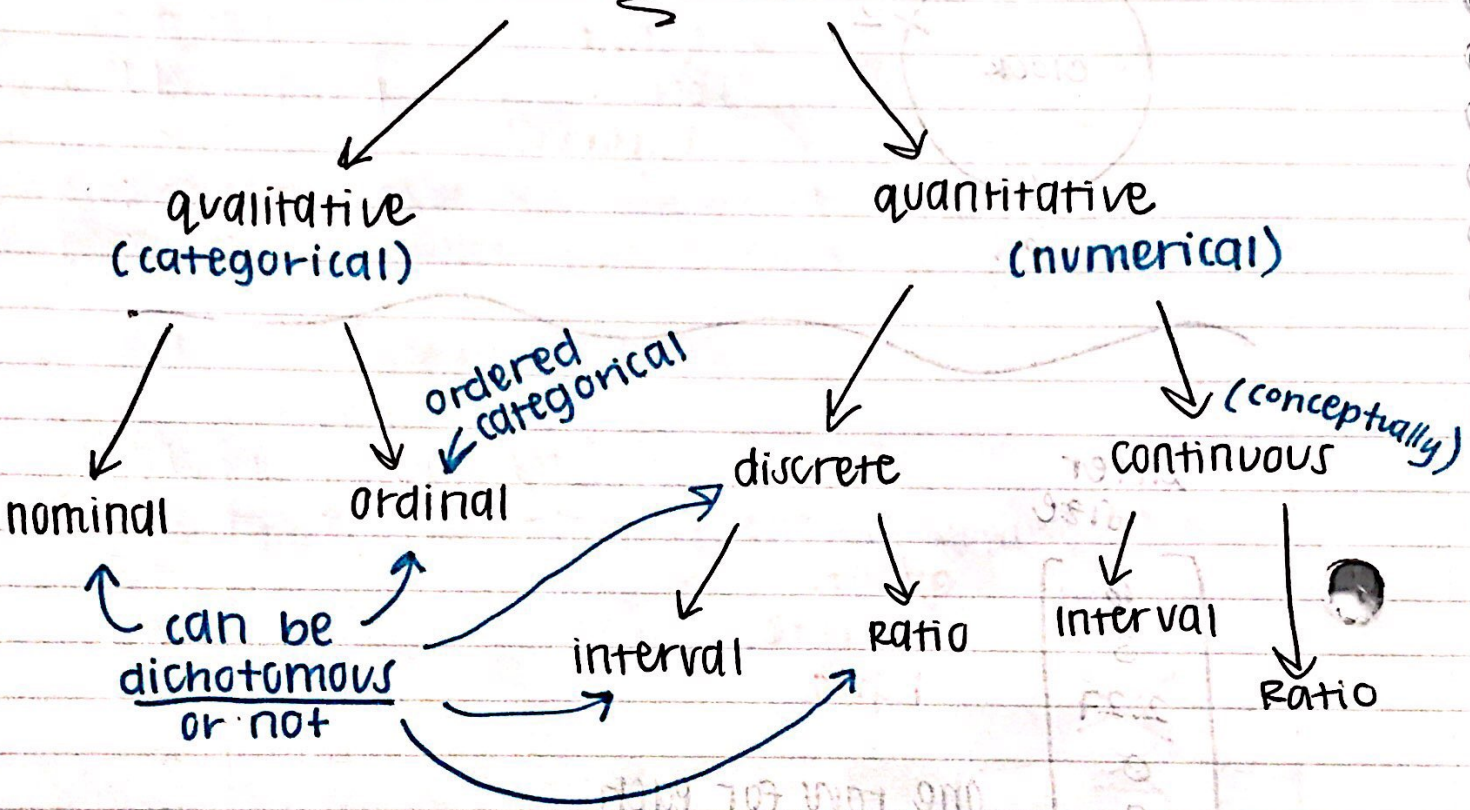
population



graphical
numerical

description
of
existing data
set

VARIABLE TYPES



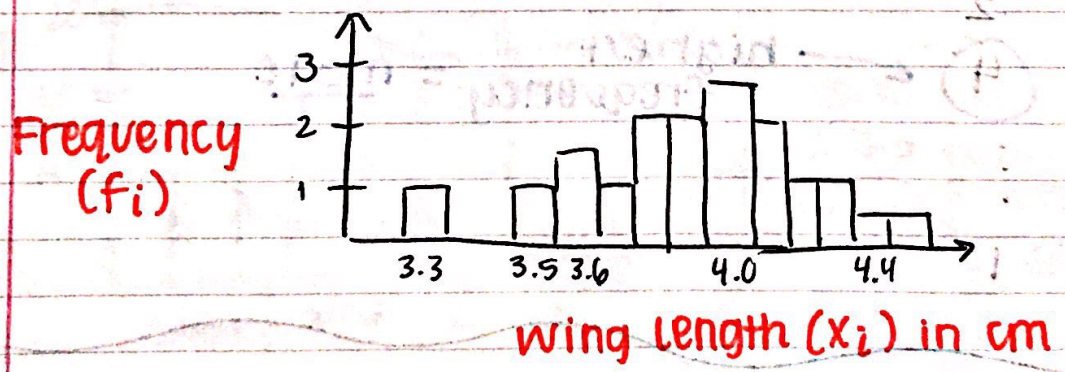
1.3 descriptive methods

→ sometimes useful and meaningful to summarize a variable by taking its mean

↳ The problem is that the mean is meaningful only for the age variable (b/c its quantitative [ratio discrete]); the other two variables are qualitative [nominal].

RAW FREQUENCY DISTRIBUTION

- sort the data from smallest to largest
- see duplicate values
- keep track of the values of the variable & the raw frequencies

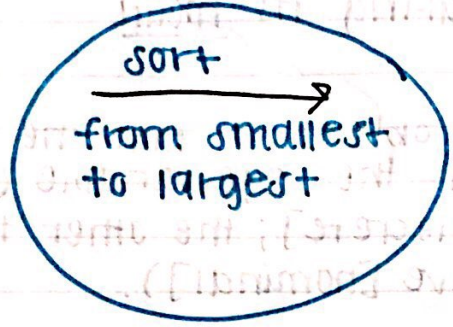


$y = \text{wing length (cm)}$

(order irrelevant)
Judgement

- 4.4
- 3.6
- 4.1
- 3.3
- ⋮
- 3.9 = $y_{24} = y_n$

$n = 24$

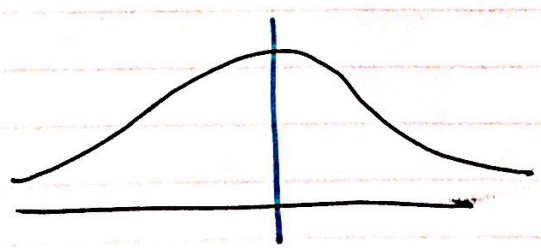


- y
- 3.3
- 3.5
- 3.6
- 3.6
- 3.7
- ⋮
- 4.5

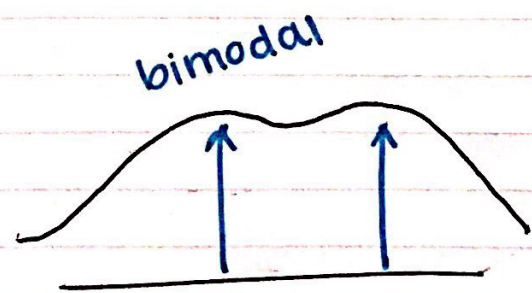
Value	count raw frequency
3.3	1
3.4	0
3.5	1
3.6	2
4.0	4
⋮	⋮
4.5	1

identical information

highest frequency = mode



mode



bimodal

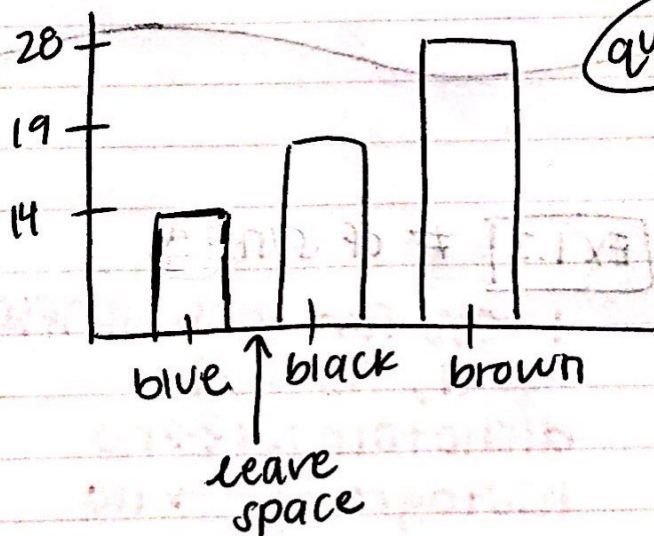
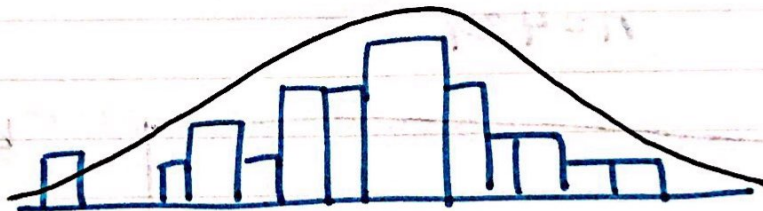
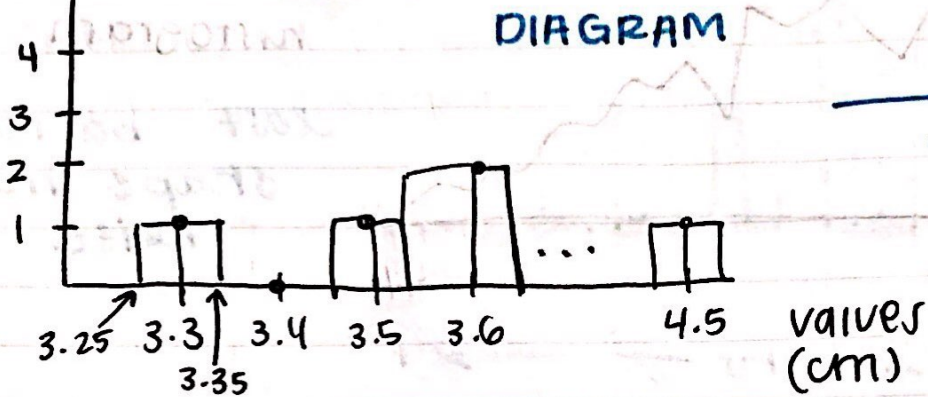
(multi)modal

(> 1)

height of adults

Raw Frequency

RAW FREQUENCY DIAGRAM



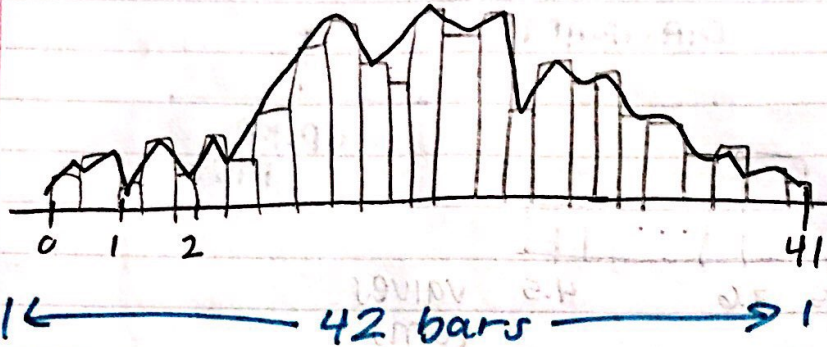
qual

blue
black
brown

blue 14
black 19
brown 28

bar graph? → yes

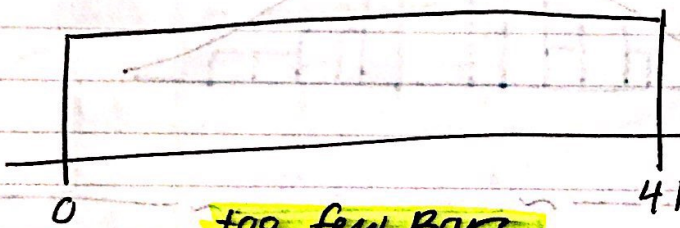
too many bars



bad histogram

lost basic shape in noise

$n=424$



too few bars

bad histogram

lost ^{all} sense of shape of distribution

more Graphical Examples

EX 1.1 | sparrow nests

1 row for each nest
qual, nominal
histogram? → no
bar graph? → yes
dichotomous? → no

EX 1.2 | # of sunfish

1 row for each sunfish
qual, ordinal
dichotomous? → no
histogram? → no
bar graph? → yes