this time: histograms, bar graphs

next time: measures of center & spread

AMS 7
10 Apr 17

go to (1)

discussion section this week (quiz)

I. Intro (week 1)

II. Descriptive Methods (graphical, numerical)

you have a data set; how do you gain insight into what it means?

variables are often named

x, y, z

y = wing length (cm)

pop

sample

the observed butterflies

wing length (cm)

4.4
3.6
⋮
3.9

at random →

n = 24

description here

1 row for each butterfly

raw data

(original order irrelevant)

4.4
3.6
4.1
⋮
3.9

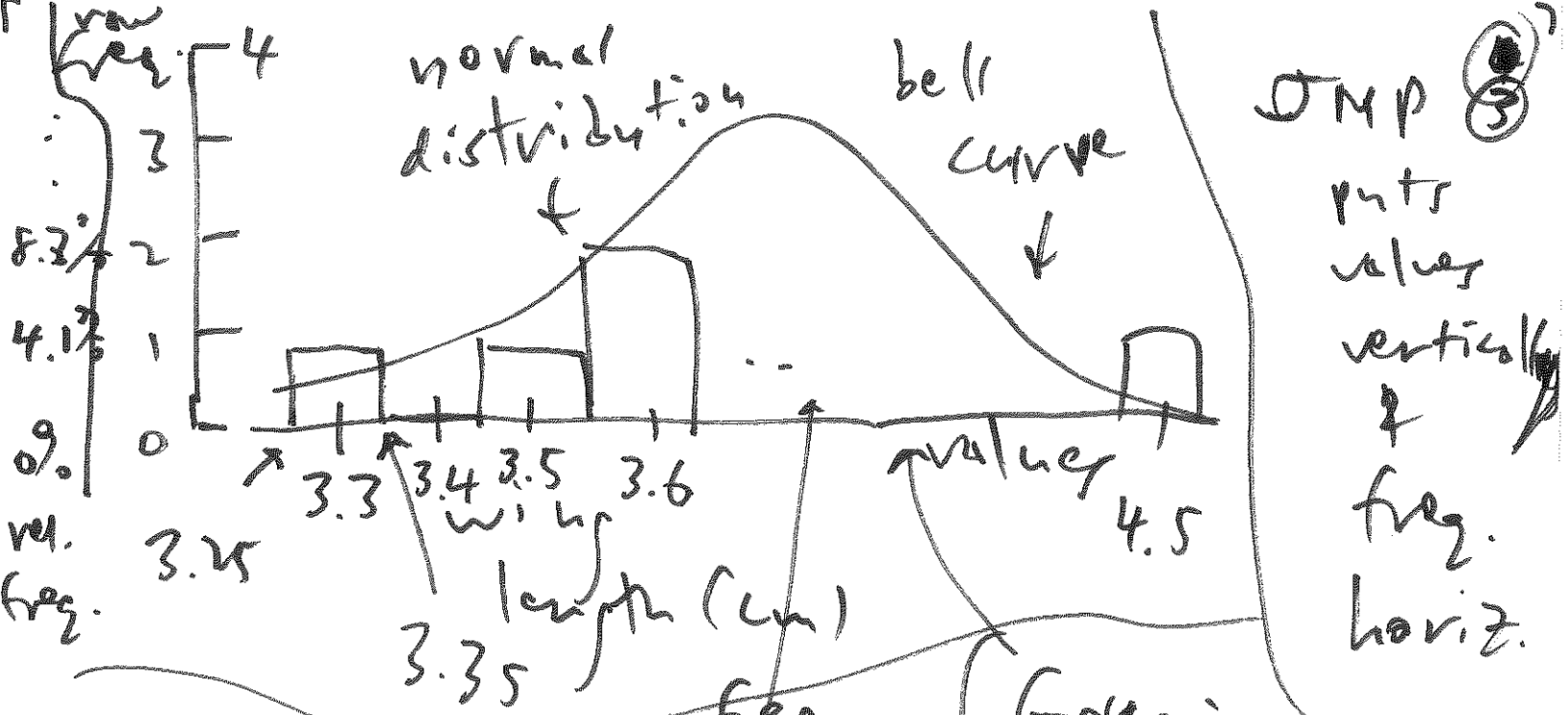
sort →

3.3 ← minimum
3.5
3.6
3.6 ← 2 significant figures (sigfigs)
⋮
4.5 ← maximum

values	count
3.3	1
3.4	0
3.5	1
3.6	3
⋮	⋮
4.5	1

count = raw frequency

raw frequency distribution



Carl Friedrich Gauss 1777-1855

Gaussian distribution

(a kind of bargraph)

single var = quantitative

conceptually continuous

but made discrete by rounding in the measurement process

histograms: only for quant. var (either d.sc. or cont.)

visualize the raw dataset

eye color
brown
black
brown
blue
⋮

$n \doteq 180$

1 row for each person in AN7 today (sec 1)

nominal

qual.

columns = variables

~~histogram~~

rows = subjects

raw freq.

bar graph

qual.

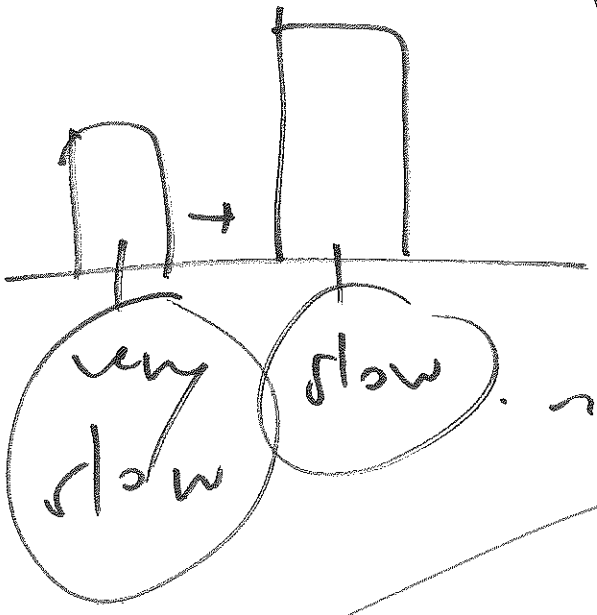
brown ⋮ brown	89
black ⋮ black	57
⋮	

hist. → quant

value	raw freq	relative freq (%)
3.3	1	$\frac{1}{24} \times 100\% \approx 4.2\%$
3.4	0	0%
3.5	1	$\frac{1}{24} \approx 4.$
3.6	2	$\frac{2}{24} \approx 8.3\%$

3 different vertical scales for histograms

① raw freq.
vs.
② relative freq.
③ vs. density scale



maze running

qual., ordinal

Q: is there just 1 unique hist. for any given data set?

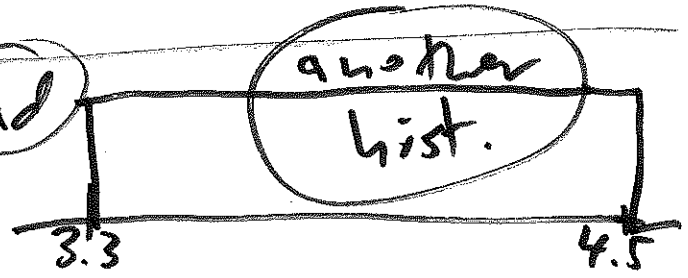
A: no - you have to make a good choice of # of bars

butterfly: n = 24

for 13 bars

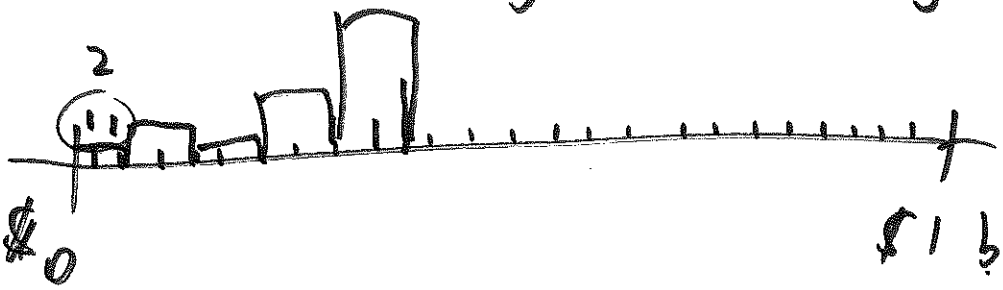
all sense of dist. shape is lost

too few bars



too few bars	bad	all sense of dist. shape is lost
just right	good	
too many bars	bad	can't see the forest for the trees — too much detail

to cut # bars in half
combine adjacent bars



location
vine
building ent
vine
⋮
X

Sparrows

$n = 56 + 60 + 46 + 49$

1 row for each nest (in site)

qual.

nominal
not dich.

bar graph? y
hist? no

---

pigment type
0 faintly speckled
2
5
0
⋮

$n = 13 + .. + 8$

1 row for each sun fish

qual. even when we code 0, ..., 4

bar graph ✓ ordinal.

hist. no