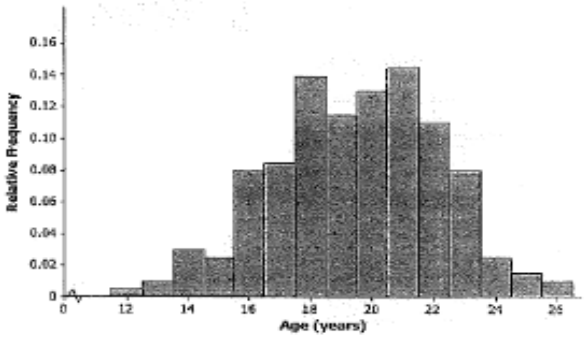
Name KEY

Algebra 2 Homework 13-1

A local utility company wanted to gather data on the age of air conditioners that people have in their homes. The company took a random sample of 200 residents of a large city and asked if the residents had an air conditioner, and if they did how old it was. Below is the distribution in the reported ages of the air conditioners.



1. Would you describe this distribution of air conditioner ages as approximately symmetric or as skewed? Explain your answer.

~ Symmetric. The left & right sides of distribution are similar.

2. Is the mean of the age distribution closer ~~to 15~~, 20, or 25? Explain your answer.

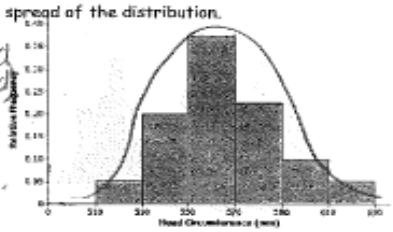
The distribution is centered at about 20, so the mean of the age distr. is closer to 20.

3. Would you use the range or the IQR to describe the spread of this distribution? Explain your answer.

range since distrib. is ~ symmetric

4. For each of the following histograms, describe the shape, and give an estimate of the mean and would you use the range or IQR to describe the spread of the distribution.

a. Distribution of head circumferences (mm)



Shape: ~ symmetric (mound shaped)

Mean: ~ 560

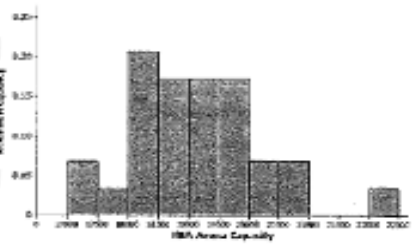
Spread: Range since distr. is ~ symmetric

b. Distribution of NBA arena seating capacity

Shape: ~ symmetric (mound shaped) or slightly skewed right

Mean: ~19,000

Spread: if symmetric → use range; if skewed → use IQR

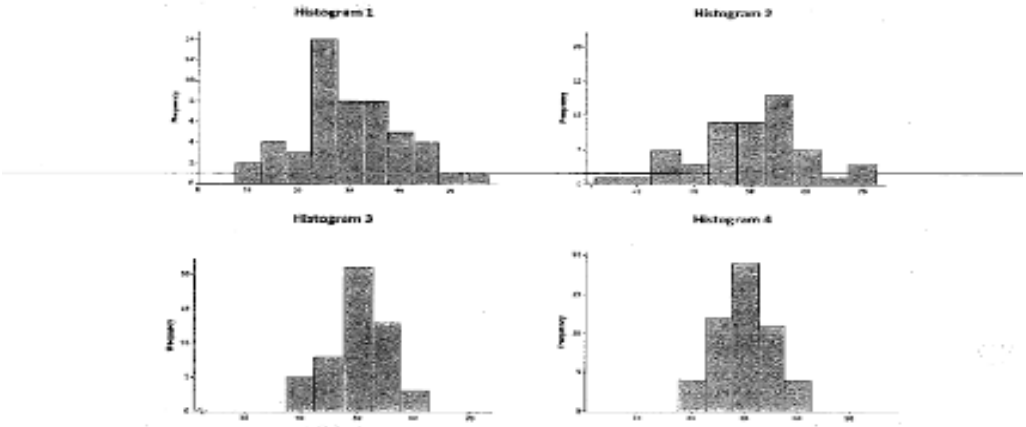


5. For each of the following, match the description of a distribution with the appropriate histogram.

Histogram	Distribution
1	C
2	B
3	A
4	D

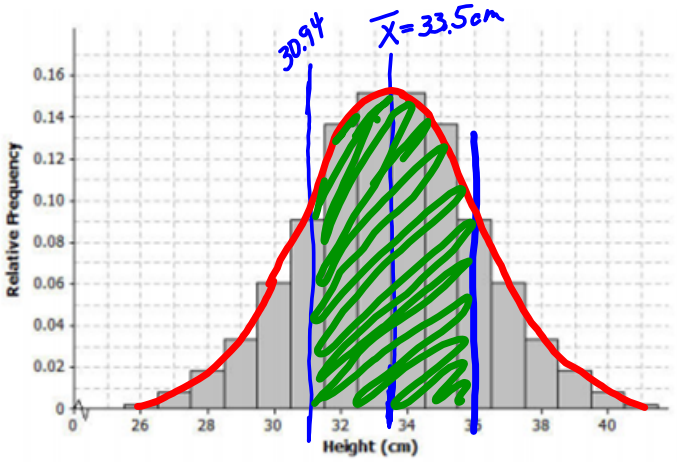
Description of distributions:

Distribution	Shape	Mean	Standard Deviation
A	Approximately symmetric, mound shaped	50	5
B	Approximately symmetric, mound shaped	50	10
C	Approximately symmetric, mound shaped	30	10
D	Approximately symmetric, mound shaped	30	5



Warm-Up:

A paleontologist studies prehistoric life and sometimes works with dinosaur fossils. Data was collected on the distribution of heights (rounded to the nearest ~~inch~~ cm) of 660 procompsognathids or "compys." The heights were determined by studying the fossil remains of the compys. The following is a relative frequency histogram of the compy heights.



1. What does the relative frequency of 0.136 mean for the height of 32 cm?

13.6% of the 660 compys were 32cm tall.
(That bar represents 31.5 to 32.49)

2. What's the width of each bar? What does the height of the bar represent?

Each bar has a width of 1 cm.

Height of bar = proportion of compy's that had that height

3. What is the area of the bar that represents the relative frequency for compys with a height of 32 cm?

$A = b \times h = 1\text{cm} \times$ Rel. Freq.

$A = 1\text{cm} \times .136\text{cm} = .136\text{cm}^2$ = Rel. Frequency

4. The mean of the distribution of compy heights is 33.5 cm, and the standard deviation is 2.56 cm. Interpret the mean and the standard deviation in this context.

The mean of 33.5 cm is the avg. height of compys in the sample, interpreted as typical compy height. A typical height of a compy is within 2.56 cm of the average of 33.5 cm.

5. Mark the mean on the graph and mark one deviation above and the below the mean.
   a. Approximately what percent of the values in this data set are within **one** standard deviation of the mean? (i.e. between 33.5-2.56 = 30.94 cm and 33.5 + 2.56 = 36.06 cm)

$.15 + .15 + .136 + .136 + .09(.5) + .09(.5) = .662$

$\sim 2/3$

6. Draw a smooth curve that comes reasonably close to passing through the midpoints of the tops of the bars in the histogram. Describe the shape of the distribution.

Symmetric (mound) shaped

7. Shade the area under the curve that represents the proportion of heights that are within one standard deviation of the mean.

8. Based on our analysis, how would you answer the questions, "How tall was a compy?"

A typical compy is approximately 31 to 36 cm tall.

*A normal curve is a smooth curve that is symmetric and bell shaped.

*Data distributions that are mound shaped are often modeled using a normal curve, and we say that such a distribution is approximately normal.

*The compy heights from our warm up is approximately normal.

Standard Deviation (Two Types):

1. Population Standard Deviation (σ_x): values in the dataset represent an entire **population** very seldom used
2. Sample Standard Deviation (s_x): values in the dataset represent a **sample** of a larger population usually have a sample of data

Example: A salesman kept track of the gas mileage for his car over a 25-week span. The mileages (miles per gallon rounded to the nearest whole number) were

23, 27, 27, 28, 25, 26, 25, 29, 26, 27, 24, 26, 26, 24, 27, 25, 28, 25, 26, 25, 29, 26, 27, 24, 26

L1

a. Use technology to find the mean and standard deviation of the mileage data. How did you use technology to assist you?

Stat - Calc - 1Var Stats

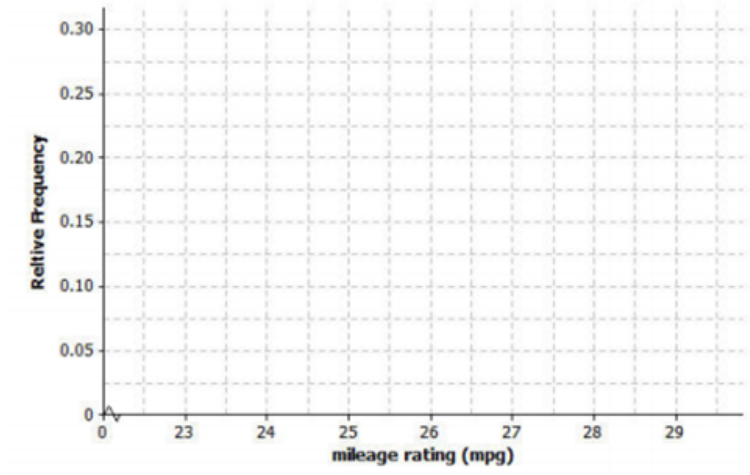
$\bar{x}$ Mean = 26.04

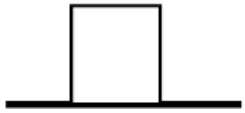
S_x Standard Deviation = 1.54

b. Calculate the relative frequency of each of the mileage values. For example, the mileage of 26 mpg has a frequency of 7. To find the relative frequency, divide 7 by 25, the total number of mileages recorded. Complete the following table.

Mileage	Frequency	Relative Frequency (%)
23	1	1/25 = .04
24	3	= .12
25	5	= .20
26	7	= .28
27	5	= .20
28	2	= .08
29	2	= .08
Total	25	1

c. Construct a relative frequency histogram using the scale below.



FYI: When you see histograms like above,  the box of 26 mpg represents #'s from 26.0 to 26.9

d. Describe the shape of the mileage distribution. Draw a smooth curve that comes reasonably close to passing through the midpoints of the tops of the bars in the histogram. Is this approximately a normal curve?

Shape: Roughly Symmetric
yes approx. normal

e. Mark the mean on the histogram. Mark one standard deviation to the left and right of the mean. Shade the area under the curve that represents the proportion of data within one standard deviation of the mean.

$\bar{x} - 1s_x = 26.04 - 1.54 = 24.5$

$\bar{x} + 1s_x = 26.04 + 1.54 = 27.58$

Remember your test grades of 100, 95, 98, 10? What did we say happened to your average after you got the 10?

Dropped. The 10 pulled the mean toward the low grade

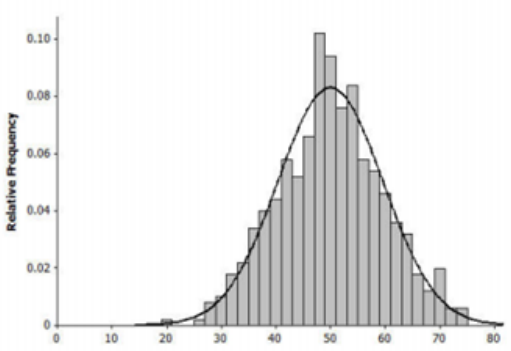
*Is the mean of a distribution that is approximately normal located near where the curve is the highest?

Yes

*Is the mean of a skewed distribution located where the curve is the highest? Why does this happen?

No - the means pulled toward the tail

*A normal curve is symmetric and bell shaped.



*Areas under a normal curve can be used to estimate the proportion of the data values that fall within a given interval.

*When a distribution is skewed, it is NOT appropriate to model the data distribution with a normal curve.