Statistical Measures

Learning Goals

- **uses of mean, median, and mode**
- **weighted mean**
- **percentiles**
- **per capita**

Collect the height of all students and order it from least to greatest.

145, 158, 165, 168, 168, 175, 176,
176, 178, 180, 186, 188, 193,

What is the **mean**?

$$\frac{\text{sum}}{13} = \frac{2256}{13} = 173.5$$

What is the **median**? 176

What is the **mode**? 176, 168

Which one is the best indicator of average?

Depends on what you want to find out.

Mean --- average amount of $ students spend on thier lunch

Median --- average class mark

Mode --- sizes of clothes most people wear (ordering purposes in a clothing store)

Weighted Mean - mean where some data points contribute more than others

ex. Your marks ---

Knowledge	25%
Application	20%
Thinking	15%
Communication	10%
CPT	15%
EXAM	15%

Weighted Mean

A college instructor uses a weighted mean to calculate her students' marks. Xin's and David's marks are shown, along with the weighting.

	Xin	David
Quiz (out of 30)	24	18
Assignment (out of 60)	40	46
Test (out of 120)	89	95
Independent Study (out of 60)	43	48
Performance Task (out of 100)	85	82
Final Exam (out of 90)	72	64

Weighting Factors
Q = quiz (15%)
A = assignment (15%)
T = test (25%)
IS = independent study (10%)
PT = performance task (10%)
EX = final exam (25%)

Who had the better percent overall mark and by how much?

Weighted Mean

A college instructor uses a weighted mean to calculate her students' marks. Xin's and David's marks are shown, along with the weighting.

	Xin	David
Quiz (out of 30)	24	18
Assignment (out of 60)	40	46
Test (out of 120)	89	95
Independent Study (out of 60)	43	48
Performance Task (out of 100)	85	82
Final Exam (out of 90)	72	64

Weighting Factors
Q = quiz (15%)
A = assignment (15%)
T = test (25%)
IS = independent study (10%)
PT = performance task (10%)
EX = final exam (25%)

Who had the better percent overall mark and by how much?

Solution

Xin

$\frac{24}{30}(15) + \frac{40}{60}(15) + \frac{89}{120}(25) + \frac{43}{60}(10)$

Mark (Weight)

$+ \frac{85}{100}(10) + \frac{72}{90}(25) = 76.2$

David

$\frac{18}{30}(15) + \frac{46}{60}(15) + \frac{95}{120}(25) + \frac{48}{60}(10)$

$+ \frac{82}{100}(10) + \frac{64}{90}(25) = 74.3$

Quartiles

A **quartile** is any of three numbers that separate a set of sorted data into 4 equal parts.

1st Quartile
-The median of the lower half of the data
-It separates the lowest 25% of the data.
-Same as the 25th percentile

2nd Quartile
-The median
-Cuts the data set in half
-Same as the 50th percentile

3rd Quartile
-The median of the upper half of the data
-It separates the highest 75% of the data
-Same as the 75th percentile

Find the 1st and 3rd quartile for the data set above on your classes heights.

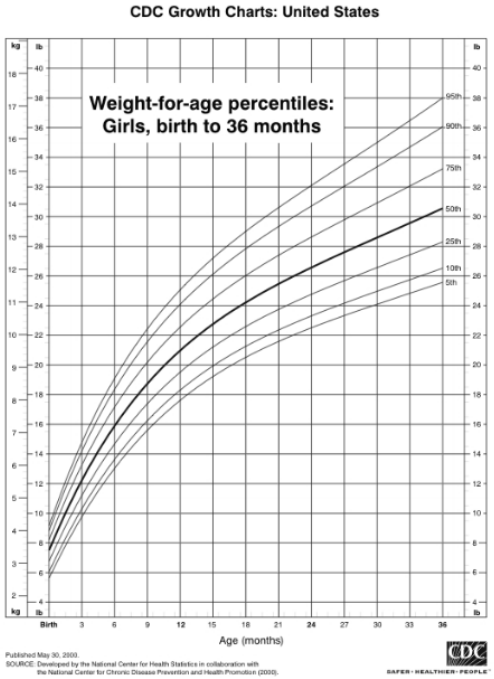
1st 165, 168 $\Rightarrow \frac{165+168}{2} = 166.5$

3rd 180, 186 $\Rightarrow \frac{180+186}{2} = 183$

Percentiles

A **percentile** tells approximately what percent of the data are below a particular value. They are a good way to rank data when you have a large data set or want to keep it private.

ex. growth charts for babies



Percentile

ex. Who in the class is in the 85th percentile for height?

Position of 85th percentile = Number of students x 0.85

$13(0.85) = 11.05$

round down

11

↑ 11th person ∴ 186

percentile that you are looking for

Determine what percentile you are in from the class data set on height.

$\frac{\text{position}}{\text{total}} = \frac{}{13}$

Examples on next page of handout

Another value that is often used when an individual's performance is compared with that of a group is a **percentile.** A score in the 65th percentile indicates that the individual scored as high or higher than 65% of the people who were tested.

Daina's mark of 68 placed her 7th out of the 34 candidates that wrote; that is, 27th from the bottom.

Stem	Leaf
3	5 8 8
4	0 4 6 6 7 8
5	0 1 2 3 5 7 8
6	1 2 2 3 3 4 4 4 5 7 8
7	2 3 6
8	2 4
9	3

Daina's mark

We can then calculate her percentile ranking.

$\frac{\text{position from bottom}}{\text{total}}$

$= \frac{27}{34}$

$= 79\%$

Daina placed in the 79th percentile. Percentiles gave her a precise statement of her achievement on the test, relative to the other candidates.

Example 2

André was told that his performance in the high jump placed him in the 87th percentile of competing athletes. What was André's rank if 134 athletes completed?

position

$\frac{\text{position}}{\text{total}} = \text{percentile}$

$\frac{x}{134} = 87\%$

$x = 116.5$

$\therefore$ position is 117th from the bottom

$\therefore$ André's rank is 17^{th}

$134 - 117$

Nineteen swimmers swam lengths of the pool to support a local charity. The number of lengths completed by each swimmer is shown below.

35	44	54	25	40
25	19	17	43	16
61	27	38	74	22
29	30	50	33	

(a) What is the range?
(b) Find the median, upper and lower quartiles, and ~~interquartile range~~.
~~(c) Draw a box-and-whisker plot.~~
(d) Raji swam 44 lengths. What is his percentile ranking?
(e) Mila swam 29 lengths. What is her percentile ranking?

16 17 19 22 [25] 25 27 29 30 [33] 35 38 40 43 [44] 50 54 61 74

range = largest - smallest
= 74 - 16
= 58

Median 33

lower quartile 25
upper quartile 44

d.) 44 laps $\Rightarrow$ upper quartile
$\therefore$ 75%

e.) 29 laps $\Rightarrow \frac{\text{position from bottom}}{\text{total}} = \frac{8}{19}$
$= 42\%$

The following scores were achieved on a recent math test.

72	63	59	68	53	64	88
44	43	72	67	48	90	54
66	65	76	81	74	79	74
86	95	52	61	68	57	37

(a) Complete a stem-and-leaf plot.
(b) Determine the median, upper quartile, lower quartile, and ~~interquartile range~~.
(c) ~~Draw a box-and-whisker plot to illustrate the data.~~
(d) Jeanna achieved 52 on the test. Indicate her score on your ~~box-and-whisker plot~~ and calculate her percentile ranking.
(e) Repeat (d) for Norma's score of 86.

a.

stem	leaf
3	7
4	3 4 8
5	2 3 4 7 9
6	1 3 4 5 6 7 8 8
7	2 2 4 4 6 9
8	1 6 8
9	0 5

b. median (66+67) / 2 = 66.5
1st quartile 54
2nd quartile 76

d. $\frac{\text{position}}{\text{total}} = \frac{5}{28} = 18\%$ e. $\frac{\text{position}}{\text{total}} = \frac{25}{28} = 89\%$

Data Reliability

When interpreting statistical information, the reliability of the source needs to be considered. List several examples under the categories below:

Reliable Sources	Skeptical Sources
scientific journals	anyone trying to sell you things
government websites	interest groups

Polls are interviews with randomly selected samples that determine opinions on topics. They will report the **margin of error** for a particular poll.

For example: "*the margin of error is ±3%, 19 times out of twenty.*"

What does this mean?

50% ⇒ 47% to 53%

Topic	*Source 1*	*Source 2*
The sound quality of a cell phone.	A commercial for the cell phone. selling	A news article comparing cell phone sound quality. ✳
The health benefits of a diet.	The diet book's introductory chapter. bias	A physician on a talk show. ✳
The views of a politician.	The politician's running mate's website. bias	The politician's website. ✳
The best tasting pizza in a city.	A public opinion award run by the local T.V. station. ✳	One pizzeria's advertisement in the paper. selling
The symptoms of a syndrome you think you might have	A medical journal ✳	A web MD website.
The number of car accidents in a year caused by poor road conditions.	Statistics Canada website ✳	Goodyear snow tires advertisement.
The benefits or adverse affects of drinking milk.	A pamphlet from an animal rights group that opposes dairy farming bias	Canada's Food Guide produced by Health Canada. ✳
The effects of logging on population of a species of bird.	A pamphlet from a wildlife protection organization. bias	A forestry company advertisement. neither bias
Possible complications of the flu shot.	A Ministry of Health Web site. ✳	A Web site run by a group that opposes immunizations. bias
The benefits of a separate Catholic School System	The blog of a parent opposing faith based schools. bias	A Church newsletter. neither bias

Seatwork

pg 201 # 1-4, 7-9