

DATA ANALYSIS & PROBABILITY

Statistics Review

MEASURES OF CENTRAL TENDENCY

Objective: IWBAT use measures of central tendency
to summarize data sets.

Vocabulary:

- **Measure of Central Tendency** – (*Mean, Median, Mode*) – Used to organize and summarize a set of data
- **Mean** – (average) find the sum of the data values and divide by the number of data values in the set

$$\text{mean} = \frac{\text{sum of data values}}{\text{total number of data values in set}}$$

- **Median** – the middle value in an ordered set of data values; for a set with an even number of data, the median is the mean of the two middle values
- **Mode** – most frequently occurring value (or values) in a data set. Data set may have no mode, one mode, or more than one mode

Vocabulary:

- Range of a data set – difference between the greatest and least data values.

Example: Use the data set to find the mean, median, mode, and range of the data set.

58	62	66	70	76	78	81	84	77	73	68	63
----	----	----	----	----	----	----	----	----	----	----	----

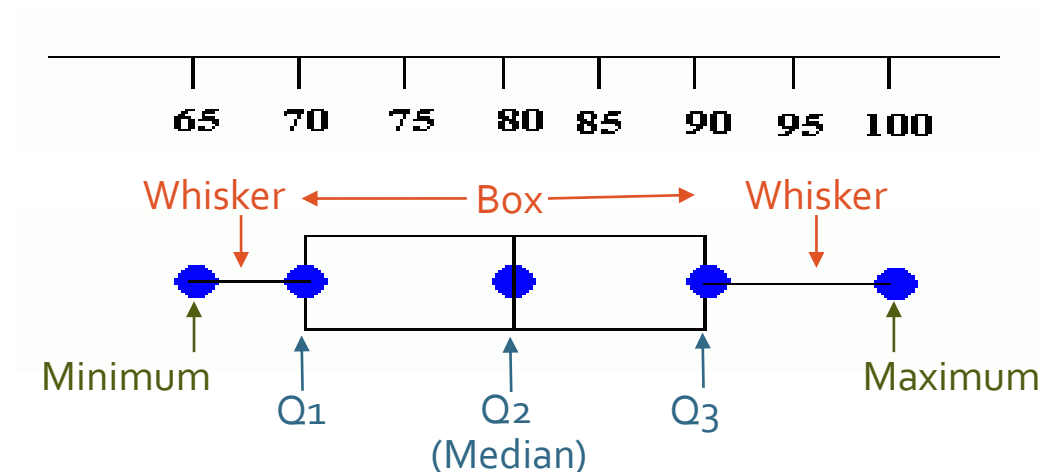
1. Order the data.
2. Calculate the mean.
3. Identify the median.
4. Is there a mode? If yes, what is it? Is there more than one? If yes, what are the others?
5. Calculate the range.

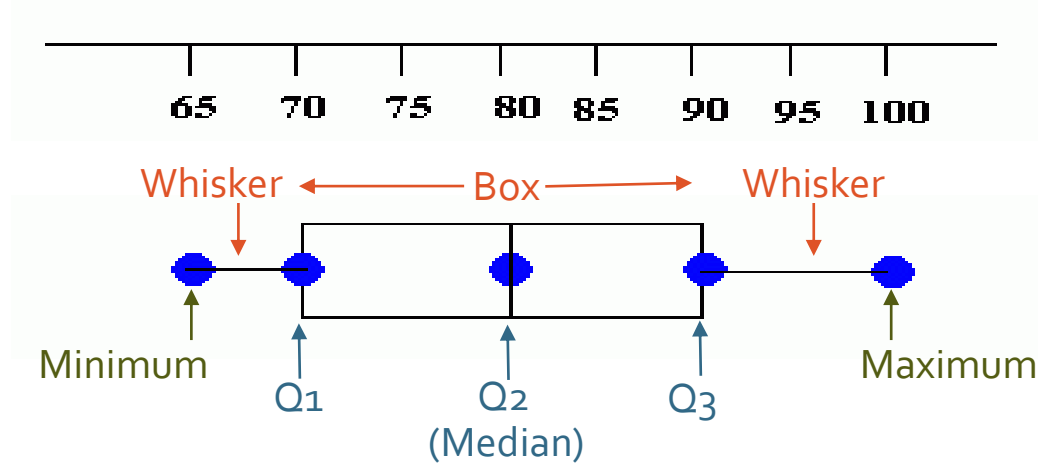
BOX-AND-WHISKER PLOTS

Objective: IWBAT make and interpret box-and-whisker plots,
find quartiles and percentiles

Vocabulary:

- **Box-and-Whisker Plot** – graph that summarizes a set of data by displaying it along a number line; it consists of 3 parts: a box, 2 whiskers
- **Quartiles** – values that divide a data set into 4 equal parts; Q2 is the median of the data set, Q1 is the median of the lower half of the data, Q3 is the median of the upper half
- **Interquartile Range** – the difference between the third (Q3) and first (Q1) quartiles





- The **LEFT Whisker** extends from the minimum to Q1. It represents about 25% of the data.
- The **BOX** extends from Q1 to Q3 and has a vertical line through the median. The length of the box represents the ***interquartile range*** and contains about 50% of the data.
- The **RIGHT Whisker** extends from Q3 to the maximum. It represents about 25% of the data.

For an odd number of data values, do not include the Median in either half when finding the 1st and 3rd Quartiles.

Example: Draw a box-and-whisker plot using the data provided below.

314	321	315	316	314	311	307	316	312	314	303
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

1. Order the data.
2. What is the minimum? Maximum?
3. Identify the median (Q_2).
4. Find Q_1 , the median of the lower half of the data.
5. Find Q_3 , the median of the upper half of the data.
6. Find Q_1 , Q_2 , Q_3 , the minimum, and the maximum on the number line.
7. Draw the box. Extend the whiskers to the minimum and to the maximum.

PROBABILITY

Objective: IWBAT find theoretical and experimental probabilities; find probabilities of mutually exclusive and overlapping events; find probabilities of independent and dependent events

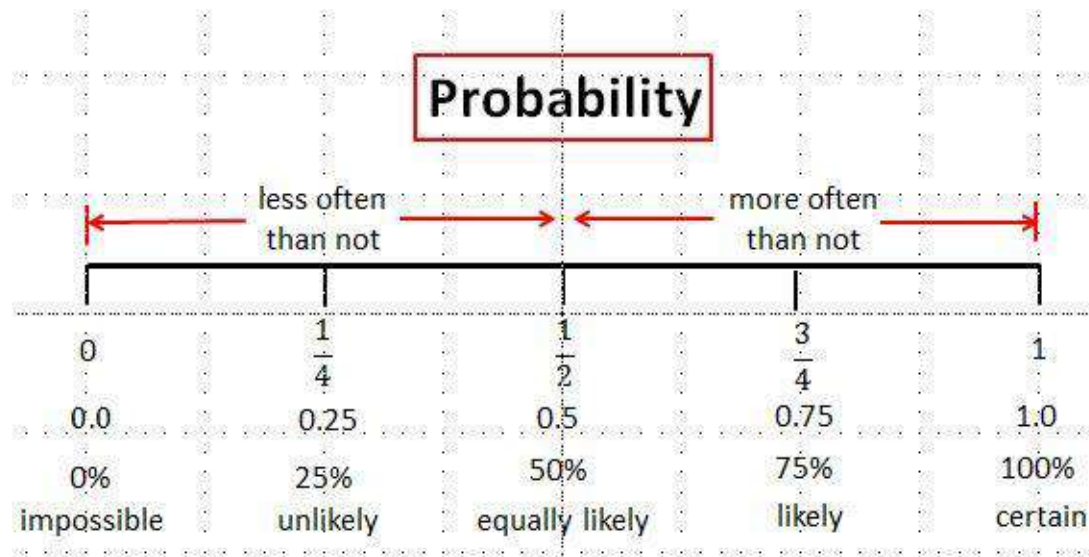
Vocabulary:

- **Outcome** – result of a single trial (example: spinning a wheel)
- **Sample Space** – all possible outcomes
- **Event** – any outcome or group of outcomes

Event	Sample Space	Favorable Outcomes
Rolling an Even #	1, 2, 3, 4, 5, 6	2, 4, 6

Vocabulary:

- **Probability** of an event, or $P(\text{event})$, tells how likely it is that an event will occur.
 - Probability can be written as a fraction, decimal, or percent
 - Probability of an event ranges from 0 to 1.



Vocabulary:

- **Theoretical Probability** – when all possible outcomes are equally likely to occur

$$P(\text{event}) = \frac{\text{number of favorable outcomes}}{\text{number of possible outcomes}}$$

Vocabulary:

- **Odds** – describe the likelihood of an event as a ratio comparing the number of favorable and unfavorable outcomes

$$\text{odds in favor} = \frac{\text{number of favorable outcomes}}{\text{number of unfavorable outcomes}}$$

$$\text{odds against} = \frac{\text{number of unfavorable outcomes}}{\text{number of favorable outcomes}}$$

Vocabulary:

- Compound Event – consists of two or more events linked by the word “and” or “or”
- Mutually Exclusive – two events that have no outcomes in common; $P(A \text{ and } B) = 0$.
- Overlapping Events – events that have at least one outcome in common

Probability of Mutually Exclusive Events:

$$P(A \text{ or } B) = P(A) + P(B)$$

Probability of Overlapping Events:

$$P(A \text{ or } B) = P(A) + P(B) - P(A \text{ and } B)$$

Vocabulary:

- **Independent Events** – the occurrence of one event does not affect the probability of the second event

Probability of Two Independent Events:

$$P(A \text{ and } B) = P(A) \bullet P(B)$$

- **Dependent Events** – the occurrence of one event affects the probability of the second event

Probability of Two Dependent Events:

$$P(A \text{ then } B) = P(A) \bullet P(B \text{ after } A)$$