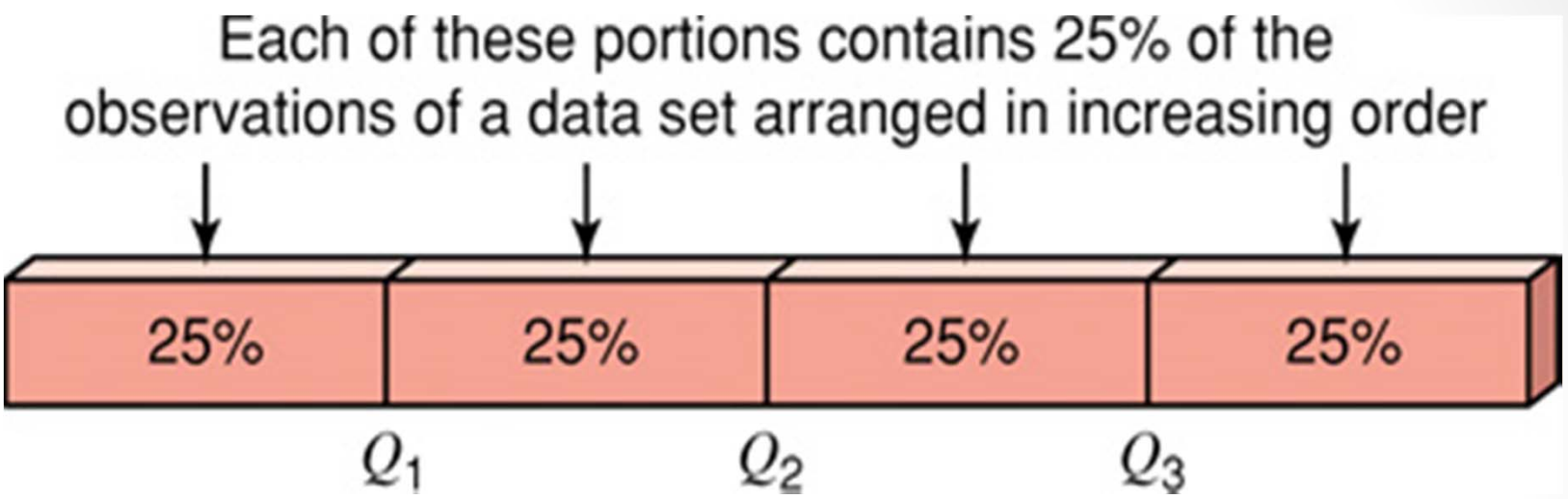
3.5 Measures of Position

Definition. The **quartiles** are the three summary measures that divide an ordered data set into four equal parts.

Notation. The first quartile = Q_1, the second quartile = Q_2, and the third quartile = Q_3.


Each of these portions contains 25% of the observations of a data set arranged in increasing order
25%
25%
25%
25%
Q_1
Q_2
Q_3

How to compute quartiles?

Note that 50% of observations lie below the second quartile, and 50% lie above. Hence, by definition, the second quartile is the median.

We do know how to compute the median.

Now, to compute Q_1, consider only the ordered observations that lie **below** the median, and find the median of this ordered dataset.

To compute Q_3, consider only the ordered observations that lie **above** the median, and find the median of this ordered dataset.

Example. The raw data are

7 9 6 8 3 1 4 3 4 8

Find the three quartiles.

Solution: The ordered dataset is

1 3 3 4 4 6 7 8 8 9

The median is $Q_2 = \frac{4+6}{2} = 5$.

Now consider only the observations that lie **below** Q_2 1 3 3 4 4

The median of this set is $Q_1 = 3$.

Now consider only the observations that lie **above** Q_2 6 7 8 8 9

The median of this set is $Q_3 = 8$.

3.6 Box-and-Whiskers Plot

Definition. The **five number summary** of a data set consists of

$$\min, Q_1, Q_2(\text{median}), Q_3, \max.$$

Example. The five number summary for our ordered data set

1 3 3 4 4 6 7 8 8 9

is

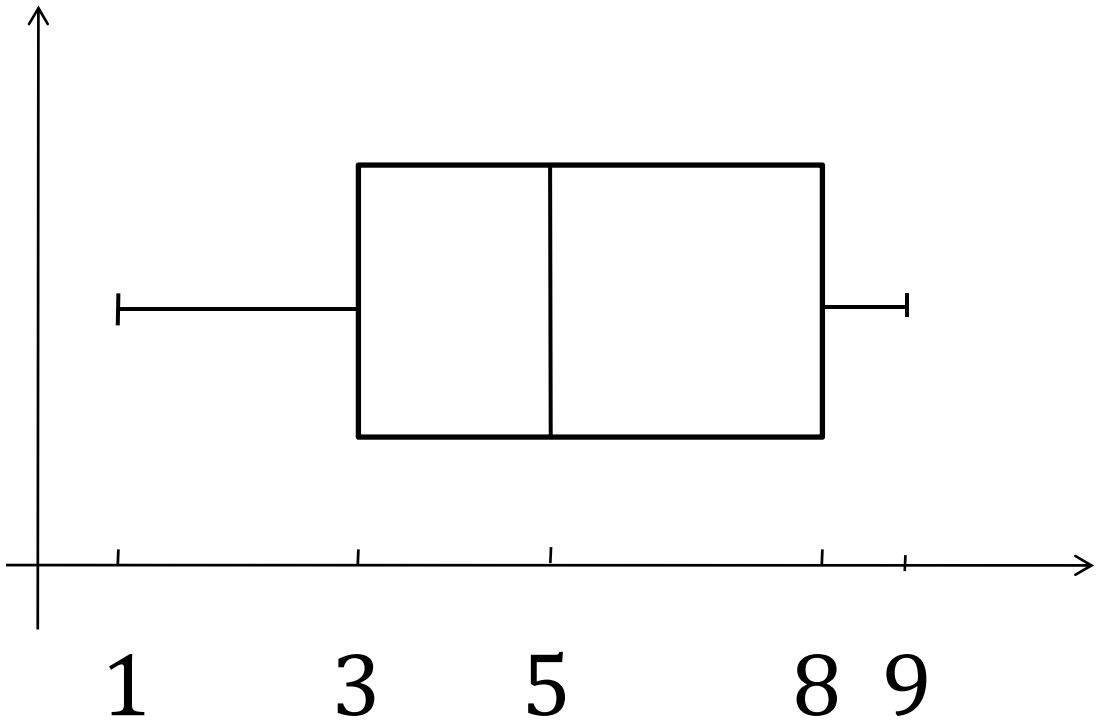
min = 1, Q_1 = 3, Q_2(median)=5, Q_3=8, max=9

Definition. A **box-and-whiskers plot** (or, simply, a **box plot**) displays the five number summary, by drawing a box between Q_1 and Q_3 with a line at Q_2, and extending "whiskers" to min and max.

Example. For our data set,

min = 1, Q_1= 3, Q_2(median)=5, Q_3=8, max=9.

The box plot looks like this:



Definition. Interquartile range is the difference between the third and first quartiles, that is, $IQR = Q_3 - Q_1$.

Example. For our data set,

$IQR = Q_3 - Q_1 = 8 - 3 = 5$.

Exercise. Construct a box plot and compute the interquartile range for the following unordered set of observations:

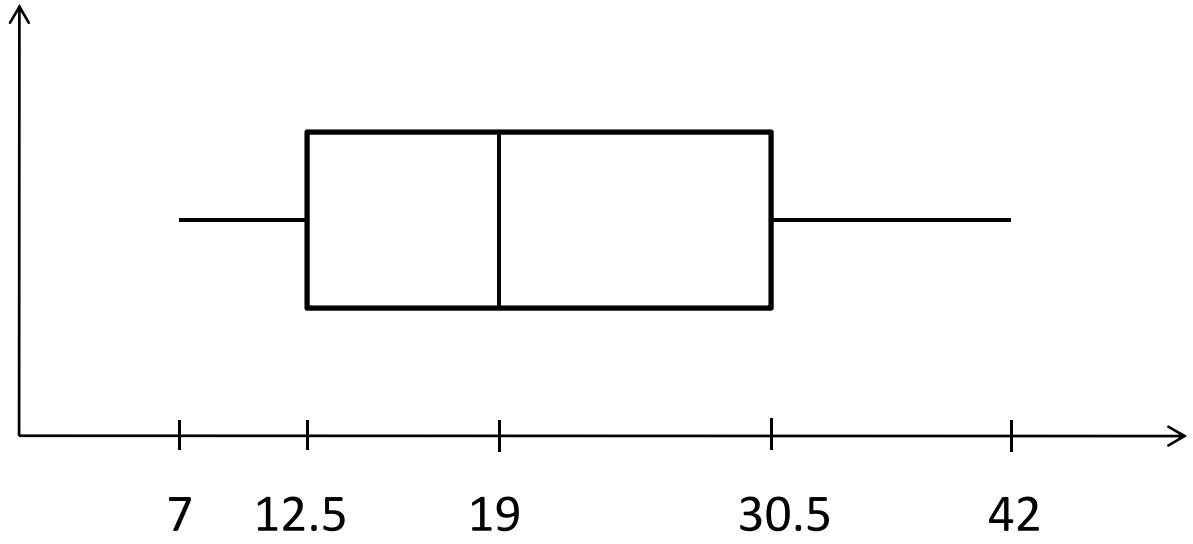
11 8 26 31 19 7 14 33 30 42 15 18 24

Solution. The ordered dataset is

7 8 11 14 15 18 19 24 26 30 31 33 42

$$\min=7,\ Q_1=\frac{11+14}{2}=12.5,\ Q_2=19,$$

$$Q_3=\frac{30+31}{2}=30.5,\ \max=42,\ \text{IQR}=30.5-12.5=18$$



4.1 Experiment, Outcomes, and Sample Space

Definition. A **random experiment** is a procedure that can be repeated as many times as needed, and has a well-defined set of possible outcomes, but outcomes are uncertain on every trial.

Examples. (1) Flipping a coin. Outcomes are either heads or tails (H or T). Exact outcome for a particular flip is unknown.

heads **tails**

(2) Rolling a die. The outcomes are either 1, 2, 3, 4, 5, or 6.

("die" – singular,
"dice" – plural)

A pair of dice, two dice, three dice, one die

Definition. A **sample space** S is a set of all possible outcomes of a random experiment.

Examples.

- Flipping a coin once. $S = \{H, T\}$
- Flipping a coin twice. $S = \{HH, HT, TH, TT\}$
- Tossing a die. $S = \{1,2,3,4,5,6\}$

Definition. An **event** is a collection of one or more outcomes of a random experiment.

Note that an event is a subset of the sample space.

Notation. Events are denoted by A, B, C, D, E, F, G (capital letters in the beginning of the alphabet), A_1, A_2, etc.

Definition. An **empty event** is an event that consists of no outcomes.

Notation. An empty event is denoted by $\varnothing$.

Examples. A coin is flipped two times.

- List the outcomes in the event *A = the second flip results in a head*.

Answer. $A = \{HH, TH\}$. The other two possible outcomes HT and TT do not satisfy this description.

- List the outcomes in the event *B = at least three heads appear*.

Answer. $B = \varnothing$. There are no outcomes that satisfy this description.

Examples. Gender of a person who comes into a room is recorded (M=male, F=female). Three persons entered the room.

- Find the sample space.

Answer.

$$S = \{MMM, MMF, MFM, FMM, MFF, FMF, FFM, FFF\}$$

- List all outcomes in the event

E = at least two females entered the room.

Answer. $E = \{MFF, FMF, FFM, FFF\}$

- Describe in words the event

$$G = \{MMM, MMF, FMM, FMF\}$$

Answer. The outcomes of the event $G = \{MMM, MMF, FMM, FMF\}$ satisfy the description: the second person who entered was male. By visual inspection, no other outcomes in the sample space $S = \{MMM, MMF, MFM, FMM, MFF, FMF, FFM, FFF\}$ satisfy this description. Therefore, *G = second person entered was male*.

Definition. A **simple event** is an event that consists of a single outcome.

Notation. Simple events are denoted by E, E_1, E_2, E_3, etc.

Example. A coin is flipped twice. Each outcome is a simple event.

$$E_1 = \{HH\}, E_2 = \{HT\}, E_3 = \{TH\}, E_4 = \{TT\}$$

Definition. A **compound event** is an event that consists of more than one outcome.

Example. A coin is flipped two times. The event *A = the second flip results in a head ={HH, TH}* is a compound event. It consists of two simple events *{HH}* and *{TH}.*