Bayesian Notation

A **set** is a collection of elements, which hold certain values. Additionally, every event has a set of outcomes that satisfy it.

The **null-set** (or **empty set**), denoted ∅, is an set which contain no values.

$$x \in A$$

Element (lower-case) Set (upper-case)

Notation:	Interpretation:	Example:
$x \in A$	"Element x is a part of set A."	$2 \in All\ even\ numbers$
$A \ni x$	"Set A contains element x."	$All\ even\ numbers \ni 2$
$x \notin A$	"Element x is NOT a part of set A."	$1 \notin All\ even\ numbers$
$\forall x$:	"For all/any x such that…"	$\forall x: x \in All\ even\ numbers$
$A \subseteq B$	"A is a subset of B"	$Even\ numbers \subseteq Integers$

Remember! Every set has at least 2 subsets.

- $A \subseteq A$
- $\emptyset \subseteq A$

Examples of Intersection

$$A \cap B = \emptyset$$

The intersection of all hearts and all diamonds is the EMPTY SET

There are no outcomes which satisfy both events simultaneously

1.00

Sets

Event		Set of outcomes (favourable outcomes)
Even		2, 4, 6...

Sets

Values of a set don't always have to be numerical

Event ➡ being a member of the European Union

France

Germany

USA

Japan

Sets

Set ➡ UPPER- CASE

Elements ➡ lower- case

Sets

"even" → X

8 → x

Sets

Any set can be either **empty** or have values in it

the empty set = the null set

$\varnothing$

1.50

Part of a set

"$x \in A$"

element set

x is an element of the set A

Part of a set

$$\text{"}A \ni x\text{"}$$

set element

A contains x

NOT part of a set

"$x \notin A$"

element set

"$A \not\ni x$"

set element

x is not in A

A does not contain x

Generalized Statement about multiple elements

$\forall$ ➡ for all/any

"$\forall x \in A$" ➡ for all x in A

Semi-colon

$\forall \; :$

$\downarrow$

<u>such that</u>

Incredibly useful when we want to make statements about a specific group of elements within a set

Semi-colon

$$\forall\, x \in A : x \text{ is even}$$

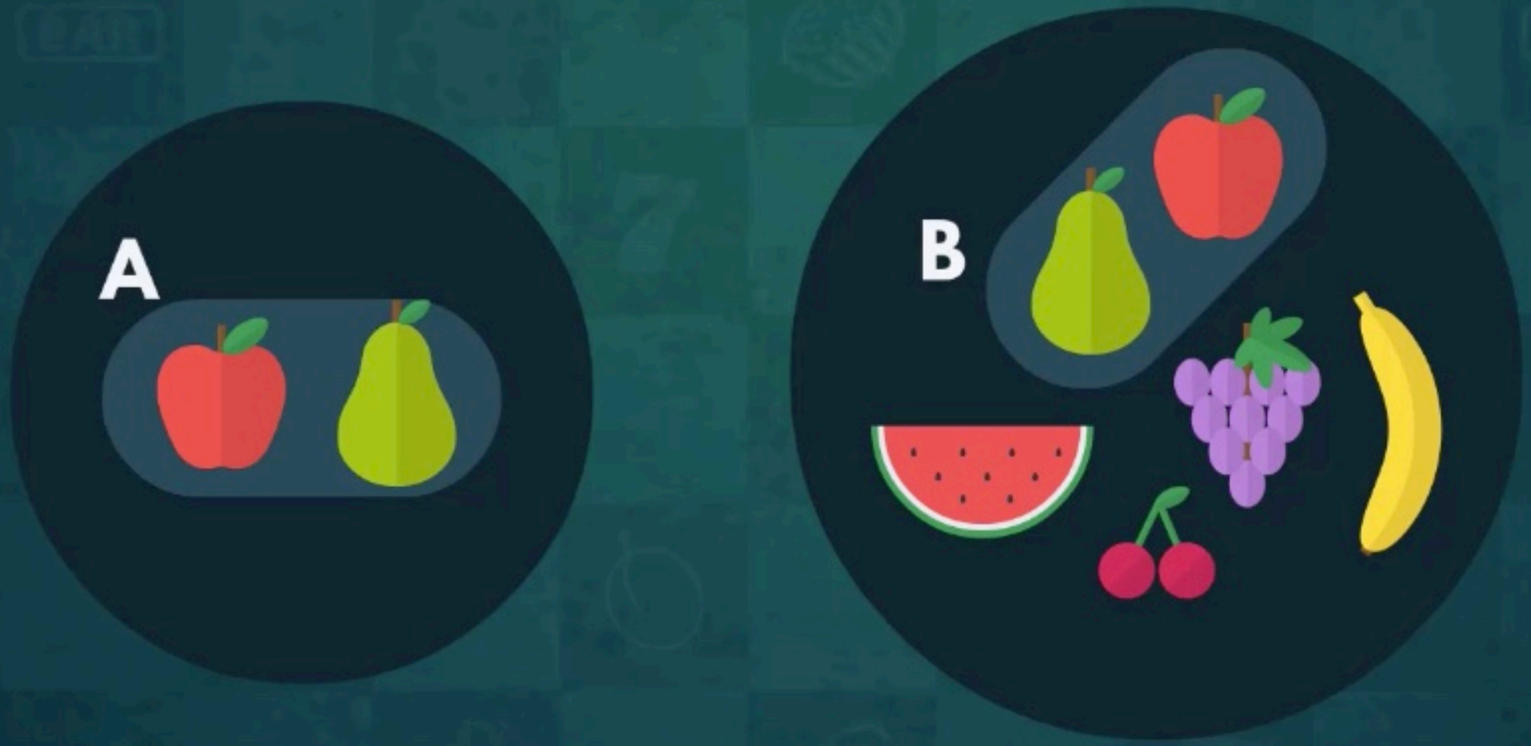
➡ <u>for all x in A, such that, x is even</u>

1.50

Subset

A set that is fully contained in another set

Every element of A is also an element of B ➡ A is a subset of B

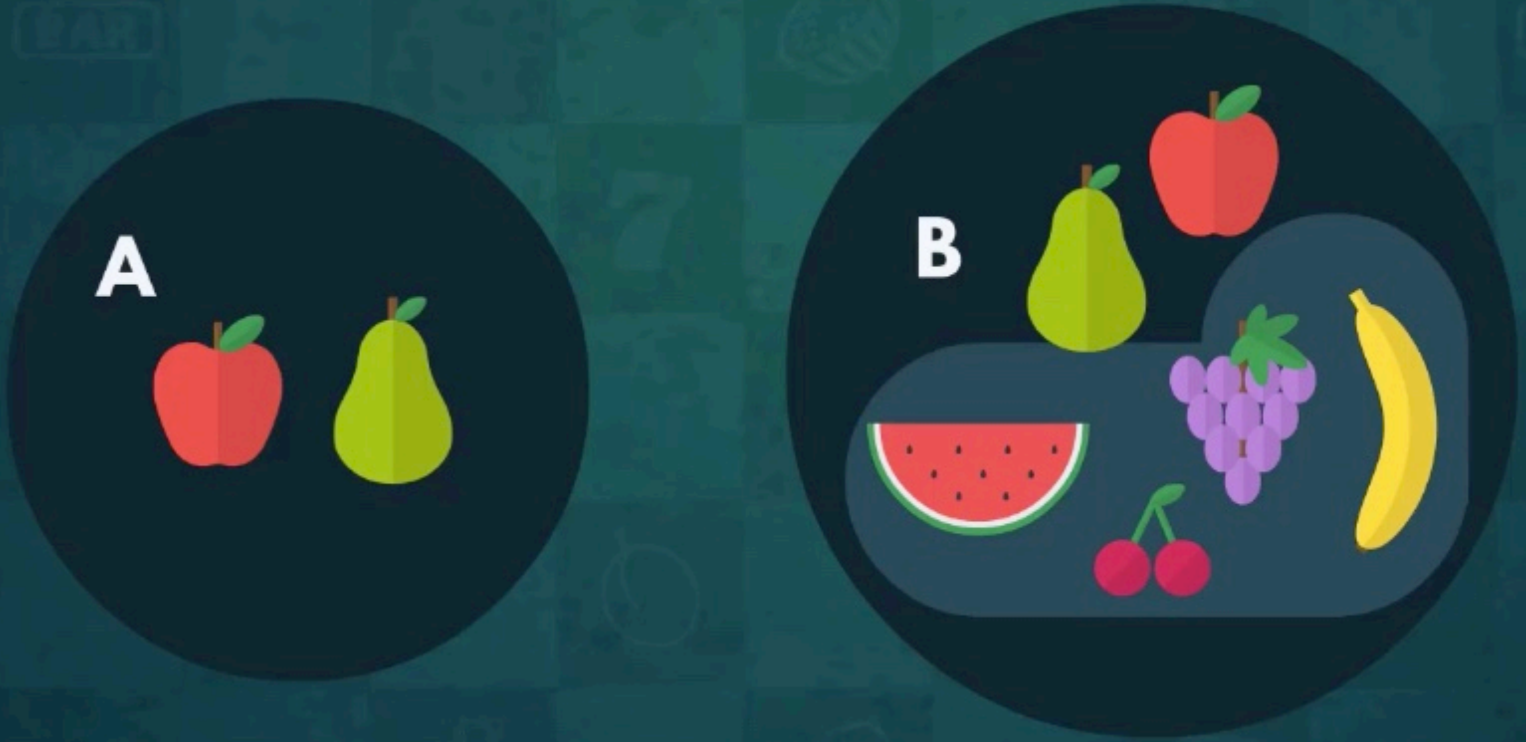


365√DataScience

Subset

A set that is fully contained in another set

Every element of A is also an element of B ➡ $A \subseteq B$



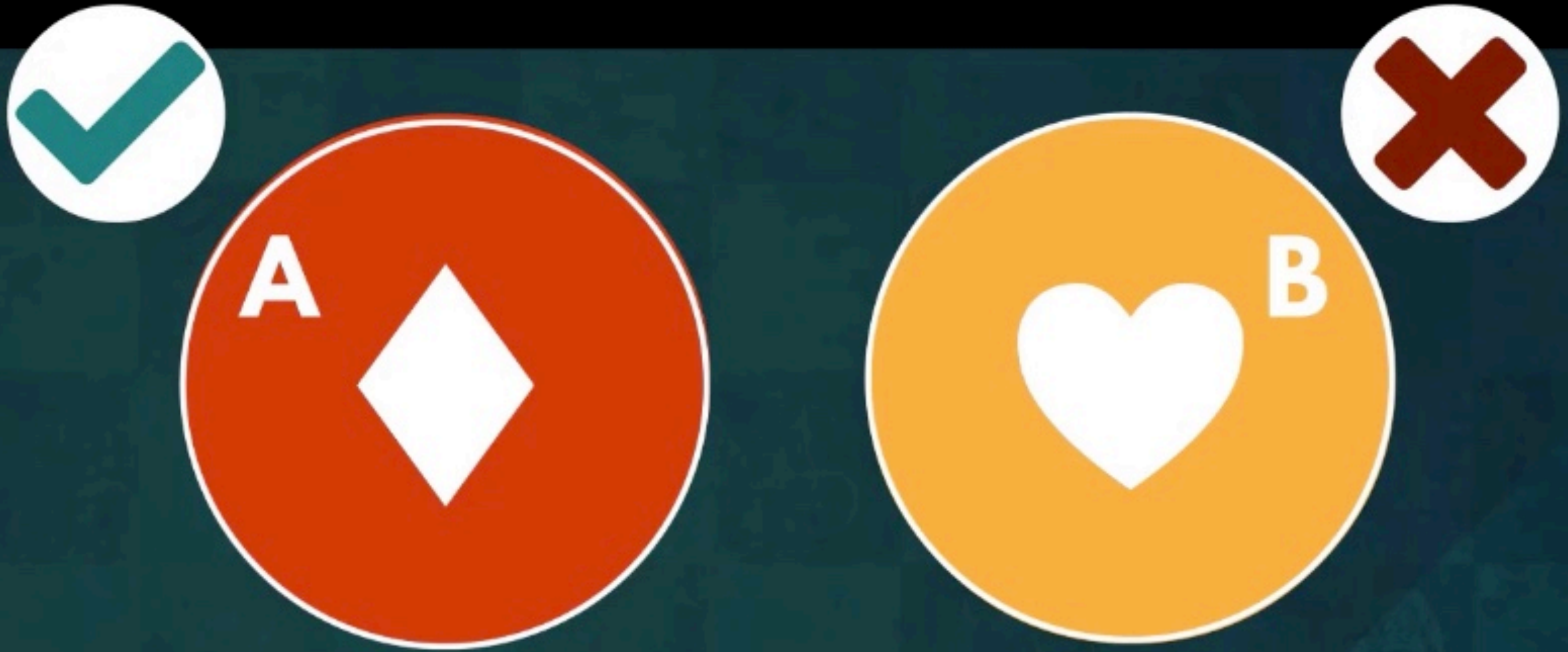
Sets

Values of a set don't always have to be numerical

Event ➡ **being a member of the European Union**

- ✔ France
- ✔ Germany
- ✘ USA
- ✘ Japan

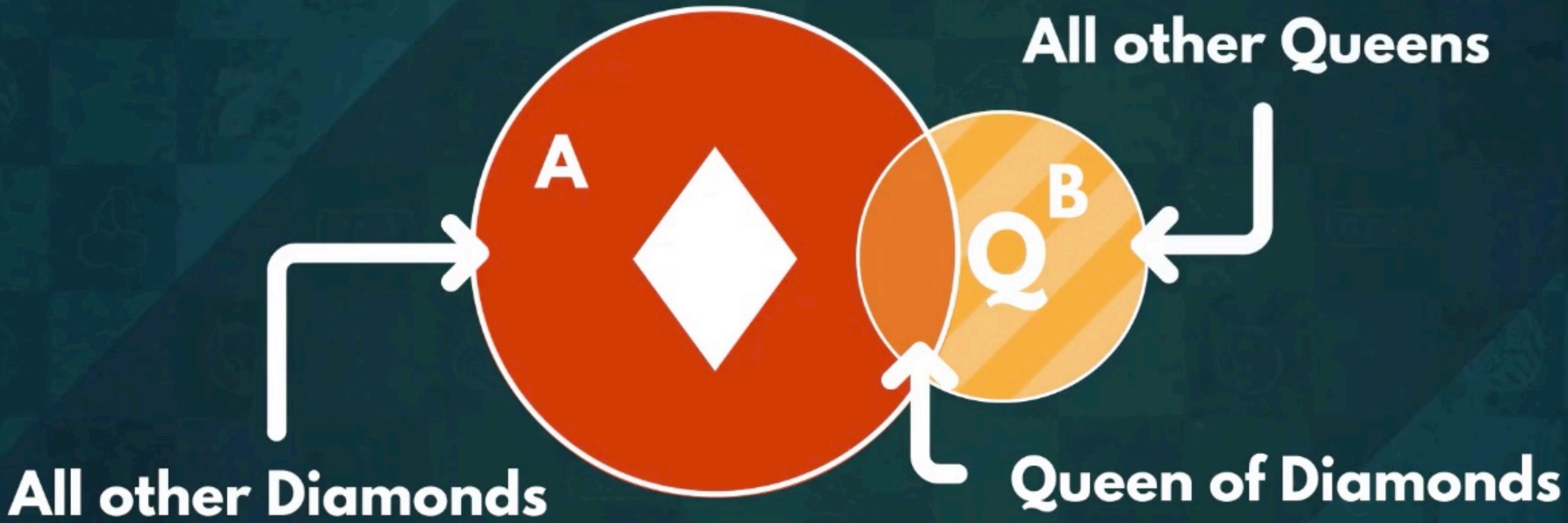
Never Touch



Each card has exactly one suit

1.60

Intersecting



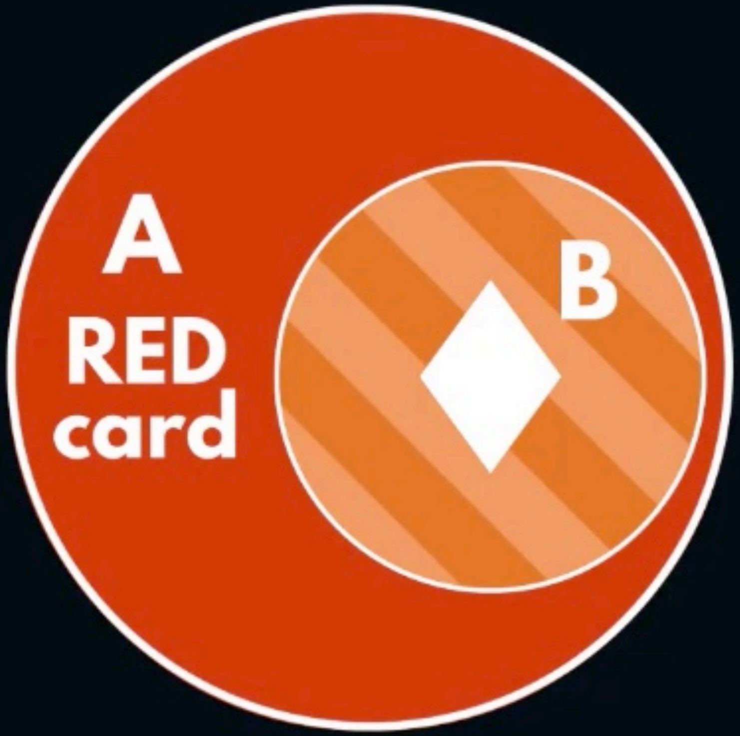
1.60

Subsets

If event A does not occur, then neither does event B

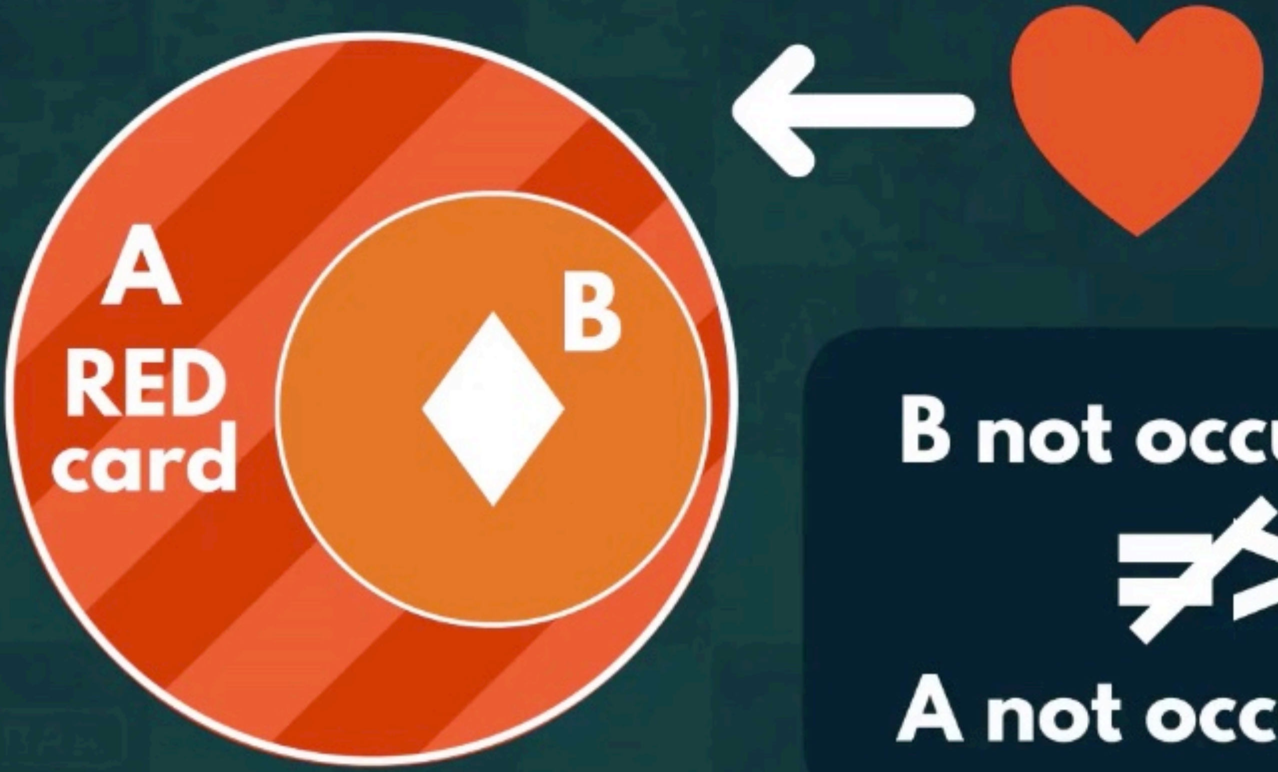
BLACK card

CANNOT be a ♦



We can only ever get a DIAMOND if we get a RED card

Subsets



B not occurring

$\not\Rightarrow$

A not occurring

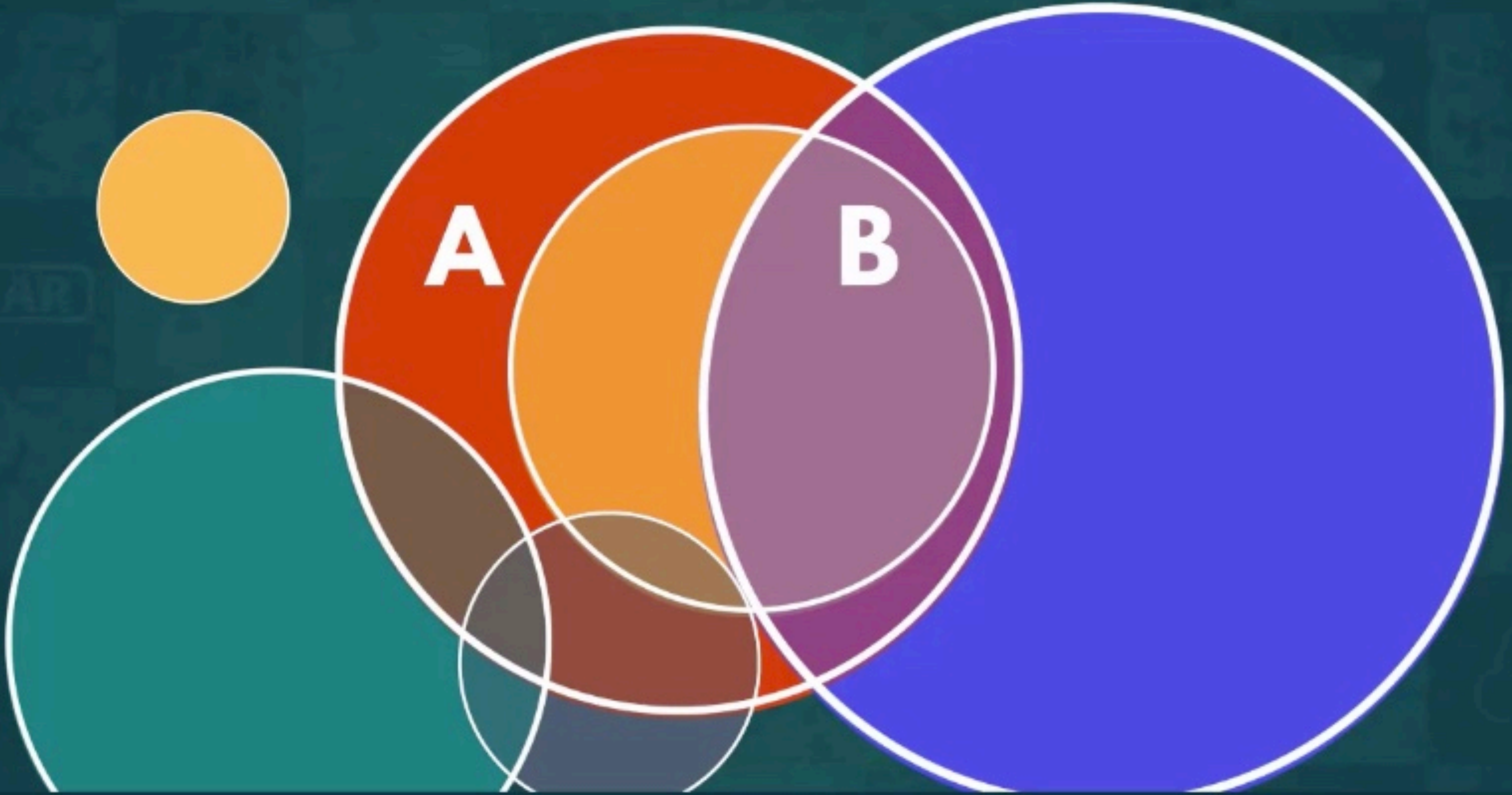
It is possible to get a RED card that ISN'T a DIAMOND

Subsets

Conclusion:

- If an outcome is NOT part of a SET,
  it CANNOT be part of any of its SUBSETS
- An outcome NOT being part of SOME subset,
  does NOT EXCLUDE it from the entirety of the greater set

Multiple sets



Relationships between each two will always be represented by one of the three ways we just went over

Examples of Intersection

$$A \cap B = \emptyset$$

The intersection of all hearts and all diamonds is the EMPTY SET

There are no outcomes which satisfy both events simultaneously

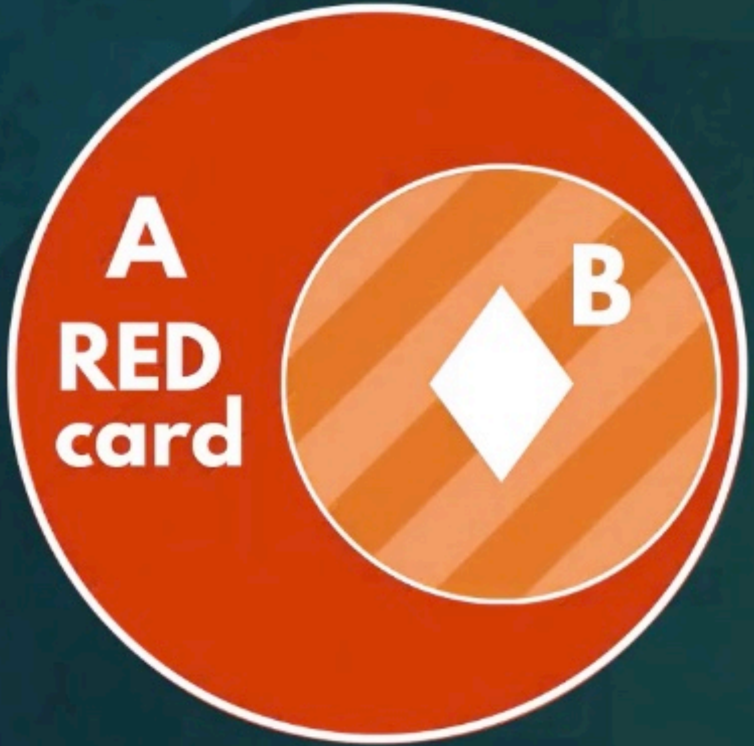
Examples of Intersection

$$A \cap B = Q\blacklozenge$$

Queen of Diamonds

the only one that satisfies being a <u>queen</u> and being a <u>diamond</u> at the <u>same time</u>

Examples of Intersection

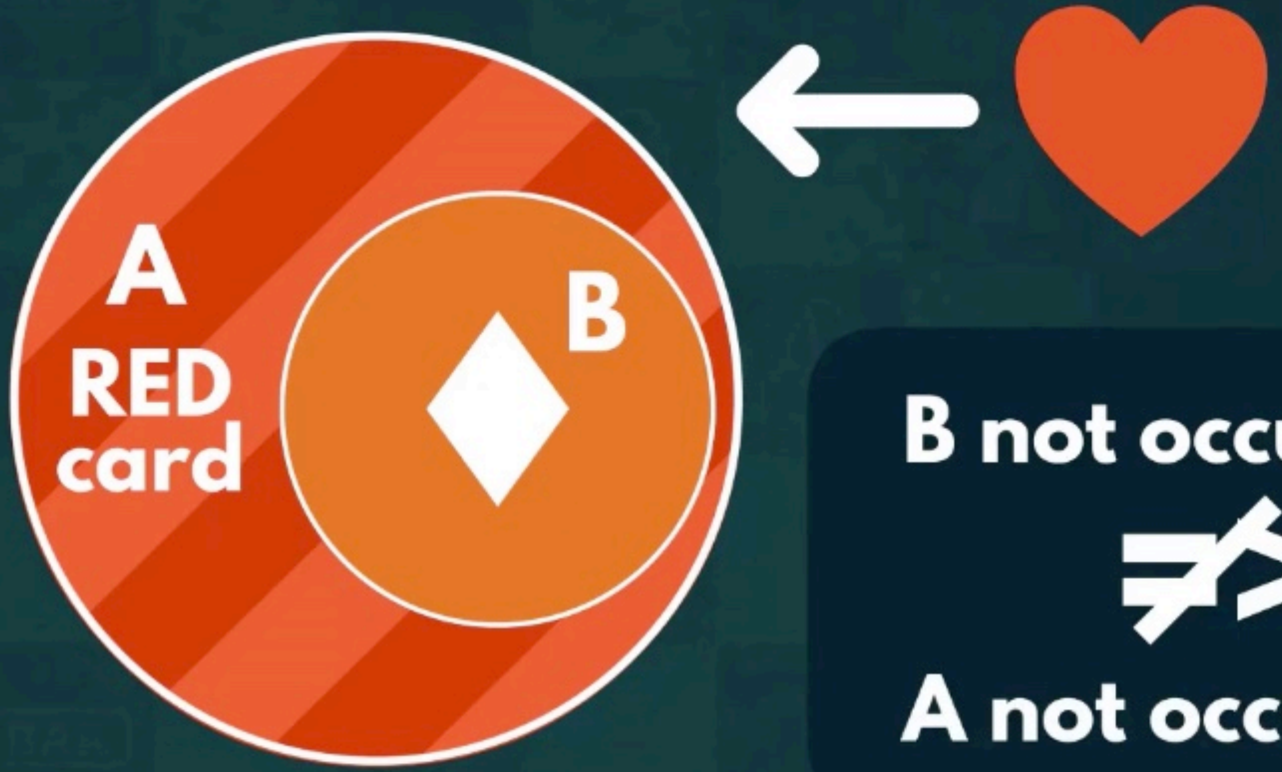


$$A \cap B = B$$

Intersection of the two would simply be "ALL DIAMONDS"

Any diamond is simultaneously RED and a DIAMOND

Subsets



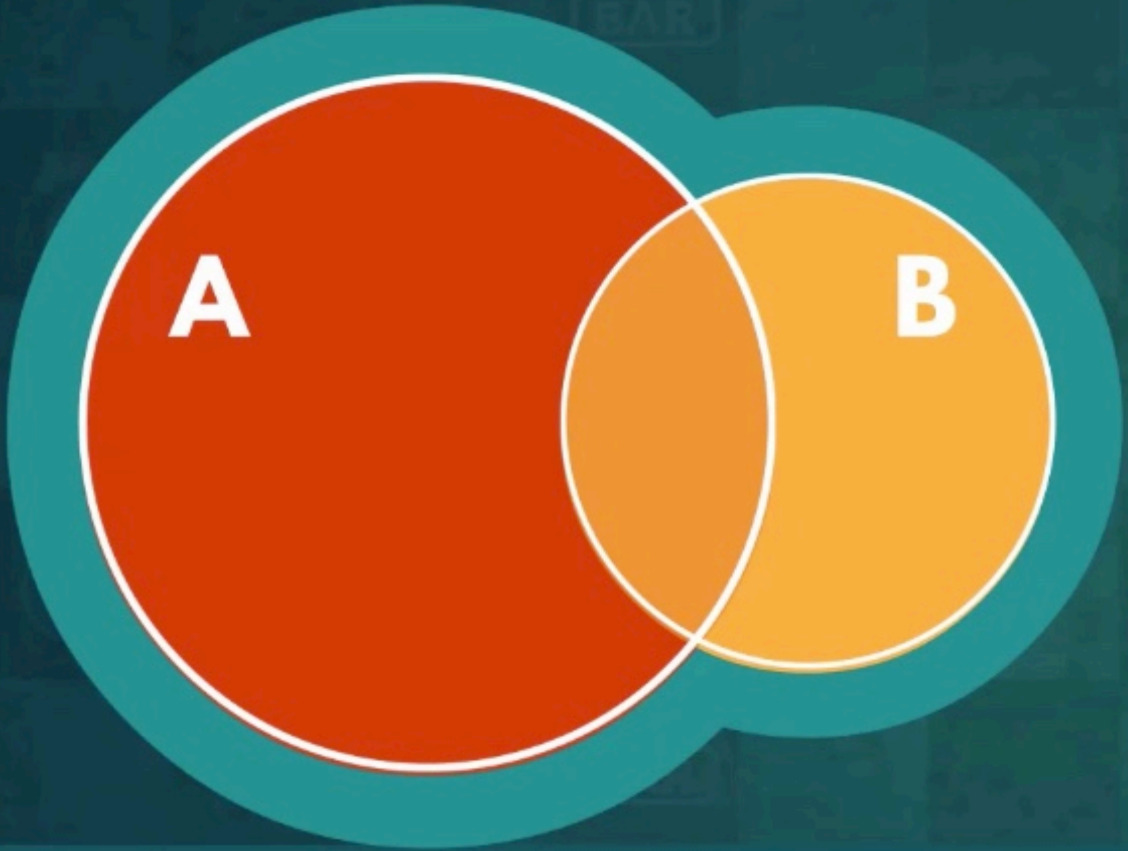
B not occurring

$\not\Rightarrow$

A not occurring

It is possible to get a RED card that ISN'T a DIAMOND

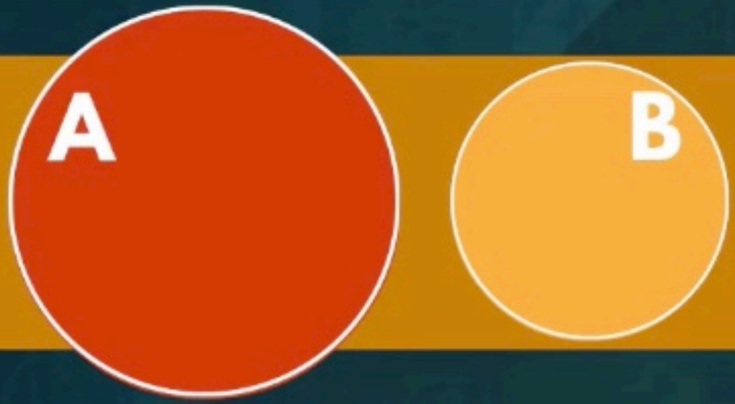
Unions



What if we only require one of them to occur?

It's the same as asking either A or B to happen

Examples of Unions



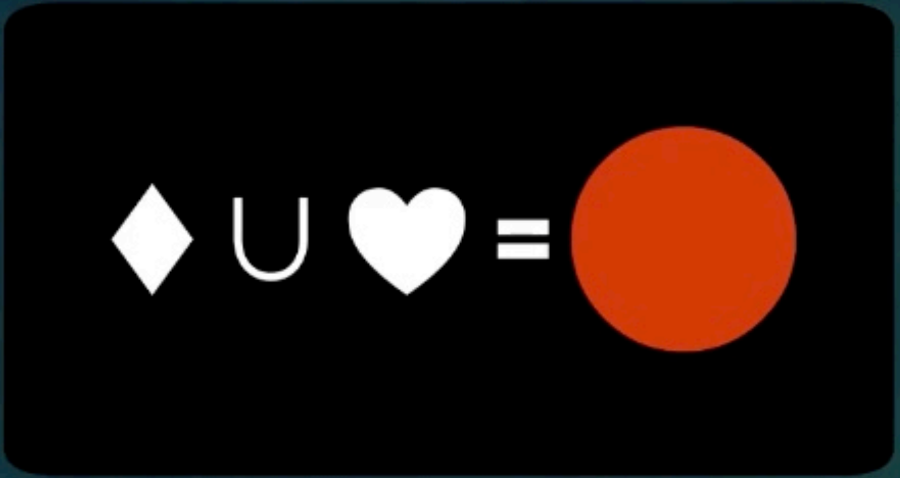
If the sets A and B do not touch at all, their union would simply be their <u>sum</u>

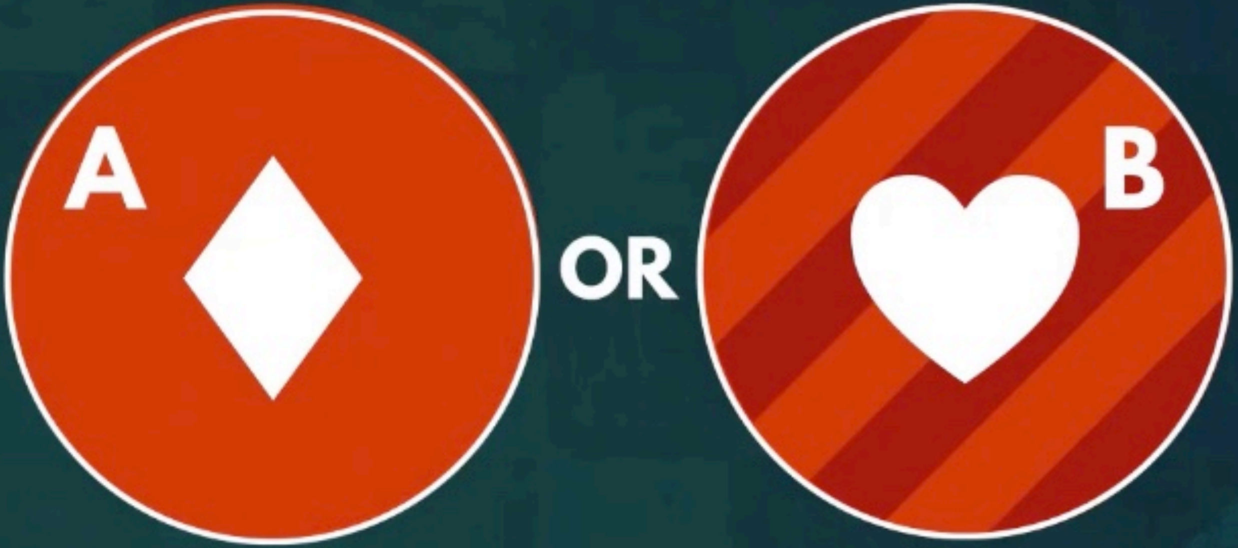
$$A \cap B = \emptyset$$

$$A \cup B = A + B$$

- No element is in both sets simultaneously
- No double-counting

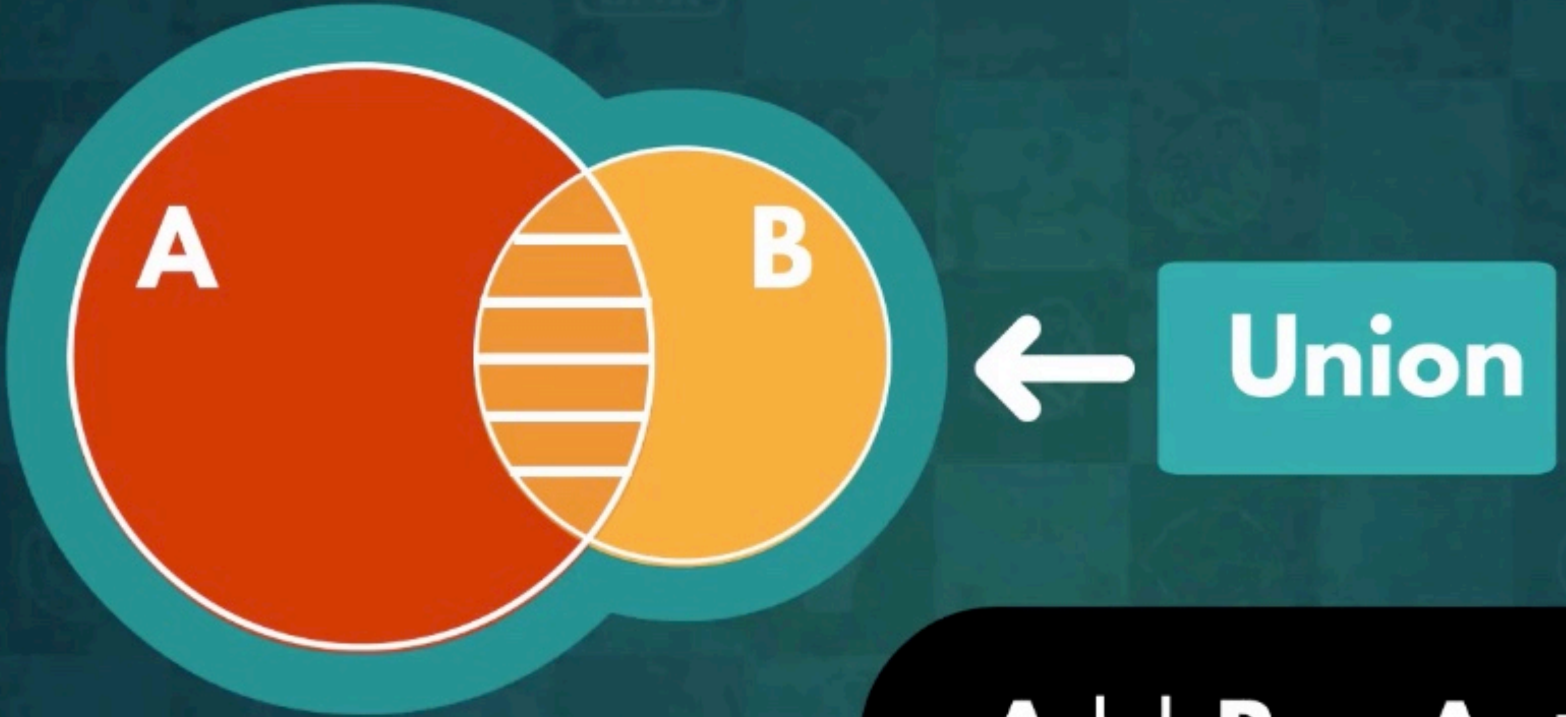
Examples of Unions





- No card can have multiple suits
- No double-counting

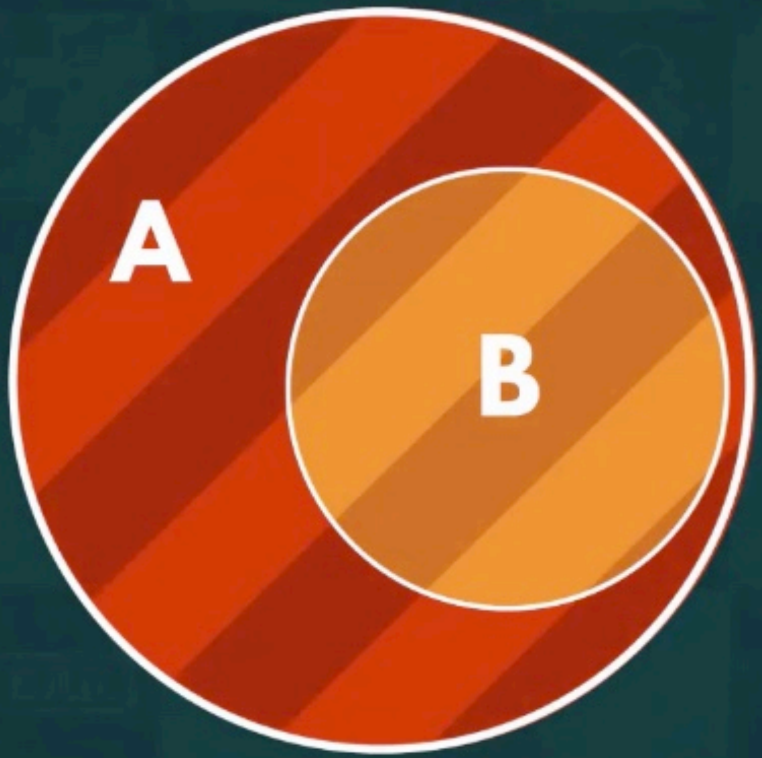
Examples of Unions



$$A \cup B = A + B - \underline{A \cap B}$$

We would be double-counting every element that is a part of the intersection

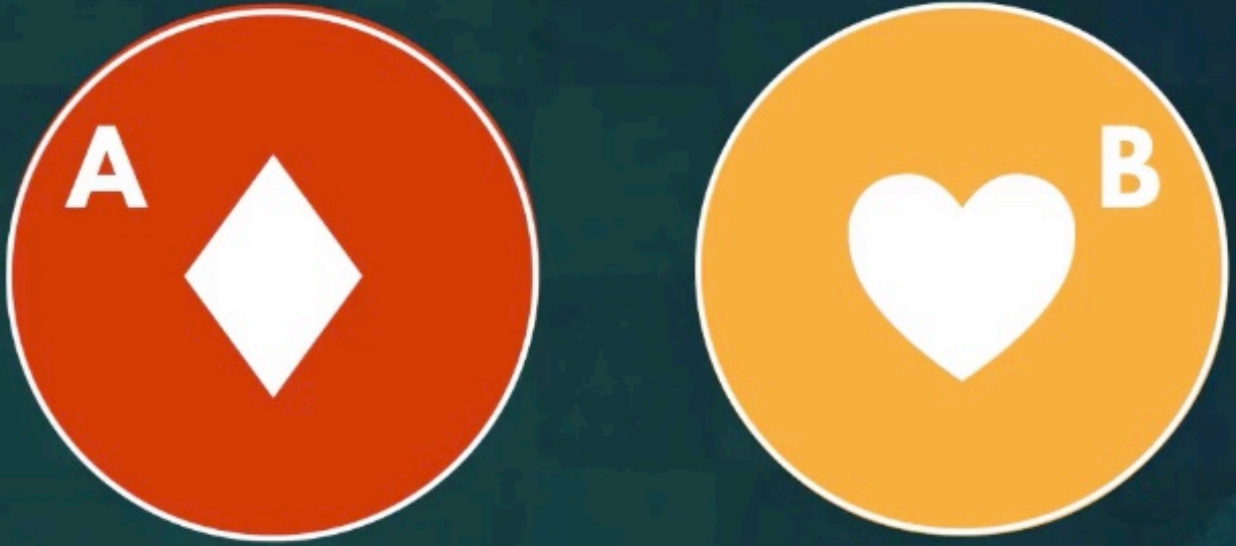
More examples of Unions



What happens if B is a subset of A?

The union would simply be the entire set A

Examples of Intersection



The intersection of all hearts and all diamonds is the EMPTY SET

There are no outcomes which satisfy both events simultaneously

Mutually Exclusive Sets

So far:

Mutually exclusive sets have the empty set as their intersection

Now:

If the intersection of any number of sets is the empty set, then they must be mutually exclusive

What about their union?

If **A** and **B** are mutually exclusive => $A \cup B = A + B$

Complements

- Sets have complements too

- **Complement Set:**

  All values that are part of the sample space, but not part of the set

1.60

Complements ≠ Mutually Exclusive

! Complements are ALWAYS mutually exclusive

! NOT all mutually exclusive sets are complements

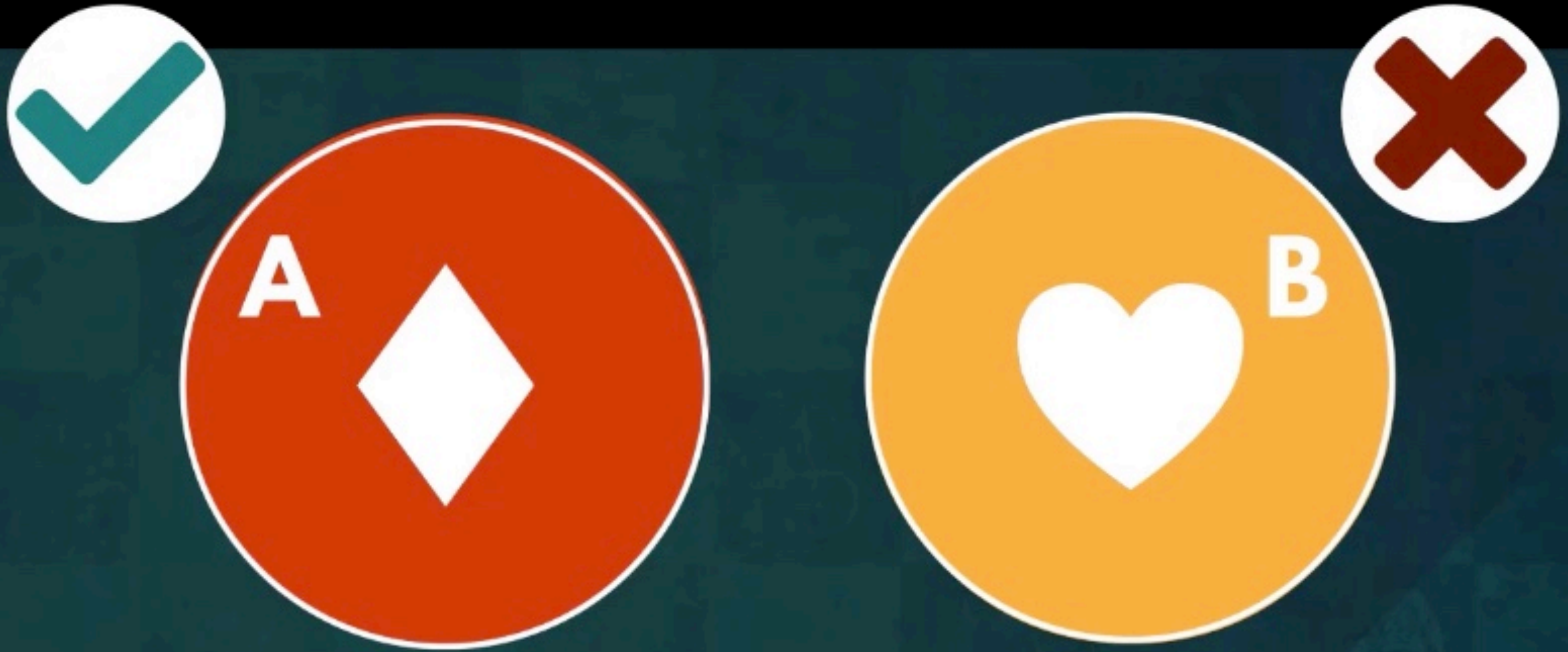
Complements $\neq$ Mutually Exclusive

Even | Ending in 5

$A \Rightarrow$ All even
$A' \Rightarrow$ All odd

$\Rightarrow A' \neq B \Rightarrow 13 \in A'$, $13 \notin B$

Never Touch



Each card has exactly one suit

Independent Events

The theoretical probability remains unaffected by other events

You always have a 50% chance of getting tails

Dependent Events

Probabilities of dependent events vary as conditions change

$$P(Q\spadesuit) = \frac{1}{52}$$

favourable outcome

elements in the sample space

Queen of Spades Example 1

Imagine we know that the card we drew was a spade

$P(Q♠)$ ↑

New sample space contains the 13 cards

$$\Rightarrow P(Q♠) = \frac{1}{13}$$

Queen of Spades Example 2

Imagine we know our card is a queen

P(Q♠) ↑

New sample space only consists of 4 cards

$$\Rightarrow P(Q♠) = \frac{1}{4}$$

1.50

Queen of Spades

Normally → $P(Q\spadesuit) = \dfrac{1}{52}$

Example 1 $\spadesuit$ → $P(Q\spadesuit) = \dfrac{1}{13}$

Example 2 Q → $P(Q\spadesuit) = \dfrac{1}{4}$

The probability of an event changes depending on the information we have

Notation

Two events: A and B

The probability of getting A, if we are given that B has occurred → $P(A|B)$ "A given B"

Queen of Spades

A → Q♠

B → ♠

$$P(A|B) = \frac{1}{13}$$

→ the probability of drawing the Queen of Spades if we know the card is a <u>spade</u>

Queen of Spades

$A \rightarrow$ Q♠

$C \rightarrow$ Q

$$P(A|C) = \frac{1}{4}$$ $\longrightarrow$ the likelihood of getting the Queen of Spades, assuming we drew a <u>queen</u>

Conditional Probability

$P(A|C) \rightarrow$ We use it to distinguish dependent from independent events

Conditional Probability

$P(A|C) \rightarrow$ We use it to distinguish dependent from independent events

Two Coin Flips

A → 

B →  on the <u>previous</u> flip

$$P(A|B) = 0.5$$

Two Coin Flips

$$P(A) = P(A|B)$$

independent

If any two events are independent

$$P(A \cap B) = P(A) \times P(B)$$

Queen of Spades

A → Q♠ B → ♠ C → Q

$$P(A) = \frac{1}{52} \neq P(A|B) = \frac{1}{13}$$

=> A and B are dependent

Notation

A → Q♠ B → ♠ C → Q

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

! If $P(B) > 0$

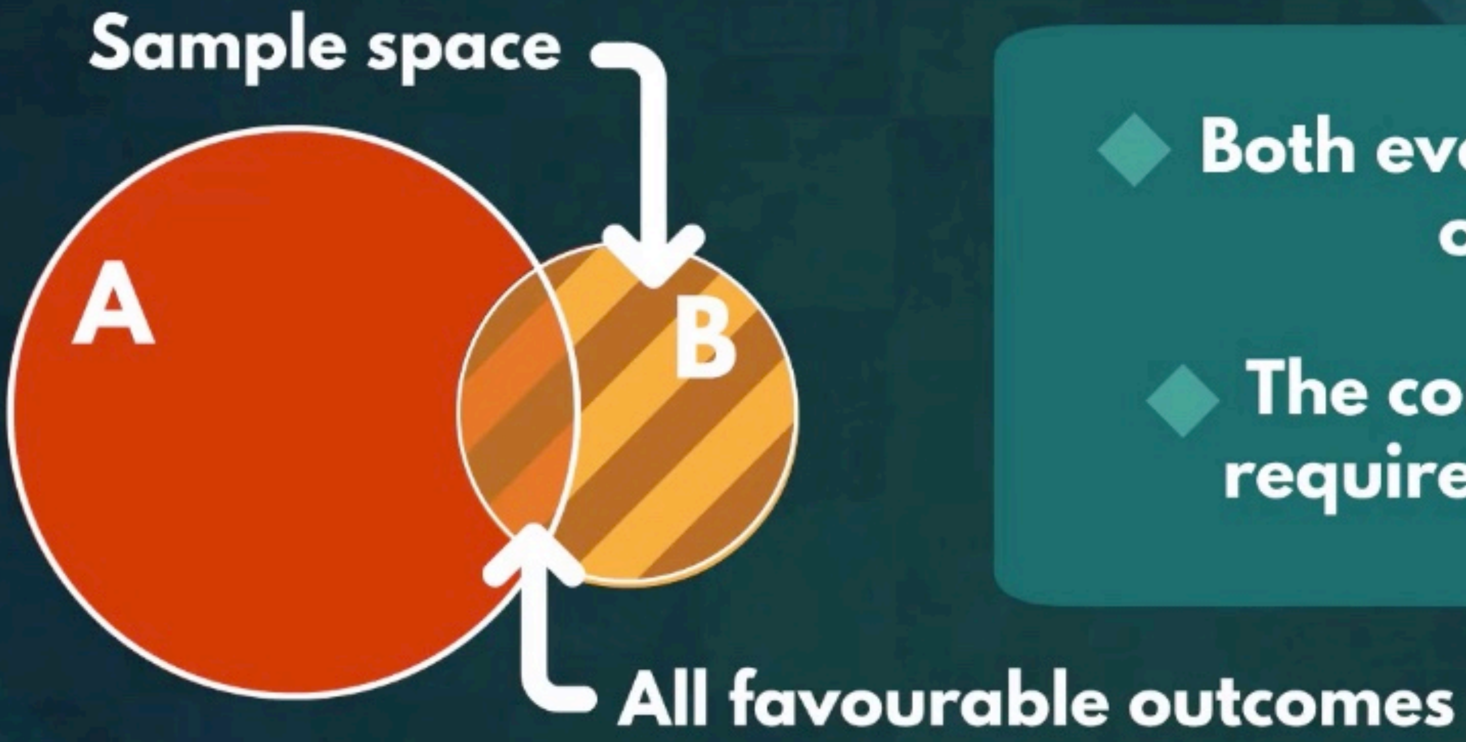
Notation

A → Q♠ B → ♠ C → Q

If $P(B) = 0$ ➡ Event B would <u>never</u> occur

$A|B$ ➡ <u>Not</u> interpretable

Notation



- Both events B and A need to occur simultaneously
- The conditional probability requires that event B occurs

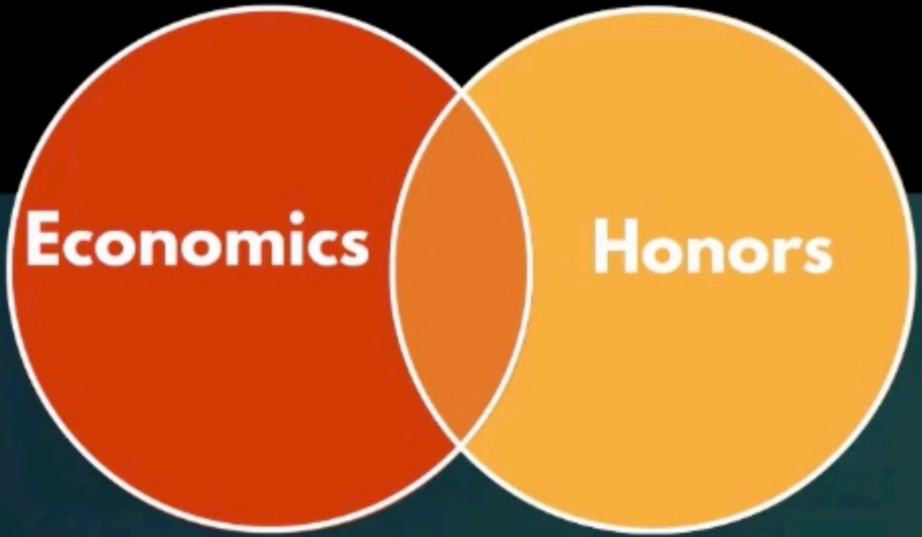
$$P(A|B) = \frac{P(A \cap B)}{P(B)} \approx P(A) = \frac{\text{favourable}}{\text{all}}$$

Importance

the order in which we write the elements is crucial

$$\underbrace{P(A|B) \qquad P(B|A)}_{\text{NOT the same}}$$

Hamilton College's Class 2018



$P(H \mid E) = 5\%$

$P(E \mid H) = 5\%$

Equal numerically

Completely different meanings

Hamilton College's Class 2018

only 4 of the 80 Economics majors, graduated with distinction

$$P(H \mid E) = \frac{4}{80}$$

4 out of the 80 students who graduated with high grades, completed a degree in Economics

$$P(E \mid H) = \frac{4}{80}$$

Vegetarian Survey

100 men and women are asked if they eat meat

			total
♀	15	32	47
♂	29	24	53
total	44	56	100

Vegetarian Survey

A → 

B → ♀

$$P(A|B) \neq P(B|A)$$

Different events

$$P(A|B) = \frac{15}{47}$$

the likelihood of a woman being vegetarian

$$P(B|A) = \frac{15}{44}$$

the likelihood of a vegetarian being woman

Vegetarian Survey

=> It is more likely for a vegetarian to be female, than for a woman NOT to eat meat

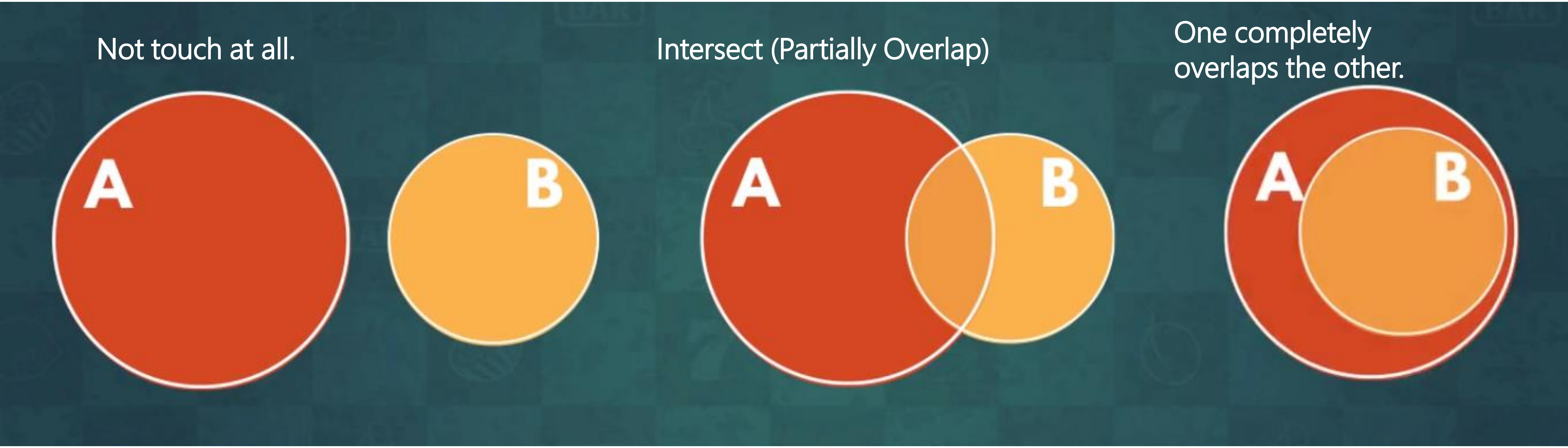
$$P(A|B) = \frac{15}{47} < P(B|A) = \frac{15}{44}$$

the likelihood of a woman being vegetarian

the likelihood of a vegetarian being woman

Multiple Events

The sets of outcomes that satisfy two events A and B can interact in one of the following **3** ways.



Examples:

A -> Diamonds
B -> Hearts

The two events can never happen simultaneously

Diamonds
Queens

The two events can occur at the same time

Red Cards
Diamond

One event can only ever occur if the other one does as well

Law of Total Probability

- $A = B_1 \cup B_2 \cup \ldots \cup B_n$

$$P(A) = P(A|B_1) \times P(B_1) + P(A|B_2) \times P(B_2) \ldots$$

Vegetarian Survey

$$P(\ ) = P(\ \mid ♂) \times P(♂) + P(\ \mid ♀) \times P(♀) =$$

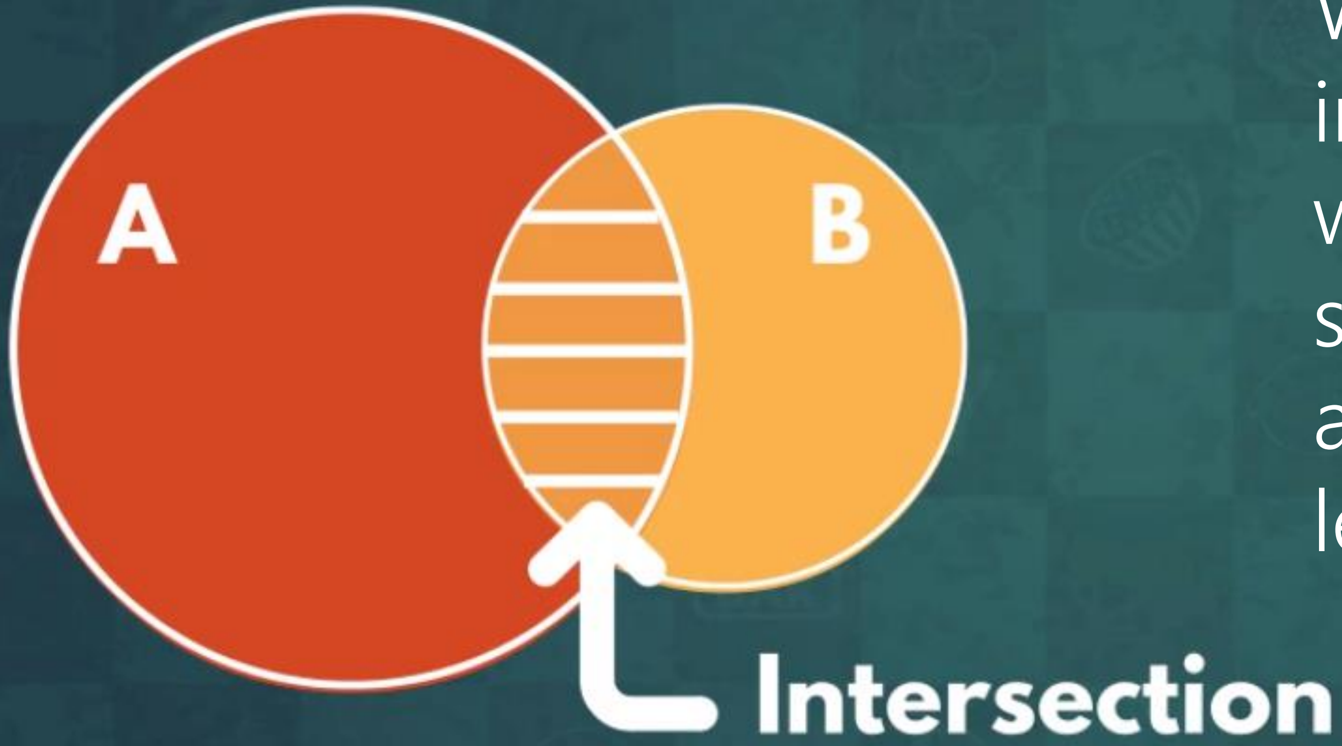
$$= \frac{29}{53} \times \frac{53}{100} + \frac{15}{47} \times \frac{47}{100} = 0.44$$

			total
♀	15	32	47
♂	29	24	53
total	44	56	100

- There is a 44% chance of someone being vegetarian

Intersection

The **intersection** of two or more events expresses the set of outcomes that satisfy all the events simultaneously. Graphically, this is the area where the sets intersect.



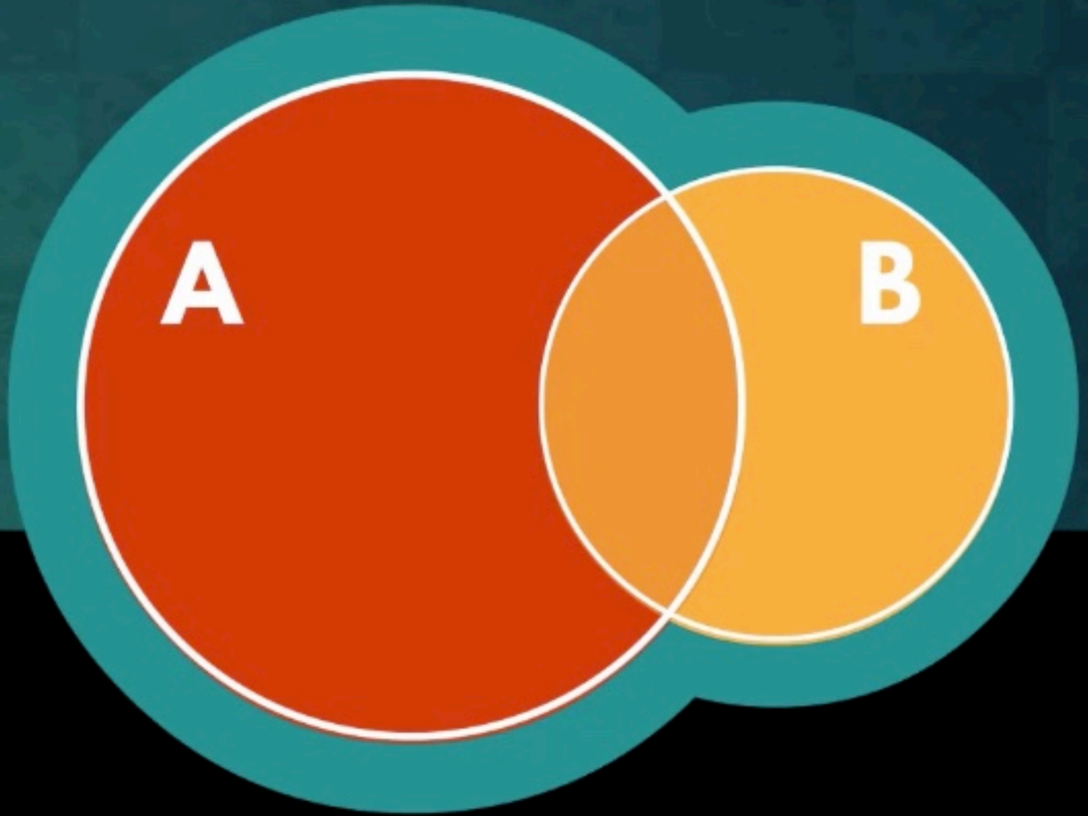
We denote the intersection of two sets with the "intersect" sign, which resembles an upside-down capital letter U:

$$\boldsymbol{A} \cap \boldsymbol{B}$$

Additive Law

Recall:

$$A \cup B = A + B - A \cap B$$



Additive Law

Definition:

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

The probability of the union of two sets
is equal to the sum of the individual probabilities of each event,
minus the probability of their intersection

Vegetarian Survey

$$P(♀ \cup \;) = P(♀) + P(\;) - P(♀ \cap \;) =$$

$$= 0.47 + 0.44 - 0.15 = 0.76$$

- There is a 76% chance that a random person from the survey is either female, vegetarian or both

			total
♀	15	32	47
♂	29	24	53
total	44	56	100

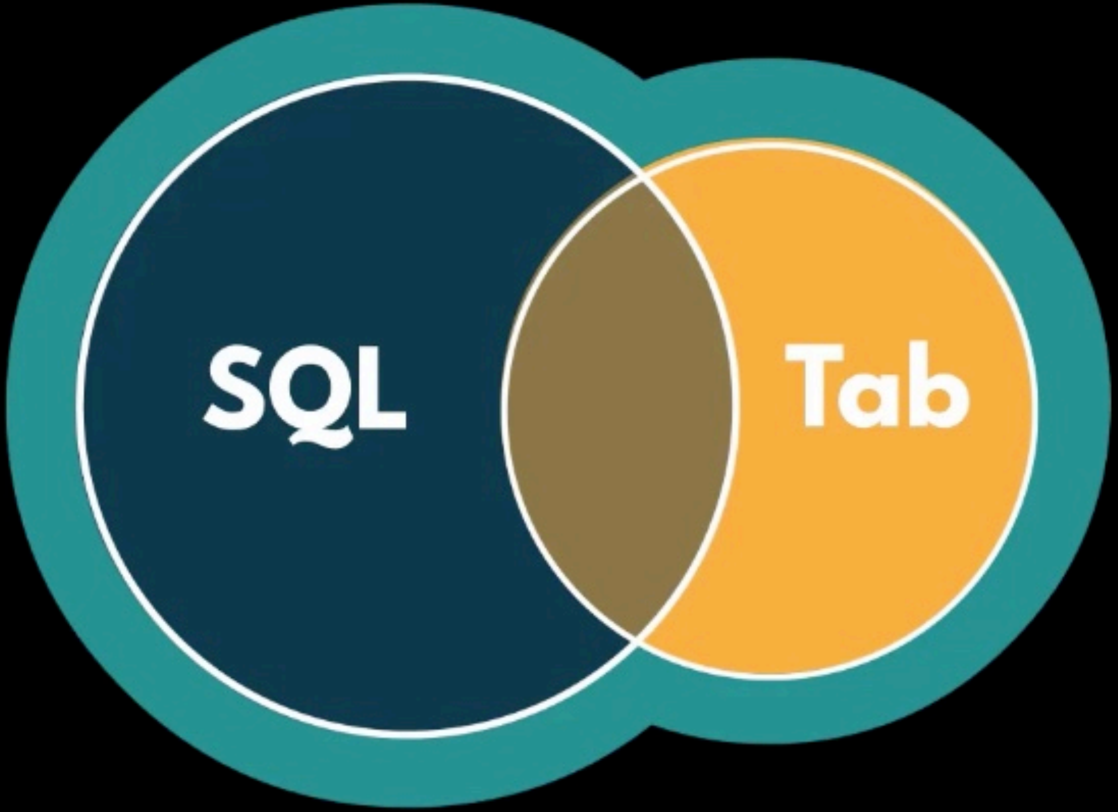
SQL and Tableau

$P(\quad) = 38\%$

$P(\quad) = 45\%$

$P(\quad \cup \quad) = 66\%$

$P(\quad \cap \quad) = ?$



SQL and Tableau

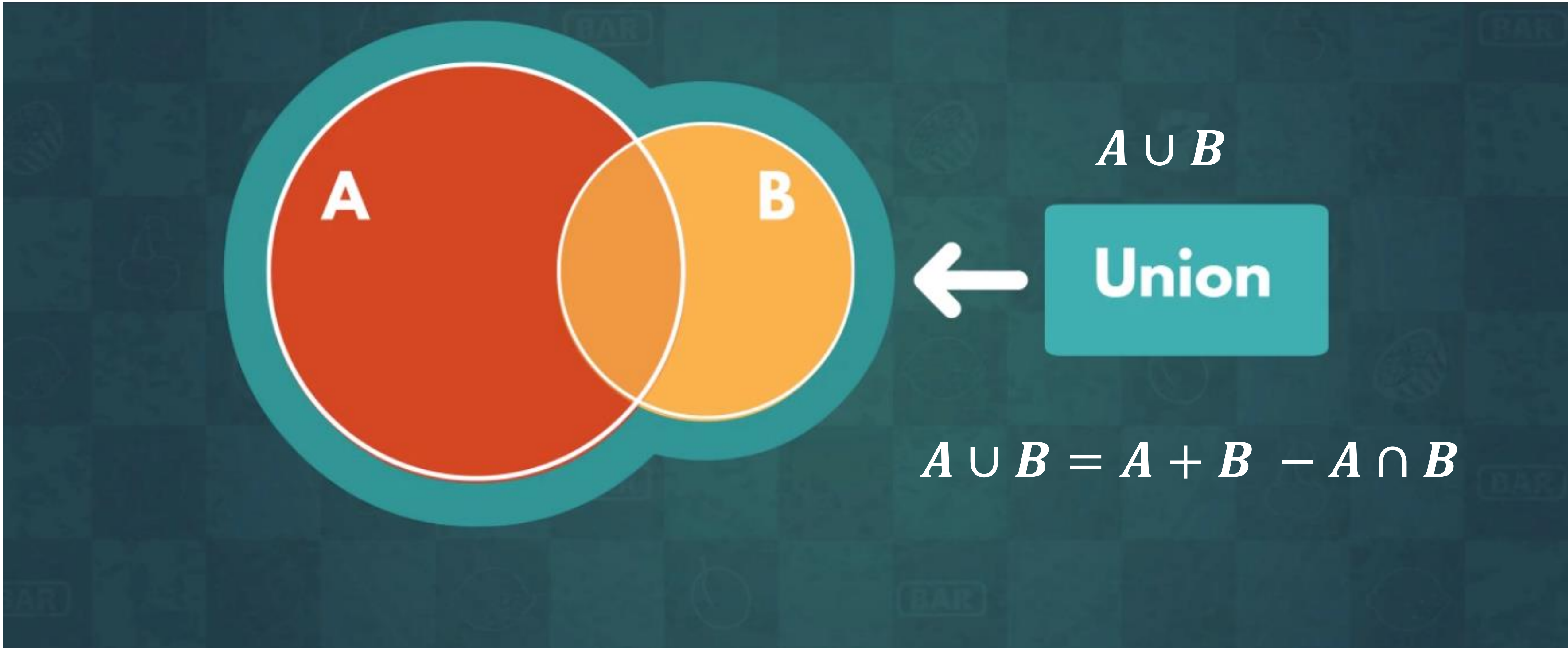
$$P(\cap) = P( ) + P( ) - P(\cup) =$$

$$= 38\% + 45\% - 66\% = 0.17$$

- A likelihood of 0.17 for somebody in the office to be able to proficiently implement SQL and Tableau

Union

The **union** of two or more events expresses the set of outcomes that satisfy at least one of the events. Graphically, this is the area that includes both sets.

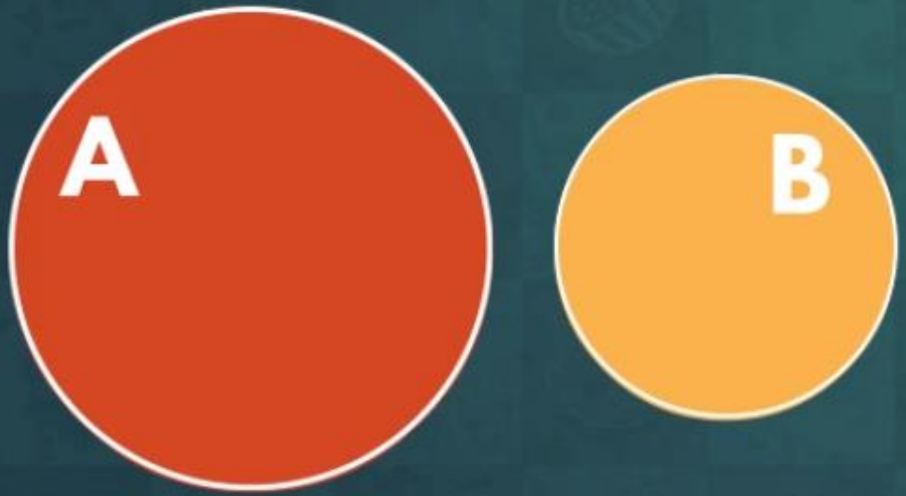


$$A \cup B$$

$$A \cup B = A + B \; - A \cap B$$

Mutually Exclusive Sets

Sets with no overlapping elements are called **mutually exclusive**. Graphically, their circles never touch.



If $A \cap B = \emptyset$, then the two sets are mutually exclusive.

Remember:
All complements are mutually exclusive, but not all mutually exclusive sets are complements.

Example:
Dogs and Cats are mutually exclusive sets, since no species is simultaneously a feline and a canine, but the two are not complements, since there exist other types of animals as well.

Independent and Dependent Events

If the likelihood of event A occurring (P(A)) is affected event B occurring, then we say that A and B are **dependent** events.

Alternatively, if it isn't – the two events are **independent**.

We express the probability of event A occurring, given event B has occurred the following way $\boldsymbol{P(A|B)}$.
We call this the conditional probability.

Independent:

- All the probabilities we have examined so far.
- The outcome of A does not depend on the outcome of B.
- $P(A|B) = P(A)$

Example

- A -> Hearts
- B -> Jacks

Dependent

- New concept.
- The outcome of A depends on the outcome of B.
- $P(A|B) \neq P(A)$

Example

- A -> Hearts
- B -> Red

Conditional Probability

For any two events A and B, such that the likelihood of B occurring is greater than 0 ($P(B) > 0$), the conditional probability formula states the following.

Probability of A, given B has occurred

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

Probability of the intersection.

Probability of event B

Intuition behind the formula:

- Only interested in the outcomes where B is satisfied.
- Only the elements in the intersection would satisfy A as well.
- Parallel to the "favoured over all" formula:
    - Intersection = "preferred outcomes"
    - B = "sample space"

Remember:

- Unlike the union or the intersection, changing the order of A and B in the conditional probability alters its meaning.
- $P(A|B)$ is not the same as $P(B|A)$, even if $P(A|B) = P(B|A)$ numerically.
- The two conditional probabilities possess **different meanings** even if they have equal values.

Law of total probability

The **law of total probability** dictates that for any set A, which is a union of many mutually exclusive sets $B_1, B_2, \ldots, B_n$, its probability equals the following sum.

$$P(A) = P(A|B_1) \times P(B_1) + P(A|B_2) \times P(B_2) + \cdots + P(A|B_n) \times P(B_n)$$

Conditional Probability of A, given B_1 has occurred.

Probability of B_1 occurring.

Conditional Probability of A, given B_2 has occurred.

Probability of B_2 occurring.

Intuition behind the formula:

- Since P(A) is the union of mutually exclusive sets, so it equals the **sum of the individual sets**.
- The **intersection** of a union and one of its subsets is the entire subset.
- We can rewrite the conditional probability formula $\boldsymbol{P(A|B) = \frac{P(A \cap B)}{P(B)}}$ to get $\boldsymbol{P(A \cap B) = P(A|B) \times P(B)}$.
- Another way to express the law of total probability is:
  - $P(A) = P(A \cap B_1) + P(A \cap B_2) + \cdots + P(A \cap B_n)$

Additive Law

The additive law calculates the probability of the union based on the probability of the individual sets it accounts for.

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

Probability of the union

Probability of the intersection

Intuition behind the formula

- Recall the formula for finding the size of the Union using the size of the Intersection:
  - $A \cup B = A + B - A \cap B$
- The probability of each one is simply its size over the size of the sample space.
- This holds true for any events A and B.

The Multiplication Rule

The multiplication rule calculates the probability of the intersection based on the conditional probability.

$$P(A \cap B) = P(A|B) \times P(B)$$

Probability of the Intersection

Conditional Probability

Probability of event B

Intuition behind the formula

- We can multiply both sides of the conditional probability formula $P(A|B) = \frac{P(A \cap B)}{P(B)}$ by $P(B)$ to get $P(A \cap B) = P(A|B) \times P(B)$.
- If event B occurs in 40% of the time ($P(B) = 0.4$) and event A occurs in 50% of the time B occurs ($P(A|B) = 0.5$), then they would simultaneously occur in 20% of the time ($\mathrm{P}(\mathrm{A}|B) \times P(B) = 0.5 \times 0.4 = 0.2$).

Multiplication Rule

$$P(A|B) \times P(B) = \frac{P(A \cap B) \times \cancel{P(B)}}{\cancel{P(B)}}$$

365√DataScience

Bayes' Law

Bayes' Law helps us understand the relationship between two events by computing the different conditional probabilities.

We also call it Bayes' Rule or Bayes' Theorem.

Conditional probability of B, given A.

Conditional probability of A, given B.

$$P(A|B) = \frac{P(B|A) \times P(A)}{P(B)}$$

Intuition behind the formula

- According to the multiplication rule $P(A \cap B) = P(A|B) \times P(B)$, so $P(B \cap A) = P(B|A) \times P(A)$.
- Since $P(A \cap B) = P(B \cap A)$, we plug in $P(B|A) \times P(A)$ for $P(A \cap B)$ in the conditional probability formula $P(A|B) = \frac{P(A \cap B)}{P(B)}$
- Bayes' Law is often used in medical or business analysis to determine which of two symptoms affects the other one more.