Calculating probabilities of two events

FOUNDATIONS OF PROBABILITY IN PYTHON

Alexander A. Ramírez M.
CEO @ Synergy Vision

Independence

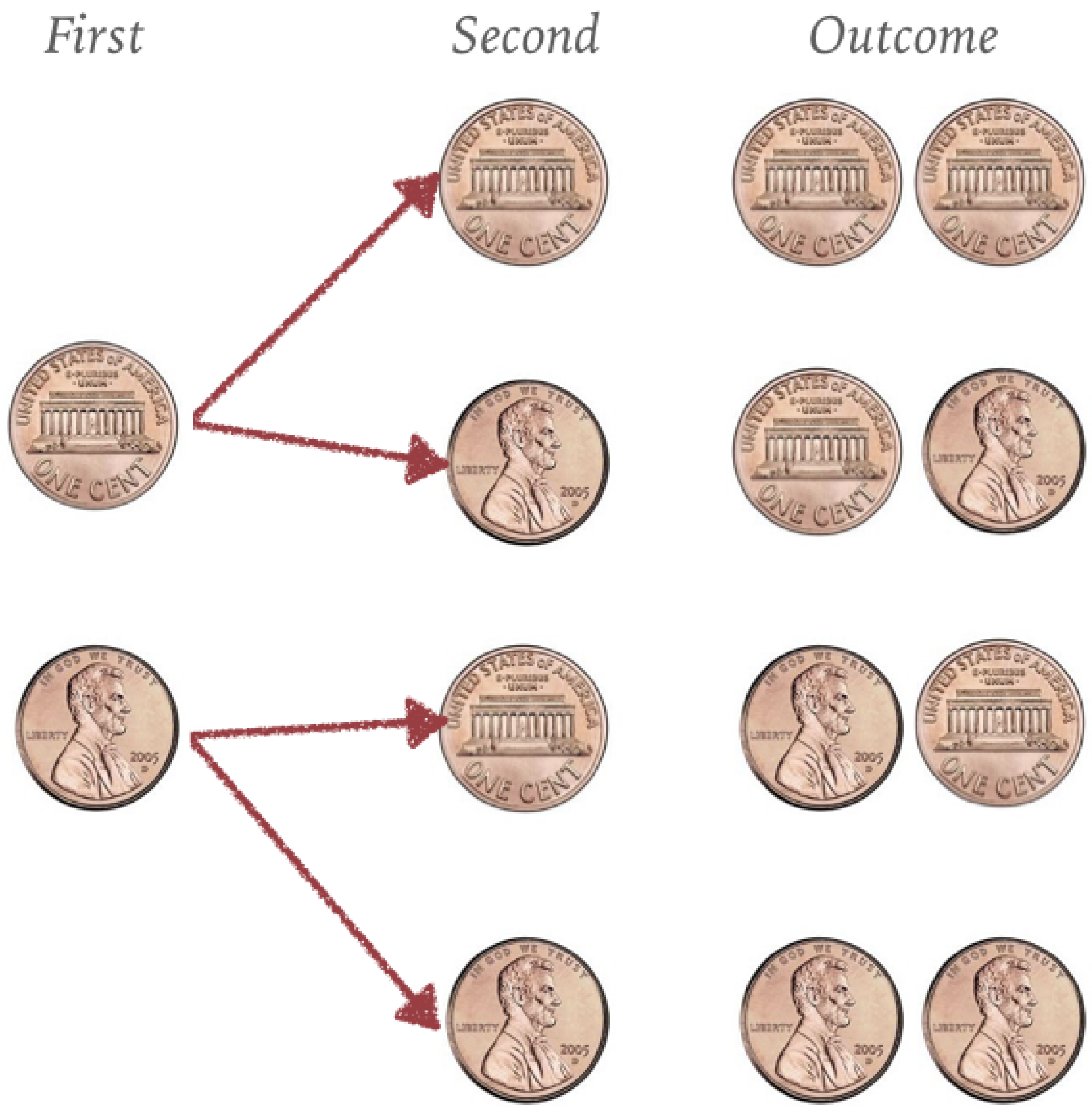
Given that A and B are events in a random experiment, the conditions for independence of A and B are:

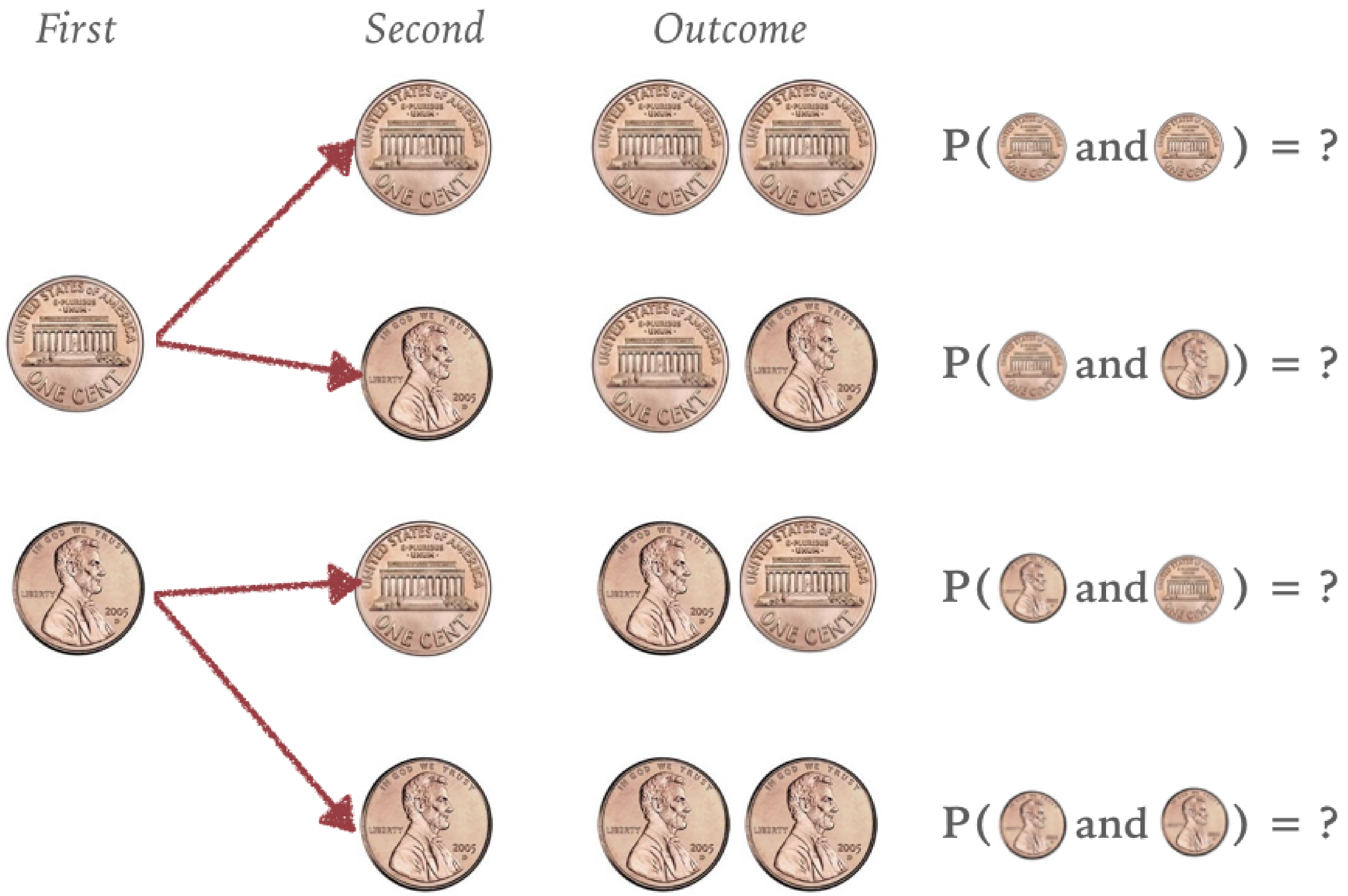
1. The order in which A and B occur does not affect their probabilities.
2. If A occurs, this does not affect the probability of B.
3. If B occurs, this does not affect the probability of A.

Fair coin

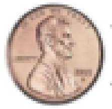
P() = 0.5

P() = 0.5

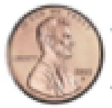

First
Second
Outcome


First
Second
Outcome
P(and) = ?
P(and) = ?
P(and) = ?
P(and) = ?

P( and ) = ?

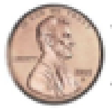
P( and ) = ?

P( and ) = ?

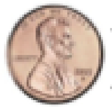
P( and ) = ?

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

P( and ) = ?

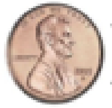
P( and ) = ?

P( and ) = ?

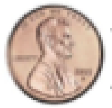
P( and ) = ?

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

P( and ) = P() P()

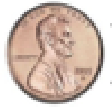
P( and ) = P() P()

P( and ) = P() P()

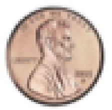
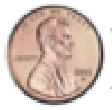
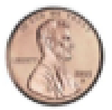
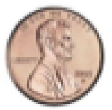
P( and ) = P() P()

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

P( and ) = P() P() = (0.5) (0.5) = 0.25

P( and ) = P() P() = (0.5) (0.5) = 0.25

P( and ) = P() P() = (0.5) (0.5) = 0.25

P( and ) = P() P() = (0.5) (0.5) = 0.25

Measuring a sample

Generate a sample that represents 1000 throws of two fair coin flips

```
from scipy.stats import binom
sample = binom.rvs(n=2, p=0.5, size=1000, random_state=1)
```

```
array([1, 1, 0, 1, 0, 0, 0, 1, 1, 1, 1, 1, 0, 2, 0, 1, 1, 1, 0, 0, 2, 2,...
```

Find repeated data

```
from scipy.stats import find_repeats
find_repeats(sample)
```

```
RepeatedResults(values=array([0., 1., 2.]), counts=array([249, 497, 254]))
```

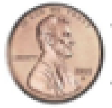
Unfair coin

P() = 0.2

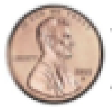
P() = 0.8

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

P( and ) = P() P() = (0.2) (0.2) = 0.04

P( and ) = P() P() = (0.2) (0.8) = 0.16

P( and ) = P() P() = (0.8) (0.2) = 0.16

P( and ) = P() P() = (0.8) (0.8) = 0.64

Measuring a biased sample

Using `biased_sample` data generated, calculate the relative frequency of each outcome

```
from scipy.stats import relfreq
relfreq(biased_sample, numbins=3).frequency
```

```
array([0.039, 0.317, 0.644])
```

Joint probability calculation

	Engine	Gear box
Fails	0.01	0.005
Works	0.99	0.995

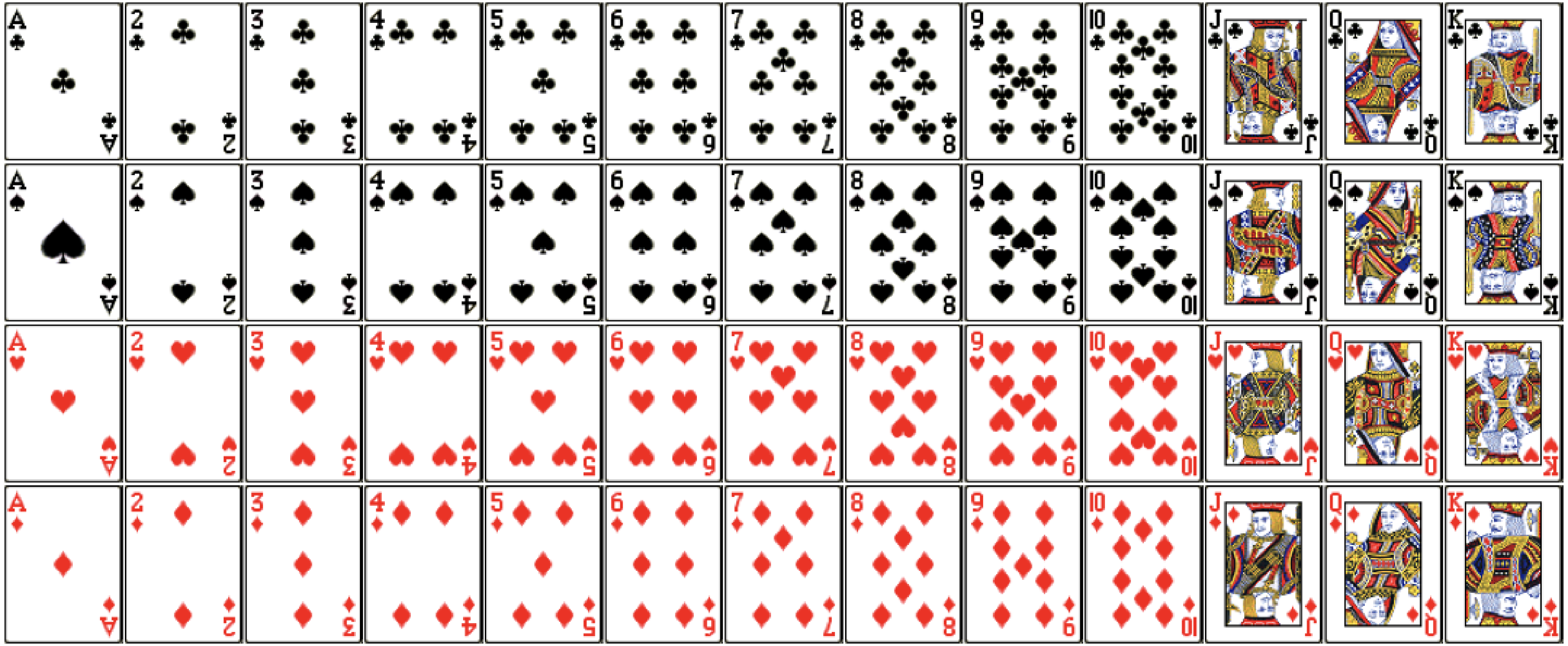
$$P(Engine\ fails\ and\ Gear\ box\ fails) = ?$$

```
P_Eng_fail = 0.01
P_GearB_fail = 0.005
P_both_fails = P_Eng_fail*P_GearB_fail
print(P_both_fails)
```

```
0.00005
```

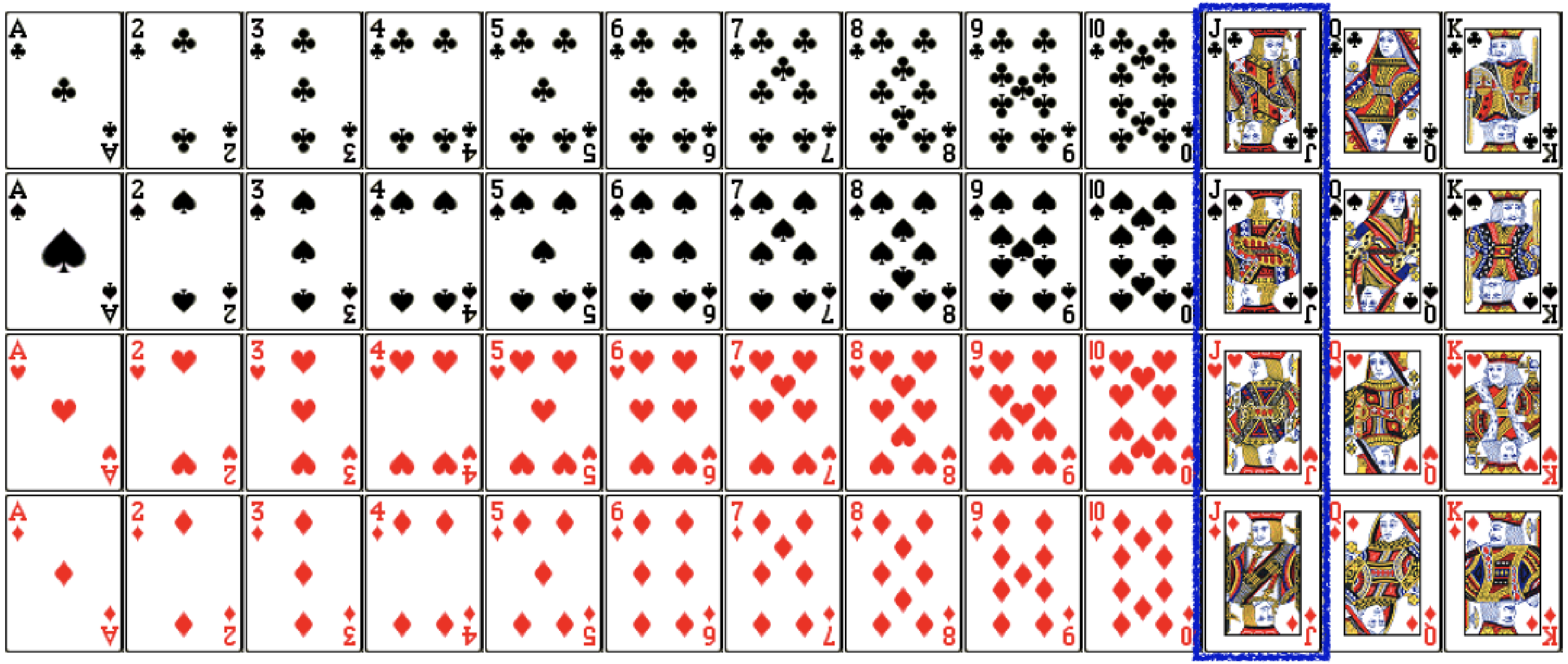
P(A or B) with cards

$$P(Jack\ or\ King) = ?$$



P(A or B) with cards (Cont.)

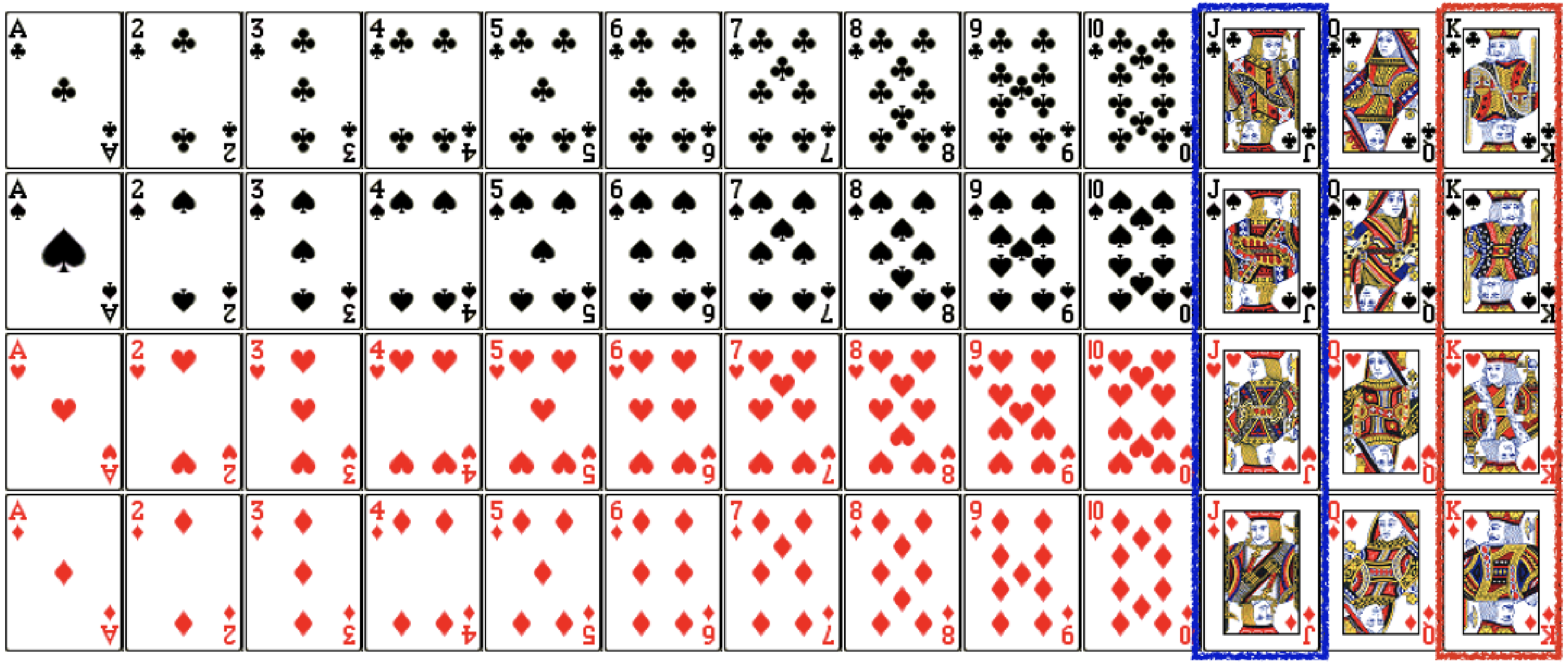
$$P(Jack\ or\ King) = P(Jack) + \ldots$$



$$P(Jack\ or\ King) = \frac{4}{52} + \ldots$$

P(A or B) with cards (Cont.)

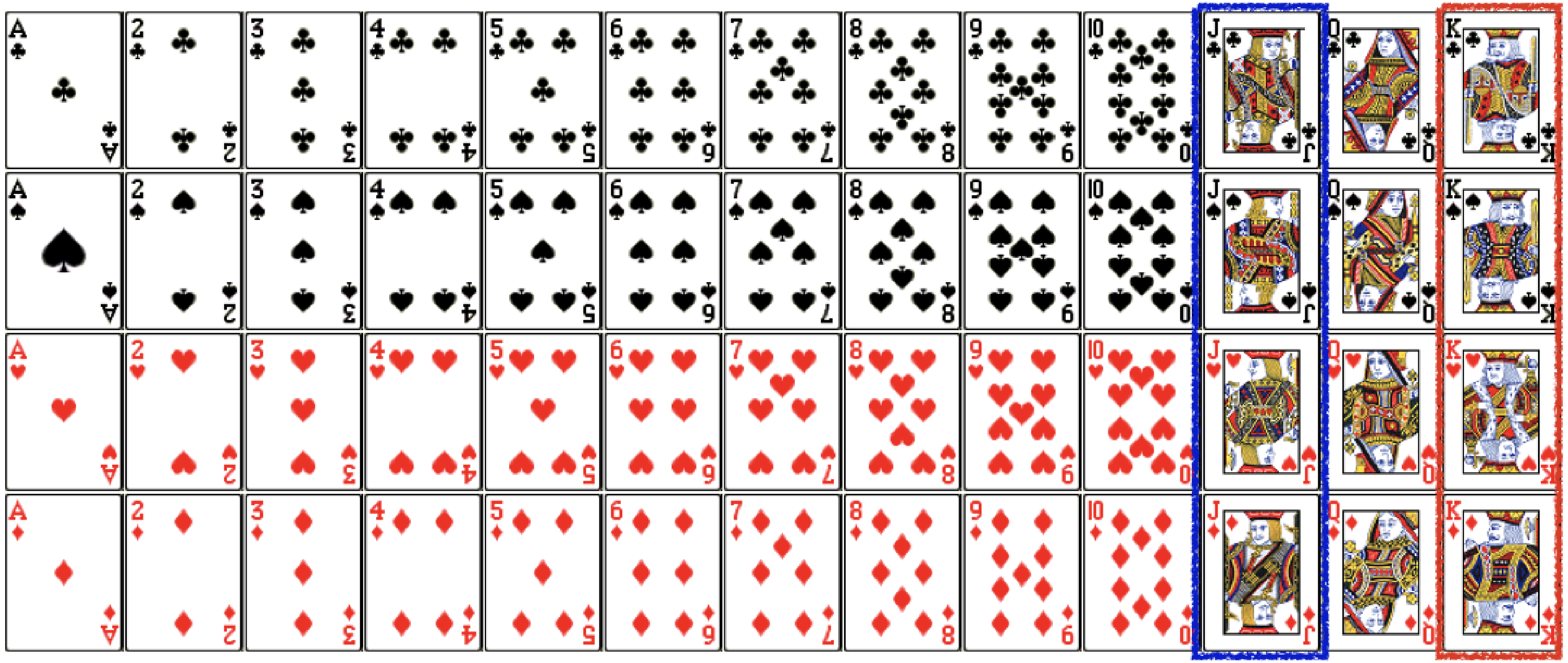
$$P(Jack\ or\ King) = P(Jack) + P(King)$$



$$P(Jack\ or\ King) = \frac{4}{52} + \frac{4}{52}$$

P(A or B) with cards (Cont.)

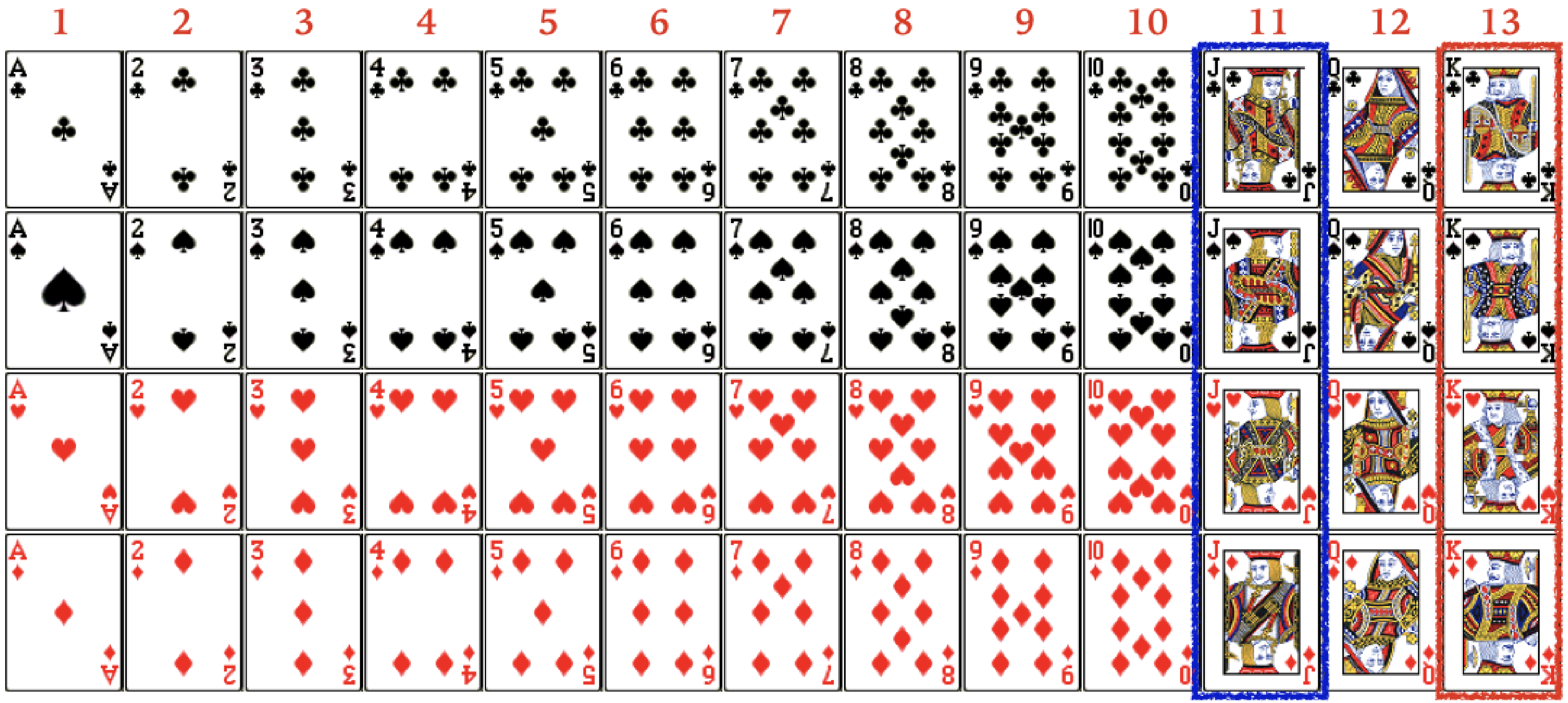
$$P(Jack\ or\ King) = P(Jack) + P(King)$$



$$P(Jack\ or\ King) = \frac{4}{52} + \frac{4}{52} = \frac{8}{52} = \frac{2}{13}$$

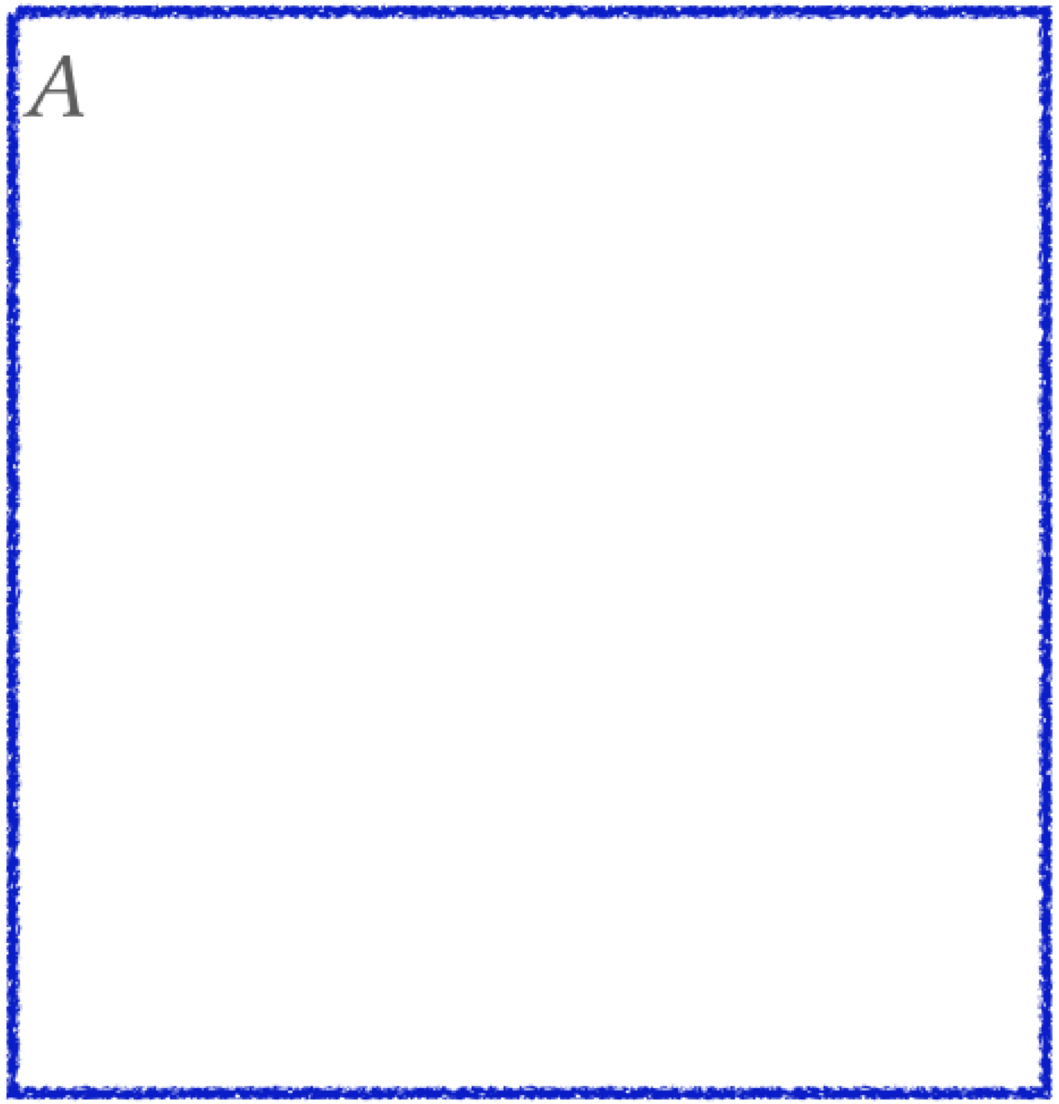
P(A or B) with cards (Cont.)

$$P(Jack\ or\ King) = P(Jack) + P(King)$$

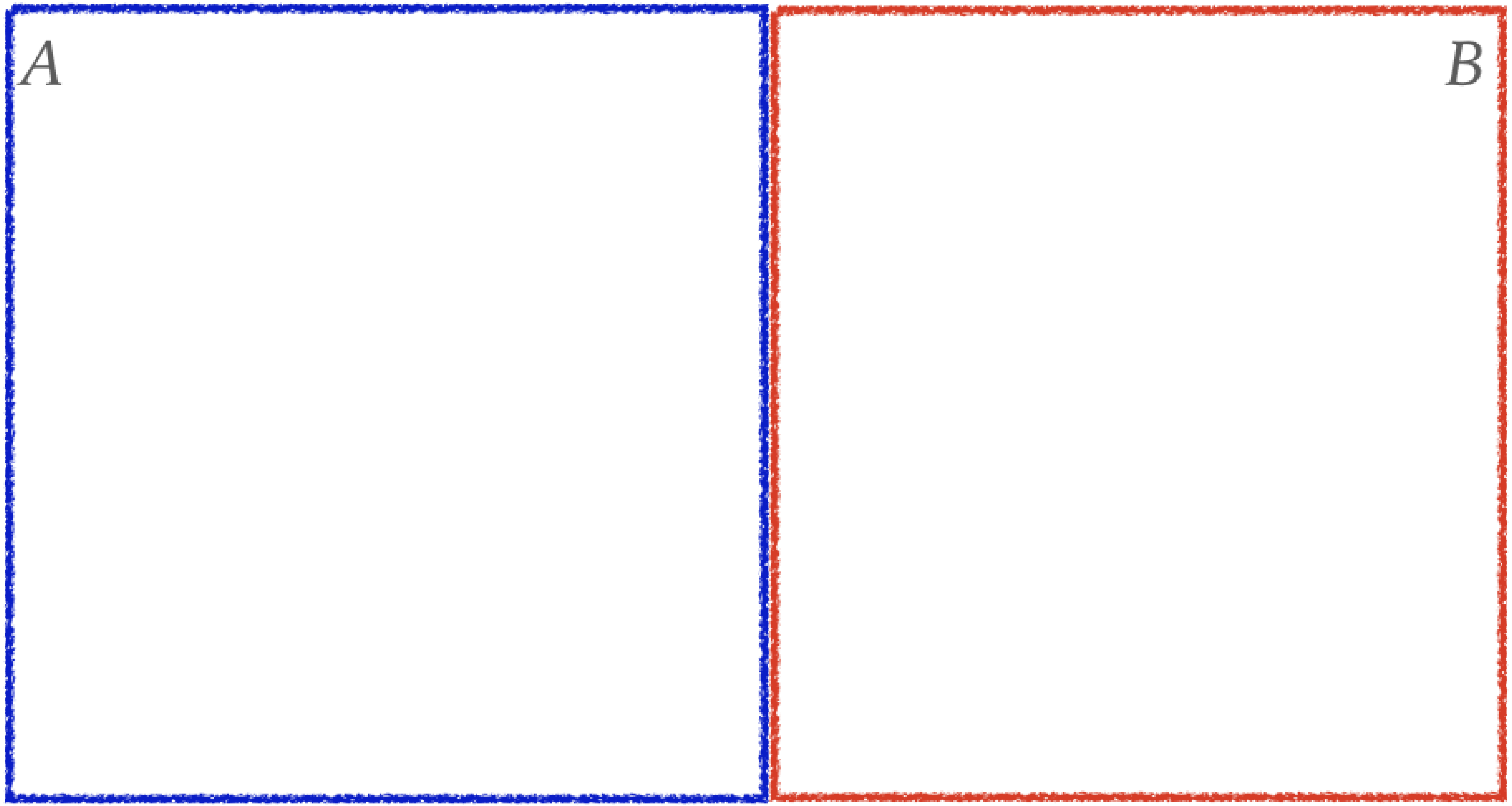


$$P(Jack\ or\ King) = \frac{4}{52} + \frac{4}{52} = \frac{8}{52} = \frac{2}{13}$$

Probability of A or B

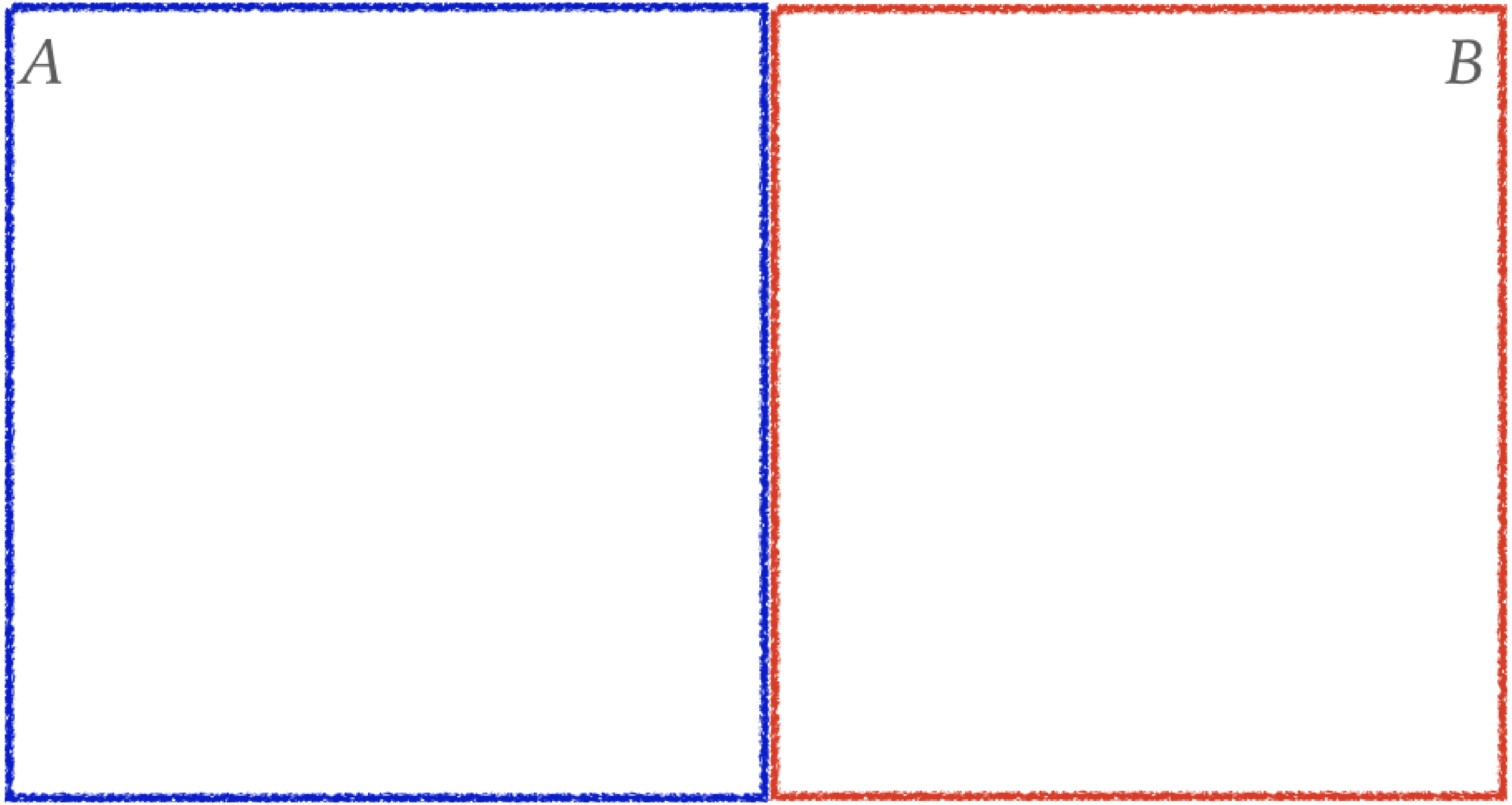


Probability of A or B (Cont.)



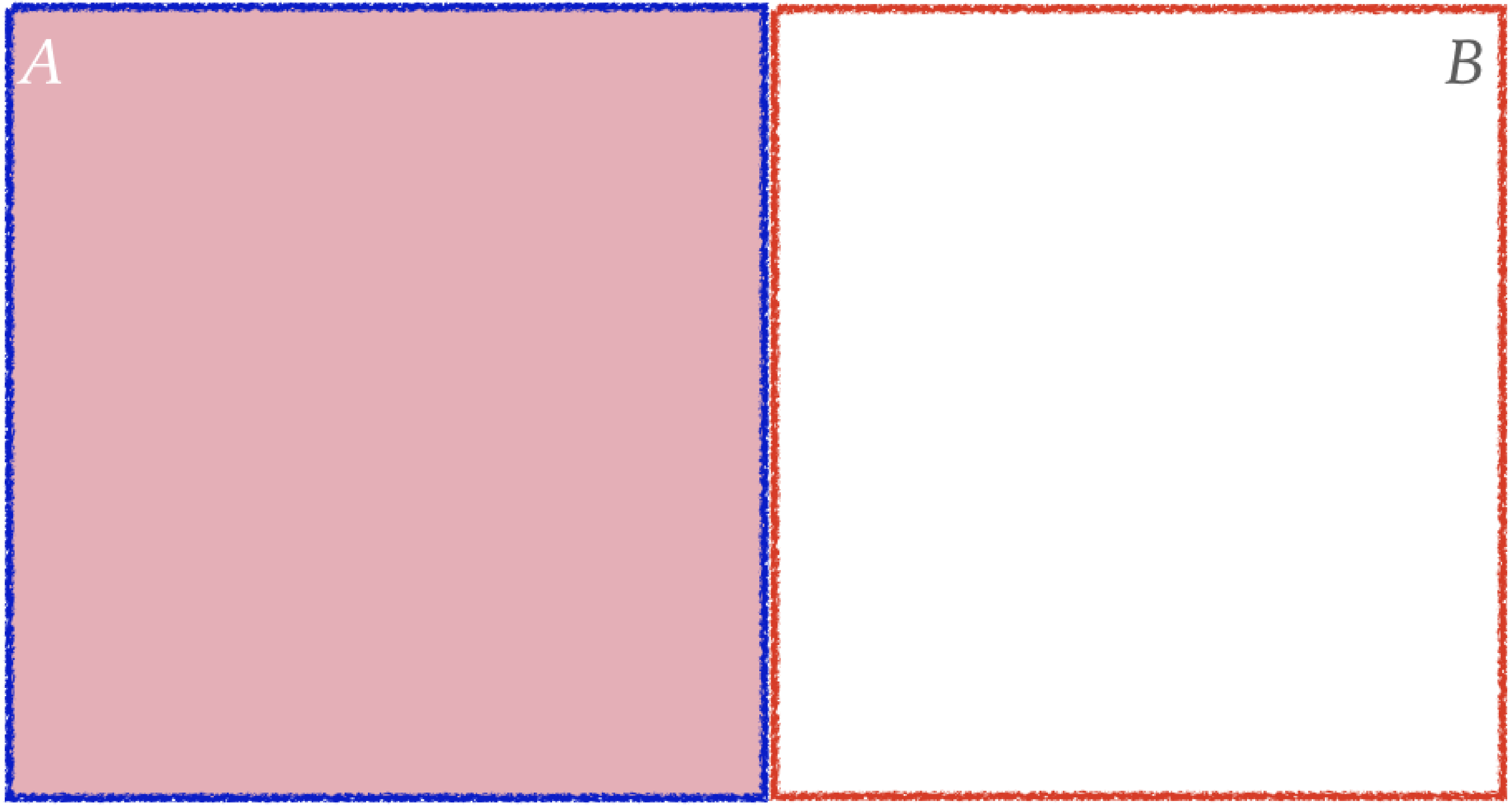
Probability of A or B (Cont.)

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



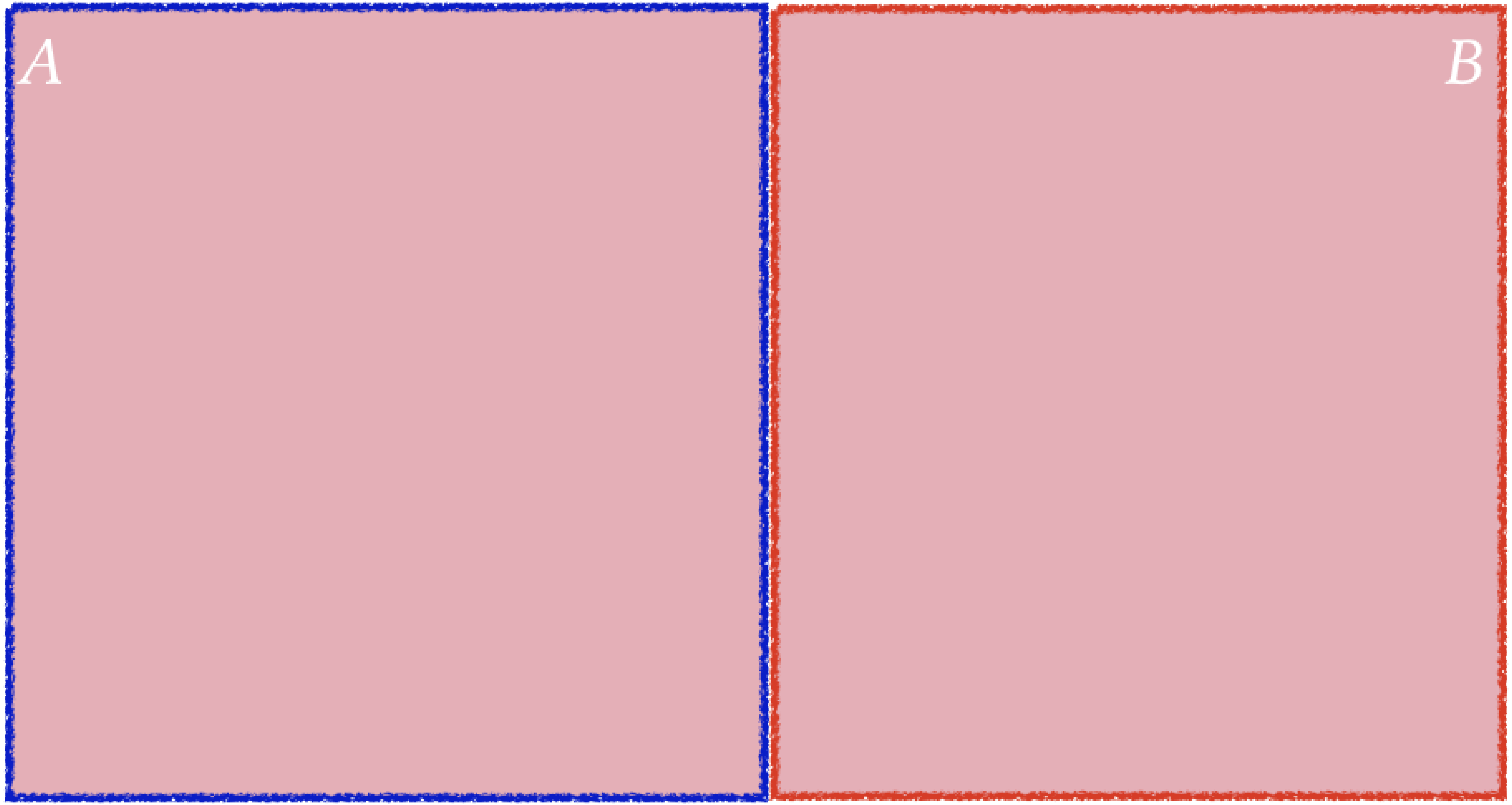
Probability of A or B (Cont.)

$$P(A\ or\ B) = P(A) + \ldots$$



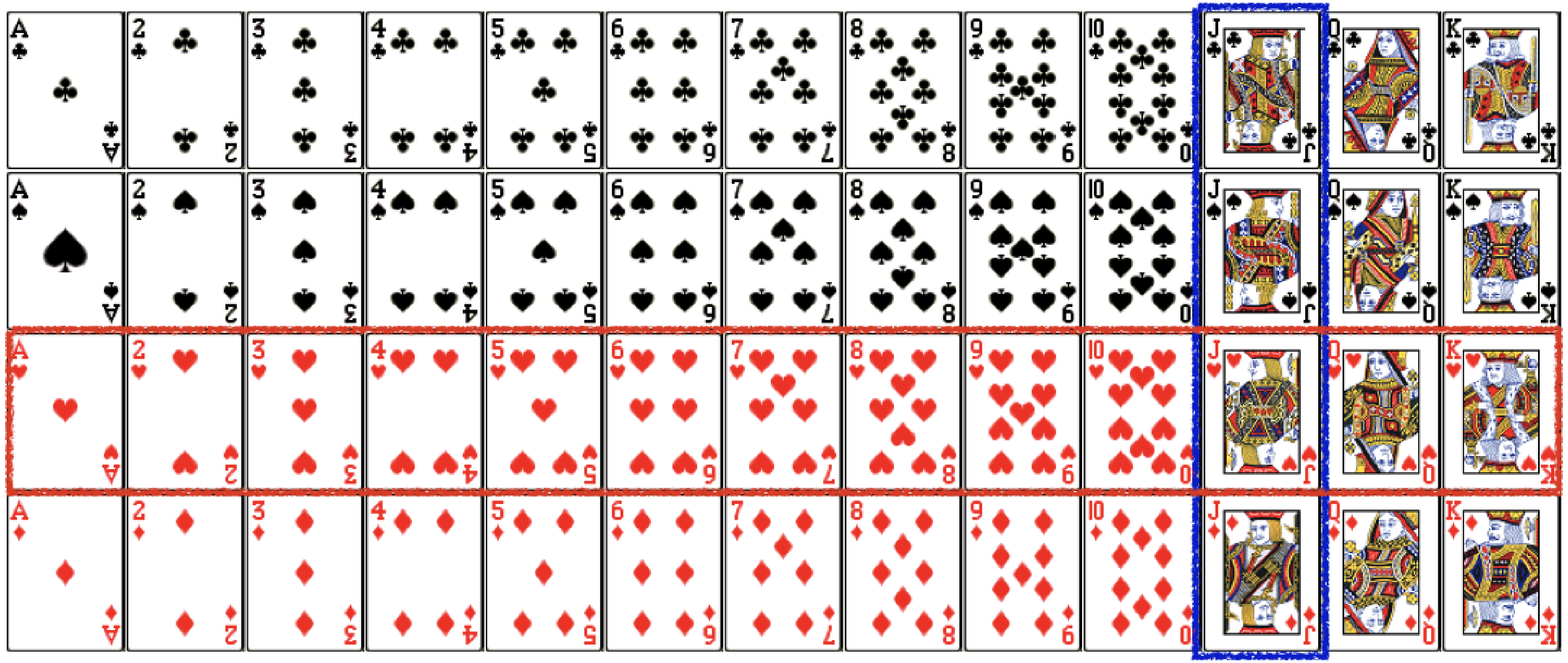
Probability of A or B (Cont.)

$$P(A\ or\ B) = P(A) + P(B)$$



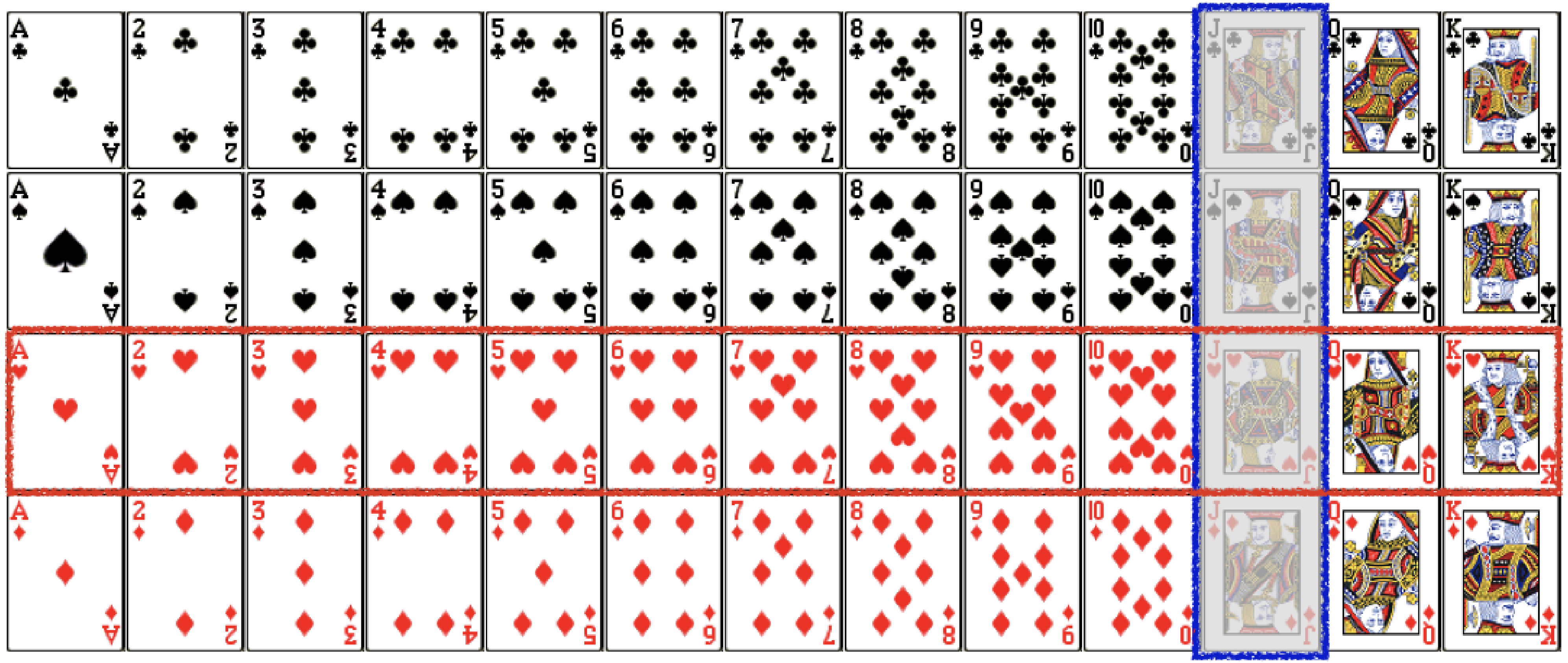
P(A or B) with overlap

$$P(Jack\ or\ Heart) = ?$$



P(A or B) with overlap (Cont.)

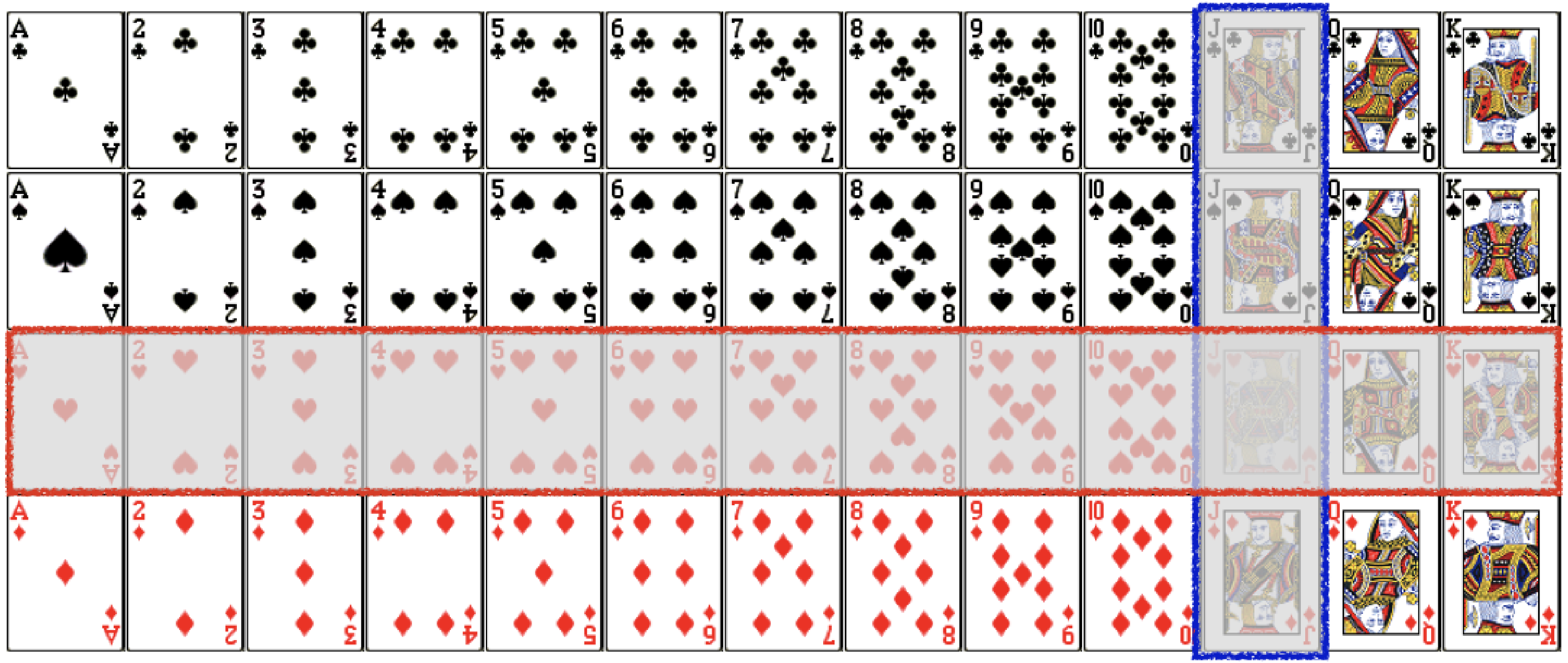
$$P(Jack\ or\ Heart) = P(Jack) + \ldots$$



$$P(Jack\ or\ Heart) = \frac{4}{52} + \ldots$$

P(A or B) with overlap (Cont.)

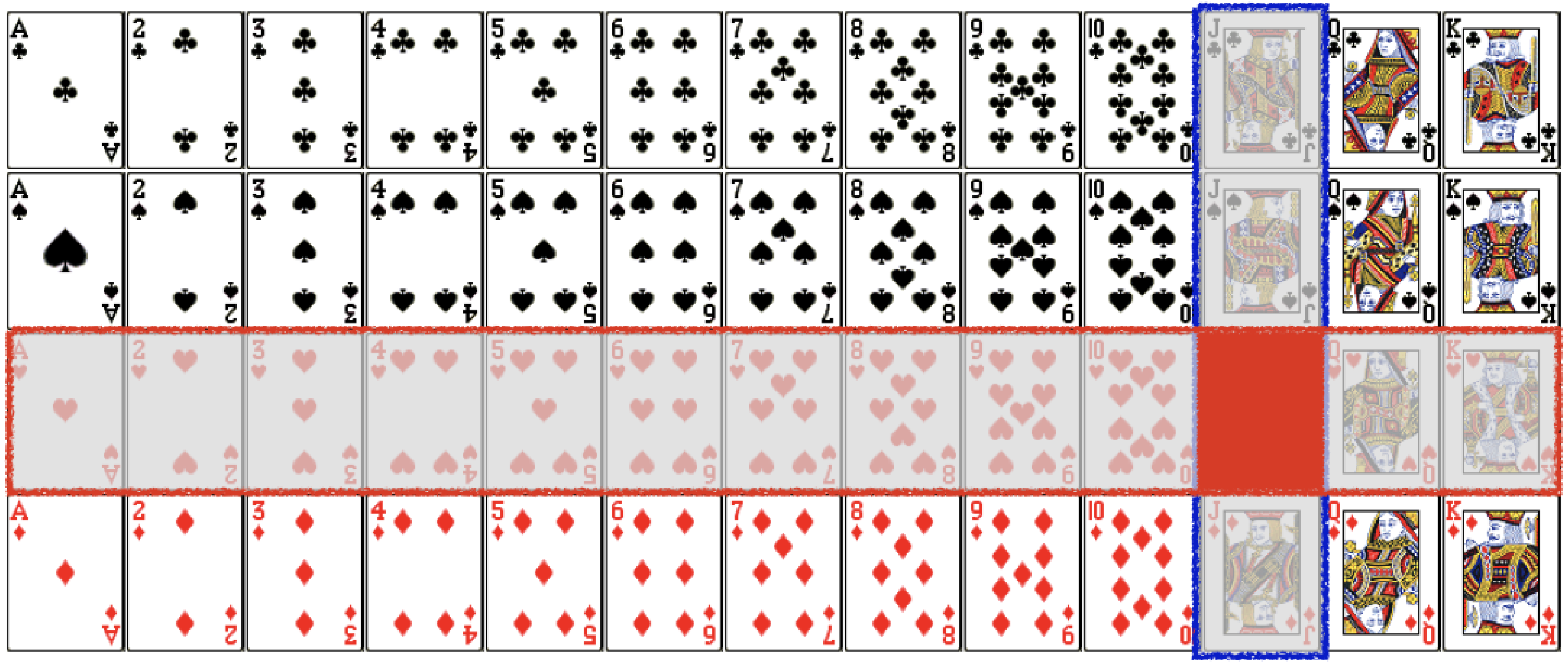
$$P(Jack\ or\ Heart) = P(Jack) + \color{red}{P(Heart)} \ldots$$



$$P(Jack\ or\ Heart) = \frac{4}{52} + \color{red}{\frac{13}{52}} \ldots$$

P(A or B) with overlap (Cont.)

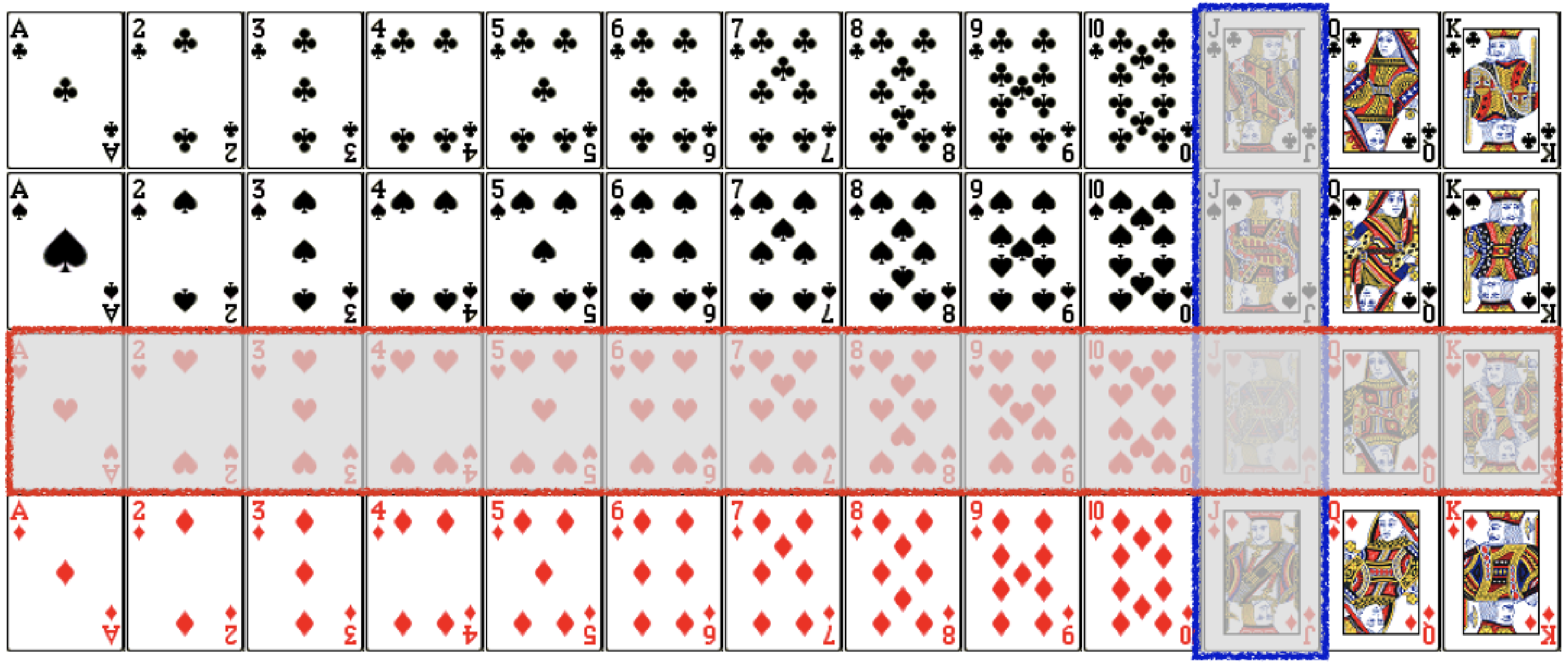
$$P(Jack\ or\ Heart) = P(Jack) + P(Heart)...$$



$$P(Jack\ or\ Heart) = \frac{4}{52} + \frac{13}{52}...$$

P(A or B) with overlap (Cont.)

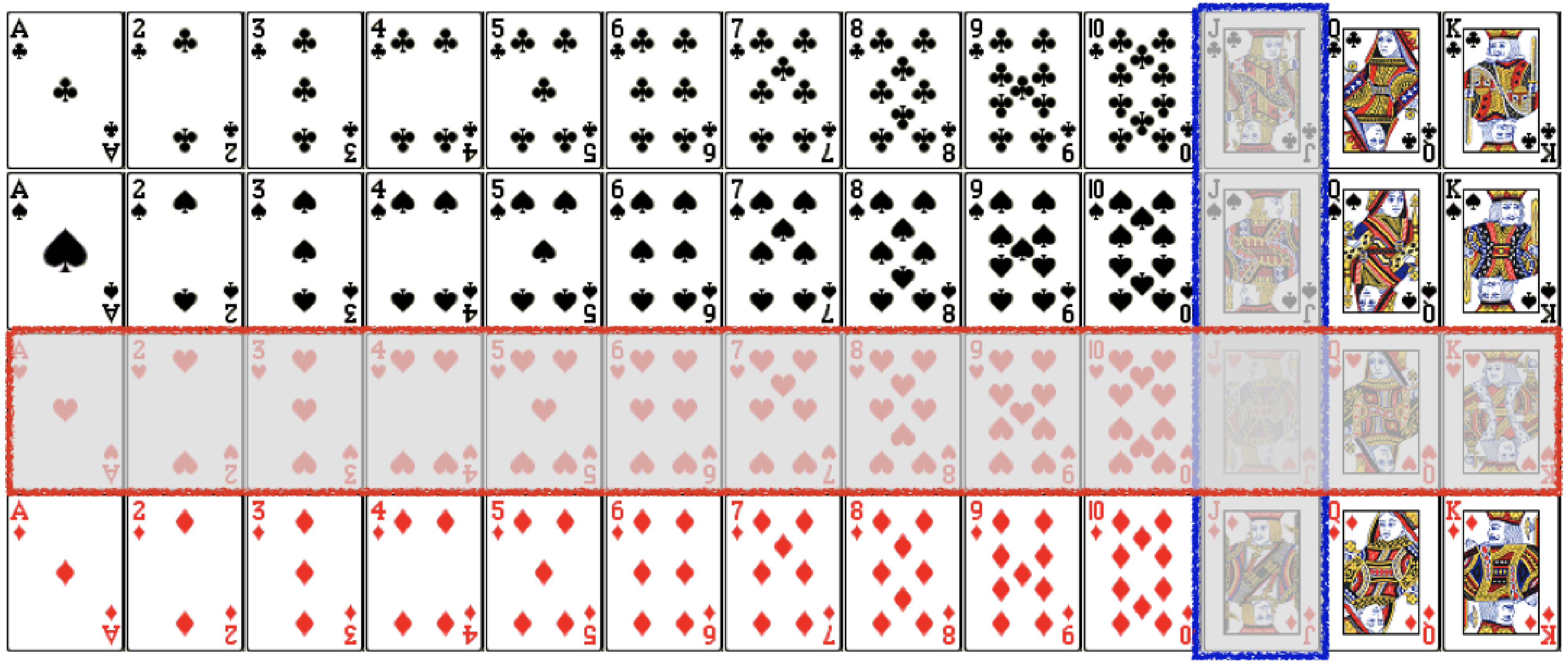
$$P(Jack\ or\ Heart) = P(Jack) + P(Heart) - P(Jack\ and\ Heart)$$



$$P(Jack\ or\ Heart) = \frac{4}{52} + \frac{13}{52} - \frac{1}{52}$$

P(A or B) with overlap (Cont.)

$$P(Jack\ or\ Heart) = P(Jack) + P(Heart) - P(Jack\ and\ Heart)$$



$$P(Jack\ or\ Heart) = \frac{4}{52} + \frac{13}{52} - \frac{1}{52} = \frac{16}{52} = \frac{4}{13}$$

Diagram of P(A or B)

Diagram of P(A or B) (Cont.)

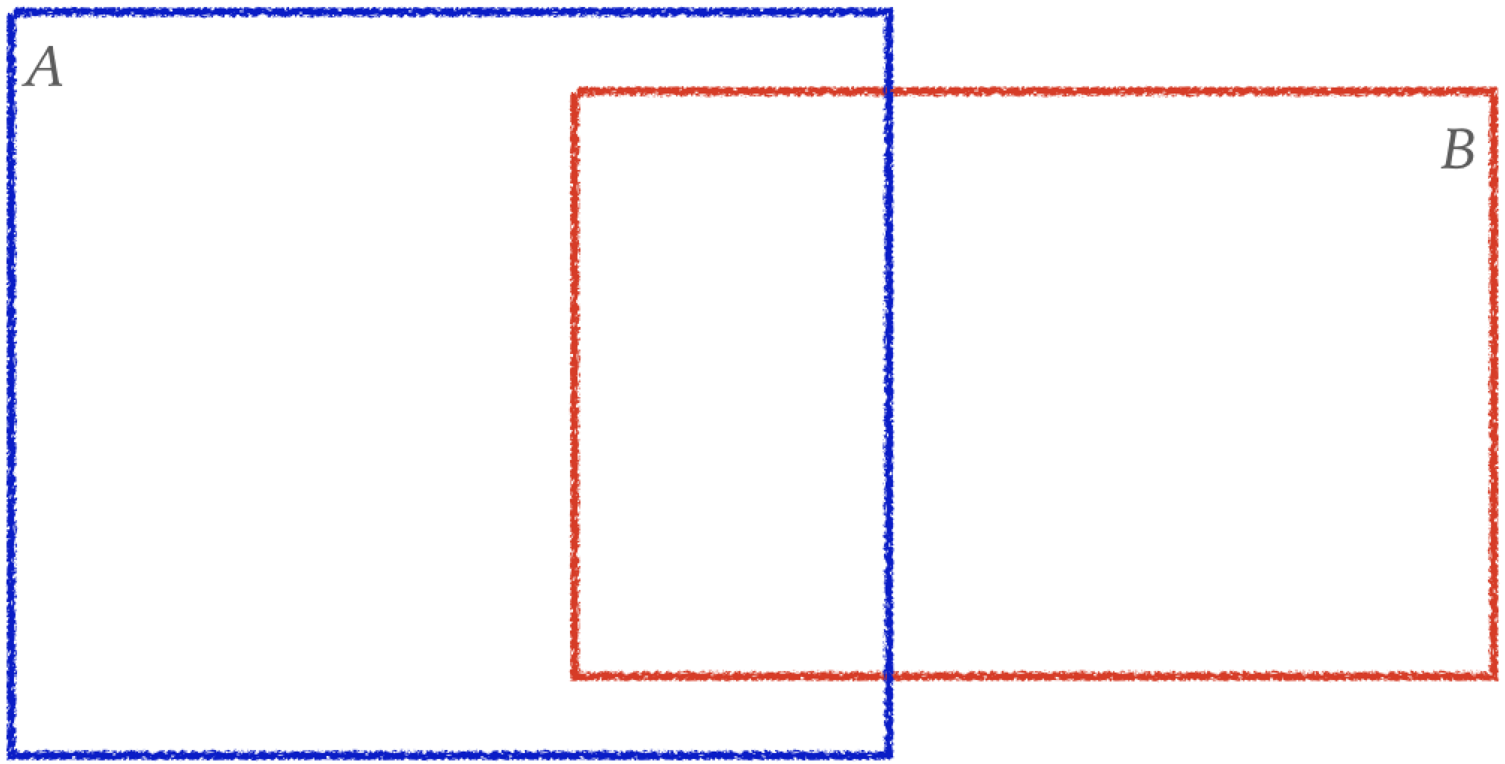


Diagram of P(A or B) (Cont.)

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

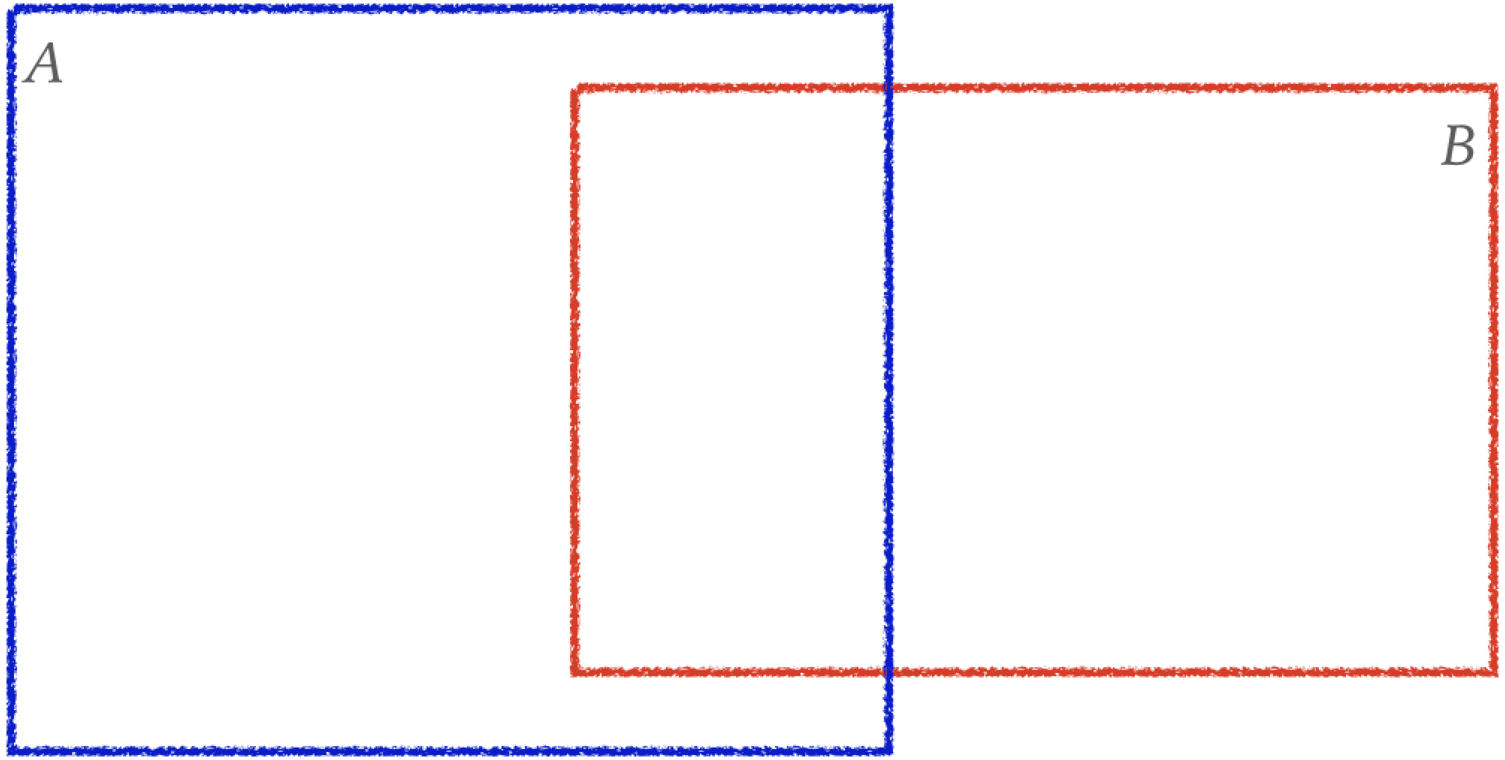


Diagram of P(A or B) (Cont.)

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

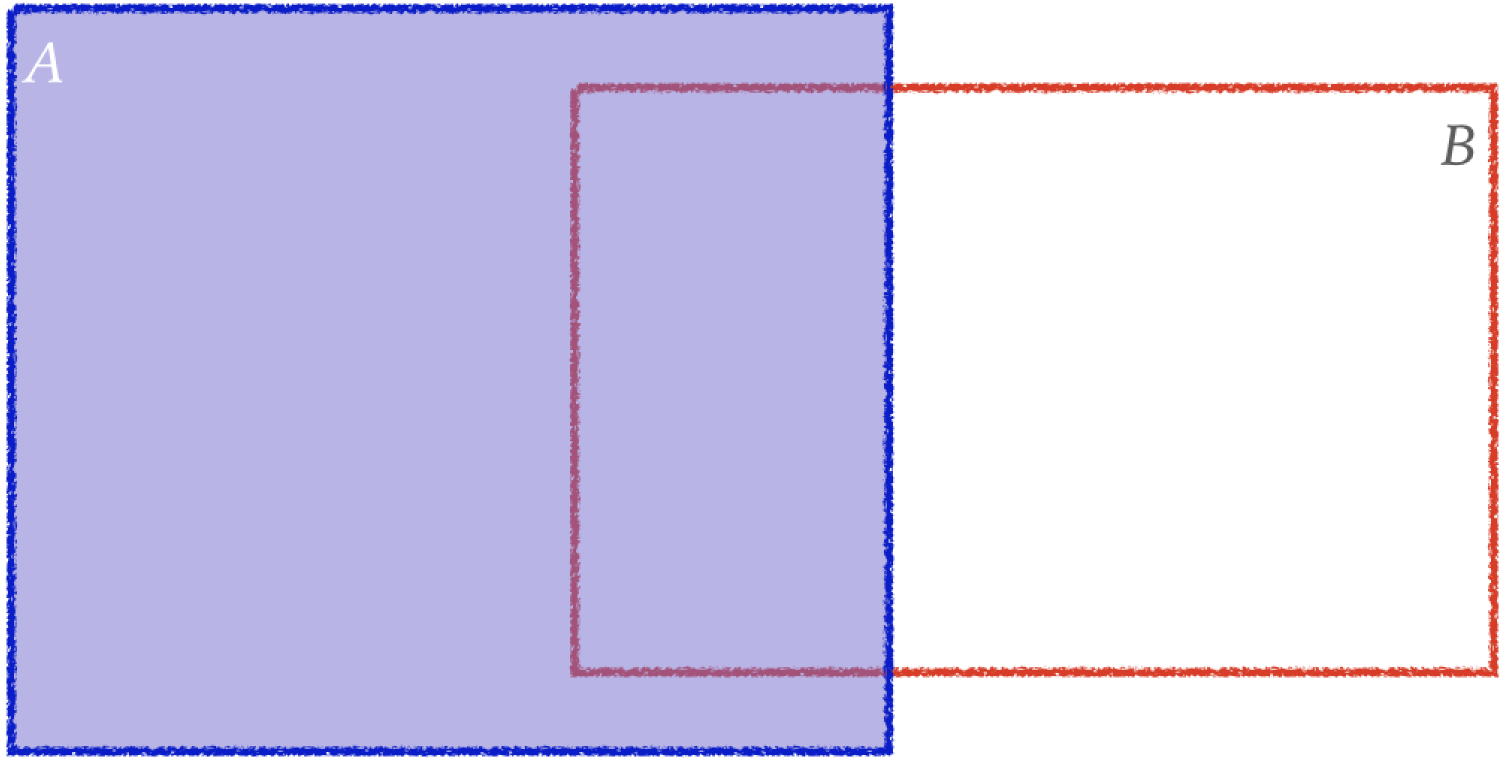


Diagram of P(A or B) (Cont.)

$$P(A\ or\ B) = P(A) + \color{red}{P(B)}\ ...$$

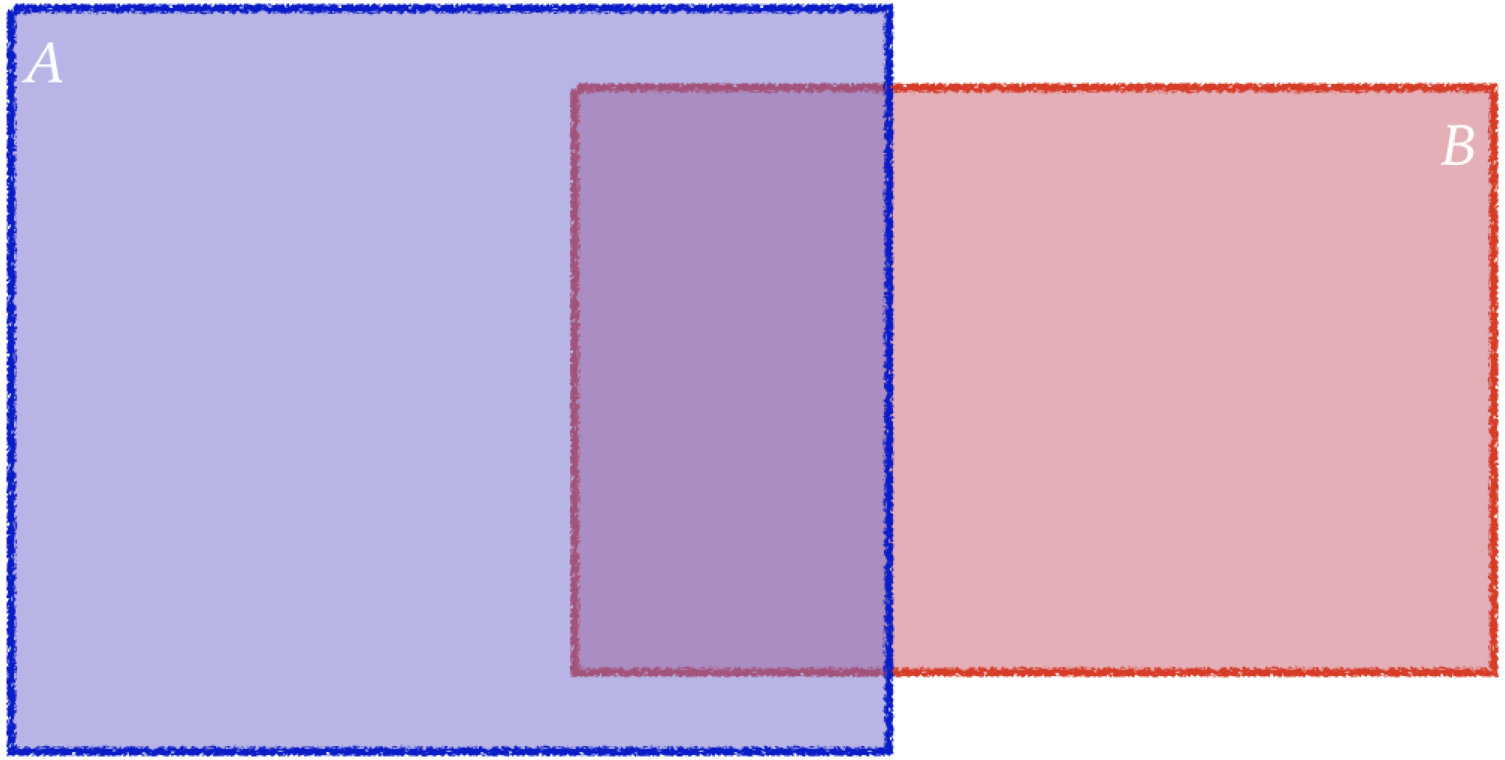


Diagram of P(A or B) (Cont.)

$$P(A\ or\ B) = P(A) + P(B)\ ...$$

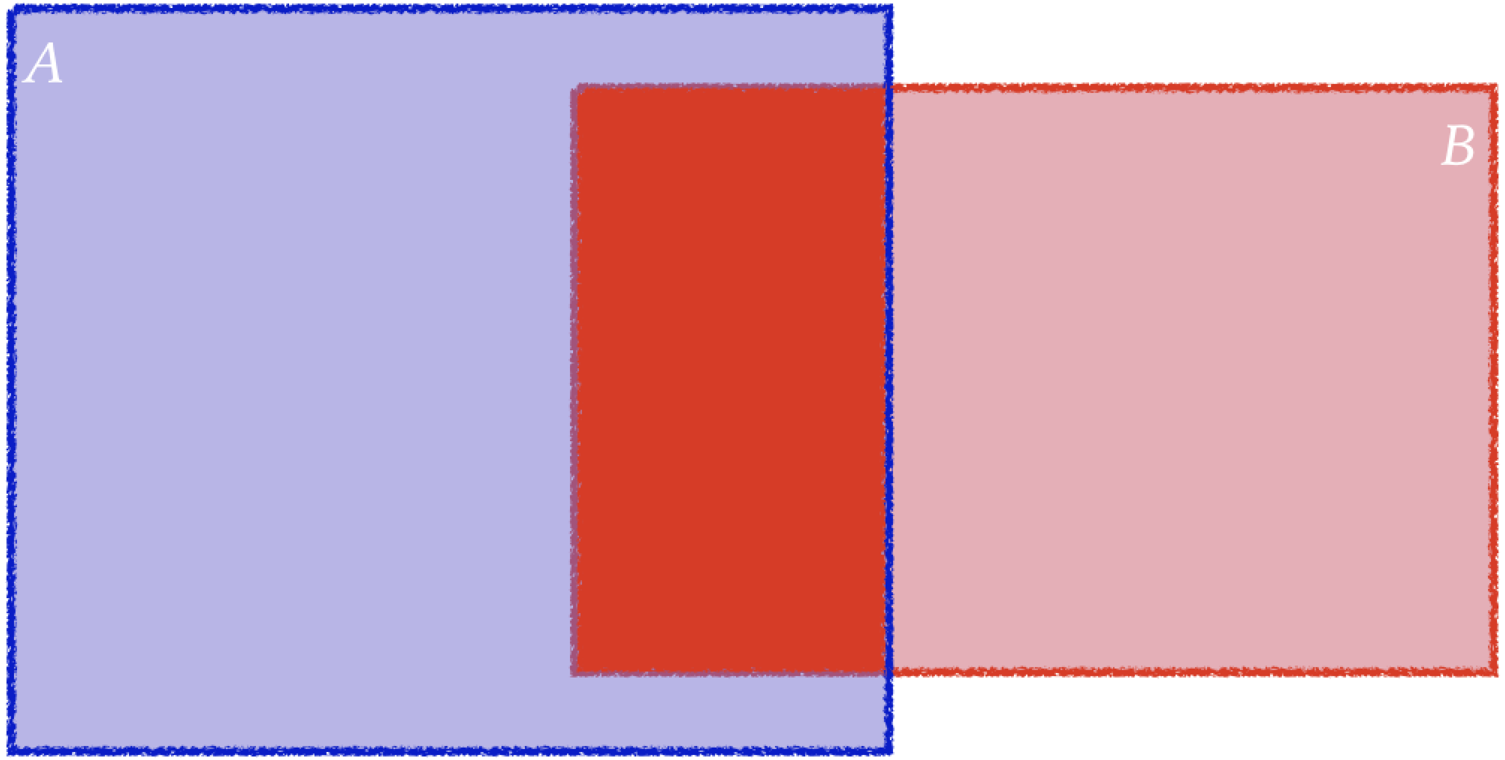


Diagram of P(A or B) (Cont.)

$$P(A \; or \; B) = P(A) + P(B) - P(A \; and \; B)$$

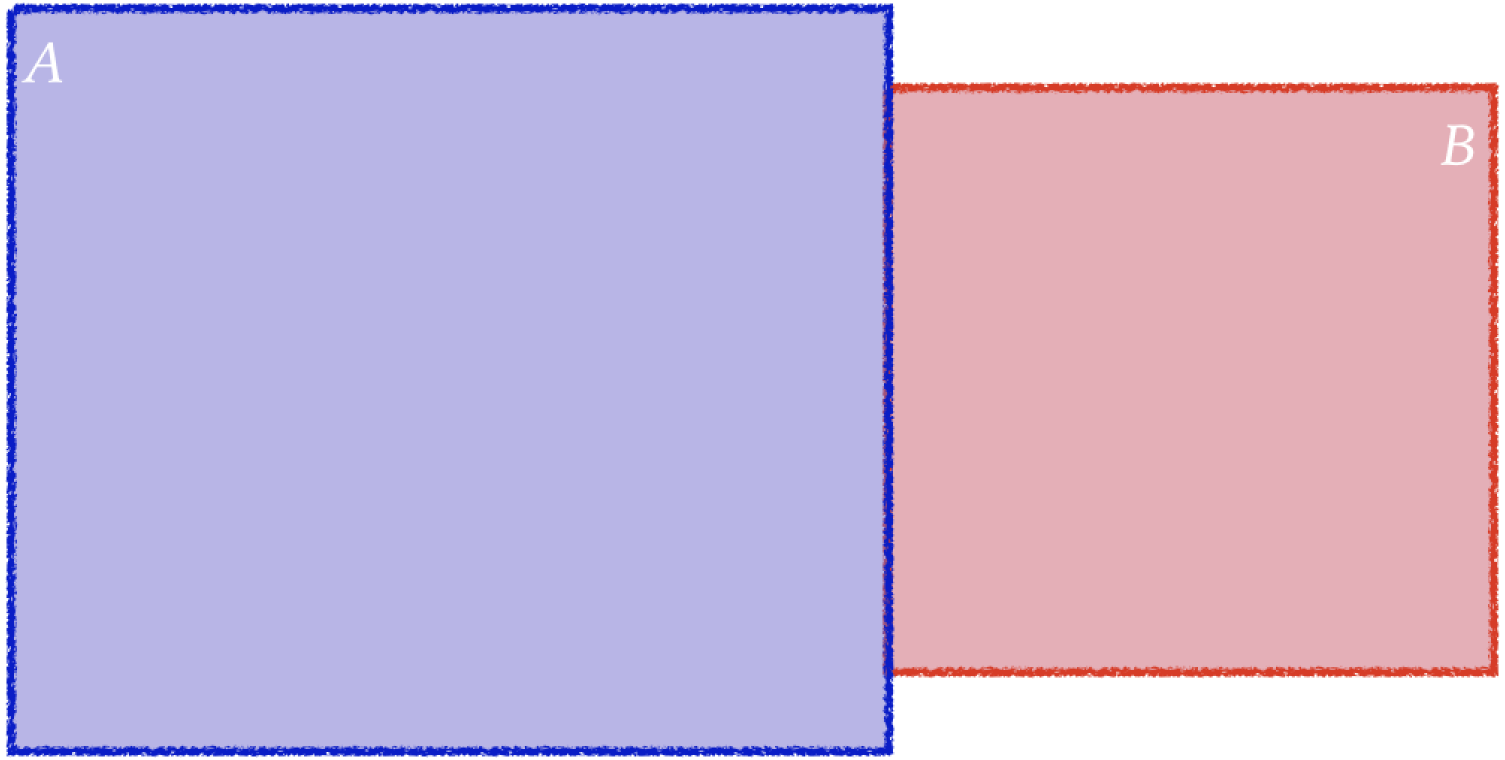
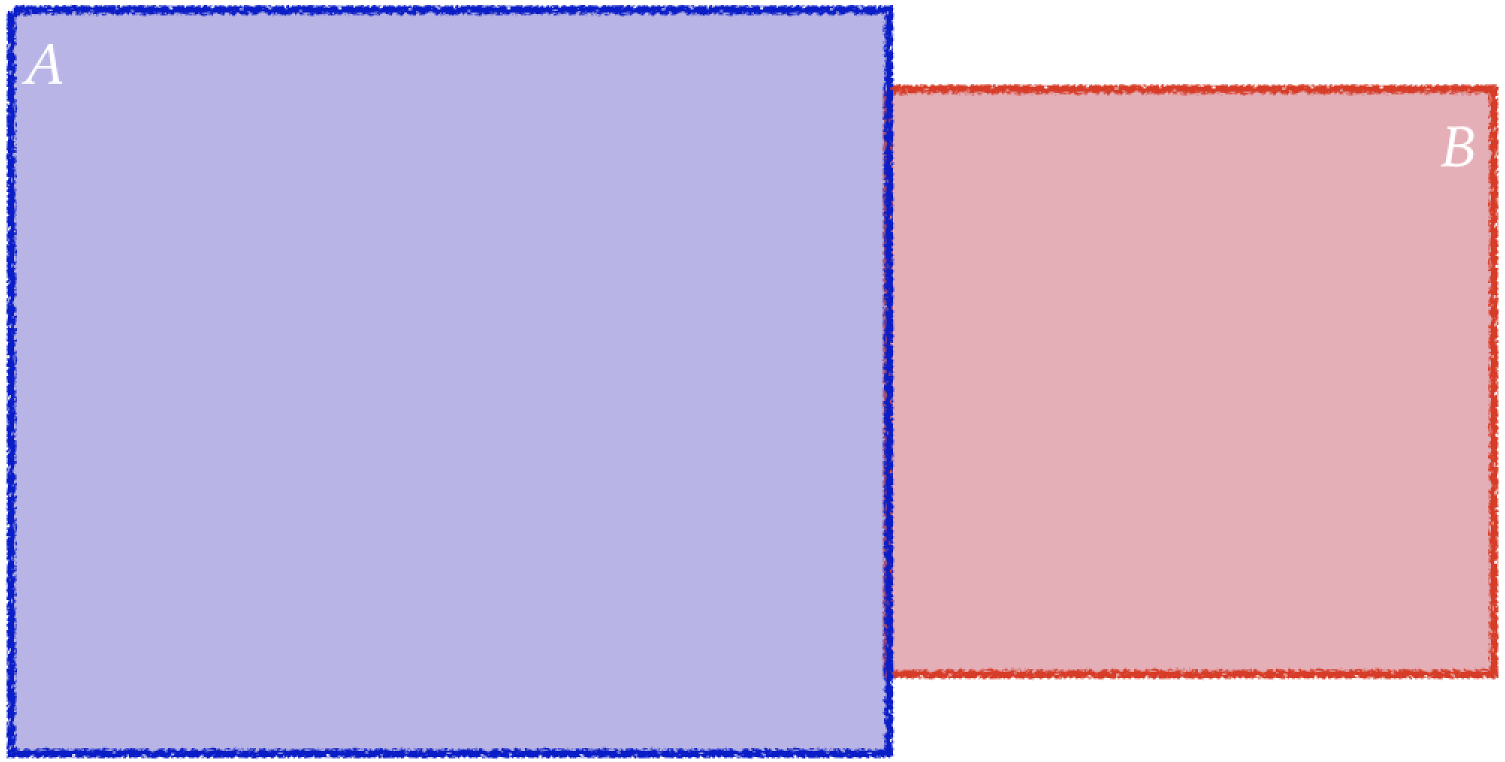


Diagram of P(A or B) (Cont.)

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



P(Jack or Heart) calculation in Python

```
P_Jack = 4/52
P_Hearts = 13/52
P_Jack_n_Hearts = 1/52
P_Jack_or_Hearts = P_Jack + P_Hearts - P_Jack_n_Hearts
print(P_Jack_or_Hearts)
```

```
0.307692307692
```

Let's calculate probabilities of two events

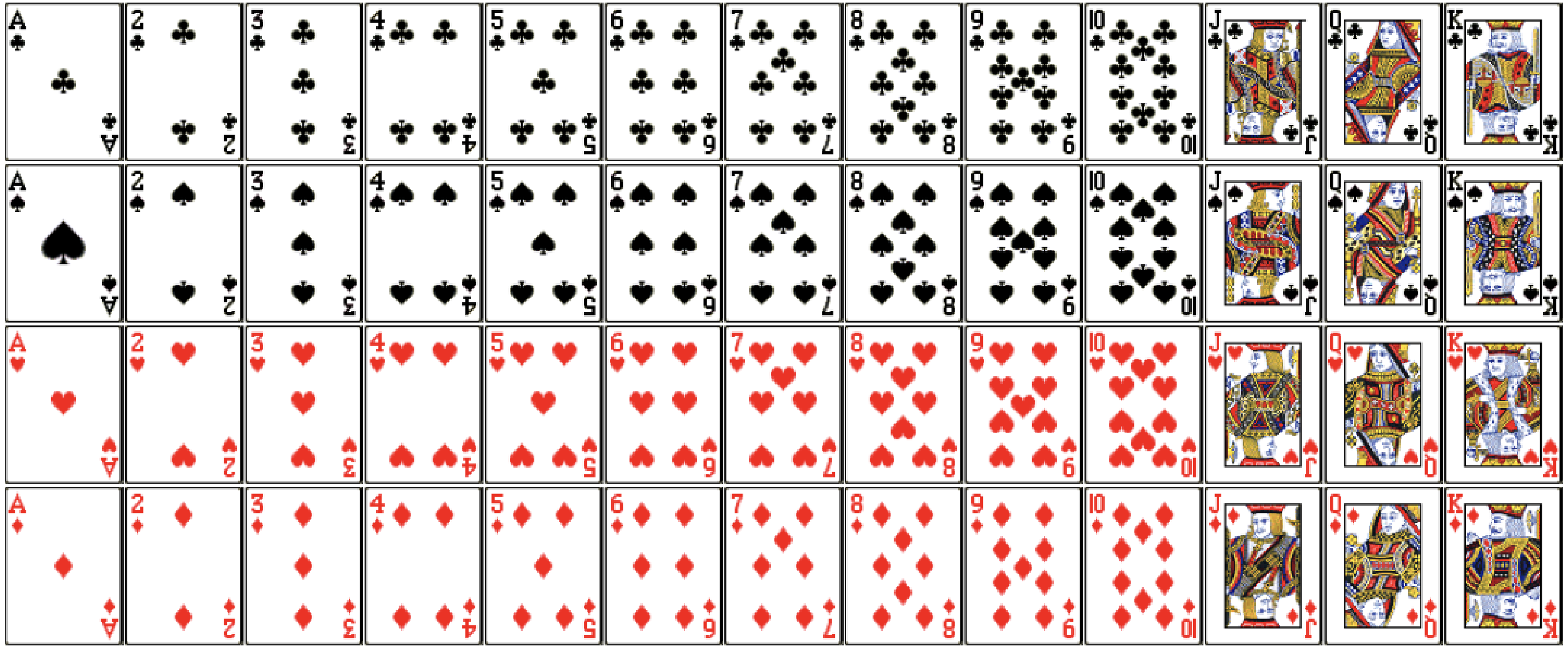
FOUNDATIONS OF PROBABILITY IN PYTHON

Conditional probabilities

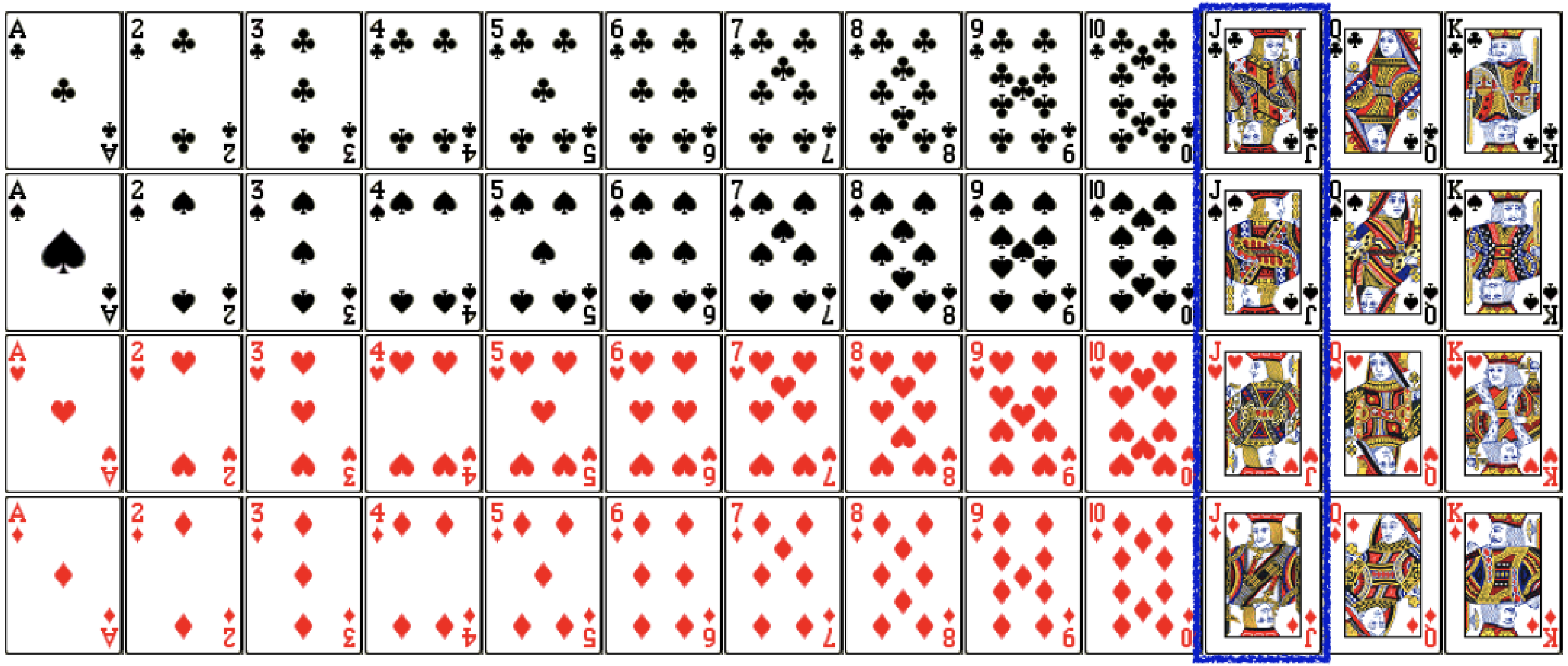
FOUNDATIONS OF PROBABILITY IN PYTHON

Alexander A. Ramírez M.
CEO @ Synergy Vision

Dependent events

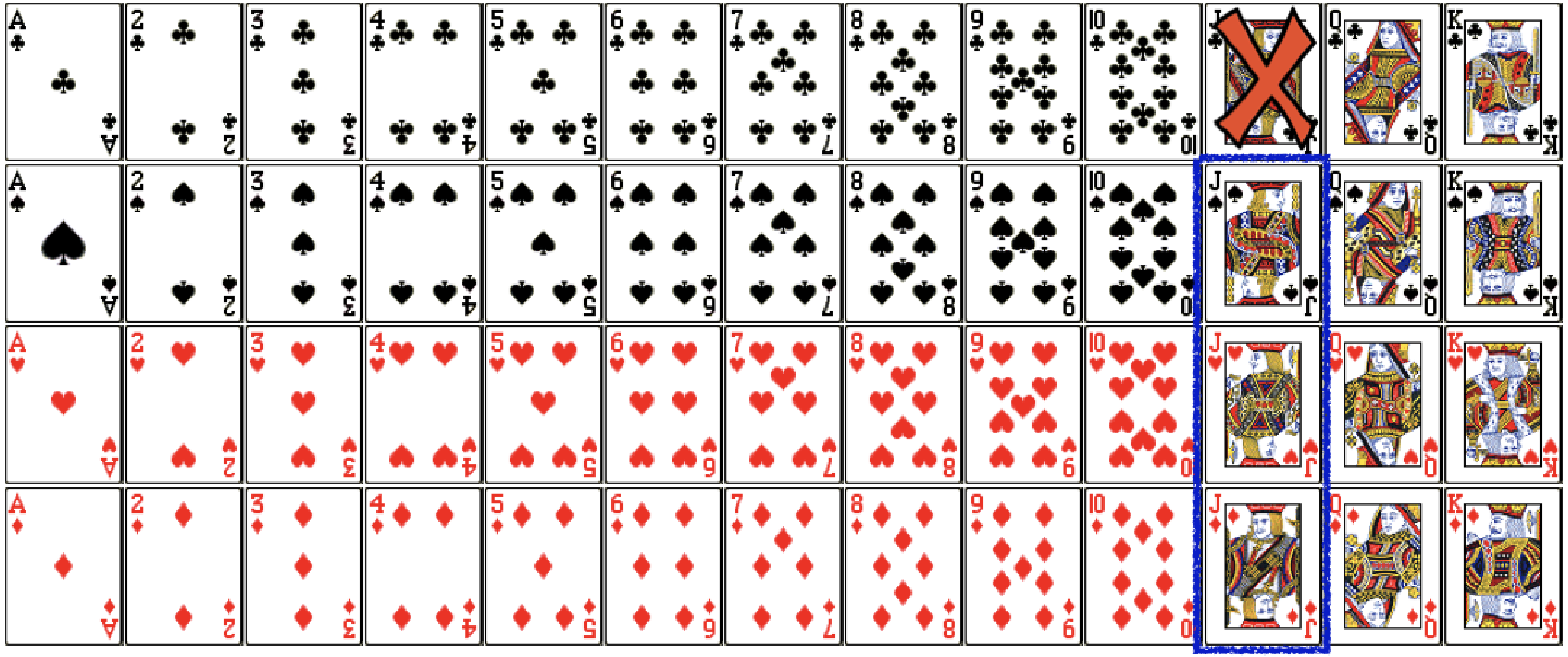


Dependent events (Cont.)



$$P(Jack) = \frac{4}{52} \simeq 7.69\%$$

Dependent events (Cont.)



$$P(Jack) = \frac{3}{51} \simeq 5.88\%$$

Conditional probability formula

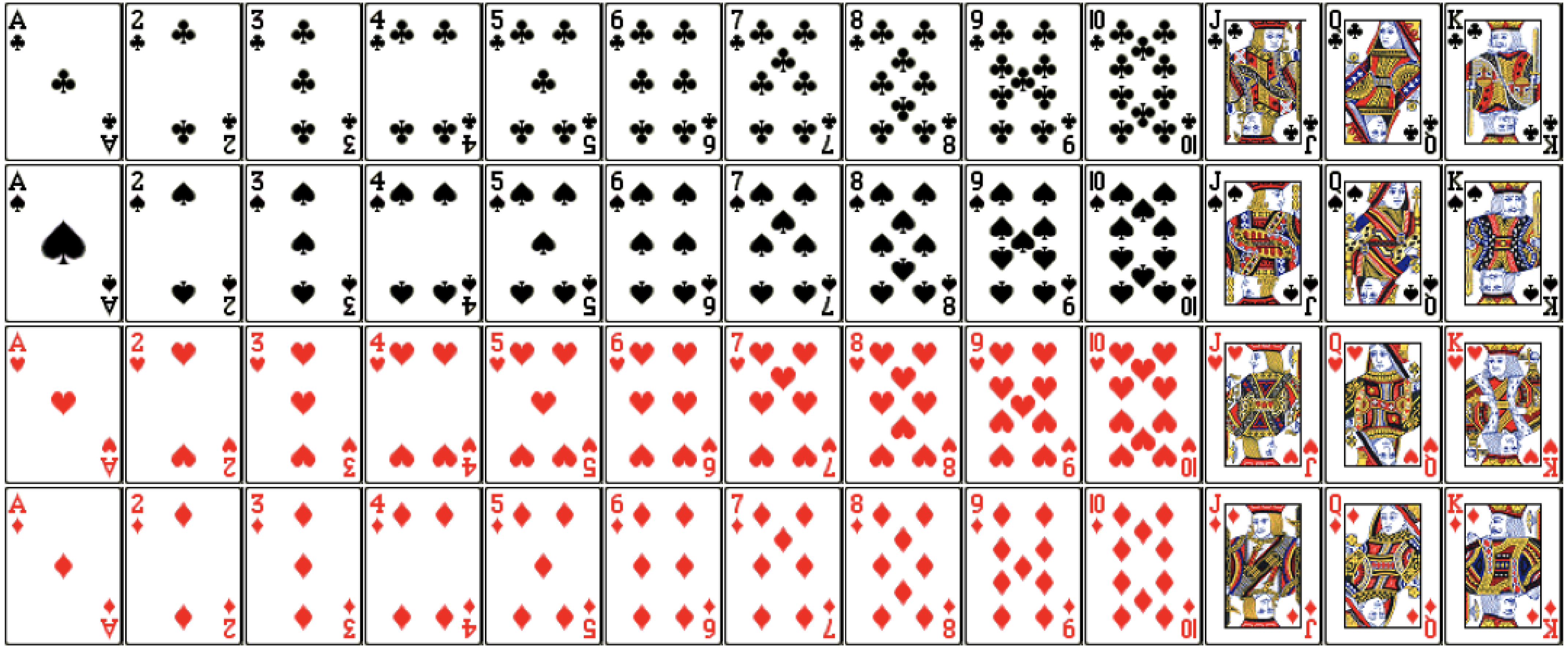
$$P(A\ and\ B) = P(A)P(B)$$

Conditional probability formula (Cont.)

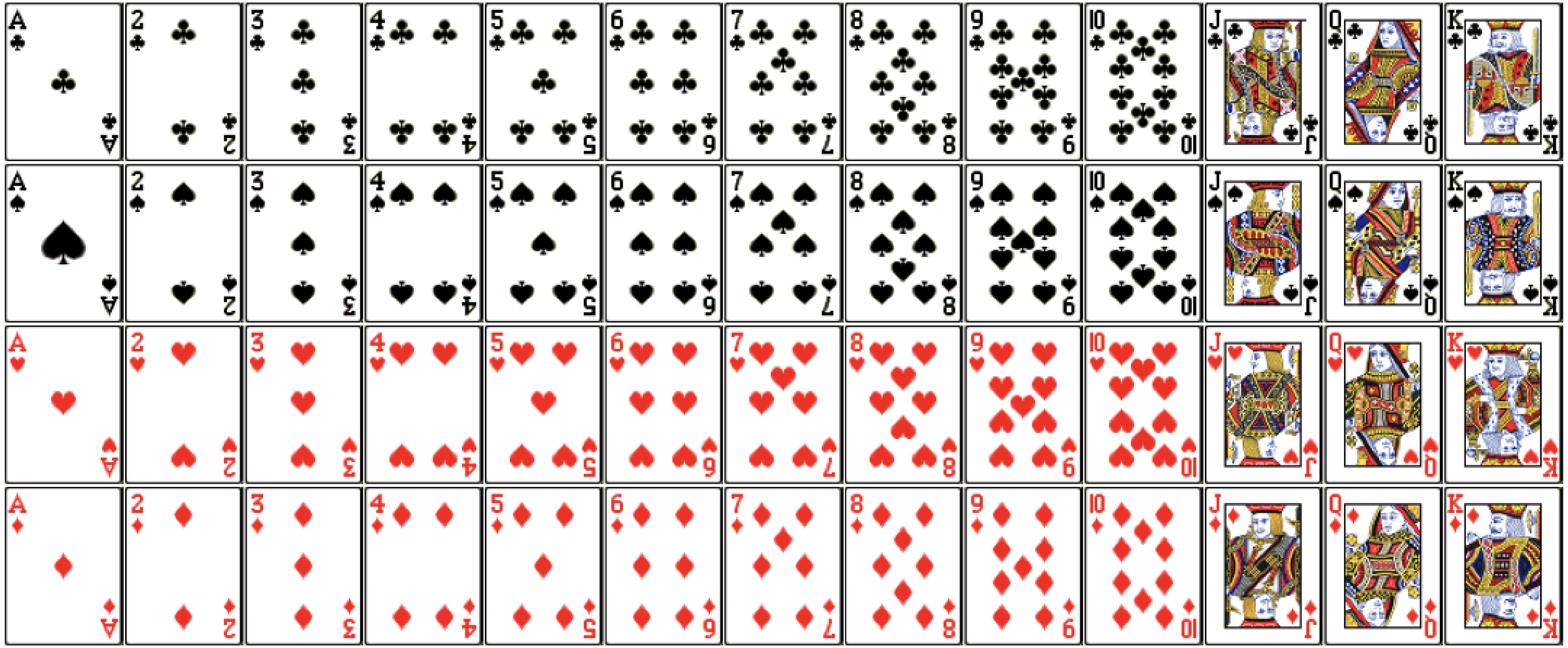
$$P(A \; and \; B) = P(A)P(B|A)$$

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

Conditional probability

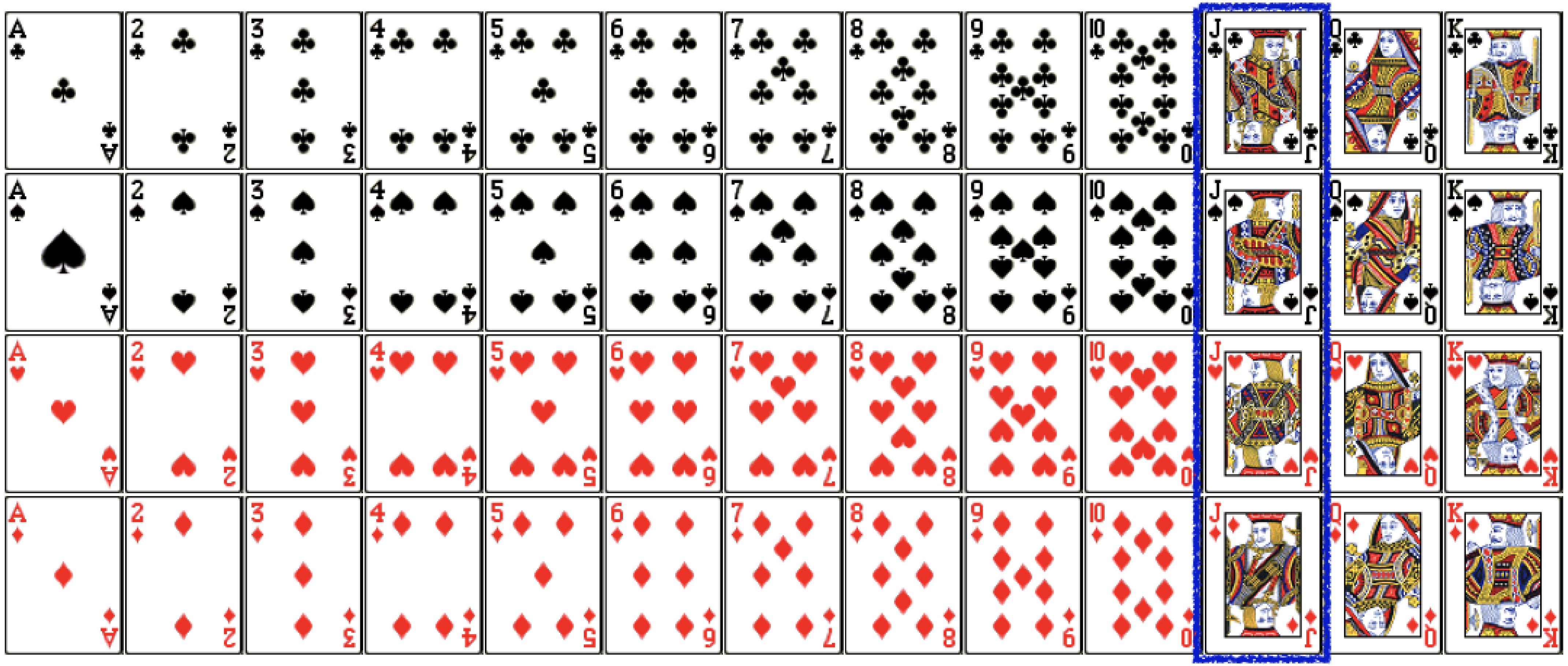


Conditional probability (Cont.)



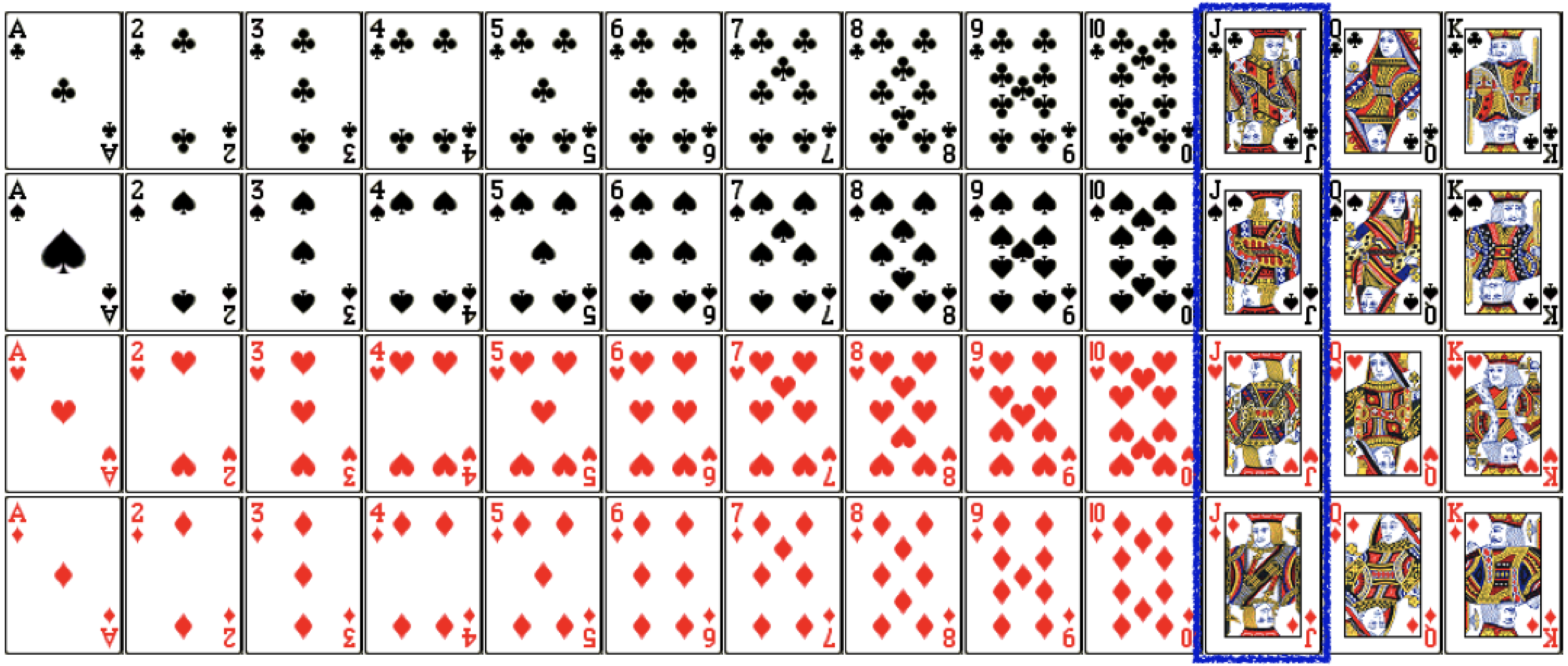
$$P(Red|Jack) = ?$$

Conditional probability (Cont.)



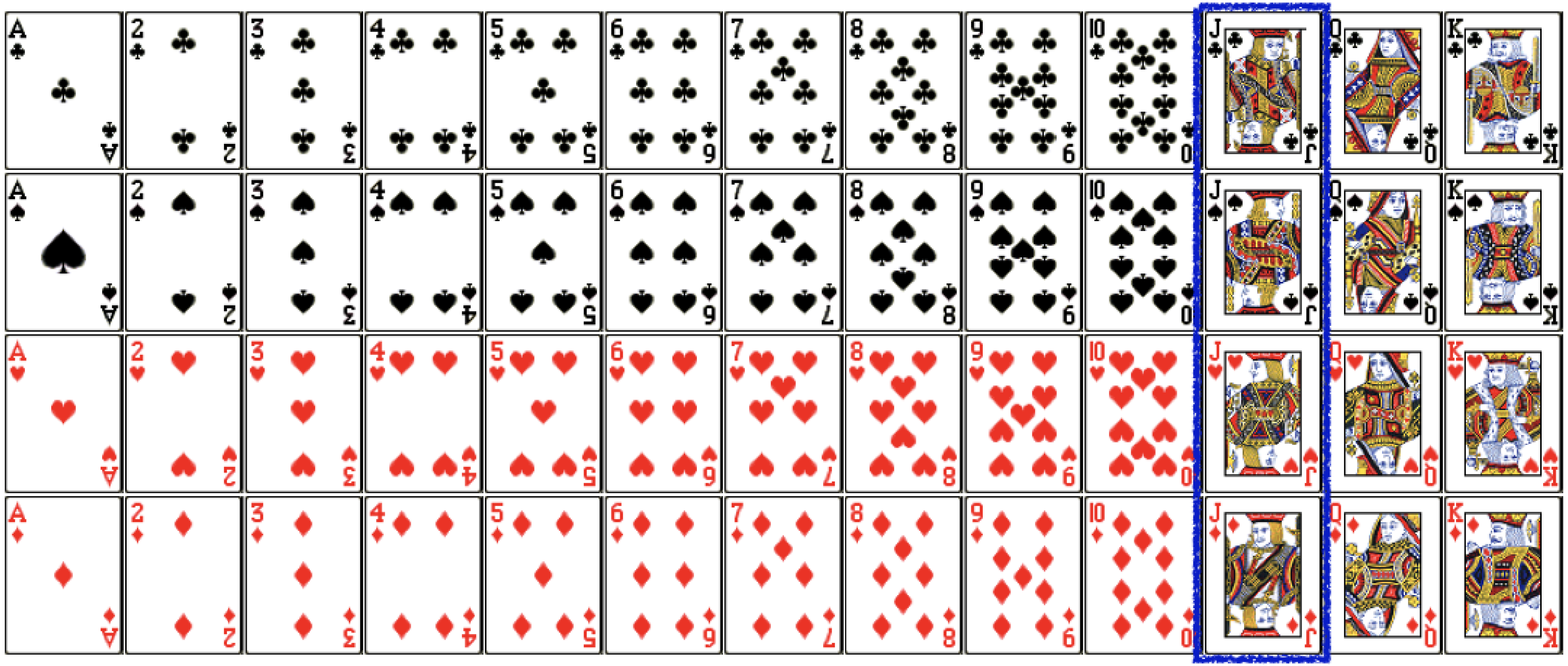
$$P(Red|Jack) = \frac{P(Jack\ and\ Red)}{P(Jack)}$$

Conditional probability (Cont.)



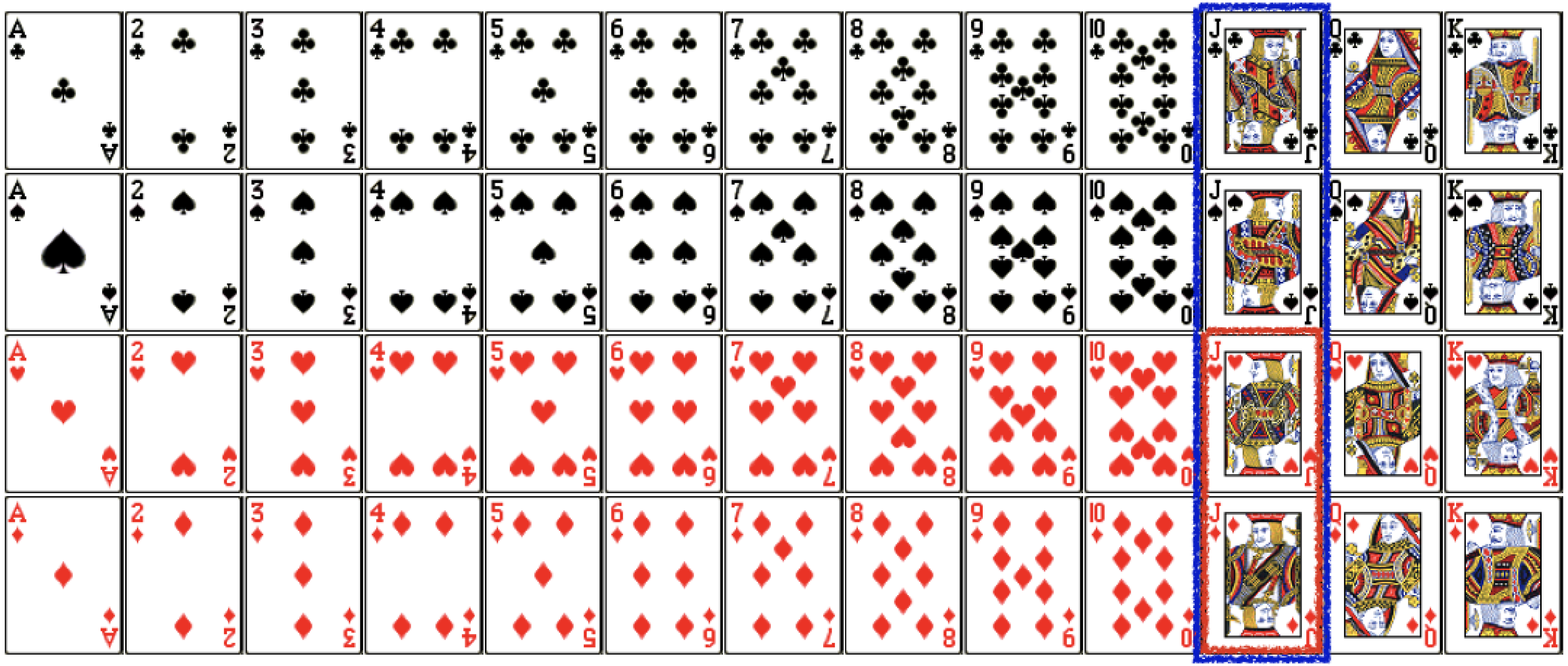
$$P(Red|Jack) = \frac{P(Jack\ and\ Red)}{P(Jack)}$$

Conditional probability (Cont.)



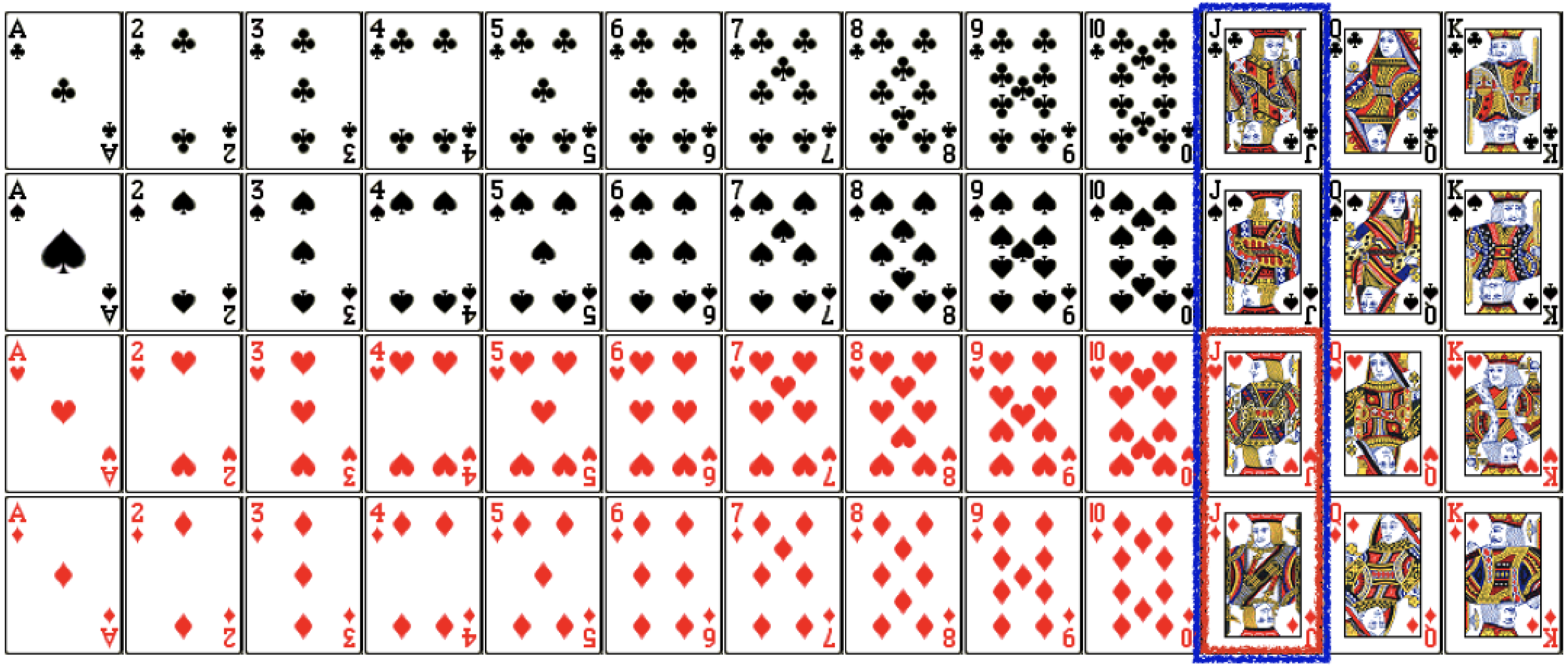
$$P(Red|Jack) = \frac{P(Jack\ and\ Red)}{P(Jack)} = \frac{X}{\frac{4}{52}}$$

Conditional probability (Cont.)



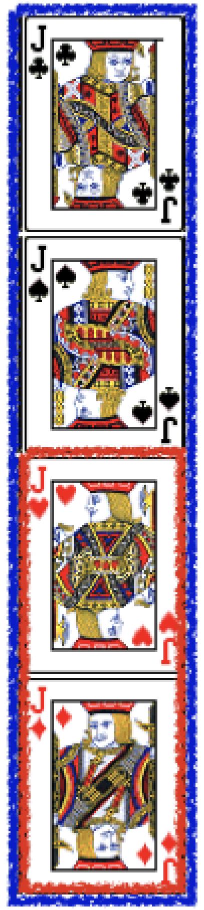
$$P(Red|Jack) = \frac{P(Jack\ and\ Red)}{P(Jack)} = \frac{\frac{2}{52}}{\frac{4}{52}}$$

Conditional probability (Cont.)



$$P(Red|Jack) = \frac{P(Jack\ and\ Red)}{P(Jack)} = \frac{\frac{2}{52}}{\frac{4}{52}} = \frac{2}{4} = \frac{1}{2}$$

Conditional probability (Cont.)



$$P(Red|Jack) = \frac{P(Jack\ and\ Red)}{P(Jack)} = \frac{\frac{2}{52}}{\frac{4}{52}} = \frac{2}{4} = \frac{1}{2}$$

P(Red | Jack) calculation in Python

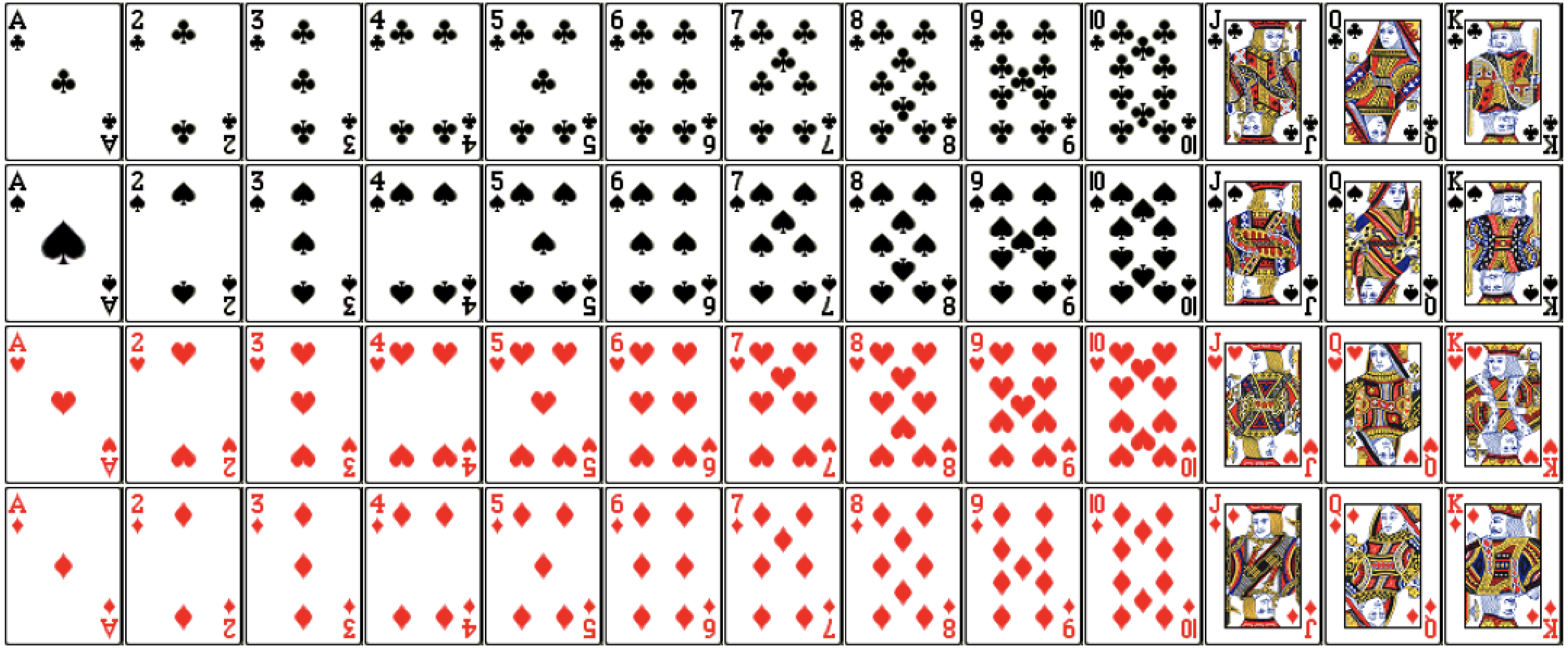
```
P_Jack = 4/52
P_Jack_n_Red = 2/52

P_Red_given_Jack = P_Jack_n_Red / P_Jack

print(P_Red_given_Jack)
```

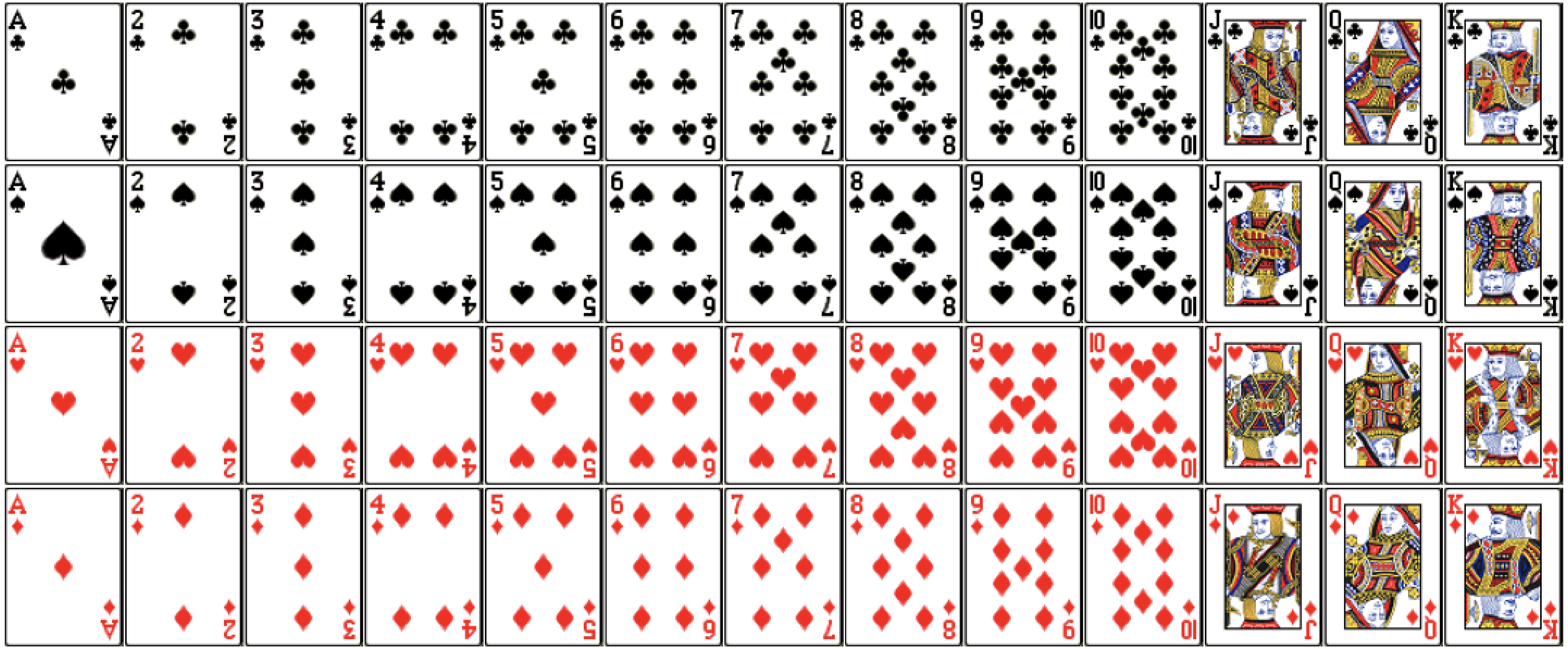
```
0.5
```

Conditional probability



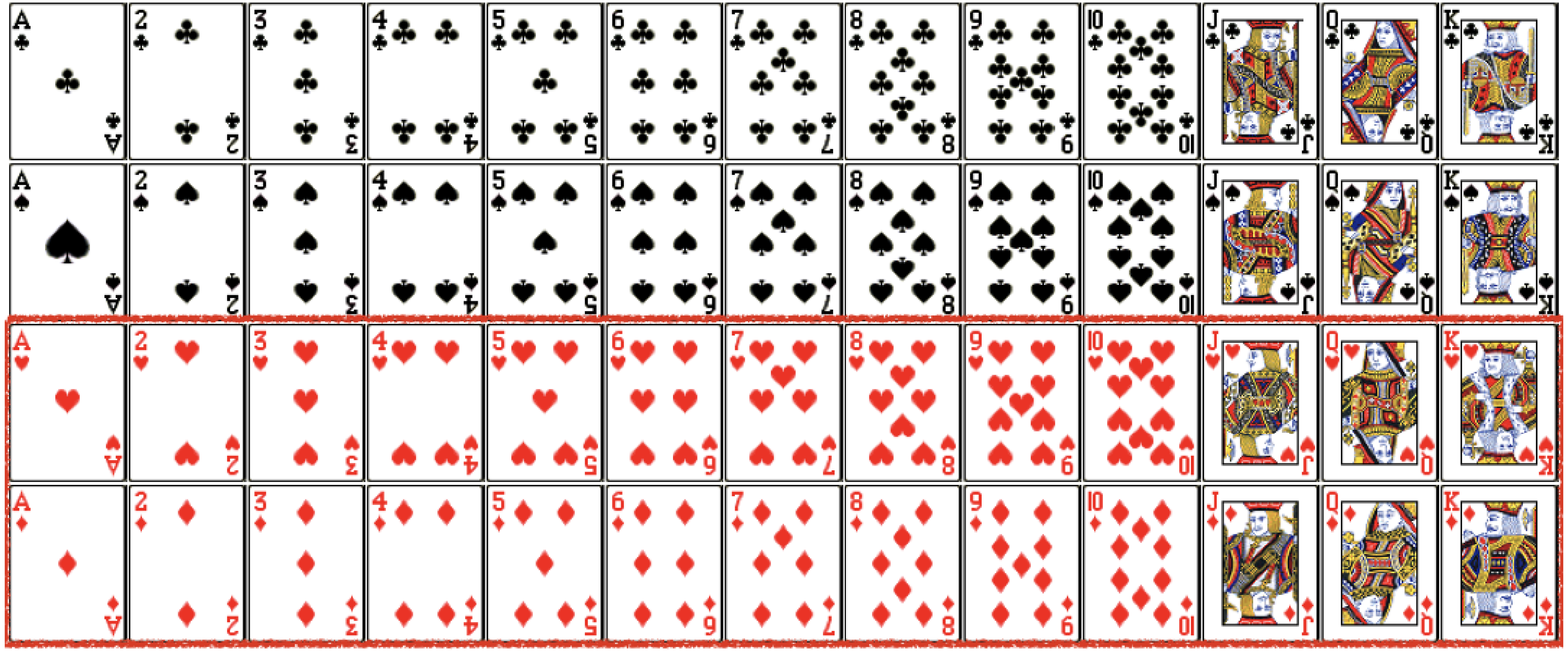
$$P(Jack|Red) = ?$$

Conditional probability (Cont.)



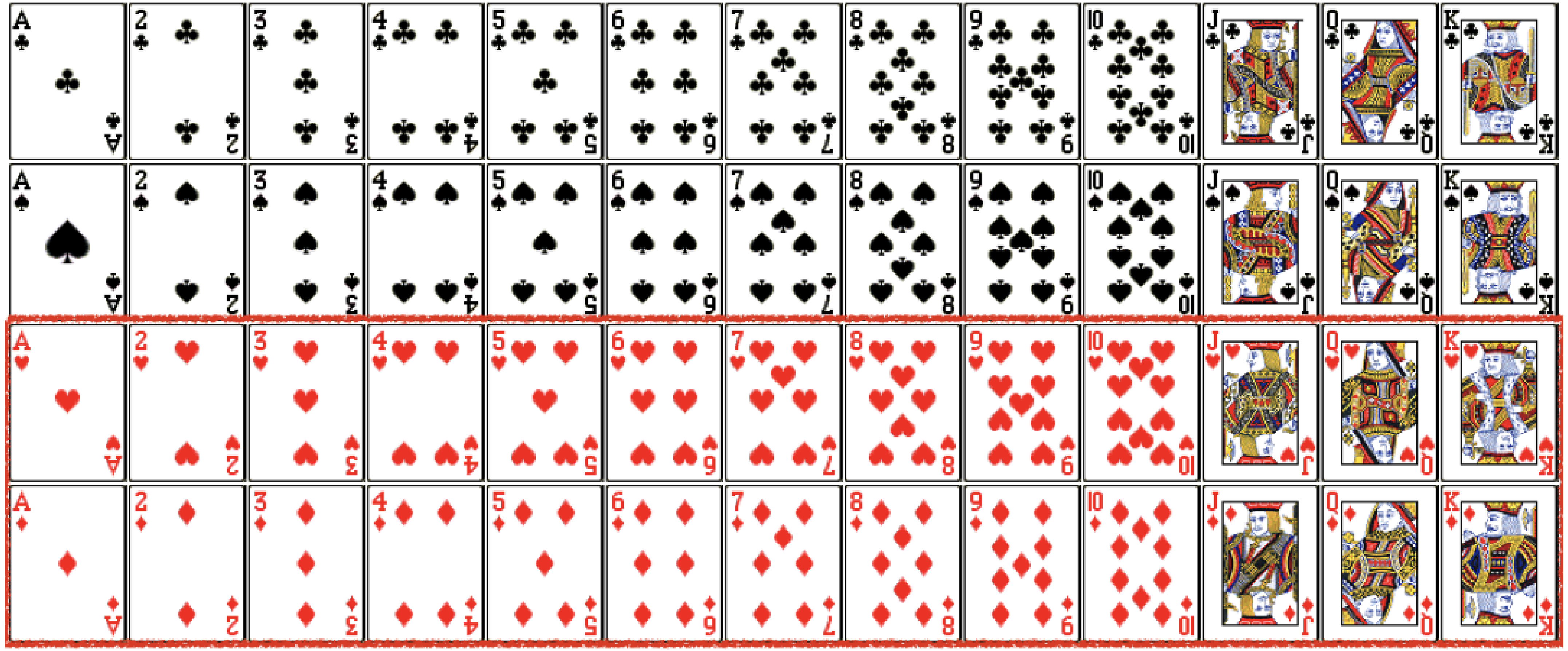
$$P(Jack|Red) = \frac{P(Red\ and\ Jack)}{P(Red)}$$

Conditional probability (Cont.)



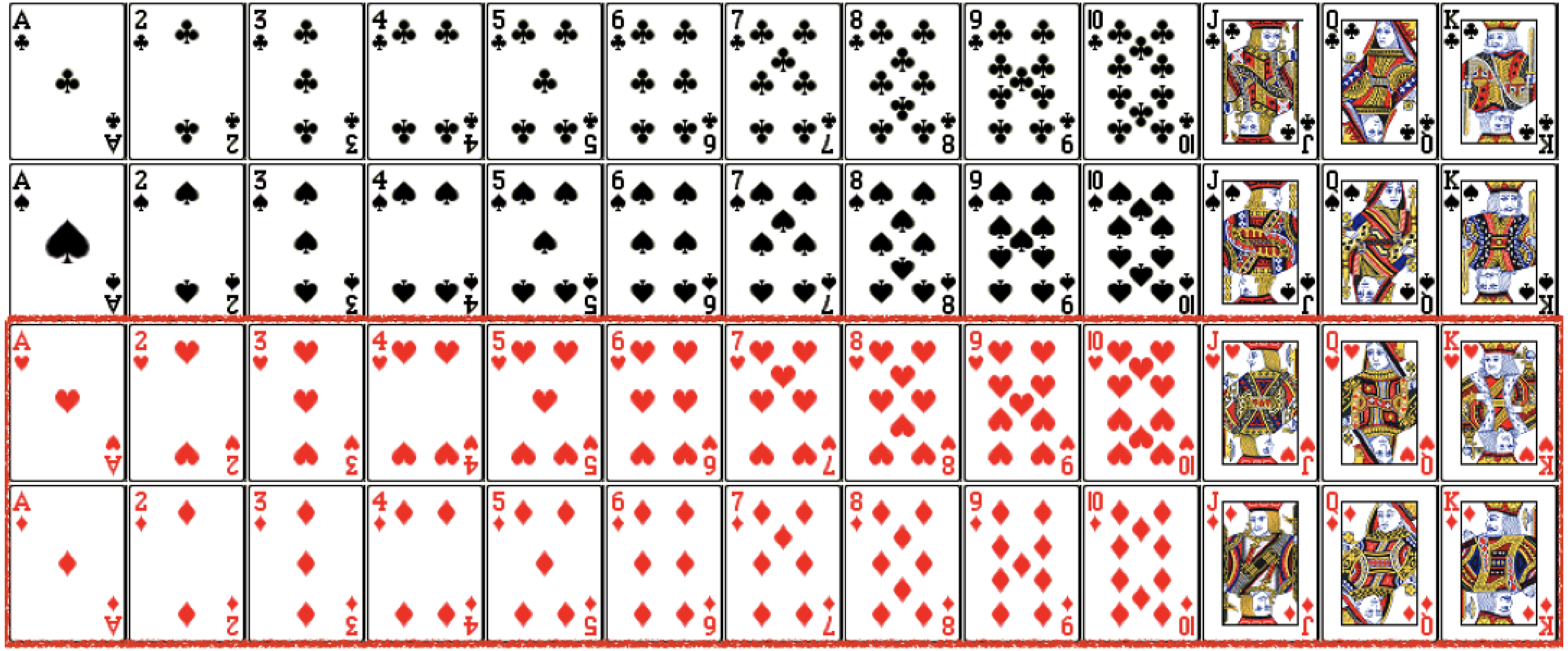
$$P(Jack|Red) = \frac{P(Red\ and\ Jack)}{P(Red)}$$

Conditional probability (Cont.)



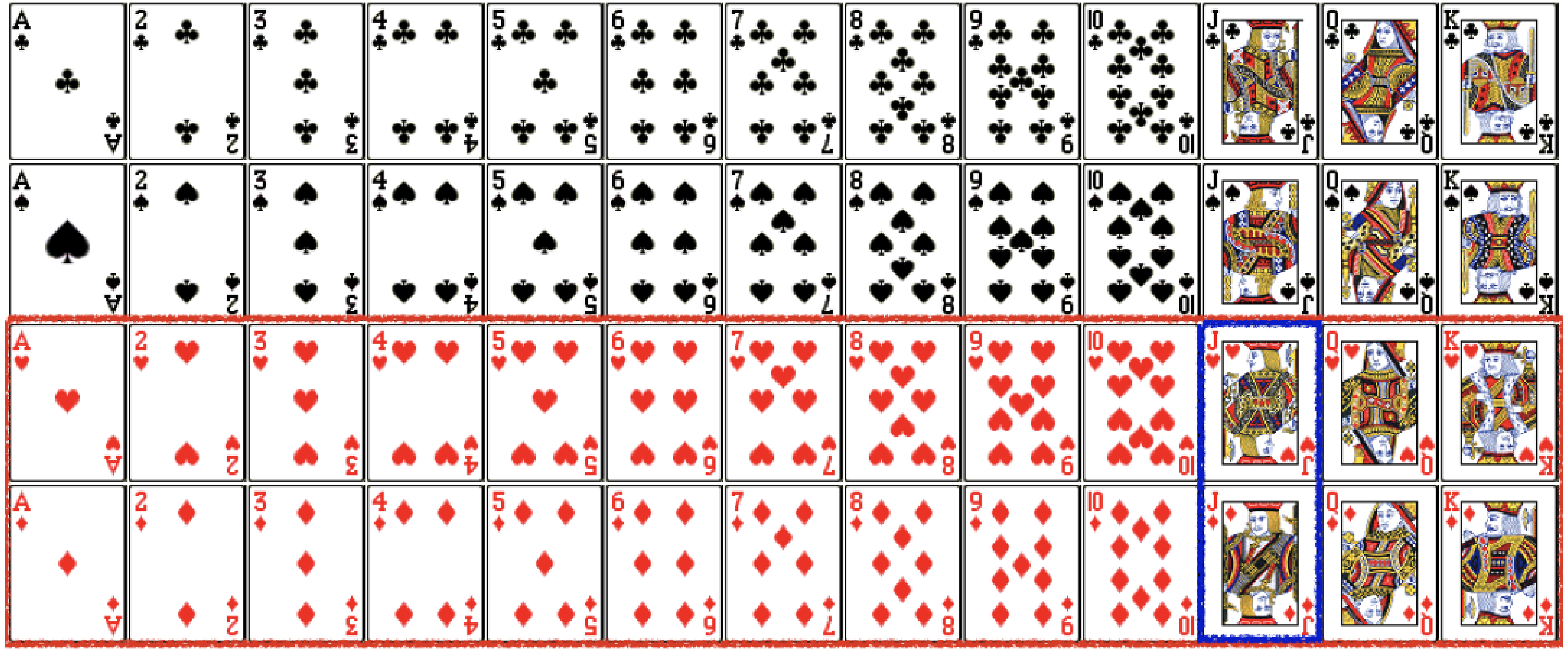
$$P(Jack|Red) = \frac{P(Red\ and\ Jack)}{P(Red)}$$

Conditional probability (Cont.)



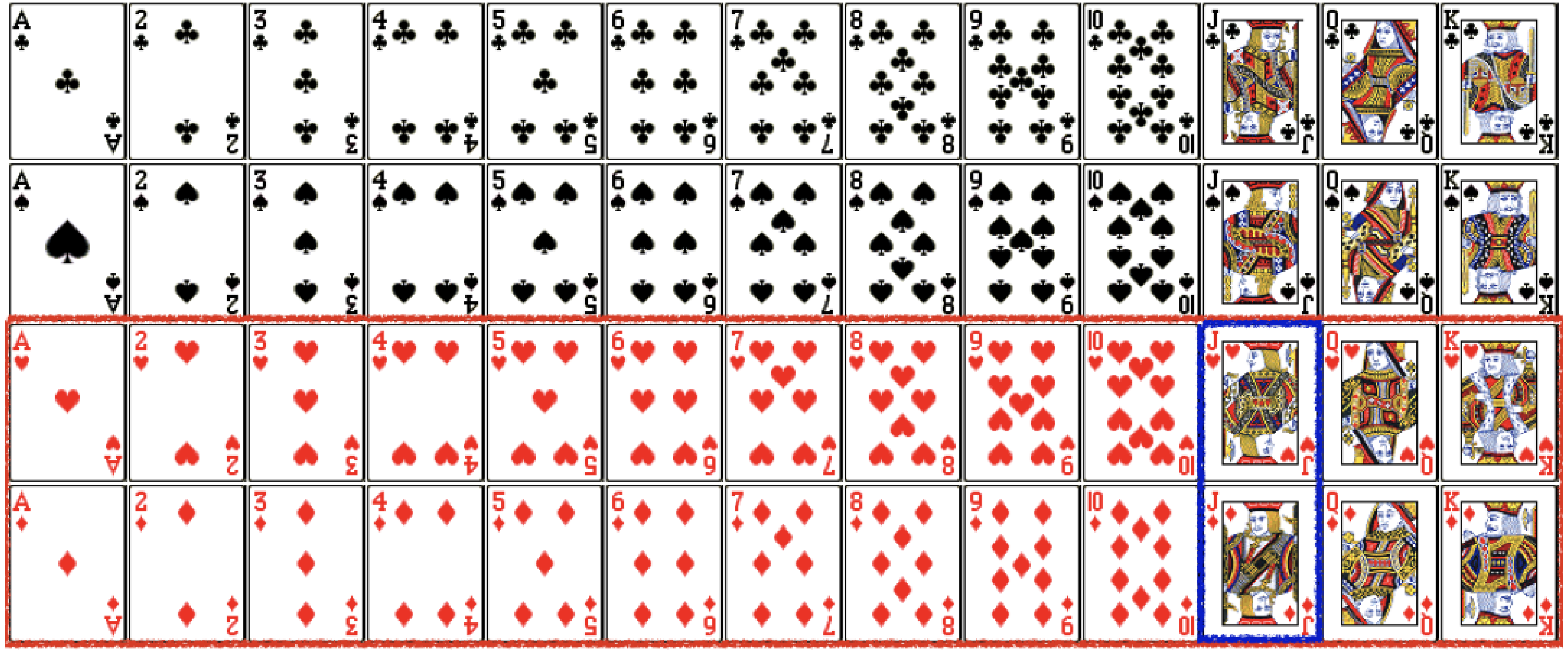
$$P(Jack|Red) = \frac{P(Red\ and\ Jack)}{P(Red)} = \frac{X}{\frac{26}{52}}$$

Conditional probability (Cont.)



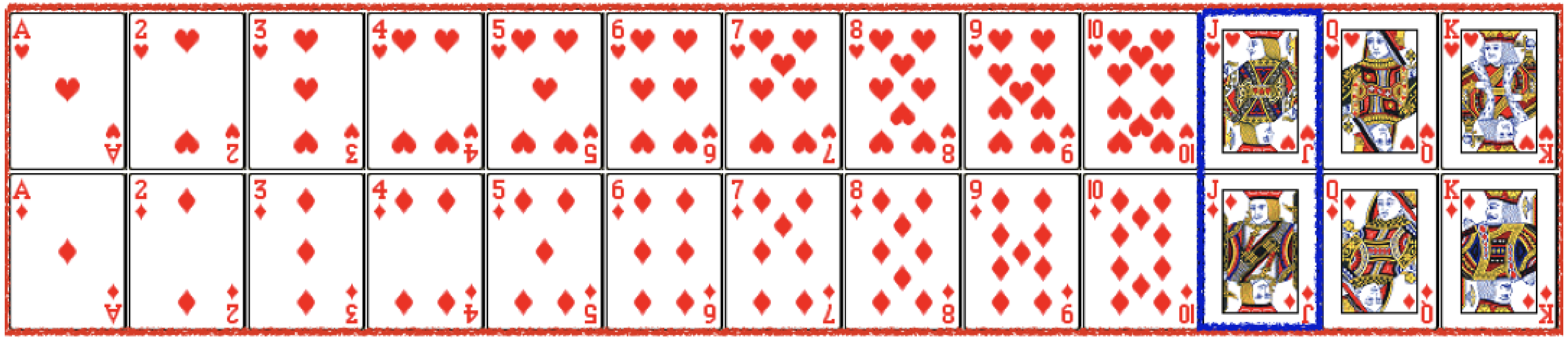
$$P(Jack|Red) = \frac{P(Red\ and\ Jack)}{P(Red)} = \frac{\frac{2}{52}}{\frac{26}{52}}$$

Conditional probability (Cont.)



$$P(Jack|Red) = \frac{P(Red\ and\ Jack)}{P(Red)} = \frac{\frac{2}{52}}{\frac{26}{52}} = \frac{2}{26} = \frac{1}{13}$$

Conditional probability (Cont.)



$$P(Jack|Red) = \frac{P(Red\ and\ Jack)}{P(Red)} = \frac{\frac{2}{52}}{\frac{26}{52}} = \frac{2}{26} = \frac{1}{13}$$

P(Jack | Red) calculation in Python

```
P_Red = 26/52
P_Red_n_Jack = 2/52

P_Jack_given_Red = P_Red_n_Jack / P_Red

print(P_of_Jack_given_Red)
```

```
0.0769230769231
```

Let's condition events to calculate probabilities

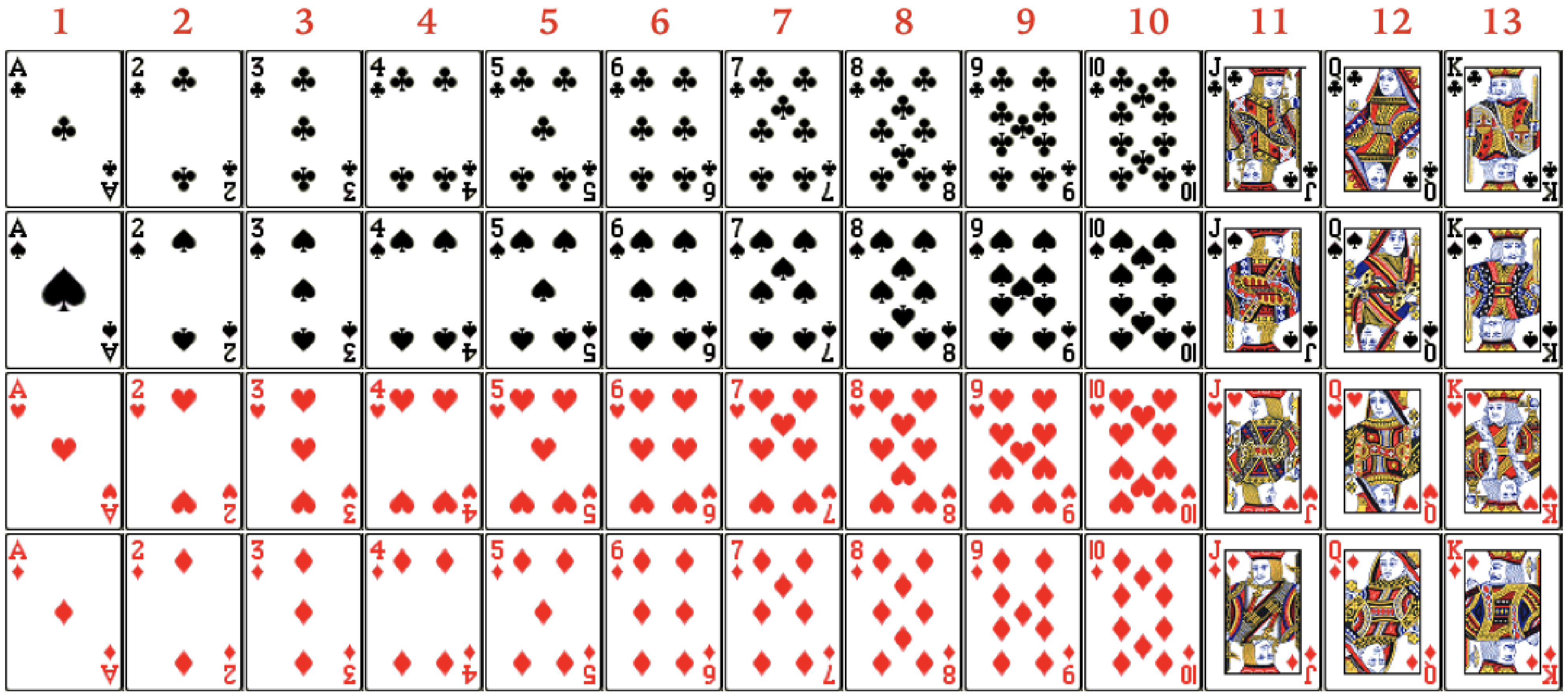
FOUNDATIONS OF PROBABILITY IN PYTHON

Total probability law

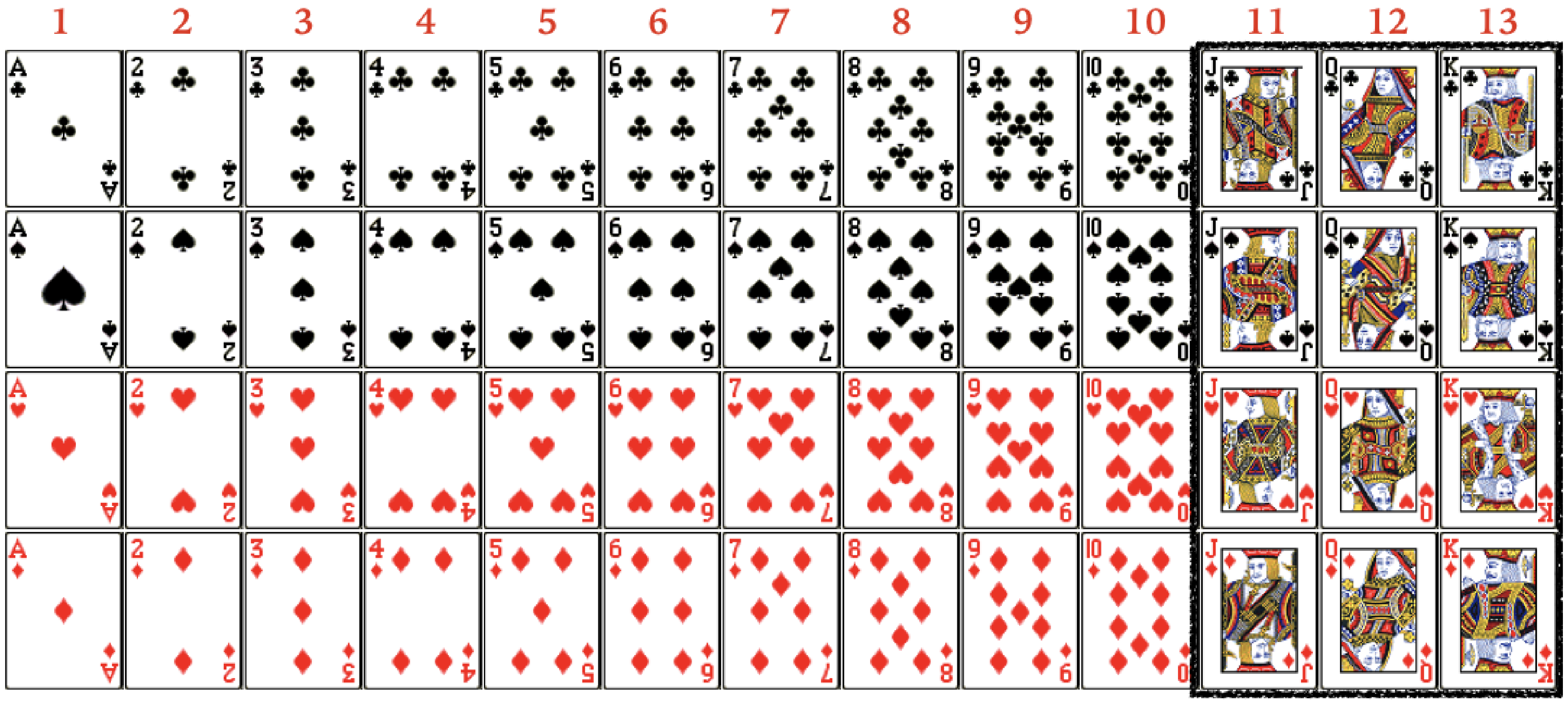
FOUNDATIONS OF PROBABILITY IN PYTHON

Alexander A. Ramírez M.
CEO @ Synergy Vision

Deck of cards example

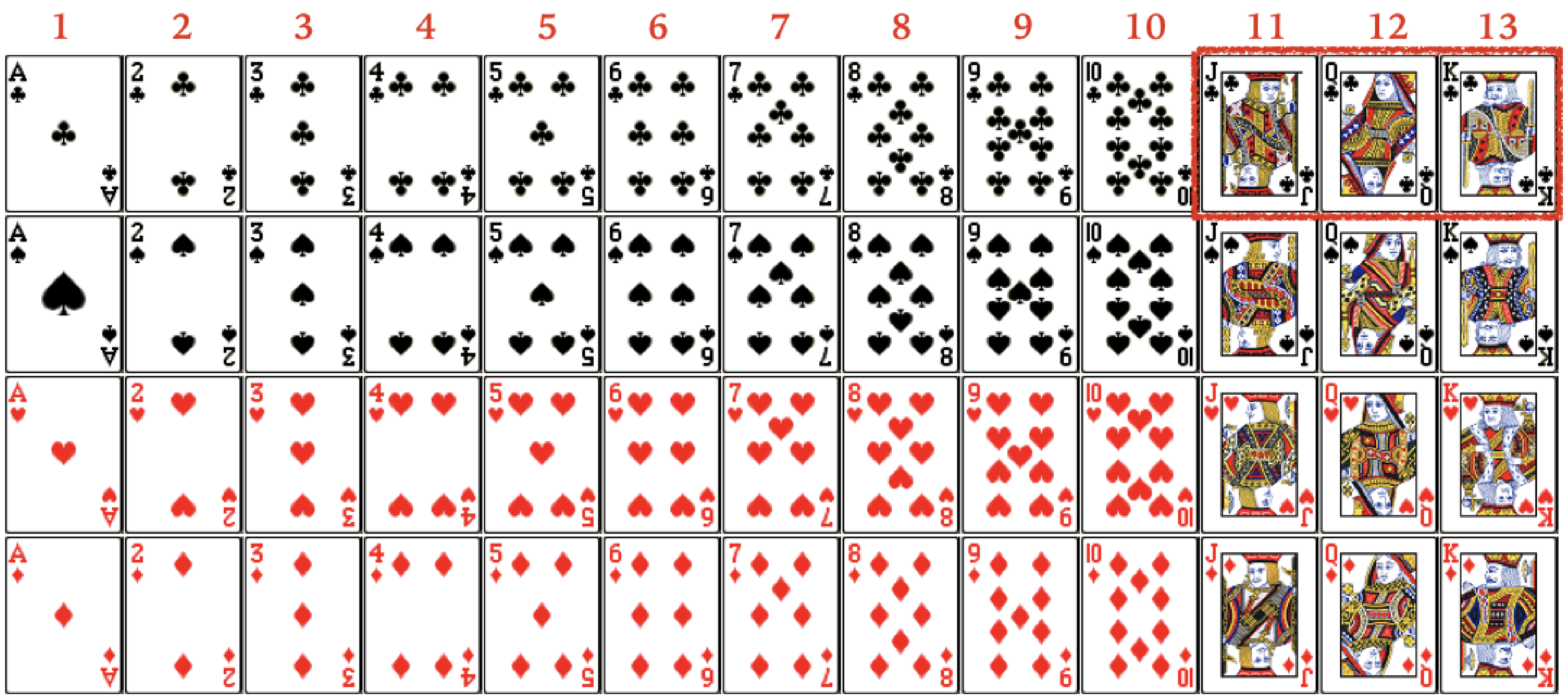


Deck of cards example (Cont.)



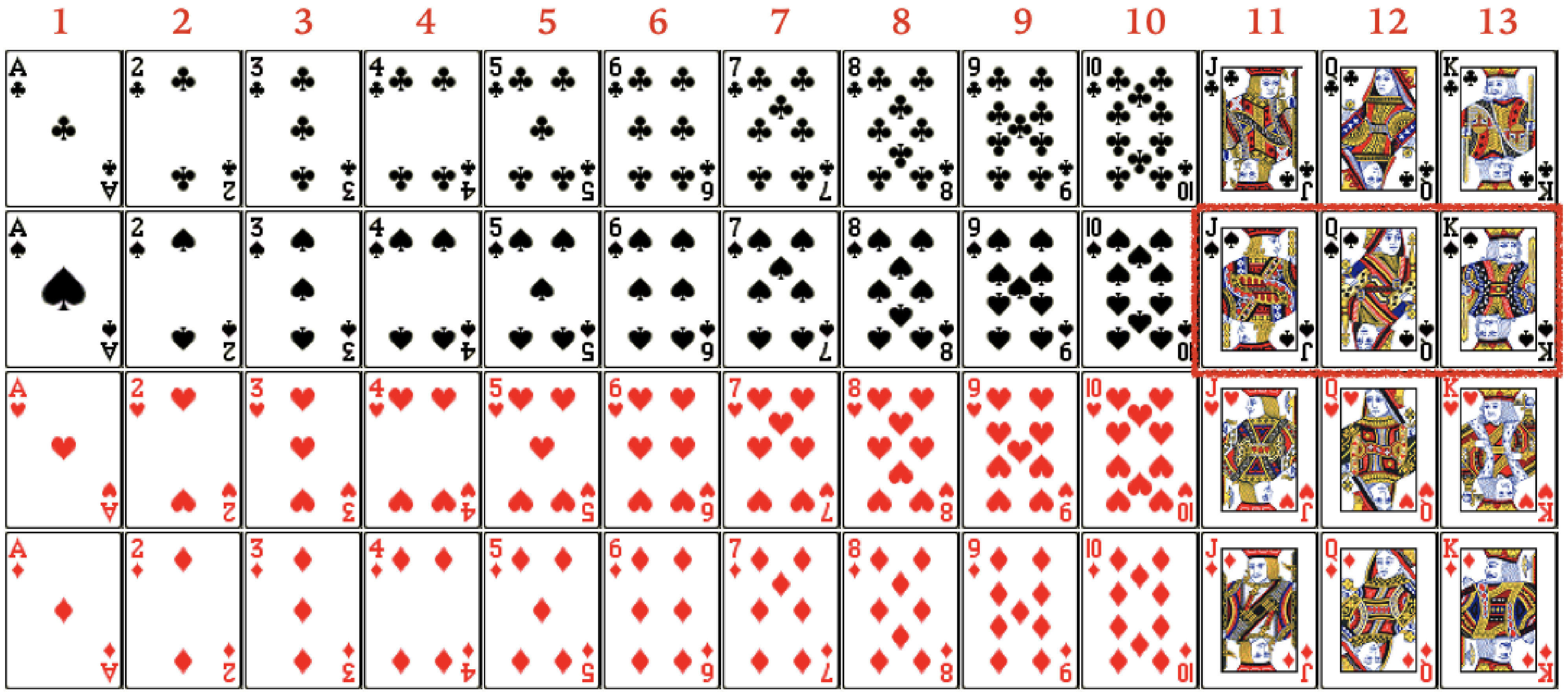
$$P(Face\ card) = ?$$

Deck of cards example (Cont.)



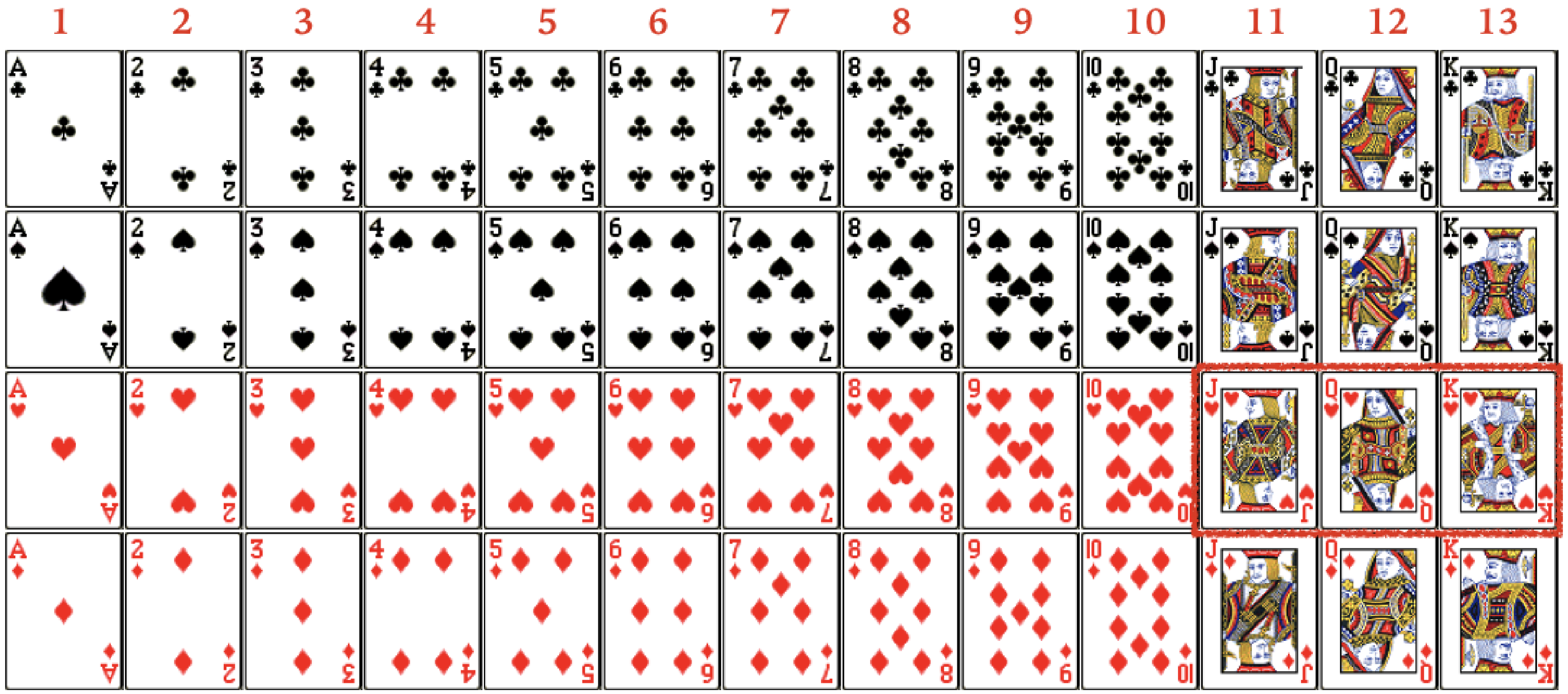
$$P(Face\ card) = P(Club\ and\ Face\ card) + \ldots$$

Deck of cards example (Cont.)



$$P(Face\ card) = P(Club\ and\ Face\ card) + P(Spade\ and\ Face\ card) + \ldots$$

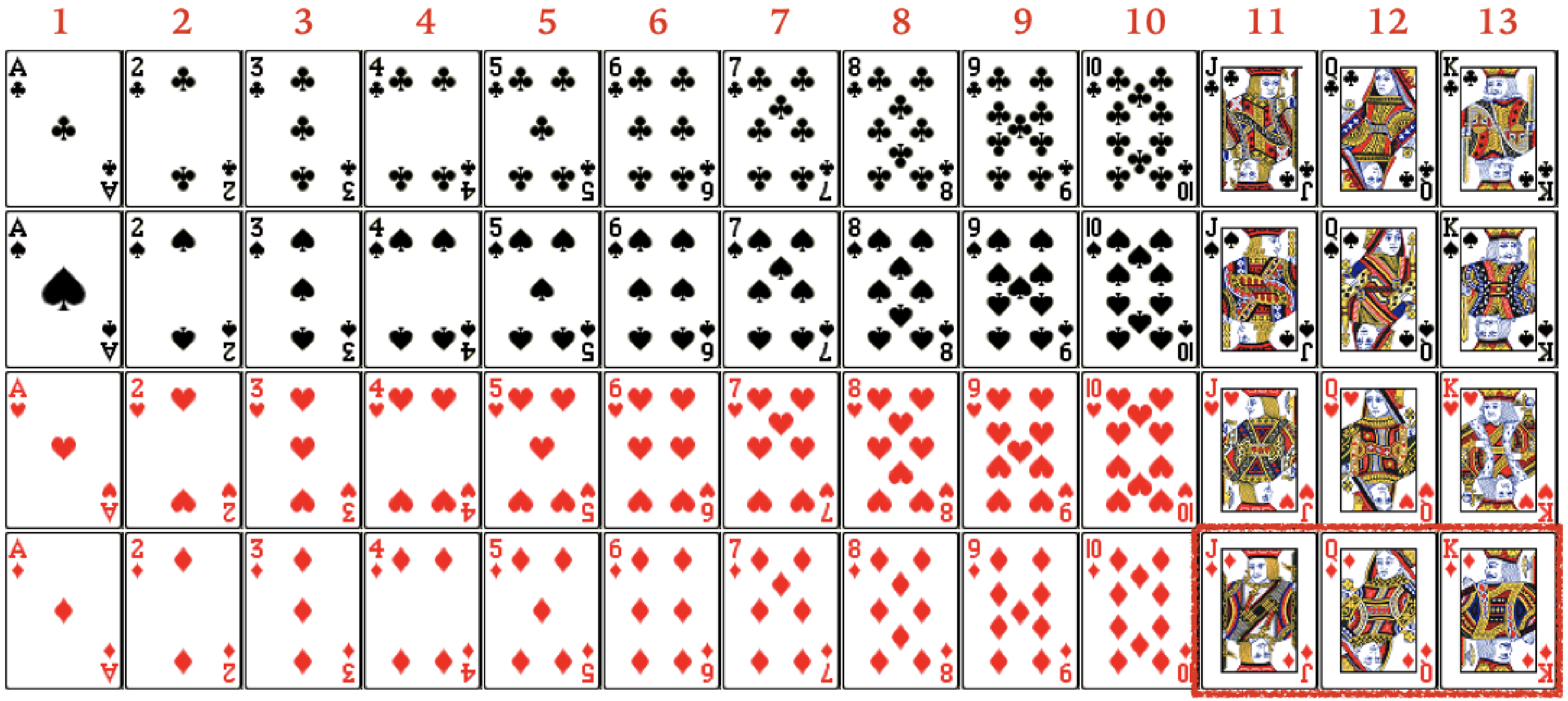
Deck of cards example (Cont.)



$$P(Face\ card) = P(Club\ and\ Face\ card) + P(Spade\ and\ Face\ card) +$$

$$\color{red}{P(Heart\ and\ Face\ card)} + \ldots$$

Deck of cards example (Cont.)



$$P(Face\ card) = P(Club\ and\ Face\ card) + P(Spade\ and\ Face\ card) +$$

$$P(Heart\ and\ Face\ card) + \color{red}{P(Diamond\ and\ Face\ card)}$$

Face card example in Python

Total probability calculation, FC is Face card in the code

```
P_Club_n_FC = 3/52
P_Spade_n_FC = 3/52
P_Heart_n_FC = 3/52
P_Diamond_n_FC = 3/52
```

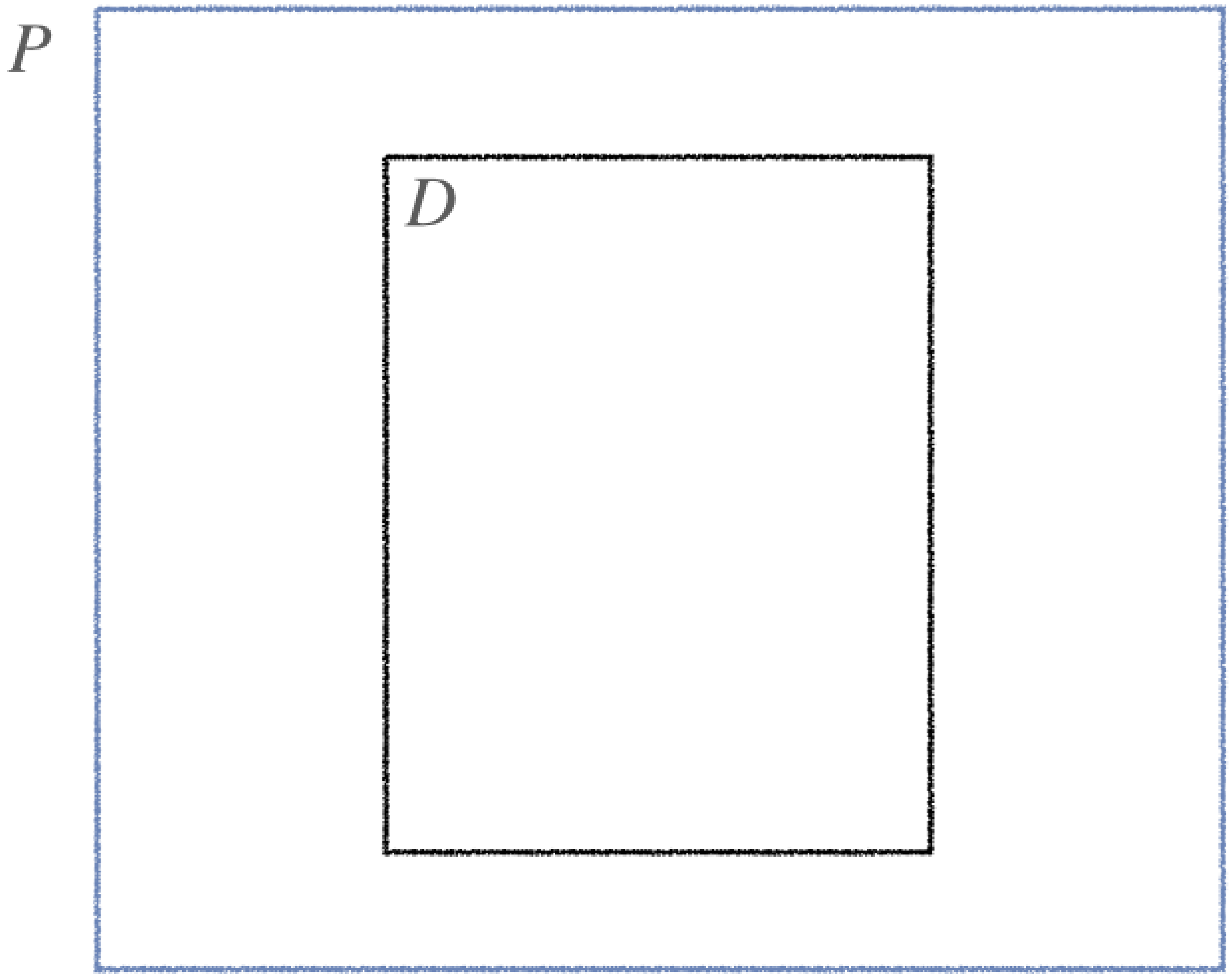
```
P_Face_card = P_Club_n_FC + P_Spade_n_FC + P_Heart_n_FC + P_Diamond_n_FC
print(P_Face_card)
```

The probability of a face card is:

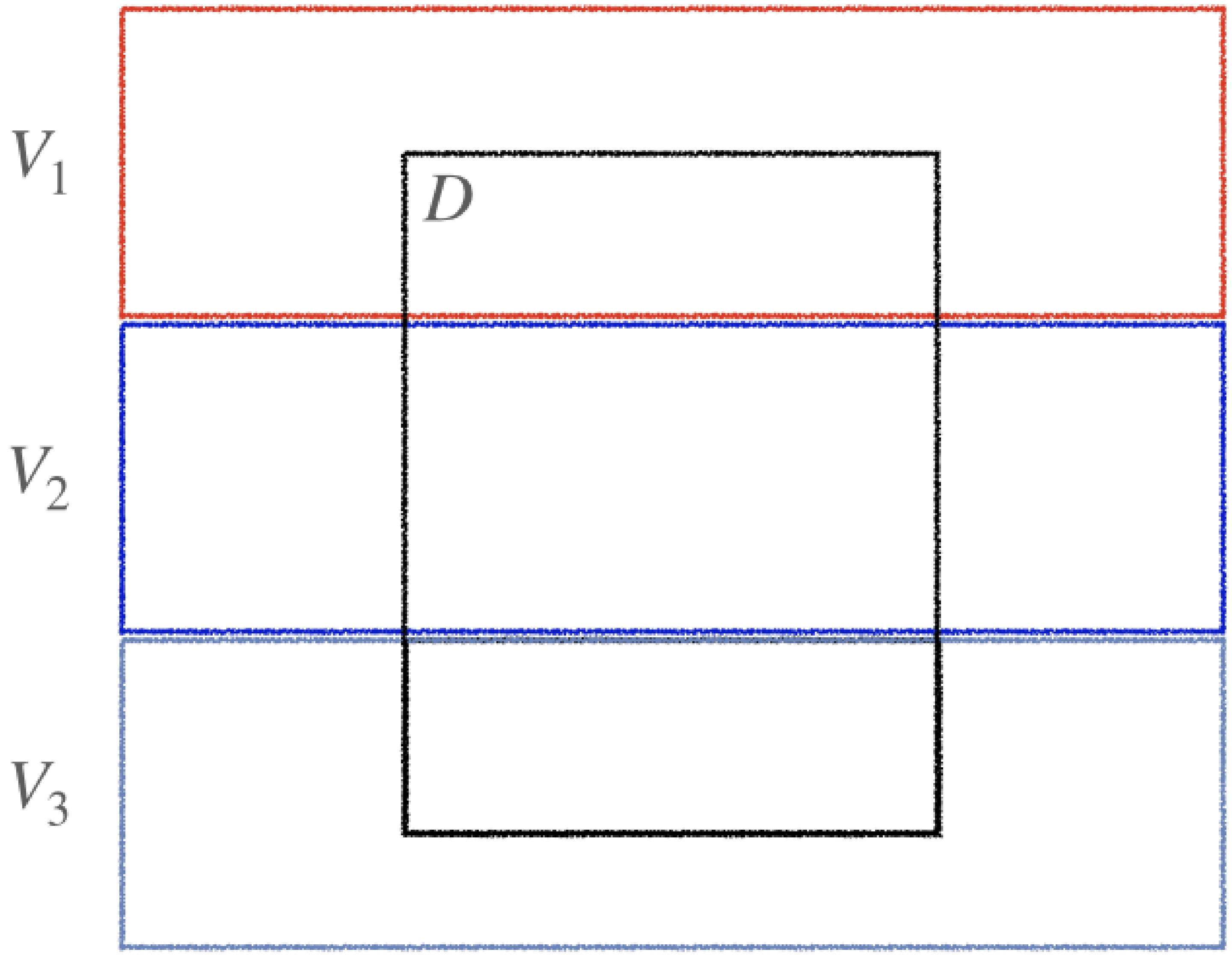
```
0.230769230769
```

Total probability

Total probability (Cont.)

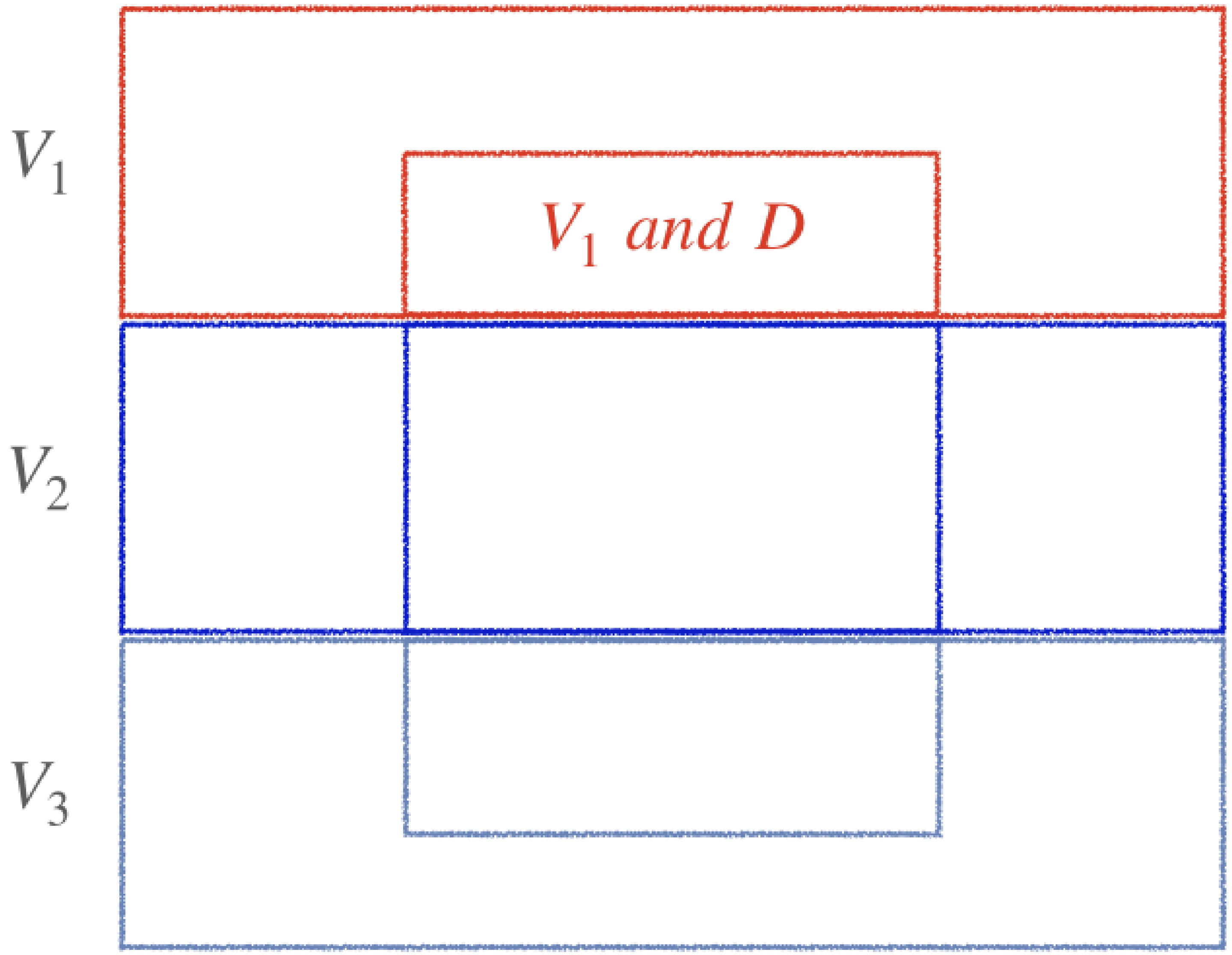


Total probability (Cont.)



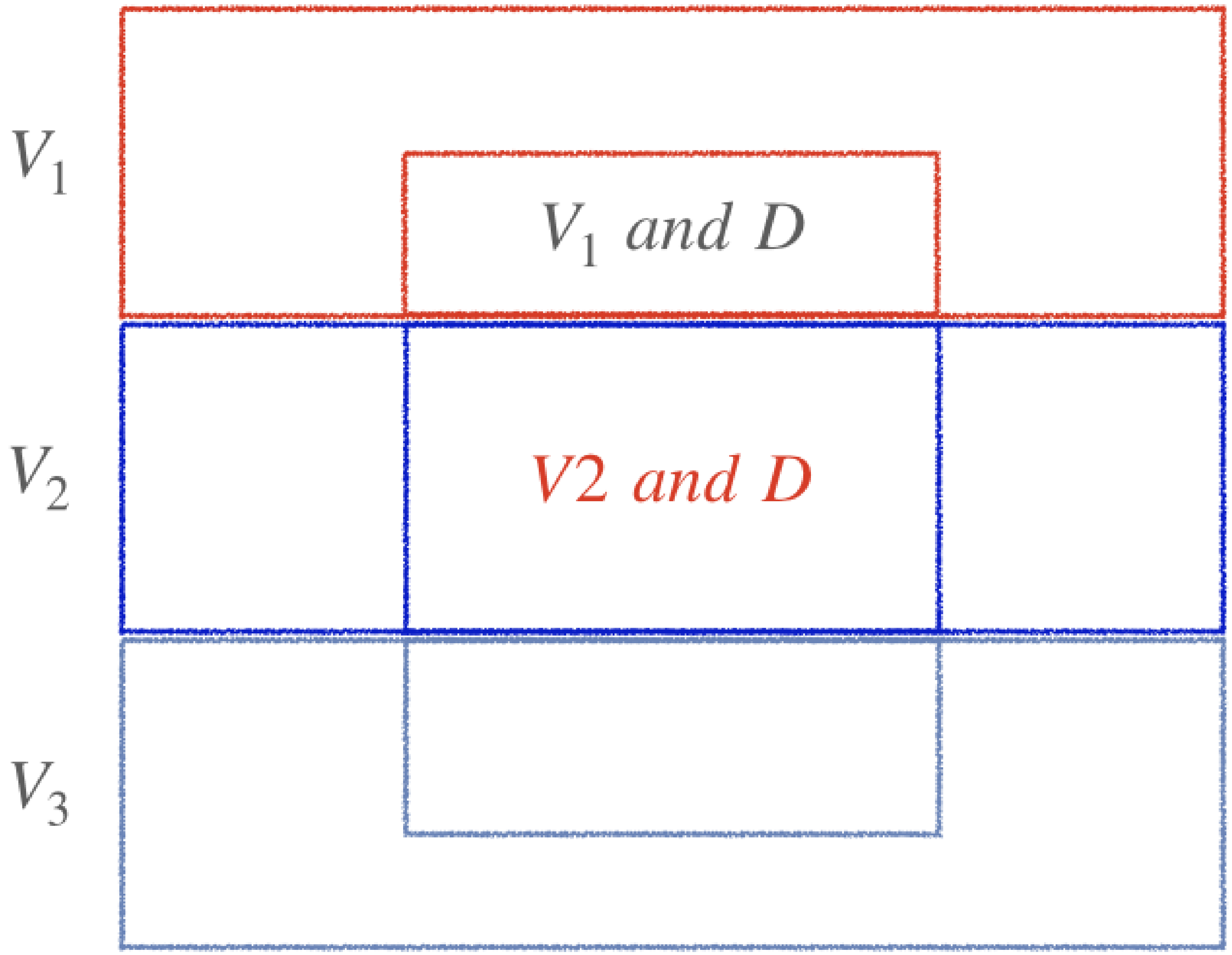
$$P(D) = ?$$

Total probability (Cont.)



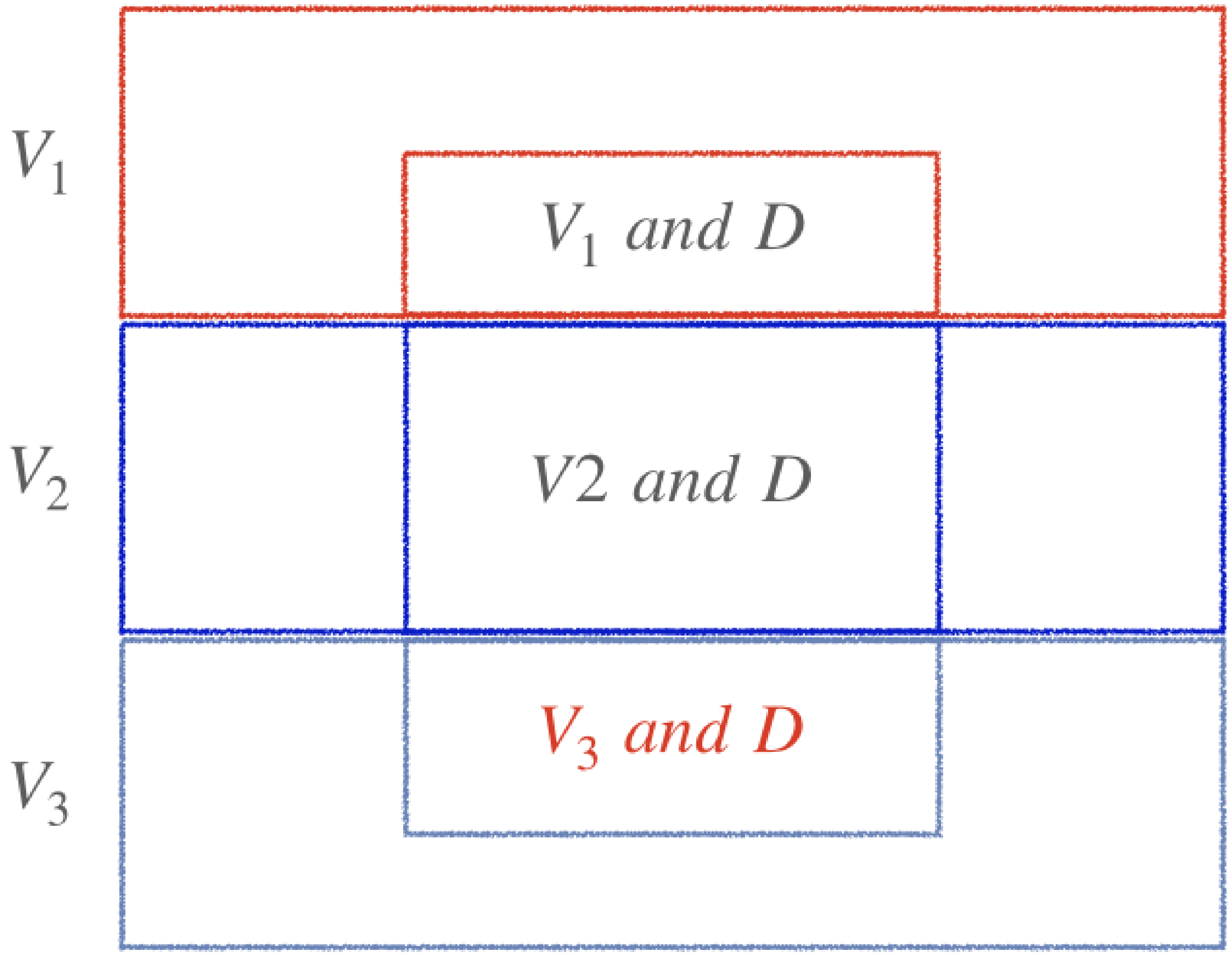
$$P(D) = P(V1 \text{ and } D) + \ldots$$

Total probability (Cont.)



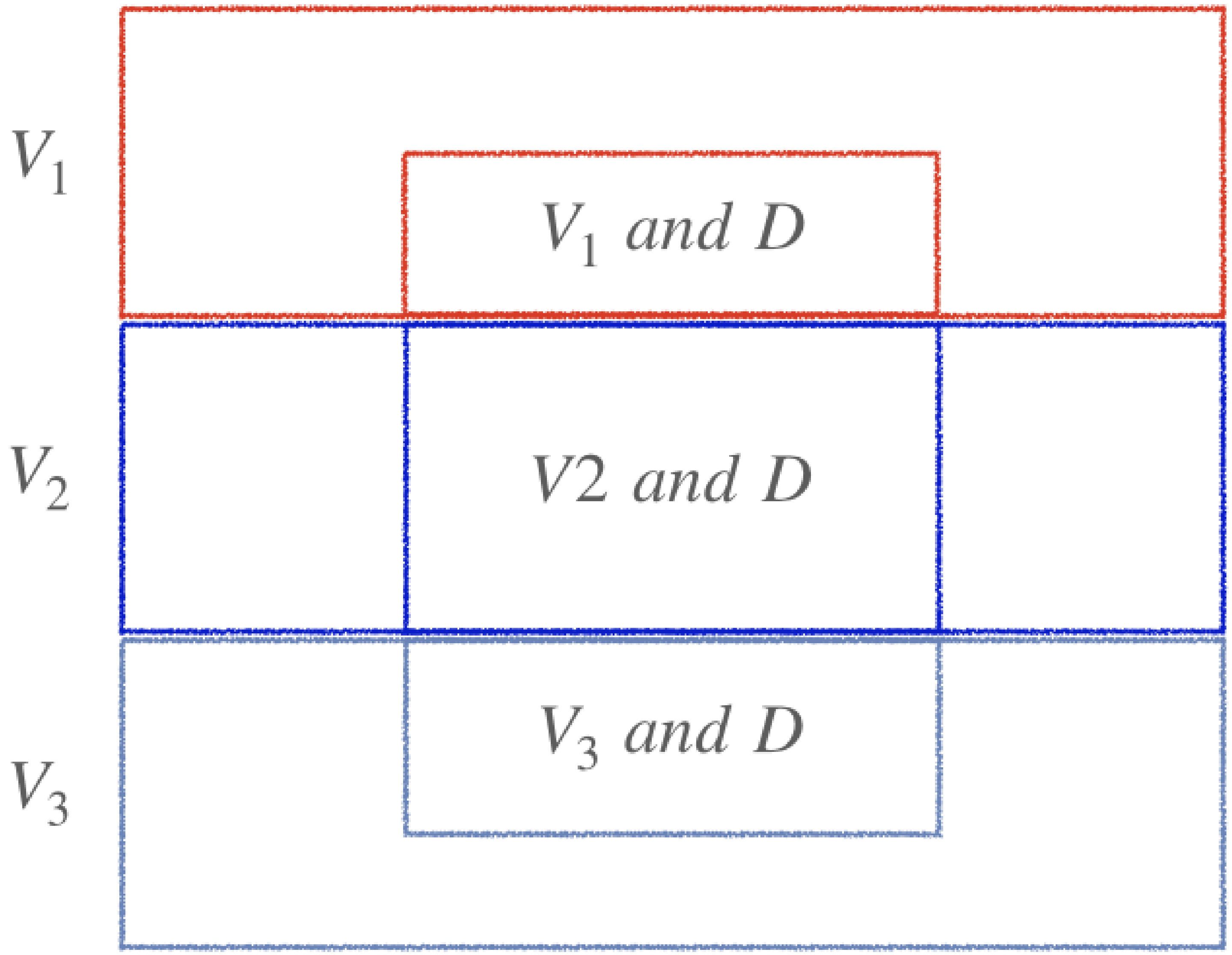
$$P(D) = P(V1\ and\ D) + P(V2\ and\ D) + \ldots$$

Total probability (Cont.)



$$P(D) = P(V1 \; and \; D) + P(V2 \; and \; D) + \color{red}{P(V3 \; and \; D)}$$

Total probability (Cont.)



$$P(D) = P(V1)P(D|V1) + P(V2)P(D|V2) + P(V3)P(D|V3)$$

Damaged parts example in Python

A certain electronic part is manufactured by three different vendors, V1, V2, and V3.

Half of the parts are produced by V1, and V2 and V3 each produce 25%. The probability of a part being damaged given that it was produced by V1 is 1%, while it's 2% for V2 and 3% for V3.

- What is the probability of a part being damaged?

Damaged parts example in Python (Cont.)

- What is the probability of a part being damaged?

```
P_V1 = 0.5
P_V2 = 0.25
P_V3 = 0.25
```

```
P_D_g_V1 = 0.01
P_D_g_V2 = 0.02
P_D_g_V3 = 0.03
```

Damaged parts example in Python (Cont.)

We apply the total probability formula

```
P_Damaged = P_V1 * P_D_g_V1 + P_V2 * P_D_g_V2 + P_V3 * P_D_g_V3
print(P_Damaged)
```

The probability of being damaged is:

```
0.0175
```

Let's start using the total probability law

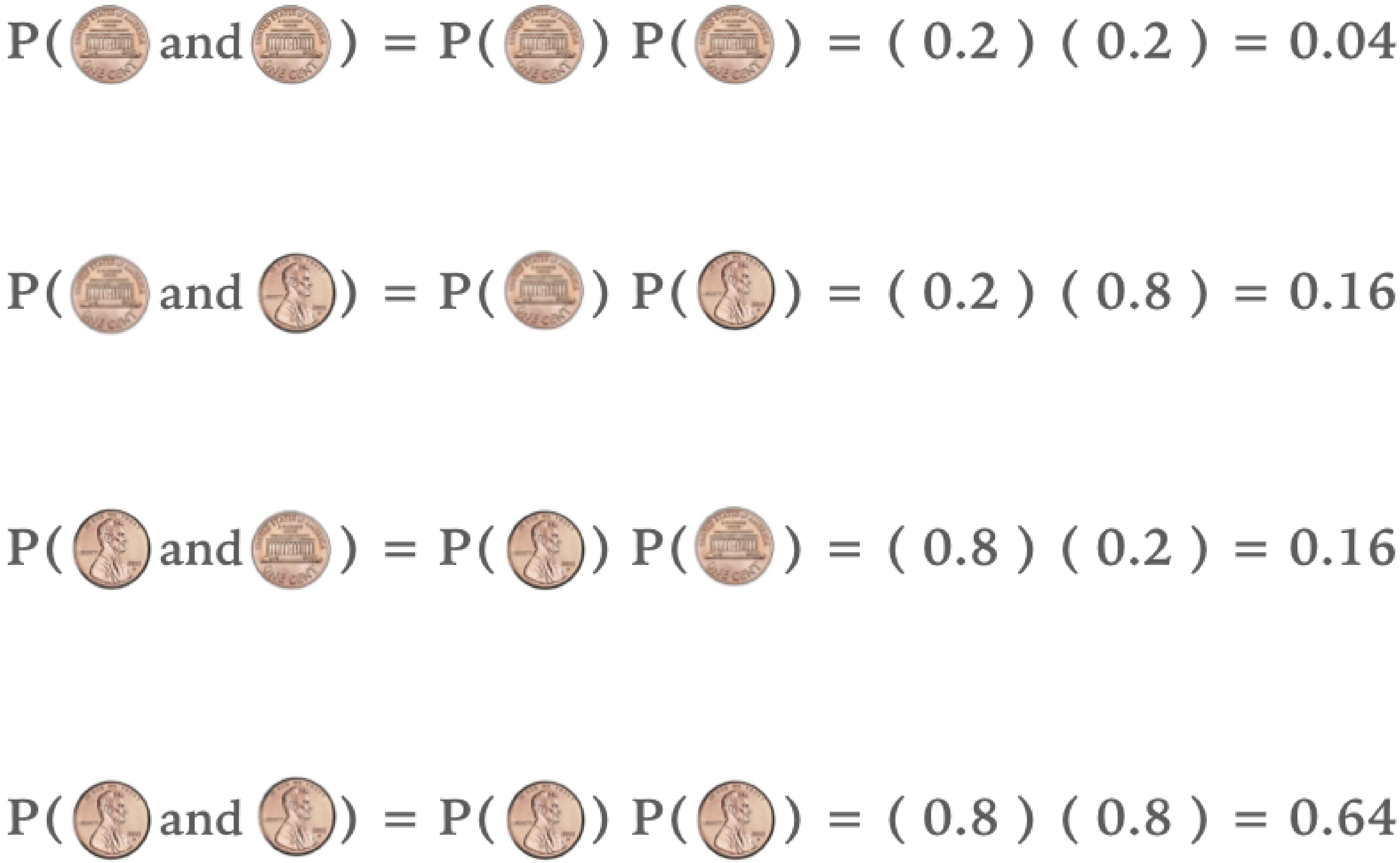
FOUNDATIONS OF PROBABILITY IN PYTHON

Bayes' rule

FOUNDATIONS OF PROBABILITY IN PYTHON

Alexander A. Ramírez M.
CEO @ Synergy Vision

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



P(A and B) for independent events

$$P(A \; and \; B) = P(A)P(B)$$

P(A and B) for dependent events

$$P(A \; and \; B) = P(A)P(B)$$

$$P(A \; and \; B) = {\color{blue}P(A)P(B|A)}$$

P(A and B) for dependent events (Cont.)

$$P(A \; and \; B) = P(A)P(B)$$

$$P(A \; and \; B) = P(A)P(B|A)$$

$$P(B \; and \; A) = P(B)P(A|B)$$

P(A and B) is equal to P(B and A)

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

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

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

P(A and B) is equal to P(B and A) (Cont.)

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

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

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

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

Bayes' relation

$$P(A \; and \; B) = P(A)P(B)$$

$$P(A \; and \; B) = P(A)P(B|A)$$

$$P(B \; and \; A) = P(B)P(A|B)$$

$$P(A)P(B|A) = P(A \; and \; B) = P(B \; and \; A) = P(B)P(A|B)$$

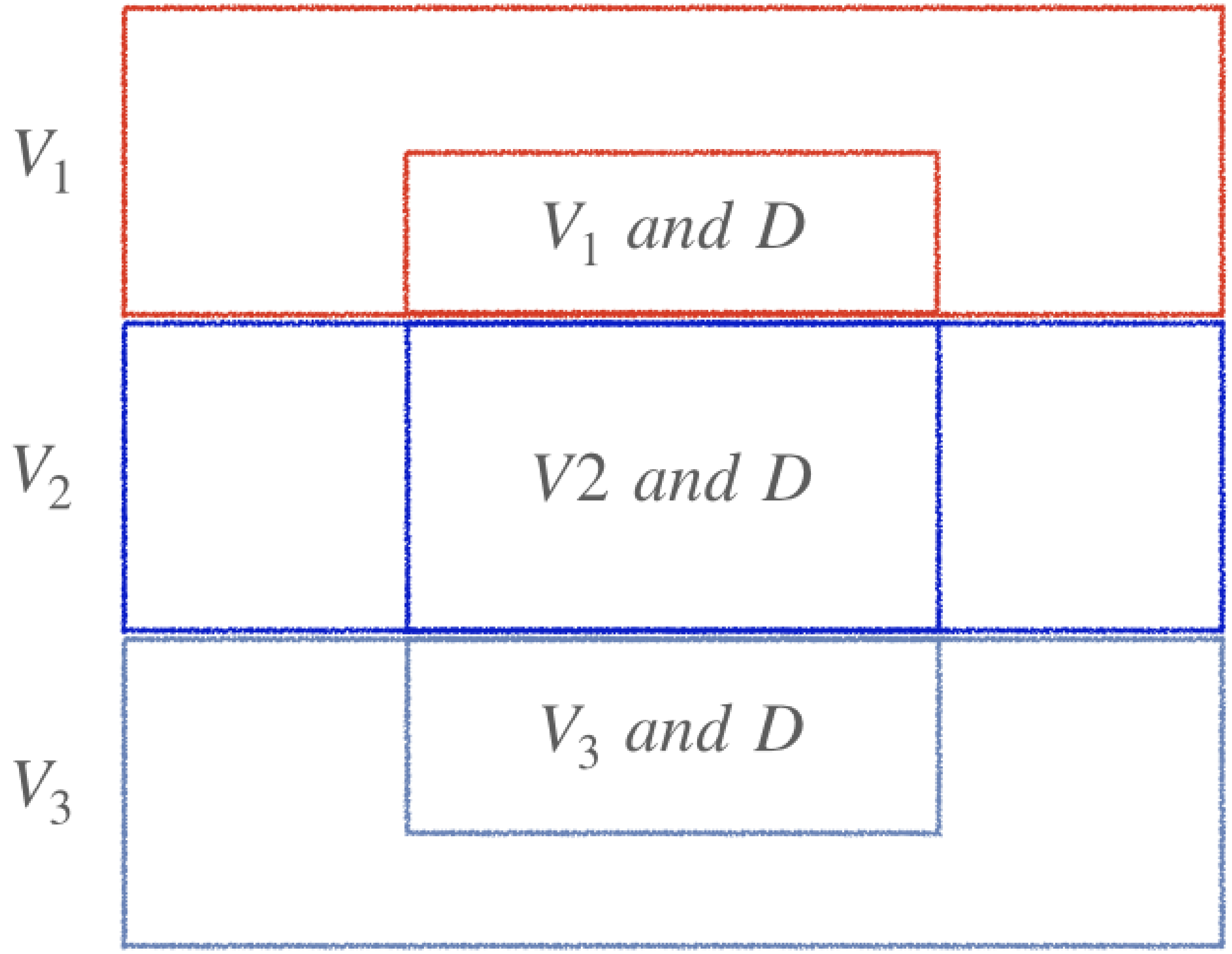
$$P(A)P(B|A) = P(B)P(A|B)$$

Bayes' rule

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

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

Total probability



$$P(D) = P(V_1 \text{ and } D) + P(V_2 \text{ and } D) + P(V_3 \text{ and } D)$$

Total probability (Cont.)

$$P(D) = P(V_1 \; and \; D) + P(V_2 \; and \; D) + P(V_3 \; and \; D)$$

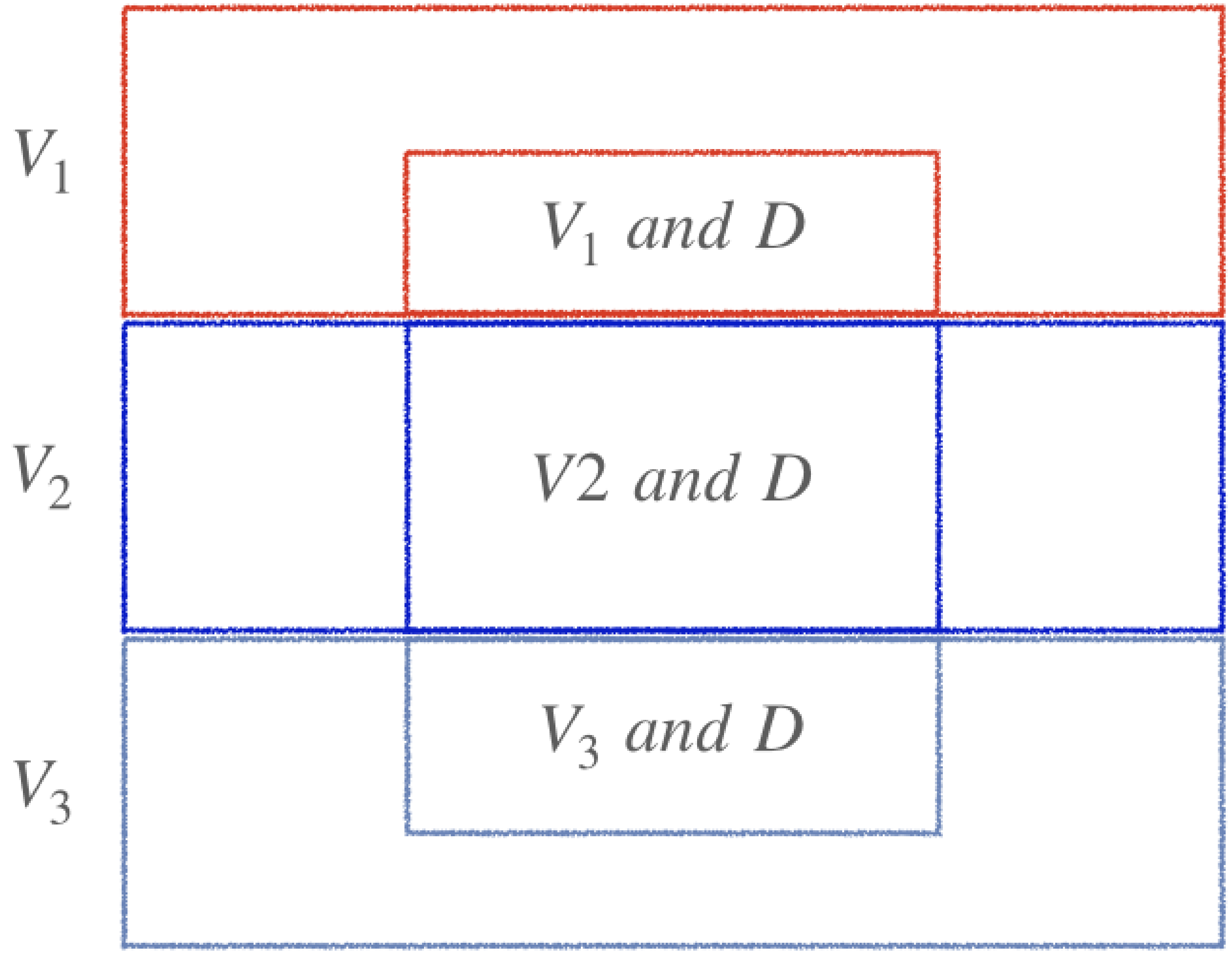
$$P(V_1 \; and \; D) = P(V_1)P(D|V_1)$$

$$P(V_2 \; and \; D) = P(V_2)P(D|V_2)$$

$$P(V_3 \; and \; D) = P(V_3)P(D|V_3)$$

$$P(D) = P(V_1)P(D|V_1) + P(V_2)P(D|V_2) + P(V_3)P(D|V_3)$$

Total probability (Cont.)



$$P(D) = P(V_1)P(D|V_1) + P(V_2)P(D|V_2) + P(V_3)P(D|V_3)$$

Bayes' formula

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

Bayes' formula (Cont.)

- Bayes' formula:

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

- The probability of a part being from vendor i, given that it is damaged:

$$P(V_i|D) = \frac{P(V_i)P(D|V_i)}{P(D)}$$

Bayes' formula (Cont.)

- Bayes' formula:

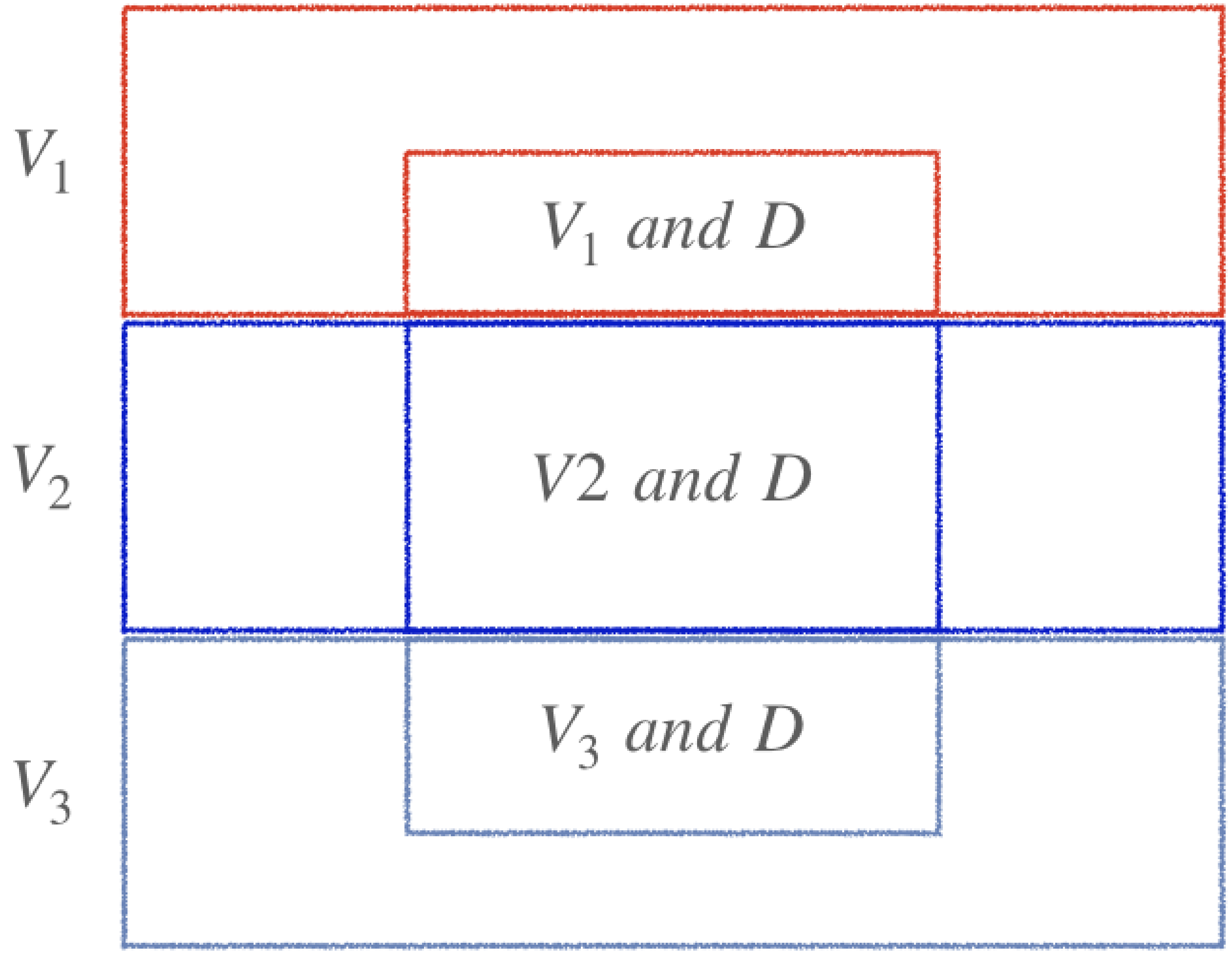
$$P(A|B) = \frac{P(A)P(B|A)}{P(B)}$$

- The probability of a part being from vendor i, given that it is damaged:

$$P(V_i|D) = \frac{P(V_i)P(D|V_i)}{P(D)}$$

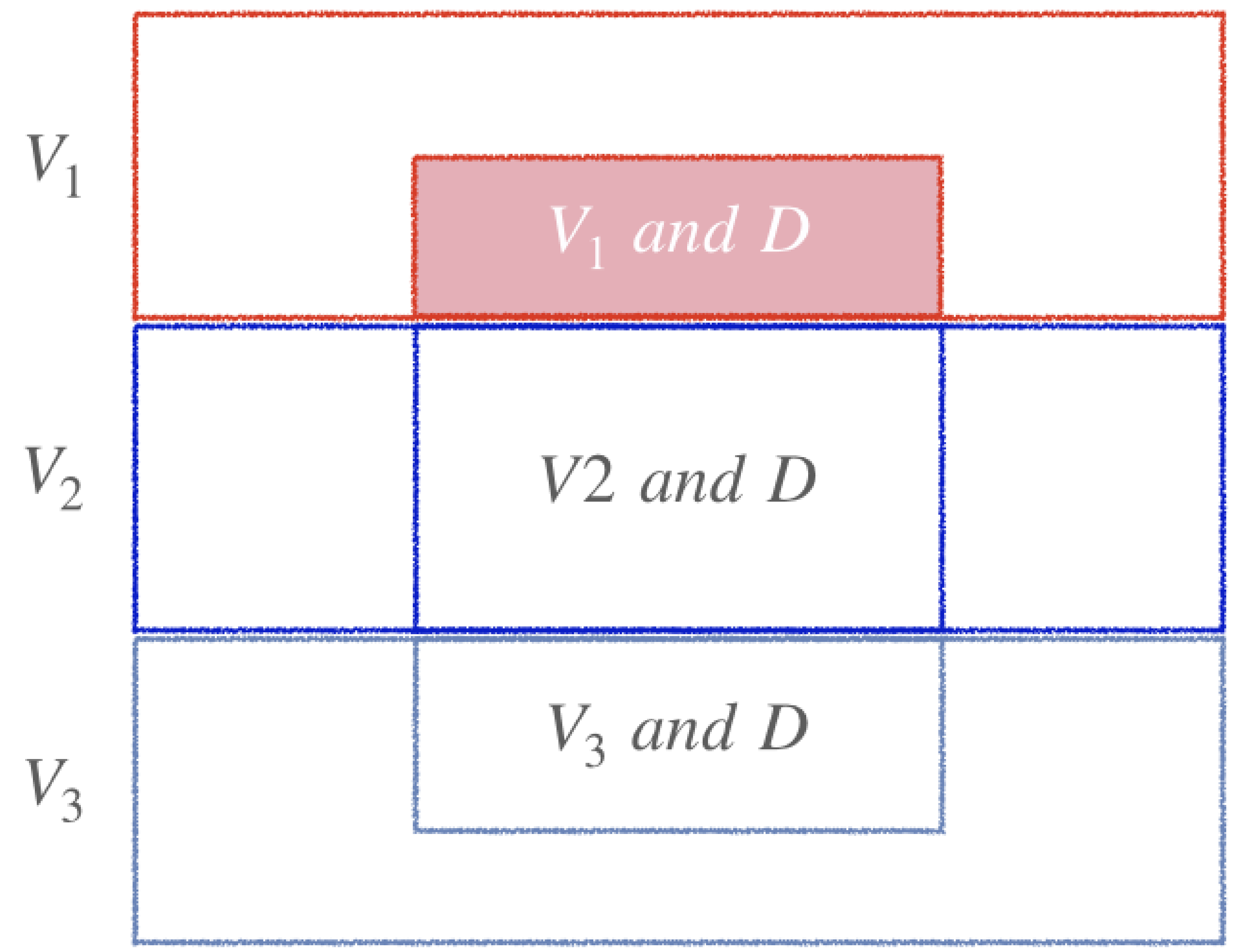
$$P(V_1|D) = \frac{P(V_1)P(D|V_1)}{P(V_1)P(D|V_1) + P(V_2)P(D|V_2) + P(V_3)P(D|V_3)}$$

Visual representation of Bayes' rule



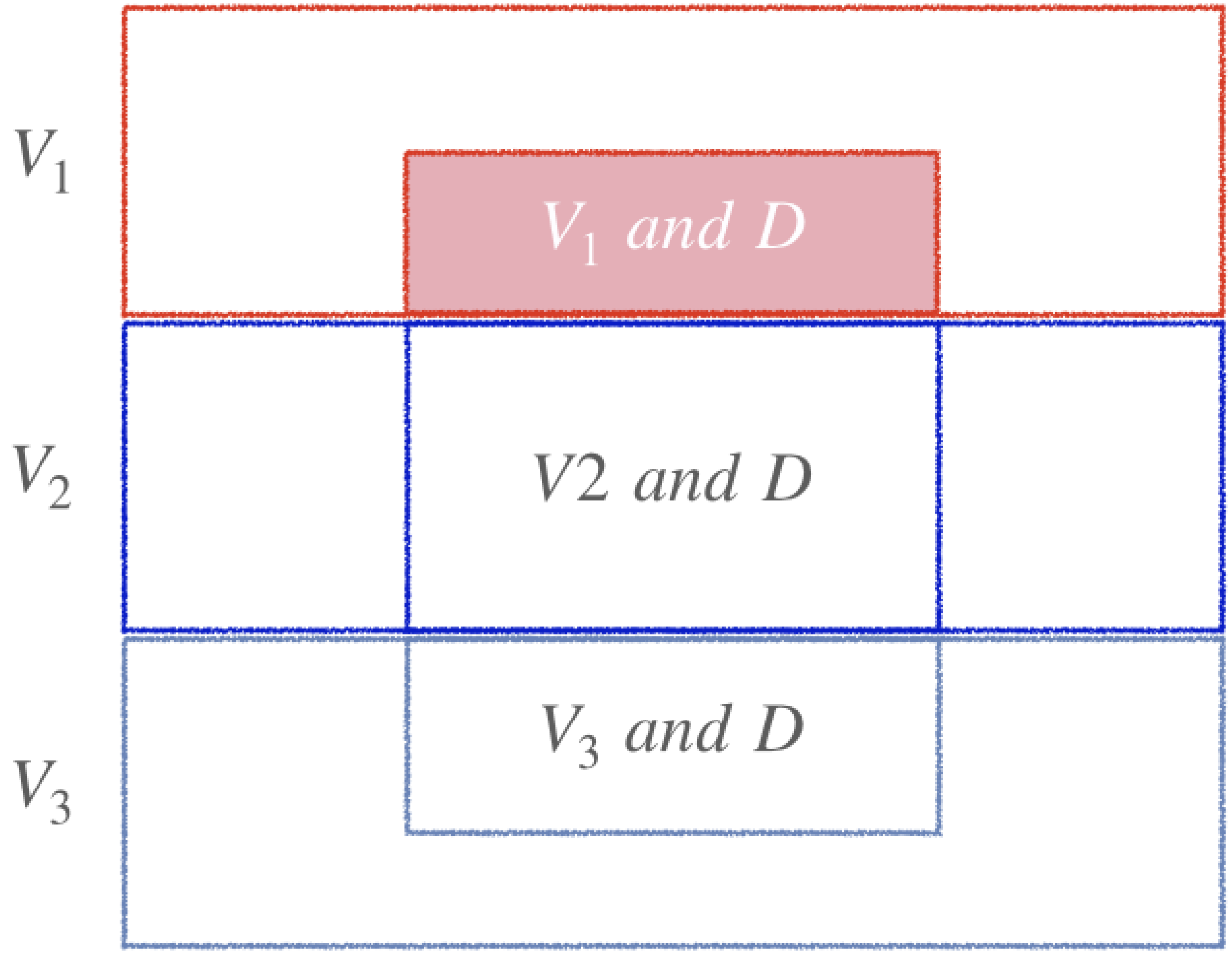
$$P(V_1|D) = \frac{P(V_1)P(D|V_1)}{P(V_1)P(D|V_1) + P(V_2)P(D|V_2) + P(V_3)P(D|V_3)}$$

Visual representation of Bayes' rule (Cont.)



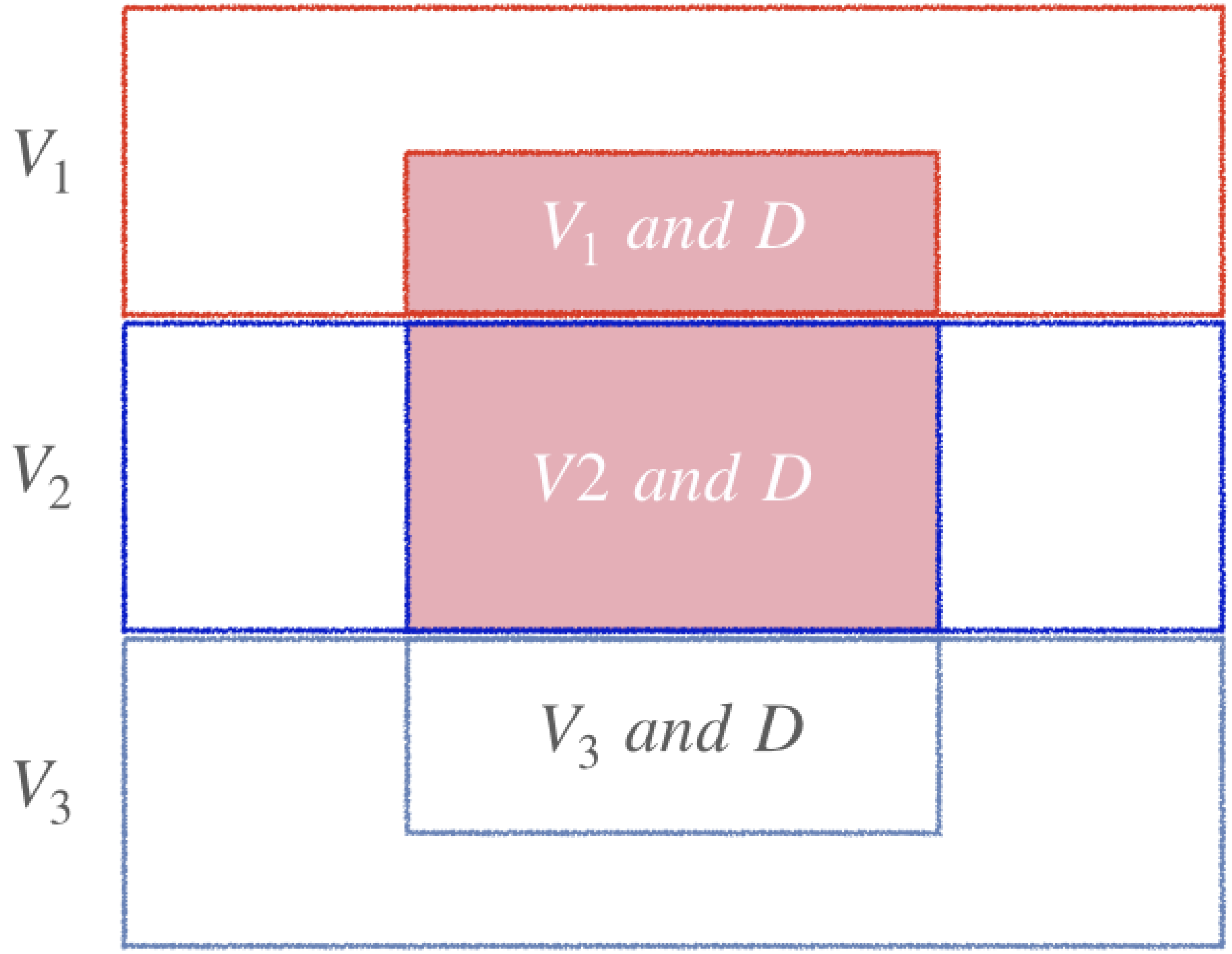
$$P(V_1|D) = \frac{P(V_1)P(D|V_1)}{P(V_1)P(D|V_1) + P(V_2)P(D|V_2) + P(V_3)P(D|V_3)}$$

Visual representation of Bayes' rule (Cont.)



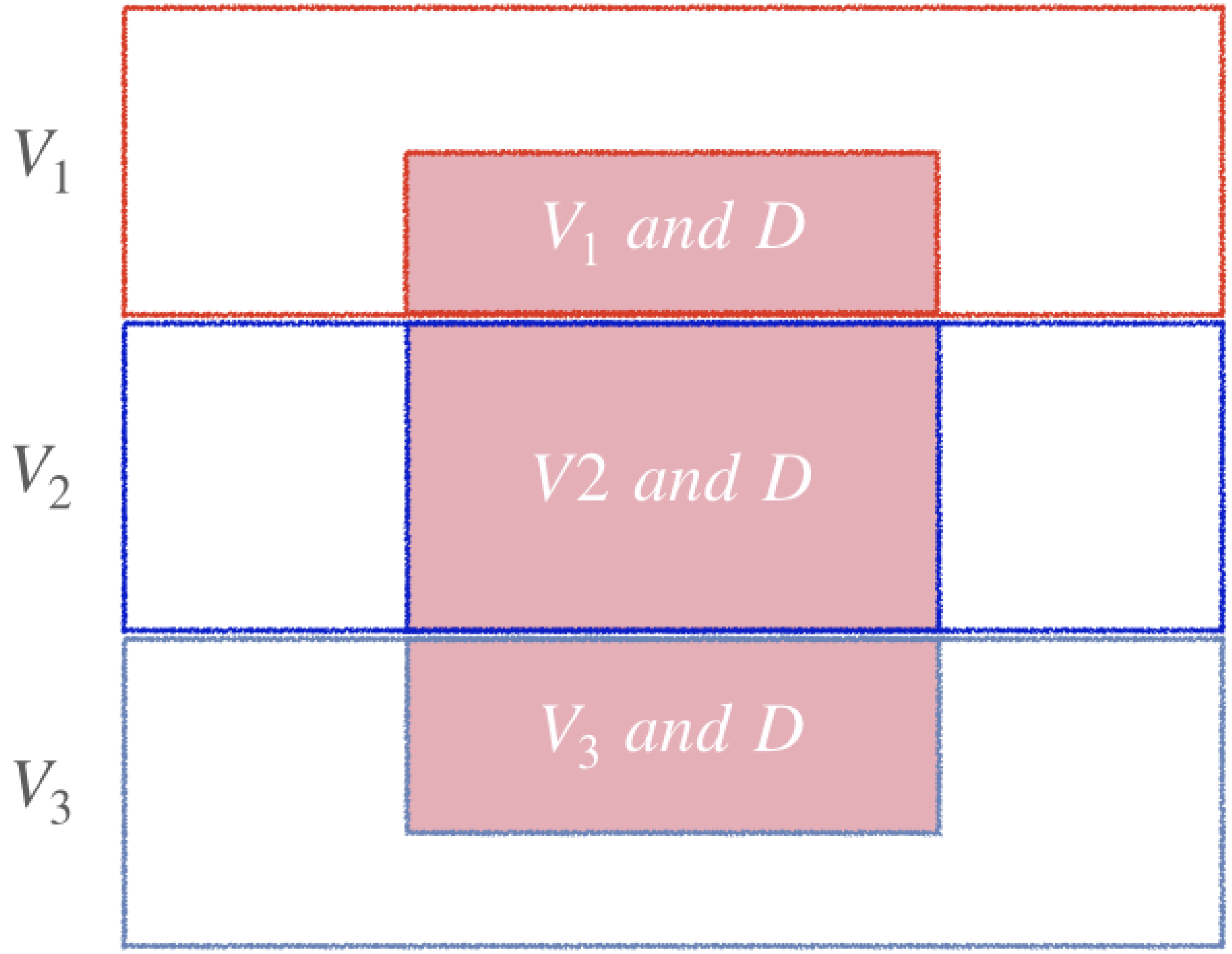
$$P(V_1|D) = \frac{P(V_1)P(D|V_1)}{P(V_1)P(D|V_1) + P(V_2)P(D|V_2) + P(V_3)P(D|V_3)}$$

Visual representation of Bayes' rule (Cont.)



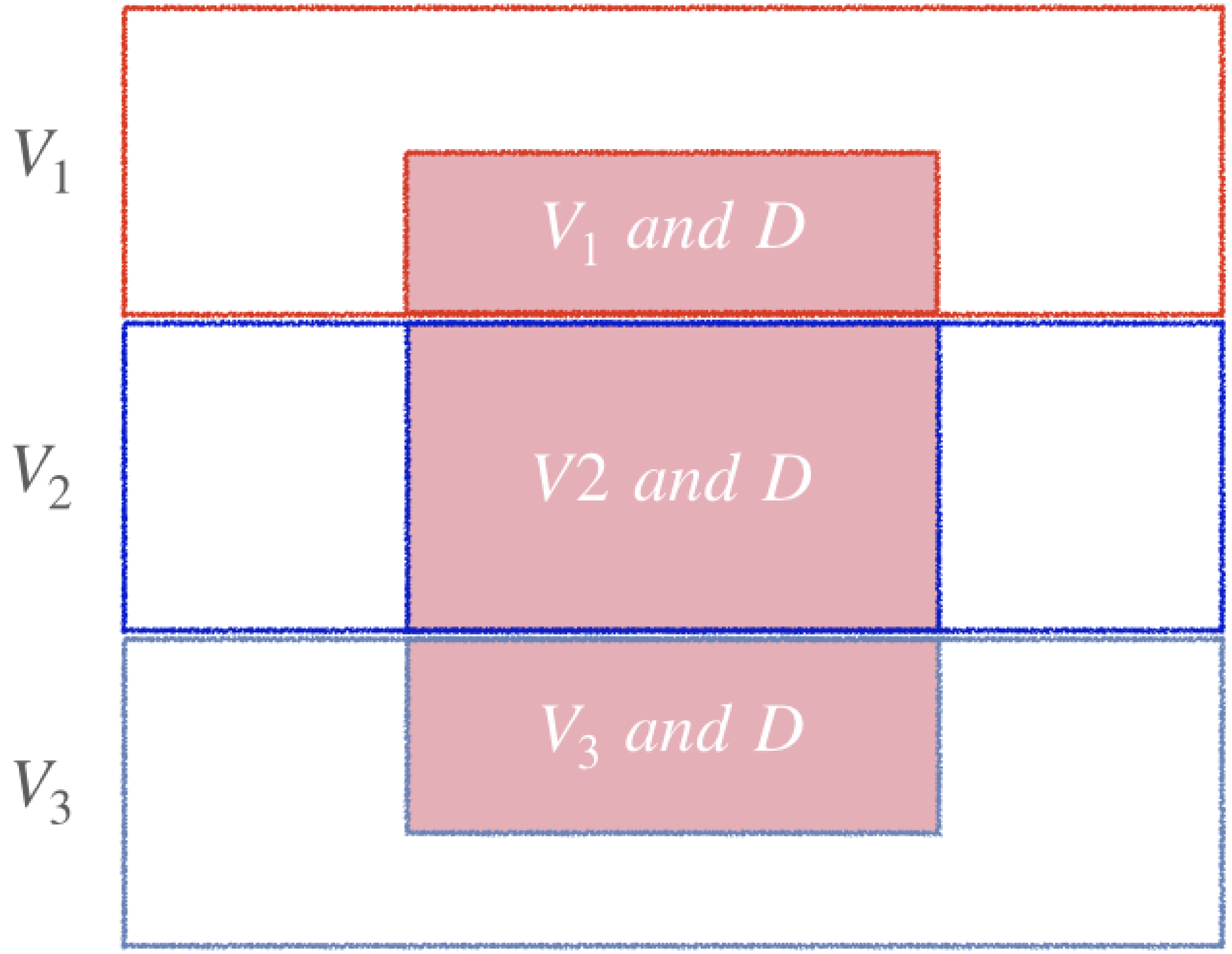
$$P(V_1|D) = \frac{P(V_1)P(D|V_1)}{P(V_1)P(D|V_1) + P(V_2)P(D|V_2) + P(V_3)P(D|V_3)}$$

Visual representation of Bayes' rule (Cont.)



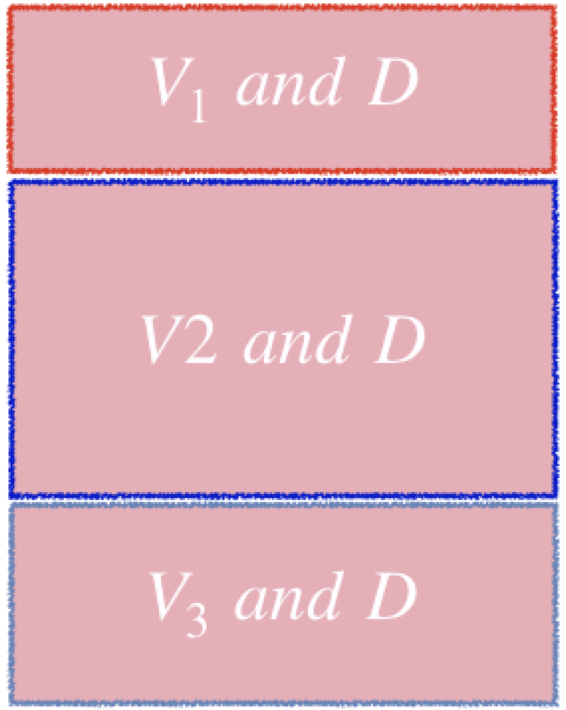
$$P(V_1|D) = \frac{P(V_1)P(D|V_1)}{P(V_1)P(D|V_1) + P(V_2)P(D|V_2) + P(V_3)P(D|V_3)}$$

Visual representation of Bayes' rule (Cont.)



$$P(V_1|D) = \frac{P(V_1)P(D|V_1)}{P(V_1)P(D|V_1) + P(V_2)P(D|V_2) + P(V_3)P(D|V_3)}$$

Visual representation of Bayes' rule (Cont.)



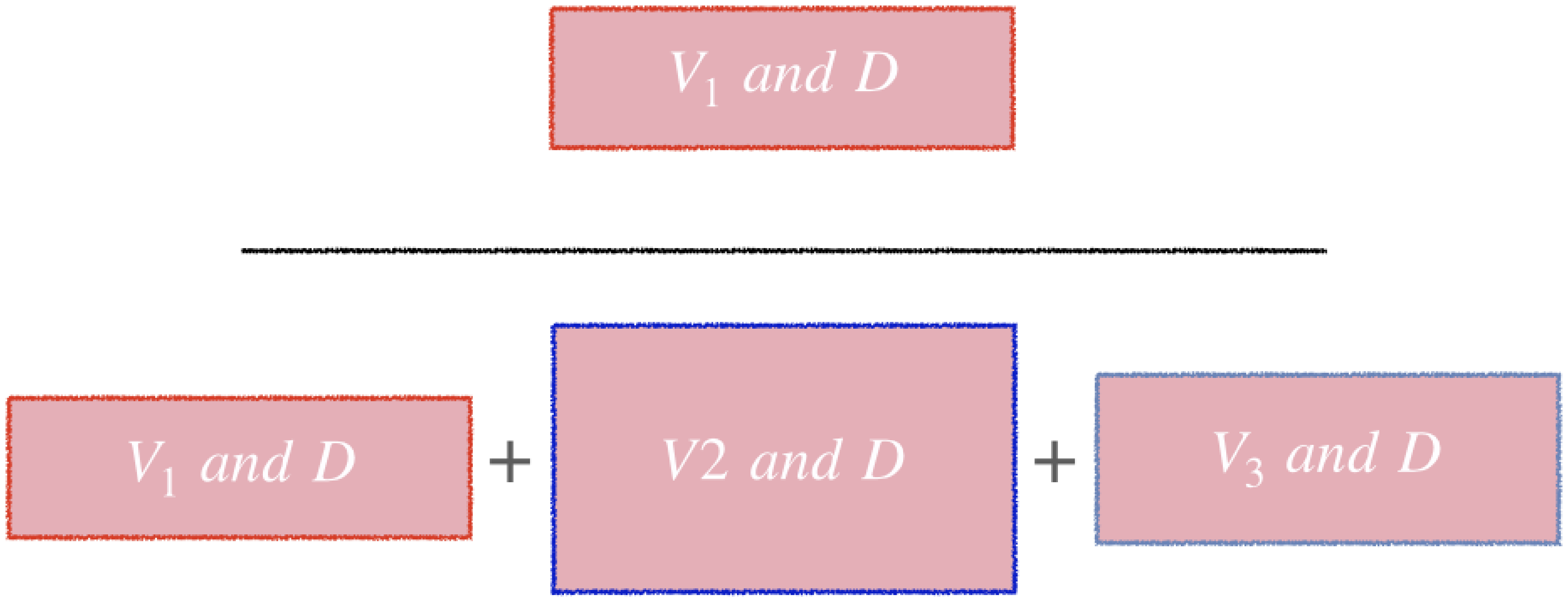
$$P(V_1|D) = \frac{P(V_1)P(D|V_1)}{P(V_1)P(D|V_1) + P(V_2)P(D|V_2) + P(V_3)P(D|V_3)}$$

Visual representation of Bayes' rule (Cont.)

$$\frac{\boxed{V_1 \; and \; D}}{\cdots}$$

$$P(V_1|D) = \frac{P(V_1)P(D|V_1)}{\cdots}$$

Visual representation of Bayes' rule (Cont.)



$$P(V_1|D) = \frac{P(V_1)P(D|V_1)}{P(V_1)P(D|V_1) + P(V_2)P(D|V_2) + P(V_3)P(D|V_3)}$$

Factories and parts example in Python

A certain electronic part is manufactured by three different vendors, V1, V2, and V3.

Half of the parts are produced by V1, and V2 and V3 each produce 25%. The probability of a part being damaged given that it was produced by V1 is 1%, while it's 2% for V2 and 3% for V3.

- Given that the part is damaged, what is the probability that it was manufactured by V1?

Factories and parts example in Python (Cont.)

- Given that the part is damaged, get the probability that it was manufactured by V1.

```
P_V1 = 0.5
P_V2 = 0.25
P_V3 = 0.25
```

```
P_D_g_V1 = 0.01
P_D_g_V2 = 0.02
P_D_g_V3 = 0.03
```

```
P_Damaged = P_V1 * P_D_g_V1 + P_V2 * P_D_g_V2 + P_V3 * P_D_g_V3
```

Factories and parts example in Python (Cont.)

```
P_V1_g_D = (P_V1 * P_D_g_V1) / P_Damaged # P(V1|D) calculation
print(P_V1_g_D)
```

A randomly selected part which is damaged is manufactured by V1 with probability:

```
0.285714285714
```

Let's exercise with Bayes

FOUNDATIONS OF PROBABILITY IN PYTHON