Dataset

Edibility	Cap Shape	Cap Color	Odor	Habitat
poisonous	convex	brown	pungent	urban
edible	convex	yellow	almond	grasses
edible	bell	white	anise	meadows
poisonous	convex	white	pungent	urban
edible	convex	gray	none	grasses
edible	convex	yellow	almond	grasses
edible	bell	white	almond	meadows
edible	bell	white	anise	meadows
poisonous	convex	white	pungent	grasses
edible	bell	yellow	almond	meadows
edible	convex	yellow	anise	grasses
edible	convex	yellow	almond	meadows
edible	bell	yellow	almond	grasses
poisonous	convex	white	pungent	urban
edible	convex	brown	none	grasses
edible	sunken	gray	none	urban
edible	flat	white	none	grasses
poisonous	convex	brown	pungent	grasses
poisonous	convex	white	pungent	urban
poisonous	convex	brown	pungent	urban
edible	bell	yellow	almond	meadows
poisonous	convex	brown	pungent	grasses
edible	bell	yellow	anise	meadows
edible	bell	white	almond	meadows
edible	bell	white	anise	meadows
poisonous	flat	white	pungent	grasses
edible	convex	yellow	almond	meadows
edible	convex	white	anise	meadows
edible	flat	brown	none	urban
edible	convex	yellow	almond	woods
edible	bell	yellow	anise	meadows
poisonous	convex	white	pungent	urban
edible	convex	yellow	anise	meadows
edible	convex	brown	anise	paths
edible	bell	yellow	anise	meadows
edible	convex	yellow	almond	woods
edible	sunken	gray	none	urban
poisonous	convex	yellow	pungent	urban
edible	convex	yellow	almond	woods
edible	bell	yellow	anise	meadows
edible	bell	yellow	almond	grasses
edible	convex	yellow	anise	paths
edible	convex	brown	none	urban
poisonous	convex	white	pungent	grasses
edible	convex	yellow	almond	meadows
edible	convex	white	almond	grasses
edible	convex	yellow	anise	meadows
edible	convex	white	anise	meadows
edible	convex	yellow	anise	paths
edible	flat	yellow	pungent	paths
edible	convex	brown	almond	grasses
edible	convex	white	anise	grasses
edible	bell	white	anise	meadows
poisonous	convex	brown	pungent	urban
poisonous	convex	white	pungent	urban
edible	bell	yellow	almond	meadows
edible	flat	gray	none	grasses
edible	bell	white	almond	grasses
edible	convex	yellow	anise	grasses
edible	convex	brown	almond	paths
edible	sunken	gray	none	urban
edible	bell	yellow	almond	meadows
edible	bell	yellow	almond	meadows
edible	bell	yellow	anise	meadows
edible	bell	white	anise	grasses
edible	flat	brown	none	grasses
edible	convex	white	almond	grasses
edible	flat	yellow	almond	grasses
edible	convex	yellow	almond	grasses
edible	convex	gray	none	urban
edible	flat	yellow	anise	woods
edible	bell	white	anise	meadows
edible	flat	yellow	anise	woods
edible	convex	brown	almond	paths
edible	bell	brown	almond	grasses
edible	bell	white	anise	woods
edible	convex	white	anise	woods
edible	flat	brown	anise	paths
poisonous	convex	brown	pungent	urban
edible	flat	brown	almond	grasses
edible	convex	brown	none	grasses
poisonous	convex	white	pungent	grasses
edible	flat	gray	none	urban
edible	convex	gray	none	grasses
edible	convex	yellow	anise	grasses
edible	convex	brown	none	grasses
edible	convex	brown	almond	grasses
edible	convex	white	almond	grasses
edible	convex	yellow	anise	grasses
edible	flat	brown	anise	grasses
edible	sunken	brown	none	urban
edible	convex	brown	none	urban
edible	bell	white	anise	grasses
edible	convex	yellow	almond	grasses
edible	convex	yellow	anise	meadows
edible	convex	brown	none	grasses
edible	convex	white	none	grasses
edible	flat	brown	anise	grasses
edible	convex	yellow	almond	grasses
edible	bell	white	almond	grasses
edible	convex	white	almond	grasses
edible	convex	white	almond	grasses
edible	bell	yellow	anise	grasses
edible	bell	yellow	anise	grasses
edible	flat	white	almond	grasses
edible	convex	yellow	almond	paths
edible	bell	white	anise	grasses
edible	convex	white	almond	meadows
edible	convex	yellow	almond	grasses
edible	bell	white	anise	meadows
edible	bell	white	almond	meadows
edible	convex	yellow	almond	meadows
edible	convex	yellow	anise	grasses
edible	sunken	gray	none	urban
edible	convex	white	almond	woods
edible	convex	yellow	almond	meadows
poisonous	convex	white	pungent	urban
edible	convex	yellow	anise	grasses
edible	convex	yellow	none	grasses
edible	sunken	gray	none	urban
edible	convex	white	none	grasses
edible	convex	yellow	anise	woods
edible	sunken	brown	none	urban
poisonous	convex	white	pungent	grasses
edible	convex	white	almond	meadows
poisonous	flat	brown	pungent	grasses
edible	flat	gray	none	grasses
edible	convex	white	none	meadows
edible	convex	white	none	grasses
edible	bell	yellow	almond	grasses
edible	flat	gray	none	grasses
edible	convex	white	almond	grasses
edible	convex	white	anise	meadows
edible	bell	white	anise	grasses
edible	bell	white	almond	meadows
edible	bell	white	anise	grasses
edible	flat	white	almond	woods
edible	convex	yellow	anise	meadows
edible	flat	white	almond	woods

More rows hidden. In total, there are 600 rows.

Variable Values

Edible	Cap Shape	Cap Color	Odor	Habitat
edible	convex	brown	almond	grasses
poisonous	bell	gray	anise	meadows
	flat	white	none	paths
	sunken	yellow	pungent	urban
				woods

Recall Bayes' Theorem:

$$P(E \mid S, C, O, H) = \frac{P(S, C, O, H \mid E)P(E)}{P(S, C, O, H)}$$

In Naive Bayes, we assume that

$$P(S, C, O, H \mid E) = P(S \mid E)P(C \mid E)P(O \mid E)P(H \mid E)$$

Then, what is $P(S, C, O, H)$?

$$P(S, C, O, H) = P(S, C, O, H \mid E = e) + P(S, C, O, H \mid E = p)$$

Now, we have everything we need:

Edibility

edible	538
poisonous	62
Total	600

Cap Shape

convex	324
bell	143
flat	111
sunken	22

Cap Color

brown	131
gray	49
white	212
yellow	208

Odor

almond	221
anise	205
none	111
pungent	63

Habitat

grasses	254
meadows	148
paths	48
urban	93
woods	57

Cap Shape | Edibility

	convex	bell	flat	sunken
edible	268	143	105	22
poisonous	56	0	6	0

Cap Color | Edibility

	brown	gray	white	yellow
edible	98	49	183	208
poisonous	33	0	29	0

Odor | Edibility

	almond	anise	none	pungent
edible	221	205	111	1
poisonous	0	0	0	62

Habitat | Edibility

	grasses	meadows	paths	urban	woods
edible	227	148	48	58	57
poisonous	27	0	0	35	0

And to calculate the probabilities:

P(Edibility)

edible	0.897
poisonous	0.103
Total	600

P(Cap Shape)

convex	0.540
bell	0.238
flat	0.185
sunken	0.037

P(Cap Color)

brown	0.218
gray	0.082
white	0.353
yellow	0.347

P(Odor-1)

almond	0.368
anise	0.342
none	0.185
pungent	0.105

P(Habitat)

grasses	0.423
meadows	0.247
paths	0.080
urban	0.155
woods	0.095

P(Cap Shape | Edibility)

	convex	bell	flat	sunken
edible	0.498	0.266	0.195	0.041
poisonous	0.903	0.000	0.097	0.000

P(Cap Color | Edibility)

	brown	gray	white	yellow
edible	0.182	0.091	0.340	0.387
poisonous	0.532	0.000	0.468	0.000

P(Odor | Edibility)

	almond	anise	none	pungent
edible	0.411	0.381	0.206	0.002
poisonous	0.000	0.000	0.000	1.000

P(Habitat | Edibility)

	grasses	meadows	paths	urban	woods
edible	0.422	0.275	0.089	0.108	0.106
poisonous	0.435	0.000	0.000	0.565	0.000

Now, let's calculate an example.

Let's say we are walking in **grass**, and we see a mushroom. The mushroom has a **flat**, **white** cap, and a **pungent** smell.

What is the probability that this mushroom is **poisonous**?

In mathematical terms, what is the following?

$$P(E = p \mid S = f, C = w, O = p, H = g) \qquad \text{For brevity, I'll call it } P(p \mid f, w, p, g)$$

From Bayes' Theorem, we have:

$$P(p \mid f, w, p, g) = \frac{P(f, w, p, g \mid p)P(p)}{P(f, w, p, g)}$$

Let's calculate the top term first. We know from our assumption that it equals:

$$P(f \mid p)P(w \mid p)P(p \mid p)P(g \mid p)P(p)$$

Notice how each term can be found by simply counting? We can calculate it by looking up the corresponding probability in the tables above. The values needed are highlighted in yellow.

$$P(f \mid p)P(w \mid p)P(p \mid p)P(g \mid p)P(p) \approx 0.097 \times 0.468 \times 1 \times 0.435 \times 0.103$$
$$\approx 0.002$$

Then the bottom term. As we have shown:

$$P(f, w, p, g) = P(f, w, p, g \mid e) + P(f, w, p, g \mid p)$$

And since we assume conditional independence between S, C, O, P given E, we have:

$$P(f, w, p, g) = P(f \mid e)P(w \mid e)P(p \mid e)P(g \mid e)P(e) + P(f \mid p)P(w \mid p)P(p \mid p)P(g \mid p)P(p)$$

We know that the second half is 0.002 — it's the same as above. What is the first half, then? We look up the table again!

The values needed are highlighted in blue.

$$P(f \mid e)P(w \mid e)P(p \mid e)P(g \mid e)P(e) \approx 0.195 \times 0.340 \times 0.002 \times 0.422 \times 0.897$$
$$\approx 0.00005$$

Now we can calculate the final answer!

$$P(p \mid f, w, p, g) = \frac{P(f, w, p, g \mid p)}{P(f, w, p, g \mid e) + P(f, w, p, g \mid p)}$$
$$\approx \frac{0.002}{0.00005 + 0.002}$$
$$\approx 0.976$$

Therefore, don't eat that mushroom!!!