Example from the first class

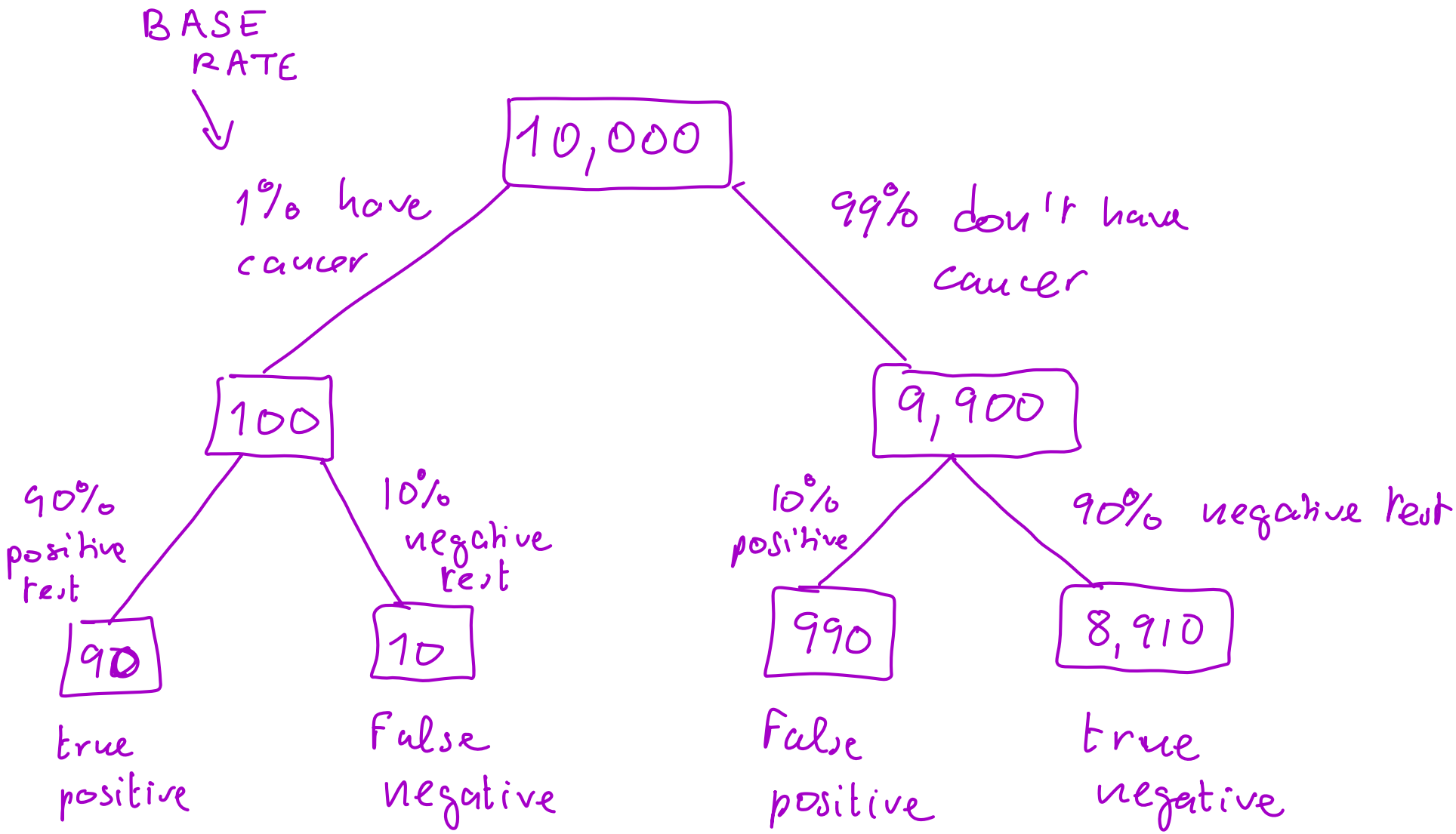
How to process information

- In the US, **1% of women of age 40 have breast cancer**.
- If a woman **has** breast cancer, the probability that she tests **positive** on a screening mammogram is **90%**.
- If she **does not have** breast cancer, the probability that she tests **negative** on a screening mammogram is **90%**.

That is, mammograms have a **90% accuracy**.

Susan is a 40-year old woman who tested **positive** *on a mammogram.*

What are the chances that she actually has breast cancer?



prob of cancer given positive test $= \frac{90}{90 + 990} = 0.0833$

Correct answer: 8.33%

CONDITIONAL REASONING: the FREQUENCY approach

- Suppose there is a new variant of COVID
- The fraction p of the population is infected
- Typical symptoms: nasal congestion
- 80% of those infected have the symptoms
- 10% of those **not** infected have the symptoms

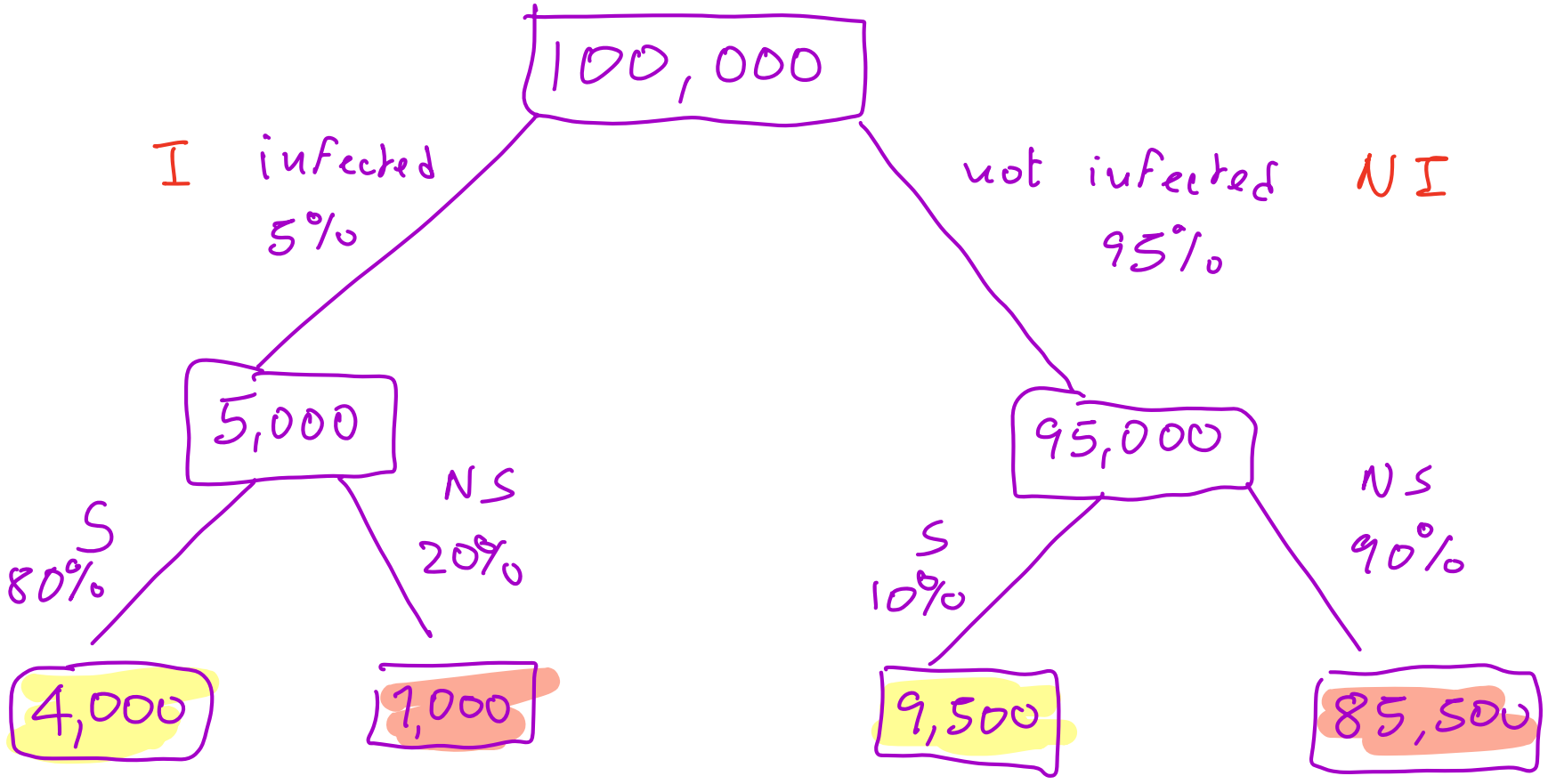
Suppose that $p = 5\%$. You wake up with nasal congestion.

How likely is it that you are infected?

- 5% of the population are infected
- 80% of those infected have the symptoms
- 10% of those **not** infected have the symptoms

S = symptoms

NS = no symptoms



$$Pr(I \mid S) = \frac{4{,}000}{4{,}000 + 9{,}500} = 29.63\%$$

↑ infected

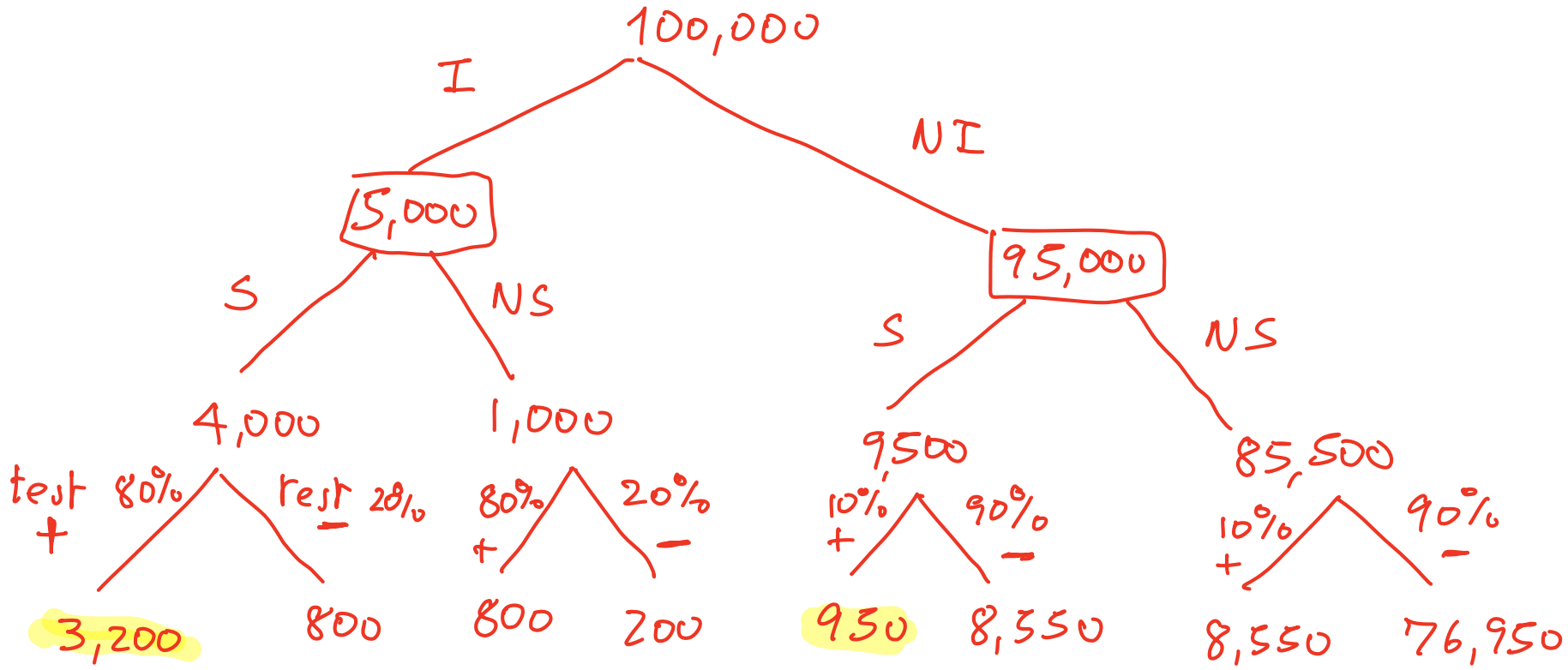
$$P(NI \mid NS) = \frac{85{,}500}{1{,}000 + 85{,}500}$$

A test is now available. The probability of testing positive is independent of whether or not you have symptoms:

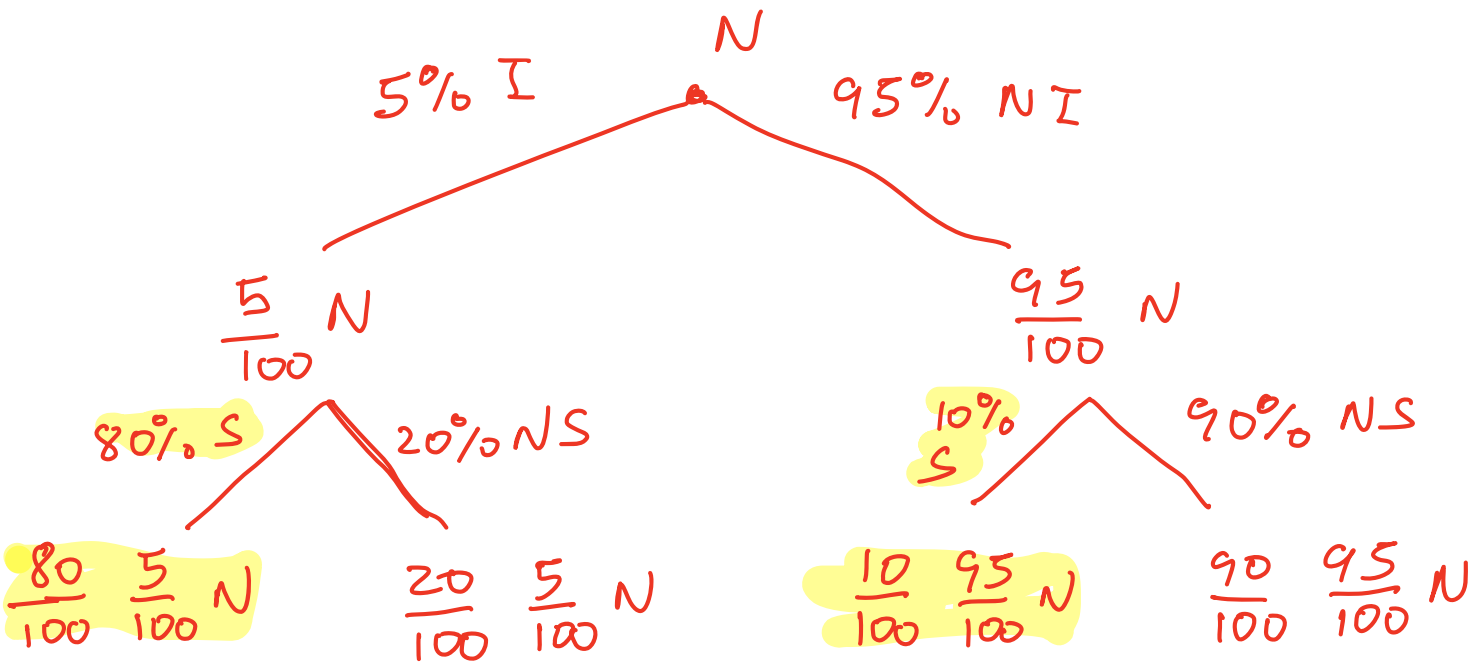
- If you are infected, the probability of testing positive is 80% (whether or not you have the symptoms)
- If you are **not** infected, the probability of testing positive is 10% (whether or not you have the symptoms)

Since you woke up with symptoms, you decided to get tested and the result was positive. How likely is it that you are infected?

- If you are infected, the probability of testing positive is 80% (whether or not you have the symptoms)
- If you are **not** infected, the probability of testing positive is 10% (whether or not you have the symptoms)



$$P(I \mid S\&+) = \frac{3,200}{3,200+950} \simeq 77.11\%$$



$$P(I \mid S) = \frac{\frac{80}{100}\frac{5}{100}N}{\frac{80}{100}\frac{5}{100}N + \frac{10}{100}\frac{95}{100}N} = \frac{80 \cdot 5}{80 \cdot 5 + 10 \cdot 95}$$

$$= 29.63\%$$

One more example

Base rate of a disease: percentage of the population that has the disease

Sensitivity of a test: percentage of those who have the disease that tests positive → probability of a true positive

Specificity of a test: percentage of those who **do not** have the disease that tests negative → probability of a true negative

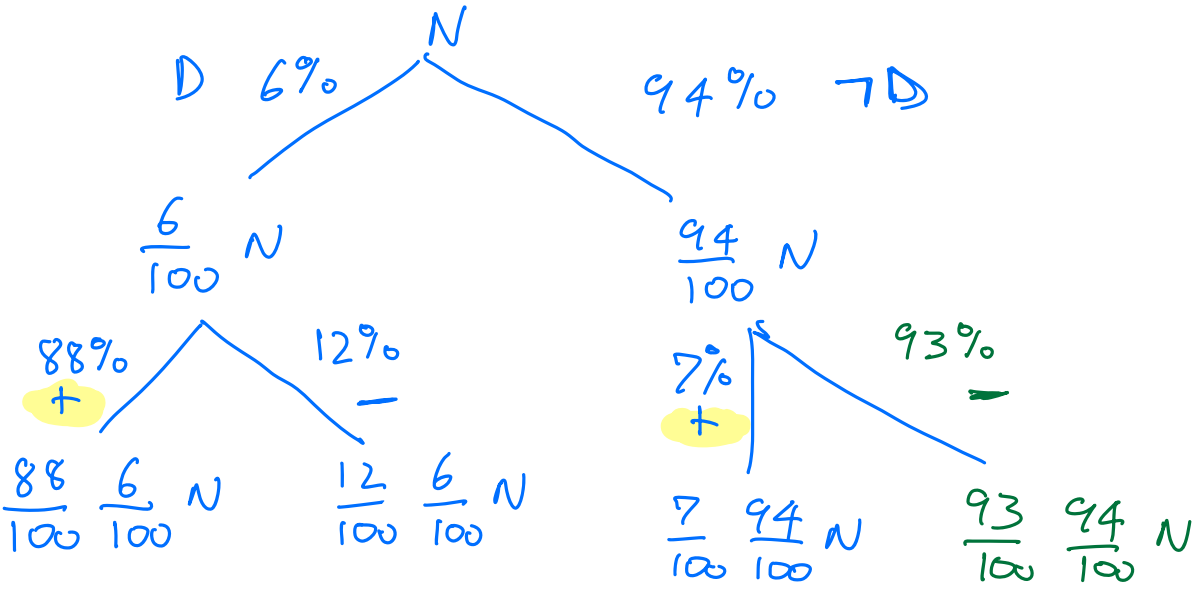
Suppose:

Base rate = 6%

Sensitivity = 88%

Specificity = 93%

Suppose you test positive. What is the probability that you have the disease?



$$P(D|+) = \frac{\frac{88}{100}\cdot\frac{6}{100}\cdot N}{\frac{88}{100}\cdot\frac{6}{100}\cdot N + \frac{7}{100}\cdot\frac{94}{100}\cdot N}$$

$$= \frac{6 \cdot 88}{6 \cdot 88 + 7 \cdot 94}$$

MORE THAN TWO CATEGORIES

Enrollment in a class

ECN	*ARE*	*PSY*	*Other*
38%	20%	12%	30%

Percentages of those who passed:

major	*ECN*	*ARE*	*PSY*	*Other*
percentage who passed	70%	60%	40%	35%

You learn that Ann passed the class. How likely is it that Ann is a PSY major?

major	*ECN*	*ARE*	*PSY*	*Other*	**Ann passed the class. How likely is it that she is a PSY major?**
enrollment	38%	20%	12%	30%	
percentage who passed	70%	60%	40%	35%	