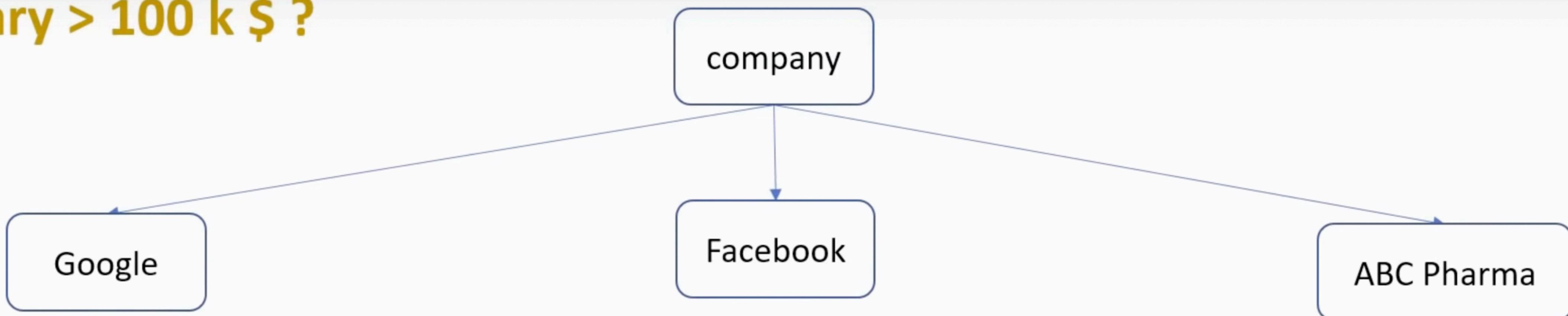


# **SC201**

## Lecture 8

Company	Job	Degree	Salary_more_then_100k
google	sales executive	bachelors	0
google	sales executive	masters	0
google	business manager	bachelors	1
google	business manager	masters	1
google	computer programmer	bachelors	0
google	computer programmer	masters	1
abc pharma	sales executive	masters	0
abc pharma	computer programmer	bachelors	0
abc pharma	business manager	bachelors	0
abc pharma	business manager	masters	1
facebook	sales executive	bachelors	1
facebook	sales executive	masters	1
facebook	business manager	bachelors	1
facebook	business manager	masters	1
facebook	computer programmer	bachelors	1
facebook	computer programmer	masters	1

# Salary > 100 k \$ ?



google	sales executive	bachelors
google	sales executive	masters
google	business manager	bachelors
google	business manager	masters
google	computer programmer	bachelors
google	computer programmer	masters

?

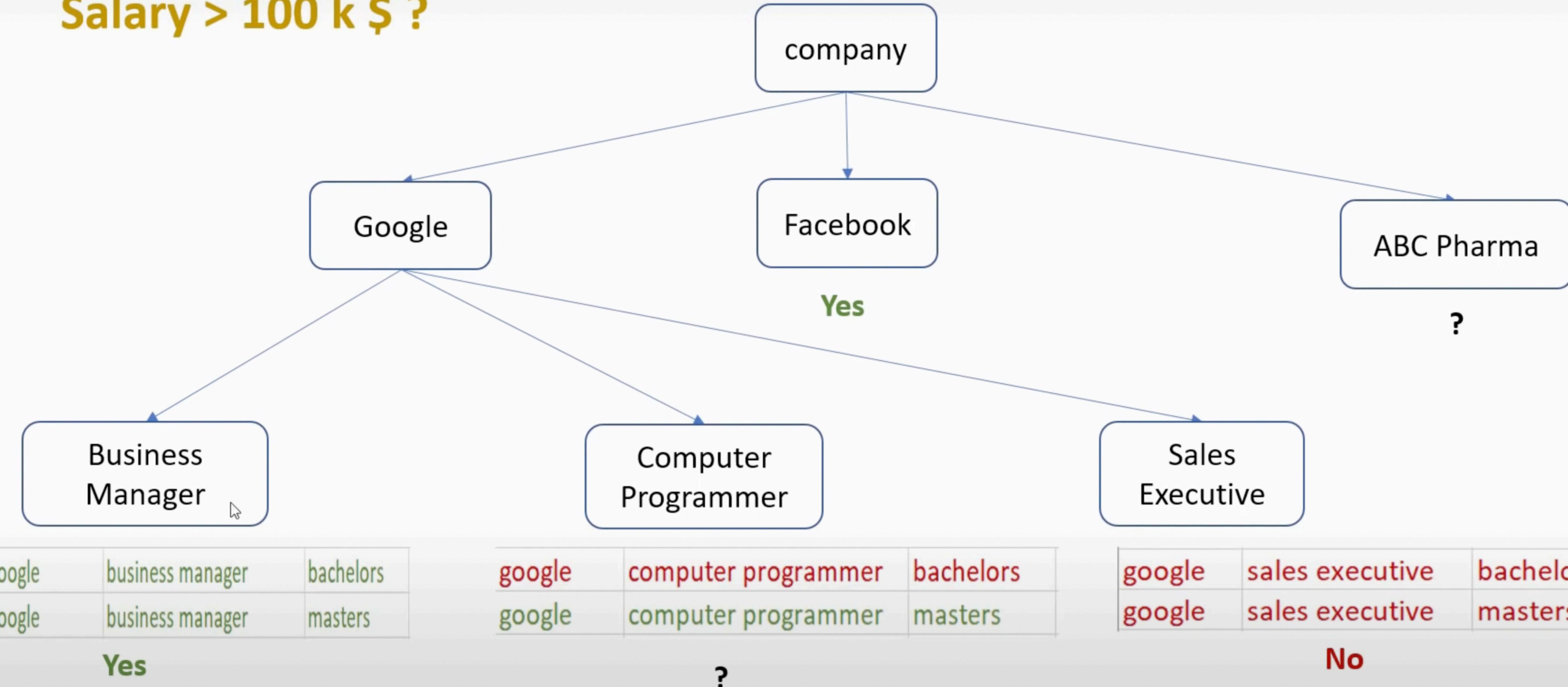
facebook	sales executive	bachelors
facebook	sales executive	masters
facebook	business manager	bachelors
facebook	business manager	masters
facebook	computer programmer	bachelors
facebook	computer programmer	masters

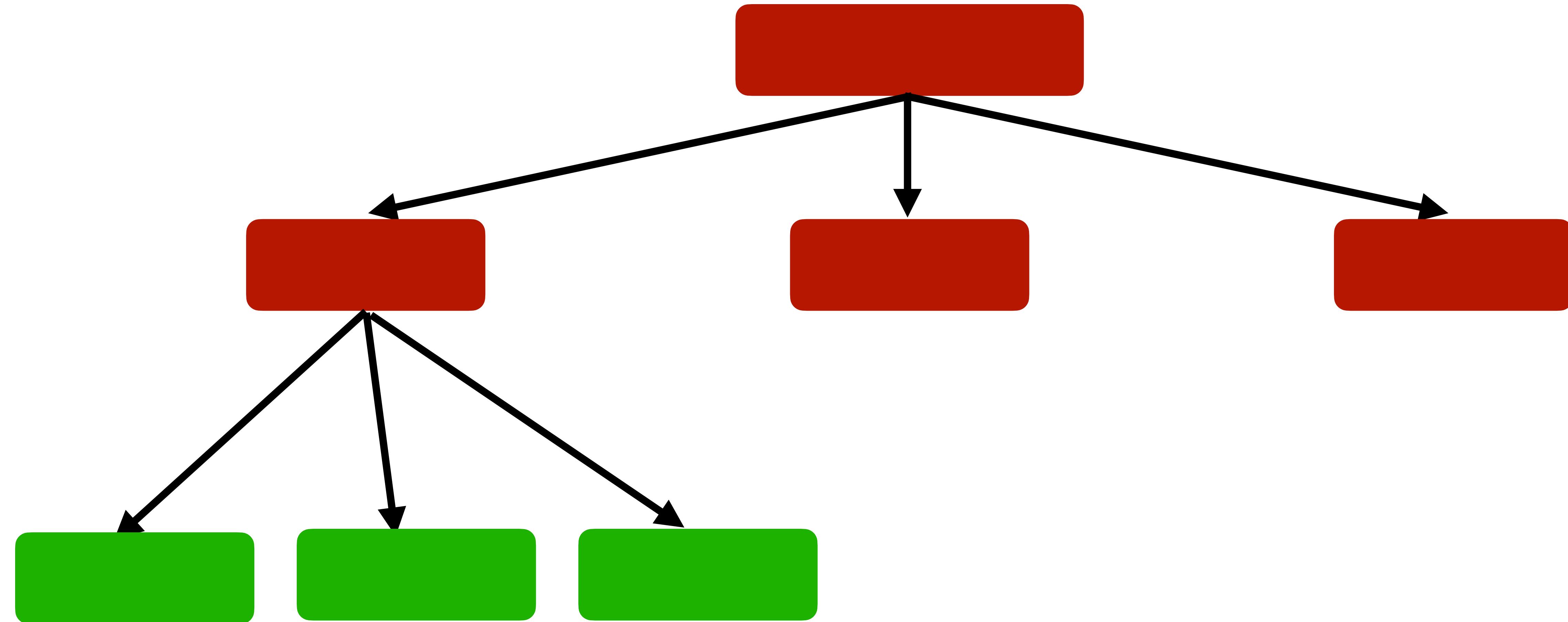
Yes

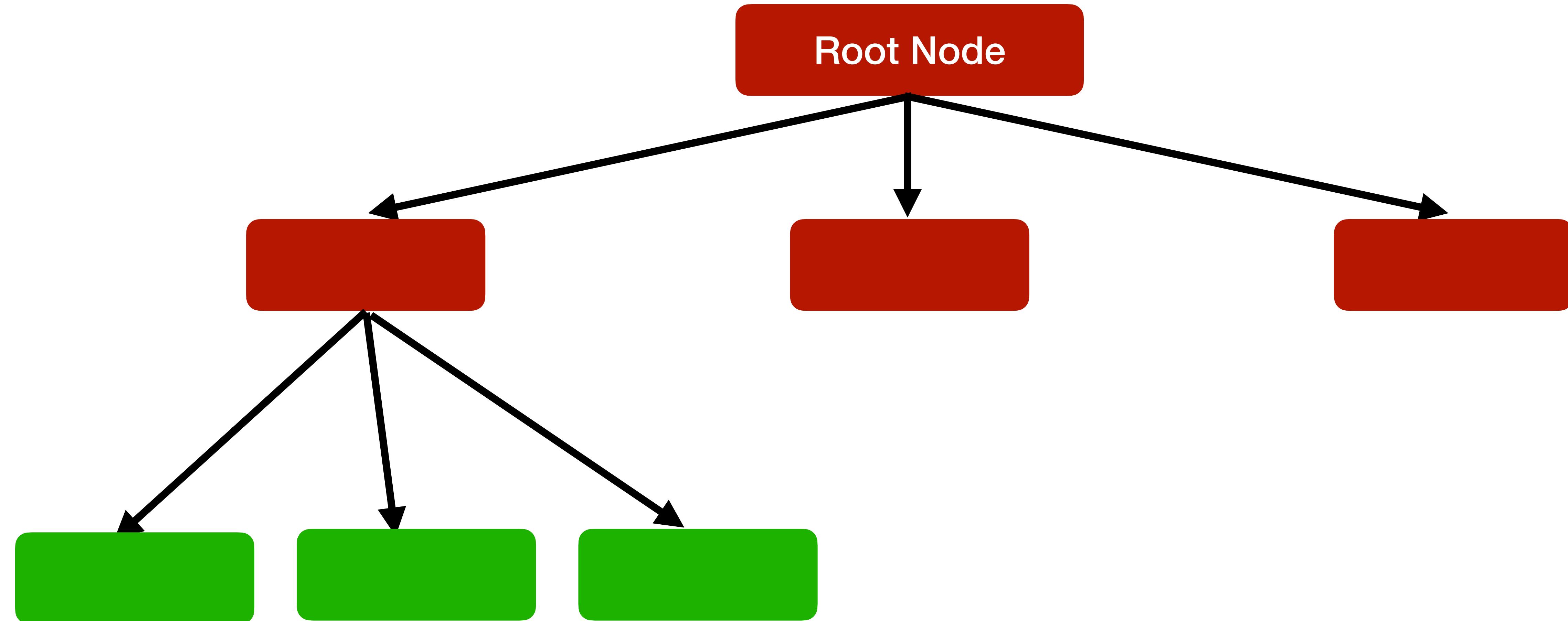
abc pharma	sales executive	masters
abc pharma	computer programmer	bachelors
abc pharma	business manager	bachelors
abc pharma	business manager	masters

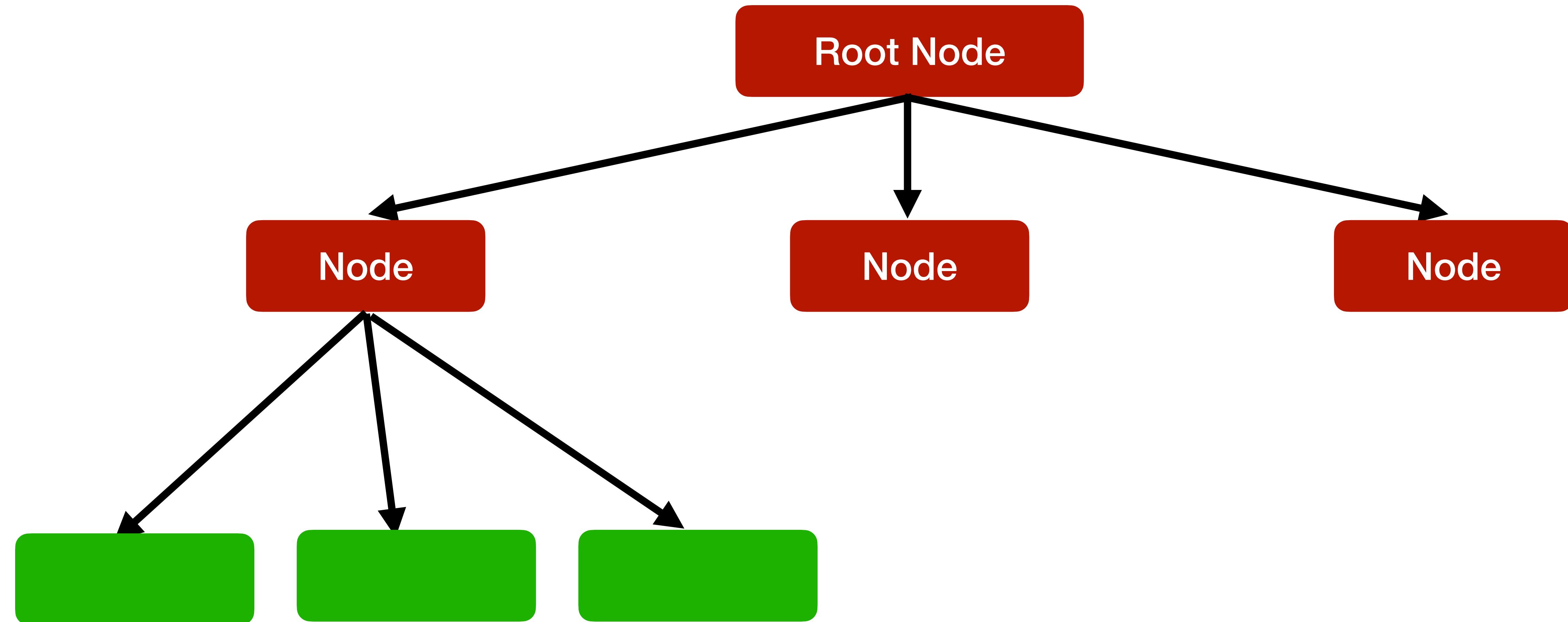
?

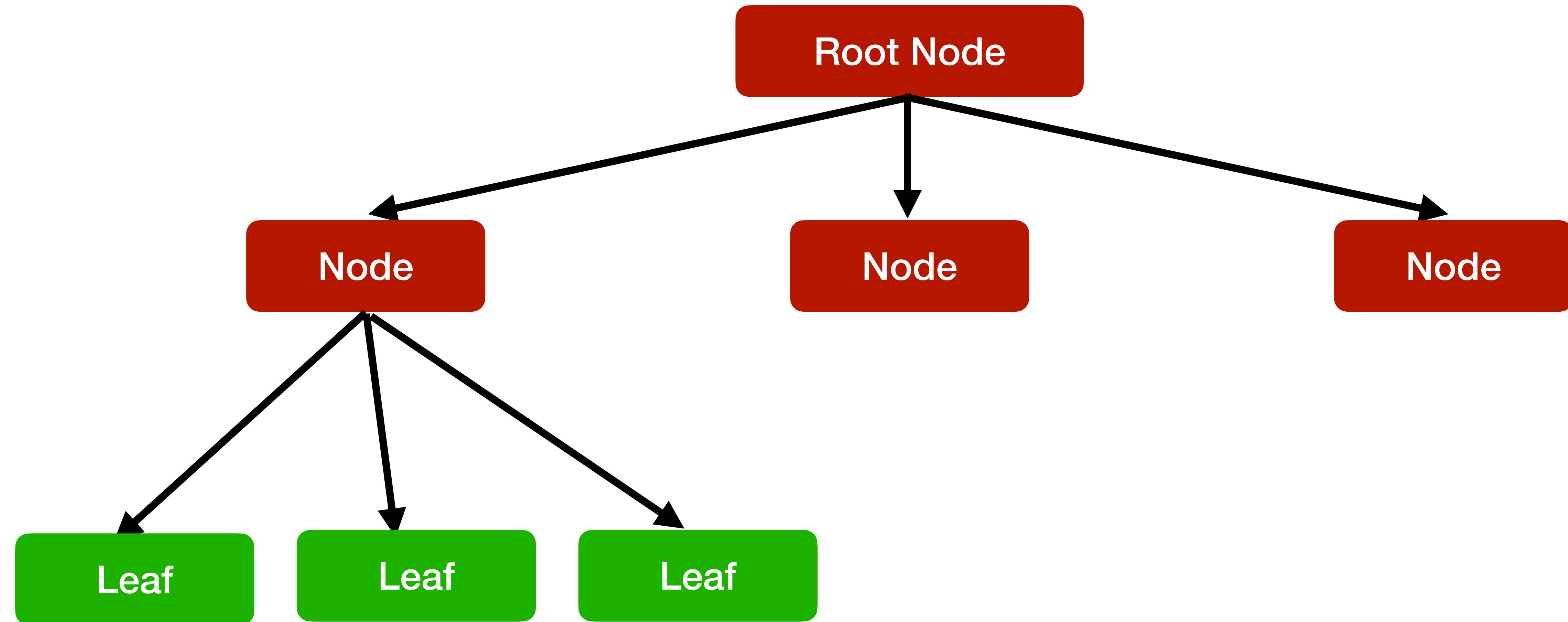
## Salary > 100 k \$ ?



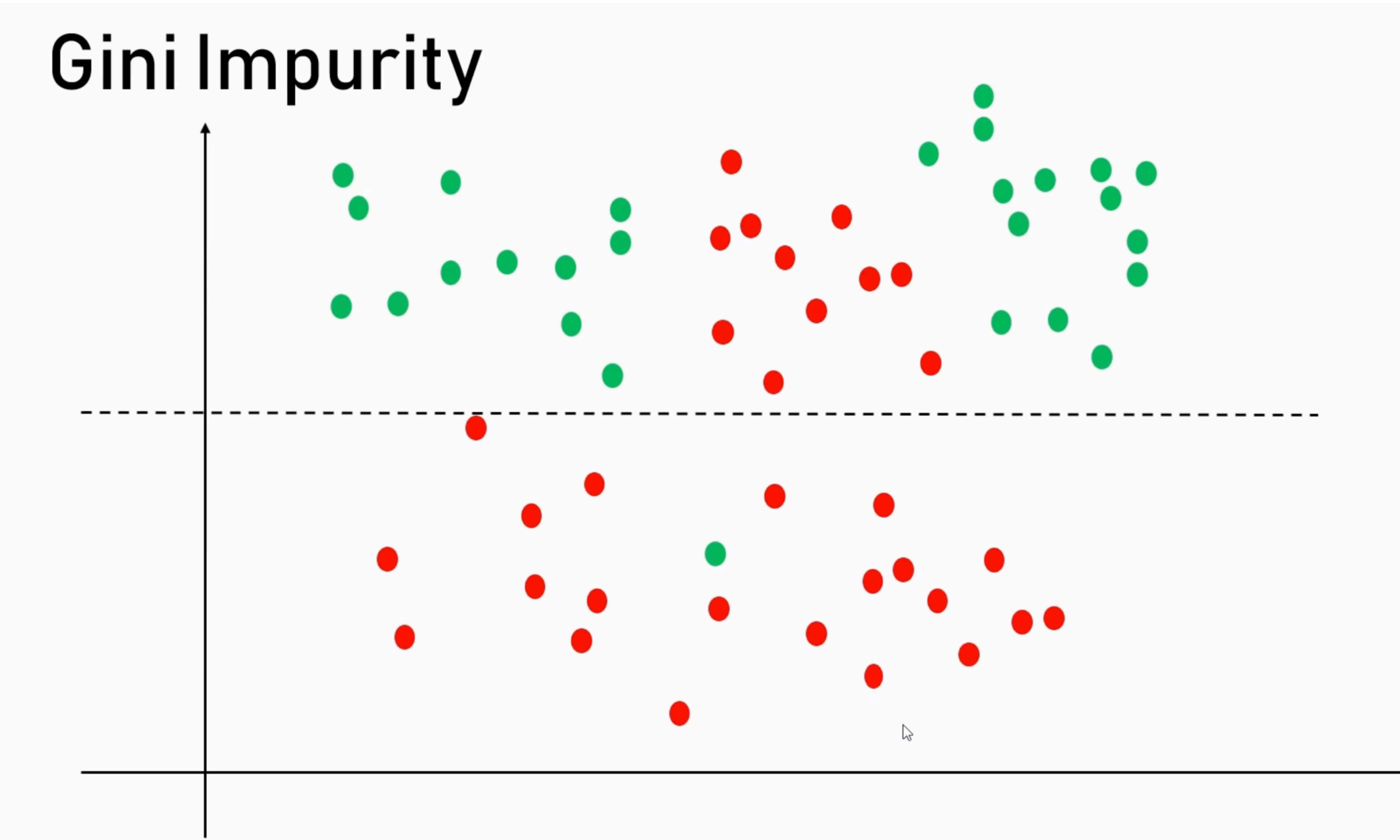






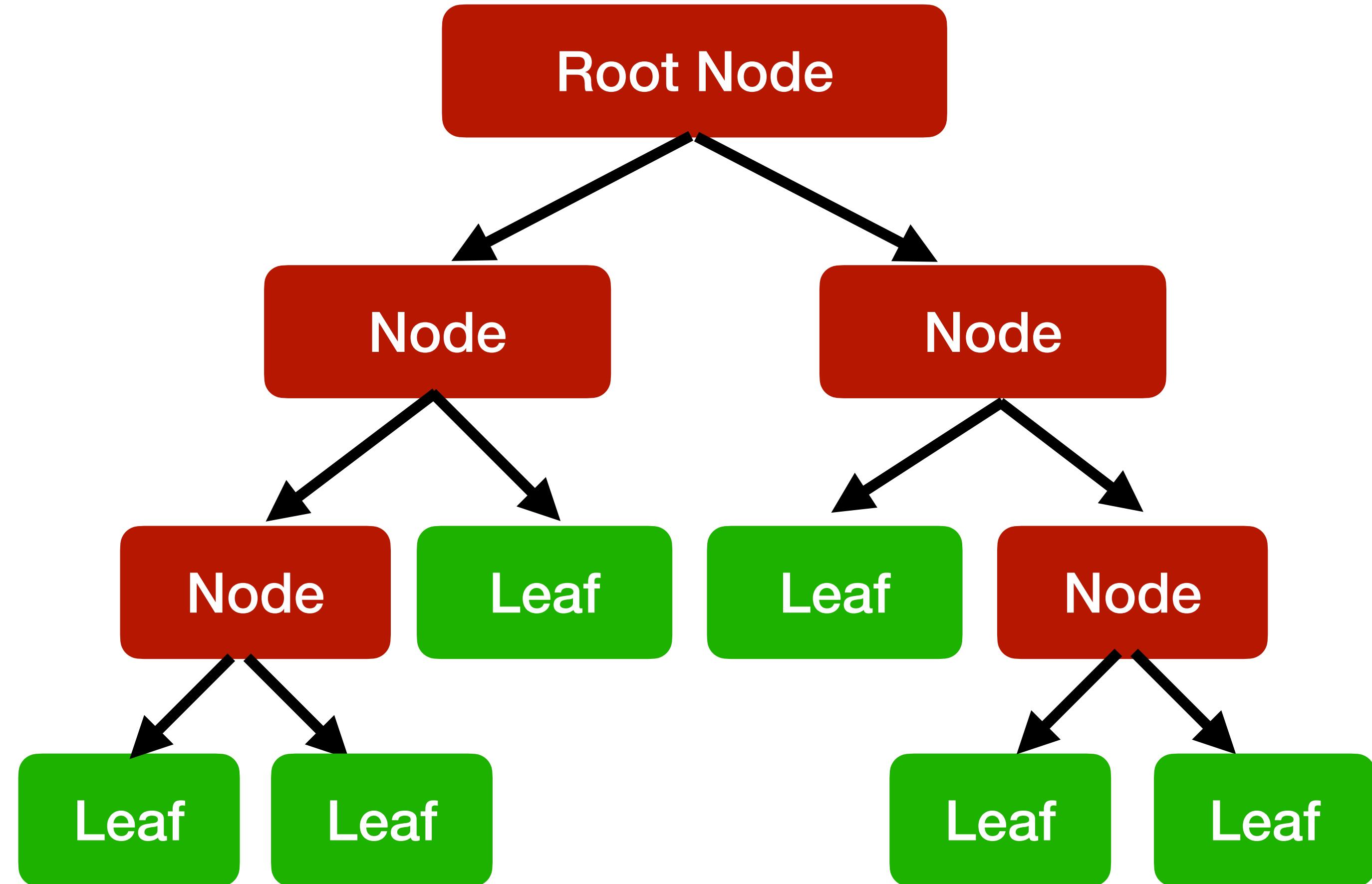


# Gini Impurity



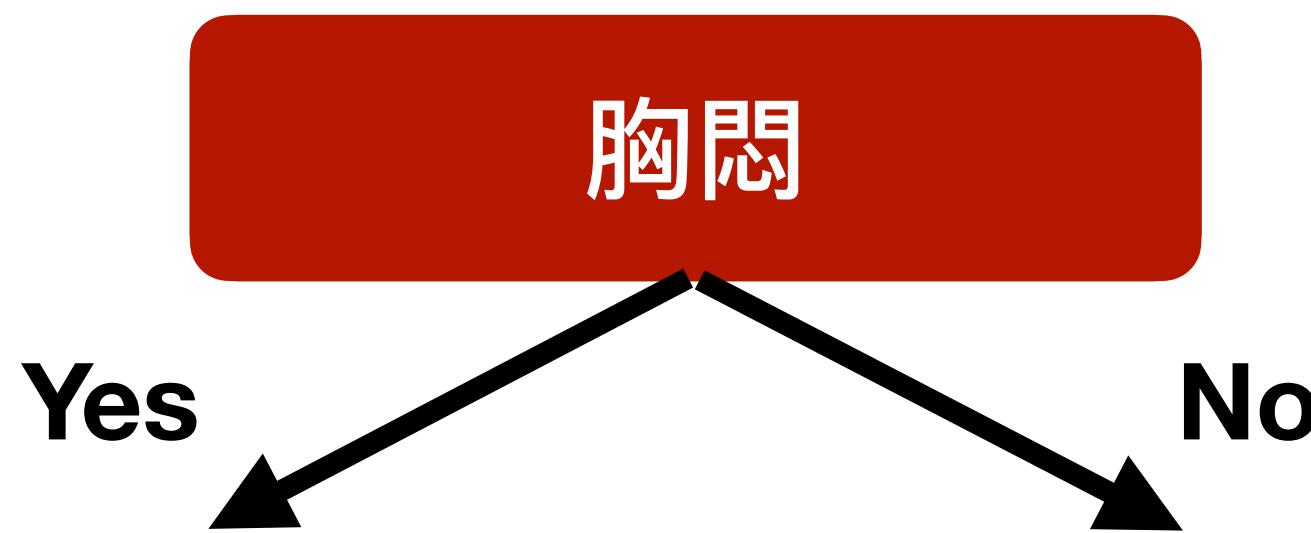
胸悶	好的血液循環	動脈阻塞	心臟疾病
No	No	No	No
Yes	Yes	Yes	Yes
Yes	Yes	No	No
Yes	No	No	Yes
...	...	...	...

胸悶	好的血液循環	動脈阻塞	心臟疾病
No	No	No	No
Yes	Yes	Yes	Yes
Yes	Yes	No	No
Yes	No	No	Yes
...	...	...	...

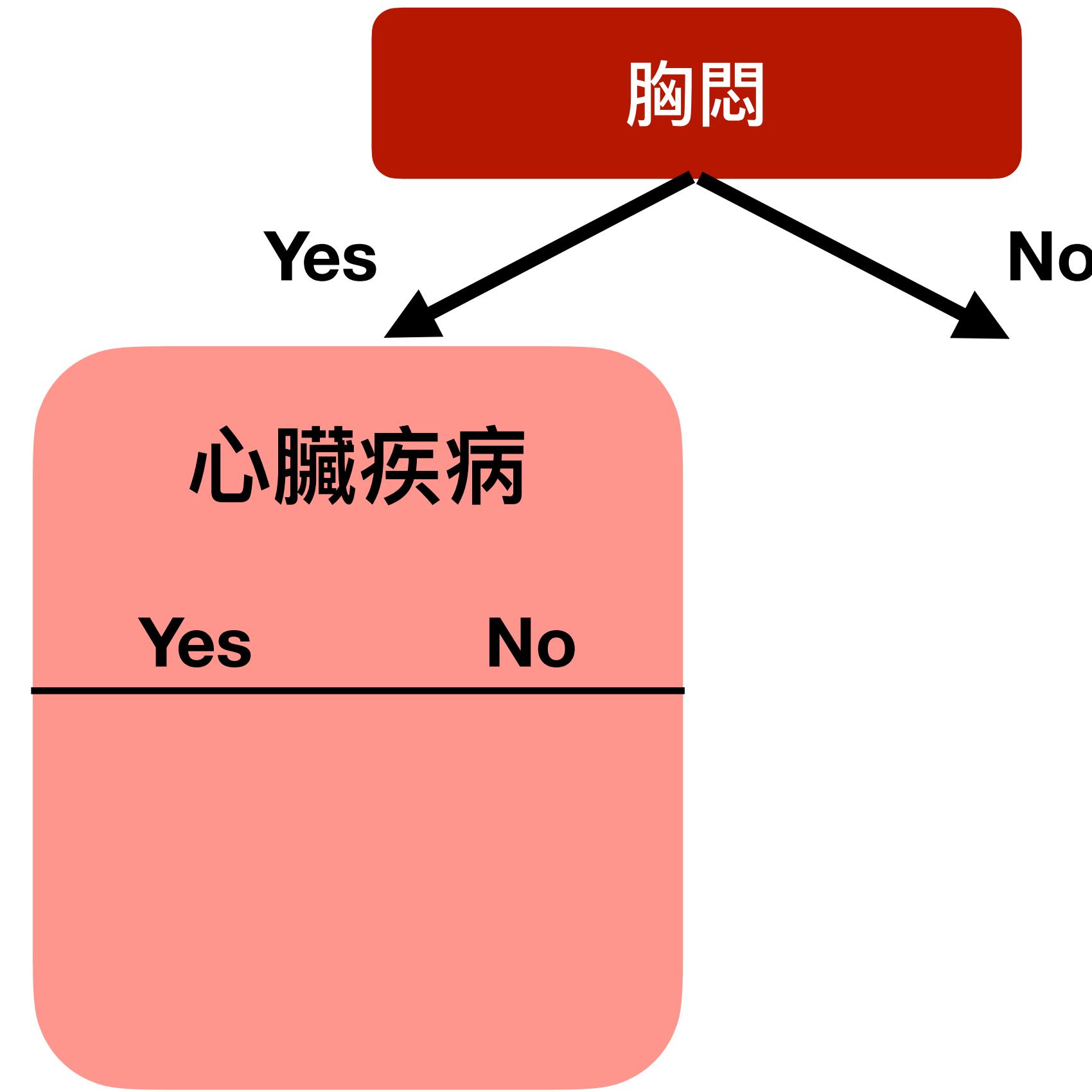


胸悶			心臟疾病
No			No
Yes			Yes
Yes			No
Yes			Yes
...			...

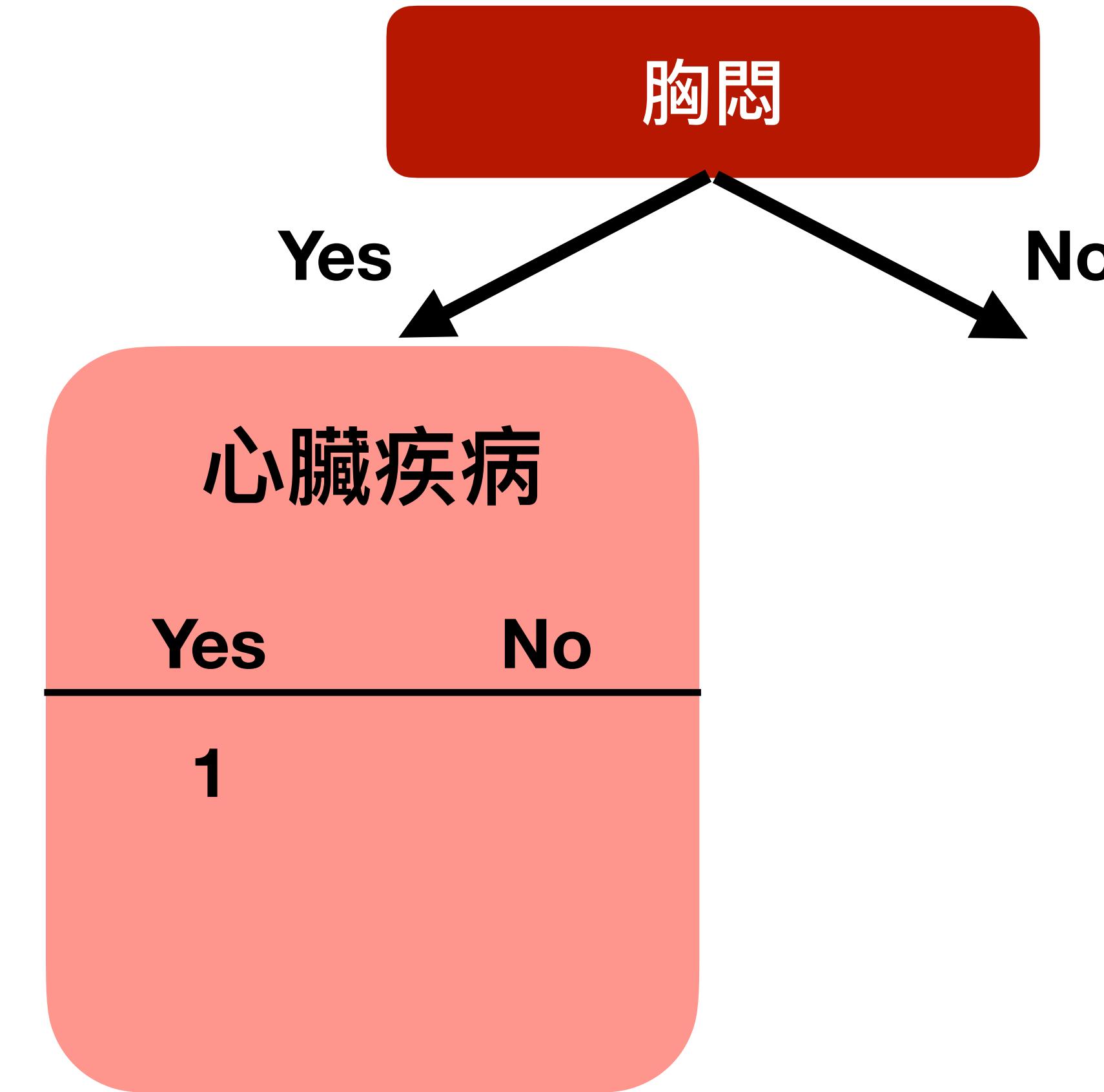
胸悶			心臟疾病
No			No
Yes			Yes
Yes			No
Yes			Yes
...			...



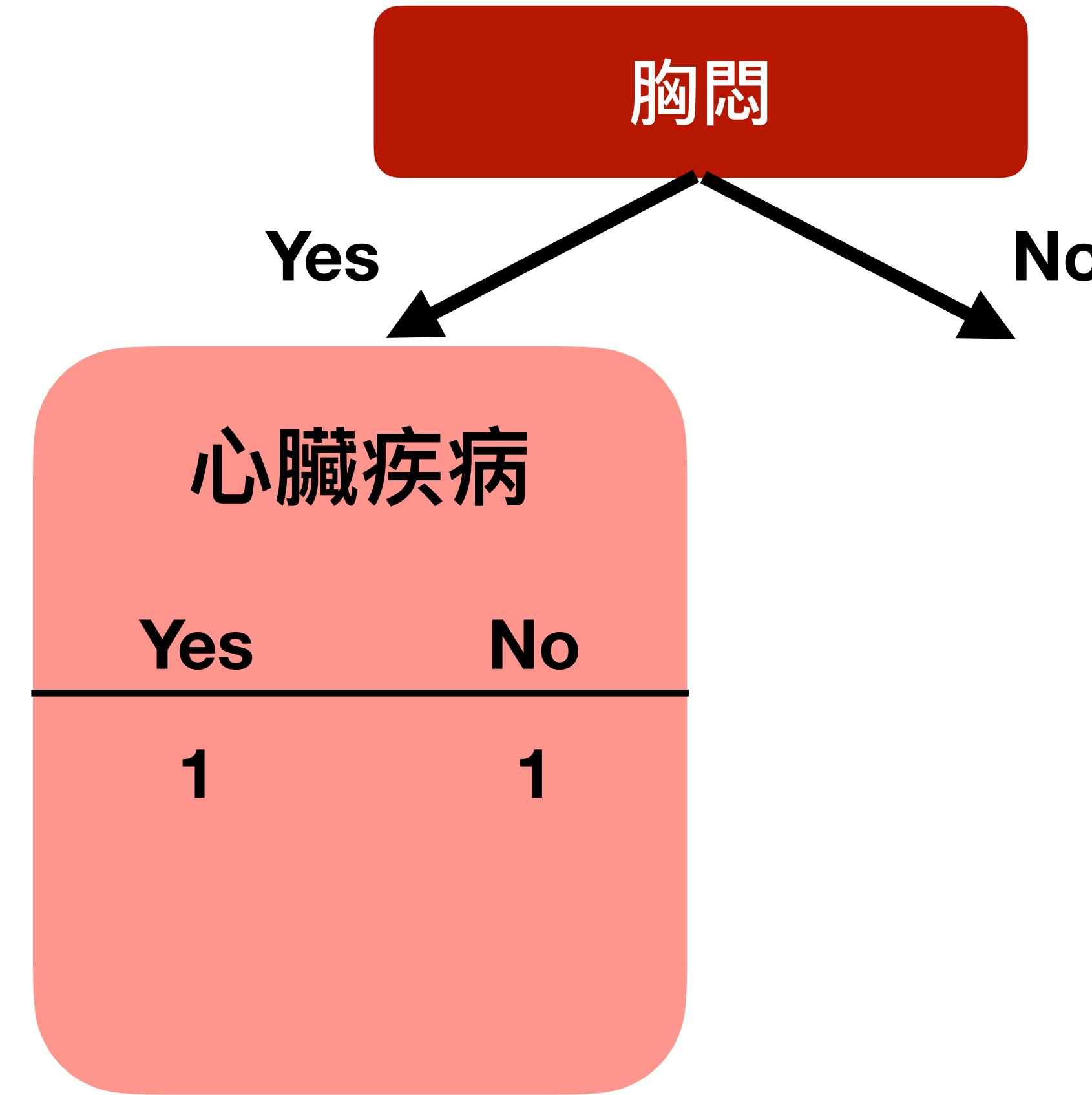
胸悶			心臟疾病
No			No
Yes			Yes
Yes			No
Yes			Yes
...			...



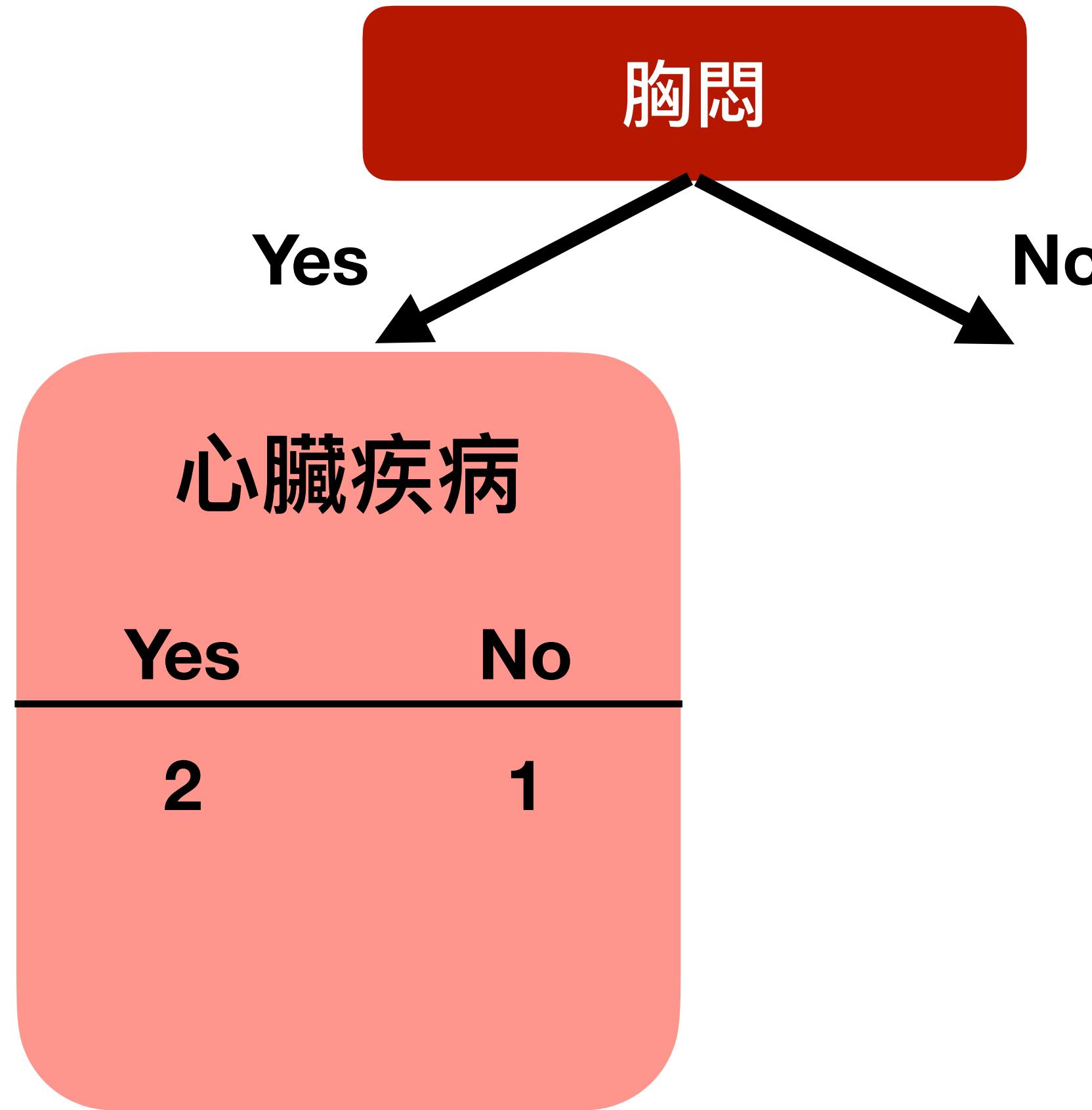
胸悶			心臟疾病
No			No
Yes			Yes
Yes			No
Yes			Yes
...			...



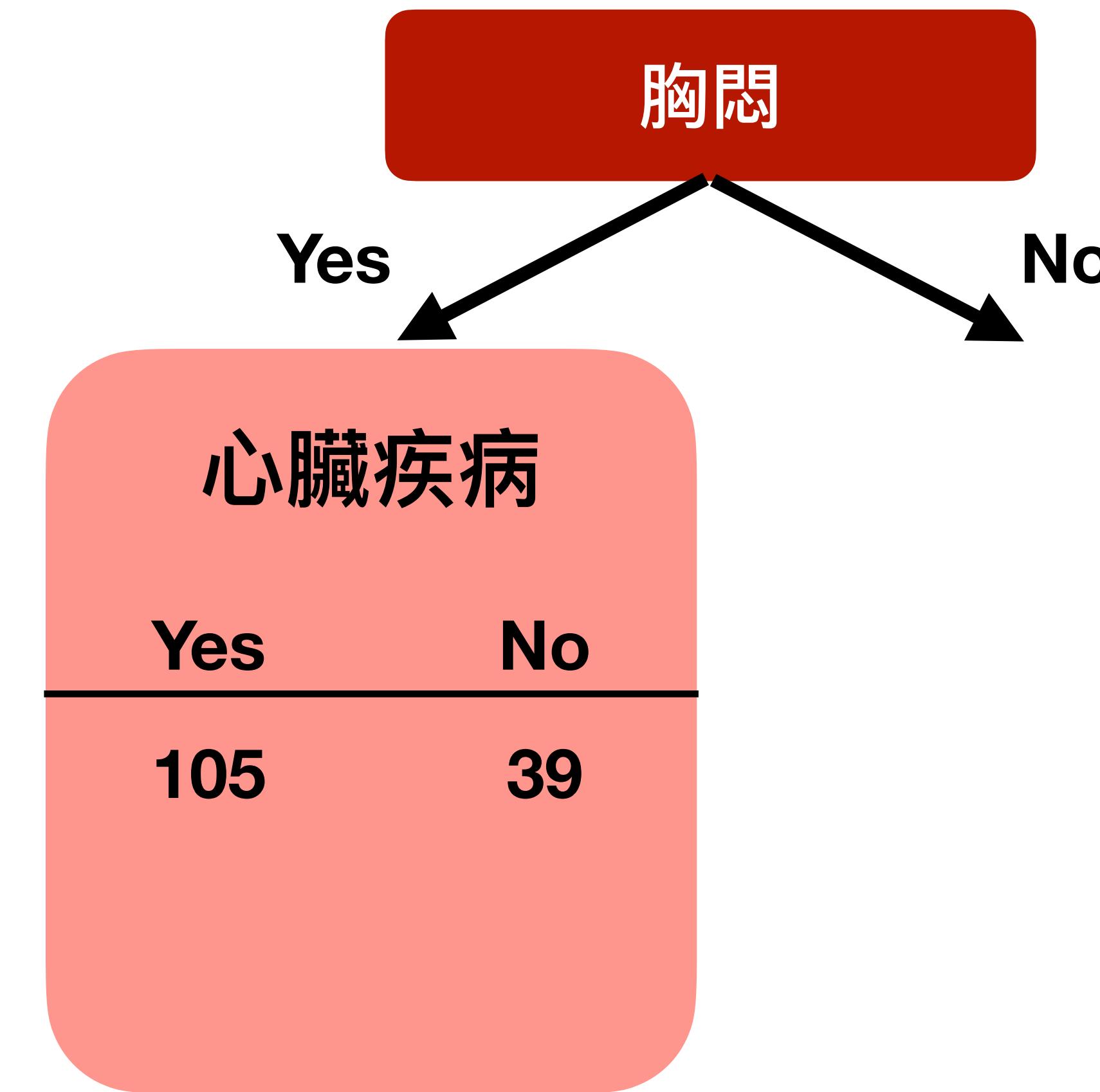
胸悶			心臟疾病
No			No
Yes			Yes
Yes			No
Yes			Yes
...			...



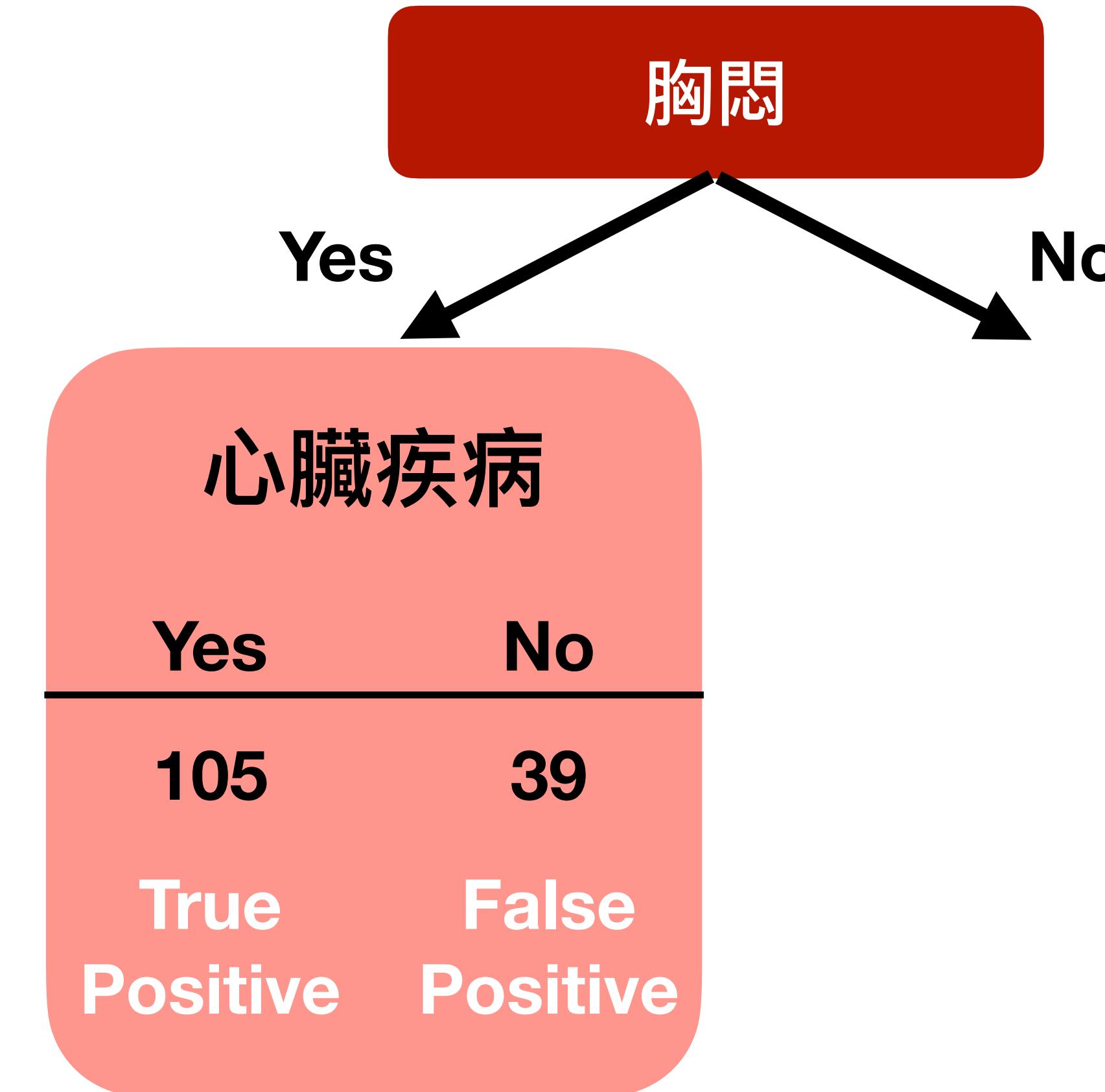
胸悶			心臟疾病
No			No
Yes			Yes
Yes			No
Yes			Yes
...			...



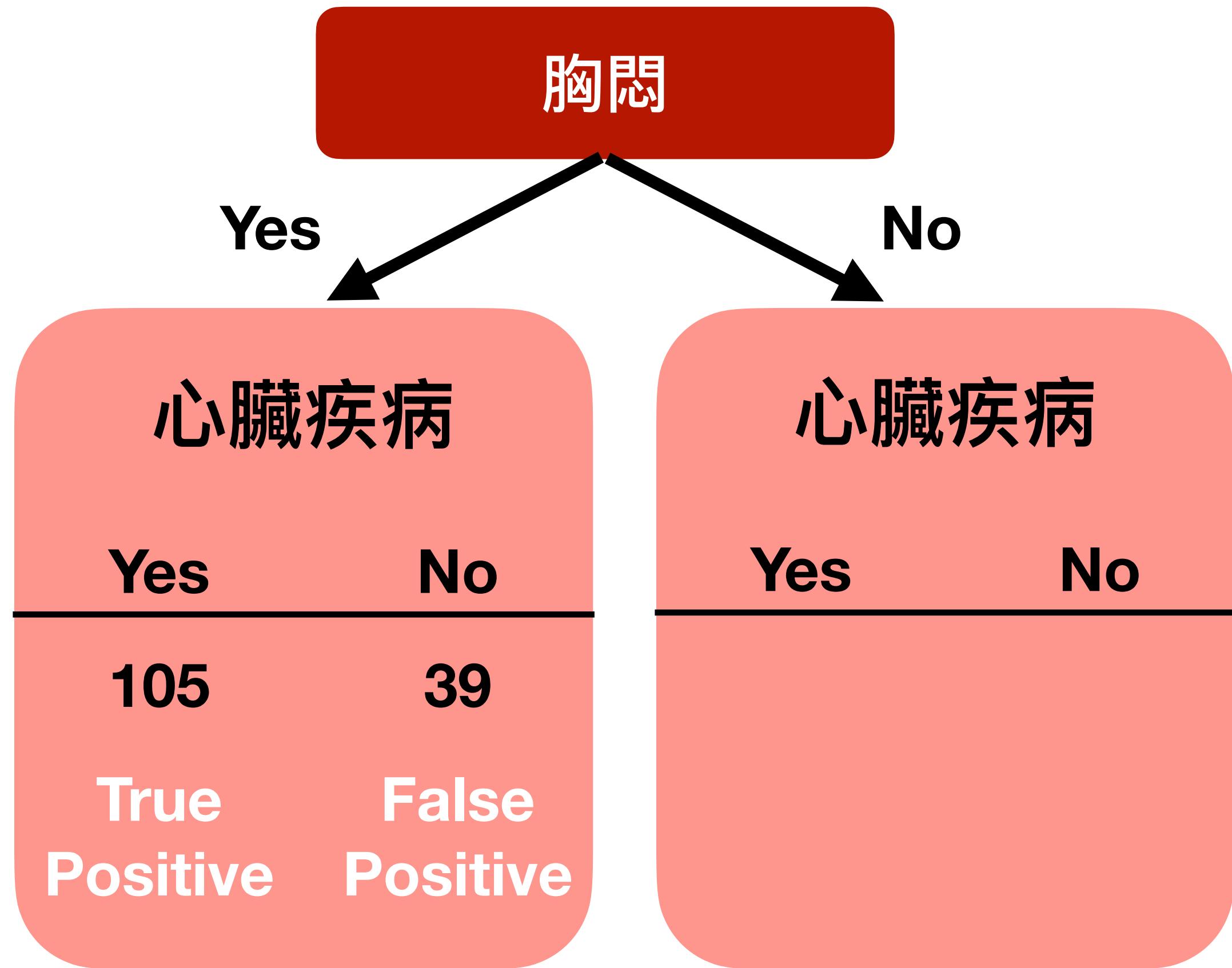
胸悶			心臟疾病
No			No
Yes			Yes
Yes			No
Yes			Yes
...			...



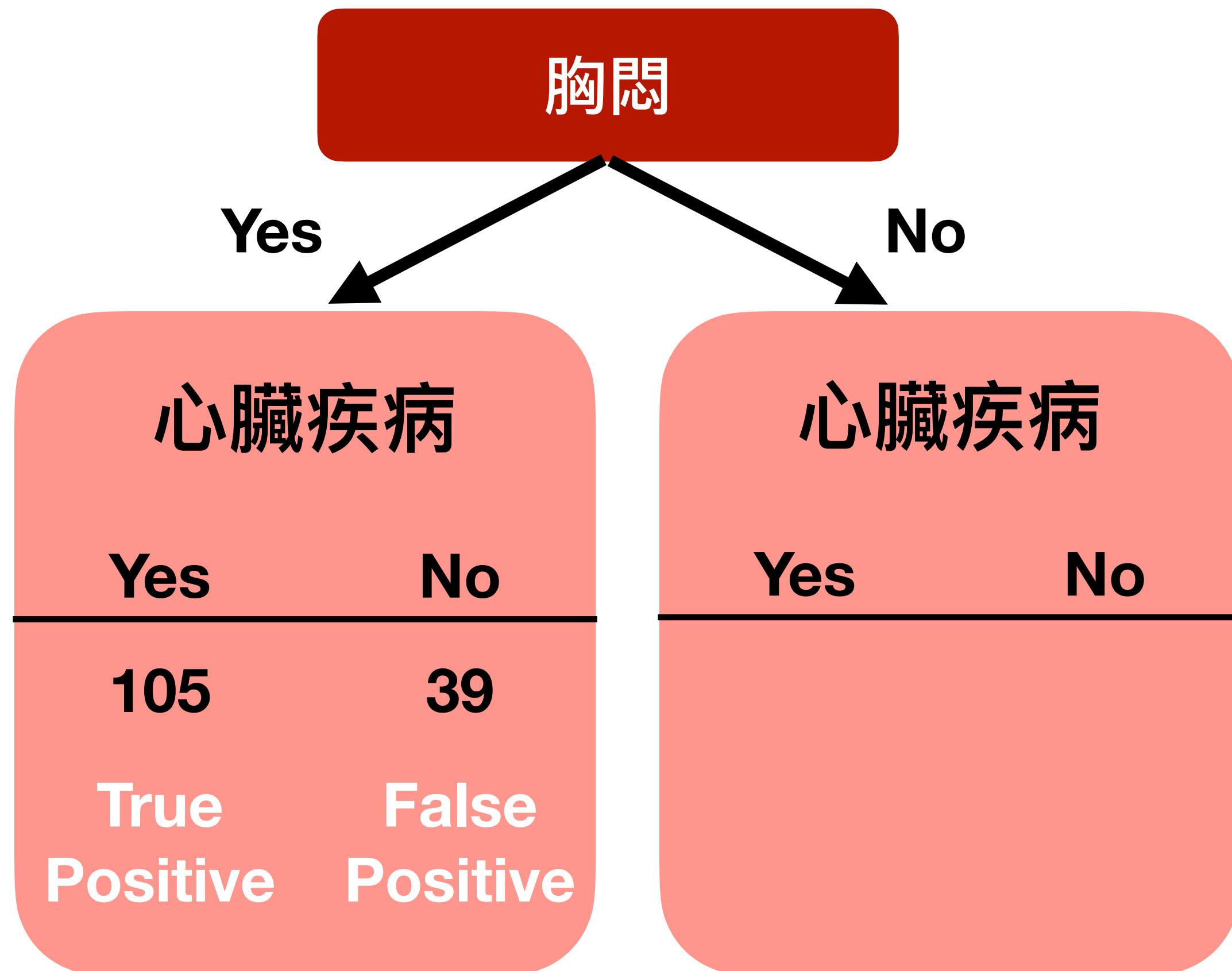
胸悶			心臟疾病
No			No
Yes			Yes
Yes			No
Yes			Yes
...			...



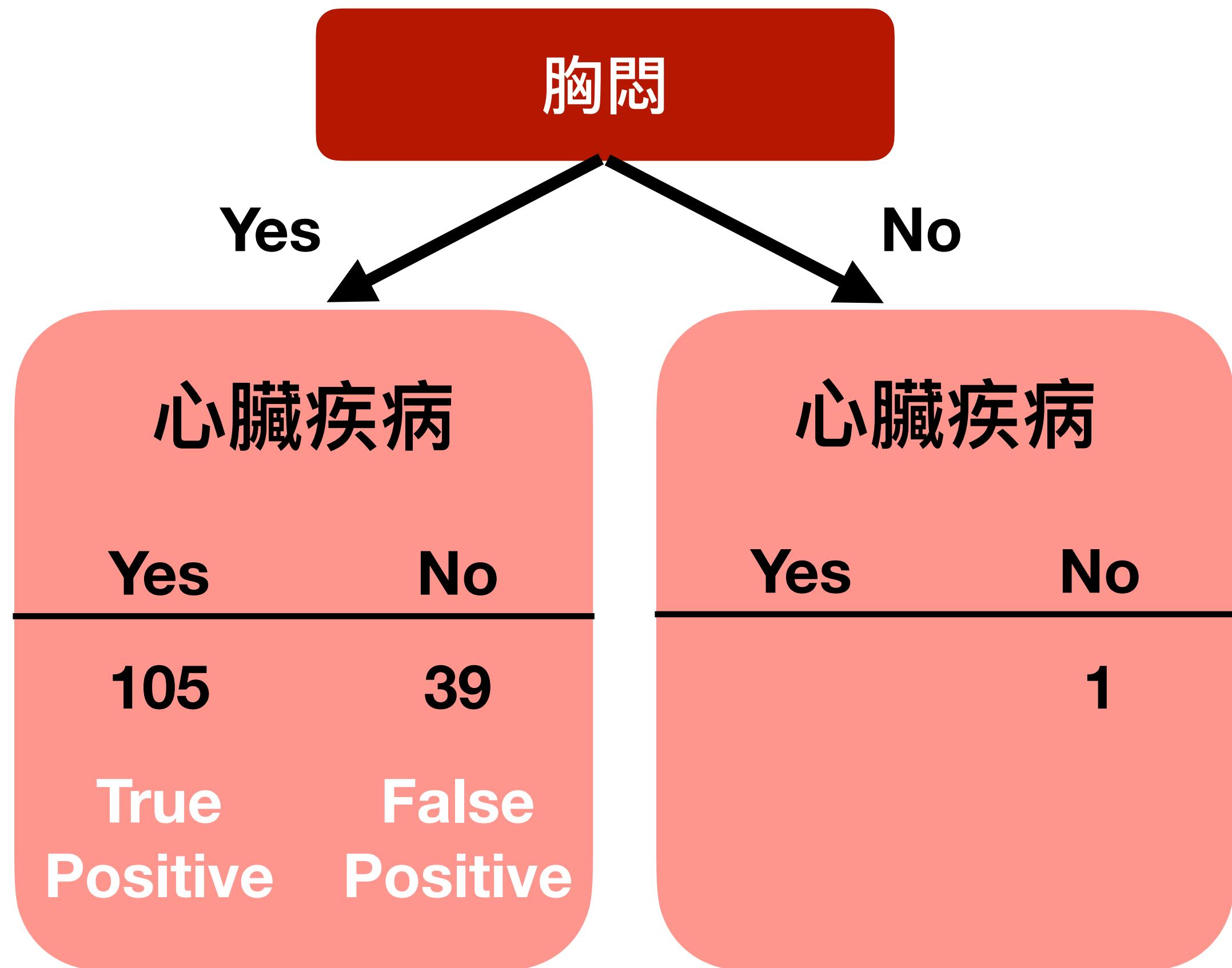
胸悶			心臟疾病
No			No
Yes			Yes
Yes			No
Yes			Yes
...			...



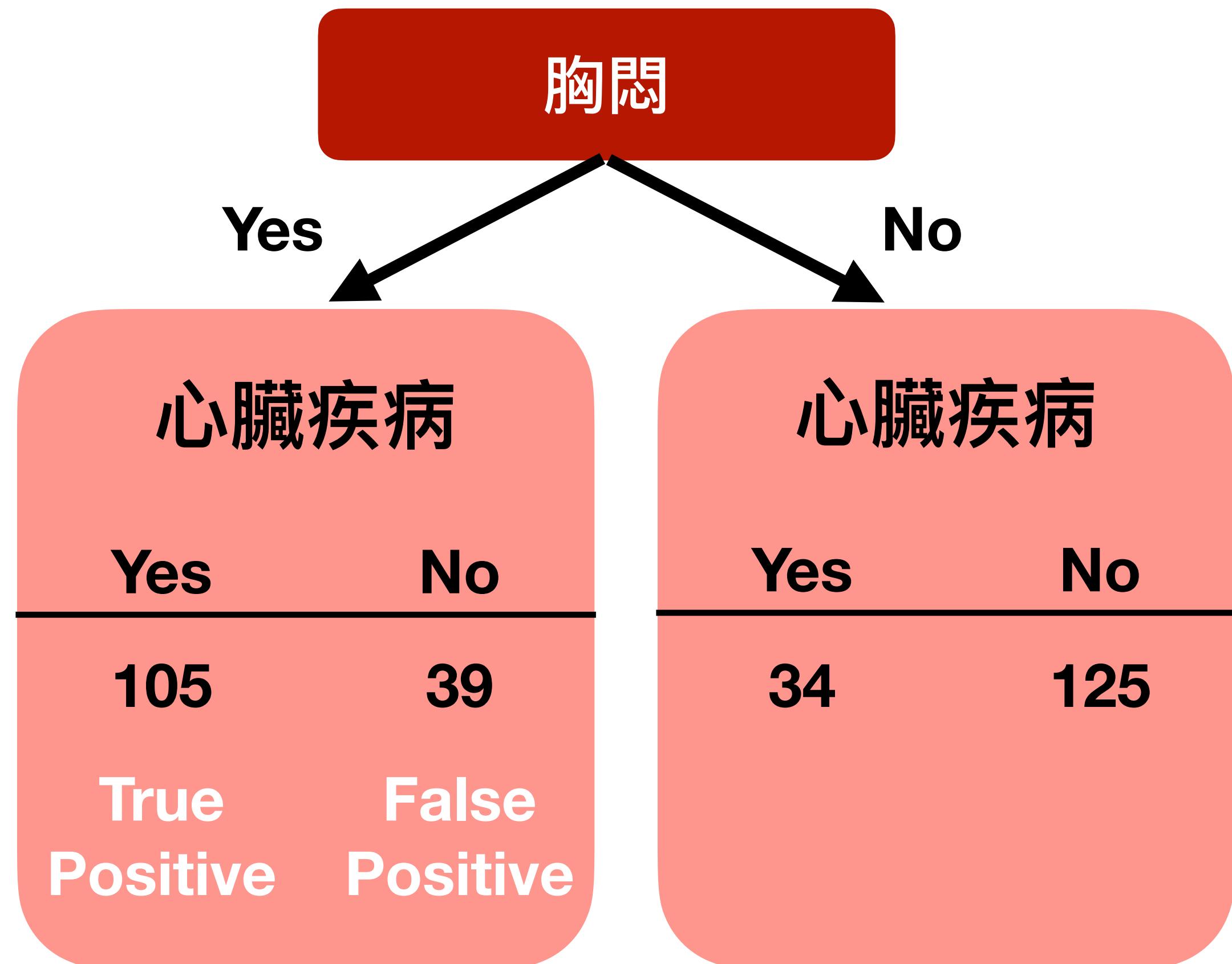
胸悶			心臟疾病
No			No
Yes			Yes
Yes			No
Yes			Yes
...			...



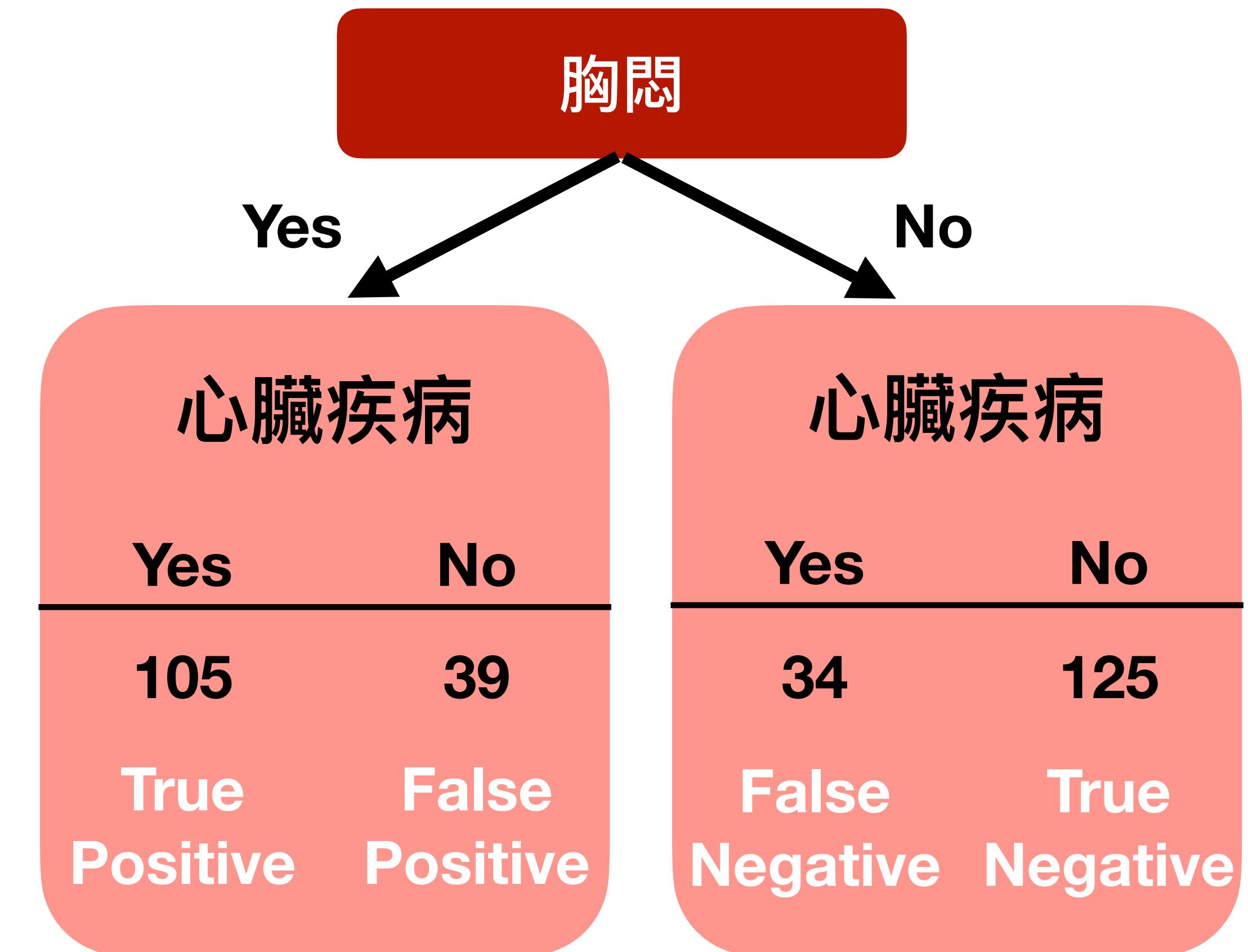
胸悶			心臟疾病
No			No
Yes			Yes
Yes			No
Yes			Yes
...			...

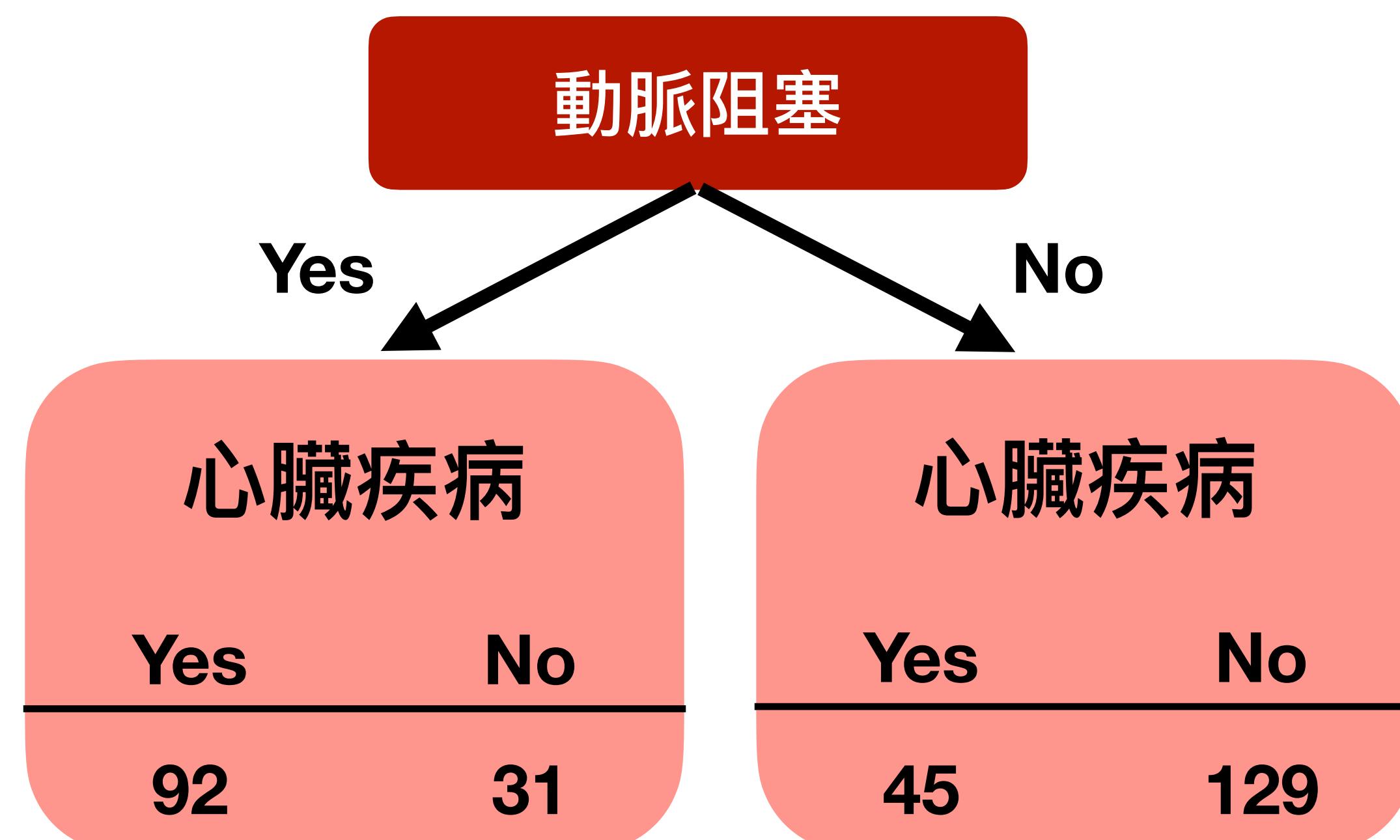
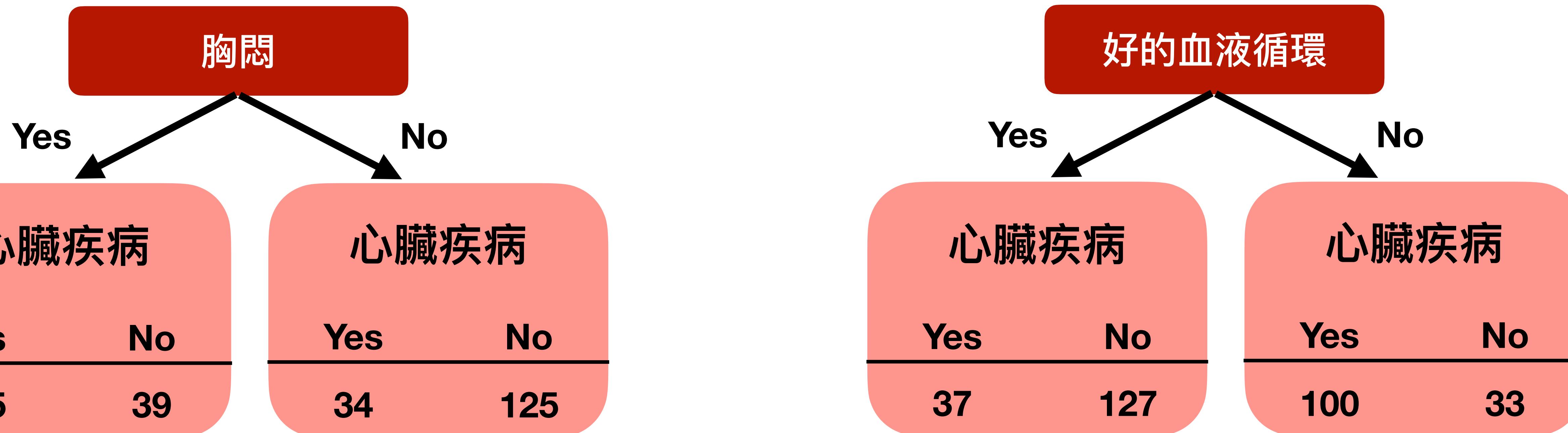


胸悶			心臟疾病
No			No
Yes			Yes
Yes			No
Yes			Yes
...			...



胸悶			心臟疾病
No			No
Yes			Yes
Yes			No
Yes			Yes
...			...

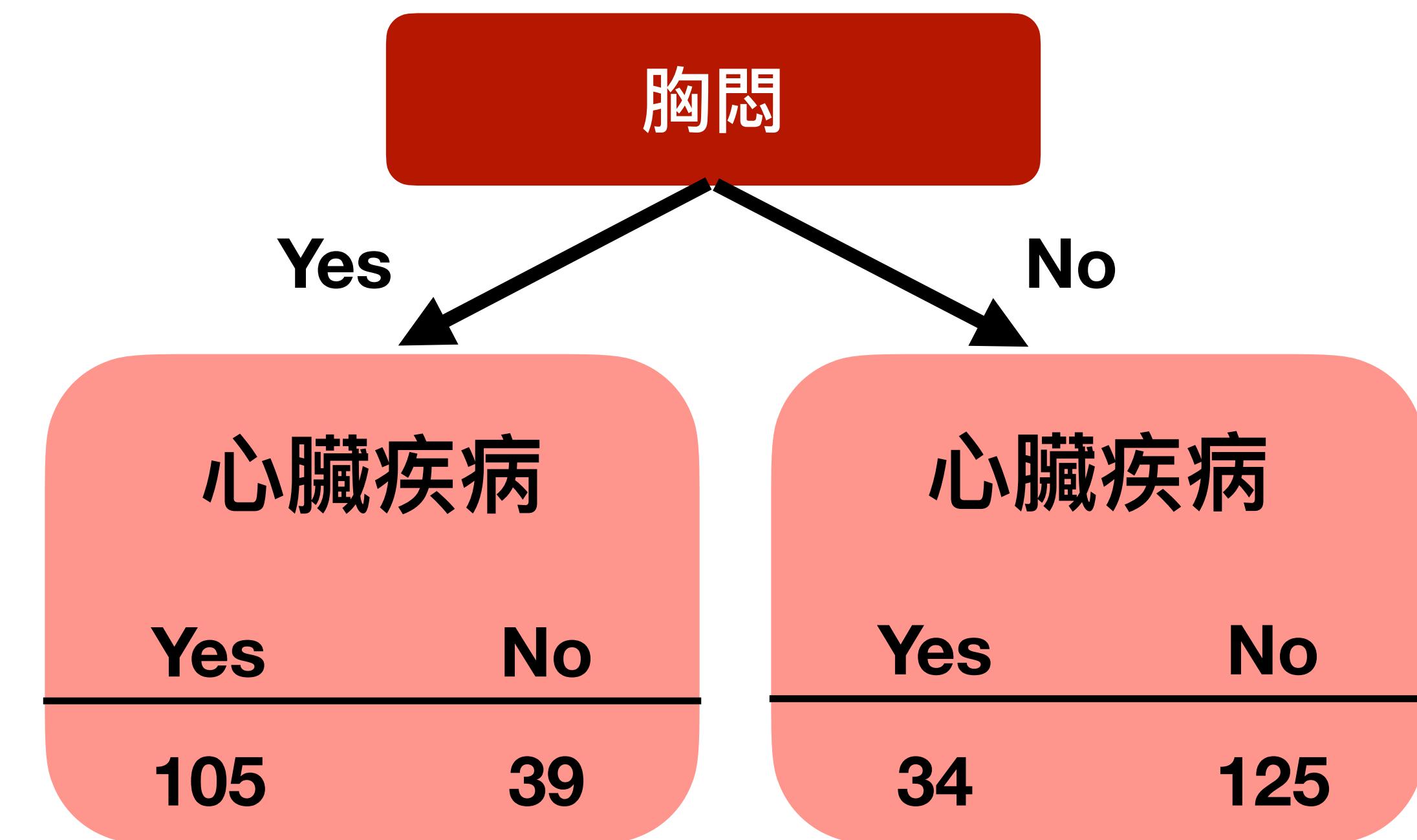


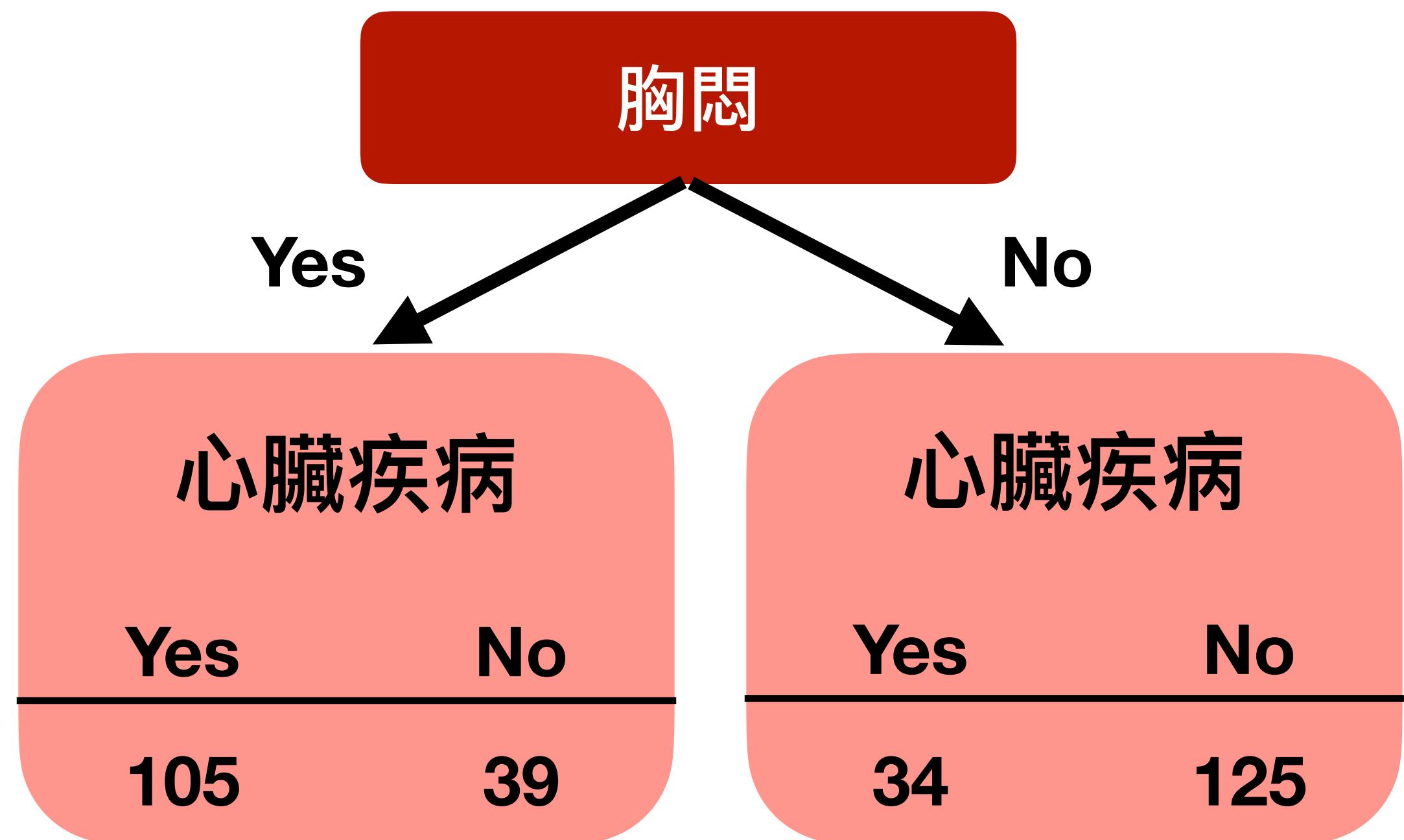


**None of the leaves are 100% “Yes” or “No”**

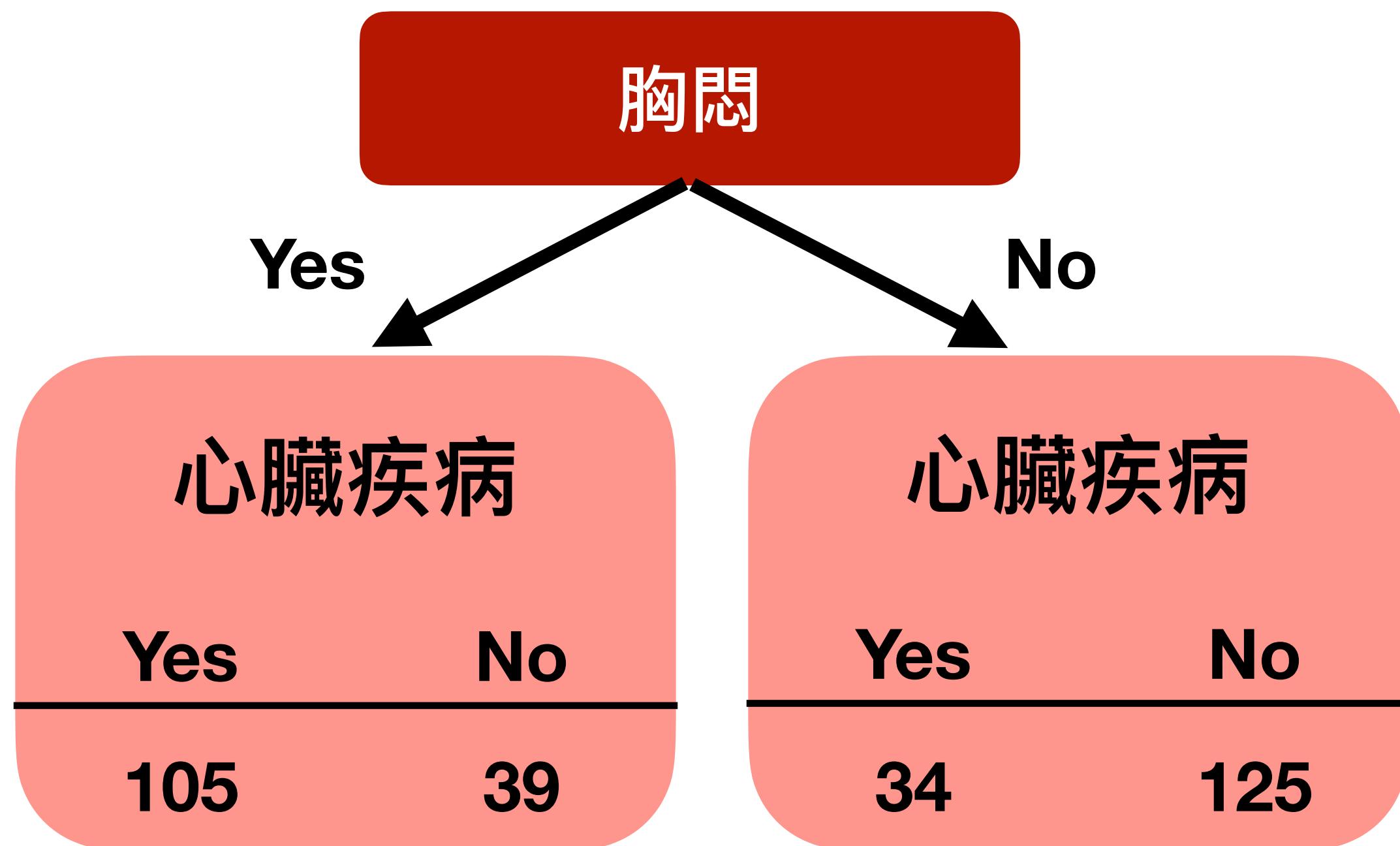
**Impurity**

**Gini Impurity =  $1 - (p(\text{yes}))^2 - (p(\text{no}))^2$**



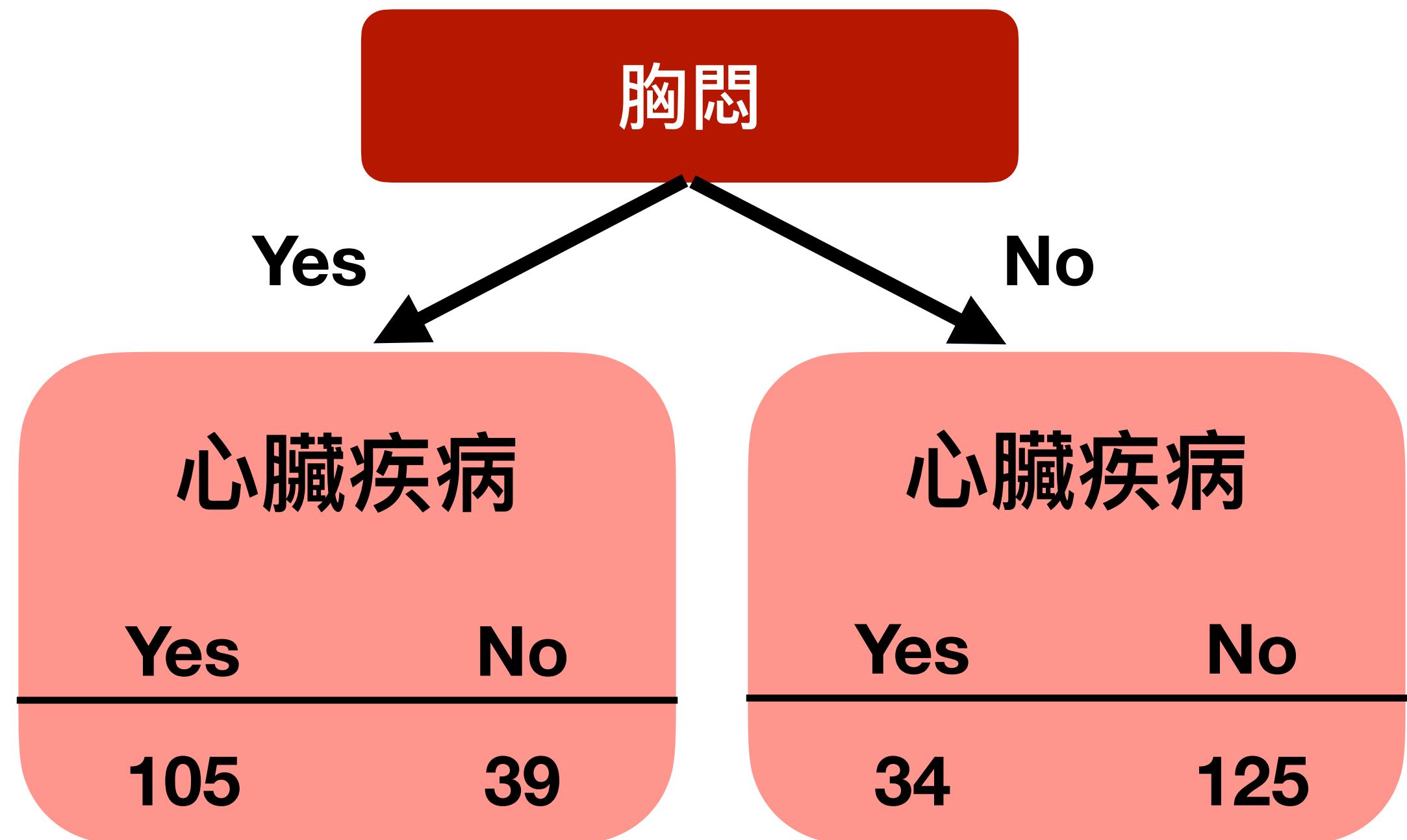


$$\text{Gini Impurity} = 1 - (105/144)^2 - (39/144)^2$$



$$\text{Gini Impurity} = 1 - (105/144)^2 - (39/144)^2$$

$$\text{Gini Impurity} = 0.395$$

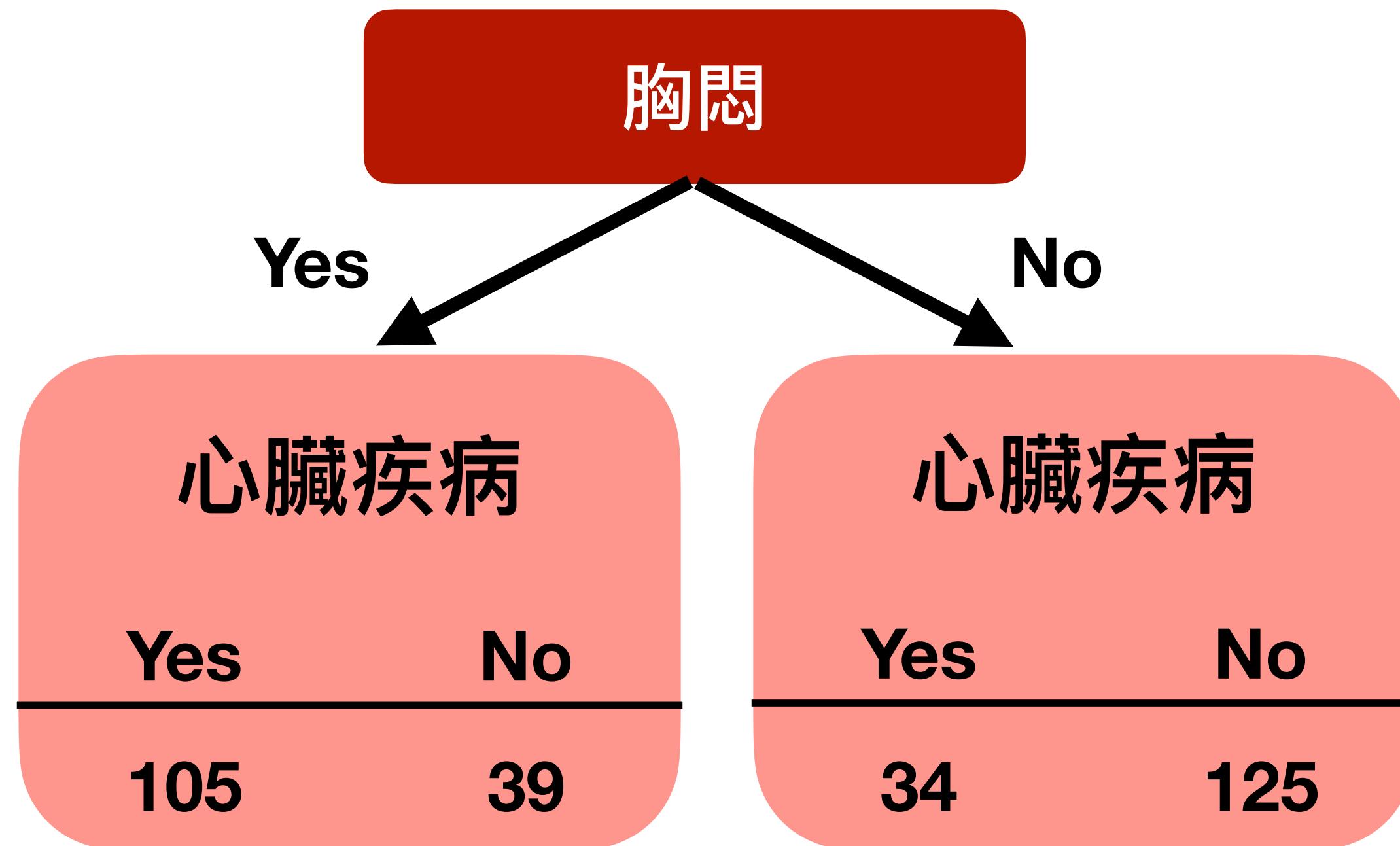


$$\text{Gini Impurity} = 1 - (105/144)^2 - (39/144)^2$$

$$\text{Gini Impurity} = 0.395$$

$$\text{Gini Impurity} = 1 - (34/159)^2 - (125/159)^2$$

$$\text{Gini Impurity} = 0.336$$



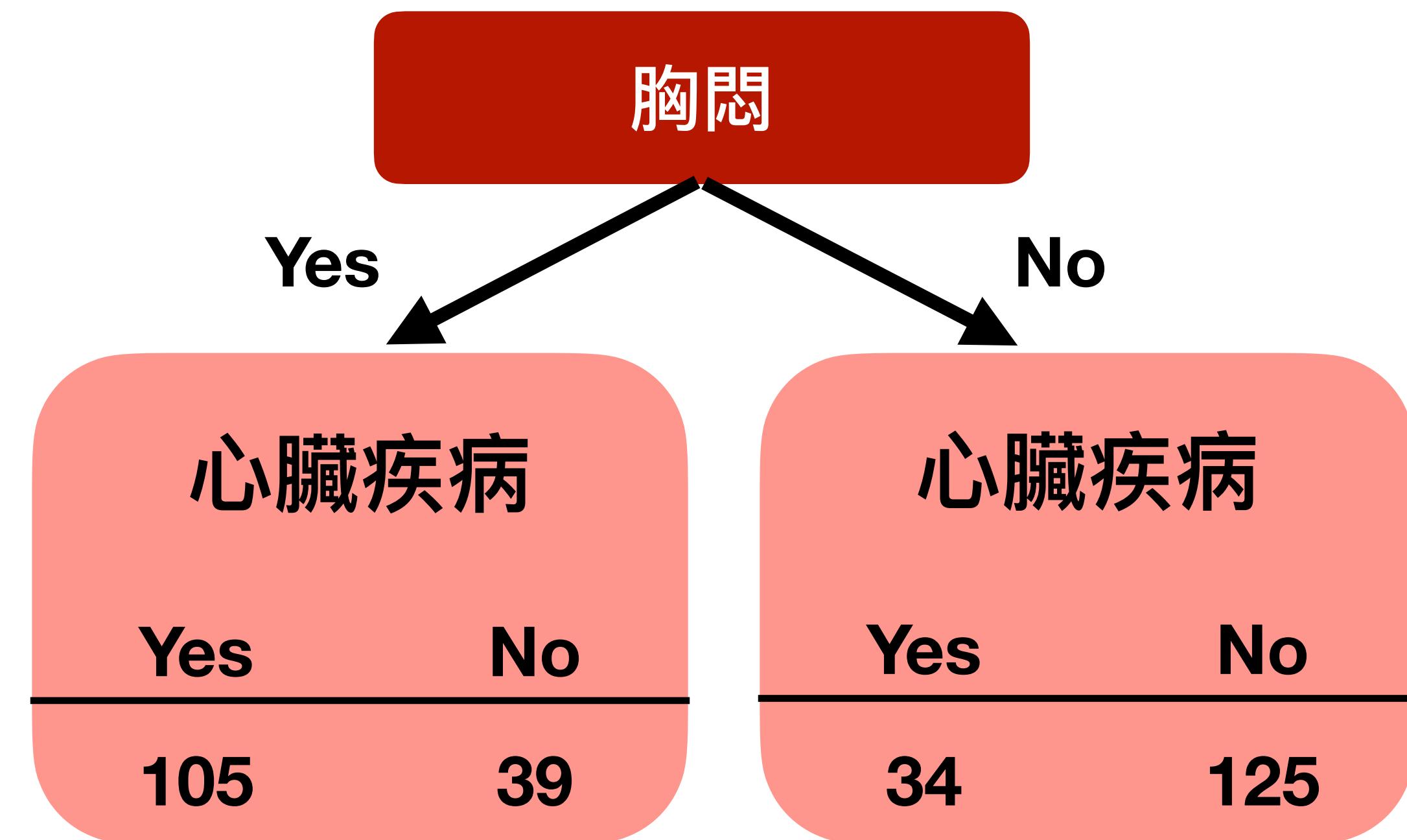
$$\text{Gini Impurity} = 1 - (105/144)^2 - (39/144)^2$$

$$\text{Gini Impurity} = 0.395$$

$$\text{Gini Impurity} = 1 - (34/159)^2 - (125/159)^2$$

$$\text{Gini Impurity} = 0.336$$

胸悶 Gini Impurity = weighted average(0.395, 0.336)



$$\text{Gini Impurity} = 1 - (105/144)^2 - (39/144)^2$$

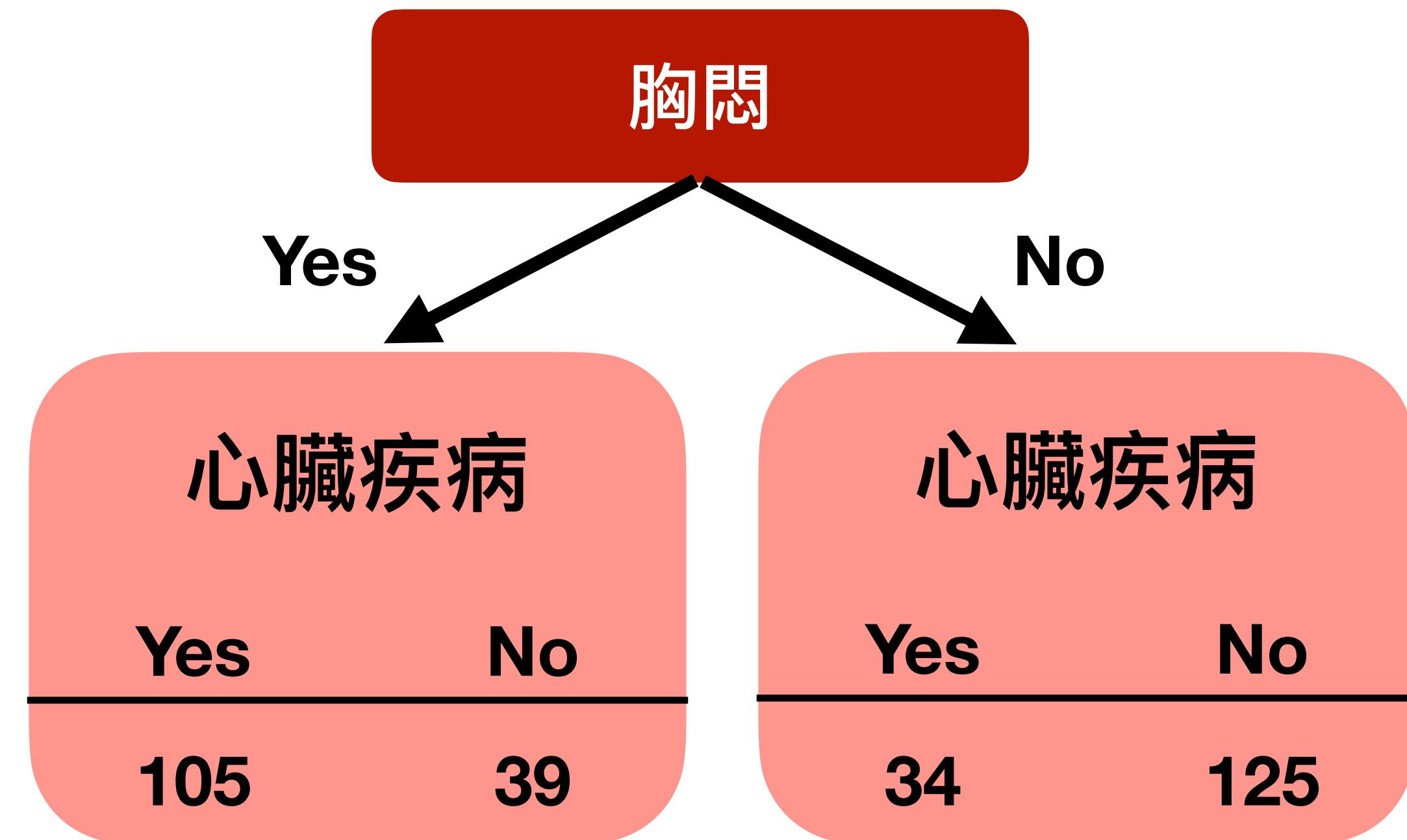
$$\text{Gini Impurity} = 0.395$$

$$\text{Gini Impurity} = 1 - (34/159)^2 - (125/159)^2$$

$$\text{Gini Impurity} = 0.336$$

胸悶 Gini Impurity = weighted average(0.395, 0.336)

$$\text{胸悶 Gini Impurity} = 0.395 * (144/303) + 0.336 * (159/303)$$



$$\text{Gini Impurity} = 1 - (105/144)^2 - (39/144)^2$$

$$\text{Gini Impurity} = 0.395$$

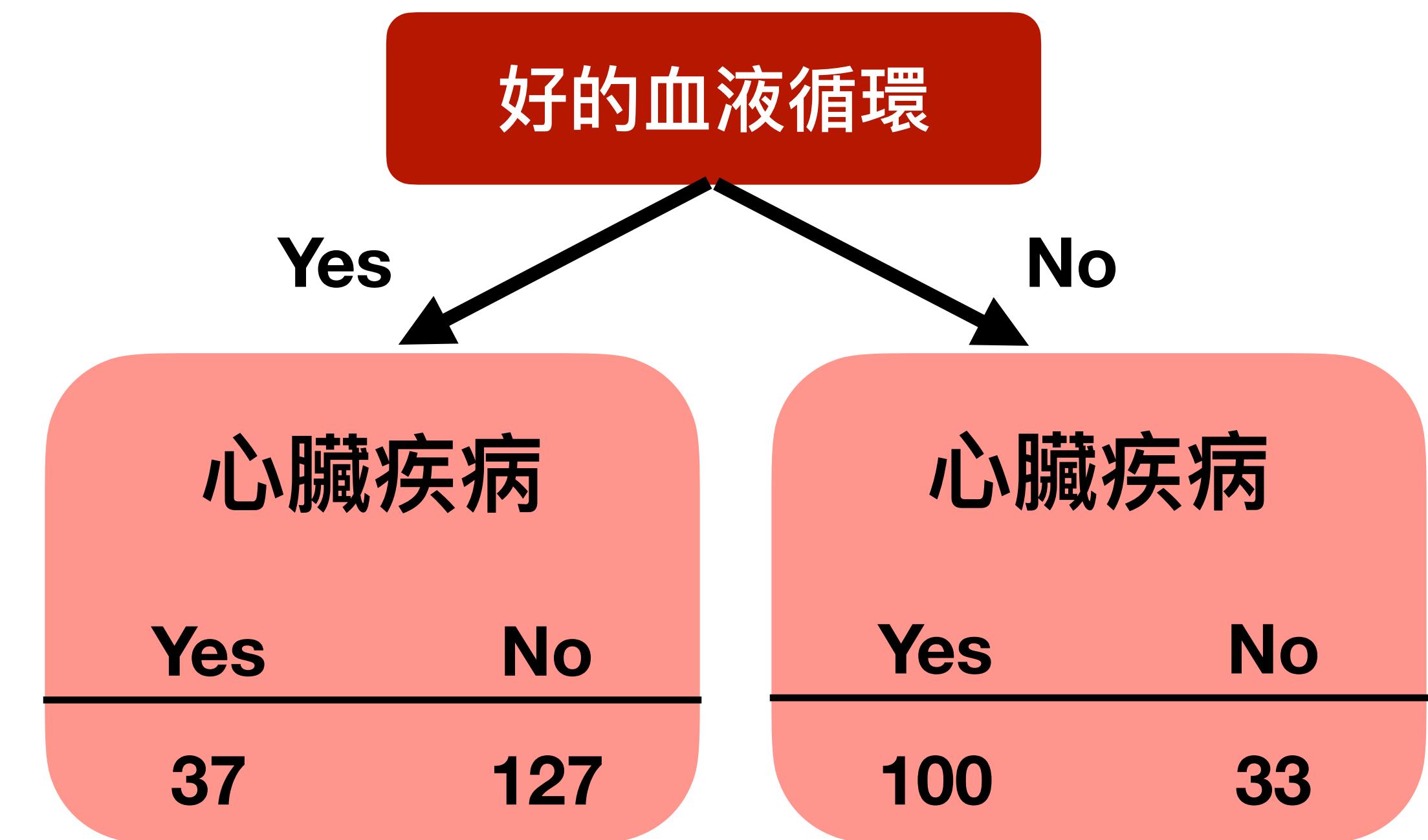
$$\text{Gini Impurity} = 1 - (34/159)^2 - (125/159)^2$$

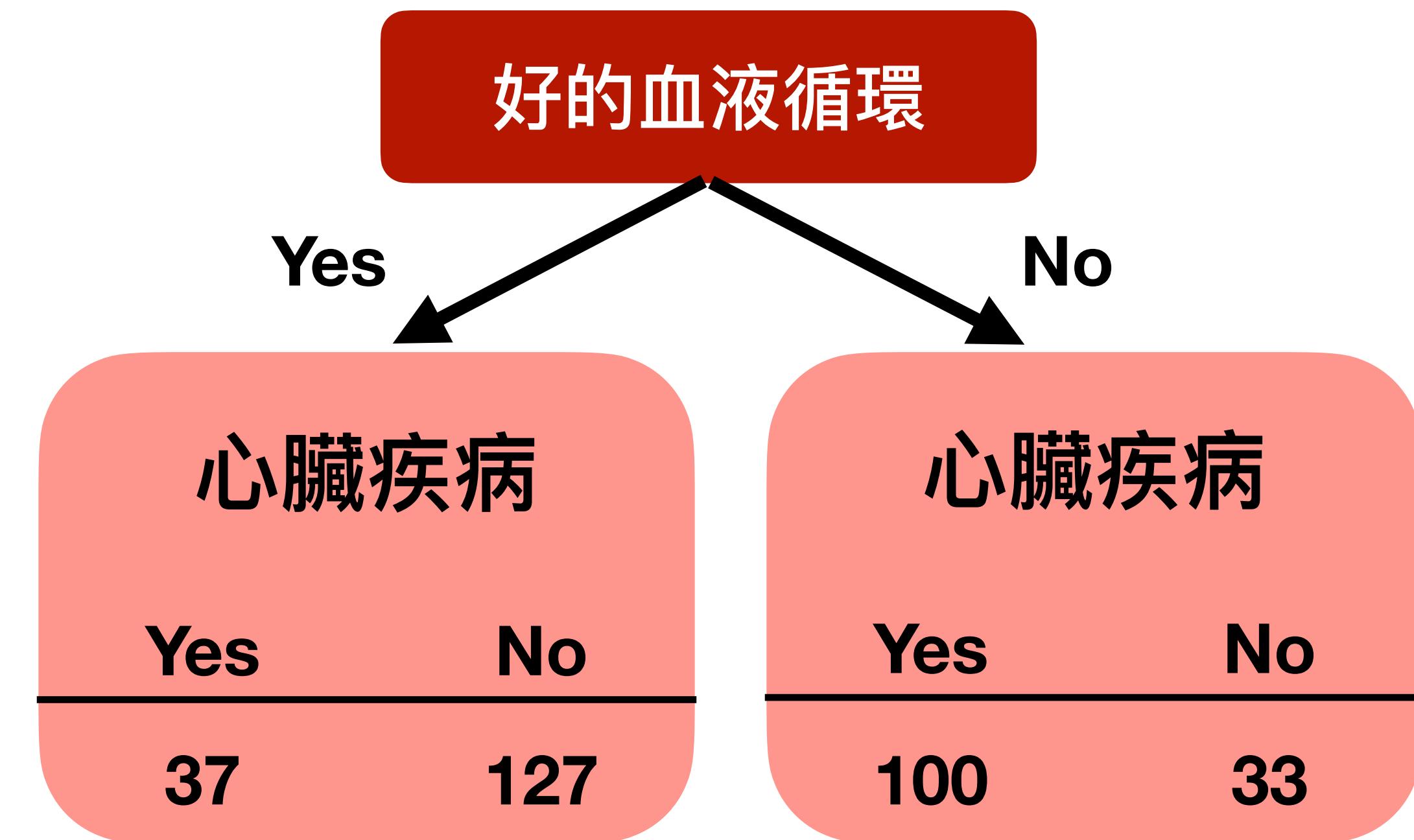
$$\text{Gini Impurity} = 0.336$$

胸悶 Gini Impurity = weighted average(0.395, 0.336)

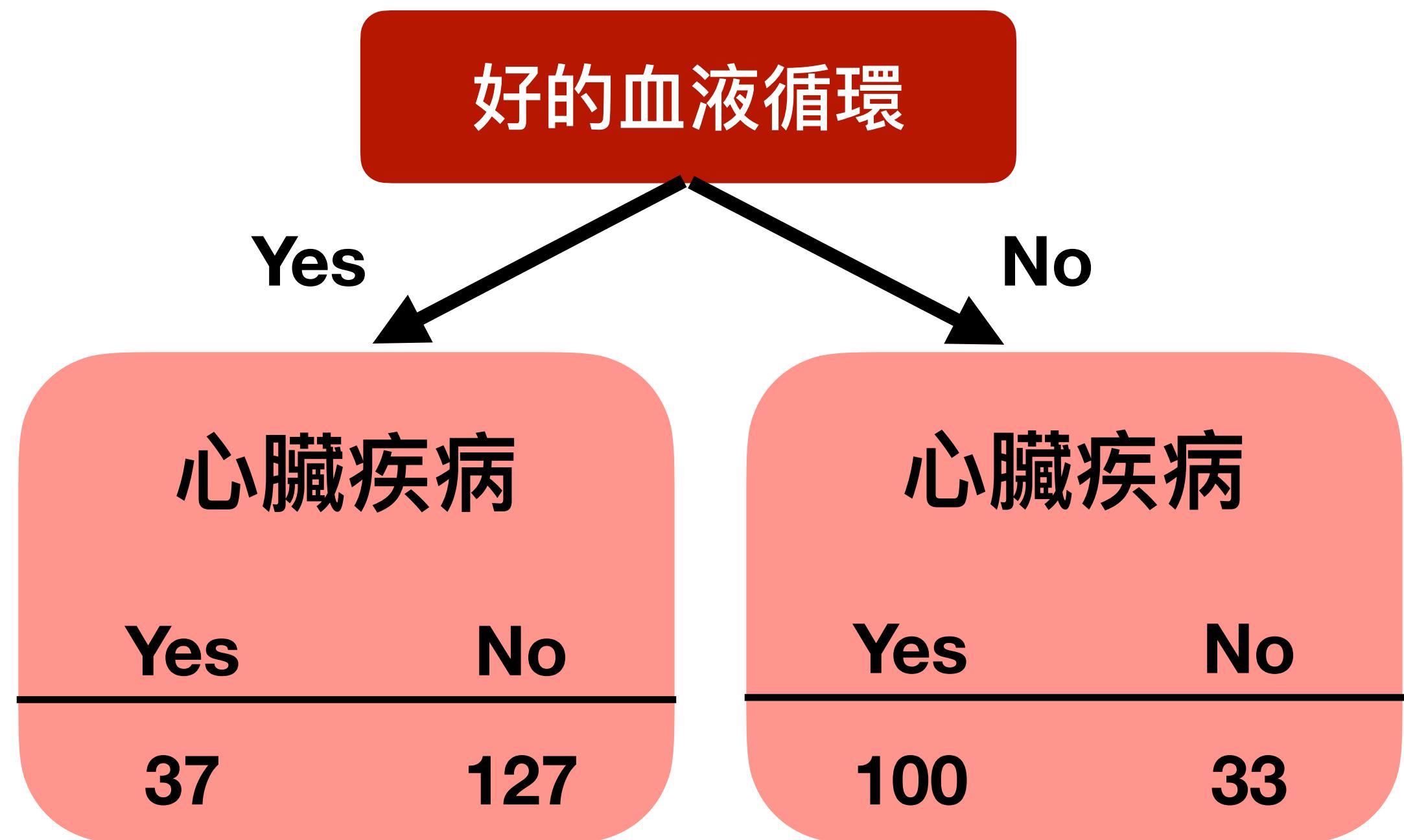
$$\text{胸悶 Gini Impurity} = 0.395 * (144/303) + 0.336 * (159/303)$$

$$\text{胸悶 Gini Impurity} = 0.364$$



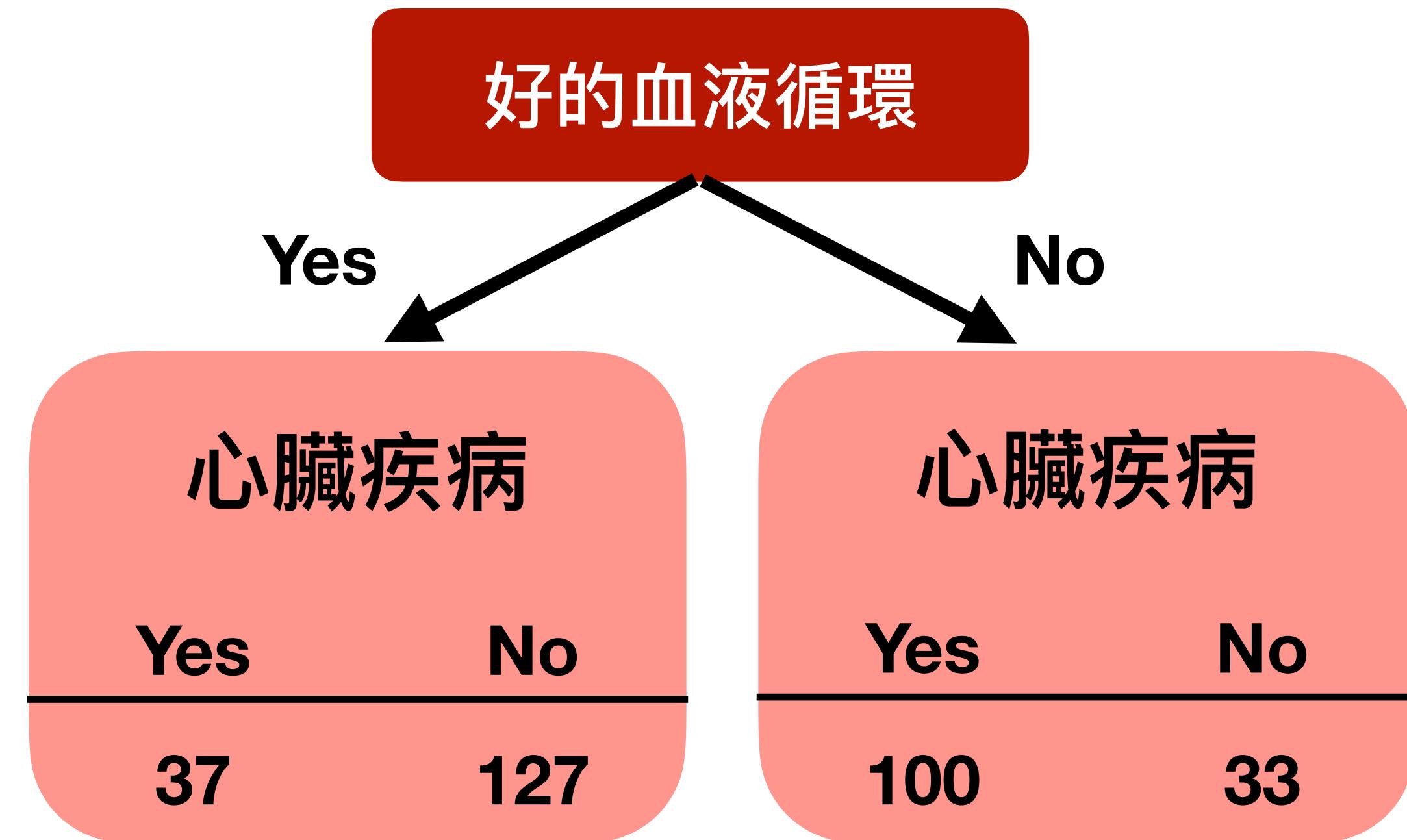


$$\text{Gini Impurity} = 1 - (37/164)^2 - (127/164)^2$$



$$\text{Gini Impurity} = 1 - (37/164)^2 - (127/164)^2$$

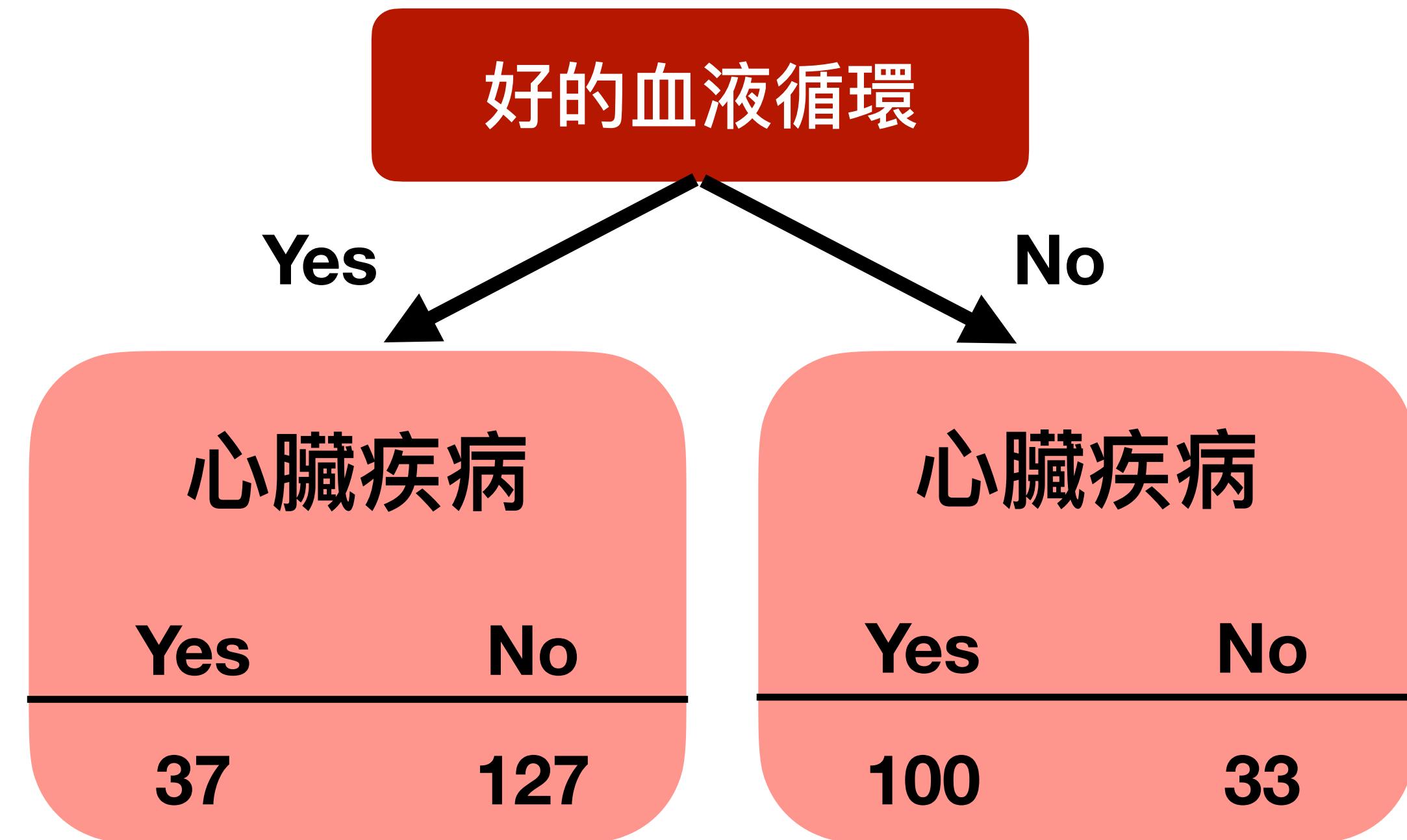
$$\text{Gini Impurity} = 0.349$$



$$\text{Gini Impurity} = 1 - (37/164)^2 - (127/164)^2$$

$$\text{Gini Impurity} = 1 - (100/133)^2 - (33/133)^2$$

$$\text{Gini Impurity} = 0.349$$

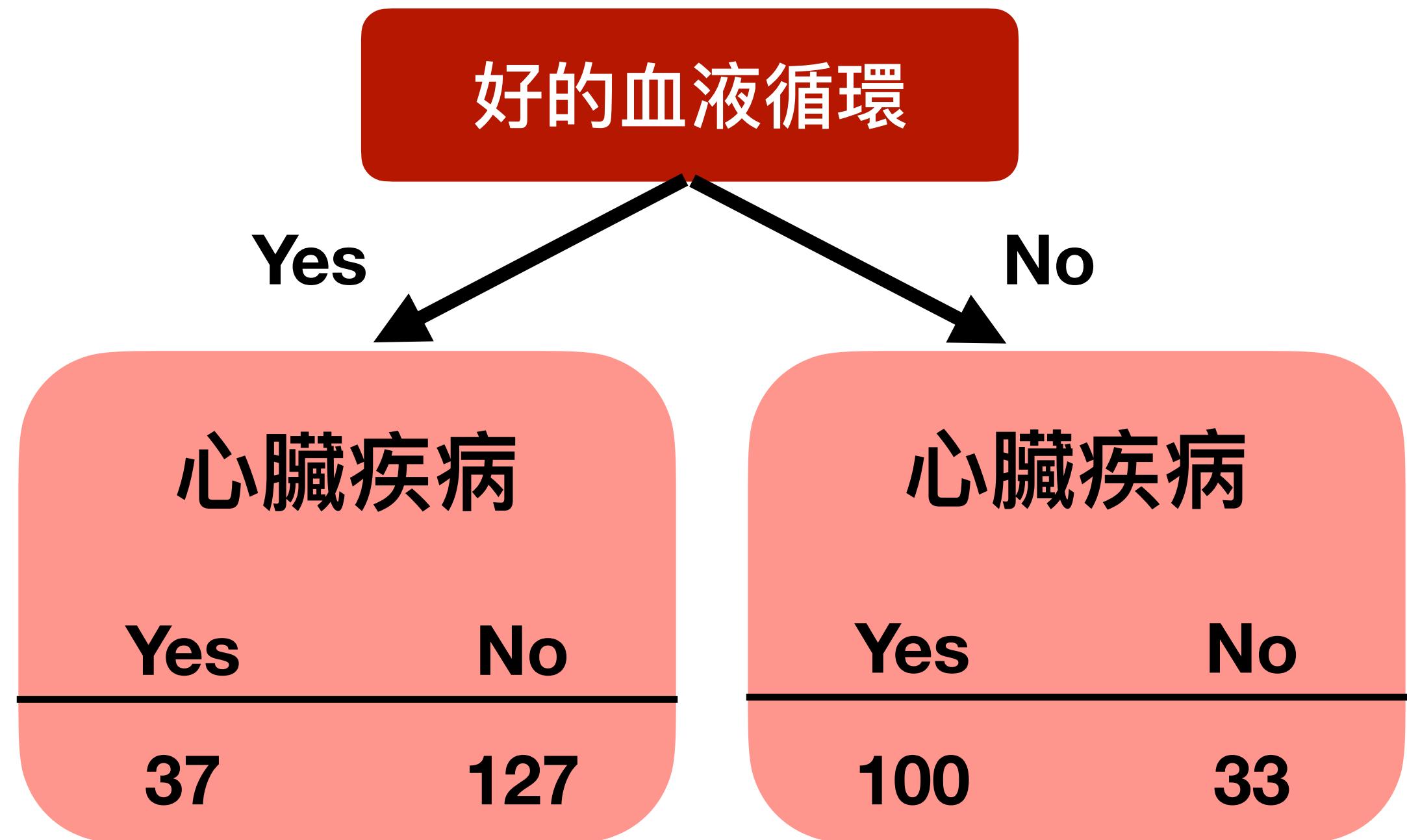


$$\text{Gini Impurity} = 1 - (37/164)^2 - (127/164)^2$$

$$\text{Gini Impurity} = 0.349$$

$$\text{Gini Impurity} = 1 - (100/133)^2 - (33/133)^2$$

$$\text{Gini Impurity} = 0.373$$



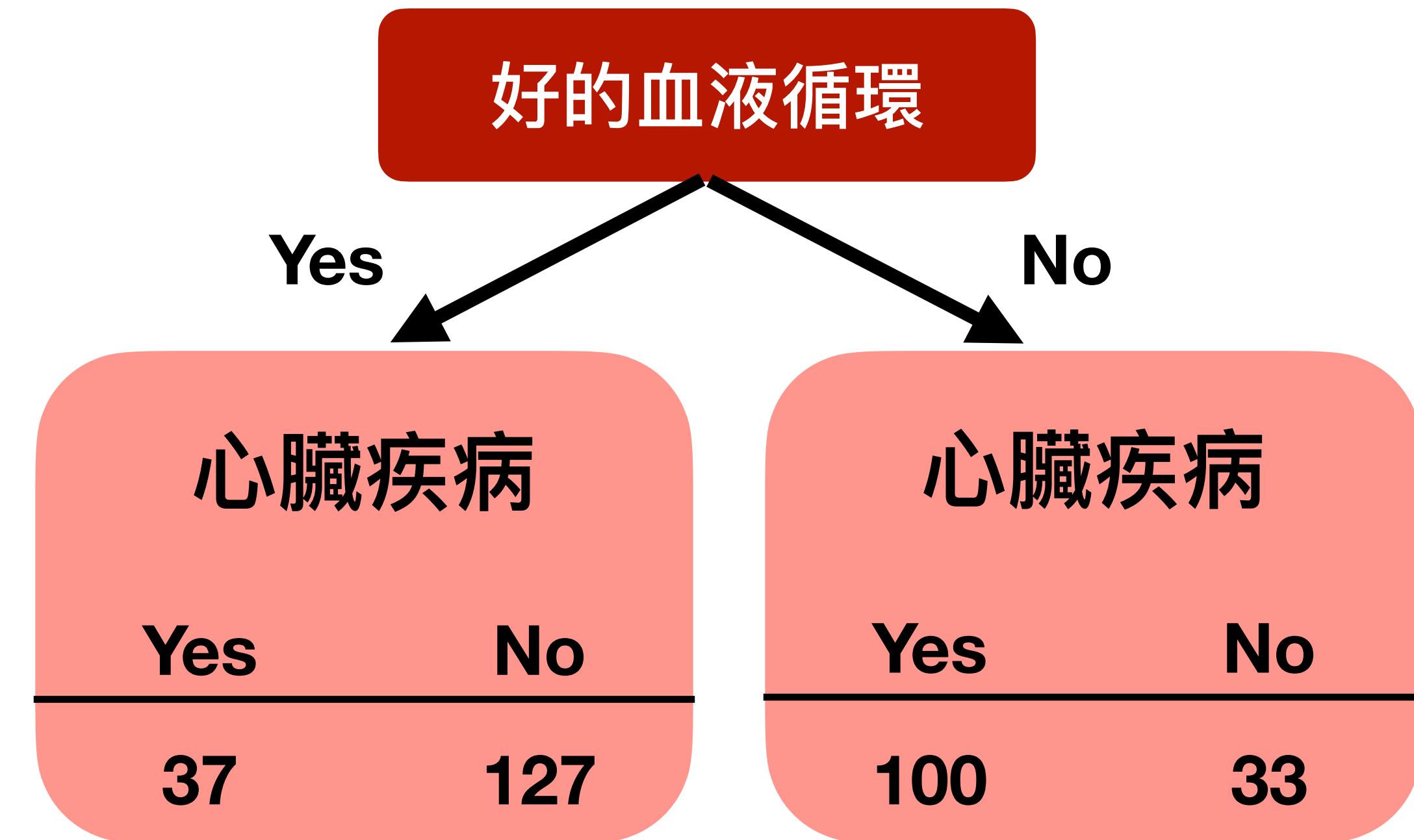
$$\text{Gini Impurity} = 1 - (37/164)^2 - (127/164)^2$$

$$\text{Gini Impurity} = 0.349$$

$$\text{Gini Impurity} = 1 - (100/133)^2 - (33/133)^2$$

$$\text{Gini Impurity} = 0.373$$

好的血液循環 Gini Impurity = weighted average(0.349, 0.373)



$$\text{Gini Impurity} = 1 - (37/164)^2 - (127/164)^2$$

$$\text{Gini Impurity} = 0.349$$

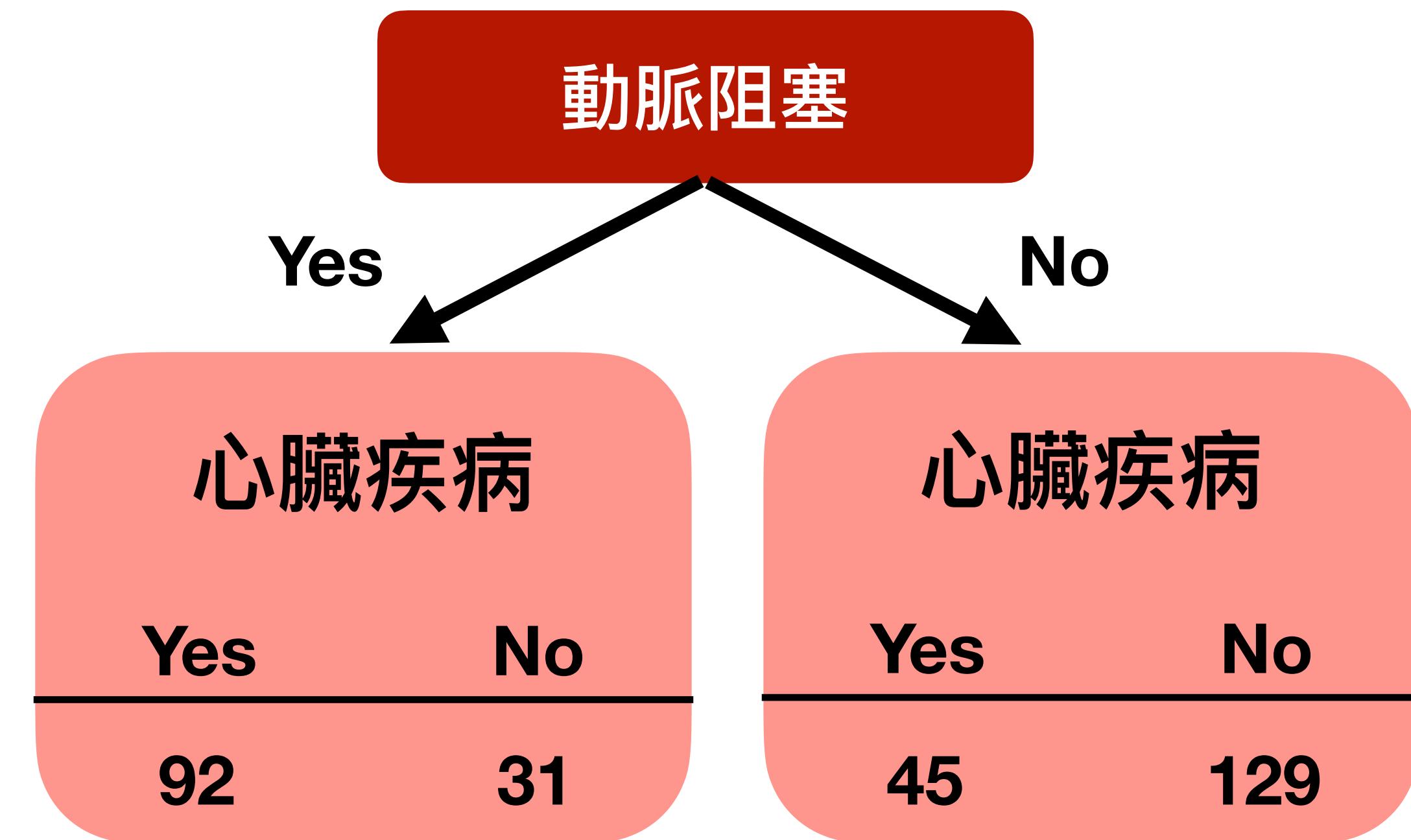
$$\text{Gini Impurity} = 1 - (100/133)^2 - (33/133)^2$$

$$\text{Gini Impurity} = 0.373$$

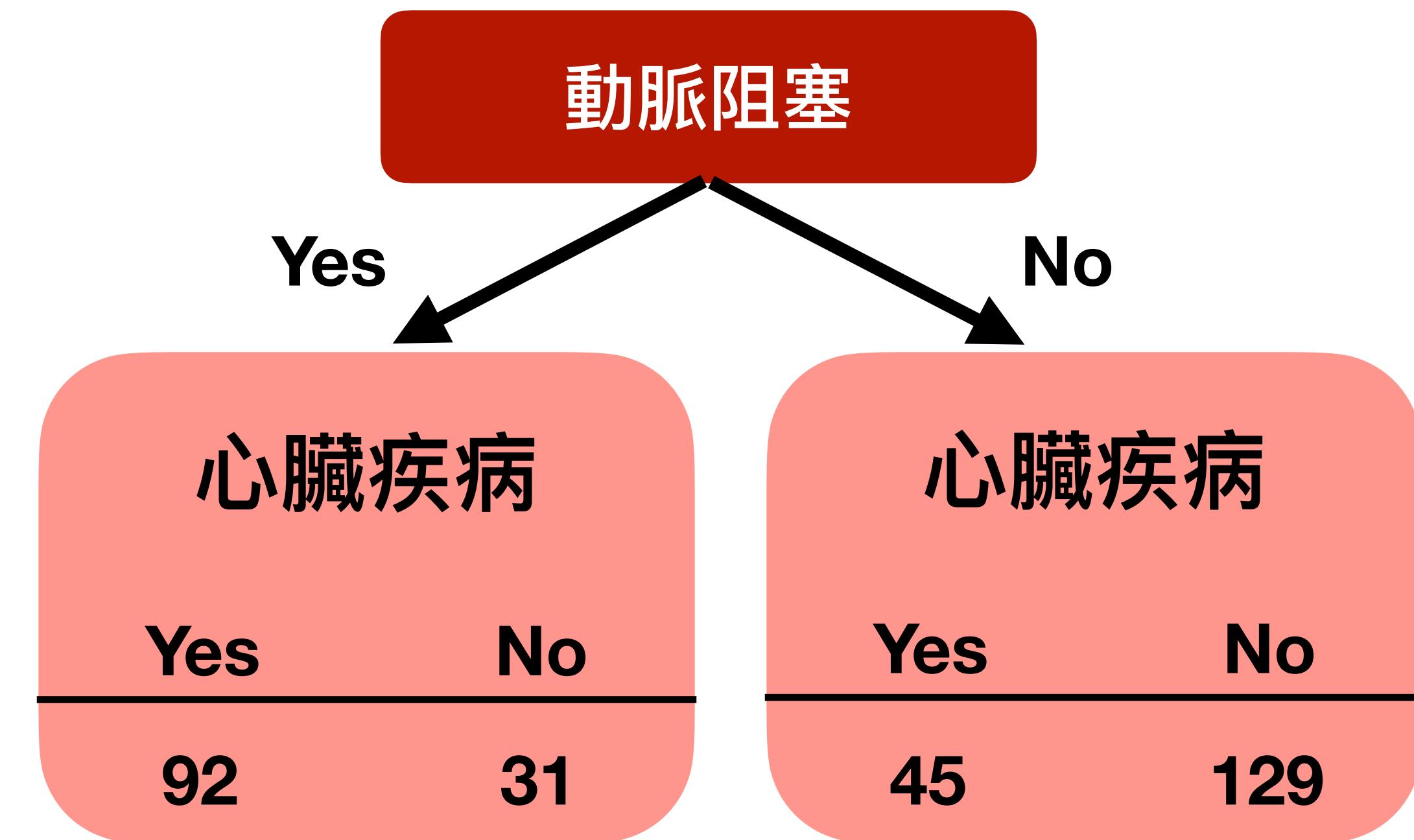
好的血液循環 Gini Impurity = weighted average(0.349, 0.373)

$$\text{好的血液循環 Gini Impurity} = 0.349 * (164/297) + 0.373 * (133/297)$$

$$\text{好的血液循環 Gini Impurity} = 0.363$$

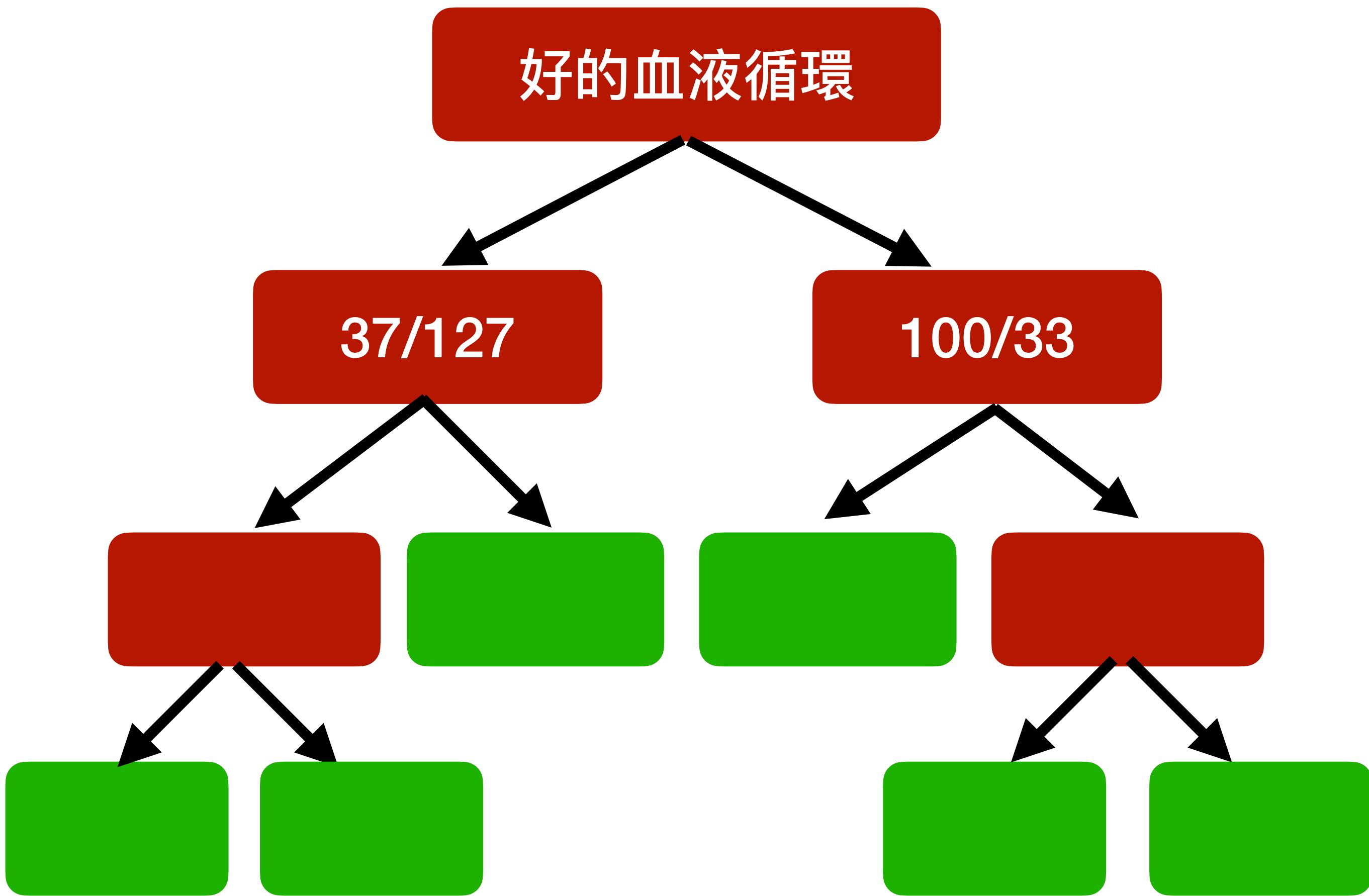


動脈阻塞 Gini Impurity = ?

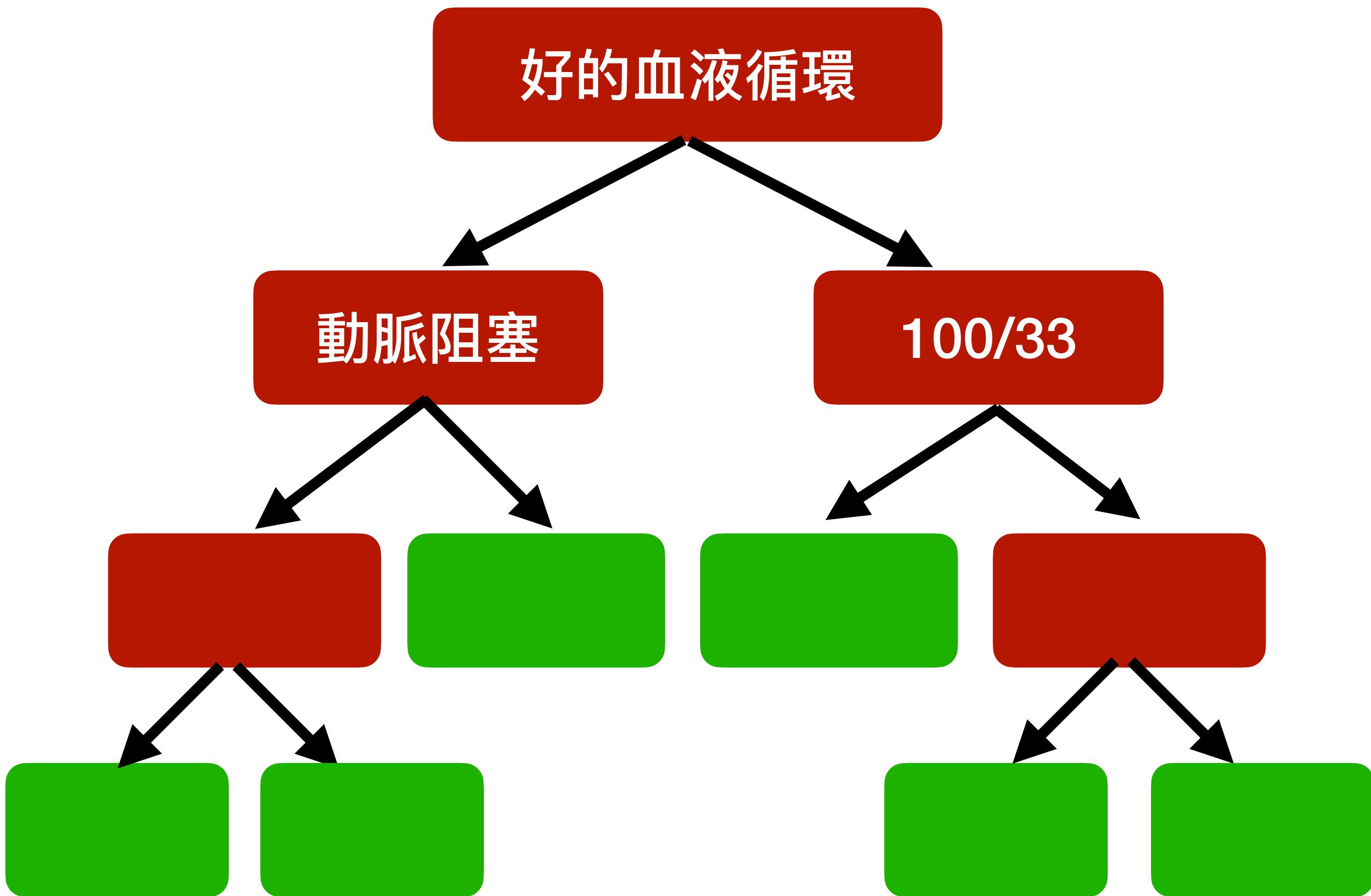


動脈阻塞 Gini Impurity = 0.38

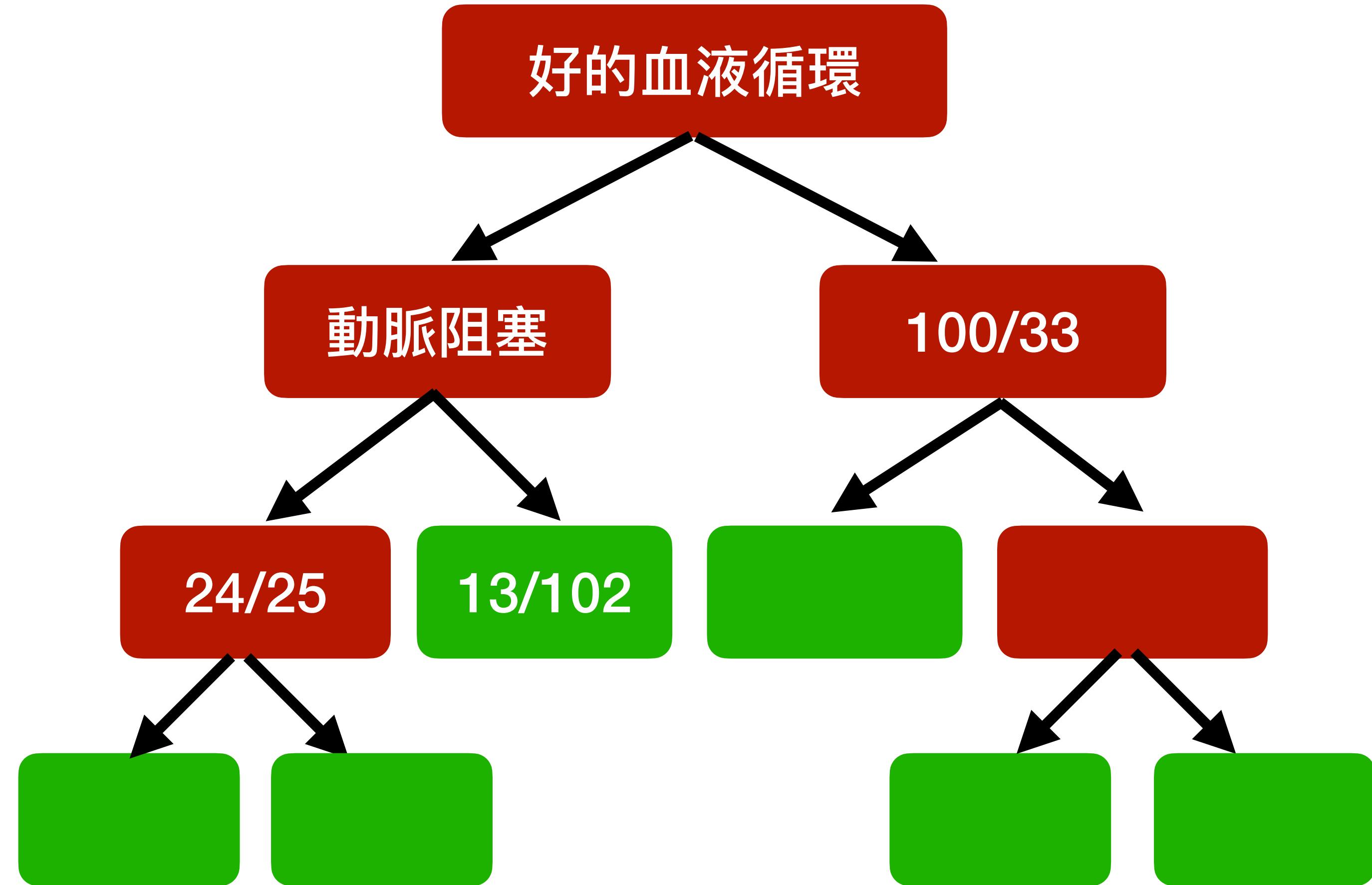
胸悶	好的血液循環	動脈阻塞	心臟疾病
No	No	No	No
Yes	Yes	Yes	Yes
Yes	Yes	No	No
Yes	No	No	Yes
...	...	...	...



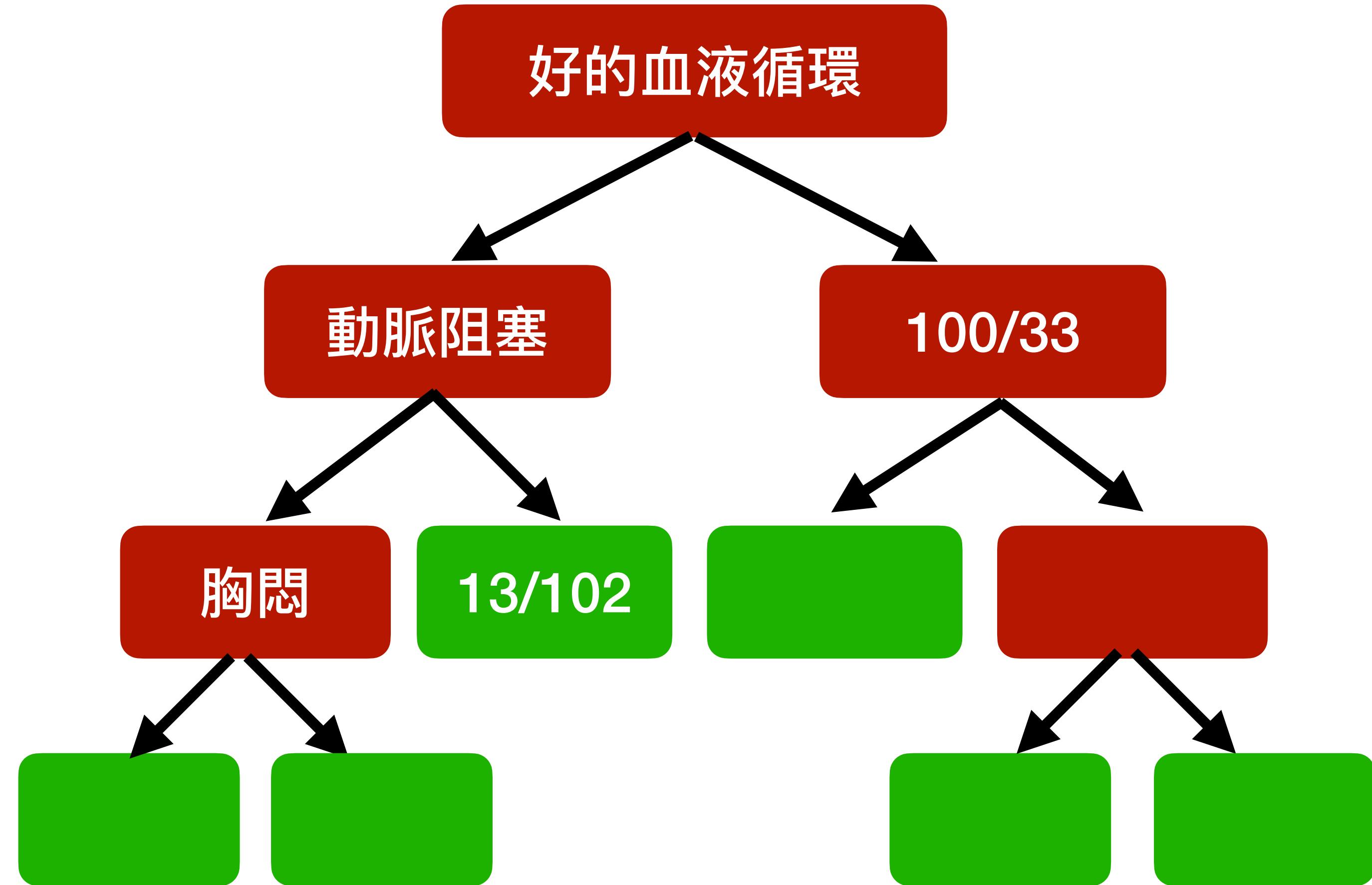
胸悶	好的血液循環	動脈阻塞	心臟疾病
No	No	No	No
Yes	Yes	Yes	Yes
Yes	Yes	No	No
Yes	No	No	Yes
...	...	...	...



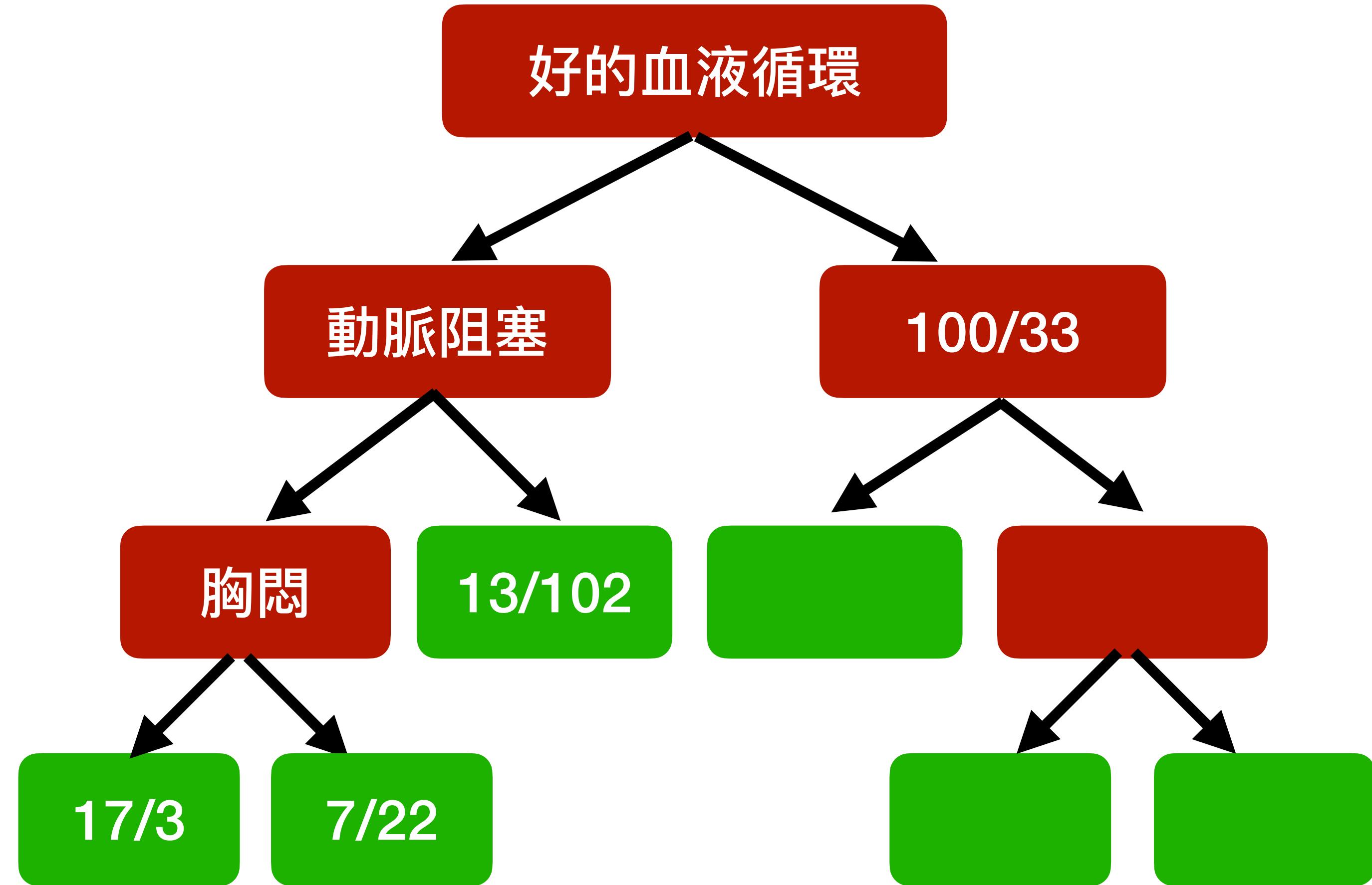
胸悶	好的血液循環	動脈阻塞	心臟疾病
No	No	No	No
Yes	Yes	Yes	Yes
Yes	Yes	No	No
Yes	No	No	Yes
...	...	...	...



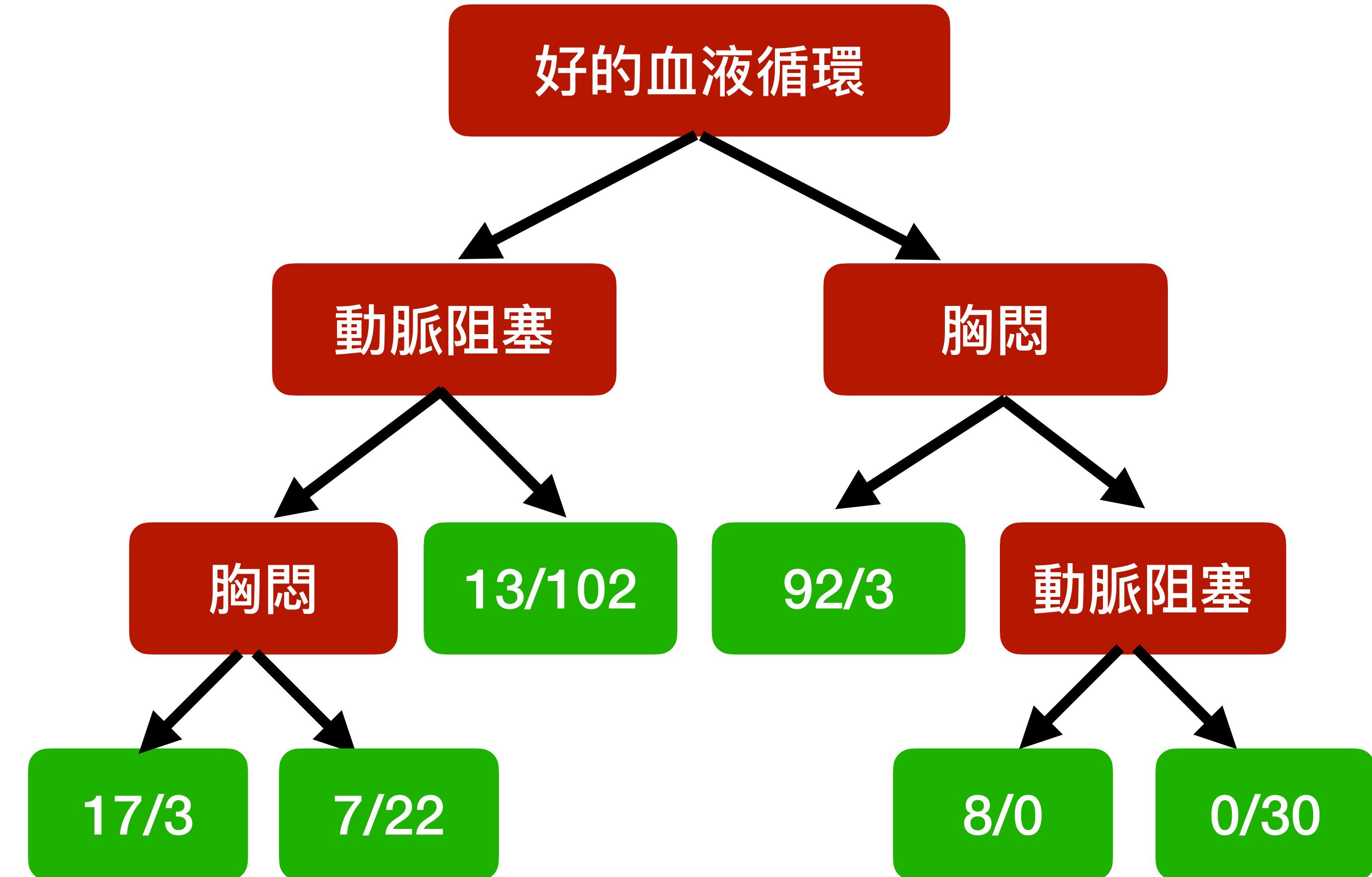
胸悶	好的血液循環	動脈阻塞	心臟疾病
No	No	No	No
Yes	Yes	Yes	Yes
Yes	Yes	No	No
Yes	No	No	Yes
...	...	...	...



胸悶	好的血液循環	動脈阻塞	心臟疾病
No	No	No	No
Yes	Yes	Yes	Yes
Yes	Yes	No	No
Yes	No	No	Yes
...	...	...	...



胸悶	好的血液循環	動脈阻塞	心臟疾病
No	No	No	No
Yes	Yes	Yes	Yes
Yes	Yes	No	No
Yes	No	No	Yes
...	...	...	...



**Let's do this!**

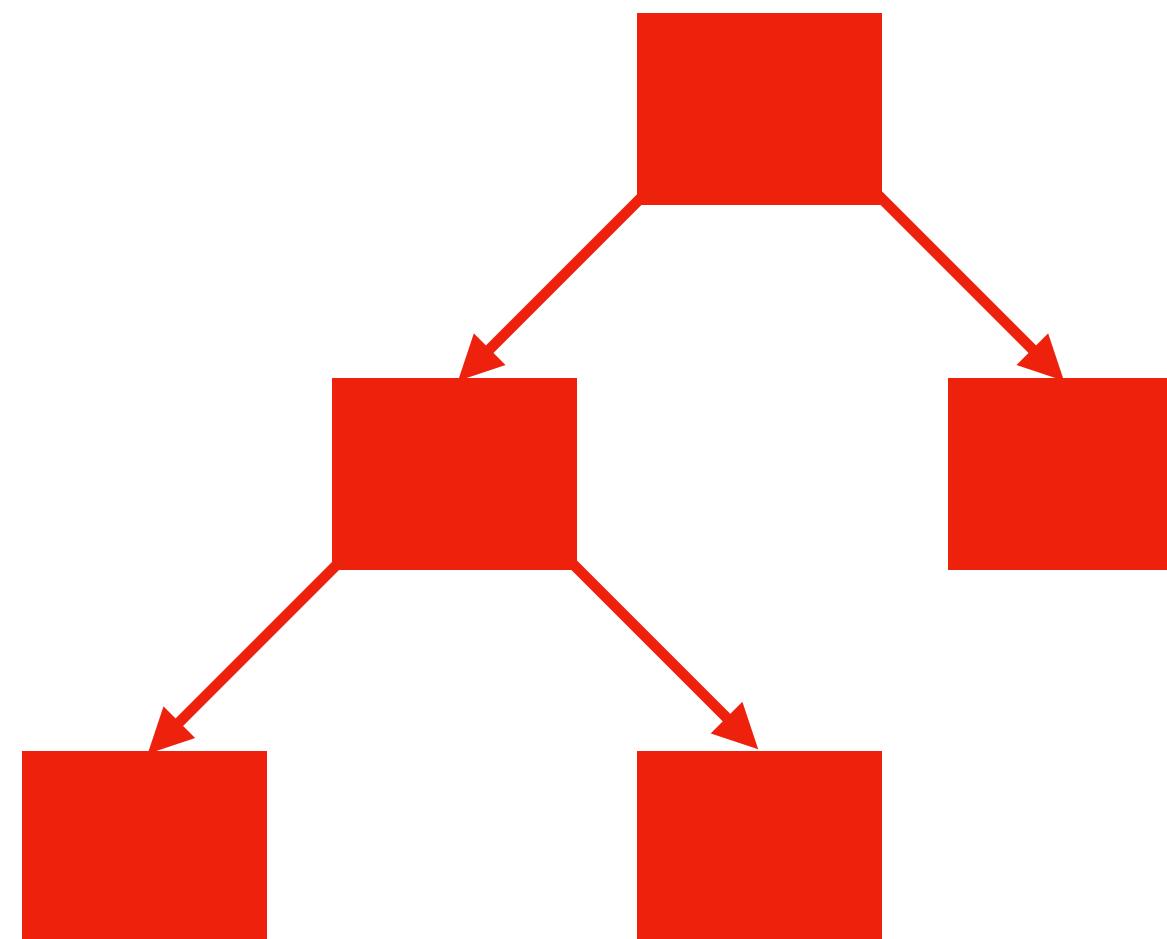


# Random Forest

# data

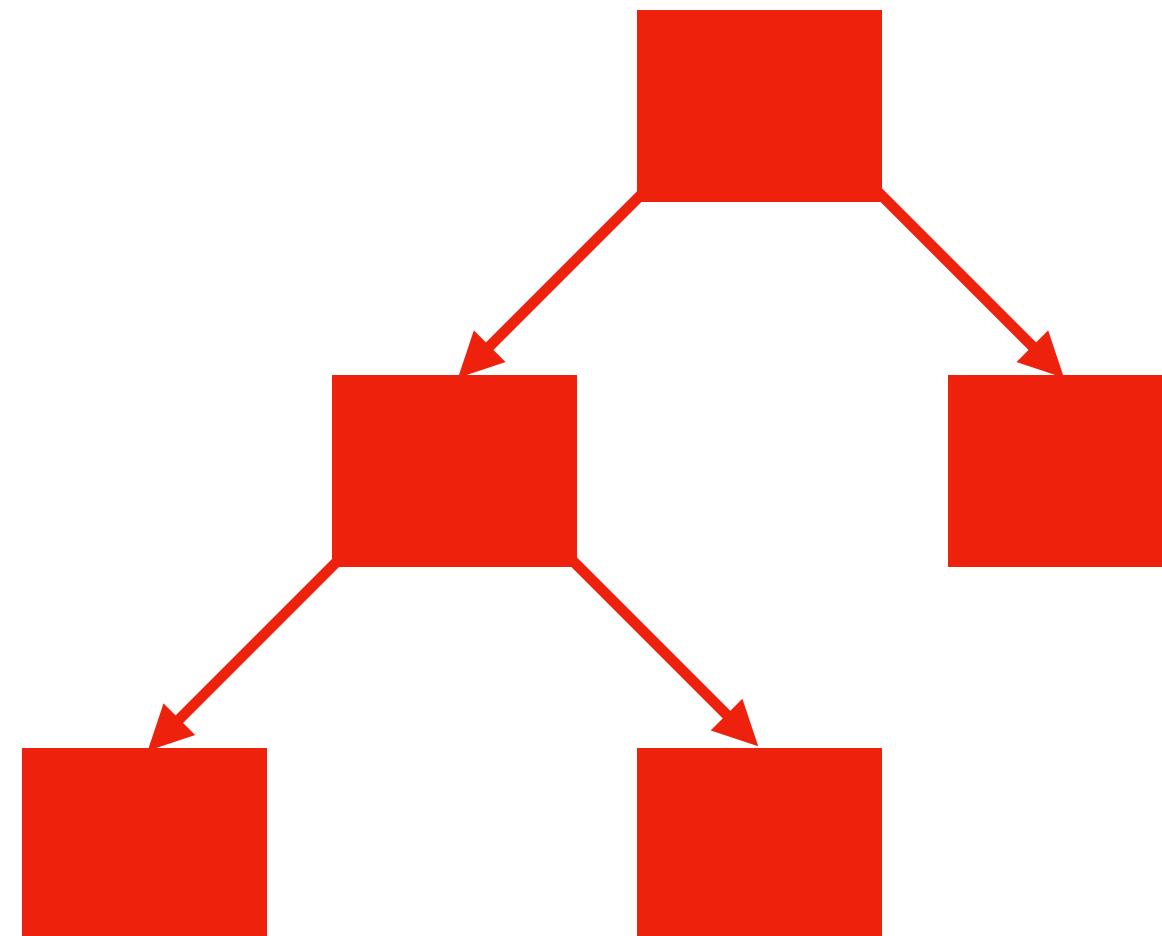
# data

## tree1

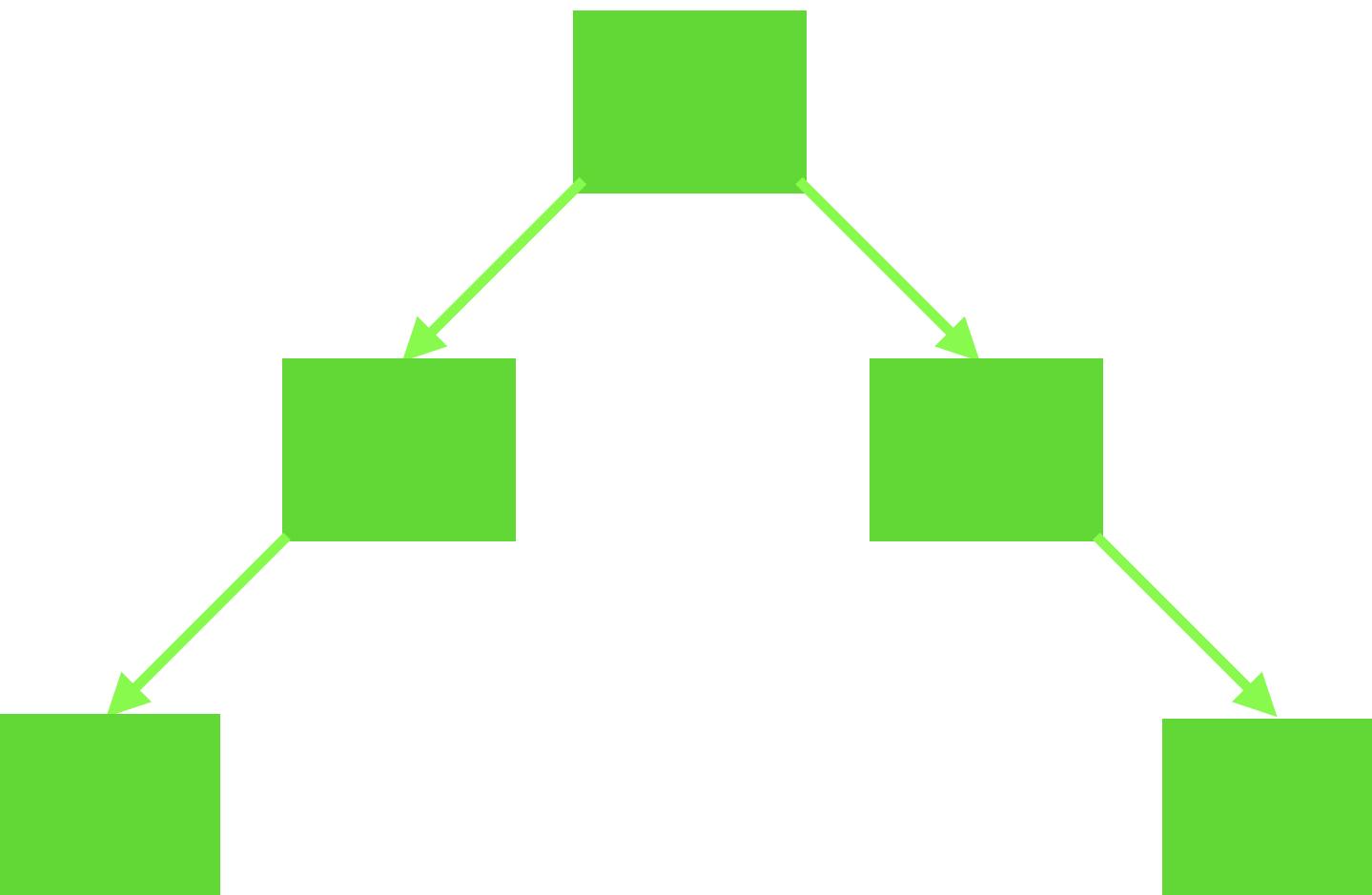


# data

tree1

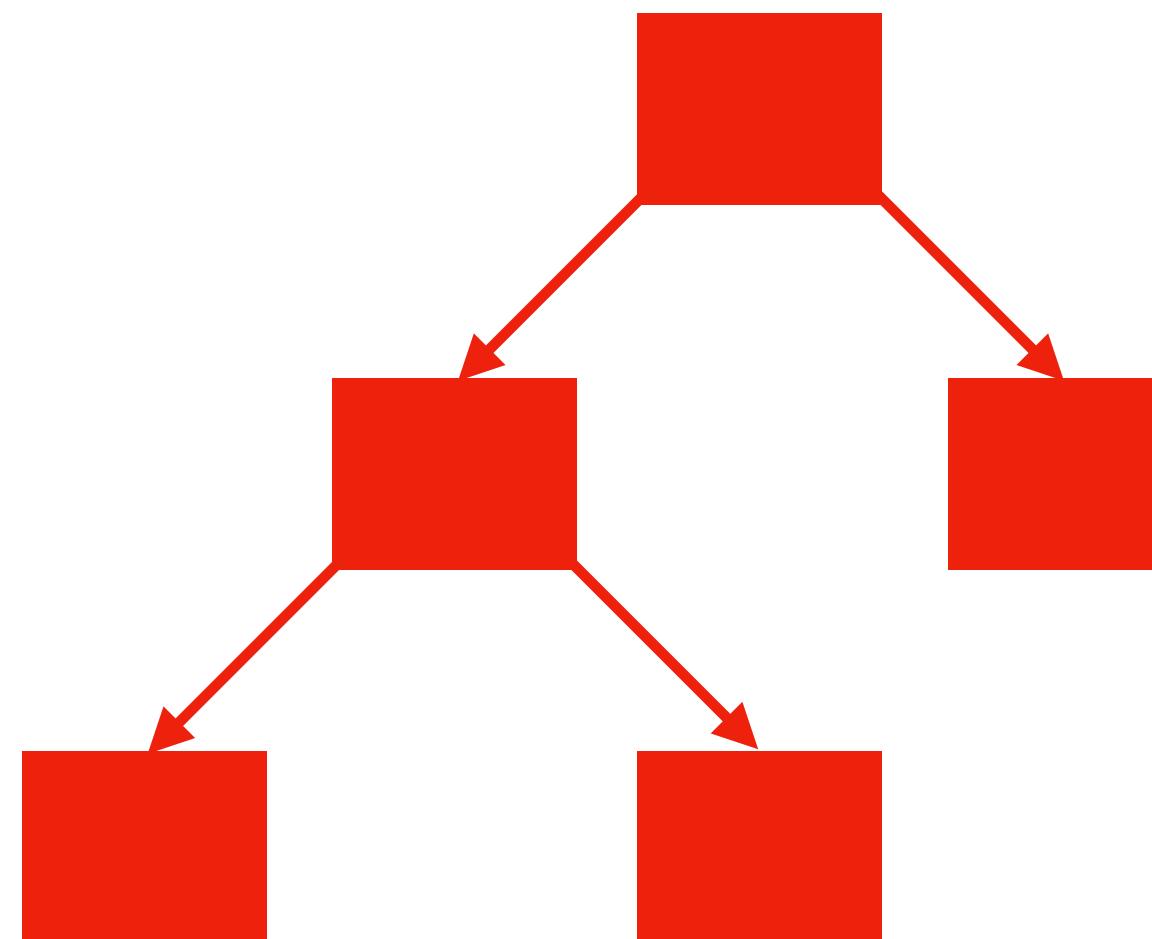


tree2

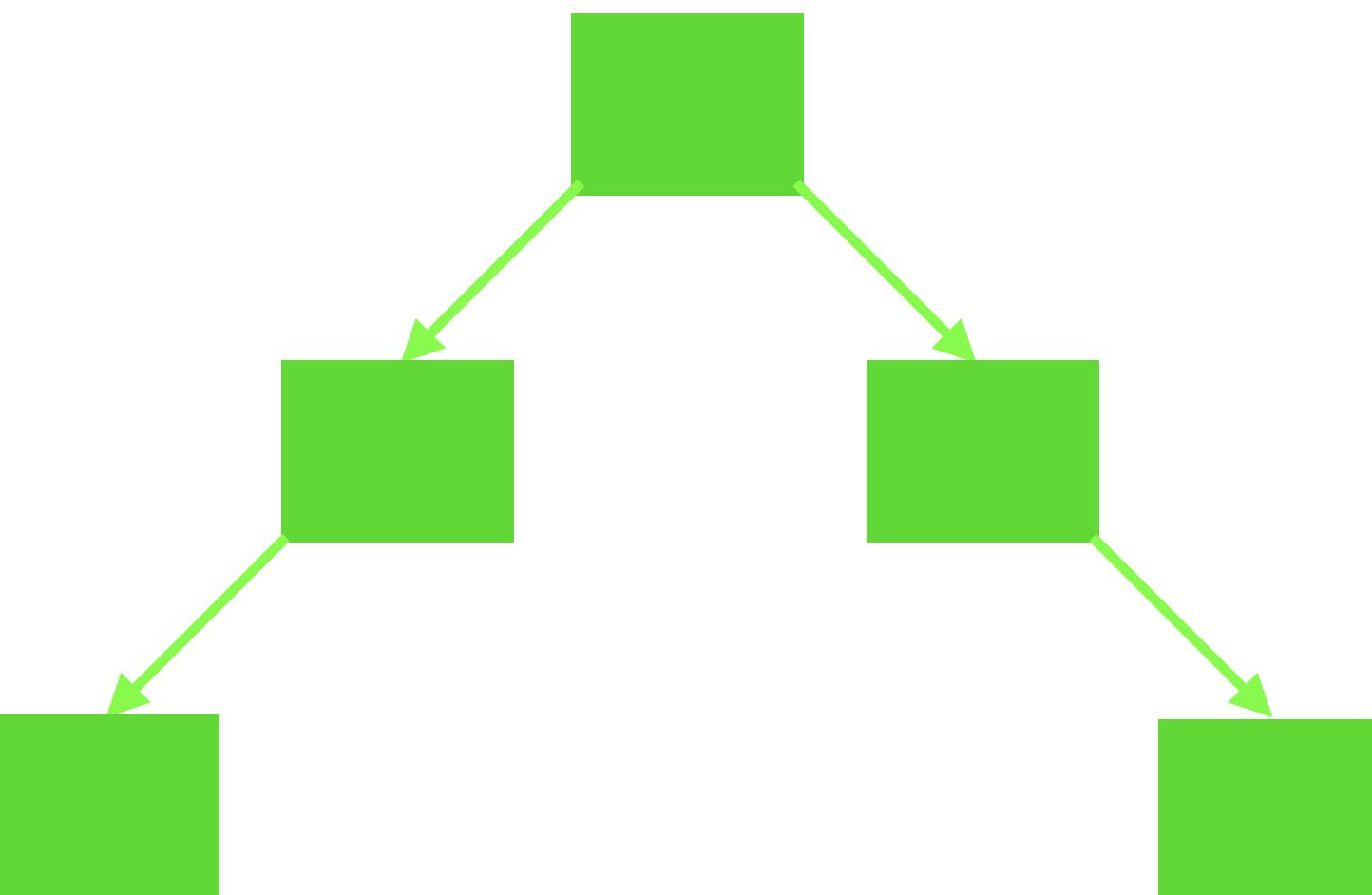


# data

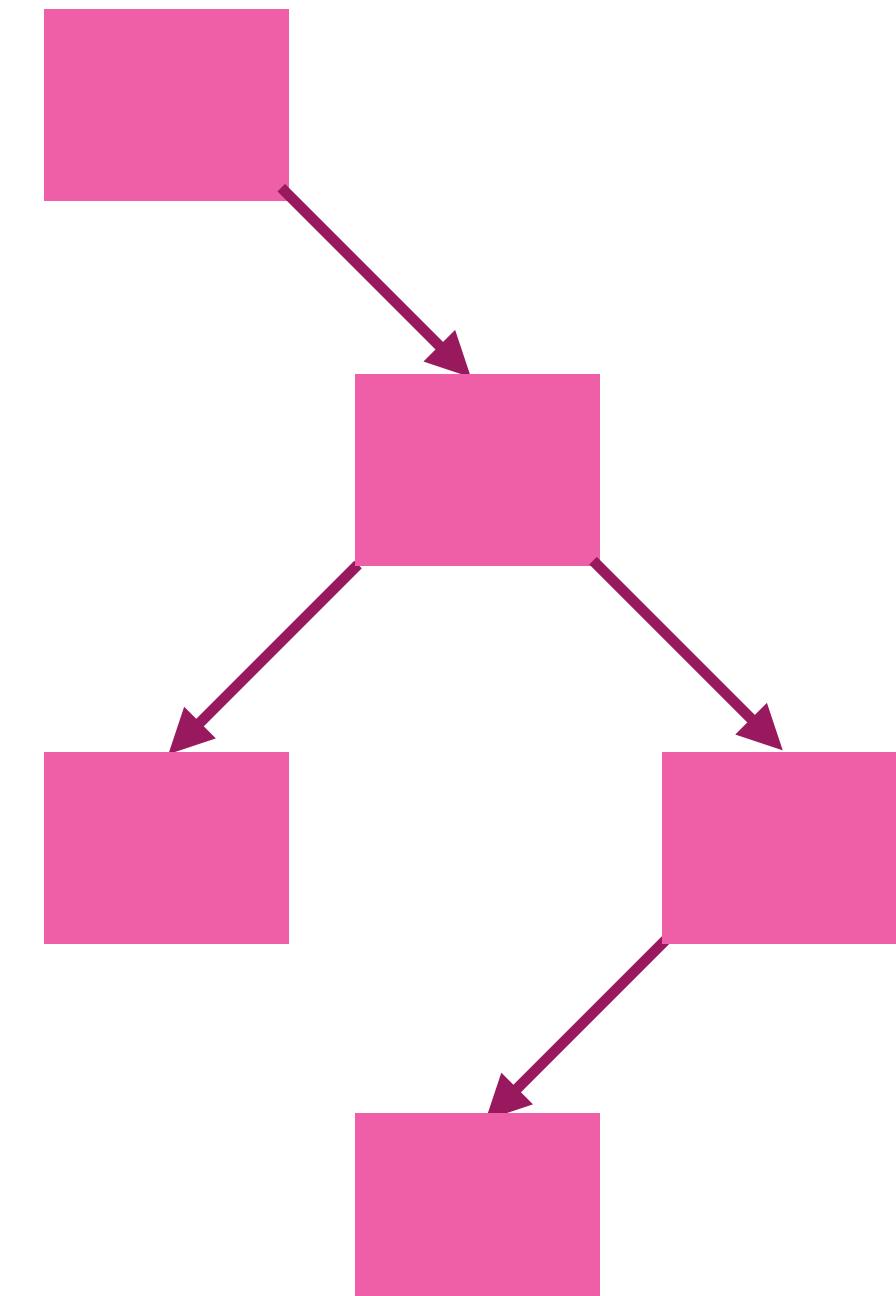
## tree1



## tree2

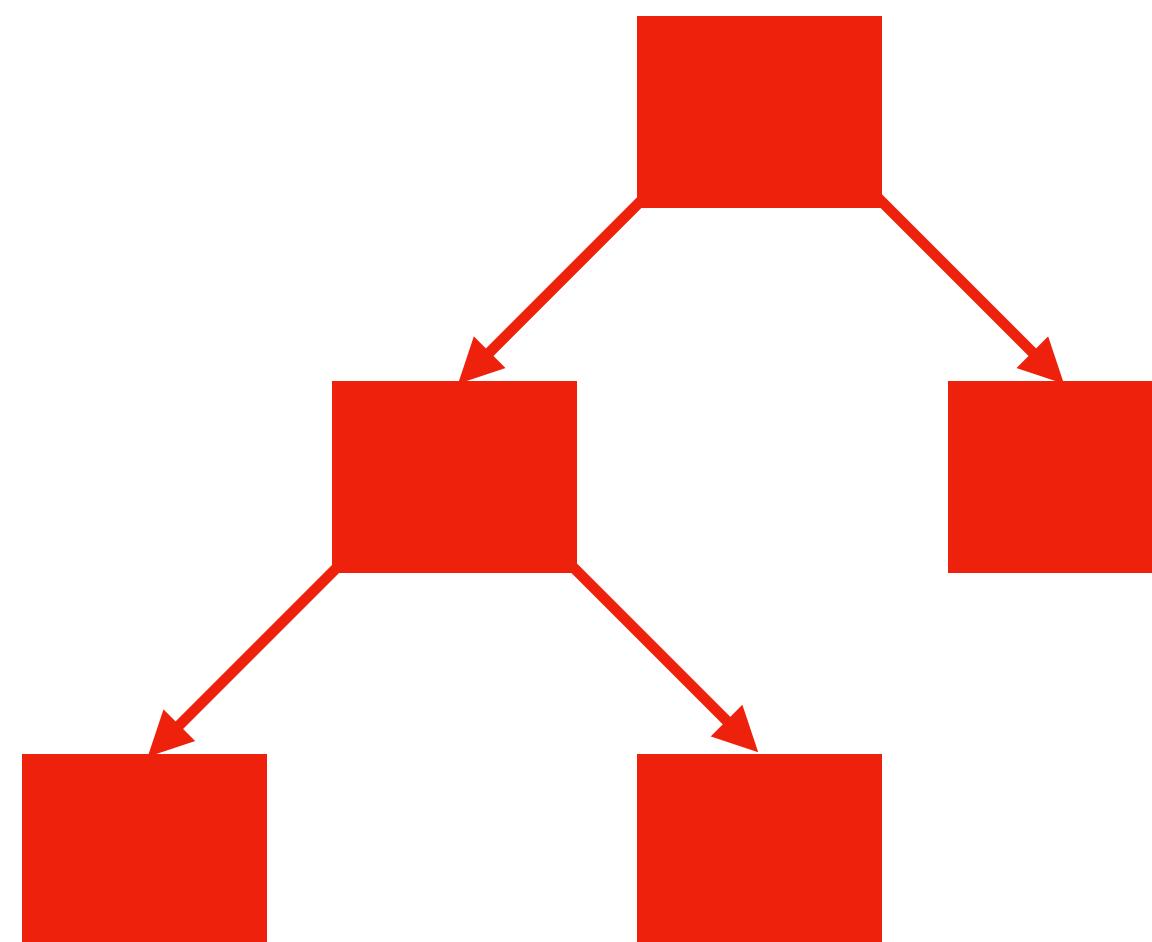


## tree3

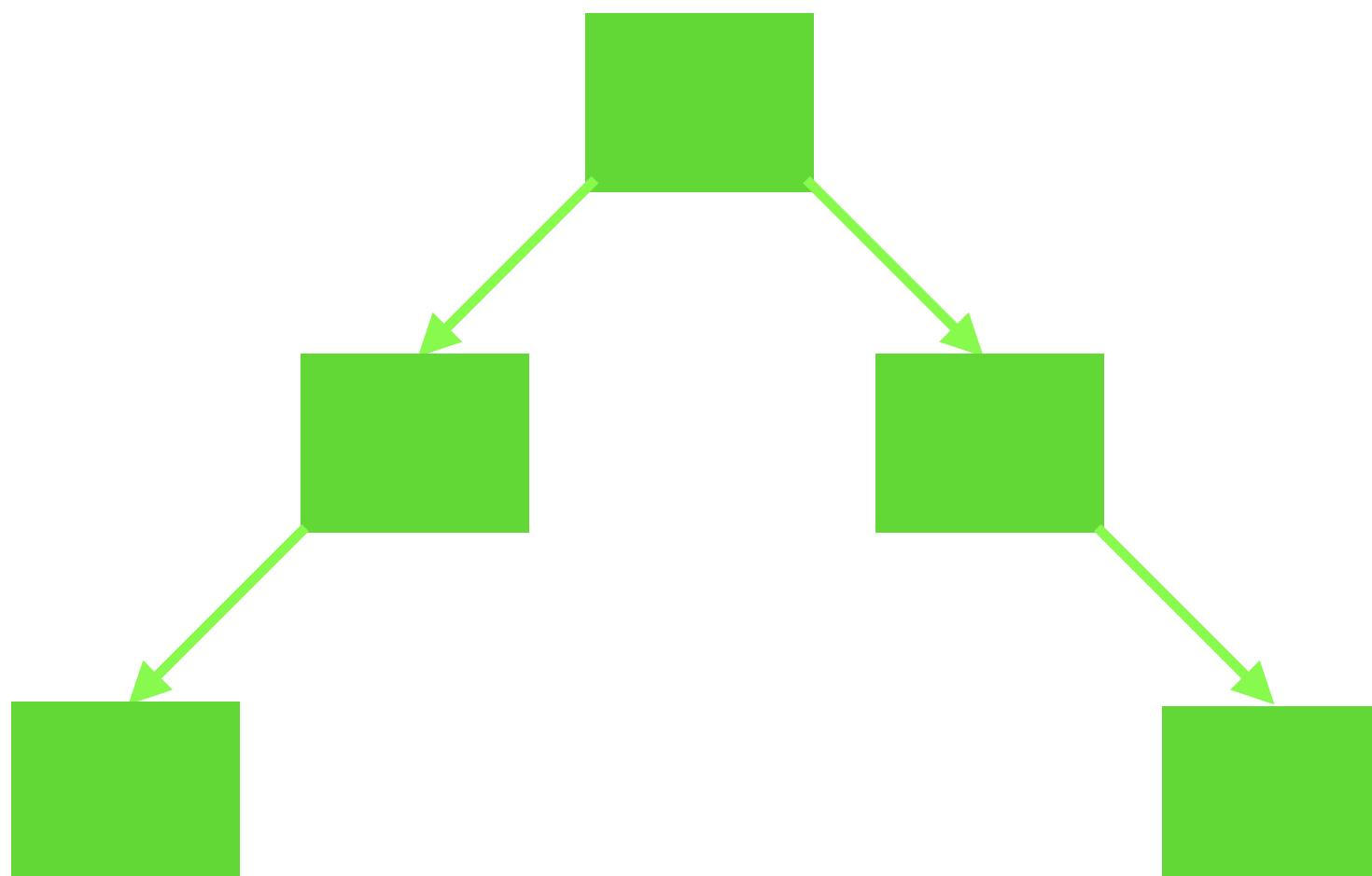


# data

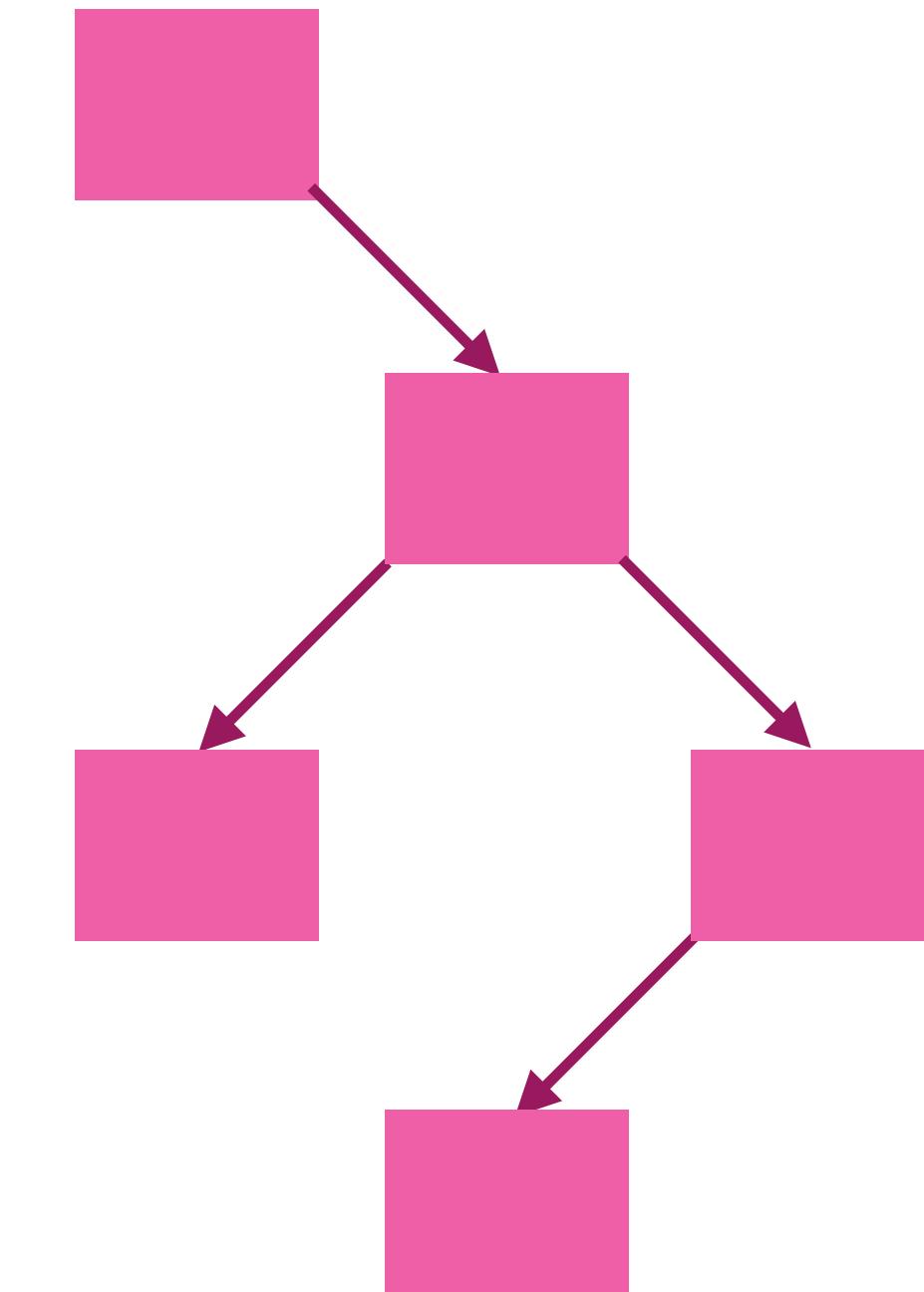
## tree1



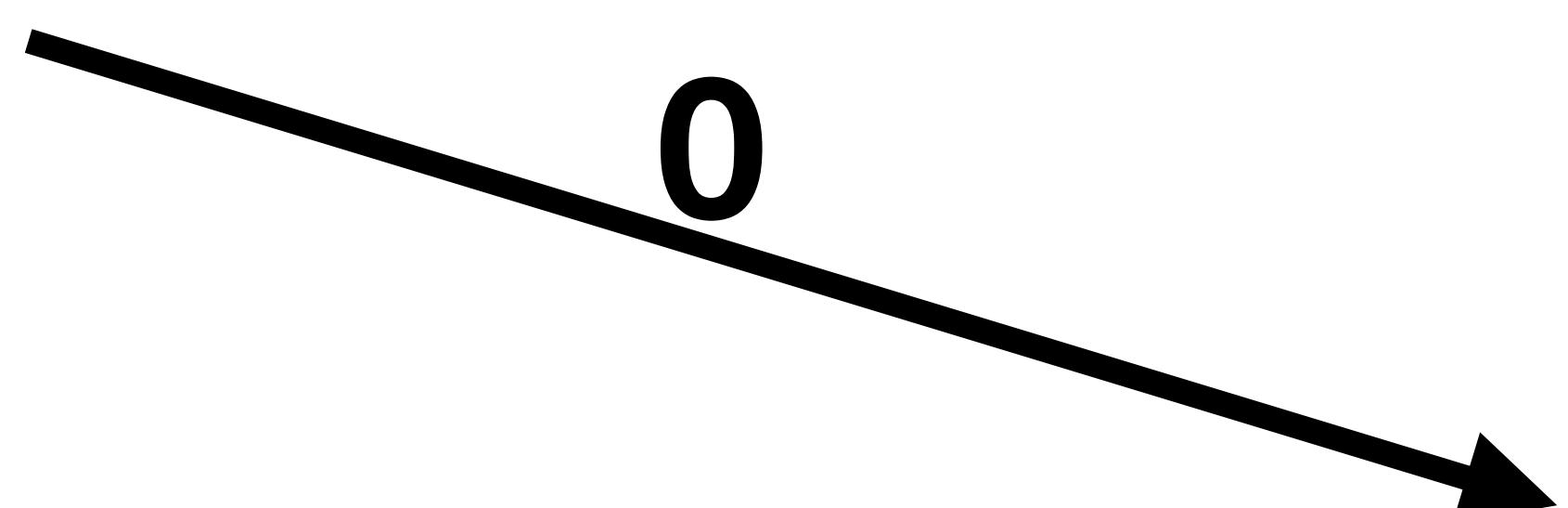
## tree2



## tree3

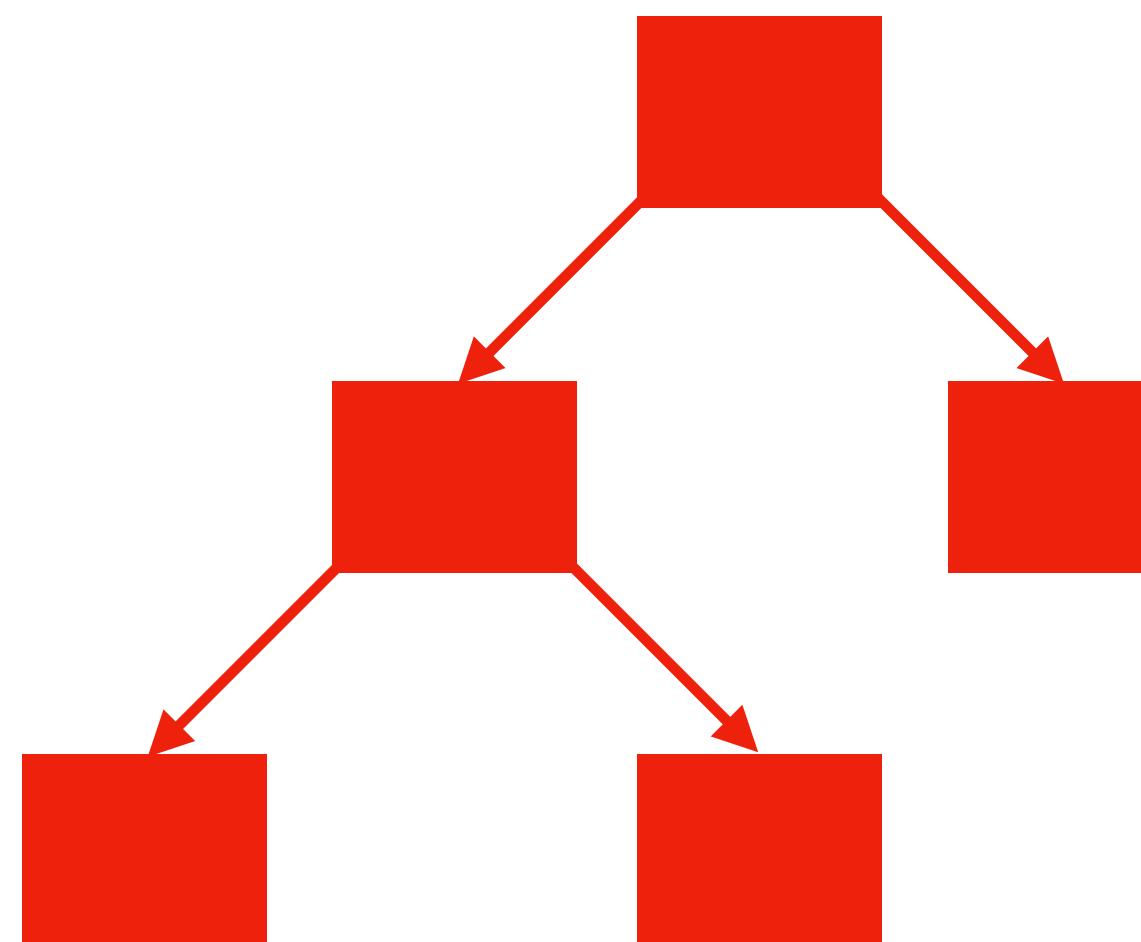


0

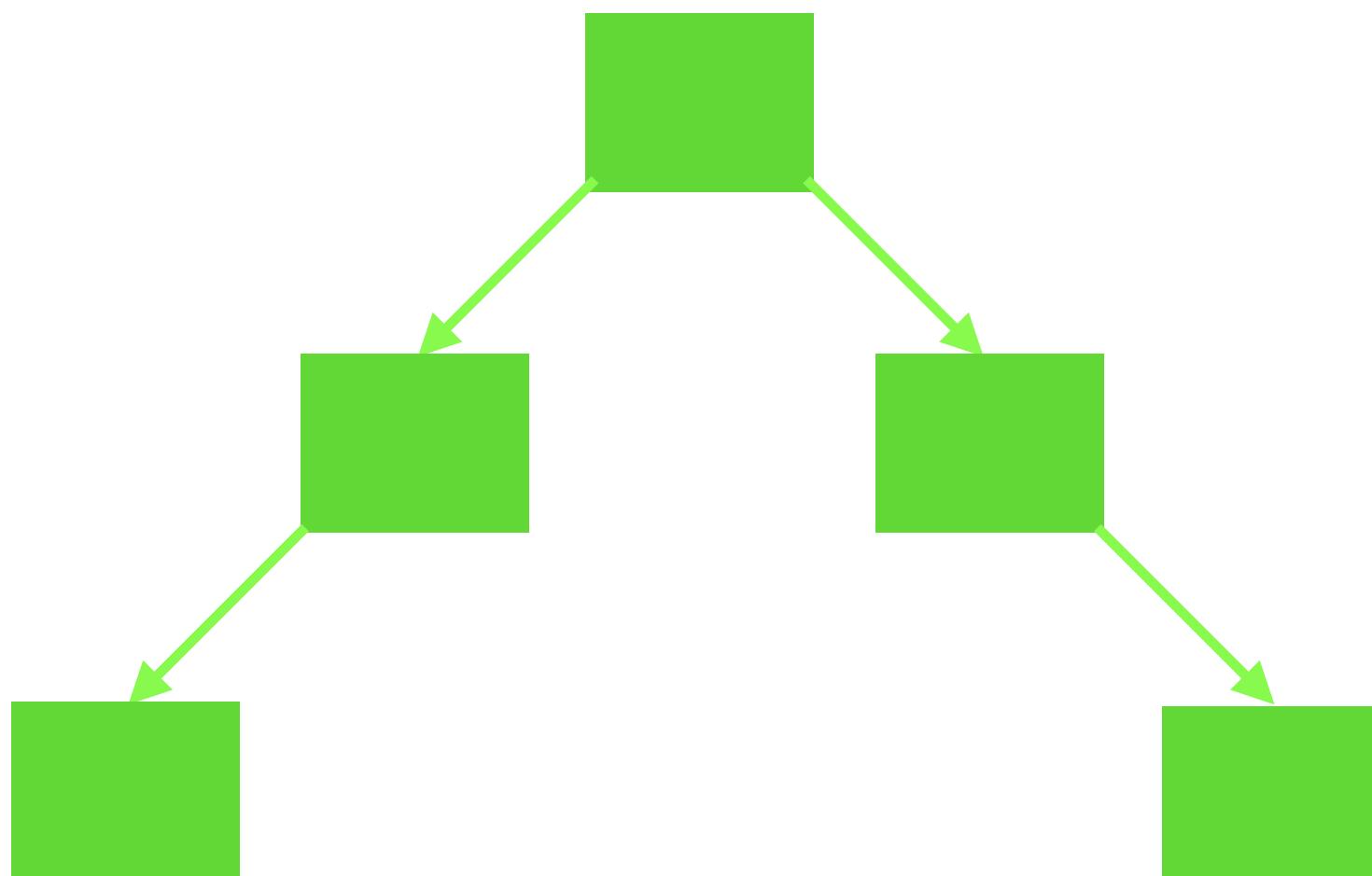


data

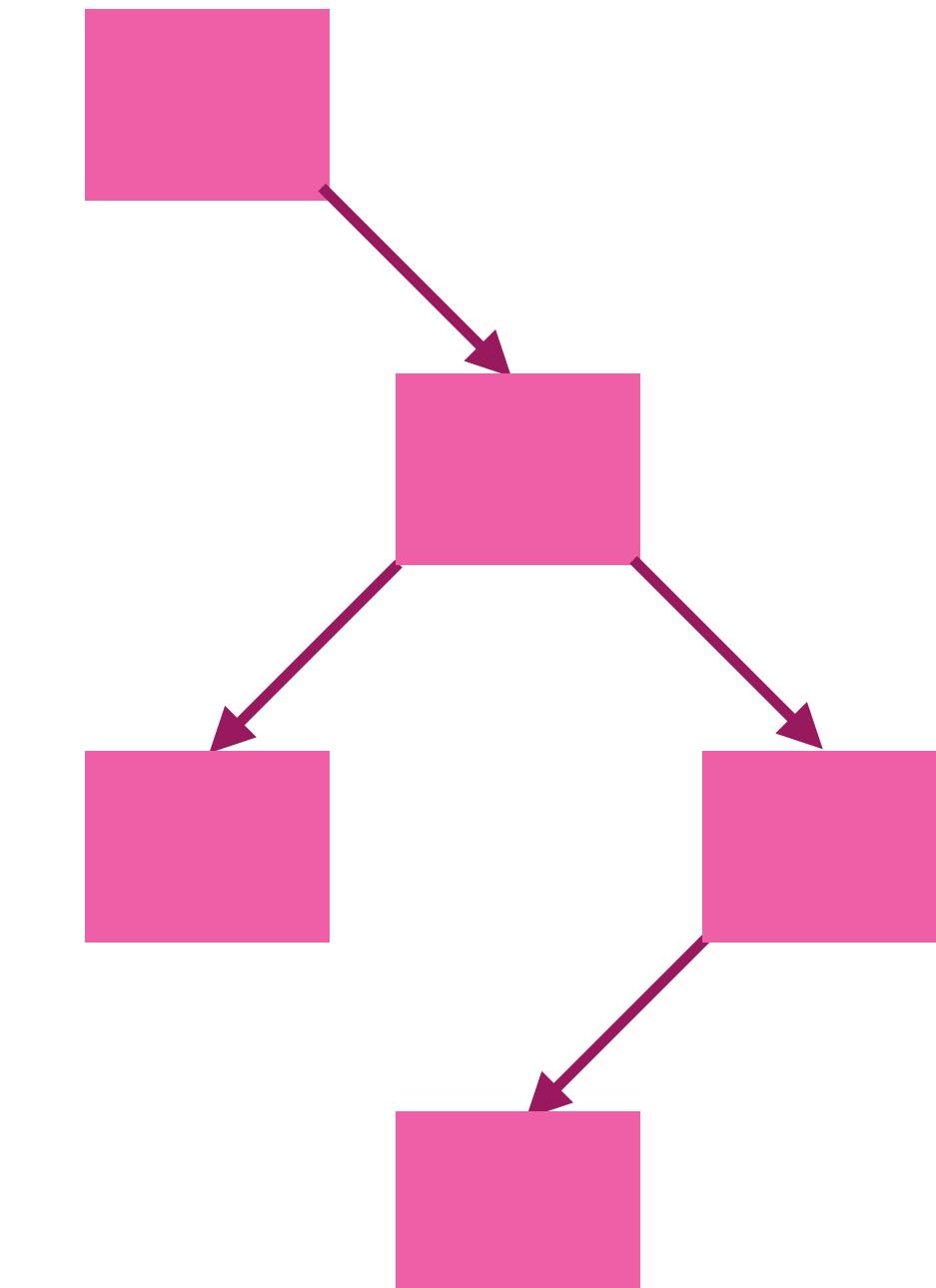
tree1



tree2



tree3

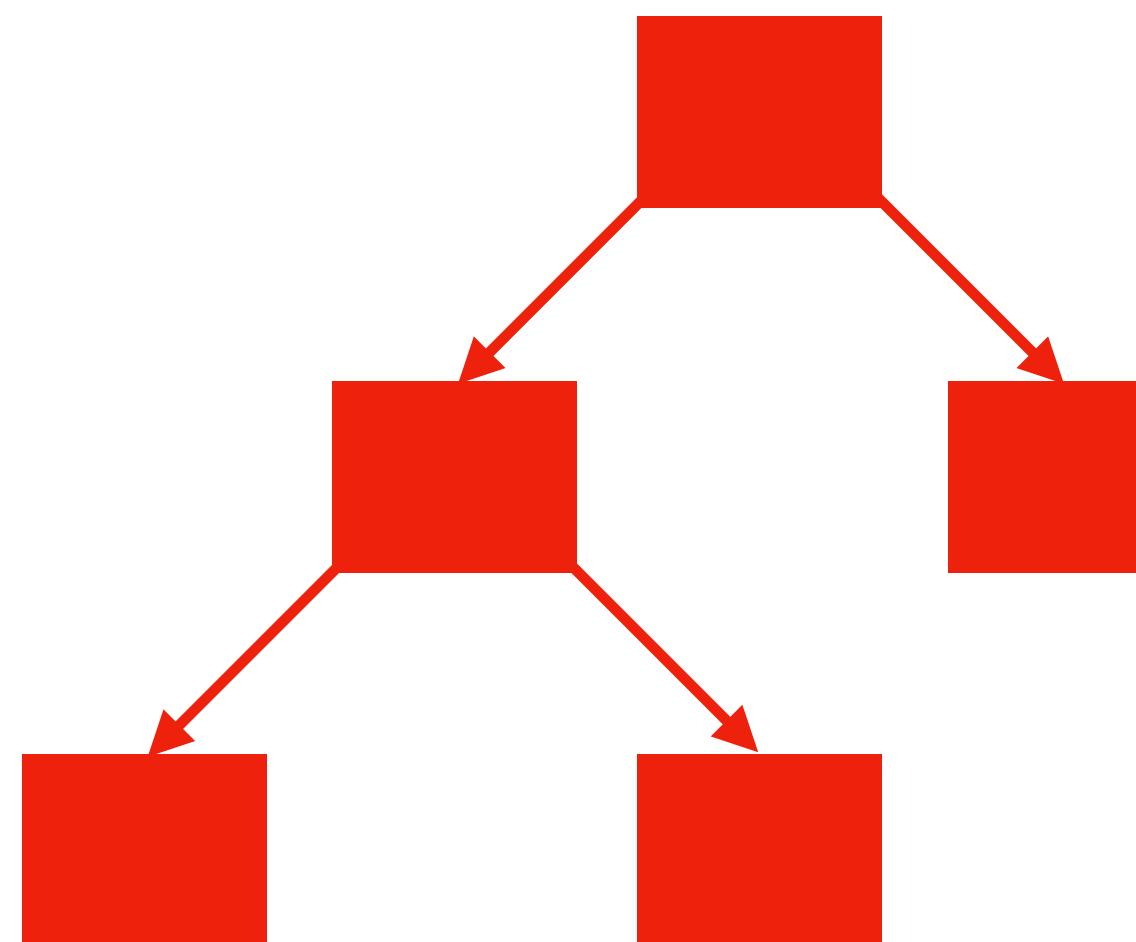


0

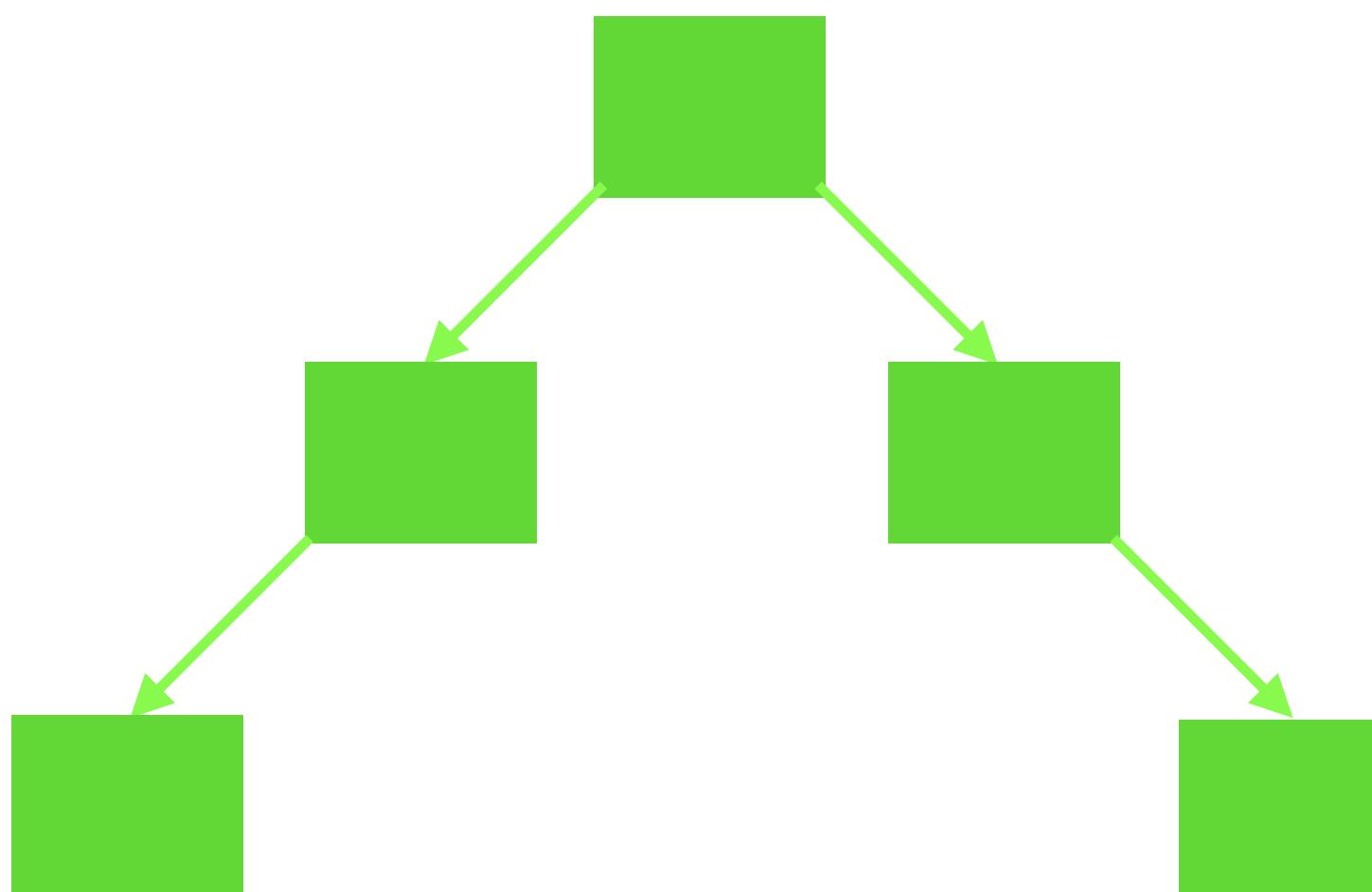
1

data

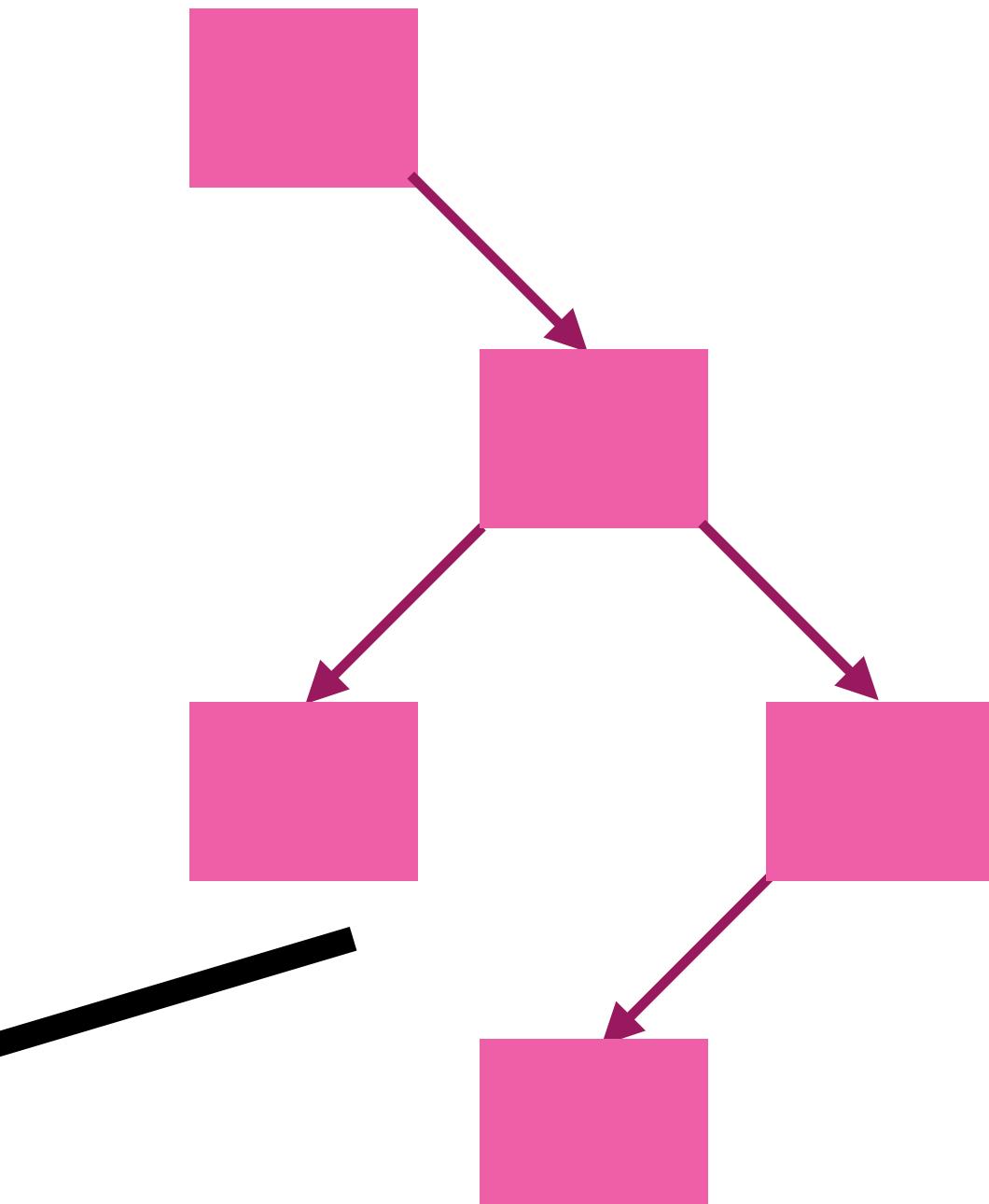
tree1



tree2



tree3



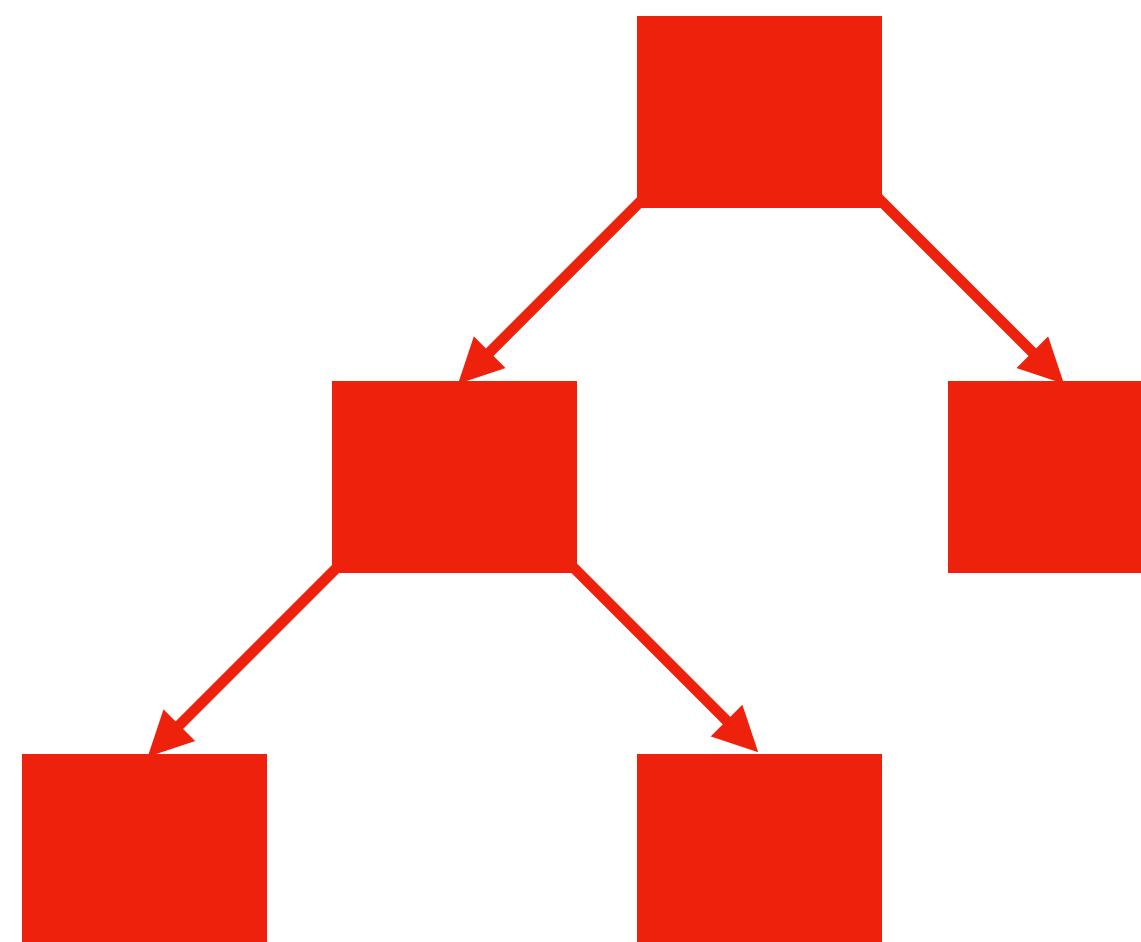
0

1

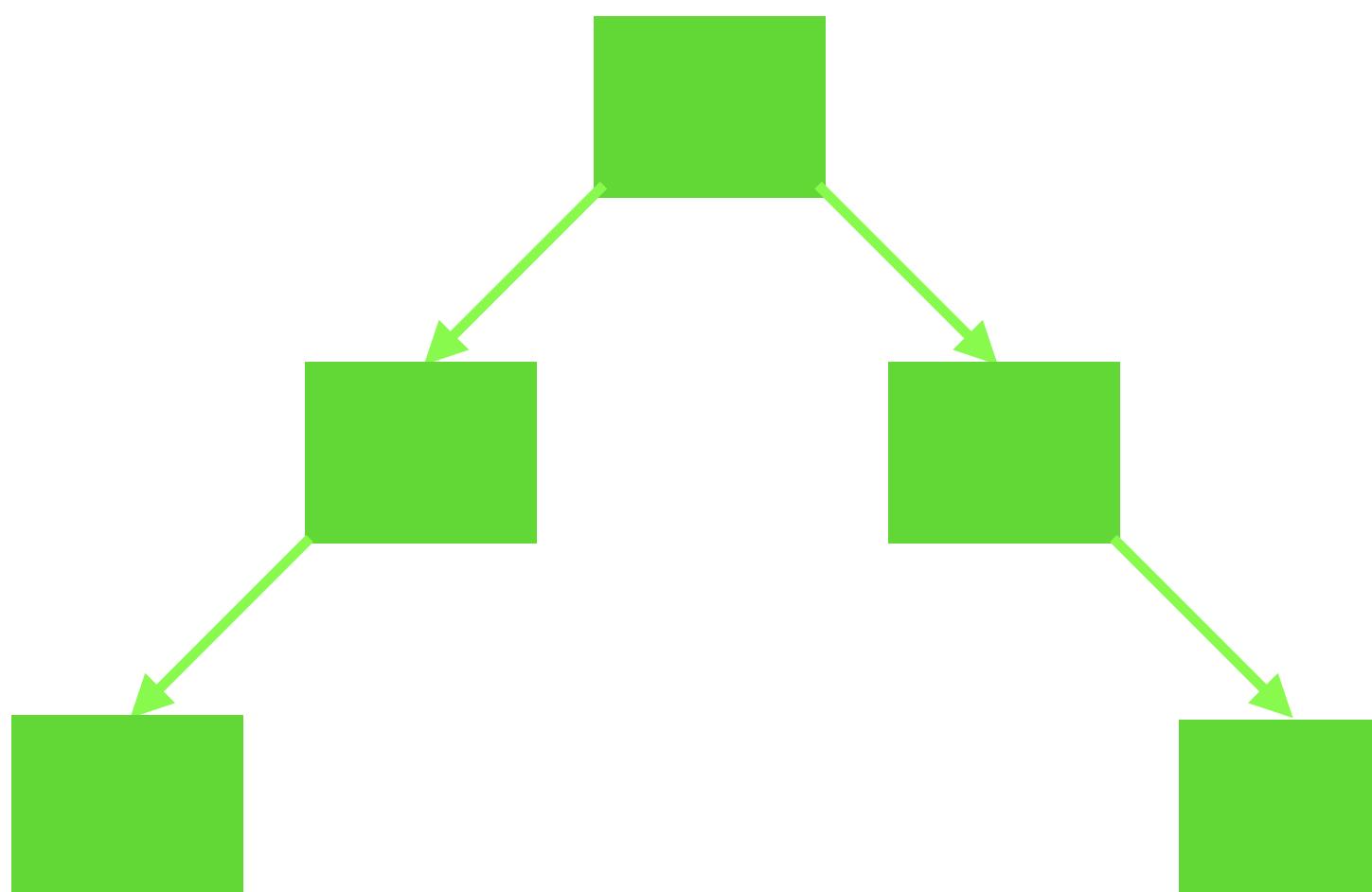
0

data

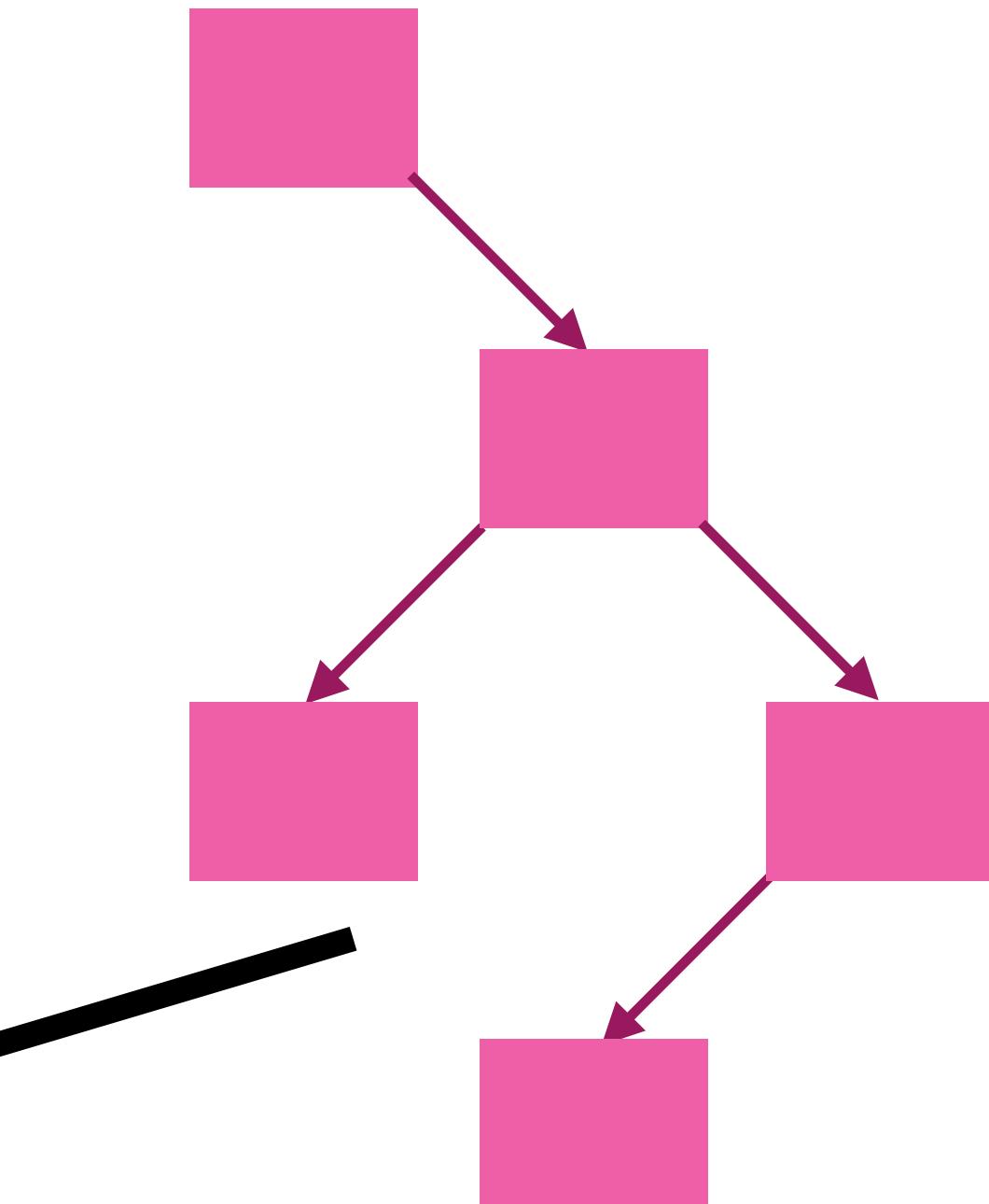
tree1



tree2



tree3



0

1

0

0

**Let's do this!**

# Bootstrapped Dataset

# Original Dataset

胸悶	好的血液循環	動脈阻塞	心臟疾病
No	No	No	No
Yes	Yes	Yes	Yes
Yes	Yes	No	No
Yes	No	No	Yes
...	...	...	...

# Original Dataset

胸悶	好的血液循環	動脈阻塞	心臟疾病
No	No	No	No
Yes	Yes	Yes	Yes
Yes	Yes	No	No
Yes	No	No	Yes
...	...	...	...

# Bootstrapped Dataset

胸悶	好的血液循環	動脈阻塞	心臟疾病
No	No	No	No

# Original Dataset

胸悶	好的血液循環	動脈阻塞	心臟疾病
No	No	No	No
Yes	Yes	Yes	Yes
Yes	Yes	No	No
Yes	No	No	Yes
...	...	...	...

# Bootstrapped Dataset

胸悶	好的血液循環	動脈阻塞	心臟疾病
No	No	No	No

# Original Dataset

胸悶	好的血液循環	動脈阻塞	心臟疾病
No	No	No	No
Yes	Yes	Yes	Yes
Yes	Yes	No	No
Yes	No	No	Yes
...	...	...	...

# Bootstrapped Dataset

胸悶	好的血液循環	動脈阻塞	心臟疾病
Yes	Yes	No	No

# Original Dataset

胸悶	好的血液循環	動脈阻塞	心臟疾病
No	No	No	No
Yes	Yes	Yes	Yes
Yes	Yes	No	No
Yes	No	No	Yes
...	...	...	...

# Bootstrapped Dataset

胸悶	好的血液循環	動脈阻塞	心臟疾病
Yes	Yes	No	No

# Original Dataset

胸悶	好的血液循環	動脈阻塞	心臟疾病
No	No	No	No
Yes	Yes	Yes	Yes
Yes	Yes	No	No
Yes	No	No	Yes
...	...	...	...

# Bootstrapped Dataset

胸悶	好的血液循環	動脈阻塞	心臟疾病
Yes	Yes	No	No
No	No	No	No

# Original Dataset

胸悶	好的血液循環	動脈阻塞	心臟疾病
No	No	No	No
Yes	Yes	Yes	Yes
Yes	Yes	No	No
Yes	No	No	Yes
...	...	...	...

# Bootstrapped Dataset

胸悶	好的血液循環	動脈阻塞	心臟疾病
Yes	Yes	No	No
No	No	No	No
...	...	...	...

# Original Dataset

胸悶	好的血液循環	動脈阻塞	心臟疾病
No	No	No	No
Yes	Yes	Yes	Yes
Yes	Yes	No	No
Yes	No	No	Yes
...	...	...	...

# Bootstrapped Dataset

胸悶	好的血液循環	動脈阻塞	心臟疾病
Yes	Yes	No	No
No	No	No	No
Yes	Yes	No	No
...	...	...	...

# Original Dataset

胸悶	好的血液循環	動脈阻塞	心臟疾病
No	No	No	No
Yes	Yes	Yes	Yes
Yes	Yes	No	No
Yes	No	No	Yes
...	...	...	...

# Bootstrapped Dataset

胸悶	好的血液循環	動脈阻塞	心臟疾病
Yes	Yes	No	No
No	No	No	No
Yes	Yes	No	No
...	...	...	...

# Original Dataset

胸悶	好的血液循環	動脈阻塞	心臟疾病
No	No	No	No
Yes	Yes	Yes	Yes
Yes	Yes	No	No
Yes	No	No	Yes
...	...	...	...

# Bootstrapped Dataset

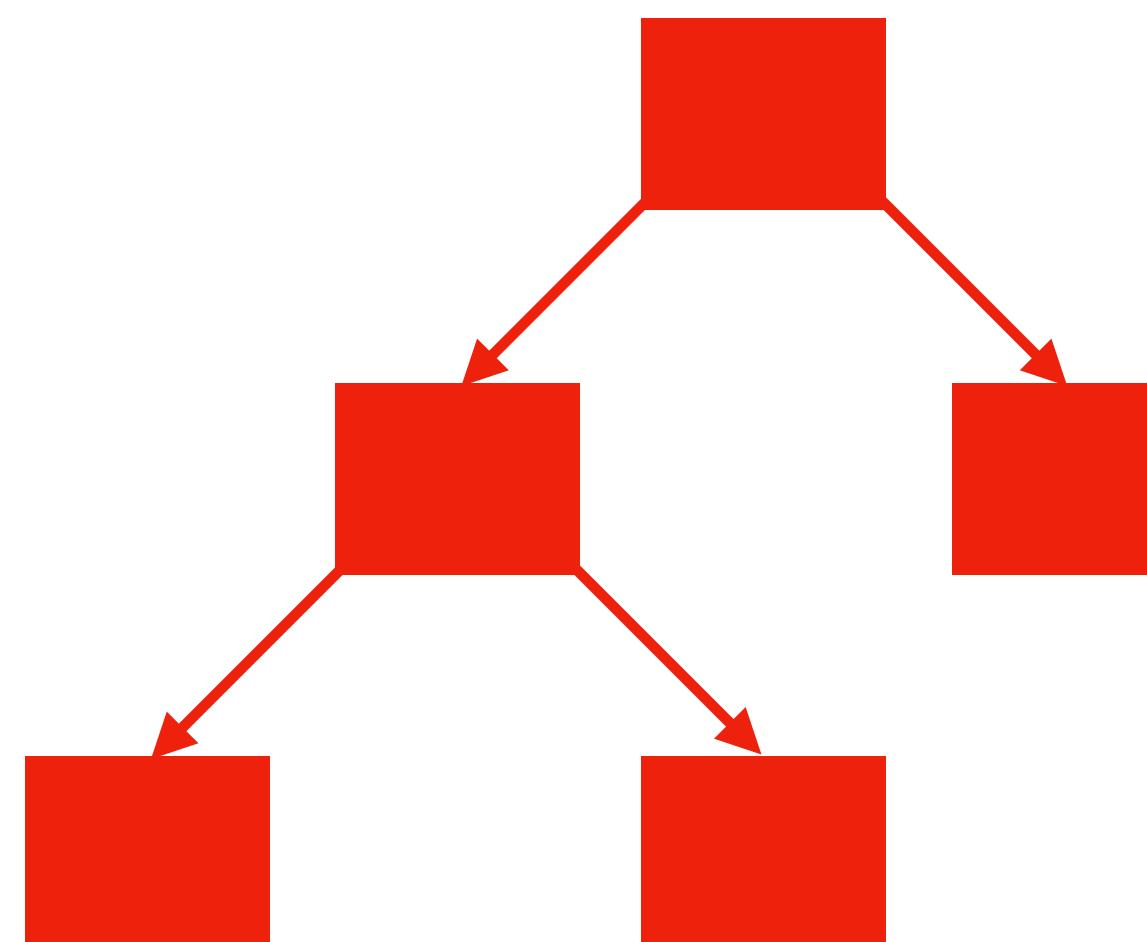
胸悶	好的血液循環	動脈阻塞	心臟疾病
Yes	Yes	No	No
No	No	No	No
Yes	Yes	No	No
Yes	No	No	Yes

# Original Dataset

胸悶	好的血液循環	動脈阻塞	心臟疾病
No	No	No	No
Yes	Yes	Yes	Yes
Yes	Yes	No	No
Yes	No	No	Yes
...	...	...	...

# Bootstrapped Dataset

tree1



# Original Dataset

胸悶	好的血液循環	動脈阻塞	心臟疾病
No	No	No	No
Yes	Yes	Yes	Yes
Yes	Yes	No	No
Yes	No	No	Yes
...	...	...	...

# Bootstrapped Dataset

胸悶	好的血液循環	動脈阻塞	心臟疾病
No	No	No	No

# Original Dataset

胸悶	好的血液循環	動脈阻塞	心臟疾病
No	No	No	No
Yes	Yes	Yes	Yes
Yes	Yes	No	No
Yes	No	No	Yes
...	...	...	...

# Bootstrapped Dataset

胸悶	好的血液循環	動脈阻塞	心臟疾病
Yes	No	No	Yes

# Original Dataset

胸悶	好的血液循環	動脈阻塞	心臟疾病
No	No	No	No
Yes	Yes	Yes	Yes
Yes	Yes	No	No
Yes	No	No	Yes
...	...	...	...

# Bootstrapped Dataset

胸悶	好的血液循環	動脈阻塞	心臟疾病
Yes	No	No	Yes
Yes	No	No	Yes
...	...	...	...

# Original Dataset

胸悶	好的血液循環	動脈阻塞	心臟疾病
No	No	No	No
Yes	Yes	Yes	Yes
Yes	Yes	No	No
Yes	No	No	Yes
...	...	...	...

# Bootstrapped Dataset

胸悶	好的血液循環	動脈阻塞	心臟疾病
Yes	No	No	Yes
Yes	No	No	Yes
No	No	No	No

# Original Dataset

胸悶	好的血液循環	動脈阻塞	心臟疾病
No	No	No	No
Yes	Yes	Yes	Yes
Yes	Yes	No	No
Yes	No	No	Yes
...	...	...	...

# Bootstrapped Dataset

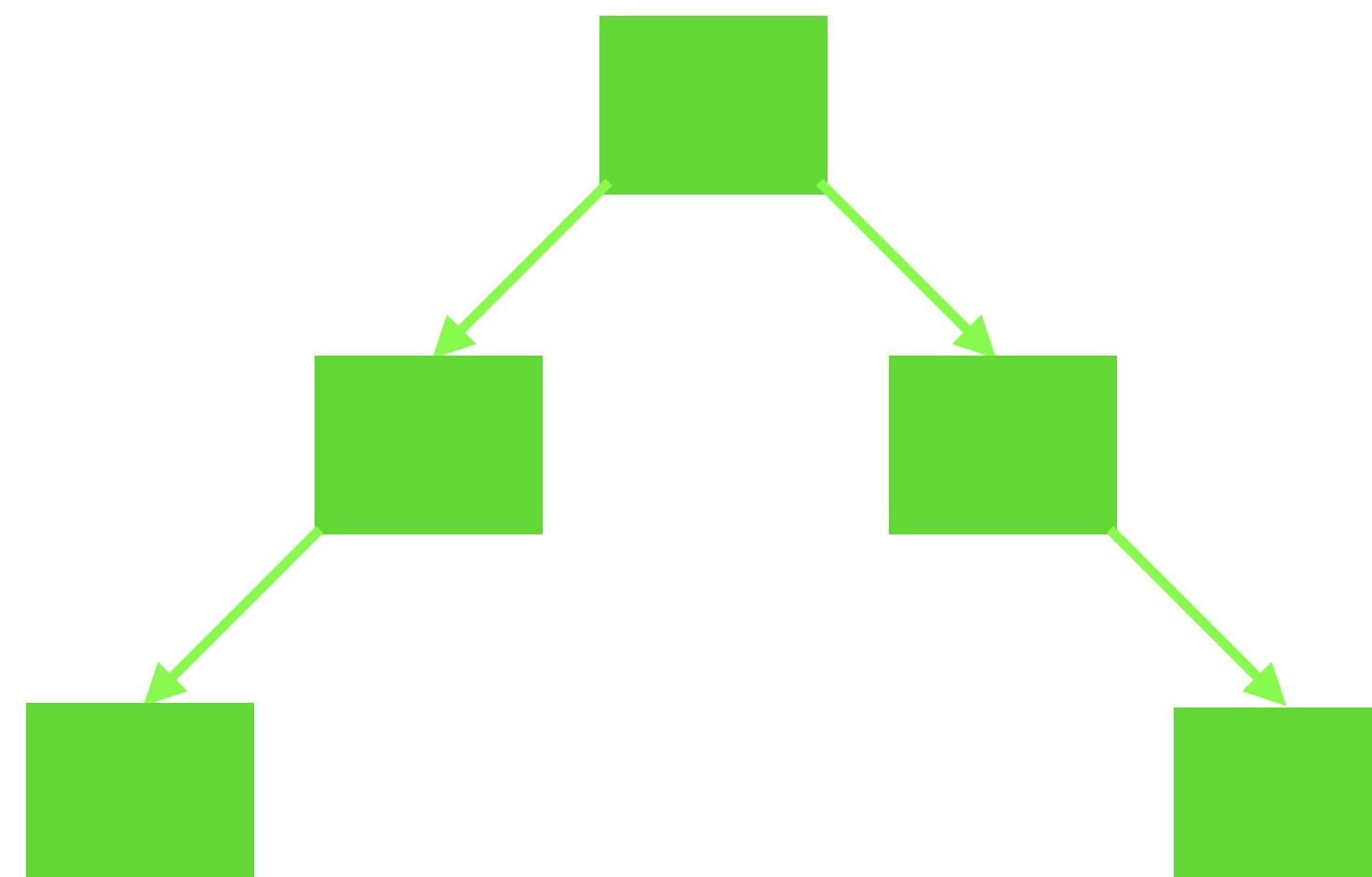
胸悶	好的血液循環	動脈阻塞	心臟疾病
Yes	No	No	Yes
Yes	No	No	Yes
No	No	No	No
Yes	No	No	Yes

# Original Dataset

胸悶	好的血液循環	動脈阻塞	心臟疾病
No	No	No	No
Yes	Yes	Yes	Yes
Yes	Yes	No	No
Yes	No	No	Yes
...	...	...	...

# Bootstrapped Dataset

tree2



# Original Dataset

胸悶	好的血液循環	動脈阻塞	心臟疾病
No	No	No	No
Yes	Yes	Yes	Yes
Yes	Yes	No	No
Yes	No	No	Yes
...	...	...	...

# Bootstrapped Dataset

胸悶	好的血液循環	動脈阻塞	心臟疾病
No	No	No	No

# Original Dataset

胸悶	好的血液循環	動脈阻塞	心臟疾病
No	No	No	No
Yes	Yes	Yes	Yes
Yes	Yes	No	No
Yes	No	No	Yes
...	...	...	...

# Bootstrapped Dataset

胸悶	好的血液循環	動脈阻塞	心臟疾病
Yes	Yes	Yes	Yes

# Original Dataset

胸悶	好的血液循環	動脈阻塞	心臟疾病
No	No	No	No
Yes	Yes	Yes	Yes
Yes	Yes	No	No
Yes	No	No	Yes
...	...	...	...

# Bootstrapped Dataset

胸悶	好的血液循環	動脈阻塞	心臟疾病
Yes	Yes	Yes	Yes
Yes	Yes	Yes	Yes
...	...	...	...

# Original Dataset

胸悶	好的血液循環	動脈阻塞	心臟疾病
No	No	No	No
Yes	Yes	Yes	Yes
Yes	Yes	No	No
Yes	No	No	Yes
...	...	...	...

# Bootstrapped Dataset

胸悶	好的血液循環	動脈阻塞	心臟疾病
Yes	Yes	Yes	Yes
Yes	Yes	Yes	Yes
Yes	Yes	No	No
Yes	Yes	No	Yes

# Original Dataset

胸悶	好的血液循環	動脈阻塞	心臟疾病
No	No	No	No
Yes	Yes	Yes	Yes
Yes	Yes	No	No
Yes	No	No	Yes
...	...	...	...

# Bootstrapped Dataset

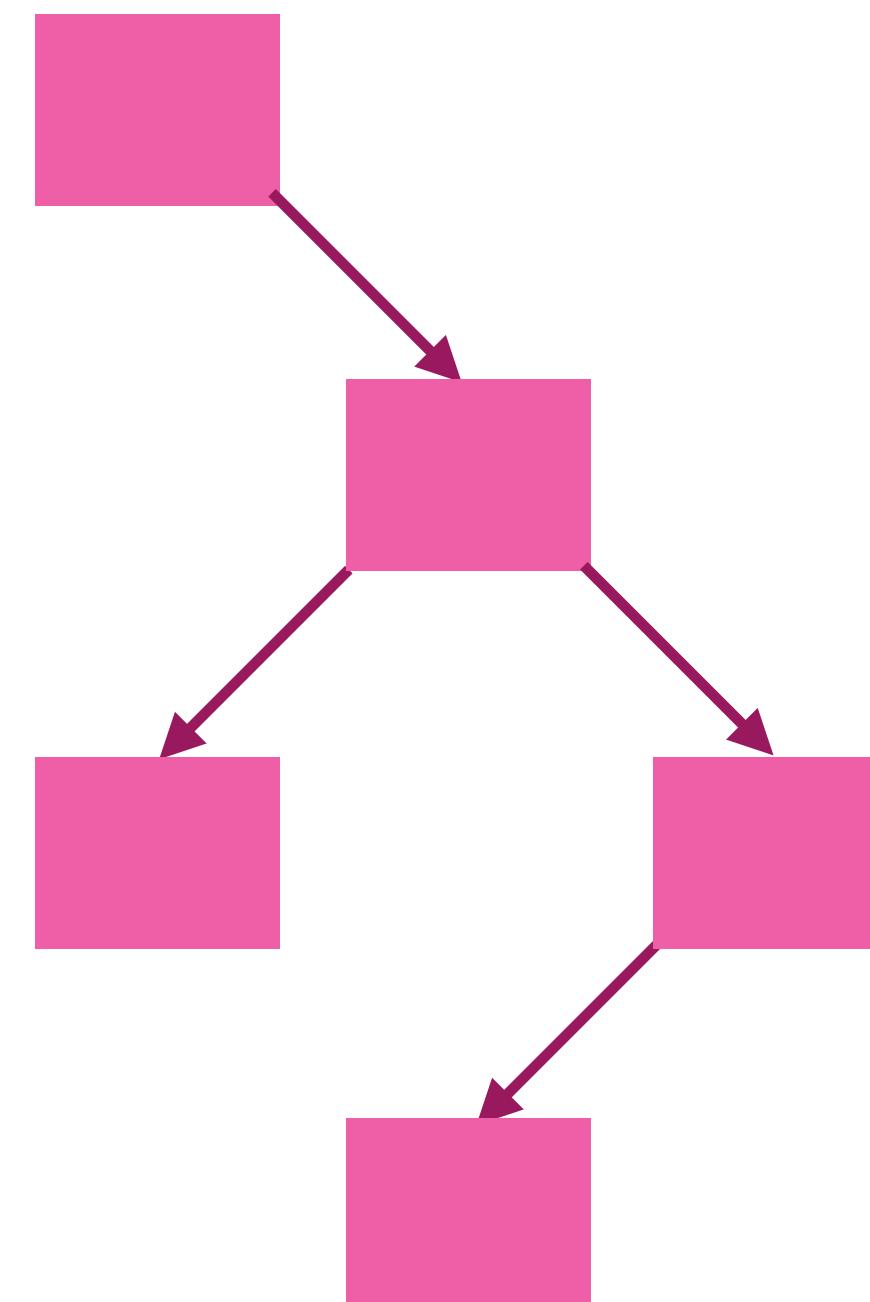
胸悶	好的血液循環	動脈阻塞	心臟疾病
Yes	Yes	Yes	Yes
Yes	Yes	No	No
Yes	No	No	Yes

# Original Dataset

胸悶	好的血液循環	動脈阻塞	心臟疾病
No	No	No	No
Yes	Yes	Yes	Yes
Yes	Yes	No	No
Yes	No	No	Yes
...	...	...	...

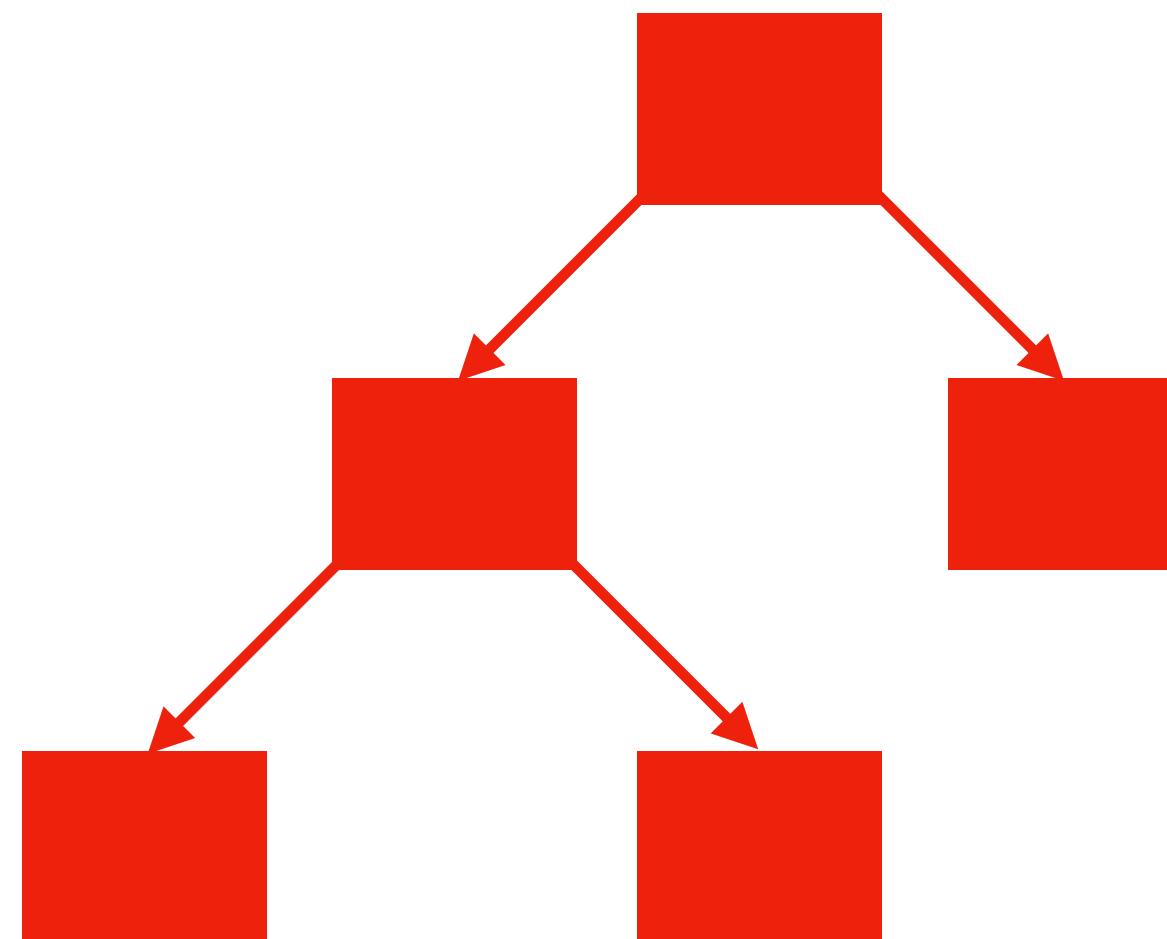
# Bootstrapped Dataset

tree3



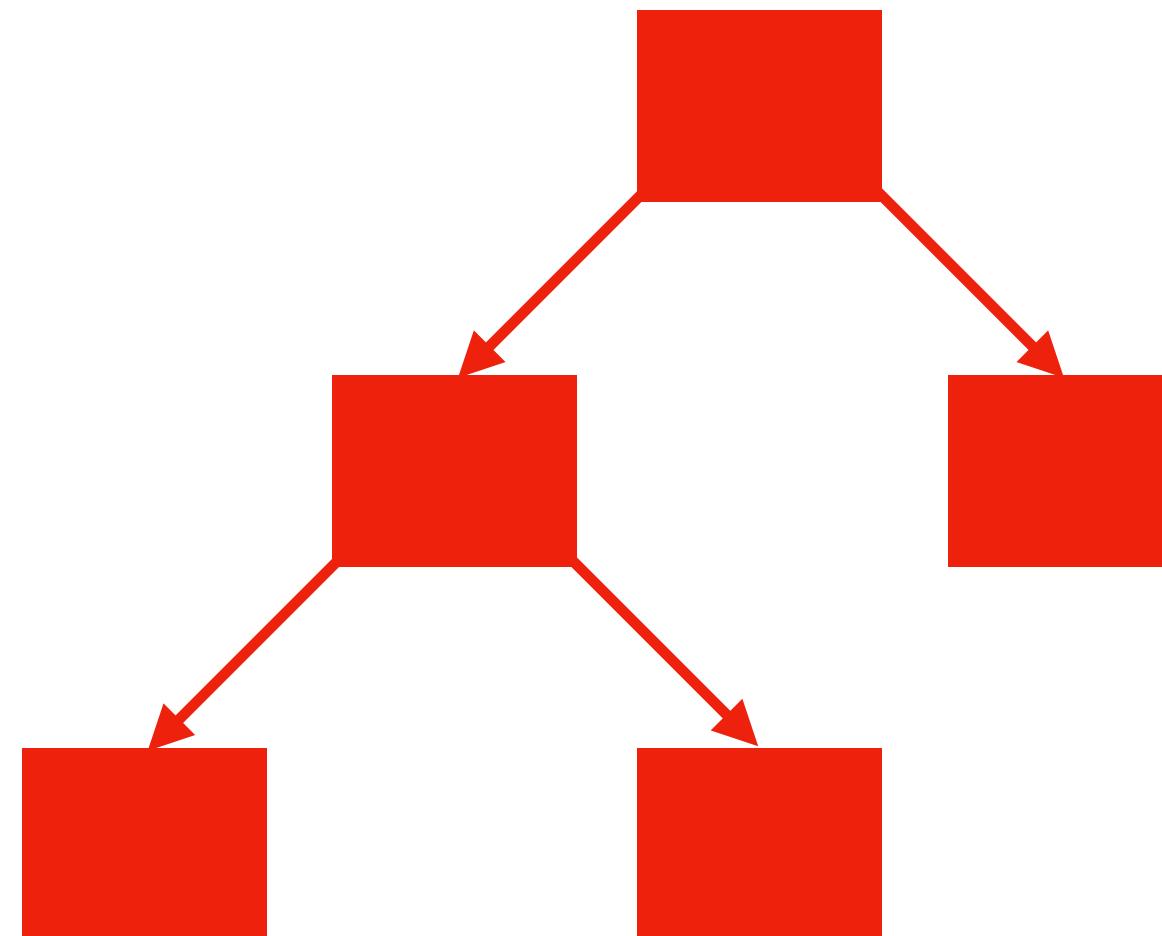
# data

## tree1

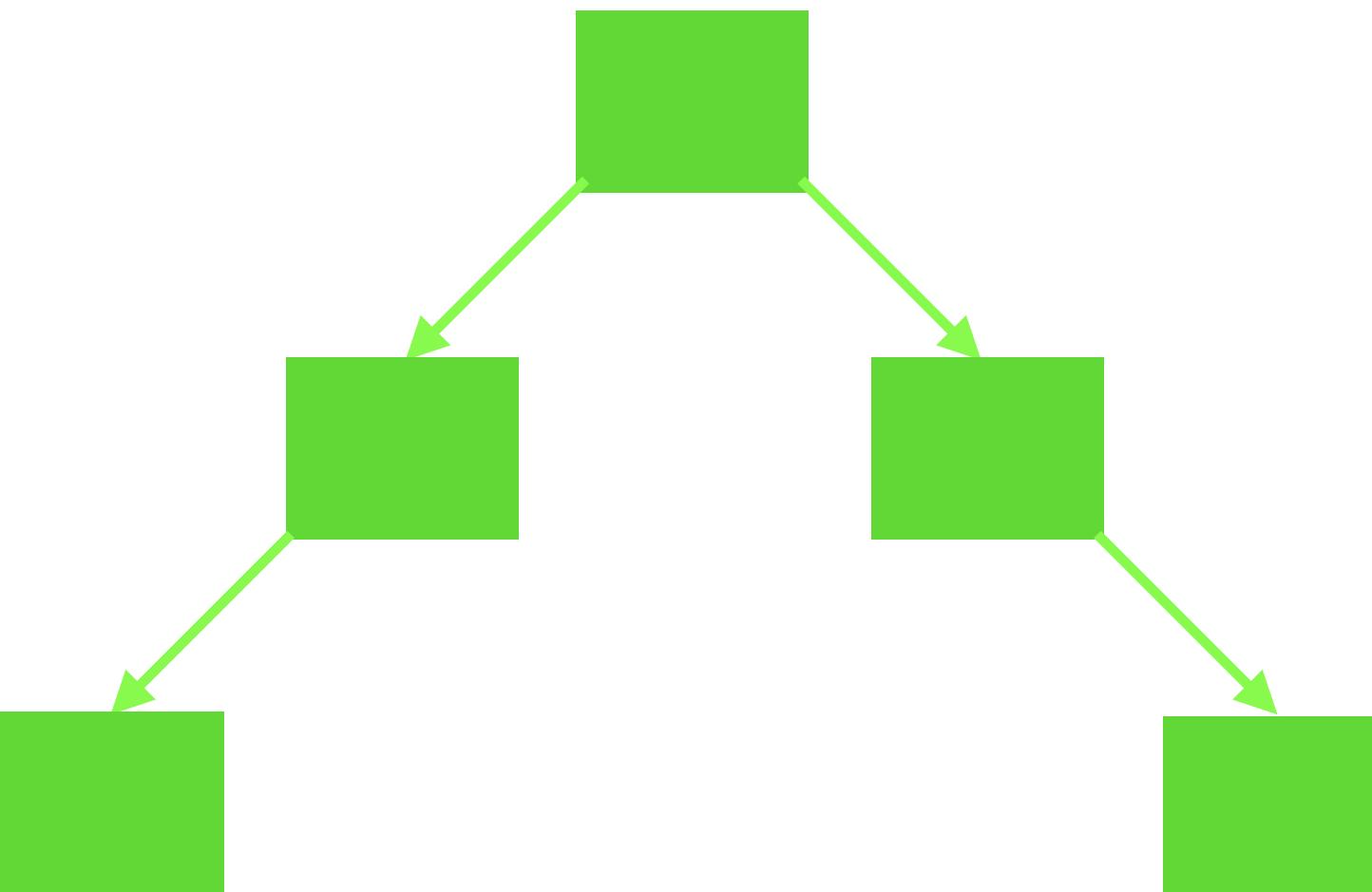


# data

tree1

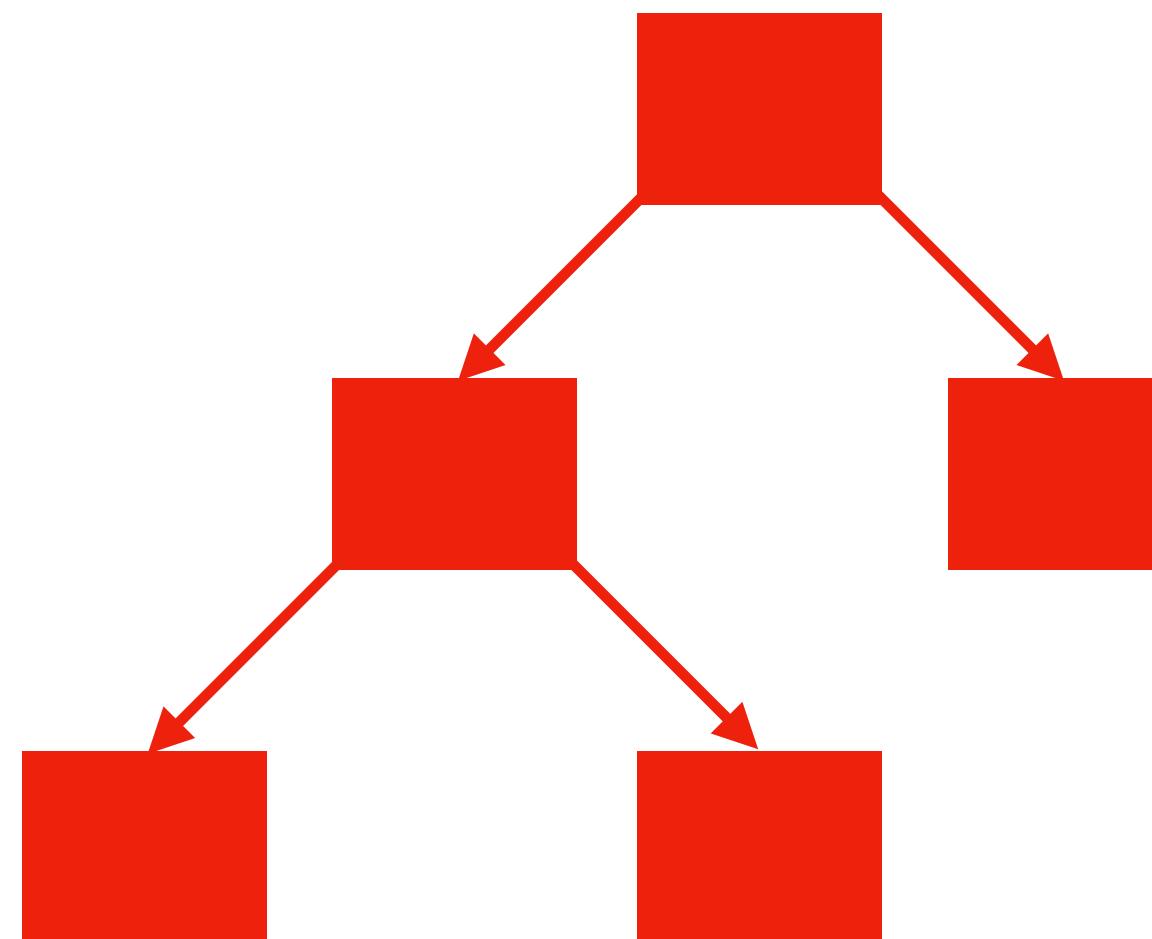


tree2

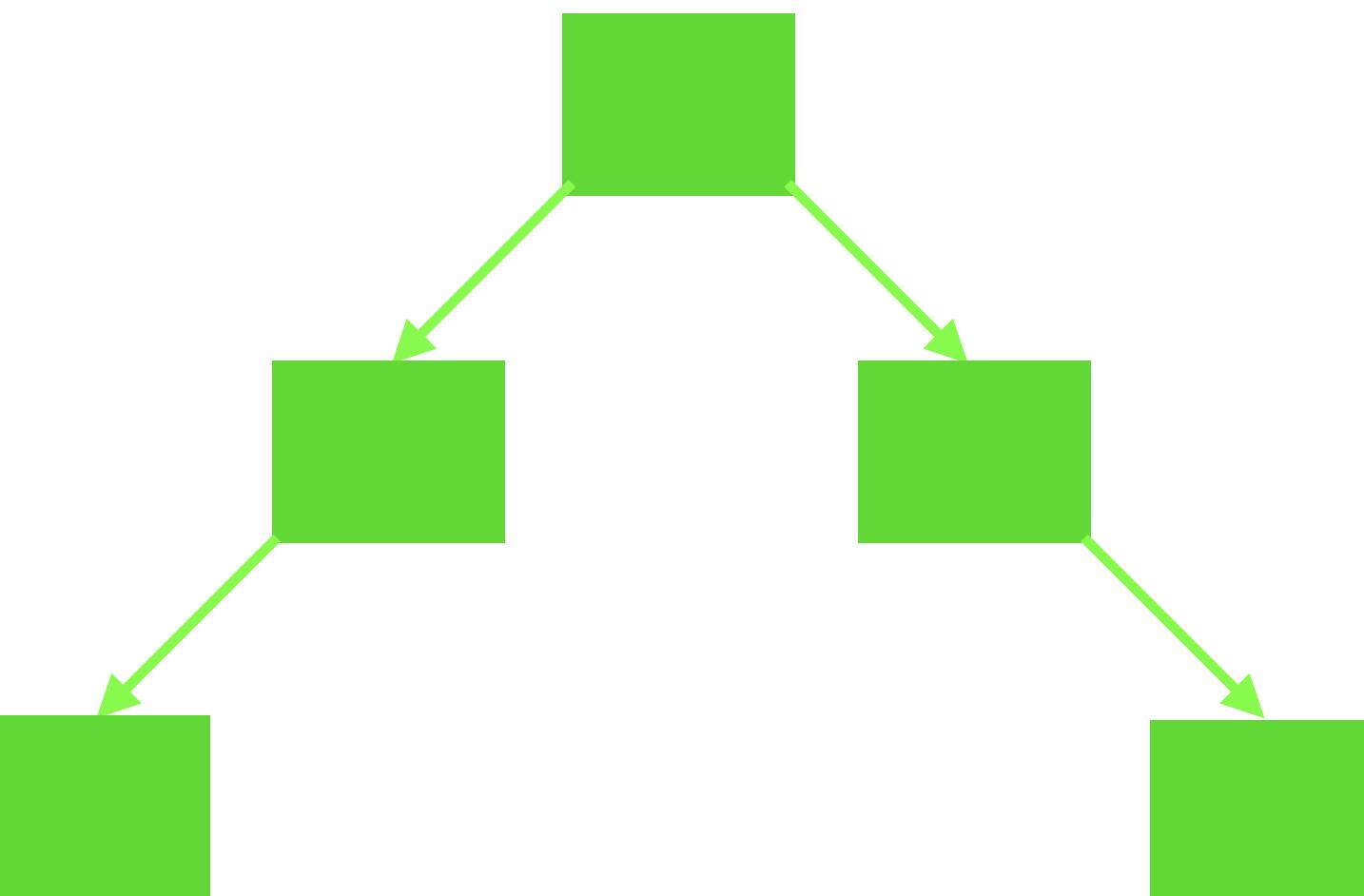


# data

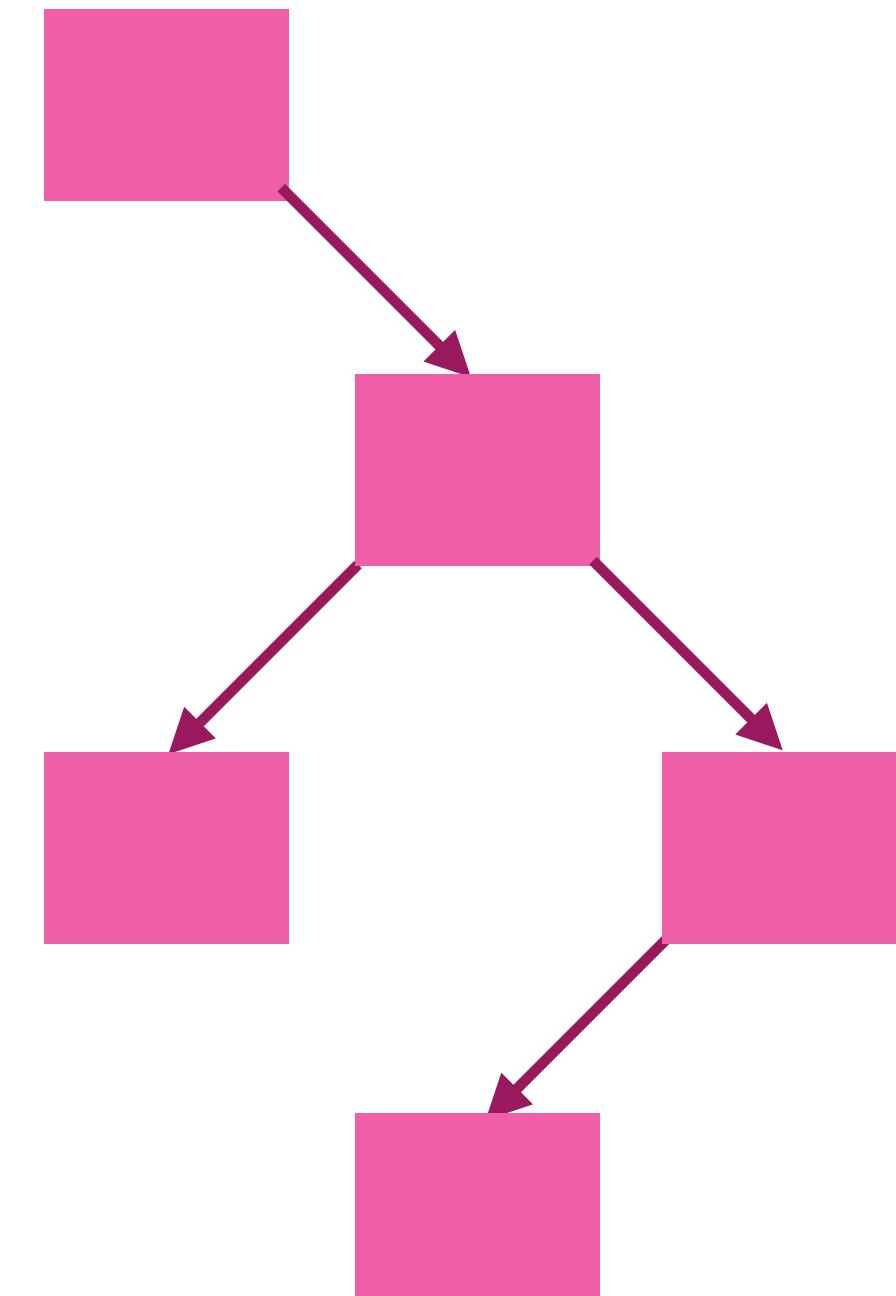
## tree1



## tree2

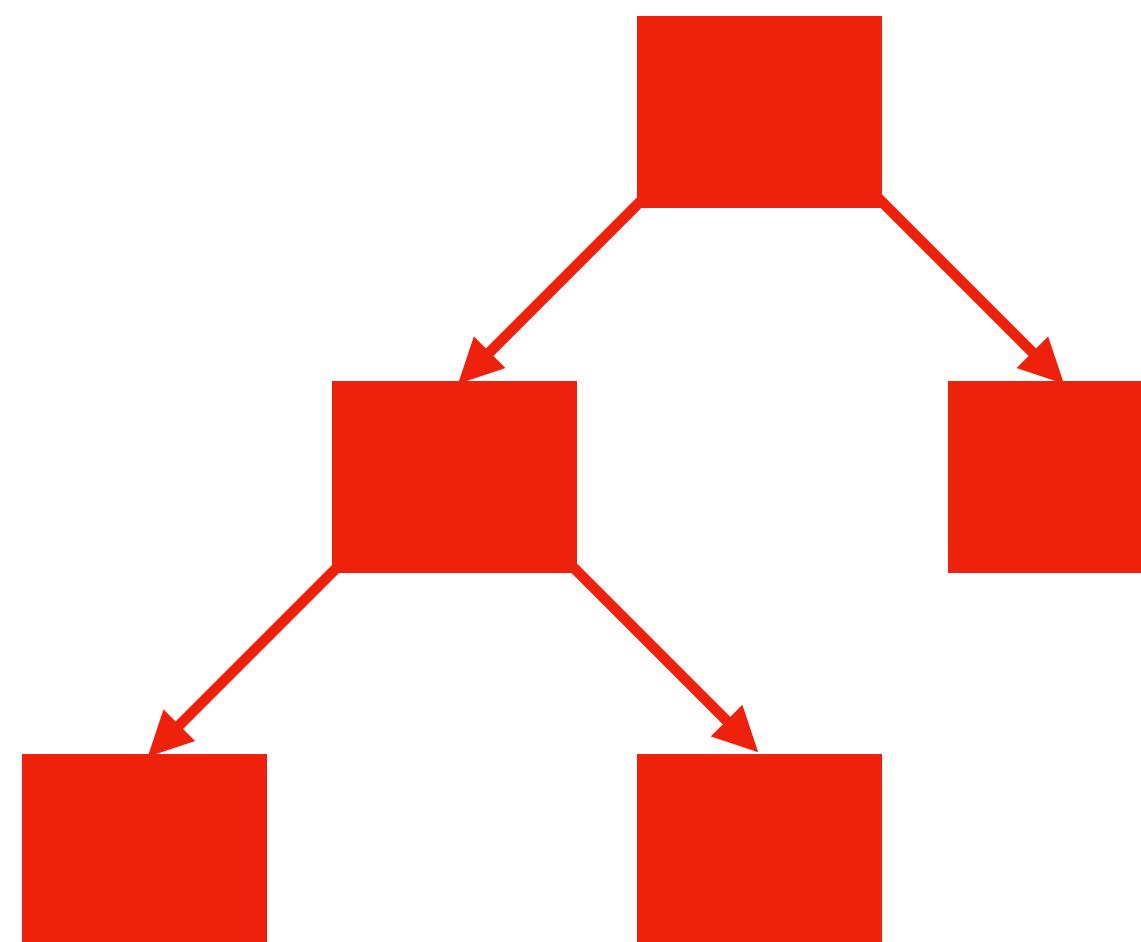


## tree3

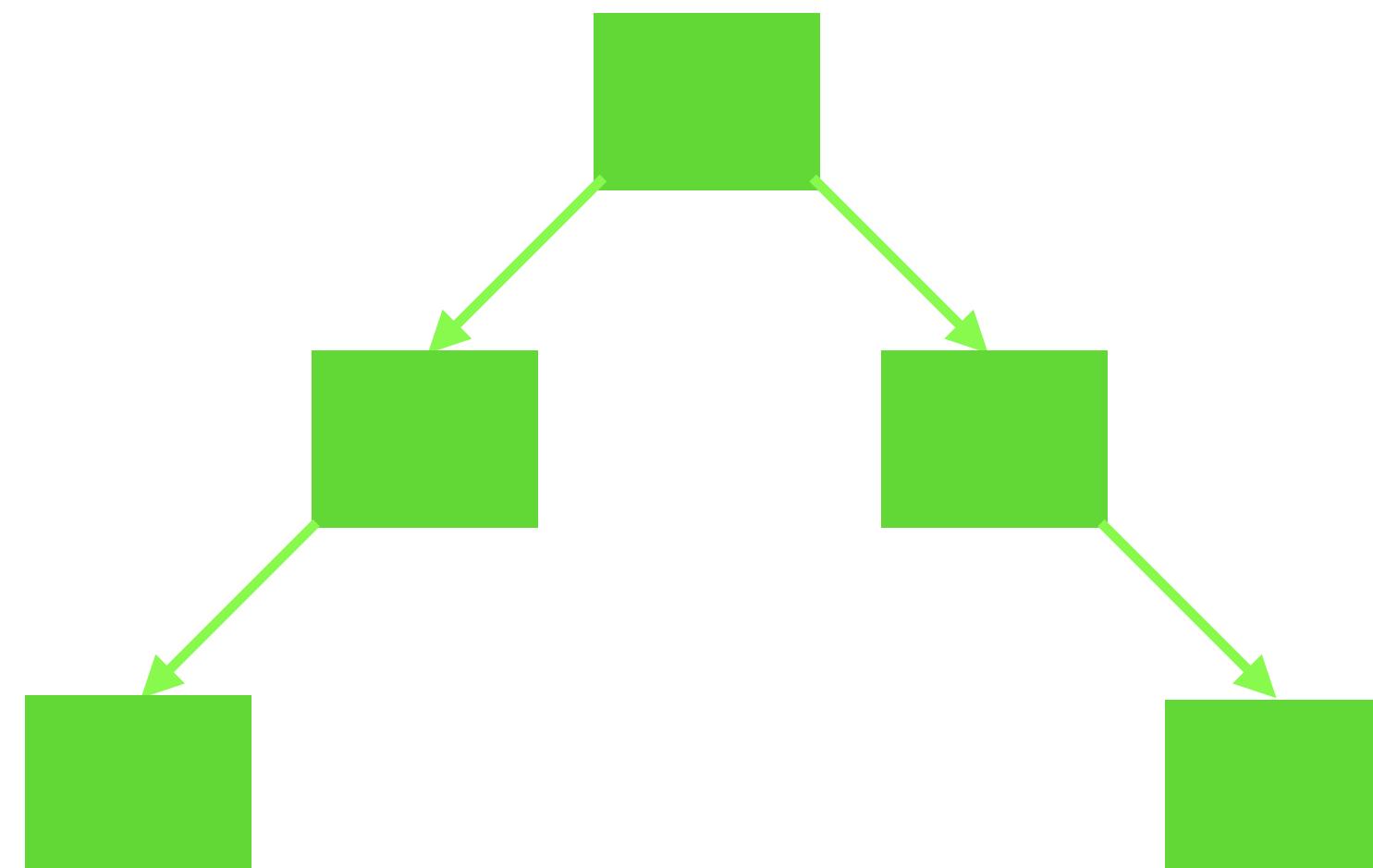


data

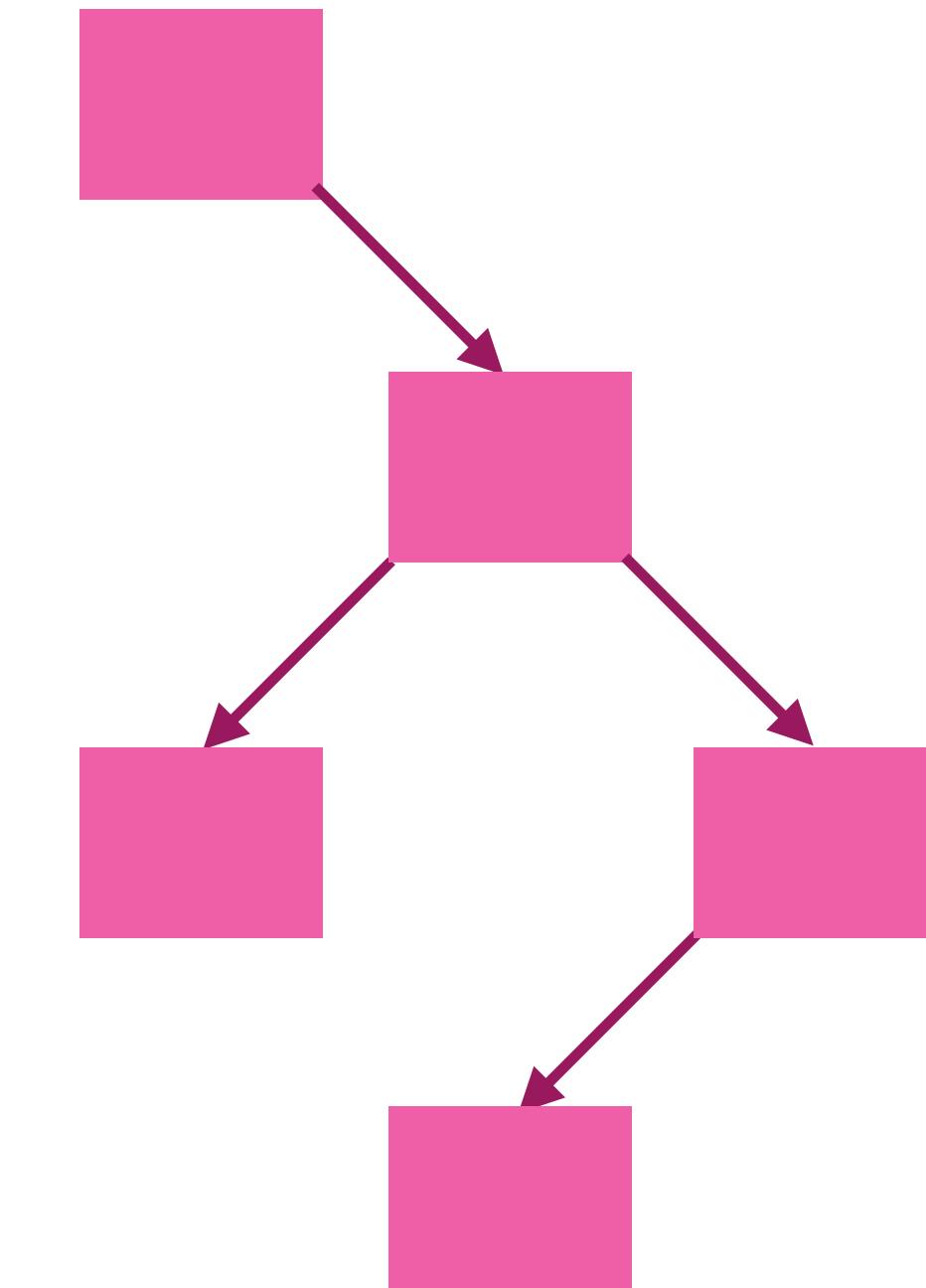
tree1



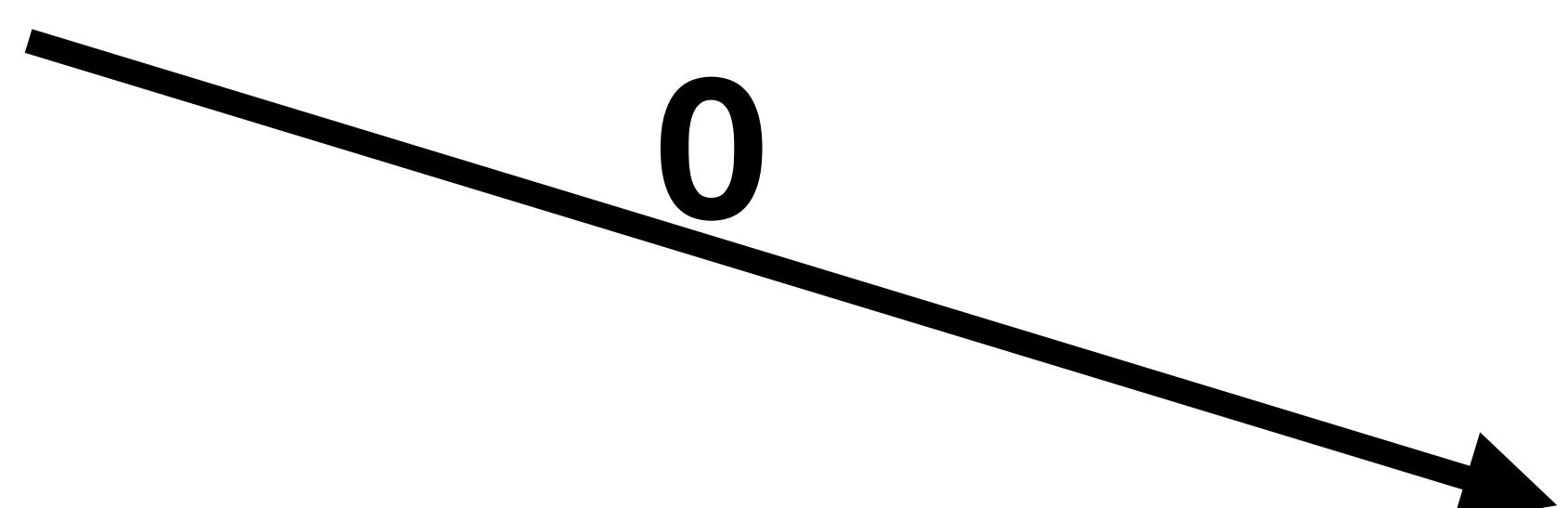
tree2



tree3

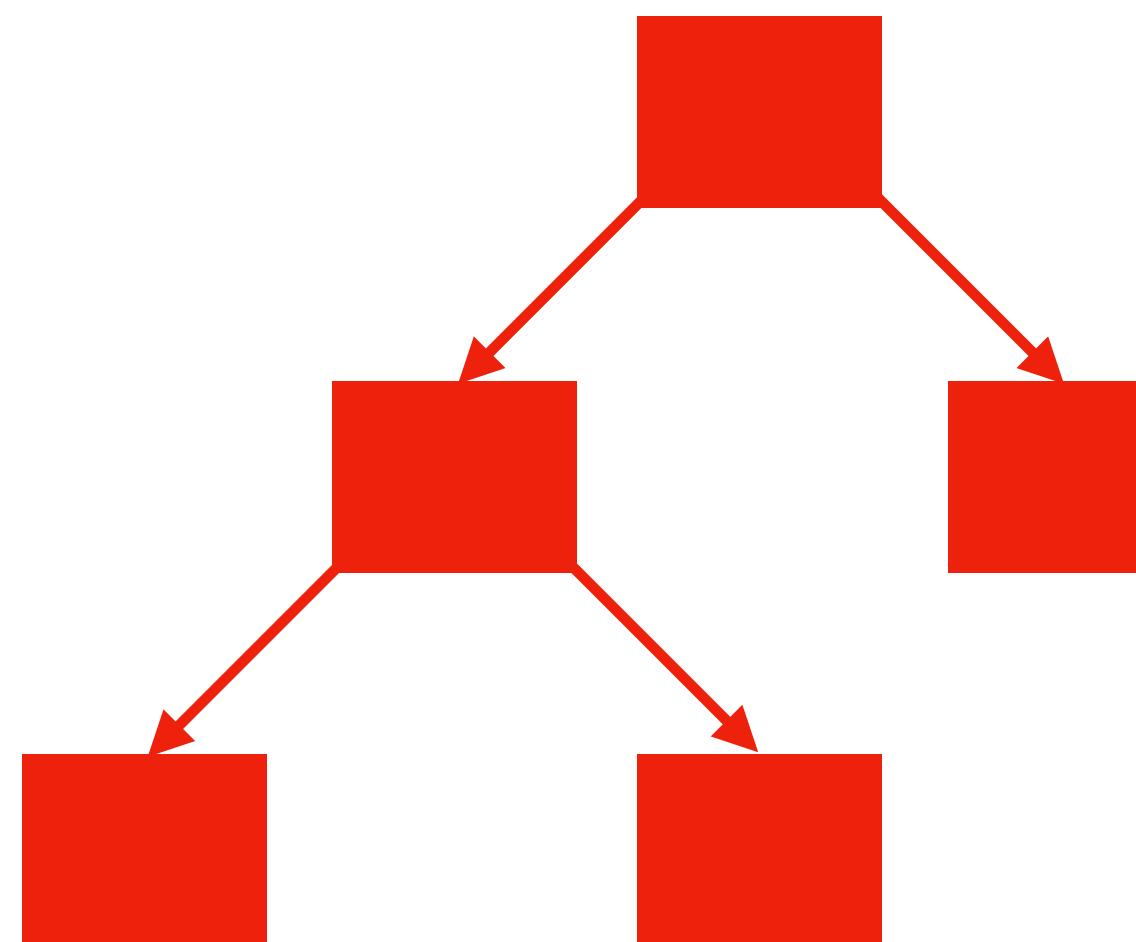


0

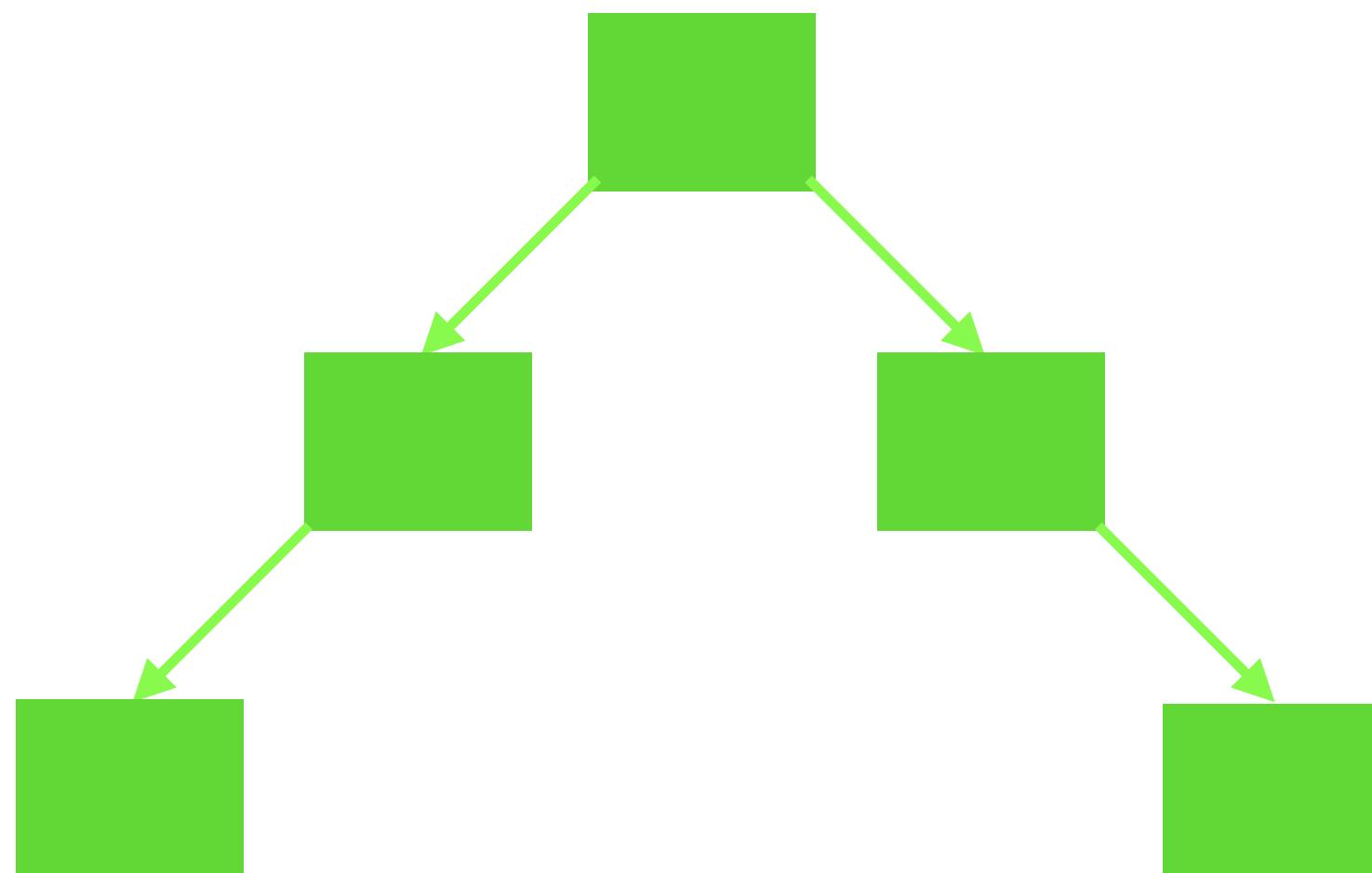


data

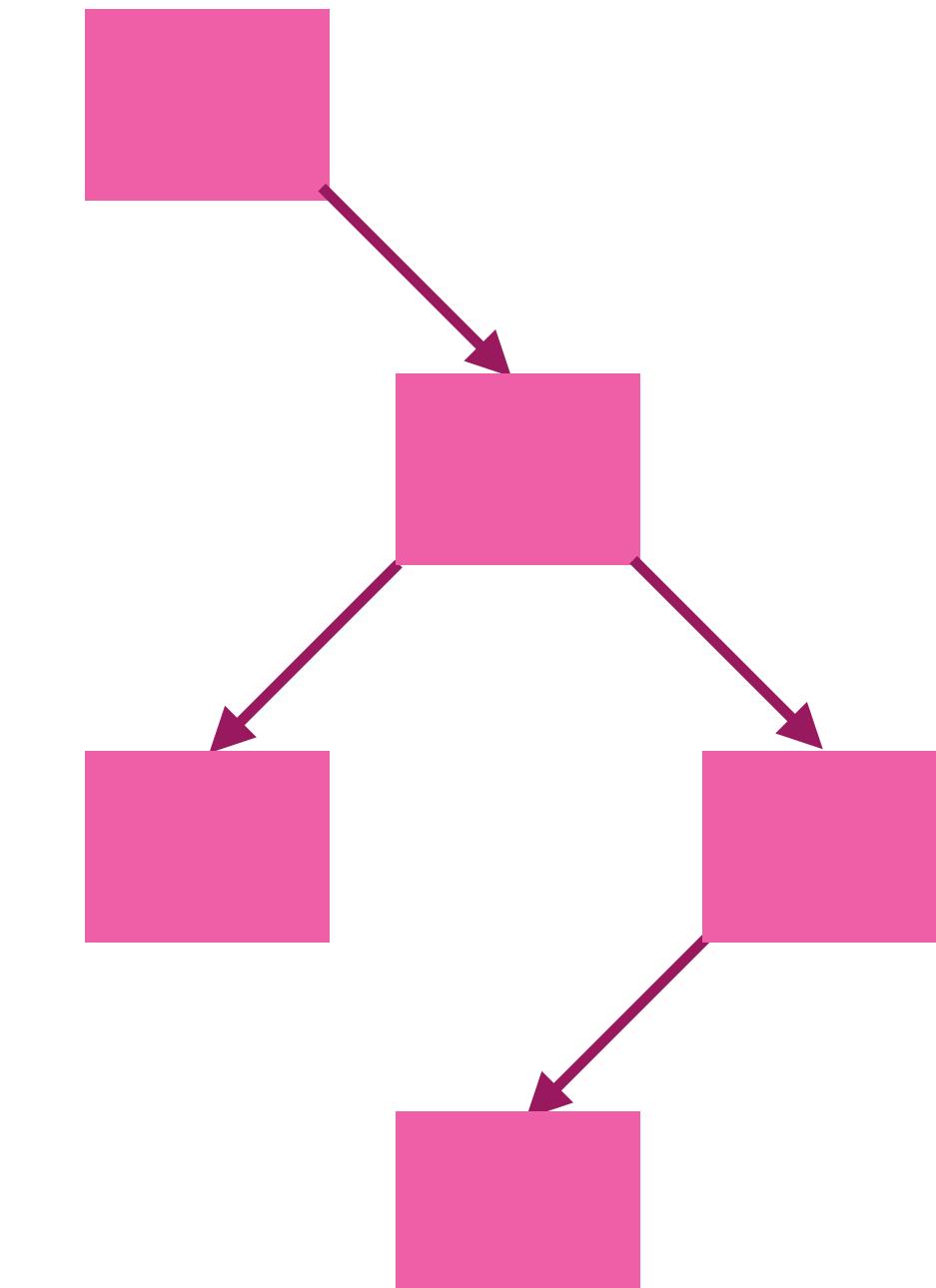
tree1



tree2



tree3

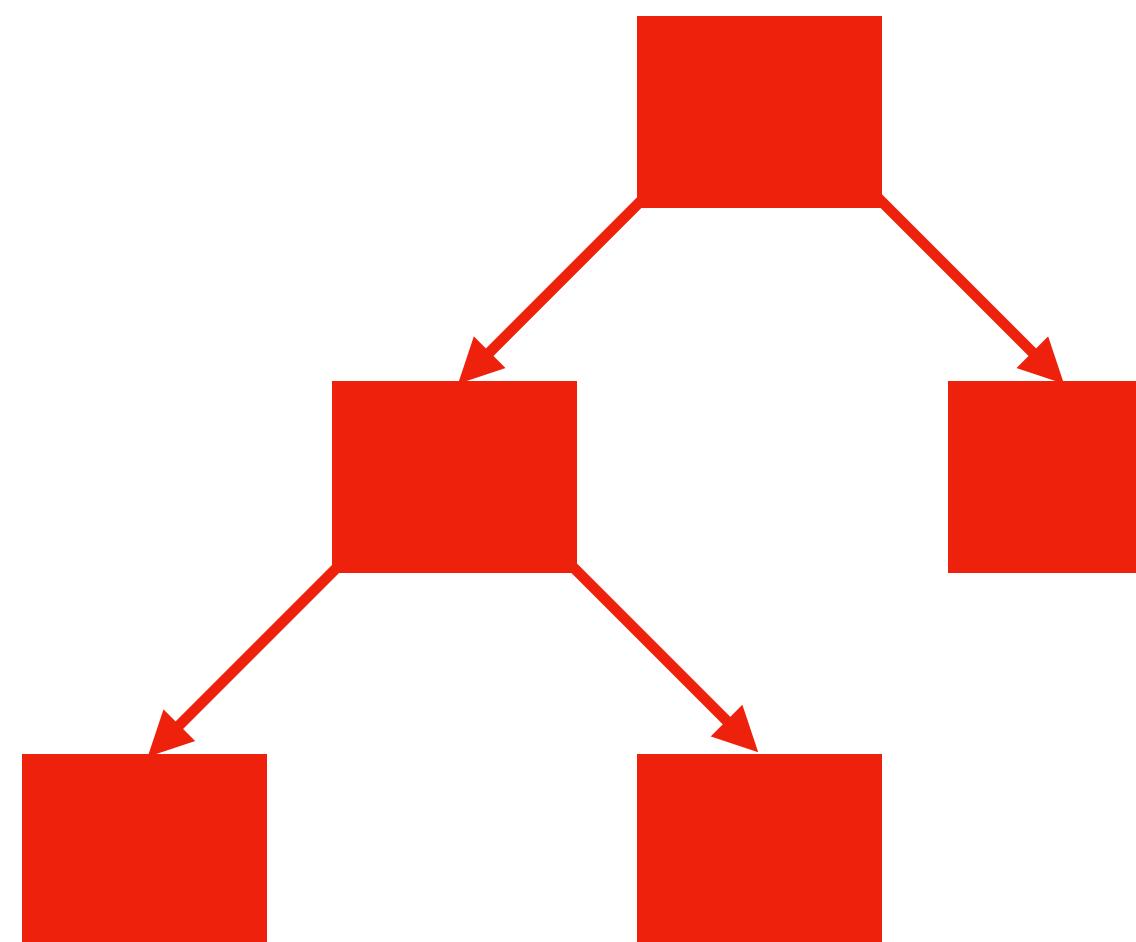


0

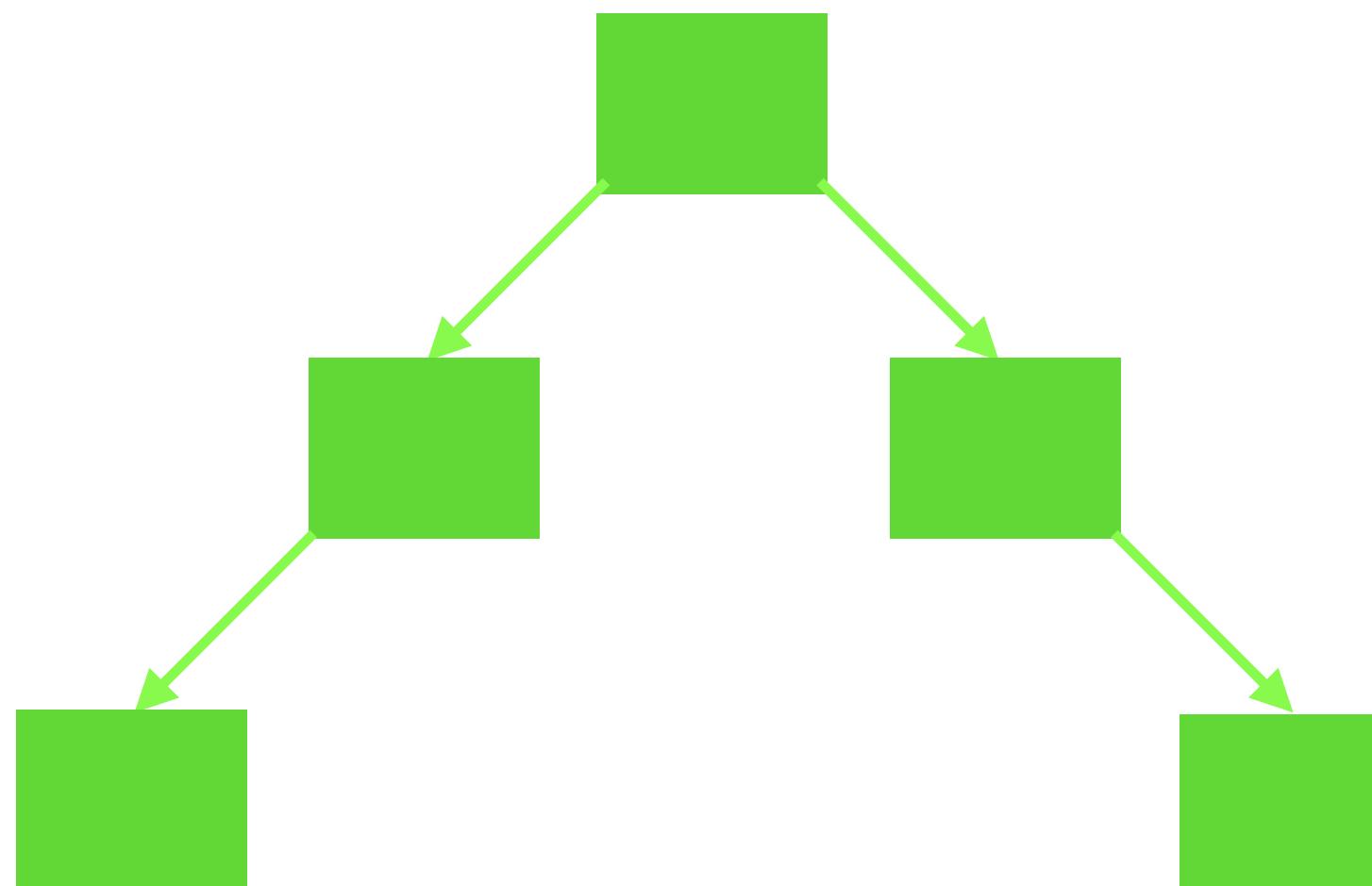
1

data

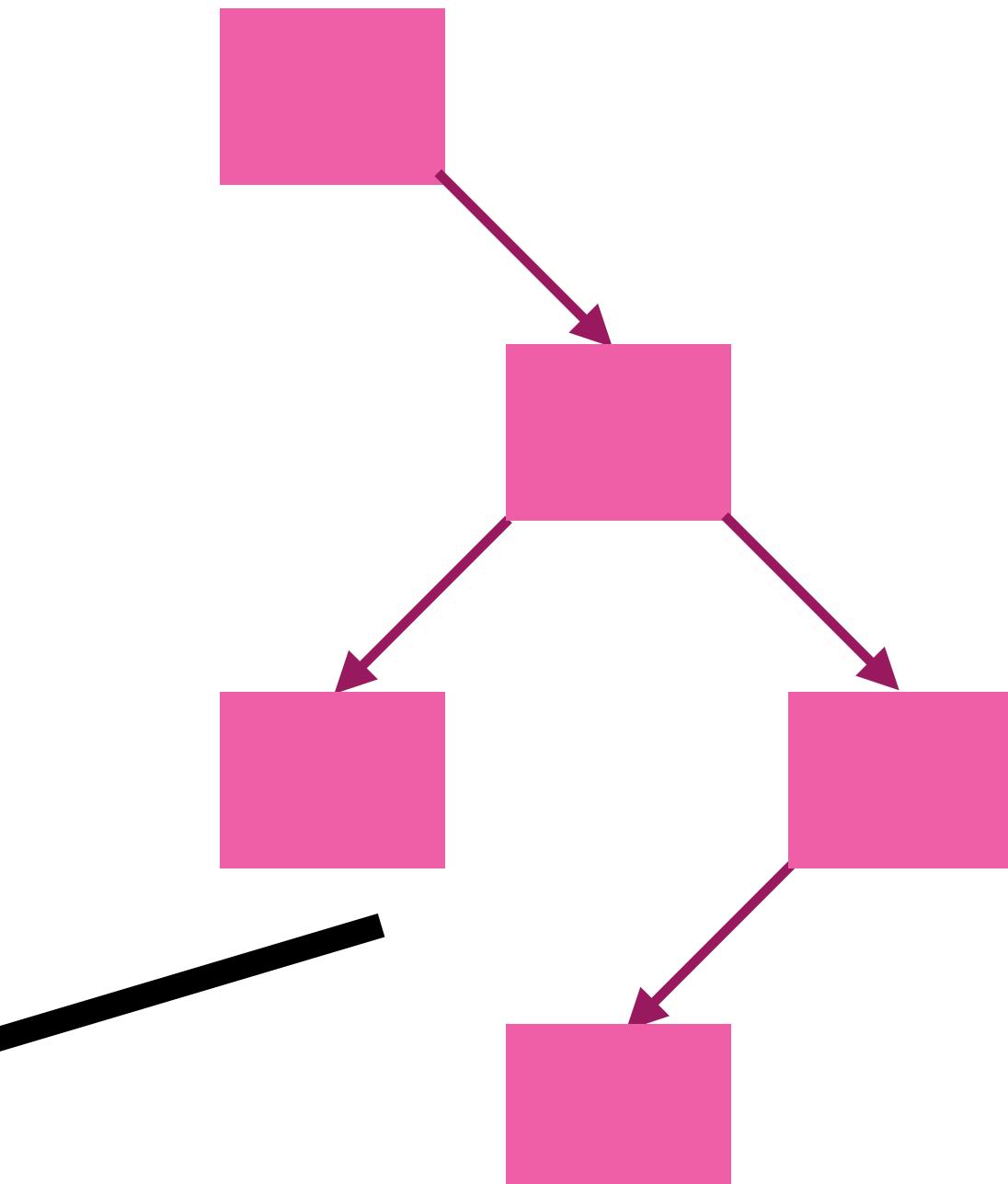
tree1



tree2



tree3



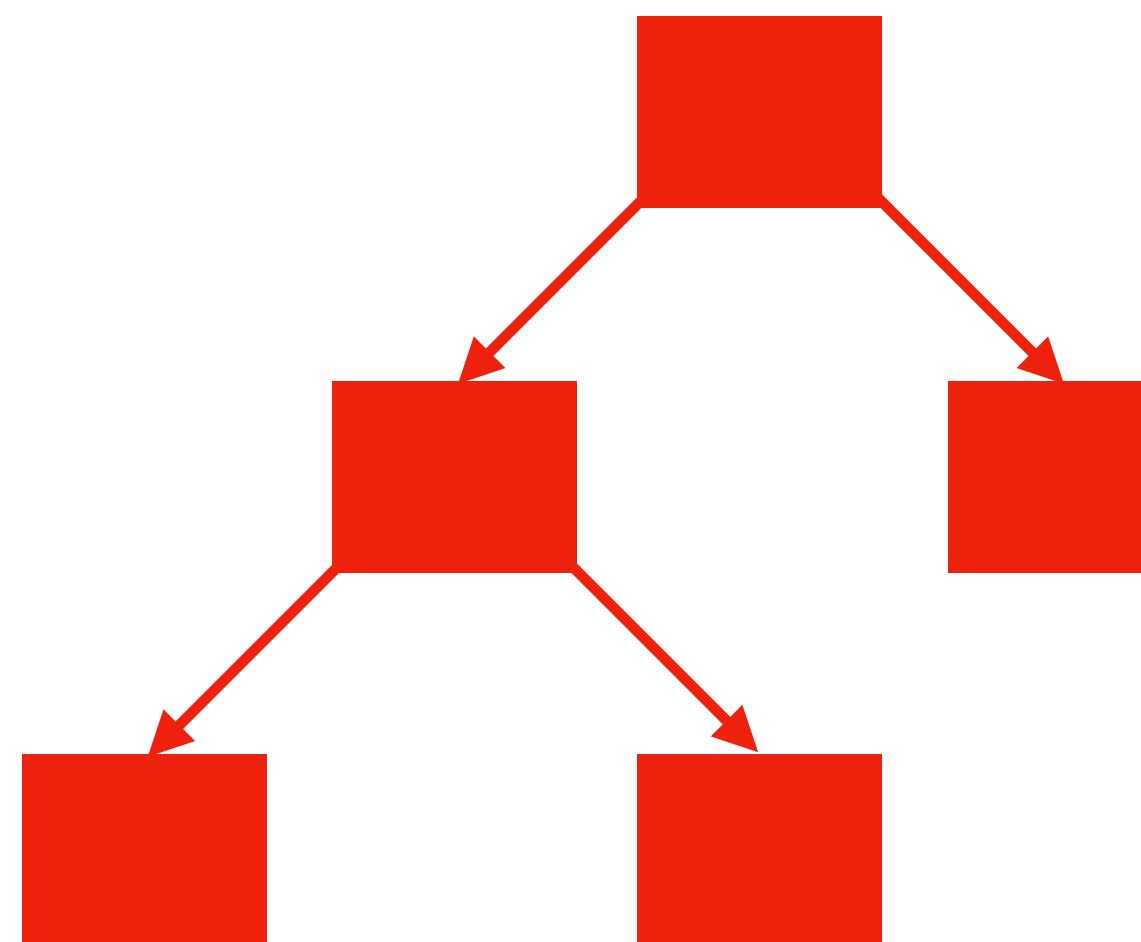
0

1

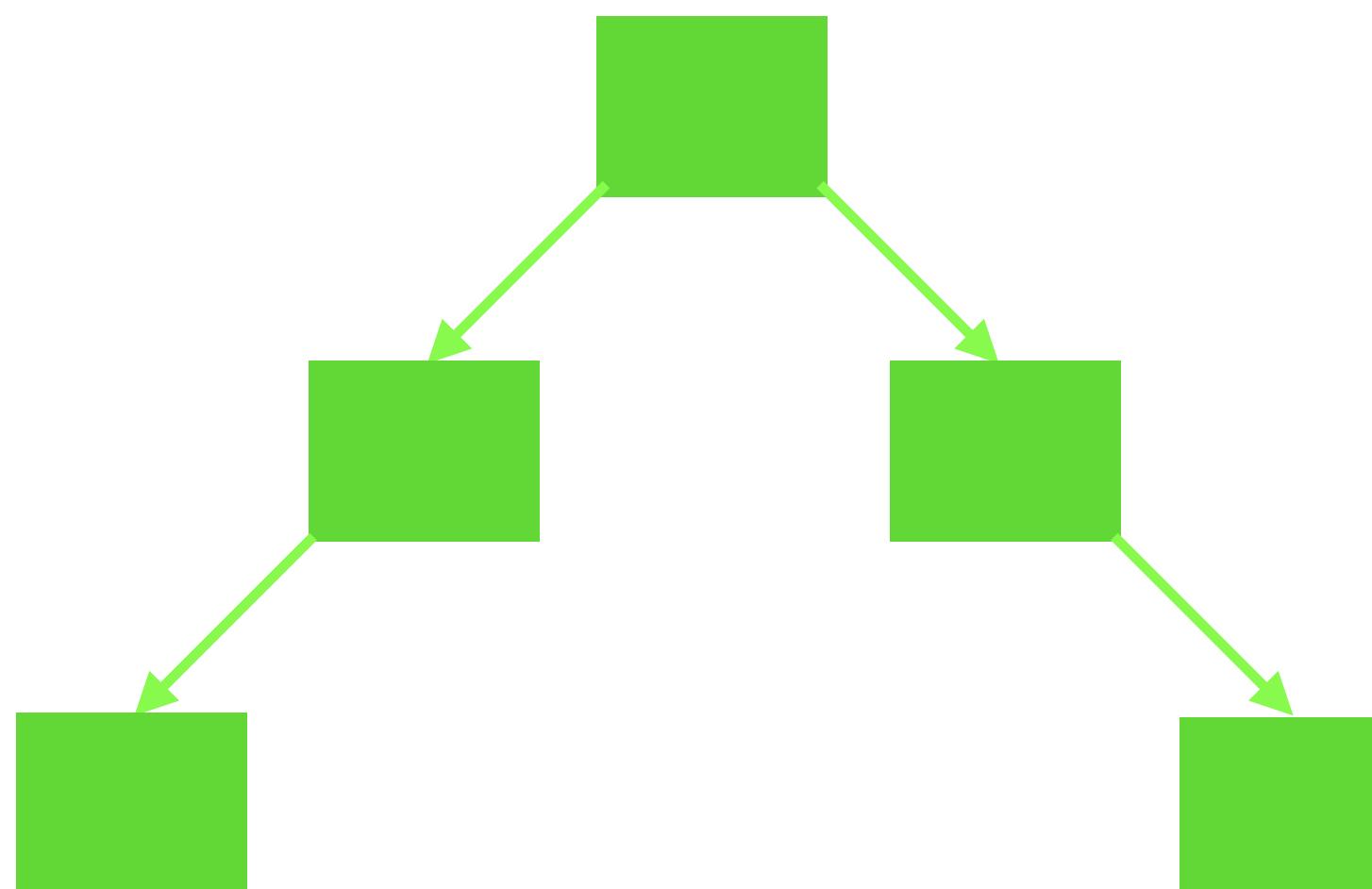
0

data

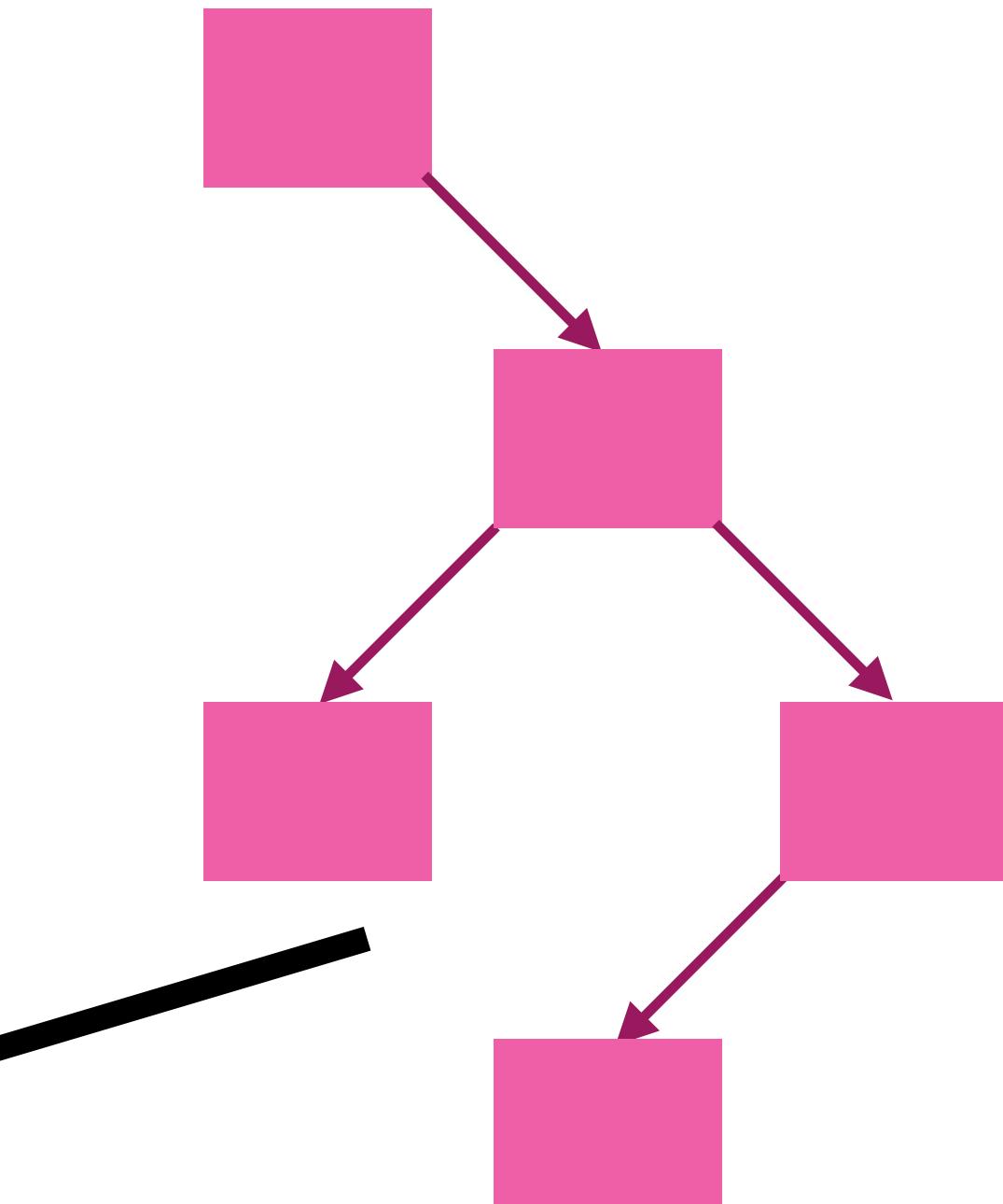
tree1



tree2



tree3



0

1

0

0

# Bagging

# BaggingClassifier

# Original Dataset

胸悶	好的血液循環	動脈阻塞	心臟疾病
No	No	No	No
Yes	Yes	Yes	Yes
Yes	Yes	No	No
Yes	No	No	Yes
...	...	...	...

# Bootstrapped Dataset1

胸悶	好的血液循環	動脈阻塞	心臟疾病
No	No	No	No

# Original Dataset

胸悶	好的血液循環	動脈阻塞	心臟疾病
No	No	No	No
Yes	Yes	Yes	Yes
Yes	Yes	No	No
Yes	No	No	Yes
...	...	...	...

# Bootstrapped Dataset1

胸悶	好的血液循環	動脈阻塞	心臟疾病
No	No	No	No

# Original Dataset

胸悶	好的血液循環	動脈阻塞	心臟疾病
No	No	No	No
Yes	Yes	Yes	Yes
Yes	Yes	No	No
Yes	No	No	Yes
...	...	...	...

# Bootstrapped Dataset1

胸悶	好的血液循環	動脈阻塞	心臟疾病
Yes	Yes	No	No

# Original Dataset

胸悶	好的血液循環	動脈阻塞	心臟疾病
No	No	No	No
Yes	Yes	Yes	Yes
Yes	Yes	No	No
Yes	No	No	Yes
...	...	...	...

# Bootstrapped Dataset1

胸悶	好的血液循環	動脈阻塞	心臟疾病
Yes	Yes	No	No

# Original Dataset

胸悶	好的血液循環	動脈阻塞	心臟疾病
No	No	No	No
Yes	Yes	Yes	Yes
Yes	Yes	No	No
Yes	No	No	Yes
...	...	...	...

# Bootstrapped Dataset1

胸悶	好的血液循環	動脈阻塞	心臟疾病
Yes	Yes	No	No
No	No	No	No

# Original Dataset

胸悶	好的血液循環	動脈阻塞	心臟疾病
No	No	No	No
Yes	Yes	Yes	Yes
Yes	Yes	No	No
Yes	No	No	Yes
...	...	...	...

# Bootstrapped Dataset1

胸悶	好的血液循環	動脈阻塞	心臟疾病
Yes	Yes	No	No
No	No	No	No
...	...	...	...

# Original Dataset

胸悶	好的血液循環	動脈阻塞	心臟疾病
No	No	No	No
Yes	Yes	Yes	Yes
Yes	Yes	No	No
Yes	No	No	Yes
...	...	...	...

# Bootstrapped Dataset1

胸悶	好的血液循環	動脈阻塞	心臟疾病
Yes	Yes	No	No
No	No	No	No
Yes	Yes	No	No
...	...	...	...

# Original Dataset

胸悶	好的血液循環	動脈阻塞	心臟疾病
No	No	No	No
Yes	Yes	Yes	Yes
Yes	Yes	No	No
Yes	No	No	Yes
...	...	...	...

# Bootstrapped Dataset1

胸悶	好的血液循環	動脈阻塞	心臟疾病
Yes	Yes	No	No
No	No	No	No
Yes	Yes	No	No

# Original Dataset

胸悶	好的血液循環	動脈阻塞	心臟疾病
No	No	No	No
Yes	Yes	Yes	Yes
Yes	Yes	No	No
Yes	No	No	Yes
...	...	...	...

# Bootstrapped Dataset1

胸悶	好的血液循環	動脈阻塞	心臟疾病
Yes	Yes	No	No
No	No	No	No
Yes	Yes	No	No
Yes	No	No	Yes

# Original Dataset

胸悶	好的血液循環	動脈阻塞	心臟疾病
No	No	No	No
Yes	Yes	Yes	Yes
Yes	Yes	No	No
Yes	No	No	Yes
...	...	...	...

# Bootstrapped Dataset2

胸悶	好的血液循環	動脈阻塞	心臟疾病
No	No	No	No

# Original Dataset

胸悶	好的血液循環	動脈阻塞	心臟疾病
No	No	No	No
Yes	Yes	Yes	Yes
Yes	Yes	No	No
Yes	No	No	Yes
...	...	...	...

# Bootstrapped Dataset2

胸悶	好的血液循環	動脈阻塞	心臟疾病
Yes	No	No	Yes

# Original Dataset

胸悶	好的血液循環	動脈阻塞	心臟疾病
No	No	No	No
Yes	Yes	Yes	Yes
Yes	Yes	No	No
Yes	No	No	Yes
...	...	...	...

# Bootstrapped Dataset2

胸悶	好的血液循環	動脈阻塞	心臟疾病
Yes	No	No	Yes
Yes	No	No	Yes
...	...	...	...

# Original Dataset

胸悶	好的血液循環	動脈阻塞	心臟疾病
No	No	No	No
Yes	Yes	Yes	Yes
Yes	Yes	No	No
Yes	No	No	Yes
...	...	...	...

# Bootstrapped Dataset2

胸悶	好的血液循環	動脈阻塞	心臟疾病
Yes	No	No	Yes
Yes	No	No	Yes
No	No	No	No

# Original Dataset

胸悶	好的血液循環	動脈阻塞	心臟疾病
No	No	No	No
Yes	Yes	Yes	Yes
Yes	Yes	No	No
Yes	No	No	Yes
...	...	...	...

# Bootstrapped Dataset2

胸悶	好的血液循環	動脈阻塞	心臟疾病
Yes	No	No	Yes
Yes	No	No	Yes
No	No	No	No
Yes	No	No	Yes

# Original Dataset

胸悶	好的血液循環	動脈阻塞	心臟疾病
No	No	No	No
Yes	Yes	Yes	Yes
Yes	Yes	No	No
Yes	No	No	Yes
...	...	...	...

# Bootstrapped Dataset3

胸悶	好的血液循環	動脈阻塞	心臟疾病
----	--------	------	------

# Original Dataset

胸悶	好的血液循環	動脈阻塞	心臟疾病
No	No	No	No
Yes	Yes	Yes	Yes
Yes	Yes	No	No
Yes	No	No	Yes
...	...	...	...

# Bootstrapped Dataset3

胸悶	好的血液循環	動脈阻塞	心臟疾病
Yes	Yes	Yes	Yes

# Original Dataset

胸悶	好的血液循環	動脈阻塞	心臟疾病
No	No	No	No
Yes	Yes	Yes	Yes
Yes	Yes	No	No
Yes	No	No	Yes
...	...	...	...

# Bootstrapped Dataset3

胸悶	好的血液循環	動脈阻塞	心臟疾病
Yes	Yes	Yes	Yes
Yes	Yes	Yes	Yes
Yes	Yes	No	No
Yes	No	No	Yes

# Original Dataset

胸悶	好的血液循環	動脈阻塞	心臟疾病
No	No	No	No
Yes	Yes	Yes	Yes
Yes	Yes	No	No
Yes	No	No	Yes
...	...	...	...

# Bootstrapped Dataset3

胸悶	好的血液循環	動脈阻塞	心臟疾病
Yes	Yes	Yes	Yes
Yes	Yes	Yes	Yes
Yes	Yes	No	No
...	...	...	...

# Original Dataset

胸悶	好的血液循環	動脈阻塞	心臟疾病
No	No	No	No
Yes	Yes	Yes	Yes
Yes	Yes	No	No
Yes	No	No	Yes
...	...	...	...

# Bootstrapped Dataset3

胸悶	好的血液循環	動脈阻塞	心臟疾病
Yes	Yes	Yes	Yes
Yes	Yes	No	No
Yes	No	No	Yes

## Bootstrapped Dataset1

胸悶	好的血液循環	動脈阻塞	心臟疾病
Yes	Yes	No	No
No	No	No	No
Yes	Yes	No	No
Yes	No	No	Yes
...	...	...	...

## Bootstrapped Dataset2

胸悶	好的血液循環	動脈阻塞	心臟疾病
Yes	No	No	Yes
Yes	No	No	Yes
No	No	No	No
Yes	No	No	Yes
...	...	...	...

## Bootstrapped Dataset3

胸悶	好的血液循環	動脈阻塞	心臟疾病
Yes	Yes	Yes	Yes
Yes	Yes	Yes	Yes
Yes	Yes	No	No
Yes	No	No	Yes
...	...	...	...

## Bootstrapped Dataset1

胸悶	好的血液循環	動脈阻塞	心臟疾病
Yes	Yes	No	No
No	No	No	No
Yes	Yes	No	No
Yes	No	No	Yes
...	...	...	...

Training

## Bootstrapped Dataset2

胸悶	好的血液循環	動脈阻塞	心臟疾病
Yes	No	No	Yes
Yes	No	No	Yes
No	No	No	No
Yes	No	No	Yes
...	...	...	...

## Bootstrapped Dataset3

胸悶	好的血液循環	動脈阻塞	心臟疾病
Yes	Yes	Yes	Yes
Yes	Yes	Yes	Yes
Yes	Yes	No	No
Yes	No	No	Yes
...	...	...	...

## Bootstrapped Dataset1

胸悶	好的血液循環	動脈阻塞	心臟疾病
Yes	Yes	No	No
No	No	No	No
Yes	Yes	No	No
Yes	No	No	Yes
...	...	...	...

Training → Classifier1

## Bootstrapped Dataset2

胸悶	好的血液循環	動脈阻塞	心臟疾病
Yes	No	No	Yes
Yes	No	No	Yes
No	No	No	No
Yes	No	No	Yes
...	...	...	...

## Bootstrapped Dataset3

胸悶	好的血液循環	動脈阻塞	心臟疾病
Yes	Yes	Yes	Yes
Yes	Yes	Yes	Yes
Yes	Yes	No	No
Yes	No	No	Yes
...	...	...	...

## Bootstrapped Dataset1

胸悶	好的血液循環	動脈阻塞	心臟疾病
Yes	Yes	No	No
No	No	No	No
Yes	Yes	No	No
Yes	No	No	Yes
...	...	...	...

Training → Classifier1

## Bootstrapped Dataset2

胸悶	好的血液循環	動脈阻塞	心臟疾病
Yes	No	No	Yes
Yes	No	No	Yes
No	No	No	No
Yes	No	No	Yes
...	...	...	...

Training →

## Bootstrapped Dataset3

胸悶	好的血液循環	動脈阻塞	心臟疾病
Yes	Yes	Yes	Yes
Yes	Yes	Yes	Yes
Yes	Yes	No	No
Yes	No	No	Yes
...	...	...	...

## Bootstrapped Dataset1

胸悶	好的血液循環	動脈阻塞	心臟疾病
Yes	Yes	No	No
No	No	No	No
Yes	Yes	No	No
Yes	No	No	Yes
...	...	...	...

Training → Classifier1

## Bootstrapped Dataset2

胸悶	好的血液循環	動脈阻塞	心臟疾病
Yes	No	No	Yes
Yes	No	No	Yes
No	No	No	No
Yes	No	No	Yes
...	...	...	...

Training → Classifier2

## Bootstrapped Dataset3

胸悶	好的血液循環	動脈阻塞	心臟疾病
Yes	Yes	Yes	Yes
Yes	Yes	Yes	Yes
Yes	Yes	No	No
Yes	No	No	Yes
...	...	...	...

## Bootstrapped Dataset1

胸悶	好的血液循環	動脈阻塞	心臟疾病
Yes	Yes	No	No
No	No	No	No
Yes	Yes	No	No
Yes	No	No	Yes
...	...	...	...

Training → Classifier1

## Bootstrapped Dataset2

胸悶	好的血液循環	動脈阻塞	心臟疾病
Yes	No	No	Yes
Yes	No	No	Yes
No	No	No	No
Yes	No	No	Yes
...	...	...	...

Training → Classifier2

## Bootstrapped Dataset3

胸悶	好的血液循環	動脈阻塞	心臟疾病
Yes	Yes	Yes	Yes
Yes	Yes	Yes	Yes
Yes	Yes	No	No
Yes	No	No	Yes
...	...	...	...

Training →

## Bootstrapped Dataset1

胸悶	好的血液循環	動脈阻塞	心臟疾病
Yes	Yes	No	No
No	No	No	No
Yes	Yes	No	No
Yes	No	No	Yes
...	...	...	...

Training → Classifier1

## Bootstrapped Dataset2

胸悶	好的血液循環	動脈阻塞	心臟疾病
Yes	No	No	Yes
Yes	No	No	Yes
No	No	No	No
Yes	No	No	Yes
...	...	...	...

Training → Classifier2

## Bootstrapped Dataset3

胸悶	好的血液循環	動脈阻塞	心臟疾病
Yes	Yes	Yes	Yes
Yes	Yes	Yes	Yes
Yes	Yes	No	No
Yes	No	No	Yes
...	...	...	...

Training → Classifier2

## Bootstrapped Dataset1

胸悶	好的血液循環	動脈阻塞	心臟疾病
Yes	Yes	No	No
No	No	No	No
Yes	Yes	No	No
Yes	No	No	Yes
...	...	...	...

## Bootstrapped Dataset2

胸悶	好的血液循環	動脈阻塞	心臟疾病
Yes	No	No	Yes
Yes	No	No	Yes
No	No	No	No
Yes	No	No	Yes
...	...	...	...

## Bootstrapped Dataset3

胸悶	好的血液循環	動脈阻塞	心臟疾病
Yes	Yes	Yes	Yes
Yes	Yes	Yes	Yes
Yes	Yes	No	No
Yes	No	No	Yes
...	...	...	...

Training → Classifier1

Training → Classifier2

Training → Classifier2

1

## Bootstrapped Dataset1

胸悶	好的血液循環	動脈阻塞	心臟疾病
Yes	Yes	No	No
No	No	No	No
Yes	Yes	No	No
Yes	No	No	Yes
...	...	...	...

## Bootstrapped Dataset2

胸悶	好的血液循環	動脈阻塞	心臟疾病
Yes	No	No	Yes
Yes	No	No	Yes
No	No	No	No
Yes	No	No	Yes
...	...	...	...

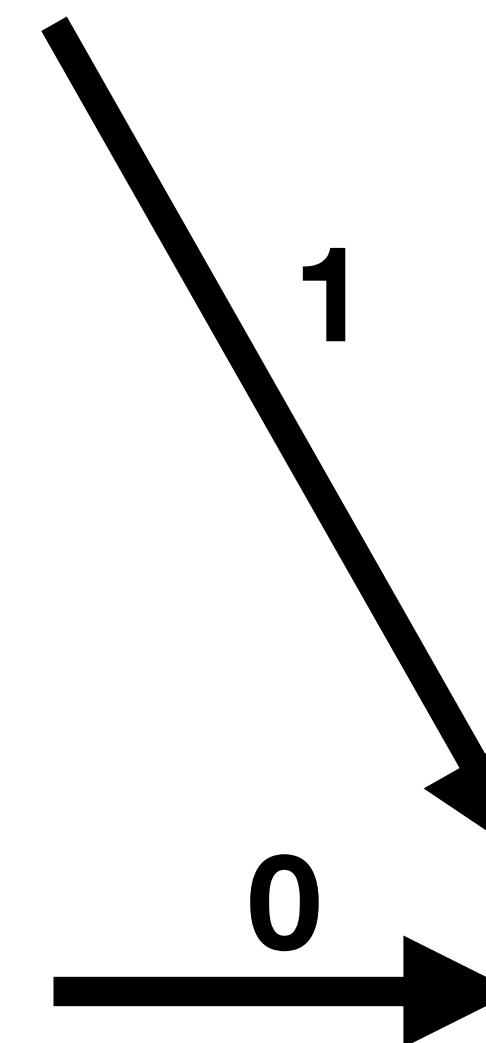
## Bootstrapped Dataset3

胸悶	好的血液循環	動脈阻塞	心臟疾病
Yes	Yes	Yes	Yes
Yes	Yes	Yes	Yes
Yes	Yes	No	No
Yes	No	No	Yes
...	...	...	...

Training → Classifier1

Training → Classifier2

Training → Classifier2



## Bootstrapped Dataset1

胸悶	好的血液循環	動脈阻塞	心臟疾病
Yes	Yes	No	No
No	No	No	No
Yes	Yes	No	No
Yes	No	No	Yes
...	...	...	...

## Bootstrapped Dataset2

胸悶	好的血液循環	動脈阻塞	心臟疾病
Yes	No	No	Yes
Yes	No	No	Yes
No	No	No	No
Yes	No	No	Yes
...	...	...	...

## Bootstrapped Dataset3

胸悶	好的血液循環	動脈阻塞	心臟疾病
Yes	Yes	Yes	Yes
Yes	Yes	Yes	Yes
Yes	Yes	No	No
Yes	No	No	Yes
...	...	...	...

Training → Classifier1

Training → Classifier2

Training → Classifier2

1

0

1

## Bootstrapped Dataset1

胸悶	好的血液循環	動脈阻塞	心臟疾病
Yes	Yes	No	No
No	No	No	No
Yes	Yes	No	No
Yes	No	No	Yes
...	...	...	...

## Bootstrapped Dataset2

胸悶	好的血液循環	動脈阻塞	心臟疾病
Yes	No	No	Yes
Yes	No	No	Yes
No	No	No	No
Yes	No	No	Yes
...	...	...	...

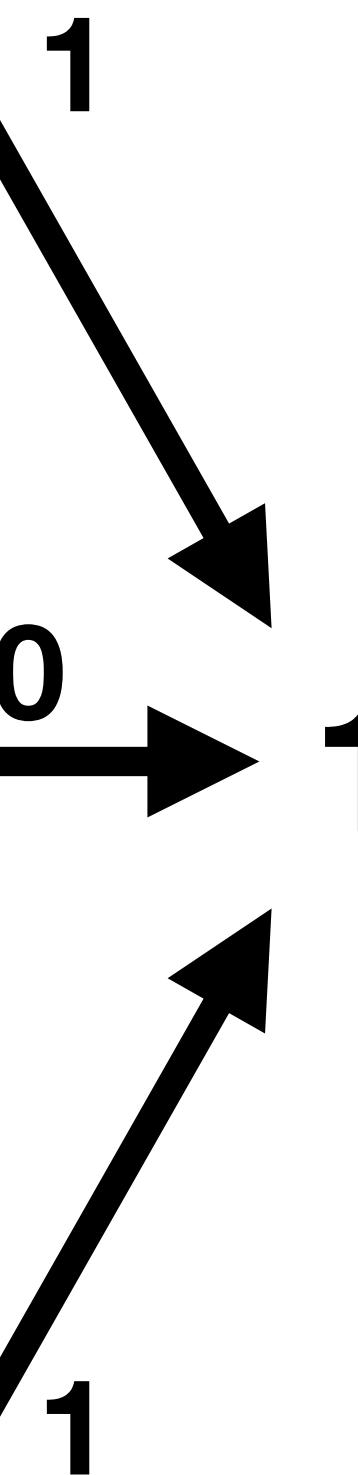
## Bootstrapped Dataset3

胸悶	好的血液循環	動脈阻塞	心臟疾病
Yes	Yes	Yes	Yes
Yes	Yes	Yes	Yes
Yes	Yes	No	No
Yes	No	No	Yes
...	...	...	...

Training → Classifier1

Training → Classifier2

Training → Classifier2



**Let's do this!**

# Ensemble Learning