

# Sway

Lobby

Rooms

Quizzes

Reports

Edit Quiz

Save and Exit

## Quiz name

Weekly quiz 8 (DUE: Sep 15 10am)

**1** | **MULTIPLE CHOICE**

Edit

In the anova method the splitting criteria is

### Answer choices

<b>A</b>	Gini
<b>B</b>	log-likelihood
<b>C</b>	SST – (SSL + SSR)
<b>D</b>	AIC



**2** | **MULTIPLE CHOICE**

Edit

The five ingredients for a regression tree are: (choose all that apply)

**Answer choices**

- |          |  |
|----------|--|
| <b>A</b> | A splitting criterion                      |
| <b>B</b> | A summary statistic or vector              |
| <b>C</b> | The error of a node.                       |
| <b>D</b> | The prediction error for a new observation |
| <b>E</b> | Any necessary initialization               |

**3** | **MULTIPLE CHOICE**

Edit

The rpart library is named from

**Answer choices**

- |          |                                    |
|----------|------------------------------------|
| <b>A</b> | Recursive PARTitioning             |
| <b>B</b> | Classification and regression tree |
| <b>C</b> | Regression and partition trees     |
| <b>D</b> | Random forests                     |

**4** | **MULTIPLE CHOICE**

Edit

How many terminal nodes in this tree?

node), split, n, deviance, yval

\* denotes terminal node

1) root 105 7118.26700 15.805220

2) Disp< 156 70 1491.62800 11.855870

4) Country=Brazil,Japan,Japan/USA,Korea,Mexico,USA 58 421.21470  
10.318470

8) Type=Small 21 50.30983 7.629048 \*

9) Type=Compact,Medium,Sporty,Van 37 132.80330 11.844890 \*

5) Country=France,Germany,Sweden 12 270.72330 19.286670 \*

3) Disp>=156 35 2351.19600 23.703910

6) HP.revs< 5550 24 980.27790 20.388710

12) Disp< 267.5 16 395.99670 17.820060 \*

13) Disp>=267.5 8 267.58000 25.526000 \*

7) HP.revs>=5550 11 531.63680 30.937090 \*

### Answer choices

<b>A</b>	3
<b>B</b>	4
<b>C</b>	5

D 6

## 5 | MULTIPLE CHOICE

Edit



In the summary of fit:

	CP	nsplit	rel error	xerror	xstd
1	0.460146	0	1.00000	1.01790	0.162343
2	0.117905	1	0.53985	0.74725	0.109078
3	0.112343	2	0.42195	0.74740	0.112607
4	0.044491	3	0.30961	0.53108	0.093947
5	0.033449	4	0.26511	0.50206	0.099952
6	0.010000	5	0.23166	0.47556	0.119522

What does the rel error correspond to?

### Answer choices

A	$1 - R^2$
B	PRESS statistic
C	$\Sigma (y_i - \bar{y})^2$
D	MSE

6

## MULTIPLE CHOICE

Edit



For this snippet of a tree

node), split, n, deviance, yval

\* denotes terminal node

3) Disp>=156 35 2351.19600 23.703910

What is the split?

**Answer choices**

<b>A</b>	Disp < 23.7203910
<b>B</b>	Disp > 2352.196
<b>C</b>	Disp < 35
<b>D</b>	Disp>=156

7

## MULTIPLE CHOICE

Edit



For this snippet of a tree

node), split, n, deviance, yval

\* denotes terminal node

3) Disp>=156 35 2351.19600 23.703910

what would be the predicted value?

**Answer choices**

<b>A</b>	156
<b>B</b>	35
<b>C</b>	2351.19600
<b>D</b>	23.703910

8

## MULTIPLE CHOICE

Edit

For this snippet of a tree

node), split, n, deviance, yval

\* denotes terminal node

3) Disp>=156 35 2351.19600 23.703910

How many cases are in this partition?

**Answer choices**

<b>A</b>	156
<b>B</b>	35
<b>C</b>	2351
<b>D</b>	23

9

## TRUE/FALSE

Edit

It is possible to overfit a data set, using a tree, by forcing the model to perfectly predict every response value. It would then not be a good model for future data sets.

**Answer**

True



## 10 | MULTIPLE CHOICE

Edit

Stopping rules for a regression tree include: (choose all that apply)

**Answer choices**

<b>A</b>	minsplit (number of cases in a subset to allow another split to happen)
<b>B</b>	minbucket (smallest number of cases allowed in a terminal node)
<b>C</b>	cp (smallest difference in statistic allowed for a split to happen)
<b>D</b>	MSE

[+ Multiple Choice](#)[+ True/False](#)[+ Short Answer](#)