Started on	Thursday, 16 December 2021, 12:05 PM
State	Finished
Completed on	Thursday, 16 December 2021, 12:41 PM
Time taken	35 mins 52 secs
Grade	**18.00** out of 25.00 (**72**%)

Information

You have 60 minutes to complete the quiz. You can skip questions, and you can "flag" questions to remember to return to them. If you wish to return to a question, please use the question navigation in the upper-right of the screen, rather than the "back" button on your browser.

This quiz is open-resource. That means that you may use: the book, your notes, the videos, your homework and my solutions, and activities and their solutions. You may also Google things to find different explanations of concepts beyond those given in the textbook.

You may NOT ask another human being for help, accept help from another human being, or post a question on the internet for others to solve. You may ask me for clarification questions, but you should not speak (verbally or virtually) to *anyone else* about the test. Violating these rules is a serious breach of academic integrity and will result in earning a 0% on the Test. In addition, the Dean for Curriculum and Academic Engagement will be notified; depending on the student's history of academic misconduct, this could lead to expulsion from the College.

A reminder of the Wooster Ethic:

I hereby join this community with a commitment to the Wooster Ethic upholding academic and personal integrity and a culture of honesty and trust in all my academic endeavors, social interactions, and official business of the College. I will submit only my own original work, and respect others and their property. I will not support by my actions or inactions the dishonest acts of others.

Question 1

Complete

Not graded

By typing my full name here, I agree to abide by the Wooster Ethic and the academic honesty policies described above.

Sarah Wright

Question 2

Correct

Mark 1.00 out of 1.00

We want to compare the amount of time spent studying for exams by first-years and seniors at CoW. Can we randomly assign students to these two groups?

Select one:

- a. No ✔
- b. Yes

Your answer is correct.

The correct answer is: No

Question 3

Correct

Mark 1.00 out of 1.00

Can changing diet reduce high blood pressure? Vegetarian diets and low-salt diets are both promising. Men with high blood pressure are assigned at random to four diets: (1) normal diet with unrestricted salt; (2) vegetarian with unrestricted salt; (3) normal with restricted salt; (4) vegetarian with restricted salt. After 8 weeks, their change in blood pressure was measured.

The response variable in this study is

Select one:

- a. the amount of salt in the subject's diet
- b. change in blood pressure after 8 weeks on the assigned diet ✔
- c. which of the four diets a subject is assigned to

Your answer is correct.

The correct answer is: change in blood pressure after 8 weeks on the assigned diet

Question 4

Incorrect

Mark 0.00 out of 1.00

Can changing diet reduce high blood pressure? Vegetarian diets and low-salt diets are both promising. Men with high blood pressure are assigned at random to four diets: (1) normal diet with unrestricted salt; (2) vegetarian with unrestricted salt; (3) normal with restricted salt; (4) vegetarian with restricted salt.

The researchers wonder if different genders will react differently to these diet changes. So in another version of this study, 500 subjects who identify as male or female were selected. First they were separated by gender, which resulted in 100 women and 400 men. Then, one-fourth of the women were randomly assigned to each diet; so 25 women in each diet. Similarly, one-fourth of the men were randomly assigned to each diet; so 100 men in each diet. This is a

Select one:

- a. block design, with two blocks
- b. matched pairs design ✖
- c. block design, with four blocks
- d. completely randomized design

Your answer is incorrect.

The correct answer is: block design, with two blocks

Question 5

Correct

Mark 1.00 out of 1.00

We want to compare the amount of time spent studying for exams by first-years and seniors at CoW. Can we randomly select students to be included in this study?

Select one:

- a. No
- b. Yes ✔

Your answer is correct.

The correct answer is: Yes

Question 6

Correct

Mark 1.00 out of 1.00

Researchers identified 242 children in the Cleveland area who had been born very prematurely (at about 29 weeks). They examined these children at age 8 and again at age 20, comparing them to another group of 233 children not born prematurely. Their report said that the "preemies" engaged in significantly less risky behavior than the others. Difference showed up in the use of alcohol and marijuana, conviction of crimes, and teenage pregnancy.

Select one:

- a. This is an observational study ✔
- b. This is an experiment
- c. There is not enough information to determine whether this is an observational study or an experiment.

The correct answer is: This is an observational study

Question 7

Correct

Mark 1.00 out of 1.00

In 1990, students and faculty planted 1000 white pine seedlings near the campus of Kenyon College. Their growth has been tracked over time. Two variables measured in 1995 were the *Cover* (categorized as 0=no thorny cover; 1=some thorny cover; 2=moderate levels of thorny cover; 3=lots of thorny cover) and the type of *Deer* damage (0=none; 1=browsed by deer).

We will use a chi-sq test for association (aka chi-sq test for independence) to investigate the relationship between these two variables (*Deer* and *Cover*).

Below is a 2-way table of the two variables of interest, along with the output from the appropriate test. What can we conclude?

```
> tally(Deer95~Cover95,data=Pines2)
      Cover95
Deer95   0   1   2   3
     0 134 138 161 155
     1  55  69  40  26
> chisq.test(tally(Deer95~Cover95,data=Pines2))

        Pearson's Chi-squared test

data:  tally(Deer95 ~ Cover95, data = Pines2)
X-squared = 23.293, df = 3, p-value = 3.509e-05
```

- a. There is **moderate** evidence of a relationship between deer damage and thorny cover.
- b. There is **strong** evidence of a relationship between deer damage and thorny cover. ✔
- c. There is **no** evidence of a relationship between deer damage and thorny cover.

Your answer is correct.

The correct answer is:
There is **strong** evidence of a relationship between deer damage and thorny cover.

Question 8

Complete

Mark 0.00 out of 1.00

In 1990, students and faculty planted 1000 white pine seedlings near the campus of Kenyon College; about half were fertilized (*Fert*=1) and half were not (*Fert*=0). Their growth has been tracked over time. In 1995, the type of *Deer* damage (0=none; 1=browsed by deer) was observed.

We will use a 2-sample proportion test to investigate the relationship between fertilization and the deer damage in 1995. Specifically, we wonder if fertilized trees were *less likely* to be browsed by deer. State the hypotheses below.

Ho:The probability of fertilized trees being browsed by deer = The probability of fertilized trees NOT being browsed by deer

Ha: The probability of fertilized trees being browsed by deer >(*less likely)* The probability of fertilized trees NOT being browsed by deer

Comment:

Question 9

Correct

Mark 1.00 out of 1.00

We have data on the passengers and crew of the British ship Titanic, which sank in 1912. We wish to assess if there is a relationship between Survival (Yes/No) and Class (1st class/2nd class/3rd class/Crew).

Could this relationship be investigated using a 2-sample proportion test?

a. yes

b. not enough information to answer

c. no ✔

Your answer is correct.

The correct answer is:
no

Question 10

Incorrect

Mark 0.00 out of 1.00

In 1990, students and faculty planted 1000 white pine seedlings near the campus of Kenyon College. Their growth has been tracked over time. Two variables measured in 1995 were the *Cover* (categorized as 0=no thorny cover; 1=some thorny cover; 2=moderate levels of thorny cover; 3=lots of thorny cover) and the type of *Deer*damage (0=none; 1=browsed by deer).

We will use a chi-sq test for association (aka chi-sq test for independence) to investigate the relationship between these two variables (*Deer* and *Cover*).

Below is a two-way table of the variables of interest. What percentage of the trees browsed by deer had moderate levels of thorny cover?

```
> tally(Deer95~Cover95,data=Pines2)
      Cover95
Deer95   0   1   2   3
     0 134 138 161 155
     1  55  69  40  26
```

- a. 5.14% ✖
- b. 19.90%
- c. 27.38%
- d. 80.10%
- e. 21.05%
- f. 20.69%

Your answer is incorrect.

The correct answer is:
21.05%

Question 11

Incorrect

Mark 0.00 out of 1.00

We have data on the passengers and crew of the British ship Titanic, which sank in 1912. We wish to assess if there is a relationship between Survival (Yes/No) and Class (1st class/2nd class/3rd class/Crew).

Could this relationship be investigated using a chi-sq test for association (also known as a chi-sq test for independence)?

- a. no
- b. not enough information to answer ✖
- c. yes

Your answer is incorrect.

The correct answer is:
yes

Question 12

Correct

Mark 1.00 out of 1.00

In an ANOVA model, we are interested in how between-group variation compares to within-group variation.

Select one:

- True ✔
- False

The correct answer is 'True'.

Question 13

Correct

Mark 1.00 out of 1.00

ANOVA can be thought of as:

- a. a generalization of the 2-sample t-test
- b. a special case of linear regression
- c. both (a) and (b) ✔
- d. neither (a) nor (b)

Your answer is correct.

The correct answer is:
both (a) and (b)

Question 14

Incorrect

Mark 0.00 out of 1.00

Suppose that you want to compare the ages of cars between the faculty, students, administrators, and staff at CoW. You take a random sample of 300 people on campus and ask how old their primary car is.

Is it appropriate to conduct ANOVA on this data? (check all that apply)

- [x] a. No, because the response variable is not quantitative. ✖
- [ ] b. No, because there are four groups to compare.
- [ ] c. No, because the sample sizes for the four groups will probably be different.
- [ ] d. Yes, assuming the conditions on residuals are met.

Your answer is incorrect.

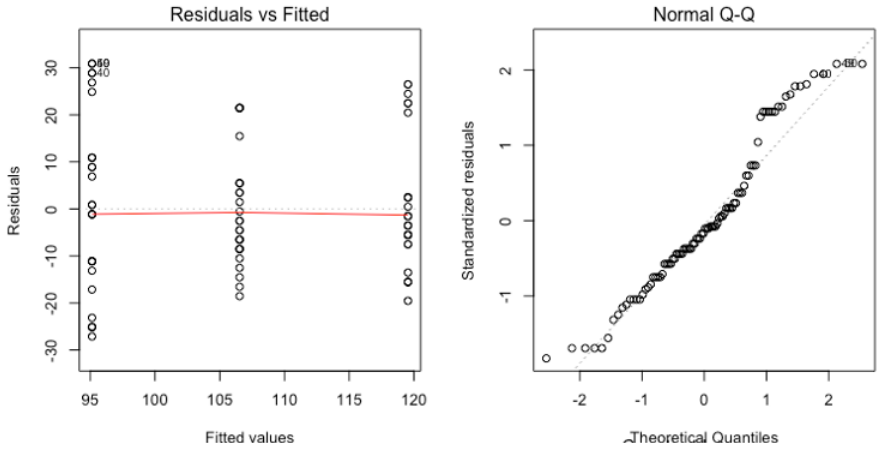
The correct answer is:
Yes, assuming the conditions on residuals are met.

Question 15

Incorrect

Mark 0.00 out of 1.00

A student, Shaka, conducted a study on college basketball players. He asked players to continually perform three different activities (1=layups, 2=free-throws, and 3=running drills) as quickly as they could. After performing each activity for five minutes, Shaka recorded the player's heart rate in beats per minute (bpm). He had five players, and he made each of them do all three activities on three different days. The days and order of activities were randomized for each player. Shaka uses a one-way ANOVA to see if heart rate depends on activity. Based on the residual plots below, do you think the conditions for the ANOVA model are met? (select all that apply)



- a. The condition of constant variance is definitely not met. ✖
- b. The condition of normal residuals is definitely not met. ✖
- c. All conditions appear to be reasonably met.
- d. These are not the right plots to assess the ANOVA conditions.

Your answer is incorrect.

The correct answer is:
All conditions appear to be reasonably met.

Question 16

Partially correct

Mark 0.50 out of 1.00

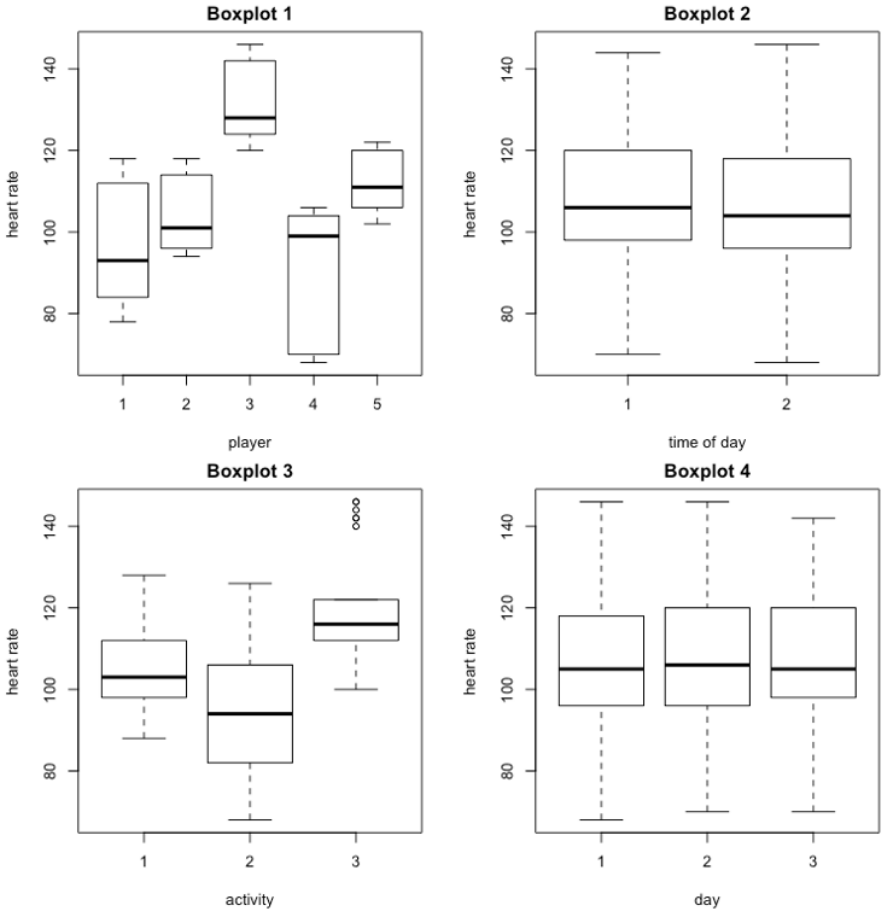
A student, Shaka, conducted a study on college basketball players. He asked players to continually perform three different activities (1=layups, 2=free-throws, and 3=running drills) as quickly as they could. After performing each activity for five minutes, Shaka recorded the player's heart rate in beats per minute (bpm). He had five players, and he made each of them do all three activities on three different days, and at two different times (morning and evening). The days and order of activities were randomized for each player. Shaka wants to see if heart rate depends on player, day, time of day, or activity. Does the exploratory data analysis below indicate that player, time, activity, or day might help explain a player's heart rate? (Choose all that apply.)

```
##   player min    Q1 median    Q3 max      mean        sd  n missing
## 1      1  78  84.0     93 112.0 118  96.66667 13.685544 18       0
## 2      2  94  96.0    101 114.0 118 104.11111  9.112944 18       0
## 3      3 120 124.5    128 141.5 146 131.33333  9.126561 18       0
## 4      4  68  70.5     99 104.0 106  91.11111 15.586151 18       0
## 5      5 102 106.0    111 119.5 122 112.11111  7.111622 18       0

##   time min Q1 median  Q3 max     mean       sd  n missing
## 1    1  70 98    106 120 144 107.2000 17.57271 45       0
## 2    2  68 96    104 118 146 106.9333 18.55998 45       0

##   activity min    Q1 median  Q3 max      mean       sd  n missing
## 1        1  88  98.0    103 112 128 106.53333 12.13639 30       0
## 2        2  68  82.5     94 106 126  95.13333 18.78872 30       0
## 3        3 100 112.5    116 122 146 119.53333 13.51049 30       0

##   day min   Q1 median    Q3 max     mean       sd  n missing
## 1   1  68 96.5    105 117.5 146 106.8667 18.72254 30       0
## 2   2  70 96.5    106 119.5 146 107.2667 18.19044 30       0
## 3   3  70 98.0    105 119.5 142 107.0667 17.59885 30       0
```



a. None of these variables appear likely to be a significant predictor of heart rate. ✖

b. *Day* is likely to be a significant predictor of heart rate.

c. *Time of Day* is likely to be a significant predictor of heart rate.

d. *Activity* is likely to be a significant predictor of heart rate.

e. *Player* is likely to be a significant predictor of heart rate. ✔

Your answer is partially correct.

You have correctly selected 1.
The correct answers are:
Player is likely to be a significant predictor of heart rate.,

Activity is likely to be a significant predictor of heart rate.

Question 17

Correct

Mark 1.00 out of 1.00

A time series is any data set where the condition of independence is violated.

Select one:

True

False ✔

The correct answer is 'False'.

Question 18

Correct

Mark 1.00 out of 1.00

Below is the seasonal means model with month as the categorical predictor for the gas production for the country of Australia, including data from January 1980 through August 1995 (Month 1 = January, Month 12 = December). What is the estimated mean gas production for January? (report to one decimal place)

```
##
## Call:
## lm(formula = recentgas ~ factor(cycle(recentgas)))
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -16249.0  -4857.2    818.1   4568.6  13804.3
##
## Coefficients:
##                             Estimate Std. Error t value Pr(>|t|)
## (Intercept)                  33631.1     1705.3  19.722  < 2e-16 ***
## factor(cycle(recentgas))2      949.8     2411.6   0.394 0.694166
## factor(cycle(recentgas))3     3991.4     2411.6   1.655 0.099690 .
## factor(cycle(recentgas))4     4716.7     2411.6   1.956 0.052067 .
## factor(cycle(recentgas))5    11988.0     2411.6   4.971 1.57e-06 ***
## factor(cycle(recentgas))6    15757.6     2411.6   6.534 6.66e-10 ***
## factor(cycle(recentgas))7    19164.7     2411.6   7.947 2.20e-13 ***
## factor(cycle(recentgas))8    17119.9     2411.6   7.099 2.97e-11 ***
## factor(cycle(recentgas))9    11360.2     2451.5   4.634 6.96e-06 ***
## factor(cycle(recentgas))10    8747.1     2451.5   3.568 0.000463 ***
## factor(cycle(recentgas))11    5773.1     2451.5   2.355 0.019625 *
## factor(cycle(recentgas))12    2044.8     2451.5   0.834 0.405345
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 6821 on 176 degrees of freedom
## Multiple R-squared:  0.479,  Adjusted R-squared:  0.4464
## F-statistic: 14.71 on 11 and 176 DF,  p-value: < 2.2e-16
```

Answer: 33631.1 ✔

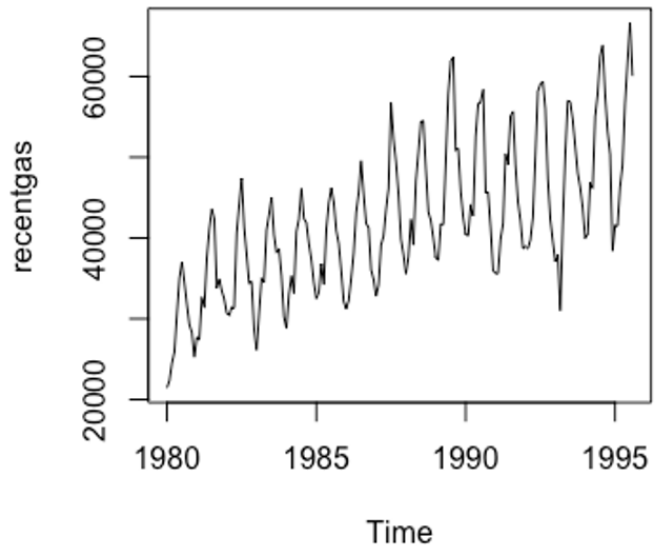
The correct answer is: 33631.1

Question 19

Correct

Mark 1.00 out of 1.00

The monthly gas production for the country of Australia from January 1980 through August 1995 is below. This plot exhibits **stationarity**.



Select one:

- True
- False ✔

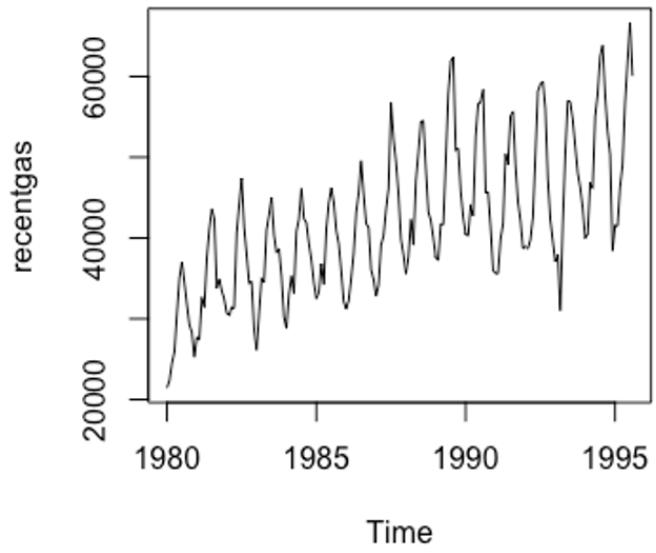
The correct answer is 'False'.

Question 20

Correct

Mark 1.00 out of 1.00

The monthly gas production for the country of Australia from January 1980 through August 1995 is below. Does this data show an overall increasing or decreasing trend?



a. No trend

b. Increasing trend ✔

c. Decreasing trend

Your answer is correct.

The correct answer is:
Increasing trend

Question 21

Complete

Mark 1.00 out of 1.00

We wish to model the gas production for the country of Australia, using data from January 1980 through August 1995. Model1 is the seasonal means model with month as the categorical predictors (Month 1 = January, Month 12 = December). Model2 is the "linear + seasonal means" model, where t = 1 for January 1980. Compare the two models. Based only on the information given, which would you prefer and why?

```
summary(model1)

##
## Call:
## lm(formula = recentgas ~ factor(cycle(recentgas)))
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -16249.0  -4857.2    818.1   4568.6  13804.3
##
## Coefficients:
##                             Estimate Std. Error t value Pr(>|t|)
## (Intercept)                  33631.1     1705.3  19.722  < 2e-16 ***
## factor(cycle(recentgas))2      949.8     2411.6   0.394 0.694166
## factor(cycle(recentgas))3     3991.4     2411.6   1.655 0.099690 .
## factor(cycle(recentgas))4     4716.7     2411.6   1.956 0.052067 .
## factor(cycle(recentgas))5    11988.0     2411.6   4.971 1.57e-06 ***
## factor(cycle(recentgas))6    15757.6     2411.6   6.534 6.66e-10 ***
## factor(cycle(recentgas))7    19164.7     2411.6   7.947 2.20e-13 ***
## factor(cycle(recentgas))8    17119.9     2411.6   7.099 2.97e-11 ***
## factor(cycle(recentgas))9    11360.2     2451.5   4.634 6.96e-06 ***
## factor(cycle(recentgas))10    8747.1     2451.5   3.568 0.000463 ***
## factor(cycle(recentgas))11    5773.1     2451.5   2.355 0.019625 *
## factor(cycle(recentgas))12    2044.8     2451.5   0.834 0.405345
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 6821 on 176 degrees of freedom
## Multiple R-squared:  0.479,  Adjusted R-squared:  0.4464
## F-statistic: 14.71 on 11 and 176 DF,  p-value: < 2.2e-16
```

```
summary(model2)

##
## Call:
## lm(formula = recentgas ~ t + factor(cycle(recentgas)))
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -13804.9  -1933.0    147.7   2038.7   9678.4
##
## Coefficients:
##                              Estimate Std. Error t value Pr(>|t|)
## (Intercept)                 23668.673    831.391  28.469  < 2e-16 ***
## t                             109.477      4.017  27.255  < 2e-16 ***
## factor(cycle(recentgas))2     840.336   1056.040   0.796 0.427260
## factor(cycle(recentgas))3    3772.421   1056.063   3.572 0.000457 ***
## factor(cycle(recentgas))4    4388.257   1056.101   4.155 5.08e-05 ***
## factor(cycle(recentgas))5   11550.093   1056.155  10.936  < 2e-16 ***
## factor(cycle(recentgas))6   15210.178   1056.223  14.401  < 2e-16 ***
## factor(cycle(recentgas))7   18507.827   1056.307  17.521  < 2e-16 ***
## factor(cycle(recentgas))8   16353.600   1056.407  15.480  < 2e-16 ***
## factor(cycle(recentgas))9   11141.251   1073.519  10.378  < 2e-16 ***
## factor(cycle(recentgas))10   8418.640   1073.556   7.842 4.19e-13 ***
## factor(cycle(recentgas))11   5335.230   1073.609   4.969 1.59e-06 ***
## factor(cycle(recentgas))12   1497.420   1073.677   1.395 0.164884
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2987 on 175 degrees of freedom
## Multiple R-squared:  0.9007, Adjusted R-squared:  0.8939
## F-statistic: 132.2 on 12 and 175 DF,  p-value: < 2.2e-16
```

Model 2 because the residuals are lower and the Fstatistict is higher.

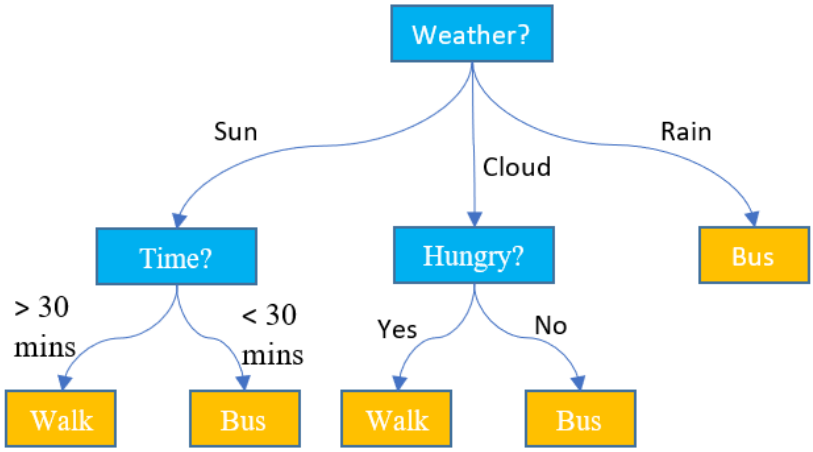
Comment:

Question 22

Correct

Mark 1.00 out of 1.00

Consider the simple decision tree below, which classifies whether Mae walks or takes the bus to work. Suppose that it is cloudy and Mae is hungry. Based on this tree, will she walk or take the bus to work?



a. Take the bus

b. Walk ✔

c. No way to know based on this information

Your answer is correct.

The correct answer is:
Walk

Question 23

Complete

Mark 1.00 out of 1.00

List two advantages of decision trees below.

They are easy to read and make predictions.

You don't have to worry about the type of response variable.

Comment:

Question 24

Correct

Mark 1.00 out of 1.00

Consider the simple decision tree below, which classifies whether Mae walks or takes the bus to work. How many explanatory variables (attributes) are used in this tree?

Previous activity

◀ Test-C-output

Jump to...

Next activity

Miller Expository approach ▶

Stay in touch

https://technology.wooster.edu

(330) 287-4357

helpdesk@wooster.edu

Data retention summary

Get the mobile app