General R Quiz

Solutions

Question 1

What does the R code *lm(length ~ age)* do?

SOLUTION: It runs a regression for the variable length in terms of the variable age, with lm standing for linear model. Running this code would produce the coefficients for the slope and the intercept.

Question 2

When changing colours in a graphical display of data in R what argument do you include in your command?

a. *colour=*
b. *choose.colour()*
c. *col=*
d. *color=*

SOLUTION: c.

Question 3

What is the quickest way to find the mean, median, minimum, maximum, first quartile and third quartile of a variable using R?

SOLUTION: *summary()*

Question 4

What is the R command for running a two sample t-test?

SOLUTION: t.test()

Question 5

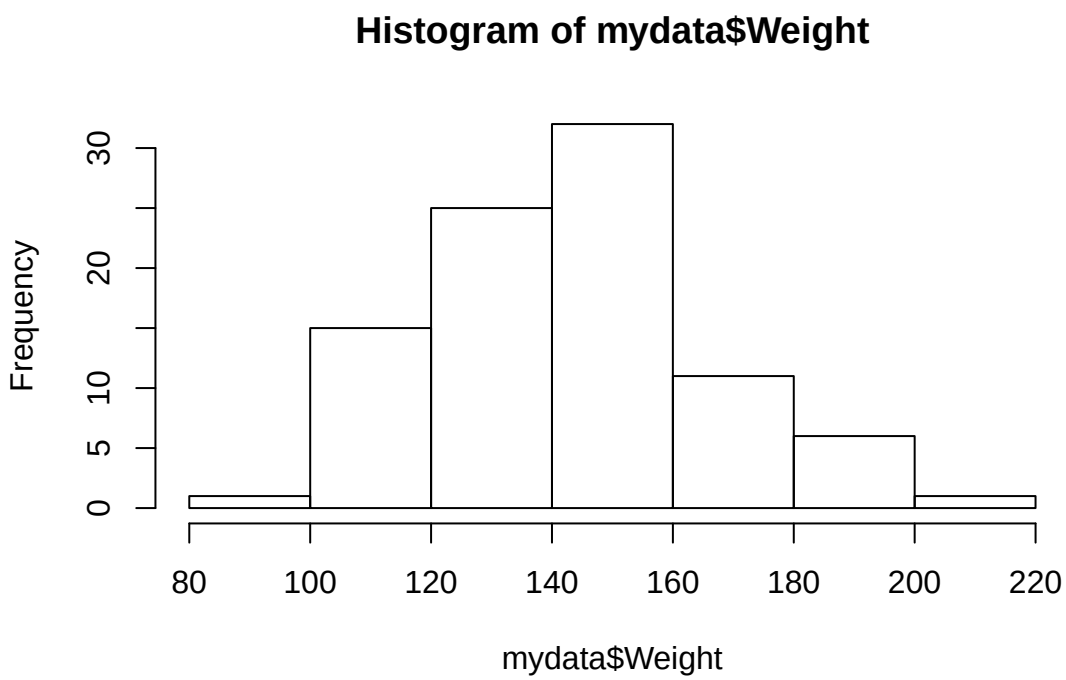
What colour are error messages in R?

SOLUTION: Red

Question 6

How would you change the following line of basic R code so that the below histogram becomes labelled "Histogram of Weights" and the x axis becomes labelled "Weights"?

```
hist(mydata$Weight)
```



SOLUTION: *hist(mydata$Weight, main="Histogram of Weights", xlab="Weight")*

Question 7

What does the R command *shapiro.test()* do?

SOLUTION: It runs the Shapiro Wilk normality test and can be used to check if a sample was drawn from an underlying distribution that is normally distributed.

Question 8

How would you change the size of the y-axis on a plot so it ranges from 0 to 40?

a. *yaxis=c(0,40)*
b. *ylim=(0,40)*
c. *yaxis=(0,40)*
d. *ylab=c(0,40)*
e. *ylim=c(0,40)*

SOLUTION: e.

Question 9

What is wrong with this line of R code?

```
mydata <- read.table(file.choose() sep=",", header=TRUE)
```

SOLUTION: This line is missing a comma (,) after the file.choose() function.

Question 10

What code would you use to find the p-value of a correlation?

a. *cor()*
b. *cor.test()*
c. *correlation()*
d. *p.cor()*

SOLUTION: b.

Question 11

What is wrong with this line of R code?

```
plot(dim ~ time, data=wound2, xlab=Wound Dimension, ylab=Healing Time)
```

SOLUTION: There should be "" around the labels for the different axes. The correct R code would be *plot(dim ~ time, data=wound2, xlab="Wound Dimension", ylab="Healing Time").*

Question 12

When exporting a plot from R, what three options do you have?

SOLUTION: Save as Image, Save as PDF and Copy to Clipboard

Question 13

What does the blue and brown magnifying glass icon at top of the script editor let you do?

a. Search the R help pages
b. Enlarge the text in an R script
c. Find and replace sections of code in an R script
d. Spell-check your code in an R script

SOLUTION: c.

Question 14

You have opened a data set in R but when you view your data frame it looks like this:

```
##       V1       V2     V3     V4     V5    V6
## 1 Gender Activity Smokes Height Weight Pulse
## 2   Male Moderate     No     66    140    64
## 3   Male Moderate     No     72    145    58
## 4   Male    A lot    Yes   73.5    160    62
## 5   Male   Slight    Yes     73    190    66
## 6   Male Moderate     No     69    155    64
```

What is the issue and how do you correct it?

SOLUTION: The issue is that the variable/column names are included as observations. To fix this, you add the argument *header=TRUE* into your command when opening the file.

Question 15

When creating a scatter plot with R, what does the argument *phc=* allow you to do?

a. Change the header of the plot
b. Change the size of one of the axis
c. State which variable you are looking at
d. Change the symbols used to represent the observations

SOLUTION: d.

Question 16

Name three things you should be looking for in the output produced when you use the *summary()* command on a regression model.

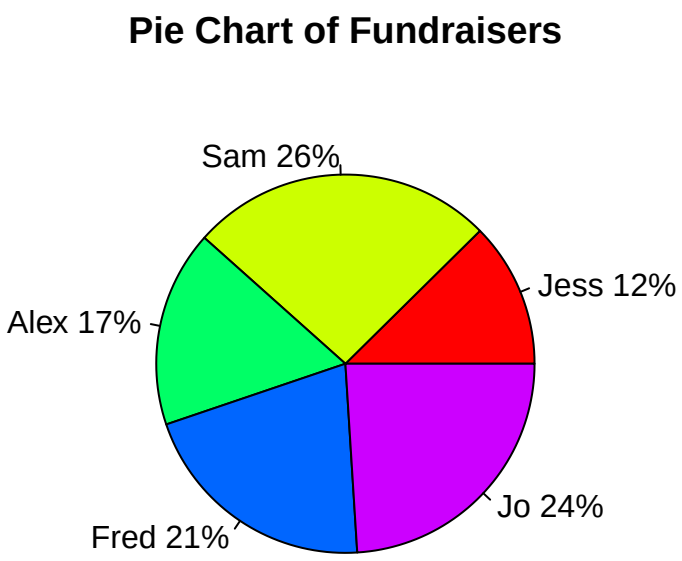
SOLUTION: The coefficients (for the slope and intercept), p-values (or the stars, as they also indicate significance level), and R-squared or adjusted R-squared.

Question 17

Data was collected on the amount of money raised by five different fundraisers:

Name	Amount
Jess	£62
Sam	£130
Alex	£84
Fred	£104
Jo	£120

What code was used to produce the following pie chart from the data above?



SOLUTION:

```
slices <- c(62, 130,84, 104,120 )
lbls <- c("Jess", "Sam", "Alex", "Fred", "Jo")
pct <- round(slices/sum(slices)*100)
lbls <- paste(lbls, pct)
lbls <- paste(lbls,"%",sep="")
pie(slices,labels = lbls, col=rainbow(length(lbls)),main="Pie Chart of Fundraisers")
```

Question 18

Using two variables called **dim** and **time** (which contain the dimension and healing time of wounds), what code would be needed to produce the following output?

Note: You can assume that the data set has been attached.

```
##
##  Pearson's product-moment correlation
##
## data:  dim and time
## t = 10.017, df = 34, p-value = 1.116e-11
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  0.7481397 0.9290068
## sample estimates:
##       cor
## 0.8642507
```

SOLUTION: *cor.test(dim,time)*

Question 19

If you were working with a data set called **mydata** and wanted to look at the variable **days**, how would you tell this to R without attaching the data set?

a. *mydata&days*
b. *days$mydata*
c. *mydata$days*
d. *mydata£days*

SOLUTION: c.

Question 20

How would you fix this line of R code?

```
setwd("C:\Users\Documents\R")
```

SOLUTION: You need to change the direction of the slashes, so that the code looks like this: *setwd("C:/Users/Documents/R").*