

Solution to Quiz 1

Question 1

Q: The function `ggplot()` requires that the value supplied to the ‘data’ argument be a

A: data frame

Question 2

Q: You run the following code and get the error printed directly after the code:

```
library(ggplot2)
library(faraway)
data(nepali)
head(nepali, 3)

## id sex wt ht mage lit died alive age
## 1 120011 1 12.8 91.2 35 0 2 5 41
## 2 120011 1 12.8 93.9 35 0 2 5 45
## 3 120011 1 13.1 95.2 35 0 2 5

ggplot(nepali, aes(x = ht, y = wt))
+ geom_point()

## Error in +geom_point(): invalid argument to unary operator
```

What happened?

A: Because you put the `+` at the start of the second line, instead of at the end of the first line, R thought the call was over and tried to run the first line by itself, and then tried to run the second line by itself, without an associated `ggplot` object.

Question 3

Q: You have a dataset of observations on study subjects which has height and weight for each subject measured at multiple time points. You want to create a scatterplot of height (x-axis) by weight (y-axis). Since there are multiple measurements per subject, you would like to plot the measurements for each subject in a separate color. There is a column in the data called `id` that gives the unique id of each study subject; this column currently has the class “numeric”. Which of the following choices is a good strategy for creating a plot where data points use color to identify the subject?

A: Convert the `id` column of the data frame to a factor, so a discrete color palette will be used, then use `ggplot2` to create a scatterplot and specify `color = id` in the `aes()` section of the `ggplot()` call.

Example code:

```
df %>%
  mutate(id = factor(id)) %>%
  ggplot(aes(x = height, y = weight, color = id)) +
  geom_point()
```

Question 4

Q: You want to save a figure from R as a pdf file named “MyFig.pdf”. Which of the following describes the proper steps to take?

A: First run `pdf("MyFig.pdf")` to open a pdf device, then run all your code to create the figure, then use `dev.off()` to close the device.

Example code:

```
pdf("MyFig.pdf")
ggplot(df, aes(x = x, y = y)) +
  geom_point()
dev.off()
```

Question 5

Q: Consider the following dataset.

```
library(lattice)
data(barley)
head(barley)
```

	yield	variety	year	site
## 1	27.00000	Manchuria	1931	University Farm
## 2	48.86667	Manchuria	1931	Waseca
## 3	27.43334	Manchuria	1931	Morris

Which ggplot2 expression would make a scatterplot of year and yield while coloring each point according to its variety?

A: `ggplot(data = barley, aes(year, yield)) + geom_point(aes(color = variety))`

Question 6

Q: Consider the following dataset.

```
data(trees)
head(trees)
```

	Girth	Height	Volume
##	8.3	70	10.3
##	8.6	65	10.3
##	8.8	63	10.2
##	10.5	72	16.4
##	10.7	81	18.8
##	10.8	83	19.7

You want to create a scatter plot showing **Girth** on the x-axis, **Height** on the y-axis, and **Volume** with point size. Which of the following code would create that plot?

A: (1)

```
trees %>%  
  ggplot(aes(x = Girth, y = Height)) +  
  geom_point(aes(size = Volume))
```

(2)

```
trees %>%  
  ggplot(aes(x = Girth, y = Height, size = Volume)) +  
  geom_point()
```

(3)

```
ggplot() +  
  geom_point(data = trees,  
            aes(x = Girth, y = Height, size = Volume))
```

(6)

```
trees %>%  
  ggplot() +  
  geom_point(aes(x = Girth, y = Height, size = Volume))
```

(8)

```
ggplot(trees, aes(x = Girth, y = Height, size = Volume)) +  
  geom_point()
```

Question 7

Q: How can you create small multiples using ggplot2?

A: (1) Add `facet_wrap` to the geom, specifying the formula to use for faceting

(2) Add `facet_grid` to the geom, specifying one or two columns to use for faceting

Question 8

Q: What does `geom_smooth(method = "lm")` do when it is added to a ggplot object?

A: It adds a fitted linear regression line to the plot

Question 9

Q: When plotting multiple panels using faceting for a single variable, what controls the order in which the panels are plotted?

A: Panels are ordered by the alphabetical order or by the order of the levels of the faceting variable

Question 10

Q: What function would you use to limit the range of the x-axis in a scatterplot?

A: `scale_x_continuous()`