**Assessment Part 2: Reshaping Data**

Use the following libraries for these questions:

library(tidyverse)  
library(dslabs)

**Question 9**

1/1 point (graded)

Examine the built-in dataset co2. This dataset comes with base R, not dslabs - just type co2 to access the dataset.

Is co2 tidy? Why or why not?

co2 is tidy data: it has one year for each row.

co2 is tidy data: each column is a different month.

co2 is not tidy: there are multiple observations per column.

co2 is not tidy: to be tidy we would have to wrangle it to have three columns (year, month and value), and then each co2 observation would have a row.

correct

You have used 1 of 2 attempts Some problems have options such as save, reset, hints, or show answer. These options follow the Submit button.

**Question 10**

1/1 point (graded)

Run the following command to define the co2\_wide object:

co2\_wide <- data.frame(matrix(co2, ncol = 12, byrow = TRUE)) %>%

setNames(1:12) %>%

mutate(year = as.character(1959:1997))

Use the gather function to make this dataset tidy. Call the column with the CO2 measurements co2 and call the month column month. Name the resulting object co2\_tidy.

Which code would return the correct tidy format?

co2\_tidy <- gather(co2\_wide,month,co2,year)

co2\_tidy <- gather(co2\_wide,co2,month,-year)

co2\_tidy <- gather(co2\_wide,co2,month,year)

co2\_tidy <- gather(co2\_wide,month,co2,-year)

correct

You have used 1 of 2 attempts Some problems have options such as save, reset, hints, or show answer. These options follow the Submit button.

**Question 11**

1/1 point (graded)

Use co2\_tidy to plot CO2 versus month with a different curve for each year:

co2\_tidy %>% ggplot(aes(as.numeric(month), co2, color = year)) + geom\_line()

What can be concluded from this plot?

CO2 concentrations increased monotonically (never decreased) from 1959 to 1997.

CO2 concentrations are highest around May and the yearly average increased from 1959 to 1997.

CO2 concentrations are highest around October and the yearly average increased from 1959 to 1997.

Yearly average CO2 concentrations have remained constant over time.

CO2 concentrations do not have a seasonal trend.

correct

You have used 1 of 2 attempts Some problems have options such as save, reset, hints, or show answer. These options follow the Submit button.

**Question 12**

1/1 point (graded)

Load the built-in admissions dataset, which contains college admission information for men and women across six majors, and remove the applicants percentage column:

library(dslabs)

data(admissions)

dat <- admissions %>% select(-applicants)

Your goal is to get the data in the shape that has one row for each major, like this:

major men women

A 62 82

B 63 68

C 37 34

D 33 35

E 28 24

F 6 7

Which command could help you to wrangle the data into the desired format?

dat\_tidy <- spread(dat, major, admitted)

dat\_tidy <- spread(dat, gender, major)

dat\_tidy <- spread(dat, gender, admitted)

dat\_tidy <- spread(dat, admitted, gender)

correct

You have used 1 of 2 attempts Some problems have options such as save, reset, hints, or show answer. These options follow the Submit button.

**Question 13**

1/1 point (graded)

Now use the admissions dataset to create the object tmp, which has columns major, gender, key and value:

tmp <- gather(admissions, key, value, admitted:applicants)

tmp

Combine the key and gender and create a new column called column\_name to get a variable with the following values: admitted\_men, admitted\_women, applicants\_men and applicants\_women. Save the new data as tmp2.

Which command could help you to wrangle the data into the desired format?

tmp2 <- spread(tmp, column\_name, key, gender)

tmp2 <- gather(tmp, column\_name, c(gender,key))

tmp2 <- unite(tmp, column\_name, c(gender, key))

tmp2 <- spread(tmp, column\_name, c(key,gender))

tmp2 <- unite(tmp, column\_name, c(key, gender))

correct

You have used 2 of 2 attempts Some problems have options such as save, reset, hints, or show answer. These options follow the Submit button.

**Question 14**

1/1 point (graded)

Which function can reshape tmp2 to a table with six rows and five columns named major, admitted\_men, admitted\_women, applicants\_men and applicants\_women?

gather

spread

separate

unite

correct

You have used 1 of 2 attempts