In this part of the assessment, you will answer several multiple choice questions that review the concepts of string processing. You can answer these questions without using R, although you may find it helpful to experiment with commands in your console.

In the second part of the assessment on the next page, you will import a real dataset and use string processing to clean it for analysis. This will require you to write code in R.

Want even more practice with regular expressions? Complete the lessons and exercises in the RegexOne ⬀ online interactive tutorial!

Question 2

1/1 point (graded)
You have the following table, `schedule` :

```
>schedule
day            staff
Monday         Mandy, Chris and Laura
Tuesday        Steve, Ruth and Frank
```

You want to turn this into a more useful data frame.

Which two commands would properly split the text in the "staff" column into each individual name? Select ALL that apply.

- [ ] `str_split(schedule$staff, ",|and")`
- [x] `str_split(schedule$staff, ", | and ")`
- [x] `str_split(schedule$staff, ",\\s|\\sand\\s")`
- [ ] `str_split(schedule$staff, "\\s?(,|and)\\s?")`

✔

Answer
Correct:
This regex will correctly split each "staff" string into three names by properly accounting for the space after the comma as well as the spaces before and after the "and", but it's not the only one.
This regex command is the same as the one above, except that the spaces are written as \\s, but it's not the only one.

Submit   You have used 1 of 2 attempts

ⓘ Answers are displayed within the problem

Question 3

1/1 point (graded)
You have the following table, `schedule` :

```
> schedule
day        staff
Monday          Mandy, Chris and Laura
Tuesday         Steve, Ruth and Frank
```

What code would successfully turn your "Schedule" table into the following tidy table?

```
> tidy
day      staff
<chr>    <chr>
Monday   Mandy
Monday   Chris
Monday   Laura
Tuesday Steve
Tuesday Ruth
Tuesday Frank
```

○ (selected)
```
        tidy <- schedule %>%
mutate(staff = str_split(staff, ", | and ")) %>%
unnest()
```

○
```
        tidy <- separate(schedule, staff, into = c("s1","s2","s3"), sep = ",") %>%
gather(key = s, value = staff, s1:s3)
```

○
```
        tidy <- schedule %>%
mutate(staff = str_split(staff, ", | and ", simplify = TRUE)) %>%
unnest()
```

✔

Answer
Correct:
The mutate command creates a column "staff". Each row in the "staff" column is a character vector of length three, with the names of each staff member. We unnest this character vector using the unnest() function from tidyr.

[Submit]   You have used 1 of 2 attempts

ℹ Answers are displayed within the problem

Question 4

1/1 point (graded)
Using the gapminder data, you want to recode countries longer than 12 letters in the region "Middle Africa" to their abbreviations in a new column, "country_short". Which code would accomplish this?

○

```
        dat <- gapminder %>% filter(region == "Middle Africa") %>%
    mutate(recode(country,
            "Central African Republic" = "CAR",
            "Congo, Dem. Rep." = "DRC",
            "Equatorial Guinea" = "Eq. Guinea"))
```

○

```
        dat <- gapminder %>% filter(region == "Middle Africa") %>%
    mutate(country_short = recode(country,
                        c("Central African Republic", "Congo, Dem. Rep.", "Equatorial Guinea"),
                        c("CAR", "DRC", "Eq. Guinea")))
```

○

```
        dat <- gapminder %>% filter(region == "Middle Africa") %>%
    mutate(country = recode(country,
                    "Central African Republic" = "CAR",
                    "Congo, Dem. Rep." = "DRC",
                    "Equatorial Guinea" = "Eq. Guinea"))
```

●

```
        dat <- gapminder %>% filter(region == "Middle Africa") %>%
    mutate(country_short = recode(country,
                        "Central African Republic" = "CAR",
                        "Congo, Dem. Rep." = "DRC",
                        "Equatorial Guinea" = "Eq. Guinea"))
```

✔

Answer
Correct: This code properly recodes each country in a new column "country_short".

| Submit |

You have used 1 of 2 attempts

ⓘ Answers are displayed within the problem