

<u>Course</u> > <u>Section 3: String Pr...</u> > <u>3.3: String Processi</u>... > Assessment Part 2: ...

## **Assessment Part 2: String Processing Part 3**

Import raw Brexit referendum polling data from Wikipedia:

```
library(rvest)
library(tidyverse)
library(stringr)
url <- "https://en.wikipedia.org/w/index.php?
title=Opinion_polling_for_the_United_Kingdom_European_Union_membership_referendum&oldid=896735054"
tab <- read_html(url) %>% html_nodes("table")
polls <- tab[[7]] %>% html_table(fill = TRUE)
```

You will use a variety of string processing techniques learned in this section to reformat these data.

## Question 5

1/1 point (graded)

Some rows in this table do not contain polls. You can identify these by the lack of the percent sign (%) in the Remain column.

```
Update polls by changing the column names to c("dates", "remain", "leave", "undecided", "lead", "samplesize", "pollster", "poll_type", "notes") and only keeping rows that have a percent sign (%) in the remain column.
```

How many rows remain in the polls data frame?

129 **✓ Answer:** 129

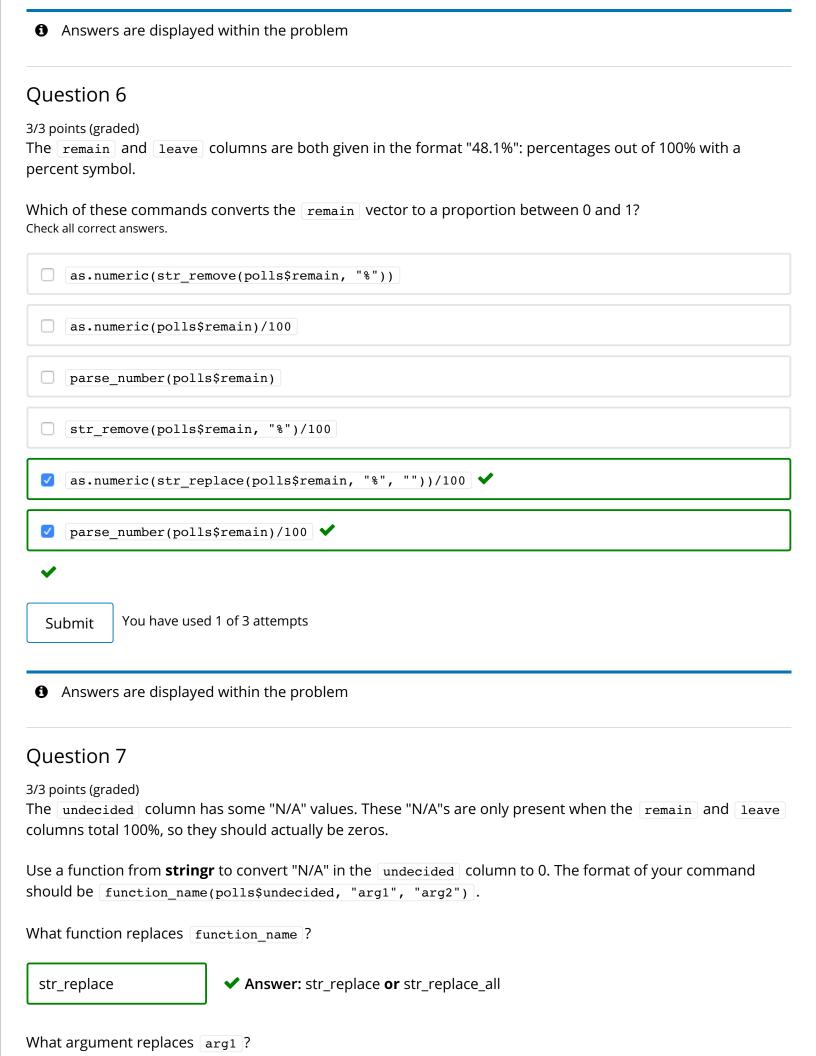
## 129

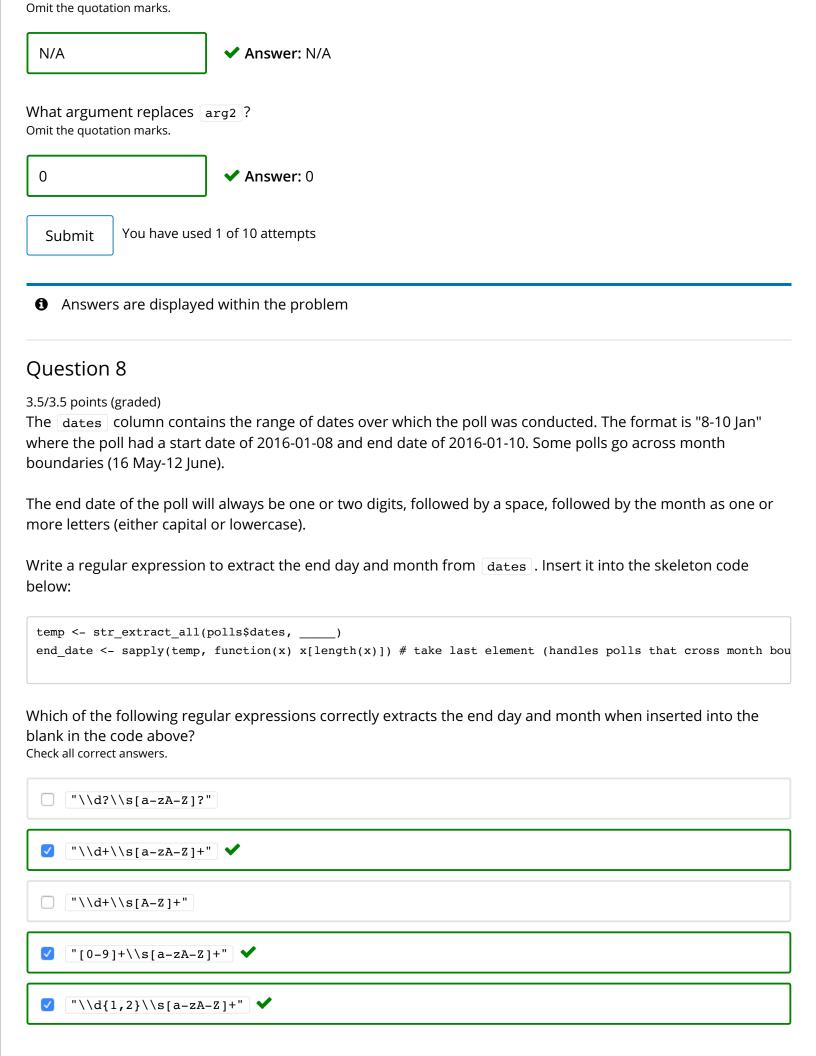
## Answer code

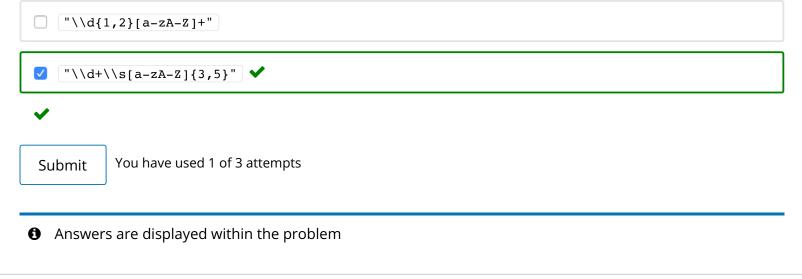
```
names(polls) <- c("dates", "remain", "leave", "undecided", "lead", "samplesize", "pollster", "poll_type",
polls <- polls[str_detect(polls$remain, "%"), -9]
nrow(polls)</pre>
```

Submit

You have used 1 of 10 attempts







© All Rights Reserved