Comprehension Check due Jun 6, 2021 06:59 +03

This assessment reviews several concepts about dates, times, and text mining. In part 1 on this page, you will practice extracting and manipulating dates in real datasets. In part 2 on the next page, you will walk through a sentiment analysis of a novel using steps covered in the previous section.

Use the following libraries and options for coding questions:

```
library(dslabs)
library(lubridate)
options(digits = 3)    # 3 significant digits
```

IMPORTANT: Some of these exercises use **dslabs** datasets that were added in a July 2019 update. Make sure your package is up to date with the command `install.packages("dslabs")`.

Question 1

1/1 point (graded)
Which of the following is the standard ISO 8601 format for dates?

○ MM-DD-YY

⦿ YYYY-MM-DD

○ YYYYMMDD

○ YY-MM-DD

✔

Answer
Correct: This is proper ISO 8601 formatting for dates.

Explanation
YYYY-MM-DD (year, month, day) is proper ISO 8601 formatting for dates.

ℹ️ Answers are displayed within the problem

Question 2

1/1 point (graded)

Which of the following commands could convert this string into the correct date format?

```
dates <- c("09-01-02", "01-12-07", "02-03-04")
```

○
```
ymd(dates)
```

○
```
mdy(dates)
```

○
```
dmy(dates)
```

🔵 It is impossible to know which format is correct without additional information.

✔️

Answer
Correct:
The formatting of these dates is ambiguous. They could be formatted as ymd, mdy or dmy. We need more information about our data to be able to select the correct command.

ⓘ Answers are displayed within the problem

Question 3

2/2 points (graded)
Load the `brexit_polls` data frame from **dslabs**:

```
data(brexit_polls)
```

How many polls had a start date (`startdate`) in April (month number 4)?

25

✔ **Answer:** 25

25

Answer code

```
sum(month(brexit_polls$startdate) == 4)
```

Use the `round_date()` function on the `enddate` column with the argument `unit="week"` . How many polls ended the week of 2016-06-12?
Read the documentation to learn more about `round_date()` .

13

✔ **Answer:** 13

13

Answer code

```
sum(round_date(brexit_polls$enddate, unit = "week") == "2016-06-12")
```

Submit

You have used 1 of 10 attempts

Question 4

1/1 point (graded)

Use the `weekdays()` function from **lubridate** to determine the weekday on which each poll ended (`enddate`).

On which weekday did the greatest number of polls end?

○ Monday

○ Tuesday

○ Wednesday

○ Thursday

○ Friday

○ Saturday

◉ Sunday

✔

Explanation

See the distribution of weekday end dates with `table(weekdays(brexit_polls$enddate))` .

| Submit | You have used 1 of 2 attempts |

Question 5

2/2 points (graded)

Load the `movielens` data frame from **dslabs**.

```
data(movielens)
```

This data frame contains a set of about 100,000 movie reviews. The `timestamp` column contains the review date as the number of seconds since 1970-01-01 (epoch time).

Convert the `timestamp` column to dates using the **lubridate** `as_datetime()` function.

Which year had the most movie reviews?

| 2000 | ✔ **Answer:** 2000 |

2000

Answer code

```
        dates <- as_datetime(movielens$timestamp)
 reviews_by_year <- table(year(dates))     # count reviews by year
 names(which.max(reviews_by_year))    # name of year with most reviews
```

Which hour of the day had the most movie reviews?

| 20 | ✔ **Answer:** 20 |

20

Answer code

```
        reviews_by_hour <- table(hour(dates))    # count reviews by hour
 names(which.max(reviews_by_hour))    # name of hour with most reviews
```

| Submit | You have used 1 of 10 attempts |