**Assessment Part 1: Dates, Times, and Text Mining**

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 update.packages("dslabs"). You can also update all packages on your system by running update.packages() with no arguments, and you should consider doing this routinely.

**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

correct

Answer

Correct:

This is proper ISO 8601 formatting for dates.

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 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.

correct

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.

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 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)? correct

25 Loading

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.

correct

13 Loading

You have used 1 of 10 attempts Some problems have options such as save, reset, hints, or show answer. These options follow the Submit button.

**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

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 5**

2.0/2.0 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? correct

2000 Loading

Which hour of the day had the most movie reviews? correct

20 Loading

You have used 1 of 10 attempts