

Check-in 5: Answers

Insert answers within the code chunks. Unless specified otherwise, do not assign your output to an object. If directed to assign output to an object, wrap the pipe in parentheses to print the output.

Problem 1: (4 pts) For the first part of this question, find the total number of letters in all of the names of the months. To do this, use `{stringr}` functions to first combine all of the month names into a single string (JanuaryFebruaryMarchAprilMayJuneJulyAugustSeptemberOctoberNovemberDecember), then output the number of characters in that string.

For the second part of this question, use `{stringr}` functions to extract the first three letters of each month name then combine them into a single string with `-` characters in between each month: Jan-Feb-Mar-Apr-May-Jun-Jul-Aug-Sep-Oct-Nov-Dec

Problem 2: (6 pts) Inside the code chunk below, write a single command (using pipes) that completes the following and assigns the output to `dog_data`:

1. Import `stevens_etal_2020_obed_data1.csv` from the following URL:
`https://decisionslab.unl.edu/data/stevens_etal_2020_obed_data1.csv`
2. Include the following variables in this order: `class`, `dog_age`, and columns that include the text `personality_`.
3. In the `dog_age` column, replace the text "< 1 year old" with "0.5".
4. In the `dog_age` column, replace the text " years old" and " year old" with no text.
5. Convert the `dog_age` column to a numeric data type.
6. Give a `glimpse()` of the data.

Problem 3: (5 pts) We want to extract semester information from the `class` column. Write a single command (using pipes) that does the following to the `dog_data` object without assigning to a new or old object.

1. Using a `{stringr}` function, create a new column called `semester` that includes only the first character in `class`.
2. Using a `{forcats}` function, replace the letters F, S, and U in `semester` with Fall, Spring, and Summer.
3. Change the levels of the `semester` factor to be in the order Summer, Fall, Spring
4. Move the `semester` column to be after `class`.
5. Output the structure of the data frame.

Problem 4: (3 pts) There are three dates listed below. Convert each of the defined objects into proper ISO 8601 standard format as date data types. Then extract the minute from `pearlharbor`, the year from `dday`, and the weekday (in full words) from `patriotsday`.

```
pearlharbor <- "7 Dec 41 7:48"
dday <- "June 6, 1944"
patriotsday <- "9/11/01"
```

Problem 5: (4 pts) Do each of the following:

- Determine the proportion of NAs in the `dog_age` column.
- Create a new column in `dog_data` called `age_group` that uses the `cut()` function to label dogs 1 and under as “Puppy” and dogs over 1 as “Adult”.
- For each class, calculate the mean rating for each of the personality measures, rounding them to one decimal place.

Problem 6: (3 pts) In your own creative way, I would like you to **extend** beyond what the previous questions have asked you to do. So build a pipeline with at least three steps that does at least one (preferably all) of the following (using this data set or another one):

1. Uses a function from this module
2. Uses a function from a previous module
3. Uses a function that we haven’t covered in the course