Explore data and R
Cleaning data
Take a closer look at the data
Weekly challenge 3

Reading: ReadingGlossary: Terms and definitions

. Duration: 5 minutes5 min

Purchase a subscription to unlock this item.

Quiz: Weekly challenge 3

10 questions

Weekly challenge 3

Quiz50 minutes • 50 min

Review Learning Objectives

Submit your assignment

Due May 7, 11:59 PM PSTMay 7, 11:59 PM PST

Attempts 3 every 24 hours

Try again

Receive grade

To Pass 80% or higher

Your grade

85%

View Feedback

We keep your highest score

Like

Dislike

Report an issue

Chat with us

Back

Weekly challenge 3

Graded Quiz. • 50 min

DueMay 7, 11:59 PM PST

Congratulations! You passed!

Grade received 85%
Latest Submission Grade 85%
To pass 80% or higher

Go to next item

1.

Question 1

A data scientist is trying to print a data frame but when you print the data frame to the console output produces too many rows and columns to be readable. What could they use instead of a data frame to make printing more readable?

1 / 1 point

A list

A tibble

A structure

A vector

Correct

2.

Question 2

A data analyst is working with a large data frame. It contains so many columns that they don’t all fit on the screen at once. The analyst wants a quick list of all of the column names to get a better idea of what is in their data. What function should they use?

1 / 1 point

- `str()`
- `head()`
- `mutate()`
- `colnames()`

Correct

3.

Question 3

You are working with the ToothGrowth dataset. You want to use the `head()` function to get a preview of the dataset. Write the code chunk that will give you this preview.

```
1
2
```

```
head(ToothGrowth)
```

Run Reset

What are the names of the columns in the ToothGrowth dataset?

1 / 1 point

VC, supp, dose

len, supp, dose

len, VC, dose

len, supp, VC

Correct

The code chunk **`head(ToothGrowth)`** gives you a preview of the dataset. Inside the parentheses of the head() function is the name of the dataset you want to preview. The code returns a view of the column names and the first few rows of the dataset. The names of the columns in the ToothGrowth dataset are len, supp, dose.

4.

Question 4

You are cleaning a data frame with improperly formatted column names. In order to clean the data frame you want to use the `clean_names()` function. Which column names will be changed using the `clean_names()` with default parameters? Select all that apply.

0.5 / 1 point

☑

☒

`column.1`

✓ Correct

☐

☐

`column4`

☐

☐

`column 2`

☑

☒

`column_3`

✗ This should not be selected

Review the video on cleaning data in R.

5.

Question 5

A data analyst is working with the penguins dataset and wants to sort the penguins by *body_mass_g* from least to greatest. When they run the following code the penguin body mass data is not displayed in the correct order.

```
penguins %>% arrange(body_mass_g)

head(penguins)
```

What can the data analyst do to fix their code?

1 / 1 point

- Add a minus sign in front of *body_mass_g* to reverse the order
- Save the results of `arrange()` to a variable that gets passed to `head()`
- Use the `print()` function instead of the `head()` function
- Correct the capitalization of `arrange()` to `Arrange()`

Correct

6.

Question 6

You are working with the penguins dataset. You want to use the summarize() and min() functions to find the minimum value for the variable *bill_depth_mm.* You write the following code:

```
penguins %>%

  drop_na() %>%

  group_by(species) %>%
```

Add the code chunk that lets you find the minimum value for the variable *bill_depth_mm.*

```
1
2
      3
4
5
6
7
```

```
library(penguins)

penguins %>%
  drop_na() %>%
  group_by(species) %>%
  summarize(min_bill_depth_mm = min(bill_depth_mm))
```

Run Reset

What is the minimum bill depth in mm for the Chinstrap species?

0 / 1 point

13.1

16.4

12.4

15.5

Incorrect

Review the video on organizing data in R for a refresher.

7.

Question 7

A data analyst is working with a data frame called *athletes*. The data frame contains a column names *record* that represents an athlete's wins and losses separated by a hyphen (`-`). They want to turn this single column into individual columns for *wins* and *losses*. Which code chunk lets the analyst split the *record* column?

1 / 1 point

- `separate(record, athletes, into=c("wins”, “losses”), sep="-")`
- `separate(athletes, record, into=c("wins”, “losses”), delim="-")`
- `separate(athletes, record, into=c("wins”, “losses”), sep="-")`
- `separate(record, athletes, into=c("wins”, “losses”), delim="-")`

Correct

8.

Question 8

A data analyst is working with a data frame named *users*. It has separate columns for first name (*first_name*) and last name (*last_name*). The analyst wants to combine the two columns into a single column called *full_name*, with the first name and last name separated by a space. What code chunk lets the analyst create the *full_name*column?

1 / 1 point

- `unite(users, "full_name", first_name, last_name, sep = " ")`

`unite(users, first_name, last_name, "full_name", sep = " ")`

`merge(users, "full_name", first_name, last_name, sep = " ")`

`unite(users, "full_name", first_name, last_name, sep = ", ")`

Correct

9.

Question 9

You are compiling an analysis of the average monthly costs for your company. What summary statistic function should you use to calculate the average?

1 / 1 point

`cor()`

`max()`

`min()`

`mean()`

Correct

10.

Question 10

A data analyst wants to find out how much the predicted outcome and the actual outcome of their data model differ. What function can they use to quickly measure this?

1 / 1 point

- [x] `bias()`
- [ ] `cor()`
- [ ] `mean()`
- [ ] `sd()`

Correct