

ETC1010

Sway

Lobby

Rooms

Quizzes

Reports

Edit Quiz

Save and Exit

Quiz name

Reading week 2 (DUE: Tue Aug 8, 9am)

1

MULTIPLE CHOICE

Edit

Tidy data is tabular format and has _____ in the columns

Answer choices

A variables

B observations

C cases

D numbers



2

TRUE/FALSE

Edit

```
#> # A tibble: 6 × 4
#>   country year cases population
#>   <chr> <int> <int>    <int>
#> 1 Afghanistan 1999   745  19987071
#> 2 Afghanistan 2000  2666  20595360
#> 3   Brazil 1999 37737 172006362
#> 4   Brazil 2000 80488 174504898
#> 5    China 1999 212258 1272915272
#> 6    China 2000 213766 1280428583
```

Is this table in tidy format?

Answer

True

3

MULTIPLE CHOICE

Edit

There are three interrelated rules which make a dataset tidy. Choose them from the list below.

Answer choices

- | | |
|----------|---|
| A | Each variable must have its own column. |
| B | Each observation must have its own row. |
| C | Each value must have its own cell. |
| D | Each cell must be allowed to have any type of value in it |

4

SHORT ANSWER

Edit

The first step is always to figure out what the variables and _____ are.

Answer

observations

5

SHORT ANSWER

Edit

To tidy a dataset with one variable spread across multiple columns, we need to _____ those columns into a new pair of variables.

Answer

gather

6

MULTIPLE CHOICE

Edit

```
table4a # cases
#> # A tibble: 3 × 3
#>   country `1999` `2000`
#> *   <chr> <int> <int>
#> 1 Afghanistan  745  2666
#> 2   Brazil 37737 80488
#> 3    China 212258 213766
```

What are the variables in this table? (Choose all that apply)

Answer choices

A	country
B	year
C	tb
D	neighbour

7

MULTIPLE CHOICE

Edit

The pair of new variables created with gather are (choose all that apply)

Answer choices

A	key
B	value
C	observation
D	variable

8

SHORT ANSWER

Edit

_____ is the opposite of gathering

Answer

Spreading

9

SHORT ANSWER

Edit

Spreading is used when an _____ is scattered across multiple rows.

Answer

observation

10

SHORT ANSWER

Edit

table3

#> # A tibble: 6 × 3

#> country year rate

#> * <chr> <int> <chr>

#> 1 Afghanistan 1999 745/19987071

#> 2 Afghanistan 2000 2666/20595360

#> 3 Brazil 1999 37737/172006362

#> 4 Brazil 2000 80488/174504898

#> 5 China 1999 212258/1272915272

#> 6 China 2000 213766/1280428583

In this table we have two variables cases and population in one column.

_____ pulls apart one column into multiple columns, by splitting wherever a separator character appears.

Answer

Separate

11

MULTIPLE CHOICE

Edit

Which of these indicates a missing value in R?

Answer choices

A	NA
B	simply a blank
C	.
D	Inf

12 | **TRUE/FALSE**[Edit](#)

An implicit missing value is simply a value not present in the data.

Answer

True

13 | **MULTIPLE CHOICE**[Edit](#)

Sometimes when a data source has primarily been used for data entry, missing values indicate that the previous value should be carried forward:

```
treatment <- tribble(
  ~ person,      ~ treatment, ~response,
  "Derrick Whitmore", 1,      7,
  NA,            2,      10,
  NA,            3,      9,
  "Katherine Burke", 1,      4
)
```

You can fill in these missing values with _____

Answer choices

A	fill()
B	complete()
C	spread()
D	na.rm=TRUE

14 | MULTIPLE CHOICE

Edit

What does the line

```
mutate(code = stringr::str_replace(code, "newrel", "new_rel"))
```

do?

Answer choices

- | | |
|---|--|
| A | puts an "_" into the text string |
| B | removes an "_" from the text string |
| C | creates a new variable called newrel |
| D | mutates a variable new_rel into newrel |

15 | MULTIPLE CHOICE

Edit

Which are valid reasons to work with non-tidy data? (Choose all that apply)

Answer choices

- | | |
|---|---|
| A | Alternative representations may have substantial performance or space advantages. |
| B | Specialised fields have evolved their own conventions for storing data that may be quite different to the conventions of tidy data. |
| C | Its never reasonable |

[+ Multiple Choice](#)[+ True/False](#)[+ Short Answer](#)