

ETC2420 5242

Sway

Lobby

Activities

Reports

Edit Quiz

Cancel

Save and Exit

Quiz name

Weekly quiz 10 (DUE: Fri Oct 6 10am)

1

MULTIPLE CHOICE

Edit

It is often said that ___% of data analysis is spent on the process of cleaning and preparing the data

Answer choices

A	40
B	50
C	60
D	80



2 | **TRUE/FALSE**

Edit

The principles of tidy data provide a standard way to organize data values within a dataset.

Answer

True

**3** | **MULTIPLE CHOICE**

Edit

The earlier versions of the tidyr package in R were (select all that apply)

Answer choices

A	reshape
B	reshape2
C	plyr
D	lubridate

**4** | **MULTIPLE CHOICE**

Edit

Like _____, tidy datasets are all alike but every messy dataset is messy in its own way.

Answer choices

A	families
B	statistics
C	mathematics
D	data science



5

MULTIPLE CHOICE

Edit

A dataset is a collection of _____, usually either numbers (if quantitative) or strings (if qualitative)

Answer choices

A	variables
B	observations
C	values
D	data



6

MULTIPLE CHOICE

Edit

Every value belongs to a _____ and a(n) _____.

Answer choices

A	variable, observation
B	row, column
C	table, dataset
D	mean, variance



7

MULTIPLE CHOICE

Edit

A _____ contains all values that measure the same underlying attribute (like height, temperature, duration) across units.

Answer choices

A	observation
B	variable
C	dataset
D	statistic

8

MULTIPLE CHOICE

Edit

An _____ contains all values measured on the same unit (like a person, or a day, or a race) across attributes.

Answer choices

A	observation
B	variable
C	dataset
D	statistic

9 | MULTIPLE CHOICE

Edit

In tidy data: (check all the apply)

Answer choices

A	Each variable forms a column.
B	Each observation forms a row.
C	Each type of observational unit forms a table.
D	Each dataset forms a table.



10 | TRUE/FALSE

Edit

Messy data is any other arrangement of the data.

Answer

True



11 | MULTIPLE CHOICE

Edit

The terminology "molten" or "melted" in this description of tidying is now called _____ in the tidyr package.

Answer choices

A	summarise
B	spread
C	gather
D	collect



12 | **TRUE/FALSE**

Edit

A common type of messy data has variables buried in the column names, e.g. for the TB data m014, m1524, m2534, m3544, m4554, m5564, m65, mu, f014,

Answer

True

**13** | **MULTIPLE CHOICE**

Edit

_____ means subsetting or removing observations based on some condition.

Answer choices

A	Sort
B	Aggregate
C	Transform
D	Filter

**14** | **MULTIPLE CHOICE**

Edit

The package that preceded the dplyr package was:

Answer choices

A	plyr
B	reshape
C	reshape2
D	lubridate



15 | TRUE/FALSE

Edit

Arrange is a verb that means to sort.

Answer

True

+ Multiple Choice

+ True/False

+ Short Answer