Default Question Block

First Name

Last Name

email

These page timer metrics will not be displayed to the recipient.
First Click: *0 seconds*
Last Click: *0 seconds*
Page Submit: *0 seconds*
Click Count: *0 clicks*

1 I have a data frame named DATA with the following variables: student_id, grade, gender, logins, comments. What code would I use to create a data frame containing **only** student_id and logins?

- gather(DATA$student_id, DATA$logins)
- dplyr::filter(DATA, student_id, logins)
- tidyr::unite(DATA, student_id, logins, sep = ",")
- dplyr::select(DATA, student_id, logins)

These page timer metrics will not be displayed to the recipient.
First Click: *0 seconds*
Last Click: *0 seconds*
Page Submit: *0 seconds*
Click Count: *0 clicks*

2 Lalitha has anonymized a data set by replacing student names with a variable called "student_id". "student_id" is a random number given to each student, what class of variable should "student_id" be in R?

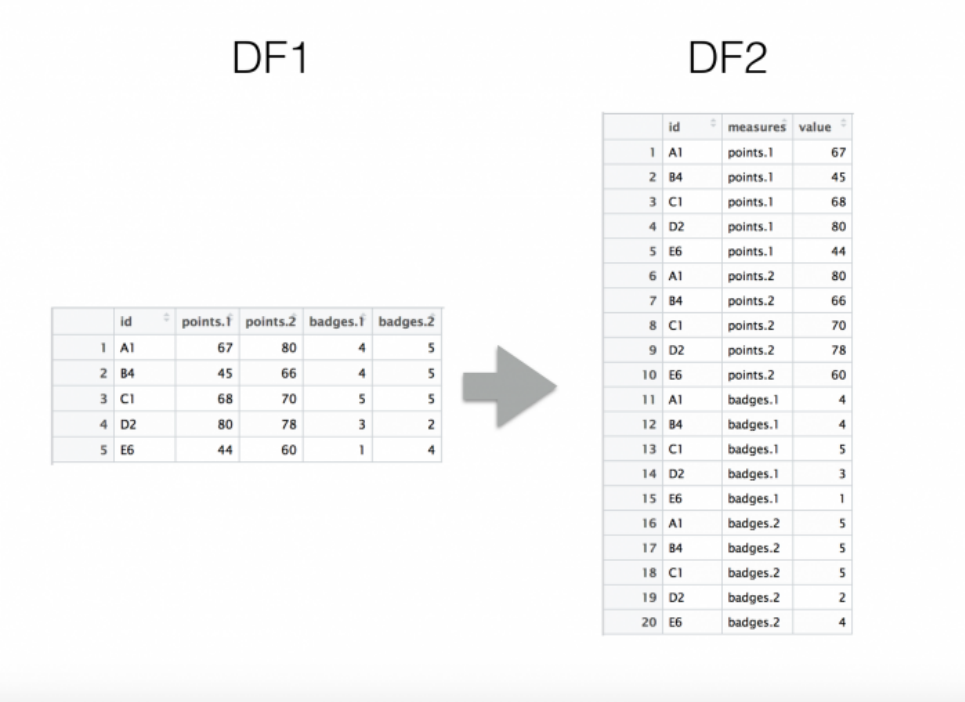
- numeric
- vector
- factor
- polynomial

These page timer metrics will not be displayed to the recipient.
First Click: *0 seconds*
Last Click: *0 seconds*
Page Submit: *0 seconds*
Click Count: *0 clicks*

3 Which R package (library) would you utilize to group students by grade, and produce the average test score for each grade?

- tidyr
- dplyr
- rpart
- ggplot2

These page timer metrics will not be displayed to the recipient.
First Click: *0 seconds*
Last Click: *0 seconds*
Page Submit: *0 seconds*
Click Count: *0 clicks*



	id	points.1	points.2	badges.1	badges.2
1	A1	67	80	4	5
2	B4	45	66	4	5
3	C1	68	70	5	5
4	D2	80	78	3	2
5	E6	44	60	1	4

	id	measures	value
1	A1	points.1	67
2	B4	points.1	45
3	C1	points.1	68
4	D2	points.1	80
5	E6	points.1	44
6	A1	points.2	80
7	B4	points.2	66
8	C1	points.2	70
9	D2	points.2	78
10	E6	points.2	60
11	A1	badges.1	4
12	B4	badges.1	4
13	C1	badges.1	5
14	D2	badges.1	3
15	E6	badges.1	1
16	A1	badges.2	5
17	B4	badges.2	5
18	C1	badges.2	5
19	D2	badges.2	2
20	E6	badges.2	4

4 What is the R code to convert DF1 into DF2?

These page timer metrics will not be displayed to the recipient.
First Click: *0 seconds*
Last Click: *0 seconds*
Page Submit: *0 seconds*
Click Count: *0 clicks*

5 A "tidy" data set is one in which:

- A variable occupies its own row, and an observation occupies its own column
- A variable occupies its own column, and an observation occupies its own row
- A variable occupies its own row and an observation also occupies its own row
- A variable occupies its own column and an observation also occupies its own column

These page timer metrics will not be displayed to the recipient.
First Click: *0 seconds*
Last Click: *0 seconds*
Page Submit: *0 seconds*
Click Count: *0 clicks*

6 You have received a .csv data file from a colleague in Europe. You are excited about it and decide to immediately upload it into R. Unfortunately you get the error, "non-numeric argument to binary operator". What might be the most likley cause?

- Some character variables have been mislabeled as numeric variables
- The file was not a .csv file
- Europeans use commas to indicate decimal places
- You used the code "read.table()" instead of "read.csv()"

These page timer metrics will not be displayed to the recipient.
First Click: *0 seconds*
Last Click: *0 seconds*
Page Submit: *0 seconds*
Click Count: *0 clicks*

7 When building CART models, what is the "splitting criteria"

- The threshold at which the algorithm decides to split a node into two descendant nodes
- The point at which the model breaks down
- The threshold used to generate a new tree
- The threshold used to determine which cost function to use

These page timer metrics will not be displayed to the recipient.
First Click: *0 seconds*
Last Click: *0 seconds*
Page Submit: *0 seconds*
Click Count: *0 clicks*

	id	score	levels	engagement	friends
1	A	90	2	22	2
2	B	100	3	23	3
3	C	87	5	15	5
4	D	55	2	28	2
5	E	76	6	25	2

8 Refer to the data frame above named DF. What package (library) would you use to build a CART model that predicts engagement from the other numeric variables?

- ggplot2
- base
- rpart
- topicmodels

9 What is the R code to fit the CART model using this library?

These page timer metrics will not be displayed to the recipient.
First Click: *0 seconds*
Last Click: *0 seconds*
Page Submit: *0 seconds*
Click Count: *0 clicks*

10 What is **recursive partitioning**?

- Back-propagation in neural networks
- When each leaf in a decision tree may be split an indefinite number of times
- When each leaf in a a decision tree uses the same splitting criterion
- When each leaf in a decision tree is dependent on the previous leaf

These page timer metrics will not be displayed to the recipient.
First Click: *0 seconds*
Last Click: *0 seconds*
Page Submit: *0 seconds*
Click Count: *0 clicks*

11 In Social Network Analysis, why is an ego network useful?

- Because large networks are inherently unstable and small networks are not
- Because students are very self interested
- Because large networks can obscure important relationships between individuals
- Because smaller networks take less time analyze

These page timer metrics will not be displayed to the recipient.
First Click: *0 seconds*
Last Click: *0 seconds*
Page Submit: *0 seconds*
Click Count: *0 clicks*

	1	2	3	4	5	6
1	0	1	1	0	0	0
2	1	0	1	1	1	0
3	1	1	0	0	0	1
4	0	1	0	0	0	0
5	0	1	0	0	0	0
6	0	0	1	0	0	0

12 What is the R code to create a graph object from the matrix above?

13 What is the R code to extract the **Bonacich Power Centrality** from this graph object?

These page timer metrics will not be displayed to the recipient.
First Click: *0 seconds*
Last Click: *0 seconds*
Page Submit: *0 seconds*
Click Count: *0 clicks*

14 What are the main two families of approach to finding subgroup structure in Social Networks?

- Bottom up and top down
- Density and degree
- Cliques and neighborhoods
- The right way and the wrong way

These page timer metrics will not be displayed to the recipient.
First Click: *0 seconds*
Last Click: *0 seconds*
Page Submit: *0 seconds*
Click Count: *0 clicks*

Define the term **maximal complete sub-graph**

15

These page timer metrics will not be displayed to the recipient.
First Click: *0 seconds*
Last Click: *0 seconds*
Page Submit: *0 seconds*
Click Count: *0 clicks*

16 You have created an igraph object called "g". What is the R code to extract the the size of the largest clique in the graph?

These page timer metrics will not be displayed to the recipient.
First Click: *0 seconds*
Last Click: *0 seconds*
Page Submit: *0 seconds*
Click Count: *0 clicks*

17 Natural Language Processing utilizes a "bag of words" approach. This refers to:

- The ways that words are grouped (bagged) to infer meaning from a text
- That grammar and word order are disregarded
- That students who study NLP have bags under their eyes from lack of sleep
- That grammar and syntax can be used to predict the meaning of sentences

These page timer metrics will not be displayed to the recipient.
First Click: *0 seconds*
Last Click: *0 seconds*
Page Submit: *0 seconds*
Click Count: *0 clicks*

18 What is the purpose of pre-processing in NLP?

These page timer metrics will not be displayed to the recipient.
First Click: *0 seconds*
Last Click: *0 seconds*
Page Submit: *0 seconds*
Click Count: *0 clicks*

In the following sentence which words would be considered **stop words**

19 **The quick brown fox jumps over the lazy dog**

These page timer metrics will not be displayed to the recipient.
First Click: *0 seconds*
Last Click: *0 seconds*
Page Submit: *0 seconds*
Click Count: *0 clicks*

What is the R code to remove whitespace from a corpus?

20

These page timer metrics will not be displayed to the recipient.
First Click: *0 seconds*
Last Click: *0 seconds*
Page Submit: *0 seconds*
Click Count: *0 clicks*

21 In LDA topic modeling a **topic** is defined as:

- Associated words within a document
- A probability distribution over a distribution of probabilites
- A distribution of words over a probability of drawing a document at random
- A probability distribution over a fixed vocabulary

These page timer metrics will not be displayed to the recipient.
First Click: *0 seconds*
Last Click: *0 seconds*
Page Submit: *0 seconds*
Click Count: *0 clicks*

What is a Term Frequency Inverse Document Frequency and what does it tell us about a corpus?

22

These page timer metrics will not be displayed to the recipient.
First Click: *0 seconds*
Last Click: *0 seconds*
Page Submit: *0 seconds*
Click Count: *0 clicks*

What is the R code to generate a Term Frequency Inverse Document Frequency from a Document Term Matrix named "dtm"?

23

These page timer metrics will not be displayed to the recipient.
First Click: *0 seconds*
Last Click: *0 seconds*
Page Submit: *0 seconds*
Click Count: *0 clicks*

Describe three uses for NLP in educational settings

24

These page timer metrics will not be displayed to the recipient.
First Click: *0 seconds*
Last Click: *0 seconds*
Page Submit: *0 seconds*
Click Count: *0 clicks*

Describe a possible danger with the application of NLP to educational data

25

These page timer metrics will not be displayed to the recipient.
First Click: *0 seconds*
Last Click: *0 seconds*
Page Submit: *0 seconds*
Click Count: *0 clicks*

	student	session	score
1	A	1	24
2	A	2	1
3	A	3	15
4	B	1	23
5	B	2	7
6	B	3	10
7	C	1	7
8	C	2	5
9	C	3	13
10	D	1	13
11	D	2	12
12	D	3	5

26 Referring to the data frame above, using the R package (library) ggplot2, what is the R code to generate a line graph where each students' score changes are visualized as their own line on a single panel?

These page timer metrics will not be displayed to the recipient.
First Click: *0 seconds*
Last Click: *0 seconds*
Page Submit: *0 seconds*
Click Count: *0 clicks*

27 In neural networks, what is the delta rule?

These page timer metrics will not be displayed to the recipient.
First Click: *0 seconds*
Last Click: *0 seconds*
Page Submit: *0 seconds*
Click Count: *0 clicks*

28 In neural networks, what is the purpose of back-propagagtion?

- To generate the most accurate prediction of the output layer
- To train the hidden layer to better predict the output of the input layer
- To adjust the weights on the synapses to improve prediction accuracy
- To create artificial intelligence

These page timer metrics will not be displayed to the recipient.
First Click: *0 seconds*
Last Click: *0 seconds*
Page Submit: *0 seconds*
Click Count: *0 clicks*

29 What benefits do you believe artificial intelligence might bring to education?

These page timer metrics will not be displayed to the recipient.
First Click: *0 seconds*
Last Click: *0 seconds*
Page Submit: *0 seconds*
Click Count: *0 clicks*

30 What dangers do you foresee with greater use of artificial intelligence in educational settings?

These page timer metrics will not be displayed to the recipient.
First Click: *0 seconds*
Last Click: *0 seconds*
Page Submit: *0 seconds*
Click Count: *0 clicks*