Pre-Workshop Exercise Exercise 1 Exercise 2 Exercise 3 Summative

Pre-Workshop Exercise

Please write the code for the following pipeline steps:

- 1. Load the tidyverse and readxl libraries.
- 2. Read in the Excel file located in: "data/medicaldata_tumorgrowth.xlsx" into a variable tumor.
- 3. Select the all the columns except Grp, and filter the rows such that Day is 0 or 20. Save this data subset into a variable tumor subset.
- 4. We want to compare baseline tumor sizes (Day 0) with tumor sizes at Day 20 between each of the groups.
 - Using tumor_subset, calculate the average
 tumor Size for each Grp and Day.
- 5. Save tumor_subset into a CSV file located in "data/tumorsubset.csv".



Next Topic

1 of 2 2021-11-04, 14:15

Pre-Workshop Exercise

Exercise 1

Exercise 2

Exercise 3

Summative

Start Over

Exercise 1

Take a look at the ebola dataset.

ebola

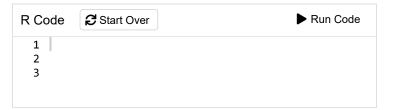
date <chr></chr>	country <chr></chr>	289 <dbl></dbl>		287 · <dbl></dbl>	286 · <dbl></dbl>
1/5/2015	Guinea	2776	NA	NA	NA
1/5/2015	SierraLeone	10030	NA	NA	NA
1/4/2015	Guinea	NA	2775	NA	NA
1/4/2015	SierraLeone	NA	9780	NA	NA
1/3/2015	Guinea	NA	NA	2769	NA
1/3/2015	Liberia	NA	NA	8166	NA
1/3/2015	SierraLeone	NA	NA	9722	NA
1/2/2015	Liberia	NA	NA	NA	8157
12/31/2014	Guinea	NA	NA	NA	NA
12/31/2014	Liberia	NA	NA	NA	NA
1 Previous	1 2 3 4	4 5	6	37 I	Vext

- Tidy the dataset such that you get the dataset below.
- You can use the last_col() to select the last column of the dataset.
- Remember to drop missing values as the last step.

date <chr></chr>	country <chr></chr>	day cases <chr> <dbl></dbl></chr>
1/5/2015	Guinea	289 2776
1/5/2015	SierraLeone	289 10030
1/4/2015	Guinea	288 2775
1/4/2015	SierraLeone	288 9780
1/3/2015	Guinea	287 2769
1/3/2015	Liberia	287 8166

ds4biomed	
Pre-Workshop Exercise	
Exercise 1	
Exercise 2	
Exercise 3	
Summative	
Start Over	

date <chr></chr>		cou <chi< th=""><th>ntry r></th><th></th><th></th><th>day <chr></chr></th><th>cases <dbl></dbl></th></chi<>	ntry r>			day <chr></chr>	cases <dbl></dbl>
1/3/2015		Sier	raLe	one		287	9722
1/2/2015		Libe	ria			286	8157
12/31/2014		Guir	nea			284	2730
12/31/2014		Libe	ria			284	8115
1 Previous	1	2	3	4	5	6 37	' Next



Previous Topic

Next Topic

Start Over

Pre-Workshop Exercise

Exercise 1
Exercise 2
Exercise 3
Summative

Exercise 2

This is a different version of the <code>ebola</code> dataset.

date <chr></chr>	da Case <dbl></dbl>	s_Guinea <dbl></dbl>	Cases_Liberia <dbl></dbl>
1/5/2015	289	2776	NA
1/4/2015	288	2775	NA
1/3/2015	287	2769	8166
1/2/2015	286	NA	8157
12/31/2014	284	2730	8115
12/28/2014	281	2706	8018
12/27/2014	280	2695	NA
12/24/2014	277	2630	7977
12/21/2014	273	2597	NA
12/20/2014	272	2571	7862
1 Previous	1 2	3 4 5	6 13 Next

- Tidy the dataset such that you get the dataset below
- Remember to drop missing values as the last step

date <chr></chr>	da; case_death <dbl≼chr></dbl≼chr>	country <chr></chr>	val <dbl></dbl>
1/5/2015	289 Cases	Guinea	2776
1/5/2015	289 Cases	SierraLeone	10030
1/5/2015	289 Deaths	Guinea	1786
1/5/2015	289 Deaths	SierraLeone	2977
1/4/2015	288 Cases	Guinea	2775
1/4/2015	288 Cases	SierraLeone	9780
1/4/2015	288 Deaths	Guinea	1781
1/4/2015	288 Deaths	SierraLeone	2943
1/3/2015	287 Cases	Guinea	2769
1/3/2015	287 Cases	Liberia	8166

1 of 2 2021-11-04, 14:16

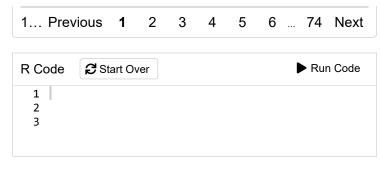
Pre-Workshop Exercise

Exercise 1

Exercise 2

Exercise 3

Summative



Previous Topic Next Topic

Start Over

Pre-Workshop Exercise

Exercise 1	
Exercise 2	
Exercise 3	
Summative	

Exercise 3

Using the same <code>ebola</code> dataset example from the previous exercise:

date <chr></chr>	da; Cas <dbl></dbl>	es_Guin <db< th=""><th></th><th>Cases</th><th>_</th><th>eria dbl></th></db<>		Cases	_	eria dbl>
1/5/2015	289	27	76			NA
1/4/2015	288	27	75			NA
1/3/2015	287	27	69		8	3166
1/2/2015	286	/	VA		8	3157
12/31/2014	284	27	30		8	3115
12/28/2014	281	27	06		8	3018
12/27/2014	280	26	95			NA
12/24/2014	277	26	30		7	977
12/21/2014	273	25	97			NA
12/20/2014	272	25	71		7	7862
1 Previous	1 2	3 4	5	6	13	Next

- Tidy the dataset such that it looks like the below example
- Remember to drop missing values as the last step

date <chr></chr>	day country <dbl×chr></dbl×chr>	Ca <dbl></dbl>	Deaths <dbl></dbl>
1/5/2015	289 Guinea	2776	1786
1/5/2015	289 SierraLeone	10030	2977
1/4/2015	288 Guinea	2775	1781
1/4/2015	288 SierraLeone	9780	2943
1/3/2015	287 Guinea	2769	1767
1/3/2015	287 Liberia	8166	3496
1/3/2015	287 SierraLeone	9722	2915
1/2/2015	286 Liberia	8157	3496
12/31/2014	284 Guinea	2730	1739

1 of 2 2021-11-04, 14:16

Start Over

ds4biomed	dateday countryCaDeaths <chr><dbl×chr><dbl><dbl></dbl></dbl></dbl×chr></chr>
	12/31/2014 284 Liberia 8115 3471
Pre-Workshop Exercise	1 Previous 1 2 3 4 5 6 37 Next
Exercise 1	R Code
Exercise 2	1 2 3
Exercise 3	
Summative	Previous Topic Next Topic

Pre-Workshop Exercise

Exercise 1
Exercise 2
Exercise 3
Summative
Start Over

Summative

This is the cmv dataset you will load:

	ag ⊦ l×dbl>	pric	r.ra			aK. . <dbl< th=""><th></th><th></th><th></th><th>_</th><th>_ne</th><th>gati</th></dbl<>				_	_ne	gati
1	61				0		1	1	rec	ipieı	nt_	posit
2	62				1		5	0	rec	ipieı	nt_	nega
3	63				0		3	0	NA			
4	33				1		2	0	rec	ipieı	nt_	posit
5	54				0		6	0	NA			
6	55				0		2	1	NA			
7	67				0		1	0	NA			
8	51				0		2	0	NA			
9	44				1		2	1	NA			
10	59				0		4	0	rec	ipieı	nt_	nega
1-1	. Previo	us '	1	2	3	4	5		6	7	N	ext

- 1. Use the readxl library to load the data/cmv.xlsx into a variable, cmv
- 2. Filter the cmv dataset such that only age > 65 are remaining. Save this to a variable, cmv_subset.
- 3. Save the cmv_subset variable to a csv file in "data/cmv_subset.csv".



4. Tidy the \mbox{cmv} dataset such that it looks like the \mbox{clean} dataset below. Save the tidy dataset into a varialbe, $\mbox{cmv_tidy}$.

ID ag ⊢ <dbl×dbl></dbl×dbl>	-		cm donor_status <dbl≭chr></dbl≭chr>
1 61	0	1	1 donor_negativ
2 62	1	5	0 donor_negative

Pre-Workshop Exercise
Exercise 1
Exercise 2
Exercise 3
Summative
Start Over

ID ag <dbl×dbl></dbl×dbl>	prior.r		aK <dbl></dbl>		donor_status
3 63		0	3	0	donor_positive
4 33		1	2	0	donor_negativ
5 54		0	6	0	donor_positive
6 55		0	2	1	donor_positive
7 67		0	1	0	donor_positive
8 51		0	2	0	donor_positive
9 44		1	2	1	donor_positive
10 59		0	4	0	donor_negativ
1-1 Previo	us 1	2 3	4 !	5	6 7 Next



5. In the ${\tt cmv_tidy}$ dataset, calculate the average age for each value of ${\tt cmv}$.



Previous Topic

2 of 3 2021-11-04, 14:16

Start Over

Pre-W	orkshop Exercise	
Exerci	se 1	
Exerci	se 2	
Exerci	se 3	
Summ	ative	