Modifying and adding to data frames for the infrequent useR

Kim Cressman

Grand Bay NERR

kimberly.cressman@msstate.edu

May 19-20, 2022

Remember how we found and identified parts of data frames

`df[row, column]`

Leave one blank to select *everything*

Selecting Columns

Base: df$column_name

Tidyverse: df %>% select(column_name)

Logical conditions

- We use `==` to specify an exact condition (is equal to)
  - < less than
  - <= less than or equal to
  - == equals
  - != *not* equal to
  - >= greater than or equal to
  - > greater than

Now to do things

Add a column

	date	time	depth_m	temp_c	temp_f
1	1/24/2022	11:45:00	0.542	9.167	48.5006
2	1/24/2022	12:00:00	0.530	9.411	48.9398
3	1/24/2022	12:15:00	0.548	9.784	49.6112

- Example: You have water temp in degrees C, and want degrees F
- Base R: $$$$
  - Create your new column on the left side of the arrow, and perform operations on the right
  - bcwq$temp_f <- (bcwq$temp_c * 9/5) + 32
- dplyr: mutate
  - bcwq <- bcwq %>%

    mutate(temp_f = (temp_c * 9/5) + 32)

Add multiple columns

	date	time	depth_m	temp_c	temp_f	depth_ft
1	1/24/2022	11:45:00	0.542	9.167	48.5006	1.778302
2	1/24/2022	12:00:00	0.530	9.411	48.9398	1.738930
3	1/24/2022	12:15:00	0.548	9.784	49.6112	1.797988
4	1/24/2022	12:30:00	0.557	9.745	49.5410	1.827517
5	1/24/2022	12:45:00	0.553	9.608	49.2944	1.814393

- Temp F, and also Depth in ft
- Base R:
  - bcwq$temp_f <- (bcwq$temp_c * 9/5) + 32
  - bcwq$depth_ft <- bcwq$depth_m * 3.281
- dplyr:
  - bcwq <- bcwq %>%

```
mutate(    temp_f = (temp_c * 9/5) + 32,
           depth_ft = depth_m * 3.281     )
```

With mutate(), you can use a column immediately after creating it!

- bcwq %>%

```
mutate(    temp_f = (temp_c * 9/5) + 32,
           depth_ft = depth_m * 3.281,
           depth_miles = depth_ft / 5280
       )
```

Credit: Allison Horst