

Code Sleuth

Pivot wider

Inference

Code example	Output
<pre>weight_by_sex <- surveys > filter(!is.na(sex)) > summarize(mean_weight = mean(weight, na.rm = TRUE), .by = c(species_id, sex)) weight_by_sex</pre>	<pre># A tibble: 49 × 3 species_id sex mean_weight <chr> <chr> <dbl> 1 NL M 166. 2 DM F 41.6 3 DM M 44.4 4 PF M 7.89 5 PE F 22.8 # i 44 more rows</pre>
<pre>weight_by_sex_wider <- weight_by_sex > pivot_wider(names_from = sex, values_from = mean_weight) weight_by_sex_wider</pre>	<pre># A tibble: 26 × 3 species_id M F <chr> <dbl> <dbl> 1 NL 166. 154. 2 DM 44.4 41.6 3 PF 7.89 7.97 4 PE 20.6 22.8 5 DS 122. 118. # i 21 more rows</pre>
<pre>taxa_by_year <- surveys > filter(year >= 1999) > inner_join(select(species, species_id, taxa), join_by(species_id)) > count(year, taxa) taxa_by_year</pre>	<pre># A tibble: 7 × 3 year taxa n <dbl> <chr> <int> 1 1999 Rodent 1108 2 2000 Bird 4 3 2000 Rodent 1505 4 2001 Bird 4 5 2001 Rodent 1536 # i 2 more rows</pre>
<pre>taxa_wider <- taxa_by_year > pivot_wider(names_from = taxa, values_from = n) taxa_wider</pre>	<pre># A tibble: 4 × 3 year Rodent Bird <dbl> <int> <int> 1 1999 1108 NA 2 2000 1505 4 3 2001 1536 4 4 2002 2186 6</pre>

What column disappeared from the output between the first and second examples? Between the third and fourth?

Topic: *Reshaping data*

What column appeared in the output between the first and second examples? Between the third and fourth?

Describe what you think “pivot_wider()” does. What do you think the parameters do?

Prediction

What columns do you think will be in the data frames produced by the following code examples?

Code example	Output
<pre>spring_counts <- surveys > filter(year >= 1998, month >= 3, month < 6) > count(year, month) spring_counts</pre>	
<pre>spring_wide <- spring_counts > pivot_wider(names_from = month, values_from = n) spring_wide</pre>	

Pivot longer

Inference

Code example	Output
<pre>weight_by_sex_wider > pivot_longer(cols = M:F, names_to = "sex", values_to = "mean_weight")</pre>	<pre># A tibble: 52 × 3 species_id sex mean_weight <chr> <chr> <dbl> 1 NL M 166. 2 NL F 154. 3 DM M 44.4 4 DM F 41.6 5 PF M 7.89 # i 47 more rows</pre>
<pre>taxa_wider > pivot_longer(cols = !year, names_to = "taxa", values_to = "n")</pre>	<pre># A tibble: 8 × 3 year taxa n <dbl> <chr> <int> 1 1999 Rodent 1108 2 1999 Bird NA 3 2000 Rodent 1505 4 2000 Bird 4 5 2001 Rodent 1536 # i 3 more rows</pre>

In “pivot_longer()”, what happens to the columns specified by the “cols” parameter?

What’s the relationship between the “cols”, “names_to”, and “values_to” parameters?

What does “pivot_longer()” do? What’s its relationship to “pivot_wider()”?

Prediction

What is the data frame returned by the following code example?

Assume the following runs first.

```
> spring_counts <- surveys |>
+   filter(year >= 1998, month >= 3, month < 6) |>
+   count(year, month)
> spring_counts
# A tibble: 14 × 3
  year month     n
  <dbl> <dbl> <int>
1  1998     3   204
2  1998     5   267
3  1999     3   127
4  1999     4   138
5  1999     5   133
# i 9 more rows
# i Use `print(n = ...)` to see more rows
> spring_wide <- spring_counts |>
+   pivot_wider(
+     names_from = month,
+     values_from = n
+   )
> spring_wide
# A tibble: 5 × 4
  year   `3`   `5`   `4`
  <dbl> <int> <int> <int>
1  1998   204   267   NA
2  1999   127   133   138
3  2000   160    74    68
4  2001   146   117    99
5  2002   164   243   202
```

Code example	Output
<pre>pivot_longer(spring_wide, 2:4, names_to = "MONTH", values_to = "COUNT")</pre>	