

For the letter G, I'd like to introduce a very useful function: `group_by`. This function lets you group data by one or more variables. By itself, it may not seem very useful, but it's great when you start manipulating and summarizing data. That's because many of the functions applied to data after you used `group_by` are done groupwise. First, let's demonstrate the effect `group_by` has on a filter. I'll load my reading dataset and group it by the Fiction flag (so 1 means the book was fiction and 0 means it was non-fiction). What was the longest book I read in each of those categories?

```
library(tidyverse)

## -- Attaching packages ----- tidyverse
1.3.0 --

##   ggplot2 3.2.1      purrr   0.3.3
##   tibble  2.1.3      dplyr   0.8.3
##   tidyr   1.0.0      stringr 1.4.0
##   readr   1.3.1      forcats 0.4.0

## -- Conflicts -----
tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()     masks stats::lag()

reads2019 <- read_csv("~/Downloads/Blogging A to Z/SaraReads2019_allrated.csv",
  col_names = TRUE)

## Parsed with column specification:
## cols(
##   Title = col_character(),
##   Pages = col_double(),
##   date_started = col_character(),
##   date_read = col_character(),
##   Book.ID = col_double(),
##   Author = col_character(),
##   AdditionalAuthors = col_character(),
##   AverageRating = col_double(),
##   OriginalPublicationYear = col_double(),
##   read_time = col_double(),
##   MyRating = col_double(),
##   Gender = col_double(),
##   Fiction = col_double(),
##   Childrens = col_double(),
##   Fantasy = col_double(),
##   SciFi = col_double(),
##   Mystery = col_double(),
##   SelfHelp = col_double()
## )

reads2019 %>%
  group_by(Fiction) %>%
  filter(Pages == max(Pages)) %>%
  select(Title, Pages, MyRating, Fiction)
```

```
## # A tibble: 2 x 4
## # Groups:   Fiction [2]
##   Title          Pages MyRating Fiction
##
## 1 How Music Works   345         5         0
## 2 It                1156        4         1
```

*It* was my longest book overall, and therefore the longest fiction book. The longest non-fiction book I read was *How Music Works*, an amazing exploration of music history both generally and personally for the author, David Byrne (from Talking Heads). I picked the book up at an airport bookstore and absolutely loved it.

I can also group by multiple variables, and use that to summarize my data. Let's see what happens when I use two of my genre variables for grouping, then generate a count of the number of books in each group.

```
reads2019 %>%
  group_by(Childrens, Fantasy) %>%
  summarise(count = n())
```

```
## # A tibble: 4 x 3
## # Groups:   Childrens [2]
##   Childrens Fantasy count
##
## 1         0         0    49
## 2         0         1    21
## 3         1         0     1
## 4         1         1    16
```

Since these genres aren't mutually exclusive, grouping by 2 variables gave me 4 groups: 49 of the books I read last year were neither children's fiction nor fantasy, 21 were fantasy not written for children, 1 was children's fiction but not fantasy, and 16 were children's fantasy. Let's update the code to have it also give me the longest book from each of those categories.

```
reads2019 %>%
  arrange(desc(Pages)) %>%
  group_by(Childrens, Fantasy) %>%
  summarise(count = n(),
            title = first(Title),
            Pages = first(Pages))
```

```
## # A tibble: 4 x 5
## # Groups:   Childrens [2]
##   Childrens Fantasy count title          Pages
##
## 1         0         0    49 The Robber Bride          528
## 2         0         1    21 It                1156
## 3         1         0     1 The Queen's Corgi: On Purpose    181
## 4         1         1    16 The Patchwork Girl of Oz (Oz, #7)   346
```

Of course *It* will end up here, since it's the longest book I read in 2019. It's also the longest fantasy not written for children (although it is written *about* children). *The Robber Bride* was the longest non-fantasy, non-children's book. *The Queen's Corgi* was the one non-fantasy children's book I read, and *The Pathwork Girl of Oz* was the longest children's fantasy.