# book_joins

*Matthew Biggs*

Make two tibbles with the book character data

```
characters_tb = tibble(
  name = c("Harry", "Ron", "Frodo", "Merry"),
  fav_food = c("Pumpkin pasties", "Chocoloate Frogs", "Lembas bread", "Mushrooms"),
  friends = c("Hermione, Ron", "Harry, Hermione", "Samwise", "Pippin"),
  book = c("Harry Potter", "Harry Potter", "Lord of the Rings", "Lord of the Rings")
)
```

```
## Warning: `list_len()` is deprecated as of rlang 0.2.0.
## Please use `new_list()` instead.
## This warning is displayed once per session.
```

```
kable(characters_tb)
```

| name | fav_food | friends | book |
|------|----------|---------|------|
| Harry | Pumpkin pasties | Hermione, Ron | Harry Potter |
| Ron | Chocoloate Frogs | Harry, Hermione | Harry Potter |
| Frodo | Lembas bread | Samwise | Lord of the Rings |
| Merry | Mushrooms | Pippin | Lord of the Rings |

```
books_tb = tibble(
  book = c("Harry Potter", "Lord of the Rings"),
  pages = c(865, 1624),
  author = c("J.K. Rowling", "J.R.R. Tolkein")
)
kable(books_tb)
```

| book | pages | author |
|------|-------|--------|
| Harry Potter | 865 | J.K. Rowling |
| Lord of the Rings | 1624 | J.R.R. Tolkein |

We want to select characters from books with greater than 1,000 pages. What steps do we need to follow?

```
char_books = characters_tb %>% left_join(books_tb, by = "book")
```

```
## Warning: `chr_along()` is deprecated as of rlang 0.2.0.
## This warning is displayed once per session.
```

```
kable(char_books)
```

| name | fav_food | friends | book | pages | author |
|------|----------|---------|------|-------|--------|
| Harry | Pumpkin pasties | Hermione, Ron | Harry Potter | 865 | J.K. Rowling |
| Ron | Chocoloate Frogs | Harry, Hermione | Harry Potter | 865 | J.K. Rowling |
| Frodo | Lembas bread | Samwise | Lord of the Rings | 1624 | J.R.R. Tolkein |
| Merry | Mushrooms | Pippin | Lord of the Rings | 1624 | J.R.R. Tolkein |

Why the duplicate entries in the book column? What would happen if we put books_tb on the left-hand side and do a join?

```
char_books2 = books_tb %>% left_join(characters_tb, by = "book")
kable(char_books2)
```

| book | pages | author | name | fav_food | friends |
|------|-------|--------|------|----------|---------|
| Harry Potter | 865 | J.K. Rowling | Harry | Pumpkin pasties | Hermione, Ron |
| Harry Potter | 865 | J.K. Rowling | Ron | Chocoloate Frogs | Harry, Hermione |
| Lord of the Rings | 1624 | J.R.R. Tolkein | Frodo | Lembas bread | Samwise |
| Lord of the Rings | 1624 | J.R.R. Tolkein | Merry | Mushrooms | Pippin |

What would happen if we added different volumes of each book, and repeat the joins?

```
booksv_tb = tibble(
  book = c("Harry Potter", "Harry Potter", "Lord of the Rings", "Lord of the Rings"),
  volume = c("And the sorcerer's stone", "And the chamber of secrets", "The fellowship of the ring", "T
  pages = c(382, 865, 732, 1624),
  author = c("J.K. Rowling", "J.K. Rowling", "J.R.R. Tolkein", "J.R.R. Tolkein")
)
kable(booksv_tb)
```

| book | volume | pages | author |
|------|--------|-------|--------|
| Harry Potter | And the sorcerer's stone | 382 | J.K. Rowling |
| Harry Potter | And the chamber of secrets | 865 | J.K. Rowling |
| Lord of the Rings | The fellowship of the ring | 732 | J.R.R. Tolkein |
| Lord of the Rings | The two towers | 1624 | J.R.R. Tolkein |

```
char_books3 = characters_tb %>% left_join(booksv_tb, by = "book")
kable(char_books3)
```

| name | fav_food | friends | book | volume | pages | author |
|------|----------|---------|------|--------|-------|--------|
| Harry | Pumpkin pasties | Hermione, Ron | Harry Potter | And the sorcerer's stone | 382 | J.K. Rowling |
| Harry | Pumpkin pasties | Hermione, Ron | Harry Potter | And the chamber of secrets | 865 | J.K. Rowling |
| Ron | Chocoloate Frogs | Harry, Hermione | Harry Potter | And the sorcerer's stone | 382 | J.K. Rowling |
| Ron | Chocoloate Frogs | Harry, Hermione | Harry Potter | And the chamber of secrets | 865 | J.K. Rowling |
| Frodo | Lembas bread | Samwise | Lord of the Rings | The fellowship of the ring | 732 | J.R.R. Tolkein |
| Frodo | Lembas bread | Samwise | Lord of the Rings | The two towers | 1624 | J.R.R. Tolkein |
| Merry | Mushrooms | Pippin | Lord of the Rings | The fellowship of the ring | 732 | J.R.R. Tolkein |
| Merry | Mushrooms | Pippin | Lord of the Rings | The two towers | 1624 | J.R.R. Tolkein |

What would happen if we have a character whose book is not in the book tibble?

```
characters2_tb = tibble(
  name = c("Harry", "Ron", "Frodo", "Merry", "Ender"),
  fav_food = c("Pumpkin pasties", "Chocoloate Frogs", "Lembas bread", "Mushrooms", "Unknown"),
  friends = c("Hermione, Ron", "Harry, Hermione", "Samwise", "Pippin", "Bean"),
  book = c("Harry Potter", "Harry Potter", "Lord of the Rings", "Lord of the Rings", "Ender's Game")
)
kable(characters2_tb)
```

| name  | fav_food          | friends          | book              |
|-------|-------------------|------------------|-------------------|
| Harry | Pumpkin pasties   | Hermione, Ron    | Harry Potter      |
| Ron   | Chocoloate Frogs  | Harry, Hermione  | Harry Potter      |
| Frodo | Lembas bread      | Samwise          | Lord of the Rings |
| Merry | Mushrooms         | Pippin           | Lord of the Rings |
| Ender | Unknown           | Bean             | Ender's Game      |

```
char_books4 = characters2_tb %>% left_join(books_tb, by = "book")
kable(char_books4)
```

| name  | fav_food         | friends         | book              | pages | author         |
|-------|------------------|-----------------|-------------------|-------|----------------|
| Harry | Pumpkin pasties  | Hermione, Ron   | Harry Potter      | 865   | J.K. Rowling   |
| Ron   | Chocolate Frogs  | Harry, Hermione | Harry Potter      | 865   | J.K. Rowling   |
| Frodo | Lembas bread     | Samwise         | Lord of the Rings | 1624  | J.R.R. Tolkein |
| Merry | Mushrooms        | Pippin          | Lord of the Rings | 1624  | J.R.R. Tolkein |
| Ender | Unknown          | Bean            | Ender's Game      | NA    | NA             |

What if we perform a right_join with the new characters tibble?

```
char_books5 = characters2_tb %>% right_join(books_tb, by = "book")
kable(char_books5)
```

| name  | fav_food         | friends         | book              | pages | author         |
|-------|------------------|-----------------|-------------------|-------|----------------|
| Harry | Pumpkin pasties  | Hermione, Ron   | Harry Potter      | 865   | J.K. Rowling   |
| Ron   | Chocolate Frogs  | Harry, Hermione | Harry Potter      | 865   | J.K. Rowling   |
| Frodo | Lembas bread     | Samwise         | Lord of the Rings | 1624  | J.R.R. Tolkein |
| Merry | Mushrooms        | Pippin          | Lord of the Rings | 1624  | J.R.R. Tolkein |

Suppose we perform a right_join with a book that doesn't have a corresponding character?

```
booksv2_tb = tibble(
  book = c("Harry Potter", "Harry Potter", "Lord of the Rings", "Lord of the Rings", "Mistborn"),
  volume = c("And the sorcerer's stone", "And the chamber of secrets", "The fellowship of the ring", "T
  pages = c(382, 865, 732, 1624, 412),
  author = c("J.K. Rowling", "J.K. Rowling", "J.R.R. Tolkein", "J.R.R. Tolkein", "Brandon Sanderson")
)
kable(booksv2_tb)
```

| book              | volume                     | pages | author            |
|-------------------|----------------------------|-------|-------------------|
| Harry Potter      | And the sorcerer's stone   | 382   | J.K. Rowling      |
| Harry Potter      | And the chamber of secrets | 865   | J.K. Rowling      |
| Lord of the Rings | The fellowship of the ring | 732   | J.R.R. Tolkein    |
| Lord of the Rings | The two towers             | 1624  | J.R.R. Tolkein    |
| Mistborn          | The final empire           | 412   | Brandon Sanderson |

```
char_books6 = characters2_tb %>% right_join(booksv2_tb, by = "book")
kable(char_books6)
```

| name | fav_food | friends | book | volume | pages | author |
|---|---|---|---|---|---|---|
| Harry | Pumpkin pasties | Hermione, Ron | Harry Potter | And the sorcerer's stone | 382 | J.K. Rowling |
| Ron | Chocoloate Frogs | Harry, Hermione | Harry Potter | And the sorcerer's stone | 382 | J.K. Rowling |
| Harry | Pumpkin pasties | Hermione, Ron | Harry Potter | And the chamber of secrets | 865 | J.K. Rowling |
| Ron | Chocoloate Frogs | Harry, Hermione | Harry Potter | And the chamber of secrets | 865 | J.K. Rowling |
| Frodo | Lembas bread | Samwise | Lord of the Rings | The fellowship of the ring | 732 | J.R.R. Tolkein |
| Merry | Mushrooms | Pippin | Lord of the Rings | The fellowship of the ring | 732 | J.R.R. Tolkein |
| Frodo | Lembas bread | Samwise | Lord of the Rings | The two towers | 1624 | J.R.R. Tolkein |
| Merry | Mushrooms | Pippin | Lord of the Rings | The two towers | 1624 | J.R.R. Tolkein |
| NA | NA | NA | Mistborn | The final empire | 412 | Brandon Sanders |

What do we expect from a full join?

```
char_books6 = characters2_tb %>% full_join(booksv2_tb, by = "book")
kable(char_books6)
```

| name | fav_food | friends | book | volume | pages | author |
|---|---|---|---|---|---|---|
| Harry | Pumpkin pasties | Hermione, Ron | Harry Potter | And the sorcerer's stone | 382 | J.K. Rowling |
| Harry | Pumpkin pasties | Hermione, Ron | Harry Potter | And the chamber of secrets | 865 | J.K. Rowling |
| Ron | Chocoloate Frogs | Harry, Hermione | Harry Potter | And the sorcerer's stone | 382 | J.K. Rowling |
| Ron | Chocoloate Frogs | Harry, Hermione | Harry Potter | And the chamber of secrets | 865 | J.K. Rowling |
| Frodo | Lembas bread | Samwise | Lord of the Rings | The fellowship of the ring | 732 | J.R.R. Tolkein |
| Frodo | Lembas bread | Samwise | Lord of the Rings | The two towers | 1624 | J.R.R. Tolkein |
| Merry | Mushrooms | Pippin | Lord of the Rings | The fellowship of the ring | 732 | J.R.R. Tolkein |
| Merry | Mushrooms | Pippin | Lord of the Rings | The two towers | 1624 | J.R.R. Tolkein |
| Ender | Unknown | Bean | Ender's Game | NA | NA | NA |
| NA | NA | NA | Mistborn | The final empire | 412 | Brandon Sanders |

What do we expect from an inner join?

```
char_books7 = characters2_tb %>% inner_join(booksv2_tb, by = "book")
kable(char_books7)
```

| name | fav_food | friends | book | volume | pages | author |
|---|---|---|---|---|---|---|
| Harry | Pumpkin pasties | Hermione, Ron | Harry Potter | And the sorcerer's stone | 382 | J.K. Rowling |
| Harry | Pumpkin pasties | Hermione, Ron | Harry Potter | And the chamber of secrets | 865 | J.K. Rowling |
| Ron | Chocoloate Frogs | Harry, Hermione | Harry Potter | And the sorcerer's stone | 382 | J.K. Rowling |
| Ron | Chocoloate Frogs | Harry, Hermione | Harry Potter | And the chamber of secrets | 865 | J.K. Rowling |
| Frodo | Lembas bread | Samwise | Lord of the Rings | The fellowship of the ring | 732 | J.R.R. Tolkein |
| Frodo | Lembas bread | Samwise | Lord of the Rings | The two towers | 1624 | J.R.R. Tolkein |
| Merry | Mushrooms | Pippin | Lord of the Rings | The fellowship of the ring | 732 | J.R.R. Tolkein |
| Merry | Mushrooms | Pippin | Lord of the Rings | The two towers | 1624 | J.R.R. Tolkein |

```
```