

Making data sets tidy: tidyr

The package tidyr has basically 2 functions:

- `gather()` — make a wide table long
- `spread()` — make a long table wide

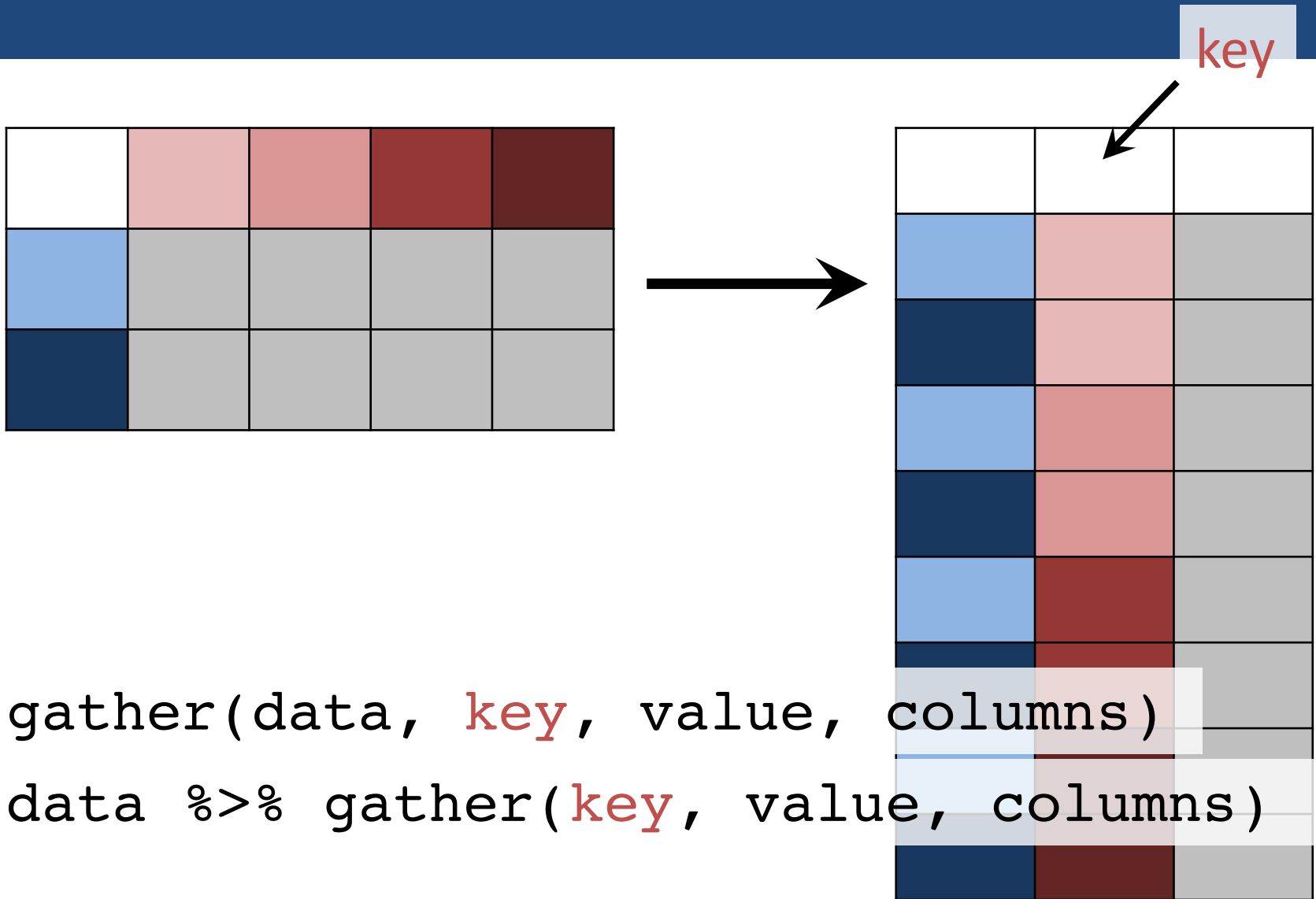
gather () : make wide table long




```
gather(data, key, value, columns)
```

```
data %>% gather(key, value, columns)
```

gather(): make wide table long



gather(): make wide table long

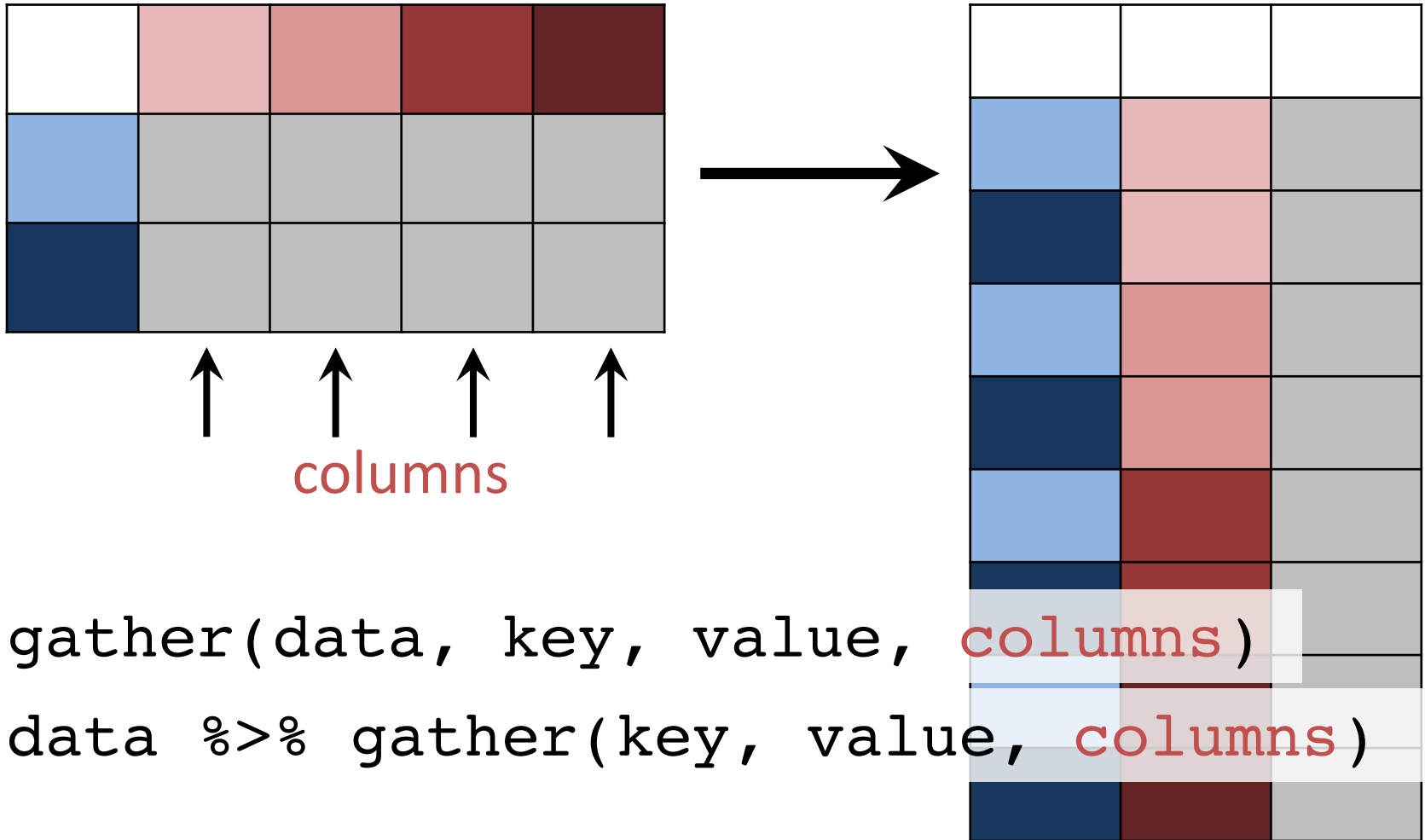
value




```
gather(data, key, value, columns)
```

```
data %>% gather(key, value, columns)
```

gather(): make wide table long



Example: Let's recreate the Sitka data from a wide table

```
> head(sitka_wide)
  tree treat t152 t174 t201 t227 t258
1     1 ozone 4.51 4.98 5.41 5.90 6.15
2     2 ozone 4.24 4.20 4.68 4.92 4.96
3     3 ozone 3.98 4.36 4.79 4.99 5.03
4     4 ozone 4.36 4.77 5.10 5.30 5.36
5     5 ozone 4.34 4.95 5.42 5.97 6.28
6     6 ozone 4.59 5.08 5.36 5.76 6.00
```

Example: Let's recreate the Sitka data from a wide table

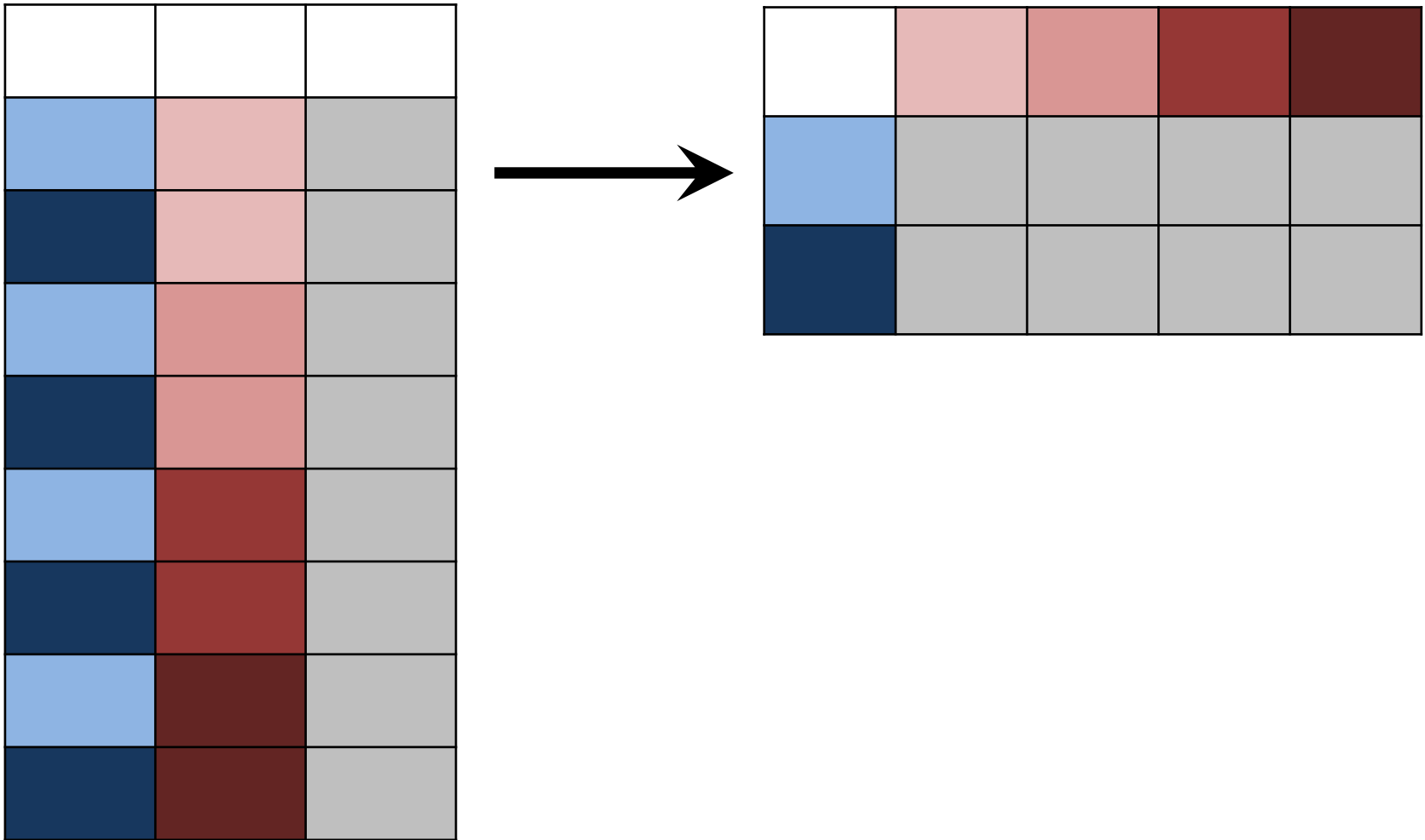
```
> head(sitka_wide)
  tree treat t152 t174 t201 t227 t258
1     1 ozone 4.51 4.98 5.41 5.90 6.15
2     2 ozone 4.24 4.20 4.68 4.92 4.96
3     3 ozone 3.98 4.36 4.79 4.99 5.03
4     4 ozone 4.36 4.77 5.10 5.30 5.36
5     5 ozone 4.34 4.95 5.42 5.97 6.28
6     6 ozone 4.59 5.08 5.36 5.76 6.00
```

```
sitka_wide %>% gather(time, size, t152:t258)
```


Example: Let's recreate the Sitka data from a wide table

```
> sitka_wide %>% gather(time, size, t152:t258) %>%  
                                     head()  
  tree treat time size  
1     1 ozone t152 4.51  
2     2 ozone t152 4.24  
3     3 ozone t152 3.98  
4     4 ozone t152 4.36  
5     5 ozone t152 4.34  
6     6 ozone t152 4.59
```

`spread()`: make long table wide



spread(): make long table wide

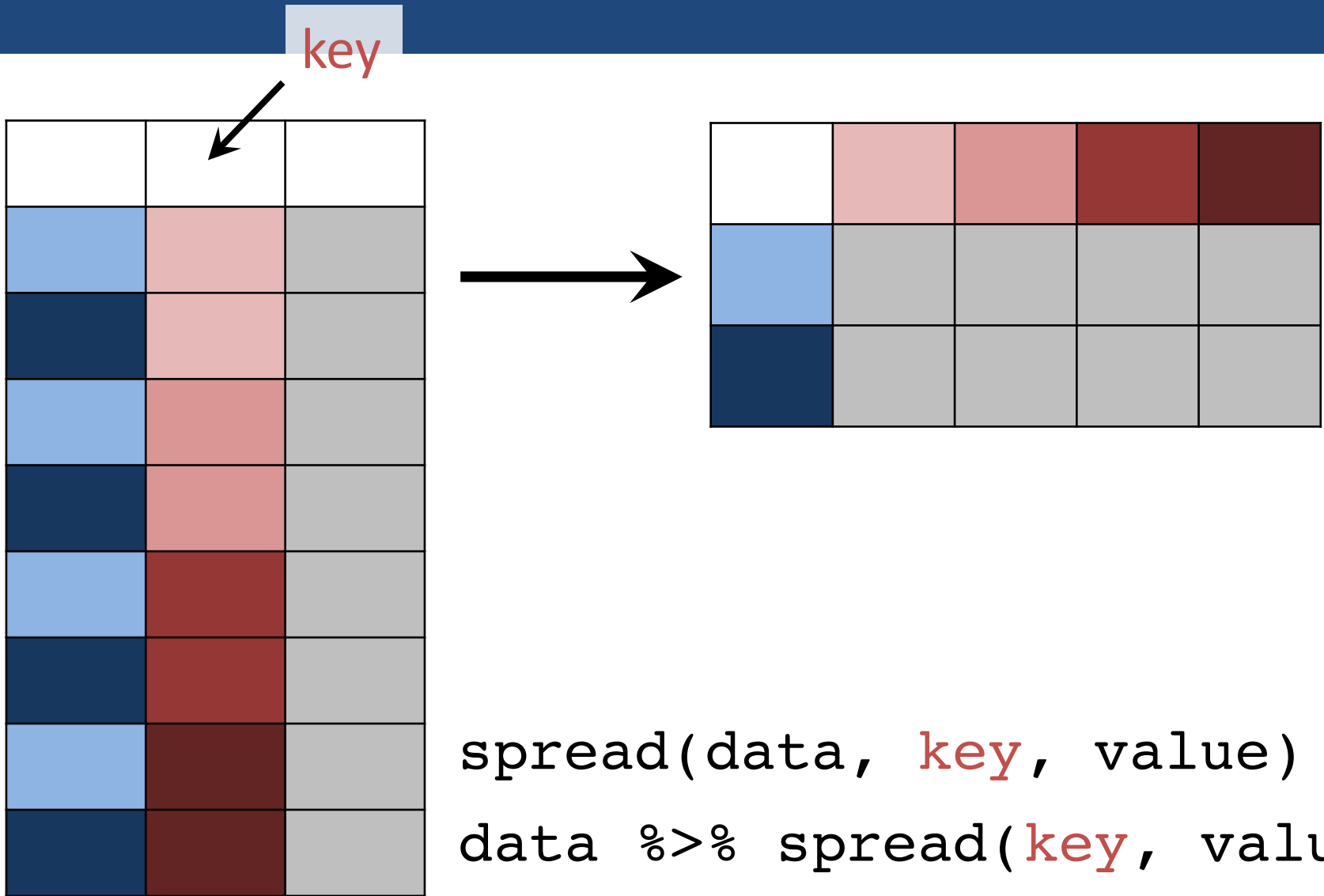
light blue	light red	gray
dark blue	light red	gray
light blue	light red	gray
dark blue	light red	gray
light blue	dark red	gray
dark blue	dark red	gray
light blue	dark red	gray
dark blue	dark red	gray



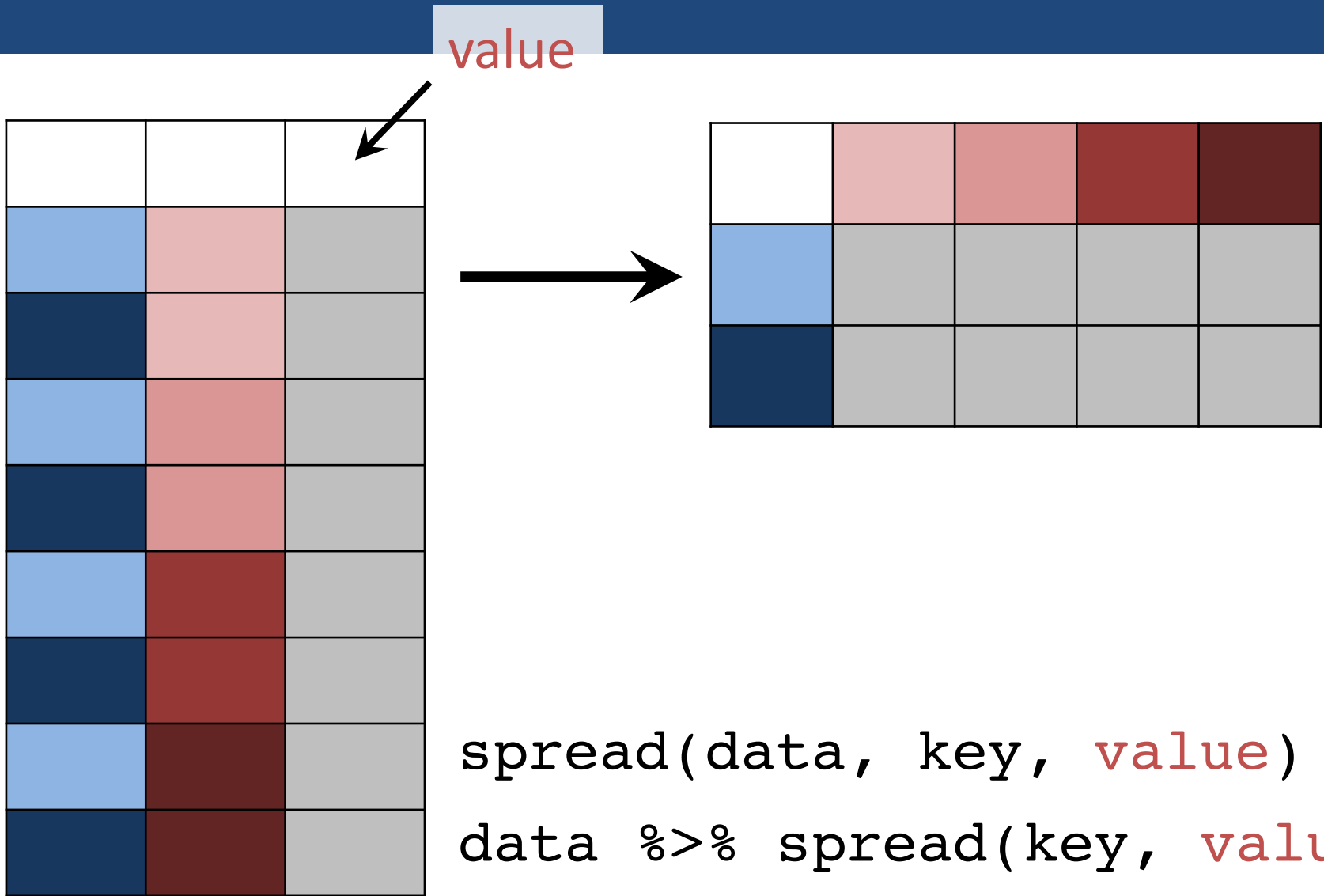
	light red	light red	dark red	dark red
light blue	gray	gray	gray	gray
dark blue	gray	gray	gray	gray

```
spread(data, key, value)  
data %>% spread(key, value)
```

spread(): make long table wide



spread(): make long table wide



Example: Let's turn the Sitka data into a wide table

```
> head(Sitka)
  size Time tree treat
1 4.51  152   1 ozone
2 4.98  174   1 ozone
3 5.41  201   1 ozone
4 5.90  227   1 ozone
5 6.15  258   1 ozone
6 4.24  152   2 ozone
```

```
Sitka %>% spread(Time, size)
```

Example: Let's turn the Sitka data into a wide table

```
> Sitka %>% spread(Time, size) %>% head()
  tree treat   152   174   201   227   258
1     1 ozone 4.51 4.98 5.41 5.90 6.15
2     2 ozone 4.24 4.20 4.68 4.92 4.96
3     3 ozone 3.98 4.36 4.79 4.99 5.03
4     4 ozone 4.36 4.77 5.10 5.30 5.36
5     5 ozone 4.34 4.95 5.42 5.97 6.28
6     6 ozone 4.59 5.08 5.36 5.76 6.00
```