

```
(base) Vasu's MacBook Pro:Desktop vasugoe1$ r
```

```
R version 3.6.1 (2019-07-05) -- "Action of the Toes"
Copyright (C) 2019 The R Foundation for Statistical Computing
Platform: x86_64-apple-darwin15.6.0 (64-bit)
```

R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.

Natural language support but running in an English locale

R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.

```
> path <- system.file('extdata', package='dslabs')
> list.files(path)
[1] "2010_bigfive_regents.xls"
[2] "carbon_emissions.csv"
[3] "fertility-two-countries-example.csv"
[4] "HRlist2.txt"
[5] "life-expectancy-and-fertility-two-countries-example.csv"
[6] "murders.csv"
[7] "olive.csv"
[8] "RD-Mortality-Report_2015-18-180531.pdf"
[9] "ssa-death-probability.csv"
> filename <- 'fertility-two-countries-example.csv'
> filepath <- file.path(path, filename)
> filepath
[1] "/Library/Frameworks/R.framework/Versions/3.6/Resources/library/dslabs/extdata/fertility-two-countries-example.csv"
> library(dslabs)
> library(tidyverse)
— Attaching packages — tidyverse 1.2.1 —
✓ ggplot2 3.2.1      ✓ purrr  0.3.2
✓ tibble  2.1.3      ✓ dplyr  0.8.3
✓ tidyr   0.8.3      ✓ stringr 1.4.0
✓ readr   1.3.1      ✓ forcats 0.4.0
— Conflicts — tidyverse_conflicts() —
✖ dplyr::filter() masks stats::filter()
✖ dplyr::lag()     masks stats::lag()
> wide_data <- read_csv(filepath)
Parsed with column specification:
cols(
  .default = col_double(),
  country = col_character()
)
See spec(...) for full column specifications.
> wide_data
# A tibble: 2 x 57
  country `1960` `1961` `1962` `1963` `1964` `1965` `1966` `1967` `1968` `1969`
  <chr>    <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl>
1 Germany  2.41  2.44  2.47  2.49  2.49  2.48  2.44  2.37  2.28  2.17
2 South ... 6.16  5.99  5.79  5.57  5.36  5.16  4.99  4.85  4.73  4.62
# ... with 46 more variables: `1970` <dbl>, `1971` <dbl>, `1972` <dbl>,
# `1973` <dbl>, `1974` <dbl>, `1975` <dbl>, `1976` <dbl>, `1977` <dbl>,
# `1978` <dbl>, `1979` <dbl>, `1980` <dbl>, `1981` <dbl>, `1982` <dbl>,
# `1983` <dbl>, `1984` <dbl>, `1985` <dbl>, `1986` <dbl>, `1987` <dbl>,
# `1988` <dbl>, `1989` <dbl>, `1990` <dbl>, `1991` <dbl>, `1992` <dbl>,
# `1993` <dbl>, `1994` <dbl>, `1995` <dbl>, `1996` <dbl>, `1997` <dbl>,
```

```

# `1998` <dbl>, `1999` <dbl>, `2000` <dbl>, `2001` <dbl>, `2002` <dbl>,
# `2003` <dbl>, `2004` <dbl>, `2005` <dbl>, `2006` <dbl>, `2007` <dbl>,
# `2008` <dbl>, `2009` <dbl>, `2010` <dbl>, `2011` <dbl>, `2012` <dbl>,
# `2013` <dbl>, `2014` <dbl>, `2015` <dbl>
> data("gapminder")
tidy_data <- gapminder %>%
  filter(country %in% c("South Korea", "Germany")) %>%
  select(country, year, fertility)
> tidy_data <- gapminder %>%
+   filter(country %in% c("South Korea", "Germany")) %>%
+   select(country, year, fertility)
> head(tidy_data)
  country year fertility
1   Germany 1960     2.41
2 South Korea 1960     6.16
3   Germany 1961     2.44
4 South Korea 1961     5.99
5   Germany 1962     2.47
6 South Korea 1962     5.79
> new_tidy_data <- wide_data %>% gather(year, fertility, -country, convert=T)
> new_tidy_data
# A tibble: 112 x 3
  country    year fertility
  <chr>    <int>    <dbl>
1 Germany    1960     2.41
2 South Korea 1960     6.16
3 Germany    1961     2.44
4 South Korea 1961     5.99
5 Germany    1962     2.47
6 South Korea 1962     5.79
7 Germany    1963     2.49
8 South Korea 1963     5.57
9 Germany    1964     2.49
10 South Korea 1964     5.36
# ... with 102 more rows
> new_tidy_data %>% ggplot(aes(year, fertility, color=country)) + geom_point()
> new_wide_data <- new_tidy_data %>% spread(year, fertility)
> new_wide_data
# A tibble: 2 x 57
  country `1960` `1961` `1962` `1963` `1964` `1965` `1966` `1967` `1968` `1969`
  <chr>    <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl>
1 Germany    2.41  2.44  2.47  2.49  2.49  2.48  2.44  2.37  2.28  2.17
2 South ...    6.16  5.99  5.79  5.57  5.36  5.16  4.99  4.85  4.73  4.62
# ... with 46 more variables: `1970` <dbl>, `1971` <dbl>, `1972` <dbl>,
# `1973` <dbl>, `1974` <dbl>, `1975` <dbl>, `1976` <dbl>, `1977` <dbl>,
# `1978` <dbl>, `1979` <dbl>, `1980` <dbl>, `1981` <dbl>, `1982` <dbl>,
# `1983` <dbl>, `1984` <dbl>, `1985` <dbl>, `1986` <dbl>, `1987` <dbl>,
# `1988` <dbl>, `1989` <dbl>, `1990` <dbl>, `1991` <dbl>, `1992` <dbl>,
# `1993` <dbl>, `1994` <dbl>, `1995` <dbl>, `1996` <dbl>, `1997` <dbl>,
# `1998` <dbl>, `1999` <dbl>, `2000` <dbl>, `2001` <dbl>, `2002` <dbl>,
# `2003` <dbl>, `2004` <dbl>, `2005` <dbl>, `2006` <dbl>, `2007` <dbl>,
# `2008` <dbl>, `2009` <dbl>, `2010` <dbl>, `2011` <dbl>, `2012` <dbl>,
# `2013` <dbl>, `2014` <dbl>, `2015` <dbl>
>

```