Creating Tables for R Markdown Reports

The basics, and making them better with kableExtra

Izzy Johnson

8 April 2021

Overview:

- Making tables in R
- R Markdown
- KableExtra Package
- Additional Resources

company	cocoa percent	company location	cocoa percent above 70%
Arete	70%	U.S.A.	0
Arete	70%	U.S.A.	0
Ocho	100%	New Zealand	1
Palette de Bine	72%	Canada	1
Sirene	73%	Canada	1
Smooth Chocolator, The	67%	Australia	0
Soma	62%	Canada	0
Soma	70%	Canada	0

Making Tables in Base R

```
table(chocolate_data_sm$company)
```

```
##
##                   Arete                    Ocho        Palette de Bine
##                      22                       6                     11
##                  Sirene Smooth Chocolator, The                    Soma
##                      11                      16                     15
```

Making Tables in Base R

```
table(chocolate_data_sm$company_loc,
      chocolate_data_sm$cocoa_pct_above_70)
```

```
##
##                  0  1
##   Australia     14  2
##   Canada        17 20
##   New Zealand    3  3
##   U.S.A.        18  4
```

Making Tables in Base R

```
table(mtcars$vs,
      mtcars$am)
```

```
##
##      0  1
##   0 12  6
##   1  7  7
```

```
table(mtcars$vs,
      mtcars$am,
      dnn = c('VS', 'Transmission: AM'))
```

```
##    Transmission: AM
## VS   0  1
##   0 12  6
##   1  7  7
```

Making Tables in R with dplyr

- Main functions: group_by() and count()
  - arrange(), summarise(), mutate(), and ungroup() are also useful

```
chocolate_data_sm %>%
  group_by(company) %>%
  count()
```

```
## # A tibble: 6 x 2
## # Groups:   company [6]
##   company                     n
##   <chr>                   <int>
## 1 Arete                      22
## 2 Ocho                        6
## 3 Palette de Bine            11
## 4 Sirene                     11
## 5 Smooth Chocolator, The     16
## 6 Soma                       15
```

Making Tables in R with dplyr

- Main functions: group_by() and count()
  - arrange(), summarise(), mutate(), and ungroup() are also useful

```
chocolate_data_sm %>%
  group_by(company) %>%
  count() %>%
  arrange(-n)
```

```
## # A tibble: 6 x 2
## # Groups:   company [6]
##   company                     n
##   <chr>                   <int>
## 1 Arete                      22
## 2 Smooth Chocolator, The     16
## 3 Soma                       15
## 4 Palette de Bine            11
## 5 Sirene                     11
## 6 Ocho                        6
```

Making Tables in R with dplyr

```
chocolate_data_sm %>%
  group_by(company_loc, cocoa_pct_above_70) %>%
  count()
```

```
## # A tibble: 8 x 3
## # Groups:   company_loc, cocoa_pct_above_70 [8]
##   company_loc cocoa_pct_above_70     n
##   <chr>                    <dbl> <int>
## 1 Australia                    0    14
## 2 Australia                    1     2
## 3 Canada                       0    17
## 4 Canada                       1    20
## 5 New Zealand                  0     3
## 6 New Zealand                  1     3
## 7 U.S.A.                       0    18
## 8 U.S.A.                       1     4
```

Making Tables in R with dplyr

```
chocolate_data_sm %>%
  group_by(company_loc, cocoa_pct_above_70) %>%
  count() %>%
  ungroup() %>%
  mutate(percent = (n/sum(n)*100) %>%
           round(2))
```

```
## # A tibble: 8 x 4
##   company_loc cocoa_pct_above_70     n percent
##   <chr>                    <dbl> <int>   <dbl>
## 1 Australia                    0    14   17.3
## 2 Australia                    1     2    2.47
## 3 Canada                       0    17   21.0
## 4 Canada                       1    20   24.7
## 5 New Zealand                  0     3    3.7
## 6 New Zealand                  1     3    3.7
## 7 U.S.A.                       0    18   22.2
## 8 U.S.A.                       1     4    4.94
```

Making Tables in R with dplyr

```
chocolate_data_sm %>%
  group_by(company_loc, cocoa_pct_above_70) %>%
  count() %>%
  mutate(percent = (n/sum(n)*100) %>%
           round(2))
```

```
## # A tibble: 8 x 4
## # Groups:   company_loc, cocoa_pct_above_70 [8]
##   company_loc cocoa_pct_above_70     n percent
##   <chr>                    <dbl> <int>   <dbl>
## 1 Australia                    0    14     100
## 2 Australia                    1     2     100
## 3 Canada                       0    17     100
## 4 Canada                       1    20     100
## 5 New Zealand                  0     3     100
## 6 New Zealand                  1     3     100
## 7 U.S.A.                       0    18     100
## 8 U.S.A.                       1     4     100
```

Making Tables in R with dplyr

```
chocolate_data_sm %>%
  group_by(company_loc, cocoa_pct_above_70) %>%
  count() %>%
  ungroup() %>%
  mutate(percent = (n/sum(n)*100) %>%
           round(2) %>%
           format(nsmall = 2) %>%
           paste0('%'))
```

```
## # A tibble: 8 x 4
##   company_loc cocoa_pct_above_70     n percent
##   <chr>                    <dbl> <int> <chr>
## 1 Australia                    0    14 "17.28%"
## 2 Australia                    1     2 " 2.47%"
## 3 Canada                       0    17 "20.99%"
## 4 Canada                       1    20 "24.69%"
## 5 New Zealand                  0     3 " 3.70%"
## 6 New Zealand                  1     3 " 3.70%"
## 7 U.S.A.                       0    18 "22.22%"
## 8 U.S.A.                       1     4 " 4.94%"
```

Making Tables in R with dplyr

```
mtcars %>%
  group_by(vs,am) %>%
  count() %>%
  ungroup() %>%
  mutate(percent = (n/sum(n)*100) %>%
           round(2) %>%
           format(nsmall = 2) %>%
           paste0('%'))
```

```
## # A tibble: 4 x 4
##      vs    am     n percent
##   <dbl> <dbl> <int> <chr>
## 1     0     0    12 37.50%
## 2     0     1     6 18.75%
## 3     1     0     7 21.88%
## 4     1     1     7 21.88%
```

R Markdown (RMD)

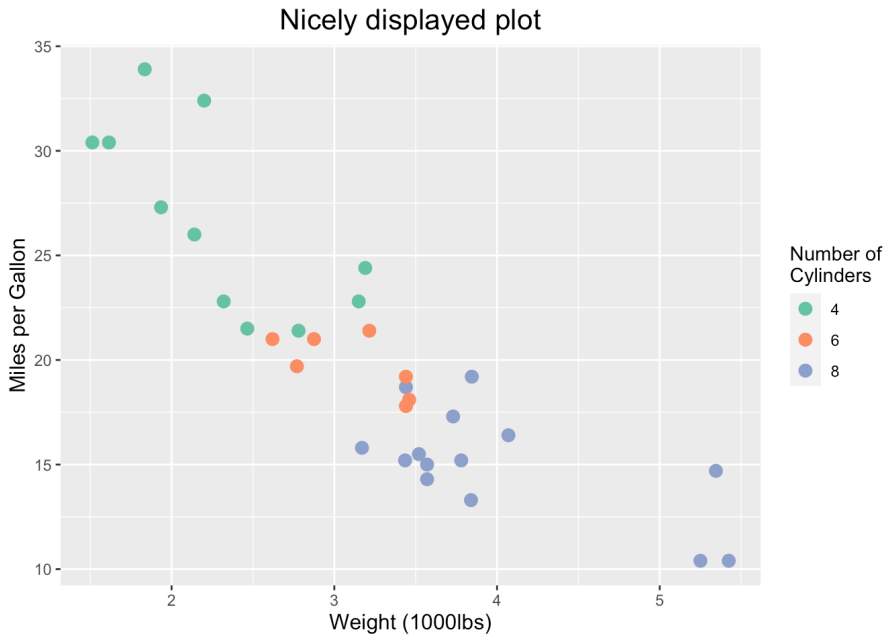
- Great for sending reproducible reports to others
- Main output files are HTML, PDF and Word

Look at How Nice This Looks!

Izzy Johnson

7 April 2021

Look at this ~lovely, code free~ plot:



Need the code? You can display it in the report, or find it in the R Markdown document that generated the report!

(Which should be easy because you left clear, concise comments, right?)

kableExtra Package

Creates nicer looking tables for your reports!

From this

```
# A tibble: 6 x 3
  company                    n percent
  <chr>                  <int> <chr>
1 Arete                     22 "27.16%"
2 Smooth Chocolator, The    16 "19.75%"
3 Soma                      15 "18.52%"
4 Palette de Bine           11 "13.58%"
5 Sirene                    11 "13.58%"
6 Ocho                       6 " 7.41%"
```

To this!

Company	Count	Percent
Arete	22	27.16%
Smooth Chocolator, The	16	19.75%
Soma	15	18.52%
Palette de Bine	11	13.58%
Sirene	11	13.58%
Ocho	6	7.41%

Using the kableExtra Package

To install the kableExtra package:

```
install.packages("kableExtra")
```

To load kableExtra:

```
library(kableExtra)
```

kableExtra: main functions

- kable() function comes from knitr to create tables
    - some parameters I often use:
        - col.names - allows you to rename your column names
        - caption - adds a caption to the top of your table
- kable_styling() allows many different options to modify tables easily*
    - Within this function some things you can do include:
        - style the font
        - Alternate the color of your rows
        - Alter the width and position of your table

** = at least more easily than having to code all of the modifications from scratch. I have still screamed at my computer while using kable_styling()*

Live coding time!

Resources

kableExtra (Package by Hao Zhu)

- An explanation of the main features of the package, written by Hau Zhu: https://cran.r-project.org/web/packages/kableExtra/vignettes/awesome_table_in_html.html
- R Cookbook explanation of kableExtra: https://bookdown.org/yihui/rmarkdown-cookbook/kableextra.html

R Markdown

- Dataquest's R Markdown Guide: https://www.dataquest.io/blog/r-markdown-guide-cheatsheet/
- R Markdown Reference Guide: https://rstudio.com/wp-content/uploads/2015/03/rmarkdown-reference.pdf

- R Markdown Cheatsheet (open R, go to 'Help', 'Cheatsheets', then click 'R Markdown Cheatsheet')

  Chocolate dataset came from Kaggle, "Chocolate Bar Ratings" uploaded by Rachael Tatman