Communication with RMarkdown

Business Science

3/14/2019

Contents

RMarkdown

Is amazing.

What can RMarkdown be used for?

1. HTML Reports & PDF Reports
2. HTML Slide Decks & PowerPoint
3. Interactive Dashboards
4. Books with `bookdown`
5. Websites with `blogdown`

Key Resources

- RMarkdown Website with Gallery
- Key Reference: RMarkdown - The Definitive Guide
- PDF Printing Setup: tinytex

```
# PDF Knitting Setup: https://yihui.name/tinytex/
# install.packages("tinytex")
# tinytex::install_tinytex()
```

How Rmarkdown Works

Header 1

Header 2

Header 3

Working with Text

Free-form text.

Make text **bold**.

Make text *italics*.

Make text ***bold + italics***.

Talk about code - the `tidyverse` is awesome

Unordered List:

- Item 1
- Item 2

Ordered List:

1. First point
2. Second point
3. More points

Tabset

Tab 1

This is Tab 1

Tab 2

This is Tab 2

Images

Figure 1: Business Science Logo

Figure 2: Business Science Logo

Code

Read in data and print to HTML. Notice effect of `df_print: paged` option for HTML.

- Try changing to `df_print: default`, or `kable` or `tibble`. PDF prints normally.
- Try changing `results = "hide"`.

```
bike_orderlines_tbl <- read_rds(path = "../00_data/bike_sales/data_wrangled/bike_orderlines.rds")

bike_orderlines_tbl
```

```
## # A tibble: 15,644 x 13
##    order_date          order_id order_line quantity price total_price model
##    <dttm>                 <dbl>      <dbl>    <dbl> <dbl>       <dbl> <chr>
```

```
##  1 2011-01-07 00:00:00        1          1        1  6070        6070 Jeky~
##  2 2011-01-07 00:00:00        1          2        1  5970        5970 Trig~
##  3 2011-01-10 00:00:00        2          1        1  2770        2770 Beas~
##  4 2011-01-10 00:00:00        2          2        1  5970        5970 Trig~
##  5 2011-01-10 00:00:00        3          1        1 10660       10660 Supe~
##  6 2011-01-10 00:00:00        3          2        1  3200        3200 Jeky~
##  7 2011-01-10 00:00:00        3          3        1 12790       12790 Supe~
##  8 2011-01-10 00:00:00        3          4        1  5330        5330 Supe~
##  9 2011-01-10 00:00:00        3          5        1  1570        1570 Syna~
## 10 2011-01-11 00:00:00        4          1        1  4800        4800 Syna~
## # ... with 15,634 more rows, and 6 more variables: category_1 <chr>,
## #   category_2 <chr>, frame_material <chr>, bikeshop_name <chr>, city <chr>,
## #   state <chr>
```

We can do data manipulations too. Try changing the YAML `code_folding` option from `none` to `hide` to `show`.

```
revenue_by_category_tbl <- bike_orderlines_tbl %>%
  select(category_2, category_1, total_price) %>%

  group_by(category_2, category_1) %>%
  summarise(total_revenue = sum(total_price)) %>%
  ungroup() %>%

  arrange(desc(total_revenue)) %>%
  mutate(category_2 = as_factor(category_2) %>% fct_rev())
```

Plots

Plotting works as expected. Try changin:

- `out.height`, `out.width` and Knitting
- Potential gotcha - Interactive plots (e.g. `plotly`) will not display in PDF

Static Plots:

- Use `ggplot2`.

```
g <- revenue_by_category_tbl %>%
  ggplot(aes(category_2, total_revenue, fill = category_1)) +

  # Geoms
  geom_col() +
  coord_flip() +

  # Formatting
  scale_fill_tq() +
  scale_y_continuous(labels = scales::dollar_format(scale = 1e-6, suffix = "M")) +
  theme_tq() +
  labs(
```

```
    title = "Total Revenue by Category",
    x = "", y = "", fill = ""
  )

g
```

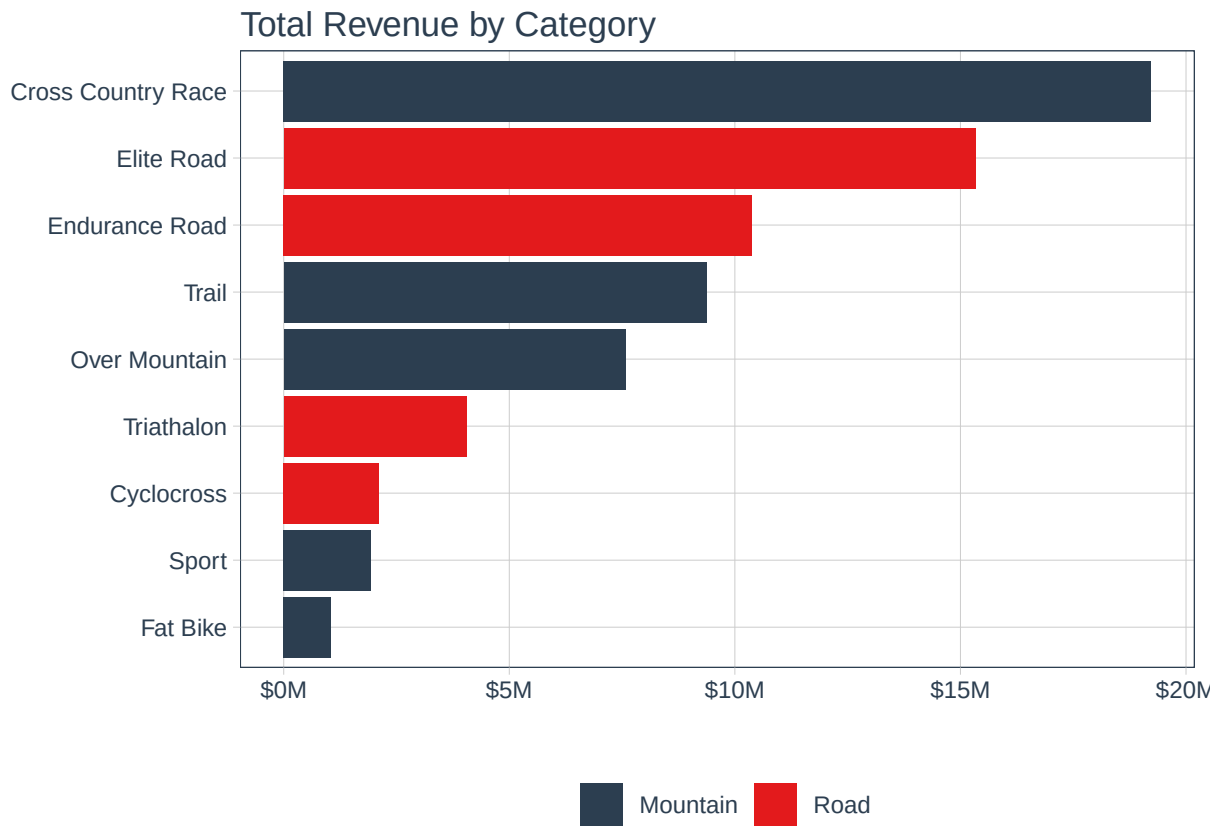

Figure 3: Revenue by Category

Interactive plots:

- Use `ggplotly()`.

```
#ggplotly(g)
```

Tables

Static Tables:

- `knitr` package - `knitr::kable()` - Simple to use, great with PDF
- `gt` package - Not on CRAN yet, but really good for static tables

```
table_formatted_tbl <- revenue_by_category_tbl %>%
  mutate(total_revenue = scales::dollar(total_revenue)) %>%
  rename_all(.funs = ~ str_replace(., "_", " ") %>%
               str_to_title())

table_formatted_tbl %>% knitr::kable()
```

Category 2	Category 1	Total Revenue
Cross Country Race	Mountain	$19,224,630
Elite Road	Road	$15,334,665
Endurance Road	Road	$10,381,060
Trail	Mountain	$9,373,460
Over Mountain	Mountain	$7,571,270
Triathalon	Road	$4,053,750
Cyclocross	Road	$2,108,120
Sport	Mountain	$1,932,755
Fat Bike	Mountain	$1,052,620

Dynamic Tables:

- Can print tables without additional formatting in HTML with the `df_print: paged` option in YAML
- Potential Gotcha: Note that this will not print with format in PDF

```
table_formatted_tbl
```

```
## # A tibble: 9 x 3
##   `Category 2`       `Category 1` `Total Revenue`
##   <fct>              <chr>        <chr>
## 1 Cross Country Race Mountain     $19,224,630
## 2 Elite Road         Road         $15,334,665
## 3 Endurance Road     Road         $10,381,060
## 4 Trail              Mountain     $9,373,460
## 5 Over Mountain      Mountain     $7,571,270
## 6 Triathalon         Road         $4,053,750
## 7 Cyclocross         Road         $2,108,120
## 8 Sport              Mountain     $1,932,755
## 9 Fat Bike           Mountain     $1,052,620
```

Footnotes

This is some text with a Footnote[1]. This is a second Footnote[2].

[1] Citation for Footnote 1
[2] Citatin for Footnote 2