Reporting with RMarkdown

Joschka Schwarz

9/15/2020

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("tintex")
# 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: NIT Logo

Figure 2: NIT 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 data
bikes_tbl      <- readRDS("C:/Users/rueta/Desktop/Data-Science-Products/data/bikes_tbl.rds")
bikeshops_tbl  <- readRDS("C:/Users/rueta/Desktop/Data-Science-Products/data/bikeshops_tbl.rds")
orderlines_tbl <- readRDS("C:/Users/rueta/Desktop/Data-Science-Products/data/orderlines_tbl.rds")

bike_orderlines_tbl <- orderlines_tbl %>%
    left_join(bikes_tbl,     by = c("product_id" = "bike_id")) %>%
    left_join(bikeshops_tbl, by = c("customer_id" = "bikeshop_id")) %>%
    mutate(total_price = price_euro * quantity)

bike_orderlines_tbl
```

```
## # A tibble: 15,644 x 23
##    order_id order_line order_date          customer_id product_id quantity model
##       <dbl>      <dbl> <dttm>                    <dbl>      <dbl>    <dbl> <chr>
##  1        1          1 2015-01-07 00:00:00           2       2681        1 Spec~
##  2        1          2 2015-01-07 00:00:00           2       2411        1 Ulti~
##  3        2          1 2015-01-10 00:00:00          10       2629        1 Neur~
##  4        2          2 2015-01-10 00:00:00          10       2137        1 Spee~
##  5        3          1 2015-01-10 00:00:00           6       2367        1 Stit~
##  6        3          2 2015-01-10 00:00:00           6       1973        1 Road~
##  7        3          3 2015-01-10 00:00:00           6       2422        1 Spee~
##  8        3          4 2015-01-10 00:00:00           6       2655        1 Infl~
##  9        3          5 2015-01-10 00:00:00           6       2247        1 Torq~
## 10        4          1 2015-01-11 00:00:00          22       2408        1 Ulti~
## # ... with 15,634 more rows, and 16 more variables: year <dbl>,
## #   frame_material <chr>, weight <dbl>, price_euro <dbl>, category_1 <chr>,
## #   category_2 <chr>, category_3 <chr>, gender <chr>, description <chr>,
## #   url <chr>, name <chr>, city <chr>, state <chr>, lat <dbl>, lng <dbl>,
## #   total_price <dbl>
```

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

```
sales_by_category_tbl <- bike_orderlines_tbl %>%
  dplyr::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())
```

PlotsN

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 <- sales_by_category_tbl %>%
  ggplot(aes(category_2, total_revenue, fill = category_1)) +

  # Geoms
  geom_col() +
  coord_flip() +
```

```
  # Formatting
  labs(
    title = "Total Revenue by Category",
    x = "", y = "", fill = ""
  )

g
```

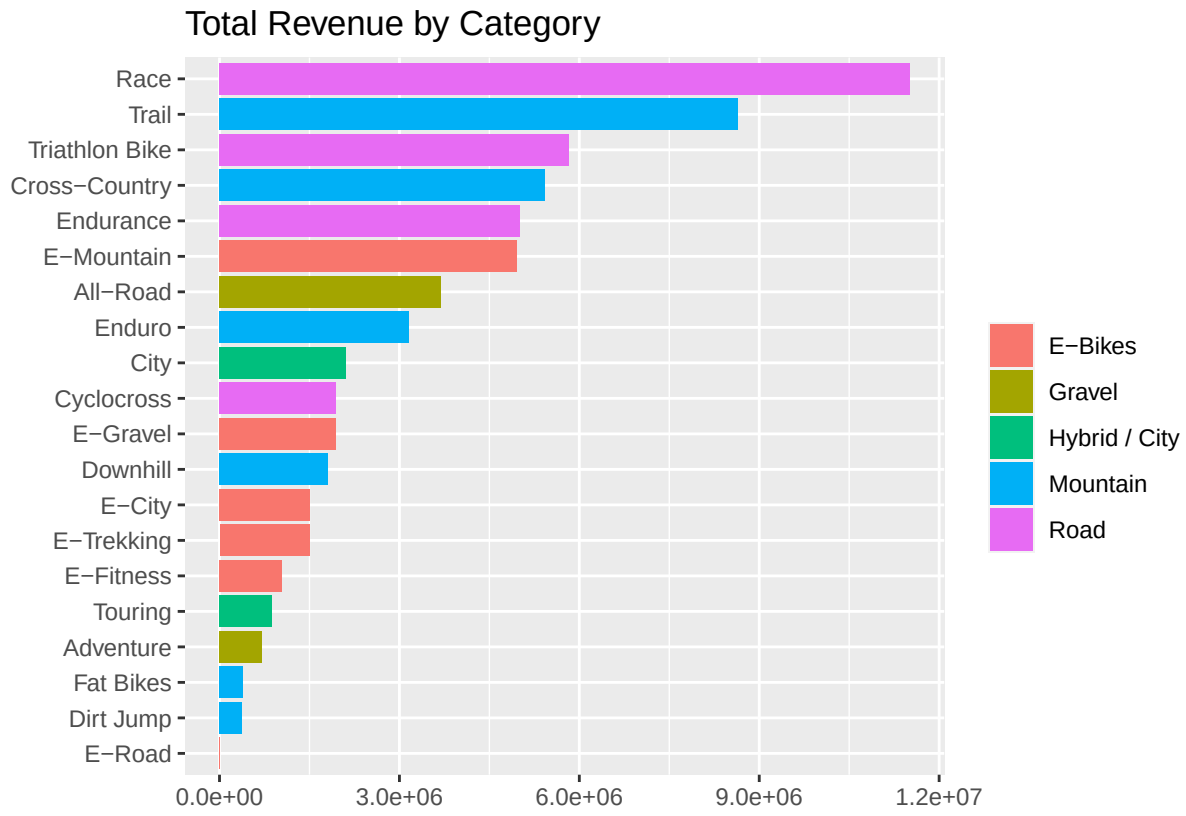

Figure 3: Revenue by Category

Interactive plots:

- Use `ggplotly()`.

```
#ggplotly(g)
```

Tables

Static Tables:

- `knitr` package - `knitr::kable()` - Simple to use, great with PDF
- `gt` package - Really good for static tables

```
table_formatted_tbl <- sales_by_category_tbl %>%
  rename_all(.funs = ~ str_replace(., "_", " ") %>%
               str_to_title())

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

Category 2	Category 1	Total Revenue
Race	Road	11509156
Trail	Mountain	8644966
Triathlon Bike	Road	5831716
Cross-Country	Mountain	5421144
Endurance	Road	5013423
E-Mountain	E-Bikes	4962946
All-Road	Gravel	3697923
Enduro	Mountain	3156837
City	Hybrid / City	2115482
Cyclocross	Road	1940532
E-Gravel	E-Bikes	1936489
Downhill	Mountain	1803970
E-City	E-Bikes	1509096
E-Trekking	E-Bikes	1500894
E-Fitness	E-Bikes	1039996
Touring	Hybrid / City	877736
Adventure	Gravel	702007
Fat Bikes	Mountain	391654
Dirt Jump	Mountain	371922
E-Road	E-Bikes	2919

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: 20 x 3
##    `Category 2`   `Category 1`  `Total Revenue`
##    <fct>          <chr>                   <dbl>
##  1 Race           Road                 11509156
##  2 Trail          Mountain              8644966
##  3 Triathlon Bike Road                  5831716
##  4 Cross-Country  Mountain              5421144
##  5 Endurance      Road                  5013423
##  6 E-Mountain     E-Bikes               4962946
##  7 All-Road       Gravel                3697923
##  8 Enduro         Mountain              3156837
##  9 City           Hybrid / City         2115482
## 10 Cyclocross     Road                  1940532
## 11 E-Gravel       E-Bikes               1936489
## 12 Downhill       Mountain              1803970
```

```
## 13 E-City          E-Bikes              1509096
## 14 E-Trekking      E-Bikes              1500894
## 15 E-Fitness       E-Bikes              1039996
## 16 Touring         Hybrid / City         877736
## 17 Adventure       Gravel                702007
## 18 Fat Bikes       Mountain              391654
## 19 Dirt Jump       Mountain              371922
## 20 E-Road          E-Bikes                 2919
```

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