Authoring Report with R Markdown

An Chu

July 08, 2018

Introduction

This is an example of using `RStudio`, `knitr` and `R Markdown` to author data analysis reporting.

The definitive guide to R Markdown can be found here:

- https://bookdown.org/yihui/rmarkdown/html-document.html

For a quick reference, see this cheatsheet:

- https://www.rstudio.com/wp-content/uploads/2015/02/rmarkdown-cheatsheet.pdf

Inspect directory

Find out the current directory:

```
getwd()
```

```
## [1] "/home/anchu/Documents/r-courses/rmarkdown_intro"
```

Check folders and files in the data directory:

```
dir(path = "~/Documents/r-courses/data", recursive = TRUE)
```

```
##  [1] "data_import/airline-safety.csv"
##  [2] "data_import/Child_Data.sav"
##  [3] "data_import/iris.sas7bdat"
##  [4] "data_import/Milk_Production.dta"
##  [5] "data_import/mtcars1.csv"
##  [6] "data_import/mtcars2.csv"
##  [7] "data_import/mtcars3.csv"
##  [8] "data_import/no_headings.csv"
##  [9] "data_import/Rotten_Tomatoes_100_Movies.xlsx"
## [10] "data_manipulation/01-data.R"
## [11] "data_manipulation/02-one-table-verbs.R"
## [12] "data_manipulation/03-pipelines.R"
## [13] "data_manipulation/04-two-table-verbs.R"
## [14] "data_manipulation/air-quality.csv"
## [15] "data_manipulation/hanoi_weather_history.csv"
## [16] "data_tidying/airports_gps.csv"
## [17] "data_tidying/data_tidying.R"
## [18] "data_tidying/diemthi_tslop10.csv"
## [19] "data_tidying/forest_fires.csv"
## [20] "data_tidying/hanoi_temp_wideform.csv"
```

```
## [21] "data_tidying/oecd_teacher_salary.csv"
## [22] "data_tidying/sales_kpi.csv"
## [23] "data_tidying/so3.csv"
## [24] "data_visualisation/ggplot2_vis.R"
## [25] "foundations/foundations.R"
## [26] "rmarkdown-intro/demo-markdown.html"
## [27] "rmarkdown-intro/rmarkdown-homepage.png"
```

Import data

Import `airline-safety.csv`:

```
library(readr)
airline_safety <- read_csv("~/Documents/r-courses/data/data_import/airline-safety.csv")
```

```
## Parsed with column specification:
## cols(
##   airline = col_character(),
##   avail_seat_km_per_week = col_double(),
##   incidents_85_99 = col_integer(),
##   fatal_accidents_85_99 = col_integer(),
##   fatalities_85_99 = col_integer(),
##   incidents_00_14 = col_integer(),
##   fatal_accidents_00_14 = col_integer(),
##   fatalities_00_14 = col_integer()
## )
```

Check data structure:

```
str(airline_safety, give.attr = FALSE)
```

```
## Classes 'tbl_df', 'tbl' and 'data.frame':    56 obs. of  8 variables:
##  $ airline               : chr  "Aer Lingus" "Aeroflot*" "Aerolineas Argentinas" "Aeromexico*" ...
##  $ avail_seat_km_per_week: num  3.21e+08 1.20e+09 3.86e+08 5.97e+08 1.87e+09 ...
##  $ incidents_85_99       : int  2 76 6 3 2 14 2 3 5 7 ...
##  $ fatal_accidents_85_99 : int  0 14 0 1 0 4 1 0 0 2 ...
##  $ fatalities_85_99      : int  0 128 0 64 0 79 329 0 0 50 ...
##  $ incidents_00_14       : int  0 6 1 5 2 6 4 5 5 4 ...
##  $ fatal_accidents_00_14 : int  0 1 0 0 0 2 1 1 1 0 ...
##  $ fatalities_00_14      : int  0 88 0 0 0 337 158 7 88 0 ...
```

Print out the first four columns:

```
knitr::kable(airline_safety[, 1:4])
```

airline	avail_seat_km_per_week	incidents_85_99	fatal_accidents_85_99
Aer Lingus	320906734	2	0
Aeroflot*	1197672318	76	14
Aerolineas Argentinas	385803648	6	0
Aeromexico*	596871813	3	1

airline	avail_seat_km_per_week	incidents_85_99	fatal_accidents_85_99
Air Canada	1865253802	2	0
Air France	3004002661	14	4
Air India*	869253552	2	1
Air New Zealand*	710174817	3	0
Alaska Airlines*	965346773	5	0
Alitalia	698012498	7	2
All Nippon Airways	1841234177	3	1
American*	5228357340	21	5
Austrian Airlines	358239823	1	0
Avianca	396922563	5	3
British Airways*	3179760952	4	0
Cathay Pacific*	2582459303	0	0
China Airlines	813216487	12	6
Condor	417982610	2	1
COPA	550491507	3	1
Delta / Northwest*	6525658894	24	12
Egyptair	557699891	8	3
El Al	335448023	1	1
Ethiopian Airlines	488560643	25	5
Finnair	506464950	1	0
Garuda Indonesia	613356665	10	3
Gulf Air	301379762	1	0
Hawaiian Airlines	493877795	0	0
Iberia	1173203126	4	1
Japan Airlines	1574217531	3	1
Kenya Airways	277414794	2	0
KLM*	1874561773	7	1
Korean Air	1734522605	12	5
LAN Airlines	1001965891	3	2
Lufthansa*	3426529504	6	1
Malaysia Airlines	1039171244	3	1
Pakistan International	348563137	8	3
Philippine Airlines	413007158	7	4
Qantas*	1917428984	1	0
Royal Air Maroc	295705339	5	3
SAS*	682971852	5	0
Saudi Arabian	859673901	7	2
Singapore Airlines	2376857805	2	2
South African	651502442	2	1
Southwest Airlines	3276525770	1	0
Sri Lankan / AirLanka	325582976	2	1
SWISS*	792601299	2	1
TACA	259373346	3	1
TAM	1509195646	8	3
TAP - Air Portugal	619130754	0	0
Thai Airways	1702802250	8	4
Turkish Airlines	1946098294	8	3
United / Continental*	7139291291	19	8
US Airways / America West*	2455687887	16	7
Vietnam Airlines	625084918	7	3
Virgin Atlantic	1005248585	1	0
Xiamen Airlines	430462962	9	1