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 |
| 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 |