London Marathon

CONTENTS

## Loading library

library(tidyverse)  
library(flextable)

## Purpose

TidyTuesday is a weekly data analysis project in which participants explore and visualize a provided dataset using R. I will be contributing to TidyTuesday by producing a report in either Word or PDF format using Quarto in R. This week’s TidyTuesday challenge involves analyzing dataset(s) on London Marathon.

## Loading data

From [github](https://github.com/rfordatascience/tidytuesday/blob/master/data/2023/2023-04-25/readme.md):

The data this week comes from Nicola Rennie’s [LondonMarathon R package](https://github.com/nrennie/LondonMarathon). This is an R package containing two data sets scraped from Wikipedia (1 November 2022) on London Marathon winners, and some general data. How the dataset was created, and some analysis, is described in Nicola’s post [“Scraping London Marathon data with {rvest}”](https://nrennie.rbind.io/blog/web-scraping-rvest-london-marathon/). Thank you for putting this dataset together @nrennie!

winners <- readr::read\_csv('https://raw.githubusercontent.com/rfordatascience/tidytuesday/master/data/2023/2023-04-25/winners.csv')  
  
no\_of\_row <- nrow(winners)  
no\_of\_col <- ncol(winners)

## Data description

For this analysis, we will be only using the winners dataset

1. winners dataset contains 163 number of rows with 5 variables

## Exploring data

1. Displaying the Nationality of London Marathon Winners (Top 5)

winners %>%  
 count(Nationality, sort = T) %>%  
 rename(Total = n) %>%  
 top\_n(5) %>%  
 flextable() %>%  
 set\_table\_properties(layout = "autofit") %>%  
 #bold  
 bold(part = "header") %>%  
 theme\_box() %>%  
 ## Header colour---  
 color(color = "white", part = "header") %>%   
 ## Header background ---  
 bg(bg = "#17365D", part = "header") %>%  
 ## Font  
 font(fontname = "Cambria", part = "all") %>%  
 ## Add a gap between code and table  
 set\_caption("")

| **Nationality** | **Total** |
| --- | --- |
| United Kingdom | 44 |
| Kenya | 30 |
| United States | 11 |
| Switzerland | 10 |
| Ethiopia | 9 |

1. Displaying the winning time

winners %>%  
 ggplot(aes(x = Year, y = Time)) +  
 geom\_line(colour = "#17365D") +  
 labs(x = "Year", y = "Winning Time", fill = "Nationality") +  
 facet\_wrap(Category ~.) +  
 theme\_minimal()

