

RMarkdown Introduction

Tatjana Kecojević

2018-09-27

R Markdown

This is my first R Markdown document. Markdown is a simple formatting syntax for authoring HTML, PDF, and MS Word documents. For more details on using R Markdown see <http://rmarkdown.rstudio.com>.

Material from today's workshop can be found on our GitHub page.

When analysing data, a starting point is to examine the characteristics of each individual variable in the data set. The way to proceed depends upon the type of variable being examined. The variables can be one of two broad types:

1. **Attribute variable:** has its outcomes described in terms of its characteristics or attributes;
2. **Measured variable:** has the resulting outcome expressed in numerical terms.

Including R Code

When you click the Knit button a document will be generated that includes both content as well as the output of any embedded R code chunks within the document. You can embed an R code chunk like this:

```
olympic <- read.csv("data/athlete_events.csv")
olympic[1:3, ]
```

```
##   ID          Name Sex Age Height Weight   Team NOC      Games
## 1  1          A Dijiang  M  24    180     80   China CHN 1992 Summer
## 2  2          A Lamusi  M  23    170     60   China CHN 2012 Summer
## 3  3 Gunnar Nielsen Aaby  M  24     NA     NA Denmark DEN 1920 Summer
##   Year Season      City      Sport      Event Medal
## 1 1992 Summer Barcelona Basketball Basketball Men's Basketball <NA>
## 2 2012 Summer   London      Judo Judo Men's Extra-Lightweight <NA>
## 3 1920 Summer Antwerpen  Football Football Men's Football <NA>
```

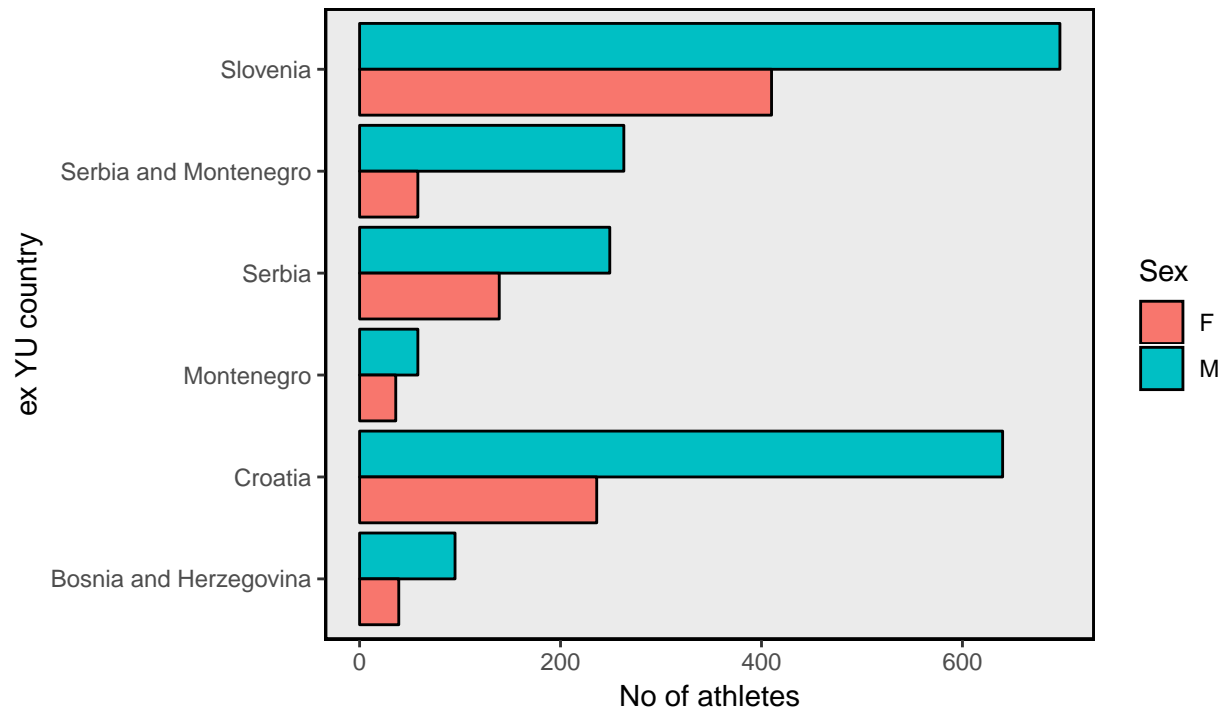
Including Plots

You can also embed plots by setting `echo = FALSE` to the code chunk to prevent printing of the R code that generates the plot. For example:

```
## # A tibble: 12 x 3
## # Groups:   Team [?]
##   Team                Sex  total
##   <fct>              <fct> <int>
## 1 Bosnia and Herzegovina F      39
## 2 Bosnia and Herzegovina M      95
## 3 Croatia             F     236
## 4 Croatia             M     640
## 5 Montenegro          F      36
## 6 Montenegro          M      58
## 7 Serbia              F     139
```

```
## 8 Serbia M 249
## 9 Serbia and Montenegro F 58
## 10 Serbia and Montenegro M 263
## 11 Slovenia F 410
## 12 Slovenia M 697
```

Comparisons of M and F representatives in exYU Teams for klikR workshop



Data from: kaggle – 120 years of Olympic history

Including Mathematical Equations

Let us fit the following model

$$lifeExp = b_0 + b_1 pop + b_2 gdpPercap$$

which we write using the LaTeX.

```
m1 <- lm(olympic$Weight ~ olympic$Age)
summary(m1)
```

```
##
## Call:
## lm(formula = olympic$Weight ~ olympic$Age)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -44.006  -9.896  -1.232   8.213 145.541
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  56.810736   0.144012   394.49  <2e-16 ***
```

```
## olympic$Age 0.554671 0.005613 98.82 <2e-16 ***
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 14.03 on 207377 degrees of freedom
## (63737 observations deleted due to missingness)
## Multiple R-squared: 0.04497, Adjusted R-squared: 0.04497
## F-statistic: 9766 on 1 and 207377 DF, p-value: < 2.2e-16
```

What do we think of this model?

Let's discuss it next time we meet up.