Tarea

2023-02-21

R Markdown

This is an 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 (http://rmarkdown.rstudio.com).

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:

Tarea

Utilizando la data who que se puede cargar con data("who"). Esta tabla recoge datos sobre casos reportados de tuberculosis en el mundo. Construya un código para que se vea la siguiente figura. Que contiene:

- Datos de niñas y niños entre 0 y 14 años.
- Datos del año 1996 en adelante (1996 - 2012).
- Geometría de nube de puntos.
- Geometría de regresión lineal.
- Ecuación de la regresión de la regresión y el coeficiente de determinación R2adj .

Coloración por años.

```
library(ggplot2)
library(tidyverse)
```

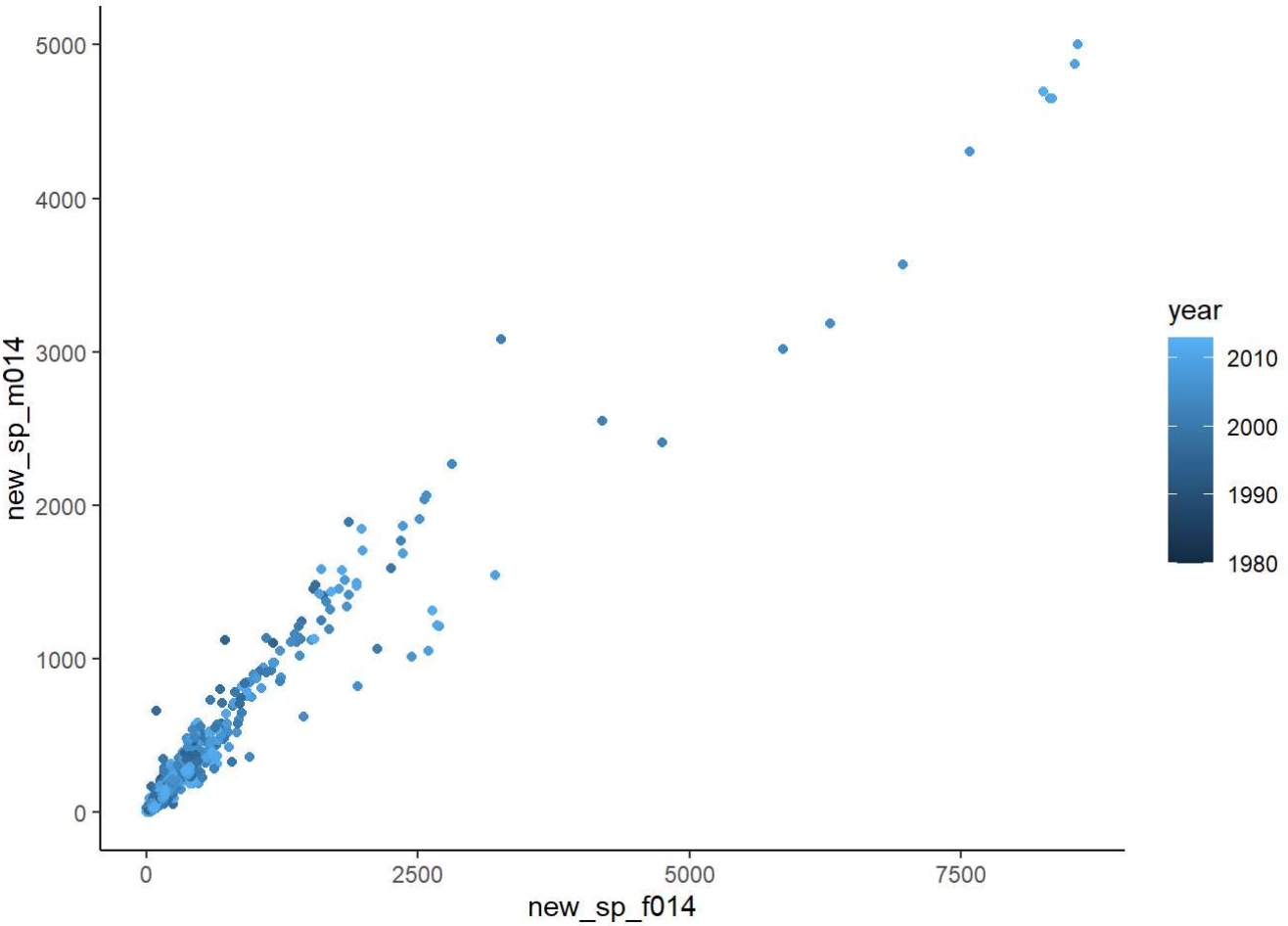
```
## ── Attaching packages ─────────────────────────────────────── tidyverse 1.3.2 ──
## ✓ tibble  3.1.8     ✓ dplyr   1.1.0
## ✓ tidyr   1.3.0     ✓ stringr 1.5.0
## ✓ readr   2.1.4     ✓ forcats 1.0.0
## ✓ purrr   1.0.1
## ── Conflicts ────────────────────────────────────────── tidyverse_conflicts() ──
## ✖ dplyr::filter() masks stats::filter()
## ✖ dplyr::lag()    masks stats::lag()
```

```
library(ggpubr)
library(ggpmisc)
```

```
## Loading required package: ggpp
##
## Attaching package: 'ggpp'
##
## The following object is masked from 'package:ggplot2':
##
##     annotate
```

```
## Cargamos el dataset who
data("who")

## Datos de niñas y niños entre 0 y 14 años.
cero_catorce <- who %>% select(country, year, new_sp_m014, new_sp_f014)

## Datos del año 1996 en adelante (1996 - 2012)
subset_96_12 <- who %>% filter (year >= 1996 & year <= 2012)

## Conjunto final de niños y niñas positivas desde el año 1996
## Eliminados los filas que contienen NA
final_data <- who %>% select(country, year, new_sp_m014, new_sp_f014) %>%
  filter (year >= 1996 & year <= 2012) %>% subset(!is.na(new_sp_m014)) %>%
  subset(!is.na(new_sp_f014))

## Geometría de nube de puntos.
ggplot(cero_catorce, aes(x=new_sp_f014, y=new_sp_m014, color = year)) + geom_point() +
  theme_classic()
```

```
## Warning: Removed 4086 rows containing missing values (`geom_point()`).
```



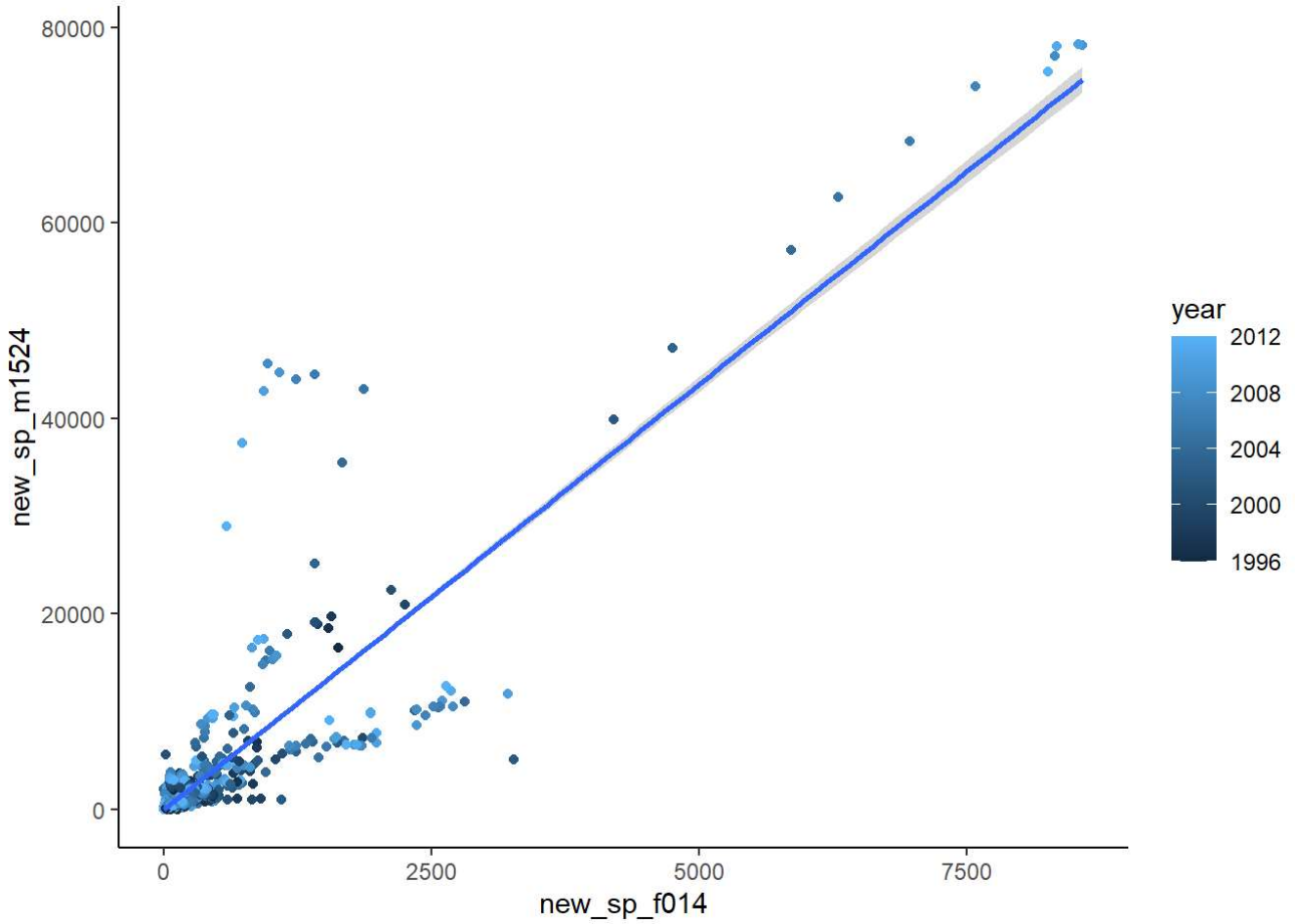

```
## Geometría de regresión lineal
ggplot(subset_96_12, aes(x=new_sp_f014, y=new_sp_m1524, color = year)) + geom_point() +
  theme_classic() + geom_smooth(method = "lm")
```

```
## `geom_smooth()` using formula = 'y ~ x'
```

```
## Warning: Removed 635 rows containing non-finite values (`stat_smooth()`).
```

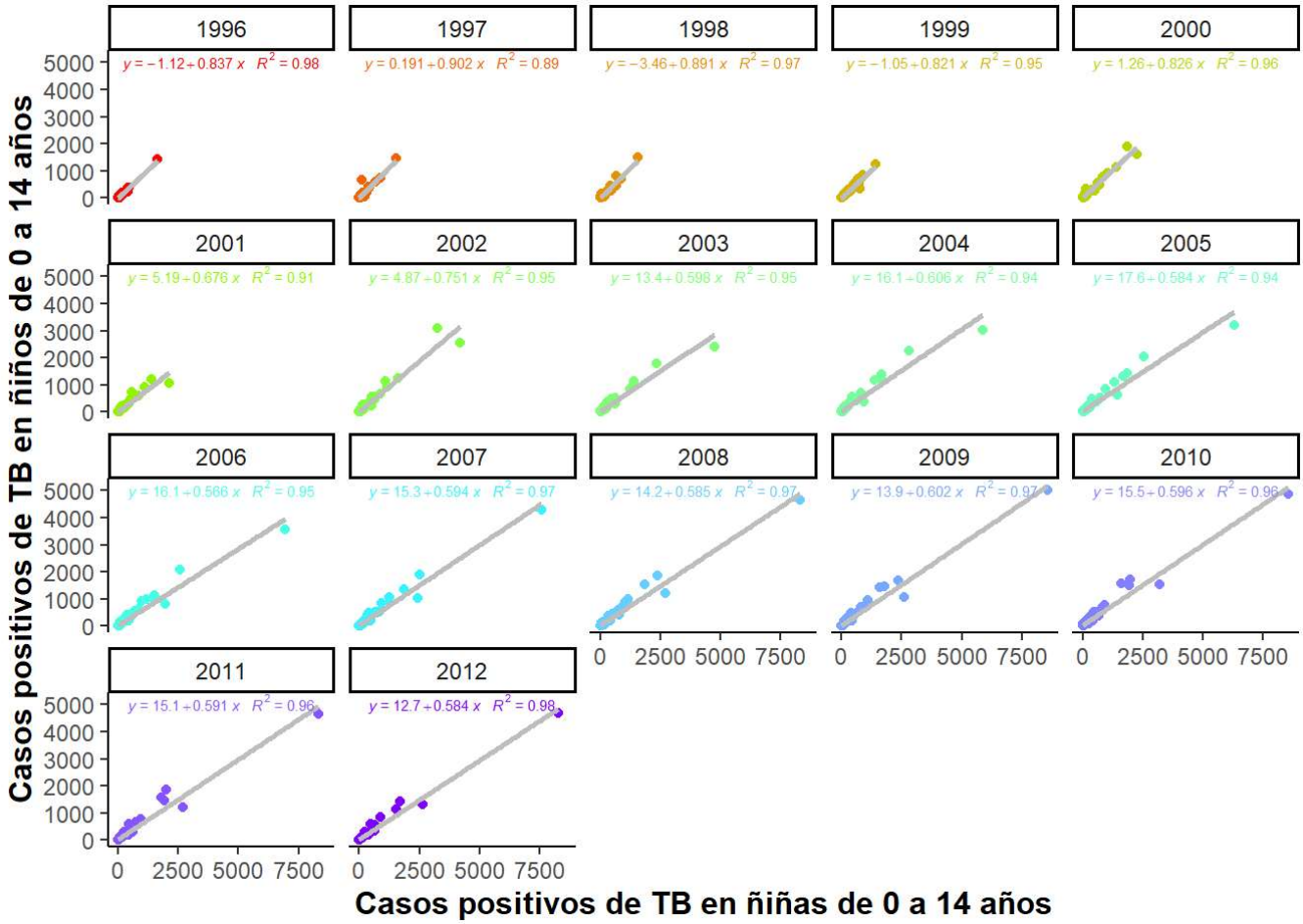
```
## Warning: The following aesthetics were dropped during statistical transformation: colour
## ℹ This can happen when ggplot fails to infer the correct grouping structure in
##   the data.
## ℹ Did you forget to specify a `group` aesthetic or to convert a numerical
##   variable into a factor?
```

```
## Warning: Removed 635 rows containing missing values (`geom_point()`).
```




```
## Ecuación de la regresión de la regresión y el coeficiente de determinación R2adj
formula <- y ~ x
ggplot(final_data, aes(x=new_sp_f014, y=new_sp_m014, color = year)) + geom_point() +
  theme_classic() +
  labs(x = "Casos positivos de TB en ñiñas de 0 a 14 años", y = "Casos positivos de TB en ñiñ
os de 0 a 14 años") +
  facet_wrap(year ~.) + geom_smooth(method = "lm", se=FALSE, color="grey", formula = formula)
+
  stat_poly_eq(aes(label = paste(..eq.label.., ..rr.label.., sep = "~~~")),
               label.x.npc = "center", label.y.npc = 5000,
               formula = formula, parse = TRUE, size = 2) +
  scale_colour_gradientn(colours=rainbow(4)) +
  theme(legend.position = "none",
        axis.title.x = element_text(face = "bold", size = 12),
        axis.title.y = element_text(face = "bold", size = 12))
```

```
## Warning: The dot-dot notation (`..eq.label..`) was deprecated in ggplot2 3.4.0.
## ℹ Please use `after_stat(eq.label)` instead.
```




```

Note that the `echo = FALSE` parameter was added to the code chunk to prevent printing of the R code that generated the plot.
```