

RMarkdown en 10 minutos - SOC4001 2024

M. Bucca

7 de Noviembre de 2024

Reportes automatizados en RMarkdown

Primeros pasos

En este documento revisaremos elementos básicos de la escritura de documentos en **RMarkdown**.

Introducción

Texto

La parte principal de un informe en **RMarkdown** suele ser texto. En un fichero `.Rmd`, todo lo que no sea encabezamiento código será interpretado como texto y lo mostrará tal cual. El texto de un documento `.Rmd` es “simplemente” texto PERO está escrito en *Markdown*. Lo que escribamos en **Rmarkdown** se mostrará tal cual en el documento final, pero es posible dar un poco de formato: negritas, cursivas, listas, enlaces de internet, etc. Para mayor detalle puedes visitar el siguiente link

Ecuaciones

En **Rmarkdown** se pueden introducir formulas matemáticas (escritas en *Látex*). Para formulas en linea se usa el signo `$` al inicio y al final de la expresión. Por ejemplo, el código `$y = f(x)$` produce la siguiente ecuación: $y = f(x)$.

Para escribir la misma ecuación en una linea independiente, se usa el signo `$$`. Por ejemplo, el código `$$y_{i} = \alpha + \beta_1 x_{i} + \beta_2 x_{i}^2 + \epsilon_i$$` produce la siguiente ecuación:

$$y_i = \alpha + \beta_1 x_i + \beta_2 x_i^2 + \epsilon_i$$

$$e = m \times c^2$$

$$\lim_{x \rightarrow \infty} \exp(-x) = 0$$

Para mayor detalle: aquí

$$f(n) = \begin{cases} n/2 & \text{if } n \text{ is even} \\ -(n+1)/2 & \text{if } n \text{ is odd} \end{cases}$$

Código (“chunks”)

RMarkdown permite introducir código de **R** en el documento de texto, evaluar tal código y mostrar los resultados directamente en el informe. A modo de ejemplo, comenzaremos mostrando un **summary** de la base de datos **Chile**, que viene incluida en el paquete **carData**.

```
mediana <- datos_chile %>% summarise(median_ingreso = quantile(income, p=0.5, na.rm = T)) %>% as.numeric()
mediana
```

```
## [1] 15000
```

En 1988 la mediana de ingreso en Chile era de 1.5×10^4 .

El trozo de arriba es un chunk de código R. Al compilar el documento, (click en el botón **knitr**, en el panel) el código se ejecutará y mostrarán los resultados en el documento final. Los chunks pueden tener diversas opciones que permiten una mayor flexibilidad en como se muestra el código y los resultados. Las opciones más usadas son:

- echo
- eval

Por ejemplo, el chunk abajo mostrará el código (**echo = TRUE**), lo evaluará y mostrará los resultados en el documento final (**eval = TRUE**). Así se ve:

```
mediana <- datos_chile %>% summarise(median_ingreso = quantile(income, p=0.5, na.rm = T)) %>% as.numeric()
mediana
```

```
## [1] 15000
```

Si sólo queremos mostrar el código (**echo = TRUE**) pero no evaluarlo (**eval = FALSE**), escribimos lo siguiente:

```
datos_chile %>% dplyr::summarise(mean(income, na.rm = T))))))))))))))))))))))))))))))))))))))))))
```

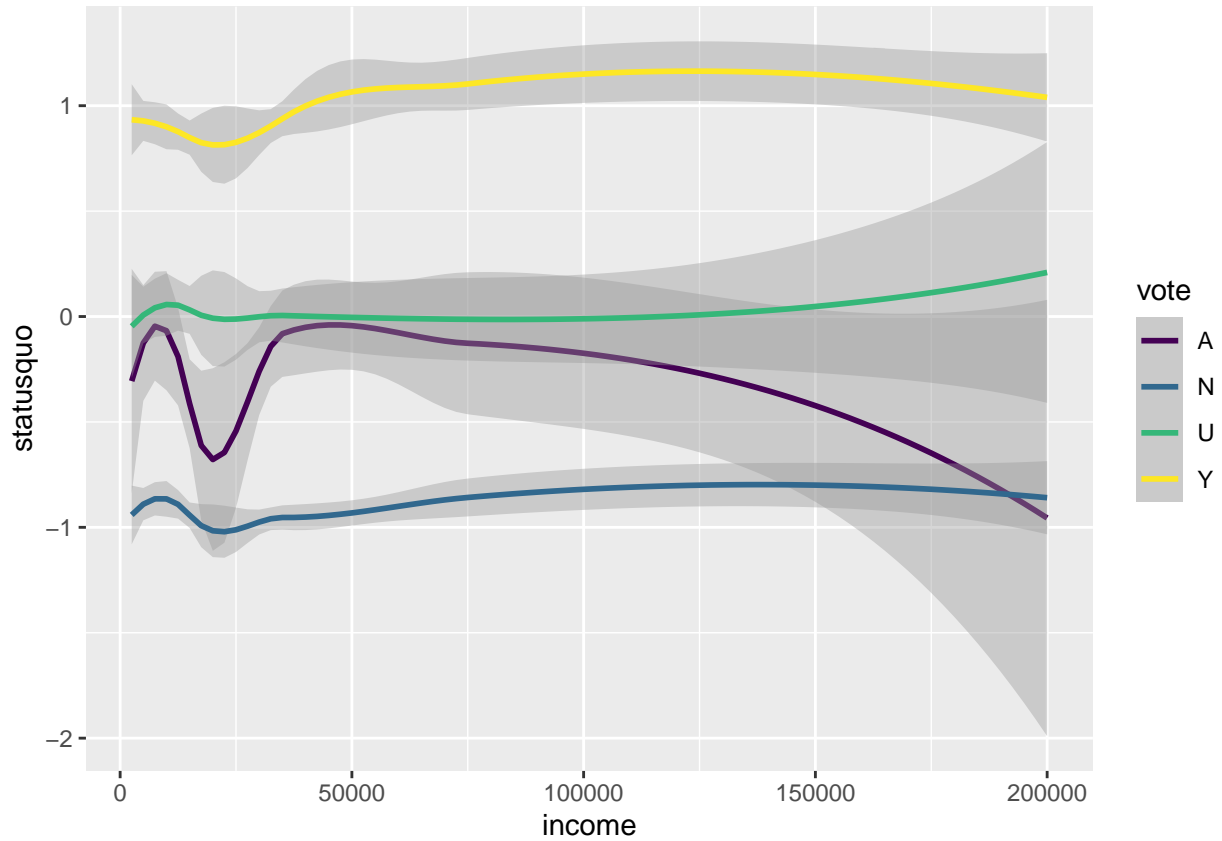
Por el contrario, si queremos evaluar el código, mostrar sus resultados, pero no mostrar el código mismo, escribimos:

Como muestra nuestro análisis el promedio de ingreso es de 3.402×10^4 . Por último, si queremos NO mostrar el código (**echo = FALSE**), SI evaluarlo (**eval = TRUE**), PERO NO se mostrar los resultados (**results = "hide"**), escribimos:

Que el código haya sido evaluado significa que el objeto “a” existirá en la memoria y podrá ser usado para posterior análisis.

Gráficos

También podemos mostrar gráficos producidos en R. Notar que los paquetes necesarios para implementar un determinado análisis (**ggplot2** y **tidyverse**, en este caso) deben ser previamente cargados.



Tablas

Este reporte usa datos de la base de datos Chile:

name	value
income	34019.9506376
statusquo	-0.0112671
age	38.2900041

Mis datos son estos:

region	population	sex	age	education	income	statusquo	vote
N	175000	M	65	P	35000	1.00820	Y
N	175000	M	29	PS	7500	-1.29617	N
N	175000	F	38	P	15000	1.23072	Y
N	175000	F	49	P	35000	-1.03163	N
N	175000	F	23	S	35000	-1.10496	N
N	175000	F	28	P	7500	-1.04685	N