Practica\_2

Marta Pulido López y Juan Manual Martín Jiménez

26/2/2020

En este documento se dan respuestas a los ejercicios de la práctica 2, el cual ha sido generado con R Markdown utilizando R versión 3.6.2 (R Core Team ([2019](#ref-RStudio))). Además, para la realización de las tablas y gráficas se han usado paquetes adicionales, tales como Tidyverse 1.3.0. (Wickham et al. ([2019](#ref-Tidyverse))), Officer 0.3.6. (Gohel ([2019](#ref-Officer))) y Flextable 0.5.8. (Gohel ([2020](#ref-Flextable))).

# 2.1. Responde a las siguientes preguntas:

# (1) ¿Qué es la Ciencia Reproducible? ¿En qué casos se consigue un 100% de la reproducibilidad?

La Ciencia Reproducible es aquella que permite garantizar que los resultados son perfectamente reproducibles ya que tanto la integración de texto, como el código y los datos incluidos permiten regenerar fácilmente el resultado final a partir de los datos originales. El análisis completo de datos y el manejo de los mismos se hace mediante código de tal forma que es posible recrear exactamente a partir de los datos originales todos los resultados y figuras incluidos y trazar de forma clara el proceso seguido para obtener los resultados y reproducirlos (Rodriguez-Sanchez, Pérez-Luque, Bartomeus, & Varela ([2016](#ref-rodriguez2016ciencia))).

Se consigue un 100% de reproducibilidad en aquellos casos en los que el análisis completo y el manejo de los datos se ha realizado mediante código.

# (2) ¿Cuáles de los beneficios que se sugieren son para ti los más importantes? Justifica tu respuesta.

Los principales beneficios de la Ciencia Reproducible es que se reduce significaticamente el riesgo de errores, es posible ejecutar tareas repetitivas sin esfuerzo y es muy fácil corregir y regenerar resultados, tablas y figuras. De esta forma, en caso de contener errores, es muy fácil detectarlos y corregirlos con sólo un botón y volver a generar todas las tablas, figuras y resultados de un trabajo, sin tener que usar diferentes cálculos y programas para su actualización (Rodriguez-Sanchez et al. ([2016](#ref-rodriguez2016ciencia))).

Por último, otro de los beneficios que consideramos muy importantes es que contiene un registro de todos los pasos seguidos en el análisis, lo que resulta fácil de compartir entre los colaboradores así como para reutilizarlo pasado un tiempo (en el que sería muy difícil repetir cada uno de los pasos seguidos).

# (3) ¿Por qué no hace todo el mundo Ciencia Reproducible?

Realizar Ciencia Reproducible requiere dominar técnicas como bases de datos, programación, git o R cuyo control requiere que el usuario dedique tiempo para aprenderlas y coger soltura.

Por ello, familiarizarse con estas herramientas en un principio puede resultar tedioso, aunque el tiempo invertido en aprender a utilizar este tipo de herramientas se ve recompensado cuando podemos regenerar todo un manuscrito, tablas y figuras incluidas con un solo clic.

## 2.2. Lee el artículo “¿Por qué usar GitHub? Diez pasos para disfrutar de GitHub y no morir en el intento” (Galiano 2018) Realiza las siguientes tareas.

Como evidencia de creación del repositorio, se puede consultar: [Repositorio](https://github.com/martapulido/practica2.git)

## 2.3. Lee “¿Por qué usar GitHub? Diez pasos para disfrutar de GitHub y no morir en el intento”. Comprueba que tienes “claro” los siguientes conceptos:

Leído el artículo “¿Por qué usar GitHub? Diez pasos para disfrutar de GitHub y no morir en el intento” (Prieto ([2018](#ref-prieto2018que))) se definen los conceptos:

* Repositorio: es un directorio público o privado en el que se guardan los archivos que contiene git y que git monitorea.
* Rama: es una especie de copia que permite realizar modificaciones en respositorios sin modificar los archivos originales. Git permite crear una especie de línea paralela de trabajo que posteriormente analizará si detecta conflicto la línea principal u otras y que se podrá unir a la rama principal posteriormente si el responsable del respositorio lo autoriza. Es útil para trabajar varias personas a la vez ya que registra las modificaciones que se realizan e informa al responsable de los cambios, quien puede o no aceptarlos.
* Pull and push: pull permite descargar las actualizaciones del repositorio y push permite subir las modificaciones al respositorio.
* Fork: permite hacer una copia de repositorios ajenos.

# 2.4. RMarkdown.

## 2.4.1. Generación de un documento de html con varias sintaxis.

Se crean opciones de sintaxis básicas de Pandoc Markdown conforme se solicita:

* Texto plano.
* *Cursiva*
* **Negrita**
* Superíndice2: r2
* \*Subíndice2: SO4
* ~~Tachado~~
* [rstudio](www.rstudio.com)
* Ecuación en línea:
* Ecuación en bloque:

# Encabezado 1

## Encabezado 2

### Encabezado 3

#### Encabezado 4

##### Encabezado 5

###### Encabezado 6

## 2.4.2. Escribe fragmentos de código con las siguientes características:

1. Peguen las palabras “Hola” y “mundo” en una línea nueva

"Hola"  
## [1] "Hola"  
  
"Mundo"  
## [1] "Mundo"

1. Peguen las palabras “Hola” y “mundo” en la misma línea de código

paste("Hola", "mundo")

## [1] "Hola mundo"

1. Peguen las palabras “Hola” y “mundo” en una línea nueva de código y que no muestren el resultado de R

paste("Hola", "mundo")

1. Peguen las palabras “Hola” y “mundo” en una línea nueva de código y que no muestren el resultado de R y que no se muestre en el documento de Word.

## 3. Muestra una tabla y una gráfica en el archivo de salida, tomando alguno de los datos que vienen en R o alguno de sus paquetes. ¡¡Se creativo!! Usando (al menos): encabezados, explicaciones, ecuaciones, líneas de código en el texto y bloques de código.

Se incluye una fórmula para el cálculo de la desviación típica y se vincula el resultado al cálculo de un campo de Iris, concretamente Sepal.With de modo que el resultado mostado es el cálculo de la desviación típica del ancho del sépalo.

= 0.4358663

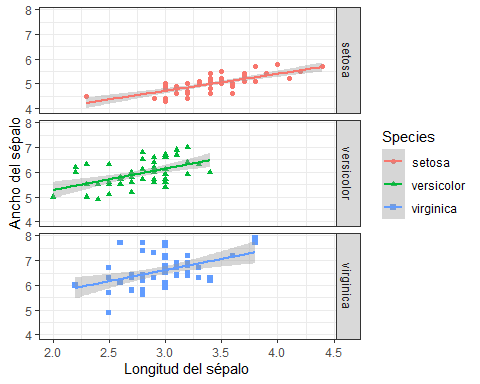
Se crea una tabla en la que se incluyan aquellas especies que tengan una longitud del sépalo entre 1 y 5 en color verde y en negrita aquellas especies que tengan un ancho de sépalo mayor a 3.

## Warning: package 'officer' was built under R version 3.6.3

## Warning: package 'forcats' was built under R version 3.6.3

| Sepal.Length | Sepal.Width | Petal.Length | Petal.Width | Species |
| --- | --- | --- | --- | --- |
| **5.1** | **3.5** | **1.4** | **0.2** | **setosa** |
| 4.9 | 3.0 | 1.4 | 0.2 | setosa |
| **4.7** | **3.2** | **1.3** | **0.2** | **setosa** |
| **4.6** | **3.1** | **1.5** | **0.2** | **setosa** |
| **5.0** | **3.6** | **1.4** | **0.2** | **setosa** |
| **7.0** | **3.2** | **4.7** | **1.4** | **versicolor** |
| **6.4** | **3.2** | **4.5** | **1.5** | **versicolor** |
| **6.9** | **3.1** | **4.9** | **1.5** | **versicolor** |
| 5.5 | 2.3 | 4.0 | 1.3 | versicolor |
| 6.5 | 2.8 | 4.6 | 1.5 | versicolor |
| **6.3** | **3.3** | **6.0** | **2.5** | **virginica** |
| 5.8 | 2.7 | 5.1 | 1.9 | virginica |
| 7.1 | 3.0 | 5.9 | 2.1 | virginica |
| 6.3 | 2.9 | 5.6 | 1.8 | virginica |
| 6.5 | 3.0 | 5.8 | 2.2 | virginica |

Se incluye un gráfico de la longitud y ancho del sépalo en función de la especie.



## Referencias:

Gohel, D. (2019). *Officer: Manipulation of microsoft word and powerpoint documents*. Retrieved from <https://CRAN.R-project.org/package=officer>

Gohel, D. (2020). *Flextable: Functions for tabular reporting*. Retrieved from <https://CRAN.R-project.org/package=flextable>

Prieto, J. G. (2018). ?‘ Por qué usar github? Diez pasos para disfrutar de github y no morir en el intento. *Revista Ecosistemas*, *27*, 140–141.

R Core Team. (2019). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from <https://www.R-project.org/>

Rodriguez-Sanchez, F., Pérez-Luque, A. J., Bartomeus, I., & Varela, S. (2016). Ciencia reproducible: Qué, por qué, cómo. *Revista Ecosistemas*, *25*, 83–92.

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L. D., François, R., … Yutani, H. (2019). Welcome to the tidyverse. *Journal of Open Source Software*, *4*, 1686. <https://doi.org/10.21105/joss.01686>