

Investigación reproducible con Quarto y R

Felipe Ortega María Jesús Algar

2024-10-24

Tabla de contenidos

Prefacio	1
Requisitos previos	2
I Quarto	3
1 Documentos científicos	4
1.1 Programación literaria	5
1.2 Investigación reproducible	6
1.2.1 Reproducibilidad y replicabilidad	8
1.2.2 Niveles de replicación	9
1.2.3 Herramientas para la replicabilidad	10
1.3 Quarto para publicaciones científicas	13
1.4 Instalación de Quarto	14
2 Tipos de documentos	16
2.1 Documentos individuales	16
2.2 Libros	16
2.3 Artículos y publicaciones	17
2.4 Presentaciones	17
2.5 Sitios web	17
2.6 <i>Dashboards</i>	18
3 Proceso de trabajo	19
3.1 Cadena de montaje de un documento	19
3.2 Producir HTML	21
3.3 Producir PDF	22
3.3.1 Personalizar documentos PDF	22
4 Documentos individuales	24
4.1 Creación del documento con RStudio	24
4.2 Estructura del documento	26
4.2.1 El preámbulo	27
4.2.2 Listado de opciones	28
4.2.3 Sintaxis Markdown básica	29

Tabla de contenidos

4.3	Creación del documento (<i>output</i>)	29
4.3.1	Previsualización	29
4.3.2	Seleccionar el tipo de documento	30
4.3.3	Opciones básicas de configuración	31
4.4	<i>Chunks</i> de código ejecutable	32
4.5	Herramientas para el autor	37
4.5.1	Secciones del documento	37
4.5.2	Ecuaciones	39
4.5.3	Tablas	40
4.5.4	Llamadas	43
4.5.5	Citas bibliográficas	43
4.5.6	Estilo general del documento	45
II	Libros con Quarto	46
5	Libros	47
5.1	Creación de proyectos de tipo <code>book</code>	47
5.2	Opciones de configuración	47
5.3	Archivo de portada	50
5.4	Herramientas de redacción	50
5.4.1	Estructura del libro	50
5.5	Gestión de referencias	52
5.6	Previsualización del proyecto	52
5.7	Publicación del proyecto	52
5.8	Plantillas y personalización	53
6	Taller: colección de apuntes	55
6.1	Opciones de diseño del documento	55
6.1.1	Contenido principal	56
6.1.2	Contenido en los márgenes	57
6.1.3	Notas al margen y citas bibliográficas	58
6.2	Anotaciones en bloques de código	59
6.3	Configuración del proyecto	60
6.3.1	Opciones globales	60
6.3.2	Idioma de los documentos	60
6.3.3	Enlaces a otros formatos	61
6.4	Ejercicio: personalizar el estilo gráfico del documento	63
6.5	Personalización de los libros	63
6.6	Plantillas y extensiones de Quarto	66
6.7	Publicación del proyecto	67
6.7.1	Método con <code>quarto-publish</code>	68

Tabla de contenidos

III Publicaciones científicas	71
7 Artículos y publicaciones científicas	72
7.1 La opción <code>keep-tex: true</code>	74
7.2 Figuras y gráficos para publicación	74
7.3 Facilitar la cita de artículos	75
7.4 Ejemplo de uso de plantillas de artículos científicos	76
7.4.1 Plantilla para revistas de Elsevier	76
8 Principios FAIR	78
8.1 Visión general	78
8.2 Publicación del código fuente y documentación técnica	80
8.3 Publicación de conjuntos de datos	82
8.4 Gestión de referencias y publicación en abierto .	82
9 Recursos adicionales	84
9.1 Quarto	84
9.2 Principios FAIR y ciencia abierta	84
Referencias	85
Apéndices	86
A Comandos y opciones de configuración	86
A.1 Resumen de sintaxis de Quarto	86
A.2 Comandos Quarto	86
A.3 Celdas de código	86
A.4 Formatos de documentos de salida	86
A.5 Opciones de configuración de proyectos	86
B Entornos de desarrollo para Quarto	87
B.1 R Studio	87
B.2 Visual Studio	87
C Paquetes R y atribuciones	88
C.1 Requisitos previos	88
C.2 Atribución de imágenes e iconos	88
D Documentos PDF con LaTeX	89
D.1 Salida en formato PDF	89
D.2 Acerca de LaTeX	89
D.2.1 Referencias de introducción a LaTeX . . .	90
D.3 Ejemplo práctico	90

Tabla de contenidos

Referencias	92
--------------------	-----------

Prefacio

Este taller explica cómo utilizar Quarto, un software de creación de documentación científica, para crear publicaciones de calidad que integren contenido de texto formateado, gráficos, tablas, así como resultados de ejecución de código software en varios lenguajes, todo ello integrado en el propio documento.

Quarto se ha convertido en una herramienta muy versátil y potente en el conjunto de herramientas de los programadores científicos, en especial por proporcionar soporte para implementar buenas prácticas de **investigación reproducible**, incluyendo los principios FAIR. El intenso movimiento iniciado desde hace años por la comunidad científica para garantizar acceso en abierto no solo al producto final (e.g. una publicación) sino también a materiales adicionales (código fuente, conjuntos de datos, figuras, procesos de trabajo, archivos de configuración, etc.) se ha convertido en un objetivo insoslayable para académicos y especialistas en muchos campos diferentes. Muchas publicaciones científicas de prestigio exigen ahora a los autores enviar estos materiales auxiliares junto con los borradores de sus manuscritos, para permitir que otros colegas reproduzcan y validen los resultados, repliquen sus estudios sobre nuevas cohortes de individuos o elementos o para contribuir a la interpretación de los resultados obtenidos.

Con la herramienta Quarto se puede combinar texto formateado (escrito en Markdown) junto con secciones de código ejecutable, todo integrado en un mismo documento. Estas secciones o *chunks* de código ejecutable pueden estar escritas en varios lenguajes: R, Python, Julia u Observable. Es incluso posible combinar en un mismo documento o colección de documentos secciones de código escritas en diferentes lenguajes de programación.

Este es un **taller práctico** que presenta ejemplos reales y comandos para crear paso a paso tus propios documentos con Quarto en poco tiempo. Además, junto a la explicación de los conceptos clave para entender este proceso también se ofrecen

Requisitos previos

recomendaciones sobre buenas prácticas de trabajo, para guiar a los aprendices de Quarto en la dirección correcta.

Puedes aprender muchos más detalles sobre cuarto en la guía en línea <https://quarto.org/docs/guide/>. En particular, en <https://quarto.org/docs/books> se documenta en detalle cómo crear libros como este utilizando Quarto.

Requisitos previos

Para poder realizar los ejemplos incluidos en este taller necesitas tener instalado R y una IDE de desarrollo para este lenguaje. Se recomienda instalar RStudio o MS Visual Code como entorno de programación.

- Instalación de R.
- Instalación de RStudio.

Adicionalmente, es necesario instalar una serie de paquetes R antes de ejecutar los ejemplos, para que todas las dependencias estén disponibles en nuestro sistema. Consulta el Apéndice Sección C.1 para comprobar el listado de paquetes R necesarios.

Parte I

Quarto

1 Documentos científicos

En su actividad diaria, los estudiantes, académicos y especialistas científicos producen gran cantidad de documentación de todo tipo: notas de laboratorio, apuntes, memorandos, informes técnicos y, sobre todo, artículos científicos para publicar sus descubrimientos y avances en un área de conocimiento. Normalmente, la creación de este tipo de documentos científicos conlleva una gran cantidad de tareas que involucran diferentes herramientas y posibles puntos de fallos.

La Figura 1.1 muestra una descripción general esquemática de un proceso de trabajo clásico para la creación de documentos científicos. Con frecuencia, el elemento principal es un archivo maestro de procesador de textos (Word, OpenOffice/LibreOffice, etc.), una página web o un fichero LaTeX (si vamos a crear un documento PDF) que recoge todo el contenido.

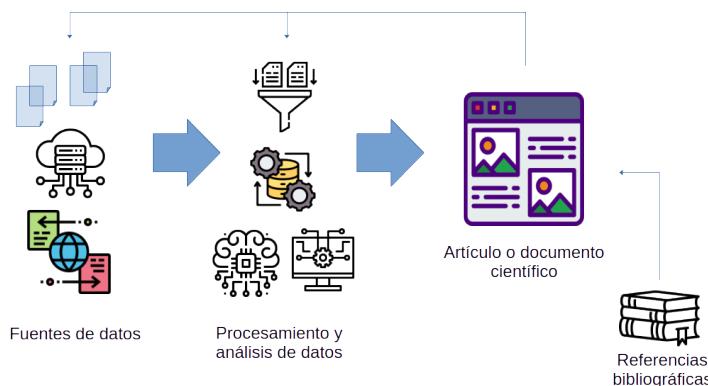


Figura 1.1: Proceso de creación de documentación y artículos científicos

Este archivo maestro se va llenando de contenido que procede de diversas fuentes, como por ejemplo:

- figuras y esquemas generados manualmente o mediante código software (como gráficos de visualización de datos);

1 Documentos científicos

- tablas y resúmenes que describen conjuntos de datos y resultados;
- resultados y evaluación del rendimiento de modelos o algoritmos; estadísticos o de aprendizaje automático;
- fórmulas y ecuaciones matemáticas;
- tablas de datos y otra información de utilidad;
- referencias bibliográficas (normalmente generadas con ayuda de algún programa de gestión de información bibliográfica).

Muchos de estos elementos fuerzan a los usuarios a ejecutar una y otra vez herramientas y programas externos, procedimientos y otras tareas para incorporar luego los nuevos resultados al fichero maestro. Debemos admitir que este proceso, en su mayor parte manual, además de tedioso puede ser muy propenso a que cometamos errores o descuidos. “¡Espera! He olvidado actualizar la Figura 1”. “¿Seguro que estos son los últimos resultados de evaluación del modelo M ?” “¿Has comprobado que hemos cargado la última versión del fichero de datos D ?” Estas son preguntas comunes que surgen en el día a día de los equipos científicos.

Sin embargo, sería genial si no fuese necesario realizar todo ese proceso manual y, en ocasiones, muy frustrante de forma manual. ¿Tenemos alguna alternativa para evitarlo? Sí, la tenemos. La respuesta a nuestras necesidades nos la brinda un concepto muy poderoso: la **programación literaria**.

1.1 Programación literaria

El concepto de programación literaria fue acuñado por el profesor Donald E. Knuth (1984). Sí, no has leído mal, hace más de 40 años. Este concepto establece que debería ser posible integrar, en un solo documento científico, texto formateado y resultados de la ejecución de código software para componer dicho documento de forma dinámica. Entonces, ¿por qué hemos tardado tanto en poner en práctica esta idea? La visión de Knuth, aunque muy adelantada a su tiempo, era correcta, pero la tecnología de la época no permitía ponerla en práctica.

Sin embargo, hoy día contamos con todos los elementos indispensables para llevar esta idea a la práctica. Es más, contamos

1 Documentos científicos

con una herramienta, Quarto, que nos va a permitir automatizar y gestionar todo el proceso de creación de documentos de programación literaria de forma rápida y fiable.

1.2 Investigación reproducible

Durante muchas décadas, el método científico se ha basado en la publicación de investigaciones que describen el resultado de análisis de datos y experimentos. En todos los casos, resulta fundamental poder confiar en las condiciones, los datos recabados, el método de análisis y de ejecución de los experimentos así como en las herramientas de diversas clase, incluído el software, que los autores de la publicación han empleado para llevarla a cabo.

Sin embargo, los numerosos avances experimentados en los últimos años en las herramientas y métodos de análisis permiten que ahora sea mucho más sencillo comprobar el resultado de estos análisis. Podríamos suponer que esto facilita mucho el trabajo de científicos y científicas, pero en realidad sucede todo lo contrario. Veamos algunos ejemplos:

- **Oncología** (Begley & Ellis, 2012): El Departamento de Biotecnología de la firma Amgen (Thousand Oaks, CA, EE.UU.) sólo pudo confirmar 6 de un total de 53 artículos de investigación emblemáticos publicados en este área. Por su parte, Bayer HealthCare (Alemania) tan sólo pudo validar un 25% de los estudios analizados.
- **Psicología** (Wicherts et al., 2006): El 73% de los autores de un total de 249 artículos publicados por la APA no respondieron en un periodo de 6 meses a las preguntas y requerimientos formulados acerca de los datos que emplearon en sus investigaciones.
- **Economía y finanzas** (Burman et al., 2010): La comparación de diferentes paquetes software aplicados en la ejecución de varios análisis de modelos financieros y estadísticos refleja que cada uno de dichos paquetes produce resultados *muy distintos* empleando *las mismas técnicas* estadísticas directamente aplicadas sobre *datos idénticos* a los empleados en la publicación original.

1 Documentos científicos

De hecho, han llegado a aparecer artículos que sugieren que buena parte de los resultados publicados en áreas como Medicina podrían no ser del todo fiables (Ioannidis, 2005). Como resultado de todos estos hallazgos recientes, se ha generado en toda la comunidad científica e investigadora una gran polémica, acompañada de una profunda crisis de confianza.

A pesar de todo, como muy bien recoge una conocida tira cómica sobre el mundo académico y la investigación (ver Figura 1.2), el proceso de desarrollo de las publicaciones científicas se basa primordialmente en la revisión continua de los métodos y resultados (empezando por los propios estudiantes y sus supervisores).

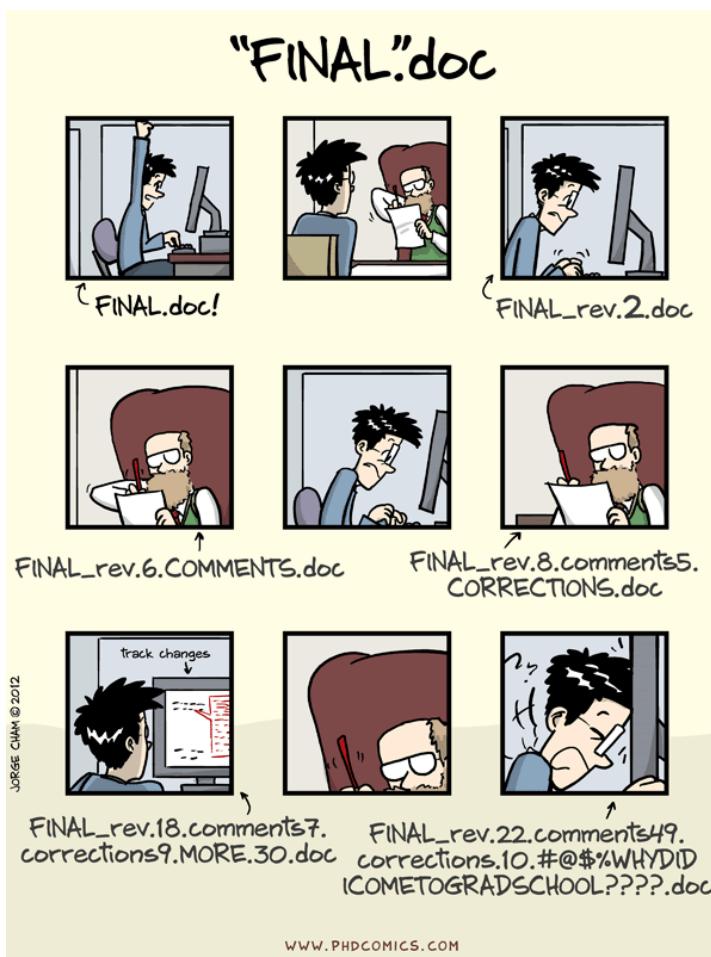


Figura 1.2: Tira cómica que representa el modelo de revisión de publicaciones científicas. Fuente: PhD comics.

1 Documentos científicos

La Figura 1.3 muestra un gráfico publicado en la prestigiosa revista Science Magazine (Brainard et al., 2018), que representa los datos sobre evolución del número de artículos de investigación retractados o retirados por diversas causas, entre 1997 y 2014. En este gráfico, podemos constatar cómo la mejora de las herramientas y la mayor disponibilidad de recursos permite analizar y revisar un mayor volumen de publicaciones y análisis, lo que permite detectar un mayor número de casos problemáticos.

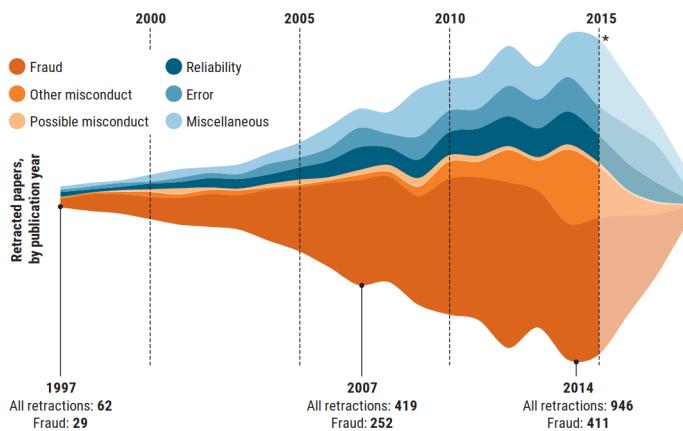


Figura 1.3: Evolución del número de publicaciones científicas retractadas o retiradas por diversas causas, entre 1997 y 2014. Fuente: Science Magazine (Brainard et al., 2018).

1.2.1 Reproducibilidad y replicabilidad

Se habla con frecuencia de *reproducir* y *replicar* un análisis de datos o un experimento científico (Leek & Peng, 2015). Sin embargo, se pueden citar numerosas evidencias que demuestran que existen definiciones incompatibles sobre estos dos términos y otros relacionados con ellos (Barba, 2018). Mucho cuidado, por tanto, porque dependiendo de la comunidad científica o el campo de conocimiento en que nos encontremos, el significado de estos dos términos puede llegar a ser incluso *enteramente opuesto* a su definición aceptada en otras áreas¹. Aquí vamos a ceñirnos a la definición aceptada en un gran número de áreas,

¹Guía rápida de KOMA-Script.

1 Documentos científicos

incluyendo estadística o computación científica (véase Barba, 2018, p. 33):

- **Reproducibilidad:** Se define como la capacidad para recomputar los resultados de un análisis, con los mismos datos que se emplearon en el análisis original, y conociendo los detalles de la secuencia (*workflow* o *pipeline*) de operaciones de componen dicho análisis. Se debe poder garantizar ciertas premisas:
 - Si usamos las mismas herramientas (e.g. R, un cierto listado de paquetes, las mismas versiones de todos los paquetes y dependencias), así como el mismo código (*scripts* de R) sobre los mismos datos, los resultados y conclusiones han de ser consistentes con los del análisis original.
 - Los autores del análisis original deben proporcionar todos los elementos (datos, código y procedimiento empleado) para permitir que el análisis sea reproducible (Barba, 2018).
- **Replicabilidad:** Se define como la capacidad para realizar un experimento o análisis independiente del original, que aborde el mismo objetivo pero sobre un conjunto de datos diferente del empleado en el estudio inicial. En caso de que los resultados no sean consistentes, será necesario realizar más réplicas y armonizar los resultados y conclusiones por medio de técnicas adecuadas, como por ejemplo el **meta-análisis**.

1.2.2 Niveles de replicación

En función de los elementos publicados por los autores del estudio original, así como del grado de detalle con el que se describe el proceso para llevar a cabo el estudio, los pasos que se han seguido y las herramientas empleadas, tenemos diferentes niveles de replicabilidad o reproducibilidad, representados en la Figura 1.4.

- *No reproducible:* No se proporcionan datos, código ni ninguna descripción concreta sobre la implementación del estudio o análisis. Muchas publicaciones científicas ya no aceptan publicar artículos en estas condiciones.

1 Documentos científicos

- *Código o Datos*: Un buen número de editoriales solicitan que los conjuntos de datos empleados en el análisis o estudio de la publicación sean accesibles mediante una URL, bien porque estén disponibles en un repositorio público o bien porque los autores del artículo lo han publicado. Así mismo, muchas publicaciones exigen que el código software para llevar a cabo el análisis también esté accesible públicamente, en un repositorio de código abierto o bien en un proyecto de un servicio de control de versiones con acceso libre.
- *Código y datos*: Lo ideal es que tanto el código como los datos estén públicamente accesibles a disposición de quien los quiera examinar o bien utilizar para reproducir los resultados (validación) o replicar el análisis con otros datos u otros casos.
- *Entorno de ejecución y datos enlazados*: Un paso más para facilitar la reproducibilidad de los estudios consiste en publicar archivos de código y metadatos con información más precisa sobre el lenguaje de programación, los paquetes software empleados y cualquier otra dependencia necesaria para llevar a cabo el mismo estudio o análisis. Otra variante para facilitar la reproducibilidad es la de encapsular el código y las dependencias en un contenedor virtual ya preconfigurado, que se pueda descargar y ejecutar directamente.
- *Gold standard*: El nivel más avanzado consistiría en documentar todos los procedimientos realizados durante el estudio o análisis, incluyendo la codificación de las tareas de obtención, limpieza y preparación de los datos, así como la generación de gráficas de visualización de resultados o cualquier otro resultado derivado del estudio.

1.2.3 Herramientas para la replicabilidad

Ciertas tecnologías y herramientas que se han refinado y sofisticado durante los últimos años están permitiendo facilitar la replicabilidad del procesamiento y análisis de datos.

- Sistemas de Control de Versiones para código software (SCV): herramientas como Git, Mercurial y servicios web como GitHub o GitLab han popularizado la creación y

1 Documentos científicos

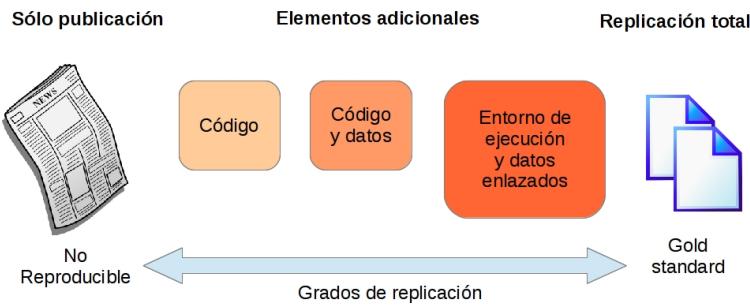


Figura 1.4: Espectro de los diferentes niveles de replicación en publicaciones científicas. Fuente: Peng (2011).

publicación de proyectos que permiten gestionar el código software que se ha creado, controlando los cambios y las versiones liberadas. Los servicios web integran, además, un buen número de herramientas para dar soporte a diferentes facetas del proceso de desarrollo de software, tales como generación de documentación, manuales y ejemplos, informes de error y peticiones de mejoras, integración continua y despliegue continuo (CI/CD), testeo sistemático del código generado, etc. Si todavía no te has planteado en qué puede beneficiarte utilizar una herramienta de control de versiones de código fuente, echa un vistazo a la Figura 1.5 en la que revivirás una situación lamentablemente muy habitual entre los investigadores y científicos que desarrollan soluciones software.

- Virtualización de software y contenedores: en un entorno tecnológico dominado por la contratación y despliegue de infraestructura de computación y servicios en arquitecturas en la nube (*cloud computing*), las herramientas de empaquetado y virtualización de aplicaciones y servicios software que pueden instalarse y desplegarse en poco tiempo han revolucionado la forma en la que se publican y gestionan los productos software, incluidos los de procesamiento y análisis de datos.
- Control de versiones de datos: De forma análoga a los SCV para código fuente, está apareciendo software para aplicar los mismos principios a los ficheros de datos. De esta forma, podemos controlar diferentes versiones de cada archivo de datos, modificaciones efectuadas en los mismos, etc. Una de estas herramientas es Data Version

1 Documentos científicos

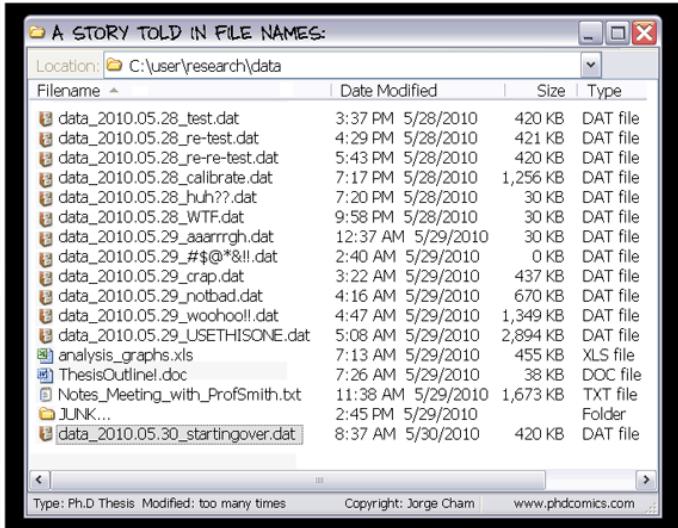


Figura 1.5: Control de versiones de software. Fuente: PhD Comics

Control (DVC), que permite versionado de datos y modelos. Como resultado, podemos saber en todo momento qué versión de los datos y qué listado de *features* se han incluido en cada modelo considerado durante el análisis, manteniendo integrada la información descriptiva sobre estos tres componentes esenciales que siempre deben ir cohesionados.

- Gestión de modelos y experimentos: Otro tipo de herramienta de gestión de proyectos de aprendizaje automático es la que permite la organización, seguimiento, comparativa y selección de los experimentos y modelos que hemos realizado. Uno de los ejemplos destacados más recientes es ML Flow, que proporciona soporte para ajuste, evaluación y optimización de modelos, despliegue de los mismos en entornos en producción, creación de un registro de modelos pre-entrenados, etc. Por supuesto, es posible combinar este tipo de herramientas con otras como DVC, creando como resultado un entorno integral de gestión de nuestros proyectos.
- Creación y gestión de *pipelines* de procesamiento de datos: el último elemento indispensable en todo proyecto de procesamiento y análisis de datos que deba cuidar la

1 Documentos científicos

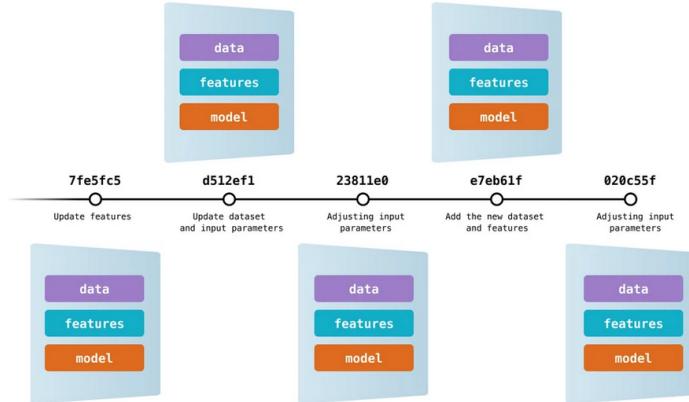


Figura 1.6: Ejemplo del versionado de datos, código y modelos que DVC es capaz de mantener. Fuente: Documentación DVC.

escalabilidad es una herramienta para creación y gestión de flujos de procesamiento y análisis o *pipelines* de datos. El conjunto de todos los *pipelines* de nuestro proyecto componen el *workflow* general del mismo. A estas herramientas se las conoce como *orquestradores* de datos o de flujos de trabajo. En esta categoría, contamos tanto con herramientas muy potentes y llenas de características como Apache Airflow o Prefect como con otras más sencillas y directas como Luigi.

Por supuesto, la comunidad de R no se ha mantenido ajena a estas nuevas tendencias, muy en particular la iniciativa R OpenSci, dentro de la cual encontramos muchos paquetes (publicados en el repositorio oficial CRAN) que cubren diversos aspectos del trabajo científico, incluyendo la gestión de *pipelines* y *workflows* mediante el paquete **targets**.

- Manual de uso del paquete R **targets**.

1.3 Quarto para publicaciones científicas

Ahora que ya conocemos el concepto fundamental sobre el que se asienta el funcionamiento de Quarto y su aplicación para conseguir un mayor nivel de reproducibilidad y transparencia en nuestro proceso científico, vamos a explicar con más detalle

1 Documentos científicos

el proceso que sigue Quarto para componer un documento. La Figura 1.7 presenta un esquema con el proceso de creación del documento y los elementos y herramientas que entran en juego para conseguirlo.

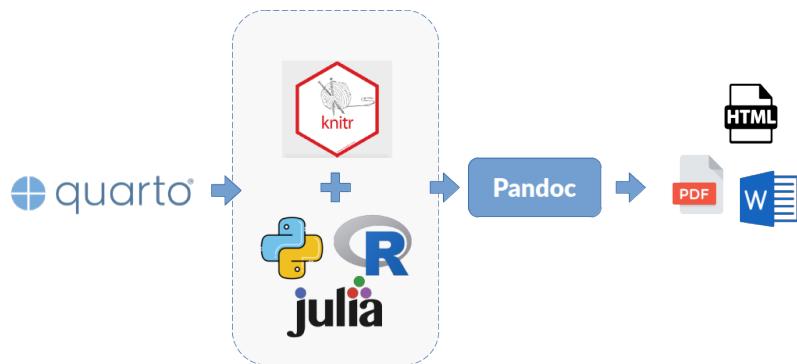


Figura 1.7: Proceso de creación de contenidos con Quarto.

- **Quarto:** un software que permite crear documentación científica siguiendo los principios de la programación literaria.
- **Knitr y lenguaje de programación:** el paquete knitr se encarga de la conexión con un intérprete de un lenguaje de programación (R, Python, Julia) que se pueda ejecutar en un entorno REPL, para poder ejecutar fragmentos de código software integrados en el documento y generar como resultado contenido en formato Markdown.
- **Markdown (contenido formateado):** lenguaje de marcado de contenido textual que permite formatear de forma sencilla la información de nuestros documentos creados con Quarto.
- **Pandoc (traductor universal de formatos documentales):** este software recibe el contenido ya formateado usando el estándar Markdown, para convertirlo en el tipo de salida seleccionado. Existen varias opciones disponibles: HTML, PDF o Word, así como diapositivas, websites o paneles interactivos (*dashboards*).

1.4 Instalación de Quarto

Para instalar la última versión del software Quarto en nuestro sistema, dirigimos nuestro navegador web a la página <https://quarto.org/install/>

1 Documentos científicos

quarto.org/docs/get-started/. Aquí, descargamos e instalamos el archivo correspondiente a nuestro sistema operativo.

En este momento, la última versión de Quarto disponible es la 1.5.57.

Requisitos para generar documentos en PDF

Por defecto, el formato de salida de los documentos generados con Quarto es HTML. Si queremos generar documentos en PDF, necesitamos tener instalada una distribución LaTeX. Para más información, consulta la Sección 3.3.

2 Tipos de documentos

En este capítulo, presentamos los principales tipos de documentos y colecciones de contenidos científicos que podemos generar con Quarto.

2.1 Documentos individuales

La forma más sencilla de trabajar con Quarto es crear un documento individual. Dicho documento podrá utilizar las secciones o *chunks* de código para leer datos de entrada o descargarlos de alguna fuente, procesarlos, analizarlos y mostrar los resultados. Se pueden añadir gráficos, tablas, ecuaciones, referencias bibliográficas y muchos otros elementos.

Los documentos tienen siempre una estructura estándar:

- *Preámbulo*: en el que se especifican opciones de configuración para la creación del documento con Quarto y sus herramientas asociadas.
- *Cuerpo*: la sección que alberga el contenido principal del documento, incluyendo secciones de texto formateado en Markdown y secciones de código ejecutable. El código software se podrá mostrar, si resulta de utilidad, o quedar oculto en el resultado final.
- *Referencias*: Al final del documento se incluyen las referencias bibliográficas, como es habitual en los textos científicos.

2.2 Libros

La evolución natural del caso anterior es reunir una colección de documentos individuales en un solo libro. *Quarto books* permite crear este tipo de documentos, estructurados en partes, capítulos y secciones. Las opciones de configuración permitirán confeccionar una portada de introducción para el sitio web que

2 *Tipos de documentos*

contiene los capítulos (un documento por capítulo) o bien los elementos necesarios para crear un libro en PDF, semejante a los publicados por una editorial.

2.3 Artículos y publicaciones

Uno de los resultados clave en todo proceso científico es la producción de artículos y publicaciones (informes técnicos, etc.) que recojan los resultados y avances científicos conseguidos. En este caso, Quarto también nos podrá ayudar, con la colaboración de otros elementos indispensables como el paquete `Rrticles`, que proporciona plantillas para generar artículos según las especificaciones de las principales publicaciones y editoriales científicas en multitud de campos de conocimiento.

2.4 Presentaciones

También es posible generar presentaciones (normalmente, en formato HTML) con diapositivas mediante Quarto. En este caso, tendríamos el soporte de varios paquetes y entornos de creación de presentaciones web a nuestra disposición, como `reveal.js` (HTML), Beamer (para LaTeX/PDF) o formato PPTX de MS Office.

Este caso no lo trataremos en este taller, pero se puede obtener más información en la guía online, disponible en <https://quarto.org/docs/presentations/>.

2.5 Sitios web

Otra opción que puede resultar interesante es crear sitios web personales (por ejemplo, para mostrar nuestro CV y una selección de trabajos destacados, publicaciones, etc.), blogs e incluso sitios web corporativos (organización, grupo de investigación) de forma rápida mediante Quarto. Existen numerosas plantillas gratuitas y de pago ya disponibles para crear sitios web con un aspecto armonizado, aunque necesitaremos aprender un poco de HTML y CSS para poder personalizar aún más nuestra web.

2 *Tipos de documentos*

Aquí tenemos un ejemplo de sitio web de un investigador en tecnología medioambiental creado con Quarto: <https://www.mm218.dev/>. Más ejemplos de diferentes tipos de sitios web generados con Quarto: <https://drganghe.github.io/quarto-academic-site-examples.html>.

Se puede conseguir más información y tutoriales para crear sitios web con Quarto en <https://quarto.org/docs/websites/>.

2.6 *Dashboards*

Por último, es posible crear cuadros de mandos o *dashboards* personalizados para monitorización de datos, análisis de modelos y resultados o bien para ejemplos y aplicaciones docentes utilizando Quarto, tal y como se describen en la guía <https://quarto.org/docs/dashboards/>.

En este caso podemos incluir entre las herramientas Shiny, un paquete software para R (también disponible para Python) con el que crear aplicaciones interactivas basadas en datos de forma rápida y sencilla.

3 Proceso de trabajo

En este apartado vamos a explicar algunos detalles más sobre el proceso de creación de los documentos en Quarto, para comprender mejor los componentes que intervienen en este proceso y las opciones de configuración que tenemos disponibles. La Figura 3.1 resume a alto nivel las fases de creación de un documento con Quarto.



Figura 3.1: Proceso de creación de contenidos con Quarto.

Fuente: RStudio.

3.1 Cadena de montaje de un documento

Podemos considerar el proceso de creación de un documento en Quarto como una cadena de montaje en la que varias herramientas software se aplican secuencialmente para producir el documento final en diferentes formatos de salida.

- Quarto: El primer agente que interviene en la interpretación del fichero con extensión .qmd es Quarto. El programa debe estar instalado previamente en nuestro equipo para que el entorno de desarrollo que decidamos emplear (RStudio, Microsoft VS Code, etc.) pueda encontrarlo. Quarto se encarga de interpretar el contenido del archivo y considerar las diferentes opciones de configuración que hemos insertado en la cabecera del documento, así como en los fragmentos de código ejecutable, para delegar las tareas de construcción de las distintas partes del documento a otras herramientas.

3 Proceso de trabajo

Quarto también se encarga de insertar de forma automática algunos elementos de autoría de documentos (como las llamadas de atención, explicadas en la Sección 4.5.4), mientras que la generación de otros elementos (numeración de figuras, tablas, citas bibliográficas, etc.) se delega a otras herramientas como Pandoc.

- Motor de ejecución de código (**engine**): Una de las grandes ventajas de la producción de documentos en Quarto es la posibilidad de insertar fragmentos de código ejecutable en nuestros documentos. Quarto puede trabajar con diferentes motores (*engines* en la terminología de Quarto) que interpretan el código y devuelven el resultado de su ejecución para integrarlo en el documento final.
 - Todos los bloques de código en lenguaje R utilizan el paquete **knitr** como *engine* de ejecución.
 - Los bloques de código ejecutable en otros lenguajes soportados en Quarto (Python, Julia, Observable) utilizan los *kernel* disponibles en la herramienta Jupyter para su ejecución, excepto en el caso explícito de que se combinen bloques de código en R y Python en el mismo documento.
 - En el caso particular de que un mismo documento combine bloques de código en R y Python, entonces se utiliza la herramienta Knitr para el código en R junto con el paquete **reticulate** de R para ejecutar los bloques escritos en Python. Esta combinación tiene la ventaja adicional de que es posible “trasvasar” variables y resultados entre los entornos de ejecución de R y Python, de forma que podamos usar variables y estructuras de datos creados en la parte de R dentro de nuestro código Python y viceversa.
- Markdown: El producto intermedio de todas las fases anteriores es un archivo en formato Markdown (extensión **.md**), que ya integra el contenido textual formateado y muchos de los resultados y elementos adicionales generados por las distintas herramientas que han entrado en juego hasta este punto. En caso de que fuese útil, este archivo intermedio se puede almacenar para ser revisado

3 Proceso de trabajo

o para utilizarlo para otros fines. A continuación, este archivo se envía a Pandoc para la última etapa de creación del documento.

- Pandoc: El proyecto Pandoc ofrece una herramienta software para traducir documentos entre diferentes formatos de representación de información. Como se puede ver rápidamente en la página principal del proyecto, la lista de formatos soportados es realmente extensa. A efectos prácticos, en Quarto se utiliza la capacidad de Pandoc para recibir como entrada un documento en formato Markdown y generar una salida en tres posibles formatos: HTML, DOC/DOCX o PDF. El archivo Markdown debe incluir una sintaxis específica para codificar determinados elementos (referencias cruzadas, citas bibliográficas, figuras, tablas, ecuaciones, etc.) que iremos presentando en los siguientes capítulos, para que Pandoc pueda interpretar estos elementos y representarlos de forma apropiada en cada formato de salida.

Por último, cabe destacar que se necesitan programas de visualización de documentos para cargar los documentos de salida, según el formato: navegador web (HTML), MS Word (archivos DOC/DOCX), visor PDF (archivos PDF).

3.2 Producir HTML

La opción por defecto para el formato de salida de los documentos Quarto es generar un documento HTML, que se puede visualizar con la mayoría de navegadores web modernos. Este formato de salida tiene varias ventajas:

- Es bastante probable que el receptor del documento ya disponga de uno o varios navegadores instalados en su sistema para ver el documento, si se lo enviamos o compartimos directamente.
- Es más sencillo publicar este tipo de documentos en la Web, utilizando alguna de las distintas plataformas disponibles para este fin:
 - Quarto Pub, un servicio de publicación de documentos que permite que sean accesibles públicamente.

3 Proceso de trabajo

- Los documentos individuales (véase Capítulo 4) y los libros o colecciones de documentos (véase Capítulo 5) se pueden publicar de forma sencilla y rápida en sitios de hospedaje de proyectos software como GitHub o GitLab, que además proporcionan servicios de control de versiones, gestión de informes de error/mejoras, documentación, testing, etc.

3.3 Producir PDF

Al contrario que en HTML, cuando generamos documentos en PDF se añade un paso adicional de compilación del documento a final de toda la cadena de montaje, utilizando LaTeX y el motor de compilación XeLaTeX para generar la salida PDF. Por tanto, si seleccionamos esta opción de salida es **imprescindible** tener **instalada una distribución TeX/LaTeX** previamente en nuestro sistema, para compilar y generar los documentos. Si no disponemos de ninguna todavía, se puede instalar TinyTeX, una distribución ligera de TeX Live que tiene mucho menor tamaño (~100 MB frente a los más de 4 GB de TeX Live completa).

3.3.1 Personalizar documentos PDF

Se pueden utilizar plantillas de documentos LaTeX predefinidas. Por defecto, Quarto utiliza varias plantillas de la colección de paquetes LaTeX koma-script.

Algunas de estas plantillas pueden funcionar de forma relativamente sencilla en Quarto, mientras que otras requieren cierta adaptación, para lo cual se necesitarán algunos conocimientos sobre LaTeX. Este es probablemente un tema más avanzado para muchos usuarios, por lo que por ahora no lo vamos a tratar en este taller introductorio.

No obstante, a modo de ejemplo, ofrecemos a continuación un listado de algunos ejemplos que ilustran las enormes posibilidades de este tipo de plantillas:

- El profesor R.J. Hyndman ha publicado plantillas de documentos Quarto de la Monash University, que pueden utilizarse como punto de partida para personalizarlas en nuestros propios proyectos.

3 Proceso de trabajo

- El repositorio Awesome Quarto Thesis recoge un listado de plantillas Quarto para generar memorias de TFG/TFM y tesis doctorales para algunas universidades. También se enlaza una plantilla genérica de extensión para Quarto, concebida para facilitar que otros usuarios puedan personalizarla según los criterios marcados por su propia institución para generar estos trabajos.

4 Documentos individuales

La manera más sencilla de comenzar a utilizar Quarto es crear documentos individuales. Se trata de documentos autocontenidos, que incorporan texto formateado y código ejecutable en un único archivo.

Para crear un documento nuevo con Quarto, simplemente podemos usar las opciones del menu de RStudio o MS Visual Code, o bien crear un archivo con extensión `.qmd`.

4.1 Creación del documento con RStudio

Antes de empezar, comprueba que has instalado el software Quarto en tu máquina. Es un programa software independiente, que tiene que estar instalado para que el resto del proceso funcione (consulta la sección Sección 1.4).

Si ya tenemos instalada una versión reciente de RStudio, necesitaremos instalar los siguientes paquetes para el ejemplo:

```
install.packages("tidyverse")
install.packages("palmerpenguins")
install.packages("quarto")
```

Ahora, en RStudio creamos un nuevo proyecto eligiendo la opción *Quarto project*, tal y como aparece en la Figura 4.1.

Podemos nombrar el directorio de nuestro proyecto como **primer-ejemplo** y pulsamos **Create Project**.

Como resultado, nos debe aparecer un nuevo proyecto abierto en pantalla, con el aspecto que se muestra en la Figura 4.2.

En concreto, en el panel superior izquierdo podemos comprobar que, por defecto se ha abierto el editor *Visual*, que permite crear documentos Quarto de forma más intuitiva. Sin embargo, para empezar a familiarizarnos desde el principio con la estructura

4 Documentos individuales

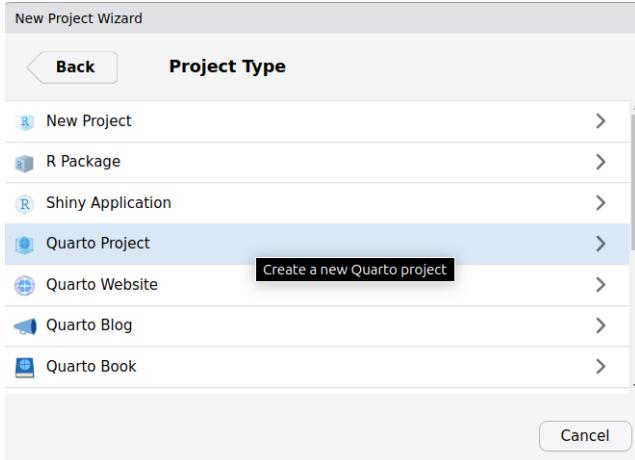


Figura 4.1: New Quarto project

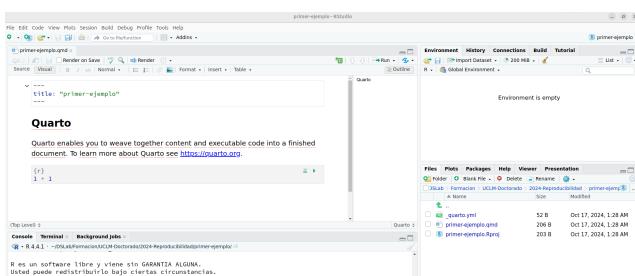


Figura 4.2: First example

4 Documentos individuales

de un documento en quarto vamos a cambiar al editor *Source* para ver el código fuente, pulsando en el botón que se muestra en la figura Figura 4.3.

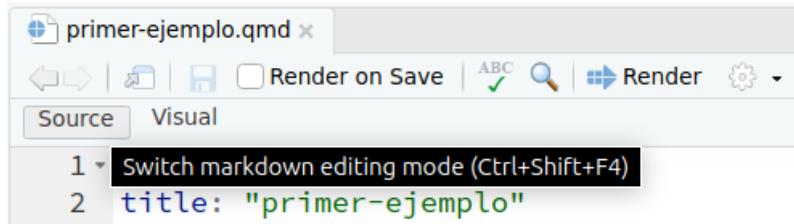


Figura 4.3: Source editor

4.2 Estructura del documento

La estructura de un documento individual en Quarto es esta.

```
---
title: "Mi primer documento"
author: John Doe
date: 2024-10-24
---

Aquí tenemos algo de contenido en formato
↳ **Markdown**.

```{r}
#| label: Etiqueta

1 + 1
```

Contenido *adicional* en Markdown.
```

El contenido del archivo consta de dos partes:

- **Preámbulo:** está delimitado por dos marcas ---. Dentro de esta área podemos asignar valores a opciones de configuración para maquetar y crear el documento, tales como el título, autor/es, fecha, etc. También podemos configurar diversas opciones relacionadas con el formato de salida de los documentos.

4 Documentos individuales

- **Cuerpo del documento:** se compone de párrafos de texto formateados utilizando la sintaxis de marcado Markdown, que veremos después. Además, también se pueden intercalar en el texto fragmentos de código ejecutable o *chunks*, que se marcan siguiendo una sintaxis especial (como vemos en el ejemplo anterior).

Cada *chunk* de código ejecutable está delimitado de la siguiente manera

```
```{r}
Código en R
```
```

Soporte para otros lenguajes de programación

Aunque en este taller nos centramos en el lenguaje R, debemos saber que Quarto también soporta otros lenguajes de programación como Python, Julia u Observable.

Podemos cambiar el lenguaje de programación de cada *chunk* indicando su nombre al comienzo, como por ejemplo:

```
```{python}
Código en Python
```
```

Sin embargo, para que pueda funcionar necesitaremos realizar algunas tareas adicionales de configuración.

4.2.1 El preámbulo

Un ejemplo básico de preámbulo es el siguiente (aunque bastaría con proporcionar únicamente un título para el documento):

```
---
title: "Mi primer documento"
author: John Doe
date: 2024-10-24
---
```

4 Documentos individuales

Por supuesto, se pueden ir añadiendo muchas más opciones que iremos explicando a lo largo de este capítulo.

4.2.2 Listado de opciones

Existe un extenso listado de opciones de configuración que podemos incluir en nuestros documentos.

- *Opciones para salida HTML*: permiten configurar diversos aspectos básicos del documento, tales como el título y subtítulo, fecha, autor (o lista de autores), resumen o DOI; opciones de formato como el tema o estilos avanzados para contenido HTML con CSS; numeración y tabla de contenidos, etc.
 - Opciones básicas para HTML con Quarto.
 - Lista completa de opciones HTML con Quarto.
- *Opciones para salida PDF*: ofrecen la posibilidad de configurar múltiples parámetros para la creación del documento en este formato, muchas de ellas similares a las de la salida en HTML. Una opción particularmente relevante es elegir el formato de documento LaTeX (opción `documentclass`), que define el aspecto general de la maquetación que se va a emplear. Por defecto, se emplean clases del metapakuete KOMA Script, como `scrartcl` o `scrbook`. También es importante indicar la opción `papersize`, en nuestro caso para garantizar que se usa un formato estándar como el A4. El formato de las citas también es relevante, pudiendo elegir, por ejemplo, el motor BibLaTeX que es más potente, con soporte multilenguaje y para codificación de caracteres UTF-8 nativa. Por último, también es importante indicar el motor de compilación. Si queremos una flexibilidad total en la maquetación del documento, se recomienda encarecidamente usar el motor XeLaTeX (opción `pdf-engine: xelatex`), que es el *valor por defecto* que utiliza Quarto.
 - Opciones básicas para PDF con Quarto.
 - Lista completa de opciones disponibles en PDF con Quarto.

4.2.3 Sintaxis Markdown básica

En el siguiente enlace puedes encontrar un rápido tutorial básico que muestra las opciones básicas de la sintaxis Markdown aceptada en documentos Quarto para formatear el contenido textual.

- Guía básica de sintaxis markdown.

4.3 Creación del documento (*output*)

Por defecto, si no indicamos nada Quarto generará un solo formato de salida del documento en HTML. Sin embargo, es posible definir más de un formato de salida incluyendo más opciones de configuración. Por supuesto, se pueden indicar diferentes opciones para generar varios formatos de salida simultáneamente, o bien para elegir el formato de salida que queremos producir en función de nuestros intereses, seleccionando el formato que necesitamos al previsualizar o al generar el documento definitivo.

4.3.1 Previsualización

Para previsualizar el documento tenemos que pulsar el botón *Render* en el menú de herramientas de la interfaz de RStudio, tal y como se muestra en la Figura 4.4.

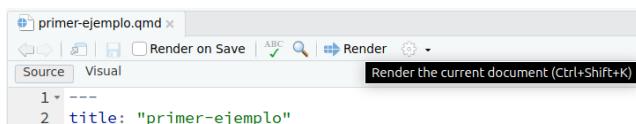


Figura 4.4: Botón *Render* para previsualizar el documento generado.

Por defecto, lo normal es que se abra nuestro navegador web principal o un panel en la interfaz de RStudio mostrando la página HTML con el documento ya generado. Pulsando sobre el icono con un engranaje junto al botón *Render* se puede seleccionar, entre otros aspectos, el tipo de previsualización que queremos que se lance tras completar la creación del documento o desactivar por completo dicha previsualización. La opciones disponibles se muestran en la Figura 4.5

4 Documentos individuales

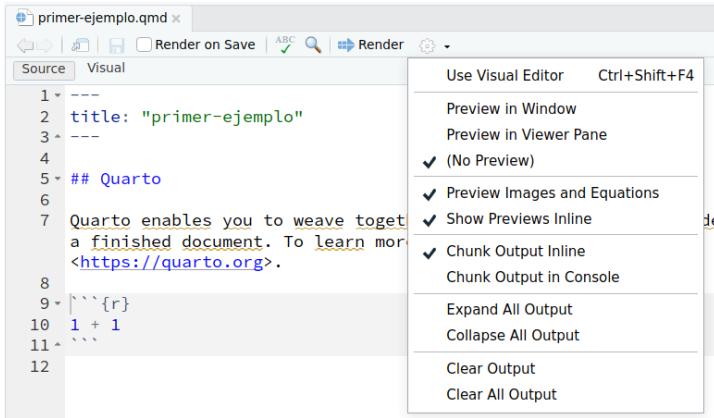


Figura 4.5: Opciones de previsualización de la operación *Render*.

4.3.2 Seleccionar el tipo de documento

Cuando tenemos varias opciones de formato de salida configuradas en nuestro documento, podemos elegir en tiempo de previsualización cuál de los formatos se elige para generar el documento. En la Figura 4.6 se puede observar un ejemplo de documento que incluye configuración para dos formatos de salida (HTML y PDF) y el cambio en el botón *Render*, en el que ahora aparece una pequeña flecha negra justo a la derecha del ícono del botón para desplegar las dos opciones de salida disponibles.

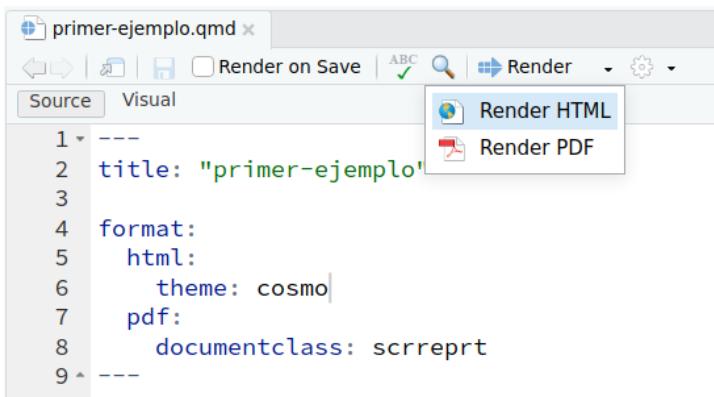


Figura 4.6: Selección de formatos de salida con la operación *Render*.

4.3.3 Opciones básicas de configuración

A continuación, se presenta un ejemplo de algunas opciones básicas de configuración que suelen ser habituales en documentos con formato de salida HTML.

```
---
title: "Mi primer documento"
author:
  - "John Doe"
  - "Mary Jane"
date: 2024-10-24

lang: es
bibliography: references.bib

format:
  html:
    theme: cosmo
    toc: true
    number-sections: true
    html-math-method: kate
    css: styles.css
  pdf:
    documentclass: scrreprt
---
RESTO DEL DOCUMENTO
```

En este ejemplo se indica, además del autor y la fecha, una lista de dos autores, el lenguaje principal del documento (español), el archivo de referencias de bibliografía (en formato `.bib`) y ya dentro de las opciones HTML, el tema de maquetación, la inclusión de una tabla de contenidos (por defecto situada en la parte superior derecha), numeración de secciones, selección del motor para renderizar ecuaciones en el documento y un archivo de estilos personalizados en formato CSS para ajustar algunas opciones finas de maquetación.

Una opción que conviene destacar es la de forzar a que todos los recursos (imágenes, información de estilos, etc.) estén integrados en el propio archivo HTML, para facilitar la compartición o publicación directa del documento sin necesidad de aportar

4 Documentos individuales

también los archivos auxiliares necesarios para mostrarlo en el navegador. Esta opción se muestra a continuación:

```
format:  
  html:  
    embed-resources: true
```

4.4 **Chunks** de código ejecutable

La característica más diferencial de los documentos creados con Quarto es la posibilidad de intercalar fragmentos de código ejecutable, llamadas *chunks* en el propio documento. Esto incluye también la opción de que dicho código genere diferentes resultados (numéricos, gráficos, tablas, animaciones, etc.) que se integren directamente en el documento. De este modo, si mantenemos actualizado el código siempre se generarán las versiones corregidas de dichos resultados.

Los fragmentos de código ejecutable tienen la siguiente estructura:

```
```{r}  
#| label: id-fragmento

Aquí va el código ejecutable
a = c(1, 2, 3, 4)
b = a^2
~~~
```

La triplete de caracteres ``` se denomina *fence* y delimita el comienzo y el final del fragmento de código. Justo a continuación del delimitador de apertura se escribe entre llaves el identificador del lenguaje de programación en el que está escrito el código de ese fragmento. Esa información se usa para elegir el resultado de sintaxis apropiado para mostrar el código de ese lenguaje y para seleccionar el intérprete que ejecuta el código y produce los resultados.

En las siguientes líneas podemos incluir una o varias **opciones de configuración** específicas para ese fragmento de código, mediante la sintaxis #| opcion: valor. Por ejemplo, en el fragmento anterior la opción #| label: id-fragmento crea

#### 4 Documentos individuales

una etiqueta (que debe ser única) para identificar a ese fragmento de código dentro del documento.

- Lista de opciones para fragmentos de código.

Algunas opciones de uso frecuente son:

- **eval: true | false | [...]**: Indica si se debe evaluar (ejecutar) el contenido de ese fragmento. Se puede pasar una lista de números de línea positivos o negativos para seleccionar explícitamente qué líneas de código se incluyen (positivos) o excluyen (negativos) de la ejecución.
- **echo: true | false | fenced | [...]**: Indica si se debe incluir el código fuente del fragmento en el documento o no. La opción **fenced** incluye también el delimitador de celda como parte de la salida. Por último, también acepta una lista de números de línea positivos o negativos para seleccionar qué líneas de código se mostrarán o no en el fragmento.
- **output: true | false | asis**: Para decidir si el resultado de la ejecución del código se incluye o no en el documento. El valor **asis** fuerza a que el resultado se trate como contenido Markdown en crudo.
- **warning: true | false**: Indica si se deben incluir los mensajes de aviso en la salida.
- **error: true | false**: Marca si los mensajes de error generados se incluyen en la salida.
- **message: true | false**: Indica si los mensajes de información generados se incluyen en la salida.

Cuando los fragmentos generan figuras, estas se insertan dentro del propio documento. Veamos un ejemplo:

```
```{r}
#| label: fig-example-cars
#| fig-cap: "Gráfico de correlación lineal positiva
  ↳ entre el kilometraje en ciudad y en carretera
  ↳ de diferentes modelos de coches."

library(ggplot2)
#| label: scatterplot
#| echo: true

ggplot(mpg, aes(x = hwy, y = cty, color = cyl)) +
```

#### 4 Documentos individuales

```
geom_point(alpha = 0.5, size = 2) +  
  scale_color_viridis_c() +  
  theme_minimal()  
~~~
```

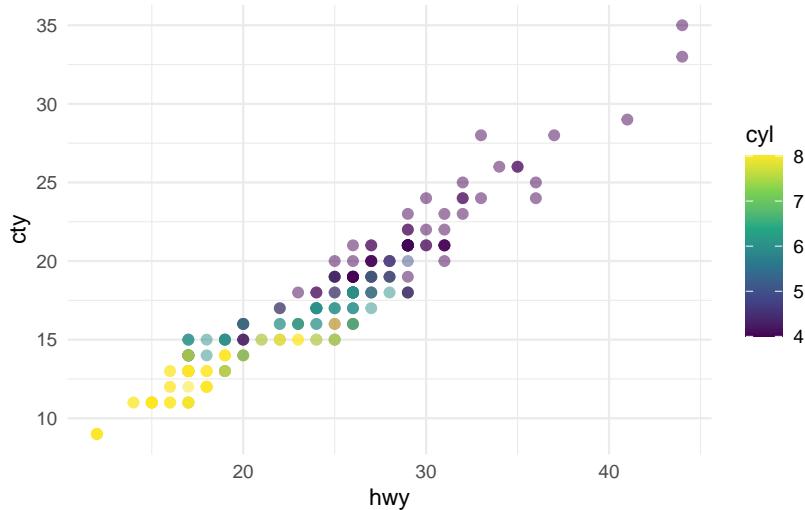


Figura 4.7: Gráfico de correlación lineal positiva entre el kilometraje en ciudad y en carretera de diferentes modelos de coches.

#### 💡 Numeración automática de figuras

Es importante que *el identificador de fragmento* que elegimos para código que genera una o varias figuras *comience por el prefijo fig-*. De ese modo, nos aseguramos de que Quarto le asigne automáticamente una numeración a la figura generada y que podamos crear referencias cruzadas (enlaces internos) a dicha figura en nuestro documento.

Como veremos más adelante, otros tipos de salida como las tablas también necesitan que se les asigne un patrón concreto en su identificador de fragmento para que se numeren de forma automática y se puedan referenciar dentro del documento.

La gestión de figuras en Quarto es bastante sofisticada, hasta el punto de que se pueden organizar de forma sencilla varias subfiguras con sus respectivas descripciones individuales, como

#### 4 Documentos individuales

se muestra en el siguiente ejemplo usando algunas opciones adicionales.

```
```{r}
#| label: fig-mpg-subplot
#| fig-cap: "Kilometraje en ciudad y en carretera
#|   de 38 modelos populares de coches."
#|
#| fig-subcap:
#|   - "Color por núm. de cilindros."
#|   - "Color por cubicaje del motor, en litros."
#| layout-ncol: 1

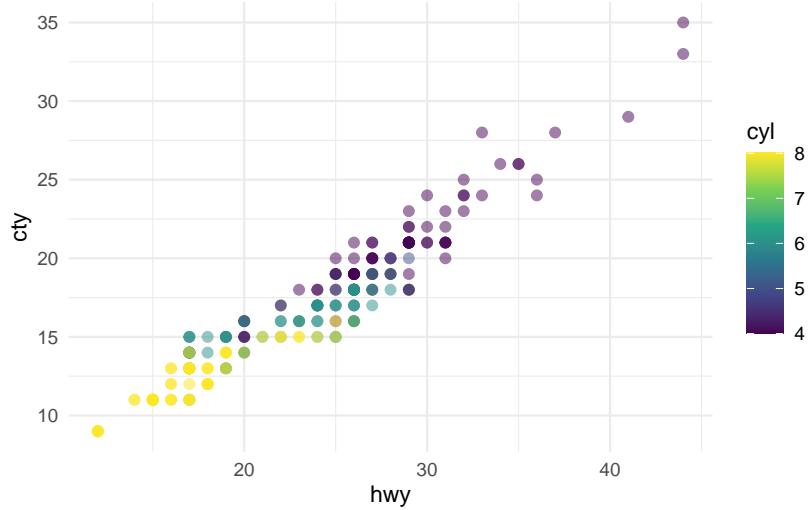
ggplot(mpg, aes(x = hwy, y = cty, color = cyl)) +
  geom_point(alpha = 0.5, size = 2) +
  scale_color_viridis_c() +
  theme_minimal()

ggplot(mpg, aes(x = hwy, y = cty, color = displ)) +
  geom_point(alpha = 0.5, size = 2) +
  scale_color_viridis_c(option = "E") +
  theme_minimal()
```
```

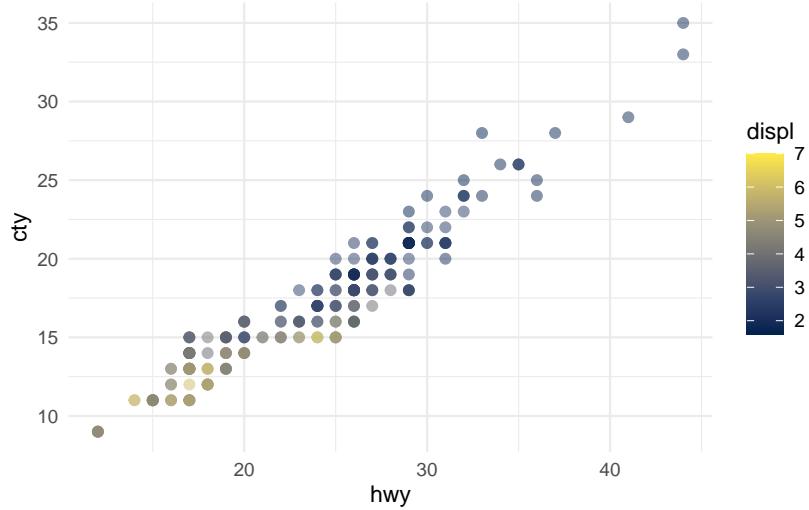
Algunas opciones frecuentes para *chunks* que generan figuras son:

- **fig-width**: Ancho de la figura.
- **fig-height**: Alto de la figura.
- **fig-cap**: String entre comillas que se insertará como descripción al pie de la figura (*caption*).
- **fig-alt**: Mensaje de texto alternativo que rellena el atributo **alt** de la imagen HTML (por ejemplo, para mejorar la accesibilidad del contenido).
- **fig-dpi**: Ajuste de la resolución de la figura (en puntos por pulgada).

El tutorial sobre fragmentos de código ejecutables de la documentación oficial presenta más información y ejemplos sobre cómo utilizar esta potente característica de Quarto.



(a) Color por núm. de cilindros.



(b) Color por cubicaje del motor, en litros.

Figura 4.8: Kilometraje en ciudad y en carretera de 38 modelos populares de coches.

## 4.5 Herramientas para el autor

Además de la capacidad de integrar código ejecutable y sus resultados en nuestros documentos científicos, Quarto incluye un buen número de recursos y herramientas para ofrecer una experiencia de autoría completa y eficiente.

### 4.5.1 Secciones del documento

Como ya vimos en el ejemplo de la Sección 4.3.3, existen dos opciones de configuración del documento HTML que nos permiten numerar las secciones e incorporar una tabla de contenidos generada automáticamente en la parte superior derecha de nuestro documento.

```
format:
 html:
 toc: true
 number-sections: true
```

Una funcionalidad importante para la creación de documentación científica es poder incluir **referencias cruzadas**, es decir, enlaces que nos lleven a otras secciones del documento. En Quarto esto se consigue siguiendo un sencillo procedimiento en dos pasos:

1. Añadimos una *etiqueta única* para identificar la sección con la sintaxis:

```
Título de sección {#sec-etiqueta}
```

2. Referenciamos en otro lugar del texto la etiqueta que hemos creado para esa sección, de forma que Quarto crea automáticamente el enlace (referencia cruzada) a dicha sección:

```
En el texto añadimos una referencia a la
 ↪ @sec-etiqueta
```

#### 4 Documentos individuales

Se puede ver un ejemplo de este tipo de referencias cruzadas creadas de forma automática al comienzo de esta misma sección. Por el contrario, si queremos que una sección del documento se excluya del esquema de numeración del resto de secciones, usamos en el título de esa sección la etiqueta especial:

```
Sección no numerada {.unnumbered}
```

Existen varias opciones adicionales que controlan la forma y estilo con que se crean y numeran las secciones. Algunas de ellas son:

- **anchor-sections**: Hace que se muestre un enlace de anclado (para enlazar directamente esa sección en otro documento) cuando se pasa el ratón por encima del título de una sección.
- **toc-depth**: Especifica cuántos niveles de profundidad en la numeración de secciones aparecen en la tabla de contenidos. Por defecto se muestran 3 niveles.
- **toc-location**: `body | left | right | left-body | right-body`: Controla la ubicación en la que aparece la tabla de contenidos en el documento.
- **toc-title**: Cadena de caracteres con el título de la tabla de contenidos.
- **toc-expand**: Indica si se deben expandir todas las secciones de la tabla de contenidos o deben quedar colapsadas para que el usuario vaya pulsando en las que quiera expandir.
- **number-depth**: Determina la profundidad máxima a la que se numeran las secciones del documento (cuidado, debería estar en consonancia con el valor asignado a la opción `toc-depth`).
- **number-offset**: Permite ajustar el número por el que se empiezan a numerar las secciones. Si queremos que el documento comience a numerar la sección de más alto nivel como “4” entonces usamos `number-offset: 3`. Si queremos que el documento empiece en una sección de nivel 2 con numeración “1.5” debemos especificar `number-offset: [1,4]`. Definir un valor para esta opción implica que automáticamente `number-sections: true`.

### 4.5.2 Ecuaciones

Otro aspecto esencial de los documentos científicos es la aparición de símbolos, fórmulas y ecuaciones matemáticas. Existen varias librerías para HTML que permiten mostrar las ecuaciones debidamente formateadas en pantalla. Por su parte, LaTeX, debido a sus orígenes, siempre ha incluido potentes y versátiles herramientas para manejar este tipo de contenido, por lo que el soporte está garantizado para documentos en PDF.

En general, la sintaxis empleada para escribir las ecuaciones es muy similar a la que se utiliza en LaTeX.

- Tutorial sobre expresiones matemáticas en LaTeX.
- Resumen de sintaxis matemática en LaTeX.

Existen dos maneras de mostrar las ecuaciones en nuestro contenido, siguiendo también una filosofía similar a la de los documentos LaTeX:

- Ecuaciones en línea con el texto: para mostrar la ecuación dentro de una línea o párrafo, a la misma altura que el resto del texto.
- Ecuaciones en modo *display*: la ecuación se muestra en un espacio aparte, entre dos párrafos de texto y con cierto margen de espacio en la zona superior e inferior.

Ejemplo de ecuación en línea: `$F = m \cdot a$`

Que produce como resultado: ejemplo de ecuación en línea:  $F = m \cdot a$ .

Ejemplo de ecuación en modo display:

`$$E = mc^2$$`

Que genera el siguiente resultado (ver debajo cómo añadir la numeración):

$$E = mc^2 \quad (4.1)$$

Si además queremos numerar nuestras ecuaciones, hay que acordarse de emplear el patrón de etiqueta identificadora única

#### 4 Documentos individuales

`eq-etiqueta` para identificarla y luego poder insertar referencias internas a dicha ecuación en el texto.

```
$$ E = mc^2 $$ {#eq-energy}
```

Como resultado, podemos insertar una a la Ecuación 4.1.

#### 4.5.3 Tablas

Las tablas son otro contenido relevante que podemos formatear de diferentes maneras en los documentos generados con Quarto.

- Introducción a la creación de tablas en Quarto.

En este caso, el editor visual nos puede simplificar mucho esta tarea. Se aconseja probarlo para ver la diferencia, ya que es una herramienta muy intuitiva. Sin embargo, siguiendo la misma línea que el resto del taller, aquí vamos a describir los detalles para crear este contenido directamente en el código Markdown del archivo.

La forma más directa de crear una tabla en Markdown es componer una tabla `pipe`, así denominada porque su sintaxis se basa en el operador `|` de la línea de comandos. Veamos un ejemplo.

```
Default	Left	Right	Center
1	2	3	4
22	23	24	25
4	3	2	1
```

El resultado de incluir el código anterior en nuestro documento es:

| Default | Left | Right | Center |
|---------|------|-------|--------|
| 1       | 2    | 3     | 4      |
| 22      | 23   | 24    | 25     |
| 4       | 3    | 2     | 1      |

#### 4 Documentos individuales

Podemos observar cómo la clave para controlar el alineamiento horizontal del contenido de la tabla es situar apropiadamente el símbolo : en la línea justo debajo de la línea de título, que separa este del cuerpo de la tabla. Si no queremos incluir título es obligatorio que incluyamos la primera línea, pero podemos dejar las celdas en blanco.

Debajo de la tabla podemos insertar la expresión : `Caption de la tabla` para incluir un mensaje descriptivo. También es posible utilizar directamente algunos elementos de estilo incluidos en las clases de Bootstrap, el entorno de estilo web que utiliza Quarto para componer las páginas (hemos visto antes cómo utilizar la opción de documento `theme: cosmo` para usar el tema Cosmo de Bootstrap). Existen diferentes efectos, y uno de los más frecuentes es colorear en gris el fondo de las filas alternativamente así como resaltar la fila en la que está posada la flecha del ratón. Estos dos efectos son `.striped` y `.hover`, respectivamente.

```
Default	Left	Right	Center
1	2	3	4
22	23	24	25
4	3	2	1
```

: `Caption de la tabla {.striped .hover}`

Tabla 4.2: `Caption de la tabla`

| Default | Left | Right | Center |
|---------|------|-------|--------|
| 1       | 2    | 3     | 4      |
| 22      | 23   | 24    | 25     |
| 4       | 3    | 2     | 1      |

Por último, de forma análoga a lo que hacemos para referenciar internamente ecuaciones y figuras en nuestro documento, también podemos etiquetar las tablas utilizando el patrón `#tbl-etiqueta` para referenciarlo como `@tbl-etiqueta` que queda formateado así: Tabla 4.3.

```
Default	Left	Right	Center
```

#### 4 Documentos individuales

|    |    |    |    |  |
|----|----|----|----|--|
| 1  | 2  | 3  | 4  |  |
| 22 | 23 | 24 | 25 |  |
| 4  | 3  | 2  | 1  |  |

```
: Caption de la tabla. {#tbl-etiqueta .striped .hover}
```

Tabla 4.3: Caption de la tabla.

| Default | Left | Right | Center |
|---------|------|-------|--------|
| 1       | 2    | 3     | 4      |
| 22      | 23   | 24    | 25     |
| 4       | 3    | 2     | 1      |

El mismo patrón de etiqueta se debe emplear en la opción de identificación de *chunks* de código `#| label: tbl-etiqueta` si luego queremos referenciar la tabla generada por ese fragmento de código con una referencia cruzada.

Se pueden consultar más detalles sobre la creación de subtablas, cambio de ubicación del caption, así como la creación de tablas *grid* que usan una sintaxis diferente y permiten incluir elementos de bloque arbitrarios en cada celda (múltiples párrafos, bloques de código, listas no numeradas o numeradas, etc.).

| Fruta   | Precio | Ventajas                          |
|---------|--------|-----------------------------------|
| Bananas | \$1.34 | - envoltorio<br>- color brillante |
| Oranges | \$2.10 | - rica en vitam. C<br>- saborsa   |

```
: Sample grid table.
```

Tabla 4.4: Tabla *grid* de ejemplo.

| Fruta   | Precio | Ventajas                                                                                             |
|---------|--------|------------------------------------------------------------------------------------------------------|
| Bananas | \$1.34 | <ul style="list-style-type: none"> <li>• envoltorio</li> <li>• color</li> <li>• brillante</li> </ul> |
| Oranges | \$2.10 | <ul style="list-style-type: none"> <li>• rica en vitam. C</li> <li>• saborsa</li> </ul>              |

#### 4.5.4 Llamadas

Es posible incluir bloques de llamadas de atención, para resaltar notas prácticas, advertencias o consejos de especial interés. Además, se suele poner un título a la llamada para hacerla aún más informativa.

```
::: {.callout-note}
Título de la nota

Existen cinco tipos diferentes de llamaedas:
`note`, `tip`, `warning`, `caution`, e `important`.
:::
```

##### **i** Título de la nota

Existen cinco tipos diferentes de llamaedas: **note**, **tip**, **warning**, **caution**, e **important**.

- Introducción al uso de llamadas en Quarto.

#### 4.5.5 Citas bibliográficas

La gestión de referencias bibliográficas en Quarto se realiza codificando la información en formato BibTeX. Esto permite utilizar cualquiera de los formatos de citas bibliográficas soportados por este paquete, o bien incluir un fichero CLS que defina un formato estándar (APA, Chicago, IEEE, etc.).

Por ejemplo, las opciones de documento

#### 4 Documentos individuales

```

```

```
title: "My Document"
bibliography: references.bib
csl: nature.csl

```

indican un fichero `references.bib` donde podemos almacenar la información sobre las referencias bibliográficas (que podemos conseguir de Google Scholar, Zotero u otras herramientas y servicios en Internet), así como un fichero de estilo de citas `nature.cls` (estilo definido por la editorial Nature).

- Repositorio CLS con estilos de citas.
- Repositorio Zotero con estilos de citas.

Dependiendo del estilo y formato de la cita, podemos utilizar una u otra sintaxis para indicar el autor y el año entre paréntesis, el autor fuera del paréntesis, números de página, capítulos, etc.

- Tabla de referencia de sintaxis para citas en Quarto.

Por último, la lista ordenada de referencias bibliográficas (según los criterios de estilo de citas que hayamos seleccionado) deben aparecer al final del documento. Para conseguirlo en un documento HTML, debemos incluir un código especial, que normalmente se coloca en una sección independiente y no numerada, tal y como se muestra en la Figura 4.9.

```
1 # References {.unnumbered}
2
3 ::: {#refs}
4 :::
```

Figura 4.9: Sintaxis para mostrar las referencias bibliográficas al final del documento.

Cuando la salida generada es en formato PDF y se usan los motores de gestión de referencias BibLaTeX o natbib, entonces la lista de referencias siempre aparece al final de documento y la etiqueta anterior se ignora. Finalmente, en el poco frecuente caso de que no queramos incluir ninguna referencia bibliográfica en nuestro documento podemos incluir en los metadatos de la cabecera del mismo la opción `suppress-bibliography: true`.

#### 4.5.6 Estilo general del documento

Hasta el momento, el documento de ejemplo que hemos mostrado así como estos mismos apuntes utilizan siempre un formato de estilo o `theme` del entorno de desarrollo web Bootstrap, llamado `cosmo`. No obstante, existe una amplia lista de temas alternativos para modificar el estilo general de nuestro documento (combinación de colores, tipografía y tamaño de fuentes, organización del contenido, aspecto de los enlaces, etc.). El proyecto Quarto se encarga de combinar regularmente los temas de estilo más populares para que estén disponibles como opción del documento.

En este directorio de temas en GitHub se puede comprobar una lista actualizada de los posibles valores que podemos asignar a la opción `theme` en la cabecera del documento. Es útil experimentar con diversas opciones para encontrar la que más se ajuste al tipo de documento generado, a su contenido y a la audiencia a la que va dirigido.

En la página web <https://bootswatch.com/> se puede acceder a una demo en línea de muchos de los temas disponibles.

## **Parte II**

# **Libros con Quarto**

# 5 Libros

Ahora que ya conocemos todos los elementos básicos para crear documentos individuales de programación literaria con Quarto, la evolución natural es preguntarnos si podemos gestionar colecciones ordenadas de documentos en un solo proyecto para, por ejemplo, crear un libro, un cuaderno de laboratorio o una bitácora de experimentos.

Los proyectos de tipo `book` en Quarto son la respuesta a estas necesidades, permitiendo agrupar y organizar varios documentos individuales en un solo sitio web o un solo volumen (PDF) para su publicación.

## 5.1 Creación de proyectos de tipo `book`

El primer paso consiste en la creación de un proyecto de tipo `book`, utilizando la IDE de nuestra elección, por ejemplo, RStudio. La Figura 5.1 muestra la interfaz de RStudio para crear un nuevo proyecto de tipo `book` con Quarto. Como es habitual, seleccionamos un nombre para el directorio que almacena el proyecto tal y como se muestra en la Figura 5.2.

- Página de inicio del tutorial sobre libros en la guía de Quarto.

## 5.2 Opciones de configuración

Al contrario que en los documentos individuales, en los proyectos como los de tipo `book` que agrupan varios documentos (archivos `.qmd`) en una sola colección podemos indicar opciones globales de configuración para todo el proyecto en un archivo independiente `_quarto.yml`, que debe estar presente en el directorio raíz de nuestro proyecto.

## 5 Libros

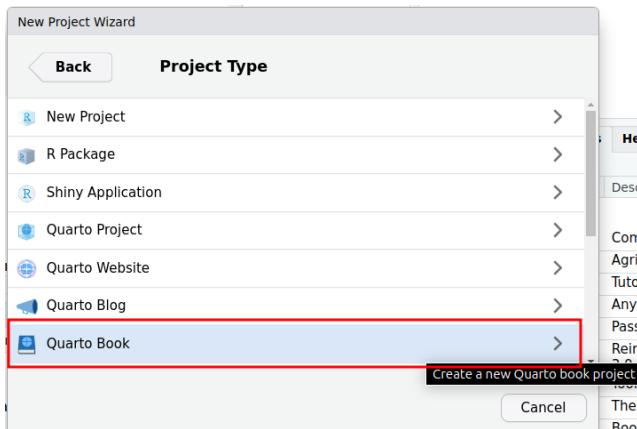


Figura 5.1: Interfaz en RStudio para la creación de un nuevo proyecto Quarto de tipo *book*.

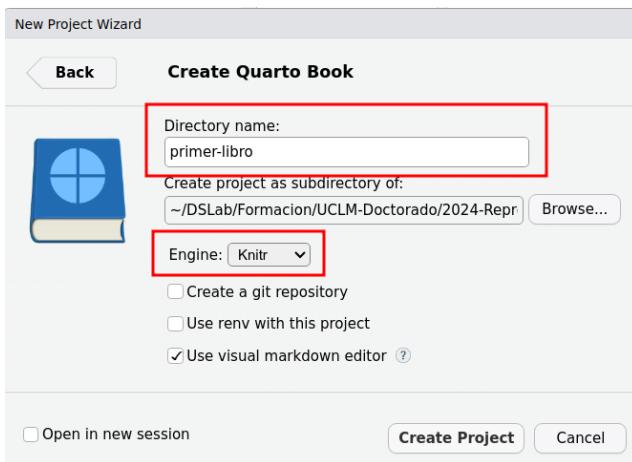


Figura 5.2: Diálogo para indicar el nombre del directorio que almacena el proyecto de tipo *book* y algunas opciones básicas de configuración, como el motor de ejecución de los bloques de código integrados en los documentos (resaltado en rojo).

## 5 Libros

Por defecto, las opciones de configuración que presenta este archivo en un proyecto de este tipo recien creado son las siguientes.

```
project:
 type: book

book:
 title: "primer-libro"
 author: "Norah Jones"
 date: "22/10/2024"
 chapters:
 - index.qmd
 - intro.qmd
 - summary.qmd
 - references.qmd

 bibliography: references.bib

format:
 html:
 theme: cosmo
 pdf:
 documentclass: scrreprt

editor: visual
```

Los nuevos elementos en este listado de opciones son:

- **project -> type: book**: Permite indicar el tipo de proyecto que queremos montar para que Quarto ajuste su comportamiento y pueda agrupar una colección de documentos. Activa la interpretación de opciones relacionadas con este tipo de proyectos.
- **book**: Nueva categoría de opciones que indica la configuración específica aplicable a todos los documentos de este proyecto, así como las opciones globales. Por ejemplo, el título, el autor y la fecha se mostrarán en una página de portada especial.
- **chapters**: Nueva subcategoría de opciones que permite insertar un listado de nombres de archivo que contienen las diferentes secciones o capítulos de nuestro libro.

Debemos tener en cuenta que las secciones o capítulos se procesarán exactamente en el mismo orden en el que aparecen en este listado, por lo que es importante prestar atención a dicha ordenación.

### 5.3 Archivo de portada

Además del nuevo archivo `_quarto.yml` se genera otro archivo llamado `index.qmd` que contiene el material que se presentará como portada del libro o colección de documentos. Un archivo general se crea también en otros tipos de proyectos, por ejemplo para la página de inicio de un sitio web generado con Quarto. Es en esta página de portada en la que se consigna la información general de la configuración (título, autores, fecha, etc.).

El contenido de este archivo es el mismo que en cualquier otro documento, siguiendo las mismas reglas de sintaxis de Quarto para archivos `.qmd` que ya hemos visto.

Debemos resaltar también que el título de esta página de portada no suele numerarse y, por tanto, es habitual que se configure como:

```
Preface {.unnumbered}
```

### 5.4 Herramientas de redacción

Todas las herramientas de redacción que ya hemos visto en el Capítulo 4 se pueden utilizar en el caso de los capítulos de un libro. Es importante resaltar que los enlaces a las secciones funcionan también desde documentos de otras secciones o capítulos diferentes del libro, aunque estén en otro archivo distinto.

#### 5.4.1 Estructura del libro

Además de los capítulos estándar, también podemos organizar el contenido del libro en partes (que agrupan capítulos relacionados entre sí), así como apéndices, presentados a continuación del contenido principal del libro para ofrecer material adicional.

## 5 Libros

Veamos un ejemplo de configuración de un libro que incluya varias partes y que podemos integrar en el resto de la configuración del proyecto, dentro del archivo `_quarto.yml`.

```
chapters:
 - index.qmd
 - preface.qmd
 - part: dice.qmd
 chapters:
 - basics.qmd
 - packages.qmd
 - part: cards.qmd
 chapters:
 - objects.qmd
 - notation.qmd
 - modifying.qmd
 - environments.qmd
```

Es importante señalar que, en este caso, la opción `part` puede aceptar un archivo con extensión `.qmd` (como en el ejemplo), o bien una cadena de caracteres entre comillas que simplemente indique un título para la parte.

Ahora presentamos un ejemplo para los apéndices.

```
book:
 title: "mybook"
 author: "Jane Doe"
 date: "5/9/2021"
 chapters:
 - index.qmd
 - intro.qmd
 - summary.qmd
 - references.qmd
 appendices:
 - tools.qmd
 - resources.qmd
```

Es importante resaltar que estas opciones de configuración generan la salida apropiada (partes y apéndices) tanto en el caso de sitios web HTML como cuando generamos un documento PDF, siguiendo en el segundo caso la sintaxis estándar de LaTeX para indicar la estructura del documento.

- Ejemplo de un conocido libro de H. Wickham creado en Quarto y accesible públicamente, que está organizado en partes que agrupan capítulos.
- Opciones de configuración adicionales para la estructura de un libro en Quarto.

## 5.5 Gestión de referencias

Como ya se ha mencionado, las referencias internas (a figuras, tablas, ecuaciones y otras secciones del documento, entre otros elementos) funcionan exactamente igual que en los documentos individuales (consulta la Sección 4.5.1), con la ventaja adicional de que en un libro que combina diferentes capítulos la numeración de todos los elementos se actualiza para reflejar el número de capítulo como prefijo del número de elemento (por ejemplo, “Figura 1.2” para la segunda figura del capítulo 1).

Es importante recalcar que para que la numeración automática de los elementos funcione adecuadamente, es imprescindible que comiencen por el prefijo adecuado (`#sec-` para secciones, `#fig-` para figuras, etc.).

## 5.6 Previsualización del proyecto

Para previsualizar en nuestra máquina local el proyecto que hemos creado, utilizando RStudio, debemos seleccionar en el panel superior derecho la pestaña *Build* y pulsar el botón *Render Book* para generar todos los formatos de salida que estén configurados en el archivo `_quarto.yml`, tal y como se muestra en la Figura 5.3. Además, también existe la opción de seleccionar solo uno de estos formatos de salida si pulsamos con cuidado en la pequeña flecha junto al botón “Render Book”, para desplegar una lista de opciones de formato de salida y seleccionar uno de ellos, como muestra la Figura 5.4

## 5.7 Publicación del proyecto

Existen diferentes [opciones de publicación] para poner nuestro libro o colección de documentos disponible para otros usuarios, incluyendo:

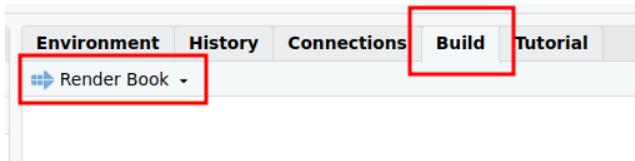


Figura 5.3: Botón para lanzar el proceso de previsualización del libro en RStudio.

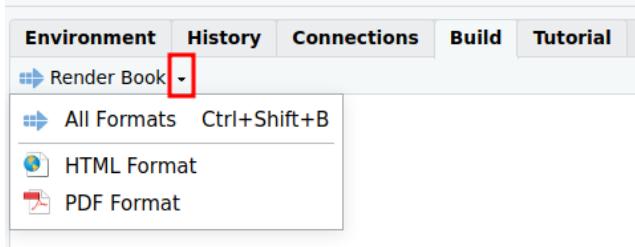


Figura 5.4: Listado de opciones de configuración que aparecen al pulsar la flecha junto al botón *Render Book*

- Quarto Pub.
- GitHub Pages, muy conveniente si queremos una solución integrada para gestión de versiones del código fuente de nuestro proyecto.
- Netlify, una plataforma de publicación web profesional que permite más opciones de ajuste y configuración.

También es posible utilizar otros servicios para publicar documentos, libros y sitios web con Quarto, incluyendo GitLab (alternativa a GitHub), aunque no estarán tan automatizados e integrados con las herramientas que ofrece Quarto como las opciones anteriores.

## 5.8 Plantillas y personalización

Al igual que hemos explicado en el caso de los documentos individuales, se pueden utilizar opciones de configuración análogas para personalizar el estilo y el tema de diseño que podemos aplicar a nuestro libro o colección de documentos, tanto en la versión de sitio web HTML como en la versión de salida en formato PDF.

## *5 Libros*

- Guía para la personalización del estilo de los libros en Quarto.
- Galería de ejemplo de algunos estilos para personalizar el aspecto de documentos Quarto.

# 6 Taller: colección de apuntes

En este capítulo, vamos a poner en práctica los conocimientos adquiridos para crear un documento de apuntes de clase, que también puede servir como cuaderno de laboratorio, cuaderno de prácticas o bitácora de experimentos.

Para ello, vamos combinar lo que ya sabemos acerca de los documentos y libros en Quarto junto con nuevas opciones de configuración y plantillas que abrirán nuevas e interesantes posibilidades de diseño. Por supuesto, con los conocimientos adicionales adecuados (no necesariamente muy avanzados), muchos usuarios serán capaces de adaptar y personalizar el diseño gráfico y la maquetación de sus documentos en PDF o LaTeX para ajustarlos a sus necesidades.

## 6.1 Opciones de diseño del documento

Además de las numerosas opciones que hemos explorado para ajustar el comportamiento de muchos elementos de nuestros documentos, Quarto todavía nos reserva muchas más alternativas para controlar el diseño de la página y la distribución de su contenido de forma que:

- Ocupe toda la región principal (centro del documento).
- Sobreponga los límites de la región principal.
- Cubra todo el ancho de la pantalla (por ejemplo, cuando queremos representar mapas o figuras muy anchas, compuestas por varias gráficas en una misma fila).
- Se sitúe en los márgenes del documento.
- Guía para personalizar el diseño del documento (HTML).
- Opciones de configuración del diseño del documento (PDF).

## 6 Taller: colección de apuntes

En este taller nos centramos en la generación de salida en HTML. Por defecto, en un documento HTML producido con Quarto tenemos un diseño estándar del espacio de la página en tres columnas:

- La columna central, *body* más ancha, para el contenido principal de la página.
- La columna izquierda, *sidebar* (barra lateral), que se suele emplear en los libros para la tabla de contenidos desplegable que permite navegar por la colección de documentos.
- La columna derecha, *margin*, que se suele utilizar para mostrar la tabla de contenidos del documento o capítulo que aparece en el *body* en ese momento, junto con otros enlaces a contenido adicional (e.g. enlace al código fuente, a la web de la organización o editorial, etc.).

Por supuesto, las dimensiones de estas tres columnas son totalmente personalizables.

### 6.1.1 Contenido principal

Si no especificamos ninguna variación en la configuración, el documento de Quarto muestra los elementos ocupando todo el ancho de la columna central (*document body*).

Si queremos que el contenido sobrepase un poco los límites de ese espacio central, pero sin llegar a ocupar todo el ancho de la pantalla, podemos usar una nueva sección *div* HTML con un estilo específico:

```
 ::= {.column-body-outset}
Outset content...
:::
```

Para ampliar aún mas el ancho de presentación, pero dejando cierto margen de espacio respecto a ambos bordes de la pantalla se puede usar en su lugar el estilo *{.column-page}*. Por último, si realmente queremos ocupar todo el ancho de la pantalla (sin dejar margen a ambos lados) se puede usar el estilo *{.column-screen}*. Estas opciones de diseño también son aplicables a la configuración de los bloques de código ejecutable. Veamos un ejemplo:

```
```{r}
#| column: page

knitr::kable(
  mtcars[1:6, 1:10]
)
```
```

|                   | mpg  | cyl | disp | hp  | drat | wt    | qsec  | vs | am | gear |
|-------------------|------|-----|------|-----|------|-------|-------|----|----|------|
| Mazda RX4         | 21.0 | 6   | 160  | 110 | 3.90 | 2.620 | 16.46 | 0  | 1  | 4    |
| Mazda RX4 Wag     | 21.0 | 6   | 160  | 110 | 3.90 | 2.875 | 17.02 | 0  | 1  | 4    |
| Datsun 710        | 22.8 | 4   | 108  | 93  | 3.85 | 2.320 | 18.61 | 1  | 1  | 4    |
| Hornet 4 Drive    | 21.4 | 6   | 258  | 110 | 3.08 | 3.215 | 19.44 | 1  | 0  | 3    |
| Hornet Sportabout | 18.7 | 8   | 360  | 175 | 3.15 | 3.440 | 17.02 | 0  | 0  | 3    |
| Valiant           | 18.1 | 6   | 225  | 105 | 2.76 | 3.460 | 20.22 | 1  | 0  | 3    |

Por último, las opciones de estilo `{.column-screen-inset}` o `{.column-screen-inset-shaded}` dan una apariencia de ocupar el ancho de la pantalla al completo pero dejando un ligero margen en los bordes (con sombreado, en el segundo caso).

### 6.1.2 Contenido en los márgenes

El estilo `{.column-margin}` permite ubicar en la columna derecha de la página cualquier elemento, ya sea un texto, una figura o gráfica, una tabla, ecuación, etc.

```
::: {.column-margin}
! [A margin image](image.png)
:::
```

Veamos un ejemplo aplicado a una ecuación que se muestra al margen:

Ahora con una figura al margen:

```
```{r}
#| label: fig-mtcars
#| fig-cap: "MPG vs horsepower, colored by
#| transmission."
```

We know from *the first fundamental theorem of calculus* that for x in $[a, b]$:

$$\frac{d}{dx} \left(\int_a^x f(u) du \right) = f(x).$$

6 Taller: colección de apuntes

```
#| column: margin

library(ggplot2)
mtcars2 <- mtcars
mtcars2$am <- factor(
  mtcars$am, labels = c('automatic', 'manual')
)
ggplot(mtcars2, aes(hp, mpg, color = am)) +
  geom_point() +
  geom_smooth(formula = y ~ x, method = "loess") +
  theme(legend.position = 'bottom')
~~~
```

Por último, probamos con una tabla ubicada en el margen:

```
~~~{r}
#| column: margin

knitr:::kable(
  mtcars[1:6, 1:3]
)
~~~
```

También es posible especificar opciones individualizadas para cada elemento cuando un bloque de código genera varios resultados.

Se recomienda consultar la [guía completa de diseño de los artículos] de Quarto para explorar más opciones adicionales de configuración.

6.1.3 Notas al margen y citas bibliográficas

Si se añaden las siguientes opciones en la cabecera del documento (o de la salida HTML del proyecto global), se modifica la ubicación de las notas al pie y las citas bibliográficas, que pasan a mostrarse en el margen derecho del documento, a la misma altura de la llamada o cita en el texto principal, siguiendo los principios de diseño de los documentos de Tufte:

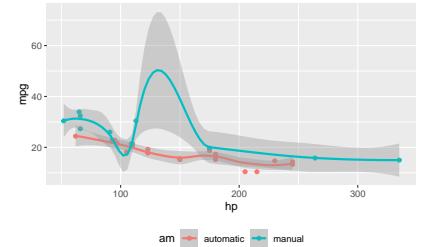


Figura 6.1: MPG vs horsepower, colored by transmission.

	mpg	cyl	disp
Mazda RX4	21.0	6	160
Mazda RX4 Wag	21.0	6	160
Datsun 710	22.8	4	108
Hornet 4 Drive	21.4	6	258
Hornet Sportabout	18.7	8	360
Valiant	18.1	6	225

```
reference-location: margin
citation-location: margin
```

Adicionalmente, también podemos poner notas puntuales (no numeradas) en el margen del documento con el estilo de párrafo `{.aside}`:

```
[This is a span that has the class `aside` which places
it in the margin without a footnote number.] {.aside}
```

This is a span that has the class
`aside` which places it in the margin
without a footnote number.

6.2 Anotaciones en bloques de código

Una característica muy útil para la creación de documentación y tutoriales es la de anotar las líneas dentro de los bloques de código ejecutable. Veamos un ejemplo:

```
```r
#| label: demo-code-annotation
#| message: false
#| output: false
library(tidyverse)
library(palmerpenguins)
penguins |>
 mutate(
 bill_ratio = bill_depth_mm / bill_length_mm, ①
 bill_area = bill_depth_mm * bill_length_mm ②
)
```
1. Take `penguins`, and then,
2. add new columns for the bill ratio and bill
   ↴ area.
```

Que produce el siguiente resultado:

```
library(tidyverse)
library(palmerpenguins)
penguins |>
  mutate(
    bill_ratio = bill_depth_mm / bill_length_mm,          ①
    bill_area   = bill_depth_mm * bill_length_mm          ②
  )
```

```
    bill_ratio = bill_depth_mm / bill_length_mm,
    bill_area  = bill_depth_mm * bill_length_mm
)
```

- ① Take `penguins`, and then,
- ② add new columns for the bill ratio and bill area.

Se pueden utilizar más opciones de configuración para controlar cómo se muestran las anotaciones de código (al hacer click sobre la anotación, al pasar el ratón por encima o la opción por defecto que hemos visto).

- Guía para anotación de líneas en bloques de código

6.3 Configuración del proyecto

6.3.1 Opciones globales

- Guía de referencia de opciones globales del diseño gráfico del documento (HTML).
- Guía de referencia de opciones globales del diseño gráfico del documento (PDF).

6.3.2 Idioma de los documentos

Una opción de configuración muy interesante para nuestros documentos y proyectos es la de especificar el idioma de los mismos. Por defecto, los documentos de Quarto se generan en inglés. Si embargo, varias herramientas, incluyendo paquetes LaTeX para localización del contenido y división de palabras, así como Pandoc y el propio Quarto pueden traducir muchas etiquetas automáticamente (“Figura”, “Tabla”, “Ecuación”, etc.) al idioma adecuado.

Para ello, tenemos que proporcionar un nuevo valor para la opción `lang` en la cabecera de nuestro documento o en el archivo de configuración `_quarto.yml` de nuestro proyecto:

```
---
title: "Mi documento en castellano"
lang: es
---
```

También es posible ofrecer traducciones personalizadas para etiquetas y campos estándar del documento o proyecto, de forma que se ajuste perfectamente a nuestras directrices de estilo o las de nuestra organización. Para más detalles, consulta la siguiente guía:

- Configuración de idioma en documentos Quarto.

En particular, si vamos a proporcionar una larga lista de traducciones alternativas conviene codificarlas en un fichero independiente y pasarlo a las opciones de configuración del documento o proyecto:

```
language: custom_translations.yml
```

Podemos llegar incluso al extremo de tener traducciones personalizadas para diferentes idiomas, que se usan en función del valor que pasemos a la opción de selección de idioma `lang`.

Respecto al **soporte multilenguaje simultáneo** (el mismo documento en varios idiomas), dicha característica todavía no está soportada en Quarto. Sin embargo, un paquete de reciente creación en la comunidad ROpenSci llamado `babledown` (aún en fase experimental), permite gestionar estas versiones multiidioma de forma simultánea (mostrando en un idioma enlaces al resto de idiomas disponibles). Más aún, el objetivo principal de este paquete es ofrecer una interfaz para el servicio de traducción automática DeepL, de forma que se traduzcan automáticamente los textos de nuestros documentos a otros idiomas.

6.3.3 Enlaces a otros formatos

Cuando generamos un documento Quarto en múltiples formatos, de manera automática se debería generar en la parte superior derecha (o en algún punto de la columna a la derecha de la página) un enlace a las otras versiones disponibles de dicho documento. La figura Figura 6.2 muestra un ejemplo del enlace que aparece en la página:

Para que esto suceda, tenemos que declarar que se genere la salida del documento en varios formatos:

Otros formatos

 PDF

Figura 6.2: Enlace a otra versión en PDF del documento creado con Quarto

```
title: Sample Page
author: Sarah Connor
date: last-modified
toc: true
format:
  html: default
  pdf: default
  ipynb: default
```

Las opciones anteriores generarían dos enlaces en la página HTML (el formato por defecto): uno para descargar la versión alternativa en PDF y otro para descargar la versión en formato *notebook* de Jupyter.

Se puede consultar la siguiente página de la guía de Quarto para descubrir más opciones de configuración, incluyendo cómo personalizar los textos de los enlaces o los formatos que se muestran para descarga:

- Guía para incluir enlaces a otros formatos del documento

Es importante saber que los proyectos de tipo `book` funcionan de una manera un tanto diferente, por ser un tipo de proyecto especial. En este caso, el enlace de descarga para versiones en otros formatos se pone junto al título del libro (en la parte superior izquierda) y se activa con la siguiente opción:

```
book:
  downloads: [pdf]
```

6.4 Ejercicio: personalizar el estilo gráfico del documento

Ahora, vamos a practicar con un ejemplo en el que se combinan varios aspectos que hemos estado tratando en este capítulo.

1. Descarga este documento de ejemplo en tu equipo: Ejemplo apuntes Tufte.
2. Crea un nuevo proyecto en RStudio, indicando como directorio de trabajo aquel en el que has guardado el ejemplo.
3. Previsualiza el documento en su versión HTML. Comprueba que hay un enlace para descargar también la versión en PDF.
4. Modifica algunos de los ejemplos para entender mejor cómo funcionan las opciones de configuración de los bloques de código ejecutables y los estilos gráficos aplicables al contenido en formato Markdown.
5. Modifica las dimensiones de la columna central (*document body*), para reducir su anchura 50 píxeles.

Referencias adicionales para el ejercicio práctico

- Galería de ejemplo de plantillas de documentos en Quarteto.
- Documento de ejemplo con los estilos gráficos de Tufte *handouts*.
 - Ejemplo de documento HTML con diseño gráfico Tufte *handouts*.

6.5 Personalización de los libros

Es importante entender la diferencia entre un proyecto configurado como libro y otro configurado como website. En el caso de un libro, se aplicarán las mismas opciones de formateo y diseño gráfico a todos los documentos que forman parte del mismo, puesto que se combinan en un solo documento (esto queda más claro cuando el formato de salida es PDF, puesto que se genera un único archivo con todo el contenido). Sin embargo, cuando creamos un sitio web es posible aplicar opciones de diseño y

6 Taller: colección de apuntes

configuración personalizadas a cada documento del sitio, puesto que se comportan de forma independiente (aunque es posible aplicar también opciones globales para todo el sitio web).

En consecuencia, en un proyecto de tipo `book` cualquier opción de renderizado (incluso las específicas de un formato de salida) deben especificarse en el archivo global `_quarto.yml`. Se pueden utilizar muchas de las herramientas aplicables a los sitios web con Quarto (barras de navegación, encabezado y pie de página, barra de búsqueda textual, etc.). No obstante, todas estas opciones hay que especificarlas bajo la etiqueta `book` en el caso de un libro, no bajo la etiqueta `website`.

Una excepción a esta regla es la opción global `output-dir`, para seleccionar un directorio de salida diferentes donde almacenar los archivos HTML, PDF, etc. del libro ya renderizado. Esta opción se debe poner debajo de la etiqueta `project`:

```
project:
  type: book
  output-dir: docs
```

Por contra, la opción `output-file`, que cambia el nombre del fichero de salida generado (por defecto es el título del documento), se pone debajo de la etiqueta `book`:

```
book:
  title: "El título de este libro es una cosa
  ↪ larguísima"
  author: "Jane Doe"
  output-file: "nombre-corto"
```

🔥 Extensiones de los archivos de salida

Cuando configuremos un nombre personalizado para el fichero de salida del libro con la opción `output-file` **no debemos añadir una extensión**, ya que ésta se añadirá automáticamente en función del tipo de fichero de salida generado.

💡 Personalización en LaTeX

Aunque no nos estamos centrando en la producción de salida LaTeX/PDF en este taller, aprovechamos este punto para resaltar que es posible también utilizar paquetes, *scripts* y otras herramientas LaTeX que ya tengamos desarrolladas previamente.

El método más directo (aunque a bajo nivel) para introducir los cambios es crear todo el libro con Quarto y compilar la salida en formato `latex`.

Listado 6.1 Terminal

```
quarto render --to latex
```

Esto nos dejará todo el código fuente de LaTeX dentro del subdirectorio `_book/book-latex` (o en el directorio de salida que hayamos configurado). Podemos acabar de editar esos ficheros y compilar con otra herramienta (e.g. un editor LaTeX o usando `make`) el documento final.

Sin embargo, si los cambios se limitan a modificar las opciones de configuración estándar del documento en LaTeX, entonces podemos usar la extensa [lista de variables de configuración para LaTeX de Pandoc], con la que podemos indicar:

- Opciones de clase (páginas a doble cara, tamaño de página).
- La opción `documentclass` para cambiar la plantilla del documento base; normalmente, Quarto usa `scrreport`, `scrbook` o `scrartcl` (clases del paquete LaTeX KOMA-Script).
- La opción `geometry` para cambiar las dimensiones de la página, los márgenes, etc.
- Fuentes y tipos de letra, color de la fuente, etc.
- Enlaces: color, estilo, ubicación, etc.

Por último, debemos tener en cuenta que si en alguno de los capítulos de nuestro libro insertamos notas en el margen, figuras al margen o cualquier otro elemento en los márgenes, entonces el área de texto de los demás capítulos se mantendrá más reducida (dejando espacio en blanco

en el margen), para poder presentar adecuadamente esos elementos en el margen de las páginas que sí los contienen. Nuevamente, esto se debe a que en el libro PDF la maquetación es común a todo el documento.

6.6 Plantillas y extensiones de Quarto

Por si no fuese suficiente con la larga lista de funcionalidades y opciones de configuración de Quarto estándar, se pueden instalar **extensiones** para ampliar todavía más su funcionalidad o modificar su comportamiento.

- Catálogo de extensiones disponibles para Quarto.

Es importante saber que antes de utilizar una extensión debemos de instalarla localmente y, además, que las extensiones se instalan **para cada carpeta de proyecto** en la que tengamos un documento individual, website, libro, etc. Por tanto, cada proyecto puede tener su propio listado de extensiones de Quarto instaladas y diferente del de otros proyectos.

Las extensiones se suelen publicar como proyectos en GitHub, por ejemplo la extensión `quarto-ext/fontawesome`. Si pulsamos en el enlace anterior vemos que una extensión es un conjunto de ficheros en un repositorio. Al instalar la extensión, estos ficheros se copiarán localmente en el directorio de nuestro proyecto, dentro de una carpeta de nombre `_extensions`.

Por ejemplo, supongamos que queremos instalar en nuestro proyecto `mi-projecto` dos nuevas extensiones creadas por `quarto-ext` en GitHub, que se llaman `fontawesome` y `video`. En una terminal del sistema, vamos al directorio de trabajo de nuestro proyecto (este paso previo es *importante*) y ejecutamos los comandos para instalarlas.

Listado 6.2 Terminal

```
cd myblog
quarto add quarto-ext/fontawesome
quarto add quarto-ext/video
```

 Extensiones en documentos sin proyecto

No es obligatorio que tengamos un proyecto creado para instalar extensiones y aplicarlas a documentos Quarto. Si tenemos uno o varios documentos en un directorio (aunque no sea un proyecto) y ejecutamos los comandos para instalar extensiones dentro de ese directorio, entonces cualquier documento situado en ese directorio puede acceder a dichas extensiones.

Las extensiones están organizadas dentro de cuatro grandes tipos:

- Códigos acortados o filtros: permiten usar una sintaxis más breve para operaciones complicadas o aplicar filtros de diversa índole para manejo de contenidos.
- Artículos de revista: extensiones que ya contienen los archivos de estilo necesarios para generar artículos científicos según el formato requerido por diversas organizaciones y editoriales (véase Capítulo 7).
- Formatos personalizados: Plantillas que aplican formatos predefinidos a nuestros documentos.
- `Reveal.js`: Plantillas y temas para la creación de diapositivas con Quarto utilizando esta biblioteca JavaScript.

6.7 Publicación del proyecto

Si queremos publicar nuestro proyecto (documento, libro, sitio web), tenemos varias opciones de publicación, como ya vimos en la Sección 5.7. De entre las opciones disponibles, una bastante directa y que no tiene coste es utilizar GitHub Pages.

GitHub Pages es un método de publicación de documentación asociada a un proyecto de GitHub. Esencialmente, tenemos que seguir un procedimiento ya definido para subir los archivos ya compilados de nuestro proyecto a GitHub y configurar ciertos parámetros de nuestro proyecto en esa plataforma. Como resultado, el servicio GitHub Pages toma ese contenido y lo publica en un sitio web, asignándole una URL para que sea accesible mediante un navegador (igual que estos mismos apuntes).

Aquí vamos a explicar el método para publicar nuestro contenido utilizando el comando `quarto publish`.

6.7.1 Método con quarto-publish

⚠️ Antes de empezar a usar Git y GitHub con RStudio

Comenzar a utilizar un Sistema de Control de Versiones (SCV) como Git y un servicio como GitHub para gestionar nuestros proyectos es una tarea factible pero no tan sencilla como puede parecer, sobre todo si no has recibido previamente formación sobre estas tecnologías o no has aprendido a usarlas por tu cuenta.

Por tanto, hacemos una llamada a la cautela para afrontar el contenido de esta sección armados de mucha *paciencia*, ganas de aprender y sobre todo *tiempo* para seguir todos los pasos descritos en los diferentes capítulos de varios tutoriales.

Para poder publicar nuestro proyecto con este método, primero hemos tenido que crear un proyecto nuevo en GitHub, clonarlo en nuestra máquina y crear un nuevo proyecto (con RStudio o MS Visual Studio) dentro del directorio del proyecto para que esté bajo control de versiones. Alternativamente, también podemos realizar este proceso a la inversa: crear primero todo el proyecto en nuestra máquina, para luego subirlo a GitHub.

En cualquiera de los dos casos, el contenido de este taller no cubre la parte de poner nuestro proyecto en un SCV (Sistema de Control de Versiones) como Git, para después subirlo a un servicio como GitHub. No obstante, dejamos a continuación unos enlaces a un tutorial paso a paso que nos guiará por este proceso.

1. Si no tenemos ya una cuenta de usuario/a en GitHub, debemos registrar una nueva cuenta.
2. Ahora, debemos instalar Git en nuestra máquina, si no estaba ya instalado previamente.
3. Para no tener problemas después cuando interactuemos con GitHub, es preciso configurar un método de autenticación para acceder a la plataforma con nuestro usuario. La forma más sencilla es mediante el navegador web y configurando un sistema de autenticación de doble factor (2FA). La opción más cómoda es instalar la app de GitHub en tu móvil y seguir las instrucciones. Una vez que esté funcionando:

6 Taller: colección de apuntes

1. Abrimos la página web de GitHub en el navegador e ingresamos con nuestro usuario y contraseña.
2. Usamos la app de GitHub en el móvil para confirmar nuestra identidad.
4. Conectamos RStudio con Git y GitHub.
5. Ahora ya podemos crear un nuevo proyecto en GitHub de varias formas:
 - Nuevo proyecto primero en GitHub.
 - Proyecto que ya existe en la máquina local y lo copio en un nuevo repositorio clonado desde GitHub.
 - Proyecto que ya existe y lo subo a un repositorio de GitHub (sin clonarlo localmente).

Ahora ya podemos trabajar en nuestro proyecto, poniendo en práctica lo explicado hasta el momento en este taller para añadir secciones, bloques de código ejecutable, figuras, tablas, ecuaciones y demás elementos y opciones de configuración. En nuestra máquina, podemos previsualizar periódicamente el resultado (bontón *Render* en RStudio). Es conveniente que cada poco tiempo, confirmemos los cambios en Git.

Confirmación manual de cambios

Al contrario que en otros servicios que controlan cambios en ficheros (como Google Docs o MS Office 365 en One-Drive), Git **no controla automáticamente** los cambios realizados en los archivos. Por eso, cada cierto tiempo, cuando hemos terminado cambios relacionados y antes de seguir cambiando muchas cosas más, debemos indicar a Git que queremos **confirmar el trabajo** hecho hasta ese punto. En el último enlace del párrafo anterior se explica cómo hacerlo desde RStudio (vídeo incluido).

Finalmente, cuando ya tenemos una versión publicable de nuestro proyecto podemos seguir las siguientes instrucciones para publicar el proyecto con el comando `quarto publish` que automatiza varias fases del proceso.

- Publicar un proyecto Quarto con GitHub Pages y el comando `quarto publish`.

⚠ Enviar los cambios de la rama `main` a GitHub

Si has prestado atención al último tutorial, verás que la publicación de los contenidos se hace en una rama diferente de la principal de tu repositorio en GitHub, llamada `gh-pages`. A esa rama sólo se van a subir **los ficheros de salida** ya generados en los formatos que hayamos configurado en nuestro proyecto (HTML, PDF, etc.).

Es muy importante darse cuenta de que los ficheros fuente de nuestro proyecto, con extensión `.qmd`, `_quarto.yml` etc. **no se han copiado al repositorio remoto** y, por tanto, todavía no tenemos una “copia de seguridad” de dichos archivos en GitHub.

Para conseguir esto, desde RStudio (pestaña Git) podemos hacer una operación `push` que envía los *commits* (cambios confirmados) desde nuestra máquina local al repositorio remoto. Alternativamente, también podemos abrir una terminal de sistema en RStudio y ejecutar (en el directorio raíz del proyecto) el siguiente comando:

Listado 6.3 Terminal

```
git push origin main
```

Si todo va correctamente, aparecerán varios mensajes en pantalla informando del progreso de la subida y confirmando que el historial de cambios en nuestro repositorio local ya está *sincronizado* con el repositorio remoto alojado en GitHub.

Parte III

Publicaciones científicas

7 Artículos y publicaciones científicas

En este capítulo vemos otra característica importante de Quarto, como es la capacidad para crear formatos de documentos extendiendo los formatos básicos de archivo de salida como `html`, `pdf` o `docx`. Uno de los usos principales que contempla esta capacidad de Quarto es el de producir documentos que cumplan los requisitos solicitados por las editoriales científicas para publicación de artículos de investigación en las revistas que editan.

Un aspecto importante es que debemos recordar que Quarto, como mínimo, siempre intenta producir salida en formatos HTML y PDF, por lo que las plantillas deberían contemplar estos dos escenarios. Además, el uso de entornos `span` y `div` para las diferentes secciones permite aplicar las herramientas de diseño y estilo solicitadas por cada plantilla editorial tanto para HTML (con CSS) como para PDF (con entornos y macros de LaTeX). Otro aspecto adicional importante es que estas plantillas pretenden unificar la codificación de la información sobre autores/as y sus filiaciones, de forma que se escriba esta información una sola vez para reutilizarla en diferentes formatos. Por último, Quarto también es capaz de gestionar los requisitos de estilos de citas bibliográficas impuestos por las editoriales científicas y revistas para sus publicaciones.

- Listado de formatos editoriales soportados en Quarto.
- Otro listado de formatos editoriales soportados en Quarto.

En ocasiones, será necesario añadir código adicional para realizar ajustes finos en las plantillas de los documentos, de forma que podamos ajustarnos a los requisitos impuestos por la editorial para esa publicación. El siguiente tutorial ofrece información sobre los diferentes puntos en los que podemos insertar código adicional para realizar estos ajustes finos.

- Ajuste de las plantillas.

Finalmente, en caso de que no exista una plantilla para el formato que necesitamos siempre es posible crear nuestro propio formato de artículo.

- Creación de nuestro propio formato de artículo.

El paso inicial consiste en crear un nuevo directorio de proyecto y ejecutar dentro del mismo el comando para descargar la plantilla estándar (sin personalizar aún) para crear formatos de publicaciones científicas en Quarto. Supongamos que el directorio de proyecto se llama `jourA`. Cuidado, porque en este ejemplo el símbolo `$` no se debe introducir en la terminal; sólo lo estamos usando para diferenciar la entrada de usuario de la respuesta que devuelve la terminal al ejecutar el comando.

Listado 7.1 Terminal

```
$ cd jourA
$ quarto create extension journal
? Extension Name > jourA

Creating extension at
↳ /home/jfelipe/quarto/dev/jourA:
- Created README.md
- Created template.qmd
- Created _extensions/jourA/jourA.lua
- Created _extensions/jourA/styles.css
- Created _extensions/jourA/_extension.yml
- Created _extensions/jourA/header.tex
- Created bibliography.bib
```

Se crean varios archivos, incluyendo:

- `_extensions/jourA/_extension.yml`, que define los formatos de salida disponibles en esta plantilla de artículo. Por ejemplo, si definimos un formato de salida `html` y otro `pdf` en esta plantilla, estos estarán disponibles en el documento Quarto como `jourA-html` y `jourA-pdf` cuando un documento Quarto utilice nuestra plantilla.
- `template.pdf`, que es el documento de ejemplo que se genera para demostrar al usuario la estructura y opciones disponibles y que empiece a trabajar sobre el mismo.

7.1 La opción `keep-tex: true`

Si revisas los ejemplos ofrecidos en la guía para creación de plantillas de publicación, verás como en las opciones de formatos de salida del fichero `template.qmd`, dentro de la salida PDF se suele incluir la opción `keep-tex: true`. Esta opción fuerza a que no se elimine el fichero LaTeX (con extensión `.tex`) que se crea como paso previo para compilar el documento final en PDF. El motivo es permitir que el usuario pueda, si así lo necesita, modificar el fichero LaTeX directamente y compilarlo de forma manual.

No obstante, hay que recordar que si se vuelve a pulsar el botón de *Render* para el formato PDF en RStudio, o bien se ejecuta `quarto render --to pdf` en la línea de comandos, dicho fichero con extensión `.tex` se **sobreescribe** y perderíamos los cambios que hayamos realizado. En consecuencia, es mejor copiarlo en otra ubicación antes de realizar los ajustes manuales o bien crear una nueva rama (si usamos control de versiones) para ajustar el fichero en la misma sin que exista el riesgo de sobreescribir los cambios realizados.

7.2 Figuras y gráficos para publicación

Una ventaja importante de utilizar Quarto para crear nuestros artículos es la de poder integrar los resultados de la ejecución de nuestro código (gráficos, tablas, resultados de evaluación de modelos y algoritmos, etc.) directamente en nuestros documentos científicos. Esto mitiga en gran medida los inconvenientes ya mencionados al principio del taller para mantener actualizados todos los elementos, permitiendo que nos aseguremos de que estamos usando la versión correcta del código sobre los datos apropiados.

- Publication Perfect: un tutorial con materiales online en abierto sobre cómo mejorar los elementos de nuestras publicaciones y artículos, creado por el grupo Harvard Chan Bioinformatics Core.

Recordemos algunos de los paquetes en R que nos permiten crear gráficos y tablas ya preparados para su publicación:

- **Hmisc**: incluye muchas funciones para descripción de datos y creación de gráficos y tablas resumen de datos y modelos, listos para publicar.
 - Ejemplos reproducibles con **Hmisc**. Se indica que los ejemplos son para R Markdown, pero en realidad funcionan también para Quarto (ya que se usa por defecto el mismo motor **knitr** para tratamiento y ejecución de los bloques de código.)
- **summarytools**: ofrece muchas funciones y herramientas para presentar resúmenes de tablas de datos y data frames, tanto en forma de tabla como en gráficas para EDA.
 - Ejemplos de uso de **summarytools**.
- **ggpubr**: paquete para ayudar a personalizar gráficos creados con **ggplot2**, de forma que queden listos para su inclusión en publicaciones científicas.
 - Colección de tutoriales y ejemplos para creación de gráficos con **ggpubr**.
- Listado y ejemplos de paquetes para crear tablas en publicaciones.

7.3 Facilitar la cita de artículos

En un contexto del mundo académico e investigador en el que el volumen de trabajos y artículos publicados ha crecido exponencialmente en los medios digitales, es muy importante facilitar al máximo el trabajo de otros investigadores/as cuando quieren citar nuestras publicaciones. Por otra parte, existen muchos servicios de referencia de citas y bases de datos que ayudan a los investigadores/as a encontrar materiales que deben revisar.

Podemos incluir metadatos en nuestro documento para facilitar la automatización total o parcial de este proceso de indexación de publicaciones y citas de nuestro trabajo.

El siguiente tutorial muestra varios ejemplos de los campos de metadatos para facilitar las citas que podemos incluir en la cabecera del documento, incluyendo información sobre la revista o publicación que recoge el artículo (como el DOI), así como

formatos específicos compatibles con motores de indexación de bibliografía científica como Google Scholar.

- Guía para creación de artículos citables en Quarto.

Por defecto, cuando incluyamos esta información en la cabecera del documento Quarto debe crear un apéndice con la información de cita en texto plano y en BibTeX.

7.4 Ejemplo de uso de plantillas de artículos científicos

7.4.1 Plantilla para revistas de Elsevier

 Requisito previo: tener instalado LaTeX

En este apartado se muestra cómo generar un *draft* de un artículo para enviar a una revista científica de la editorial Elsevier. Para que este proceso funcione, recuerda que previamente debes tener instalada una distribución de LaTeX en tu equipo (como TeX Live) o bien instalar la distribución mínima TinyTeX en RStudio.

En este caso, los pasos a seguir son bastante sencillos:

1. Creamos un nuevo directorio para guardar nuestro proyecto, por ejemplo, una carpeta con el nombre `ejemplo-elsevier`.
2. En la terminal (por ejemplo en RStudio, pestaña en el panel inferior izquierdo), entramos en la carpeta recién creada y ejecutamos un comando para descargar la plantilla y empezar a usarla:

Listado 7.2 Terminal

```
cd ejemplo-elsevier
quarto use template quarto-journals/elsevier
```

3. Creamos un nuevo proyecto en RStudio sobre el directorio ya existente en el que hemos descargado los archivos de la plantilla.

7 Artículos y publicaciones científicas

4. Abrimos el archivo `ejemplo-elsevier.qmd` y pulsamos el botón *Render* para generar el la salida en PDF.

Si ya teníamos creado previamente nuestro proyecto con un documento genérico de Quarto no tenemos que empezar desde cero. Basta con abrir una terminal dentro del directorio del proyecto y ejecutar:

Listado 7.3 Terminal

```
quarto add quarto-journals/elsevier
```

Por último, si queremos generar el artículo *draft* en formato PDF desde la línea de comandos, ejecutamos:

Listado 7.4 Terminal

```
quarto render article.qmd --to elsevier-pdf
```

Fíjate bien en la opción de formato de salida que se especifica en el archivo `.qmd` para generar el artículo *draft*:

Listado 7.5 ejemplo-elsevier.qmd

```
format:
  elsevier-pdf:
    keep-tex: true
```

8 Principios FAIR

8.1 Visión general

Desde hace años, dentro de la comunidad científica se ha desarrollado un movimiento imparable para favorecer el acceso a toda la información relacionada con los trabajos, experimentos y publicaciones científicas, de forma que se facilite su validación y reproducción/replicación por parte de otros investigadores/as o expertos/as interesados.

Este movimiento se ha concentrado especialmente, en su fase inicial, en garantizar al menos el acceso a la materia prima necesaria para desarrollar muchos de estos proyectos: los datos. Los principios fundamentales que deben cumplirse para que los datos puedan ser indexados y reutilizados al máximo se denominan **Principios FAIR** y vienen recogidos, entre otras fuentes, en los documentos de guía de la Comisión Europea para los investigadores que participan en proyectos financiados por dicha organización. Estos principios FAIR fueron establecidos inicialmente por Wilkinson et al. (2016) y son:

- **Findable:** Los datos y metadatos (datos que describen a los datos, como su formato, contenido, significado, enlace con otros datos, etc.) deben recibir un identificador único global y persistente que permita localizarlos de forma directa. El estándar más utilizado hoy en día es el sistema DOI (ISO 26324). Los metadatos deben incluir de forma clara y explícita el identificador del dato al que describen, y tanto datos como metadatos deben estar registrados o indexados en fuentes que permitan su búsqueda y recuperación.
- **Accessible:** Datos y metadatos deben poder obtenerse mediante un protocolo de comunicación estándar y abierto. Los metadatos deben seguir permaneciendo accesibles aún cuando los datos ya no estén disponibles.

8 Principios FAIR

- **Interoperable:** Los datos y metadatos deben usar formatos estándar y abiertos de representación del conocimiento, vocabularios que sigan principios FAIR y deben incluir referencias a otros datos y metadatos con los que estén relacionados.
- **Reusable:** Los datos y metadatos están descritos de manera rica y precisa, con múltiples atributos relevantes que faciliten su utilización por otros usuarios/as.

Es importante recalcar la trascendencia de publicar en abierto y siguiendo buenas prácticas de investigación revisable y reproducible/replicable. Por ejemplo, la Comisión Nacional Evaluadora de la Actividad Investigadora (CNEAI) ha publicado en 2023-2024 nuevos criterios para evaluación de las publicaciones válidas para poder ser evaluadas en la concesión de un sexenio de investigación. Así mismo, en todas las convocatorias recientes financiadas por los distintos ministerios del Gobierno de España se exige que los resultados de investigación, especialmente todas las publicaciones, *datasets*, software y procedimientos sean públicamente accesibles y cumplan con los principios FAIR.

Sin embargo, todavía queda camino por recorrer. Un reciente estudio (Kumar et al., 2024) analizó el grado de cumplimiento de los principios FAIR (el llamado FAIRness) de los resultados de investigación publicados por proyectos multi-actor financiados recientemente dentro del programa marco H2020 europeo y relacionados con el sector agroalimentario. Como principales conclusiones, menos de un 10% de los proyectos analizados lograron cumplir con los principios FAIR, aunque dichos principios se cumplían en mayor medida en los artículos de investigación publicados en revistas y conferencias, así como en libros.

En general, otra de las conclusiones es que el sector agroalimentario y rural europeo se está convirtiendo en dependiente de los datos de manera cada vez más intensa y que la aplicación de los principios FAIR contribuye a mejorar la toma de decisiones y a explotar mejor los resultados de innovación derivados de estos proyectos. Sin embargo, también se constata que la comunidad de investigación para el desarrollo del sector agroalimentario y rural todavía tiene una experiencia limitada en la aplicación de estos principios.

8.2 Publicación del código fuente y documentación técnica

- Plataformas como GitHub y GitLab facilitan en gran medida la gestión y publicación de proyectos software dentro de iniciativas y proyectos de investigación, así como la publicación y mantenimiento de centros de documentación técnica digital sobre herramientas, procedimientos y buenas prácticas relacionadas con la actividad de dichas iniciativas.
- Ejemplos del proyecto SoilWise, financiado dentro del programa HE de la Comunidad Europea.
 - Espacio de co-creación digital de SoilWise en GitHub. Se ha conseguido creando una organización (sin coste), de forma que se puedan crear repositorios en esta plataforma, agrupados bajo el paraguas del proyecto, para diferentes fines: listados de software de interés, documentación técnica, manuales de uso, arquitectura de la plataforma propuesta, etc.
 - Centro de documentación. Aquí se usa la solución MkDocs, que funciona con el lenguaje Python (Quarto tiene grandes ventajas en este aspecto).
 - Repositorio con las herramientas para recolección de datos, documentadas en la página que describe el proceso de obtención de datos en la infraestructura del proyecto.
- Repositorios abiertos de datos y material de investigación como Zenodo y Figshare pueden ayudar a publicitar y citar material de investigación. Por ejemplo, Zenodo puede emitir un identificador persistente (DOI) para un repositorio de software en GitHub apuntando a una versión en particular del software que se haya publicado y etiquetado desde dicho proyecto en GitHub. Esto permite incluirlo en una publicación científica para saber con certeza qué versión exacta del código software se ha utilizado para realizar el trabajo recogido en esa publicación.

Como ejemplo, la Figura 8.1 muestra la página de Zenodo correspondiente al software PyMPDATA, que apunta al repositorio original del software alojado en GitHub, mostrado a su

8 Principios FAIR

vez en la Figura 8.2. Podemos ver las referencias cruzadas de metadatos que mantienen la coherencia de conexión en ambos sentidos.

Figura 8.1: Página del proyecto PyMPDATA (v1.2.0) en Zendo

Figura 8.2: Página del proyecto PyMPDATA (v1.2.0) en GitHub

- La página del proyecto en Zenodo apunta a una versión concreta del repositorio en GitHub, y mantiene un listado de todas las versiones del mismo proyecto anteriormente referenciadas. Cada nueva versión recibe un DOI diferente, para diferenciarlas únicamente.
 - La página de descripción del repositorio en GitHub exhibe, entre otras etiquetas, el DOI para esa versión en particular, el DOI al artículo científico publicado en el JOSS explicando este software, así como otras etiquetas para atribución de las fuentes de financiación del proyecto.

8.3 Publicación de conjuntos de datos

- Zenodo es, probablemente junto con Figshare, uno de los repositorios de datos abiertos más conocidos y ampliamente utilizados en investigación, sobre todo en Europa, al ser un proyecto integrado en OpenAIRE, una organización sin ánimo de lucro creada por la Unión Europea para favorecer la ciencia abierta.
- Publicaciones científicas en acceso abierto (*open access*).
- Importante enlazar el DOI de un *dataset en un artículo*, de forma que se puedan trazar (en sentido inverso) los DOIs de las publicaciones que emplean dicho dataset*.

8.4 Gestión de referencias y publicación en abierto

Además de la asignación de un DOI a artículos, conjuntos de datos, software y otros elementos de los trabajos de investigación, existen algunas herramientas adicionales que facilitan la identificación de los trabajos y la atribución de su autoría (muchas de ellas con soporte explícito en Quarto)

- El ORCID es un identificador gratuito, único y persistente para individuos que se involucran en actividades de investigación, innovación y académicas. Permite identificar rápidamente la identidad de un autor/a en una publicación o trabajo científico o la persona encargada de publicar y mantener un recurso (*dataset*, repositorio de software, etc.). Además, el ORCID permite generar listado de todas las contribuciones realizadas por el individuo identificado.
- Existe una amplia lista de archivos de *preprints* que permiten la publicación de trabajos preliminares que no han pasado aún un proceso de revisión por pares para su publicación en una revista. Cada vez más editoriales aceptan (e incluso fomentan) la publicación de estos documentos para dejar constancia rápida de los avances de investigación, debido a los elevados tiempos que requiere el proceso de revisión y publicación en muchas revistas de prestigio.

8 Principios FAIR

- PLOS es una editorial Open Access sin ánimo de lucro, que edita y publica varias revistas digitales de elevado impacto y gran difusión en muchas áreas incluyendo, transformación y sostenibilidad.

Un número creciente de editoriales están adoptando igualmente principios de publicación abierta (Open Access), aunque normalmente repercutiendo un coste significativo en los autores o las instituciones a las que están afiliados para sufragar los gastos de publicación.

- Guía para crear artículos citables con Quarto.

9 Recursos adicionales

9.1 Quarto

- Getting started: <https://quarto.org/docs/get-started/>.
- Guía completa (online): <https://quarto.org/docs/guide/>.
- Referencia de funciones y opciones (online): <https://quarto.org/docs/reference/>.
- Galería de proyectos de ejemplo: <https://quarto.org/docs/gallery/>.

9.2 Principios FAIR y ciencia abierta

- Artículo sobre principios FAIR: <https://www.nature.com/articles/sdata201618>.
- Mandato sobre gestión de datos científicos en el Programa Horizonte Europa (CE): <https://www.openaire.eu/how-to-comply-with-horizon-europe-mandate-for-rdm>.
 - Todos los proyectos HE deben publicar al comienzo del plan de trabajo un Data Management Plan (DMP), explicando con gran detalle cómo se van a obtener, procesar, analizar y gestionar los datos utilizados y generados en las actividades del proyecto.
 - * DMP TOOL: Herramienta online para crear DMPs de forma estructurada.
 - “*How to make your data FAIR*”: <https://www.openaire.eu/how-to-make-your-data-fair>.
 - Data Management Plan (University) of Cambridge.
- Comunidad ROpenSci: <https://ropensci.org/es/>.
- Comunidad ROpenSpain: <https://ropenspain.es/>.

Referencias

- Barba, L. A. (2018). Terminologies for reproducible research. *arXiv preprint arXiv:1802.03311*.
- Begley, C., & Ellis, L. (2012). *Drug development: Raise standards for preclinical cancer research*. *Nature*. [Online]. 483 (7391).
- Brainard, J., You, J., et al. (2018). What a massive database of retracted papers reveals about science publishing's «death penalty». *Science*, 25(1), 1-5.
- Burman, L. E., Reed, W. R., & Alm, J. (2010). A call for replication studies. *Public Finance Review*, 38(6), 787-793.
- Ioannidis, J. P. (2005). Why most published research findings are false. *PLoS Medicine*, 2(8), e124.
- Knuth, D. E. (1984). Literate Programming. *Comput. J.*, 27(2), 97-111. <https://doi.org/10.1093/comjnl/27.2.97>
- Kumar, P., Hendriks, T., Panoutsopoulos, H., & Brewster, C. (2024). Investigating FAIR data principles compliance in horizon 2020 funded Agri-food and rural development multi-actor projects. *Agricultural Systems*, 214, 103822. <https://doi.org/https://doi.org/10.1016/j.agsy.2023.103822>
- Leek, J. T., & Peng, R. D. (2015). Reproducible research can still be wrong: Adopting a prevention approach. *Proceedings of the National Academy of Sciences*, 112(6), 1645-1646.
- Peng, R. D. (2011). Reproducible research in computational science. *Science*, 334(6060), 1226-1227.
- Wicherts, J. M., Borsboom, D., Kats, J., & Molenaar, D. (2006). The poor availability of psychological research data for reanalysis. *American Psychologist*, 61(7), 726.
- Wilkinson, M. D., Dumontier, M., Aalbersberg, Ij. J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.-W., Silva Santos, L. B. da, Bourne, P. E., et al. (2016). The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data*, 3(1), 1-9. <https://doi.org/https://doi.org/10.1038/sdata.2016.18>

A Comandos y opciones de configuración

A.1 Resumen de sintaxis de Quarto

- Archivo de ejemplo sintaxis básica de Quarto.

A.2 Comandos Quarto

- Listado de comandos y opciones para publicar con Quarto.
- *Cheatsheet*: publicación de documentos con Quarto(PDF).

A.3 Celdas de código

- Opciones de configuración para bloques de código.

A.4 Formatos de documentos de salida

- Opciones de configuración para documentos HTML.
- Opciones de configuración para documentos en PDF.

A.5 Opciones de configuración de proyectos

- Opciones globales de configuración.
- Opciones de configuración para libros.
- Opciones de configuración para manuscritos.
- Opciones de configuración para websites.
- Configuración de referencias cruzadas.
- Configuración de referencias bibliográficas.

B Entornos de desarrollo para Quarto

B.1 R Studio

Es la IDE de referencia que hemos utilizado en el taller para explicar la creación y gestión de proyectos de documentación científica con Quarto y R.

- Introducción al uso de Quarto en RStudio.

B.2 Visual Studio

Es una IDE alternativa, creada por Microsoft, y con extensiones para dar soporte a múltiples lenguajes y plataformas de desarrollo.

En concreto, tiene una extensión para Quarto.

- Extensión de VS Code para Quarto.
- Introducción al uso de Quarto con Microsoft VS Code.

C Paquetes R y atribuciones

C.1 Requisitos previos

Para ejecutar los ejemplos incluidos en este taller, se necesita tener instalado R y una IDE de desarrollo para este lenguaje, como por ejemplo RStudio o Microsoft VS Code.

También tenemos que instalar el software Quarto, que es independiente de R y de la IDE de desarrollo que utilicemos.

Adicionalmente, se necesitan los siguientes paquetes de R:

- `tidyverse`.
- `palmerpenguins`.

Para generar documentos de salida en formato PDF se necesita tener instalada en el sistema una distribución LaTeX. Quarto ya incluye algunos comandos para facilitar la instalación de TinyTeX, así como para comprobar que este u otro motor de compilación de LaTeX están disponibles en el sistema.

- Instalación de TinyTeX con Quarto.
- Comprobación de herramientas TeX/LaTeX instaladas.

C.2 Atribución de imágenes e iconos

- R logo by The R Foundation.
- Python file icons created by Flat Icons - Flaticon.
- Html icons created by Freepik - Flaticon.
- Microsoft word icons created by Pixel perfect - Flaticon.
- Pdf icons created by Dmitry Miroliubov - Flaticon.

D Documentos PDF con LaTeX

D.1 Salida en formato PDF

Como ya hemos explicado, seleccionando el formato:

```
format:  
  pdf:
```

podemos indicar que se genere una salida de nuestro documento o proyecto con Quarto en formato PDF. Para ello, Quarto se apoya directamente en Pandoc y en la existencia de una distribución LaTeX en nuestro equipo para compilar el archivo con extensión `.tex` que se genera.

D.2 Acerca de LaTeX

LaTeX (con esa grafía) es un software de maquetación de documentos de alta calidad creado originalmente por L. Lamport sobre TeX, un sistema de tipografía ideado por el profesor Donald E. Knuth.

Para poder compilar un archivo LaTeX con extensión `.tex` necesitamos tener instalada una distribución LaTeX, como MiKTeX, TeX Live o TinyTeX que es una versión reducida de TeX Live que únicamente incluye los paquetes más empleados habitualmente y ocupa unos 100 MB de espacio de almacenamiento.

- Instrucciones de instalación de TinyTeX para Quarto.

Dentro de cada distribución de LaTeX hay diferentes motores de compilación (*compilation engines*), que son programas que toman los ficheros con extensión `.tex` y generan un documento final maquetado. Algunos de ellos son `pdflatex`, `xelatex` o

`luatex`. Por defecto, Quarto usa el motor `xelatex`, una opción moderna que incluye soporte nativo para UTF-8, múltiples tipografías en formato OTF (basta con que estén instaladas como fuentes en nuestro equipo) y muchas otras ventajas.

Por último, para manejar las referencias bibliográficas se necesita un gestor de bibliografía. La opción clásica era BibTeX (los archivos `.bib` de referencias están codificados en un formato compatible con este gestor). Sin embargo, en la actualidad se recomienda el uso del paquete BibLaTeX, mucho más potente, robusto y que ofrece un soporte multilenguaje óptimo para la generación de referencias bibliográficas según las normas aplicables a cada idioma. Para usarlo en nuestra salida en PDF de un proyecto Quarto, indicamos la siguiente opción:

```
format:
  pdf:
    cite-method: bibtex
```

D.2.1 Referencias de introducción a LaTeX

- Learn LaTeX in 30 minutes. Este manual lo ha creado el equipo de Overleaf, un servicio web para edición colaborativa de documentos en LaTeX.
- The Not So Short Introduction to LaTeX2 ε .

D.3 Ejemplo práctico

- Los apuntes de este taller están creados con Quarto y se genera salida tanto en formato HTML como en PDF. El documento PDF se puede descargar pulsando sobre el ícono  junto al título del libro en la parte superior izquierda de la pantalla.
 - El archivo `_quarto.yml` incluye varias opciones bajo el apartado `format:`, dentro de la opción `pdf:`. Entre ellas tenemos:
 - * `documentclass`: permite indicar la plantilla de documento principal. Algunas opciones son

D Documentos PDF con LaTeX

`scrartcl`, `scrreprt` o `scrbook`, correspondientes a las clases KOMA-Script ¹, `memoir` para libros y tesis ², o la clase básica `book`. Existen muchos otros paquetes LaTeX que ofrecen plantillas más especializadas para tesis, trabajos y otros tipos de documentos.

- * `classoptions`: son las opciones que se pasan para configurar la plantilla principal, dentro del comando LaTeX `\documentclass`.
- * `hyperrefoptions`: permite insertar opciones para modificar el comportamiento del paquete `hyperref`, que controla todos los tipos de enlaces (incluyendo a bibliografía y referencias cruzadas) del documento.
- * `include-in-header`: abre la posibilidad de insertar comandos LaTeX adicionales para cargar otros paquetes o ajustar otros parámetros de configuración.
- Dentro del apartado `book`: de la configuración, la opción `downloads`: `[pdf]` provoca que se añada el icono  junto al título del libro como enlace de descarga de la versión en PDF.
- Para las referencias bibliográficas, las opciones de alto nivel `bibliography`: y `csl`: permiten indicar un archivo de referencias bibliográficas en formato BibTeX y un estilo de formateo de las referencias. Esta información es aplicable también a la salida en PDF.

Pandoc explica en el apartado de su documentación dedicado a crear documentos PDF muchas de las opciones soportadas. La lista completa de variables de configuración para LaTeX también está disponibles. Todas ellas están directamente soportadas en Quarto.

¹Guía rápida de KOMA-Script.

²Manual de la clase `memoir`.

Referencias

- Barba, L. A. (2018). Terminologies for reproducible research. *arXiv preprint arXiv:1802.03311*.
- Begley, C., & Ellis, L. (2012). *Drug development: Raise standards for preclinical cancer research*. *Nature*. [Online]. 483 (7391).
- Brainard, J., You, J., et al. (2018). What a massive database of retracted papers reveals about science publishing's «death penalty». *Science*, 25(1), 1-5.
- Burman, L. E., Reed, W. R., & Alm, J. (2010). A call for replication studies. *Public Finance Review*, 38(6), 787-793.
- Ioannidis, J. P. (2005). Why most published research findings are false. *PLoS Medicine*, 2(8), e124.
- Knuth, D. E. (1984). Literate Programming. *Comput. J.*, 27(2), 97-111. <https://doi.org/10.1093/comjnl/27.2.97>
- Kumar, P., Hendriks, T., Panoutsopoulos, H., & Brewster, C. (2024). Investigating FAIR data principles compliance in horizon 2020 funded Agri-food and rural development multi-actor projects. *Agricultural Systems*, 214, 103822. <https://doi.org/https://doi.org/10.1016/j.agsy.2023.103822>
- Leek, J. T., & Peng, R. D. (2015). Reproducible research can still be wrong: Adopting a prevention approach. *Proceedings of the National Academy of Sciences*, 112(6), 1645-1646.
- Peng, R. D. (2011). Reproducible research in computational science. *Science*, 334(6060), 1226-1227.
- Wicherts, J. M., Borsboom, D., Kats, J., & Molenaar, D. (2006). The poor availability of psychological research data for reanalysis. *American Psychologist*, 61(7), 726.
- Wilkinson, M. D., Dumontier, M., Aalbersberg, Ij. J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.-W., Silva Santos, L. B. da, Bourne, P. E., et al. (2016). The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data*, 3(1), 1-9. <https://doi.org/https://doi.org/10.1038/sdata.2016.18>