

Coding Milestones 2

2024-08-02

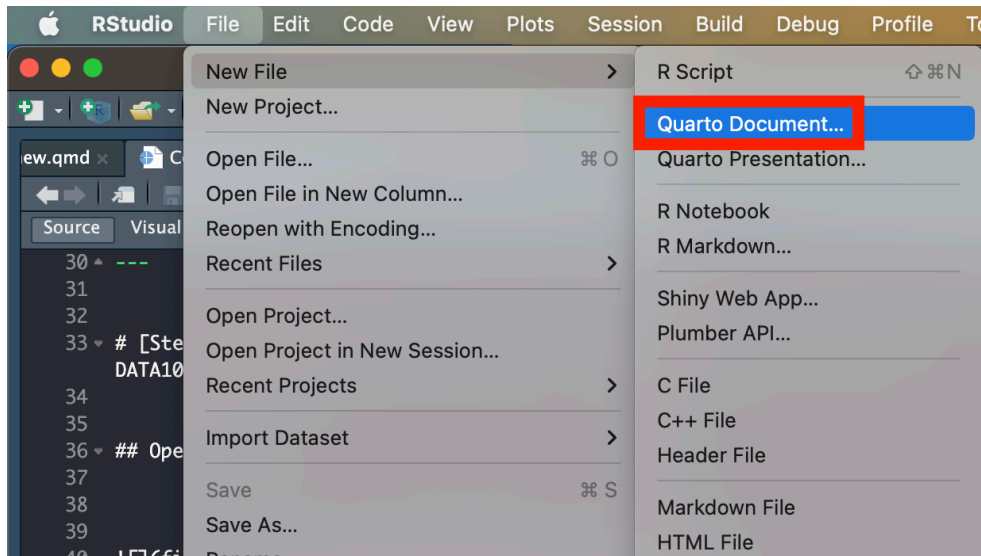
Coding Milestones 2

- Open RStudio, create a .qmd file & store it in your DATA1001 folder
 - Create a .qmd in RStudio
 - Store in your DATA1001 folder
- Clean the .qmd file, by keeping the YAML and removing the other lines.
 - Clean the .qmd, keeping the YAML
 - Create a code chunk
- Perform IDA on the iris data, using `str()`, `head()` and `tail()`
 - IDA on iris using `str()`, `head()` and `tail()`
- Create a `ggplot` of the iris data, from a blank canvas to a full plot with title
 - Example 1: Making a `ggplot` simple boxplot
 - Example 2: Making a `ggplot` sliced histogram with title
- Edit & knit your .qmd file to produce an .html file
 - Render your .qmd to create a .html

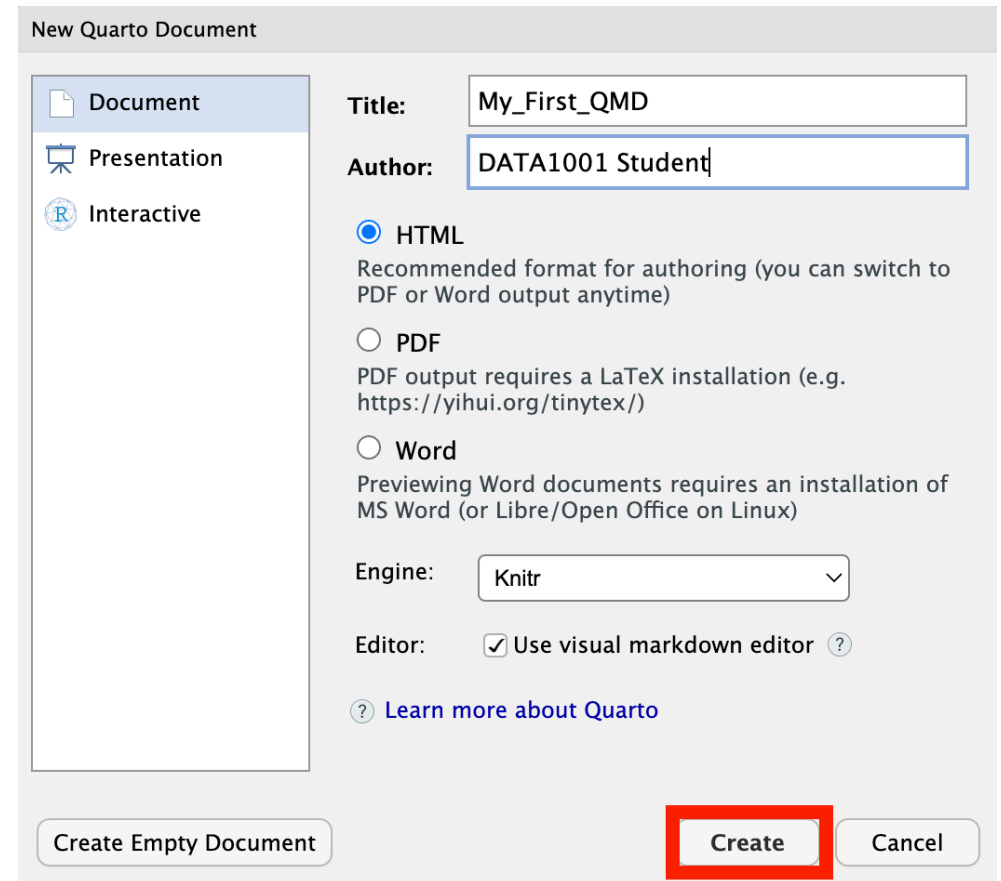
**Open RStudio, create a .qmd file &
store it in your DATA1001 folder**

Create a .qmd in RStudio

Open a Quarto Document (.qmd) in RStudio.

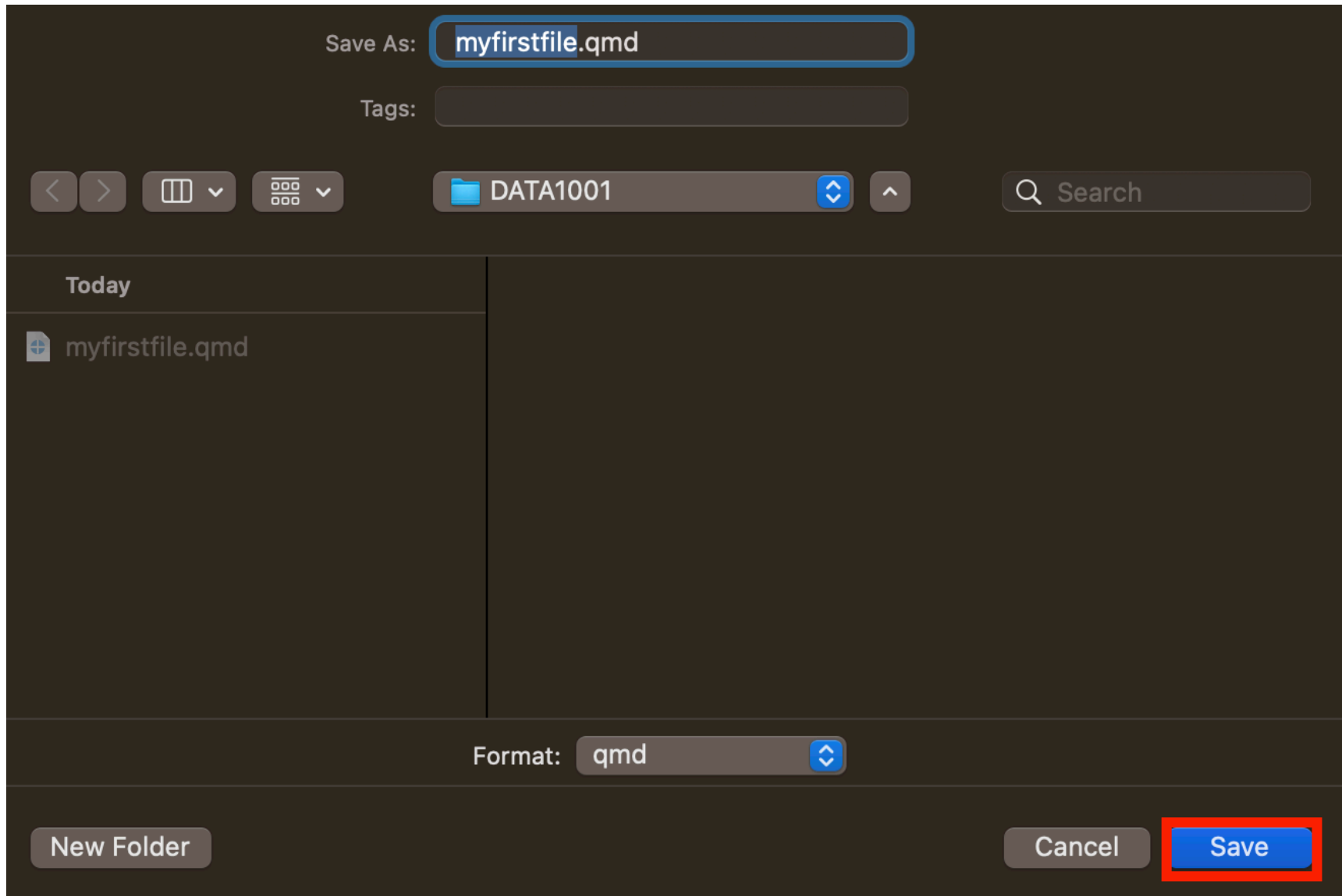


Give your .qmd a title and author, then click create.



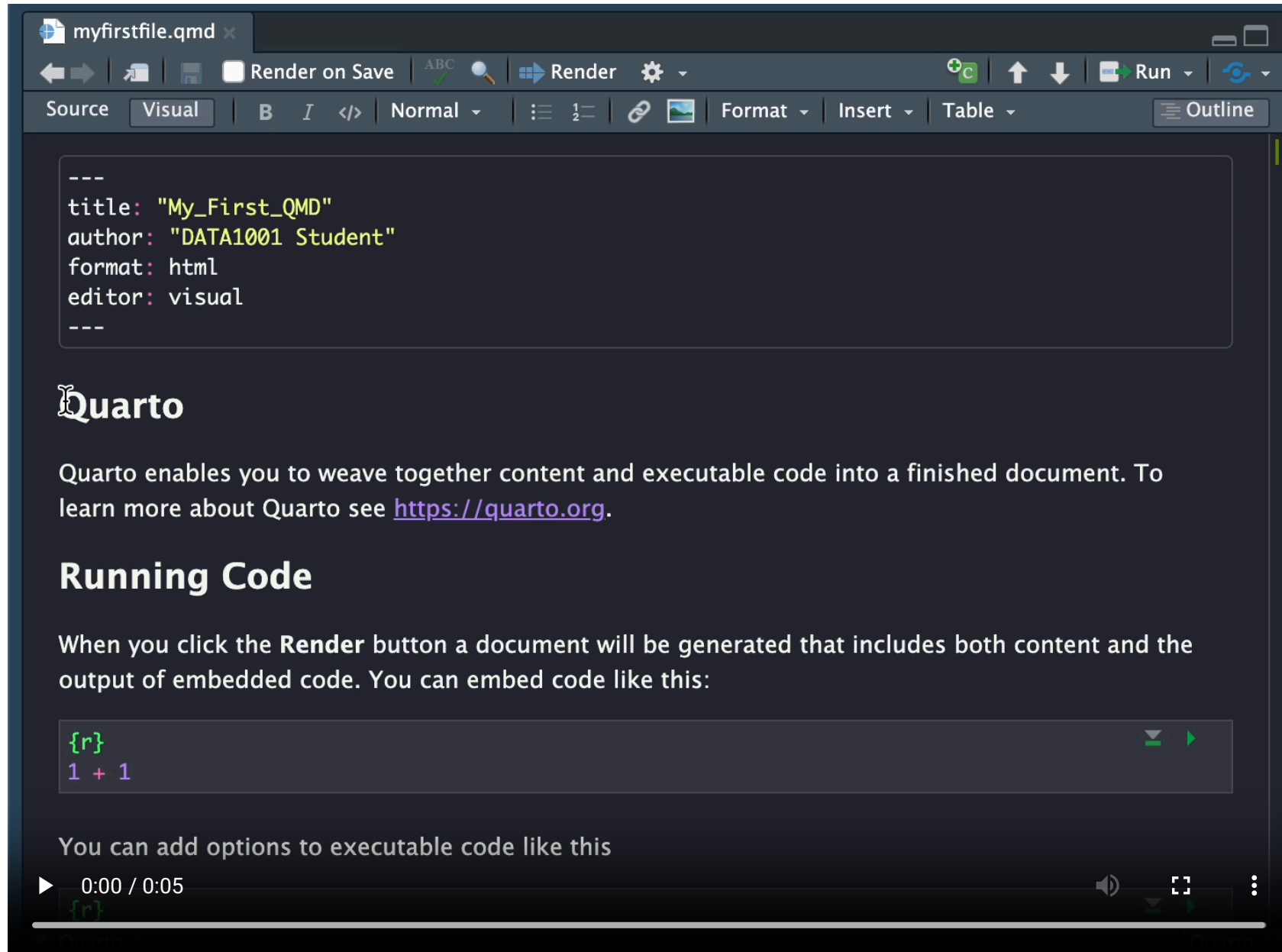
Store in your DATA1001 folder

Name your file and save it in your DATA1001 folder.



**Clean the .qmd file, by keeping the
YAML and removing the other lines.**

Clean the .qmd, keeping the YAML



The screenshot shows the Quarto editor interface for a file named `myfirstfile.qmd`. The top toolbar includes buttons for navigation, rendering, and running. The editor is in 'Visual' mode, showing the YAML metadata at the top of the file:

```
---  
title: "My_First_QMD"  
author: "DATA1001 Student"  
format: html  
editor: visual  
---
```

Below the metadata, the content of the file is displayed. It starts with the Quarto logo, followed by a paragraph explaining that Quarto weaves content and executable code into a finished document, with a link to <https://quarto.org>. This is followed by a section titled 'Running Code' which explains that clicking the 'Render' button generates a document including both content and code output. An example of executable code is shown in a code block:

```
{r}  
1 + 1
```

Below the code block, it states 'You can add options to executable code like this'. At the bottom of the editor, there is a video player showing a 0:00 / 0:05 duration.

Create a code chunk

Switch over to “Source”.

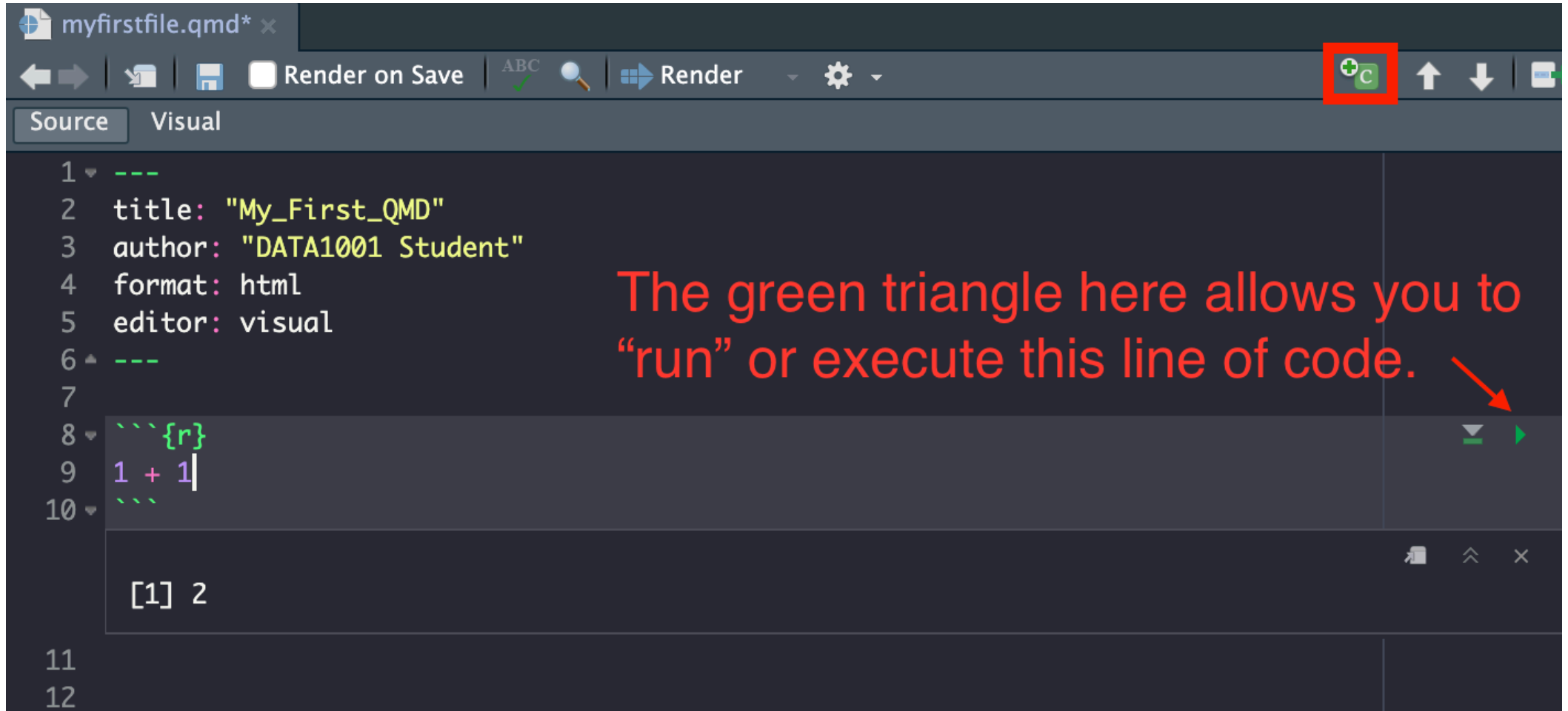


The screenshot shows a code editor window for a file named 'myfirstfile.qmd'. The interface includes a toolbar with navigation and editing icons, a 'Render on Save' checkbox, and a 'Render' button. Below the toolbar, there are two tabs: 'Source' and 'Visual'. The 'Source' tab is currently selected and highlighted with a red rectangular box. The code area displays a YAML frontmatter block with the following content:

```
1  ---
2  title: "My_First_QMD"
3  author: "DATA1001 Student"
4  format: html
5  editor: visual
6  ---
7
8  |
9
10
```


Create a code chunk

Click the green “C” button to create a code chunk, where you will write your code!



The screenshot shows the RStudio interface with a file named `myfirstfile.qmd*` open. The toolbar at the top includes buttons for navigation, saving, and rendering. A green button with a white 'C' and a plus sign is highlighted with a red box. Below the toolbar, the 'Source' tab is active, displaying a code chunk. The chunk starts with a dashed line, followed by a title, author, format, and editor. The code inside the chunk is `1 + 1`. A red arrow points to the green triangle button on the right side of the chunk, which is used to execute the code. The output of the code is shown in a console window at the bottom, displaying `[1] 2`.

```
1 ---  
2 title: "My_First_QMD"  
3 author: "DATA1001 Student"  
4 format: html  
5 editor: visual  
6 ---  
7  
8 ```{r}  
9 1 + 1  
10 ```  
11  
12
```

The green triangle here allows you to “run” or execute this line of code.

```
[1] 2
```

**Perform IDA on the iris data, using
`str()`, `head()` and `tail()`**

IDA on iris using `str()`, `head()` and `tail()`

```
str(iris)
```

```
'data.frame':  150 obs. of  5 variables:
 $ Sepal.Length: num  5.1 4.9 4.7 4.6 5 5.4 4.6 5 4.4 4.9 ...
 $ Sepal.Width : num  3.5 3 3.2 3.1 3.6 3.9 3.4 3.4 2.9 3.1 ...
 $ Petal.Length: num  1.4 1.4 1.3 1.5 1.4 1.7 1.4 1.5 1.4 1.5 ...
 $ Petal.Width : num  0.2 0.2 0.2 0.2 0.2 0.4 0.3 0.2 0.2 0.1 ...
 $ Species      : Factor w/ 3 levels "setosa","versicolor",...: 1 1 1 1 1 1 1 1 1 1
 ...
```

```
head(iris)
```

	Sepal.Length	Sepal.Width	Petal.Length	Petal.Width	Species
1	5.1	3.5	1.4	0.2	setosa
2	4.9	3.0	1.4	0.2	setosa
3	4.7	3.2	1.3	0.2	setosa
4	4.6	3.1	1.5	0.2	setosa
5	5.0	3.6	1.4	0.2	setosa
6	5.4	3.9	1.7	0.4	setosa

```
tail(iris)
```

	Sepal.Length	Sepal.Width	Petal.Length	Petal.Width	Species
145	6.7	3.3	5.7	2.5	virginica
146	6.7	3.0	5.2	2.3	virginica
147	6.3	2.5	5.0	1.9	virginica
148	6.5	3.0	5.2	2.0	virginica
149	6.2	3.4	5.4	2.3	virginica
150	5.9	3.0	5.1	1.8	virginica

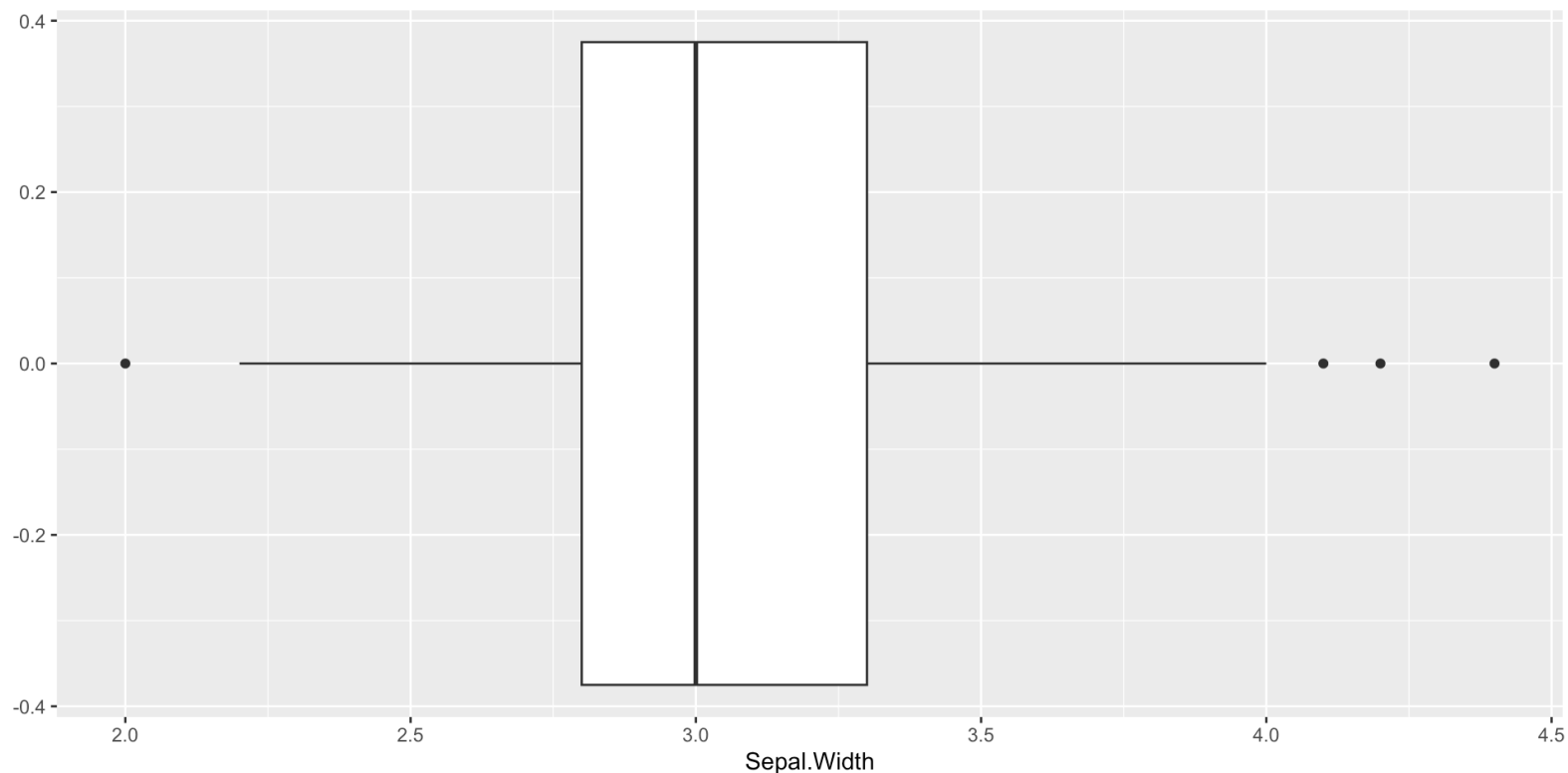
Create a **ggplot** of the iris data,
from a blank canvas to a full plot with
title

Example 1: Making a `ggplot` simple boxplot

Let's say we want to see a simple boxplot of **Sepal Width**. This means we will use 1 quantitative variable.

```
library(tidyverse) # Load the required library

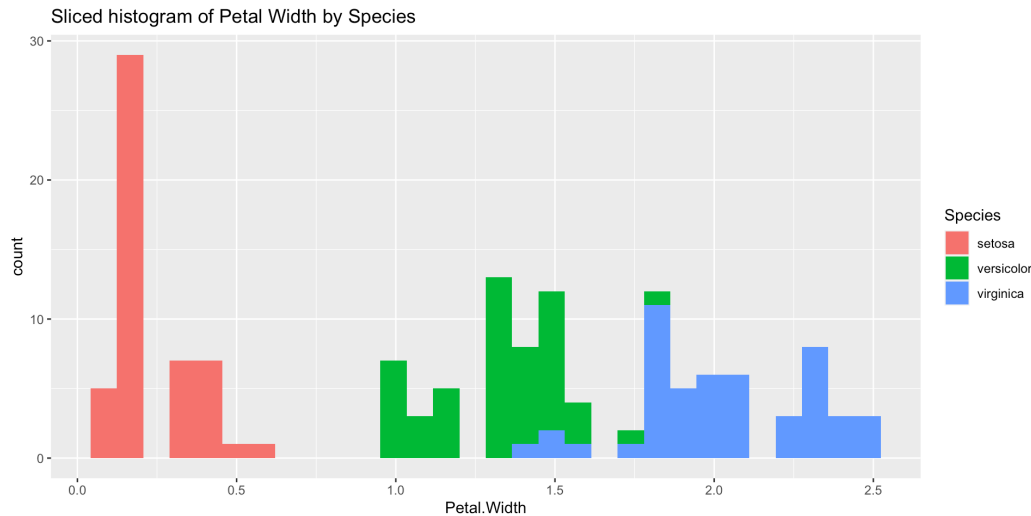
ggplot(iris, aes(x = Sepal.Width)) +
  geom_boxplot() # Making a boxplot
```



Example 2: Making a `ggplot` sliced histogram with title

Let's say we want to see a sliced histogram with the variables **Petal Width** (quantitative) and **Species** (qualitative).

```
ggplot(iris, aes(x = Petal.Width, fill = Species)) +  
  geom_histogram() + # Making a histogram  
  labs(title = "Sliced histogram of Petal Width by Species") # Adding a title
```



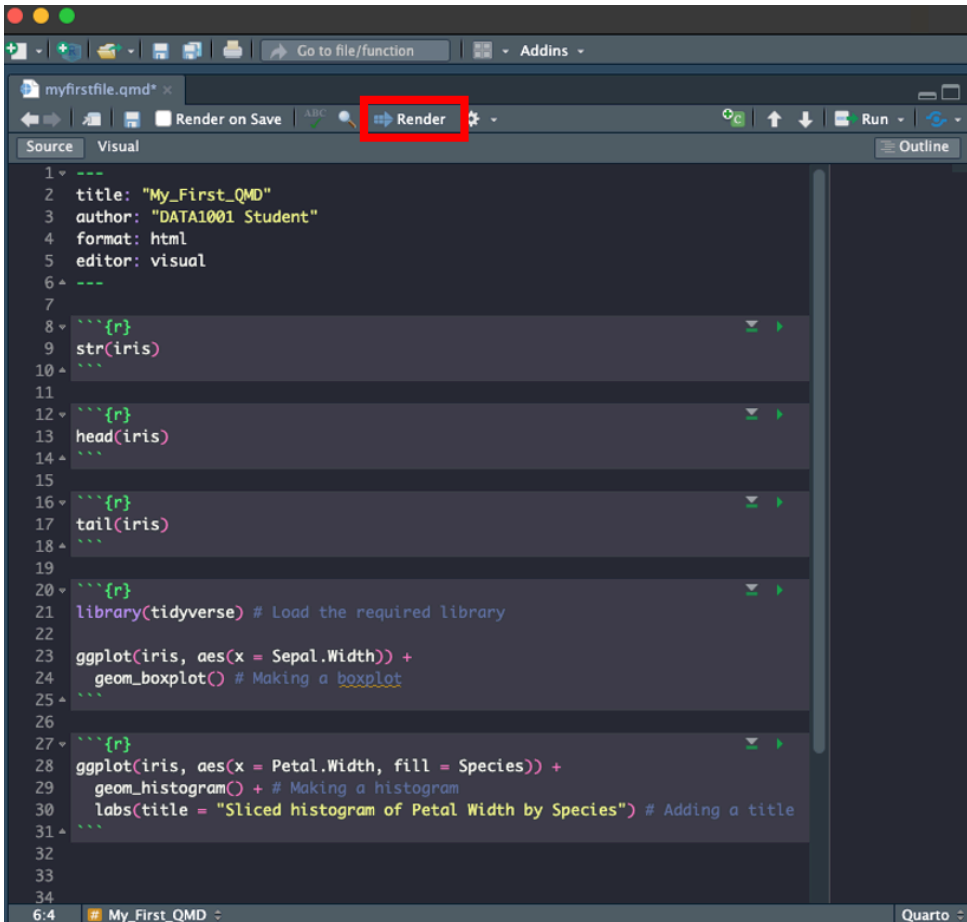
💡 Now it's your turn! Try making another graph using any of the variables you found from `str()`. Try it on another dataset too!

**Edit & knit your .qmd file to produce
an .html file**

Render your .qmd to create a .html

Render your .qmd file to produce .html output.

Now navigate to your DATA1001 folder and you will have a .html file!



The screenshot shows the Quarto editor interface. The top toolbar includes a 'Render' button, which is highlighted with a red rectangle. Below the toolbar, the 'Source' tab is active, displaying R code for a Quarto document. The code includes metadata (title, author, format, editor) and several R chunks for data manipulation and visualization using ggplot2.

```
1 ---  
2 title: "My_First_QMD"  
3 author: "DATA1001 Student"  
4 format: html  
5 editor: visual  
6 ---  
7  
8 ```{r}  
9 str(iris)  
10 ```  
11  
12 ```{r}  
13 head(iris)  
14 ```  
15  
16 ```{r}  
17 tail(iris)  
18 ```  
19  
20 ```{r}  
21 library(tidyverse) # Load the required library  
22  
23 ggplot(iris, aes(x = Sepal.Width)) +  
24   geom_boxplot() # Making a boxplot  
25 ```  
26  
27 ```{r}  
28 ggplot(iris, aes(x = Petal.Width, fill = Species)) +  
29   geom_histogram() # Making a histogram  
30   labs(title = "Sliced histogram of Petal Width by Species") # Adding a title  
31 ```  
32  
33  
34  
6:4 My_First_QMD
```

