

## Overview

Summary

Learning Objectives

Reproducibility

Creating an R Markdown Document in R Studio

R Markdown Syntax Basics

YAML Header

Headings

Text Formatting, Links, and Math Type

Lists

Tables

Chunks

In-line Code

HTML Preview

Further Reading

# R Bootcamp - Course 4: R Markdown

James Baglin

Last updated: 13 July, 2020

## Overview

### Summary

This course will introduce you to the basics of R Markdown and RStudio Notebooks. These documents enable you to create professional and reproducible reports through RStudio.

### Learning Objectives

By the end of this course, you will have completed the following:

- Understand the importance of reproducibility and the role of R Markdown.
- Install and create an R Markdown document in RStudio.

- Knit/preview and save an R Markdown document in HTML or PDF.
- Use basic R Markdown syntax to create reproducible reports.
- Know where to find further information for using R Markdown.

## Reproducibility

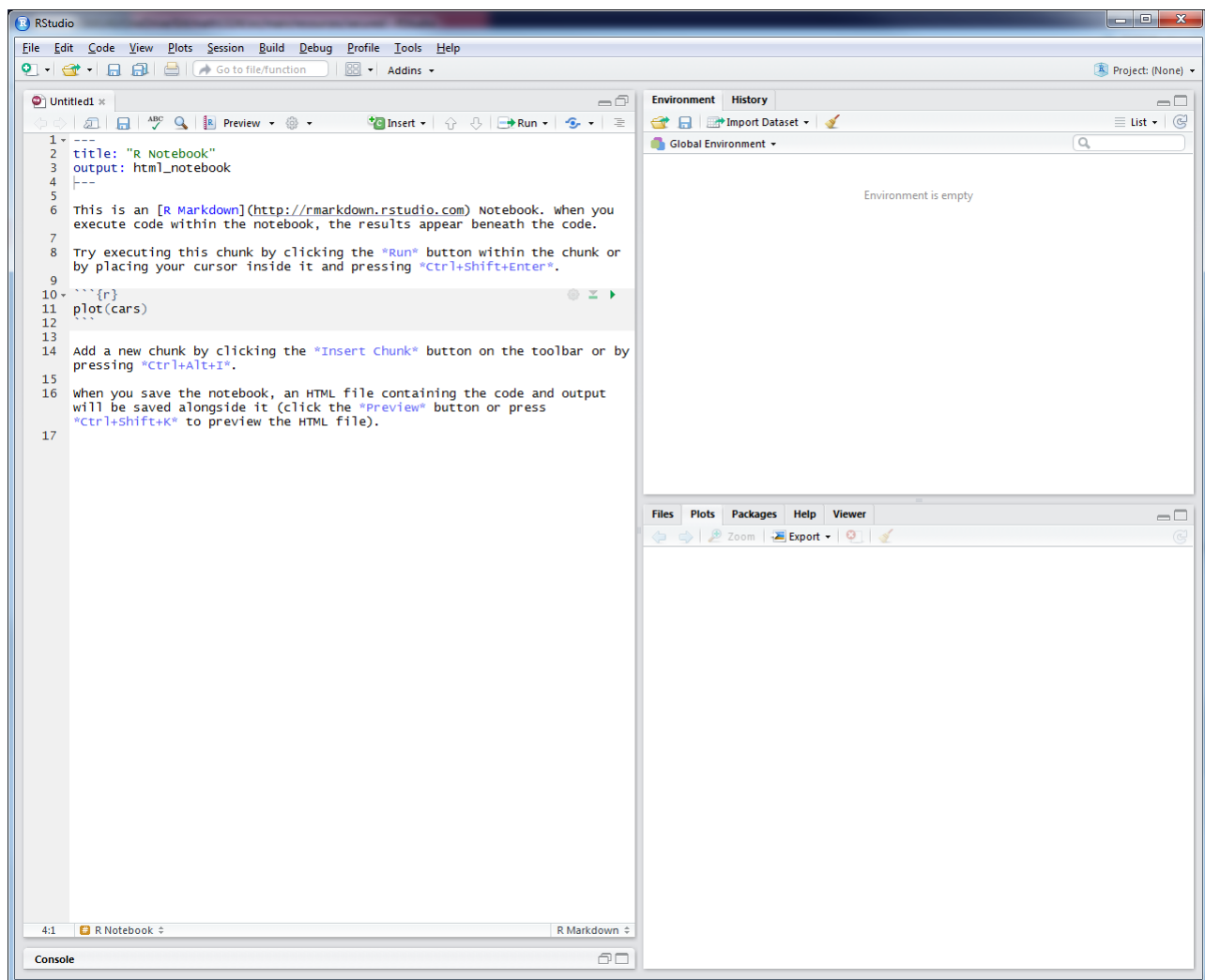
Your statistical analysis must be reproducible, not only for yourself, but more importantly for others. Each stage of your analysis must be documented and explainable to others. The ultimate test of reproducibility is to give someone your R Markdown file and data and for them to replicate all of your analysis by running it through R. In practice, things are never this simple, but that's certainly the goal we should have in mind.

## Creating an R Markdown Document in R Studio

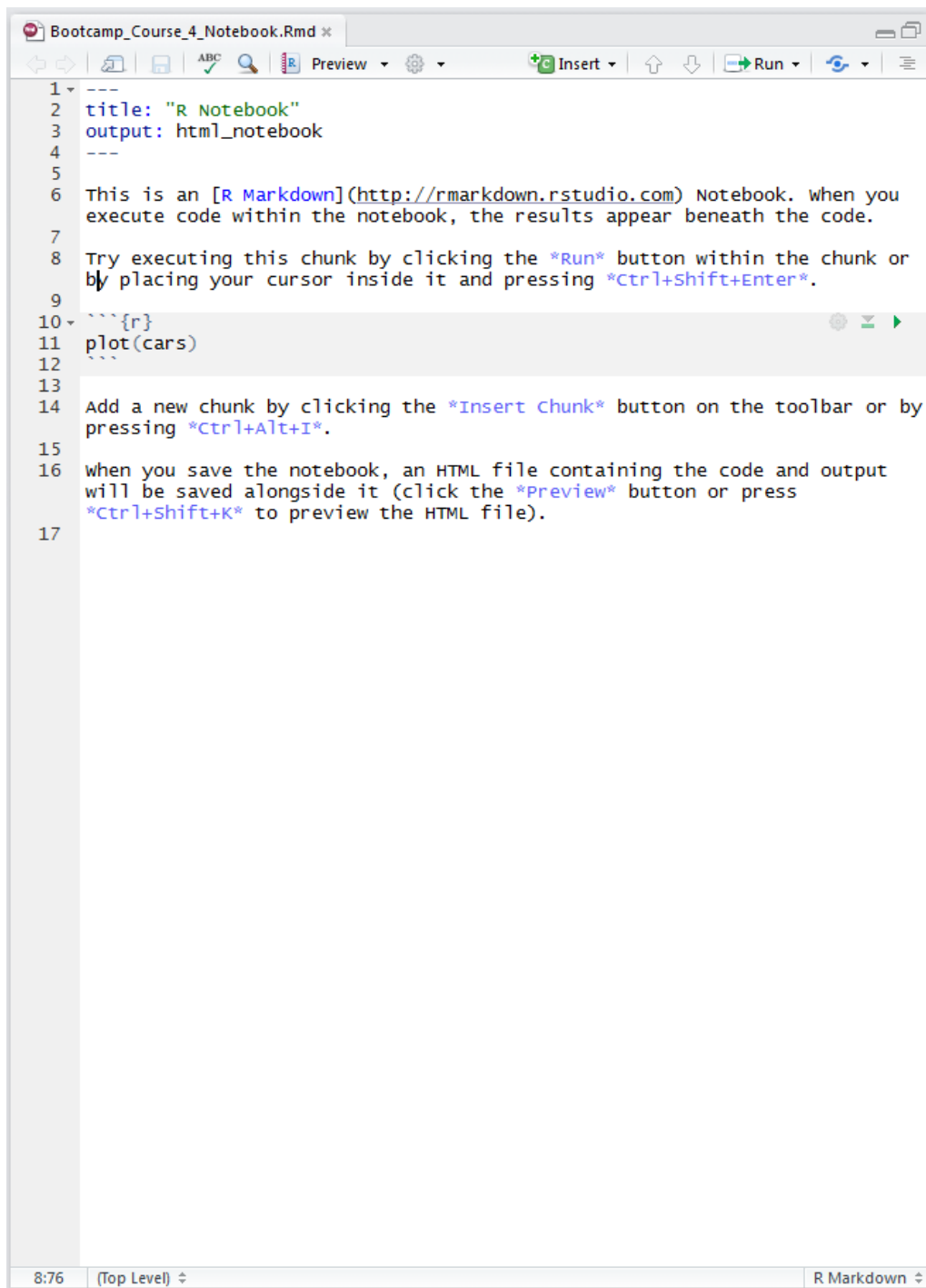
First, ensure you have an up-to-date version of RStudio (<https://www.rstudio.com/>). In RStudio go to **File** → **New File** → **R Notebook**. You could also select **R Markdown** (R Notebooks are a type of Markdown file), but for this course, we will stick to the Notebook output type. Allow RStudio to install or update any necessary packages. If you have trouble, run the following installation which will force all the required packages to also install:

```
install.packages("rmarkdown")
```

A new notebook will appear in the RStudio script window. This usually forces your console window to disappear, so don't worry. As we will discover, we won't need it.



Save your Notebook. It will save with an `.Rmd` extension which tells you it is an R Markdown file. The template includes some basic examples of R Markdown.

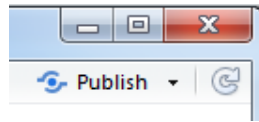


An RStudio Notebook, or any other R Markdown file, can be both published, e.g. as a HTML, or run like an ordinary R script file. Most of the file looks like a word document. The screenshot above shows the R Notebook template included when you create a new R Notebook. The first paragraph is descriptive text. However, you can also see R code. R code is written in “Chunks”. You can see the R code chunk in the screenshot above. It’s highlighted in grey and sits in between `{r}` and syntax. For example, the R code chunk:

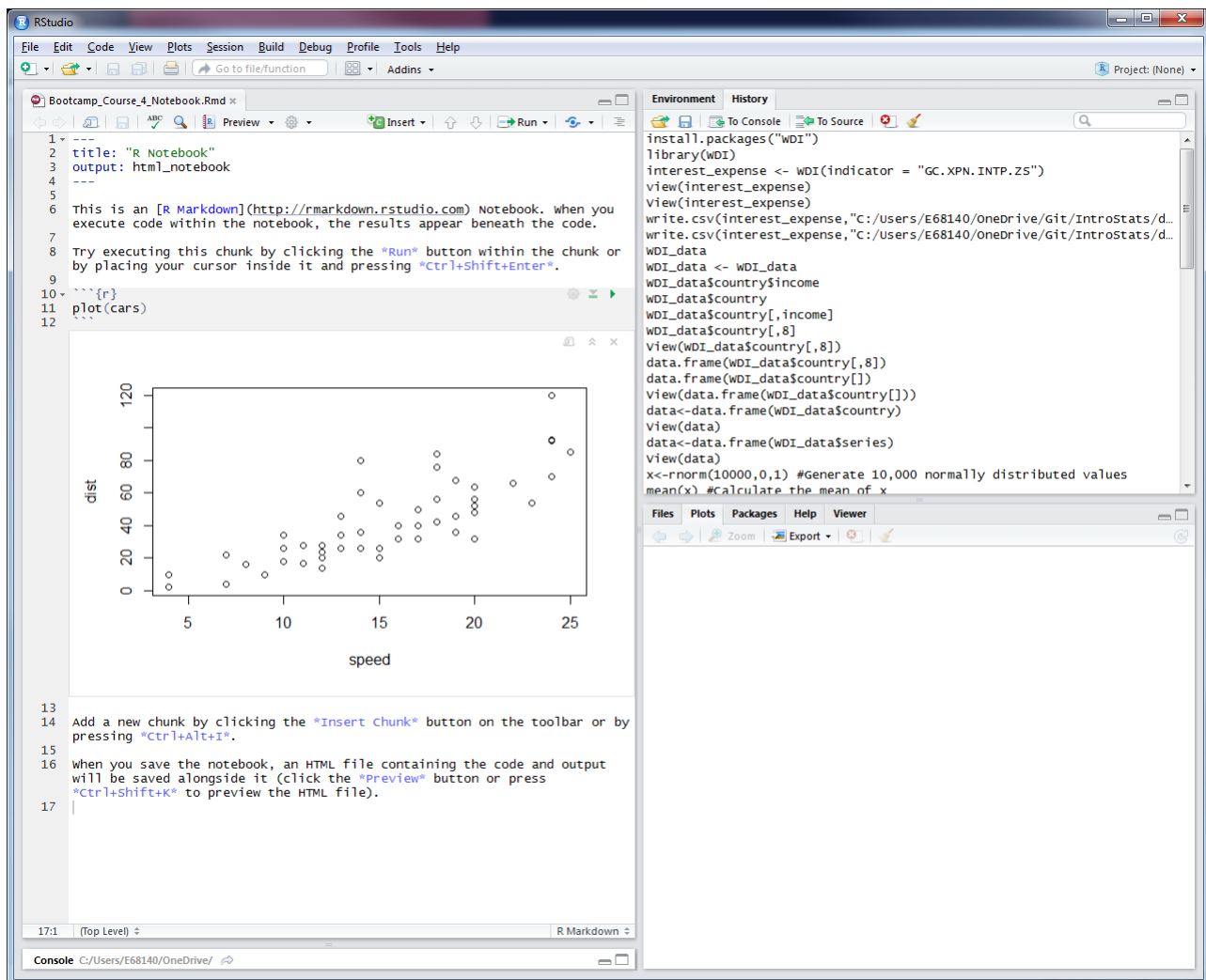
```
{r}  
plot(cars)
```

performs a `plot` function on the dataset object `cars`. If we want to run the R code in an R Notebook, we press the green arrow “run” button that appears in the top right-hand side of the chunk. Or, we place the cursor in the chunk and click on the **Chunks** menu

(top right hand corner of the script screen). This menu gives us plenty of options. We can use Insert chunk to automatically type out the `{r}` and ````` syntax, jump to different chunks, and run previous code, current chunk or next chunk. You also have the option to run all chunks together in the R Notebook Document. This is what we would do if we wanted to replicate someone else's analysis.



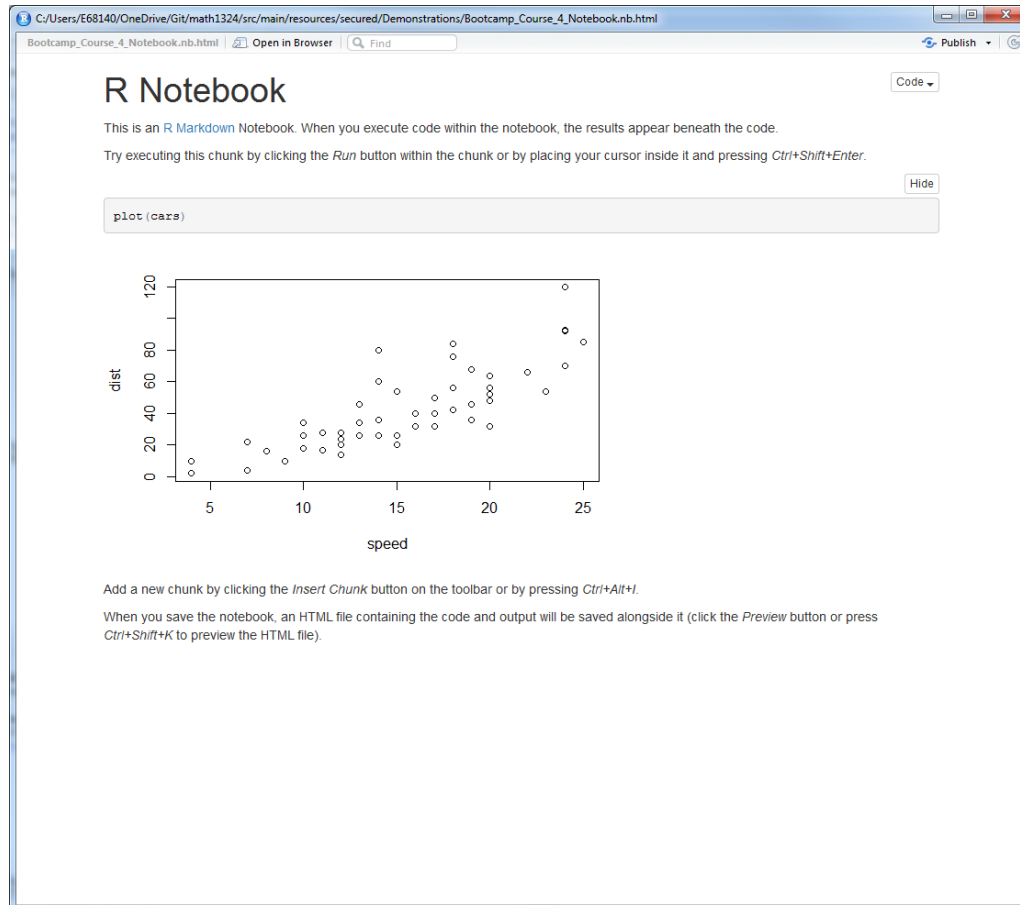
Let's run the first R chunk above.



Chunks are sent to the console and executed like a script. The fantastic thing about Notebooks is that output appears underneath the chunk instead of in the console or in the plot window. This significantly improves RStudio workflow and the enjoyment of using R. In the screenshot above you can see the plot of the `cars` data.

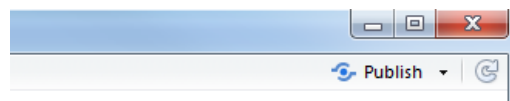
R Markdown documents are fully reproducible. This means you could send the data and the R Markdown report to another investigator and they should be able to reproduce everything you did. Reproducibility is crucial in statistics, analytics and data science. However, in practice, it is never that simple. Due to a multitude of different factors, you will always have some issues with reproducing someone else's report. A common issue is different versions of packages.

If we've finished analysing our data and writing our report, we click the **Preview** button to create a HTML report (see the screenshot below). Things like R output and plots will be automatically added and scaled to the document. This can save you a lot of editing time. It also includes the option to hide the code. This is a nice touch if you are sharing your report with stakeholders that might not want to see your code. There are many other output options available including PDF and Microsoft Word, but these formats lose interactive feature and can be tricky to compile.



A HTML document will be saved to your working directory where you saved your R Markdown file and a preview will appear in the RStudio browser. You can open this file in any web browser or word processor and save in a different format (e.g. PDF or .doc).

You also have the option to **Publish** this document directly to a hosting service. This allows you to publish your Notebooks online or provide secure access to your team or clients.



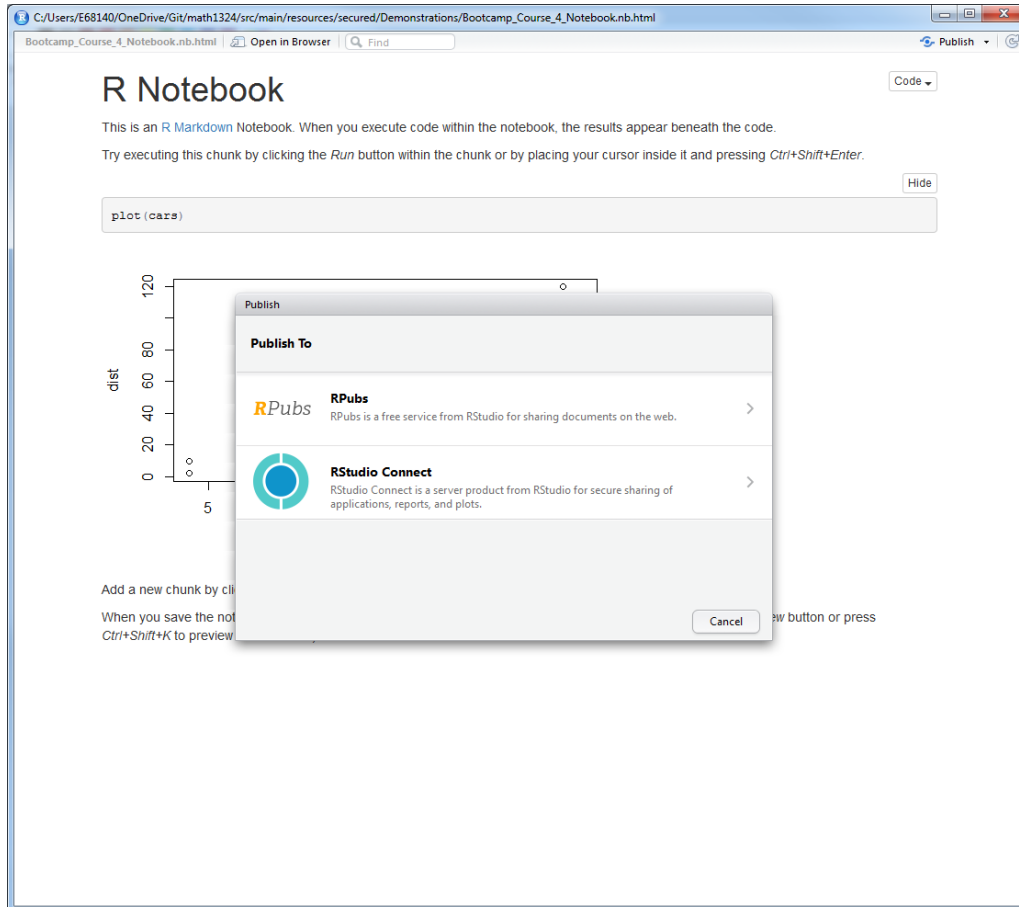
RStudio offers a few different options:

- RPubS (<http://rpubs.com/>): Free, open access. Only use this for open data projects.
- Shiny Apps (<https://www.shinyapps.io/>): A free account include 5 applications (notebooks) that can be public or secured. Paid accounts get more. Shiny Apps can also host interactive Shiny Applications.

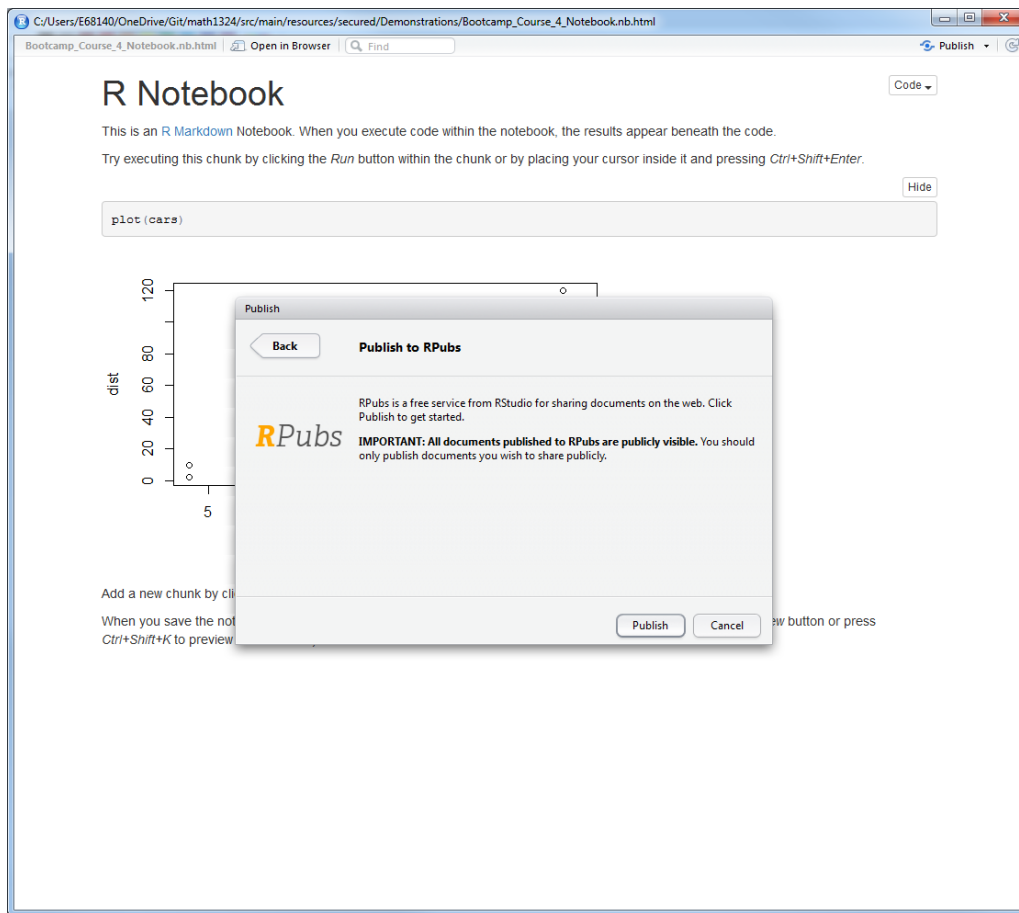
If you run your own server, you can use the following products to securely host and share your notebooks online.

- R Shiny Server (<https://www.rstudio.com/products/shiny/shiny-server/>): Free open source server application for hosting R Markdown and Shiny apps online. RStudio also offers a Pro version with lots of extra goodies.
- RStudio Connect (<https://www.rstudio.com/products/connect/>): Premium, paid hosting server for R Markdown reports, dashboards, Shiny apps and more.

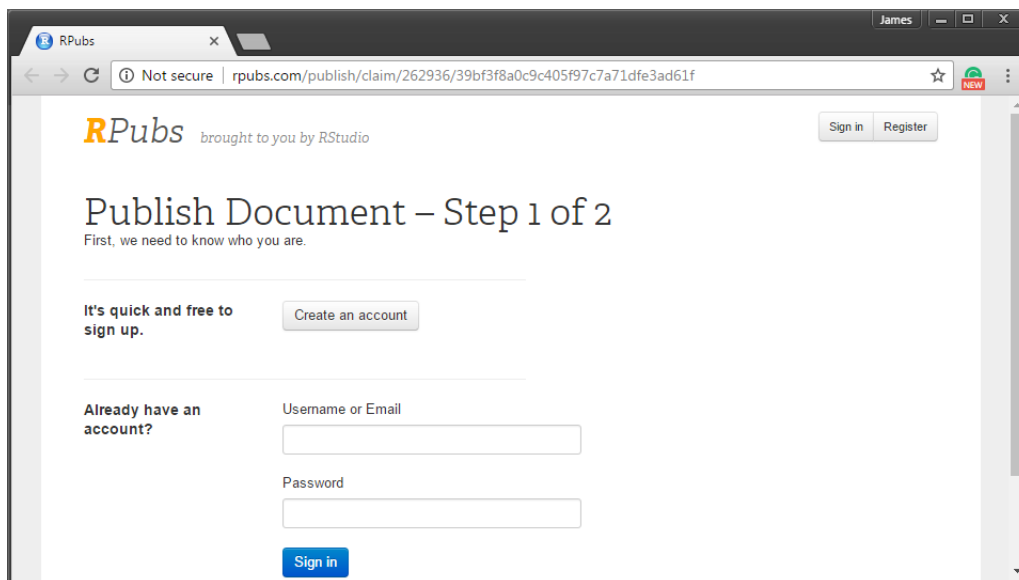
For your final assignment, you will need to publish your work to RPubs. Click **Publish**.



Select RPubs.



Click **Publish**



Sign-in or create a new account. Your R Notebook will be published in no time. Once published, you can copy the URL from the published report to share the report with others. RPubs can also publish slides (see the links at the bottom of this page).

## R Markdown Syntax Basics

R Markdown is a powerful, light markup language. Using simple code, you can learn to format and present beautiful reports that integrate with your R code and statistical analysis. For detailed guides please see the Reference Guide (<http://www.rstudio.com/wp->



content/uploads/2015/03/rmarkdown-reference.pdf) and Cheatsheet (http://www.rstudio.com/wp-content/uploads/2016/03/rmarkdown-cheatsheet-2.0.pdf) produced by RStudio.

The following code will demonstrate some of the basics of R Markdown.

## YAML Header

Start your documents with a YAML Header. There are many options that can be changed. The following header includes the basics. Note the use of `---` to start and finish the header.

```
---
title: "R Bootcamp: Course 4"
author: "James Baglin"
date: "Tuesday, March 28, 2017"
output: html_notebook
---
```

## Headings

Using different levels of headings is a great way to structure your reports.

```
# Header 1

## Header 2

### Header 3

#### Header 4 etc...
```

## Text Formatting, Links, and Math Type

There are simple markdown codes used to format text, add hyperlinks and write mathematical formulas.

To write a paragraph, use plain text. You can use the following syntax to *italicise*, **bold**, superscript<sup>2</sup>, hyperlink, [link](http://www.rmit.edu.au/), and write inline equations using Latex code,  $z = \frac{x - \mu}{\sigma}$ .

Which will render as...

To write a paragraph, use plain text. You can use the following syntax to *italicise*, **bold**, superscript<sup>2</sup>, hyperlink, link (<http://www.rmit.edu.au/>), and write in-line equations using LaTeX code (<https://en.wikibooks.org/wiki/LaTeX/Mathematics>),  
$$z = \frac{x - \mu}{\sigma}.$$

## Lists

Here are some examples of lists. To get second levels working, ensure you use a double tab indent.

```
* Unordered list
* Item 2
    + Sub-item 1
    + Sub-item 2

1. Ordered list
2. Item 2
    + Sub-item 1
    + Sub-item 2
```

Which will appear as...

- Unordered list
- Item 2
  - Sub-item 1
  - Sub-item 2
- 1. Ordered list
- 2. Item 2
  - Sub-item 1
  - Sub-item 2

## Tables

Tables can be manually entered.

**\*\*Table 1:\*\*** Mean and Median for Speed and Distance

```
Variable|Mean |Median
-----|-----|-----
Speed   |15.40|15.0
Distance|42.98|36.0
```

Renders as...

**Table 1:** Mean and Median for Speed and Distance

Variable	Mean	Median
Speed	15.40	15.0
Distance	42.98	36.0

You can also use the `knitr::kable()` function to convert tables and data frames from R into nicely formatted HTML tables.

```
```${r}
library(knitr)
knitr::kable(anscombe, caption = "Anscombe's Quartet Data")
```
```

This appears like...

```
## Warning: package 'knitr' was built under R version 3.5.3
```

Anscombe's Quartet Data

| x1 | x2 | x3 | x4 | y1   | y2   | y3    | y4    |
|----|----|----|----|------|------|-------|-------|
| 10 | 10 | 10 | 8  | 8.04 | 9.14 | 7.46  | 6.58  |
| 8  | 8  | 8  | 8  | 6.95 | 8.14 | 6.77  | 5.76  |
| 13 | 13 | 13 | 8  | 7.58 | 8.74 | 12.74 | 7.71  |
| 9  | 9  | 9  | 8  | 8.81 | 8.77 | 7.11  | 8.84  |
| 11 | 11 | 11 | 8  | 8.33 | 9.26 | 7.81  | 8.47  |
| 14 | 14 | 14 | 8  | 9.96 | 8.10 | 8.84  | 7.04  |
| 6  | 6  | 6  | 8  | 7.24 | 6.13 | 6.08  | 5.25  |
| 4  | 4  | 4  | 19 | 4.26 | 3.10 | 5.39  | 12.50 |

| x1 | x2 | x3 | x4 | y1    | y2   | y3   | y4   |
|----|----|----|----|-------|------|------|------|
| 12 | 12 | 12 | 8  | 10.84 | 9.13 | 8.15 | 5.56 |
| 7  | 7  | 7  | 8  | 4.82  | 7.26 | 6.42 | 7.91 |
| 5  | 5  | 5  | 8  | 5.68  | 4.74 | 5.73 | 6.89 |

## Chunks

To include R code with console output use the “```{r}`” chunk syntax.

```
``{r}
summary(cars)
``
```

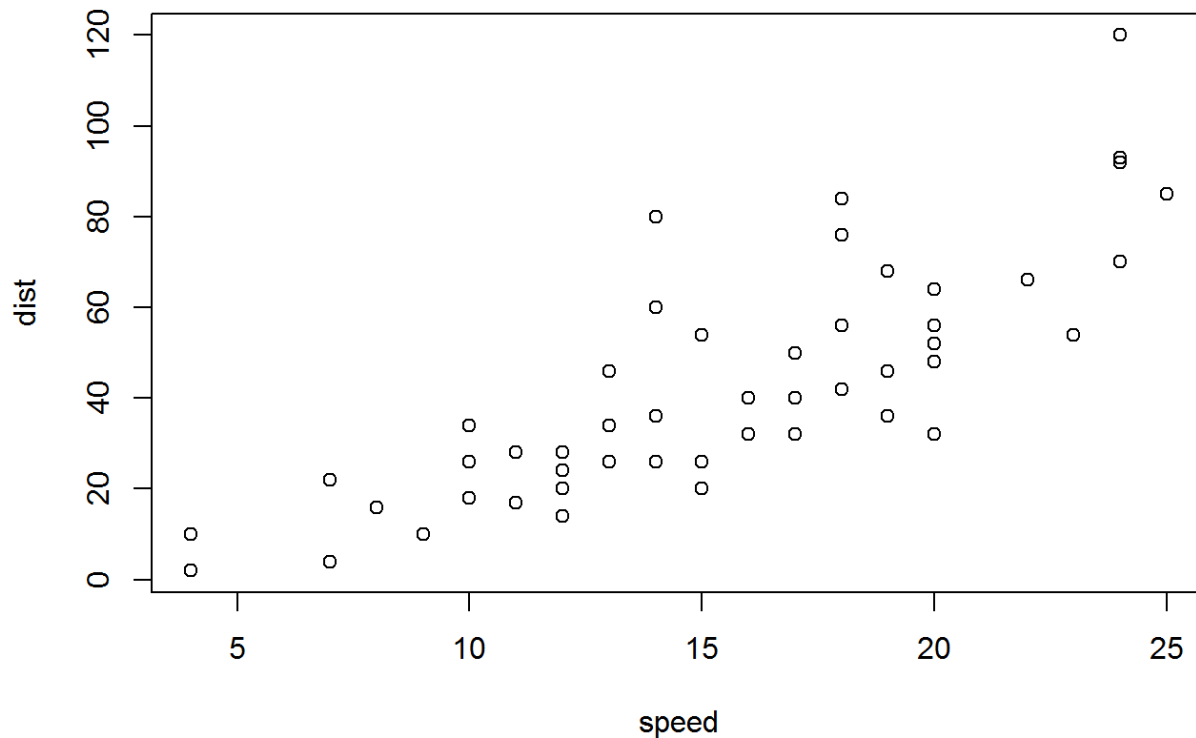
This returns...

```
##      speed      dist
##  Min.   : 4.0    Min.   :  2.00
## 1st Qu.:12.0    1st Qu.: 26.00
##  Median :15.0    Median : 36.00
##  Mean   :15.4    Mean   : 42.98
## 3rd Qu.:19.0    3rd Qu.: 56.00
##  Max.   :25.0    Max.   :120.00
```

To report output without the R code, use `echo = FALSE` option,

```
``{r, echo=FALSE}
plot(cars)
``
```

This reports only the output...



## In-line Code

Using R code in-line allows you to create dynamic documents that can update values in your report by re-running the analysis. Here's an example...

```
You can also use R Markdown syntax to report R code inline. For example,  
the mean speed was 15.4, *SD* = 5.29, *n* = 50. Ensure you enclose the sy  
ntax using two backward single quotes `.
```

This will render as follows:

You can also use R Markdown syntax to report R code inline. For example, the mean speed was 15.4, *SD* = 5.29, *n* = 50. Ensure you enclose the syntax using two backward single quotes `.

Very handy if you have dynamic reports that are updated on a regular basis. Re-rendering the Markdown files will automatically update all the values.

## HTML Preview

You can download a copy of the R Markdown Notebook with the code above to try for yourself here ([Demonstrations/Bootcamp\\_Course\\_4\\_Notebook\\_Example.Rmd](#)). The window below shows you what the notebook will look like once it is previewed or

# R Bootcamp: Course 4

Code ▼

*James Baglin*

*Tuesday, March 28, 2017*

## Header 1

## Header 2

### Header 3

#### Header 4 etc...

To write a paragraph, use plain text. You can use the following syntax to *italicise*, **bold**, superscript<sup>2</sup>, hyperlink, link (<http://www.rmit.edu.au/>), and write inline equations using Latex code,

$$z = \frac{x - \mu}{\sigma}.$$

- Unordered list
  - Item 2
    - Sub-item 1
    - Sub-item 2
1. Ordered list
  2. Item 2
    - Sub-item 1
    - Sub-item 2

**Table 1:** Mean and Median for Speed and Distance

| Variable | Mean  | Median |
|----------|-------|--------|
| Speed    | 15.40 | 15.0   |
| Distance | 42.98 | 36.0   |

## Further Reading

There is a lot of amazing things you can do with RStudio and R Markdown (<http://rmarkdown.rstudio.com/formats.html>). Here are some links to help you learn more.

- R Markdown Cheatsheet (<https://www.rstudio.com/wp-content/uploads/2016/03/rmarkdown-cheatsheet-2.0.pdf>)
- R Markdown Websites ([http://rmarkdown.rstudio.com/rmarkdown\\_websites.html](http://rmarkdown.rstudio.com/rmarkdown_websites.html)) - This website is an example of R Markdown Website.

- R Markdown Presentations (<http://rmarkdown.rstudio.com/lesson-11.html>). A variety of formats are offered including beamer, ioslides, slidy and revealjs.
- Shiny Applications (<https://shiny.rstudio.com/>): Interactive applications and data visualisations.
- Dashboards (<http://rmarkdown.rstudio.com/lesson-12.html>): Create beautiful dashboards quickly using the power of R and Markdown.
- Interactive Documents (<http://rmarkdown.rstudio.com/lesson-14.html>): Embed a wide range of interactive features into your HTML R Markdown pages.
- Plotly (<https://plot.ly/r/getting-started/>): Embed interactive, web-based, D3.js driven plots into your R Markdown reports.