

資料科學 Data Science 系列

R 語言資料分析實務 (3)

在 RStudio 中撰寫文件

姓名：羅左欣

日期：2016/10/24(一)



本著作係採用創用 CC 姓名標示-非商業性-相同方式分享 3.0 台灣 授權條款授權。

部落格：<http://shouzo.github.io/>

Agenda

- (一) Prepare : 預備工作**
- (二) Basic : 基本介紹與操作**
- (三) Theme : 撰寫文件**
- (四) Reference : 學習資源**

(一) Prepare : 預備工作

(一) Prepare :
預備工作

(一) Prepare：預備工作

A screenshot of the RStudio application window. The title bar reads 'RStudio'. The menu bar includes 'File', 'Edit', 'Code', 'View', 'Plots', 'Session', 'Build', 'Debug', 'Tools', and 'Help'. Below the menu bar is a toolbar with icons for file operations and a search bar. The main console area displays the following text:

```
R version 3.2.3 (2015-12-10) -- "Wooden Christmas-Tree"  
Copyright (C) 2015 The R Foundation for Statistical Computing  
Platform: x86_64-pc-linux-gnu (64-bit)  
  
R 是免費軟體，不提供任何擔保。  
在某些條件下您可以將其自由散布。  
用 'license()' 或 'licence()' 來獲得散布的詳細條件。  
  
R 是個合作計劃，有許多人為之做出了貢獻。  
用 'contributors()' 來看詳細的情況並且  
用 'citation()' 會告訴您如何在出版品中正確地參照 R 或 R 套件。  
  
用 'demo()' 來看一些示範程式，用 'help()' 來檢視線上輔助檔案，或  
用 'help.start()' 透過 HTML 瀏覽器來看輔助檔案。  
用 'q()' 離開 R。  
  
[Workspace loaded from ~/.RData]  
>
```

The right-hand side of the window shows the 'Environment' and 'History' panes. The 'Environment' pane is currently empty, displaying the text 'Environment is empty'. Below it are panes for 'Files', 'Plots', 'Packages', and 'Help', each with its own set of icons and controls.

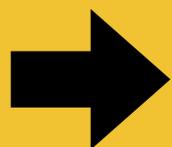
在這個系列的簡報中，主要以 "RStudio" 做為主要軟體。

(二) Basic : 基本介紹與操作

(二) Basic : 基本介紹與操作

(二) Basic : 基本介紹與操作

Markdown 概觀



<http://markdown.tw/>

(二) Basic：基本介紹與操作

動手練習： 使用 HackMD

➔ 平台網址：<https://hackmd.io/>

(三) Theme : 撰寫文件

(三) Theme : 撰寫文件

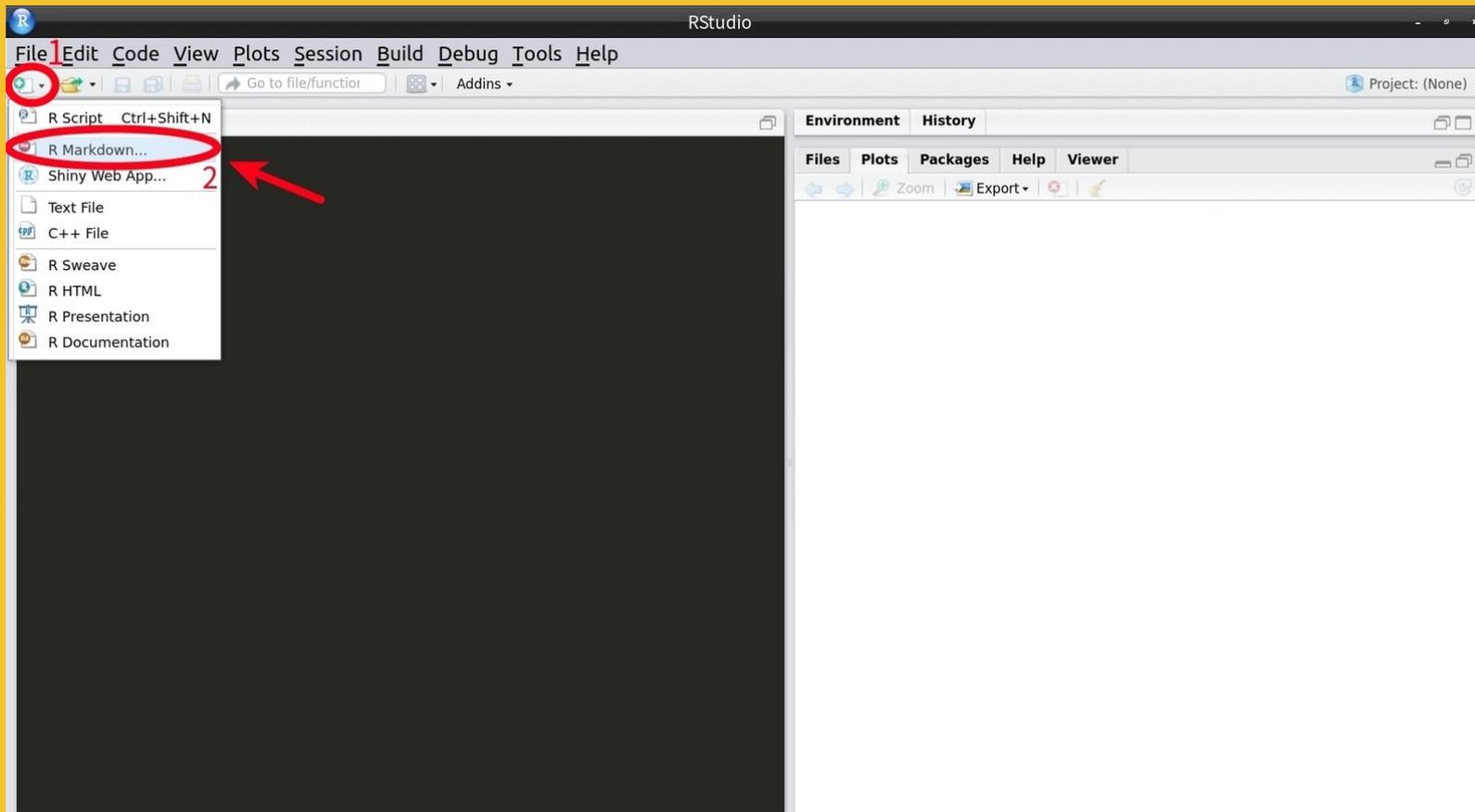
(三) Theme：撰寫文件

R Markdown

➔ 功能導覽：<https://goo.gl/MruFBA>

(三) Theme：撰寫文件

選擇 R Markdown ...



(三) Theme : 撰寫文件

輸入標題、作者...

New R Markdown

Document
Presentation
Shiny
From Template

Title: 20161024_Demo

Author: shouzo

Default Output Format:

HTML
Recommended format for authoring (you can switch to PDF or Word output anytime).

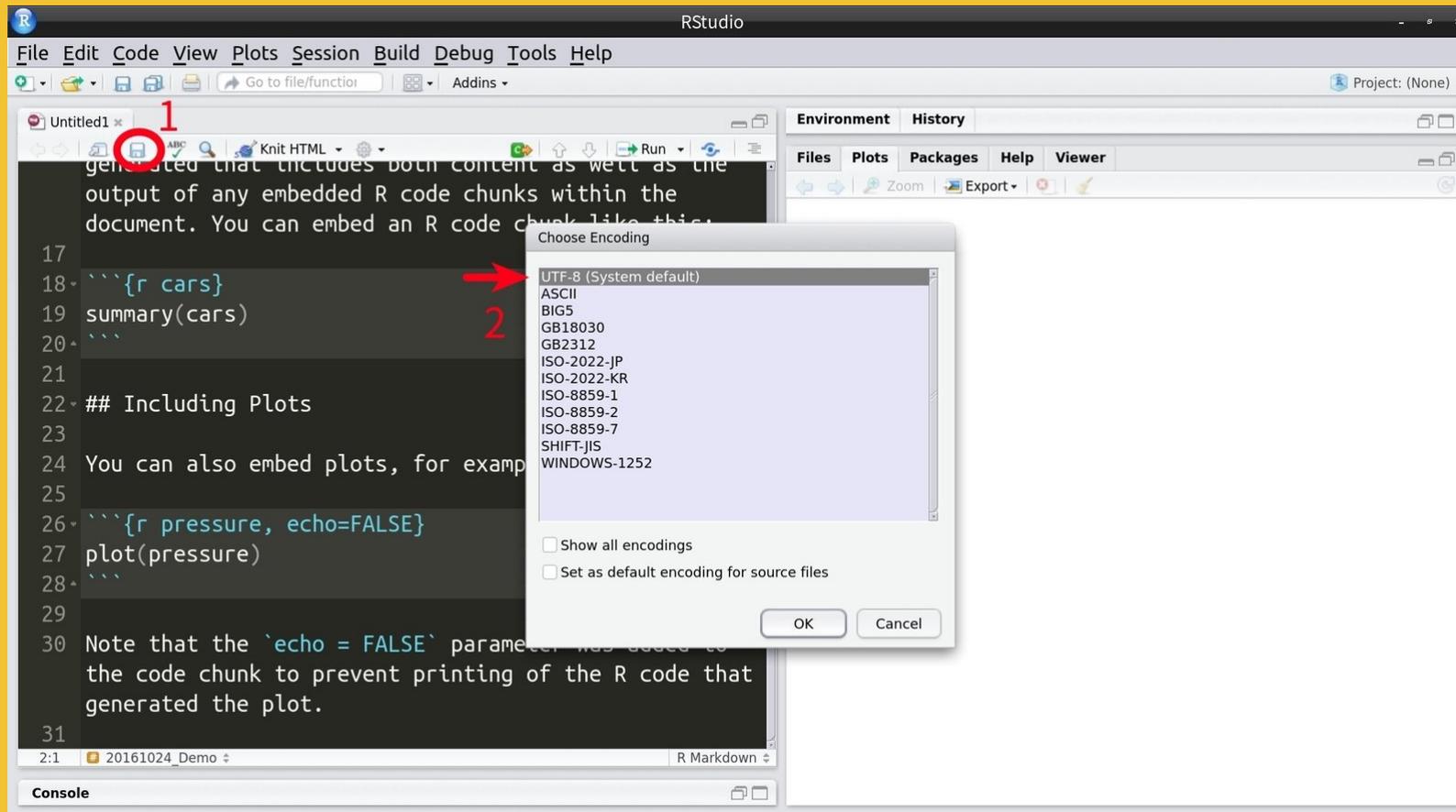
PDF
PDF output requires TeX (MiKTeX on Windows, MacTeX 2013+ on OS X, TeX Live 2013+ on Linux).

Word
Previewing Word documents requires an installation of MS Word (or Libre/Open Office on Linux).

OK Cancel

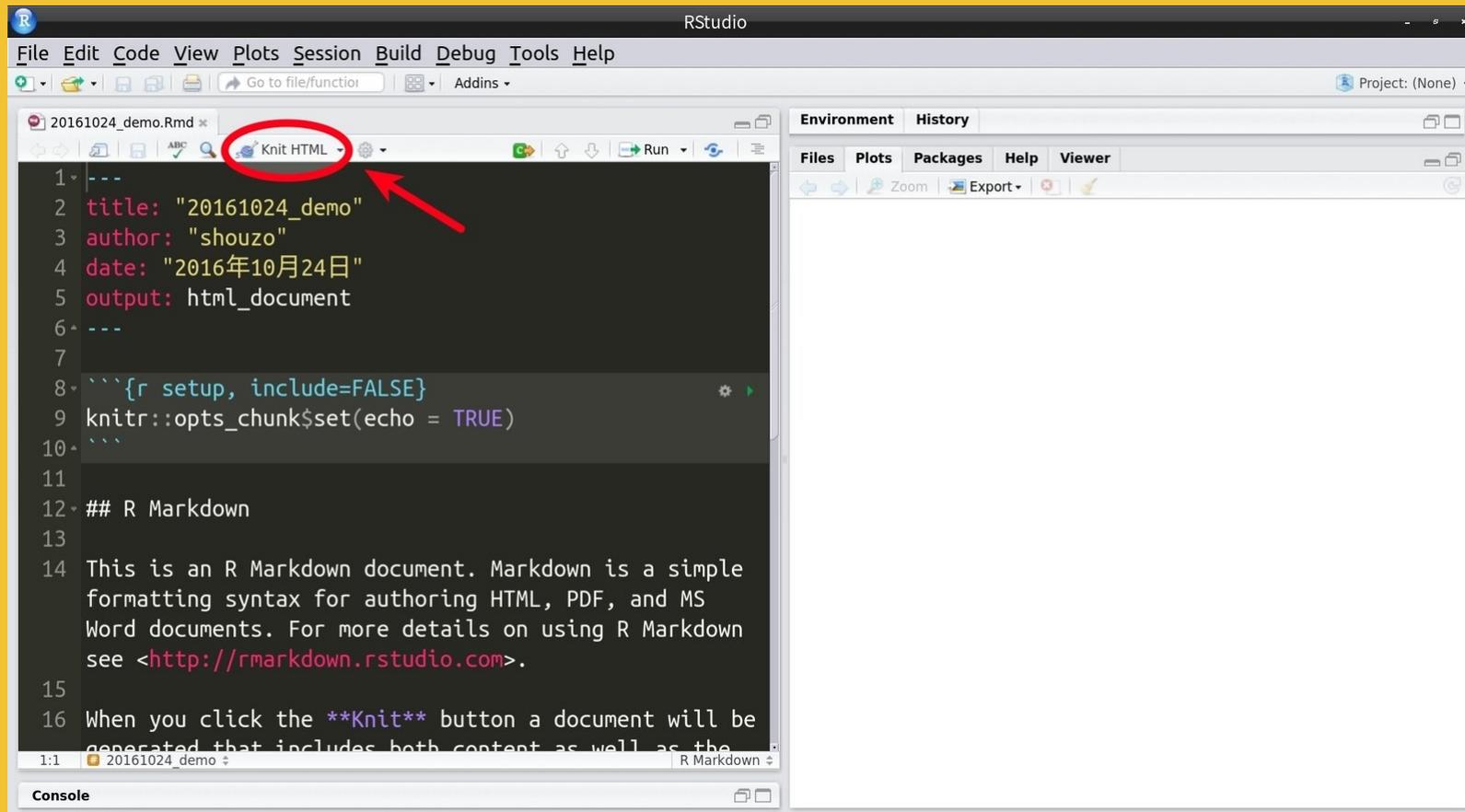
(三) Theme : 撰寫文件

儲存檔案 ...



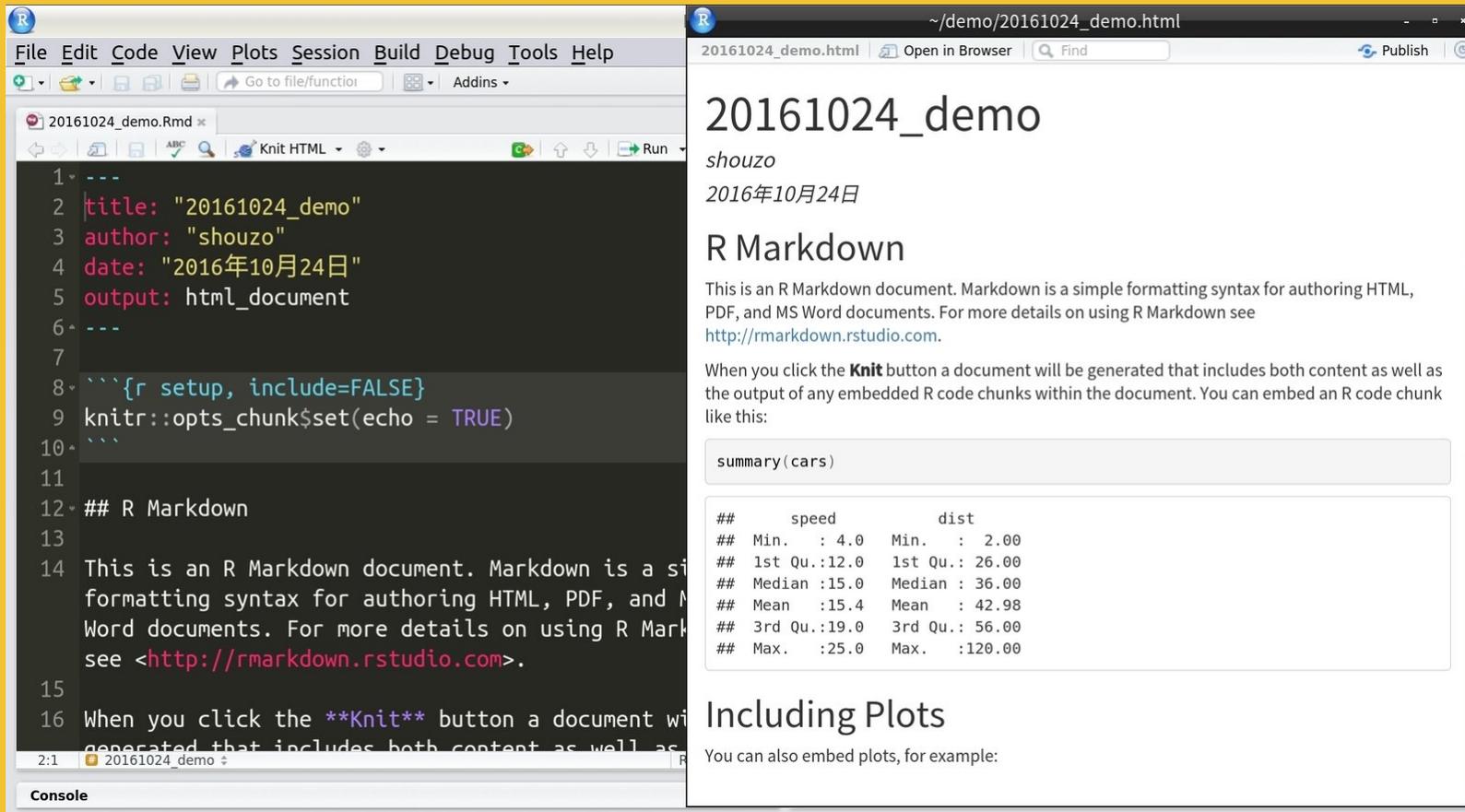
(三) Theme : 撰寫文件

點選 Knit HTML ...



(三) Theme : 撰寫文件

輸出結果



The image shows two side-by-side windows from the RStudio interface. The left window displays the source R Markdown file, and the right window shows the rendered HTML output.

Left Window (Source Code):

```
1 ---
2 title: "20161024_demo"
3 author: "shouzo"
4 date: "2016年10月24日"
5 output: html_document
6 ---
7
8 ```{r setup, include=FALSE}
9 knitr::opts_chunk$set(echo = TRUE)
10 ```
11
12 ## R Markdown
13
14 This is an R Markdown document. Markdown is a simple
15 formatting syntax for authoring HTML, PDF, and MS Word
16 documents. For more details on using R Markdown see
17 <http://rmarkdown.rstudio.com>.
18
19 When you click the Knit button a document will be
20 generated that includes both content as well as the
21 output of any embedded R code chunks within the document.
22 You can embed an R code chunk like this:
23
24 ```{r}
25 summary(cars)
26 ```
```

Right Window (Rendered Output):

20161024_demo

shouzo
2016年10月24日

R Markdown

This is an R Markdown document. Markdown is a simple formatting syntax for authoring HTML, PDF, and MS Word documents. For more details on using R Markdown see <http://rmarkdown.rstudio.com>.

When you click the **Knit** button a document will be generated that includes both content as well as the output of any embedded R code chunks within the document. You can embed an R code chunk like this:

```
summary(cars)
```

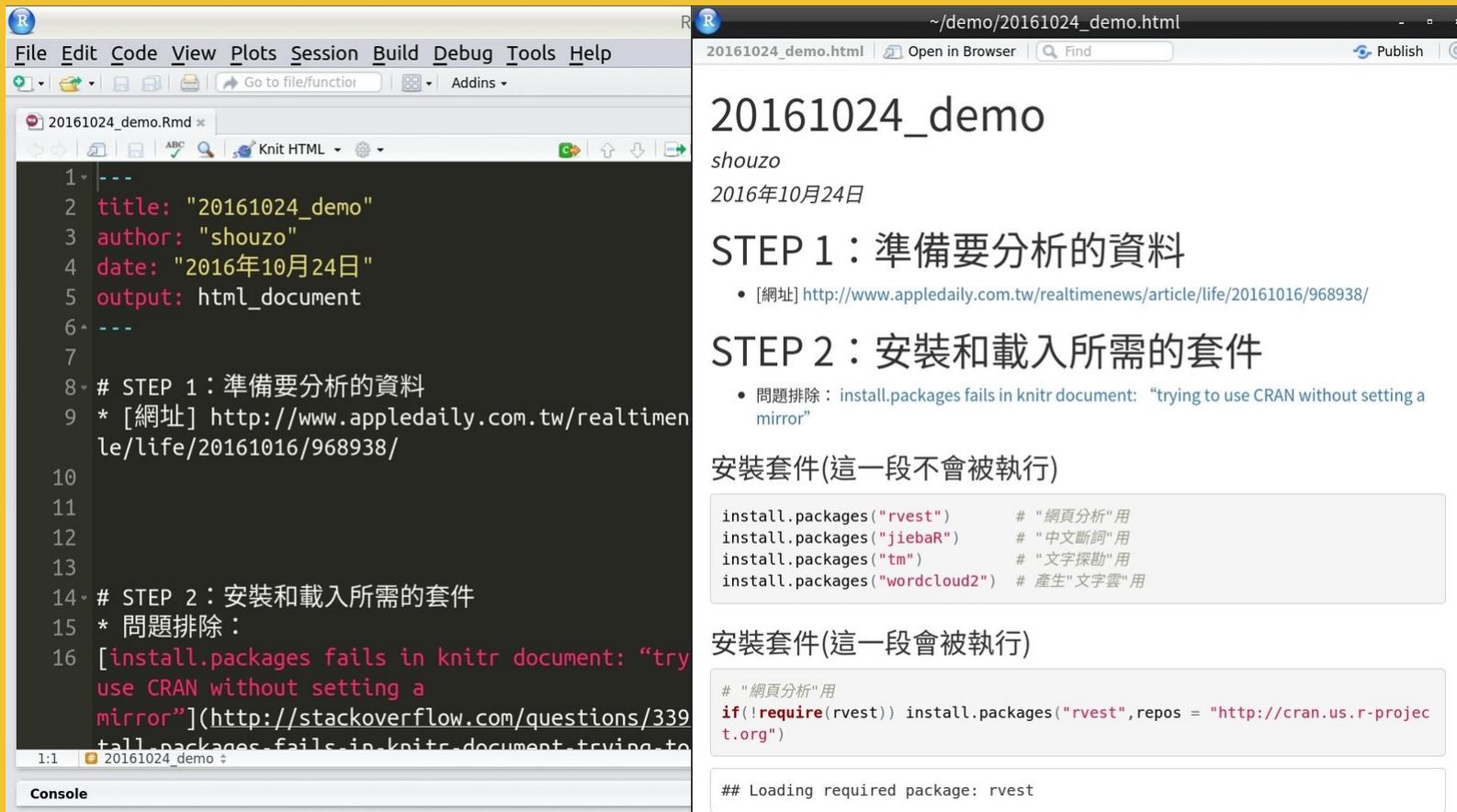
##	speed	dist
## Min.	: 4.0	Min. : 2.00
## 1st Qu.:	:12.0	1st Qu.: 26.00
## Median :	:15.0	Median : 36.00
## Mean :	:15.4	Mean : 42.98
## 3rd Qu.:	:19.0	3rd Qu.: 56.00
## Max. :	:25.0	Max. : 120.00

Including Plots

You can also embed plots, for example:

(三) Theme：撰寫文件

試著撰寫自己的文章...



The image shows two windows from RStudio. The left window displays the R Markdown source code for a document titled '20161024_demo.Rmd'. The code includes a YAML header with title, author, date, and output format. It is divided into two steps: Step 1 (preparation of data) and Step 2 (installation and loading of packages). Step 2 includes a troubleshooting note about CRAN and a code block for installing 'rvest', 'jiebaR', 'tm', and 'wordcloud2' packages. The right window shows the rendered HTML output of the document, which includes the title, author, date, and the two steps. The code blocks for installing packages are shown as pre-rendered code snippets.

```
1 ---
2 title: "20161024_demo"
3 author: "shouzo"
4 date: "2016年10月24日"
5 output: html_document
6 ---
7
8 # STEP 1：準備要分析的資料
9 * [網址] http://www.appledaily.com.tw/realtimenews/article/life/20161016/968938/
10
11
12
13
14 # STEP 2：安裝和載入所需的套件
15 * 問題排除：
16 [install.packages fails in knitr document: "trying to use CRAN without setting a mirror"](http://stackoverflow.com/questions/33911111/install-packages-fails-in-knitr-document-trying-to)
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
```

20161024_demo

shouzo

2016年10月24日

STEP 1：準備要分析的資料

- [網址] <http://www.appledaily.com.tw/realtimenews/article/life/20161016/968938/>

STEP 2：安裝和載入所需的套件

- 問題排除：install.packages fails in knitr document: "trying to use CRAN without setting a mirror"

安裝套件(這一段不會被執行)

```
install.packages("rvest") # "網頁分析"用
install.packages("jiebaR") # "中文斷詞"用
install.packages("tm") # "文字探勘"用
install.packages("wordcloud2") # 產生"文字雲"用
```

安裝套件(這一段會被執行)

```
# "網頁分析"用
if(!require(rvest)) install.packages("rvest", repos = "http://cran.us.r-project.org")

## Loading required package: rvest
```

(三) Theme：撰寫文件

補充資料：

RStudio R 簡報製作教學：使用 RMarkdown 與 Slidify

➡ 網址：<https://goo.gl/DGIhXS>

(四) Reference : 學習資源

**(四) Reference :
學習資源**

(四) Reference：學習資源

線上教材

一、中文教材

1. **R 語言翻轉教室** - Wush Wu、Chih Cheng Liang、Johnson Hsieh

<http://datascienceandr.org/>

2. **手把手教你 R 語言資料分析實務** - 張毓倫&陳柏亨

<http://goo.gl/18mwug>

3. **R 軟體與資料探勘之開發與應用** - 陳志華

<https://goo.gl/NPdzzP>



二、英文教材

1. **DataCamp**

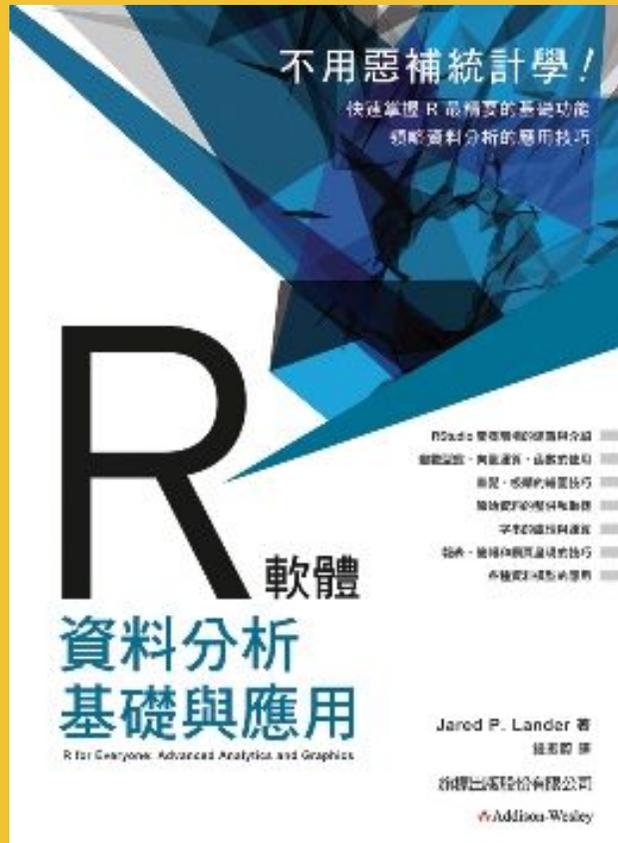
<https://www.datacamp.com/>

2. **R for Data Science**

<http://r4ds.had.co.nz/>

(四) Reference：學習資源

推薦書籍



R 軟體資料分析基礎與應用

作者：Jared P. Lander

譯者：鍾振蔚

出版社：旗標

(四) Reference：學習資源

相關社群

台灣資料科學年會

<https://www.facebook.com/twdsconf/>



Taiwan R User Group

<https://www.facebook.com/Tw.R.User/>

資料視覺化 / Data Visualization

<https://www.facebook.com/data.visualize/>



Q & A