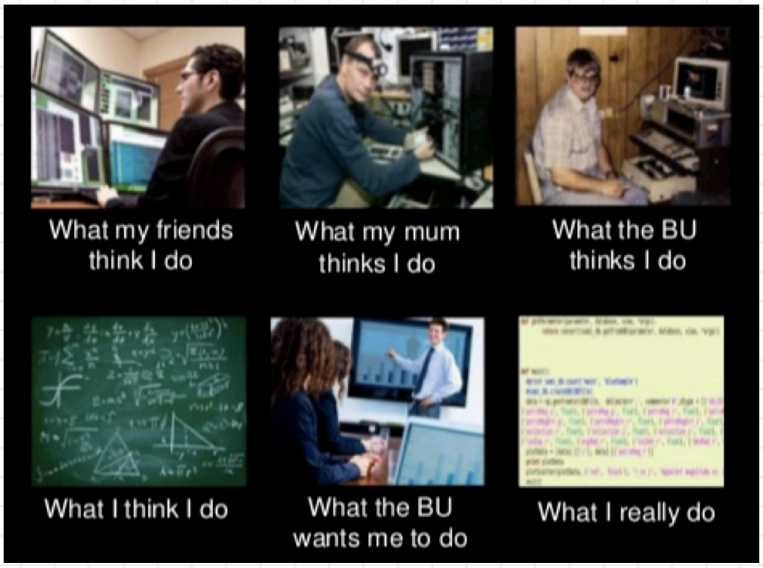
Introduction to R markdown
Dynamic documents for R

Wush Wu, Johnson Hsieh

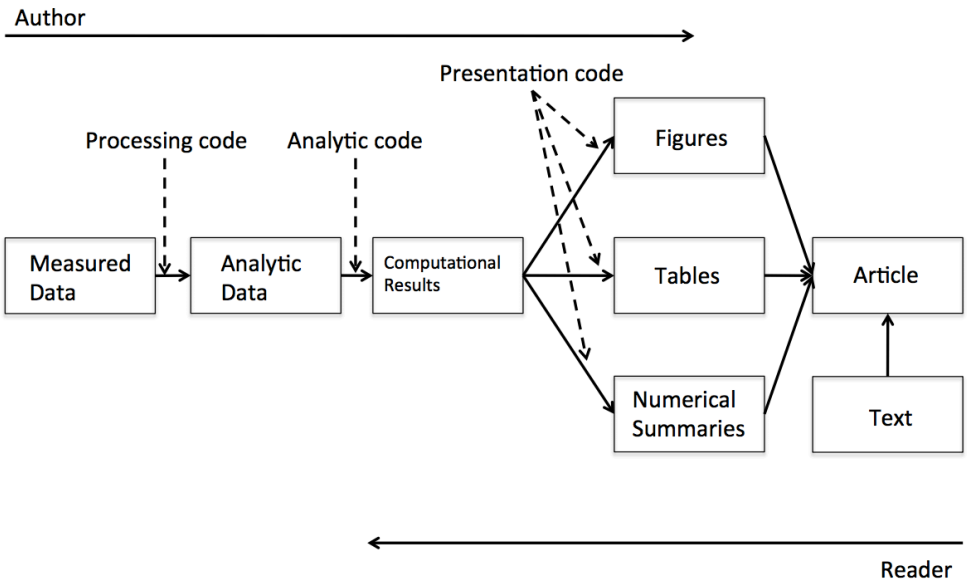
dsp.im

What does data analyst do



http://goo.gl/52THNf

What does data analyst actually do



https://www.coursera.org/course/repdata

Why you need R Markdown

- 資料分析做不完
- 說明文件沒空寫
- 教授要求換方法
- 老闆說要改流程
- 客戶要我做做看

http://goo.gl/YiJrlI

Why you need R Markdown

http://goo.gl/rwrhpK

What is R Markdown

A convenient tool to generate reproducible document.

- Markdown
  - Remove HTML tag for higher readibility.
  - Inline HTML is avaliable.
- R markdown
  - Markdown + embedded R code chunks
  - Rmd -> md -> html(docx, pdf)

Why R Markdown

- 製作reproducible的報告、投影片
- 想寫數學式子好展現自己的專業 $e = mc^2$
- 只有一份source code，不需要額外複製圖片到報告中
- 需求更改時，可以動態改變報告內容
- 增加資料分析演算法的可讀性
- IDE? RStudio提供支援

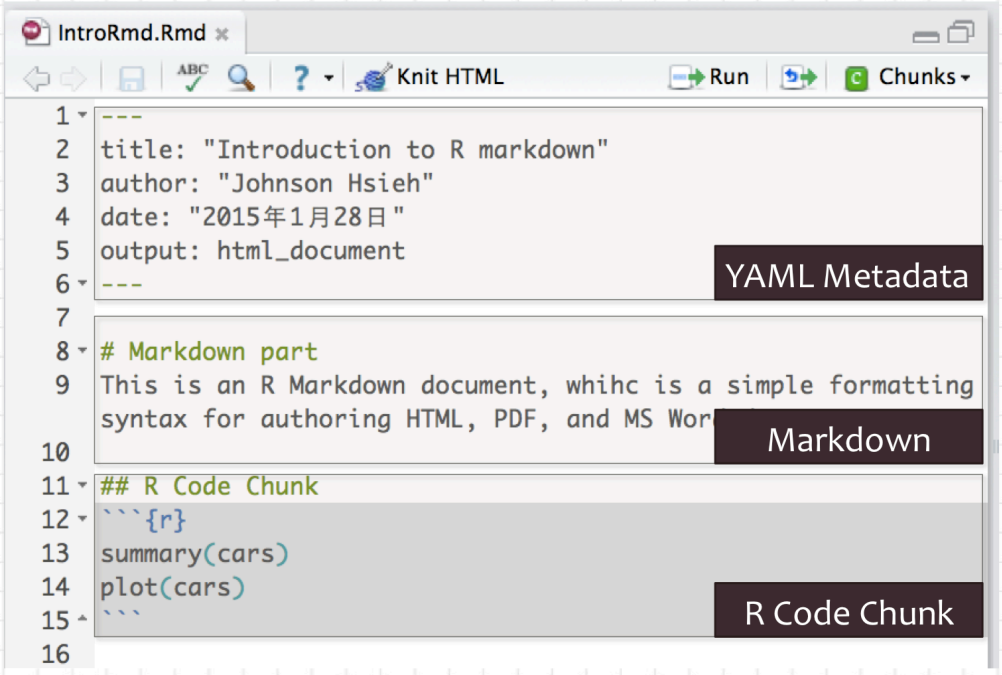
Installation

- 最新版的RStudio已經包含R Markdown功能
- 你也可以透過以下指令安裝R Markdown套件：

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

R Markdown 快速導覽

Overview



Markdown

example.Rmd

Knit HTML | Chunks

```
1  Header 1
2  ------------------------------------------------
3  This is an R Markdown document. Markdown is a
   simple formatting syntax for authoring web pages.
4
5  Use an asterisk mark, to provide emphasis such as
   *italics* and **bold**.
6
7  Create lists with a dash:
8  - Item 1
9  - Item 2
10 - Item 3
11
12 You can write `in-line` code with a back-tick.
13
14 ```
15 Code blocks display
16 with fixed-width font
17 ```
18
19 > Blockquotes are offset
20
```

RStudio: Preview HTML

Preview: ~/example.html

Header 1

This is an R Markdown document. Markdown is a simple formatting syntax for authoring web pages.

Use an asterisk mark, to provide emphasis such as *italics* and **bold**.

Create lists with a dash:

- Item 1
- Item 2
- Item 3

You can write `in-line` code with a back-tick.

```
Code blocks display
with fixed-width font
```

> Blockquotes are offset

R Code Chunks

chunks.Rmd

Knit HTML | Chunks

```
1  R Code Chunks
2  ==========================
3
4  With R Markdown, you can insert R code
   chunks including plots:
5
6  ```{r qplot, fig.width=4, fig.height=3,
   message=FALSE}
7  # quick summary and plot
8  library(ggplot2)
9  summary(cars)
10 qplot(speed, dist, data=cars) +
11       geom_smooth()
12 ```
13
```

RStudio: Preview HTML

Preview: ~/chunks.html | Save As | Publish

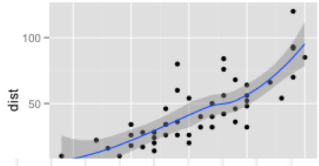
R Code Chunks

With R Markdown, you can insert R code chunks including plots:

```
# quick summary and plot
library(ggplot2)
summary(cars)
```

```
##      speed           dist
##  Min.   : 4.0   Min.   :  2
##  1st Qu.:12.0   1st Qu.: 26
##  Median :15.0   Median : 36
##  Mean   :15.4   Mean   : 43
##  3rd Qu.:19.0   3rd Qu.: 56
##  Max.   :25.0   Max.   :120
```

```
qplot(speed, dist, data = cars) + geom_smooth()
```



Inline R Code and Equations

- 利用 `r` 在markdown中插入R程式
- 插入 LaTeX 公式的方法：
  - 行內`$ equation $`
  - 段落 `$$ equation $$`

```
1
2  這是DSP推出的第`r 3+1`門課程
3
4
5  熵指標的公式為$-\sum{p_i \log{p_i}}$，表示系統的亂度
6
```


Rendering Output

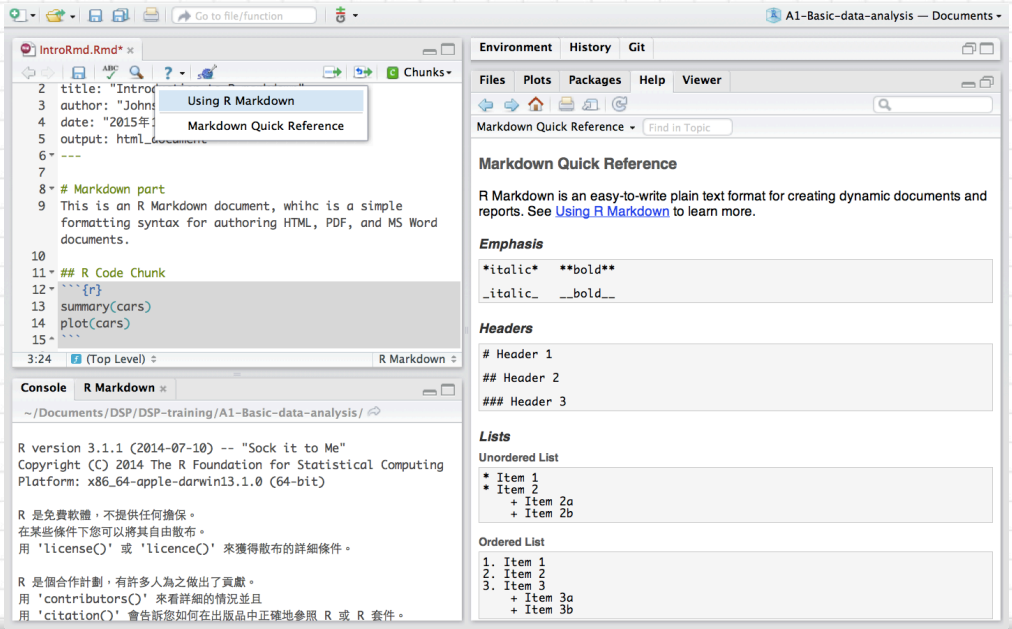
- RStudio: "Knit" command (Ctrl+Shift+K)
- Command line: rmarkdown::render function

```
rmarkdown::render("input.Rmd")
```

Markdown Basics

Markdown Quick Reference

在RStudio中，在UI界面中點選`help (?)`可以查閱Markdown語法



R Code Chunks

Overview

R code will be evaluated and printed

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

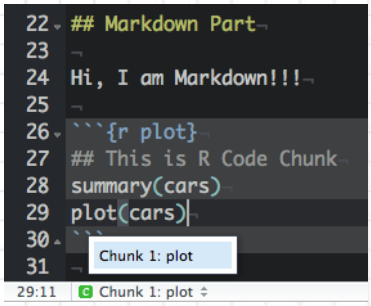
```
summary(cars$dist)
```

```
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
      2      26      36      43      56     120
```

Named R code chunk.

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

- Easy Navigation in RStudio



Basic Chunk Options

- `echo`(TRUE): whether to include R source code in the output file
- `eval`(TRUE): whether to evaluate the code chunk
- `message`(TRUE): whether to preserve messages emitted by message()
- `include`(TRUE): if include=FALSE, nothing will be written into the output document, but the code is still evaluated and plot files are generated
- `warning`(TRUE): whether to preserve warnings in the output
- `comment`("##"): set to comment notation

Set global chunk options:

```
knitr::opts_chunk$set()
```

Exercise

利用R Markdown 製作《一周天氣預報》書面報告。

- 原始出處
- 參考範本
- 範例資料

Exercise Q1

利用R Markdown 製作《一周天氣預報》書面報告。

- 計算01/28日當日的最高溫與最低溫度

```
# Hint：
# 1. 下載weather-utf8.csv到自己的電腦上
# 2. 在R chunk中，利用read.csv()讀取檔案進行分析
# 3. 找出01/28當日最高溫 max()
# 4. 找出01/28當日最低溫 min()
# 5. use inline R chunk `r max(...)`
```

Table Output

Set `results='asis'` to write raw results from R into the output document

- `knitr::kable`

````
```{r, results='asis'}
knitr::kable(women)
```
````


Exercise Q2

利用R Markdown 製作《一周天氣預報》書面報告。

- 製作未來七天天氣預報表

```
# Hint：
# 你可能需要dplyr套件
# 可以先用filter把白天、晚上分開處理
# 利用 paste(低溫,高溫,sep="-") 來製作溫度區間, i.e. 16-17
# 利用colnames, rownames來對整理好的資料表的行與列命名
```

Exercise Q3

利用R Markdown 製作《一周天氣預報》書面報告。

- 製作未來七天天氣預報圖

```
# Hint：
# 你可能需要ggplot2套件
# Mac顯示中文需設置字型
# http://equation85.github.io/blog/graph-font-of-r-in-mac-os-x/
# par(family='STHeiti')
```


Exercise

利用R Markdown 製作《一周天氣預報》書面報告。

- 原始出處
- 參考範本
- 範例資料
- 參考解答

Appendiex

About Document Content

You can add R Markdown and HTML in the YAML content.

```
---
title: "Introduction to R Markdown"
author: "Wush Wu, Johnson Hsieh"
date: "2015-02-01"
output: html_document
---
```

Generate Markdown and HTML

```
```{r results='asis', echo=FALSE}
library(whisker)
temp = '<span class=""></span>'
numbers = c("", "2", "3")
colors = c("red", "blue", "green", "yellow", "gray")
for (color in colors){
    cat("- ")
    for (number in numbers){
        out = whisker.render(temp)
        cat(out)
    }
    cat("\n")
}
```
```

Some Useful HTML

- iframe: displaying a web page within a web page

```
<iframe src="http://twconf.data-sci.org/" height=600 width=800></iframe>
```

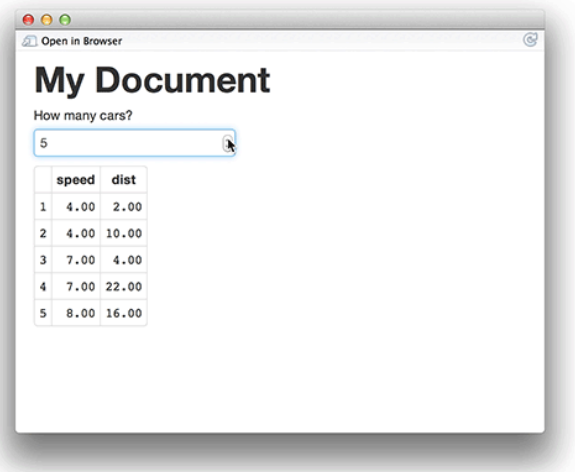
- img: inserting images into an HTML document.

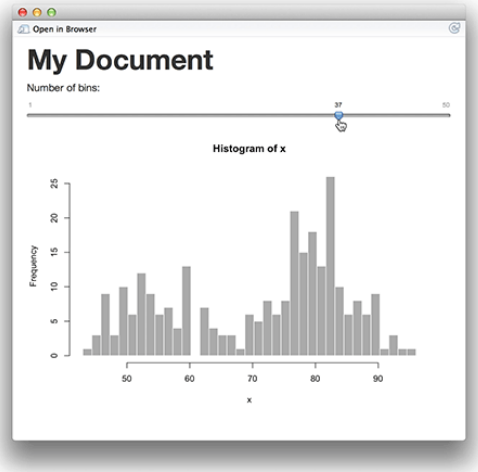
  Much easier for adjusting width and height.

```
<img src="assets/img/dsp-logo.png" alt="logo">
```

Interactive Documents

It's possible to embed a Shiny application within a document.





References

- R Markdown Cheat Sheet
- R Markdown
- knitr
- RStudio Documentation
- Reproducible Research
- Shiny Articles