

# 給初學夥伴 (M) R學習地圖

Dana Huang Lucy Chen



## 分析,就像學做菜

#### 材料 & 工具 R-basic



#### 基本套餐

完整的專案







花式套餐

進階分析 & 呈現





### 做菜可能發生什麼事?

#### 買錯材料、切到手 怎麼殺魚&切肉?!



其中一個配菜出問題 、搭配不起來



被爐子燙到、 味道跟想像不一樣



浪費好食材、 畫龍點睛 or 悲劇?





## 2017 R-basic Schedule

主題	2/6	3/6	3/27	4/24	5/29	6/26	7/31	8月	9月	10月	11月	12月
1 R軟體小故事										0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0		
2 安裝與環境介紹												
3基本運算&常用指令	0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	•								0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0		
4 運用Library/Package												
5 一般數據結構與格式	## ## ## ## ## ## ## ## ## ## ## ## ##				•	•				0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0		
6 讀取各類資料							•					
7 敘述統計分析	**************************************								•	9 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0		
8 資料預處理												
9 抽樣												•



## R-basic 主題

	主題	內容
1	R軟體小故事	R的由來 R與其他統計軟體的比較
2	安裝與環境介紹	Rstudio安裝程序 其他安裝事項 (不同分析目的)
3	基本運算&常用指令	基本運算 執行、暫停、關閉 好用的查詢:?, help, example 呼叫執行其他檔案的 source
4	Library & Package	安裝及呼叫 package 好用的套件
5	一般數據結構 8資料格式	向量、矩陣、陣列 資料格式:因子、資料格式、清單格式



## R-basic 主題

主題	<b>內容</b>
6 讀取各類資料	csv · excel · txt
7 敘述統計分析	summary 基本統計圖表 資料衡量(Measurement) 相關性分析
8 資料預處理	如何補齊資料 刪除部分資料
9 抽樣	抽樣方法 SQL指令 將資料分訓練和測試二部分、cross validation



## R-basic 研習教學指引





### 每月15分鐘怎麼報告

#### A. 界定範疇:讓聽眾得到什麼?

(EX)

- ●基本觀念與架構
- ●流程與操作
- ●實際例子
- ●工具
- ●參考資料











### 每月15分鐘怎麼報告

B. 決定呈現方式: 什麼樣的接觸體驗?







#### 每月15分鐘怎麼報告

- 分享操作過程中的發現
- 優先選擇流通性高的實用套件、簡單範例
- 時間!濃縮為15分鐘精華
- 活動前一天將material或slide上傳給所有觀眾 (也許前來的當天路上都還可以馬上看喔XD)



### 你可以參考...

• 之前R-Basic教材

https://github.com/TaiwanRUserGroup/DSC2015Tutorialhttps://github.com/dboyliao/DSC2015\_RBasic

- R語言翻轉教室 <a href="http://datascienceandr.org/">http://datascienceandr.org/</a>
- Code School http://tryr.codeschool.com/
- Data Camp
   https://www.datacamp.com/
- Quick R
   http://www.statmethods.net/



### 你可以用下面資料來練習...

- 著名的初/中/高階練習資料 <a href="https://www.analyticsvidhya.com/blog/2016/10/17-ultimate-data-science-projects-to-boost-your-knowledge-and-skills/">https://www.analyticsvidhya.com/blog/2016/10/17-ultimate-data-science-projects-to-boost-your-knowledge-and-skills/</a>
- 加州大學機器學習資料庫 http://archive.ics.uci.edu/ml/
- 美國開放資料平臺 http://www.data.gov
- Data.taipei <u>http://data.taipei/</u>
- Stanford dataset <u>http://snap.stanford.edu/data/</u>

#### 如何快速找到想用的套件

http://www.meetup.com/Taiwan-R/messages/boards/thread/27399932



### 有問題的時候?

#### 1. 嘗試自己找解 (e.g. dplyr select)

- Google
- Stackoverflow
- PTT R\_Language版,
- Github

#### 2. 勇敢在版面上發問

- 先說明自己找解遇到的困難(做了哪些嘗試)
- 將問題從stackoverflow轉貼到社群頁面, etc.

#### 3. 如何提昇回答率

- 多問幾個人
- 給解答問題人動機,例如以給talk交換
- 將所學得的內容書寫公開

#### 4. 試著提供協助

● 先了解問問題人的程度、遇到的狀況、思索解法





## Let's Demo R-basic





## Thank you