R Basic
Tableau with R

Dana(昱璇)

2018.6.25

Intergration Between R and Tableau
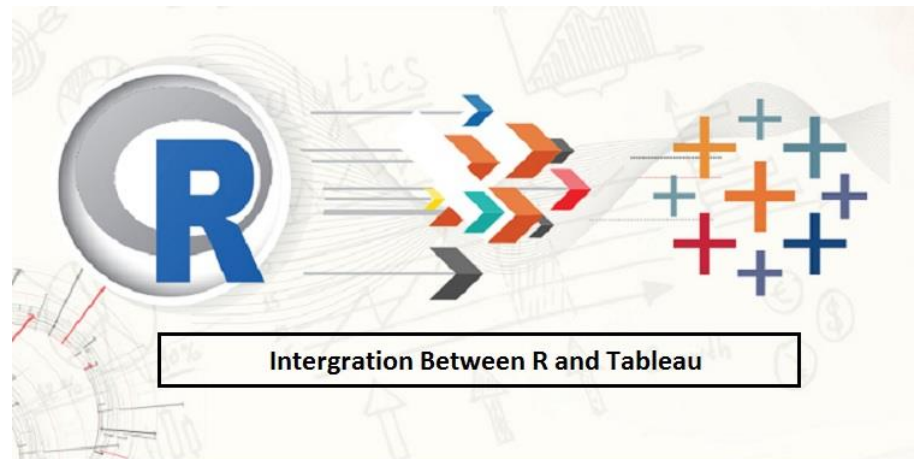
為何有Tableau還需要R

Tableau

✓ Super-easy to use
✓ Beautiful visualizations out-of-the-box
✓ Very powerful
✓ Very fast
✓ Lots of data connectors
✓ Short learning curve
✓ Big community
✗ Data manipulation is a pain
✗ Limited to visualizations
✗ Sharing workbooks (for confidential data) is a challenge (expensive)

R

✓ Huge community
✓ Lots of libraries
✓ Can do pretty much everything
  ✓ ?web extraction
  ✓ machine learning
  ✓ statistics (duh)
  ✓ data visualization
  ✓ interactive reporting
  ✓ web-apps
✗ Difficult to learn
✗ Need to know/learn programming
✗ Errors are hard to pin-point

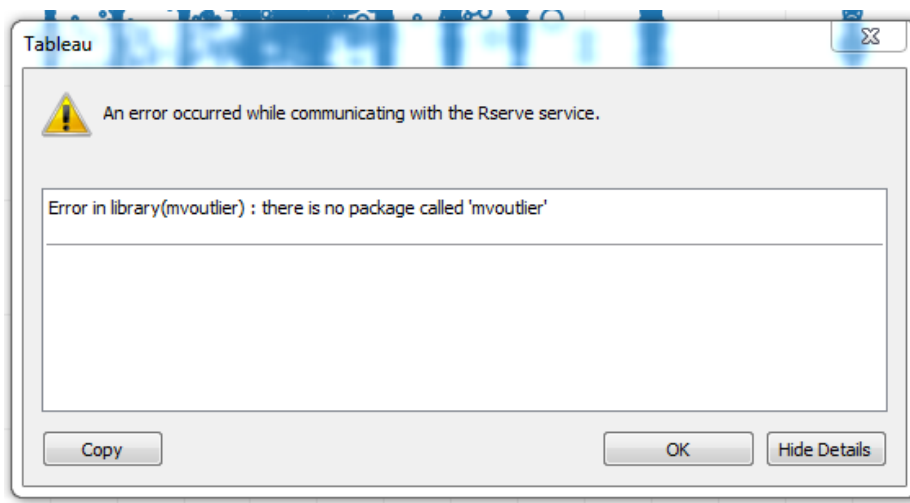Dynamically recompute the results without manually having to recreate the data set for R

Reference: http://nandeshwar.info/data-science-2/tableau-vs-r/

如何連線

執行R程式碼，並到tableau啟動連線

記得要用的套件要先在R軟體安裝

以排除outlier案例來說，未安裝mvoutlier套件時，執行tableau outlier變數會出現該錯誤訊息

於Rstudio執行語法及安裝完成畫面

```
all.packages("mvoutlier")
```

```
package 'sROC' successfully unpacked and MD5 sums checked
package 'VIM' successfully unpacked and MD5 sums checked
package 'zCompositions' successfully unpacked and MD5 sums checked
package 'sgeostat' successfully unpacked and MD5 sums checked
package 'robCompositions' successfully unpacked and MD5 sums checked
package 'mvoutlier' successfully unpacked and MD5 sums checked

The downloaded binary packages are in
        C:\Users\danahuang1\AppData\Local\Temp\RtmpOq0LJ0\downloaded_p
ackages
>
```

基本公式 (Tableau 8.1後版本)

依分析輸出的格式分為:

- SCRIPT_REAL –Decimal Number
- SCRIPT_INT – Whole Number
- SCRIPT_STR - Text
- SCRIPT_BOOL - True/False
- 以SCRIPT_INT為例

Clusters

```
Results are computed along Table (across).
SCRIPT_INT('result <-
kmeans(x = data.frame(.arg1,.arg2,.arg3,.arg4), 3);          'insert R script here'
result$cluster',
SUM([Sepal length]),SUM([Sepal width]),SUM([Petal length]),
SUM([Petal width]))|                                          [someTableauAggregateMeasure]
```

☺ 要會寫R code才能完整使用喔☺

Reference: https://community.tableau.com/docs/DOC-5313

基本語法彙整

- 各種靜態的表示法

test_real

Results are computed along Table (across).
```
SCRIPT_REAL('0.1',max(1))
```

test_int

Results are computed along Table (across).
```
SCRIPT_INT('0.1',max(1))
```

test_string

Results are computed along Table (across).
```
SCRIPT_STR('x<-"Hello world"',MIN(0))
```

test_boolen

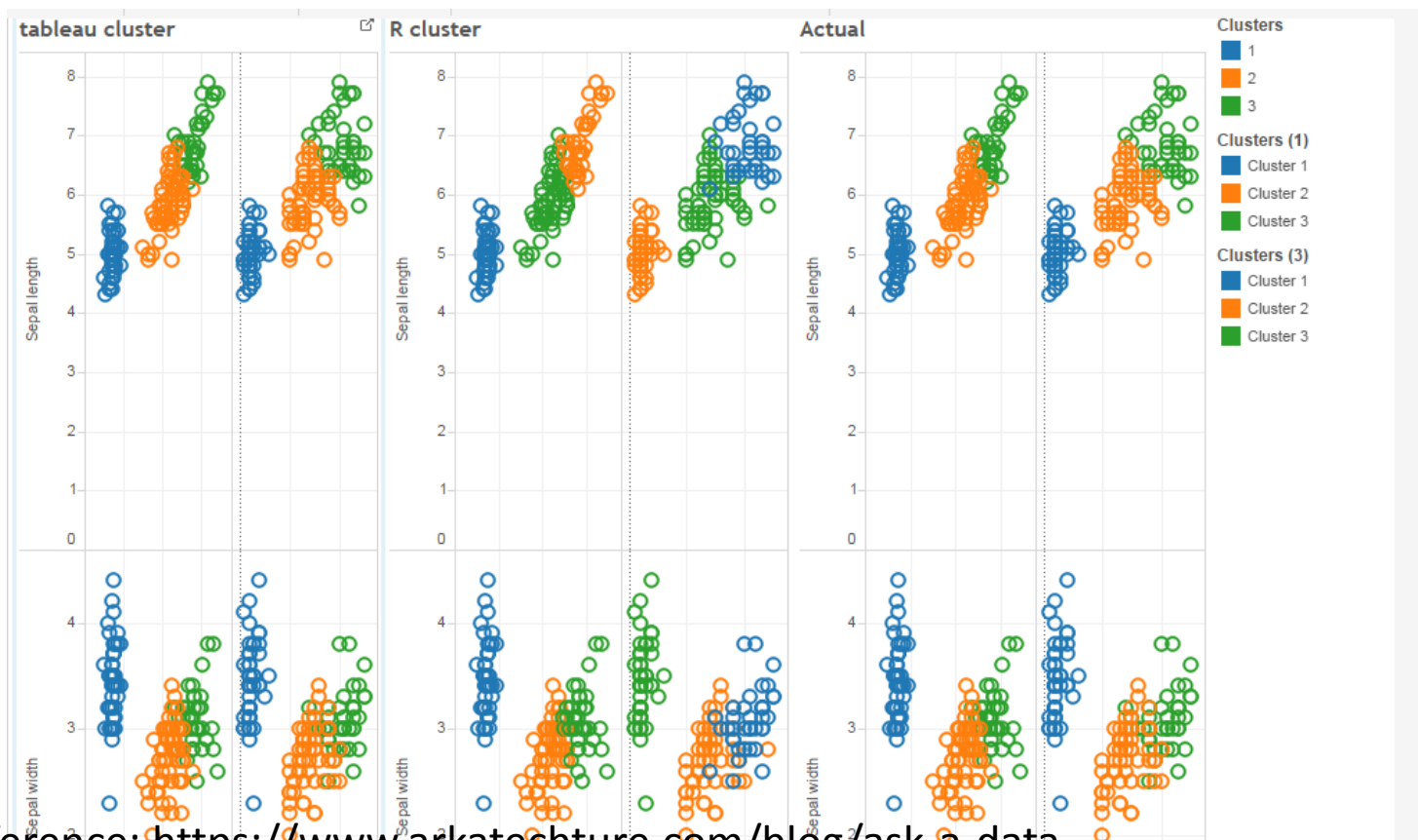Results are computed along Table (across).
```
SCRIPT_BOOL('TRUE',max(1))
```

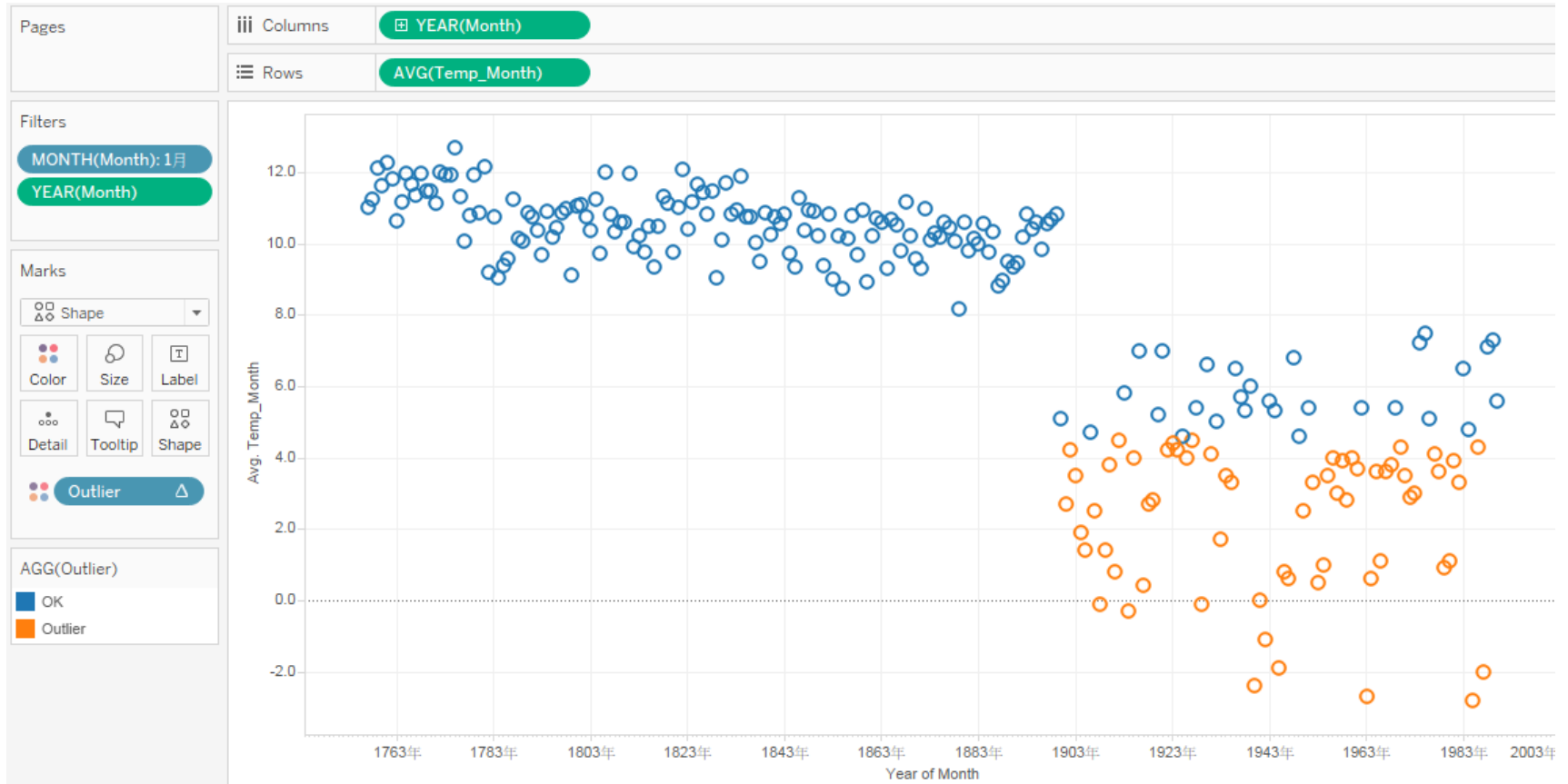# Let's Demo

案例分享-R與tableau分群比較

- 使用kmeans做分群 (使用iris Data)

Reference: https://www.arkatechture.com/blog/ask-a-data-scientist-how-does-clustering-in-r-vs.-tableau-work

案例分享-排除outlier

- 使用R套件區別outlier (使用Paris temperature Data)

Reference: https://community.tableau.com/thread/138417

案例分享-排除outlier

- 若無outlier，會出現此錯誤訊息

Tableau

⚠ An error occurred while communicating with the RServe service.

Error in svd(xs) : infinite or missing values in 'x'

Copy   OK   Hide Details