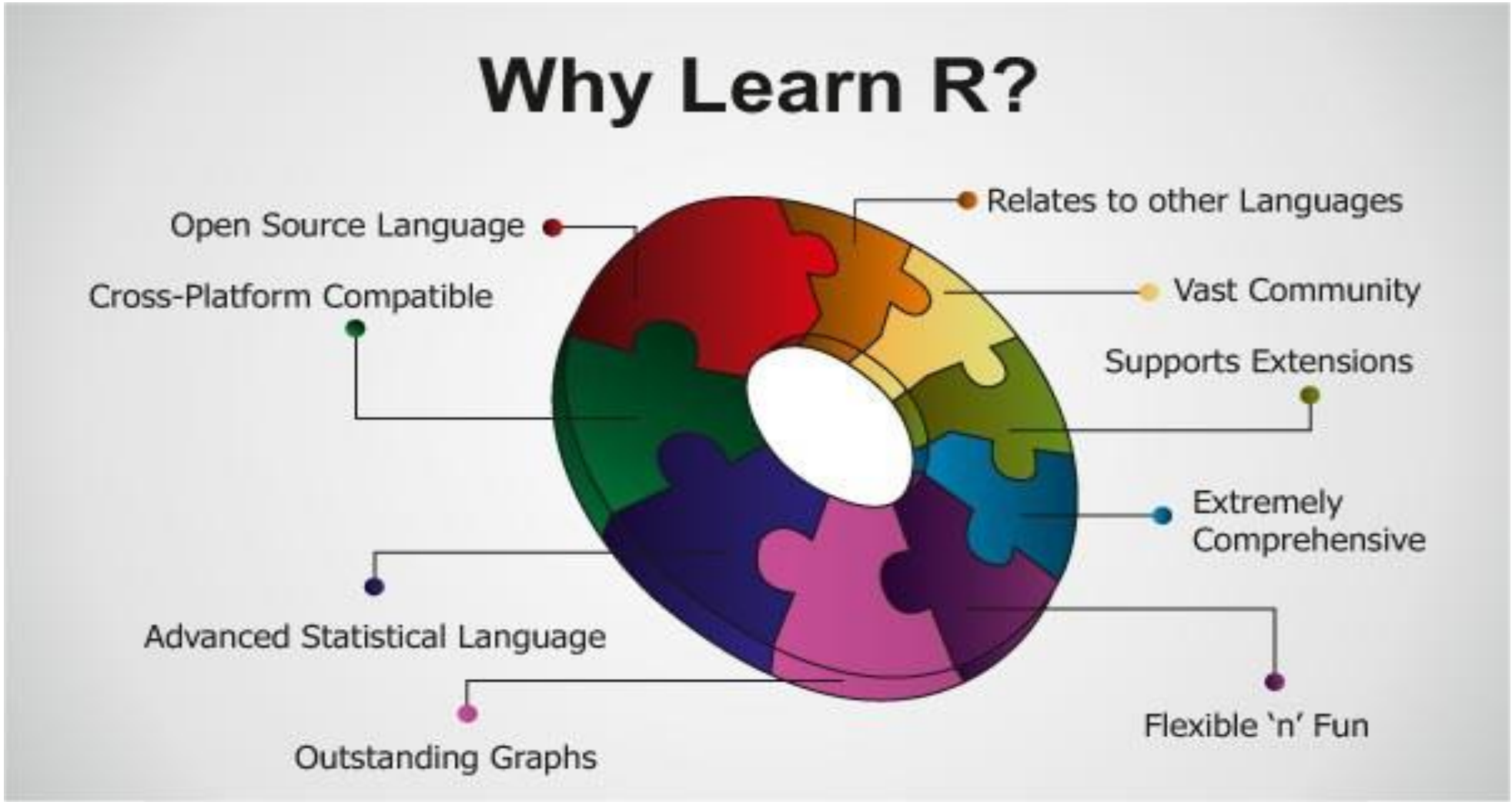
R Studio, R & Tableau

A brief introduction

Why Learn R?


Open Source Language
Cross-Platform Compatible
Advanced Statistical Language
Outstanding Graphs
Relates to other Languages
Vast Community
Supports Extensions
Extremely Comprehensive
Flexible 'n' Fun

What is R?

The facts:

- R is a language and environment for statistical computing and graphics
- Freely available and maintained by volunteers
- R is extensible; can be expanded by installing “packages”

How to get it:

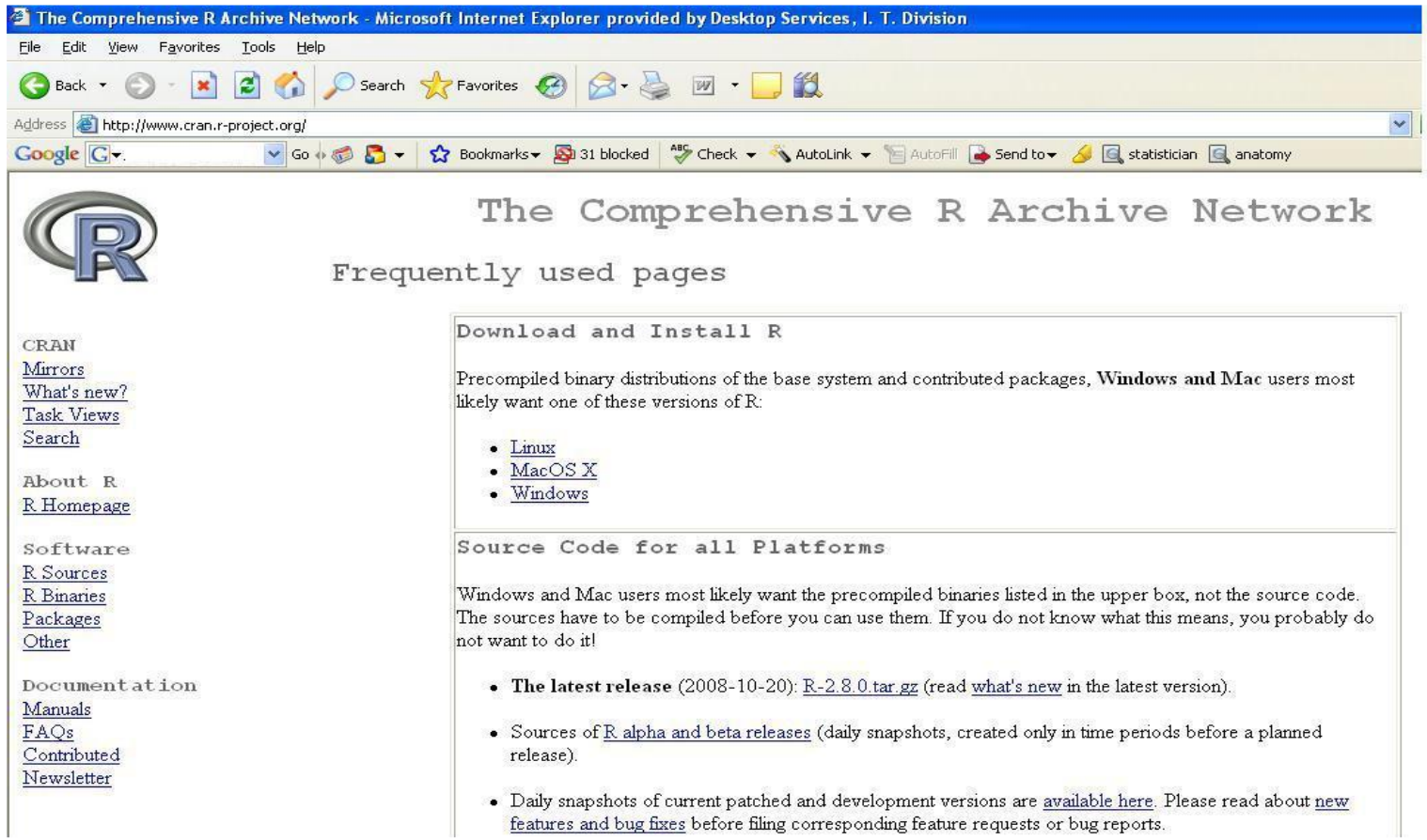
- http://www.r-project.org/ (or Google “Download R”)
- Available for Windows, Mac, Linux
- Free to install, no catches

Also highly recommended:

- R Studio: a free IDE for R
- http://www.rstudio.com/
- If you install R and R Studio, then you only need to run R Studio

Installing R

- R must be installed on your system! If it is not, go to www.cran.r-project.org



- Click on
- Windows > base > R-version-win32.exe > Run
- and follow the instructions to install the programme

RStudio screen

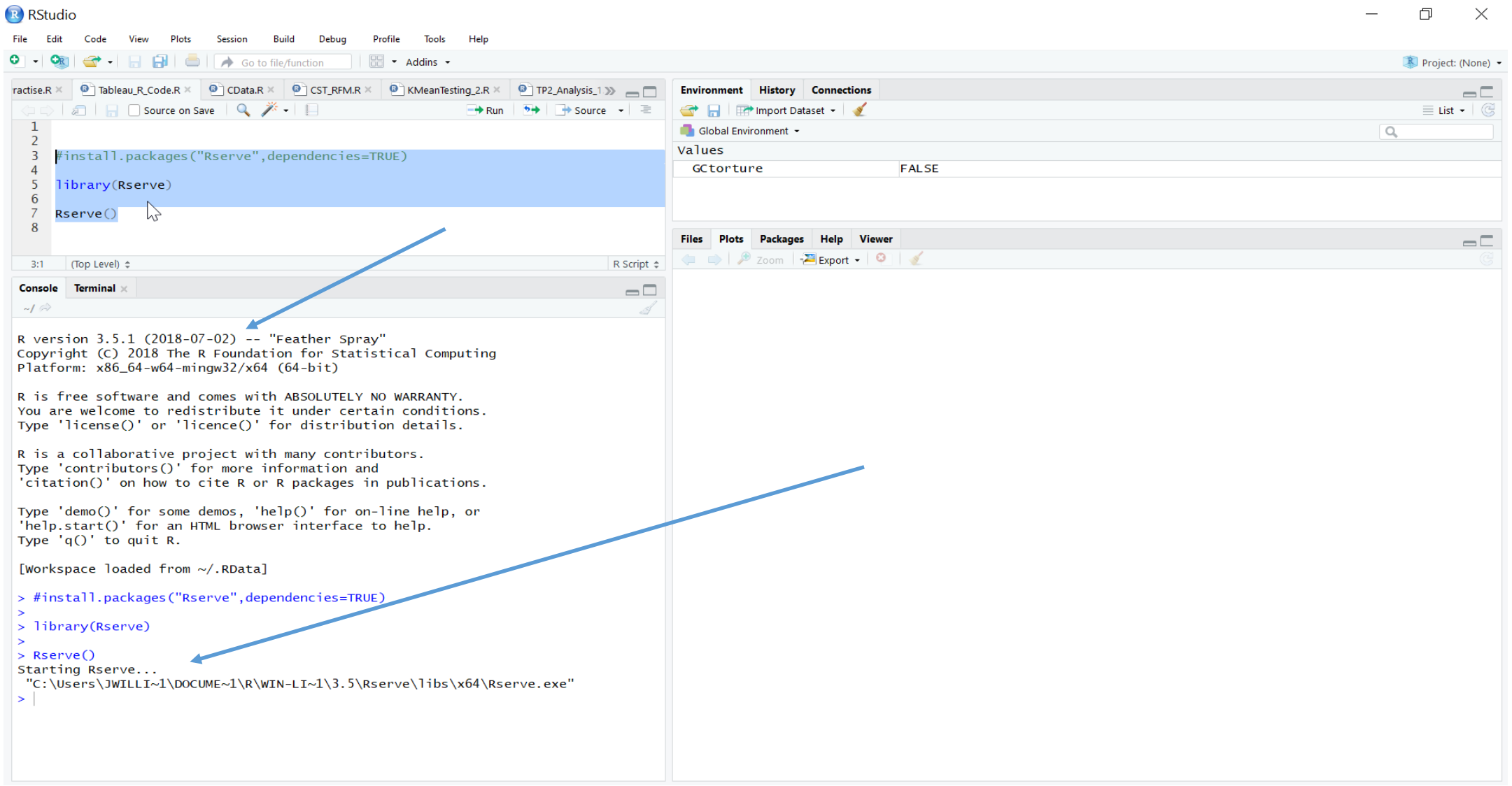
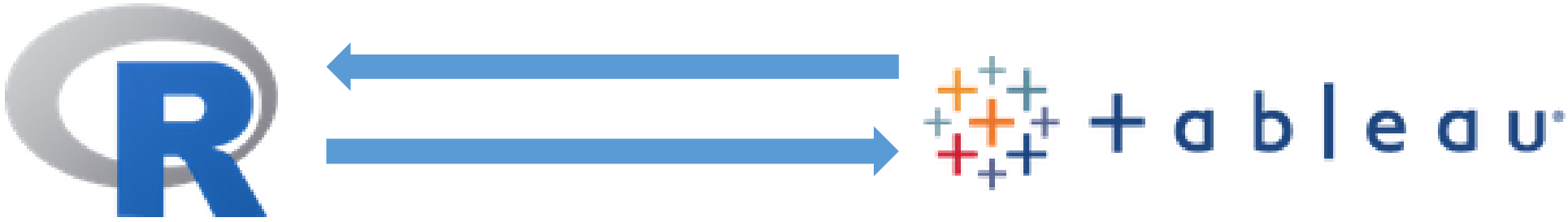


Tableau and R

What are the benefits of using Tableau and R?

Access R packages or functions for quantitative analysis



Take advantages of Tableau's data visualization capacities

Tableau and R Integration

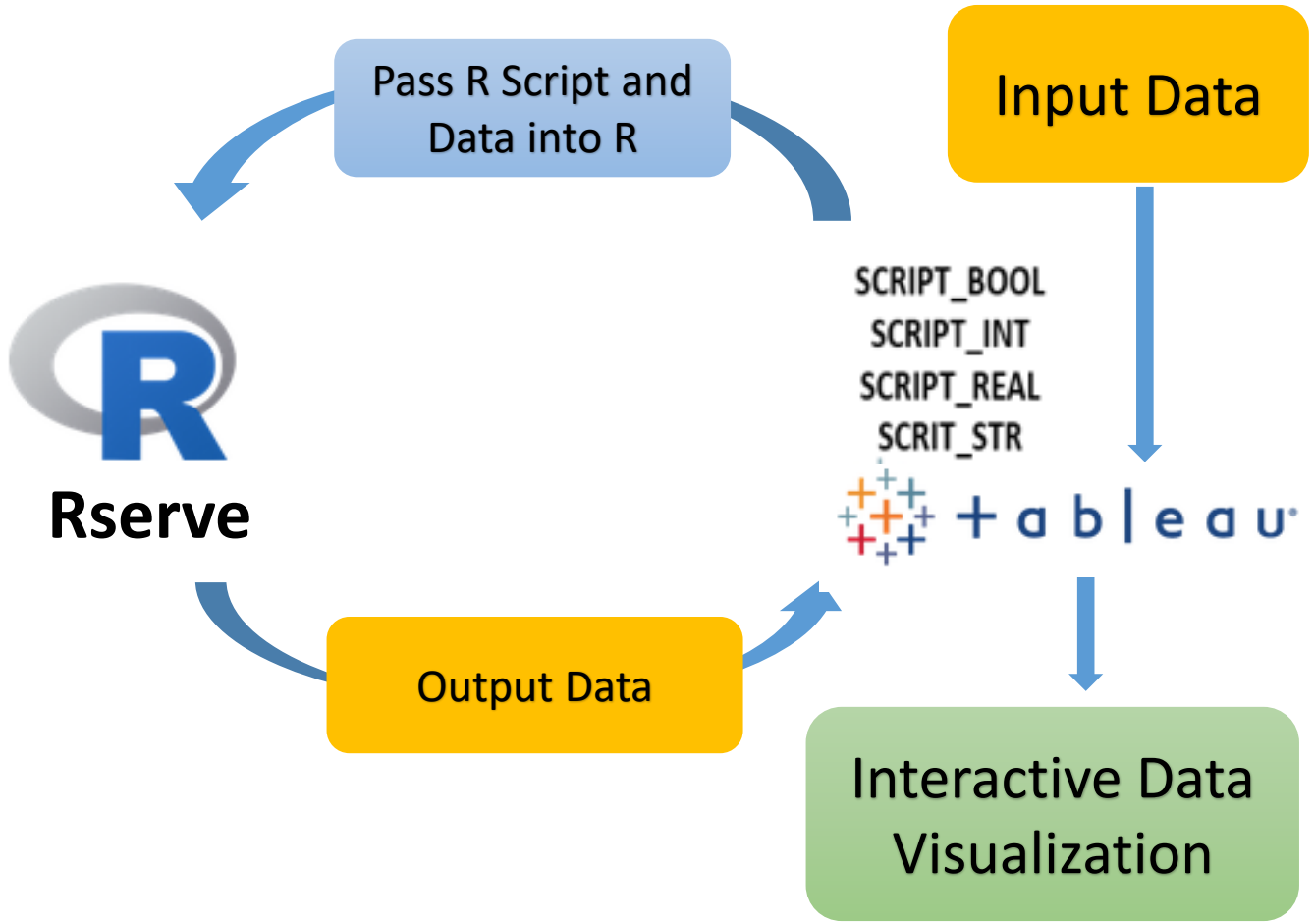
How is Tableau integrating with R?

- Install Rserve Package

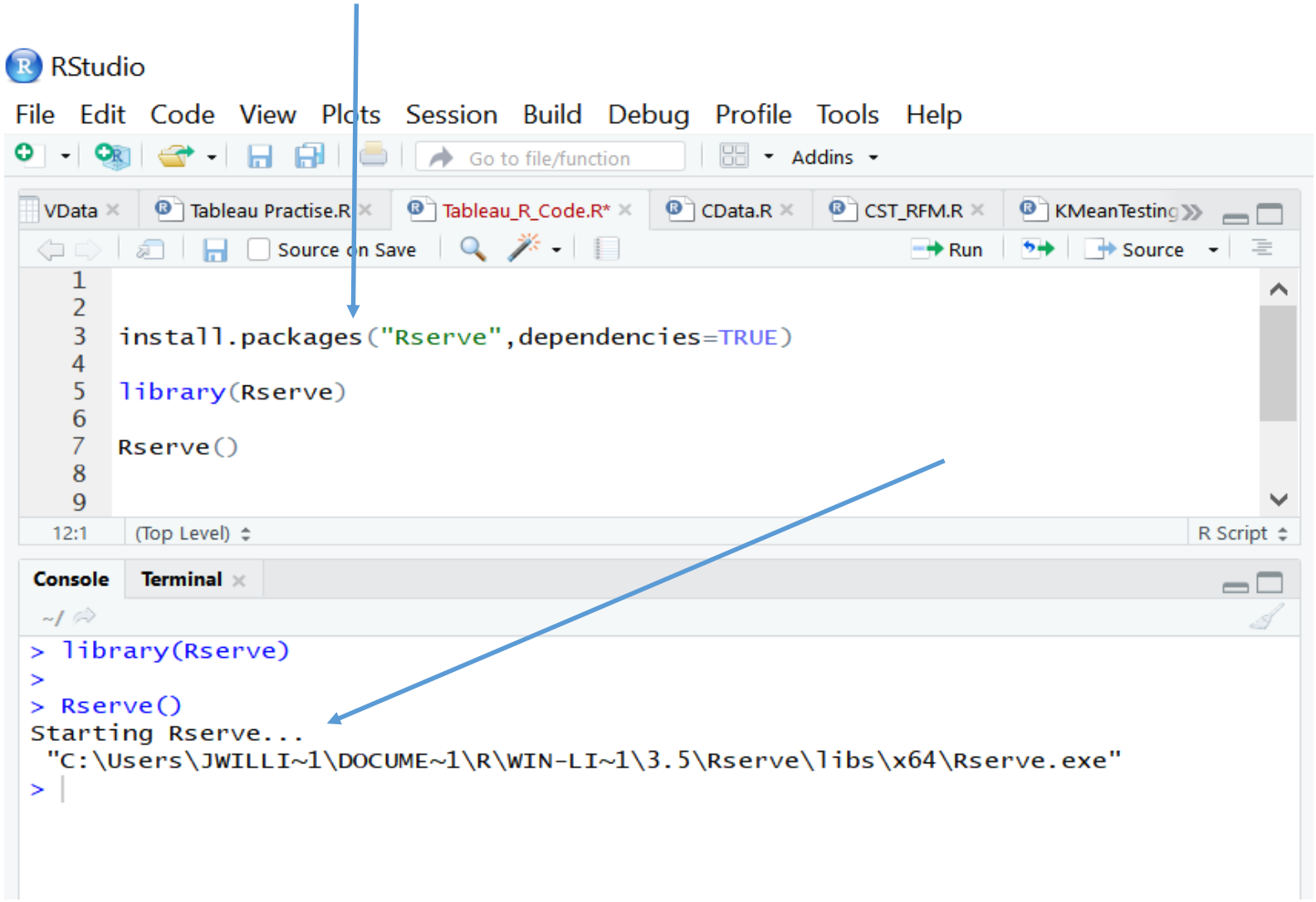
```
install.packages("Rserve")
library(Rserve)
Rserve()
```

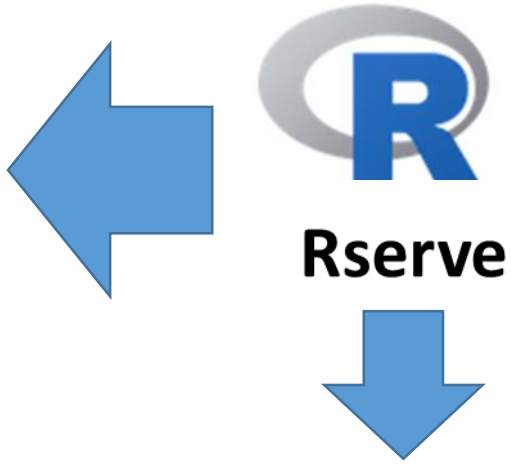
- Create Calculation Field in Tableau

- Use one of the four functions:
  SCRIPT_BOOL: Return a **Boolean**
  SCRIPT_INT: Return an **Integer**
  SCRIPT_REAL: Return a **Real**
  SCRIPT_STR: Return a **String**



Rserve Package





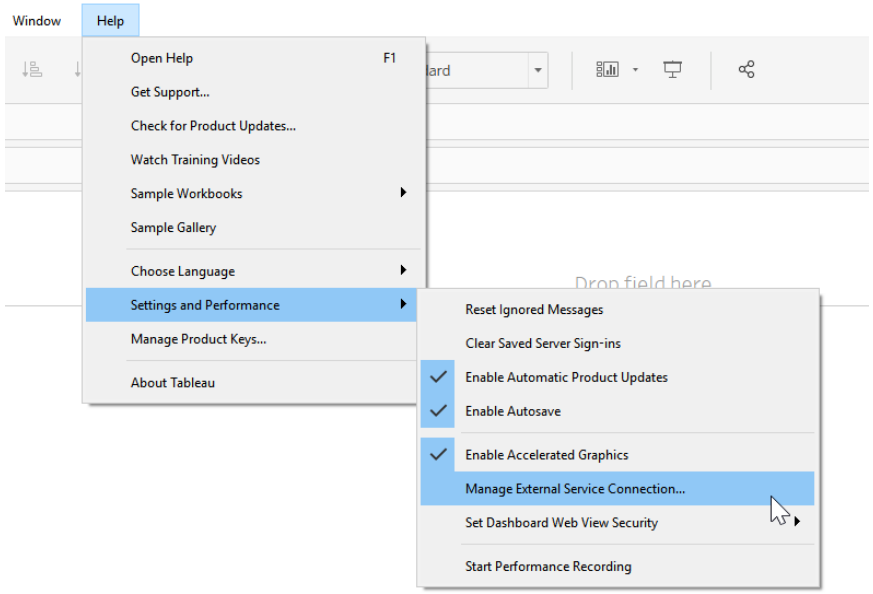
Very Important!

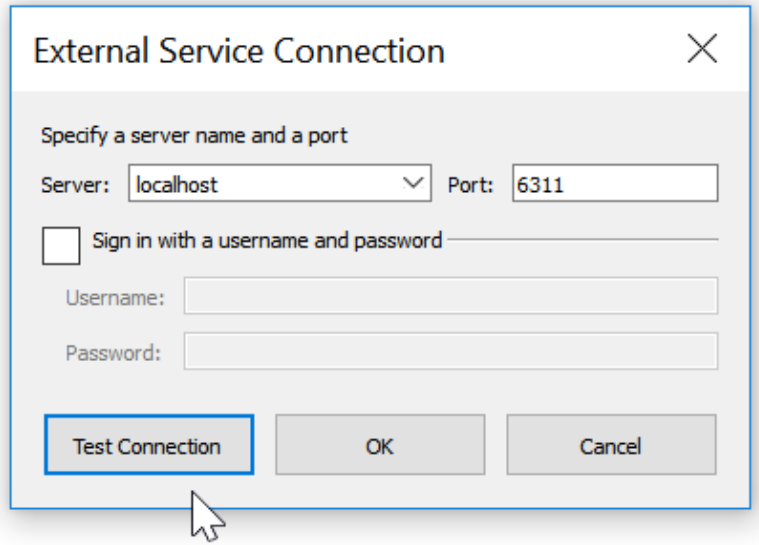
Remember to keep Rstudio open to use Rserve in Tableau.

Tableau (1)

Go to Help -> Setting and Performance -> Manage External Server Connection and set it.

Make sure you choose "localhost" and test it. That's all.





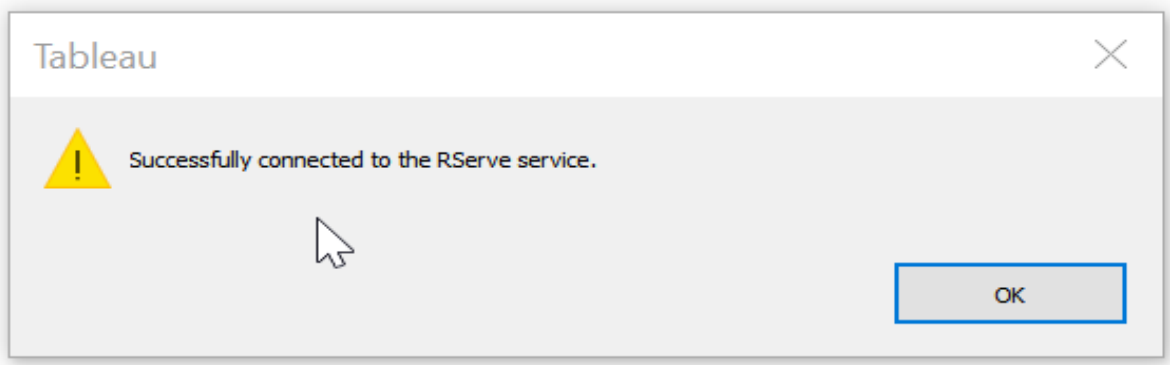
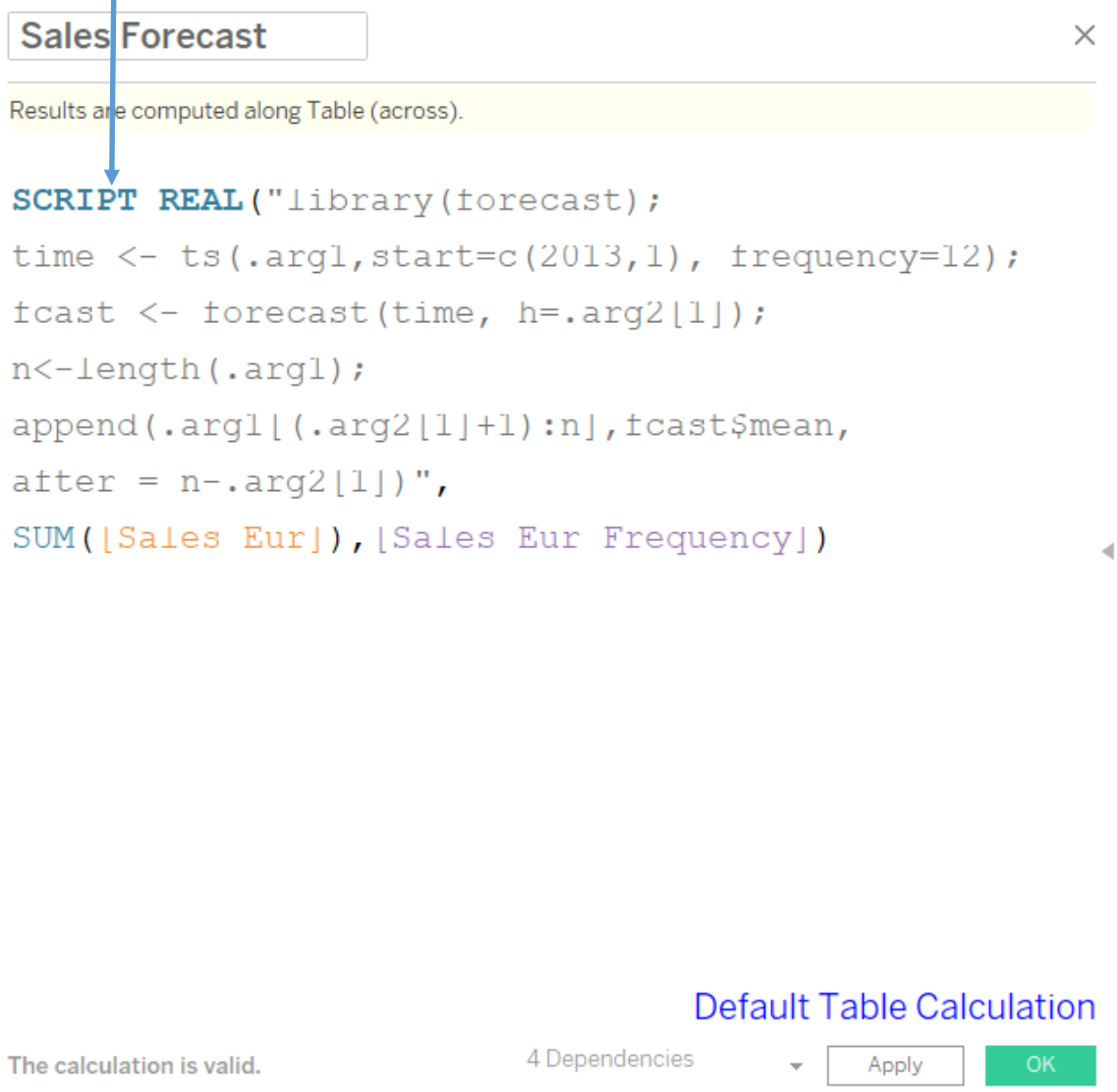


Tableau (2)


Sales Forecast
Results are computed along Table (across).
SCRIPT_REAL("library(forecast);
time <- ts(.arg1,start=c(2013,1), frequency=12);
fcast <- forecast(time, h=.arg2[1]);
n<-length(.arg1);
append(.arg1[(.arg2[1]+1):n],fcast$mean,
after = n-.arg2[1])",
SUM([Sales Eur]),[Sales Eur Frequency])
Default Table Calculation
The calculation is valid.
4 Dependencies
Apply
OK

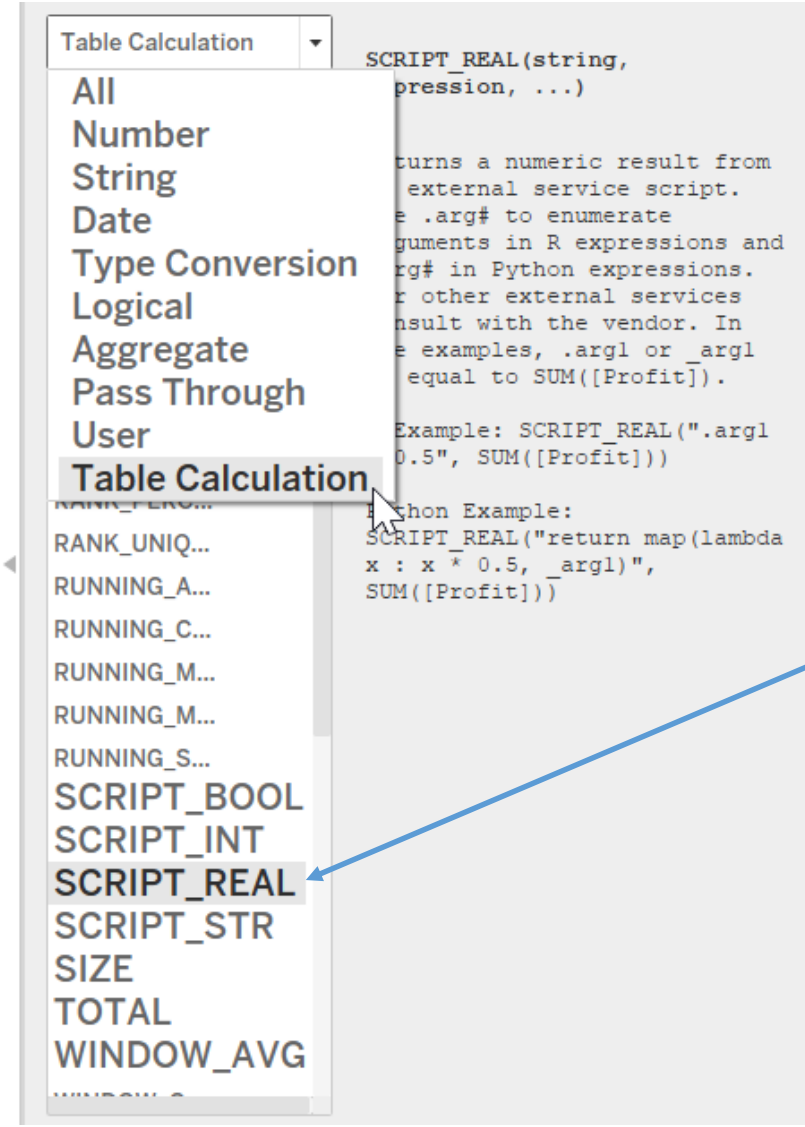


Table Calculation
All
Number
String
Date
Type Conversion
Logical
Aggregate
Pass Through
User
Table Calculation
RANK_UNIQ...
RUNNING_A...
RUNNING_C...
RUNNING_M...
RUNNING_M...
RUNNING_S...
SCRIPT_BOOL
SCRIPT_INT
SCRIPT_REAL
SCRIPT_STR
SIZE
TOTAL
WINDOW_AVG
SCRIPT_REAL(string,
pression, ...)
turns a numeric result from
external service script.
e .arg# to enumerate
guments in R expressions and
rg# in Python expressions.
r other external services
nsult with the vendor. In
e examples, .arg1 or _arg1
equal to SUM([Profit]).
Example: SCRIPT_REAL(".arg1
0.5", SUM([Profit]))
hon Example:
SCRIPT_REAL("return map(lambda
x : x * 0.5, _arg1)",
SUM([Profit]))

Tableau (3)

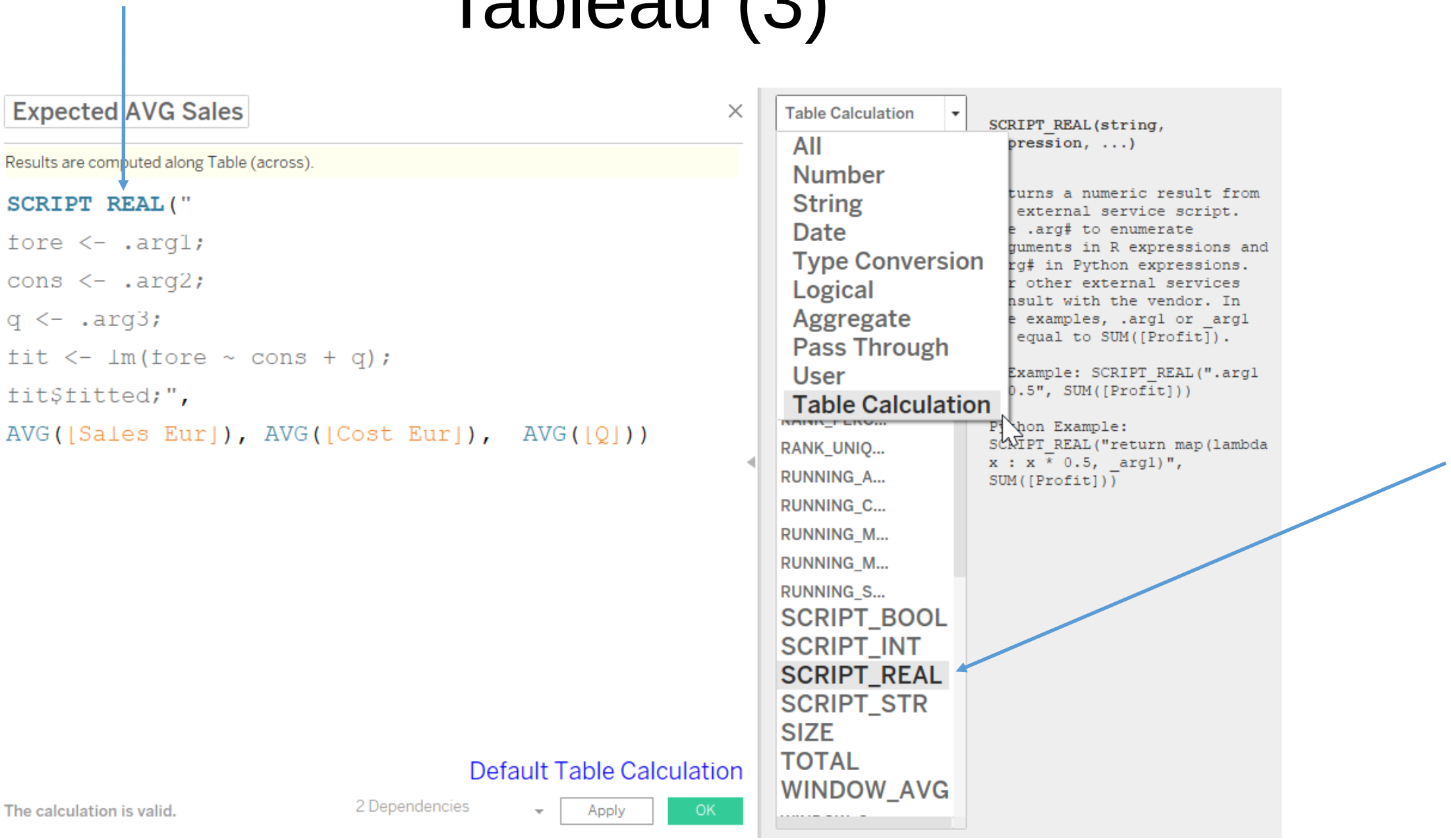

Expected AVG Sales
Results are computed along Table (across).
SCRIPT_REAL("
tore <- .arg1;
cons <- .arg2;
q <- .arg3;
tit <- lm(tore ~ cons + q);
tit$titted;",
AVG([Sales Eur]), AVG([Cost Eur]), AVG([Q]))
Default Table Calculation
The calculation is valid.
2 Dependencies
Apply
OK
Table Calculation
All
Number
String
Date
Type Conversion
Logical
Aggregate
Pass Through
User
Table Calculation
RANK_UNIQ...
RUNNING_A...
RUNNING_C...
RUNNING_M...
RUNNING_M...
RUNNING_S...
SCRIPT_BOOL
SCRIPT_INT
SCRIPT_REAL
SCRIPT_STR
SIZE
TOTAL
WINDOW_AVG
SCRIPT_REAL(string,
pression, ...)
turns a numeric result from
external service script.
.arg# to enumerate
guments in R expressions and
rg# in Python expressions.
other external services
nsult with the vendor. In
examples, .arg1 or _arg1
equal to SUM([Profit]).
Example: SCRIPT_REAL(".arg1
0.5", SUM([Profit]))
hon Example:
SCRIPT_REAL("return map(lambda
x : x * 0.5, _arg1)",
SUM([Profit]))

Tableau (4)

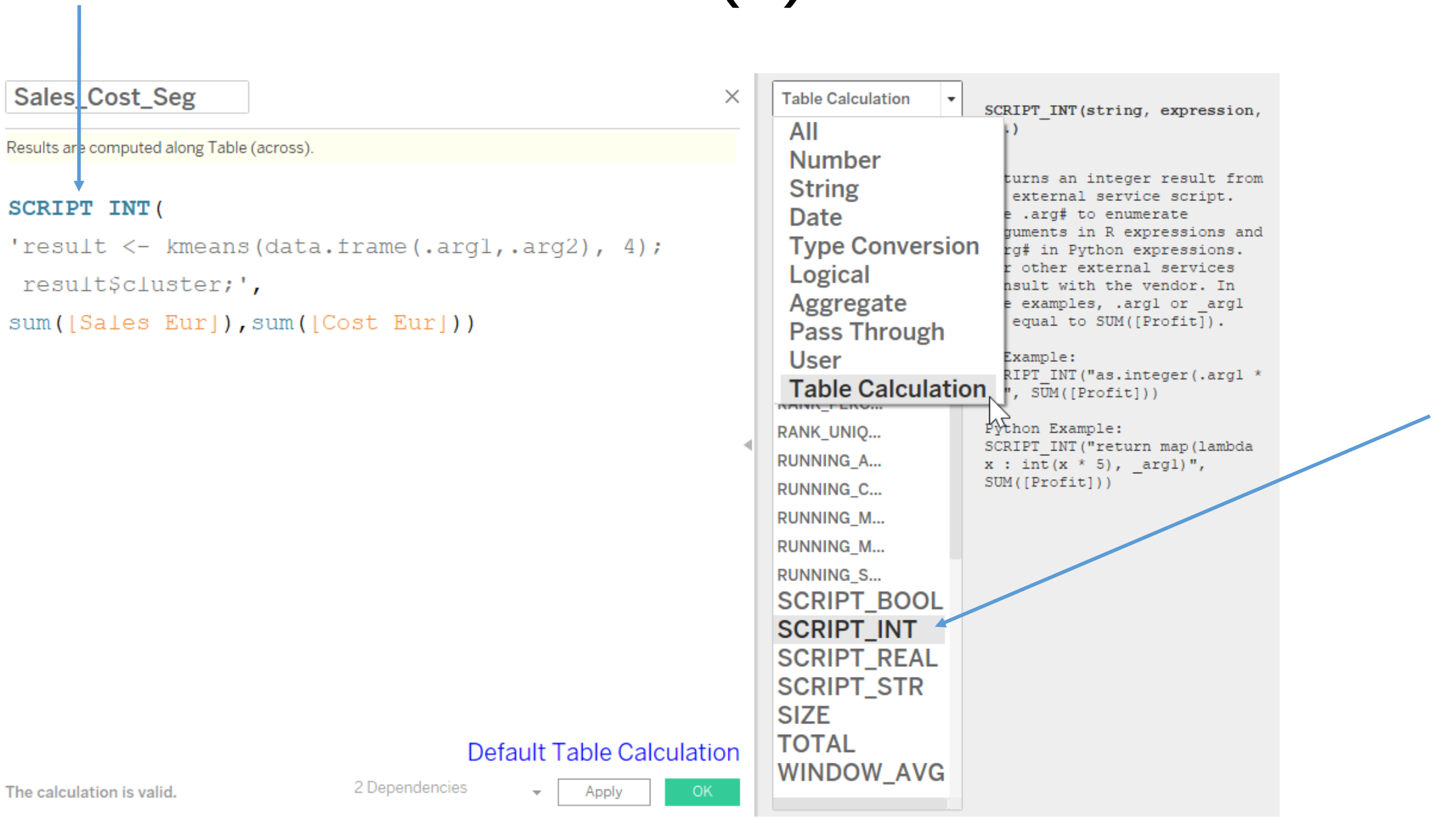

Sales_Cost_Seg
Results are computed along Table (across).
SCRIPT INT(
'result <- kmeans(data.frame(.arg1,.arg2), 4);
result$cluster;',
sum([Sales Eur]),sum([Cost Eur]))
Default Table Calculation
The calculation is valid.
2 Dependencies
Apply
OK
Table Calculation
All
Number
String
Date
Type Conversion
Logical
Aggregate
Pass Through
User
Table Calculation
RANK_UNIQ...
RUNNING_A...
RUNNING_C...
RUNNING_M...
RUNNING_M...
RUNNING_S...
SCRIPT_BOOL
SCRIPT_INT
SCRIPT_REAL
SCRIPT_STR
SIZE
TOTAL
WINDOW_AVG
SCRIPT_INT(string, expression,
turns an integer result from
external service script.
.arg# to enumerate
guments in R expressions and
rg# in Python expressions.
other external services
nsult with the vendor. In
examples, .argl or _argl
equal to SUM([Profit]).
Example:
RIPT_INT("as.integer(.argl *
", SUM([Profit]))
Python Example:
SCRIPT_INT("return map(lambda
x : int(x * 5), _argl)",
SUM([Profit]))

Tableau (5)

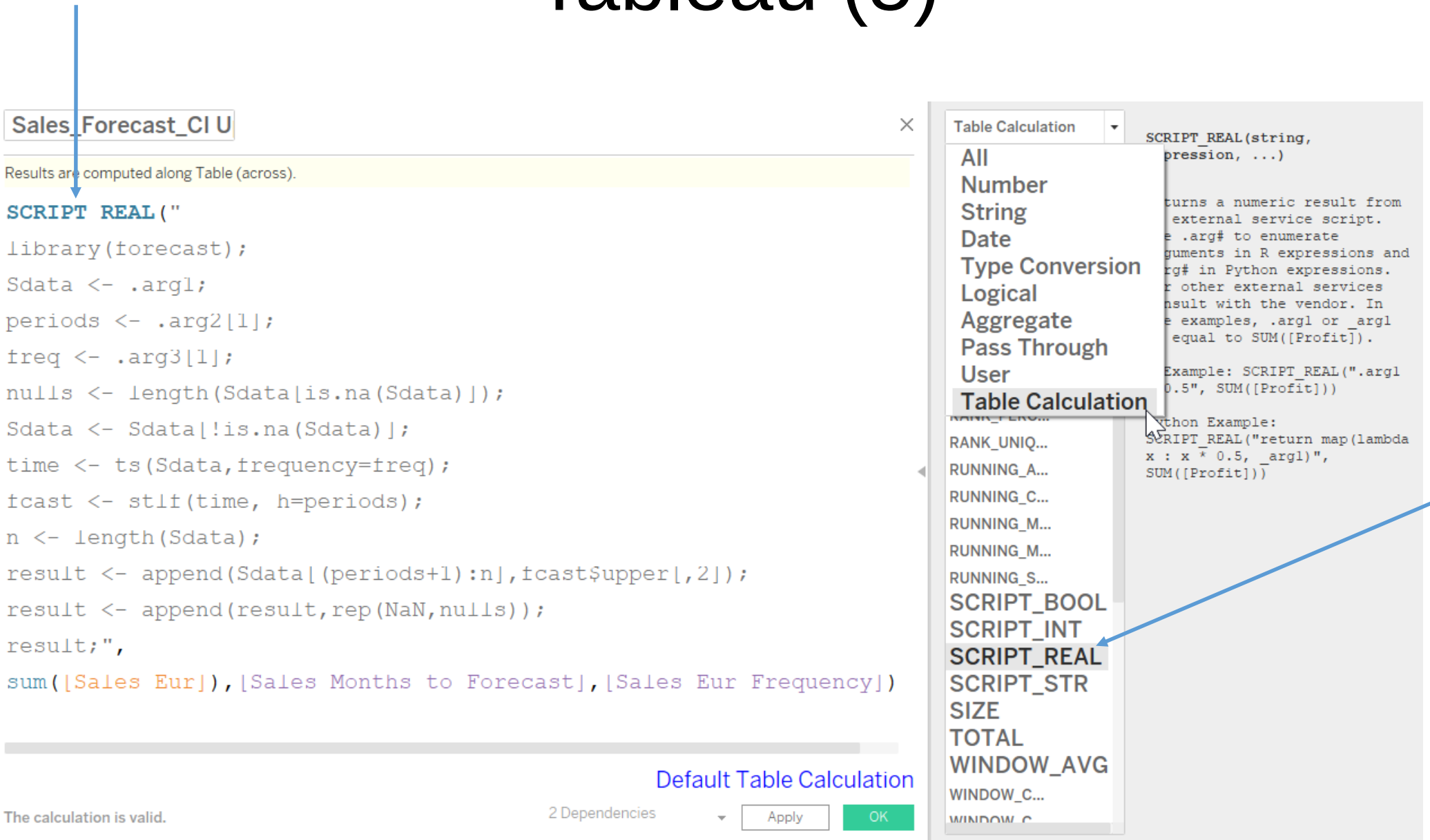

Sales_Forecast_CI U
Results are computed along Table (across).
SCRIPT_REAL("
library(forecast);
Sdata <- .arg1;
periods <- .arg2[1];
freq <- .arg3[1];
nulls <- length(Sdata[is.na(Sdata)]);
Sdata <- Sdata[!is.na(Sdata)];
time <- ts(Sdata,frequency=freq);
fcast <- stlf(time, h=periods);
n <- length(Sdata);
result <- append(Sdata[(periods+1):n],fcast$upper[,2]);
result <- append(result,rep(NaN,nulls));
result;",
sum([Sales Eur]),[Sales Months to Forecast],[Sales Eur Frequency])
Default Table Calculation
The calculation is valid.
2 Dependencies
Apply
OK
Table Calculation
All
Number
String
Date
Type Conversion
Logical
Aggregate
Pass Through
User
Table Calculation
RANK_UNIQ...
RUNNING_A...
RUNNING_C...
RUNNING_M...
RUNNING_M...
RUNNING_S...
SCRIPT_BOOL
SCRIPT_INT
SCRIPT_REAL
SCRIPT_STR
SIZE
TOTAL
WINDOW_AVG
WINDOW_C...
SCRIPT_REAL(string,
pression, ...)
turns a numeric result from
external service script.
e .arg# to enumerate
guments in R expressions and
rg# in Python expressions.
r other external services
nsult with the vendor. In
e examples, .arg1 or _arg1
equal to SUM([Profit]).
Example: SCRIPT_REAL(".arg1
0.5", SUM([Profit]))
thon Example:
SCRIPT_REAL("return map(lambda
x : x * 0.5, _arg1)",
SUM([Profit]))

Tableau (6)

Sales_Complex_TS_A

Results are computed along Table (across).

```
SCRIPT_REAL("
    library(forecast)

    ## Creating vectors

    hold.orig <- .arg4
    len.orig <- length( hold.orig )
    len.new <- len.orig - hold.orig[1]

    year.orig <- .arg2
    month.orig <- .arg1
    sales.orig <- .arg3

    ## Sorting the Data

    date.orig <- year.orig + month.orig / 12
    dat.orig <- cbind(year.orig, month.orig, sales.orig)[sort(date.orig,
                    index.return = TRUE)$ix,]
    dat.new <- dat.orig[1:len.new,]

    ## Fitting the Time Series

    timeser <- ts(dat.new[,3], start = c(dat.new[1,1], dat.new[1,2]),
            end = c(dat.new[len.new,1], dat.new[len.new,2]), frequency = 12)
    rep(pact(timeser, plot=FALSE)$act,5)[1:len.orig]"

, ATTR( MONTH( [Date] ) ), ATTR( YEAR( [Date] ) ), SUM( [Sales Eur] ),
 [Months to Forecast] )
```

Default Table Calculation

The calculation is valid.

2 Dependencies | Apply | OK

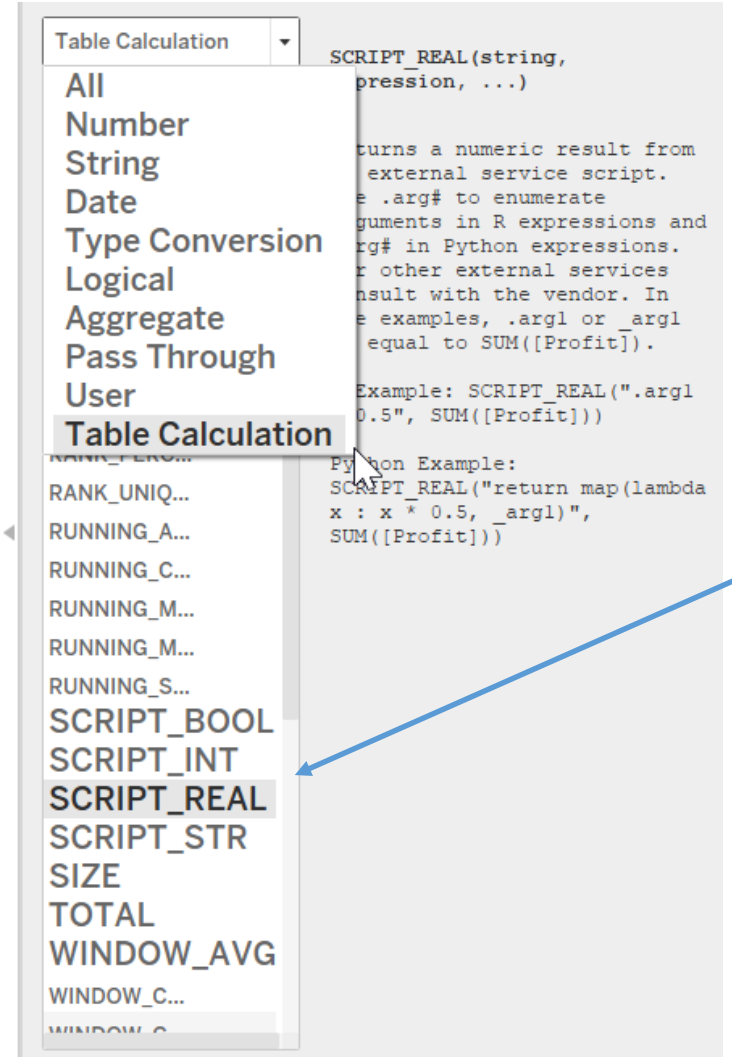


Tableau (7)

Clustering (Norm)

Results are computed along Table (across).

```
SCRIPT_INT("
## Sets the seed
set.seed(.arg5[1]);

V <- (.arg1 - mean(.arg1))/sd(.arg1);
A <- (.arg2 - mean(.arg2))/sd(.arg2);
B <- (.arg3 - mean(.arg3))/sd(.arg3);
C <- (.arg4 - mean(.arg4))/sd(.arg4);

dat <- cbind(V, A, B, C);

num <- 4

## create cluster

result<-kmeans(dat, num); result$cluster;",

SUM([V]), SUM(A), SUM(B), SUM(C),[Seed])
```

Default Table Calculation

The calculation is valid. 2 Dependencies Apply OK

Table Cal...

All
Number
String
Date
Type Conversion
Logical
Aggregate
User
Table Calculation

RUNNING_M...
RUNNING_S...
SCRIPT_BOOL
SCRIPT_INT
SCRIPT_REAL
SCRIPT_STR
SIZE
TOTAL
WINDOW_AVG

```
SCRIPT_INT(string,
pression, ...)

turns an integer
sult from an
ternal service
ipt. Use .arg#
enumerate
guments in R
pressions and
g# in Python
pressions. For
her external
services consult
with the vendor.
In the examples,
.arg1 or _arg1 is
equal to
SUM([Profit]).

R Example:
SCRIPT_INT("as.int
eger(.arg1 * 5)",
SUM([Profit]))
```

Tableau (8)

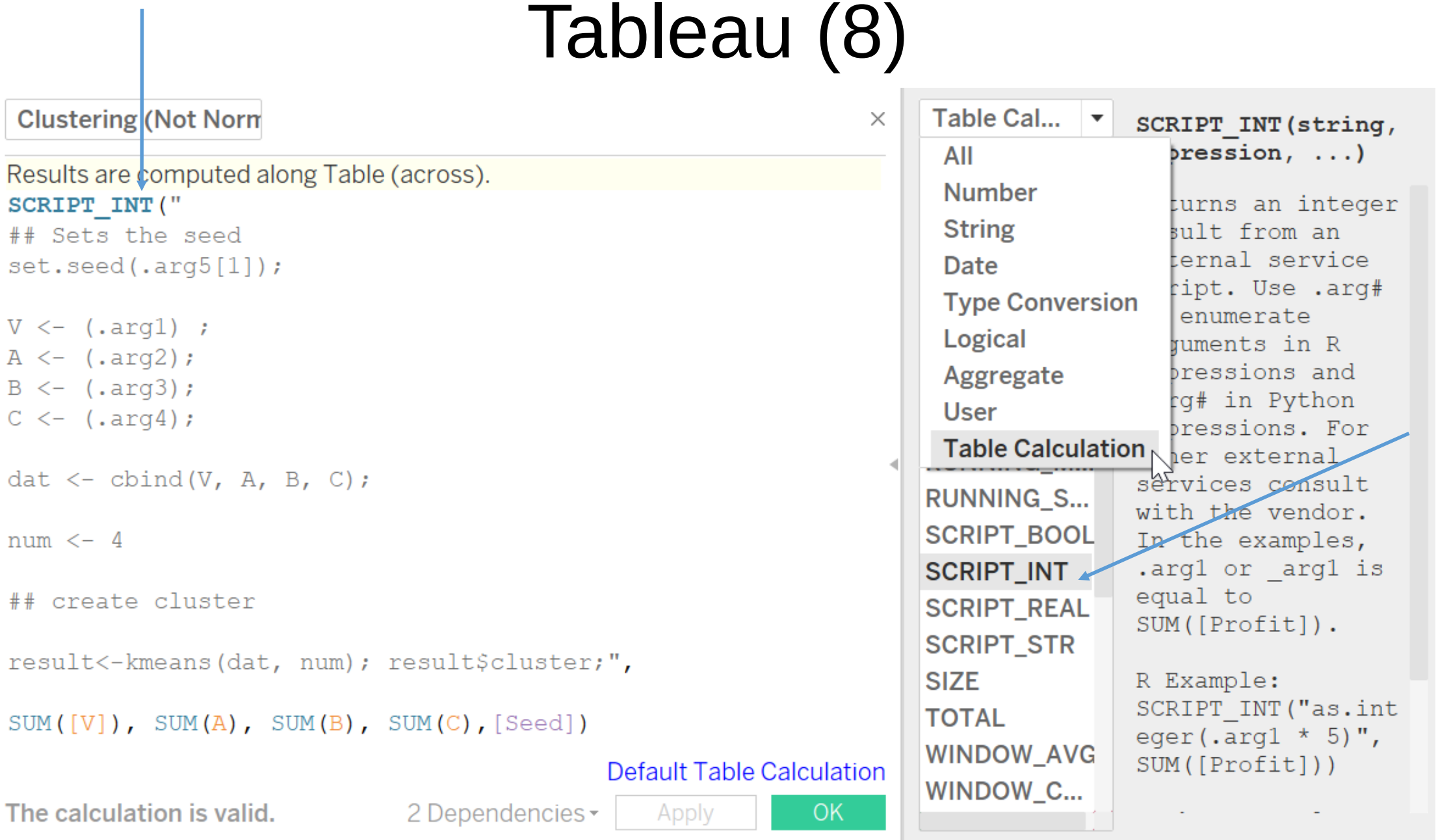

Clustering (Not Norm
Results are computed along Table (across).
SCRIPT_INT("
Sets the seed
set.seed(.arg5[1]);
V <- (.arg1) ;
A <- (.arg2);
B <- (.arg3);
C <- (.arg4);
dat <- cbind(V, A, B, C);
num <- 4
create cluster
result<-kmeans(dat, num); result$cluster;",
SUM([V]), SUM(A), SUM(B), SUM(C),[Seed])
Default Table Calculation
The calculation is valid.
2 Dependencies
Apply
OK
Table Cal...
All
Number
String
Date
Type Conversion
Logical
Aggregate
User
Table Calculation
RUNNING_S...
SCRIPT_BOOL
SCRIPT_INT
SCRIPT_REAL
SCRIPT_STR
SIZE
TOTAL
WINDOW_AVG
WINDOW_C...
SCRIPT_INT(string,
pression, ...)
turns an integer
sult from an
ternal service
ript. Use .arg#
enumerate
guments in R
pressions and
rg# in Python
pressions. For
her external
services consult
with the vendor.
In the examples,
.arg1 or _arg1 is
equal to
SUM([Profit]).
R Example:
SCRIPT_INT("as.int
eger(.arg1 * 5)",
SUM([Profit]))

Tableau (9)

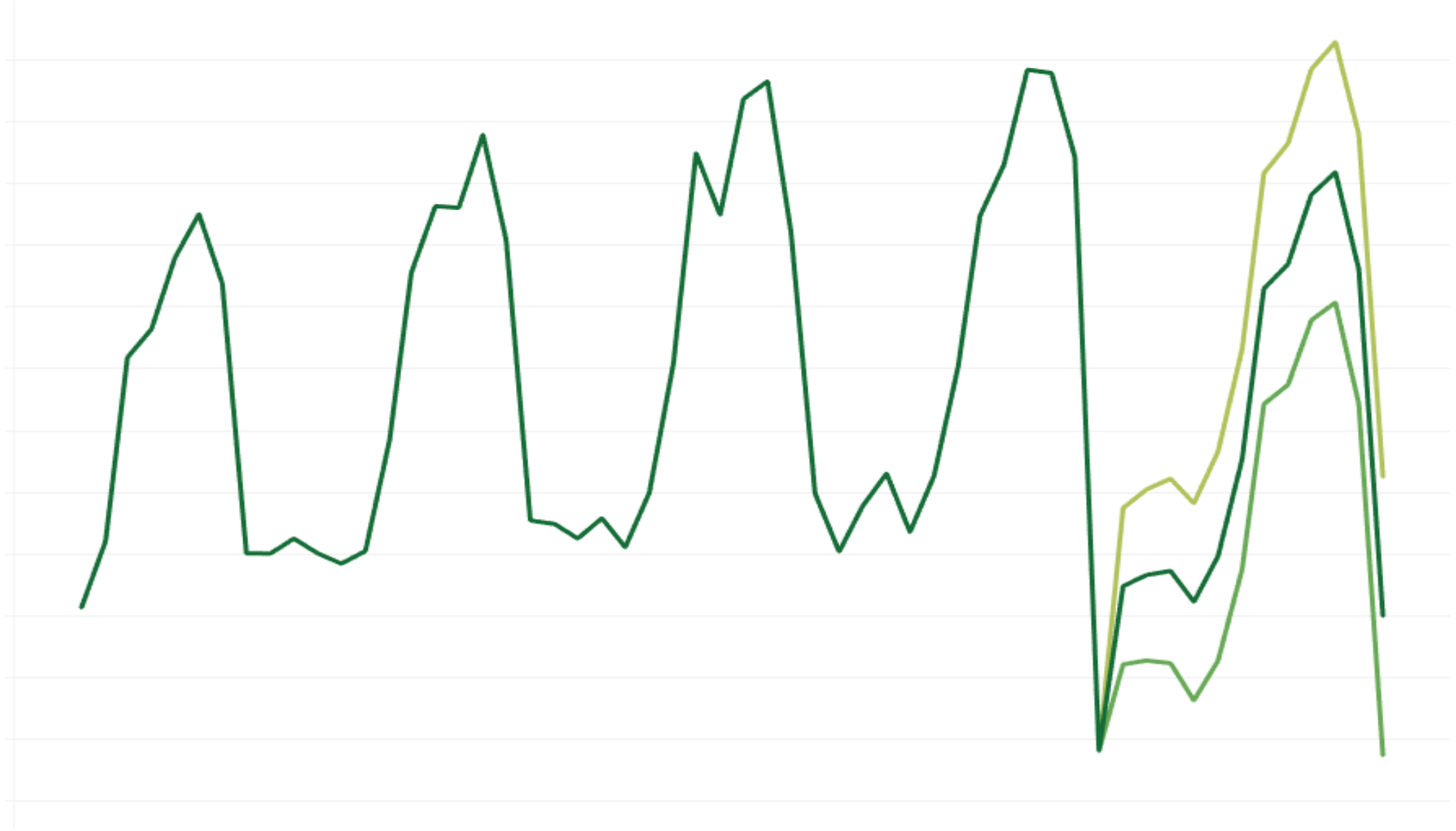


Tableau (10)

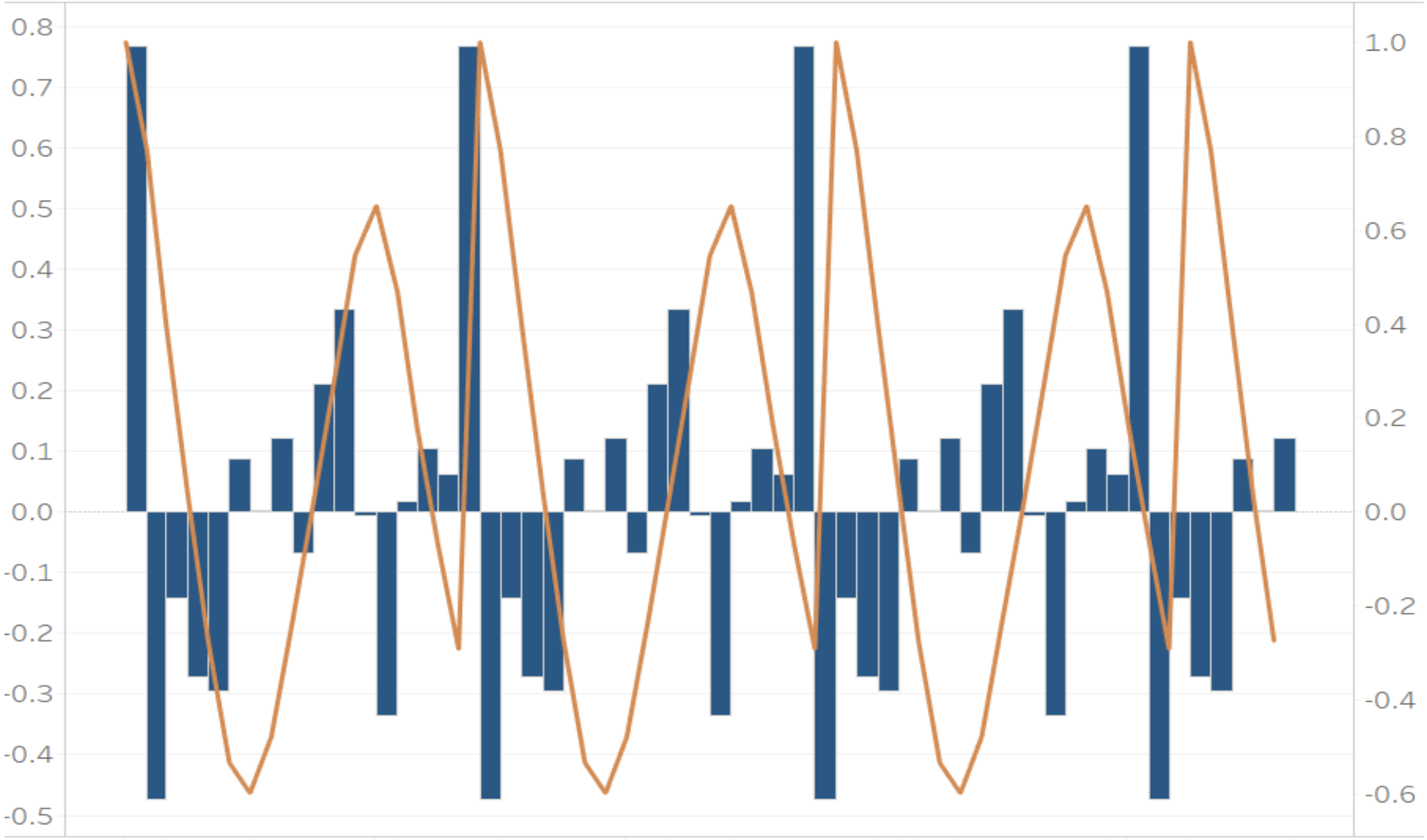


Tableau (11)

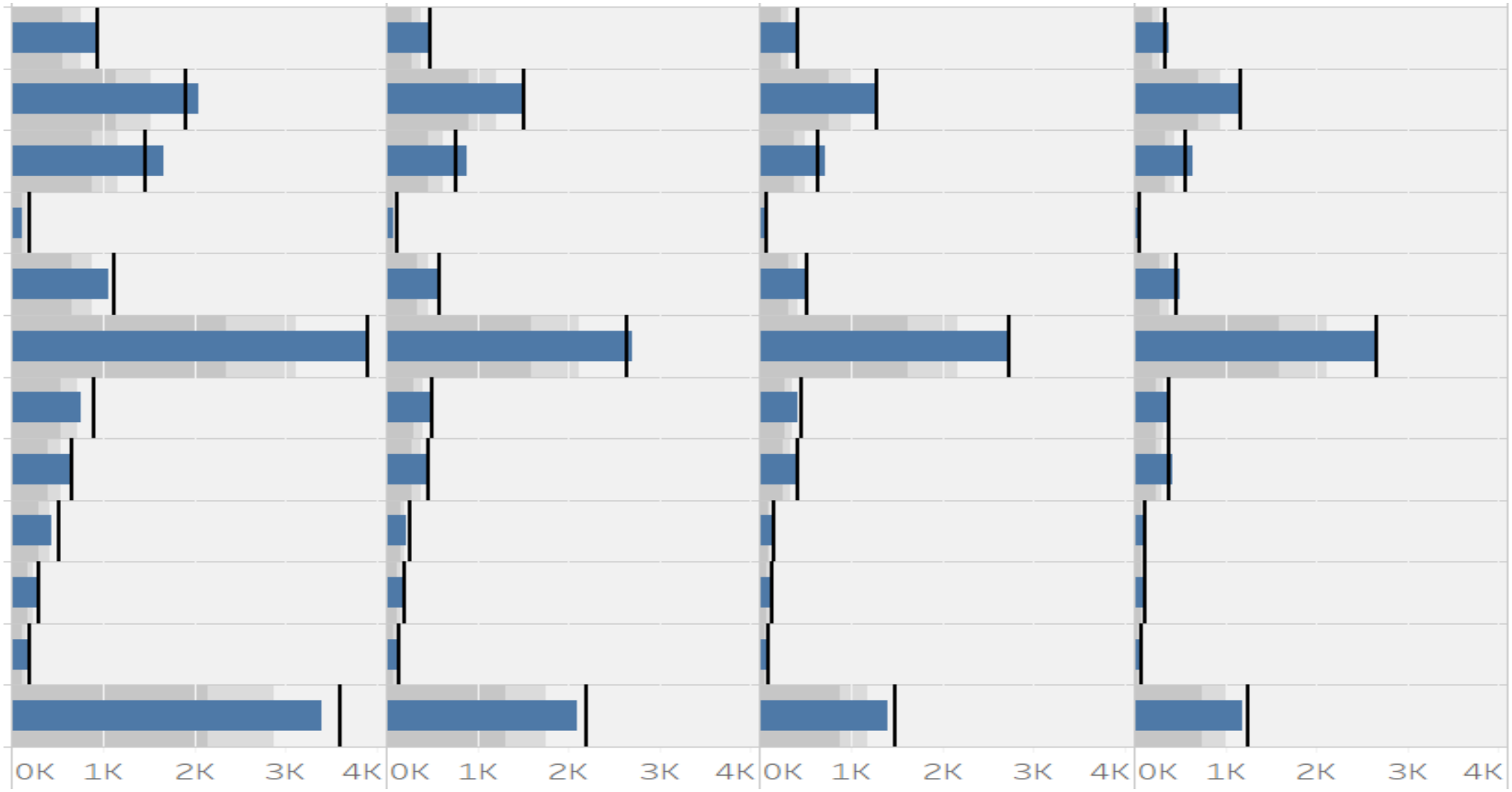


Tableau (12)

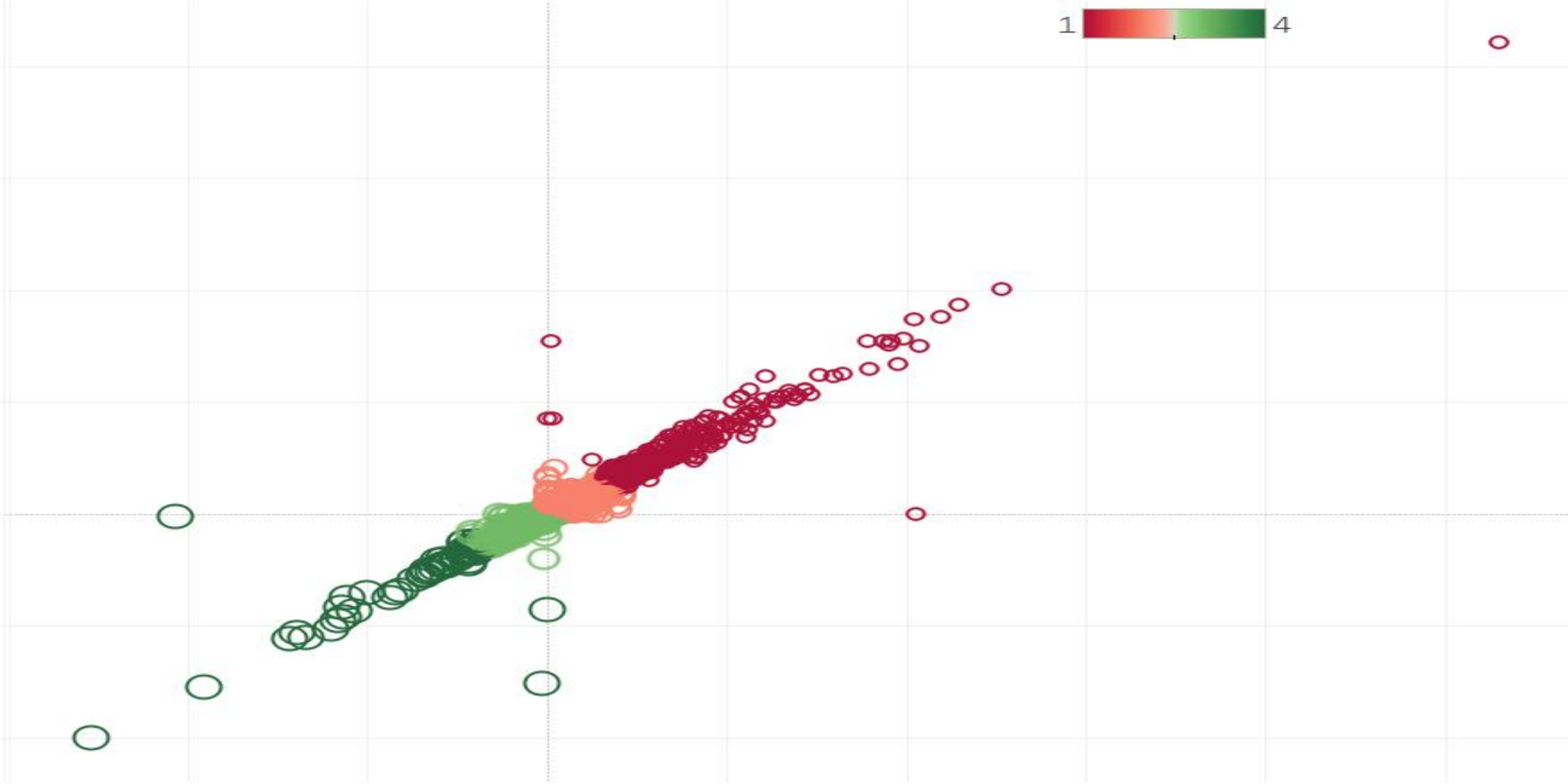


Tableau (13)

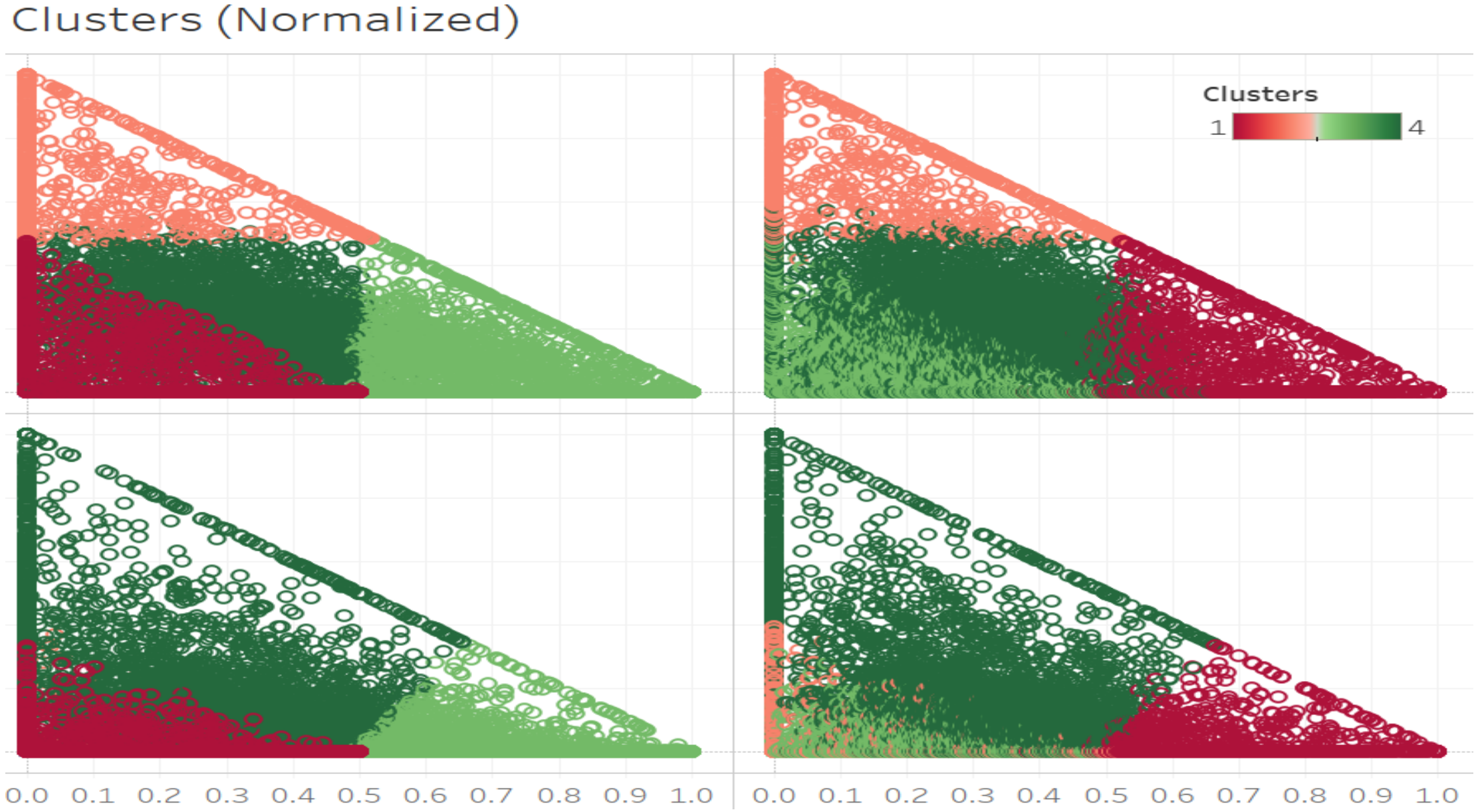

Clusters (Normalized)
Clusters
1
4
0.0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1.0
0.0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1.0

Tableau (14)

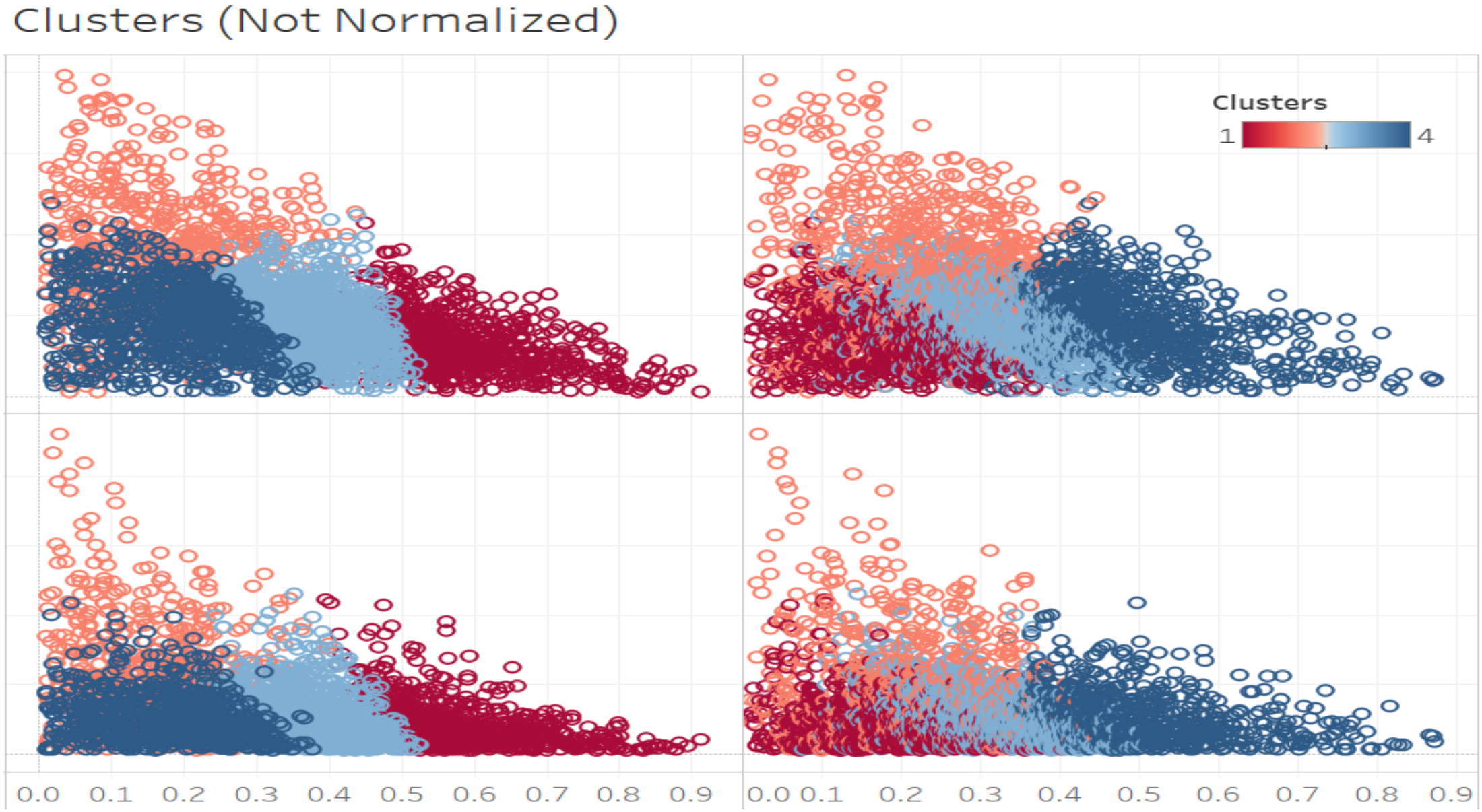
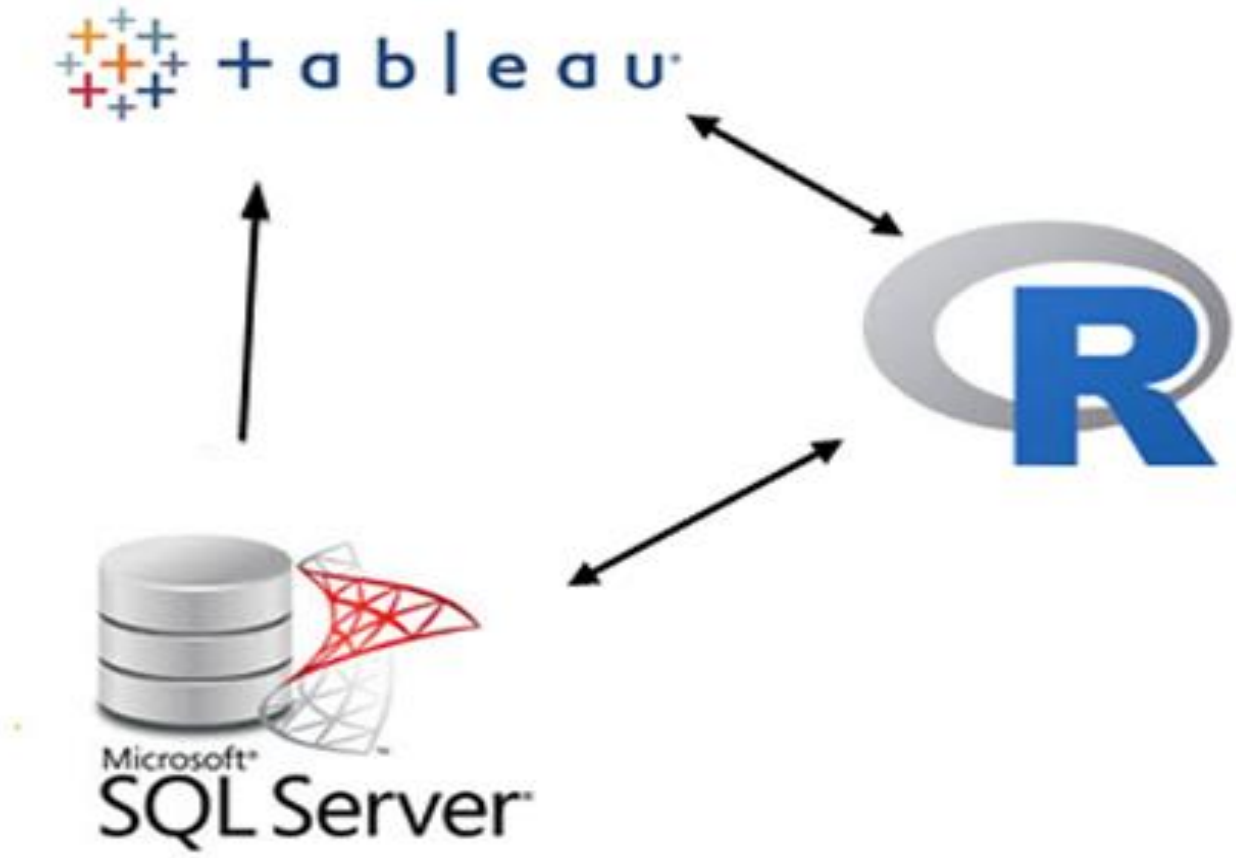

Clusters (Not Normalized)
Clusters
1
4
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9

Tableau + R + MS SQL Server



Thank you