This document is step by step guide of integrating R with Tableau and display correlations calculated in R on sample Tableau chart.

Integrating Tableau and R

Jayesh Zope (363039)

What is R?

R is a language and an environment for most latest statistical and machine learning libraries and awesome graphics. It is open source, distributed under GNU license and so, a popular among academia and industry alike. It provides a wide variety of statistical (linear and nonlinear modelling, classical statistical tests, time-series analysis, classification, clustering etc) and graphical techniques, and is highly extensible. The rich ecosystem that R enjoys and the talent gap that it fills being famous among student, makes it the language of future.

Together, R and Tableau could be really potent couple that today's data science has to offer to solve any organization's end to end data discovery needs.

What are the benefits of using Tableau and R integration?

Tableau and R integration enables us to -

- Access R packages or functions for quantitative analysis
- Take advantages of Tableau's data visualization capacities

There are four new built-in functions that are used to call specific R models and functions. The functions are

1. SCRIPT_BOOL: Return a Boolean
2. SCRIPT_INT: Return an Integer
3. SCRIPT_REAL: Return a Real
4. SCRIPT_STR: Return a String

These functions are different only in the type of result they return.

Setup Steps:
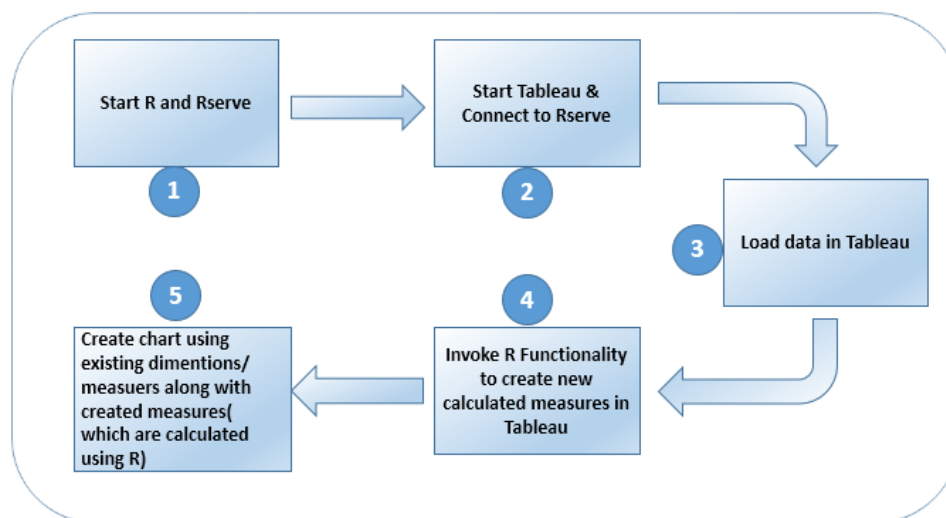
1. Start R and Rserve:

    a. Download and Install R

    b. Download Rserve package using command install.packages("Rserve")

    c. After successful installation load the library: library(Rserve).

    d. Start R Server using command: Rserve()

```
> library(Rserve)
> Rserve()
Starting Rserve...
 "C:\Users\Jayesh\DOCUME~1\R\WIN-LI~1\3.2\Rserve\libs\x64\Rserve.exe"
>
```
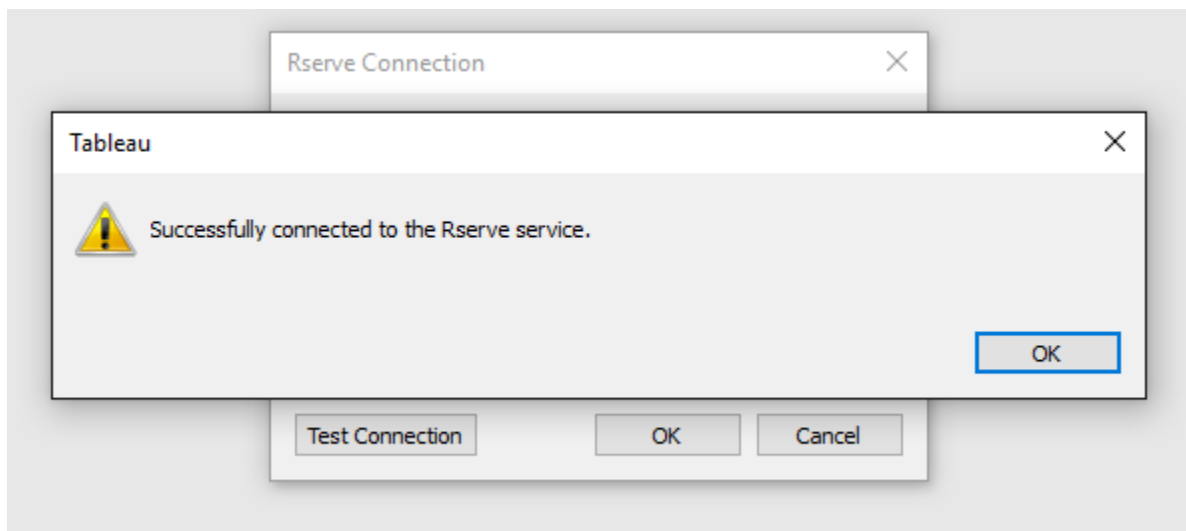
2. Start Tableau & connect to Rserve:
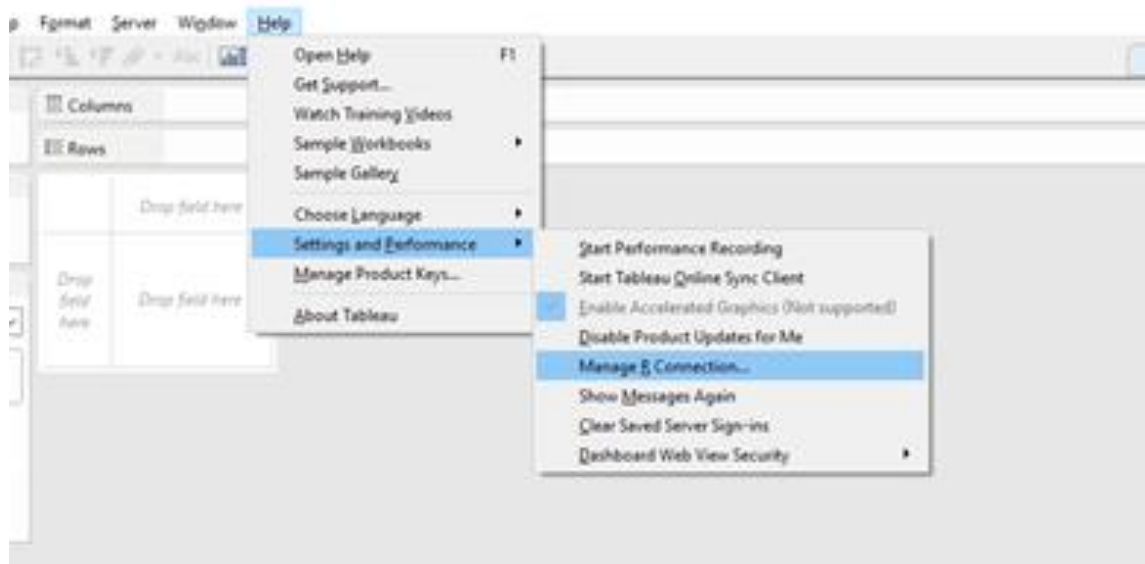
✓ Click Help->Settings and Performance-> Manage R Connections

✓ Enter required details on Rserve Connection wizard.

    • Specify a Server name and Port.

    • Enter Username Password
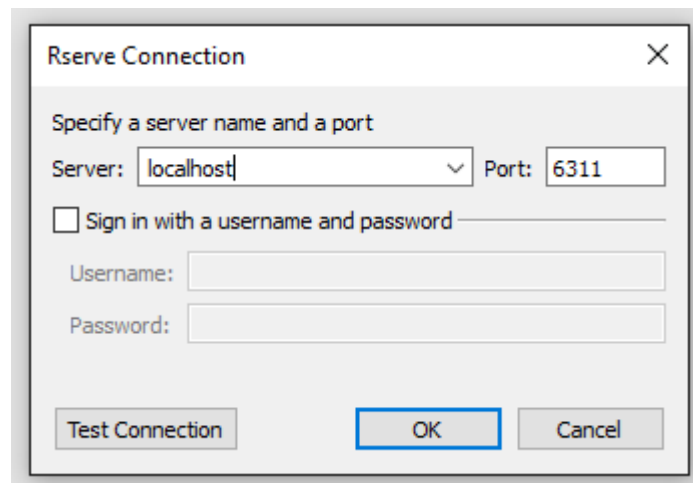
✓ Click on 'Test Connection' button.

✓ After successful connection Tableau displays a popup "Successfully connected to Rserve service"

3. Load data in Tableau:

Tableau and Rserve are now connected, we can load data to create visualizations in Tableau and invoke R functionality.

Let's take Sample Superstore Data file for better understating of how to use R with Tableau.

In this example, we will use a scattered plot chart, where user will have options to select measures for plotting the chart. Chart should display Correlation Coefficient in tooltip for selected measures.

- ✓ Create a Parameter called "Measure1"

- ✓ Add values 'Sales', 'Quantity' to parameter Measure1.

Create Parameter   ✕

Name: Measure1

Properties

Data Type: String ▼

Current value: Sales ▼

Display format: [   ] ⌄

Allowable values: ◯ All   ⦿ List   ◯ Range

List of values

Value	Display As
Sales	Sales
Quantity	Quantity
Add	

Add from Parameter ▶
Add from Field ▶
Paste from Clipboard

Clear All

OK   Cancel

✓ Similarly create Measure2, and add values 'Profit', 'Discount'

Edit Parameter [Measure2] ✕

Name: Measure2 | Comment >>

Properties

Data Type: String ▼

Current value: Profit ▼

Display format: ▼

Allowable values: ○ All ● List ○ Range

List of values

Value	Display As
Profit	Profit
Discount	Discount
Add	

Add from Parameter ▶
Add from Field ▶
Paste from Clipboard

Clear All

OK | Cancel

✓ Create calculated measure fields, with case statements as follows. We will be using these calculations on scattered plot.

Measure1

```
Case
[Measure1]
when "Sales" then SUM([Sales])
when "Quantity" then SUM([Quantity])
END
```

Measure2

```
CASE
[Measure2]
when "Profit" then SUM([Profit])
when "Discount" then SUM([Discount])
END
```

4. Create calculated filed using R functions:

✓ Create a calculated field called 'Correlation' using SCRIPT_REAL function as follows-
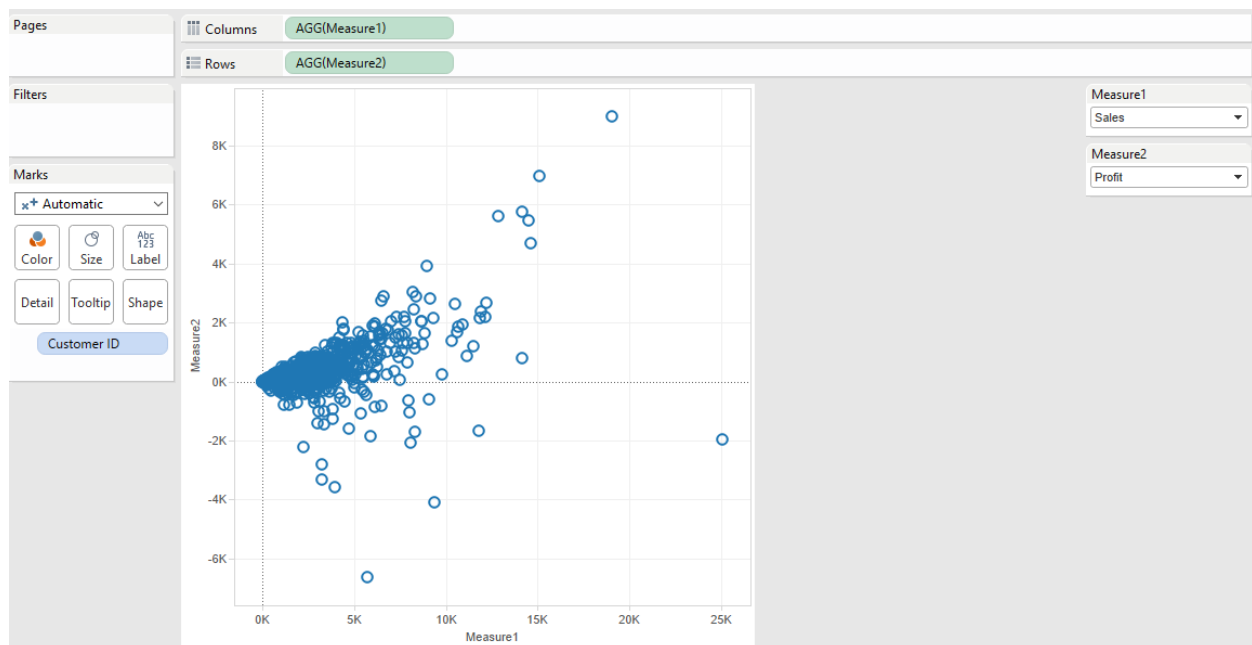
Correlation

```
Results are computed along Table (Across).
SCRIPT_REAL("cor(.arg1, .arg2)", ([Measure1]), ([Measure2]))
```
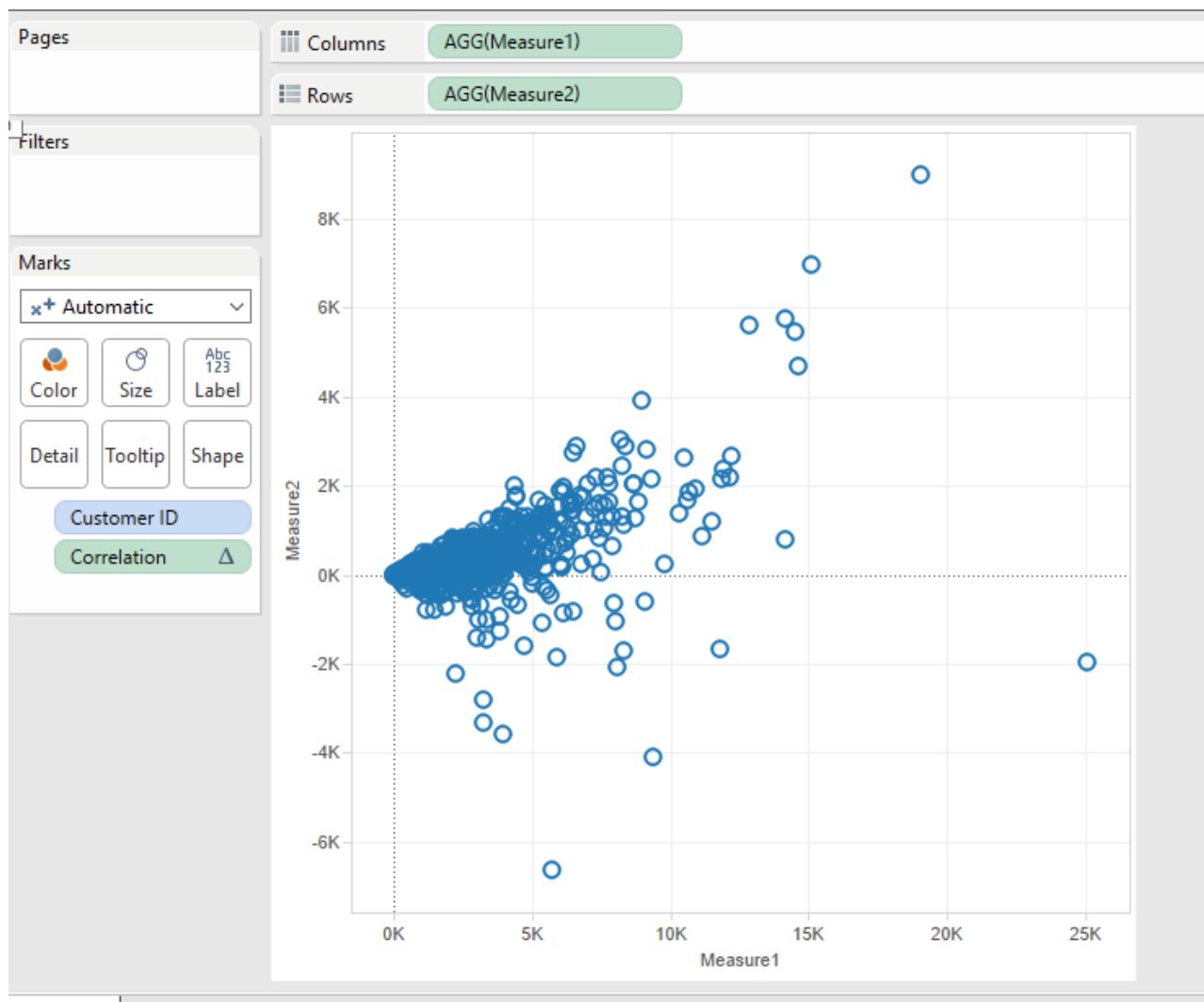
This will calculate Pearson Coefficient of selected measures.

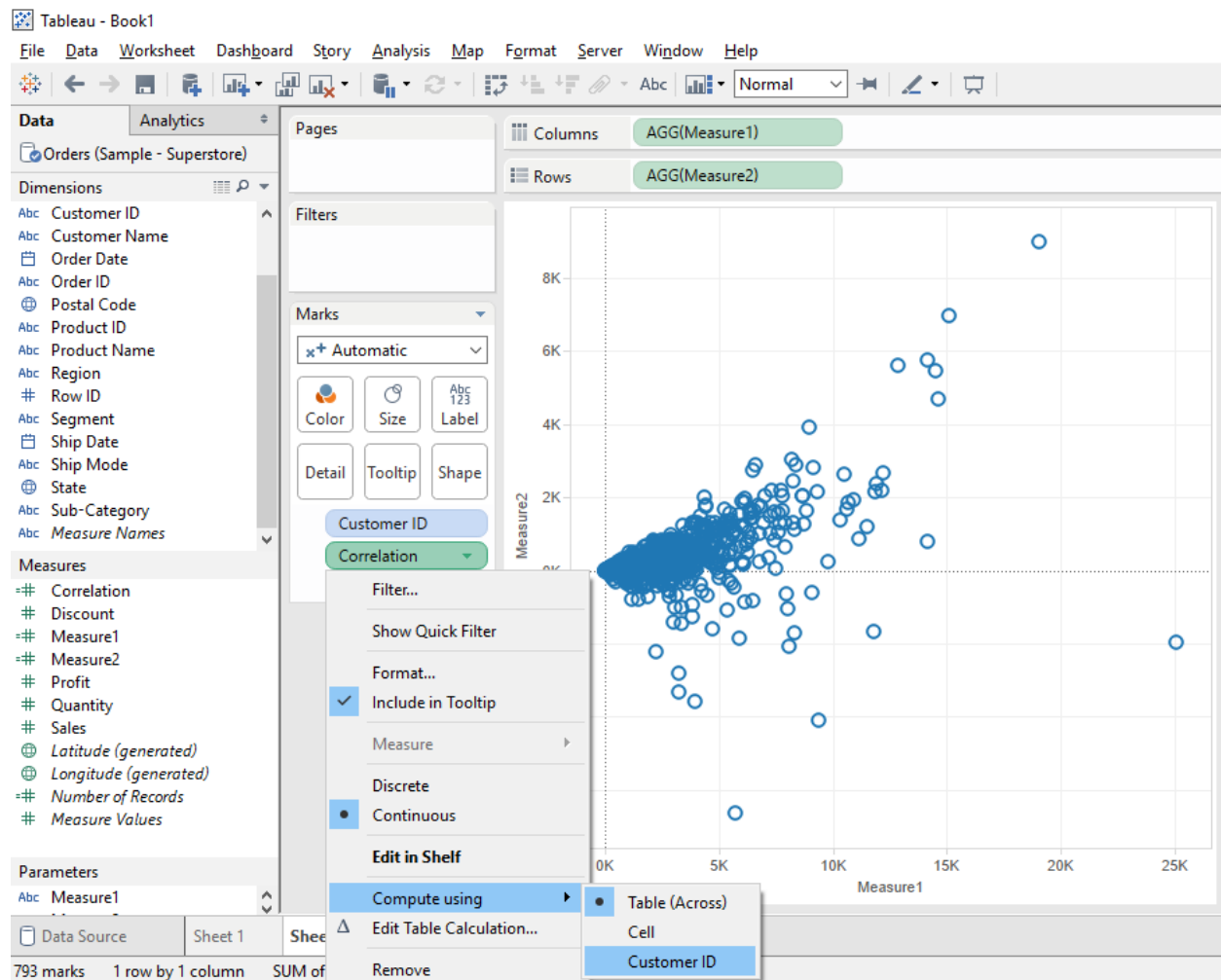## 5. Create chart using existing dimentions/measuers along with created measures (which are calculated using R)

✓ Create scattered plot chart as follows. Keep Customer ID on Details

✓ Drag 'Correlation' calculation on Marks shelf as Detail.

✓ Compute "Correlation" using Customer ID.

✓ Tooltip will display correlation coefficient of selected measures as shown in screenshot below.