On-Demand Training: Using R within Tableau Transcript

Welcome to this R Integration video. You can download the Exercises workbook to follow along in your own copy of Tableau. However, R and Rserve must be installed on your machine to fully follow along with this video. Refer to the video How to Integrate R and Tableau for more information.

This video is intended to show how to leverage R to extend Tableau's capabilities and to visualize outputs from R. This is not a tutorial on R.

Writing R in Tableau

R functions are computed as table calculations in Tableau. In order to let Tableau know a calculation needs to go to Rserve, it must be passed using one of four functions:

- SCRIPT_BOOL

- SCRIPT_INT

- SCRIPT_REAL

- SCRIPT_STR

The argument for these functions is the R expression, followed by the Tableau reference parameters.

For example, here is a function which simply adds 5 and 7

- The R script (using assignments) is here

To communicate the same from Tableau,

- we'd start with **SCRIPT_INT** because we know our result will be an integer

- Then, we'd enter **(".arg1 + .arg2",**

- Followed by the assignments of the measures from Tableau

- Our data set simply contains two fields, five and seven

- Because this is a table calculation, the fields must be aggregated, so we're using minimum of each of the fields

- And if we evaluate this calculation, we get 12

Outlier Analysis

Here, we have a view of profit by the Quarter of Order Date. We want to color code when Profit was exceptionally high or low. R has a package for evaluating multivariate outliers, mvoutlier.

We have a calculated field here which will ask R what values are outliers

- To start, we have a Tableau IF statement, such that if the R script result is 0, label it outlier, otherwise call it normal.

- Inside that, we have the SCRIPT_REAL function, which lets Tableau know we'll be asking R for real number

- Let's take a look at the R script itself.

- We call the library mvoutlier package and run any functions from that package

- Next, we establish the placeholder, .arg1

- Tableau is agnostic to the R script within the quotes. Because of this, it's often a good idea to test all scripts within R or Rstudio before using them within Tableau

- Finally, let R know which measure, here we use SUM([Profit]), to use for .arg1

Now let's bring Outlier to Color, and we can see that we do have some outliers, colored in orange.

What if we want to see how our outlier analysis changes over time? As we drill down, we see the outliers automatically. Note that we don't have to modify or re-write our calculation or R script; Tableau's functionality works seamlessly with R.

Cluster Analysis

Next, we'll go over an example of cluster analysis using R. This example uses Fisher's Iris Data, clustering three species of Iris based on the dimensions of their petals and sepals.

We've built scatterplots of these 4 measurements and we want to cluster the points into three groups and see how well that matches the different species of Irises.

Let's take a look at the R script in the calculated field Cluster.

- We're asking for an integer result

- And the R script uses k-means clustering to create three clusters based on four measures

- And the measures from Tableau are those petal and sepal measurements

When we bring Clusters to color, we see the results of that cluster analysis from R visually. We also have the Species information, and if we bring that to Color instead of Cluster, we see that our analysis was pretty accurate.

The details of these calculations can be examined by downloading the workbook asset for this video.

Conclusion

Thank you for watching this R integration training video. We invite you to continue with the On-Demand Training videos to learn more about using Tableau.