Module 6 - Tableau for Data Science, July 2017

This week's Session: Predictive analysis using R

Part Ten

Dataset (press CTRL+ CLICK to open the link): Download

Dataset description: Sample super store data

Source: Tableau

Instructions:

- In order to let tableau know that the calculations need to go to R, it must be passed through one of the 4 functions: SCRIPT_BOOL , SCRIPT_INT , SCRIPT_REAL , SCRIPT_STR
- R Functions are computed as Table calculations in Tableau.
- Since these are table calculations, all the **Fields being passed to R must be aggregated** for example Sum (PROFIT), MIN (Profit), Max (Profit) etc.

Some basic Exercises

1. Double the shipping cost for all the products

```
SCRIPT_REAL("2*(.arg1)",SUM([]))
```

2. Calculate cost of products

```
SCRIPT_REAL("(.arg1 - .arg2)",SUM([]),SUM([]))
```

3. Identify profits in different product sub categories

```
SCRIPT_BOOL("(.arg1 > 0)",SUM([])
```

4. Concatenate strings values of product categories and product names

```
SCRIPT_STR(" paste(.arg1 ,.arg2)",
attr([]), attr([])
```

5. Divide profits by half

```
SCRIPT_INT("(.arg1 / 2)",SUM([])
```

Performing simple linear regression

6. Do price discount influence profit forecast

```
SCRIPT_REAL("
fit <- lm(.arg1 ~ .arg2 )
fit$fitted
"
,
SUM([Profit]),
Aggregation([Discount])

)
```

7. What other independent attributes would you consider to predict profit forecast

8. What other independent attributes would you consider to predict Revenue forecast