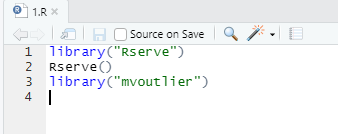
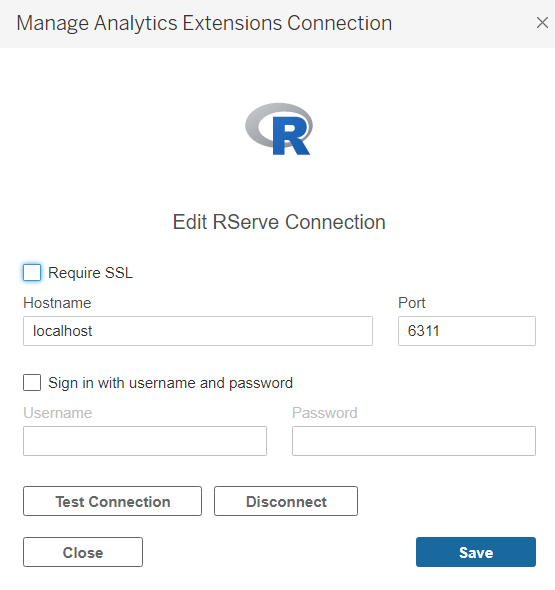
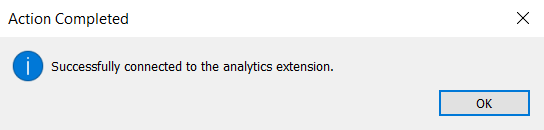
**Aryaman Mishra**

**19BCE1027**

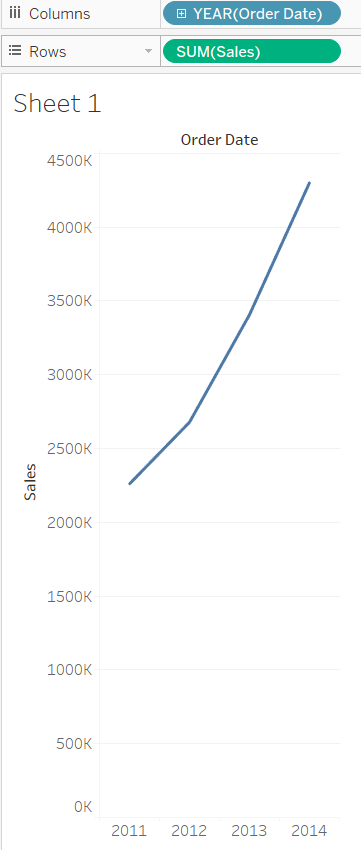
Find the outliers in superstore dataset using R in Tableau and write the steps for the following: o Connect to Rserve from Tableau

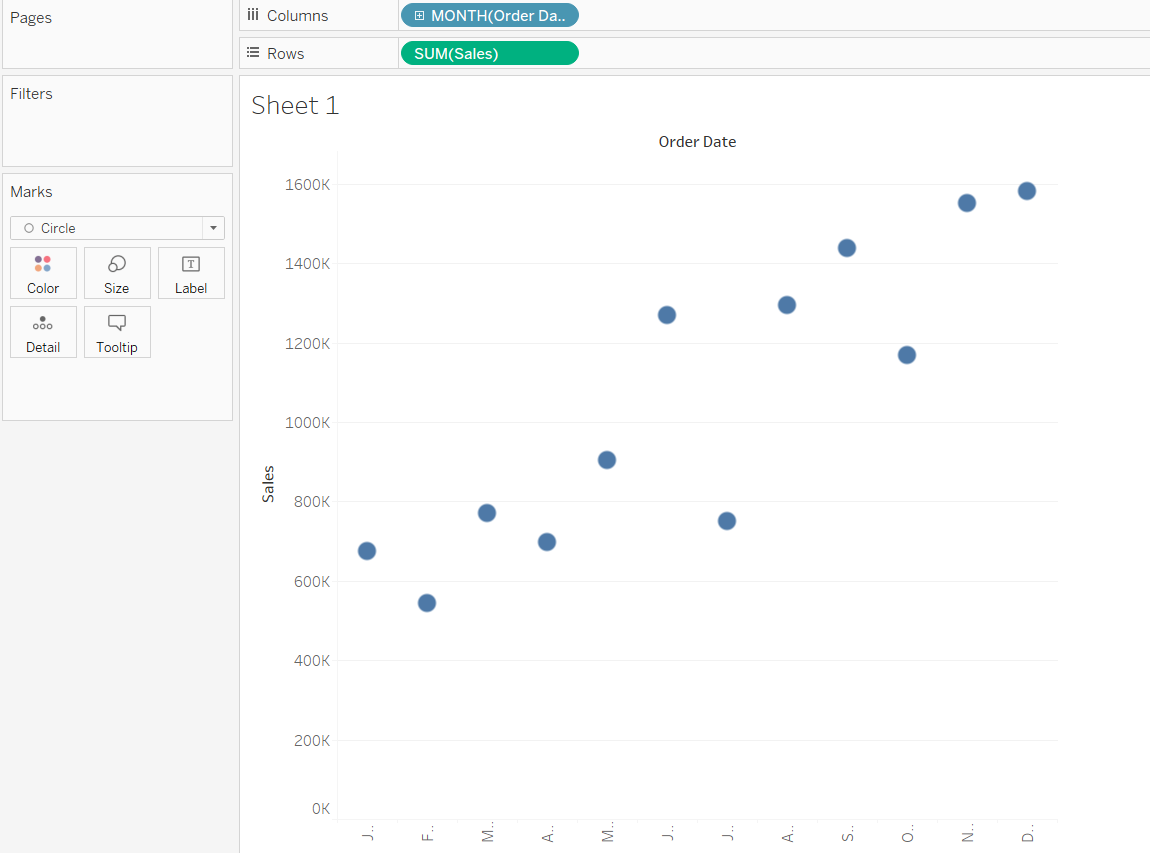
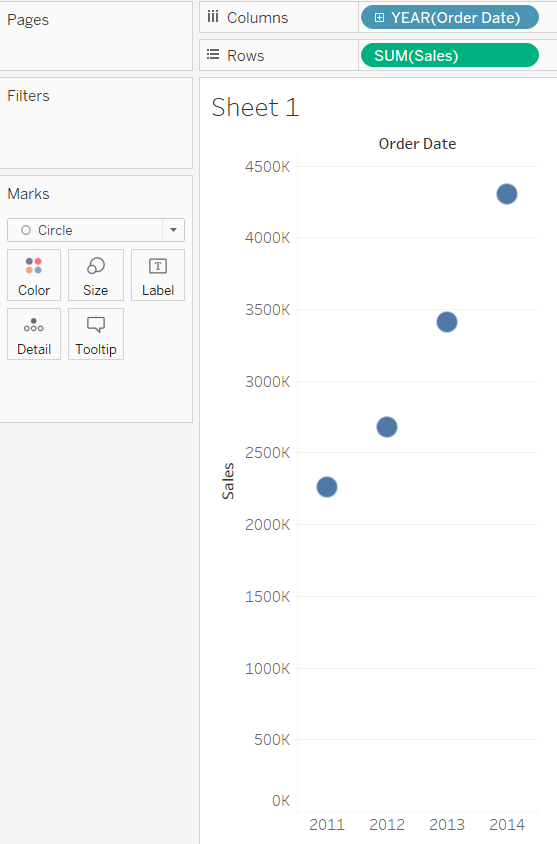




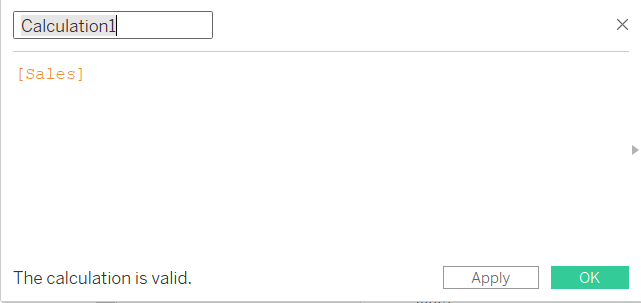


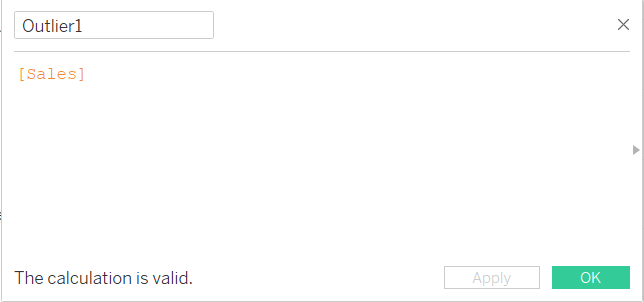
o Create a chart as per requirement in Tableau



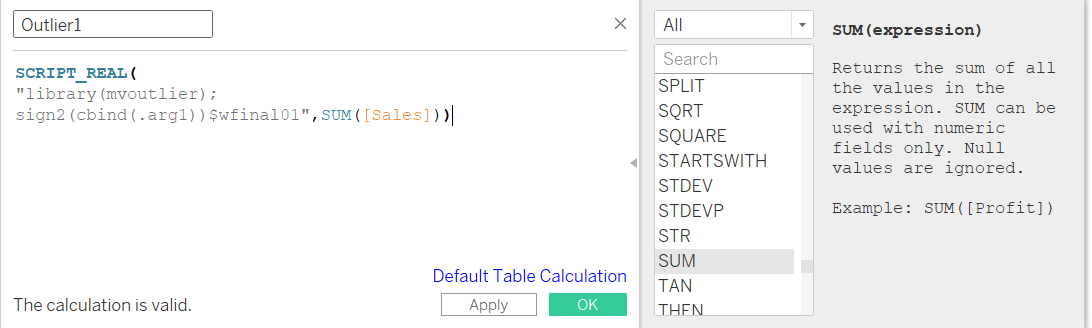


o Create a calculative field

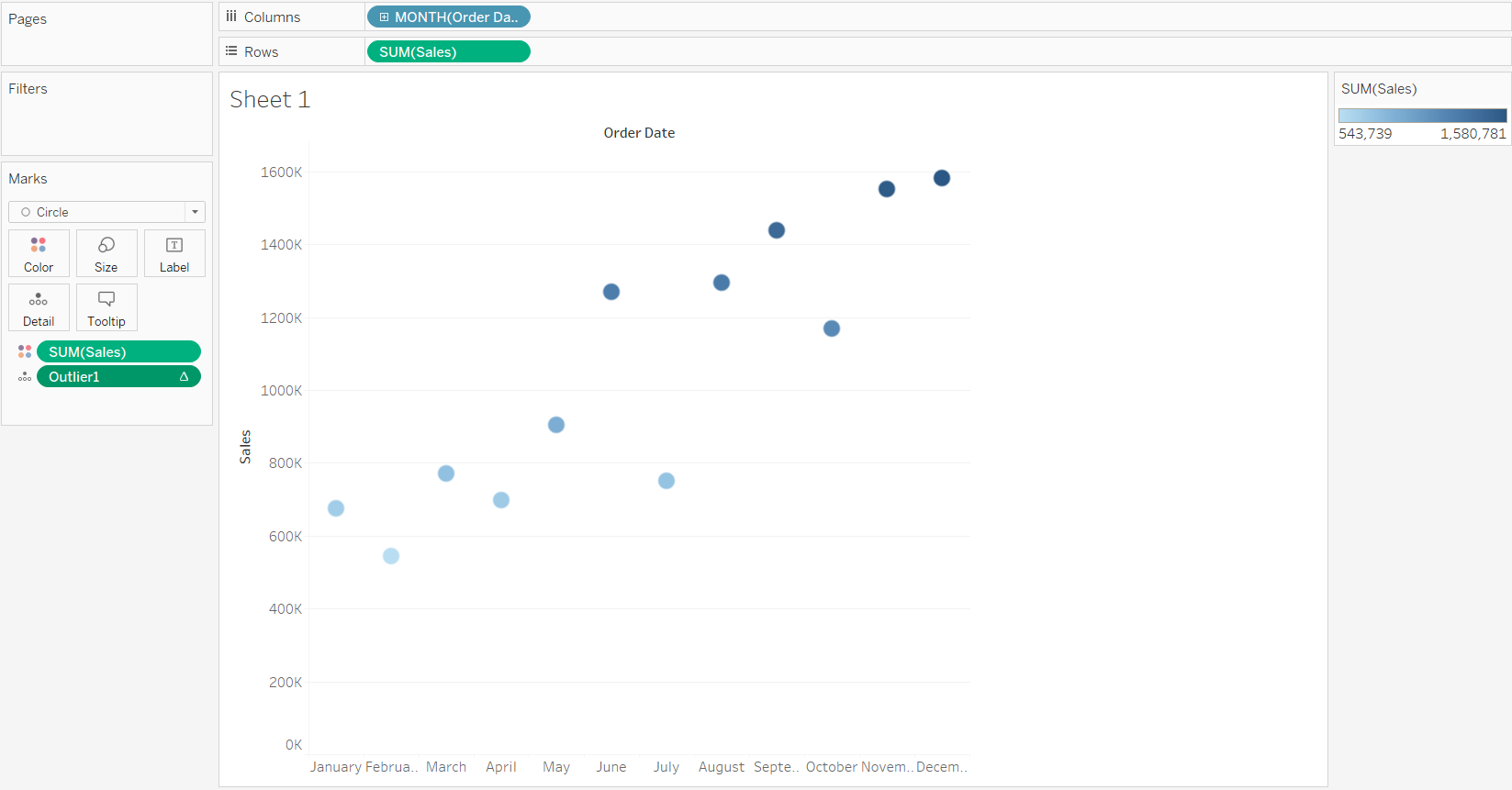




o Explanation to the calculative field



o Drop the field in color shelf to see the outliers

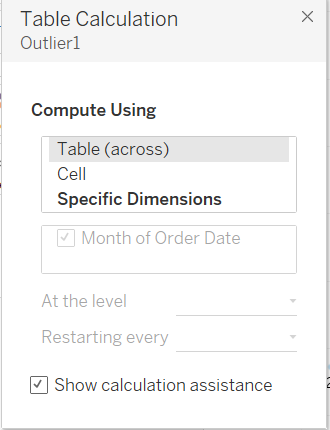


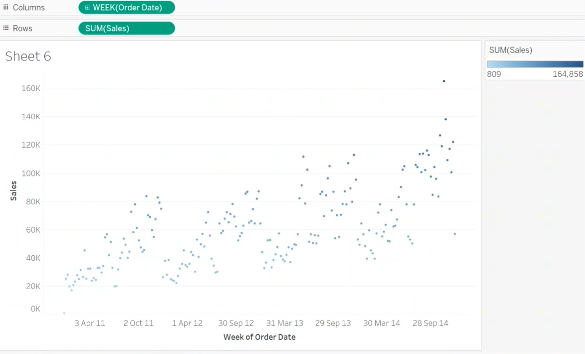
o Inference

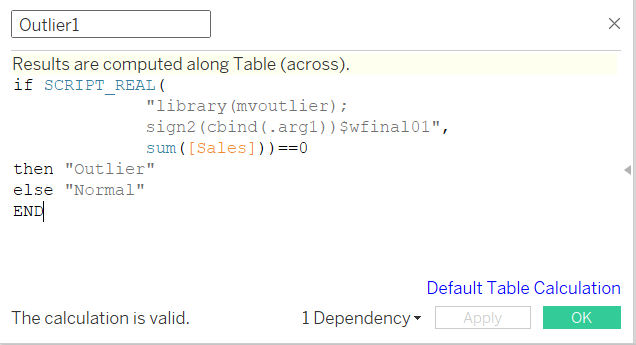
Extreme values are aggregated marks that are outliers, based on a model of the visualized marks. The selected mark is considered to contain an extreme value if a record value is in the tails of the distribution of the expected values for the data.

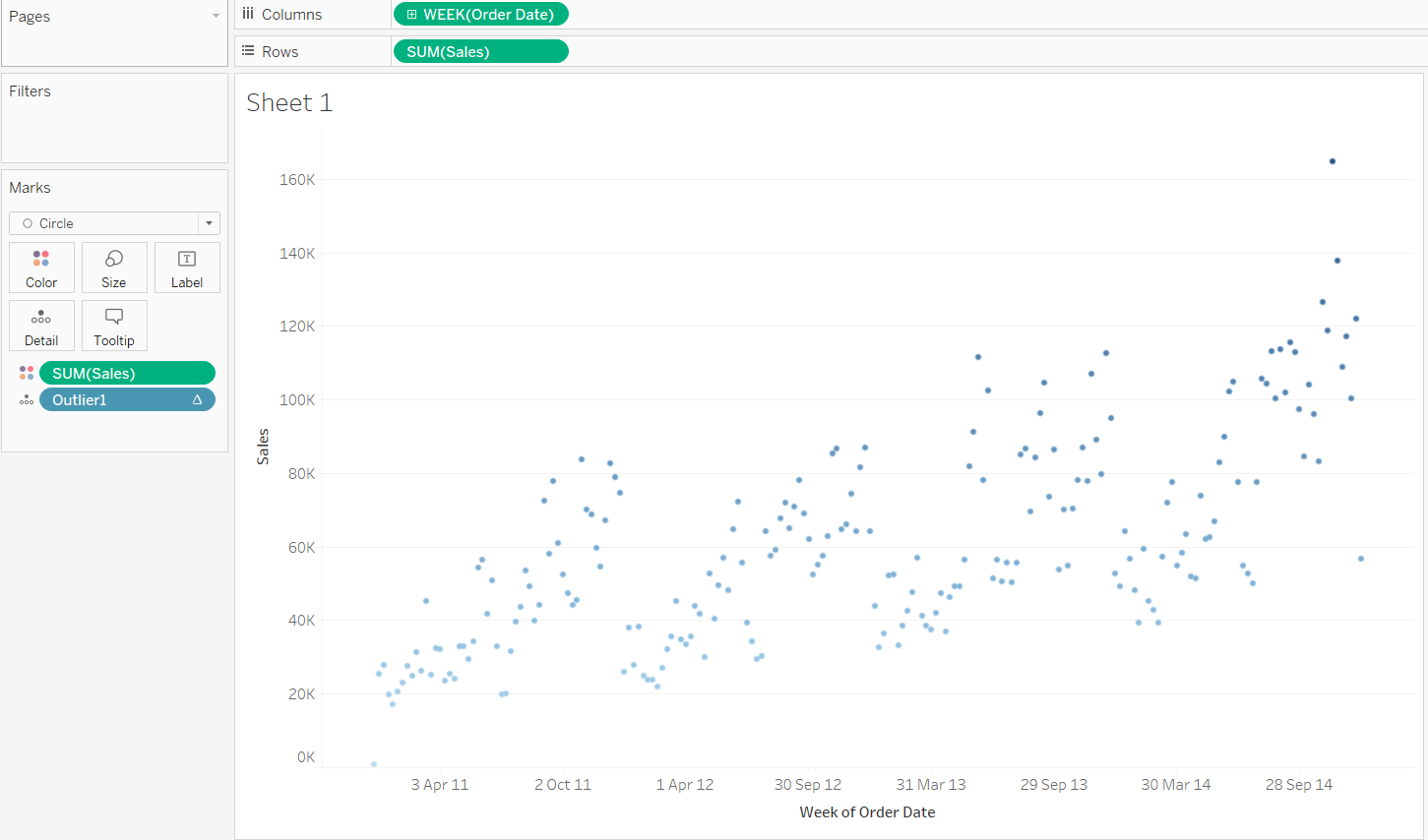
An extreme value is determined by comparing the aggregate mark with and without the extreme value. If the mark becomes less surprising by removing a value, then it receives a higher score.

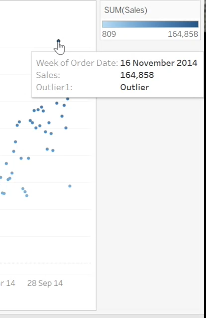
When a mark has extreme values, it doesn't automatically mean it has outliers, or that you should exclude those records from the view. That choice is up to you depending on your analysis. The explanation is simply pointing out an interesting extreme value in the mark. For example, it could reveal a mistyped value in a record where a banana cost 10 dollars instead of 10 cents. Or, it could reveal that a particular sales person had a great quarter.

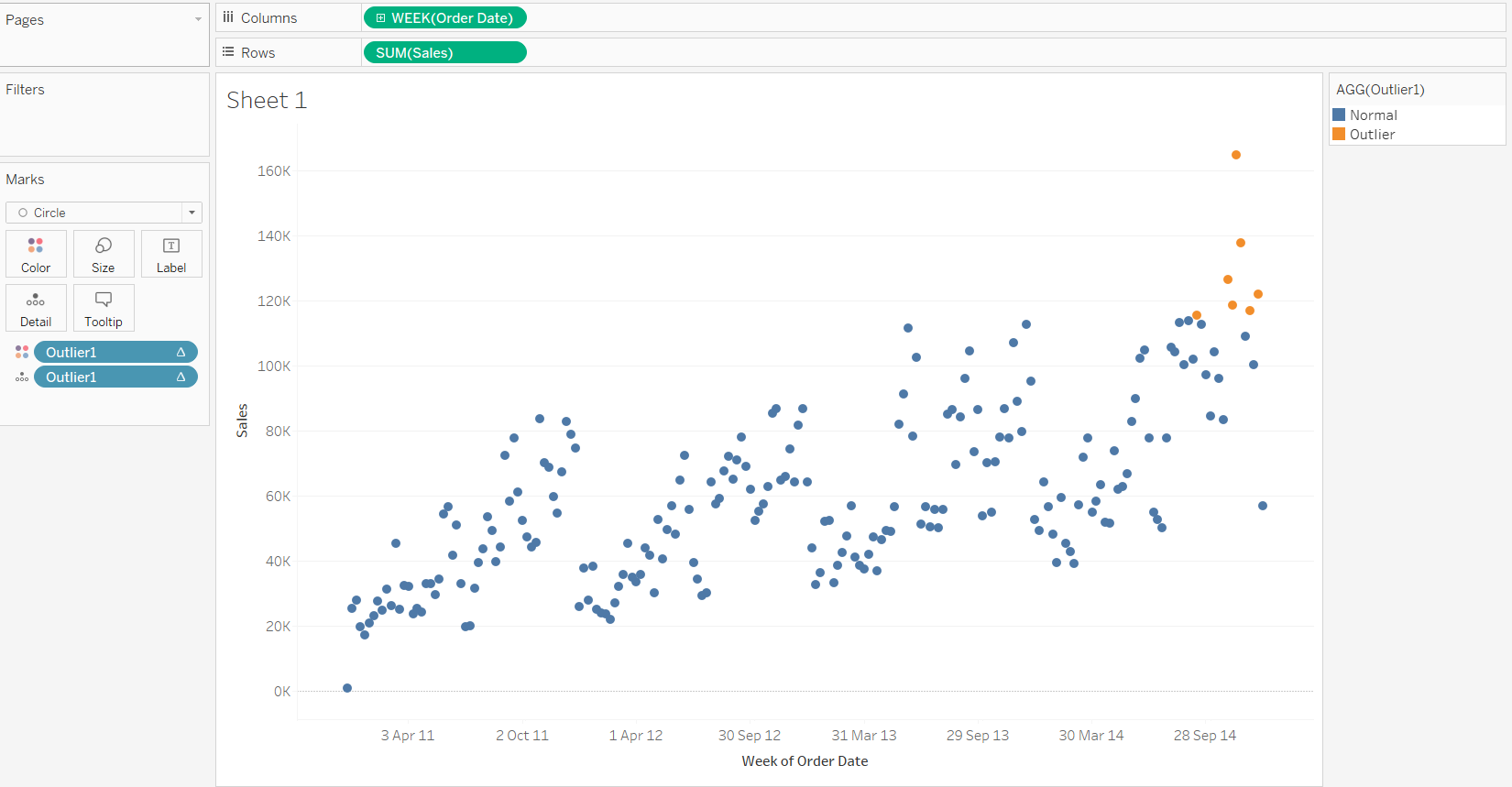








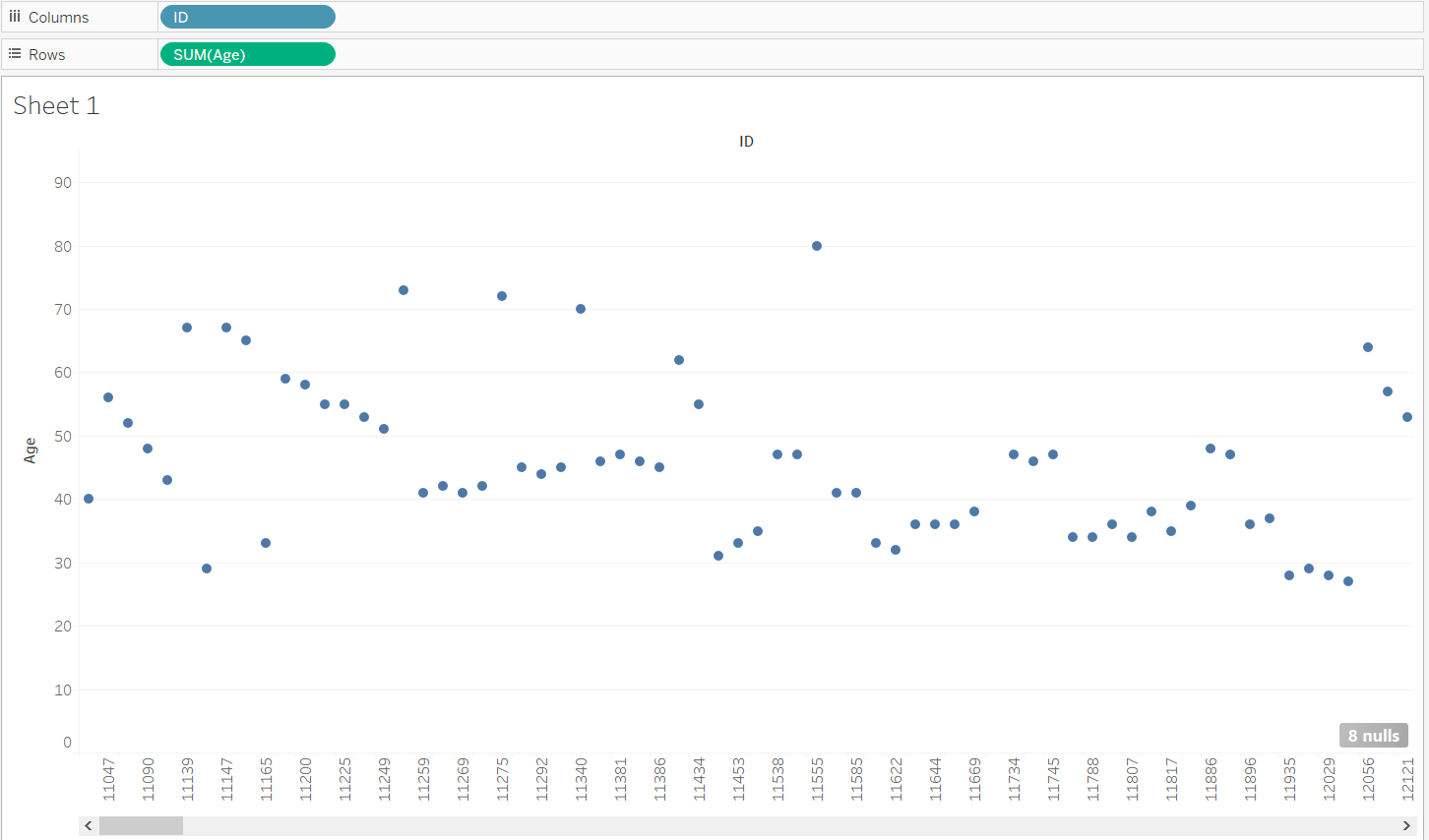


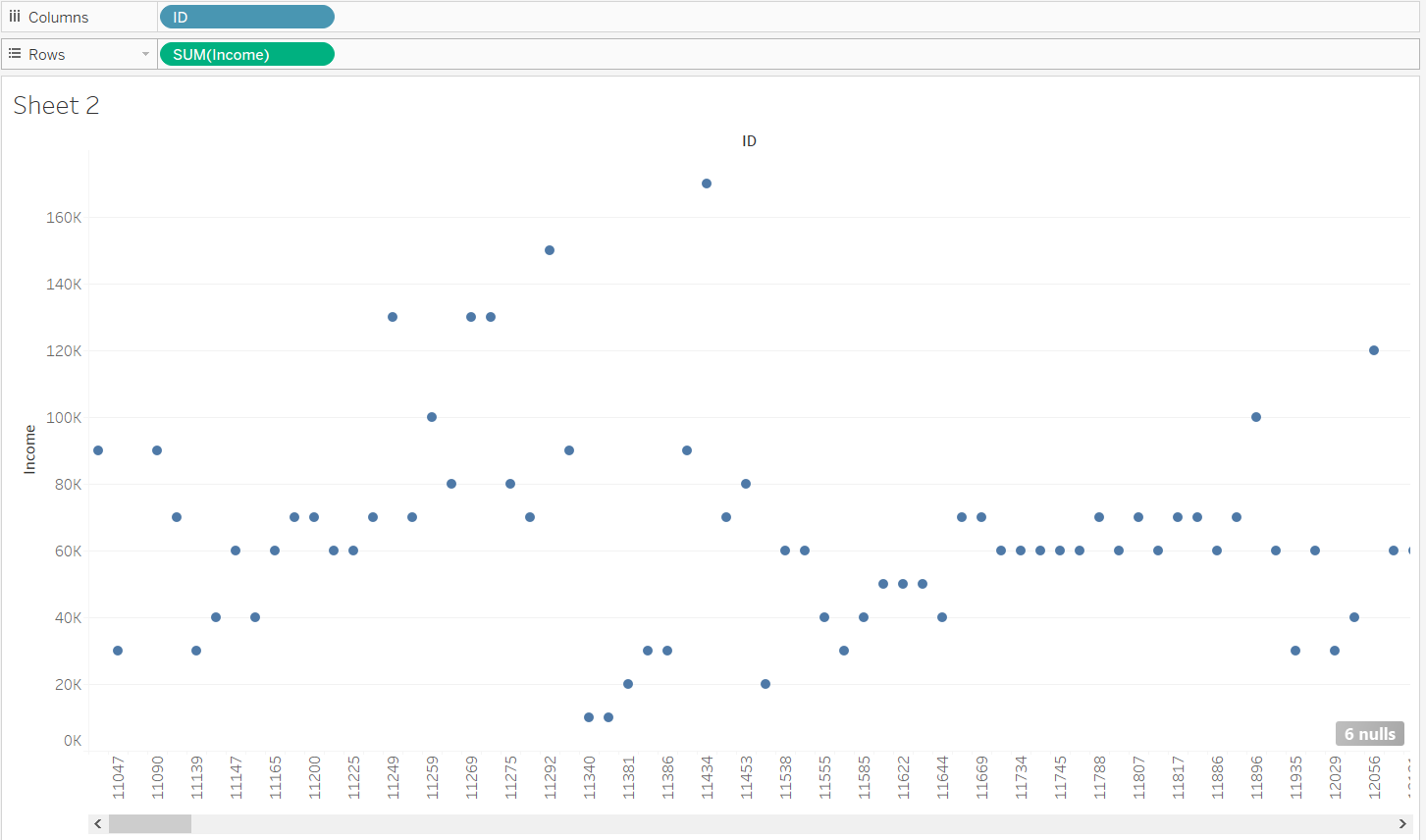


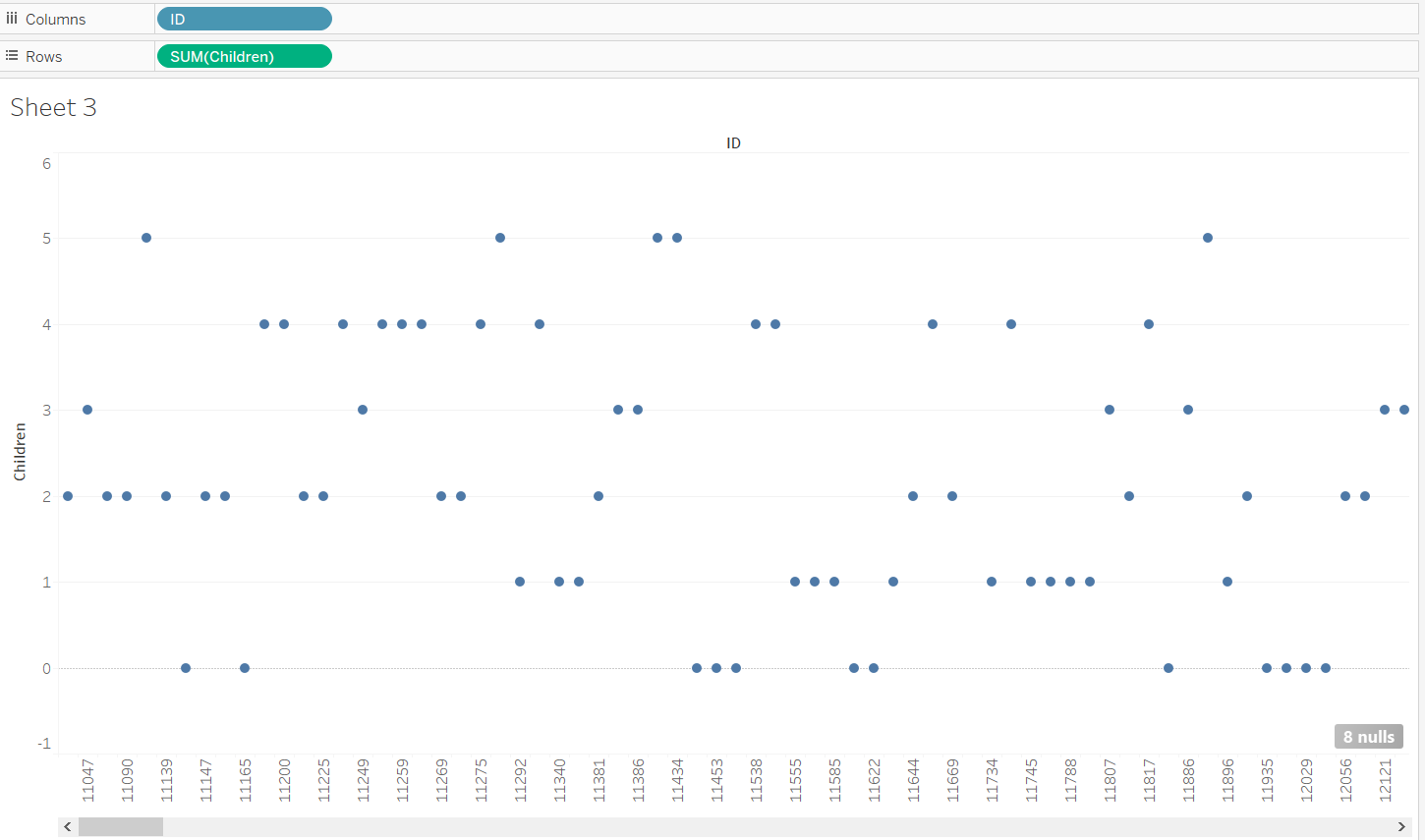
• Formulate the clusters based on age, income, children and cars from Bike buyers dataset using R in Tableau and write the steps for the following:

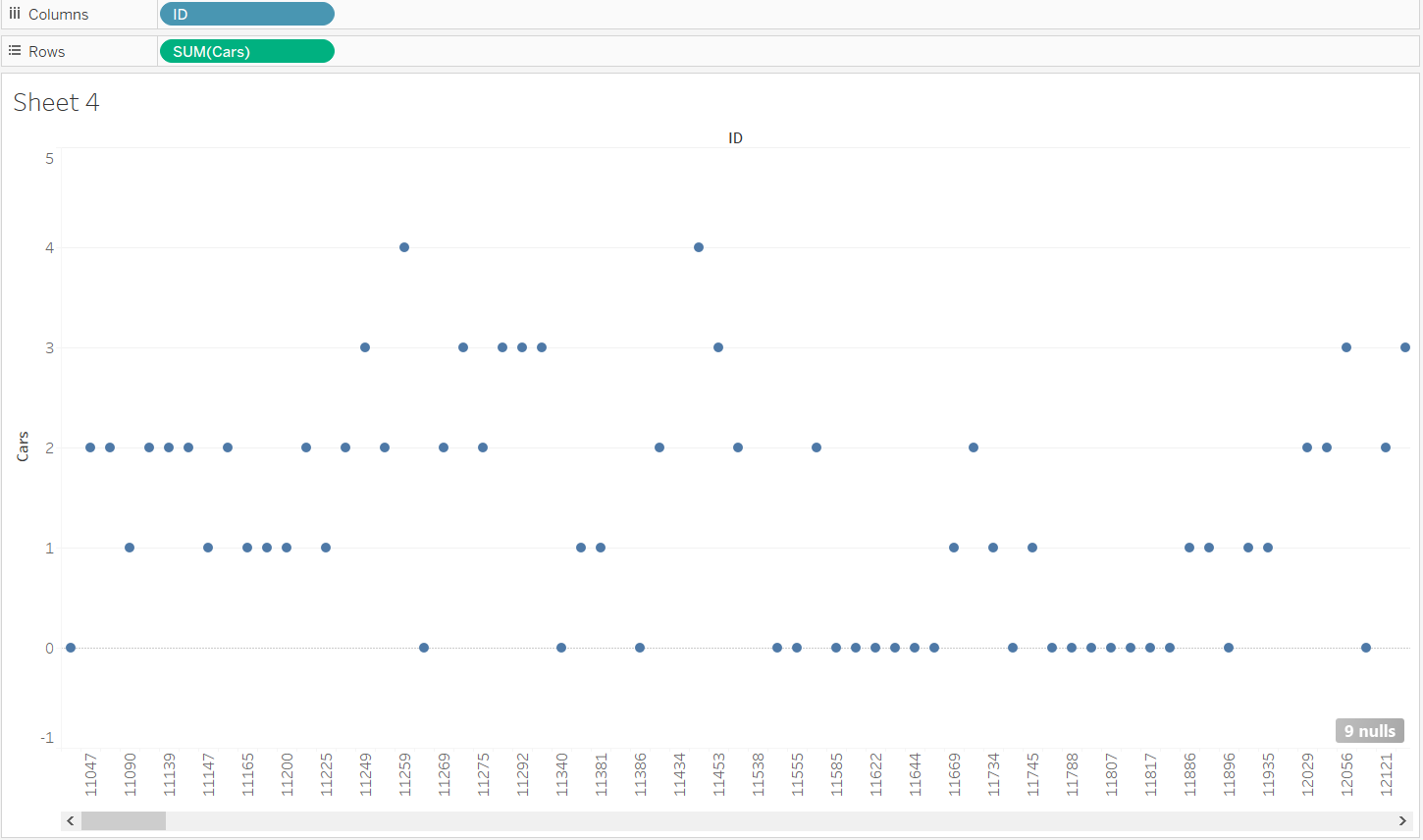
o Connect to Rserve from Tableau

o Create a chart as per requirement in Tableau





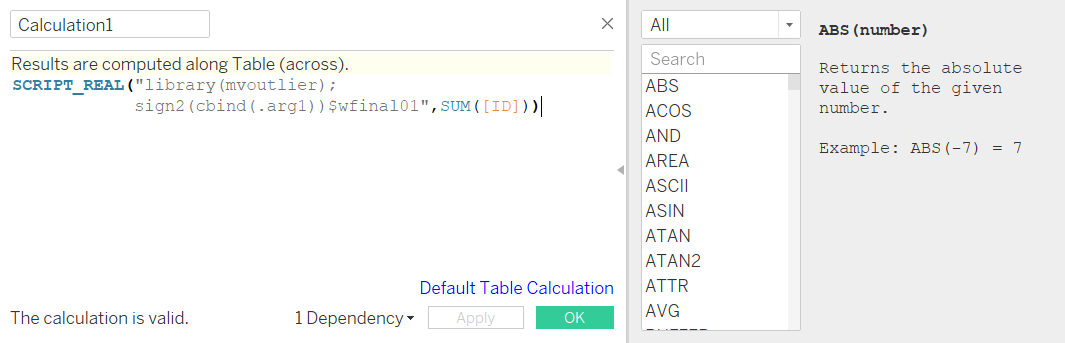




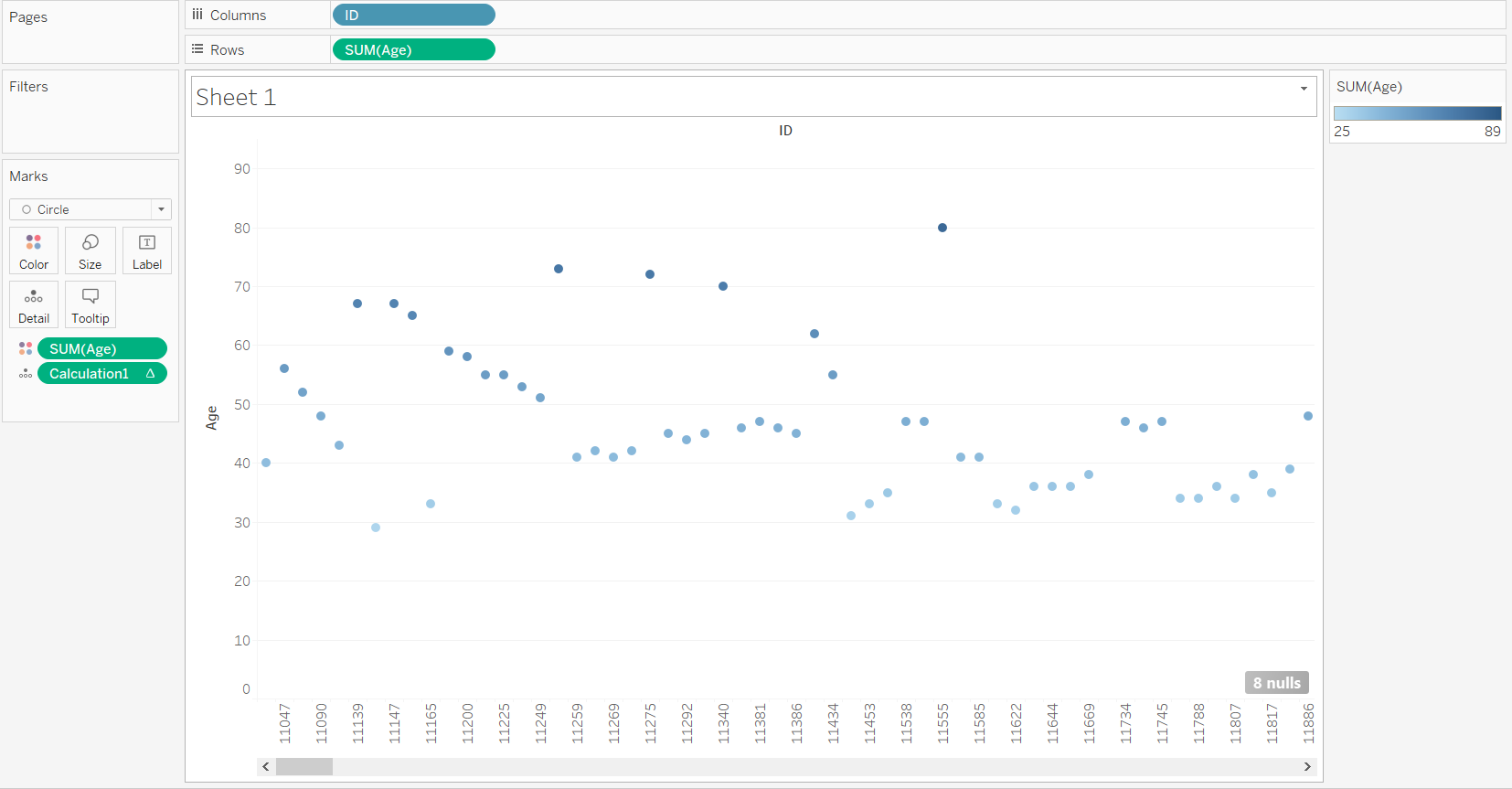
o Create a calculative field

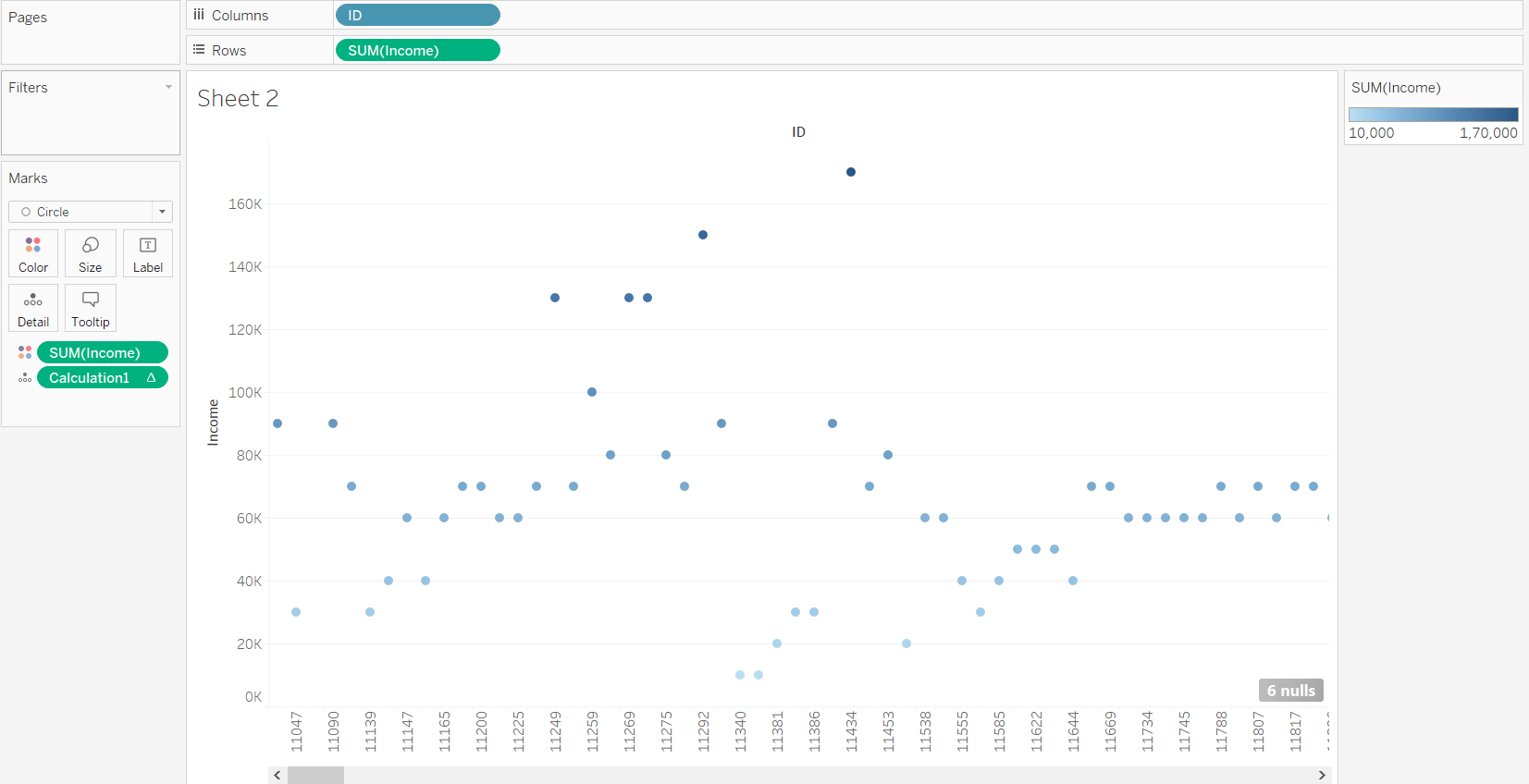
SCRIPT\_REAL("library(mvoutlier);

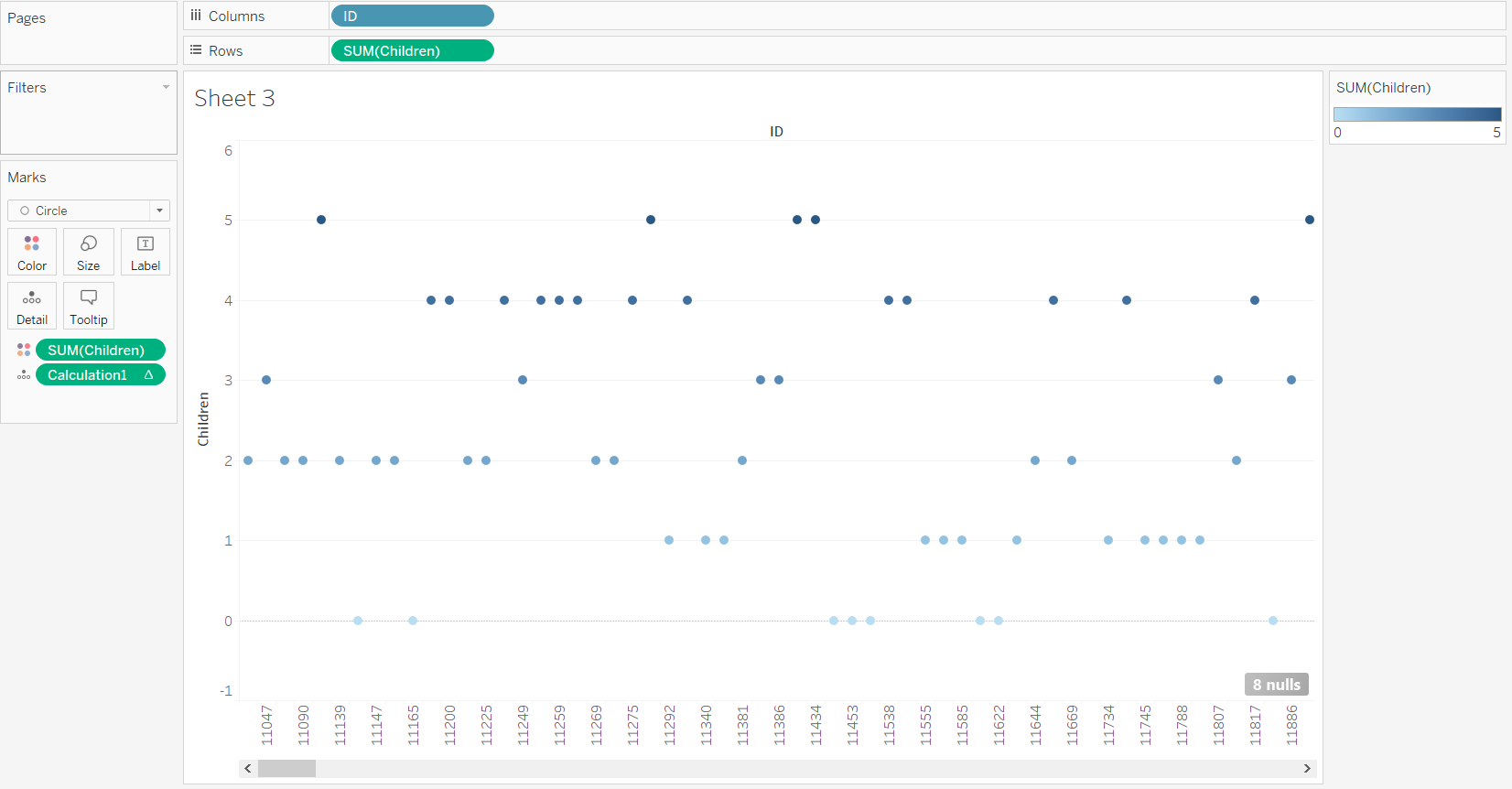
sign2(cbind(.arg1))$wfinal01",SUM([(field name]))

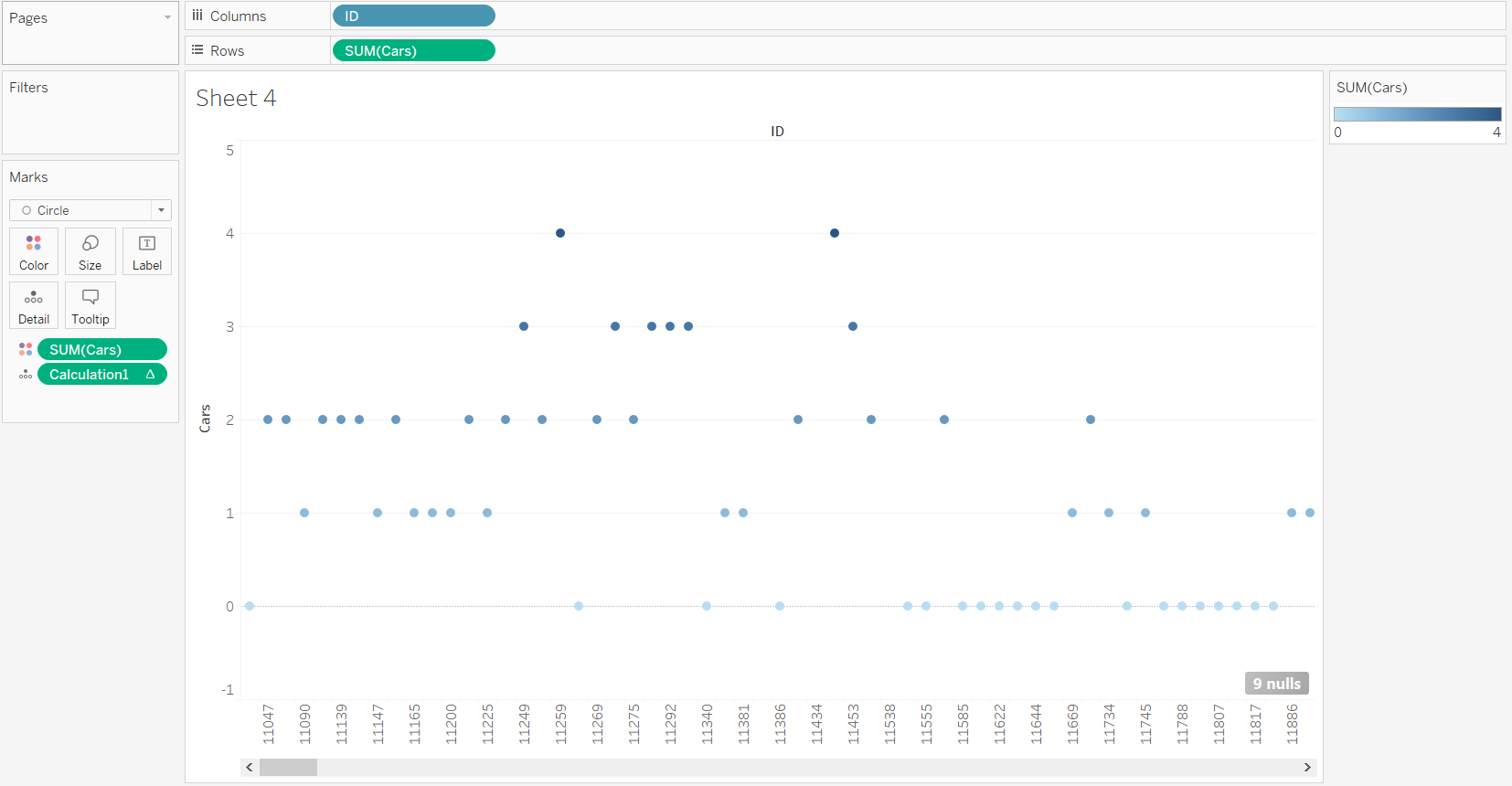
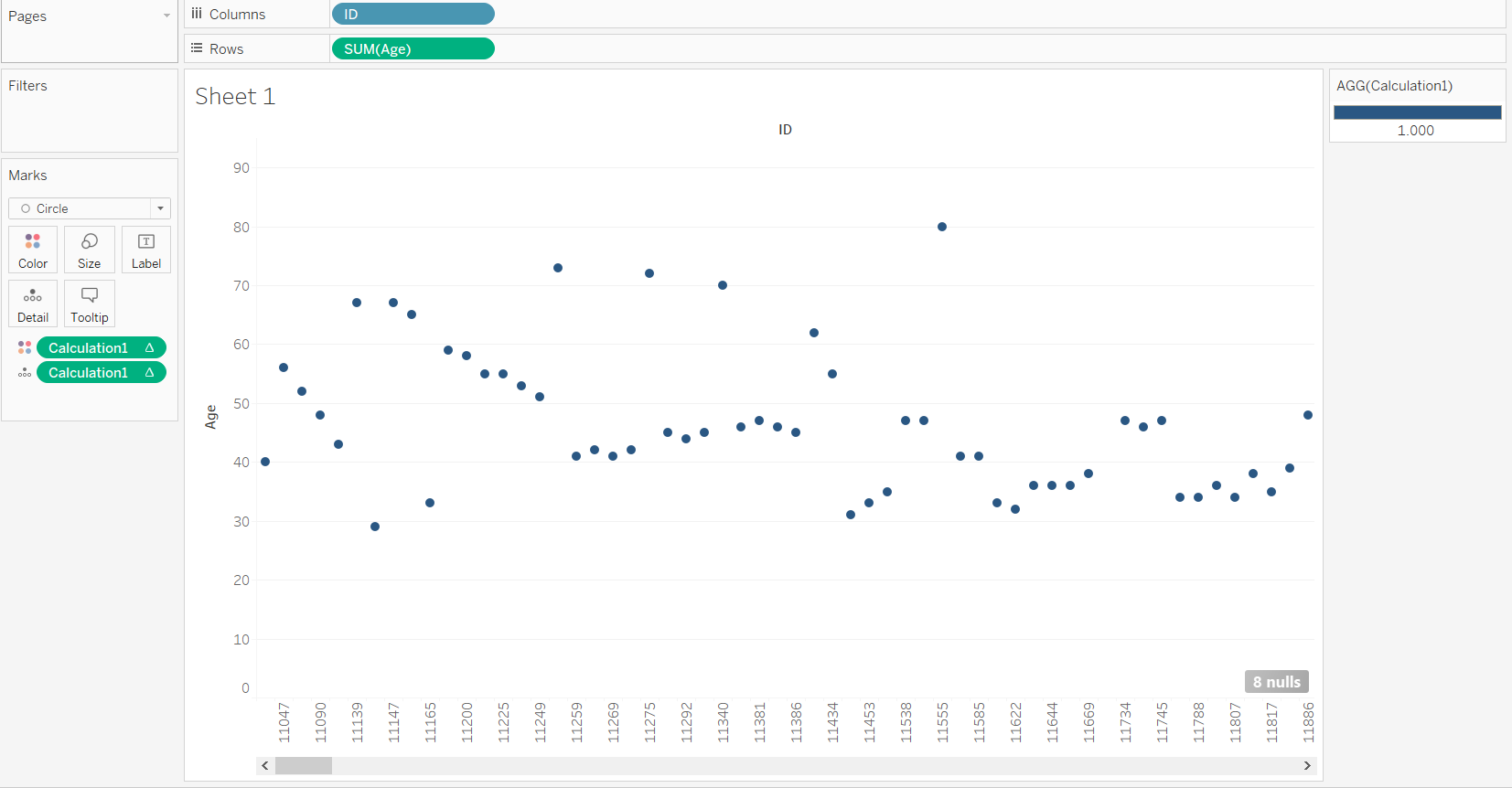
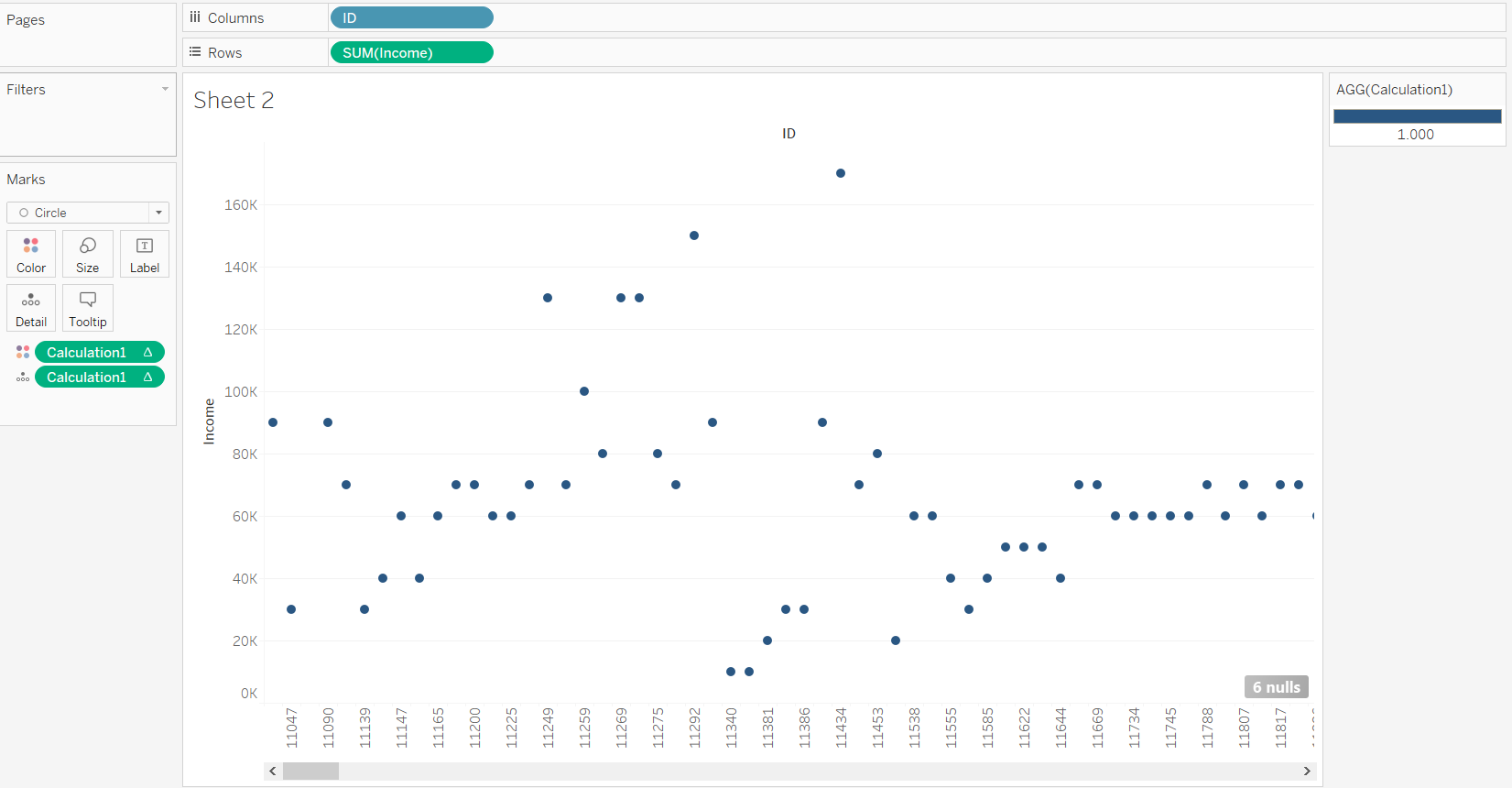
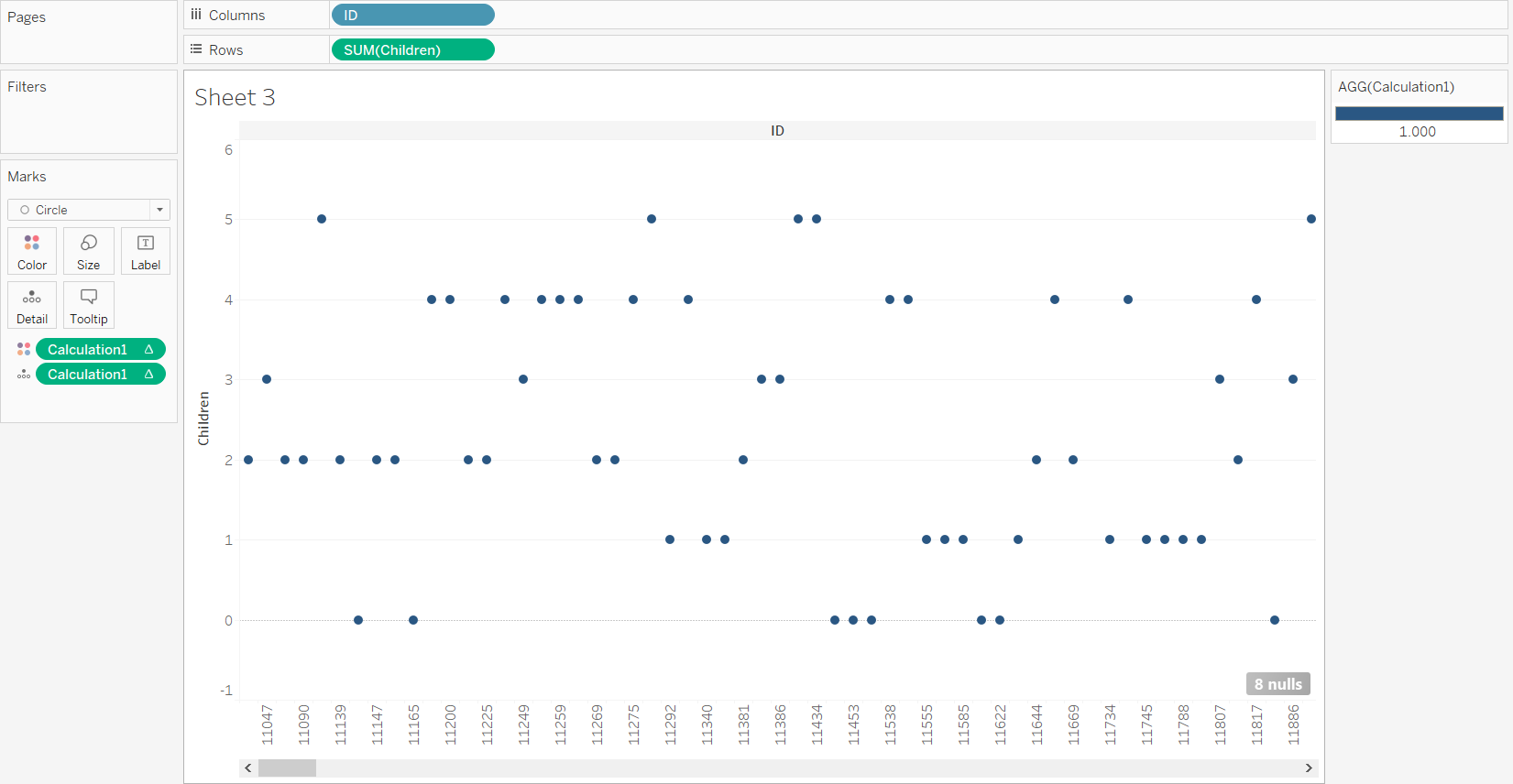
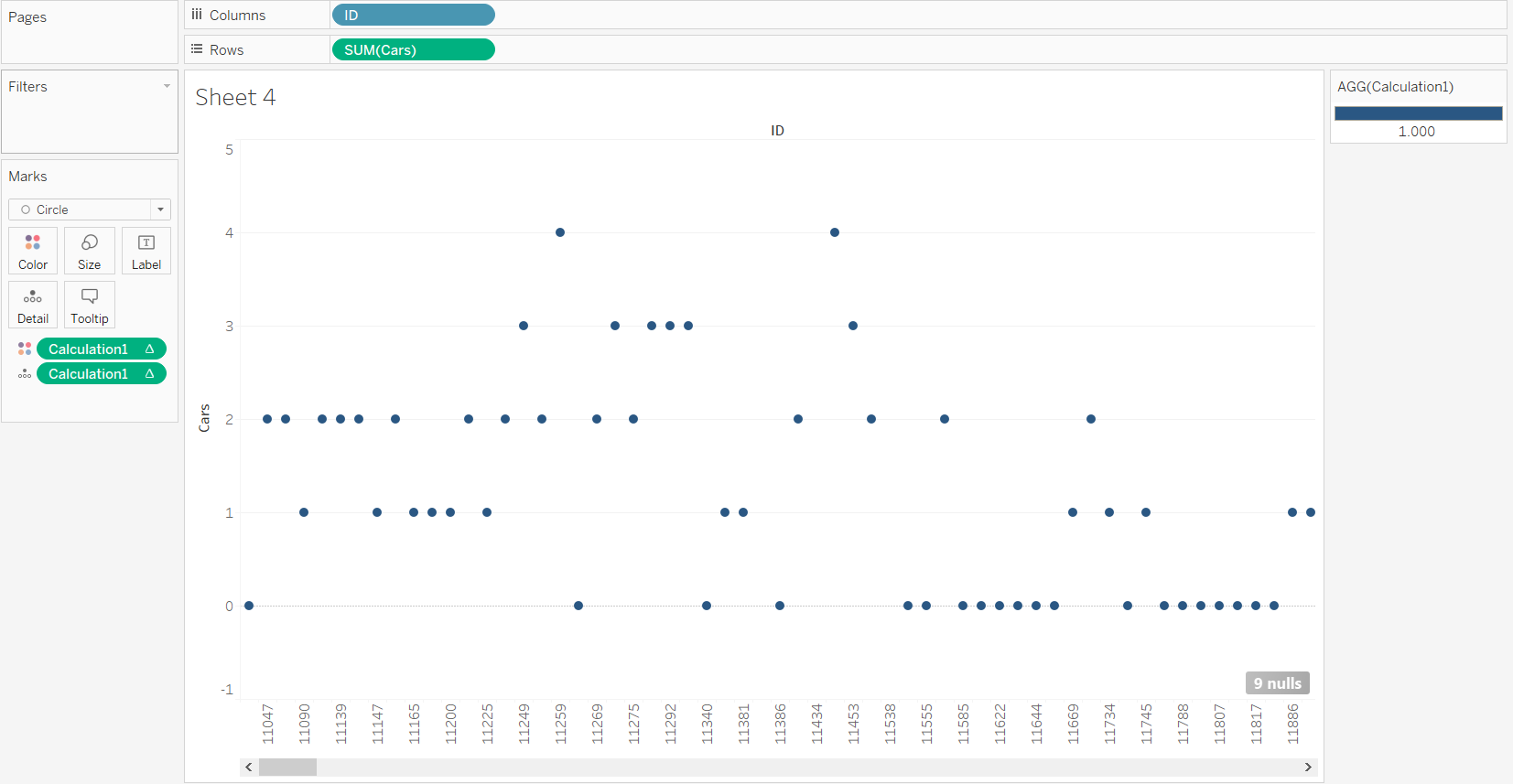


o Drop the field in color shelf to see the clusters







o Inference

Clustering is available in Tableau Desktop, but is not available for authoring on the web (Tableau Server, Tableau Online). Clustering is also not available when any of the following conditions apply:

* When you are using a cube (multidimensional) data source.
* When there is a blended dimension in the view.
* When there are no fields that can be used as variables (inputs) for clustering in the view.
* When there are no dimensions present in an aggregated view.

When any of those conditions apply, you will not be able to drag **Clusters** from the Analytics pane to the view.

In addition, the following field types cannot be used as variables (inputs) for clustering:

* Table calculations
* Blended calculations
* Ad-hoc calculations
* Generated latitude/longitude values
* Groups
* Sets
* Bins
* Parameters
* Dates
* Measure Names/Measure Values