



Accelerate Your Advanced Analytics R, Python & MATLAB

Advanced Analytics for Data
Scientists

Erik Polano

Associate Solution Architect

Erwin van Laar

Manager, Product Consulting

Agenda

Tableau's Built in Analytics

External Services Connection

Live Calculations

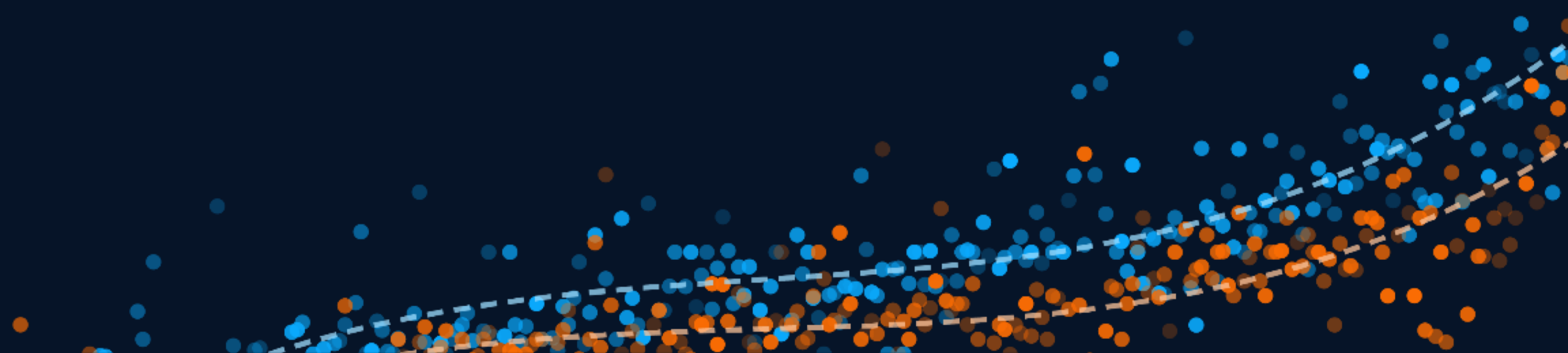
Materialized Calculations

Summary





Introduction

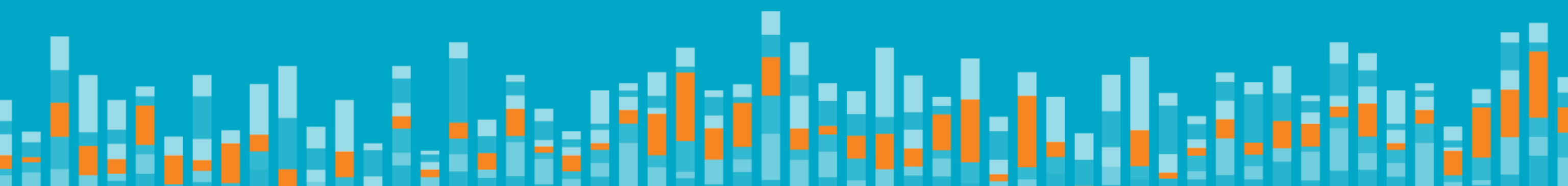


Erwin van Laar

Team Lead, Product Consultancy, Tableau

Erik Polano

Associate Solution Architect, Tableau



Objective

Create interactive dashboards...

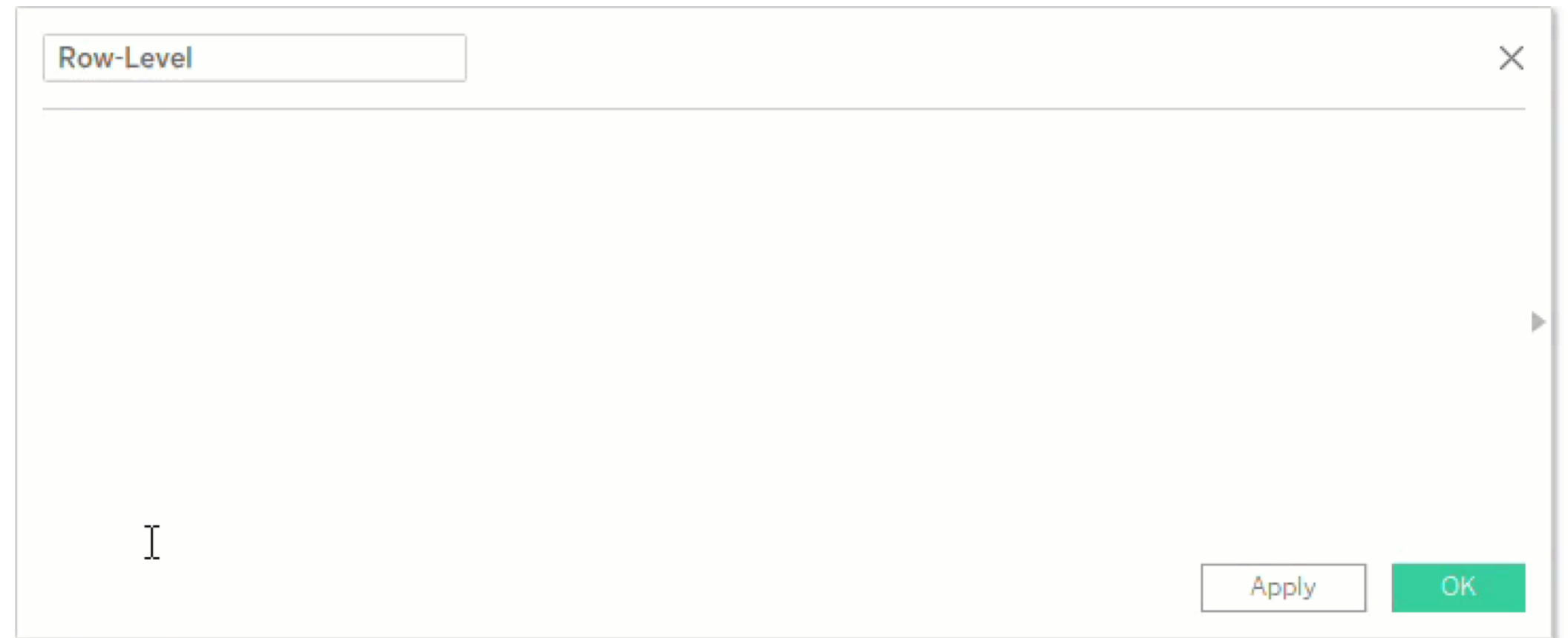
Without the end user needing to understand code...

That enables them to answer questions...

That are not contained in the underlying data source

Built In Analytics in Tableau

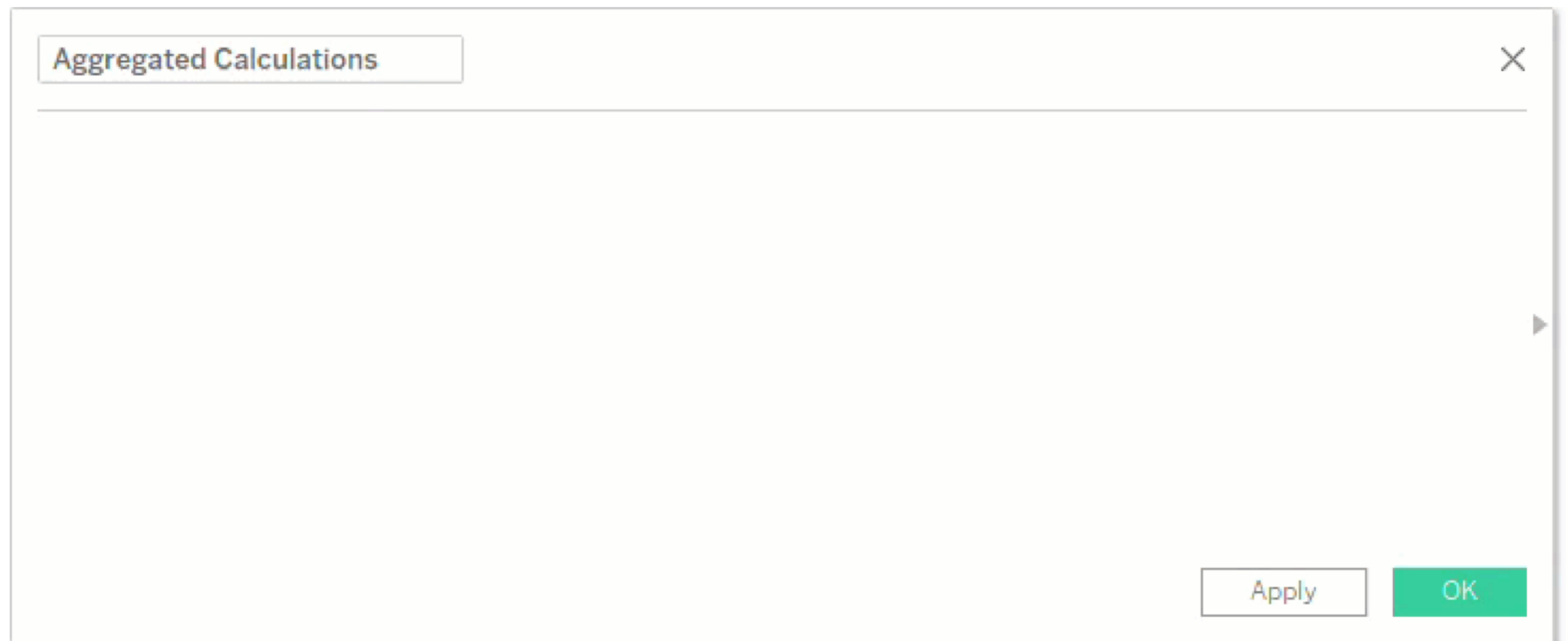
Row Level Calculations



Built In Analytics in Tableau

Row Level Calculations

Aggregated Calculations

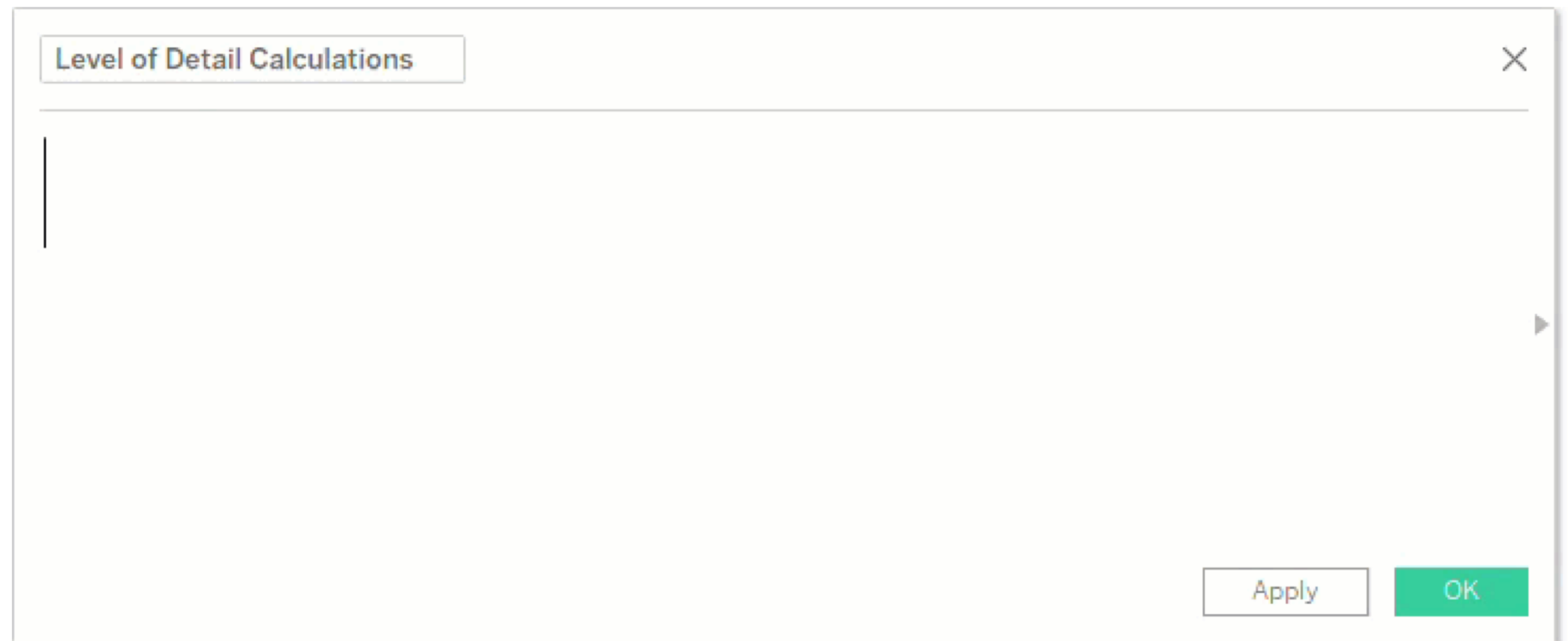


Built In Analytics in Tableau

Row Level Calculations

Aggregated Calculations

Level of Detail



Built In Analytics in Tableau

Row Level Calculations

Aggregated Calculations

Level of Detail

Table Calculations

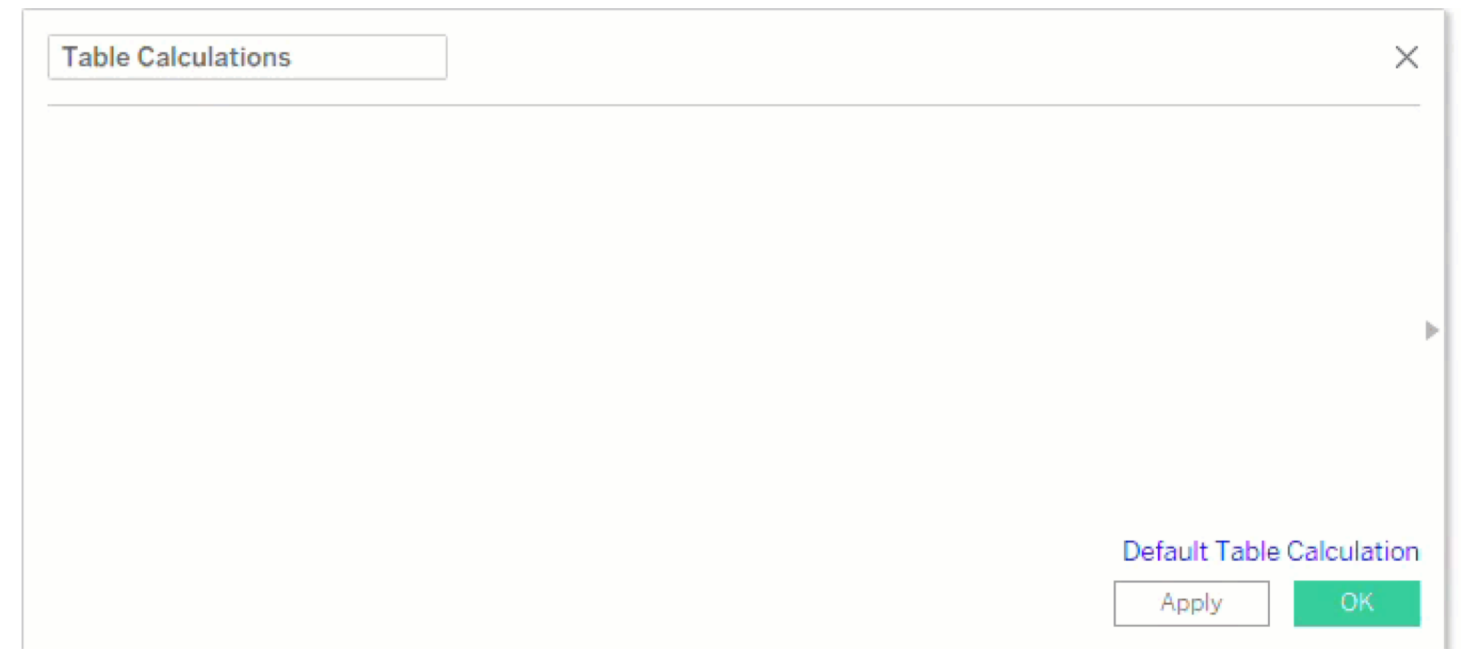
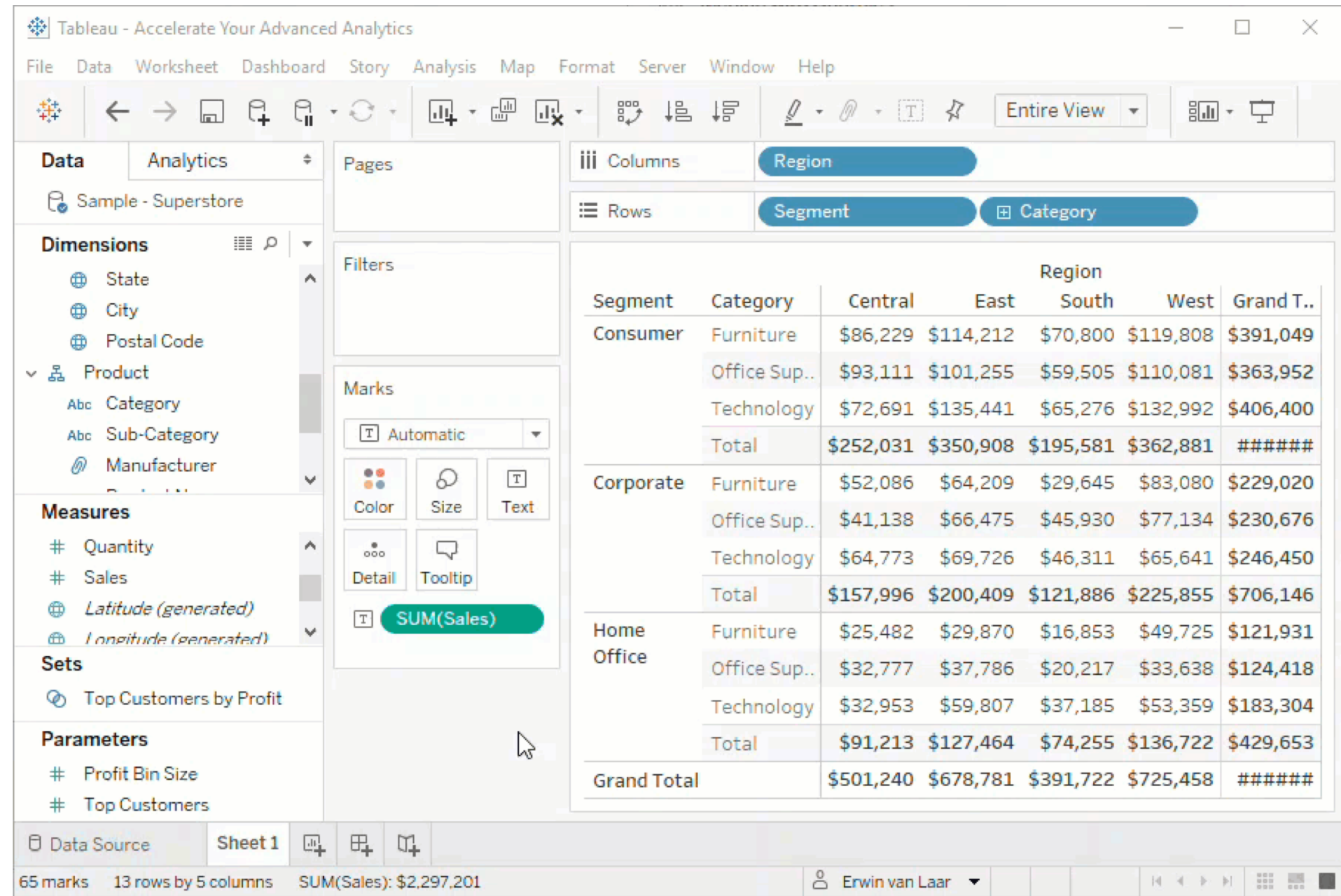


Tableau Calculations

Based on the
Level of
Aggregation in
the View



Built In Analytics in Tableau

Row Level Calculations

Aggregated Calculations

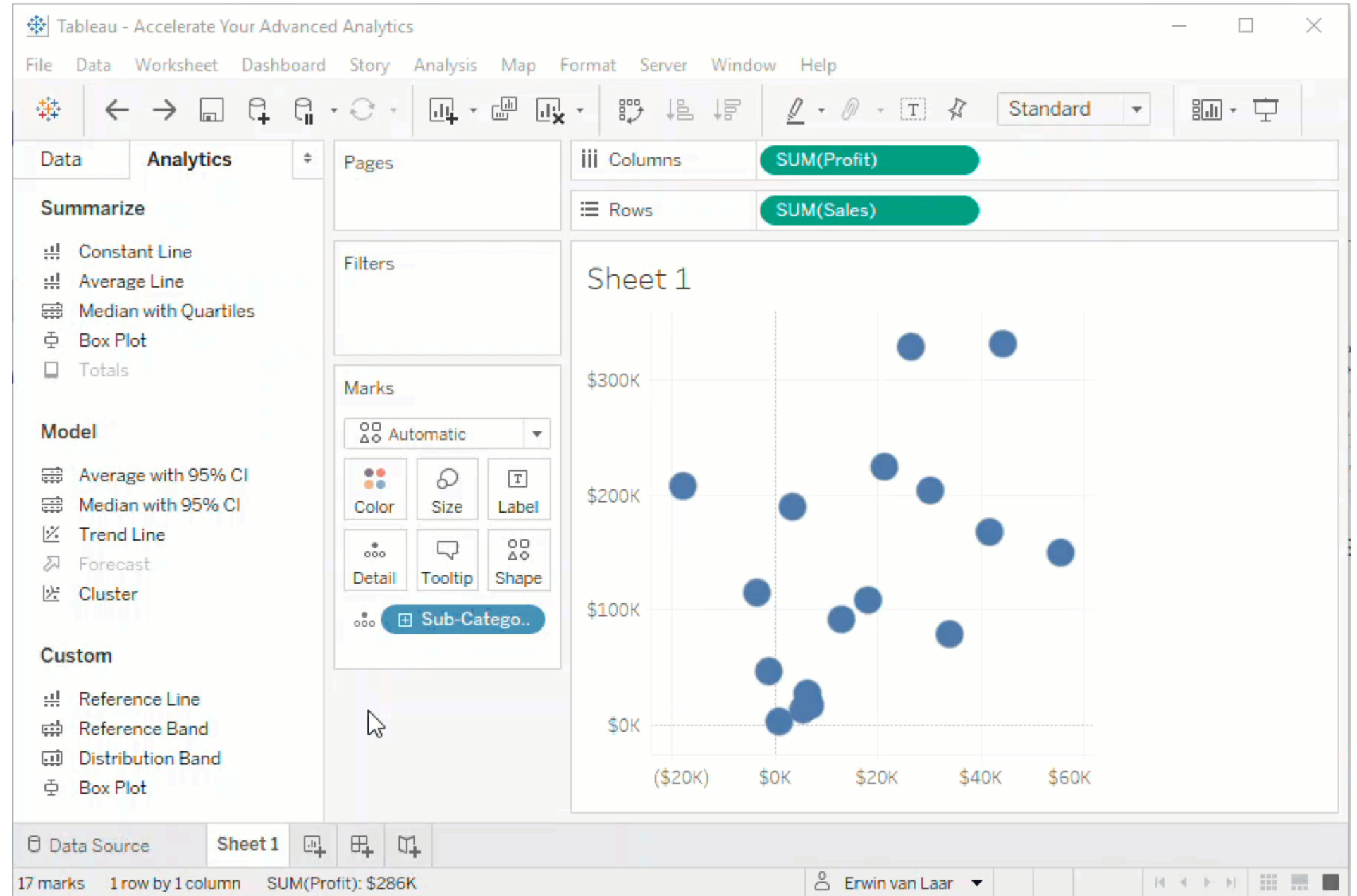
- Level of Detail Expressions

- Table Calculations

- Analytics Objects (ref lines, trend lines, forecast, clusters)

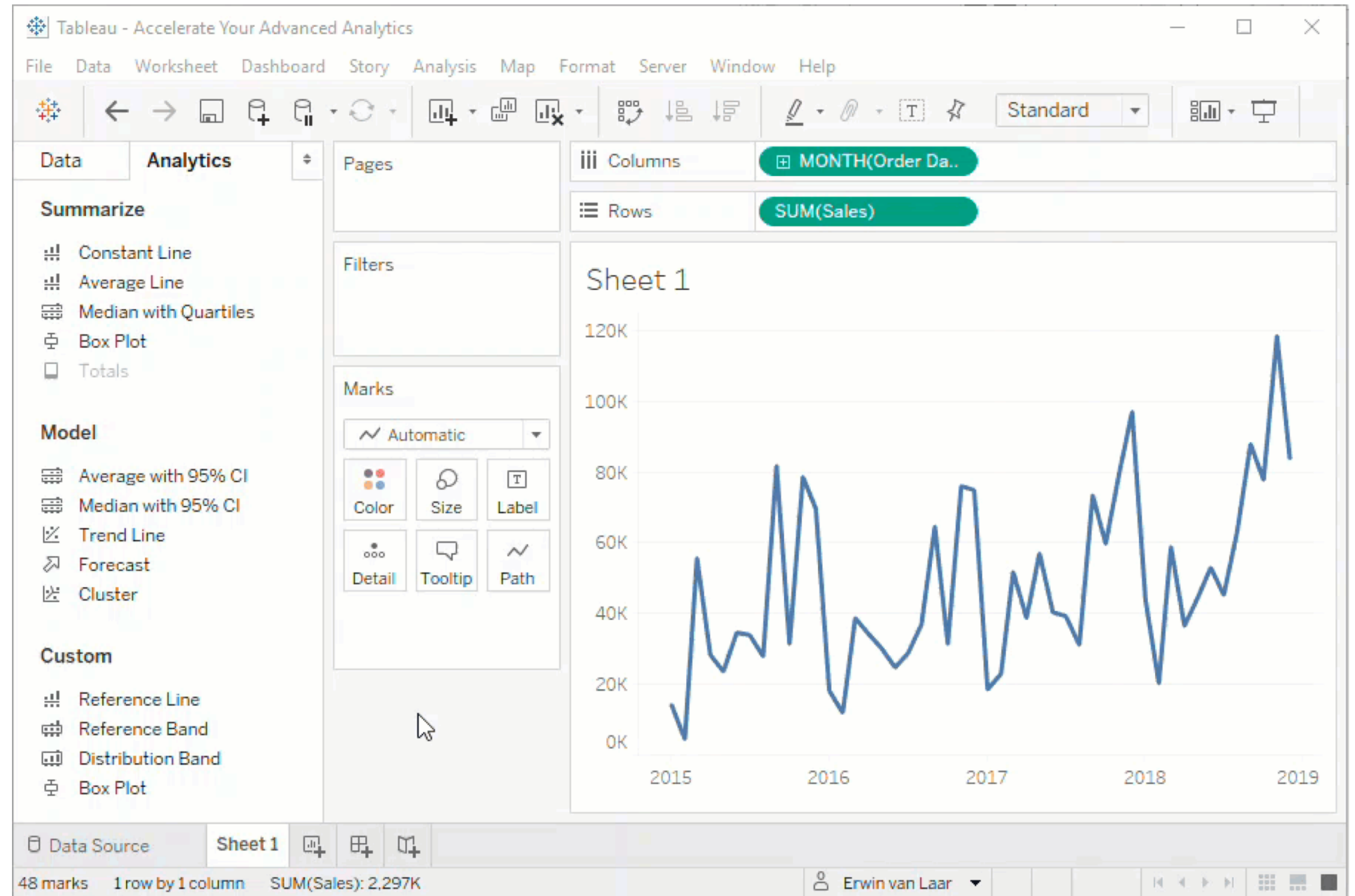
Analytic Objects

Drag & Drop
Analytical
features



Analytic Objects

Drag & Drop Analytical features



But What If I Want More?

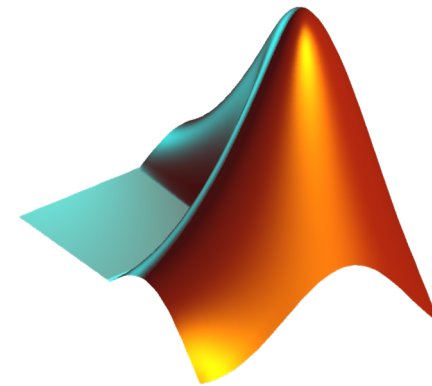
Different Model than Standard Tableau Functions

More complex models

Models trained on your data

Need more flexibility

External Services



Objective

Create interactive dashboards...

Without the end user needing to understand code...

That enables them to answer questions...

That are not contained in the underlying data source

By leveraging an Advanced Analytical Language



External Services

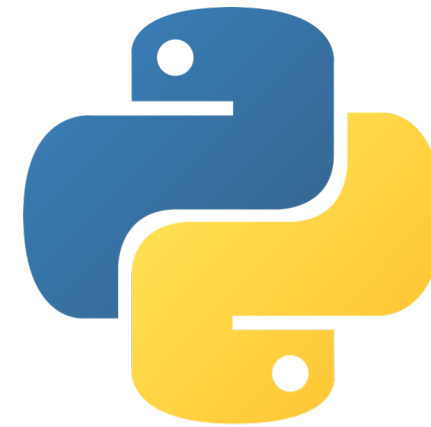


External Services Connection

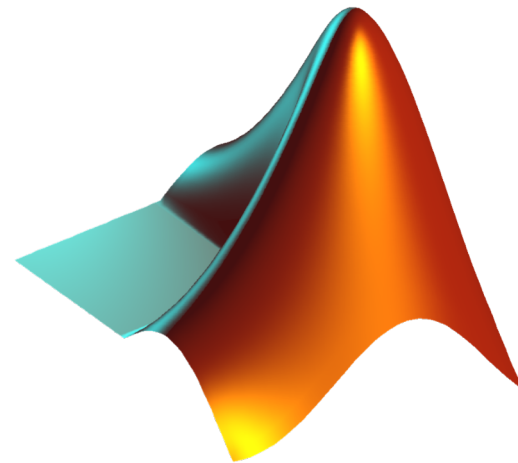
R (8.1+)



Python (10.1+)



MATLAB (10.4+)



SCRIPT Functions

SCRIPT_BOOL

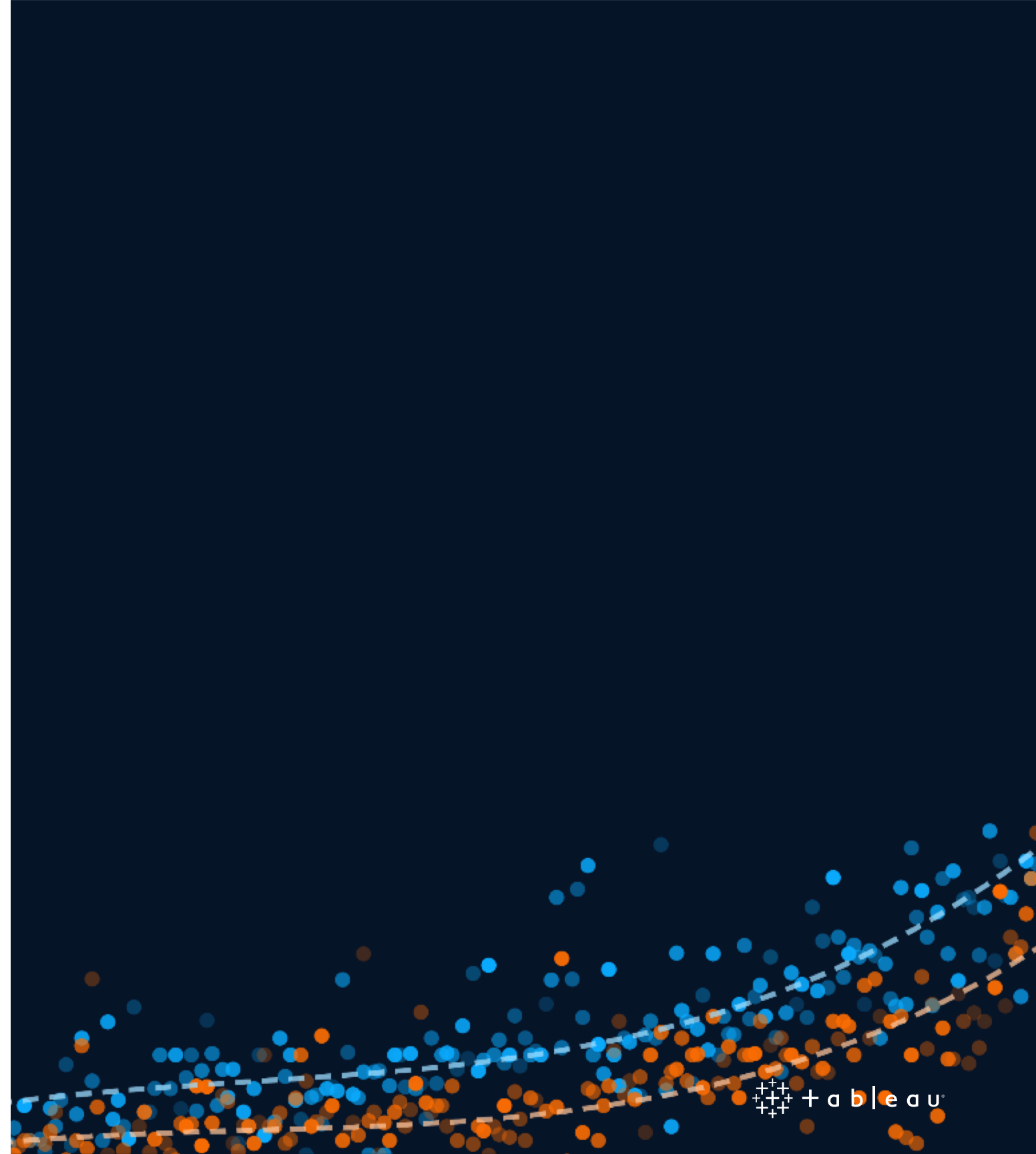
SCRIPT_INT

SCRIPT_REAL

SCRIPT_STR

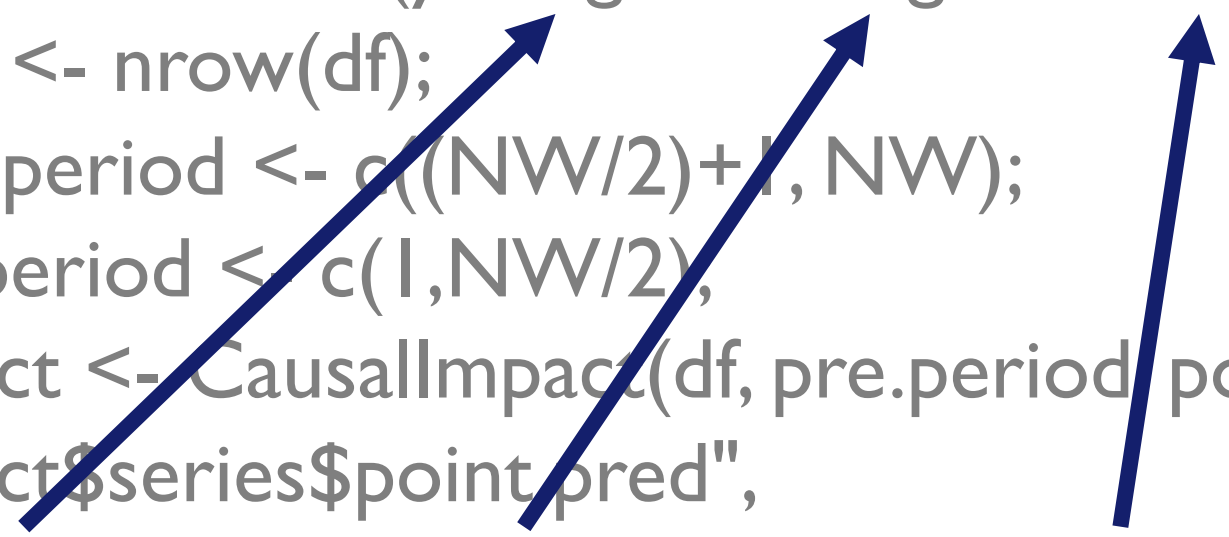
Script defines the type of outcome that Tableau expects

They will be run as a Table Calculation

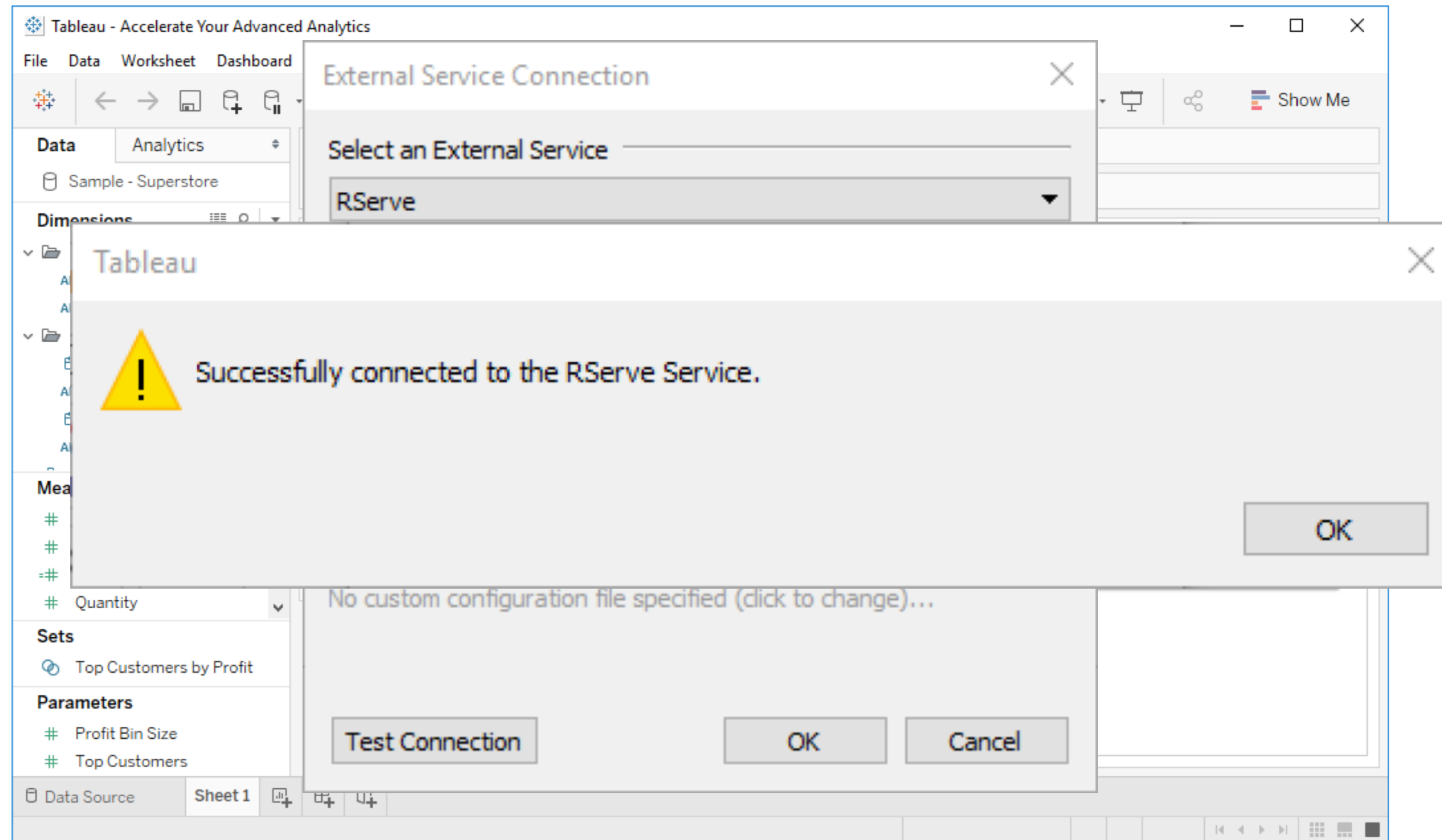


Example

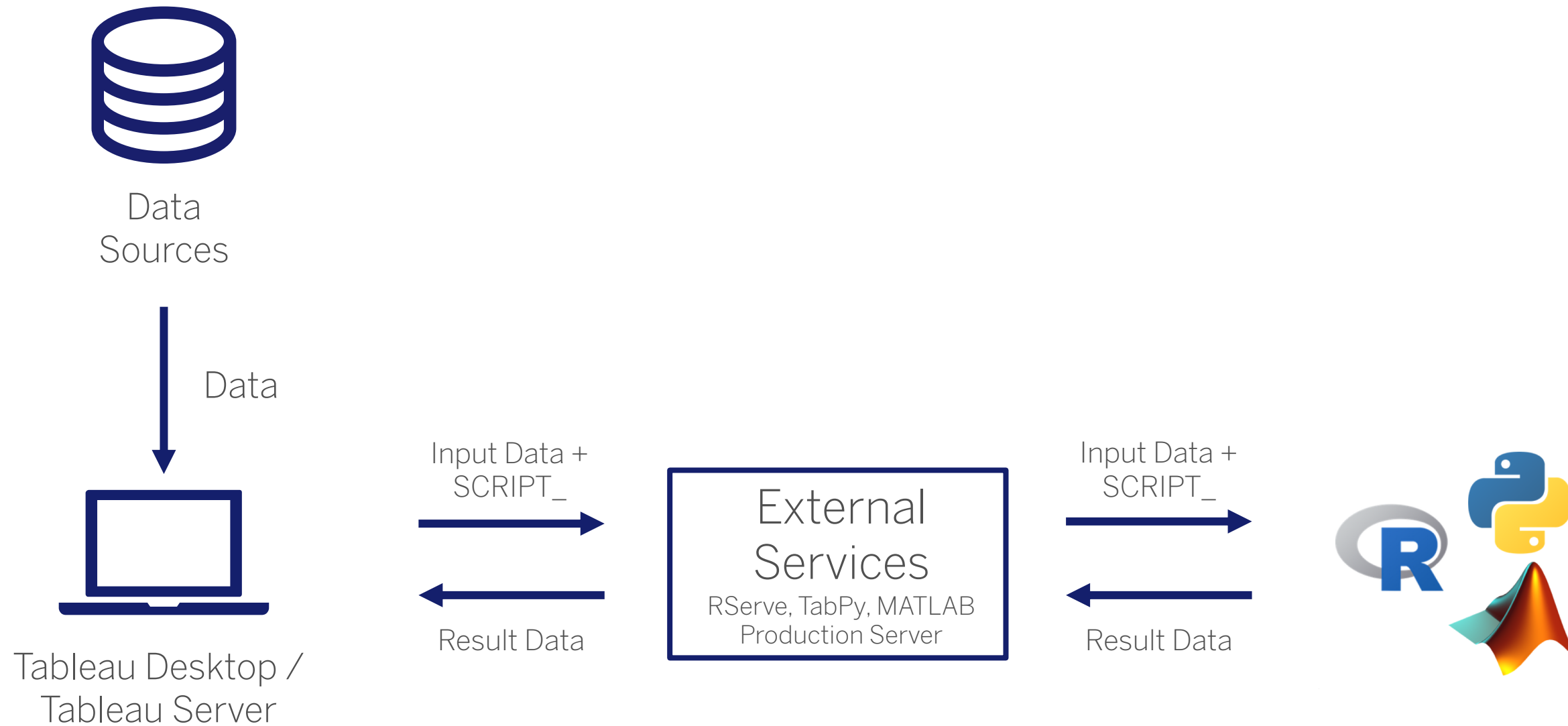
```
SCRIPT_REAL("library(CausallImpact);  
df <- data.frame(y=.arg1,x1=.arg2,x2=.arg3);  
NW <- nrow(df);  
post.period <- c((NW/2)+1, NW);  
pre.period <- c(1, NW/2);  
impact <- CausallImpact(df, pre.period, post.period);  
impact$series$point_pred",  
SUM([Profit]), SUM([Quantity]), SUM([Sales]))
```



Manage External Services Connection



Connecting Tableau to an External Service





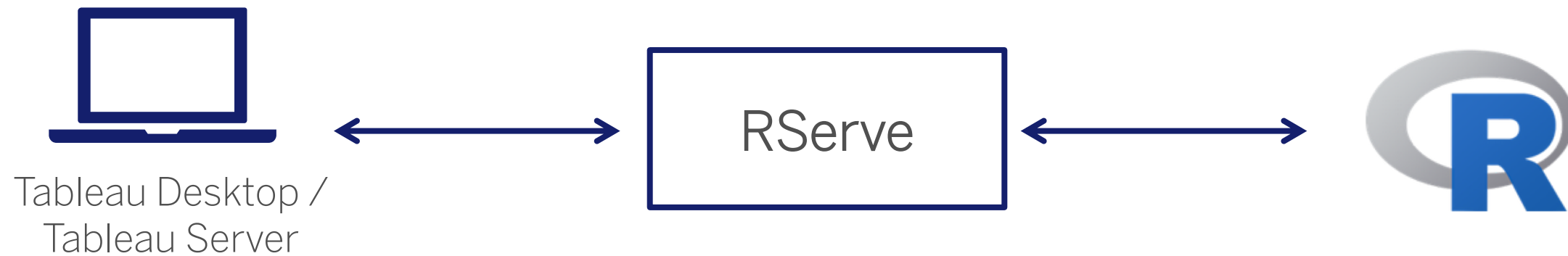
R



R - RServe

TCP/IP server

Allows other programs to use facilities of R



Examples

Tableau helps you visually identify pattern in data

But are those patterns significantly different?

→ T-Test



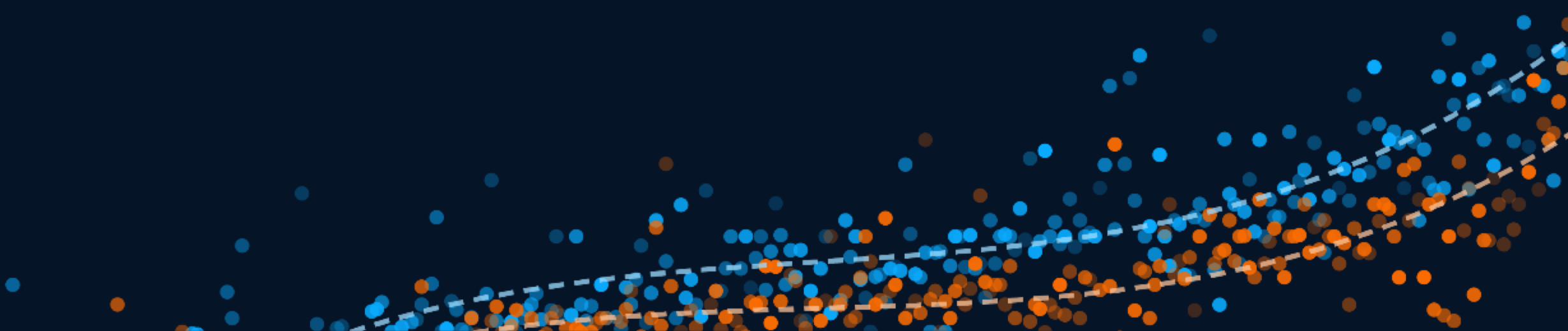
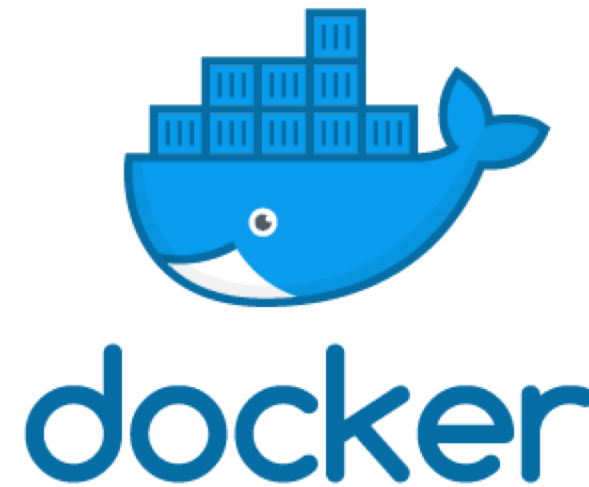
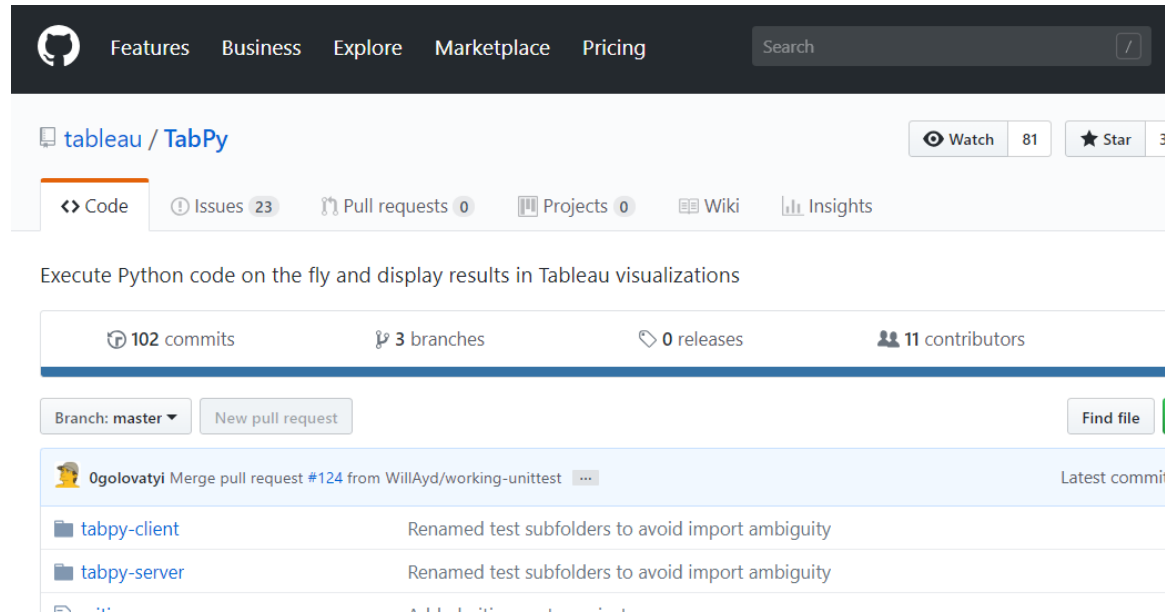


Tableau Python Server (TabPy)

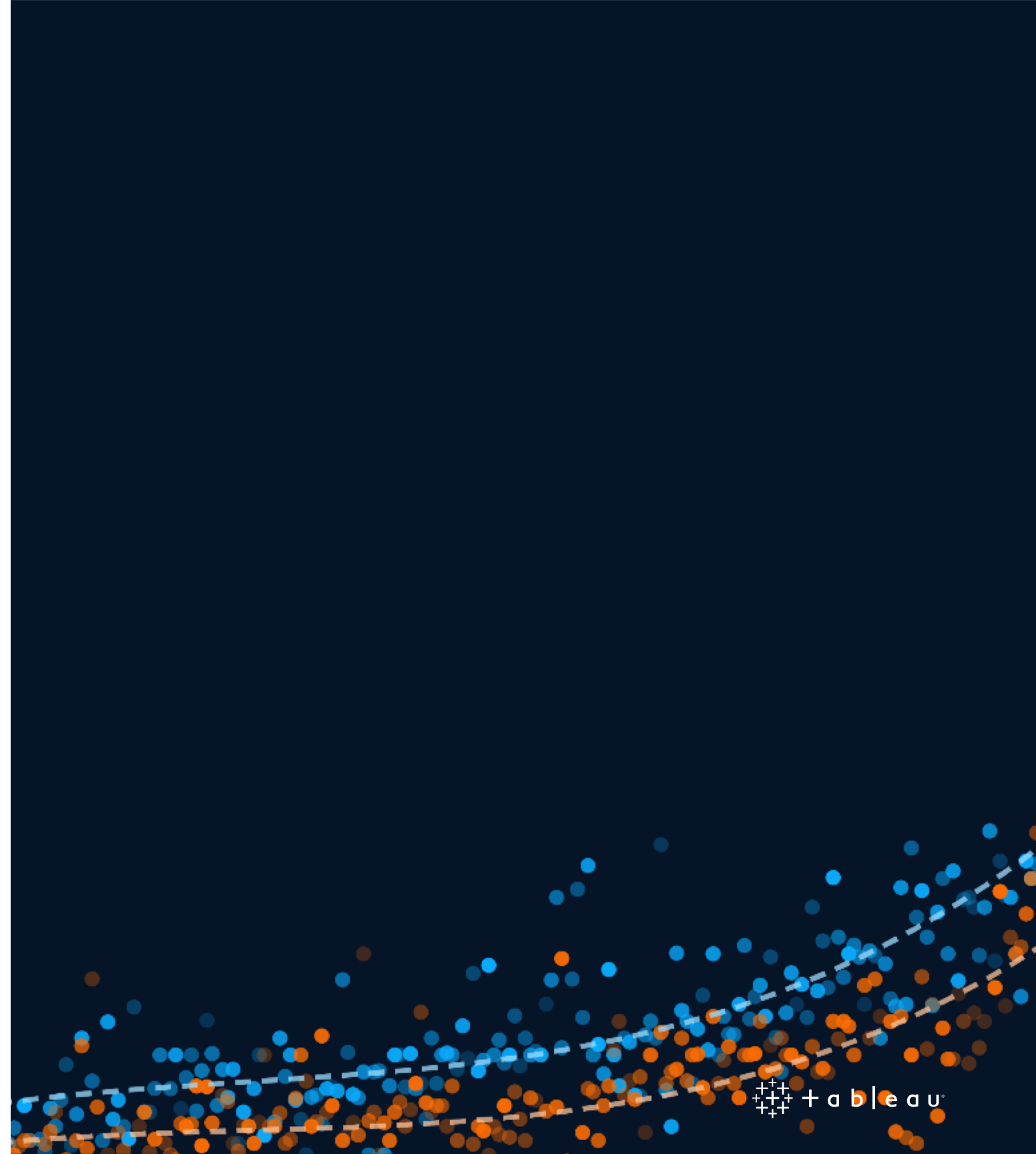


numpy (1.11.2)
pandas (0.19.1)
scikit-learn (0.17.1)
scipy (0.18.1)
textblob (0.11.1)
nltk (3.2.1)
vaderSentiment (0.5)
geopy (1.11.0)
requests (2.12.4)
reverse_geocoder (1.5.1)



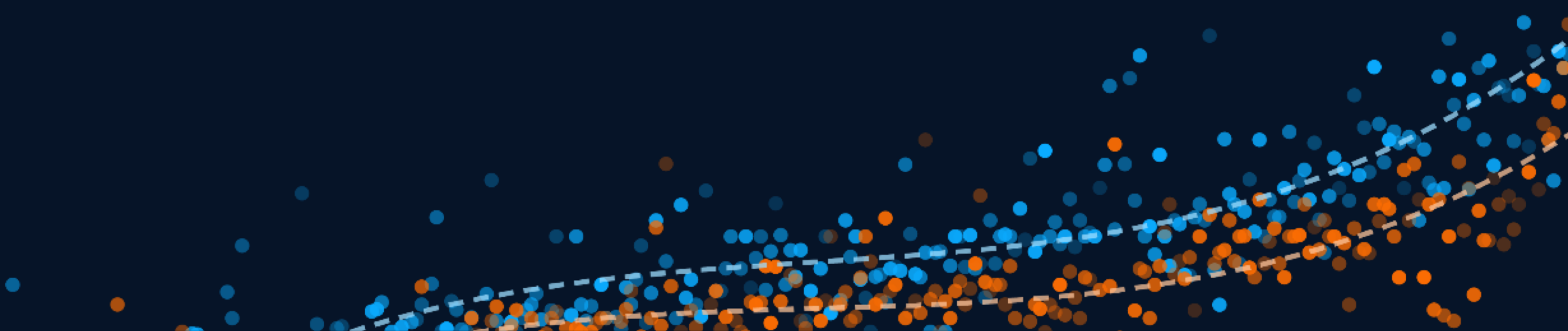
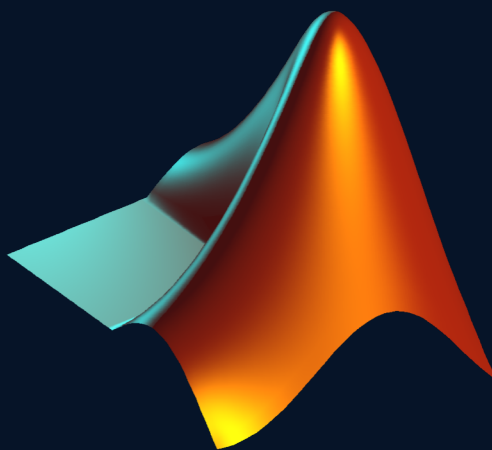
Python - Example

Twitter Sentiment Analysis





MATLAB



MATLAB – Two Different Options

MATLAB allows to connect to models in two different ways:

External Service Connection

Web Data Connector

Option 1: External Service Connection

Define Functions in MATLAB Production Server

Call Functions through Calculated Fields

Possible to filter and re-run calculations on different levels of aggregation

Advantage: Flexibility.

The image shows a composite screenshot illustrating the integration of MATLAB and Tableau. On the left, the MATLAB R2017a editor displays a script for solving a Traveling Salesman Problem (TSP) using linear programming. The script includes comments and code for allocating sparse matrices, setting constraints, and solving the problem. Below the editor, the MATLAB Compiler interface is visible, showing options for packaging the MATLAB program for deployment as standalone applications, Hadoop MapReduce programs, shared libraries, or MATLAB Production Server components. The command window shows the execution of the 'deploytool' command.

On the right, the Tableau interface displays a map visualization titled "Shortest Flight Path". The map shows a network of cities across the United States, with green lines connecting them to represent the shortest path. The Tableau interface includes a Columns shelf with "AVG(Longitude)" and a Rows shelf with "AVG(Latitude)". A tooltip for the "ShortestPath" calculated field is visible, showing the underlying calculation: `SCRIPT_INT('TSP/GetLatLongVector', AVG([Longitude]), AVG([Latitude]))`. The tooltip also indicates that the calculation is valid and shows the number of sheets affected.

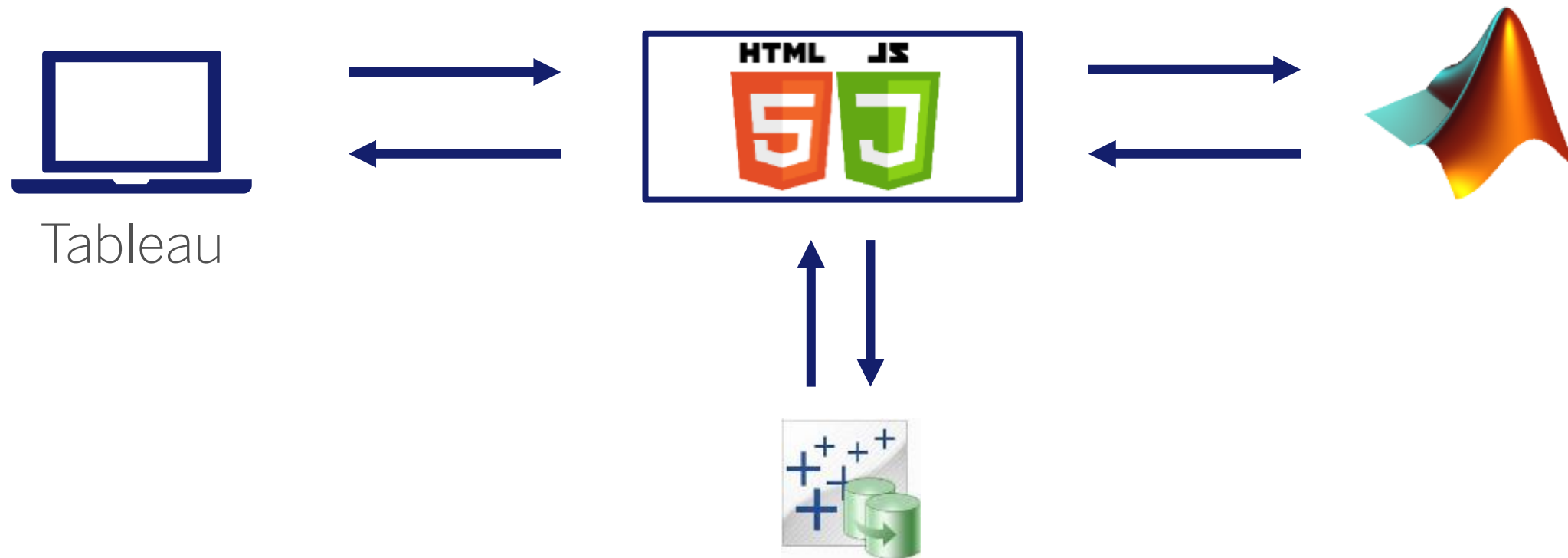
Option 2: Web Data Connector

Request Data Source at Initial Data Source Stage

Request sent once to MATLAB and all results written on row level into extract

Advantage: Performance

Advantage: All data can be used regardless of aggregation level



MATLAB - Example

Shortest Salesperson Route

The image displays a workflow involving MATLAB and Tableau. On the left, the MATLAB R2017a editor shows a script for solving a Traveling Salesman Problem (TSP) using linear programming. The script includes comments and code for allocating sparse matrices, adding constraints, and solving the problem using `intlinprog`.

```
%% Linear programming solution with subtour  
A = spalloc(0, lendist, 0); % Allocate a sparse matrix  
b = [];  
while numtours > 1 % repeat until there is only one tour  
    % Add the subtour constraints  
    b = [b; zeros(numtours, 1)];  
    A = [A; spalloc(numtours, lendist, nStops)];  
    for stopCount = 1:numtours  
        rowIdx = size(A, 1)+1;  
        % ...  
    end  
end  
  
% Optimize using integer linear programming  
[x_tsp, ~, ~, ~] = intlinprog(dist, intcon, ...
```

On the right, the Tableau interface shows a map titled "Shortest Flight Path" over the United States. The map displays a route connecting 28 numbered cities. The Tableau interface includes a Columns shelf with `AVG(Longitude)` and a Rows shelf with `AVG(Latitude)`. A tooltip for the "ShortestPath" field shows the calculation: `SCRIPT_INT('TSP/GetLatLongVector', AVG([Longitude]), AVG([Latitude]))`. The tooltip also indicates that the calculation is valid and shows the number of sheets affected.



Materialized Calculations



Categorization

Twitter Sentiment Analysis

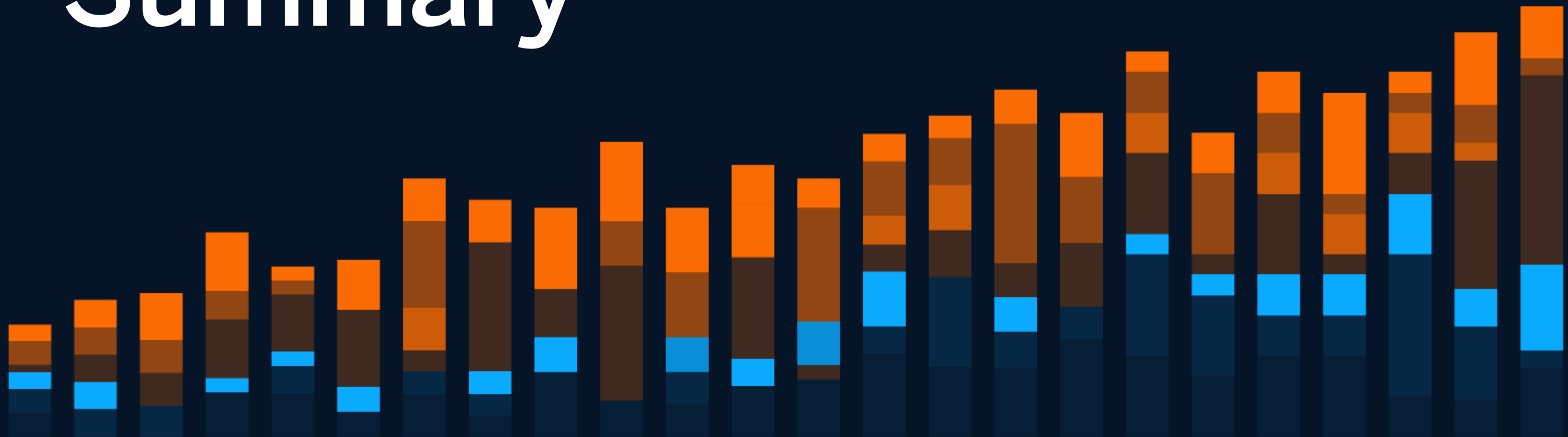
Customer Identification

Reclassification

Row-Level Calculations



Summary



What Do External Services Bring to Tableau?

Deeper Statistics

Machine Learning

Productionalize Predictive Modeling

More Expressive Power

What Does Tableau Bring to External Services?

Data Connectivity

Explore at the Speed of Thought

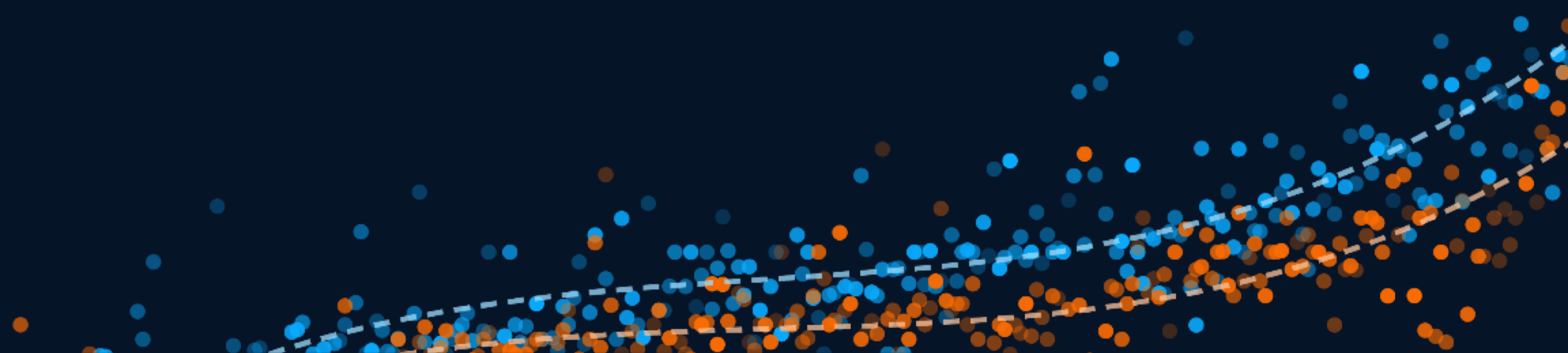
Interactive Visual Storytelling

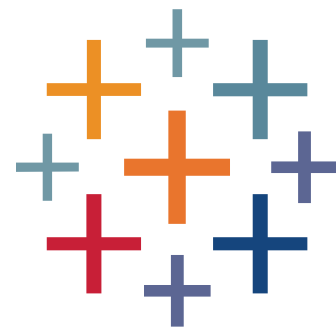
Any Questions?



THANK YOU!

Erik Polano & Erwin van Laar





+ a b l e a u[®]