SPSS **Modeler** *18.2.2*

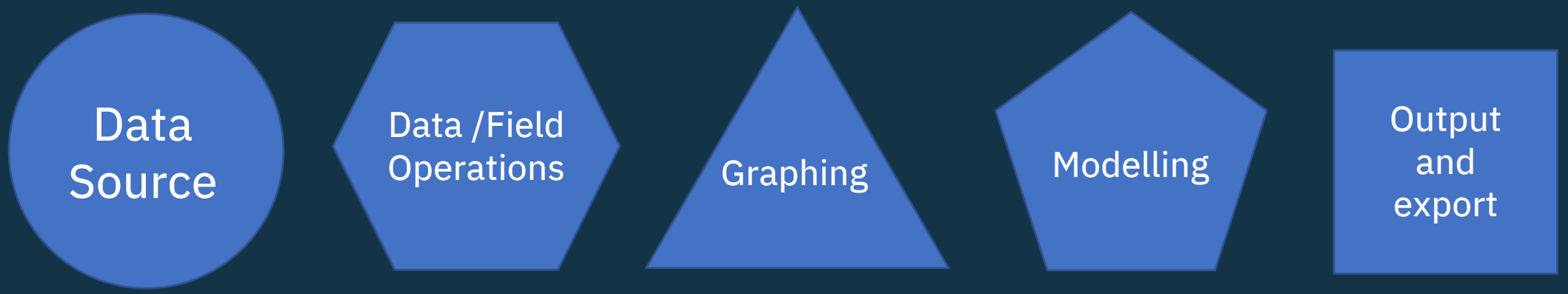
Download sample files from:
github.com/krishnac7/spssSamples

Krishna Balaga
krbalaga@in.ibm.com

A Quick Recap

Looks too complex to
use but it **really is not**

The **shape** of the Node tells you what it does



Download sample files from:
github.com/krishnac7/spssSamples

Creating connections

Middle click on the Node and
draw a connection to next node

Click on the node and
press F2 on your keyboard

Common Data import nodes

Used for .*sav* files which is most common format in spss

Used for other text encoded files like .*csv* .*xls*

Roles

Input

Features used to make the prediction

Target

Label we are trying to predict

None

Values to be Ignored

Measurement

Continuous

Numerical
Eg:
1,2,3,4,5....

Categorical

Contains groups
Eg:
Cat,Dog
Class1,class2

Ordinal

Categories with ranking
Eg:
Rating 1,2,3
Bad, Good, Excellent

Nominal

Categories without ranking
Eg:
Delhi, Mumbai, Pune
Monday,Tuesday

Flag

Has only 2 values
Eg:
Yes, No
True, False

Typeless

Values to be Ignored

Classification

Classification is when we predict a **Label**

Eg:

- *Is this email spam or not?*
- *Is this picture of a Dog or a Cat?*
- *Will this customer purchase my premium plan or not?*

Regression

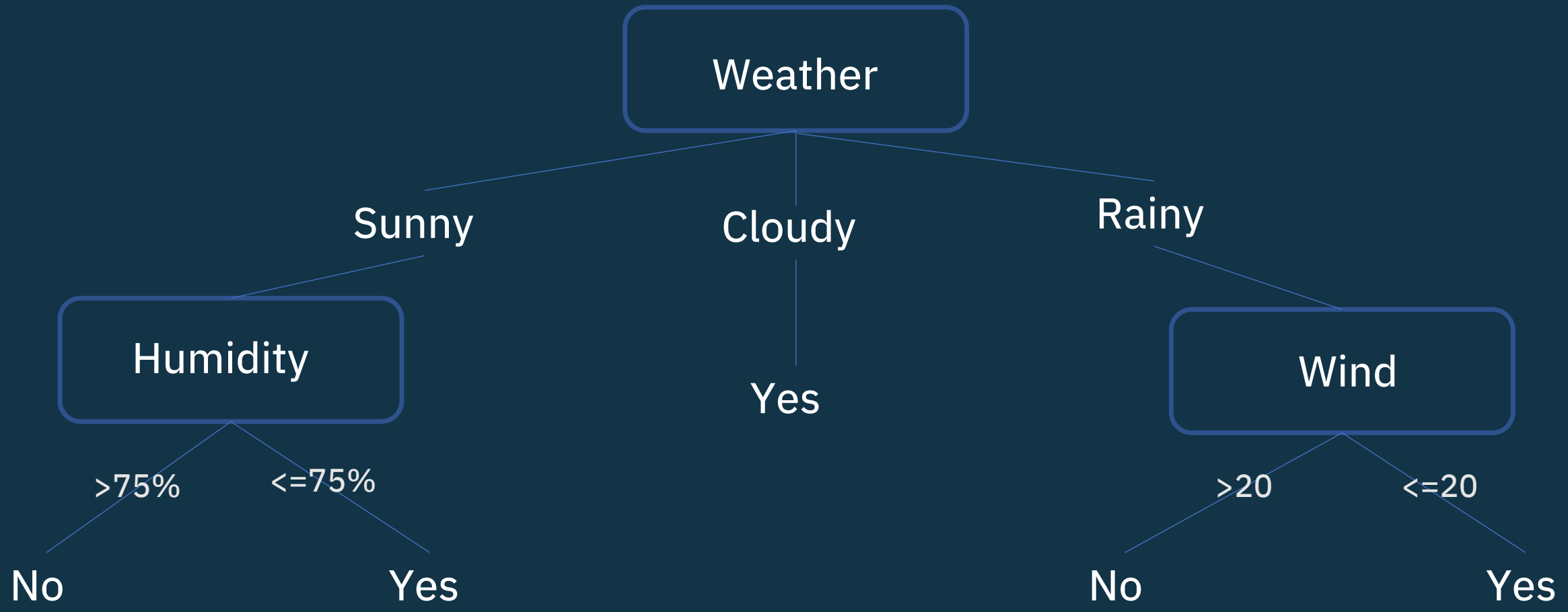
Regression is when we predict a **Quantity**

Eg:

- *What can be the price of this stock by tomorrow?*
- *What can be the temperature tomorrow at 6 am?*
- *What can be the price of this house with 3 bedrooms if I know the price of a similar flat with 2 bedrooms*

Decision Trees

- Decision tree algorithm can be used for solving **Regression** and **Classification** problems too.



Ensemble Models

Ensemble models are combination of several base models in order to produce one optimal predictive model

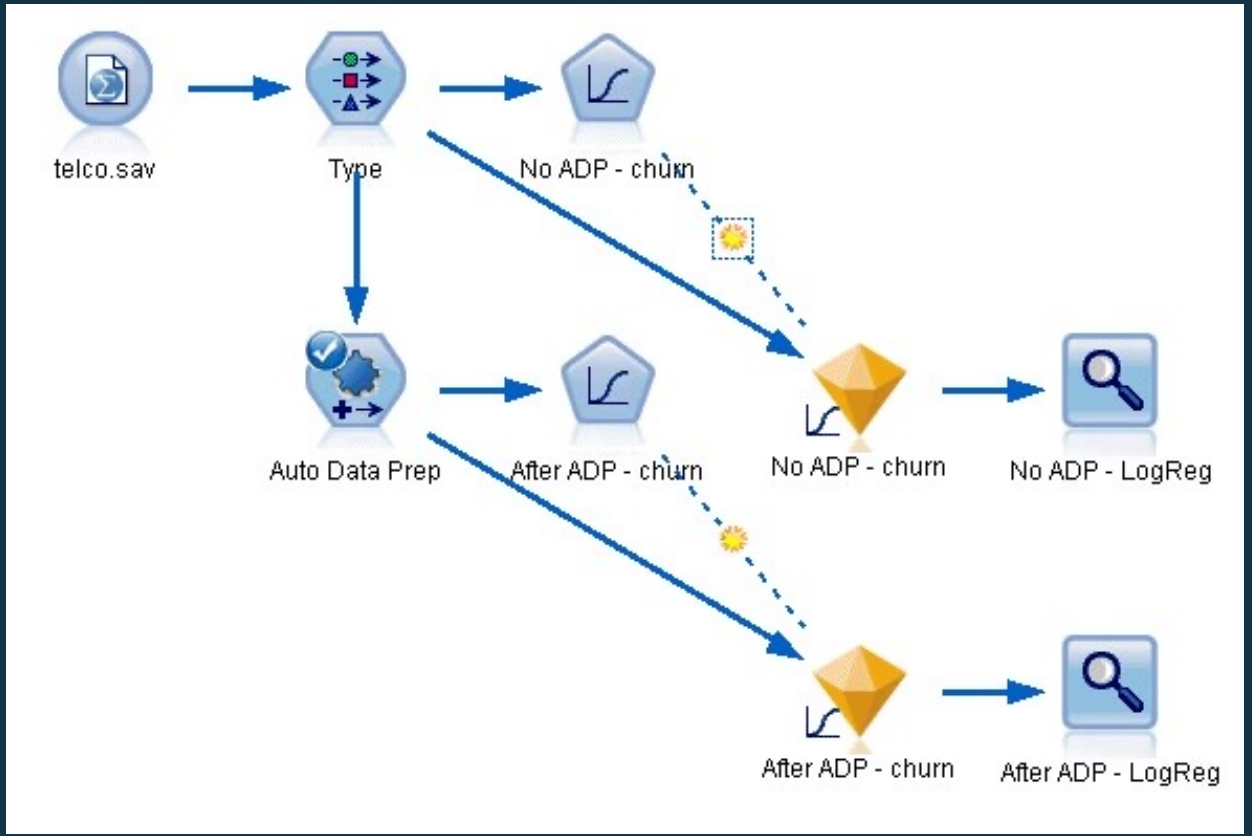
automatically perform the Basic Data Preparation operations

Nodes used to auto select the best suited model and generate top performing models

Download sample files from:
github.com/krishnac7/spssSamples

Automated Data Preparation

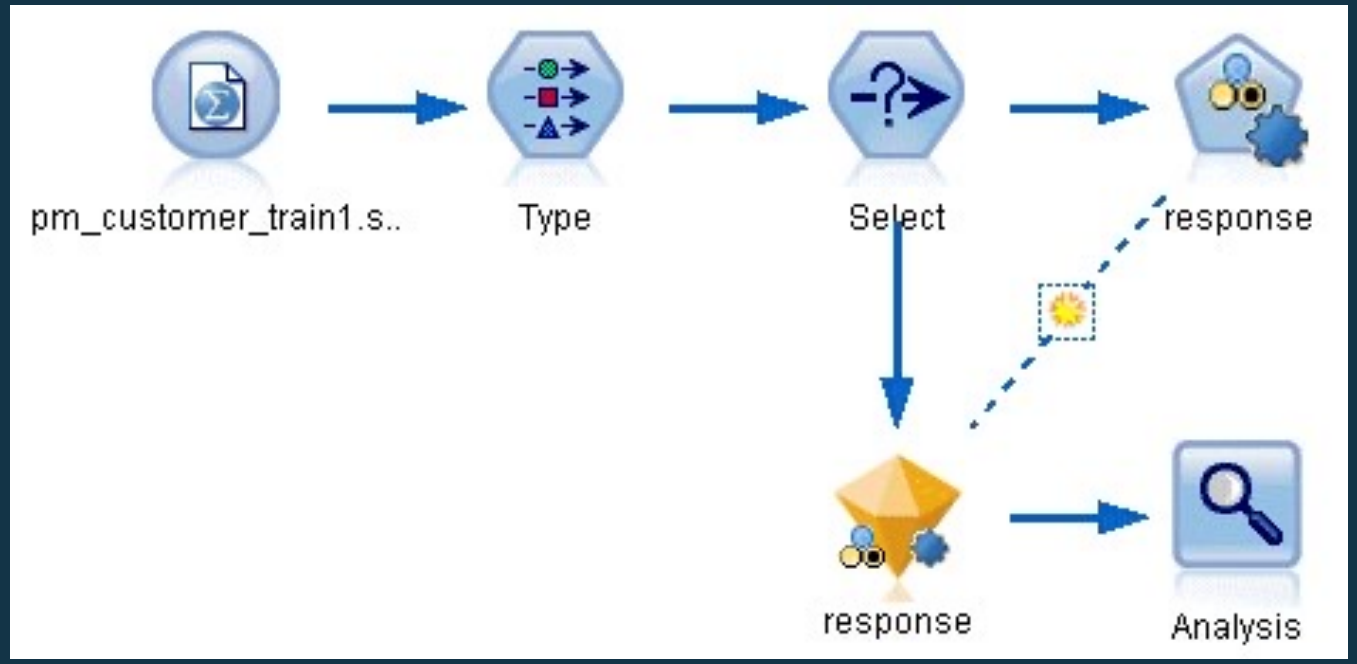
Compare and contrast model accuracies prior and post Data Prep



Download sample files from:
github.com/krishnac7/spssSamples

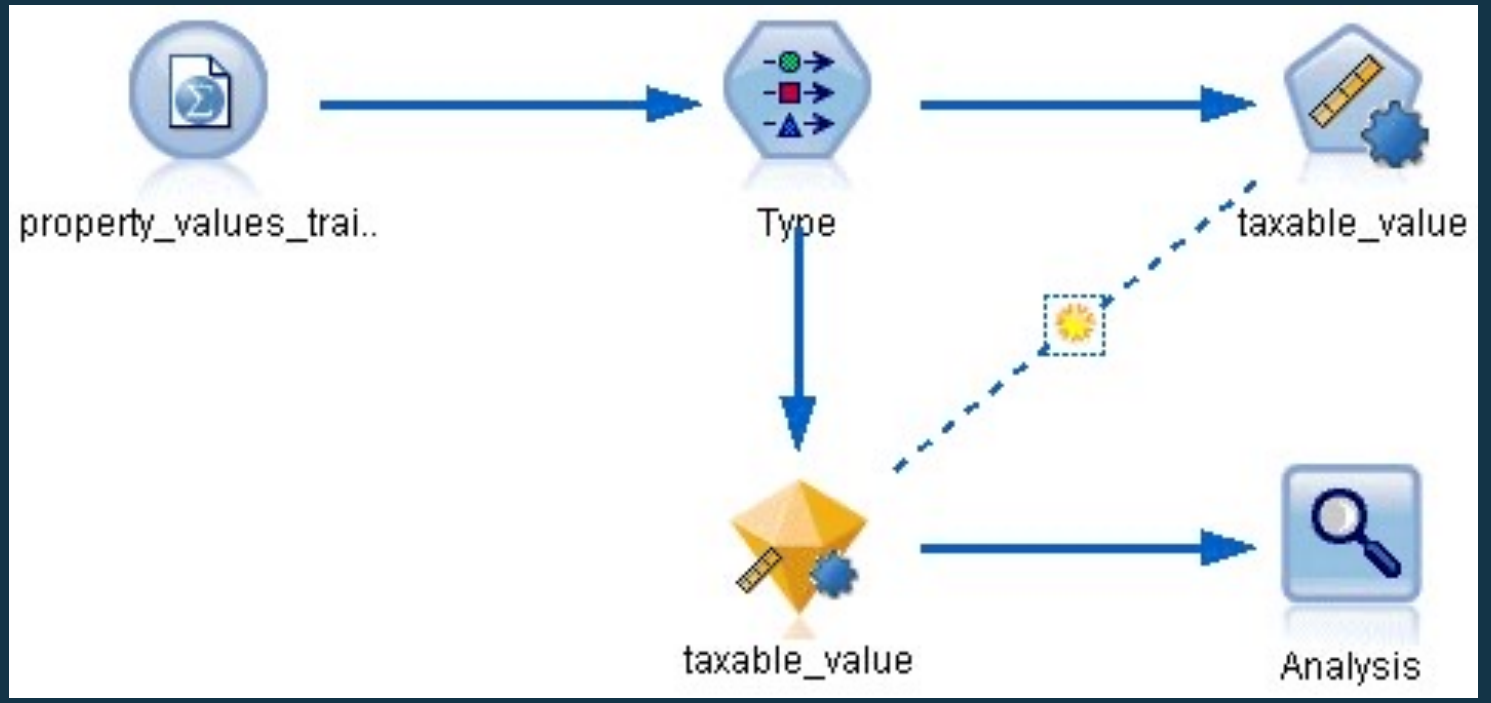
Automated modelling for Flag Target

A Company that wants to achieve more profitable results by matching the right offer to each customer.



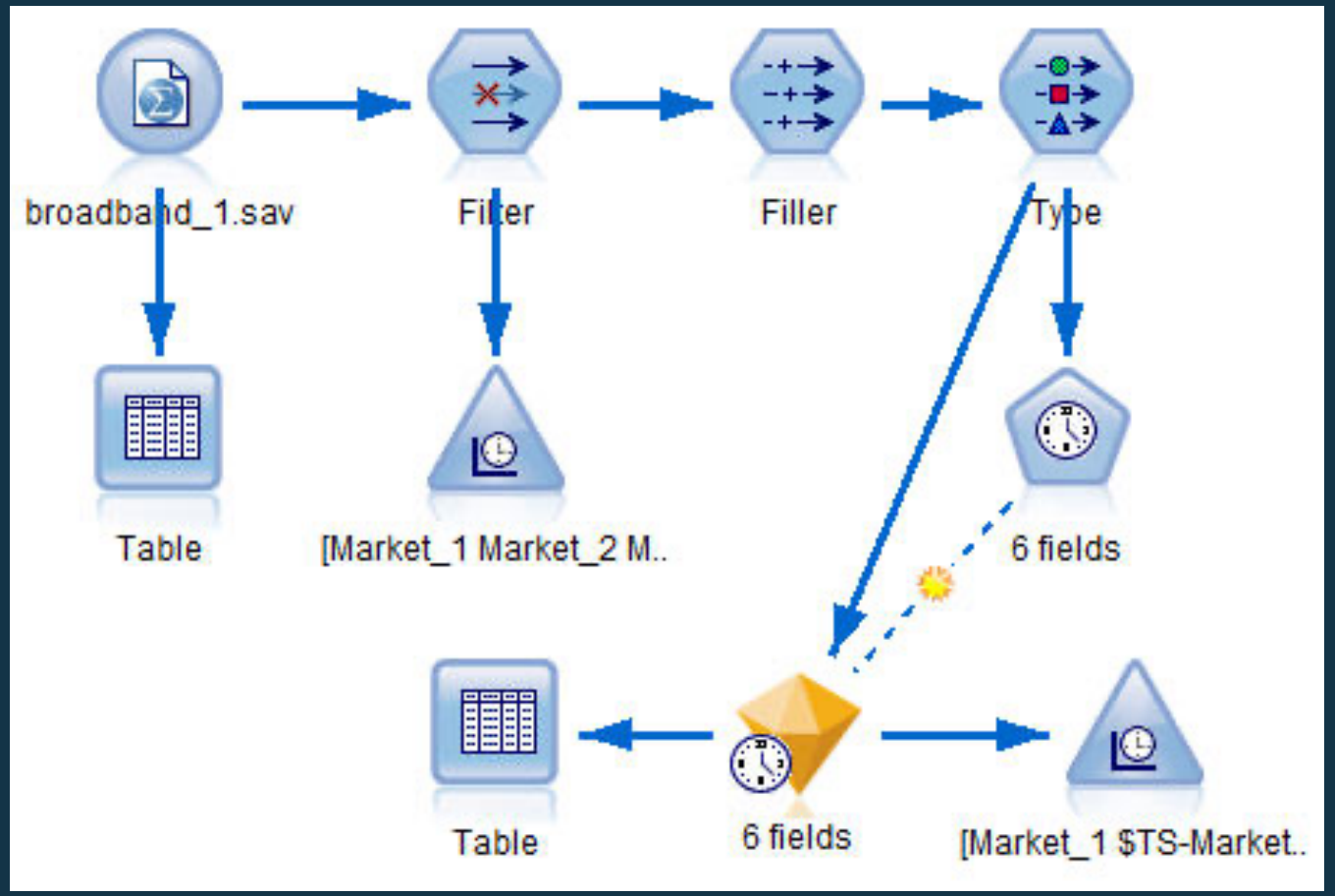
Automated Modelling for Continuous Target

We will build a model that predicts property values based on building type, neighbourhood, size, and other known factors.



Modelling Timeseries Data

Produce forecasts of user subscriptions in order to predict utilization of bandwidth



Thank you!

github.com/krishnac7/spssSamples

Krishna Balaga
krbalaga@in.ibm.com