Data Science Project for WAVE Volleyball

By Prithvi Balaji

12/30/2019

Goals

Natasha Balac <nbalac@eng.ucsd.edu> Nov 4, 2019, 6:43 AM

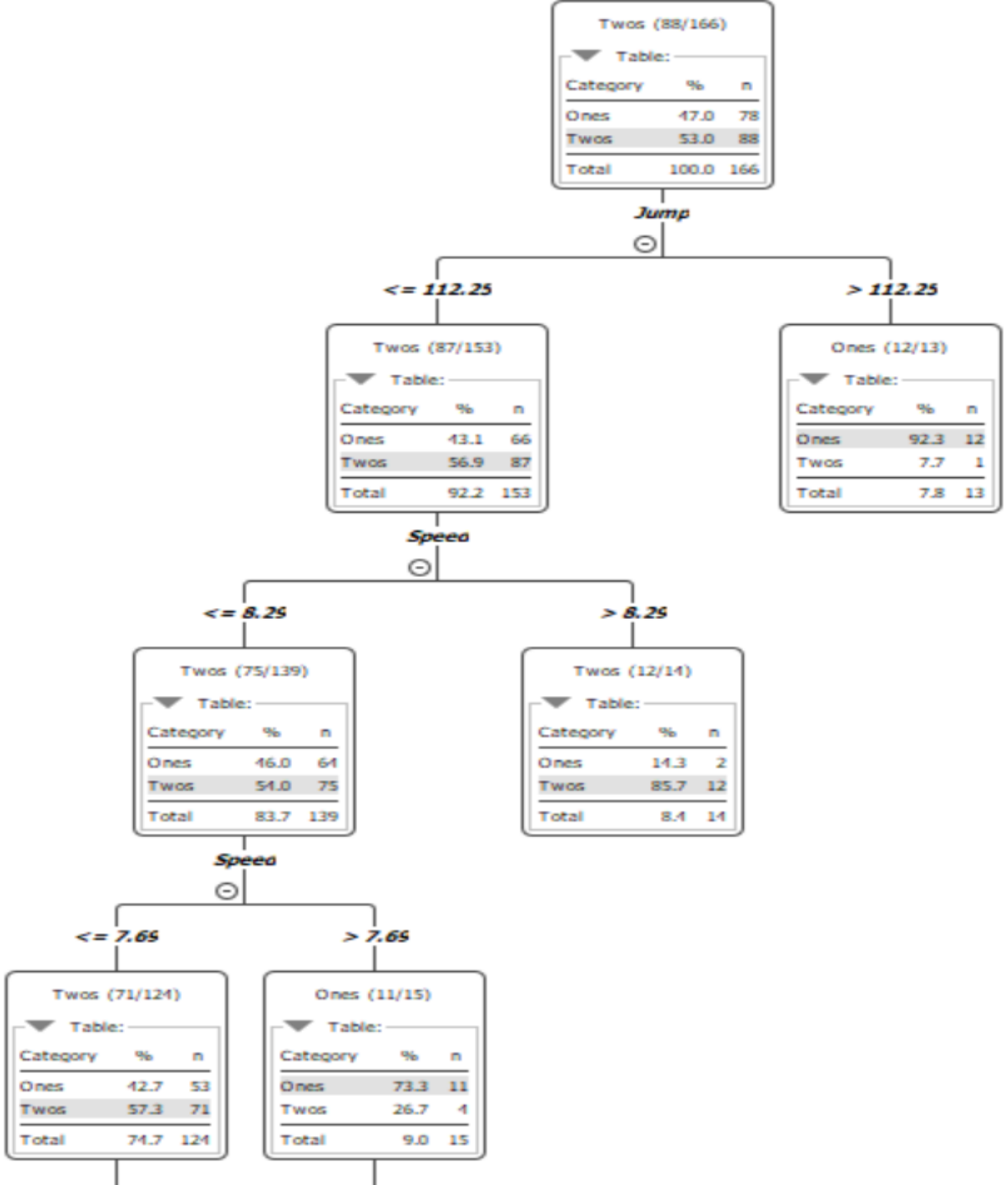
to me

Hi Prithvi,
Looks great. I am out of town most of this and next week - back on Thursday each week and can set time to chat.
Is there are way to show a larger number of athletes on the chart?
the other task you could work on is to build a predictive model to find out the characteristics of athletes that are in ones and twos team. maybe consider clustering and decision trees.
Best regards,
-Natasha

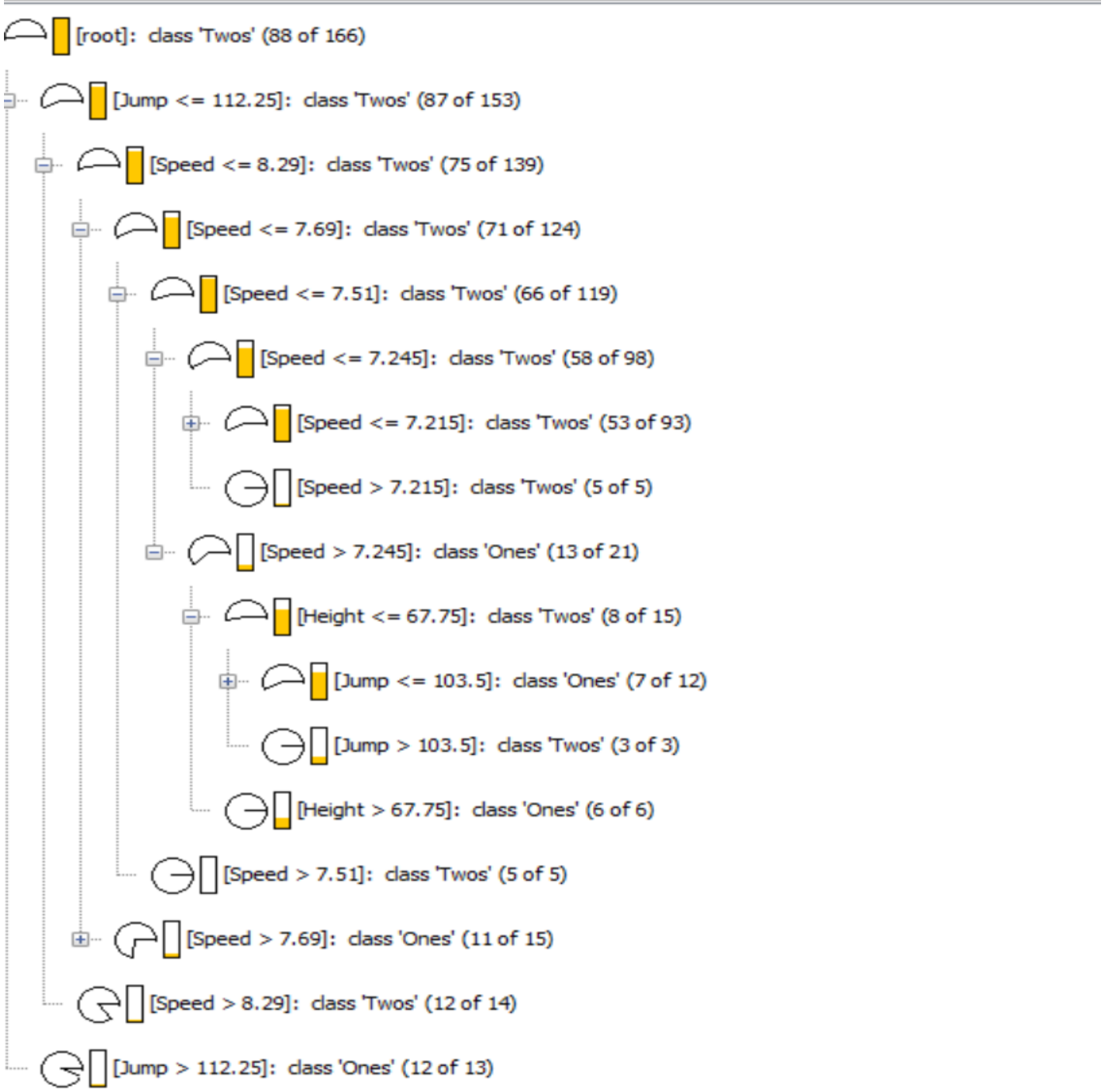
Goals

- Dataset
- KNIME workflow
- Predictive models: Decision Tree and Clustering
- Team – Ones & Twos
- Attributes: Jump, App Vertical, Speed, Height and Reach

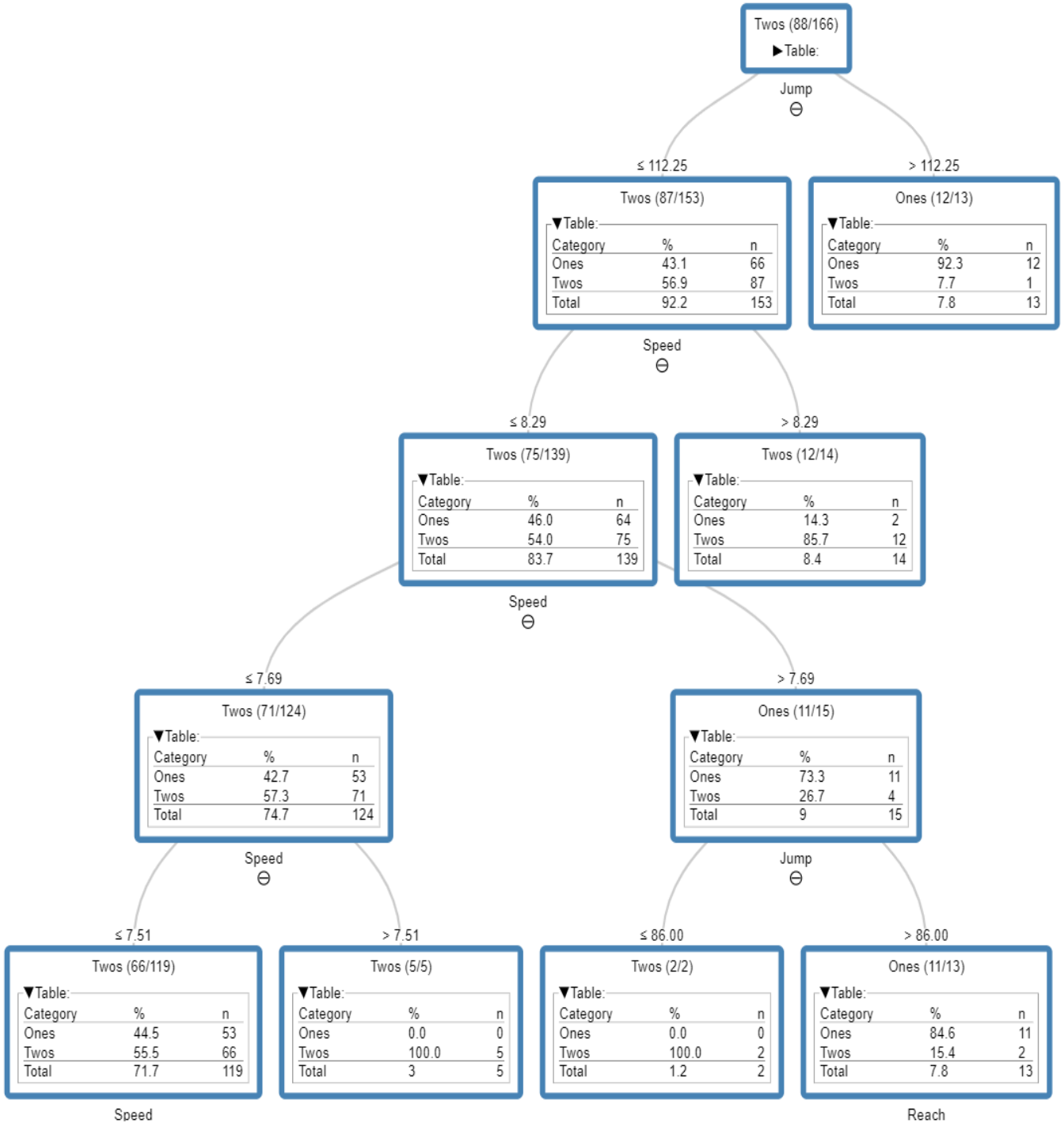
Decision Tree

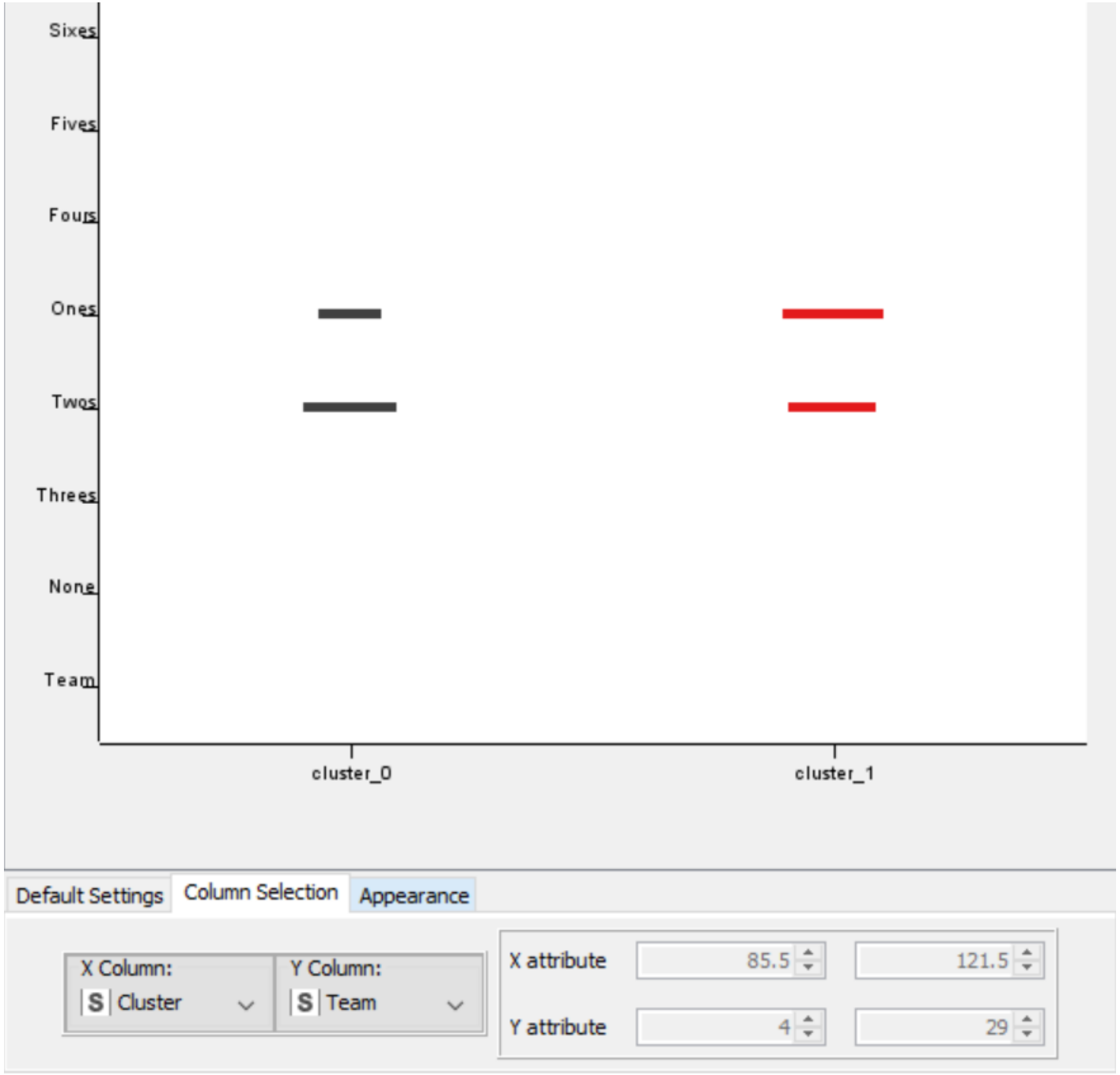


Simple Decision Tree



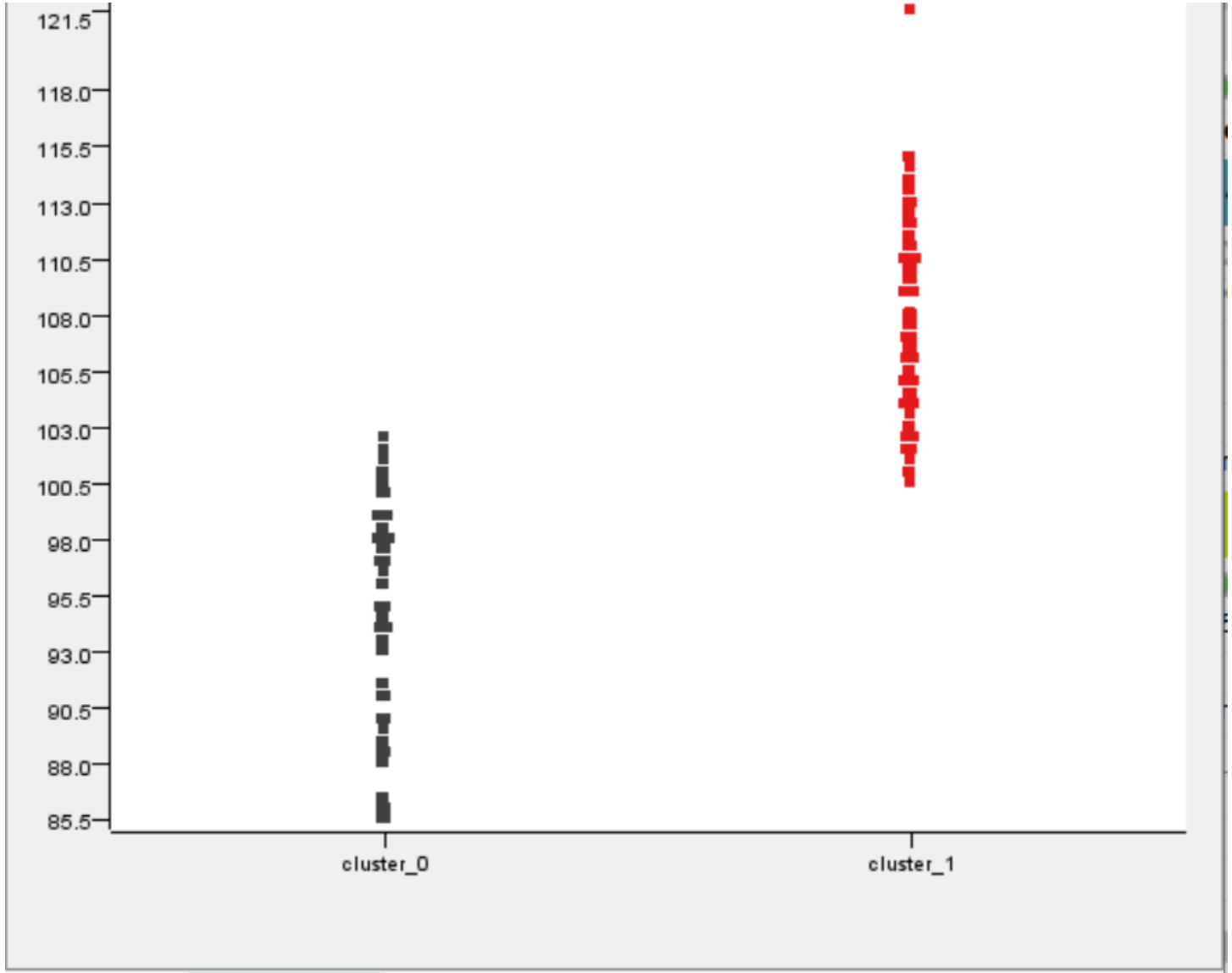
Interactive Decision Tree





Scatter Plot

Cluster v Team



Default Settings | Column Selection | Appearance

X Column: S Cluster

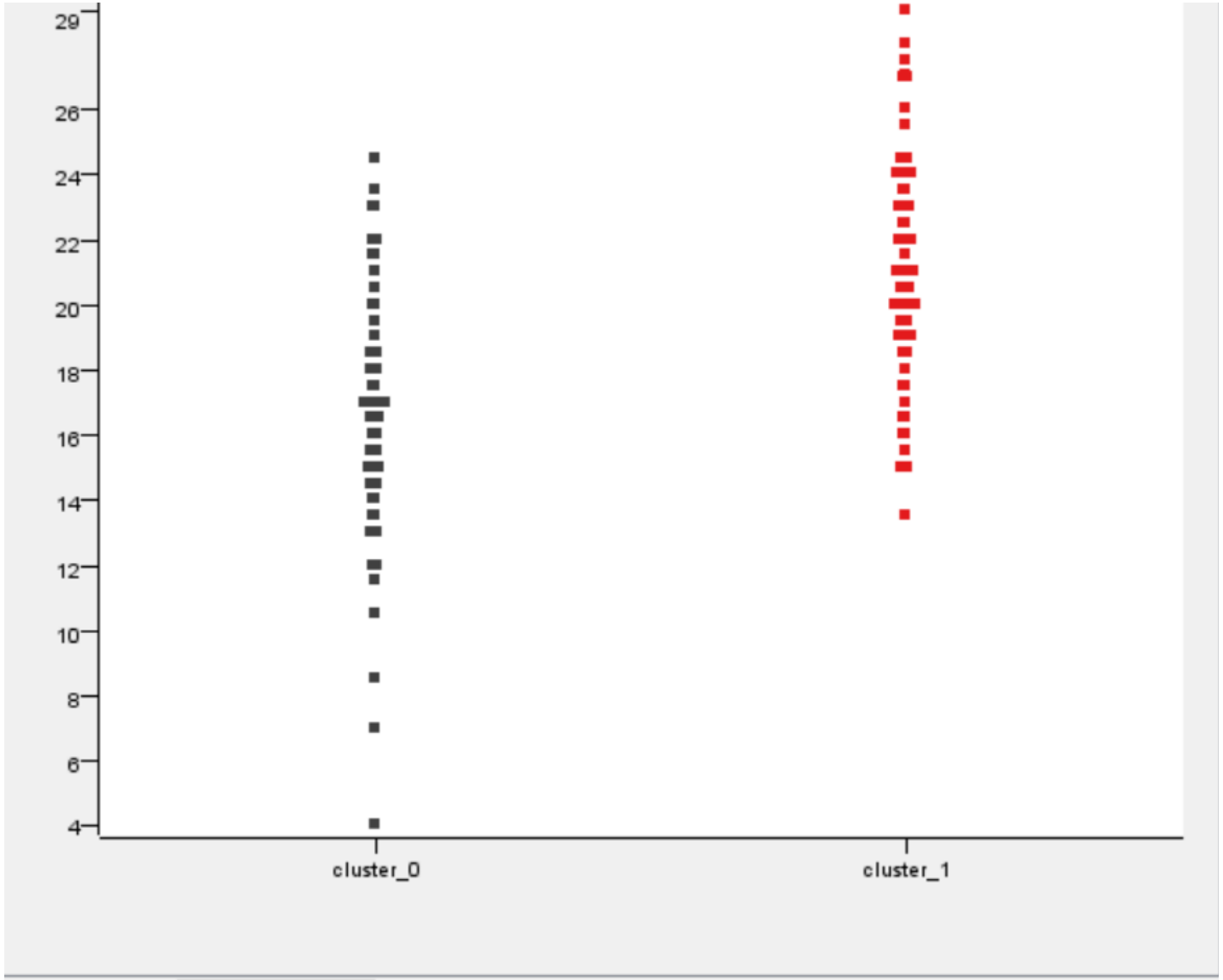
Y Column: D Jump

X attribute 85.5 121.5

Y attribute 85.5 121.5

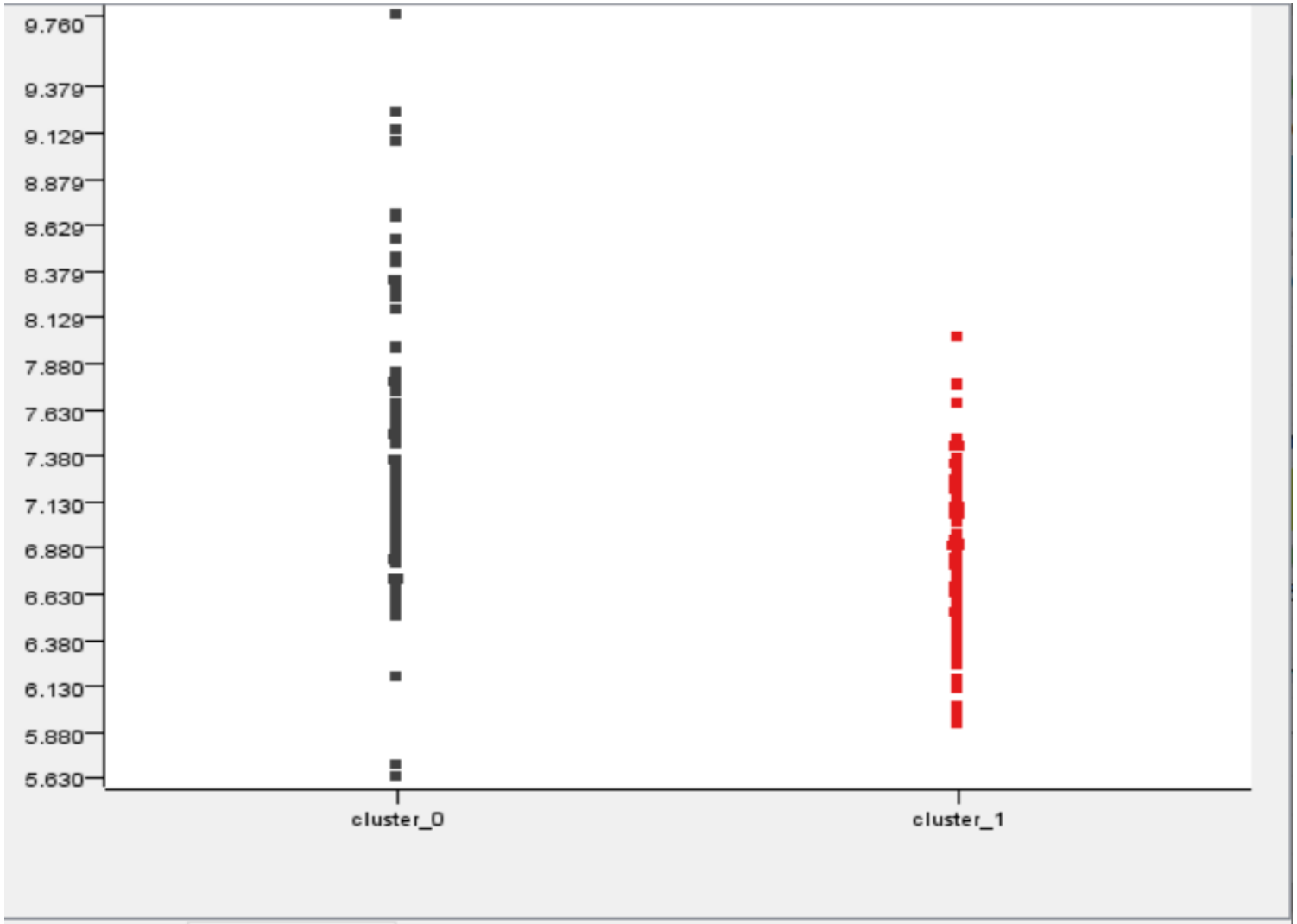
Scatter Plot

Cluster v Jump



Scatter Plot

Cluster v App Vertical



Default Settings | Column Selection | Appearance

X Column: S Cluster

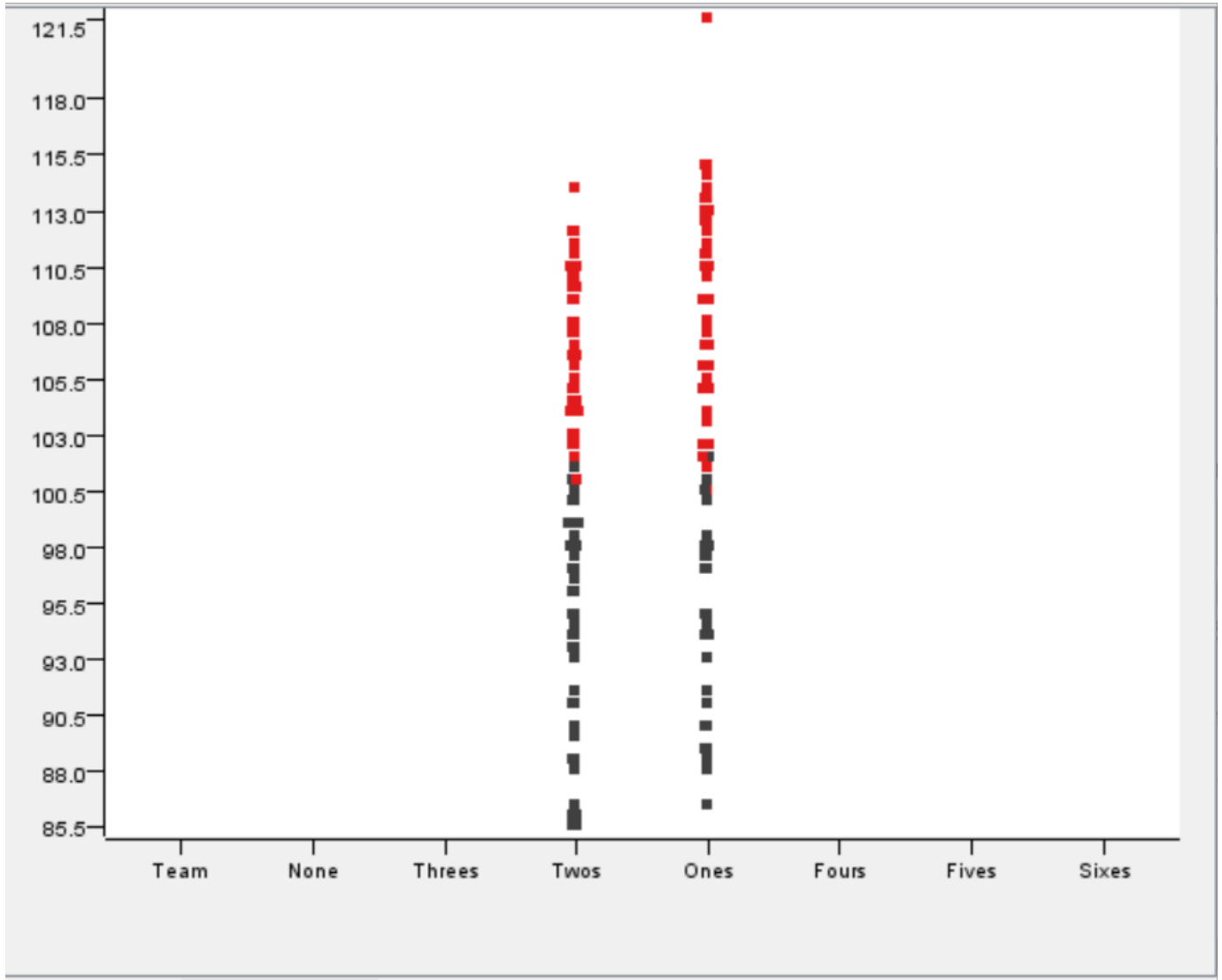
Y Column: D Speed

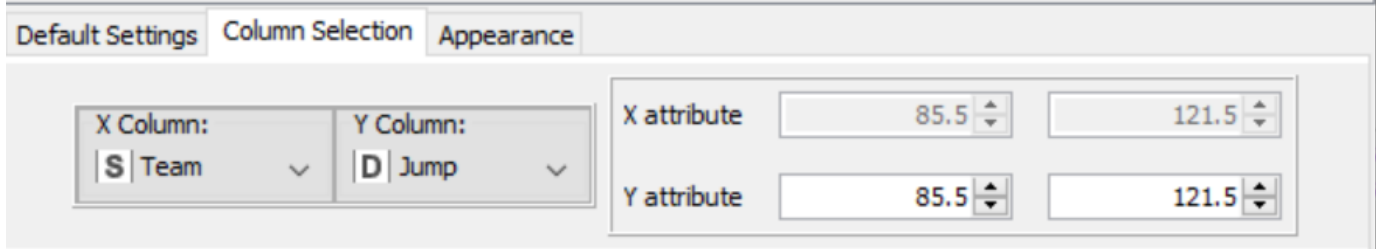
X attribute 85.5 121.5

Y attribute 5.63 9.76

Scatter Plot

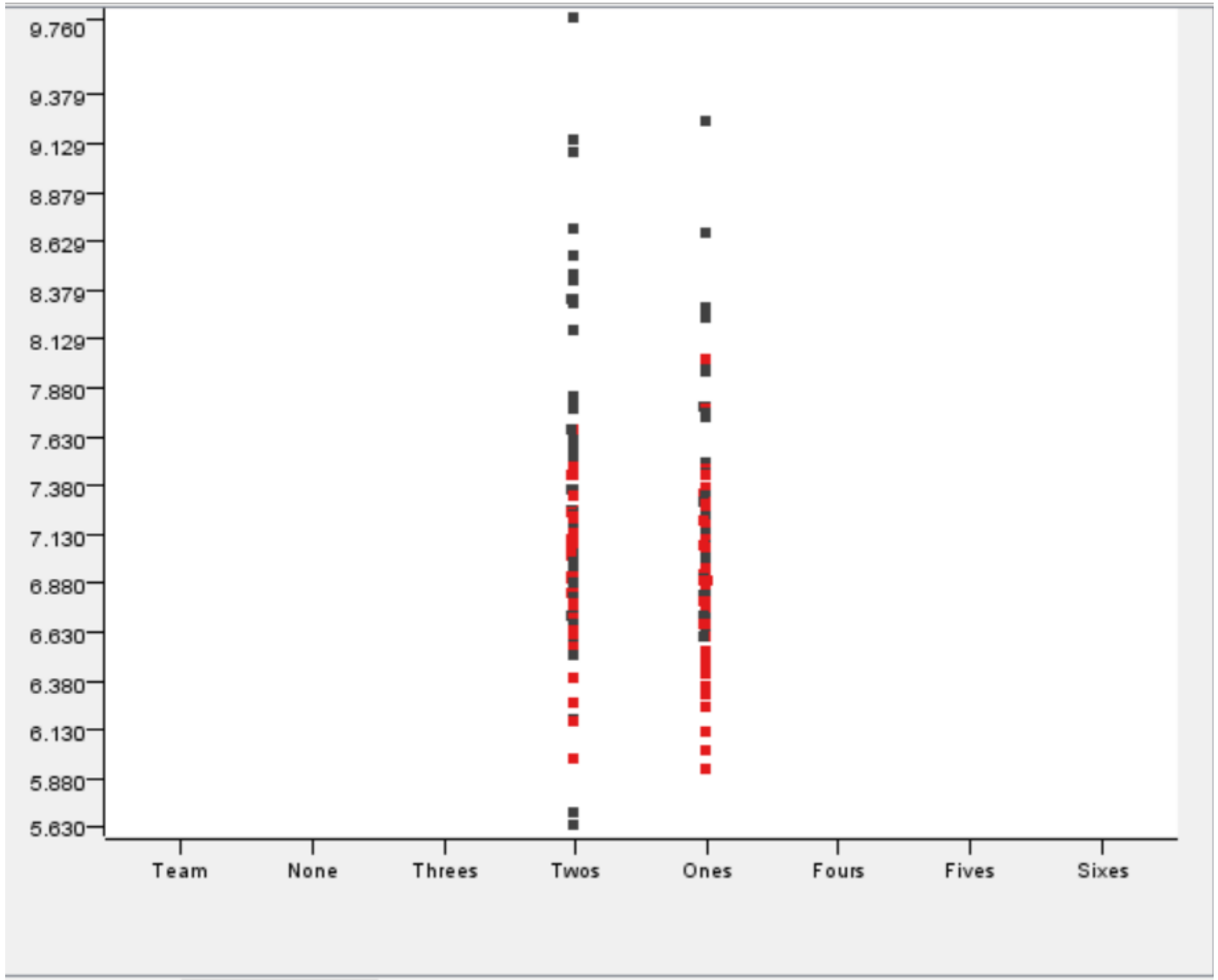
Cluster v Speed





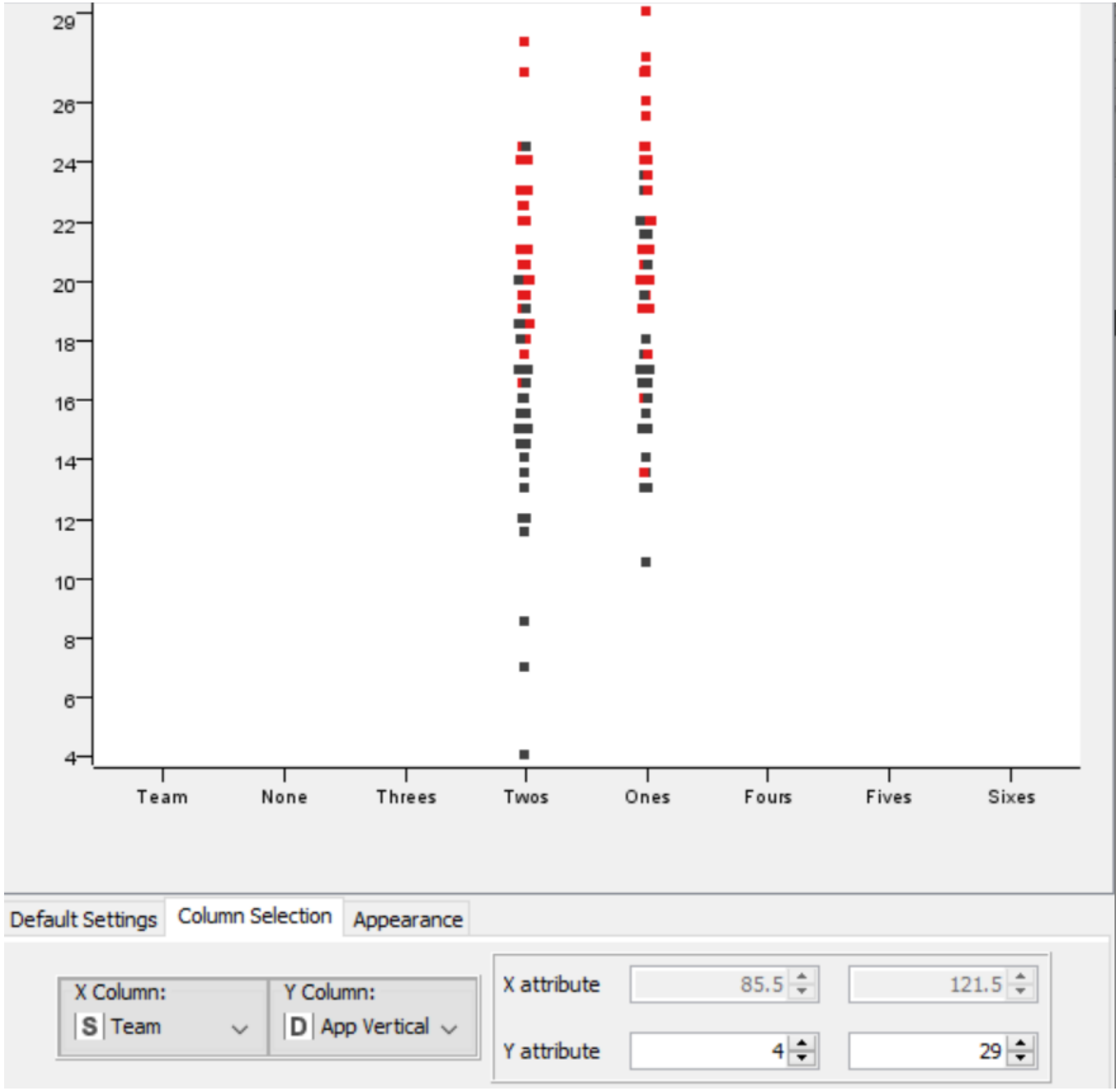
Scatter Plot

Team v Jump



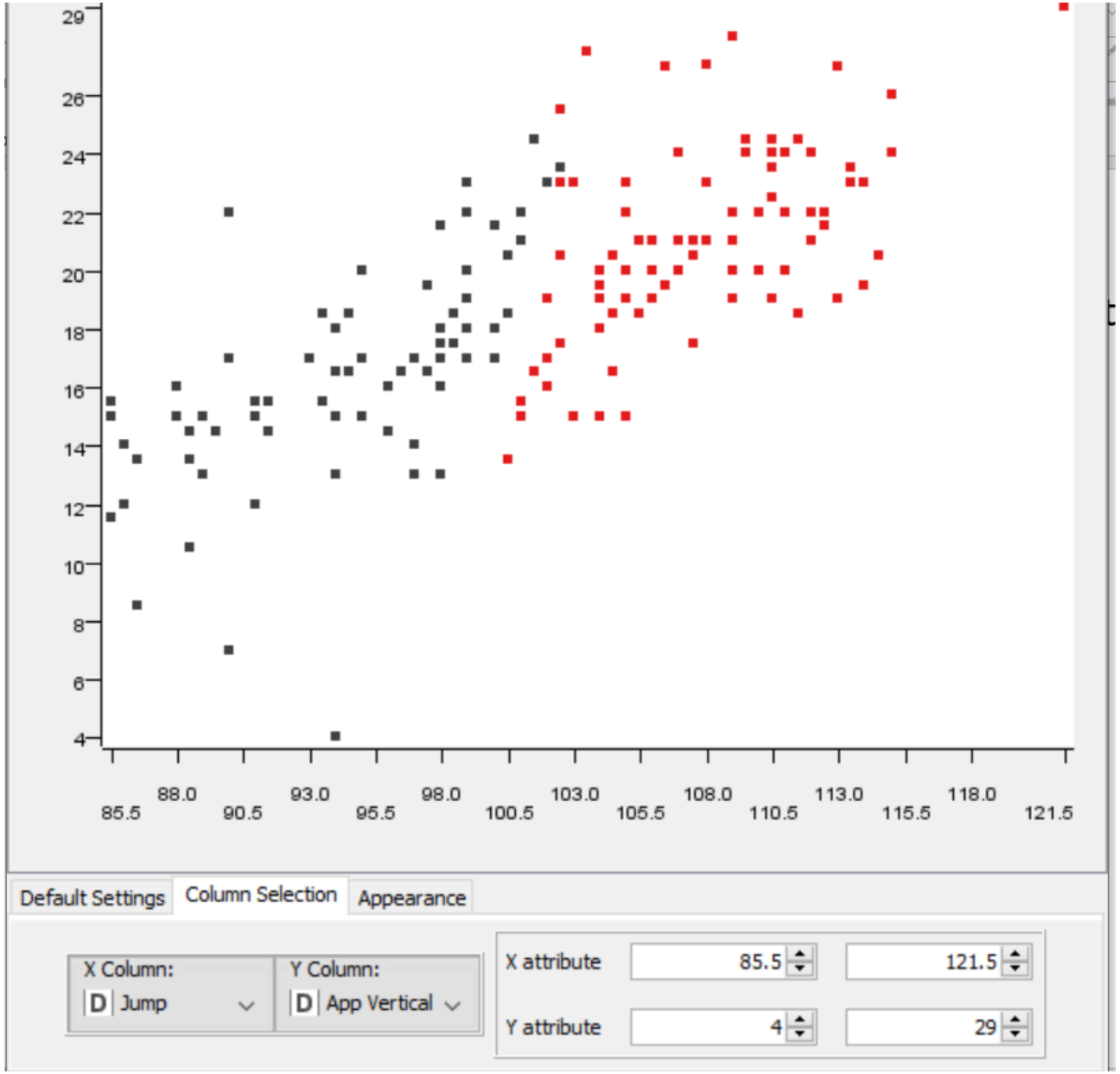
Scatter Plot

Team v Speed



Scatter Plot

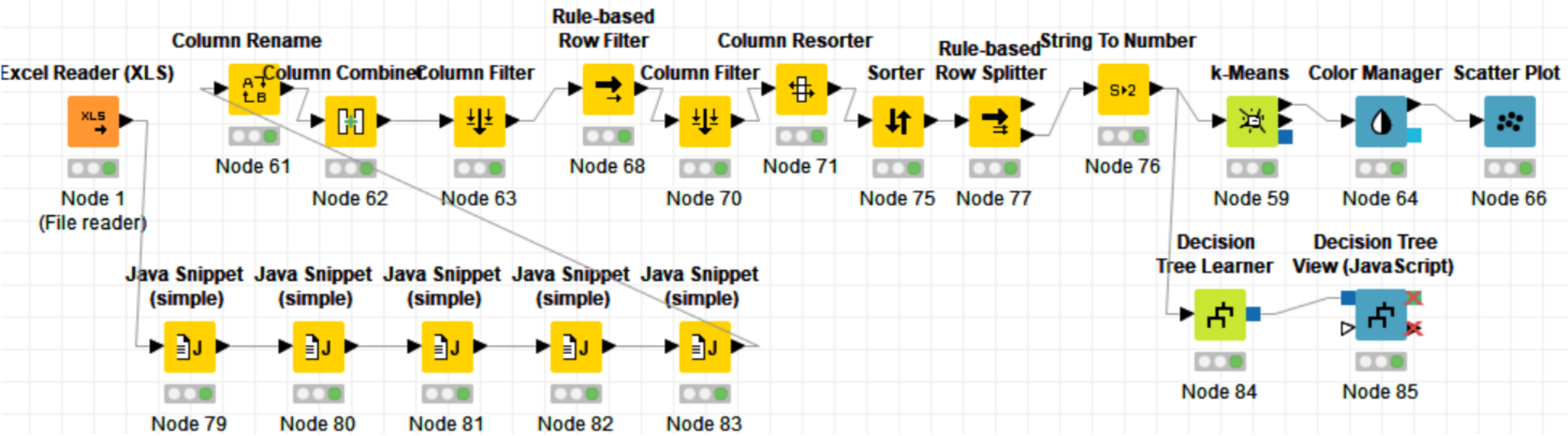
Team v App Vertical



Scatter Plot

Jump v App Vertical

Work Flow


Excel Reader (XLS)
Node 1
(File reader)
Column Rename
Node 61
Column Combiner
Node 62
Column Filter
Node 63
Rule-based
Row Filter
Node 68
Column Filter
Node 70
Column Resorter
Node 71
Sorter
Node 75
Rule-based
Row Splitter
Node 77
String To Number
Node 76
k-Means
Node 59
Color Manager
Node 64
Scatter Plot
Node 66
Java Snippet
(simple)
Node 79
Java Snippet
(simple)
Node 80
Java Snippet
(simple)
Node 81
Java Snippet
(simple)
Node 82
Java Snippet
(simple)
Node 83
Decision
Tree Learner
Node 84
Decision Tree
View (JavaScript)
Node 85

My observations so far from Decision Tree..

- Making the Ones team is predicated by the Jump attribute (Near 92 percent on Ones team had >112.25)
- Speed was merely a indicator of Ones/Twos team
- If an athlete had lower than a 7.7 Speed time, it came back to their Reach and their Jump

My observations so far from Clustering..

- In the Cluster v Speed, there is much more variance in the Twos Team
- In the Team v Jump, the Twos team tended to fill the bottom of the graph and Ones team tended to fill the top, resulting in Jump being a clear indicator of team decisions
- In the Cluster v App Vertical, there is many outliers in the Twos Team

Next steps..

- Need Dr. Balac's input & suggestions
- Expand model to include game attributes (Hit, Throw, Serve etc)
- Test predictive models with sample data
- Complete visualization of athletes from 10u to 16u in Team 1s and 2s