

Task 2: Formats for New Stores



1. What methodology did you use to predict the best store format for the new stores? Why did you choose that methodology? (Remember to Use a 20% validation sample with Random Seed = 3 to test differences in models.)

What are the three most important variables that help explain the relationship between demographic indicators and store formats? Please include a visualization.

I considered *Cluster* as target variable and selected all demographic variables as predictors. Compared the performance of Decision Tree, Forest and Boosted models regarding Overall Accuracy, F1 Score and the accuracy of each class. Boosted model presented the best performance.

I chose Boosted Model.

Fit and error measures					
Model	Accuracy	F1	Accuracy_1	Accuracy_2	Accuracy_3
Boosted Model	0.7647	0.8333	0.5000	1.0000	1.0000
Forest Model	0.7059	0.7917	0.3750	1.0000	1.0000
Decision Tree Model	0.7059	0.7083	0.6250	1.0000	0.5000

Confusion matrix of Boosted.Model			
	Actual_1	Actual_2	Actual_3
Predicted_1	4	0	0
Predicted_2	2	5	0
Predicted_3	2	0	4

Confusion matrix of Decision.Tree.Model			
	Actual_1	Actual_2	Actual_3
Predicted_1	5	0	2
Predicted_2	2	5	0
Predicted_3	1	0	2

Confusion matrix of Forest.Model			
	Actual_1	Actual_2	Actual_3
Predicted_1	3	0	0
Predicted_2	3	5	0
Predicted_3	2	0	4

Image 8, Model Comparison Report

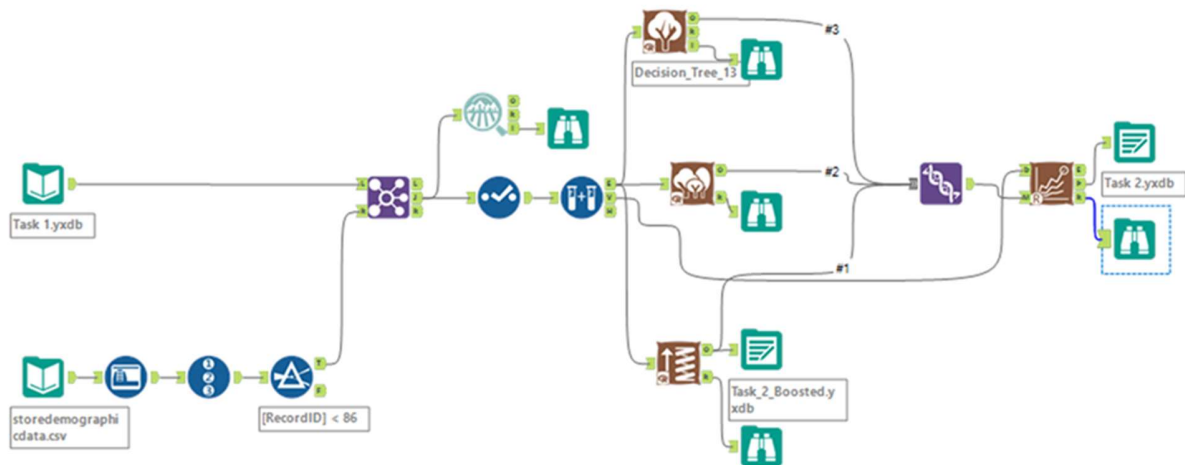


Image 9, Defining the best model using Alteryx

The three most important variables that help explain the relationship between demographic indicators and store formats are *Age0to9*, *HVal750KPlus* and *Age65Plus*.

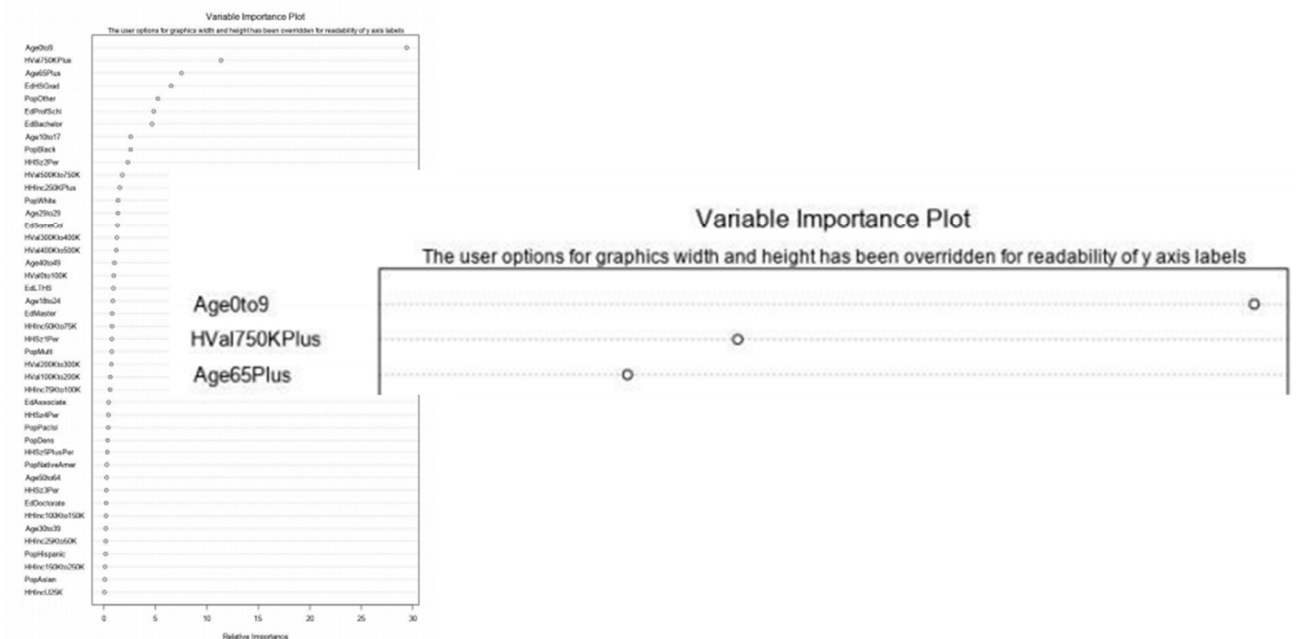


Image 10, Three most important variables

2. What format do each of the 10 new stores fall into? Please fill in the table below.

I applied Boosted model output, New Store data, Score tool and Formula tool resulting in 10% in cluster 1, 30% in cluster 3 and 60% in cluster 2.

Store Number	Segment
S0086	1
S0087	2
S0088	3
S0089	2
S0090	2
S0091	3
S0092	2
S0093	3
S0094	2
S0095	2

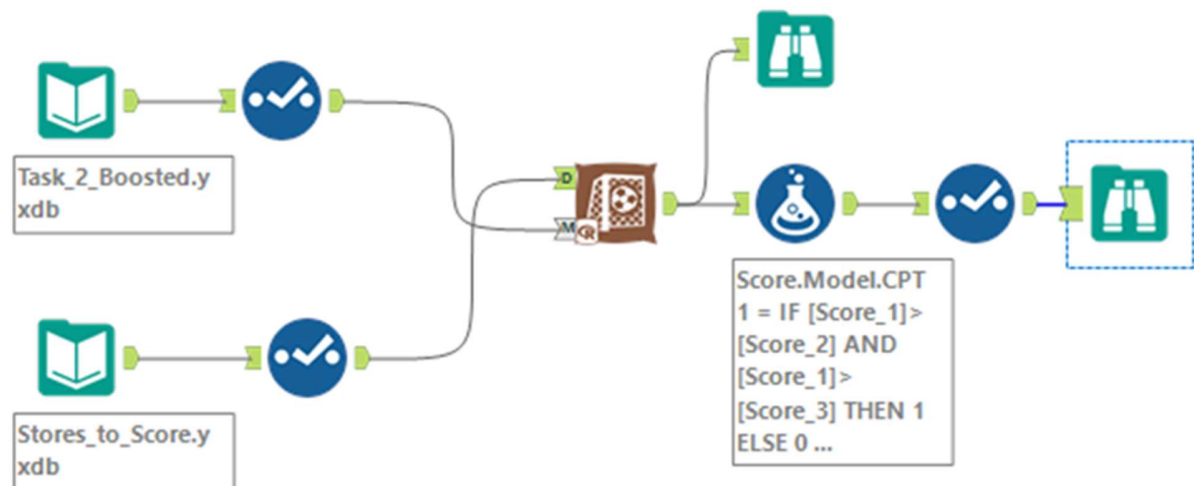


Image 11, Scoring and defining Store Formats for New Stores using Alteryx