Team 2 - Homework Two

Assignment 3: KJ 8.1-8.3; KJ 8.7

Juliann McEachern

11/8/19

Dependencies

```
# Forecast libraries
libraries("mlbench", "AppliedPredictiveModeling", "party", "randomForest",
    "caret")

# Formatting Libraries
libraries("default", "knitr", "kableExtra", "tidyverse")

# Plotting Libraries
libraries("ggplot2", "grid", "ggfortify")
```

(1) Kuhn & Johnson 8.1

Recreate the simulated data from Exercise 7.2:

```
set.seed(200)
simulated <- mlbench.friedman1(200, sd = 1)
simulated <- cbind(simulated$x, simulated$y)
simulated <- as.data.frame(simulated)
colnames(simulated)[ncol(simulated)] <- "y"
```

(a). Fit a random forest model to all of the predictors, then estimate the variable importance scores. Did the random forest model significantly use the uninformative predictors (V6-V10)?

```
model1 <- randomForest(y ~ ., data = simulated, importance = T,
    ntree = 1000)
rfImp1 <- varImp(model1, scale = F)
```

The random forest model predominantly favored the predictor variables V1-V5, with the exception of V3. The uninformative predictors (V6-10) all had significantly lower overall importance.

Table 1: Variable Importance for Random Forest Model1

	V1	V4	V2	V5	V3	V6	V7	V10	V9	V8
Overall	8.73	7.62	6.42	2.02	0.76	0.17	-0.01	-0.07	-0.1	-0.17

(b). Now add an additional predictor that is highly correlated with one of the informative predictors. Fit another random forest model to these data. Did the importance score for V1 change? What happens when you add another predictor that is also highly correlated with V1? For example:

```r
simulated$duplicate1 <- simulated$V1 + rnorm(200) * 0.1
model2 <- randomForest(y ~ ., data = simulated, importance = T,
    ntree = 1000)
rfImp2 <- varImp(model2, scale = F)
```

The correlation of the simulated V1 and duplicate variable was 0.9460206. Adding a highly correlated value, such as this duplicate, decreased the overall variable importance of V1.

Table 2: Variable Importance for Random Forest Model2

	V4	V2	V1	duplicate1	V5	V3	V6	V10	V9	V7	V8
Overall	7.05	6.07	5.69	4.28	1.87	0.63	0.14	0.03	0.01	-0.01	-0.04

(c). Use the `cforest` function in the party package to fit a random forest model using conditional inference trees. The party package function `varimp` can calculate predictor importance. The `conditional` argument of that function toggles between the traditional importance measure and the modified version described in Strobl et al. (2007). Do these importances show the same pattern as the traditional random forest model?

The `cforest` approach shows the same pattern as the traditional random forest model. Both models have similar overall importance for V4,V1, and V2, however the importance of variable V5 decreases slightly with the `cforest` approach. Like `rforest`, the `cforest` model also decreases the overall importance of V1 when the duplicate variable is evaluated in the model. The overall importance decreases more with the `cforest` method.

Table 3: Variable Importance for Conditional Inference Trees Models

	V4	V2	duplicate1	V1	V5	V3	V7	V6	V10	V9	V8
Original	6.66	5.19	NA	5.50	1.21	0.03	-0.01	-0.01	-0.01	0.00	0.00
Duplicate	6.05	4.82	1.92	1.92	1.06	0.02	0.01	0.00	-0.01	-0.01	-0.02

(d). Repeat this process with different tree models, such as boosted trees and Cubist. Does the same pattern occur?

We repeated this process a final time using a gradient boosted model. We used a simple model with 1000 trees The boosted trees approach also picked V1-V5 to be the variables with the most influence. Again, we see V5 and V3 have lower importance than V1, V2, and V4. The duplicate value in the `gbm` model also responds similarly to our previous approaches. The importance of V1 and the duplicate variable both decrease. Unlike the other models, however, the `gbm` model only slightly lowers the importance of V1 and the duplicate only has half the level of importance, compared to the original.

Table 4: Variable Relative Influence of Gradient Boosted Models

	V4	V2	V1	V3	V5	duplicate1	V7	V6	V10	V9	V8
Original	25.10	19.60	23.43	9.10	11.22	NA	2.94	2.53	1.59	2.4	2.09
Duplicate	24.94	19.43	16.85	10.28	10.10	8.16	3.16	2.18	1.96	1.5	1.46

(2) Kuhn & Johnson 8.2

Use a simulation to show tree bias with different granularities.

Basic regression trees split predictor variables into small groupings based on the response variable. According to the text, "predictors with a higher number of distinct values are favored over more granular predictors." The simulation below helps us visualize this tree bias among different granularities.

In our simulation:

- V1 contains 100 random samples of 2 numbers
- V2 contains 100 random samples of 100 numbers
- V3 contains 100 random samples of 1000 numbers
- V4 contains 100 random samples of 5000 numbers

Our response variable, y, was derived using the formula: $y = V1 * V2 * V3 + e$, where e is a random error.

Table 5: Head of Simulation Table

V1	V2	V3	V4	y
1	13	946	493	12298.61
2	47	151	1164	14195.20
1	17	749	2128	12732.55
2	55	344	3788	37840.56
2	59	645	2281	76109.04
1	87	477	3513	41498.56

Despite V1-V3 having equal weight in our model, V3 had the highest overall variable importance. This helps illustrate the bias in trees as V1 and V2 were both of much smaller granularity.

	Overall
V2	32.829698
V3	30.738404
V1	25.108499
V4	4.484885

(3) Kuhn & Johnson 8.3

In stochastic gradient boosting the bagging fraction and learning rate will govern the construction of the trees as they are guided by the gradient. Although the optimal values of these parameters should be obtained through the tuning process, it is helpful to understand how the magnitudes of these parameters affect magnitudes of variable importance. Figure 8.24 provides the variable importance plots for boosting using two extreme values for the bagging fraction (0.1 and 0.9) and the learning rate (0.1 and 0.9) for the solubility data. The left-hand plot has both parameters set to 0.1, and the right-hand plot has both set to 0.9:

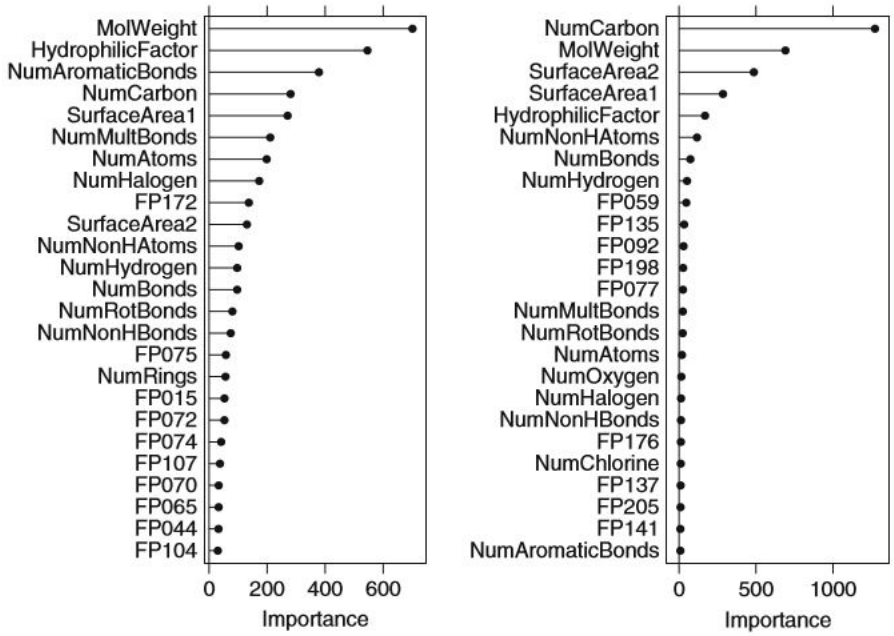


Fig. 8.24: A comparison of variable importance magnitudes for differing values of the bagging fraction and shrinkage parameters. Both tuning parameters are set to 0.1 in the *left* figure. Both are set to 0.9 in the *right* figure

(a). Why does the model on the right focus its importance on just the first few of predictors, whereas the model on the left spreads importance across more predictors?

Stochastic gradient boosting distributes data differently than other classification trees. Boosted trees contribute to the model unequally and have a distribution structure dependent on past trees and tree depth. The dependency aspect of boosted trees creates a much steeper cut off in variable importance than other tree models. From the text, we know that the learning rate is the fraction of the current predicted value added to the previous iteration's predicted value. And, bagging reduces the variance of a prediction. Thus, the bagging and shrinking parameters would greatly affect the influence of boosted trees on the predictor variable importance.

(b). Which model do you think would be more predictive of other samples?

The higher bagging and learning rate tuning parameter of 0.9 could make the prediction model more stable and more accurate. We would have to run tests to ensure that the increased value did not over-fit the data to our model.

(c). How would increasing interaction depth affect the slope of predictor importance for either model in Fig.8.24?

Increased interaction depth would increase the spread of variable importance as each tree would have more features to be considered during the splitting process.

(4) Kuhn & Johnson 8.7

Refer to Exercises 6.3 and 7.5 which describe a chemical manufacturing process. Use the same data imputation, data splitting, and pre-processing steps as before and train several tree-based models:

(a). Which tree-based regression model gives the optimal resampling and test set performance?

Resampling & Test Performance

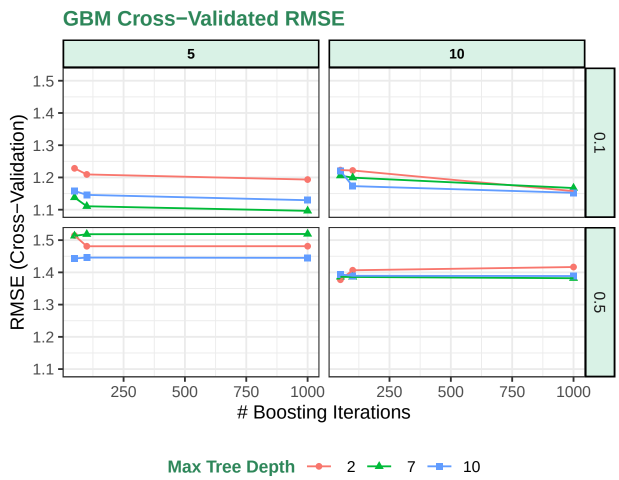
For this question, we evaluated random forest, gradient boosted, and cubist models on our chemical manufacturing process data. We found that the cubist model gave the best results, however the gradient boosted performed similiarly.

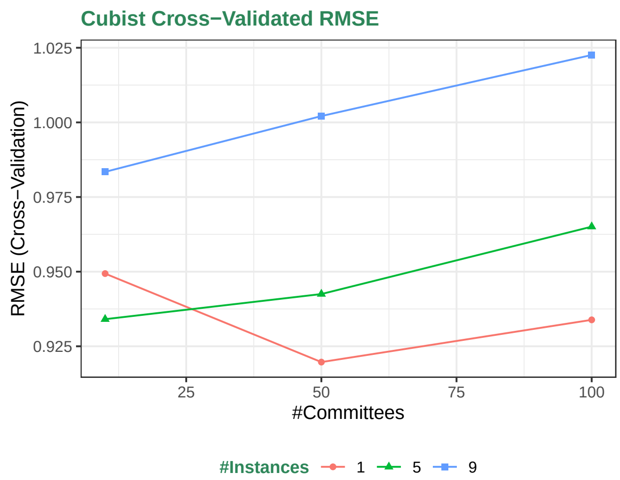
Table 6: Model Performance

	RMSE	RSquared	MAE
rfTrain	1.1225	0.6288	0.8835
rfTest	1.1553	0.8461	0.8699
gbmTrain	1.0964	0.6171	0.8875
gbmTest	1.0916	0.7947	0.8108
cbTrain	**0.9197**	**0.7503**	**0.6608**
cbTest	**1.0916**	**0.7947**	**0.8108**

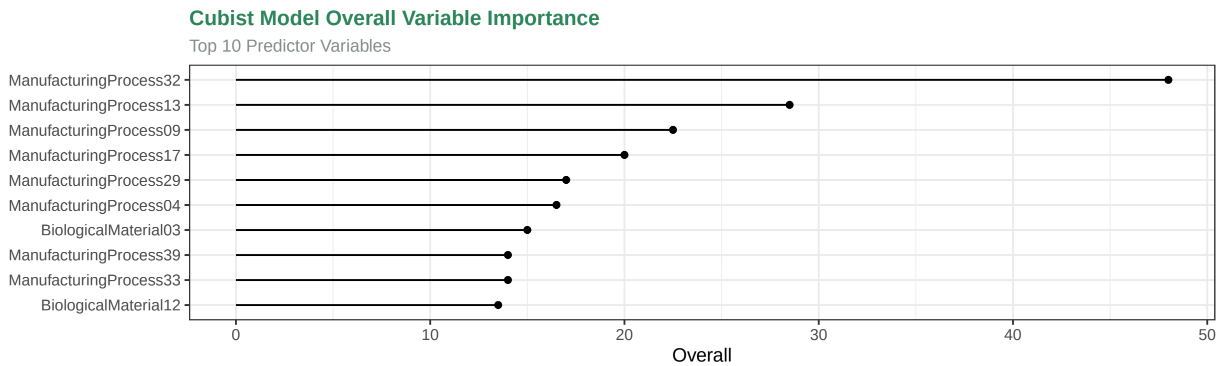
RMSE Plot:

We compared the RMSE metric between our cubist and gradient boosted below and determined the cubist was our optimal model.





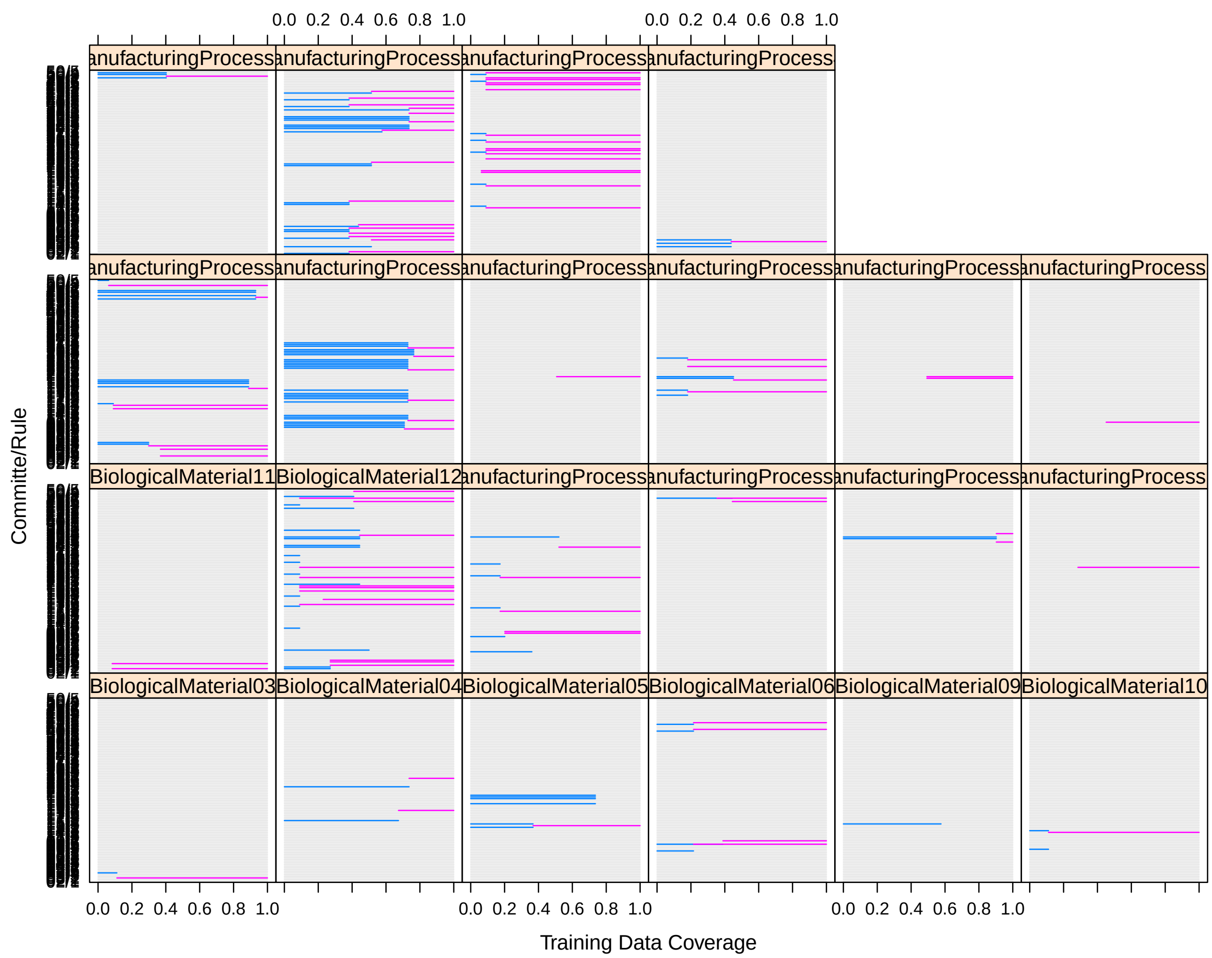
(b). Which predictors are most important in the optimal tree-based regression model? Do either the biological or process variables dominate the list? How do the top 10 important predictors compare to the top 10 predictors from the optimal linear and nonlinear models?



Manufacturing processes dominate the top 10 list of important predictors for the cubist model. The high presence of manufacturing is inline with our non-linear models, whereas our optimal linear model favored biological materials.

(c). Plot the optimal single tree with the distribution of yield in the terminal nodes. Does this view of the data provide additional knowledge about the biological or process predictors and their relationship with yield?

I was unable to find a package which plotted the terminal nodes of the cubist package. The final model summary, however, provides the rules used for the distriubution of yield in the terminal nodes. The cubist model also allows you to view the splits as a dotplot shown below. For each split, a panel is created for each variable to show how the committee/rule was applied. I have researched the cubist method and can not find a way (ie party package) to nicely plot / view what this question is asking. I will leave this answer for now and I am hoping to work as a team to find a better solution :)




```
Call:
cubist.default(x = x, y = y, committees = param$committees)


Cubist [Release 2.07 GPL Edition]  Tue Nov 12 00:33:17 2019
---------------------------------


    Target attribute `outcome'

Read 144 cases (57 attributes) from undefined.data

Model 1:

  Rule 1/1: [144 cases, mean 40.168, range 35.25 to 44.16, est err 0.919]

    outcome = -14.502 + 0.29 ManufacturingProcess32
              + 0.53 ManufacturingProcess09 - 0.241 ManufacturingProcess33

Model 2:

  Rule 2/1: [55 cases, mean 38.931, range 36.77 to 41.43, est err 0.977]
```

```
    if
    ManufacturingProcess32 <= 156
    then
    outcome = -5.979 + 1.333 ManufacturingProcess29
              - 0.00577 ManufacturingProcess27 - 0.18 BiologicalMaterial02
              + 0.153 BiologicalMaterial03 + 0.07 ManufacturingProcess32
              - 0.37 ManufacturingProcess13 + 0.048 ManufacturingProcess04
              - 0.21 ManufacturingProcess21 - 0.027 ManufacturingProcess28
              + 0.067 BiologicalMaterial05

  Rule 2/2: [89 cases, mean 40.932, range 35.25 to 44.16, est err 0.874]

    if
    ManufacturingProcess32 > 156
    then
    outcome = -25 + 2.007 ManufacturingProcess29
              - 0.00869 ManufacturingProcess27 - 0.272 BiologicalMaterial02
              + 0.179 ManufacturingProcess32 - 1 ManufacturingProcess13
              + 0.23 BiologicalMaterial03 - 0.114 ManufacturingProcess28
              + 0.073 ManufacturingProcess04 + 0.86 ManufacturingProcess45
              - 0.047 ManufacturingProcess24 - 0.32 ManufacturingProcess21
              + 0.1 BiologicalMaterial05

Model 3:

  Rule 3/1: [21 cases, mean 38.723, range 36.77 to 42.73, est err 1.042]

    if
    BiologicalMaterial03 > 63.44
    BiologicalMaterial11 > 140.84
    BiologicalMaterial12 <= 19.73
    then
    outcome = 262.361 - 1.117 BiologicalMaterial11
              + 0.669 BiologicalMaterial03 - 0.0172 ManufacturingProcess19
              - 0.025 ManufacturingProcess32

  Rule 3/2: [39 cases, mean 38.724, range 35.25 to 42.73, est err 1.259]

    if
    BiologicalMaterial12 <= 19.73
    then
    outcome = 37.402 - 0.112 BiologicalMaterial11
              + 0.31 ManufacturingProcess09 - 0.175 ManufacturingProcess33
              + 0.063 ManufacturingProcess32 + 0.079 BiologicalMaterial03

  Rule 3/3: [13 cases, mean 39.391, range 37.86 to 41.43, est err 4.204]

    if
    BiologicalMaterial12 > 19.73
    ManufacturingProcess13 > 34.2
    ManufacturingProcess32 <= 158
    ManufacturingProcess43 <= 0.7
    then
```

```
    outcome = 416.183 - 0.861 ManufacturingProcess32
              - 4.38 ManufacturingProcess13 + 3.45 ManufacturingProcess43
              - 0.1011 ManufacturingProcess05 + 1.55 BiologicalMaterial01

  Rule 3/4: [6 cases, mean 39.553, range 37.64 to 40.66, est err 1.740]

    if
    BiologicalMaterial03 <= 63.44
    BiologicalMaterial11 > 140.84
    then
    outcome = 265.769 - 1.037 BiologicalMaterial11
              + 0.227 BiologicalMaterial03 - 0.0118 ManufacturingProcess19
              - 0.077 ManufacturingProcess32 - 0.137 ManufacturingProcess33

  Rule 3/5: [37 cases, mean 40.552, range 37.86 to 44.16, est err 1.228]

    if
    BiologicalMaterial12 > 19.73
    ManufacturingProcess43 <= 0.7
    then
    outcome = 72.761 + 9.7 ManufacturingProcess43
              + 0.24 ManufacturingProcess32 - 0.305 BiologicalMaterial02
              - 1.18 ManufacturingProcess13 - 0.096 ManufacturingProcess02
              - 0.267 ManufacturingProcess33

  Rule 3/6: [68 cases, mean 40.787, range 38.37 to 43.84, est err 0.845]

    if
    BiologicalMaterial12 > 19.73
    ManufacturingProcess43 > 0.7
    then
    outcome = 12.471 + 0.185 ManufacturingProcess32
              - 0.21 ManufacturingProcess33 + 0.27 ManufacturingProcess09

  Rule 3/7: [11 cases, mean 41.254, range 35.25 to 43.88, est err 1.158]

    if
    ManufacturingProcess13 > 34.2
    ManufacturingProcess32 > 158
    ManufacturingProcess43 <= 0.7
    then
    outcome = 144.699 - 2.51 ManufacturingProcess13
              + 0.367 ManufacturingProcess32 - 0.467 BiologicalMaterial02
              - 0.147 ManufacturingProcess02 - 0.408 ManufacturingProcess33
              - 0.023 ManufacturingProcess05 + 0.38 BiologicalMaterial01

Model 4:

  Rule 4/1: [55 cases, mean 38.931, range 36.77 to 41.43, est err 1.120]

    if
    ManufacturingProcess32 <= 156
    then
    outcome = -20.874 + 0.774 ManufacturingProcess29
```

```
            - 0.00254 ManufacturingProcess26
            + 0.125 ManufacturingProcess32 - 0.45 ManufacturingProcess17
            + 0.059 ManufacturingProcess04

  Rule 4/2: [55 cases, mean 40.397, range 35.25 to 43.88, est err 1.086]

    if
    ManufacturingProcess13 > 33.9
    ManufacturingProcess32 > 156
    then
    outcome = -142.974 - 1.257 BiologicalMaterial04
            + 3.08 BiologicalMaterial10 + 0.987 ManufacturingProcess29
            + 0.297 ManufacturingProcess32 + 0.306 BiologicalMaterial03
            + 0.67 ManufacturingProcess09 + 0.117 ManufacturingProcess04
            - 0.00133 ManufacturingProcess26 - 1.29 BiologicalMaterial09
            - 0.72 BiologicalMaterial01 - 0.23 ManufacturingProcess17

  Rule 4/3: [43 cases, mean 41.308, range 38.42 to 44.16, est err 0.891]

    if
    ManufacturingProcess13 <= 33.9
    then
    outcome = -150.864 - 1.333 BiologicalMaterial04
            + 1.21 ManufacturingProcess29 + 3.26 BiologicalMaterial10
            + 0.299 ManufacturingProcess32 + 0.325 BiologicalMaterial03
            - 0.00191 ManufacturingProcess26 + 0.56 ManufacturingProcess09
            + 0.135 ManufacturingProcess04 - 1.36 BiologicalMaterial09
            - 0.77 BiologicalMaterial01 - 0.34 ManufacturingProcess17
            - 0.03 ManufacturingProcess13

  Rule 4/4: [18 cases, mean 41.345, range 39.4 to 44.16, est err 1.040]

    if
    ManufacturingProcess01 <= 11.1
    ManufacturingProcess13 <= 33.9
    ManufacturingProcess32 > 156
    then
    outcome = 139.501 - 3.94 ManufacturingProcess13
            + 1.19 BiologicalMaterial01 + 0.339 ManufacturingProcess29
            + 0.107 BiologicalMaterial05 + 0.0032 ManufacturingProcess14

Model 5:

  Rule 5/1: [144 cases, mean 40.168, range 35.25 to 44.16, est err 1.097]

    outcome = 0.897 + 0.341 ManufacturingProcess32
            - 0.556 ManufacturingProcess33 + 0.46 ManufacturingProcess09

Model 6:

  Rule 6/1: [38 cases, mean 38.592, range 36.77 to 40.66, est err 1.657]

    if
    BiologicalMaterial12 <= 20.04
```

```
    ManufacturingProcess32 <= 156
    then
    outcome = 40.48 - 0.163 BiologicalMaterial11
              + 0.028 ManufacturingProcess32 - 0.15 ManufacturingProcess13
              + 0.014 ManufacturingProcess04 + 0.0012 ManufacturingProcess15
              + 0.012 BiologicalMaterial06 - 0.09 ManufacturingProcess37

  Rule 6/2: [55 cases, mean 38.931, range 36.77 to 41.43, est err 2.676]

    if
    ManufacturingProcess32 <= 156
    then
    outcome = -39.785 - 2.05 BiologicalMaterial12
              + 0.0164 ManufacturingProcess15 + 0.049 ManufacturingProcess32
              - 0.26 ManufacturingProcess13 + 0.025 ManufacturingProcess04
              + 0.02 BiologicalMaterial06 - 0.15 ManufacturingProcess37

  Rule 6/3: [89 cases, mean 40.932, range 35.25 to 44.16, est err 1.141]

    if
    ManufacturingProcess32 > 156
    then
    outcome = -127.363 + 0.237 ManufacturingProcess32
              - 1.33 ManufacturingProcess13 + 0.128 ManufacturingProcess04
              + 0.006 ManufacturingProcess15 + 0.059 BiologicalMaterial06
              + 0.0037 ManufacturingProcess14 - 0.45 ManufacturingProcess37

Model 7:

  Rule 7/1: [144 cases, mean 40.168, range 35.25 to 44.16, est err 1.061]

    outcome = -44.396 + 0.15 ManufacturingProcess32
              + 13.9 ManufacturingProcess34 + 0.44 ManufacturingProcess09
              + 0.118 BiologicalMaterial06

Model 8:

  Rule 8/1: [63 cases, mean 39.067, range 36.77 to 42.31, est err 0.897]

    if
    ManufacturingProcess32 <= 157
    then
    outcome = 11.278 + 1.262 ManufacturingProcess29
              - 0.00549 ManufacturingProcess27
              + 0.0147 ManufacturingProcess15 + 0.132 ManufacturingProcess32
              - 0.0125 ManufacturingProcess14 - 0.164 BiologicalMaterial02
              - 0.59 ManufacturingProcess13 + 0.122 BiologicalMaterial03
              + 0.064 ManufacturingProcess04 - 0.044 ManufacturingProcess28
              - 0.089 ManufacturingProcess33 - 0.3 ManufacturingProcess07

  Rule 8/2: [81 cases, mean 41.024, range 35.25 to 44.16, est err 0.972]

    if
    ManufacturingProcess32 > 157
```

```
    then
    outcome = -16.756 + 2.087 ManufacturingProcess29
              - 0.00908 ManufacturingProcess27
              + 0.263 ManufacturingProcess32 - 1.17 ManufacturingProcess13
              - 0.27 BiologicalMaterial02 + 0.136 ManufacturingProcess04
              + 0.202 BiologicalMaterial03 - 0.072 ManufacturingProcess28
              - 0.148 ManufacturingProcess33 + 0.006 ManufacturingProcess14
              - 0.5 ManufacturingProcess07

Model 9:

  Rule 9/1: [31 cases, mean 38.515, range 36.77 to 40.66, est err 1.345]

    if
    BiologicalMaterial06 <= 45.48
    then
    outcome = 75.765 - 0.239 BiologicalMaterial11
              - 1.49 ManufacturingProcess44 - 0.05 ManufacturingProcess13

  Rule 9/2: [16 cases, mean 39.242, range 37.86 to 41.43, est err 1.335]

    if
    BiologicalMaterial10 <= 2.27
    then
    outcome = 33.561 + 9.17 BiologicalMaterial10
              - 7.25 ManufacturingProcess44

  Rule 9/3: [42 cases, mean 39.790, range 35.25 to 43.38, est err 1.052]

    if
    ManufacturingProcess17 > 34.8
    then
    outcome = 99.417 - 0.02256 ManufacturingProcess25
              + 1.087 ManufacturingProcess29 + 0.181 ManufacturingProcess32
              + 0.089 ManufacturingProcess02 + 0.22 ManufacturingProcess30
              - 0.00035 ManufacturingProcess26 + 0.039 BiologicalMaterial03
              - 0.034 BiologicalMaterial02 - 0.06 ManufacturingProcess13

  Rule 9/4: [14 cases, mean 39.869, range 37.64 to 42.03, est err 1.002]

    if
    ManufacturingProcess01 <= 10.6
    ManufacturingProcess17 <= 34.8
    then
    outcome = 33.454 - 0.00179 ManufacturingProcess26
              + 0.00233 ManufacturingProcess27 - 0.26 ManufacturingProcess01
              + 0.11 ManufacturingProcess09 + 0.026 ManufacturingProcess32
              - 0.057 BiologicalMaterial04 + 0.02 ManufacturingProcess30

  Rule 9/5: [20 cases, mean 40.031, range 37.86 to 42.73, est err 1.252]

    if
    BiologicalMaterial06 > 45.48
    BiologicalMaterial06 <= 47.12
```

```
    ManufacturingProcess17 <= 34.8
    then
    outcome = 293.612 - 10.67 ManufacturingProcess39
              - 3.839 BiologicalMaterial06 + 0.03 ManufacturingProcess30
              - 5e-005 ManufacturingProcess26 + 0.006 BiologicalMaterial03
              + 0.004 ManufacturingProcess32 - 0.005 BiologicalMaterial02

  Rule 9/6: [88 cases, mean 40.396, range 37.14 to 44.16, est err 1.438]

    if
    ManufacturingProcess01 > 10.6
    ManufacturingProcess17 <= 34.8
    then
    outcome = 33.593 - 0.02384 ManufacturingProcess26
              + 0.01278 ManufacturingProcess27 + 1.04 ManufacturingProcess09
              - 0.848 BiologicalMaterial04 + 1.92 BiologicalMaterial10
              + 0.216 ManufacturingProcess32 + 0.77 ManufacturingProcess13
              - 0.55 ManufacturingProcess17 + 0.57 ManufacturingProcess30
              + 0.087 ManufacturingProcess24 - 0.28 ManufacturingProcess01
              + 0.098 BiologicalMaterial06 + 0.059 BiologicalMaterial03
              - 0.051 BiologicalMaterial02

  Rule 9/7: [19 cases, mean 40.783, range 38.37 to 44.16, est err 1.727]

    if
    BiologicalMaterial06 > 47.12
    ManufacturingProcess01 > 10.6
    ManufacturingProcess17 <= 34.8
    ManufacturingProcess27 > 4580
    then
    outcome = 34.322 - 0.0325 ManufacturingProcess26
              + 0.02669 ManufacturingProcess27 + 1.75 ManufacturingProcess09
              - 1.189 BiologicalMaterial04 + 2.82 BiologicalMaterial10
              + 0.228 ManufacturingProcess32 - 0.72 ManufacturingProcess01
              + 1.13 ManufacturingProcess13 - 0.81 ManufacturingProcess17
              + 0.127 ManufacturingProcess24 + 0.075 BiologicalMaterial11
              - 0.047 ManufacturingProcess04 - 0.01 BiologicalMaterial06

Model 10:

  Rule 10/1: [144 cases, mean 40.168, range 35.25 to 44.16, est err 1.058]

    outcome = -136.226 + 0.264 ManufacturingProcess32
              - 0.8 ManufacturingProcess13 + 0.088 ManufacturingProcess04
              + 0.31 ManufacturingProcess09 + 0.0074 ManufacturingProcess19
              + 0.0044 ManufacturingProcess14

Model 11:

  Rule 11/1: [39 cases, mean 39.879, range 35.25 to 43.38, est err 0.923]

    if
    ManufacturingProcess17 > 34.9
    then
```

```
    outcome = 77.226 - 0.00824 ManufacturingProcess27
              - 0.85 ManufacturingProcess17 + 0.175 ManufacturingProcess32
              + 0.257 BiologicalMaterial04

  Rule 11/2: [10 cases, mean 40.038, range 38.63 to 42.73, est err 1.982]

    if
    BiologicalMaterial12 <= 19.1
    ManufacturingProcess17 <= 34.9
    then
    outcome = -78.916 + 6.39 BiologicalMaterial12

  Rule 11/3: [94 cases, mean 40.186, range 36.77 to 44.16, est err 1.226]

    if
    ManufacturingProcess17 <= 34.9
    ManufacturingProcess39 > 6.9
    then
    outcome = 131.707 - 6.43 ManufacturingProcess39
              - 1.04 ManufacturingProcess17 - 0.13 ManufacturingProcess02
              - 0.388 ManufacturingProcess33 + 0.155 ManufacturingProcess32
              + 0.352 ManufacturingProcess29
              - 0.00118 ManufacturingProcess26 - 0.53 ManufacturingProcess13
              - 0.00119 ManufacturingProcess16 + 1.1 BiologicalMaterial09
              + 0.033 BiologicalMaterial06

  Rule 11/4: [11 cases, mean 41.042, range 38.48 to 43.88, est err 2.939]

    if
    ManufacturingProcess17 <= 34.9
    ManufacturingProcess39 <= 6.9
    then
    outcome = 189.04 - 0.02465 ManufacturingProcess26
              + 0.71 ManufacturingProcess39 + 0.115 ManufacturingProcess29
              - 0.00039 ManufacturingProcess16
              - 0.031 ManufacturingProcess33 - 0.05 ManufacturingProcess17
              + 0.011 ManufacturingProcess32 + 0.011 BiologicalMaterial06

Model 12:

  Rule 12/1: [144 cases, mean 40.168, range 35.25 to 44.16, est err 0.957]

    outcome = -27.3 + 0.259 ManufacturingProcess32
              + 0.58 ManufacturingProcess09

Model 13:

  Rule 13/1: [144 cases, mean 40.168, range 35.25 to 44.16, est err 1.119]

    outcome = 35.817 + 2.553 ManufacturingProcess29
              - 0.0111 ManufacturingProcess27 - 0.301 BiologicalMaterial02
              + 0.295 BiologicalMaterial03 - 0.65 ManufacturingProcess17
              + 9.3 ManufacturingProcess34
```

```
Model 14:

  Rule 14/1: [41 cases, mean 38.769, range 36.77 to 40.87, est err 1.725]

    if
    BiologicalMaterial10 > 2.27
    ManufacturingProcess32 <= 156
    then
    outcome = 48.372 - 0.504 BiologicalMaterial11
              + 0.87 BiologicalMaterial04 + 2.01 BiologicalMaterial12
              + 0.022 ManufacturingProcess32 + 0.04 ManufacturingProcess09
              + 0.008 ManufacturingProcess04 - 0.05 BiologicalMaterial09

  Rule 14/2: [14 cases, mean 39.404, range 37.86 to 41.43, est err 1.106]

    if
    BiologicalMaterial10 <= 2.27
    ManufacturingProcess32 <= 156
    then
    outcome = -255.113 + 0.05299 ManufacturingProcess20
              + 0.366 BiologicalMaterial11

  Rule 14/3: [89 cases, mean 40.932, range 35.25 to 44.16, est err 1.191]

    if
    ManufacturingProcess32 > 156
    then
    outcome = -190.005 + 0.366 ManufacturingProcess32
              + 0.67 ManufacturingProcess09 + 0.155 ManufacturingProcess04
              - 0.2 BiologicalMaterial09

Model 15:

  Rule 15/1: [49 cases, mean 39.895, range 37.86 to 43.88, est err 0.928]

    if
    BiologicalMaterial05 <= 18.04
    ManufacturingProcess13 > 32.9
    then
    outcome = 47.052 - 0.258 BiologicalMaterial06
              + 0.208 BiologicalMaterial03 + 0.447 BiologicalMaterial05
              - 1.72 BiologicalMaterial09 + 0.92 BiologicalMaterial01

  Rule 15/2: [91 cases, mean 40.220, range 35.25 to 44.16, est err 1.272]

    if
    BiologicalMaterial05 > 18.04
    then
    outcome = 17.951 - 0.478 ManufacturingProcess33
              - 1.16 ManufacturingProcess13 + 0.166 ManufacturingProcess32
              + 0.355 BiologicalMaterial05 + 0.0098 ManufacturingProcess15

  Rule 15/3: [16 cases, mean 40.531, range 37.86 to 43.88, est err 2.120]
```

```
    if
    BiologicalMaterial05 <= 18.04
    BiologicalMaterial09 <= 12.92
    ManufacturingProcess13 > 32.9
    then
    outcome = 3.604 + 0.657 BiologicalMaterial03
              - 0.424 BiologicalMaterial06 + 0.734 BiologicalMaterial05
              - 2.82 BiologicalMaterial09 + 1.52 BiologicalMaterial01
              + 0.63 ManufacturingProcess09

  Rule 15/4: [13 cases, mean 41.981, range 38.48 to 44.16, est err 1.497]

    if
    ManufacturingProcess13 <= 32.9
    then
    outcome = 37.714 + 0.574 ManufacturingProcess28
              - 0.03 ManufacturingProcess13

Model 16:

  Rule 16/1: [144 cases, mean 40.168, range 35.25 to 44.16, est err 1.115]

    outcome = -129.386 + 0.371 ManufacturingProcess32
              + 0.65 ManufacturingProcess09 + 0.087 ManufacturingProcess04

Model 17:

  Rule 17/1: [67 cases, mean 39.868, range 36.77 to 44.16, est err 1.624]

    if
    BiologicalMaterial04 <= 12.7
    ManufacturingProcess01 > 10.4
    ManufacturingProcess17 <= 34.9
    then
    outcome = 121.475 - 2.42 ManufacturingProcess13

  Rule 17/2: [39 cases, mean 39.879, range 35.25 to 43.38, est err 1.025]

    if
    ManufacturingProcess17 > 34.9
    then
    outcome = 64.185 - 0.67 ManufacturingProcess17
              + 0.086 BiologicalMaterial02 - 0.15 ManufacturingProcess13

  Rule 17/3: [12 cases, mean 39.897, range 37.64 to 42.03, est err 1.509]

    if
    ManufacturingProcess01 <= 10.4
    ManufacturingProcess17 <= 34.9
    then
    outcome = 38.855

  Rule 17/4: [10 cases, mean 40.038, range 38.63 to 42.73, est err 2.113]
```

```
    if
    BiologicalMaterial12 <= 19.1
    ManufacturingProcess17 <= 34.9
    then
    outcome = -31.98 + 3.96 BiologicalMaterial12
              - 0.03 ManufacturingProcess13

  Rule 17/5: [20 cases, mean 40.257, range 37.73 to 43.42, est err 1.621]

    if
    BiologicalMaterial12 > 19.1
    ManufacturingProcess17 <= 34.9
    ManufacturingProcess21 <= -0.8
    ManufacturingProcess39 > 6.9
    then
    outcome = 124.616 - 10.31 ManufacturingProcess39
              + 3.9 ManufacturingProcess21 - 0.21 ManufacturingProcess13

  Rule 17/6: [11 cases, mean 41.042, range 38.48 to 43.88, est err 3.566]

    if
    ManufacturingProcess17 <= 34.9
    ManufacturingProcess39 <= 6.9
    then
    outcome = 84.587 - 2.61 ManufacturingProcess30
              + 0.92 ManufacturingProcess39 - 0.66 ManufacturingProcess17

  Rule 17/7: [42 cases, mean 41.073, range 38 to 43.62, est err 1.547]

    if
    BiologicalMaterial04 > 12.7
    ManufacturingProcess21 > -0.8
    then
    outcome = -0.115 + 1.11 ManufacturingProcess09
              + 0.704 ManufacturingProcess29 - 1.45 ManufacturingProcess21
              - 0.76 BiologicalMaterial12 - 0.25 ManufacturingProcess13

  Rule 17/8: [14 cases, mean 41.412, range 39.49 to 43.84, est err 1.996]

    if
    BiologicalMaterial12 > 19.67
    ManufacturingProcess17 <= 34.9
    ManufacturingProcess21 <= -0.8
    then
    outcome = 64.509 - 0.895 BiologicalMaterial04
              - 0.29 ManufacturingProcess13 - 0.14 ManufacturingProcess39
              + 0.05 ManufacturingProcess21

Model 18:

  Rule 18/1: [144 cases, mean 40.168, range 35.25 to 44.16, est err 1.112]

    outcome = -138.858 + 0.341 ManufacturingProcess32
              + 0.56 ManufacturingProcess09 + 0.107 ManufacturingProcess04
```

```
Model 19:

  Rule 19/1: [16 cases, mean 39.364, range 35.25 to 42.68, est err 1.234]

    if
    ManufacturingProcess13 > 35.4
    then
    outcome = 102.15 - 2.05 ManufacturingProcess13
              + 0.216 BiologicalMaterial06

  Rule 19/2: [10 cases, mean 40.038, range 38.63 to 42.73, est err 1.555]

    if
    BiologicalMaterial12 <= 19.1
    ManufacturingProcess13 <= 35.4
    then
    outcome = -39.409 + 4.33 BiologicalMaterial12
              - 0.03 ManufacturingProcess13

  Rule 19/3: [106 cases, mean 40.173, range 37.51 to 44.16, est err 1.295]

    if
    BiologicalMaterial05 <= 19.51
    then
    outcome = 165.077 - 4.85 ManufacturingProcess39
              + 2.166 ManufacturingProcess29
              - 0.00835 ManufacturingProcess25 - 0.437 BiologicalMaterial11
              + 2.6 BiologicalMaterial12 + 0.3 BiologicalMaterial03
              + 0.627 BiologicalMaterial05 - 0.0188 ManufacturingProcess15
              - 1.16 ManufacturingProcess21

  Rule 19/4: [128 cases, mean 40.268, range 36.77 to 44.16, est err 1.869]

    if
    ManufacturingProcess13 <= 35.4
    then
    outcome = 46.709 + 0.45 ManufacturingProcess39
              - 0.24 ManufacturingProcess13

  Rule 19/5: [109 cases, mean 40.281, range 36.77 to 44.16, est err 1.661]

    if
    BiologicalMaterial12 > 19.1
    ManufacturingProcess13 <= 35.4
    ManufacturingProcess39 > 6.8
    then
    outcome = 101.864 - 9.48 ManufacturingProcess39
              + 0.01128 ManufacturingProcess16 - 2.23 ManufacturingProcess13
              + 1.51 BiologicalMaterial12

  Rule 19/6: [50 cases, mean 40.353, range 37.51 to 44.16, est err 1.125]

    if
```

```
    BiologicalMaterial05 <= 19.51
    ManufacturingProcess13 <= 35.4
    ManufacturingProcess21 > -0.4
    ManufacturingProcess39 > 6.8
    then
    outcome = 136.427 - 5.02 ManufacturingProcess39
              + 0.896 ManufacturingProcess29
              - 0.00345 ManufacturingProcess25
              - 0.274 ManufacturingProcess22 - 0.181 BiologicalMaterial11
              + 0.89 BiologicalMaterial12 + 0.124 BiologicalMaterial03
              - 0.51 ManufacturingProcess13 + 0.26 BiologicalMaterial05
              - 0.0078 ManufacturingProcess15 - 0.48 ManufacturingProcess21

  Rule 19/7: [24 cases, mean 40.370, range 37.86 to 43.84, est err 1.273]

    if
    BiologicalMaterial05 <= 19.51
    BiologicalMaterial12 > 19.1
    ManufacturingProcess21 <= -0.4
    ManufacturingProcess22 > 4
    then
    outcome = 238.069 - 6.54 ManufacturingProcess39
              + 3.63 ManufacturingProcess29 - 5.96 ManufacturingProcess21
              - 0.46 ManufacturingProcess22 - 0.0336 ManufacturingProcess19
              - 0.00114 ManufacturingProcess25 - 0.06 BiologicalMaterial11
              + 0.29 BiologicalMaterial12 + 0.041 BiologicalMaterial03
              + 0.086 BiologicalMaterial05 - 0.0026 ManufacturingProcess15
              - 0.04 ManufacturingProcess13

  Rule 19/8: [9 cases, mean 40.682, range 38.37 to 43.84, est err 1.103]

    if
    BiologicalMaterial05 <= 19.51
    BiologicalMaterial12 > 19.1
    ManufacturingProcess19 > 6022
    ManufacturingProcess21 <= -0.4
    ManufacturingProcess22 > 4
    then
    outcome = 232.529 - 6.5 ManufacturingProcess39
              + 3.517 ManufacturingProcess29 - 5.77 ManufacturingProcess21
              - 0.459 ManufacturingProcess22 - 0.0325 ManufacturingProcess19
              - 0.00111 ManufacturingProcess25 - 0.058 BiologicalMaterial11
              + 0.29 BiologicalMaterial12 + 0.04 BiologicalMaterial03
              + 0.084 BiologicalMaterial05 - 0.0025 ManufacturingProcess15
              - 0.07 ManufacturingProcess13

Model 20:

  Rule 20/1: [144 cases, mean 40.168, range 35.25 to 44.16, est err 1.048]

    outcome = -35.535 + 0.315 ManufacturingProcess32
              + 0.56 ManufacturingProcess09

Model 21:
```

```
  Rule 21/1: [144 cases, mean 40.168, range 35.25 to 44.16, est err 1.286]

    outcome = 11.032 + 1.563 ManufacturingProcess29
              - 0.00688 ManufacturingProcess27 - 0.39 BiologicalMaterial02
              + 0.332 BiologicalMaterial03 - 0.482 ManufacturingProcess33
              + 0.192 ManufacturingProcess32 - 0.94 ManufacturingProcess13
              + 0.0102 ManufacturingProcess15

Model 22:

  Rule 22/1: [43 cases, mean 38.626, range 36.77 to 40.66, est err 1.887]

    if
    BiologicalMaterial12 <= 19.96
    ManufacturingProcess32 <= 158
    then
    outcome = 53.423 - 2.02 BiologicalMaterial12 + 0.4 BiologicalMaterial06
              + 0.032 ManufacturingProcess32

  Rule 22/2: [74 cases, mean 39.157, range 36.77 to 42.31, est err 1.168]

    if
    ManufacturingProcess32 <= 158
    then
    outcome = -128.715 + 0.178 ManufacturingProcess32
              + 0.137 ManufacturingProcess04 + 0.28 ManufacturingProcess09

  Rule 22/3: [70 cases, mean 41.236, range 35.25 to 44.16, est err 1.283]

    if
    ManufacturingProcess32 > 158
    then
    outcome = -101.227 + 0.31 ManufacturingProcess32
              - 1.24 ManufacturingProcess17 + 0.207 ManufacturingProcess04
              - 0.12 ManufacturingProcess02 - 0.0124 ManufacturingProcess15
              + 0.44 ManufacturingProcess13 + 0.09 ManufacturingProcess09

Model 23:

  Rule 23/1: [144 cases, mean 40.168, range 35.25 to 44.16, est err 1.326]

    outcome = 74.879 - 1.01 ManufacturingProcess13

Model 24:

  Rule 24/1: [144 cases, mean 40.168, range 35.25 to 44.16, est err 1.201]

    outcome = -45.473 + 0.377 ManufacturingProcess32
              + 0.57 ManufacturingProcess09

Model 25:

  Rule 25/1: [39 cases, mean 39.879, range 35.25 to 43.38, est err 0.802]
```

```
  if
  ManufacturingProcess17 > 34.9
  then
  outcome = -12.902 + 1.373 ManufacturingProcess29
            - 0.00598 ManufacturingProcess27 - 0.87 ManufacturingProcess17
            - 0.358 ManufacturingProcess33 + 0.134 ManufacturingProcess32
            + 0.091 ManufacturingProcess04

Rule 25/2: [57 cases, mean 39.911, range 36.77 to 44.16, est err 1.267]

  if
  BiologicalMaterial04 <= 13.02
  BiologicalMaterial12 > 19.1
  ManufacturingProcess01 > 10.4
  ManufacturingProcess17 <= 34.9
  ManufacturingProcess39 > 6.9
  then
  outcome = 81.112 - 1.88 ManufacturingProcess17
            - 1.3 ManufacturingProcess39 + 2.38 BiologicalMaterial12
            - 0.365 BiologicalMaterial06 + 0.000411 ManufacturingProcess12
            + 0.24 ManufacturingProcess21 + 0.03 ManufacturingProcess10

Rule 25/3: [11 cases, mean 40.009, range 37.64 to 42.03, est err 2.316]

  if
  ManufacturingProcess01 <= 10.4
  ManufacturingProcess17 <= 34.9
  ManufacturingProcess21 > -0.8
  then
  outcome = 103.628 - 2.63 ManufacturingProcess39
            - 1.66 ManufacturingProcess17 + 0.684 ManufacturingProcess29
            - 0.0021 ManufacturingProcess26 + 1.27 ManufacturingProcess21
            + 0.77 BiologicalMaterial12 - 0.000201 ManufacturingProcess12
            + 0.07 BiologicalMaterial03 - 0.106 ManufacturingProcess33
            - 0.052 BiologicalMaterial11 - 0.00023 ManufacturingProcess27
            + 0.005 ManufacturingProcess32 + 0.004 ManufacturingProcess04

Rule 25/4: [10 cases, mean 40.038, range 38.63 to 42.73, est err 3.103]

  if
  BiologicalMaterial12 <= 19.1
  ManufacturingProcess17 <= 34.9
  then
  outcome = -13.635 + 3.39 BiologicalMaterial12
            - 1.07 ManufacturingProcess17 + 0.447 ManufacturingProcess29
            - 0.00194 ManufacturingProcess27 - 0.29 ManufacturingProcess39
            - 0.117 ManufacturingProcess33 + 0.044 ManufacturingProcess32
            + 0.029 ManufacturingProcess04 + 0.06 ManufacturingProcess10
            + 0.0005 ManufacturingProcess15

Rule 25/5: [94 cases, mean 40.186, range 36.77 to 44.16, est err 1.159]

  if
```

```
    ManufacturingProcess17 <= 34.9
    ManufacturingProcess39 > 6.9
    then
    outcome = 98.774 - 6.8 ManufacturingProcess39
              - 0.96 ManufacturingProcess17 + 0.75 ManufacturingProcess10
              + 0.78 BiologicalMaterial12

  Rule 25/6: [11 cases, mean 41.042, range 38.48 to 43.88, est err 3.032]

    if
    ManufacturingProcess17 <= 34.9
    ManufacturingProcess39 <= 6.9
    then
    outcome = 146.69 - 0.01586 ManufacturingProcess26
              + 0.7 ManufacturingProcess39 - 1.15 ManufacturingProcess10
              + 0.109 ManufacturingProcess29
              - 0.00016 ManufacturingProcess27 - 0.02 ManufacturingProcess33
              - 0.04 ManufacturingProcess17 + 0.008 BiologicalMaterial03
              - 0.006 BiologicalMaterial11 + 0.004 ManufacturingProcess32

  Rule 25/7: [22 cases, mean 41.237, range 39.4 to 43.62, est err 1.056]

    if
    BiologicalMaterial04 > 13.02
    ManufacturingProcess17 <= 34.9
    ManufacturingProcess21 > -0.8
    ManufacturingProcess39 > 6.9
    then
    outcome = 42.697 - 3.56 ManufacturingProcess39
              - 1.1 ManufacturingProcess17 + 0.663 ManufacturingProcess29
              - 0.00162 ManufacturingProcess27 + 0.72 ManufacturingProcess10
              - 0.00097 ManufacturingProcess26 - 0.14 ManufacturingProcess33
              + 0.0056 ManufacturingProcess15 + 0.036 ManufacturingProcess32
              + 0.025 ManufacturingProcess04 + 0.032 BiologicalMaterial03
              - 0.024 BiologicalMaterial11 + 0.13 ManufacturingProcess21
              + 0.11 BiologicalMaterial12

  Rule 25/8: [8 cases, mean 41.659, range 40.54 to 43.42, est err 3.710]

    if
    BiologicalMaterial12 > 19.1
    ManufacturingProcess10 > 8.7
    ManufacturingProcess21 <= -0.8
    ManufacturingProcess39 > 6.9
    then
    outcome = 48.971 - 2.2 ManufacturingProcess39
              + 0.5 ManufacturingProcess10 + 0.143 ManufacturingProcess29
              - 0.00062 ManufacturingProcess27
              - 0.037 ManufacturingProcess33 + 0.014 ManufacturingProcess32
              - 0.05 ManufacturingProcess17 + 0.009 ManufacturingProcess04

Model 26:

  Rule 26/1: [144 cases, mean 40.168, range 35.25 to 44.16, est err 1.114]
```

```
    outcome = -39.044 + 0.337 ManufacturingProcess32
              + 0.57 ManufacturingProcess09

Model 27:

  Rule 27/1: [144 cases, mean 40.168, range 35.25 to 44.16, est err 1.354]

    outcome = -56.453 + 2.927 ManufacturingProcess29
              - 0.00962 ManufacturingProcess26 + 0.47 BiologicalMaterial06
              - 0.311 BiologicalMaterial02 + 0.131 ManufacturingProcess02
              + 0.122 ManufacturingProcess04 - 0.75 ManufacturingProcess13

Model 28:

  Rule 28/1: [144 cases, mean 40.168, range 35.25 to 44.16, est err 1.180]

    outcome = -42.763 + 0.532 ManufacturingProcess32
              - 0.456 ManufacturingProcess33 + 0.61 ManufacturingProcess09

Model 29:

  Rule 29/1: [34 cases, mean 39.903, range 35.25 to 43.38, est err 1.146]

    if
    ManufacturingProcess17 > 35
    then
    outcome = -4.516 - 0.89 ManufacturingProcess17
              + 0.0099 ManufacturingProcess19 + 0.22 ManufacturingProcess10
              + 0.0027 ManufacturingProcess14 + 0.014 BiologicalMaterial03

  Rule 29/2: [14 cases, mean 39.951, range 37.64 to 42.03, est err 4.330]

    if
    ManufacturingProcess01 <= 10.4
    ManufacturingProcess17 <= 35
    then
    outcome = 59.487 - 9.12 ManufacturingProcess39
              + 0.0369 ManufacturingProcess05 - 0.468 BiologicalMaterial05
              + 0.46 ManufacturingProcess17

  Rule 29/3: [10 cases, mean 40.038, range 38.63 to 42.73, est err 3.667]

    if
    BiologicalMaterial12 <= 19.1
    ManufacturingProcess17 <= 35
    then
    outcome = -48.319 - 5.06 BiologicalMaterial02
              + 3.18 BiologicalMaterial03 + 7.73 BiologicalMaterial12

  Rule 29/4: [99 cases, mean 40.162, range 36.77 to 44.16, est err 1.355]

    if
    ManufacturingProcess17 <= 35
```

```
    ManufacturingProcess39 > 6.9
    then
    outcome = 133.602 - 5.89 ManufacturingProcess39
              + 1.818 ManufacturingProcess29
              - 0.00514 ManufacturingProcess25 - 1.14 ManufacturingProcess17
              + 0.94 BiologicalMaterial12 - 0.0072 ManufacturingProcess14
              - 0.44 ManufacturingProcess10 + 0.18 ManufacturingProcess01
              - 0.039 ManufacturingProcess32

  Rule 29/5: [11 cases, mean 41.042, range 38.48 to 43.88, est err 1.597]

    if
    ManufacturingProcess17 <= 35
    ManufacturingProcess39 <= 6.9
    then
    outcome = 24.272 - 0.01383 ManufacturingProcess27
              + 0.465 ManufacturingProcess29 + 0.95 ManufacturingProcess10
              + 0.0132 ManufacturingProcess14 + 0.169 BiologicalMaterial03
              + 0.45 ManufacturingProcess39 - 0.15 BiologicalMaterial02
              - 0.31 BiologicalMaterial12

Model 30:

  Rule 30/1: [144 cases, mean 40.168, range 35.25 to 44.16, est err 1.213]

    outcome = -151.164 + 0.556 ManufacturingProcess32
              - 0.446 ManufacturingProcess33
              - 0.00217 ManufacturingProcess26 + 0.63 ManufacturingProcess09
              + 0.116 ManufacturingProcess04 + 0.401 ManufacturingProcess29

Model 31:

  Rule 31/1: [39 cases, mean 39.879, range 35.25 to 43.38, est err 1.054]

    if
    ManufacturingProcess17 > 34.9
    then
    outcome = 46.508 - 0.73 ManufacturingProcess17
              + 0.8 ManufacturingProcess43 + 0.113 ManufacturingProcess32
              + 0.08 ManufacturingProcess29 - 0.00021 ManufacturingProcess26

  Rule 31/2: [10 cases, mean 40.038, range 38.63 to 42.73, est err 2.068]

    if
    BiologicalMaterial12 <= 19.1
    ManufacturingProcess17 <= 34.9
    then
    outcome = -26.34 + 3.65 BiologicalMaterial12
              + 0.052 ManufacturingProcess29
              - 0.00015 ManufacturingProcess26 - 0.03 ManufacturingProcess17

  Rule 31/3: [94 cases, mean 40.186, range 36.77 to 44.16, est err 1.574]

    if
```

```
    ManufacturingProcess17 <= 34.9
    ManufacturingProcess39 > 6.9
    then
    outcome = 126.943 - 7 ManufacturingProcess39
              + 2.658 ManufacturingProcess29
              - 0.00784 ManufacturingProcess26 - 1.31 ManufacturingProcess17
              + 0.204 BiologicalMaterial06 - 0.126 ManufacturingProcess32
              + 0.91 BiologicalMaterial09

  Rule 31/4: [11 cases, mean 41.042, range 38.48 to 43.88, est err 2.406]

    if
    ManufacturingProcess17 <= 34.9
    ManufacturingProcess39 <= 6.9
    then
    outcome = 41.171 - 3.365 ManufacturingProcess29
              + 3.15 BiologicalMaterial12 + 1.06 ManufacturingProcess39

Model 32:

  Rule 32/1: [83 cases, mean 39.239, range 35.25 to 43.12, est err 1.081]

    if
    ManufacturingProcess32 <= 159
    then
    outcome = -18.206 + 0.41 ManufacturingProcess32
              - 0.393 ManufacturingProcess33 + 0.54 ManufacturingProcess09
              - 0.00088 ManufacturingProcess26 - 0.03 ManufacturingProcess13

  Rule 32/2: [61 cases, mean 41.432, range 36.83 to 44.16, est err 1.131]

    if
    ManufacturingProcess32 > 159
    then
    outcome = -20.583 + 0.248 ManufacturingProcess32
              + 0.48 ManufacturingProcess09

Model 33:

  Rule 33/1: [144 cases, mean 40.168, range 35.25 to 44.16, est err 1.158]

    outcome = -55.69 + 1.811 ManufacturingProcess29
              - 0.00516 ManufacturingProcess26 - 0.92 ManufacturingProcess17
              + 0.113 ManufacturingProcess04 + 0.107 ManufacturingProcess32

Model 34:

  Rule 34/1: [36 cases, mean 38.540, range 36.77 to 40.66, est err 1.636]

    if
    BiologicalMaterial12 <= 19.96
    ManufacturingProcess01 > 11.4
    ManufacturingProcess32 <= 161
    then
```

```
  outcome = -19.596 - 0.601 ManufacturingProcess33
            + 0.26 ManufacturingProcess32 + 1.01 ManufacturingProcess17
            + 0.76 ManufacturingProcess09 - 0.00163 ManufacturingProcess27
            + 0.00145 ManufacturingProcess25 - 0.51 BiologicalMaterial12
            - 0.059 ManufacturingProcess28 + 0.15 ManufacturingProcess43
            + 0.035 BiologicalMaterial06 - 0.12 ManufacturingProcess13
            - 0.026 BiologicalMaterial03

Rule 34/2: [53 cases, mean 38.882, range 35.25 to 42.58, est err 2.086]

  if
  BiologicalMaterial12 <= 19.96
  ManufacturingProcess32 <= 161
  then
  outcome = 160.299 - 2.69 ManufacturingProcess17
            - 1.86 ManufacturingProcess09 + 0.479 ManufacturingProcess32
            - 0.175 ManufacturingProcess33 - 0.29 BiologicalMaterial12
            - 0.00048 ManufacturingProcess27
            + 0.00043 ManufacturingProcess25 - 0.03 ManufacturingProcess28
            - 0.04 ManufacturingProcess13 + 0.04 ManufacturingProcess43
            + 0.01 BiologicalMaterial06 - 0.008 BiologicalMaterial03

Rule 34/3: [106 cases, mean 39.576, range 35.25 to 44.16, est err 1.182]

  if
  ManufacturingProcess32 <= 161
  then
  outcome = 57.18 - 0.00976 ManufacturingProcess27
            + 0.00832 ManufacturingProcess25
            + 0.347 ManufacturingProcess32 - 0.49 ManufacturingProcess33
            + 0.85 ManufacturingProcess43 + 0.201 BiologicalMaterial06
            - 0.72 ManufacturingProcess13 - 0.149 BiologicalMaterial03
            - 0.074 ManufacturingProcess28 - 0.32 ManufacturingProcess17

Rule 34/4: [14 cases, mean 40.898, range 38.03 to 43.84, est err 1.913]

  if
  ManufacturingProcess09 > 47.33
  then
  outcome = -196.084 + 3.67 ManufacturingProcess09
            + 0.594 ManufacturingProcess32 - 1.281 BiologicalMaterial04
            - 0.00276 ManufacturingProcess27
            + 0.00241 ManufacturingProcess25 - 0.14 ManufacturingProcess33
            + 0.059 BiologicalMaterial06 - 0.21 ManufacturingProcess13
            - 0.049 BiologicalMaterial03

Rule 34/5: [38 cases, mean 41.819, range 38.95 to 43.88, est err 1.119]

  if
  ManufacturingProcess32 > 161
  then
  outcome = 25.487 + 0.113 ManufacturingProcess32
            - 0.121 ManufacturingProcess33 + 0.14 ManufacturingProcess09
```

```
Model 35:

  Rule 35/1: [144 cases, mean 40.168, range 35.25 to 44.16, est err 1.160]

    outcome = -108.54 + 2.496 ManufacturingProcess29
              - 0.00977 ManufacturingProcess27
              + 0.255 ManufacturingProcess32 - 0.91 ManufacturingProcess17
              - 0.276 BiologicalMaterial02 + 0.294 BiologicalMaterial06
              + 0.167 ManufacturingProcess04 - 0.326 ManufacturingProcess33

Model 36:

  Rule 36/1: [49 cases, mean 38.822, range 35.25 to 42.58, est err 1.340]

    if
    BiologicalMaterial12 <= 19.96
    ManufacturingProcess09 <= 47.33
    ManufacturingProcess32 <= 161
    then
    outcome = 120.543 - 0.00747 ManufacturingProcess25
              - 0.374 BiologicalMaterial11 + 0.186 BiologicalMaterial03
              - 0.33 ManufacturingProcess01 - 0.03 ManufacturingProcess13

  Rule 36/2: [13 cases, mean 39.602, range 35.25 to 42.58, est err 2.044]

    if
    BiologicalMaterial12 <= 19.96
    ManufacturingProcess01 <= 11.4
    ManufacturingProcess09 <= 47.33
    ManufacturingProcess32 <= 161
    then
    outcome = -112.6 + 2.15 ManufacturingProcess01
              + 0.573 ManufacturingProcess32 + 2.06 BiologicalMaterial12

  Rule 36/3: [53 cases, mean 40.270, range 38.35 to 44.16, est err 1.413]

    if
    BiologicalMaterial12 > 19.96
    ManufacturingProcess32 <= 161
    then
    outcome = 60.168 + 19.84 ManufacturingProcess17
              - 20.62 ManufacturingProcess13 - 20.61 ManufacturingProcess21
              + 0.232 ManufacturingProcess32 - 0.198 BiologicalMaterial03
              - 0.54 BiologicalMaterial12 - 0.00052 ManufacturingProcess26
              - 0.0002 ManufacturingProcess25

  Rule 36/4: [14 cases, mean 40.898, range 38.03 to 43.84, est err 2.070]

    if
    ManufacturingProcess09 > 47.33
    then
    outcome = -252.993 + 5.36 ManufacturingProcess09
              + 0.457 ManufacturingProcess32 - 2.18 ManufacturingProcess11
              - 2.04 BiologicalMaterial01
```

```
  Rule 36/5: [38 cases, mean 41.819, range 38.95 to 43.88, est err 1.261]

    if
    ManufacturingProcess32 > 161
    then
    outcome = 49.259 - 0.146 ManufacturingProcess22
              - 0.17 ManufacturingProcess13

Model 37:

  Rule 37/1: [144 cases, mean 40.168, range 35.25 to 44.16, est err 1.280]

    outcome = -90.216 + 2.248 ManufacturingProcess29
              - 0.00823 ManufacturingProcess25 + 0.361 BiologicalMaterial03
              - 0.344 BiologicalMaterial02 + 0.619 ManufacturingProcess42
              - 2.6 ManufacturingProcess44 - 0.68 ManufacturingProcess17
              + 0.146 ManufacturingProcess32 + 0.125 ManufacturingProcess04
              + 0.25 ManufacturingProcess09 - 0.8 BiologicalMaterial09

Model 38:

  Rule 38/1: [64 cases, mean 39.358, range 35.25 to 43.88, est err 1.399]

    if
    BiologicalMaterial12 <= 19.96
    then
    outcome = 21.437 - 423.58 ManufacturingProcess17
              + 423.43 ManufacturingProcess13
              + 423.54 ManufacturingProcess21 + 0.557 ManufacturingProcess32
              - 0.675 ManufacturingProcess33 - 0.211 BiologicalMaterial11
              + 0.75 ManufacturingProcess11 + 0.24 BiologicalMaterial12
              - 0.14 ManufacturingProcess30 + 0.051 BiologicalMaterial04
              - 0.006 ManufacturingProcess35 - 0.01 ManufacturingProcess24

  Rule 38/2: [106 cases, mean 39.576, range 35.25 to 44.16, est err 1.759]

    if
    ManufacturingProcess32 <= 161
    then
    outcome = 43.353 - 846.55 ManufacturingProcess17
              + 846.25 ManufacturingProcess13
              + 846.47 ManufacturingProcess21 + 0.253 ManufacturingProcess32
              - 0.287 ManufacturingProcess33 - 0.103 BiologicalMaterial11
              + 0.22 ManufacturingProcess09 - 0.28 ManufacturingProcess30
              + 0.39 ManufacturingProcess11 + 0.102 BiologicalMaterial04
              - 0.013 ManufacturingProcess35 - 0.2 BiologicalMaterial12
              - 0.02 ManufacturingProcess24

  Rule 38/3: [38 cases, mean 41.819, range 38.95 to 43.88, est err 1.377]

    if
    ManufacturingProcess32 > 161
    then
```

```
    outcome = 53.931 - 0.32 ManufacturingProcess13

Model 39:

  Rule 39/1: [55 cases, mean 38.931, range 36.77 to 41.43, est err 1.240]

    if
    ManufacturingProcess32 <= 156
    then
    outcome = -0.452 + 0.0074 ManufacturingProcess15
              - 0.18 ManufacturingProcess13

  Rule 39/2: [89 cases, mean 40.932, range 35.25 to 44.16, est err 1.295]

    if
    ManufacturingProcess32 > 156
    then
    outcome = -81.57 + 0.0181 ManufacturingProcess27
              + 0.523 BiologicalMaterial03 - 0.451 BiologicalMaterial02
              + 1.57 ManufacturingProcess30 + 0.23 ManufacturingProcess32
              - 0.159 ManufacturingProcess28 - 0.63 ManufacturingProcess13

Model 40:

  Rule 40/1: [31 cases, mean 38.515, range 36.77 to 40.66, est err 0.759]

    if
    BiologicalMaterial06 <= 45.48
    then
    outcome = 90.391 - 0.0386 ManufacturingProcess05
              - 0.1 ManufacturingProcess32 - 0.33 ManufacturingProcess13
              + 4.6 ManufacturingProcess34 + 0.038 BiologicalMaterial02

  Rule 40/2: [113 cases, mean 40.621, range 35.25 to 44.16, est err 1.308]

    if
    BiologicalMaterial06 > 45.48
    then
    outcome = -40.371 + 0.164 ManufacturingProcess32
              + 0.33 ManufacturingProcess09 + 0.042 ManufacturingProcess04
              - 0.24 ManufacturingProcess13 + 3.3 ManufacturingProcess34
              + 0.027 BiologicalMaterial02

Model 41:

  Rule 41/1: [55 cases, mean 38.931, range 36.77 to 41.43, est err 0.964]

    if
    ManufacturingProcess32 <= 156
    then
    outcome = -16.167 + 0.178 ManufacturingProcess32
              + 0.0125 ManufacturingProcess15
              - 0.0112 ManufacturingProcess14 - 0.202 ManufacturingProcess33
              - 0.092 BiologicalMaterial02 - 0.24 ManufacturingProcess17
```

```
            + 0.069 BiologicalMaterial03 + 0.029 ManufacturingProcess04

  Rule 41/2: [89 cases, mean 40.932, range 35.25 to 44.16, est err 1.274]

    if
    ManufacturingProcess32 > 156
    then
    outcome = -77.387 + 0.352 ManufacturingProcess32
              + 0.0216 ManufacturingProcess15
              - 0.0195 ManufacturingProcess14 - 0.349 ManufacturingProcess33
              - 0.61 ManufacturingProcess17 + 0.00171 ManufacturingProcess27
              - 0.159 BiologicalMaterial02 + 0.119 BiologicalMaterial03
              + 0.066 ManufacturingProcess04

Model 42:

  Rule 42/1: [31 cases, mean 38.515, range 36.77 to 40.66, est err 0.816]

    if
    BiologicalMaterial06 <= 45.48
    then
    outcome = 66.89 - 0.449 BiologicalMaterial06
              - 0.26 ManufacturingProcess13

  Rule 42/2: [113 cases, mean 40.621, range 35.25 to 44.16, est err 1.461]

    if
    BiologicalMaterial06 > 45.48
    then
    outcome = 69.484 - 0.81 ManufacturingProcess13

Model 43:

  Rule 43/1: [74 cases, mean 39.157, range 36.77 to 42.31, est err 1.162]

    if
    ManufacturingProcess32 <= 158
    then
    outcome = -46.486 + 0.343 ManufacturingProcess32
              + 0.0236 ManufacturingProcess15 - 0.296 BiologicalMaterial02
              - 0.386 ManufacturingProcess33 - 0.0137 ManufacturingProcess14
              + 0.176 BiologicalMaterial03 - 0.46 ManufacturingProcess13

  Rule 43/2: [70 cases, mean 41.236, range 35.25 to 44.16, est err 1.180]

    if
    ManufacturingProcess32 > 158
    then
    outcome = -70.226 + 0.454 ManufacturingProcess32
              + 0.64 ManufacturingProcess09 - 1.18 BiologicalMaterial08
              + 0.0117 ManufacturingProcess15 - 0.147 BiologicalMaterial02
              - 0.192 ManufacturingProcess33 + 1.11 BiologicalMaterial09
              - 0.000249 ManufacturingProcess12
              - 0.0068 ManufacturingProcess14 + 0.088 BiologicalMaterial03
```

```
            - 0.23 ManufacturingProcess13

Model 44:

  Rule 44/1: [144 cases, mean 40.168, range 35.25 to 44.16, est err 1.282]

    outcome = 63.582 - 0.0127 ManufacturingProcess27
              + 2.151 ManufacturingProcess29 + 0.2 ManufacturingProcess31
              - 0.87 ManufacturingProcess13 - 0.123 ManufacturingProcess28
              + 0.138 BiologicalMaterial03 - 0.082 ManufacturingProcess22

Model 45:

  Rule 45/1: [144 cases, mean 40.168, range 35.25 to 44.16, est err 1.088]

    outcome = -30.565 + 0.29 ManufacturingProcess32
              + 0.53 ManufacturingProcess09

Model 46:

  Rule 46/1: [51 cases, mean 39.029, range 36.77 to 42.73, est err 1.273]

    if
    BiologicalMaterial12 <= 19.93
    ManufacturingProcess13 <= 35.5
    ManufacturingProcess39 > 6.9
    then
    outcome = 122.619 - 12.84 ManufacturingProcess39
              + 0.271 BiologicalMaterial11 - 0.85 ManufacturingProcess13
              + 0.034 ManufacturingProcess29
              - 0.00015 ManufacturingProcess27 - 0.005 BiologicalMaterial02

  Rule 46/2: [10 cases, mean 39.580, range 35.25 to 42.68, est err 1.846]

    if
    ManufacturingProcess13 > 35.5
    then
    outcome = 138.657 - 2.74 ManufacturingProcess13
              - 2.65 ManufacturingProcess37 + 0.24 ManufacturingProcess01

  Rule 46/3: [11 cases, mean 39.853, range 38 to 42.73, est err 3.915]

    if
    BiologicalMaterial12 <= 19.1
    ManufacturingProcess13 <= 35.5
    then
    outcome = 0.718 - 6.43 ManufacturingProcess39
              + 10.27 BiologicalMaterial12 - 3.06 ManufacturingProcess13
              + 0.354 ManufacturingProcess29
              - 0.00153 ManufacturingProcess27 + 0.074 BiologicalMaterial03
              - 0.054 BiologicalMaterial02 - 0.027 BiologicalMaterial06
              - 0.016 BiologicalMaterial11 + 0.09 ManufacturingProcess11
              + 0.033 BiologicalMaterial05 - 0.007 ManufacturingProcess24
```

```
Rule 46/4: [131 cases, mean 40.134, range 35.25 to 44.16, est err 1.133]

  if
  ManufacturingProcess39 > 6.9
  then
  outcome = 85.845 - 2.66 ManufacturingProcess39
            + 1.631 ManufacturingProcess29
            - 0.00707 ManufacturingProcess27 + 0.343 BiologicalMaterial03
            - 0.248 BiologicalMaterial02 - 0.72 ManufacturingProcess13
            - 0.081 ManufacturingProcess02 - 0.123 BiologicalMaterial06
            - 0.073 BiologicalMaterial11 + 0.42 ManufacturingProcess11
            + 0.153 BiologicalMaterial05 - 0.03 ManufacturingProcess24

Rule 46/5: [23 cases, mean 40.724, range 38.37 to 44.16, est err 1.835]

  if
  BiologicalMaterial12 > 19.93
  ManufacturingProcess02 > 20.5
  ManufacturingProcess13 <= 35.5
  ManufacturingProcess39 > 6.9
  then
  outcome = 197.729 - 3.55 BiologicalMaterial12
            - 2.32 ManufacturingProcess13 - 0.381 BiologicalMaterial02
            + 0.203 BiologicalMaterial03 - 0.41 ManufacturingProcess17
            + 11.9 ManufacturingProcess03

Rule 46/6: [12 cases, mean 40.818, range 38.35 to 43.88, est err 1.112]

  if
  ManufacturingProcess13 <= 35.5
  ManufacturingProcess39 <= 6.9
  then
  outcome = 41.725 + 3.751 ManufacturingProcess29
            - 0.01627 ManufacturingProcess27 + 0.789 BiologicalMaterial03
            - 0.57 BiologicalMaterial02 - 0.284 BiologicalMaterial06
            - 0.167 BiologicalMaterial11 + 0.98 ManufacturingProcess11
            + 0.351 BiologicalMaterial05 - 0.07 ManufacturingProcess24
            + 0.02 ManufacturingProcess39

Rule 46/7: [9 cases, mean 41.890, range 39.71 to 43.33, est err 0.740]

  if
  BiologicalMaterial12 > 19.1
  ManufacturingProcess02 > 19.5
  ManufacturingProcess02 <= 20.5
  ManufacturingProcess39 > 6.9
  then
  outcome = 67.006 + 2.27 ManufacturingProcess29
            - 0.00988 ManufacturingProcess27 - 0.345 BiologicalMaterial02
            + 0.309 BiologicalMaterial03 - 0.57 ManufacturingProcess17
            - 0.47 ManufacturingProcess39 - 0.15 ManufacturingProcess13
            + 0.03 BiologicalMaterial06 - 0.07 BiologicalMaterial12
            - 0.011 BiologicalMaterial11 + 0.06 ManufacturingProcess11
            + 0.022 BiologicalMaterial05 - 0.004 ManufacturingProcess24
```

```
             + 0.8 ManufacturingProcess03

Model 47:

  Rule 47/1: [144 cases, mean 40.168, range 35.25 to 44.16, est err 1.056]

    outcome = -134.79 + 0.317 ManufacturingProcess32
              + 0.57 ManufacturingProcess09 + 0.088 ManufacturingProcess04
              + 7.6 ManufacturingProcess34 - 0.26 ManufacturingProcess10

Model 48:

  Rule 48/1: [144 cases, mean 40.168, range 35.25 to 44.16, est err 1.155]

    outcome = 70.625 + 2.179 ManufacturingProcess29
              - 0.00948 ManufacturingProcess27 - 0.91 ManufacturingProcess13
              + 0.043 ManufacturingProcess02

Model 49:

  Rule 49/1: [144 cases, mean 40.168, range 35.25 to 44.16, est err 1.180]

    outcome = -139.577 + 0.363 ManufacturingProcess32
              + 0.53 ManufacturingProcess09 + 0.105 ManufacturingProcess04
              - 0.57 ManufacturingProcess21

Model 50:

  Rule 50/1: [38 cases, mean 39.027, range 36.77 to 42.73, est err 1.232]

    if
    BiologicalMaterial12 <= 19.93
    ManufacturingProcess28 <= 0
    ManufacturingProcess39 > 6.9
    then
    outcome = 6.941 - 18.13 ManufacturingProcess39
              - 0.01452 ManufacturingProcess27
              + 0.01105 ManufacturingProcess26
              + 0.352 ManufacturingProcess06 + 1.19 ManufacturingProcess21
              - 1.04 BiologicalMaterial12 + 0.0168 ManufacturingProcess19
              + 1.1 ManufacturingProcess11

  Rule 50/2: [77 cases, mean 40.287, range 35.25 to 43.62, est err 1.036]

    if
    ManufacturingProcess13 > 32.8
    ManufacturingProcess28 > 0
    then
    outcome = 38.476 + 1.778 ManufacturingProcess29
              - 0.00747 ManufacturingProcess27
              + 0.532 ManufacturingProcess28 + 0.655 BiologicalMaterial04
              - 0.451 ManufacturingProcess33 - 1.65 BiologicalMaterial10
              - 0.259 BiologicalMaterial02 + 0.147 ManufacturingProcess32
              + 0.198 BiologicalMaterial03 + 0.73 ManufacturingProcess43
```

```
            + 0.56 ManufacturingProcess11 - 0.026 ManufacturingProcess35
            + 0.36 ManufacturingProcess21 - 0.58 ManufacturingProcess37
            + 0.15 ManufacturingProcess09 - 0.06 ManufacturingProcess13

  Rule 50/3: [8 cases, mean 40.454, range 38.48 to 43.88, est err 4.760]

    if
    ManufacturingProcess28 <= 0
    ManufacturingProcess39 <= 6.9
    then
    outcome = 61.505 - 0.01275 ManufacturingProcess27
            + 0.54 ManufacturingProcess29 + 0.66 ManufacturingProcess39
            + 0.84 ManufacturingProcess13 + 0.223 BiologicalMaterial04
            - 0.141 ManufacturingProcess33 - 0.56 BiologicalMaterial10
            - 0.081 BiologicalMaterial02 + 0.046 ManufacturingProcess32
            + 0.062 BiologicalMaterial03 + 0.25 ManufacturingProcess43
            - 0.036 ManufacturingProcess28 + 0.19 ManufacturingProcess11
            - 0.009 ManufacturingProcess35 + 0.12 ManufacturingProcess21
            - 0.2 ManufacturingProcess37

  Rule 50/4: [12 cases, mean 40.857, range 39.42 to 43.42, est err 1.341]

    if
    BiologicalMaterial12 > 19.93
    ManufacturingProcess28 <= 0
    ManufacturingProcess39 > 6.9
    then
    outcome = 219.681 - 5.78 BiologicalMaterial12
            - 1.32 ManufacturingProcess39 - 1.63 ManufacturingProcess13
            - 0.00259 ManufacturingProcess27
            + 0.00195 ManufacturingProcess26 - 0.213 BiologicalMaterial03
            + 0.2 ManufacturingProcess11 + 0.003 ManufacturingProcess19

  Rule 50/5: [9 cases, mean 42.791, range 41.43 to 44.16, est err 2.137]

    if
    ManufacturingProcess13 <= 32.8
    then
    outcome = 51.131 - 0.24 ManufacturingProcess24
            + 0.542 ManufacturingProcess29
            - 0.00235 ManufacturingProcess27
            - 0.102 ManufacturingProcess33 - 0.057 BiologicalMaterial02
            + 0.036 ManufacturingProcess32 + 0.049 BiologicalMaterial03
            - 0.18 ManufacturingProcess13


Evaluation on training data (144 cases):

    Average  |error|              0.553
    Relative |error|               0.38
    Correlation coefficient        0.89


    Attribute usage:
```

```
	Conds  Model

	 25%    71%    ManufacturingProcess32
	 13%    27%    ManufacturingProcess17
	 12%    16%    ManufacturingProcess39
	 12%    15%    BiologicalMaterial12
	  9%    48%    ManufacturingProcess13
	  4%    18%    BiologicalMaterial06
	  4%     4%    ManufacturingProcess01
	  4%     8%    BiologicalMaterial05
	  2%    12%    ManufacturingProcess21
	  2%     7%    BiologicalMaterial04
	  2%    10%    ManufacturingProcess28
	  2%     4%    ManufacturingProcess43
	  1%    44%    ManufacturingProcess09
	         4%    BiologicalMaterial10
	         3%    ManufacturingProcess22
	         8%    ManufacturingProcess02
	        30%    BiologicalMaterial03
	        11%    BiologicalMaterial11
	        24%    ManufacturingProcess27
	         8%    BiologicalMaterial09
	         3%    ManufacturingProcess19
	         6%    ManufacturingProcess10
	        34%    ManufacturingProcess29
	        33%    ManufacturingProcess04
	        28%    ManufacturingProcess33
	        26%    BiologicalMaterial02
	        15%    ManufacturingProcess26
	        13%    ManufacturingProcess15
	        10%    ManufacturingProcess14
	         9%    ManufacturingProcess25
	         7%    ManufacturingProcess34
	         6%    ManufacturingProcess24
	         6%    ManufacturingProcess11
	         5%    ManufacturingProcess30
	         3%    ManufacturingProcess37
	         3%    ManufacturingProcess35
	         3%    BiologicalMaterial01
	         3%    ManufacturingProcess16
	         2%    ManufacturingProcess44
	         2%    ManufacturingProcess07
	         2%    ManufacturingProcess31
	         2%    ManufacturingProcess42
	         2%    ManufacturingProcess12
	         1%    ManufacturingProcess45


Time: 0.3 secs
```

R Code

```
# Libraries
library(mlbench)
library(AppliedPredictiveModeling)
library(party)
library(randomForest)
library(caret)
library(tidyverse)
library(gbm)
library(mice)
library(recipes)
library(Cubist)

# Set Seed
set.seed(200)

# (8.1)
simulated <- mlbench.friedman1(200, sd = 1)
simulated <- cbind(simulated$x, simulated$y)
simulated <- as.data.frame(simulated)
colnames(simulated)[ncol(simulated)] <- "y"

# (8.1a)
model1 <- randomForest(y ~ ., data = simulated, importance = TRUE,
    ntree = 1000)
rfImp1 <- varImp(model1, scale = FALSE)

# (8.1b)
simulated$duplicate1 <- simulated$V1 + rnorm(200) * 0.1
model2 <- randomForest(y ~ ., data = simulated, importance = T,
    ntree = 1000)
rfImp2 <- varImp(model2, scale = F)

# (8.1c) Rename variables for ease of reference
sim_original <- select(simulated, -duplicate1)
sim_duplicate <- simulated

model3 <- cforest(y ~ ., data = sim_original, controls = cforest_unbiased(ntree = 1000))
cfImp3 <- varimp(model3, conditional = T)
cfImp3df <- as.data.frame(cfImp3) %>% rownames_to_column("Variable") %>%
    rename(Overall = "cfImp3")
model4 <- cforest(y ~ ., data = sim_duplicate, controls = cforest_unbiased(ntree = 1000))
cfImp4 <- varimp(model4, conditional = T)
cfImp4df <- as.data.frame(cfImp4) %>% rownames_to_column("Variable") %>%
    rename(Overall = "cfImp4")
cfImpTbl <- right_join(cfImp3df, cfImp4df, by = "Variable") %>%
    column_to_rownames("Variable") %>% rename(Original = Overall.x,
    Duplicate = Overall.y)

# (8.1d)
model5 <- gbm(y ~ ., data = sim_original, distribution = "gaussian",
    n.trees = 1000, cv.folds = 5)
```

```
gbmImp5 <- summary(model5, plot = F) %>% mutate(var = as.character(var))
gbmImp5df <- as.data.frame(gbmImp5) %>% arrange(desc(rel.inf))
model6 <- gbm(y ~ ., data = sim_duplicate, distribution = "gaussian",
    n.trees = 1000, cv.folds = 5)
gbmImp6 <- summary(model6, plot = F) %>% mutate(var = as.character(var))
gbmImp6df <- as.data.frame(gbmImp6) %>% arrange(desc(rel.inf))
gbmImpTbl <- right_join(gbmImp5df, gbmImp6df, by = "var") %>%
    column_to_rownames("var") %>% rename(Original = rel.inf.x,
    Duplicate = rel.inf.y)

# (8.2)
random_predictor <- data.frame(V1 = sample(1:2, 100, replace = TRUE),
    V2 = sample(1:100, 100, replace = TRUE), V3 = sample(1:1000,
        100, replace = TRUE), V4 = sample(1:5000, 100, replace = TRUE))
sim_df <- random_predictor %>% mutate(y = V1 * V2 * V3 + rnorm(100))
sim_rf <- randomForest(y ~ ., data = sim_df, importance = TRUE,
    ntree = 1000)
sim_varImp <- varImp(sim_rf, scale = T)

# (8.7)

# FOR FINAL HW SUBMISSION, DO NOT REPEAT STEPS: JUST CALL
# VARIABLES FROM PRIOR ASSIGNMENT.

data(ChemicalManufacturingProcess)
CMP_Impute <- mice(ChemicalManufacturingProcess, m = 5, printFlag = F)
CMP_DF <- mice::complete(CMP_Impute, 2)

# Create Partition for Train/Test Splits
trainingRows <- createDataPartition(CMP_DF$Yield, p = 0.8, list = FALSE)

# Split Train/Test Data
train <- CMP_DF[trainingRows, ]
test <- CMP_DF[-trainingRows, ]

# Pre-Process Recipe
rec <- recipes::recipe(CMP_DF, Yield ~ .)
rec <- rec %>% step_nzv(all_predictors(), options = list(freq_cut = 95/5,
    unique_cut = 10))
prep_rec = prep(rec, training = CMP_DF)
CMP_DF_TF = bake(prep_rec, CMP_DF)

# Create Partition for Train/Test Splits
trainingRows <- createDataPartition(CMP_DF_TF$Yield, p = 0.8,
    list = FALSE)

# Split Train/Test Data
train <- CMP_DF_TF[trainingRows, ]
test <- CMP_DF_TF[-trainingRows, ]

# (8.7a) Controls
tr_control <- trainControl(method = "cv", number = 10, repeats = 10)
# Random Foreset
```

```r
rfModel <- train(Yield ~ ., data = train, method = "rf", importance = T,
    trControl = tr_control, tuneLength = 5)
rfTrainMetric <- data.frame(RMSE = rfModel$results$RMSE, RSquared = rfModel$results$Rsquared,
    MAE = rfModel$results$MAE) %>% filter(RMSE == min(RMSE))
rfPred <- predict(rfModel, newdata = test)
rfPerf <- postResample(pred = rfPred, obs = test$Yield)
rfPlot <- ggplot(rfModel) + labs(title = "Random Forest Model Cross-Validated RMSE Profile") +
    theme(legend.position = "bottom")

# Gradient Boosted
gbmGrid <- expand.grid(n.trees = c(50, 100, 1000), interaction.depth = c(2,
    7, 10), shrinkage = c(0.1, 0.5), n.minobsinnode = c(5, 10))
gbmModel <- train(Yield ~ ., data = train, method = "gbm", trControl = tr_control,
    tuneGrid = gbmGrid, tuneLength = 5)
gbmTrainMetric <- data.frame(RMSE = gbmModel$results$RMSE, RSquared = gbmModel$results$Rsquared,
    MAE = gbmModel$results$MAE) %>% filter(RMSE == min(RMSE))
gbmPred <- predict(gbmModel, newdata = test)
gbmPerf <- postResample(pred = gbmPred, obs = test$Yield)
gbmPlot <- ggplot(gbmModel) + labs(title = "GBM Cross-Validated RMSE") +
    theme(legend.position = "bottom")

# Cubist
cbGrid <- expand.grid(committees = c(10, 50, 100), neighbors = c(1,
    5, 9))
cbModel <- train(Yield ~ ., data = train, method = "cubist",
    trControl = tr_control, tuneGrid = cbGrid, tuneLength = 5)
cbTrainMetric <- data.frame(RMSE = cbModel$results$RMSE, RSquared = cbModel$results$Rsquared,
    MAE = cbModel$results$MAE) %>% filter(RMSE == min(RMSE))
cbPred <- predict(gbmModel, newdata = test)
cbPerf <- postResample(pred = gbmPred, obs = test$Yield)
cbPlot <- ggplot(cbModel) + labs(title = "Cubist Cross-Validated RMSE") +
    theme(legend.position = "bottom")

# Metrics Table
performance <- rbind(rfTrain = rfTrainMetric, rfTest = rfPerf,
    gbmTrain = gbmTrainMetric, gbmTest = gbmPerf, cbTrain = cbTrainMetric,
    cbTest = cbPerf)

# (8.7b)
cbImp <- varImp(cbModel, scale = F)
cbImpPlot <- cbImp$importance %>% as.data.frame() %>% rownames_to_column("Variable") %>%
    top_n(10, Overall) %>% ggplot(aes(x = reorder(Variable, Overall),
    y = Overall)) + geom_point() + geom_segment(aes(x = Variable,
    xend = Variable, y = 0, yend = Overall)) + coord_flip() +
    labs(y = "Overall", x = "", title = "Cubist Model Overall Variable Importance",
        subtitle = "Top 10 Predictor Variables")

# (8.7c)
cbDotPlot <- dotplot(cbModel$finalModel)
summary(cbModel)
```