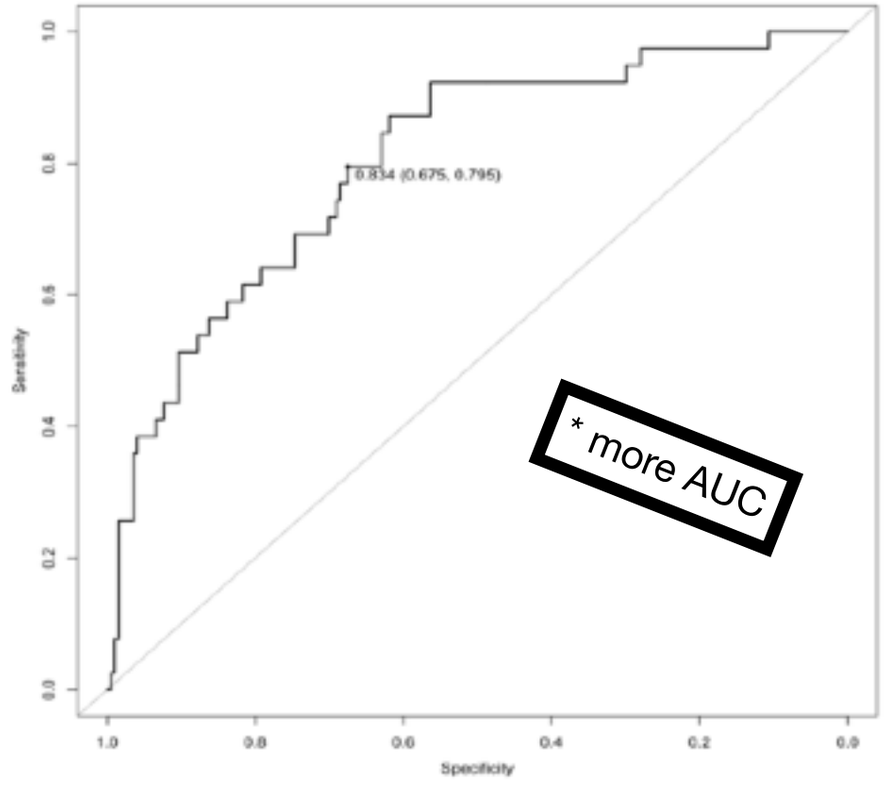
```
Deviance Residuals:
    Min       1Q   Median       3Q      Max
-2.0260  -0.5795  -0.4078  -0.2680   2.8794

Coefficients:
             Estimate Std. Error z value Pr(>|z|)
(Intercept) -6.265e+00  6.454e-01  -9.707  < 2e-16 ***
CMU.Hrs      5.280e-05  1.469e-05   3.593 0.000327 ***
Fe           1.944e-01  4.025e-02   4.830 1.36e-06 ***
Si           4.491e-01  1.249e-01   3.595 0.000324 ***
Na           1.428e-01  3.996e-02   3.574 0.000351 ***
ST           2.764e-02  1.266e-02   2.184 0.028994 *
PFc          1.541e+00  5.713e-01   2.697 0.007007 **
`Fe:Si`     -2.358e-02  8.633e-03  -2.731 0.006307 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 854.26  on 947  degrees of freedom
Residual deviance: 703.34  on 940  degrees of freedom
AIC: 719.34

Number of Fisher Scoring iterations: 6
```

* I know significance isn't everything but hey, it's at least a little reassuring



* more AUC

… model's final form

```
Confusion Matrix and Statistics

          Reference
Prediction   A Not A
     A     193    29
     Not A   4    10

               Accuracy : 0.8602
                 95% CI : (0.8093, 0.9018)
    No Information Rate : 0.8347
    P-Value [Acc > NIR] : 0.1678

                  Kappa : 0.3178

 Mcnemar's Test P-Value : 2.943e-05

            Sensitivity : 0.9797
            Specificity : 0.2564
         Pos Pred Value : 0.8694
         Neg Pred Value : 0.7143
             Prevalence : 0.8347
         Detection Rate : 0.8178
   Detection Prevalence : 0.9407
      Balanced Accuracy : 0.6181
```

unhealthy >= 50%

```
Confusion Matrix and Statistics

          Reference
Prediction   A Not A
     A     175    19
     Not A  22    20

               Accuracy : 0.8263
                 95% CI : (0.7718, 0.8724)
    No Information Rate : 0.8347
    P-Value [Acc > NIR] : 0.6750

                  Kappa : 0.3891

 Mcnemar's Test P-Value : 0.7548

            Sensitivity : 0.8883
            Specificity : 0.5128
         Pos Pred Value : 0.9021
         Neg Pred Value : 0.4762
             Prevalence : 0.8347
         Detection Rate : 0.7415
   Detection Prevalence : 0.8220
      Balanced Accuracy : 0.7006
```

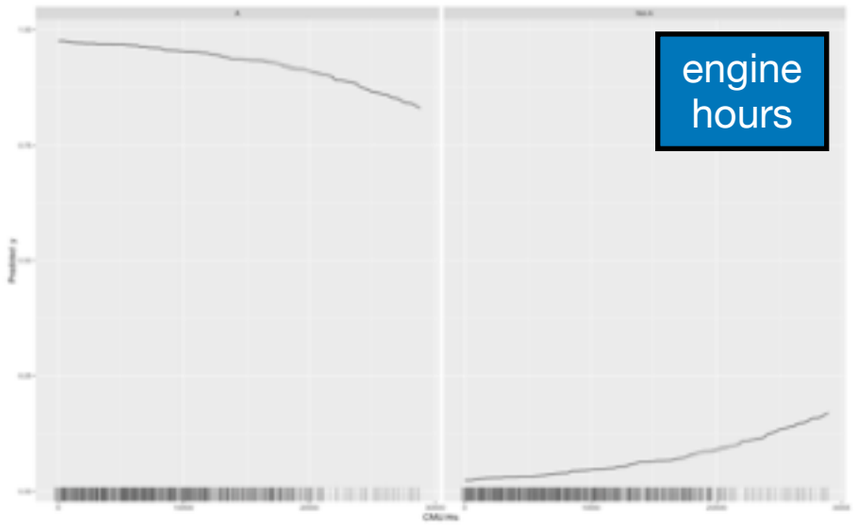
unhealthy >= 30%

```
Confusion Matrix and Statistics

          Reference
Prediction   A Not A
     A      98     3
     Not A  99    36

               Accuracy : 0.5678
                 95% CI : (0.502, 0.6319)
    No Information Rate : 0.8347
    P-Value [Acc > NIR] : 1

                  Kappa : 0.2116

 Mcnemar's Test P-Value : <2e-16

            Sensitivity : 0.4975
            Specificity : 0.9231
         Pos Pred Value : 0.9703
         Neg Pred Value : 0.2667
             Prevalence : 0.8347
         Detection Rate : 0.4153
   Detection Prevalence : 0.4280
      Balanced Accuracy : 0.7103
```

unhealthy >= 15%

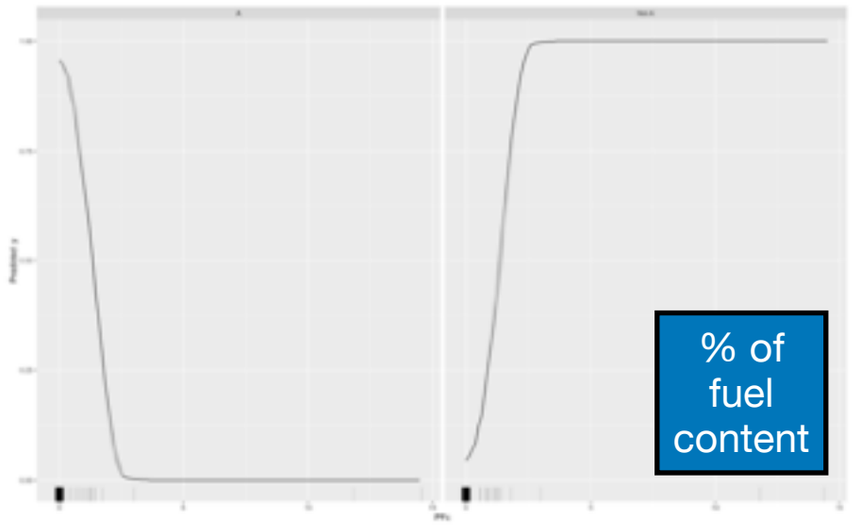
* better performance in general, especially when adjusting the probability cut-off

… simpler interpretation

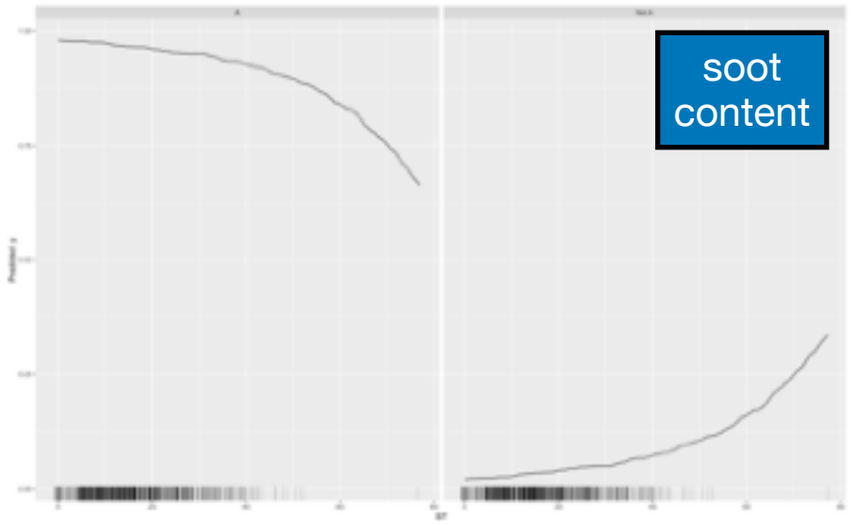


engine hours

Probability of a sample being **“healthy”** is on the **left** of each plot pair!

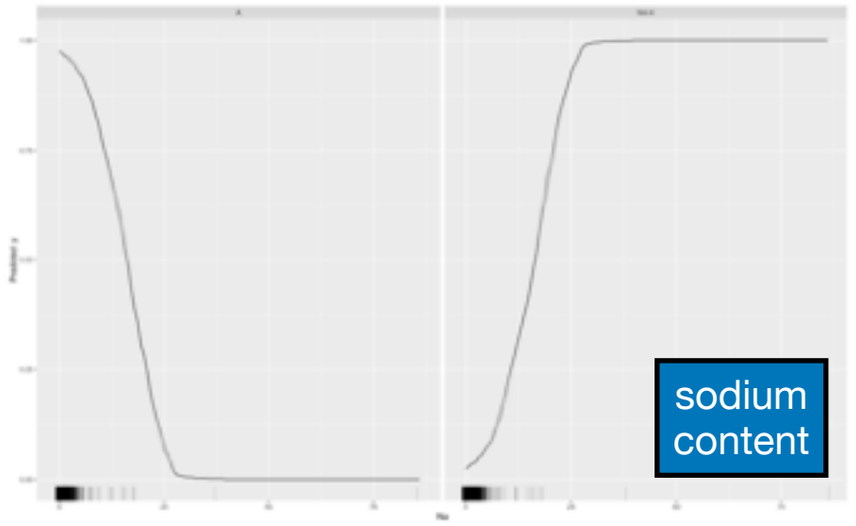


% of fuel content

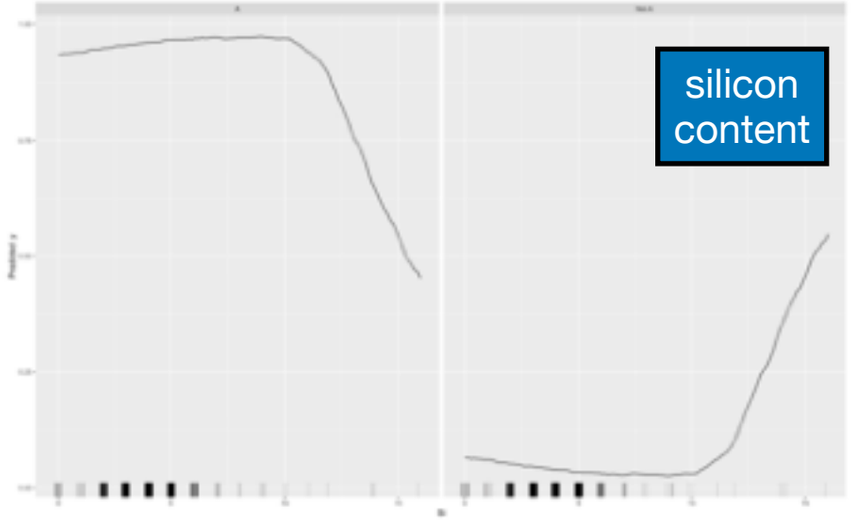


soot content

Probability of a sample being **“unhealthy”** is on the **right** of each plot pair!

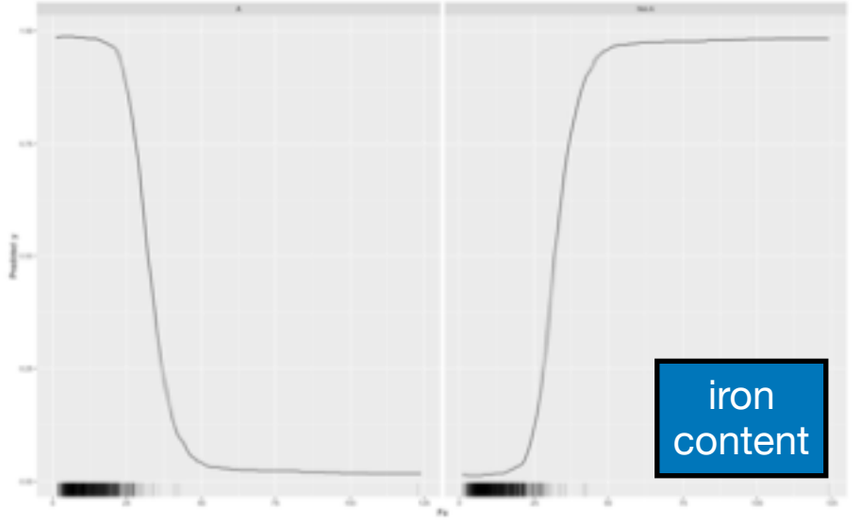


sodium content



silicon content

As you move from **left to right** the variable in question is increasing along the X-axis!



iron content