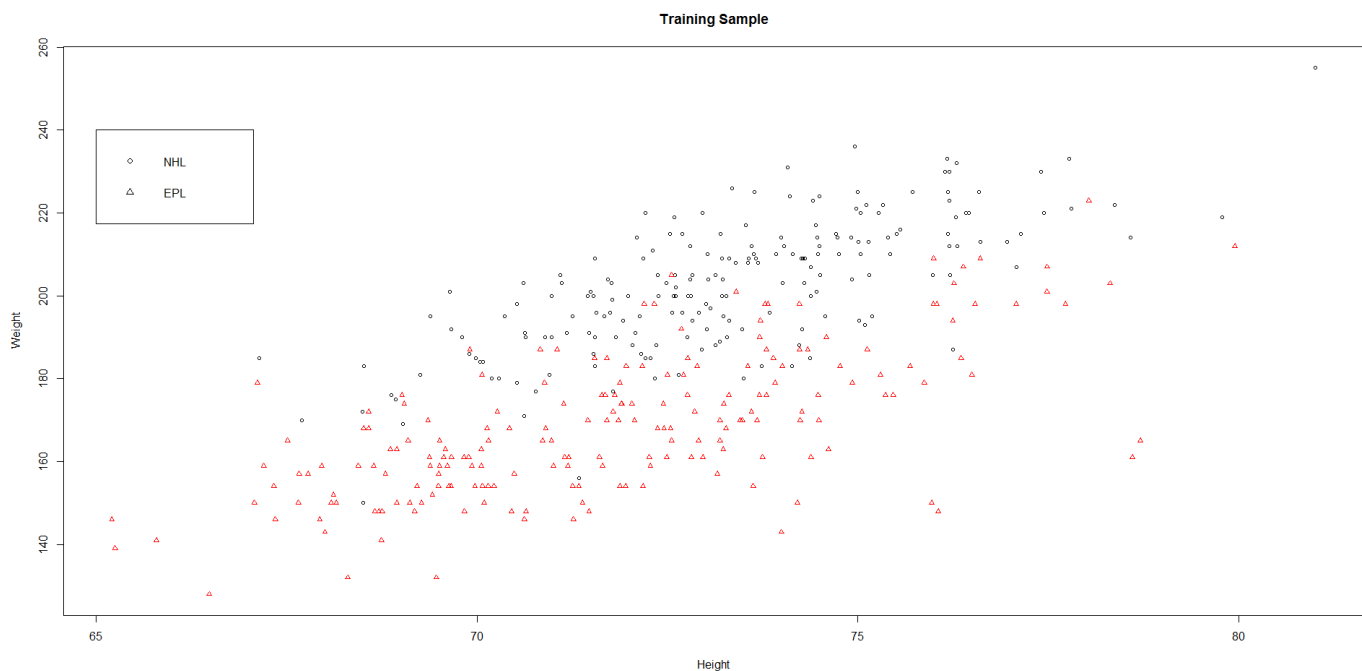


NHL/EPL Heights/Weights – Fisher’s Discriminant Function

Data: Groups: National Hockey League ($N_{NHL}=717$) and English Premier League ($N_{EPL} = 526$)

Variables: Height + $U(-0.5,0.5)$ (inches) and Weight (pounds)



Procedure: Take Random Samples of $n_{NHL} = n_{EPL} = 200$ from each population and obtain Fisher’s Discriminant Function, then apply the function and decision rule to the holdout samples ($517 + 326 = 843$ players).

$$\hat{\mathbf{a}} = \mathbf{S}_{\text{Pooled}}^{-1} (\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2) \quad \hat{m} = \frac{1}{2} (\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2)' \mathbf{S}_{\text{Pooled}}^{-1} (\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2) \quad \hat{y}_0 = \hat{\mathbf{a}}' \mathbf{x}_0 \quad \text{Classify as league 1 if } \hat{y}_0 \geq \hat{m}, \text{ 2 ow}$$

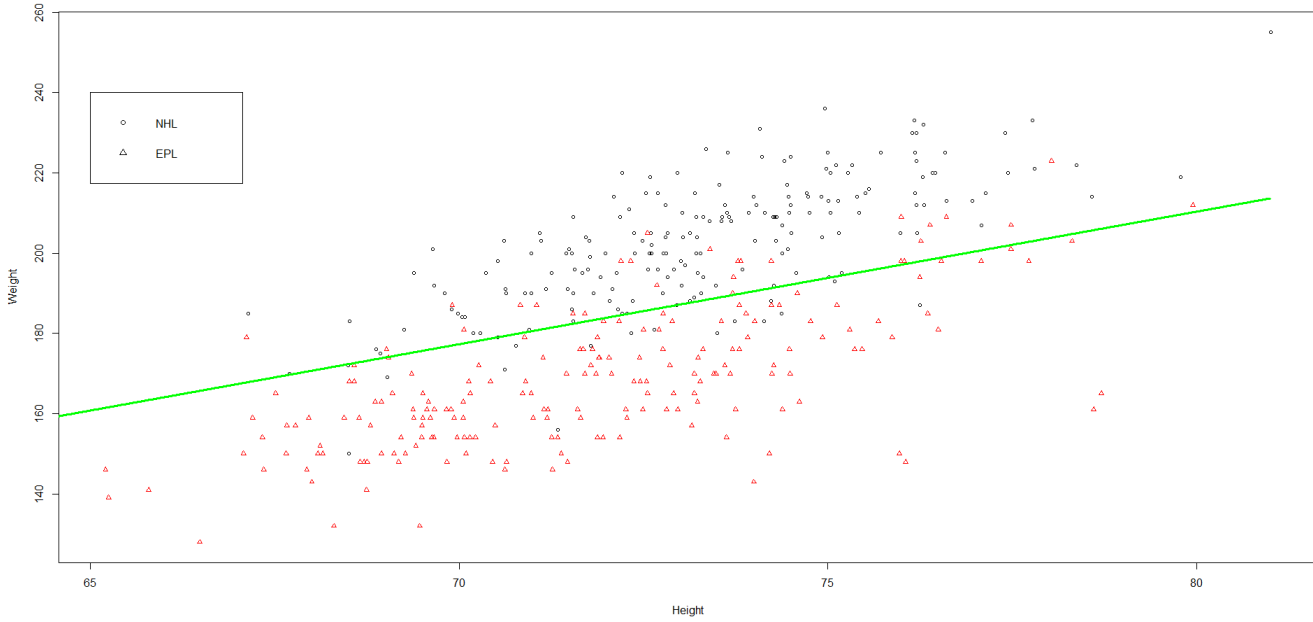
Xbar1	Xbar2	Diff	Spooled	
73.294	71.811	1.483	6.677	29.874
202.645	168.795	33.850	29.874	276.956

```
> (ahat <- solve(Sp) %*% (xbar1-xbar2))
-0.6274986
0.1899074
> (mhat <- 0.5 * (t(xbar1-xbar2) %*% solve(Sp) %*% (xbar1+xbar2)))
-10.25683
```

"Line of Discrimination": Height on Horizontal, Weight on Vertical Axis

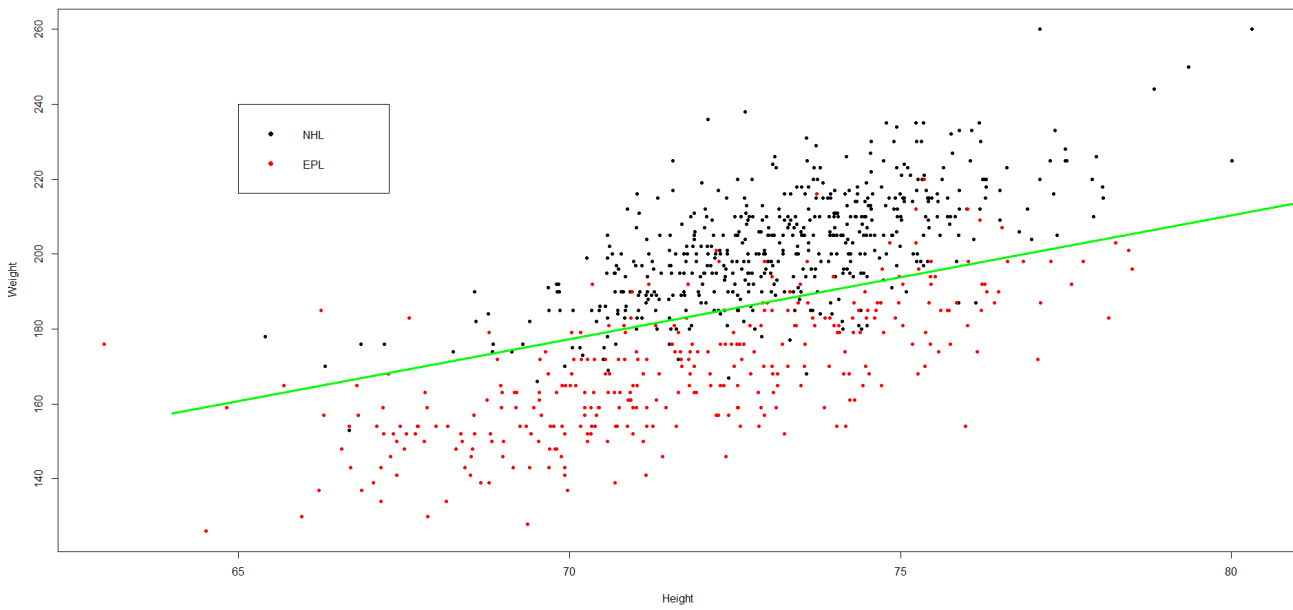
$$-0.6275H + 0.1899W = -10.2568 \Rightarrow W = \frac{-10.2568 - (-0.6275H)}{0.1899} = -54.0096 + 3.3042H$$

Training Sample with Discrimination Line



Classification of Holdout Sample

Holdout Sample with Discrimination Line



Confusion Matrix and Classification Probabilities

```
> (classtab <- table(league, classify))
      classify
league  1   2
      1 466  51
      2  36 290
>
>
> prop.table(classtab)
      classify
league      1      2
      1 0.55278766 0.06049822
      2 0.04270463 0.34400949
> prop.table(classtab,1)
      classify
league      1      2
      1 0.90135397 0.09864603
      2 0.11042945 0.88957055
> prop.table(classtab,2)
      classify
league      1      2
      1 0.92828685 0.14956012
      2 0.07171315 0.85043988
>
> (E.AER <- (classtab[1,2]+classtab[2,1])/sum(classtab))
[1] 0.1032028
```