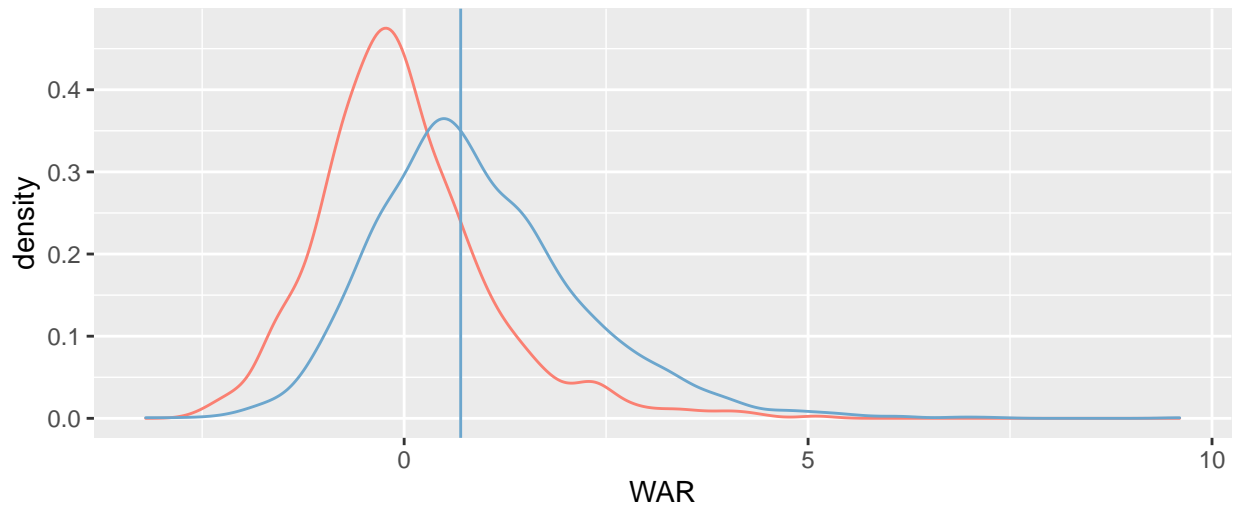


Modeling

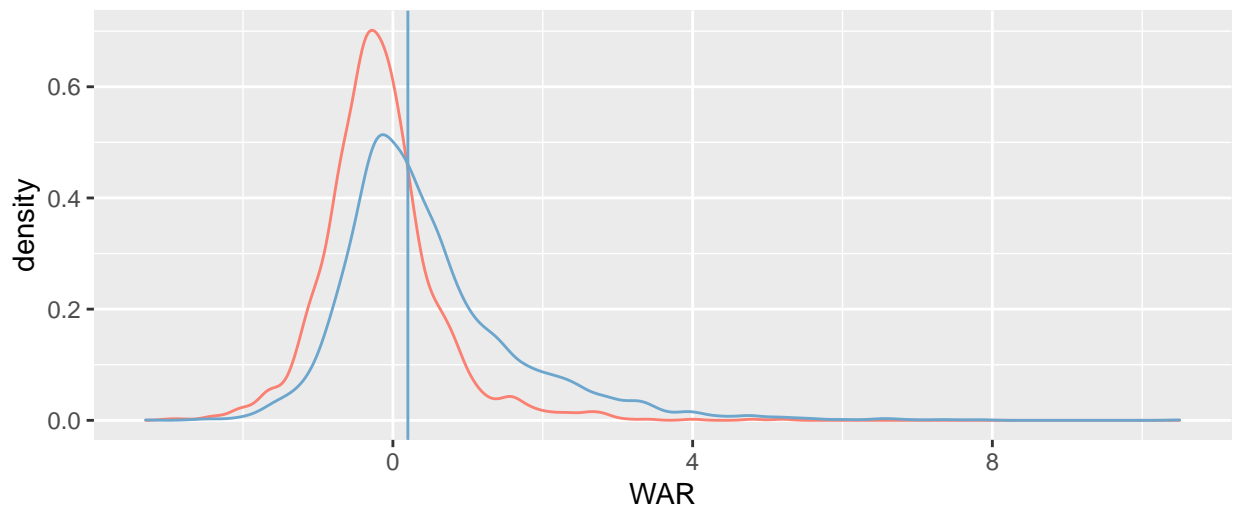
Group 6

The Couldabeen Classification Problem

Pitchers



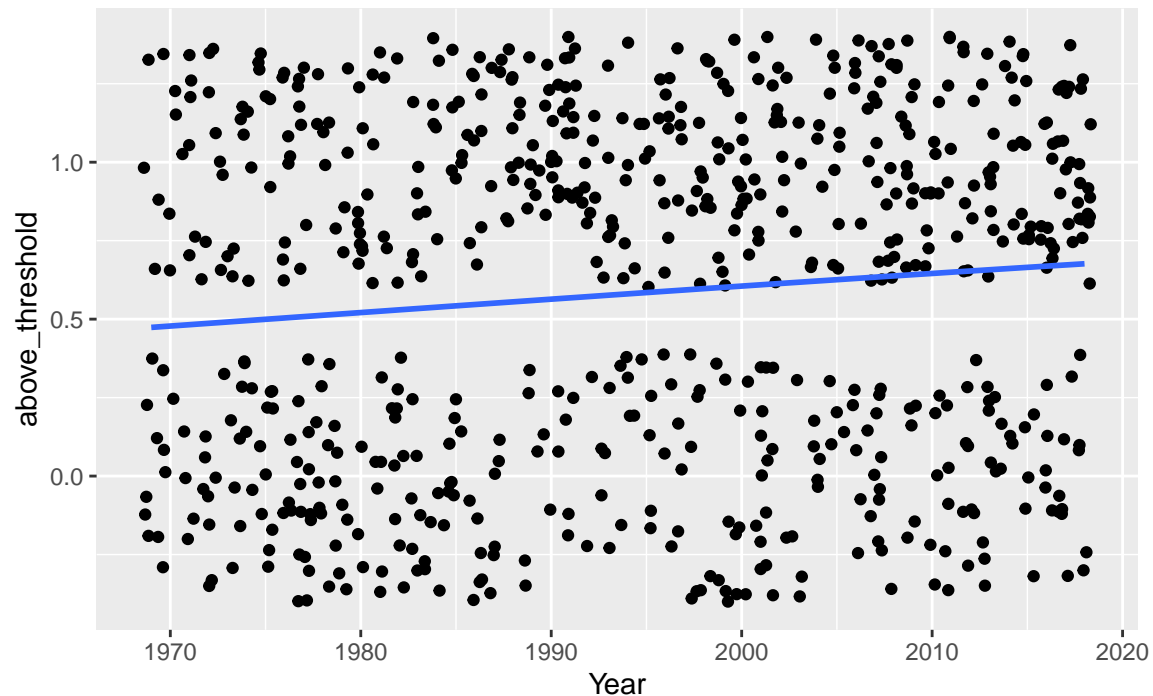
Position



```
#####  
#           Counting: Couldabeens           #  
#####  
# Combine the threshold-classified retiree datasets  
retirees <- rbind(pit_ret,pos_ret)  
# Count couldabeens  
couldabeens <- count_cbns(retirees)
```

```
## # A tibble: 6 x 2  
##   Year  cbns  
##   <dbl> <int>  
## 1  1969    25  
## 2  1970    27  
## 3  1971    33  
## 4  1972    34  
## 5  1973    32  
## 6  1974    32
```

First Look: A Logistic Model



```
##  
## Call:  
## glm(formula = above_threshold ~ Year, family = "binomial", data = dataset)  
##  
## Deviance Residuals:  
##      Min       1Q   Median       3Q      Max   
## -1.5021  -1.2503   0.9162   1.0500   1.2224   
##  
## Coefficients:  
##              Estimate Std. Error z value Pr(>|z|)      
## (Intercept) -33.95461   10.62870  -3.195  0.00140 **  
## Year         0.01719    0.00533   3.225  0.00126 **
```

```
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##      Null deviance: 948.93  on 697  degrees of freedom
## Residual deviance: 938.39  on 696  degrees of freedom
## AIC: 942.39
##
## Number of Fisher Scoring iterations: 4

##   Year cbns retirees      prop
## 1 1969   25         32 0.7812500
## 2 1970   27         35 0.7714286
## 3 1971   33         44 0.7500000
## 4 1972   34         49 0.6938776
## 5 1973   32         41 0.7804878
## 6 1974   32         45 0.7111111

#=====#
#      Modeling      #
#=====#
# Partition dataset into years before and after rule
couldabeens_pre <- prerule(couldabeens)
couldabeens_post <- postrule(couldabeens)
# Obtain linear model for pre-rule years
model_pre <- linear_model(couldabeens_pre)
coefs_pre <- model_pre$coefficients
# Obtain linear model for post-rule years
model_post <- linear_model(couldabeens_post)
coefs_post <- model_post$coefficients
```

Couldabeens: Pre-rule Era (1969-2002)

```
##  
## Call:  
## lm(formula = prop ~ I(Year), data = dataset)  
##  
## Residuals:  
##      Min       1Q   Median       3Q      Max   
## -0.16527 -0.04650  0.00206  0.05467  0.14437   
##  
## Coefficients:  
##              Estimate Std. Error t value Pr(>|t|)      
## (Intercept) -1.370105   2.515373  -0.545   0.590      
## I(Year)      0.001062   0.001267   0.838   0.408      
##  
## Residual standard error: 0.07247 on 32 degrees of freedom  
## Multiple R-squared:  0.02147,    Adjusted R-squared:  -0.009108   
## F-statistic: 0.7022 on 1 and 32 DF,  p-value: 0.4083
```

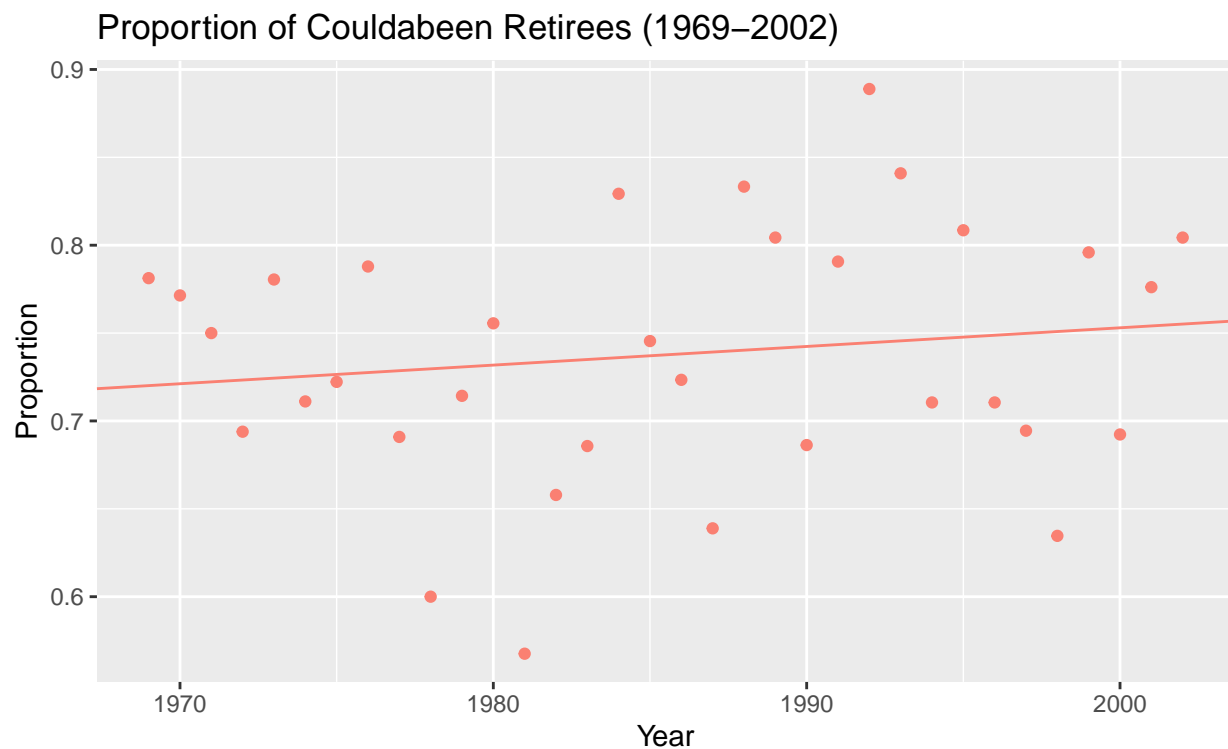


Figure 1: Proportion of Retirees who were Couldabeens prior to the implementation of the Luxury Tax

Couldabeens: Post-rule Era (2003-2018)

```
##  
## Call:  
## lm(formula = prop ~ I(Year), data = dataset)  
##  
## Residuals:  
##      Min       1Q   Median       3Q      Max   
## -0.119317 -0.059649  0.007258  0.052156  0.191109   
##  
## Coefficients:  
##              Estimate Std. Error t value Pr(>|t|)      
## (Intercept) -1.783306  10.019728  -0.178    0.861      
## I(Year)      0.001235   0.004984   0.248    0.808      
##  
## Residual standard error: 0.09189 on 14 degrees of freedom  
## Multiple R-squared:  0.004364,    Adjusted R-squared:  -0.06675   
## F-statistic: 0.06136 on 1 and 14 DF,  p-value: 0.8079
```

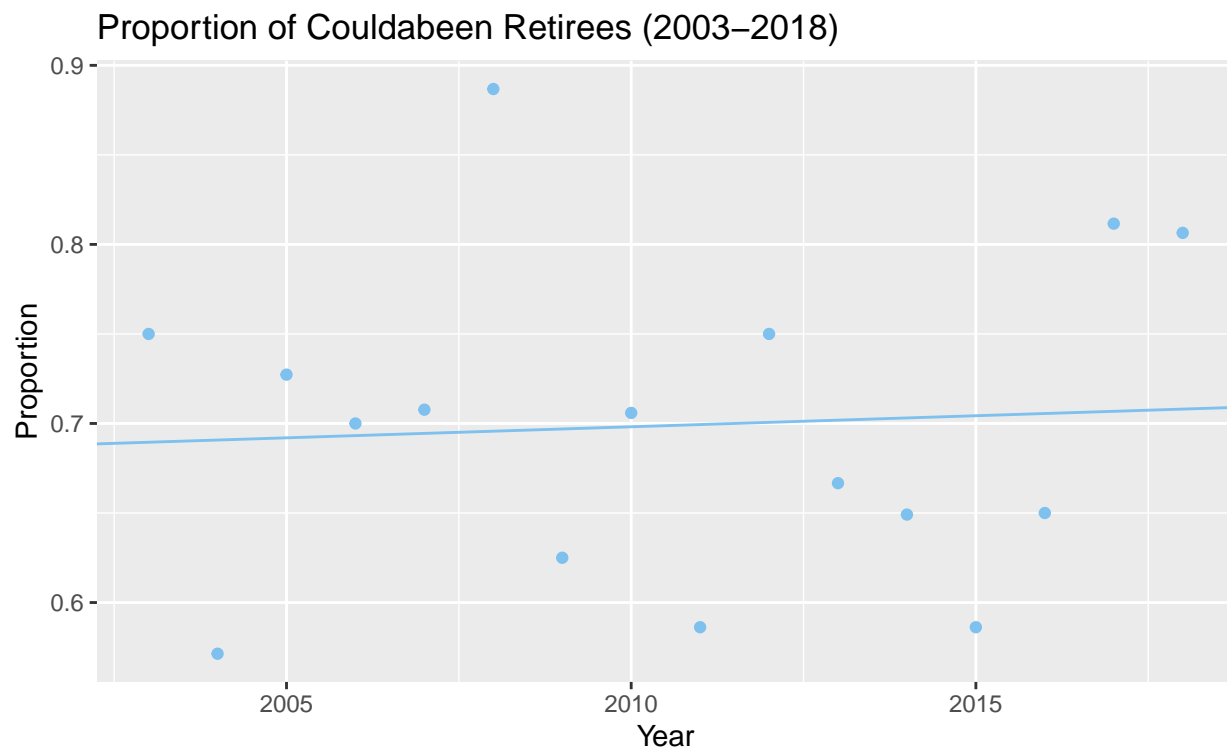


Figure 2: Proportion of Retirees who were Couldabeens after the implementation of the Luxury Tax