

Model 404

Load and Prepare Data

```
Students <- read.table("Students.dat", header = TRUE)
attach(Students)
head(Students)
```

```
##  subject gender age hsgpa cogpa dhome dres tv sport news aids veg affil ideol
## 1      1      0  32   2.2   3.5    0  5.0  3    5    0    0    0    2    6
## 2      2      1  23   2.1   3.5  1200  0.3 15    7    5    6    1    1    2
## 3      3      1  27   3.3   3.0  1300  1.5  0    4    3    0    1    1    2
## 4      4      1  35   3.5   3.2  1500  8.0  5    5    6    3    0    3    4
## 5      5      0  23   3.1   3.5  1600 10.0  6    6    3    0    0    3    1
## 6      6      0  39   3.5   3.5   350  3.0  4    5    7    0    1    1    2
##  relig  abor  affirm  life
## 1      2      0      0    1
## 2      1      1      1    3
## 3      2      1      1    3
## 4      1      1      1    2
## 5      0      1      0    2
## 6      1      1      1    3
```

```
library(tidyverse)
```

```
## Warning: package 'ggplot2' was built under R version 4.5.2
```

```
## -- Attaching core tidyverse packages ----- tidyverse 2.0.0 --
## v dplyr      1.1.4      v readr      2.1.5
## v forcats    1.0.1      v stringr   1.6.0
## v ggplot2    4.0.1      v tibble    3.3.0
## v lubridate  1.9.4      v tidyr     1.3.1
## v purrr      1.2.0
## -- Conflicts ----- tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
## i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become errors
```

```
students <- read.table("Students.dat", header = TRUE)
students$abor <- as.factor(students$abor)

# Select numeric variables except abor
numeric_vars <- names(students)[sapply(students, is.numeric)]
numeric_vars <- setdiff(numeric_vars, "abor")
```

```

# Standardize all numeric variables
students_scaled <- students
students_scaled[numeric_vars] <- scale(students[numeric_vars])

# Convert to long format
long_data <- students_scaled %>%
  pivot_longer(cols = numeric_vars,
               names_to = "variable",
               values_to = "value")

```

```

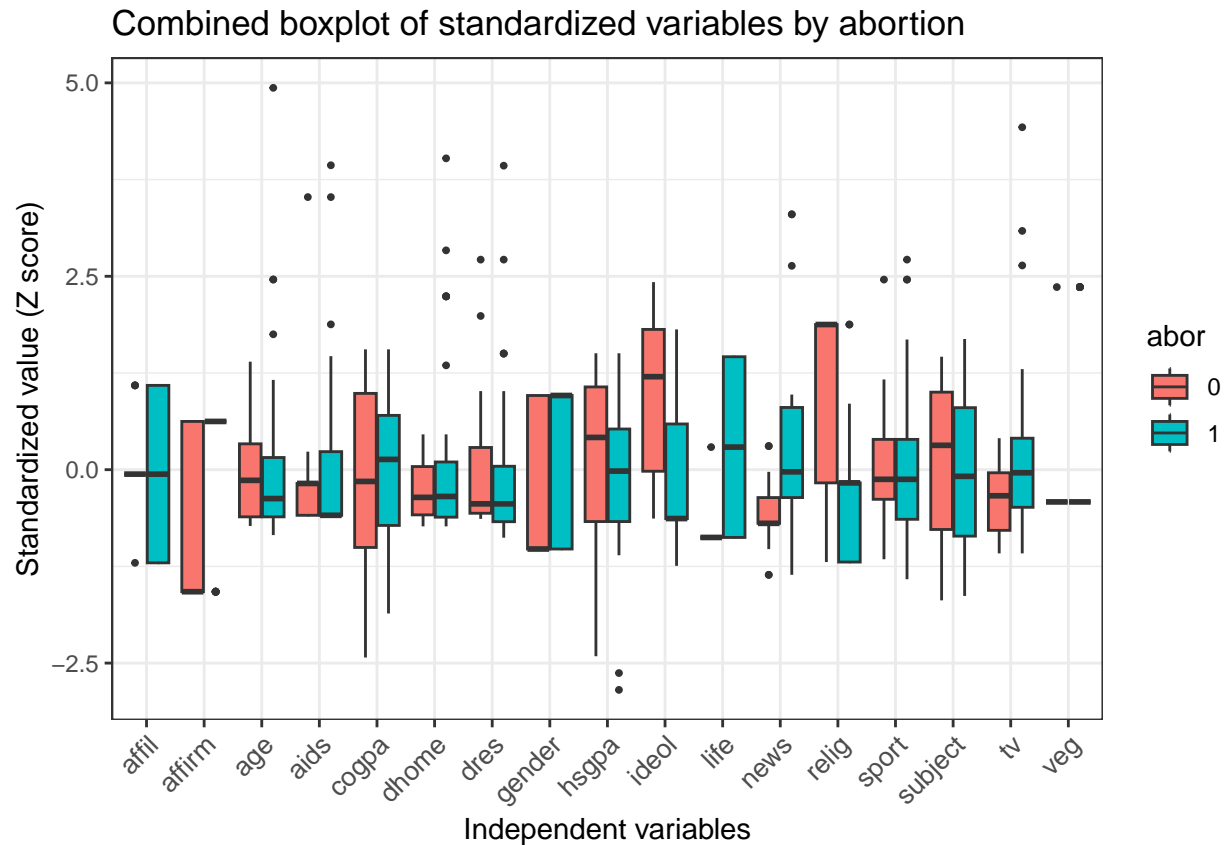
## Warning: Using an external vector in selections was deprecated in tidysselect 1.1.0.
## i Please use 'all_of()' or 'any_of()' instead.
##   # Was:
##   data %>% select(numeric_vars)
##
##   # Now:
##   data %>% select(all_of(numeric_vars))
##
## See <https://tidysselect.r-lib.org/reference/faq-external-vector.html>.
## This warning is displayed once every 8 hours.
## Call 'lifecycle::last_lifecycle_warnings()' to see where this warning was
## generated.

```

```

# Plot
ggplot(long_data, aes(x = variable, y = value, fill = abort)) +
  geom_boxplot(position = position_dodge(width = 0.8), outlier.size = 0.7) +
  labs(
    x = "Independent variables",
    y = "Standardized value (Z score)",
    title = "Combined boxplot of standardized variables by abortion"
  ) +
  theme_bw() +
  theme(axis.text.x = element_text(angle = 45, hjust = 1, size = 10))

```



```
# Create affil2: Republican dummy variable
affil2 <- ifelse(affil == 2, 1, 0) # 1 = Republican, 0 = otherwise

# Create religdummy: Religious dummy variable
religdummy <- ifelse(relig == 3, 1, 0) # 1 = relig is 3, 0 = otherwise

# Check the distributions
table(affil2)
```

```
## affil2
## 0 1
## 45 15
```

```
table(religdummy)
```

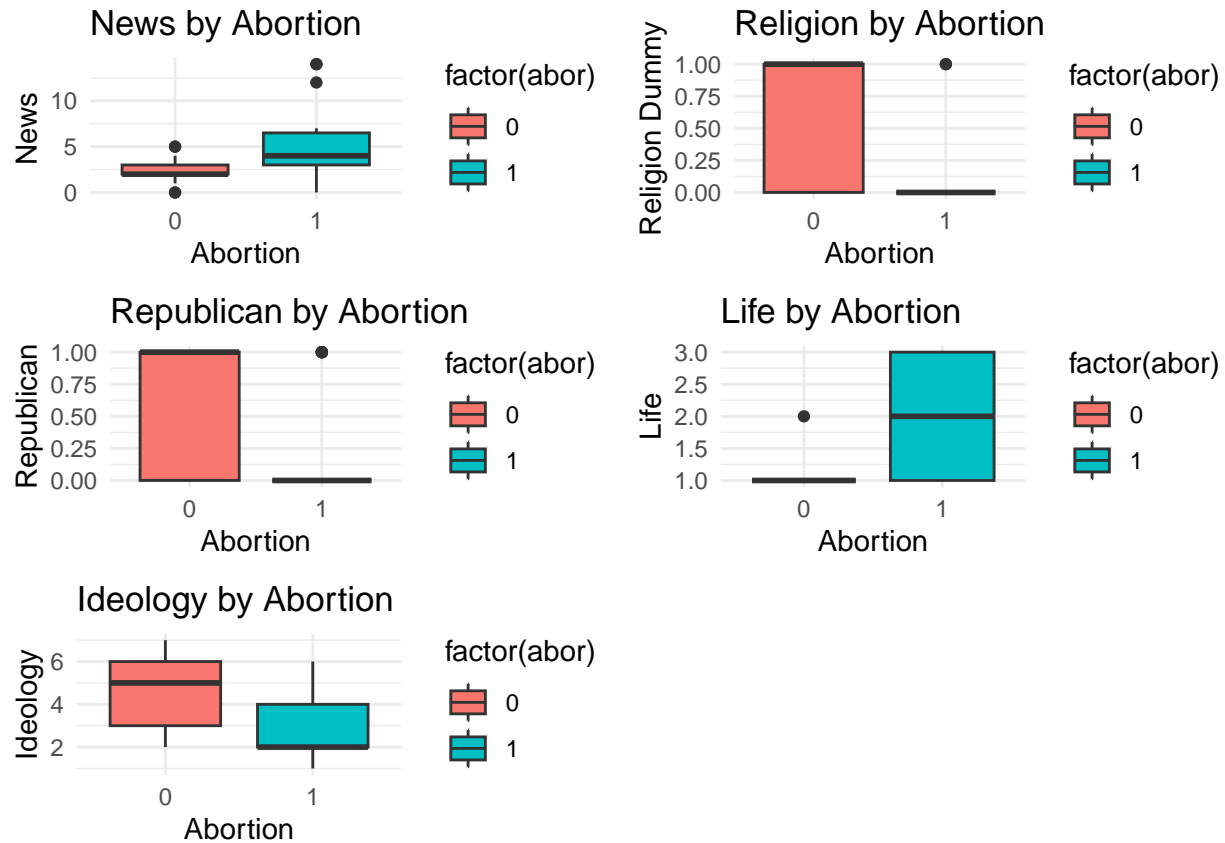
```
## religdummy
## 0 1
## 51 9
```

```
library(gridExtra)
```

```
##
## Attaching package: 'gridExtra'
```

```
## The following object is masked from 'package:dplyr':  
##  
##      combine
```

```
# News by Abortion  
p1 <- ggplot(Students, aes(x = factor(abor), y = news, fill = factor(abor))) +  
  geom_boxplot() +  
  labs(title = "News by Abortion", x = "Abortion", y = "News") +  
  theme_minimal()  
  
# religion by Abortion  
p2 <- ggplot(Students, aes(x = factor(abor), y = religdummy, fill = factor(abor))) +  
  geom_boxplot() +  
  labs(title = "Religion by Abortion", x = "Abortion", y = "Religion Dummy") +  
  theme_minimal()  
  
# Affiliation by Abortion  
p3 <- ggplot(Students, aes(x = factor(abor), y = affil2, fill = factor(abor))) +  
  geom_boxplot() +  
  labs(title = "Republican by Abortion", x = "Abortion", y = "Republican") +  
  theme_minimal()  
  
# Life by Abortion  
p4 <- ggplot(Students, aes(x = factor(abor), y = life, fill = factor(abor))) +  
  geom_boxplot() +  
  labs(title = "Life by Abortion", x = "Abortion", y = "Life") +  
  theme_minimal()  
  
# Ideol by Abortion  
p5 <- ggplot(Students, aes(x = factor(abor), y = ideol, fill = factor(abor))) +  
  geom_boxplot() +  
  labs(title = "Ideology by Abortion", x = "Abortion", y = "Ideology") +  
  theme_minimal()  
  
grid.arrange(p1, p2, p3, p4, p5, ncol = 2)
```



```
# Check the distribution of the response variable
table(abor)
```

```
## abor
## 0 1
## 13 47
```

```
# Number of observations in each category
length(abor[which(abor==1)])
```

```
## [1] 47
```

```
length(abor[which(abor==0)])
```

```
## [1] 13
```

Interpretation: The data set is fairly imbalanced, with roughly 3.6 the amount of students who supported abortion in the first 3 months of pregnancy, than those who did not support it. Based off of guidelines, the data set should contain at least 10 outcomes of each type for every explanatory variable, meaning that our model should only have 1 explanatory variable. This is conservative and might be unrealistic since we can't collect more data and still need to make an accurate model, which might have more terms.

Manual Stepwise Regression Process

##Purposeful Selection

##Fit the Null Model (Model 1)

The null model contains only the intercept.

```
Model1 <- glm(abor ~ 1, family=binomial, data=Students)
summary(Model1)

##
## Call:
## glm(formula = abor ~ 1, family = binomial, data = Students)
##
## Coefficients:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)  1.2852      0.3134   4.101 4.11e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##      Null deviance: 62.719  on 59  degrees of freedom
## Residual deviance: 62.719  on 59  degrees of freedom
## AIC: 64.719
##
## Number of Fisher Scoring iterations: 4
```

Test Each Predictor Individually

Fit separate models with each predictor as the sole explanatory variable.

```
# Test all predictors individually
psgender= glm(abor~factor(gender),family=binomial, data=Students)
#summary(psgender) #p-value .2863
psage= glm(abor~age,family=binomial, data=Students)
#summary(psage) #p-value 0.946
pshsgpa= glm(abor~hsgpa,family=binomial, data=Students)
#summary(pshsgpa) #p-value 0.788
pscogpa= glm(abor~cogpa,family=binomial, data=Students)
#summary(pscogpa) #p-value 0.534
psdhome= glm(abor~dhome,family=binomial, data=Students)
#summary(psdhome) #p-value 0.3081
psdres= glm(abor~dres,family=binomial, data=Students)
#summary(psdres) #p-value 0.37114
pstv= glm(abor~tv,family=binomial, data=Students)
#summary(pstv) #p-value 0.143
pssport= glm(abor~sport,family=binomial, data=Students)
#summary(pssport) #p-value 0.58502
psnews= glm(abor~news,family=binomial, data=Students)
#summary(psnews) #p-value 0.0226
psaids= glm(abor~aids,family=binomial, data=Students)
```

```

#summary(psaids) #p-value 0.832161
psveg= glm(abor~veg,family=binomial, data=Students)
#summary(psveg) #p-value 0.417411
psaffil= glm(abor~factor(affil),family=binomial, data=Students)
#summary(psaffil) #p-value 0.00316 for affil2
psideol= glm(abor~ideol,family=binomial, data=Students)
#summary(psideol) #p-value= 0.000498
psrelig= glm(abor~factor(relig),family=binomial, data=Students)
#summary(psrelig) #p-value 0.00295 for relig3
psaffirm= glm(abor~affirm,family=binomial, data=Students)
#summary(psaffirm) #p-value 0.0269
pslife= glm(abor~factor(life),family=binomial, data=Students)
#summary(pslife) #p-value 0.0666 for life2, high for other classes of life

```

```

# Likelihood Ratio Tests vs Null Model
cat("=== Testing Each Predictor vs Null Model ===\n\n")

```

```

## === Testing Each Predictor vs Null Model ===

```

```

cat("Gender:\n")

```

```

## Gender:

```

```

anova(Model1, psgender, test="LRT")

```

```

## Analysis of Deviance Table
##
## Model 1: abor ~ 1
## Model 2: abor ~ factor(gender)
##   Resid. Df Resid. Dev Df Deviance Pr(>Chi)
## 1         59      62.719
## 2         58      61.554  1   1.1648  0.2805

```

```

cat("\nAge:\n")

```

```

##
## Age:

```

```

anova(Model1, psage, test="LRT")

```

```

## Analysis of Deviance Table
##
## Model 1: abor ~ 1
## Model 2: abor ~ age
##   Resid. Df Resid. Dev Df Deviance Pr(>Chi)
## 1         59      62.719
## 2         58      62.714  1 0.0046236  0.9458

```

```
cat("\nHS GPA:\n")
```

```
##  
## HS GPA:
```

```
anova(Model1, pshsgpa, test="LRT")
```

```
## Analysis of Deviance Table  
##  
## Model 1: abor ~ 1  
## Model 2: abor ~ hsgpa  
##   Resid. Df Resid. Dev Df Deviance Pr(>Chi)  
## 1         59      62.719  
## 2         58      62.645  1  0.073479  0.7863
```

```
cat("\nCollege GPA:\n")
```

```
##  
## College GPA:
```

```
anova(Model1, pscogpa, test="LRT")
```

```
## Analysis of Deviance Table  
##  
## Model 1: abor ~ 1  
## Model 2: abor ~ cogpa  
##   Resid. Df Resid. Dev Df Deviance Pr(>Chi)  
## 1         59      62.719  
## 2         58      62.333  1  0.38598  0.5344
```

```
cat("\nDistance from Home:\n")
```

```
##  
## Distance from Home:
```

```
anova(Model1, psdhome, test="LRT")
```

```
## Analysis of Deviance Table  
##  
## Model 1: abor ~ 1  
## Model 2: abor ~ dhome  
##   Resid. Df Resid. Dev Df Deviance Pr(>Chi)  
## 1         59      62.719  
## 2         58      61.339  1   1.3794  0.2402
```

```
cat("\nDistance from Residence:\n")
```

```
##  
## Distance from Residence:
```

```
anova(Model1, psdres, test="LRT")
```

```
## Analysis of Deviance Table
##
## Model 1: abor ~ 1
## Model 2: abor ~ dres
##   Resid. Df Resid. Dev Df Deviance Pr(>Chi)
## 1         59      62.719
## 2         58      61.955  1   0.7636   0.3822
```

```
cat("\nTV:\n")
```

```
##
## TV:
```

```
anova(Model1, pstv, test="LRT")
```

```
## Analysis of Deviance Table
##
## Model 1: abor ~ 1
## Model 2: abor ~ tv
##   Resid. Df Resid. Dev Df Deviance Pr(>Chi)
## 1         59      62.719
## 2         58      59.773  1   2.9454  0.08612 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
cat("\nSport:\n")
```

```
##
## Sport:
```

```
anova(Model1, pssport, test="LRT")
```

```
## Analysis of Deviance Table
##
## Model 1: abor ~ 1
## Model 2: abor ~ sport
##   Resid. Df Resid. Dev Df Deviance Pr(>Chi)
## 1         59      62.719
## 2         58      62.427  1   0.29205   0.5889
```

```
cat("\nNews (SIGNIFICANT):\n")
```

```
##
## News (SIGNIFICANT):
```

```
anova(Model1, psnews, test="LRT")
```

```
## Analysis of Deviance Table
##
## Model 1: abor ~ 1
## Model 2: abor ~ news
##   Resid. Df Resid. Dev Df Deviance Pr(>Chi)
## 1         59      62.719
## 2         58      55.389  1   7.3299 0.006782 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
cat("\nAIDS:\n")
```

```
##
## AIDS:
```

```
anova(Model1, psaid, test="LRT")
```

```
## Analysis of Deviance Table
##
## Model 1: abor ~ 1
## Model 2: abor ~ aids
##   Resid. Df Resid. Dev Df Deviance Pr(>Chi)
## 1         59      62.719
## 2         58      62.672  1 0.046696  0.8289
```

```
cat("\nVegetarian:\n")
```

```
##
## Vegetarian:
```

```
anova(Model1, psveg, test="LRT")
```

```
## Analysis of Deviance Table
##
## Model 1: abor ~ 1
## Model 2: abor ~ veg
##   Resid. Df Resid. Dev Df Deviance Pr(>Chi)
## 1         59      62.719
## 2         58      61.930  1  0.78917  0.3744
```

```
cat("\nAffiliation (SIGNIFICANT):\n")
```

```
##
## Affiliation (SIGNIFICANT):
```

```
anova(Model1, psaffil, test="LRT")
```

```
## Analysis of Deviance Table
##
## Model 1: abor ~ 1
## Model 2: abor ~ factor(affil)
##   Resid. Df Resid. Dev Df Deviance Pr(>Chi)
## 1         59      62.719
## 2         57      46.316  2   16.403 0.0002743 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
cat("\nIdeology (SIGNIFICANT):\n")
```

```
##
## Ideology (SIGNIFICANT):
```

```
anova(Model1, psideol, test="LRT")
```

```
## Analysis of Deviance Table
##
## Model 1: abor ~ 1
## Model 2: abor ~ ideol
##   Resid. Df Resid. Dev Df Deviance Pr(>Chi)
## 1         59      62.719
## 2         58      45.464  1   17.255 3.269e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
cat("\nReligion (SIGNIFICANT):\n")
```

```
##
## Religion (SIGNIFICANT):
```

```
anova(Model1, psrelig, test="LRT")
```

```
## Analysis of Deviance Table
##
## Model 1: abor ~ 1
## Model 2: abor ~ factor(relig)
##   Resid. Df Resid. Dev Df Deviance Pr(>Chi)
## 1         59      62.719
## 2         56      45.893  3   16.826 0.0007676 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
cat("\nAffirmative Action:\n")
```

```
##
## Affirmative Action:
```

```
anova(Model1, psaffirm, test="LRT")
```

```
## Analysis of Deviance Table
##
## Model 1: abor ~ 1
## Model 2: abor ~ affirm
##   Resid. Df Resid. Dev Df Deviance Pr(>Chi)
## 1         59      62.719
## 2         58      57.789  1   4.9298  0.0264 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
cat("\nLife (SIGNIFICANT):\n")
```

```
##
## Life (SIGNIFICANT):
```

```
anova(Model1, pslife, test="LRT")
```

```
## Analysis of Deviance Table
##
## Model 1: abor ~ 1
## Model 2: abor ~ factor(life)
##   Resid. Df Resid. Dev Df Deviance  Pr(>Chi)
## 1         59      62.719
## 2         57      48.432  2   14.287 0.0007899 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Interpretation: Based on the tests above, we focus on the significant predictors ($p < 0.2$): - news - affil - relig - ideol - life - affirm

#Create Dummy Variables: Based off of the summaries for the sole predictor models, we will make the following dummy variables, to account for more significant classes

```
# Create affil2: Republican dummy variable
affil2 <- ifelse(affil == 2, 1, 0) # 1 = Republican, 0 = otherwise

# Create religdummy: Religious dummy variable
religdummy <- ifelse(relig == 3, 1, 0) # 1 = relig is 3, 0 = otherwise

lifedummy=ifelse(Students$life==1,1,0)
```

Model 2: News Only

```
Model2 <- glm(abor ~ news, family=binomial, data=Students)
summary(Model2)
```

```
##
## Call:
## glm(formula = abor ~ news, family = binomial, data = Students)
##
## Coefficients:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)  -0.0385     0.5849  -0.066   0.9475
## news           0.4032     0.1769   2.280   0.0226 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##    Null deviance: 62.719  on 59  degrees of freedom
## Residual deviance: 55.389  on 58  degrees of freedom
## AIC: 59.389
##
## Number of Fisher Scoring iterations: 5
```

Model 3: Affil2 Only (Republican)

```
Model3 <- glm(abor ~ affil2, family=binomial, data=Students)
summary(Model3)
```

```
##
## Call:
## glm(formula = abor ~ affil2, family = binomial, data = Students)
##
## Coefficients:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)   2.3273     0.5238   4.443 8.88e-06 ***
## affil2        -2.7327     0.7431  -3.678 0.000235 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##    Null deviance: 62.719  on 59  degrees of freedom
## Residual deviance: 47.187  on 58  degrees of freedom
## AIC: 51.187
##
## Number of Fisher Scoring iterations: 5
```

Model 4: Religdummy Only

```
Model4 <- glm(abor ~ religdummy, family=binomial, data=Students)
summary(Model4)
```

```
##
```

```
## Call:
## glm(formula = abor ~ religdummy, family = binomial, data = Students)
##
## Coefficients:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)   2.0149     0.4346   4.636 3.55e-06 ***
## religdummy   -3.2677     0.9120  -3.583 0.00034 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##    Null deviance: 62.719  on 59  degrees of freedom
## Residual deviance: 46.480  on 58  degrees of freedom
## AIC: 50.48
##
## Number of Fisher Scoring iterations: 4
```

Model 5: Ideol Only

```
Model5 <- glm(abor ~ ideol, family=binomial, data=Students)
summary(Model5)
```

```
##
## Call:
## glm(formula = abor ~ ideol, family = binomial, data = Students)
##
## Coefficients:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)   4.4205     1.0649   4.151 3.31e-05 ***
## ideol        -0.8789     0.2524  -3.482 0.000498 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##    Null deviance: 62.719  on 59  degrees of freedom
## Residual deviance: 45.464  on 58  degrees of freedom
## AIC: 49.464
##
## Number of Fisher Scoring iterations: 5
```

Model 6: Life Only

```
Model6 <- glm(abor ~ lifedummy, family=binomial, data=Students)
summary(Model6)
```

```
##
## Call:
```

```
## glm(formula = abor ~ lifedummy, family = binomial, data = Students)
##
## Coefficients:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)   3.332      1.018   3.274  0.00106 **
## lifedummy    -2.873      1.082  -2.654  0.00796 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##    Null deviance: 62.719  on 59  degrees of freedom
## Residual deviance: 50.080  on 58  degrees of freedom
## AIC: 54.08
##
## Number of Fisher Scoring iterations: 6
```

Model 7: Affirm Only

```
Model7 <- glm(abor ~ affirm, family=binomial, data=Students)
summary(Model7)
```

```
##
## Call:
## glm(formula = abor ~ affirm, family = binomial, data = Students)
##
## Coefficients:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)   0.3567      0.4928   0.724   0.4692
## affirm        1.4625      0.6607   2.213   0.0269 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##    Null deviance: 62.719  on 59  degrees of freedom
## Residual deviance: 57.789  on 58  degrees of freedom
## AIC: 61.789
##
## Number of Fisher Scoring iterations: 4
```

Step 4: Compare Single-Predictor Models to Null Model

H_0 : Null Model is better

H_a : Model with variable is better

```
cat("=== Comparing Significant Single-Predictor Models to Null ===\n\n")
```

```
## === Comparing Significant Single-Predictor Models to Null ===
```

```
cat("News vs Null:\n")
```

```
## News vs Null:
```

```
anova(Model1, Model2, test="LRT")
```

```
## Analysis of Deviance Table
```

```
##
```

```
## Model 1: abor ~ 1
```

```
## Model 2: abor ~ news
```

```
##   Resid. Df Resid. Dev Df Deviance Pr(>Chi)
```

```
## 1         59      62.719
```

```
## 2         58      55.389  1   7.3299 0.006782 **
```

```
## ---
```

```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
cat("\nAffil2 vs Null:\n")
```

```
##
```

```
## Affil2 vs Null:
```

```
anova(Model1, Model3, test="LRT")
```

```
## Analysis of Deviance Table
```

```
##
```

```
## Model 1: abor ~ 1
```

```
## Model 2: abor ~ affil2
```

```
##   Resid. Df Resid. Dev Df Deviance  Pr(>Chi)
```

```
## 1         59      62.719
```

```
## 2         58      47.187  1   15.532 8.112e-05 ***
```

```
## ---
```

```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
cat("\nReligdummy vs Null:\n")
```

```
##
```

```
## Religdummy vs Null:
```

```
anova(Model1, Model4, test="LRT")
```

```
## Analysis of Deviance Table
```

```
##
```

```
## Model 1: abor ~ 1
```

```
## Model 2: abor ~ religdummy
```

```
##   Resid. Df Resid. Dev Df Deviance  Pr(>Chi)
```

```
## 1         59      62.719
```

```
## 2         58      46.480  1   16.239 5.584e-05 ***
```

```
## ---
```

```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
cat("\nIdeol vs Null:\n")
```

```
##  
## Ideol vs Null:
```

```
anova(Model1, Model5, test="LRT")
```

```
## Analysis of Deviance Table  
##  
## Model 1: abor ~ 1  
## Model 2: abor ~ ideol  
##   Resid. Df Resid. Dev Df Deviance Pr(>Chi)  
## 1         59      62.719  
## 2         58      45.464  1   17.255 3.269e-05 ***  
## ---  
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
cat("\nLife vs Null:\n")
```

```
##  
## Life vs Null:
```

```
anova(Model1, Model6, test="LRT")
```

```
## Analysis of Deviance Table  
##  
## Model 1: abor ~ 1  
## Model 2: abor ~ lifedummy  
##   Resid. Df Resid. Dev Df Deviance Pr(>Chi)  
## 1         59      62.719  
## 2         58      50.080  1   12.638 0.0003779 ***  
## ---  
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
cat("\nAffirm vs Null:\n")
```

```
##  
## Affirm vs Null:
```

```
anova(Model1, Model7, test="LRT")
```

```
## Analysis of Deviance Table  
##  
## Model 1: abor ~ 1  
## Model 2: abor ~ affirm  
##   Resid. Df Resid. Dev Df Deviance Pr(>Chi)  
## 1         59      62.719  
## 2         58      57.789  1    4.9298  0.0264 *  
## ---  
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Decision: All p-values are less than the alpha of .05, so we will keep all of these variables for the initial main effects model. Based off of the p-values, we have enough statistical evidence to say that the models with the sole explanatory variable are better.

Model	Explanatory Variables					Deviance	df	AIC
	1	None				62.72	59	64.72
	2	<i>news</i>				55.39	58	59.39
	3	<i>affil2</i>				47.19	58	51.19
	4	<i>relig</i>				46.48	58	50.48
	5	<i>affirm</i>				45.46	58	61.79
	6	<i>Ideol</i>				50.08	58	49.46
	7	<i>lifedummy</i>				57.79	58	54.08

#Step 1 Final: Build initial main effects model

```
init= glm(abor~news+affil2+religdummy+affirm+ideol+lifedummy,family=binomial, data=Students)
summary(init)
```

```
##
## Call:
## glm(formula = abor ~ news + affil2 + religdummy + affirm + ideol +
##      lifedummy, family = binomial, data = Students)
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)   1.9321     1.9818   0.975   0.3296
## news          1.1117     0.5434   2.046   0.0408 *
## affil2       -2.2343     1.7347  -1.288   0.1977
## religdummy   -2.2882     1.3209  -1.732   0.0832 .
## affirm        0.4405     1.1572   0.381   0.7034
## ideol        -0.6865     0.6222  -1.103   0.2699
## lifedummy    -0.4689     1.5098  -0.311   0.7561
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##      Null deviance: 62.719  on 59  degrees of freedom
## Residual deviance: 25.104  on 53  degrees of freedom
## AIC: 39.104
##
## Number of Fisher Scoring iterations: 7
```

#Step 2- Backwards Elimination

H_0 : Simple Model is better

H_a : Complex Model with news is better

```
remnews= glm(abor~+affil2+religdummy+affirm+ideol+lifedummy,family=binomial, data=Students)
anova(remnews,init, test="LRT")
```

```
## Analysis of Deviance Table
##
## Model 1: abor ~ +affil2 + religdummy + affirm + ideol + lifedummy
## Model 2: abor ~ news + affil2 + religdummy + affirm + ideol + lifedummy
##   Resid. Df Resid. Dev Df Deviance Pr(>Chi)
## 1         54      35.023
## 2         53      25.104  1   9.9188 0.001636 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
#or
Gnews=remnews$deviance-init$deviance
Gnews
```

```
## [1] 9.918825
```

```
dfnews=remnews$df.residual-init$df.residual
dfnews
```

```
## [1] 1
```

```
pvaluenews=pchisq(Gnews,dfnews,lower.tail = FALSE)
pvaluenews
```

```
## [1] 0.001635968
```

#Backwards Elimination- keep news, test affil2

H_0 : Simple Model is better

H_a : Complex Model with affil2 is better

```
baseline1= glm(abor~news+affil2+religdummy+affirm+ideol+lifedummy,family=binomial, data=Students)
remaffil2= glm(abor~news+religdummy+affirm+ideol+lifedummy,family=binomial, data=Students)
anova(remaffil2, baseline1, test="LRT")
```

```
## Analysis of Deviance Table
##
## Model 1: abor ~ news + religdummy + affirm + ideol + lifedummy
## Model 2: abor ~ news + affil2 + religdummy + affirm + ideol + lifedummy
##   Resid. Df Resid. Dev Df Deviance Pr(>Chi)
## 1         54      26.834
## 2         53      25.104  1   1.7301  0.1884
```

```
Gaffil2=remaffil2$deviance-baseline1$deviance
Gaffil2
```

```
## [1] 1.73011
```

```
dfaffil2=remaffil2$df.residual-baseline1$df.residual
dfaffil2
```

```
## [1] 1
```

```
pvalueaffil2=pchisq(Gaffil2,dfaffil2,lower.tail = FALSE)
pvalueaffil2
```

```
## [1] 0.1883968
```

##Backwards Elimination- drop affil2, test relig

H_0 : Simple Model is better

H_a : Complex Model with religdummy is better

```
baseline2= glm(abor~news+religdummy+affirm+ideol+lifedummy,family=binomial, data=Students)
remrelig= glm(abor~news+affirm+ideol+lifedummy,family=binomial, data=Students)
anova(remrelig, baseline2, test="LRT")#p-value=0.05561, consider removing relig??
```

Analysis of Deviance Table

##

Model 1: abor ~ news + affirm + ideol + lifedummy

Model 2: abor ~ news + religdummy + affirm + ideol + lifedummy

Resid. Df Resid. Dev Df Deviance Pr(>Chi)

1 55 30.498

2 54 26.834 1 3.6636 0.05561 .

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

##Backwards Elimination- keep relig, test affirm

H_0 : Simple Model is better

H_a : Complex Model with affirm is better

```
baseline4= glm(abor~news+religdummy+affirm+ideol+lifedummy,family=binomial, data=Students)
remaffirm= glm(abor~news+religdummy+ideol+lifedummy,family=binomial, data=Students)
anova(remaffirm, baseline4, test="LRT")
```

Analysis of Deviance Table

##

Model 1: abor ~ news + religdummy + ideol + lifedummy

Model 2: abor ~ news + religdummy + affirm + ideol + lifedummy

Resid. Df Resid. Dev Df Deviance Pr(>Chi)

1 55 27.121

2 54 26.834 1 0.28647 0.5925

##Backwards Elimination- remove affirm, test ideol

H_0 : Simple Model is better

H_a : Complex Model with ideol is better

```
baseline5= glm(abor~news+religdummy+ideol+lifedummy,family=binomial, data=Students)
remideol= glm(abor~news+religdummy+lifedummy,family=binomial, data=Students)
anova(remideol, baseline5, test="LRT")
```

```
## Analysis of Deviance Table
##
## Model 1: abor ~ news + religdummy + lifedummy
## Model 2: abor ~ news + religdummy + ideol + lifedummy
##   Resid. Df Resid. Dev Df Deviance Pr(>Chi)
## 1          56      36.819
## 2          55      27.121  1   9.6986 0.001844 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

##Backwards Elimination- keep ideol, test lifedummy

H_0 : Simple Model is better

H_a : Complex Model with lifedummy is better

```
baseline6= glm(abor~news+religdummy+ideol+lifedummy,family=binomial, data=Students)
remlifedum= glm(abor~news+ideol+religdummy,family=binomial, data=Students)
anova(remlifedum, baseline6, test="LRT")
```

```
## Analysis of Deviance Table
##
## Model 1: abor ~ news + ideol + religdummy
## Model 2: abor ~ news + religdummy + ideol + lifedummy
##   Resid. Df Resid. Dev Df Deviance Pr(>Chi)
## 1          56      27.153
## 2          55      27.121  1  0.032333   0.8573
```

##Backwards Elimination- drop lifedummy

H_0 : Simple Model is better

H_a : Complex Model (modelafterback) is better

```
modelafterback= glm(abor~news+religdummy+ideol,family=binomial, data=Students)
summary(modelafterback)
```

```
##
## Call:
## glm(formula = abor ~ news + religdummy + ideol, family = binomial,
##      data = Students)
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)   3.0262     1.2583   2.405  0.01617 *
## news          1.1386     0.4757   2.394  0.01669 *
## religdummy    -2.4099     1.1408  -2.112  0.03464 *
## ideol         -1.2497     0.4809  -2.599  0.00936 **
## ---
```

```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##    Null deviance: 62.719  on 59  degrees of freedom
## Residual deviance: 27.153  on 56  degrees of freedom
## AIC: 35.153
##
## Number of Fisher Scoring iterations: 7
```

```
#to verify significance
mabdrop1 =glm(abor~news+religdummy,family=binomial, data=Students)
anova(mabdrop1, modelafterback, test= "LRT")
```

```
## Analysis of Deviance Table
##
## Model 1: abor ~ news + religdummy
## Model 2: abor ~ news + religdummy + ideol
##   Resid. Df Resid. Dev Df Deviance  Pr(>Chi)
## 1          57      39.271
## 2          56      27.153  1   12.117 0.0004996 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
mabdrop2 =glm(abor~news+ideol,family=binomial, data=Students)
anova(mabdrop2, modelafterback, test= "LRT")
```

```
## Analysis of Deviance Table
##
## Model 1: abor ~ news + ideol
## Model 2: abor ~ news + religdummy + ideol
##   Resid. Df Resid. Dev Df Deviance Pr(>Chi)
## 1          57      32.014
## 2          56      27.153  1    4.8611 0.02747 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
mabdrop3 =glm(abor~religdummy+ideol,family=binomial, data=Students)
anova(mabdrop3, modelafterback, test= "LRT")
```

```
## Analysis of Deviance Table
##
## Model 1: abor ~ religdummy + ideol
## Model 2: abor ~ news + religdummy + ideol
##   Resid. Df Resid. Dev Df Deviance  Pr(>Chi)
## 1          57      39.139
## 2          56      27.153  1   11.985 0.0005362 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
#Table 2 shows results from backward elimination
```

Model	Explanatory Variables	Deviance	p-value	Models compared	Deviance Difference
1	<i>news+affil2+religdummy+affirm+ideol+lifedummy</i>	27.15	-	-	-
2	<i>affil2+religdummy+affirm+ideol+lifedummy</i>	35.92	0.0016	(2) - (1)	9.9188
1	<i>news+affil2+religdummy+affirm+ideol+lifedummy</i>	27.15	-	-	-
3	<i>news+religdummy+affirm+ideol+lifedummy</i>	26.83	0.1884	(3) - (1)	1.7301
3	<i>news+religdummy+affirm+ideol+lifedummy</i>	26.83	-	-	-
4	<i>news+affirm+ideol+lifedummy</i>	30.50	0.0556	(4) - (3)	3.6636
4	<i>news+affirm+ideol+lifedummy</i>	30.50	-	-	-
5	<i>news+religdummy+ideol+lifedummy</i>	27.12	0.5925	(5) - (4)	0.2865
5	<i>news+religdummy+ideol+lifedummy</i>	27.12	-	-	-
6	<i>news+religdummy+lifedummy</i>	36.81	0.0018	(6) - (5)	9.6986
5	<i>news+religdummy+ideol+lifedummy</i>	27.12	-	-	-
7	<i>news+religdummy+ideol</i>	27.15	0.8573	(7) - (5)	0.0323

##Step 3: Add back in variables not included in Step 1

#Add gender

H_0 : Simple Model is better

H_a : Complex Model with gender is better

```
modelafterback= glm(abor~news+religdummy+ideol,family=binomial, data=Students)
addgender= glm(abor~gender+news+religdummy+ideol,family=binomial, data=Students)
anova(modelafterback, addgender, test="LRT")
```

Analysis of Deviance Table

##

Model 1: *abor ~ news + religdummy + ideol*

Model 2: *abor ~ gender + news + religdummy + ideol*

Resid. Df Resid. Dev Df Deviance Pr(>Chi)

1 56 27.153

2 55 27.096 1 0.057307 0.8108

Interpretation: Gender not significant

#Add age

H_0 : Simple Model is better

H_a : Complex Model with age is better

```
modelafterback= glm(abor~news+religdummy+ideol,family=binomial, data=Students)
addage= glm(abor~age+news+religdummy+ideol,family=binomial, data=Students)
anova(addage, modelafterback, test="LRT")
```

Analysis of Deviance Table

##

Model 1: *abor ~ age + news + religdummy + ideol*

Model 2: *abor ~ news + religdummy + ideol*

Resid. Df Resid. Dev Df Deviance Pr(>Chi)

1 55 26.523

2 56 27.153 -1 -0.62999 0.4274

Interpretation: Age not significant

#Add hsgpa

H_0 : Simple Model is better

H_a : Complex Model with hsgpa is better

```
modelafterback= glm(abor~news+religdummy+ideol,family=binomial, data=Students)
addhsgpa= glm(abor~hsgpa+news+religdummy+ideol,family=binomial, data=Students)
anova(addhsgpa, modelafterback, test="LRT")
```

```
## Analysis of Deviance Table
```

```
##
```

```
## Model 1: abor ~ hsgpa + news + religdummy + ideol
```

```
## Model 2: abor ~ news + religdummy + ideol
```

```
##   Resid. Df Resid. Dev Df Deviance Pr(>Chi)
```

```
## 1         55      24.474
```

```
## 2         56      27.153 -1   -2.6787   0.1017
```

Interpretation: hsgpa not significant based on .1 alpha, however we will keep it in mind later

#Add cogpa

H_0 : Simple Model is better

H_a : Complex Model with cogpa is better

```
modelafterback= glm(abor~news+religdummy+ideol,family=binomial, data=Students)
addcogpa= glm(abor~cogpa+news+religdummy+ideol,family=binomial, data=Students)
anova(addcogpa, modelafterback, test="LRT")
```

```
## Analysis of Deviance Table
```

```
##
```

```
## Model 1: abor ~ cogpa + news + religdummy + ideol
```

```
## Model 2: abor ~ news + religdummy + ideol
```

```
##   Resid. Df Resid. Dev Df   Deviance Pr(>Chi)
```

```
## 1         55      27.113
```

```
## 2         56      27.153 -1  -0.039971   0.8415
```

Interpretation: cogpa not significant

#Add dhome

H_0 : Simple Model is better

H_a : Complex Model with dhome is better

```
modelafterback= glm(abor~news+religdummy+ideol,family=binomial, data=Students)
adddhome= glm(abor~dhome+news+religdummy+ideol,family=binomial, data=Students)
anova(adddhome, modelafterback, test="LRT")
```

```
## Analysis of Deviance Table
```

```
##
```

```
## Model 1: abor ~ dhome + news + religdummy + ideol
```

```
## Model 2: abor ~ news + religdummy + ideol
```

```
##   Resid. Df Resid. Dev Df   Deviance Pr(>Chi)
```

```
## 1         55      27.151
```

```
## 2         56      27.153 -1  -0.0024696   0.9604
```

Interpretation: dhome not significant

#Add dres

H_0 : Simple Model is better

H_a : Complex Model with dres is better

```
modelafterback= glm(abor~news+religdummy+ideol,family=binomial, data=Students)
addres= glm(abor~dres+news+religdummy+ideol,family=binomial, data=Students)
anova(addres, modelafterback, test="LRT")
```

```
## Analysis of Deviance Table
```

```
##
```

```
## Model 1: abor ~ dres + news + religdummy + ideol
```

```
## Model 2: abor ~ news + religdummy + ideol
```

```
##   Resid. Df Resid. Dev Df Deviance Pr(>Chi)
```

```
## 1         55      26.465
```

```
## 2         56      27.153 -1  -0.68786  0.4069
```

Interpretation: dres not significant

#Add tv

H_0 : Simple Model is better

H_a : Complex Model with religdummy is better

```
modelafterback= glm(abor~news+religdummy+ideol,family=binomial, data=Students)
addtv= glm(abor~tv+news+religdummy+ideol,family=binomial, data=Students)
anova(addtv, modelafterback, test="LRT")
```

```
## Analysis of Deviance Table
```

```
##
```

```
## Model 1: abor ~ tv + news + religdummy + ideol
```

```
## Model 2: abor ~ news + religdummy + ideol
```

```
##   Resid. Df Resid. Dev Df Deviance Pr(>Chi)
```

```
## 1         55      26.290
```

```
## 2         56      27.153 -1  -0.86369  0.3527
```

Interpretation: tv not significant

#Add sport

H_0 : Simple Model is better

H_a : Complex Model with sport is better

```
modelafterback= glm(abor~news+religdummy+ideol,family=binomial, data=Students)
addsport= glm(abor~sport+news+religdummy+ideol,family=binomial, data=Students)
anova(addsport, modelafterback, test="LRT")
```

```
## Analysis of Deviance Table
```

```
##
```

```
## Model 1: abor ~ sport + news + religdummy + ideol
```

```
## Model 2: abor ~ news + religdummy + ideol
```

```
##   Resid. Df Resid. Dev Df Deviance Pr(>Chi)
```

```
## 1         55      27.129
```

```
## 2         56      27.153 -1  -0.024006  0.8769
```

Interpretation: sport not significant

#Add aids

H_0 : Simple Model is better

H_a : Complex Model with aids is better

```
modelafterback= glm(abor~news+religdummy+ideol,family=binomial, data=Students)
addaids= glm(abor~aids+news+religdummy+ideol,family=binomial, data=Students)
anova(addaids, modelafterback, test="LRT")
```

Analysis of Deviance Table

##

Model 1: abor ~ aids + news + religdummy + ideol

Model 2: abor ~ news + religdummy + ideol

Resid. Df Resid. Dev Df Deviance Pr(>Chi)

1 55 26.911

2 56 27.153 -1 -0.24264 0.6223

Interpretation: aids not significant

#Add veg

H_0 : Simple Model is better

H_a : Complex Model with veg is better

```
modelafterback= glm(abor~news+religdummy+ideol,family=binomial, data=Students)
addveg= glm(abor~veg+news+religdummy+ideol,family=binomial, data=Students)
anova(addveg, modelafterback, test="LRT")
```

Analysis of Deviance Table

##

Model 1: abor ~ veg + news + religdummy + ideol

Model 2: abor ~ news + religdummy + ideol

Resid. Df Resid. Dev Df Deviance Pr(>Chi)

1 55 27.055

2 56 27.153 -1 -0.097849 0.7544

Interpretation: veg not significant

#Checking stepAIC for fun

```
#library(MASS)
```

```
#stepAIC(init)
```

#Step 4: Add interaction

H_0 : Simple Model is better

H_a : Model with interaction is better

```
testfinal = glm(abor~news+religdummy+ideol,family=binomial, data=Students)
testfinalint1 = glm(abor~news+religdummy+ideol+news*religdummy,family=binomial, data=Students)
summary(testfinalint1)
```

```
##
## Call:
## glm(formula = abor ~ news + religdummy + ideol + news * religdummy,
##      family = binomial, data = Students)
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    2.1832     1.3184   1.656   0.0977 .
## news           1.4064     0.5591   2.516   0.0119 *
## religdummy     3.5276     4.7793   0.738   0.4605
## ideol          -1.1413     0.5008  -2.279   0.0227 *
## news:religdummy -2.2859     1.7626  -1.297   0.1947
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##      Null deviance: 62.719  on 59  degrees of freedom
## Residual deviance: 23.594  on 55  degrees of freedom
## AIC: 33.594
##
## Number of Fisher Scoring iterations: 8
```

```
anova(testfinal, testfinalint1)
```

```
## Analysis of Deviance Table
##
## Model 1: abor ~ news + religdummy + ideol
## Model 2: abor ~ news + religdummy + ideol + news * religdummy
##   Resid. Df Resid. Dev Df Deviance Pr(>Chi)
## 1         56      27.153
## 2         55      23.594  1    3.5588  0.05923 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
testfinalint2 = glm(abor~news+religdummy+ideol+news*ideol,family=binomial, data=Students)
summary(testfinalint2)
```

```
##
## Call:
## glm(formula = abor ~ news + religdummy + ideol + news * ideol,
##      family = binomial, data = Students)
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    3.14611     1.95833   1.607   0.1082
## news           1.08042     0.85877   1.258   0.2084
## religdummy     -2.41574     1.14113  -2.117   0.0343 *
## ideol          -1.28776     0.67704  -1.902   0.0572 .
## news:ideol      0.01534     0.18918   0.081   0.9354
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
```

```
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 62.719 on 59 degrees of freedom
## Residual deviance: 27.147 on 55 degrees of freedom
## AIC: 37.147
##
## Number of Fisher Scoring iterations: 7
```

```
anova(testfinal, testfinalint2)
```

```
## Analysis of Deviance Table
##
## Model 1: abor ~ news + religdummy + ideol
## Model 2: abor ~ news + religdummy + ideol + news * ideol
## Resid. Df Resid. Dev Df Deviance Pr(>Chi)
## 1 56 27.153
## 2 55 27.147 1 0.0065932 0.9353
```

```
testfinalint3 = glm(abor~news+religdummy+ideol+religdummy*ideol,family=binomial, data=Students)
```

```
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
```

```
summary(testfinalint3)
```

```
##
## Call:
## glm(formula = abor ~ news + religdummy + ideol + religdummy *
## ideol, family = binomial, data = Students)
##
## Coefficients:
## Estimate Std. Error z value Pr(>|z|)
## (Intercept) 1.8459 1.3384 1.379 0.1678
## news 1.2891 0.5374 2.399 0.0165 *
## religdummy 51.8086 7306.3336 0.007 0.9943
## ideol -0.9583 0.5144 -1.863 0.0625 .
## religdummy:ideol -18.2156 2435.4448 -0.007 0.9940
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 62.719 on 59 degrees of freedom
## Residual deviance: 22.743 on 55 degrees of freedom
## AIC: 32.743
##
## Number of Fisher Scoring iterations: 19
```

```
anova(testfinal, testfinalint3)
```

```
## Analysis of Deviance Table
##
```

```
## Model 1: abor ~ news + religdummy + ideol
## Model 2: abor ~ news + religdummy + ideol + religdummy * ideol
##   Resid. Df Resid. Dev Df Deviance Pr(>Chi)
## 1         56      27.153
## 2         55      22.743  1    4.4101  0.03573 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Model	Explanatory Variables	Deviance	df	AIC
1	<i>news+religdummy+ideol</i>	27.15	56	35.15
2	<i>news+religdummy+ideol+(news:religdummy)</i>	23.59	55	33.59
3	<i>news+religdummy+ideol+(news:ideol)</i>	27.15	55	37.15
4	<i>news+religdummy+ideol+(religdummy:ideol)</i>	22.74	55	32.74

Final Model Selection

```
Model404 <- glm(abor~news+religdummy+ideol,family=binomial, data=Students)
summary(Model404)
```

```
##
## Call:
## glm(formula = abor ~ news + religdummy + ideol, family = binomial,
##      data = Students)
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)   3.0262     1.2583   2.405  0.01617 *
## news          1.1386     0.4757   2.394  0.01669 *
## religdummy   -2.4099     1.1408  -2.112  0.03464 *
## ideol        -1.2497     0.4809  -2.599  0.00936 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##      Null deviance: 62.719  on 59  degrees of freedom
## Residual deviance: 27.153  on 56  degrees of freedom
## AIC: 35.153
##
## Number of Fisher Scoring iterations: 7
```

Goodness of Fit Test

H_0 : The model provides an adequate fit to the data

H_a : The model does not adequately fit the data

```

Goodness <- function(x) {
  G1 <- x$deviance
  df <- x$df.residual
  pvalue <- pchisq(G1, df, lower.tail = FALSE)
  cat("Deviance =", G1, "df =", df, " p-value =", pvalue, "\n")
}

Goodness(Model404)

```

```
## Deviance = 27.1532 df = 56 p-value = 0.9995998
```

Interpretation: p-value > 0.05, we fail to reject H_0 , suggesting the model fits adequately.

Global Test for Main Effects

Test whether all predictors jointly have an effect:

$$H_0 : \beta_{news} = \beta_{affil2} = \beta_{religdummy} = 0$$

$$H_a : \text{At least one } \beta_i \neq 0$$

```
anova(Model11, Model404, test="LRT")
```

```

## Analysis of Deviance Table
##
## Model 1: abor ~ 1
## Model 2: abor ~ news + religdummy + ideol
##   Resid. Df Resid. Dev Df Deviance Pr(>Chi)
## 1         59      62.719
## 2         56      27.153  3   35.566 9.252e-08 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

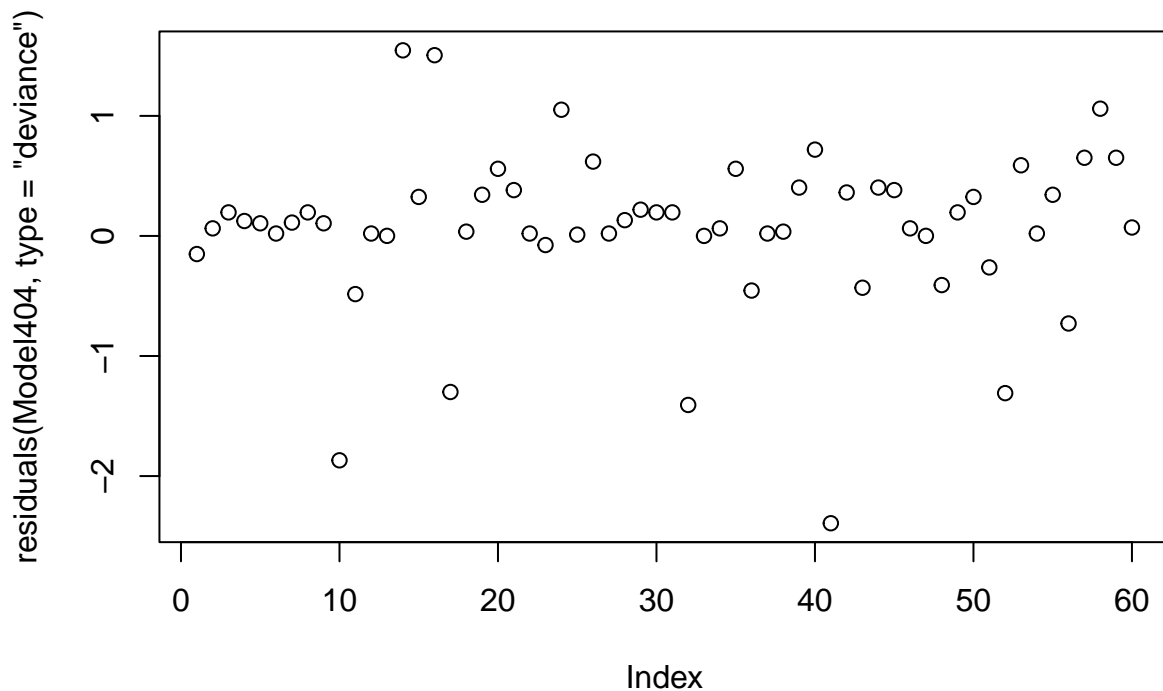
```

Decision: p-value < $\alpha = 0.05$, we reject H_0 .

Conclusion: We have enough evidence that at least one explanatory variable has an effect on abortion attitudes.

#Residual test

```
plot(residuals(Model404, type = "deviance"))
```



Train Test Split

```
# We cant do as we created dummy variables
```

```
#ROC
```

```
library(pROC) # Model with interaction
```

```
## Type 'citation("pROC")' for a citation.
```

```
##
```

```
## Attaching package: 'pROC'
```

```
## The following objects are masked from 'package:stats':
```

```
##
```

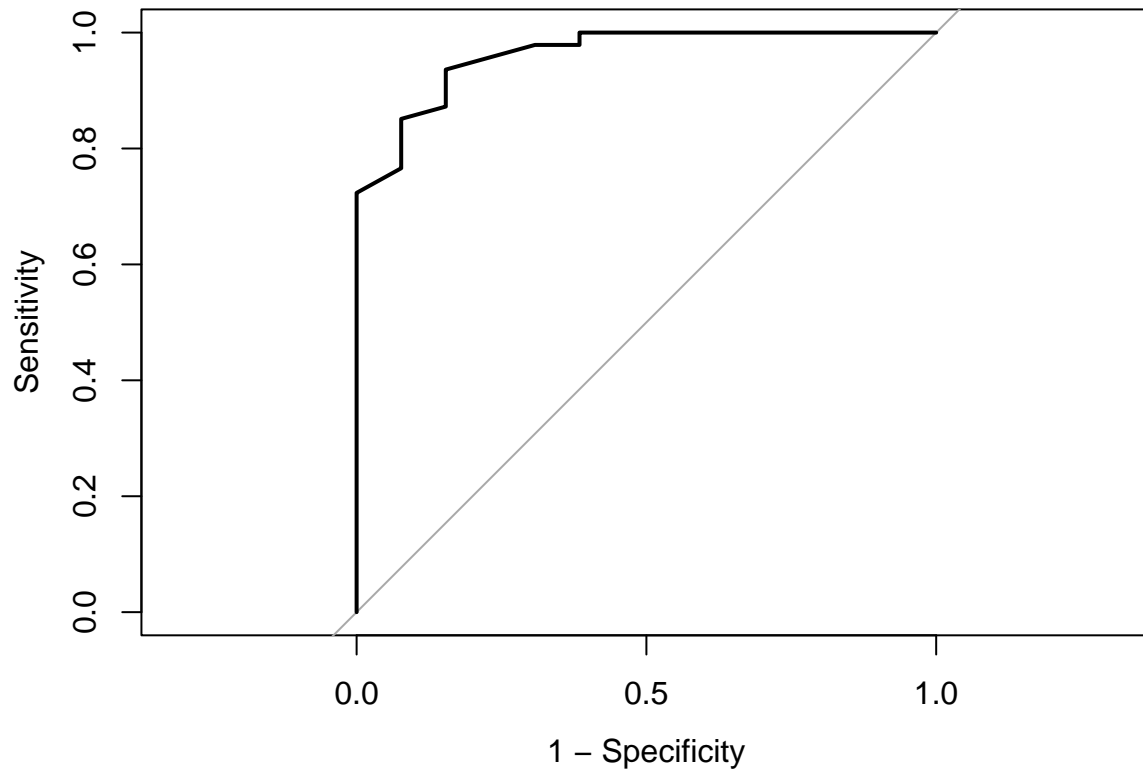
```
## cov, smooth, var
```

```
rocplot <- roc(Students$abor ~ fitted(testfinalint3), data=Students)
```

```
## Setting levels: control = 0, case = 1
```

```
## Setting direction: controls < cases
```

```
plot.roc(rocplot, legacy.axes=TRUE)
```



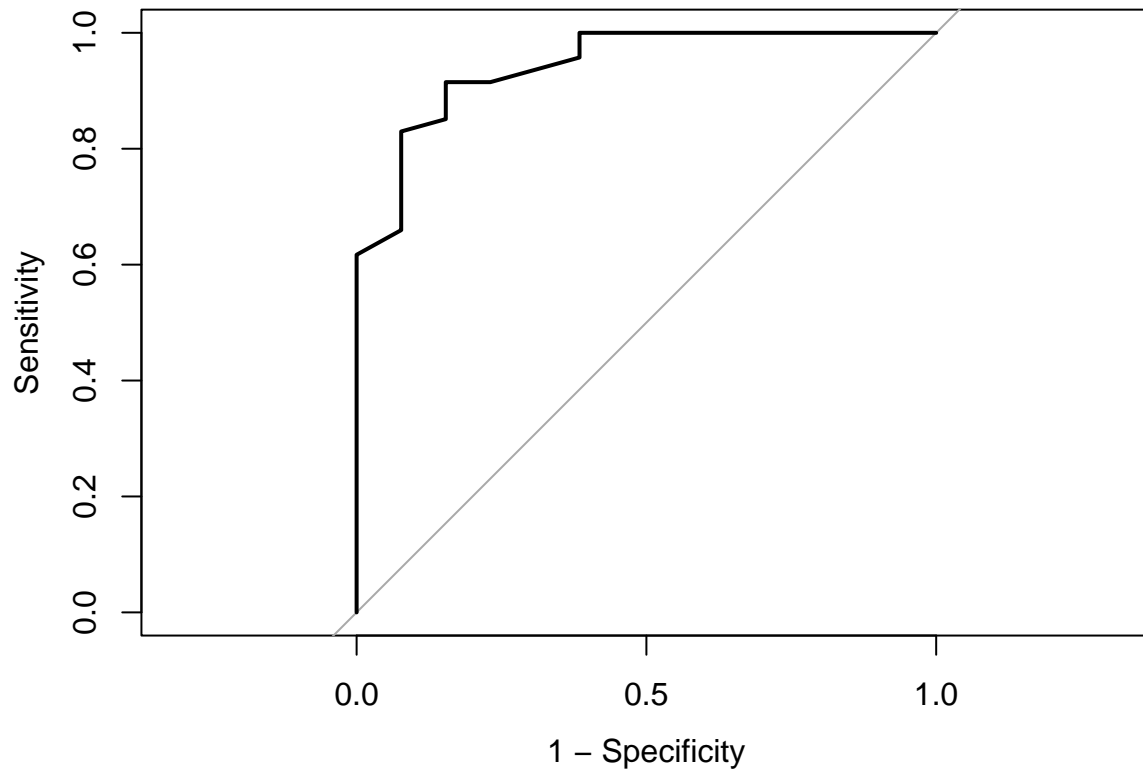
```
auc(rocplot)
```

```
## Area under the curve: 0.9615
```

```
library(pROC) #Model 404 final model without interaction  
rocplot <- roc(Students$abor ~ fitted(testfinal), data=Students)
```

```
## Setting levels: control = 0, case = 1  
## Setting direction: controls < cases
```

```
plot.roc(rocplot, legacy.axes=TRUE)
```



```
auc(rocplot)
```

```
## Area under the curve: 0.9435
```

```
summary(testfinal)
```

```
##
## Call:
## glm(formula = abort ~ news + religdummy + ideol, family = binomial,
##      data = Students)
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)   3.0262     1.2583   2.405  0.01617 *
## news          1.1386     0.4757   2.394  0.01669 *
## religdummy   -2.4099     1.1408  -2.112  0.03464 *
## ideol        -1.2497     0.4809  -2.599  0.00936 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##      Null deviance: 62.719  on 59  degrees of freedom
## Residual deviance: 27.153  on 56  degrees of freedom
```

```
## AIC: 35.153
##
## Number of Fisher Scoring iterations: 7
```

```
summary(testfinalint3)
```

```
##
## Call:
## glm(formula = abor ~ news + religdummy + ideol + religdummy *
##       ideol, family = binomial, data = Students)
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)      1.8459      1.3384   1.379   0.1678
## news              1.2891      0.5374   2.399   0.0165 *
## religdummy       51.8086    7306.3336   0.007   0.9943
## ideol            -0.9583      0.5144  -1.863   0.0625 .
## religdummy:ideol -18.2156    2435.4448  -0.007   0.9940
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##    Null deviance: 62.719  on 59  degrees of freedom
## Residual deviance: 22.743  on 55  degrees of freedom
## AIC: 32.743
##
## Number of Fisher Scoring iterations: 19
```

```
detach(Students)
```

Conclusion: Both models show excellent predictive power with AUC values above 0.9. The difference in AUC between the two models is 0.018, which is negligible.

For practical purposes, both models perform similarly, and the choice between them should consider the balance between model complexity and interpretability. Given the similar predictive performance and the results, the simpler model without interaction (model404) may be preferred for its interpretability.