# Output Version A

tally(~Elect2020, data=states)

## Elect2020  
## D R   
## 25 25

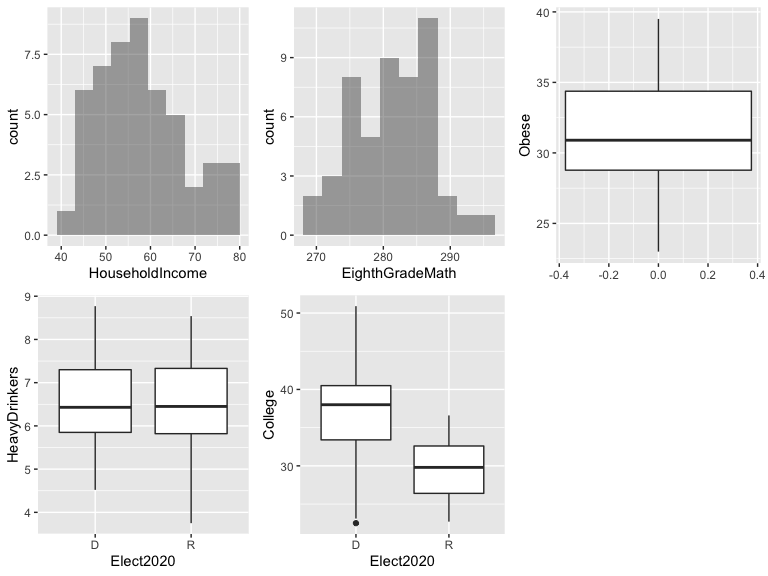
tally(~Region, data=states)

## Region  
## MW NE S W   
## 13 11 13 13

tally(Elect2020~EnoughFruit, data=states)

## EnoughFruit  
## Elect2020 HighFruit LowFruit  
## D 17 8  
## R 2 23

g1 <- gf\_histogram(~HouseholdIncome,bins=10,data=states); g2 <-gf\_histogram(~EighthGradeMath, bins=10,data=states)  
g4 <-gf\_boxplot(~Obese,data=states); g5 <-gf\_boxplot(HeavyDrinkers~Elect2020, data=states)  
g6 <- gf\_boxplot(College~Elect2020, data=states)  
grid.arrange(g1, g2, g4, g5, g6, ncol=3, nrow=2)



## Model 1

model1 <- glm(Elect2020~EnoughFruit+HouseholdIncome+College+EighthGradeMath+Obese, family=binomial, data=states); summary(model1)

##   
## Call:  
## glm(formula = Elect2020 ~ EnoughFruit + HouseholdIncome + College +   
## EighthGradeMath + Obese, family = binomial, data = states)  
##   
## Deviance Residuals:   
## Min 1Q Median 3Q Max   
## -1.76158 -0.34358 0.02668 0.51421 1.83729   
##   
## Coefficients:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) -122.60203 49.95234 -2.454 0.0141 \*  
## EnoughFruitLowFruit 1.02753 1.31418 0.782 0.4343   
## HouseholdIncome 0.04994 0.07888 0.633 0.5266   
## College -0.51757 0.22493 -2.301 0.0214 \*  
## EighthGradeMath 0.41470 0.17401 2.383 0.0172 \*  
## Obese 0.60365 0.27276 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: 69.315 on 49 degrees of freedom  
## Residual deviance: 31.140 on 44 degrees of freedom  
## AIC: 43.14  
##   
## Number of Fisher Scoring iterations: 6

anova(model1, test="Chisq")

## Analysis of Deviance Table  
##   
## Model: binomial, link: logit  
##   
## Response: Elect2020  
##   
## Terms added sequentially (first to last)  
##   
##   
## Df Deviance Resid. Df Resid. Dev Pr(>Chi)   
## NULL 49 69.315   
## EnoughFruit 1 21.1245 48 48.190 4.304e-06 \*\*\*  
## HouseholdIncome 1 4.6738 47 43.516 0.03063 \*   
## College 1 0.4193 46 43.097 0.51728   
## EighthGradeMath 1 4.5842 45 38.513 0.03227 \*   
## Obese 1 7.3732 44 31.140 0.00662 \*\*   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

1-pchisq(69.315-31.14, df=5)

## [1] 3.479797e-07

## Model 2

model2<- glm(Elect2020~College+EighthGradeMath+Obese, family=binomial, data=states)   
summary(model2)

##   
## Call:  
## glm(formula = Elect2020 ~ College + EighthGradeMath + Obese,   
## family = binomial, data = states)  
##   
## Deviance Residuals:   
## Min 1Q Median 3Q Max   
## -2.02399 -0.38416 0.01832 0.50614 1.85647   
##   
## Coefficients:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) -113.3785 44.3293 -2.558 0.01054 \*   
## College -0.5027 0.1840 -2.732 0.00629 \*\*  
## EighthGradeMath 0.3936 0.1617 2.434 0.01492 \*   
## Obese 0.5970 0.2026 2.947 0.00321 \*\*  
## ---  
##   
## Null deviance: 69.315 on 49 degrees of freedom  
## Residual deviance: 32.348 on 46 degrees of freedom  
## AIC: 40.348  
##   
## Number of Fisher Scoring iterations: 6

anova(model2, test="Chisq")

## Analysis of Deviance Table  
##   
## Response: Elect2020  
##   
## Terms added sequentially (first to last)  
##   
##   
## Df Deviance Resid. Df Resid. Dev Pr(>Chi)   
## NULL 49 69.315   
## College 1 17.5298 48 51.785 2.828e-05 \*\*\*  
## EighthGradeMath 1 5.3671 47 46.418 0.0205195 \*   
## Obese 1 14.0699 46 32.348 0.0001761 \*\*\*  
## ---

1-pchisq(69.3-32.3, df=3)

## [1] 4.60138e-08

anova(model2, model1, test="Chisq")

## Analysis of Deviance Table  
##   
## Model 1: Elect2020 ~ College + EighthGradeMath + Obese  
## Model 2: Elect2020 ~ EnoughFruit + HouseholdIncome + College + EighthGradeMath +   
## Obese  
## Resid. Df Resid. Dev Df Deviance Pr(>Chi)  
## 1 46 32.348   
## 2 44 31.140 2 1.2081 0.5466

## Model 3

model3<- glm(Elect2020~College+Obese, family=binomial, data=states)  
summary(model3)

##   
## Call:  
## glm(formula = Elect2020 ~ College + Obese, family = binomial,   
## data = states)  
##   
## Deviance Residuals:   
## Min 1Q Median 3Q Max   
## -2.16411 -0.51599 0.01598 0.71920 1.88606   
##   
## Coefficients:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) -5.8365 5.2336 -1.115 0.26477   
## College -0.1977 0.0845 -2.340 0.01929 \*   
## Obese 0.3915 0.1438 2.722 0.00649 \*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Null deviance: 69.315 on 49 degrees of freedom  
## Residual deviance: 41.781 on 47 degrees of freedom  
## AIC: 47.781  
##   
## Number of Fisher Scoring iterations: 5

anova(model3, test="Chisq")

## Analysis of Deviance Table  
##   
## Model: binomial, link: logit  
##   
## Response: Elect2020  
##   
## Terms added sequentially (first to last)  
##   
##   
## Df Deviance Resid. Df Resid. Dev Pr(>Chi)   
## NULL 49 69.315   
## College 1 17.530 48 51.785 2.828e-05 \*\*\*  
## Obese 1 10.004 47 41.781 0.001562 \*\*   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

1-pchisq(69.3-41.8, df=2)

## [1] 1.067704e-06

confint(model3)

## Waiting for profiling to be done...

## 2.5 % 97.5 %  
## (Intercept) -16.6375245 4.26306336  
## College -0.3882890 -0.04806454  
## Obese 0.1373755 0.71187065

## Model 4

model4<- glm(Elect2020~Region, family=binomial, data=states)  
summary(model4)

##   
## Call:  
## glm(formula = Elect2020 ~ Region, family = binomial, data = states)  
##   
## Deviance Residuals:   
## Min 1Q Median 3Q Max   
## -1.9348 -0.9854 0.2889 0.8576 1.3824   
##   
## Coefficients:  
## Estimate Std. Error z value Pr(>|z|)  
## (Intercept) 0.8109 0.6009 1.349 0.177  
## RegionNE -19.3770 1966.6496 -0.010 0.992  
## RegionS 0.8938 0.9757 0.916 0.360  
## RegionW -1.2809 0.8283 -1.546 0.122  
##   
## (Dispersion parameter for binomial family taken to be 1)  
##   
## Null deviance: 69.315 on 49 degrees of freedom  
## Residual deviance: 44.534 on 46 degrees of freedom  
## AIC: 52.534  
##   
## Number of Fisher Scoring iterations: 17

anova(model4, test="Chisq")

## Analysis of Deviance Table  
##   
## Model: binomial, link: logit  
##   
## Response: Elect2020  
##   
## Terms added sequentially (first to last)  
##   
##   
## Df Deviance Resid. Df Resid. Dev Pr(>Chi)   
## NULL 49 69.315   
## Region 3 24.781 46 44.534 1.716e-05 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

pred.success1 <- ifelse(fitted(model1)>0.5,1,0)  
tally(~pred.success1+model1$y, format="proportion")

## model1$y  
## pred.success1 0 1  
## 0 0.40 0.06  
## 1 0.10 0.44

pred.success2 <- ifelse(fitted(model2)>0.5,1,0)  
tally(~pred.success2+model2$y, format="proportion")

## model2$y  
## pred.success2 0 1  
## 0 0.38 0.06  
## 1 0.12 0.44

pred.success3 <- ifelse(fitted(model3)>0.5,1,0)  
tally(~pred.success3+model3$y, format="proportion")

## model3$y  
## pred.success3 0 1  
## 0 0.38 0.06  
## 1 0.12 0.44

pred.success4 <- ifelse(fitted(model4)>0.5,1,0)  
tally(~pred.success4+model4$y, format="proportion")

## model4$y  
## pred.success4 0 1  
## 0 0.38 0.10  
## 1 0.12 0.40

# Output Version B

tally(~majority\_vax, data=states)

## majority\_vax  
## No Yes   
## 29 21

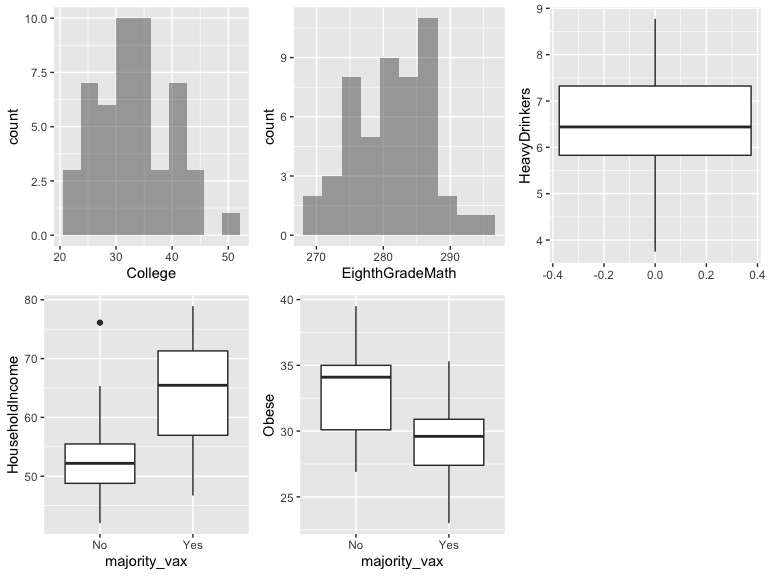
tally(~EnoughFruit, data=states)

## EnoughFruit  
## HighFruit LowFruit   
## 19 31

tally(majority\_vax~Region, data=states)

## Region  
## majority\_vax MW NE S W  
## No 10 0 12 7  
## Yes 3 11 1 6

g1 <- gf\_histogram(~College,bins=10, data=states); g2 <- gf\_histogram(~EighthGradeMath, bins=10, data=states)  
g3 <- gf\_boxplot(~HeavyDrinkers, data=states); g4 <- gf\_boxplot(HouseholdIncome~majority\_vax, data=states)   
g5 <- gf\_boxplot(Obese~majority\_vax, data=states)  
grid.arrange(g1, g2, g3, g4, g5, ncol=3, nrow=2)



## Model 1

model1 <- glm(majority\_vax~EnoughFruit+HouseholdIncome+College+EighthGradeMath+Obese, family=binomial, data=states); summary(model1)

##   
## Call:  
## glm(formula = majority\_vax ~ EnoughFruit + HouseholdIncome +   
## College + EighthGradeMath + Obese, family = binomial, data = states)  
##   
## Deviance Residuals:   
## Min 1Q Median 3Q Max   
## -2.6727 -0.4305 -0.1833 0.3045 2.0725   
##   
## Coefficients:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) 74.45524 44.44945 1.675 0.0939 .  
## EnoughFruitLowFruit -1.61633 1.22808 -1.316 0.1881   
## HouseholdIncome 0.02727 0.07838 0.348 0.7279   
## College 0.42922 0.20332 2.111 0.0348 \*  
## EighthGradeMath -0.29446 0.16294 -1.807 0.0707 .  
## Obese -0.21361 0.23166 -0.922 0.3565   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## (Dispersion parameter for binomial family taken to be 1)  
##   
## Null deviance: 68.029 on 49 degrees of freedom  
## Residual deviance: 30.264 on 44 degrees of freedom  
## AIC: 42.264  
##   
## Number of Fisher Scoring iterations: 6

anova(model1, test="Chisq")

## Analysis of Deviance Table  
##   
## Model: binomial, link: logit  
##   
## Response: majority\_vax  
##   
## Terms added sequentially (first to last)  
##   
##   
## Df Deviance Resid. Df Resid. Dev Pr(>Chi)   
## NULL 49 68.029   
## EnoughFruit 1 24.0632 48 43.966 9.322e-07 \*\*\*  
## HouseholdIncome 1 6.8422 47 37.124 0.008903 \*\*   
## College 1 2.0510 46 35.073 0.152104   
## EighthGradeMath 1 3.9764 45 31.096 0.046141 \*   
## Obese 1 0.8319 44 30.264 0.361738   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

1-pchisq(68.03-30.26, df=5)

## [1] 4.196683e-07

## Model 2

model2<- glm(majority\_vax~College+EighthGradeMath+Obese, family=binomial, data=states); summary(model2)

##   
## Call:  
## glm(formula = majority\_vax ~ College + EighthGradeMath + Obese,   
## family = binomial, data = states)  
##   
## Deviance Residuals:   
## Min 1Q Median 3Q Max   
## -2.4524 -0.5061 -0.1777 0.2885 2.0582   
##   
## Coefficients:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) 76.2501 40.9731 1.861 0.06275 .   
## College 0.5011 0.1742 2.876 0.00403 \*\*  
## EighthGradeMath -0.2902 0.1533 -1.894 0.05826 .   
## Obese -0.3653 0.1724 -2.118 0.03415 \*   
## ---  
##   
## Null deviance: 68.029 on 49 degrees of freedom  
## Residual deviance: 32.090 on 46 degrees of freedom  
## AIC: 40.09  
##   
## Number of Fisher Scoring iterations: 6

anova(model2, test="Chisq")

## Analysis of Deviance Table  
##   
## Response: majority\_vax  
##   
## Terms added sequentially (first to last)  
##   
##   
## Df Deviance Resid. Df Resid. Dev Pr(>Chi)   
## NULL 49 68.029   
## College 1 25.2289 48 42.800 5.091e-07 \*\*\*  
## EighthGradeMath 1 4.8201 47 37.980 0.02813 \*   
## Obese 1 5.8904 46 32.090 0.01522 \*   
## ---

1-pchisq(68.03-32.09, df=3)

## [1] 7.71033e-08

anova(model2, model1, test="Chisq")

## Analysis of Deviance Table  
##   
## Model 1: majority\_vax ~ College + EighthGradeMath + Obese  
## Model 2: majority\_vax ~ EnoughFruit + HouseholdIncome + College + EighthGradeMath +   
## Obese  
## Resid. Df Resid. Dev Df Deviance Pr(>Chi)  
## 1 46 32.090   
## 2 44 30.264 2 1.8254 0.4014

## Model 3

model3<- glm(majority\_vax~College+Obese, family=binomial, data=states)  
summary(model3)

##   
## Call:  
## glm(formula = majority\_vax ~ College + Obese, family = binomial,   
## data = states)  
##   
## Deviance Residuals:   
## Min 1Q Median 3Q Max   
## -1.8983 -0.6147 -0.1754 0.4919 2.7603   
##   
## Coefficients:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) -0.6411 5.6005 -0.114 0.90886   
## College 0.3002 0.1038 2.891 0.00384 \*\*  
## Obese -0.3121 0.1424 -2.192 0.02842 \*   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## (Dispersion parameter for binomial family taken to be 1)  
##   
## Null deviance: 68.029 on 49 degrees of freedom  
## Residual deviance: 37.026 on 47 degrees of freedom  
## AIC: 43.026  
##   
## Number of Fisher Scoring iterations: 6

anova(model3, test="Chisq")

## Analysis of Deviance Table  
##   
## Model: binomial, link: logit  
##   
## Response: majority\_vax  
##   
## Terms added sequentially (first to last)  
##   
##   
## Df Deviance Resid. Df Resid. Dev Pr(>Chi)   
## NULL 49 68.029   
## College 1 25.2289 48 42.800 5.091e-07 \*\*\*  
## Obese 1 5.7744 47 37.026 0.01626 \*   
## ---

1-pchisq(68.03-37.03, df=2)

## [1] 1.855391e-07

confint(model3)

## Waiting for profiling to be done...

## 2.5 % 97.5 %  
## (Intercept) -11.9978335 10.47524976  
## College 0.1253805 0.54612142  
## Obese -0.6276552 -0.05461934

## Model 4

model4<- glm(majority\_vax~Region, family=binomial, data=states)  
summary(model4)

##   
## Call:  
## glm(formula = majority\_vax ~ Region, family = binomial, data = states)  
##   
## Deviance Residuals:   
## Min 1Q Median 3Q Max   
## -1.11269 -0.72438 -0.40011 0.00013 2.26493   
##   
## Coefficients:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) -1.2040 0.6583 -1.829 0.0674 .  
## RegionNE 19.7700 1966.6496 0.010 0.9920   
## RegionS -1.2809 1.2315 -1.040 0.2983   
## RegionW 1.0498 0.8619 1.218 0.2232   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## (Dispersion parameter for binomial family taken to be 1)  
##   
## Null deviance: 68.029 on 49 degrees of freedom  
## Residual deviance: 39.041 on 46 degrees of freedom  
## AIC: 47.041  
##   
## Number of Fisher Scoring iterations: 17

anova(model4, test="Chisq")

## Analysis of Deviance Table  
##   
## Model: binomial, link: logit  
##   
## Response: majority\_vax  
##   
## Terms added sequentially (first to last)  
##   
##   
## Df Deviance Resid. Df Resid. Dev Pr(>Chi)   
## NULL 49 68.029   
## Region 3 28.988 46 39.041 2.252e-06 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

pred.success1 <- ifelse(fitted(model1)>0.5,1,0)  
tally(~pred.success1+model1$y, format="proportion")

## model1$y  
## pred.success1 0 1  
## 0 0.54 0.06  
## 1 0.04 0.36

pred.success2 <- ifelse(fitted(model2)>0.5,1,0)  
tally(~pred.success2+model2$y, format="proportion")

## model2$y  
## pred.success2 0 1  
## 0 0.56 0.08  
## 1 0.02 0.34

pred.success3 <- ifelse(fitted(model3)>0.5,1,0)  
tally(~pred.success3+model3$y, format="proportion")

## model3$y  
## pred.success3 0 1  
## 0 0.52 0.10  
## 1 0.06 0.32

pred.success4 <- ifelse(fitted(model4)>0.5,1,0)  
tally(~pred.success4+model4$y, format="proportion")

## model4$y  
## pred.success4 0 1  
## 0 0.58 0.20  
## 1 0.00 0.22