Untitled

Samantha Rodriguez

4/4/2022

```
data <- hospitalData
head(hospitalData)
```

```
##       County County.Value State.Rate Year Content.Area       Date
## 1     mohave           46      37.22 2019       Asthma 2019-04-10
## 2       gila           46      37.22 2019       Asthma 2019-04-10
## 3 santa cruz           45      37.22 2019       Asthma 2019-04-10
## 4   maricopa           39      37.22 2019       Asthma 2019-04-10
## 5    cochise           39      37.22 2019       Asthma 2019-04-10
## 6       yuma           34      37.22 2019       Asthma 2019-04-10
##      County.Year
## 1     mohave2019
## 2       gila2019
## 3 santa cruz2019
## 4   maricopa2019
## 5    cochise2019
## 6       yuma2019
```

```
str(hospitalData)
```

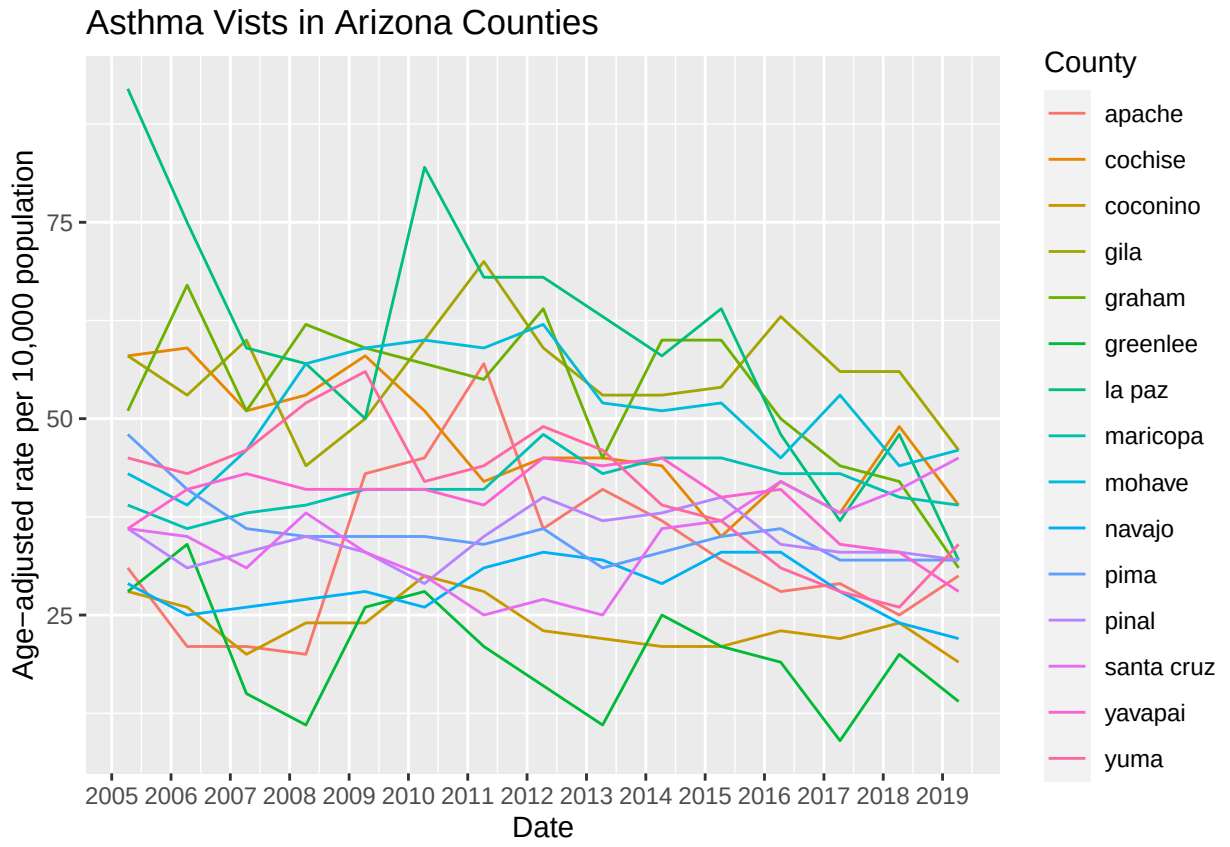
```
## 'data.frame':    1535 obs. of  7 variables:
##  $ County      : Factor w/ 15 levels "apache","cochise",..: 9 4 13 8 2 15 7 11 12 5 ...
##  $ County.Value: int  46 46 45 39 39 34 32 32 32 31 ...
##  $ State.Rate  : num  37.2 37.2 37.2 37.2 37.2 ...
##  $ Year        : chr  "2019" "2019" "2019" "2019" ...
##  $ Content.Area: Factor w/ 9 levels "Asthma","Carbon Monoxide Poisoning",..: 1 1 1 1 1 1 1 1 1 1 ...
##  $ Date        : Date, format: "2019-04-10" "2019-04-10" ...
##  $ County.Year : chr  "mohave2019" "gila2019" "santa cruz2019" "maricopa2019" ...
```

```
data.frame(variable = names(hospitalData),
           class = sapply(hospitalData, typeof),
           first.values = sapply(hospitalData, function(x) paste0(head(x),
                                                                  collapse = ",")),
           row.names = NULL) %>%
  kable(caption = "1535 obs. of 7 variables")
```

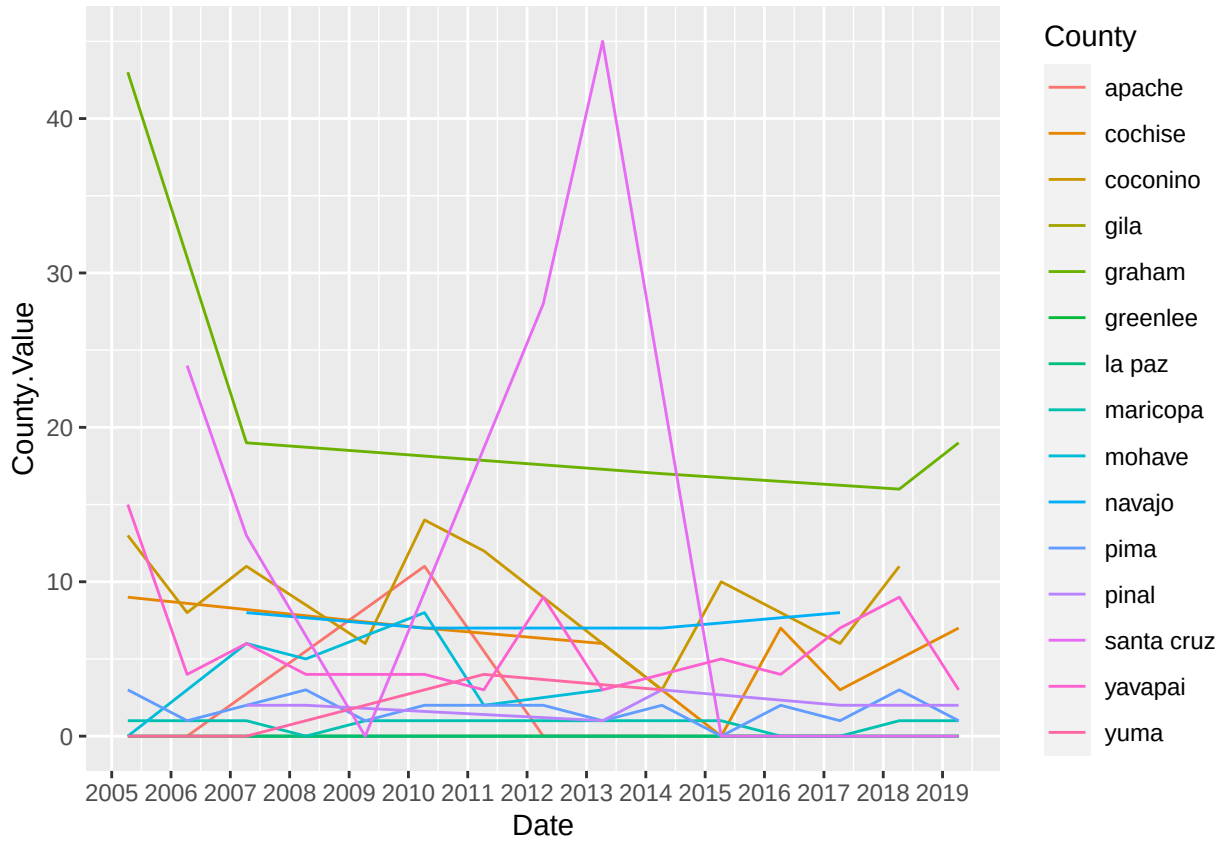
```
asthma <- hospitalData[hospitalData$Content.Area == "Asthma",]
ggplot(asthma, aes(x=Date, y=County.Value, color=County)) + geom_line() +
  scale_x_date(date_labels = "%Y", date_breaks = "1 year") +
  ggtitle("Asthma Vists in Arizona Counties") +
  ylab("Age-adjusted rate per 10,000 population")
```

Table 1: 1535 obs. of 7 variables

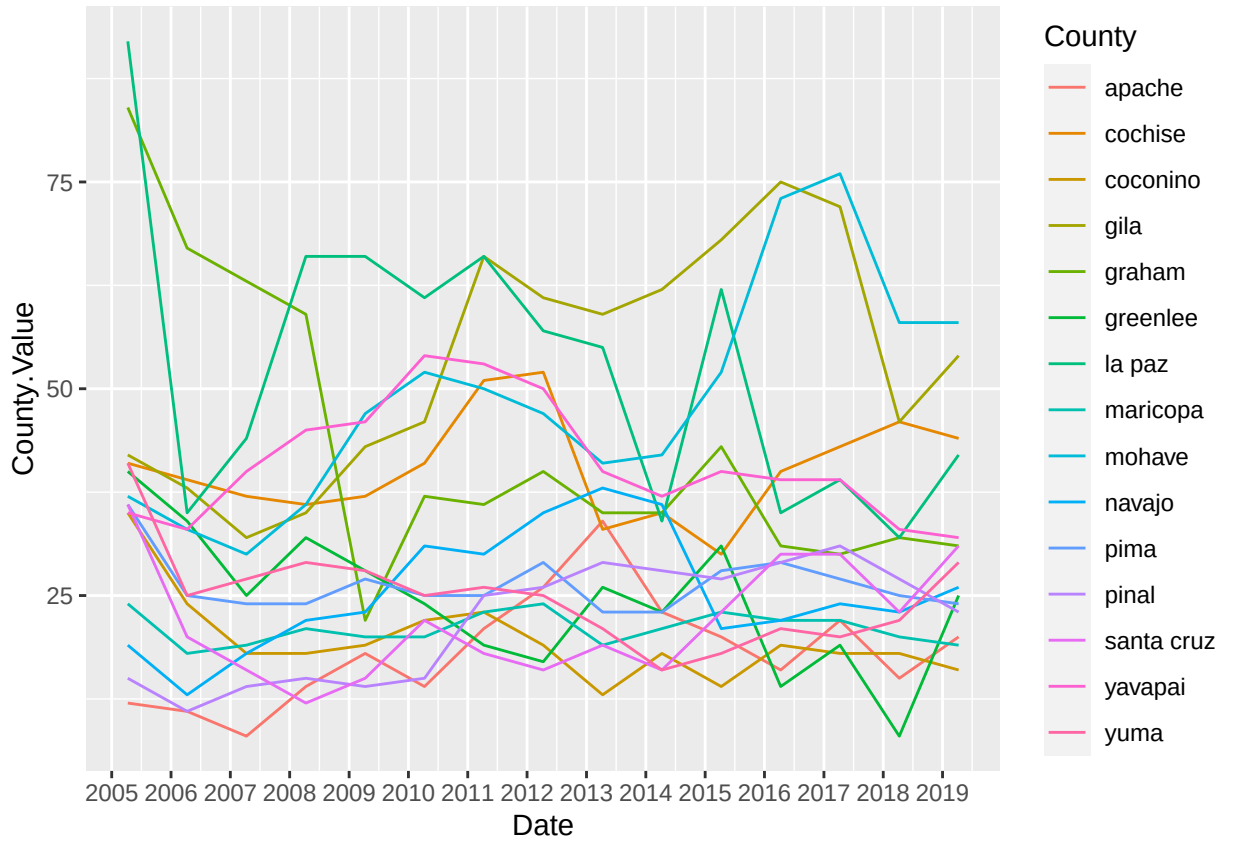
variable	class	first.values
County	integer	mohave,gila,santa cruz,maricopa,cochise,yuma
County.Value	integer	46,46,45,39,39,34
State.Rate	double	37.22,37.22,37.22,37.22,37.22,37.22
Year	character	2019,2019,2019,2019,2019,2019
Content.Area	integer	Asthma,Asthma,Asthma,Asthma,Asthma,Asthma
Date	double	2019-04-10,2019-04-10,2019-04-10,2019-04-10,2019-04-10,2019-04-10
County.Year	character	mohave2019,gila2019,santa cruz2019,maricopa2019,cochise2019,yuma2019



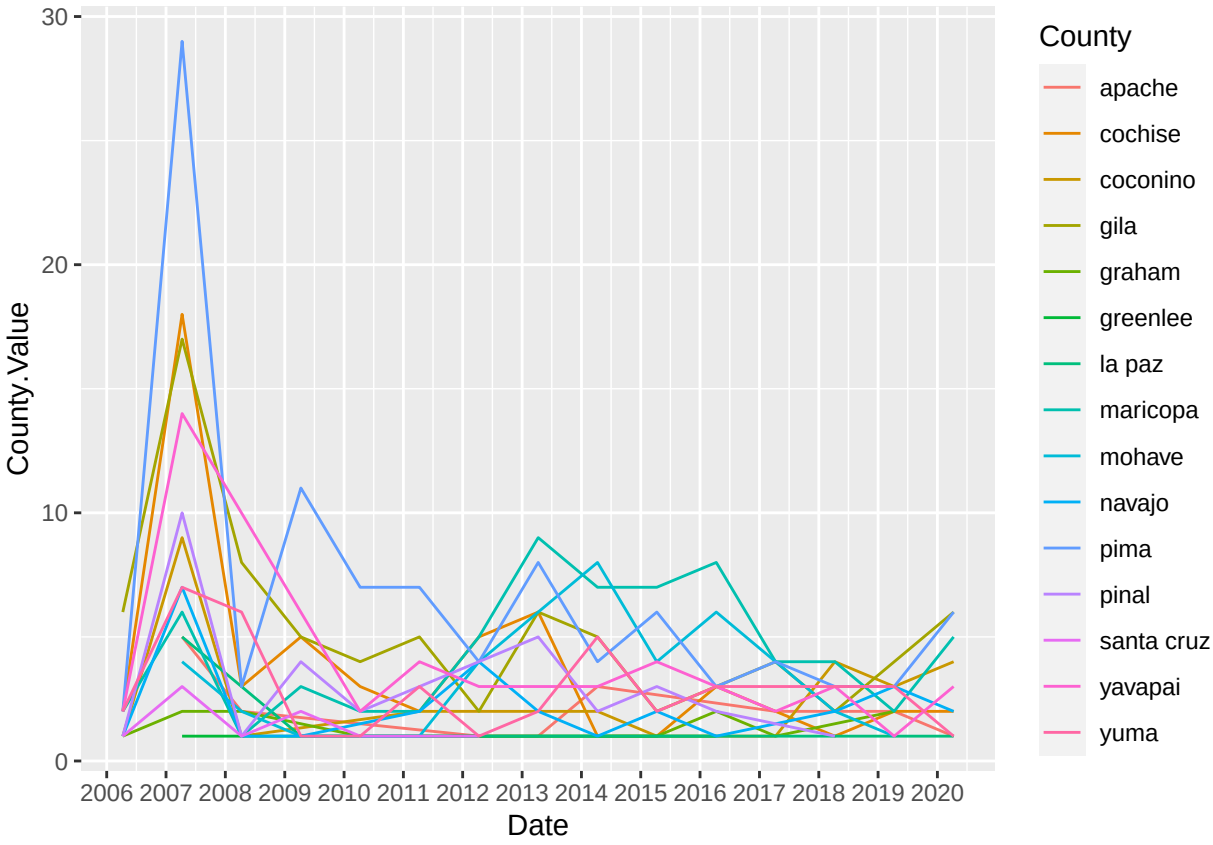

```
cmp <- hospitalData[hospitalData$Content.Area == "Carbon Monoxide Poisoning",]
ggplot(cmp, aes(x=Date, y=County.Value, color=County)) + geom_line() +
  scale_x_date(date_labels = "%Y", date_breaks = "1 year")
```




```
copd <- hospitalData[hospitalData$Content.Area == "Chronic Obstructive Pulmonary Disease (COPD)",]
ggplot(copd, aes(x=Date, y=County.Value, color=County)) + geom_line() +
  scale_x_date(date_labels = "%Y", date_breaks = "1 year")
```

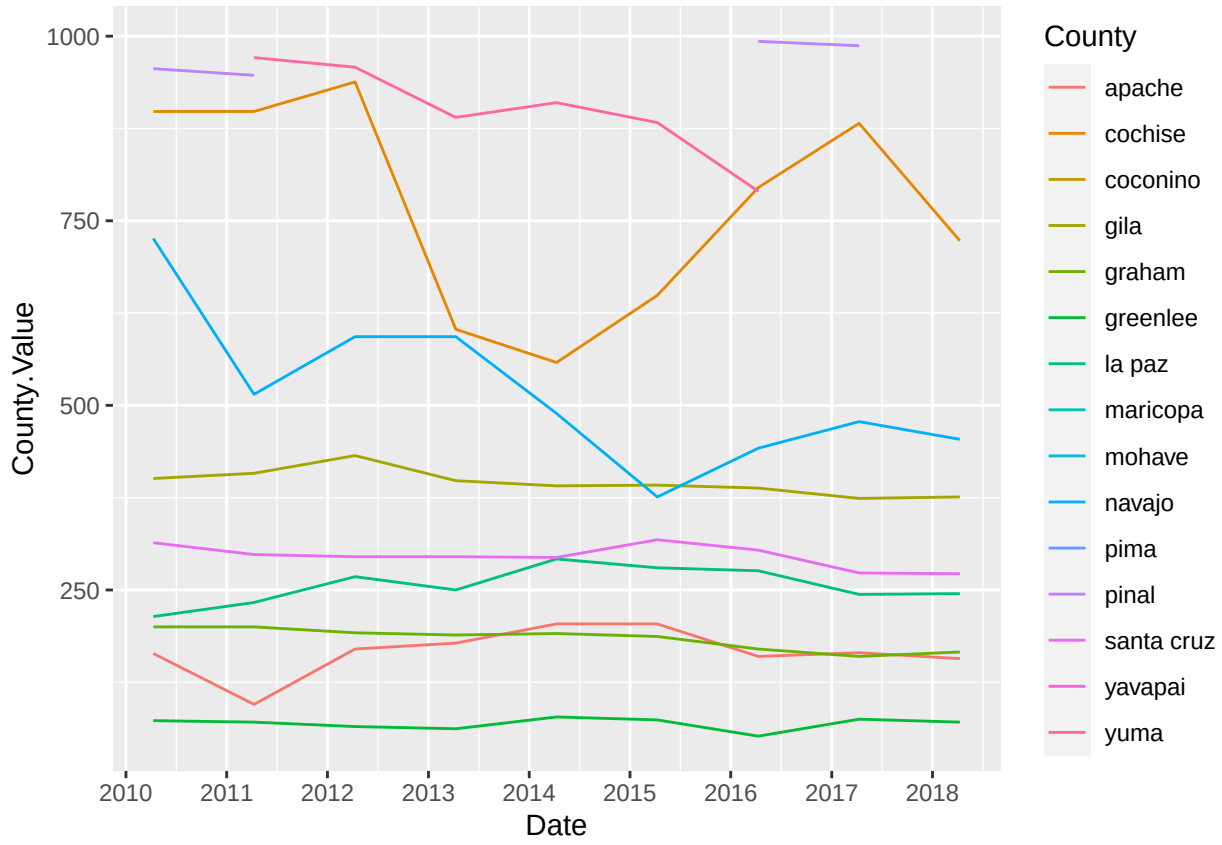



```
dwq <- hospitalData[hospitalData$Content.Area == "Drinking Water Quality",]
ggplot(dwq, aes(x=Date, y=County.Value, color=County)) + geom_line() +
  scale_x_date(date_labels = "%Y", date_breaks = "1 year")
```

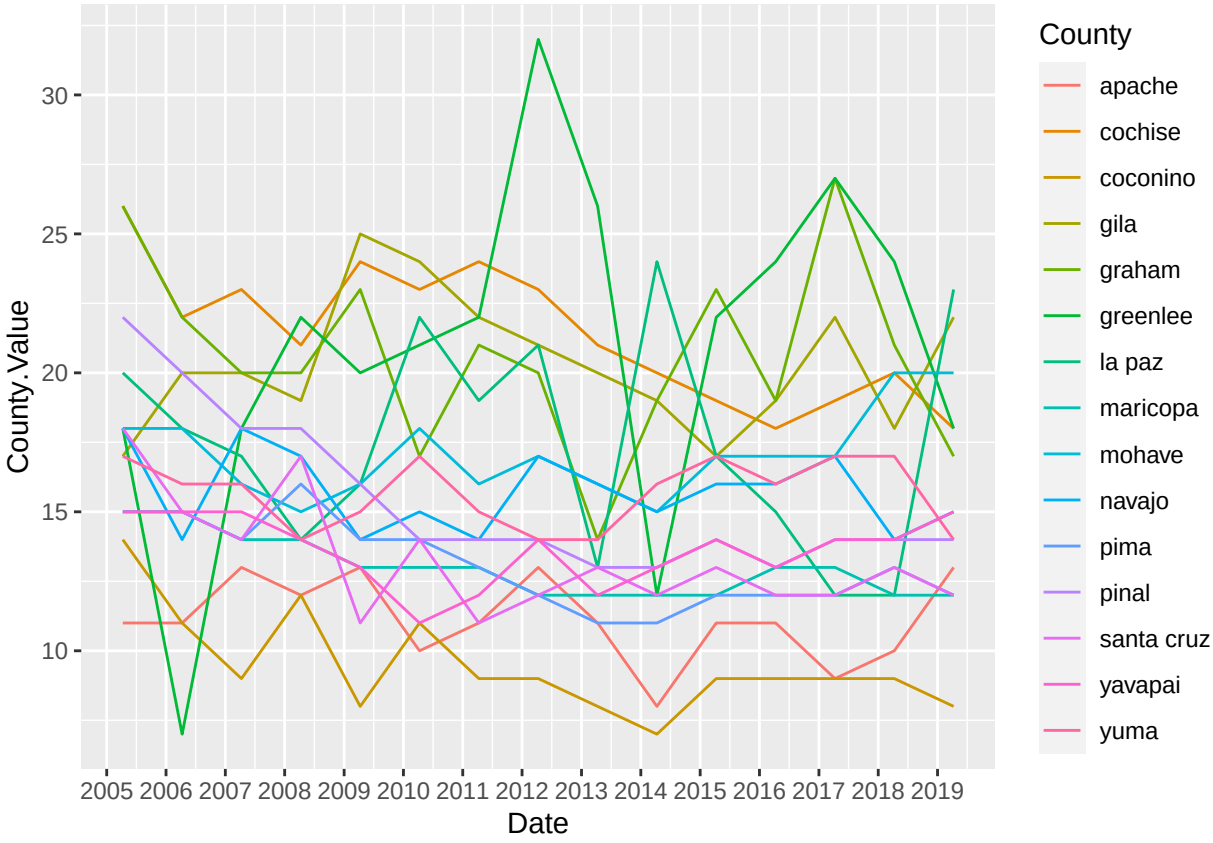



```
fs <- hospitalData[hospitalData$Content.Area == "Food Safety",]
ggplot(fs, aes(x=Date, y=County.Value, color=County)) + geom_line() +
  scale_x_date(date_labels = "%Y", date_breaks = "1 year")
```

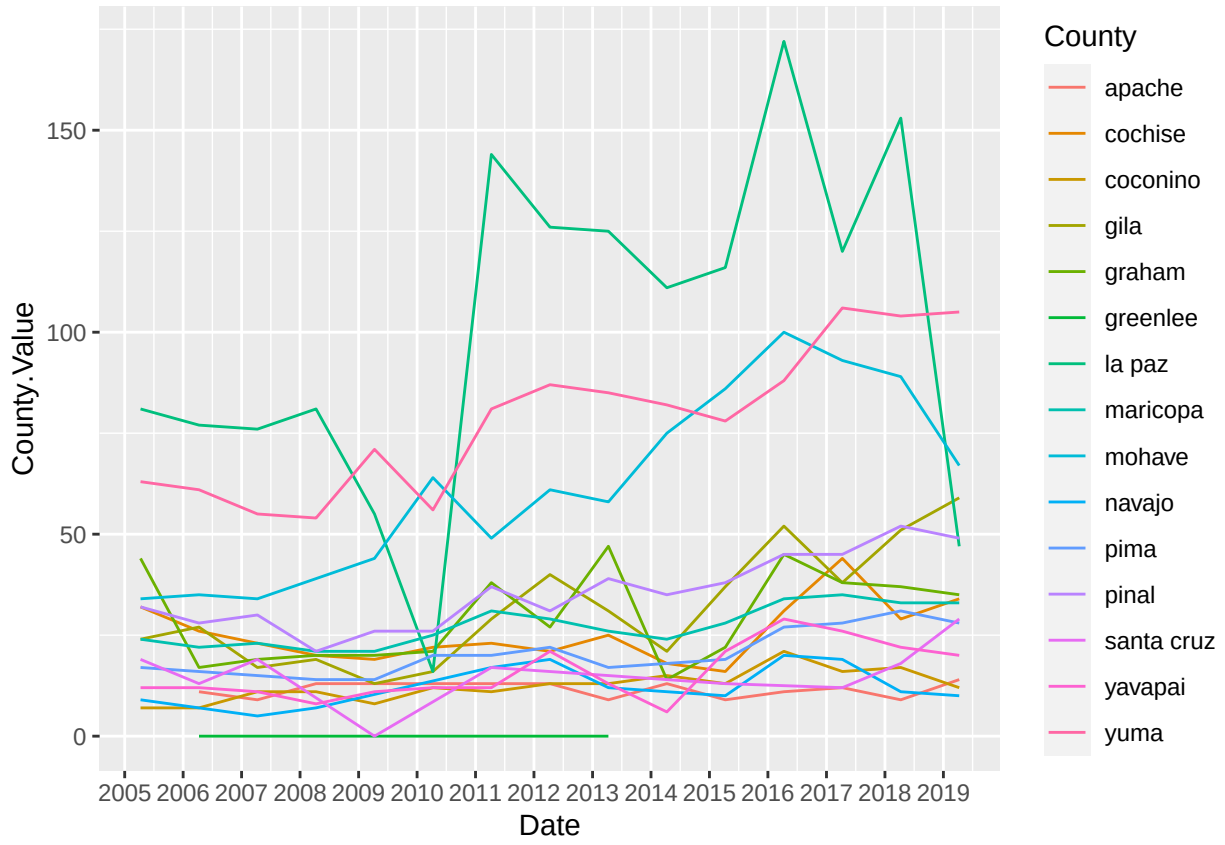
```
## Warning: Removed 47 row(s) containing missing values (geom_path).
```



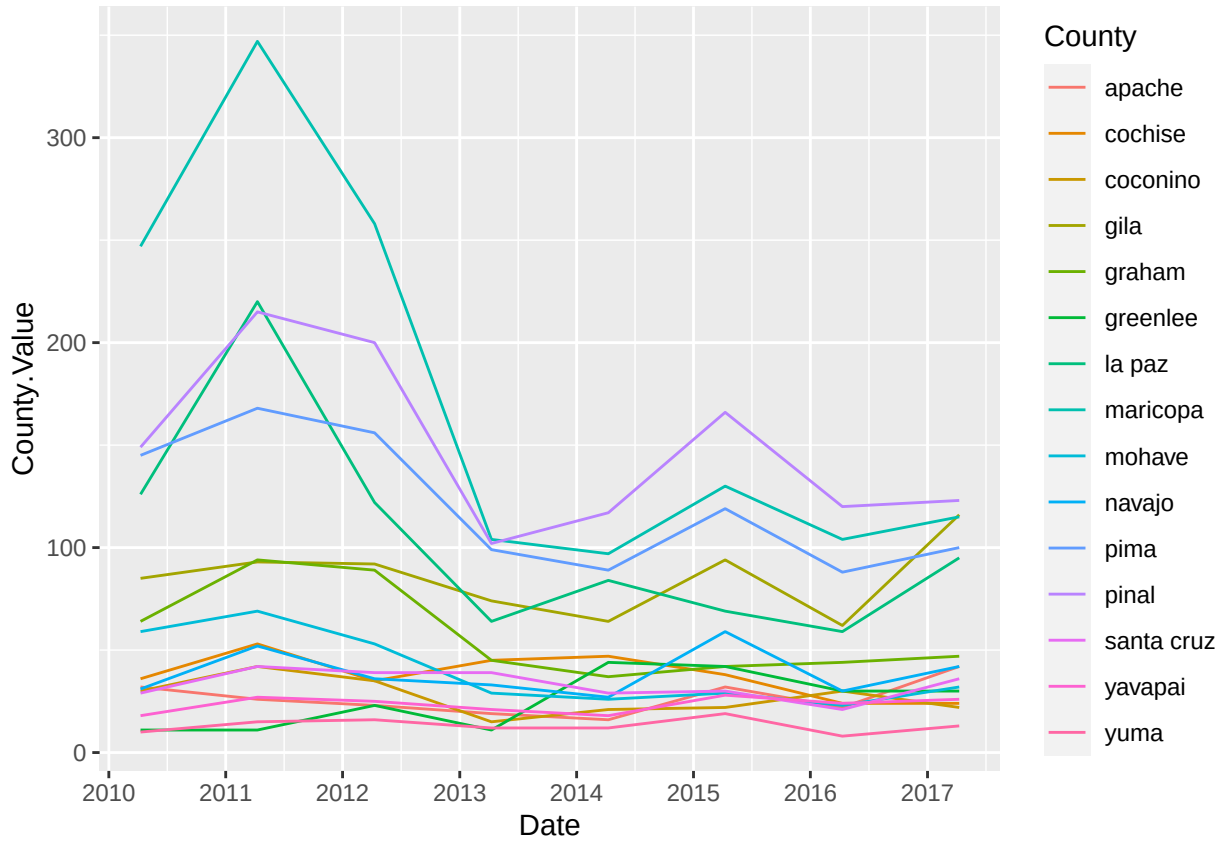

```
hd <- hospitalData[hospitalData$Content.Area == "Heart Disease",]
ggplot(hd, aes(x=Date, y=County.Value, color=County)) + geom_line() +
  scale_x_date(date_labels = "%Y", date_breaks = "1 year")
```



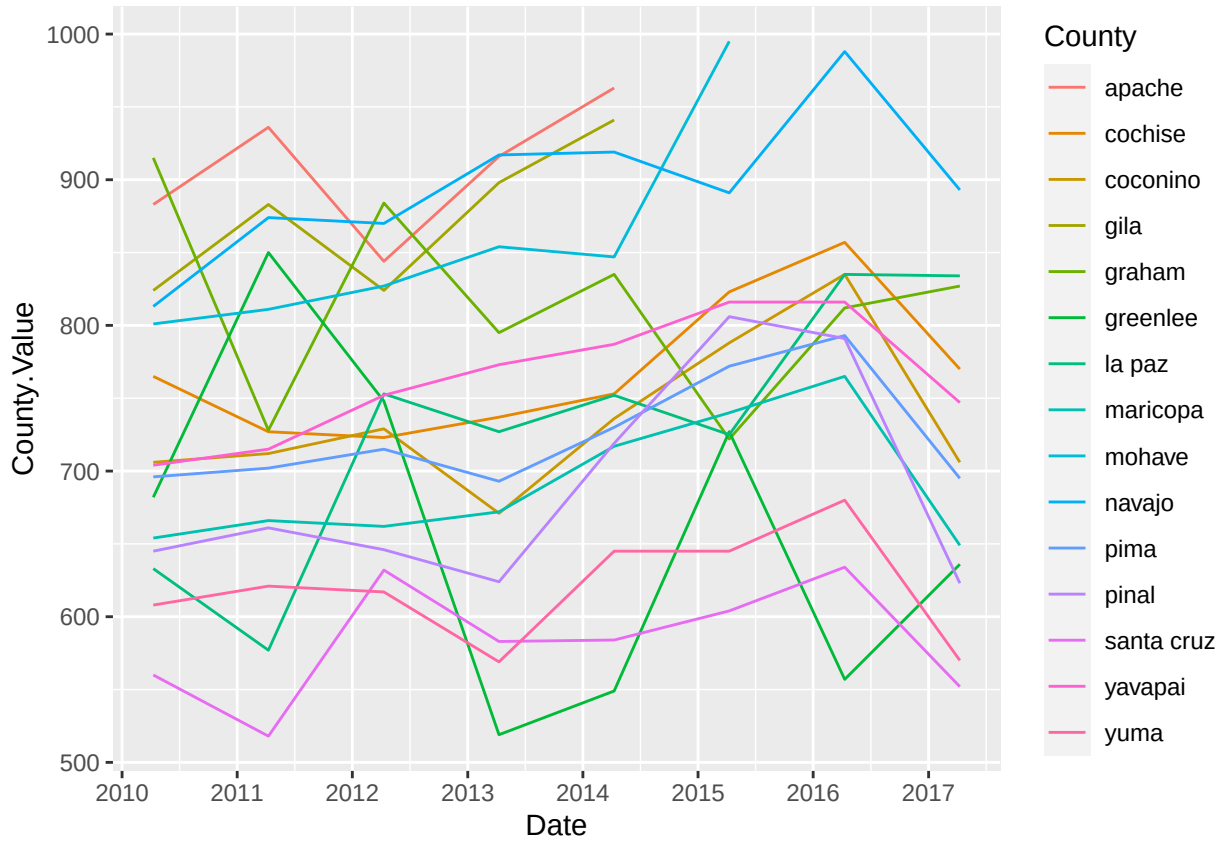

```
hsi <- hospitalData[hospitalData$Content.Area == "Heat Stress Illness",]
ggplot(hsi, aes(x=Date, y=County.Value, color=County)) + geom_line() +
  scale_x_date(date_labels = "%Y", date_breaks = "1 year")
```




```
id <- hospitalData[hospitalData$Content.Area == "Infectious Diseases",]
ggplot(id, aes(x=Date, y=County.Value, color=County)) + geom_line() +
  scale_x_date(date_labels = "%Y", date_breaks = "1 year")
```




```
m <- hospitalData[hospitalData$Content.Area == "Mortality",]
ggplot(m, aes(x=Date, y=County.Value, color=County)) + geom_line() +
  scale_x_date(date_labels = "%Y", date_breaks = "1 year")
```




```
asthma <- hospitalData[hospitalData$Content.Area == "Asthma" &
                         hospitalData$County == "COCONINO",]
asthma <- asthma %>% arrange(Date) %>% pull(County.Value)
plot(asthma)
acf(asthma)
pacf(asthma)
obj <- arima(asthma, order=c(1,0,0))
acf(obj$residuals)
predict(obj, n.ahead = 1)
```

```
hospitalNew <- hospitalData[!(hospitalData$County == "gila" |
                         hospitalData$County == "graham" |
                         hospitalData$County == "greenlee" |
                         hospitalData$County == "la paz" |
                           hospitalData$County == "santa cruz"),]
hospitalNew <- hospitalNew[!(hospitalNew$Date == "2005-04-10" |
                                hospitalNew$Date == "2006-04-10" |
                                hospitalNew$Date == "2007-04-10" |
                                hospitalNew$Date == "2008-04-10" |
                                hospitalNew$Date == "2009-04-10" |
                              hospitalNew$Date == "2020-04-10"),]
asthma <- hospitalNew[hospitalData$Content.Area == "Asthma",]
```

```
hospital.wider <- hospitalNew %>% select(-State.Rate)
hospital.wider <- pivot_wider(hospital.wider, names_from = Content.Area,
```

```
                                values_from = County.Value)
```

```
all.data <- full_join(hospital.wider, censusData, by="County.Year") %>%
  relocate(County.x, Year.x, County.y, Year.y, Date.x, Date.y)
```

```
asthma.wider <- hospital.wider %>% select(c(County.Year, Asthma))
asthma.data <- merge(asthma.wider, censusData)
cmp.wider <- hospital.wider %>% select(c(County.Year, `Carbon Monoxide Poisoning`))
cmp.data <- merge(cmp.wider, censusData)
copd.wider <- hospital.wider %>%
  select(c(County.Year, `Chronic Obstructive Pulmonary Disease (COPD)`))
copd.data <- merge(copd.wider, censusData)
dwq.wider <- hospital.wider %>%
  select(c(County.Year, `Drinking Water Quality`))
dwq.data <- merge(dwq.wider, censusData)
fs.wider <- hospital.wider %>%
  select(c(County.Year, `Food Safety`))
fs.data <- merge(fs.wider, censusData)
hd.wider <- hospital.wider %>%
  select(c(County.Year, `Heart Disease`))
hd.data <- merge(hd.wider, censusData)
hsi.wider <- hospital.wider %>%
  select(c(County.Year, `Heat Stress Illness`))
hsi.data <- merge(hsi.wider, censusData)
id.wider <- hospital.wider %>%
  select(c(County.Year, `Infectious Diseases`))
id.data <- merge(id.wider, censusData)
m.wider <- hospital.wider %>%
  select(c(County.Year, `Mortality`))
m.data <- merge(m.wider, censusData)
```

```
# colnames(cor.data) <- c('x1', 'x2', 'x3', 'x4', 'x5', 'x6', 'x7', 'x8', 'x9',
#                         'x10', 'x11', 'x12', 'x13', 'x14', 'x15', 'x16', 'x17',
#                         'x18', 'x19', 'x20', 'x21', 'x22', 'x23', 'x24',
#                         'x25', 'x26', 'x27', 'x28', 'x29', 'x30', 'x31', 'x32',
#                         'x33', 'x34', 'x35',
#                         'x36', 'x37', 'x38', 'x39', 'x40', 'x41', 'x42', 'x43',
#                         'x44', 'x45', 'x46',
#                         'x47', 'x48', 'x49', 'x50', 'x51', 'x52', 'x53', 'x54',
#                         'x55', 'x56', 'x57',
#                         'x58', 'x59', 'x60', 'x61', 'x62', 'x63', 'x64', 'x65',
#                         'x66', 'x67', 'x68',
#                         'x69', 'x70', 'x71', 'x72', 'x73', 'x74', 'x75', 'x76',
#                         'x77', 'x78', 'x79',
#                         'x80', 'x81', 'x82', 'x83', 'x84', 'x85', 'x86', 'x87',
#                         'x88', 'x89', 'x90',
#                         'x91', 'x92')
```

```
cor.data <- all.data %>% select(-c(County.Year, Year, County))
cor.asthma <- asthma.data %>% select(-c(County.Year, Year, County))
cor.cmp <- cmp.data %>% select(-c(County.Year, Year, County))
cor.copd <- copd.data %>% select(-c(County.Year, Year, County))
```

```
cor.dwq <- dwq.data %>% select(-c(County.Year, Year, County))
cor.fs <- fs.data %>% select(-c(County.Year, Year, County))
cor.hd <- hd.data %>% select(-c(County.Year, Year, County))
cor.hsi <- hsi.data %>% select(-c(County.Year, Year, County))
cor.id <- id.data %>% select(-c(County.Year, Year, County))
cor.m <- m.data %>% select(-c(County.Year, Year, County))
corrplot::corrplot(cor(cor.asthma, use = 'pairwise.complete.obs'), tl.cex = 0.8)
corrplot::corrplot(cor(cor.cmp, use = 'pairwise.complete.obs'), tl.cex = 0.8)
corrplot::corrplot(cor(cor.copd, use = 'pairwise.complete.obs'), tl.cex = 0.8)
corrplot::corrplot(cor(cor.dwq, use = 'pairwise.complete.obs'), tl.cex = 0.8)
corrplot::corrplot(cor(cor.fs, use = 'pairwise.complete.obs'), tl.cex = 0.8)
corrplot::corrplot(cor(cor.hd, use = 'pairwise.complete.obs'), tl.cex = 0.8)
corrplot::corrplot(cor(cor.hsi, use = 'pairwise.complete.obs'), tl.cex = 0.8)
corrplot::corrplot(cor(cor.id, use = 'pairwise.complete.obs'), tl.cex = 0.8)
corrplot::corrplot(cor(cor.m, use = 'pairwise.complete.obs'), tl.cex = 0.8)
corrplot::corrplot(cor(cor.data, use = 'pairwise.complete.obs'), tl.cex = 0.8)
cor.v <- cor(cor.data)
pos.cor <- cor.v[cor.v > 0.5]
```

```
test <- cor.test(hospital.wider$Asthma, censusData$Total.Population)
test$p.value
test$estimate
cor.test(hospital.wider[,4], censusData$Total.Population)
```

```
all.data$Asthma <- as.numeric(all.data$Asthma)
all.data <- all.data %>% select(-c(County.x, Year.x, Date.x, `Food Safety`))
colnames(all.data)[1] <- "County"
colnames(all.data)[2] <- "Year"
colnames(all.data)[3] <- "Date"
all.data$Year <- as.factor(all.data$Year)
names(all.data) <- gsub(" ", ".", names(all.data))
colnames(all.data)[7] <- "COPD"
```

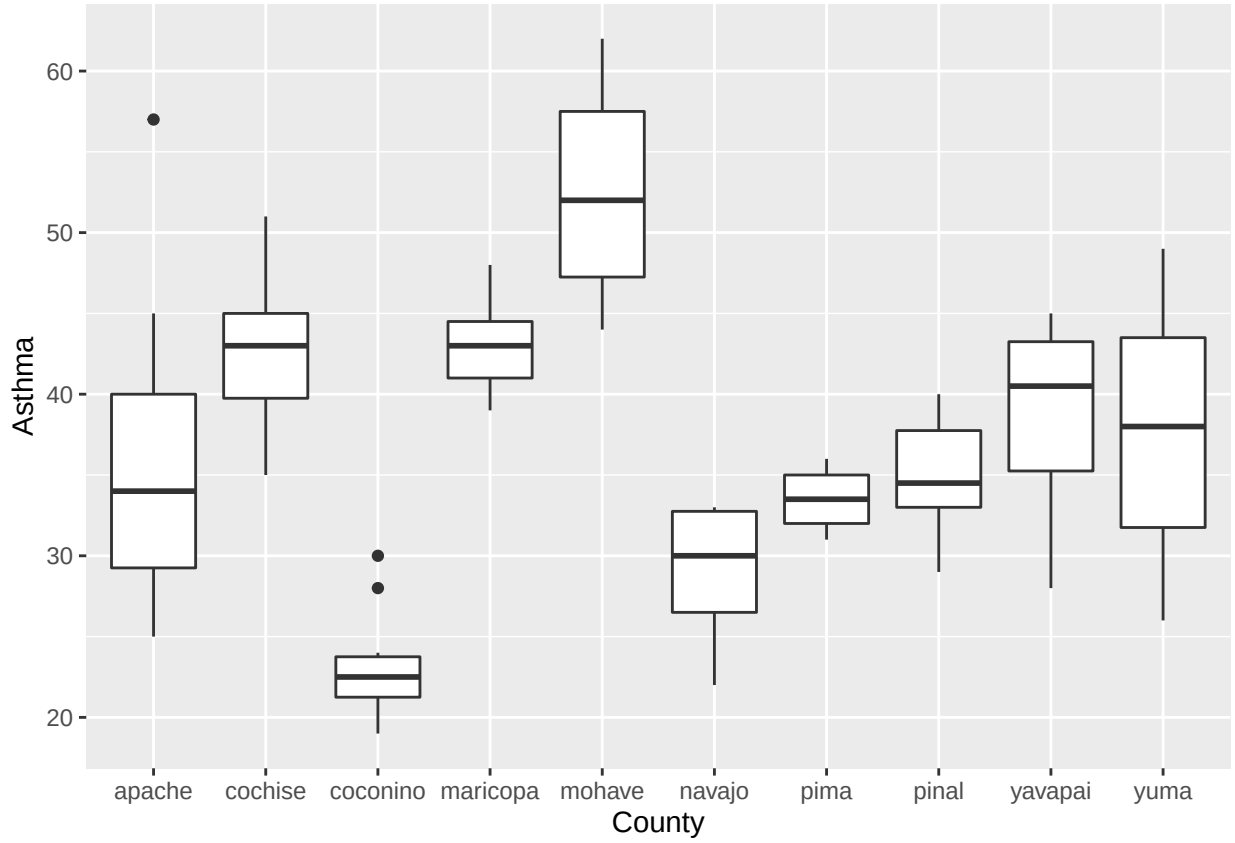
```
correlation <- data.frame(Content.Area = character(),
                          Demographic = character(),
                          p.value = integer(),
                          corr = integer())
for(cont in 5:12){
  for(demo in 13:95){
    #print(paste(cont, demo))
    hold <- cor.test(pull(all.data[,cont]), pull(all.data[,demo]))
    corr.df <- data.frame(Content.Area = c(colnames(all.data)[cont]),
                          Demographic = c(colnames(all.data)[demo]),
                          p.value = c(hold$p.value),
                          corr = c(hold$estimate))
    if(is.na(correlation[1,1])){
      correlation <- corr.df
    }
    else {
      correlation <- rbind(correlation, corr.df)
    }

  }
```

```
}
```

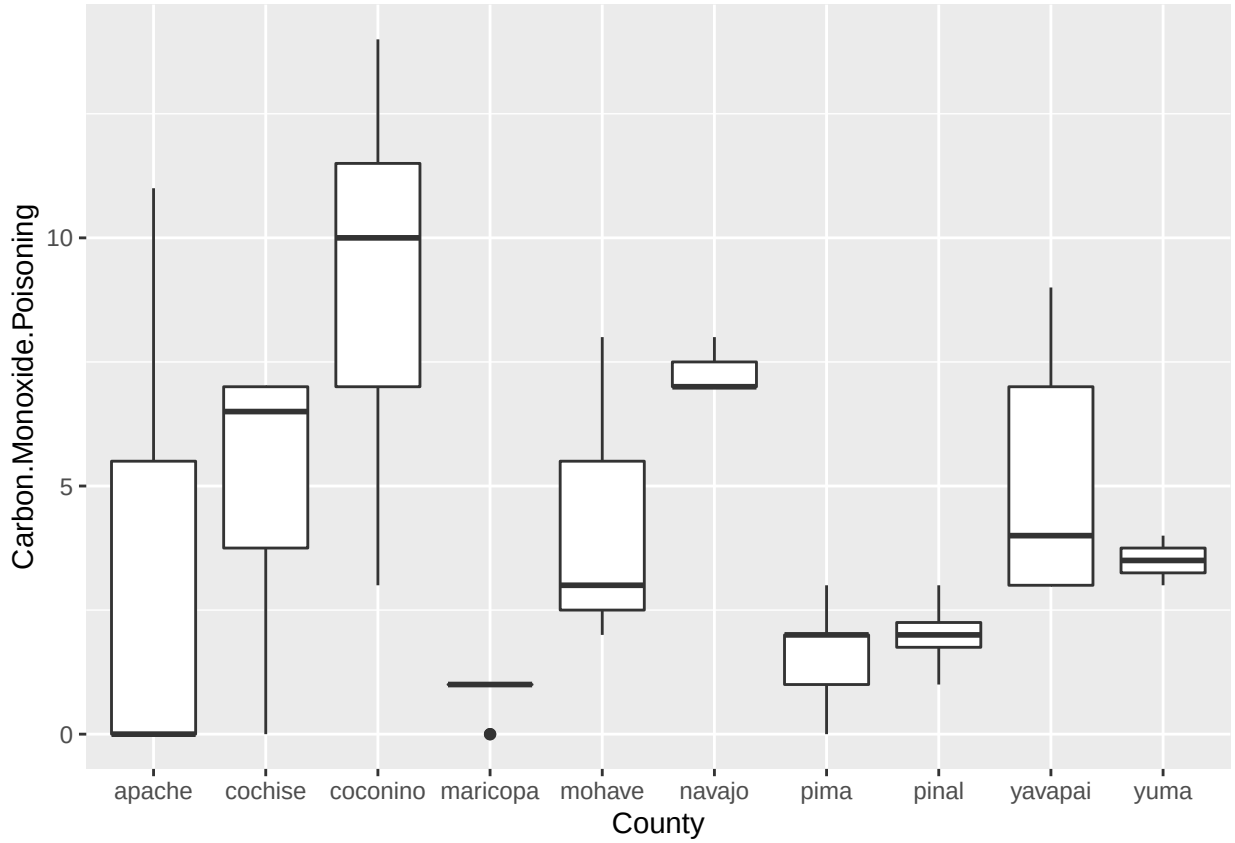
```
correlation$adj.p.value <- p.adjust(correlation$p.value, method="bonferroni")
```

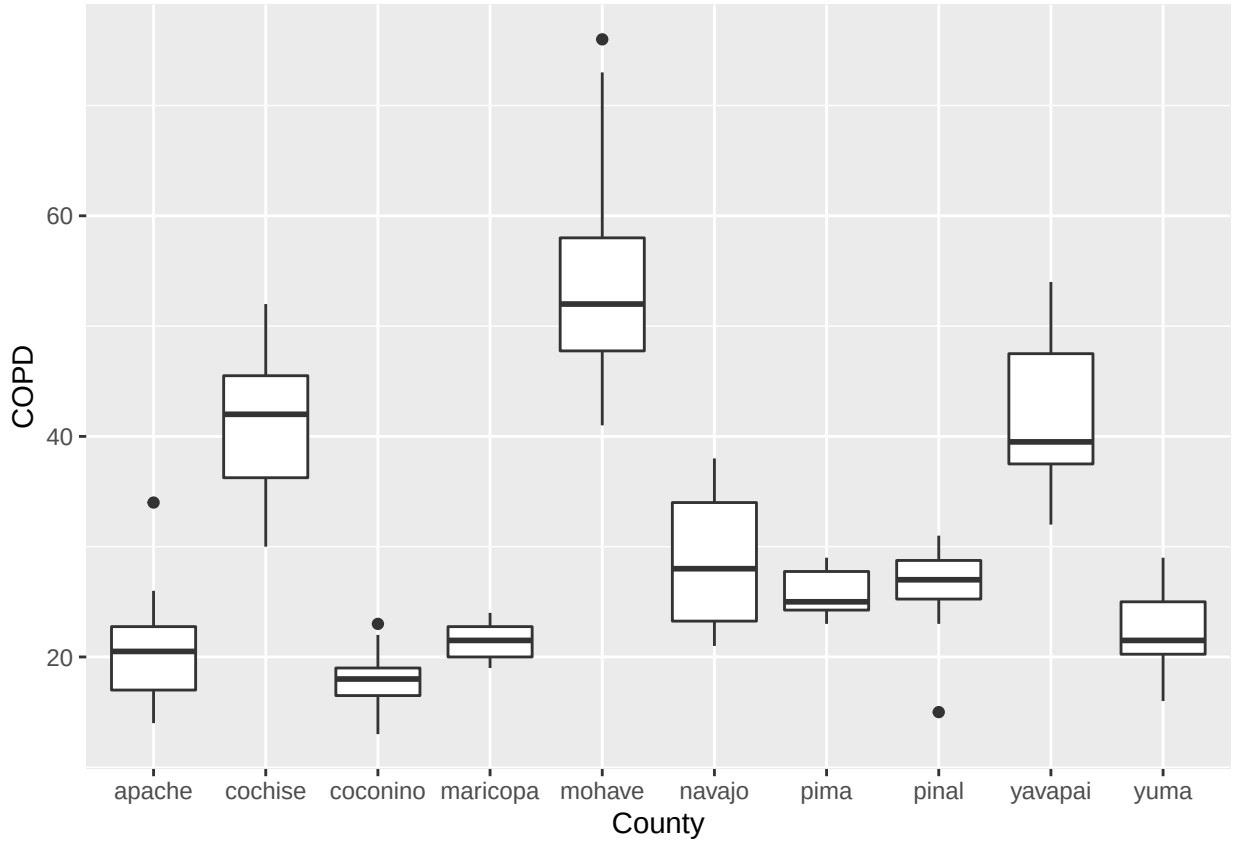
```
adj.cor <- correlation %>% filter(adj.p.value < 0.05) %>% arrange(corr)
```

```
#ggplot(all.data, aes(x=County, y=Asthma)) + geom_boxplot()
for (index in 5:12) {
  print(ggplot(all.data, aes_string(x="County", y=colnames(all.data)[index])) +
    geom_boxplot())
}
```

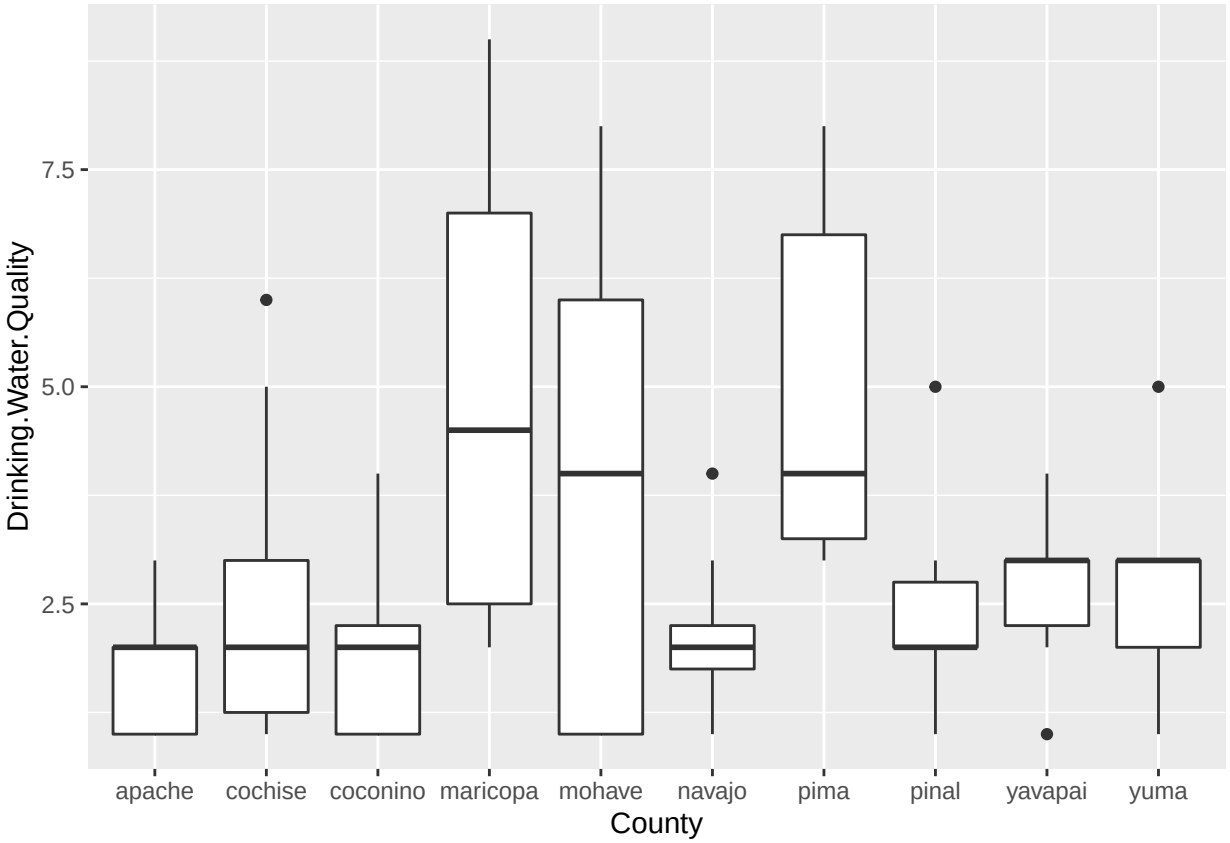


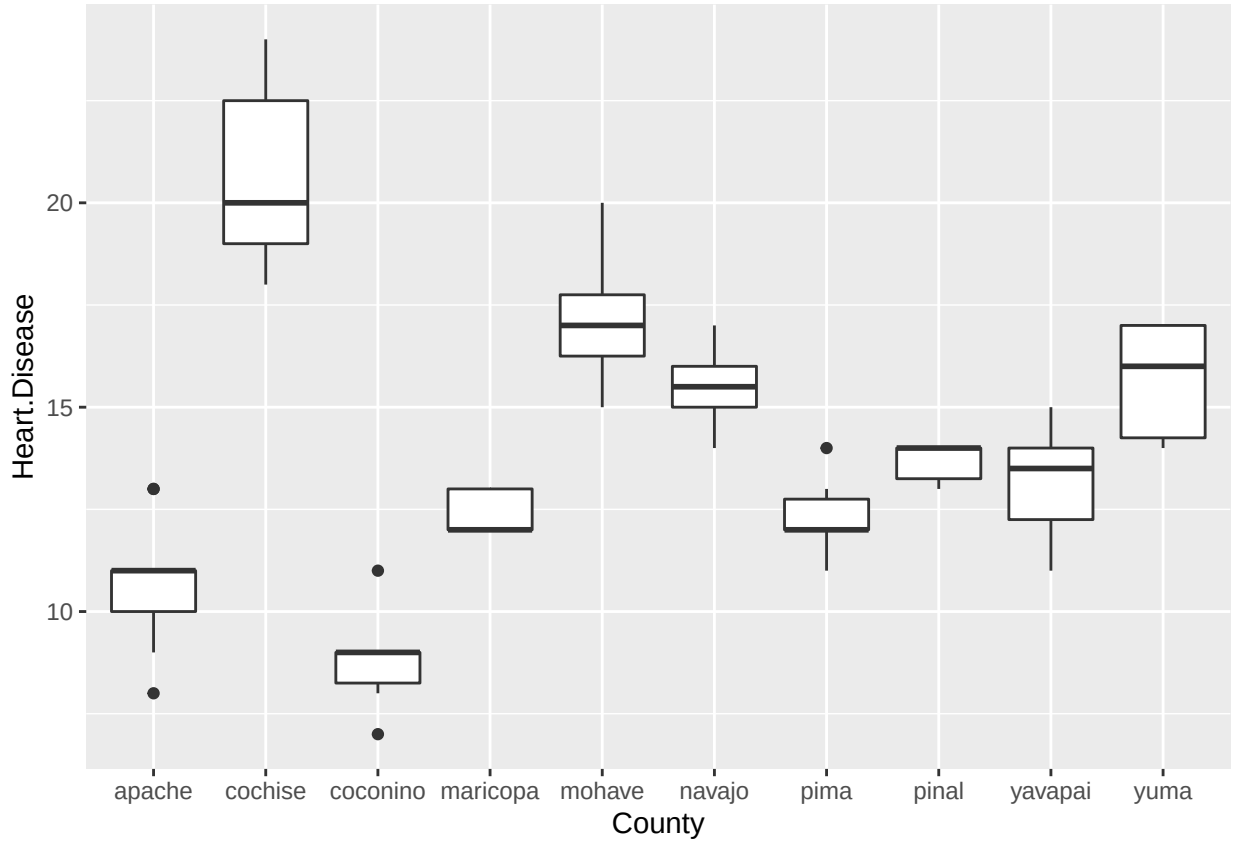

```
## Warning: Removed 43 rows containing non-finite values (stat_boxplot).
```


Carbon.Monoxide.Poisoning
10
5
0
apache
cochise
coconino
maricopa
mohave
navajo
pima
pinal
yavapai
yuma
County

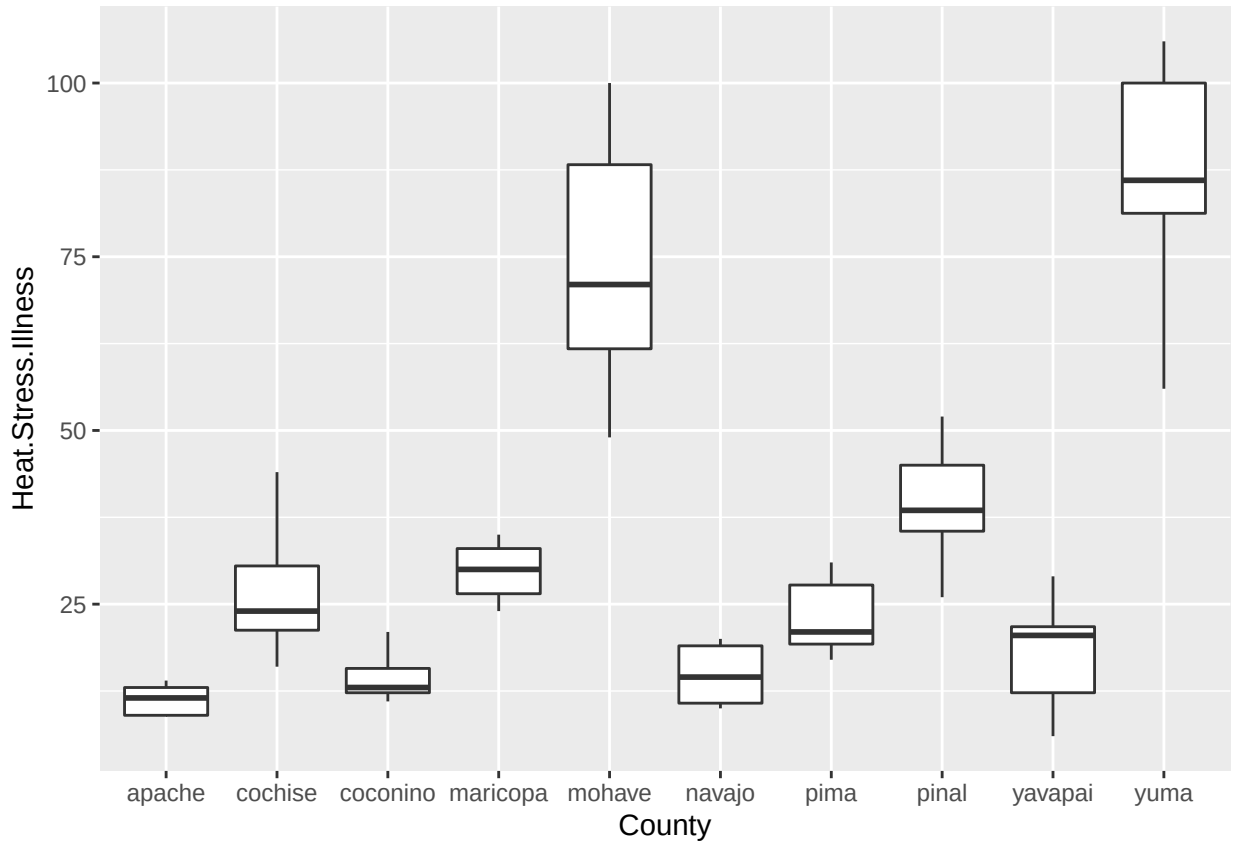



```
## Warning: Removed 14 rows containing non-finite values (stat_boxplot).
```

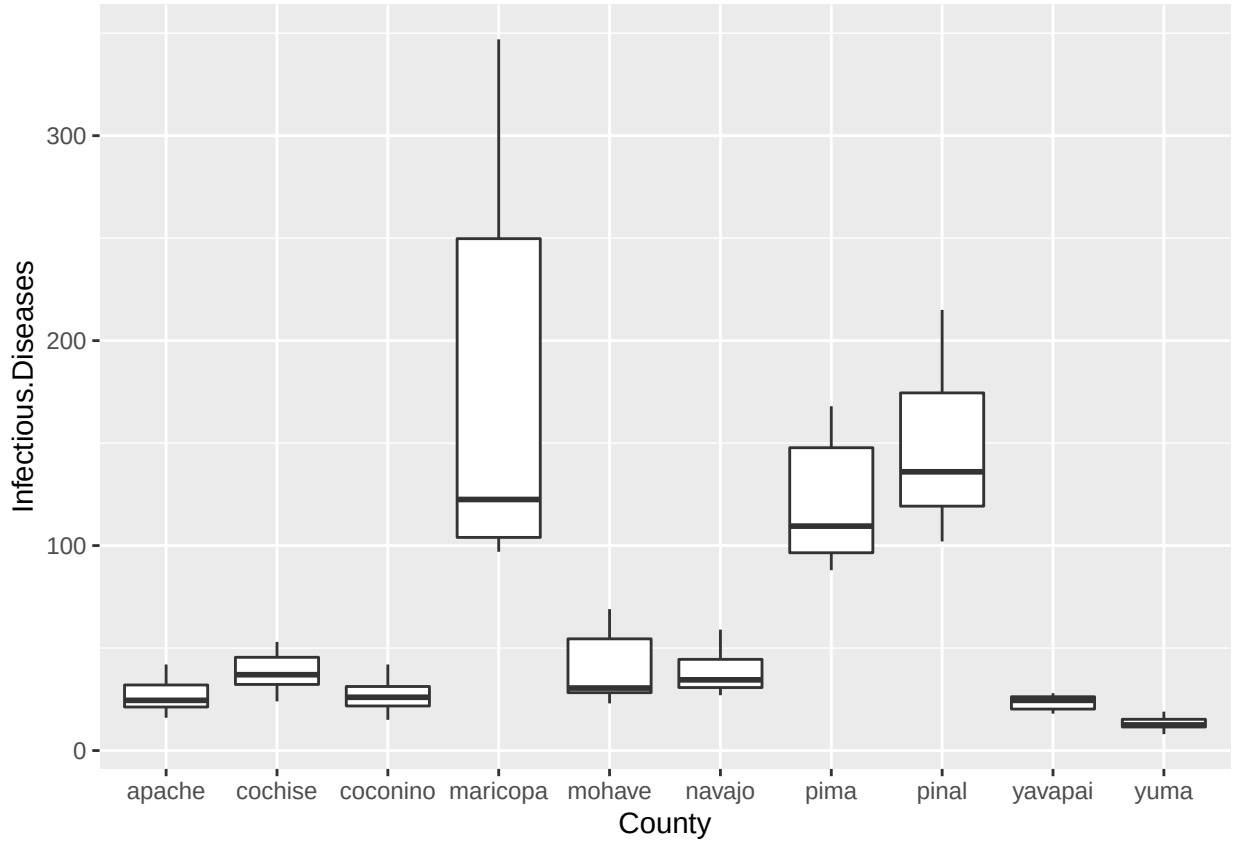

Drinking.Water.Quality
7.5
5.0
2.5
apache
cochise
coconino
maricopa
mohave
navajo
pima
pinal
yavapai
yuma
County



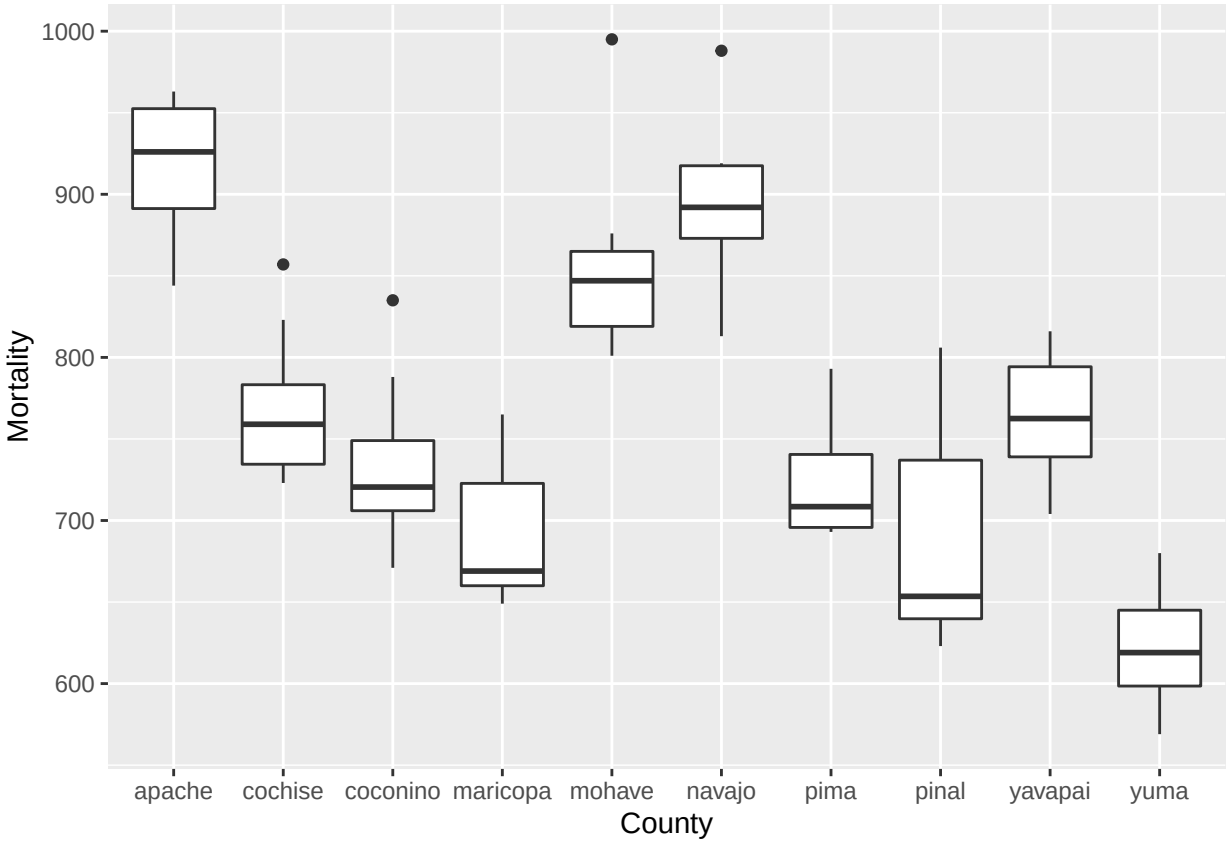

```
## Warning: Removed 4 rows containing non-finite values (stat_boxplot).
```




```
## Warning: Removed 20 rows containing non-finite values (stat_boxplot).
```




```
## Warning: Removed 23 rows containing non-finite values (stat_boxplot).
```

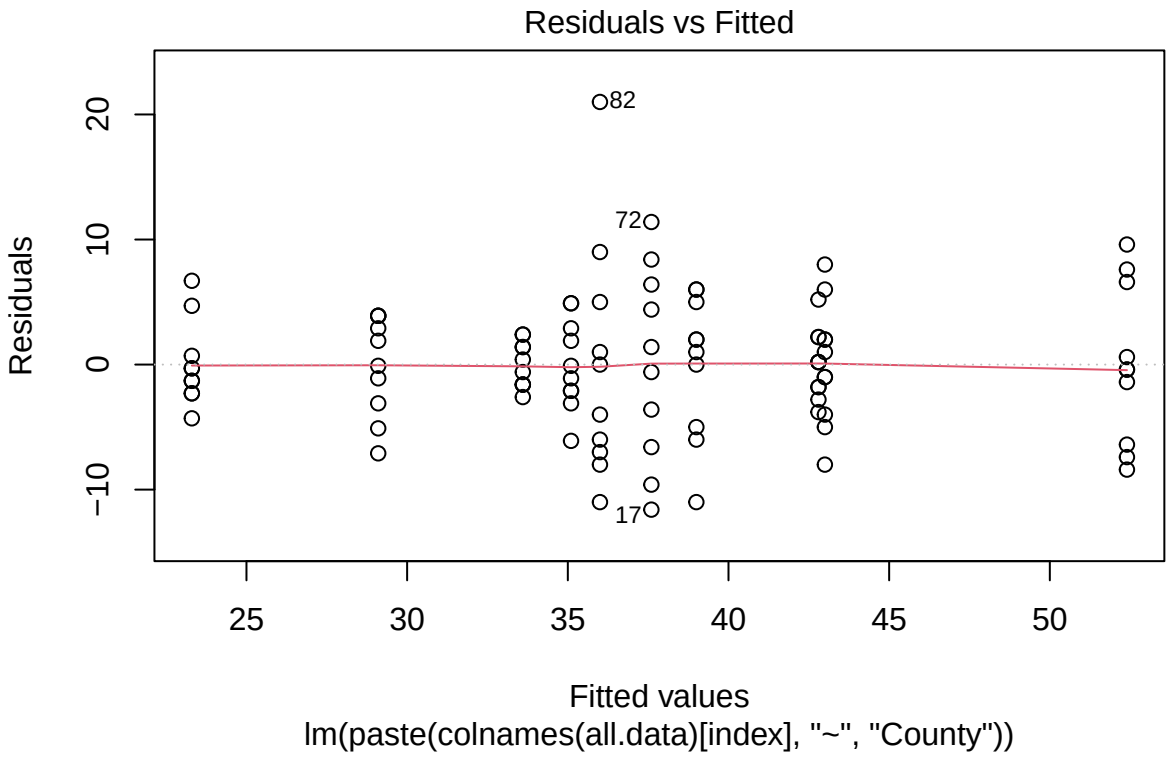


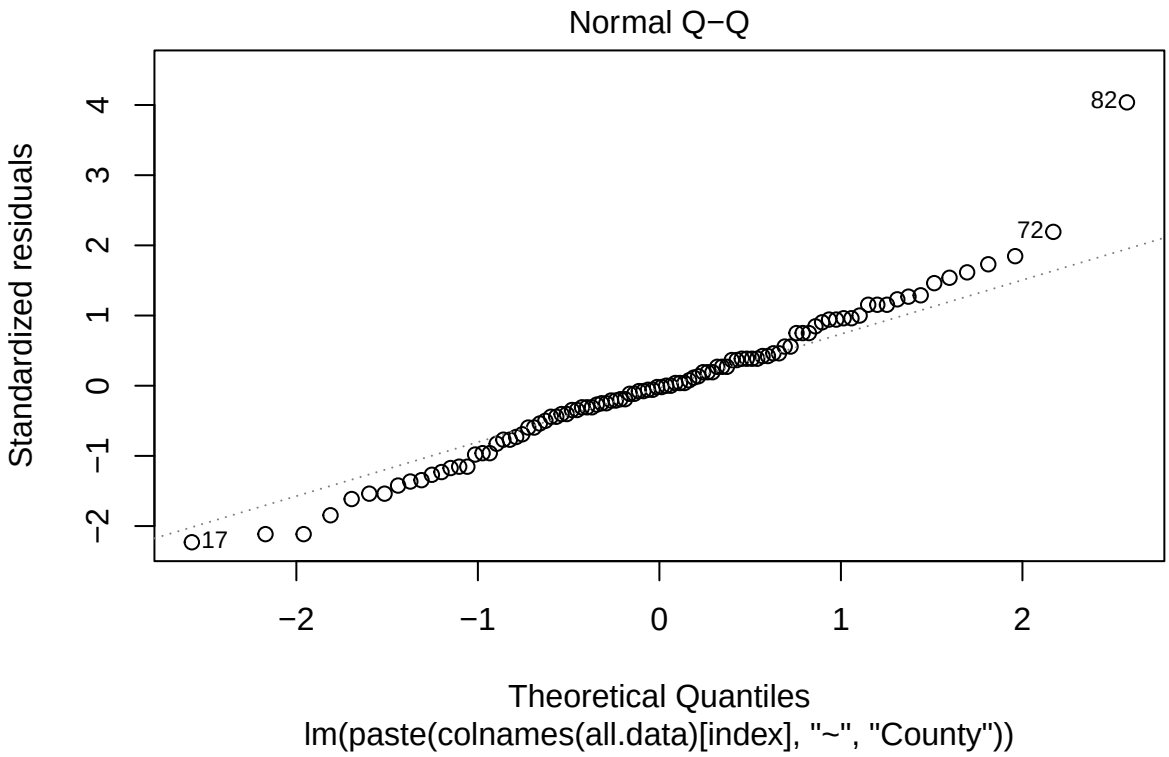

```
obj <- lm(Asthma ~ County, data=all.data)
summary(obj)
plot(obj)
anova(obj)
emmeans(obj, pairwise ~ County)
```

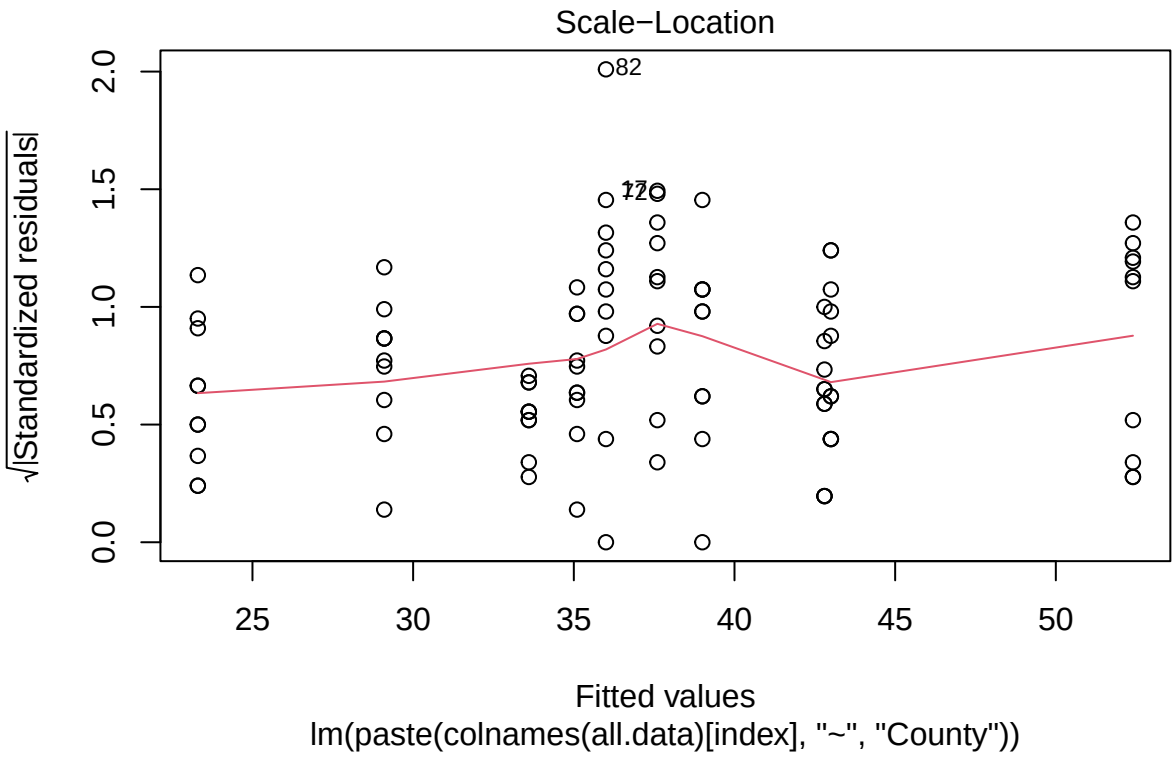
```
content.area.p <- data.frame(
  County = factor(),
  p.value = integer()
)
for (index in 5:12) {
  obj <- lm(paste(colnames(all.data)[index], "~", 'County'), data=all.data)
  sum.obj <- summary(obj)
  f <- sum.obj$fstatistic
  p <- pf(f[1], f[2], f[3], lower.tail=F)
  p.df <- data.frame(County = colnames(all.data)[index],
                     p.value = p)
  if(index == 5)
  {
    content.area.p <- p.df
  }
  else
  {
    content.area.p <- rbind(content.area.p, p.df)
  }
```

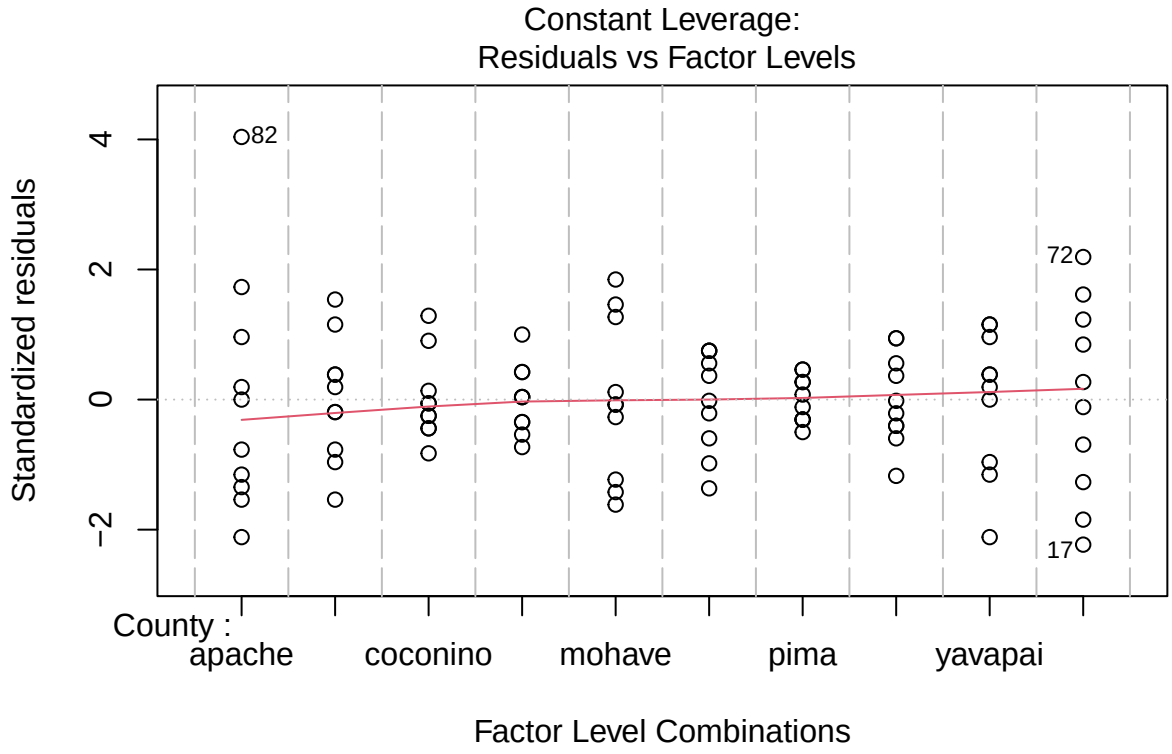
```
  print(sum.obj)
  print(anova(obj))
  print(plot(obj))
  #print(emmeans(obj, pairwise ~ County))
}
```

```
##
## Call:
## lm(formula = paste(colnames(all.data)[index], "~", "County"),
##     data = all.data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -11.600  -2.875  -0.100   2.525  21.000
##
## Coefficients:
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept)      36.000      1.734  20.767  < 2e-16 ***
## Countycochise     7.000      2.452   2.855  0.00534 **
## Countycoconino  -12.700      2.452  -5.180 1.34e-06 ***
## Countymaricopa    6.800      2.452   2.774  0.00674 **
## Countymohave     16.400      2.452   6.689 1.85e-09 ***
## Countynavajo     -6.900      2.452  -2.814  0.00600 **
## Countypima       -2.400      2.452  -0.979  0.33023
## Countypinal      -0.900      2.452  -0.367  0.71440
## Countyyavapai     3.000      2.452   1.224  0.22427
## Countyyuma        1.600      2.452   0.653  0.51566
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 5.482 on 90 degrees of freedom
## Multiple R-squared:  0.6809, Adjusted R-squared:  0.649
## F-statistic: 21.34 on 9 and 90 DF,  p-value: < 2.2e-16
##
## Analysis of Variance Table
##
## Response: Asthma
##           Df Sum Sq Mean Sq F value    Pr(>F)
## County     9 5770.7  641.19  21.336 < 2.2e-16 ***
## Residuals 90 2704.7   30.05
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```


Residuals vs Fitted
Residuals
Fitted values
lm(paste(colnames(all.data)[index], "~", "County"))
82
72
17
20
10
0
−10
25
30
35
40
45
50

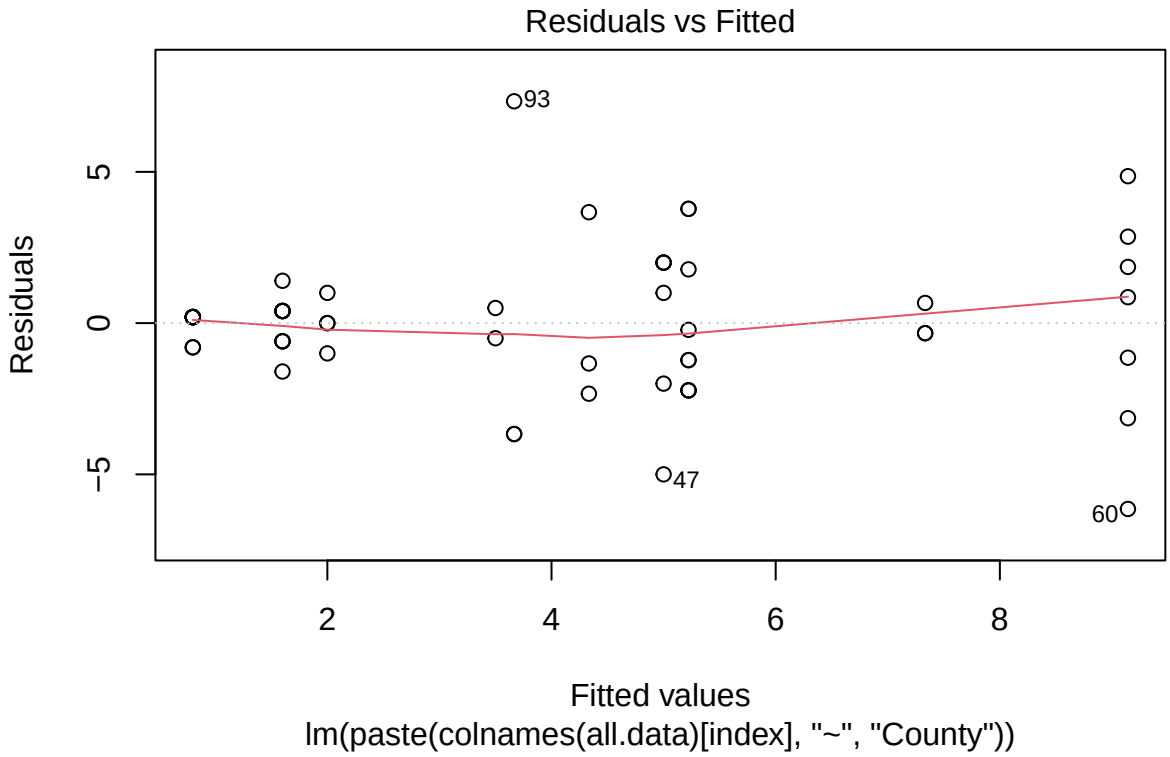

Normal Q−Q
Standardized residuals
Theoretical Quantiles
lm(paste(colnames(all.data)[index], "~", "County"))
82
72
17
−2
−1
0
1
2
3
4

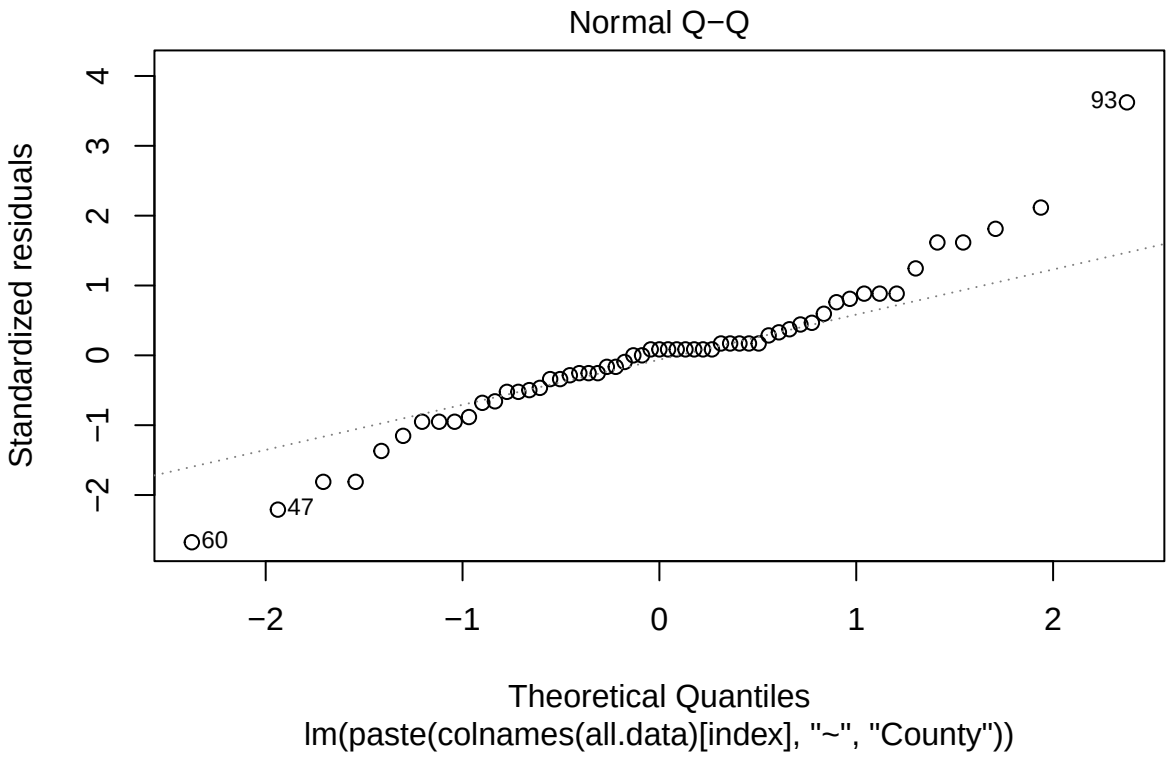

Scale−Location
√|Standardized residuals|
0.0
0.5
1.0
1.5
2.0
82
120
25
30
35
40
45
50
Fitted values
lm(paste(colnames(all.data)[index], "~", "County"))

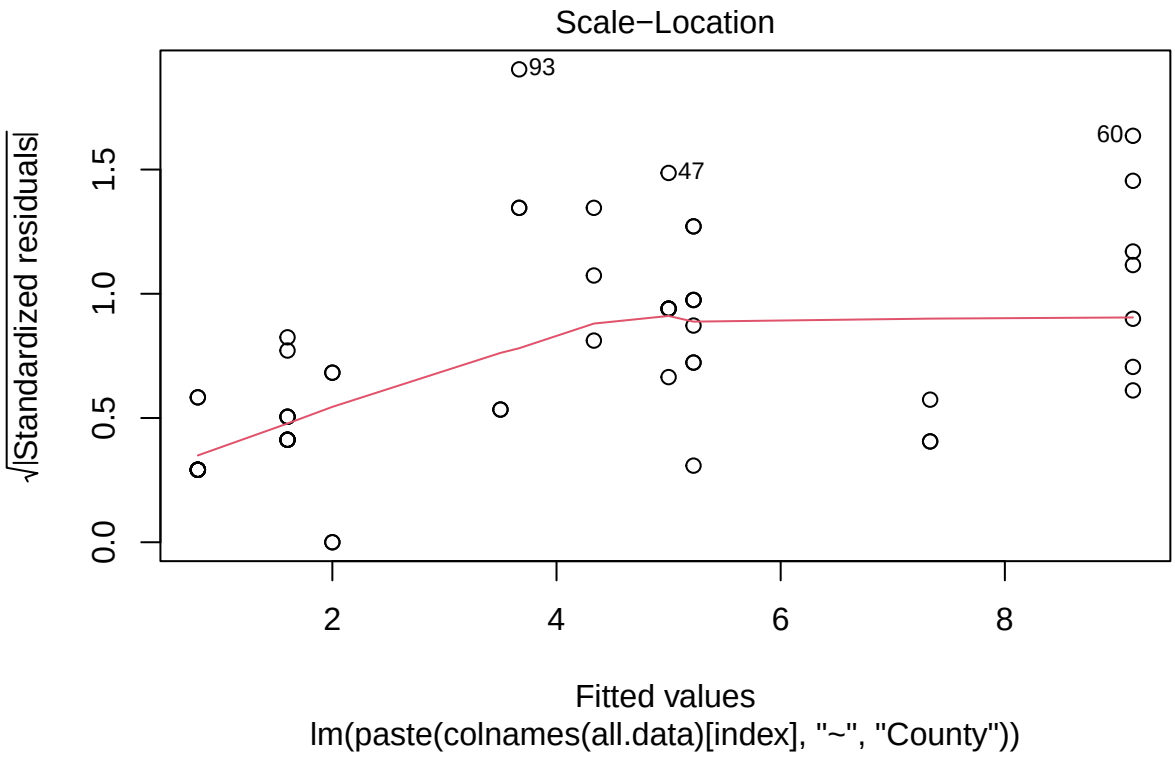


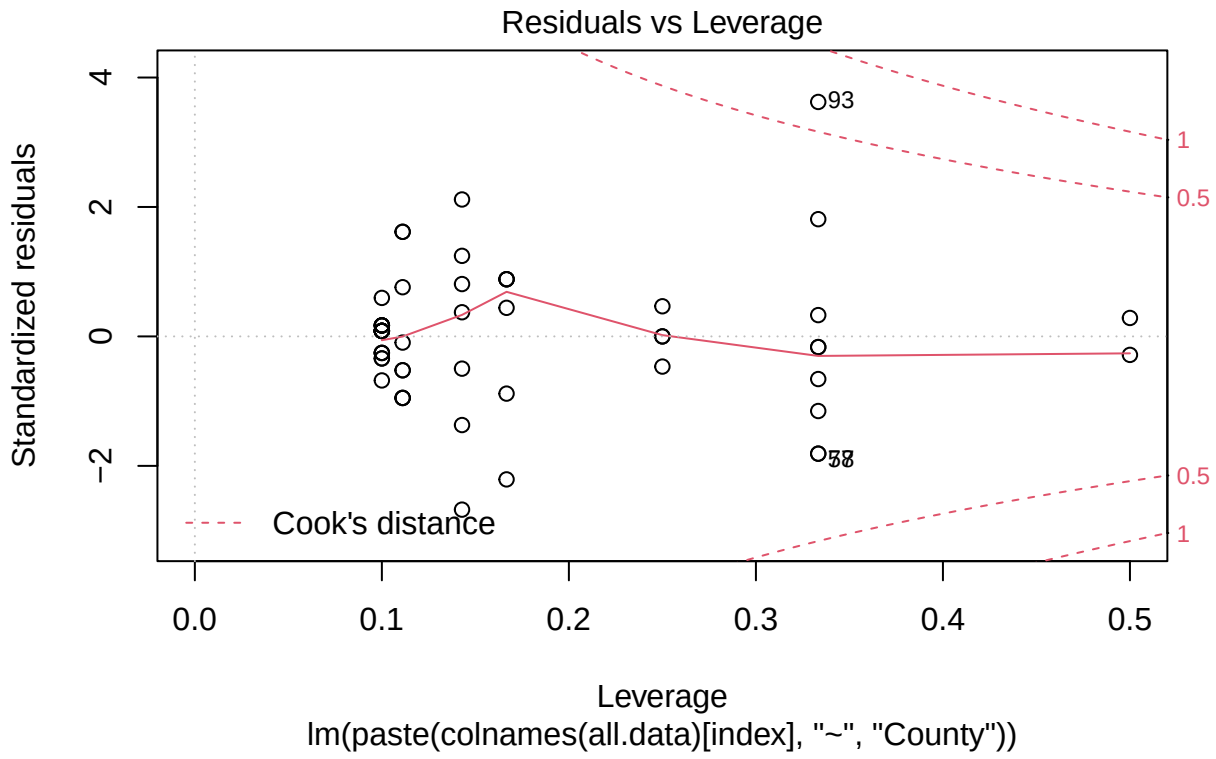

```
## NULL
##
## Call:
## lm(formula = paste(colnames(all.data)[index], "~", "County"),
##     data = all.data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -6.1429 -1.1429  0.2000  0.8571  7.3333
##
## Coefficients:
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept)      3.6667     1.4314   2.562  0.01369 *
## Countycochise    1.3333     1.7532   0.761  0.45073
## Countycoconino   5.4762     1.7109   3.201  0.00246 **
## Countymaricopa  -2.8667     1.6321  -1.756  0.08553 .
## Countymohave     0.6667     2.0244   0.329  0.74338
## Countynavajo     3.6667     2.0244   1.811  0.07649 .
## Countypima      -2.0667     1.6321  -1.266  0.21166
## Countypinal     -1.6667     1.8936  -0.880  0.38326
## Countyyavapai    1.5556     1.6529   0.941  0.35146
## Countyyuma      -0.1667     2.2633  -0.074  0.94161
## ---
## Signif. codes:  0 ’***’ 0.001 ’**’ 0.01 ’*’ 0.05 ’.’ 0.1 ’ ’ 1
##
## Residual standard error: 2.479 on 47 degrees of freedom
```

```
##   (43 observations deleted due to missingness)
## Multiple R-squared:  0.5896, Adjusted R-squared:  0.511
## F-statistic: 7.502 on 9 and 47 DF,  p-value: 1.054e-06
##
## Analysis of Variance Table
##
## Response: Carbon.Monoxide.Poisoning
##           Df Sum Sq Mean Sq F value    Pr(>F)
## County     9 415.02  46.113  7.5016 1.054e-06 ***
## Residuals 47 288.91   6.147
## ---
## Signif. codes:  0 ’***’ 0.001 ’**’ 0.01 ’*’ 0.05 ’.’ 0.1 ’ ’ 1
```



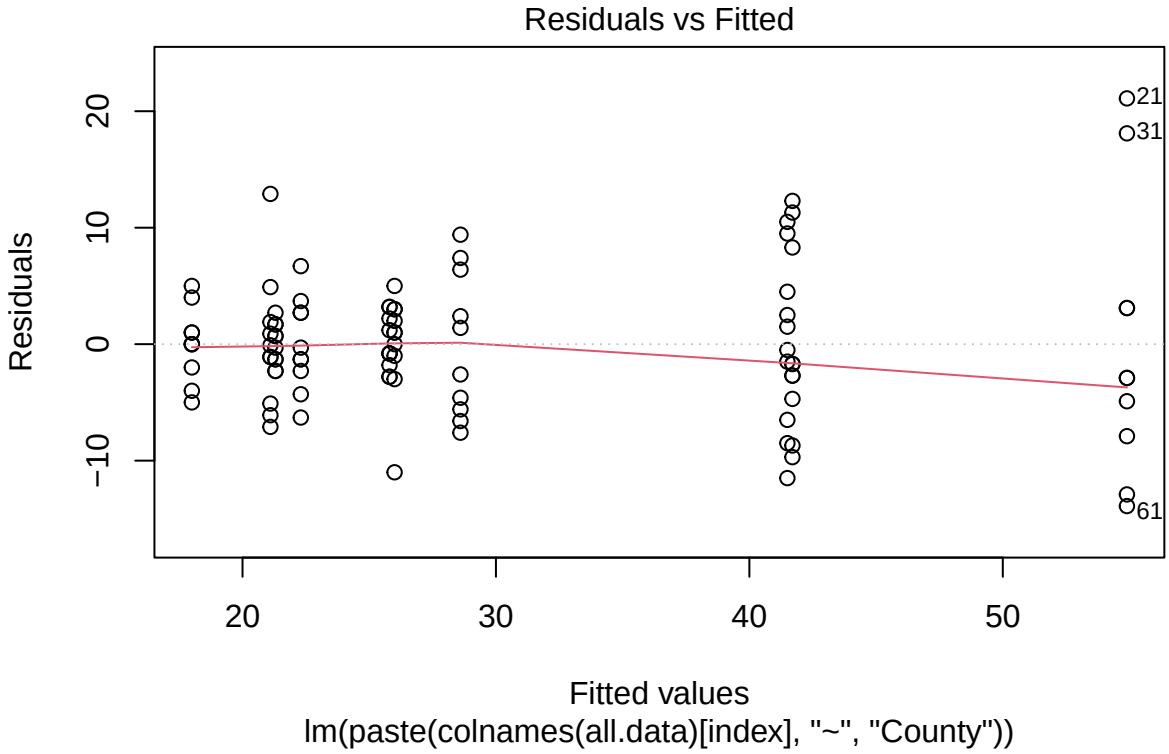

Normal Q−Q
Standardized residuals
Theoretical Quantiles
lm(paste(colnames(all.data)[index], "~", "County"))
93
47
60

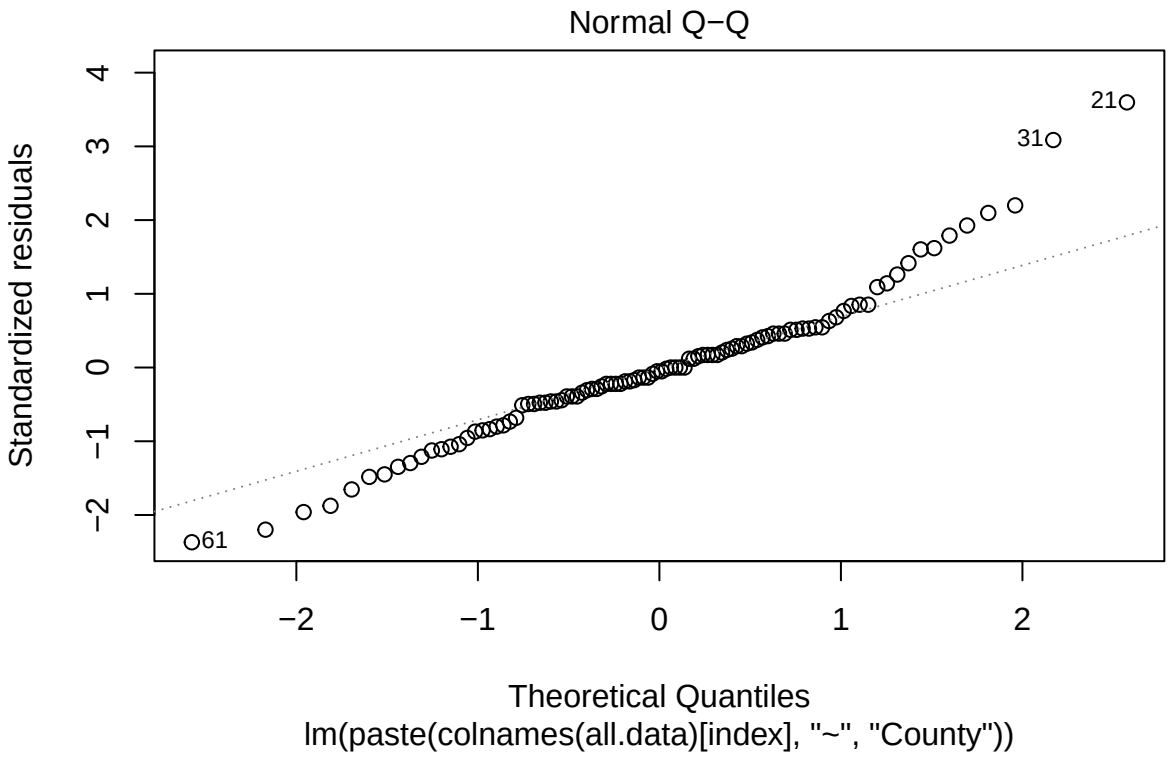

Scale−Location
93
47
60
√|Standardized residuals|
0.0
0.5
1.0
1.5
2
4
6
8
Fitted values
lm(paste(colnames(all.data)[index], "~", "County"))

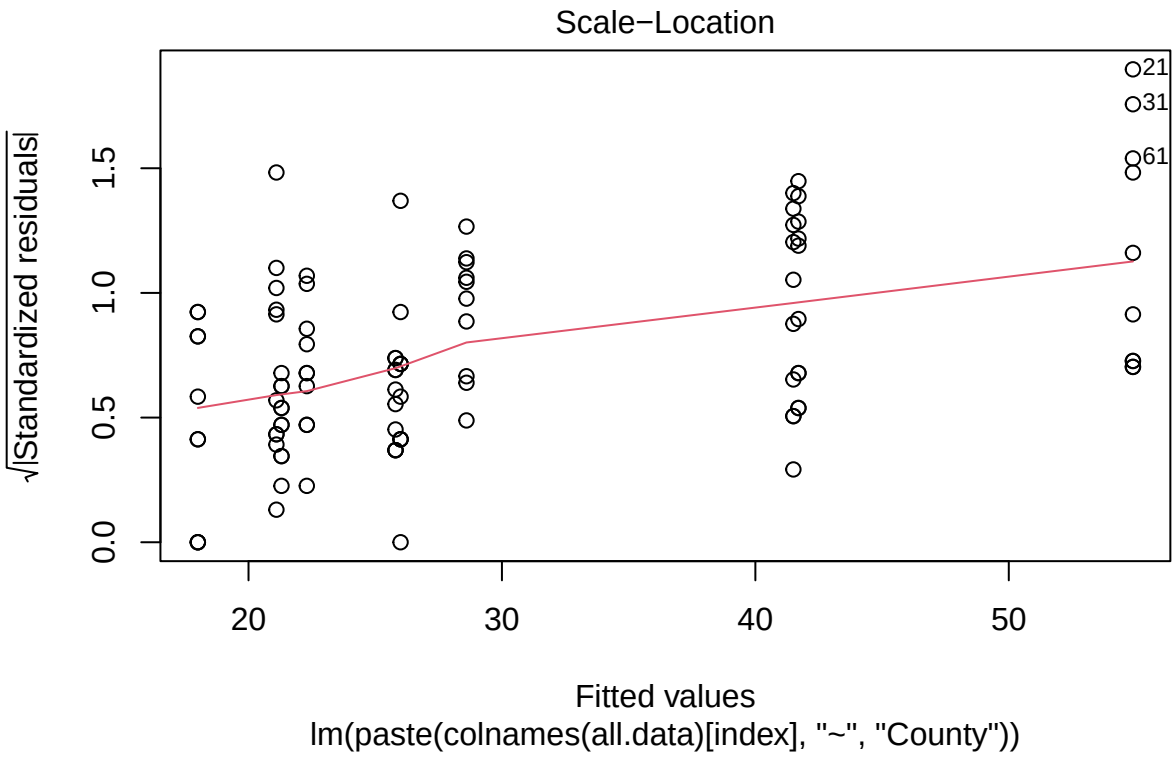


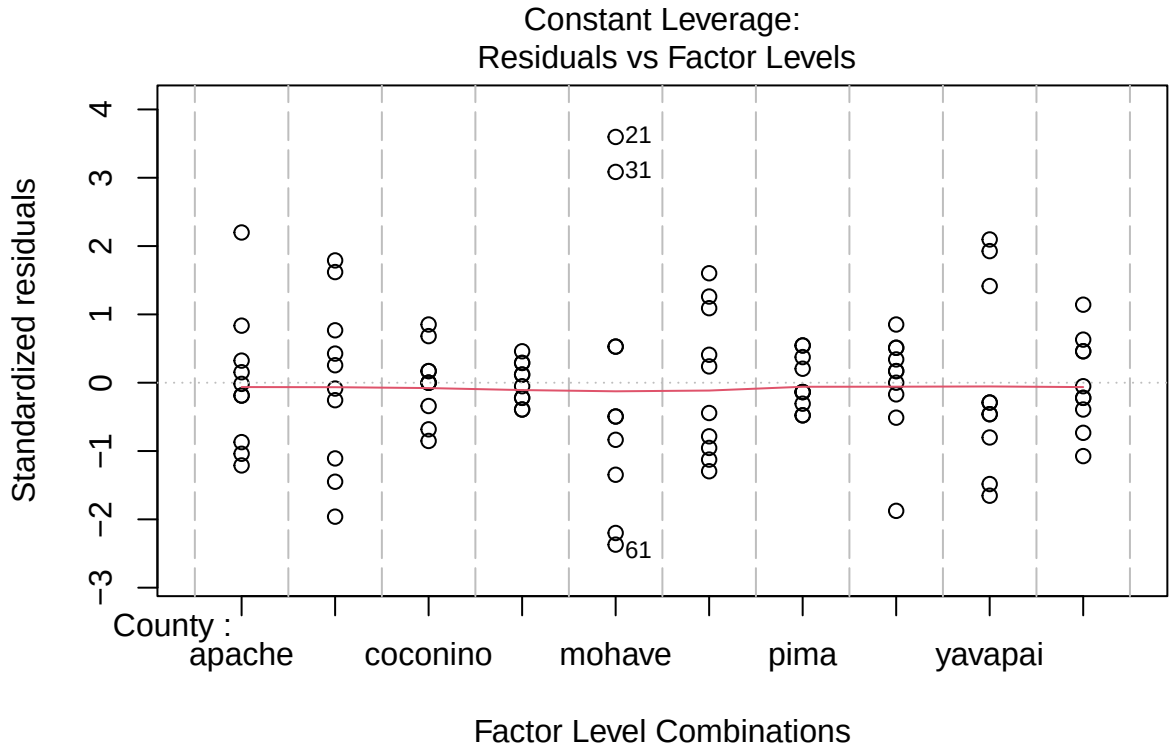

```
## NULL
##
## Call:
## lm(formula = paste(colnames(all.data)[index], "~", "County"),
##     data = all.data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -13.900  -2.825  -0.300   2.700  21.100
##
## Coefficients:
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept)      21.100      1.955  10.792  < 2e-16 ***
## Countycochise    20.400      2.765   7.378 7.66e-11 ***
## Countycoconino   -3.100      2.765  -1.121   0.2652
## Countymaricopa    0.200      2.765   0.072   0.9425
## Countymohave     33.800      2.765  12.224  < 2e-16 ***
## Countynavajo      7.500      2.765   2.712   0.0080 **
## Countypima        4.700      2.765   1.700   0.0926 .
## Countypinal       4.900      2.765   1.772   0.0798 .
## Countyyavapai    20.600      2.765   7.450 5.46e-11 ***
## Countyyuma        1.200      2.765   0.434   0.6653
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 6.183 on 90 degrees of freedom
```

```
## Multiple R-squared:  0.7885, Adjusted R-squared:  0.7674
## F-statistic: 37.28 on 9 and 90 DF,  p-value: < 2.2e-16
##
## Analysis of Variance Table
##
## Response: COPD
##           Df  Sum Sq Mean Sq F value    Pr(>F)
## County     9 12828.0 1425.33  37.284 < 2.2e-16 ***
## Residuals 90  3440.6   38.23
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```



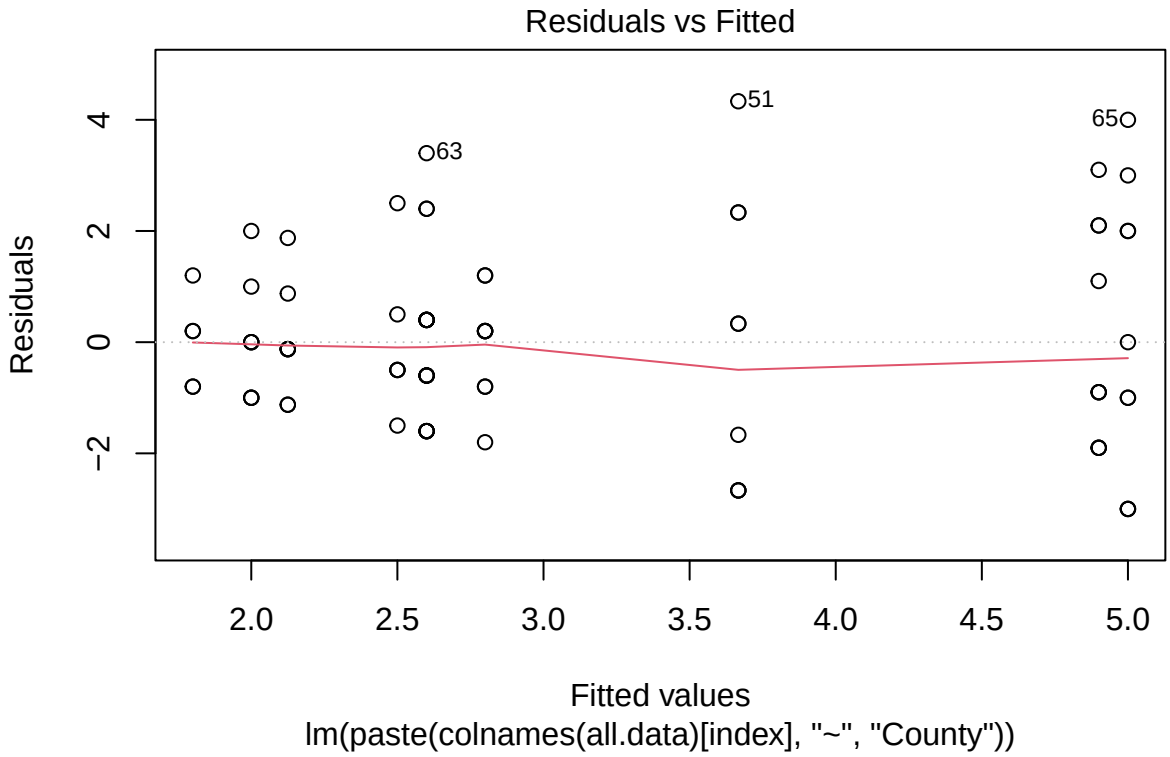

Normal Q−Q
Standardized residuals
Theoretical Quantiles
lm(paste(colnames(all.data)[index], "~", "County"))
21
31
61

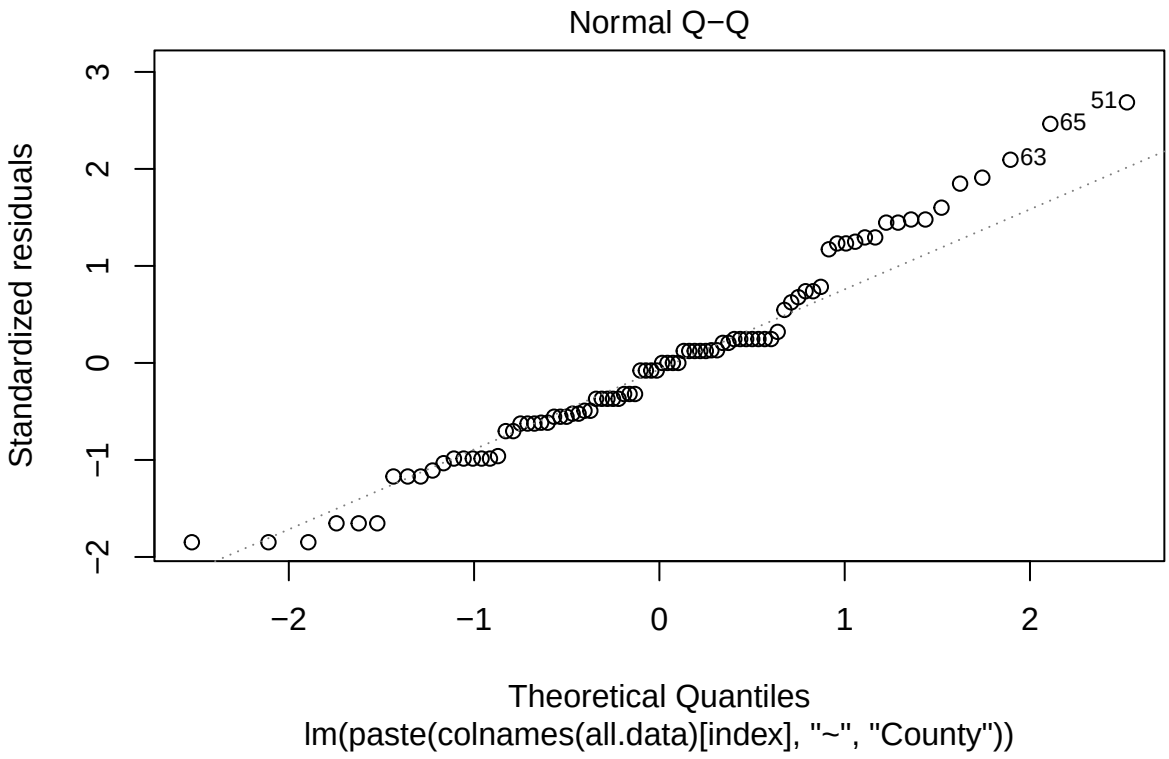

Scale−Location
√|Standardized residuals|
0.0
0.5
1.0
1.5
20
30
40
50
21
31
61
Fitted values
lm(paste(colnames(all.data)[index], "~", "County"))

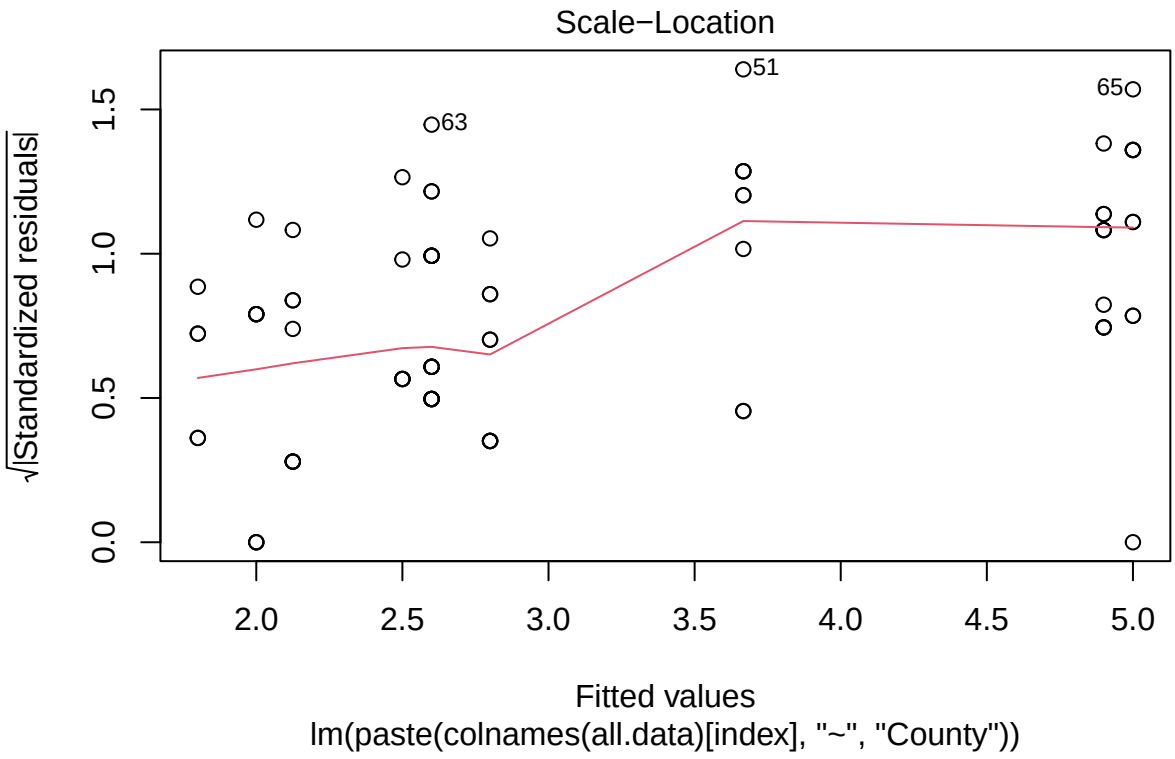


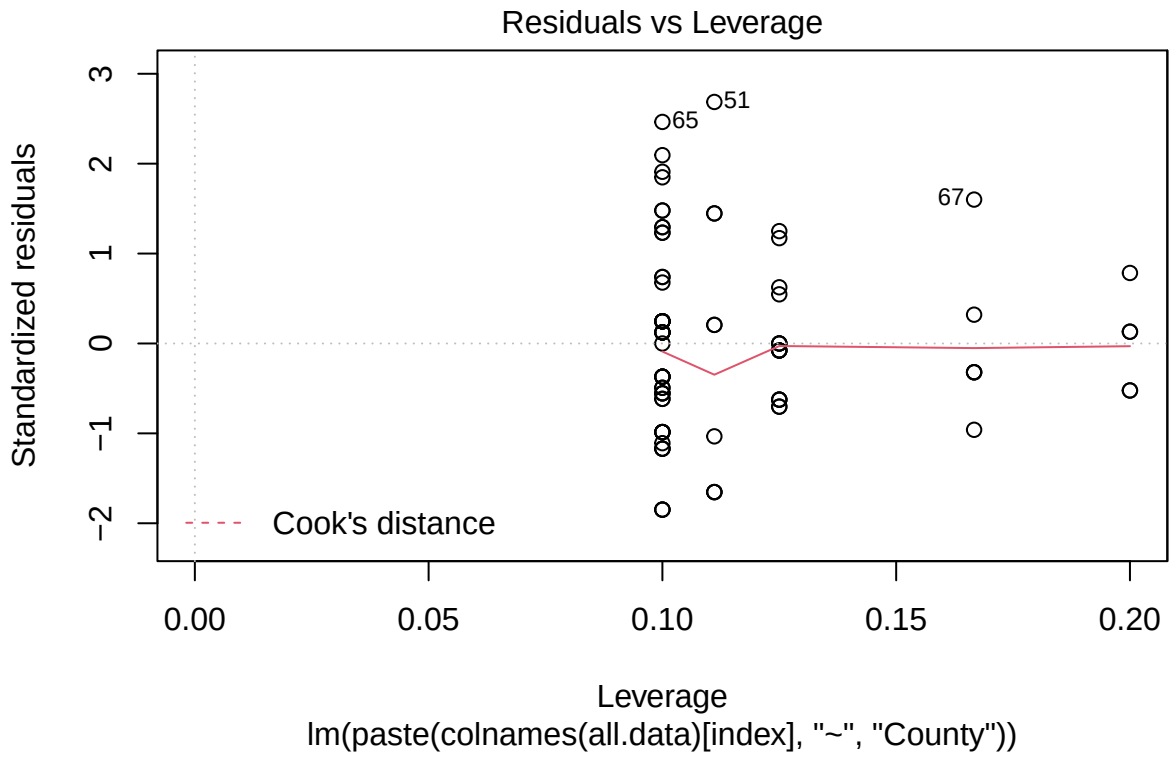

```
## NULL
##
## Call:
## lm(formula = paste(colnames(all.data)[index], "~", "County"),
##     data = all.data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -3.0000 -1.0000 -0.0625  0.7812  4.3333
##
## Coefficients:
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.8000     0.7652   2.352  0.02124 *
## Countycochise    0.8000     0.9371   0.854  0.39596
## Countycoconino   0.2000     0.9754   0.205  0.83808
## Countymaricopa   3.2000     0.9371   3.415  0.00103 **
## Countymohave     1.8667     0.9543   1.956  0.05414 .
## Countynavajo     0.3250     0.9754   0.333  0.73990
## Countypima       3.1000     0.9371   3.308  0.00144 **
## Countypinal      0.7000     1.0360   0.676  0.50131
## Countyyavapai    1.0000     0.9371   1.067  0.28930
## Countyyuma       0.8000     0.9371   0.854  0.39596
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.711 on 76 degrees of freedom
```

```
##   (14 observations deleted due to missingness)
## Multiple R-squared:  0.3209, Adjusted R-squared:  0.2405
## F-statistic:  3.99 on 9 and 76 DF,  p-value: 0.0003327
##
## Analysis of Variance Table
##
## Response: Drinking.Water.Quality
##           Df Sum Sq Mean Sq F value    Pr(>F)
## County     9 105.12 11.6798  3.9899 0.0003327 ***
## Residuals 76 222.47  2.9273
## ---
## Signif. codes:  0 ’***’ 0.001 ’**’ 0.01 ’*’ 0.05 ’.’ 0.1 ’ ’ 1
```



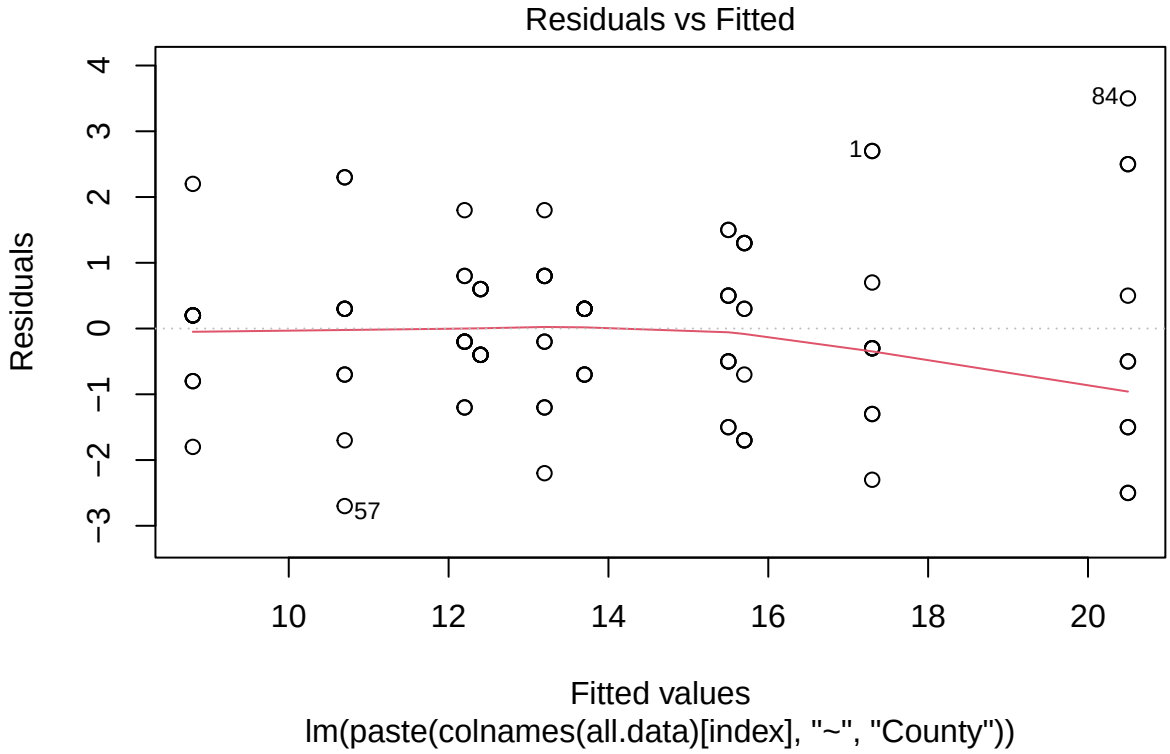

Normal Q−Q
Standardized residuals
Theoretical Quantiles
lm(paste(colnames(all.data)[index], "~", "County"))
51
65
63
−2
−1
0
1
2
3

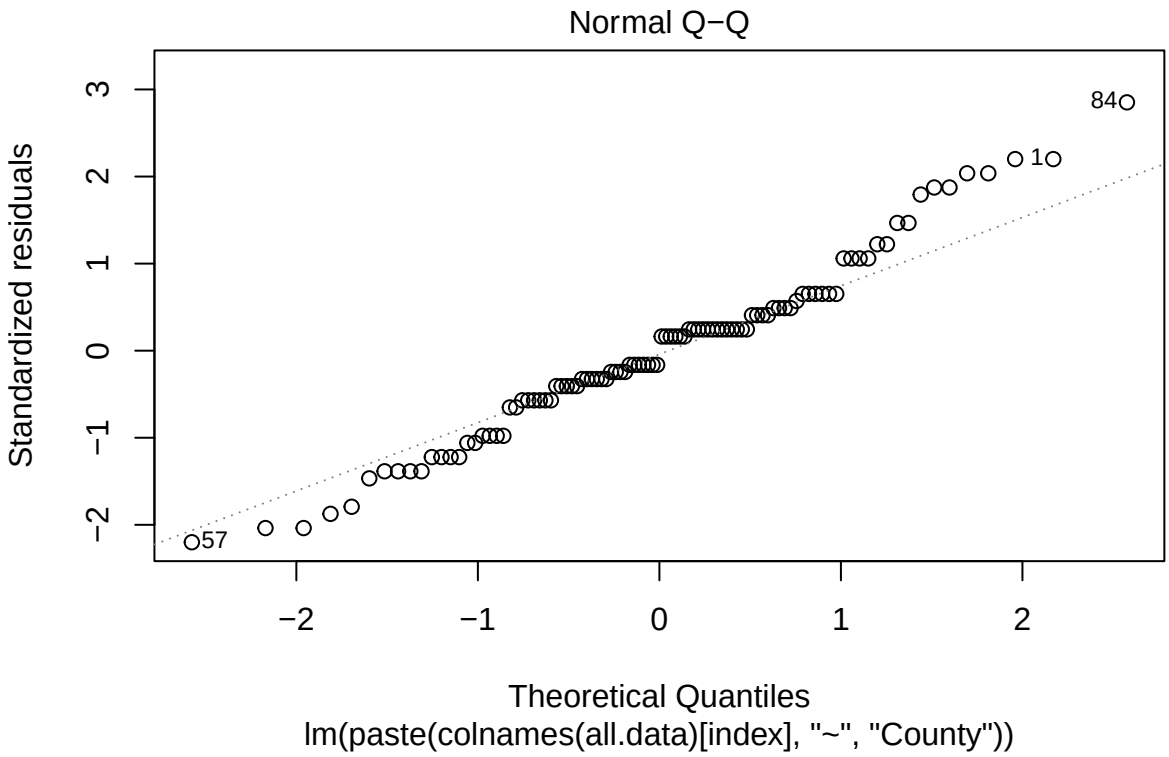

Scale−Location
√|Standardized residuals|
0.0
0.5
1.0
1.5
2.0
2.5
3.0
3.5
4.0
4.5
5.0
63
51
65
Fitted values
lm(paste(colnames(all.data)[index], "~", "County"))

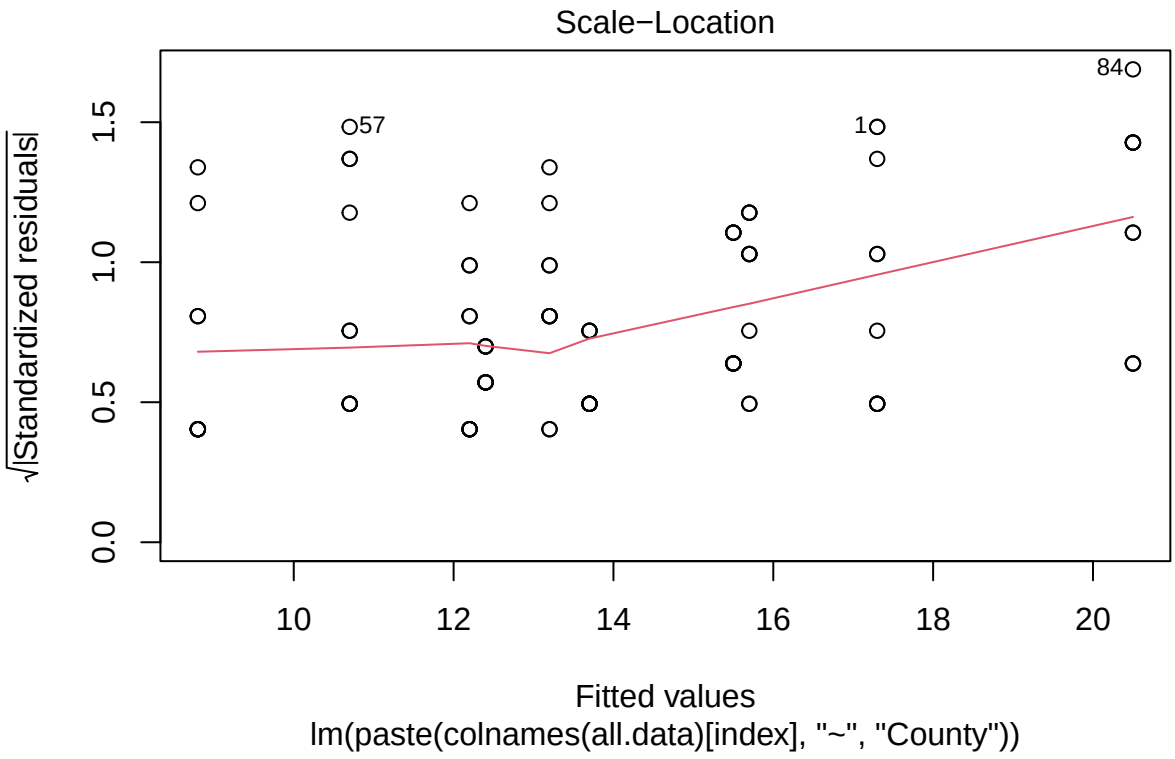


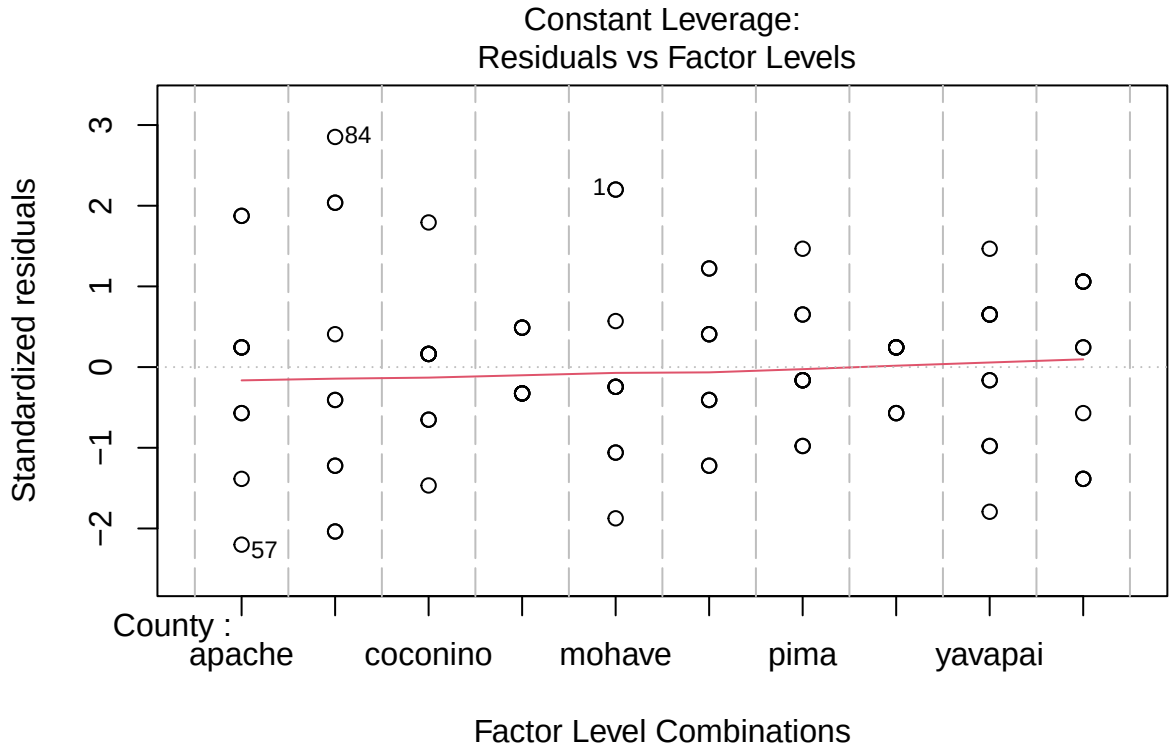

```
## NULL
##
## Call:
## lm(formula = paste(colnames(all.data)[index], "~", "County"),
##     data = all.data)
##
## Residuals:
##    Min     1Q Median     3Q    Max
##   -2.7   -0.7    0.0    0.6    3.5
##
## Coefficients:
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept)     10.7000     0.4091  26.157  < 2e-16 ***
## Countycochise    9.8000     0.5785  16.940  < 2e-16 ***
## Countycoconino  -1.9000     0.5785  -3.284  0.00146 **
## Countymaricopa   1.7000     0.5785   2.939  0.00419 **
## Countymohave     6.6000     0.5785  11.409  < 2e-16 ***
## Countynavajo     4.8000     0.5785   8.297 9.92e-13 ***
## Countypima       1.5000     0.5785   2.593  0.01111 *
## Countypinal      3.0000     0.5785   5.186 1.31e-06 ***
## Countyyavapai    2.5000     0.5785   4.321 3.99e-05 ***
## Countyyuma       5.0000     0.5785   8.643 1.90e-13 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.294 on 90 degrees of freedom
```

```
## Multiple R-squared:  0.8722, Adjusted R-squared:  0.8594
## F-statistic: 68.22 on 9 and 90 DF,  p-value: < 2.2e-16
##
## Analysis of Variance Table
##
## Response: Heart.Disease
##           Df Sum Sq Mean Sq F value    Pr(>F)
## County     9 1027.4 114.156  68.221 < 2.2e-16 ***
## Residuals 90  150.6   1.673
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```



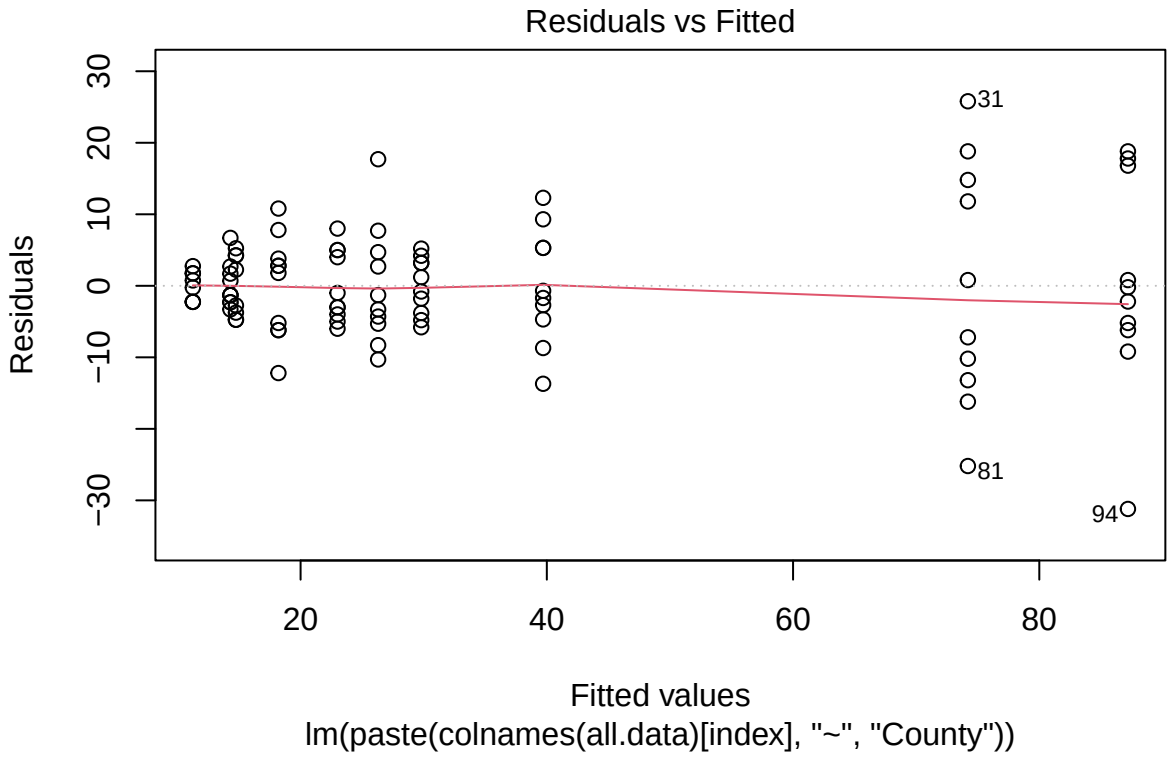

Normal Q−Q
Standardized residuals
Theoretical Quantiles
lm(paste(colnames(all.data)[index], "~", "County"))
84
1
57
−2
−1
0
1
2
3

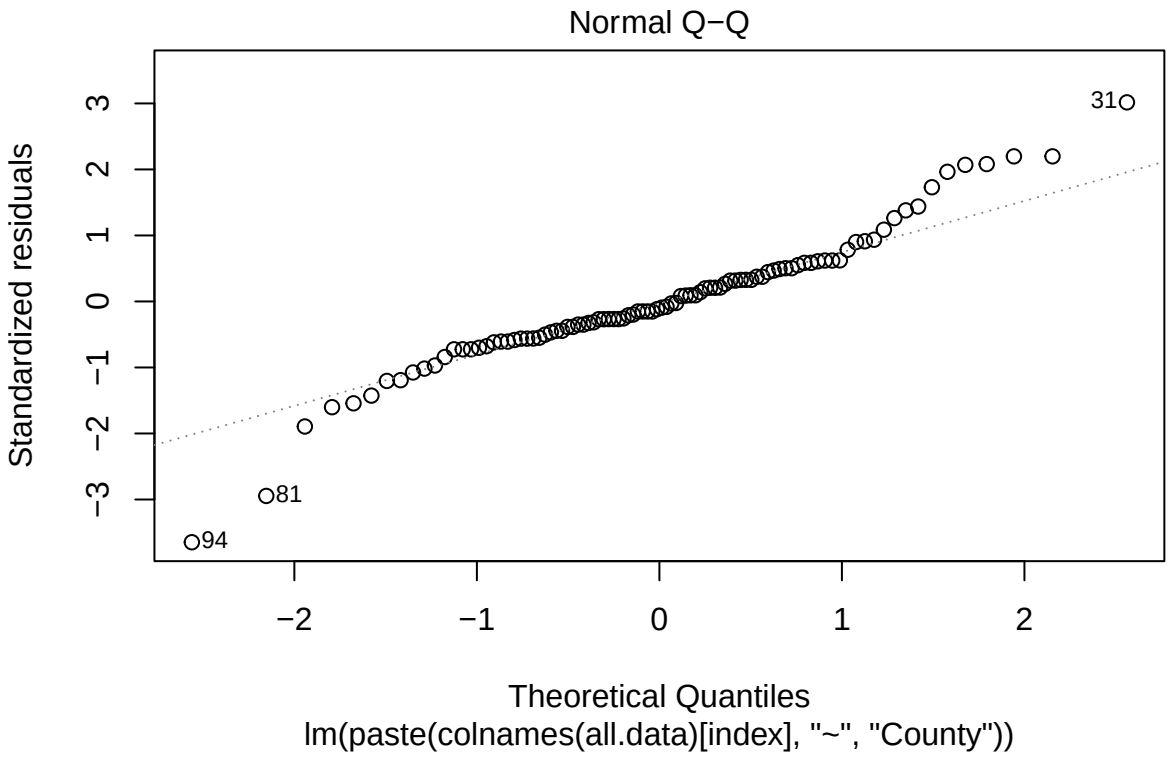

Scale−Location
√|Standardized residuals|
0.0
0.5
1.0
1.5
57
1
84
10
12
14
16
18
20
Fitted values
lm(paste(colnames(all.data)[index], "~", "County"))

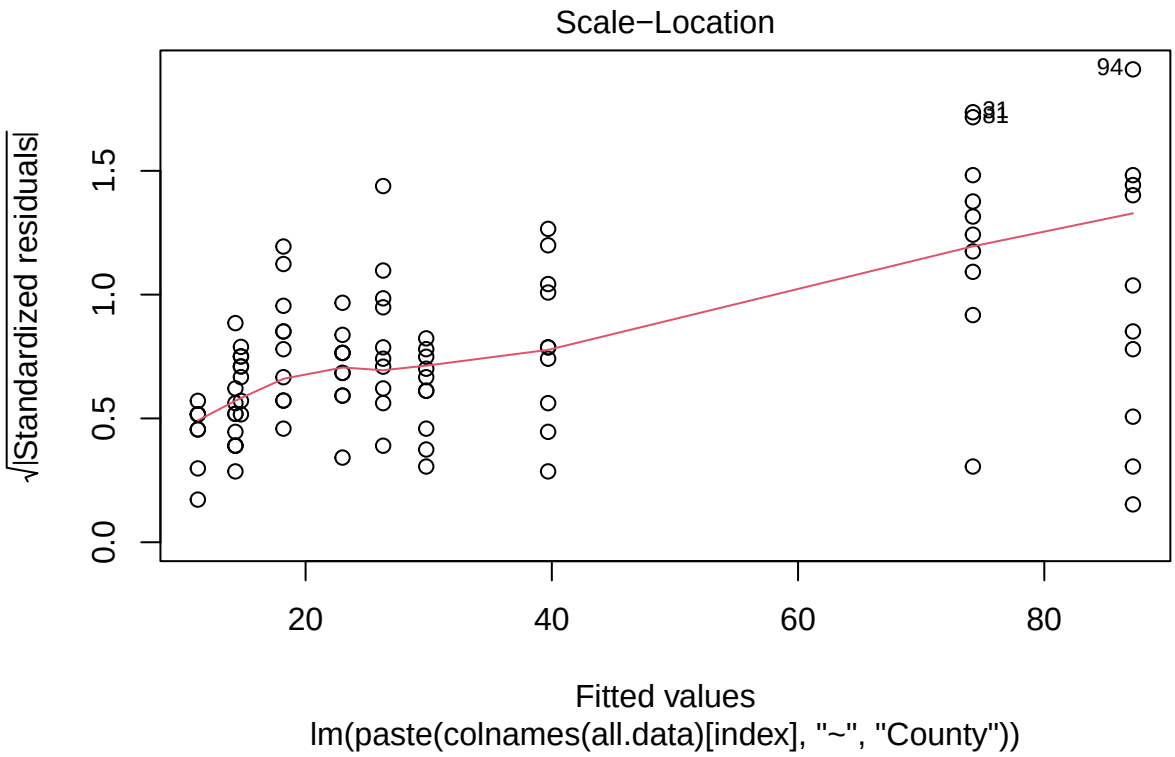


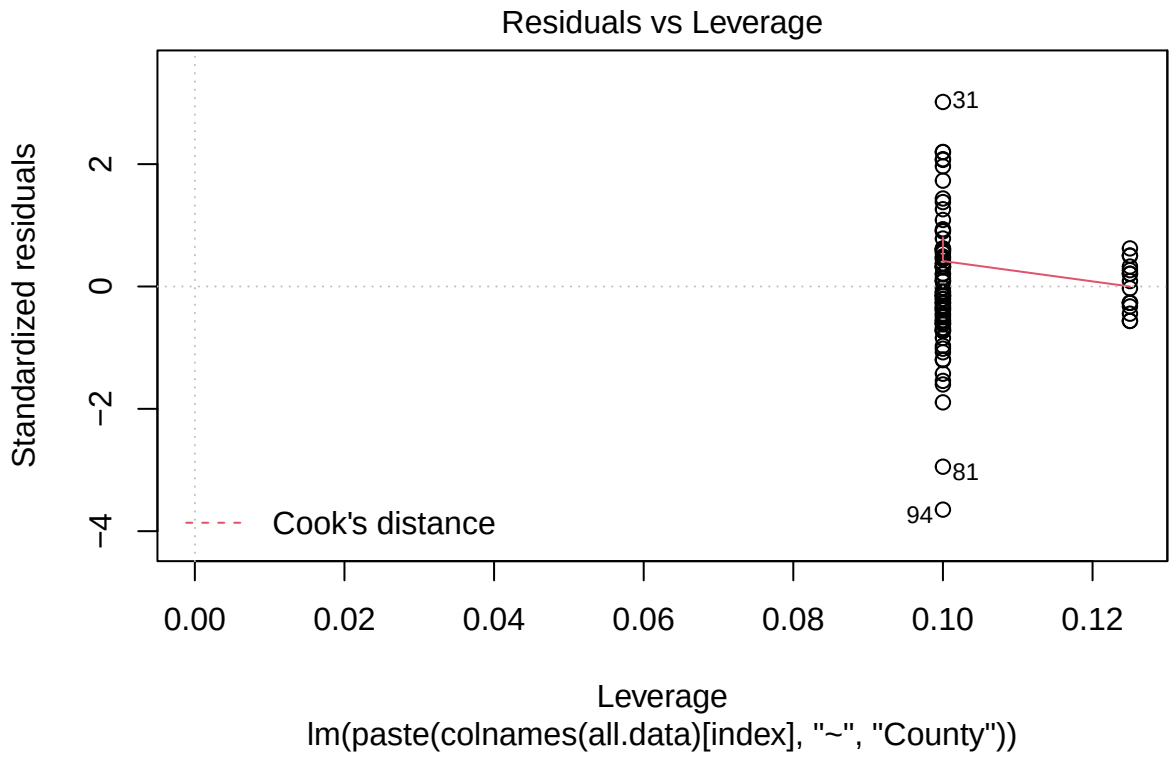

```
## NULL
##
## Call:
## lm(formula = paste(colnames(all.data)[index], "~", "County"),
##     data = all.data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -31.200  -4.713  -0.900   4.213  25.800
##
## Coefficients:
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept)      11.250      3.188   3.529 0.000672 ***
## Countycochise    15.050      4.277   3.519 0.000695 ***
## Countycoconino    3.050      4.277   0.713 0.477698
## Countymaricopa   18.550      4.277   4.337 3.91e-05 ***
## Countymohave     62.950      4.277  14.718  < 2e-16 ***
## Countynavajo      3.500      4.508   0.776 0.439674
## Countypima       11.750      4.277   2.747 0.007319 **
## Countypinal      28.450      4.277   6.652 2.55e-09 ***
## Countyyavapai     6.950      4.277   1.625 0.107824
## Countyyuma       75.950      4.277  17.758  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 9.017 on 86 degrees of freedom
```

```
##   (4 observations deleted due to missingness)
## Multiple R-squared:  0.8959, Adjusted R-squared:  0.885
## F-statistic: 82.27 on 9 and 86 DF,  p-value: < 2.2e-16
##
## Analysis of Variance Table
##
## Response: Heat.Stress.Illness
##           Df Sum Sq Mean Sq F value    Pr(>F)
## County     9  60197  6688.5  82.271 < 2.2e-16 ***
## Residuals 86   6992    81.3
## ---
## Signif. codes:  0 ’***’ 0.001 ’**’ 0.01 ’*’ 0.05 ’.’ 0.1 ’ ’ 1
```



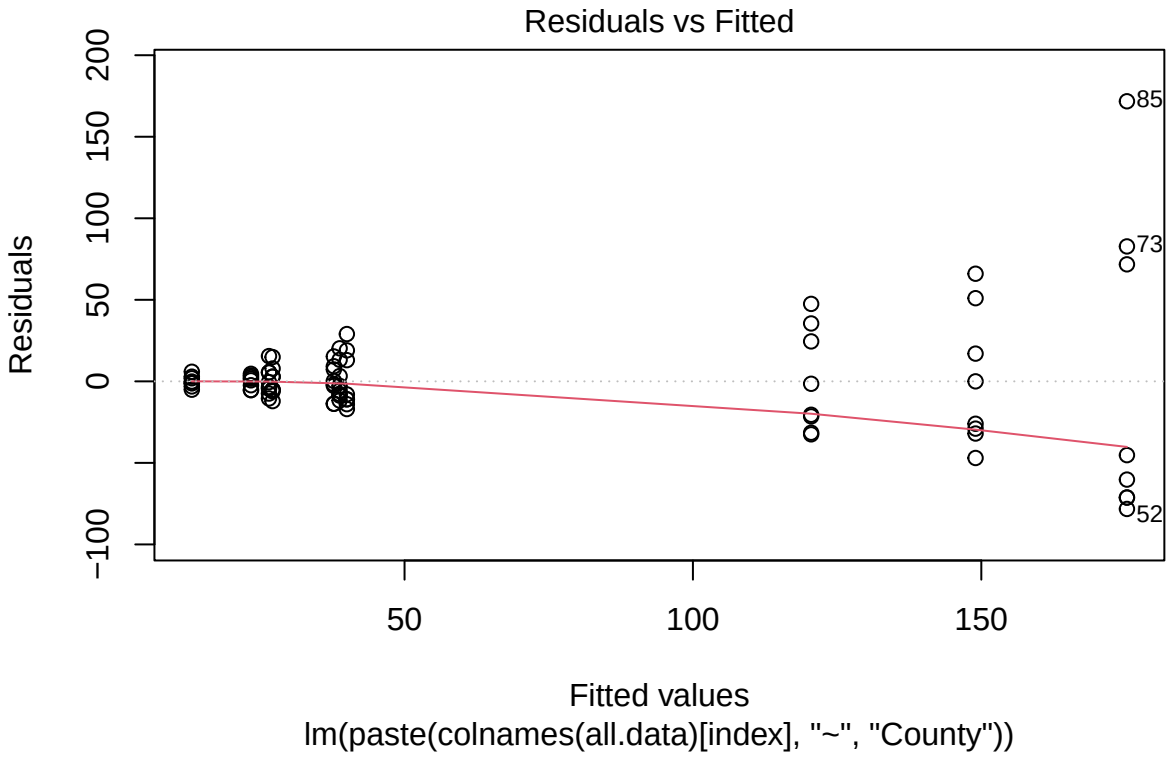

Normal Q−Q
Standardized residuals
Theoretical Quantiles
lm(paste(colnames(all.data)[index], "~", "County"))
31
81
94
−3
−2
−1
0
1
2
3

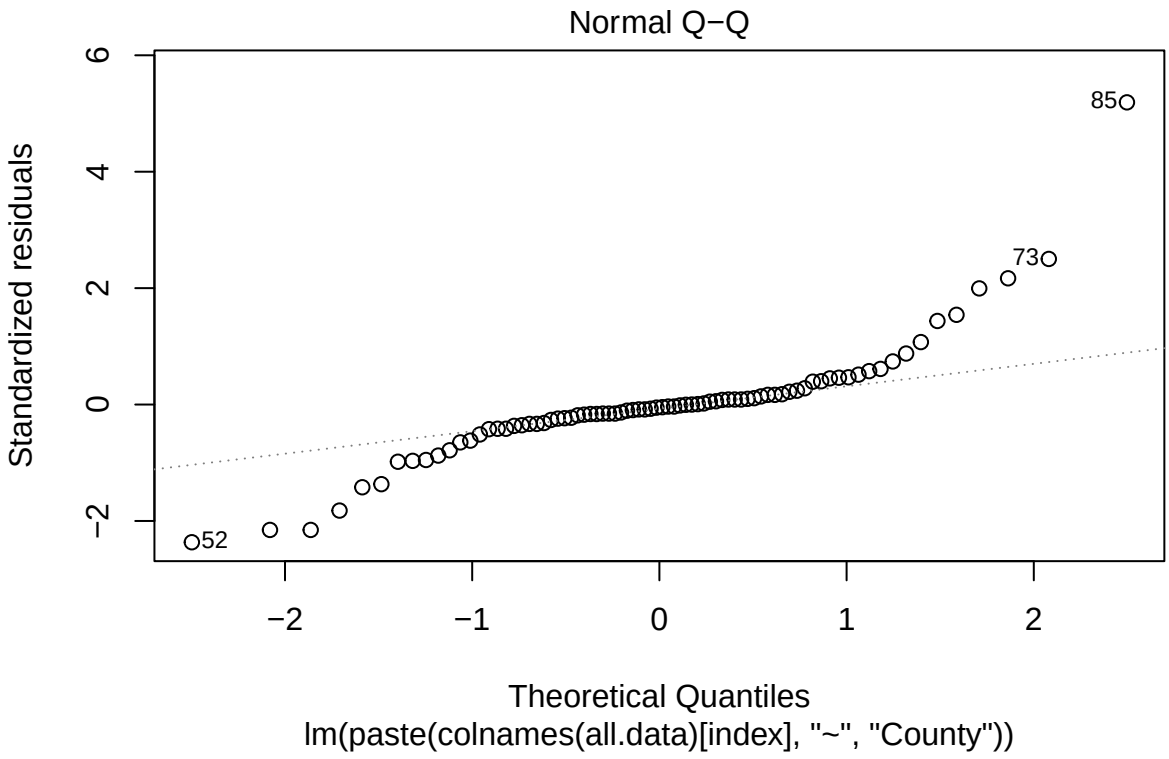

Scale−Location
94
√|Standardized residuals|
0.0
0.5
1.0
1.5
20
40
60
80
Fitted values
lm(paste(colnames(all.data)[index], "~", "County"))

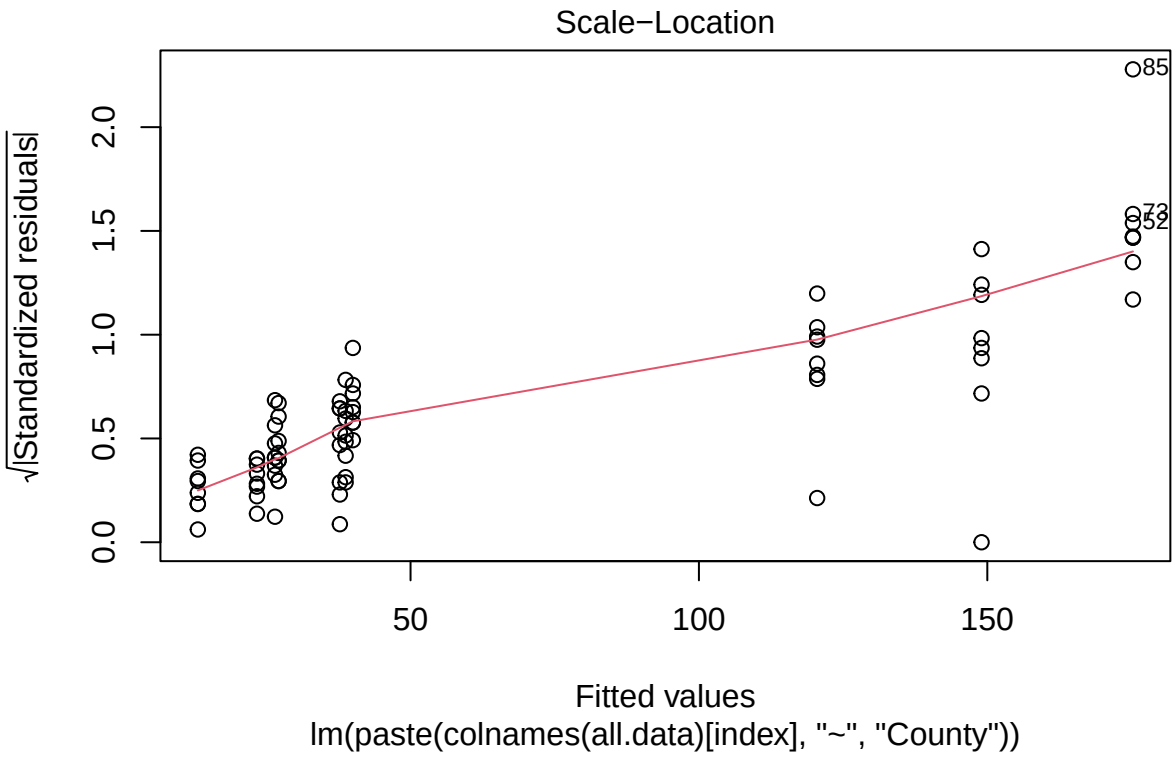


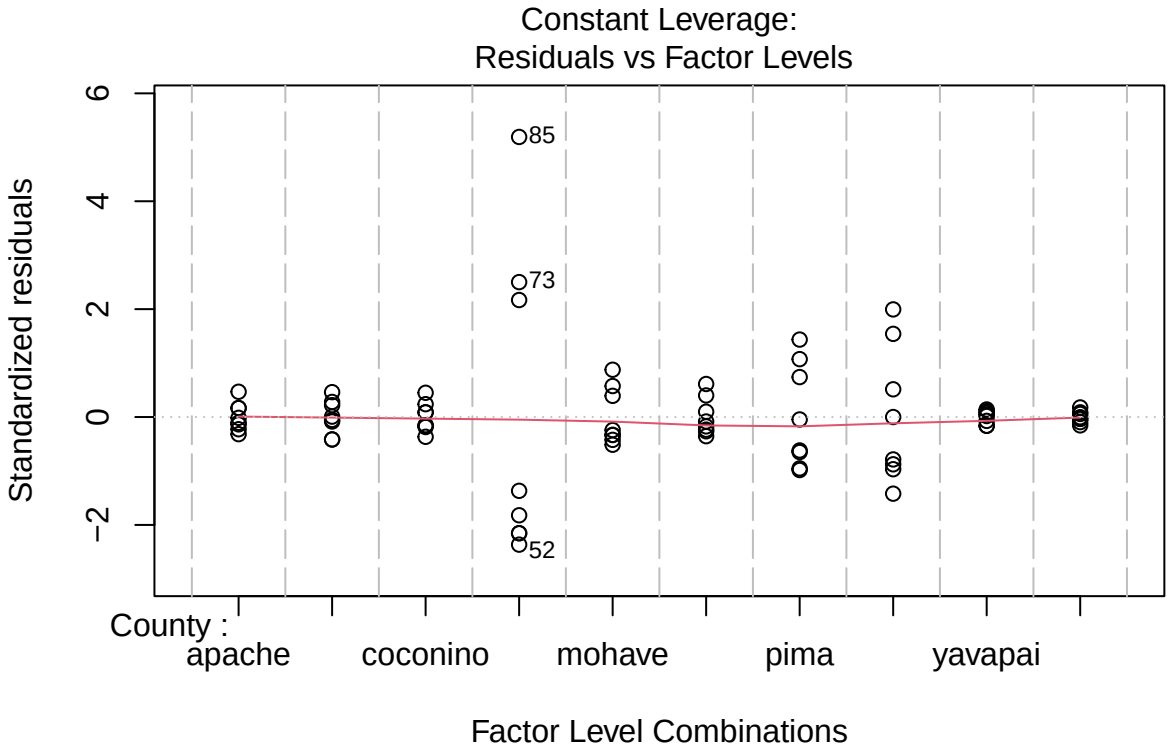

```
## NULL
##
## Call:
## lm(formula = paste(colnames(all.data)[index], "~", "County"),
##     data = all.data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -78.250 -11.000  -1.625   6.219 171.750
##
## Coefficients:
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept)      26.500     12.504   2.119   0.0376 *
## Countycochise    11.250     17.683   0.636   0.5267
## Countycoconino    0.625     17.683   0.035   0.9719
## Countymaricopa  148.750     17.683   8.412 3.18e-12 ***
## Countymohave     13.500     17.683   0.763   0.4478
## Countynavajo     12.250     17.683   0.693   0.4908
## Countypima       94.000     17.683   5.316 1.20e-06 ***
## Countypinal     122.500     17.683   6.927 1.69e-09 ***
## Countyyavapai    -3.125     17.683  -0.177   0.8602
## Countyyuma      -13.375     17.683  -0.756   0.4520
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 35.37 on 70 degrees of freedom
```

```
##   (20 observations deleted due to missingness)
## Multiple R-squared:  0.7433, Adjusted R-squared:  0.7103
## F-statistic: 22.52 on 9 and 70 DF,  p-value: < 2.2e-16
##
## Analysis of Variance Table
##
## Response: Infectious.Diseases
##           Df Sum Sq Mean Sq F value    Pr(>F)
## County     9 253505 28167.3  22.519 < 2.2e-16 ***
## Residuals 70  87556  1250.8
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```



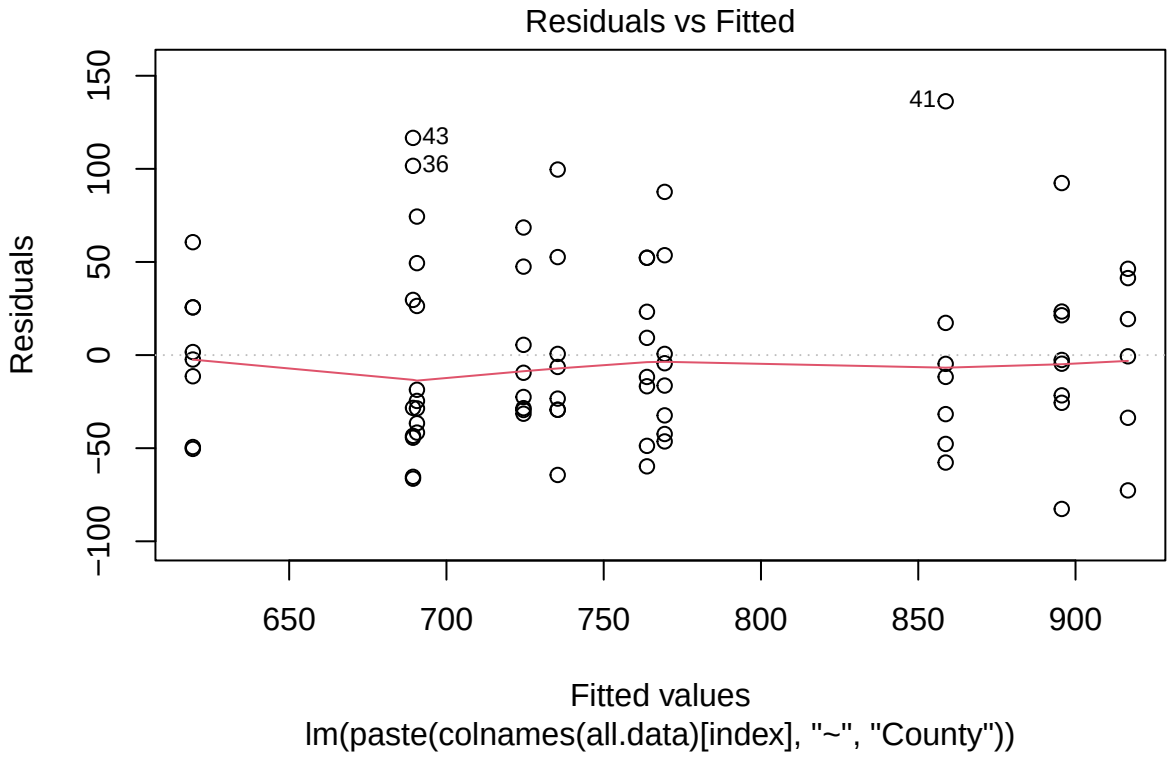

Normal Q−Q
Standardized residuals
Theoretical Quantiles
lm(paste(colnames(all.data)[index], "~", "County"))
85
73
52
−2
0
2
4
6
−2
−1
0
1
2

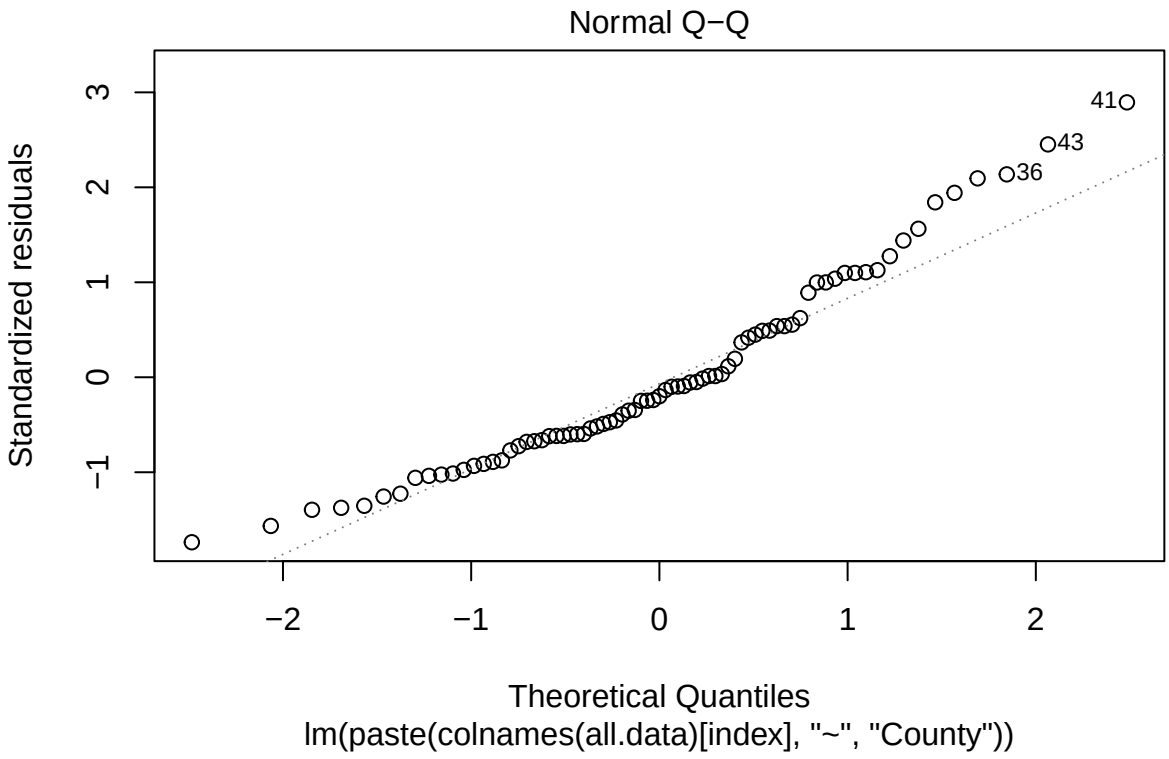

Scale−Location
√|Standardized residuals|
0.0
0.5
1.0
1.5
2.0
50
100
150
85
73
52
Fitted values
lm(paste(colnames(all.data)[index], "~", "County"))

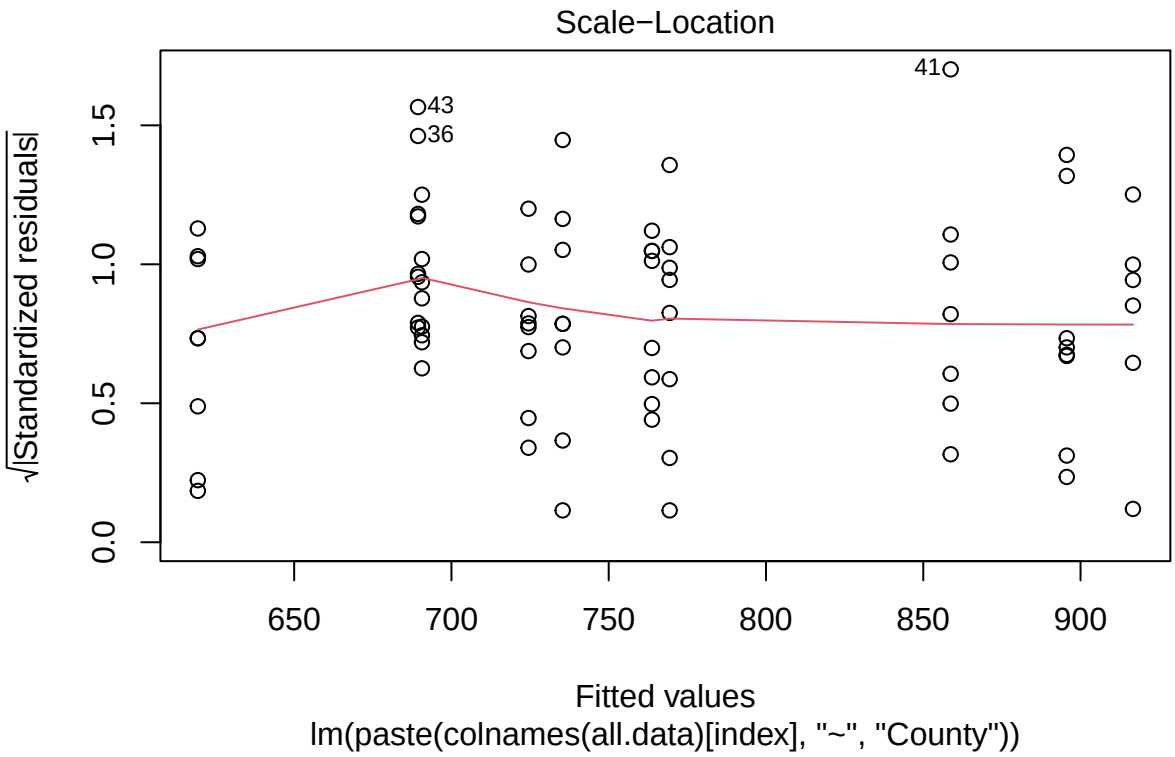


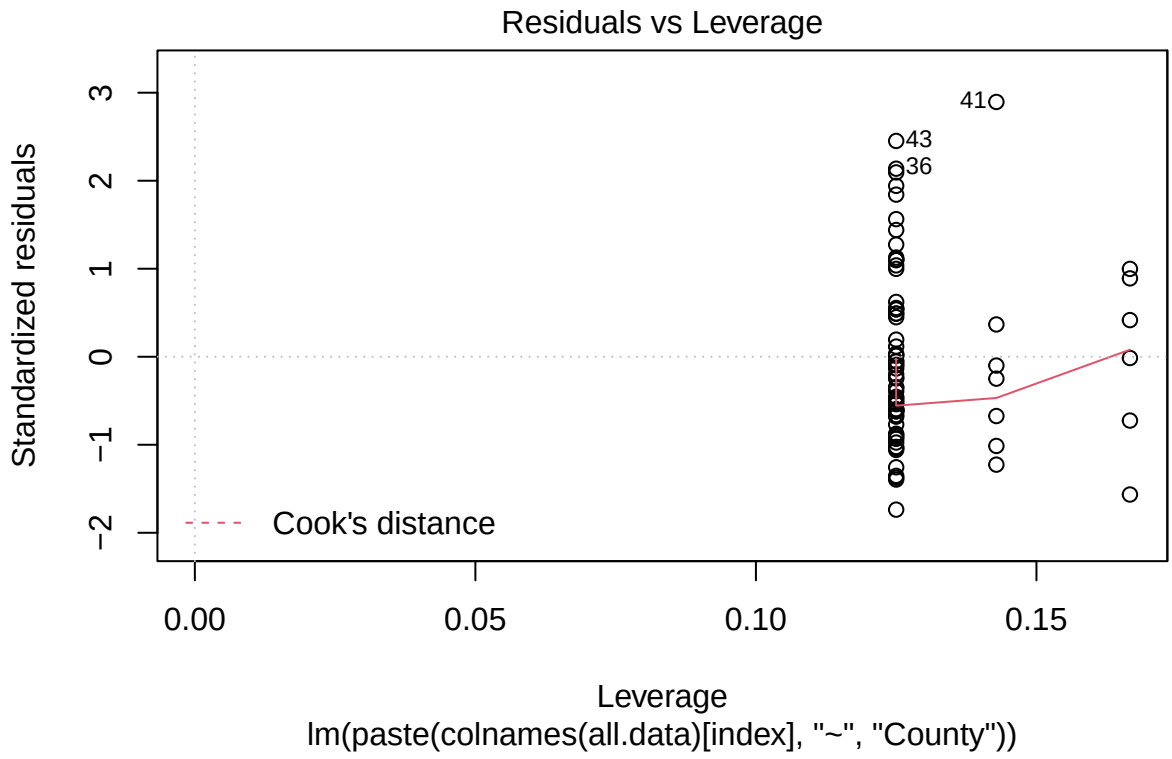

```
## NULL
##
## Call:
## lm(formula = paste(colnames(all.data)[index], "~", "County"),
##     data = all.data)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -82.62 -31.71  -9.50  25.62 136.29
##
## Coefficients:
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept)      916.67      20.76  44.156  < 2e-16 ***
## Countycochise   -147.29      27.46  -5.363 1.09e-06 ***
## Countycoconino  -181.29      27.46  -6.601 7.78e-09 ***
## Countymaricopa  -226.04      27.46  -8.231 9.23e-12 ***
## Countymohave     -57.95      28.29  -2.048   0.0444 *
## Countynavajo     -21.04      27.46  -0.766   0.4463
## Countypima      -192.17      27.46  -6.997 1.53e-09 ***
## Countypinal     -227.29      27.46  -8.276 7.64e-12 ***
## Countyyavapai   -152.92      27.46  -5.568 4.91e-07 ***
## Countyyuma      -297.29      27.46 -10.825 2.34e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 50.85 on 67 degrees of freedom
```

```
##   (23 observations deleted due to missingness)
## Multiple R-squared:  0.7802, Adjusted R-squared:  0.7506
## F-statistic: 26.42 on 9 and 67 DF,  p-value: < 2.2e-16
##
## Analysis of Variance Table
##
## Response: Mortality
##           Df Sum Sq Mean Sq F value    Pr(>F)
## County     9 614872   68319  26.421 < 2.2e-16 ***
## Residuals 67 173250    2586
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```




Normal Q−Q
Standardized residuals
Theoretical Quantiles
lm(paste(colnames(all.data)[index], "~", "County"))
41
43
36
−2
−1
0
1
2
−1
0
1
2
3


Scale−Location
√|Standardized residuals|
0.0
0.5
1.0
1.5
650
700
750
800
850
900
43
36
41
Fitted values
lm(paste(colnames(all.data)[index], "~", "County"))

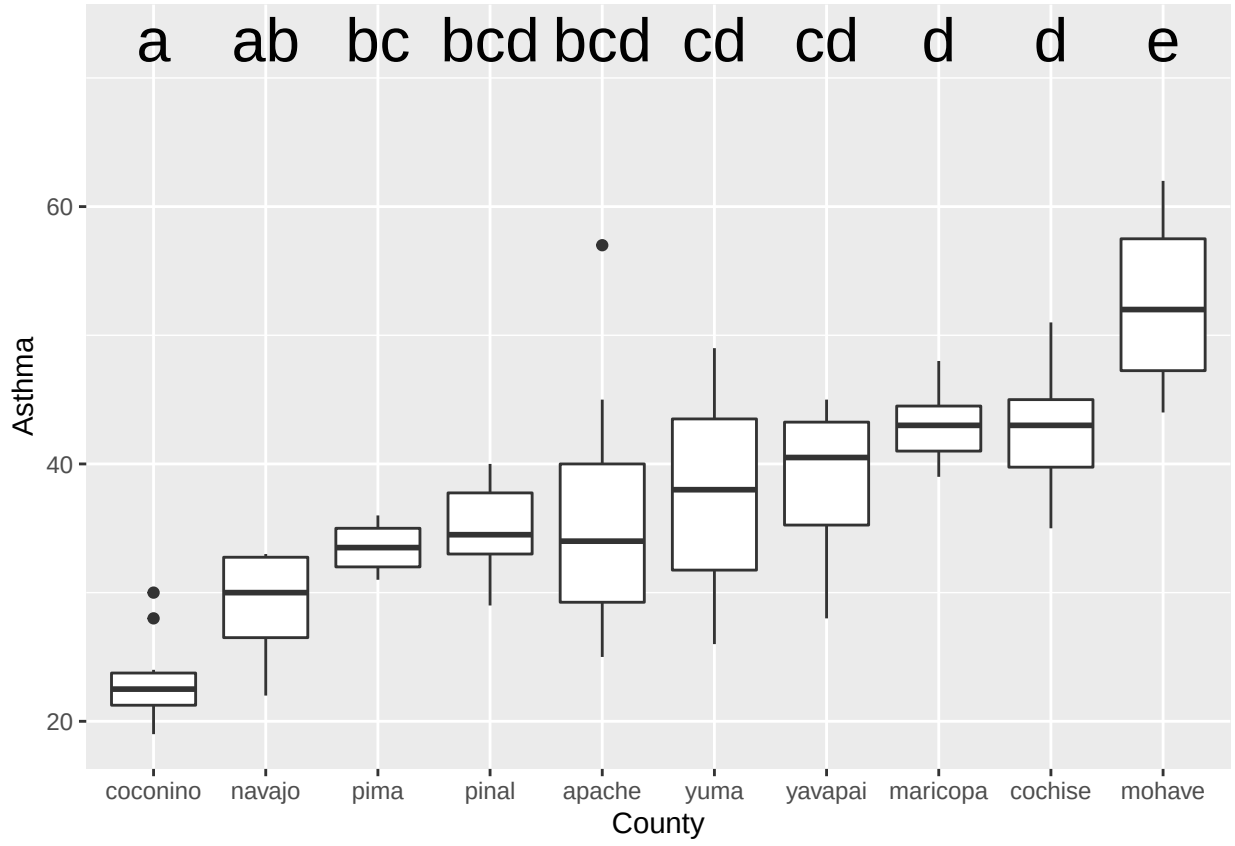



```
## NULL
```

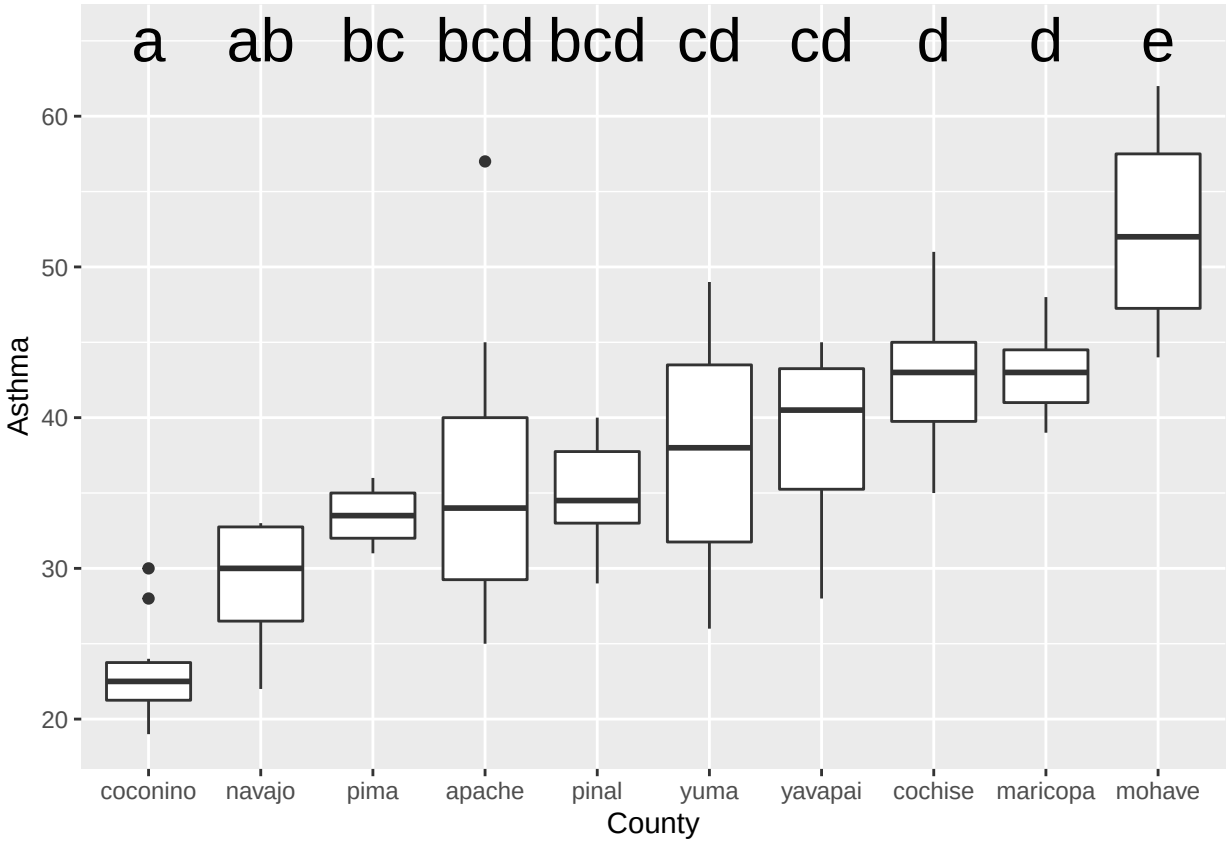
```
content.area.p <- content.area.p %>% arrange(p.value)
```

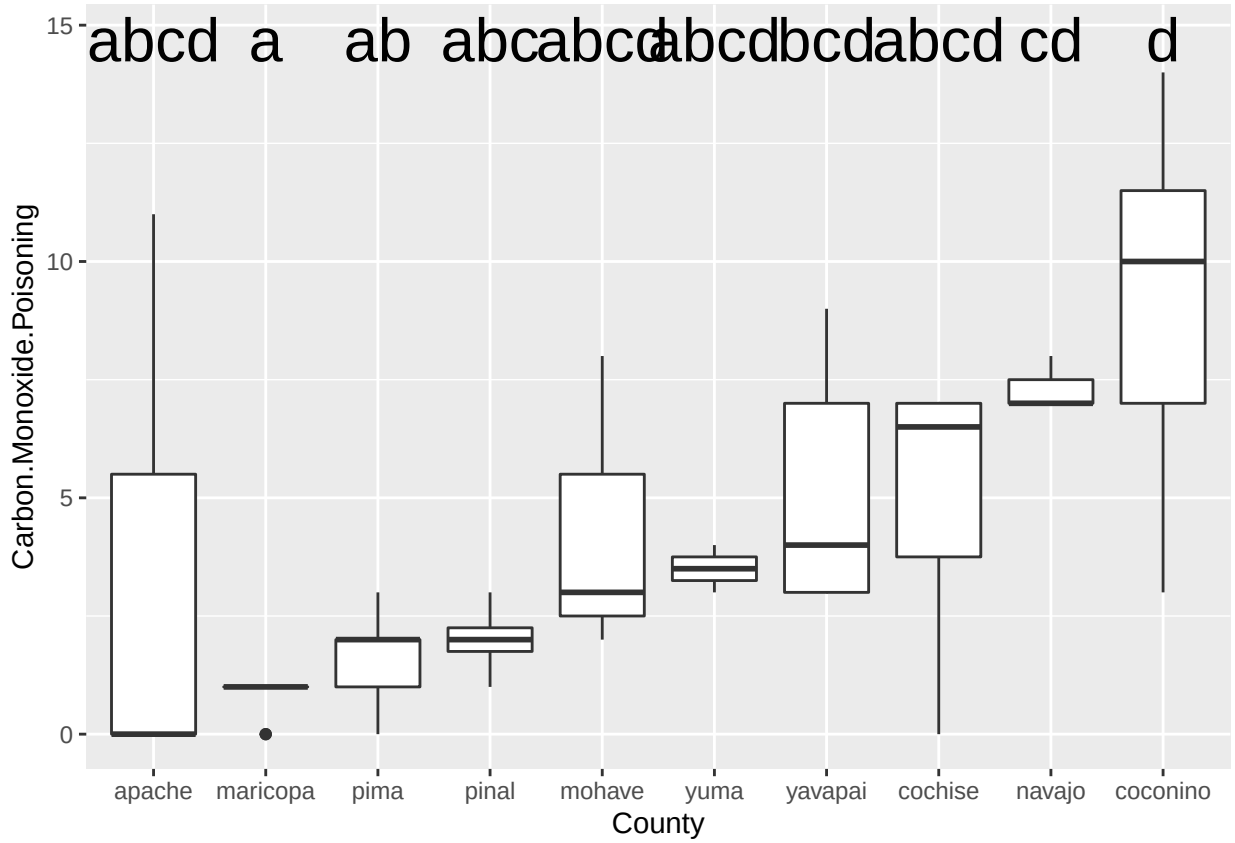
```
obj <- lm(Asthma ~ County, data=all.data)
letter.data.asthma <- emmeans(obj, specs = ~ County) %>%
  multcomp::cld(Letters = letters, level = 0.95) %>%
  mutate(.group = str_remove_all(.group, '\\s')) %>%
  mutate(y=73)
```

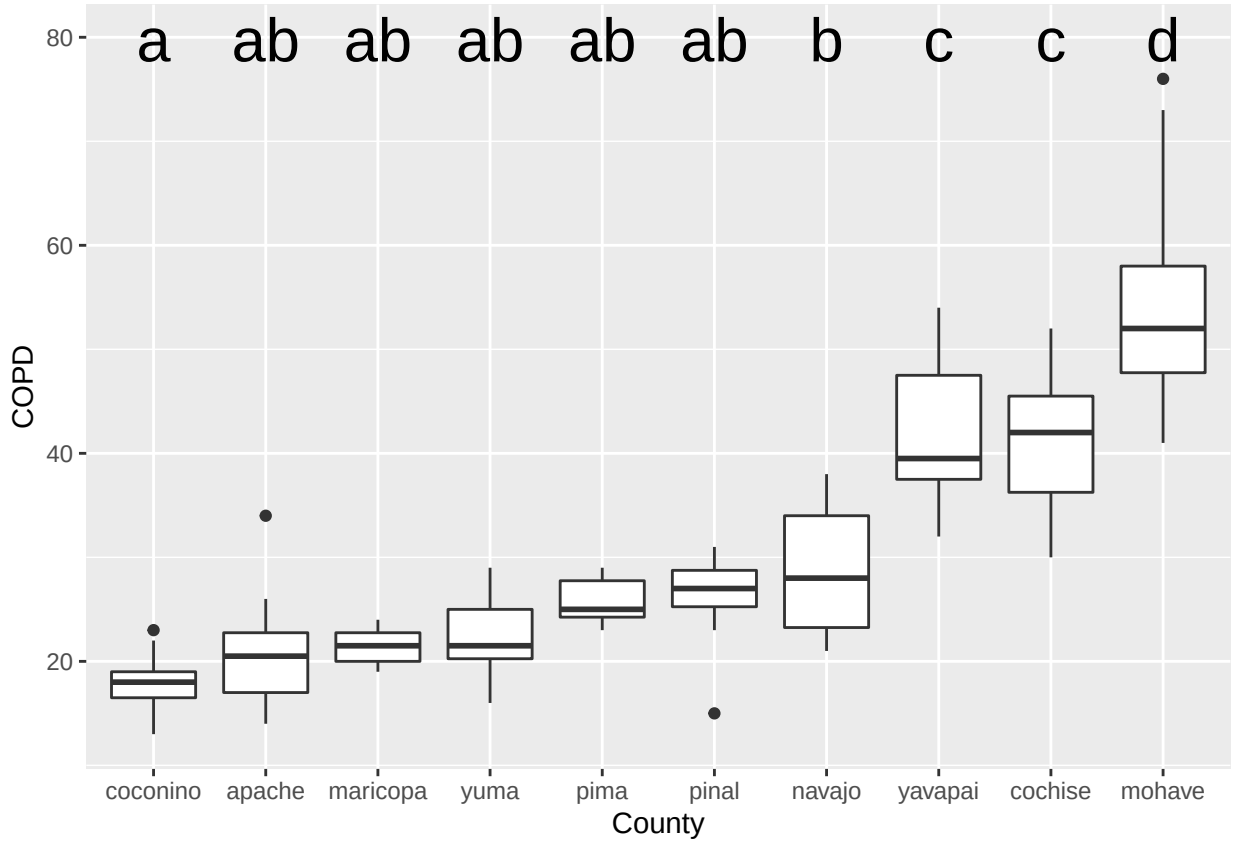
```
ggplot(all.data, aes(x=reorder(County, Asthma), y=Asthma)) +
  geom_boxplot() +
  labs(x = 'County', y = 'Asthma') +
  geom_text(data=letter.data.asthma, aes(x=County, y=y, label=.group), size=8)
```

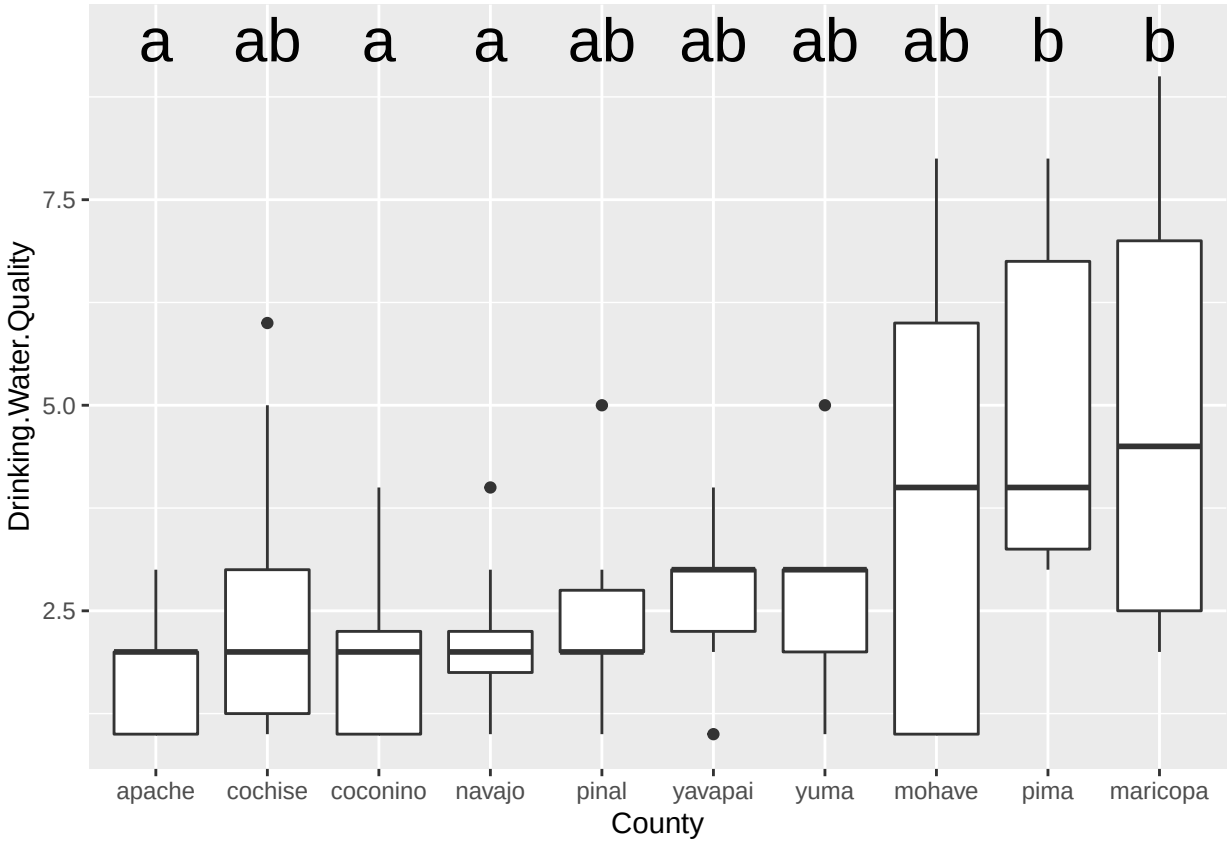


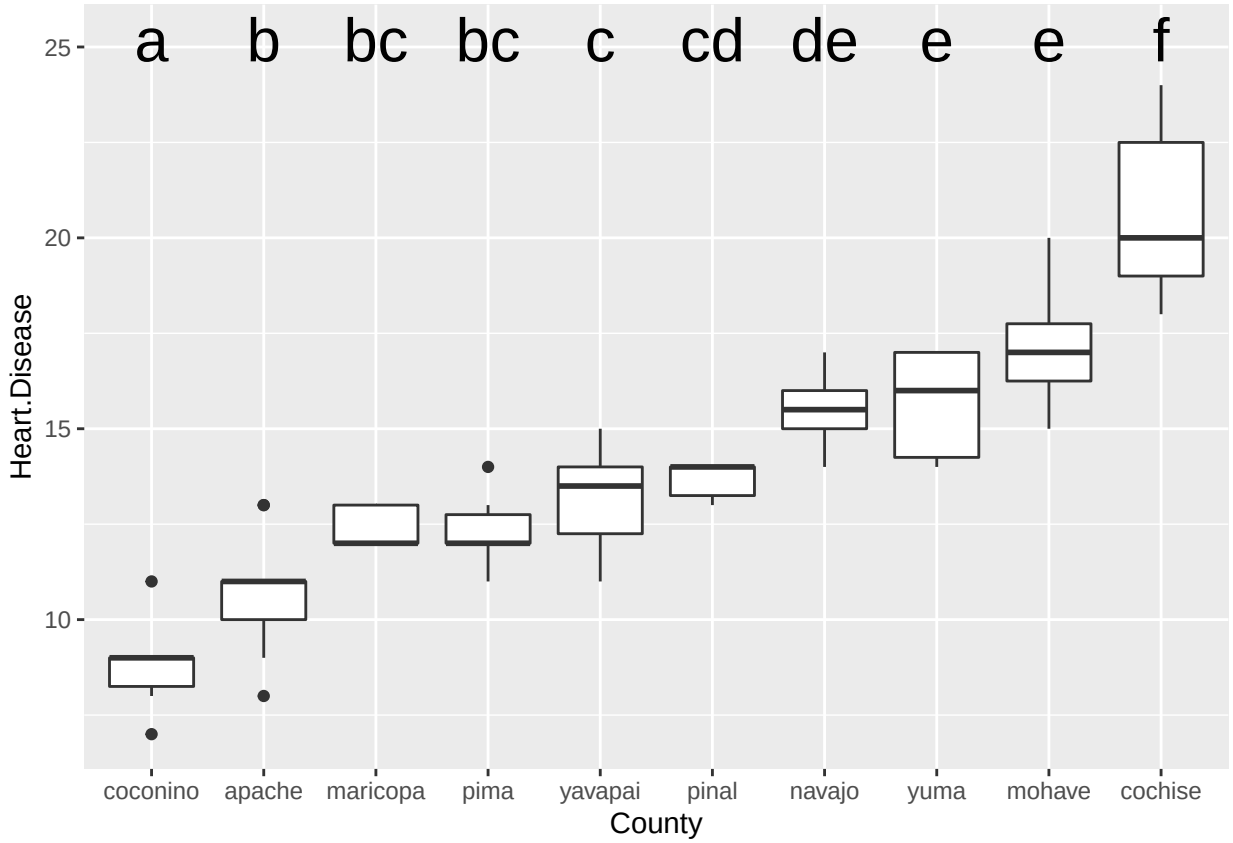

```
for(iter in 5:12)
{
  obj <- lm(paste(colnames(all.data)[iter], "~", 'County'), data=all.data)
  letter.data <- emmeans(obj, specs = ~ County) %>%
    multcomp::cld(Letters = letters, level = 0.95) %>%
    mutate(.group = str_remove_all(.group, '\\s')) %>%
    mutate(y = max(all.data[,iter], na.rm=T) + (max(all.data[,iter], na.rm=T) * .05))
  hold.data <- data.frame(
    County = all.data$County,
    Value = all.data[,iter]
  )
  hold.data <- hold.data %>% na.omit()
  print(ggplot(hold.data,
               aes(x=reorder(County, !!sym(colnames(hold.data)[2]), FUN = median),
                   y=!!sym(colnames(hold.data)[2]))) +
          geom_boxplot() +
          labs(x = "County") +
          geom_text(data = letter.data, aes(x=County, y=y, label=.group),
                    size = 8))
}
```

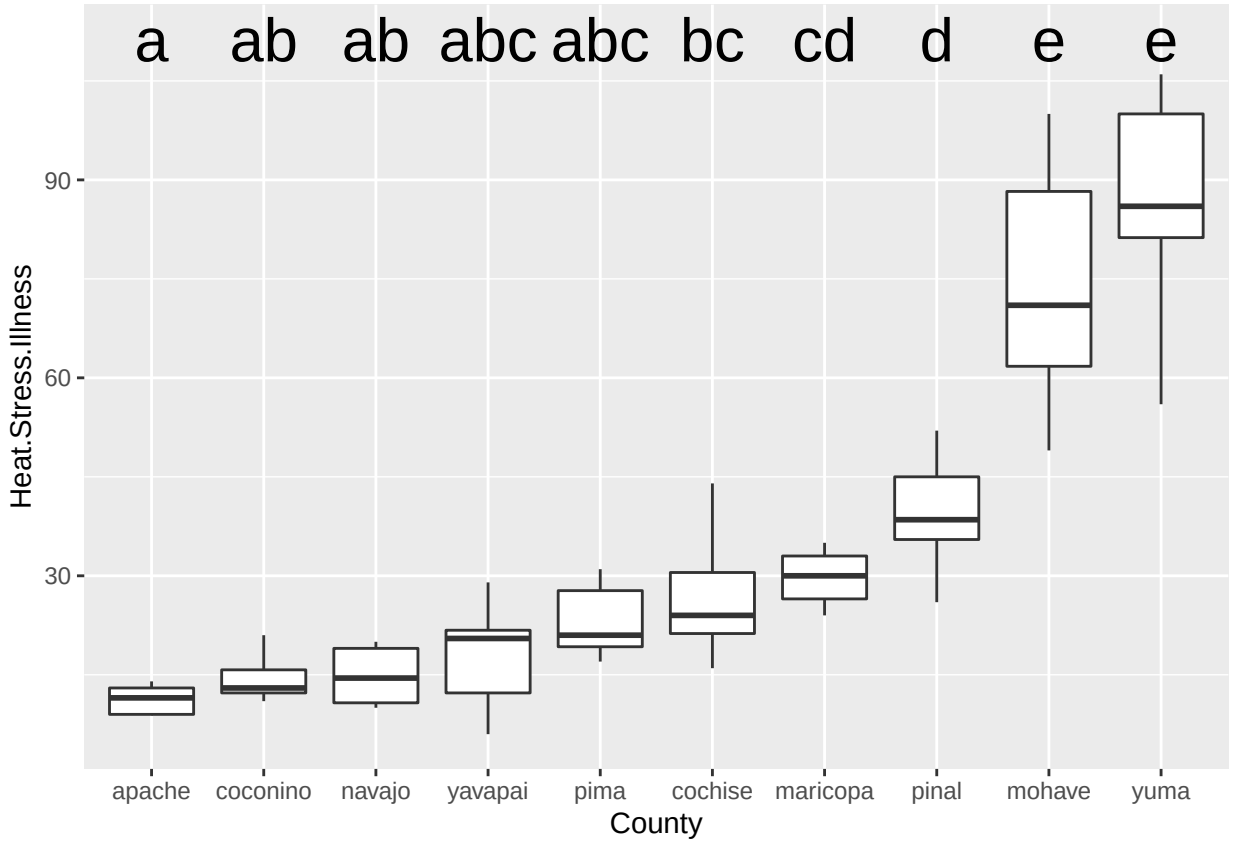

a
ab
bc
bcd
bcd
cd
cd
d
d
e
60
50
40
30
20
Asthma
coconino
navajo
pima
apache
pinal
yuma
yavapai
cochise
maricopa
mohave
County

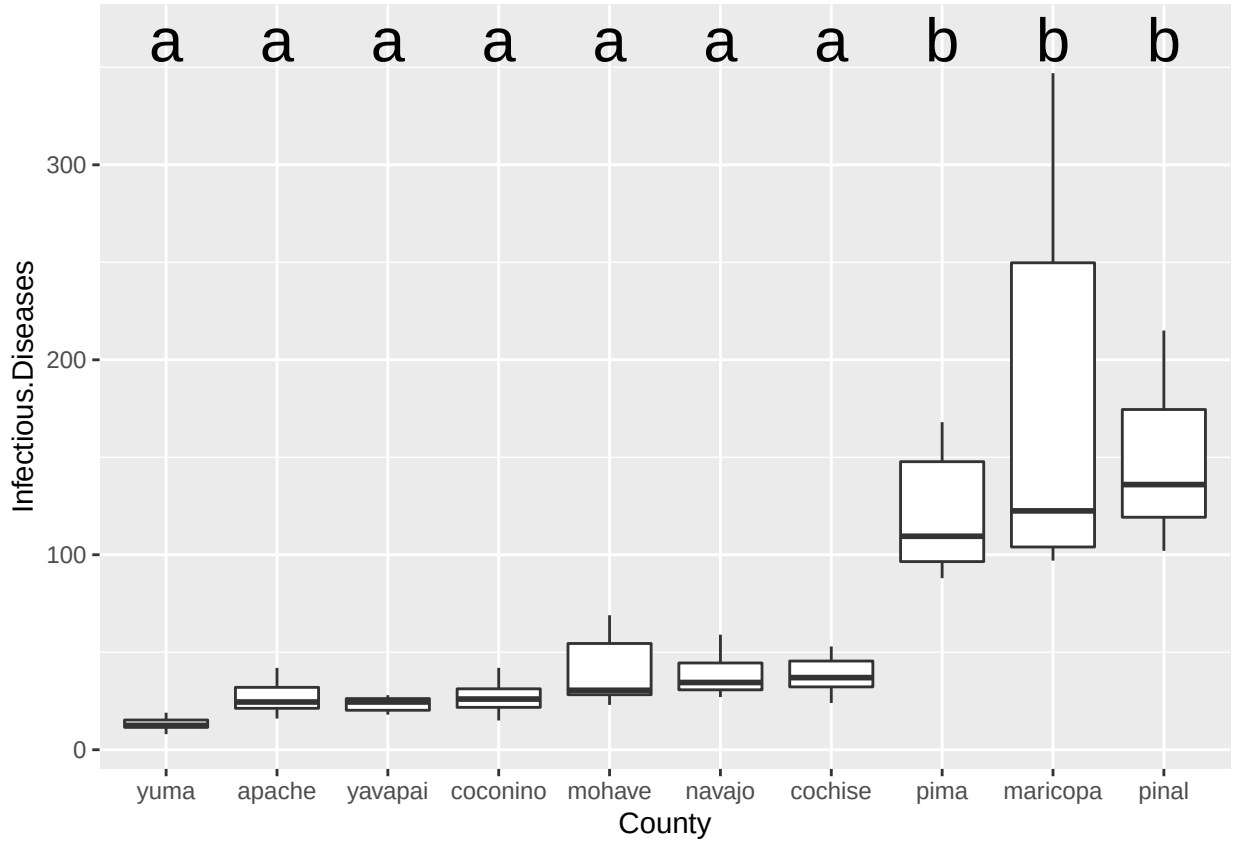

abcd
a
ab
abc
abcd
abcd
bcd
abcd
cd
d
15
10
5
0
Carbon.Monoxide.Poisoning
apache
maricopa
pima
pinal
mohave
yuma
yavapai
cochise
navajo
coconino
County

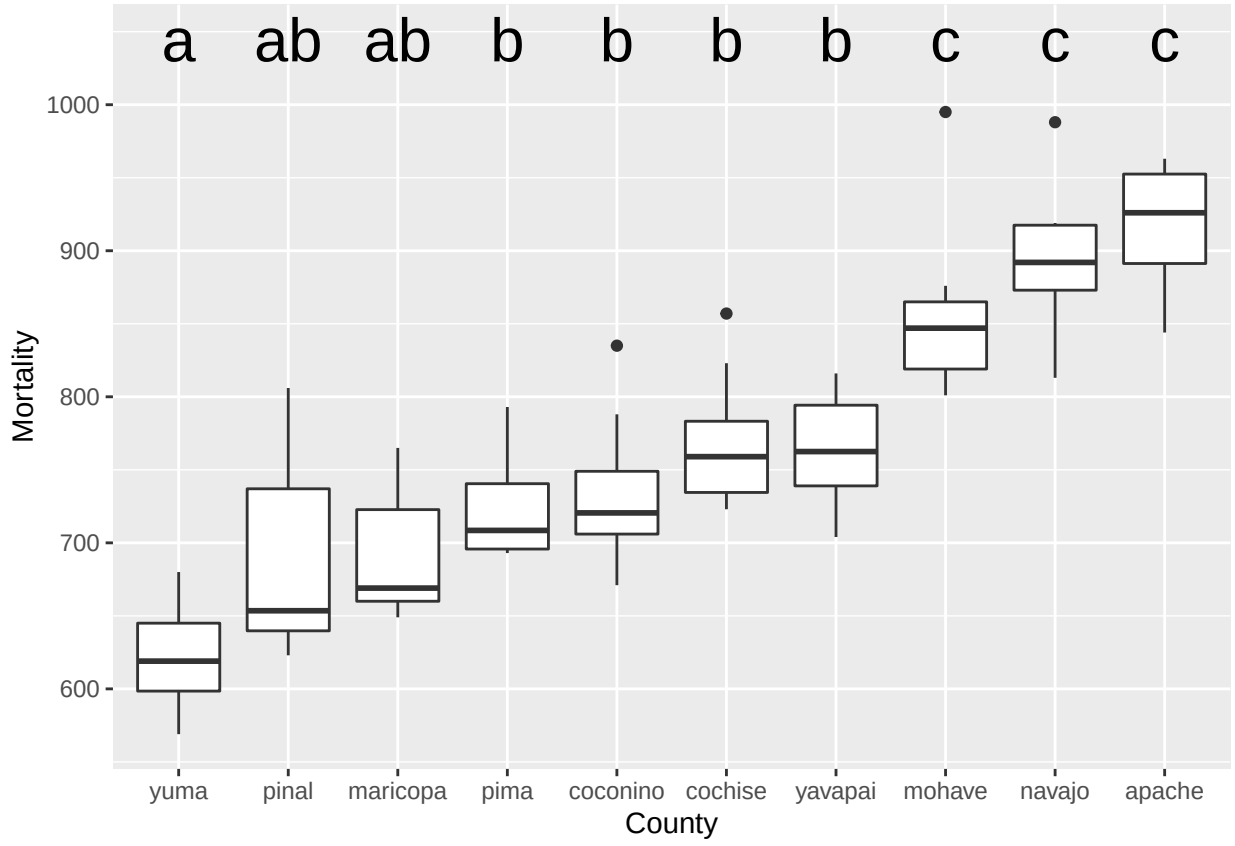

a
ab
ab
ab
ab
ab
b
c
c
d
80
60
40
20
COPD
coconino
apache
maricopa
yuma
pima
pinal
navajo
yavapai
cochise
mohave
County


a
ab
a
a
ab
ab
ab
ab
b
b
7.5
5.0
2.5
Drinking.Water.Quality
apache
cochise
coconino
navajo
pinal
yavapai
yuma
mohave
pima
maricopa
County


a
b
bc
bc
c
cd
de
e
e
f
25
20
15
10
Heart.Disease
coconino
apache
maricopa
pima
yavapai
pinal
navajo
yuma
mohave
cochise
County

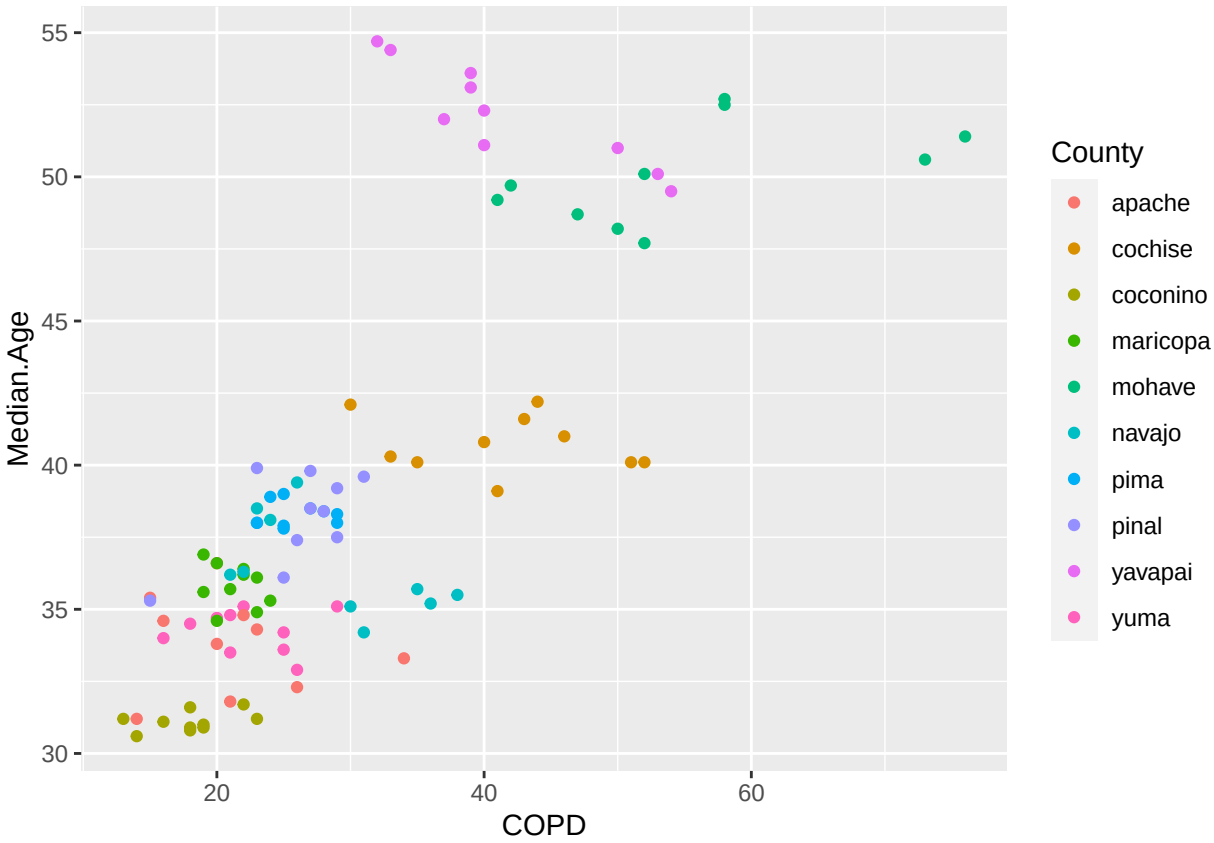

a
ab
ab
abc
abc
bc
cd
d
e
e
90
60
30
Heat.Stress.Illness
apache
coconino
navajo
yavapai
pima
cochise
maricopa
pinal
mohave
yuma
County


a a a a a a a b b b
300
200
100
0
Infectious.Diseases
yuma
apache
yavapai
coconino
mohave
navajo
cochise
pima
maricopa
pinal
County



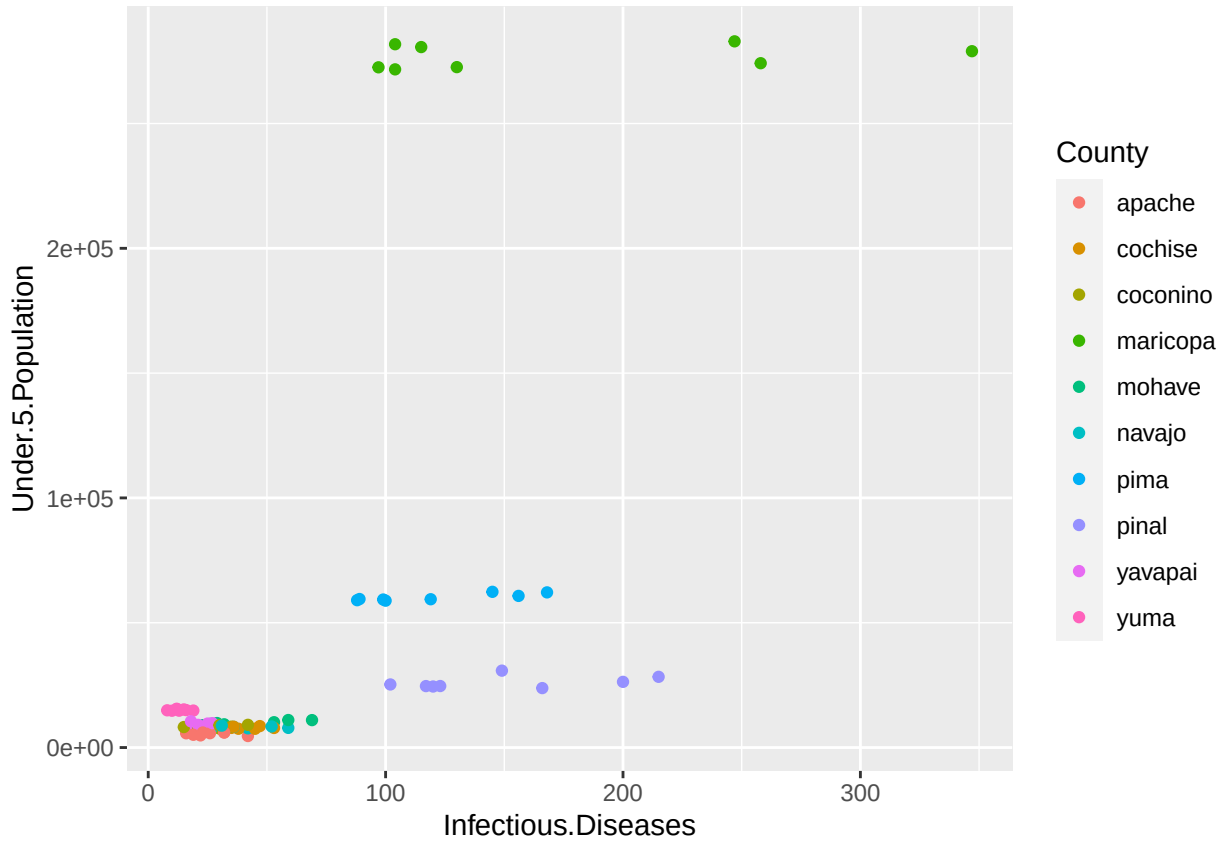

```
final.p.value <- adj.cor %>%
  group_by(Content.Area) %>% slice(which.min(adj.p.value)) %>%
  arrange(adj.p.value)
```

```
ggplot(all.data, aes(x=COPD, y=Median.Age, color=County)) + geom_point()
```

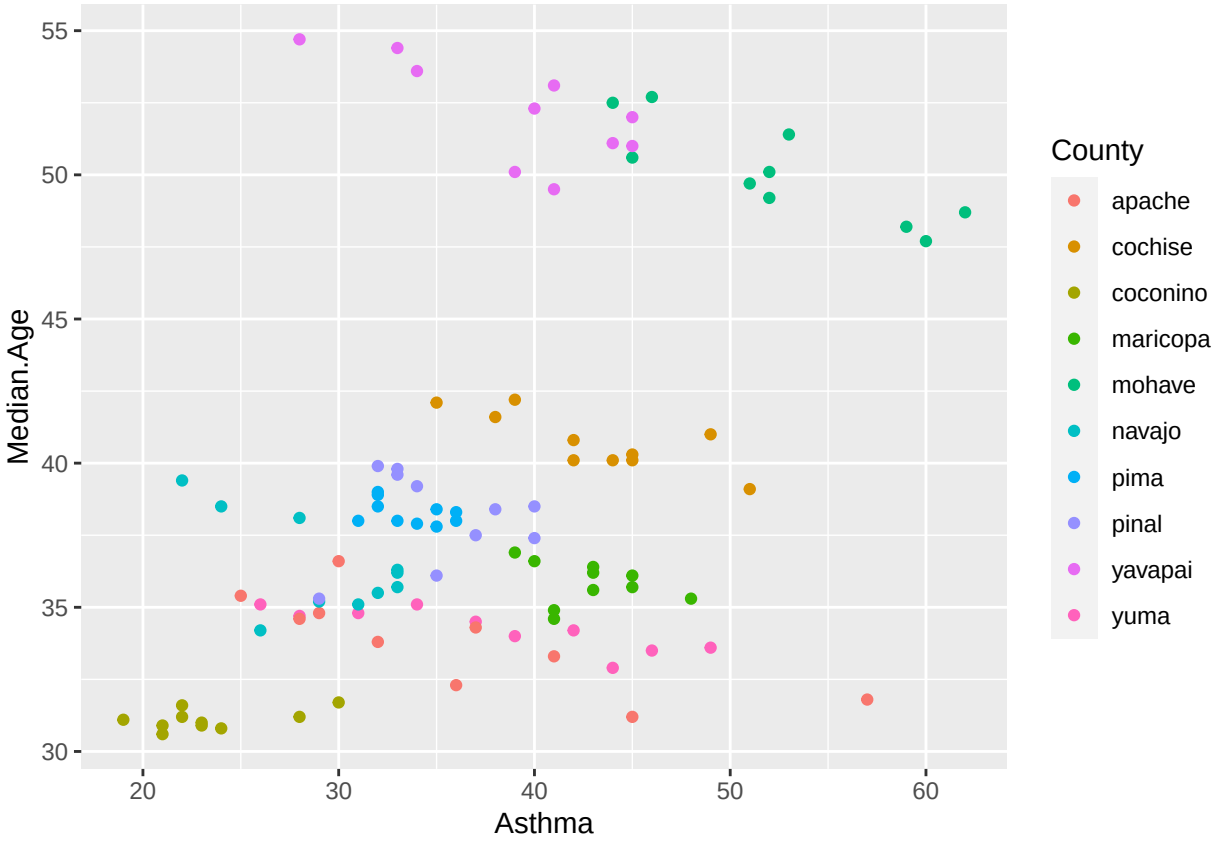



```
ggplot(all.data, aes(x=Infectious.Diseases, y=Under.5.Population, color=County)) +
  geom_point()
```

```
## Warning: Removed 20 rows containing missing values (geom_point).
```




```
ggplot(all.data, aes(x=Asthma, y=Median.Age, color=County)) + geom_point()
```

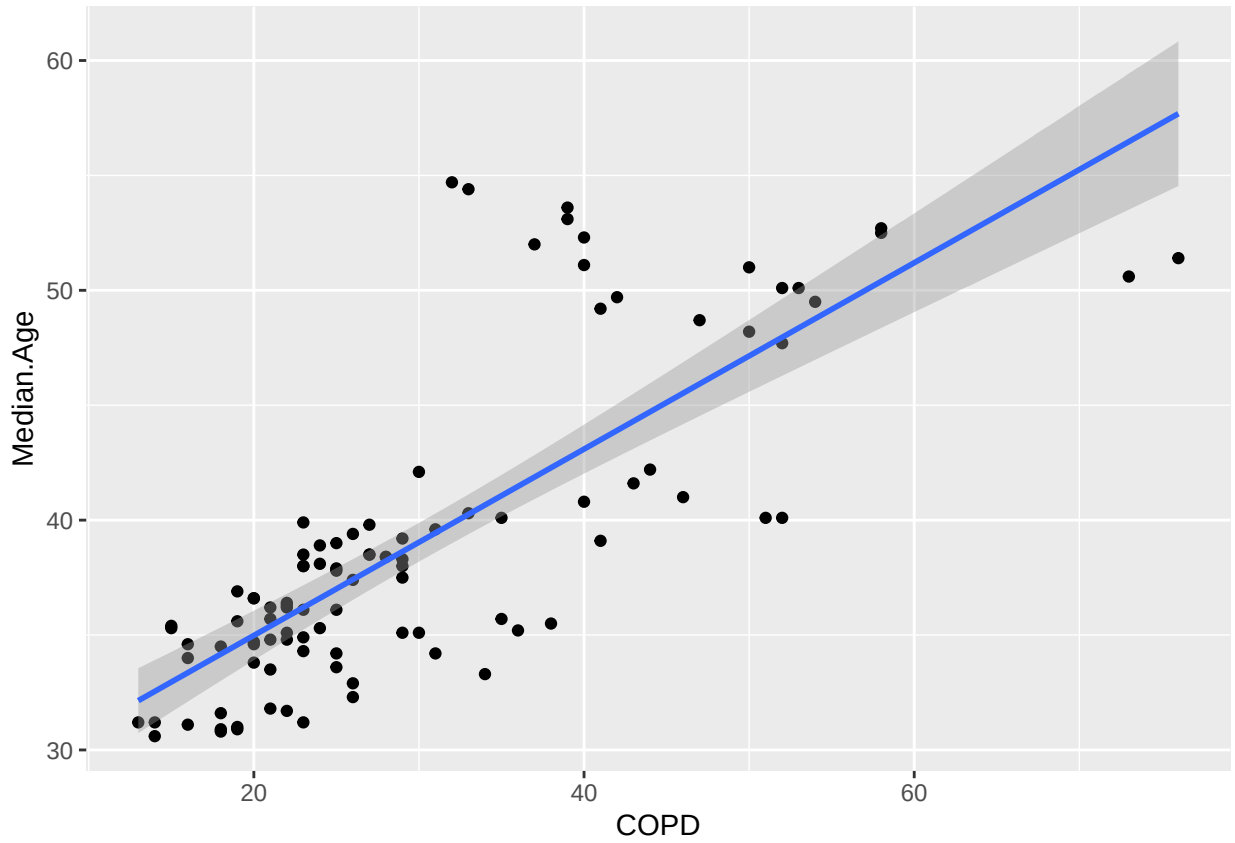



```
for (index in 1:6) {
  print(ggplot(all.data,
               aes_string(x=pull(final.p.value[index,1]),
                          y=pull(final.p.value[index,2]))) +
          geom_point() + geom_smooth(method="lm", model='y~x'))
}
```

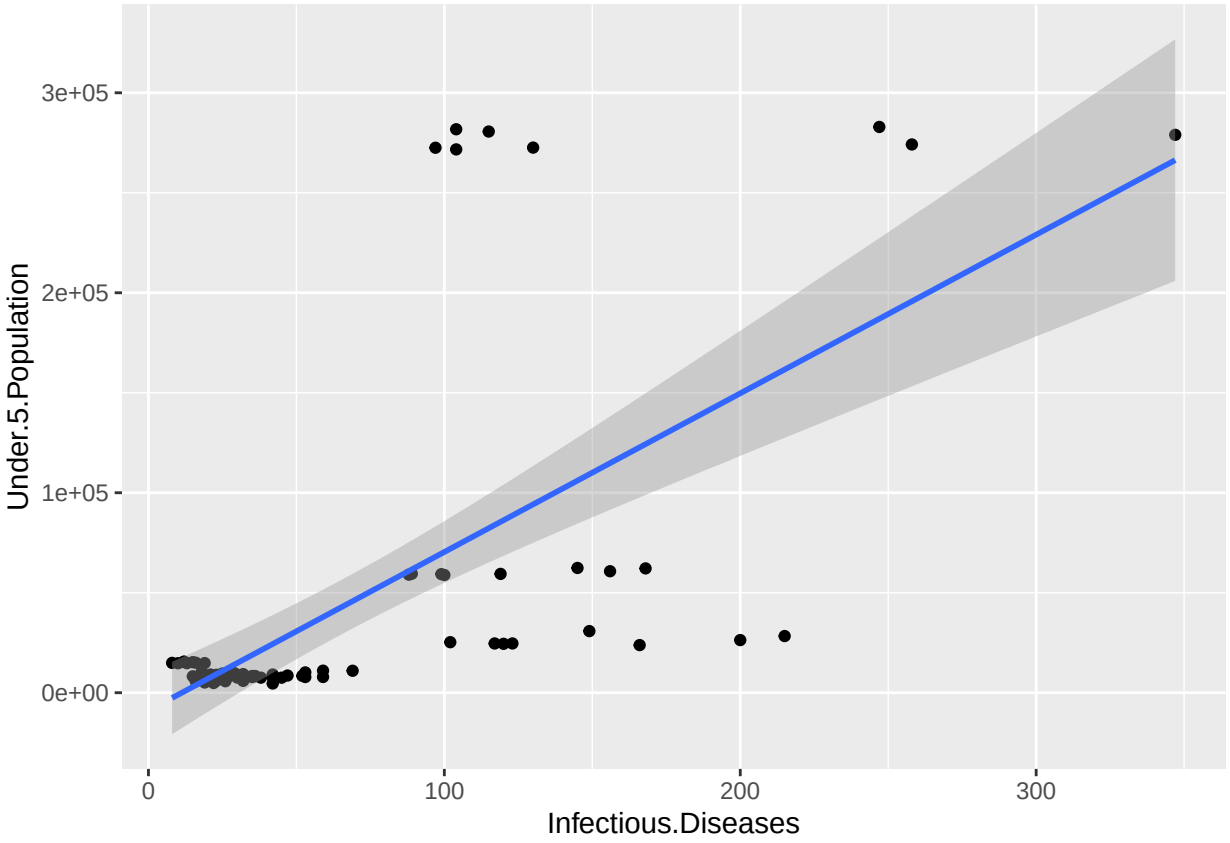
```
## Warning: Ignoring unknown parameters: model
```

```
## `geom_smooth()` using formula 'y ~ x'
```

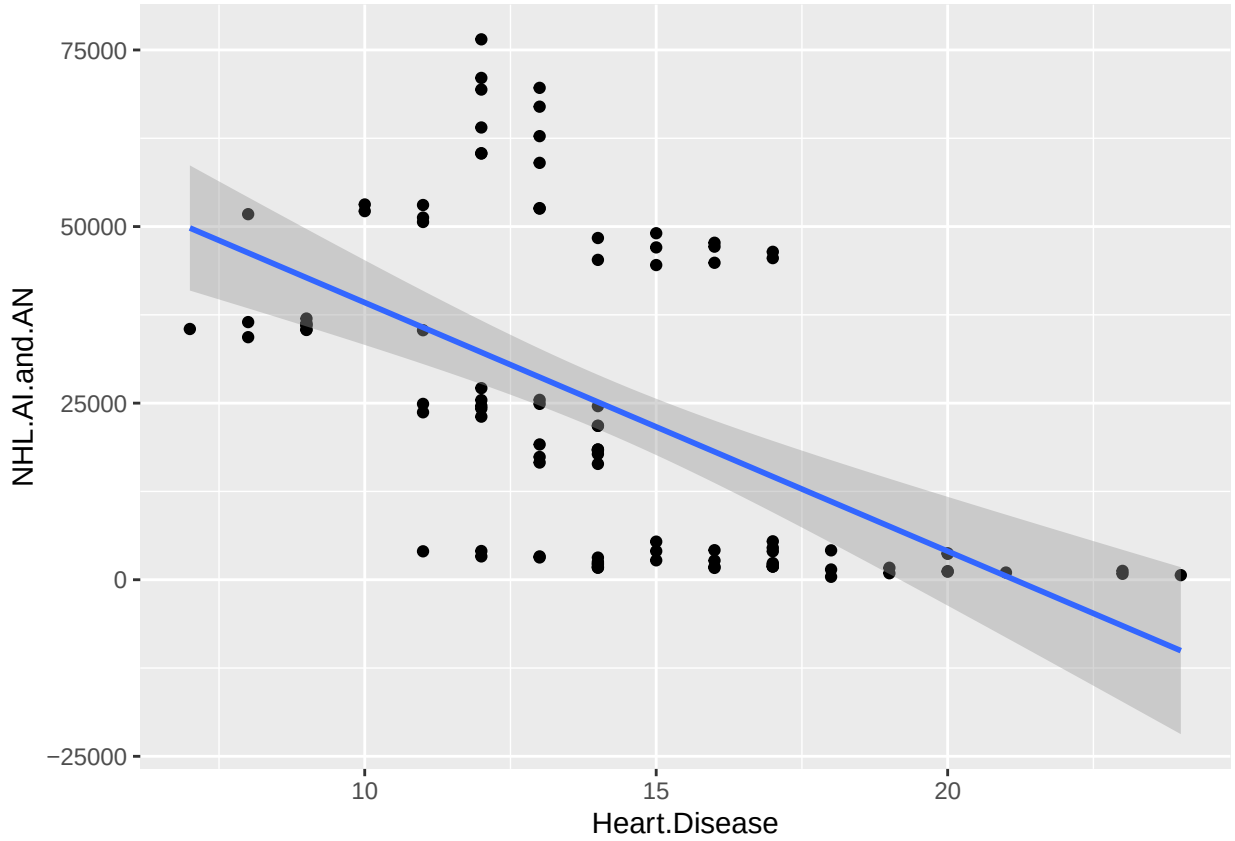
```
## Warning: Ignoring unknown parameters: model
```



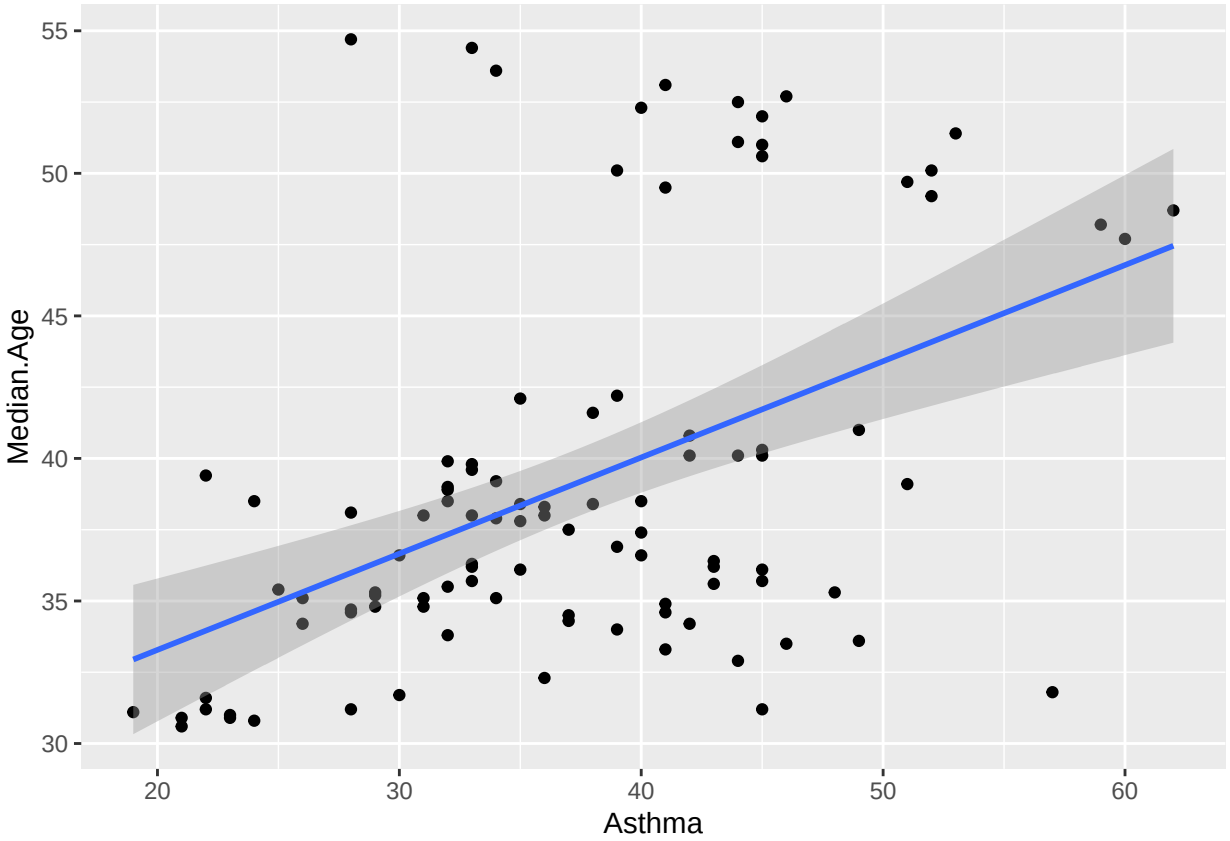

```
## `geom_smooth()` using formula 'y ~ x'

## Warning: Removed 20 rows containing non-finite values (stat_smooth).

## Warning: Removed 20 rows containing missing values (geom_point).

## Warning: Ignoring unknown parameters: model
```



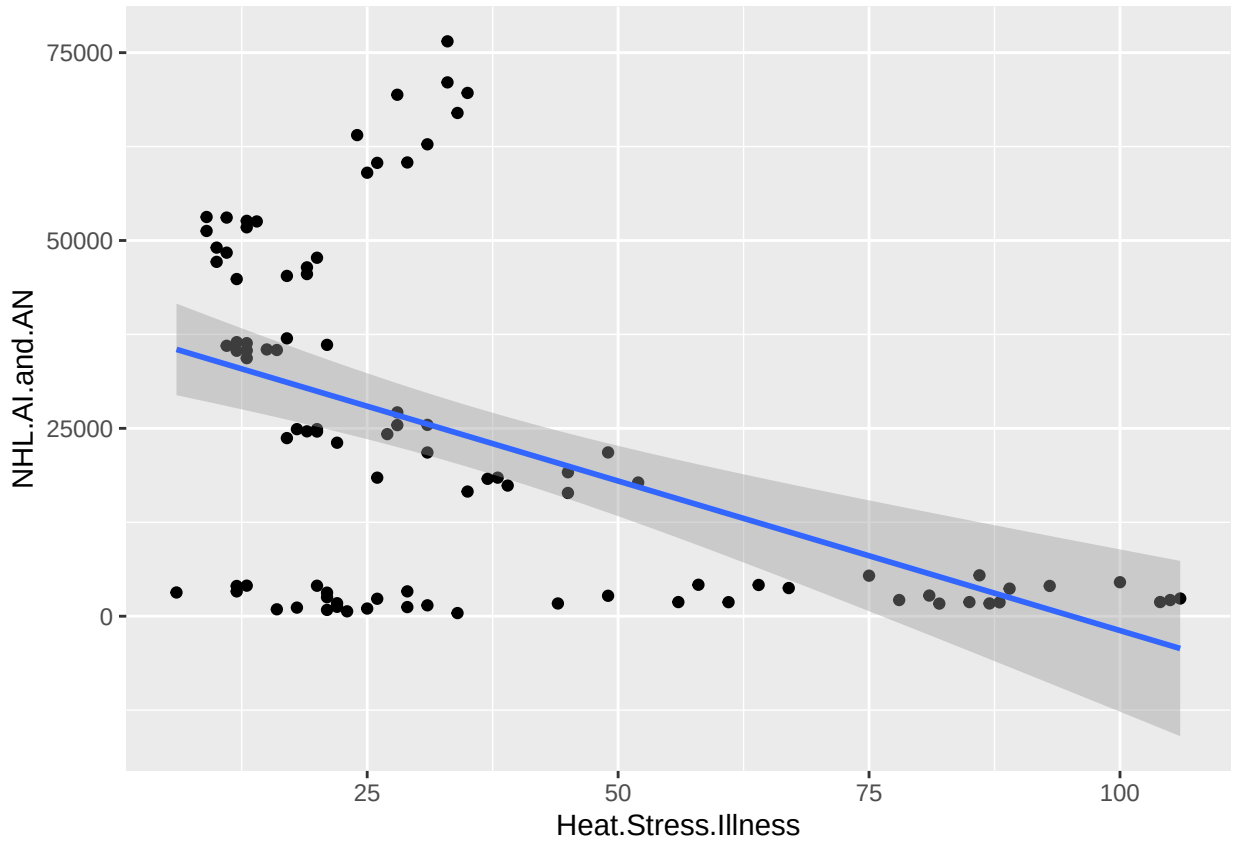

```
## `geom_smooth()` using formula 'y ~ x'

## Warning: Removed 2 rows containing non-finite values (stat_smooth).

## Warning: Removed 2 rows containing missing values (geom_point).

## Warning: Ignoring unknown parameters: model
```



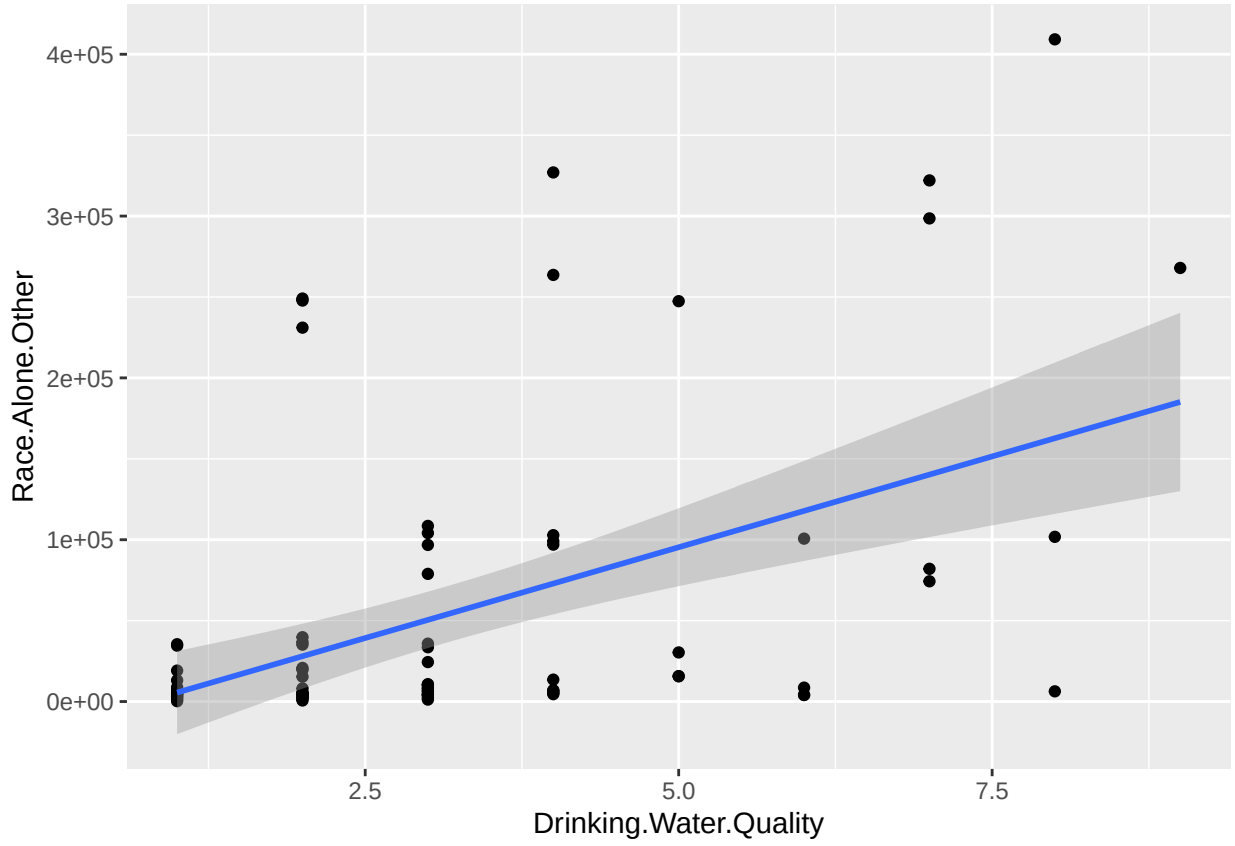

```
## `geom_smooth()` using formula 'y ~ x'

## Warning: Ignoring unknown parameters: model
```




```
## `geom_smooth()` using formula 'y ~ x'

## Warning: Removed 6 rows containing non-finite values (stat_smooth).

## Warning: Removed 6 rows containing missing values (geom_point).

## Warning: Ignoring unknown parameters: model
```




```
## `geom_smooth()` using formula 'y ~ x'

## Warning: Removed 14 rows containing non-finite values (stat_smooth).

## Warning: Removed 14 rows containing missing values (geom_point).
```



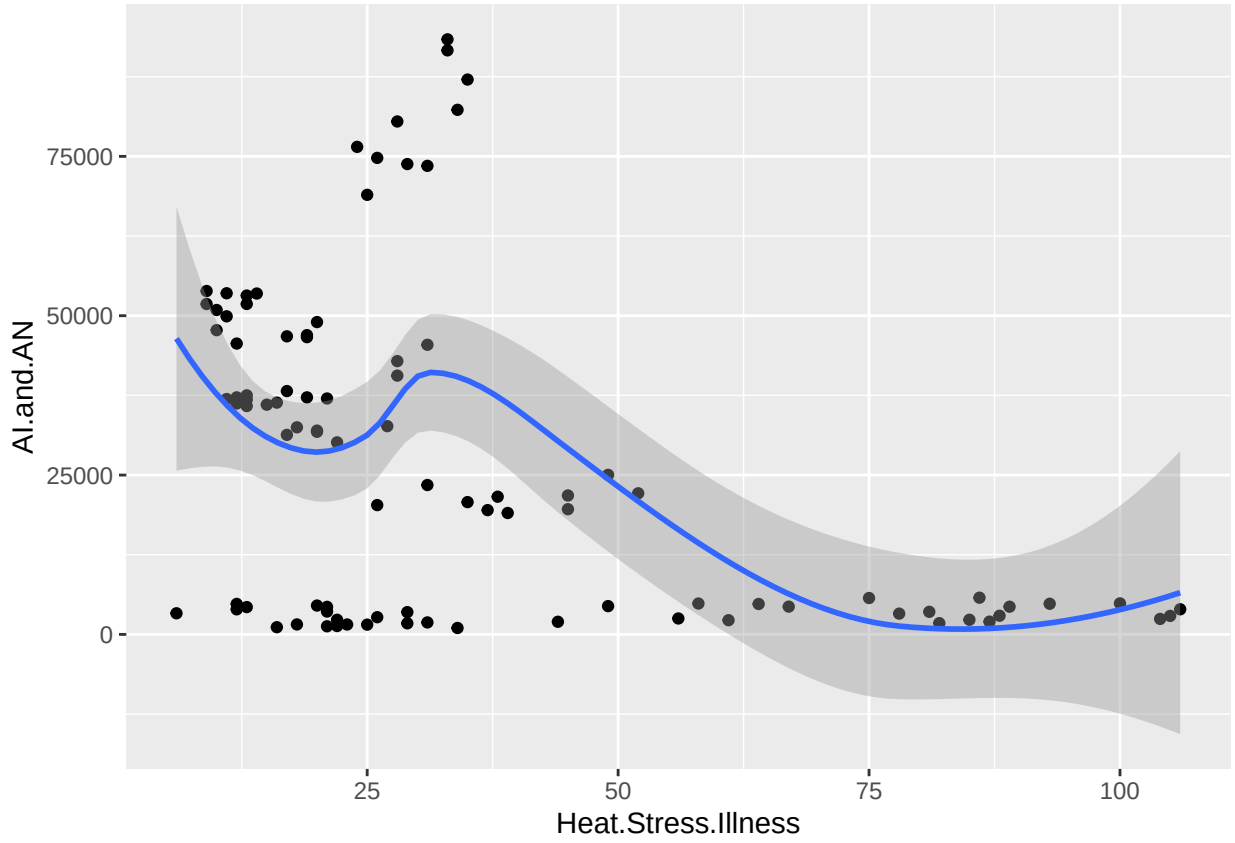

```
ggplot(all.data,
       aes(x=Heat.Stress.Illness, y=AI.and.AN)) +
  geom_point() + geom_smooth(method="loess", model='y~x')
```

```
## Warning: Ignoring unknown parameters: model
```

```
## `geom_smooth()` using formula 'y ~ x'
```

```
## Warning: Removed 6 rows containing non-finite values (stat_smooth).
```

```
## Warning: Removed 6 rows containing missing values (geom_point).
```




```
hospital.info <- str(hospital.wider)
write.table(hospital.info, file = "hospital.txt", sep = ",", quote = FALSE, row.names = T)
```

```
data.frame(variable = names(hospital.wider),
           class = sapply(hospital.wider, typeof),
           first.values = sapply(hospital.wider, function(x) paste0(head(x),
                                                                    collapse = ",")),
           row.names = NULL) %>%
  kable(caption = "100 obs. of 13 variables")
```

```
data.frame(variable = names(censusData),
           class = sapply(censusData, typeof),
           first.values = sapply(censusData, function(x) paste0(head(x),
                                                                collapse = ",")),
           row.names = NULL) %>%
  kable()
```

Table 2: 100 obs. of 13 variables

variable	class	first.values
County	integer	mohave,maricopa,cochise,yuma,pima,pinal
Year	character	2019,2019,2019,2019,2019,2019
Date	double	2019-04-10,2019-04-10,2019-04-10,2019-04-10,2019-04-10,20
County.Year	character	mohave2019,maricopa2019,cochise2019,yuma2019,pima201
Asthma	integer	46,39,39,34,32,32
Carbon Monoxide Poisoning	integer	NA,1,7,NA,1,2
Chronic Obstructive Pulmonary Disease (COPD)	integer	58,19,44,29,24,23
Drinking Water Quality	integer	1,2,2,3,3,NA
Food Safety	integer	NA,NA,NA,NA,NA,NA
Heart Disease	integer	20,12,18,14,12,14
Heat Stress Illness	integer	67,33,34,105,28,49
Infectious Diseases	integer	NA,NA,NA,NA,NA,NA
Mortality	integer	NA,NA,NA,NA,NA,NA

variable	class	first.values
Year	character	2019,2019,2019,2019,2019,2019
County	integer	apache,cochise,coconino,maricopa,mohave,navajo
Date	double	2019-04-10,2019-04-10,2019-04-10,2019-04-10,2019-04-10,2019-04-10
County.Year	character	apache2019,cochise2019,coconino2019,maricopa2019,mohave2019,navajo2019
Total.Population	double	71887,125922,143476,4485414,212181,110924
Male.Total.Population	double	36435,64204,71036,2217116,106919,54994
Female.Total.Population	double	35452,61718,72440,2268298,105262,55930
Sex.Ratio	double	102.8,104,98.1,97.7,101.6,98.3
Under.5.Population	double	4558,6855,6900,276119,8997,7448
5.to.9.Population	double	5138,7951,7749,283710,9216,8115
10.to.14.Population	double	6089,7437,8629,312364,11202,9016
15.to.19.Population	double	5631,7424,17156,299470,10381,7021
20.to.24.Population	double	4216,7822,18360,296675,9174,5838
25.to.34.Population	double	9316,15467,19997,658682,20109,12793
35.to.44.Population	double	7153,14188,15283,583814,19281,13028
45.to.54.Population	double	8286,13124,14245,555903,22856,12009
55.to.59.Population	double	4905,7725,9092,271428,13249,7727
60.to.64.Population	double	4992,8990,7479,250782,21889,7369
65.to.74.Population	double	7110,16326,11890,402314,37168,12703
75.to.84.Population	double	3551,9222,5699,221756,23634,6473
Over.85.Population	double	942,3391,997,72397,5025,1384
Median.Age	double	36.6,42.2,31.1,36.9,52.7,39.4
Under.18.Population	double	19100,27003,28005,1052439,35605,29122
Over.16.Population	double	54969,101567,118789,3553180,181081,84586
Over.18.Population	double	52787,98919,115471,3432975,176576,81802
Over.21.Population	double	49300,95081,97207,3254644,170466,78557
Over.62.Population	double	14961,34619,23241,842012,78366,24828
Over.65.Population	double	11603,28939,18586,696467,65827,20560
Over.18	double	52787,98919,115471,3432975,176576,81802
Over.18.Male	double	26304,50636,56074,1680405,89165,40681
Over.18.Female	double	26483,48283,59397,1752570,87411,41121
Over.18.Sex.Ratio	double	99.3,104.9,94.4,95.9,102,98.9
Over.65.Male	double	5433,14060,8595,313899,32587,9748
Over.65.Female	double	6170,14879,9991,382568,33240,10812
Over.65.Sex.Ration	double	88.1,94.5,86,82.1,98,90.2
One.Race.Total.Population	double	70612,120411,137246,4322940,204279,107609
Two.Or.More.Races.Population	double	1275,5511,6230,162474,7902,3315
White	double	15109,108177,91649,3547155,188756,51262
Black.Or.African.American	double	721,6163,1364,266128,1496,1309
AI.and.AN	double	53480,1006,37187,93358,4358,50892

```
data.frame(variable = names(all.data),
           class = sapply(all.data, typeof),
           first.values = sapply(all.data, function(x) paste0(head(x),
                                                              collapse = ",")),
           row.names = NULL) %>%
  kable()
```

variable	class	first.values
County	integer	mohave,maricopa,cochise,yuma,pima,pinal
Year	integer	2019,2019,2019,2019,2019,2019
Date	double	2019-04-10,2019-04-10,2019-04-10,2019-04-10,2019-04-10,2019-04-10
County.Year	character	mohave2019,maricopa2019,cochise2019,yuma2019,pima2019,pinal2019
Asthma	double	46,39,39,34,32,32
Carbon.Monoxide.Poisoning	integer	NA,1,7,NA,1,2
COPD	integer	58,19,44,29,24,23
Drinking.Water.Quality	integer	1,2,2,3,3,NA
Heart.Disease	integer	20,12,18,14,12,14
Heat.Stress.Illness	integer	67,33,34,105,28,49
Infectious.Diseases	integer	NA,NA,NA,NA,NA,NA
Mortality	integer	NA,NA,NA,NA,NA,NA
Total.Population	double	212181,4485414,125922,213787,1047279,462789
Male.Total.Population	double	106919,2217116,64204,110189,516110,241369
Female.Total.Population	double	105262,2268298,61718,103598,531169,221420
Sex.Ratio	double	101.6,97.7,104,106.4,97.2,109
Under.5.Population	double	8997,276119,6855,15099,57113,25490
5.to.9.Population	double	9216,283710,7951,14298,60599,28129
10.to.14.Population	double	11202,312364,7437,15550,61314,31894
15.to.19.Population	double	10381,299470,7424,14941,69026,28912
20.to.24.Population	double	9174,296675,7822,16812,88778,24346
25.to.34.Population	double	20109,658682,15467,30027,135885,60718
35.to.44.Population	double	19281,583814,14188,22967,120304,60986
45.to.54.Population	double	22856,555903,13124,21689,112541,50925
55.to.59.Population	double	13249,271428,7725,10460,63631,29151
60.to.64.Population	double	21889,250782,8990,11136,65492,26572
65.to.74.Population	double	37168,402314,16326,19820,119874,56517
75.to.84.Population	double	23634,221756,9222,16790,69655,30749
Over.85.Population	double	5025,72397,3391,4198,23067,8400
Median.Age	double	52.7,36.9,42.2,35.1,38.9,39.9
Under.18.Population	double	35605,1052439,27003,53571,215606,102573
Over.16.Population	double	181081,3553180,101567,164744,853555,370055
Over.18.Population	double	176576,3432975,98919,160216,831673,360216
Over.21.Population	double	170466,3254644,95081,150446,778821,343421
Over.62.Population	double	78366,842012,34619,47136,252031,111333
Over.65.Population	double	65827,696467,28939,40808,212596,95666
Over.18	double	176576,3432975,98919,160216,831673,360216
Over.18.Male	double	89165,1680405,50636,83103,405576,188926
Over.18.Female	double	87411,1752570,48283,77113,426097,171290
Over.18.Sex.Ratio	double	102,95.9,104.9,107.8,95.2,110.3
Over.65.Male	double	32587,313899,14060,19321,96614,45770
Over.65.Female	double	33240,382568,14879,21487,115982,49896
Over.65.Sex.Ration	double	98,82.1,94.5,89.9,83.3,91.7
One.Race.Total.Population	double	204279,4322940,120411,207267,987957,442591
Two.Or.More.Races.Population	double	7902,162474,5511,6520,59322,20198
White	double	188756,3547155,108177,189117,795391,366928
Black.Or.African.American	double	1496,266128,6163,3244,38343,21264
AI.and.AN	double	4358,93358,1006,2900,40603,25035
Cherokee.TG	double	NA,1294,NA,NA,0,NA
Chippewa.TG	double	NA,518,NA,NA,606,NA
Navajo.TG	double	NA,39306,NA,NA,2814,NA
Sioux.TG	double	NA,1130,NA,NA,504,NA
Asian	double	1923,192301,2795,2388,29276,6642
Asian.Indian	double	NA,63846,NA,173,4411,1482
Chinese	double	NA,33121,NA,61,7446,661
Filipino	double	NA,34749,NA,928,6429,2228
Japanese	double	NA,6943,NA,405,1293,321

```
content.area.p %>%
  mutate_if(is.numeric, funs(as.character(signif(., 3)))) %>%
  kable(.)
```

```
## Warning: `funs()` was deprecated in dplyr 0.8.0.
## Please use a list of either functions or lambdas:
##
##   # Simple named list:
##   list(mean = mean, median = median)
##
##   # Auto named with `tibble::lst()`:
##   tibble::lst(mean, median)
##
##   # Using lambdas
##   list(~ mean(., trim = .2), ~ median(., na.rm = TRUE))
## This warning is displayed once every 8 hours.
## Call `lifecycle::last_lifecycle_warnings()` to see where this warning was generated.
```

	County	p.value
value5	Heat.Stress.Illness	2.04e-38
value4	Heart.Disease	2.47e-36
value2	COPD	1.2e-26
value	Asthma	8.02e-19
value7	Mortality	9.23e-19
value6	Infectious.Diseases	2.13e-17
value1	Carbon.Monoxide.Poisoning	1.05e-06
value3	Drinking.Water.Quality	0.000333

```
adj.cor[adj.cor$Content.Area == "Heat.Stress.Illness",] %>% kable()
```

	Content.Area	Demographic	p.value	corr	adj.p.value
cor487	Heat.Stress.Illness	NHL.AI.and.AN	1.7e-06	-0.4709142	0.0011076
cor450	Heat.Stress.Illness	AI.and.AN	1.0e-05	-0.4382536	0.0066211

```
adj.cor[adj.cor$Content.Area == "Heart.Disease",] %>%
  mutate_if(is.numeric, funs(as.character(signif(., 3)))) %>% kable()
```

	Content.Area	Demographic	p.value	corr	adj.p.value
cor404	Heart.Disease	NHL.AI.and.AN	1.4e-08	-0.535	9.33e-06
cor367	Heart.Disease	AI.and.AN	3.05e-08	-0.524	2.03e-05
cor392	Heart.Disease	Race.Alone.AI.and.AN	5.46e-07	-0.476	0.000362
cor349	Heart.Disease	Median.Age	3.17e-05	0.403	0.0211
cor362	Heart.Disease	Over.65.Sex.Ration	7.85e-06	0.718	0.00521

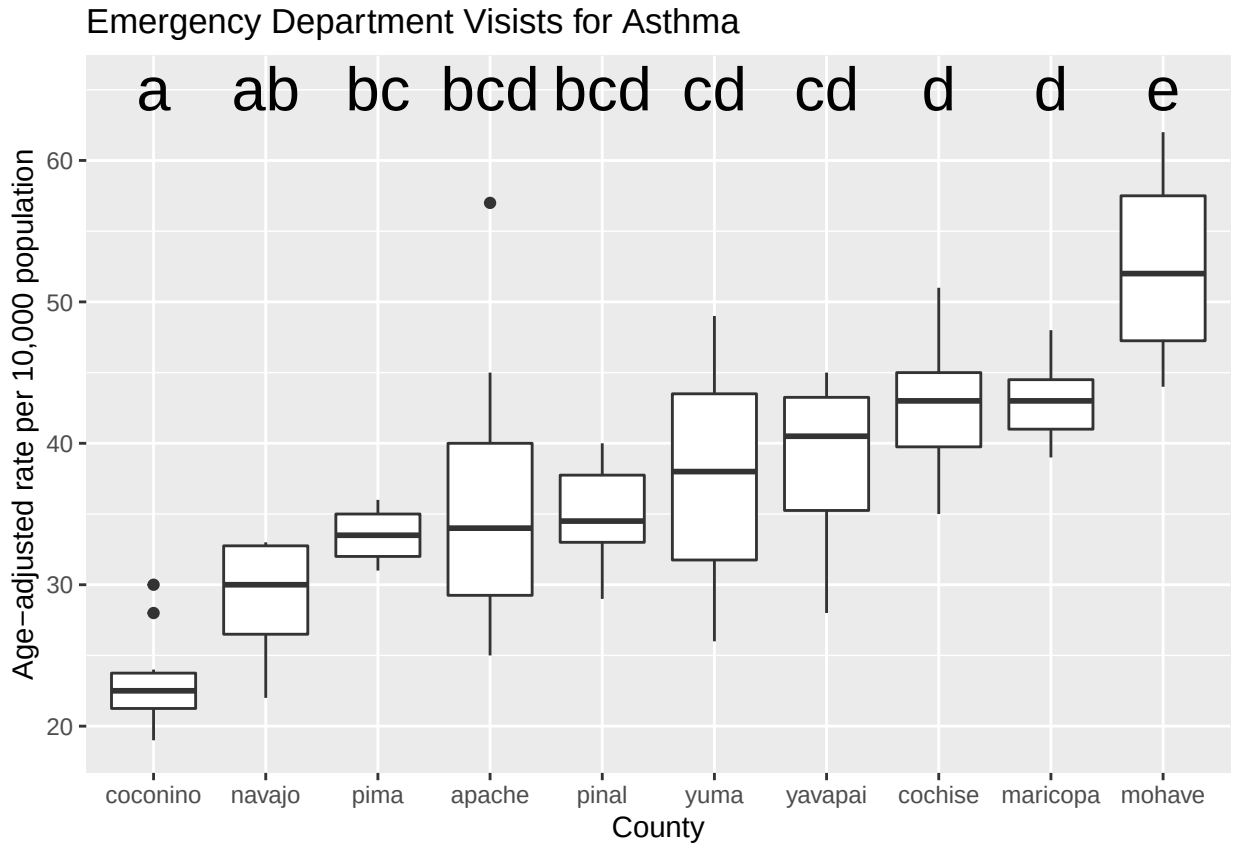
```
adj.cor[adj.cor$Content.Area == "COPD",] %>%
  mutate_if(is.numeric, funs(as.character(signif(., 3)))) %>% kable()
```

	Content.Area	Demographic	p.value	corr	adj.p.value
cor238	COPD	NHL.AI.and.AN	2.88e-09	-0.556	1.91e-06
cor201	COPD	AI.and.AN	5.1e-09	-0.548	3.39e-06
cor226	COPD	Race.Alone.AI.and.AN	1.04e-07	-0.502	6.93e-05
cor196	COPD	Over.65.Sex.Ration	1e-05	0.712	0.00667
cor183	COPD	Median.Age	2.63e-21	0.776	1.75e-18

```
adj.cor[adj.cor$Content.Area == "Asthma",] %>%
  mutate_if(is.numeric, funs(as.character(signif(., 3)))) %>% kable()
```

	Content.Area	Demographic	p.value	corr	adj.p.value
cor17	Asthma	Median.Age	1.02e-06	0.466	0.00068
cor46	Asthma	Vietnamese	7.41e-05	0.585	0.0492

```
obj <- lm(Asthma ~ County, data=all.data)
letter.data <- emmeans(obj, specs = ~ County) %>%
  multcomp::cld(Letters = letters, level = 0.95) %>%
  mutate(.group = str_remove_all(.group, '\\s')) %>%
  mutate(y = max(all.data$Asthma, na.rm=T) +
           (max(all.data$Asthma, na.rm=T) * .05))
hold.data <- data.frame(
  County = all.data$County,
  Value = all.data$Asthma
)
hold.data <- hold.data %>% na.omit()
ggplot(hold.data,
               aes(x=reorder(County, Value, FUN = median),
                   y=Value)) +
          geom_boxplot() +
          labs(x = "County",
               y = "Age-adjusted rate per 10,000 population",
               title = "Emergency Department Visists for Asthma") +
          geom_text(data = letter.data, aes(x=County, y=y, label=.group),
                    size = 8)
```



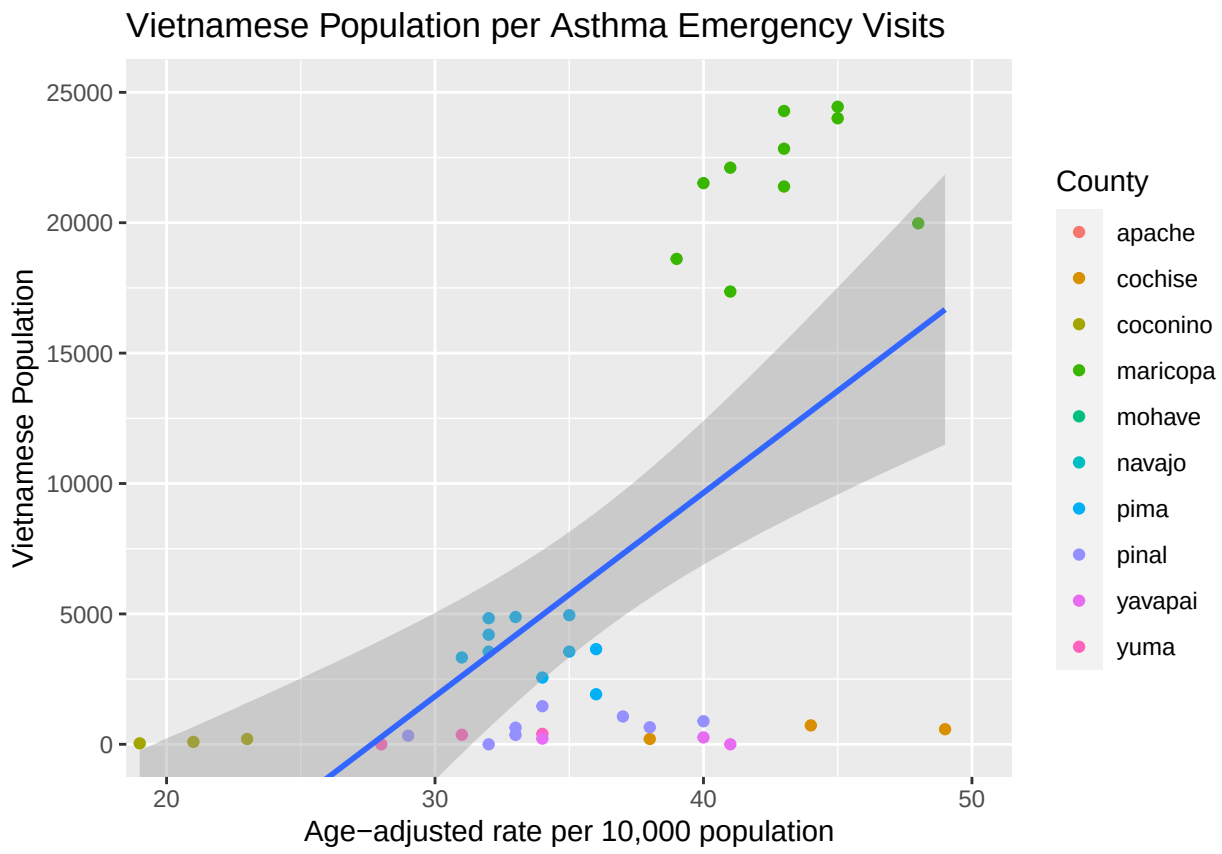

```
ggplot(all.data,
               aes(x=Asthma,
                           y=Vietnamese)) +
          geom_point(aes(color=County)) + geom_smooth(method="lm", model='y~x') +
  labs(x="Age-adjusted rate per 10,000 population",
       y="Vietnamese Population",
       title="Vietnamese Population per Asthma Emergency Visits") +
  coord_cartesian(xlim = c(20,50),
                  ylim = c(0,25000))
```

```
## Warning: Ignoring unknown parameters: model
```

```
## `geom_smooth()` using formula 'y ~ x'
```

```
## Warning: Removed 60 rows containing non-finite values (stat_smooth).
```

```
## Warning: Removed 60 rows containing missing values (geom_point).
```


Vietnamese Population per Asthma Emergency Visits
Vietnamese Population
25000
20000
15000
10000
5000
0
20
30
40
50
Age−adjusted rate per 10,000 population
County
apache
cochise
coconino
maricopa
mohave
navajo
pima
pinal
yavapai
yuma