

Week 9 Assignment

Jessalyn Chuang

#(1) Logistic Regression Download the dataset `jrn_lizard.csv` from Canvas. This dataset contains data on lizards captured using pitfall traps by researchers at the Jornada Basin Long Term Ecological Research site in southern New Mexico. For this portion of the assignment, you must build a logistic regression model to investigate the following research question:

Do snout-to-vent length (`SV_length`), sex (`sex`), and vegetation zone (`zone`) at time of capture significantly predict if a lizard tail is recorded as whole (`tail = W`)?

Before performing this regression, filter the dataset only for side-blotched lizards (*Uta stansburiana* or `spp = UTST`). Also, for context, the four possible vegetation zones are black grama grassland (`zone = G`), creosotebush shrubland (`zone = C`), mesquite duneland (`zone = M`), and tarbush shrubland (`zone = T`).

Using the workflow in today's lab as a guide, work through the steps of performing a logistic regression using these data, including exploratory data visualizations. You should report the results of your final model in text form (1-2 sentences) as well as provide 2 predictive figures showing the change in probability of a lizard tail being whole as lizard length increases, with one including lines colored by sex and another including lines colored by vegetation zone.

```
#install.packages("naniar", dependencies = TRUE)
#install.packages("DHARMa")
#install.packages("gtsummary")
```

```
library(here)
```

```
## here() starts at /home/guest/Statistical_Modeling_Sp25
```

```
library(tidyverse)
```

```
## -- Attaching core tidyverse packages ----- tidyverse 2.0.0 --
## v dplyr      1.1.4      v readr      2.1.5
## v forcats    1.0.0      v stringr    1.5.1
## v ggplot2    3.5.1      v tibble     3.2.1
## v lubridate  1.9.3      v tidyr      1.3.1
## v purrr      1.0.2
```

```
## -- 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
```

```
library(naniar)
library(DHARMa)
```

```
## Warning in check_dep_version(): ABI version mismatch:
## lme4 was built with Matrix ABI version 2
## Current Matrix ABI version is 1
## Please re-install lme4 from source or restore original 'Matrix' package
```

```
## This is DHARMA 0.4.7. For overview type '?DHARMA'. For recent changes, type news(package = 'DHARMA')
```

```
library(gtsummary)
library(broom)

lizard_data <- read_csv("./Assignments/week9/jrn_lizard.csv")
```

```
## Rows: 4091 Columns: 14
## -- Column specification -----
## Delimiter: ","
## chr (7): zone, site, plot, spp, sex, rcap, tail
## dbl (6): pit, toe_num, SV_length, total_length, weight, pc
## date (1): date
##
## i Use 'spec()' to retrieve the full column specification for this data.
## i Specify the column types or set 'show_col_types = FALSE' to quiet this message.
```

```
utst_lizards <- lizard_data %>%
  filter(spp == "UTST") %>%
  filter(zone %in% c('G', 'C', 'M', 'T')) %>%
  select(SV_length, zone, sex, tail)
utst_lizards_clean <- na.omit(utst_lizards)
```

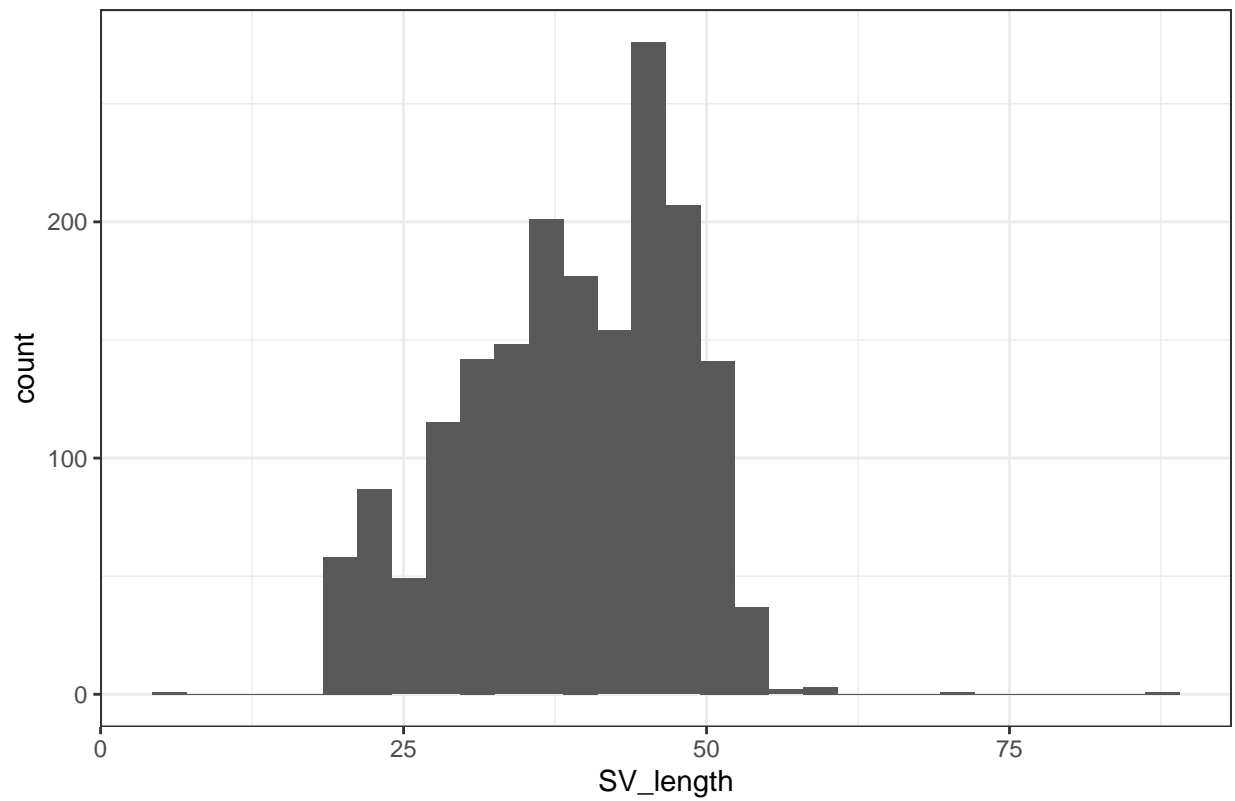
```
## Visualizing Data
```

```
hist_SV_length <- ggplot(utst_lizards_clean, aes(x = SV_length)) +
  geom_histogram() +
  theme_bw() +
  ggtitle("Count of lizards for each Snout-to-Vent Length")

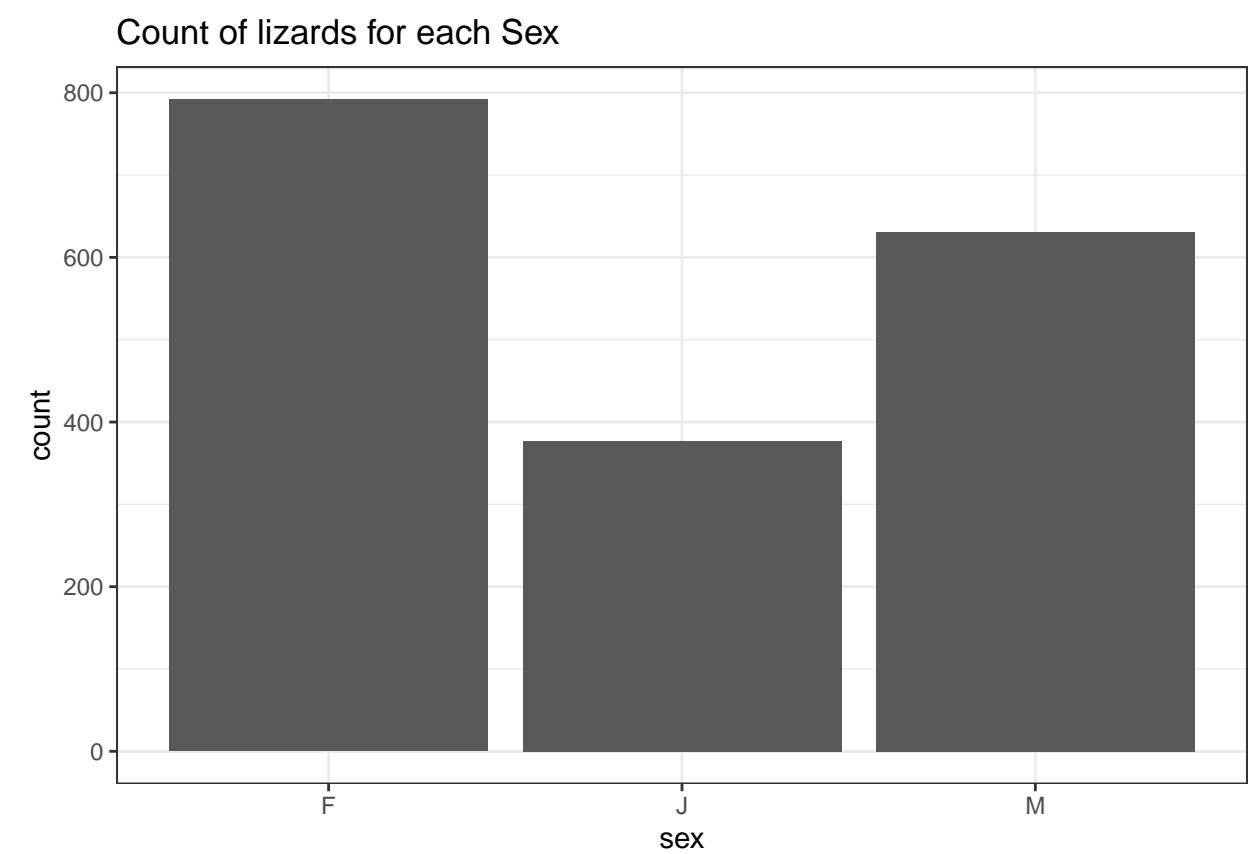
hist_SV_length
```

```
## 'stat_bin()' using 'bins = 30'. Pick better value with 'binwidth'.
```

Count of lizards for each Snout-to-Vent Length

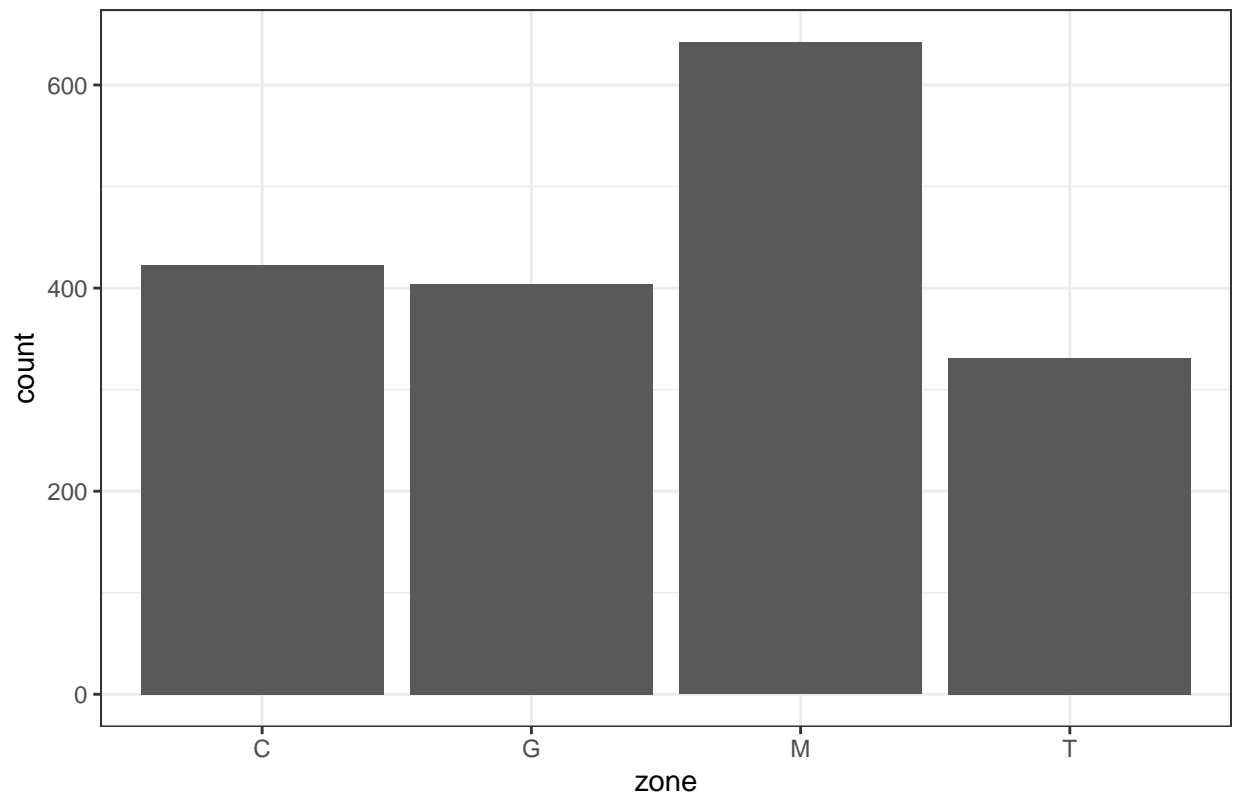


```
bar_sex <- ggplot(utst_lizards_clean, aes(x = sex)) +  
  geom_bar() +  
  theme_bw() +  
  ggtitle("Count of lizards for each Sex")  
bar_sex
```



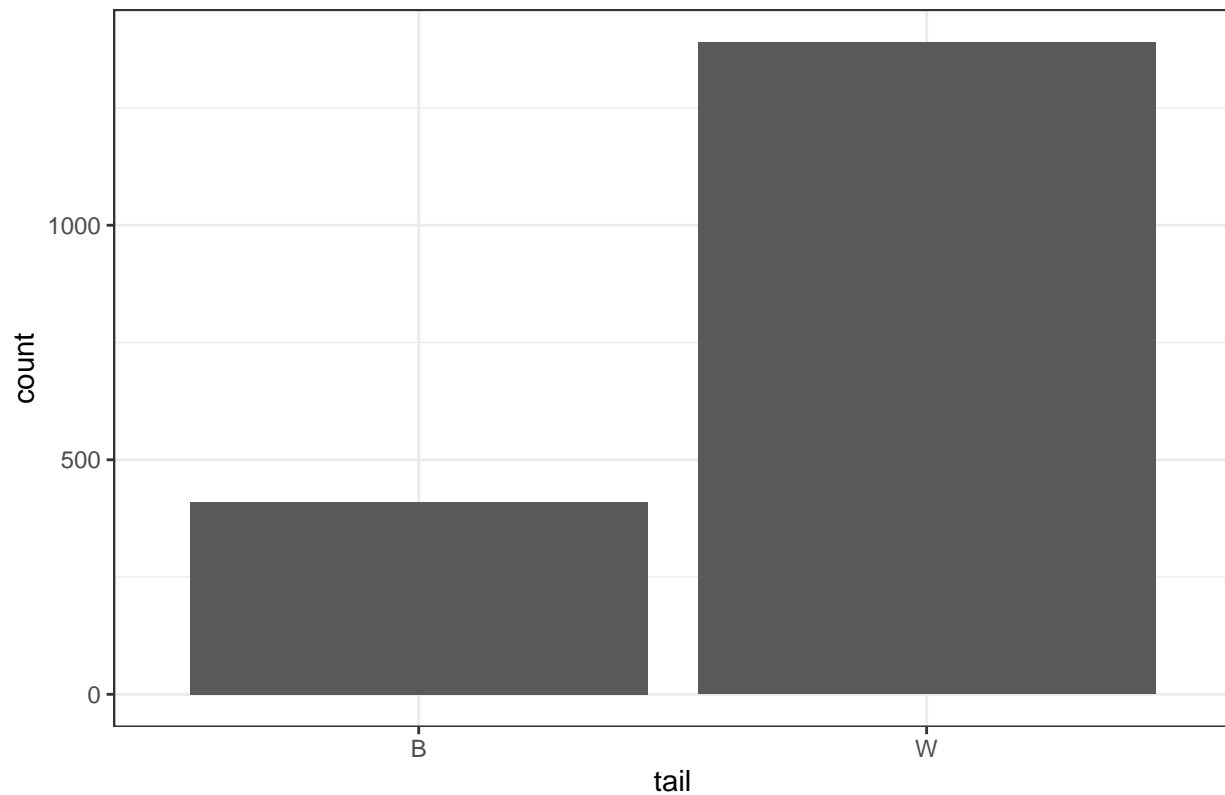
```
bar_zone <- ggplot(utst_lizards_clean, aes(x = zone)) +  
  geom_bar() +  
  theme_bw() +  
  ggtitle("Count of lizards in each Zone")  
bar_zone
```

Count of lizards in each Zone



```
bar_tail <- ggplot(utst_lizards_clean, aes(x = tail)) +  
  geom_bar() +  
  theme_bw() +  
  ggtitle("Count of lizards for each Tail Type")  
bar_tail
```

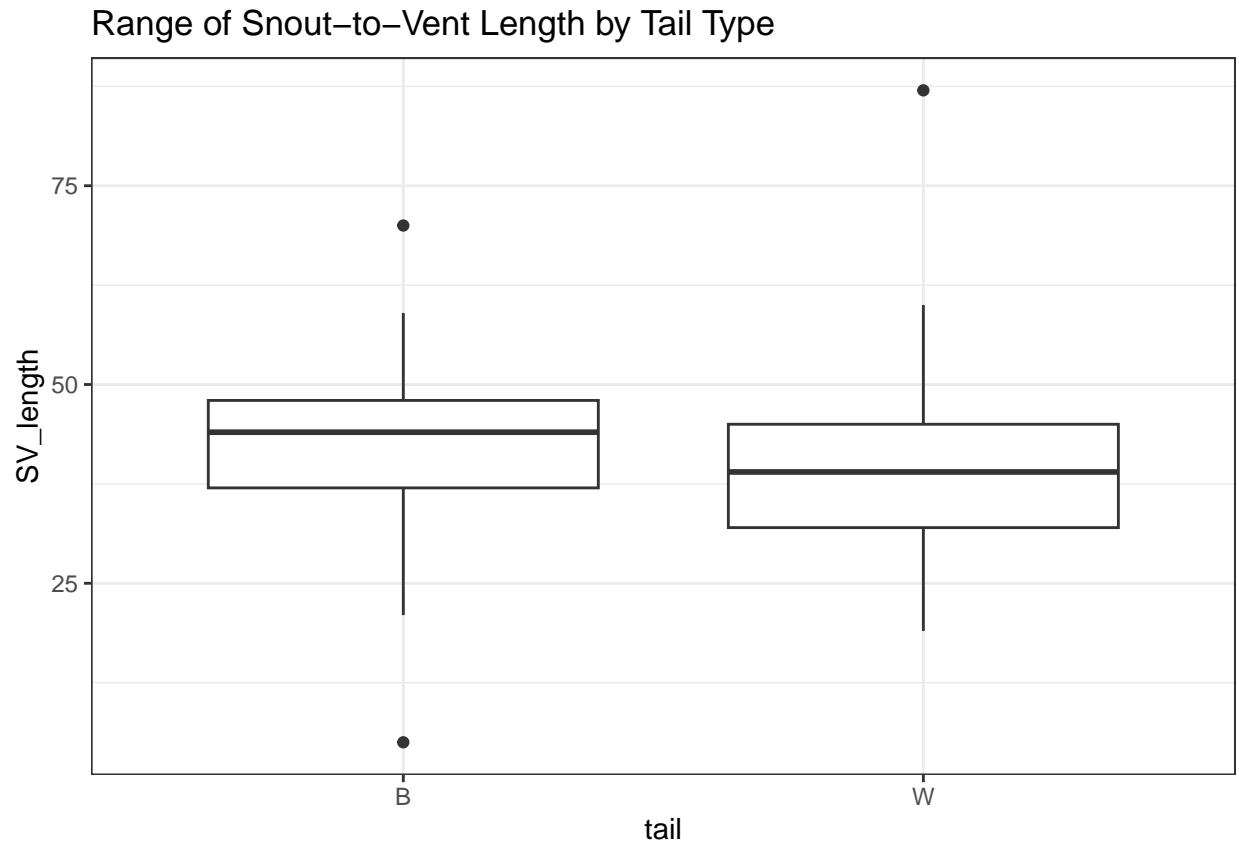
Count of lizards for each Tail Type



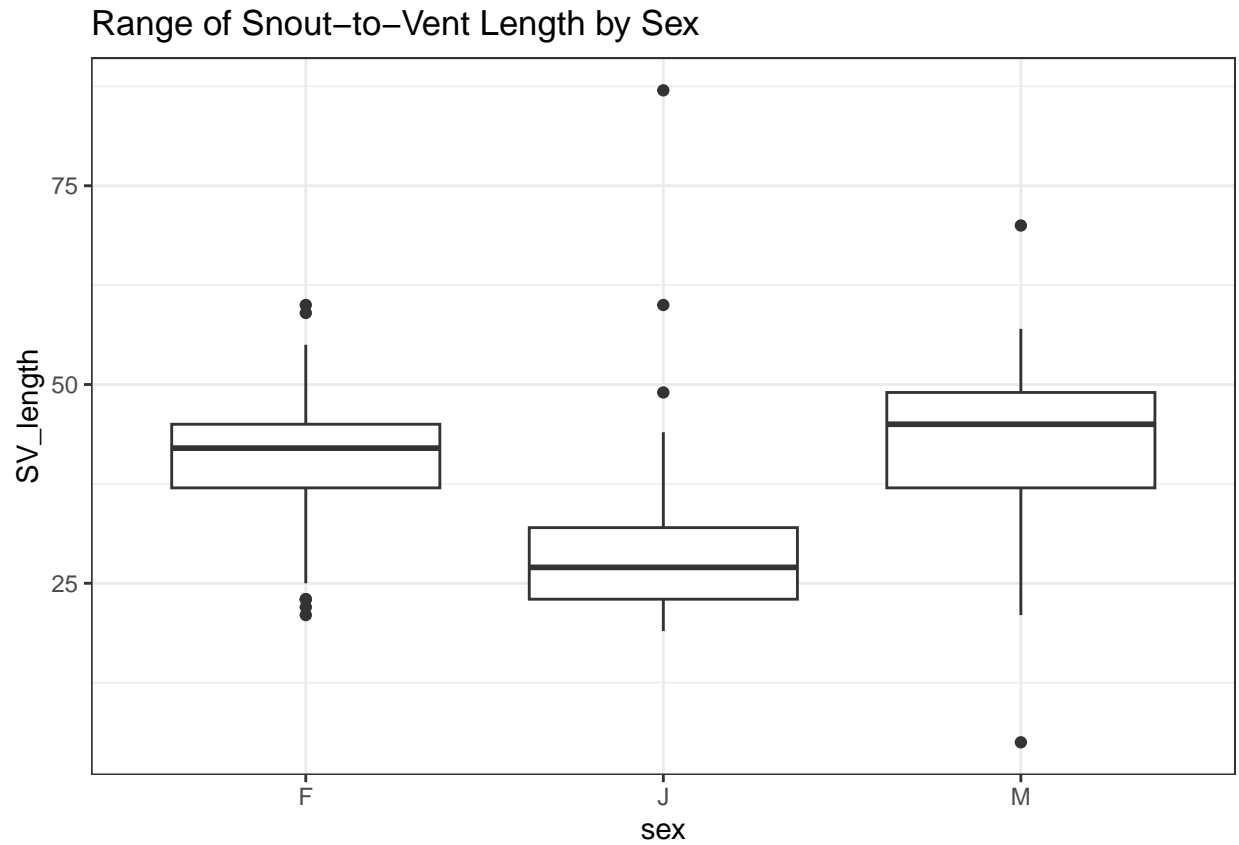
Observations from distributions: 1)SV_length: concentrated between 25 and 55, some potential outliers (values above 70 and below 15). 2) Sex: three major categories with F being most common 3) Zone: Zone M (mesquite duneland) has the highest count 4) Tail: There are much more lizards recorded with a Whole tail. There is a strong class imbalance here.

Relationships between predictors (SV_length, Sex, Zone) and the outcome variable (Tail)

```
boxplot_1 <- ggplot(utst_lizards_clean, aes(x = tail, y = SV_length)) +
  geom_boxplot() +
  theme_bw() +
  ggtitle("Range of Snout-to-Vent Length by Tail Type")
boxplot_1
```

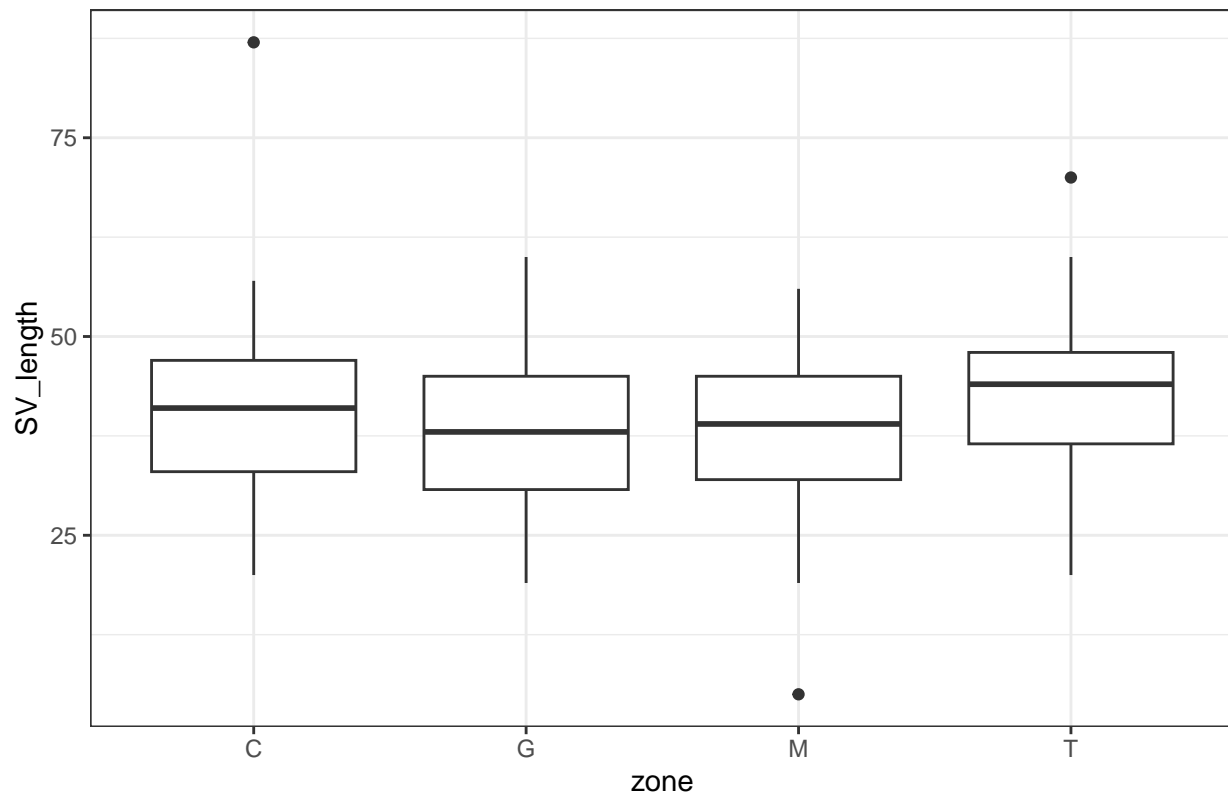


```
boxplot_2 <- ggplot(utst_lizards_clean, aes(x = sex, y = SV_length)) +  
  geom_boxplot() +  
  theme_bw() +  
  ggtitle("Range of Snout-to-Vent Length by Sex")  
boxplot_2
```



```
boxplot_3 <- ggplot(utst_lizards_clean, aes(x = zone, y = SV_length)) +  
  geom_boxplot() +  
  theme_bw() +  
  ggtitle("Tail") +  
  ggtitle("Range of Snout-to-Vent Length by Zone")  
boxplot_3
```


Range of Snout-to-Vent Length by Zone



These three box plots display the distribution of snout-to-vent (SV) length in lizards, grouped by tail type, sex, and zone. The first plot shows that lizards with B tend to have slightly longer SV lengths than those with whole tails, though the difference is small and both groups include a few outliers. The second plot reveals that juveniles (J) have the shortest SV lengths, while males (M) and females (F) show greater lengths with overlapping ranges: males have a slightly higher median than females. The third plot compares SV length by geographic zone, showing relatively similar distributions in zones C, G, and M, with zone T standing out as having the highest median and upper range of SV lengths. Overall, SV length appears to vary meaningfully by sex and zone, and to a lesser extent by tail condition.

##Fitting regression model

```
#Set tail type column to be a factor with a reference level of B
utst_lizards_clean$tail <- factor(utst_lizards_clean$tail, levels = c("B", "W"))
utst_lizards_clean$sex <- factor(utst_lizards_clean$sex, levels = c("J", "F", "M"))
utst_lizards_clean$zone <- factor(utst_lizards_clean$zone, levels = c("C", "G", "M", "T"))

tail_type_mod <- glm(tail ~ SV_length + sex + zone, data = utst_lizards_clean,
                     family = "binomial")
```

##Evaluating model diagnostics

```
summary(tail_type_mod)
```

##

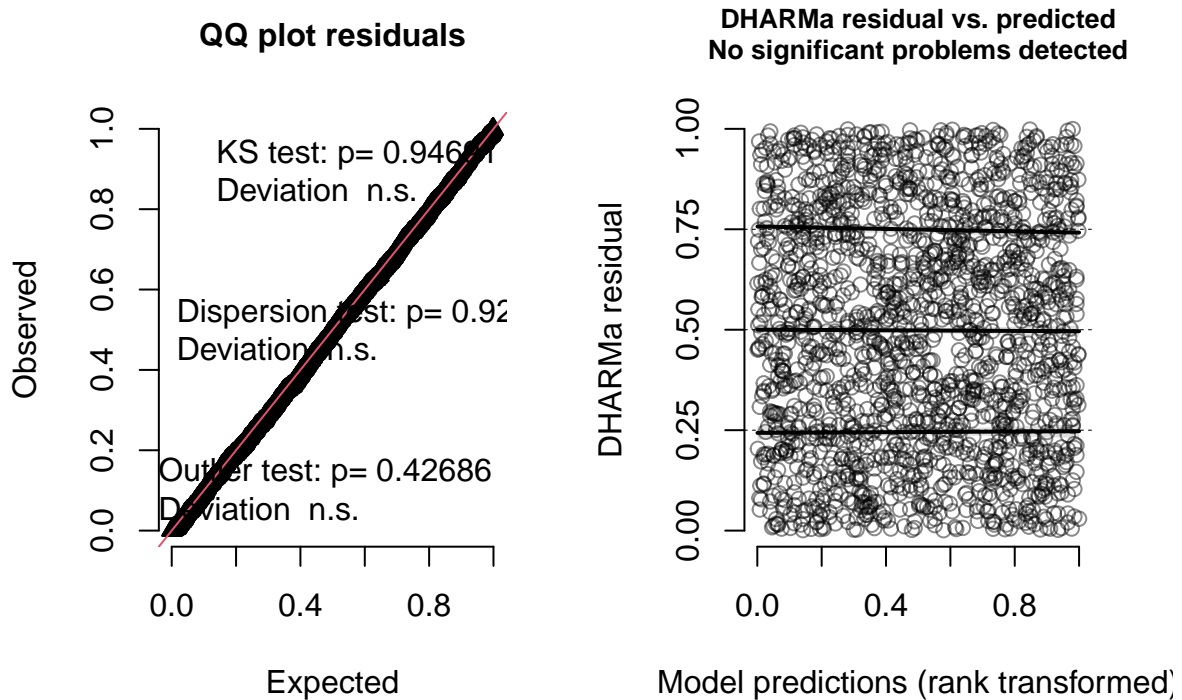
Call:

```
## glm(formula = tail ~ SV_length + sex + zone, family = "binomial",
```

```
## data = utst_lizards_clean)
##
## Coefficients:
## Estimate Std. Error z value Pr(>|z|)
## (Intercept) 3.493940 0.332420 10.511 < 2e-16 ***
## SV_length -0.045230 0.008877 -5.095 3.49e-07 ***
## sexF -0.203871 0.206033 -0.990 0.32241
## sexM -0.138956 0.223664 -0.621 0.53442
## zoneG -0.565369 0.173745 -3.254 0.00114 **
## zoneM -0.301313 0.162032 -1.860 0.06294 .
## zoneT -0.404349 0.179938 -2.247 0.02463 *
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 1931.7 on 1799 degrees of freedom
## Residual deviance: 1867.3 on 1793 degrees of freedom
## AIC: 1881.3
##
## Number of Fisher Scoring iterations: 4

simulateResiduals(tail_type_mod) %>% plot()
```

DHARMa residual



The residual diagnostics indicate a good fit of the logistic regression model. The QQ plot on the left shows that the observed residuals closely follow the expected uniform distribution, with no significant deviation

(KS test: $p = 0.255$). The residual vs. predicted plot on the right shows that residuals are evenly scattered without clear patterns or trends, suggesting no major issues with heteroscedasticity. Overall, the residuals appear to follow what we should expect.

Interpreting model and communicating results

The results of a logistic regression indicate that the log-odds of having a whole tail significantly decrease with increasing Snout-to-Vent length ($B = -0.045$, $p < 0.001$). Compared to lizards in the reference vegetation zone C, lizards in zone G ($B = -0.565$, $p = 0.001$) and zone T ($B = -0.404$, $p = 0.025$) have significantly lower log-odds of having a whole tail. There is no significant difference in tail status for lizards in zone M ($p = 0.063$), or between juveniles and either sex (sexF or sexM).

#2 Predictive Figures showing the change in probability of a lizard tail being whole as lizard length increases, with one including lines colored by sex and another including lines colored by vegetation zone

```
SV_length_vector <- rep(seq(from = 0, to = 90), 3)

# Keep zone constant at "C"
zone_vector <- rep("C", length(SV_length_vector))

# Simulate sex data repeating "F", "J", "M"
sex_vector <- rep(c("F", "J", "M"), each = length(seq(from = 0, to = 90)))

data_pred <- data.frame(SV_length_vector, zone_vector, sex_vector)
colnames(data_pred) <- c("SV_length", "zone", "sex")

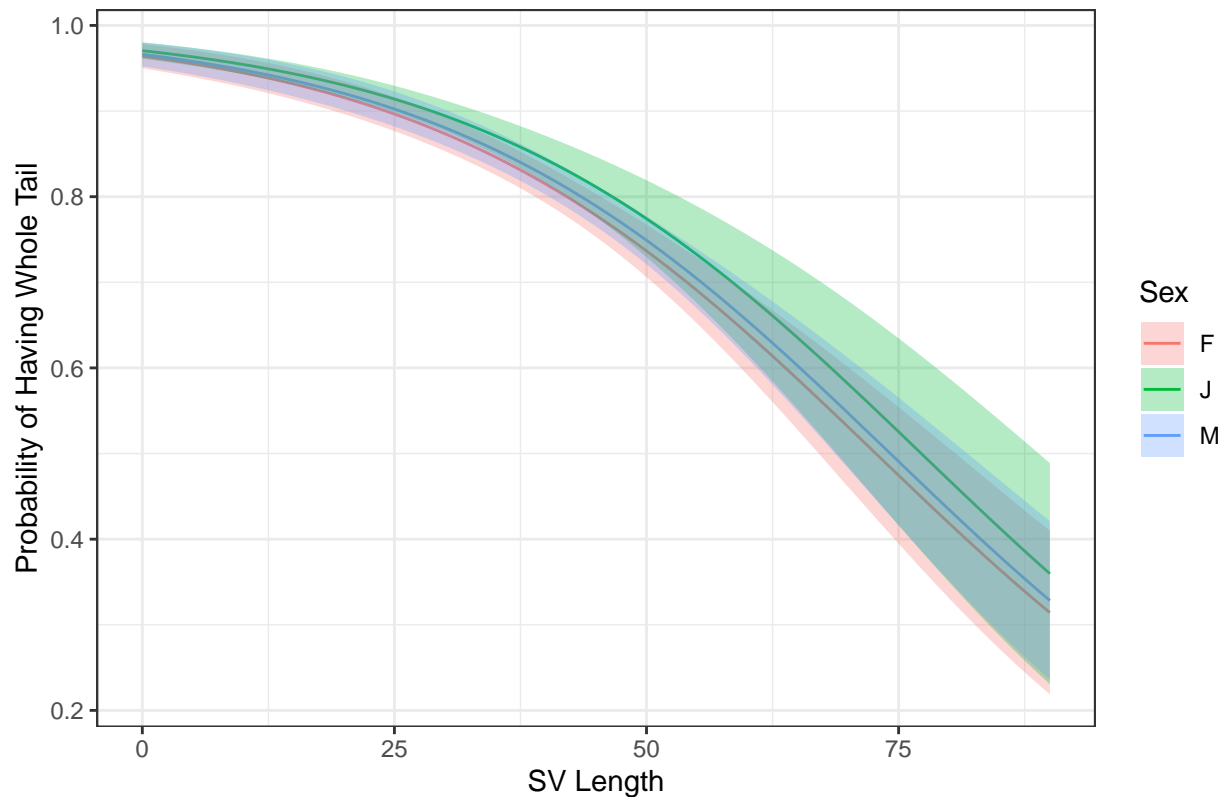
prediction <- predict(tail_type_mod,
                      newdata = data_pred,
                      type = "response",
                      se.fit = TRUE)

data_fig <- data.frame(data_pred,
                      prediction$fit,
                      prediction$se.fit)

# Rename columns.
colnames(data_fig) <- c("SV_length", "zone", "sex", "probability", "se")

# Graph probability of a lizard tail being whole as lizard length increases, with
# lines colored by sex
ggplot(data_fig, aes(x = SV_length, y = probability)) +
  geom_line(aes(color = sex)) +
  geom_ribbon(aes(ymin = probability - se,
                 ymax = probability + se,
                 fill = sex), alpha = 0.3) +
  labs(x = "SV Length",
       y = "Probability of Having Whole Tail",
       color = "Sex",
       fill = "Sex") +
  theme_bw() +
  ggtitle("Probability of Whole Tail with Change in Lizard Length by Sex")
```

Probability of Whole Tail with Change in Lizard Length by Sex



This figure shows the modeled probability of a lizard having a whole tail as a function of snout-to-vent length, with separate curves for females (F), juveniles (J), and males (M). Across all sex categories, the probability of having a whole tail decreases steadily as lizard length increases. While all three groups follow a similar downward trend, juveniles consistently show a slightly higher probability of a whole tail at larger lengths, whereas females show the lowest probabilities across much of the SV length range. However, the confidence intervals for each group (shaded areas) broadly overlap, alluding to differences between sexes not being statistically significant (which is what was found from the GLM results).

```
SV_length_vector <- rep(seq(from = 0, to = 90), 4)

# Simulate zone data repeating "C", "G", "M", "T"
zone_vector <- rep(c("C", "G", "M", "T"), each = length(seq(from = 0, to = 90)))

# Keep sex constant at "F"
sex_vector <- rep("J", length(SV_length_vector))

data_pred <- data.frame(SV_length_vector, zone_vector, sex_vector)
colnames(data_pred) <- c("SV_length", "zone", "sex")

prediction <- predict(tail_type_mod,
                      newdata = data_pred,
                      type = "response",
                      se.fit = TRUE)

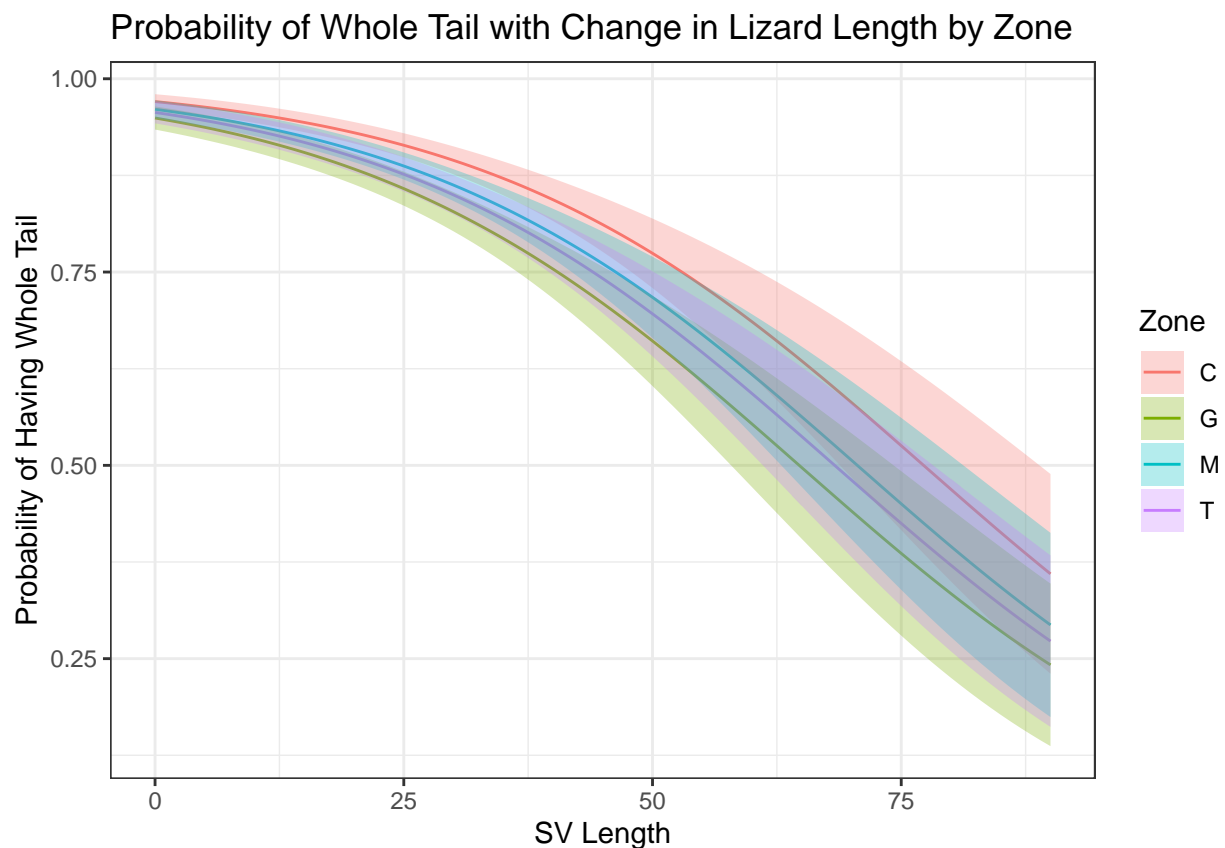
data_fig <- data.frame(data_pred,
                      prediction$fit,
```

```

prediction$se.fit)

# Rename columns.
colnames(data_fig) <- c("SV_length", "zone", "sex", "probability", "se")
#Graph probability of a lizard tail being whole as lizard length increases, with
#lines colored by zone
ggplot(data_fig, aes(x = SV_length, y = probability)) +
  geom_line(aes(color = zone)) +
  geom_ribbon(aes(ymin = probability - se,
                 ymax = probability + se,
                 fill = zone), alpha = 0.3) +
  labs(x = "SV Length",
       y = "Probability of Having Whole Tail",
       color = "Zone",
       fill = "Zone") +
  theme_bw() +
  ggtitle("Probability of Whole Tail with Change in Lizard Length by Zone")

```



This figure illustrates how the probability of having a whole tail changes with SV length across four vegetation zones (C = creosotebush shrubland, G = black grama grassland, M = mesquite duneland, T = tarbush shrubland). In all zones, the probability of whole tail declines as lizard length increases. Lizards in Zone C consistently show the highest probability of having a whole tail across the SV length range, while those in Zones G, M, and T have lower probabilities, with Zone G showing the steepest decline. Although the confidence intervals overlap, the separation between the curves indicates potential differences in tail loss patterns by zone.

#(2) Poisson Regression

Download the dataset `jrn_lizard_npp.csv` from Canvas. This dataset combines counts of lizards captured using pitfall traps with percent cover of different plant species at the Jornada Basin Long Term Ecological Research. For this portion of the assignment, you must build a poisson regression model to investigate the following research question:

Do season (season) and plant species percent cover (BOER, LATR, PRGL, SCBR) significantly predict lizard counts (lizard_count)?

These data have already been filtered only for side-blotched lizards (*Uta stansburiana*). Seasons include fall (F), winter (W), and summer (S). In addition, the four plant species whose percent covers are reported here are creosote bush (*Larrea tridentata* or LATR), burrograss (*Scleropogon brevifolius* or SCBR), honey mesquite (*Prosopis glandulosa* or PRGL), and black grama grass (*Bouteloua eriopoda* or BOER). Note, “percent cover” refers to the percent of a given quadrat, or sampling area, that a species occupies.

Using the workflow in today’s lab as a guide, work through the steps of performing a poisson regression using these data, including exploratory data visualizations. You are not expected to include all four plant species in your model structure but should use your initial data visualizations to help determine which you feel are best suited to help predict lizard counts. You should report the results of your final model in text form (1-2 sentences). If you choose to remove outliers during the modeling process, be sure to report how many you removed and your justification for doing so.

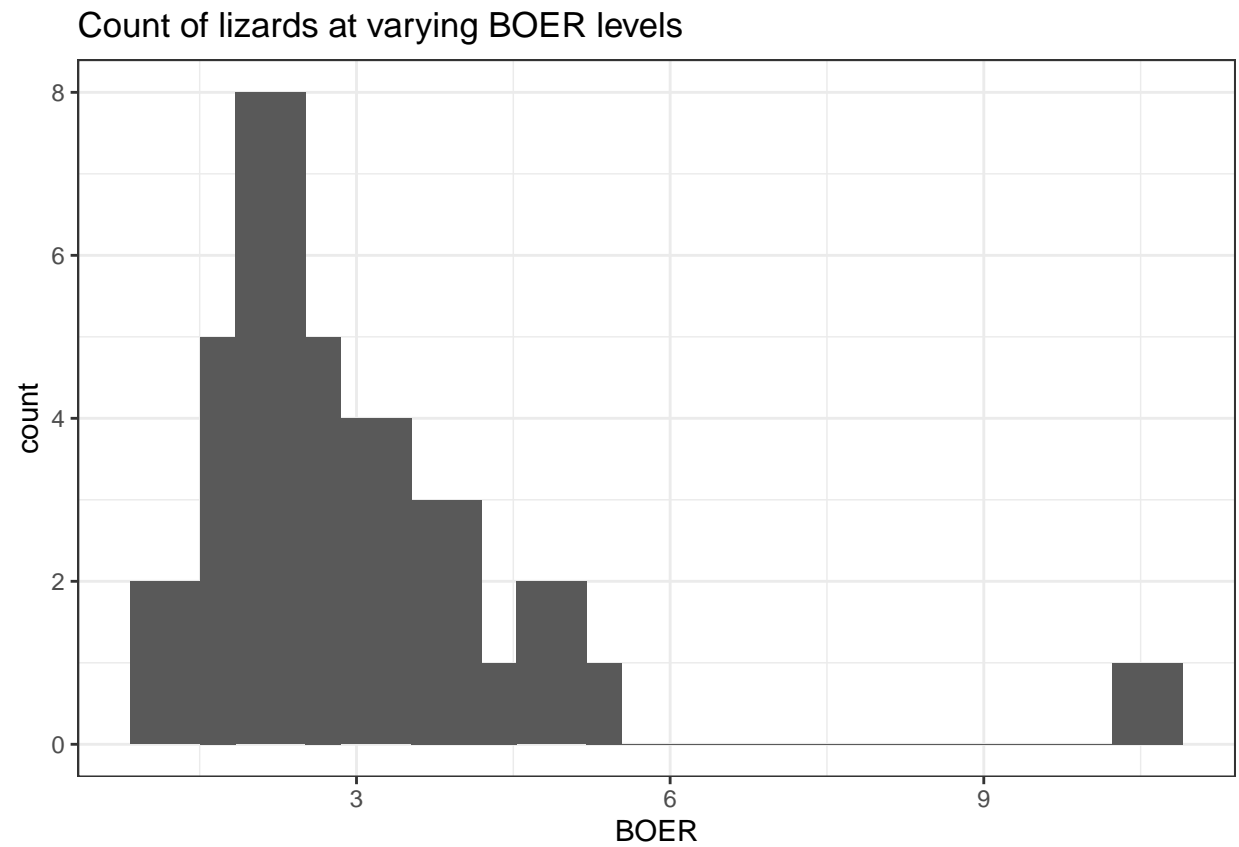
```
lizard_npp <- read_csv("../Assignments/week9/jrn_lizard_npp.csv")

## Rows: 52 Columns: 7
## -- Column specification -----
## Delimiter: ","
## chr (1): season
## dbl (6): sample_year, lizard_count, BOER, LATR, PRGL, SCBR
##
## i Use 'spec()' to retrieve the full column specification for this data.
## i Specify the column types or set 'show_col_types = FALSE' to quiet this message.

## Visualizing Data

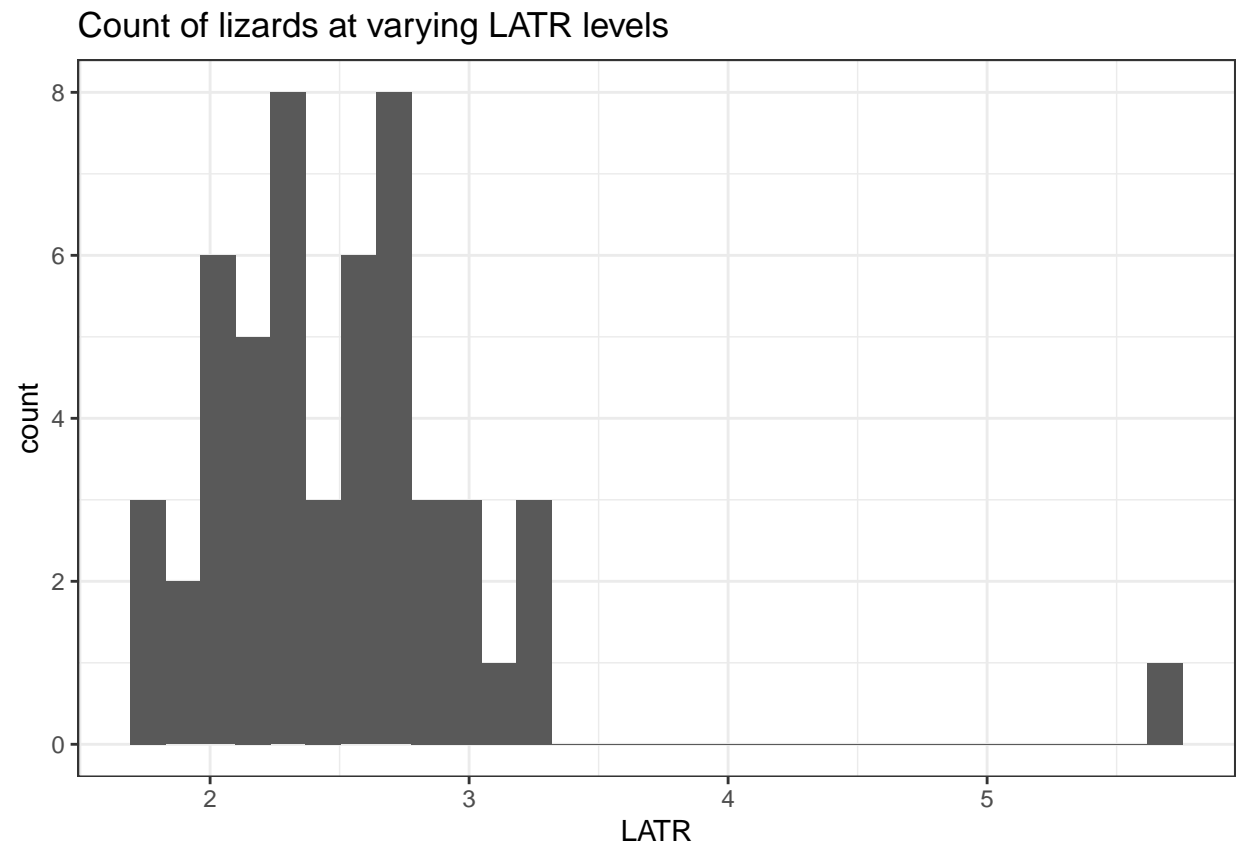
hist_BOER <- ggplot(lizard_npp, aes(x = BOER)) +
  geom_histogram() +
  theme_bw() +
  ggtitle("Count of lizards at varying BOER levels")
hist_BOER

## 'stat_bin()' using 'bins = 30'. Pick better value with 'binwidth'.
```



```
hist_LATR <- ggplot(lizard_npp, aes(x = LATR)) +  
  geom_histogram() +  
  theme_bw() +  
  ggtitle("Count of lizards at varying LATR levels")  
hist_LATR
```

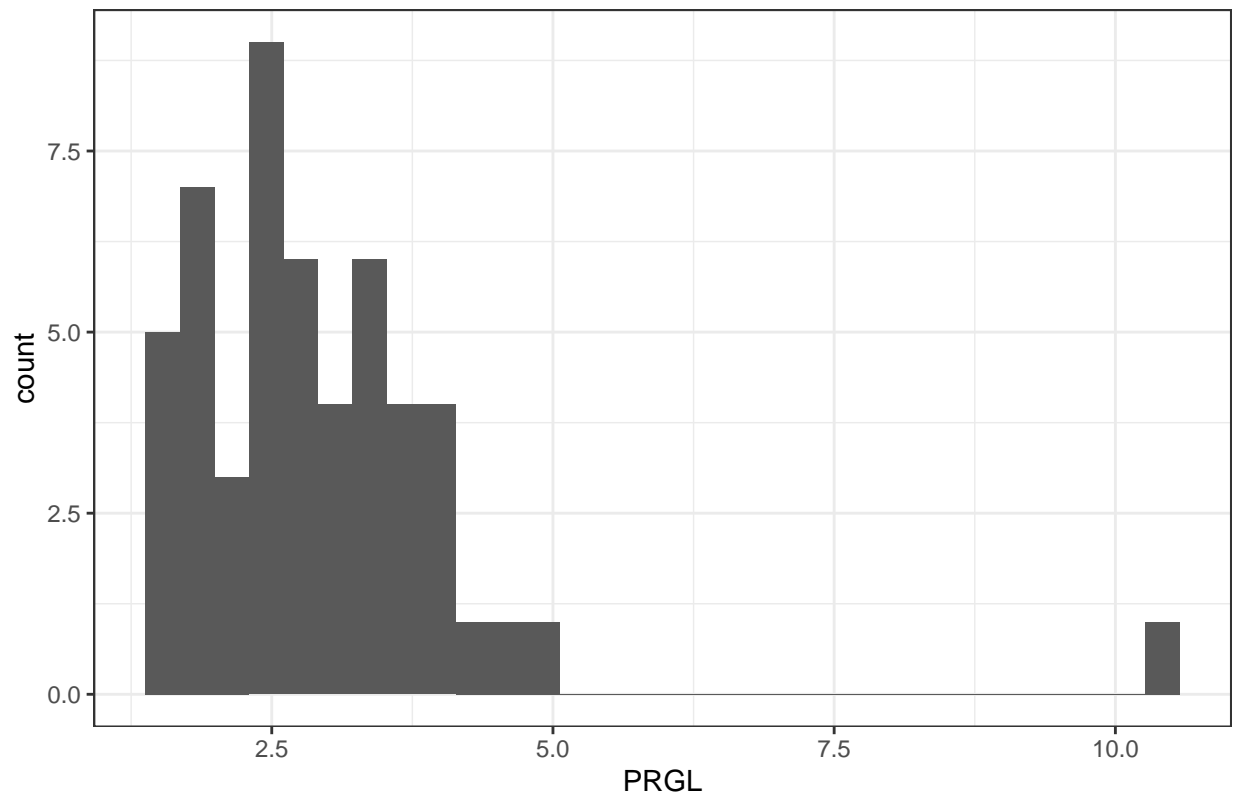
```
## 'stat_bin()' using 'bins = 30'. Pick better value with 'binwidth'.
```



```
hist_PRGL <- ggplot(lizard_npp, aes(x = PRGL)) +  
  geom_histogram() +  
  theme_bw() +  
  ggtitle("Count of lizards at varying PRGL levels")  
hist_PRGL
```

```
## 'stat_bin()' using 'bins = 30'. Pick better value with 'binwidth'.
```


Count of lizards at varying PRGL levels

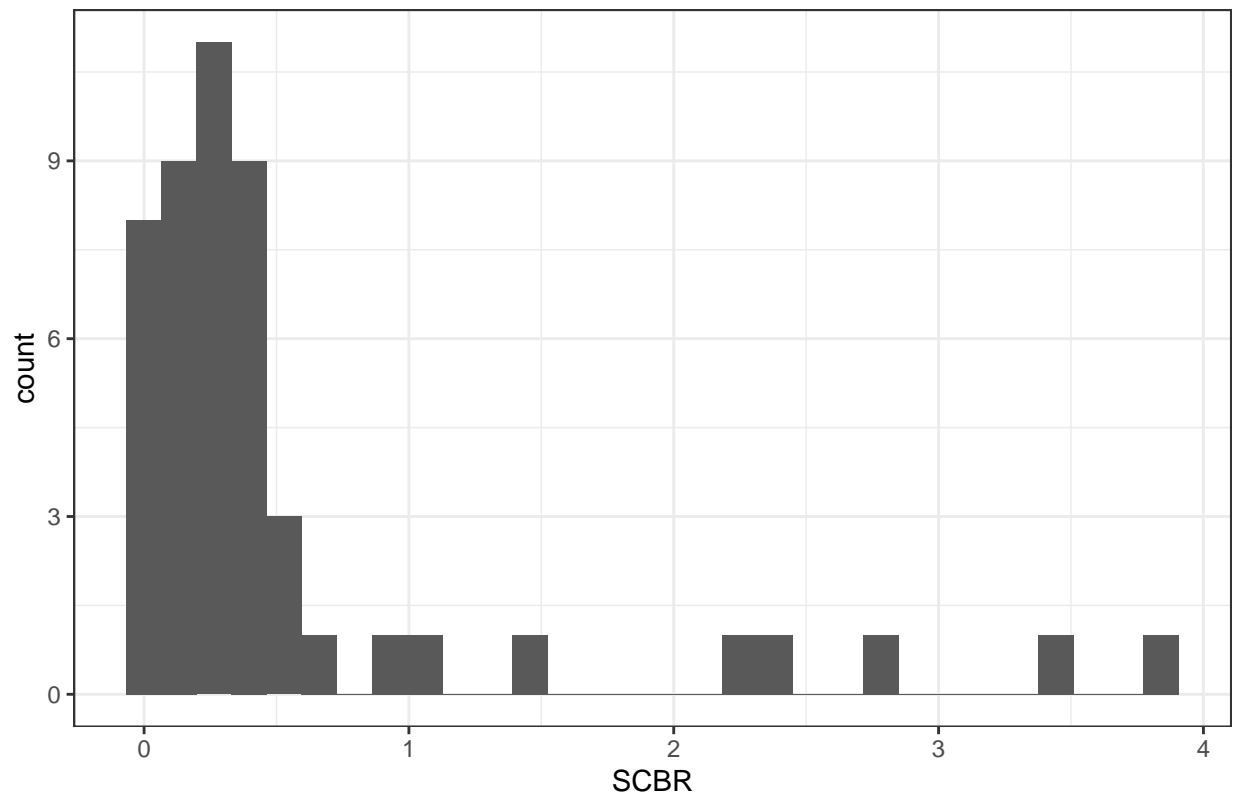


```
hist_SCBR <- ggplot(lizard_npp, aes(x = SCBR)) +  
  geom_histogram() +  
  theme_bw() +  
  ggtitle("Count of lizards at varying SCBR levels")  
hist_SCBR
```

```
## 'stat_bin()' using 'bins = 30'. Pick better value with 'binwidth'.
```

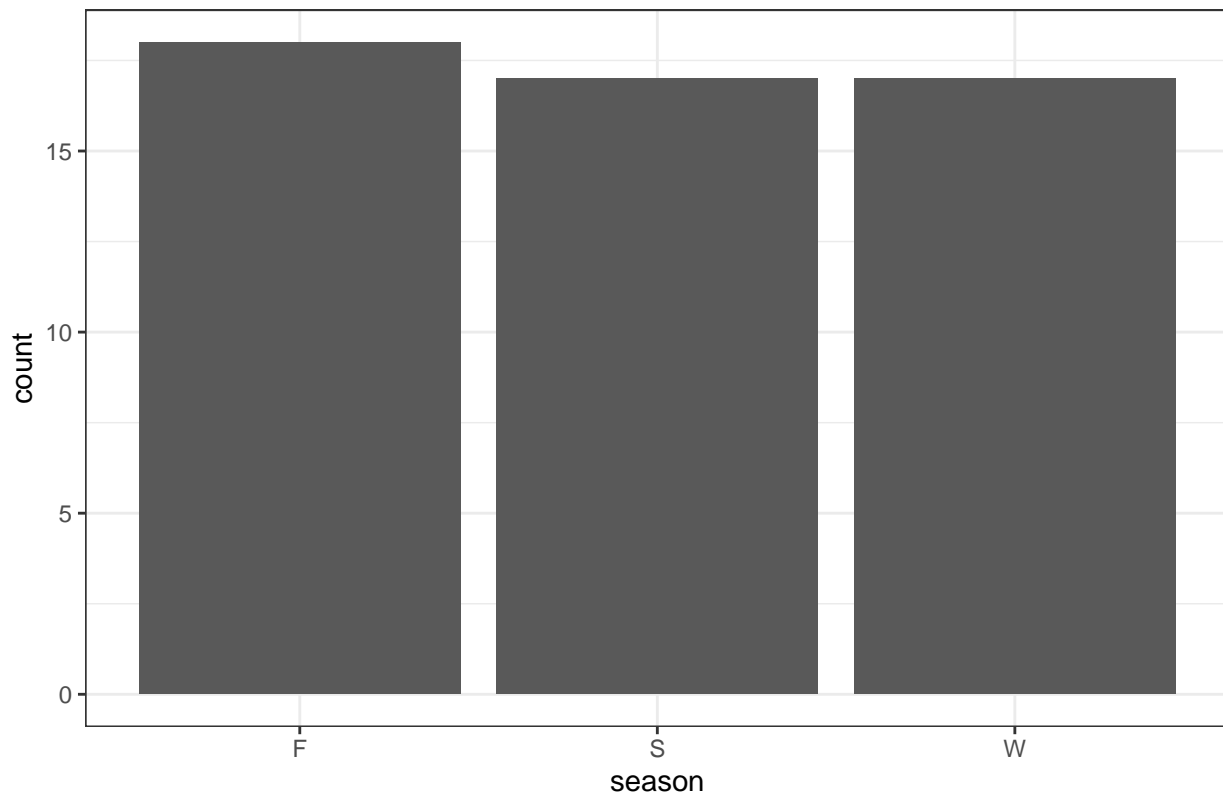
```
## Warning: Removed 3 rows containing non-finite outside the scale range  
## ('stat_bin()').
```

Count of lizards at varying SCBR levels



```
bar_season <- ggplot(lizard_npp, aes(x = season)) +  
  geom_bar() +  
  theme_bw() +  
  ggtitle("Count of lizards Collected in each Season")  
bar_season
```

Count of lizards Collected in each Season



BOER, LATR, and PRGL all seem to have values concentrated on one end with a few number of outliers at the higher end. SCBR has a bit more counts spread out across percentages, though counts are more concentrated between 0-1%. Counts across seasons are fairly balanced, with Fall only seeing slightly more counts.

Investigating multi-collinearity:

```
cor.test(lizard_npp$BOER, lizard_npp$LATR)
```

```
##
## Pearson's product-moment correlation
##
## data: lizard_npp$BOER and lizard_npp$LATR
## t = 5.3965, df = 50, p-value = 1.865e-06
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  0.4000164 0.7546451
## sample estimates:
##      cor
## 0.6066879
```

```
cor.test(lizard_npp$BOER, lizard_npp$PRGL)
```

```
##
## Pearson's product-moment correlation
```

```
##
## data: lizard_npp$BOER and lizard_npp$PRGL
## t = 5.6654, df = 50, p-value = 7.212e-07
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
## 0.4248704 0.7672511
## sample estimates:
##      cor
## 0.6252721
```

```
cor.test(lizard_npp$BOER, lizard_npp$SCBR, use = "complete.obs")
```

```
##
## Pearson's product-moment correlation
##
## data: lizard_npp$BOER and lizard_npp$SCBR
## t = 8.4492, df = 47, p-value = 5.486e-11
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
## 0.6337121 0.8681592
## sample estimates:
##      cor
## 0.7765324
```

```
cor.test(lizard_npp$LATR, lizard_npp$PRGL)
```

```
##
## Pearson's product-moment correlation
##
## data: lizard_npp$LATR and lizard_npp$PRGL
## t = 9.5588, df = 50, p-value = 7.176e-13
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
## 0.6802929 0.8830936
## sample estimates:
##      cor
## 0.80394
```

```
cor.test(lizard_npp$LATR, lizard_npp$SCBR, use = "complete.obs")
```

```
##
## Pearson's product-moment correlation
##
## data: lizard_npp$LATR and lizard_npp$SCBR
## t = 3.7874, df = 47, p-value = 0.0004319
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
## 0.2342117 0.6732166
## sample estimates:
##      cor
## 0.4835612
```

```
cor.test(lizard_npp$PRGL, lizard_npp$SCBR, use = "complete.obs")
```

```
##
## Pearson's product-moment correlation
##
## data: lizard_npp$PRGL and lizard_npp$SCBR
## t = 5.4155, df = 47, p-value = 2.038e-06
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
## 0.4101664 0.7673181
## sample estimates:
## cor
## 0.6198689
```

After checking all the possible pairwise combinations of the species, it looks like LATR and SCBR are the only two species that are not highly correlated with each other (others have correlation values > 0.62) so I will only include these two species in my final model.

##Fitting regression model

```
lizard_npp$season <- factor(lizard_npp$season, levels = c("F", "S", "W"))

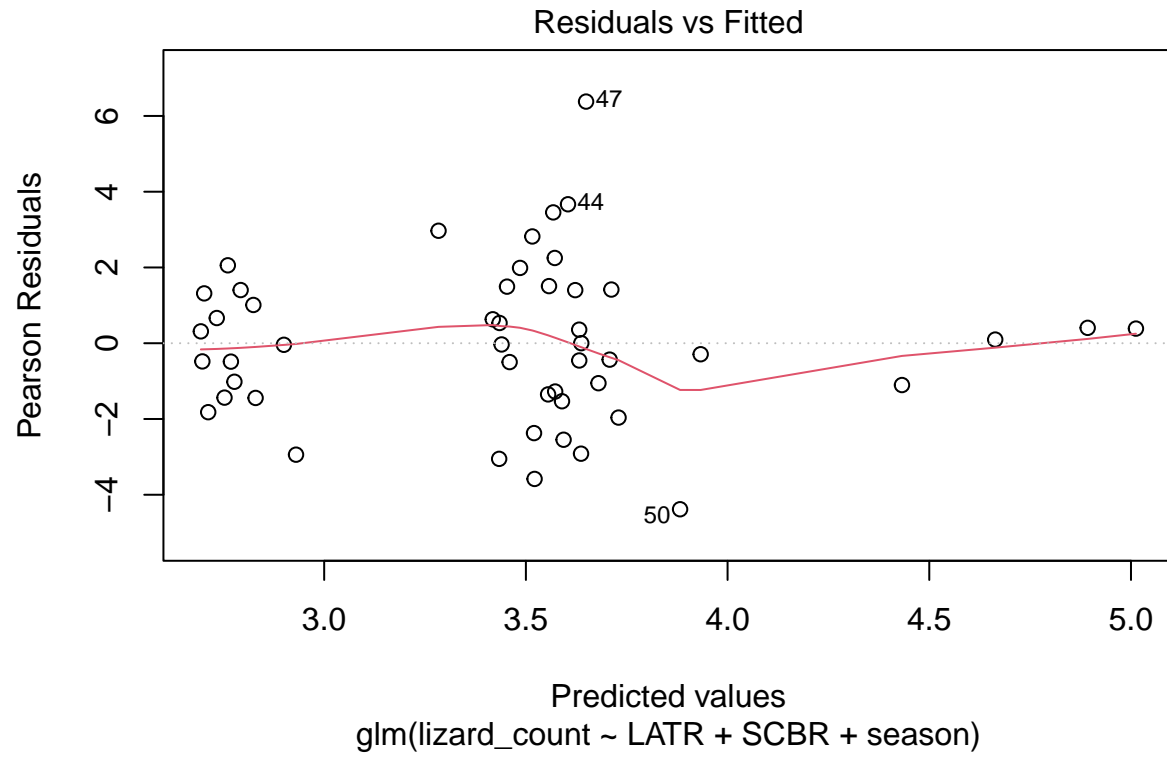
lizard_count_mod <- glm(lizard_count ~ LATR + SCBR + season, data = lizard_npp,
                        family = "poisson")
```

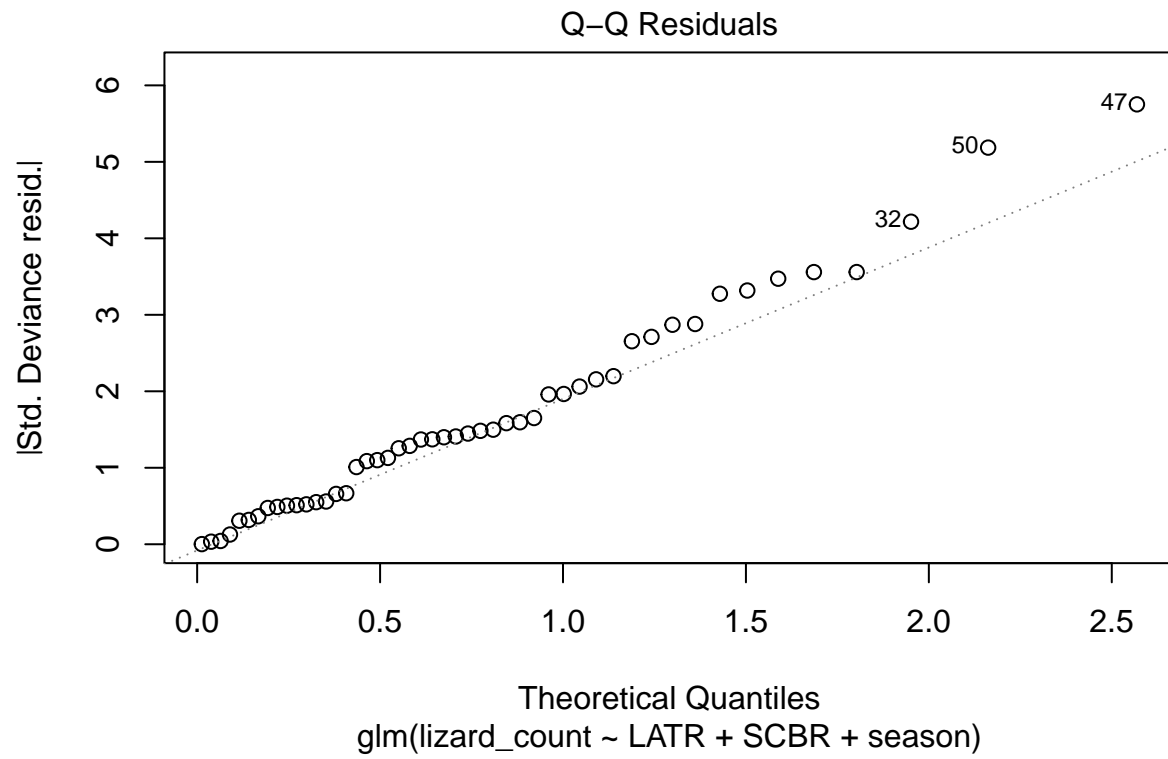
##Evaluating model diagnostics

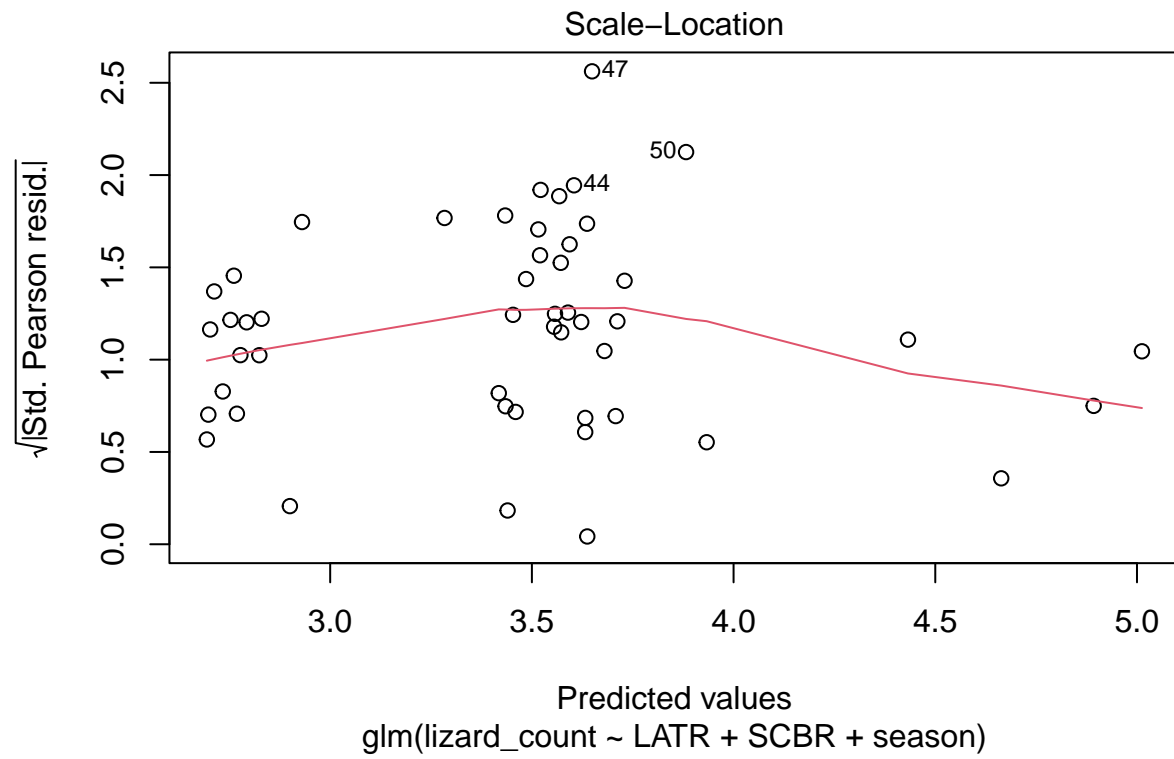
```
summary(lizard_count_mod)
```

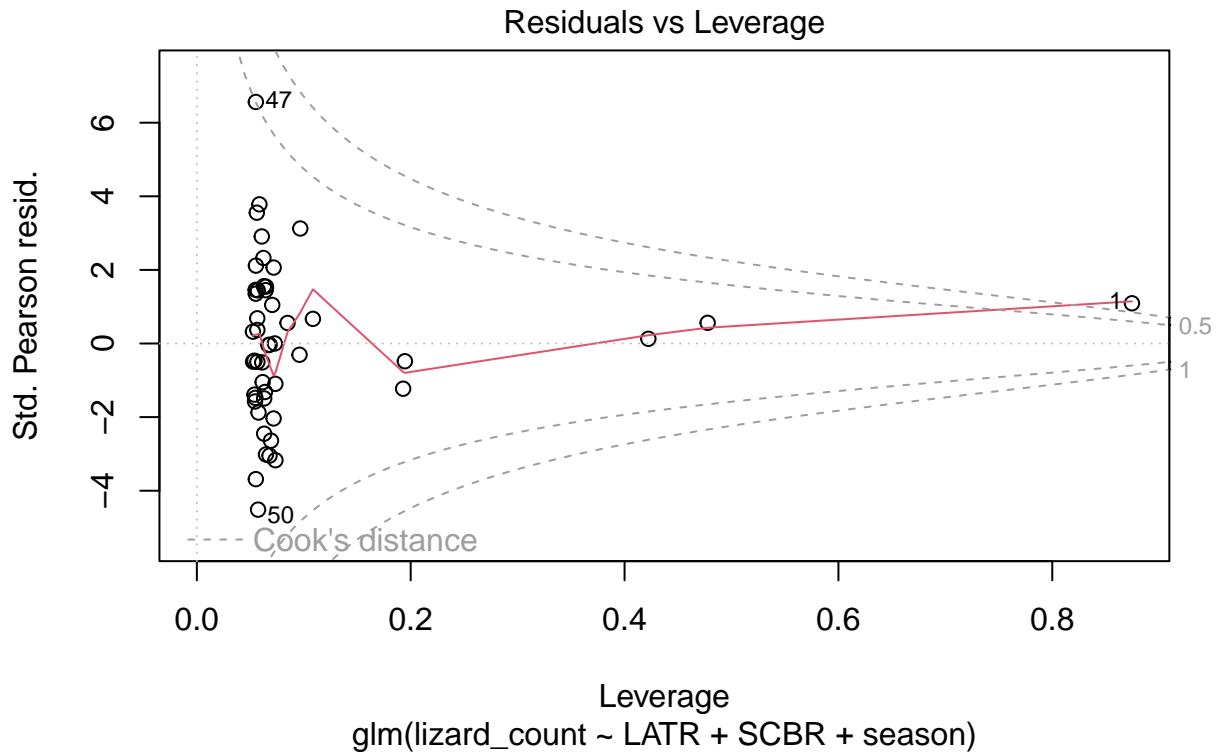
```
##
## Call:
## glm(formula = lizard_count ~ LATR + SCBR + season, family = "poisson",
##      data = lizard_npp)
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)  3.56475    0.08892  40.089  <2e-16 ***
## LATR        -0.04079    0.03286  -1.241    0.215
## SCBR         0.43595    0.02338  18.650  <2e-16 ***
## seasonS     -0.79630    0.06872 -11.588  <2e-16 ***
## seasonW     -0.02633    0.05315  -0.495    0.620
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for poisson family taken to be 1)
##
## Null deviance: 911.44 on 48 degrees of freedom
## Residual deviance: 207.10 on 44 degrees of freedom
## (3 observations deleted due to missingness)
## AIC: 472.76
##
## Number of Fisher Scoring iterations: 4
```

```
plot(lizard_count_mod)
```









Based on the plot of the residuals, point 47 is considered an outlier on the Residuals vs Fitted, Scale-Location, and Residuals vs Leverage plots. Point 1 is also considered an outlier on the Residuals vs Leverage plot. These two points are removed, and I refitted the model on the dataset without the outliers:

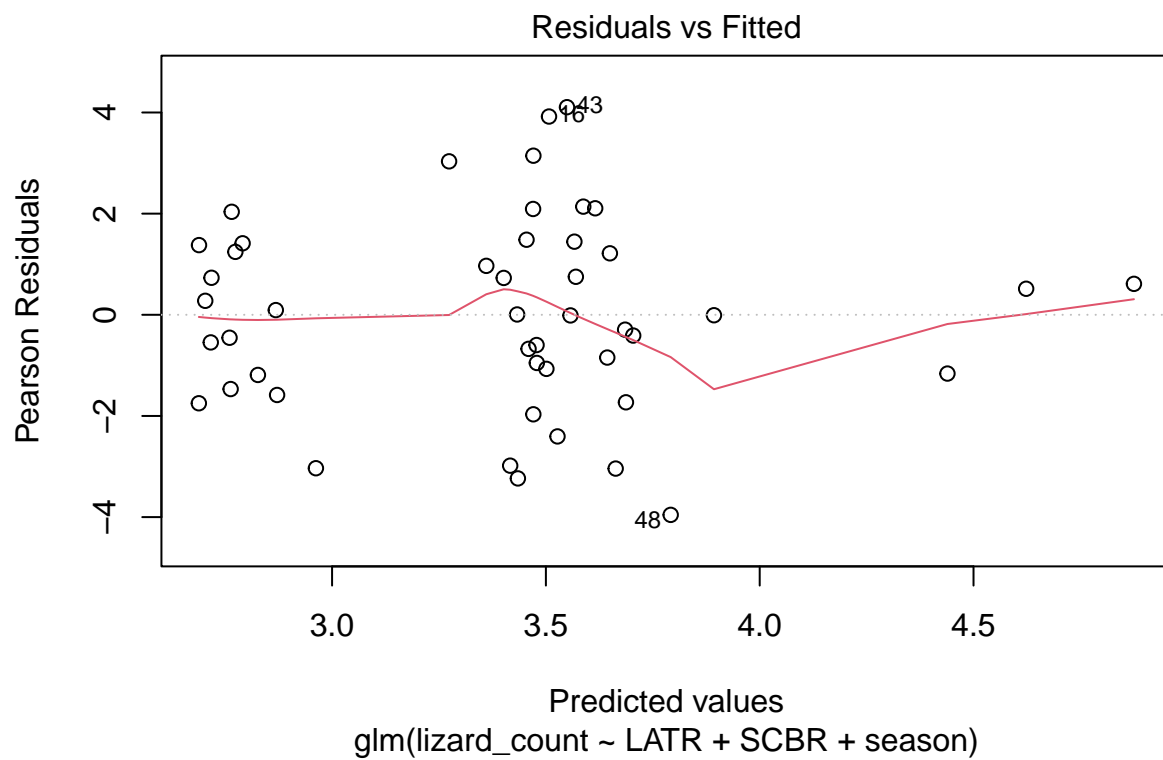
```
#remove outliers using rownumbers
lizard_npp_outliers_removed <- lizard_npp[-c(1, 47),]
lizard_count_mod_2 <- glm(lizard_count ~ LATR + SCBR + season, data = lizard_npp_outliers_removed,
                          family = "poisson")

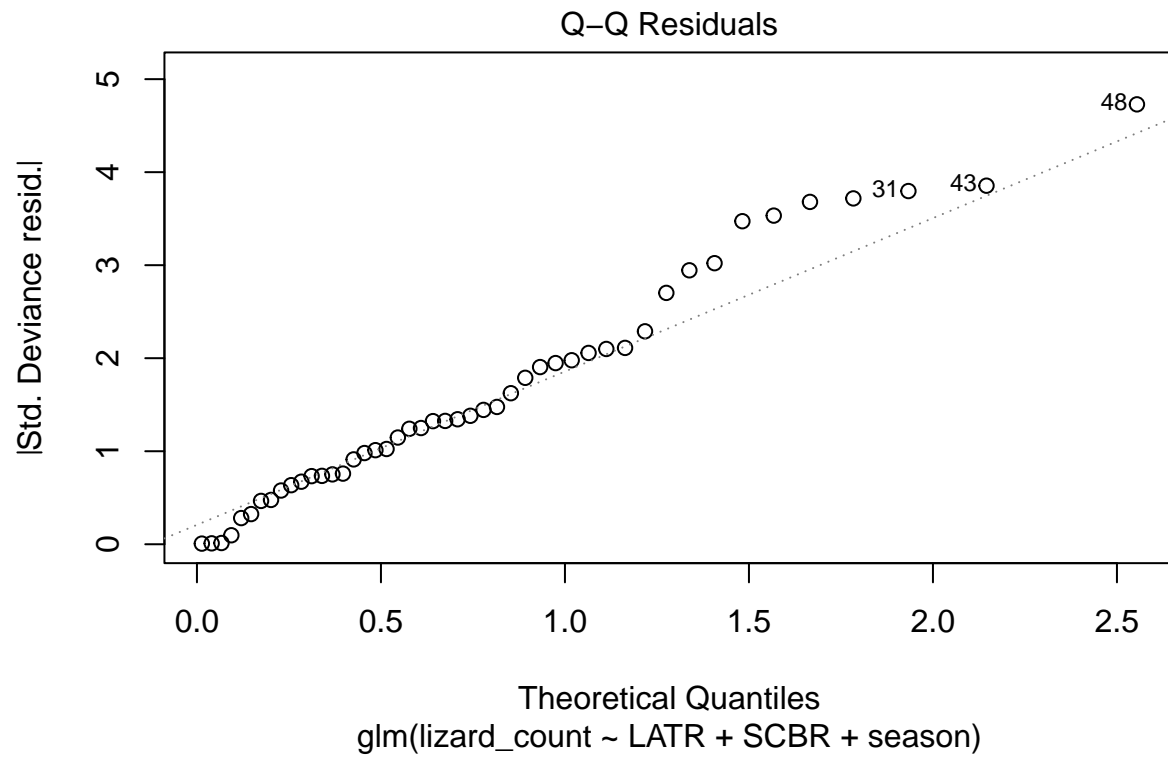
summary(lizard_count_mod_2)
```

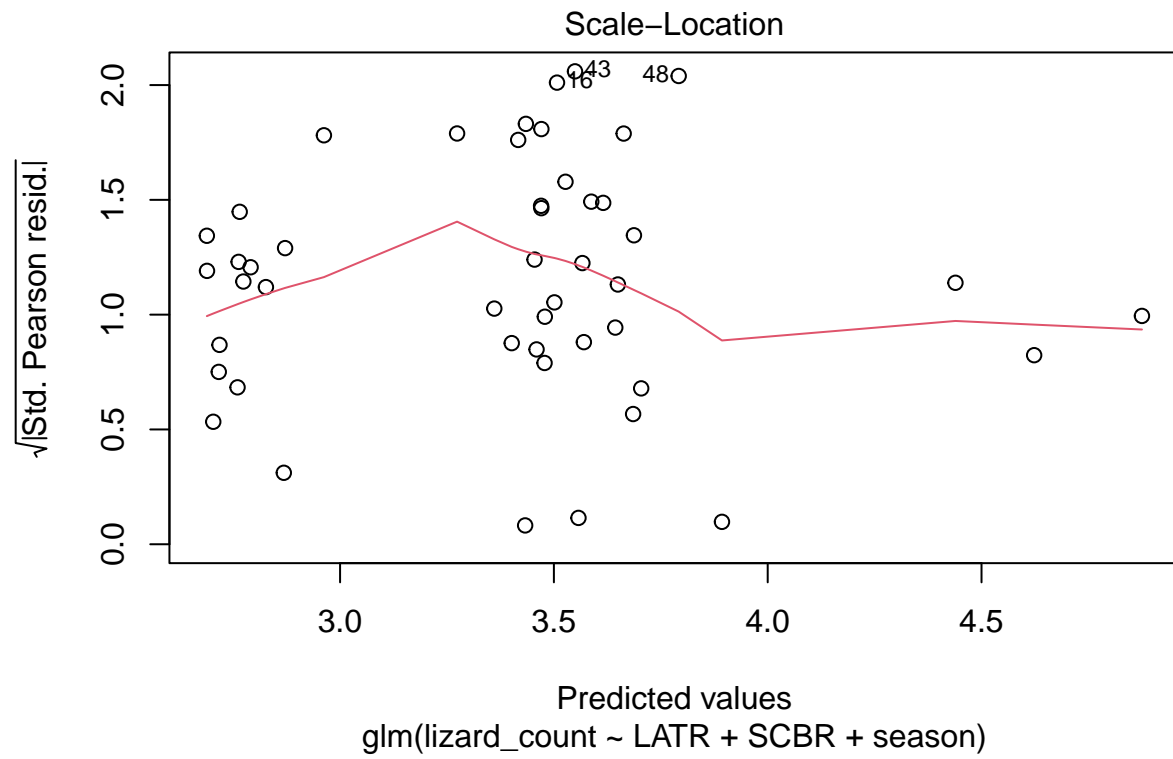
```
##
## Call:
## glm(formula = lizard_count ~ LATR + SCBR + season, family = "poisson",
##      data = lizard_npp_outliers_removed)
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)  3.65630    0.16530  22.119  <2e-16 ***
## LATR         -0.10795    0.06764  -1.596    0.110
## SCBR          0.44183    0.02388  18.503  <2e-16 ***
## seasonS      -0.72215    0.07098 -10.173  <2e-16 ***
## seasonW       0.04546    0.05616   0.809    0.418
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
```

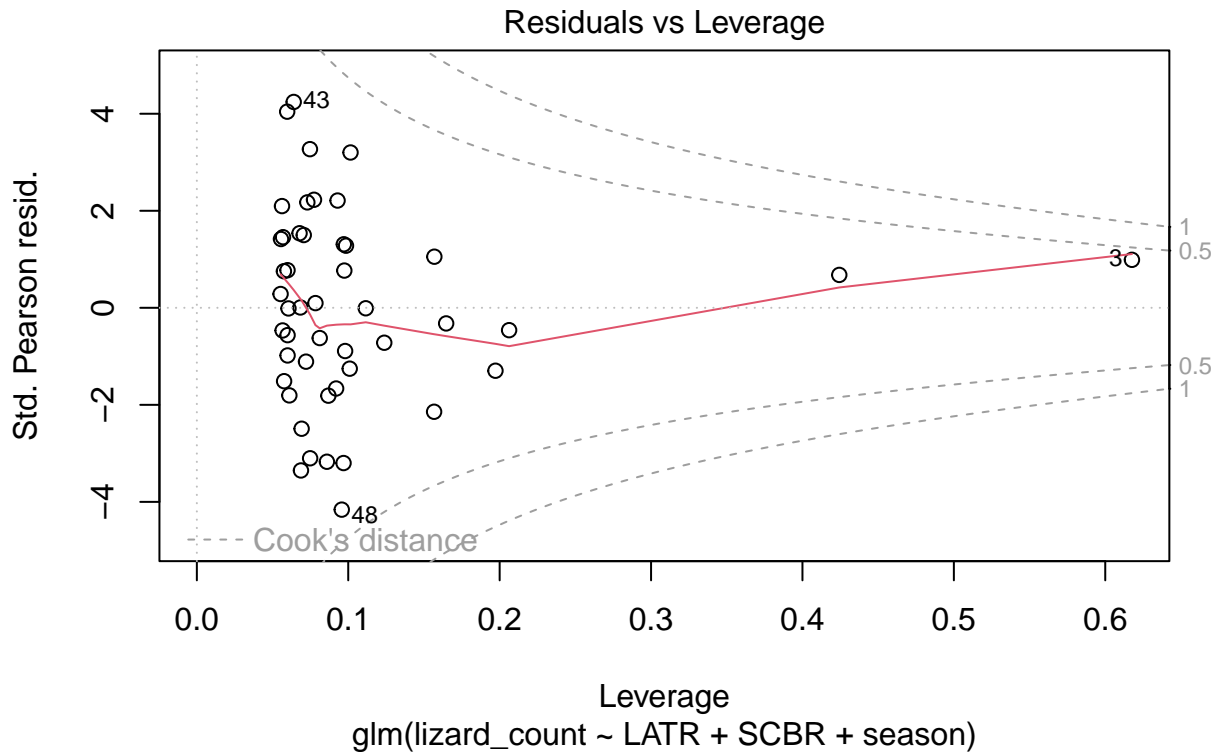
```
## (Dispersion parameter for poisson family taken to be 1)
##
## Null deviance: 655.92 on 46 degrees of freedom
## Residual deviance: 171.78 on 42 degrees of freedom
## (3 observations deleted due to missingness)
## AIC: 424.36
##
## Number of Fisher Scoring iterations: 4
```

```
plot(lizard_count_mod_2)
```









The model still seems to have slightly non-normal residuals, uneven variance, and a few influential points. Point 48 consistently shows up as an extreme outlier across all three plots, so I will remove this last point and refit the model:

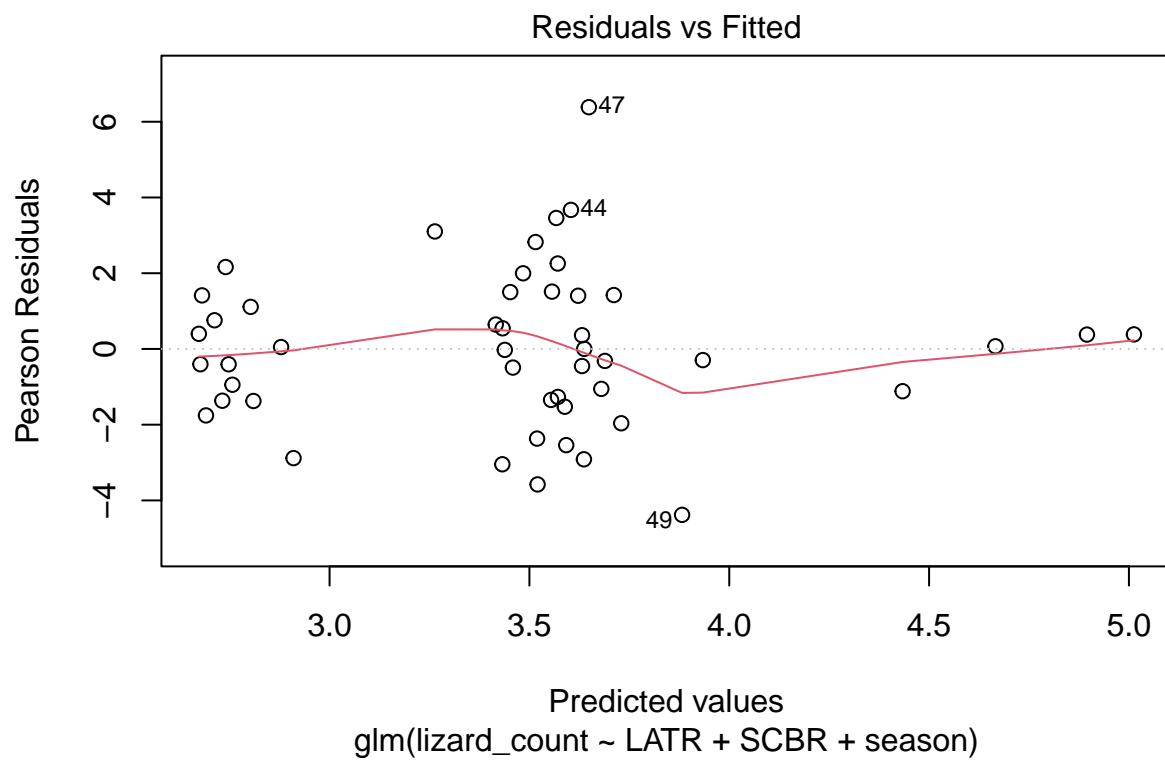
```
#remove outliers using rownumbers
lizard_npp_outliers_removed_2 <- lizard_npp[-c(48),]
lizard_count_mod_3 <- glm(lizard_count ~ LATR + SCBR + season, data = lizard_npp_outliers_removed_2,
                          family = "poisson")

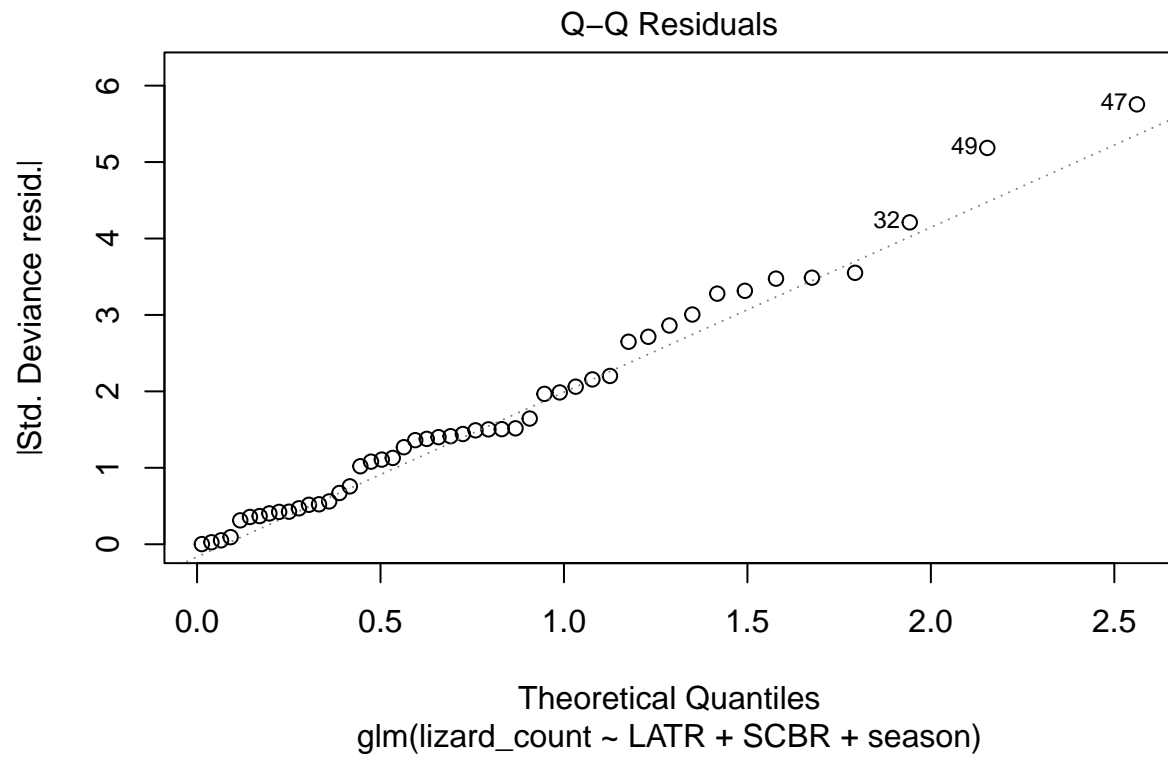
summary(lizard_count_mod_3)
```

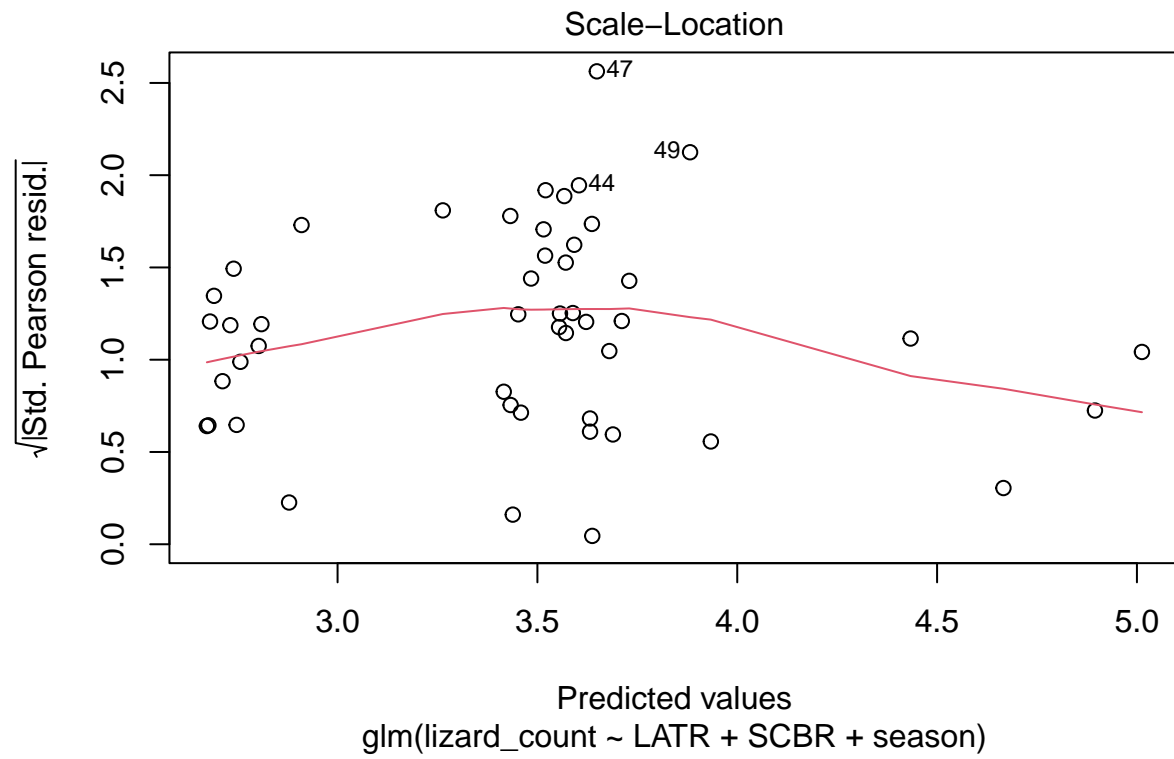
```
##
## Call:
## glm(formula = lizard_count ~ LATR + SCBR + season, family = "poisson",
##      data = lizard_npp_outliers_removed_2)
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)  3.56632    0.08893  40.101  <2e-16 ***
## LATR         -0.04185    0.03288  -1.273    0.203
## SCBR          0.43717    0.02340  18.682  <2e-16 ***
## seasonS      -0.81674    0.07067 -11.558  <2e-16 ***
## seasonW      -0.02656    0.05315  -0.500    0.617
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
```

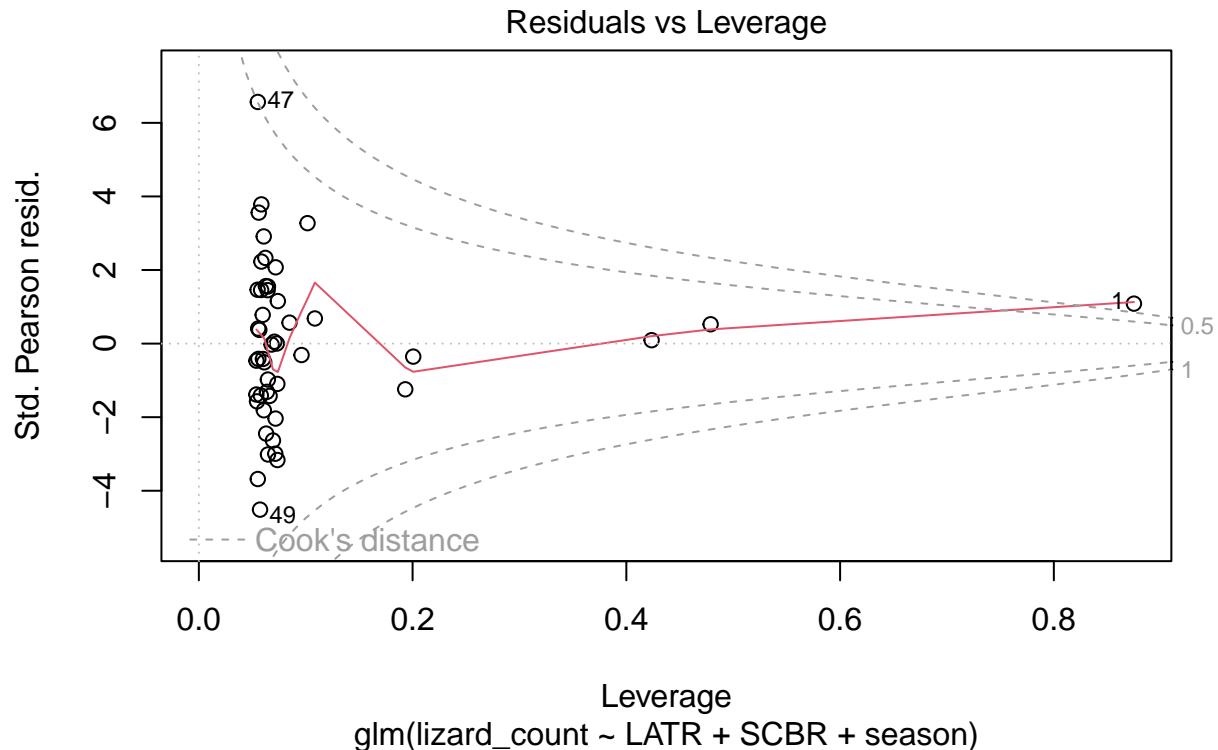
```
## (Dispersion parameter for poisson family taken to be 1)
##
## Null deviance: 904.01 on 47 degrees of freedom
## Residual deviance: 205.20 on 43 degrees of freedom
## (3 observations deleted due to missingness)
## AIC: 465.93
##
## Number of Fisher Scoring iterations: 4
```

```
plot(lizard_count_mod_3)
```









The residual plots overall look better. There are now 2 points considered outliers on the Residuals vs Leverage plot, but I am opting to keep those as i've already removed three observations, and the other three plots look more stable with residual residuals appearing more homoscedastic and normally distributed than before.

##Intepreting model and communicating results

The results of a poisson regression suggest that both the winter season ($B = -0.027$, $p = 0.62$) and percent cover of creosote bush ($B = -0.042$, $p = 0.20$) do not have a strong effect on lizard counts. However, the summer season does have a significant effect ($B = -0.817$, $p < 0.001$), and so does the percent cover of burrograss ($B = 0.437$, $p < 0.001$). I removed observations 1, 47, and 48 since they appeared as outliers in the residual plots and may have skewed data results inaccurately to show higher correlations when in reality they are weaker. Note, these coefficients are for estimation of log-transformed counts.

#(3) GitHub customization Having created a repository that functions as your homepage on GitHub, go ahead and cus- tomize it in whatever way you think is most helpful for you and for future collaborators to see. At a minimum, you must include at least one image (it does not need to be a photo of you, but it should be properly attributed if you did not take it), a few sentences of text, and a hyperlink to another site (e.g., LinkedIn). For this assignment, all you have to do is include a hyperlink to your GitHub profile

My GitHub profile: <https://github.com/jessalynlc/jessalynlc>