```r
rm(list=ls())
library(tidyverse)
```

```
## -- Attaching packages ---------------------------------------- tidyverse 1.3.2 --
## v ggplot2 3.3.6      v purrr   0.3.4
## v tibble  3.1.8      v dplyr   1.0.10
## v tidyr   1.2.1      v stringr 1.4.1
## v readr   2.1.3      v forcats 0.5.2
## -- Conflicts ------------------------------------------- tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

```r
library(broom)
library(rstanarm)
```

```
## Loading required package: Rcpp
## This is rstanarm version 2.21.3
## - See https://mc-stan.org/rstanarm/articles/priors for changes to default priors!
## - Default priors may change, so it's safest to specify priors, even if equivalent to the defaults.
## - For execution on a local, multicore CPU with excess RAM we recommend calling
##   options(mc.cores = parallel::detectCores())
```

```r
data(roaches, package="rstanarm")
```

```r
## 1
poisson_model <- glm(y ~ roach1 + treatment + senior,
                     data=roaches,
                     offset=log(exposure2),
                     family="poisson")
```

```r
## 2
summary(roaches)
```

```
##        y               roach1         treatment          senior
##  Min.   :  0.00   Min.   :  0.00   Min.   :0.0000   Min.   :0.0000
##  1st Qu.:  0.00   1st Qu.:  1.00   1st Qu.:0.0000   1st Qu.:0.0000
##  Median :  3.00   Median :  7.00   Median :1.0000   Median :0.0000
##  Mean   : 25.65   Mean   : 42.19   Mean   :0.6031   Mean   :0.3053
##  3rd Qu.: 24.00   3rd Qu.: 50.50   3rd Qu.:1.0000   3rd Qu.:1.0000
##  Max.   :357.00   Max.   :450.00   Max.   :1.0000   Max.   :1.0000
##    exposure2
##  Min.   :0.200
##  1st Qu.:1.000
##  Median :1.000
##  Mean   :1.021
##  3rd Qu.:1.000
##  Max.   :4.286
```

```r
grid <- expand_grid(roach1 =seq(min(roaches$roach1), max(roaches$roach1), length.out=5),
                    treatment=c(min(roaches$treatment), max(roaches$treatment)),
```

```
                   senior=c(min(roaches$senior), max(roaches$senior)),
                   exposure2=seq(min(roaches$exposure2), max(roaches$exposure2), length.out=10))

ggplot(data=augment(poisson_model, newdata=grid),
       aes(x=treatment, y=.fitted)) +
  geom_jitter(width=0.05, height=0.0, alpha=0.5) +
  geom_smooth(method="loess", formula="y~x", se=F) +
  labs(title="Predicted Roach Count vs. Treatment",
       y="Fitted Values",
       x="Treatment (Y/N)")
```

```
## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : pseudoinverse used at -0.005
```

```
## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : neighborhood radius 1.005
```
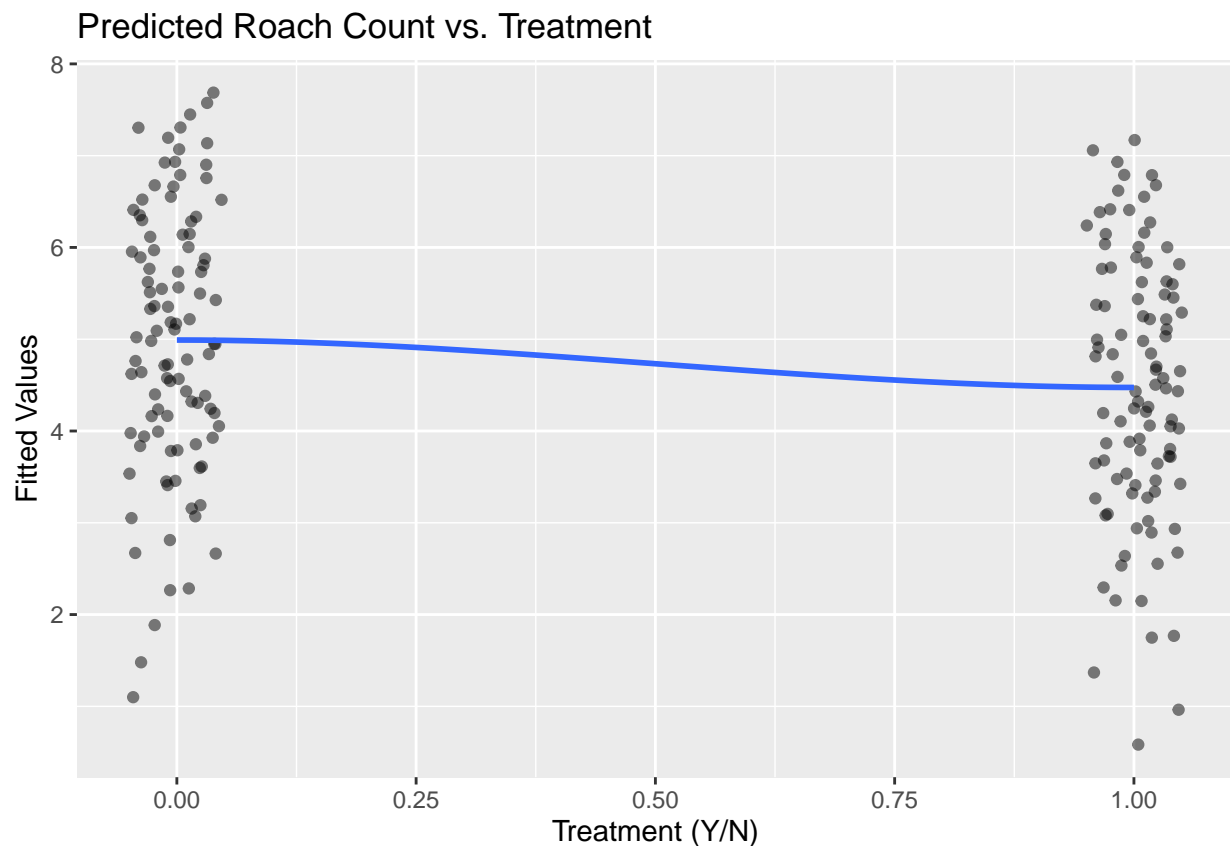
```
## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : reciprocal condition number 0
```

```
## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : There are other near singularities as well. 1.01
```
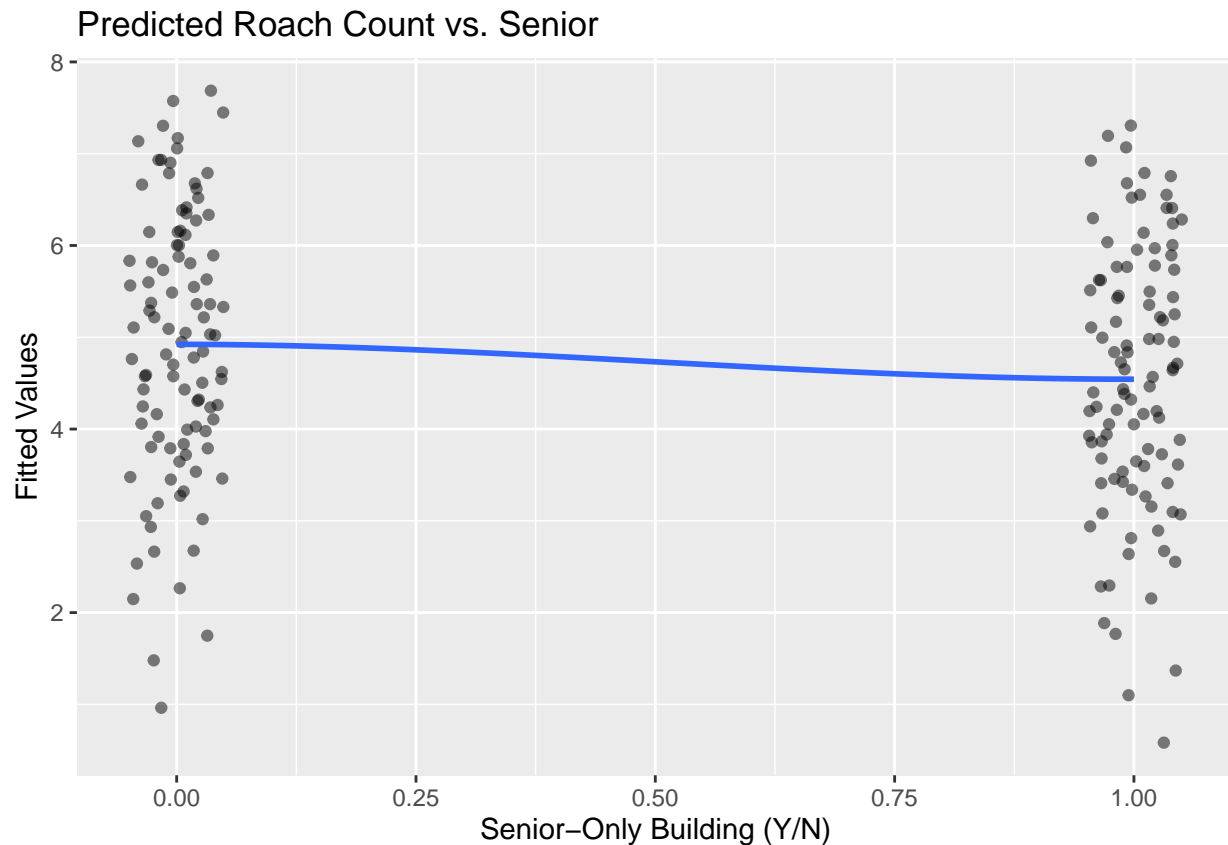
```
ggplot(data=augment(poisson_model, newdata=grid),
       aes(x=senior, y=.fitted)) +
  geom_jitter(width=0.05, height=0.0, alpha=0.5) +
  geom_smooth(method="loess", formula="y~x", se=F) +
  labs(title="Predicted Roach Count vs. Senior",
       y="Fitted Values",
       x="Senior-Only Building (Y/N)")
```

```
## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : pseudoinverse used at -0.005

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : neighborhood radius 1.005

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : reciprocal condition number 0

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : There are other near singularities as well. 1.01
```



```
ggplot(data=augment(poisson_model, newdata=grid),
       aes(x=treatment, y=.fitted)) +
  geom_jitter() +
```

```
  facet_grid(cols = vars(roach1)) +
  geom_smooth(method="loess", formula="y~x", se=F) +
  labs(title="Predicted Roach Count vs. Treatment\nFaceted by Roach Count before Treatment",
       x="Treatment (Y/N)",
       y="Fitted Values")
```

Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
parametric, : pseudoinverse used at -0.005

Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
parametric, : neighborhood radius 1.005

Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
parametric, : reciprocal condition number 0

Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
parametric, : There are other near singularities as well. 1.01

Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
parametric, : pseudoinverse used at -0.005

Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
parametric, : neighborhood radius 1.005

Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
parametric, : reciprocal condition number 0

Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
parametric, : There are other near singularities as well. 1.01

Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
parametric, : pseudoinverse used at -0.005

Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
parametric, : neighborhood radius 1.005

Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
parametric, : reciprocal condition number 0

Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
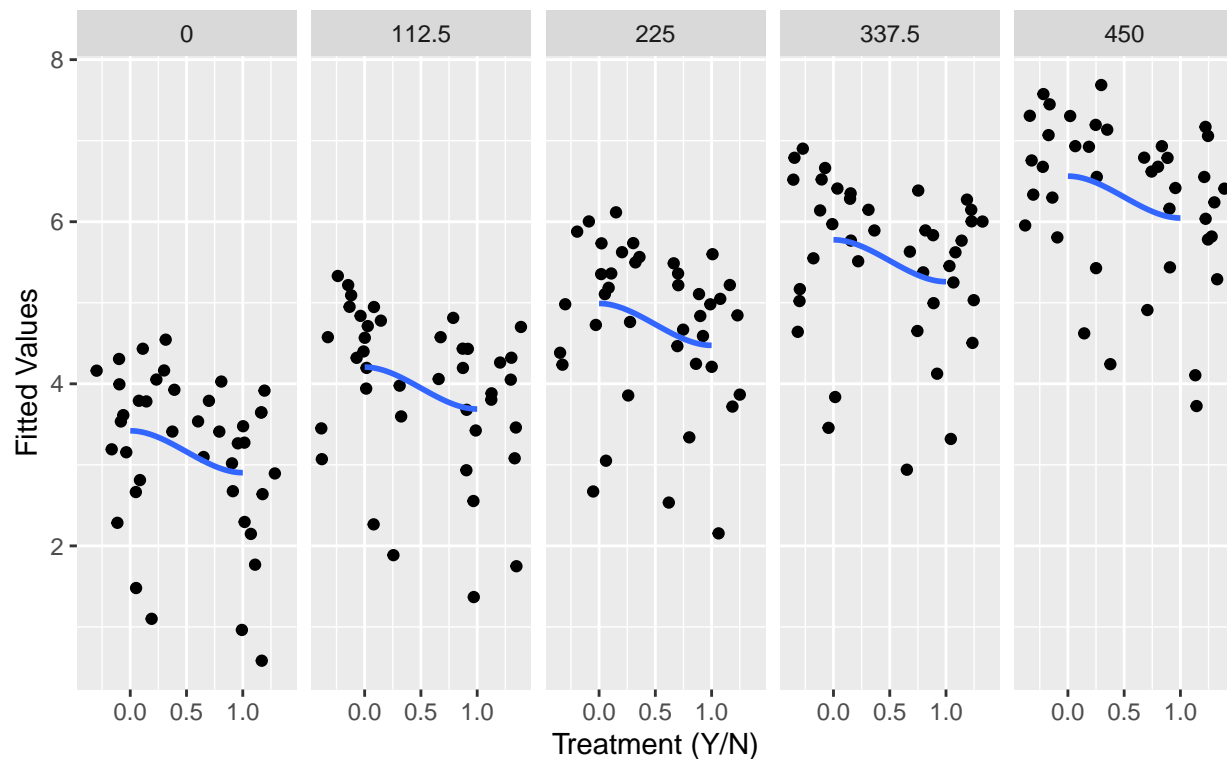## parametric, : There are other near singularities as well. 1.01

Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
parametric, : pseudoinverse used at -0.005

Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
parametric, : neighborhood radius 1.005

Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
parametric, : reciprocal condition number 0

```
## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : There are other near singularities as well. 1.01
```

```
## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : pseudoinverse used at -0.005
```

```
## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : neighborhood radius 1.005
```

```
## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : reciprocal condition number 0
```

```
## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : There are other near singularities as well. 1.01
```


Predicted Roach Count vs. Treatment
Faceted by Roach Count before Treatment

```
ggplot(data=augment(poisson_model, newdata=grid),
       aes(x=senior, y=.fitted)) +
  geom_jitter() +
  facet_grid(cols = vars(roach1)) +
  geom_smooth(method="loess", formula="y~x", se=F) +
  labs(title="Predicted Roach Count vs. Senior\nFaceted by Roach Count before Treatment",
       x="Senior-Only Building (Y/N)",
       y="Fitted Values")
```

```
## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : pseudoinverse used at -0.005

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : neighborhood radius 1.005

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : reciprocal condition number 0

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : There are other near singularities as well. 1.01

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : pseudoinverse used at -0.005

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : neighborhood radius 1.005

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : reciprocal condition number 0

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : There are other near singularities as well. 1.01

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : pseudoinverse used at -0.005

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : neighborhood radius 1.005

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : reciprocal condition number 0

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : There are other near singularities as well. 1.01

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : pseudoinverse used at -0.005

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : neighborhood radius 1.005

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : reciprocal condition number 0

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : There are other near singularities as well. 1.01

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : pseudoinverse used at -0.005
```

```
## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : neighborhood radius 1.005

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : reciprocal condition number 0

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : There are other near singularities as well. 1.01
```

Predicted Roach Count vs. Senior
### Faceted by Roach Count before Treatment
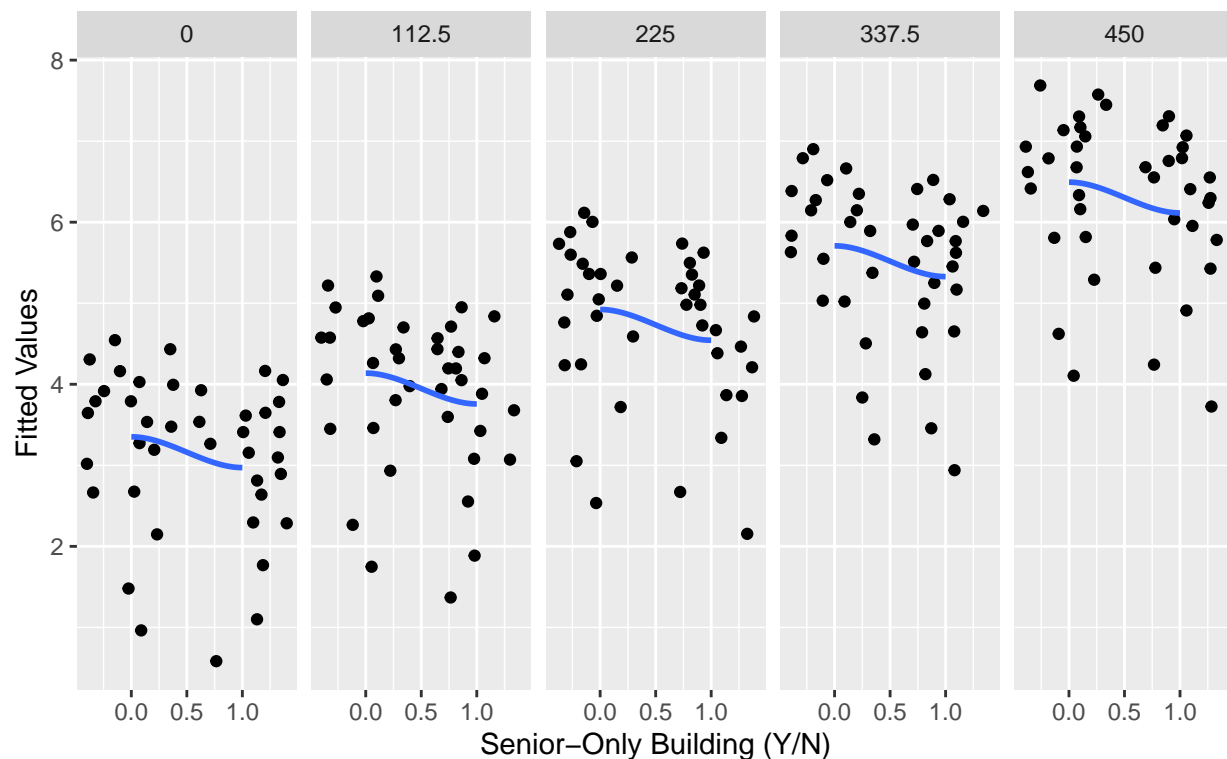
Interpretation: The plot of the fitted values vs. treatment displays that receiving a treatment results in a decrease in the the predicted number of roaches that were trapped, although the decrease is a slight decrease. The plot of the fitted values vs. the building being a senior-only building displays the same relationship. The coplots that display these two simple relationships controlling for the roach count before treatment demonstrate the importance of conditioning on this preliminary count. Across all five facets of original roach count, the relationship between receiving a treatment and the fitted values is more negative than when examining the relationship between receiving a treatment and the fitted values on their own. The same is true of the relationship between fitted values and a building being a seniors-only building. Since no interaction is specified in the model, the relationship is similar in each facet in both coplots.

```
## 3.
ggplot(data=data.frame(pearson_residuals = residuals(poisson_model, type="pearson"),
                       log_fitted_values = log(fitted.values(poisson_model) + 1)),
       aes(x=log_fitted_values, y=pearson_residuals)) +
  geom_point() +
  labs(title="Residual Plot",
```

```
            y="Pearson Residual",
            x="Fitted Values (Logged)")
```

Residual Plot
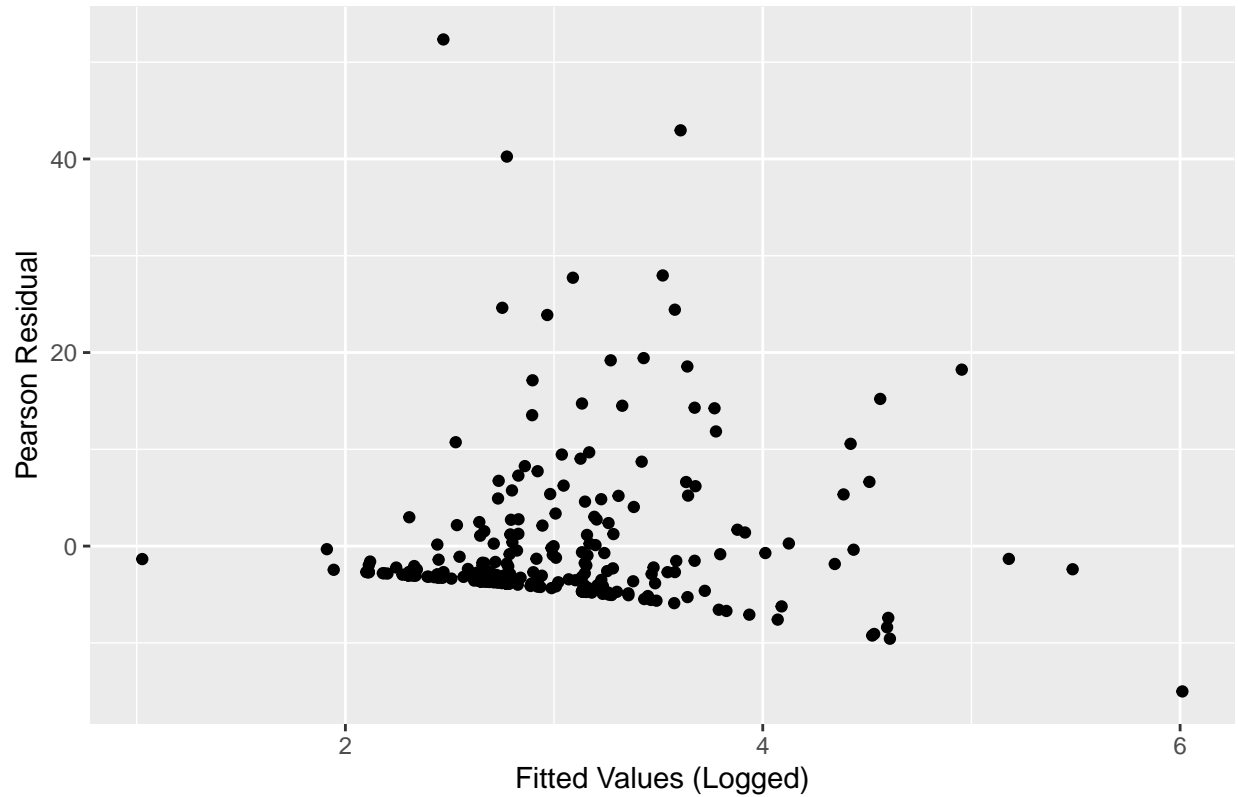
Answer: In the theory behind the poisson model, the variance is equal to the expected value. So, in a plot of the residuals vs. the fitted values we expect the variance of the residuals to increase as the fitted values increase. This is not the case here as the highest variance occurs at moderate fitted values. Overdispersion may be the cause and a disperson parameter could be incorporated to solve it.