

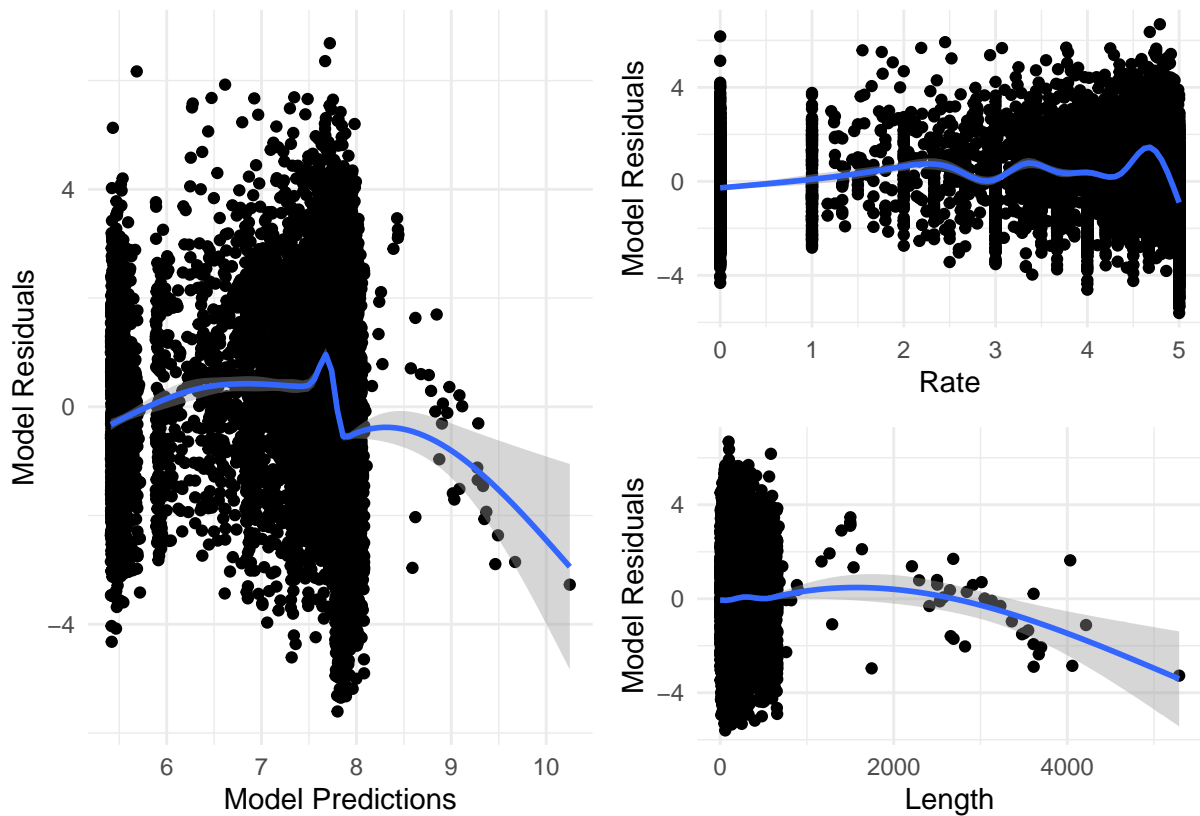
Evaluating the Classical Linear Model Assumptions

Q1.3 Linear Conditional Expectation

First the linear regression model was created.

```
model_12 <- lm(log(views) ~ rate + length, data= videos)
```

To assess whether there is linear conditional expectation we can examine two types of plots, model residuals vs predicted values and model residuals vs each feature's values.



Analysis of the left most plot reveals some non-linear relationships in the data. Beyond prediction values of ~ 9 the model's residuals become negative only, meaning the linear model predicts values higher than the actual data points. This indicates some non-linear behavior in the joint density function that is not captured in the linear model.

Analysis of the upper right plot reveals residuals fluctuating about a mean of zero, resembling roughly linear conditional expectation with the Rate feature. However, analysis of the lower right plot again reveals the conditional expectation of the residuals drops below zero at Lengths greater than 3000 seconds.

Therefore, we conclude that the assumption of Linear Conditional Expectation is not met because of the non-linear behavior seen in the residual plots above.