

# R Notebook

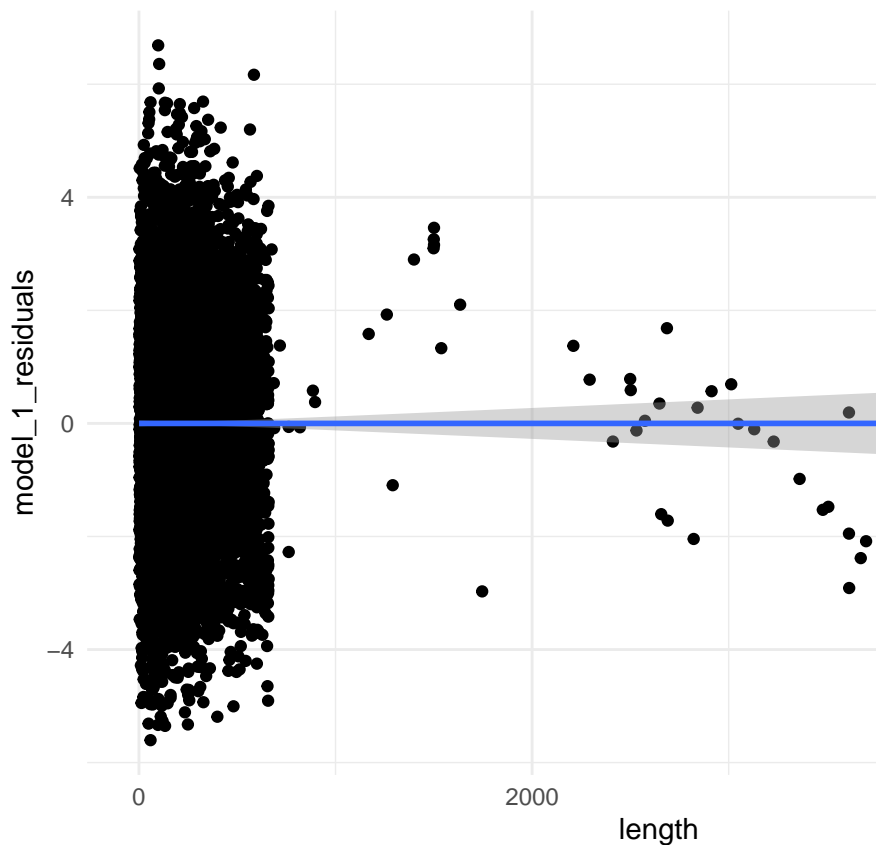
```
knitr::opts_chunk$set(echo = FALSE)
knitr::opts_chunk$set(message=FALSE)
```

```
library(tidyverse)
```

```
## -- Attaching packages ----- tidyverse 1.3.2 --
## v ggplot2 3.4.0      v purrr   0.3.5
## 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()
```

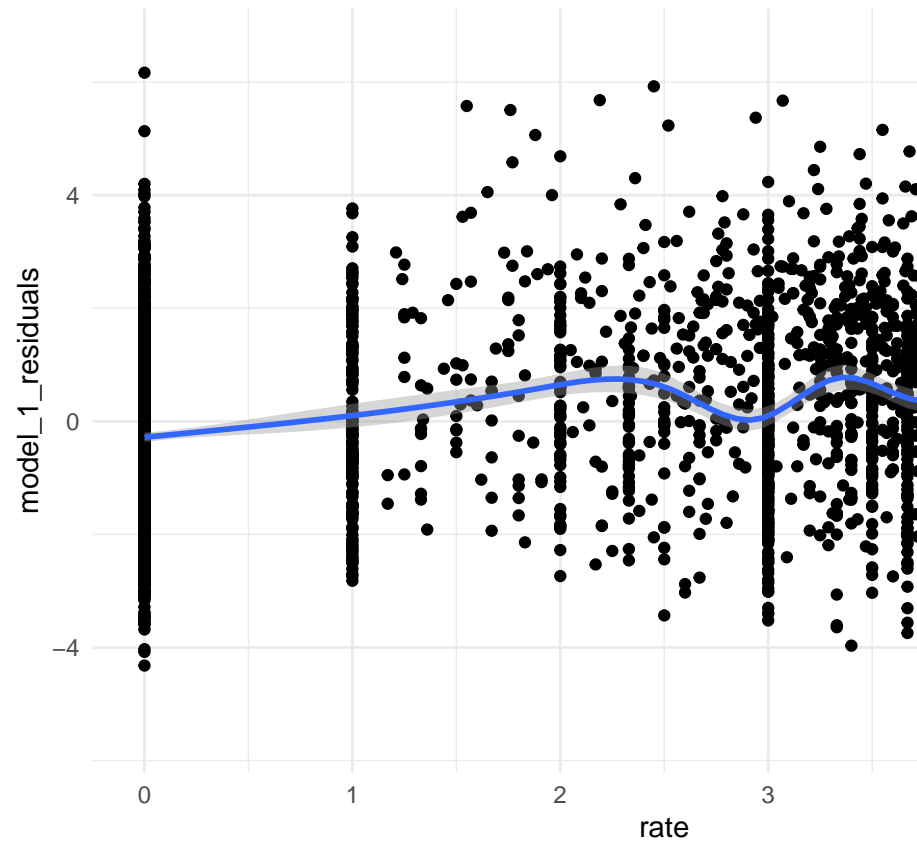
```
library(patchwork)
```

```
knitr::opts_chunk$set(message=FALSE)
theme_set(theme_minimal())
```



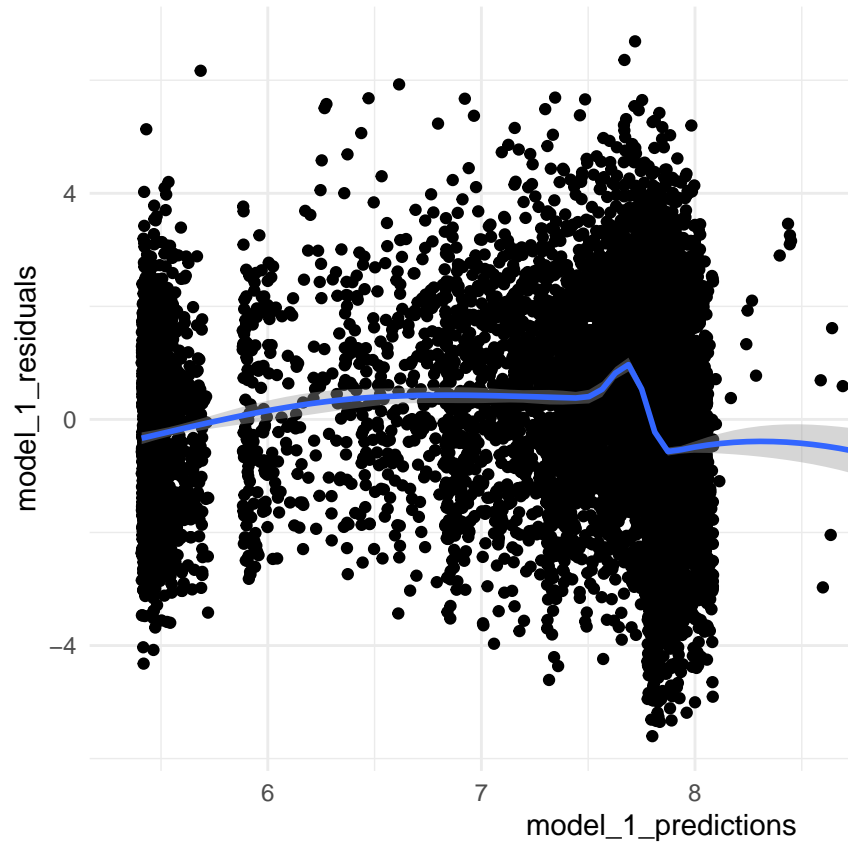
Below is the plot for: Feature length vs Residuals

As we can see in the above plot, the model residuals are clearly distributed about zero where bulk of the points are lie. So, we believe that the linear conditional expectation condition is not violated based on what we see within this feature.



Below is the plot for: Feature rate vs residuals

In the plot above, it doesn't appear that the model predictions are too high or low depending on the rate. We do see lots of points cluster on exact rating value, which is likely because the video had a single rating score. However, if we look at the region of the plot where majority of points lie, we see that the residuals remain distributed about zero. So, we believe that the linear conditional expectation condition is not violated based on what we see here.



Below is the plot for: Model predictions vs residuals

Looking at the plot above, it does look like the model is slightly under predicting for range where the predictions are supposed to be very small or large, but it seems to be very close to zero (not over or under predicting) where bulk of points lie on the plot. It is hard to say whether this is because those points are outliers, or if there is actual non-linearity to be expected, when it comes to smaller/larger predictions, unless we have more data. So, we believe that the linear conditional expectation condition is not violated.

So, we conclude that the Linear Conditional Expectation is met.