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()
```
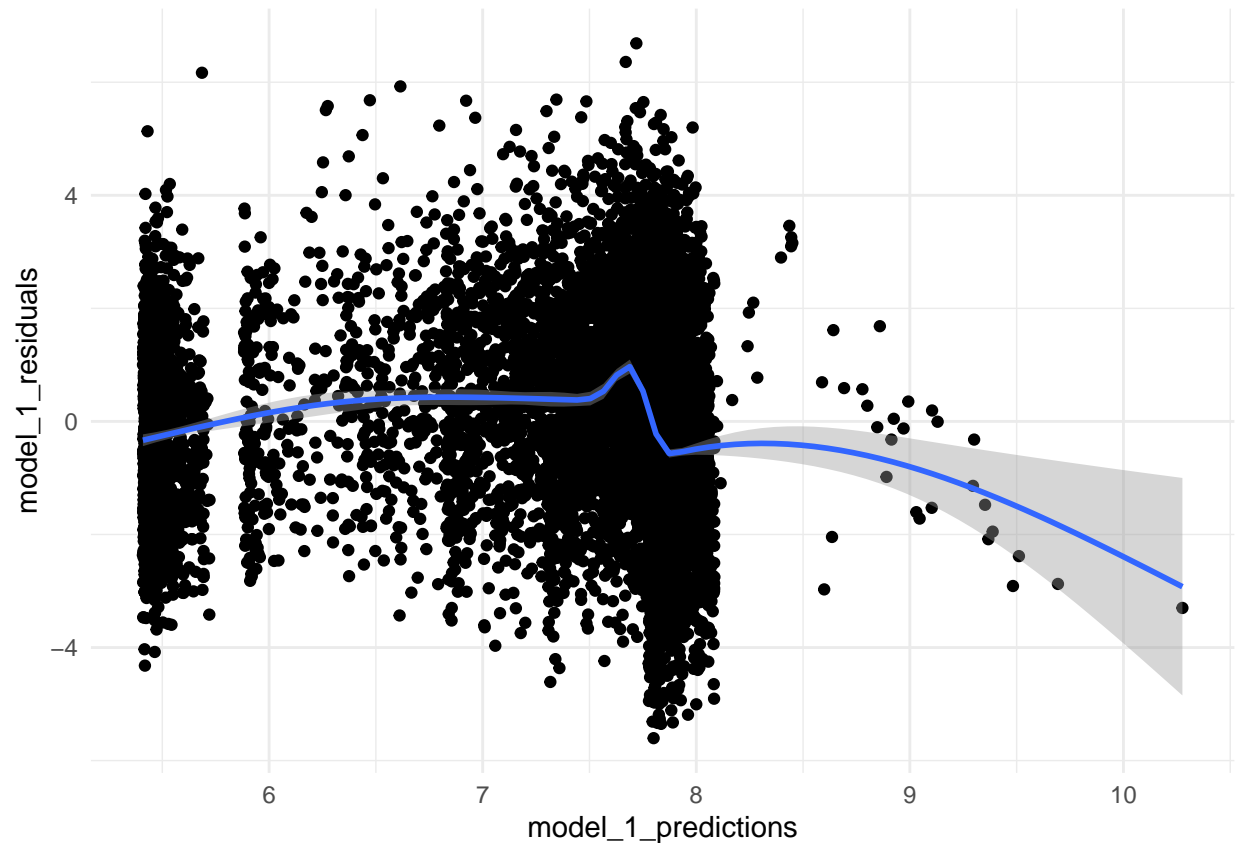
```
library(patchwork)

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

Building the model: lm(log(views) ~ length + rate, data = df)

Plotting the model predictions vs residuals to perform an ocular test to verify homoscedasticity.

From the above plot we can see that the model tends to under-predict for larger predictions, leading to negative residuals for upper range of the plot on the x.

Conducting the Breusch-Pagan test to look for evidence of heteroscedasticity.

```
##
##  studentized Breusch-Pagan test
##
## data:  model_1
## BP = 127.16, df = 2, p-value < 2.2e-16
```

We get a very small p-value, so we have strong evidence to reject the null hypothesis and conclude that there is some heteroscedasticity in the way residuals are distributed.