HW week 12

w203: Statistics for Data Science

w203 teaching team

OLS Inference

The file videos.txt contains data scraped from Youtube.com.

1. Fit a linear model predicting the number of views (views), from the length of a video (length) and its average user rating (rate).

2. Using diagnostic plots, background knowledge, and statistical tests, assess all 6 assumptions of the CLM. When an assumption is violated, state what response you will take.

3. Generate a printout of your model coefficients, complete with standard errors that are valid given your diagnostics. Comment on both the practical and statistical significance of your coefficients.