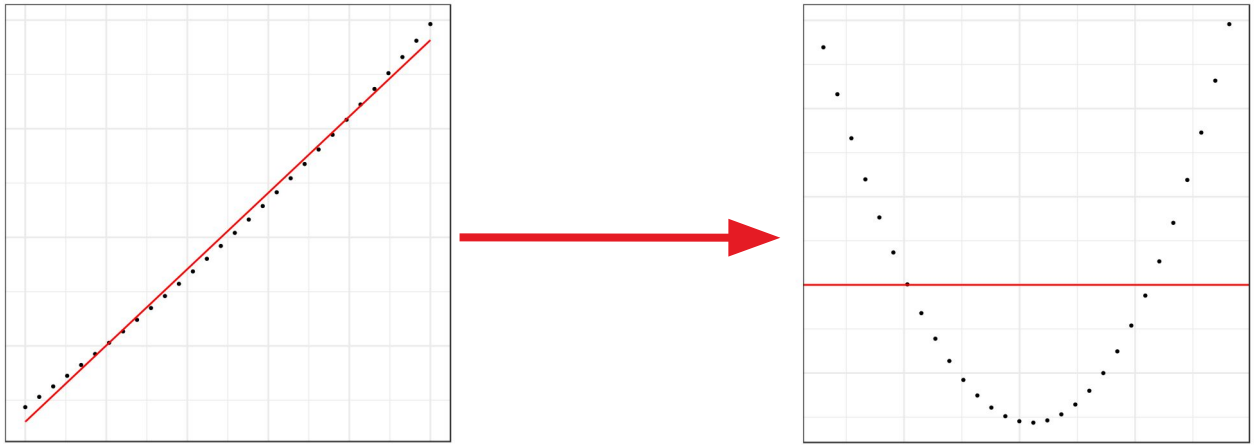
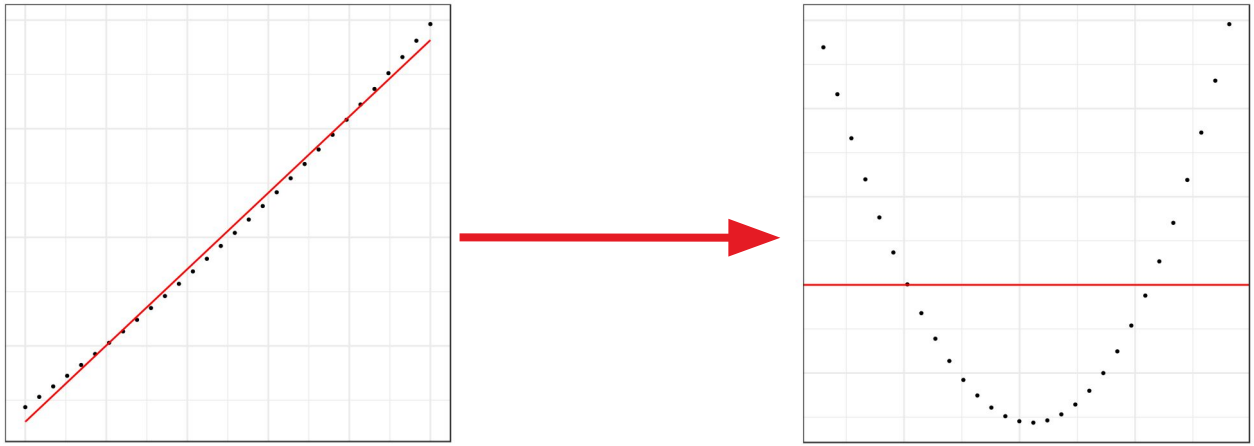
Visualization of some common LINE violations

L - Linearity violated

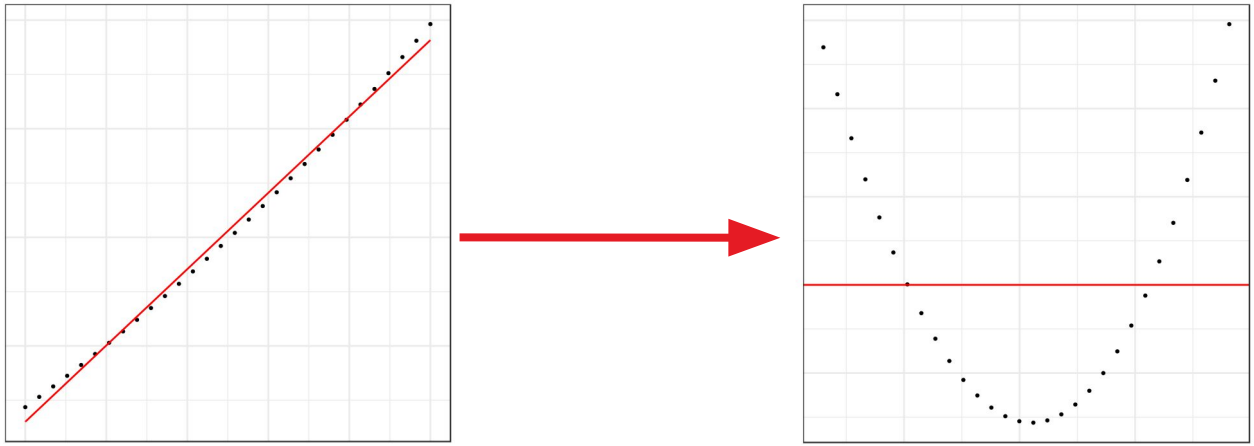


L - Linearity violated



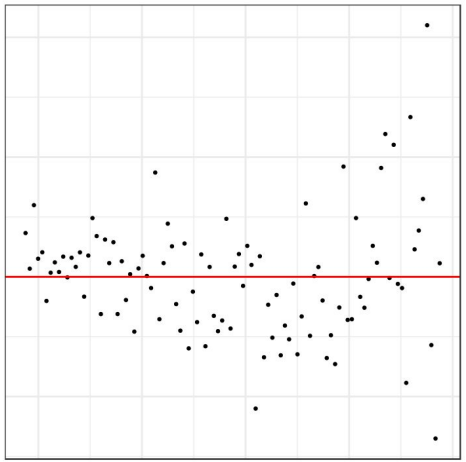
Looking at a residual plot magnifies nonlinearities in our model. The residuals contain the “leftover” relationships after our model has been applied. If there are patterns in the residuals, that means there are patterns in our data that our model have not accounted for.

L - Linearity violated



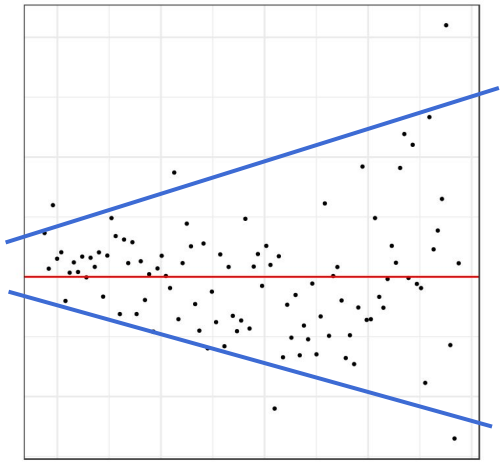
The Fix? We have unaccounted for relationships in our data. The fix could be adding nonlinearities to either an *x*-variable, our *y*-variable, or both. The next step is to perform additional EDA on our data and tease out individual relationships between the *x*s and *y*.

E - Equal variances violated - heteroscedasticity!



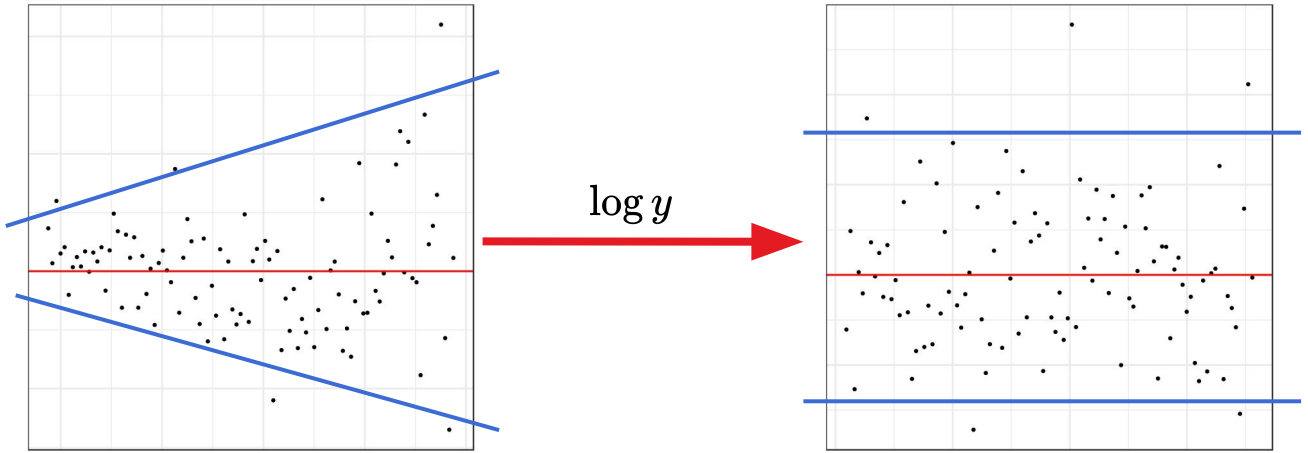
Heteroscedasticity is characterized by a **fan-shape** (although it usually looks more like a horn or a trumpet).

E - Equal variances violated - heteroscedasticity!



Heteroscedasticity implies our model is getting worse for larger predicted values. That is, our model is only “good” for only a portion of our data.

E - Equal variances violated - heteroscedasticity!



The Fix? Simple! Taking the (natural) logarithm of our *y*-variable and rerunning the regression usually solves this problem. This usually solves the **N assumption** as well!

