

# WU #16 - Outliers 1

Math 158 - Jo Hardin

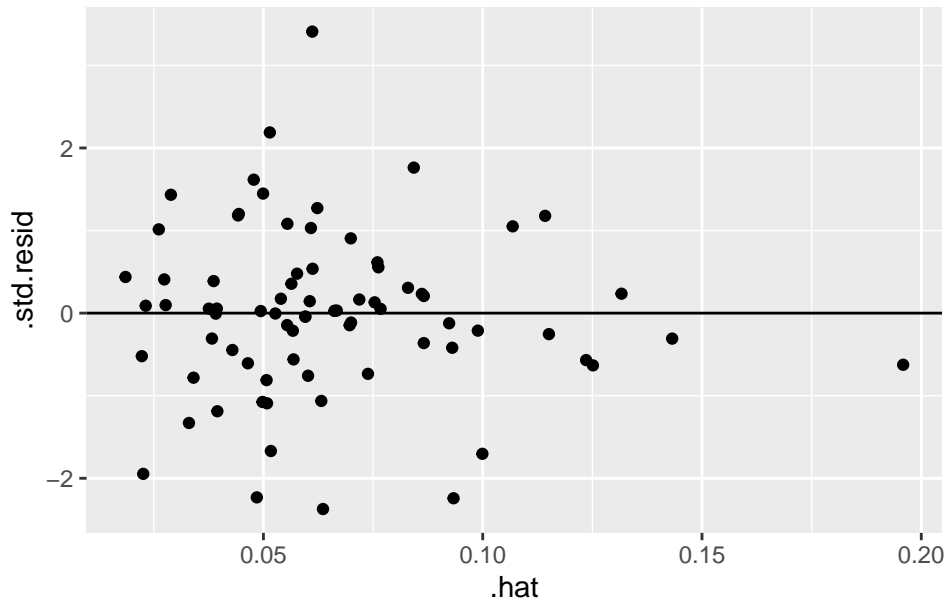
Tuesday 3/29/2022

Name: \_\_\_\_\_

Names of people you worked with: \_\_\_\_\_

The following plot describes leverage and studentized residuals. Does it seem like an of the observations will be particularly influential in the model? Explain.

Outliers for Starbucks Data



**Solution:**

The point at the top of the plot has a high residual but low leverage. The point at the right of the plot has high leverage but a very low residual.

Notice that Cook's distance can be written as:

$$D_i = \frac{\sum_{j=1}^n (\hat{Y}_j - \hat{Y}_{j(i)})^2}{pMSE} = \frac{e_i^2}{pMSE} \left[ \frac{h_{ii}}{(1 - h_{ii})^2} \right]$$

In order to get influence seen by Cook's distance (for example), both  $e_i$  and  $h_{ii}$  should be moderate to large. So probably neither item is particularly influential.

**Outliers for Starbucks Data**

