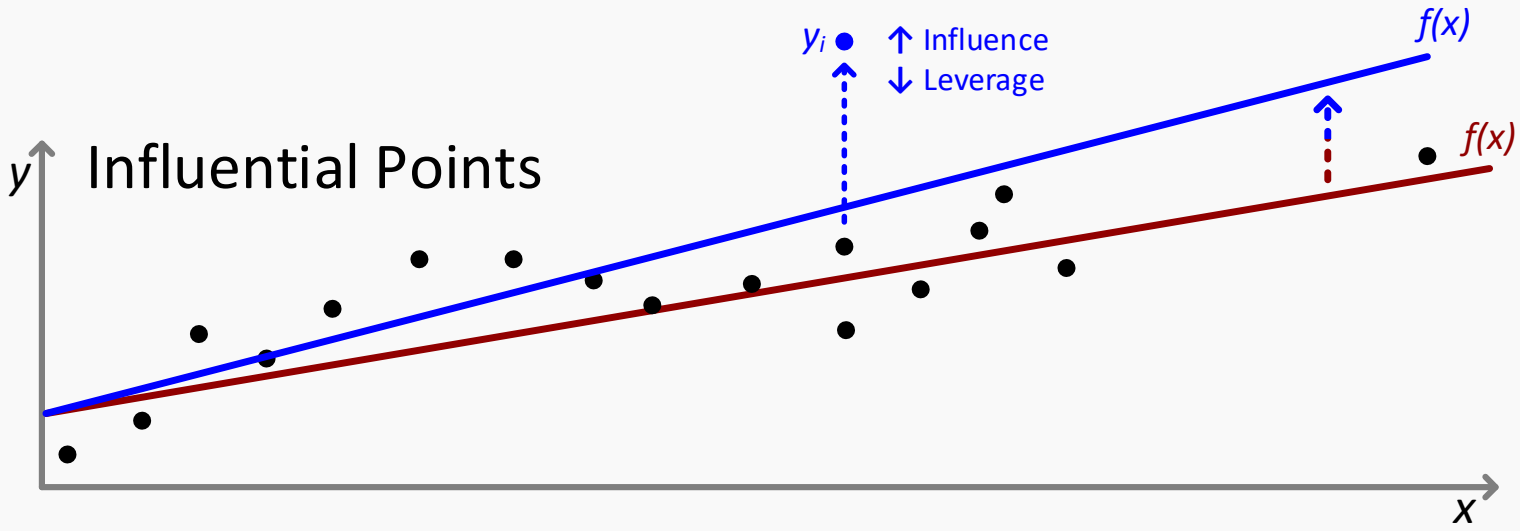
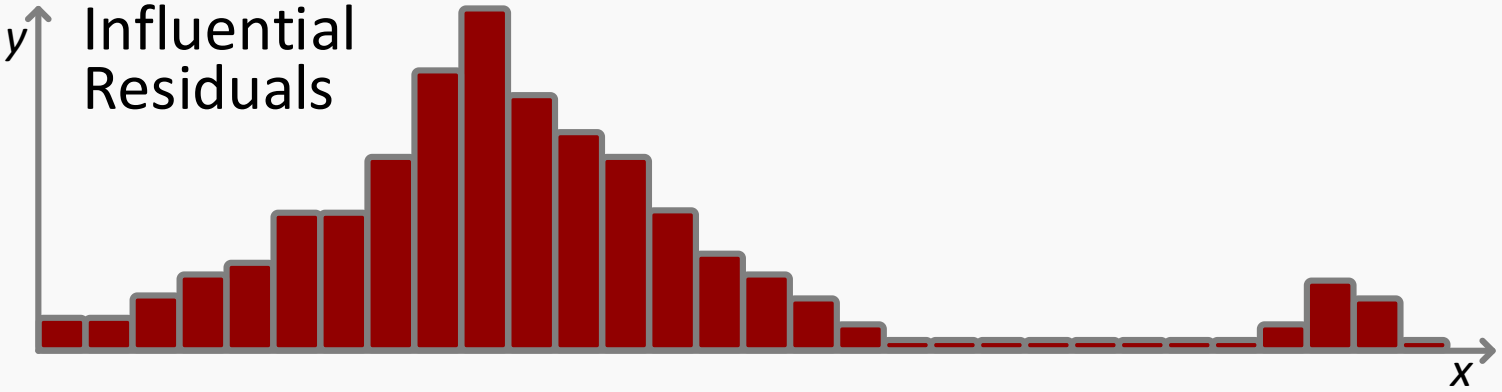
influential ≋ points

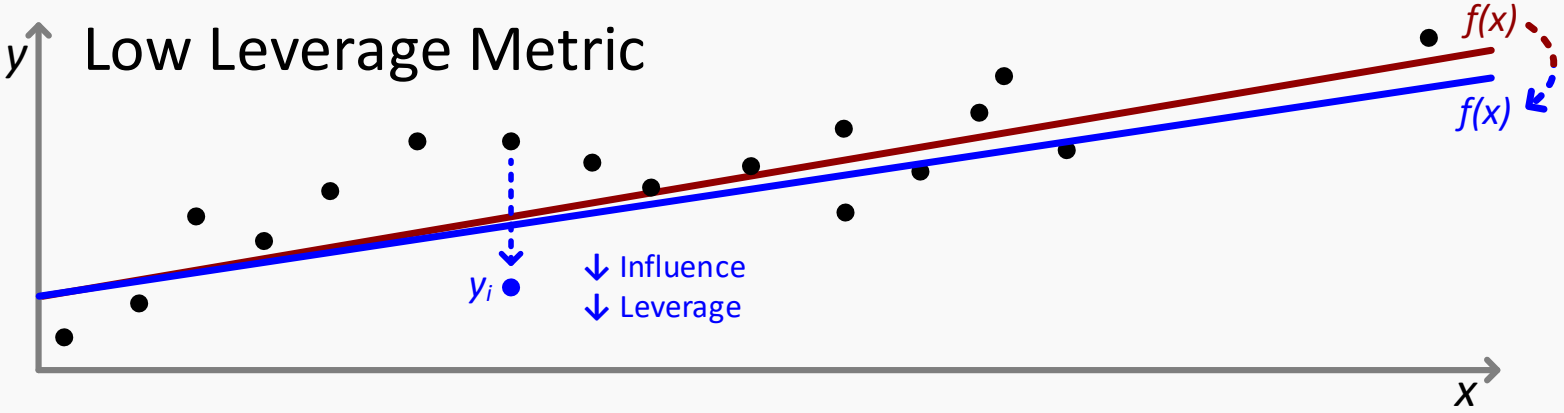
An **Influential Point** changes a model's predicted values significantly when omitted or altered:



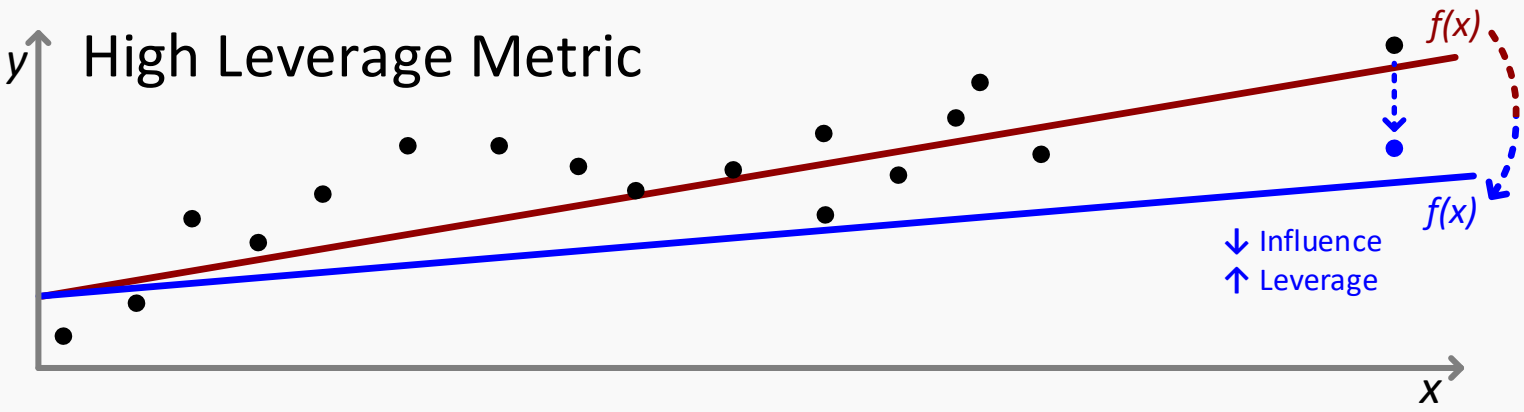
Equally, **Influential Points** significantly impact a plot of residuals seen in the histogram below:



Alternatively, **Leverage** is where a point can be altered in multiple ways without having a significant impact on the model's predictions or the plotted residual distribution as seen in the illustration below:



However, depending on the location of the point in question, **Leverage** can be substially increased. Any alteration of such a point will cause the model to alter in a more significant nature seen below:



Formally, if point x_i is moved, and $f(x_i)$ moves proportionally, the proportionality constant is referred to as the leverage of point i.

Leverage depends on x_i but does not depend on y_i. Leverage depends on how far away x_i is from the mean of the points x_i *'s*.

It is important to note the **Influential Points** do not necessarily have a high amount of **Leverage**. Alternatively, High **Leverage** points are not necessarily **Influential Points**.

Ultimately, the **Leverage** of point x_i measures the impact of x_i on $f(x_i)$.

Investigating the nature of outliers typically involves asking questions such as:

- Is the data correct and represented as such?
- Is the model complete and performing sufficiently?