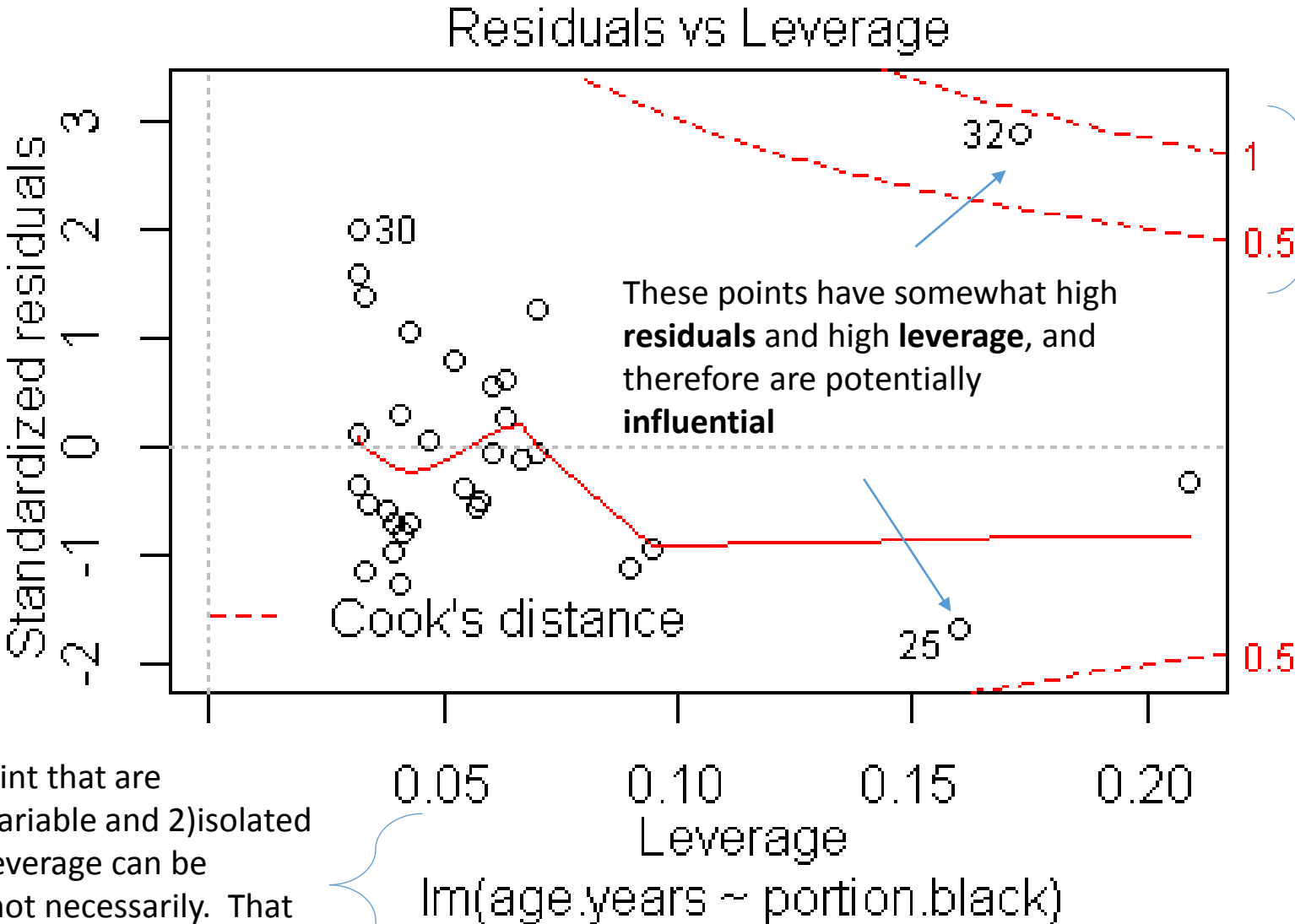**Diagnostics for the potential impact of "influential" points and outliers:**
Residuals vs. Leverage Plots and Cook's Distance
R code: plot(fitted.model, which = 5)

Dashed red lines represent cut-off values for **Cook's distance**.
Points outside of the 0.5 line, and especially 1.0, are potentially **"influential"**. A data point is influential (has "high influence") if your results would be different if you had not collected that data point.

A data point's "**residual**" is the distance between it and a fitted regression line. A data point with a large residual lies far from the line. A data point with a small residual (near 0) lies near the line. Data points can have large residuals and low **leverage**, while **influential** data points always have large residuals
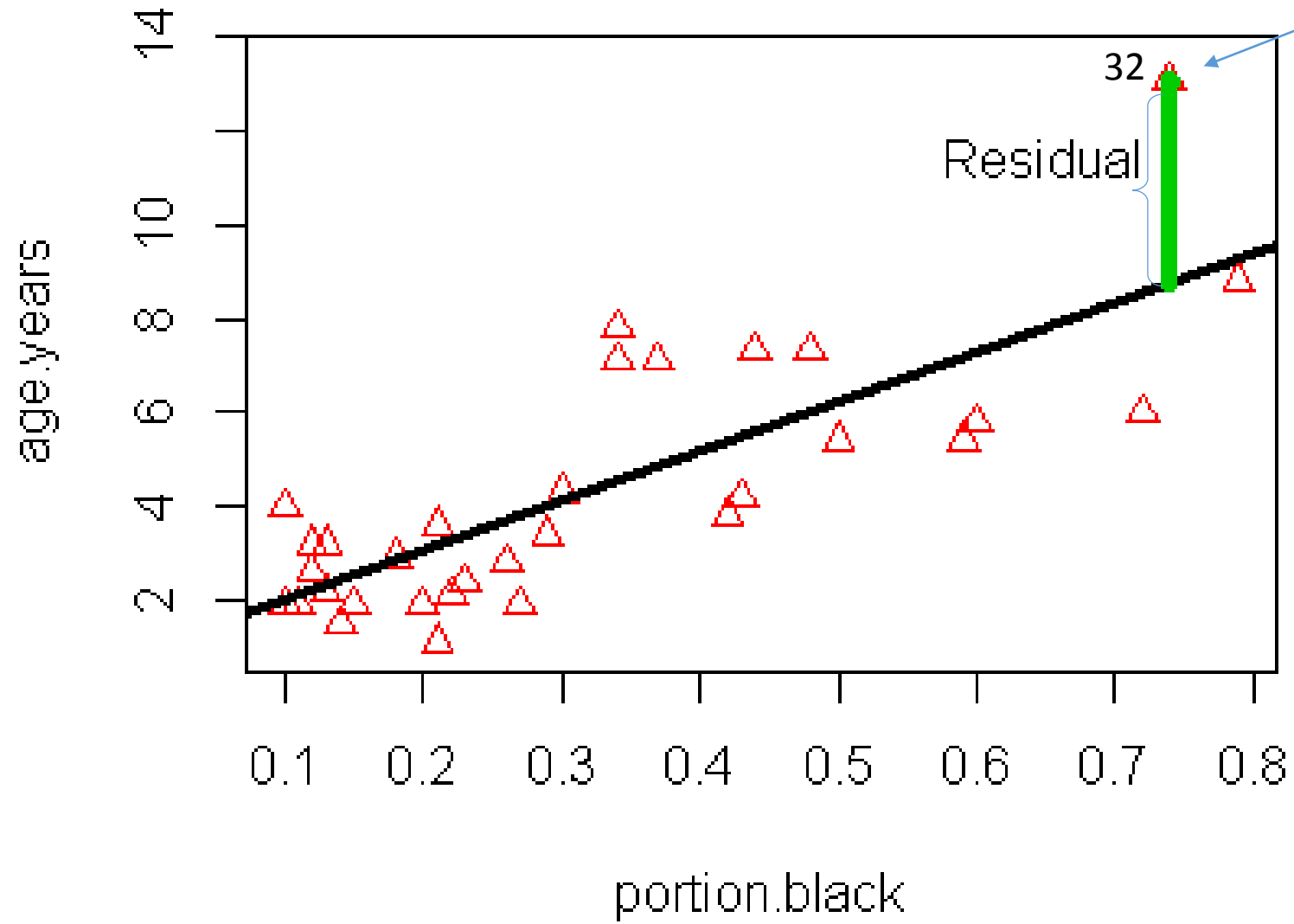

Residuals vs Leverage

These points have somewhat high **residuals** and high **leverage**, and therefore are potentially **influential**

Cook's distance

lm(age.years ~ portion.black)

Leverage increases for data point that are 1)far from the mean of the x-variable and 2)isolated from other data point. High leverage can be related to high influence, but not necessarily. That is, points with high leverage are not "outliers" and are not necessarily "influential"

Raw data use for diagnostic plot

This is point 32, which is indicated as being potentially **influential** (high **residual**, high **leverage)**