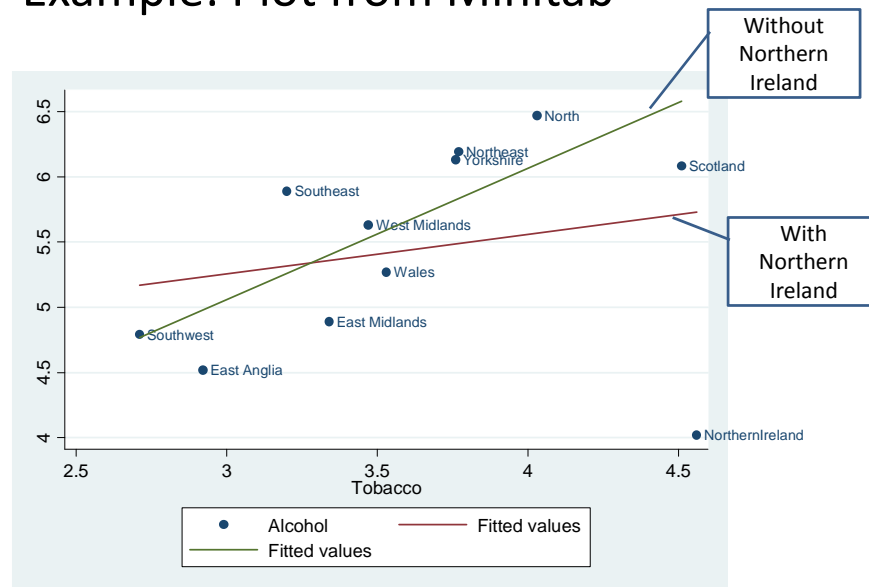


Example

Region	Alcohol	Tobacco
North	6.47	4.03
Yorkshire	6.13	3.76
Northeast	6.19	3.77
East Midlands	4.89	3.34
West Midlands	5.63	3.47
East Anglia	4.52	2.92
Southeast	5.89	3.2
Southwest	4.79	2.71
Wales	5.27	3.53
Scotland	6.08	4.51
Northern Ireland	4.02	4.56

Example: Plot from Minitab



```

> BritRegions <- lm(Alcohol~Tobacco, data=BritishRegions)
> summary(BritRegions)
Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)   4.3512     1.6067   2.708  0.0241 *
Tobacco        0.3019     0.4388   0.688  0.5087
Residual standard error: 0.8196 on 9 degrees of freedom
Multiple R-squared:  0.04998,    Adjusted R-squared:  -
0.05557
F-statistic: 0.4735 on 1 and 9 DF,  p-value: 0.5087

```

For leverage, flag values with $h_i > 4/11 = .363$

None are flagged b/c Scotland and No. Ireland are *both* far from the center.

```

> options(digits=3)
> hatvalues(BritRegions)
  1      2      3      4      5      6      7      8      9     10     11
0.1395 0.0967 0.0975 0.1131 0.0972 0.2306 0.1410 0.3273 0.0931 0.3188 0.3451

```

Standardized residuals: Flag if absolute value > 2; No. Ireland is moderate

```

> rstandard(BritRegions)
  1      2      3      4      5      6      7      8      9     10     11
1.186  0.826  0.900 -0.608  0.297 -0.992  0.754 -0.564 -0.188  0.543 -2.575

```

Studentized residuals: Flag if absolute value > 2; No. Ireland is extreme

```

> rstudent(BritRegions)
  1      2      3      4      5      6      7      8      9     10     11
1.218  0.810  0.889 -0.586  0.281 -0.990  0.734 -0.542 -0.178  0.520 -4.732

```

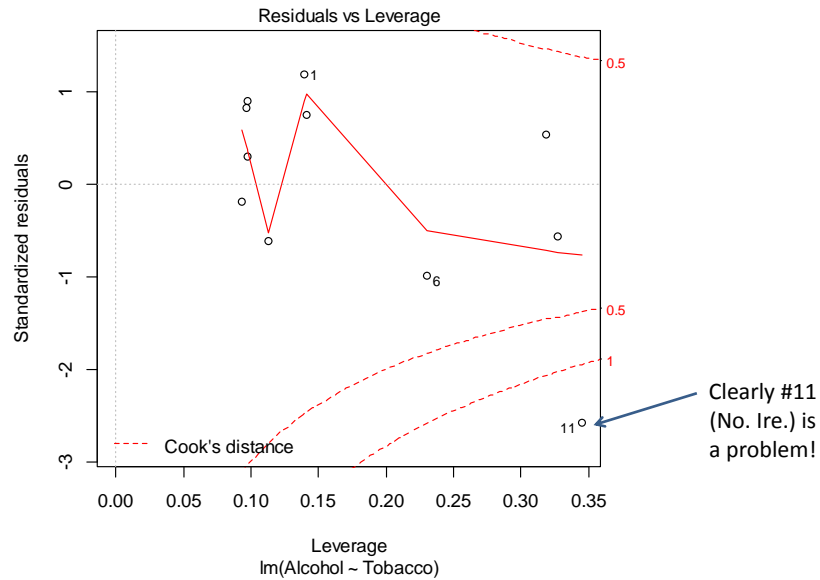
Cook's distance: Flag if $D > 0.5$: No. Ireland is very extreme!

```

> cooks.distance(BritRegions)
  1      2      3      4      5      6      7      8      9     10
0.11410 0.03652 0.04373 0.02360 0.00474 0.14733 0.04665 0.07749 0.00182 0.06892
 11
1.74723

```

Plot(BritRegions) produces 4 plots, including this one



Remove Northern Ireland

```
> Nolreland=subset(BritishRegions,subset=Region!="NorthernIreland")
```

```
> Nolre<-lm(Alcohol~Tobacco,data=Nolreland)
```

Coefficients:

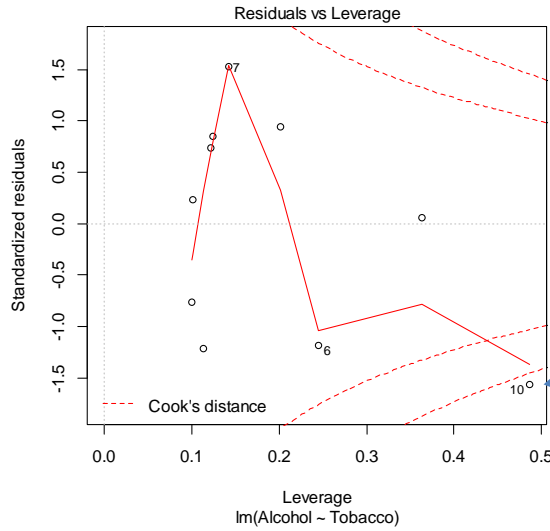
	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	2.041	1.001	2.04	0.0759 .
Tobacco	1.006	0.281	3.58	0.0072 **

Residual standard error: 0.446 on 8 degrees of freedom

Multiple R-squared: 0.615, Adjusted R-squared: 0.567

F-statistic: 12.8 on 1 and 8 DF, p-value: 0.00723

Without Northern Ireland, Scotland is a problem



Now Scotland has high leverage and high Cook's D