Goodness of Fit: R^2 vs. Adjusted R^2

You can think of the difference between R^2 and adj. R^2 this way.

- The R^2 assumes that *every* independent variable in the model helps to explain variation in the DV. So, it tells you the percentage of explained variation as if *all* IVs in the model affect the DV (as if each IV passes the t-test.)
- The adj. R^2
  - tells you the percentage of variation explained by only those IVs that truly affect the DV (only those IVs that pass the t-test) **AND**
  - penalizes you for adding independent variable(s) that do not belong in the model
- So, you can expect that the value of the adj. R^2 will be ≤ value of R^2.

Notice what happens to R^2 and adj. R^2 as add an independent variable.

Model #1.

Dependent Variable: SALARY
Included observations: 20

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	73.98441	4.010613	18.44716	0.0000
EXPERIENCE (years)	0.945455	0.470826	2.008077	0.0608
GENDER (0, 1)	-12.2012	2.854127	-4.274932	0.0005

R-squared	**0.750652**	F-statistic	25.5889
Adjusted R-squared	**0.721317**	Prob(F-statistic)	0.000007

Model #2. (Add SCORE to Model #1)

Dependent Variable: SALARY
Included observations: 20

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	62.90383	16.68642	3.769763	0.0017
EXPERIENCE (years)	0.950914	0.478424	1.987597	0.0642
GENDER (0, 1)	-11.06503	3.340895	-3.311995	0.0044
SCORE (0 -100)	0.124637	0.18201	0.684779	0.5033 ←

R-squared	**0.797752**	F-statistic	16.68265
Adjusted R-squared	**0.67233**	Prob(F-statistic)	0.000035

SALARY = annual salary in $1000
GENDER = 1 if female, 0 if male
EXPERIENCE = years in industry
SCORE = score on programmer's aptitude test (0 – 100)