started class with finishing up the aclfest.csv exercise. Try to predict who will be performing at ACL with information on what other music fests bands have performed at.

One of the groups found through step-wise selection that Coachella, Lollapalooza, and OutsHL were positive predictors of ACL headliners and BonnHL was a negative predictor.

exp(-2.77)/{1+exp(-2.77)} = 0.05897
Test the intercept. We find that if bands went to no other music fests, they only have around a 6% chance of performing at ACL.

When using the glm function, you have to make sure you use binomial outcomes.

lmk function → $P(y_i = 1) = e^{\Psi_i} / (1 + e^{\Psi_i})$
linear predictor → $\Psi_i = B_0 + B_1x_{i1} + B_2x_{i2} + \ldots. + B_px_{ip}$

binary predictors and binary response →

	Dependent var 1	Dependent var 2
Independent var 1		
Independent var 2		

step-wise selection isn't perfect. It could remove a variable that you may care about.

way to assess if a variable contributes significantly–shuffle it or use a permutation test

trends – overall movement over time.
- key strategy = regress on a time index.
- Let t be the period number.
- $y_t = B_0 + B_1t + Ɛt$

time series = linear trend*time index
- B_1 = rise/run = change y/ change x

seasonalities – variation within that year
- most iphones sold in December. Least sold in January
- introduce seasonal dummy variables