## Main Instructions

1. Open the California Schools dataset (use the ".sav" version if it is easier)
2. Test whether reading performance differs significantly between the "Los Angeles" county and the "San Diego" count. (Tip: first recode county into a numeric variable using Transform - Automatic Recode; then Analyze - means - independent groups t-test)
3. Define the null and alternative hypothesis
4. Calculate cohen's d
5. Interpret the confidence intervals

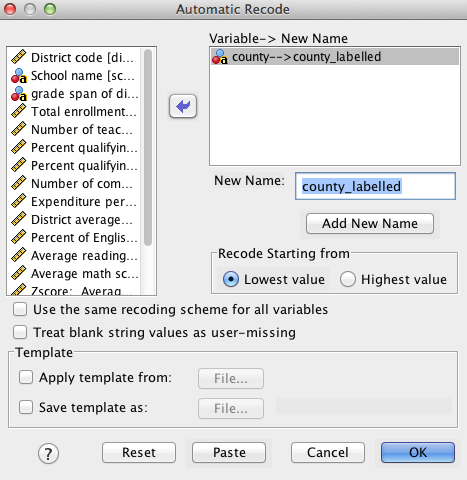
## G-Power Homework Exercise

1. Install G-Power 3
2. Work out the statistical power of the San Diego versus Los Angeles analysis assuming the sample group means are the population group means.
3. In G-Power 3 work out what sample size you would need to examine differences between two groups when you are expecting a half standard deviation difference between the group means.

# Answers

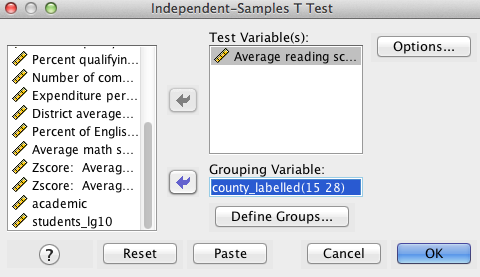
## Test whether reading performance differs significantly between the "Los Angeles" county and the "San Diego" count. (Tip: first recode county into a numeric variable using Transform - Automatic Recode; then Analyze - means - independent groups t-test)

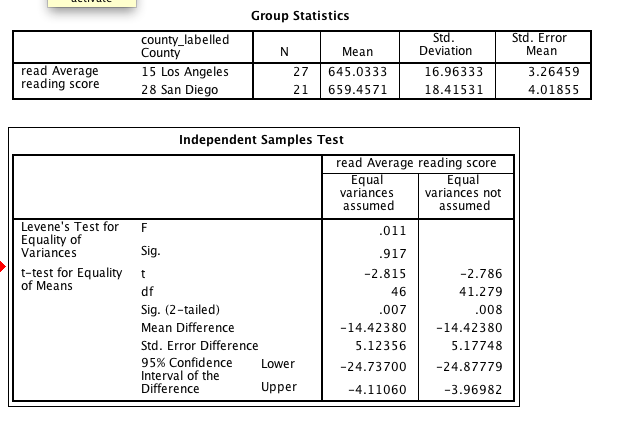
Some SPSS procedures only accept numeric variables with value labels instead of text variables. To automatically create a numeric variable with text labels that corresponds to a text variable, use transform - automatic recode.



Then take note of the numbers assigned to Los Angeles and San Diego.

Then analyze - compare means - independent samples t-test





## Define the null and alternative hypothesis

H0: San Diego mean equals Los Angeles mean

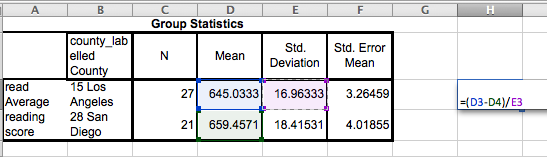
H1: San Diego mean does not equal Los Angeles mean

## Calculate cohen's d

There are few different formulas for calculating Cohen's d based on how the standard deviation is calculated. A more sophisticated option would be to take the standard deviation as the square root of MS error, but for now, let's just take the standard deviation in Los Angeles.

Thus, the equation is (645.0 - 6459.4) / 16.9

A nice option in these cases is to paste into Excel and do the calculation exactly there. This can also make it useful when you have many such calculations to perform



This gives us a d of -0.85. That is to say San Diego schools performed substantially better.

## Interpret the confidence intervals

We can be 95% confident that the true value is between -4.1 and -24.7, which corresponds to d's of around -.24 and -1.5.

## G-Power Homework Exercise

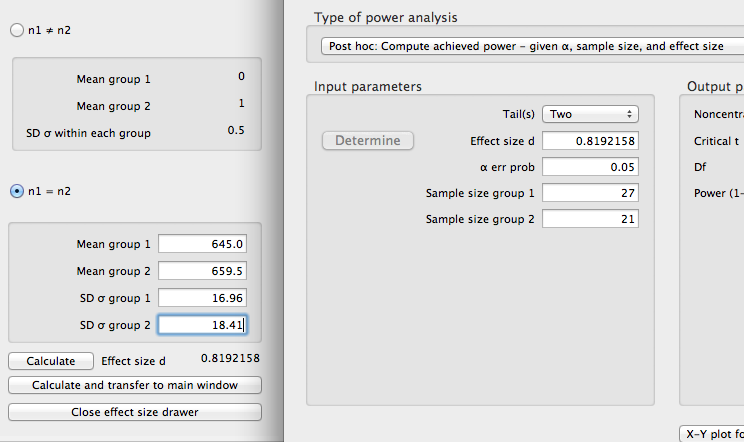
## Install G-Power 3

Just google G-Power 3 and follow the instructions. G Power 3 is available for both Mac and Windows.

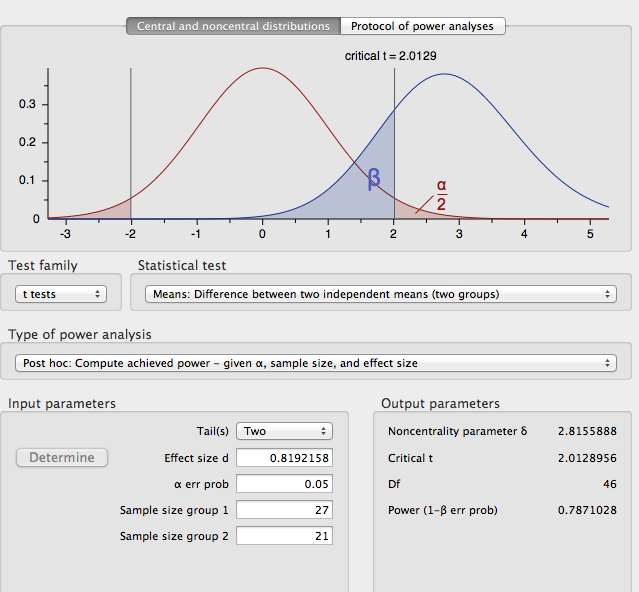
http://www.psycho.uni-duesseldorf.de/abteilungen/aap/gpower3/

## Work out the statistical power of the San Diego versus Los Angeles analysis assuming the sample group means are the population group means.

* In G-Power 3 go to Tests - Means - Two independent groups
* Choose post hoc
* Choose two-tail
* Enter sample sizes from data
* Click Determine and enter information to caclulate effect size



* Click calculate and transfer to main window
* Click calculate on main window



This shows that the post-hoc power was .787. I.e., reasonable power to detect such a large effect.

## In G-Power 3 work out what sample size you would need to examine differences between two groups when you are expecting a half standard deviation difference between the group means.

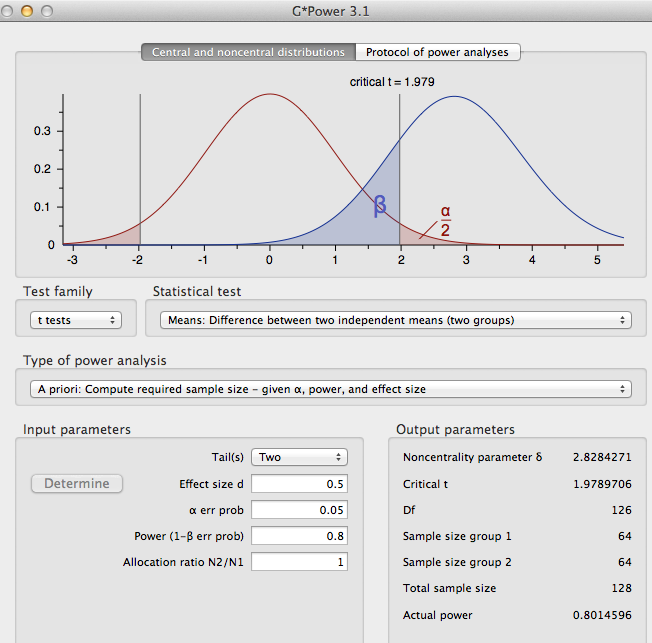
Change the power analysis "to a priori"

Furthermore, at this point you have to make a number of options, presumably 2-tail, d=0.5, alpha=.05.

The ratio of sample sizes could be 1 if you expect equal group sizes.

A common rule of thumb for adequate power is .80 (i.e., 80%).

But you might also want to see what sample size would be required for say 95% or 99% power.



This suggests that 128 participants are required for 80% power.