```
##### T TESTS #####

# A t-test compares two distributions to test the null hypothesis that the two means are not
# different. t-Tests are classic, related to ANOVAs, and still legit for simple comparisons. Here
# we explore a number of binary (two condition) potential drivers of birth weights in North
# Carolina.

# In 2004, the state of North Carolina released a large data set containing information on births
# recorded there. This data set is useful to researchers studying predictors of newborn weights,
# which itself predicts health. We will test the following three hypotheses:

# A. Smoking by pregnant women causes their newborn babies to weigh less at birth than babies
#     born to mothers who do not smoke. This would support efforts to reduce smoking.
# B. Babies born to married women weigh more on average than those born to unmarried
#     mothers (note that same-sex marriage was not legal in NC in 2004). This would support
#     need-based prenatal programs (where single- vs. dual income is assumed to affect access to
#     health care, etc.).
# C. Babies born to white mothers weigh more on average than those born to non-white mothers.
#     This would support prenatal programs in minority-dominated areas to improve
#     newborn health (because race remains a proxy for economic disadvantages, health care
#     access, etc.).

### Notice that we focus on quantitative responses to categorical predictors.

# Load the ncbirths data set, then attach it for convenience, and view it. Variables are:
#    fage = father's age
#    mage = mother's age
#    mature = category for mother's age
#    weeks = pregnancy interval
#    premie = category for weeks
#    visits = number of prenatal doctor visits
#    marital = legal marital status
#    gained = weight gained while pregnant (lbs)
#    weight = baby's weight at delivery (lbs)
#    lowbirthweight = category for weight
#    gender = baby's sex
#    habit = mother is smoker or not
#    whitemom = mother is white or not

    summary(ncbirths) # to inspect the data - any problems with the data columns we will use?

#  What to do with NAs? - include this line in commands below: na.rm=TRUE
# this essentially says “NA removal = true” and omits NAs from analyses

### Assumptions
# Do the data we use here (weights) fit assumptions of normality [and homogeneous variance]
```

```
# for each comparison we will make (habit, marital, whitemom)?
# If not, make data fit assumptions using transformations, as you already learned to do.

# Here's the scoop on t-tests. Unlike some statistical packages, the default assumes unequal
# variance (convenient!) and applies the Welsh df modification. The basic commands are:

# independent 2-group t-test
t.test(y~x) # where y is numeric and x is a binary factor

# independent 2-group t-test
t.test(y1,y2) # where y1 and y2 are numeric

# paired t-test
t.test(y1,y2,paired=TRUE) # where y1 & y2 are numeric

# one sample t-test
t.test(y,mu=0) # Ho: mu=0

# options include:
#    the var.equal = TRUE statement, inside (), to specify equal variances
#    the alternative="less" or alternative="greater" option to specify a one tailed test, as opposed
#    to the default alternative="two.sided".  Notice that the order of subtraction for that option is
#    alphabetical (e.g., nonsmoker – smoker).

# Now test the three hypotheses using t-tests, where YOU CHOOSE appropriately among the
options above.

# Try one more hypothesis that you make up.

# One last consideration: we have now tried (at least) 4 different t-tests on birth weight. With
# enough attempts, we might eventually stumble on a significant effect at random. Thus a
# Bonferroni correction: where we adjust the critical p-value to find signficance for the number
# of t-tests we conduct.

# So if we stick to just the three hypotheses (A-C), the Bonferonni correction would lead to a
# critical p-vale of 0.05 / 3 = 0.0167. Thus any one test would have to attain a p-value of 0.0167
# or less to be considered signficant, rather than the customary 0.05. Did this change any of your
# interpretations?
```