Two-Sample *t* Test in SAS® Studio

We are going to conduct a two-sample t test (a.k.a. independent *t* test) to compare means between two independent groups. The statistical software provides output for us to validate assumptions of the parametric test.

Dataset Description

We will be using the BIRTH data for illustration**.** Babies with low birth weights (defined to be less than 2500 grams) are a concern because of their potential medical problems. Health researchers used this data to identify possible contributing factors to low birth weight and recommend strategies to reduce the number of low-birth-weight babies. There are 189 cases with 11 variables including:

- ID: Identification Code
- LOW: Low Birth Weight (0: birth weight>=2500g, 1: birth weight<2500g)
- AGE: Age of the Mother in Years
- LWT: Weight in Pounds at the Last Menstrual Period
- ETH: Ethnicity Category (1, 2, 3)
- SMOKE: Smoking Status During Pregnancy (1: yes, 0: no)
- PTL: History of Premature Labor (0: none, 1: one, and so on)
- HT: History of Hypertension (1: yes, 0: no)
- UI: Presence of Uterine Irritability (1: yes, 0: no)
- FTV: Number of Physician Visits During the First Trimester (0: none, 1: one, and so on)
- BWT: Birth Weight in Grams

Two-sample t test

For this example, we will be investigating the question "Does a baby's birth weight (BWT) differ among smoking mothers (SMOKE=1) versus non-smoking mothers (SMOKE=0)?".

Therefore, the hypothesis statements are:

H_0: The mean BWT with non-smoking mom is equal to the mean BWT with smoking mom.

H_A: The mean BWT with non-smoking mom is not equal to the mean BWT with smoking mom.

In this question, we have one continuous variable, BWT, and one categorical variable that has only two groups (SMOKE). Therefore, we can use the *t*-test to identify if there is a difference between these two groups.

1. Bring the birth dataset to the software.
2. Expand Tasks and Utilities, expand Tasks, expand Statistics, and then select *t*-Tests.
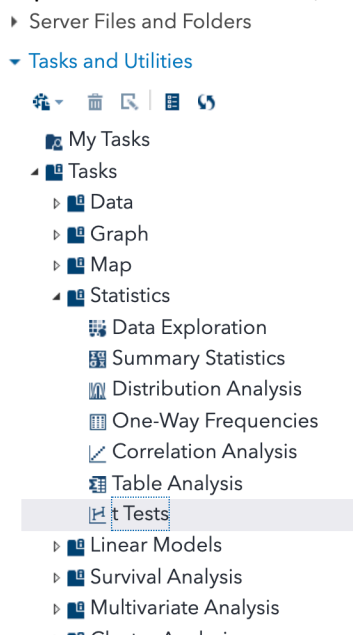
3. Select birth file under Data. Under *t* test of Roles, select Two-sample test. Select BWT as the Analysis variable, and Smoke as the Groups variable.
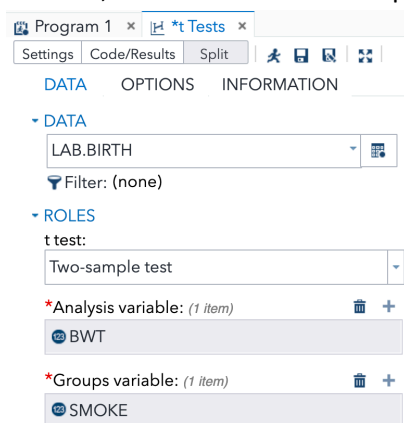
4. Go to Options and select the two-tailed test with the alternative hypothesis that the difference between the means is 0. Select test for normality to check the normality assumption. Under plots, select the Histogram and box plot, and Normality plot

5. To add your name in the footnote, go to the code and click Edit.
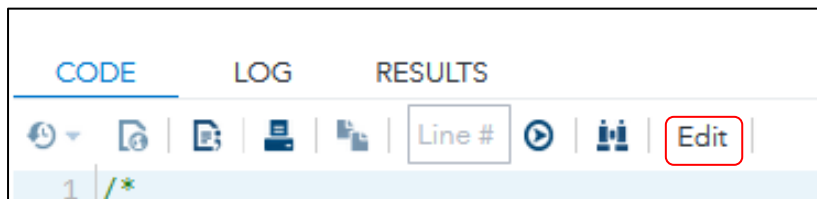
6. Insert the following code to the t test code portion between 'var BWT;' and 'run;'.

```
/* t test */
proc ttest data=LAB.BIRTH sides=2 h0=0 plots(only showh0)=(summaryPlot qqplot);
    class SMOKE;
    var BWT;
    Footnote "First and Last Name";
run;
```

7. Run the code. The first two sections are the results of test of Normality for the two Smoke groups. The output provides several tests for normality, all of which compare the scores in the sample to a normally distributed set of scores with the same mean and standard deviation using the null hypothesis that "the distribution is normal". The Shapiro-Wilk test is the most common test for normality and the p-value of this test is 0.41. Therefore, we fail to reject the null hypothesis and determine that it is appropriate to treat this distribution as if it were normally distributed at significance level of 0.05.

Variable: BWT (Birth Weight, Grams)
SMOKE = 0

Tests for Normality				
Test		Statistic	p Value	
Shapiro-Wilk	W	0.987102	Pr < W	0.3436
Kolmogorov-Smirnov	D	0.060256	Pr > D	>0.1500
Cramer-von Mises	W-Sq	0.07215	Pr > W-Sq	>0.2500
Anderson-Darling	A-Sq	0.495093	Pr > A-Sq	0.2188

Variable: BWT (Birth Weight, Grams)
SMOKE = 1

| Tests for Normality				
Test		Statistic	p Value	
Shapiro-Wilk	W	0.982771	Pr < W	0.4103
Kolmogorov-Smirnov	D	0.068722	Pr > D	>0.1500
Cramer-von Mises	W-Sq	0.04179	Pr > W-Sq	>0.2500
Anderson-Darling	A-Sq	0.309712	Pr > A-Sq	>0.2500

The Shapiro-Wilk test is based on the correlation between the data and the corresponding normal scores and provides better power when compared to the Kolmogrov-Smirnov test. Power is the most frequent measure of the value of a test for normality because it describes the ability to detect whether a sample comes from a non-normal distribution.

The Kolmogrov-Smirnov test is an empirical distribution function in which the theoretical cumulative distribution function of the test distribution is contrasted with the empirical distribution function of the data. This test is highly sensitive to extreme values and tends to have lower power. This test is not recommended when parameters are estimated from the data, regardless of sample size.

The Cramér-von Mises's test is an empirical distribution function omnibus test for the composite hypothesis of normality. It uses the summed squared differences between observed and expected cumulative proportions as the test statistic.

The Anderson–Darling test is a statistical test of whether a given sample of data is drawn from a given probability distribution. When applied to testing whether a normal distribution adequately describes a set of data, it is one of powerful statistical tools for detecting most departures from normality.

8. We can simply check the Q-Q plot of the data for normality assumption. The Q-Q plots don't show large deviation from normal distribution for BWT in either SMOKE group.

First and Last Name
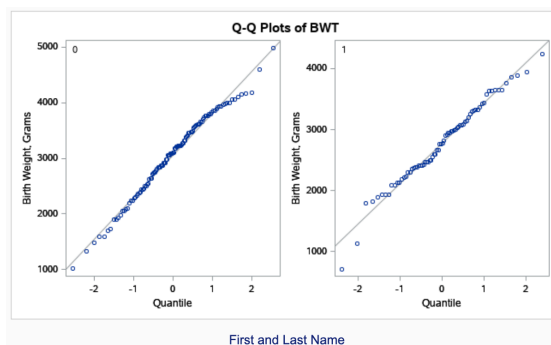
9. The results from the two-samples *t*-test are shown below.

Variable: BWT (Birth Weight, Grams)

SMOKE	Method	N	Mean	Std Dev	Std Err	Minimum	Maximum
0		115	3055.0	752.4	70.1625	1021.0	4990.0
1		74	2773.2	660.1	76.7322	709.0	4238.0
Diff (1-2)	Pooled		281.7	717.8	107.0		
Diff (1-2)	Satterthwaite		281.7		104.0		

SMOKE	Method	Mean	95% CL Mean		Std Dev	95% CL Std Dev	
0		3055.0	2916.0	3193.9	752.4	666.1	864.6
1		2773.2	2620.3	2926.2	660.1	568.2	787.7
Diff (1-2)	Pooled	281.7	70.6927	492.7	717.8	651.8	798.7
Diff (1-2)	Satterthwaite	281.7	76.4668	487.0			

Method	Variances	DF	t Value	Pr > |t|
Pooled	Equal	187	2.63	0.0092
Satterthwaite	Unequal	170	2.71	0.0074

Equality of Variances

Method	Num DF	Den DF	F Value	Pr > F
Folded F	114	73	1.30	0.2290

Under the null that the mean group differences is zero, we have an observed 281.7 differences between newborns with non-smoking moms (SMOKE=0) vs with smoking counterparts (SMOKE=1). How likely to get this 281.7 differences in our sample while the true difference is 0? Not likely by referring to the p-value which is .0074 for Satterthwaite method.

As for the t-statistics, we have two p-values associated with two t statistics: t Value=2.63 under Pooled method and t Value=2.71 under Satterthwaite method. Which one to choose? It does not matter in our case as both show significance that that the observed group difference (281.7) cannot be just by chance. However, pooled t test is only valid if we confirm equal variances. In case of lack of testing homogeneity of the variances, we can always use Satterthwaite results. We can also refer to the box plots to visually check equality of the variances of the two groups. It doesn't show too much difference.
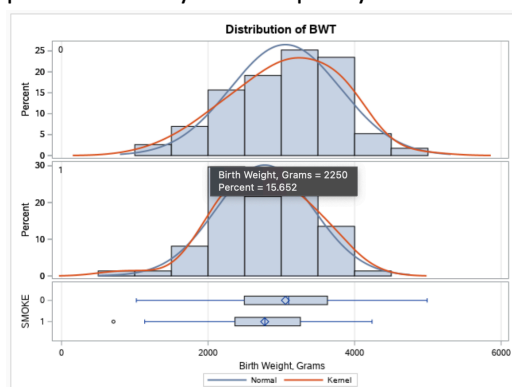
Why are the two methods different? They use different formulas to calculate the statistics. We can see that the pooled method and Satterthwaite method show different standard error and degree of

freedom (DF): DF=107 and SE=107 with Pooled method (under the assumption of equal group variances) and DF=104 and SE=104 with Satterthwaite method.

Overall, we will conclude at significance level of 0.05 that baby's birth weight differs among smoking mothers versus non-smokingmothers with p-value of 0.0074.