I receive data that contain information about infant birth weight, including details about the mother's gender, marital status, age, baby's gender, smoking habits, weight gain during pregnancy, number of prenatal visits, and education level. I plan to investigate if there is a correlation between the mother's condition and the infant's weight at birth. I plan to conduct some statistical analysis to investigate a potential correlation. I hypothesize that there is a link between the mother's health and the infant's birth weight, and the gender of the baby does not impact the baby's weight. To support this hypothesis, I will focus on the mother's smoking habits and the gender of the baby.

I have conducted a statistical analysis to investigate a potential correlation between infant gender and weight. I assert that there is no correlation between these two factors.

Null Hypothesis (Ho): there is no significant relationship(correlation) between the weight of the infant's birth and the baby's gender.
Alternative hypothesis (Ha): there is a significant relationship between the weight of the infant's birth and the gender of the child.
From my statistical analysis, with a P-value of 0.0001<alpha indicating strong evidence against the null hypothesis, it appears that there is a correlation between the gender of an infant and their weight.

Another statistical analysis examines the relationship between a mother's smoking and the gender of her infant.

Null Hypothesis (Ho): there is a significant relationship(correlation) between the weight of the infant's birth and the mom's smoking.
Alternative hypothesis (Ha): there is no significant relationship between the weight of the infant's birth and the gender of the child.
After conducting a statistical analysis, I can confirm that the claim is invalid. The P-value was 0.0004, which is less than the alpha of 0.05, leading me to reject the null hypothesis. Therefore, I found no relationship between the mother's smoking habit and the infant's birth weight.

Independent T-test for samples Weight by infant gender

I conducted a statistical analysis to explore the potential correlation between a baby's gender and their birth weight. To assist me in my analysis, I employed an independent t-test. Was there a relationship between these two variables? There were only two possible outcomes: either there was a connection or there was not. Initially, I hypothesized that there was no correlation, but I remained open to the possibility that there was one.

To ensure the authenticity of my test, I verified that the data I used followed a normal distribution, which accurately represented the population. Additionally, I confirmed that the groups I compared were independent to prevent any overlap.

When I examined the test results, I focused on the p-value. The p-value indicates the likelihood that the observed results were due to chance. In this instance, the p-value was incredibly low, indicating that there is indeed a correlation between a baby's gender and their birth weight. I support my examination with the statistical analysis down below.

Null Hypothesis (Ho): there is no significant relationship(correlation) between the weight of the infant's birth and the baby's gender.
Alternative hypothesis (Ha): there is a significant relationship between the weight of the infant's birth and the gender of the child.
My samples are independent. I see they are normally distributed based on a Q-Q plot, and the samples are more than 30.

The TTEST Procedure

Variable: Weight (Weight)

Boy	Method	N	Mean	Std Dev	Std Err	Minimum	Maximum
0		24208	3310.6	547.7	3.5204	240.0	6350.0
1		25792	3427.3	577.7	3.5970	284.0	5970.0
Diff (1-2)	Pooled		-116.7	563.4	5.0416		
Diff (1-2)	Satterthwaite		-116.7		5.0331		

Boy	Method	Mean	95% CL Mean		Std Dev	95% CL Std Dev	
0		3310.6	3303.7	3317.5	547.7	542.9	552.7
1		3427.3	3420.2	3434.3	577.7	572.7	582.7
Diff (1-2)	Pooled	-116.7	-126.6	-106.8	563.4	559.9	566.9
Diff (1-2)	Satterthwaite	-116.7	-126.6	-106.8			

Method	Variances	DF	t Value	Pr > |t|
Pooled	Equal	49998	-23.15	<.0001
Satterthwaite	Unequal	49993	-23.18	<.0001

95% of Confidence level

We can confidently state with 95% certainty that there is a significant correlation between an infant's gender and their birth weight.

By default, Sas performs a two-tail T-Test

Assuming Equal variances

Assuming Unequal variances

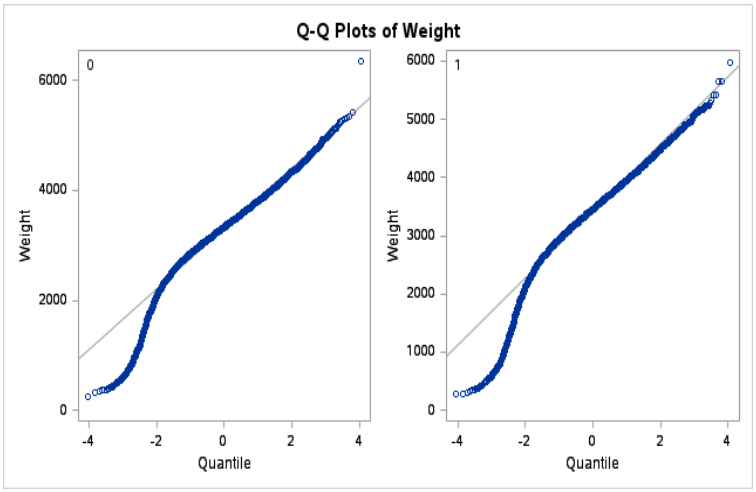
Equality of Variances				
Method	Num DF	Den DF	F Value	Pr > F
Folded F	25791	24207	1.11	<.0001

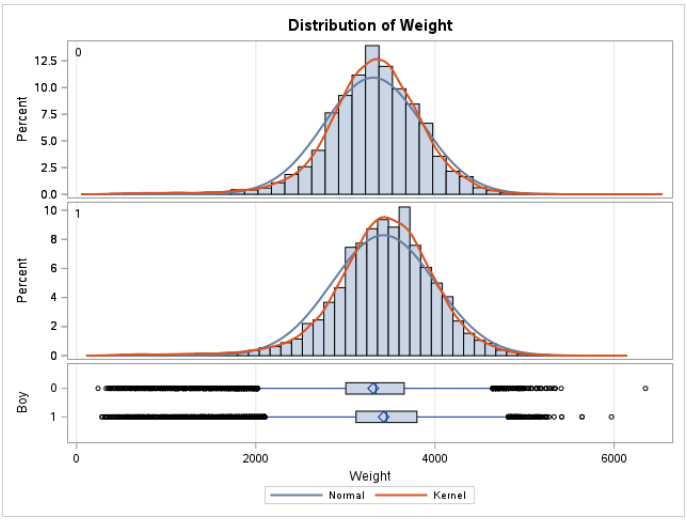
P-value<0.05, Based on P-value number, we reject the Null hypothesis.

The output above prints the t-statistic (T value=-23.18 for unequal variances, T value= -23.15 for equal variances) and the degrees of freedom 49993 for unequal variances and 49998 for equal variances (n - 1). The p-value is 0.0001, less than 0.05, so we reject the null hypothesis in favour of the Alternative hypothesis.

Interpretation of the output

Since the p-value is 0.0001, it is less than Alpha (0.0001<0.05). We reject the Null hypothesis at the 5% significance level, which says there is a significant relationship between the weight of the infant's birth and the gender of the child.
Based on statistics, we say there is a significant relationship between the weight of the infant's birth and the gender of the child, with a confidence level of 95%.





Independent T-test for samples Weight by Mom Smoke habit

I performed a T-test to examine the relationship between the weight of infants' birth and the variable Mom-smoke to see if there is any relationship between the two independent variables.

Null Hypothesis (Ho): a significant relationship(correlation) exists between the infant's birth weight and the mom-smoke habit.

Alternative hypothesis (Ha): no significant relationship exists between the infant's birth weight and the mom-Smoke habit.

My samples are independent. I see they are normally distributed based on a Q-Q plot, and the samples are more than 30.

The TTEST Procedure

Variable: Weight (Weight)

MomSmoke	Method	N	Mean	Std Dev	Std Err	Minimum	Maximum
0		43467	3402.3	558.0	2.6766	240.0	6350.0
1		6533	3160.9	576.8	7.1358	312.0	5245.0
Diff (1-2)	Pooled		241.5	560.5	7.4376		
Diff (1-2)	Satterthwaite		241.5		7.6213		

95% of Confidence level

We can confidently state with 95% certainty that there is no significant correlation between baby birth weight and mom smoke habit.

MomSmoke	Method	Mean	95% CL Mean		Std Dev	95% CL Std Dev	
0		3402.3	3397.1	3407.6	558.0	554.3	561.8
1		3160.9	3146.9	3174.8	576.8	567.0	586.8
Diff (1-2)	Pooled	241.5	226.9	256.0	560.5	557.1	564.0
Diff (1-2)	Satterthwaite	241.5	226.5	256.4			

Method	Variances	DF	t Value	Pr > |t|
Pooled	Equal	49998	32.46	<.0001
Satterthwaite	Unequal	8474.1	31.68	<.0001

Assuming Equal variances

By default, Sas performs a two-tail T-Test

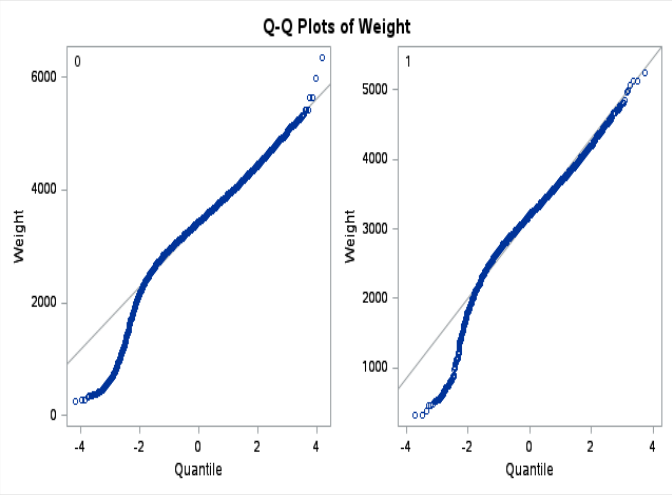
Assuming Unequal variances

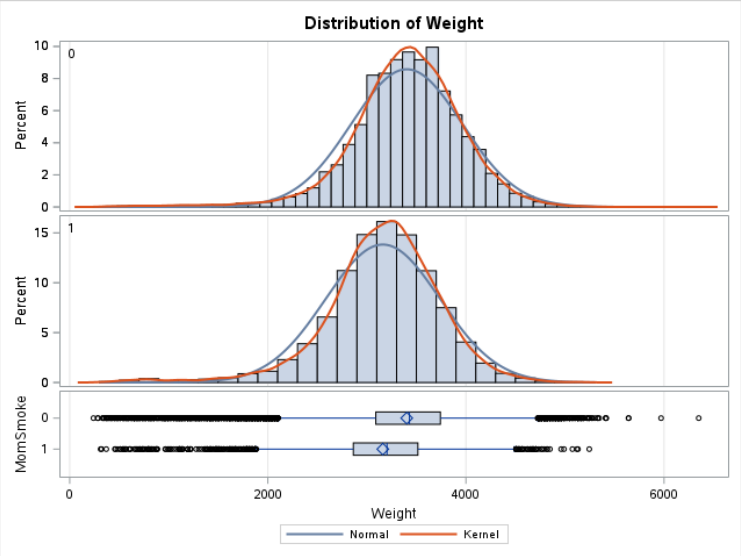
Equality of Variances

Method	Num DF	Den DF	F Value	Pr > F
Folded F	6532	43466	1.07	0.0004

P-value<0.05, Based on P-value number, we reject the Null hypothesis.

The output above prints the t-statistic (T value=31.68 for unequal variances, T value= 32.46 for equal variances) and the degrees of freedom 49998 for equal variances and 8474.1 for unequal variances (n - 1). The P-value is 0.0004, less than 0.05, so we reject the null hypothesis in favour of the Alternative hypothesis.





Interpretation of the output

Since the p-value is 0.0004, it is less than Alpha (0.0004<0.05). We reject the Null hypothesis at the 5% significance level, which says a significant relationship exists between the weight of the infant's birth and the mom-smoke habit.
Based on statistics, we say there is no significant relationship between the weight of the infant's birth and the mom-smoke habit, with a confidence level of 95%.