Assignment 1

STATS 252
LAB S5

Jingyu Xiang
1537572

1.

This study is an observational study because there is no imposing of treatments and no manipulation or control of any variables or conditions. Since it's an observational study and random sampling should have not been implemented because all mammals in this study were from one zoo, so population inference can not be made for the whole mammality. Also, it is impossible to establish a causal link between brain size and litter size using this data because causal inference can only be applied to experimental study.

2.(a)

Descriptives:

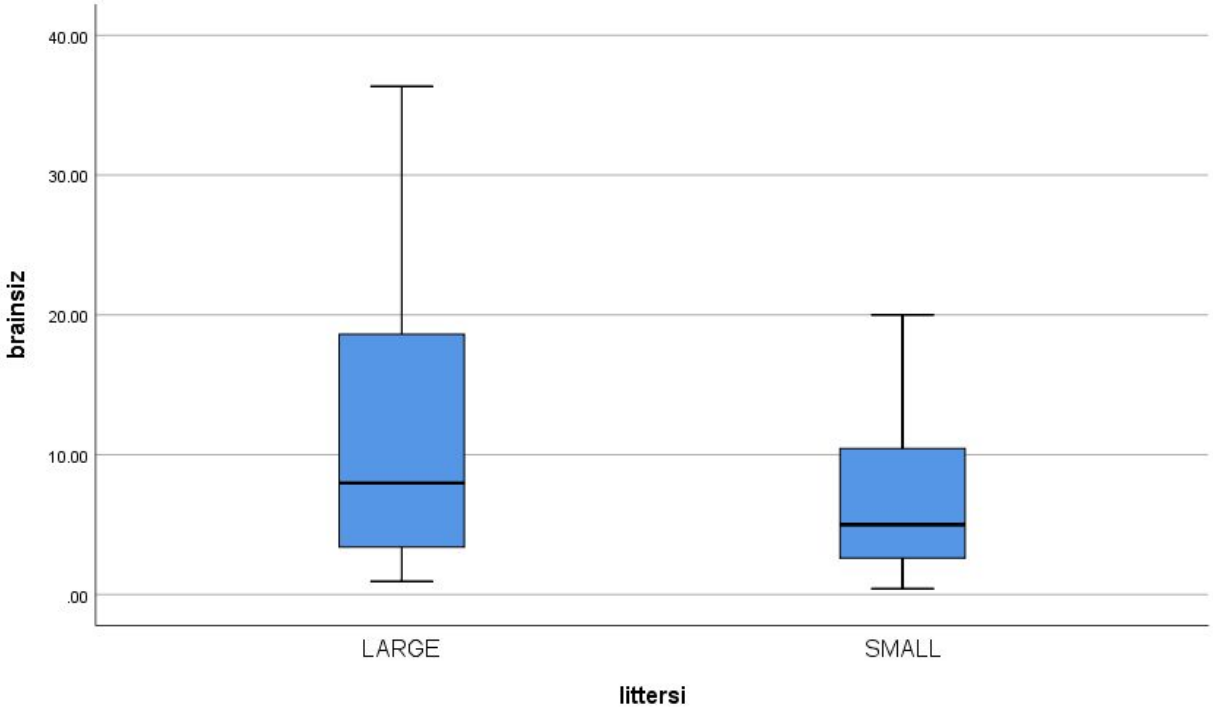
	littersi		Statistic	Std. Error
brainsiz	LARGE	Mean	10.9684	1.46640
		95% Confidence Interval for Mean	Lower Bound	8.0131
			Upper Bound	13.9238
		5% Trimmed Mean	10.1645	
		Median	7.9700	
		Variance	96.765	
		Std. Deviation	9.83692	
		Minimum	.94	
		Maximum	36.35	
		Range	35.41	
		Interquartile Range	15.37	
		Skewness	1.092	.354
		Kurtosis	.228	.695
	SMALL	Mean	6.8859	0.76459
		95% Confidence Interval for Mean	Lower Bound	5.3501
			Upper Bound	8.4216
		5% Trimmed Mean	6.5247	
		Median	5.0000	
		Variance	29.815	
		Std. Deviation	5.46030	
		Minimum	.42	
		Maximum	20.00	
		Range	19.58	
		Interquartile Range	8.00	
		Skewness	.816	.333
		Kurtosis	-.355	.656

Both the mean and the standard deviation of the brain size for the large litter-size group(10.9684 and 9.83692) are greater than the mean and standard deviation of the small litter-size group(6.8859 and 5.46030). The standard error of mean for the large litter-size group

is 1.46640, and the standard error for the small litter-size group is 0.76459. The standard error of the mean is a measure of sample means around the population mean.

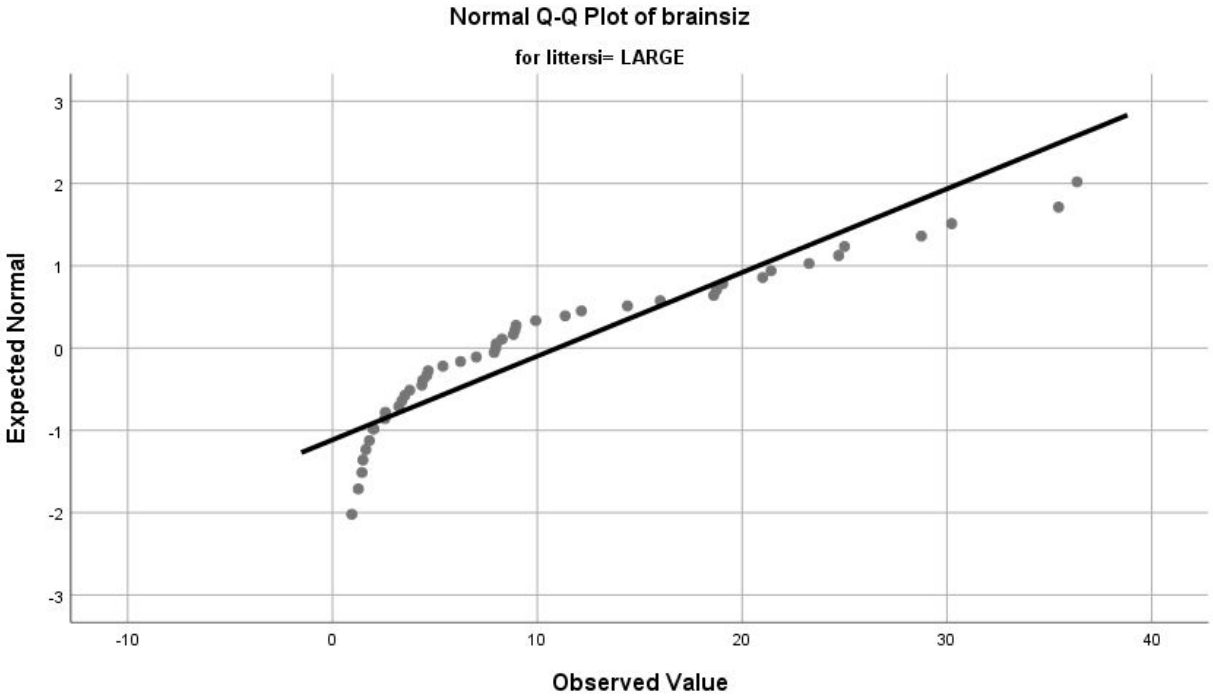
(b)For the large litter-size group, the 95% confidence interval for the mean relative brain weight is (8.0131, 13.9238). For the small litter-size group, the 95% confidence interval for the mean relative brain weight is(5.3501, 8.4216). The confidence interval for the large litter-size group is wider because the the standard error for the larger litter-size group is greater, and confidence interval = (estimate - critical value*standard error, estimate + critical value*standard error). And by observing, the two confidence intervals overlap.

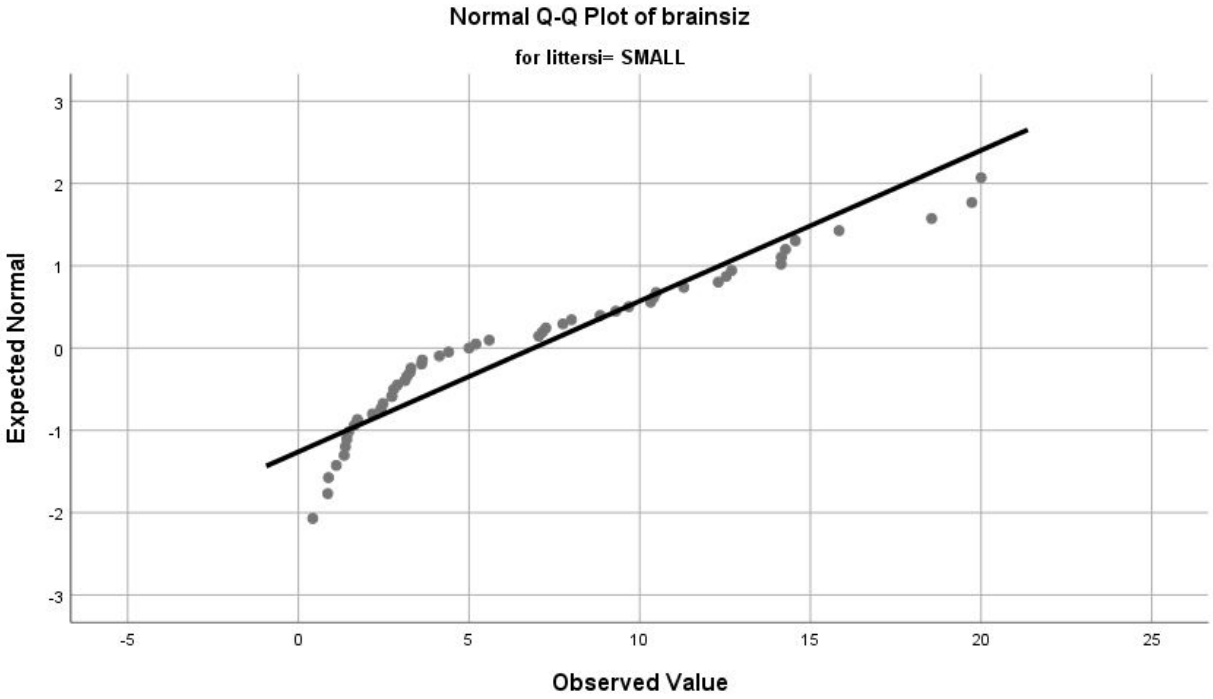
(c)



Both of the two distributions show right skewness. Since both of the two distributions are right-skewed, median is a better measure of central tendency, because it is a resistant measure which is more robust to extreme values and skewness. There are no outliers.

(d)





There is evidence that the assumption of normality may be seriously violated in both cases. Because in both plots, not all data points fall approximately in a straight line(not roughly linear).

(e)Data are not collected by simple random samples; Normality can not be assumed according to the result of (d); Samples are independent . Therefore, two assumptions are violated.

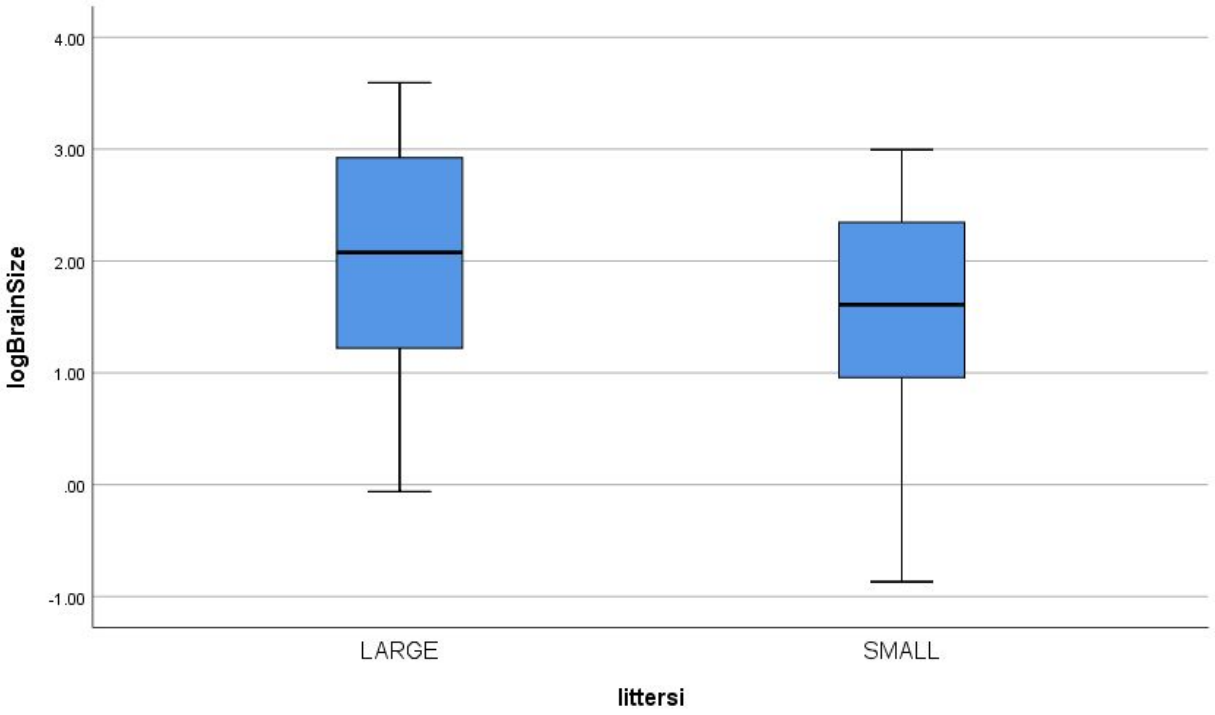
3.(a)

Descriptives

	littersi		Statistic	Std. Error
logBrainSize	LARGE	Mean	1.9494	.15150
		95% Confidence Interval for Mean	Lower Bound	1.6441
			Upper Bound	2.2548
		5% Trimmed Mean	1.9617	
		Median	2.0757	
		Variance	1.033	
		Std. Deviation	1.01629	
		Minimum	-.06	
		Maximum	3.59	
		Range	3.66	
		Interquartile Range	1.73	
		Skewness	-.162	.354
		Kurtosis	-1.022	.695
	SMALL	Mean	1.5525	.13334
		95% Confidence Interval for Mean	Lower Bound	1.2846
			Upper Bound	1.8203
		5% Trimmed Mean	1.5834	
		Median	1.6094	
		Variance	.907	
		Std. Deviation	.95223	
		Minimum	-.87	
		Maximum	3.00	
		Range	3.86	
		Interquartile Range	1.44	
		Skewness	-.390	.333
		Kurtosis	-.645	.656

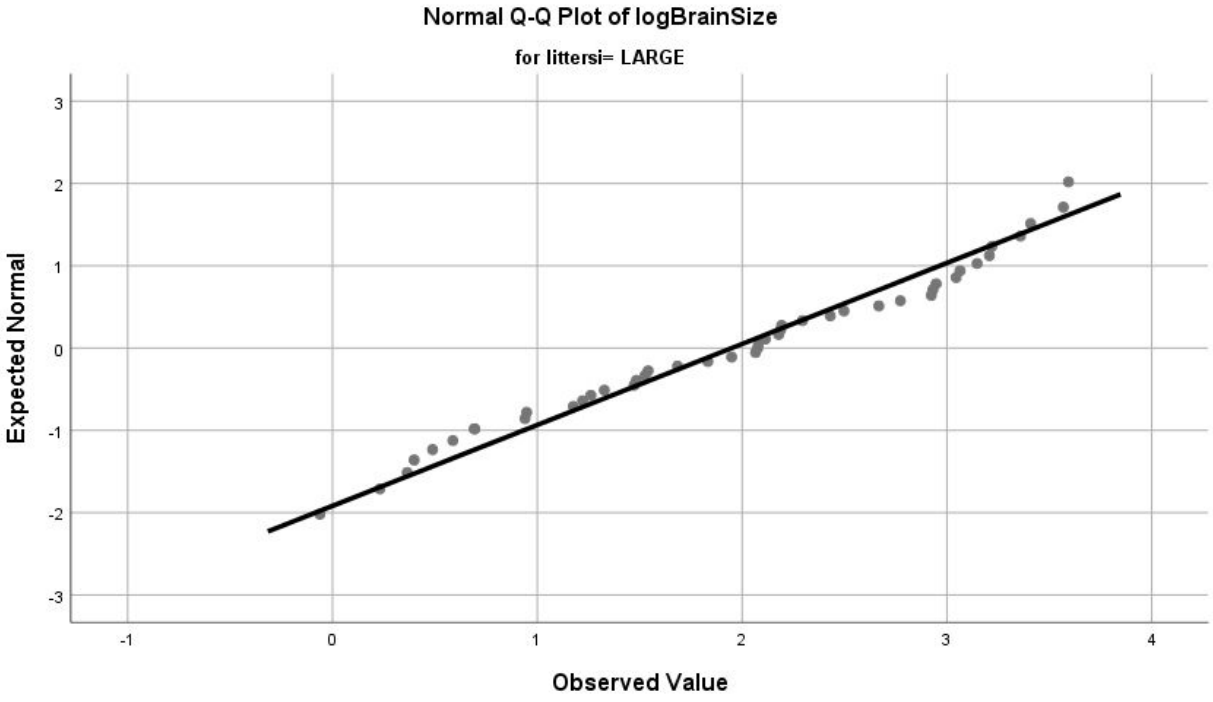
After log transformation, the the mean brain size for the large-littersized group(1.9494) is larger than the mean brain size of the small litter-sized group(1.5525). The standard deviation of the large-littersized group(1.01629) is larger than the value of small-littersized group(0.95223). And for the standard error of mean, the large litter-sized group again has a larger value of that.

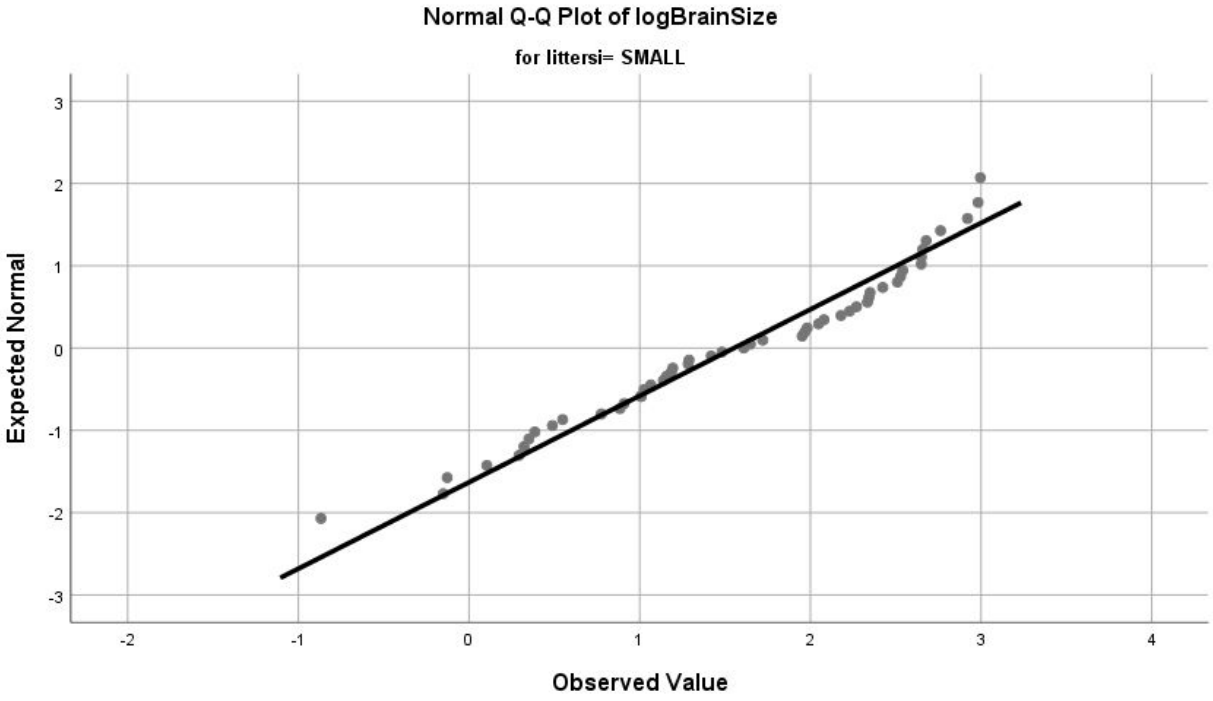
(b)



After log transformation. The large litter-sized group has a larger mean of brian size and the small litter-sized group has a smaller mean of brian size. The small litter-sized group's distribution shows left skewness. And for the large litter-sized group, it is also a little bit left skewed.

(c)





There is no evidence that the assumption of normality may be seriously violated in either case. Because in both plots, all data points fall approximately in a straight line(roughly linear).

(d)After log-transformation, the data distribution become more normal.

4.(a)The difference between the average of log-relative brain weight for large litter size group and small litter size group is -0.3969. The antilog of this value is 1.4872. Which means the ratio of median of the large litter-sized group and small litter-sized group is 1.4872.

(b)H0: meanLARGE is equal to meanSMALL
Ha: meanLAEGE is not equal to meanSMALL
the distribution of the test statistic under the null hypothesis is t distribution
Test statistic t = 1.975
P-val = 0.051
Conclusion: p-val is greater than 0.05, so reject H0. Therefore, there is sufficient evidence that meanLarge is not equal to meanSmall.

(c)
Mean difference(LARGE-SMALL) = 4.08256
SE(mean diff)=1.59939
95% confidence interval for original scale: (0.90694, 7.25819)
We are 95% percent confident that the mean difference of the brain size of the LARGE group and the SMALL group is between 0.90694 and 7.25819. This calculation is consistent with the conclusion of part(b) because this original confidence interval does not contain 0.

5.
SE(mean diff) = sp*root(1/n1+1/n2)
sp=root{[(n1-1)s1^2+(n2-1)s2^2)]/(n1+n2-2)}
n1=45 n2=51 s1=1.01629 s2=0.95223
SE(mean diff)=0.20099
t= (mean1-mean2)/SE(mean diff)
mean1=1.9494 mean2=1.5522
t=(1.9494-1.5522)/0.20099=1.975
This result is consistent with SPSS outcome.

6.(a)

Hypothesis Test Summary

	Null Hypothesis	Test	Sig.	Decision
1	The distribution of brainsiz is the same across categories of littersi.	Independent-Samples Mann-Whitney U Test	.065	Retain the null hypothesis.

Asymptotic significances are displayed. The significance level is .05.

(b)p-val = 0.065>0.05
So H0 is rejected. There is sufficient evidence to say that the distribution of brain size is not the same across categories of litter size(SMALL and LARGE).

7.Diet, body weight and gestation lengths.