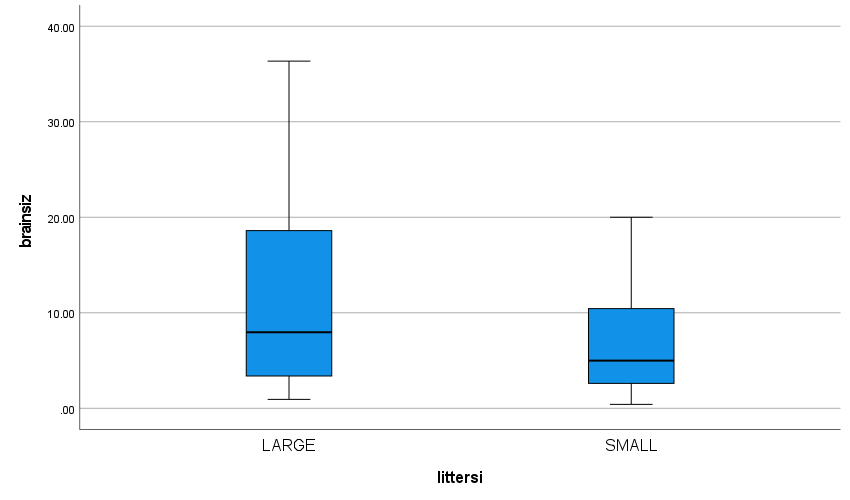
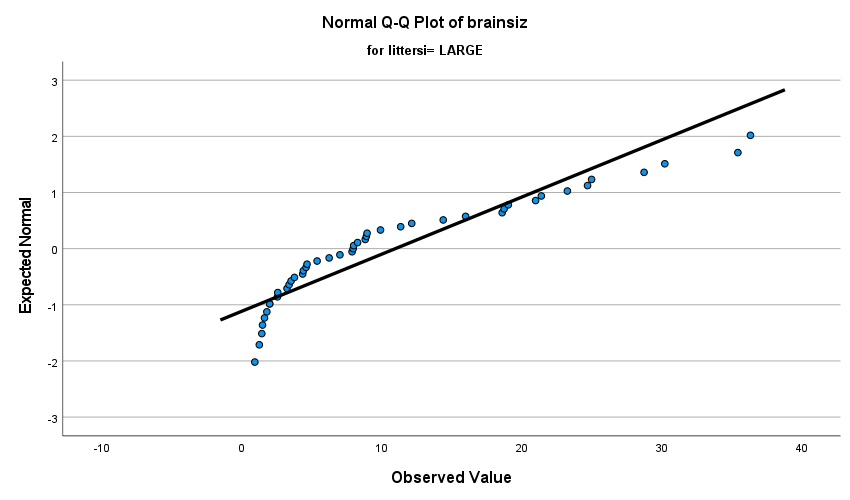
1. This type of study is an observational study. The reason for this is because we are not assigning any form of treatments to different groups nor are we actually controlling any type of variables within this study. We are also just collecting data randomly from mammals as it occurs in nature rather than us controlling anything. However, as it is an observational study, we can not make a causal inference (establish a cause-and-effect relation between brain sizes and litter sizes). On the other hand, we can make population inferences if the data from the mammals were selected randomly. This means that as long as we have randomly selected data, we can generalize the data and findings to the population of mammals.

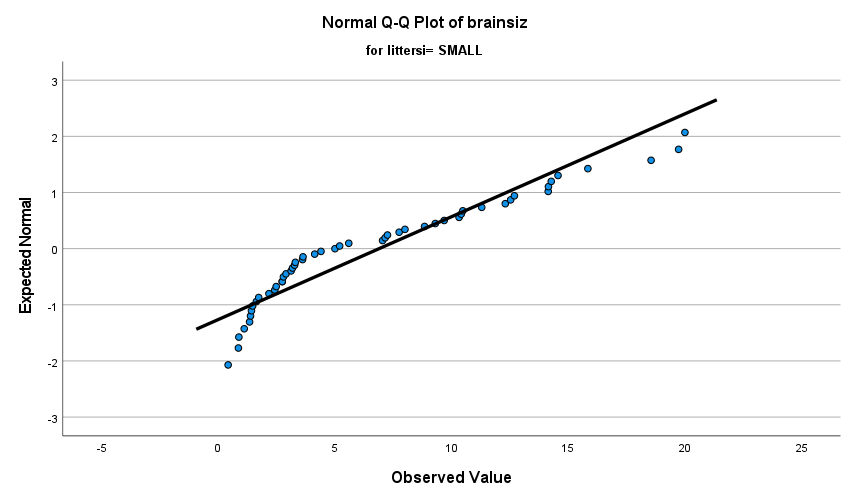
|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Descriptives** | | | | | |
|  | littersi | | | Statistic | Std. Error |
| brainsiz | LARGE | Mean | | 10.9684 | 1.46640 |
| 95% Confidence Interval for Mean | Lower Bound | 8.0131 |  |
| Upper Bound | 13.9238 |  |
| 5% Trimmed Mean | | 10.1645 |  |
| 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 | .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 |

* 1. As seen by the descriptive statistics above, for large litter sizes, we have a mean of 10.9684 with a standard deviation of 9.83692 alongside a standard error of 1.46640. For the small litter sizes, we have a mean of 6.8859 with a standard deviation of 5.46030 and a standard error of 0.76459. That means that the larger litter sizes have a larger mean brain size and has a much wider range of brain sizes from the mean compared to that of the smaller litter sizes which have a smaller mean yet the values are much closer to the mean value than the large brain sizes. However, as the ratio of the standard deviations of the larger divided smaller value is less than 2(the exact value is 1.801534714), we can say the 2 deviations are quite close to each other. The standard error here tells us the amount of error(accuracy) we should expect when we are estimated the population mean based off the sample mean.



As seen by the boxplot above, we can see the large litter sizes have a larger median brain size and a larger interquartile range than that of the small litter sizes. Both large and small litter sizes are skewed to the right as Q3-Q2>Q2-Q and due to this, the best measure of central tendency is the median. As the mean is best used in cases where there is generally symmetric data, when there is skewed data, the best would be to use the median. There also appear to be no outliers within the data.



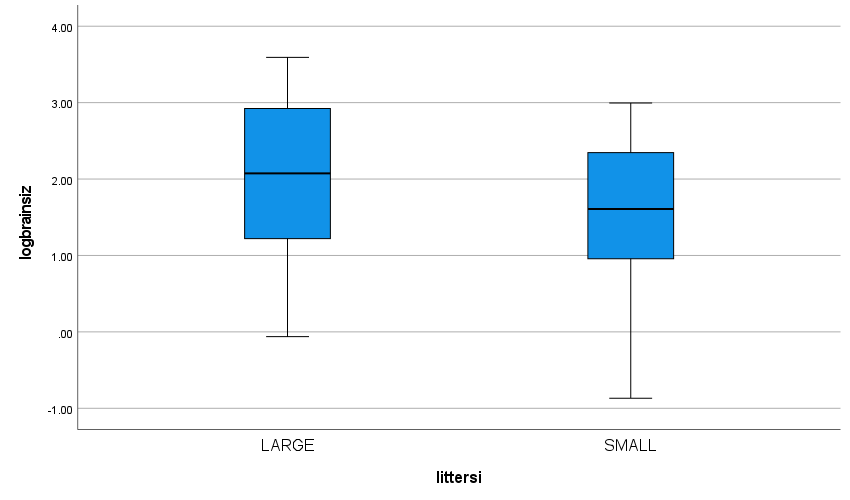


As seen with the two normality plots both tend to have a few points further away from the line and points trailing off at both ends. This indicates to us that normality is being violated within these two data plots. We can also infer from these plots that there is a right skewness.

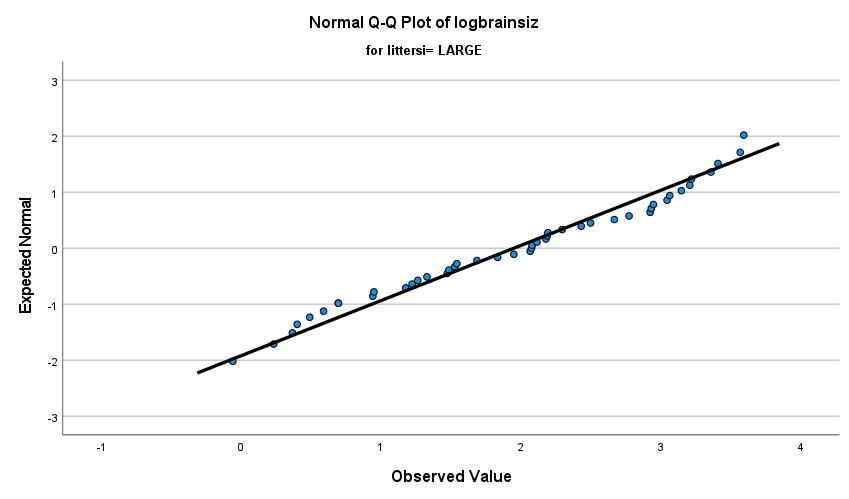
* 1. The assumptions needed to use t tools are: that random sampling must have been done, the data distribution is normal and the samples are independent from each other. However, as the data is not normally distributed as seen above due to there being a right skewness for both large and small litter sizes.

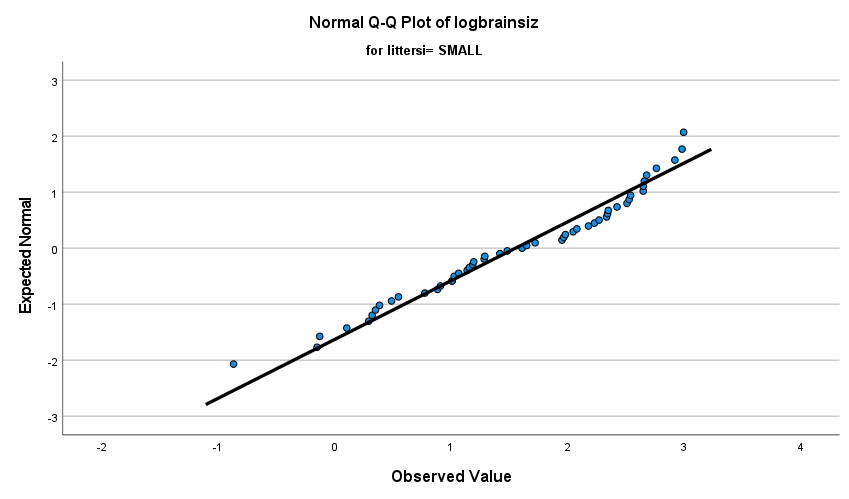
|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Descriptives** | | | | | |
|  | littersi | | | Statistic | Std. Error |
| logbrainsiz | LARGE | Mean | | 1.9494 | .15150 |
| 95% Confidence Interval for Mean | Lower Bound | 1.6441 |  |
| Upper Bound | 2.2548 |  |
| 5% Trimmed Mean | | 1.9617 |  |
| 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 |

The larger litter size has the larger mean of 1.9494 and standard deviation of 1.01629 in comparison to the smaller litter size which has a mean of 1.5525 and standard deviation of 0.95223. However, the two standard deviations are quite close to each other due to the ratio of larger over the smaller one is 1.067914264. The larger litter size has a standard error of 0.15150 while the smaller litter size has a standard error of 0.13334



As seen with the box plots above, the distribution appears to be symmetric for both box plots. This means that Q3-Q2=Q2-Q1. However, the larger litter size has the larger interquartile range in comparison to the smaller litter sizes. The larger litter size also has a larger median than the smaller litter size.





The normality plots show us that the points tend to lie close to the line with seldom few points lying far away from the line. This indicates to us the data is normally distributed and is symmetric (as confirmed above with the box plots.

* 1. The effect of the log transformation on the data has allowed our data to become more symmetric and normally distributed. It also brought our standard deviations for both small and large litter sizes closer to each other. The interquartile ranges also appear to have been brought closer together than they were before the log transformation.
  2. 1.9494-1.5525=0.3969. e0.3969=1.487207202. This tells us the median for the large litter brain sizes is about 1.487 times larger than the median brain sizes of small litter sizes
  3. We Shall perform a two-sample mean test

H0: µLarge= µSmall Ha: µLarge≠ µSmall

We shall be using t distributions with a df=94

t= -1.975

P-value = 0.026

α=0.01

As our P-value> α value, we fail to reject our null hypothesis.

In the 1% significance level, we do not have enough evidence from our data to conclude there is a difference between large litter brain sizes and small litter brain sizes.

* 1. 95% confidence interval on log scale = (-0.79605,0.00211)

95% confidence interval on original scale = (0.45111, 1.00211)

As the number 1 is within our confidence interval, we fail to reject the null hypothesis.

In the 5% significance level, we do not have enough evidence from our data to conclude there is a difference between large litter brain sizes and small litter brain sizes This conclusion is consistent with our answer for part b).

* 1. H0: µLarge= µSmall Ha: µLarge≠ µSmall

P: Value = 0.065

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

* 1. Therefore we will fail to reject the null hypothesis as P value is greater than our alpha level of 0.05. That means in the 5% significance level, our data does not present enough evidence to conclude there is a difference in brain sizes between large and small litter sizes. This is also the same conclusion we have received for part 4.

1. Other potential factor that could have been noted down is life expectancy. Another potential factor could include if they are a herbivore or an carnivore.