**Statistical Report Writing Sample No.1.**

**Introduction.** A federal regulatory agency is investigating an advertised claim that a certain device can increase the gasoline mileage of car. Ten of these devices are purchased and installed in cars. Gasoline mileage (mpg) for each of the cars is recorded both before and after installation. The data are listed in the following table.

 Table: Gasoline mileage (mpg)

|  |  |  |  |
| --- | --- | --- | --- |
| Car | Before | After | Change |
| 1 | 19.1 | 25.8 | 6.7 |
| 2 | 29.9 | 23.7 | -6.2 |
| 3 | 17.6 | 28.7 | 11.1 |
| 4 | 20.2 | 25.4 | 5.2 |
| 5 | 23.5 | 32.8 | 9.3 |
| 6 | 26.8 | 19.2 | -7.6 |
| 7 | 21.7 | 29.6 | 7.9 |
| 8 | 25.7 | 22.3 | -3.4 |
| 9 | 19.5 | 25.7 | 6.2 |
| 10 | 28.2 | 20.1 | -8.1 |

The columns “Before” and “After” show the mileage (mpg) before and after the installation of device,

the column “Change” represents the improvement of mileage.

 Table: Summary statistics for gasoline mileage (mpg)

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Variable | Mean | SD | Q1 | Median | Q3 |
| Before | 23.22 | 4.25 | 19.68 | 22.60 | 26.53 |
| After | 25.33 | 4.25 | 22.65 | 25.55 | 27.98 |
| Change | 2.11 | 7.54 | -5.50 | 5.70 | 7.60 |

In this study we want to determine whether there is a significant gain in mileage after the devices were installed.

**Data analysis and statistical test.**  The following figures visualize data for the change in gasoline mileage. Note that the sample size is small. Thus, the reliability of statistical test is dependent on the ability to assume that the data come from a normal distribution. In particular, we need to assess the normality for the change of mileage since it is used for a statistical test. The histogram displays neither a center nor symmetry of distribution. Rather, it is bimodal, suggesting that there are two distinct populations to which the effect of device could be different.



The QQ plot also indicates that the plots do not follow the straight line. Thus, the normality of data cannot be assumed, making our statistical test less reliable.



Here we test the hypotheses for the population mean improvement of mileage, which we simply call “mean change” hereafter. Regarding the advertised claim, we can set the null hypotheses that the mean change is equal to zero, and the alternative hypothesis that the mean change is greater than zero. Since the standard deviation is not known, we use the test statistic with sample variance, and compare it with the critical point obtained from the t-distribution with 9 degrees of freedom (one-sided t-test). Given the significance level 0.05, the test statistic 0.885 is smaller than the critical value 1.833 on the right-tailed region, suggesting that the observed sample mean is not “unusual” under the assumption of null hypothesis. We can also find the 95% confidence interval (−3.28, 7.50) for the mean change, leaving the possibility that the mean change could be zero.

**Conclusion.** The p-value of 0.1995 indicates that the result is not significant. Therefore, there is not sufficient evidence to support any gain in mileage after the devices were installed. The sample size is small, and therefore, a new study with larger sample size is recommended. We also observed the possibility of two distinct populations in the histogram, and the need for further investigation on a specific class of cars for which the device would work effectively. Thus, in the new study we also recommend to collect data from a particular class of cars, which may reduce the variability of sample distribution.