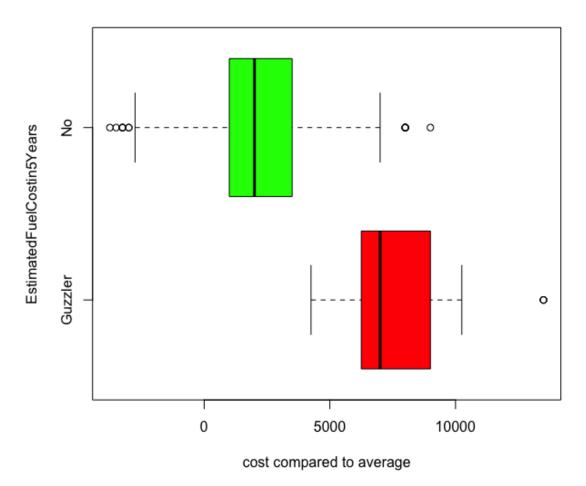# Data Analysis Five

Author:
Thomas Noelcke

Instructor:
Katie Jager

I. PART 1

A.

Guzlers Vs Non Guzlers

B.

The distribution of the data for non guzzlers vs guzzlers does suggest that the non guzzlers on average burn less fuel than guzzlers. This is illustrated by the fact that the first quartile of the guzzlers is out side the 3rd quartile of the non guzzlers. The average consumption of the non guzzlers is also much lower than the the guzzlers. I do find it interesting that there are non guzzlers that consume fuel such that maybe they should be guzzlers.

C.

	Mean	Standard Deviation	Sample Size
GUZZLER	7861.111	2202.326	36
NON-GUZZLER	2202.326	2111.156	860

D.

Guzzlers are entirely composed of luxury cars and super cars. These type of cars tend to be high performance cars with large engines.

II. PART 2

A. Calculate

1) : Shown Below is the Calculation of the test statistic:

$$t = \frac{(X_bar1 - X_bar2) - \sigma}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

$$= \frac{(7861.111 - 2202.326) - 0}{\sqrt{\frac{2170.784^2}{36} + \frac{2202.326^2}{860}}} = 15.3143$$

B.

I'm choosing to use conservative degrees of freedom. In this case the degrees of freedom are as follows:

$$min(n_1 - 1, n_2 - 1) = 35$$

C.

The p Value for our test statistic will be approximately 0. If we look at the t table we will find that the nearest value that matches up is 99.9 at 35 degrees of freedom is only around 3.5.

D.

Shown below is my calculation for the 99% confidence interval.

$$(X_bar1 - X_bar2) \pm t^* \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

$$= (7861.111 - 2202.326) \pm 2.7238 * \sqrt{\frac{2170.784^2}{36} + \frac{2202.784^2}{860}} = [4652.54, 6664.8]$$

E.

Stats From R:

t statistic	15.34
p-Value	2.2e-16
99% Interval	[4658.28, 6659.30]

The values that R Calculated and the values that I found by hand are similar however there are some differences. The differences that you see are due to the fact that I chose the conservative estimation for degrees of freedom rather than the Satterwaite method.

F. Conclude

   1) : There is convincing evidence that the difference between the Guzzlers and non Guzzlers is not zero.

   2) : We reject the null hypothesis in this case with a test statistic of 15.34 and a p-value of 2.2e-16.

   3) : On average Guzzlers cost $5658.79 to operate over 5 years when compared to non guzzlers with a 99% confidence interval of 4658.28 to 6659.30.