Final Project

Project Idea #16 : Fuel efficiency

Melina Padron

2022-12-06

Research Question:

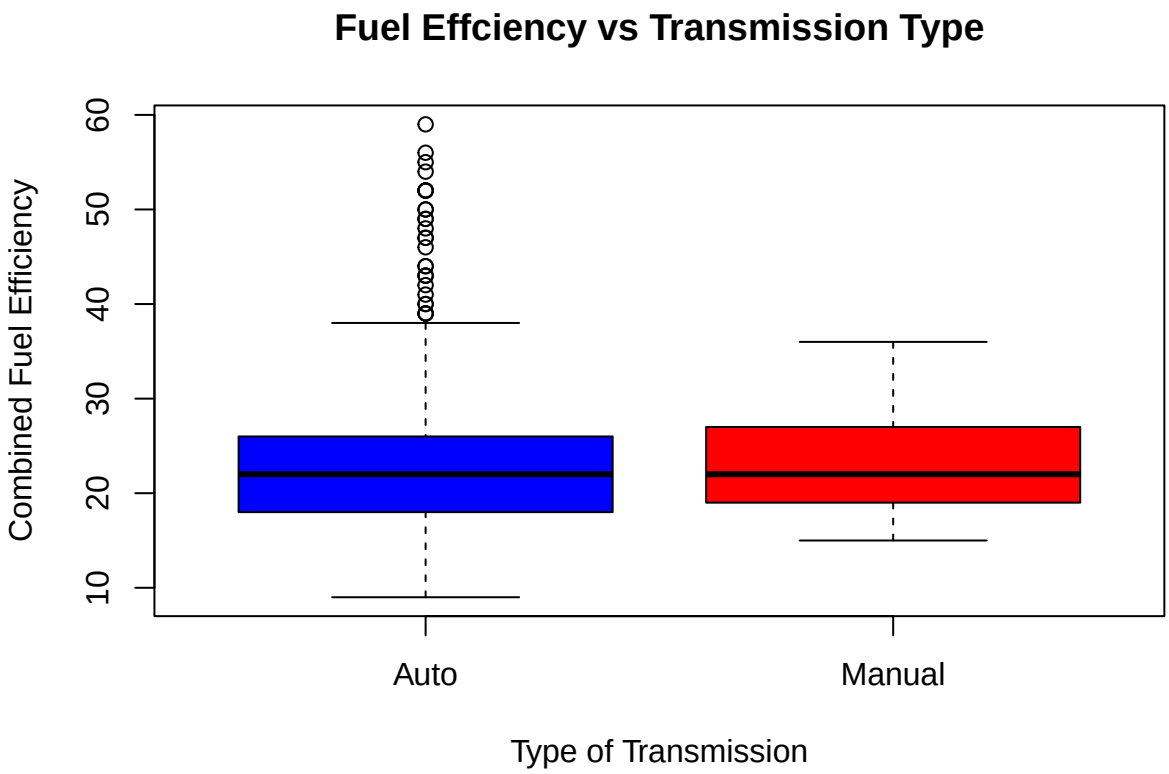
Is there a significant difference between the fuel efficiency of cars with automatic transmission and those with manual transmission?

Visualize the Data:

The fuel economy estimate used is the combined fuel efficiency. A "combined" estimate that represents a combination of city driving (55%) and highway driving (45%).

```
car.data = read.csv("~/cardata.2022.csv")

boxplot(car.data$Comb.FE[car.data$Transmission == "Auto"],
        car.data$Comb.FE[car.data$Transmission == "Manual"],
        names = c("Auto", "Manual"),
        xlab = "Type of Transmission",
        ylab = "Combined Fuel Efficiency",
        main = "Fuel Effciency vs Transmission Type",
        col=(c("blue","red"))
        )
```



Determining the Statistical Procedure:

Proposed Procedure: T Procedure | Two Sample T-Test

Check Conditions:

1.) Independence: The combined fuel economy of one car in the sample of auto transmission vehicles should be independent of the another, and the combined fuel economy of one car in the sample of manual transmission vehicles should be independent of the another. Moreover, the the combined fuel economy of the auto transmission vehicles and the manual transmission vehicles should be independent.

2.) Skewness: Using the box plots from the "Visualize the Data" section, we know that the auto transmission data has many outliers which means it is skewed. We can also note that the whiskers on the manual transmission box plot are slightly asymmetrical, meaning the data is also skewed. However, since our sample size for both auto transmission vehicles and manual transmission vehicles are over 100, our sample sizes are large enough to utilize the T-procedure regardless of any skewness.

Thus, we can utilize the T-procedure.

Hypothesis Testing:

Since we are working with a dependent quantitative variable (Combined Fuel Economy) and an independent categorical variable (with two groups: Auto transmission and Manual transmission), we can use a two sample t-test to determine if the two population means are equal or not.

The following hypothesis test will use a 5% significance level.

$$H_0 : \mu_a = \mu_m \text{ vs. } H_a : \mu_a \neq \mu_o$$

```
auto.data = car.data$Comb.FE[car.data$Transmission == "Auto"]
manual.data = car.data$Comb.FE[car.data$Transmission == "Manual"]

tt = t.test(auto.data,manual.data)
tt
```

```
##
##  Welch Two Sample t-test
##
## data:  auto.data and manual.data
## t = -0.27758, df = 107.08, p-value = 0.7819
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -1.381105  1.041830
## sample estimates:
## mean of x mean of y
##  22.94531  23.11494
```

Conclusion:

With a significance level of 5% and a large p-value of 0.7819, we fail to reject the null hypothesis that there is no significant difference between the means of combined fuel economy of auto transmission vehicles and manual transmission vehicles.