Lab 2: Analysis of Results

David Barragan Alcantar

University of Essex, Colchester Campus
CE888 Data Science and Decision Making

Introduction

A business is looking at changing its current vehicle fleet and replacing their vehicles with ones used by their competitors. They have captured the MPG of some of the cars in both fleets.

The business analysts come up with a comparison algorithm that requires the upper and lower bounds for the mean (i.e. bootstrap) in order to say which fleet is better.

Results

A scatter plot is shown in Figure 1 to visualize the values from both fleets. Although it should be clarified that there is no relationship between an element from the current fleet and an element from the new fleet. It is displayed in this way only for visual purposes.

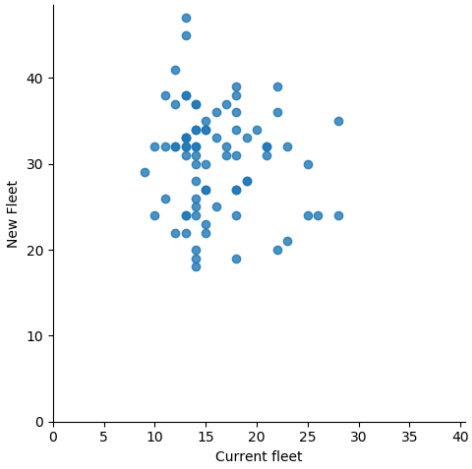

Figure 1: Scatter plot of both fleets.

The distribution of the MPG of both fleets are shown in Figure 2 and Figure 3 presented as absolute frequency histograms.

Finally, Figure 4 and Figure 5 show the upper and lower mean bounds of the current fleet and the new fleet respectively. These bounds were calculated with the help of bootstrap through 50,000 iterations.

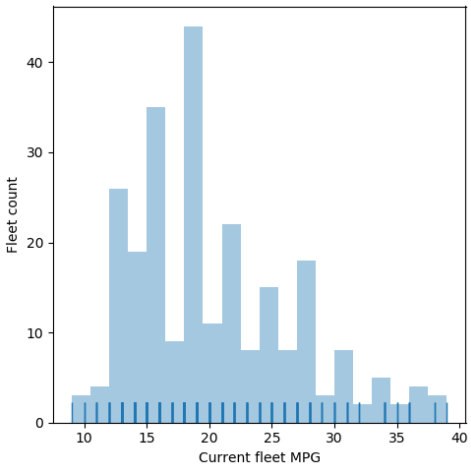

Figure 2: Histogram of current fleet MPG.

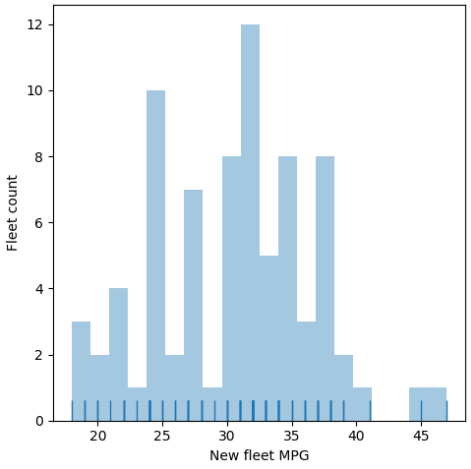

Figure 3: Histogram of new fleet MPG.

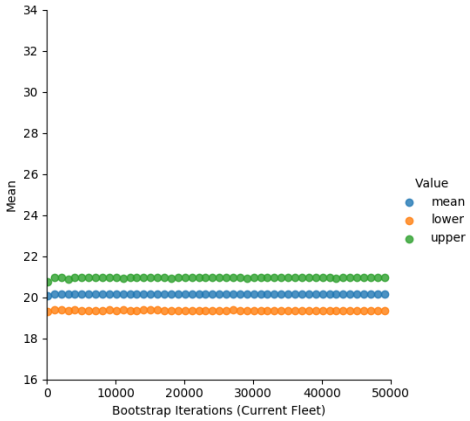

Figure 4: Upper and lower mean bound of the current fleet.

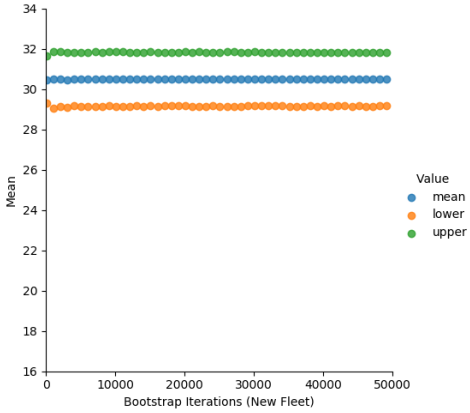

Figure 5: Upper and lower mean bound of the new fleet.

Analysis

From Figure 4 and Figure 4, it can be seen that the mean bound of the current fleet MPG goes from approximately 19 to 21, whereas the mean bound of the new fleet MPG goes from about 29 to 32.

Since there is no overlap in the boundaries of both fleets, It is concluded according to the data collected that the new fleet of vehicles can travel more miles per gallon, based on the bootstrapping analysis through 50,000 iterations.