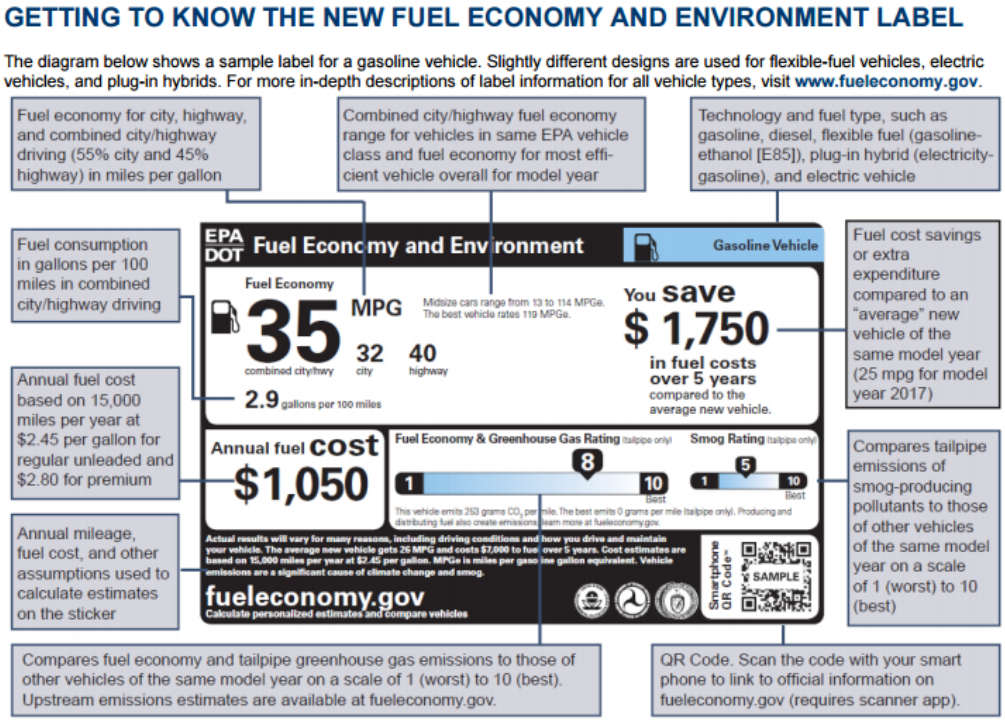
ENHANCING STATISTICS TEACHER EDUCATION WITH E-MODULES

A.2.f. Investigating CO_2 Emissions in Vehicles

Deciding on which car is best to purchase can be a daunting task. There are many factors to consider when deciding on a new car: type, average miles per gallon (city and highway), annual fuel costs, carbon dioxide emissions, etc. In this investigation we will take a careful look at the estimated rate at which vehicles release carbon dioxide (CO_2) into the air, which EPA claims is about 404 grams per mile for passenger vehicles. The label below is used on all new vehicles in the U.S. to display several key characteristics of vehicles. See the image below to better understand information displayed on the label. What attributes do you consider the most important when buying a car?



Lee, H.S., Mojica, G.F., Azmy, C., Barker, H., & Harrison, T.R. (2019). *Investigating CO2 Emissions in Vehicles*. Enhancing Statistics Teacher Education with E-Modules. Retrieved from: https://fi-esteem.s3.amazonaws.com/module_a/part2/vehiclestask.pdf

Part I. Getting to Know a Few Vehicles

Let's get familiar with the different attributes of a vehicle that is reported by the official U.S. fuel economy government website.

1. Go to www.fueleconomy.gov. Find the link on the homepage that says "**Compare Side by Side**". Choose two recent models of vehicles of interest to compare. Once you've chosen your two cars hit the "Compare" button. On the compare side by side page be sure to take note of features under the fuel economy, energy and environment (where CO_2 emissions is found), safety and specs tabs.
    a. When comparing the two cars you chose, what difference and similarities stand out to you?

    b. How do the carbon emissions compare for the two vehicles you chose?

    c. How do the tailpipe CO_2 values compare with the reported overall typical emissions for passenger vehicles of 404 grams/mile?

    d. What do you think might contribute to any differences or similarities in CO_2 emissions among different types of vehicles?

We could investigate the overall population of US manufactured vehicles to discover the *average* Tailpipe CO_2 emissions. Considering the fact that there may be differences in Tailpipe CO_2 emissions among different types of vehicles, those differences might influence how we choose to represent the average. In addition, we usually do not have access to a data from an entire population, so we must rely on information we obtain from samples.

Part II. Investigating Tailpipe CO_2 emissions from samples

2. **Make a prediction**. Thinking back to the cars you selected to investigate and any other knowledge you may have of CO_2 emissions make an estimate for the average amount of CO_2 emissions for this dataset. When deciding on what the average may be, think about if the average CO_2 emissions may lie in the high end of all CO_2 emissions , the lower end, or somewhere near the middle.

Lee, H.S., Mojica, G.F., Azmy, C., Barker, H., & Harrison, T.R. (2019). *Investigating CO2 Emissions in Vehicles*. Enhancing Statistics Teacher Education with E-Modules. Retrieved from: https://fi-esteem.s3.amazonaws.com/module_a/part2/vehiclestask.pdf

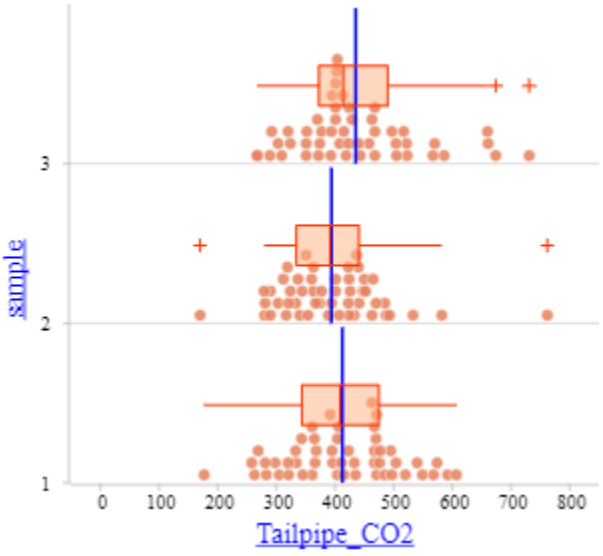
Open this CODAP document with Sampler set up and ready to randomly choose vehicles from a population of 1273 passenger fuel-based vehicles manufactured in the U.S. (electric vehicles are not included).

3. **Collect a sample.** Now you will estimate the typical CO_2 emissions of all vehicles manufactured in 2018 using a sample of 50 vehicles. Use the Sampler to select 50 items, and collect 1 sample. The sampled data will appear in the hierarchical table with 3 levels (Experiments, Samples, Items). Each vehicle (at the Items level) has a blank column labeled Value and then 18 attributes with data values in each column.

   a. Investigate the *tailpipe* CO_2 emissions (Amount of grams of CO_2 per mile the vehicle emits) for this sample. What is a reasonable estimate for the tailpipe CO_2 emissions for all passenger vehicles based on your sample? Justify your answer with screenshots and an explanation below.

   b. What other factors seem related to CO_2 emissions? Use other quantitative and categorical attributes to investigate any relationships with tailpipe CO_2. Describe what you found and insert screenshots of graphs to illustrate your points.

4. **Compare samples.** Examine the distributions for the CO_2 emissions for three other random samples of 50 vehicles. A boxplot and the mean are shown for each sample. How do these samples compare to your sample?



Lee, H.S., Mojica, G.F., Azmy, C., Barker, H., & Harrison, T.R. (2019). *Investigating CO2 Emissions in Vehicles*. Enhancing Statistics Teacher Education with E-Modules. Retrieved from: https://fi-esteem.s3.amazonaws.com/module_a/part2/vehiclestask.pdf

5. How would you revise your estimate of the typical tailpipe CO_2 emissions based on the information from the additional 3 samples?

6. If we took a sample of 50 vehicles with a mean of 437 grams per mile, how might that impact your estimate of the average Tailpipe CO_2 emissions?

Part III. Reasoning About Means from Many Samples

Sometimes we want to report the average of a certain attribute. When we take a sample to answer a question about the average or typical value for an attribute, our answer should take into account variability. If we took another sample, would we see the same average? How confident are we that the sample we took represents the entire population?

7. Let's continue to focus on CO_2 emissions. Each sample will likely result in different means for CO_2 emissions. In your CODAP sampler table, compute the ***mean*** CO_2 emissions of ***your sample*** (at the Sample level in the table)

What is the computed sample mean for your sample of 50 vehicles? _____

8. Investigate the distribution of the means after taking many more samples. To do this, first clear your data and "Refresh List."
Take many samples of size 50 and graph the distribution of the sample means of CO_2 emissions, using the column you used to compute them.

a. Describe the overall shape of the distribution of sample means of CO_2.

b. Interpret this distribution to estimate a typical Tailpipe CO_2 emissions for US manufactured vehicles in 2018. Justify your claim with screenshots and an explanation below.

9. How might sample size impact our claims? Collect many samples of size 50, 150, and 400. Separate your graph of the sample means by Sample Size (at the Experiments level in the table). Insert screenshots to support answers to the following questions.

Lee, H.S., Mojica, G.F., Azmy, C., Barker, H., & Harrison, T.R. (2019). *Investigating CO2 Emissions in Vehicles*. Enhancing Statistics Teacher Education with E-Modules. Retrieved from: https://fi-esteem.s3.amazonaws.com/module_a/part2/vehiclestask.pdf

a. How does sample size impact the distribution of the sample means?

b. How does sample size impact your ability and confidence to estimate the average Tailpipe CO_2 emissions?

Part IV. Extension--Examining the population

We will now open a dataset that contains the population of 1,273 vehicles manufactured in 2018 for which the EPA has tested. Use this link to download the csv file. Open a new CODAP document and import the dataset. Once you drag the dataset into CODAP you should have 1,273 cases.

10. As you did in Part II, investigate the tailpipe CO_2 emissions, this time for the entire population.

a. What is the mean tailpipe CO_2 emissions for all passenger vehicles?

b. Look back to the sample you collected for question 3 and 7. How does your sample mean for tailpipe CO_2 emissions compare to the mean for tailpipe CO_2 emissions for all vehicles? How does the mean tailpipe CO_2 emissions for 2018 compare with the value reported by the EPA? Justify your answer with an explanation below.

11. You estimated the average tailpipe CO_2 emissions by first taking a single sample of size 50, taking several samples of size 50 then estimating the sample means, and finally by looking at the entire population of passenger vehicles. Describe the distribution of the single sample, the distribution of sample means, and the distribution of the population. Which distributions are similar to each other? Justify your answer below with an explanation and screenshots.

12. If you wanted to minimize your carbon footprint from the vehicle you drive, what factors would you look for in a car to ensure a lower CO_2 emission rate?

Lee, H.S., Mojica, G.F., Azmy, C., Barker, H., & Harrison, T.R. (2019). *Investigating CO2 Emissions in Vehicles*. Enhancing Statistics Teacher Education with E-Modules. Retrieved from: https://fi-esteem.s3.amazonaws.com/module_a/part2/vehiclestask.pdf

2018 Vehicles Dataset Attributes and Descriptions

Attribute	Description
Mfr Name	Manufacturer name
Division	Company name
Carline	Model name of the vehicle
Hybrid	Identifies whether the vehicle has a hybrid engine (Y, N) such that it utilizes more than one form of onboard energy to achieve propulsion. A hybrid will have a traditional engine and fuel tank, as well as one or more electric motors and a battery pack
# Cyl	Number of cylinders in an engine
City FE	Estimated miles per gallon in city driving
Hwy FE	Estimated miles per gallon in highway driving
Comb FE	Estimated miles per gallon in a combination of city driving (55%) and highway driving (45%)
Tailpipe CO2	Amount of grams of CO2 per mile the vehicle emits. The best emits 0 grams of CO2 per mile (tailpipe only). Producing and distributing fuel also creates emissions that aren't part of this calculation.
Transmission	Identifies vehicles as manual or automatic transmission
No_Gears	Number of transmission gears
AnnualFuel_Cost	Estimated annual fuel cost assuming 15,000 miles per year (55% city and 45% highway) and average fuel price
Money saved over 5 years	Fuel cost savings or extra expenditure compared to an "average" new vehicle of the same model year (25 mpg for model year 2018)
DriveDesc	Drivetrain (2-wheel, 4-wheel, or all-wheel)
Type	Identifies the vehicle as car, SUV, van, minivan, or truck
FE Rating	Compares fuel economy to those of other vehicles of the same model year on a scale of 1 (worst) to 10 (best).
GHG Rating	Compares tailpipe greenhouse gas emissions to those of other vehicles of the same model year on a scale of 1 (worst) to 10 (best).
Smog Rating	Compared tailpipe emissions of smog-producing pollutants to those of other vehicles of the same model year on a scale of 1 (worst) to 10 (best).

Lee, H.S., Mojica, G.F., Azmy, C., Barker, H., & Harrison, T.R. (2019). *Investigating CO2 Emissions in Vehicles*. Enhancing Statistics Teacher Education with E-Modules. Retrieved from: https://fi-esteem.s3.amazonaws.com/module_a/part2/vehiclestask.pdf