DATA-101 IN-CLASS ASSIGNMENT SEPTEMBER 8, 2021

Consider the automobile data set posted on blackboard. The Automobile dataset has a different characteristic of an auto such as body-style, wheel-base, engine-type, price, mileage, horsepower and many more.

1. **In Python**, answer the following:

   (a) (3 points) Using the pandas library, read the csv datafile and create a data-frame called `autos`

   (b) (3 points) Report the mean `price`

   (c) (3 points) Report the median `price`

   (d) (3 points) A common rule of thumb to determine the skewness of a numeric dataset is to compare the mean and the median using the following rules:

   - If the mean $>$ the median $\Rightarrow$ right-skewed distribution
   - If the mean $\approx$ the median $\Rightarrow$ symmetric distribution
   - If the mean $<$ the median $\Rightarrow$ left-skewed distribution

   Using the above rules, what is the skewness of `price`?

2. Health care issues are receiving much attention in both academic and political arenas. A sociologist recently conducted a survey of citizens over 60 years of age whose net worth is too high to qualify for Medicaid. The ages of 10 senior citizens were as follows:

   60 61 62 63 64 65 66 68 68 69

   In R, answer the following:

   (a) (3 points) Create a vector called `age` that contains the 10 senior ages

   (b) (3 points) Compute the mean `age`

   (c) (3 points) Compute the variance `age`

   (d) (3 points) Compute mean of the log(`age`)