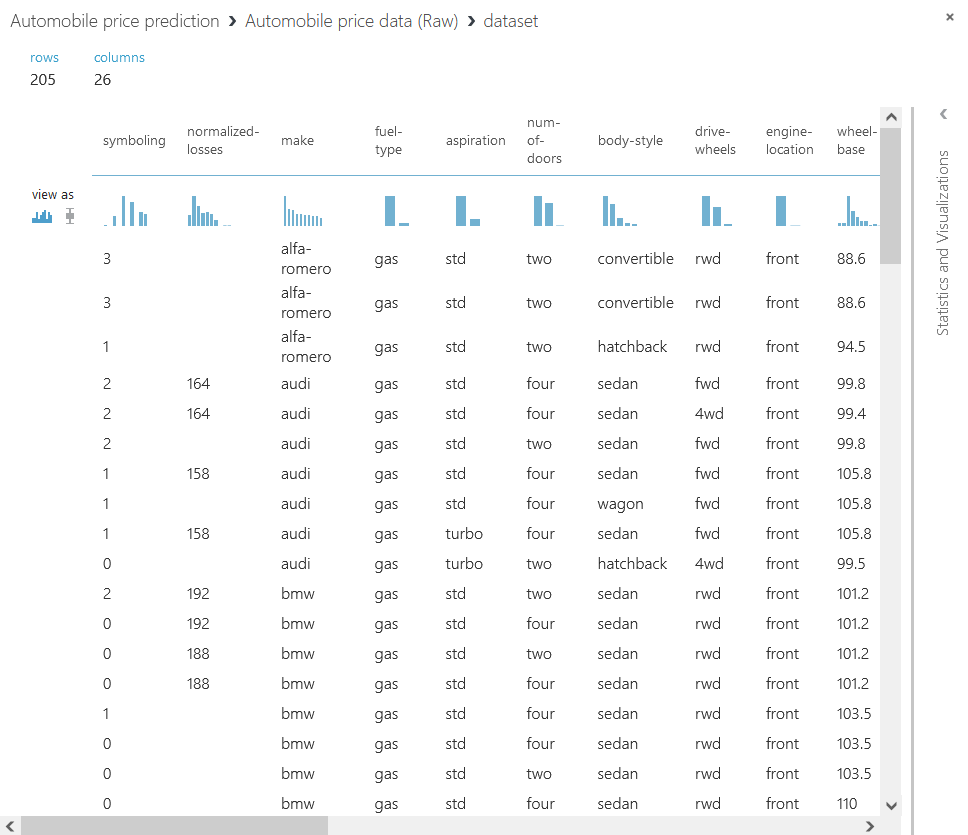
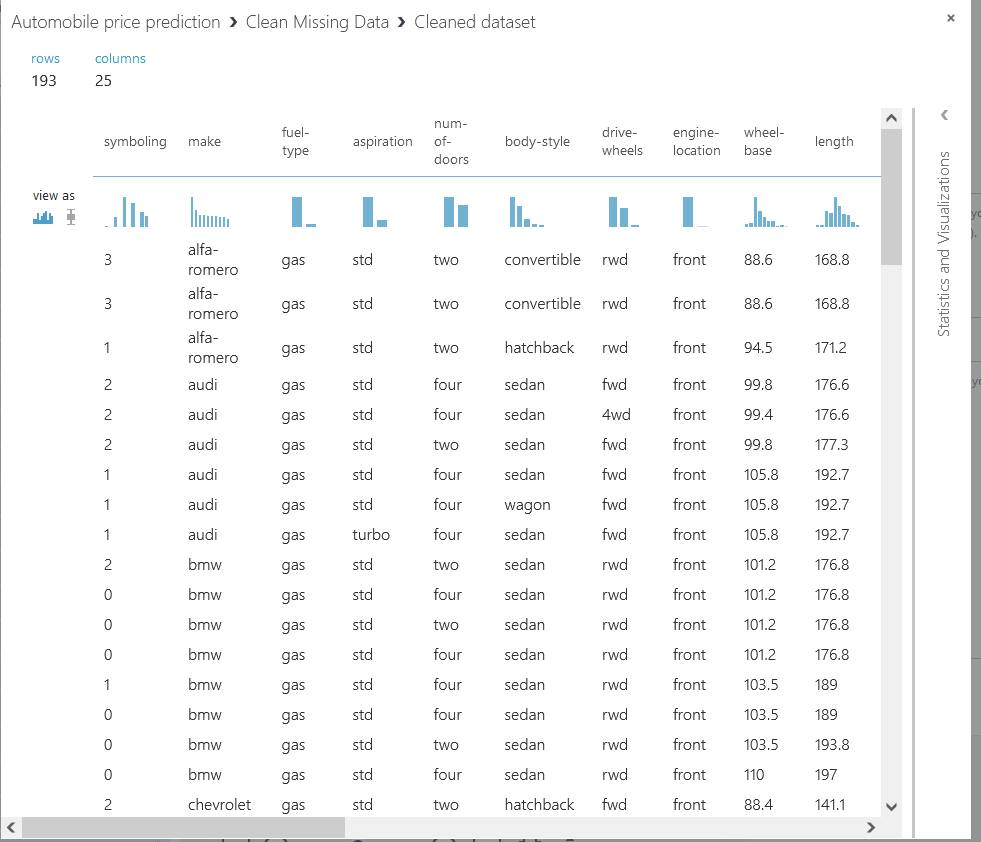
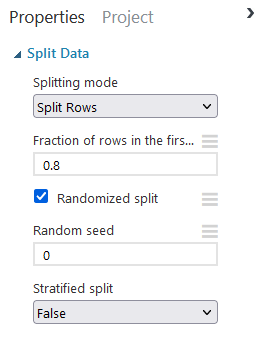
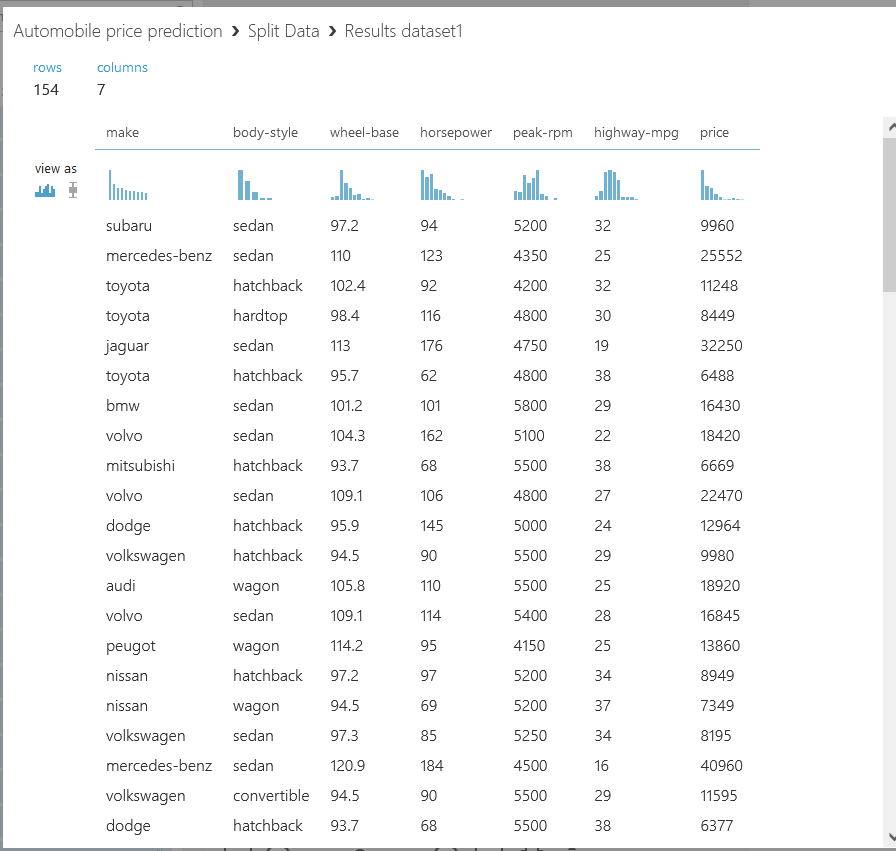
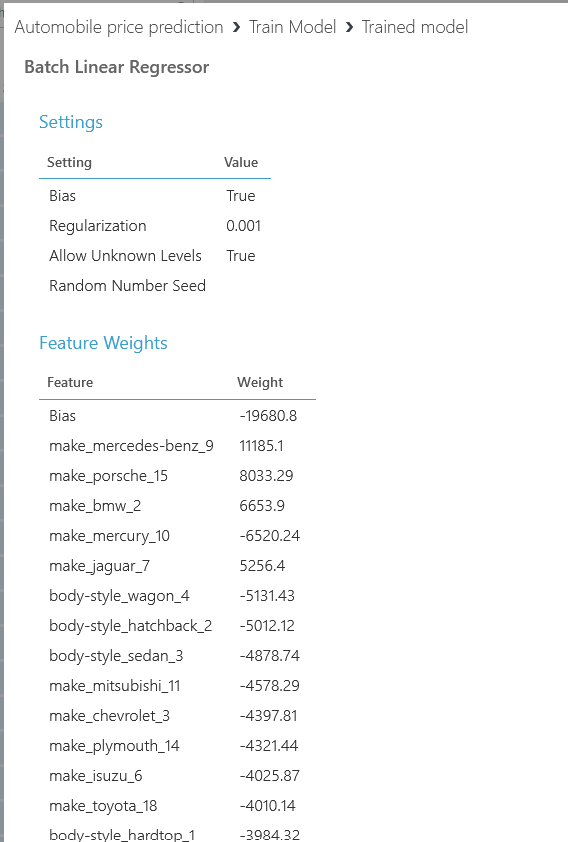
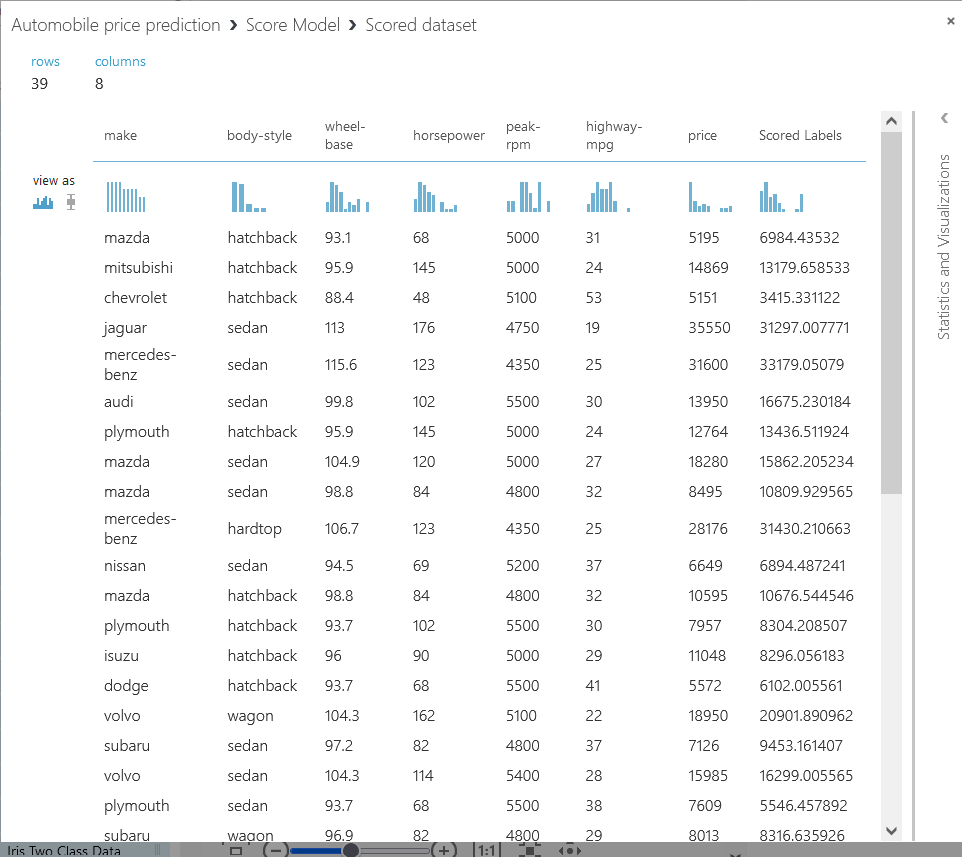
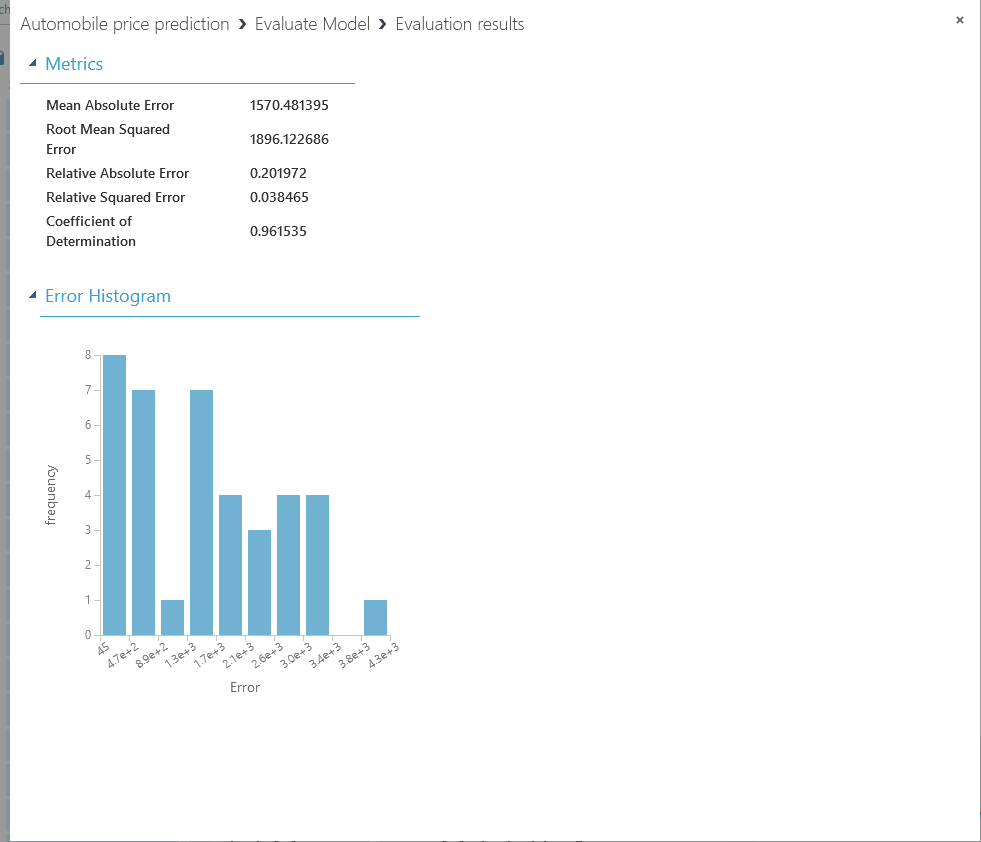
a. Import the entire dataset (Automobile Price Data (Raw))  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
b. Clean the missing data by dropping rows with missing values (select all columns in the dataset and do not "exclude the normalized losses" from the original tutorial). Leave the maximum missing value ratio to 1.  
  


c. Partition and sample the data. (Note: do not use “Split Data”)  
  
  
  


d. Create a new model: Linear Regression (add the default Linear regression, i.e., do not change any values here)  
  
  
  


e. Finally, perform cross-validation on the dataset. (Hint: use the price column here)

  
  
In Lab 4, I discovered a new data analysis tool called Microsoft Machine Learning Studio. This tool allows for data manipulation without the need for programming skills. I learned how to upload, clean data, execute linear regression, and apply cross-validation using this platform. I found it quite handy for conducting analyses. It took me a few minutes to get the hang of it, especially since Microsoft has discontinued support for this tool, which led to some login and operational delays during my lab. Despite these challenges, navigating through them was part of the learning process. I thoroughly enjoyed the experience and the interface. I hope Microsoft continues to support this incredible software.