**5. Problem Statement**

Data collected for several hundred used General Motors (GM) cars allows us to develop a multivariate regression model to determine car values based on a variety of characteristics such as mileage, make, model, cruise control, and so on.

• **Price**: suggested retail price of the used GM car.

• Mileage: number of miles the car has been driven

• Make: manufacturer of the car such as Cadillac, Pontiac, and Chevrolet

• Cylinder: number of cylinders in the engine

• Liter: a more specific measure of engine size

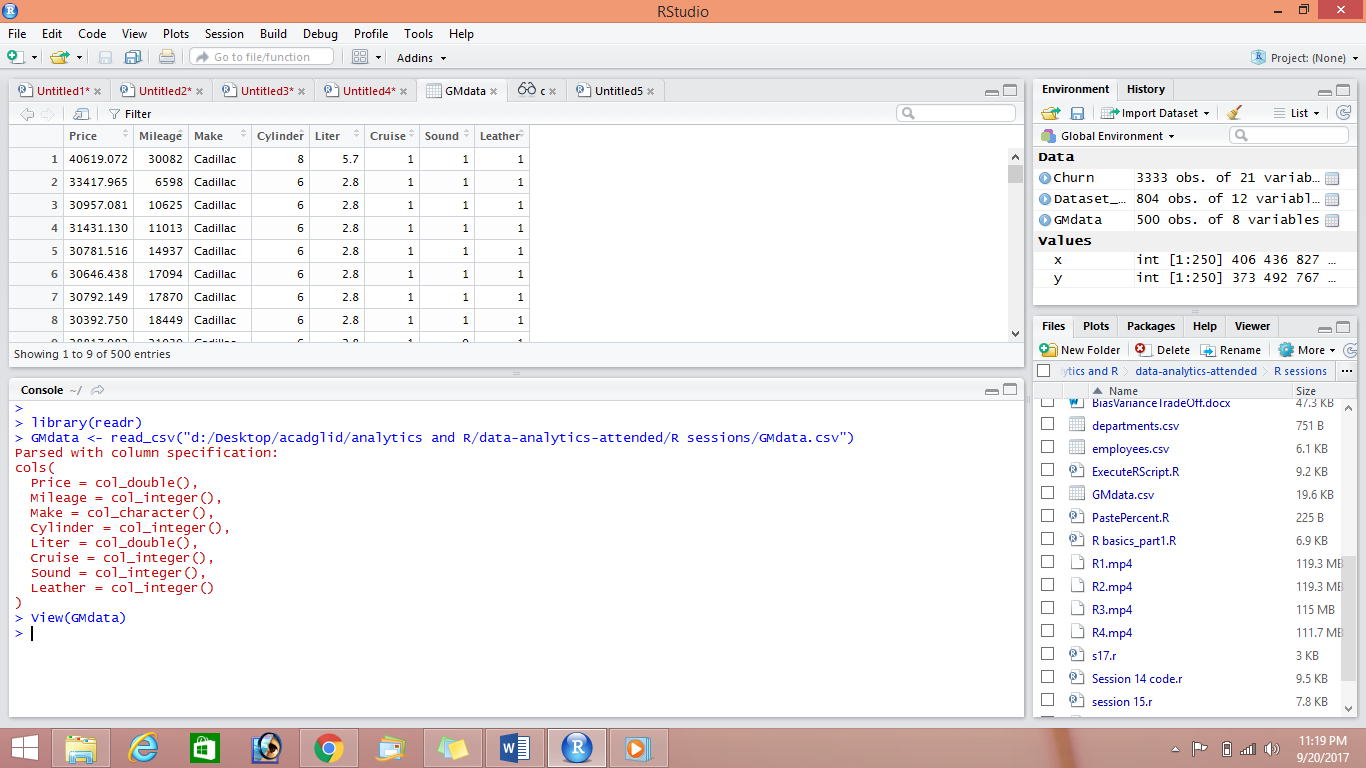
• Cruise: indicator variable representing whether the car has cruise control (1 = cruise)

• Sound: indicator variable representing whether the car has upgraded speakers (1 = upgraded)

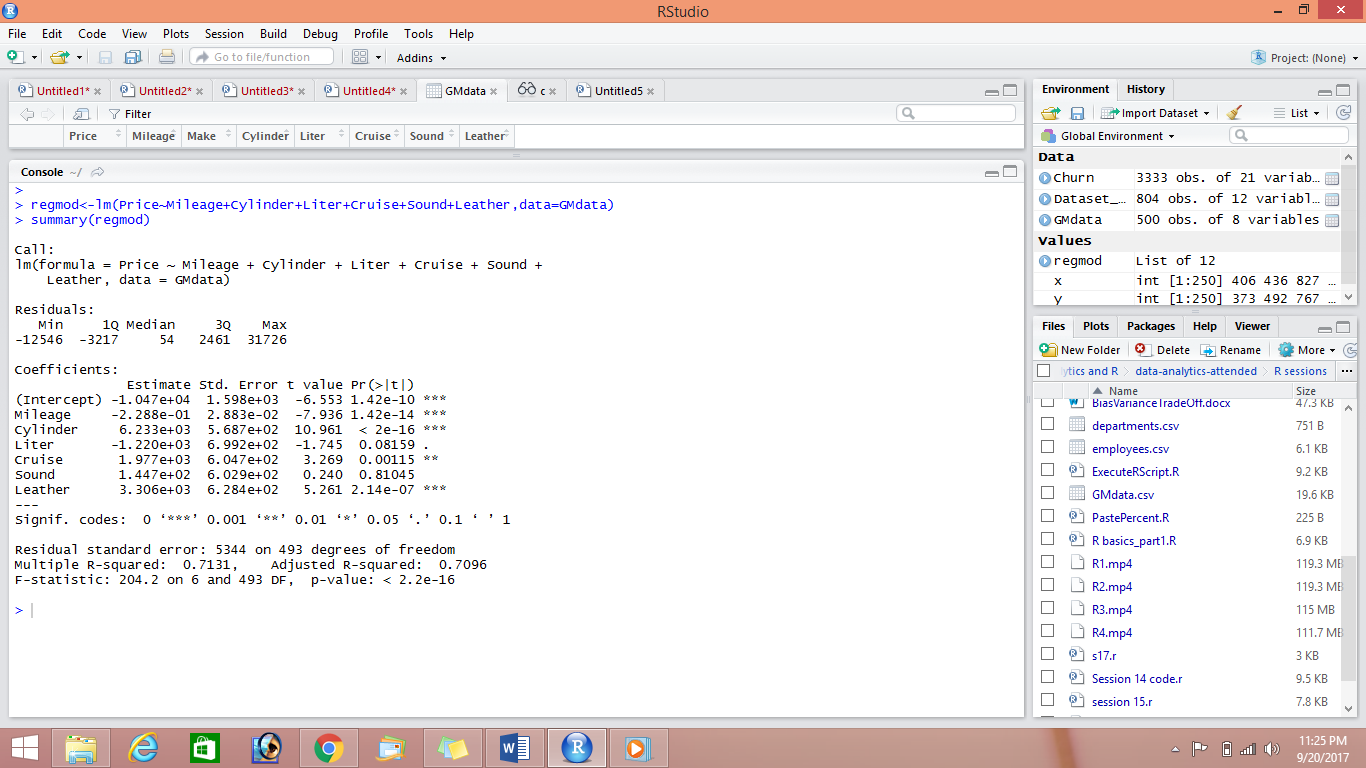
• Leather: indicator variable representing whether the car has leather seats (1 = leather)

**Answer**

Importing GMData file



**Step 2:running multiple regression**



P value is significant (2.2e-16) which suggests the model is a good fit(p<.05)at 95% confidence level

Model will be

Price=-1.047e+04-2.288e-01(mileage)+ 6.233e+03(cylinder)

+ 1.977e+03(cruise)+ 3.306e+03(Leather)

**Multiple R-squared or  coefficient of determination: 0.7131**

Is quite high . this shows the model is very near to the model fit.

R-squared is always between 0 and 100%:

Conclusion: mileage , cylinder ,cruise and leather predicts price very well and has a **very good model fit**