**ASSIGNMENT:**  
1. Find the ‘best’ sales target price for predicting sales -  
2. Determing values of age, gender, income level or commute distance i.e. miles/wk that predict best sales  
3. Determing the appropriate neural net - how many hidden layers, what variables to leave out?  
4. Does it help to use 80/20 (or any other) split for training/test sets?  
5. Recommend an advetising approach based on your findings

#1.

#the best sales are depended several dimensions, so it's hard to say the specific price, but there is a model to suggest the reasonable price.

#2

#all things below are the part of the codes. I put here for easy to look.

#linearModel = glm(sales~ age + gender + miles + debt + income, data = #trainDataSet)

#summary(linearModel)

#print(linearModel)

#plot(linearModel)

#this part gives the idea of how to determine it. It suggests that only debt and income works for the price.

#linearModel = glm(sales~ debt + income, data = trainDataSet)

#summary(linearModel)

#print(linearModel)

#plot(linearModel)

# then it suggests price = 124 + 0.3282\*debt +0.9451\*income

#3

#Maybe the more hiden layer, the better outcome presents, but my laptop is too sick to deal with amount caculations. Too many hiden layers may also caused overfit. In the neural network, it suggests us to delete the gender valuable.

#4

#The more data we give to train set, the better model performs. However, if we put so much data into train set, it will be easy to overfit. In my own opinion, 70% is enough.

#5

#It’s always good to recommend cars to high income guys as well as high debt huys!