09-Assignment-BrundageK

Kelley Brundage

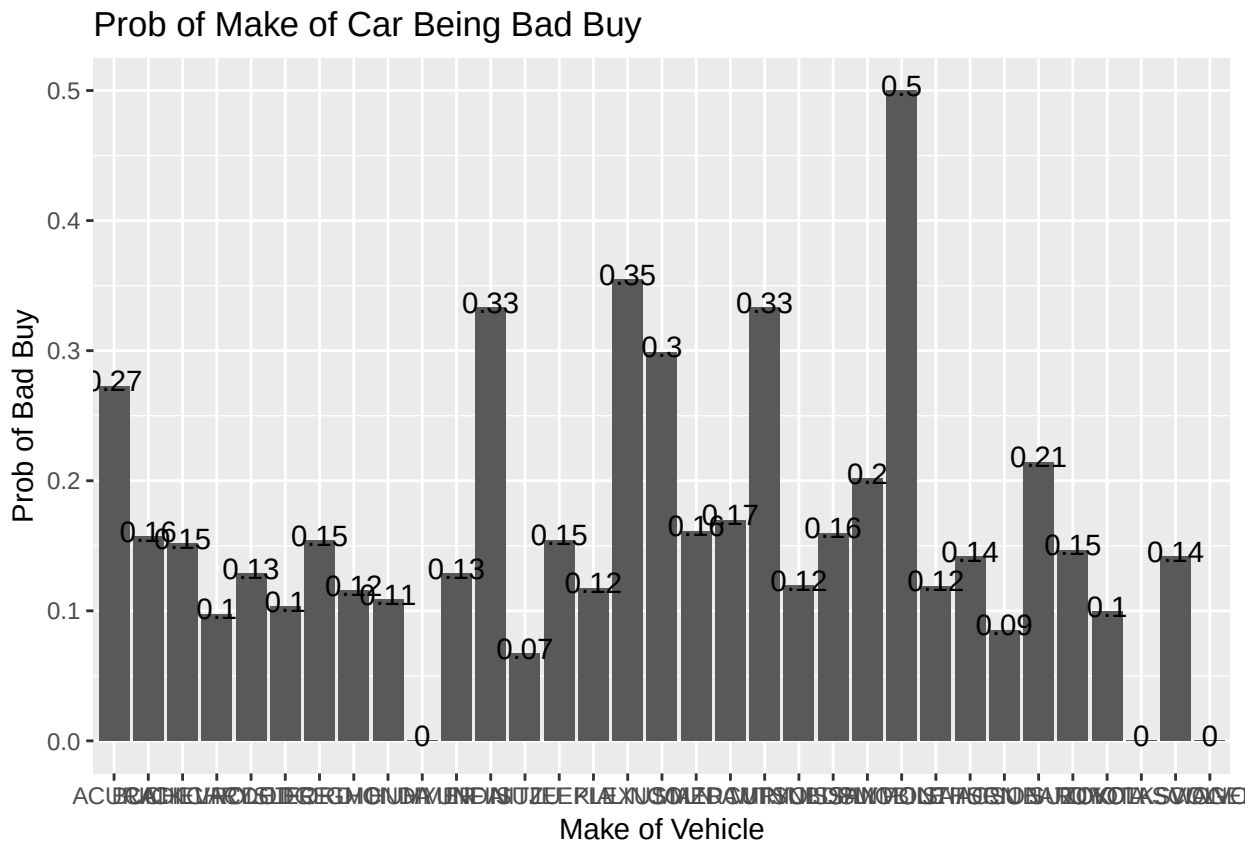
6/30/2019

For this assignment, you'll be using the lemons dataset, which is a subset of the dataset used for a Kaggle competition described here: "https://www.kaggle.com/c/DontGetKicked/data"

Question #1: Using the lemons dataset, plot the probability of a car being a bad buy by make.

If we want to make this a little better, we can change the row and column titles

Then we can plot this using our familiar ggplot commands:



Question #2: Create a table that shows the probability of a car being a bad buy by make.

	Not a Bad Buy	Is a Bad Buy
ACURA	72.73	27.27
BUICK	84.31	15.69
CADILLAC	84.85	15.15
CHEVROLET	90.25	9.75
CHRYSLER	87.14	12.86
DODGE	89.68	10.32
FORD	84.59	15.41
GMC	88.44	11.56
HONDA	89.13	10.87

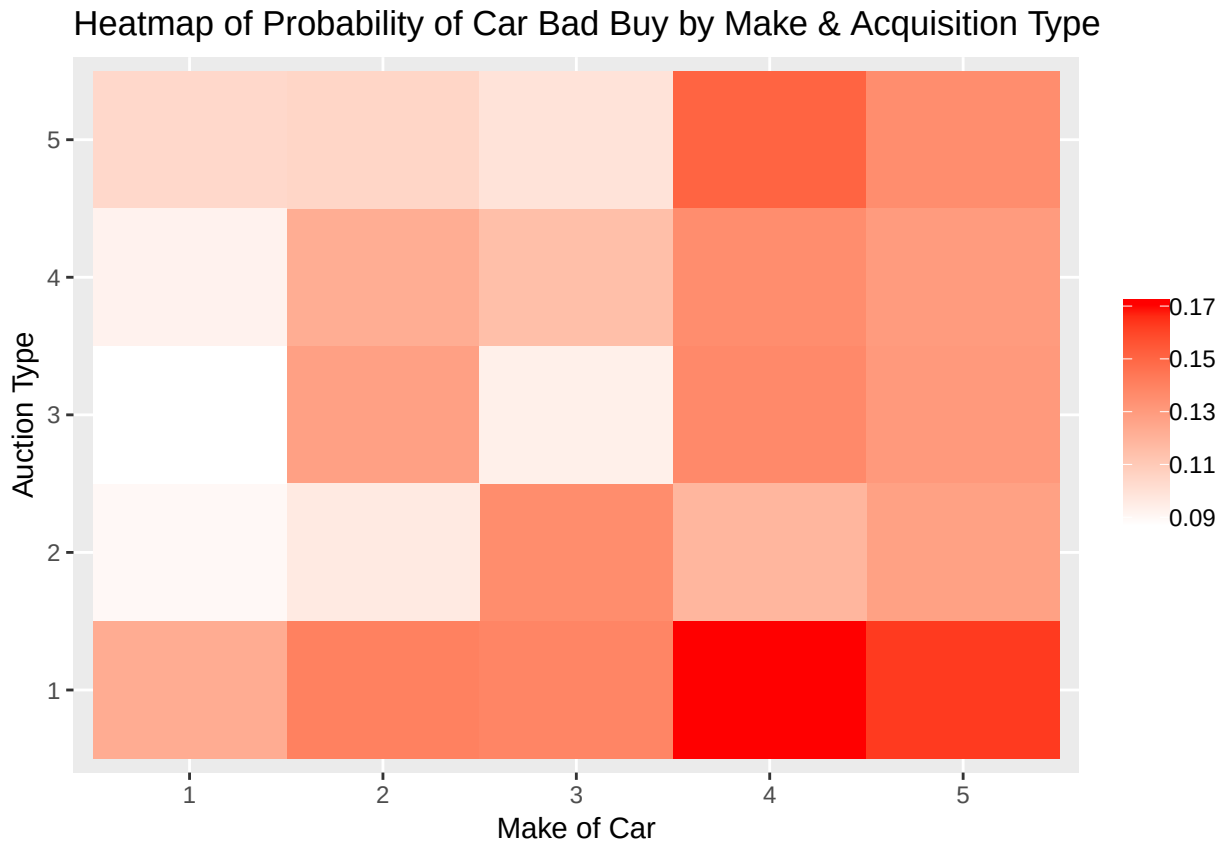
	Not a Bad Buy	Is a Bad Buy
HUMMER	100.00	0.00
HYUNDAI	87.13	12.87
INFINITI	66.67	33.33
ISUZU	93.28	6.72
JEEP	84.55	15.45
KIA	88.24	11.76
LEXUS	64.52	35.48
LINCOLN	70.10	29.90
MAZDA	83.86	16.14
MERCURY	83.02	16.98
MINI	66.67	33.33
MITSUBISHI	88.06	11.94
NISSAN	84.03	15.97
OLDSMOBILE	79.84	20.16
PLYMOUTH	50.00	50.00
PONTIAC	88.09	11.91
SATURN	85.85	14.15
SCION	91.47	8.53
SUBARU	78.57	21.43
SUZUKI	85.32	14.68
TOYOTA	90.03	9.97
TOYOTA SCION	100.00	0.00
VOLKSWAGEN	85.82	14.18
VOLVO	100.00	0.00

Question #3: Create a heatmap of the probability of a car being a bad buy by make and acquisition type.

Then we'll create a summary dataset that shows the probabilities of the outcome across all of the combined categories of the two independent variables.

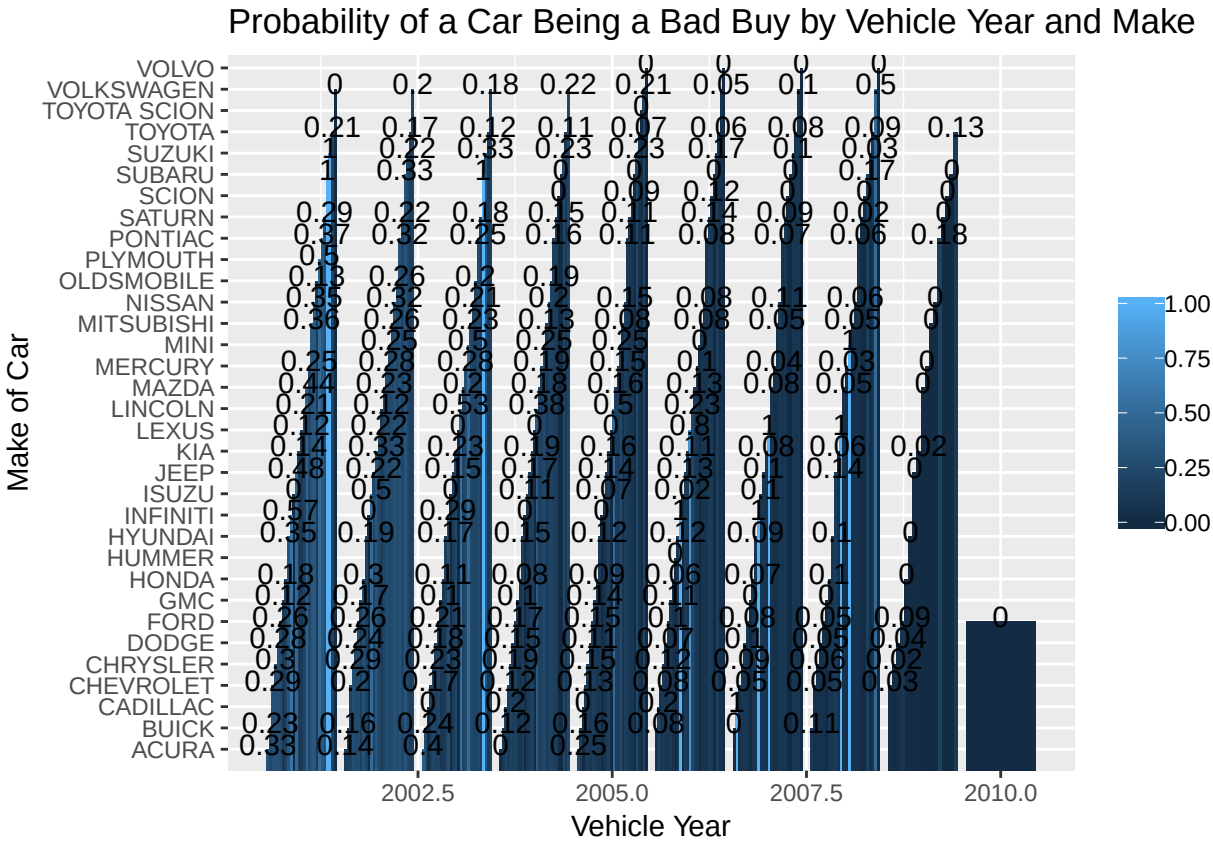
Missing data isn't important, so we'll drop it.

Now we're ready to plot!



Question #4: Create a plot of your choosing that shows the probability of a car being a bad buy by year and make.

Now we can plot the probability of a car being a bad buy by year and make


Probability of a Car Being a Bad Buy by Vehicle Year and Make
Make of Car
VOLVO
VOLKSWAGEN
TOYOTA SCION
TOYOTA
SUZUKI
SUBARU
SCION
SATURN
PONTIAC
PLYMOUTH
OLDSMOBILE
NISSAN
MITSUBISHI
MINI
MERCURY
MAZDA
LINCOLN
LEXUS
KIA
JEEP
ISUZU
INFINITI
HYUNDAI
HUMMER
HONDA
GMC
FORD
DODGE
CHRYSLER
CHEVROLET
CADILLAC
BUICK
ACURA
2002.5
2005.0
2007.5
2010.0
Vehicle Year
1.00
0.75
0.50
0.25
0.00