

Lab # 8 Scatterplots

Scatterplots show value for two variables.

Using size or shade, a third variable can be shown.

This is a great way to visualize the relationship between two variables. If you drew a line through the “best fit” of the points, that is the regression line. This is a visual representation of linear regression.

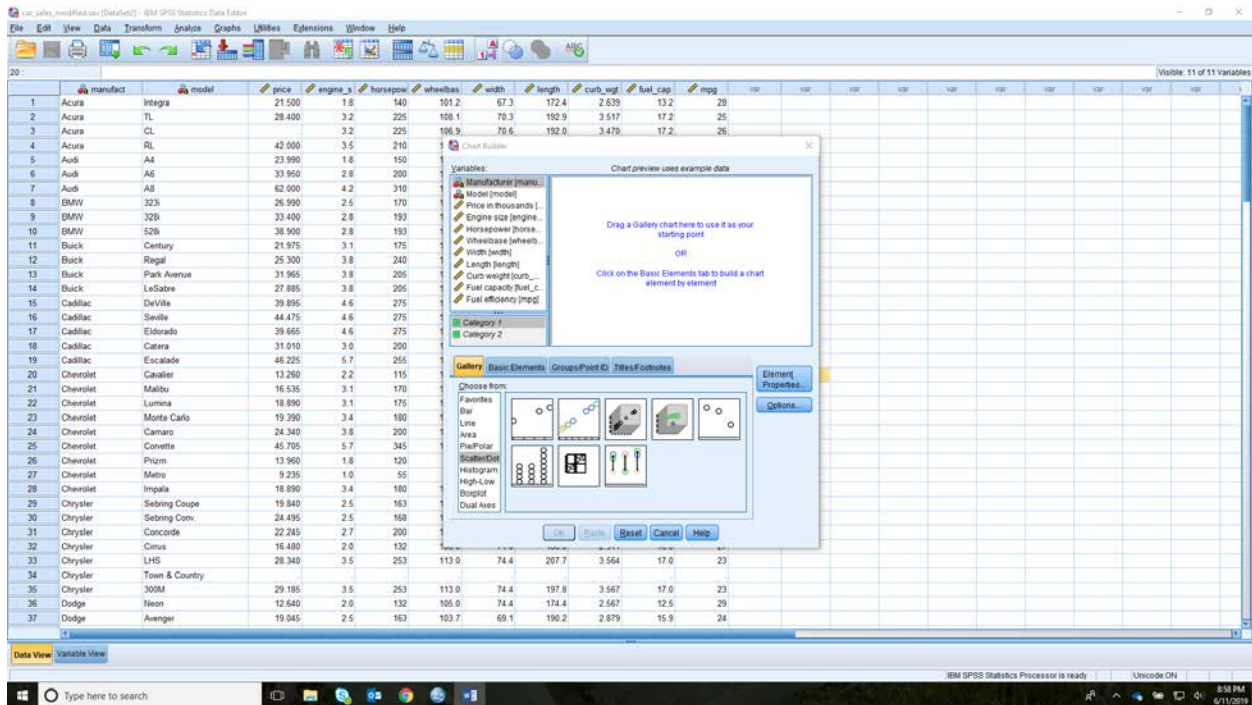
<https://www.youtube.com/watch?v=g3Dxlyfr1k8>

Open the car_sales_modified.sav file, and review the data.

What variables do we have. What would be a good predictor of sale price?

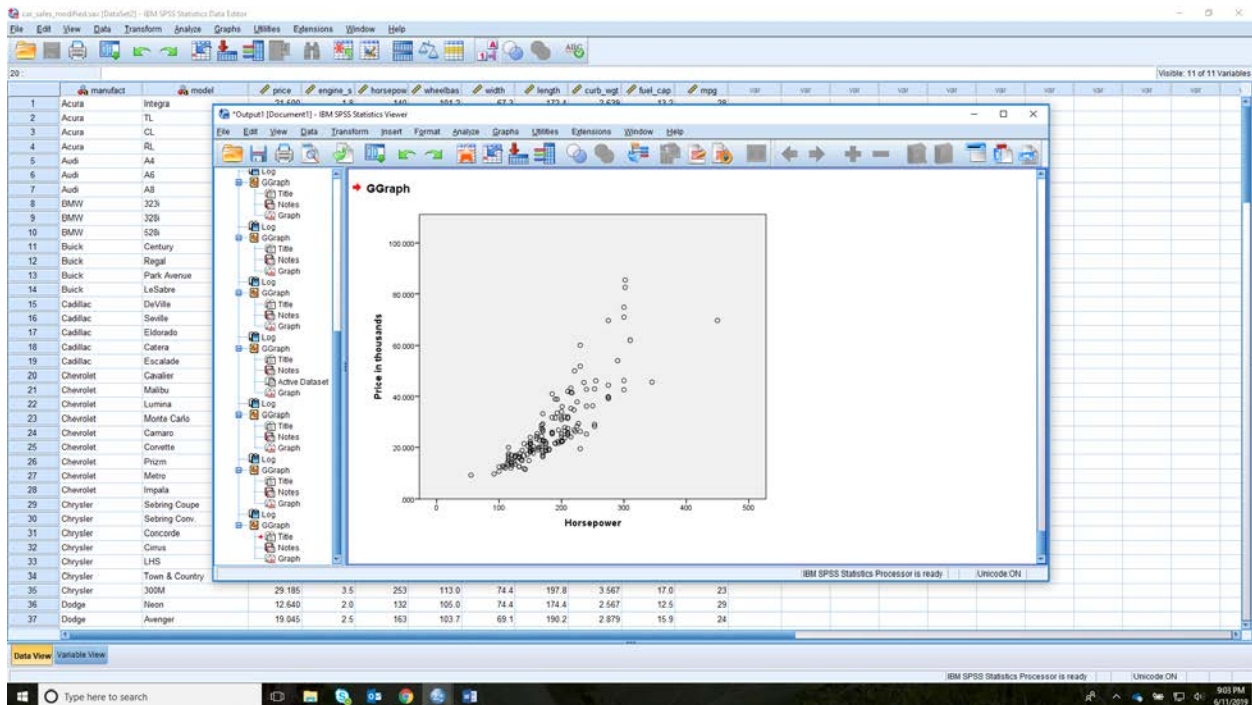
	manufact	model	price	engine_s	horsepow	wheelbas	width	length	curb_wgt	fuel_cap	mpg
1	Acura	Integra	21 500	1.8	140	101.2	67.3	172.4	2 639	13.2	28
2	Acura	TL	28 400	3.2	225	108.1	70.3	192.9	3 517	17.2	25
3	Acura	CL	42 000	3.2	225	106.9	70.6	192.0	3 410	17.2	26
4	Acura	RL	42 000	3.5	210	114.6	71.4	196.6	3 850	18.0	22
5	Audi	A4	23 990	1.8	150	102.6	68.2	178.0	2 988	16.4	27
6	Audi	A6	33 960	2.8	200	108.7	76.1	192.0	3 561	18.5	22
7	Audi	A8	62 000	4.2	310	113.0	74.0	198.2	3 902	23.7	21
8	BMW	323	26 990	2.5	170	107.3	68.4	176.0	3 179	16.6	26
9	BMW	328	33 400	2.8	193	107.3	68.5	176.0	3 197	16.6	24
10	BMW	528	38 900	2.8	193	111.4	70.9	188.0	2 472	19.5	25
11	Buick	Century	21 975	3.1	175	109.0	72.7	194.6	3 368	17.5	25
12	Buick	Regal	25 300	3.8	240	109.0	72.7	196.2	3 543	17.5	23
13	Buick	Park Avenue	31 965	3.8	205	113.8	74.7	206.8	3 778	19.5	24
14	Buick	LeSabre	27 885	3.8	205	112.2	73.6	200.0	3 591	17.5	25
15	Cadillac	DeVille	39 895	4.6	275	115.3	74.5	207.2	3 978	18.5	22
16	Cadillac	Seville	44 475	4.6	275	112.2	75.0	201.0	3 770	18.0	22
17	Cadillac	Eldorado	39 665	4.6	275	108.0	75.5	200.6	3 843	19.0	22
18	Cadillac	Catera	31 010	3.0	200	107.4	70.3	194.8	3 770	18.0	22
19	Cadillac	Escalade	46 225	5.7	265	117.5	77.0	201.2	5 672	30.0	15
20	Chevrolet	Cavalier	13 260	2.2	115	104.1	67.9	180.9	2 676	14.3	27
21	Chevrolet	Malibu	16 535	3.1	170	107.0	69.4	190.4	3 051	16.0	25
22	Chevrolet	Lumina	18 890	3.1	175	107.5	72.5	200.9	3 330	16.6	25
23	Chevrolet	Monte Carlo	19 390	3.4	180	110.5	72.7	197.9	3 340	17.0	27
24	Chevrolet	Camaro	24 340	3.8	200	101.1	74.1	193.2	3 500	16.8	25
25	Chevrolet	Corvette	45 705	5.7	345	104.5	73.6	179.7	3 210	19.1	22
26	Chevrolet	Puizam	13 960	1.8	120	97.1	66.7	174.3	2 398	13.2	33
27	Chevrolet	Metro	9 235	1.0	55	93.1	62.6	149.4	1 895	10.3	45
28	Chevrolet	Impala	18 890	3.4	180	110.5	73.0	200.0	3 389	17.0	27
29	Chrysler	Sebring Coupe	19 840	2.5	163	103.7	69.7	190.9	2 967	15.9	24
30	Chrysler	Sebring Conv	24 495	2.5	168	106.0	69.2	193.0	3 332	16.0	24
31	Chrysler	Concorde	22 245	2.7	200	113.0	74.4	209.1	3 452	17.0	26
32	Chrysler	Cirrus	16 480	2.0	132	108.0	71.0	186.0	2 911	16.0	27
33	Chrysler	LHS	28 340	3.5	253	113.0	74.4	207.7	3 564	17.0	23
34	Chrysler	Town & Country									
35	Chrysler	300M	29 185	3.5	253	113.0	74.4	197.8	3 567	17.0	23
36	Dodge	Neon	12 640	2.0	132	105.0	74.4	174.4	2 567	12.5	29
37	Dodge	Avenir	19 045	2.5	163	103.7	69.1	190.2	2 879	15.9	24

Click Graphs>Chart Builder

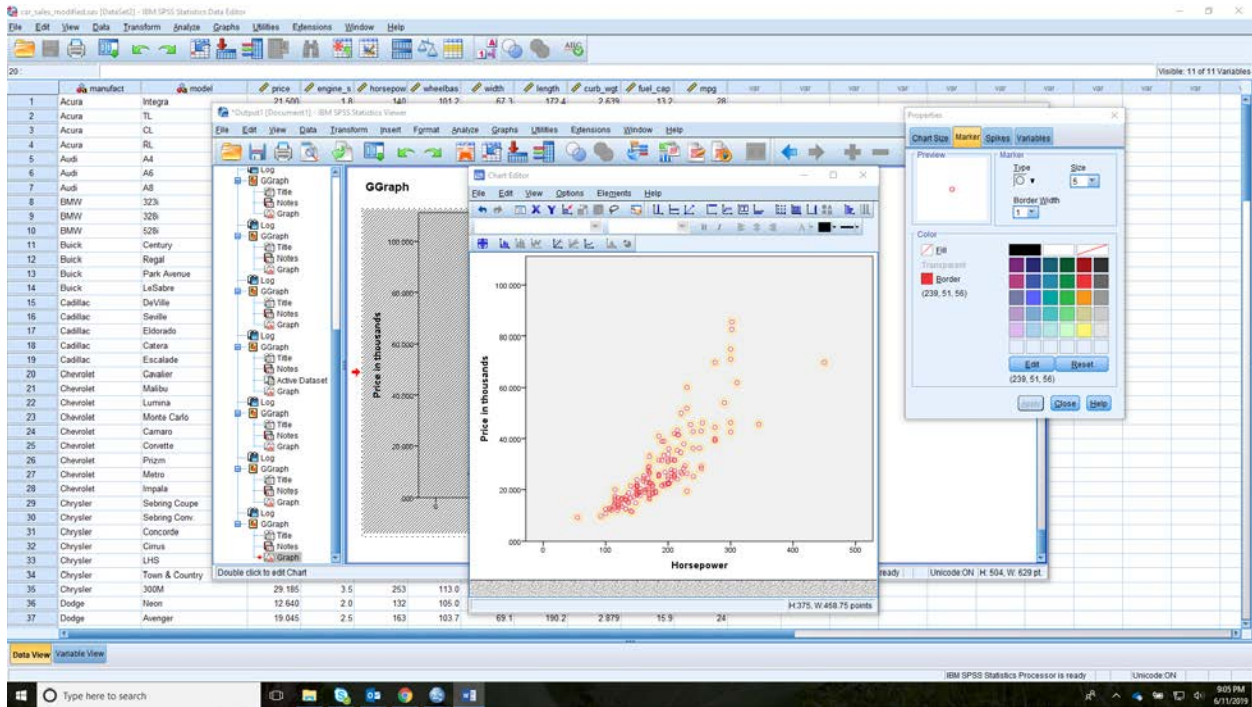


Click scatter/dot and drag the first (Simple Scatter) to the canvas. Price in thousands to the Y-Axis.

Experiment to see which would be the best predictor of price in thousands. Add horsepower.

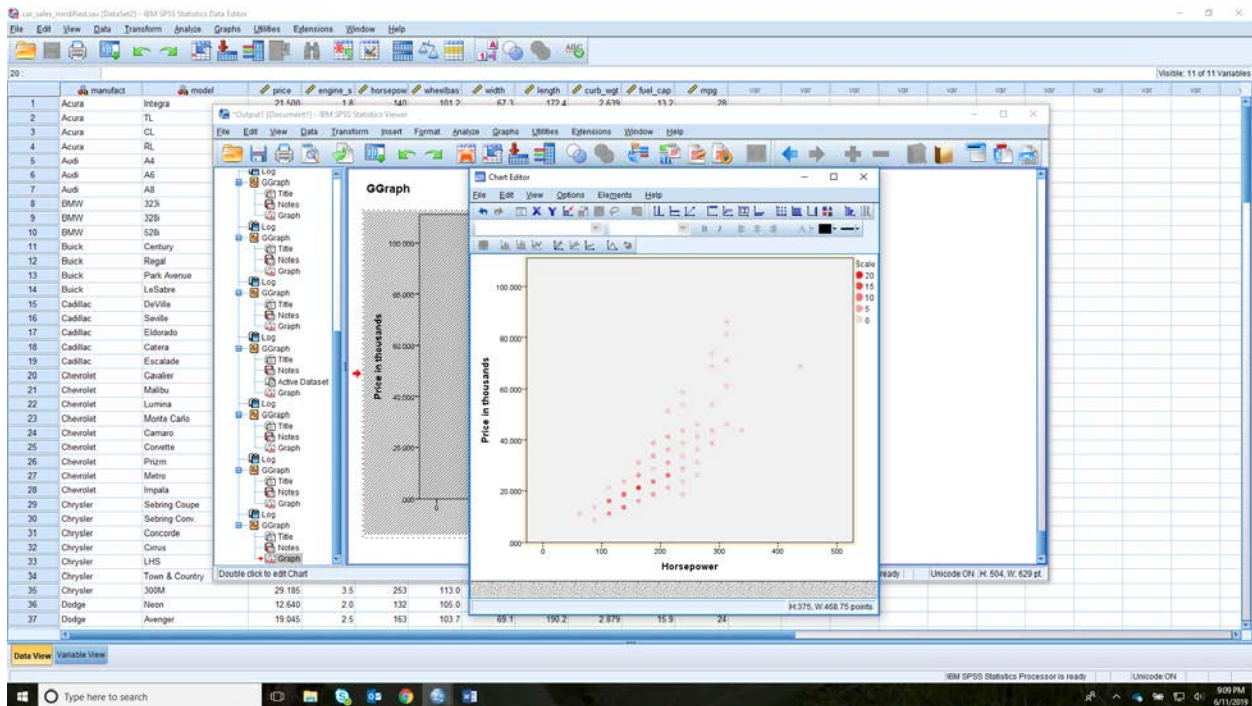


Double click to enter the “chart editor”. Double click the points to edit the properties. Red borders > apply.

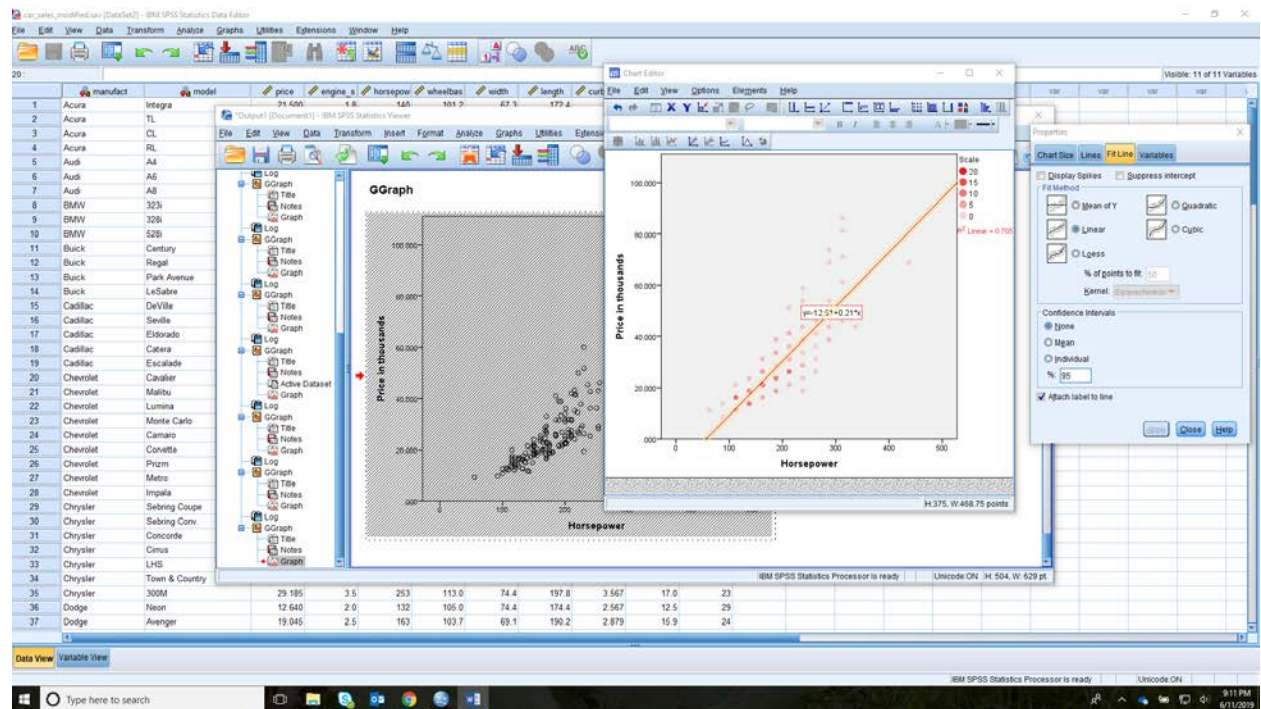


Then from the chart editor click Options>Bin Element. Click “color intensity” then “apply”.

This allows us to show another variable – the frequency of observations, shown by the intensity of colour.



In the "Chart Editor" click on "add fit line".



From the output window, right click on the graph and export ONLY THE GRAPH as a PDF file.