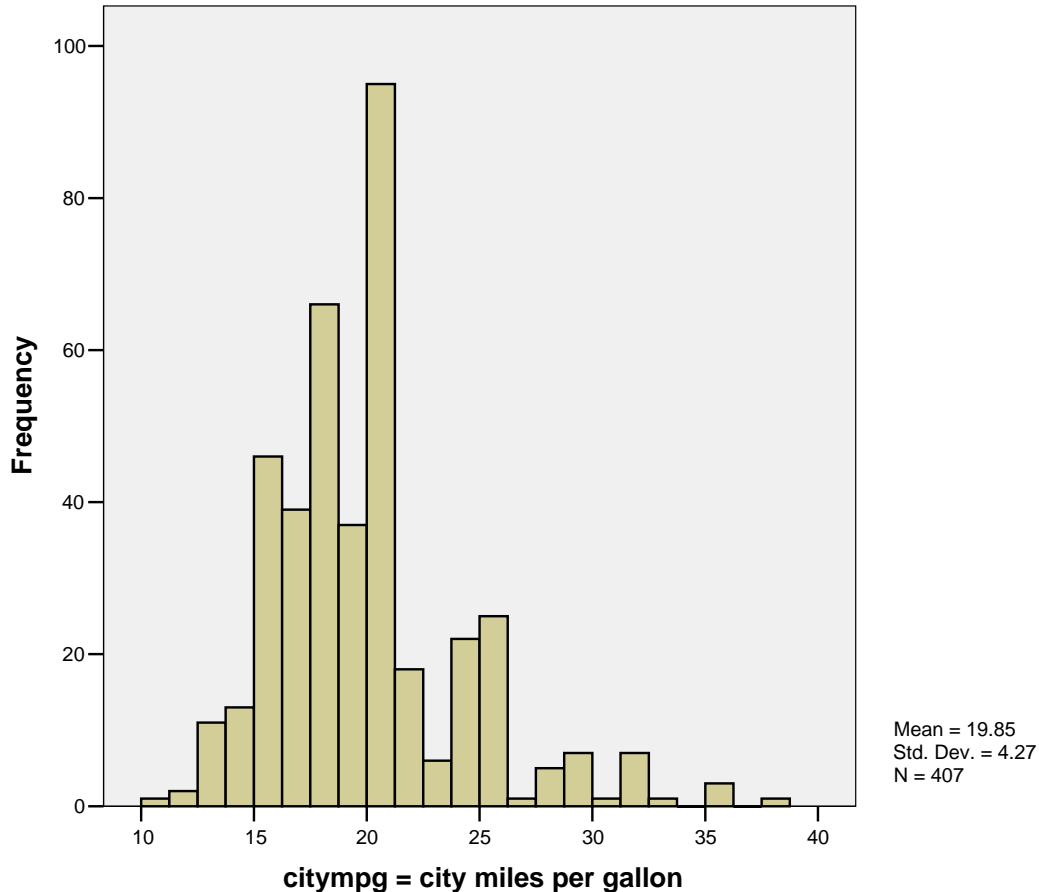


DSC 330 – Homework Example

Even-numbered problems from Appendix A

A.2 (a) Histogram for car city miles per gallon:



(b) The histogram shows city MPG figures ranging from the low 10s to the high 30s, with more values tending to cluster around 18-20 MPG rather than at the extremes. The distribution seems slightly positively skewed, with a few high values quite a lot higher than the rest.

(c) Mean = 19.85, median = 19.

(d) For skewed data, the median is a better summary of the central tendency of the data than the mean. For example, when data are skewed to the right, a few very large values cause the mean to be relatively large in comparison to the median (which is not affected to the same extent by these large values). In this case, the data are moderately skewed so the median might be marginally preferred to the mean. However, the mean and median are quite close together suggesting that the amount of skew is fairly minor.

(e) 25th percentile (first quartile) = 17, 50th percentile (second quartile, or median) = 19, 75th percentile (third quartile) = 21.

(f) The middle 50% of the salaries falls between 17 and 21.

A.4 (a) Cross-tabulation of type and drive:

type * drive Crosstabulation

			drive			Total
			All-wheel drive	Front-wheel drive	Rear-wheel drive	
type	Minivan	Count	3	16	1	20
		% within type	15.0%	80.0%	5.0%	100.0%
		% within drive	3.4%	7.5%	.9%	4.9%
	Pick-up truck	Count	11	0	12	23
		% within type	47.8%	.0%	52.2%	100.0%
		% within drive	12.4%	.0%	11.3%	5.7%
	Sedan	Count	24	153	54	231
		% within type	10.4%	66.2%	23.4%	100.0%
		% within drive	27.0%	72.2%	50.9%	56.8%
	Sports car	Count	5	8	32	45
		% within type	11.1%	17.8%	71.1%	100.0%
		% within drive	5.6%	3.8%	30.2%	11.1%
	Sport utility vehicle	Count	37	22	0	59
		% within type	62.7%	37.3%	.0%	100.0%
		% within drive	41.6%	10.4%	.0%	14.5%
	Wagon	Count	9	13	7	29
		% within type	31.0%	44.8%	24.1%	100.0%
		% within drive	10.1%	6.1%	6.6%	7.1%
Total		Count	89	212	106	407
		% within type	21.9%	52.1%	26.0%	100.0%
		% within drive	100.0%	100.0%	100.0%	100.0%

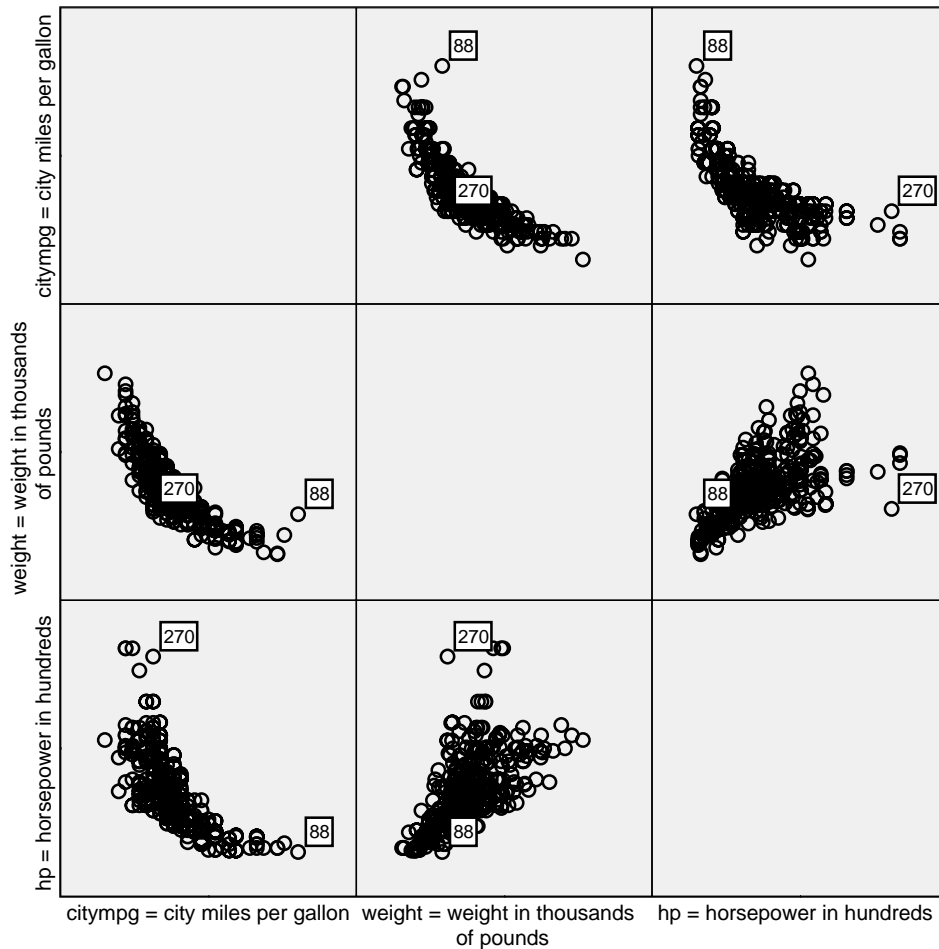
(b) There are 153 front-wheel drive sedans.

(c) 66.2% of sedans are front-wheel drive.

(d) 72.2% of front-wheel drive vehicles are sedans.

(e) By far the most common combination of type and drive is front-wheel drive sedans, with over 150 vehicles (nearly 38% of the dataset). There are also more than 50 rear-wheel drive sedans, and more than 20 all-wheel drive sedans. There are nearly 60 sport utility vehicles, more than 60% of which have all-wheel drive, with the remainder having front-wheel drive. There are 45 sports cars, more than 70% of which are rear-wheel drive. There are nearly 30 wagons, with all types of drive fairly evenly represented. There are also 23 pick-up trucks, evenly split between all-wheel drive and rear-wheel drive. Finally, there are 20 minivans, mostly front-wheel drive.

A.6 (a) Scatterplot matrix for car city miles per gallon, weight, and horsepower:



(b) There are fairly strong negative relationships between citympg and each of weight and hp, but the relationships are curved---steeper for low values of weight and hp and becoming shallower as weight and hp increase. There is a positive, fairly linear relationship between weight and hp, with the relationship becoming more spread out as the values of weight and hp increase. A few vehicles “stick out” from the dominant patterns in the plots. For example, vehicle 88 (a Volkswagen Jetta) has the highest value of city MPG, but it doesn't have a very low weight. Also, vehicle 270 (a Porsche 911) has a very high horsepower relative to its modest weight.