Dulaney, Stewart
MATH 54
Section 4053
SID: 1545566

10/10

HW 2

4.1

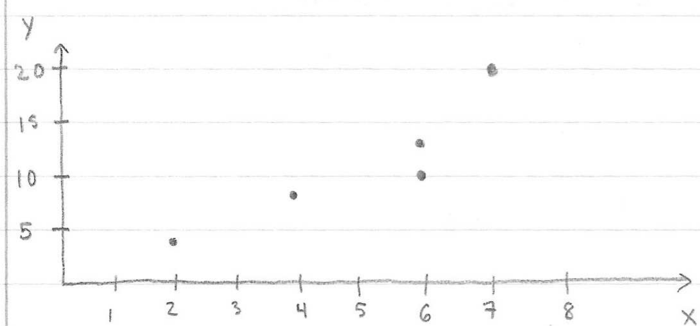
1. Univariate data is data in which a single variable is measured. Bivariate data is data in which two variables are measured on an individual.

5. -1

9. no linear relation

13. a) III b) IV
c) II d) I

17.
a)



b) $\bar{x} = 5$ $\bar{y} = 11$
$s_x = 2$ $s_y = 6$

x	y	$\frac{x_i - \bar{x}}{s_x}$	$\frac{y_i - \bar{y}}{s_y}$	$\left(\frac{x_i - \bar{x}}{s_x}\right)\left(\frac{y_i - \bar{y}}{s_y}\right)$
2	4	-1.5	-1.1667	1.7501
4	8	-0.5	-0.5	0.25
6	10	0.5	-0.1667	-0.0834
6	13	0.5	0.3333	0.1667
7	20	1	1.5	1.5

$$\Sigma \left(\frac{x_i - \bar{x}}{s_x}\right)\left(\frac{y_i - \bar{y}}{s_y}\right) = 3.5834$$

$$r = \frac{3.5834}{5-1} = 0.896$$

c) linear relation b/c $|r| = |0.896| = 0.896 > 0.878$

21. a) Positive because children under age 3 likely need diapers.

b) Negative because consumers are less likely to buy as interest rates increase.

c) Negative because exercise lowers cholesterol because you lose weight.

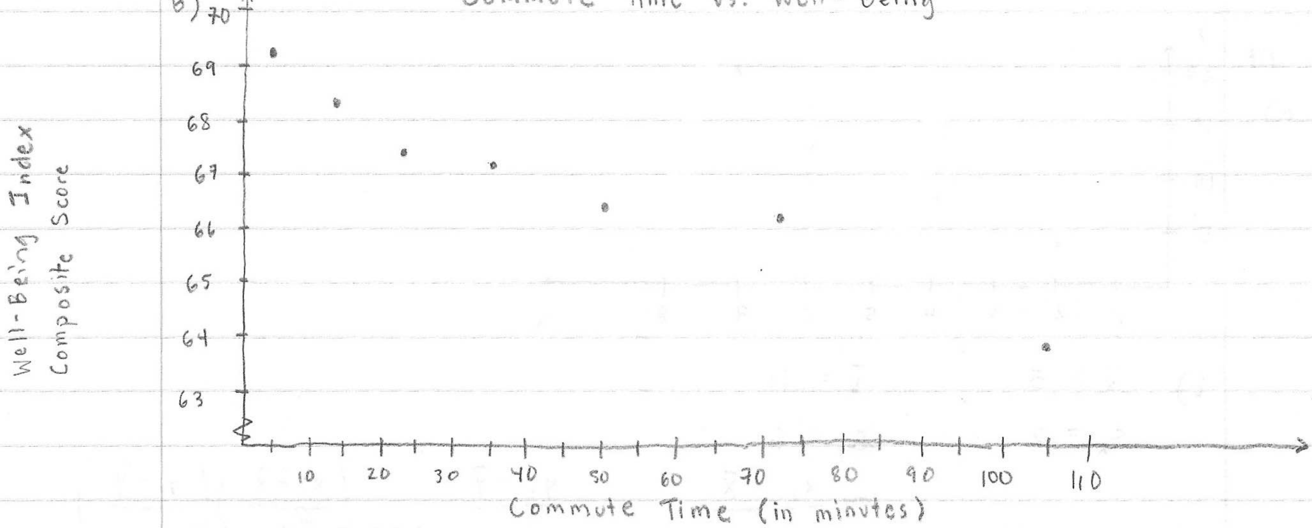
d) Negative because less customers will eat there as the price of a Big Mac increases.

e) No correlation

25. a) explanatory: commute time

response: well-being score

b) Commute Time vs. Well-Being



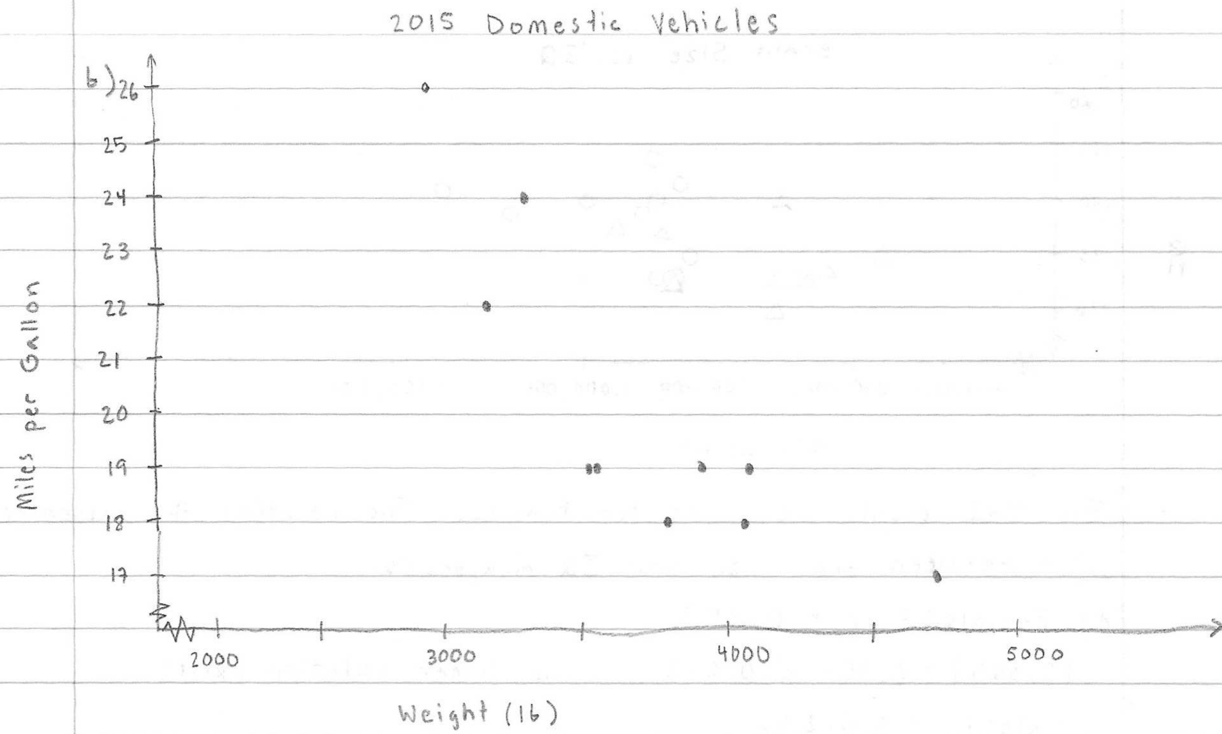
c) $r = -0.981$

d) Yes b/c $|-0.981| = 0.981 > 0.754$. There is a negative linear relation.

29. a) explanatory: weight of car

response: miles per gallon

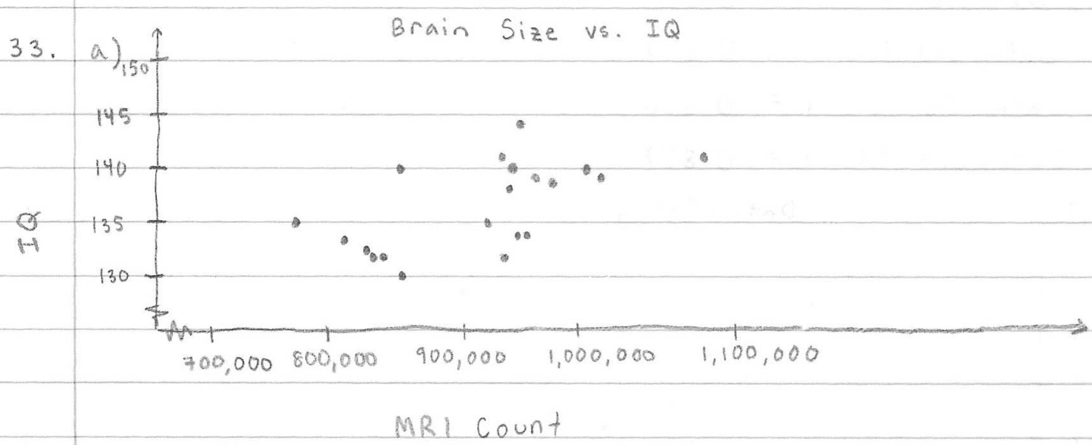
b)



c) $r = -0.842$

d) Yes b/c $|-0.842| = 0.842 > 0.632$. There is a negative linear relation.

33. a)

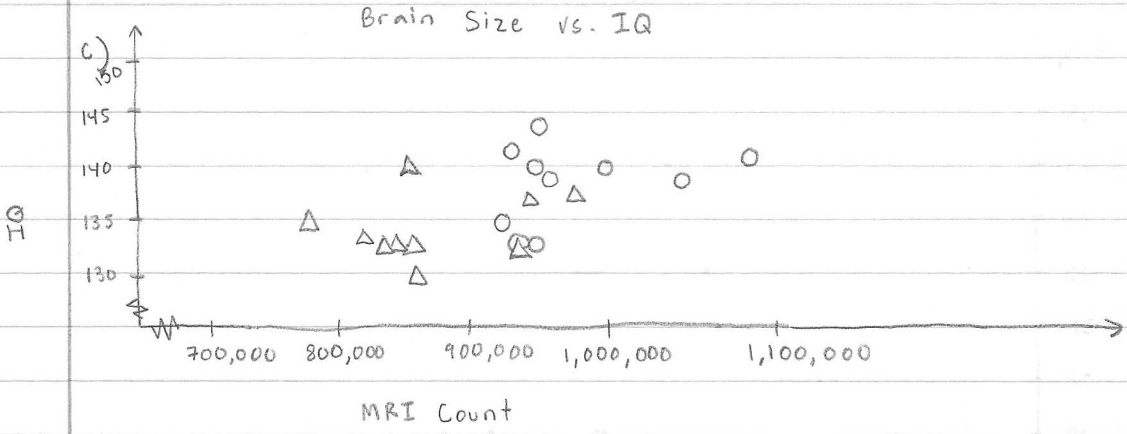


There may be a loose positive linear relation.

b) $r = 0.548$

Yes b/c $|0.548| = 0.548 > 0.444$. There is a positive association.

c)



The MRI counts are lower for females. The relation that appeared to exist between brain size and IQ disappears.

d) Females: $r = 0.359$

$|0.359| = 0.359 < 0.632$, $\therefore$ no linear relation exists

Males: $r = 0.236$

$|0.236| = 0.236 < 0.632$, $\therefore$ no linear relation exists

The moral is to be vigilant about lurking variables.

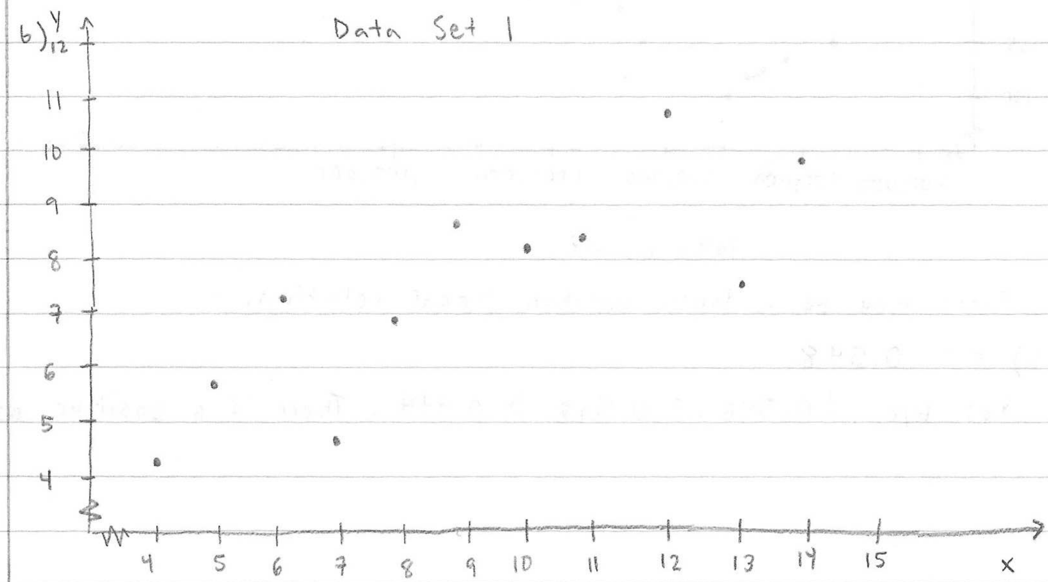
37. a) Data Set 1: $r = 0.816$

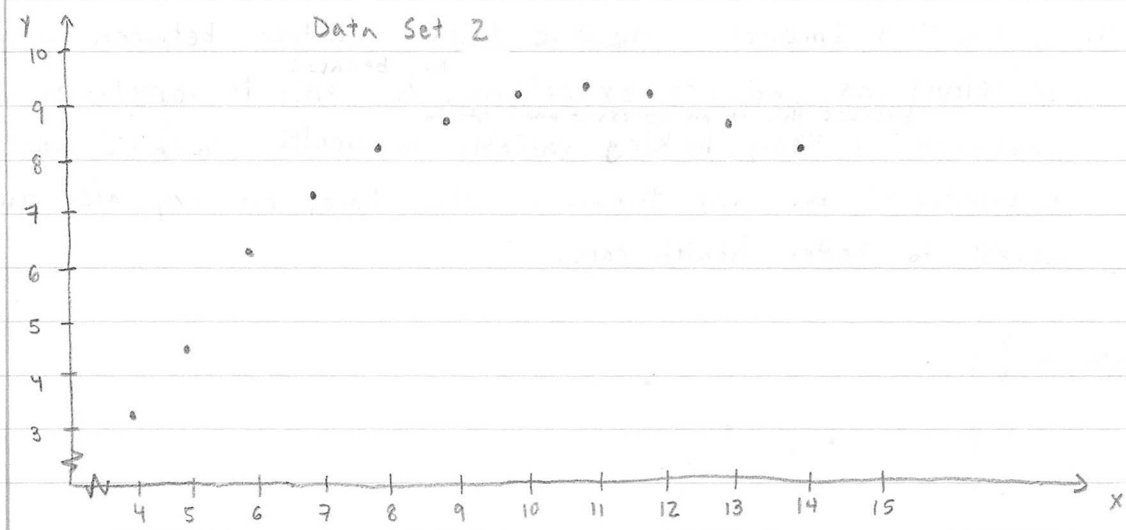
Data Set 2: $r = 0.817$

Data Set 3: $r = 0.816$

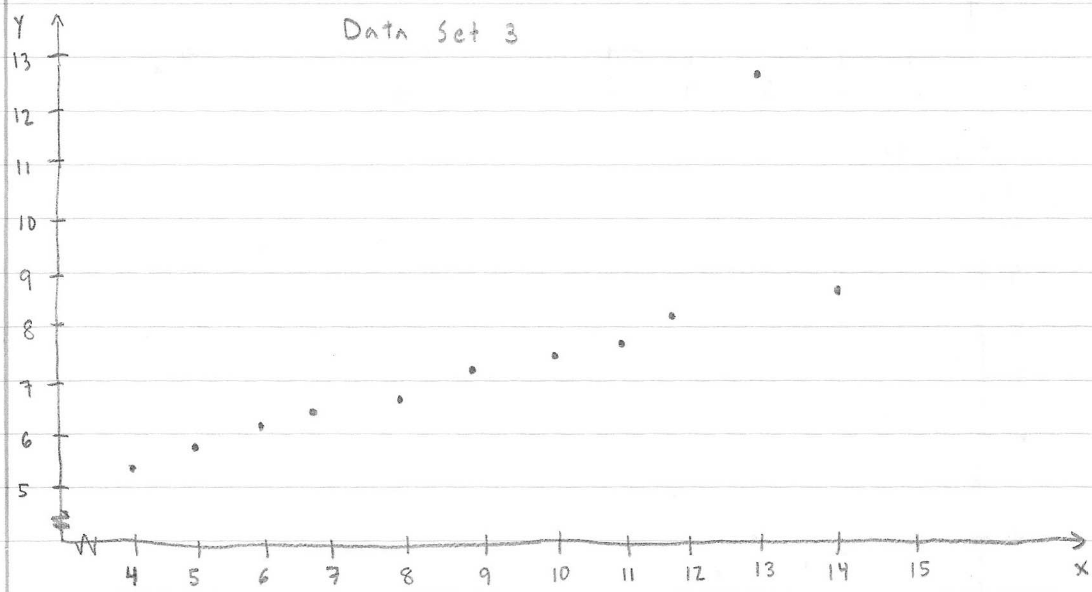
Data Set 4: $r = 0.817$

b)

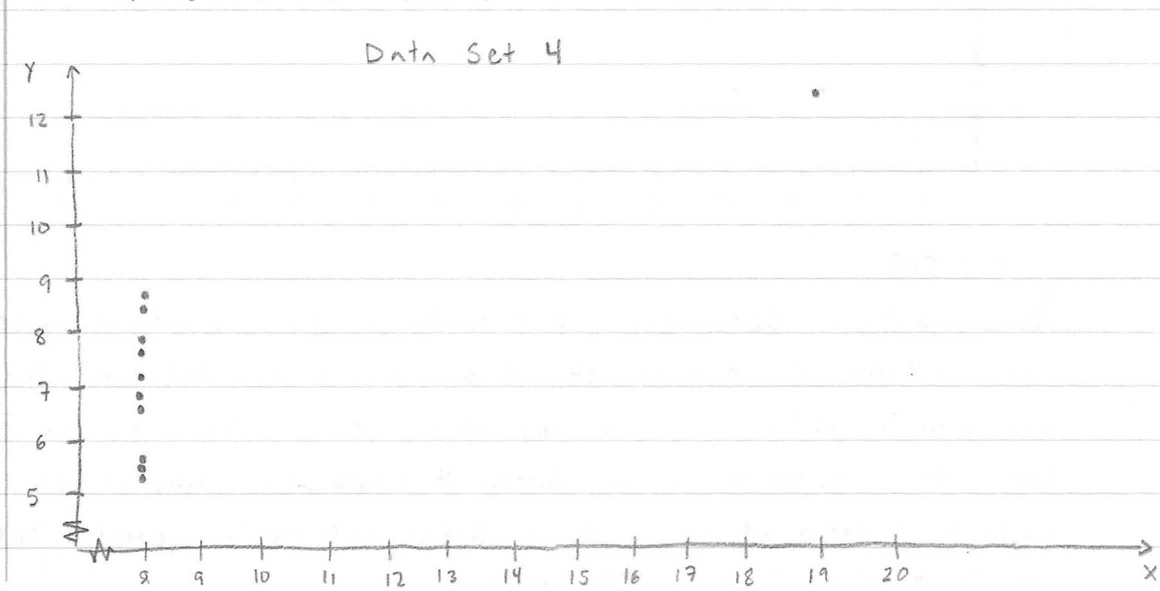




Data Set 2
Y
10
9
8
7
6
5
4
3
4
5
6
7
8
9
10
11
12
13
14
15
X



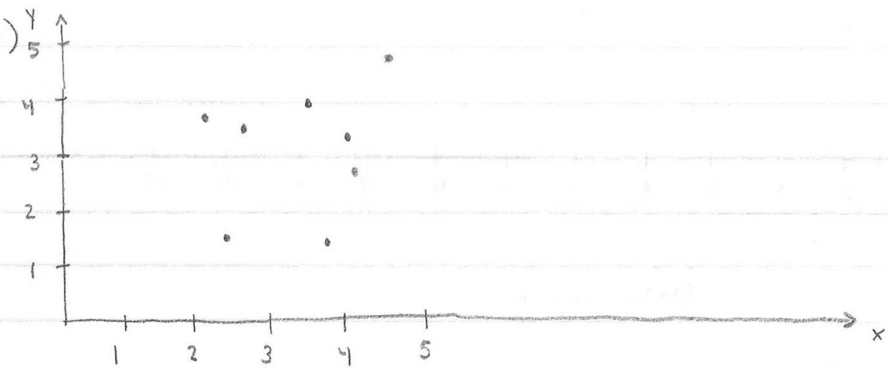
Data Set 3
Y
13
12
11
10
9
8
7
6
5
4
5
6
7
8
9
10
11
12
13
14
15
X



Data Set 4
Y
12
11
10
9
8
7
6
5
8
9
10
11
12
13
14
15
16
17
18
19
20
X

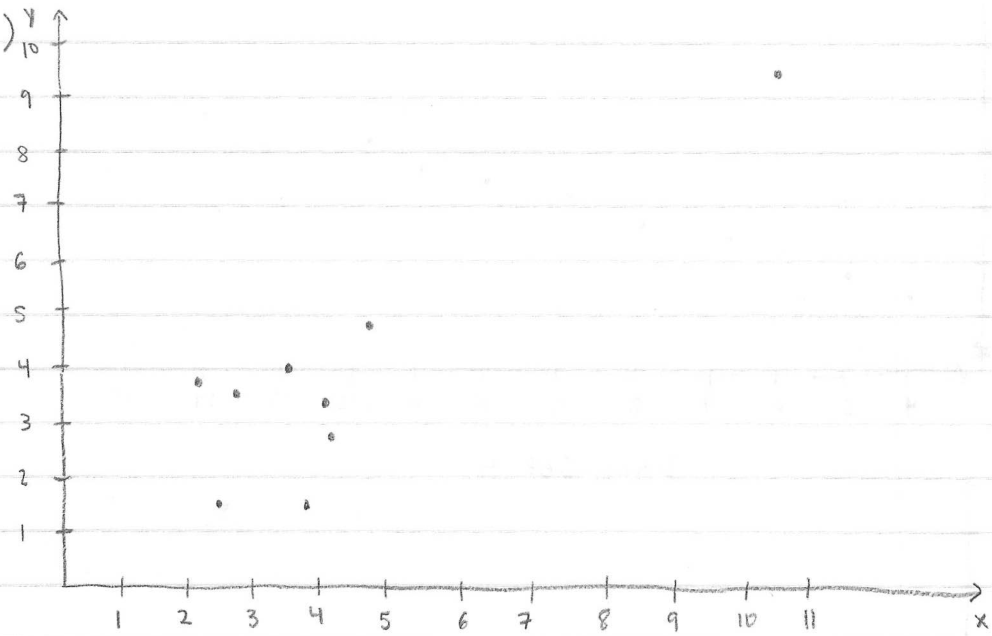
41. $r = 0.599$ implies a positive linear relation between number of televisions and life expectancy. No, because this is correlation not causation because this is an observational study. A likely lurking variable is wealth, because the wealthier a country is, the more televisions they have, but they also have access to better health care.

45. a)



$r = 0.228$

b)



$r = 0.860$

The additional data point has a large effect on r because it is not resistant. Correlations should always be reported with scatter diagrams in order to spot any observations that do not follow the overall pattern of the data. A single point may affect r, making it seem as if a strong linear relationship exists between two variables, when it does not.

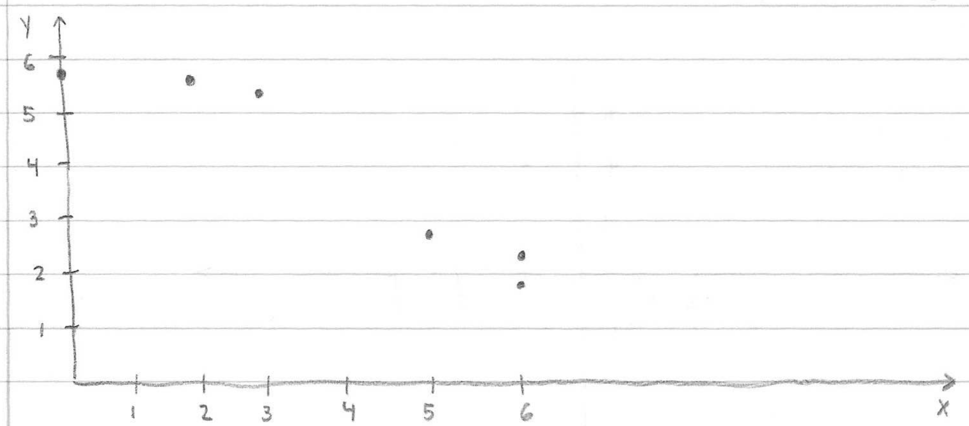
49. A perfect positive linear relation exists between the two variables and the scatter diagram looks like a straight line.

53. Correlation describes a relation between two variables in an observational study. Causation describes a conditional (if - then) relation in an experimental study. It is only appropriate to state that the correlation implies causation if the data examined are the result of an experiment.

4.2 1. residual

5.

a)



negative association

b) $\bar{x} = 3.6667$ $\quad \bar{y} = 3.9333$ $\quad r = {}^{-}0.9477$

$s_x = 2.4221$ $\quad s_y = 1.8239$

$b_1 = r \cdot \dfrac{s_y}{s_x}$ $\quad b_0 = \bar{y} - b_1 \bar{x}$

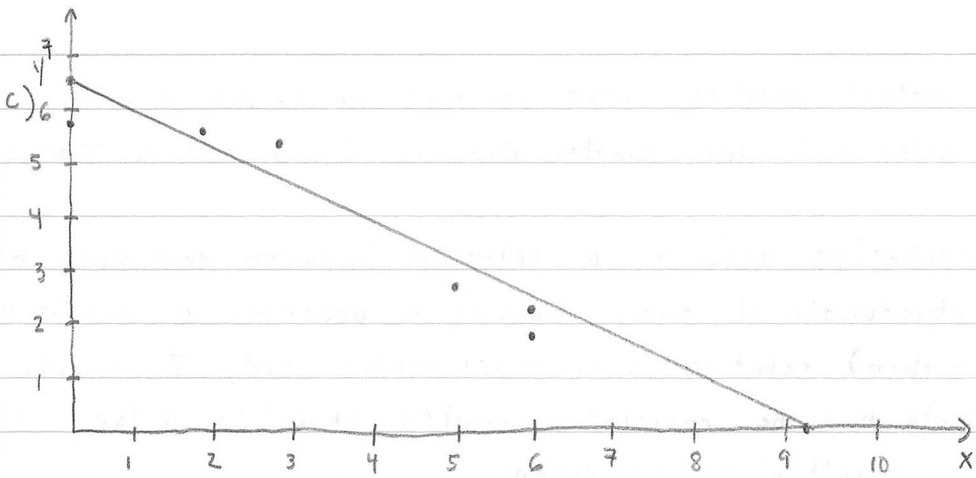
$b_0 = 3.9333 - ({}^{-}0.7136)(3.6667)$

$b_1 = ({}^{-}0.9477) \cdot \dfrac{(1.8239)}{(2.4221)}$ $\quad b_0 = 6.5499$

$b_1 = {}^{-}0.7136$

$\hat{y} = {}^{-}0.7136x + 6.5499$

c)

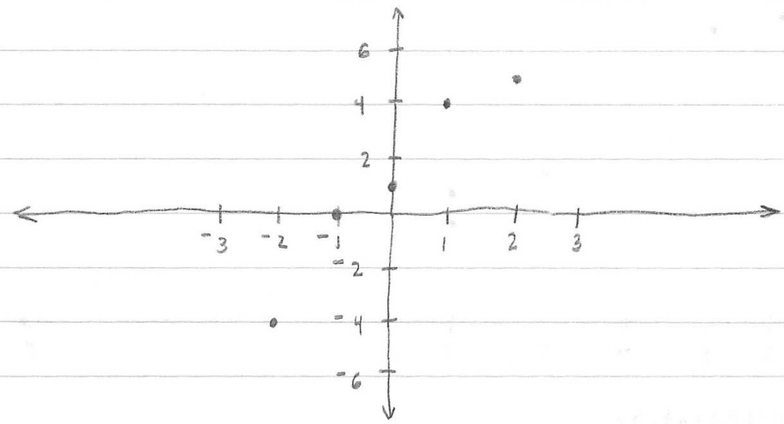


$0 = {}^{-}0.7136x + 6.5499$

${}^{-}6.5499 = {}^{-}0.7136x$

$x = 9.1787$

9. a)



b) $({}^{-}2, {}^{-}4)$, $(2, 5)$

$m = \frac{\text{rise}}{\text{run}} = \frac{5-({}^{-}4)}{2-({}^{-}2)} = \frac{5+4}{2+2} = \frac{9}{4}$

$y = \frac{9}{4}x + b$

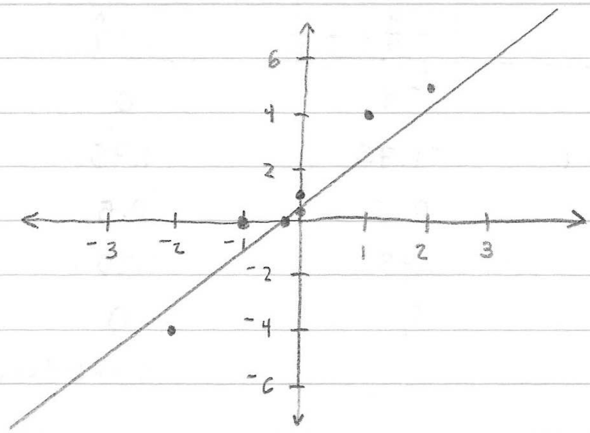
$-4 = \frac{9}{4}(-2) + b$

$-4 = \frac{{}^{-}18}{4} + b$

$b = \frac{{}^{-}16}{4} + \frac{18}{4} = \frac{2}{4} = \frac{1}{2}$

$y = \frac{9}{4}x + \frac{1}{2}$

c)



$$0 = \frac{9}{4}x + \frac{1}{2}$$

$$-\frac{1}{2} = \frac{9}{4}x$$

$$x = -\frac{1}{2}\left(\frac{4}{9}\right) = \frac{-4}{18} = \frac{-2}{9}$$

d) $\bar{x} = 0$ $\quad\bar{y} = 1.2$ $\quad r = 0.9761$

$s_x = 1.5811$ $\quad s_y = 3.5637$

$$b_1 = (0.9761) \cdot \frac{(3.5637)}{(1.5811)}$$

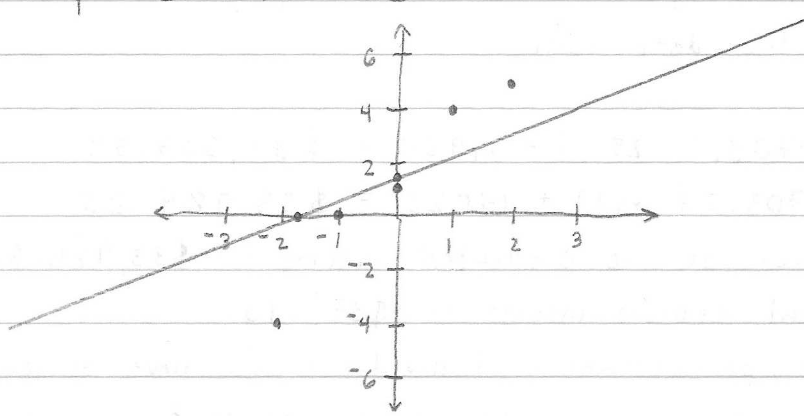
$$b_0 = 1.2 - (2.2001)(0)$$

$$b_0 = 1.2$$

$$b_1 = 2.2001$$

$$\hat{y} = 2.2001x + 1.2$$

e)



$$0 = 2.2001x + 1.2$$

$$-1.2 = 2.2001x$$

$$x = -1.8334$$

f)

x	y	$\hat{y} = \frac{9}{4}x + \frac{1}{2}$	Residual $y - \hat{y}$	Residual^2 $(y-\hat{y})^2$
-2	-4	-4	0	0
-1	0	-1.75	1.75	3.0625
0	1	0.5	0.5	0.25
1	4	2.75	1.25	1.5625
2	5	5	0	0

$\Sigma \text{Residual}^2 = 4.875$

g)

x	y	$\hat{y} = 2.2001x + 1.2$	$y - \hat{y}$	$(y-\hat{y})^2$
-2	-4	-3.2002	-0.7998	0.6397
-1	0	-1.0001	1.0001	1.0002
0	1	1.2	-0.2	0.04
1	4	3.4001	0.5999	0.3599
2	5	5.6002	-0.6002	0.3602

$\Sigma \text{Residual}^2 = 2.4$

h) The sum of the squared residuals is smaller for the least-squares regression line found in part (d) and therefore it fits the data better.

13. a) $\hat{y} = 703.5(25) + 14920 = \$32,507.50$

b) $\hat{y} = 703.5(27.1) + 14920 = \$33,984.85$

Higher because the predicted income is \$33,984.85 while the actual median income is \$37,193.

c) If the percentage of individuals who have at least a bachelor's degree increases by one percent, the median income increase by \$703.50, on average.

d) This would be outside the scope of the model because there are no observations near 0%.

17. a) $\hat{y} = -0.0479x + 69.0296$

b) If commute time increases by one minute, the well-being score decreases by 0.0479, on average. The average well-being score for a person with a 0-minute commute is 69.0296. Someone who works from home will have a 0-minute commute, so the y-intercept has a meaningful interpretation.

c) $\hat{y} = -0.0479(30) + 69.0296 = 67.5926$

d) $\hat{y} = -0.0479(20) + 69.0296 = 68.0716$

No, she is less well-off because her score is less than the mean score.

21. a) $\hat{y} = -0.0047x + 37.3569$

b) If weight of a car decreases by one pound, the miles per gallon decreases by 0.0047, on average. It is not appropriate to interpret the y-intercept because a 0-pound car does not exist.

c) $\hat{y} = -0.0047(3649) + 37.3569 = 20.2066$

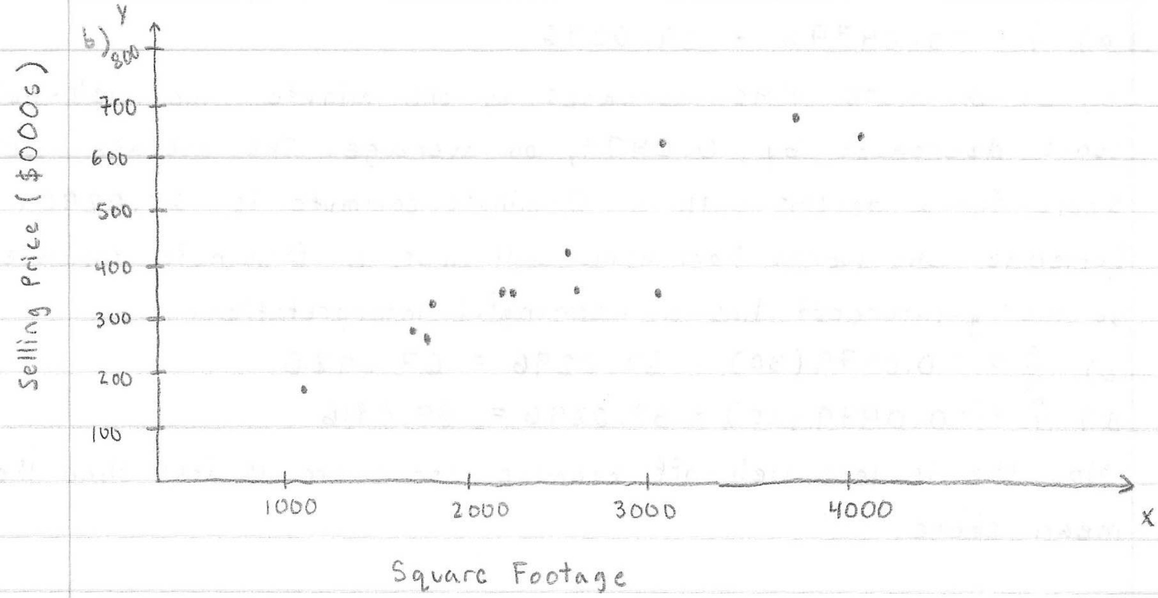
Below average (slightly)

d) No because it is a different type of car (hybrid).

25. In both cases, $\hat{y} = \bar{y} = 49.7\%$ because in 4.1 Problem 31 stock performance and CEO compensation were found not to have a linear relation.

29. a) square footage

b)



c) 0.9026

d) $|0.9026| > 0.576$

Yes, positive association

e) $\hat{y} = 0.1563x + 14.8740$

f) If square footage increases by one square foot, selling price increases by 0.1563 ($000s) = $156.30, on average.

g) No because a house that is 0-square feet does not exist.

h) $\hat{y} = 0.1563(1465) + 14.8740 = 243.8535$

Above average, some factors that could affect price are location and year built.

33. Legendre explains that the concept is straightforward-- by minimizing the sum of the squares of the errors, the result of the least-squares criterion is the best approximation of the truth, or a line that best fits the data.

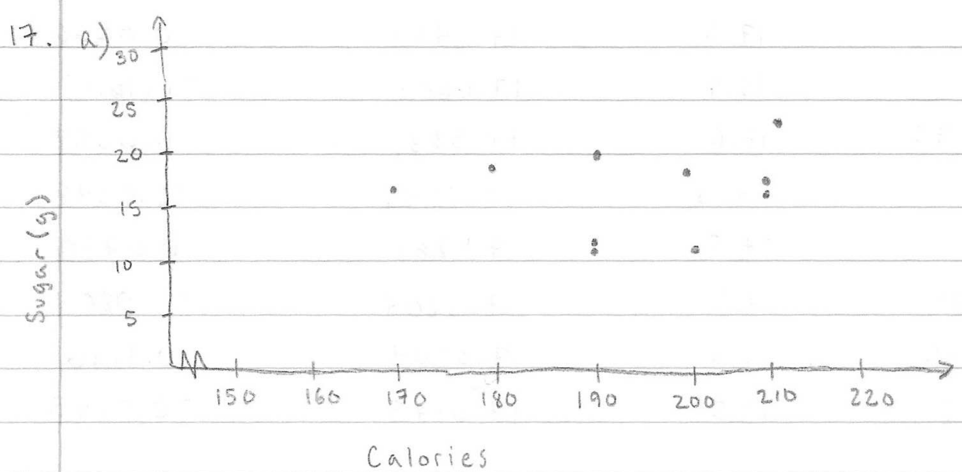
4.3

1. coefficient of determination

5. The variance of the residuals is not constant.

9. a) III b) II
c) IV d) I

13. No influential b/c the slope and y-intercept change only slightly.

17. a)


Positive association

b) r = 0.2610
|0.2610| < 0.576
No linear relation

c)


r = 0.6492
|0.6492| > 0.576 → Yes linear relation → Yes influential observation

d) Influential observations can cause the correlation coefficient to increase significantly, increasing the apparent strength of the linear relation between two variables.

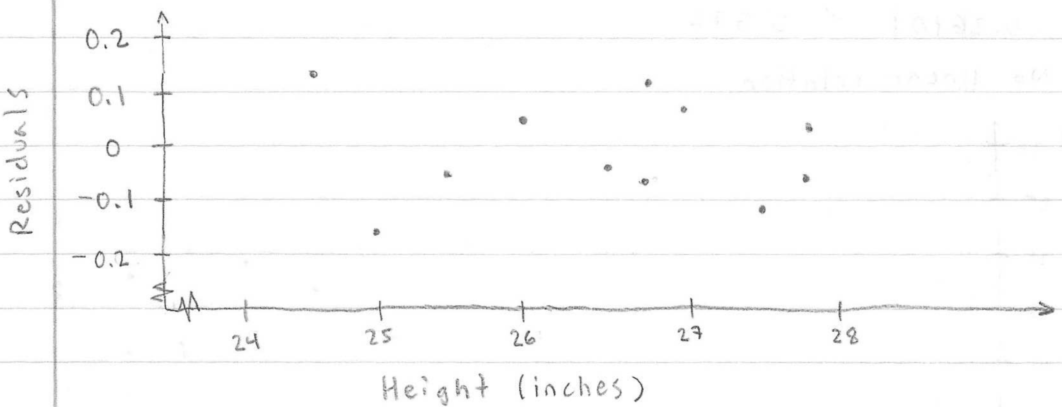
21. a) $R^2 = 0.83$

b)

Height	Head Circumference	$\hat{y}$	Residual $(y - \hat{y})$
27.75	17.5	17.5631	-0.0631
24.5	17.1	16.9694	0.1306
25.5	17.1	17.1521	-0.0521
26	17.3	17.2434	0.0566
25	16.9	17.0607	-0.1607
27.75	17.6	17.5631	0.0369
26.5	17.3	17.3348	-0.0348
27	17.5	17.4261	0.0739
26.75	17.3	17.3804	-0.0804
26.75	17.5	17.3804	0.1196
27.5	17.5	17.5175	-0.0175

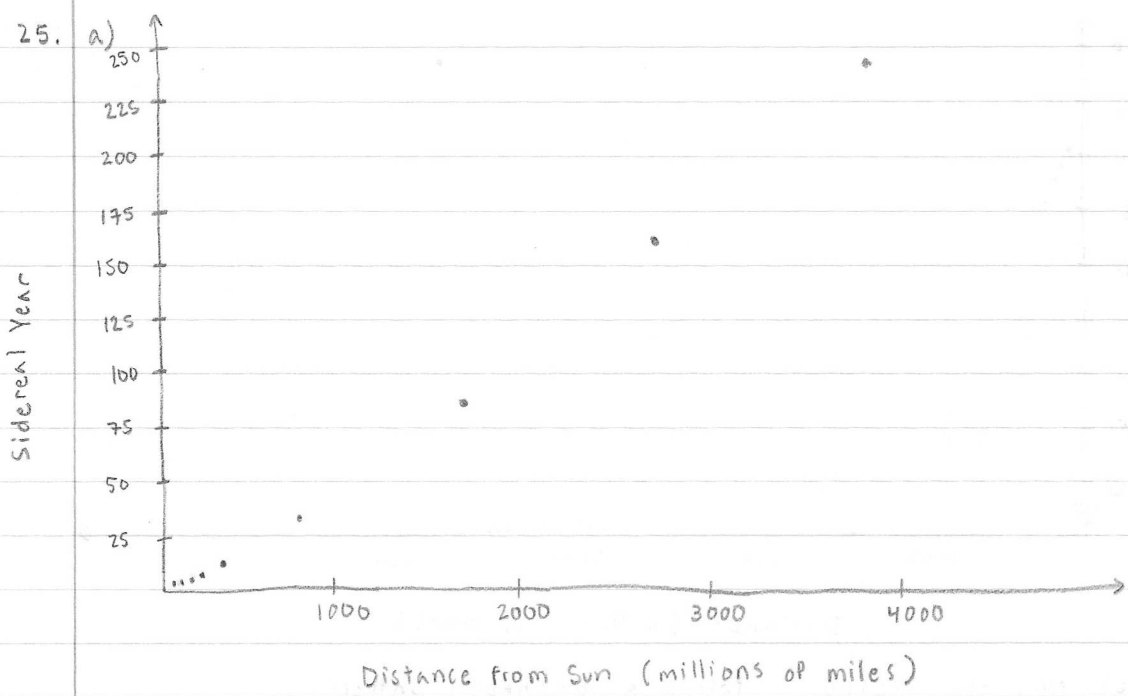
$\hat{y} = 0.1827x + 12.4932$

Residuals vs. Height



c) 83% of the variation in head circumference is explained by the least-squares regression equation. The linear model is appropriate based on the residual plot.

25. a)



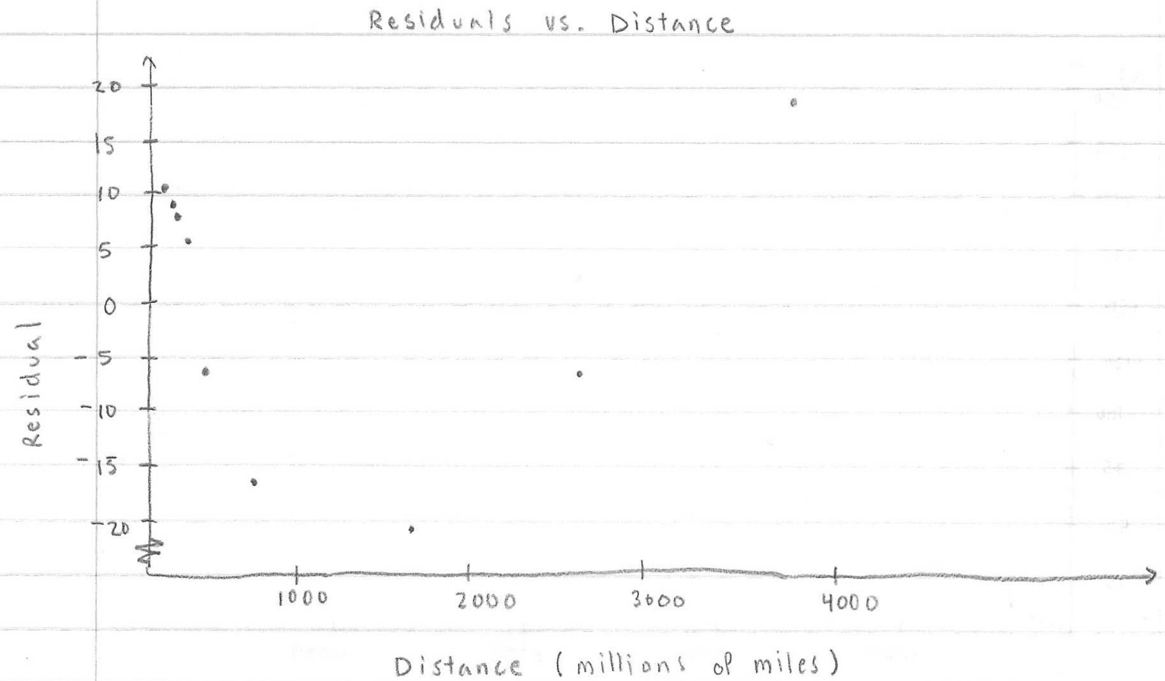
b) $r = 0.9889$

$|0.9889| > 0.666 \rightarrow$ Yes linear relation

c) $\hat{y} = 0.0657x - 12.4967$

d)

Distance	Sidereal Year	$\hat{y}$	$(y - \hat{y})$
36	0.24	$^-10.1315$	10.3715
67	0.62	$^-8.0948$	8.7148
93	1.00	$^-6.3866$	7.3866
142	1.88	$^-3.1673$	5.0473
483	11.9	19.2364	$^-7.3364$
887	29.5	45.7792	$^-16.2792$
1785	84.0	104.7778	$^-20.7778$
2797	165.0	171.2662	$^-6.2662$
3675	248.0	228.9508	19.0492



e) No, the residuals follow a V-shaped pattern.

29. $\hat{y} = 0.9245x - 64.0746$ (w/ Boardman)

$\hat{y} = 0.9245(4618) - 64.0746 = 4205.2664$

$4813 > 4205.2664$

$\hat{y} = 0.5029x + 22.9377$ (w/o Boardman)

Energy, x	Carbon, Y	$\hat{y}$	$(y - \hat{y})$
4618	4813	4205.2664	607.7336
2636	1311	2372.9074	-1061.9074
535	377	430.5329	-53.5329
357	341	265.9719	75.0281
370	238	277.9904	-39.9904
15	44	-50.2071	94.2071
25	18	-40.9621	58.9621
22	15	-43.7356	58.7356
13	10	-52.0561	62.0561
64	6	-4.9066	10.9066
13	5	-52.0561	57.0561
33	3	-33.5661	36.5661
24	2	-41.8866	43.8866
17	2	-48.3581	50.3581

Residuals:

$IQR = Q_3 - Q_1 = 62.0561 - 10.9066 = 51.1495$

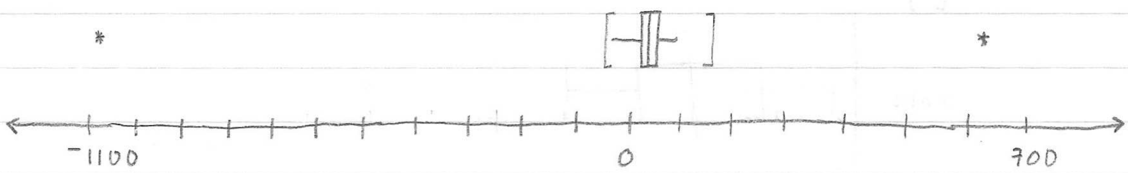
Lower fence $= Q_1 - 1.5(IQR) = 10.9066 - 1.5(51.1495) = {}^-65.8177$

Upper fence $= Q_3 + 1.5(IQR) = 62.0561 + 1.5(51.1495) = 138.7801$

$Q_1 = 10.9066$

$Q_3 = 62.0561$

$M = 53.7071$



The Boardman plant is unusual b/c it emits more carbon than expected from the overall trend. The observation is an outlier and influential.

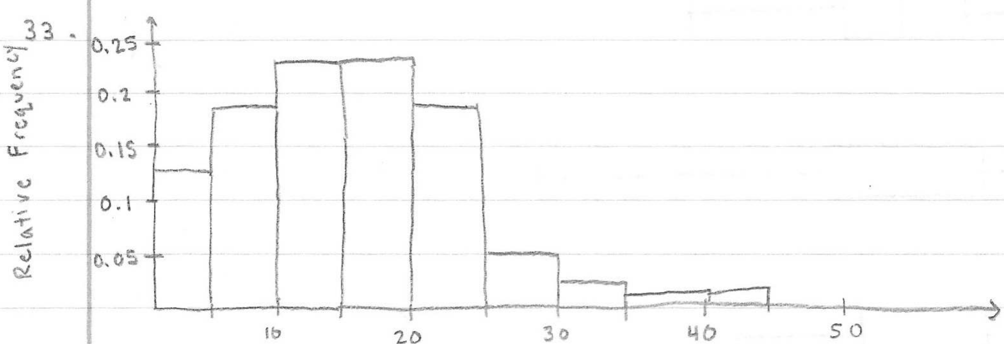
$\hat{y} = 0.5029(2636) + 22.9377 = 1348.5821$

$y - \hat{y} = 1311 - 1348.5821 = {}^-37.5821$

The Hermiston plant is not unusual as it follows the overall trend of the data. Although it has a large residual initially, this effect goes away if the Boardman observation is removed from the data.

What Percent of Income Should Be Paid in Federal Income Tax?

33.



Tax Rate (%)

The graph is skewed right. Two classes have a relative frequency of 0.23: 10 - 14.9, 15 - 19.9.

b) $\bar{x} = 12.3\%$

$M = 10\%$

c) $s = 7.6\%$

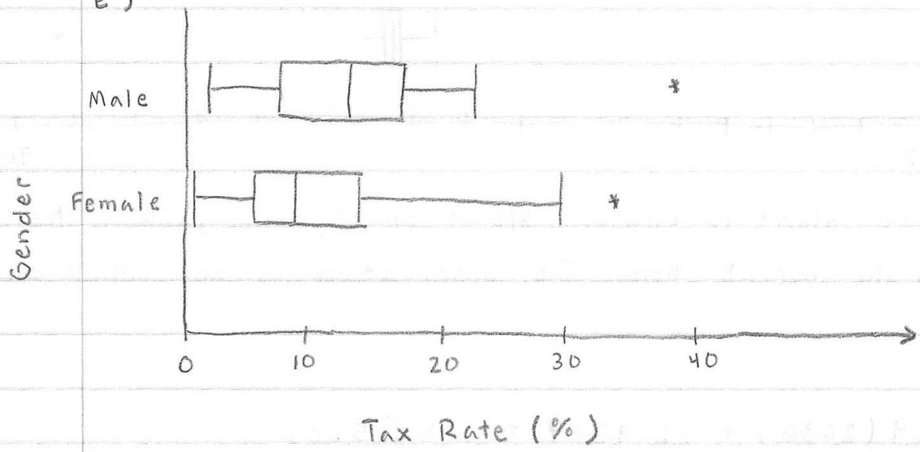
$IQR = 12\%$

d) $LF = -13\%$

$UF = 35\%$

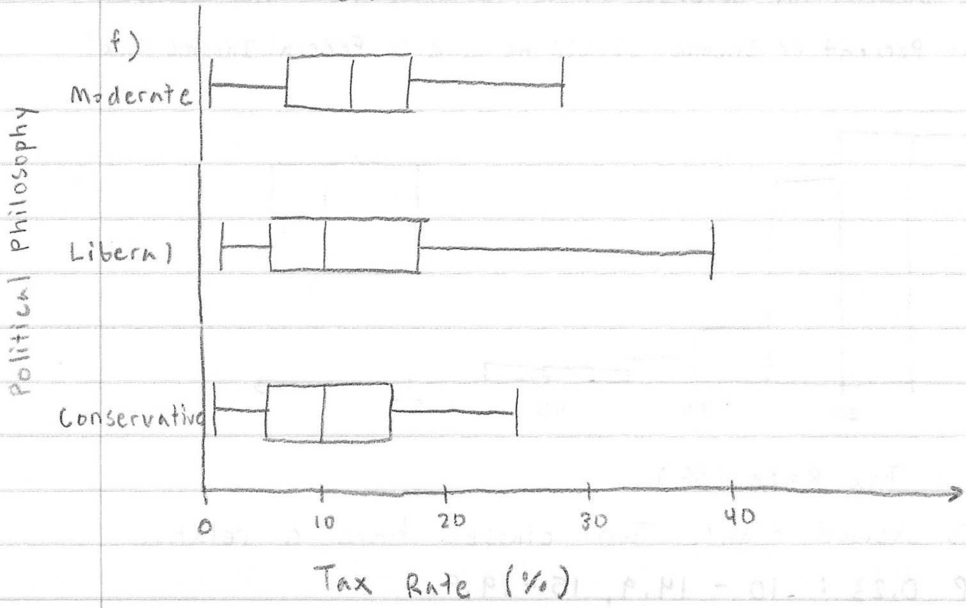
40% is the only outlier.

e)



It seems males are willing to pay more in federal income tax. The median for males is 15%, while the median for females is 10%. However, the range of responses for females is greater than that of males, although both genders have similar interquartile ranges. Both genders have an outlier.

f)



None of the boxplots have outliers. Moderates have the highest median (around 12%). Liberals have the highest tax rate (40%) and the most dispersion (as measured by the IQR and range). Conservatives have the least dispersion (as measured by the IQR and range). Interestingly, liberals and conservatives have the same median tax rate.

4.4

1. A marginal distribution of a variable is a frequency or relative frequency distribution of either the row or column variable in the contingency table. A conditional distribution lists the relative frequency of each category of the response variable, given a specific value of the explanatory variable in the contingency table.

5. a)

	X_1	X_2	X_3	Marginal Distribution
Y_1	20	25	30	75
Y_2	30	25	50	105
Marginal Distribution	50	50	80	180

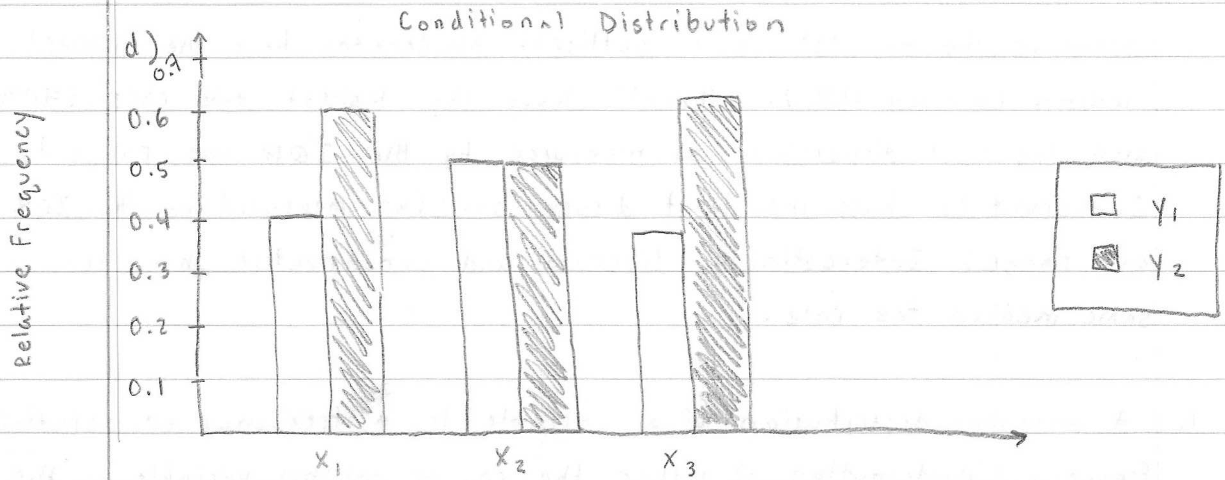
b)

	X_1	X_2	X_3	Rel. Freq. Marginal Dist.
Y_1	20	25	30	0.417
Y_2	30	25	50	0.583
Rel. Freq. Marginal Dist.	0.278	0.278	0.444	1

c)

	X_1	X_2	X_3
Y_1	0.400	0.500	0.375
Y_2	0.600	0.500	0.625
Total	1	1	1

d)



9. a)

	Female	Male	Freq. Marginal Dist.
Republican	105	115	220
Democrat	150	103	253
Independent	150	179	329
Freq. Marginal Dist.	405	397	802

b)

	Female	Male	Rel. Freq. Marginal Dist.
Republican	105	115	0.274
Democrat	150	103	0.315
Independent	150	179	0.410
Rel. Freq. Marginal Dist.	0.505	0.495	1

c) 0.410

d)

	Female	Male
Republican	0.259	0.290
Democrat	0.370	0.259
Independent	0.370	0.451
Totals	1	1

Party Affiliation

e)

f) Yes, males are more likely to be Independents and less likely to be Democrats.

13. a) Smokers: 0.239

Nonsmokers: 0.314

This implies that it is healthier to smoke.

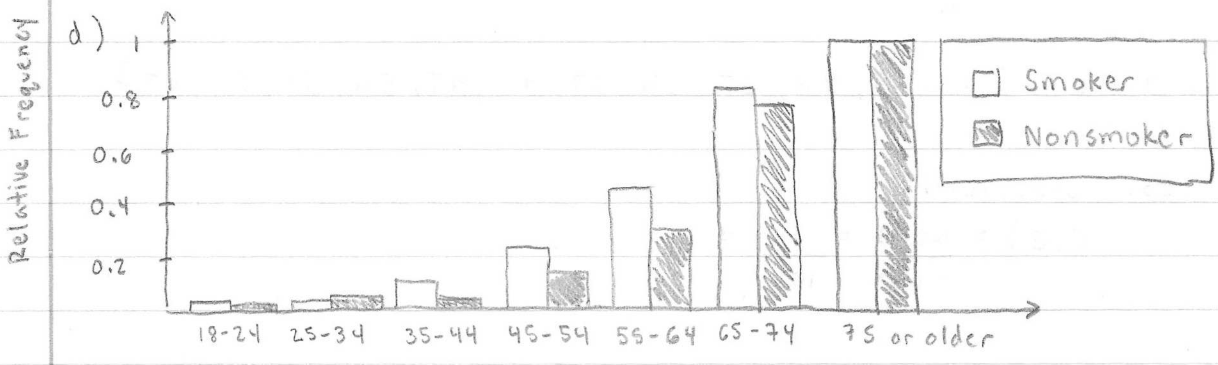
b) Smokers: 0.036

Nonsmokers: 0.016

c)

	18-24	25-34	35-44	45-54	55-64	65-74	75 or older
Smoker	0.036	0.024	0.128	0.208	0.443	0.806	1
Nonsmoker	0.016	0.032	0.058	0.154	0.331	0.783	1

d)



e) When taking age into account, the direction of association changed. In almost all age groups, smokers had a higher death rate than nonsmokers. The most notable exception is for the 25-34 age group, the largest age group for nonsmokers. A possible explanation could be rigorous physical activity that nonsmokers are more likely to participate in than smokers.

5.1

1. Zero. No, when a probability is based on an empirical experiment, a probability of zero does not mean the event cannot occur.

3. True

5. experiment

7. Rule 1: all probabilities are ≥ 0 and ≤ 1
Rule 2: the sum of all probabilities is 1
The outcome "blue" is an impossible event

9. b/c $P(\text{Green}) < 0$

11. 0, 0.01, 0.35, 1

15. Empirical
$P(E) = \frac{95}{100} = 0.95$

19. $S = \{1H, 1T, 2H, 2T, 3H, 3T, 4H, 4T, 5H, 5T, 6H, 6T\}$

23. Classical
$P(E) = \frac{N(E)}{N(S)} = \frac{3}{10} = 0.3$

25. Classical
$E = \{2, 4, 6, 8\}$
$P(E) = \frac{4}{10} = 0.4$

29. a) $P(\text{caught}) = \frac{85}{1000} = 0.085$

b) $P(\text{dropped}) = \frac{296}{1000} = 0.296$

c) $P(\text{hat}) = \frac{2}{85} = 0.024$

If 1000 caught home runs are randomly selected, we would expect 24 to be caught in a hat.

d) $P(\text{failed hat}) = \frac{8}{296} = 0.027$

If 1000 dropped home runs are randomly selected, we would expect 27 to be failed hat attempts.

31. a) $S = \{0, 00, 1, 2, 3, \ldots, 34, 35, 36\}$

b) $P(8) = \frac{1}{38} = 0.0263$

If we spun the wheel 1000 times, we would expect about 26 of those times to result in the ball landing in slot 8.

c) $P(\text{odd}) = \frac{18}{38} = 0.4737$

If we spun the wheel 100 times, we would expect about 47 of the spins to result in an odd number.

Total = 4776

35. a)

Response	Probability
Never	0.026
Rarely	0.068
Sometimes	0.116
Most of the time	0.263
Always	0.527

b) Yes b/c $P(\text{never}) < 0.05$

39. A, B, C, F

5.2

1. Two events are disjoint (aka mutually exclusive) if they have no outcomes in common.

5. E and F = {5, 6, 7}
not mutually exclusive

9. E and G = {}
yes mutually exclusive

13. $P(E \text{ or } F) = 0.25 + 0.45 - 0.15 = 0.55$

17. $P(E^c) = 1 - 0.25 = 0.75$

21. $P(\text{Titleist or Maxfli}) = \frac{9}{20} + \frac{8}{20} = \frac{17}{20} = 0.85$

25. a) Yes, nonnegative and sum is 1
b) $P(\text{rifle or shotgun}) = 0.023 + 0.025 = 0.048$
If 1000 murders in 2013 were randomly selected, we would expect 48 of them to be the result of a rifle or shotgun.
c) $P(\text{handgun, rifle, or shotgun}) = 0.472 + 0.023 + 0.025 = 0.52$
If 100 murders in 2013 were randomly selected, we would expect 52 of them to be the result of a handgun, rifle, or shotgun
d) $P(\text{not a gun}) = P(\text{knives}) + P(\text{hands}) + P(\text{other})$
$= 0.122 + 0.056 + 0.132 = 0.31$
If 100 murders in 2013 were randomly selected, we would expect 31 of them to be the result of a weapon other than a firearm.
e) Yes b/c $P(\text{shotgun}) < 0.05$

29. a) $P(5-5.9) = \frac{23}{125} = 0.184$

b) $P(\text{not } 5-5.9) = 1 - 0.184 = 0.816$

c) $P(<9) = \frac{3+12+16+23+23+21+17+5}{125} = 0.96$

$= 1 - P(9-10) = 1 - \frac{5}{125} = 0.04 = 0.96$

d) $P(>8) = P(8-8.9) + P(9-10) = \frac{5}{125} + \frac{5}{125} = 0.04 + 0.04 = 0.08$

If 100 hospitals in Illinois are randomly selected, we would expect 8 to have received reduced Medicare payments. Given the fact this is due to a poor patient track record, we would like the proportion to be close to zero, so this is not unusual enough (unfortunately).

31. a) $P(\text{heart or club}) = P(\text{heart}) + P(\text{club}) = \frac{13}{52} + \frac{13}{52} = \frac{1}{2} = 0.5$

b) $P(\text{heart or club or diamond}) = \frac{13}{52} + \frac{13}{52} + \frac{13}{52} = \frac{3}{4} = 0.75$

c) $P(\text{ace or heart}) = P(\text{ace}) + P(\text{heart}) - P(\text{ace and heart})$

$= \frac{4}{52} + \frac{13}{52} - \frac{1}{52} = \frac{16}{52} = 0.308$

39. a) $P(\text{died from cancer}) = 0.007$

b) $P(\text{current smoker}) = 0.057$

c) $P(\text{died from cancer and current smoker}) = 0.001$

d) $P(\text{died from cancer or current cigar smoker})$

$= P(\text{died from cancer}) + P(\text{current smoker}) - P(\text{died from cancer and current smoker})$

$= 0.007 + 0.057 - 0.001$

$= 0.063$

5.3

1. independent

3. Addition

5. $P(E) \cdot P(F)$

7. a) Dependent, speeding increases chance of being pulled over
b) Dependent, eating fast food regularly can cause weight gain
c) Independent, they are unrelated

9. $P(E \text{ and } F) = P(E) \cdot P(F) = 0.3 \cdot 0.6 = 0.18$

11. $P(\text{five heads in a row}) = P(\text{head 1st and head 2nd} \ldots \text{and head 5th})$
$= P(\text{head 1st}) \cdot P(\text{head 2nd}) \cdot \ldots \cdot P(\text{head 5th})$
$= \left(\frac{1}{2}\right)^5 = 0.03125$

If we flipped a coin five times, 100 different times, we would expect to observe 5 heads in a row about 3 times.

17. a) $P(\text{two will live to 41 years}) = (0.99757)^2 = 0.99515$
b) $P(\text{five will live to 41 years}) = (0.99757)^5 = 0.98791$
c) $P(\text{at least one of five dies}) = 1 - 0.98791 = 0.01209$
Unusual b/c < 0.05

23. a) $P(\text{system does not fail}) = 1 - P(\text{all 3 fail}) = 1 - (0.03)^3 = 0.999973$
b) Six components

27. $P(\text{at least 1}) = 1 - P(\text{none detect missile}) = 1 - (0.10)^4 = 0.9999$
I'd feel safer if there were even more satellites.

31. a) $P(\text{male and bets on pro sports}) = 0.484 \cdot 0.17 = 0.0823$
b) $P(\text{male or bets on pro sports}) = 0.484 + 0.17 - 0.0823 = 0.5717$
c) The independence assumption is not correct
d) $P(\text{male or bets on pro sports}) = 0.484 + 0.17 - 0.106 = 0.548$
The initial calculation overestimated the probability

5.4

1. F occurring, E has occurred

3. $P(F \mid E) = \frac{P(E \text{ and } F)}{P(E)} = \frac{0.6}{0.8} = 0.75$

5. $P(F \mid E) = \frac{N(E \text{ and } F)}{N(E)} = \frac{420}{740} = 0.5676$

7. $P(E \text{ and } F) = P(F \mid E) \cdot P(E) = 0.4 \cdot 0.8 = 0.32$

9. No

11. $P(\text{club}) = \frac{13}{52} = 0.25$

$P(\text{club} \mid \text{black}) = \frac{P(\text{club and black})}{P(\text{black})} = \frac{0.25}{0.5} = 0.5$

13. $P(\text{rainy} \mid \text{cloudy}) = \frac{0.21}{0.37} = 0.5676$

15. $P(\text{unemployed} \mid \text{dropout}) = \frac{0.021}{0.08} = 0.2625$

17. a) $P(35\text{-}44 \mid \text{more likely}) = \frac{0.1523}{0.6153} = 0.2475$

b) $P(\text{more likely} \mid 35\text{-}44) = \frac{0.1523}{0.2481} = 0.6139$

c) $P(\text{more likely} \mid 18\text{-}34) = \frac{0.1102}{0.2509} = 0.4392$

$P(\text{more likely}) = 0.6153$

No, they are less likely compared to in general

19. a) $P(\text{Female} \mid \text{Sunday}) = \dfrac{0.0582}{0.1636} = 0.3557$

b) $P(\text{Sunday} \mid \text{Female}) = \dfrac{0.0582}{0.3506} = 0.1660$

c) $P(\text{Male} \mid \text{Sunday}) = \dfrac{4143}{6430} = 0.64$

$P(\text{Male} \mid \text{Monday}) = \dfrac{3178}{4883} = 0.65$

$P(\text{Male} \mid \text{Tuesday}) = \dfrac{3280}{5019} = 0.65$

$P(\text{Male} \mid \text{Wed}) = \dfrac{3197}{4926} = 0.65$

$P(\text{Male} \mid \text{Thurs}) = \dfrac{3389}{5228} = 0.65$

$P(\text{Male} \mid \text{Fri}) = \dfrac{3975}{6154} = 0.65$

$P(\text{Male} \mid \text{Sat}) = \dfrac{4749}{7266} = 0.65$

No, it only has a range of 0.64 - 0.65 (essentially the same) For any given day of the week

21. $P(\text{both work}) = \frac{4}{6} \cdot \frac{3}{5} = 0.4$ (Gen. Mult. Rule)

$P(\text{at least } 1 \text{ does not work}) = 1 - 0.4 = 0.6$

27. a) $P(\text{like both}) = \frac{5}{13} \cdot \frac{4}{12} = 0.128$

Not unusual (> 0.05)

b) $P(\text{like neither}) = \frac{8}{13} \cdot \frac{7}{12} = 0.359$

c) $P(\text{like exactly one}) = 1 - 0.128 - 0.359 = 0.513$

d) $P(\text{like both}) = \frac{5}{13} \cdot \frac{5}{13} = 0.148$

$P(\text{like neither}) = \frac{8}{13} \cdot \frac{8}{13} = 0.379$

$P(\text{like exactly one}) = 1 - 0.148 - 0.379 = 0.473$

31. $P(\text{female and smokes}) = P(\text{female} \mid \text{smokes}) \cdot P(\text{smokes})$

$= 0.445 \cdot 0.203 = 0.0903$

It would not be unusual (> 0.05)

5.5

1. permutation

5. $5! = 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1 = 120$

9. $0! = 1$

13. ${}_4P_4 = \frac{4!}{(4-4)!} = \frac{4 \cdot 3 \cdot 2 \cdot 1}{0!} = \frac{24}{1} = 24$

17. ${}_8P_3 = \frac{8!}{(8-3)!} = \frac{8 \cdot 7 \cdot 6 \cdot \cancel{5!}}{\cancel{5!}} = 336$

21. ${}_{10}C_2 = \frac{10!}{2!(10-2)!} = \frac{10 \cdot 9 \cdot \cancel{8!}}{2 \cdot 1 \cdot \cancel{8!}} = \frac{90}{2} = 45$

25. ${}_{48}C_3 = \frac{48!}{3!(48-3)!} = \frac{48 \cdot 47 \cdot 46 \cdot \cancel{45!}}{3 \cdot 2 \cdot 1 \cdot \cancel{45!}} = 34592$

29. ab, ac, ad, ae, bc, bd, be, cd, ce, de

${}_5C_2 = 10$

31. ${}_6C_1 \cdot {}_4C_1 = 6 \cdot 4 = 24$

33. ${}_{12}P_{12} = 479,001,600$

45. ${}_{40}P_3 = 59280$

51. ${}_{50}C_5 = 2{,}118{,}760$

55. Sequence of length $n = 10$

$n_1 = 3, \ n_2 = 2, \ n_3 = 2, \ n_4 = 3$

$$\frac{10!}{3!\,2!\,2!\,3!} = \frac{10 \cdot 9 \cdot 8 \cdot 7 \cdot 6 \cdot 5 \cdot 4 \cdot \cancel{3!}}{\cancel{3!} \cdot 2 \cdot 1 \cdot 2 \cdot 1 \cdot 3 \cdot 2 \cdot 1} = 25{,}200$$

61. a) $P(\text{all students}) = \frac{8}{18} \cdot \frac{7}{17} \cdot \frac{6}{16} \cdot \frac{5}{15} \cdot \frac{4}{14} = 0.0065$ (Gen. Mult. Rule)

b) $P(\text{all faculty}) = \frac{10}{18} \cdot \frac{9}{17} \cdot \frac{8}{16} \cdot \frac{7}{15} \cdot \frac{6}{14} = 0.0294$

c) $P(\text{2 students and 3 faculty}) = \frac{8}{18} \cdot \frac{7}{17} \cdot \frac{10}{16} \cdot \frac{9}{15} \cdot \frac{8}{14} = 0.0392$