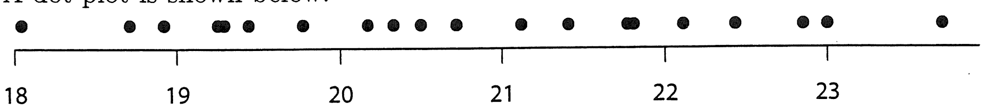
1.2 (a) Mean=20.7675 and Median=20.610.

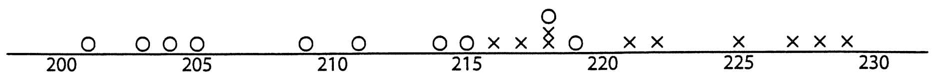
(b) $\bar{x}_{\text{tr10}} = 20.743$.

(c) A dot plot is shown below.



(d) No. They are all close to each other.

1.3 (a) A dot plot is shown below.



In the figure, "×" represents the "No aging" group and "o" represents the "Aging" group.

(b) Yes; tensile strength is greatly reduced due to the aging process.

(c) $\text{Mean}_{\text{Aging}} = 209.90$, and $\text{Mean}_{\text{No aging}} = 222.10$.

(d) $\text{Median}_{\text{Aging}} = 210.00$, and $\text{Median}_{\text{No aging}} = 221.50$. The means and medians for each group are similar to each other.

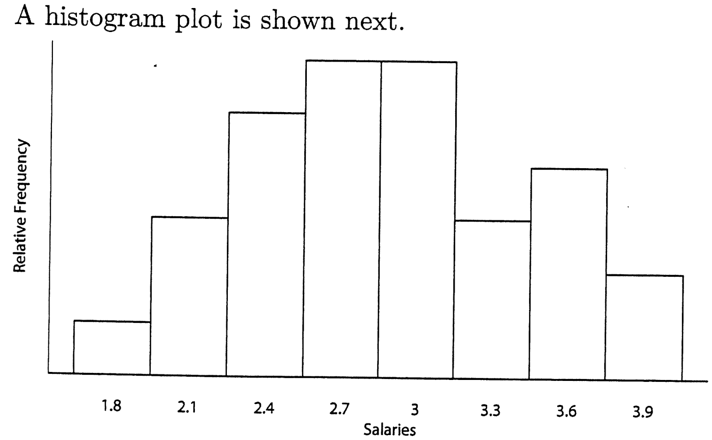
1.8 $s^2 = \frac{1}{20-1}[(18.71 - 20.7675)^2 + (21.41 - 20.7675)^2 + \cdots + (21.12 - 20.7675)^2] = 2.5329$;
$s = \sqrt{2.5345} = 1.5915$.

1.9 (a) $s^2_{\text{No Aging}} = \frac{1}{10-1}[(227 - 222.10)^2 + (222 - 222.10)^2 + \cdots + (221 - 222.10)^2] = 23.66$;
$s_{\text{No Aging}} = \sqrt{23.62} = 4.86$.
$s^2_{\text{Aging}} = \frac{1}{10-1}[(219 - 209.90)^2 + (214 - 209.90)^2 + \cdots + (205 - 209.90)^2] = 42.10$;
$s_{\text{Aging}} = \sqrt{42.12} = 6.49$.

(b) Based on the numbers in (a), the variation in "Aging" is smaller that the variation in "No Aging" although the difference is not so apparent in the plot.

1.24 (a) $\bar{X} = 2.8973$ and $s = 0.5415$.

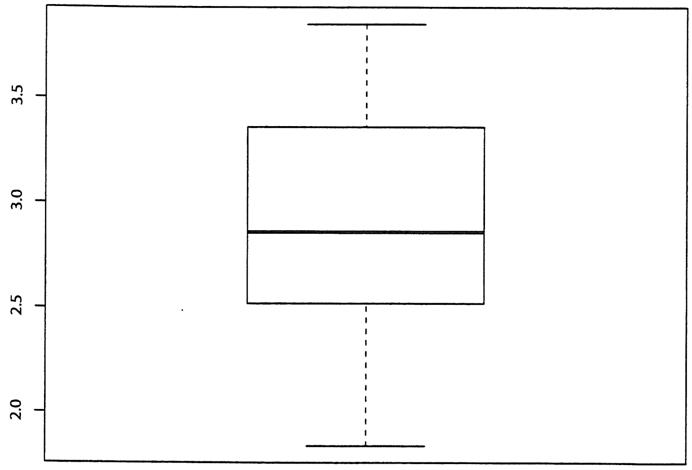
(b) A histogram plot is shown next.



(c) Use the double-stem-and-leaf plot, we have the following.

Stem	Leaf	Frequency
1	(84)	1
2*	(05)(10)(14)(37)(44)(45)	6
2	(52)(52)(67)(68)(71)(75)(77)(83)(89)(91)(99)	11
3*	(10)(13)(14)(22)(36)(37)	6
3	(51)(54)(57)(71)(79)(85)	6

1.29 A box plot is shown next.

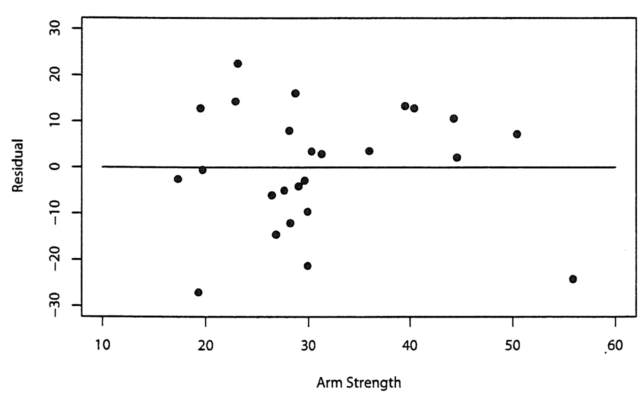


11.1 (a) $\sum_i x_i = 778.7$, $\sum_i y_i = 2050.0$, $\sum_i x_i^2 = 26,591.63$, $\sum_i x_i y_i = 65,164.04$, $n = 25$. Therefore,

$$b_1 = \frac{(25)(65,164.04) - (778.7)(2050.0)}{(25)(26,591.63) - (778.7)^2} = 0.561,$$
$$b_0 = \frac{2050 - (0.5609)(778.7)}{25} = 64.529.$$

(b) Using the equation $\hat{y} = 64.529 + 0.561x$ with $x = 30$, we find $\hat{y} = 64.529 + (0.561)(30) = 81.359$.

(c) Residuals appear to be random as desired.



11.4 (a) $\hat{y} = -1.70 + 1.81x$.

(b) $\hat{x} = (54 + 1.70)/1.81 = 30.77$.

11.14 From the data summary, we obtain

$$b_1 = \frac{(12)(318) - [(4)(12)][(12)(12)]}{(12)(232) - [(4)(12)]^2} = -6.45,$$
$$b_0 = 12 - (-6.45)(4) = 37.8.$$

Hence, $\hat{y} = 37.8 - 6.45x$. It appears that attending professional meetings would not result in publishing more papers.

Problem 1 The following data are the ordered annualized returns for a group of 14 stocks.

-1.2 3.9 8.2 9.0 10.0 11.0 11.5 12.5 13.0 14.8 15.5 16.2 16.7 18.0

1. Draw the histogram to see what is the shape of the distribution.
2. What are the mean, median, standard deviation and IQR?
3. What are the 5-number-summary? Draw the boxplot.
4. What is the central 50% region?
5. What is the percentage of data points between minimum and Q_1?
6. Is it symmetric?

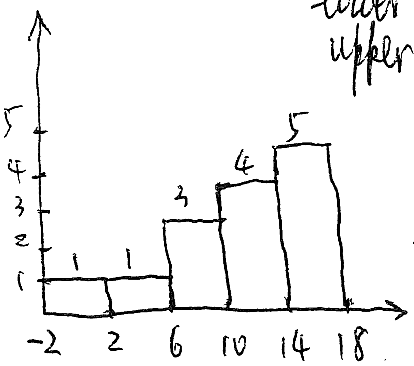
(Hint: Calculator: STAT→EDIT→Put Data Set in One List→STAT→CALC→Tell the Calculator of the List→1-Var Stats)

① $n=14 \approx 2^{b-1} \Rightarrow b=5$

lower $=-2$, upper $=18$

$20/5=4$

(-2, 2] 1
(2, 6] 1
(6 10] 3
(10 14] 4
(14, 18] 5



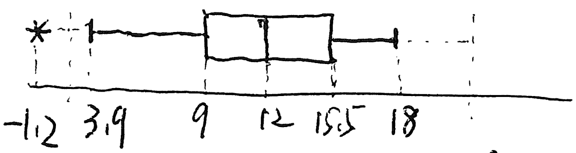
② $\bar{x}=11.364$ median $=12$

$s=5.262$ $IQR=Q_3-Q_1=15.5-9=6.5$

③ (min $=-1.2$, $Q_1=9$, median $=12$, $Q_3=15.5$, max $=18$)

lower fence $=Q_1-1.5IQR=9-9.75=-0.75$

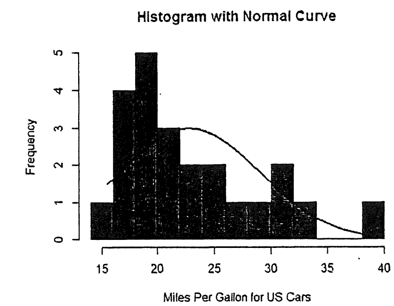
upper fence $=Q_3+1.5IQR=15.5+9.75=25.25$



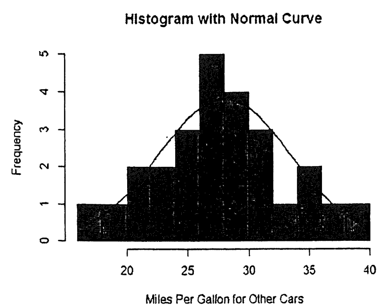
④ [9, 15.5] ⑤ 25%

⑥ Not symmetric.

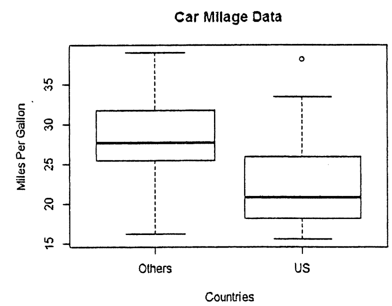
Problem 2 A consumer organization compared gas mileage figures for several models of cars made in the United States with autos manufactured in other countries. The histograms and the boxplots are given below:



(a) Histogram for MPG of Cars made in US



(b) Histogram for MPG of Cars made in Other countries



(c) Boxplots for Comparison

Figure 1: Histogram and Boxplot

1. Describe the shape of the distribution of MPG for cars made in US.
2. What's the relationship between the mean MPG and median MPG for cars made in US.
3. Compare which one performs better, cars in US or cars in other countries.

(1) skew to the right.

(2) mean > median

(3) Cars in other countries perform better since they have higher median MPG.

Problem 3:

1. What is the regression line equation? [regression line: $\hat{y} = \hat{\alpha} + \hat{\beta}x$]

2. What is the correlation coefficient? If we recorded time at the table in hours rather than in minutes, how would the correlation change? [properties of correlation coefficient]

3. What fraction of variation of Calories is accounted for by lunchtime? [coefficient of determination]

4. What is the predicted calories the child consumes given that the child spends at the table for 33.9 minutes? [Make prediction: given a particular value of predictor variable, find out the prediction value of the response variable.]

5. What's the residual for the case of lunchtime 33.9 minutes? [Analysis of residual: Residual=Actual-Predicted]

6. Explain the meaning of intercept and slope.

7. One analyst concluded that "it is clear from this correlation that toddlers who spend more time at the table eat less. Evidently something about being at the table causes them to lose their appetites." Is this conclusion appropriate or not? [Keep in mind that **Correlation is NOT Causation!**]

~~8. Give a 90% CI for the slope.~~

~~9. Is the slope significantly negative? Perform a test at the 1% level.~~

② $r = \dfrac{S_{xy}^2}{S_x S_y} = \dfrac{-135.091}{30.07 \times 7.04} = -0.638$

① $\hat{\beta} = r\dfrac{S_y}{S_x} = -0.638\,\dfrac{30.07}{7.04} = -2.722$

$\hat{\alpha} = \bar{y} - \hat{\beta}\bar{x} = 451.87 + 2.722 \times 34.81$
$= 546.62$

③ $R^2 = r^2 = 0.638^2 = \cancel{0.3661}\ 0.407 \Rightarrow 40.7\%$

④ $\hat{y} = 546.62 - 2.722x \Rightarrow \hat{y} = 546.62 - 2.722 \times 33.9$
$= 454.344$

⑤ $e = y - \hat{y} = 479 - 454.344 = 24.6558$

⑥ when lunchtime is 0, the expected calories consumped ~~response~~ is 546.62. But it's ~~mea~~ meaningless since $x=0$ is far away from the region for x datapoints. when lunchtime increase by 1 minute, the ~~change of~~ expected calories consumed will ~~rate~~ decrease 2.722.

⑦ Not appropriate!