Assignment 2 Solution

1.

Excel output

	Exam
平均數	74.32
標準誤	1.659466
中間值	73
眾數	66
標準差	11.7342
變異數	137.6914
峰度	-0.76047
偏態	0.119988
範圍	47
最小值	49
最大值	96
總和	3716
個數	50

SPSS output

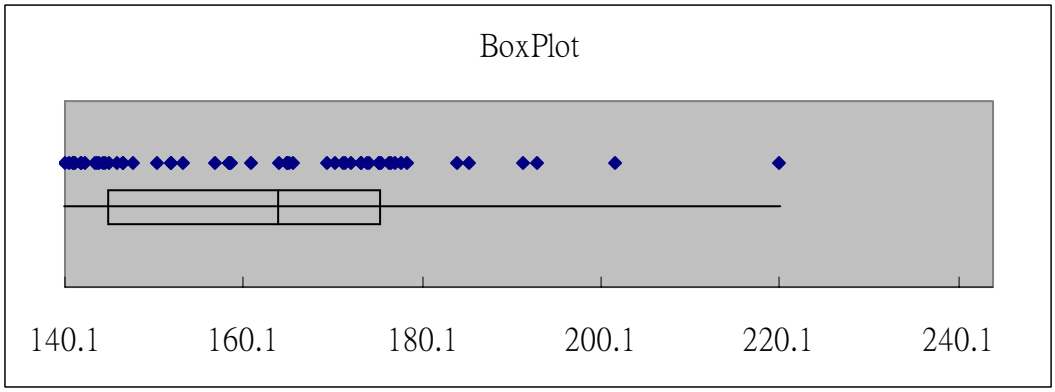
Statistics

Exam		
N	Valid	50
	Missing	0
Mean		74.3200
Median		73.0000
Mode		66.00
Std. Deviation		11.73420
Variance		137.691

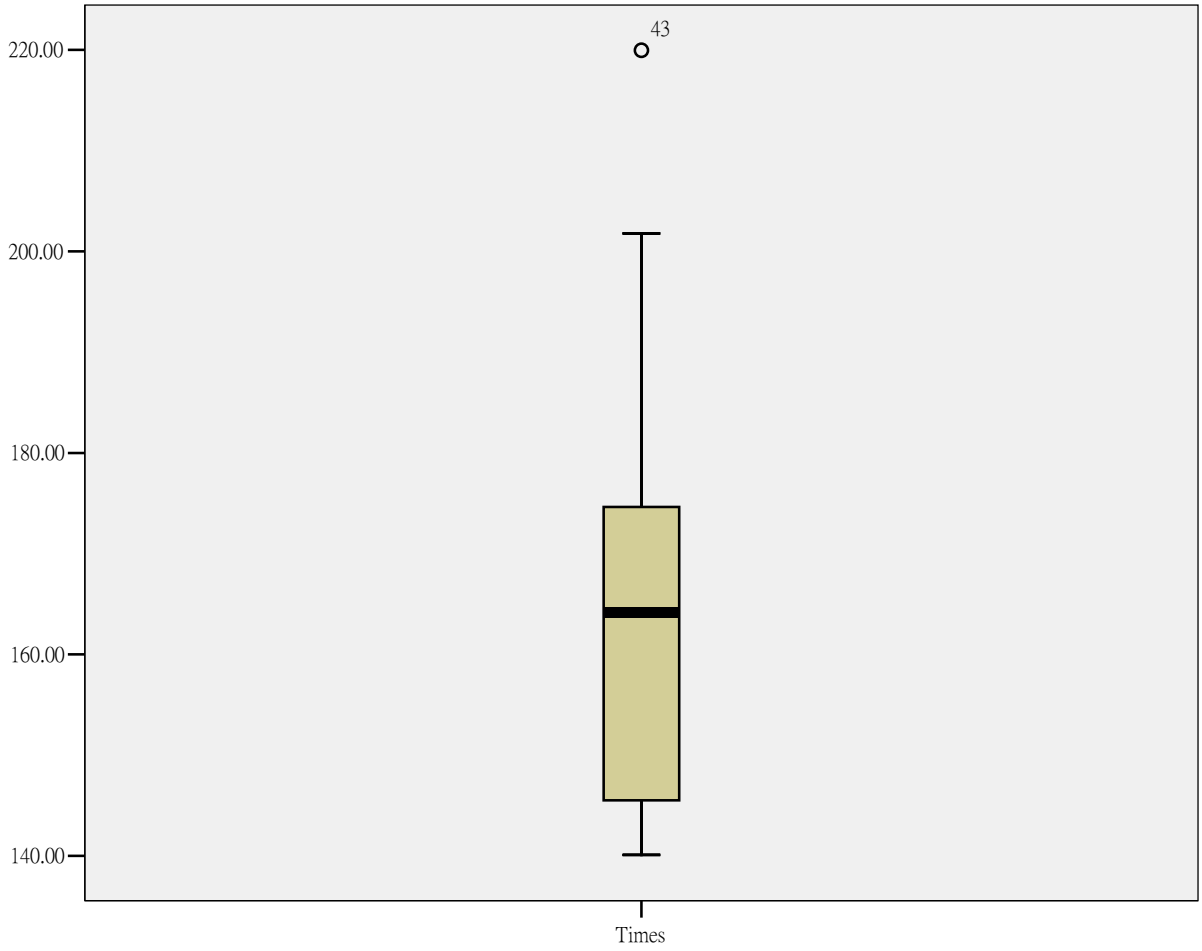
2.

a.

Excel output



SPSS output



b.

From Excel

Times

Smallest = 140.1

Q1 = 145.11

Median = 164.17

Q3 = 175.18

Largest = 219.96

IQR = 30.07

Outliers:

From SPSS

Statistics

Times

N	Valid	55
	Missing	0
Percentiles	25	145.1100
	50	164.1700
	75	175.1800

c.

145.11-1.5*IQR=100.005

175.18+1.5*IQR=220.285

All the data are within the whiskers; therefore, there are no outliers.

d.

The data are positively skewed. One-quarter of the times are below 145.11 and one-quarter are above 175.18.

3.

a.

Excel output

Covariance

	Internet	Education
Internet	88.21614	
Education	11.552	3.672

Coefficient of Correlation

	Internet	Education
Internet	1	
Education	0.641847	1

Because Excel computes the population parameters, you should adjust the values of covariance by multiplying each by 250/249. The table of covariance becomes

	Internet	Education
Internet	88.57043	
Education	11.59839	3.686747

The table of coefficient of correlation also becomes

	Internet	Education
Internet	1	
Education	0.641847	1

SPSS output

Correlations

		Internet	Education
Internet	Pearson Correlation	1	.642**
	Sig. (2-tailed)		.000
	Sum of Squares and Cross-products	22054.036	2888.000
	Covariance	88.570	11.598
	N	250	250
Education	Pearson Correlation	.642**	1
	Sig. (2-tailed)	.000	
	Sum of Squares and Cross-products	2888.000	918.000
	Covariance	11.598	3.687
	N	250	250

**. Correlation is significant at the 0.01 level (2-tailed).

b.

From the statistics, you can see that there is a moderate and positive relationship between Internet use and Education.

4.

a. $\bar{x} = \frac{\Sigma x_i}{n} = \frac{266}{7} = \$38 \text{ per day}$

$s^2 = \frac{\Sigma(x_i - \bar{x})^2}{n-1} = \frac{582}{6} = 97$

$s = \sqrt{97} = \$9.85$

b. The mean car-rental rate per day is $38 for both Eastern and Western cities. However, Eastern cities show a greater variation in rates per day. This greater variation is most likely due to the inclusion of the most expensive city (New York) in the Eastern city sample.

5.

a.

For marketing majors

Mean=36.3, median=35.5, mode=34.2

For accounting majors

Mean=45.7, median=44.7, mode=all

b.

For marketing majors

$(10+1)\times\frac{25}{100} = 2.75, (10+1)\times\frac{75}{100} = 8.25$

First quartile=30.6+0.75*(34.2-30.6)=33.3

Third quartile=39.5+0.25*(42.4-39.5)=40.225

For accounting majors

$(16+1)\times\frac{25}{100} = 4.25, (16+1)\times\frac{75}{100} = 12.75$

First quartile=40.8+0.25*(41.1-40.8)=40.875

Third quartile=49.7+0.75*(49.9-49.7)=49.85

c.

Yes, it can be seen that the mean starting salary for accounting majors is larger than that for marketing majors. The difference is also significant from the first and third quartile of both samples.

6.

Order number:28, 42, 45, 48, 49, 50, 55, 55, 58, 60

a. Range=60-28=32

IQR= $Q_3 - Q_1 = 56.5 - 43.5 = 13$

b. $\bar{x} = \frac{435}{9} = 48.33$

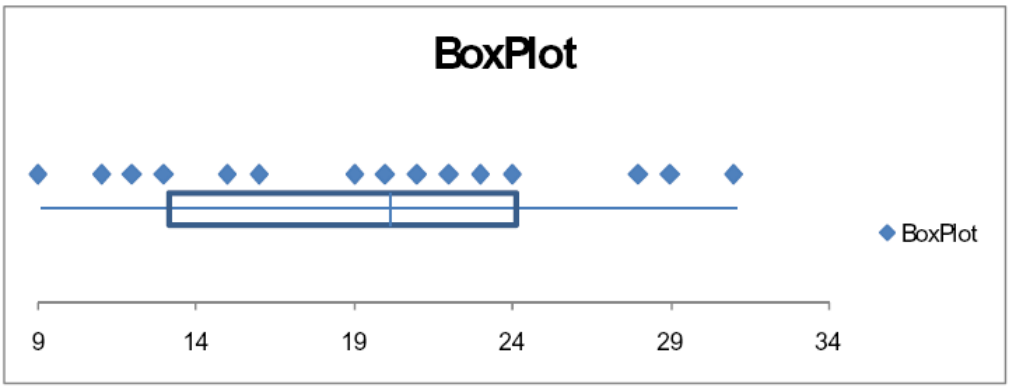
$\Sigma(x_i - \bar{x})^2 = 742$

$s^2 = \frac{\Sigma(x_i - \bar{x})^2}{n-1} = \frac{742}{8} = 92.75$

$s = \sqrt{92.75} = 9.63$

c. The average air quality is about the same. But, the variability is greater in Anaheim.

7.



Smallest=9, Q1=13, Q2=20, Q3=24, Largest=31, IQR=11, No outliers

8.

x_i	y_i	x_i^2	y_i^2	$x_i y_i$
40	77	1,600	5,929	3,080
42	63	1,764	3,969	2,646
37	79	1,369	6,241	2,923
47	86	2,209	7,396	4,041
25	51	625	2,601	1,276
44	78	1,936	6,084	3,432
41	83	1,681	6,889	3,403
48	90	2,304	8,100	4,320
35	65	1,225	4,225	2,275
28	47	784	2,209	1,316
Total 387	719	15,497	53,643	28,712

$$\sum_{i=1}^{n} x_i = 387 \quad \sum_{i=1}^{n} y_i = 719 \quad \sum_{i=1}^{n} x_i^2 = 15{,}497 \quad \sum_{i=1}^{n} y_i^2 = 53{,}643 \quad \sum_{i=1}^{n} x_i y_i = 28{,}712$$

a.

$$s_{xy} = \frac{1}{n-1}\left[\sum_{i=1}^{n} x_i y_i - \frac{\sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{n}\right] = \frac{1}{10-1}\left[28{,}712 - \frac{(387)(719)}{10}\right] = 98.52$$

$$s_x^2 = \frac{1}{n-1}\left[\sum_{i=1}^{n} x_i^2 - \frac{\left(\sum_{i=1}^{n} x_i\right)^2}{n}\right] = \frac{1}{10-1}\left[15{,}497 - \frac{(387)^2}{10}\right] = 57.79$$

$$s_y^2 = \frac{1}{n-1}\left[\sum_{i=1}^{n} y_i^2 - \frac{\left(\sum_{i=1}^{n} y_i\right)^2}{n}\right] = \frac{1}{10-1}\left[53{,}643 - \frac{(719)^2}{10}\right] = 216.32$$

b.

$$r = \frac{s_{xy}}{s_x s_y} = \frac{98.52}{\sqrt{(57.79)(216.32)}} = .8811$$

c.

There is a strong positive linear relationship between marks and study time.