

166 Business Statistics

19. Compute coefficient of variation from the following table :

Weekly Income (in Rs.)	No. of employees	Weekly Income (in Rs.)	No. of employees
1300-1399	30	1700-1799	60
1400-1499	46	1800-1899	50
1500-1599	58	1900-1999	20
1600-1699	76		

20. The following table gives the fluctuations in the prices of shares of two companies A and B. Find out which of them show greater variability. Comment on the result.

Price (in Rs.)	Share A	Share B	Price (in Rs.)	Share A	Share B
318	2,542	324	2,545		
322	2,522	315	2,530		
325	2,534	308	2,556		
312	2,532	319	2,530		

[Share A : C.V. = 1.75 ; Share B : C.V. = 0.42]

21. From an analysis of monthly wages paid to workers in two organisations C and D, the following results were obtained:

	C	D
No. of workers	550	600
Average monthly wages	2260	2348
Variance of the distribution of wages	100	144

Obtain the average monthly wages and the variability in individual wages of all the workers in the organisations taken together.

22. From the following table giving data regarding income of workers in two factories draw a graph (Lorenz curve) to show which factory has greater inequalities of income :

Income (Rs.)	Factory A	Factory B
Below 1500	6,000	5,000
1,500-2,000	4,250	4,500
2,000-3,000	3,600	4,800
3,000-4,000	1,500	2,200
4,000-5,000	650	1,500

[Factory A]

23. A distribution consists of three parts, characterised as follows :

Part	Number of items	Arithmetic average	S.D.
1	1200	5	3
2	2250	10	4
3	3300	15	5

Show that the arithmetic average of the whole distribution is 10.67 and its standard deviation is 5.83 approximately.

24. Lives of two models of refrigerators in a recent survey are :

Life (No. of years)	Model A	Model B
0-2	5	2
2-4	16	7
4-6	13	12
6-8	7	19
8-10	5	9
10-12	4	1

What is the average life of each of these refrigerators ? Which model has greater uniformity ?

[Model A: $\bar{X} = 5.12$, C.V. = 54.9; Model B: $\bar{X} = 6.16$, C.V. = 36.2]
(MBA, Bharthidasan Univ., 2001; IAS, 2002; C.S.E., 2002)

25. The mean of two samples of size 50 and 100 respectively is 54.1 and 50.3 and the standard deviations are 8 and 7. Obtain the standard deviation of the sample of size 150 obtained by combining the two samples.

[7.55]

Jewel's Care Collected

26. The mean and standard deviation of 20 items is found to be 10 and 2 respectively. At the time of checking it was found that one item 8 was incorrect. Calculate the mean and standard deviation if item 8 was incorrect. Calculate the mean and standard deviation if (i) the wrong item is omitted, and (ii) it is replaced by 12.

[10.1, 2.26, 10.2.2]

27. The mean and standard deviation of 1,000 observations of a frequency distribution (grouped in intervals 0-10, 10-20, etc.) were found to be 35 and 8. Later it was discovered that in calculating these values the errata which was supplied with the data was not considered. The errata read as follows :

- Group 0-10 for frequency 25 read frequency 52
- Group 10-20 for frequency 75 read frequency 57
- Group 20-30 for frequency 121 read frequency 21
- Group 30-40 for frequency 137 read frequency 73
- Group 40-50 for frequency 59 read frequency 95.

Calculate the correct mean and standard deviation.

28. An analysis of the weekly wages paid to workers in two firms A and B belonging to the same industry, gives the following results:

	Firm A	Firm B
Number of wage-earners	550	650
Average daily wages	100	95
Standard deviation of the distribution of wages	$\sqrt{90}$	$\sqrt{120}$

- (a) Which firm A or B pays out large amount as daily wages ?
- (b) In which firm A or B is there greater variability in individual wages ?
- (c) What are the measures of (i) average daily wages and (ii) standard deviation in the distribution of individual wages of all workers in the two firms taken together ?

(MBA, M.D. Univ.; Diploma in Mgt., AIMA, Dec. 1999)

29. A factory produces two types of electric lamps A and B. In an experiment relating to their life, the following results were obtained :

Length of life (in hours)	No. of lamps	
	A	B
500-700	5	4
700-900	11	30
900-1,100	26	12
1,100-1,300	10	8
1,300-1,500	8	6

Compare the variability of the life of the two types using coefficient of variation.

[$CV_A = 21.64$; $CV_B = 23.41$]

30. In a small town, a survey was conducted in respect of profits made by retail shops. The following results were obtained:

Profit or Loss (in '000 Rs.)	No. of shops	Profit or Loss (in '000 Rs.)	No. of shops
-4 to -3	4	1 to 2	56
-3 to -2	10	2 to 3	40
-2 to -1	22	3 to 4	24
-1 to 0		28	4 to 518
0 to 1		38	5 to 610

Calculate (i) the average profit made by a retail shop.

(ii) total profit made by all shops, and (iii) the coefficient of variation of earnings.

[(i) 1348; (ii) 3,37,000; (iii) 152.8]

31. The following is the distribution of amounts spent for research and development and for marketing for the year 2003 by 10 drug firms and cosmetic firms :

Expenditure (Lakhs of rupees)	Drug Cos. R and D	Cosmetic Cos. Marketing	R and D/Marketing
10-202	3	5	0
20-302	2	3	2
30-403	3	2	4
40-503	2	0	4

Jewel's Care Collected

168 Business Statistics

- (a) Compute the arithmetic mean for each type of company, for each type of expenditure.
 (b) Compute the standard deviation for each type of company, for each type of expenditure.
 (c) What conclusion do you draw from the results?

(MBA, Kurukshetra Univ.)

$[\bar{X} = 32, 29, 22, 37; \sigma = 11, 11.14, 7.81, 7.48]$

32. The life of two types of tyres in a sample survey is given below :

Life (in km)	Type A	Type B
5,000-10,000	18	15
10,000-15,000	22	24
15,000-20,000	26	30
20,000-25,000	25	18
25,000-30,000	9	13

- (a) Which of the two types of tyre give a higher average life?
 (b) If prices are same for both the types, which type would you prefer and why?

(MBA, Delhi Univ., 1999)

[(a) type B, (b) Type B]

33. A collar manufacturer is considering the production of new style of collar to attract young men. The following statistics of neck circumference are available based on measurement of a typical group:

Mid-value

(in inches)	12.5	13.0	13.5	14.0	14.5	15.0	15.5	16.0
No. of Students	4	19	30	63	66	29	18	1

Compute the arithmetic mean and standard deviation and comment on the results.

$[14.22, 0.70]$

34. Calculate the arithmetic mean, median and standard deviation for the following distributions

Height (inches)	No. of persons	Height (inches)	No. of persons
60 less than 63	4	69 less than 72	33
63 less than 66	14	72 less than 75	8
66 less than 69	59	75 less than 78	2

$[68.32; 68.14; 2.79]$

35. The following data relate to the number of bonds applied for number of applicants and number of bonds allotted to each applicant by the Rayon Silk Mfg. (Wvg.) Co. Ltd. in Jan. 2004:

No. of Bonds applied for	No. of Applicants	No. of bonds allotted to each applicant
5-15	61,685	5
20-45	18,879	10
50-105	7,230	15
110-185	647	20
190-300	177	25
305-325	79	30

Calculate the average number of bonds allotted to each applicant and the standard deviation.

36. The standard deviation of a distribution of 100 values was Rs. 2. If the sum of the squares of the actual values was Rs. 3,600, what was the mean of this distribution?

$[5.66]$

37. 32 trials of a process to finish a certain job revealed the following information :

Mean time taken to complete the job = 1 hr. 20 mts.

Standard deviation = 16 mts.

Another set of 8 trials gave mean time as 100 minutes and standard deviation 25 minutes.

Find the combined mean and standard deviation.

$[\bar{X}_c = 84, \sigma_c = 19.84]$

Jewel's Care Collected

38. The following table relates to the profits and losses of 100 firms. Calculate the average profits and the standard deviation of profits :

Profits & Loss	No. of firms	Profits & Loss	No. of firms
5,000-6,000	8	0-1,000	6
4,000-5,000	12	(-) 1,000-0	5
3,000-4,000	30	(-) 2,000-(-)1,000	8
2,000-3,000	10	(-) 3,000-(-)2,000	9
1,000-2,000	5	(-) 4,000-(-)3,000	7

$[\bar{x} = 531, \sigma = 214.8, C.V. = 40.5\%]$

39. A study of 241 authors revealed the following data on the distribution of age :

Age (years)	Number of Authors
up to 30	20
up to 40	73
up to 50	80
up to 60	44
up to 70	22
up to 80	2

Compute the mean and coefficient of variation of the distribution.

$[\bar{x} = 44.21, \sigma = 11.18, C.V. = 25.29]$

Jewel's Care Collected

40. A survey of domestic consumption of electricity gave the following distribution of the no. of units consumed:

Number of units	Number of consumers	Number of units	Number of consumers
0-200	9	800-1000	45
200-400	18	1000-1500	38
400-600	27	1500-2000	20
600-800	32	2000 and above	2

Use a graphical method to calculate as accurately as possible the two quartiles and hence find the quartile deviation.

$[Q_1 = 570, Q_3 = 1250, Q.D. = 340]$

41. The Shareholder Research Bureau of India conducted recently a research study on the price behaviour of three leading industrial shares A, B, C for the period 1996 to 1999, the results of which are published as following in the quarterly journal :

Share	Average Price (Rs.)	Standard Deviation (Rs.)	Current selling price (Rs.)
A	18.00	5.40	36.00
B	22.50	4.50	34.75
C	24.00	6.00	39.00

- (i) Which share in your opinion appears to be more stable in value ?
- (ii) If you are the holder of all three shares which one would you like to dispose of at present, and why ?

$[C.V.: (A) = 30, (B) = 20, (C) = 25]$

42. Find the missing value from the following table :

Sub-group	N	\bar{x}	Variance
A	?	25	9
B	250	?	96
C	300	15	?
	750	16	31.733

$[N_1 = 200, \bar{x}_2 = 10, \sigma^2 = 25]$

43. The number examined, the mean weight and the standard deviation in each group of examination and two medical examiners are given below. Find the weight and standard deviation of both groups taken together.

Medical Examiner	Number Examined	Mean Weight	Standard Deviation
A	50	113 pounds	6.5 pounds
B	60	120 pounds	8.2 pounds

170 Business Statistics

44. Blood serum cholesterol levels of 10 persons are as under :

240, 260, 290, 245, 255, 288, 272, 263, 277, 250.

Calculate standard deviation with the help of assumed mean.

[$\sigma = 16.48$]

45. Mean and standard deviation of the following continuous series are 31 and 15.94 respectively. The distribution after step deviations is as follows :

d :	-3	-2	-1	0	1	2	3
f :	10	15	25	25	10	11	5

Determine the actual class-intervals.

(MBA, Vikram Univ., 2002)

[$i = 10, A = 35, 0-10, 10-20, 20-30, 30-40$ etc.]

46. The mean of 5 observations is 15 and the variance is 9. If two more observations having values -3 and 10 are combined with these 5 observations, what will be the mean and variance of 7 observations ?
47. Daily rated employees in SLM-Devital, an engineering firm, earn Rs. 66 per day. The workers estimate that on the average they turn out 30 pieces per day with a standard deviation of six pieces per day. Under a suggested piece rate plan how much will they ask per piece if they wish to earn more than their present daily income 90 per cent of the time ?
48. The coefficients of variation of wages of male workers and female workers are 55 per cent and 70 per cent respectively, while the standard deviations are 220 and 15.4 respectively. Calculate the overall average wage of all workers given that 80 per cent of the workers are male.
49. Calculate the standard deviation from the following data :

Temperature (C)	No. of days	Temperature (C)	No. of days
-40 to -30	10	10 to 20	65
-30 to -20	24	20 to 30	180
-20 to -10	30		14
-10 to -0	42		

(MBA, Jodhpur Univ., 1996)

50. Calculate coefficient of variation from the following data :

No. of days absent:	5-10	10-15	15-20	20-25	25-30
No. of Students :	25	140	250	108	5221

[C.V. = 41.32%]

51. The index number of prices of Cotton and Coal shares in 2003 were as under :

Month	Jan.	Feb.	March	April	May	June	July	Aug	Sept.	Oct.
Cotton	188	178	173	164	172	184	184	185	211	217
Coal	131	130	130	129	129	120	127	127	130	137

Which of the shares you consider more variable in price?

(MBA, Osmania Univ.)

52. (a) The mean and standard deviation of 17 observations were found to be 25 and 5 respectively. Later on it was found that two values 51 and 31 were wrongly read as 35 and 13 respectively. Find the correct mean and standard deviation.

(MBA, M.D. Univ., 1994)

[Correct : $\bar{y} = 27, \sigma = 6.96$]

- (b) In a survey, data on daily wages paid to workers of two factories A and B are as follows :

(No. of workers)

Daily wages :	20-30	30-40	40-50	50-60	60-70	70-80	80-90
Factory A :	15	30	44	50	30	14	7
Factory B :	25	40	60	35	20	15	5

Find out :

- (a) Which factory pays higher average wages ? By how much ?
 (b) Wages of which have greater variability.
 (c) Monthly wage bill of both Factories (month = 25 days)

(MBA, M.D. Univ., 1994)

The number of employees, wages per employee and the variance of the wages per employee for two factories is given below :

	Factory A	Factory B
No. of employees	100	150
Average wage per employee per month (Rs.)	3,200	2,800
Variance of the wages per employee per month (Rs.)	625	720

- (a) In which factory is there greater variation in the distribution of wages per employee ?
 (b) Suppose in factory B, the wages of an employee were wrongly noted as Rs. 3050 instead of 3650, what would be the correct variance for factory B ? (MBA, Kuruman Univ., 2003)
11. Name the various measures of dispersion. How would you compare the performance of two companies which reported profits for last five years as follows :

Company I	: 4.0	4.1	4.3	4.0	4.1	
Company II	: 7.3	-3.7	8.4	-2.5	11.0	(MBA, M.D. Univ., 2000)

12. Calculate variance and coefficient of variation from the following data :

Profits (Rs. crores)	No. of Cos.	Profits (Rs. crores)	No. of Cos.
Less than 10	8	Less than 40	70
" " 20	20	" " 50	90
" " 30	40	" " 68	100

(MBA, Guru Jambhwar Univ., 2003)

13. Assume that your pathology lab. provides the following details of blood test carried out on 100 patients for diabetes :

Blood Sugar	: 90-100	100-110	110-120	120-130	130-140	140-150	150-160
Level (testing)	: 16	20	22	25	10	4	3

(MBA, IICA, 2002)

Calculate coefficient of variation. What inference do you draw.

($\bar{y} = 116.7, \sigma = 15.17, C.V. = 13\%$)

14. For two firms A and B belonging to the same industry, the following details are available.

	Firm A	Firm B
No. of Employees	100	200
Average wage monthly	Rs. 2400	Rs. 1800
S.D.	Rs. 60	Rs. 80

- (i) Which firm pays out larger amount as wages ?
 (ii) Which firm shows greater variability in the distribution of wages ?
 (iii) Find average monthly wage and standard deviation of all the employees in both the firms.

(MBA, D.U. Cal., 2003)

Jewel's Care Collected

PROBLEMS

1-A: Answer the following questions, each question carries one mark:

- (i) What is skewness?
- (ii) Point out the role of studying skewness.
- (iii) Name the various methods of finding skewness.
- (iv) What are kurtosis?
- (v) What are moments?
- (vi) How are the values of β_1 and β_2 calculated?
- (vii) Give the formula for finding Karl Pearson's coefficient of skewness.
- (viii) What is Bowley's method of finding skewness?
- (ix) What is symmetrical distribution?
- (x) Distinguish between positive and negative skewness.

1-B: Answer the following questions, each question carries four marks:

- (i) Distinguish between positively and negatively skewed distribution.
 - (ii) In what type of situations Karl Pearson's or Bowley's method should be preferred?
 - (iii) Would the various methods of studying skewness lead to same answer? If not, give reasons.
 - (iv) What are the various methods of studying kurtosis?
 - (v) Explain the terms leptokurtic, platykurtic and mesokurtic with a suitable diagram.
1. (a) Explain briefly the different methods of measuring skewness. *(MBA, Kumaun Univ., 2000)*
 - (b) What do you understand by the terms skewness and kurtosis? Point out their role in analysing a frequency distribution. *(MBA, Delhi Univ., 1994)*
 2. Take any suitable imaginary data and explain how would you measure skewness and kurtosis.
 3. Distinguish between Karl Pearson's and Bowley's measure of skewness. Which one of these would you prefer and why? *(MBA, Delhi Univ., 2000)*
 4. In what way measures of central tendency, variation, skewness, and kurtosis are complementary to one another in understanding a frequency distribution? Elucidate. *(MBA, Osmania Univ.; MBA, Sukhadia Univ., 1995)*
 5. Define 'Moments'. How can you find out skewness and kurtosis of a distribution from moments about the mean?
 6. Explain clearly how the moments help in describing the characteristics of a frequency distribution. *(MBA, Delhi Univ., 1997)*
 7. Explain clearly how the measures of skewness and kurtosis can be used in describing a frequency distribution.
 8. What is meant by 'moments' of a distribution? Show how moments are used to describe the characteristics of a distribution, i.e., central tendency, dispersion, skewness and kurtosis.
 9. What are the raw and the central moments of a distribution? Show that the central moments are invariant under change of origin but not under change of scale.
 10. Define raw and central moments of a frequency distribution. Express the second, third and fourth central moments in terms of raw moments.
 11. (a) Explain the terms 'Skewness' and 'Kurtosis' used in connection with the frequency distribution of a continuous variable.
Give the different measures of skewness (any two of the measures to be given) and kurtosis.
 - (b) Define and discuss the 'quartiles' of a distribution. How are they used for measuring variation and skewness?
 12. Define moments. Establish the relationship between the moments about mean in terms of moments about any arbitrary point and vice-versa.
 13. (a) Define moments. How are they helpful in study of the different aspects of the formation of a frequency distribution?
 - (b) "A frequency distribution can be described almost completely by the first four moments and the two measures based on the moments." Examine.
 14. (a) Explain the third and fourth central moment in terms of the first four moments about the origin.
 - (b) Distinguish between variation and skewness and point out the various methods of measuring skewness.
 - (c) Explain the term 'skewness'. What purpose does a measure of skewness serve? Comment on some of the well-known measures of skewness.

Jewel's Care Collected

16. (a) Distinguish between skewness and kurtosis.
 (b) Briefly mention the tests which can be applied to determine the presence of skewness.
17. (a) How do measures of central tendency, dispersion, skewness and kurtosis help in analysing a frequency distribution? Explain with the help of an example. (MBA, Sukhadia Univ., 1999)
 (b) Find out coefficient of skewness from the following table giving wages of 240 persons :

Wages (Rs.)	No. of persons	Wages (Rs.)	No. of persons
2000-2200	12	2800-3000	50
2200-2400	18	3000-3200	45
2400-2600	35	3200-3400	30
2600-2800	42	3400-3600	8

(Sk = - 0.267)

18. Calculate Karl Pearson's coefficient of skewness from the following data :

Profit (Rs. Lakhs)	No. of Cos.	Profit (Rs. Lakhs)	No. of Cos.
400-450	8	600-650	62
450-500	10	650-700	32
500-550	30	700-750	15
550-600	45	750-800	8

19. The following data represent the percentage of ash content in a particular variety of coal as determined by test on 280 wagon loads :

Percentage of ash content	Frequency	Percentage of ash content	Frequency
Less than 6.0	0	10.0-10.9	84
6.0-6.9	1	11.0-11.9	45
7.0-7.9	7	12.0-12.9	28
8.0-8.9	28	13.0-13.9	7
9.0-9.9	78	14.0-14.9	2

Calculate the quartile coefficient of skewness. Also compare the proportion of the total frequency outside the limit $\bar{X} \pm 2\sigma$ for the distribution.

[Sk=0.05; 2.3]

20. From the following data of daily travelling allowance (in Rs.) of salesmen, calculate coefficient of skewness and comment on its value :

Travelling allowance (per day)	No. of salesmen	Travelling allowance (per day)	No. of salesmen
110-115	4	135-140	90
115-120	10	140-145	52
120-125	26	145-150	33
125-130	49	150-155	17
130-135	72	155-160	7

21. From the following data pertaining to profits (Rs. lakhs) for 50 companies, calculate moments β_1 and β_2 :

Profits (Rs. Lakhs)	No. of Companies
70-90	8
90-110	11
110-130	18
130-150	9
150-170	4

$[\mu_1 = 528, \mu_2 = 960, \mu_3 = 642816, \beta_1 = 0.006, \beta_2 = 2.31]$

Jewel's Care Collected

22. A record was kept over a period of 6 months by a sales manager to determine the average number of calls made per day by his six salesmen. The results are shown below :

Salesmen	:	A	B	C	D	E	F
Average number of calls per day	:	8	10	12	15	7	5

- (i) Compute a measure of skewness. Is the distribution symmetrical?
 (ii) Compute a measure of kurtosis. What does this measure mean?

$[\beta_1=0.11; \beta_2=1.97]$

23. Locate the mode and calculate mean and standard deviation of the following distribution and using your results comment on the skewness of the distribution :

Scores	Frequency	Scores	Frequency
10-15	2	35-40	6
15-20	8	40-45	4
20-25	6	45-50	3
25-30	12	50-55	1
30-35	7	55-60	1

$[\bar{X} = 30.1; Mo. = 27.73; \sigma = 10.45; Sk = 0.227]$

(MBA, Delhi Univ., 2002)

24. You are given the following information before and after the settlement of an industrial dispute :

	Before settlement of dispute	After settlement of dispute
No. of workers	1100	950
Average wage (Rs.)	2350	2400
Standard deviation (Rs.)	425	400
Median wage (Rs.)	2375	2325

Comment on the gains and losses from the point of view of workers and that of management.

25. The arithmetic mean of a distribution is 5. The second and the third moments about the mean are 20 and 140 respectively. Find the third moment of the distribution about 10.

[-285]

26. For the frequency distribution given below, calculate the coefficient of skewness based on the quartiles :

Class limits	Frequency	Class limits	Frequency
10-19	5	50-59	25
20-29	9	60-69	5
30-39	14	70-79	8
40-49	20	80-89	4

27. (a) For a distribution, Bowley's coefficient of skewness is -0.48 , $Q_3 = 10.2$ and Median = 14.4. What is the quartile coefficient of distribution?

- (b) Karl Pearson's coefficient of skewness of a distribution is $+0.4$. Its standard deviation is 10 and mean 40.5. Find the mode and median of the distribution.

- (c) Find coefficient of skewness from the information given below :

$Q_1 = 60, Q_3 = 75, Med. = 68.$

- (d) The following information was obtained from the records of a factory relating to wages; $\bar{X} = 275, Med. = 260, \sigma = 45.8$

Give as much information as you can about the distribution of wages.

$[(a) 0.22 (b) 39.17 (c) -0.07 (d) Sk = 0.98]$

28. The first three moments of a distribution about the value 7 calculated from a set of 9 observations are 0.2, 19.4 and -41.0 . Find the measures of central tendency and dispersion and also the third moment about origin.

$[7.2, 4.4, -52.624]$

29. The first four moments of a distribution about $A = 4$ are 1, 4, 10 and 45. Obtain the various characteristics of the distribution on the basis of the information given. Comment upon the nature of the distribution.

$[\beta_1 = 0, \beta_2 = 2.897]$

30. (a) State the use of quartiles for measuring dispersion and skewness.

- (b) Calculate Bowley's coefficient of skewness from the following data :

Mid-value	75	100	125	150	175	200	225	250
Frequency	35	40	48	100	125	80	50	22

[-0.032]

Jewel's Care Collected

31. A prospective buyer tested the bursting pressure of the sample of polythene bags received from a manufacturer. The results are given below:

Bursting pressure (in lbs.)	5-10	10-15	15-20	20-25	25-30	30-35
No. of bags	2	20	30	50	6	2

The buyer calculated the mean and mode of the sample as 20.2 lbs. and 21.5 lbs. respectively.

Calculate (i) coefficient of variation, (ii) Karl Pearson's coefficient of skewness for bursting pressure.

32. From the following data, calculate coefficient of variation and coefficient of skewness:

Age (in years)	25-30	30-35	35-40	40-45	45-50	50-55	55-60
No. of employees	9	18	30	40	10	7	6

33. The frequency distribution of weekly wages (in Rs.) in a certain factory is as follows:

Weekly wages	No. of workers	Weekly wages	No. of workers
423-427	2	448-452	16
428-432	6	453-457	12
433-437	9	458-462	6
438-442	14	463-467	2
443-447	32	468-472	1

Find Karl Pearson's coefficient of skewness and interpret its value. [Sk_p = 0.0572]

(MBA, Delhi Univ., 1997)

34. A survey was conducted by a manufacturing company to enquire the maximum price at which persons would be willing to buy their product. The following table gives the stated prices (in rupees) by persons:

Price (in Rs.)	80-90	90-100	100-110	110-120	120-130
No. of persons	11	29	18	27	15

Calculate Bowley's coefficient of skewness and interpret its value.

(MBA, Delhi Univ., 2002)

35. The standard deviation of a symmetrical distribution is 3. What must be the value of fourth moment about the mean in order that the distribution be mesokurtic?

36. Calculate coefficient of variation and Karl Pearson's coefficient of skewness from the data given below:

Sales (Rs. crores) Less than	40	50	60	70	80
No. of Companies	8	20	50	72	80

[Coeff. of Variation = 19.55, Coeff. of Sk = -0.06]

37. Assume that a firm has selected a random sample of 100 from its production line and has obtained the data shown in the table below:

Class-interval	Frequency	Class-interval	Frequency
130-134	3	150-154	19
135-139	12	155-159	12
140-144	21	160-164	5
145-149	28	Total	100

Compute Karl Pearson's Coefficient of Skewness. [Coeff. of Sk = -0.572]

(MBA, Mangalore Univ., 1998)

38. (a) A moderately skewed distribution has mean and median as 25 and 26 respectively. Then its mode approximately equals.....

(b) Whether the following statement is true or false: If a distribution has negative skewness then its mean is greater than mode.

39. Calculate the first four moments about mean and find the values of β_1 and β_2 and comment on the result:

Profits (Rs. lakhs)	0-10	10-20	20-30	30-40	40-50	50-60	60-70
No. of Companies	8	12	20	30	15	10	5

(MBA, Kumaun Univ., 1990)

40. From the following data pertaining to the income of 5,800 persons, find Bowley's coefficient of skewness and interpret its value:

Income (Rs.)	No. of persons	Income (Rs.)	No. of persons
Below 10,000	170	40,000-50,000	1,350
10,000-20,000	630	50,000-60,000	1,000
20,000-30,000	1,900	60,000 and above	400
30,000-40,000	1,250		

[Coeff. of Sk = -0.067]

(MBA, Kurukshetra Univ., 2000)

Jewel's Care Collected

41. Compute the first 3 moments about the arithmetic mean from the following data :
- | | | | | | | | |
|------------------|---|----|----|----|----|----|----|
| Variable value : | 5 | 10 | 15 | 20 | 25 | 30 | 35 |
| Frequency : | 8 | 15 | 20 | 32 | 23 | 17 | 5 |
- (MBA, Lucknow Univ., 2001)

42. The following distribution gives the pattern of overtime work done in a month by 100 employees of a company :
- | | | | | | | |
|--------------------|-------|-------|-------|-------|-------|-------|
| Overtime hours : | 10-15 | 15-20 | 20-25 | 25-30 | 30-35 | 35-40 |
| No. of employees : | 11 | 20 | 35 | 20 | 8 | 6 |
- Compute mean, mode, standard deviation and coefficient of skewness.
[23.1, 22.5, 6.4915, 0.0924]
- (MBA, DU, 1994)

43. The following table gives the distribution of monthly wages of 500 workers in a factory:
- | Monthly wages
(Rs. hundred) | No. of
workers | Monthly wages
(Rs. hundred) | No. of
workers |
|--------------------------------|-------------------|--------------------------------|-------------------|
| 15-20 | 10 | 30-35 | 220 |
| 20-25 | 25 | 35-40 | 70 |
| 25-30 | 145 | 40-45 | 30 |
- Compute Karl Pearson's and Bowley's coefficient of skewness. Interpret your values.
[SK_p = -0.023, SK_b = -0.1022]
- (MBA, DU, 2002)

44. Calculate Karl Pearson's coefficient of skewness from the data given below :
- | Marks | No. of candidates | Marks | No. of candidates |
|-------|-------------------|-------|-------------------|
| 70-80 | 11 | 30-40 | 21 |
| 60-70 | 22 | 20-30 | 11 |
| 50-60 | 30 | 10-20 | 6 |
| 40-50 | 35 | 0-10 | 5 |
- [-0.026]
- (MBA, Kumaun Univ., 2001)

45. Calculate β_1 and β_2 from the following distribution and interpret the results :
- | Age | Frequency | Age | Frequency |
|-------|-----------|-------|-----------|
| 25-30 | 2 | 45-50 | 25 |
| 30-35 | 8 | 50-55 | 16 |
| 35-40 | 18 | 55-60 | 7 |
| 40-45 | 27 | 60-65 | 2 |
- [$\beta_1 = 0.034$, $\beta_2 = 2.59$]
- (MBA, Kumaun Univ., 2003)

Jewel's Care Collected

PROBLEMS

1-A: Answer the following questions, each question carries one mark:

- (i) What are the properties of correlation coefficient? (MBA, Madurai-Kamaraj, 2001)
- (ii) What are the limitations of correlation analysis? (MBA, Madurai-Kamaraj, April 2001)
- (iii) State the formula for coefficient of correlation in terms of regression coefficients. (MBA, Madurai-Kamaraj, April 2003)
- (iv) What is meant by correlation? (M.Com., M.K. Univ., Nov. 2002; MBA, Madurai-Kamaraj Nov. 2003)
- (v) What are the limits of coefficient of correlation? (M.Com., Madurai-Kamaraj, 2002)
- (vi) What is the use of scatter diagram?
- (vii) What is 'Rank correlation' ? (MBA, Madurai-Kamaraj, Nov. 2003)
- (viii) Write down the formula for rank correlation coefficient. (M.Com., M.K. Univ., 2001)
- (ix) Interpret the following value of r : $r = 0$, $r = -1$, $r = +1$, $r = 0.25$.
- (x) How can 'r' be determined through regression coefficients?

1-B: Answer the following questions. Each question carries four marks:

- (i) The coefficient of correlation between the variables x and y is 0.4, their covariance is 16. The variance of x is 9. Find the standard deviation of y . (M. Com., M.K. Univ., April 2002)
- (ii) Briefly explain the various types of correlation. (M.Com., M.K. Univ., Nov. 2002)
- (iii) What do you understand by correlation? Describe the uses of the study of correlation. (M.A. Eco., M.K. Univ., April 2003)
- (iv) Define correlation between two variables. What is the value of 'r' interpreted? (MBA, Madras Univ., Nov. 2003)
- (v) Does correlation always signify a cause and effect relationship between variables?
2. (a) Explain the meaning and significance of the term correlation. (MBA, Delhi Univ., 2000)
- (b) What is correlation? Clearly explain with suitable illustration its role in taking some business problem. (MBA, Delhi Univ., 2002)
3. Define the coefficient of correlation. What is it intended to measure? How would you interpret the sign and magnitude of a calculated r ? Consider particularly the values of $r = 0$, $r = +1$ and $r = -1$.
4. What is a scatter diagram? How does it help in studying the correlation between two variables, in respect of both its direction and degree? (MBA, Delhi Univ., 2002)
5. (a) What is Spearman's rank correlation coefficient? Bring out its usefulness. How does the coefficient differ from Karl Pearson's coefficient of correlation?
- (b) Explain briefly the different methods of measuring correlation.
6. (a) Does correlation always signify a cause and effect relationship between the variables? (MBA, Osmania, 2002)
- (b) Does a high positive correlation between the increase in cigarette smoking and the increase in lung cancer prove that one causes the other?
7. (a) Define correlation coefficient 'r' and give its limits. What interpretation would you give if told that the correlation between the number of truck accidents per year and the age of the driver is (-) 0.60 if only drivers with at least one accident are considered? (MBA, AMU, 1993)
- (b) What is a scatter diagram? How do you interpret a scatter diagram? (MPhil., Kurukshetra Univ., 1994)
8. (a) What is correlation? Does it always signify cause-and-effect relationship?
- (b) What is coefficient of Rank correlation? Bring out its usefulness. How does this coefficient differ from coefficient of correlation? (MBA, Delhi Univ., 1994, 2003)
9. (a) Prove that the correlation coefficient is unaffected by the change of origin and scale.
- (b) How is Scatter Diagram helpful in the study of correlation?
10. (a) Explain how covariance of X and Y is related to the coefficient of simple correlation between X and Y .
- (b) What is meant by correlation? Distinguish between positive, negative and zero correlation. (MBA, BIT, Ranchi Univ., 1992; MBA, DU, 1999)
- (c) Explain critically any two methods of measuring correlation. (MBA, Osmania, 1999)
11. Find Karl Pearson's coefficient of correlation from the following index numbers and interpret it:

Wages	:	100	101	103	102	104	99	97	98	96	96
Cost of living	:	98	99	99	97	95	02	95	94	90	91

$[r = 0.85]$

12. Find Karl Pearson's coefficient of correlation between capital employed and profit obtained from the following data.

Capital employed (Rs. crores)	Profits obtained (Rs. crores)	Capital employed (Rs. crores)	Profits obtained (Rs. crores)
10	2	60	15
20	4	70	14
30	8	80	20
40	5	90	22
50	10	100	50

[$r = 0.85$]

13. Using the following data :

- (a) Calculate the coefficient of correlation.
 (b) Estimate the percentage of the group with lung cancer in a country where 15 per cent of the group smoke heavily.

Country	% of group smoking heavily	% of group with lung cancer
A	10	5
B	20	15
C	20	20
D	30	25
E	30	20

[$r = 0.91$]

14. From the following data, calculate coefficient of correlation between the percentage yield on securities and wholesale price indices for certain years :

Year	1997	1998	1999	2000	2001	2002	2003
% Yield on securities	5.0	5.1	5.2	4.9	4.8	5.3	5.4
Index No. of wholesale prices	140	138	126	132	140	135	132

What inference do you draw from the result ?

[$r = -0.16$]

15. Find the correlation by Karl Pearson's method between the two kinds of assessment of postgraduate student's performance (marks out of 100) :

Roll No. of students	1	2	3	4	5	6	7	8	9	10
Internal assessment	45	62	67	32	12	38	47	67	42	85
External assessment	39	48	65	32	20	35	45	77	30	62

[$r = 0.88$]

16. Two housewives, Mrs. Neena and Mrs. Meena, asked to express their preferences for different kinds of detergents, gave the following replies :

Detergent	A	B	C	D	E	F	G	H	I	J
Neena	1	2	4	3	7	8	6	5	9	10
Meena	1	4	2	3	5	7	6	8	9	10

To what extent the preferences of these two ladies go together ?

[$R = +0.89$]

17. An office contains 10 clerks. The longer-serving clerks feel that they should have a seniority increment based on length of service built into their salary structure. An assessment of their efficiency by their departmental manager and the personnel department produces a ranking of efficiency. This is shown below together with a ranking of their length of service. Do the data support the clerk's claim for seniority increment ?

Ranking according to length of service	1	2	3	4	5	6	7	8	9	10
Ranking according to efficiency	2	5	3	10	6	4	8	9	7	1

[$R = +0.164$]

18. The following table gives the frequency, according to age groups, of marks obtained by 68 students in a general knowledge test. Measure the degree of relationship between age and general knowledge.

Test Marks	Age in Years			
	21	22	23	24
200 - 250	4	4	2	1
250 - 300	3	5	4	2
300 - 350	2	6	8	5
350 - 400	1	5	6	10

[$r = 0.415$]

Jewel's Care Collected

19. Find coefficient of correlation between output and cost per scooter from the following data :

Output of scooter (in '000s)	: 3.5	4.0	5.2	6.3	6.8	7.4	8.5	9.0
Cost per scooter (in '000 Rs.)	: 12.0	11.8	11.2	10.6	10.3	9.8	9.3	9.2

[$r=0.0996$]

20. Find the coefficient of correlation between price and sales from the following data :

Price (Rs.)	: 103	98	85	92	90	84	88	90	93	95
Sales (Units)	: 500	610	700	630	670	800	800	750	700	680

[$r=0.85$]

21. Calculate correlation coefficient from the following two-way table, with X representing the average salary of families selected at random in a given area and Y representing the average expenditure on entertainment (movies, magazines, etc.):

Expenditure on entertainment (in 00's Rs.)	Average salary (in 00's Rs.)				
	100-150	150-200	200-250	250-300	300-350
0-10	5	4	5	2	4
10-20	2	7	3	7	1
20-30	-	6	-	4	5
30-40	8	-	4	-	8
40-50	-	7	3	5	10

[$r=0.205$]

22. A psychologist wanted to compare two methods A and B of teaching. He selected a random sample of 22 students. He grouped them into 11 pairs so that the students in a pair have approximately equal scores on an intelligence test. In each pair, one student was taught by method A and the other by method B and examined after the course. The marks obtained by them are tabulated below :

Pair	: 1	2	3	4	5	6	7	8	9	10	11
A	: 24	29	19	14	30	19	27	30	20	28	11
B	: 37	35	16	26	23	27	19	20	16	11	21

(i) Find the correlation coefficient between the two sets of scores.

(MBA, HPU, 1999)

(ii) Find the rank correlation coefficient.

[$r=0.175$]

23. The mileage (Y) that can be obtained from a certain gasoline depends on the amount (X) of certain chemical in the gasoline. The value of ten observations, where X and Y are measured in appropriate units are shown in the table below :

Amount (X)	Mileage (Y)	Amount (X)	Mileage (Y)
0.10	10.98	0.60	14.63
0.20	11.14	0.70	15.66
0.30	13.17	0.80	13.71
0.40	13.34	0.90	15.43
0.50	14.39	1.00	18.36

Find the coefficient of correlation between X and Y and represent the data by a graph.

24. Calculate the coefficient of correlation between age and sum assured from the data given below and comment on the value :

Age	Sum assured (in lakh Rs.)				Total
	5	10	15	20	
20-30	5	3	4	6	15
30-40	2	3	3	5	10
40-50	-	2	2	3	7
50-60	-	2	2	2	18
Total	5	8	3	2	50
	7	15	12	16	

[$r=0.3442$]

(MBA, Delhi Univ., 1998)

Jewel's Care Collected

25. Compute the coefficient of correlation between dividends and prices of securities as given below :

Security Prices (in Rs.)	Annual Dividends (in hundred Rs.)					
	6-8	8-10	10-12	12-14	14-16	16-18
130-140	—	—	1	3	4	2
120-130	—	1	3	3	3	1
110-120	—	1	2	3	2	—
100-110	—	2	3	2	—	—
90-100	2	2	1	1	—	—
80-90	3	1	1	—	—	—
70-80	2	1	—	—	—	—

[$r = +0.71$]

26. The top executives of Sonal Electrical rank managerial candidates on the basis of what they know about each candidate. In order to determine if there is any consistency in the ranking obtained in this manner, two vice-presidents were asked to rank the same ten candidates. Compute the coefficient of rank correlation from the following two sets of ranks :

Candidate	A	B	C	D	E	F	G	H	I	J
V-P 1 :	3	1	8	1	4	9	5	7	10	6
V-P 2 :	2	5	9	1	6	10	3	4	8	7

[$R = +0.746$]

27. Seven methods of imparting business education were ranked by the MBA students of two universities as follows :

Method of teaching	I	II	III	IV	V	VI	VII
Rank by Students of Univ. A :	2	1	5	3	4	7	6
Rank by Students of Univ. B :	1	3	2	4	7	5	6

Calculate rank correlation coefficient and comment on its value.

[$R = +0.5$]

(MBA, Osmania Univ., MBA, South Gujarat Univ., 2002)

28. (a) Coefficient of correlation between X and Y for 20 items is 0.3, mean of X is 15 and that of $Y=20$, standard deviations are 4 and 5 respectively. At the time of calculation one item 26 was wrongly taken as 17 in case of X series and 35 instead of 30 in case of Y series. Find the correct value of correlation coefficient.

[Correct value of correlation coefficient is 0.504.]

(b) In order to find the correlation coefficient between two variables X and Y from 12 pairs of observations, the following calculations were made :

$$\Sigma X = 30, \Sigma Y = 5, \Sigma X^2 = 670, \Sigma Y^2 = 285, \Sigma XY = 334.$$

On subsequent verification it was found that the pair ($X = 11, Y = 4$) was copied wrongly, the correct value being ($X = 10, Y = 14$). Find the correct value of correlation coefficient.

[$r = 0.78$]

29. A Statistician while calculating the correlation coefficient between two variates X and Y from 25 pairs of observations obtained the following results :

$$n = 25, \Sigma X = 125, \Sigma X^2 = 650, \Sigma Y = 100, \Sigma Y^2 = 400, \Sigma XY = 508.$$

It was, however, later discovered at the time of checking that he had copied down the pairs as

X	6	8
Y	14	6

While the correct values were

X	8	6
Y	12	8

Obtain the correct value of the correlation coefficient.

[$r = 0.667$]

(MBA, Sukhadia Univ., 1995; MBA, Kumaun Univ., 2000; MBA, Anna Univ., 2002)

30. The following data relate to the prices and supplies of a commodity during a period of eight years :

Price (Rs./kg) :	10	12	18	16	15	19	18	17
Supply (100 kg) :	30	35	45	44	42	43	47	46

Calculate the coefficient of correlation between the two series.

[$r = -0.09$]

(MBA, Punjab Univ., 2002)

Jewel's Care Collected

31. Calculate the coefficient of correlation between family income and its percentage spent on food for the following data :

Family Income (in Rs.)	Food Expenditure (in percentage)				
	10-15	15-20	20-25	25-30	30-35
12000-13000	2	3	1	4	—
13000-14000	3	4	2	1	5
14000-15000	4	1	5	12	8
15000-16000	1	2	3	—	4
16000-17000	5	6	—	3	1

[$r = 0.1048$]

(MBA, Delhi Univ., 1993)

32. Calculate the coefficient of correlation and probable error of r between the values of X and Y given below :

X :	78	98	96	69	59	79	68	61
Y :	125	137	156	112	107	136	123	108

[$r = 0.955$, P.E.r. = 0.021]

(M.Com., Sukhadia Univ., 1996)

33. Find the coefficient of correlation for the following bivariate frequency distribution :

Marks in Physics	Marks in Mathematics						Total
	40-49	50-59	60-69	70-79	80-89	90-99	
90-99				2	4	4	10
80-89			1	4	6	5	16
70-79			5	10	8	1	24
60-69	1	4	9	5	2		21
50-59	3	6	6	2			17
40-49	3	5	4				12
Total	7	15	25	23	20	10	100

[$r = +0.769$]

(M. Com., M.D. Univ., 1997)

34. The bivariate frequency distribution based on monthly salary and age of 100 employees working in some large-scale commercial organisation is as under :

Age (Years)	Monthly Salary (in 000's Rs.)			
	8-10	10-12	12-14	14-16
20 and less than 30	16	6	—	—
30 and less than 40	4	10	4	4
40 and less than 50	—	4	18	12
50 and less than 60	—	—	10	12

Compute Karl Pearson's coefficient of correlation between age and monthly salary of employees and comment on its value.

[$r = +0.763$]

35. A survey regarding income and savings provided the following data :

Income (Rs.)	Saving (Rs.)			
	500	1000	1500	2000
4000	8	4	—	—
6000	—	12	24	6
8000	—	9	7	2
10000	—	—	10	5
12000	—	—	9	4

Compute Karl Pearson's coefficient of correlation and interpret its value.

(MBA, Kurukshetra Univ., 1997)

[$r = +0.522$]

36. Calculate the coefficient of correlation from the following data and interpret the value.

Advertising expenditure (Rs. lakhs) :	10	12	13	23	27	30
Sales turnover (Rs. crores) :	40	42	46	48	50	56

(MBA, Delhi Univ., 2002)

[$r = +0.956$]

37. You are given the following data of marks obtained by 11 students in statistics in two tests, one before and the other after special coaching :

First Test (Before coaching) :	23	20	19	21	18	20	18	17	23	16	19
Second Test (After coaching) :	24	19	22	18	20	22	20	20	23	20	17

(M.Com., Delhi Univ., 2000)

Do the marks indicate that the special coaching has benefited the students ?

[$r = +0.477$]

38. The scores of students in an examination in Mathematics and Statistics are given below :

Student No.	1	2	3	4	5	6	7	8
Marks in Mathematics	70	48	58	55	54	50	60	52
Marks in Statistics	62	47	53	60	55	68	51	48

Find : (i) Correlation coefficient, and
(ii) Rank correlation coefficient and compare the two values.

39. The following data show the marks of 10 students in Mathematics and Statistics in an examination :

Marks in Mathematics	45	70	65	30	90	40	50	75	85	60
Marks in Statistics	35	90	70	40	95	40	60	80	80	50

Find Karl Pearson's coefficient of correlation and its probable error.

(MBA, Vikram Univ., 2001)

40. A researcher collected the following information for two variables x and y :

No. of Pairs = 20, $r = 0.5$, $\bar{x} = 15$, $\bar{y} = 20$, $\sigma_x = 4$, $\sigma_y = 5$

Later it was found that one pair of value has been wrongly taken as $\frac{x}{16}$, $\frac{y}{30}$ whereas the correct values were $\frac{x}{26}$, $\frac{y}{35}$. Find the correct value of r .

[$r = 0.559$]

(MBA, MD Univ., 2001)

41. Calculate the Karl Pearson's Coefficient of Correlation between age and playing habits from the data given below. Comment on the value :

Age	20	21	22	23	24	25
No. of students	500	400	300	240	200	160
Regular players	400	300	180	96	60	24

(MBA, Osmania Univ., 1990)

[$r = -0.991$]

42. The following bivariate frequency distribution relates to the age and salary of 100 computer operators working in an organisation. Find the coefficient of correlation and interpret its value.

Age (Yrs)	Salary (Rs.)			
	5000 - 6000	6000 - 7000	7000 - 8000	8000 - 9000
20 - 30	4	6	5	2
30 - 40	2	5	8	3
40 - 50	8	12	20	2
50 - 60	—	8	12	1

[$r = 0.057$]

Jewel's Care Collected

(MBA, Delhi Univ., 1991)

43. Compute the rank correlation coefficient from the following data :

Series X	115	109	112	87	98	98	120	100	98	118
Series Y	75	73	85	70	76	65	82	73	68	60

(MBA, KU, 2001)

44. From the following data calculate coefficient of correlation between age and playing habit. How do you interpret the result?

Age group	No. of Employees	No. of regular players
20 - 30	50	20
30 - 40	120	60
40 - 50	80	24
50 - 60	40	4
60 - 70	20	1

(MBA, Guru Jambheshwar Univ., 2001)

45. Calculate the coefficient of correlation from the following data :

x	65	66	67	67	68	69	70	72
y	67	68	65	68	72	72	69	71

46. Calculate the coefficient of correlation from the following data :

X	100	200	300	400	500	600	700
Y	30	50	60	80	100	110	130

(M. Com., Madurai Kamaraj, 2001)

4. Calculate the coefficient of correlation for the following :

X :	5	10	5	11	12	4	3	2	7	1
Y :	1	6	2	8	5	1	4	6	5	2

5. Two designs A and B gave the following output in 9 trials of each. Which is a better design ? Why ?

A :	16	16	53	15	31	17	14	30	20
B :	18	27	23	21	22	26	39	17	28

6. Calculate Pearson's coefficient of correlation from the following taking 100 and 50 as the assumed average of X and Y respectively :

X :	104	111	104	114	118	117	105	108	106	100	104	105
Y :	97	55	47	45	45	50	64	63	66	62	69	61

7. Find the correlation coefficient from the following data :

X :	53	25	19	37	42	10	15
Y :	9	6	5	7	7	4	5

8. The marking of 10 trainees in two skills, programming and analysis are as follows. What is the coefficient of rank correlation?

Programming :	3	5	8	4	7	10	2	1	6	9
Analysis :	6	4	9	8	1	2	3	10	5	7

(MBA, Bharathidasan Univ., 2004)

Jewel's Care Collected

Illustration 26. In trying to evaluate the effectiveness in its advertising campaign, a firm compiled the following information:

Year	1996	1997	1998	1999	2000	2001	2002	2003
Adv. Expenditure ('000 Rs.)	12	15	15	23	24	38	42	48
Sales (Lakh Rs.)	5.0	5.6	5.8	7.0	7.2	8.8	9.2	9.5

Calculate the regression equation of sales on advertising expenditure. Estimate the probable sales when advertising expenditure is Rs. 60 thousand. (MBA, Bharathidasan Univ., Nov. 2001)

Solution :

CALCULATION OF REGRESSION EQUATION

X	(X-24)	d_x^2	Y	(Y-7.0)	d_y^2	$d_x d_y$
12	-12	144	5.0	-2.0	4.00	24.0
15	-9	81	5.6	-1.4	1.96	12.6
15	-9	81	5.8	-1.2	1.44	10.8
23	-1	1	7.0	0	0	0
24	0	0	7.2	+0.2	.04	0
38	+14	196	8.8	+1.8	3.24	25.2
42	+18	324	9.2	+2.2	4.84	39.6
48	+24	576	9.5	+2.5	6.25	60.0
$\Sigma X = 217$	$\Sigma d_x = 25$	$\Sigma d_x^2 = 1403$	$\Sigma Y = 58.1$	$\Sigma d_y = 2.1$	$\Sigma d_y^2 = 21.77$	$\Sigma d_x d_y = 172.2$

$$\bar{X} = \frac{\Sigma X}{N} = \frac{217}{8} = 27.125; \bar{Y} = \frac{\Sigma Y}{N} = \frac{58.1}{8} = 7.26$$

Regression equation of sales on advertisement exp. is given by :

$$(Y - \bar{Y}) = b_{yx}(X - \bar{X})$$

where

$$b_{yx} = \frac{N \Sigma d_x d_y - (\Sigma d_x)(\Sigma d_y)}{N \Sigma d_x^2 - (\Sigma d_x)^2}$$

$$= \frac{8(172.2) - (25)(2.1)}{8(1403) - (25)^2}$$

$$= \frac{1377.6 - 52.5}{11224 - 625} = \frac{1325.1}{10599} = 0.125$$

Substituting the values, we have

$$Y - 7.2625 = 0.125(X - 27.125)$$

$$Y - 7.2625 = 0.125X - 3.3906$$

$$Y = 3.8719 + 0.1250X$$

When X = 60, the estimated value of Y shall be :

$$Y = 3.8719 + 0.1250(60) = 3.8719 + 7.5 \approx 11.37.$$

Jewel's Care Collected

PROBLEMS

1-A : Answer the following questions, each question carries one mark:

- (i) What is regression ?
- (ii) What is the use of studying regression ?
- (iii) When will regression coefficients become coefficient of correlation ?
- (iv) Write down the two regression equations.
- (v) Write down the formula for regression coefficient of x and y ?
- (vi) What do you understand by the term 'regression line' ?
- (vii) What are regression coefficients ?
- (viii) Can both the regression coefficients exceed one ?
- (ix) Are regression coefficients independent of change of scale and origin or only origin ?
- (x) In the regression equation of y on x how do you interpret the values of 'a' and 'b' ?
- (xi) Who had coined the term 'regression' ?

(MBA, Madurai-Kamaraj, 2001)
 (MBA, Madurai-Kamaraj, Nov. 2003)
 (M.Com., Madurai-Kamaraj, 2002)
 (M.Com., M.K. Univ., 2001)
 (M.Com., M.K. Univ., Nov. 2002)

- Answer the following questions, each question carries four marks:
- (i) Distinguish between 'correlation' and 'regression analysis'. (M. Com., M.K. Univ., Nov. 2002)
 - (ii) What are regression coefficients? How do you interpret them?
 - (iii) What are the important characteristics of regression coefficients?
 - (iv) If two regression coefficients are -1.2 and -0.8 , what would be the value of r ?
 - (v) What are the important uses of regression analysis?

2. (a) Explain the concept of regression and point out its usefulness in dealing with business problems. (MBA, Delhi Univ.)
- (b) Distinguish between correlation and regression. Also point out the properties of regression coefficients. (MBA, Sukhadia Univ., 1995)

3. (a) Compare and contrast the role of correlation and regression in studying the interdependence of two variates.
- (b) Explain the concept of regression and point out its importance in business forecasting. (MBA, Delhi Univ., 1998)

4. Under what conditions can there be one regression line? Explain. (MBA, HPU, 1996)

5. "The regression line gives only the best estimate of the value of quantity in question. We may assess the degree of uncertainty in this estimate by calculating a quantity known as the standard error of estimate". Elucidate.

6. Do you agree with the view that regression equations are irreversible, i.e., we cannot find out the regression of X on Y from that of Y on X ?

7. (a) Point out the usefulness of regression analysis in business and industry.
- (b) What is linear regression? When is it used? (MBA, Madurai-Kamaraj Univ., 1994)
- (c) Discuss the role of correlation and regression analysis in business. Illustrate. (MBA, Bharathidasan Univ., 2002)

8. What are regression lines? With the help of an example, illustrate how they help in business decision-making. (MBA, Delhi Univ., 1998)

9. Point out the role of regression analysis in business decision-making. What are the important properties of regression coefficients? (MBA, Osmania Univ.; MBA, D.U., 1999)

10. (a) Distinguish between correlation and regression analysis and point out their role in business. (MBA, Delhi Univ.; MBA, HPU, 1998)

(b) What are regression coefficients? State some of the important properties of regression coefficients. (Diploma in Mgt., AIMA, MBA, Osmania)

(c) What is regression? How is this concept useful to business forecasting? (MBA, Jodhpur Univ., 1999)

(d) State the utility of regression in economic analysis. (MBA, Madras Univ., 2002)

11. The following data give the hardness (X) and tensile strength (Y) of 7 samples of metal in certain units. Find the linear regression equation of Y on X .

X :	146	152	158	164	170	176	182
Y :	75	78	77	89	82	85	86

$[Y = 29.45 + 0.31X]$

12. The average daily wage for working class in Nagpur is Rs. 12 and for that in Delhi Rs. 18, their respective standard deviations are Rs. 2 and Rs. 3 and the coefficient of correlation is 0.67. Find the most likely wage in Delhi corresponding to the wage of Rs. 20 in Nagpur. (MBA, Delhi Univ., 1998)

$[Y_{20} = 26.04]$

13. There are two series of index numbers D for disposable personal income and S for a salary of the company. The mean and standard deviations of the D series are 120 and 15 respectively and of the S series 115 and 10. The coefficient of correlation between the two series is 0.75. From the given information obtain a linear equation for estimating the values of S for different values of D . How will you interpret the values of S corresponding to different values of D obtained from the equation? Can the same equation be used for estimating values of D for different values of S ? (MBA, Delhi Univ., 1998)

$[S = 0.5; D = 55; \text{No}]$

14. The following calculations have been made for closing prices of 12 stocks (X) on the Bombay Stock Exchange on a certain day along with the volume of sales in thousand of shares (Y). From these calculations find the regression equations.

$\Sigma X = 580, \quad \Sigma Y = 370, \quad \Sigma XY = 11,494$
 $\Sigma X^2 = 41,658, \quad \Sigma Y^2 = 17,206$

$[Y = 53.55 - 0.47X, X = 79.16 - 1.1Y]$

15. Given the following data, what will be the possible yield when the rainfall is 29" ?

	Rainfall	Production
Mean	29"	40 units per acre
S.D.	3"	6 units per acre

Coefficient of correlation between rainfall and production = 0.8.
[40 units]

16. In the following table are recorded data showing the test scores made by salesmen on an intelligence test and their weekly sales:

Salesmen	1	2	3	4	5	6	7	8	9	10
Test Scores	45	75	50	60	80	90	85	40	80	55
Sales ('000)	2.0	6.5	3.5	5.0	4.5	6.0	6.5	2.5	5.5	4.5

Calculate the regression line of sales on test score and estimate the most probable weekly sales volume if a salesman has a score of 70.

$$[Y = -0.541 + 0.078X, 4919]$$

17. The following marks have been obtained by a group of students in Statistics (out of 100):

Paper I :	80	45	55	56	58	60	65	68	70	75	85
Paper II :	82	56	50	48	60	62	64	65	70	74	90

Compute the coefficient of correlation for the above data. Find the lines of regression and examine the relationship.

$$[r = 0.75, Y = -1 + 0.75X, X = 4.25 + 0.75Y]$$

18. The following table gives marks out of 50 awarded in a French and a German test to the same group of boys. Assume there is a linear relation between the sets of marks, calculate the equations of the lines of regression.

French :	10	10	18	25	28	33	34	39	42	43
German :	11	22	22	19	35	27	33	40	42	47

$$[Y = 6.25 + 0.13X, X = -0.34 + 0.96Y]$$

19. You are given the following result of the height (X) and weight (Y) of 1,000 managers:

Mean (X)	= 68.00"
Mean (Y)	= 150 lbs
Standard deviation (X)	= 2.50"
Standard deviation (Y)	= 20 lbs

Coefficient of correlation between X and Y = 0.6. Estimate from the above data the height of a manager whose weight is 200 lbs. (MBA, Kurukshetra Univ. 2001)

20. The following table shows the mean and standard deviation of the prices of two shares on a stock exchange:

Share	Mean (in Rs.)	Standard deviation (in Rs.)
A Ltd.	39.5	10.8
B Ltd.	47.5	16.8

If the coefficient of correlation between the prices of two shares is 0.42, find the most likely price of share A corresponding to a price of Rs. 55 observed in the case of share B. (BBE, M.K. Univ. 2001)

21. Catalogues listing textbooks were examined to discover the relationship between the cost of a book and number of pages it contains. The perusal gives the following data for ten books:

Pages :	700	540	210	625	380	910	610	420	750	400
Price (Rs.) :	12	11	5	10	7	15	9	8	12	9

- Obtain the line of regression for estimating the price of a book.
- What is your estimate for the price of a book containing 500 pages ?
- What increase would you expect for a book if it is decided to increase the number of pages of the book by 100 ?
- Calculate the standard error of the estimate.

22. From the data given below find the two regression equations.

Age of wife	Age of Husband			Total
	20-25	25-30	30-35	
16-20	4	9	—	13
20-24	1	4	—	6
24-28	4	4	1	11
Total	9	17	3	30

(M.Phil, Kurukshetra Univ. 2001)

23. The data given below relate to the scores obtained by 9 salesmen in an intelligence test and their weekly sales, in lakhs of rupees :

Salesman	1	2	3	4	5	6	7	8	9
Test Score	50	60	50	60	80	50	80	40	70
Sales (Rs. lakhs)	3	6	4	5	6	3	7	5	6

Obtain regression equation of sales on the intelligence test scores. If a salesman has obtained a score of 65, what would be his expected weekly sales ?

$[Y = .075X + 0.5, \text{Rs. } 5.375 \text{ lakhs}]$

24. The following figures relate to advertisement expenditure and sales :

Adv. Exp. (in lakh of Rs.)	60	62	65	70	73	75	71
Sales (in crores of Rs.)	10	11	13	15	16	19	14

Estimate (i) the sales for advertisement expenditure of Rs. 80 lakhs and (ii) the advertisement for a sales target of Rs. 25 crores.

$[20.1; 87.75]$

25. You are given the following data about the sales and advertisement expenditure of a firm :

	Sales (Rs. crores)	Advertisement Expenditure (Rs. crores)
Arithmetic Mean	50 \bar{x}	10 \bar{y}
Standard Deviation	10 σ_x	2 σ_y
Coefficient of Correlation	$+0.9 R$	

- (a) Calculate the two regression equations.
 - (b) Estimate the likely sales for a proposed advertisement expenditure of Rs. 13.5 crores.
 - (c) What should be the advertisement budget if the company wants to achieve a sales target of Rs. 70 crores ?
- (MBA, DU, 1998)

$[(a) Y = 4.5X + 5, X = .18Y + 1. (b) 65.75 \text{ crores. (c) } 13.6 \text{ crores}]$

26. The following bivariate frequency distribution relates to sales turnover (in lakh Rs.) and money spent on advertising budget (in thousand Rs.). Obtain the two regression equations.

Sales Turnover (in lakh Rs.)	Advertising budget (in thousand Rs.)			
	50-60	60-70	70-80	80-90
25-50	2	1	2	5
50-75	3	4	7	6
75-100	1	5	8	6
100-125	2	7	9	2

Estimate (i) the sales turnover corresponding to advertising budget of Rs. 150 thousand, (ii) the advertising budget to achieve a sales turnover of Rs. 200 lakhs, and (iii) compute the coefficient of correlation.

27. The following data give the test scores and sales made by nine salesmen during the last one year :

Test Scores	14	19	24	21	26	22	15	20	19
Sales ('000 Rs.)	31	36	48	37	50	45	33	41	39

Obtain (i) the regression equation of test scores on sales, (ii) the regression equation of sales on test scores, and (iii) coefficient of correlation.

$[X = -2.312 + 0.5578 Y, (ii) Y = 7.834 + 1.6083 X, (iii) r = 0.947]$

28. A study of share prices of Textile group and Fertiliser group of companies yielded the following results :

	Textiles	Fertilisers
Mean	12.8	985.0
Standard Deviation	1.6	70.1
Coefficient of Correlation	$+0.52$	

The financial expert has estimated the likely price of textiles shares at the close of the next accounting year as 92. What would be your estimate of the likely price of fertiliser shares at the corresponding time ?

29. Following are the data on business turnover and staff of a company for eight years from 1996 to 2003 :

	1996	1997	1998	1999	2000	2001	2002	2003
Business Turnover (Rs. crores)	45	50	60	75	80	110	150	170
Staff	2,600	3,000	3,100	3,530	3,850	4,300	5,870	7,150

Jewell's Date Collected

Fit a proper regression equation to estimate manpower in terms of business turnover. Estimate the staff requirement when the business turnover reaches Rs. 200 crores.

$[Y = 33.24X + 1100.3; 7748.3]$

30. The data on sales and promotion expenditure on a product for 10 years are given below :

Sales (Rs. lakhs) :	8	10	9	12	10	11	12	13	14	15
Promotion Exp. (Rs. thousands) :	2	2	3	4	5	5	5	6	7	8

Use two-variable regression model to estimate the effect of promotion on sales. Forecast the sales for next year when the company hopes to spend Rs. 10 thousand on promotion.

$[X = 0.815Y - 4.591; Y = 1.003X + 6.686; Y_{10} = 16.716]$

31. Table below shows the power and top speeds of different brands of sports cars :

Brand	A	B	C	D	E	F
Power X [kW]	70	63	72	60	66	70
Speed Y [km/h]	155	150	180	135	156	168
Brand	G	H	I	J	K	L
Power X [kW]	74	65	62	67	65	68
Speed Y [km/h]	178	160	132	145	139	152

- (i) Find the best linear relationship that fits the given data.
- (ii) Estimate the speed of a car that has a power of 63 kW and find a 95% confidence interval for this estimate.
- (iii) Determine how much of the variability in speed may be explained by the regression hypothesis.

32. Calculate the coefficient of correlation from the following data :

X :	1	2	3	4	5	6	7	8	9
Y :	9	8	10	12	11	13	14	16	15

Also obtain the regression equations and find an estimate of Y which should correspond on an average to X=6.2.

$[Y = .95X + 7.25; Y_{6.2} = 13.14]$

(MBA, Madurai-Kamaraj Univ., 1998)

33. Family income and its percentage spent on food gave the following bivariate frequency table :

Food Expenditure (in%)	Monthly Family Income (in hundred Rs.)				
	25-35	35-45	45-55	55-65	65-70
15-20	8	9	12	13	8
20-25	6	3	6	11	14
25-30	—	7	9	—	4
30-35	5	8	10	14	13

- (i) Estimate the family income for a food expenditure of 40%.
- (ii) What amount should be spent on food expenditure for a monthly family income of Rs. 10,000.
- (iii) Compute coefficient of correlation.

34. You are given below the following information about advertisement and sales.

	Adv. Exp. (X) (Rs. crores)	Sales (Y) (Rs. crores)
Mean	20	120
S.D.	5	25
Correlation coefficient	+0.8	

- (i) Calculate the two regression equations.
- (ii) Find the likely sales when advertisement expenditure is Rs. 25 crores.
- (iii) What should be the advertisement budget if the company wants to attain sales target of Rs. 150 crores?

$[Y = 4X + 40; X = .16Y + 0.8; Y_{25} = 140; X_{150} = 24.8]$

(MBA, DU)

35. From the following data obtain the regression equation. Also find the correlation coefficient with the help of regression coefficient :

X :	6	2	10	4	8
Y :	9	11	5	8	7

$[Y = 11.9 - 0.65X; X = 16.4 - 1.3Y, r = -0.919]$

(MBA, Kurukshetra Univ., MBA, Vikram Univ.)

Jewel's Care Collected

36. The monthly expenditure on advertisement and sales of a firm are given for 2003. It is generally found that expenditure on advertisement has its impact after two months. Allowing for time lag:
- calculate the correlation between expenditure on advertisement and sales.
 - estimate the sales of the firm in February 2005.

Months/Year (2003)	Expenditure on Advertisement (Rs.)	Sales (Rs.)
January	50	1200
February	60	1500
March	70	1600
April	90	2000
May	120	2200
June	150	2500
July	140	2400
August	160	2600
September	170	2800
October	190	2900
November	200	3100
December	250	3900

Jewel's Care Collected

37. The following figures relate to advertisement expenditure and sales :

Advertisement (in Rs. lakhs) :	60	62	65	70	73	75	71
Sales (in Rs. crores) :	10	11	13	15	16	19	14

Estimate (i) the sales for advertisement expenditure of Rs. 80 lakhs; and (ii) the advertisement expenditure for a sales target of Rs. 25 crores.

38. Given the regression equation of Y on X and X on Y are respectively $Y = 2X$ and $6X - Y = 4$ and the second moment of X about the origin is 3. Find (i) the correlation coefficient, and (ii) standard deviation of Y .

39. Find the regression coefficient of Y on X from the following regression equations :

$$5X = 22 + Y$$

$$64X = 24 + 45Y$$

Is it possible to calculate the standard deviation of Y from the given information? Answer with reason.

40. A financial analyst has gathered the following data about the relationship between income and investment in securities in respect of 8 randomly selected families :

Income (Rs. '000) :	8	12	9	24	143	37	19	16
Per cent invested in securities :	36	25	33	15	28	19	20	22

- Develop an estimating equation that best describes these data.
- Find the coefficient of determination and interpret it.
- Calculate the standard error of estimate for this relationship.
- Find an approximate 90 per cent confidence interval for the percentage of income invested in securities by a family earning Rs. 25,000 annually. (MFC, Delhi Univ., 1997)

41. From the data given below find :

- The two regression equations.
- The coefficient of correlation between marks in Economics and Statistics.
- The most likely marks in Statistics when the marks in Economics are 30.

Marks in Economics (X) :	25	28	35	32	31	36	29	38	34	32
Marks in Statistics (Y) :	43	46	49	41	36	32	31	30	33	39

42. A financial analyst obtained the following information relating to return on security A and that of market portfolio M for the past 8 years :

Year :	1	2	3	4	5	6	7	8
Return on A :	10	15	18	14	16	16	18	4
Return on M :	12	14	13	10	9	13	14	7

- Develop an estimating equation that best describes these data.
- Find the coefficient of determination and interpret it.
- Determine the percentage of total variation in security return being explained by the return on the market portfolio. (MFC, Delhi Univ., 1998)

43. Given the bivariate data :

X	1	5	3	2	1	1	7	3
Y	6	1	0	0	1	2	1	5

(i) Fit a regression equation of Y on X.

(ii) If a person has scored 8 on X variable, what would be his score on Y variable ?

(MBA, Osmania, 1994)

44. Personnel Manager of a large industrial unit is interested to find a measure that can be used to fix the wages (yearly) of skilled workers. On experimental basis, the data on the length of service and their yearly wages (in Rs. '000) from a group of 10 randomly selected skilled workers are given below :

Length of service (X)	11	7	9	5	8	6	10	12	3	4
Yearly wages (Y)	14	11	10	9	13	10	14	16	6	7

(a) Develop the regression equation of wage (Y) on the length of service X.

(b) On the basis of (a) what initial pay the personnel manager should give to a skilled worker who has put in thirteen years of service on a similar basis, in another industry.

$[Y = 3.455 + 1.006 X; Y = 16.533]$

(DJM, IGNOU, 2002)

45. In a laboratory experiment on correlation research study, the equation to the two regression lines were $2X - Y + 1 = 0$ and $3X - 2Y + 7 = 0$. Find (i) the means of X and Y. Also work out the values of the regression coefficients and coefficient of correlation between the two variables X and Y.

$[\bar{X} = 5, \bar{Y} = 11; b_{xy} = 0.5, b_{yx} = 1.5; r = 0.866]$

46. An industrial engineer collected the following data on experience & performance rating of 8 operators :

Operators	1	2	3	4	5	6	7	8
Experience (years)	16	12	18	4	3	10	5	12
Performance Rating	87	88	89	68	58	80	70	85

(a) Does the data give evidence that experience improves performance ?

(b) Estimate the performance rating of an operator having (a) 9 years and (b) 15 years of experience.

$[Y = 69.67 + 1.133 X]$

(MBA, M.D. Univ., 1994; MBA, Kumaun Univ., 2002)

47. The following table gives the age of cars of certain make and the annual maintenance costs. Find (i) the coefficient of correlation between the variables and (ii) Regression equation for costs related to age.

Age of Cars (in years)	2	4	6	8
Maintenance costs (in hundred Rs.)	0	20	25	30

(MBA, HPU, 2002)

48. A firm administers a test to sales trainees before they go into the field. The management of the firm is interested in determining the relationship between the test scores and the sales made by the trainees at the end of one year in the field. The following data were collected for ten sales personnel who have been in the field for one year :

Sales Person Number	Test Score	Number of Units Sold
1	2.6	95
2	3.7	140
3	2.4	85
4	4.5	180
5	2.6	100
6	5.0	195
7	2.8	115
8	3.0	136
9	4.0	175
10	3.4	150

(i) Find the regression line which would be used to predict sales from trainees test scores.

(ii) Predict the number of units which would be sold by trainee who received an average test score. (MBA, DU, 2002)

49. For the data given below :

Production (in units)	Average	S.D.
Capacity utilisation (%)	35	10
	85	8

Coefficient of correlation = 0.6

Obtain the two regression equations.

Jewel's Care Collected

- Estimate the production when the capacity utilisation is 70 per cent. (MBA, D.U., 2003)
50. Explain why there are two regression lines? What happens if the two lines are identical? For the data given below, find the relevant line of regression to estimate the price, if supply is 25 million tonnes.

Supply (m.t.)	:	5	10	12	15	18
Price (Rs./kg.)	:	16	15	12	12	10

(M. Com., A.M.U., 2001)

51. The following table shows the ages (X) and blood pressure (Y) of 8 persons :

X	:	52	63	45	36	72	65	47	25
Y	:	62	53	51	25	79	43	60	33

Obtain the regression equation of Y on X and find out the expected blood pressure of a person who is 49 years old.

(M. Com., Madurai-Kamaraj, 2002)

52. Determine the equation of the straight line which best fits the following data :

X	Y
10	19
12	22
13	24
16	27
17	29
20	33
25	37

53. Regression calculations were carried out as follows :

$$\sum X = 32, \sum Y = 24, \sum XY = 218$$

$$\sum X^2 = 296, \sum Y^2 = 162.5, n = 4$$

Find the lines of regression and coefficient of correlation and comment.

(MBA, M.D. Univ., 2000)

54. From the following data obtain the two regression equations :

Sales	:	91	97	103	121	67	124	52	73	111	57
Purchases	:	97	75	69	97	70	91	39	61	83	47

(MBA, Madurai Kamaraj Univ., Nov. 2001)

55. Obtain the regression of Y on X and X on Y from the following data and estimate the blood pressure when the age is 50.

Age	Blood Pressure	Age	Blood Pressure
50	147	55	150
42	125	49	145
72	160	38	115
36	118	42	140
63	149	68	150
47	128	60	155

(MBA, Bharathidasan Univ., 2001)

56. From the data given below, find the two regression equations and the most likely marks in statistics when marks in Economics are 30.

Marks in Economics	:	25	28	35	32	31	36	24	38	34	32
Marks in Statistics	:	43	46	49	41	36	32	31	30	33	39

(MBA, M.K. Univ., 2003)

57. Cost accountants often estimate overheads based on the level of production. At BFL company, the data collected are as follows. Find the best fit equation between production and overhead costs. Predict overheads when 50 units are produced.

Overhead	:	191	170	272	155	280	173	234	116	153	178
Production units	:	40	42	53	35	56	39	48	30	37	40

(MBA, Bharathidasan Univ., April, 2003)

Jewel's Care Collected

(MBA., IGNOU, Dec. 2001)

Illustration 31. Calculate from the following data, the Fisher's Ideal Index Number for the year 2005 :

Commodity Selected	2004		2005	
	Price (Rs.)	Expenditure on quantity consumed (Rs.)	Price (Rs.)	Expenditure on quantity consumed (Rs.)
A	8	200	65	1950
B	20	1400	30	1650
C	5	80	20	900
D	10	360	15	300
E	27	2160	10	600

Solution : First find quantity by dividing expenditure by price.

CALCULATION OF FISHER'S IDEAL INDEX

Commodity	P_0	q_0	P_1	q_1	P_1q_0	P_0q_1	P_1q_1	P_0q_1
A	8	25	65	30	1625	200	1950	240
B	20	70	30	55	2100	1400	1650	1100
C	5	16	20	45	320	80	900	225
D	10	36	15	20	540	360	300	200
E	27	80	10	60	800	2160	600	1620
					$\Sigma P_1q_0 = 5385$	$\Sigma P_0q_0 = 4200$	$\Sigma P_1q_1 = 5400$	$\Sigma P_0q_1 = 3385$

$$P_{01} = \sqrt{\frac{\Sigma P_1q_0 \times \Sigma P_1q_1}{\Sigma P_0q_0 \times \Sigma P_0q_1}} \times 100$$

$$= \sqrt{\frac{5385 \times 5400}{4200 \times 3385}} \times 100 = 1.430 \times 100 = 143.$$

Illustration 32. Construct Fisher's Ideal Index from the following data and show that it satisfies time reversal and factor reversal test:

Commodity	2004		2005	
	Price	Value	Price	Value
A	10	100	12	144
B	15	75	20	120
C	8	80	10	110
D	20	60	25	50
E	50	500	60	540

CALCULATION OF FISHER'S IDEAL INDEX

Commodity	P_0	q_0	P_1	q_1	P_1q_0	P_0q_1	P_1q_1	P_0q_1
A	10	10	12	12	120	100	144	120
B	15	5	20	6	100	75	120	90
C	8	10	10	11	100	80	110	88
D	20	3	25	2	75	60	50	40
E	50	10	60	9	600	500	540	450
					$\Sigma P_1q_0 = 995$	$\Sigma P_0q_0 = 815$	$\Sigma P_1q_1 = 964$	$\Sigma P_0q_1 = 788$

Fisher's Ideal Index or

$$P_{01} = \sqrt{\frac{\Sigma P_1q_0 \times \Sigma P_1q_1}{\Sigma P_0q_0 \times \Sigma P_0q_1}} \times 100$$

$$= \sqrt{\frac{995 \times 964}{815 \times 788}} \times 100 = \sqrt{1.4935} \times 100 = 1.222 \times 100 = 122.2$$

Time Reversal Test : Time reversal test is satisfied when :

$$P_{01} \times P_{10} = 1$$

$$P_{10} = \sqrt{\frac{\sum p_1 q_0}{\sum p_1 q_1} \times \frac{\sum p_0 q_0}{\sum p_0 q_1}} = \sqrt{\frac{788}{964} \times \frac{815}{995}}$$

$$P_{01} \times P_{10} = \sqrt{\frac{995}{815} \times \frac{964}{788} \times \frac{788}{964} \times \frac{815}{995}} = \sqrt{1} = 1$$

Hence time reversal test is satisfied.

Factor Reversal Test : Factor reversal test is satisfied when :

$$P_{01} \times Q_{01} = \frac{\sum p_1 q_1}{\sum p_0 q_0}$$

$$Q_{01} = \sqrt{\frac{\sum q_1 p_0}{\sum p_0 q_0} \times \frac{\sum q_1 p_1}{\sum q_0 p_1}} = \sqrt{\frac{788}{815} \times \frac{964}{995}}$$

$$P_{01} \times Q_{01} = \sqrt{\frac{995}{815} \times \frac{964}{788} \times \frac{788}{815} \times \frac{964}{995}} = \frac{964}{815}$$

$\frac{\sum p_1 q_1}{\sum p_0 q_0}$ is also equal to $\frac{964}{815}$. Hence factor reversal test is satisfied by the given data.

Illustration 33. From the data given below, calculate Fisher's Ideal Index and show that it satisfies time reversal test.

Commodity	2004		2005	
	Price	Quantity	Price	Quantity
A	12	20	14	30
B	14	13	20	15
C	10	12	15	20
D	6	8	4	10
E	8	5	6	5

Solution.

CALCULATION OF FISHER'S IDEAL INDEX

Commodity	2004		2005		$p_1 q_0$	$p_0 q_0$	$p_1 q_1$	$p_0 q_1$
	p_0	q_0	p_1	q_1				
A	12	20	14	30	280	240	420	360
B	14	13	20	15	260	182	300	210
C	10	12	15	20	180	120	300	200
D	6	8	4	10	32	48	40	60
E	8	5	6	5	30	40	30	60
					$\sum p_1 q_0 = 782$	$\sum p_0 q_0 = 630$	$\sum p_1 q_1 = 1090$	$\sum p_0 q_1 = 870$

Fisher's Ideal Index or $P_{01} = \sqrt{\frac{\sum p_1 q_0}{\sum p_0 q_0} \times \frac{\sum p_1 q_1}{\sum p_0 q_1}} \times 100$

$$= \sqrt{\frac{782}{630} \times \frac{1090}{870}} \times 100 = 1.247 \times 100 = 124.7$$

Time Reversal Test : Time Reversal Test is satisfied when

$$P_{01} \times P_{10} = 1$$

$$P_{10} = \sqrt{\frac{\sum p_0 q_1}{\sum p_1 q_1} \times \frac{\sum p_0 q_0}{\sum p_1 q_0}} = \sqrt{\frac{870}{1090} \times \frac{630}{782}}$$

PROBLEMS

1-A : Answer the following questions, each question carries one mark:

(MBA, Madurai-Kamraj, 2000)

- (i) What is index number ?
- (ii) Give two important uses of index numbers.
- (iii) Name two important problems that arise while constructing index number.
- (iv) Give the formula for Fisher's Ideal Index Number.
- (v) What is time reversal test ?
- (vi) What is Laspeyres method of constructing index numbers.
- (vii) What is quantity index ?
- (viii) What is base shifting ?
- (ix) Which average is most appropriate for constructing index numbers.
- (x) Give any two limitations of index numbers.

1-B : Answer the following questions, each question carries four marks:

(M.A. Econ., Madras Univ., 2000)

- (i) Describe the problem faced in the construction of index numbers.
 - (ii) Differentiate between Time reversal test and Factor reversal test.
 - (iii) What is Fisher's Ideal Index ? Why is it called ideal ?
 - (iv) Briefly explain the concept of splicing and deflating.
 - (v) What are fixed base and chain base indices ? Explain with suitable example.
2. (a) What is an index number ? Describe briefly its applications in business and industry.
 - (b) Discuss briefly the importance and the use of index numbers in business.
3. (a) What are the uses of index numbers ? What are the problems in their construction? (MBA, Vikram Univ., 1994)
 - (b) What are index numbers ? How are they constructed ? Explain the role of weights in the construction of general price index numbers.
 - (c) Explain the nature and uses of index numbers.
4. What is an index number ? Examine the various problems involved in the construction of an index number. Discuss in brief the uses of an index number.
 5. What is an index number ? Explain the terms price relative, quantity relative and value relative with reference to a single commodity and deduce the factor reversal property.
 6. Describe the steps involved in the computation of Fisher's Ideal Index Number. What are its advantages and disadvantages ?
 7. What is Fisher's Ideal Index ? Why is it called ideal ? Show that it satisfies both the time reversal test as well as the factor reversal test. (MBA, Sukhadia Univ.; MBA, HPU 1990)
 8. Laspeyres' price index generally shows an upward trend in the price changes while Paasche's method shows a downward trend on them. Elucidate the statement. (MBA, Delhi Univ., 1997)
 9. (a) "Index numbers are signs and guide-posts along the business highway, that indicate to the businessman how he should drive or manage his affairs". Explain the above statement and also point out the relative advantages of the various types of averages as applied to index numbers. Which would you prefer and why ?
 - (b) What is Fisher's Index ? Why is it called Ideal ?
10. Discuss the following statements :
 - (i) Compute
 - (ii) "The purpose determines the type of index number to use."
 - (iii) "An index number is a special type of average."
 - (iv) "There is no such thing as unweighted index numbers."
 - (v) "The choice of a suitable base period is at best a temporary solution." Why ?
 - (vi) "Theoretically, geometric mean is the best average in the construction of index numbers but in practice most arithmetic mean is used." Why ?
 11. (a) "Since the value of the base is always 100, it does not make any difference which period is selected as the base in which to construct an index." Comment.

- (b) If you are employed to construct a price index for a department store that sells thousands of items (a) how would you decide on which items to include? (b) how would you define the price? (c) what weights would you use? (d) which formula would you select?
12. (a) Define index numbers. Describe the construction of wholesale price index number elucidating the following points:
- Selection of commodities,
 - Selection of the prices and the market,
 - Selection of the base year,
 - Selection of the average,
 - Decision on the system of weighting.
- (b) What is an ideal index? How does the Fisher formula for ideal index satisfy the following two tests:
- Time Reversal Test, and
 - Factor Reversal Test.
13. (a) Distinguish clearly between fixed base and chain base index number and point out their relative merits and demerits.
 (b) Explain Time Reversal Test and Factor Reversal Test with the help of a suitable example.
- (MBA, HPU; MBA, Osmania Univ., 1998)
14. What are time reversal and factor reversal tests? Does the following index number formula satisfy these tests?

$$I = \sqrt{\frac{\sum p_x q_0}{\sum p_0 q_0} \times \frac{\sum p_x q_x}{\sum p_0 q_x}} \times 100$$

15. What is an index number? Discuss its importance in business and industry.

Explain:

- Time reversal test,
 - Factor reversal test, and
 - Circular test as applied to index number.
16. What do you understand by reversibility of index numbers? Explain time reversal and factor reversal test in this context.
17. (a) It is said that index numbers are a specialized type of averages. How far do you agree with this statement. Explain briefly time reversal and factor reversal tests.
 (MBA, Osmania, 1998)
- (b) What are the factor reversal and circular tests of consistency in the selection of an appropriate index formula? Verify whether Fisher's Ideal Index satisfies such tests.
 (CA, May, 1999)
- (c) The following are the prices of six different commodities for 2004 and 2005. Compute a price index by (a) simple aggregative method and (b) average of price relative method by using both arithmetic mean as well as geometric mean.

Commodity	Unit	Price in 2004		Price in 2005	
		(Rs.)	(Rs.)	(Rs.)	(Rs.)
Wheat	Quintal	1900	2200		
Rice	"	1500	2000		
Pulses	"	2000	3000		
Ghee	1 kilo	120	122		
Butter	"	130	136		
Potatoes	"	11	12		

18. Construct an appropriate index for purposes of comparison from the following data:

Commodity	A		B		C	
	Price	Qty.	Price	Qty.	Price	Qty.
Year						
2004	4	50	3	10	2	5
2005	10	40	8	8	4	4

Jewel's Care Collected

19. The following table gives the per capita income and cost of living index number of a particular community. Deflate the per capita income by taking into account the rise in the cost of living :

Year	Per capita income	Cost of Living Index No. Base
1996	800	100
1997	900	150
1998	950	180
1999	1020	200
2000	1150	220
2001	1200	250
2002	1500	300
2003	1600	400

20. Calculate Laspeyres' and Paasche's Index number from the following data :

Items	Qty.	Base Year		Current Year	
		Price per Kg.	Qty.	Price per Kg.	Qty.
Bread	10	Rs. 22.50	12	Rs. 25.00	
Meat	8	Rs. 80.00	9	Rs. 90.00	
Tea	2	Rs. 100.00	4	Rs. 120.00	

21. Construct from the following data spliced index continuous with index A and a spliced index continuous with index B:

Year	Index A	Index B
1999	100	
2000	95	
2001	110	
2002	125	110
2003		105
2004		94

22. (a) In the construction of a certain consumer price index the following group index numbers were found. Calculate the consumer price index by using (i) the weighted arithmetic mean, and (ii) the weighted geometric mean :

Groups	Index	Weights
Food	300	5
Fuel and Lighting	250	1
Clothing	280	1
House Rent	200	2
Miscellaneous	150	1

(b) In calculating a certain cost of living index number, the following weights were used : Food 15, clothing 0, rent 4, fuel and light 2, miscellaneous 1. Calculate the index for a date when the average percentage increases in prices of items in the various groups over the base period were 32, 54, 47, 74 and 58 respectively.

23. Using the following data, show that Fisher's Ideal formula satisfies the Factor Reversal Test :

Commodity	Price Per Unit (Rs.)		Number of Units	
	Base Period	Current Period	Base Period	Current Period
A	6	10	50	56
B	2	2	100	120
C	4	6	60	60
D	10	12	30	24
E	8	12	40	36

[139.79]

24. Using the above food index and the information given below calculate the cost of living index number :

Groups	Food	Clothing	Fuel & Light	House Rent	Miscellaneous
Index	-	310	220	150	300
Weight	60	5	8	9	18

25. Given below are the data on prices of some consumer goods and the weights attached to the various items. Compute price index numbers for the year 2005 (Base: 2004 = 100) using (i) simple average, and (ii) weighted average of price relatives.

Item	Unit	Price (Rs.)		Weight
		2004	2005	
Wheat	kg.	10.00	11.00	2
Milk	litre	15.00	16.00	5
Sugar	kg.	16.00	17.00	8
Shoes Pair	Rs.	500.00	550.00	1

26. An enquiry into the budgets of the middle-class families of a certain city revealed that on an average the percentage expenses on the different groups were—Food 45, rent 15, clothing 12, fuel and light 8 and miscellaneous 20. The group index numbers for the current year as compared with a fixed base period were respectively 410, 150, 343, 248 and 285. Calculate the consumer price index number for the current year. Mr. X was getting Rs. 240 in the base period and Rs. 480 in the current year. State how much he ought to have received as extra allowance to maintain his former standard of living. (MBA, HPU, 1997)

27. The following are the group index numbers and the group weights of an average working class family's budget. Construct the cost of living index number by assuming the weight :

Group	Index number	Weight
Food	152	48
Fuel and Lighting	110	6
Clothing	130	8
House rent	100	15
Miscellaneous	90	

[129.73]

28. From the chain base index numbers given below prepare fixed base index numbers.

Year	2000	2001	2002	2003	2004
Chain base Index No.	80	140	130	110	90

[80, 112, 145.6, 160.16, 144.1]

29. From the following data prepare index number for real wages of workers :

Year	1999	2000	2001	2002	2003	2004
Wages (Rs.)	2000	2500	3110	3600	3900	4000
Index number	100	160	280	300	330	340

30. Calculate the Fisher's ideal index. Does this data satisfy time and factor reversal tests :

Commodity	Price	Qty.	Price	Qty.
A	5	4	4	5
B	5	10	6	10
C	8	15	8	20
D	10	8	12	6
E	2	6	2	8

[106.63, yes]

31. The following table gives the average wholesale price of five groups of commodities for the years 2000 to 2004. Compute chain base index numbers.

Commodity	2000	2001	2002	2003	2004
A	2	3	4	2	7
B	3	6	9	4	3
C	4	12	20	8	22
D	5	7	22	16	18
E	3	8	11	14	12

320 Business Statistics

32. Compute the index numbers of prices from the following data by applying : (a) Laspeyres', (b) Paasche's (c) Fisher's, and (d) Bowley's method.

Commodity	2004		2005	
	Price	Quantity	Price	Quantity
A	3	8	6	9
B	5	9	8	10
C	6	15	7	12
D	4	20	5	15

[(a) 135.98; (b) 140.19; (c) 138.07; (d) 138.09]

33. Prepare index numbers (1997 = 100) from the Link relatives given below :

Year	1998	1999	2000	2001	2002	2003	2004
Link relatives	105	75	71	105	98	90	90

34. Calculate Laspeyres', Paasche's, and Fisher's Ideal Index from the following data :

Commodity	Price	Value	Price	Value
A	10	100	8	96
B	16	96	14	98
C	12	36	10	40
D	15	60	5	25

[P_{01} = 73.29; 72.96; 73.12]

35. Prepare price index numbers for 2004 with 1994 as base year from the following data by using (i) Laspeyres' (ii) Paasche's and (iii) Fisher's method. (Correct up to 4 places of decimal)

Year	Article							
	I		II		III		IV	
	P	Q	P	Q	P	Q	P	Q
1995	12.50	9	9.63	7.75	6	5.00	5	
2005	18.75	9	7.75	8.80	10	10	6.507	

[P: Price; Q: Quantity]

With the help of above data prove that the Time Reversal Test is satisfied by Fisher's formula, but not necessarily by the Laspeyres' and Paasche's index numbers.

36. Construct Fisher's Ideal Index No. for the following data and show that it satisfies the time reversal and factor reversal tests :

Commodity	Base Year		Current Year	
	Price	Qty.	Price	Qty.
A	6	30	15	40
B	5	40	10	55
C	10	25	12	20
D	4	15	3	30
E	2	50	5	28

37. From the following prices of these groups of commodities for the years 1999 to 2003, find the chain base index numbers chained to 1999 :

Groups	1999	2000	2001	2002	2003
I	4	6	8	10	12
II	16	20	24	30	36
III	8	10	16	20	24

(MBA, Madurai Kamaraj Univ., April 2002)

(MBA, M.K. Univ., May 2002)

PROBLEMS

1-A : Answer the following questions, each question carries one marks:

- (i) What do you mean by Time series ?
- (ii) Write a short note on "Secular trend".
- (iii) What is seasonal variation ?
- (iv) Write down the most important factors causing seasonal variations.
- (v) What are the normal equations for the straight line $Y = a + bx$?
- (vi) What are cyclical fluctuations ?
- (vii) What is Business forecasting ?
- (viii) Name a few methods of Business forecasting.
- (ix) How irregular variations are caused ?
- (x) "Forewarned is Forearmed." Comment.

(MBA, Madurai-Kamaraj, Nov 2000)
(MBA, Madurai-Kamaraj, Nov 2001)
(MBA, Madurai-Kamaraj, Nov 2002)
(M. Com., Madurai-Kamaraj, 2003)

1-B : Answer the following questions, each question carries four marks:

- (i) What are seasonal variations ? Bring out the factors that cause seasonal variations.
(M.A. Econ., Madras Univ., April 2003)
- (ii) Narrate the merits and limitations in the use of moving average method. (M.A. Econ., Madras Univ., 2003)
- (iii) Distinguish between the additive and multiplicative models of time series analysis.
- (iv) Suggest the important adjustments to the model before analysing time series.
- (v) Explain the freehand (graphic method) of measuring trend.
2. (a) What is Business Forecasting ? Explain its role and limitations.
(b) How does analysis of time series help in making business forecast ?
(c) What is forecasting ? Discuss in brief the various theories and methods of business forecasting.
(MBA, Delhi Univ., 1998)
3. (a) Explain clearly the different components into which a time series may be analysed. Explain any method in isolating trend values in a time series. (MBA, Delhi Univ., MBA, Vikram Univ., 1994)
(b) Explain clearly the meaning of Time Series Analysis. Mention its important components. Explain these components with examples, indicating the importance of each component in business. (B.Com., Andhra Univ., 2003)
(c) Describe the seasonal variation and cyclical fluctuations in a time series. (MBA, Anna Univ., 2003)
4. Explain what do you understand by Time Series. Why is Time Series considered to be an effective tool of forecasting ?
(MBA, BHU, 2002)
5. (a) What is business forecasting ? What are the assumptions on which business forecasts are made ? Describe the techniques of forecasting that are commonly employed by big business houses.
(b) Explain briefly the additive and multiplicative models of time series. Which of these models is more popular in practice and why ?
(MBA, Osmania, 1989)
6. (a) Critically examine the various methods that are used for measuring trend. Which method do you think is the best and why ?
(b) Explain briefly the different methods of measuring trend. (MBA, Madras Univ., 2002)
7. (a) How seasonal variations are accounted for in the analysis of Time Series ?
(b) What are the common methods in use for eliminating seasonality from a time series data ? Explain any one method taking imaginary figures.
8. Critically examine the various methods that are used for business forecasting. Why is time series considered to be an effective tool for forecasting analysis ? Explain.
9. Explain the following terms in the study of time series :
(i) Secular trend, (ii) Seasonal variation, (iii) Cyclical fluctuations.
10. (a) What do you understand by 'seasonal variation' in time series data ? Explain their uses.
(b) Why do we measure seasonal variations in a time series ?
(c) How would you eliminate seasonal influences ? Illustrate with the help of an example.
(d) Explain clearly with the help of an illustration how seasonal index is useful in planning sales or production for specific periods. Are there any limitations of seasonal index ?
11. (a) Explain the method of Moving Averages in estimating the trend of a time series. What are the disadvantages in using this method ?
(b) Explain the concept of 'auto correlation' and its use for time series analysis. Give an example of a single variable with two different time lags.
(MBA, IGNOU, 2001)

12. (a) Why do we deseasonalize data? Explain the ratio-to-moving average method to compile the seasonal index.
 (b) Explain the following statements:
 (i) "... the business analyst who uses moving averages to smoothen his data while in the process of trying to discover business cycles, is likely to come up with some non-existent cycles."
 (ii) "There is nothing sacred in computing seasonal indices by the method of moving average using exclusively monthly data."
 (iii) "Despite great limitations of statistical forecasting, the forecasting techniques are invaluable to the economist, the businessman and the Government."
13. Suppose you are provided with a given time series data and asked to analyse its general pattern and fluctuations. Describe in detail the steps you would follow in determining the pattern of trend and whether a seasonal and/or a cyclical component contributed to movements in the series.
14. (a) (i) "A key assumption in the classical method of time series analysis is that each of the component movements in the time series can be isolated individually from a series". Do you agree with this statement? Does this assumption create any serious limitation to such analysis?
 (ii) "A 12-month moving average of time series data removes trend and cycle." Do you agree? Why or why not?
 (b) Examine critically the time-lag and the action and the reaction theory of business forecasting. Which of these, in your opinion, is better and why?
15. Answer the following by a brief statement on each:
 (i) Why must short-term forecasts be more precise than long-term ones?
 (ii) What is the major objective of seasonal analysis?
 (iii) What purpose does a seasonal index solve?
16. (a) What is the difference between seasonal fluctuations and cyclical variations in a time series data.
 (b) Illustrate the historical analogy theory of business forecasting.
 (c) What is a time series? What are its components? Which components of the series is mainly applicable in the following cases?
 (i) A fire in a factory delaying production for one month.
 (ii) Formation of rocks.
 (iii) Decrease in the employment in sugar factory during the off-season.
 (iv) Sale of New Year greeting cards.
 (v) Fall in death rate due to advances in science.
 (vi) An after Deepawali sales in a departmental store.
 (vii) A need for increased rice production due to a constant increase in population.
17. Indicate three categories of forecasting models and list out five techniques from each category. Describe Delphi technique in detail.
18. (a) Critically examine the time-lag and the action and reaction theory of business forecasting. Which of these two is better and why?
 (b) While fitting a straight line trend of the type $Y = a + bX$, what is signified by Y , X , a and b ?
19. (a) Discuss the role of forecasting as a business tool.
 (b) How do we manage long-range forecasting and technical change for any organisation?
 (c) Write short notes on Delphi method and Historical analogy method for business forecasting.
 (d) Explain how can we use market surveys as a method of forecasting. Illustrate. (MBA, Kurukshetra Univ., 1997)
 (e) Write a lucid note on Box and Jenkin's method of forecasting. (MBA, Osmania, 1999)
 (f) Explain with appropriate example different methods of estimating seasonal variations. (MBA, Jamia Millia, 2003)
 (g) What do you understand by Naive (Time series) Quantitative Models of forecasting?
20. Business today generate a large amount of data continuously. This data may be used to gain information about the system. For one such system, it is known that the relation between variables is non-linear, i.e., in the form $y = ax^b$, where a and b are constants. Use a transformation to make it linear and discuss how would you use the method of least squares to fit a straight line to the transformed linear model.
21. Apply the method of semi-averages for determining trend to the following data and estimate the value for 2006:

Year	Sales (Thousand units)	Year	Sales (Thousand units)
2000	20	2003	30
2001	24	2004	28
2002	22	2005	32

If the actual figure of sales for 2006 is 35,000 units, how do you account for difference between the figure you obtain and the actual figure given to you?

Jewel's Care Collected

22. Plot the following data on graph paper and ascertain trend by the method of semi-averages :

Year	Sales (million tonnes)	Year	Sales (million tonnes)
1999	100	2003	108
2000	120	2004	102
2001	95	2005	112
2002	105		

23. Apply the method of semi-average to depict the long-term tendency of following data and estimate the value for 2008

Year	Production (million tonnes)	Year	Production (million tonnes)
1998	40	2002	51
1999	44	2003	50
2000	42	2004	54
2001	48	2005	56

24. The following series relate to the profits of a commercial concern for 8 years :

Year	Profits Rs.	Year	Profits Rs.
1998	15,420	2002	26,120
1999	14,420	2003	31,950
2000	15,520	2004	35,360
2001	21,020	2005	35,670

Find the trend of profits. Assume a three-year cycle and ignore decimals.

25. Find out the trend values for the following time series of production by the method of moving average using 5-point time period for your purpose. State briefly the procedure that would have been adopted if you were to choose a 4-point time period. How does one choose the proper 'period' of the moving average?

Year	Production (m. tonnes)	Year	Production (m. tonnes)	Year	Production (m. tonnes)
1988	351	1994	410	2000	502
1989	366	1995	420	2001	540
1990	361	1996	450	2002	557
1991	362	1997	500	2003	571
1992	400	1998	518	2004	586
1993	419	1999	455	2005	612

26. Below are given the figures of production of a sugar factory :

Year	Production (thousand tonnes)	Year	Production (thousand tonnes)
2000	92	2003	92
2001	83	2004	92
2002	94	2005	110

Apply the method of least squares to determine the trend values. Also find out the short-term fluctuations.
 $[Y = 95 + 1473.X]$

27. Fit a straight line trend by the method of least squares :

Year	Milk consumption (million litres)	Year	Milk consumption (million litres)
1997	102.3	2002	118.7
1998	101.9	2003	124.5
1999	105.8	2004	129.9
2000	112.0	2005	134.8
2001	114.8		

$$[Y = 116.1 + 4.3.X]$$