MES Abasaheb Garware College (Autonomous), Pune

(Affiliated to Savitribai Phule Pune University)

S. Y. B. Sc. (Computer Science)

S. Y. B. C. A.

As per NEP Structure

Semester-IV

CS-292-MNP/CA-292-MNP

Data Representation and Condensation

Work Book

Name:_____

Roll No.: _____ Division:_____

Academic Year: _____

Editors:

Dr. Yogeshri Gaidhani - MES Abasaheb Garware College, Pune
Mr. Akshay Shelke - MES Abasaheb Garware College, Pune

Reviewer:

Mrs. Chitra Nagrkar - MES Abasaheb Garware College, Pune

Introduction

1. **About the work book:**
   This workbook is intended to be used by S.Y.B.Sc. (Computer Science) and S.Y.B.C.A. students for Data Representation and Condensation in Semester–IV. This workbook is designed by considering all the practical concepts / topics mentioned in syllabus.

2. **The objectives of this workbook are:**
   1) Defining the scope of the course.
   2) To have continuous assessment of the course and students.
   3) Providing ready reference for the students during practical implementation.
   4) Provide more options to students so that they can have good practice before facing the examination.
   5) Catering to the demand of slow and fast learners and accordingly providing the practice assignments to them.

3. Instructions to the students

Please read the following instructions carefully and follow them.

1. Students are expected to carry this book every time they come to the lab for practical.
2. Students should prepare oneself beforehand for the Assignment by reading the relevant material.
3. Instructor will specify which problems to solve in the lab during the allotted slot and student should complete them and get verified by the instructor. However student should spend additional hours in Lab and at home to cover as many problems as possible.
4. Students will be assessed for each exercise on a scale from 0 to 5.

Not Done	0
Incomplete	1
Late Complete	2
Needs Improvement	3
Complete	4
Well Done	5

4. Instruction to the Instructors

1) Explain the assignment and related concepts in around ten minutes using white board if required or by demonstrating the software.
2) You should evaluate each assignment carried out by a student on a scale of 5 as specified above by ticking appropriate box.
3) The value should also be entered on assignment completion page of the respective Lab course.

M.E.S. Abasaheb Garware College, Pune-04.

S.Y.B.Sc.(Comp.Sc.)/ S.Y.B.C.A. Data Representation and Condensation

Certificate

This is to certify that Mr./Ms._____ ,
Roll No. _____ of Class _____ has satisfactorily completed all the assignments of Data Representation and Condensation in Semester IV of the academic year 2024-25 .

Date: Batch In-charge

Internal Examiner External Examiner

Assignment Completion Sheet

Assignment No	Description	Marks (out of 2)
1		
2		
3		
4		
5		
6		
7		
8		
9		

Assignment No. 1- Introduction

Q.1 The following data give the marks of 50 students in a test:

21,50,35,39,48,46,36,54,42,30,29,42,32,40,34,31,35,37,52,44,39,45,37,33,51,53,52,46,43,47,41,26,52,48,25,34,37,33,36,27,54,36,41,33,23,39,28,44,45,38.

a) Prepare a frequency table taking class-interval 20-24, 25-29,..., etc.

b) Draw i) Histogram ii) Ogive curves

Q.2 The weights in grams of 50 mobile phones picked out at random are as follow:

156,167, 176, 182,169,167,145, 193,187,125,123,125,150, 192,186,170, 126,168,130,149, 139, 119,155,178, 200,186,184, 199,173,194, 150,141,136,163, 190,115,198,160,178,190, 157,181, 131,175, 184, 194, 210, 180, 148, 182.

a) Form the grouped frequency table by dividing the variate range into intervals of equal width, each corresponding to 20 grams.

b) Find the no of mobiles below weight 160grams. Also find the number of mobiles above weight 180 grams. Estimate the number of mobiles with weight at least 120 grams but less than 160 grams.

c) Mobiles with weight less than or equal to 160 are called ‘light weight’, with weight greater than 160 and less than or equal to 180 are ’normal weight’ and with weight greater than 180 are called ‘heavy weight’. Draw Pie chart for classification of mobiles as light, normal and heavy weight.

Q.3 Sales of a company for 2021-2024 are stated below. Represent the data using subdivided bar diagram.

Year	North	South	East	West
2021	15	20	18	12
2022	10	30	25	20
2023	20	25	30	25
2024	25	20	40	60

Q.4 An examination was held in a college for F.Y.BCS class and it resulted as follows:

Number of students getting at least 20 marks was 81, at least 40 marks was 60, at least 60 marks was 45, at least 80 marks was 15.

a) Tabulate the above information in a frequency table.

b) Also draw histogram and more than Ogive.

Q.5 Three regions Vidarbha, Marathwada and Western Maharashtra were surveyed for damage to the crops due to unseasonal rains. It was observed that the area under cultivation in Vidarbha, Marathwada and Western Maharashtra was 80, 65 and 85 lakh hectors respectively. Area of crops damanged in Vidarbha was 20 lakh hectors, in Marathwada 13 lakh hectors and in Western Maharashtra 17 lakh hectors. Prepare a table which summarizes the above information. Draw appropriate diagram.

Assignment Evaluation

0: Not Done [ ]	1: Incomplete [ ]	2: Late Complete [ ]
3: Needs Improvement [ ]	4: Complete [ ]	5: Well Done [ ]

Signature of Instructor

Assignment No. 2 - Graphical Representation of Data

Q.1 Represent the following information by an appropriate diagram using Excel.

i) According to Maharashtra economic survey 2023-24 GDP numbers are as given below.

District	Pune	Mumbai	Satara	Nagpur	Wardha	Solapur	Dhule	Nanded
GDP (in thousand Cr.)	418	716	81	181	33	121	45	67

ii) The following data represents the daily rainfall (in mm) recorded over 60 days

Class Interval (Rainfall in mm)	**Frequency**
0-9	6
10-19	12
20-29	10
30-39	8
40-49	4

iii) A survey of 50 students recorded the time (in minutes) they took to solve a problem

Class Interval (Time taken in minutes)	0-9	10-19	20-29	30-39	40-49	50-59	60-69
No. of students	0	4	6	12	15	8	5

Q.2 Draw histogram and Ogives for the following frequency distribution of marks in statistics.

Marks	0-20	20-40	4-60	60-80	80-100
No. of students	2	18	42	28	5

Q.3 Represent the following data by means of a histogram.

Class	10-15	15-20	20-25	25-30	30-40	40-50	60-80
Frequency	7	19	27	15	12	12	8

Q.4 The following table shows the marks obtained by 50 students in a test.

Ages (in years)	0-9	10-19	20-29	30-39	40-49	50-59	60-69
Frequency	8	12	15	20	25	10	5

a) Prepare a cumulative frequency table.

b) Draw the "more than" ogive curve.

Q.5 The following table shows the marks obtained by 50 students in a test.

Marks	0-10	11-20	21-30	31-40	41-50	51-60	61-70	71-80
No of students	2	2	5	8	12	10	7	2

a) Prepare a cumulative frequency table.

b) Draw the "less than" ogive curve.

Q.6 The following data shows the marks obtained by 15 students in a test:

55, 63, 72, 54, 78, 89, 65, 72, 66, 88, 93, 77, 59, 62, 68.

a) Construct a stem-and-leaf chart for the marks obtained.

Q.7 The profit (in lakhs of Rs.) of 30 companies for the year 2023-24 are given below.

Prepare a stem and leaf chart which stem has maximum leaves.

20, 22, 35, 42, 37, 42, 48, 53, 49, 65, 39, 48, 67, 18, 16,

23, 27, 35, 49 ,63, 65, 55, 45, 58, 57, 69, 25, 29, 58, 65.

Assignment Evaluation

0: Not Done [ ] 1: Incomplete [ ] 2: Late Complete [ ]

3: Needs Improvement [ ] 4: Complete [ ] 5: Well Done [ ]

Signature of Instructor

Assignment No. 3 - Measures of Central Tendencies

Q. 1 In a set of 50 items, arranged in ascending order of magnitude the values of 24^{th}, 25^{th} and 26^{th} items are 40, 42 and 45 respectively. Find the median. Also find the median if the number of observations was 51.

Q. 2 The weekly incomes of 20 families in a village are as follows: 1200, 1000, 1100, 1250, 950, 1300, 1350, 1150, 1200, 1050, 1200, 1000, 1100, 900, 1200, 1300, 1350, 1250, 950, 1000. Find the mean, median and mode of the income distribution using excel.

Q. 3 A study of a certain operation shows the following distribution for 100 workers. Calculate the mean, median and mode using excel.

Class Interval (In seconds)	221-240	241-260	261-280	281-300	301-320	321 340	341-360	361-380	381-400
Frequency	6	9	11	14	20	15	10	8	7

Q. 4 Using excel calculate median and mode for the following frequency distribution:

x_i	Less than 1000	1000-2000	2000-3000	3000-4000	4000-5000	Above 5000
f_i	08	14	13	25	07	03

Q. 5 Calculate arithmetic mean mode median of the following frequency distribution using excel.

Class	10-15	15-20	20-25	25-30	30-35	35-40	40-45
Frequency	5	6	8	9	7	5	4

Q. 6 A set of 10 values has arithmetic mean 20. Find the arithmetic mean if

a) Each value is doubled and then increased by 2.
b) Each value is increased by 5 and then doubled.
c) Each value is decreased by 5.
d) Each value is increased by 3.

Q. 7 Find the combined arithmetic mean of the salary given that:

Group	**Sales**	**Production**	**Accounts**	**Security**	**HR**
No. of employees	75	100	15	40	15
Arithmetic mean of salaries	16000	25100	19000	15000	19500

Assignment Evaluation

0: Not Done [ ]	1: Incomplete [ ]	2: Late Complete [ ]
3: Needs Improvement [ ]	4: Complete [ ]	5: Well Done [ ]

Signature of Instructor

Assignment No. 4 - Measures of Central Tendencies using Excel

Solve the following problems using excel.

Q. 1 The weekly incomes of 22 families in a village are as follows: 1200, 1000, 1100, 1250, 950, 1300, 1000, 1350, 1150, 1200, 1050, 1200, 1000, 1100, 900, 1200, 1300, 1000, 1350, 1250, 950, 1000. Find the mean, median and mode of the income distribution.

Q. 2 A study of a certain operation shows the following distribution for 100 workers. Calculate the mean, median and mode using excel.

Class Interval (In seconds)	221-240	241-260	261-280	281-300	301-320	321 340	341-360	361-380	381-400
Frequency	6	9	11	14	20	15	10	8	7

Q. 3 Using excel calculate median and mode for the following frequency distribution:

x_i	Less than 1000	1000-2000	2000-3000	3000-4000	4000-5000	Above 5000
f_i	08	14	13	25	07	03

Q. 4 Calculate arithmetic mean mode median of the following frequency distribution using excel.

Class	10-15	15-20	20-25	25-30	30-35	35-40	40-45
Frequency	5	6	8	9	7	5	4

Q. 5 A company surveyed 100 employees to determine the number of hours they work per week. The following frequency distribution table was created based on the data:

Hours Worked per Week	**Frequency (Number of Employees)**
30 - 39	10
40 - 49	25
50 - 59	35
70 - 79	15
80 - 89	10

a) Calculate the **mean** number of hours worked per week.
b) Calculate the **median** number of hours worked per week.
c) Calculate the **mode** of hours worked per week.

Assignment Evaluation

0: Not Done [ ] 1: Incomplete [ ] 2: Late Complete [ ]

3: Needs Improvement [ ] 4: Complete [ ] 5: Well Done [ ]

Signature of Instructor

Assignment No. 5 - Measures of Dispersion

Q.1 Find range and coefficient of range for the weekly incomes of 20 families in a village are as follows: 1200, 1000, 1100, 1250, 950, 1300, 1350, 1150, 1200, 1050, 1200, 1000, 1100, 900, 1200, 1300, 1350, 1250, 950, 1000.

Q.2 Find Quartile deviation and coefficient of quartile deviation for the profit (in lakhs of Rs.) of 12 companies for the year 2023-24 are given below.
20, 22, 35, 42, 37, 42, 48, 53, 49, 65, 45, 50.

Q.3 Find variance and standard deviation for the following frequency distribution of marks in a test.

Marks	0-20	20-40	40-60	60-80	80-100
No. of students	2	18	42	28	5

Q.4 Find standard deviation and coefficient of variation for the following distribution.

Profit (in Crore Rs.)	10-20	20-30	30-40	40-50	50-60
Number of Companies	8	12	20	6	4

Q.5 The following table shows the marks obtained by 100 students in two tests. Find coefficient of variations for both the tests and compare.

Marks	0- 9	10- 19	20- 29	30- 39	40- 49
Number of students –Test 1	10	20	25	30	15
Number of students –Test 2	5	20	30	35	10

Assignment Evaluation

0: Not Done [ ] 1: Incomplete [ ] 2: Late Complete [ ]

3: Needs Improvement [ ] 4: Complete [ ] 5: Well Done [ ]

Signature of Instructor

Assignment No. 6 Measures of Dispersion using Excel

Solve the following problems using excel.

Q.6 Find range and coefficient of range for the weekly incomes of 20 families in a village are as follows: 1200, 1000, 1100, 1250, 950, 1300, 1350, 1150, 1200, 1050, 1200, 1000, 1100, 900, 1200, 1300, 1350, 1250, 950, 1000.

Q.7 Find Quartile deviation and coefficient of quartile deviation for the profit (in lakhs of Rs.) of 12 companies for the year 2023-24 are given below.
20, 22, 35, 42, 37, 42, 48, 53, 49, 65, 45, 50.

Q.8 Find variance and standard deviation for the following frequency distribution of marks in a test.

Marks	0-20	20-40	40-60	60-80	80-100
No. of students	2	18	42	28	5

Q.9 Find standard deviation and coefficient of variation for the following distribution.

Profit (in Crore Rs.)	10-20	20-30	30-40	40-50	50-60
Number of Companies	8	12	20	6	4

Q.10 The following table shows the marks obtained by 100 students in two tests. Find coefficient of variations for both the tests and compare.

Marks	1- 9	10- 19	20- 29	30- 39	40- 49
Number of students –Test 1	10	20	25	30	15
Number of students –Test 2	5	20	30	35	10

Assignment Evaluation

0: Not Done [ ] 1: Incomplete [ ] 2: Late Complete [ ]

3: Needs Improvement [ ] 4: Complete [ ] 5: Well Done [ ]

Signature of Instructor

Assignment No. 6 - Measures of Dispersion using Excel

Solve the following problems using excel.

Q.1 Find range and coefficient of range for the weekly incomes of 20 families in a village are as follows: 1200, 1000, 1100, 1250, 950, 1300, 1350, 1150, 1200, 1050, 1200, 1000, 1100, 900, 1200, 1300, 1350, 1250, 950, 1000.

Q.2 Find Quartile deviation and coefficient of quartile deviation for the profit (in lakhs of Rs.) of 12 companies for the year 2023-24 are given below.
20, 22, 35, 42, 37, 42, 48, 53, 49, 65, 45, 50.

Q.3 Find variance and standard deviation for the following frequency distribution of marks in a test.

Marks	0-20	20-40	40-60	60-80	80-100
No. of students	2	18	42	28	5

Q.4 Find standard deviation and coefficient of variation for the following distribution.

Profit (in Crore Rs.)	10-20	20-30	30-40	40-50	50-60
Number of Companies	8	12	20	6	4

Q.5 The following table shows the marks obtained by 100 students in two tests. Find coefficient of variations for both the tests and compare.

Marks	2- 9	10- 19	20- 29	30- 39	40- 49
Number of students –Test 1	10	20	25	30	15
Number of students –Test 2	5	20	30	35	10

Assignment Evaluation

0: Not Done [ ] 1: Incomplete [ ] 2: Late Complete [ ]

3: Needs Improvement [ ] 4: Complete [ ] 5: Well Done [ ]

Signature of Instructor

Assignment No. 7 - Raw and Central Moments

Q1. Find first four central moments for the set of numbers 2, 4, 6, 8.

Q2. Find first four central moments for the following frequency distribution.

Class	0-10	10-20	20-30	30-40
Frequency	1	3	4	2

Q3. Find first four raw moments of the distribution 2, 20, 40, 80. Hence find the central moments.

Q4. The first three raw moments of a distribution are 2, 25 and 80 respectively. Calculate the first three central moments.

Assignment Evaluation

0: Not Done [ ] 1: Incomplete [ ] 2: Late Complete [ ]

3: Needs Improvement [ ] 4: Complete [ ] 5: Well Done [ ]

Signature of Instructor

Assignment No. 8 - Raw and Central Moments using Excel

Solve the following problems using Excel.

Q1. Find first four central moments for the set of numbers 2, 4, 6, 8.

Q2. Find first four central moments for the following frequency distribution.

Class	0-10	10-20	20-30	30-40
Frequency	1	3	4	2

Q3. Find first four raw moments of the distribution 2, 20, 40, 80. Hence find the central moments.

Q4. Find first four central moments for the following frequency distribution.

Class	0-20	20-40	40-60	60-80	80-100
Frequency	2	18	42	28	5

Assignment Evaluation

0: Not Done [ ] 1: Incomplete [ ] 2: Late Complete [ ]

3: Needs Improvement [ ] 4: Complete [ ] 5: Well Done [ ]

Signature of Instructor

Assignment No. 9 – Skewness and Kurtosis

Q. 1 If mode of a frequency distribution is larger than mean by 4.5 and its variance is 121, find Karl Pearson's coefficient of skewness.

Q. 2 If median of a frequency distribution is larger than mean by 5 and its variance is 16, compute Karl Pearson's coefficient of skewness.

Q. 3 If mean, standard deviation and Sk_p of a frequency distribution are 58, 20, and 0.32, then find its median and mode.

Q. 4 Obtain Karl Pearson's coefficient of skewness Sk_p for the following frequency distribution

height (in inches)	No. of persons
59-61	9
61-63	20
63-65	35
65-67	10
67-69	6

Q. 5 Obtain Bowley's coefficient of skewness Sk_b for the following frequency distribution

height (in inches)	No. of persons
59-61	9
61-63	20
63-65	35
65-67	10
67-69	6

Q. 6 For the two distribution A and B following is the summary statistics compare skewness of two distributions using Bowley's coefficient of skewness.

Quartiles \ distributions	A	B
Q1	63.65	61.49
Q2	80.1	79.8
Q3	125	124.5

Q. 7 For the two distribution A and B following is the summary statistics compare skewness of two distributions using Karl person's coefficient of skewness.

For group A	$\overline{X_1} = 55$	$\sigma_1 = 15.4$	Mode=58.72
For group B	$\overline{X_2} = 53$	$\sigma_2 = 15.4$	Mode=48.83

Q. 8 Find first four central moments of the given frequency distribution, hence comment on its skewness and kurtosis.

Class	0-10	10-20	20-30	30-40
Frequency	1	3	4	2

Q. 9 The first four raw moment of a frequency distribution are 2, 25, 80, 122 comment on the skewness and kurtosis of this frequency distribution.

Q. 10 Given that $\beta_1 = 0.5$, $\beta_2 = 2.5$ and $\mu_2 = 2$, find μ_3 and μ_4. If $\mu_1' = 0.5$, $\mu_2' = 19.672$, $\mu_3' = 29.258$, and $\mu_4' = 856.502$. Is the distribution platykurtic? Justify your answer.

Assignment Evaluation

0: Not Done [ ] 1: Incomplete [ ] 2: Late Complete [ ]

3: Needs Improvement [ ] 4: Complete [ ] 5: Well Done [ ]

Signature of Instructor