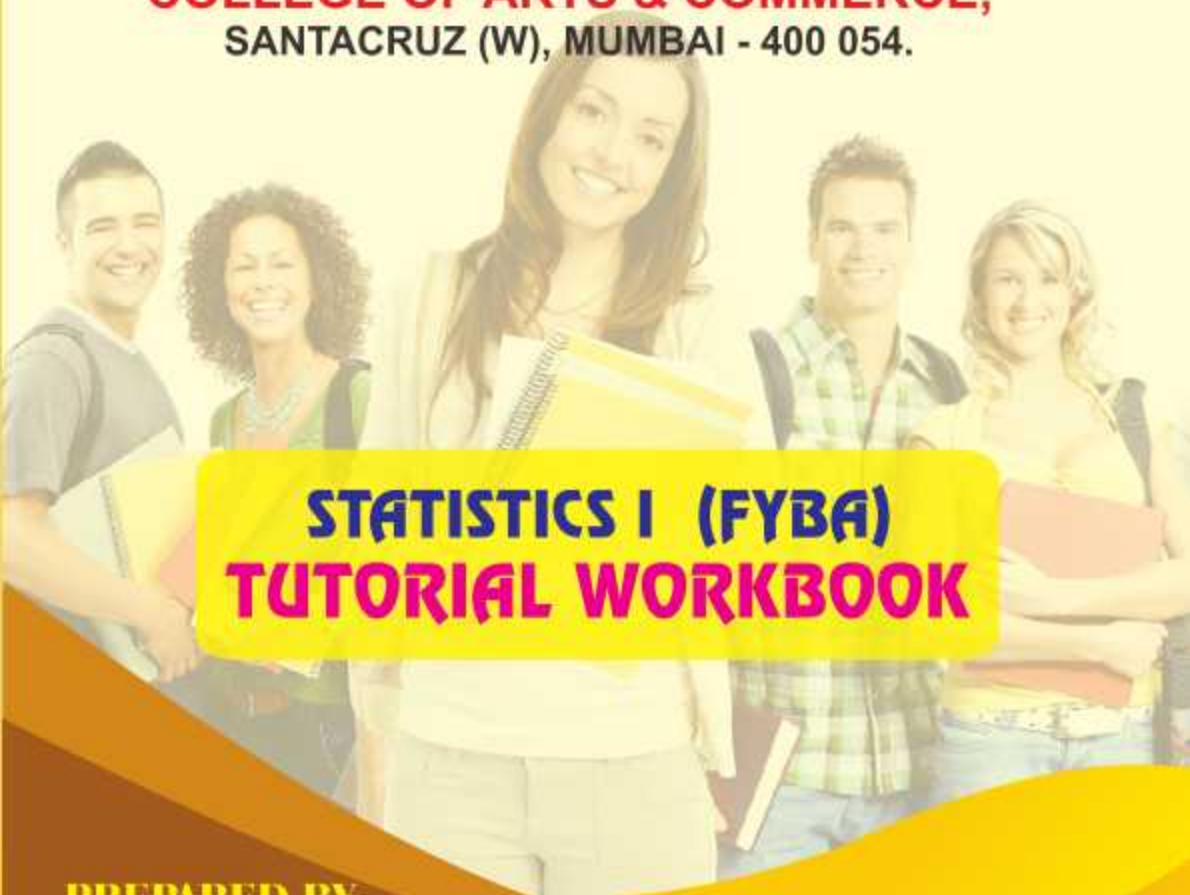# Sadhana Education Society

L. S. RAHEJA

COLLEGE OF ARTS & COMMERCE,
SANTACRUZ (W), MUMBAI - 400 054.

STATISTICS I (FYBA)
TUTORIAL WORKBOOK

PREPARED BY
Dr. Seema Ukidve
Chinmay Kishore Joshi

DEPARTMENT OF MATHEMATICS, STATISTICS & COMPUTERS

STATISTICS I

TUTORIAL WORKBOOK

PREPARED BY:
Dr. Seema Ukidve
Chinmay Kishore Joshi

Why this tutorial handbook is introduced?

"Mathematics is not about numbers, equations, computations or algorithms: it is about understanding."

_____**William Paul Thurston**

"Statistics is the grammar of science."

_____**Karl Pearson**

The Field of Statistics

The field of statistics is the science of learning from data. Statisticians offer essential insight in determining which data and conclusions are trustworthy. Statisticians know how to solve scientific mysteries and how to avoid traps that can trip up investigators.

When statistical principles are correctly applied, statistical analyses tend to produce accurate results. What's more, the analyses even account for real-world uncertainty in order to calculate the probability of being incorrect.

To produce conclusions that you can trust, statisticians must ensure that all stages of a study are correct. Statisticians know how to:

- Design studies that can answer the question at hand
- Collect trustworthy data
- Analyze data appropriately and check assumptions
- Draw reliable conclusions

It has been observed that students enrolling for F.Y.B.A lack basics of Mathematics and Statistics as some of them did not opt for Statistics in F.Y.J.C and S.Y.J.C, due to which they lose connect with mathematical concepts and rigour.

To boost the confidence of students and to make them understand Statistics lessons taught in the class and to provide them hand on practice of standard questions this tutorial handbook has been introduced.

This tutorial handbook contains:

- ✓ Latest Syllabus of Statistical techniques paper.
- ✓ Paper Pattern
- ✓ Reference Books
- ✓ Unit wise questions for practice with enough space to solve them
- ✓ Graph Papers

We hope this handbook will inculcate the problem solving aptitude among students and remove their Mathematics and Statistics phobia.

Why Revision?

There is a Rapid expansion of knowledge in subject matter areas and improved instructional method during last decade. There are considerable curricular revisions happening at the high school level. Application of Mathematics and Statistics are widely used in industry and business. Keeping this in mind, a revision of syllabus required in accordance with the growth of subject of at the high school level and emerging needs of industry and its application.

Objective:

The main objective of this course is to introduce mathematics and statistics to undergraduate students of commerce, so that they can use them in the field of commerce and industry to solve the real life problems.

Distribution of topics and lectures

SEMESTER I Course Code	Credits
Statistics I	**2 Credits (45 lectures)**
Unit I: Types of data and data condensation: Population, Sample, SRS, SRSWR, SRSWOR, types of scales, primary and secondary data, tabulation, Association of attributes, Yule's coefficient of association, Yule's coefficient of colligation.	**15 Lectures**
Unit II: Classification of data and measures of central tendency: Univariate frequency distribution, Bivariate frequency distribution, Cumulative frequency distribution, Graphical representation, Histogram, frequency polygon, Ogives, Steam and leaf diagram. Measures of central tendency: Requirement of good measure, Median, Mode, Partition Values, AM, GM, HM, Empirical relation between Mean, median and mode, Merits and demerits.	**15 Lectures**
Unit III: Measures of Dispersion, Skewness & Kurtosis: Dispersion, Requirements of good measure. Absolute and relative measures of dispersion, Standard Deviation, Variance, Combined variance, raw moments & Central moments and relationship between them, Their properties, Skewness and Kurtosis. Measures of Skewness and Kurtosis based on moments, Box plot.	**15 Lectures**

Assessment of Practical Core Courses Per Semester per course:

1. Semester work, Documentation, Journal _____ 10 Marks.

2. Semester End Practical Examination _____ 40 Marks

Semester End Examination Theory:

At the end of the semester, Theory examination of three hours' duration and 100 marks based on the three units shall be held for each course. Pattern of **Theory question** paper at the end of the semester for the course:

There shall be Five Questions of twenty marks each.

Question 1 based on all Three units. Ten sub-questions of two marks each.

Question 2 based on Unit I (Attempt any TWO out of THREE)

Question 3 based on Unit II (Attempt any TWO out of THREE)

Question 4 based on Unit III (Attempt any TWO out of THREE)

Question 5 based on all Three Units combined. (Attempt any TWO out of THREE)

Practicals:

At the end of the semester, Practical examination of **2** hours' duration and 40 marks shall be held for the course. Marks for term work in each paper should be given out of 10. (5 for viva and 5 for journal)

Pattern of **Practical question** paper at the end of the semester for **each course**:

There shall be Four Questions of ten marks each. Students should attempt all questions.

Question 1 based on Unit I,

Question 2 based on Unit II,

Question 3 based on Unit III,

Question 4 based on all Three Units combined.

Students should attempt **any two** sub questions out of the **three** in each Question.

Workload Theory: 3 lectures per week per course.

Practicals: 3 lecture periods per course per week per batch.

Descriptive Statistics – I

Unit I: Types of data and Data condensation

Q. 1. Answer the following questions

1. What are the merits and demerits of Tables?

2. Discuss the importance of classification in statistical analysis.

3. Write a short note on Association of attributes.

4. Write a short note on Measures of association.

5. What is contingency table? Give contingency tables for two attributes and three attributes.

Q. 2. Solve following problems:

1. The weights in kg of 30 workers in a factory are as follows:
    61.4, 67.0, 56.2, 71.0, 64.0, 69.0, 63.5, 63.0, 62.4, 72.5, 61.2, 64.5, 58.5,
    62.1, 64.0, 69.5, 58.0, 63.2, 50.0, 60.8, 64.5, 60.0, 57.5, 62.5, 57.6, 62.5,
    68.4, 52.6, 59.5, 63.0
    Classify and tabulate the data.

2. The marks obtained by 50 students of a class in a test are as follow:

26 49 33 42 53 35 47 39 61 43

44 31 59 43 25 52 45 39 48 57

52 68 47 38 33 43 45 29 46 37

34 44 28 54 48 36 42 57 36 49

32 43 59 47 27 51 38 69 64 57

Classify the data taking C.I. 20-29, 30-39, 40-49, 50-59, 60-69

3. Convert the class intervals 10-18, 20-28, 30-38, 40-48 into exclusive type.

4. The marks obtained by 100 students of a class in a certain subject are given below:

Marks	20-30	30-40	40-50	50-60	60-70	70-80
No. of students	8	20	30	24	12	6

    a)   Prepare Less than cumulative frequency table.

    b)   Prepare More than cumulative frequency table

    c)   Construct Histogram, frequency polygon, frequency curve, Ogives

    d)   Find the graphical value of mode.

5. Construct relative frequency table for the data

x:	0	1	2	3	4
f:	2	5	11	6	1

6. The following data gives the age in years 0f 60 staff member in an office. Prepare the relative frequency table.

Age in years	20-30	30-40	40-50	50-60	60-70
No. of students	7	15	20	12	6

7. N = 450, (A) = 265, (B) = 242, (C) = 250, (AB) = 135, (AC) = 132, (BC) = 144, (ABC) = 88. Find the other class frequencies.

8. From the following data find the remaining frequencies. In a survey of 400 people it was observed out of 84 lefties there were 48 males. In all there were 160 females.

9. From the following data find coefficient of association between Inoculation and affection.

	Affected	Not Affected
Inoculated	12	26
Non-Inoculated	16	6

Unit II: Classification of data & measures of Central Tendency

Q. 1. Define the following terms with one example:

1. Class interval

2. Class width

3. Exclusive class interval

4. Inclusive class interval

5. Open end class intervals

6. Class limits

7. Class boundaries

8. Class mark

9. Relative frequency distribution

10.Cumulative frequency distribution

11.Arithmetic mean

12. Weighted mean

13. Median

14. Mode

Q. 2. Answer the following questions.

1. Write a short note on Frequency distribution for discrete variable and continuous variable.

2. What is the cumulative frequency distribution? What are it's types.?

3. What are the advantages and disadvantages of graphs?

4. What are the requisites of good measures of central tendency?

5. How a mode can be located graphically by using Histogram?

6. Explain the mathematical relation between Mean, Mode and Median.

Q. 3. Write a short note on:

1. Histogram

2. Frequency Polygon

3. Frequency Curve

4. Ogives

5. Steam and leaf diagram

6. Types of Measures of central tendency

Q. 4. Solve following problems:

1. The following data gives distribution of income tax paid by 500 employees of a company.
   Find
   i. Q_1, Q_3, Quartile Deviation.
   ii. Standard Deviation.
   iii. Variance.
   iv. Coefficient of variance.

Income Tax (Rs.'000)	03-04	04-05	05-06	06-07	07-08	08-09	09-10

| No of employees | 40 | 60 | 80 | 120 | 100 | 60 | 40 |

2. Calculate Standard deviation, Variance, D_4, D_7, P_{60} & P_{30} from the following data.

Marks	20	25	30	35	45
No. of students	5	14	20	11	10

3. A hospital has collected data regarding number of days spent in hospital for a particular disease. Calculate the arithmetic mean for number of days spent in the hospital.

No. of days spent	1	2	3	4	5	6	7	8
No. of patients	5	6	5	10	8	4	3	2

4. Find the arithmetic mean, median and mode for the following data, representing marks of 80 students.

Marks	0-10	10-20	20-30	30-40	40-50
No. of students	12	13	21	19	15

5. The following data gives the consumption of electricity. Find the average consumption.

No. of units	0-200	200-400	400-600	600-800	800-1000	1000-2000	1200-1400
No. of consumers	9	18	27	35	32	28	11

6. The following frequency distribution represents weights in grams of mangoes of a given variety, in a box. Find the average weight of mangoes.

Weight in gms	410-419	420-429	430-439	440-449	450-459	460-469	470-479
No. Of mangoes	14	20	42	54	45	18	7

7. Find the range and median for the followings sets of observation 53, 43, 30, 55, 75, 50, 32, 39, 62.

8. Find the median for the followings data representing the age in the year of children.

Age in the years	3	4	5	6	7	8	9	10
No. of children	14	20	40	54	40	18	7	7

9. Following are the marks of the 15 students in a certain test. Find the modal marks.   18,22,25,42,39,35,25,33,34,25,29,37,35,25 and 40.

10. Find the mode of the following data, representing size of readymade pants.

Size of pants in cms	60	65	70	75	80	85	90
No. of pants	11	15	25	40	20	15	10

11. The following data gives frequency distribution of marks of some students. The average marks are 33. Using this find the numbers of students with marks between 40 and 50.

Marks	0-10	10-20	20-30	30-40	40-50	50-60
No. of students	5	10	25	30	---	10

Unit III: Measures of Dispersion, Skewness & Kurtosis

Q. 1. Define the following terms with one example:

1. Range

2. Coefficient of range

3. Quartile deviation

4. Coefficient of quartile definition

5. Mean deviation

6. Coefficient of variation

Q. 2. Answer the following questions.

1. What are the requisites of good measures of dispersion?

2. Define standard deviation as measure and the corresponding relative measure of dispersion.

3. Explain the effect of change of origin and scale on standard deviation.

Q. 3. Write a short note on:

1. Explain the difference between Absolute measures of dispersion and relative measures of dispersion.

2. State and prove properties of variance.

Q. 4. Solve the following problems:

1. The following data gives distribution of income tax paid by 500 employees of a company.
   Find

   v. Q_1, Q_3, Quartile Deviation.

   vi.   Standard Deviation.

   vii.   Variance.

viii.   Coefficient of variance.

Income Tax (Rs.'000)	03-04	04-05	05-06	06-07	07-08	08-09	09-10
No of employees	40	60	80	120	100	60	40

2. Calculate Standard deviation, Variance, D_4, D_7, P_{60} & P_{30} from the following data.

Marks	20	25	30	35	45
No. of students	5	14	20	11	10