Minutes of Meeting

GE31 and HS31

Date: 10th August, 2021
Venue: Online on Zoom between 15:00 to 16:30 Hours
Convener: Dr Vishruti Gupta
Semester: Third Semester

Course Code	HS31	GE31
Nature	Skill-Enhancement Elective Courses (SEC)	Generic Elective
Lectures	2 Theory + 4 Practicals	5 Theory + 2 Practicals

Attended by:

1	Dr. D Appala Naidu	ARSD
2	Shweta Nanda	ARSD
3	Dr. Deepika Goel	Aryabhatta College
4	N. Manichandra Singh	Aryabhatta College
5	Ashok Kumar	Daulat Ram College
6	Nitish Kashyap	Daulat Ram College
7	Rakesh Kumar	Dyal Singh College
8	Bhavna Seth	Dyal Singh College (M)
9	Siddharth Rathore	Gargi College
10	Sonam Aggarwal	Hansraj College
11	Nivedita Mullick	Hindu College
12	Manvi Jain	Indraprastha College for Women
13	Dr. Prabhjot Kaur	Indraprastha College for Women
14	Alisha George	Jesus and Mary College
15	Pummy	Kalindi College
16	Har Simrat Kaur	Lady Shri Ram College for Women
17	Surabhi Gupta	Lady Shri Ram College for Women
18	Dr. Hena Oak	Miranda House
19	Deepika Kandpal	PGDAV College (Morning)
20	Dr. Enakshi Sinha Ray Chaudhury	Rajdhani college
21	Dr. Apoorva Gupta	Ramjas College

22	Manisha Jayant	S. P. M. College
23	Dr. Kalpna Negi	S.P.M College
24	Mansi Sachdeva	Satyawati College (E)
25	Dr. Rupali Sharma	SGTB Khalsa College
26	Dr.Aditi Aeron Bansal	Shaheed Bhagat Singh College
27	Nikita Gupta	Shivaji college
28	Sushmitha Naidu	SRCC
29	Akhum David Longkumer	SRCC
30	Abhinav Parashar	Sri Aurobindo College (Evening)
31	Sneha Bhardwaj	Sri Venkateswara College
32	Yogita Yadav	Sri Venkateswara College
33	Amit Kumar Jha	Sri venkateswara college
34	Poonam Kalra	St. Stephen's College

Course Objective

This is a skill enhancement course for data analysis. The students will be given hands on training on using statistical and computing software to better visualize and understand data concepts.

The course is designed to be delivered through 2 classroom lectures and 4 computer lab classes per week.

Course Learning Outcomes

The course will use data simulations and publicly available data sources to help students learn about data types, their organization and visual representation. They will learn how to compute summary statistics and do some basic statistical inference.

Unit 1

Introduction to the course: How can the representation and analysis of data help us study real-world problems. Publicly available data sets.

- Levine *et. al* pgs 25-35, Chapter 1 Sections 1.1-1.3
- Devore, Chapter 1, pgs 32-33 (trimmed mean)
- Data base of Indian Economy, RBI (www.dbie.rbi.org.in), world bank data set
- On data and its representation and a large data set
- https://www.ted.com/talks/hans_rosling_the_best_stats_you_ve_ever_seen (Gapminder.org)
- Tattar *et. al.* Chapter 1 (pg 6 -7)

Unit 2

Using Data: Available statistical software, steps in data storage, organization and Cleaning

- Levine *et. al.* Chapter 1, Section 1.4 onwards; Chapter 2, Sections 2.1-2.2
- Gardener Chapter 1(pg1-24), Chapter 2(till pg61), Chapter 9 (pg 312-313)
- Tattar *et. al.* Chapter2 (pg 15-18, 41-46)

Unit 3

Visualization and Representation: Alternative forms of presenting summarizing and presenting data.

- Levine *et. al.* Chapter2, Section 2.3 onwards; Chapter3.
- Levine *et. al.* Chapter4, page 169
- Tattar *et. al.* Chapter 5 (pg105-109)
- Gardener, Chapter 5 (pg 154-158), Chapter 7 (pg215-217)

Unit 4

Simple estimation techniques and tests for statistical inference. Hypothesis testing.

- Levine *et. al.* Chapter5 (pg 199,203-4 and relevant functions from the Excel Guide of Chapter 5)
- Levine *et. al.* Chapter6 (pg 225, 228-229 and relevant functions from the Excel Guide of Chapter 6)
- Levine *et. al.* Chapter7; Chapter8, Sections 8.1-8.4 and pg 292-293; Chapter9; Chapter10, Sections 10.1,10.4, Pg361- Summary onwards, Relevant parts of Excel guide
- Gardener Chapter 6 (pg 181-183)

*For Levine et al. refer to the relevant sections on MS Excel at the end of the respective chapters.

References

1. Levine, D., Stephan, D., Szabat, K. (2017). *Statistics for managers using Microsoft Excel*, *8th ed*. Pearson.

2. Tattar, P.*,* Ramaiah, S., Manjunath, B. (2018). *A course in statistics with R.* Wiley.

3. Mark Gardener: Beginning **R The Statistical Programming Language,** Wiley (2012)

4. Jay L. Devore: Probability and Statistics for Engineering and the Sciences, 8th edition, Cengage Learning (2012)

5. Rstudio should also be downloaded for the treatment of data through the following link https://rstudio.com/, and for this book by Wickham will be helpful https://rstudio.com/resources/books/

Reference for Teachers

Hadley Wickham & Garrett Grolemund: R for Data Science, (2017).

Teaching Learning Process

Lectures and tutorials

Assessment Methods

Internal assessment and final examination as per CBCS rules

Internal Assessment will be worth 25 marks of which 5 marks will be for attendance and 20 marks for a hands-on project for the project the students are expected to use secondary sources of data available in public domain (eg. Indian economy data, World Bank data etc.) and analyse it using **at least** one of the software taught (Excel and R) preferably both(though for different parts of the project).

External Assessment will be 65 theory and 10 practical.

Points to be noted:

1. This course has been assigned 2 lectures and 4 practicals.

2. The college should provide 4 lab hours per week to conduct these practicals.

4. The University end of semester exam will be worth 75 marks which will be conducted in the lab. It will be a combination of theory and practical questions with greater weightage to the latter.

Minutes of the Meeting

A virtual meeting of the faculty members teaching the paper was held on Tuesday, August 10, 2021. The faculty members present in the meeting agreed on the following points:

1. The first 4 - 5 classes to be dedicate to introducing basics of R/ R-Studio/R-Cloud and Microsoft Excel. Certain commands could be taught in these classes to help the students get started. These may not necessarily be application of the statistical concepts.
2. This year following topics will be covered in the current syllabus

    a. Basic concepts of probability, Simple Probability, Joint Probability, and the General Addition Rule in Unit 3 (Levine *et. al.* Chapter4) Page 169
    b. Discrete Probability Distribution, Binomial Distribution, Poisson Distribution, Uniform Distribution and Normal Distribution in Unit 4 (Levine *et. al.* Chapter5 and Chapter6) Chapter 5 (pg 199,203-4 and relevant functions from the Excel Guide of Chapter 5) Chapter 6(pg 225, 228-229 and relevant functions from the Excel Guide of Chapter 6)
3. The coverage of R would be extended beyond basic this year. Application of the concepts to be taught both in Excel and R.
4. There would be suggested list of packages/commands (preferably compatible with R- Cloud) which would be taught in the course.

5. It was decided in the meeting that the teacher would prepare a list of suggested commands/packages, so that teaching is uniform across colleges.
6. The evaluation criteria remains the same (as discussed in meeting on 11 November 2020). (Document attached).
7. With reference to point 4, if a student uses a different approach/package/ command in the examination and it is correct then the student would not be penalised. In other words, no penalty for using a correct method which is not taught in the class.
8. The end-of-semester exam should not rely on students having access to computers since this will not be uniformly true.
9. The evaluation and question paper should be balanced and test skills both in Excel and R.
10. Emphasis on interpretation of the results obtained both while teaching and evaluation.
11. GE31 and HS31 would have the same course outline and evaluation criteria.
12. Suggested Resources for the teachers for learning like:
    a. Datacamp
    b. Coursera
    c. FDP courses
13. Teaching is not restricted to the references above since there are a number of online resources available and these keep changing. The readings are to indicate topics covered. These could be supplemented and substituted with other material.
14. The faculty should keep soft copies of practical exam question paper and answer scripts for each student in records as evidence for the university.
15. PHSTAT – this is a paid add on (with the book). This would not be used in the course.
16. Versions in Excel and Google Sheets. Excel 2007 Kindly ensure that the command gives desired result in various versions of Excel.

Data Analysis (SEC)
Minutes of the meeting
Wednesday 11.11.20 at 4 p.m

A virtual meeting of the faculty members teaching the paper was held on Wednesday, 11.11.20 to discuss the composition of theory and practical exams, the breakup for which has already been officially communicated as 65(Theory) and 10 (Practical).

The faculty members present in the meeting agreed on the following points:

1. The practical exam of 10 marks will be internally taken by the concerned faculty teaching the paper through online mode.
2. The practical component should test students' for 'hands on using Excel and R'.
3. The faculty should keep softcopies of practical exam question paper and answer scripts for each student in records as evidence for the university.

4. The house unanimously agreed to include theory questions including numerical for 25 marks and the remaining 40 marks for checking students' analytical skills and ability to interpret results probably through a small caselet or screenshot of excel workbook.
5. The members further discussed the nature of questions falling under the two categories as discussed above.
6. The question paper should be set incorporating the fact that students need not have a laptop facility to appear in the theory exam.
7. The members suggested to have an internal choice in the paper.
8. The weightage of marks for each unit remained undecided.
9. The members emphasized on sharing some of the sample questions in the interest of students.