DELHI SCHOOL OF ECONOMICS

DEPARTMENT OF ECONOMICS

Minutes of Meeting

Subject: B.A. (Prog) Economics Discipline, Fifth Semester (CBCS)
Course: Data Analysis (PS51), Skill-Enhancement Elective Courses (SEC) - Credit: 4
Date: 6th August, 2021
Venue: Online
Convener: Prof. Rohini Somanathan

The course meeting was attended by following teachers:

1. Dr Renu Kumari Verma – Motilal Nehru College Evening
2. Dr Appala Naidu – ARSD College
3. Rakesh Kumar – Dyal Singh College
4. Gautam Kumar Jha – PGDAV College (Eve)
5. Devangana Jha – Maitreyi College
6. Ajay Gupta- Shyamlal College Evening
7. Sukhvinder Kaur – SGND Khalsa College
8. Sakshi Jindal – Mata Sundri College for Women
9. Abhinav Parashar – Sri Aurobindo College Evening
10. Madhuri Singh – Kalindi College
11. Bhavna Seth – Dyal Sigh College
12. Dr Enakshi Sinha Ray Chaudhary – Rajdhani college
13. Mr Lalit- Vivekanand College
14. Rohit – Bhagini Nivedita College
15. Swati Yadav - Bhagini Nivedita College
16. Suneyana Sharma – Ram Lal Anand College
17. Rakhi Solanki - Sri Aurobindo College
18. Amrat Lal Meena - Motilal Nehru College
19. Akanksha Aggarwal – Jesus & Mary College
20. Vickey Mahariya – Maharaja Agrasen College
21. Akanksha Saini – Kamla Nehru College
22. Shweta nanda- ARSD College
23. Dr Arun Kumar- Delhi College of Arts and Commerce
24. Manavi Jain- IP College for Women
25. Rakhi Gupta- Deen Dayal Upadhyaya College
26. Loveleen Gupta- Bharti College
27. Sonu Kumar- Ramanujan College
28. Swarup Santra-Satyawati College
29. Jasmine- Jesus and Marry College
30. Nivedita Mullick – Hindu College
31. Dr Promila Sehrawat- Aditi Mahavidyalaya
32. Akshay Garg - PGDAV College
33. Vaishali Banshal

34. Divisha – Laxmibai College

Minutes of the Meeting:

A virtual meeting of the faculty members teaching the paper was held on Friday, August 06, 2021. The faculty members present in the meeting agreed on the following points:

1. Internal Assessment will be worth 25 marks of which 5 marks will be for attendance and 20 marks for a hands-on project for which the students are expected to use secondary data sources available in public domain (e.g. data sets from the RBI and the World Bank.) and analyze them using either of the two software packages used in the course: R and Microsoft Excel.
2. The University end-of-semester exam will be worth 75 marks, the breakup of which is 65 (Theory) and 10 (Practical). The practical exam of 10 marks will be conducted internally by the concerned faculty.
3. The faculty should keep soft copies of practical exam question paper and answer scripts for each student in records as evidence for the university.
4. The end-of-semester exam should not rely on students having access to computers since this will not be uniformly true. Students should be able to answer all questions using either software. (R or Excel).
5. Unit 6 will not be a part of the end-of-semester examinations. It will be only evaluated as part of internal assessment.
6. Teaching is not restricted to the references below 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.

Course Objectives

This is a skill enhancement course for data analysis. Students will be instructed on the use of spreadsheet and statistical software to analyse data. Software used for the course will be either MS Excel or R based on what is available.
The course is designed to be delivered through 4 classroom lectures per week.

Course Learning Outcomes

Students will learn to input, visually represent and analyse data. They will learn to compute summary statistics and do some basic statistical inference.

Unit 1

Introduction to available software and how it deals with data

Levine et al. Chapter 1 Sections 1.1-1.3; Appendix B: B.1 to B.5.

Tattar et. al. Chapter 1, Sections 1.1,1.2 (Page 1-5).

Gardener: Chapter 1 (pp. 1 – 9); Chapter 2 (except “Viewing Name Objects”)

Unit 2

Data cleaning: checking for outliers, cleaning variable names, consistency checks
Suggestion: Also include data organisation: categorical and numerical

Levine et. al. Chapter 1, Section 1.4 onwards; Chapter 2, Sections 2.1-2.2.

Tattar et. al. Chapter 1, Sections 1.4,1.5 and 1.6 (Page 6-10).

Abedin & Das: Chapter 1 (pp. 7 – 13, Variables, categorical and Numerical),

Gardener: Chapter 3 (pp. 82 – 85, Names)

Tattar: Chapter 14 (only section 14.3.2, for outlier)

van der Loo and Jonge: (pp. 1-8) [for Consistency Checks, data validation]

Unit 3

Data visualisation: scatter plots, line graphs, box plots and other graphical formats

Levine et. al. Chapter2, Section 2.3 to 2.5.

Tattar et. al. Chapter 2(page 15-18,29-31, 41-46).

Gardener: Chapter 7 (pp 215 – 239; Box plots, Scatter plots, Pair Plots (optional), Line chart, Pie chart), (pp. 245 – 256; Bar Chart)

Unit 4

Calculating and representing summary statistics and lines of best fit

Levine et. al. Chapter3.

Tattar et. al. Chapter 3, Section 3.1-3.6 (page 49-61), Chapter 4, Section 4.1-4.3(page 67-91).

Gardener: Chapter 4 (pp. 104 – 116; summary statistics), Chapter 10 (pp. 331 – 333; best fit line)

Unit 5

Elements of statistical inference: calculating and plotting confidence intervals; tests of population differences in population statistics

Levine et. al. Chapter7; Chapter8, Sections 8.1-8.4 and pp. 292-293; Chapter9; Sections 9.1 to 9.3, 9.6.

Gardener: Chapter 10 (pp. 330, 351-355; only confidence intervals)

Gardener: Chapter 6 (pp. 181-187; t-test).

Unit 6

Miscellaneous other topics: elements of writing simple programs for repetitive tasks, etc.

Suggestion: teaching how to automate tasks using Macro Recorder in MS Excel, how to create recurring tasks in MS Excel (available on Microsoft.com).

Gardener: Chapter 12 (pp. 417 – 420; Creating own function), (pp. 421 -428; Making Source Code).

*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. Jaynal Abedin and Kishor Kumar Das: Data Manipulation with R, 2nd Edition (2015), PACKT Publishing

5. Mark P. J. van der Loo and Edwin de Jonge: Data Validation Infrastructure for R, *Journal of Statistical Software,* Vol. 97, No. 10. [https://cran.r-project.org/web/packages/validate/vignettes/JSS_3483.pdf]