CP–201:DATA SCIENCE &BUSINESS ANALYTICS (CBCS)

Objectives

The objective of this paper is to make the students familiar with basic techniques used in Data Science. With the advancement in technology of storage capacity and computational power, now it is possible to deal with huge amounts of data. Innovations in social network analysis techniques and machine learning (‘big data’) have opened up a new area of methodological research. This course offers a review of these techniques in conceptual level, because they offer new ways of research design for social scientists.The main focus, however, is in their applications in business decision-making.

Course Contents	**L**	**T**	**P**
1. Data Analytics: Data & Information, Population & Sample, Types ofsampling, Sample size determination, Scaling & Measurement, Selection of appropriate test for hypothesis.Time Series, Cross Section and Panel. Organization and presentation of data. Measures of Central Tendencies, Dispersion, Skewness & Kurtosis. Co-variance & Correlation. Curve fitting and method of Least Square, Regression Equation, and Coefficient of Determination.	13	0	0
2. Data Base Management System: What is database system, purpose of database system, view of data, relational databases, database architecture, transaction management, The importance of data models, Basic building blocks, Business rules, The evolution of data models, Degrees of data abstraction.Database design and ER Model: overview, ER-Model, Constraints, ER-Diagrams, ERD Issues, weak entity sets, Codd’s rules, Relational Schemas, Introduction to UML,Relational Database design: features of good relational database design, atomic domain and Normalization (1NF, 2NF, 3NF, BCNF).	14	0	0
3. Advanced Data Science: Big Data: Paradigm Shift, Cloud Computing, Cloud operating system, Hadoop architecture, HDFS and YARN, Map Reduce programming model and scheduling algorithms, Spark for lightning fast cluster computing, Spark Streaming for real-time streaming data processing, Big Data security and privacy challenges in the cloud, Big Data Applications.	13	0	0

Suggested Readings

1. Provost, F. and Fawcett T. (2013), “Data Science for Business”, O’Rielly, Sebastopol, CA.
2. Chandha, N.K. Statistics for Behavioral and Social Scientists, Reliance Publishing House, Delhi.
3. Big data. Architettura, tecnologie e metodi per l'utilizzo di grandibasi di dati, A. Rezzani, Apogeo Education, 2013
4. Mayer-Schonberger, V. and Cukier, K. (2014), “Big Data”, Mariner Books, Boston, MA.
5. An introduction to Database Systems, C J Date, Addition-Wesley.
6. Database System Concepts, Abraham Silberschatz, Henry F. Korth& S. Sudarshan, McGraw Hill.
7. Gupta, S.P and Gupta M.P. Business Statistics, Sultan Chand, New Delhi.

The list of cases and specific references including recent articles and reports will be announced in the class at the time of launching of the course.