



**YESHWANTRAO CHAVAN COLLEGE OF ENGINEERING**  
(An Autonomous Institution affiliated to R T M Nagpur University Nagpur)  
Accredited by NAAC (1<sup>st</sup> Cycle) with 'A' Grade (Score 3.25 on 4 Point Scale)

Wanadongri, Hingna Road, Nagpur-441110

**Department of Computer Technology**  
**[Honors in BSDA (NPTEL Course)]**



**B.E. Honors in Business and Statistical Data  
Analytics**  
**SoE & Syllabus 2021-22**



Nagar Yuwak Shikshan Sanstha's

# Yeshwantrao Chavan College of Engineering

(An Autonomous Institution affiliated to Rashtrasant Tukadoji Maharaj Nagpur University)

## Department of Computer Technology

### SoE and Syllabus

#### B.E Honors in Business and Statistical Data Analytics (NPTEL Course)

SoE No.  
HON-101

### B.E Honors in Business and Statistical Data Analytics Information Brochure of Honor Program

1. Title of Program: **Honors Course in Business and Statistical Data Analytics**
2. Type of Program: **Honor**
3. Department offering the program: Computer Technology.
4. Collaboration: NPTEL / SWAYAM based
5. Department/s eligible to opt for the program:

**Students of Department of Computer technology are only eligible to opt for this program. Department of Computer Technology**

6. General information about courses in program:

83% of India's data-driven companies are more resilient and confident during the pandemic than non-data-driven companies (The Economic Times, 2020). This statement very well justifies the need of understanding the business analytics.

Business Analytics is the process by which businesses use statistical methods and technologies for analyzing historical data in order to gain new insight and improve strategic decision-making.

Statistical data analysis can be considered as the pillar for any type of data analysis. Statistical data analysis is a procedure of performing various statistical operations. It is a kind of quantitative research, which seeks to quantify the data, and typically, applies some form of statistical analysis.

This program is designed not only to explain what each model does or functions but also explores how businesses use them, whether it is to gather insights, solve problems, or predict outcomes.

This program tries to cover maximum courses which can prepare the student to work in this domain. Along with the course on Business Statistics, this program also includes Big Data Computing and Algorithms for Big Data. It also covers the Descriptive Statistics with 'R' Software and Data Analytics with Python.

The program is designed with the courses from NPTEL; hence student will get flexibility in attending the classes and the assessment process.

7. Advance knowledge or research orientation of Program:

The scope of the contents of this program is from some basic concepts in data analysis to the advance data analysis, business analytics and big data analytics. The course on big data gives the advance knowledge in the field of data science. Also, the courses from the program will help to solve the research problems in the field of business analytics and data science.

		May 2021	1.00	Applicable for AY2021-22 Onwards
Chairperson	Dean (Acad. Matters)	Date of Release	Version	



Nagar Yuwak Shikshan Sanstha's

# Yeshwantrao Chavan College of Engineering

(An Autonomous Institution affiliated to Rashtrasant Tukadoji Maharaj Nagpur University)

## Department of Computer Technology SoE and Syllabus

SoE No.  
HON-101

### B.E Honors in Business and Statistical Data Analytics (NPTEL Course)

#### 8. Employability potential of program:

Business analytics, widely known for its diagnostic, predictive and prescriptive prowess, has gained tremendous traction in the last year as governments and healthcare institutions embraced this progressive approach to overcome organizational obstacles and provide long-lasting solutions.

As per Mckinsey, business analytical capabilities are projected to create \$9.5 trillion to \$15.4 trillion annual economic impacts across nineteen industries worldwide, of which 40% can be credited to AI implementation.

Business analytics is expected to significantly impact marketing, sales, customer experience, finance, risk management, HR, and social media practices, among others.

Business professionals with an in-depth understanding of business analytics are better positioned to be industry leaders.

This program equip the students with the working knowledge of Business Analytics and Statistical Data Analysis.

#### 9. Departmental Steering committee: For proper publicity / conduct of program

SN	Name of the Faculty Member	Post	Designation	e-mail ID	Contact Number
1	Dr. G. M. Dhopavkar	HoD & Chairman	Asst. Prof.	hod_ct@ycce.edu	9822087970
2	Dr. P. A. Deshkar	Member	Asst. Prof.	padeshkar@ycce.edu	9923401052
3	Dr. K. R. Singh	Member	Asso. Prof.	singhkavita19@gmail.com	8275783031
4	Dr. S. D. Kamble	Member	Asso. Prof.	shailesh_2kin@rediffmail.com	9158886477
5	Dr. R. D. Wajgi	Member	Asst. Prof.	rdwajgi@ycce.edu	9970238062
6	Dr. L. B. Damahe	Member	Asst. Prof.	lalitdamahe3379@ycce.edu	9823289971
7	Prof. N. M. Mangrulkar	Member	Asst. Prof.	nmangrulkar@ycce.edu	7767888776

#### 10. Program Coordinator:

SN	Name of the Faculty Member	Post	Designation	e-mail ID	Contact Number
1	Dr. Prarthana A. Deshkar	Coordinator	Asst. Prof.	padeshkar@ycce.edu	9923401052

		May 2021	1.00	Applicable for AY2021-22 Onwards
Chairperson	Dean (Acad. Matters)	Date of Release	Version	



Nagar Yuwak Shikshan Sanstha's

# Yeshwantrao Chavan College of Engineering

(An Autonomous Institution affiliated to Rashtrasant Tukadoji Maharaj Nagpur University)

## Department of Computer Technology

### SoE and Syllabus

#### B.E Honors in Business and Statistical Data Analytics (NPTEL Course)

SoE No.  
HON-101

### Honors in Business and Statistical Data Analytics (NPTEL Course)

SN	Sem	Sub Code	Subject	T/P	Credit Hours				Credits	% Weightage				ESE
					L	T	P	Hrs		MSEs	TA	CA	ESE	Duration Hours
1	5	CTHN01	Data Analytics with Python	T					3	<p>This is SWAYAM / NPTEL based program and COURSES with 12–14-week syllabus are expected to be available on SWAYAM/NPTEL platform.</p> <p>If they are not available before the commencement of semester, Similar / Equivalent Subjects shall be notified by BoS of the Department.</p> <p>Chairman BoS will notify all the subjects which are 12-14 week duration before the commencement of academic session.</p>				
2	5	CTHN02	Business Statistics	T					3					
3	6	CTHN03	Big Data Computing	T					3					
4	6	CTHN04	Scalable Data Science	T					3					
5	7	CTHN11	Algorithms for Big Data	T					3					
6	7	CTHN12	Descriptive statistics with R software	T					3					
<b>TOTAL</b>									<b>18</b>					

**MSEs\* = Three MSEs of 15 Marks each will conducted and marks of better 2 of these 3 MSEs will be considered for Continuous Assessment**

**TA \*\* = for Theory : 20 marks on lecture quizzes, 8 marks on assignments, 2 marks on class performance**

**TA\*\* = for Practical : MSPA will be 15 marks each**

		May 2021	1.00	Applicable for AY2021-22 Onwards
Chairperson	Dean (Acad. Matters)	Date of Release	Version	



Nagar Yuwak Shikshan Sanstha's

# Yeshwantrao Chavan College of Engineering

(An Autonomous Institution affiliated to Rashtrasant Tukadoji Maharaj Nagpur University)

## Department of Computer Technology

### SoE and Syllabus

#### B.E Honors in Business and Statistical Data Analytics (NPTEL Course)

SoE No.  
HON-101

## Syllabus of Honors in Business and Statistical Data Analytics (NPTEL Course)

### V Semester

CTHN01	Data Analytics with Python						Credits = 3
Evaluation Scheme	MSE-I*	MSE-II*	MSE-III*	TA	ESE	Total	ESE Duration
*Best Two out of three MSE's would be considered							

#### Prerequisites

#### Contents of the course:

Introduction to data analytics and Python fundamentals  
Introduction to probability  
Sampling and sampling distributions  
Hypothesis testing  
Two sample testing and introduction to ANOVA  
Two-way ANOVA and linear regression  
Linear regression and multiple regression  
Concepts of MLE and Logistic regression  
ROC and Regression Analysis Model Building  
C Test and introduction to cluster analysis  
Clustering analysis  
Classification and Regression Trees (CART)

		May 2021	1.00	Applicable for AY2021-22 Onwards
Chairperson	Dean (Acad. Matters)	Date of Release	Version	



Nagar Yuwak Shikshan Sanstha's

# Yeshwantrao Chavan College of Engineering

(An Autonomous Institution affiliated to Rashtrasant Tukadoji Maharaj Nagpur University)

## Department of Computer Technology

### SoE and Syllabus

#### B.E Honors in Business and Statistical Data Analytics (NPTEL Course)

SoE No.  
HON-101

### V Semester

CTHN02	Business Statistics						Credits = 3
Evaluation Scheme	MSE-I*	MSE-II*	MSE-III*	TA	ESE	Total	ESE Duration
*Best Two out of three MSE's would be considered							

#### Prerequisites

#### Contents of the course:

Introduction, data collection and presenting data in tables  
Numerical descriptive measures and basic probability  
Discrete and continuous probability distributions  
Sampling and sampling distributions  
Confidence interval estimation  
One sample tests and hypothesis testing  
Two sample tests means  
Two sample tests proportions and variance  
ANOVA  
Chi-Square tests  
Simple linear regression  
Multiple regression basics

		May 2021	1.00	Applicable for AY2021-22 Onwards
Chairperson	Dean (Acad. Matters)	Date of Release	Version	





Nagar Yuwak Shikshan Sanstha's

# Yeshwantrao Chavan College of Engineering

(An Autonomous Institution affiliated to Rashtrasant Tukadoji Maharaj Nagpur University)

## Department of Computer Technology

### SoE and Syllabus

#### B.E Honors in Business and Statistical Data Analytics (NPTEL Course)

SoE No.  
HON-101

### VI Semester

CTHN03	<b>Big Data Computing</b>						Credits = 3
Evaluation Scheme	MSE-I*	MSE-II*	MSE-III*	TA	ESE	Total	ESE Duration
*Best Two out of three MSE's would be considered							

#### Prerequisites

#### Contents of the course:

Introduction to Big Data: Why Big Data and Where did it come from?, Characteristics of Big Data- Volume, Variety, Velocity, Veracity, Valence, Value, Challenges and applications of Big Data

Introduction to Enabling Technologies for Big Data, Introduction to Big Data Stack, Introduction to some Big Data distribution packages

Introduction to Big Data Platforms, Overview of Apache Spark, HDFS, YARN, Introduction to MapReduce, MapReduce Programming Model with Spark, MapReduce Example: Word Count, Page Rank etc.

Introduction to Big Data Storage Platforms for Large Scale Data Storage, CAP Theorem, Eventual Consistency, Consistency Trade-O-s, ACID and BASE, Introduction to Zookeeper and Paxos, Introduction to Cassandra, Cassandra Internals, Introduction to HBase, HBase Internals

Introduction to Big Data Streaming Platforms for Fast Data, Introduction to Big Data Streaming Systems, Big Data Pipelines for Real-Time computing, Introduction to Spark Streaming, Kafka, Streaming Ecosystem

Introduction to Big Data Applications (Machine Learning), Overview of Big Data Machine Learning, Mahout

Introduction, Big Data Machine learning Algorithms in Mahout- kmeans, Naïve Bayes etc.

Introduction of Big data Machine learning with Spark, Big Data Machine Learning Algorithms in Spark-

Introduction to Spark MLlib, Introduction to Deep Learning for Big Data

Introduction to Big Data Applications (Graph Processing), Introduction to Pregel, Introduction to Giraph, Introduction to Spark GraphX

		May 2021	1.00	Applicable for AY2021-22 Onwards
Chairperson	Dean (Acad. Matters)	Date of Release	Version	



Nagar Yuwak Shikshan Sanstha's

# Yeshwantrao Chavan College of Engineering

(An Autonomous Institution affiliated to Rashtrasant Tukadoji Maharaj Nagpur University)

## Department of Computer Technology

### SoE and Syllabus

#### B.E Honors in Business and Statistical Data Analytics (NPTEL Course)

SoE No.  
HON-101

### VI Semester

<b>CTHN04</b>	<b>Scalable Data Science</b>						<b>Credits = 3</b>
<b>Evaluation Scheme</b> *Best Two out of three MSE's would be considered	<b>MSE-I*</b>	<b>MSE-II*</b>	<b>MSE-III*</b>	<b>TA</b>	<b>ESE</b>	<b>Total</b>	<b>ESE Duration</b>

#### Prerequisites

#### Contents of the course:

Background: Introduction Probability: Concentration inequalities Linear algebra: PCA, SVD  
Optimization: Basics, Convex, GD Machine Learning: Supervised, generalization, feature learning, clustering.

Memory-efficient data structures: Hash functions, universal / perfect hash families Bloom filters  
Sketches for distinct count Misra-Gries sketch Statistical Mechanics an overview.

Memory-efficient data structures (contd.): Count Sketch, Count-Min Sketch Approximate near neighbors search: Introduction, kd-trees etc LSH families, MinHash for Jaccard, SimHash for L2.

Approximate near neighbors search: Extensions e.g. multi-probe, b-bit hashing, Data dependent variants  
Randomized Numerical Linear Algebra Random projection.

Randomized Numerical Linear Algebra CUR Decomposition Sparse RP, Subspace RP, Kitchen Sink.

Map-reduce and related paradigms Map reduce - Programming examples - (page rank, k-means, matrix multiplication) Big data: computation goes to data. + Hadoop ecosystem.

Map-reduce and related paradigms (Contd.) Scala + Spark Distributed Machine Learning and  
Optimization: Introduction SGD + Proof.

Distributed Machine Learning and Optimization: ADMM + applications Clustering Conclusion.

		May 2021	1.00	Applicable for AY2021-22 Onwards
Chairperson	Dean (Acad. Matters)	Date of Release	Version	





Nagar Yuwak Shikshan Sanstha's

# Yeshwantrao Chavan College of Engineering

(An Autonomous Institution affiliated to Rashtrasant Tukadoji Maharaj Nagpur University)

## Department of Computer Technology

### SoE and Syllabus

#### B.E Honors in Business and Statistical Data Analytics (NPTEL Course)

SoE No.  
HON-101

### VII Semester

CTHN11	Algorithms for Big Data						Credits = 3
Evaluation Scheme	MSE-I*	MSE-II*	MSE-III*	TA	ESE	Total	ESE Duration
*Best Two out of three MSE's would be considered							

#### Prerequisites

#### Contents of the course:

Intro to Probability Theory  
Tail bounds with Applications  
Markov Chains and Random Walks  
Randomized Algorithms against an Oblivious Adversary  
Pairwise Independence and Universal Hashing  
The Streaming Model  
Approximate Counting  
Approximate Median  
Flajolet Martin -- Distinct Sampling  
Alon-Mattias-Szegedy Sketch  
Bloom Filters  
Count-min Sketch  
Property Testing Model  
Local search and testing connectivity  
Enforce and Test Technique: Biclique and Bipartiteness Testing  
Random Walks and Testing Bipartiteness & Expansion  
Regularity Lemma and Testing Triangle Freeness  
Boolean Functions, BLR test for Linearity.

		May 2021	1.00	Applicable for AY2021-22 Onwards
Chairperson	Dean (Acad. Matters)	Date of Release	Version	



Nagar Yuwak Shikshan Sanstha's

# Yeshwantrao Chavan College of Engineering

(An Autonomous Institution affiliated to Rashtrasant Tukadoji Maharaj Nagpur University)

## Department of Computer Technology

### SoE and Syllabus

#### B.E Honors in Business and Statistical Data Analytics (NPTEL Course)

SoE No.  
HON-101

### VII Semester

CTHN12	Descriptive statistics with R software						Credits = 3
Evaluation Scheme	MSE-I*	MSE-II*	MSE-III*	TA	ESE	Total	ESE Duration
*Best Two out of three MSE's would be considered							

#### Prerequisites

#### Contents of the course:

Calculations with R Software  
Introduction to Descriptive Statistics, frequency distribution  
Graphics and Plots  
Central Tendency of Data  
Variation in Data  
Moments, Association of Variables  
Association of Variables  
Association of Variables, Fitting of Linear Models

		May 2021	1.00	Applicable for AY2021-22 Onwards
Chairperson	Dean (Acad. Matters)	Date of Release	Version	