

**SVKM's NMIMS**  
**Mukesh Patel School of Technology Management & Engineering**

<b>Program: B. Tech Data Science (Business Analytics)</b>				<b>Semester: III</b>	
<b>Course/Module : Managing Uncertainty</b>				<b>Module Code: BTDS03008</b>	
<b>Teaching Scheme</b>				<b>Evaluation Scheme</b>	
<b>Lecture (Hours per week)</b>	<b>Practical (Hours per week)</b>	<b>Tutorial (Hours per week)</b>	<b>Credit</b>	<b>Internal Continuous Assessment (ICA) (Marks -50)</b>	<b>Term End Examinations (TEE) (Marks -100 in Question Paper)</b>
3	2	0	4	Marks Scaled to 50	Marks Scaled to 50
<b>Objective:</b> <ul style="list-style-type: none"> <li>Students will be familiar with basic rules of probability and will be able to use them in modeling uncertainty in obtaining and recording data. They will be able to utilize graphical and numerical summaries of data in understanding data generating processes. To enable the students to analyze data more effectively using MS Excel</li> <li>To increase the student's ability in problem solving</li> </ul>					
<b>Outcomes:</b> After completion of the course, students would be able to : <ul style="list-style-type: none"> <li>Students will be able to learn basic probability and statistics and apply them to the analysis of real data sets from business fields.</li> </ul>					
<b>Detailed Syllabus: (per session plan)</b>					
<b>Unit</b>	<b>Description</b>				<b>Duration</b>
1	<b>(1) Classification and tabulation of Data:</b> Meaning and objective of classification, Types of classification, formation of discrete and continuous distribution.				2
	<b>(2) Data Classification and Data Presentation :</b> Histogram, Frequency distribution, Quantitative Data Graphs (Histograms , Frequency Polygons, Ogives, Dot Plots, Stem-and-Leaf Plots) ; Qualitative Data Graphs (Pie Charts ,Bar Graphs, Pareto Charts ) ; Graphical Depiction of Two-Variable Numerical Data: Scatter Plots				4
	<b>(3) Descriptive Statistics:</b> Measures of Central Tendencies – Grouped and Ungrouped Data; Mean, Sample Mean– Weighted mean, Geometric Mean, Harmonic Mean; Median – Quartiles, Deciles, and Percentiles; Mode, Box Plot;				4
	Measures of Variability– Dispersion, Range, Standard deviation, Chebyshev's theorem; Population v/s sample variance and standard deviation, Skewness; Kurtosis.				4



Signature  
 (Prepared by Concerned Faculty/HOD)

**SVKM's NMIMS**  
**Mukesh Patel School of Technology Management & Engineering**

2	<b>(1) Probability Distribution:</b> Introduction to Probability and Probability Distribution: Methods of Assigning probabilities, Probability Space, conditions of probability model, Events, simple and compound, Laws of probability, Probability density function, Cumulative distribution function, Expected values of Mean and Variance. Marginal, union, joint and conditional probabilities, Bayes' Theorem	4
	<b>(2) Discrete Probability Distribution:</b> Basics of Binomial Distribution pdf, Multinomial Distribution, Negative Binomial Distribution, cdf, Poisson Distribution pdf, cdf, Hypergeometric Distribution pf, cdf. Continuous Probability Distributions: Relative frequency, distribution and pdf, Exponential pdf, cdf, Normal distribution – Normal pdf, cdf, Standard Normal Distribution, Normal Approximation to the binomial.	4
3	<b>Sampling Distribution:</b> (1) Introduction, Central Limit Theorem; Population frequency distribution vs. Sampling Distributions; Sampling distribution of the sample mean – Estimating population Mean and Standard deviation; Sampling distribution of the sample proportion – Estimating proportions in a binomial population.	4
	(2) Sampling distributions of the difference between sample means/ proportion – Mean and Standard Deviation, Conditions for estimation.	4
4	<b>Hypothesis Testing:</b> (1) Large Sample estimation of the population parameters and Hypothesis testing: Basics of Estimating the populations mean and difference; estimating the proportion and difference; large sample test for population mean, difference; large sample test for proportion, difference.	4
	(2) Estimation of a population variance: Sampling distribution of variance, estimation.	4
	(3) Inferences from small sample: Student's t distribution; Small sample t test for following – A population mean, A difference between two means, Confidence interval.	4



Signature  
(Prepared by Concerned Faculty/HOD)

**SVKM's NMIMS**  
**Mukesh Patel School of Technology Management & Engineering**

	(4) Rejection and Non-rejection region, Type I and Type II errors, testing hypothesis about a population mean using the Z- statistic, using p-values to test Hypothesis	3
Total		45
Prescribed Text :		
<div>1. Richard, L &amp; David, R. (2013). <i>Statistics For Management</i>, Pearson</div> <div>2. Gupta, S. P (2012). <i>Statistical Methods</i>. Sultan Chand &amp; Sons</div>		
References:		
<div>1. Gujarati, D (2011). <i>Basic Econometrics</i>. McGraw Hill</div> <div>2. William, M. (1993). <i>Statistics for Management and Economics</i>. Duxbery Press</div> <div>3. Ken Black (2010). <i>Business Statistics</i>. E-book</div>		
Internet references: NIL		
Any other information: NIL		
Total Marks of Internal Continuous Assessment (ICA): 50 Marks		
Distribution of ICA Marks:		
Description of ICA	Marks	
Test Marks	20	
Term Work Marks	30	
Total Marks :	50	
Details of Term work:		
<div>• Class Test/Assignments/ Case Studies / Projects / Presentations</div>		



\_\_\_\_\_  
**Signature**  
 (Prepared by Concerned Faculty/HOD)

**SVKM's NMIMS**  
**Mukesh Patel School of Technology Management & Engineering**

<b>Program: B. Tech. –Data Science (Business Analytics)</b>				<b>Semester : IV</b>	
<b>Course/Module: Statistical Method - I</b>				<b>Module Code: BTDS04009</b>	
<b>Teaching Scheme</b>				<b>Evaluation Scheme</b>	
<b>Lecture (Hours per week)</b>	<b>Practical (Hours per week)</b>	<b>Tutorial (Hours per week)</b>	<b>Credit</b>	<b>Internal Continuous Assessment (ICA) (Marks -50)</b>	<b>Term End Examinations (TEE) (Marks -100 in Question Paper)</b>
3	2	0	4	Marks Scaled to 50	Marks Scaled to 50
<b>Pre-requisite:</b> Managing Uncertainty (SEM - III)					
<b>Objectives:</b> <ul style="list-style-type: none"> <li>To provide advanced statistical background for analysing data and drawing inferences from that analysis</li> <li>Predicative Analytics using liner and generalized liner model</li> </ul>					
<b>Outcomes:</b> After completion of the course, students would be able to : <ul style="list-style-type: none"> <li>Students will be able to learn advanced statistical technique and apply them to the analysis of real data sets from different fields.</li> </ul>					
<b>Detailed Syllabus:</b>					
<b>Unit</b>	<b>Description</b>				<b>Duration</b>
1	<b>ANOVA/MANOVA:</b> Chi-Square as a test of independent, Chi-square as a Test of goodness of fit: Testing the Appropriateness of a Distribution, Analysis of Variance, Multivariate analysis of variance				3
2	<b>Regression Model:</b> <ol style="list-style-type: none"> <li>Least squares and linear regression: Introduction; Notation; Ordinary least squares; Regression to the mean; Linear regression; Residuals; Regression inference</li> <li>Multivariable regression: Multivariate regression; Multivariate examples; Adjustment; Residual variation and diagnostics; Multiple variables , Interaction Terms, Non-linear Transformations of the Predictors, Qualitative Predictors</li> <li>Multiple Regression Analysis: The Problem of Estimation and the Problem of Inference</li> <li>Dummy Variable Regression Models</li> <li>Multi-collinearity, Heteroscedasticity, Autocorrelation</li> <li>Econometric Modelling: Model Specification and Diagnostic Testing</li> <li>Correlation and Covariance Analysis</li> <li>Canonical Analysis, Canonical Roots/variates</li> </ol>				3 4 3 3 3 4

\_\_\_\_\_  
**Signature**  
 (Prepared by Concerned Faculty/HOD)

\_\_\_\_\_  
**Signature**  
 (Approved by Dean)

**SVKM's NMIMS**  
**Mukesh Patel School of Technology Management & Engineering**

3	Extension of regression analysis: Ridge Regression, The Lasso	5
	Nonlinear Regression Models: Approaches to <b>Estimating Nonlinear</b> Regression models	5
4	<b>Generalized linear models:</b>	3
	Logistic Regression, Binary outcomes, Count outcomes, Multiple Logistic Regression	3
<b>Total</b>		<b>45</b>
<b>Text Books:</b> <ol style="list-style-type: none"> <li>1. An Introduction to Statistical learning with application in R . Hastie T, Robert T. (2014). Springer Science Business Media: New York</li> <li>2. Gujarati, D (2011). Basic Econometrics. McGraw Hill</li> </ol>		
<b>Reference Books:</b> <ol style="list-style-type: none"> <li>1. Statistics for Management, Seventh Edition, by Richard I. Levin, David S. Rubin, Pearson</li> <li>2. An Introduction to Categorical Data Analysis. Agresti, A. (2012). John Wiley &amp; sons</li> <li>3. The Element of Statistical Learning, Data mining, Inference and Prediction. Hastie, T, Tibshirani, R, &amp; Friedman, J. (2011). New York: Springer Series in Statistics.</li> <li>4. Hair, Black, Babin, Anderson and Tatham (2009). Multivariate Data Analysis, Pearson</li> </ol>		
<b>Any other information: NIL</b>		
<b>Details of Internal Continuous Assessment (ICA):</b> <b>Test Marks: 20</b> <b>Term Work Marks: 30</b>		
<b>Details of Term work:</b> <ul style="list-style-type: none"> <li>• Practical based on 10 Experiments</li> <li>• Two class tests.</li> <li>• Minimum two assignments</li> </ul>		

\_\_\_\_\_  
**Signature**  
**(Prepared by Concerned Faculty/HOD)**

\_\_\_\_\_  
**Signature**  
**(Approved by Dean)**

**SVKM's NMIMS**  
**Mukesh Patel School of Technology Management & Engineering**

<b>Program: B. Tech Data Science (Business Analytics)</b>				<b>Semester: V</b>	
<b>Course/Module: Statistical Method - II</b>				<b>Module Code: BTDS05008</b>	
<b>Teaching Scheme</b>				<b>Evaluation Scheme</b>	
<b>Lecture (Hours per week)</b>	<b>Practical (Hours per week)</b>	<b>Tutorial (Hours per week)</b>	<b>Credit</b>	<b>Internal Continuous Assessment (ICA) (Marks -50)</b>	<b>Term End Examinations (TEE) (Marks -100 in Question Paper)</b>
3	2	0	4	Marks Scaled to 50	Marks Scaled to 50
<b>Pre-requisite:</b> Statistical Method-I					
<b>Objectives:</b> To introduce and provide some core and necessary data mining techniques so that students understand how to work with large data sets and apply the appropriate data mining technique to answer business questions					
<b>Outcomes:</b> After completion of the course, students would be able to: 1. Students will able to learn a number of well-defined data mining tasks such as classification, estimation, prediction, affinity grouping and clustering, and data visualization are discussed					
<b>Detailed Syllabus:</b>					
<b>Unit</b>	<b>Description</b>				<b>Duration</b>
1	<b>Classification:</b>				
	a) Using Bayes' Theorem for Classification, Procedure of Discriminant Analysis, Linear Discriminant Analysis, Estimating Misclassification Probabilities, Quadratic Discriminant Analysis				6
	b) Cluster Analysis: Measures of Association for Continuous Variables, Measures of Association for Binary Variables, Agglomerative Hierarchical Clustering, Ward's Method, K-Means Procedure, K-Nearest-Neighbours				6
	c) Principal Components Analysis (PCA) and Factor Model: Procedure Principal Component Analysis (PCA), Maximum Likelihood Estimation Method, Factor Rotations, Varimax Rotation, Estimation of Factor Scores.				6
2	Resampling Methods				6



\_\_\_\_\_  
**Signature**  
 (Prepared by Concerned Faculty/HOD)

**SVKM's NMIMS**  
**Mukesh Patel School of Technology Management & Engineering**

3	Tree-Based Methods: The Basics of Decision Trees, Bagging, Random Forests, Boosting	8
	Support Vector Machines: MaximalMargin Classifier, Support Vector Classifiers, SVMs with More than Two Classes	7
4	Conjoint analysis	6
<b>Total</b>		<b>45</b>

**Text Books:**

1. An Introduction to Statistical learning with application in R . Hastie T, Robert T. (2014). Springer Science Business Media: New York
2. Hair, Black, Babin, Anderson and Tatham (2009). Multivariate Data Analysis, Pearson

**Reference Books:**

1. Statistics for Management, Seventh Edition, by Richard I. Levin, David S. Rubin, Pearson
2. An Introduction to Categorical Data Analysis. Agresti, A. (2012). John Wiley & sons
3. The Element of Statistical Learning, Data mining, Inference and Prediction. Hastie, T, Tibshirani, R, & Friedman, J. (2011). New York: Springer Series in Statistics.
4. Gujarati, Damodar N, and Dawn C. Porter. Basic Econometrics. Boston, Mass: McGraw-Hill, 2009

**Any other information: NIL**

**Total Marks of Internal Continuous Assessment (ICA): 50 Marks**

**Distribution of ICA Marks:**

Description of ICA	Marks
Test Marks	20
Term Work Marks	30
<b>Total Marks :</b>	<b>50</b>

**Term Work:**

1. Practical based on 10 Experiments
2. Two class tests.
3. Minimum two assignments



\_\_\_\_\_  
**Signature**  
**(Prepared by Concerned Faculty/HOD)**