HSNC UNIVERSITY, MUMBAI

Board of Faculty of Science & Technology

Board of Studies in the Subjects of Statistics and Data Science& Business Analytics

1) NAME OFCHAIRPERSON/CO-CHAIRPERSON/COORDINATOR:-

a) **Dr Asha Jindal,**Associate Professor and Head of Department, Department of Statistics, K. C. college, HSNC University Churchgate, Mumbai –400 020. Email ID-asha.jindal@kccollege.edu.in Mobile no-9821235627

2) TWO TO FIVE TEACHERS EACH HAVING MINIMUM FIVE YEARS TEACHING EXPERIENCE AMONGST THE FULL TIME TEACHERS OF THE DEPARTMENTS, IN THE RELEVANTSUBJECT.

a) **Dr. S. B. Muley,**Assistant Professor, Department of Statistics, K. C. college, HSNC University Churchgate, Mumbai – 400 020. Email IDsakharam.muley@kccollege.edu.in, Mobile No-9323817918

b) **Mrs.PratikshaKadam,**Assistant Professor, Department of Statistics, K. C. college, HSNC University Churchgate, Mumbai – 400 020. Email IDpratiksha.kadam@kccollege.edu.in, Mobile No-7507162816

c) **Ms.ShailajaRane,**Assistant Professor, Department of Statistics, K. C. college, HSNC University Churchgate, Mumbai – 400 020. Email IDshailaja.rane@kccollege.edu.in, Mobile No-7506986359

3) ONE PROFESSOR / ASSOCIATE PROFESSOR FROM OTHER UNIVERSITIES OR PROFESSOR / ASSOCIATE PROFESSOR FROM COLLEGES MANAGED BY PARENTBODY;

a) **Dr AnjumAra Ahmed;** I/C Principal, Rizvi College, Mumbai. Email ID**anjumahmed8@gmail.com, Mobile No-** 8451046220

4) FOUR EXTERNAL EXPERTS FROM INDUSTRY / RESEARCH / EMINENT SCHOLAR IN THE FIELD RELEVANT TO THE SUBJECT NOMINATED BY THE PARENTBODY;

a. **Prof. Suresh Kumar Sharma,** Professor, Department of Statistics, Panjab University, Chandigarh.
Email IDssharma643@yahoo.co.in, **Mobile No**-9815911381

b. **Mr Mukesh Jain,**Chief Technological Officer, Capgemini. Email IDmdjain@hotmail.com, **Mobile No**-7972637347.

c. **Dr Santosh Gite,**Associate Professor, Dept. of Statistics, University of Mumbai, Mumbai. Email ID santgite@yahoo.com, **Mobile No-** 9167157717.

d. **Mr Prashant Kumar Nair,** Director, GeoSpatial Analytics Global Lead, Intelligent Analytics, Nielsen Connect,Email IDprashantkumar.nair@nielsen.com , **Mobile No**-9833747057.

5. Top rankers of the Final Year Graduate and Final Year Post Graduate examination of previous year of the concerned subject as invitee members for discussions on framing or revision of syllabus of that subject or group of subjects for one year.

a) **Ms.Mohaddasah Patel**(undergraduate student 18-19) EmailId-mohaddasah.98@gmail.com; Mobile no-9833781878

b) **Ms.Divya Srivastava**(undergraduate student18-19) EmailID-divyasrivastav20@gmail.com ; Mobile no-8879240305

B.Com. Honours Programme in Statistics

There is always scope for improvement. Perfection is a moving goal. Improvement is measured through business results. There is an unknown improvement sequence that will bring best results in your unique business situation. Statistical Data driven approaches help in finding root causes of problems and fixing them permanently.

This Course is based on motto of learning while doing and is of 20 Credits. There will be Sic Semesters. Each Semester course work is of 3 Credits. Two Credits are assigned for Internship in Companies like Reliance, Accenture, CapGeminietc in R & D Departments. A 15days internship of one credit should be undertaken in any two semesters out of Sem III, Sem IV and Sem V only and two internship cannot be completed in single semester.

Learning Outcomes:

i. Students will learn Statistics from basics concepts to creating basic models for predictions
ii. Students will learn how Statistics is actually used large corporations like Microsoft, Jio, Amazon and other top companies globally
iii. In the hands-on session, Students will gain skills on identifying opportunities for Analytics.
iv. Students will acquire analytics skill needed for R & D for smooth run of any business.
v. Students will learn to solve business problems using a structured approach to improve an Organisation's performance.

Semester I: **Data Collection and Visualisation with MS-Excel**

Course Code	Title	Credits
	Data Collection and Visualisation with MS-Excel	**3 Credits (45 Hours)**

Unit I: Types of Data and Data Condensation(15 Hours)

I. Installation---------- 2 Hours

Installation of add-in**MegaStat** in Excel Library fromhttps://megastat.software.informer.com and activation of **DataAnalysis ToolPak**.

II. Introduction, Data Types---------- 3 Hours

- Definition of Statistics, Application and scope of Statistics
- Concept of Population and Sample. Finite, Infinite Population, Parameter and Statistic.
- Types of data: Qualitative and quantitative data; Time series data and cross section data, discrete and continuous data.
- Different types of scales: nominal, ordinal, interval and ratio.
- Concept of population and sample. Finite and Infinite population

III. Primary data and Secondary Data---------- 4 Hours

concept of a questionnaire and a schedule, Secondary data and Collection

IV. Experiential Learning---------- 6 Hours

1) Designing of a questionnaire and/or a schedule
2) Collection and Preparation of Excel Response sheet

Unit II: Data Visualization (15 Hours)

I. **Introduction** ---------- 2 Hours

The philosophy of Data Visualization, Deconstructing and understanding the expression data visualization, Understanding how data analysis and visualization complement each other, A drill-down into what constitutes data with examples

II. Experiential Learning with Pivot Table---------- 4 Hours

An analysis of a "data table", The "Row View", The "Column View", What might interest a data analyst about rows i.e. instances or samples, What might interest a data analyst about columns i.e. attributes or variables with suitable examples 2 Hours

III. Visualisation---------- 2 Hours

Where does Data Visualization enter the frame of Data Analysis? , The motive force behind visualization – A PICTURE SPEAKS LOUDER THAN A THOUSAND WORDS, The interplay of variables that leads to visualization,Univariate analysis, Bivariate analysis , Multivariate analysis and their visualization

IV. The WHY, WHAT and HOW of Visualization---------- 5 Hours

Why visualize (what questions do we want visualization to answer),What is being visualized (the data)?, How should we visualize (the various types of plots and graphs) V. Experiential Learning with Univariate Visualization---------- 1) Constructing and interpreting: Histogram, Pie chart, Bar chart. Box and Whisker plot , stem leaf diagram 2) Going beyond Univariate Visualization: Pareto Chart, Fishbone Chart	
Unit III: Data Visualization(15 Hours)	
I. Pair Plots, Bivariate Visualization ---------- visualizing the relationship of two variables, The cause-effect relationship, Concept of X and Y variables, Concept of independent and dependent variables, Choice of pair of variables, which is X and which is Y, what questions will the combination answer	1 Hours
II. The matrix of 4 combinations---------- ➢ Categorical (X) vs Categorical (Y) ➢ Numeric (X) vs Numeric (Y) ➢ Categorical (X) vs Numeric (Y) ➢ Numeric (X) vs Categorical (Y)	1 Hours
III. Experiential Learning with Bivariate Visualization---------- 1) The Grouped Bar chart, The Stacked Bar chart, Ogives, Frequency Curve and Frequency Polygon 2) Going beyond Bivariate Visualization: Thermometer Chart, Funnel Chart, Waterfall chart, Pyramids, Power Maps 3) The Scatterplot: Using the scatterplot for bivariate visualization (two numeric variables), 4) More visualizations: Power Curve, Exponential Curve, Logarithmic Curve, Pareto Curve and relationship with Coefficient of Determination R^2. 5) Time series visualizations using the line chart	9 Hours
IV. Infographics---------- Concept, Types,Experiential Learning for storytelling using Excel deshboard and infographics.	4 Hours

Reference Books:

1. Medhi J.: Statistical Methods, An Introductory Text, Second Edition, New Age International Ltd.
2. Spiegel M.R.: Theory and Problems of Statistics, Schaum's Publications series. Tata McGraw-Hill.
3. Kothari C.R. : Research Methodology, Wiley Eastern Limited.

Reference Websites:

i). www.statsci.org/datasets.html

ii). www.statweb.calpoly.edu/bchance/stat-stuff.html

iii). www.amstat.org/publications/jse/jse-data-archive.html

iv). www.amstat.org/publications/chance

v). www.math.uah.edu/stat

vi). www.amstat.org/publications/stats

vii). www.stat.ucla.edu/cases

viii). https://data.gov.in/

ix). https://www.connectmath.com/training/MegaStat_User_Guide.pdf

x). https://youtu.be/bxr0iL6NI1g

xi). https://www.moresteam.com/toolbox/pareto-chart-manufacturing.cfm

xii). https://youtu.be/ySbhsX-y8zE

Semester II: **Data Analytics-I**

Course Code	Title	Credits
	Data Analytics- I **Introduction to Six Sigma with MS-Excel**	**3 Credits (45 Hours)**

Unit I : SIX SIGMA(10 Hours)

I. Introduction---------- 4 Hours

History and concept, Basic Principles, Goals, six sigma v/s TQM, ISO 9000, Traditional Management, Quality defined, VOC and CTQ, Quality measurement to six sigma.

II. Experiential Learning---------- 3 Hours

Seven tools of quality and its application: 1)Histogram or Stem and Leaf display. 2) Check sheet. 3) Pareto Chart. 4) Cause and Effect diagram (Fish bone Diagram) 5) Defect concentration diagram. 6) Scatter diagram. 7) Control charts (**Only concept of control chart**).

III. DMAIC with case study---------- 2Hours

IV. Introduction to Lean Six Sigma ---------- 1 Hour

Unit II : INTRODUCTION TO BASIC STATISTICS(15 Hours)

I. Descriptive Statistics---------- 2 Hours

Averages, Measures of Variation, Skewness, Kurtosis, Box and Whisker plots

II. Data Distribution---------- 2 Hours

Normal Distribution, CLT theorem, Sampling distribution of mean)

III. Hypothesis Testing---------- 1 Hour

IV. Experiential Learning ---------- 10 Hours

- Parametric Test: Applications of Large Sample Test, Applications of Small Sample Test,Application of Chi-Square Test and F test,
- *Non Parametric Test: Mann-Whitney U test, Kruskal-Wallis test, Moods median test,*
- Design of experiments*: One way and Two way ANOVA.*
- Regression analysis using Scatter Diagram.

Unit III : CONTROL CHARTS (20 Hours)		
I.	Introduction, Chance and assignable causes, Statistical basis of the control chart: Basic principles of control chart (Shewhart Control Charts), Choice of control limits	2 Hours
II.	Control chart for Attributes---------- Theory of P, np, c and u charts, p-chart with variable sample size, Experiential Learning: Plotting above charts and Interpretation, *Problems involving setting up standards for future use*	8 Hours
III.	Control chart for Variables---------- X-Bar, R, S[*sample standard deviation*] Experiential Learning: Plotting above charts and Interpretation, *Problems involving setting up standards for future use*	6 Hours
IV.	Introduction to process capability ---------- concept, Specification limits natural tolerance limits and their comparisions, estimate of percent defectives, Capability ratio and Capability indices (Cp), Capability performance indices Cpk with respect to machine and process interpretation, relationship between i. Cp and Cpk ii. Defective parts per million and Cp	4 Hours

References:

1) Fundamental of Mathematical Statistics, Gupta and Kapoor.
2) Probability and Random process by T. Veerarajan.
3) Six Sigma For Business Excellence, (2005), Penelope Przekop, McGraw-HillSix Sigma Handbook, by Pyzdek, McGraw Hill Education;4 edition (1 July 2017).
4) The Certified Six Sigma Green Belt Handbook, Roderick A. Munro and GovindarajanRamu , American Society for Quality (ASQ),
5) What Is Design For Six Sigma,(2005), Roland Cavanagh, Robert Neuman, Peter Pande, Tata McGraw-Hill.
6) The Six Sigma Way: How GE, Motorola, And Other Top Companies Are Honing Their Performance, (2000), Peter S. Pande, Robert P. Neuman, Roland R. Cavanagh, McGraw-Hill
7) What Is Lean Six Sigma,(2004), Mike George, Dave Rowlands, Bill Kastle, McGraw-Hill8.

8) Six Sigma Deployment,(2003), Cary W. Adams, Charles E Wilson Jrs, Praveen Gupta, Elsevier Science.
9) Six Sigma For Beginners: Pocket Book(2018), Rajiv Tiwari Kindle Edition
10) Introduction to Statistical Quality Control(2009), Montgomery, Douglas, C ,Sixth Edition, John Wiley &Sons.Inc.
11) Statistical Quality Control: E.L.Grant. 2nd edition, McGraw Hill, 1988.
12) Quality Control and Industrial Statistics: Duncan. 3rd edition, D.Taraporewala sons & company.
13) Quality Control: Theory and Applications: Bertrand L. Hansen, (1973),Prentice Hall of IndiaPvt. Ltd.
14) Introduction to Statistical Quality Control(2009), Montgomery, Douglas, C. , Sixth Edition, John Wiley & Sons, Inc.
15) Quality Control (1976), I.V. Burr, Mardekkar, New York, 16.Fundamentals of Applied Statistics , Gupta and Kapoor.

Web Sites

i). https://sixsigmastudyguide.com/run-chart/
ii). https://kissflow.com/project/agile/5-principles-of-lean/
iii). https://quality-one.com/grr/

HSNC University, Mumbai

Board of Studies

In the Subject of Statistics and Data Science & Business Analytics

1) Dr. Asha Jindal- Chairperson of Ad-hoc BOS
2) Dr. S.B. Muley- Member
3) Ms PratikshaKadam- Member-
4) Ms ShailajaRane- Member
5) Dr. Annum Area Ahmed I/C Principal, Rizvi College- Member (Associate Professor from other University)
6) Dr.Suresh Sharma, Department of Statistics, Panjab University- Member(Eminent Scholar)
7) Dr. Santosh Gite,Associate professor, University of Mumbai -Member (Research Institute)
8) Mr.Mukesh Jain, CTO, CapGemini – Industry Expert (Data Scientist - Contributionn in Society at Large from Industry)
9) Mr. Prashant Nair, Director ,Nelson- Industry Expert
10) Rankers and bright students of Department of Statistics, K.C.College,

i. Ms.Mohadassah Patel
ii. Ms.DivyaSrivastav.