BHARATHIAR UNIVERSITY: COIMBATORE – 641 046

DIPLOMA IN DATA ANALYTICS

(For the candidates admitted from the academic year 2020-2021 onwards)

Diploma in Data Analytics is aimed at providing basic and comprehensive knowledge at theoretical as well as practical level to the students. Data Analytics refers to the techniques to analyze data to enhance productivity and business gain. The role of data scientist is now a buzz worthy career. It has staying power in the marketplace and provides opportunities for people who study data analytics to make valuable contributions to their companies and societies at large. A diploma degree in data analytics would impart students with important transferable skills, such as project management, critical-thinking and problem-solving, and open up opportunities across industries as data managers, data consultants, consumer and market knowledge managers, chief data officers, and big data architects.

Programme Objectives

1. Develop skills in statistics, data analysis, and visualization
2. Be prepared to solve challenging problems involving large, diverse data sets from different application domains

Programme Outcomes

- Demonstrate proficiency with statistical analysis of data.
- Develop the ability to build and assess data-based models.
- Execute statistical analyses with professional statistical software.
- Demonstrate skill in data management.
- Apply data science concepts and methods to solveproblems in real-world contexts and will communicate these solutions effectively

Eligibility for Admission

Candidates for admission to the Diploma course shall be required to have passed the Higher Secondary examinations (Academic or Vocational) conducted by the Government of Tamil Nadu or any other examination accepted as equivalent thereto by the syndicate.

Duration of the course

The course shall extend over a period of six months comprising of one semester.

Medium of instruction and examinations

The medium of instruction and examinations shall be in English.

SCHEME OF EXAMINATION

	Title of The Paper	Instruction Hours / Week	Exam Duration	Total Marks	Credits
Theory 1	Statistical and Forecasting Analysis	4	3 Hours	100	4
Theory 2	Artificial Intelligence	4	3 Hours	100	4
Theory 3	Data Mining	4	3 Hours	100	4
Practical 1	Excel For Statistical Data Analysis	3	3 Hours	100	4
			MAXIMUM MARKS	**400**	**16**

Examinations

- A candidate will be permitted to appear for the University Examination for any year if he / she secures minimum 75% of attendance in the number of instructional days.
- Examination shall be conducted at the end of course.

QUESTION PAPER PATTERN

Distribution of Marks – Theory Semester Exam				
Total Marks	**External**		**Internal (Max Marks)**	**overall Passing minimum** (Internal + External)
	Maximum	Passing Minimum for external alone		
100	75	30	25	40
Distribution of Marks – Practical Semester Exam				
100	60*	24	40	40

* 60 marks break up – 25 by external examiner, 25 by internal examiner, 10 for report by internal examiner

Maximum 100 Marks			
Section A	Answer all the questions (10 questions of one-word answer Q.Nos.1 to 10)	10X1=10	10 questions – 2 each from every unit
Section B	Answer all the questions - 5 questions either or type, ONE question from each unit) [Q.Nos.11 a (or) b … 15 a (or) b]	5X6=30	5 questions – 1 each from every unit
Section C	Answer all the questions (5 question either or type) ONE question from each unit	5X12=60	5 questions – 1 each from every unit

Passing Minimum

a) A candidate shall be declared to have passed in a paper if he or she obtains not less than 40% of marks in that paper.
b) A candidate failing to secure the minimum marks prescribed shall be required to reappear for the examination in that paper and obtain not less than the minimum marks required for passing the paper.

Classification of Successful Candidates

a) Successful candidates passing the examinations by securing not less than 60% of total marks in all subjects shall be declared to have passed in First class.

b) Successful candidates passing the examinations by securing not less than 50% of total marks in all subjects shall be declared to have passed in second class.

c) All other successful candidates shall be declared to have passed the examinations in Third class.

Theory I - STATISTICAL AND FORECASTING ANALYSIS

Course Objective:

a) To demonstrate statistical tools for diagrammatic and graphical representation of data and also computation of various measure of central tendency and dispersion.
b) To make students formulate and solve problems occur in the real world

Learning Outcomes:

a) Able to choose the right analysis technique for business forecasting
b) Understand the meaning of association between two variables and use regression analysis in prediction

UNIT – I

Introduction of Statistics - Definition of Statistics - descriptive and inferential statistics- Variables - collection of data- primary and secondary –classification-tabulation-diagrammatic and graphical presentation.

UNIT – II

Measures of central tendency - Measures of central tendency-mean, median, mode – measures of dispersion - range - quartile deviation, mean deviation-coefficient of variation.

UNIT - III

Correlation and Regression Analysis - Correlation-scatter diagram -Karl Pearson's coefficient of correlation -Spearman's rank correlation coefficient-regression analysis and its uses.

UNIT - IV

Time series and Trend Analysis - Time series - business forecasting estimating trend, graphic, semi average, moving average and method of least squares(linear)

UNIT - V

Index numbers - Index numbers-uses – characteristics of Index numbers – types - aggregate method – simple average of relative method - unweighted and weighted index numbers- tests of an index number- cost of living index number.

Books Recommended:

1. Fundamental of Statistics, S.C Gupta , 2012, 7th Edition, Himalaya Publishing, Delhi.
2. Business Mathematics and Statistics, Navaneetham

Theory II - ARTIFICIAL INTELLIGENCE

Course Objective:

a) To understand the basics of intelligent agents and searching strategies
b) To demonstrate working knowledge of reasoning in the presence of uncertain knowledge

Learning Outcomes:

a) Able to apply various decision making methods in real world environment
b) Able to apply and analyze the application in view of artificial intelligence

UNIT I

Artificial Intelligence - Introduction- intelligent agents- - solving problems by searching- adversarial search

UNIT II

Knowledge and Reasoning - Logical Agents - -First Order Logic-inference in first order Logic - Knowledge representation

UNIT III

Planning and Reasoning - Planning and acting in real world - uncertain Knowledge - Quantifying uncertainty

UNIT IV

Decision making and learning - Making simple decisions-making complex Decisions-Learning from Examples-Knowledge in Learning

UNIT V

Communicating-Perceiving and Acting - Communication-Natural Language Processing – components – difficulties in NLU – step in NLP - -Perception

Books Recommended:

1. Stuart Russell and Peter Norvig, Artificial Intelligence: A Modern Approach, Pearson Education Series, Prentice hall Publishers, Third Edition
2. Russell: Artificial Intelligence 3e: A Modern Approach

Theory III -DATA MINING

Course Objective

a) To demonstrate the need for data preprocessing and suggest appropriate methods to produce proper data sources for mining
b) To analyze the suitability of design technique to find mining solution with efficient time, cost and memory requirement

Learning Outcomes

a) Able to construct a legitimate mining solution with the help of design technique guidelines and validate the suitability of the techniques applied.
b) Students could devise efficient managerial decisions based on mathematical models for real time business applications.

UNIT I

Introduction - Definition and need of data mining - kinds of data and patterns - applications and issues - types of data-data objects and attribute type-measuring data similarity and dissimilarity.

UNIT II

Data Preprocessing - Data Preprocessing-overview-data cleaning-data integration-data transformation-and data discretization

UNIT III

Data warehouse and OLAP (Online Analytical Processing) Technology - Data warehouse-basic concepts-data warehouse modeling-data warehouse **-** implementation

UNIT IV

Data Mining Techniques - Mining frequent patterns and associations-basic concepts-frequent item set mining methods-pattern mining in multilevel and multidimensional space-detection of outliers

UNIT V

Data Mining Trends - Mining sequence data-statistical data mining-visual and audio data mining

Books Recommended:

1. Data Mining-Concepts and Techniques, Jiawei Han, Micheline Kamber and Jian Pei, Third Edition, Elsevier Publications, 2012
2. Data Mining Techniques, Arun K Pujari, Universities Press.

Practical I - EXCEL FOR STATISTICAL DATA ANALYSIS

Course Objective

a) To create applications and tools dealing with complex problems
b) To achieve success in the master spreadsheet analysis

Learning Outcomes

a) Essential business skills in Data Analysis and Business Modeling will be developed
b) Future trends may be predicted for high level decision making

UNIT I

Introduction - Introduction to MS Excel - Entering data - Tabulation of data - Diagrammatic representation of data - Graphical representation of data.

UNIT II

Measures of Central Tendency - Evaluation of mean, median, mode - Evaluation of range - Quartile deviation - Mean deviation - Coefficient of variation.

UNIT III

Correlation and Regression Analysis – Correlation - Scatter diagram - Karl Pearson's coefficient of correlation - Spearman's rank correlation - Regression analysis

UNIT IV

Time series - Estimating trend - Graphical method - Semi average method - Moving average method - Method of least squares

UNIT V

Index numbers - Unweighted index numbers - Weighted index numbers - Tests of an index number - Cost of living index number **–** Minor Project-Strategy design for Business issues

Books Recommended:

1. Hector guerrero, Excel Data Analysis, Springer-Verlang Berlin Heidelberg, 2010.
2. Wayne L. Winston: Microsoft Excel 2016 - Data Analysis and Business Modeling.