

Programme Project Report (PPR)
for
Distance Learning Programme under School of Distance Education

Post Graduate Diploma in Data and Business Analytics (PGDDBA)

Course Co-ordinator: Dr.Sajimon Abraham

Academic support by
School of Management and Business Studies (SMBS)
Mahatma Gandhi University
Kottayam, Kerala

POST GRADUATE DIPLOMA IN DATA AND BUSINESS ANALYTICS (PGDDBA)
(Distance Learning Programme - Diploma Programme)

Programme Project Report

Mahatma Gandhi University started the School of Distance Education in 1989 with a vision to provide the opportunity of quality education to all realms of society. Since the beginning, thousands of students availed this opportunity for higher education throughout Kerala to a great extent and also outside the state to some extent. But after the new directions of UGC in 2014, University had stopped all its Off-Campus Centres of the School of Distance Education inside and outside the State.

Now it is the new endeavour to revamp the functioning of the school with different types of Diploma and Certificate programmes very relevant to the contemporary society, in addition to the conventional Graduate and Post Graduate programmes with the academic and infrastructural support of the eminent Schools and interdisciplinary interuniversity Centres of the University. All these Schools/ Centres have already conducted similar Programmes or Post Graduate Programmes in the same area. This Post Graduate Diploma Programme has been designed by the School of Management and Business Studies and to be conducted by the School of Distance Education with the academic support of the School.

School of Management and Business Studies is a regular statutory Department for Management studies in Mahatma Gandhi University. The School had a humble beginning on 25th March 1999 with a two year full-time MBA program for the management aspirants with the objective of molding enterprising youth into career oriented and competent management professionals. With dedicated and high calibre professional expertise and state-of-art infrastructure, the institution imparts the best of theory and practical knowledge to carve a niche for itself in the challenging two year MBA, MPhil and PhD programmes on Management and Business Studies. Ours is one of the prominent Management Institutes in South India. The MBA program offered by SMBS is aimed at creating business leaders and entrepreneurs by leveraging on its strength in technology, computing and social sciences. The department is currently engaged in a diverse set of activities including teaching, academic research, management development programs, and public sector projects. The department places heavy emphasis on experiential and process oriented learning,

and the pedagogical tools include extensive use of case studies, simulation exercises, industry oriented project works to facilitate the same. Besides honing up the skills of individual decision making, enough emphasis is laid on developing team skills and value focused decision making. Continuous industry interaction, seminars and live projects are a regular part of the curriculum. Organizational environments are simulated to sharpen the skills of decision making, leadership and team building. Teamwork, group assignments, case studies, participation in class discussions and real business issues are strong features of the management program at SMBS.

(a) Programme's mission & objectives :

Business Analytics is the process of converting data into insights. It is “the extensive use of data, statistical and quantitative analysis, explanatory and predictive models, and fact based management to drive decisions and actions.” With the increase in the availability of data, Analytics has now become a major differentiator in both the top line and the bottom line of any organization. It is hence not surprising that research has shown that data-driven companies perform 5%-6% better per annum. The objective of the course is to help participants to understand Large Data, its elements, and usage, from multiple perspectives—as analysts, decision maker, corporate leaders, investment banker, consumer analyst and entrepreneurs.

(b) Relevance of the program with HEI's Mission and Goals :

The enormous amount of data generated digitally can be used by business enterprises to generate new insights, enable better decision making and improve processes of organizations. Business analytics refers to the analysis of data using statistical, machine learning and quantitative techniques with the purpose of understanding past performance of the business and generating new insights for the future. According to a forecast made by IDC, the global spending on business analytics services is expected to rise from US\$ 51.6 billion in 2014 to US\$ 89.6 billion in 2018. Despite the growing significance of business analytics, the supply of trained analytics professionals is lagging far behind the demand for such professionals. The objective of the Programme is to provide a wide spectrum of knowledge, skills and technological advances while fostering literacy in the broadest sense, . which go in-line with the mission of the Institution - To provide skilled manpower to the professional, industrial and service sectors to meet global demands.

(c) Nature of prospective target group of learners:

Students with a minimum of Bachelor's Degree with Statistics, Mathematics, Management, Economics, Commerce, Engineering, Computer Science, IT, and related disciplines. Graduate Students studied mathematics and statistics as one of the subjects either in plus two level or graduate level can also apply. Students pursuing their Master's degree and Working professionals can also apply.

(d) Appropriateness of programme to be conducted in Open and Distance Learning mode to acquire specific skills and competence :

As the programme is targeting for working people and those engaged in regular studies the only way to deliver the programme is through week end contact classes and through distance learning mode like on-line lectures and sharing of video and audio files. Today's internet and networking availability is strong in our country which is reachable to most of the common man we can effectively utilize this facility as a medium for course delivery, evaluation and for other administrative requirements. On completion of this course the students can work as

- Data Mining Expert
- Business Intelligence Expert
- Data Scientists
- Business Analyst and Domain expert

(e) Instructional Design :

- i. Duration of the Programme: One Year –Two Semesters
- ii. Eligibility: Students with a minimum of Bachelor's Degree with Statistics, Mathematics, Management, Economics, Commerce, Engineering, Computer Science, IT, and related disciplines. Graduate Students studied mathematics and statistics as one of the subjects either in plus two level or graduate level

Number of Courses : 8

Scheme and Evaluation

Course Code	Course Type	Contact Classes (Hrs)	Course Name	Credits	IA Marks	ESE Marks	Total Marks
PGD DBA-101	Common Core course (Theory)	12	Data Base Management System	4	20	80	100
PGD DBA-102	Common Core course (Theory)	12	Exploratory Data Analysis	4	20	80	100
PGD DBA-103	Common Core course (Theory)	12	Data Visualization and Reporting	4	20	80	100
PGD DBA-104	Common Core course	12	Data Mining	4	20	80	100
PGD DBA-105	Common Core course	12	R Programming	4	20	80	100
PGD DBA-106	Common Core course	12	Business Forecasting	4	20	80	100

PGD DBA- 107	Common Core course	12	Big Data Technologies	4	20	80	100
PGD DBA- 108	Common Core course (Practical)	60	Practical & Project Work	Practical -2 Project Work -2	20	80	100
Total		144		32	160	640	800

(f) Procedure for admissions, Curriculum Transaction and Evaluation

Admission to the programme will be done by the University through a common procedure for all the programmes under the School of Distance Education. Bachelor's Degree with Statistics, Mathematics, Management, Economics, Commerce, Engineering, Computer Science, IT, and related disciplines. Graduate who have studied mathematics and statistics as one of the subjects either in plus two level or graduate level can also apply. Fee structure will be decided by the University. The School will prepare an academic calendar/activity planner and will be circulated among all the learners at the time of admission itself. The academic calendar will include all the significant activities, important dates, schedule of submission of assignments, schedule of contact classes, schedule of examinations, etc. The course contents will be delivered through online and print formats. For practicals, 20% will be virtual and remaining will be by direct laboratory work. This course will have three types of graded activities that will be included in overall course grade. These include: **Assignments:** Answer monthly quizzes that account for 20% of the course total grade. Each quiz includes 20 multiple choice questions that examines your understanding of the learning materials. **Final project:** At the end of the course the candidate will be asked to complete a final project. It will consists of a written report that focuses on the utilization of computational /analytical techniques for various applications.

Evaluation of the courses shall be done by the faculty themselves on the basis of internal assessment and end semester examinations. 20% of the marks will be decided by the internal evaluations and the remaining 80% by the end semester examinations which will be done by the University. The performance of a student in each course is evaluated in terms of percentage of marks with a provision for conversion to grade points.

Each student shall be required to do one Assignment/Book Review/Debate/Seminar/Presentation of case study for each course. Assignments/Book Review after valuation shall be returned to the students. The teacher shall define the expected quality of the above in terms of structure, content, presentation and the like, and inform the same to the students.

Grading System will be followed for the evaluation on a ten point scale. The details of the grading system are given in the following Table.

Percentage Equivalence of Grade:

Range of % of Marks	Grade Letter	Performance	Grade Point
95 - ≤ 100	O	Outstanding	10
85 - < 95	A plus	Excellent	9
75 - < 85	A only	Very Good	8
65 - < 75	B plus	Good	7
55 - < 65	B only	Above Average	6
45 - < 55	C	Average	5
40 - < 45	P	Pass	4
< 40	F	Fail	0
Absent	Ab	Absent	0

'P' grade is required for a minimum pass in a course. The minimum GPA required for a pass in the Diploma programme is 4.

Calculation of Grade Point Average (GPA) :

Credit Points for the Course = (No. of Credits assigned for the course x Grade Point secured for that course).

GPA indicates the performance of a student in the programme. GPA is based on the total **credit points** earned by a student in all the courses divided by the total number of credits assigned to the courses required in the programme.

Note: GPA is computed only if the candidate passes in all the required courses (gets a minimum required grade for a pass in all the required courses as per the curriculum).

GPA =

$$\frac{\text{Total credit points earned by the student from all the required courses of the programme}}{\text{Total credits of all courses required in the programme}}$$

This formula shall be printed on the Grade Card issued to the student with a note that it could be used to convert the grades into mark-percentages. (The details of the grading system as indicated above shall also be printed on the Grade Card).

Conversion of GPA to Grade

GPA	Grade
10	O
9.0 - < 10	A plus
8.0 - < 9	A only
7.0 - < 8	B plus
6.0 - < 7	B only
5.0 - < 6	C
4.0 - < 5	P
< 4	F
Absent	Ab

Conversion of GPA to percentage

$$\text{Equivalent Percentage} = (\text{GPA obtained}) \times 10$$

(g) Details of Laboratory support required for the programme

The computational facility available in School of Management and Business Studies shall be used. The computing facility available in the campus as well as the regional centers can also

be used for this purpose. Some external computing facilities may be hired based on the number of enrolment.

Mahatma Gandhi University Library and Information System consists of University Library, libraries of the Schools and 4 study centre Libraries. The University Library was established in 1989. The University Library which is situated in the main campus occupies purpose-built accommodation, and provides a variety of facilities and has a user-friendly environment. These include individual work spaces, room for group study and teaching, audio-visual access and online information retrieval system. The building of the University Library is 2000 sq.m in area consisting of the cellar, the ground floor and the first floor.

Academic as well as public users are given the facility to use the library. Special category membership is provided to journalists. The library is providing service from 8 am to 8 pm in three shift timings for its staff. The library functions on an average of 345 days in a year. The libraries of teaching departments are open during working hours of the Schools. Reading space is provided in all the three floors housing the various sections of the library. The library provides reading facility to the visually impaired users too. For this, an electronic lab custom made for visually and physically challenged users has been set up during 2016.

The University Library has a Library Advisory Committee. It is an 18 member committee with Vice-Chancellor as Chairman and University Librarian as Convener.

The library has a collection of 59,000 books, 232 journals, 2,135 Ph.D. theses and has access to 15000+ e-journals under E-Shodh Sindhu. The activities of the Library are comprehensively automated using open source library management software KOHA. OPAC, Journal Article Index, By monthly Bibliography compilation and Literature Search Service are also available

The library is a member of the INFLIBNET Centre, Ahmedabad as well as DELNET (Developing Library Network). As a member of these networks, the library provides access to the resources of other major libraries in the country. In addition to the access to UGC INFONET consortium, it has access to major online databases, such as EBSCO, ProQuest dissertations and theses, Oxford Scholarship Online, IEEE All Society Periodicals Package etc. Mahatma Gandhi University had won the State IT Award during the year 2009 in the e-learning category for its university online theses digital library. The various department libraries have a good collection of subject specific books and journals.

A. MAHATMA GANDHI UNIVERSITY LIBRARY	
Category	No.
Books	59000
Journals	232
Bound Journals	7500
Ph.D Theses	2135
E-Journals (in UGC-Infonet, renamed as E-ShodhSindhu)	15000
Online databases (in UGC Infonet)	11
Online Archives subscribed	185 Titles
Online databases subscribed	4
E-books	7338

DVDs: Educational Videos	293
--------------------------	-----

B	Name of School/Centre	Total No. of books
	School of Management and Business Studies	7549

(h) Cost estimate of the programme and the provisions:

Budget estimate (for 100 students)

S.No.	Item	Amount (Rs. in Lakhs)
1.	Manpower	5
2.	Study material/on line materials	3
3.	Laboratory/Virtual lab	5
4.	Internal assessment/on-line quizzes	0.5
5.	End semester examination	1.5
	Total	14.00

Total Programme fee: Rs.15000/-

(i) Quality assurance mechanism and expected programme outcomes

The quality of the programme will be ensured through strict monitoring by an executive committee including the Co-ordinator of the programme, the subject experts, Director, School of Distance Education and Head of the School of Management and Business Studies. The Co-ordinator of the programme shall ensure the regular student feedback of courses, teachers and programme in the prescribed format towards the end of the semester and the same shall be analysed to draw conclusions for effecting improvement. Periodical review meetings on the programme efficacy will be held in which the remarks of teachers on curriculum, syllabi and methods of teaching and evaluation will be given due importance. Moreover, the progress and the quality of the programme will be monitored by the Internal Quality Assurance Cell of the University from the outcome and feedback of the learners as well as the proper documentation maintained in the Centre.

The P.G. Diploma programs will educate the aspirants who want to make an impact in the corporate and academic world in the domain of data analytics as data scientist and researcher, big data leads/ administrators/ managers, business analysts and data visualization specialist. The course is also suitable for those who are already working in analytics to enhance their theoretical and conceptual knowledge as well as those with analytical aptitude and would like to start career in big data analytics in different business sectors.

.....

Syllabus

POST GRADUATE DIPLOMA IN DATA AND BUSINESS ANALYTICS (PGDDBA)

PGDDBA-101 Data Base Management System

Principles, Tools and Platforms / (Database Management Systems): Database concepts, Basic components of DBMS, sources of data, logging, cleaning data, data representation, data models – (hierarchical, network, XML), and Stores, NoSQL database, design for performance / quality parameters, documents and information retrieval related tools – (Postgres, OLTP, OLAP, Hadoop, Mapreduce).

PGDDBA-102 Exploratory Data Analysis:

Mathematics for Data Analytics: Basic probability theory, distributions and their properties, Simple and multiple regression analysis, hypothesis testing and sampling, estimation theory, least square methods. Descriptive statistics – uni-variate and bi-variate, residual analysis, confidence and prediction intervals regression, associations, sequencing, introduction to forecasting, design of experiments and performing basic statistical analysis of data experiments (both field and laboratory) to investigate business issues, tools for conducting basic statistics (Excel, R, SPSS)

PGDDBA-103 Data Visualization and Reporting:

Purpose of visualization, Multidimensional visualization, tree visualization, graph visualization and time series data visualization techniques, visual perception, cognitive issues, evaluation as well as other theory and design principles behind information visualization, understanding analytics output and their usage, basic interaction techniques such as selection and distortion, evaluation, examples of information visualization applications and systems, user tasks and analysis

PGDDBA-104 Data Mining:

Clustering, Association rules, factor analysis, scale development, survival analysis, data reduction using PCA, scoring new data and model implementation, improving predictive models, association and market basket analysis, advanced regression models: concepts and applications, conjoint and discrete choice analysis, design and analysis of experiment.

PGDDBA-105 R Programming

History and overview of R , Install and configuration of R programming environment , Basic language elements and data structures , R+Knitr+Markdown+GitHub , Data input/output , Data storage formats , Subsetting objects , Vectorization , Control structures , Functions , Scoping Rules , Loop functions , Graphics and visualization , Grammar of data manipulation (dplyr and related tools) , Debugging/profiling , Statistical simulation.

PDDBA-106 Business Forecasting :

The Importance of Forecasting-Time Series Data-Component Factors of the Time-Series Model Trend Analysis-Seasonal and Cyclical Behaviour-Smoothing of Annual Time Series: Moving averages, Exponential smoothing -Least-Squares Trend Fitting and Forecasting: Linear, quadratic and exponential models , Autocorrelation and Auto regression-Autoregressive Models - ARIMA time-series Model Time-Series Forecasting of Monthly or Quarterly Data-Accuracy Statistics and Forecast Model Selection-Families of Forecasting Models –Hierarchical Forecasting-Adjustments to Statistical Forecasts, Event Variables-Outlier Variables and Other Model Inputs-Using Event Variables, Based on Calendar Effects-Combined Model Forecasts-Honest Assessment

PDDBA-107. Big Data Technologies:

Big data definition, enterprise / structured data, social / unstructured data, unstructured data needs for analytics, Big data programming (Hadoop / HDFS, Map-reduce, event stream processing, complex event processing), evolution, purpose and use, application data stores, (NSQL databases, in-memory databases), data computing appliance (DCA) and OLAP, massive parallel processing, in-memory computing / analytics, data science, enterprise / external search, HDFS – Overview and concepts, data flow (read and write), interface to HDFS (HTTP, CLI and Java API), high availability and Name Node federation, Map Reduce developing and deploying programs, optimization techniques, Map Reduce Anatomy, Data flow framework programming Map Reduce best practices and debugging.

PDDBA-108. Practical & Project Work

A business Analytics solution to the issues of typical company has to be identified and have to be implemented using Tools studied. A project report has to be submitted at the end of the programme.

Books Recommended:

1. Vignesh Prajapati, “*Big Data Analytics with R and Hadoop*”, 1st Edition, Shroff / Packt Publications
3. Chuck Lam, “*Hadoop in Action*”, Dreamtech Press Publisher.
4. Michele Chambers, Michael Minelli, Ambiga Dhiraj., “*Big Data Big Analytics, Emerging Business Intelligence and Analytic Trends for Today's Businesses*” , 1st Edition, Wiley Publications
5. Gert H. N. Laursen, Jesper Thorlund, “*Business Analytics for Managers*” Taking Business Intelligence Beyond Reporting, Wiley Publications.
6. Damodar Gujarati & Dawn Porter, Sangeetha Gunasekar, “*Basic Econometrics*”, 5th Edition McGraw Hill Education (India) Private Limited
7. Levine, Stephan, Krehbiel and Berenson., “*Statistics for Managers using Microsoft excel*”, PHI Learning Private Limited, 2010.