SCHEME OF EXAMINATION

AND

SYLLABUS

FOR

POST-GRADUATE DIPLOMA IN DATA ANALYTICS (PGDDA)

APPROVED BY

UNIVERSITY GRANTS COMMISSION (UGC)

UNDER

NATIONAL SKILL QUALIFICATIONS FRAMEWORK (NSQF)

(LEVEL – 8)

WITH EFFECT FROM 2022-2023 ONWARDS

UNIVERSITY SCHOOL OF MANAGEMENT STUDIES

GURU GOBIND SINGH INDRAPRASTHA UNIVERSITY

SECTOR 16 C, DWARKA, NEW DELHI

www.ipu.ac.in

GURU GOBIND SINGH
INDRAPRASTHA
UNIVERSITY

GURU GOBIND SINGH INDRAPRASTHA UNIVERSITY, NEW DELHI

POST-GRADUATE DIPLOMA IN DATA ANALYTICS

(PGDDA)

Evaluation

The evaluation for theory papers shall involve 25 marks for continuous evaluation and 75 marks for final examination. For lab or practical/project based courses, the internal marks shall be 40 and the external marks will be 60.

Criteria for Continuous Evaluation

The continuous evaluation of the students shall be marked out of 25 marks for theory papers and out of 40 marks for practical. The criteria for continuous evaluation shall be decided by the faculty delivering the course under intimation to the Dean (USMS)/ Coordinator (PG Diploma Programs, USMS).

Final Examination

The final examination shall be conducted by the university or relevant sector skill council/assessment board in consultation with the industry partner(s). The MOOC/NPTEL/Similar courses shall be completed by the student on his/her own, and the documentary proof of completion shall be submitted to the university from concerned approved examining agency.

Maximum and Minimum Credits for the Program

Total number of Credits for the PGDDA program is 60. For the award of the diploma, the student is required to complete 56 credits.

Multiple Exit

The student who completes the first semester (with at least 28 credits) may opt to exit from the program. Such student shall be awarded 'Six-Month Certificate Course in Data Analytics'.

Note: The scheme and Syllabus is as per the ordinance 11 of the University

Pravin Chandra

Code	Course	Credits			Examining Body
		Theory	Practical/Tutorial	Total	
PGDDA-101	Information Systems Management	2	0	2	GGSIPU
PGDDA-103	Database Management Systems	2	0	2	GGSIPU
PGDDA-105	Data Preparation (Importing Data and Pre-Processing Data)	1	2	3	Skill Board/GGSIPU
PGDDA-107	Data Exploration	2	2	4	Skill Board/GGSIPU
PGDDA-109	Decision Sciences	2	2	4	GGSIPU
PGDDA-111	Introduction to Analytics and R Programming	2	0	2	Skill Board/GGSIPU
PGDDA-113	Inter-personal Communication (Work effectively with colleagues) NUES	1	1	2	NUES
PGDDA-115	Choice based Online Open Course (MOOCs/NPTEL/ similar course)	0	2	2	-
PGDDA-117	Project-1		6	6	GGSIPU/Skill Board
PGDDA-151	Database Management Systems (Lab)	0	1	1	Skill Board/GGSIPU
PGDDA-153	Introduction to Analytics and R Programming (Lab)	0	2	2	Skill Board/GGSIPU
	Semester Credits	**12**	**18**	**30**	
	Theory-12 Credits, Skill based 18 Credits				**168 hrs (Theory) 504 hrs (Practical)**

Pravin Chandra

SEMESTER II **Total No of Credits: 30**

Code	Course	Credits			Examining Body
		Theory	Practical/ Tutorial	Total	
PGDDA-102	Advanced Business Analytics and Predictive Modelling	2	2	4	Skill Board/GGSIPU
PGDDA-104	Data Modelling with Python	2	0	2	Skill Board/GGSIPU
PGDDA-106	Risk Assessment and Mitigation	2	1	3	Skill Board/GGSIPU
PGDDA-108	Business Performance Modelling	2	2	4	Skill Board/GGSIPU
PGDDA-110	Data Visualization	1	2	3	Skill Board/GGSIPU
PGDDA-112	Choice based Online Open Course (MOOCs/NPTEL/ similar course)	0	2	2	-
PGDDA-114	Managing Client Communication	2	0	2	GGSIPU
PGDDA-116	Project – 2		8	8	Skill Board/GGSIPU
PGDDA-152	Data Modelling with Python (Lab)	0	2	2	Skill Board/GGSIPU
	Semester Credits	**11**	**19**	**30**	
	Theory-11 Credits, Skill based 19 Credits				**154 hrs (Theory) 532 hrs (Practical)**

Total Programme Credits:30+30=60 **Total Programme Hrs. 1358**

Credits required to complete the Programme:56

Note: 1 credit=1 hour of teaching.

Pravin Chandra

POSTGRADUATE DIPLOMA IN DATA ANALYTICS (PGDDA)

Programme Outcomes and Course Outcomes

Programme Outcome:

PO1: Understand and analyse principles applied to contexts and environments of data analytics.

PO2: Design tested and effective advanced analytics models and simulations for decision making.

PO3: Create integrated views of data collected from multiple sources of an enterprise.

PO4: Understand and effectively communicate the results of data analytics to decision makers.

Information Systems Management

Course Code: PGDDA-101 **T-2 P-0 Credits-2**

Course Outcome:

Students who have completed this course would have learned to:

CO1: Understand the concept of information system and its importance for the business analytics.

CO2: Understand concept of data management, telecommunication, internet, and wireless technology.

CO3: Analyze key information systems and enterprise applications, concept of E-commerce and Decision making.

CO4: Analyze and understand building information systems - system analysis and design, and business value of information system.

Database Management Systems

Course Code: PGDDA-103 **T-2 P-0 Credits-2**

Course Outcome:

Students who have completed this course would have learned to:

CO1: To understand the different types of databases DBMS, their purposes, advantages, and relative importance of different database models.

Dr. Pravin Chandra

CO2: To implement SQL for creation, modify and display data from DBMS, understand the basic concept of ER Models, design issues and key constraints, also Reduction of E-R Schema to Tables.

CO3: To impart knowledge about Oracle, its structure and PL/SQL commands, cursors, triggers procedure and functions.

CO4: To impart knowledge about the structure of relational database, normalization, functional dependencies, and their application, and learn database transaction, concurrency control and methods to manage data integrity.

Data Preparation (Importing Data and Pre-Processing Data)

Course Code: PGDDA-105 **T-1 P-2 Credits-3**

Course Outcome:

Students who have completed this course would have learned to:

CO1: Explain a typical process for data collection and transformation methods from different sources.

CO2: Set informed and realistic utilization of meta data.

CO3: Collect raw data and construct a data set.

CO4: Sample and split your data set with considerations for imbalanced data.

CO5: Transform numerical and categorical data.

Data Exploration

Course Code: PGDDA-107 **T-2 P-2 Credits-4**

Course Outcome:

Students who have completed this course would have learned to

CO1: Understand the basic principles of exploratory analysis.

CO2: Know modern extensions to data exploration, including working with "problem data".

CO3: Know the basic principles behind working with all types of data for building all types of models.

CO4: Be able to criticize constructively and determine existing issues with applied linear models in published work.

CO5: Be able to explore the advantages and disadvantages of various approaches to exploratory analysis.

Decision Sciences

Course Code: PGDDA-109 **T-2 P-2 Credits-4**
Course Outcome:

Students who have completed this course would have learned to:

CO1: Greater insight into decision-making processes by use of decision models

CO2: Effective decisions making to solves real-world problems safely and efficiently thru Simulation techniques.

CO3: Possess a range of different perspectives on what counts as an 'effective' decision

CO4: Be better equipped to understand and influence the decision-making processes of other individuals and groups

CO5: understand better how people perceive and decide about risk.

Introduction to Analytics and R Programming

Course Code: PGDDA-111 **T-2 P-0 Credits-2**
Course Outcome:

Students who have completed this course would have learned to:

CO1: Critically thinking on import, manage and structure data files for using business analytics.

CO2: Apply analytical knowledge with the R interface and language for different fields.

CO3: Provide leadership in analytics in existing datasets into R or create new ones.

CO4: Cultivating cognitive skills acquired on existing data and performs all conventional statistical analysis tests. using R knowledge on data management.

Inter-personal Communication (Work effectively with colleagues)

Course Code: PGDDA-113 **T-1 P-1 Credits-2**
Course Outcome:

Students who have completed this course would have learned to:

CO1: Effectively communicate through verbal/oral communication and improve the listening skills.

CO2: Become more effective individual through goal/target setting, self-motivation and practicing creative thinking.

CO3: Function effectively in multi-cultural and heterogeneous teams through the knowledge of teamwork, Inter-personal relationships, conflict management and leadership quality.

CO4: Demonstrate the ability of effective persuasion and negotiation.

SEMESTER II

Advanced Business Analytics and Predictive Learning

Course Code: PGDDA-102 **T-2 P-2 Credits-4**
Course Outcome:

Students who have completed this course would have learned to:

CO1: Understand the process of formulating business objectives, data selection, and process to successfully design, build, evaluate and implement predictive models.

CO2: Compare the underlying predictive modelling techniques.

CO3: Select appropriate predictive modelling approaches to identify cases to progress with.

CO4: Apply predictive modelling approaches using a suitable packages.

Data Modelling with PYTHON

Course Code: PGDDA-104 **T-2 P-0 Credits-2**
Course Outcome:

Students who have completed this course would have learned to:

CO1: Understand Python as a useful scripting language for data analysis.

CO2: To have hands-on training of Statistical Data Analysis through Python Programming

CO3: To Design and implement object-oriented applications.

CO4: To develop the ability to write data mining applications using Python

Risk Assessment and Mitigation

Course Code: PGDDA-106 **T-2 P-1 Credits-3**

Course Outcome:

Students who have completed this course would have learned to:
CO1: Demonstrate knowledge of the range of business and data related risks faced by organisations.

CO2: Understand the algorithm approach to risk management through risk identification, risk measurement and risk management (or mitigation)

CO3: Understand modelling risk mitigation measures.

CO4: Understand operational risk and how to counter measures to manage it.

Pravin Chandra

Business Performance Modelling

Course Code: PGDDA-108 **T-2 P-2 Credits-4**
Course Outcome:

Students who have completed this course would have learned to:

CO1: To understand Performance metrics for algorithms on different business outcomes

CO2: To model performance computation as per specified business outcome and mapping it.

CO3: Understand Different Optimization algorithms and analysis.

CO4: Working on BPM uses cases.

Data Visualization

Course Code: PGDDA-110 **T-1 P-2 Credits-3**
Course Outcome:

Students who have completed this course would have learned to:

CO1: Understand the importance of data visualization and the design and use of many visual components.

CO2: Learn to wisely use various visualization structures such as tables, spatial data, time-varying data, tree and network, etc.

CO3: Learn the basics of colours, views, and other popular and important visualization-based issues.

CO4: Learn basic algorithms in data visualization.

Managing Client Communication

Course Code: PGDDA-114 **T-2 P-0 Credits-2**
Course Outcome:

Students who have completed this course would have learned to:

CO1: Demonstrate competence in verbal business communication.

CO2: Demonstrate competence in the fundamentals of business writing.

CO3: Demonstrate confidence in dealing with client.

CO4: Demonstrate the ability to negotiate with clients confidently and effectively.

Pravin Chandra

PO-CO Mapping

Semester-I

Course Name and Code	CO Number	PO1	PO2	PO3	PO4
PGDDA-101 Information Systems Management	CO1	2	3	2	3
	CO2	3	2	3	2
	CO3	2	3	3	1
	CO4	3	2	2	3
PGDDA 103 Database Management System	CO1	3	1	2	1
	CO2	3	2	2	3
	CO3	1	3	2	3
	CO4	3	2	3	2
PGDDA-105 Data Preparation	CO1	3	3	1	3
	CO2	2	3	1	3
	CO3	2	3	3	2
	CO4	3	2	3	1
PGDDA- 107 Data Exploration	CO1	3	3	1	1
	CO2	2	2	2	3
	CO3	1	3	2	3
	CO4	3	2	2	3
PGDDA 109 Decision Sciences	CO1	3	2	3	2
	CO2	2	2	3	3
	CO3	2	3	2	3
	CO4	3	2	3	2
PGDDA-111 Introduction to Analytics and R Programming	CO1	3	2	3	2
	CO2	2	2	3	2
	CO3	2	2	2	3
	CO4	2	3	3	2
PGDDA-113 Inter-personal Communication (Work effectively with colleagues)	CO1	2	1	1	2
	CO2	1	1	1	2
	CO3	2	1	1	1
	CO4	1	1	1	2

Semester -II

Course Name and Code	CO Number	PO1	PO2	PO3	PO4
PGDDA- 102 Advanced Business Analytics and Predictive Learning	CO1	3	2	1	3
	CO2	3	3	2	2
	CO3	1	3	1	3
	CO4	2	1	3	3
PGDDA- 104 Data Modelling with	CO1	3	3	2	2
	CO2	2	3	1	2

Pravin Chandra

PYTHON	CO3	2	3	3	3
	CO4	2	3	2	3
PGDDA - 106 Risk Assessment and Mitigation	CO1	2	3	2	3
	CO2	2	2	1	3
	CO3	3	2	2	3
	CO4	3	2	1	3
PGDDA - 108 Business Performance Modelling	CO1	3	3	1	3
	CO2	2	2	2	3
	CO3	3	2	3	2
	CO4	3	2	3	3
PGDDA - 110 Data Visualization	CO1	3	2	1	2
	CO2	2	3	1	3
	CO3	3	2	3	2
	CO4	2	3	1	3
PGDDA – 114 Managing Client Communication	CO1	2	1	1	2
	CO2	1	1	1	2
	CO3	2	1	1	1
	CO4	1	1	1	2

Pravin Chandra

SEMESTER I

GURU GOBIND SINGH INDRAPRASTHA UNIVERSITY, NEW DELHI

POSTGRADUATE DIPLOMA IN DATA ANALYTICS (PGDDA)

Information Systems Management

Course Code: PGDDA-101 **L-2 P-0 Credits-2**

Objective: The objective of this course is to expose the students to the managerial issues relating to information systems planning, design, development, and its linkage with the business strategy and help them identify and evaluate various options in this regard.

Course Content

Unit I

Meaning and Role of Information Systems: Information Systems: Digital Enterprises, Role of Information Systems, Types of Information Systems: Operations Support Systems, Management Support Systems, Expert Systems and Knowledge Management Systems. Competitive Strategy Concepts, Strategic Role of Information Systems. **(7 T Hours)**

Unit II

Planning for Information Systems: Identification of Applications, Business Planning Systems and Critical Success Factors, Method of Identifying Applications, Procurement Options – Make or Buy decisions, Outsourcing as an Option. **(7T Hours)**

Unit III

Building Information Systems: Concepts of Systems Analysis and Design, SDLC Approach, Prototyping, Spiral method. Logical and Physical Design. Implementation Strategies of Information Systems. **(7T hours)**

Unit IV

Emerging Concepts and Issues in Information Systems: Enterprise Systems: Supply Chain Management, Customer Relationship Management, ERP, Knowledge Management, Benefits and Challenges of Enterprise Systems. Enhancing Decision Making. **(7T Hours)**

Text Books:

1. James A O'Brien, George M Marakas and Ramesh Behl (2013). Management Information Systems, Tenth Edition, Tata McGraw Hill Education, New Delhi.
2. Kenneth C. Laudon and Jane P. Laudon (2016). Management Information Systems, Fourteenth Edition, Pearson, New Delhi.

References:

1. Turban, E., and Linda Volonino (2017). Information Technology for Management: Transforming Organizations in the Digital Economy, Seventh Edition. Wiley.

Pravin Chandra

2. D. P. Goyal (2014). Management Information Systems-Managerial Perspectives, Third Edition, Macmillan, New Delhi.
3. S.A. Kelkar (2019). Management Information Systems-A concise Study, Second Edition, PHI Learning, New Delhi. (eBook)
4. Sahil Raj (2013), Management Information Systems, Pearson Education, New Delhi

GURU GOBIND SINGH INDRAPRASTHA UNIVERSITY, NEW DELHI

POSTGRADUATE DIPLOMA IN DATA ANALYTICS (PGDDA)

Database Management Systems

Course Code: PGDDA-103 **L-2 P-0 Credits-2**

Objective: This course will help students to understand how databases can be used to store an organization's information.

Course Content

Unit I

Introduction to DBMS: Purpose, Advantages and Disadvantages of DBMS, Schemas and Instances, DBMS Architecture , Data Independence, Data abstraction, Data models – Hierarchical, Network, Relational, Object-Oriented and Object Relational. Data dictionary, metadata, RDBMS, Codd's rules. **(6T Hours)**

Unit II

ER-Model: Basic concepts, Types of attributes, Design Issues, Mapping Constraints, Keys, E-R Diagram :Design of an ER Database Schema, Reduction of E-R Schema to Tables, DDL and DML of SQL: Background, Basic Structure, Set Operations, Aggregate Functions, Null Values, Nested Sub Queries, Derived Relations, Views, Joined Relations, Domain Constraints, Referential Integrity, String Commands, Numeric Function, Date Function, Translate and Decode Function, Modification of Database, Data manipulation. **(8T Hours)**

Unit III

Oracle: Basic Architecture, Introduction to PL/SQL (Conditional, Logic, Loops, Exceptional Handling, Triggers, Procedures, Functions, Cursor). **(8T Hours)**

Unit IV

Structure of Relational Databases: Relational Algebra, Functional Dependencies, Normal forms NF1, NF2, NF3 and BCNF, Multivalued Dependencies and Fourth Normal Form, Join Dependencies and Fifth Normal Form. Transaction: ACID Properties, Transaction State, Concurrency: Locks, Deadlock Condition, Two- Phase Locking Protocol. **(6T Hours)**

Text Books

1. Silberschatz, A, Korth H and Sudarshan S (2013), Database System Concepts, 6/e, McGraw- Hill Education.
2. Elmsari R. and Navathe S. (2013), Fundamentals of Database Systems, 6/e, Pearson Education.

References:

1. Koch, G. & Loney, K. (2008), Oracle 11g, The complete reference. Mc GrawHill Education.
2. Bipin C. Desai (2000), Introduction to Database Management System, Galgotia Publication.
3. Singh, Shio Kumar, Database Systems: Concepts, Design and Applications, 2/e, Pearson Education.
4. Rob. Peter (2010), Data base system concepts, 1/e, Cengage Learning India Pvt. Ltd.

Pravin Chandra

GURU GOBIND SINGH INDRAPRASTHA UNIVERSITY, NEW DELHI

POSTGRADUATE DIPLOMA IN DATA ANALYTICS (PGDDA)

Data Preparation (Importing Data and Pre-Processing Data)

Course Code: PGDDA-105 **L-1 P-2 Credits-3**

Objective: This course will help students to learn importing, preparing, the data before processing.

Course Content

Unit I

Data Preparation: Data import: Open Sources of data, paid data sources, uses and characteristics of open and paid data sources, knowledge development, types of data, enterprise data, consumer data. Reading and importing data from different formats, metadata – meaning and purpose. **(4T + 8P Hours)**

Unit II

Data Mapping: Organising and mapping metadata as per analysis requirement, tools for importing data from databases and data stores, storing data in datasets and data frames, data measurement, features of data. **(3T + 6P Hours)**

Unit III

Data Pre-processing: Processed and unprocessed data, difference, anomalies in the unprocessed data, impact of unprocessed data on analytical operations, tools for pre-processing data, properties of processing tools **(3T + 6P Hours)**

Unit IV

Data Cleaning: Identification and removal of missing values, incorrect data types, techniques and functions for cleaning unprocessed data, transforming incorrect data, approaches to normalize datasets, feature scaling. **(4T + 8P Hours)**

Text Books:

1. Pyle, Dorian Data Preparation for Data Mining, Morgan Koufmann Publishers (Latest Edition)

References:

1. Robert Hoyt and Robert Muenchen (2020). Data Preparation and Exploration, Informatics Education.

Pravin Chandra

GURU GOBIND SINGH INDRAPRASTHA UNIVERSITY, NEW DELHI

POSTGRADUATE DIPLOMA IN DATA ANALYTICS (PGDDA)

Data Exploration

Course Code: PGDDA-107 **L-2 P-2 Credits-4**

Objective: This course will familiarize with the tools used for data exploration and hypothesis testing.

Course Content

Unit I

Data Exploration: meaning, importance, limitations in exploring, types of data, tools for data exploration, properties of exploration tools, selection of right tolls for data exploration for different types of data, guidelines for data exploration **(7T + 7P Hours)**

Unit II

Data summarization: descriptive statistics - mean, median, mode, range, variance, frequency, skewness, normality of data, dimension reduction approaches – Principal Component Analysis, Linear Discriminant Analysis and Non-negative Matrix Factorization **(7T + 7P Hours)**

Unit III

Data Analysis: Analysing data relationship using scatter diagrams and other graphical techniques, using clustering to evaluate correlations between different data points, principles of hypothesis testing, drawing inferences from the results of data analysis. **(7T + 7P Hours)**

Unit IV

Prescriptive actions: meaning, importance, types of prescriptive actions based on data analysis, recommendations of prescriptive actions. **(7T + 7P Hours)**

Text Books:

1. Pimpler, Eric. Data Visualization and Exploration with R: A Practical Guide to Using R, RStudio and Tidyverse for Data Visualization Exploration and Data Science Applications (2017). Geospatial Training Services.
2. Robert Hoyt and Robert Muenchen (2020). Data Preparation and Exploration, Informatics Education

References:

1. Kumar, Suresh Mukhiya and Ahmed, Usman. Hands-On Exploratory Data Analysis with Python: Perform EDA techniques to understand, summarize, and investigate your data
2. Pyle, Dorian Data Preparation for Data Mining, Morgan Koufmann Publishers (Latest Edition)

Pravin Chandra

GURU GOBIND SINGH INDRAPRASTHA UNIVERSITY, NEW DELHI

POSTGRADUATE DIPLOMA IN DATA ANALYTICS (PGDDA)

Decision Sciences

Course Code: PGDDA-109 **L-2 P-2 Credits-4**

Objective: The purpose of this course is to make student learn and practice various decision models used for data analysis.

Course Content

Unit I

Decision Models: Quantitative, Qualitative. Simulation, Forecasting, Prescriptive Model (Quantitative), Linear Programming, Problem Formulation and Solution, Duality, Sensitivity Analysis, Goal Programming. **(7T + 7P Hours)**

Unit II

Simulation modeling: Nature of Simulation Systems, Models and Simulation, Simulation of a Single Server Queuing System, Simulation of an Inventory System, Parallel /Distributed Simulation, Steps in a Simulation Study, Other types of Simulation, Forecasting Process, Forecasting Techniques, Moving Averages and Exponential Smoothing, Forecasting with Regression Methods (elementary) **(7T + 7P Hours)**

Unit III

Qualitative Modeling Concepts: Multicriteria Decision Making (MCDM), ISM (Interpretive Structural Model), IRP (Interpretive Ranking Process), AHP (Analytical Hierarchy Process), ANP (Analytical Network Process), TOPSIS (Technique for Order preference by Similarity to Ideal Solution) **(7T + 7P Hours)**

Unit IV

Decision Making: Supply Chain Management Decisions, Sourcing Analytics, Solving procurement, Manufacturing decisions **(7T + 7P Hours)**

Text Books:

1. J. Holton Wilson and Barry Keating. Business Forecasting, Fourth Edition, (McGraw Hill/Irwin, 2001) ISBN 0-07-252646-7
2. Sharma, J.K. Operations Research, Sixth Edition (Luxmi Publications)

References:

1. Veerachamy, R. and Kumar, Ravi V. (2020). Operations Research, Wiley Publications
2. Hamdy, Taha. Operations research, Ninth Edition (Pearson Education)

Pravin Chandra

GURU GOBIND SINGH INDRAPRASTHA UNIVERSITY, NEW DELHI

POSTGRADUATE DIPLOMA IN DATA ANALYTICS (PGDDA)

Introduction to Analytics and R Programming

Course Code: PGDDA-111 **L-2, Credits-2**

Objective: This course will help students in analysing the data with the help of R Programming technique.

Course Content

Unit I

Introducing to R: R Data Structures, Help functions in R Vectors, Common Vector operations Using all and any Vectorised operations NA and NULL values Filtering Vectorised if-then else Vector Equality Vector Element names, data frames - Creating Data Frames Matrix-like operations in frames Merging Data Frames Applying functions to Data frames Factors and Tables factors and levels Common functions used with factors Working with tables - Other factors and table related functions - Control statements Arithmetic and Boolean operators and values, Recursion Replacement functions Tools for composing function code Math and Simulations in R. **(8T Hours)**

Unit II

Matrices, Arrays and Lists : Creating matrices - Matrix operations, Applying Functions to Matrix Rows and Columns Adding and deleting rows and columns, Vector/Matrix Distinction, Avoiding Dimension Reduction, Higher Dimensional arrays, lists, creating lists, General list operations, accessing list components and values, applying functions to lists, recursive lists. **(6T Hours)**

Unit III

Statistics : Descriptive Statistics (summary Measures) using R, Graphs and charts, Binomial distribution Poisson distribution, Normal distribution, Hypothesis Testing, Analysis of Variance (One way ANOVA, Two way ANOVA), Correlation, Simple and Multiple Linear Regression Analysis Logistic Regression, Time Series Analysis, Factor Analysis, Cluster Analysis. **(7T Hours)**

Unit IV

Advanced R Programming : Interfacing R to Other Languages, Text mining, Neural Networks, Monte Carlo methods, classification, Market Basket Analysis. **(7T Hours)**

Text Books:

1. Motwani, Bharti. (2019). Data Analytics with R, Wiley Publications.
2. Chellappen, Subhashini and Acharya, Seema (2019). Big Data and Analytics, 2nd Edition, Wiley Publications.

References:

1. Ruiz, Diego Modejar (2020). An Introduction to Data Analysis in R: Hands – on coding, Data Mining, Visualization and Statistics from Scratch. Springer Publications
2. Heumann, Christian, Schomaker, Michael Shalabh (2020). Introduction to Statistics and Data Analysis, Springer Publication.

Pravin Chandra

GURU GOBIND SINGH INDRAPRASTHA UNIVERSITY, NEW DELHI

POSTGRADUATE DIPLOMA IN DATA ANALYTICS (PGDDA)

Inter-personal Communication (Work Effectively with Colleagues)

Course Code: PGDDA-113 **L-1 P-1 Credits-2**

Objective: The course aims to provide basic concept, knowledge, and skills with regard to interpersonal communication. The course helps in exploring the inherent challenges and provides techniques to overcome hurdles in effective communication thereby expanding one's professional network and career opportunities.

Course Content

Unit I

Introduction: Defining interpersonal communication, Foundations of Interpersonal Communication, The process of interpersonal communication, Models of interpersonal communication, Culture and Interpersonal Communication, Perception of the Self and Others.

(4T + 4P Hours)

Unit II

Interpersonal Messages & Interpersonal Relationships: Verbal & Non-verbal Messages, Conversational & Emotional Messages, Interpersonal Relationship Stages, Theories, and Communication, Interpersonal Conflict and Conflict Management, Interpersonal Power and Influence.

(4T + 4P Hours)

Unit III

Practicing assertiveness: Practical session on the art of self-disclosure, persuasion, negotiation, questioning and reflection, mindful listening and practicing assertiveness. **(3T + 3P Hours)**

Unit IV

Transactional Analysis: Interpersonal skills for the virtual world, Contemporary case studies.

(3T + 3P Hours)

Text Books:

1. DeVito, Joseph A. The Interpersonal Communication Book, 15th Edition, Pearson, 2018.
2. Hargie, O., Skilled interpersonal communication: Research Theory and Practice , 6th Edition, Routledge, 2016

References:

1. Wood, Julia T., Interpersonal Communication: Everyday Encounters, Cengage Learning, 9th Edition, 2018
2. Solomon, D., and Theiss, J., Interpersonal communication: Putting Theory Into Practice, Routledge, 2013
3. Berko, R., Aitken, Joan E., and Wolvin, A., ICOMM: Interpersonal Concepts and Competencies : Foundations of Interpersonal Communication, Rowman & Littlefield Publishers, 2010

GURU GOBIND SINGH INDRAPRASTHA UNIVERSITY, NEW DELHI
POSTGRADUATE DIPLOMA IN DATA ANALYTICS (PGDDA)

Choice Based Online Open Course
(MOOCs/NPTEL/Similar Course)

Course Code: PGDDA-115 **T/P - 2, Credits – 2**

Objective- The purpose of the course is to give flexibility to the student to adopt self-learning in a selected course of from any domain considered important for the program being pursued by the candidate. The details of the chosen course should be informed by the student well in time in the beginning of the semester to the concerned Dean/Coordinator of the Program. The selected course should have the similar structure and weightage. The decision of the Dean/Coordinator shall be final in this respect.

The student has to undertake the course from online digital resources and submit the successful completion certificate to the University to take the benefit of credits earned for completion of the program.

GURU GOBIND SINGH INDRAPRASTHA UNIVERSITY, NEW DELHI
POSTGRADUATE DIPLOMA IN DATA ANALYTICS (PGDDA)

Project-1

Course Code: PGDDA-117 **T/P - 6, Credits – 6**

Objective- The objective of this project is to enable the students to apply the knowledge acquired during the semester through various courses, to be applied and prepared in the form of a project.

The suggested structure of the project is as under:

1) Descriptive title of the study
2) Nature of the study
   - Problem to be examined
   - Significance and need for the study
   - Background information available
   - Scope of the study – extent and limitations
   - To whom will it be useful?
3) Hypothesis, if any, to be tested
4) Data Sources and Collection procedure
5) Methodology for data analysis
6) Schedule – target dates for completing
   - Review of Literature (Library/Internet research)
   - Primary research
   - Data analysis
   - Findings and conclusions
   - Future scope
7) Bibliography

Pravin Chandra

GURU GOBIND SINGH INDRAPRASTHA UNIVERSITY, NEW DELHI

POSTGRADUATE DIPLOMA IN DATA ANALYTICS (PGDDA)

Database Management Systems (Lab)

Course Code: PGDDA-151 **L –0 P-1 Credits: 1**

Course Contents This course will be based on PGDDA-103 Database Management Systems Course and is part of it.

GURU GOBIND SINGH INDRAPRASTHA UNIVERSITY, NEW DELHI

POSTGRADUATE DIPLOMA IN DATA ANALYTICS (PGDDA)

Introduction to Analytics and R Programming (Lab)

Course Code: PGDDA-153 **L –0 P-2 Credits: 2**

Lab will be based on Paper PGDDA-111 and will basically cover the following: R Data Structures, Correlation and Regression, Neural Networks, Market Basket Analysis and other models in R Programming language.

Students are required to maintain a record of all the exercises done by them in a Lab file duly signed by the faculty.

Pravin Chandra

SEMESTER II

GURU GOBIND SINGH INDRAPRASTHA UNIVERSITY, NEW DELHI

POSTGRADUATE DIPLOMA IN DATA ANALYTICS (PGDDA)

Advanced Business Analytics and Predictive Learning

Course Code: PGDDA-102 **L-2 P-2 Credits-4**

Objective: The students will be able to understand the process of formulating business objectives, data selection/collection, preparation, and process to successfully design, build, evaluate, compare, select, and implement predictive models for a various business application.

Course Content

Unit I

Data Understanding and Preparation: Introduction, Reading data from various sources, Data visualization, Distributions and summary statistics, Relationships among variables, Segmentation, Outlier detection, Automated Data Preparation, Combining data files, Aggregate Data, Duplicate Removal, Sampling DATA, Missing Values. **(7T + 7P Hours)**

Unit II

Model development & techniques: Data Partitioning, Model selection, Model Development Techniques, Linear Regression: Simple and Multiple Regression, Model Validation, Classification using SPSS: Two stage, Hierarchical, k-Means, KNN, Decision Trees, Linear Discriminant Analysis. **(7T + 7P Hours)**

Unit III

Advanced Modelling: Neural networks, Cox Regression, Support vector machine, Bayesian Networks, Association rules. **(7T + 7P Hours)**

Unit IV

Model Evaluation and Deployment: Introduction, Rule Induction Using CHAID, Automating Models for Categorical and Continuous targets, Comparing and Combining Models, Evaluation Charts for Model Comparison, Deploying Model, Assessing Model Performance, Model Updation. **(7T + 7P Hours)**

Text Books:

1. Bari, A., Chaouchi, M., Jung, T. Analytics for Dummies, 2nd Edition, 2016.
2. Siegel, Eric, Predictive Analytics: The Power to Predict Who Will Click, Buy, Lie, or Die, Wiley, 2016

References:

1. Theobald, Oliver, Data Analytics for Absolute Beginners Cengage Learning, 2nd Edition, 2019
2. Albright, Christian S., and Winston, Wayne, L., Business Analytics: Data Analysis & Decision Making, Cengage Learning, 6th Edition.

Pravin Chandra

GURU GOBIND SINGH INDRAPRASTHA UNIVERSITY, NEW DELHI

POSTGRADUATE DIPLOMA IN DATA ANALYTICS (PGDDA)

Data Modelling with PYTHON

Course Code: PGDDA-104 **L-2, Credits-2**

Objective: The students will be able to learn about the information needs of Management and shall also get hands on training of statistical data analysis through Python Programming

Course Content

Unit I

Introducing to Python: Installing Python and Anaconda Introduction to Jupyter (I-Python) notebook - Environment setup Datatypes Variables and Types Installing libraries Strings Lists Tuples Dictionaries List Comprehensions Dictionary Comprehensions Regular Expressions, Control Flow (if-then statements, looping) Organizing code (functions, modules, packages) _ Reading and writing data from local files (.txt,.csv,.xls,.json, etc) Scraping tables from webpages (.html) read table method Introduction to Numpy and 2D plotting Understanding the N-dimensional data structure Creating arrays Indexing arrays by slicing or more generally with indices or masks Basic operations and manipulations on N-dimensional arrays. **(7T Hours)**

Unit II

Plotting with matplotlib: Scatter plots, line plots, box plots, bar charts and histograms, Working with Pandas data structures: Series and Data Frames, Accessing your data: indexing, slicing, fancy indexing, Boolean indexing, Data wrangling, including dealing with dates and times and missing data, Adding, dropping, selecting, creating, and combining rows and columns, Pandas powerful group by method, Reshaping, pivoting, and transforming your data Simple and rolling statistics. **(7T Hours)**

Unit III

Introduction to Machine Learning: AI ML Concepts, Learning algorithms, Supervised learning Linear Regression, Logistic Regression, Decision Trees, Ensemble Learning, KNN, Bayesian Techniques, Support Vector Machines, Time Series Analysis, Neural Networks, Unsupervised learning, Cluster analysis. **(7T Hours)**

Unit IV

Introduction to Text Mining: Text Processing using Base Python and Pandas, Regular expressions Text Processing with specialized modules like NLTK, sklearn, etc., Sentiment Analysis, Word cloud analysis, Segmentation using K-Means/Hierarchical Clustering, Classification (Span/Not spam), Basics of deep learning and neural networks Optimizing a neural network with backward propagation Building deep learning models with keras, Fine-tuning keras models, Introduction to TensorFlow, Convolutional Neural Networks(CNN), Recurrent Neural Networks(RNN) _ Unsupervised Learning – Autoencoders. **(7T Hours)**

Text Books:

1. Miller, Thomas, W. Modelling Techniques in Predictive Analytics with Python and R: A Guide to Data Science, Pearson, 2014.
2. McKinney, William, Python for Data Analysis: Data Wrangling with Pandas, NumPy, and IPython, 2nd Edition, O'Reilly, 2017

Pravin Chandra

References:

1. Downey, Allen and Elkner, Jeffrey and Meyers, Chris Learning with Python, Dreamtech Press, 2015.
2. Nageswara, Rao R. (2018). Core Python Programming, 2nd Edition, Dreamtech Publications.

GURU GOBIND SINGH INDRAPRASTHA UNIVERSITY, NEW DELHI

POSTGRADUATE DIPLOMA IN DATA ANALYTICS (PGDDA)

Risk Assessment and Mitigation

Course Code: PGDDA-106 **L-2 P-1 Credits-3**

Objective: The students will be able to learn the techniques for risk assessment and mitigation for the different models.

Course Content

Unit I

Introduction: Describe the various factors that contribute to algorithmic risk such as flawed data or assumptions, coding errors, insufficient sample sizes, Comprehend the impact that risk factors might have on the outcome of the algorithmic model **(6T +3P Hours)**

Unit II

Risk Estimation: Compute deviation from expected outcomes of model by testing it with multiple inputs; Apply different techniques to estimate the risks involved when the model deviates from expected outcomes **(6T +3P Hours)**

Unit III

Mitigation: Categorize the various mitigation measures that can be introduced to counter each type of model risk, select suitable checks and mitigation measures to counter the risk. **(8T +4P Hours)**

Unit IV

Measures of Mitigation: Translate mitigation measures into a structured corrective action that can be communicated to the rest of the organization **(8T +4P Hours)**

Text Books:

1. Trendowicz, Adam, Software Cost Estimation, Benchmarking, and Risk Assessment: The software Decision-Makers' Guide to Predictable Software Development, Springer, 2015.
2. Fundamentals of Risk Management: Understanding, Evaluating and Implementing Effective Risk Management by Paul Hopkin, kobo publication.
3. Business Risk and Simulation Modelling in Practice, By MICHAEL REES , ISBN: 978-1-118-90404-6 , Wiley publication.

References:

1. Risk Analysis Foundations, Models, and Methods, by Louis Anthony Cox Jr. (Author), Publisher : Springer; 2002nd edition, ISBN-10 : 9780792376156
2. Kramer, Anne, Model□Based Testing Essentials: Guide to the ISTQB Certified Model□Based Tester Foundation Level, John Wiley & Sons, 2016

Pravin Chandra

GURU GOBIND SINGH INDRAPRASTHA UNIVERSITY, NEW DELHI

POSTGRADUATE DIPLOMA IN DATA ANALYTICS (PGDDA)

Business Performance Modelling

Course Code: PGDDA-108 **L-2 P-2 Credits-4**

Objective: The Students will be able to categorize the different performance metrics based on different business outcomes, compute the performance of the model, describe different hyperparameters that can maximize model performance, apply different techniques to identify hyperparameters, use different optimization algorithms and apply the concepts behind hyperparameter tuning, batch normalization.

Course Content

Unit I

Introduction: Performance metrics for algorithms on different business outcomes, categorization of performance metrices **(7T +7P Hours)**

Unit II

Model performance: Computation as per specified business outcome; Hyperparameters description for maximizing model performance; techniques for identifying hyperparameters – grid search grid search, random search, Bayesian optimization **(7T +7P Hours)**

Unit III

Optimization algorithms: meaning, features, purpose, types - minibatch gradient descent, RMSprop, Adam **(7T +7P Hours)**

Unit IV

Hyperparameter tuning: meaning, concepts behind hyperparameter tuning and their application, batch normalization **(7T +7P Hours)**

Text Books:

1. Loguna, Manuel, Marklund, Johan, Business Process Modeling, Simulation and Design, Pearson Publication, 2015.
2. Havey, Michael, Essentials Business Process Modeling, O'Reilly Media Inc, 2005

References:

1. Southekal, H. Prashanth, Data for Business Performance: Model to Transform Business Data into an Enterprise, Technics Publications, 2020.

Pravin Chandra

GURU GOBIND SINGH INDRAPRASTHA UNIVERSITY, NEW DELHI

POSTGRADUATE DIPLOMA IN DATA ANALYTICS (PGDDA)

Data Visualization

Course Code: PGDDA-110 **L-1 P-2 Credits-3**

Objective: The students will be able to learn different visualization tolls and will be able to apply them on real life data.

Course Content

Unit I

Introduction: Visualization – meaning, importance, results analysis, categorization of different business outcomes, Meeting business outcomes through visual reports, Selecting the right visual tool. **(3T + 6P Hours)**

Unit II

Data Analysis: Audience analysis for visualization, Identifying the right delivery mode and format to report the results of a data analysis **(3T + 6P Hours)**

Unit III

Data Summarization: Charting techniques: Basic and Advanced Charts, different types of visualizations, challenges in information visualizations, Role of perception, summarizing narratives **(4T + 8P Hours)**

Unit IV

Visualization: Preparing visuals, formatting charts and graphs, types of axis, different types of calculations, dynamic charts, dashboards and story boards. **(4T + 8P Hours)**

Text Books:

1. Nussbaumer, Knaflic, Cole (2015). Storytelling with Data: A Data Visualization, Wiley Publication
2. Healy, Kieran (2019). Data Visualization – A Practical Introduction, Princeton University Press

References:

1. Jones, Ben (2014). Communicating Data with Tableau: Designing, Developing and Delivering Data, O'Reilly Publications.
2. Wilke, Claus O. (2019). Fundamentals of Data Visualization: A Primer on Making Informative and Compelling Figures, O'Reilly Publications.

Pravin Chandra

GURU GOBIND SINGH INDRAPRASTHA UNIVERSITY, NEW DELHI
POSTGRADUATE DIPLOMA IN DATA ANALYTICS (PGDDA)

Choice Based Online Open Course
(MOOCs/NPTEL/Similar Course)

Course Code: PGDDA-112 **T/P - 2, Credits – 2**

Objective- The purpose of the course is to give a flexibility to the student to adopt self-learning in a selected course of from any domain considered important for the program being pursued by the candidate. The details of the chosen course should be informed by the student well in time in the beginning of the semester to the concerned Dean/Coordinator of the Program. The selected course should have the similar structure and weightage. The decision of the Dean/Coordinator shall be final in this respect.

The student has to undertake the course from online digital resources and submit the successful completion certificate to the University to take the benefit of credits earned for completion of the program.

Pravin Chandra

GURU GOBIND SINGH INDRAPRASTHA UNIVERSITY, NEW DELHI

POSTGRADUATE DIPLOMA IN DATA ANALYTICS (PGDDA)

Managing Client Communication

Course Code: PGDDA-114 **L-2 P-0 Credits-2**

Objective: The students will learn to make effective communication with clients for amicable long-term relations.

Course Content

Unit I

Writing process: Understanding three steps- planning, writing and completing business messages. Writing routine, positive, negative, and persuasive messages. Exercises- on writing effective letters, emails, blogs, reports. **(7T Hours)**

Unit II

Business Presentations: Developing Presentations, Enhancing Presentations. Understanding how to deliver presentations. Practical sessions on how to deliver presentations keeping in mind, body language and other relevant nonverbal communication methods. **(7T Hours)**

Unit III

Employment Messages and preparing for an interview: Writing resumes and application letters. **(7T Hours)**

Unit IV

Practical sessions: Practical sessions on mock interviews and group discussions **(7T Hours)**

Textbooks:

1. Courtland L. Bovée et. al. (2017) Business Communication Today
2. Lesikar R et.al. (2017) Business Communication: Making Connections in a Digital World

References:

1. Murphy H et.al. (2017) Effective Business Communication
2. Verma S (2014) Business Communication: Essential Strategies for 21st Century Managers

GURU GOBIND SINGH INDRAPRASTHA UNIVERSITY, NEW DELHI

POSTGRADUATE DIPLOMA IN DATA ANALYTICS (PGDDA)

Project-2

Course Code: PGDDA- 116 **T/P - 8, Credits – 8**

Objective- The objective of this project is to enable the students to apply the knowledge acquired during the semester through various courses, to be applied and prepared in the form of a project.

The suggested structure of the project is as under:

1) Descriptive title of the study
2) Nature of the study
   - Problem to be examined
   - Significance and need for the study
   - Background information available
   - Scope of the study – extent and limitations
   - To whom will it be useful?
3) Hypothesis, if any, to be tested
4) Data Sources and Collection procedure
5) Methodology for data analysis
6) Schedule – target dates for completing
   - Review of Literature (Library/Internet research)
   - Primary research
   - Data analysis
   - Findings and conclusions
   - Future scope
7) Bibliography

GURU GOBIND SINGH INDRAPRASTHA UNIVERSITY, NEW DELHI

POSTGRADUATE DIPLOMA IN DATA ANALYTICS (PGDDA)

Data Modelling with Python (Lab)

Course Code: PGDDA- 152 **P - 2, Credits – 2**

Lab will be based on Paper PGDDA 104 and will basically cover the following: Introduction to J. Notebook, illustrations, charts using matplotlib in Python, text mining and various machine learning models.

Students are required to maintain a record of all the exercises done by them in a Lab file duly signed by the faculty.