**Academic Year - 2015**

**PGDM in Business Analytics**

**Duration of the Course: 1 Year Full-time (2 semesters)**

**Course Objective**

Most industry sectors have recognized the value that Business Analytics can provide in not only driving compelling Business Solutions, but also in helping them to differentiate themselves to customers, investors and regulators. The global Business Analytics is a USD 105 billion market, growing at a CAGR of 8%. India, with its surfeit of talent, has become the Analytics hub for organizations across the world. Large corporate like Walmart, Target, Citibank, ICICI Bank, Airtel, Vodafone are increasingly adopting analytics in their processes. Consulting giants like PwC, IBM, Accenture, Infosys have large teams offering Analytics solutions to their clients. All of these translate into a huge global and domestic demand for Business Analytics professionals.

The curriculum is designed to provide an in-depth hands-on understanding of Business Analytics’ techniques to analyze and solve problems conceptually and practically.

**Eligibility Criteria**

1. Minimum 50% marks in Graduation
2. Reasonably good quantitative aptitude
3. Minimum 60% score in entrance test

**Semester I**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Subject Code** | **Subject** | **Credits** | **Marks** | | **Total Marks** |
| **IA** | **Term End** |
| PGDM101 | Business Foundation | 6 | 30 | 70 | 100 |
| PGDM102 | DBMS and Data Warehousing | 6 | 30 | 70 | 100 |
| PGDM103 | Statistics for Business | 6 | 30 | 70 | 100 |
| PGDM104 | Introduction to Business Analytics | 6 | 30 | 70 | 100 |
| PGDM105 | Spreadsheet Modelling and Decision Analysis | 6 | 30 | 70 | 100 |
| PGDM106 | Research Methodology | 6 | 30 | 70 | 100 |
|  | **TOTAL** | **36** |  | | **600** |

**Semester II**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Subject Code** | **Subject** | **Credits** | **Marks** | | **Total Marks** |
| **IA** | **Term End** |
| PGDM201 | Multivariate Data Analysis – 1 | 6 | 30 | 70 | 100 |
| PGDM202 | Multivariate Data Analysis – 2 | 6 | 30 | 70 | 100 |
| PGDM203 | Big Data Analytics | 6 | 30 | 70 | 100 |
| PGDM204 | Data Mining and Optimization | 6 | 30 | 70 | 100 |
| PGDM205 | Project | 6 | 30 | 70 | 100 |
|  | **TOTAL** | **30** |  | | **500** |

1 credit = 10 learning hours

**Semester I**

**PGDM101: Business Foundation**

Learning hours: 60

Credits: 6



**Learning Objectives:**

At the end of this Subject, students will:

* Familiarise with the basics of marketing at the strategic & execution levels.
* Understand the rational, sensory & emotional triggers in consumer buying & selling.
* Familiarise with fundamental concepts of Financial Management
* Learn how individual behaviour affects organizational behaviour and vice versa.
* Understand the crucial role of people in any business and to explore different aspects of building this vital capability.

**Module 1 10** **Hours**

**An Overview of Marketing Management**

Introduction to Marketing Management, What is Marketing?, Understanding the Marketplace and Consumer Needs, Designing a Customer Driven Marketing Strategy, Construct an Integrated Marketing Program, Building Customer Relationships, Capturing Value from Customers, The Changing Landscape of Marketing, Consumer Behaviour and Business Buyer Behaviour, Harley-Davidson Consumers: A Case Study, The Buyer Decision Process, Business Buying Process.

**Module 2 14** **Hours**

**Market Segmentation, and Products, Services and Pricing Decisions**

Marketing Strategy, Market Segmentation and Product Positioning, Market Targeting, Target Market Strategies, Product Positioning and Differentiation, Choosing a Differentiation and Positioning Strategy, Changing the Product Positioning, Consumer Needs, Wants, and Consumer Insights, Benefits Sought By Consumers, The Impact of Emotional Benefits, Examples of Consumer Insights, Consumer Insight, Possible Sources of Insights.

Product and Price, Products and Services, Levels of Product and Services, Consumer Products, Industrial Products, Product and Service Decisions, Product and Service Attributes, Branding, Packaging, Labelling, Product Support Services; Product Line Decisions, Product Mix Decisions; Services Marketing, Branding Strategy, Building Strong Brands, New Product Development and Product Life Cycle (PLC), Pricing, Factors to Consider When Setting Prices, Customer Perceptions of Value, Company and Product Costs, Public Policy and Pricing.

**Module 3** **10 Hours**

**Introduction to Financial Management**

* Definition, meaning and role of Financial Management, Goals of Business finance, Profit Maximization vs. Wealth Maximization, Profit Maximization – Approvals and Objections, Wealth Maximization, Growth Maximization, Agency problems.
* Time Value of Money: Time line, Cash Flow Sign Convention, Calculation of Annuities, Inflation Adjusted Interest Rates, Present Value, Net Present Value, Excel Application.

**Module 4 14 Hours**

**Cost of Capital** **and** **Financial Statements Analysis**

Cost of Capital:

* Cost of Debt: Irredeemable and Redeemable Debenture, Preference Shares, Common Stock
* Cost of Equity: Dividend Yield Method, Dividend Yield Plus Growth Rate Method, Earning Yield Method, Realised Yield Method.
* Calculating Weighted Average Cost Of Capital: the adjusted-beta CAPM, APM, the Bond Yield plus Risk Premium Approach.
* Cost of Retaining Earnings, Weighted Average Cost of Capital– Meaning, Limitations and Considerations in Calculating WACC, Adjusting WACC for Risk.

Analysis of Financial Statements and Cash Flow Analysis:

* Common size, trend ratios and comparative techniques, Ratio analysis, profitability ratios, turnover ratios, market test ratios, liquidity and solvency ratios Inter and Intra firm comparisons
* Format and analysis [With special reference to AS 3].

**Module 5 12 Hours**

**Organizational Behaviour and HRM**

Importance of people in an Organization, Theories of work motivation [Herzberg, Maslow, Adam Equity Theory, Porter and Lawler model, Vroom], Perceptual process, Special HR issues in Motivation, Extrinsic and Intrinsic level of motivation, Factors influencing perception, Importance of perceptions in decision making, Perceptual organization: Influence of past experiences, the plasticity of perception, Components of attitude [ABC Model]. Introduction to HRM, What is HRM today? Functions of HRM, Steps involved in HR planning, Factors affecting HRP. Sources of Recruitment. Factors affecting recruitment, Importance of performance appraisal, External and internal factors affecting HR management.

**Reference Books:**

1. Kotler, P., Keller, K. L. , Koshy, A., Jha, M. (2013). *Marketing Management: A South Asian Perspective*. India: Pearson Education.
2. Rajan, S. (2005). *Marketing Management.* India: Tata McGraw-Hill Education.
3. Khan & Jain. (2014). *Financial Management*, (7th ed.). TMH.
4. Pandey, I. M. (2013). *Financial Management*, (10th ed.). Vikas Publications.
5. Chandra, Prasanna. (2010). *Financial Management, (*5th ed.). Tata McGraw Hill,.
6. [Stephen, P. R](http://www.flipkart.com/author/stephen-p-robbins)., [Timothy A. J.](http://www.flipkart.com/author/timothy-a-judge) (2010). *Organizational Behaviour*. India: Pearson.
7. Pareek, U. (2011). *Understanding Organizational Behavior*. India: Oxford University Press.

**PGDM102: DBMS and Data Warehousing**

Learning hours: 60

Credits: 6



**Learning Objectives:**

The subject gives an understanding on how technology and architecture collaborate together to serve the business requirements of different users in the organization. The course predominantly explains the importance of data management in the organization for transactional and analytical objective. The operational data from database management systems are used for day-to-day business transactions and serve the users on current information specifically whereas the goal of Data Warehousing is to provide the users with homogenized, consistent and comprehensive view of the organization to support planning, forecasting and decision making processes at an enterprise level. This subject enables the students to the fundamentals of Data Warehousing and Databases, and how to leverage them for maximising individual and organisational effectiveness.

**Unit 1 12 hours**

**Introduction to Database Management Systems**

Introduction to databases, the relational model, database design, normalization process, parallel and distributed databases, object oriented databases: concept, web technology and DBMS, transaction management.

**Unit 2 12 hours**

**Introduction to Data Warehousing**

Data warehousing concepts, Databases versus Data Warehousing, Business need for data warehousing, architecture of Data Warehouse, building a Data Warehouse, properties of data in Data Warehouse, importance of Meta Data, Data Marts, critical success factors of Data Warehouse, trends in Data Warehousing.

**Unit 3 12 hours**

**Data Preparation for Data Warehousing**

Mapping Data Warehouse with Business, dimensional modeling, Data Extraction, Transformation and Loading Tools (ETL), importance of data quality in Data Warehousing.

**Unit 4 12 hours**

**Data Warehouse and Analysis**

Categorizing users of Data Warehouse and their business requirement, reporting and query tools, On-Line Analytical Processing (OLAP) in Data Warehouse, role of Data Warehousing on web applications, introduction to Data Mining, Data Visualization.

**Unit 5 12 hours**

**Data Warehouse Implementation and Maintenance**

Introduction to Business Intelligence Applications, expanding Data Warehouse/ Business Intelligence System, deployment, growth and maintenance of Data Warehouse.

**Reference Books:**

1. Mundy, Joy. et al. (2008). *The Data Warehouse Lifecycle Toolkit*. Indianapolis: Wiley publishing Inc.
2. Ponniah, Paulraj. (2011). *Data Warehousing: Fundamentals for IT Professionals*. New Delhi: Wiley India Pvt Ltd.
3. Prabhu, C. S. R. (2008). *Data Warehousing: Concepts, Techniques, Products and Applications*. New Delhi: PHI Publications.
4. Inmon, W. H. (2005). *Buidling the Data Warehouse*. London: Wiley Publications.

**PGDM103: Statistics for Business**

Learning hours: 60

Credits: 6



**Learning Objectives:**

The objective of this subject is to give students a conceptual introduction to the field of statistics and its many applications. This is to enable the students to familiarise themselves with descriptive statistics, probability and sampling distributions, hypothesis testing, forecasting, nonparametric methods, decision analysis, time value of money, simple and compound interest, discounting and compounding, present value and future value.

**Unit 1**

**Descriptive Statistics 12 hours**

Introduction, summarizing Data, EDA, Crosstabulation, measures of location, measures of variability, measures of distribution shape, measures of association, applications with Minitab and Excel.

Probability Distributions: Introduction, Bayes’ Theorem, Discrete Probability Distributions, Continuous Probability Distributions, Binomial Distribution, Poisson Distribution, Approximation of Binominal Distribution to Normal Distribution.

**Unit 2**

**Sampling Distributions** **12 hours**

Sampling and Sampling Distributions, Central Limit Theorem, Concept of Standard Error, Confidence Intervals, Applications with Minitab and Excel.

Hypothesis Testing :Hypothesis Testing and Decision Making, Statistical Inferences about means and proportions and variances, Test of Goodness of Fit and Independence, Applications with Minitab and Excel.

**Unit 3**

**Additional Topics 12 hours**

Index Numbers, Forecasting, Decision Analysis, Control Charts, Applications with Minitab and Excel.

Nonparametric Methods: Related samples, Wilcoxon Signed Rank, Mann-Whitney test, Friedman AVOVA test.

**Unit 4**

**Financial Mathematics 12 hours**

Time Value of Money, Simple and Compound Interest, Discounting and Compounding, Present Value and Future Value of single sum and annuity, Loan payments and Amortisation, Applications with Excel.

**Unit 5**

**Correlation and Regression 12 hours**

Analyses of Relationship, Ordinary Least Square Regression, Standard error of estimate, Adjusted R square and goodness of fit.

**Reference Books:**

1. Anderson, David R., Thomas A. Williams and Dennis J. Sweeney. (2012). *Statistics for Business and Economics*. New Delhi: South Western.
2. Levin, Richard I. and David S. Rubin (1994). *Statistics for Management.* New Delhi: Prentice Hall.
3. Waller, Derek. (2008). *Statistics for Business*. London: BH Publications.
4. Lee, Cheng. et al. (2013). *Statistics for Business and Financial Economics*. New York: Heidelberg Dordrecht.

**PGDM104: Introduction to Business Analytics**

Learning hours: 60

Credits: 6



**Learning Objectives:**

The objective of this subject is to provide foundational knowledge associated with the domain of business analytics, to familiarise the students with all concepts including what problems the technology of Data Warehouse (DW) /Business Intelligence (BI) /Advanced Analytics (AA) solve for businesses and when an organisation is ready for DW/BI/AA.

**Unit 1**

**IT and Business Analytics 12 hours**

Business View of Information Technology Applications, Business Enterprise Organization, Its Functions, and Core Business Processes, Baldrige Business Excellence Framework (Optional Reading) , Key Purpose of Using IT in Business, The Connected World: Characteristics of Internet-ready IT Applications, Enterprise Applications (ERP/CRM, etc.) and Bespoke IT Applications, Information Users and Their Requirements.

**Unit 2**

**Digital Data, OLTP and OLAP 12 hours**

Types of Digital Data, Getting to Know Structured Data, Getting to Know Unstructured Data, Getting to Know Semi-Structured Data, Difference Between Semi-Structured and Structured Data. Introduction to OLTP and OLAP, OLTP (On-Line Transaction Processing), OLAP (On-Line Analytical Processing), Different OLAP Architectures, OLTP and OLAP, Data Models for OLTP and OLAP, Role of OLAP Tools in the BI Architecture, Should OLAP be Performed Directly on Operational Databases?, A Peek into the OLAP Operations on Multidimensional Data, Leveraging ERP Data Using Analytics.

**Unit 3**

**Business Intelligence 12 hours**

Getting Started with Business Intelligence, Using Analytical Information for Decision Support, Information Sources Before Dawn of BI?, Business Intelligence (BI) Defined, Evolution of BI and Role of DSS, EIS, MIS, and Digital Dashboards, Need for BI at Virtually all Levels, BI for Past, Present, and Future, The BI Value Chain, Introduction to Business Analytics.BI Definitions and Concepts, BI Component Framework, Who is BI for?, BI Users, Business Intelligence Applications, BI roles and Responsibilities, popular BI tools.

**Unit 4**

**Data Integration and Modeling 12 hours**

Basics of Data Integration, Need for Data Warehouse, Definition of Data Warehouse, What is a Data Man?, What is Then an ODS?, Ralph Kimball's Approach vs. Who Inmon's Approach, Goals of a Data Warehouse, What Constitutes a Data Warehouse?, What is Data Integration?, Data Integration Technologies, Data Quality, Data Profiling. Multidimensional Data Modeling, Types of Data Model, Data Modeling Techniques, Fact Table, Dimension Table, Typical Dimensional Models, Dimensional Modeling Life Cycle.

**Unit 5**

**Performance Management and Enterprise Reporting 12 hours**

Understanding Measures and Performance Measurement System Terminology, Navigating a Business Enterprise, Role of Metrics, and Metrics Supply Chain "Fact based Decision Making" and KPIS , KPI Usage in Companies, Where Do Business Metrics and KPIS Come From, Connecting the Dots: Measures to Business Decisions. Basics of Enterprise Reporting, Reporting Perspectives Common to All Levels of Enterprise, Report Standardization and Presentation Practices, Enterprise Reporting Characteristics in OLAP World, Balanced Scorecard, Dashboards, How Do You Create Dashboards?, Scorecards vs. Dashboards, The Buzz Behind Analysis.

**Reference Books:**

1. Prasad, R. N., and Acharya, Seema. (2011). *Fundamentals of Business Analytics*, 1st Edition. New Delhi: Wiley India.
2. Shmueli, Galit, Patel, Nitin R. and Bruce, Peter C. (2011). *Data Mining for Business Intelligence*. New Delhi: Wiley-India.
3. Kimball, Ralph and Ross, Margy. (2011). *Practical Tools for Data Warehousing and Business Intelligence*, 2nd Edition. New Delhi: Wiley-India.
4. Biere, Mike. (2003*). Business Intelligence for the Enterprise*, First Edition. IBM Press.

**PGDM105: Spreadsheet Modeling and Decision Analysis**

Learning hours: 60

Credits: 6



**Learning Objectives:**

To teach students the theoretical approach to information technology and management information systems within a business/computer science environment. Practical business computer applications will emphasize word processing, spreadsheets, databases, presentation software, and electronic commerce of the Internet.

**Unit 1**

**Introduction to Excel 12 hours**

* Excel Introduction, Excel 2010, Excel 2007 Introduction, Excel 2003 Introduction.
* Basic formulae: Entering and editing data, Creating and copying formulae, Creating functions easily.
* Formatting your work: Inserting/deleting rows/columns, Changing fonts, Colours and borders, Merging and aligning cells.
* Printing: Page break preview, Using page layout view, Headers and footers, Freezing print titles.
* Charts: Selecting data, Quick ways to create charts, Formatting your chart
* Basic tables: Table styles, Using calculated columns, Header rows and total rows, Sorting and simple filtering.
* Using range names: Using IF conditional functions, Using VLOOKUP functions, Creating, sorting and filtering lists/tables of data, Creating pivot tables.

**Unit 2**

**Advanced Excel 1 12 hours**

* Number formatting: Creating custom formats, The four parts of a format, Scaling numbers.
* Dates and times: How dates and times are stored, Useful date/time functions, Formatting dates and times.
* Conditional formatting: Creating/using cell rules, Data bars and colour sets, Styles and themes, How themes work, Using the default styles, Creating custom styles.
* Validation and protection: Setting cell validation, Protecting cells/worksheets, Grouping and outlining, Cell comments.
* Range names and absolute references: Absolute references ($ symbol), Fixing only the row/column, Creating range names, Labelling ranges automatically.
* IF and LOOKUP functions: The conditional (IF) function, Nested IFs are evil, Lookup tables.

**Unit 3**

**Advanced Excel II 12 hours**

* Advanced lookup functions: Using MATCH to find values, Combining MATCH and INDEX, Using OFFSET, The feared INDIRECT function, Linking to other data, Linking to Access, Linking to web sites, Multiple worksheets, Copying, moving and renaming, The dangerous group mode, Summing across worksheets, 3-D range names.
* Basic tables: Table styles, Using calculated columns, Header rows and total rows, Sorting and simple filtering, Advanced Tables, Removing duplicates, Advanced filters, Creating table styles.
* Pivot tables: Creating pivot tables, Swapping rows, columns and pages, Grouping fields, Drill-down, Slicers, Pivot table slicers, Changing slicer properties.
* Charts: Selecting data, Quick ways to create charts, Formatting your chart, Advanced charts, Creating chart templates, Combination charts, Picture charts, Custom chart types, Regression and trendlines, Sparklines, Creating sparklines, Changing chart types, Formatting sparklines, Consolidation , Using formulae, Using arrays, Using data consolidation, Using pivot tables, Recording Macros, Absolute/relative recording, Where macros are stored, Assigning macros to buttons.

**Unit 4**

**Excel Business Modelling 12 hours**

* Building models: Freezing windows, What makes a good formula, Separating inputs and calculations, Avoiding the IF function, Maintaining constant periodicity.
* Masking: Avoiding the IF function, Creating a mask.
* Range names and absolute references: Absolute references ($ symbol), Fixing only the row/column, Creating range names, Labelling ranges automatically, Styles and themes, How themes work, Using the default styles, Creating custom styles.
* Number formatting: Creating custom formats, The four parts of a format, Scaling numbers.
* Dates and times: How dates and times are stored, Useful date/time functions, Formatting dates and times.
* Conditional formatting: Creating/using cell rules, Data bars and colour sets.
* Validation and protection: Setting cell validation, Protecting cells/worksheets, Cell comments, Grouping and outlining.
* Outlining: Creating outlines, Improving how symbols appear, Multiple worksheets, Copying, moving and renaming, The dangerous group mode, Summing across worksheets, 3-D range names.
* Array formulae: Creating array formulae, Editing and deleting array formulae, Strengths and weaknesses.
* Cashflow calculations: Avoiding circularity, Separating interest paid and received, Dealing with compound interest rates, Getting quarterly summary figures.
* Investment appraisal: Net present value, Internal rate of return, Non-periodic cashflows.
* IF and LOOKUP functions: The conditional (IF) function, Nested IFs are evil, Lookup tables.
* Advanced lookup functions: Using MATCH to find values, Combining MATCH and INDEX, Using OFFSET, The feared INDIRECT function, Advanced formulae tricks, Using SUMPRODUCT, Using dynamic ranges.
* Other functions: Text functions in Excel, Inspection functions, INFO and CELL
* Scenarios: The scenario drop-down tool, What-if analysis, Creating a scenario report, Goal-seeking, Targeting a given value, Setting number of iterations.
* Data tables: One-way data tables, Two-way data tables.

**Unit 5**

**Power Pivot for Excel 12 hours**

* What Power Pivot is, Installing Power Pivot, The Power Pivot ribbon.
* Connecting to tables: Connecting to SQL Server, Filtering rows and columns, Changing a table after import.
* Creating a data model: Hiding columns and tables, Renaming columns and tables
* Relationships: Diagram view, Auto-detection, Manually creating, Pivot tables, Creating pivot tables, Sorting pivot tables, Formatting pivot tables, Slicers.
* Connecting to Excel: Creating Excel tables, Linking to Excel tables.
* Connecting to other data: Linking to Access, Pasting tables, Linking to SSRS reports.
* Calculated columns: Creating aggregator columns, IF and SWITCH, Using the BLANK function, Using RELATED to link tables.
* Measures: Measures vs. calculated columns, Creating aggregation measures, Problems with ratios, X suffix functions (eg SUMX).
* Filtering tables: The FILTER function, The CALCULATE function, ALL and ALLEXCEPT, Using the EARLIER function.
* Creating a calendar: Why use a calendar table?, Creating in Excel, Creating in SQL, Multiple date tables.
* Date calculations: Working days, Period to date, Prior and parallel years, Semi-additive measures.
* Hierarchies: Creating a hierarchy, Using in pivot tables.
* Key Performance Indicators: Three parts of a KPI, Creating the measures needed, Creating a KPI, KPIs with absolute targets.

**Reference Books:**

1. [Walkenbach](http://www.flipkart.com/author/john-walkenbach), J. (2013). *Microsoft Excel 2010 Bible*. New Delhi: Wiley India Pvt Ltd.
2. Elmasri, R., Navathe, S.B. (2003). *Fundamentals of Data Base Systems*. New Delhi: Pearson.
3. Saxena, S. (2000). *A first course in computers*. New Delhi: Vikas Publishing House Pvt. Ltd.
4. O'Brein, J. (1991). *Management Information System*. New Delhi: Tata Mcgraw Hill.

**PGDM106: Research Methodology**

Learning hours: 60

Credits: 6



**Learning Objectives:**

This subject will make the students familiar with the concepts of Research Methods starting from the variable types, types of research, scale of measurement to actual application with SPSS. The subject will equip the students with the skills of running appropriate SPSS commands and with that of decoding SPSS output for hypothesis testing and for Correlation and Regression Analysis.

**Unit 1**

**Basics of Research**  **12 hours**

Introduction to SPSS: SPSS windows, Creation and saving of SPSS files, Variable view, Data view, Split and select command.

Research Types: Exploratory and Experimental Research, Qualitative and Quantitative Research, Selection of Scale, Comparative and Non Comparative Scaling Techniques, Likert Scale, Mathematically Derived Scales, Choosing a Scaling Technique, Randomness. Run test for randomness.

**Unit 2**

**Questionnaire Design** **12 hours**

Questionnaire Design Process, Quality of a good questionnaire, type of interviewing method, individual question content, choosing question structure, choosing question wording, Closed and Open question, determining the order of question.

**Unit 3**

**Descriptive Statistics**  **12 hours**

Univariate descriptive analysis: Types of scales, Nominal, Ordinal, Interval and Ratio Scales. Categorical and Continuous Data. Bar Charts, Pie Charts, Box Plots, Stem and Leaf Diagram, Histogram, Mean Plot. Normalcy tests: Shapiro-Wilk and Kolmogorov Smirnov Test of Normality.

Normality assessment through Histogram, Probability Plot, Q-Q Plot. Outlier’s assessment. SPSS application.

**Unit 4**

**Hypothesis Testing**  **12 hours**

Hypothesis Testing I: Central Limit Theorem, Types I and Type II Error, Level of Significance and Confidence Interval. Null and Alternate Hypotheses, Parametric and Non-parametric tsets. One Sample T-test, Paired Sample T-test, Independent Sample T test. SPSS application.

Hypothesis Testing II: One way ANOVA, Chi-square test, Related Sample Wilcoxon Signed Rank Test, Mann-Whitney Test, Kruskal-Wallis and Friedman ANOVA test. SPSS application.

**Unit 5**

**Correlation and Regression Analysis**  **12 hours**

Correlation and Simple Linear Regression Analysis: Relationship through scatter plots, Coefficient of correlation, Assumptions of Linear Regression, Standard Error of estimate, Adjusted R Square, Goodness of fit, Residual Analysis. Durbin-Watson statistics and Multicollinearity.

**Reference Books:**

1. Bajpai, N. (2009). *Business Statistics*. New Delhi: Pearson.
2. Nargundkar, R. (2003). *Marketing Research.* New Delhi: Tata McGrawHill.
3. Malhotra, N.(2008). *Marketing Research*. New Delhi: Pearson.
4. Kothari, C. R. (2004). *Research Methodology Methods and Techniques*. New Delhi: New age International Publishers.
5. Chaudhary, C. M. (1991). *Research Methodology*. Jaipur: RBSA Publishers.

**Semester II**

**PGDM201: Multivariate Data Analysis – 1**

Learning hours: 60

Credits: 6

**Learning Objectives:**

By the end of the subject, students should be able to understand the theory behind the statistics, Select the appropriate methods in function of the research question, Apply those methods to their data, Interpret and report the results from the analysis, Develop critical thinking of statistics.

**Unit 1**

**Overview of Multivariate Statistics 12 hours**

What is Multivariate Analysis?, Multivariate Analysis in statistical terms, some basic concepts of Multivariate Analysis, a classification of Multivariate Techniques, types of Multivariate Techniques**,** Canonical Correlation, Guidelines for Multivariate Analyses and interpretation, a structured approach to Multivariate model building.

**Unit 2**

**Data Cleaning and Multivariate Techniques 12 hours**

Graphical examination of the Data, missing data, outliers, testing the assumptions of Multivariate Analysis, incorporating Nonmetric data with dummy variables, a managerial overview of the results.

**Unit 3**

**Factor Analysis 12 hours**

What is factor analysis, a hypothetical example of Factor Analysis, Factor Analysis decision process, an illustrative example, a managerial overview.

**Unit 4**

**Multiple Regression 12 hours**

What is Multiple regression analysis?, an example of simple and multiple regression analysis, a decision process for multiple regression analysis, Regression: Mediation & Moderation, a managerial overview.

**Unit 5**

**Multiple Discriminant Analysis and Logistic Regression 12 hours**

What are Discriminant Analysis and Logistic Regression?, analogy ith regression and MANOVA, hypothetical example of Discriminant Analysis, the decision process for Discriminant Analysis, logistic regression: regression with a Binary dependent variable, Dichotomous Dependent Variables, an illustrative example of logistic regression, a managerial overview.

**Reference Books:**

1. Hair, J. F. et al. (2006). *Multivariate Data Analysis*, 6th edition. NJ: Prentice Hall.
2. Aiken, L. S., & West, S. G. (1991). *Multiple Regression: Testing and Interpreting Interactions*. Newbury Park, CA: Sage.
3. Menard, S. (2002). *Applied Logistic Regression Analysis*. Thousand Oaks, CA: Sage.
4. Tabachnick, B. and Fidell, L (2007). *Using Multivariate Statistics*, 5th edition. New York: Allyn & Bacon.
5. Cohen, J., & Cohen, P. (1983). *Applied Multiple Regression/Correlation Analysis for the Behavioural Sciences*, 2nd edition. Hillsdale, N.J.: Erlbaum.
6. Stevens, J. (1992). *Applied Multivariate Statistics for the Social Sciences*, 2nd edition. Hillsdale, N.J.: L. Erlbaum Associates.

**PGDM202: Multivariate Data Analysis – 2**

Learning hours: 60

Credits: 6

**Learning Objectives:**

By the end of the subject, students should be able to understand the theory behind the statistics, Select the appropriate methods in function of the research question, Apply those methods to their data, Interpret and report the results from the analysis, Develop critical thinking of statistics.

**Unit 1**

**Cluster Analysis 12 hours**

What is Cluster Analysis, Cluster Analysis as a multivariate technique, conceptual development with Cluster Analysis, necessity of conceptual support in Cluster Analysis, how does Cluster Analysis work? Objective versus subjective considerations, Cluster Analysis decision process, objective of Cluster Analysis, research design in Cluster Analysis, assumptions in Cluster Analysis, deriving clusters and assessing overall fit, interpretation of the clusters, validation and profiling of the clusters, an illustrative example.

**Unit 2**

**Multidimensional Scaling and Correspondence Analysis 12 hours**

What is Multidimensional Scaling? MDS works, comparing MDS to other interdependence techniques, a decision framework for perceptual mapping Overview of Multidimensional Scaling, Correspondence Analysis, illustrations of MDS and Correspondence Analysis.

**Unit 3**

**MANOVA 12 hours**

MANOVA: Extending Univariate Methods for Assessing Group Differences, a hypothetical illustration of MANOVA, a decision process for MANOVA, illustration of MANOVA analysis, example 1: difference between two independent groups, example 2: difference between K independent groups, example 3: a factorial design for MANOVA with tow independent variables, a managerial overview of the results.

**Unit 4**

**Structural Equation Modeling Part I 12 hours**

What is Structural Equation Modeling?, SEM and other multivariate techniques, the role of theory in structural equation modelling, history of SEM, a sample example of SEM, developing a modelling strategy, six stages in structural equation modelling, what is confirmatory factor analysis, SEM stages for testing measurement theory validation with CFA, advanced topics in CFA, advanced CFA illustration.

**Unit 5**

**Structural Equation Modeling Part II 12 hours**

What is structural model, a sample example of structural model, an overview of theory testing with SEM, stages in testing structural theory, SEM illustration.

**Reference Books:**

1. Hair, J. F. et al. (2006). *Multivariate Data Analysis*, 6th edition. NJ: Prentice Hall.
2. Aiken, L. S., & West, S. G. (1991). *Multiple Regression: Testing and Interpreting Interactions*. Newbury Park, CA: Sage.
3. Menard, S. (2002). *Applied Logistic Regression Analysis*. Thousand Oaks, CA: Sage.
4. Tabachnick, B. and Fidell, L (2007). *Using Multivariate Statistics*, 5th edition. New York: Allyn & Bacon.
5. Cohen, J., & Cohen, P. (1983). *Applied Multiple Regression/Correlation Analysis for the Behavioural Sciences*, 2nd edition. Hillsdale, N.J.: Erlbaum.
6. Stevens, J. (1992). *Applied Multivariate Statistics for the Social Sciences*, 2nd edition. Hillsdale, N.J.: L. Erlbaum Associates.

**PGDM203: Big Data Analytics**

Learning hours: 60

Credits: 6

**Learning Objectives:**

By the end of the course, the student will be able to:

* Deploy a structured lifecycle approach to data science and big data analytics projects.
* Select techniques and tools to analyze big data and create statistical models.

**Unit 1**

**Introduction 12 hours**

Big Data Overview, Definition with Real Time Examples, How BigData is generated with Real Time Generation, Use of BigData, Future of BigData!, the challenges for processing big data, technologies supporting big data,

**Unit 2**

**Hadoop 12 hours**

Why Hadoop?, What is Hadoop?, Hadoop vs RDBMS, Hadoop vs BigData, Brief history of Hadoop, Problems with traditional large-scale systems, Requirements for a new approach, Anatomy of a Hadoop cluster

**Unit 3**

**Hadoop Distributed File System (HDFS) 12 hours**

Concepts & Architecture, Data Flow (File Read , File Write), Fault Tolerance, Shell Commands, Java Base API, Data Flow Archives, Coherency, Data Integrity, Role of Secondary NameNode

**Unit 4**

**MapReduce 12 hours**

Theory, Data Flow (Map – Shuffle - Reduce), MapRed vs MapReduce APIs, Programming [ Mapper, Reducer, Combiner, Partitioner]

**Unit 5**

**HIVE, PIG and HBase 12 hours**

Architecture, Installation, Configuration, Hive vs RDBMS, Tables, DDL & DML, Partitioning & Bucketing, Hive Web Interface, Why Pig, Use case of Pig, Pig Components

Data Model, Pig Latin

RDBMS Vs NoSQL, HBase Introduction, HBase Components Scanner, Filter Hbase POC, Introduction to MongoDB.

**Reference Books:**

1. Tan, Pang-Ning, Steinbach, Michael and Kumar, Vipin. (2005). *Introduction to Data Mining*. Addison-Wesley.
2. Lin, Jimmy and Dyer, Chris. (2010). *Data-Intensive Text Processing with MapReduce.* Morgan & Claypool Publishers.
3. Rajaraman, Anand and Ullman, Jeff. (2008). *Mining of Massive Datasets*. New York: Cambridge Press.
4. Han, Jiawei and Kamber, Micheline. (2000). *Data Mining: Concepts and Techniques*. The Morgan Kaufmann Series in Data Management Systems, Jim Gray, Series Editor Morgan Kaufmann Publishers.

**PGDM204: Data Mining and Optimization**

Learning hours: 60

Credits: 6

**Learning Objectives:**

At the end of this Subject, students will:

* Explain what the goals and objectives of data mining are and how to conduct a data mining project.
* Have sound knowledge of popular classification techniques, such as decision trees, support vector machines and nearest-neighbor approaches.
* Familiarise with Machine learning models.

**Unit 1 12 hours**

Data Mining for Business Intelligence, Data Mining Concepts and Definitions, Characteristics, and Benefits, How Data Mining Works, Data Mining Applications. Classification of Data Science solutions for Business Problems, Data Requirements.

**Unit 2 12 hours**

Preparing data for mining, Preprocessing of data: Identifying invalid data items, outliers, handling missing values, derived variables, binning, data reduction techniques etc., Unsupervised vs supervised techniques. Examples, Case study: Workhorse of Industry: Logistic regression – use and limitations and Why Machine Learning Tools?, Preparing Model Data: Training, Validation and testing data sets, Handling imbalanced sample problem, Evaluation of Model: Various measures to assess the performance of the model such as Confusion matrix, ROC, Lift chart, decile analysis, Cross-Validation etc., Case Study

**Unit 3 12 hours**

Machine learning models vs classical models : Concept of Machine Learning tools, Tree-based models: Classification and regression Tress, CHAID, Pruning of trees, Robustness of Trees, Issues in Tree-based models. Use of trees in feature selection, Case study-I : Classification Trees, Case study-2: Regression trees

**Unit 4 12 hours**

Introduction: brief history and logic of Artificial Neural Network, Training of neural network using Back Propagation method, Radial Basis Function Network, Issues in Training and Implementation of Neural Network in practice: Preparing model data and Starting values for training Neural Network, Over fitting and number of hidden layers, Interpreting the output, Case study.

**Unit 5 12 hours**

Logic behind Support Vector Machines and Types of Kernels, Computations of support vectors, SVM for classification and regression, Issues in SVM, Case study : Comparison of NN,SVM and Logistic regression, Random forests and its implementation, Variable Importance, Measures for accuracy of Random forest, Comparison of Random forest with Trees and SVM: Case study, Occam’s Razor, Model Selection and Deployment Issues.

**Reference Books:**

1. Turban, Sharda Efraim; Ramesh, Dursun Delen and King, David. (2011). *Business Intelligence: A Managerial Approach*, 2nd Edition. Publisher: Prentice Hall.
2. Han, Jiawei and Kamber, Micheline. (2012). *Data Mining: Concepts and Techniques*, 3rd edition. Morgan Kaufman Publishers.
3. Tang, P.N., Steinback, M. and Kumar, V. (2006). *Introduction to Data Mining*. Addison Wesley.
4. Myatt, Glenn and Johnson, Wayne. (2009). *Making Sense of Data II.* John Wiley& Sons.
5. Rajaraman, Anand. (2011). *Mining of Massive Datasets.* New York: Cambridge University Press.