

# Nagindas Khandwala College



**Revised Syllabus  
And  
Question Paper Pattern  
Of Course  
Of  
Master of Science Information Technology  
(MSc IT) Programme**

**(Department Of IT)  
Part I  
*Semester I***

**Under Autonomy**

***(To be implemented from Academic Year- 2017-2018)***

**Masters In Information Technology (MSc IT) Program**  
*Under Choice Based Credit, Grading and Semester System*  
**Course Structure**

**MSC IT**

*(To be implemented from Academic Year- 2017-2018)*

MSC IT – SEMESTER I							
Course Code	Course	Hrs. of Instruction/Week	Exam Duration (Hours)	Maximum Marks			Credits
				CIE	SEE	Total	
1711PITDM	Data Mining with Introduction to Data Science	4	2 1/2 Hours	40	60	100	04
1712PITDS	Distributed Systems	4	2 1/2 Hours	40	60	100	04
1713PITDA	Data Analysis Tools	4	2 1/2 Hours	40	60	100	04
1714PITST	Software Testing	4	2 1/2 Hours	40	60	100	04
1711PITPR	Data Mining with Introduction to Data Science Practical	4	2 Hours		50	50	02
1712PITPR	Distributed Systems Practical	4	2 Hours		50	50	02
1713PITPR	Data Analysis Tools Practical	4	2 Hours		50	50	02
1714PITPR	Software Testing Practical	4	2 Hours		50	50	02
	<b>TOTAL</b>	<b>32</b>					<b>24</b>

Course Code :	Course	Hrs. of Instruc tion/ week	Exam Duratio n (Hours)	Maximum Marks			Credits
				CIE	SEE	Total	
1711PITDM	<b>Description: Data Mining with Introduction to Data Science</b>	3	2 ½ hrs	25	75	100	4

Sr. No.	Modules / Units
<b>1</b>	<b>UNIT 1</b>
	<p><b>Introduction:</b> Basics of data mining, related concepts, Data mining Techniques. <b>Data:</b> Introduction, Attributes, Data Sets, and Data Storage, Issues Concerning the Amount and Quality of Data,</p> <p><b>Knowledge Representation:</b></p> <p>Data Representation and their Categories: General Insights, Categories of Knowledge Representation, Granularity of Data and Knowledge Representation Schemes, Sets and Interval Analysis, Fuzzy Sets as Human-Centric Information Granules, Shadowed Sets, Rough Sets, Characterization of Knowledge Representation Schemes, Levels of Granularity and Perception Perspectives, The Concept of Granularity in Rules</p>
<b>2</b>	<b>UNIT 2</b>
	<p><b>Data Preprocessing:</b> Descriptive Data Summarization, Data Cleaning, Data Integration and Transformation, Data Reduction, Data Discretization and Concept Hierarchy Generation.</p> <p><b>Mining Frequent Patterns, Associations, and Correlations:</b> Basic Concepts, Efficient and Scalable Frequent Item set Mining Methods, Mining Various Kinds of Association Rules, From Association Mining to Correlation Analysis, Constraint-Based Association Mining</p>
<b>3</b>	<b>UNIT 3</b>
	<p><b>Classification and Prediction:</b> What Is Classification? What Is Prediction? Issues Regarding Classification and Prediction, Classification by Decision Tree Induction, Bayesian Classification, Rule-Based Classification, Classification by Back-propagation, Support Vector Machines, Associative</p>

	<p>Classification: Classification by Association Rule Analysis, Lazy Learners, Other Classification Methods, Prediction, Accuracy and Error Measures, Evaluating the Accuracy of a Classifier or Predictor, Ensemble Methods Increasing the Accuracy, Model Selection</p> <p><b>Cluster Analysis:</b> What Is Cluster Analysis?, Types of Data in Cluster Analysis, A Categorization of Major Clustering Methods, Partitioning Methods, Hierarchical Methods, Density-Based Methods, Grid-Based Methods, Model-Based Clustering Methods, Clustering High-Dimensional Data, Constraint-Based Cluster Analysis, Outlier Analysis</p>
<b>4</b>	<b>UNIT 4</b>
	<p><b>Graph Mining, Social Network Analysis, and Multirelational Data Mining:</b> Graph Mining, Social Network Analysis, Multirelational Data Mining. <b>Mining Object, Spatial, Multimedia, Text, and Web Data:</b> Multidimensional Analysis and Descriptive Mining of Complex Data Objects, Spatial Data Mining, Multimedia Data Mining, Text Mining, Mining the World Wide Web.</p>
<b>5</b>	<b>UNIT 5</b>
	<p><b>Introduction to Data Science:</b></p> <p>Data science in Big Data world, Data Science process , Machine Learning</p>

<b>Reference Books</b>
<b>Description: Data Mining with Introduction to Data Science</b>
<p><b>Text Books:</b></p> <ol style="list-style-type: none"> <li>1. M. H. Dunham. Data Mining: Introductory and Advanced Topics. Pearson Education. 2010.</li> <li>2. Krzysztof J. Cios, W. Pedrycz, R. W. Swiniarski, L.A. Kurgan, " <b>Data Mining</b>" A Knowledge Discovery Approach", Springer</li> <li>3. J. Han and M. Kamber, " Data Mining: Concepts and Techniques", Second Edition, Elsevier, Reprinted</li> </ol>

4. Davy Cielen Arno D.B. Meysman and Mohamed Ali, "Introducing Data Science", Dreamtech press

**References:**

1. Dr.CarolynK.Hamm,"Oracle Data Mining",RampantTechPress, SPD.
2. C.Ballard, Dynamic Warehousing and Data Mining Made Easy,ReddBooks,IBM (SPD)
3. H.Witten and E. Frank. Data Mining: Practical Machine Learning Tools and Techniques. Morgan Kaufmann. 2005.
4. D. Hand, H. Mannila and P. Smyth. Principles of Data Mining. Prentice-Hall. 2001.
5. Z.Tang and J MacLennan, "Data Mining with SQL Server 2005", Wiley
6. Jason Nell, "Machine Learning for Big Data", Wiley

**Practical (1711PITPR)**

1. Show the implementation of Naïve Bayes algorithm.
2. Show the implementation of Decision Tree.
3. Show the implementation of Time Series Algorithm.
4. Show the implementation of Clustering Algorithm.
5. Show the implementation of k-nearest neighbor.
6. Show the implementation of Apriori Algorithm
7. Show the implementation of Association Algorithm.
8. Show the implementation of Text Mining.
9. Show the implementation of Multimedia Mining.
10. Show the implementation of Spatial Mining.

Course Code:	Course	Hrs. of Instruction/week	Exam Duration (Hours)	Maximum Marks			Credits
				CIE	SEE	Total	
<b>1712PITDS</b>	Distributed Systems	3	2 ½ hrs	25	75	100	4

Sr. No.	Modules / Units
1	<p data-bbox="375 282 486 315"><b>UNIT 1</b></p> <p data-bbox="375 349 1300 488"><b>Characterization Of Distributed Systems:</b> Introduction, Examples of Distributed Systems, Trends In Distributed Systems, Focus On Resource Sharing, Challenges, Case Study: The World Wide Web.</p> <p data-bbox="375 521 1300 555"><b>System Models:</b> Physical Models, Architectural Models, Fundamental Models</p>
2	<p data-bbox="375 591 486 624"><b>UNIT 2</b></p> <p data-bbox="375 658 1204 824"><b>Networking And Internetworking:</b> Types Of Network, Network Principles, Internet Protocols, Case Studies: Ethernet, Wifi And Bluetooth.</p> <p data-bbox="375 857 1236 1093"><b>Interprocess Communication:</b> The Api For The Internet Protocols, External Data Representation And Marshalling, Multicast Communication, Network Virtualization: Overlay Networks, Case Study: MPI</p> <p data-bbox="375 1126 1204 1265">JAVA RMI : Creating Distributed Applications Using RMI and JDBC: Understanding Remote Method Invocation (RMI), Creating a Multitier Database Application Using RMI.</p>
3	<p data-bbox="375 1301 486 1335"><b>UNIT 3</b></p> <p data-bbox="375 1368 1300 1507"><b>Name Services:</b> Name services and the Domain Name System, Directory services, Case study: The Global Name Service, Case study: The X.500 Directory Service.</p> <p data-bbox="375 1541 1157 1706"><b>Time And Global States:</b> Clocks, events and process states , Synchronizing physical clocks , Logical time and logical clocks, Global states, Distributed debugging</p> <p data-bbox="375 1740 1189 1906"><b>Coordination And Agreement:</b> Distributed mutual exclusion Elections Coordination and agreement in group communication, Consensus and related problems</p>

<b>4</b>	<b>UNIT 4</b>
	<p><b>Transactions and Concurrency Control</b></p> <p>Introduction, Transactions, Nested transactions, Locks Optimistic concurrency control.</p> <p>Timestamp ordering, Comparison of methods for concurrency control.</p> <p><b>Distributed Transactions</b></p> <p>Introduction, Flat and nested distributed transactions, Atomic commit process, Concurrency control in distributed transactions, Distributed deadlocks, Transaction recovery.</p> <p><b>Replication</b></p> <p>Introduction, System model and group communication, Fault-tolerant services.</p> <p>Case study: The gossip architecture, CODA.</p>
<b>5</b>	<b>UNIT 5</b>
	<p><b>PEER TO PEER SERVICES AND FILE SYSTEM</b> Peer-to-peer Systems – Introduction – Napster and its legacy – Peer-to-peer – Middleware – Routing overlays. Overlay case studies: Pastry, Tapestry- Distributed File Systems –Introduction – File service architecture – Andrew File system. File System: Features-File model -File accessing models – File sharing semantics Naming: Identifiers, Addresses, Name Resolution – Name Space Implementation – Name Caches – LDAP.</p> <p><b>PROCESS &amp; RESOURCE MANAGEMENT</b></p> <p>Process Management: Process Migration: Features, Mechanism – Threads: Models, Issues, Implementation. Resource Management: Introduction- Features of Scheduling Algorithms –Task Assignment Approach – Load Balancing Approach – Load Sharing Approach.</p>

## Reference Books

### Microprocessor Architecture

#### Text book:

1. George Coulouris, Jean Dollimore, Tim Kindberg, Gordon Blair , Distributed Systems - Concepts and Design (Unit I-Unit 5 )
2. Pradeep K Sinha, “Distributed Operating Systems: Concepts and Design”, Prentice Hall of India, 2007.( Unit 5)
3. Dynamic web programming : using Java, JavaScript, and Informix / Graham Harrison. 2000 ISBN: 0130861847.
4. <http://catalogue.pearsoned.co.uk/samplechapter/0130861847.pdf> ( Unit 2)

#### References :

1. Tanenbaum A.S., Van Steen M., “Distributed Systems: Principles and Paradigms”, Pearson Education, 2007.
2. Liu M.L., “Distributed Computing, Principles and Applications”, Pearson Education, 2004.
- 3 Nancy A Lynch, “Distributed Algorithms”, Morgan Kaufman Publishers, USA, 2003.

### Practical (1712PITPR)

1. Write a program for implementing Client Server communication model.
2. Write a program to show the object communication using RMI.
3. Show the implementation of Remote Procedure Call.
4. Write a program to execute any one mutual exclusion algorithm.
5. Write a program to implement any one election algorithm.
6. Show the implementation of any one clock synchronization algorithm.
7. Write a program to implement two phase commit protocol.
8. Database handling using RMI  
Design and develop a distributed Hotel booking application using Java RMI. A distributed hotel booking system consists of the hotel server and the client machines. The server manages hotel rooms booking information. A customer can invoke the following operations at his machine
  - i) Book the room for the specific guest
  - ii) Cancel the booking of a guest
  - iii) Enquire the check in date for the specified customer/guest.
9. Show the distributed file system implementation with manets in NS2 simulator



Course Code	Course	Hrs. of Instruction/Week	Exam Duration (Hours)	Maximum Marks			Credits
				CIE	SEE	Total	
1713PITDA	Core 1: Data Analysis Tools	3	2 1/2 Hours	25	75	100	4

Sr. No.	Modules / Units
1	<b>UNIT 1</b>
	<p><b>Introduction to R:</b> R Basics, Download R and RStudio, Structure of R, R help, Using R functions, Common mistakes of R beginners. Arithmetic with R, Variable assignment, Basic data types in R.</p> <p><b>Vectors:</b> What is a vector, create vector, naming a vector, vector selection</p> <p><b>Matrix:</b> What is a matrix, Naming a matrix, adding row/column, selection of matrix elements, arithmetic with matrices</p>
2	<b>UNIT 2</b>
	<p><b>Factor:</b> introduction to factors, summarizing a factor, ordered factors</p> <p><b>Lists:</b> Need, creation, selecting elements from a list</p> <p><b>Plotting Graphs:</b> R Datasets and Data Frames, Importing CSV files, R Base graphs</p>
3	<b>UNIT 3</b>
	<p><b>PART II : STATISTICS</b></p> <p>Statistics in Modern day: Application of statistics in different fields</p> <p><b>Distributions for description :</b> Moments ,Sample distributions, Using the sample distributions , Non-parametric description</p> <p><b>Linear projections:</b> Principal component analysis, OLS and friends, Discrete variables, Multilevel modeling</p>
4	<b>UNIT 4</b>
	<p><b>Hypothesis testing with the CLT:</b> The Central Limit Theorem, Meet the Gaussian family, Testing a hypothesis, ANOVA, Regression, Goodness of fit.</p>
5	<b>UNIT 5</b>

	<p><b>Maximum likelihood estimation:</b> Log likelihood and friends, Description: Maximum likelihood estimators, Missing data, Testing with likelihoods</p> <p><b>Monte Carlo :</b> Random number generation, Description: Finding statistics for a distribution, Inference: Finding statistics for a parameter, Drawing a distribution, Non-parametric testing</p>
--	---

Reference Books
Data Analysis Tools
<p><b>Reference books:</b></p> <ol style="list-style-type: none"> <li>1. Computational Statistics, James E. Gentle, Springer</li> <li>2. Computational Statistics, Second Edition, Geof H. Givens and Jennifer A. Hoeting, Wiley Publications</li> <li>3. <a href="https://www.rstudio.com/online-learning/">https://www.rstudio.com/online-learning/</a></li> </ol>

Practical (1713PITPR)
<ol style="list-style-type: none"> <li>1. Implementing matrix and vectors</li> <li>2. Summarize a factor</li> <li>3. Graph Plotting</li> <li>4. Implement the statistical distributions</li> <li>5. Implement regression and goodness of fit</li> <li>6. Implement testing with likelihood</li> <li>7. Generate random numbers using Monte Carlo method</li> <li>8. Implementing Non-Parametric testing</li> <li>9. Drawing an Inference</li> <li>10. Implement Non-parametric Testing</li> </ol>

Course Code:	Course	Hrs. of Instruction/week	Exam Duration (Hours)	Maximum Marks			Credits
				CIE	SEE	Total	
1714PITST	Software Testing	3	2 ½ hrs	25	75	100	4

Sr. No.	Modules / Units
1	<b>UNIT 1</b>
	<p><b>Test Basics:</b> Introduction, Testing in the Software Lifecycle, Specific Systems, Metrics and Measurement, Ethics</p> <p><b>Testing Processes:</b> Introduction, Test Process Models, Test Planning and Control, Test Analysis and Design, Non-functional Test Objectives, Identifying and Documenting Test Conditions, Test Oracles, Standards, Static Tests, Metrics, Test Implementation and Execution, Test Procedure Readiness, Test Environment Readiness, Blended Test Strategies, Starting Test Execution, Running a Single Test Procedure, Logging Test Results, Use of Amateur Testers, Standards, Metrics, Evaluating Exit Criteria and Reporting, Test Suite, Defect Breakdown, Confirmation Test Failure Rate, System Test Exit Review, Standards, Evaluating Exit Criteria and Reporting Exercise, System Test Exit Review, Test Closure Activities</p>
2	<b>UNIT 2</b>
	<p><b>Test Management:</b> Introduction, Test Management Documentation, Test Plan Documentation Templates, Test Estimation, Scheduling and Test Planning, Test Progress Monitoring and Control, Business Value of Testing, Distributed, Outsourced, and Insourced Testing, RiskBased Testing, Risk Management, Risk Identification, Risk Analysis or Risk Assessment, Risk Mitigation or Risk Control, Risk Identification and Assessment Results, Risk-Based Testing throughout the Lifecycle, Risk-Aware Testing Standards, Risk Based Testing Exercise, Project Risk By-Products, Requirements Defect By-Products, Test Case Sequencing Guidelines, Failure Mode and Effects Analysis, Test Management Issues</p>
3	<b>UNIT 3</b>
	<p><b>Test Techniques</b> Introduction, Specification-Based, Equivalence Partitioning, Avoiding Equivalence Partitioning Errors, Composing Test Cases with Equivalence Partitioning, Equivalence Partitioning Exercise, Boundary Value Analysis, Examples of Equivalence Partitioning and Boundary Values, Non-functional Boundaries, Functional Boundaries, Integers, Floating Point Numbers, Testing Floating Point Numbers, Number of Boundaries, Boundary Value Exercise, Decision Tables, Collapsing</p>

	<p>Columns in the, Combining Decision Table Testing with Other Techniques, Nonexclusive Rules in Decision Tables, 4 Decision Table Exercise, Decision Table Exercise Debrief, State-Based Testing and State Transition Diagrams, Superstates and Substates, State Transition Tables, Switch Coverage, State Testing with Other Techniques, State Testing Exercise, State Testing Exercise Debrief, RequirementsBased Testing Exercise, Requirements-Based Testing Exercise Debrief, Structure-Based, Control-Flow Testing, Building Control-Flow Graphs, Statement Coverage, Decision Coverage, Loop Coverage, Hexadecimal Converter Exercise, Hexadecimal Converter Exercise Debrief, Condition Coverage, Decision/Condition Coverage, Modified Condition/Decision Coverage(MC/DC), Multiple Condition Coverage, Control-Flow Exercise, Control-Flow Exercise Debrief, Path Testing, LCSAJ, Basis Path/Cyclomatic Complexity Testing, Cyclomatic Complexity Exercise, Cyclomatic Complexity Exercise Debrief, Final Word on Structural Testing, Structure-Based Testing Exercise, Structure-Based Testing Exercise Debrief, Defect- and Experience-Based, Defect Taxonomies, Error Guessing, Checklist Testing, Exploratory Testing, Test Charters, Exploratory Testing Exercise, Software Attacks, An Example of Effective Attacks, Other Attacks, Software Attack Exercise, Software Attack Exercise Debrief, Specification-, Defect-, and Experience-Based Exercise, Specification-, Defect-,and Experience-Based Exercise Debrief, Common Themes, Static Analysis, Complexity Analysis, Code Parsing Tools, Standards and Guidelines, Data-Flow Analysis, Set-Use Pairs, Set-Use Pair Example, Data-Flow Exercise, Data-Flow Exercise Debrief, Data-Flow Strategies, Static Analysis for Integration Testing, Call-Graph Based Integration Testing, McCabe Design Predicate Approach to Integration Testing, Hex Converter Example, McCabe Design Predicate Exercise, McCabe Design Predicate Exercise Debrief, Dynamic Analysis, Memory Leak Detection, Wild Pointer Detection, API Misuse Detection.</p>
<b>4</b>	<b>UNIT 4</b>
	<p><b>Tests of Software Characteristics</b> Introduction, Quality Attributes for Domain Testing, Accuracy, Suitability, Interoperability, Usability, Usability Test Exercise, Usability Test Exercise Debrief, Quality Attributes for Technical Testing, Technical Security, Security Issues, Timely Information, Reliability, Efficiency, Multiple Flavours of Efficiency Testing, Modelling the System, Efficiency Measurements, Examples of Efficiency Bugs, Exercise: Security, Reliability and Efficiency, Exercise: Security, Reliability, and Efficiency Debrief, Maintainability, Subcharacteristics of Maintainability, Portability, Maintainability and Portability Exercise.</p> <p><b>Reviews</b> Introduction, The Principles of Reviews, Types of Reviews, Introducing Reviews, Success Factors for Reviews, Deutsch’s Design Review Checklist, Marick’s Code Review Checklist, The Open Laszlo Code Review Checklist, Code Review Exercise, Deutsch Checklist Review Exercise.</p> <p><b>Incident Management</b> Introduction, When Can a Defect Be Detected? Defect Lifecycle, Defect Fields, Metrics and Incident Management, Communicating Incidents, Incident Management Exercise.</p>

5	<b>UNIT 5</b>
	<p><b>Standards and Test Process Improvement</b> Introduction, Standards Considerations, Test Improvement Process, Improving the Test Process, Improving the Test Process with TMM, Improving the Test Process with TPI, Improving the Test Process with CTP, Improving the Test Process with STEP, Capability Maturity Model Integration, CMMI, Test Improvement Process Exercise.</p> <p><b>Test Techniques</b> Introduction, Test Tool Concepts, The Business Case for Automation, General Test Automation Strategies, An Integrated Test System Example, Test Tool Categories, Test Management Tools, Test Execution Tools, Debugging, Troubleshooting, Fault Seeding, and Injection Tools, Static and Dynamic Analysis Tools, Performance Testing Tools, Monitoring Tools, Web Testing Tools, Simulators and Emulators, Keyword-Driven Test Automation, Capture/Replay Exercise, Capture/Replay Exercise Debrief, Evolving from Capture/Replay, The Simple Framework Architecture, Data-Driven Architecture, Keyword-Driven Architecture, Keyword Exercise, Performance Testing, Performance Testing Exercise.</p> <p><b>People Skills and Team Composition</b> Introduction, Individual Skills, Test Team Dynamics, Fitting Testing within an Organization, Motivation, Communication.</p>

Reference Books
Software Testing
<p><b>Reference books:</b></p> <ol style="list-style-type: none"> <li>1. Advanced Software Testing—Vol. 3 by Rex Black and Jamie L. Mitchell, Rocky Nook Publication</li> <li>2. Advanced Software Testing Vol. 2 by Rex Black, Rocky Nook Publication, 2008</li> <li>3. Foundations of Software Testing ISTQB Certification by Rex Black, Erik van Veenendaal, Dorothy Graham</li> </ol>

Practical (1714PITPR)
<ol style="list-style-type: none"> <li>1. Evaluating Test Exit Criteria and Reporting</li> <li>2. Static testing using tool</li> <li>3. Rate Quality Attributes for Domain and Technical Testing</li> <li>4. Perform Review</li> <li>5. Incident Management</li> <li>6. Black Box Testing Technique</li> <li>7. White Box Testing Technique</li> <li>8. Performance Testing</li> <li>9. Using Testing Tool Selenium</li> <li>10. Using Selenium Webdriver</li> <li>11. Using Testing Tool ZAPTEST</li> </ol>

## Evaluation Scheme

### I. Internal Exam-40 Marks

**(i) Test– 30 Marks** - Duration 60 mins

It will be conducted either as a written test or using any open source learning management system such as Moodle (Modular object-oriented dynamic learning environment) Or a test based on an equivalent online course on the contents of the concerned course(subject) offered by or build using MOOC (Massive Open Online Course) platform.

**(ii) 10 Marks** – Presentation and active participation in routine class instructional deliveries

Overall conduct as a responsible student, manners, skill in articulation, leadership qualities demonstrated through organizing co-curricular activities, etc.

### II. External Examination- 60 Marks

**(i) Duration** - 2.5 Hours.

**(ii) Theory question paper pattern:-**

All questions are compulsory.		
Question	Based on	Marks
Q.1	Unit I	12
Q.2	Unit II	12
Q.3	Unit III	12
Q.4	Unit IV	12
Q.5	Unit V	12

- All questions shall be compulsory with internal choice within the questions.
- Each Question may be sub-divided into sub questions as a, b, c, d & e, etc & the allocation of Marks depends on the weightage of the topic.

### III. Practical Examination – 50 marks (Duration: 2 Hours)

- Each practical course carries 50 Marks : 40 marks + 05 marks (journal)+ 05 marks(viva)
- Minimum 75% practical from each core/allied course are required to be completed and written in the journal.

**(Certified Journal is compulsory for appearing at the time of Practical Exam)**