

Table of contents :

| | |
|--|---------|
| Organisation and Features..... | Page 7 |
| Review Questions..... | Page 8 |
| Chapter Summary..... | Page 9 |
| Contents..... | Page 12 |
| Databases in Context..... | Page 23 |
| 1.1 Introduction..... | Page 24 |
| 1.3 The Rise of Organisations..... | Page 25 |
| 1.5 Internationalisation..... | Page 26 |
| 1.6 Industrialisation..... | Page 27 |
| 1.7 Mass Transport..... | Page 28 |
| 1.8 Communication..... | Page 29 |
| 1.9 Stocks and Shares..... | Page 30 |
| 1.11 The Challenges of Multinational Operations..... | Page 32 |
| 1.12 The Data Asset..... | Page 33 |
| 1.13 Electronic Storage..... | Page 36 |
| 1.14 Big Data..... | Page 37 |
| 1.16 Data, Data Everywhere..... | Page 38 |
| 1.18.2 Group Work Research Activities..... | Page 40 |
| Further Reading..... | Page 41 |
| 2.1 Introduction..... | Page 43 |
| 2.3 Sequential Systems..... | Page 44 |
| 2.5 Origins of Modern Databases..... | Page 46 |
| 2.6 Transaction Processing and ACID..... | Page 47 |
| 2.7 Two-Phase Commit..... | Page 48 |
| 2.8 Hierarchical Databases..... | Page 49 |
| 2.9 Network Databases..... | Page 50 |
| 2.10 Relational Databases..... | Page 51 |
| 2.11 Object Oriented Databases..... | Page 52 |
| 2.12 Data Warehouse..... | Page 53 |
| 2.13 The Gartner Hype Cycle..... | Page 54 |
| 2.14 Big Data..... | Page 55 |
| 2.16 The Need for Speed..... | Page 56 |
| 2.18 NoSQL..... | Page 57 |
| 2.19 Spatial Databases..... | Page 58 |
| 2.21 Distributed Databases..... | Page 59 |
| 2.22 XML..... | Page 60 |
| 2.25.1 Review Questions..... | Page 62 |
| Further Reading..... | Page 63 |
| 3.1 The Fundamental Building Block..... | Page 64 |
| 3.2.1 In-Memory Structures..... | Page 65 |
| 3.2.2 Walking Through a Straightforward Read..... | Page 66 |
| 3.2.3 Server Processes..... | Page 67 |
| 3.2.4 Permanent Structures..... | Page 69 |
| 3.3 Data Storage..... | Page 70 |
| 3.3.1 Row Chaining and Migration..... | Page 72 |
| 3.5 Control, Redo and Undo..... | Page 73 |
| 3.6 Log and Trace Files..... | Page 74 |

| | |
|--|----------|
| 3.7 Stages of Start-up and Shutdown..... | Page 75 |
| 3.8 Locking..... | Page 76 |
| 3.9 Moving Data..... | Page 77 |
| 3.10 Import and Export..... | Page 79 |
| 3.11 Distributed Databases..... | Page 80 |
| 3.13 Review Questions..... | Page 83 |
| 3.14 Group Work Research Activities..... | Page 85 |
| Reference..... | Page 87 |
| Database Types..... | Page 88 |
| 4.1 Origins..... | Page 89 |
| 4.2 Normalisation..... | Page 90 |
| 4.2.1 First Normal Form (1NF)..... | Page 91 |
| 4.3 Second Normal Form (2NF)..... | Page 92 |
| 4.4 Third Normal Form (3NF)..... | Page 93 |
| 4.5 Beyond Third Normal Form..... | Page 95 |
| 4.7 Use Case Modelling..... | Page 97 |
| 4.8 Further Modelling Techniques..... | Page 103 |
| 4.9 Notation..... | Page 105 |
| 4.10 Converting a Design into a Relational Database..... | Page 107 |
| 4.12 Create the Tables..... | Page 108 |
| 4.14 Populate the Tables..... | Page 111 |
| 4.16 Joins..... | Page 112 |
| 4.17 More Complex Data Retrieval..... | Page 115 |
| 4.18 UPDATE and DELETE..... | Page 116 |
| 4.20 Group Work Research Activity..... | Page 117 |
| Further Reading..... | Page 118 |
| 5.1 Databases and the Web..... | Page 119 |
| 5.2 The NoSQL Movement..... | Page 120 |
| 5.2.1 What Is Meant by NoSQL?..... | Page 122 |
| 5.3 Differences in Philosophy..... | Page 123 |
| 5.4 Basically Available, Soft State, Eventually Consistent (BASE)..... | Page 125 |
| 5.5 Column-Based Approach..... | Page 126 |
| 5.6 Examples of Column-Based Using Cassandra..... | Page 127 |
| 5.6.2 Data Sources..... | Page 128 |
| 5.6.3 Getting Started..... | Page 129 |
| 5.6.4 Creating the Column Family..... | Page 132 |
| 5.6.6 Retrieving Data..... | Page 133 |
| 5.6.8 Command Line Script..... | Page 136 |
| 5.7 CQL..... | Page 137 |
| 5.7.1 Interactive CQL..... | Page 139 |
| 5.7.3 Timings..... | Page 141 |
| 5.8 Document-Based Approach..... | Page 142 |
| 5.8.1 Examples of Document-Based Using MongoDB..... | Page 143 |
| 5.8.3 Getting Started..... | Page 144 |
| 5.8.5 Creating a Collection..... | Page 145 |
| 5.8.6 Simple Inserting and Reading of Data..... | Page 146 |
| 5.8.7 More on Retrieving Data..... | Page 149 |
| 5.8.8 Indexing..... | Page 150 |

| | |
|---|----------|
| 5.8.9 Updating Data..... | Page 151 |
| 5.8.10 Moving Bulk Data into Mongo..... | Page 152 |
| 5.9 Summary..... | Page 153 |
| 5.11.1 Sample Solutions..... | Page 154 |
| 5.11.2 MongoDB Crib..... | Page 155 |
| References..... | Page 156 |
| Data Sources..... | Page 157 |
| 6.1 What is Big Data?..... | Page 158 |
| 6.2.1 The Four Vs..... | Page 160 |
| 6.2.2 The Cloud Effect..... | Page 162 |
| 6.3.1 So Why Isn't Big Data just Called Data Warehousing 2?..... | Page 164 |
| 6.3.2 What is a Data Scientist?..... | Page 167 |
| 6.4 Big Data Tools..... | Page 168 |
| 6.4.1 MapReduce..... | Page 169 |
| 6.4.2.1 If You Should Want to Explore Hadoop..... | Page 170 |
| 6.5 Getting Hands-On with MapReduce..... | Page 171 |
| 6.6 Using MongoDB's db.collection.mapReduce() Method..... | Page 172 |
| 6.6.1 And if You Have Time to Test Your MongoDB and JS Skills..... | Page 177 |
| 6.6.2 Sample Solutions..... | Page 178 |
| 6.9 Group Work Research Activities..... | Page 179 |
| References..... | Page 180 |
| 7.1 Querying Data..... | Page 181 |
| 7.2 Problems with Relational Databases..... | Page 182 |
| 7.3 What Is an Object?..... | Page 183 |
| 7.4 An Object Oriented Solution..... | Page 184 |
| 7.5 XML..... | Page 187 |
| 7.7 What Is Object Relational?..... | Page 189 |
| 7.8 Classes..... | Page 190 |
| 7.9 Pointers..... | Page 192 |
| 7.9.1 Hierarchies and Inheritance..... | Page 195 |
| 7.9.2 Aggregation..... | Page 197 |
| 7.9.3 Encapsulation and Polymorphism..... | Page 198 |
| 7.9.4 Polymorphism..... | Page 199 |
| 7.9.5 Support for Object Oriented and Object Relational Database Development..... | Page 200 |
| 7.9.6.2 Group Work Research Activities..... | Page 201 |
| References..... | Page 202 |
| Further Reading..... | Page 203 |
| 8.1 Introduction..... | Page 204 |
| 8.2 Origins..... | Page 205 |
| 8.4 Interim Solution—Create a RAM Disk..... | Page 206 |
| 8.5 Interim Solution—Solid State Drive (SSD)..... | Page 207 |
| 8.6 In-Memory Databases—Some Misconceptions..... | Page 208 |
| 8.7 In-Memory Relational Database—The Oracle TimesTen Approach..... | Page 209 |
| 8.8 In-Memory Column Based Storage—The SAP HANA Approach..... | Page 211 |
| 8.9 In-Memory On-Line Transaction Processing—The Starcounter Approach..... | Page 213 |
| 8.9.2 In Memory Databases and Personal Computers..... | Page 214 |
| 8.10.1 Review Questions..... | Page 216 |
| References..... | Page 217 |

| | |
|---|----------|
| Further Reading..... | Page 218 |
| 9.1 Introduction..... | Page 219 |
| 9.2 Hierarchical Data Model..... | Page 220 |
| 9.4 Hierarchical Data Paths..... | Page 221 |
| 9.6 Hierarchical Queries..... | Page 222 |
| 9.8 Applications of Hierarchical Databases..... | Page 223 |
| 9.9 Summary..... | Page 224 |
| References..... | Page 225 |
| 10.1 Introduction..... | Page 227 |
| 10.4 Distributed Data Storage..... | Page 229 |
| 10.5 Distributed Query Processing..... | Page 230 |
| 10.7 Examples of Distributed Databases..... | Page 231 |
| 10.9 Security of Distributed Databases..... | Page 232 |
| 10.10 Examples of Applications of Distributed Databases..... | Page 234 |
| References..... | Page 235 |
| 11.1 Introduction..... | Page 237 |
| 11.4 Examples of Graph Databases..... | Page 240 |
| 11.5 Graph Databases Search Algorithms..... | Page 241 |
| 11.7 Graph Visualization..... | Page 242 |
| 11.8 Applications of Graph Databases..... | Page 243 |
| 11.10 Review Questions..... | Page 246 |
| References..... | Page 247 |
| What Database Professionals Worry About..... | Page 250 |
| 12.1 What Do We Mean by Scalability?..... | Page 251 |
| 12.2 Coping with Growing Numbers of Users..... | Page 252 |
| 12.2.1 So Why Can't Access Cope with Lots of Users?..... | Page 253 |
| 12.2.2 Client/Server..... | Page 254 |
| 12.2.3 Scalability Eras..... | Page 256 |
| 12.2.3.1 Separating Users from Sessions..... | Page 257 |
| 12.2.3.2 Adding Tiers..... | Page 258 |
| 12.3 Coping with Growing Volumes of Data..... | Page 259 |
| 12.3.2 Vertical and Horizontal Scaling..... | Page 260 |
| 12.3.3 Database Scaling Issues..... | Page 261 |
| 12.3.4 Single Server Solutions..... | Page 262 |
| 12.3.4.1 Partitioning..... | Page 263 |
| 12.3.5 Distributed RDBMS: Shared Nothing Versus Shared Disk..... | Page 264 |
| 12.3.6 Horizontal Scaling Solutions..... | Page 265 |
| 12.3.7 Scaling Down for Mobile..... | Page 267 |
| 12.5 Review Questions..... | Page 268 |
| References..... | Page 269 |
| 13.1 What Do We Mean by Availability?..... | Page 270 |
| 13.2 Keeping the System Running—Immediate Solutions to Short Term Problems..... | Page 271 |
| 13.2.1.1 Database Software Related Problems..... | Page 272 |
| 13.2.1.2 Systems Design Problems and User Mistakes..... | Page 273 |
| 13.2.1.3 Technology Failures..... | Page 274 |
| 13.2.1.4 CPU, RAM and Motherboards..... | Page 275 |
| 13.2.1.5 Using RAID to Improve Availability..... | Page 276 |
| 13.2.1.7 Proactive Management..... | Page 277 |

| | |
|--|----------|
| 13.3 Back-Up and Recovery..... | Page 278 |
| 13.3.1.1 Recovery..... | Page 281 |
| 13.3.1.2 RMAN and Enterprise Manager..... | Page 282 |
| 13.4 Disaster Recovery (DR)..... | Page 283 |
| 13.4.2 High Availability for Critical Systems..... | Page 286 |
| 13.4.4 The Challenge or Opportunity of Mobile..... | Page 288 |
| 13.7 Group Work Research Activities..... | Page 289 |
| References..... | Page 290 |
| 14.1 What Do We Mean by Performance?..... | Page 291 |
| 14.2 A Simplified RDBMS Architecture..... | Page 293 |
| 14.3.1 Block Size..... | Page 295 |
| 14.3.3 Alternatives to the HDD..... | Page 296 |
| 14.3.4 Operating System (OS)..... | Page 297 |
| 14.3.5 Database Server Processes..... | Page 298 |
| 14.3.7 Schema Level: Data Types, Location and Volumes..... | Page 299 |
| 14.3.8 SQL Optimisation..... | Page 300 |
| 14.3.9 Indexes..... | Page 301 |
| 14.3.9.1 B-Tree Indexes..... | Page 302 |
| 14.3.9.2 Non B-Tree Indexes—Bitmap Indexes..... | Page 303 |
| 14.3.9.3 Non B-Tree Indexes—Reverse Key Index..... | Page 304 |
| 14.3.9.4 Non B-Tree Indexes—Function-Based Indexes..... | Page 305 |
| 14.3.11.1 Design with Performance in Mind..... | Page 306 |
| 14.4.1 What is Database Tuning?..... | Page 308 |
| 14.4.2 Benchmarking..... | Page 309 |
| 14.5.1 Tuning and Performance Tools..... | Page 310 |
| 14.5.1.1 SQLPLUS Tools..... | Page 312 |
| 14.5.2 Using the Built-In Advisers..... | Page 325 |
| 14.6 Summary..... | Page 327 |
| Appendix: Creation Scripts and Hints..... | Page 328 |
| Reference..... | Page 332 |
| 15.1 Introduction..... | Page 333 |
| 15.2 Physical Security..... | Page 334 |
| 15.4 Privilege Abuse..... | Page 336 |
| 15.5 Platform Weaknesses..... | Page 341 |
| 15.6 SQL Injection..... | Page 342 |
| 15.7 Weak Audit..... | Page 343 |
| 15.9 Authentication Vulnerabilities..... | Page 345 |
| 15.11 Mobile Device Based Threats..... | Page 346 |
| 15.12 Security Issues in Cloud Based Databases..... | Page 347 |
| 15.13 Policies and Procedures..... | Page 348 |
| 15.14 A Security Checklist..... | Page 349 |
| 15.15 Review Questions..... | Page 350 |
| References..... | Page 351 |
| 16.1 Introduction..... | Page 353 |
| 16.2.1 Changes in User Requirements..... | Page 354 |
| 16.3 Examples of Adaptive Capabilities of Databases..... | Page 355 |
| 16.3.2 Version Control on Databases..... | Page 356 |
| 16.4.1 Indexing Strategies..... | Page 357 |

| | |
|---|----------|
| 16.4.3 Handling Data Variety..... | Page 358 |
| 16.4.6 Scalability..... | Page 359 |
| 16.5.1 Heterogeneous Database Systems..... | Page 360 |
| 16.5.4 Database Virtualization..... | Page 361 |
| 16.6.1 Social Media Applications..... | Page 362 |
| 16.8 Review Questions..... | Page 363 |
| References..... | Page 364 |
| Advanced Applications of Specialised Databases..... | Page 366 |
| 17.1 Introduction..... | Page 367 |
| 17.2 Blockchain Architecture..... | Page 368 |
| 17.3 Blockchain Protocols..... | Page 370 |
| 17.4 Blockchain Performance..... | Page 371 |
| 17.5 Blockchain Security..... | Page 372 |
| 17.6 Legal Issues in Blockchain..... | Page 374 |
| 17.7 Blockchain Data Analytics..... | Page 375 |
| 17.9 Blockchain Applications in Healthcare..... | Page 376 |
| 17.11 Blockchain Applications in the Transportation Industry..... | Page 377 |
| 17.14 Review Questions..... | Page 378 |
| References..... | Page 379 |
| 18.1 Introduction..... | Page 382 |
| 18.2 Significance of Biological Databases..... | Page 384 |
| 18.4 Types of Biological Databases..... | Page 385 |
| 18.5 Types of Data in Biological Databases..... | Page 386 |
| 18.7 GenBank..... | Page 387 |
| 18.9 UniProt..... | Page 388 |
| 18.12 Review Questions..... | Page 389 |
| References..... | Page 390 |
| 19.1 Introduction..... | Page 392 |
| 19.3 Types of GIS Databases..... | Page 393 |
| 19.4 GIS Databases Data Formats..... | Page 395 |
| 19.6 GIS Water Erosion..... | Page 396 |
| 19.9 Weather Forecasting..... | Page 397 |
| 19.11 Linking GPS Data to GIS Databases..... | Page 398 |
| 19.14 Group Work Research Activities..... | Page 399 |
| References..... | Page 400 |
| Index..... | Page 402 |