



DeepSphere.AI
Enterprise AI and IIoT for Analytics

Advanced Data Preparation and Exploratory Analysis

Part

2

Your first step to becoming a
Data Scientist and Data Engineer

Learning Level: Foundation



Author

Jothi Periasamy

Co-Founder | DeepSphere.AI

Board Member | University of California

Board Member, Chief Data Scientist & Lead Instructor
University of California
Silicon Valley & Davis, CA, USA

Artificial Intelligence Learning Facilitator
MIT CSAIL & MIT Sloan
Boston, MA, USA

Head of Artificial Intelligence
Experfy (Harvard Innovation Lab)
Boston, MA, USA

Academic Council for AI
Universiti Tenaga Nasional University
Selangor, Malaysia

- 17+ years of management consulting and end-to-end AI (ML, DL, RPA, NLP) experience with Deloitte, E&Y, and KPMG
- Over 50+ clients across industries: Retail, Consumer, Health, Energy, Oil & Gas
- Teaches AI at MIT – Head of AI, Architect at Experfy, Harvard Innovation Lab
- Co-innovates industry solution along with SAP America
- Published on finance, AI, Big Data, IIoT, Cloud, and SAP Appliances
- Delivered many real-time research projects. Globally recognized as a keynote speaker at around 50+ conferences
- Recently invited by Prime Minister's office of Dubai, Malaysia, and Montreal
- Stanford alumni

TABLE OF CONTENTS

Unit 1	Data Preparation and Exploration	04
Unit 2	Data Collection	21
Unit 3	Data Cleansing and Data Blending	32
Unit 4	Information Modeling and Data Virtualization	49
Unit 5	Data Catalog / Business Glossary	62
Unit 6	Data Governance and Data Quality	72
Unit 7	Exploratory Analysis and Data Visualisation	87
Unit 8	Industry Use Case Implementation	103
Unit 9	Developmental Standards and Best Practices	118



Unit -1

Data Preparation and Exploration

