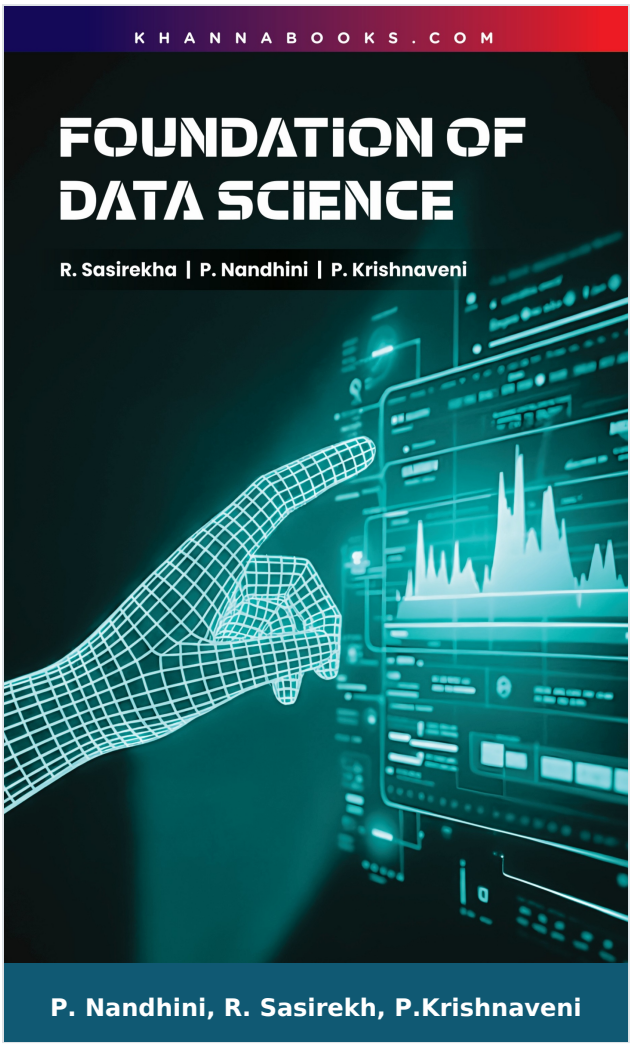


Foundation of Data Science

Author :	P. Nandhini, R. Sasirekh, P.Krishnaveni
ISBN 13 :	978-93-55385-57-4
ISBN 10 :	9355385579
E-ISBN 13 :	978-93-55385-57-4
Edition :	First
Pages :	168
Type of book :	Physical
Year :	2026
Language :	English
Publisher :	Khanna Publishing House
M.R.P :	Rs 248.00
Categories :	New Arrivals, Computer Science Engineering, Book Store, Sathyabama Series
Condition Type :	New
Country Origin :	India

Product Description

Foundation of Data Science is a comprehensive textbook designed to provide readers with a strong foundation in modern data science concepts, techniques, and applications. In today's data-driven world, the ability to collect, process, clean, analyze, and interpret large volumes of data is essential across industries. This book serves as a structured and practical guide to the complete data analysis pipeline, with a particular emphasis on Natural Language Processing (NLP) and Big Data Analytics.

The text begins with the fundamentals of data acquisition, covering methods for collecting data from multiple sources, including internal databases, external repositories, APIs, and web scraping techniques. It then explores critical data preprocessing activities such as data cleaning, transformation, quality control, handling missing values, and noise reduction, enabling readers to prepare datasets for accurate analysis and modeling.

A significant portion of the book is dedicated to Natural Language Processing, introducing key concepts such as Bag-of-Words, Regular Expressions, Tokenization, Stemming, Lemmatization, TF-IDF, and Sentiment Analysis. Through practical examples and case studies, readers learn how to transform unstructured textual information into meaningful insights for decision-making and business intelligence.

The book further extends into Exploratory Data Analysis (EDA), data visualization techniques, and modern Big Data frameworks such as Hadoop and Spark. Combining theoretical understanding with practical implementation, this resource is suitable for students, researchers, data analysts, and professionals seeking to build expertise in data science and text analytics. By integrating NLP methodologies with contemporary Big Data concepts, the book provides a complete toolkit for extracting value from data and succeeding in a data-centric environment.

Salient Features:

Khanna Publishing House

4C/4344, Ansari Road, Daryaganj, New Delhi-110002

Email: contact@khannabooks.com | Tel: 011-2324 44 47 - 48 | Mobile: +91-99109 09320

- Data Sourcing & APIs: Explains data acquisition techniques from internal and external sources, including web scraping and the use of APIs such as Google Maps and IBM Watson for real-time data collection.
- Data Cleaning & Wrangling: Covers systematic approaches to data preprocessing, transformation, data munging, handling noisy datasets, and resolving inconsistencies to improve data quality.
- Core NLP Toolkit: Introduces essential Natural Language Processing concepts, including Bag-of-Words, Regular Expressions, Tokenization, Sentence Splitting, and pattern matching techniques.
- Linguistic Normalization: Clearly differentiates between stemming and lemmatization, demonstrating how linguistic rules and morphological analysis are applied to derive meaningful root forms of words.
- Feature Engineering: Discusses advanced text representation techniques such as TF-IDF (Term Frequency–Inverse Document Frequency), stop-word removal, and feature extraction methods for enhanced analytical performance.
- Practical Visualization: Provides hands-on guidance on data visualization and Exploratory Data Analysis (EDA), including the use of professional tools such as Tableau for effective data presentation.
- Big Data Ecosystem: Introduces the principles, characteristics, and frameworks of Big Data, including technologies such as Hadoop and Spark for managing and processing large-scale datasets.
- Real-World NLP Case Study: Presents a comprehensive Sentiment Analysis case study demonstrating how NLP techniques can be applied to extract opinions, emotions, and actionable insights from textual data across diverse domains.

Khanna Publishing House

4C/4344, Ansari Road, Daryaganj, New Delhi-110002

Email: contact@khannabooks.com | Tel: 011-2324 44 47 - 48 | Mobile: +91-99109 09320

Table Of Contents

- Introduction to Data Acquisition
- Data Pre-Processing
- Reliable Data: Quality Control And Transformation Methods
- Introduction to NLP
- Introduction to Big Data

Author

P. Nandhini

P. Nandhini P. Krishnaveni

R. Sasirekh

R. Sasirekha P. Nandhini P. Krishnaveni

P.Krishnaveni

R. Sasirekha P. Nandhini P. Krishnaveni

Khanna Publishing House

4C/4344, Ansari Road, Daryaganj, New Delhi-110002

Email: contact@khannabooks.com | Tel: 011-2324 44 47 - 48 | Mobile: +91-99109 09320