ITSB | Eco Industry Oriented University

sinarmas

ADVANCED DATA SCIENCE PROGRAMMING

Study Case Using R and Python

R

Writer:

Bakti Siregar, M.Sc., CDS.

D'SCIENCELABS
Smart Idea

Kampus Merdeka
INDONESIA JAYA

First Edition

Advanced Data Science Programming
Study Case Using R and Python

Bakti Siregar, M.Sc.,CDS

Table of contents

In today's digital era, data is a strategic asset that drives decision-making, innovation, and competitive advantage across industries. Advanced Data Science Programming equips professionals with the skills to design scalable workflows, build predictive models, and deploy intelligent systems that transform raw data into actionable insights.

This module goes beyond the fundamentals, focusing on advanced concepts and practices in programming for data science. It introduces modularization and object-oriented programming (OOP) as the foundation for writing clean, reusable, and maintainable code. Readers will also explore API integration for accessing diverse data sources, as well as advanced data wrangling and feature engineering techniques to prepare high-quality datasets for analysis.

The module further delves into predictive modeling and interactive visualization, enabling the creation of models that not only generate accurate forecasts but also communicate results effectively to stakeholders. Emphasis is placed on debugging, testing, and workflow automation, ensuring that projects are reliable, efficient, and reproducible. Finally, learners will gain practical experience in deployment and model monitoring, mastering the tools and methods required to bring data science solutions into production environments and track their long-term performance.

By completing this module, readers will be equipped with the end-to-end programming capabilities needed to tackle real-world data challenges and deliver impactful solutions in research, industry, and beyond.

Preface

About the Writer

[Bakti Siregar, M.Sc., CDS](#) works as a Lecturer at the [ITSB Data Science Program](#). He earned his Master's degree from the Department of Applied Mathematics at National Sun Yat Sen University, Taiwan. In addition to teaching, Bakti also works as a Freelance Data Scientist for leading companies such as [JNE](#), [Samora Group](#), [Pertamina](#), and [PT. Green City Traffic](#).

He has a strong enthusiasm for projects (and teaching) in the fields of Big Data Analytics, Machine Learning, Optimization, and Time Series Analysis, particularly in finance and investment. His core expertise lies in statistical programming languages such as R Studio and Python. He is also experienced in implementing database systems like MySQL/NoSQL for data management and is proficient in using Big Data tools such as Spark and Hadoop.

Some of his projects can be viewed here: [Rpubs](#), [Github](#), [Website](#), and [Kaggle](#)

Acknowledgments

Advanced Data Science Programming plays a vital role in building scalable, reliable, and impactful data-driven solutions. This module is designed to strengthen programming capabilities beyond the basics and covers key areas such as:

- Writing modular and maintainable code through Modularization & OOP
- Leveraging APIs and data integration for diverse data sources
- Applying advanced wrangling and feature engineering to prepare high-quality datasets
- Developing predictive models and interactive visualizations for actionable insights
- Ensuring robust debugging, testing, workflow automation, and deployment for real-world use

This book is intended for learners who already possess fundamental programming knowledge and wish to advance their skills in designing, implementing, and deploying end-to-end data science solutions. I am deeply grateful to the learners, colleagues, and mentors whose engagement, feedback, and discussions enriched the development of this material. It is my hope that this book serves as both a practical reference and a roadmap for applying advanced programming techniques in data science projects across research, industry, and innovation.

Feedback & Suggestions

Your feedback is invaluable in enhancing the quality and relevance of this module. We warmly encourage readers to share their thoughts on the content, structure, clarity, and practical applicability of the materials. Suggestions for expanding the coverage—whether in advanced techniques, case studies, or tools—are highly appreciated.

With your input, we strive to continually refine this E-book into a more comprehensive and practical resource for Advanced Data Science Programming, supporting both academic and professional applications. Thank you for your active participation and contributions to the growth of this material!

For feedback and suggestions, feel free to contact:

- dsciencelabs@outlook.com

- siregarbakti@gmail.com

- siregarbakti@itsb.ac.id

Chapter 1

Advanced Programming

Chapter 2

Modularization & OOP

Chapter 3

API & Data Integration

Chapter 4

Advanced Data Wrangling

Chapter 5

Feature Engineering

Chapter 6

Predictive Modeling

Chapter 7

Interactive Visualization

Chapter 8

Debugging & Testing

Chapter 9

Workflow Automation

Chapter 10

Deploy & Monitoring