Data Analysis with Scala

ERNESTO .NET

Scala's advanced techniques to solve real-world problems in data analysis and gain valuable insights from your data

What You Will Learn

- Techniques to determine the validity and confidence level of data
- Apply quartiles and n-tiles to datasets to see how data is distributed into many buckets
- Create data pipelines that combine multiple data lifecycle steps
- Use built-in features to gain a deeper understanding of the data

Labs

Labs for this course are available at path shared below.

Path: `~/work/data-analysis-scala`

1. Scala Overview
2. Data Analysis
3. Data Ingestion
4. Data Exploration and Visualization
5. Applying Statistics and Hypothesis Testing
6. Introduction to Spark for Distributed Data Analysis
7. Traditional Machine Learning for Data Analysis
8. Near Real-Time Data Analysis Using Streaming
9. Working with Data at Scale
10. Spark Administration