

## CLOUDERA DATA SCIENCE COURSE CONTENT

- ❖ **DATA SCIENCE OVERVIEW:**
  - What Is Data Science?
  - The Growing Need for Data Science
  - The Role of a Data Scientist
- ❖ **USE CASES:**
  - Finance
  - Retail
  - Advertising
  - Defense and Intelligence
  - Telecommunications and Utilities
  - Healthcare and Pharmaceuticals
- ❖ **PROJECT LIFECYCLE:**
  - Steps in the Project Lifecycle
  - Lab Scenario Explanation
- ❖ **DATA ACQUISITION:**
  - Where to Source Data
  - Acquisition Techniques
- ❖ **EVALUATING INPUT DATA:**
  - Data Formats
  - Data Quantity
  - Data Quality
- ❖ **DATA TRANSFORMATION:**
  - Anonymization
  - File Format Conversion
  - Joining Datasets
- ❖ **DATA ANALYSIS AND STATISTICAL METHODS:**
  - Relationship between Statistics and Probability
  - Descriptive Statistics
  - Inferential Statistics
- ❖ **FUNDAMENTALS OF MACHINE LEARNING:**
  - Overview
  - The Three Cs of Machine Learning
  - Spotlight: Naïve Bayes Classifiers
  - Importance of Data and Algorithms
- ❖ **RECOMMENDER OVERVIEW:**
  - What Is a Recommender System?
  - Types of Collaborative Filtering
  - Limitations of Recommender Systems
  - Fundamental Concepts

❖ **INTRODUCTION TO APACHE MAHOUT:**

- What Apache Mahout Is (and Is Not)
- A Brief History of Mahout
- Availability and Installation
- Demonstration: Using Mahout's Item-Based Recommender

❖ **IMPLEMENTING RECOMMENDERS WITH APACHE MAHOUT :**

- Overview
- Similarity Metrics for Binary Preferences
- Similarity Metrics for Numeric Preferences
- Scoring

❖ **EXPERIMENTATION AND EVALUATION:**

- Measuring Recommender Effectiveness
- Designing Effective Experiments
- Conducting an Effective Experiment
- User Interfaces for Recommenders

❖ **PRODUCTION DEPLOYMENT AND BEYOND :**

- Deploying to Production
- Tips and Techniques for Working at Scale
- Summarizing and Visualizing Results
- Considerations for Improvement
- Next Steps for Recommenders

❖ **CONCLUSION:**

- Appendix A : Hadoop Overview
- Appendix B: Mathematical Formulas
- Appendix C : Language and Tool
- Reference