

**The Evolution of Data Science Roles**

**Course Number:** DATA-130WA  
**Duration:** 1 day

**Overview**

This Evolution of Data Science Roles training course teaches the various roles, responsibilities, and interactions among data engineers, machine learning engineers, data analysts, AI engineers, DevOps engineers, and data architects. In addition, this course covers the skills and tools required for each role and discusses the future of data science roles.

**Prerequisites**

* Basic knowledge of statistics and programming concepts (e.g., familiarity with common programming languages like Python or R)
* An understanding of data analysis fundamentals, such as data cleaning and visualization techniques
* Some awareness of machine learning concepts and algorithms is beneficial but not required
* Prior exposure to basic data management principles and tools (e.g., databases and SQL) is helpful but not required

**Materials**

All Data Science training students receive comprehensive courseware.

**Software Needed on Each Student PC**

A modern web browser and an Internet connection.

**Objectives**

* Understand the evolution and significance of Data Science roles
* Identify the responsibilities and skills required for each specialized role
* Gain insights into the tasks, tools, and techniques used in each role
* Explore the collaborations and interactions between different Data Science roles
* Discuss the impact of these roles on data-driven decision-making and business outcomes
* Recognize potential career opportunities and paths within the Data Science domain

**Outline**

* Introduction to Data Science Roles
  + Overview of the history and evolution of Data Science roles
  + Understanding the intersection of statistical modeling, computer science, and data analysis
  + Importance and relevance of specialized roles in the Data Science domain
* The Data Scientist Role
  + Exploring the origins and responsibilities of Data Scientists
  + Data extraction, cleaning, and analysis techniques
  + Incorporating machine learning models and predictive analytics
  + A/B testing and implementing solutions in production applications
* The Data Engineer Role
  + Understanding the need for a bridge between raw data and Data Scientists
  + Responsibilities of Data Engineers in data collection, storage, and processing
  + Working with big data tools like Hadoop and Spark
  + Managing databases and ETL (Extract, Transform, Load) tools
* The Machine Learning Engineer Role
  + Introduction to Machine Learning Engineers and their role in AI advancements
  + Designing, implementing, and maintaining machine learning systems
  + Leveraging advanced statistical skills and deep learning techniques
  + Collaboration with Data Scientists to take models from conception to production
* The Data Analyst Role
  + The specialization of Data Analysts within the Data Science field
  + Interpreting data and performing statistical analysis
  + Providing ongoing reports and actionable insights
  + Distinction from Data Scientists in complexity of models vs. operational decision-making focus
* The AI Engineer Role
  + Overview of the recent emergence of AI Engineers in the Data Science ecosystem
  + Combining software engineering skills with AI and machine learning principles
  + Building and maintaining AI systems, programming and training models
* DevOps and Architects in Data Science
  + Understanding the roles of DevOps Engineers and Data Architects
  + Streamlining operations and processes, automating deployments
  + Data management infrastructure design and ensuring data quality/security
  + Integration of these roles in maintaining and improving data science operations
* Future Trends
* Recap of the evolving Data Science roles and their importance
* Discussion on the growing significance of data in various industries
* Trends and potential future developments in Data Science roles
* Conclusion