Data Stewardship and Quality

Course Content

Course Description:

This course focuses on data engineering, data quality assessment, and data stewardship, providing participants with an in-depth understanding of how to manage and maintain high-quality data in their organization. Participants will learn how to design and implement data engineering pipelines to ingest, transform, and load data from various sources, as well as how to assess and improve data quality using a range of methods and techniques.

The course covers data stewardship and governance, including establishing data governance policies and procedures, defining roles and responsibilities for data stewardship, and integrating data quality and stewardship into business processes and data management systems. Participants will also learn how to measure the success of data quality and stewardship initiatives, and how to manage change and overcome resistance to change in these initiatives.

Upon completion of the course, participants will have a solid understanding of data engineering, data quality assessment and improvement, and data stewardship and governance, and will be able to design and implement effective data engineering pipelines and manage high-quality data in their organization.

(This is the 3rd course of the Analytics Bootcamp)

Course Objectives:

Upon completion of this course, participants will be able to:

- Design and implement data engineering pipelines to ingest, transform, and load data from various sources
- Assess data quality and identify areas for improvement
- Implement methods for improving data quality
- Establish effective data stewardship and governance practices
- Integrate data quality and stewardship into business processes and data management systems
- Address data quality and stewardship in project management
- Manage change in data quality and stewardship initiatives
- Measure the success of data quality and stewardship initiatives

Data Stewardship and Quality

Prerequisites:

This course is designed for data professionals, including data engineers, data analysts, data scientists, and other data management professionals who are responsible for managing and maintaining high-quality data in their organization. Participants should have a good understanding of core data concepts and basic data management principles.

Target Audience:

The audience for this course is data professionals who want to learn how to design and implement effective data engineering pipelines and manage high-quality data in their organization. This includes data engineers, data analysts, data scientists, and other data management professionals who are responsible for managing and maintaining high-quality data in their organization. Participants should also be interested in data stewardship and governance, and integrating data quality and stewardship into business processes and data management systems.

Topics:

Each module will have hands-on activities, real-world examples, and assessments to reinforce learning and ensure understanding.

Module 1: Introduction to Data Engineering and Data Quality
- Understanding data engineering and its importance
- Overview of data quality and its impact on data engineering
- Introduction to data engineering pipelines

Module 2: Data Ingestion and Transformation
- Methods for ingesting and transforming data (in tools like snowflake, Kafka, etc.)
- Techniques for cleaning and preparing data (using techniques in SQL, Python, etc.)
- Managing data pipelines and workflow orchestration

Module 3: Data Quality Assessment and Improvement
- Techniques for assessing data quality
- Methods for improving data quality (in mesh framework and self-service tools like dataiku)
- Establishing data quality metrics (in common data catalogs)

Module 4: Data Stewardship and Governance
- Overview of data stewardship and governance
- Defining roles and responsibilities for data stewardship
- Establishing data governance policies and procedures

Module 5: Integrating Data Quality and Stewardship into Business Processes
- Integrating data quality and stewardship into business processes
- Implementing data quality and stewardship in data management systems
- Addressing data quality and stewardship in project management

Module 6: Managing Change and Measuring Success
- Overcoming resistance to change in data quality and stewardship initiatives
- Establishing a framework for measuring the success of data quality and stewardship initiatives
- Continuous improvement for data quality and stewardship processes

Register for this class by visiting us at:
www.tcworkshop.com or calling us at 800-639-3535