
Data Ops – One Day Course Agenda

Overview

Data Quality and Best Practice Controls in ETL processes – Data quality plagues most organisations which can in turn lead to poor decision making. This course focuses on data quality and best practice for ETL processes, with cloud technologies making it easier to access data, while efficiently and effectively ingesting and preparing data from multiple sources. We will deep dive into how to set up a strong culture of data quality and reliability through setting your processes up for success, making continuous improvements and building in automated testing.

Target Audience

This course is aimed at participants involved in the facilitation of data to consumers, looking to implement data quality controls across your organisation.

Learning Outcomes

At the end of this course delegates should be able to:

- Understand terminology and roles pertinent to data platform teams
 - Understand the fundamental relevance and mission of data platform teams
 - Identify common pitfalls and problems befalling data platform teams
 - Have an arsenal of CI/CD strategies and controls to enable consistent delivery at high quality and avoid arduous manual processes
 - Apply a framework to handle issues with data accuracy and quality
-

#	Topic
1.	Background Concepts <ul style="list-style-type: none">• Data ecosystem• ETL fundamentals• Flow of a Modern Data Platform• People of a Modern Data Platform Team
2.	The Seams – Where it can fail <ul style="list-style-type: none">• Incomplete Requirements Gathering• Slow Data Integration• Poor Data Quality• Improper Data Modelling*• Taking Shortcuts on Testing• Failure to aggressively automate *Dimensional Modelling is an entire course on its own
3.	Opening the Bottlenecks <ul style="list-style-type: none">• Requirements Gathering Strategies• Data Platform Code• Team Management• Team Empowerment• DI Checklist
4.	Closing the Gaps <ul style="list-style-type: none">• Start with a test• Creating your pipeline framework• Striving for relevant Data Model Design*• Data Controls• Aggressive Automation *Dimensional Modelling is an entire course on its own
5.	Data Ops Scenarios and Discussion
6.	Summary and Close