

# Data Pipeline Testing

# Learning Objectives

- ▶ Data Pipeline tests categories
- ▶ Standard tests types

# Data Pipeline Testing

- ▶ Data quality tests: test the quality of the data
  - ▶ Apply check constraints to Delta tables
- ▶ Standard tests: test the code logic
  - ▶ Unit testing
  - ▶ Integration testing
  - ▶ End-to-end testing

# Unit testing

- ▶ Approach to testing individual units of code, such as functions.
- ▶ If you make any changes to them in the future, you can determine whether they still work as you expect them to.
- ▶ This helps you find problems with your code faster and earlier in the development life cycle.

# Unit Test Assertions

- ▶ An assertion is a statement that enables you to test the assumptions you have made in your code
- ▶ **assert** func() == expected\_value

# Integration Testing

- ▶ Approach to testing the interaction between subsystems of an application
- ▶ Software modules are integrated logically and tested as a group.

# End-to-End Testing

- ▶ Approach to ensure that your application can run properly under real-world scenarios
- ▶ Simulate a user experience from start to finish