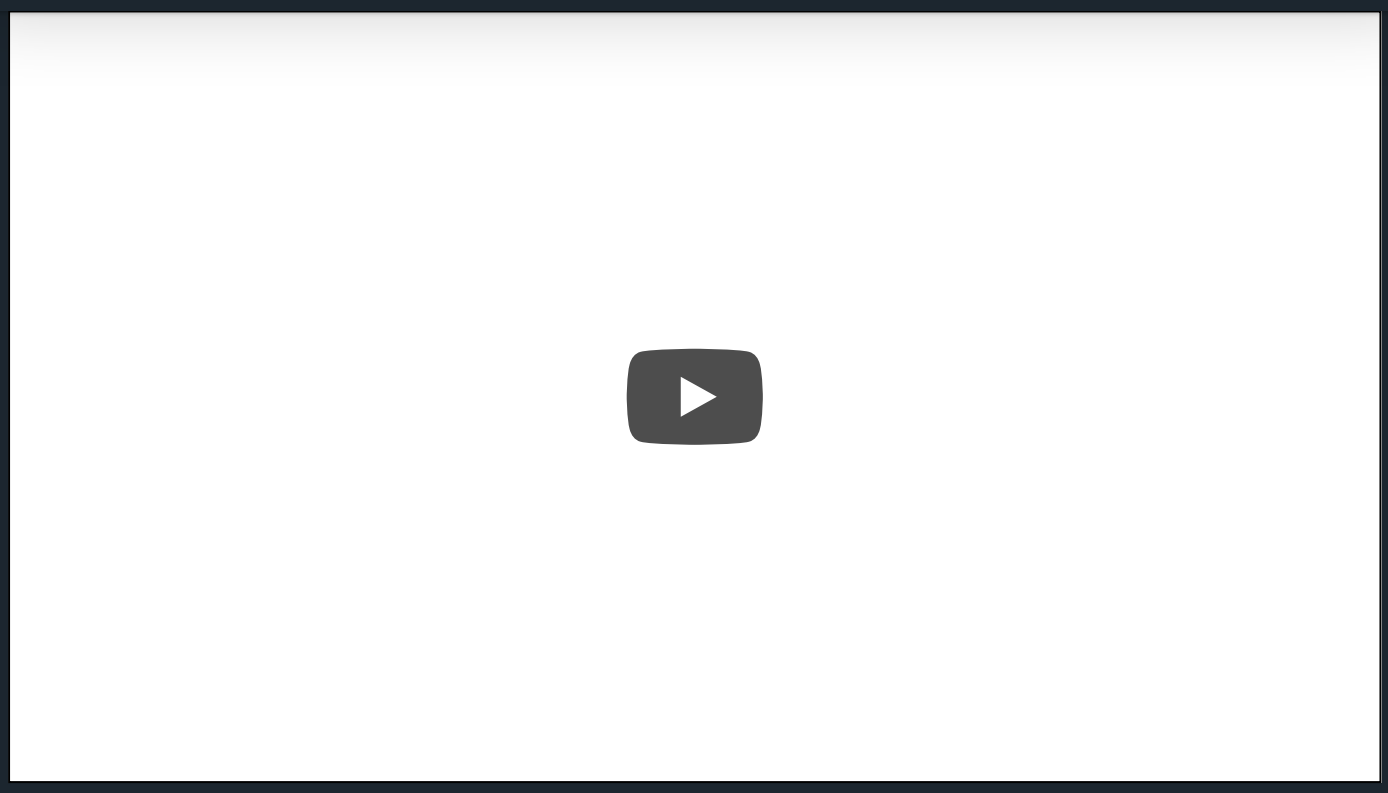

Testing And Data Science

- Problems that could occur in data science aren't always easily detectable; you might have values being encoded incorrectly, features being used inappropriately, unexpected data breaking assumptions
- To catch these errors, you have to check for the quality and accuracy of your *analysis* in addition to the quality of your *code*. Proper testing is necessary to avoid unexpected surprises and have confidence in your results.
- **TEST DRIVEN DEVELOPMENT:** a development process where you write tests for tasks before you even write the code to implement those tasks.
- **UNIT TEST:** a type of test that covers a "unit" of code, usually a single function, independently from the rest of the program.

Resources:

- Four Ways Data Science Goes Wrong and How Test Driven Data Analysis Can Help: Blog Post
- Ned Batchelder: Getting Started Testing: Slide Deck and Presentation Video

NEXT