

Data Engineering Career Track

Open-ended Capstone step 8: Deploy Your Code for Testing

Estimated Time: 15-20 Hours

In this step, you'll deploy your code to Azure compute resources for testing, and make any necessary modifications to your code in order for it to run smoothly. Consider this step as analogous to deploying your code to a testing environment. You'll deploy your code, build a robust test suit, check the edge cases which may break your code, and see if it is performing at the standard you expect it to. In Step Nine, you'll deploy your finalized code to a production environment.

If you have not designed unit tests for your code, do so now. All code used in the data pipeline should be submitted with unit tests, as per the requirements of this project. If you need a refresher on unit testing, please refer to subunit 4.5.

At the end of this step, make sure that the code that you have deployed passes all unit tests and can process your dataset utilizing Azure's parallel computing capabilities.

Deliverables:

- Update the project's GitHub repo with the work you complete for this step of the project and update the readme markdown file
- Report of unit test outlining the following metrics:
 - Number of Passed/Failed tests
 - Code coverage percentage

Update Your Slide Deck:

Add your work from this step to your slide deck, including:

- An overview of your test suite, and any edge cases you wrote tests to cover
- Details about any changes you had to make to your code as a result of your testing or deployment
- A snapshot of your test results
- If applicable: any changes to your deployment architecture as a result of your testing deployment