



## Data Engineering Career Track

### Open-ended Capstone Step Nine: Deploy Production Code & Process Dataset

---

#### Estimated Time: 8-10 Hours

In this step of your Capstone, you'll **deploy the code you optimized in the last step to a fresh set of Azure compute resources**. The resources you deploy your code to should match the architecture you mapped in Step Seven. Make sure that you create appropriately sized resources, be careful not to make them too small or too large **as that can easily lead to massive charges from Azure. Make sure that you delete or terminate temporary resources, which your pipeline does not continuously require to function, as a cost-saving strategy.**

Your deployed code should process your entire dataset, which can be stored either in an Azure Storage Solution or streamed via API calls, and should write the processed data to an Azure Storage Solution as specified in your deployment architecture.

**Deliverables:**

- Update the project's GitHub repo with the work you completed in this step and update the readme file.

**Slide deck updates:**

From this step, you should include:

- An overview or flow chart demonstrating how you migrated from test code to the final code for this step
- An overview of your finalized production deployment architecture and code