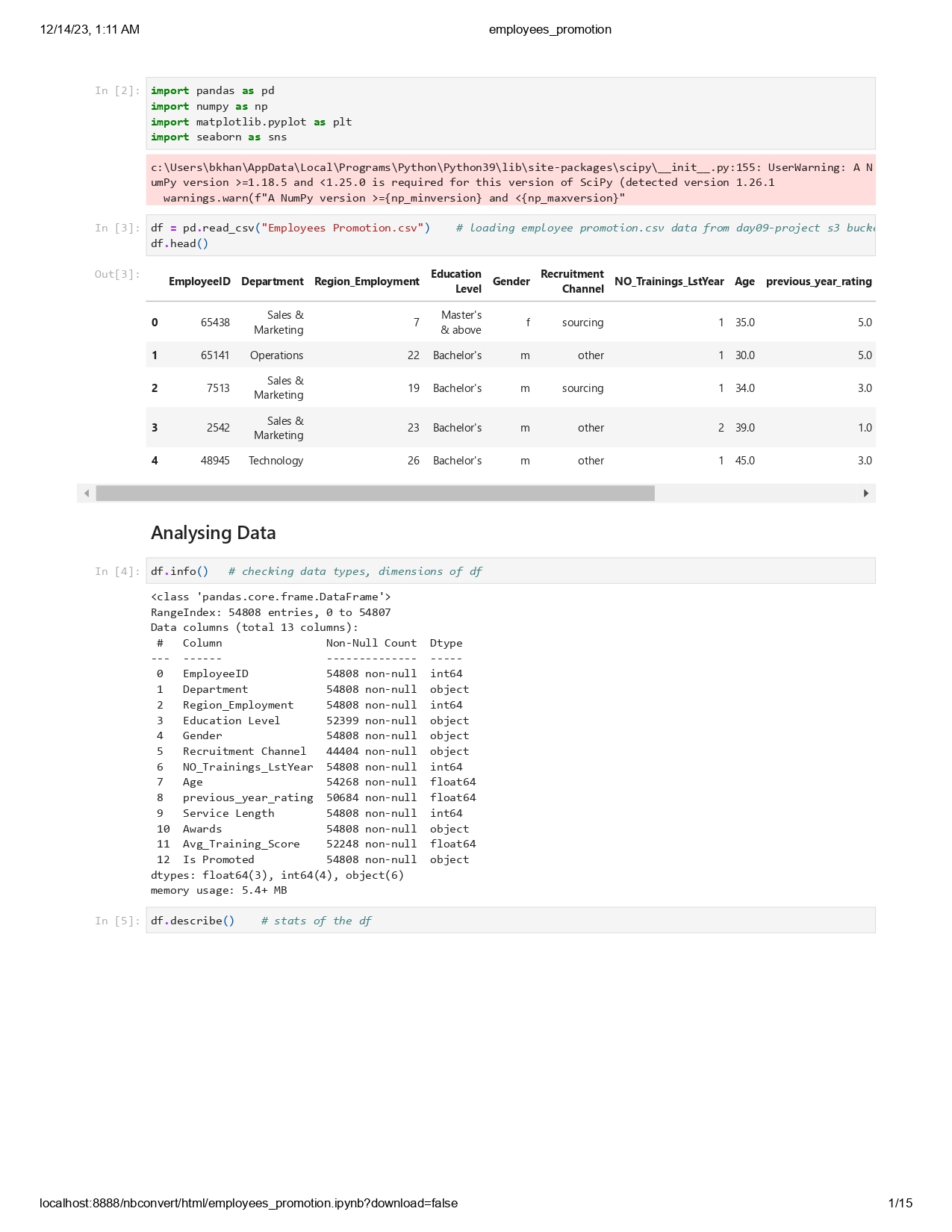
Artificial Intelligence in Cloud Computing

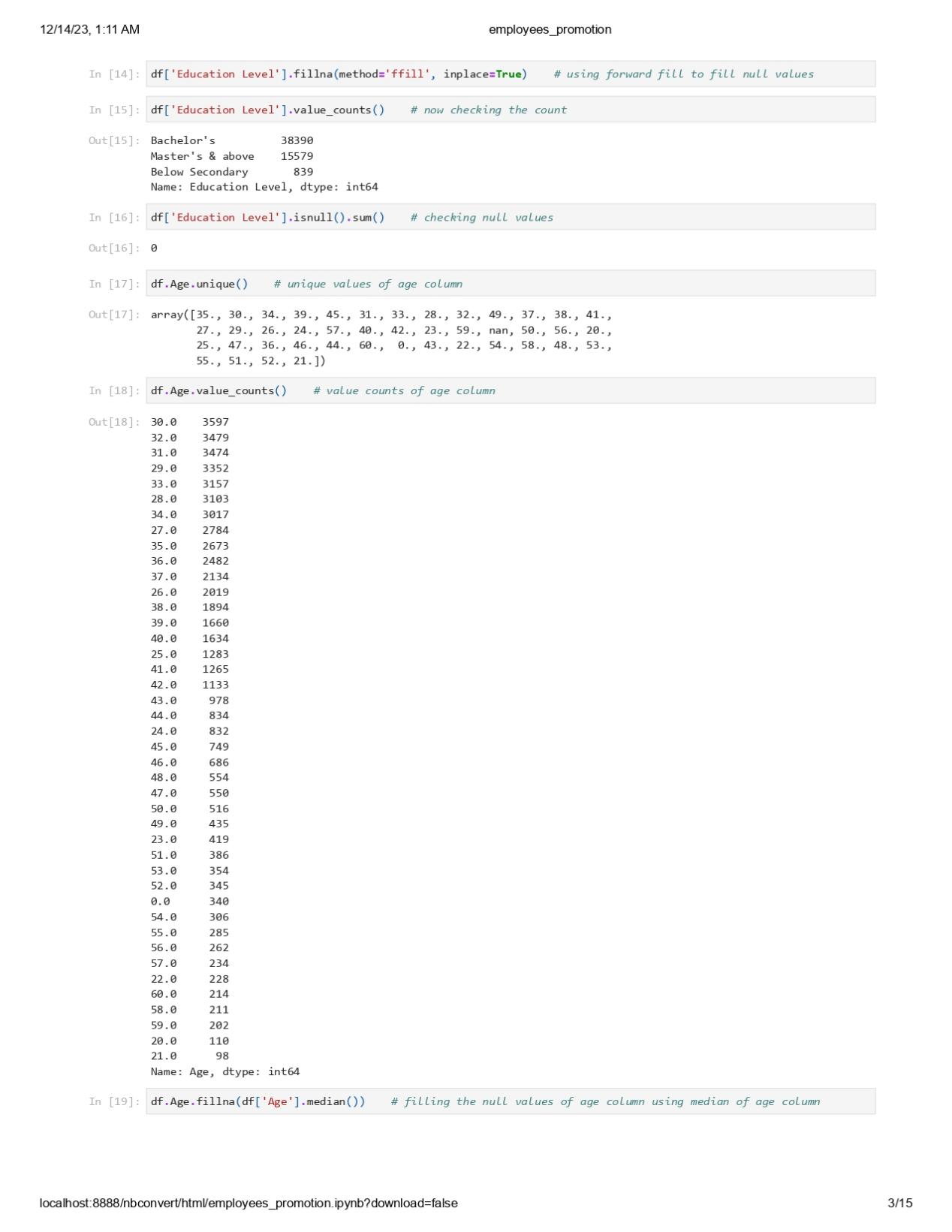
Final Project

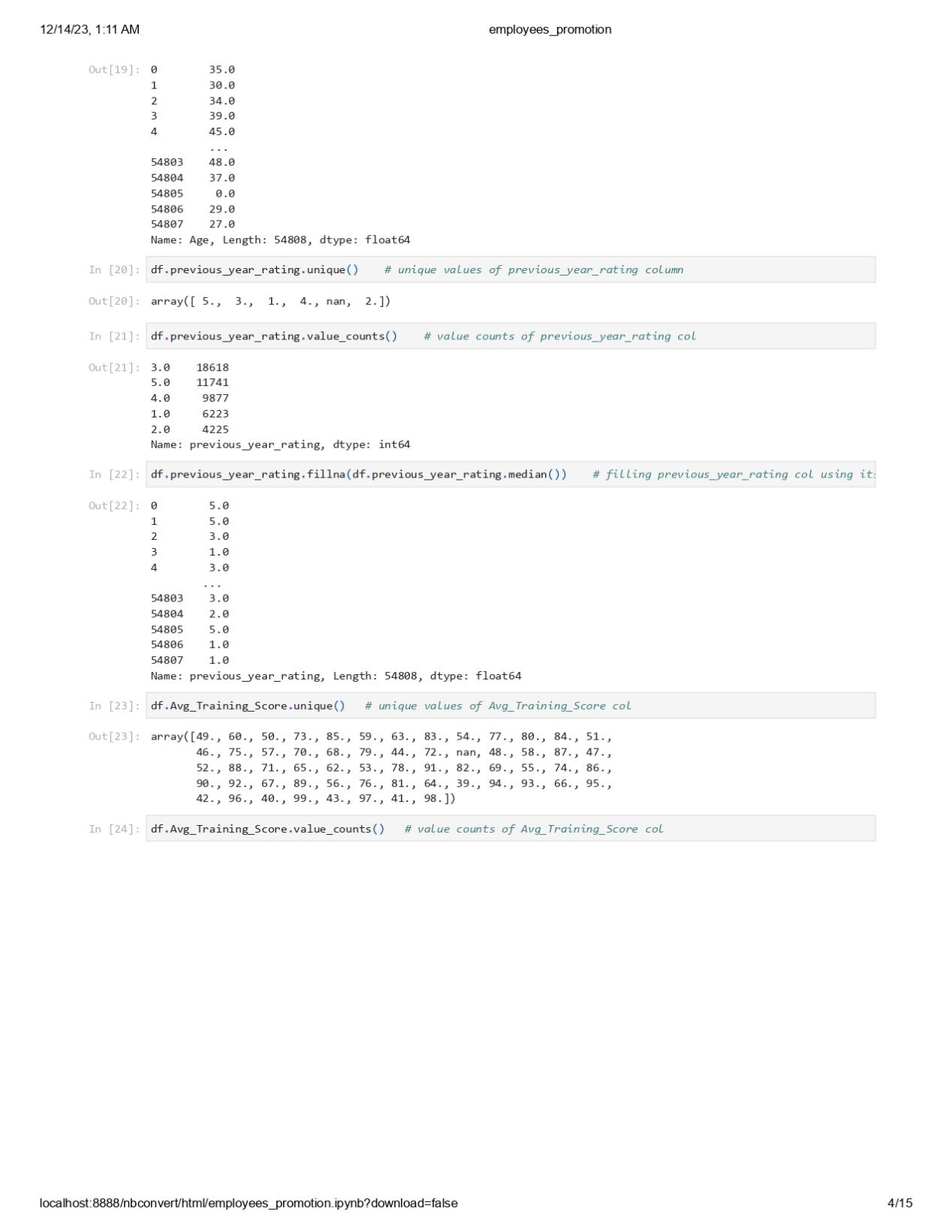
Name: Mohammed Bilal Shakil Ahmed Khan

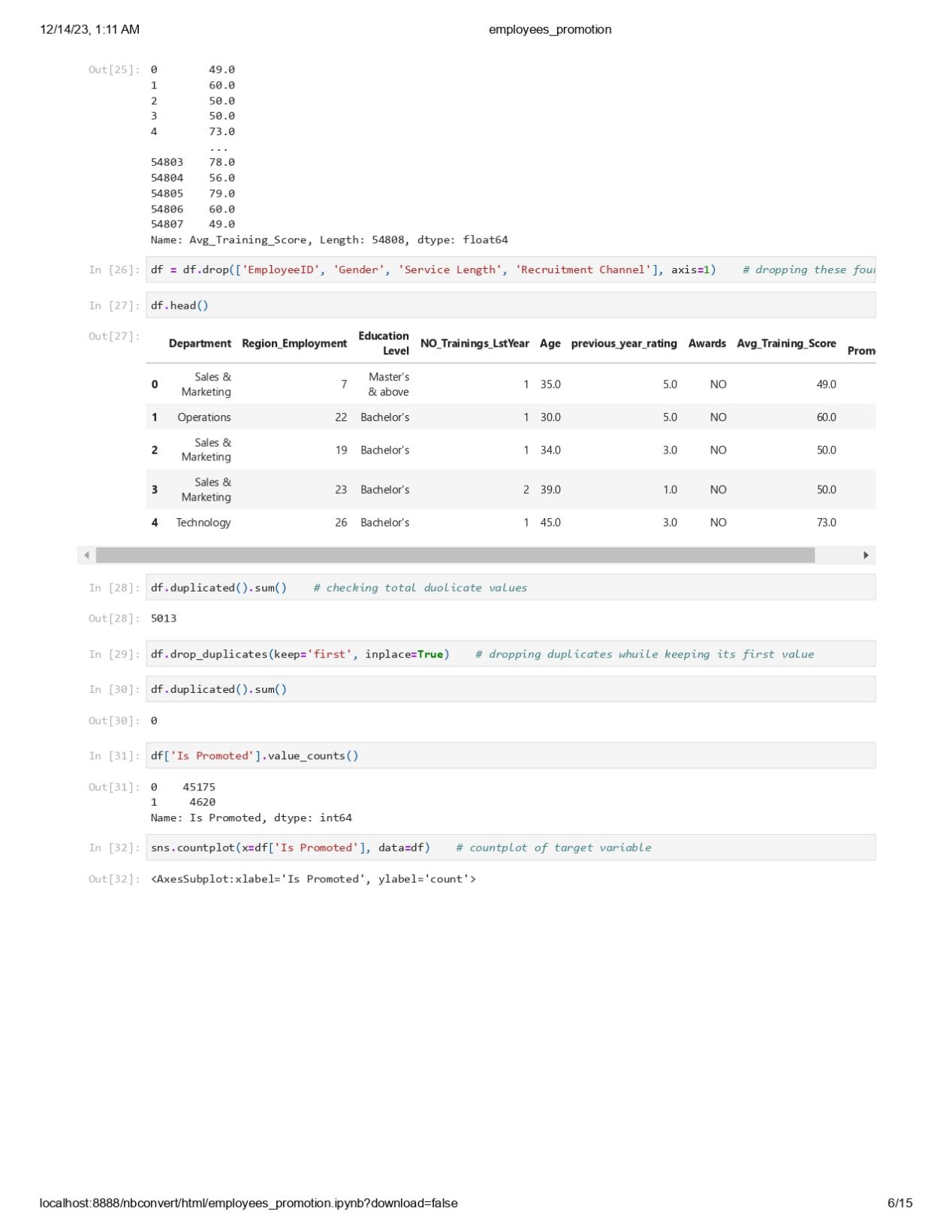
ID: N01637402























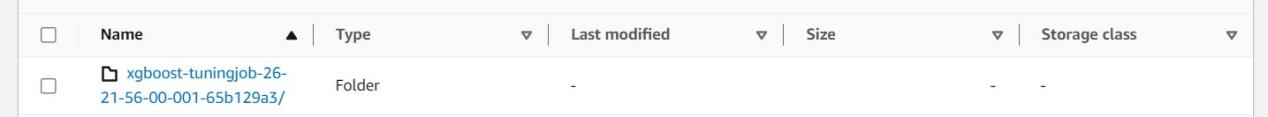


Now, we will perform Inference endpoint using-canary traffic shifting.

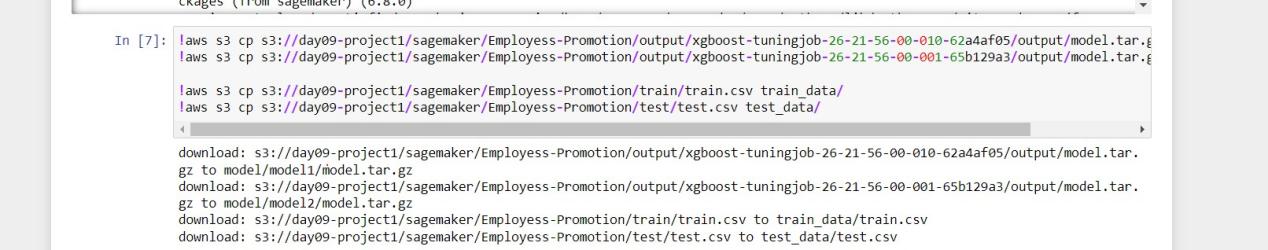
1. Select two models:

xgboost-tuningjob-26-21-56-00-010-62a4af05 is my best model.

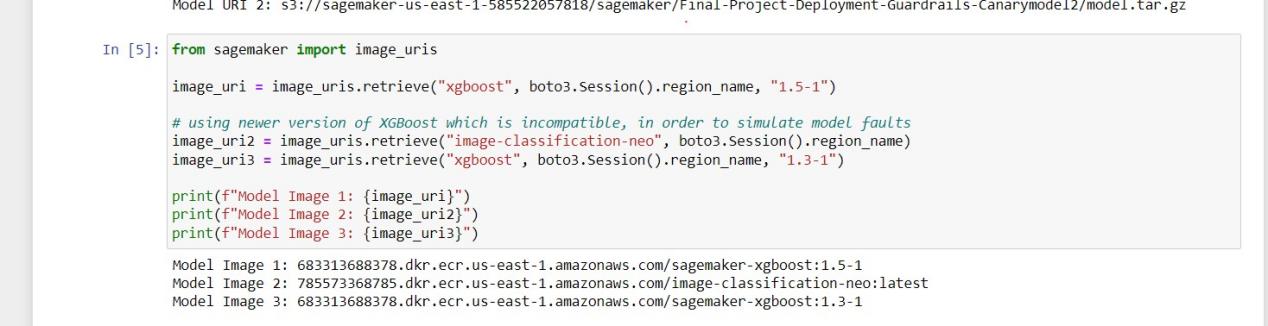
xgboost-tuningjob-26-21-56-00-001-65b129a3 is my second model.



1. Copy those two models and test.csv file to current bucket.



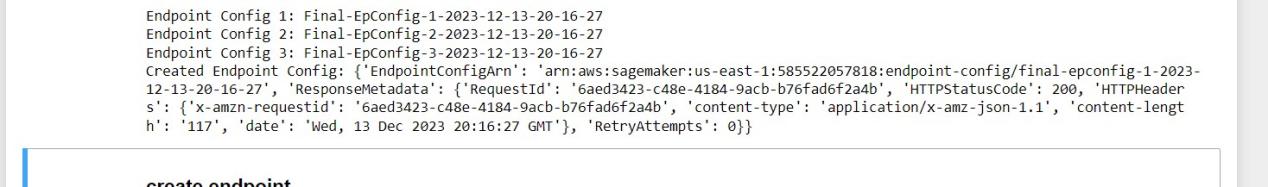
1. Using xgboost version 1.5-1 for the best, image-classification-neo for the second model, and xgboost version 1.3-1 for the third model. I tried to use linear learner and xgboost for the second model but it didn’t failed when I was invoking the endpoint.



1. Creating three models.

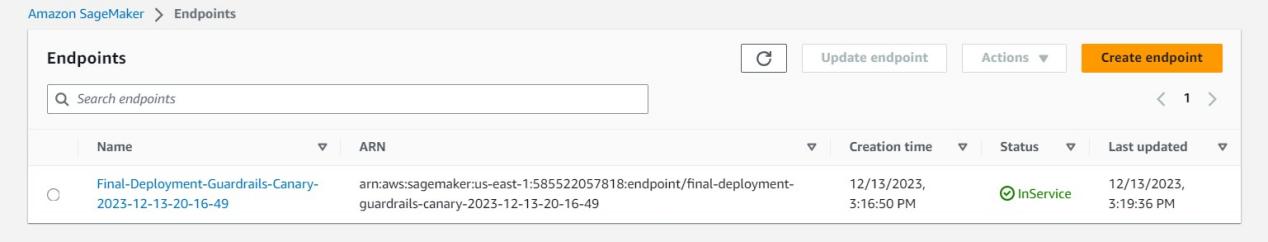


1. Creating three endpoint configuration.

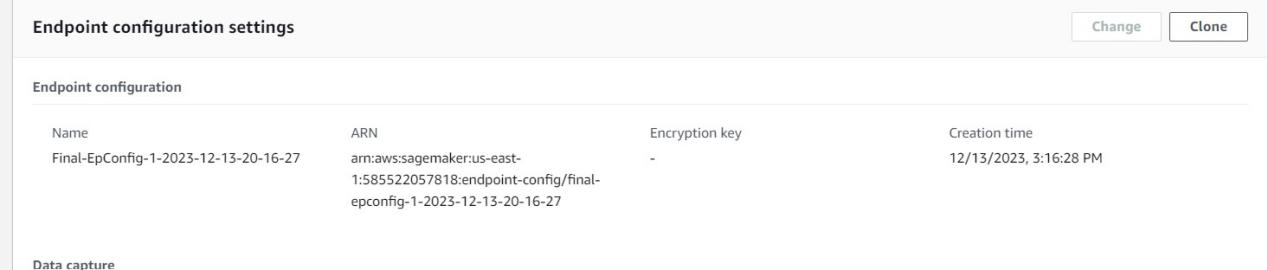


1. Creating an endpoint





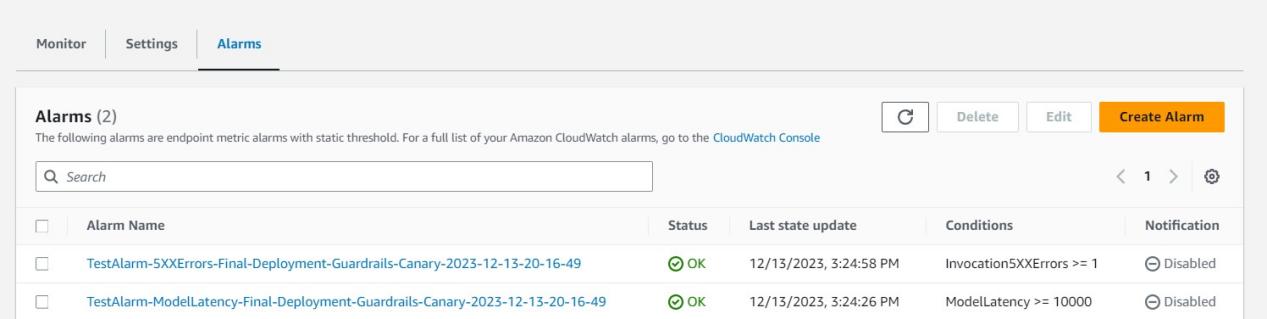
1. This endpoint is configured to endpoint configuration1 (‘Final-EPConfig-1’).



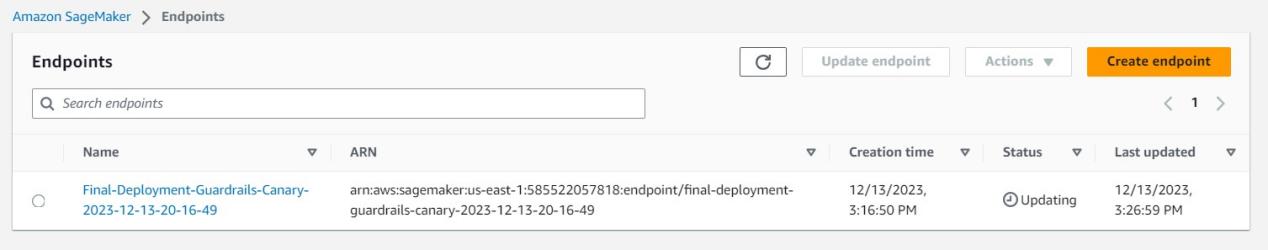
1. Sending traffic to endpoint which has endpoint configuration1 (‘Final-EPConfig-1’).



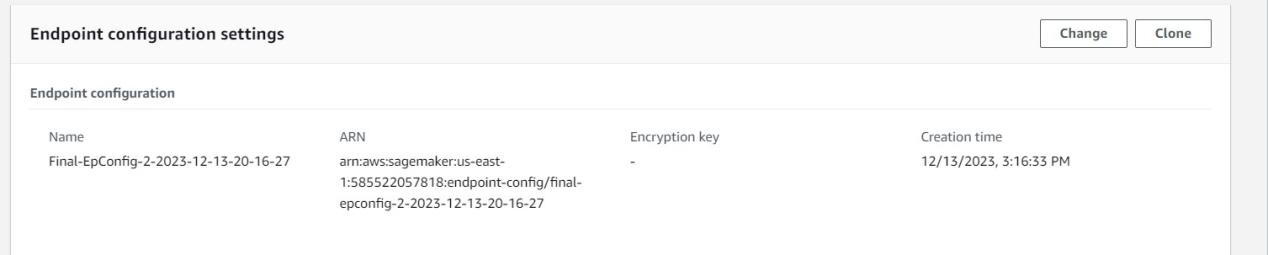
1. Creating alarms.



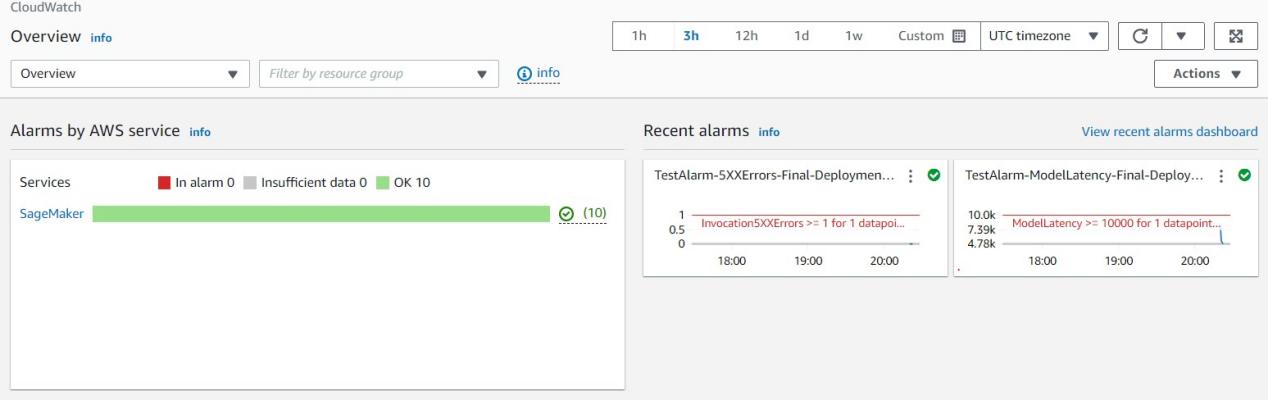
1. Creating rollback and updating the endpoint to set the endpoint configuration to ‘Final-EPConfig-2’.



1. Endpoint set to ‘Final-EPConfig-2’ after updating.



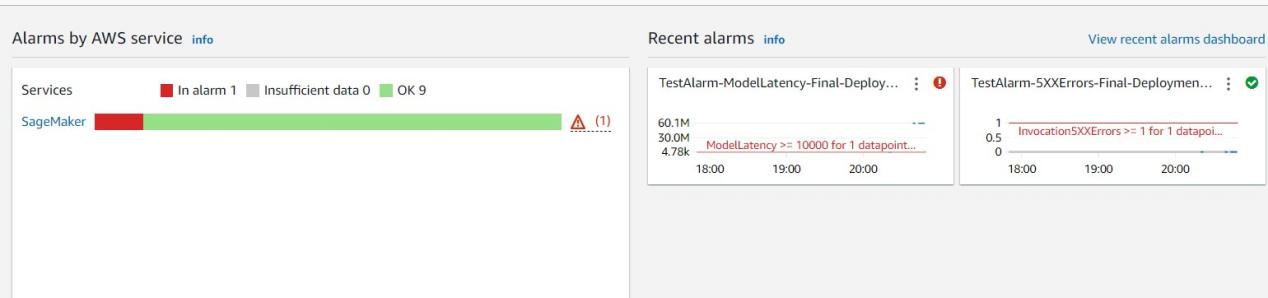
1. Checking alarm before invoking.



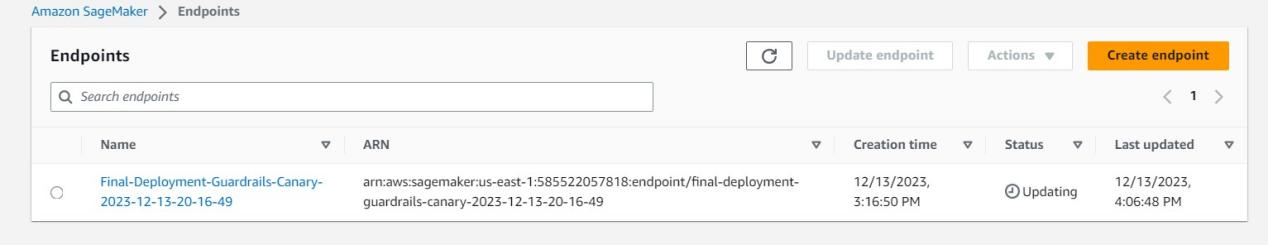
1. Invoking endpoint for second configuration which should fail and perform rollback.

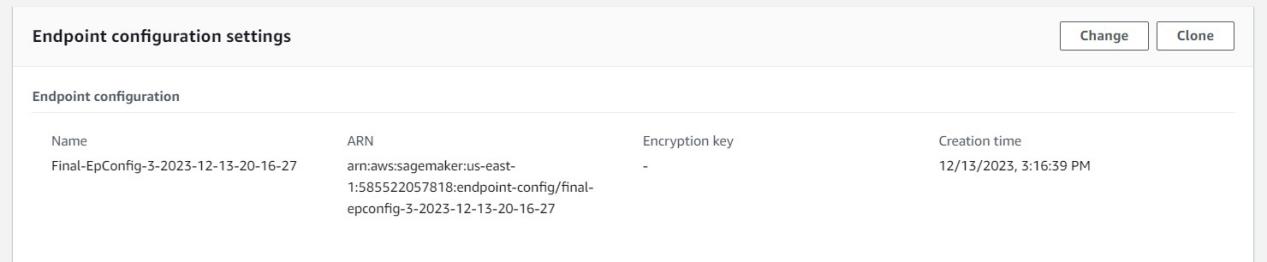


1. Checking alarm in the cloudwatch for the errors.



1. Updating endpoint to ‘Final-EPConfig-3’

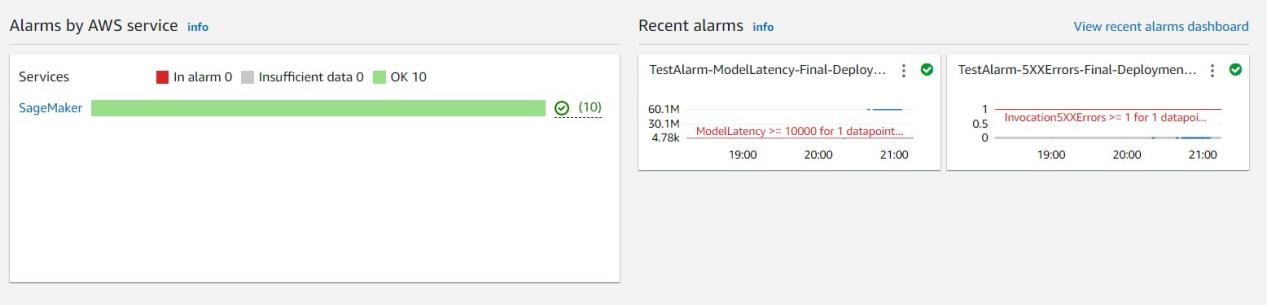




1. Invoking endpoint for the third configuration.



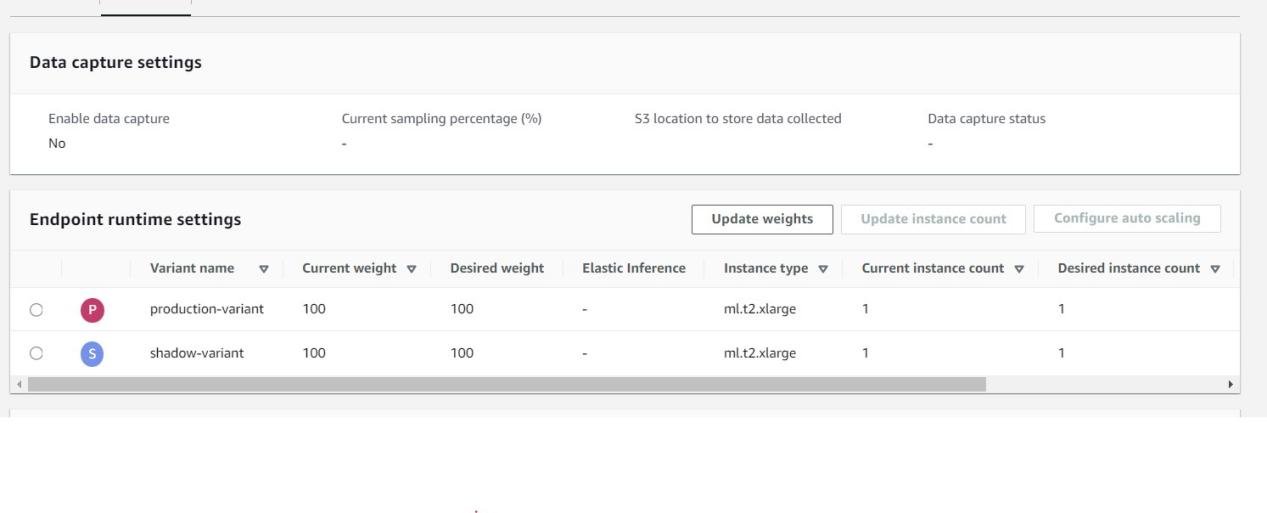
1. Checking cloudwatch again for the alarms.



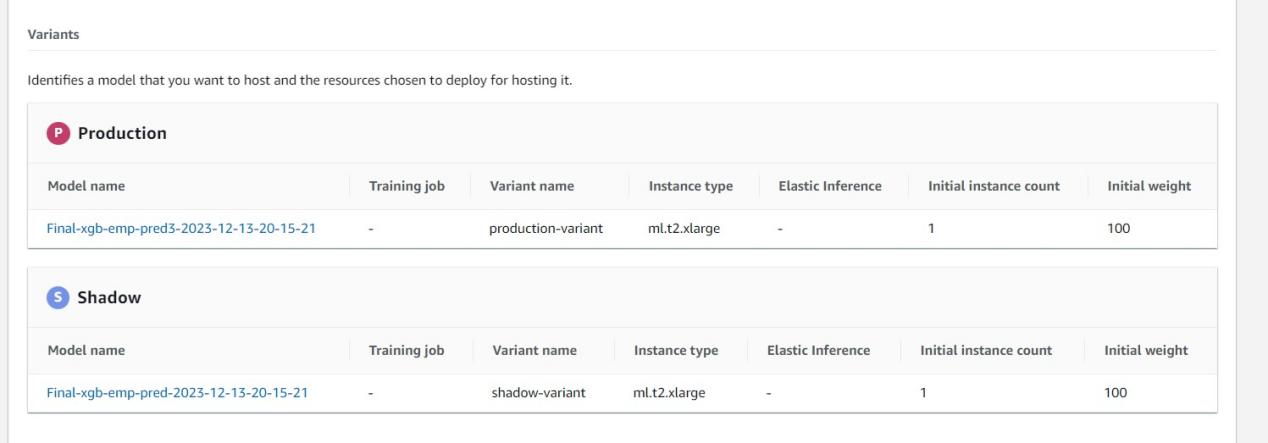
1. Creating Shadow test.



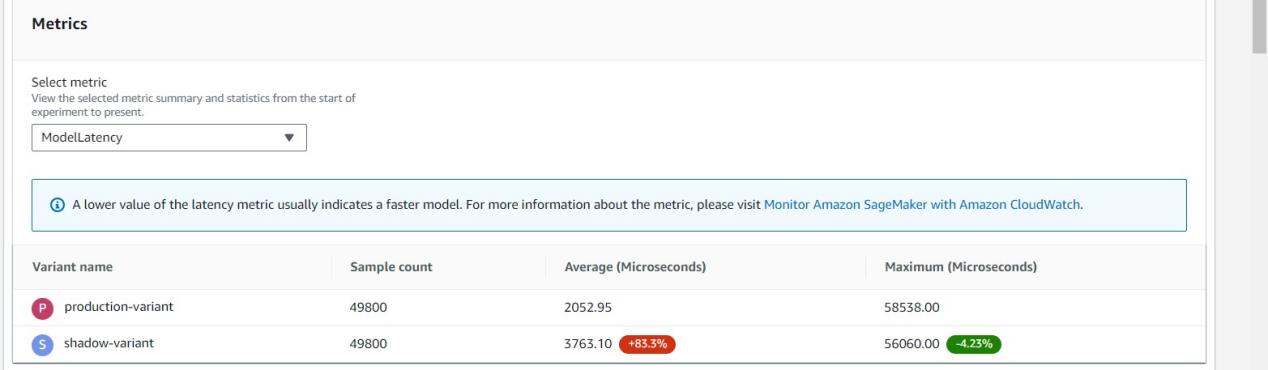
1. Endpoint runtime settings.



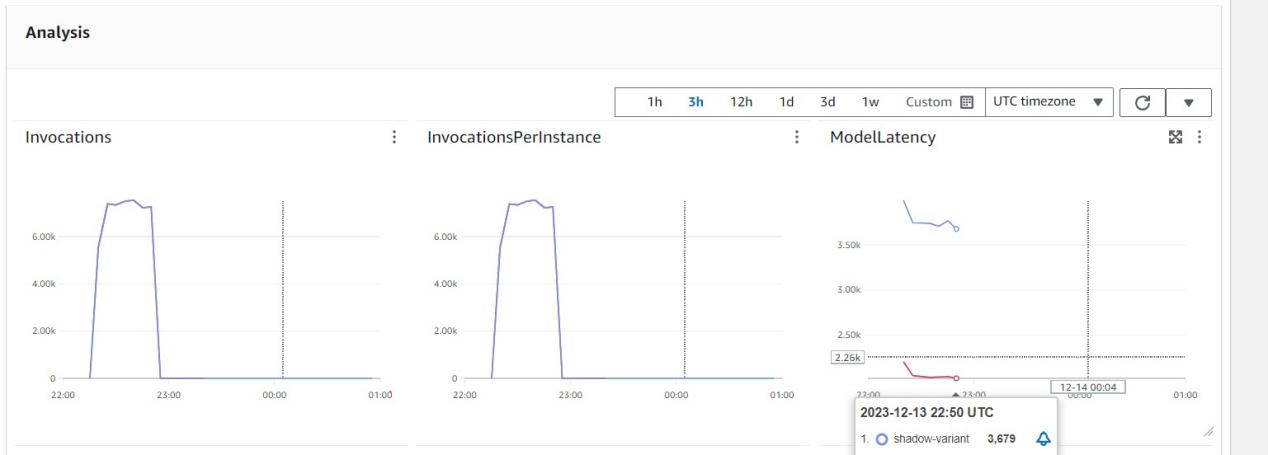
1. My production and shadow variant models. Shadow variant has my best model whereas production variant has the second best.



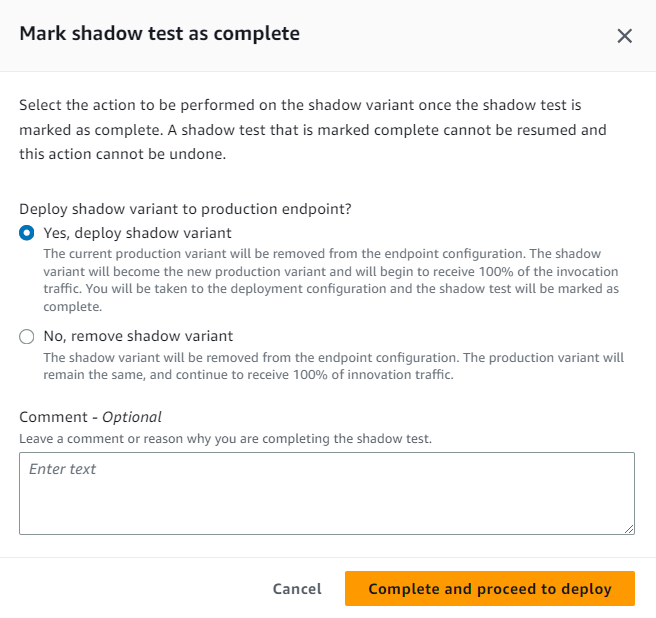
1. Production and shadow variant metrics.



1. Shadow Test Analysis.



1. Deploying shadow test.



1. Deployed.

