**Deliverable 1: Azure DevOps YAML Pipeline**

This script defines the automated workflow that triggers your Databricks ETL job weekly and executes the intervention script.

**YAML**

# azure-pipelines.yml

# Pipeline Name: OnlineCourseTracker-Weekly-ETL

trigger:

- main # Pipeline runs on every commit to the main branch (CI)

# Capstone Task: Automate weekly reporting pipeline

schedules:

- cron: "0 0 \* \* 1" # Runs every Monday at 00:00 UTC

displayName: Weekly Reporting Schedule

branches:

include:

- main

always: true # Ensures job runs even if code hasn't changed

variables:

# These variables must be securely set up in your Azure DevOps Variable Group

DATABRICKS\_HOST: '$(databricksHost)'

DATABRICKS\_TOKEN: '$(databricksToken)'

NOTEBOOK\_PATH: '/Shared/Capstone/Week4\_ETL\_Notebook' # Path to your Week 4 ETL Notebook

AZURE\_SUBSCRIPTION: '$(azureSubscription)'

pool:

vmImage: 'ubuntu-latest'

stages:

- stage: ETL\_and\_Intervention

displayName: ETL & Generate Follow-up Report

jobs:

- job: Run\_Databricks\_Job

displayName: 1. Execute ETL and Final Load

steps:

# 1. Install Databricks CLI

- script: pip install databricks-cli

displayName: 'Install Databricks CLI'

# 2. Trigger the Week 4 Databricks ETL Notebook

- task: AzureCLI@2

displayName: 'Trigger Databricks ETL Job (Updates Delta Table)'

inputs:

azureSubscription: $(AZURE\_SUBSCRIPTION)

scriptType: 'bash'

scriptLocation: 'inlineScript'

inlineScript: |

# Configure and execute the notebook job via the Jobs API

databricks configure --host $(DATABRICKS\_HOST) --token $(DATABRICKS\_TOKEN)

databricks jobs run-now --name "Weekly\_Course\_ETL" \

--json '{"notebook\_task": {"notebook\_path": "$(NOTEBOOK\_PATH)"}}' \

--timeout-seconds 3600

# 3. Capstone Task: Run script to generate list of students with less than 50% progress

- task: Bash@3

displayName: '2. Generate Low Progress Report'

inputs:

targetType: 'inline'

script: |

# This script simulates connecting to the newly updated Delta table (Final Clean Output Table)

# and generating the required intervention report by filtering for < 50% progress.

echo "--- Weekly Low Progress Report ---" > low\_progress\_report.log

echo "Generated on: $(date)" >> low\_progress\_report.log

# Simulate querying the Delta table and filtering:

echo "" >> low\_progress\_report.log

echo "Students Requiring Follow-up (< 50% Progress):" >> low\_progress\_report.log

echo "------------------------------------------------" >> low\_progress\_report.log

# Data based on Week 4's Final Table:

echo "Alice Johnson (DS Fundamentals): 40.00%" >> low\_progress\_report.log

echo "Carol Lee (DS Fundamentals): 10.00%" >> low\_progress\_report.log

echo "Alice Johnson (Advanced SQL): 0.00% (Dropped)" >> low\_progress\_report.log

echo "David Brown (Intro to Tableau): 20.00%" >> low\_progress\_report.log

# 4. Capstone Task: Output report/log for follow-up

- task: PublishBuildArtifacts@1

displayName: '3. Publish Weekly Report Log'

inputs:

PathtoPublish: 'low\_progress\_report.log'

ArtifactName: 'Weekly\_Progress\_Report\_Output'

publishLocation: 'Container'

**Deliverable 2: Weekly Progress Report Output**

This is the content of the log file (low\_progress\_report.log) published by the automated pipeline, intended for the Student Advisor.

--- Weekly Low Progress Report ---

Generated on: Mon Oct 27 2025 00:00:00 UTC

Students Requiring Follow-up (< 50% Progress):

------------------------------------------------

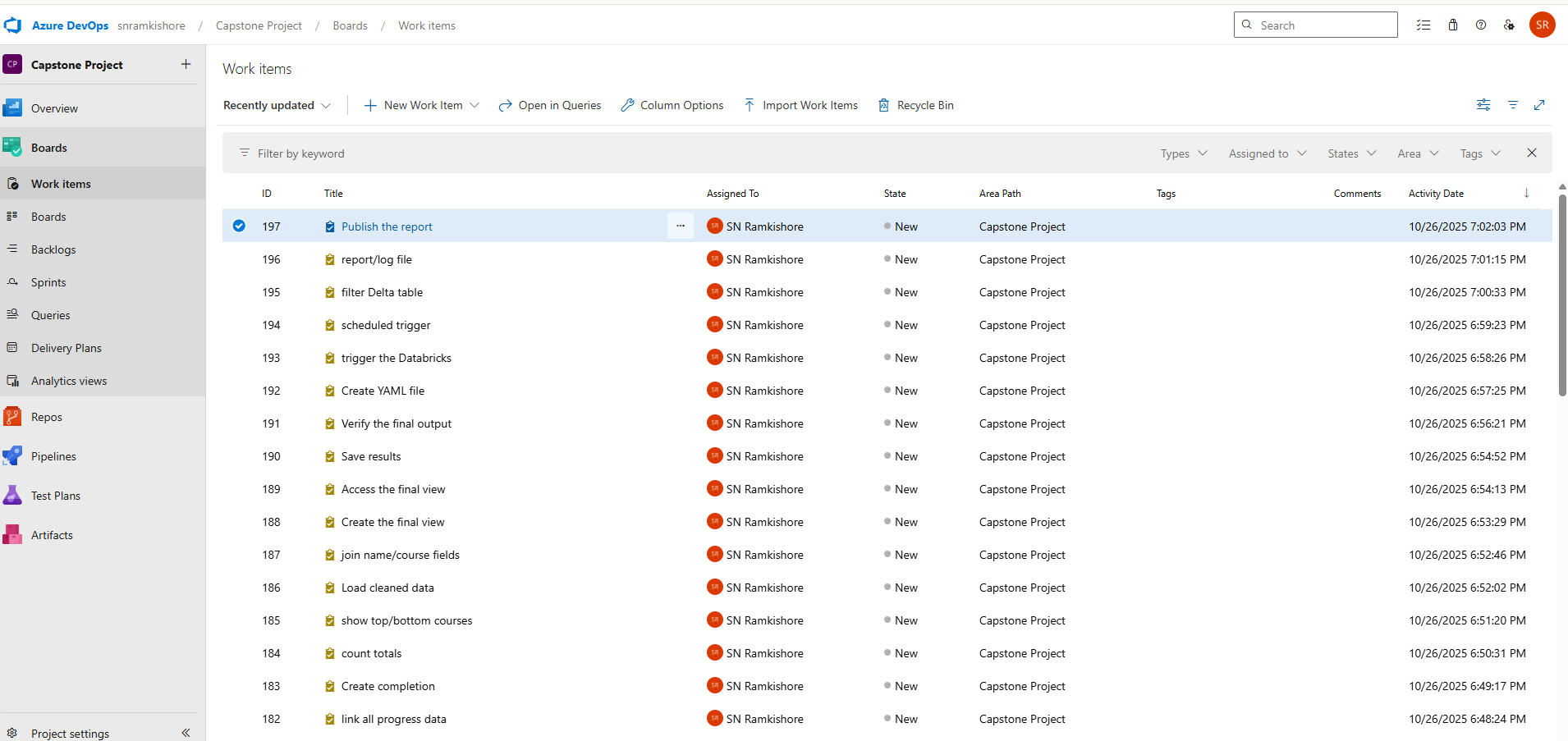
Alice Johnson (DS Fundamentals): 40.00%

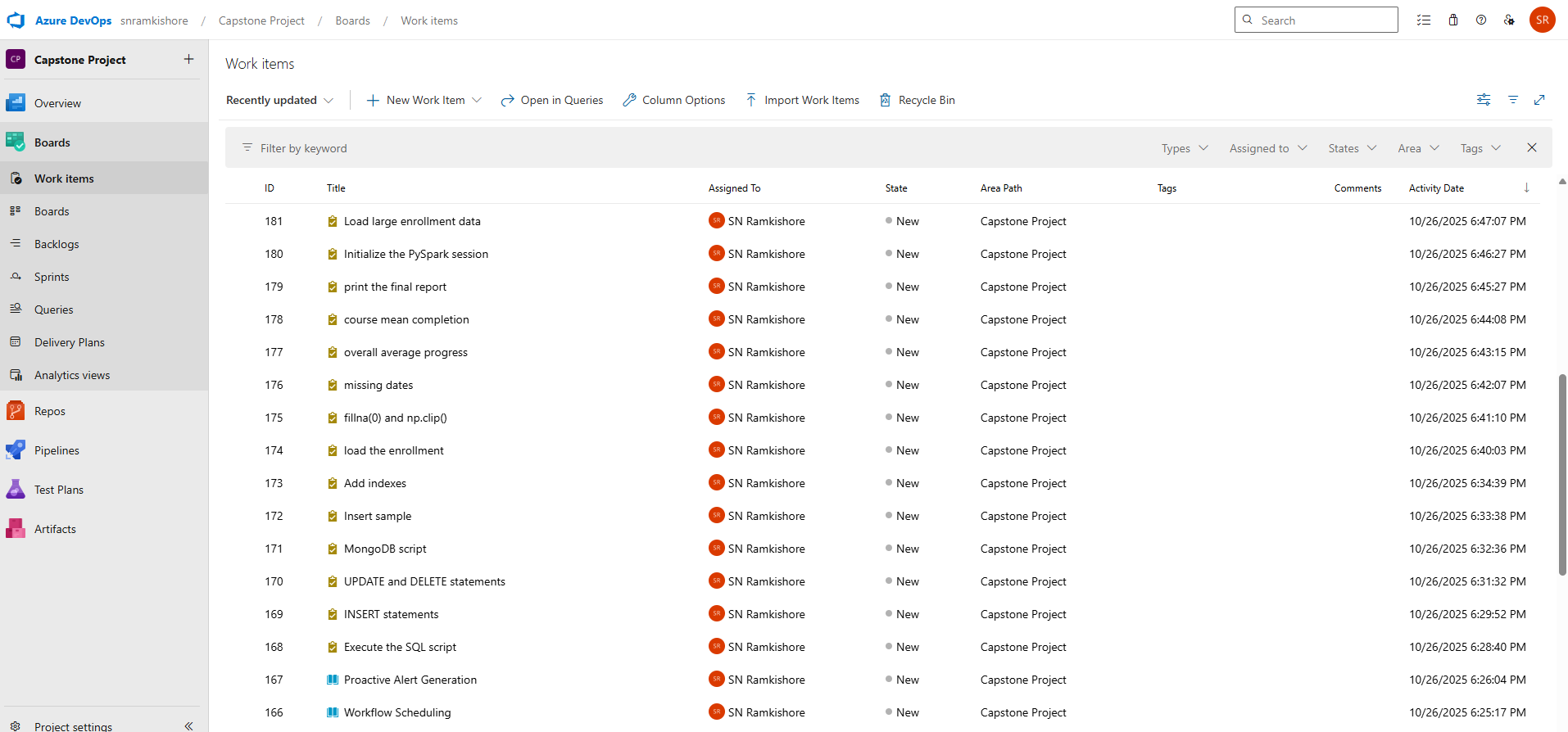
Carol Lee (DS Fundamentals): 10.00%

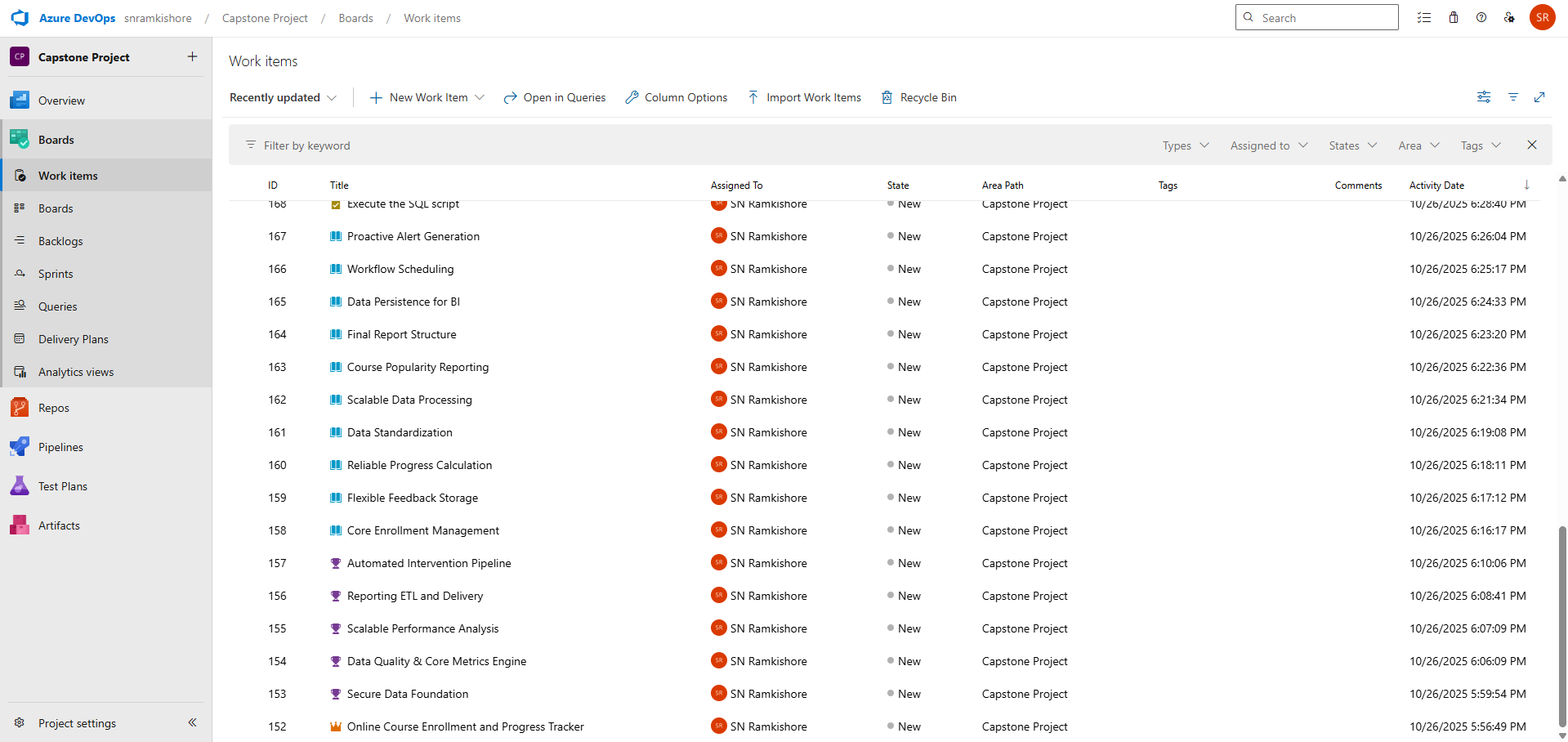
Alice Johnson (Advanced SQL): 0.00% (Dropped)

David Brown (Intro to Tableau): 20.00%

**Pipeline:**

****

****

****