Logo, company name

Description automatically generated

**Ganit Capstone Project 3**

**Instructions:**

Welcome to Ganit capstone project!

We hope you find this challenge interesting. We won't provide any kind of starter code, so feel free to configure your own environment to achieve the proposed goals.

I have created a Data folder which consists of multiple datasets related to movies

Objective of Capstone Project

* Store all these data files in an AWS S3 bucket. Do the entire data profiling to understand the trends in data. Build QC checklist based on the data profiling done. Create data catalogs preferably using glue
* Build custom data processing scripts to clean, process & aggregate data. The scripts should have failure alerts and should also provide a summary report over email
* The processed can be loaded in S3 or moved to a data warehouse. The processed data should be able to be directly consumed by the visualization tool
* Build a live dashboard to show the various trends. You can create any visualizations that might be able to actor/character/platform comparison

The entire infrastructure should be developed on AWS. You are free to use any services on AWS. You can also use any visualization tool of your choice

Kindly share the git repository with us after completion of the challenge. The git repositories should also capture screenshots of successful outcomes, dashboards & documentation.

Good luck, and happy coding!