Automation and Cloud Deployment

Git Repos URL
https://github.com/adajami/rainfall

To automate and deploy this analysis in a cloud environment (AWS, Google Cloud, Azure), we need to use the following:

Tools and Technologies:
-Python for scripting and data processing.
-GDAL (Geospatial Data Abstraction Library) for raster and vector operations.
-Pandas for data manipulation and aggregation.
-GeoPandas for geospatial data handling.
-STAC API for accessing and querying raster data.
-Cloud Optimized GeoTIFFs (COGs) for efficient data storage and retrieval.

Automation Steps:
-Use Python scripts orchestrated by Airflow or AWS Step Functions for workflow management.
-Implement Lambda functions (AWS) for specific tasks like percentile calculation, raster/vector
-operations, and population exposure computation, triggered by events like new data uploads.

Containers:
-Use Docker containers to ensure reproducibility and ease of deployment.

Deployment:
-Utilize Cloud Storage (S3) for storing input and output data.
-Deploy scripts and functions as serverless applications for scalability and cost-effectiveness.
-Set up API endpoints for triggering and monitoring the automated workflows.
-Monitoring: Implement AWS CloudWatch for monitoring the pipeline and logging errors.