DIC2025 Assignment 3 - LocalStack implementation for Reviews Text Analysis

Overview

This repository contains the serverless review analysis application designed for LocalStack using lambda functions with triggers on specific S3 buckets creating a data processing pipeline

Requirements

Prerequisites

- **Python 3.12+** (3.12+ compatible)
- **Docker Desktop** (for LocalStack)
- **Git** (for bash scripts on Windows)

Platform-Specific Setup

Windows Setup

Environment Setup (Windows)

```
# 1. Navigate to project directory
cd C:\path\to\DIC_3

# 2. Create and activate virtual environment
python -m venv .venv
.venv\Scripts\Activate.ps1

# 3. Install dependencies
pip install -r requirements.txt

# 4. Start Docker Desktop (ensure it's running)
docker --version

# 6. For full AWS LocalStack testing (requires Git Bash):
& "C:\Program Files\Git\bin\bash.exe" ./setup_environment.sh
python package_lambdas.py
& "C:\Program Files\Git\bin\bash.exe" ./setup_aws_lambdas.sh
& "C:\Program Files\Git\bin\bash.exe" ./setup_triggers.sh
```

macOS Setup

Environment Setup (macOS)

```
# 1. Navigate to project directory
cd /path/to/DIC_3

# 2. Create and activate virtual environment
python3 -m venv .venv
source .venv/bin/activate

# 3. Install dependencies
pip install -r requirements.txt

# 4. Start Docker Desktop and verify
docker --version

# 6. Full AWS LocalStack automated setup
chmod +x setup_environment.sh
./setup_environment.sh
python package_lambdas.py
chmod +x setup_aws_lambdas.sh
./setup_aws_lambdas.sh
chmod +x setup_triggers.sh
./setup_triggers.sh
```

Linux Setup

Environment Setup (Linux)

```
# 1. Navigate to project directory
cd /path/to/DIC_3

# 2. Create and activate virtual environment
python3 -m venv .venv
source .venv/bin/activate

# 3. Install dependencies
pip install -r requirements.txt

# 4. Verify Docker
docker --version

# 6. Full AWS LocalStack automated setup
chmod +x setup_environment.sh setup_aws_lambdas.sh setup_eventbridge.sh
./setup_environment.sh
python package_lambdas.py
./setup_aws_lambdas.sh
./setup_triggers.sh
```

Testing & Analysis

Run Unit/Integration Tests

```
# Run all tests
./setup_environment.sh
pip install -r ./src/tests/requirements.txt
python -m nltk.downloader corpora punkt punkt_tab vader_lexicon stopwords wordnet
averaged_perceptron_tagger
python ./src/tests/test_preprocessing.py
python ./src/tests/test_profanity_check.py
python ./src/tests/test_sentiment_analysis.py
python ./src/tests/test_integration.py
```

Run AWS LocalStack Pipeline Test

```
# Test complete pipeline with real dataset
# needs setting up proper AWS credentials

export AWS_ACCESS_KEY_ID=test
export AWS_SECRET_ACCESS_KEY=test
export AWS_DEFAULT_REGION="us-east-1"
export AWS_ENDPOINT_URL="http://localhost:4566"

python ./upload_reviews.py

# to check the logs from localstack:
localstack logs -f
```

Create report

After the end of execution seen in localstack logs we can generate the report - assignment_results.json

```
# Generate the report based on the reviews in buckets and DynamoDB Table
python ./generate_report.py
```

Stop execution

```
# Clean up
localstack stop
```