ASSIGNMENT 02 // QUOTE ASSEMBLER

This project asks students to create a data pipeline that completes three tasks: (1) makes an HTTP POST request to an API endpoint to populate a series of messages in an SQS queue in AWS; (2) monitors that queue for the eventual delivery of all messages, then retrieves them, stores their data, and deletes them; (3) assembles a complete message based on the contents of the messages and submits the answer to a separate SQS queue. Students will write this pipeline using the Prefect framework, which will run locally on your laptop. Bonus points are available for students who also write the same pipeline in the form of an Airflow DAG. (i.e. both formats, not just one.)

QUESTIONS ARE DESIGNED TO EVALUATE STAGES OF DATA PROCESSING AND SKILL LEVELS FROM LOW TO HIGH. THESE MIMIC THE REQUIREMENTS A DATA ENGINEER IS PRESENTED IN INDUSTRY OR RESEARCH PROJECTS.

Student Name: _Rameez Rauf_

	YES	NO	TOTAL
GENERAL REQUIREMENTS			13 points
PROJECT SUBMISSION: ALL REQUIRED FILES (PYTHON CODE, SQS SUBMISSION.) ARE SUBMITTED BY THE DEADLINE.	✔	☐	2
CODE EXECUTION: THE PROVIDED CODE RUNS WITHOUT SIGNIFICANT ERRORS (E.G., SYNTAX ERRORS, OR UNHANDLED EXCEPTIONS THAT CRASH THE PROGRAM).	✔	☐	2
FORKED SUBMISSION: THE SUBMITTED REPOSITORY IS A FORK OF THE SOURCE.	✔	☐	1
COMMENTS: SUBMISSION INCLUDES INLINE COMMENTS FOR EACH TASK / CLASS / FUNCTION.	✔	☐	2
LANGUAGE: PROJECT IS WRITTEN IN PYTHON, SQL. OR BASH (WITHIN AIRFLOW DAG TASKS AS A `BashOperator`).	✔	☐	1
ERROR HANDLING: PYTHON INCLUDES PROPER ERROR HANDLING THROUGHOUT.	✔	☐	2
LOGGING: PYTHON INCLUDES LOGGING SPECIFIC TO THE PIPELINE PLATFORM, PREFECT OR AIRFLOW. PROJECTS SHOULD NOT IMPORT THE PYTHON "LOGGING" PACKAGE.	✔	☐	2
NO CRUFT: YOUR REPOSITORY SHOULD NOT CONTAIN ADDITIONAL FILES BEYOND THE PREFECT PIPELINE AND AIRFLOW DAG.	✔	☐	1
API REQUEST			4 points
API URL: THE CODE USES THE CORRECT ENDPOINT FOR THE SCATTER METHOD OF THE API.	✔	☐	1
SINGLE EXECUTION: THE ENDPOINT SHOULD ONLY BE CALLED ONCE PER FLOW EXECUTION.	✔	☐	2
POST METHOD: THE REQUEST TO THE API MUST USE THE "POST" METHOD ONLY.	✔	☐	1
MONITORING QUEUE			9 points
MONITORS WITH PRECISION: WORKFLOW INCLUDES METHOD FOR REPEATEDLY MONITORING MESSAGE QUANTITY IN SQS QUEUE.	✔	☐	5
BONUS FOR INNOVATIVE DESIGN THAT DOES NOT SIMPLY WAIT 900 SECS. CONSIDER A DESIGN THAT CHECKS AVAILABLE MESSAGES REGULARLY BEFORE PROCEEDING TO FETCHING MESSAGES.	✔	☐	4

COLLECT MESSAGES
11 points

WORKFLOW FETCHES ALL MESSAGES SUCCESSFULLY	☑	☐	2
WORKFLOW PARSES MESSAGE CONTENT PROPERLY	☑	☐	2
WORKFLOW PERSISTENTLY STORES PARSED MESSAGE CONTENT	☑	☐	3
WORKFLOW IMPLEMENTS ERROR HANDLING IN CASE OF NULL REPLY TO "RECEIVE_MESSAGES" METHOD.	☑	☐	2
WORKFLOW DELETES ALL MESSAGES SUCCESSFULLY	☑	☐	2

SUBMISSION
8 pts

SQS MESSAGE: CODE SUBMITS REASSEMBLED PHRASE, UVA COMPUTING ID, AND WORKFLOW PLATFORM VIA SQS MESSAGE TO INSTRUCTOR QUEUE.	☑	☐	4
ACCURACY: PHRASE SUBMITTED IS CORRECTLY ORDERED.			4

EXTENDED PROCESS
5 pts

BROADEN THE SCOPE OF THIS PROJECT BY WRITING AN ADDITIONAL WORKFLOW AS AN APACHE AIRFLOW DAG. THIS DAG SHOULD PERFORM THE SAME STEPS AS THE PREFECT FLOW AND REPRODUCE THE OUTPUT. THIS IS IN ADDITION TO WRITING A PREFECT FLOW, NOT A REPLACEMENT.	☑	☐	5