MSDS442
Project – Phase 6

Submission:

Submit your Project-Phase 6 as a SINGLE ZIP file on Canvas by the due date.
Your submitted ZIP file must have the name:
Project_Phase_6_Your_LastName.zip

Deliverables:

Your ZIP file for the exercise submission must include the following:

- All source code that you wrote, compiled and built on your personal computer.
- Panopto video recording of a live run of your code on your personal development computer.

Requirements specification:

Utilize **Kafka**, **Java**, **Python**, and **Neo4j GDS** library to implement the following requirements based on the requirements specification and architecture document for the OnMart Superstore real-time data streaming application that you reviewed in Phase 1:

1. After you create the CSV log file for at least 10,000 real-time transactions, perform the following tasks on the data log
    1) Detect fake or BOT-generated ratings/reviews utilizing the following rules:
        - High volume of ratings of a certain product in a short time span
        - Multiple ratings of the same product from the same user in 2 or more zip codes in a short period of time

    2) Build a classifier to detect credit card frauds utilizing the following rules:
        - If a credit card is used in a short period of time for multiple deliveries in different zip codes, then fraudulent transaction
        - If transaction amount > $1,000 and credit card is used for first time in the delivery zip-code, then fraudulent transaction
        - If transaction amount <= $1,000 and credit card is used for first time in the delivery zip-code, then not fraudulent transaction
        - If transaction amount > $1,000 and credit card is not used for first time in the delivery zip-code, then not fraudulent transaction
        - If transaction amount <= $1,000 and credit card is not used for first time in the zip-code delivery, then not fraudulent transaction

2. Use **Neo4j GDS** library in your implementation of the disaster recovery and contingency plan for OnMart logistics and supply chain network during the hurricane season in state of Florida, Find a backup warehouse for every serving warehouse in the state of Florida; the backup warehouse must be the **farthest** warehouse that has the **lowest number of shipments** within **150-200 miles radius** for every serving warehouse.