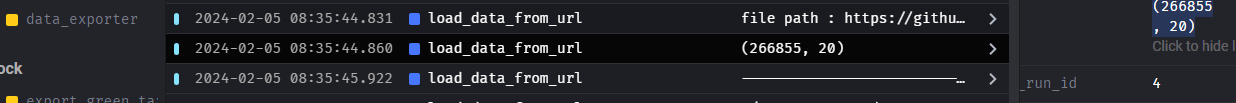
**Workflow Orchestration (Mage) – Homework**

Question 1. Data Loading

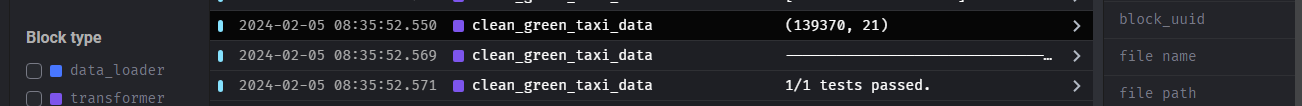
Once the dataset is loaded, what's the shape of the data?

**Ans: 266,855 rows x 20 columns**

****

Question 2. Data Transformation

Upon filtering the dataset where the passenger count is greater than 0 and the trip distance is greater than zero, how many rows are left?



Ans : 139,370 rows

Question 3. Data Transformation

Which of the following creates a new column lpep\_pickup\_date by converting lpep\_pickup\_datetime to a date?

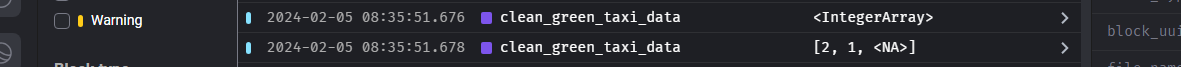
Ans: data['lpep\_pickup\_date'] = data['lpep\_pickup\_datetime'].dt.date

Question 4. Data Transformation

What are the existing values of VendorID in the dataset?

Ans: 1 or 2

<NA> stands for NULL values



Question 5. Data Transformation

How many columns need to be renamed to snake case?

Ans: 4

(VendorID, RatecodeID, PULocationID, DOLocationID)

Question 6. Data Exporting

Once exported, how many partitions (folders) are present in Google Cloud?

Ans : 96