PROBLEM STATEMENT

You are given a collection of fitness classes and it's bookings data.

Each of the classes has the following fields.

id = unique class id
name = name of the class
capacity = maximum no of bookings that is allowed for the class

And each of the bookings has the following fields.

id = unique booking id
class_id = id of the class
status = Indicates if the booking is successful or not. Allowed values ("BOOKED"/"FAILED")

Example input:

classes
---
c1 = {"1", "Stretch & Form session", 4}
c2 = {"2", "Reshape 45 Min", 10}

bookings
---
b1 = {"10", "1", "BOOKED"}
b2 = {"12", "1", "BOOKED"}
b3 = {"13", "1", "FAILED"}
b4 = {"14", "1", "BOOKED"}
b5 = {"15", "1", "FAILED"}
b6 = {"18", "2", "BOOKED"}

The booking performance of a class is considered good if the booking ratio is greater than 0.7 . Booking ratio is defined as "total number of successful bookings of a class" / "capacity of class". Only successful bookings (i,e with status = 'BOOKED') should be included for this ratio.

For the above example, the capacity of "Stretch & Form session" class is 4, total number of bookings is 3 and the booking ratio is .75 (i.e) 3/4

QUESTION: Write code to find the classes with good booking performance.

Example Output: { "Stretch & Form session”, 0.75 }

You can use any language of choice, but Python/Scala/Java would be preferred. Input for classes and bookings can be read from a text/csv/json file or any other input source that you prefer. A fully working tested code is expected. Jupyter notebook or an executable code that runs on a Spark cluster will be a plus.