
Scenario – Students & Courses

Step 1: Create DataFrames

```
# Student data
students_data = [
    (1, "Rahul Sharma", 20, "Bangalore"),
    (2, "Priya Singh", 21, "Delhi"),
    (3, "Aman Kumar", 19, "Hyderabad"),
    (4, "Sneha Reddy", 22, "Chennai"),
    (5, "Arjun Mehta", 23, "Mumbai"),
    (6, "Divya Nair", 20, None)   # Student without city
]

students_cols = ["student_id", "name", "age", "city"]
students_df = spark.createDataFrame(students_data, students_cols)

# Course data
courses_data = [
    (101, "Python", "Programming"),
    (102, "Data Science", "Analytics"),
    (103, "Databases", "Technology"),
    (104, "Business Studies", "Management")
]

courses_cols = ["course_id", "course_name", "category"]
courses_df = spark.createDataFrame(courses_data, courses_cols)

# Enrollment data
enrollment_data = [
    (1, 101, "A"),
    (2, 101, "B"),
    (3, 102, "A"),
    (4, 103, "C"),
    (5, 102, "B"),
    (7, 104, "A")   # Enrollment with non-existent student
]

enrollment_cols = ["student_id", "course_id", "grade"]
enrollment_df = spark.createDataFrame(enrollment_data, enrollment_cols)

# Show all DataFrames
students_df.show()
courses_df.show()
enrollment_df.show()
```

Step 2: Activities to Complete

Now that the DataFrames are created, your batch should solve these tasks

1. **Transformation Tasks**
    - Select all student names and their cities.
    - Find students who are older than 20.
    - List all courses under the "Analytics" category.
2. **Aggregation Tasks**
    - Count how many students are enrolled in each course.
    - Find the average age of students per city.
    - Get the maximum and minimum age of students.
3. **Join Tasks**
    - Join students with enrollments to see which student took which course.
    - Left join enrollments with courses to get course details.
    - Find students who are not enrolled in any course.
    - Find courses with no students enrolled.
4. **SQL Tasks**
    - Write a query to get all students with their course names and grades.
    - Find the number of students who got grade "A" in each course.
    - Find the top city with the most students enrolled in courses.
