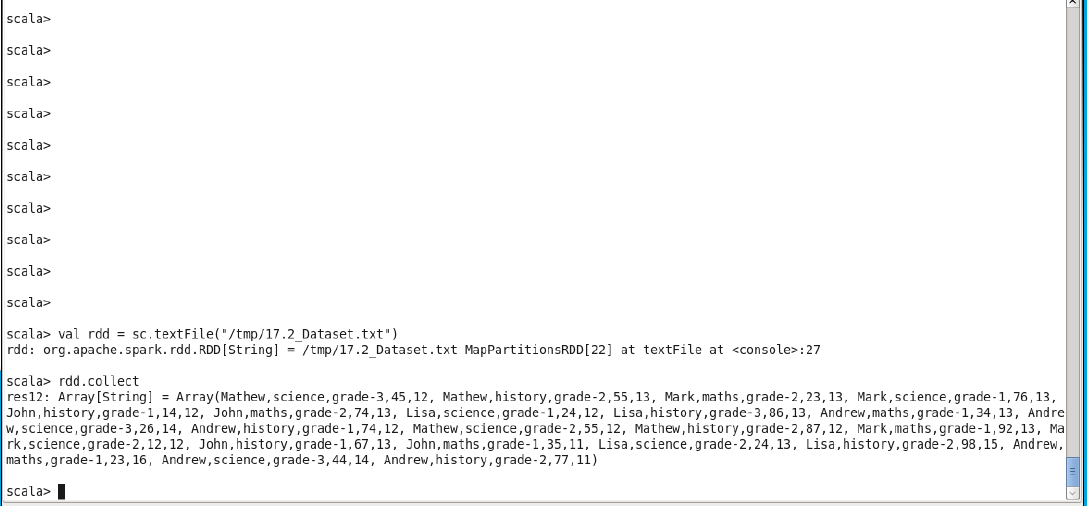
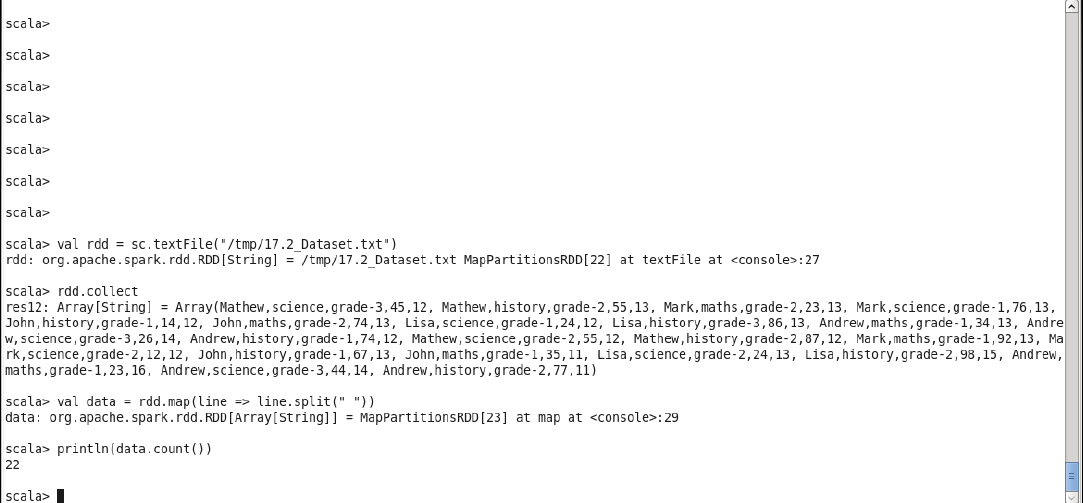
**Problem Statement 1:**

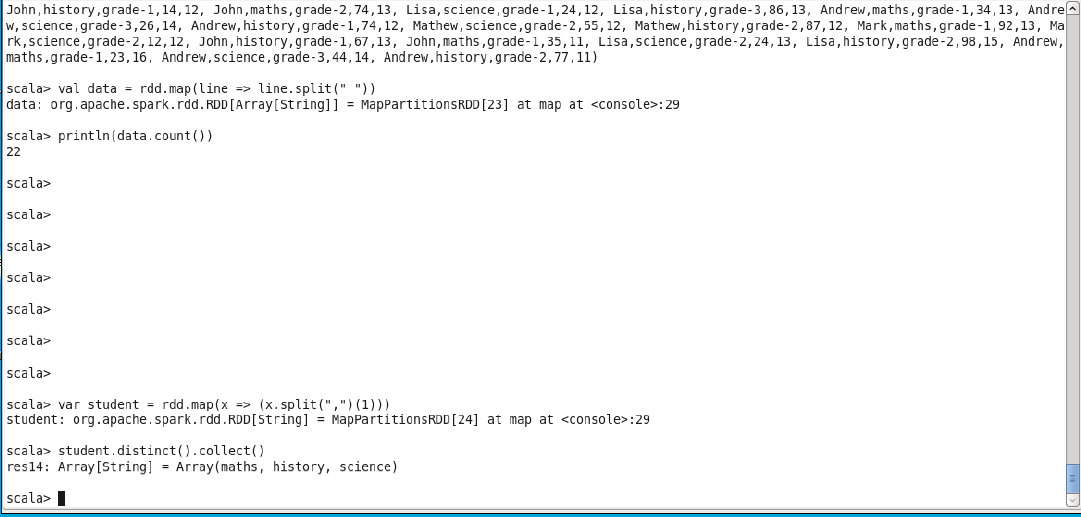
1. **Read the text file, and create a tupled rdd**

****

1. **Find the count of total number of rows present.**

****

1. **What is the distinct number of subjects present in the entire school**

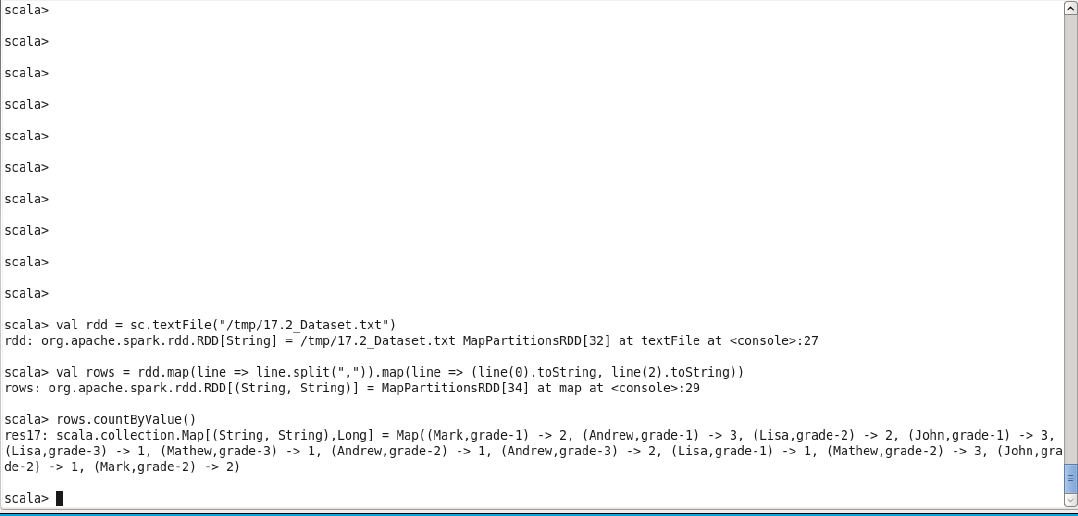
****

**4. What is the count of the number of students in the school, whose name is Mathew and marks is 55**

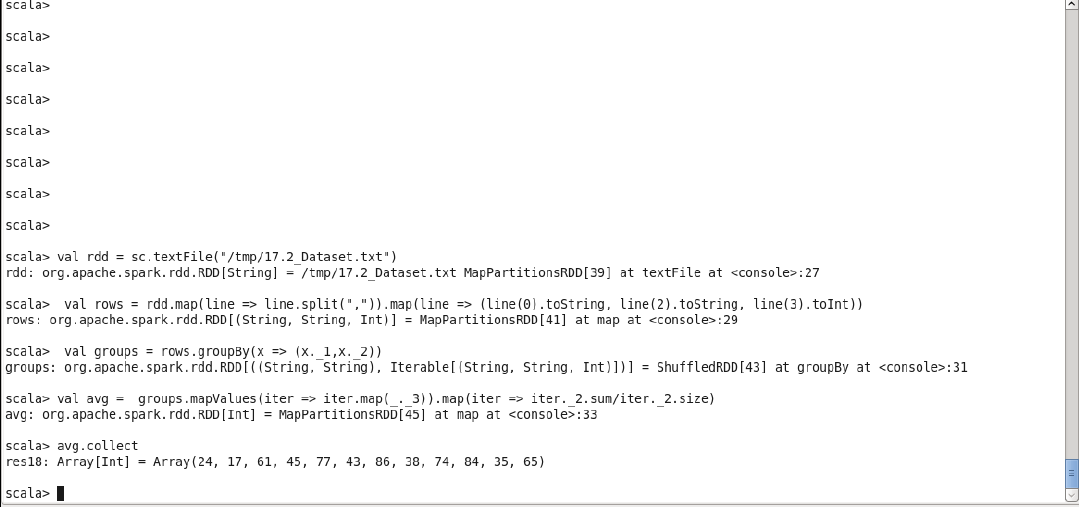
****

**Problem Statement 2:**

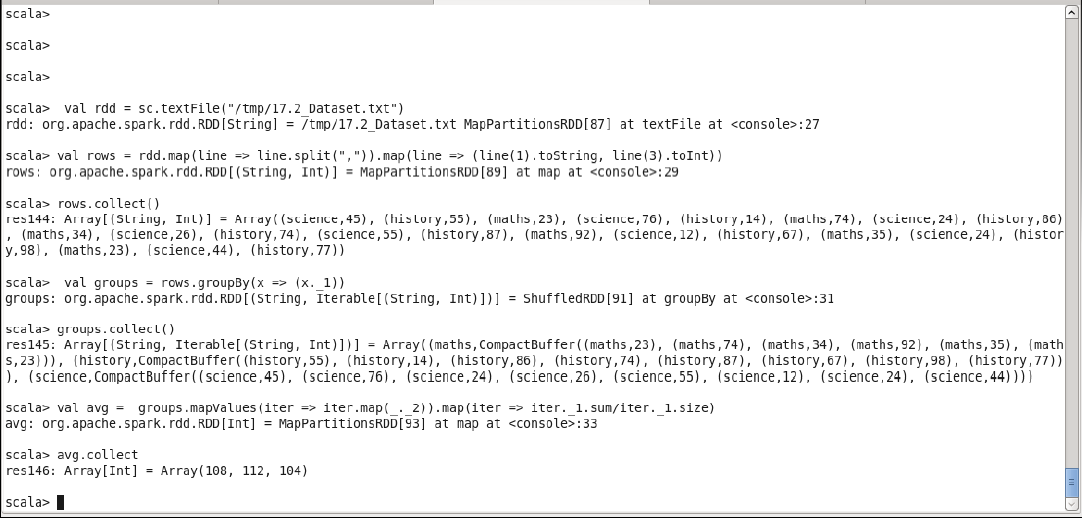
1. **What is the count of students per grade in the school?**

****

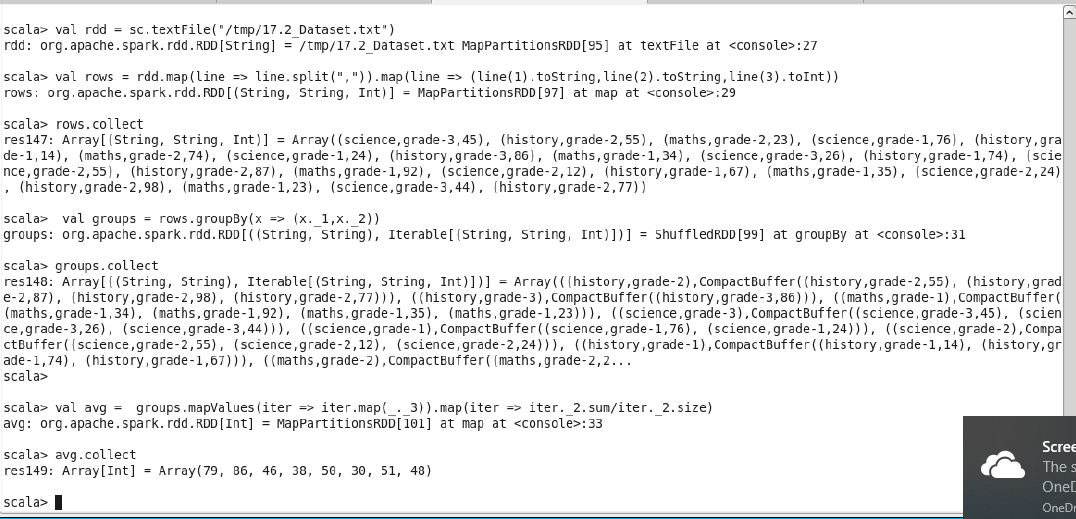
**2. Find the average of each student (Note - Mathew is grade-1, is different from Mathew in some other grade!)**

****

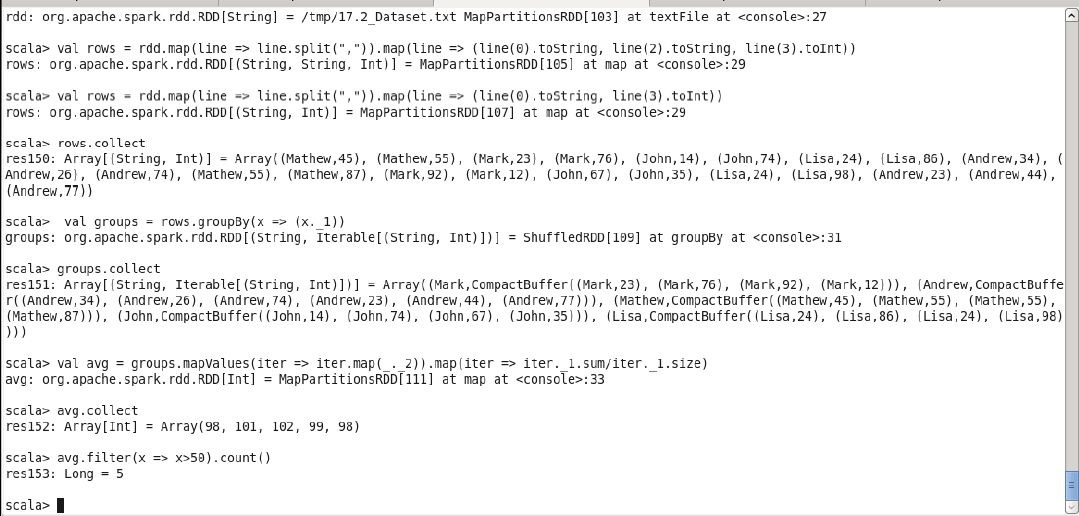
1. **What is the average score of students in each subject across all grades?**

****

1. **What is the average score of students in each subject per grade?**

****

1. **For all students in grade-2, how many have average score greater than 50?**

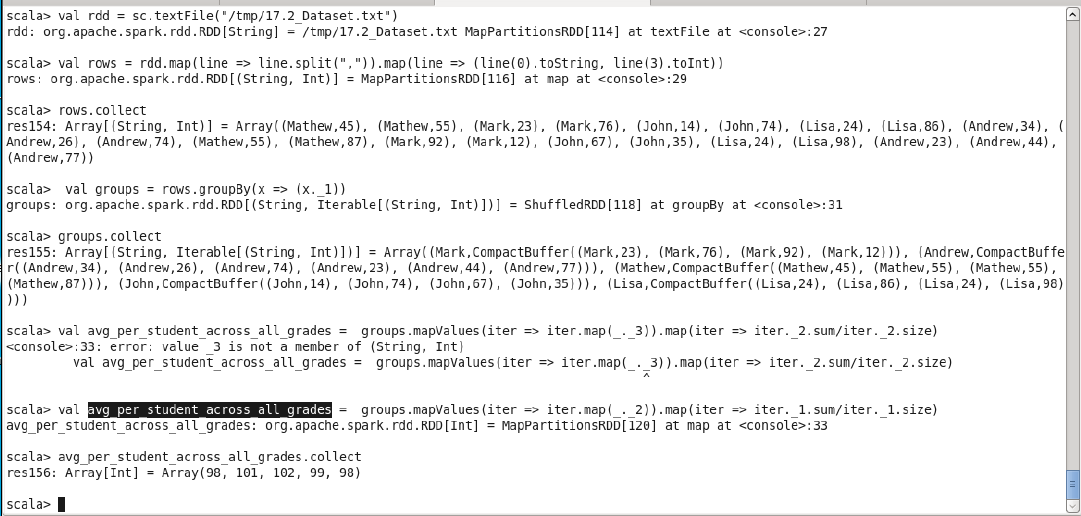
****

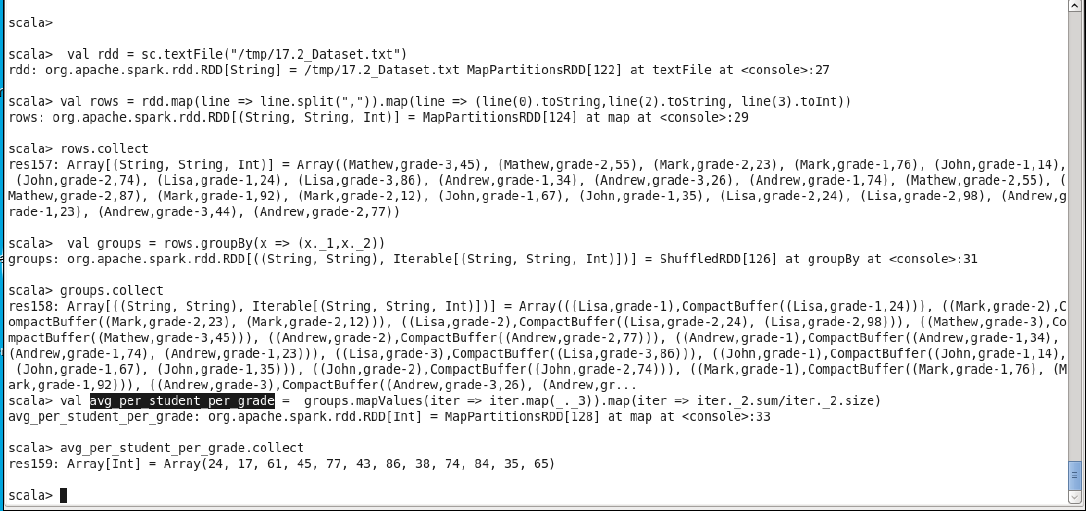
**Problem Statement 3:**

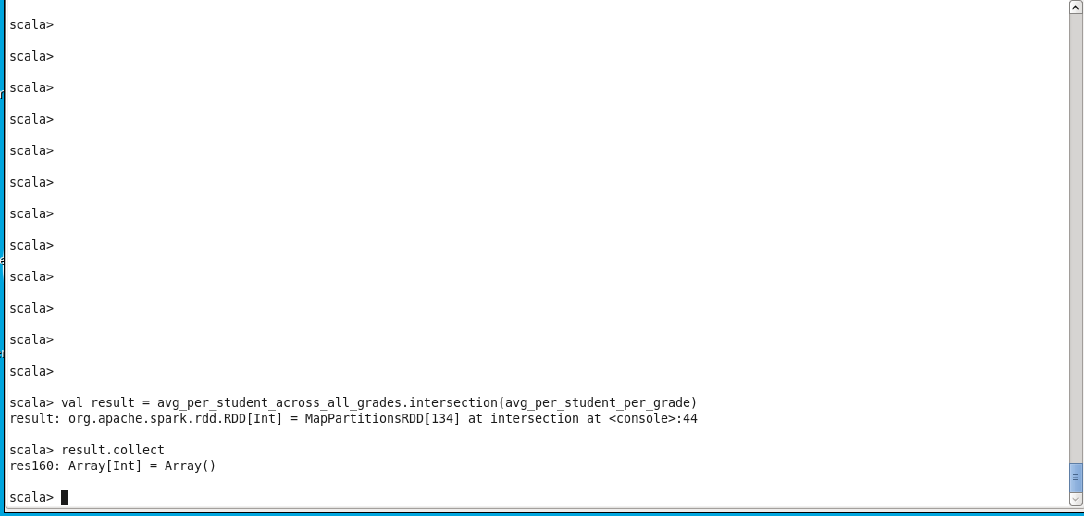
**Are there any students in the college that satisfy the below criteria :**

1. **Average score per student\_name across all grades is same as average score per student\_name per grade**

**Hint - Use Intersection Property**

****

****

****