**Question 3:** Harsha, a Cricket commentator, wants to know which players have played over 200 matches and have a highest score of 150 or above. Write a Spark SQL query using Cloudera or use data bricks?

1. Cricket Data is loaded into spark cluster. The unknown special characters are removed using regexp\_replace module in spark as shown.

Graphical user interface, text, application

Description automatically generatedText

Description automatically generated with medium confidence

1. A temporary table is created and a query on players who played 200 matches and highest score with 150 or above is executed using “where” condition.

Graphical user interface, application

Description automatically generated

Table

Description automatically generated

1. Number of records is 22.

Table

Description automatically generated

Graphical user interface, text, application, email

Description automatically generated

Link:

<https://databricks-prod-cloudfront.cloud.databricks.com/public/4027ec902e239c93eaaa8714f173bcfc/3166984684234423/2287749801263727/8643682393510715/latest.html>