Meets Specifications

Congratulations on completing the Cassandra project. You have demonstrated significant understanding of table design in Apache Cassandra. Keep up the good work.

Here are some references to learn more about Apache Cassandra:

Apache Cassandra Data Modeling and Query Best Practices
https://www.red-gate.com/simple-talk/sql/nosql-databases/apache-cassandra-data-modeling-and-query-best-practices/
Top 5 reasons to use the Apache Cassandra Database
https://towardsdatascience.com/top-5-reasons-to-use-the-apache-cassandra-database-d541c6448557
Introduction to Apache Cassandra
https://www.geeksforgeeks.org/introduction-to-apache-cassandra/

ETL

Pipeline Processing

Student creates `event_data_new.csv` file.

Student uses the appropriate datatype within the `CREATE` statement.
The data types used are appropriate.

Data Modeling

Student creates the correct Apache Cassandra tables for each of the three queries.
The `CREATE TABLE` statement should include the appropriate table.
Good job following the one table per query rule of Apache Cassandra. With this single table-single query approach, queries can perform faster.

Student demonstrates good understanding of data modeling by generating correct SELECT statements to generate the result being asked for in the question.

The SELECT statement should NOT use `ALLOW FILTERING` to generate the results.
Appropriate SELECT statements are used. Use of "SELECT * " is avoided. Well done!

Student should use table names that reflect the query and the result it will generate. Table names should include alphanumeric characters and underscores, and table names must start with a letter.
Table names are good. We should use table names that reflect the query and the result it will generate. For e.g., for query 2, an appropriate table name should reflect song playlist_in_session (e.g., name could be song_playlist_session). This is important because this describes the data model.

The sequence in which columns appear should reflect how the data is partitioned and the order of the data within the partitions.
The sequence of the columns in the CREATE and INSERT statements follow the order of the COMPOSITE PRIMARY KEY and CLUSTERING columns. Well done!

PRIMARY KEYS

The combination of the PARTITION KEY alone or with the addition of CLUSTERING COLUMNS should be used appropriately to uniquely identify each row.

Appropriate PRIMARY KEYs are used.

Presentation

The notebooks should include a description of the query the data is modeled after.

- Please explain in one or two sentences why those primary key and clustering columns are used .
- You could also include headers right above the SELECT statement cell to highlight the responses to the questions.

Here is an example description for Query 1:

Description: Here the Primary Key has two fields: sessionId is the partition key, and itemInSession is clustering key.

Partitioning is done by sessionId and within that partition, rows are ordered by the itemInSession.

Code should be organized well into the different queries. Any in-line comments that were clearly part of the project instructions should be removed so the notebook provides a professional look.

Nice job removing all the `TODO` statements. The code looks clean and professional.