UDACITY

Return to Classroom

DISCUSS ON STUDENT HUB

Data Modeling with Cassandra

REVIEW

HISTORY

Meets Specifications

Congratulations on completing **Data Modeling with Cassandra** project. Nice job using the appropriate queries. You have demonstrated a 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.

- The csv file has the expected number (6821) of records.

Student uses the appropriate datatype within the `CREATE` statement.

- The data type `float` is ideal for the column length.
    - Ref: https://itprimer.wordpress.com/2016/04/12/cassandra-tips/
    - https://www.oreilly.com/library/view/learning-apache-cassandra/9781787127296/dcd51887-3fad-4e4e-b1d1-491ef3d227b6.xhtml
- You can learn more about the Cassandra data types from this documentation: https://docs.datastax.com/en/cql/3.3/cql/cql_reference/cql_data_types_c.html

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.

- Please remove the `song` from the SELECT statement of Query 3.

Make sure the SELECT statements request only the data required.
For the Query 3, SELECT statement should not require anything more than user name first and last name in the SELECT statement

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.

- Nice choice of table names. We should use table names that reflect the query and the result it will generate.

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 follows the order of the COMPOSITE PRIMARY KEY and CLUSTERING columns.

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.

- Nice job with the query descriptions.

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.

- All the `TO-DO` instructions are removed.
- The notebook looks clean and organized.

DOWNLOAD PROJECT

RETURN TO PATH

Rate this review

START