UDACITY

< Volver al aula

Data Modeling with Cassandra

REVISIÓN

HISTORIAL

Cumple las especificaciones

Dear student,

Excellent job on this project!
All the requested changes have been implemented and you've successfully completed this project. 👏👏

- You've created all three tables with correct data types and excellent choices of primary keys.
- The notebook is clean and organized, all tables have a description of how the primary key was chosen and you've selected the necessary columns to answer the three questions.
- The tables have proper names and the column's sequence reflects how the tables are partitioned.

I hope you've enjoyed and learned a lot from it.

Good luck with your next Data endeavors!😀

ETL Pipeline Processing

Student creates `event_data_new.csv` file.

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

Data Modeling

Student creates the correct Apache Cassandra tables for each of the three queries. The `CREATE TABLE` statement should include the appropriate table.

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.

☑ **Great job selecting only the requested columns for each table.**
This is very important because Cassandra usually contains lots of data and also receives many concurrent INSERTs. And since we're talking about a lot of data, selecting **all** columns every time will put an unnecessary load in the Cassandra cluster.

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.

The sequence in which columns appear should reflect how the data is partitioned and the order of the data within the partitions.

☑ **Awesome job.** The column order in the `CREATE` and `INSERT` statements correspond to the `PRIMARY KEY` definitions.

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.

Presentation

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

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

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.

DESCARGAR PROYECTO

VOLVER A RUTA

Califique esta revisión

INICIAR