

[Return to Classroom](#)[DISCUSS ON STUDENT HUB](#)

Data Modeling with Cassandra

REVIEW

CODE REVIEW

HISTORY

Meets Specifications

Dear Student,

Congratulations on clearing the project.

You should be proud of this achievement. Check this discussion to learn more about [choosing a partition key in cassandra](#)

I'd recommend you to check out [this](#) discussion on Quora that talks about all the big companies that are using Cassandar in the real word. I'm sure you will find it interesting.

I am happy to pass this.

Regards.

ETL Pipeline Processing



Student creates `event_data_new.csv` file.

Nice! the `event_data_new.csv` file is properly generated.



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

Good work. You have made proper use of the various datatypes available in Cassandra.

Rate this review



Data Modeling

✓	Student creates the correct Apache Cassandra tables for each of the three queries. The <code>CREATE TABLE</code> statement should include the appropriate table.
	<p>Nicely done.</p> <p>In practice, this generally means you will use roughly one table per query pattern. If you need to support multiple query patterns, you usually need more than one table.</p> <p>To put this another way, each table should pre-build the "answer" to a high-level query that you need to support. If you need different types of answers, you usually need different tables. This is how you optimize for reads.</p> <p>Remember, data duplication is okay. Many of your tables may repeat the same data.</p>
✓	Student demonstrates good understanding of data modeling by generating correct <code>SELECT</code> statements to generate the result being asked for in the question.
	<p>The <code>SELECT</code> statement should NOT use <code>ALLOW FILTERING</code> to generate the results.</p> <p>The select statements are on point. Good job only selecting the columns that are required instead of doing a <code>select *</code>. This helps in performance when you are loading from a large table with loads of column. If you <code>select *</code>, you'll end up loading a lot of columns that you won't use and they would take up precious memory.</p>
✓	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.
	<p>Nice! The table names reflect the query it answers. Always remember, in Cassandra, every table represents a query. So, try to come up with names that represent the query.</p>
✓	The sequence in which columns appear should reflect how the data is partitioned and the order of the data within the partitions.
	<p>Awesome work maintaining the order of <code>CREATE</code> and <code>INSERT</code> statements to follow the order of the <code>COMPOSITE PRIMARY KEY</code> and <code>CLUSTERING</code> columns. This is important because Apache Cassandra is a partition row store, which means the partition key determines which any particular row is stored on which node. In case of composite partition key, partitions are distributed across the nodes of the cluster and how they are chunked for write purposes. Any clustering column(s) would determine the order in which the data is sorted within the partition.</p>

PRIMARY KEYS

✓	The combination of the <code>PARTITION KEY</code> alone or with the addition of <code>CLUSTERING COLUMNS</code> should be used appropriately to uniquely identify each row.
	<p>The combination of <code>PRIMARY KEYS</code> have been appropriately used in all the three queries and uniquely identifies each row in the tables. Good job.</p> <p>Good Read: Difference between partition key, composite key and clustering key in Cassandra?</p>

Presentation

✓	The notebooks should include a description of the query the data is modeled after.
	<p>Good work. You have also detailed how the partition key, composite key was arrived at. Excellent!</p>
✓	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.
	<p>Good job removing the <code>TODO</code> statements and in-line comments that were clearly part of the project instructions. The notebook provides a professional look now.</p>