

WHERE clause

- Data Modeling in Apache Cassandra is query focused, and that focus needs to be on the WHERE clause
- Failure to include a WHERE clause will result in an error

Additional Resource

AVOID using "ALLOW FILTERING": Here is a reference to the [DataStax document](#) that explains ALLOW FILTERING and why you should not use it.

Commonly Asked Questions:

Why do we need to use where statement since we are not concerned about analytics? Is it only for debugging purposes?

The `WHERE` statement is allowing us to do the fast reads. With Apache Cassandra, we are talking about big data -- think terabytes of data -- we are making it fast for read purposes. Since the data is spread across all the nodes. By using the `WHERE` statement, we know which node to go to, from which node to get that data and serve it back. For e.g., imagine we have 10 years of data on 10 nodes or servers. So 1 year's data is on a separate node. By using the `WHERE year = 1` statement we know which node to go fast to pull the data from.