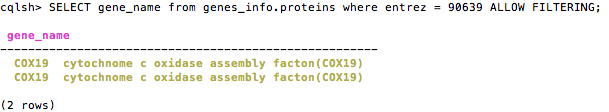
REPORT:

1) Need to store info about genes. Certain genes can have more information than others stored in additional columns. Hence, Cassandra is a natural choice for this part of the project.

2) The data is stored in a table where each row can have different number of columns. Each row represents a gene, and columns its properties.

3) SELECT ‘column’ from genes\_info.proteins WHERE ‘conditions’ ALLOW FILTERING;

4) Example:

5) Cassandra does not have JOIN analogues, and as a result, tables tend to be denormalized.