

What are statistics?

- When the optimizer is given a set of query commands, the optimizer will select the best and most efficient way to execute that query. To return the most efficient execution plan, the optimizer will rely on statistics on the table and columns to attain that result.
- So, statistics are objects that contain **statistical information about the distribution of values in one or more columns** of a table or indexed view. What do we mean by 'distribution of values in one or more columns'? As we shall see with the DBCC command, can view the 'internals' of the statistics.

Why do we need them?

- The Query Optimizer uses these statistics to estimate the number of rows or the *cardinality* returned in the query result.
- The Query Optimizer uses this cardinality estimates to choose either the index seek operator instead of the more resource-intensive index scan operator and return the most efficient execution plan.

When to create statistics

- SQL Server automatically creates statistics when you create a table with PK
- Or when you create an index on an existing table
- When you rebuild an index it automatically creates new statistics
- By default, SQL Server has the auto create statistics enabled (don't turn it off)