Maintaining Statistics

Gail Shaw
TECHNICAL LEAD

@SQLintheWild http://sqlinthewild.co.za

Overview

- What are statistics and why are they important?
- Consequences of out of date statistics
- Maintenance options
- Guidelines

What are Statistics

Aggregated data about data in table

Measure of uniqueness of column

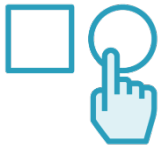
Distribution of data within columns

Not kept in sync with the table

Why are they important?

Statistics are used by the query optimizer in the generation of execution plans to:

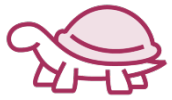
Estimate rows affected and data sizes



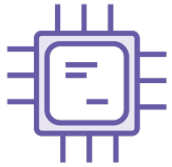
Choose appropriate operators for the number of rows

Calculate memory needed to run query

Effects of incorrect statisics



Slow queries



High CPU

Excessive IOs

Spills to TempDB

Demo

Examine what information statistics contain

Look at the effects of incorrect statistics on a query plan

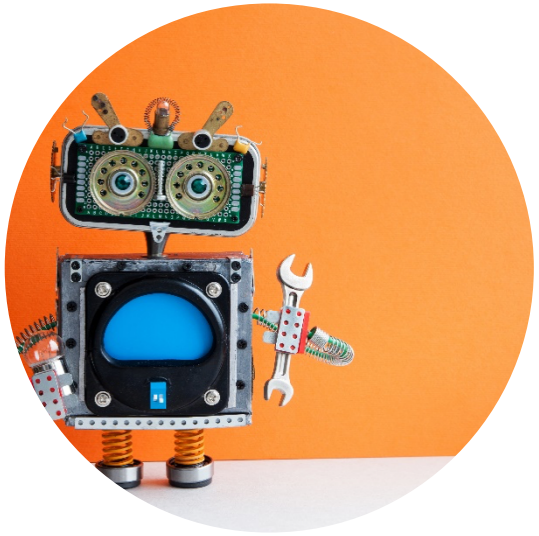
Automatic stats updates

Triggered after a certain number of changes to the underlying table

Update occurs the next time the stats are needed

Sampled update on larger tables

Options for keeping stats up to date



Auto updates

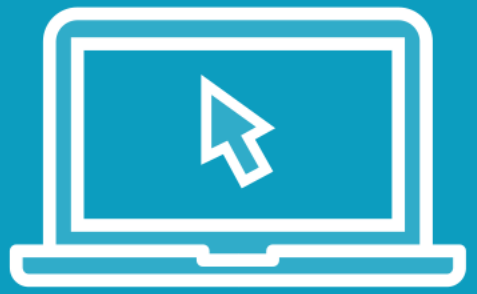
Index Maintenance

Manual updates

Syntax for Statistics Updates

```
UPDATE STATISTICS TableName [ALL | <index name>]
WITH
FULLSCAN | SAMPLE |RESAMPLE
[NORECOMPUTE]
INCREMENTAL = {ON|OFF}
```

Demo



Updating out of date statistics

Guidelines for stats maintenance

Do some!

If you have the time, update all stats with full scan every night

Hard to over-do stats maintenance

Be careful of doing sampled updates after index maintenance

Summary

- What statistics are and how they're used
- Impact of out of date statistics
- Maintenance options
- Guidelines