Troubleshooting Bad Assumptions

Joe Sack
Joe@SQLskills.com
http://www.SQLskills.com/blogs/Joe

pluralsight
hardcore developer training

Module Introduction

- **When assumptions about row estimates for logical operations are incorrect, query performance can suffer**

- **In this module we'll walk through common areas where bad assumptions can occur that negatively impact overall query performance**
  - I'll demonstrate several problems and their associated solutions

Model Assumptions

- **The query optimizer makes assumptions that may or may not reflect the current state of your data**
- **Independence**
  - Filters are uncorrelated in absence of statistics indicating otherwise
- **Uniformity**
  - Values in a histogram step are evenly distributed (spread) and have the same frequency
- **Inclusion**
  - When using a column-equal-constant predicate, it is assumed the value actually exists
- **Containment**
  - When estimating an equality join, it is assumed that there is a maximum overlap of distinct values (think “PK-to-FK” relationship”)

Troubleshooting Questions (1)

- Are my statistics missing?
- Are my statistics stale?
- Is the sampling adequate? (jagged distributions, “lossy” histograms)
- Would multi-column statistics help?
- Is this a parameter sensitivity issue?
- Are table variables and/or MSTVFs causing problems?

Troubleshooting Questions (2)

- Am I usually only querying the most recent rows?
- Are data type conversions causing issues?
- Am I comparing columns from the same table?
- Am I accessing remote data sources?
- Are my predicates being buried in complexity?
- Am I trying to do too much within a single query? (deep tree)
- For more on this topic, see the *SQL Server: Troubleshooting Query Plan Quality Issues* course, http://bit.ly/PlanQuality

Module Summary

- In this module we walked through common areas where bad assumptions can occur that negatively impact overall query performance
- You saw several problems related to bad assumptions and their associated solutions

- Now assuming we've addressed bad assumptions, in the next module we'll begin coverage of query construction issues and the problems related to imprecision