Question
Kevin Furze · May 14, 2016

How do I force outliers to be ignored - 2016.1

I have an sql that is refusing to play fairly.

I have

- tuned affected tables
- made sure I'm on the latest version (2016.1)
- got enough global buffers
- tried forcing the sql to use outliers correctly

Still nothing. here's the SQL
SELECT id , dealerid -> extnumber AS exNum , dealerid -> deliveryorder AS ord , dealerid -> deliverytown AS town , consTypeID , constypeid -> GroupInitials AS gty , barcode , constypeid , constypeid -> Initials AS ty , constypeid -> MaxDaysInWarehouse AS maxWh , daysInWh , dealerid , statusActual , wgtCurrent , packType -> ShortDesc , jdType , statusCode , mustGo , wgtAccurate FROM sqluser . newconsignments WHERE dealerid -> routeid = 4 AND statusCode IN (80,240) AND consolidateID = 0

Its the consolidateID=0 that's the problem, It's a similar issue talked of in https://community.intersystems.com/post/global-summit-2016-find-and-fix-slow-query the issue talked of "closed cases" - in my case, I there are many more values of zero than other values (ratio 500,000:150)

I get a rowcount of 31 and , performance of 4.625 and global counts of just over 1,000,990 using that sql

the plan clearly shows it's dropping down the consolidateID index first, then the dealerid->routeid, then the statusCode but has to drop down 500,000+ consolidateID's and then find the dealers and stautsCodes.

I've looked at the indicies (mostly bitmaps) and the selectivities after tuning tables,

```
<Property name="consolidateID">
<AverageFieldSize>1</AverageFieldSize>
<OutlierSelectivity>.998115:0</OutlierSelectivity>
<Selectivity>0.0357%</Selectivity>
</Property>
```

```
<Property name="statusCode">
<AverageFieldSize>3</AverageFieldSize>
<OutlierSelectivity>.582469:300</OutlierSelectivity>
<Selectivity>4.1753%</Selectivity>
</Property>
```

```
<Property name="dealerID">
<AverageFieldSize>3</AverageFieldSize>
<Selectivity>0.6289%</Selectivity>
</Property>
```

```
<Property name="RouteID">
<AverageFieldSize>1.63</AverageFieldSize>
<Selectivity>4.1667%</Selectivity>
</Property>
```

and all the selectivities suggest that the consolidationID (0.0357%) should be used last as the other selectivities are much higher.

I though I would follow the advice and change the wehere clause to AND ((consolidateID = 0)) by using the double brackets but all that did getter a slightly better speed (probably because the data's in the buffers now) and the global count's went down to 1,000,831 (a drop of about 100) and it's still dropping down the consolidateID first (wrongly)

I use this SQL inside a fetch chunk of code and so to make it so much more efficient, I deleted the AND ((consolidateID = 0)) clause and the fetched just the 30-50 rows and ignored those where consolidateID<>0

so of course, the question is:

how do I force the AND ((consolidateID = 0)) to be used as the LAST filter, not the first

kevin

#SQL #Caché

Source URL:https://community.intersystems.com/post/how-do-i-force-outliers-be-ignored-20161