**‘AccessBank DBA Internship Program’**

**Project 15: Updating Statistics**

Today you received an email from one of the developers who had a slow-performing query. When you checked the statistics, you found out that those statistics for the index, which is used while executing the query, was lastly updated 5 months ago. It means there is a serious problem with the statistics. Fix the problem. Here is the problematic query:

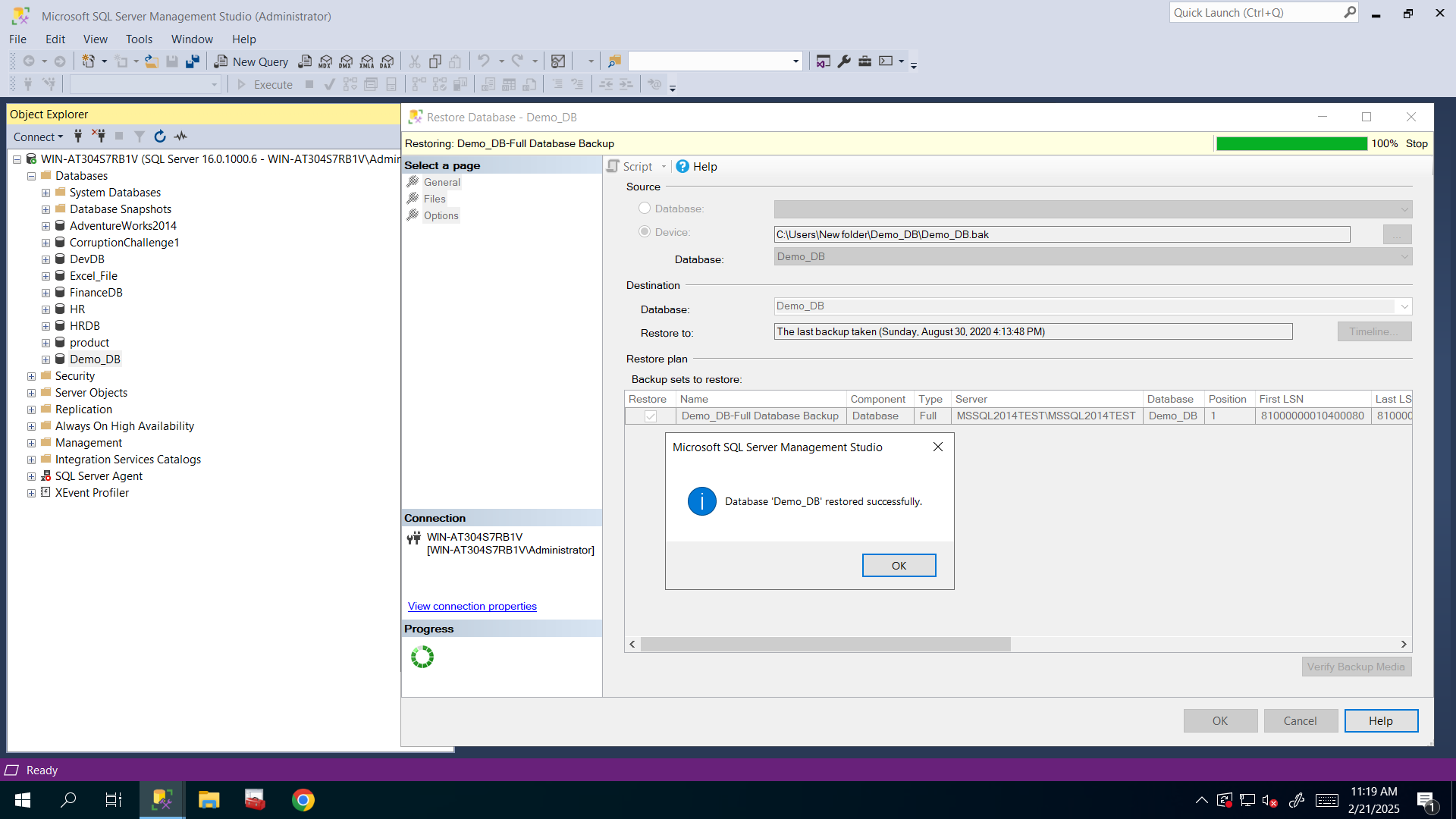
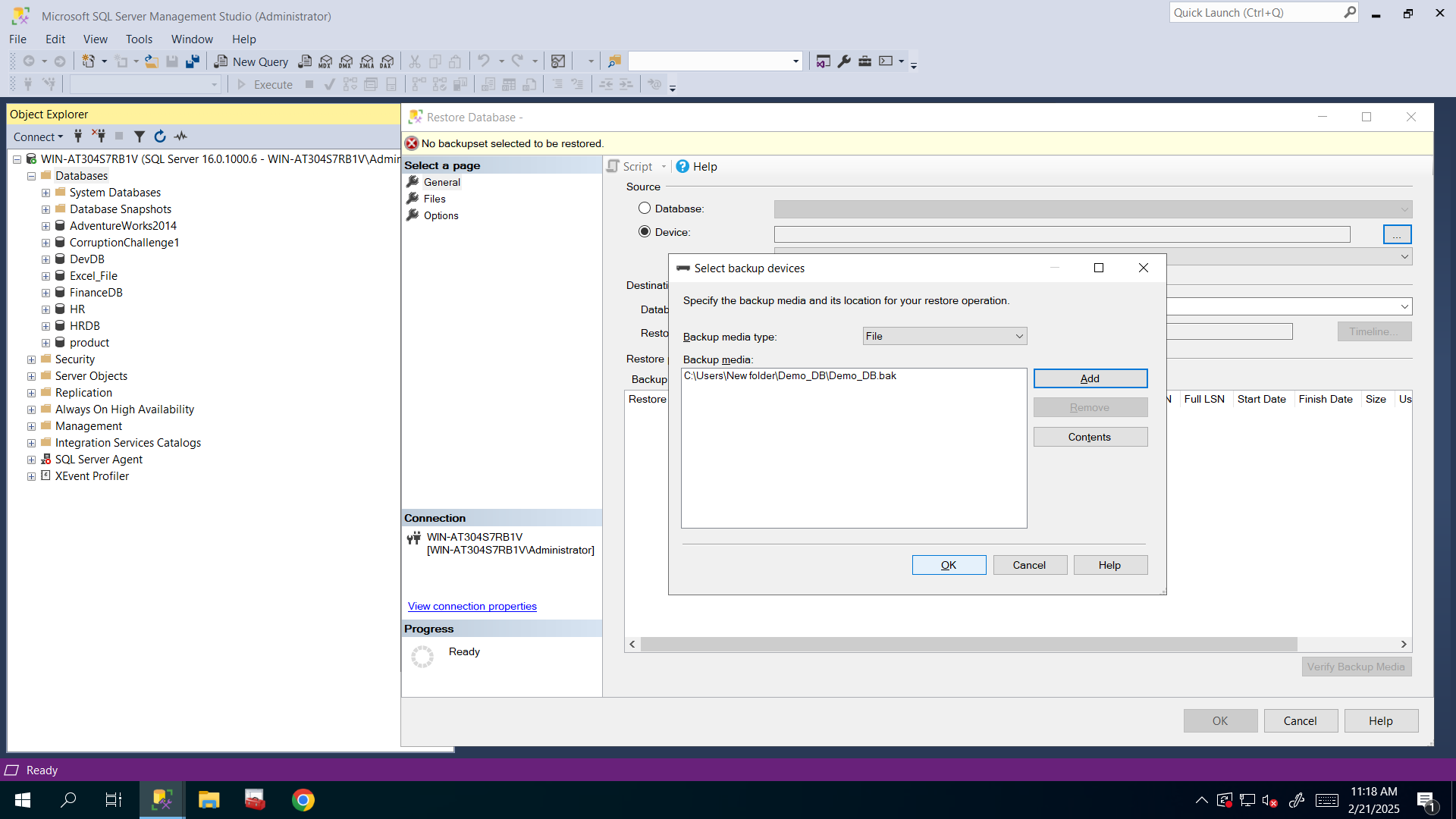
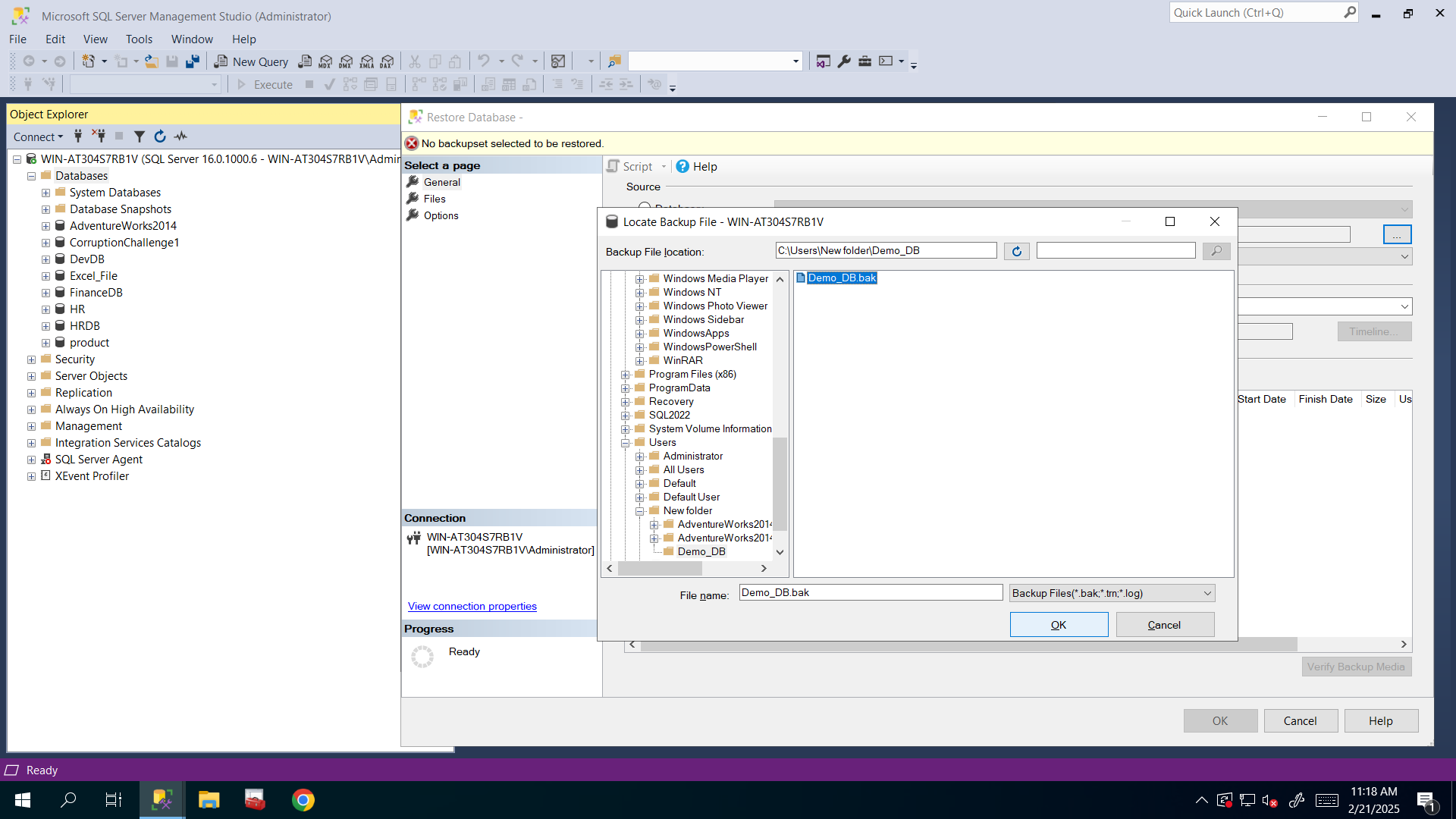
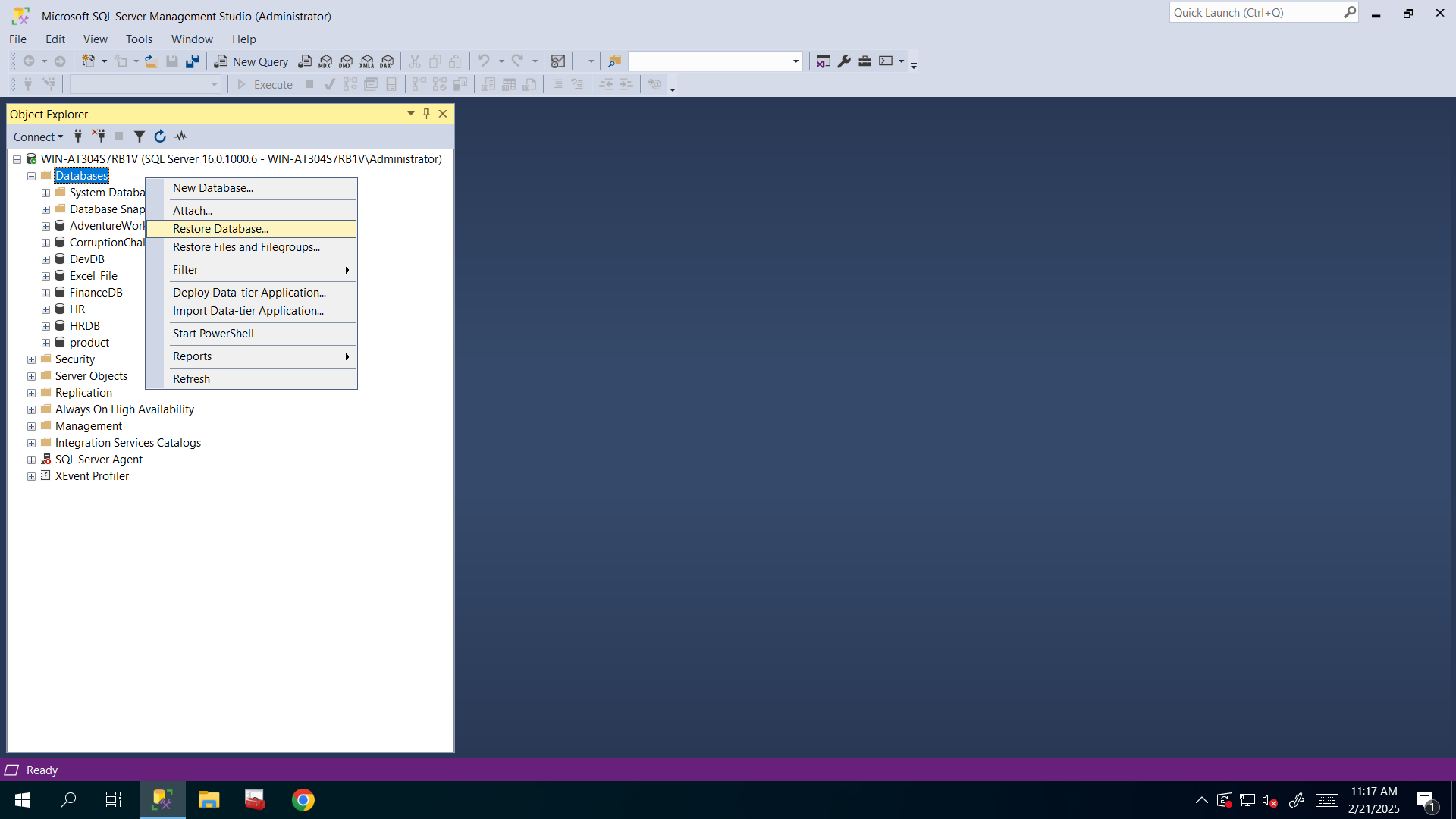
SELECT count(RandNum1) FROM STAT\_TEST WHERE RandNum2 < 5000000

group by RandNum1

1. restore DEMO\_DB database:

<https://drive.google.com/file/d/1sX-Dxq2wjTUsNxfPBNfCN-nU1NL_nsiZ/view?usp=sharing>

**First of all, I restored the Demo\_DB database by using provided backup file after downloading it on my PC.**



1. Identify outdated statistics using the below query:

USE Demo\_DB

GO

SELECT o.name 'Table Name', i.name 'Index Name', i.type 'Index Type'

,STATS\_DATE(i.[object\_id], i.[index\_id]) 'Last Statistics Update'

FROM sys.indexes i, sys.objects o

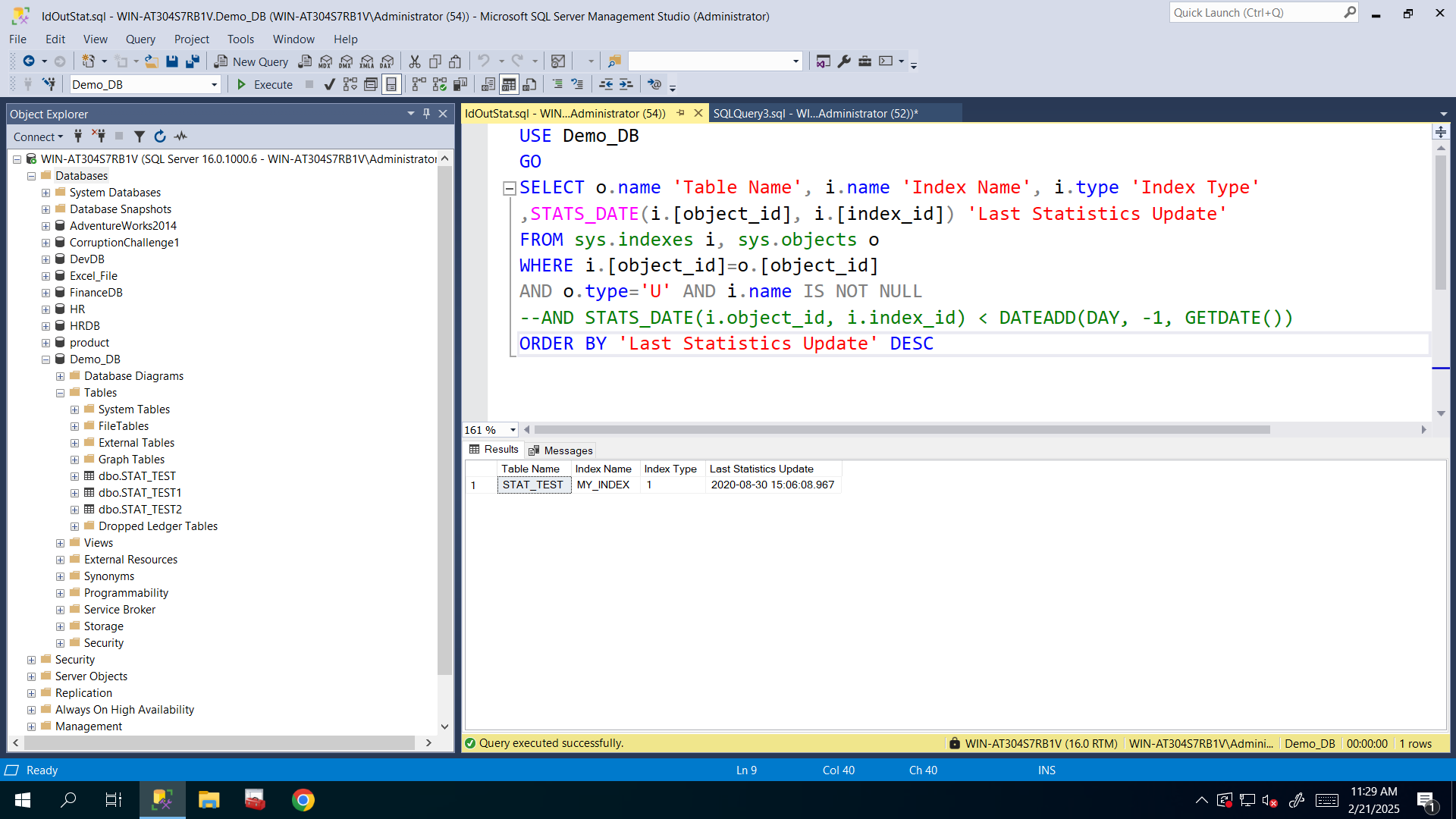
WHERE i.[object\_id]=o.[object\_id]

AND o.type='U' AND i.name IS NOT NULL

--AND STATS\_DATE(i.object\_id, i.index\_id) < DATEADD(DAY, -1, GETDATE())

ORDER BY 'Last Statistics Update' DESC

**After executing above query, I got that STAT\_TEST table’s MY\_INDEX index’s statistics are outdated so I need to update them.**

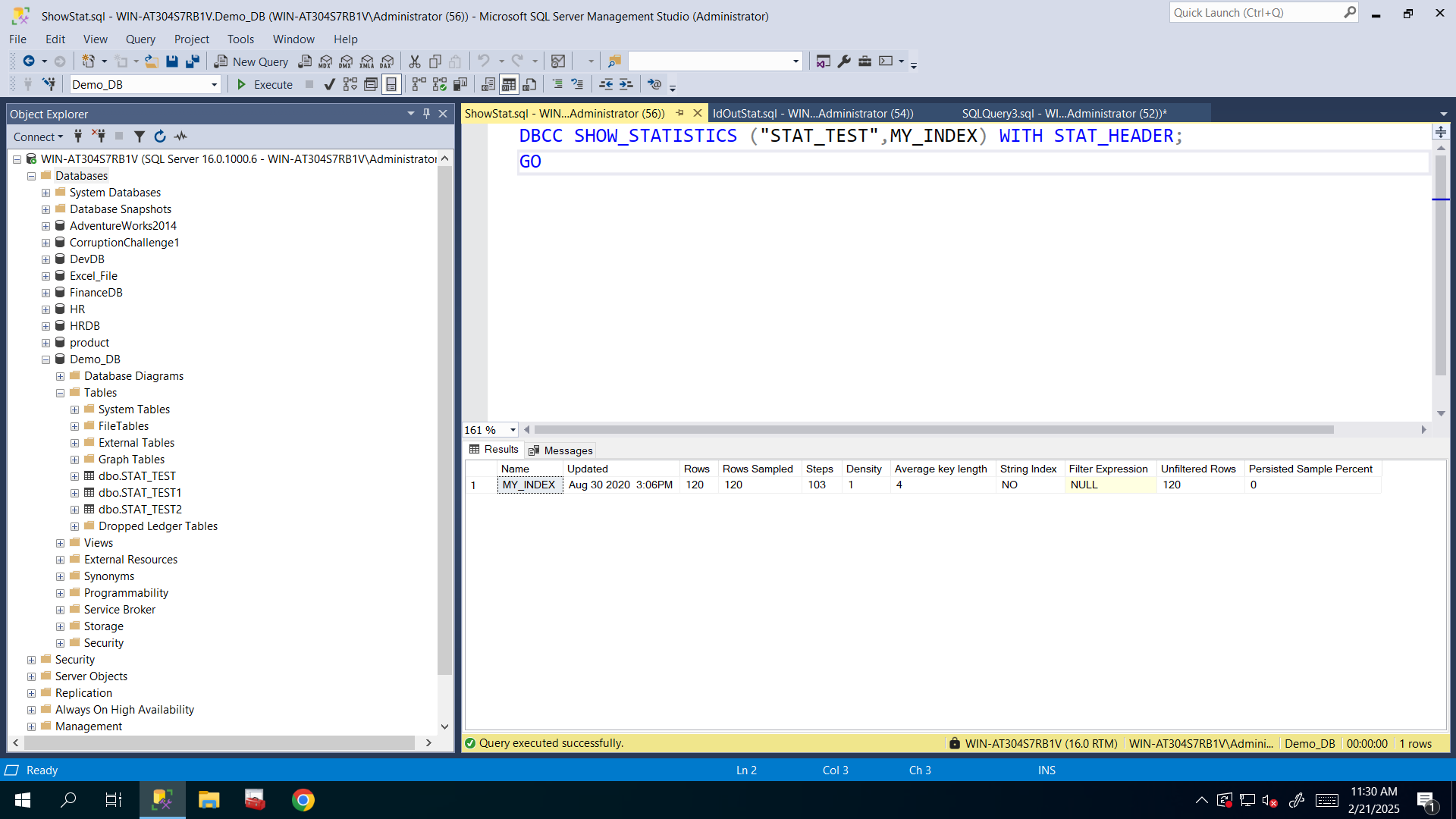


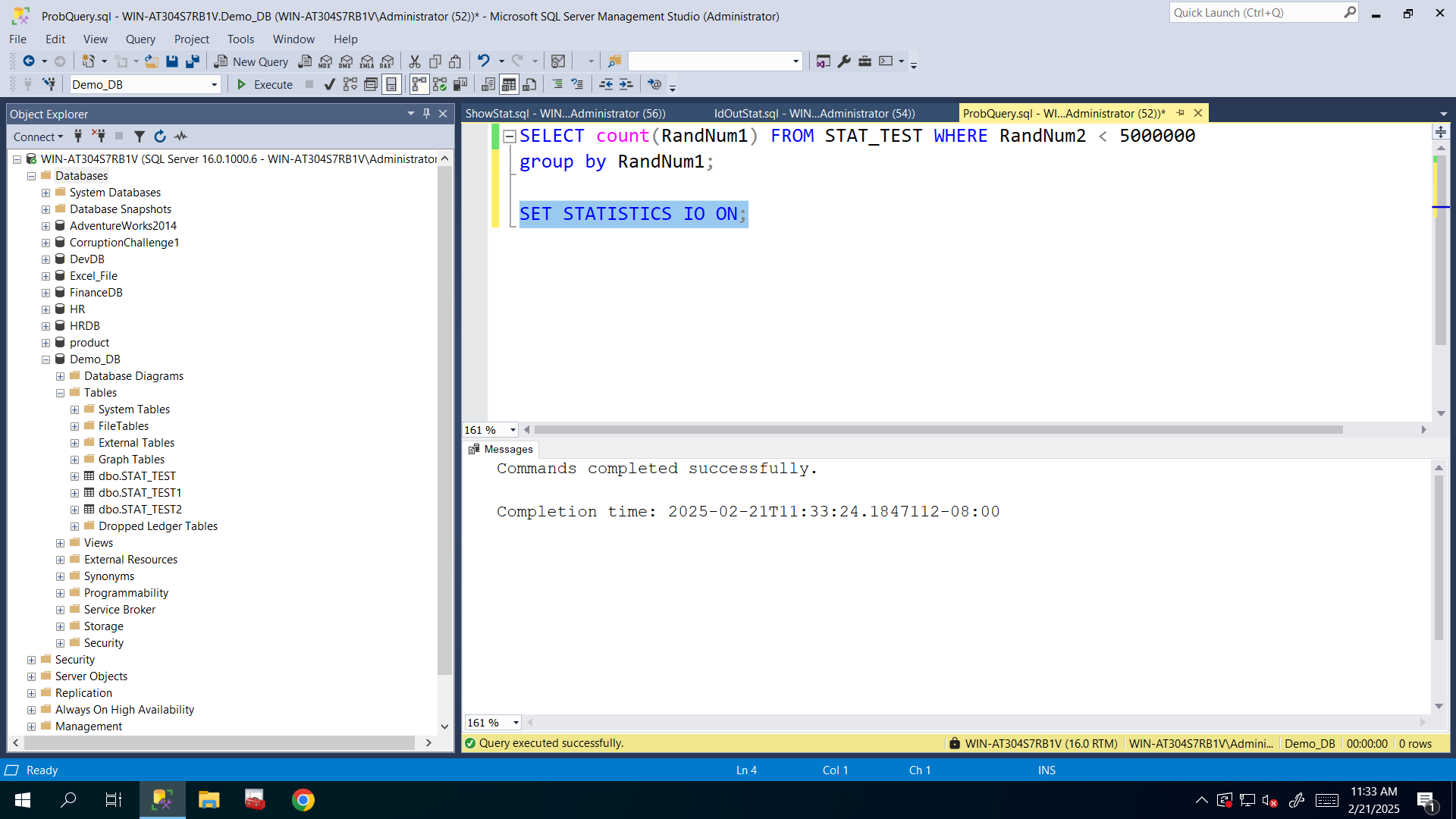
1. Go further and find out for how many rows does statistics keep information using the below query:

DBCC SHOW\_STATISTICS ("TABLE\_NAME",INDEX\_NAME) WITH STAT\_HEADER;

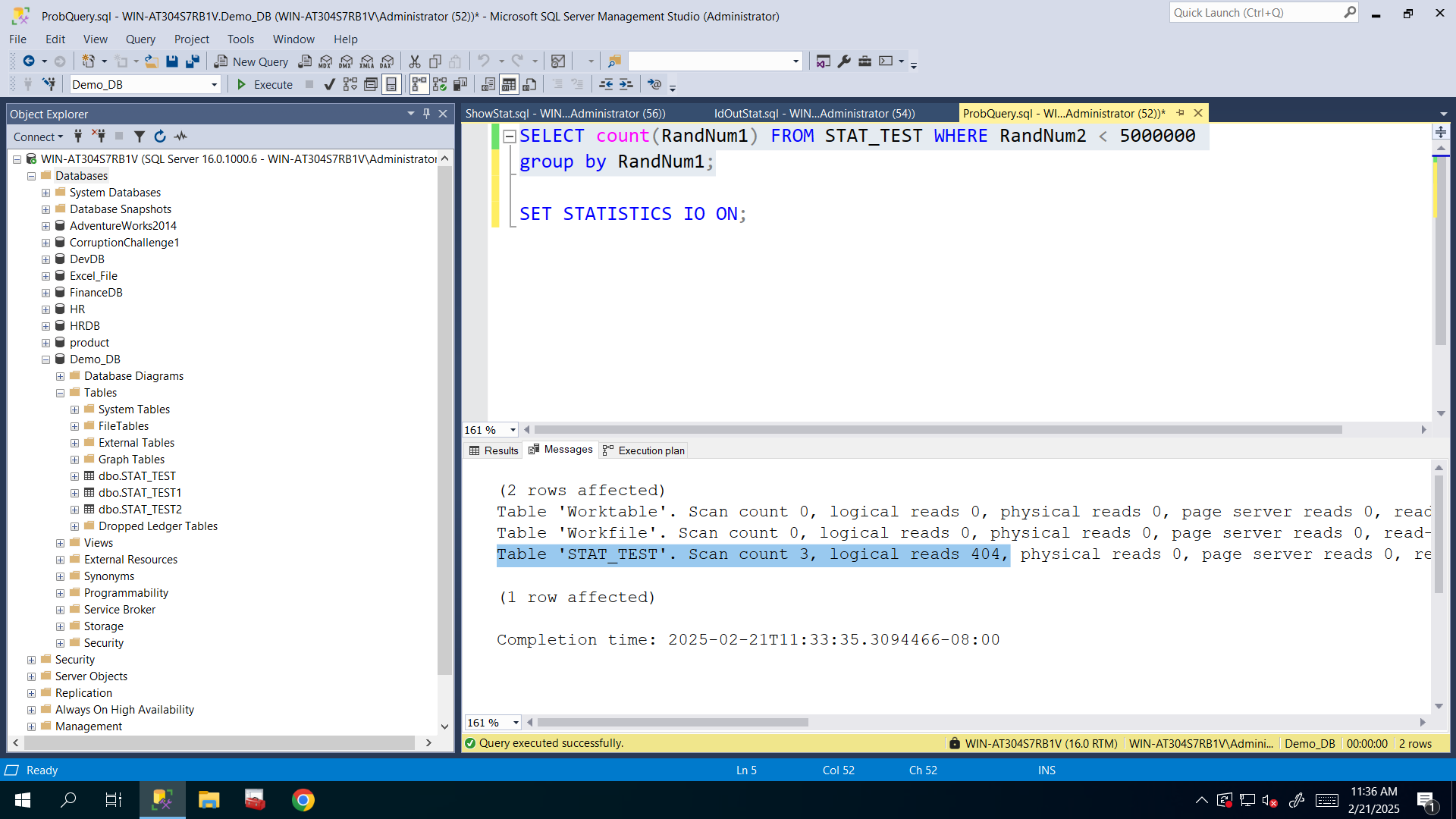
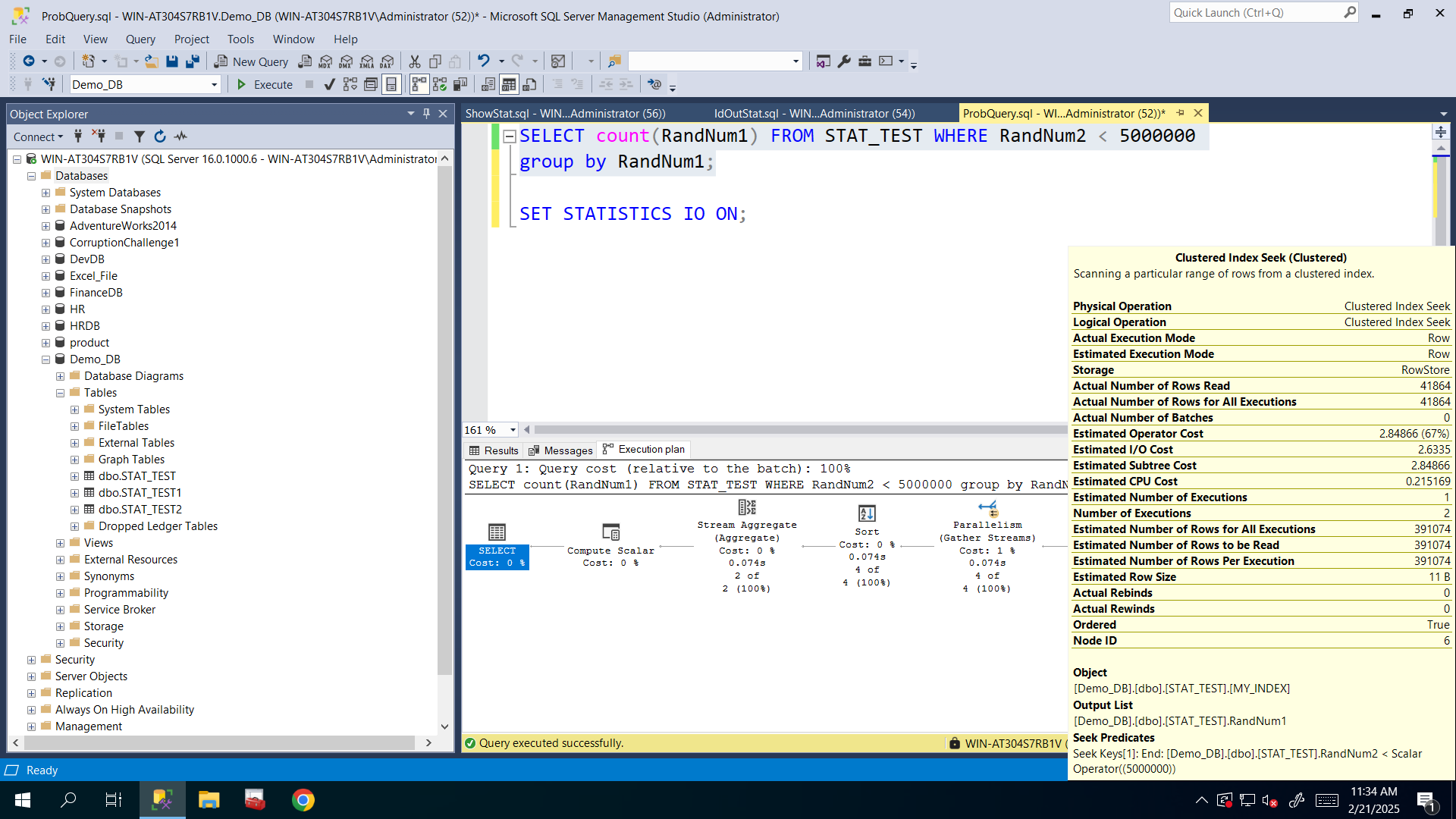
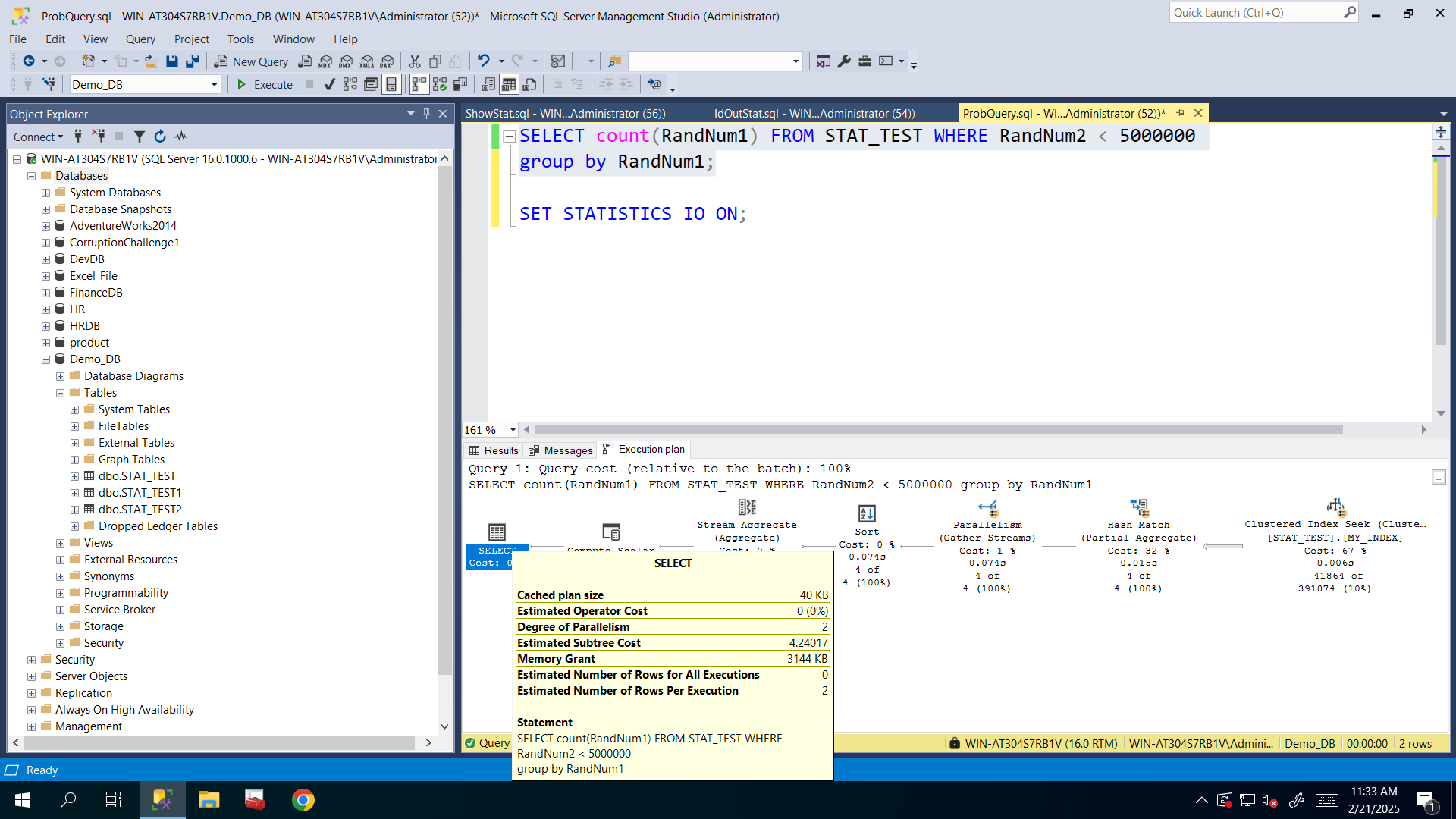
GO

**After executing above query, as you can see, statistics keep 120 rows information.**



**To see scan counts and logical reads, I run the following query after seeing it in the clues:**

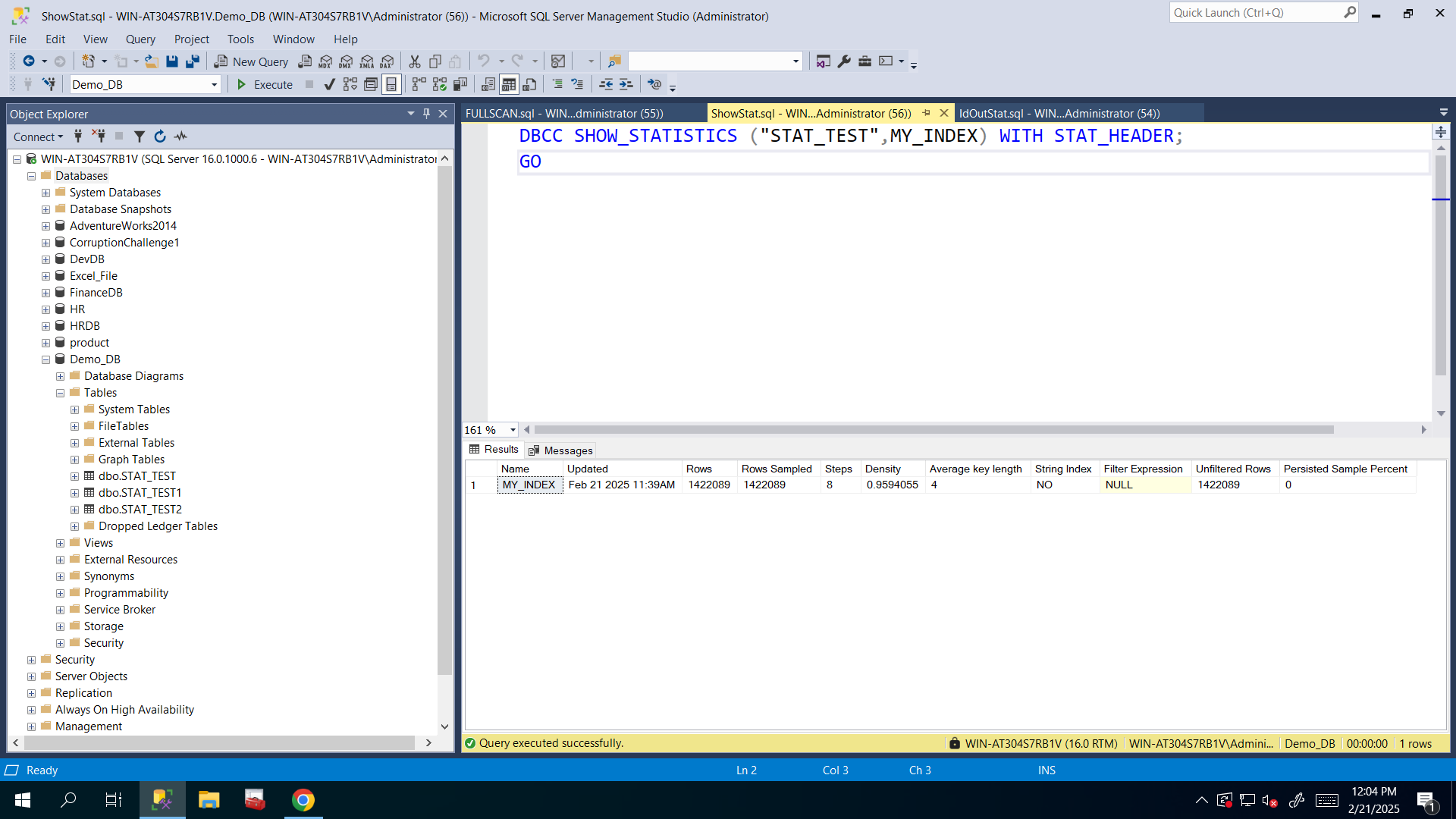
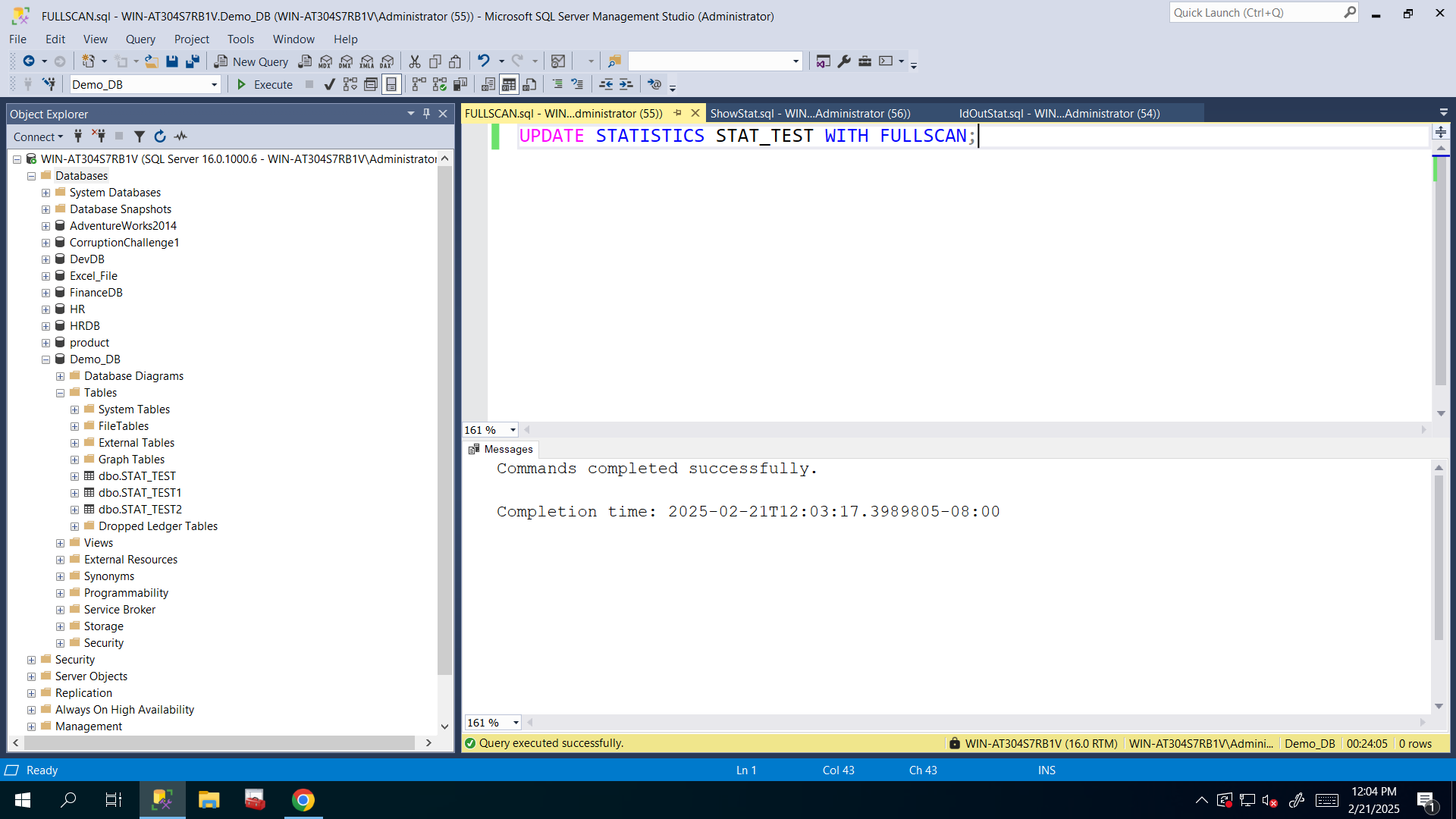
Run the provided query and fill up the following table based on the execution plan:



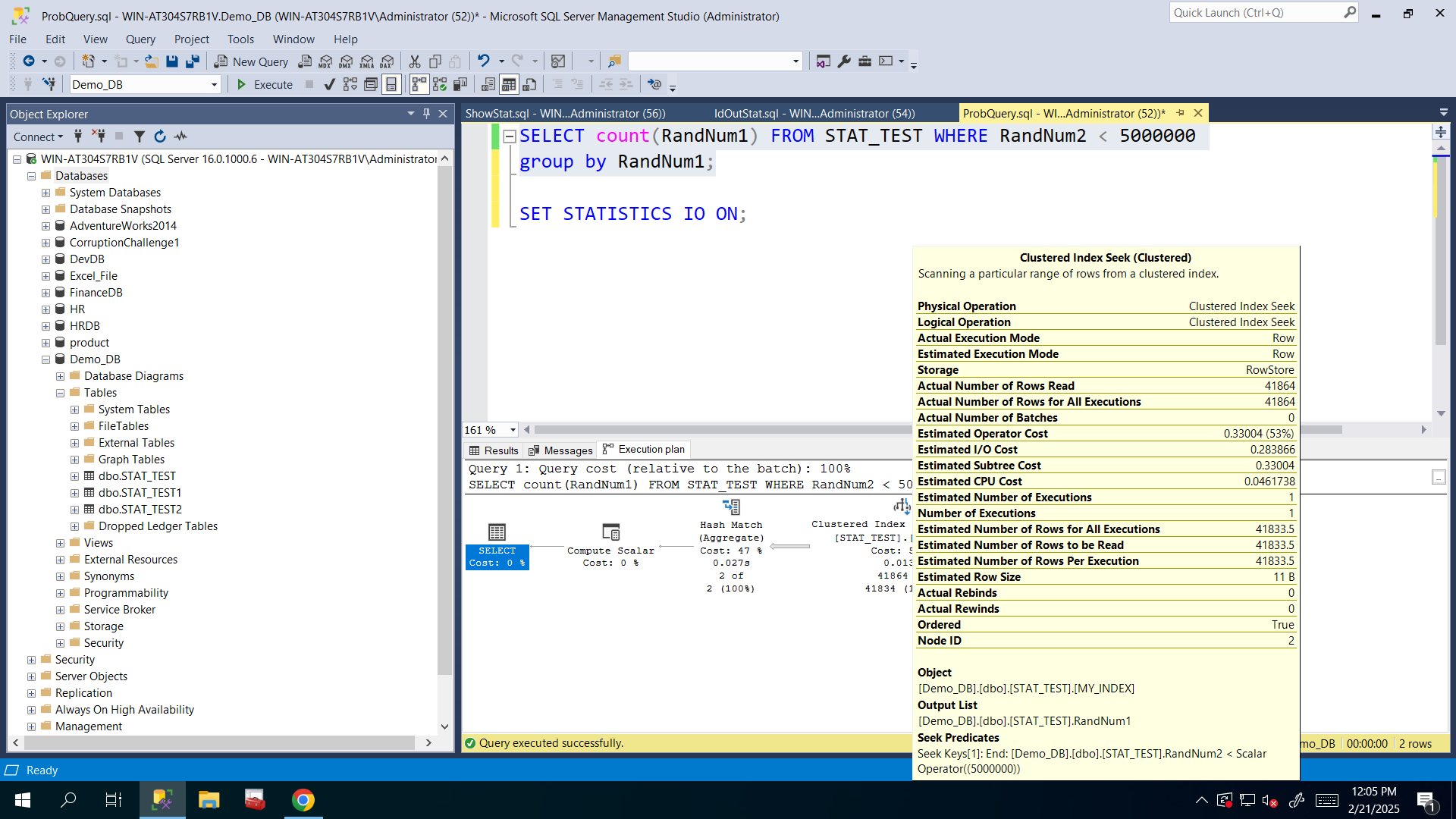
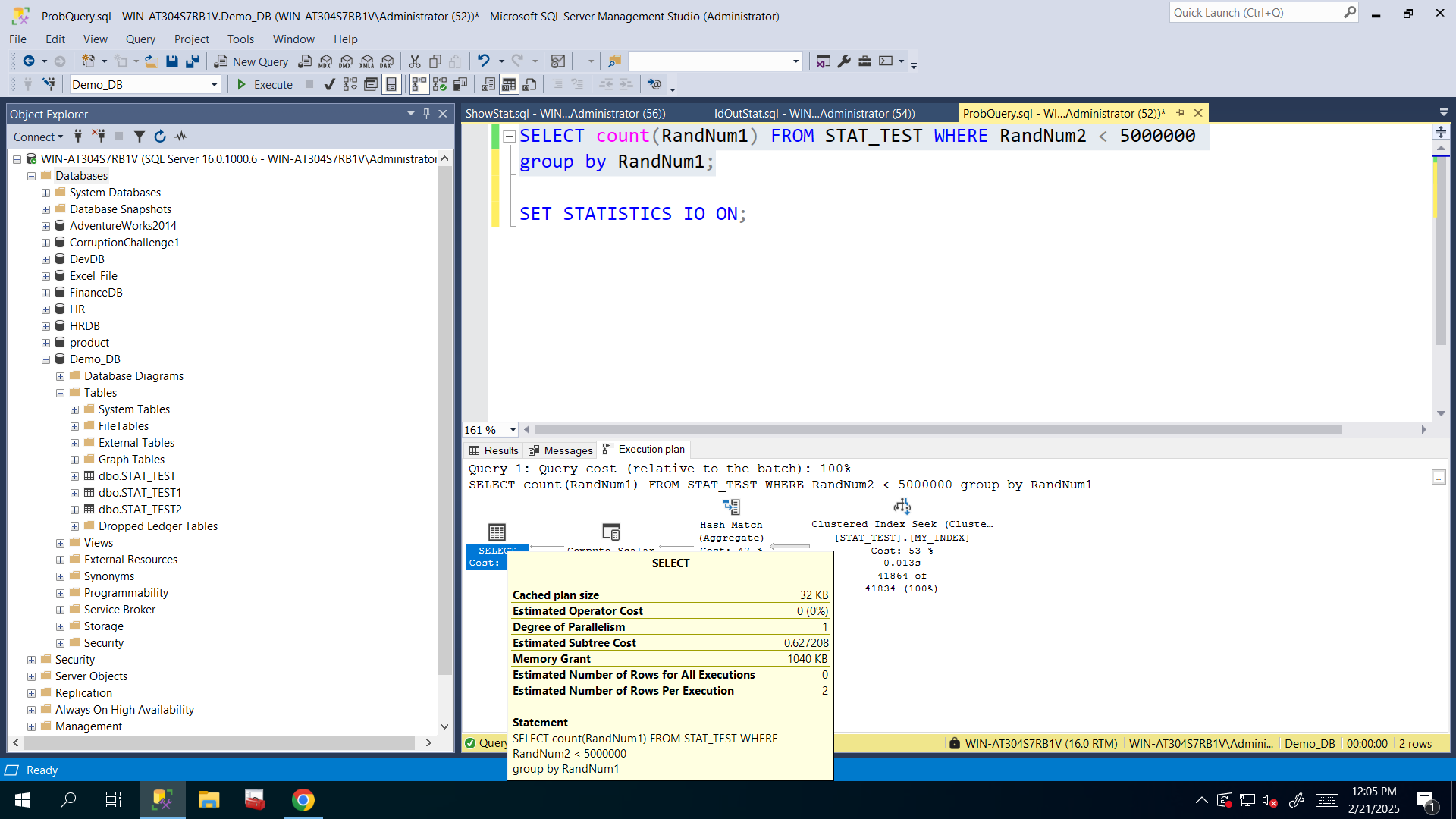
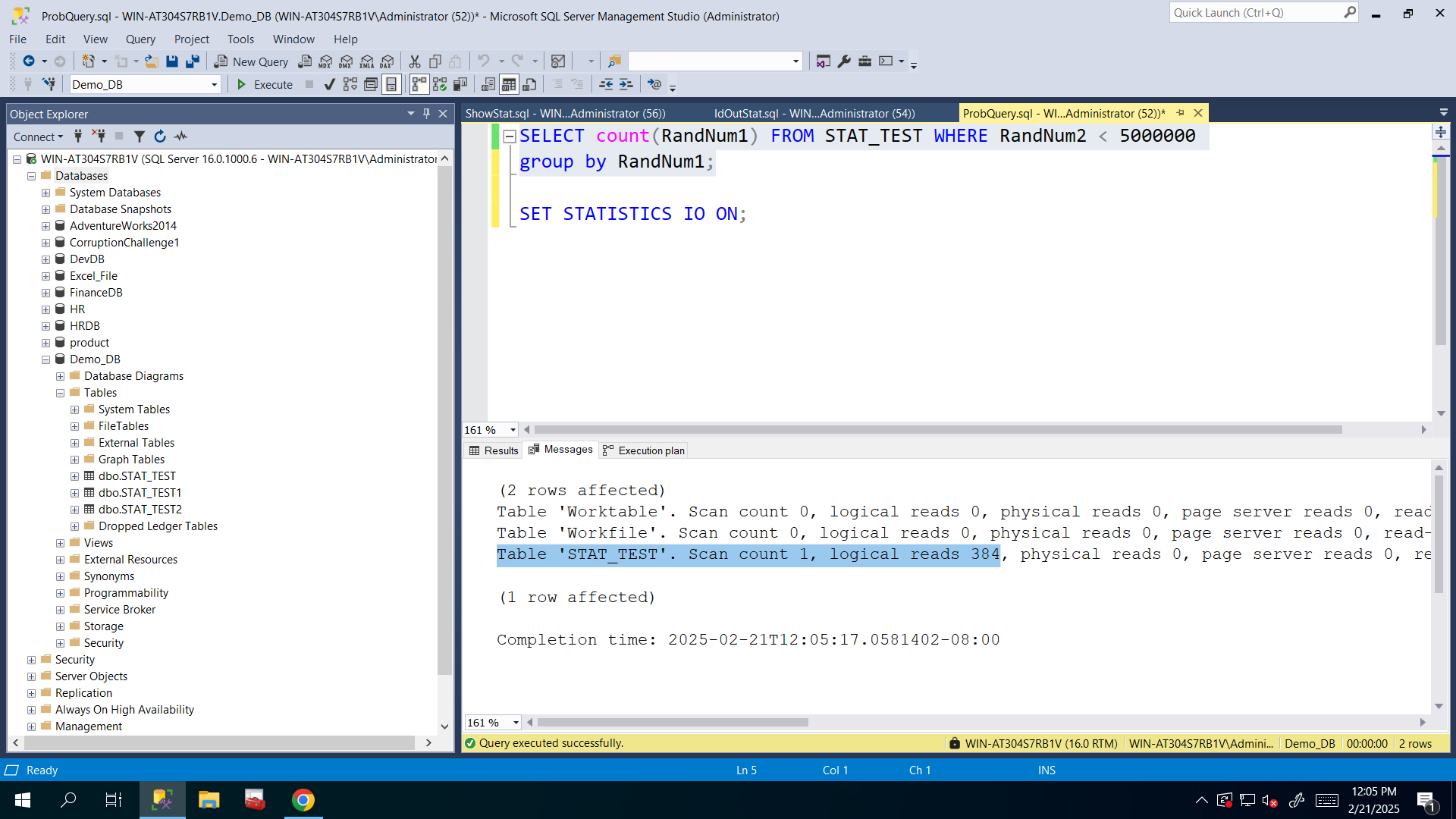
|  |  |
| --- | --- |
| **Property Name** | **Changes before updating statistics** |
| Estimated Subtree Cost for the query (no measure unit) | **4.24017** |
| Memory Grant for the query (for performing sort and join operations, for example here memory grant is used by GROUP BY) (in KB) | **3144** |
| Logical reads for the query (in pages) | **404** |
| Scan count for the query (in times) | **3** |
| Estimated and Actual Number of rows for the index | **391074 (Estimated)**  **41864 (Actual)** |

1. Update outdated statistics with FULLSCAN

**I updated STAT\_TEST table’s statistics and checked to make sure it’s updated correctly:**



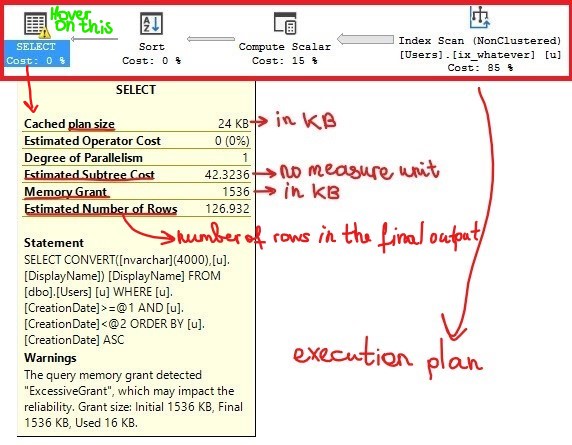
1. Run the provided query again and fill up the following table based on the new execution plan:



|  |  |
| --- | --- |
| **Property Name** | **Changes after updating statistics** |
| Estimated Subtree Cost for the query (no measure unit) | **0.627208** |
| Memory Grant for the query (for performing sort and join operations, for example here memory grant is used by GROUP BY) (in KB) | **1040** |
| Logical reads for the query (in pages) | **384** |
| Scan count for the query (in times) | **1** |
| Estimated and Actual Number of rows for index | **41833.5 ~ 41834 (Estimated)**  **41864 (Actual)**  **That’s the result of updating statistics. I need to update them otherwise they will differ from each other a lot. After updating, numbers are very close to each other (estimated and actual).** |

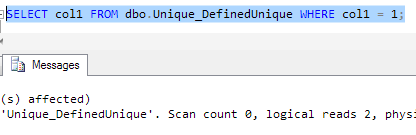
CLUES

1. For Estimated Subtree Cost, Memory Grant use pay attention to the following things (this is just an example):



1. For Scan count, Logical Reads run the following command and switch to the Messages tab as shown below:

**SET STATISTICS IO ON**



1. Actual Number of rows are known only after query is executed. It shows you what happened during the query execution. For showing Actual Number of rows, you need to enable Actual Execution Plan and then run your query:

