

Database Indeksering 2.HF

It & Data, Odense

Indeksering af en database

- Bruges til optimering af søge resultater
- Er både godt og skidt:
 - Kun kolonner, der hyppigt søges på (og som IKKE er en primær nøgle) kan med fordel indgå i et indeks.
 - Bivirkningen ved et indeks er, at det koster tid ved opdatering, da det også skal opdateres
- Der er to typer af indeks
 - Clustered, styre den rækkefølge data gemmes i
 - Nonclustered, heap-style, arbejder med "row locator"
- Primær nøgler (Primary key)
 - Generer automatisk et indeks af typen: Clustered
- Unik (Unique)
 - Generer automatisk et indeks af type: Nonclustered

Læs Microsofts egen beskrivelse her:

<https://docs.microsoft.com/en-us/sql/relational-databases/indexes/clustered-and-nonclustered-indexes-described?view=sql-server-2017#how-indexes-are-used-by-the-query-optimizer>

Indeks: SQL statement

Syntaks	Eksempel
CREATE INDEX <i>index_name</i> ON <i>table_name</i> (<i>column1</i> , <i>column2</i> , ...);	CREATE INDEX idx_lastname ON Persons (LastName);
	CREATE INDEX idx_pname ON Persons (LastName, FirstName);
CREATE UNIQUE INDEX <i>index_name</i> ON <i>table_name</i> (<i>column1</i> , <i>column2</i> , ...);	CREATE UNIQUE INDEX idx_email ON Persons (Email);
DROP INDEX <i>table_name.index_name</i> ;	DROP INDEX Persons.idx_pname;

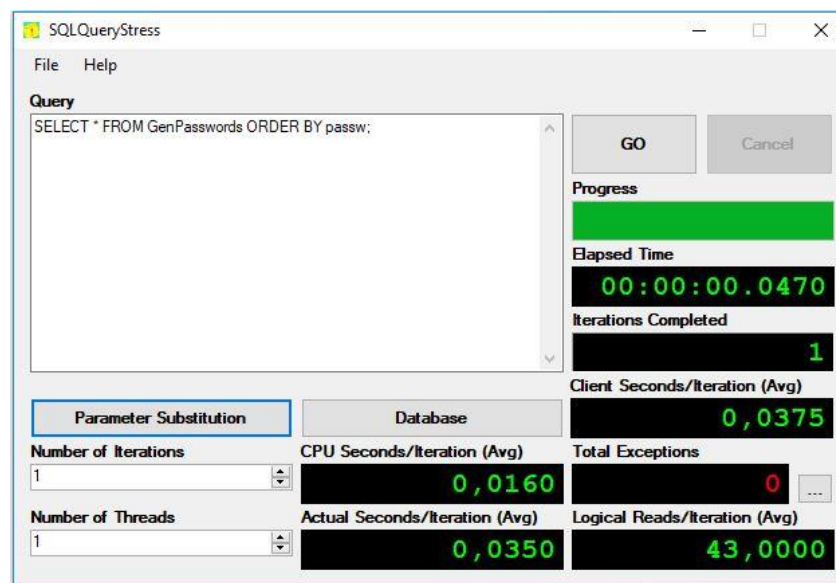
https://www.w3schools.com/sql/sql_create_index.asp

Opgave: Performance test

- I skal kreere en database til testen med to næsten ens tabeller.
- Den ene tabel skal have et felt med et indeks (ikke primær nøglen).
- Den anden tabel har et tilsvarende felt uden indeks.
- Begge tabeller skal fyldes med data (minimum 10.000 rækker, gerne flere).
- Test forskellige Queries på begge tabeller og dokumenter dine resultater: F.eks.
 - `SELECT * FROM tbl1 (og tbl2)`
 - `SELECT * FROM tbl1 ORDER BY col1 (med indeks)`
 - `SELECT * FROM tbl2 ORDER BY col1 (uden indeks)`
 - `SELECT * FROM tbl1 WHERE col1 ...`
- Test eventuelt på en tabel med en ekstra kolonne med indeks men hvor der ikke søges på feltet – gør det en forskel med det ekstra indeks?

Performance målinger

- Der er rigtig mange måde at foretage disse målinger på
- Indbyggede features i MS SQL serveren
- Massere af programmer, der kan tilkøbes
- Mit valg faldt på dette program: SQLQueryStress
 - Det er gratis
 - Simpelt og nemt at bruge

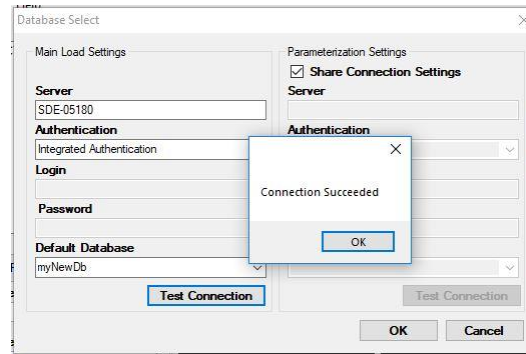


Kan hentes via linket:

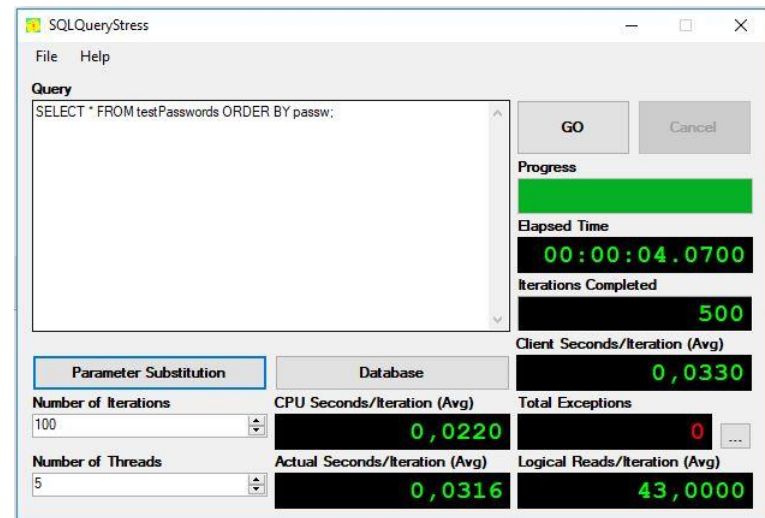
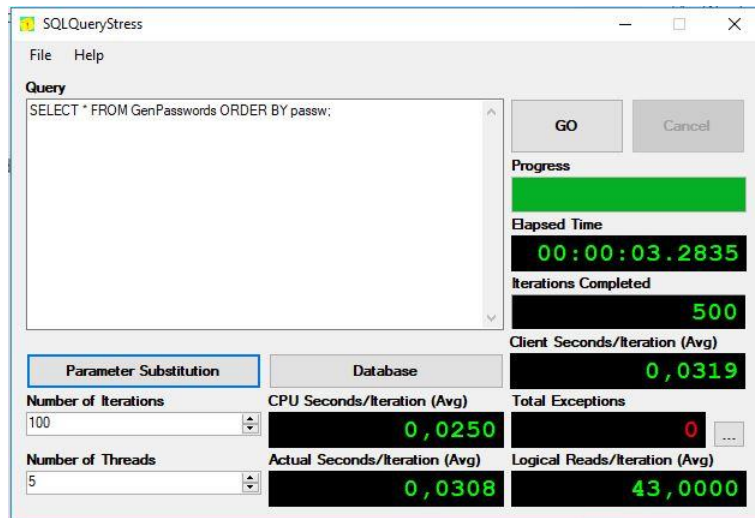
<https://www.itsupportguides.com/knowledge-base/sql-server/sql-quick-and-easy-way-to-performance-test-sql-statements/>

SQLQueryStress Test

- Nem at sætte op



- Hurtigt og nemt at teste med flere iterationer og tråde
- Nemt at sammenligne de forskellige resultater



Masser af data – import data

Der er stadig færdige database med indhold, der kan hentes på nettet, som f.eks.:

- AdventureWorks sample database
- WideWorldImporters sample database

Link:

<https://docs.microsoft.com/en-us/sql/samples/sql-samples-where-are?view=sql-server-2017>

Masser af data

Der er flere metoder til at oprette massevis af data

- Tabellen kan kreeres og fyldes sådan:

```
CREATE TABLE [Location] (  
    [id] INT IDENTITY(1,1),  
    [Date] DATETIME DEFAULT GETDATE (),  
    [City] CHAR (75) DEFAULT 'Odense');  
  
-- Insert 10.000 random values  
INSERT INTO Location DEFAULT VALUES ;  
GO 10000  
GO
```

- Problemet er bare, at alle data er næsten ens – Tidspunktet for datoen vil selvfølgelig variere lidt.

En smart måde at få data genereret via kode finder du her:

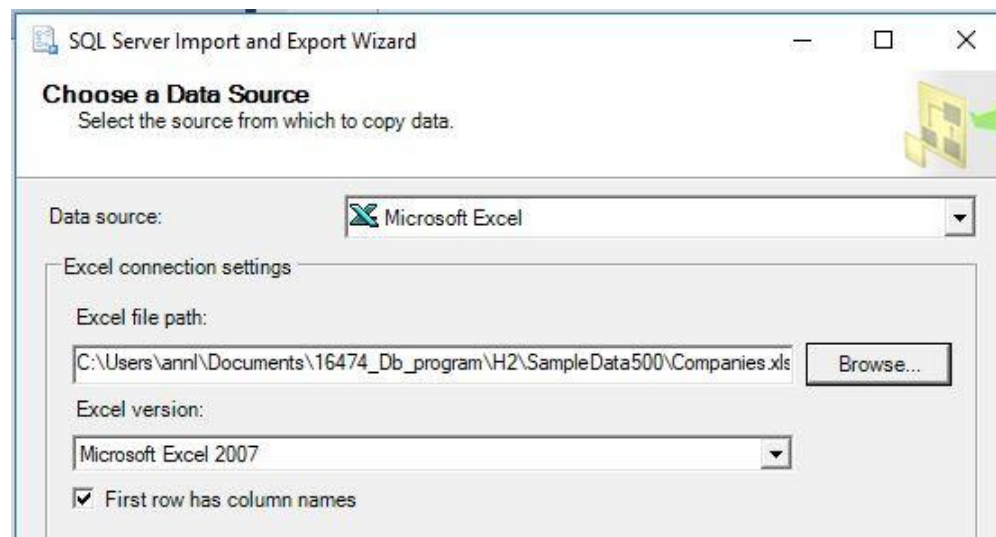
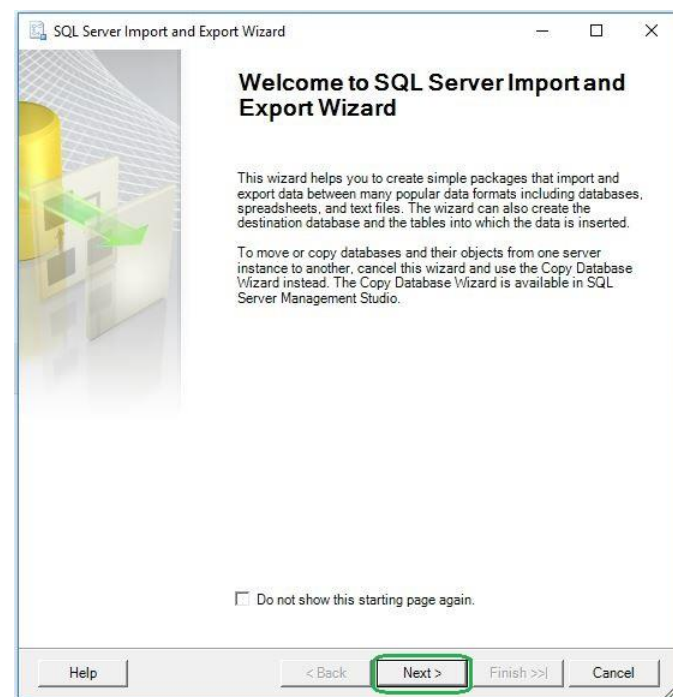
- <https://www.mssqltips.com/sqlservertip/5303/simple-sql-server-function-to-generate-random-8-character-password/>
- Prøv at se om du kan implementere den kode, Eli demonstrer på siden.

Note: Jeg vil ikke anbefale ikke-krypterede koder, men som en tilfældighedsgenerator er den fin.

Masser af data via Excel ark

Mini-guide:

1. Opret eventuelt en tabel først
2. Højre-klik på databasen -> Tasks -> Import data
3. Data source = Excel ark (2007) -> Browse for at lokalisere fil (tjek "First row has column names" hvis det er sandt).



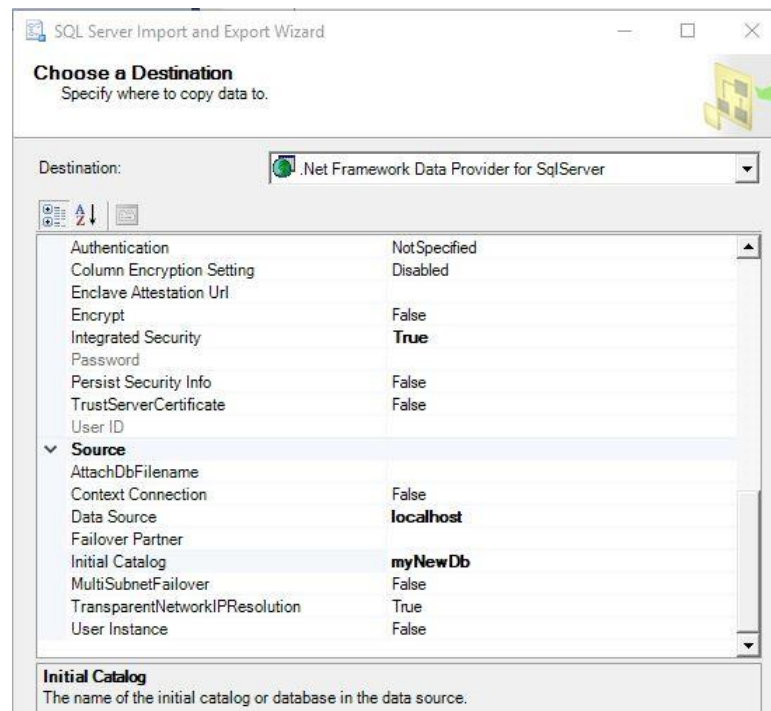
Masser af data via Excel ark

Forsættelse på mini-guide:

4. Destination = SqlServer

- Integrated Security = True
- Data source = localhost (127.0.0.1)
- Initial catalog = din database

5. Tjek “Copy data from one or more tables or views”



Masser af data via Excel ark

6. Destination: Vælg korrekt tabel eller opret en ny

- Edit mappings: Tjek at formaterne er korrekte (INT, VARCHAR osv.)

Source: 'Company\$'
Destination: 'dbo.'company2'

☐ Create destination table
☐ Delete rows in destination table ☐ Drop and re-create destination table
☒ Append rows to the destination table ☐ Enable identity insert

Mappings:

Source	Destination	Type	Nullable	Size	Precision	Scale
id	<ignore>					
first_name	first_name	nvarchar	✓	100		
last_name	last_name	nvarchar	✓	100		
company_name	company_name	nvarchar	✓	250		
address	address	nvarchar	✓	100		
city	city	nvarchar	✓	100		
county	county	nvarchar	✓	30		
state	states	nvarchar	✓	3		
post	post	nvarchar	✓	15		
phone1	phone1	nvarchar	✓	12		
phone2	phone2	nvarchar	✓	12		
email	email	nvarchar	✓	100		
web	web	nvarchar	✓	100		

Source column: id Double (15)

SQL Server Import and Export Wizard

Specify Table Copy or Query
Specify whether to copy one or more tables and views or to copy the results of a query from the data source.

☒ **Copy data from one or more tables or views**
Use this option to copy all the data from the existing tables or views in the source database.

Evt. preview source

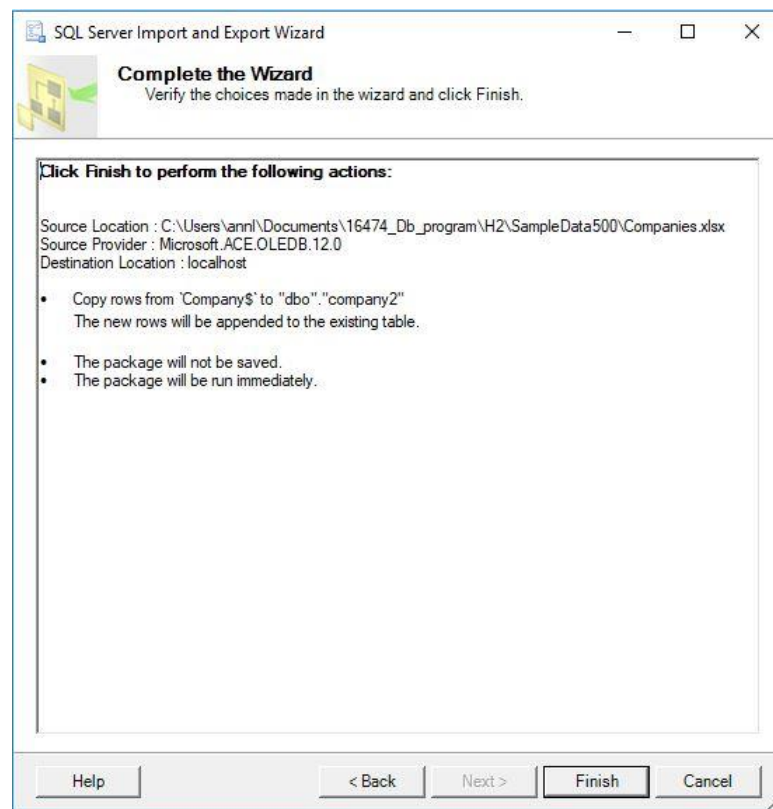
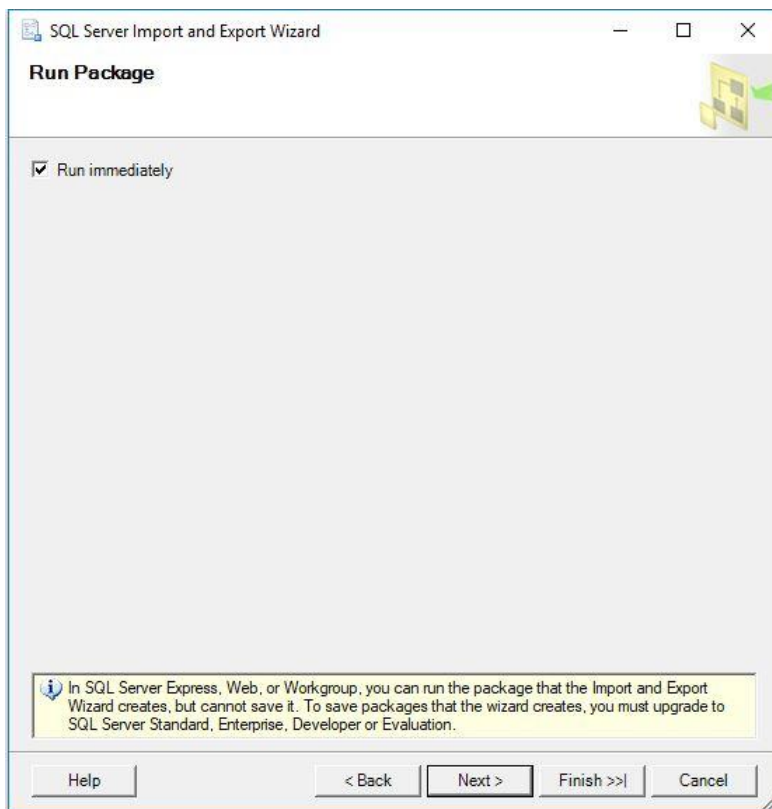
Preview Data

Source: `SELECT * FROM Company$`

id	first_name	last_name	company_name	address	city
1	Rebecca	Didio	Brandt, Jonathan F Esq	171 E 24th St	L
2	Stevie	Hallo	Landrum Temporary Services	22222 Acoma St	P
3	Mariko	Stayer	Inabinet, Macre Esq	534 Schoenbom St #51	H
4	Gerardo	Woodka	Morris Downing & Sherred	69206 Jackson Ave	T
5	Mayra	Bena	Buel, David L Esq	808 Glen Cove Ave	L
6	Idella	Scotland	Artesian Ice & Cold Storage Co	373 Lafayette St	C
7	Sherill	Klar	Midway Hotel	87 Sylvan Ave	N
8	Ena	Desjardiws	Selsor, Robert J Esq	60562 Ky Rt 321	B
9	Vince	Siena	Vincent J Petti & Co	70 S 18th Pl	P

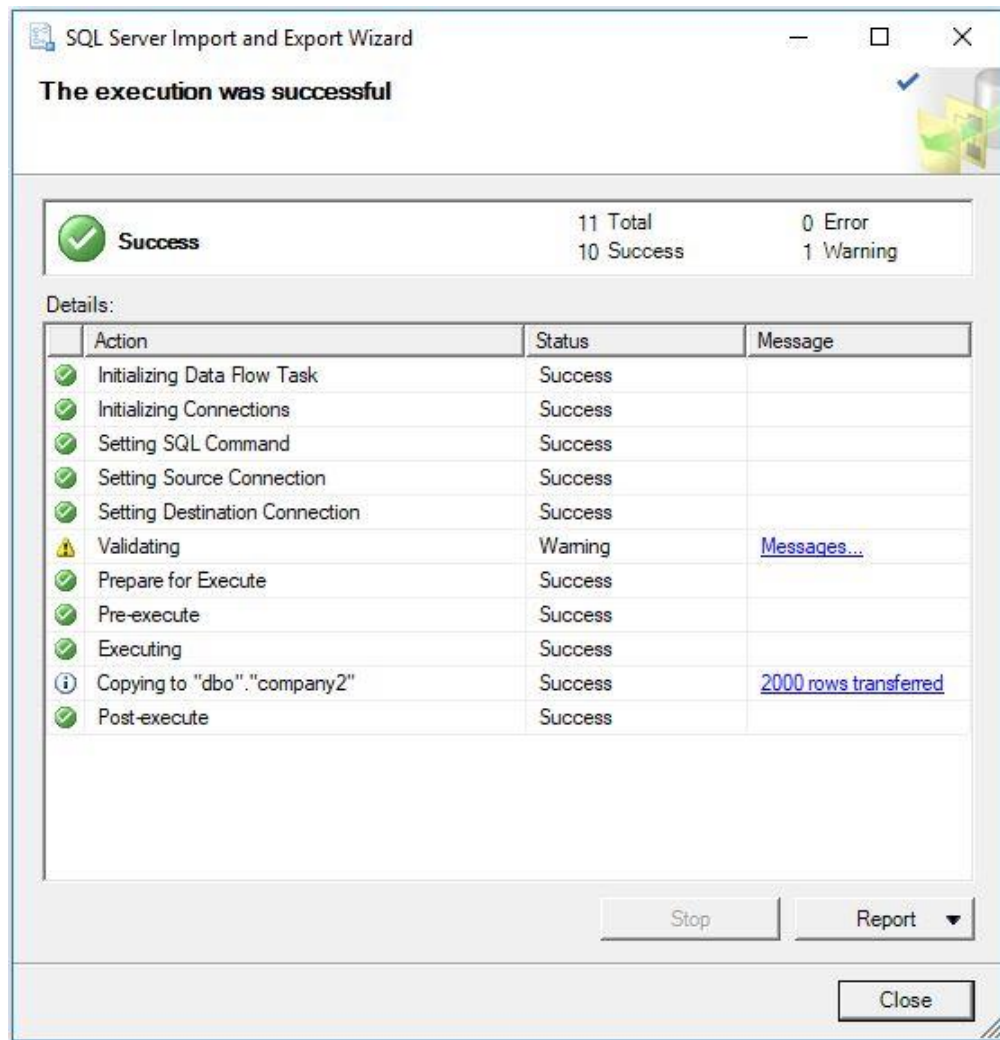
Masser af data via Excel ark

7. Kør med det samme



Masser af data via Excel ark

8. Resultatet skal helst være noget a la det her:



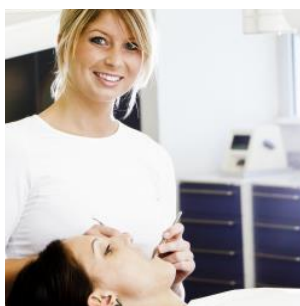
**Hvis du ønsker at teste din viden
yderligere, så tag testen:**

<https://use-the-index-luke.com/3-minute-test/sql-server>

**Læs eventuelt, hvad Markus skriver om
performance-testing**

<https://use-the-index-luke.com/sql/example-schema/sql-server/performance-testing-scalability>

Slut på den praktiske opgave.



Campus M – Vi uddanner Danmarks dygtigste