

Angela Henry



Data Types Do Matter

Angela Henry

Angela is a DBA/BI Developer, living in High Point, NC and loves what she does. She's worked with all versions of SQL Server & worn all the hats that come with dealing with SQL Server throughout the years: developer, administrator, data architect and BI developer.

MSCE: Business Intelligence

Microsoft Data Platform MVP

She is also the local group leader for the PASS Triad SQL User Group in Greensboro, NC.

Tweets at @SQLSwimmer

Blogs at SQLSwimmer.wordpress.com

In her spare time you can probably find her in or at the pool, she's an avid US Masters & USA Swimming Swimmer, Coach and Instructor.



Overview

Data Quality

Storage

Performance



SQL Server Data Types

user-defined data types

sql_variant

xml

datetimeoffset

datetime2

datetime

smalldatetime

date

time

float

real

decimal

money

smallmoney

bigint

int

smallint

tinyint

bit

ntext

text

image

timestamp

uniqueidentifier

nvarchar (including nvarchar(max))

nchar

varchar (including varchar(max))

char

varbinary (including varbinary(max))

binary



Data Quality

Character Data Types - Everyone's favorite

- For Temporal Data (Dates/Times)
- For Numbers



A large, stylized teal graphic on the left side of the slide, consisting of several overlapping, rounded rectangular shapes that create a sense of depth and movement.

Demo - Data Quality

Data Quality

Comparison Operators

Sorting

Garbage In = Garbage Out



Storage

They're just numbers, right?!

It's just string data, who cares?

A date's a date



Just Numbers

Data type	Range	Storage
bigint	-2^{63} to $2^{63}-1$ -9,223,372,036,854,775,808 to 9,223,372,036,854,775,807 (6 commas!!)	8 Bytes
int	-2^{31} to $2^{31}-1$ (-2,147,483,648) to (2,147,483,647)	4 Bytes
smallint	-2^{15} to $2^{15}-1$ (-32,768) to (32,767)	2 Bytes
tinyint	0 to 255	1 Byte



A large, stylized teal graphic element on the left side of the slide, consisting of several overlapping, rounded rectangular shapes that create a sense of depth and movement, resembling a stylized letter 'D' or a series of connected loops.

Demo – Just Numbers

Not Just Numbers

Data type	Range	Storage	Impact
bigint	-2^{63} to $2^{63}-1$ -9,223,372,036,854,775,808 to 9,223,372,036,854,775,807 (6 commas!!)	8 Bytes	200% > int 400% > smallint 800% > tinyint
int	-2^{31} to $2^{31}-1$ (-2,147,483,648) to (2,147,483,647)	4 Bytes	200% > smallint 400% > tinyint
smallint	-2^{15} to $2^{15}-1$ (-32,768) to (32,767)	2 Bytes	200% > tinyint
tinyint	0 to 255	1 Byte	



Just String Data

Data Type	Range	Storage
Char	Fixed Length 1 to 8000	Length of data
Varchar	Variable Length 1 to 8000 or MAX → $2^{31} - 1$ Bytes (2GB)	Length of data + 2 Bytes
nChar	Fixed Length 1 to 4000	2 * Length of data
nVarchar	Variable Length 1 to 4000 or MAX → $2^{31} - 1$ Bytes (2GB)	2 * Length of data + 2 Bytes





Demo – Just String Data

Not Just String Data

Data Type	Range	Storage	Impact
Char	Fixed Length 1 to 8000	Length of data	200% > varchar with same lengths
Varchar	Variable Length 1 to 8000 or MAX --> $2^{31} - 1$ Bytes (2GB)	Length of data + 2 Bytes	
nChar	Fixed Length 1 to 4000	2 * Length of data	200% > nvarchar with same lengths
nVarchar	Variable Length 1 to 4000 Or MAX → $2^{31} - 1$ Bytes (2GB)	2 * Length of data + 2 Bytes	



A Date's A Date

Data Type	Range	Storage	Accuracy
Datetime2	0001-01-01 thru 9999-12-31	6-8 Bytes	100 Nanoseconds
Datetime	1753-01-01 thru 9999-12-31	8 Bytes	Milliseconds rounded to .000, .003 or .007
SmallDatetime	1900-01-01 thru 2079-06-06	4 Bytes	1 Minute
Date	0001-01-01 thru 9999-12-31	3 Bytes	1 Day





Demo – A Date's a Date

Not All Dates Are Created Equal

Data Type	Range	Storage	Accuracy	Impact
Datetime2	0001-01-01 thru 9999-12-31	6-8 Bytes	100 Nanoseconds	Up to 230% > Date
Datetime	1753-01-01 thru 9999-12-31	8 Bytes	Milliseconds rounded to .000, .003 or .007	230% > Date
SmallDatetime	1900-01-01 thru 2079-06-06	4 Bytes	1 Minute	30% > Date
Date	0001-01-01 thru 9999-12-31	3 Bytes	1 Day	



And The Winner Is.. ~ 10MM Rows

Data Types	Data Only	Data With Index
Date → Datetime2	50MB	100MB
TinyInt → BigInt	70MB	140MB
Varchar(255) → Char(255)	2.5GB	5GB



Performance

Execution Plans

I/O



Execution Plans

SARGability - 'SARG' stands for **S**earch **ARG**ument, and it *means* that the predicate can be executed using an index seek

Predicate is <expression> <operator> <expression>

- DateAdd(day, 1, myDateCol) = @myDate is NOT SARGable
- myDateCol = @myDateVariable IS SARGable when they are the same data type



Conversion Precedence

1 user-defined data types (highest)	11 real	21 text
2 sql_variant	12 decimal	22 image
3 xml	13 money	23 timestamp
4 datetimeoffset	14 smallmoney	24 uniqueidentifier
5 datetime2	15 bigint	25 nvarchar (including nvarchar(max))
6 datetime	16 int	26 nchar
7 smalldatetime	17 smallint	27 varchar (including varchar(max))
8 date	18 tinyint	28 char
9 time	19 bit	29 varbinary (including varbinary(max))
10 float	20 ntext	30 binary (lowest)



A decorative graphic on the left side of the slide, consisting of several overlapping, curved teal shapes that resemble stylized, rounded lines or a partial circular pattern.

Demo – Performance Execution Plans

Fat Tables

```
CREATE TABLE Demo.SuppliersOriginal(  
    SupplierID int NOT NULL,  
    SupplierName nvarchar(100) NOT NULL,  
    SupplierCategoryID int NOT NULL,  
    PrimaryContactPersonID int NOT NULL,  
    AlternateContactPersonID int NOT NULL,  
    DeliveryMethodID int NULL,  
    DeliveryCityID int NOT NULL,  
    PostalCityID int NOT NULL,  
    SupplierReference nvarchar(20) NULL,  
    BankAccountName nvarchar(50) NULL,  
    BankAccountBranch nvarchar(50) NULL,  
    BankAccountCode nvarchar(20) NULL,  
    BankAccountNumber nvarchar(20) NULL,  
    BankInternationalCode nvarchar(20) NULL,  
    PaymentDays int NOT NULL,  
    InternalComments nvarchar(max) NULL,  
    PhoneNumber nvarchar(20) NOT NULL,  
    FaxNumber nvarchar(20) NOT NULL,  
    WebsiteURL nvarchar(256) NOT NULL,  
    DeliveryAddressLine1 nvarchar(60) NOT NULL,  
    DeliveryAddressLine2 nvarchar(60) NULL,  
    DeliveryPostalCode nvarchar(10) NOT NULL,  
    PostalAddressLine1 nvarchar(60) NOT NULL,  
    PostalAddressLine2 nvarchar(60) NULL,  
    PostalPostalCode nvarchar(10) NOT NULL,  
    LastEditedBy int NOT NULL)
```

```
CREATE TABLE Demo.SuppliersFat(  
    SupplierID int NOT NULL,  
    SupplierName char(255) NOT NULL,  
    SupplierCategoryID int NOT NULL,  
    PrimaryContactPersonID int NOT NULL,  
    AlternateContactPersonID int NOT NULL,  
    DeliveryMethodID int NULL,  
    DeliveryCityID int NOT NULL,  
    PostalCityID int NOT NULL,  
    SupplierReference char(255) NULL,  
    BankAccountName char(255) NULL,  
    BankAccountBranch char(255) NULL,  
    BankAccountCode char(255) NULL,  
    BankAccountNumber char(255) NULL,  
    BankInternationalCode char(255) NULL,  
    PaymentDays int NOT NULL,  
    InternalComments char(2000) NULL,  
    PhoneNumber char(25) NOT NULL,  
    FaxNumber char(255) NOT NULL,  
    WebsiteURL char(256) NOT NULL,  
    DeliveryAddressLine1 char(255) NOT NULL,  
    DeliveryAddressLine2 char(255) NULL,  
    DeliveryPostalCode char(255) NOT NULL,  
    PostalAddressLine1 char(255) NOT NULL,  
    PostalAddressLine2 char(255) NULL,  
    PostalPostalCode char(255) NOT NULL,  
    LastEditedBy int NOT NULL)
```



A large, stylized teal graphic on the left side of the slide, consisting of several overlapping, rounded, parallel lines that curve from the top-left towards the bottom-left, resembling a stylized letter 'D' or a decorative flourish.

Demo – Performance IO

Impact on Performance

Implicit Data Type Conversions

Table Scans/Index Scans

More Pages Mean More IO



References

[How Can Data-Type Choice Effect Performance by Paul Randal](#)

[What is a SARGible Predicate by Gail Shaw](#)

[Varchar\(max\) forum Thread](#)

SQL Server Documentation



Contact Information

Twitter: @SQLSwimmer

Blog: sqlswimmer.wordpress.com

Linked In: [in/angelahenrydba](https://www.linkedin.com/in/angelahenrydba)

Email: henry.angela@gmail.com

