Dipping Your Toes Into the SQL Pond

A *hands-on* follow-up to "SQL for GIS PPL"
Mike Dolbow| GIS Coordinator

MINNESOTA IT SERVICES

- Some repeat info from "SQL for GIS PPL" early on (~10 minutes)

- I am NOT a Database Administrator (DBA)

- I am NOT a SQL (or spatial SQL) expert

RERUNS?

TIME FOR A BIO BREAK

- What is SQL and what does it mean for a "GIS person"?

- Why / When use a relational/enterprise database?

- Hands-On exercise with SQL Examples

- This WILL be a "technical" session (but I'll go slow with the hands-on; don't be afraid to try!). It's for you if:

  - You don't mind the lack of pretty pictures

  - Saving time and reducing duplication gets you excited

- SQL = "Structured Query Language"

- Wikipedia: "a domain-specific language used in programming and designed for managing data held in a relational database management system (RDBMS)…
  It is particularly useful in handling structured data, i.e. data incorporating relations among entities and variables."

- "Structured data" = Tables! (A layer is just a table with a geometry column.)

Not as scary as it might seem

You've likely been defining WHERE clauses in ArcGIS Desktop and ArcGIS Pro for a while

Select by Attributes

Enter a WHERE clause to select records in the table window.

Method : Create a new selection

SCHNAME
ADDRESS
CITY
ZIP
SCH_TYPE

=	< >	Like	'Clarkfield'
>	> =	And	'Clear Lake'
<	< =	Or	'Clearbrook'
_	%	()	Not	'Clearwater'
			'Cleveland'

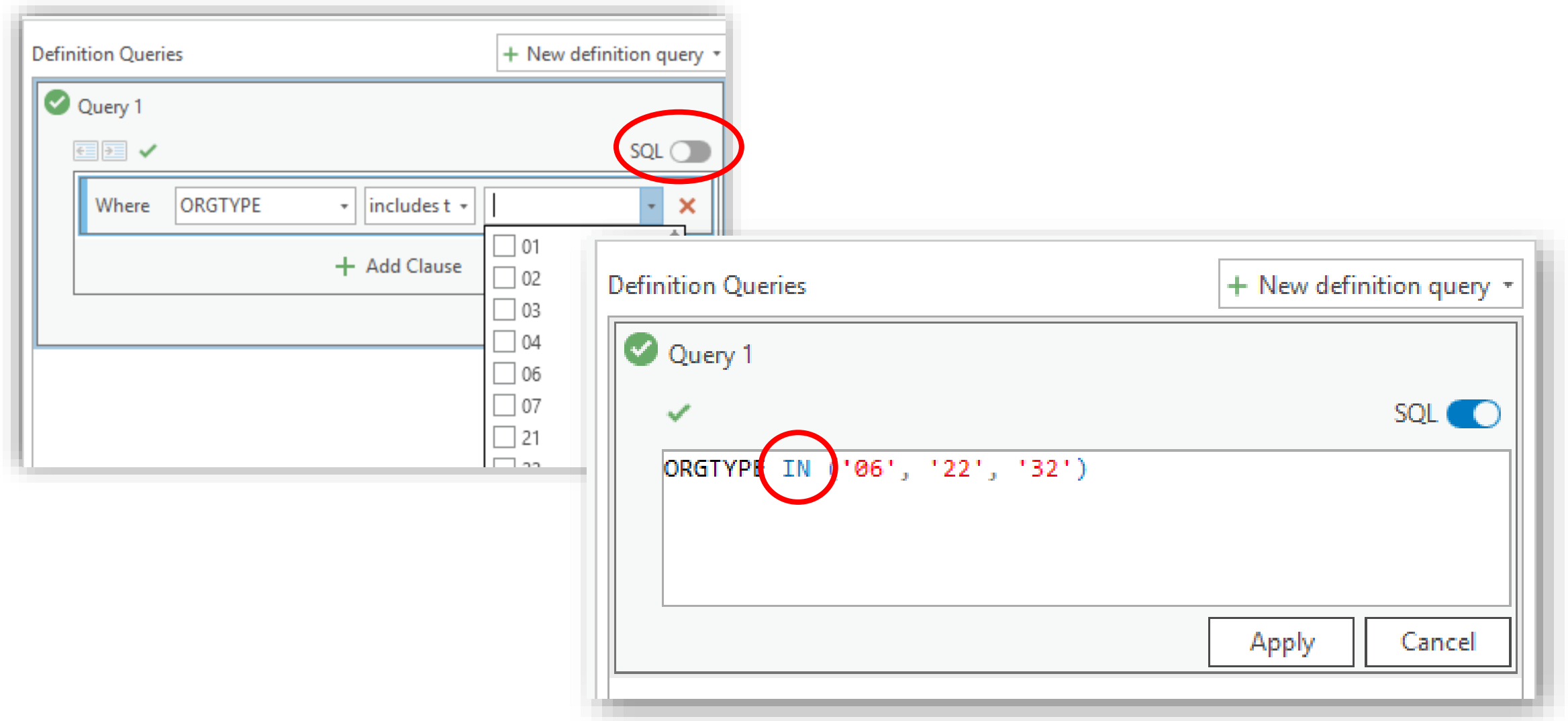Is In Null Get Unique Values Go To:

SELECT * FROM sde.SDE.strds_school_program_locs_base WHERE:

CITY = 'Clearwater'

Clear Verify Help Load... Save...

Apply Close

Query: School District Boundaries (ID: 0)

Where:

You can even use SQL tricks within Arc

Why use databases? Why is SQL an important skill?

Health policy

Covid: how Excel may have caused loss of 16,000 test results in England

Public Health England data error blamed on limitations of Microsoft spreadsheet

- **Coronavirus – latest updates**
- **See all our coronavirus coverage**

Alex Hern *UK technology editor*

🐦 @alexhern

Tue 6 Oct 2020 03.21 EDT

f 🐦 ✉

2,622

Pete Timmins 🇪🇺 @petertimmins3 · Oct 5

16,000 covid cases omitted from the official data because an Excel spreadsheet failed to upload onto an Access database

Test and trace system is using Microsoft office to manage the UK's data. It's the equivalent of your bank keeping your records using a quill pen and parchment.
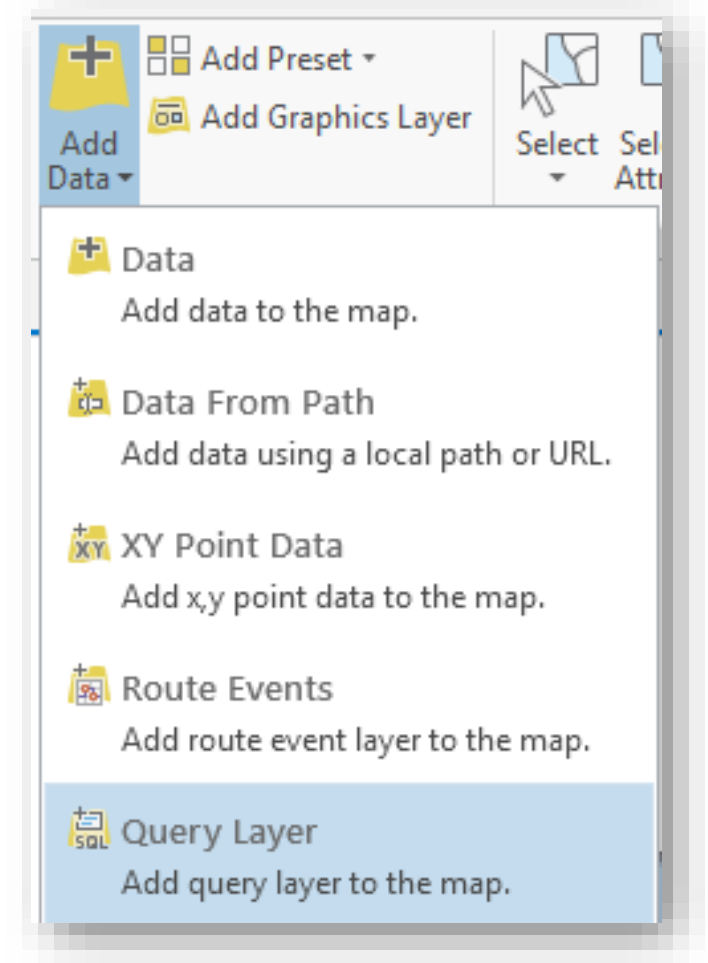
💬 181   🔁 1.5K   ♡ 3.5K   ⬆

⚠ More than 50,000 potentially infectious people may have been missed by contact tracers after 15,841 positive tests were left off the daily figures. Photograph: Simon Leigh/Alamy

A million-row limit on Microsoft's Excel spreadsheet software may have led to Public Health England misplacing nearly 16,000 Covid test results, it is understood.
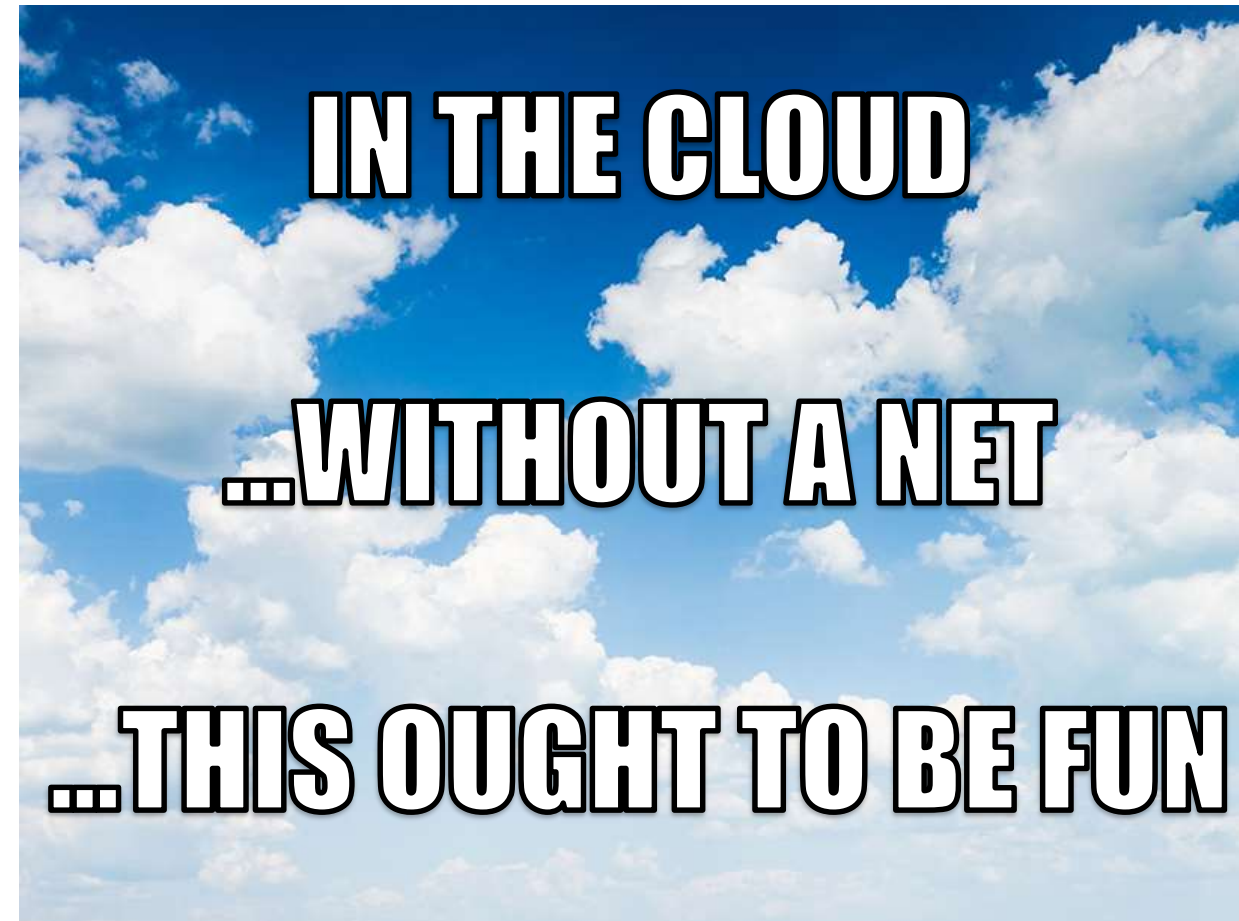
- A *query* is a way to ask a question of the database and get an answer

- A *query layer* in a desktop GIS stores that query in memory and gives you a live look at the data

- A *database view* is a stored query that can behave just like a regular spatial table


```
select
    column1,
    column2,
    shape
from <table>
where column1 = "yes";
```

- Open up two browser tabs

- Direct one to sqlfiddle.com

- Direct the other to
  https://tinyurl.com/sqlpnd

- (will redirect to
  github.com/mmdolbow/SQLPond)

Thank You!

Mike Dolbow

mike.dolbow@state.mn.us - @mmdolbow

651-582-8789