

How to create (LOTS!) of sample time-series data with PostgreSQL `generate_series()`

Ryan Booz, Director of Developer Advocacy

 @ryanbooz

 ryan@timescale.com

[**github.com/ryanbooz/presentations**](https://github.com/ryanbooz/presentations)

**Why create
sample data?**



**Why create sample
data with SQL?**







Fun!



Benefits of `generate_series()`

- No external tools necessary
- No data is *required* from the database itself
- Data can be queried in memory or inserted into a table



Two cautions when using `generate_series()`

- Single threaded
- Uses local server resources (memory/cpu), particularly the more you "customize" the results



DEMO!

- 01** How to use `generate_series()`
- 02** Make the data look more realistic
- 03** Create shaped



Calculating total rows

**(Length of time) x (readings per hour)
x (Number of "things" being tracked)**

Range of readings	Length of interval	Number of "devices"	Total rows
1 year	1 hour	4	35,040
1 year	10 minutes	100	5,256,000
6 months	5 minutes	1,000	52,560,000

**What questions do
you have?**

Thank you!