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
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) \times (readings per hour) \times (Number of "things" being tracked)
<table>
<thead>
<tr>
<th>Range of readings</th>
<th>Length of interval</th>
<th>Number of &quot;devices&quot;</th>
<th>Total rows</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 year</td>
<td>1 hour</td>
<td>4</td>
<td>35,040</td>
</tr>
<tr>
<td>1 year</td>
<td>10 minutes</td>
<td>100</td>
<td>5,256,000</td>
</tr>
<tr>
<td>6 months</td>
<td>5 minutes</td>
<td>1,000</td>
<td>52,560,000</td>
</tr>
</tbody>
</table>
What questions do you have?
Thank you!