Why I Cringe When Everyone Wants Data In Excel

James McNally (CLA, LabVIEW Champion)
Wiresmith Technology
Make You Think
Outline
Florence Nightingale

• “The lady with the data”
• Revolutionised hospital care with careful data collection and visualisation.
Agenda

FILE FORMATS

INSPECTION TOOLS
# File Formats

<table>
<thead>
<tr>
<th>Format</th>
<th>CSV</th>
<th>TDMS</th>
<th>Binary</th>
</tr>
</thead>
<tbody>
<tr>
<td>Write Speed</td>
<td>Good</td>
<td>Better</td>
<td>Best</td>
</tr>
<tr>
<td>Precision</td>
<td>Good</td>
<td>Best</td>
<td>Best</td>
</tr>
<tr>
<td>Read In Anything</td>
<td>Yes</td>
<td>No (General purpose readers available though)</td>
<td>No (Custom reader required)</td>
</tr>
</tbody>
</table>
CSV Files

• Need to include your own meta data
• Need to decide on precision up front (what is the accuracy of the source?)
• No Standards for Time/Date
TDMS

• Hierarchical Structure
• "Database + Binary File"
• How you write impacts performance.
## Technical Comparison

<table>
<thead>
<tr>
<th>Spec</th>
<th>CSV</th>
<th>TDMS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Size</td>
<td>Varies: 2.6: 10 bytes/sample 2.2: 6 bytes/sample</td>
<td>Double: 8 bytes/sample + headers</td>
</tr>
<tr>
<td>Write Speed</td>
<td>1.3MS/s</td>
<td>57MS/s*</td>
</tr>
<tr>
<td>Random Read</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Easy</td>
<td>Yes</td>
<td>No</td>
</tr>
</tbody>
</table>

* After change to interleaved. On decimated performance was 410kS/s
Reading Files - Excel

- CSV
  - Easy – Supported Out Of Box
- TDMS
  - Requires plugin

Limitations
- Up to 1,048,576 Rows
  - 1k: 16 minutes
  - 10k: 1m 44 sec
  - 100k: 10(.5) sec
- Date/Time can be tricky
### Input

- 2018-11-05 13:10:30.123456
- 2018-11-05 13:10:30
- 2018-11-05T13:10:30.123456Z
- 05-11-2018 13:10:30.123456
- 2018/11/05 13:10:30.123456
- 2018/11/05 13:10:30.123456
- 2018-11-05,13:10:30.123456
- 2018-11-05,13:10:30.123456
- 05/11/2018 13:10:30.123456
- 05-11-2018 13:10:30.123456

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td>10:30.1</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>05/11/2018 13:10</td>
</tr>
<tr>
<td>3</td>
<td>2018-11-05T13:10:30.123456Z</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td>10:30.1</td>
</tr>
<tr>
<td>5</td>
<td></td>
<td>10:30.1</td>
</tr>
<tr>
<td>6</td>
<td></td>
<td>05/11/2018 10:30.1</td>
</tr>
<tr>
<td>7</td>
<td></td>
<td>05/11/2018 10:30.1</td>
</tr>
<tr>
<td>8</td>
<td></td>
<td>10:30.1</td>
</tr>
<tr>
<td>9</td>
<td></td>
<td>10:30.1</td>
</tr>
<tr>
<td>10</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Demo - Data Inspection

• Excel
• Diadem
Alternatives?
Summary

• Keep people comfortable with CSV.
• Make their life easier with TDMS (and DIAadem training)
Badge Collector?

Identify gaps in your knowledge and find related learning resources with badges.

Documenting Measurement Data

Acquire and store data and metadata for easy management, analysis, and reporting.

ni.com/badges
21st – 22nd August 2019, Birmingham, UK
Gdevcon.com
Office and Home/Hobby Software Usage Survey

$25 Amazon Gift Card for first 200 qualified respondents who complete the survey

We want to hear from those of you who like to tinker with software development for home projects. How does this influence the tools you use at work and how important are the online tools and ecosystem?

bit.ly/nisurvey2019

Survey will take about 15 minutes
Before you go, take the survey.