Pitfalls of Historical Data Use in Software Estimation

David DeWitt – Director Commercial and International Programs
Presented by: Brian Glauser – V.P. Business Development
Mark Twain Famously Said...
'There are three kinds of lies: lies, damned lies, and statistics'
Problem Is – He Never Said it... His Autobiography Quotes Benjamin Disraeli, a 19th century British Prime Minister

It was Prime Minister Arthur Balfour Quoting Professor Joseph Munro!
Around the Same Time Mark Twain was Misremembering a Famous Quote

“Let data and facts do the talking”

Frederick Taylor: The Principals of Scientific Management 1901.
Can We Trust the Data?

Maybe... But People Can be A Bigger Problem

“In God We Trust... all Others Bring Data”

- W. Edwards Deming
Four Common Fallacies Made in Historical Data Selection
And You Won’t Even See It Coming

Narrative Fallacy
Causal Analysis
Silent Evidence
Ludic Fallacy
Narrative Fallacy

“…limited ability to look at sequences of facts without weaving an explanation into them, or, equivalently, forcing a logical link, an arrow of relationship upon them”
The Narrative Fallacy

A Quick Example

- Sort Pieces
- Build The Frame
- Assemble Large Parts
- Fill In The Holes
- Hunt for Missing Pieces
- Assume Some Missing
1. Neckerchiefs
2. Dogs Paws
3. Tower
4. Umbrellas
5. Sunglasses
Our Trust in Narratives Can Shape How Data is Selected
Our Trust in Narratives Can Shape How Data is Selected
Nobel Prize winner Daniel Kahneman provided a solution to the impact of an “Inside View” in his book Thinking Fast and Slow.
Famous Authors Hold Their Heads…

It’s Hard to See the Narrative Coming!

Steven King
JK Rowling
Oscar Wilde
Kurt Vonnegut
Truman Capote
Virginia Woolf
How Can Causal Errors Slip In?

**SAME ENVIRONMENT**

We specifically selected only ERP and Business Mission Critical Environment data.

**SAME PROBLEM DOMAIN**

We said we were only looking for any projects that were built for Financial Transactions.

**SIMILAR SCALE**

We limited our data to only completed projects of our anticipated size in Function Points.

**ASSUMED CORRELATION**

The Narrative lead us to believe all these data points were of equal value for analysis. And they may be!
Correlation Can Make Us Assume Causation

- Schedule Includes Three months of Sequestration
- Same Vendor and Same Team!
- Offshore Team Working uncompensated Hours not captured
- Missing Test Phase
How To Prevent Causal Analysis Errors?

Again we turn to the Outside View. Need to have someone that understands what the data means – not just where it falls in a historical trend line.
How Can We Trust the Data?

Better Question is... Do We Have All the Data?

“In God We Trust... all Others Bring Data”

- W. Edwards Deming
Building A Historical Data Archive

ID Completed Projects
We look for programs that have a sufficient available data to store in a repository.

Sort Into Domains
We specify what bucket of data the projects belong and we create identifiers for easy extraction.

Grade Data Quality
We provide some indicators as to the integrity of the data and any limitations to usage – typically a quality flag.

Promote the Data
We Notify others of the data availability and encourage usage of the repository.
We look for programs that have a sufficient available data to store in a repository. We specify what bucket of data the projects belong and we create identifiers for easy extraction. We provide some indicators as to the integrity of the data and any limitations to usage – typically a quality flag. We Notify others of the data availability and encourage usage of the repository.
Transparency Into All the Evidence Forces Difficult Conversations
Fundamental Problems With Historical Data

- It’s costly to obtain
- Difficult to catalog and store
- Can often be wrong
- Can be intentionally laced with inaccuracies

Fallacy of Silent Evidence

Drives the conversation about data quality
Ludic Fallacy – Latin for Games

“basing studies of chance on the narrow world of games and dice”
The Pareto Principle:
80% of Results Come from 20% of the Production

Life is not Fair
Has The Historical Data Been Gamed?

“Randomness has an additional layer of uncertainty concerning the rules of the game in real life.”

“Men follow their sentiments and their self-interest, but it pleases them to imagine that they follow reason.”
Most Common Omissions From Historical Data Ranked In Order Of Significance

1) Unpaid overtime by exempt staff
2) Charging time to the wrong project
3) User effort on software projects
4) Management effort on software projects
5) Specialist effort on software projects
   - Human factors specialists
   - Database administration specialists
   - Integration specialists
   - Quality assurance specialists
   - Technical writing specialists
   - Education specialists
   - Hardware or engineering specialists
   - Marketing specialists
   - Metrics and function point specialists
6) Effort spent prior to cost tracking start up
7) Inclusion/exclusion of non-project tasks
   - Departmental meetings
   - Courses and education
   - Travel

Source: Capers Jones, *Errors And Omissions In Software Historical Data: Separating Fact From Fiction*, August 17, 2009
### Typical Results

**Reviewing Customer Historical Data**

<table>
<thead>
<tr>
<th>Activities Performed</th>
<th>Completeness of historical data</th>
</tr>
</thead>
<tbody>
<tr>
<td>01 Requirements</td>
<td>Missing or Incomplete</td>
</tr>
<tr>
<td>02 Prototyping</td>
<td>Missing or Incomplete</td>
</tr>
<tr>
<td>03 Architecture</td>
<td>Missing or Incomplete</td>
</tr>
<tr>
<td>04 Project planning</td>
<td>Missing or Incomplete</td>
</tr>
<tr>
<td>05 Initial analysis and design</td>
<td>Missing or Incomplete</td>
</tr>
<tr>
<td>06 Detail design</td>
<td>Missing or Incomplete</td>
</tr>
<tr>
<td>07 Design reviews</td>
<td>Missing or Incomplete</td>
</tr>
<tr>
<td>08 Coding</td>
<td>Complete</td>
</tr>
<tr>
<td>09 Reusable code acquisition</td>
<td>Missing or Incomplete</td>
</tr>
<tr>
<td>10 Purchased package acquisition</td>
<td>Missing or Incomplete</td>
</tr>
<tr>
<td>11 Code inspections</td>
<td>Missing or Incomplete</td>
</tr>
<tr>
<td>12 Independent verification and validation</td>
<td>Complete</td>
</tr>
<tr>
<td>13 Configuration management</td>
<td>Missing or Incomplete</td>
</tr>
<tr>
<td>14 Integration</td>
<td>Missing or Incomplete</td>
</tr>
<tr>
<td>15 User documentation</td>
<td>Missing or Incomplete</td>
</tr>
<tr>
<td>16 Unit testing</td>
<td>Incomplete</td>
</tr>
<tr>
<td>17 Function testing</td>
<td>Incomplete</td>
</tr>
<tr>
<td>18 Integration testing</td>
<td>Incomplete</td>
</tr>
<tr>
<td>19 System testing</td>
<td>Incomplete</td>
</tr>
<tr>
<td>20 Field testing</td>
<td>Missing or Incomplete</td>
</tr>
<tr>
<td>21 Acceptance testing</td>
<td>Missing or Incomplete</td>
</tr>
<tr>
<td>22 Independent testing</td>
<td>Complete</td>
</tr>
<tr>
<td>23 Quality assurance</td>
<td>Missing or Incomplete</td>
</tr>
<tr>
<td>24 Installation and training</td>
<td>Missing or Incomplete</td>
</tr>
<tr>
<td>25 Project management</td>
<td>Missing or Incomplete</td>
</tr>
<tr>
<td>26 Total project resources, costs</td>
<td>Incomplete</td>
</tr>
</tbody>
</table>

Source: Capers Jones, *Errors And Omissions In Software Historical Data: Separating Fact From Fiction*, August 17, 2009
Incentives For Fraud Have Been Somewhat Mitigated

June 14, 2025

**Operations Employee Timesheet**

<table>
<thead>
<tr>
<th>Task</th>
<th>Loc</th>
<th>WO#</th>
<th>Work Description</th>
<th>Job Title</th>
<th>Title #</th>
<th>Sun</th>
<th>Mon</th>
<th>Tue</th>
<th>Wed</th>
<th>Thu</th>
<th>Fri</th>
<th>Sat</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>SD</td>
<td>32984</td>
<td>Coding Optical Scanner CSC#3</td>
<td>Programmer</td>
<td>SW</td>
<td>6.00</td>
<td>7.00</td>
<td>7.00</td>
<td>5.00</td>
<td>8.00</td>
<td>33.00</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>SD</td>
<td>32984</td>
<td>Unit Testing CSC #3</td>
<td>Programmer</td>
<td>SW</td>
<td>1.00</td>
<td>2.00</td>
<td></td>
<td></td>
<td></td>
<td>3.00</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>SD</td>
<td>32985</td>
<td>Bug Fixing CSC #3</td>
<td>Programmer</td>
<td>SW</td>
<td>1.00</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td>2.00</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>SD</td>
<td>400329</td>
<td>Design Support</td>
<td>Programmer</td>
<td>SW</td>
<td></td>
<td></td>
<td>1.00</td>
<td>1.00</td>
<td></td>
<td>2.00</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>SD</td>
<td>335329</td>
<td>Requirements Review</td>
<td>Programmer</td>
<td>SW</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Total Regular Hours:**

8.00  8.00  8.00  8.00  8.00  40.00

---

Joe - Please move hours for task 1 across task 4 and 5. The Boss
Overcoming Four Fallacies Made in Historical Data Selection

**Narrative Fallacy**
- Deploy An Outside View

**Causal Analysis**
- Deploy An Outside View

**Silent Evidence**
- Bring All Your Data

**Ludic Fallacy**
- Trust... But Verify!
Thank For Your Time Today

Galorath
Incorporated
info@galorath.com
GALORATH.COM