Lessons Learned From Software Maintenance and Support Datasets

Arlene Minkiewicz
Chief Scientist
PRICE Systems LLC
Arlene.Minkiewicz@pricesystems.com
PRICE®

Cost Analytics

TrueExplorer  TrueFindings  PRICE® Models  TruePlanner  TrueMapper  TrueBOE  TrueXLS

Search & extract data from the PCA Ecosystem
Manage & Analyze Data Sets
Predictive Models
Integration Framework
Customer Data Mapping
Basis-of-Estimate Generator
Access PCA Engine from Excel

Estimate with Confidence™
Agenda

• Introduction
  – Software Sustainment
  – ISBSG

• Data Collection and Analysis Process

• ISBSG Data Study
  – Study Objectives
  – Data Collection and Preparation
  – Modeling and Evaluation
  – Deployment

• Lessons Learned

• Future Work
Introduction

• Software Sustainment has become an increasingly important part of the operation of complex systems
  – Software is more malleable than hardware, easier to change ‘on the fly’
  – The life of many software projects have greatly exceeded their expected lifetime
  – Changes need to be made to the software to evolve it as hardware and software technology evolves
  – Software is easier to deploy than hardware – sometimes used to correct hardware inadequacies

• While many studies have been done over time on the estimation of software sustainment costs, few have much publicly available data behind them

• This study presents lessons learned from the ISBSG database
Software Sustainment

• Many definitions for software sustainment

  “Software sustainment involves orchestrating the processes, practices, technical resources, information and workforce competencies for systems and software engineering, to enable system to continue mission operations and also to be able to be enhanced to meet evolving threat and capability needs.”

  – Software Engineering Institute at Carnegie Mellon University (SEI/CMU)

• Software maintenance is “the process of modifying a software system after delivery to correct faults, improve performance and adapt to changing environments.”

  – Institute of Electrical and Electronics Engineering’s (IEEE) Standard 12207

• Software Sustainment is generally considered to be all activities associated with a software application after it is first released to the public
Software Sustainment

• For the purposes of this study, Software Sustainment will encompass the activities outline in the ISBSG Attributes associated with effort collection
  – Adaptive Maintenance
  – Corrective Maintenance
  – Perfective Maintenance
  – Preventative Maintenance
  – Maintenance Management
  – Problem Investigation Support
  – Queries and Quick Service Support
  – User Help and Advice Support
  – Support Management
ISBSG

• International Software Benchmark Standards Group
  – Not for profit
  – Founded in Australia in 1997
  – Mission to help the Information Technology (IT) industry improve software processes and Practices

• All ISBSG Data is:
  – Validated and rated in accordance with the ISBSG Quality Guidelines
  – Representative of the industry
  – Independent and Trusted
  – Captured from a range of organizations and industries
  – Anonymously stored in the repository

• In August 2020 Version 7 of the ISBSG Maintenance and Support (M&S) database was release

• Study data was collected from this database

www.isbsg.org | ISBSG
Delivering IT Confidence
Data Collection and Analysis
Data Collection and Analysis Process

• Data collection is a daunting task

• Data mining is the process of finding anomalies, patterns and correlations within data sets to predict outcomes and to turn raw data into useful information.

• Sound data mining practices applied sensibly are an important success criteria for data collection and analysis projects
Data Collection and Analysis Process

- Understand the data challenge
- Data Understanding and Preparation
- Modeling
- Evaluation
- Deployment

Iterative Process
ISBSG Data Study
The Questions To Be Answered

• What lessons can be learned from the ISBSG M&S Data?

• How can these lessons be used to provide software maintenance and support estimation guidance to the cost community?

• What useful benchmarks can be gleaned from the industries to support estimation efforts?

• How do Source Lines of Code compare to Functional Size Measurements with respect to effort and defect estimation?
Data Preparation and Understanding

• ISBSG M&S Database
  – 1177 Data Points
  – 114 potential features (attributes)

• Two datasets to be collected
  – Application Sizing Approach is either IFPUG or NESMA
  – Entered KSLOC is greater than 0

• Acquire the right tools for the next steps
  – RapidMiner
  – TrueFindings®
  – Microsoft® Excel
Data Preparation and Understanding

• Operators in the RM Process accomplish the following tasks
  • Retrieve the ISBSG M&S Data
    • Since the data from the ISBSG was in the Excel format it was easily imported into RapidMiner for use in this process.
  • Filter out all records for which Total Maintenance and Support Hours = 0.
  • Filter to develop the two Datasets
    • Application Size Approach = IFPUG or NESMA
    • Entered KSLOC > 0
  • For the records that remain set the following values to 0 if they are missing
    • Adaptive Maintenance Hours
    • Corrective Maintenance Hours
    • Perfective Maintenance Hours
    • Preventative Maintenance Hours
    • Problem Investigation Support Hours
    • Queries and Quick Service Support Hours
    • User Help and Advice Support
    • Minor Defects
    • Major Defects
    • Extreme Defects
  • Removed records where the Age of Benchmark attribute was missing.
Operators in the RM Process accomplish the following tasks (cont.):

- Eliminate attributes that are not needed for this study or are scarcely populated. Appendix C contains a list of the attributes eliminated along with the rationale.
- Filter out records missing Industry Sector.
- Eliminate records with Total Defects Missing.
- Replace Missing Percent %P with 100 – %P is the percent of Primary Language in the application.
- Eliminate records where the %P is equal to 0 as this makes the data record suspicious.
- Convert Data Quality Factors to Numerical values to make numerical analyses possible.
- Replace Application Type and Primary Language with Numerical values to make numerical analyses possible.
- Replace Benchmark Start and End Date with Numerical Values.
- Write resulting example set to a CSV file for importation into TrueFindings®.
Benchmarks for each Data Set

• Benchmarks across both data sets (IFPUG and KSLOC) were compiled for:
  – Application Size
  – Hours (Adaptive, Corrective, Preventive, Perfective, Management, Problem Investigation, Queries and Quick Service, User Help and Advice, Maintenance, Support and Total M&S)
  – Hours Per Size Unit for each type of Hours mentioned above
  – Defect counts (Minor, Major & Extreme)

• Benchmarks include:
  – Sample Size
  – Min
  – Max
  – 25%
  – 75%
  – Median
  – Mode
  – Standard Deviation
  – Coefficient of Variation
### Sample Benchmark Data

<table>
<thead>
<tr>
<th>Attribute</th>
<th>Sample Size</th>
<th>Min</th>
<th>Max</th>
<th>25%</th>
<th>75%</th>
<th>Mean</th>
<th>Median</th>
<th>Mode</th>
<th>Standard Deviation</th>
<th>Coefficient of Variation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age at Benchmark</td>
<td>226</td>
<td>1</td>
<td>31</td>
<td>4</td>
<td>12</td>
<td>8.85</td>
<td>8</td>
<td>3</td>
<td>6.018</td>
<td>0.68</td>
</tr>
<tr>
<td>Adaptive Hours Per FP</td>
<td>30</td>
<td>0.002</td>
<td>3.894</td>
<td>0.055</td>
<td>0.423</td>
<td>0.44</td>
<td>0.197</td>
<td>0.002</td>
<td>0.768</td>
<td>1.746</td>
</tr>
<tr>
<td>Adaptive Hours</td>
<td>30</td>
<td>6</td>
<td>1032</td>
<td>21.75</td>
<td>388.75</td>
<td>235</td>
<td>69</td>
<td>18</td>
<td>304.77</td>
<td>1.297</td>
</tr>
<tr>
<td>Application Size</td>
<td>226</td>
<td>24</td>
<td>55312</td>
<td>459</td>
<td>3477.8</td>
<td>2931.24</td>
<td>1261.5</td>
<td>234</td>
<td>4893.831</td>
<td>1.67</td>
</tr>
<tr>
<td>Corrective FP Rate</td>
<td>41</td>
<td>0.002</td>
<td>4.292</td>
<td>0.038</td>
<td>0.327</td>
<td>0.356</td>
<td>0.095</td>
<td>0.002</td>
<td>0.742</td>
<td>2.083</td>
</tr>
<tr>
<td>Corrective Maintenance Hours</td>
<td>41</td>
<td>2</td>
<td>4224</td>
<td>24</td>
<td>132</td>
<td>320.707</td>
<td>48</td>
<td>6</td>
<td>843.871</td>
<td>2.631</td>
</tr>
<tr>
<td>Extreme Defects per FP</td>
<td>183</td>
<td>3.10E-04</td>
<td>0.066</td>
<td>0.002</td>
<td>0.006</td>
<td>0.005</td>
<td>0.004</td>
<td>0.002</td>
<td>0.005</td>
<td>1.12</td>
</tr>
<tr>
<td>Extreme Defects</td>
<td>183</td>
<td>1</td>
<td>152</td>
<td>2</td>
<td>15.5</td>
<td>12.778</td>
<td>5</td>
<td>2</td>
<td>18.525</td>
<td>1.45</td>
</tr>
<tr>
<td>Maintenance FP Rate</td>
<td>208</td>
<td>0.1</td>
<td>6.3</td>
<td>0.2</td>
<td>0.411</td>
<td>0.414</td>
<td>0.307</td>
<td>0.2</td>
<td>0.624</td>
<td>1.506</td>
</tr>
<tr>
<td>Major Defects FP Rate</td>
<td>185</td>
<td>3.10E-04</td>
<td>1</td>
<td>0.006</td>
<td>0.012</td>
<td>0.024</td>
<td>0.008</td>
<td>0.008</td>
<td>0.111</td>
<td>4.58</td>
</tr>
<tr>
<td>Major Defects</td>
<td>185</td>
<td>1</td>
<td>326.51</td>
<td>5</td>
<td>32</td>
<td>27.4</td>
<td>14</td>
<td>3</td>
<td>41.697</td>
<td>1.522</td>
</tr>
<tr>
<td>Management Maintenance Hours</td>
<td>172</td>
<td>2</td>
<td>3139</td>
<td>66.25</td>
<td>283.5</td>
<td>215.326</td>
<td>149</td>
<td>52</td>
<td>291.266</td>
<td>1.353</td>
</tr>
<tr>
<td>Minor Defects Per FP</td>
<td>193</td>
<td>3.10E-04</td>
<td>0.667</td>
<td>0.016</td>
<td>0.036</td>
<td>0.037</td>
<td>0.022</td>
<td>0.05</td>
<td>0.065</td>
<td>1.729</td>
</tr>
<tr>
<td>Minor Defects</td>
<td>193</td>
<td>1</td>
<td>954</td>
<td>19</td>
<td>82</td>
<td>69.658</td>
<td>41</td>
<td>12</td>
<td>121.581</td>
<td>1.745</td>
</tr>
<tr>
<td>Perfective Hours per FP</td>
<td>26</td>
<td>0.006</td>
<td>0.43</td>
<td>0.03</td>
<td>0.094</td>
<td>0.078</td>
<td>0.065</td>
<td>0.086</td>
<td>0.084</td>
<td>1.076</td>
</tr>
<tr>
<td>Preventive Maintenance Hours</td>
<td>34</td>
<td>1</td>
<td>240</td>
<td>6.75</td>
<td>76.5</td>
<td>60.853</td>
<td>31</td>
<td>6</td>
<td>71.711</td>
<td>1.178</td>
</tr>
<tr>
<td>Preventive FP Rate</td>
<td>34</td>
<td>0.002</td>
<td>0.747</td>
<td>0.034</td>
<td>0.113</td>
<td>0.116</td>
<td>0.068</td>
<td>0.1</td>
<td>0.16</td>
<td>1.377</td>
</tr>
<tr>
<td>Problem Investigation FP Rate</td>
<td>31</td>
<td>0.003</td>
<td>2.113</td>
<td>0.041</td>
<td>0.234</td>
<td>0.226</td>
<td>0.098</td>
<td>0.003</td>
<td>0.408</td>
<td>1.802</td>
</tr>
<tr>
<td>Problem Investigation Hours</td>
<td>31</td>
<td>8</td>
<td>1237</td>
<td>23.5</td>
<td>106.5</td>
<td>112.258</td>
<td>60</td>
<td>13</td>
<td>220.839</td>
<td>1.967</td>
</tr>
<tr>
<td>Queries &amp; Quick Service Support Hours</td>
<td>33</td>
<td>1</td>
<td>3752</td>
<td>13</td>
<td>132</td>
<td>243.909</td>
<td>67</td>
<td>13</td>
<td>656.753</td>
<td>2.693</td>
</tr>
<tr>
<td>Queries and Quick Service FP Rate</td>
<td>33</td>
<td>0.003</td>
<td>1.1</td>
<td>0.055</td>
<td>0.271</td>
<td>0.222</td>
<td>0.12</td>
<td>0.003</td>
<td>0.274</td>
<td>1.233</td>
</tr>
<tr>
<td>Support FP Rate</td>
<td>204</td>
<td>0.1</td>
<td>3.3</td>
<td>0.186</td>
<td>0.378</td>
<td>0.339</td>
<td>0.3</td>
<td>0.2</td>
<td>0.337</td>
<td>0.992</td>
</tr>
<tr>
<td>total Defects FP Rate</td>
<td>219</td>
<td>0.001</td>
<td>1.667</td>
<td>0.023</td>
<td>0.05</td>
<td>0.059</td>
<td>0.034</td>
<td>0.043</td>
<td>0.157</td>
<td>2.647</td>
</tr>
<tr>
<td>Total Defects</td>
<td>219</td>
<td>1</td>
<td>1311</td>
<td>22.5</td>
<td>111.5</td>
<td>96.831</td>
<td>51</td>
<td>32</td>
<td>159.462</td>
<td>1.647</td>
</tr>
<tr>
<td>Total Maintenance and Support Hours</td>
<td>226</td>
<td>11</td>
<td>22423</td>
<td>329.3</td>
<td>1759</td>
<td>1511.04</td>
<td>881.5</td>
<td>52</td>
<td>2457.856</td>
<td>1.627</td>
</tr>
<tr>
<td>Total Maintenance Hours</td>
<td>219</td>
<td>6</td>
<td>13182</td>
<td>181</td>
<td>936.5</td>
<td>780.174</td>
<td>452</td>
<td>126</td>
<td>1327.703</td>
<td>1.702</td>
</tr>
<tr>
<td>Total Support Hours</td>
<td>208</td>
<td>8</td>
<td>11212</td>
<td>170.8</td>
<td>841.25</td>
<td>687.611</td>
<td>453.5</td>
<td>40</td>
<td>1021.549</td>
<td>1.486</td>
</tr>
<tr>
<td>User Help and Advise Support Hours</td>
<td>27</td>
<td>3</td>
<td>495</td>
<td>12.5</td>
<td>99.5</td>
<td>72.074</td>
<td>33</td>
<td>13</td>
<td>100.605</td>
<td>1.396</td>
</tr>
<tr>
<td>User Help and Advice FP Rate</td>
<td>27</td>
<td>0.004</td>
<td>1.1</td>
<td>0.027</td>
<td>0.147</td>
<td>0.128</td>
<td>0.051</td>
<td>0.004</td>
<td>0.22</td>
<td>1.72</td>
</tr>
</tbody>
</table>
Additional Benchmark Scenarios

• IFPUG/NESMA Dataset
  – Primary Language
    • .Net
    • COBOL
    • Java
    • Oracle
  – Application Type
    • Business Applications
    • CRM/ERP
    • Document Management
    • Web Application
  – Industry Sector
    • Education
    • Financial
    • Government
    • Logistics
    • Utilities
  – Organization Type
    • Banking
    • Transport

• KSLOC Dataset
  – Primary Language
    • COBOL
    • Java
    • JCL
    • Telon
  – Industry Sector
    • Banking
    • Education
    • Government
    • Wholesale and Retail

IFPUG/NESMA ORGANIZATION TYPE DATA POINTS

- Banking: 16%
- Education: 15%
- Energy/Utilities: 12%
- Government: 24%
- Other: 24%
- Transport: 9%

Presented for the ICEAA 2021 Online Workshop - www.iceaaonline.com
# Sample Benchmark Data

## IFPUG/NESMA Industry Sector Data Points

<table>
<thead>
<tr>
<th>Sector</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Education</td>
<td>8%</td>
</tr>
<tr>
<td>Financial</td>
<td>10%</td>
</tr>
<tr>
<td>Government</td>
<td>12%</td>
</tr>
<tr>
<td>Logistics</td>
<td>4%</td>
</tr>
<tr>
<td>Utilities</td>
<td>8%</td>
</tr>
<tr>
<td>Other</td>
<td>8%</td>
</tr>
</tbody>
</table>

**Total**: 50%

## Attribute Sample Size Min Max 25% 75% Mean Median Mode Standard Deviation Coefficient of Variation

<table>
<thead>
<tr>
<th>Attribute</th>
<th>Sample Size</th>
<th>Min</th>
<th>Max</th>
<th>25%</th>
<th>75%</th>
<th>Mean</th>
<th>Median</th>
<th>Mode</th>
<th>Standard Deviation</th>
<th>Coefficient of Variation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age at Benchmark</td>
<td>54</td>
<td>1</td>
<td>31</td>
<td>6</td>
<td>12</td>
<td>9.981</td>
<td>9</td>
<td>6</td>
<td>6.299</td>
<td>0.631</td>
</tr>
<tr>
<td>Application Size</td>
<td>54</td>
<td>24</td>
<td>18500</td>
<td>444.75</td>
<td>3542</td>
<td>2724.5</td>
<td>871.5</td>
<td>453</td>
<td>3578.376</td>
<td>1.313</td>
</tr>
<tr>
<td>Maintenance FP Rate</td>
<td>45</td>
<td>0.1</td>
<td>2.1</td>
<td>0.206</td>
<td>0.374</td>
<td>0.383</td>
<td>0.287</td>
<td>0.2</td>
<td>0.357</td>
<td>0.931</td>
</tr>
<tr>
<td>Support FP Rate</td>
<td>38</td>
<td>0.112</td>
<td>0.732</td>
<td>0.197</td>
<td>0.363</td>
<td>0.307</td>
<td>0.29</td>
<td>0.112</td>
<td>0.142</td>
<td>0.462</td>
</tr>
<tr>
<td>Total Defects FP Rate</td>
<td>54</td>
<td>0.003</td>
<td>1.667</td>
<td>0.022</td>
<td>0.071</td>
<td>0.122</td>
<td>0.037</td>
<td>0.043</td>
<td>0.305</td>
<td>2.499</td>
</tr>
<tr>
<td>Total Defects</td>
<td>54</td>
<td>5.415</td>
<td>1037</td>
<td>28.25</td>
<td>121.3</td>
<td>101.194</td>
<td>48.5</td>
<td>32</td>
<td>155.578</td>
<td>1.537</td>
</tr>
<tr>
<td>Total Maintenance &amp; Support Hours</td>
<td>54</td>
<td>13</td>
<td>15338</td>
<td>273.5</td>
<td>1824</td>
<td>1515.13</td>
<td>809</td>
<td>13</td>
<td>2385.777</td>
<td>1.575</td>
</tr>
<tr>
<td>Total Maintenance Hours</td>
<td>52</td>
<td>13</td>
<td>4657</td>
<td>137.5</td>
<td>941</td>
<td>718.635</td>
<td>364.5</td>
<td>225</td>
<td>926.848</td>
<td>1.29</td>
</tr>
<tr>
<td>Total Maintenance &amp; Support FP Rate</td>
<td>54</td>
<td>0.038</td>
<td>4.75</td>
<td>0.367</td>
<td>0.759</td>
<td>0.728</td>
<td>0.606</td>
<td>0.038</td>
<td>0.732</td>
<td>1.005</td>
</tr>
<tr>
<td>Total Support Hours</td>
<td>38</td>
<td>91</td>
<td>3153</td>
<td>235.25</td>
<td>967.5</td>
<td>734.579</td>
<td>495.5</td>
<td>91</td>
<td>662.119</td>
<td>0.901</td>
</tr>
</tbody>
</table>

**Presented for the ICEAA 2021 Online Workshop - www.iceaaonline.com**
Modeling and Evaluation

• Visualization and Correlation Analysis rendered several initial conclusions
  – There was not very much interesting to learn from the KSLOC data set
    • Possibly due to the primary focus of ISBSG Data collection – which is Functional Size
    • Further analysis on the KSLOC data set was postponed to a future time when more data is available
  – There appeared to be inconsistencies between the Total Hours for Support and Maintenance as compared to the categorial hours
  – The relationships around Size and Defects told a more compelling story about Support than Maintenance

• Modeling was performed using RapidMiner. Regression Analysis was the primary technique
## Modeling and Evaluation

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Maintenance hours</td>
<td>1</td>
<td>0.140</td>
<td>0.152</td>
<td>0.253</td>
<td>0.045</td>
<td>0.061</td>
<td>0.810</td>
<td>0.129</td>
<td>0.002</td>
<td>0.101</td>
<td>-0.033</td>
<td>0.538</td>
<td>0.363</td>
<td>0.457</td>
<td>0.144</td>
<td>0.835</td>
<td>0.708</td>
<td>0.495</td>
<td>0.708</td>
</tr>
<tr>
<td>Preventive Maintenance hours</td>
<td>0.140</td>
<td>1</td>
<td>0.782</td>
<td>0.555</td>
<td>0.454</td>
<td>-0.140</td>
<td>0.441</td>
<td>0.198</td>
<td>0.382</td>
<td>0.194</td>
<td>-0.115</td>
<td>-0.113</td>
<td>-0.009</td>
<td>-0.033</td>
<td>0.054</td>
<td>0.053</td>
<td>0.053</td>
<td>-0.037</td>
<td>0.052</td>
</tr>
<tr>
<td>Corrective Maintenance hours</td>
<td>0.152</td>
<td>0.782</td>
<td>1</td>
<td>0.625</td>
<td>0.522</td>
<td>-0.159</td>
<td>-0.034</td>
<td>0.550</td>
<td>0.175</td>
<td>0.513</td>
<td>-0.138</td>
<td>-0.133</td>
<td>-0.024</td>
<td>-0.103</td>
<td>0.054</td>
<td>0.029</td>
<td>-0.055</td>
<td>0.027</td>
<td></td>
</tr>
<tr>
<td>Adaptive Maintenance hours</td>
<td>0.253</td>
<td>0.555</td>
<td>0.025</td>
<td>1</td>
<td>0.102</td>
<td>-0.094</td>
<td>0.023</td>
<td>0.577</td>
<td>0.152</td>
<td>0.501</td>
<td>0.005</td>
<td>-0.057</td>
<td>-0.044</td>
<td>0.129</td>
<td>-0.037</td>
<td>0.137</td>
<td>0.144</td>
<td>0.079</td>
<td>0.142</td>
</tr>
<tr>
<td>Management Maintenance hours</td>
<td>0.045</td>
<td>0.454</td>
<td>0.522</td>
<td>0.102</td>
<td>1</td>
<td>-0.140</td>
<td>-0.058</td>
<td>0.217</td>
<td>0.155</td>
<td>0.351</td>
<td>0.075</td>
<td>-0.132</td>
<td>-0.131</td>
<td>-0.116</td>
<td>-0.091</td>
<td>-0.111</td>
<td>-0.071</td>
<td>-0.112</td>
<td>-0.069</td>
</tr>
<tr>
<td>Total Support hours</td>
<td>0.661</td>
<td>-0.140</td>
<td>-0.159</td>
<td>-0.094</td>
<td>-0.140</td>
<td>1</td>
<td>0.910</td>
<td>-0.705</td>
<td>-0.083</td>
<td>-0.127</td>
<td>-0.055</td>
<td>0.798</td>
<td>0.586</td>
<td>0.810</td>
<td>0.263</td>
<td>0.692</td>
<td>0.907</td>
<td>0.572</td>
<td>0.914</td>
</tr>
<tr>
<td>Problem Investigation Support hours</td>
<td>0.013</td>
<td>0.411</td>
<td>0.550</td>
<td>0.577</td>
<td>0.217</td>
<td>-0.105</td>
<td>0.083</td>
<td>1</td>
<td>0.219</td>
<td>0.855</td>
<td>-0.005</td>
<td>-0.071</td>
<td>-0.008</td>
<td>-0.056</td>
<td>-0.096</td>
<td>0.094</td>
<td>-0.033</td>
<td>-0.068</td>
<td>-0.034</td>
</tr>
<tr>
<td>Quoano &amp; Quick Service Support hours</td>
<td>0.002</td>
<td>0.108</td>
<td>0.175</td>
<td>0.156</td>
<td>0.156</td>
<td>-0.083</td>
<td>0.213</td>
<td>0.219</td>
<td>1</td>
<td>0.200</td>
<td>-0.011</td>
<td>0.014</td>
<td>0.452</td>
<td>0.244</td>
<td>0.022</td>
<td>0.081</td>
<td>0.034</td>
<td>0.297</td>
<td>-0.007</td>
</tr>
<tr>
<td>User Help &amp; Advice Support hours</td>
<td>0.101</td>
<td>0.302</td>
<td>0.513</td>
<td>0.501</td>
<td>0.361</td>
<td>-0.127</td>
<td>0.053</td>
<td>0.855</td>
<td>0.209</td>
<td>1</td>
<td>-0.001</td>
<td>-0.096</td>
<td>-0.119</td>
<td>-0.080</td>
<td>-0.111</td>
<td>0.065</td>
<td>-0.046</td>
<td>-0.080</td>
<td>-0.046</td>
</tr>
<tr>
<td>Management Support Hours</td>
<td>-0.003</td>
<td>0.164</td>
<td>-0.025</td>
<td>0.005</td>
<td>0.075</td>
<td>-0.065</td>
<td>-0.058</td>
<td>-0.006</td>
<td>-0.011</td>
<td>-0.001</td>
<td>-1</td>
<td>-0.029</td>
<td>-0.012</td>
<td>-0.030</td>
<td>-0.068</td>
<td>-0.046</td>
<td>-0.024</td>
<td>-0.035</td>
<td>-0.027</td>
</tr>
<tr>
<td>Extreme Defects</td>
<td>0.539</td>
<td>-0.115</td>
<td>-0.130</td>
<td>-0.057</td>
<td>-0.132</td>
<td>0.798</td>
<td>0.768</td>
<td>-0.071</td>
<td>0.014</td>
<td>-0.006</td>
<td>-0.029</td>
<td>1</td>
<td>0.759</td>
<td>0.538</td>
<td>0.119</td>
<td>0.572</td>
<td>0.851</td>
<td>0.750</td>
<td>0.651</td>
</tr>
<tr>
<td>Major Defects</td>
<td>0.383</td>
<td>-0.113</td>
<td>-0.133</td>
<td>-0.044</td>
<td>-0.131</td>
<td>0.586</td>
<td>0.674</td>
<td>-0.090</td>
<td>0.452</td>
<td>-0.119</td>
<td>-0.012</td>
<td>0.759</td>
<td>1</td>
<td>0.724</td>
<td>0.117</td>
<td>0.509</td>
<td>0.882</td>
<td>0.654</td>
<td>0.656</td>
</tr>
<tr>
<td>Minor Defects</td>
<td>0.457</td>
<td>-0.009</td>
<td>-0.024</td>
<td>0.129</td>
<td>-0.116</td>
<td>0.610</td>
<td>0.642</td>
<td>-0.050</td>
<td>0.244</td>
<td>-0.080</td>
<td>-0.038</td>
<td>0.538</td>
<td>0.724</td>
<td>1</td>
<td>0.159</td>
<td>0.712</td>
<td>0.733</td>
<td>0.973</td>
<td>0.703</td>
</tr>
<tr>
<td>Age at Benchmark</td>
<td>0.144</td>
<td>-0.033</td>
<td>-0.103</td>
<td>-0.027</td>
<td>-0.091</td>
<td>0.203</td>
<td>0.230</td>
<td>-0.096</td>
<td>0.022</td>
<td>-0.111</td>
<td>-0.008</td>
<td>0.119</td>
<td>0.117</td>
<td>0.160</td>
<td>1</td>
<td>0.194</td>
<td>0.190</td>
<td>0.161</td>
<td>0.195</td>
</tr>
<tr>
<td>Total Maintenance &amp; Support hours</td>
<td>0.835</td>
<td>0.054</td>
<td>0.054</td>
<td>0.137</td>
<td>-0.011</td>
<td>0.022</td>
<td>0.007</td>
<td>0.094</td>
<td>0.081</td>
<td>0.005</td>
<td>-0.040</td>
<td>0.572</td>
<td>0.569</td>
<td>0.712</td>
<td>0.194</td>
<td>1</td>
<td>0.702</td>
<td>0.712</td>
<td>0.735</td>
</tr>
<tr>
<td>Application size</td>
<td>0.708</td>
<td>0.053</td>
<td>0.020</td>
<td>0.144</td>
<td>-0.071</td>
<td>0.067</td>
<td>0.890</td>
<td>-0.033</td>
<td>0.034</td>
<td>-0.046</td>
<td>-0.024</td>
<td>0.851</td>
<td>0.828</td>
<td>0.733</td>
<td>0.198</td>
<td>0.752</td>
<td>1</td>
<td>0.795</td>
<td>0.888</td>
</tr>
<tr>
<td>Total Defects</td>
<td>0.495</td>
<td>-0.037</td>
<td>-0.055</td>
<td>0.079</td>
<td>-0.112</td>
<td>0.672</td>
<td>0.720</td>
<td>-0.088</td>
<td>0.297</td>
<td>-0.069</td>
<td>-0.035</td>
<td>0.758</td>
<td>0.854</td>
<td>0.973</td>
<td>0.161</td>
<td>0.712</td>
<td>0.795</td>
<td>1</td>
<td>0.787</td>
</tr>
<tr>
<td>CalcApplicationSize</td>
<td>0.708</td>
<td>0.052</td>
<td>0.027</td>
<td>0.142</td>
<td>-0.069</td>
<td>0.914</td>
<td>0.884</td>
<td>-0.034</td>
<td>-0.007</td>
<td>-0.048</td>
<td>-0.027</td>
<td>0.851</td>
<td>0.658</td>
<td>0.703</td>
<td>0.195</td>
<td>0.735</td>
<td>0.998</td>
<td>0.767</td>
<td>1</td>
</tr>
</tbody>
</table>
Modeling and Evaluation

Presented for the ICEAA 2021 Online Workshop - www.iceaaonline.com
Modeling and Evaluation

- Filter example is used to remove values of 0 from the data set for the attributes that are slated to be independent and dependent variables in the analysis.
  - For example - some examples do not have Support Hours because it was not reported by the submitter (possibly the organization does not track that, or another department handles it) so that data point would not be a valuable addition to a study looking at drivers for Support Hours.

- Select Attributes selects those attributes that represent the dependent and independent variables for a particular analysis.

- Set Role labels the Dependent variable for a particular analysis.

- The Cross Validation Operator performs a Linear Regression on the entire data set and returns performance parameters.

- The Split Data Operator splits the data into a training set (75% of examples) and a scoring set (25% of examples).

- The Model Created by the Linear Regression Operator and the scoring set are sent to the Apply Model which determines how well the model performs on data not in the training set.
# Modeling and Evaluation

<table>
<thead>
<tr>
<th>Dependent Variable</th>
<th>Independent Variable 1</th>
<th>Independent Variable 2</th>
<th>Independent Variable 3</th>
<th>Equation</th>
<th>( R^2 ) (equation)</th>
<th>Training Count</th>
<th>Scoring Count</th>
<th>( R^2 ) (prediction)</th>
<th>Pred(30)</th>
<th>Pred(50)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Maintenance Hours</td>
<td>Application Size(app)</td>
<td>Extreme Defects (exd)</td>
<td>Major Defects (mad)</td>
<td>( 165.806 + 0.2469 \times \text{app} - 12.171 \times \text{exd} - 2.604 \times \text{mad} )</td>
<td>0.949</td>
<td>134</td>
<td>45</td>
<td>0.888</td>
<td>62%</td>
<td>89%</td>
</tr>
<tr>
<td>Total Maintenance &amp; Support Hours</td>
<td>Application Size(app)</td>
<td>Extreme Defects (exd)</td>
<td>Major Defects (mad)</td>
<td>( 226.824 + 0.3977 \times \text{app} - 0.54477 \times \text{exd} - 3.6202 \times \text{mad} )</td>
<td>0.927</td>
<td>135</td>
<td>45</td>
<td>0.483</td>
<td>58%</td>
<td>76%</td>
</tr>
<tr>
<td>Total Support Hours</td>
<td>Total Maintenance Hours (totMaint)</td>
<td></td>
<td></td>
<td>( 202.89 + 0.621 \times \text{totMaint} )</td>
<td>0.910</td>
<td>152</td>
<td>51</td>
<td>0.959</td>
<td>47%</td>
<td>76%</td>
</tr>
<tr>
<td>Total Support Hours</td>
<td>Application Size(app)</td>
<td>Extreme Defects (exd)</td>
<td>Major Defects (mad)</td>
<td>( 125.929 + 0.2366 \times \text{app} - 5.015 \times \text{exd} - 5.125 \times \text{mad} )</td>
<td>0.879</td>
<td>117</td>
<td>39</td>
<td>0.879</td>
<td>60%</td>
<td>88%</td>
</tr>
<tr>
<td>Total Support Hours</td>
<td>Application Size(app)</td>
<td></td>
<td></td>
<td>( 85.746 + 0.193 \times \text{app} )</td>
<td>0.855</td>
<td>164</td>
<td>55</td>
<td>0.528</td>
<td>56%</td>
<td>78%</td>
</tr>
<tr>
<td>Extreme Defects</td>
<td>Application Size(app)</td>
<td></td>
<td></td>
<td>( 2.81 + 0.0029 \times \text{app} )</td>
<td>0.723</td>
<td>137</td>
<td>46</td>
<td>0.792</td>
<td>30%</td>
<td>52%</td>
</tr>
<tr>
<td>Total Support Hours</td>
<td>Total Defects (td)</td>
<td></td>
<td></td>
<td>( 181.055 + 5.759 \times \text{td} )</td>
<td>0.700</td>
<td>151</td>
<td>50</td>
<td>0.571</td>
<td>30%</td>
<td>54%</td>
</tr>
<tr>
<td>Total Maintenance Hours</td>
<td>Total Defects (td)</td>
<td></td>
<td></td>
<td>( 383.085 + 4.828 \times \text{td} )</td>
<td>0.660</td>
<td>158</td>
<td>53</td>
<td>0.593</td>
<td>22%</td>
<td>45%</td>
</tr>
<tr>
<td>Total Support Hours</td>
<td>Extreme Defects (exd)</td>
<td>Major Defects (mad)</td>
<td></td>
<td>( 90.07 + 10.798 \times \text{exd} + 20.812 \times \text{mad} )</td>
<td>0.659</td>
<td>131</td>
<td>43</td>
<td>0.684</td>
<td>28%</td>
<td>65%</td>
</tr>
<tr>
<td>Total Defects</td>
<td>Application Size(app)</td>
<td></td>
<td></td>
<td>( 314.571 + 31.084 )</td>
<td>0.632</td>
<td>164</td>
<td>55</td>
<td>0.643</td>
<td>40%</td>
<td>64%</td>
</tr>
<tr>
<td>Support Hours</td>
<td>Major Defects (mad)</td>
<td></td>
<td></td>
<td>( 89.802 + 26.7155 \times \text{mad} )</td>
<td>0.614</td>
<td>131</td>
<td>43</td>
<td>0.674</td>
<td>27%</td>
<td>60%</td>
</tr>
<tr>
<td>Total Maintenance &amp; Support Hours</td>
<td>Application Size(app)</td>
<td></td>
<td></td>
<td>( 451.95 + 0.382 \times \text{app} )</td>
<td>0.570</td>
<td>170</td>
<td>56</td>
<td>0.796</td>
<td>43%</td>
<td>70%</td>
</tr>
<tr>
<td>Total Maintenance Hours</td>
<td>Extreme Defects (exd)</td>
<td>Major Defects (mad)</td>
<td></td>
<td>( 160.733 + 52.734 \times \text{exd} - 1.757 \times \text{mad} )</td>
<td>0.569</td>
<td>134</td>
<td>45</td>
<td>0.388</td>
<td>42%</td>
<td>51%</td>
</tr>
<tr>
<td>Total Maintenance Hours</td>
<td>Extreme Defects (exd)</td>
<td></td>
<td></td>
<td>( 187.36 + 46.755 \times \text{exd} )</td>
<td>0.537</td>
<td>135</td>
<td>45</td>
<td>0.239</td>
<td>22%</td>
<td>56%</td>
</tr>
<tr>
<td>Total Support Hours</td>
<td>Extreme Defects (exd)</td>
<td></td>
<td></td>
<td>( 263.9 + 32.074 \times \text{exd} )</td>
<td>0.528</td>
<td>131</td>
<td>44</td>
<td>0.740</td>
<td>25%</td>
<td>41%</td>
</tr>
<tr>
<td>Maintenance Hours</td>
<td>Application Size(app)</td>
<td></td>
<td></td>
<td>( 285.421 + 0.173 \times \text{app} )</td>
<td>0.510</td>
<td>164</td>
<td>55</td>
<td>0.988</td>
<td>41%</td>
<td>41%</td>
</tr>
<tr>
<td>Total Maintenance &amp; Support Hours</td>
<td>Total Defects (td)</td>
<td></td>
<td></td>
<td>( 565.86 + 11.33 \times \text{td} )</td>
<td>0.510</td>
<td>164</td>
<td>55</td>
<td>0.825</td>
<td>35%</td>
<td>51%</td>
</tr>
<tr>
<td>Major Defects</td>
<td>Application Size(app)</td>
<td></td>
<td></td>
<td>( 10.00 + 0.0054 \times \text{APP} )</td>
<td>0.446</td>
<td>139</td>
<td>46</td>
<td>0.393</td>
<td>19%</td>
<td>50%</td>
</tr>
</tbody>
</table>
Lessons Learned & Future Works
Lessons Learned to support estimation

• In cases where the domains and project characteristics are similar:
  ▪ Industry specific benchmarks
  ▪ Programming language specific benchmarks
  ▪ Organizational benchmarks
  ▪ Rules of thumb for allocating maintenance and support efforts throughout the lifecycle
  ▪ Rules of thumb for predicting defect types for a project

• The resulting CERs, though not ideal, provide some insight into areas where the things we know about our project might help discover things unknown or only suspected:
  ▪ Potential predictors for Software Support
  ▪ Potential predictors for Software Maintenance
  ▪ Potential predictors for software program Defects
Lessons Learned about SW Data Collection

- Qualitative fields should not be freeform
- Application Set Information would be much more useful if companies and projects could be assigned (anonymously) for cross tracking
- Relationship between Total Hours and Hours for Sub-activities requires clarification
- Enhancement percent might be better replaced with a discrimination between baseline and new capability
- Latent defects would be a good attribute to collect
Future Work

• This research is worthy of a second pass, preferably with new data added to the mix

• Gain further clarity on discrepancies identified between Maintenance Hours and Support Hours and their component hour counts

• Additional data would provide an opportunity to create additional benchmark scenarios

• Application of machine learning techniques such as K-nearest neighbor and neural networks to this dataset
Questions?