Lessons Learned from Software Maintenance and Support Datasets

Abstract

Effort applied to the maintenance of software applications is thought to be between 60-90% [1] of the total ownership costs of the application. Yet there are significantly fewer studies focused on maintenance as compared to development, especially focused on public domain datasets. This paper presents results from a study focused on maintenance effort data from the International Software Benchmarking Standards Group - Maintenance and Support (ISBSG M&S) database.

Introduction

Estimating software sustainment costs continues to be an issue for many organizations, government agencies and their contractors. Software is not like hardware in that it does not wear out after a certain number of uses. Software is a much more malleable 'thing' than hardware; software developers are often asked to stretch and mold software solutions to make accommodations for limitations in the hardware or other software applications that are part of a system.

For the purpose of this study, software sustainment effort will be tracked against the following activities that occur once a software application has been first deployed. The clock on software sustainment starts when the software is first delivered into production and continues until the software application goes out of service. The categories of effort collected represent the data collected from the contributors to the ISBSG M&S database. These include:

- Maintenance Effort
  - Adaptive Maintenance
  - Corrective Maintenance
  - Perfective Maintenance
  - Preventive Maintenance
- Support Effort
  - Problem Investigation Support
  - Queries and Quick Service Support
  - User Help and Advice Support

The first section discusses in more detail what software sustainment is in the context of this study and reviews related research in this area. The second session provides an overview of the data mining process applied to this study. Following this is a brief description of the ISBSG organization and the data available for this study through this organization. The following sections describe how the data mining steps have been applied to this data as well as details of the results.

Software Sustainment

More and more systems are reliant on software for successful operations. There are many reasons for this. First and foremost is the need to keep up with ever improving technology options in both the hardware and software world. Due to budget constraints and the availability of money for research and development efforts, less new software is being developed while legacy software applications are being

corrected, enhanced, adapted and modernized in an effort to meet new threats, mission requirements, coalition configurations, etc .[1]. Software is often modified to accommodate changing requirements because it is easier to deploy than hardware. Due to this increased reliance on software and the need to make it last longer, software sustainment is a significant concern to all involved in fielding software intensive systems, consuming up to 60-90% of total costs (and effort) for many programs.

According to the Software Engineering Institute at Carnegie Mellon University (SEI/CMU) [2]:

"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."

According to the Institute of Electrical and Electronics Engineering's (IEEE) Standard 12207 [2]:

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

The software sustainment activities that are part of this study are those efforts collected for the ISBSG M&S Database. These activities are:

- Adaptive Maintenance – Enhancement necessary to accommodate changes in the environment in which a software product must operate.
- Corrective Maintenance – The reactive modification of a software product performed after delivery to correct discovered problem.
- Perfective Maintenance – The modification of a software application, after delivery, to improve performance or maintainability.
- Preventative Maintenance – The modification of a software application after delivery to detect and correct latent faults in the software product before they become effective faults.
- Maintenance Management – The time that cannot be directly attributed to the maintenance of an application (e.g., Administration)
- Problem Investigation Support – This is carried out to determine if a reported incident is in fact a defect, or an error in user documentation, training, or a user error.
- Queries & Quick Service Support – One tine questions that are not part of the application and can be delivered by the support team on request of the user
- User Hep & Advice Support – General support that is not related to an incident and does not involve any data extraction or manipulation.
- Support Management – The time that cannot be directly attributed to the support of an application (e.g., administration)

Many studies of software maintenance/software sustainment data towards better estimation of effort and/or cost have been done in the past. While an exhaustive search of the literature in this area has not been conducted as part of this study, it is clear that further research may be valuable to the industry. Papers such as [4],[5], and [8] present the results of regression-based analysis of data sets that don't appear to be public. Several authors have discussed the use of Open Source Software (OSS) and change histories as the basis for maintenance effort prediction as found in [7],[10] and [11]. [6] presents a review of estimation problems in Software Maintenance. The only work similar to this study is found in [9] where

the authors used a subset of the ISBSG M&S data to predict software maintenance and support costs. This study differs from their work in that it uses the 2020 version of the ISBSG database and includes more recent data.

Data Collection and Analysis Process

Performing data collection correctly is not easy. And it is often problematic even when the 'data' is right at your fingertips because there is often drama associated with taking data, especially data collected from disparate sources, and turning it into information and (hopefully) knowledge. Regardless of the problem being solved, there are many important steps that need to be taken in order to make sure that the data collection and analysis process is structured to efficiently collect the best set of data to satisfy your data quest. According to Wikipedia [3] The ability to apply sound data mining techniques provides an important tool in the toolkit of the cost analyst or cost estimator interested in conquering large and diverse data sets.

While the data mining and analysis process is intended to be iterative as more is learned about the data, the activities that are required for a successful data mining project are detailed below:

- **Understanding the current data challenge**

  This phase represents a significant part of any such project. There needs to be a purpose and the purpose needs to be understood and accepted by project stakeholders. To put this more simply, the first step in a data mining exercise is to understand what question (or questions) need to be answered and how these answers might be used. One would not start building software without first asking what the requirements this software must fulfill. Similarly, one should not start collecting and analyzing data without asking what problem(s) the data is expected to solve

- **Data Understanding and Preparation**

  Once the focus of the study has been identified, the next obvious step requires investigation into where and how that data might be available. Chances are good that the data required to answer a business-related question will requiring harvesting data from multiple sources within an organization or an industry. It is important that a data mining team begin to determine the data items that are likely to be required to answer the question(s) of the study and that are likely to be available for collection.

  Figuring out what data to collect and gathering it is just the start of the process. Data is often ugly, especially cost/effort data from software projects. In a perfect world, the data analyst would be able to collect spreadsheets replete with endless rows and columns rich with quantitative and qualitative information relative to the study being executed. While in a perfect world this may happen, it often doesn't; certainly, it can be especially rare in the cost estimating community. While there are many numerical values for analysis, there is often also important non-numerical context data that has the potential to be as valuable as some of the numerical data in helping understand and analyze the data. There are also many instances where datasets are incomplete, missing data, or suspicious in nature. Datasets may require filtering to remove pieces of data that are uninteresting or irrelevant or insufficiently available in the dataset to be of value for the analysis at hand.

- **Modeling**

Models can be used for classification or prediction, depending on the question to be answered. If one is trying to determine the best audience to target advertising for a new vehicle, they may want to create a model that classifies car owners based on previous purchases. If one is trying to predict the costs of developing or sustaining software, they will want to create a predictive model to accomplished this based on outcomes of similar software projects. Having said this, the cost model builders still may want to do some classification modeling before addressing the predictive problem, because as noted earlier, there are often context details that drive stratifications in data sets.

- **Evaluation**

  Clearly, once a model has been developed, it is important to make sure that the model makes sense. This evaluation involves testing the model developed against a set of data with known outcomes and ensuring that it behaves properly. It is important to review via statistical tests, the 'goodness' of any models developed in order to ensure credibility and to provide context, as to when it is, and is not, appropriate to use these models. It is desirable to hold back a part of the data set, when possible, to use as a test of the model developed to ensure that the model adequately responds to as many possible variations as possible. Equally important, the data mining team should always remember that sometimes the data is going to tell you something that is completely implausible; common sense should be an important tool any data mining evaluation tool set.

- **Deployment**

  Once the team is happy with the model created, the next big step is to introduce it to the stakeholders who need answers to the question originally posed. This phase may involve creating a physical implementation that makes the model easy for the end user to apply, documenting this implementation and training the end users on the proper ways to use it. End users must fully understand the limitations of model use as well; a model developed using only fixed wing aircraft data may be unsuitable for predicting the costs of a rotorcraft. Deployment may also involve some evangelization on the part of the data mining team to build acceptance and confidence.

The Data

The International Software Benchmarking Standards Group (ISBSG) is a not for profit organization that was founded in Australia in 1997 by a group of software metrics association with the intention to help the Information Technology (IT) industry improve software processes and practices. Since their inception they have been collecting industry data from organizations all over the world. 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 [12]

In August 2020 they release Version 7 of the Maintenance and Support database containing 1177 data points, from which the data sets used for this study have been extracted. For more information about the ISBSG go to www,isbsg.org. Data from the ISBSG database can be obtained by anyone although there is a fee to purchase the data. Since it requires purchase, this author can not publish the data points that were used in this study. However, each datapoint in the ISBSG database is assigned a random identification number. Appendix B lists all the data points used for each of the two datasets studied so that any interested reader with access to the study can reference them with respect to this work.

Data Study

Understanding the questions to be answered

The intent of this study is to determine what lessons can be learned from the ISBSG M&S data and to identify ways to use these lessons to provide software maintenance and support estimation guidance to the cost community. The goal is to provide as appropriate –benchmarks aligned with industries and programming languages, cost estimating relationships (CERs)for size, effort and defects, as well as to provide other useful guidance for the cost estimators.

Although the ISBSG data set is focused on Functional Size Measures (Function Points, COSMIC Function Points, etc.), the repository also allows for the collection oi KSLOC information for each of the projects reported. Another question to be answered is whether there is any indication whether KSLOC or FMs are decidedly better predictors of productivity for sustainment projects.

Preparing and understanding the data

The ISBSG M&S Database contains 1177 data points each with a potential 114 attributes or features. Appendix A contains a table that list all of the possible attribute as well as the number of data points for which that attribute is missing. Inspection of this table will help the reader understand why the datasets used in this analysis are significantly smaller than the entirety of the database.

Two datasets were collected for potential analysis – one which contained all the data points that listed IFPUG or NESMA as the Application Size Approach and one which contained a value greater than 0 for Entered KSLOC. The reason for these two choices for analysis was that IFPUG and NESMA are considered to be very comparable size measures and between them make up almost 60% of those projects that indicated an Application Size Approach in their submissions. The Entered KSLOC was selected for the KSLOC study rather than the Derived KSLOC because it was preferable that the analysis be based on real (hopefully) measured values than ones that have been derive through backfiring or some other technique.

Once the boundaries of the data study are established the next step is to prepare the data for analysis. Even though the two data sets are much smaller than the total amount of data points in the database, each is too large for data preparation and manipulation without a good set of tools. The tools used for this study include:

- RapidMiner – a data science software platform that provides an integrated environment for data preparation, machine learning, deep learning, text mining and predictive analytics. RapidMiner has an Open Source Community Version that can be downloaded and used for free (more information available at https://rapidminer.com/)

- TrueFindings® - a data analytics tool developed by PRICE® System specifically designed for data collection and analysis for cost estimators and analysts (more information available at www.pricesystem.com)
- Microsoft® Excel

Figure 1 shows the RapidMiner process developed for the first pass of data preparation.

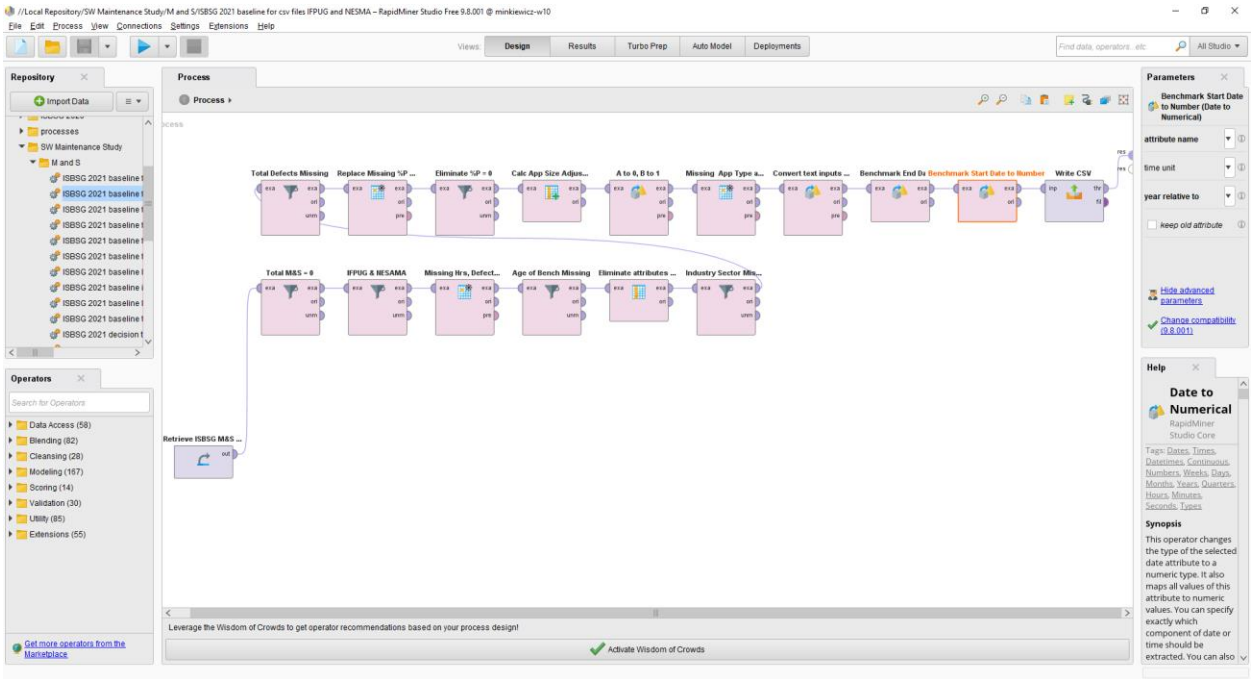


Figure 1 - RapidMiner Process for data Preparation

The steps in the process are as follows:

- 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. If Total M&S Hours is 0 that means no value was entered for either Maintenance Hours or Support Hours making these records unusable for this study
- Filter so that all data points that do not have Application Size Approach set to IFPUG or NESMA (or Entered KSLOC for the KSLOC dataset)
- For the records that remain set the following values to 0 if they are missing. While it is not always legitimate to set missing values to 0, in this case it was deemed sensible since many submitters only fill in the total value and many projects may not collect data for some of the specific values being collected (adaptive, preventive, etc.). The same logic applies for Defect related attributes.
    - Adaptive Maintenance Hours
    - Corrective Maintenance Hours
    - Extreme Defects
    - Major Defects
    - Management Maintenance Hours
    - Management Support Hours

  - Minor Defects
  - Perfective Maintenance Hours
  - Preventative Maintenance Hours
  - Problem Investigation Support Hours
  - Queries & Quick Service Support Hours
  - Total Maintenance Hours
  - Total Support Hours
  - User Help & Advice Support Hours
- Remove records where the Age of Benchmark attribute was missing. This was done because the calculation described in the Notes for this input (see Appendix A) seemed to address all possibilities of inferring this value from existing fields so there was no reasonable way to determine a good value for substituting the missing value.
- Eliminate attributes that are not needed for this study or are so scarce that including them does not make sense. Appendix C contains a list of the attributes eliminated along with the rationale.
- Filter out records missing Industry Sector. Industry Sector was well represented in the population and is likely to provide information.
- Eliminate records with Total Defects Missing
- Replace Missing Percent %P with 100 – %P is the percent of Primary Language in the application. A visual scan indicated that all records with %P missing also had no Secondary language.
- 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 Missing 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 both data sets

Once the data is prepared, the first step is to develop benchmarks by industry, application type, and programming language for effort, size, defects, and associated productivity rates. These benchmarks are intended to provide estimators with values appropriate for sanity checks and uncertainty analysis around software maintenance estimates. Data was loaded into TrueFIndings which was used to create the values for Tables 1 and 2. Figure 2 shows an example of the TrueFindings analysis capability.

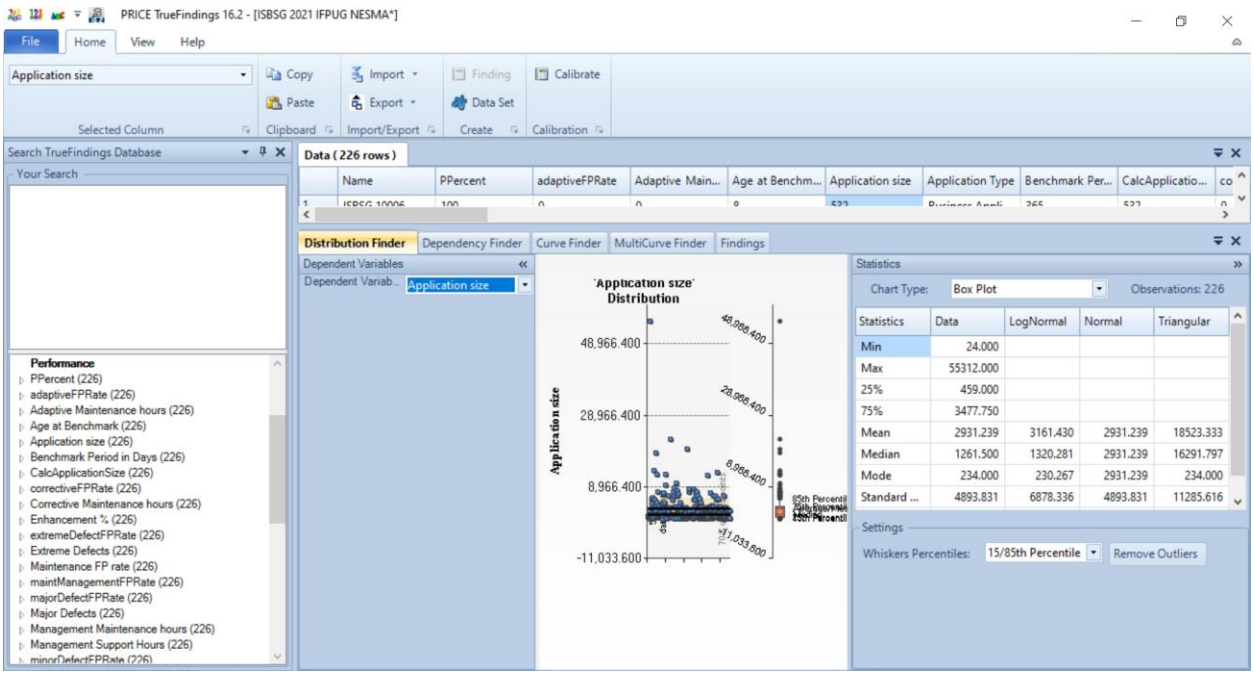


Figure 2- TrueFindings Benchmark Analysis for Application Size

Attribute	Sample Size	Min	Max	25%	75%	Mean	Median	Mode	Standard Deviation	Coefficient of Variation
Age at Benchmark	226	1	31	4	12	8.85	8	3	6.018	0.68
Adaptive Hours Per FP	30	0.002	3.894	0.055	0.423	0.44	0.197	0.002	0.768	1.746
Adaptive Hours	30	6	1032	21.75	388.75	235	69	18	304.77	1.297
Application Size	226	24	55312	459	3477.8	2931.24	1261.5	234	4893.831	1.67
Corrective FP Rate	41	0.002	4.292	0.038	0.327	0.356	0.095	0.002	0.742	2.083
Corrective Maintenance Hours	41	2	4224	24	132	320.707	48	6	843.871	2.631
Extreme Defects per FP	183	1.43E-04	0.066	0.002	0.006	0.005	0.004	0.002	0.005	1.12
Extreme Defects	183	1	152	2	15.5	12.778	5	2	18.525	1.45
Maintenance FP Rate	208	0.1	6.3	0.2	0.411	0.414	0.307	0.2	0.624	1.506
Major Defect FP Rate	185	3.10E-04	1	0.006	0.012	0.024	0.008	0.008	0.111	4.58
Major Defects	185	1	326.51	5	32	27.4	14	3	41.697	1.522
Management Maintenance Hours	172	2	3139	66.25	283.5	215.326	149	52	291.266	1.353
Minor Defects Per FP	193	3.10E-04	0.667	0.016	0.036	0.037	0.022	0.05	0.065	1.729
Minor Defects	193	1	954	19	82	69.658	41	12	121.581	1.745
Perfective Hours per FP	26	0.006	0.43	0.03	0.094	0.078	0.065	0.086	0.084	1.076
Preventive Maintenance Hours	34	1	240	6.75	76.5	60.853	31	6	71.711	1.178
Preventive FP Rate	34	0.002	0.747	0.034	0.113	0.116	0.068	0.1	0.16	1.377
Problem Investiageion FP Rate	31	0.003	2.113	0.041	0.234	0.226	0.098	0.003	0.408	1.802
Problem Imvestigation Hours	31	8	1237	23.5	106.5	112.258	60	13	220.839	1.967
Queries & Quick Service Support Hours	33	1	3752	13	132	243.909	67	13	656.753	2.693
Queries and Quick Service FP Rate	33	0.003	1.1	0.055	0.271	0.222	0.12	0.003	0.274	1.233
Support FP Rate	204	0.1	3.3	0.186	0.378	0.339	0.3	0.2	0.337	0.992
total Defects FP Rate	219	0.001	1.667	0.023	0.05	0.059	0.034	0.043	0.157	2.647
Total Defects	219	1	1311	22.5	111.5	96.831	51	32	159.462	1.647
Total Maintenance and Support Hours	226	11	22423	329.3	1759	1511.04	881.5	52	2457.856	1.627
Total Maintenance Hours	219	6	13182	181	936.5	780.174	452	126	1327.703	1.702
Total Support Hours	208	8	11212	170.8	841.25	687.611	453.5	40	1021.549	1.486
User Help and Advise Support Hours	27	3	495	12.5	99.5	72.074	33	13	100.605	1.396
User Help and Advice FP Rate	27	0.004	1.1	0.027	0.147	0.128	0.051	0.004	0.22	1.72

Table 1 - Benchmark values for Effort, Size, and Function Points Rates for IFPUG/NESMA dataset

Attribute	Sample Size	Min	Max	25%	75%	Mean	Median	Mode	Standard Deviation	Coefficient of Variation
Age at Benchmark	172	0	20	2.2	8	5.634	4	4	4.616	0.819
Adaptive Hours Per KSLOC	38	0.034	16.645	0.307	3.527	2.44	0.952	0.818	3.423	1.403
Adaptive Hours	38	6	15017	37.5	520.5	735.553	146	18	2424.936	3.297
Entered KSLOC	172	0.023	810000	12.7	208.77	5284.84	52.964	10	61751.61	11.685
Corrective FP Rate	67	0.006	30.976	0.325	4.353	3.809	1.924	0.006	5.56	1.46
Corrective Maintenance Hours	67	2	30976	37	1025	1566.08	142	6	4327.271	2.763
Extreme Defects per KSLOC	41	1.84E-05	2.079	0.008	0.304	0.235	0.067	1.84E-05	0.384	1.634
Extreme Defects	41	0.685	729.9	5	81.72	68.132	14.875	1	126.131	1.851
Maintenance KSLOC Rate	131	0.1	60.2	0.5	1.9	2.56	1	0.2	6.726	2.628
Maintenance KSLOC Rate	155	0.018	1298.04	2.759	31.271	45.925	9.308	1.059	148.677	3.237
Major Defect KSLOC Rate	44	6.91E-05	10.474	0.037	0.806	0.977	0.201	6.91E-05	2.045	2.093
Major Defects	44	1	1113.68	19.67	143.94	144.261	60.009	1	239.909	1.663
Minor Defects Per KSLOC	44	2.19E-04	4.467	0.024	0.427	0.47	0.117	2.19E-04	0.873	1.859
Minor Defects	44	1	4811	7.825	106.5	231.952	27.739	1	765.74	3.301
Perfective Hours per KSLOC	92	0.004	346.918	0.873	22.309	23.924	6.387	0.004	52.972	2.214
Perfective Hours	92	3	2500	65.75	393	361.761	130.5	6	534.15	1.477
Preventive Maintenance Hours	44	1	816	9	190.5	131.386	62.5	6	189.599	1.443
Preventive KSLOC Rate	44	0.012	3.727	0.185	0.791	0.67	0.431	0.012	0.809	1.208
Problem Investiageion KSLOC Rate	56	0.021	19.4	0.515	2.117	2.527	1.027	0.021	4.4	1.741
Problem Imvestigation Hours	56	2	19360	48.5	593	886.875	136.5	13	2748.464	3.099
Queries & Quick Service Support Hours	36	1	3872	24	204.5	347.778	95.5	13	760.742	2.187
Queries and Quick Service KSLOC Rate	36	0.037	58.2	0.31	1.23	2.559	0.913	0.037	9.577	3.743
Support KSLOC Rate	100	0.07	4173.91	0.987	6.631	58.806	2.287	0.422	425.019	7.228
total Defects FP Rate	136	3.07E-04	17.02	0.071	0.883	1.001	0.249	3.07E-04	2.099	2.096
Total Defects	136	1	5677	4.65	80	166.125	18.626	1	560.565	3.374
Total Maintenance and Support Hours	172	11	162597	236.3	2715.3	4976.36	553.5	21	16461.559	3.308
Total Maintenance and Support Hours	172	5.84E-04	5632	3.737	42.844	113.469	11.751	5.84E-04	550.532	4.852
Total Support Hours	100	2	23232	64.5	1040.5	1548.67	270.5	40	3236.819	2.09
User Help and Advise Support Hours	31	3	1418	13	134	190.774	66	13	338.915	1.777
User Help and Advice KSLOC Rate	31	0.037	32.333	0.214	1.345	1.843	0.619	0.037	5.712	3.099

Table 2- Benchmark values for Effort, Size, and Function Points Rates for KSLOC dataset

Perusal of these tables should give the reader an idea of the layout of the data sets. The IFPUG/NESMA dataset has 226 datapoints where the KSLOC dataset contains 172 data points. Since both datasets only contain records that have Total Maintenance and Support Hours and Application Size > 0 it is obvious that these attributes are well represented. It is also apparent that values for Total Maintenance Hours, Total Support Hours, and all of the Defect related attributes are fairly well populated. What is disappointing about what both these tables reveal is that the hours for the specific maintenance and support activities (Corrective Maintenance, Adaptive Maintenance User Help & Advice Support, etc.) are poorly

represented, making it clear that detailed analysis around these sub-activities is likely not the best use of analysis time.

Further scrutiny of the data using the filtering capabilities of TrueFindings did make it possible to create sets of benchmarking values for specific industries, organizations, and programming languages. The categories with adequate data to create such benchmarks are:

- 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 & Retail

Figures 3 and Table 2 show a sample of the stratified benchmarking values for the Industry Sector for ISBSG/NESMA for Government projects. Appendix D contains statistics and benchmarks for the other categories outlined above.

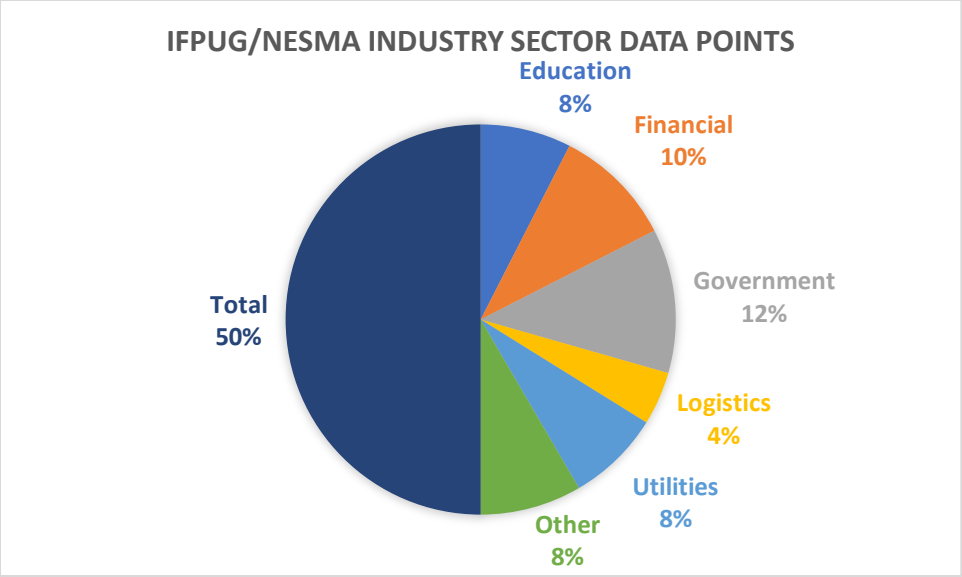


Figure 3 Industry Sector Data Points for IFPUG/NESMA Dataset

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

Table 3 - IFPUG/NESMA Industry Sector Benchmarks

Another study was conducted to try to ascertain how effort was typically split between the various sub activities of the maintenance and support activities as well as how the various types of defects were generally spread across the total number of defects. The results led to the discovery that only a fairly small percentage of the total hour counts and total defects were broken into sub categories. It also led to an observation that not all of the Total Maintenance or Total Support attributes that did have sub-activities represented the sum of these sub activities. In other words, the sum of Total Maintenance was greater than the sum of Adapted, Corrective, Perfective, Preventative and Management even when these values were present with similar effect for the Support Hours. Tables 4 through 7 provide a distribution for maintenance and support sub activities but only in the context of those activities where they were reported. It also led to the conclusion that some of the submitters had maintenance and support effort that did not fall into any of these categories and thus reported that as part of the Total Maintenance and Total Support Hours. For this reason, as well as due to the scarcity of sub activity information, the modeling for effort was focused on the Total Maintenance Hours, Total Support Hours and Total Maintenance and Support Hours.

Activity	Hours	Count	Percent of Total Hours	Percent of Total Records
Maintenance				
Adaptive	7050	30	12%	13%
Corrective	13149	41	21%	18%
Perfective	1877	26	3%	12%
Preventative	2069	34	3%	15%
Maintenance Management	37036	172	61%	76%
Total Maintenance Calculated	61181	216		96%
Total Maintenance Entered	170858	219		97%
Support				
Problem Investigation	3480	31	26%	14%
Queries & Quick Service	8049	33	60%	15%
User Help and Advice Support Hours	1946	27	14%	12%
Support Mangement	4	2	0%	1%
Total Support Calculated	13479	35	100%	15%
Total Support Entered	143023	208		92%

Table 4 - IFPUG/NESMA effort by maintenance and support sub-activities.

Defect Type	Total	Count	Percent of Total Defects	Percent of Total Records
Extreme	2338.46	183	11%	81%
Major	5069.09	185	24%	82%
Minor	13444.09	193	64%	85%
Total Calculated	20851.63			
Total Entered	21206.05	219		97%

Table 5 - IFPUG/NESMA Defects by sub-categories.

Activity	Hours	Count	Percent of Total Hours	Percent of Total Records
Maintenance				
Adaptive	27951	38	15%	22%
Corrective	104927	67	57%	39%
Perfective	33282	92	18%	53%
Preventative	4994	44	3%	26%
Maintenance Management	11409	5	6%	3%
Total Maintenance Calculated	182563	135		78%
Total Maintenance Entered	420078	155		
Support				
Problem Investigation	49665	56	73%	
Queries & Quick Service	12520	36	18%	
User Help and Advice Support Hours	5914	31	9%	
Support Mangement	2	1	0%	
Total Support Calculated	68101	59		
Total Support Entered	154867	100	58%	

Table 6- Table 4 - KSOC effort by maintenance and support sub-activities.

Defect Type	Total	Count	Percent of Total Defects	Percent of Total Records
Extreme	2793	41	14%	24%
Major	6347.00	44	33%	26%
Minor	10206	44	53%	26%
Total Calculated	19346			26%
Total Entered	22593	136		

Table 7 - KSLOC Defects by sub-categories.

Modeling and evaluation

Once the data is cleansed and normalized and perused, the next step is to embark on analysis to determine what predictive value these data sets contained. This focus of this phase was primarily on regression analysis. From preliminary review of correlation matrices for the two datasets, the first observation was that there was very little correlation of interest in the KSLOC dataset. This author believes that this may be due to the fact that the primary focus of most submitters is Functional Size measures, it is quite possible that many of the KSLOC counts are either calculated through backfiring or gotten from a code counter making it likely that nuances around reuse of baseline capability which are more likely to have inform the

Functional Size measures. Deeming this an area for further study, the KSLOC part of the modeling will be tackled in a future study when there is more data or more clarity into the Enterred KSLOC attribute. The KSLOC benchmarks presented earlier should still be useful if used with full knowledge of their statistical information.

While the full matrix was studied as part of the research effort, an abbreviated version of the IFPUG/NESMA Correlation matrix (eliminated attributes that displayed little to no correlation for ease of visualization) can be seen in Figure 4. It is interesting to note here that the study mentioned earlier in the paper which studied an earlier version of the ISBSG database [9] found that the most significant drivers of effort were Application Size, Defects and Team Size. The dataset analyzed in this exercise had very few samples that contains Team Size and it seemed to have very low correlation to any of the effort categories.

Attributes	Total M...	Perfecti...	Prevent...	Correcti...	Adaptiv...	Manage...	Total Su...	Proble...	Queries...	User He...	Manage...	Extrem...	Major D...	Minor D...	Age at ...	Total M...	Applicat...	Total De...	CalcAp...
Total Maintenance hours	1	0.140	0.152	0.253	0.045	0.661	0.810	0.129	0.002	0.101	-0.033	0.538	0.383	0.457	0.144	0.835	0.708	0.495	0.708
Perfective Maintenance hours	0.140	1	0.782	0.555	0.454	-0.140	-0.025	0.441	0.198	0.392	0.164	-0.115	-0.113	-0.009	-0.033	0.054	0.053	-0.037	0.052
Preventative Maintenance hours	0.152	0.782	1	0.625	0.522	-0.159	-0.034	0.550	0.175	0.513	-0.025	-0.138	-0.133	-0.024	-0.103	0.054	0.026	-0.055	0.027
Corrective Maintenance hours	0.253	0.555	0.625	1	0.102	-0.094	0.023	0.577	0.152	0.501	0.005	-0.057	-0.044	0.129	-0.027	0.137	0.144	0.079	0.142
Adaptive Maintenance hours	0.045	0.454	0.522	0.102	1	-0.140	-0.058	0.217	0.155	0.351	0.075	-0.132	-0.131	-0.116	-0.091	-0.011	-0.071	-0.112	-0.069
Management Maintenance hours	0.661	-0.140	-0.159	-0.094	-0.140	1	0.910	-0.105	-0.083	-0.127	-0.055	0.798	0.586	0.610	0.263	0.692	0.907	0.672	0.914
Total Support hours	0.810	-0.025	-0.034	0.023	-0.058	0.910	1	0.083	0.213	0.053	-0.058	0.768	0.674	0.642	0.230	0.807	0.890	0.720	0.884
Problem Investigation Support hours	0.129	0.441	0.550	0.577	0.217	-0.105	0.083	1	0.219	0.855	-0.005	-0.071	-0.098	-0.058	-0.096	0.094	-0.033	-0.068	-0.034
Queries & Quick Service Support hours	0.002	0.198	0.175	0.152	0.155	-0.083	0.213	0.219	1	0.209	-0.011	0.014	0.452	0.244	0.022	0.081	0.034	0.297	-0.007
User Help & Advice Support hours	0.101	0.392	0.513	0.501	0.351	-0.127	0.053	0.855	0.209	1	-0.001	-0.096	-0.119	-0.080	-0.111	0.065	-0.046	-0.089	-0.046
Management Support Hours	-0.033	0.164	-0.025	0.005	0.075	-0.055	-0.058	-0.005	-0.011	-0.001	1	-0.029	-0.012	-0.038	-0.068	-0.046	-0.024	-0.035	-0.027
Extreme Defects	0.538	-0.115	-0.138	-0.057	-0.132	0.798	0.768	-0.071	0.014	-0.096	-0.029	1	0.759	0.638	0.119	0.572	0.851	0.758	0.851
Major Defects	0.383	-0.113	-0.133	-0.044	-0.131	0.586	0.674	-0.098	0.452	-0.119	-0.012	0.759	1	0.724	0.117	0.509	0.682	0.854	0.656
Minor Defects	0.457	-0.009	-0.024	0.129	-0.116	0.610	0.642	-0.058	0.244	-0.080	-0.038	0.638	0.724	1	0.168	0.712	0.733	0.973	0.703
Age at Benchmark	0.144	-0.033	-0.103	-0.027	-0.091	0.263	0.230	-0.096	0.022	-0.111	-0.068	0.119	0.117	0.168	1	0.194	0.198	0.161	0.195
Total Maintenance & Support hours	0.835	0.054	0.054	0.137	-0.011	0.692	0.807	0.094	0.081	0.065	-0.046	0.572	0.509	0.712	0.194	1	0.752	0.712	0.735
Application size	0.708	0.053	0.026	0.144	-0.071	0.907	0.890	-0.033	0.034	-0.046	-0.024	0.851	0.682	0.733	0.198	0.752	1	0.795	0.998
Total Defects	0.495	-0.037	-0.055	0.079	-0.112	0.672	0.720	-0.068	0.297	-0.089	-0.035	0.758	0.854	0.973	0.161	0.712	0.795	1	0.767
CalcApplicationSize	0.708	0.052	0.027	0.142	-0.069	0.914	0.884	-0.034	-0.007	-0.046	-0.027	0.851	0.656	0.703	0.195	0.735	0.998	0.767	1

Figure 4 - Abbreviated Correlation Matrix for IFPUG/NESMA Dataset

RapidMiner was used to perform this analysis and was also an instrumental part of the regression analysis activity. Figure 5 shows a generalization of the Regression Process developed for this dataset.

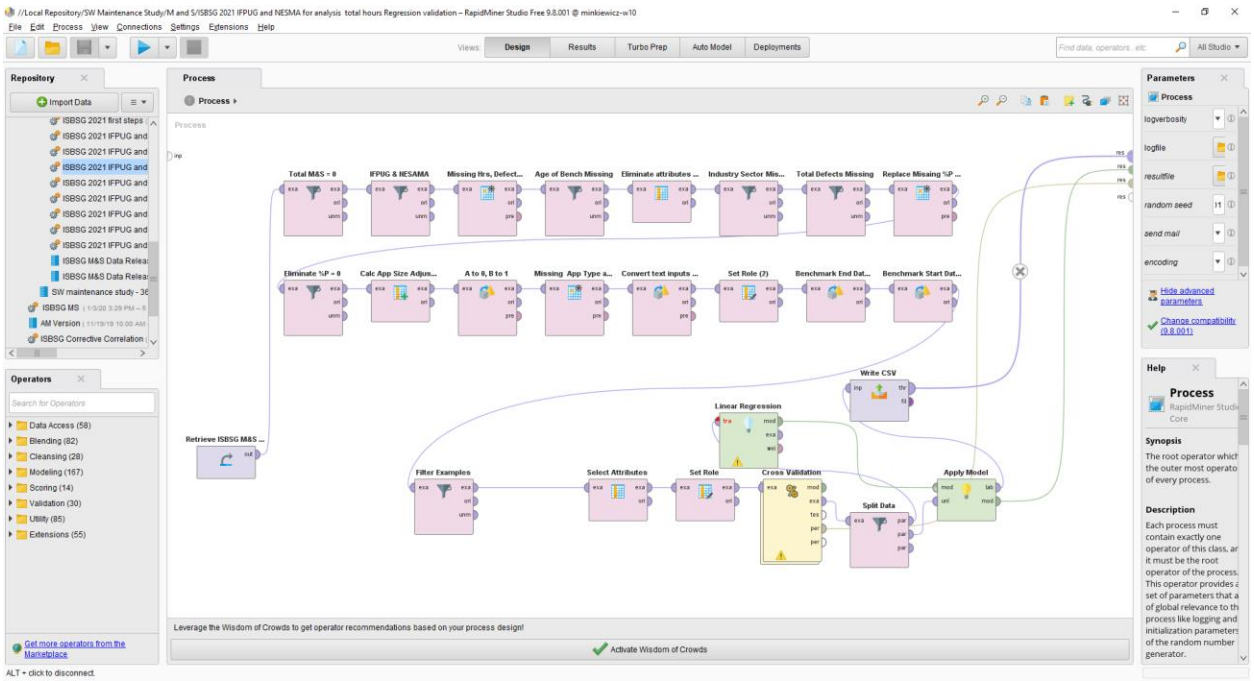


Figure 5 - RapidMiner Process for Linear Regression

The operators in the top two lines of this process should be familiar from the data preparation discussed earlier in this paper. The new operators that have been added to the process are described here

- 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 examples 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

This model was applied to various combinations of Independent and Dependent Variables guided by the information provided by the Correlation Matrix. Table 8 shows an analysis of the results of these iterations. The table provides analysis based on the performance of the model against the entire data set as well as the results that are provided by the application of the model to the scoring set. The columns in the spreadsheet are:

- Dependent Variable – the value to be predicted.
- Independent Variables – the values to be considered in the prediction.

- Equation – the output of the Linear Regression Model
- R2 (Equation) – the performance of the model on the entire dataset
- Training count – the number of data points in the training set
- Scoring Count – the number of data points in the scoring set
- R2 (prediction) – the performance of the model on the scoring set
- Pred(30) – Percentage of predictions that were with +/- 30 Percent of the Actual Value
- Pred(50) – Percentage of predictions that were within +/- 40 Percent of the Actual Value

				Training Results			Scoring Results		
Dependent Variable	Independent Variable 1	Independent Variable 2	Independent Variable 3	Equation	R2 (equation)	Training Count	Scoring Count	R2 (prediction)	Pred(30)	Pred(50)
Total Maintenance Hours	Application Size(app)	Extreme Defects (exd)	Major Defects (mad)	165.806+0.2469*app-12.171*exd-2.604*mad	0.949	134	45	0.888	62%	89%
Total Mainenance & Support Hours	Application Size(app)	Extreme Defects (exd)	Major Defects (mad)	226.824+0.3977*app-0.54477*exd-3.6202*mad	0.927	135	45	0.483	58%	76%
Total Support Hours	Total Maintenance Hours(totMaint)			202.89+0.621*totMaint	0.910	152	51	0.959	47%	76%
Total Support Hours	Application Size(app)	Extreme Defects (exd)	Major Defects (mad)	125.929 +0.2366*app-5.015*exd-5.125*mad	0.879	117	39	0.879	60%	88%
Total Support Hours	Application Size(app)			85.746+0.193*app	0.855	164	55	0.528	56%	78%
Extreme Defects	Application Size(app)			2.81+0.0029*app	0.723	137	46	0.792	30%	52%
Total Support Hours	Total Defects (td)			181.055+5.759*td	0.700	151	50	0.571	30%	54%
Total Maintenance Hours	Total Defects (td)			383.085 + 4.828*td	0.660	158	53	0.593	22%	45%
Total Support Hours	Extreme Defects (exd)	Major Defects (mad)		90.07+10.798*exd+20.812*mad	0.659	131	43	0.684	28%	65%
Total Defects	Application Size(app)			314.571+31.084	0.632	164	55	0.643	40%	64%
Support Hours	Major Defects (mad)			89.802+26.7155*mad	0.614	131	43	0.674	27%	60%
Total Maintenance & Support Hours	Application Size(app)			451.95+0.382*app	0.570	170	56	0.796	43%	70%
Total Maintenance Hours	Extreme Defects (exd)	Major Defects (mad)		160.733+52.734*exd-1.757*mad	0.569	134	45	0.388	42%	51%
Total Maintenance Hours	Extreme Defects (exd)			187.36+46.755*exd	0.537	135	45	0.239	22%	56%
Total Support Hours	Extreme Defects (exd)			263.9+32.074*exd	0.528	131	44	0.740	25%	41%
Maintenance Hours	Application Size(app)			285.4212+0.173*app	0.510	164	55	0.988	41%	41%
Total Maintenance & Support Hours	Total Defects (td)			565.86+11.33*td	0.510	164	55	0.825	35%	51%
Major Defects	Application Size(app)			10.00+0.0054*APP	0.446	139	46	0.393	19%	50%

Table 8 IFPU/NESMA Regression Results

The regressions are sorted by the model that performed best across the entire dataset. It is interesting to note that the Total Support Hours are most strongly influence by Application Size but the application of Defect Information makes the relationship stronger. It is not surprising to note that Total Defects are fairly good at predicting maintenance hours even though corrective maintenance is not a significantly greater percentage of the total maintenance costs (per the percentage information shown in Table 4. This is likely to reflect the fact that all maintenance activities are more expensive if the code is lower quality. It also makes sense that Total Defects show some predictive power for Total Support Hours as lower quality software also requires additional support.

It's also important to note that the results truly speak for themselves. Even though the dataset is fairly large, the fact that the data comes from disparate sources and although the definitions of the attributes are well spelled out in the notes, it is quite likely that different submitters interpret them differently. The most important thing to note is how different statistics represent analysis differently. It is not unusual for the performance of the scoring test to be lower than the performance of the model on the total dataset, it's more disturbing when the scoring results outperform the training results. Inspection of the data can explain this, but inspection is sometimes required. It is also important to note how different measures of goodness give very different interpretations of many of these models

Deployment

This paper and the subsequent presentation represent the first step in deployment for the results of this research While the results are not sterling, there is information from this study that can be applied to software maintenance estimates in appropriate situations. Clearly the use needs to be tempered by an understanding what the data does and does not tell us. But certainly, additional benchmarks and CERS can become part of an estimators toolbox to be pulled out when they fit the need.

After presentation and refinements on feedback, as well as incorporation of any additional data identified, the hope is that these benchmarks and possibly some of the CERs can find their way into guidance, rules of thumbs or findings for the cost community.

Lessons Learned

Starting with the obvious, the study did result in a better understanding of what the possibilities are for the ISBSG M&DS database. While clearly this dataset has limitations, it definitely gives a basis for the development of benchmarks and rules of thumbs that could prove useful in sanity checks for defensibility 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

The ISBSG database, though fraught with limitations so many other cost or effort collection repositories also suffer from, still, in my opinion, is an excellent source for quality data for the software industry to use to provide benchmarks to help learn about the industry as a whole and how individual industries can improve their software processes and practice. There are however some limitations that complicated this study for me and is likely to confound others trying to maximize the value they get from this plentiful, validated source. Not just ISBSG, but any organization looking to begin or improve a data collection program may want to consider these as well.

- Several of the qualitative fields, such as Application Type and Application Description are completely freeform. While it is appropriate to provide space for submitters to add extra narrative to their submissions, having a fixed pick list for classifications important to benchmark creation and possibly CER creation would add value.
- Information about Application Set, Application Number and Application Number of would be incredibly more useful if the entries some the same project and/or the same organization were included (in an anonymous term of course) so that more analysis might be possible within types of organization. These attributes seem to be fairy well populated, but the inter-related ness of the concept is hard to work around.

- Referencing the issue where Total Maintenance Hours is not always equal to the sum of the aum of the hours of the sub-activities, maybe there is a need to allow for Other Maintenance Hours as an attribute. Understanding that it is difficult to get people to contribute data and that taking what one can get is often the best option, it would be easier to use and understand if all hours were overtly specified.
- Instead of enhancement % it might be useful to characterize Application Size as Baseline and New capability
- Since there projects are being performed on existing systems, having an attribute of latent defects (those that were known to be in the product at deployment) would be useful for analysis

Further Work

This research certainly is worthy of a second pass, preferably with some newer data added to the mix. Increased clarity on the relationships between totals and sub-categories for effort and defects would potentially enhance the analysis and improve the results. Additional data would provide an opportunity to extend the categories of benchmarks as well. As part of this study, this author began to run some excursions using some of the Machine Learning capabilities in RapidMiner such as K-Nearest Neighbor and Neural Networks with some interesting results but due to time constraints and learning curve the results are not at this time ready for prime time. Future work in this area will certainly focus on determining some innovative approaches to learning more from this data.

Cited Works

[1] McLendon, et. al., "Addressing Software Sustainment Costs for the DoD", Crosstalk Magazine, January/February 2014, available at (retrieved 3/2018): http://static1.1.sqspcdn.com/static/f/702523/24156564/1388991346767/201401-McLendon.pdf?token=2sB1RBJakWI3Azmvvrh9vNGWs%2BY%3D

[2] "DAU Acquisition Encyclopedia, Software Sustainment", available at https://www.dau.mil/acquipedia/Pages/ArticleDetails.aspx?aid=544442b8-70d3-4a5c-afd4-42ffab0a9f41, December, 2017, retrieved March 2018

[3] https://en.wikipedia.org/wiki/Data_mining#Situation_in_the_United_States

[4] I. Buchmann, S. Frischbier and D. Putz, "Towards an Estimation Model for Software Maintenance Costs," *2011 15th European Conference on Software Maintenance and Reengineering*, Oldenburg, Germany, 2011, pp. 313-316, doi: 10.1109/CSMR.2011.45 [12] www.isbsg.org

[5] V. Nguyen, "Improved size and effort estimation models for software maintenance," 2010 IEEE International Conference on Software Maintenance, Timisoara, Romania, 2010, pp. 1-2, doi: 10.1109/ICSM.2010.5609554.

[6] S. S. Ali, M. Shoaib Zafar and M. T. Saeed, "Effort Estimation Problems in Software Maintenance – A Survey," *2020 3rd International Conference on Computing, Mathematics and Engineering Technologies (iCoMET)*, Sukkur, Pakistan, 2020, pp. 1-9, doi: 10.1109/iCoMET48670.2020.9073823.

[7] H. Wu, L. Shi, C. Chen, Q. Wang and B. Boehm, "Maintenance Effort Estimation for Open Source Software: A Systematic Literature Review," 2016 IEEE International Conference on Software Maintenance and Evolution (ICSME), Raleigh, NC, USA, 2016, pp. 32-43, doi: 10.1109/ICSME.2016.87.

[8] V. Nguyen, B. Boehm and P. Danphitsanuphan, "Assessing and Estimating Corrective, Enhancive, and Reductive Maintenance Tasks: A Controlled Experiment," *2009 16th Asia-Pacific Software Engineering Conference*, Batu Ferringhi, Malaysia, 2009, pp. 381-388, doi: 10.1109/APSEC.2009.49.

[9] J. Aziz, F. Ahmed and M. S. Laghari, "Empirical Analysis of Team and Application Size on Software Maintenance and Support Activities," 2009 International Conference on Information Management and Engineering, Kuala Lumpur, Malaysia, 2009, pp. 47-51, doi: 10.1109/ICIME.2009.51.

[10] B. Dit, A. Holtzhauer, D. Poshyvanyk and H. Kagdi, "A dataset from change history to support evaluation of software maintenance tasks," *2013 10th Working Conference on Mining Software Repositories (MSR)*, San Francisco, CA, USA, 2013, pp. 131-134, doi: 10.1109/MSR.2013.6624019.

[11] A. Kaur, K. Kaur and K. Pathak, "Software maintainability prediction by data mining of software code metrics," *2014 International Conference on Data Mining and Intelligent Computing (ICDMIC)*, Delhi, India, 2014, pp. 1-6, doi: 10.1109/ICDMIC.2014.6954262.

Appendix A – ISBSG M&S Attributes (for the 487 with a Quality Rating of A or B

Attribute	Comment	Nbr missing from the 487 A&B rows
ISBSG ID	ISBSG ID A primary key, for identifying submissions to the ISBSG repository. (These Identification numbers have been 'randomised' to remove any chance of identifying a submitter).	0
Data Quality Rating	ISBSG Rating This field contains an ISBSG rating code of A, B, C or D applied to the data by the ISBSG quality reviewers to denote the following: A = The data submitted was assessed as being sound with nothing being identified that might affect its integrity. B = The submission appears fundamentally sound but there are some factors which could affect the integrity of the submitted data. C = Due to significant data not being provided, it was not possible to assess the integrity of the submitted data. D = Due to one factor or a combination of factors, little credibility should be given to the submitted data.	0
Year	Year of Data Derived from Benchmark Period If no Benchmark Period date known, it is the year of ISBSG receipt of the data.	0
Age at Benchmark	Age of Application when benchmark was performed Derived from the Application Launchdate and Benchmark Period Start If no Benchmark Period start date known, then the date of submission to the ISBSG is used to calculate age. Age is not calculated where Application Launchdate is not known, or where the Launchdate is after the Benchmark Period start date.	73
Scaling factor	Scaling required To allow meaningful comparison of application data with the ISBSG repository, effort must be available for a yearly Benchmark Period. Where the period of benchmark of the data submitted is not a period of 12 months, a scaling factor, deived from your period of benchmark, hasbeen applied to effort values to enable this comparison to the repository. This is the scaling factor required to normalise data in Repository to a one year Benchmark Period.	0
Total Maintenance hours	Calculated Maintenance hours The total maintenance work effort in hours for a 12 month period, converted to hours where recorded otherwise. Sum of perfective, preventative, corrective, adaptive and management hours.	34
Perfective Maintenance hours	Perfective Maintenance Perfective maintanance is the modification of a software application, after delivery, to improve performance or maintainability. Values are in hours for a 12 month period.	356
Preventative Maintenance hours	Preventative Maintenance The modification of a software application after delivery to detect and correct latent faults in the software product before they become effective faults. Values are in hours for a 12 month periods.	408
Corrective Maintenance hours	Corrective Maintenance The reactive modification of a software product performed after delivery to correct discovered problems. Values are in hours for a 12 month period	348
Adaptive Maintenance hours	Adaptive Maintenance Enhancements necessary to accommodate changes in the environment in which a software product must operate. Values are in hours for a 12 month period.	423
Management Maintenance hours	Maintenance Management The time that cannot be directly attributed to the maintenance of an application (eg. Administration). Values are in hours for a 12 month period.	292
Total Support hours	Calculated Support hours The total support work effort in hours for a 12 month period, converted to hours where recorded otherwise. Sum of problem investigation, queries & quick support, user help & advice and management.	185
Problem Investigation Support hours	Problem Investigation: This is carried out to determine if a reported incident is in fact a defect, or an error in user documentation, training or a user error. Values are in hours for a 12 month period.	401
Queries & Quick Service Support hours	Queries & Quick Service One-time questions that are not part of the application and can be delivered by the support team on request of the user. Values are in hours for a 12 month period.	423
User Help & Advice Support hours	User Help & Advice General support that is not related to an incident and does not involve any data extraction or manipulation. Values are in hours for a 12 month period.	431
Management Support Hours	Management Support The time that cannot be directly attributed to the support of an application (eg. Administration). Values are in hours for a 12 month period	475
Total Maintenance & Support hours	Calculated Total hours The total work effort in hours for a 12 month period, converted to hours where recorded otherwise.	3
Maintenance FP rate	Maintenance FP rate Application maintenance rate calculated as Maintenance Effort (in hours) normalised to one year period, divided by Application Size (in FP) Maintenance rate is expressed as: hrs per fp	80
Support FP rate	Support FP rate Application support rate calculated as Support Effort (in hours) normalised to one year period, divided by Application Size (in FP) Support rate is expressed as: hrs per fp	198
Functional rate	Functional rate Application rate calculated as Total Effort (in hours) normalised to one year period, divided by Application Size (in FP) Functional rate is expressed as: hrs per fp	63
Maintenance KSLOC rate	Maintenance KSLOC rate Application maintenance rate calculated as Maintenance Effort (in hours) normalised to one year period, divided by Application Size (in KSLOC) Maintenance KSLOC rate is expressed as: hrs per KSLOC	289
Support KSLOC rate	Maintenance KSLOC rate Application support rate calculated as Support Effort (in hours) normalised to one year period, divided by Application Size (in KSLOC) Maintenance KSLOC rate is expressed as: hrs per KSLOC	360

KSLOC rate	KSLOC rate Application rate calculated as Total Application Effort (in hours) normalised to one year period, divided by Application Size (in KSLOC) KSLOC rate is expressed as: hrs per KSLOC	263
Enhancement %	Enhancement % Percentage of maintenance time spent on small enhancements Enhancement % is calculated only where the breakdown of maintenance effort is available.	190
Team Size	Team Size Application team size of the maintenance and support team	351
User Organisation type	User Organisation type The industry(s) in which the software's end users work The type of business in which the organisation is mainly involved (e.g.. Banking, Insurance, Manufacturing, Financial services, Service industry and so on)	0
Industry sector	Industry sector Grouping key for submissions by industry type of the organisation running the software This is derived from the organisation type, and is a sub-set of organisation type	8
Benchmark start date	Benchmark start date Start date of benchmark period The recommendation is that a benchmark submission should contain data for a period of 12 months, though other periods are submitted. Data for alternative periods are scaled appropriately before analysis	0
Benchmark end date	Benchmark end date End date of benchmark period The recommendation is that a benchmark submission should contain data for a period of 12 months, though other periods are submitted. Data for alternative periods are scaled appropriately before analysis	0
FTE hours	FTE hours Number of worked hours per Full Time Equivalent (FTE), for the Maintenance and Support Staff Year This is the number of hours worked by a full time member of maintenance and support staff in a complete year. This measure is used to 'normalise' data so that comparisons can be made. With data being contributed by different companies and indeed from different countries it is necessary to establish an unambiguous measure for effort. An hour has been adopted, being a constant period wherever data originates	208
Staff size	Staff size Complete Software Maintenance and Support Staff size The number of actual staff employed (including those that are not FTE) upon the maintenance and support activities set out in the activity model for the whole of that part of the organisation being measured	288
Portfolio	Portfolio Total Number of applications in the Portfolio This is the total of all applications in the portfolio of the organisation or department submitting the data	456
Applications - small	Applications - small Number of small applications in the Portfolio If the Portfolio size is not known, it may be useful to group applications by size: small <400 FP medium up to 1000 FP large > 1000 FP	473
Applications - medium	Applications - medium Number of medium applications in the Portfolio If the Portfolio size is not known, it may be useful to group applications by size: small <400 FP medium up to 1000 FP large > 1000 FP	471
Applications - large	Applications - large Number of large applications in the Portfolio If the Portfolio size is not known, it may be useful to group applications by size: small <400 FP medium up to 1000 FP large > 1000 FP	472
Portfolio size	Portfolio size The total combined size of all applications in the organisation This is one of the 'normative' measures used to draw comparisons. The size should be expressed in one of the full or interim status sizing methods accepted by ISBSG (see Portfolio size approach)	458
Portfolio size approach	Portfolio size approach The sizing method used to size applications in the organisation This should be one of the full or interim status sizing methods accepted by ISBSG: . IFPUG Function Points . COSMIC Function Points . NESMA Function Points . FiSMA Function Points . MKII Function Points . Feature Points . Lines of code measured in KSLOC	458
Total Maintenance FTE	Total Maintenance FTE Total Software Application Maintenance Effort for the organisation This is the sum of: . Perfective Maintenance . Preventative Maintenance . Corrective Maintenance . Adaptive Maintenance	119
Total Support FTE	Total Support FTE Total Software Support Effort for the organisation This is the sum of: . Problem Investigation . Help and Advice . Queries and Quick services	118
Total M&S FTE	Total M&S FTE Total Software Application Maintenance and Support Effort for the organisation This is the sum of: . Perfective Maintenance . Preventative Maintenance . Corrective Maintenance . Adaptive Maintenance . Problem Investigation . Help and Advice . Queries and Quick services	122
Application Description	Application Description Description of the software application As submitted to give some idea for research analysis	327

Application size	Application size The functional size of the application which is a major factor in maintenance effort The size should be expressed in one of the full or interim status sizing methods accepted by ISBSG (see Application size approach)	51
Application size approach	Application size approach The sizing method used to size the application This should be one of the full or interim status sizing methods accepted by ISBSG: . IFPUG Function Points . COSMIC Function Points . NESMA Function Points . FiSMA Function Points . MKII Function Points . Feature Points Note this is Functional Size, not the Technical Size which is in lines of code	33
Other size	Other size The functional size of the application where the method used is not listed by ISBSG The size has not been expressed in one of the full or interim status sizing methods accepted by ISBSG (see Other size approach)	467
Other size approach	Other size approach The sizing method used to size the application The size has not been expressed one of the full or interim status sizing methods accepted by ISBSG: . IFPUG Function Points . COSMIC Function Points . NESMA Function Points . FiSMA Function Points . MKII Function Points . Feature Points Note this is Functional Size, not the Technical Size	441
Launch date	Launch date The application launch date which is the date or year of the last major rewrite or re-launch This, combined with start date of benchmark period gives the application age In general it appears that the older an application is the greater is the maintenance effort required to keep it in service. There are known to be applications that are very old but are not being enhanced; in these cases maintenance effort is usually quite low	242
Not important process	Not important process Indicates where the application is not an important process to the business This set of data identifies the importance of the application for the business. Process here means a business process within the organisation	480
Important process	Important process Indicates where the application is an important process to the business This set of data identifies the importance of the application for the business. Process here means a business process within the organisation	309
Critical process	Critical process Indicates where the application is a critical process to the business This set of data identifies the importance of the application for the business. Process here means a business process within the organisation	256
Business target	Business target Indicates where the application is important for business targets This set of data identifies the importance of the application for the business. Business target here means any target set by the business or organisation	296
Critical target	Critical target Indicates where the application is a critical for business targets This set of data identifies the importance of the application for the business. Business target here means any target set by the business or organisation	271
Benchmark Period (Days)	Benchmark Period (Days) The number of days during which the maintenance and support effort was recorded.	0
Total Maint Effort	Total Maint Effort Effort devoted to the application, on various maintenance activities of: . Perfective maintenance . Preventative maintenance . Corrective maintenance . Adaptive maintenance These maintenance activities are defined in the Glossary of Terms that forms part of the Data Collection Questionnaire	35
Unit of Maint	Unit of Maint The unit of the maintenance effort The effort should be expressed as FTE staff years (with FTE hours/year given) or hours Where expressed in days or months (or where FTE and FTE hours/year is not given) a conversion is used: Staff Day = Hours divided by 6 Staff Months = Hours divided by 120 Staff Years = Hours divided by 1440 It is essential that all effort values are expressed in the same unit and that this unit is indicated	10
Total Support Effort	Total Support Effort Effort devoted to the application, on various support activities of: . Problem Investigation . Queries & Quick Service . User Help and Advice These support activities are defined in the Glossary of Terms that forms part of the Data Collection Questionnaire	156
Unit of Support	Unit of Support The unit of the support effort The effort should be expressed as FTE staff years (with FTE hours/year given) or hours Where expressed in days or months (or where FTE and FTE hours/year is not given) a conversion is used: not given) a conversion is used: Staff Day = Hours divided by 6 Staff Months = Hours divided by 120 Staff Years = Hours divided by 1440 It is essential that all effort values are expressed in the same unit and that this unit is indicated	98
Total M&S Effort	Total Maint Effort Effort devoted to the application, on various maintenance and support activities of: . Perfective maintenance . Preventative maintenance . Corrective maintenance . Adaptive maintenance . Problem Investigation . Queries & Quick Service . User Help and Advice These activities are defined in the Glossary of Terms that forms part of the Data Collection Questionnaire	0
Unit of effort	Unit of effort The unit of the total effort The effort should be expressed as FTE staff years (with FTE hours/year given) or hours Where expressed in days or months (or where FTE and FTE hours/year is not given) a conversion is used: Staff Day = Hours divided by 6 Staff Months = Hours divided by 120 Staff Years = Hours divided by 1440 It is essential that all effort values are expressed in the same unit and that this unit is indicated	35

Product versions	Product versions Number of versions of the product The higher the number of versions of the product, the higher is the expectation of maintenance/support effort, as this would influence testing effort	430
Operating systems	Operating systems Number of Operating systems supported The higher the number of Operating systems simultaneously supported by the product, the higher is the expectation of maintenance/support effort, as this would influence testing effort	468
Data bases	Data bases Number of data bases supported The higher the number of Databases supported by the product, the higher is the expectation of maintenance/support effort, as this would influence testing effort	468
Third Party Products	Third Party Products The contribution of third party products such as packages to the application being supported. It is expected that parameterised packages may be easier to adapt than bespoke applications. If third party products are used information may be supplied on: · Language(s) used, · Whether package or bespoke · If code is supplier maintained or maintained in-house · If parameters are maintained by the users or in-house or not at all	452
Product name	Product name If third party products are used this is the name of the product	482
Language used	Language used The language used in third party products Where more than one is recorded, these are numbered (1) (2) up to (4)	475
Third Party Package	Third Party Package If third party products are used this indicates whether this is: . Package or bespoke	481
Code maintained	Code maintained If third party products are used this indicates if code is: . maintained by the supplier or . maintained in-house	481
Parameter maintained	Parameter maintained If third party products are used this indicates if parameters are: . maintained by the users or . in-house or . not at all	481
Third Party Ownership	Third Party Product Ownership This indicates the extent of ownership of third party product Partial Ownership could be a possibility where for a very critical application, high severity defect correction is not out-sourced. This is relevant because in the case of partial ownership, then for the same KLOC, the effort expended in maintenance & support could be lower (e.g.. If the application is partially owned such that maintenance and support is 40% provided by the application provider and 60% provided by the organisation, then the maintenance and support effort for this application incurred by the organisation will be divided by 0.6 to give the full effort.	463
Partial Effort Scaled	Partial Effort Scaled If third party product ownership is either total or partial, then effort size and so on should be scaled This indicates if the effort is scaled or not	487
Hardware	Hardware The hardware used in the primary platform Primary platform is the platform on which the bulk of the application resides. It is recognised that in many modern distributed applications there may be several levels. Four levels may be described. Where more than one level is recorded, these are numbered (1) (2) up to (4)	219
HW number	HW number This is the number of hardware units at this level of the platform Four levels may be described. Where more than one level is recorded, these are numbered (1) (2) up to (4)	478
Operating System	Operating System The operating system on which the larger proportion of the application function is hosted As an example in a client server where there is little functionality residing on the client, the relevant operating system will be that which is operating on the server Four levels may be described. Where more than one level is recorded, these are numbered (1) (2) up to (4)	436
DBMS	Database Management System DBMS is the database management system used Four levels may be described. Where more than one level is recorded, these are numbered (1) (2) up to (4)	405
Platform Role	Platform Role Platform role is used to describe, as far as possible, the role the platform The role or function provided by a hardware platform to an application. In client/server or distributed software applications, one or more computers provide services to software operating on other computers. Platform role examples are: Server, Client, Core System, Warehouse, Infrastructure Four levels may be described. Where more than one level is recorded, these are numbered (1) (2) up to (4)	458

Primary language	Primary language The primary programming language that is used to program the bulk of the application	94
KSLOC (P)	KSLOC (P) The primary programming language lines of code Amount of code in number of KSLOC of the primary language (Note: this is thousand lines of code, not units of lines of code)	185
% (P)	% (P) The primary programming language expressed as a percentage figure Amount of code expressed as a percentage figure of the total lines of code	292
Secondary language	Secondary language The secondary programming language that is used to program the application	324
KSLOC (S)	KSLOC (S) The secondary programming language lines of code Amount of code in number of KSLOC of the secondary language (Note: this is thousand lines of code, not units of lines of code)	290
% (S)	% (S) The secondary programming language expressed as a percentage figure Amount of code expressed as a percentage figure of the total lines of code	292
Third language	Third language The third programming language that is used to program the application	376
KSLOC (3)	KSLOC (3) The third programming language lines of code Amount of code in number of KSLOC of the third language (Note: this is thousand lines of code, not units of lines of code)	318
% (3)	% (3) The third programming language expressed as a percentage figure Amount of code expressed as a percentage figure of the total lines of code	323
4th language	4th language The fourth programming language that is used to program the application	457
KSLOC (4)	KSLOC (4) The fourth programming language lines of code Amount of code in number of KSLOC of the fourth language (Note: this is thousand lines of code, not units of lines of code)	382
% (4)	% (4) The fourth programming language expressed as a percentage figure Amount of code expressed as a percentage figure of the total lines of code	387
5th language	5th language The fifth programming language that is used to program the application	481
KSLOC (5)	KSLOC (5) The fifth programming language lines of code Amount of code in number of KSLOC of the fifth language (Note: this is thousand lines of code, not units of lines of code)	403
% (5)	% (5) The fifth programming language expressed as a percentage figure Amount of code expressed as a percentage figure of the total lines of code	405
6th language	6th language The sixth programming language that is used to program the application	482
KSLOC (6)	KSLOC (6) The sixth programming language lines of code Amount of code in number of KSLOC of the sixth language (Note: this is thousand lines of code, not units of lines of code)	404
% (6)	% (6) The sixth programming language expressed as a percentage figure Amount of code expressed as a percentage figure of the total lines of code	406
7th language	7th language The seventh programming language that is used to program the application	484
KSLOC (7)	KSLOC (7) The seventh programming language lines of code Amount of code in number of KSLOC of the seventh language (Note: this is thousand lines of code, not units of lines of code)	404
% (7)	% (7) The seventh programming language expressed as a percentage figure Amount of code expressed as a percentage figure of the total lines of code	408
8th language	8th language The eighth programming language that is used to program the application	485
KSLOC (8)	KSLOC (8) The eighth programming language lines of code Amount of code in number of KSLOC of the eighth language (Note: this is thousand lines of code, not units of lines of code)	405
% (8)	% (8) The eighth programming language expressed as a percentage figure Amount of code expressed as a percentage figure of the total lines of code	406
9th language	9th language The ninth programming language that is used to program the application	485
KSLOC (9)	KSLOC (9) The ninth programming language lines of code Amount of code in number of KSLOC of the ninth language (Note: this is thousand lines of code, not units of lines of code)	405
% (9)	% (9) The ninth programming language expressed as a percentage figure Amount of code expressed as a percentage figure of the total lines of code	406
10th language	10th language The tenth programming language that is used to program the application	485
KSLOC (10)	KSLOC (10) The tenth programming language lines of code Amount of code in number of KSLOC of the tenth language (Note: this is thousand lines of code, not units of lines of code)	405
% (10)	% (10) The tenth programming language expressed as a percentage figure Amount of code expressed as a percentage figure of the total lines of code	406

Appendix B – ISBSG Data Points used in IFPUG/NESMA and KSLOC Studies

IFPUG/NESMA		
ISBSG 10006	ISBSG 12456	ISBSG 14430
ISBSG 10062	ISBSG 12479	ISBSG 14433
ISBSG 10087	ISBSG 12502	ISBSG 14446
ISBSG 10090	ISBSG 12510	ISBSG 14497
ISBSG 10117	ISBSG 12524	ISBSG 14505
ISBSG 10121	ISBSG 12545	ISBSG 14513
ISBSG 10198	ISBSG 12555	ISBSG 14543
ISBSG 10200	ISBSG 12575	ISBSG 14554
ISBSG 10228	ISBSG 12630	ISBSG 14582
ISBSG 10331	ISBSG 12650	ISBSG 14615
ISBSG 10337	ISBSG 12655	ISBSG 14622
ISBSG 10341	ISBSG 12671	ISBSG 14635
ISBSG 10346	ISBSG 12702	ISBSG 14671
ISBSG 10377	ISBSG 12717	ISBSG 14692
ISBSG 10409	ISBSG 12748	ISBSG 14710
ISBSG 10467	ISBSG 12778	ISBSG 14735
ISBSG 10480	ISBSG 12790	ISBSG 14863
ISBSG 10512	ISBSG 12801	ISBSG 14889
ISBSG 10517	ISBSG 12892	ISBSG 14904
ISBSG 10566	ISBSG 12904	ISBSG 14914
ISBSG 10589	ISBSG 12925	ISBSG 14928
ISBSG 10631	ISBSG 12928	ISBSG 14962
ISBSG 10631	ISBSG 12932	ISBSG 15028
ISBSG 10646	ISBSG 12977	ISBSG 15047
ISBSG 10665	ISBSG 12987	ISBSG 15076
ISBSG 10681	ISBSG 13003	ISBSG 15106
ISBSG 10889	ISBSG 13047	ISBSG 15149
ISBSG 10955	ISBSG 13049	ISBSG 15234
ISBSG 10971	ISBSG 13134	ISBSG 15315
ISBSG 11120	ISBSG 13157	ISBSG 15399
ISBSG 11150	ISBSG 13167	ISBSG 15420
ISBSG 11163	ISBSG 13172	ISBSG 15452
ISBSG 11229	ISBSG 13204	ISBSG 15454
ISBSG 11257	ISBSG 13234	ISBSG 15477
ISBSG 11268	ISBSG 13279	ISBSG 15492
ISBSG 11285	ISBSG 13335	ISBSG 15509
ISBSG 11350	ISBSG 13352	ISBSG 15533
ISBSG 11397	ISBSG 13412	ISBSG 15574
ISBSG 11424	ISBSG 13469	ISBSG 15722
ISBSG 11533	ISBSG 13539	ISBSG 15723
ISBSG 11536	ISBSG 13557	ISBSG 15724
ISBSG 11596	ISBSG 13565	ISBSG 15743
ISBSG 11674	ISBSG 13569	ISBSG 15743
ISBSG 11690	ISBSG 13577	ISBSG 15892
ISBSG 11693	ISBSG 13648	ISBSG 16200
ISBSG 11704	ISBSG 13681	ISBSG 16655
ISBSG 11727	ISBSG 13709	ISBSG 17223
ISBSG 11746	ISBSG 13722	ISBSG 17528
ISBSG 11749	ISBSG 13731	ISBSG 18875
ISBSG 11771	ISBSG 13754	ISBSG 19528
ISBSG 11793	ISBSG 13755	ISBSG 19601
ISBSG 11808	ISBSG 13828	ISBSG 19805
ISBSG 11844	ISBSG 13856	ISBSG 20073
ISBSG 11850	ISBSG 13860	ISBSG 20219
ISBSG 11854	ISBSG 13887	ISBSG 20554
ISBSG 11859	ISBSG 13981	ISBSG 20938
ISBSG 11861	ISBSG 13987	ISBSG 22115
ISBSG 11881	ISBSG 14010	ISBSG 22821
ISBSG 11921	ISBSG 14013	ISBSG 23373
ISBSG 11938	ISBSG 14042	ISBSG 24307
ISBSG 12004	ISBSG 14043	ISBSG 26311
ISBSG 12058	ISBSG 14069	ISBSG 27043
ISBSG 12066	ISBSG 14071	ISBSG 27722
ISBSG 12076	ISBSG 14078	ISBSG 28002
ISBSG 12120	ISBSG 14270	ISBSG 28396
ISBSG 12152	ISBSG 14302	ISBSG 28700
ISBSG 12167	ISBSG 14307	ISBSG 28952
ISBSG 12176	ISBSG 14316	ISBSG 29632
ISBSG 12188	ISBSG 14323	ISBSG 29713
ISBSG 12343	ISBSG 14324	ISBSG 29880
ISBSG 12361	ISBSG 14328	ISBSG 29952
ISBSG 12371	ISBSG 14331	ISBSG 30398
ISBSG 12379	ISBSG 14335	ISBSG 30857
ISBSG 12391	ISBSG 14362	ISBSG 31126
ISBSG 12424	ISBSG 14424	ISBSG 31349
		ISBSG 32285

KSLOC		
ISBSG 10087	ISBSG 20067	ISBSG 30398
ISBSG 10117	ISBSG 20073	ISBSG 30571
ISBSG 10221	ISBSG 20139	ISBSG 30724
ISBSG 10624	ISBSG 20142	ISBSG 30786
ISBSG 11258	ISBSG 20153	ISBSG 30791
ISBSG 11674	ISBSG 20166	ISBSG 30833
ISBSG 11684	ISBSG 20185	ISBSG 30857
ISBSG 11727	ISBSG 20223	ISBSG 31115
ISBSG 11852	ISBSG 20318	ISBSG 31126
ISBSG 11996	ISBSG 20508	ISBSG 31148
ISBSG 12066	ISBSG 20554	ISBSG 31349
ISBSG 12173	ISBSG 20567	ISBSG 31488
ISBSG 12338	ISBSG 20901	ISBSG 31546
ISBSG 12343	ISBSG 20938	ISBSG 31878
ISBSG 12360	ISBSG 20977	ISBSG 32150
ISBSG 12371	ISBSG 21071	ISBSG 32417
ISBSG 12650	ISBSG 21216	ISBSG 32430
ISBSG 12790	ISBSG 21672	ISBSG 32494
ISBSG 12892	ISBSG 21968	
ISBSG 12961	ISBSG 21994	
ISBSG 12985	ISBSG 22115	
ISBSG 13709	ISBSG 22221	
ISBSG 13740	ISBSG 22297	
ISBSG 13988	ISBSG 22373	
ISBSG 14053	ISBSG 22497	
ISBSG 14323	ISBSG 22608	
ISBSG 14418	ISBSG 22666	
ISBSG 14582	ISBSG 22731	
ISBSG 14710	ISBSG 22821	
ISBSG 14887	ISBSG 23069	
ISBSG 14956	ISBSG 23227	
ISBSG 15021	ISBSG 23345	
ISBSG 15138	ISBSG 23373	
ISBSG 15222	ISBSG 23840	
ISBSG 15302	ISBSG 24000	
ISBSG 15385	ISBSG 24276	
ISBSG 15435	ISBSG 24432	
ISBSG 15551	ISBSG 24529	
ISBSG 15700	ISBSG 24812	
ISBSG 15722	ISBSG 24877	
ISBSG 15786	ISBSG 25227	
ISBSG 15892	ISBSG 25393	
ISBSG 16146	ISBSG 25501	
ISBSG 16200	ISBSG 25545	
ISBSG 16308	ISBSG 25678	
ISBSG 16440	ISBSG 25909	
ISBSG 16494	ISBSG 26030	
ISBSG 16655	ISBSG 26311	
ISBSG 16676	ISBSG 26393	
ISBSG 16808	ISBSG 26739	
ISBSG 16849	ISBSG 26971	
ISBSG 16967	ISBSG 27043	
ISBSG 17093	ISBSG 27307	
ISBSG 17223	ISBSG 27541	
ISBSG 17528	ISBSG 27722	
ISBSG 17674	ISBSG 27793	
ISBSG 17793	ISBSG 27859	
ISBSG 17885	ISBSG 28002	
ISBSG 17941	ISBSG 28106	
ISBSG 18054	ISBSG 28396	
ISBSG 18131	ISBSG 28407	
ISBSG 18261	ISBSG 28682	
ISBSG 18799	ISBSG 28700	
ISBSG 18875	ISBSG 28803	
ISBSG 18959	ISBSG 28952	
ISBSG 19195	ISBSG 29565	
ISBSG 19273	ISBSG 29620	
ISBSG 19337	ISBSG 29632	
ISBSG 19418	ISBSG 29713	
ISBSG 19528	ISBSG 29880	
ISBSG 19601	ISBSG 29891	
ISBSG 19613	ISBSG 29952	
ISBSG 19672	ISBSG 29954	
ISBSG 19719	ISBSG 30255	
ISBSG 19795	ISBSG 30257	
ISBSG 19808	ISBSG 30309	
ISBSG 19941	ISBSG 30393	

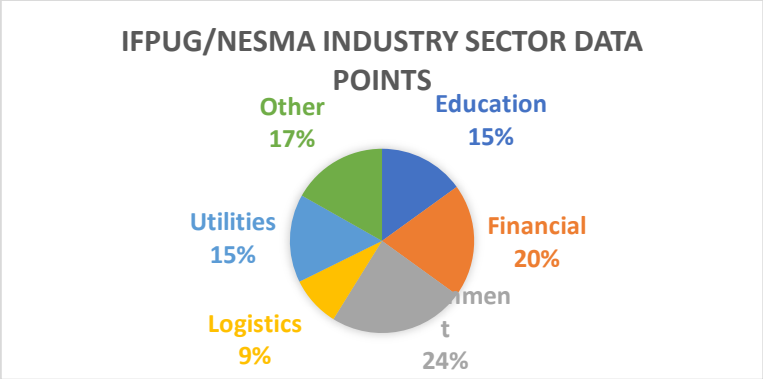
Appendix C – Attributes eliminated from all analyses.

Attributes	Reason for elimination
ISBSG ID	Not useful for analysis
Application Description	No conformity, entries very freeform
Application Large	Scarcely populated across the dataset
Application Medium	Scarcely populated across the dataset
Application Small	Scarcely populated across the dataset
Code Maintained	Scarcely populated across the dataset
Data Bases	Scarcely populated across the dataset
DBMS	Scarcely populated across the dataset
FTE Hrs	Not necessary as all effort has been converted and scaled to hours per 12 month period
Hardware	Scarcely populated across the dataset
Hardware Number	Scarcely populated across the dataset
KSLOC Rate	Scarcely populated across the dataset
Lanugauge Used	Scarcely populated across the dataset, this applies to the language used in a Third Party App, not for the application code
Msinteance KSLOC Rate	Scarcely populated across the dataset
Operating System	Scarcely populated across the dataset
Operating System	Scarcely populated across the dataset, this applies to the language used in a Third Party App, not for the application code
Other Size	Scarcely populated across the dataset
Other Size Approach	Scarcely populated across the dataset
Parameters Maintains	Scarcely populated across the dataset, this applies to the language used in a Third Party App, not for the application code
Partial Effort Scaled	Scarcely populated across the dataset, this applies to the language used in a Third Party App, not for the application code
Porfolio Size Approach	Scarcely populated across the dataset
Portfolio	Scarcely populated across the dataset
Portfolio Size	Scarcely populated across the dataset
Product Name	Scarcely populated across the dataset, this applies to the language used in a Third Party App, not for the application code
Product Versions	Scarcely populated across the dataset, this applies to the language used in a Third Party App, not for the application code
Secondary - 10th Language stuff	Scarcely populated across the dataset
Support KSLOC Rate	Scarcely populated across the dataset
Third Party	Scarcely populated across the dataset
Third Party Ownership	Scarcely populated across the dataset
Third Party Package	Scarcely populated across the dataset
Total M&S Effort	This applies to the value that is entered by the user, all analysis uses the effort in hours scaled to a 12 hour period
Total M&S Units	This applies to the value that is entered by the user, all analysis uses the effort in hours scaled to a 12 hour period
Total Maintenance FTE	This applies to the value that is entered by the user, all analysis uses the effort in hours scaled to a 12 hour period
Total Support Effort	This applies to the value that is entered by the user, all analysis uses the effort in hours scaled to a 12 hour period
Total Support FTE	This applies to the value that is entered by the user, all analysis uses the effort in hours scaled to a 12 hour period
Total Support Units	This applies to the value that is entered by the user, all analysis uses the effort in hours scaled to a 12 hour period

Appendix D – Benchmarks for Industry Sector, Application Types, Primary Languages

D.1 – IFPUG/NESMA by Industry Sector

IFPUG/NESMA Industry	Number of Data Points
Education	34
Financial	45
Government	54
Logistics	20
Utilities	35
Other	38
Total	226



Education										
Attribute	Sample Size	Min	Max	25%	75%	Mean	Median	Mode	Standard Deviation	Coefficient of Variation
Age at Benchmark	34	1	15	4	10.75	7.235	5.5	4	4.606	0.637
Application Size	34	60	3900	231.25	1478	1021.74	450	550	1162.802	1.138
Maintenance FP Rate	28	0.1	5.9	0.3	0.725	0.766	0.45	0.3	1.102	1.439
Support FP Rate	32	0.1	3.3	0.2	0.5	0.536	0.307	0.1	0.73	1.363
Total Defects FP Rate	27	0.001	0.057	0.007	0.034	0.02	0.015	0.001	0.017	0.866
Total Defects	27	1	156	3	18.5	19.519	9	1	32.714	1.676
Total Maintenance & Support Hours	34	11	2197	143	1129	632.765	355	11	619.95	0.98
Total Maintenance Hours	31	6	1665	90	597	425.935	171	96	490.27	1.151
Total Maintenance & Support FP Rate	34	0.031	6.43	0.458	1.222	1.136	0.726	0.031	1.338	1.178
Total Support Hours	34	11	805	46.25	384.3	244.412	185	40	216.21	0.885

Financail										
Attribute	Sample Size	Min	Max	25%	75%	Mean	Median	Mode	Standard Deviation	Coefficient of Variation
Age at Benchmark	45	1	26	3	11	7.96	6	3	6.291	0.79
Application Size	45	231	55312	734	3545	4222.89	2352	231	8596.118	2.036
Maintenance FP Rate	44	0.118	0.81	0.195	0.383	0.307	0.28	0.118	0.154	0.5
Support FP Rate	43	0.113	0.779	0.184	0.367	0.295	0.302	0.113	0.142	0.481
Total Defects FP Rate	45	0.01	0.129	0.026	0.045	0.04	0.034	0.01	0.022	0.564
Total Defects	45	9	1311	33	120	117.936	69	9	198.407	1.682
Total Maintenance & Support Hours	45	175	22423	441	1787	2094.11	1147	175	3673.79	1.754
Total Maintenance Hours	44	89	11211	217.75	919.5	970.818	599	196	1738.272	1.791
Total Maintenance & Support FP Rate	45	0.231	19.041	0.386	0.742	1.007	0.607	0.231	2.764	2.745
Total Support Hours	44	86	11212	236.5	861	929.841	489	581	1739.593	1.871

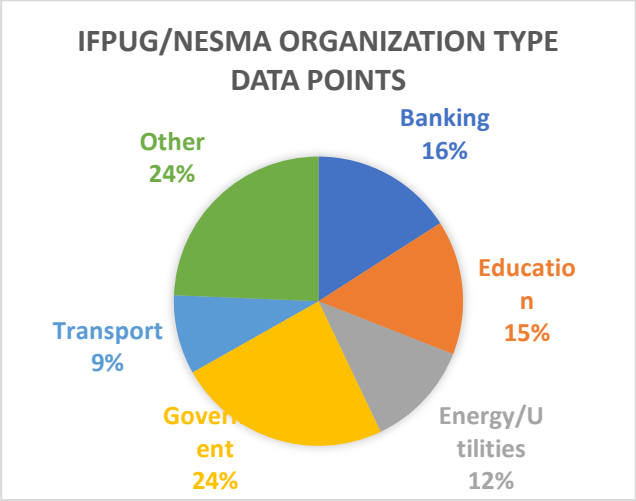
Government										
Attribute	Sample Size	Min	Max	25%	75%	Mean	Median	Mode	Standard Deviation	Coefficient of Variation
Age at Benchmark	54	1	31	6	12	9.981	9	6	6.299	0.631
Application Size	54	24	18500	444.75	3542	2724.5	871.5	453	3578.376	1.313
Maintenance FP Rate	45	0.1	2.1	0.206	0.374	0.383	0.287	0.2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

Logistics										
Attribute	Sample Size	Min	Max	25%	75%	Mean	Median	Mode	Standard Deviation	Coefficient of Variation
Age at Benchmark	20	2	28	3	14.25	10	7	3	8.429	0.843
Application Size	20	172	9352	523.5	4369	2700.15	1658.5	172	2774.759	1.028
Maintenance FP Rate	20	0.126	0.576	0.2	0.428	0.341	0.337	0.126	0.128	0.376
Support FP Rate	20	0.121	0.477	0.194	0.377	0.303	0.314	0.121	0.11	0.364
Total Defects FP Rate	20	0.014	0.343	0.027	0.054	0.059	0.037	0.014	0.072	1.222
Total Defects	20	16	275	32.5	117.5	82.2	59	59	67.367	0.82
Total Maintenance & Support Hours	20	181	3682	361.75	1966	1347.75	1066.5	181	1158.416	0.86
Total Maintenance Hours	20	99	2175	182.5	1059	725.4	589.5	99	632.77	0.872
Total Maintenance & Support FP Rate	20	0.246	1.052	0.392	0.81	0.644	0.692	0.246	0.224	0.348
Total Support Hours	20	82	1841	190.75	817.8	622.35	439.5	82	542.374	0.871

Utilities										
Attribute	Sample Size	Min	Max	25%	75%	Mean	Median	Mode	Standard Deviation	Coefficient of Variation
Age at Benchmark	35	2	17	6	14	9.857	10	15	4.453	0.452
Application Size	35	234	22342	808.5	4517	3444.83	1821	234	4493.143	1.304
Maintenance FP Rate	35	0.127	0.696	0.213	0.403	0.32	0.333	0.127	0.126	0.396
Support FP Rate	35	0.111	0.723	0.196	0.388	0.301	0.263	0.111	0.137	0.457
Total Defects FP Rate	35	0.016	0.073	0.03	0.047	0.038	0.036	0.016	0.013	0.35
Total Defects	35	9	1048	25	129.5	117.343	54	15	185.19	1.578
Total Maintenance & Support Hours	35	186	9058	364	1967	1507.63	1079	186	1627.572	
Total Maintenance Hours	35	82	4529	197.5	1036	783.6	518	192	826.503	1.055
Total Maintenance & Support FP Rate	35	0.238	1.419	0.404	0.776	0.62	0.649	0.238	0.259	0.418
Total Support Hours	35	88	4529	175	914.5	724.029	561	88	805.09	1.112

D.2 – IFPUG/NESMA by Organization Type

User Organization Type	Number of Data Points
Banking	36
Education	34
Energy/Utilities	27
Government	54
Transport	20
Other	55
Total	226



Banking

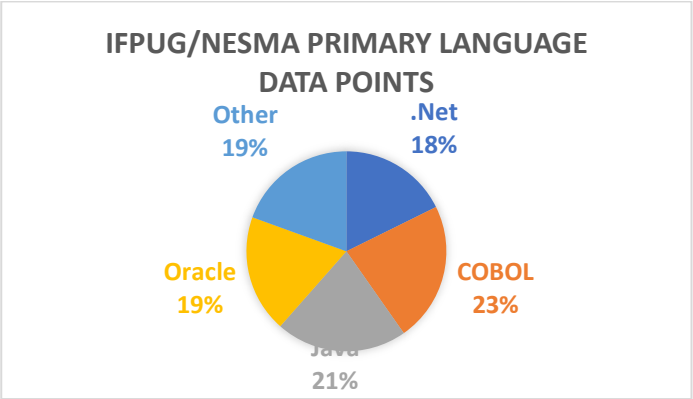
Attribute	Sample Size	Min	Max	25%	75%	Mean	Median	Mode	Standard Deviation	Coefficient of Variation
Age at Benchmark	36	1	26	3	12	8.361	6	3	6.962	0.833
Application Size	36	231	55312	775.25	3873.75	4615.889	2415	231	9423.054	2.041
Maintenance FP Rate	36	0.118	0.745	0.194	0.383	0.298	0.28	0.118	0.142	0.478
Support FP Rate	35	0.113	0.779	0.169	0.367	0.286	0.313	0.113	0.136	0.475
Total Defects FP Rate	36	0.006	0.129	0.027	0.043	0.04	0.034	0.006	0.024	0.605
Total Defects	36	8	1311	36.5	126.5	130	75.5	9	218.209	1.679
Total Maintenance & Support Hours	36	175	22423	465	1787.25	2024.667	1124.5	175	3783.717	1.869
Total Maintenance Hours	36	89	11211	227.5	952.25	1037.722	599	89	1889.907	1.821
Total Maintenance & Support FP Rate	36	0.231	1.525	0.342	0.731	0.577	0.634	0.231	0.275	0.477
Total Support Hours	36	8	11212	252.25	888.5	986.944	489	581	1895.259	1.92

Transport

Attribute	Sample Size	Min	Max	25%	75%	Mean	Median	Mode	Standard Deviation	Coefficient of Variation
Age at Benchmark	20	2	28	3	14.25	10	7	3	8.429	0.843
Application Size	20	172	9352	523.5	4369	2700.15	1658.5	172	2774.759	1.028
Maintenance FP Rate	20	0.126	0.576	0.2	0.428	0.341	0.337	0.126	0.128	0.376
Support FP Rate	20	0.121	0.477	0.194	0.377	0.303	0.314	0.121	0.11	0.364
Total Defects FP Rate	20	0.014	0.343	0.027	0.054	0.059	0.037	0.014	0.072	1.222
Total Defects	20	16	275	32.5	117.5	82.2	59	59	67.367	0.82
Total Maintenance & Support Hours	20	181	3682	361.75	1966	1347.75	1066.5	181	1158.416	0.86
Total Maintenance Hours	20	99	2175	182.5	1058.75	725.4	589.5	99	632.77	0.872
Total Maintenance & Support FP Rate	20	0.246	1.052	0.392	0.81	0.644	0.692	0.246	0.224	0.348
Total Support Hours	20	82	1841	190.75	817.75	622.35	439.5	82	542.374	0.871

D.3 – IFPUG/NESMA Benchmarks by Primary Language

Primary Language	Number of Data Points
.Net	40
COBOL	51
Java	48
Oracle	43
Other	44
	226



.Net

Attribute	Sample Size	Min	Max	25%	75%	Mean	Median	Mode	Standard Deviation	Coefficient of Variation
Age at Benchmark	40	1	21	3	8	5.925	5	3	4.047	0.683
Application Size	40	135	12987	543.25	3343.5	2568.875	1823.5	876	2678.126	1.043
Maintenance FP Rate	40	0.1	0.423	0.185	0.239	0.234	0.198	0.1	0.092	0.392
Support FP Rate	39	0.146	0.376	0.157	0.198	0.208	0.171	0.146	0.079	0.381
Total Defects FP Rate	40	0.016	0.129	0.031	0.045	0.041	0.038	0.016	0.019	0.461
Total Defects	40	6	551	27.75	123.5	89	58.5	11	97.451	1.095
Total Maintenance & Support Hours	40	100	3780	296.5	1219.75	871.8	528	343	781.224	0.896
Total Maintenance Hours	40	51	1871	160	641.25	455.275	311.5	192	382.671	0.841
Total Maintenance & Support FP Rate	40	0.165	0.799	0.342	0.429	0.438	0.367	0.165	0.172	0.392
Total Support Hours	40	44	1909	120.75	585.25	416.525	243	151	401.602	0.964

COBOL

Attribute	Sample Size	Min	Max	25%	75%	Mean	Median	Mode	Standard Deviation	Coefficient of Variation
Age at Benchmark	51	8	31	12	20.5	16.137	15	15	5.759	0.357
Application Size	51	523	55312	1600.5	5421.5	5492.059	2642	523	8486.645	1.545
Maintenance FP Rate	50	0.195	0.624	0.225	0.465	0.361	0.369	0.195	0.125	0.345
Support FP Rate	50	0.195	0.503	0.23	0.376	0.32	0.341	0.195	0.086	0.27
Total Defects FP Rate	51	0.003	0.141	0.024	0.043	0.035	0.031	0.003	0.021	0.6
Total Defects	51	10	1311	55	133	150.951	86	86	224.436	1.487
Total Maintenance & Support Hours	51	464	22423	1264.5	2999.5	2788.706	1899	464	3303.232	1.185
Total Maintenance Hours	50	264	11211	687.75	1559	1456.08	1016	264	1654.438	1.136
Total Maintenance & Support FP Rate	51	0.269	1.075	0.446	0.825	0.672	0.696	0.269	0.206	0.306
Total Support Hours	50	182	11212	638	1433.5	1378.94	915	182	1670.868	1.212

Java

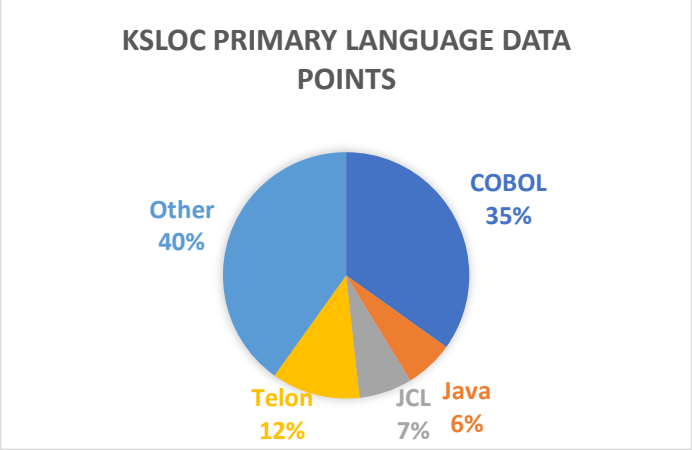
Attribute	Sample Size	Min	Max	25%	75%	Mean	Median	Mode	Standard Deviation	Coefficient of Variation
Age at Benchmark	48	1	15	3	9	6.292	6	3	3.82	0.607
Application Size	48	24	12543	358.75	3387.5	2318.521	743.5	267	2945.582	1.27
Maintenance FP Rate	42	0.118	0.483	0.141	0.334	0.268	0.305	0.34	0.11	0.41
Support FP Rate	39	0.113	0.6	0.14	0.37	0.275	0.258	0.263	0.132	0.481
Total Defects FP Rate	48	0.006	1.667	0.034	0.072	0.135	0.04	0.006	0.32	2.367
Total Defects	48	8	908.544	26.75	108.5	103.428	38.5	32	154.277	1.492
Total Maintenance & Support Hours	48	51	3997	230.75	1090.5	758.792	378	454	813.137	1.072
Total Maintenance Hours	48	51	1627	111.75	556.25	370.5	245	107	347.548	0.938
Total Maintenance & Support FP Rate	48	0.113	4.75	0.262	0.672	0.594	0.592	0.603	0.654	1.101
Total Support Hours	42	8	3752	139.5	593.5	443.762	209.5	198	607.822	1.37

Oracle

Attribute	Sample Size	Min	Max	25%	75%	Mean	Median	Mode	Standard Deviation	Coefficient of Variation
Age at Benchmark	43	2	17	4.5	8.5	7.116	6	5	4.007	0.563
Application Size	43	172	8458	666	3294.5	2359.535	953	854	2483.586	1.053
Maintenance FP Rate	42	0.196	0.81	0.235	0.4	0.346	0.29	0.196	0.157	0.452
Support FP Rate	42	0.174	0.779	0.222	0.398	0.359	0.357	0.174	0.16	0.447
Total Defects FP Rate	43	0.01	0.343	0.023	0.055	0.044	0.032	0.01	0.05	1.123
Total Defects	43	7	251	21.5	104	68.186	54	15	59.726	0.876
Total Maintenance & Support Hours	43	181	3479	469.5	1751	1201.535	854	530	946.991	0.788
Total Maintenance Hours	42	99	1809	213.25	906.5	610.929	393.5	196	509.537	0.834
Total Maintenance & Support FP Rate	43	0.37	1.709	0.46	0.806	0.729	0.646	0.37	0.345	0.473
Total Support Hours	42	82	1670	250.25	848.25	590.738	437.5	183	449.767	0.761

D.4 KSLOC Benchmarks by Primary Language

Primary Language	Number of Data Points
COBOL	60
Java	11
JCL	12
Telon	20
Other	69
	172



COBOL

Attribute	Sample Size	Min	Max	25%	75%	Mean	Median	Mode	Standard Deviation	Coefficient of Variation
Age at Benchmark	60	0.1	20	2.4	4.9	4.572	3.5	4.2	4.049	0.886
Application Size	60	1.656	810000	6.804	67.462	14348.64	17.555	1.656	104498.371	7.283
Maintenance KSLOC Rate	57	0.6	1298.036	9.186	49.815	59.133	25.469	0.6	172.068	2.91
Support KSLOC Rate	24	0.326	158.037	1.384	12.999	14.995	4.161	0.326	32.797	2.187
Total Defects KSLOC Rate	45	3.07E-04	4.355	0.068	0.653	0.558	0.238	3.07E-04	0.884	1.585
Total Defects	45	1	5677	3	24	157.956	8	1	844.143	5.344
Total Maintenance & Support Hours	60	19	56320	239.75	2729.5	4613.2	553.5	19	10635.228	2.305
Total Maintenance Hours	57	19	27940	214	1763	2753.965	507	600	6064.788	2.202
Total Maintenance & Support KSLOC Rate	60	5.84E-04	5632	10.08	59.118	156.101	26.664	5.84E-04	738.424	4.73
Total Support Hours	24	12	11661	85.75	3789.75	2201.417	307.5	46	3293.112	1.496

Java

Attribute	Sample Size	Min	Max	25%	75%	Mean	Median	Mode	Standard Deviation	Coefficient of Variation
Age at Benchmark	11	0.9	14	4.95	9.5	7.055	7	0.9	4.106	0.582
KSLOC	11	11.088	650	16.898	97	124.406	53	11.088	196.933	1.583
Maintenance KSLOC Rate	9	0.86	16.527	2.643	8.962	5.748	3.423	0.86	4.981	0.867
Total Defects KSLOC Rate	11	0.07	17.02	2.374	4.834	4.442	3.266	0.07	4.687	1.055
Total Defects	11	8	1847.3	43	749.5	455	62	8	666.927	1.466
Total Maintenance & Support Hours	11	51	162597	73.5	6287	17226.36	124	51	48485.99	2.815
Total Maintenance Hours	9	51	5825	58	237	761.556	114	51	1900.805	2.496
Total Maintenance & Support FP Rate	11	0.86	494.216	2.847	17.696	66.182	4.69	0.86	150.483	2.274

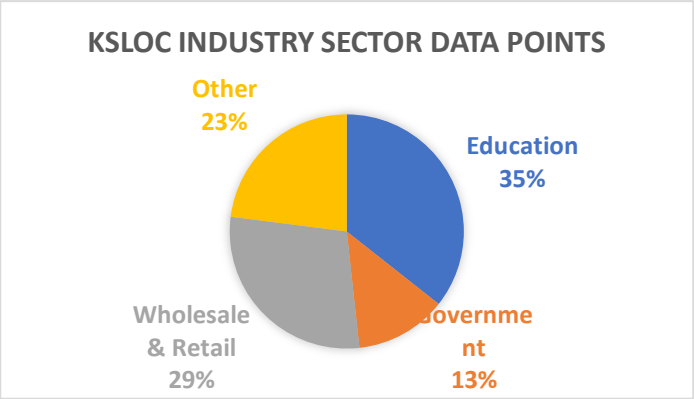
Telon

Attribute	Sample Size	Min	Max	25%	75%	Mean	Median	Mode	Standard Deviation	Coefficient of Variation
Age at Benchmark	20	0.2	6	0.775	2.375	2.05	1.85	0.7	1.606	0.784
KSLOC	20	4.785	226.771	20.071	52.332	44.447	33.765	4.785	47.77	1.075
Maintenance KSLOC Rate	18	1.828	94.671	5.566	21.696	24.008	12.704	1.828	27.692	1.153
Support KSLOC Rate	7	0.093	47.171	0.421	6.2	8.834	1.331	0.093	17.324	1.961
Total Defects KSLOC Rate	12	0.022	0.419	0.075	0.303	0.188	0.171	0.022	0.137	0.728
Total Defects	12	1	47	2.5	16.5	12.083	7.5	1	13.548	1.121
Total Maintenance & Support Hours	20	0.093	111.851	4.084	20.241	24.699	10.846	0.093	33.46	1.355
Total Maintenance Hours	18	82	1717	208.5	504.25	523.167	404.5	82	483.311	0.924
Total Maintenance & Support KSLOC Rate	20	0.093	111.851	4.084	20.241	24.699	10.846	0.093	33.46	1.355
Total Support Hours	7	16	326	21	73.5	80.571	33	21	112.085	1.391

JCL										
Attribute	Sample Size	Min	Max	25%	75%	Mean	Median	Mode	Standard Deviation	Coefficient of Variation
Age at Benchmark	12	1	13.5	1.85	6.8	4.692	4.15	1	3.586	0.764
KSLOC	12	0.023	30	0.988	10.179	8.022	2.999	0.023	10.858	1.354
Maintenance KSLOC Rate	10	6.467	978.339	34.545	327.249	279.796	156.414	6.467	342.591	1.224
Support KSLOC Rate	6	6.229	4173.913	12.299	680.006	864.284	61.865	6.229	1655.841	1.916
Total Defects KSLOC Rate	7	0.267	7.22	0.777	4.548	3.017	2.982	0.267	2.702	0.895
Total Defects	7	2	19	3	9.5	7.571	7	3	5.996	0.792
Total Maintenance & Support Hours	12	96	3492	265.75	1494.25	1105.583	380.5	96	1307.582	1.183
Total Maintenance Hours	10	127	3243	255.25	492.25	644.4	315.5	127	935.156	1.451
Total Maintenance & Support KSLOC Rate	12	35.925	4173.913	50.065	805.149	665.305	266.362	35.925	1156.884	1.739
Total Support Hours	6	96	3298	116	2255.25	1137.167	200.5	96	1537.177	1.352

D.5 KSLOC Benchmarks by Industry Sector

KSLOC Industry Sector	Number of Data Points
Banking	85
Education	31
Government	11
Wholesale & Retail	25
Other	20
Total	172



Banking										
Attribute	Sample Size	Min	Max	25%	75%	Mean	Median	Mode	Standard Deviation	Coefficient of Variation
Age at Benchmark	85	0.2	18	1.8	4.2	3.395	2.9	2.9	2.636	0.776
KSLOC	85	0.023	11000	6.267	47.203	193.51	15.579	0.023	1229.553	6.354
Maintenance KSLOC Rate	79	1.491	1298.036	10.294	54.565	83.453	28.616	1.491	201.607	2.416
Support KSLOC Rate	31	0.07	4173.913	1.507	17.067	177.022	7.267	0.07	758.065	4.282
Total Defects KSLOC Rate	57	0.017	7.22	0.107	0.869	0.847	0.3	0.017	1.437	1.696
Total Defects	57	1	5677	2	19	113.68	7	1	750.438	6.601
Total Maintenance & Support Hours	85	19	56320	191	1267	2089.329	434	21	7107.24	3.402
Total Maintenance Hours	79	19	27940	202.5	905.5	1172.633	418	304	3240.421	2.763
Total Maintenance & Support FP Rate	85	0.093	5632	10.856	83.852	208.382	31.566	0.093	771.487	3.702
Total Support Hours	31	8	11661	39.5	292.5	923.677	101	46	2320.267	2.512

Education										
Attribute	Sample Size	Min	Max	25%	75%	Mean	Median	Mode	Standard Deviation	Coefficient of Variation
Age at Benchmark	31	1	15	4	11.5	7.419	5	4	4.689	0.632
KSLOC	31	3.411	752.3	44.435	193.143	172.893	93	3.411	206.962	1.197
Maintenance KSLOC Rate	28	0.018	25.355	1.018	4.451	3.868	2.107	1.059	5.236	1.354
Support KSLOC Rate	31	0.112	9.968	0.985	3.114	2.383	1.818	0.112	2.143	0.899
Total Defects KSLOC Rate	24	0.005	0.88	0.031	0.193	0.159	0.074	0.005	0.209	1.312
Total Defects	24	1	42	3	16.5	11.375	8	1	11.382	1.001
Total Maintenance & Support Hours	31	11	2197	116.5	1103	617.581	294	11	633.371	1.026
Total Maintenance Hours	28	6	1665	77.25	595.5	426.714	166.5	96	511.322	1.198
Total Maintenance & Support KSLOC Rate	31	0.146	27.484	2.535	6.554	5.876	3.882	0.146	6.121	1.042
Total Support Hours	31	11	805	43	336.5	232.161	173	40	212.7	0.916

Government

Attribute	Sample Size	Min	Max	25%	75%	Mean	Median	Mode	Standard Deviation	Coefficient of Variation
Age at Benchmark	11	0.9	20	5.5	9.5	8.355	7	6	5.441	0.651
KSLOC	11	11.088	4040	16.898	900	668.133	53	11.088	1215.962	1.82
Maintenance KSLOC Rate	9	0.6	16.527	2.34	4.924	5.073	3.423	0.6	5.02	0.99
Total Defects KSLOC Rate	11	0.005	7.486	0.171	3.276	2.337	2.364	0.005	2.571	1.1
Total Defects	11	7	1037	34	75.5	159.194	47.135	7	303.835	1.909
Total Maintenance & Support Hours	11	51	15338	73.5	1318	2283	124	51	4718.748	2.067
Total Maintenance Hours	9	51	3939	58	237	584.889	114	51	1269.744	2.171
Total Maintenance & Support KSLOC Rate	11	51	15338	73.5	1318	2283	124	51	4718.748	2.067

Wholesale & Retail

Attribute	Sample Size	Min	Max	25%	75%	Mean	Median	Mode	Standard Deviation	Coefficient of Variation
Age at Benchmark	25	0	16	3	12	8.016	9	1	5.356	0.668
KSLOC	25	10.718	8061.947	106.088	785.746	736.261	230	10.718	1613.157	2.191
Maintenance KSLOC Rate	23	0.201	24.53	1.789	9.009	6.055	3.048	0.201	6.649	1.098
Support KSLOC Rate	21	0.073	63.398	0.779	3.603	5.967	1.944	0.073	13.801	2.313
Total Defects KSLOC Rate	25	0.042	9.781	0.211	1.33	1.27	0.71	0.042	1.977	1.557
Total Defects	25	4.866	1613.898	79	360.888	284.802	175.176	4.866	361.747	1.27
Total Maintenance & Support Hours	25	12	10419	718	3985	2689.64	2041	12	2496.539	0.928
Total Maintenance Hours	23	10	4990	387	2383	1512.783	916	10	1555.67	1.028
Total Maintenance & Support KSLOC Rate	25	0.494	66.823	2.883	13.509	10.613	5.184	0.494	13.98	1.317
Total Support Hours	21	2	9885	169	1318	1539.524	628	2	2470.746	1.605