Practical Software and Systems Measurement

A foundation for objective project management

Building Models

PSM Users Group Conference

Department of Defense
U.S. Army

Building Models from Your PSM Data

- ***This workshop addressed the evolution of measurement practice beyond the collection of data.***
- ***Small-scale empirical models were discussed***
  - ***Variation in the data***
  - ***Confidence intervals***
  - ***Model performance:***
    - ***Magnitude Relative Error***
    - ***Prediction Level (PRED)***

Objectives of the Workshop

- ***Analyze Two sets of data***
  - ***The first set will be used to learn a technique: SEL Size, Effort, Driver data***
  - ***The second set will be used to practice the technique: PSM Speaker Time data***
- ***As statistical concepts were needed they were discussed***
- ***Intent was to show how to create small models by example using <u>nothing more than Excel</u>***

Need for Models

- ***Models are useful for forecasting, performance analysis, and decision-making***
  - ***WBS is narrowly addressed with current estimation models***
  - ***Strength of cause and effect relationships***
  - ***Impact of decision making: Personnel turnover***
- ***Establish data requirements (model parameters)***
- ***Explain assignable causes of variation and their degree of influence***
- ***Used to validate data***
  - ***Poor data definitions and collection consistency***
  - ***Poor processes that produce the data***

WBS Help

How is the effort estimated for the rest of these?

Software Cost Estimation Models

3.1 Program Management
3.1.1 Planning & Mgt
3.1.2 Program Control
3.1.3 Contract Management
3.1.4 Contractor Laboratory
3.2 System Engineering
3.2.1 SysReq'ts
3.2.2 Design and Integration
3.2.4 Sup. & Maintainability Eng.
3.2.5 QA
3.2.6 CM
3.2.7 Human Factors
3.2.8 Security
3.3 HW/SW Design, Development and Prod
3.3.1 HW Design & Dev
3.3.2 SW Design & Dev
3.3.3 HW/SW Integration & Checkout
3.5 Test and Evaluation
3.5.1 Sys T&E
3.5.4 Site Accep
3.6 Documentation
3.7 Support

Model Building Steps

- *1. Decide what relationship you would like to investigate*
- *2. Identify assignable causes of variation (drivers)*
- *3. Collect Data*
- *4. Normalize data and check for consistency*
- *5. Build model and evaluate*
- *6. Add or remove drivers*
- *7. Repeat steps 3 - 7 until conclusion*
- *8. Pilot model*

Output

- *We practiced modeling on data gathered at this conference: Speaker Times*
- *Data*
- *Speaker name*
- *Number of slides in presentation*
- *Number of text / table slides presented*
- *Number of figure slides presented*
- *Start and Stop times*

Speaker Time Data

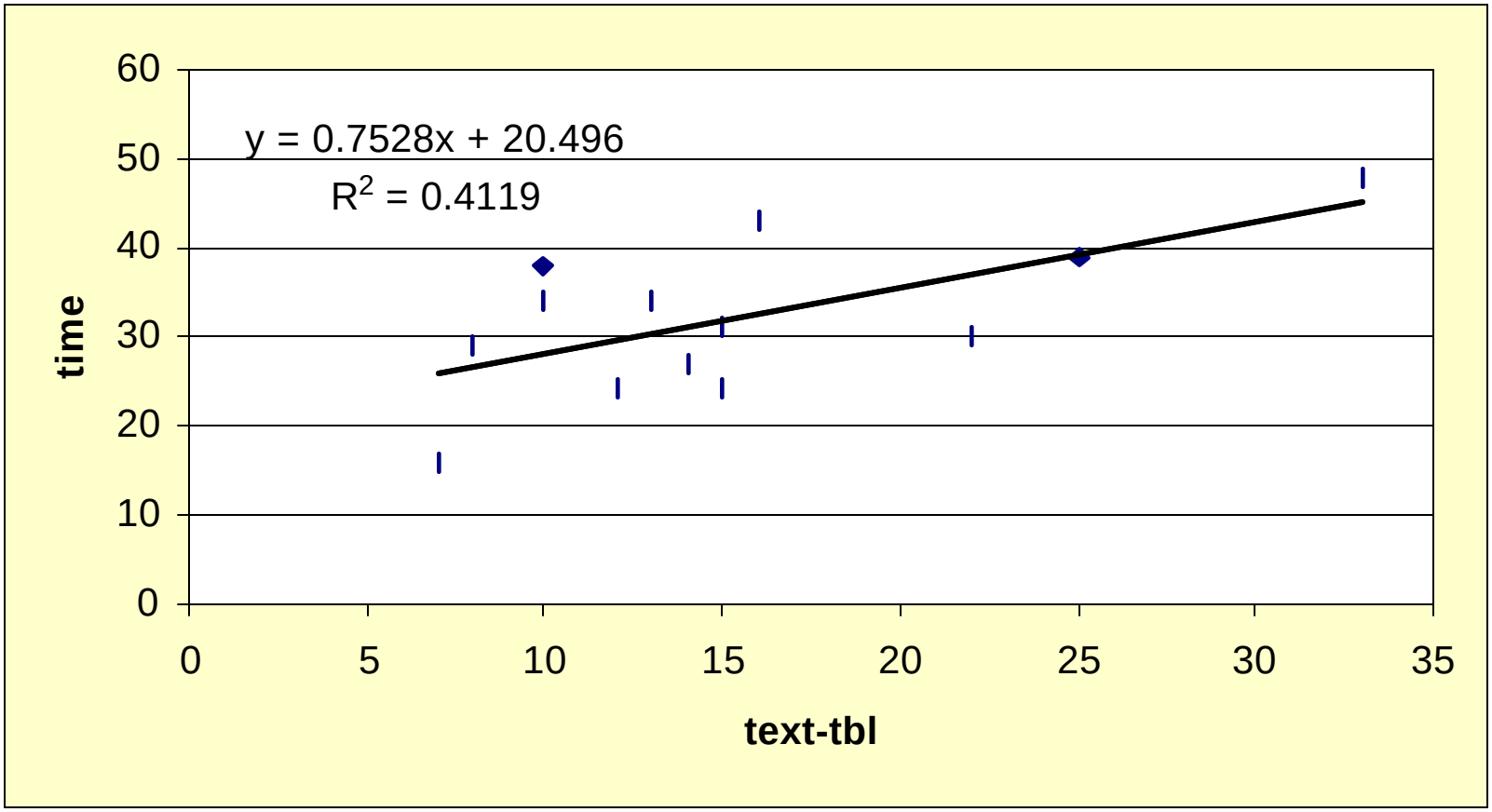
Tuesday Speakers

Speaker No.	Presenter	M/F	Total Slides	Text-Tbl	Figures	Time
2	McGarry	0	33	25	5	0:39
3	Jarzombek/	1	28	16	9	0:43
4	Sturgess	1	33	12	18	0:24
5	Croll	1	28	15	3	0:24
6	Clark,B	1	23	10	8	0:34
7	Buys	0	23	15	0	0:31
8	Schreiber	0	16	8	8	0:29
9	MacIver	1	21	10	8	0:38
10	Majewski	1	28	7	5	0:16
Wednesday Speakers						
11	Hernandez /Janusz	0.5	26	22	3	0:30
12	Roedler	1	24	13	6	0:34
13	Morris	1	18	0	18	0:21
14	Mercurio	1	19	14	3	0:27
15	Clark,E	0	43	33	6	0:48

Correlation Analysis of Data

	M/F	Total Slides	Text-Tbl	Figures	Start	Stop	Time
M/F	1.00						
Total Slides	-0.26	1.00					
Text-Tbl	-0.56	0.77	1.00				
Figures	0.37	-0.03	-0.46	1.00			
Start	0.00	-0.21	-0.38	-0.14	1.00		
Stop	-0.03	-0.19	-0.34	-0.15	1.00	1.00	
Time	-0.40	0.43	0.69	-0.19	-0.23	-0.18	1.00

Scatter Plot of the data



Analysis Results

Productivity	est. Time	MRE	PRED(.30)
0.64	52.08	-0.34	
0.37	33.33	0.22	1
0.50	25.00	-0.04	1
0.63	31.25	-0.30	1
0.29	20.83	0.39	
0.48	31.25	-0.01	1
0.28	16.67	0.43	
0.26	20.83	0.45	
0.44	14.58	0.09	1
0.73	45.83	-0.53	
0.38	27.08	0.20	1
0.52	29.17	-0.08	1
0.69	68.75	-0.43	
			0.54

Productivity	
Mean	0.48
Standard Error	0.04
Median	0.48
Standard Deviation	0.16
Range	0.47
Minimum	0.26
Maximum	0.73
Sum	6.21
Count	13.00
Confidence Level(80.0%)	0.06

Workshop Participants

- ***Brad Clark - leader***
- ***Betsy Clark - scribe / cheerleader***
- ***Mary Anne McGarry***
- ***Renee Schriber***
- ***Chris Miller***

Summary

- ***We had fun analyzing data***
  - ***Instructional data set: SEL Data***
  - ***Practice data set: Speaker times***
- ***Text / Table slides were the best predictor of time***
- ***With 80% confidence, the mean delivery time for each slide is between 1:51 - 2:23 min.***
- ***Women may talk longer (need more data to validate this preliminary finding)***

Conclusions, Recommendations, and Results

- ***Using only the Excel spreadsheet, we performed data analysis and created small models***
- ***The "hands-on" approach was a lot of fun***
  - ***Attendees learned the skills by performing the spreadsheet operations***
  - ***We all learned a few new Excel capabilities***
- ***We recommend more of the "hands-on" type workshops in the future***