HERREN
insights that ignite

ICEAA

2018 Professional Development and Training Workshop

The Art of Employing Data Science to Improve Cost Data Analysis

Economic/Data Analysis Track (EA10)

International Cost Estimating and Analysis Association (ICEAA)
Phoenix, AZ

Shavaiz Saood | Richard Shea

June 2018

HERREN

Presentation Agenda

▶ Introduction
  - Data Science Discussion
  - Problem Statement

▶ Discussion Data Organization Problem

▶ Working Through the Problem

▶ Solutions

▶ Summary

Herren Associates, Inc.   **International Cost Estimating and Analysis Association (ICEAA)**   1
Phoenix, AZ | June 2018

Presented at the 2018 ICEAA Professional Development & Training Workshop - www.iceaaonline.com

Data Science Discussion

▶ Historical data is the backbone of a good estimate

- *Good data provides credibility, accuracy, and defensibility*

▶ Cost estimators must be able to discern data quality

▶ Data collection is a top priority for cost estimators

▶ Contextual completeness is crucial

▶ Important to cost estimating, nature of data (types, formats, stories)

▶ Not always conducive to analysis – hard to manipulate, work with, glean important info

▶ Too much data can be just as hard to use as too little data – need to know how to filter and manipulate to find what you're seeking

Problem Statement

▶ We acquired a large data extract from an Enterprise Resource Planning (ERP) database

▶ Data supported a program office that was geared towards tracking funding execution by project through the Planning, Programming, Budgeting and Execution (PPBE) process

▶ Trying to find per ship ordnance and electronic systems costs

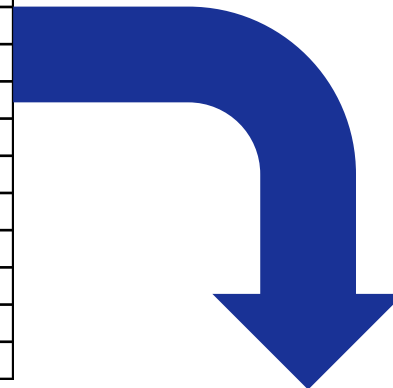▶ Electronics and ordnance system map to multiple ships

Problem Statement

▶ There were multiple problems we had to solve for this project:

– The ERP database was organized in a contract execution hierarchy

– Single or multiple children would sum to the parent

– Lower level parent(s) summing to a higher level parent

– The work breakdown structure was ten-levels deep

– ERP Data was for tracking funding execution through the PPBE process

– Government Furnished Equipment (GFE) estimate

Presented at the 2018 ICEAA Professional Development & Training Workshop - www.iceaaonline.com

Problem Statement

ERP WBS	Sum of Children	Parent/Child
11	1443	Parent
11	1443	Parent
1101	1325	Parent
110101	16	Parent
11010101	9	Parent
1101010101	1	Parent
110101010101	0	Child
1101010102	1	Parent
110101010201	0	Child
110101013	4	Parent
110101010301	0	Child
110101010302	0	Child
110101010303	0	Child
11010101304	0	Child
11010102	5	Parent
1101010201	2	Parent
110101020101	0	Child
110101020102	0	Child
1101010202	1	Parent
110101020201	0	Child

▶ Scrubbed data to identify the child elements that rolled up into parent elements using Excel functions

▶ Allowed us to reduce our rows of data from over 6000 to under 3000

WBS Desc Lvl 3	WBS Desc Lvl 4	WBS Desc Lvl 5	WBS Desc Lvl 6	WBS Desc Lvl 7
Program Element	Hardware or System Element	Ship Description	Performer	Task or Work to be Performed

HERREN

Important Things to Consider

▸ Planned Costs and Budget are not consistently done at the child level

▸ Commitments + Obligations + Actual Costs = Assigned Costs

▸ Pivot tables are filtered to show only the Electronics and Ordnance

▸ Mapping of ships and systems to the child level

HERREN

Raw Data Table Sample

Commitments + Obligations + Actual Costs = Assigned Costs

PS-(Clean) 1	Combined	Description	Planned Cost	Budget	Commitments	Obligations	Actual Costs	Assigned Costs	Available Budget	Year
BS-840000	11	Gov Performer #1	6,213,376,860	2,485,997,842	38,659,828	296,428,636	1,737,741,480	2,072,829,944	413,167,898	FY11
BS-840000	110101	DIRECT CITE	102,479,772	49,366,000	1,003,898	4,277,719	34,626,277	39,907,894	9,458,106	FY11
BS-840000	11010101	REIMBURSABLE	34,479,772	5,825,880	650,269	643,448	20,154,680	21,448,397	(15,622,517)	FY11
BS-840000	1101010101	CONTRACT SVCS	5,405,772	5,458,065	-	418,173	4,987,599	5,405,772	52,293	FY11
BS-840000	110101010101	FIELD SVCS	5,405,772	5,458,065	-	418,173	4,987,599	5,405,772	52,293	FY11
BS-840000	1101010102	REIMBURSABLE	-	367,815	6,241	-	301,074	307,315	60,500	FY11
BS-840000	110101010201	CONTRACT SVCS	-	367,815	6,241	-	301,074	307,315	60,500	FY11
BS-840000	1101010103	FIELD SVCS	-	-	644,028	225,275	14,866,007	15,735,310	(15,735,310)	FY11

HERREN

Presented at the 2018 ICEAA Professional Development & Training Workshop - www.iceaaonline.com

Data Table Expanded Sample

Data Provided   Data Extracted

Parent/Child	Description	Assigned Costs	Year	Ship	Electronics and Ordnance	Performer	DIRECT CITE	REIMBURSABLE
Parent	System HARDWARE Ship #3	$133,265,390	FY13	Ship #3	System 2	Hardware		
Child	Government CONTRACTS	$132,317,960	FY13	Ship #3	System 2	Gov't Contracts		
Parent	Government Performer #2	$861,689	FY13	Ship #3	System 2	Government Performer #2		
Child	REIMBURSABLE	$252,252	FY13	Ship #3	System 2	Government Performer #2		REIMBURSABLE
Child	DIRECT CITE	$609,437	FY13	Ship #3	System 2	Government Performer #2	DIRECT CITE	
Parent	Government Performer #3	$58,224	FY13	Ship #3	System 2	Government Performer #3		
Child	REIMBURSABLE	$58,224	FY13	Ship #3	System 2	Government Performer #3		REIMBURSABLE
Parent	Government Performer #4	$27,517	FY13	Ship #3	System 2	Government Performer #4		
Child	REIMBURSABLE	$10,517	FY13	Ship #3	System 2	Government Performer #4		REIMBURSABLE
Child	DIRECT CITE	$17,000	FY13	Ship #3	System 2	Government Performer #4	DIRECT CITE	

HERREN

Problem Solution

Original Dataset Pivot

Parent/Child	(All)	

Row Labels	Sum of Assigned Costs
⊟ FY13	$267,478,210
⊟ Ship #3	$267,478,210
⊟ System 2	$267,478,210
⊟ Government Performer #1	$132,317,960
Government CONTRACTS	$132,317,960
⊟ Government Performer #2	$1,723,378
DIRECT CITE	$609,437
Government Performer #2	$861,689
REIMBURSABLE	$252,252
⊟ Government Performer #3	$116,448
Government Performer #3	$58,224
REIMBURSABLE	$58,224
⊟ Government Performer #4	$55,034
DIRECT CITE	$17,000
Government Performer #4	$27,517
REIMBURSABLE	$10,517
⊟ Hardware	$133,265,390
System HARDWARE Ship #3	$133,265,390
Grand Total	**$267,478,210**

▶ Using Excel pivot table functions on organized data we can filter out the Parent Costs resulting in the Child Cost in the total and thereby reducing cost duplication.

New Dataset Pivot

Parent/Child	Child	

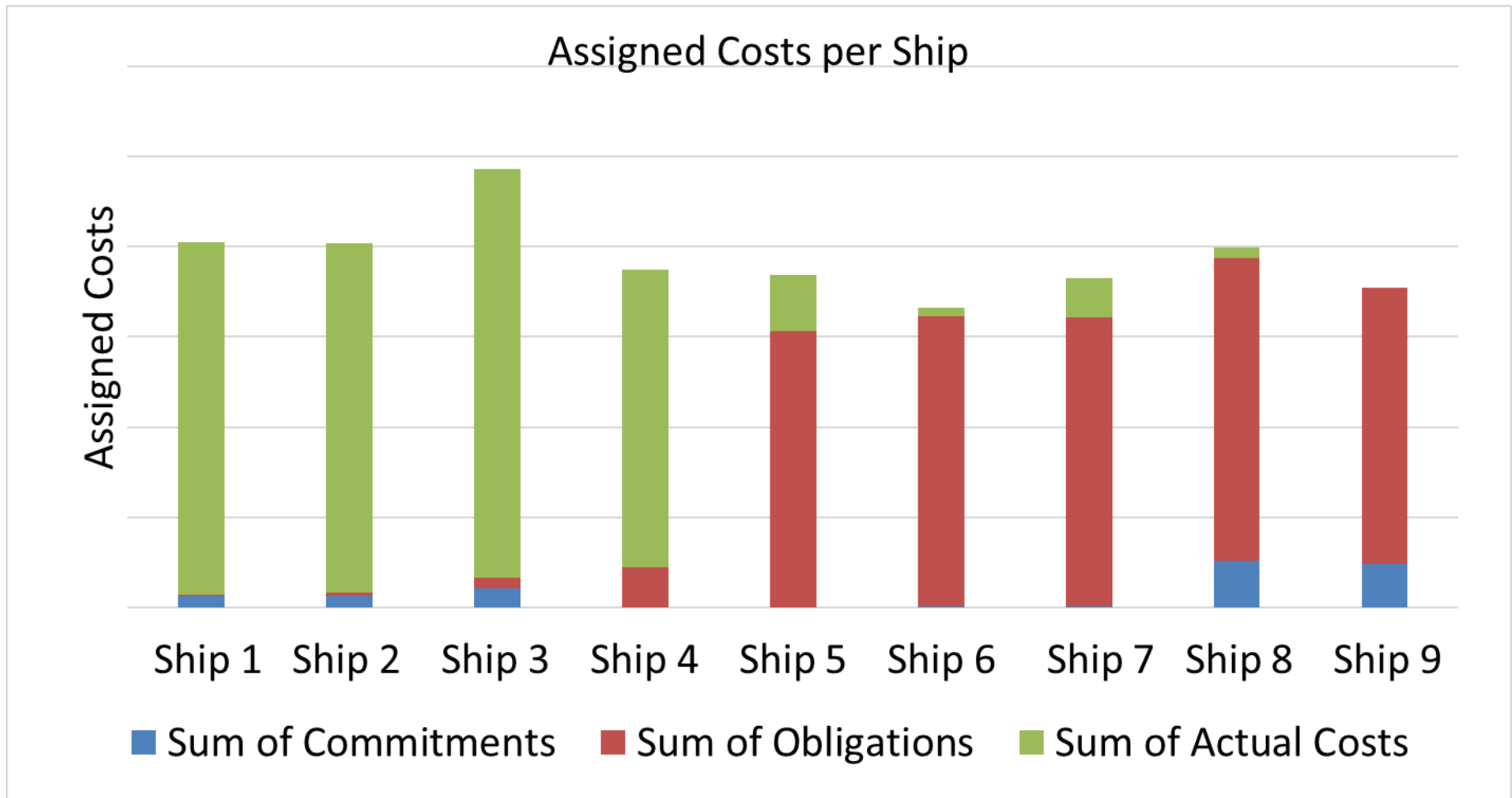| Row Labels	Sum of Assigned Costs
⊟ FY13	$133,265,390
⊟ Ship #3	$133,265,390
⊟ System 2	$133,265,390
⊟ Government Performer #1	$132,317,960
Government CONTRACTS	$132,317,960
⊟ Government Performer #2	$861,689
DIRECT CITE	$609,437
REIMBURSABLE	$252,252
⊟ Government Performer #3	$58,224
REIMBURSABLE	$58,224
⊟ Government Performer #4	$27,517
DIRECT CITE	$17,000
REIMBURSABLE	$10,517
Grand Total	**$133,265,390**

Problem Solution

▶ Solution Approach

- Identify cost at child level associated with a specific ship

- Breakout system costs applies to multiple ships to single ship

- Extract details out of a single data field
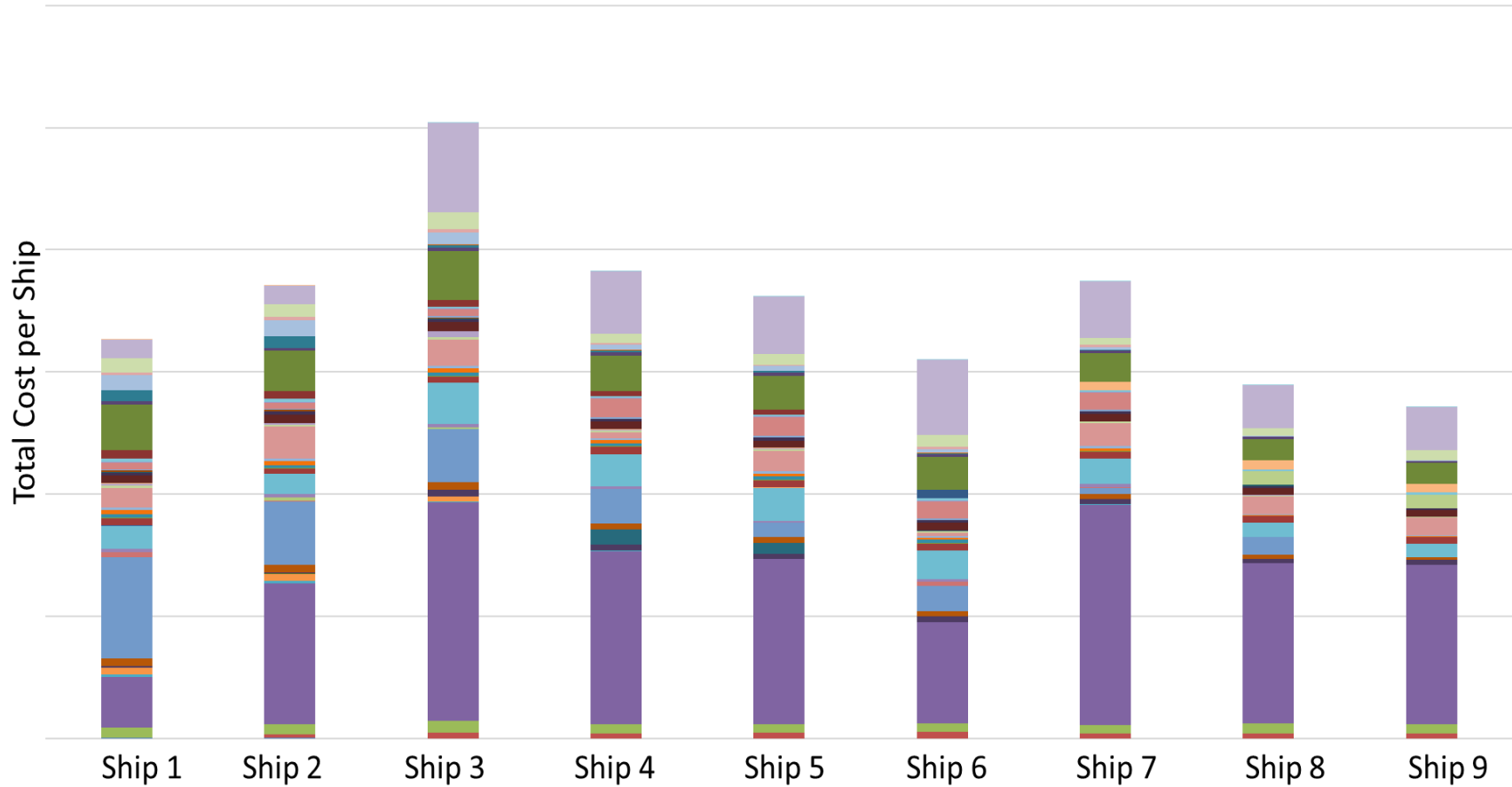 - System
  - Ships
  - Tasking
  - Performer

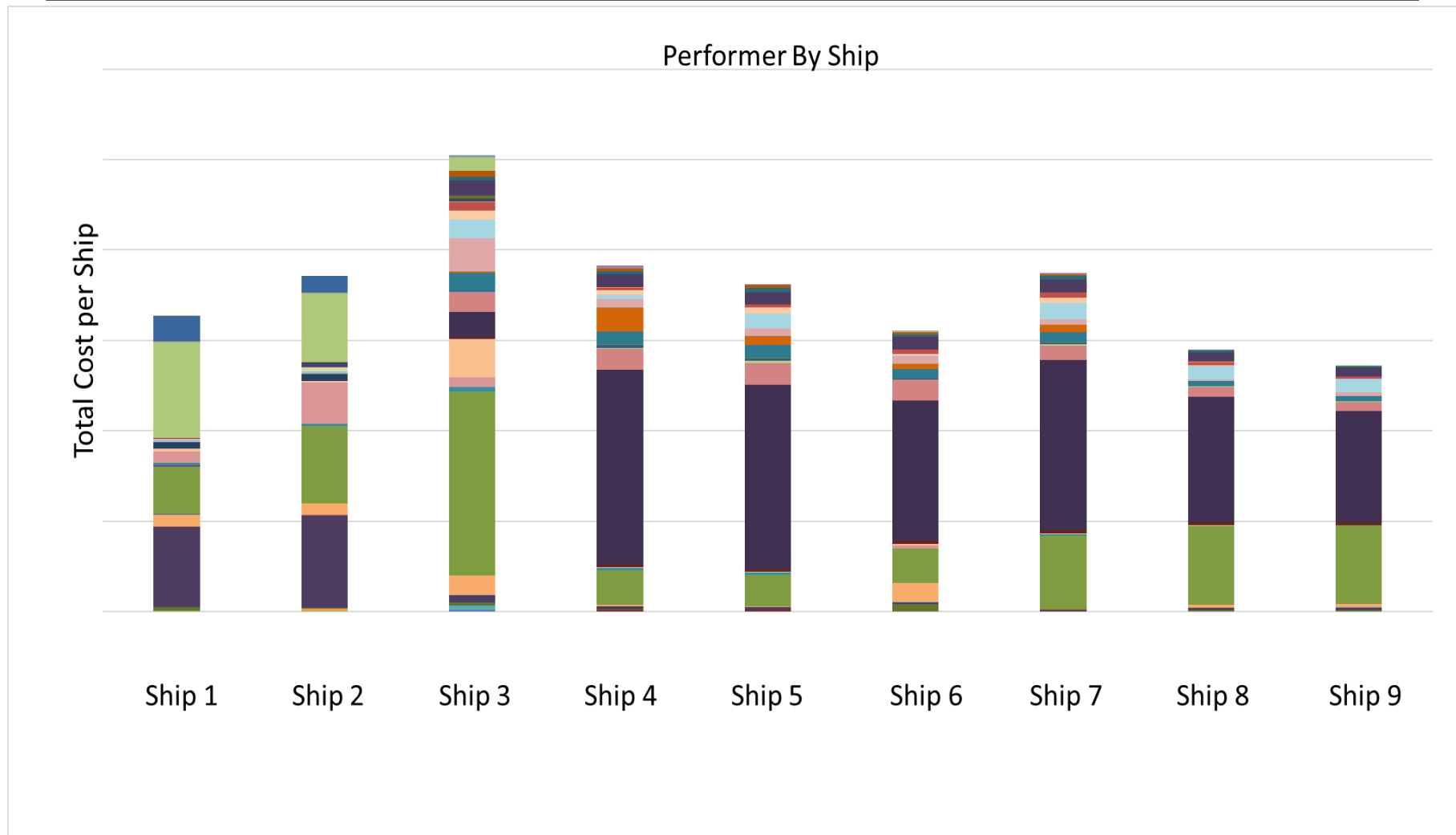Presented at the 2018 ICEAA Professional Development & Training Workshop - www.iceaaonline.com

HERREN

Problem Solution Illustration

Presented at the 2018 ICEAA Professional Development & Training Workshop - www.iceaaonline.com
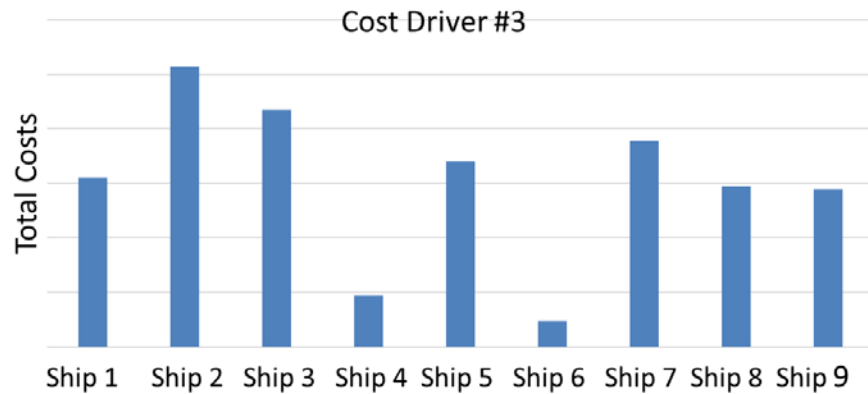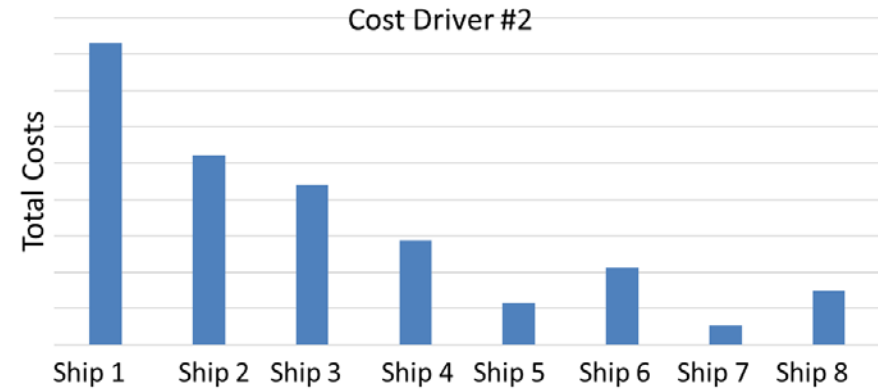
HERREN

Problem Solution Illustration

Problem Solution Illustration

Presented at the 2018 ICEAA Professional Development & Training Workshop - www.iceaaonline.com

Problem Solution Illustration

Presented at the 2018 ICEAA Professional Development & Training Workshop - www.iceaaonline.com
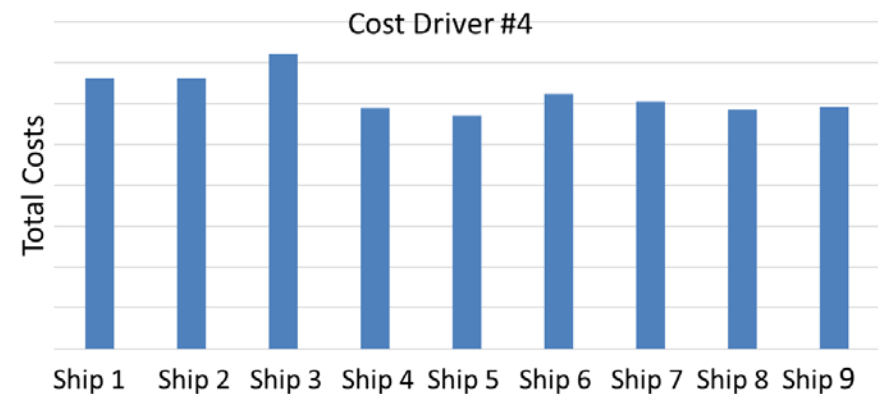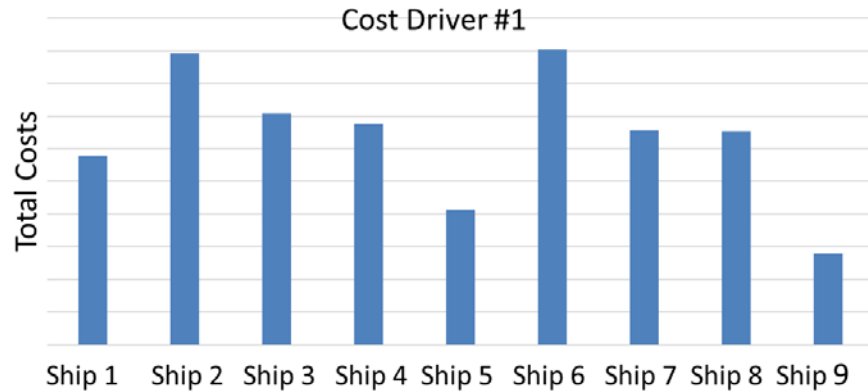
General Application

▶ This case was specific to us, but general approach and rules can be applied everywhere

▶ Maybe you've done this before and didn't realize it, reformatted data to make more useful

▶ Structured/organized data looks nice but can be much less useful than database format

▶ Database format allows analyst to use Excel functions/pivot tables however he/she chooses to glean information from the data

▶ Endless possibilities/views for use of database formatted data, not true of structured data

Closing

▸ This analysis and comparison then provided the foundation for which claims made in a Government Furnished Equipment (GFE) model could be tested

▸ The transformation of this data allowed analysts to garner valuable insights into subtle inconsistencies and improve the accuracy of the estimate

▸ If you're ever on the sending end of large data sets, be cognizant of the format you put it in

▸ A pretty, structured format tells one story, but not all of them, and may not be the most beneficial to the user

Stay Connected

(in) linkedin.com/company/herren-associates-inc-

Authors

Shavaiz Saood I Consultant
Shavaiz.Saood@jlha.com

Richard Shea I Consultant
Phone: (202) 400-3471
Richard.Shea@jlha.com

About Herren

Founded in 1989, Herren Associates is an engineering and management consulting firm with a proven record of maximizing the value of every taxpayer dollar. As trusted advisors to federal executives, we partner with clients to drive operational improvements and manage performance - maximizing efficiency and cost effectiveness.