Data Collection, Normalization & Organization

4th Annual Workshop Canada
Ottawa, ON

Tuesday, May 1st, 2018

Abstract

This presentation will address what is arguably the key to credible, defensible cost analysis – the underlying data. Regardless of the specific project for which an analyst is either developing an estimate or assessing another analyst's estimate, the source of the data upon which the analyst's work is based must be fully researched in order to gain an understanding of its strengths/weaknesses relative to fidelity, completeness, applicability, accessibility, etc.

There are a variety of data types critical to quality cost analysis, including cost, technical and programmatic information. When it comes to cost data, there are not surprisingly varying degrees of quality depending on the origin of the data. These quality differences shape the data collection strategy, specifically the prioritization of data sources that an analyst should pursue to accomplish a particular cost analysis task. This presentation will provide an overview of the data types and sources that are available to cost analysts and discuss the subject of quality in relation to cost data.

Additionally, this presentation will address a process that is essential to making the data analysis-ready, specifically data normalization. Collection of quality source data is necessary, but not sufficient. Normalization for escalation, quantity and other factors will ensure absolute data quality, thereby comparability/consistency across data points.

Finally and time permitting, this presentation will introduce some important concepts relative to data organization.

Outline

- Setting the stage
- Data collection
  - Data types
  - Data sources
  - Example in context of method development
- Data normalization
- Summary

The Big Picture

Quality cost estimates facilitate authoritative decision-maker knowledge and informed decision making

Decision-makers Must Trust Our Work

- As analysts, we are asking decision-makers to place their trust/confidence in our analysis and resulting estimates
- Thus, the first requirement of an estimate is that it be credible/defensible
- The principal means to establish credibility/defensibility (and corresponding trust) is to explain in very specific terms the path from *data/facts* to *methods/models* to *estimates* (i.e., the evidentiary chain)
  - Clarity of this evidentiary chain is paramount
  - Clarity enables persuasion leading to trust
- Our work as analysts must be persuasive and compelling
  - Many of the consumers of our estimates are exposed to an endless stream of advocates with agendas
  - So, when we tell these consumers “no, project x is going to cost 50% more than you expected”, we had better be prepared to sell/defend our position

Credibility --> Confidence --> Trust

Data Engenders Trust …

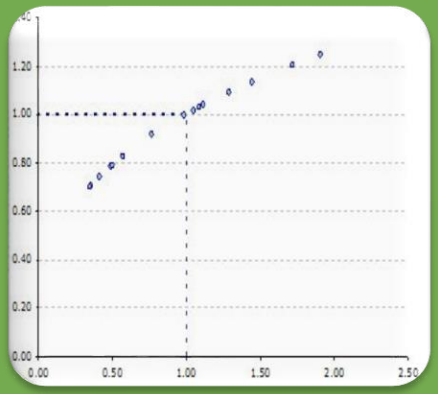
- Data is the foundation for estimate credibility and defensibility
  - An estimate not grounded in data can be viewed a guess or, at best, analyst opinion/judgement
  - An estimate is only as good as it's underlying data

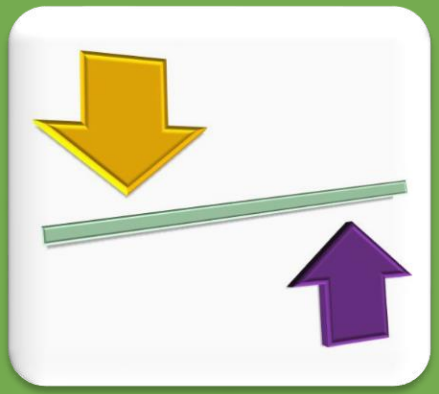
- Data collection and normalization must be the top priority of cost analysts
  - The equivalent of doing the all important prep work (scraping, sanding, priming) prior to painting the exterior of a house

Data is the lifeblood of cost analysts

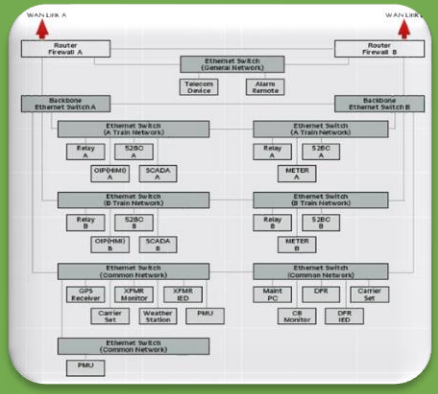
… In Our Analytical Methods

Top-down





Bottom-up



Parametric Cost Estimating Relationships

- Mathematical equations typically developed via regression analysis
- Relates actual costs of relevant programs to technical & programmatic characteristics

Analogy

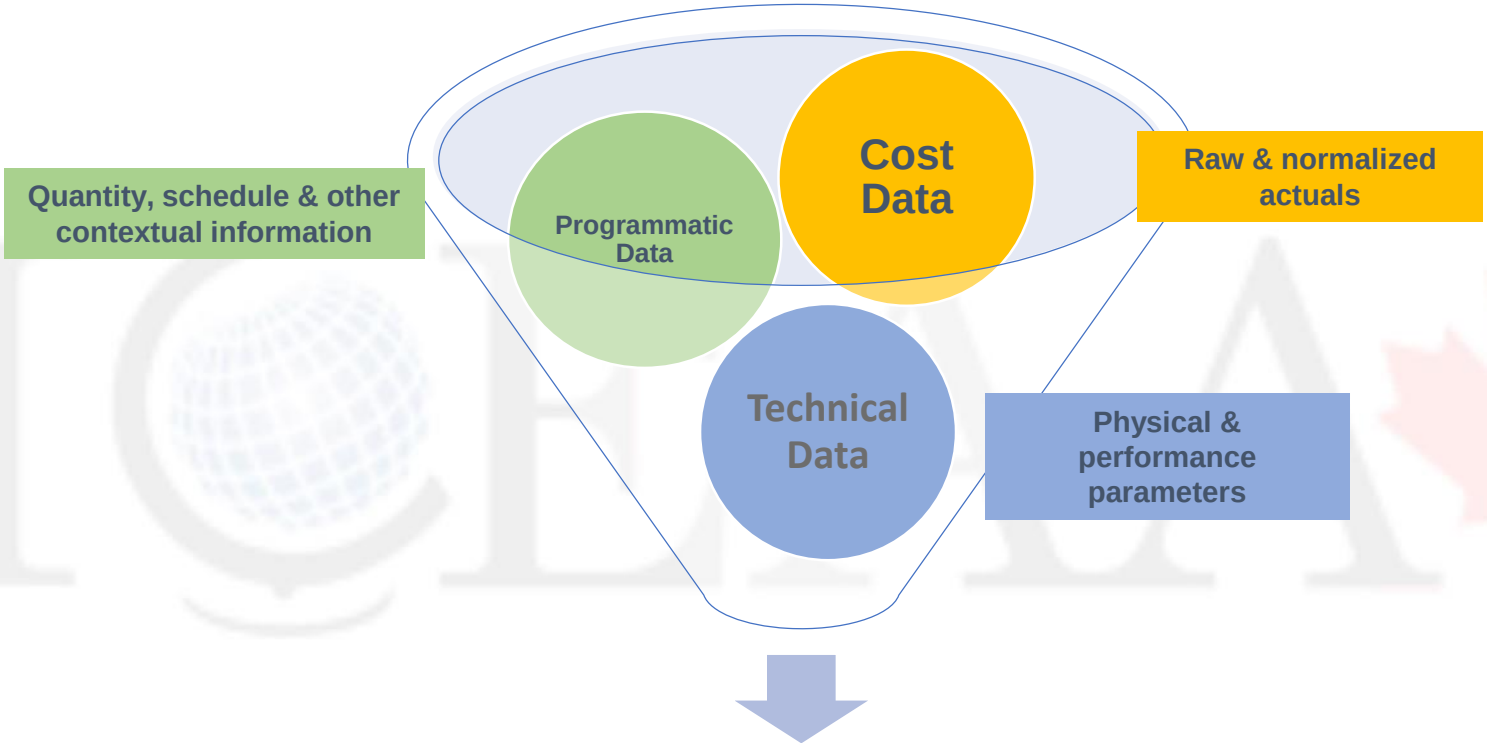
- Uses actual cost of a relevant program
- Adjusts that cost for the new program's technical and programmatic differences

Engineering Build-up

- An industrial engineering approach
- Addresses separate segments of labor and material at the detailed level and aggregates to a total cost element

This presentation is focused on data underlying the parametric & analogy methods

Data Types – Overview

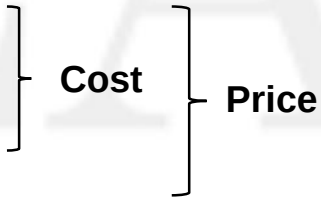


Foundation for credible & defensible estimates

High performing cost organizations/analysts do what it takes to get the 'right' data

Data Types – Cost Data

- Dollars associated with the labor and materials required for a project
  - Hour data is a nice to have that provides useful context to labor cost data
- Dollars can be expressed as *cost* or *price;* important to understand what's included (or not) in the values
  - Direct cost
  - Overhead cost
  - General & administrative (G&A) cost
  - Fee or profit



- Whether cost or price, important to understand the timeframe reflected in the values for the purpose of normalizing the data to a logical reference point

The most reliable cost data is actuals for completed contracts/projects

Data Types – Technical Data

- Physical characteristics of a system
  - Size
  - Weight
  - Volume
  - Power
  - Etc.
- Performance characteristics of a system
  - Speed
  - Depth
  - Frequency
  - Range
  - Mean time between failure
  - Operational availability
  - Etc.

Provide context to cost data and enable cross-system comparisons

Data Types – Programmatic Data

- Schedule
  - Development time in months
  - Time to first flight in months
  - Time to ship launch
- Quantity
  - Number of developmental units (i.e., prototypes)
  - Number of test articles
  - Number of test flights
  - Number of production units
- Acquisition strategy
  - Sole source
  - Competition
    - Dual source
    - Winner-take-all
- Contract type
  - Cost-plus
  - Fixed-price

Provide context to cost data and enable cross-system comparisons

Data Sources – Cost Data

- The most authoritative data is actual price at completion for a given contract
  - Represents *what actually happened*
  - And not *what should have happened* (i.e., a contract value)
  - And not *what could have happened* (i.e., a bid)

- In the U.S., the best sources for this contract data are
  - Internal contractor accounting system records
  - Deliverable (i.e., contractually required) cost reports
    - Earned value management (EVM) reporting
    - Cost and Software Data Reporting (CSDR), which is specific to defense contracts

- In the absence of these sources, the next 'best' sources (in order of descending goodness) are
  - Contract line item price data (i.e., contracts)
  - Government finance/accounting system data
  - Government budget data

Actuals are the closest thing to truth

Data Sources – Cost Data Quality

- Quality is best characterized in terms of how the data is reported and what data is reported; the graphic below represents desired data features

Reporting Structure (How)

- Hierarchical, product oriented work breakdown structure (WBS)*
- Hardware, software, services, data & facilities cost reported at multiple levels of indenture

Reporting Visibility (What)

- Recurring vs non-recurring cost*
- Labor vs. material cost
- Direct vs. indirect cost
- Functional cost (engr, mfg, QC & tooling)
- Prime vs. subcontractor cost

- A product-oriented WBS enables cross-system/project comparisons
- Recurring vs. non-recurring cost visibility is critical to identifying hardware build cost vs. hardware design cost; this distinction is essential to proper understanding and application of the data

Internal accounting records & CSDRs typically reflect most-to-all of these attributes

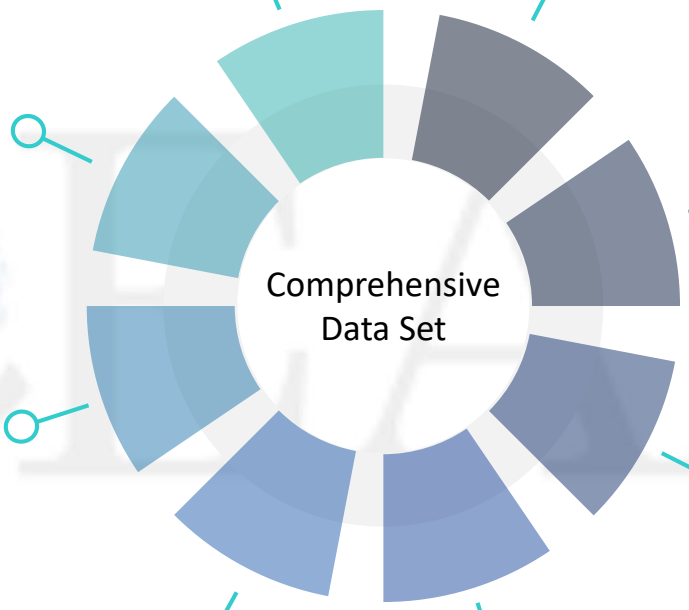
Data Sources – Example Cost Reporting Structure

WBS Element Code	WBS Element Name
1.0	MIM-1000A Xample Missile System
1.1	Air Vehicle
1.1.1	Propulsion
1.1.2	Payload
1.1.2.1	Warhead
1.1.2.2	Safe Arm Device
1.1.2.3	Payload Integration, Assembly, Test, & Checkout
1.1.3	Airframe
1.1.4	Reentry System
1.1.5	Post Boost System
1.1.6	Guidance and Control
1.1.6.1	Guidance Section
1.1.6.2	Control/Actuators System
1.1.6.3	G&C Application Software
1.1.6.4	G & C Integration, Assembly, Test & Checkout
1.1.7	Ordnance Initiation Set
1.1.8	Airborne Test Equipment
1.1.9	Airborne Training Equipment
1.1.10	Auxiliary Equipment
1.1.11	Integration, Assembly, Test and Checkout
1.2	Command and Launch
1.2.1	Surveillance, Identification and Tracking Sensors
1.2.2	Launch and Guidance Control
1.2.3	Communications
1.2.4	Command and Launch Applications Software
1.2.5	Command and Launch System Software
1.2.6	Launcher Equipment
1.2.6.1	Canister
1.2.6.2	Canister/Missile Integration, Assembly, Test, & Checkout
1.2.6.3	Command & Launch Unit
1.2.7	Auxiliary Equipment
1.2.8	Booster Adapter
1.3	System Engineering/Program Management

WBS Element Code	WBS Element Name
1.4	System Test and Evaluation
1.4.1	Development Test and Evaluation
1.4.2	Operational Test and Evaluation
1.4.3	Mock-ups / System Integration Labs (SILs)
1.4.4	Test and Evaluation Support
1.4.5	Test Facilities
1.5	Training
1.5.1	Equipment
1.5.2	Services
1.5.3	Facilities
1.6	Data
1.6.1	Technical Publications
1.6.2	Engineering Data
1.6.3	Management Data
1.6.4	Support Data
1.6.5	Data Depository
1.7	Peculiar Support Equipment
1.7.1	Test and Measurement Equipment
1.7.2	Support and Handling Equipment
1.8	Common Support Equipment
1.8.1	Test and Measurement Equipment
1.8.2	Support and Handling Equipment
1.9	Operational/Site Activation
1.10	Industrial Facilities
1.11	Initial Spares and Repair Parts

A hierarchical, product-oriented WBS facilitates clarity & completeness

Data Sources – CSDR Overview



CSDR Plan (DD 2794)
Defines WBS Structure and Schedule for reporting

Functional Cost-Hour Report (DD 1921-1)
Cost and hour data for specific elements broken down by functional category

Contract Work Breakdown Structure Dictionary (CWBS)
Lists all contract data reporting elements with definitions of technical, work, and cost content (recurring vs. nonrecurring by functional category)

Progress Curve Report (DD 1921-2)
Lot or Unit reporting of direct recurring costs to date for specific hardware elements

Software Resource Data Reports (SRDRs)
Size, schedule & effort on data on software development

Contractor Business Data Report (DD 1921-3)
Direct & Indirect cost, hour & employee data by functional category for a Contractor Business Unit

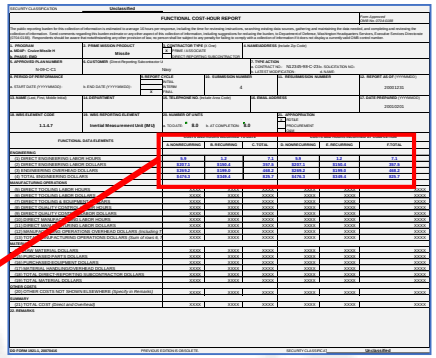
Cost Data Summary Report (DD 1921)
Summary cost data for all elements on approved contract CSDR plan

Sustainment Report (DD 1921-5)
Nonrecurring & recurring costs and hours reported against a sustainment cost element structure & functional categories

An institutional approach to cost data collection mandated by DoD leadership

Using CSDRs to Develop Methodology (p. 1 of 3)

Assume DoD needs to develop a new inertial measurement unit (IMU) for a new cruise missile. This example addresses collection of historical missile program data to enable estimation of the new missile's IMU.



CSDR Functional Cost-Hour Report (DD 1921-1)

18. WBS ELEMENT CODE	19. WBS REPORTING ELEMENT	20. NUMBER OF UNITS	21. APPROPRIATION
1.1.4.7	Inertial Measurement Unit (IMU)	a. TO DATE: 8.0 b. AT COMPLETION: 8.0	X RDT&E PROCUREMENT O&M

FUNCTIONAL DATA ELEMENTS	COSTS AND HOURS INCURRED TO DATE			COSTS AND HOURS INCURRED AT COMPLETION		
	A. NONRECURRING	B. RECURRING	C. TOTAL	D. NONRECURRING	E. RECURRING	F.TOTAL
ENGINEERING						
(1) DIRECT ENGINEERING LABOR HOURS	5.9	1.2	7.1	5.9	1.2	7.1
(2) DIRECT ENGINEERING LABOR DOLLARS	$207.1	$150.4	357.5	$207.1	$150.4	357.5
(3) ENGINEERING OVERHEAD DOLLARS	$269.2	$199.0	468.2	$269.2	$199.0	468.2
(4) TOTAL ENGINEERING DOLLARS	$476.3	$349.4	825.7	$476.3	$349.4	825.7

Develop historical data set required to derive a parametric relationship for estimating design engineering hours for a new IMU, where:

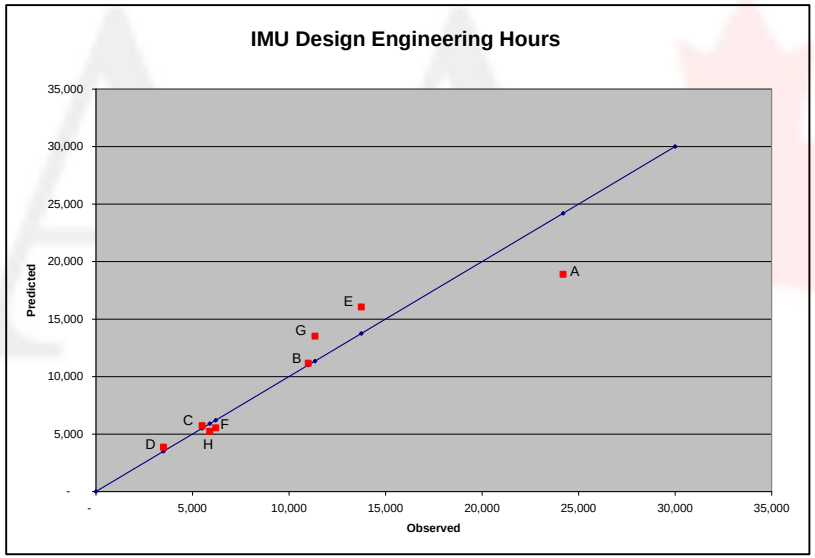
Design Engr Hrs = f (Bias, Volume, Schedule)

	Inertial Measurement Unit			
	Nonrecurring Engineering Hours	Bias	Volume	Schedule
Program A				
B				
C				
D				
E				
F				
G				
H	5,900			

Using CSDRs to Develop Methodology (p. 2 of 3)

IMU Design Engineering Hours = 46.5 * (Bias)$^{-0.47}$ * (Volume)$^{0.43}$ * (Schedule)$^{1.12}$

	Inertial Measurement Unit			
	Nonrecurring Engineering Hours	Bias (degrees)	Volume (cu. cm.)	Schedule (months)
Program A	24,200	1	30	54
B	11,000	2	40	40
C	5,500	2.5	20	32
D	3,500	10	30	34
E	13,750	1	28	48
F	6,200	2	30	24
G	11,350	1	30	40
H	5,900	5	25	36



Data collected from programmatic and schedule documents

Data collected from technical documents

Data collected from CSDR Form 1921-1 (Functional Cost-Hour Report)

Using CSDRs to Develop Methodology (p. 3 of 3)

IMU Design Engineering Hours = 46.5 * (Bias)$^{-0.47}$ * (Volume)$^{0.43}$ * (Schedule)$^{1.12}$

INERTIAL MEASUREMENT UNIT

	Nonrecurring Engineering Hours	Bias (degrees)	Volume (cu. Cm.)	Schedule (months)
Generic Supersonic Cruise Missile	11,356.7	1	24	40
	ESTIMATE			

	11,356.7	Hrs
$	40.00	Engineering Labor Rate
$	454,268.0	Direct Engineering Labor Dollars
	115%	Engineering Overhead Rate
$	522,408.2	Engineering Overhead Dollars
$	120,000.00	Material Dollars
$	46,787.0	ODC Dollars
$	1,143,464.4	Total Engineering Dollars

Data Normalization

- Cost
  - Correct for price level changes
  - Cost vs. price
  - $M vs. $K
- Quantity
  - Unit vs. lot cost
  - Full-up vs. partial unit
- Sizing
  - Weight (kg vs. lbs)
  - Length (meters vs. feet)
  - Physical vs. logical lines of code
- Content
  - Recurring vs. non-recurring
  - Fixed vs. variable
- State of Development Maturity
  - Technology readiness level
  - Prototype vs. low-rate initial production vs. full production
- Key Groupings
  - By mission application (e.g., bomber vs. fighter vs. helicopter vs. trainer)
  - By operating environment (e.g., manned space vs. unmanned space vs. aerospace vs. shipboard vs. commercial)

Essential to ensuring data consistency & comparability

In Summary

- Data is the lifeblood of cost analysts
- The best cost data is actuals for completed contracts
- High performing cost analysts are highly skilled in and devote the requisite time to collecting and normalizing data
  - They understand the right questions to ask to ascertain what a specific value means (or doesn't) and includes (or doesn't)
  - They understand what if any adjustments the data requires to ensure utility, consistency and comparability
- Though difficult to quantify ROI, the value of an institutional approach to collecting authoritative cost data is significant
  - Authoritative data enables credible and defensible estimates that facilitate better decisions re. budgets and contracts
  - Authoritative data enables credible and defensible estimates that engender decision maker confidence/trust

Cost analysts must be vigilant about getting the 'right' data

Data: The Means to Avoid this Situation