Dimensional INSIGHT

HIMSS NEW ENGLAND Chapter

Advanced Analytics: *Back to Basics?*

Data Driven Decision Making in the 21st Century

George Dealy
Vice President, Healthcare Solutions
Dimensional Insight, Inc.

HIMSS Regional Conference, Portland, Maine
November 21, 2019

Dimensional INSIGHT

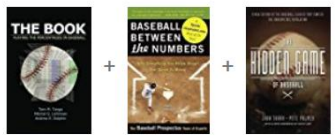
Data Driven Decision Making

Data Driven Decision Making is Pervasive

Google Maps

amazon

Frequently bought together



Total price: $52.76

Add all three to Cart

Add all three to List

One of these items ships sooner than the other. Show details

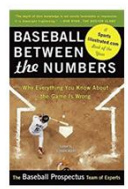
This item: The Book: Playing The Percentages In Baseball by Tom Tango Paperback $21.95

Baseball Between the Numbers: Why Everything You Know About the Game Is Wrong by The Baseball Prospectus Team of Experts Paperback $11.87

The Hidden Game of Baseball: A Revolutionary Approach to Baseball and Its Statistics by John Thorn Paperback $18.94

Customers who bought this item also bought

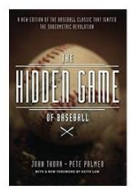
Page 1 of 25



Baseball Between the Numbers: Why Everything You Know About the...

The Baseball...

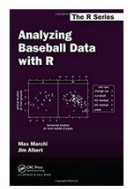
52



The Hidden Game of Baseball: A Revolutionary Approach to Baseball...

John Thorn

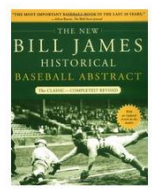
15



Analyzing Baseball Data with R (Chapman & Hall/CRC The R Series)

› Max Marchi

29



The New Bill James Historical Baseball Abstract

› Bill James

149

tripadvisor

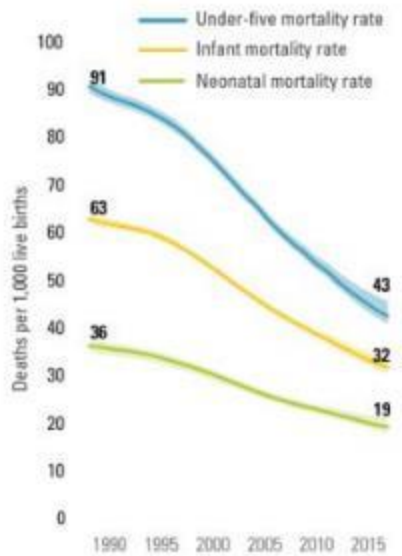
DRAFT KINGS

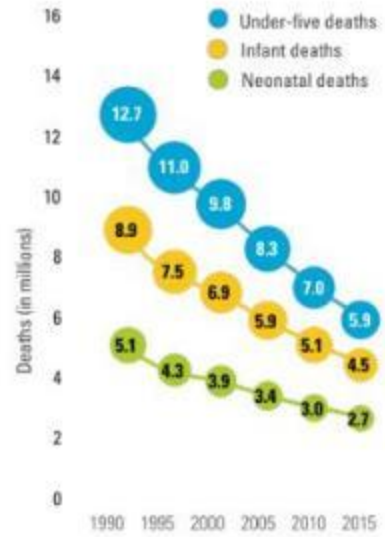
Expedia

DATA.GOV

Impact of Data Driven Decisions in Healthcare

The under-five mortality rate has fallen by more than half since 1990. That represents 4.4 million children who didn't die!





- Breastfeed for at least the first six months
- Deliver antibiotics immediately if a baby becomes ill.
- Use hand-pumped oxygen masks and basic resuscitation training.
- Increase skin-to-skin contact with the baby
- Keep the baby warm and dry.
- Keep the umbilical cord clean to prevent infection.

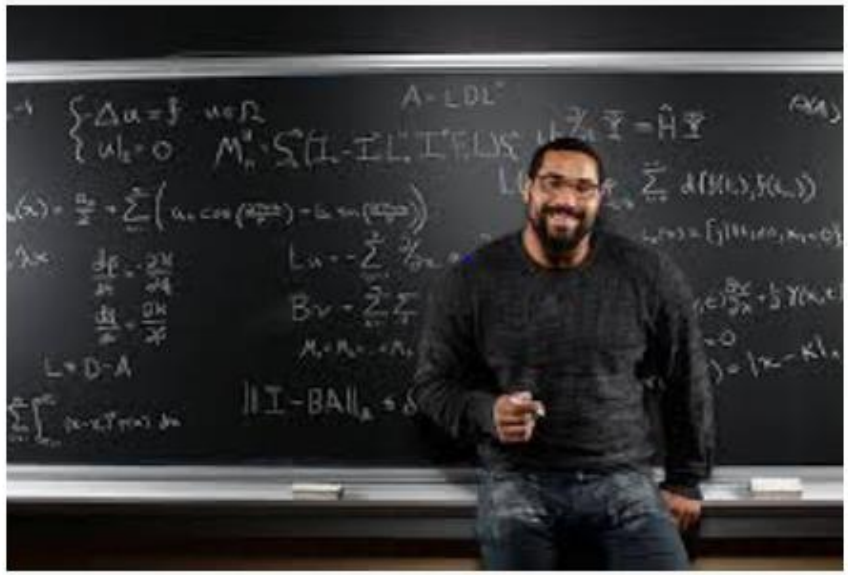
Globally, the infant mortality rate has decreased from an estimated rate of 63 deaths per 1000 live births in 1990 to 32 deaths per 1000 live births in 2015. Annual infant deaths have declined from 8.9 million in 1990 to 4.5 million in 2015.

Source: *unicef*

John Urschel

Former NFL Lineman, Baltimore Ravens

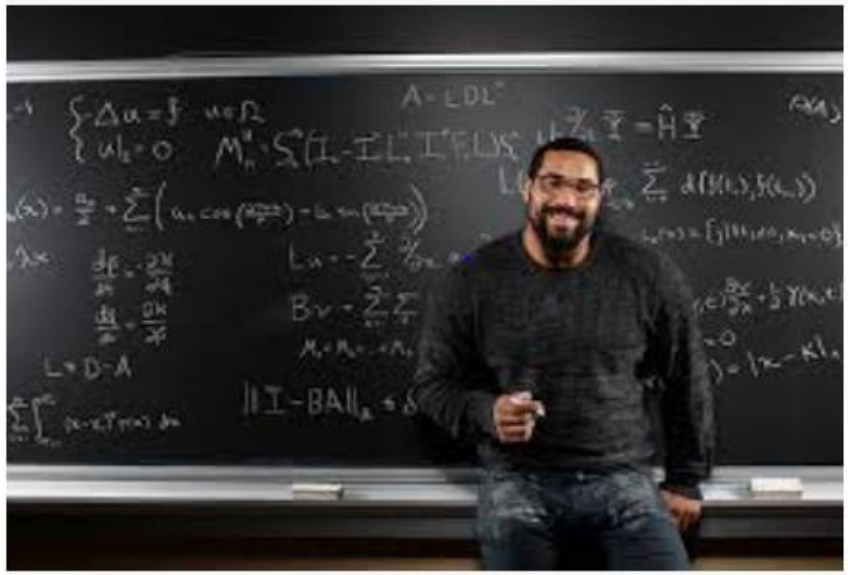
PhD Candidate in Mathematics, MIT



John Urschel

Former NFL Lineman, Baltimore Ravens

PhD Candidate in Mathematics, MIT



“Being mathematically illiterate is quite a dangerous thing.”

$$P(A \mid B) = \frac{P(B \mid A) * P(A)}{P(B)}$$

Bayes Theorem

$$P(H \mid E) = \frac{P(E \mid H) * P(H)}{P(E)}$$

H = Hypothesis

E = Evidence

Bayes Theorem

$$\overbrace{P(H \mid E)}^{\text{Posterior}} = \frac{P(E \mid H) * \overbrace{P(H)}^{\text{Prior}}}{P(E)}$$

H = Hypothesis

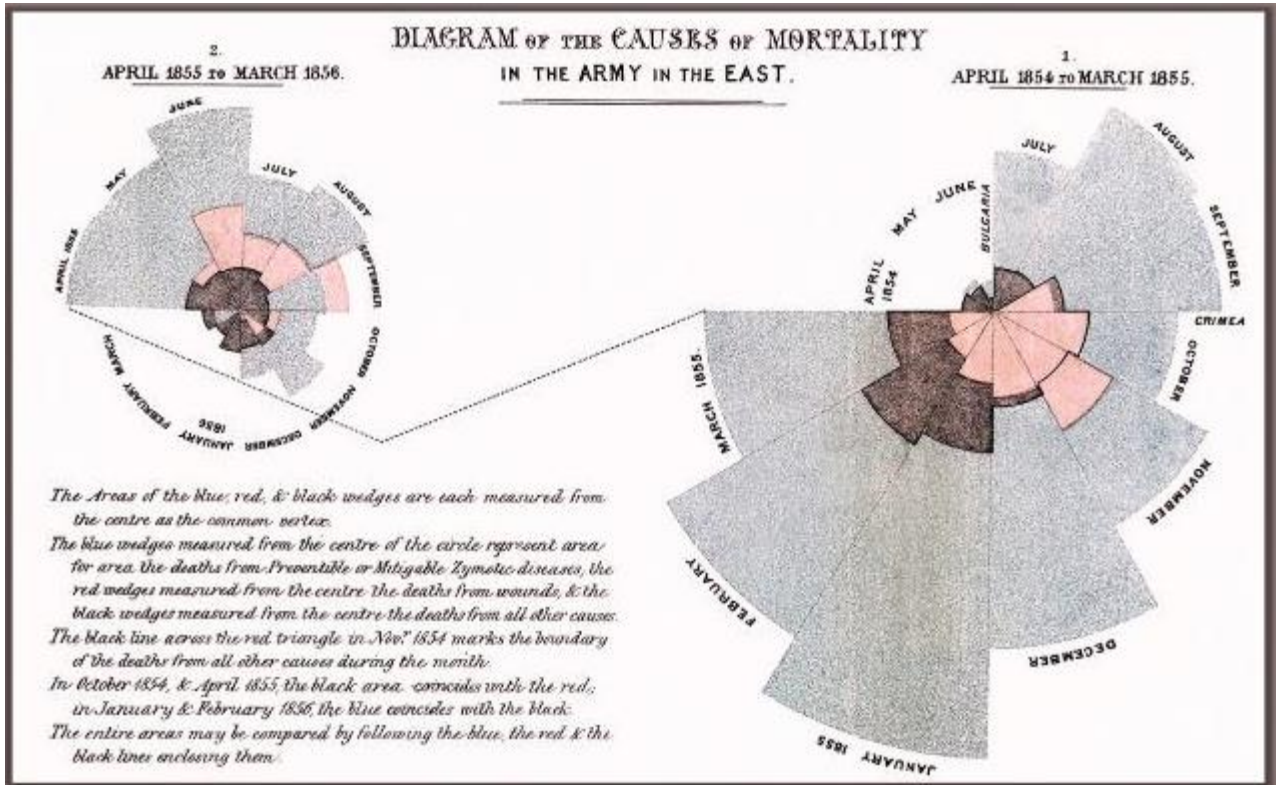
E = Evidence

A Long Standing Precedent in Healthcare

Florence Nightingale

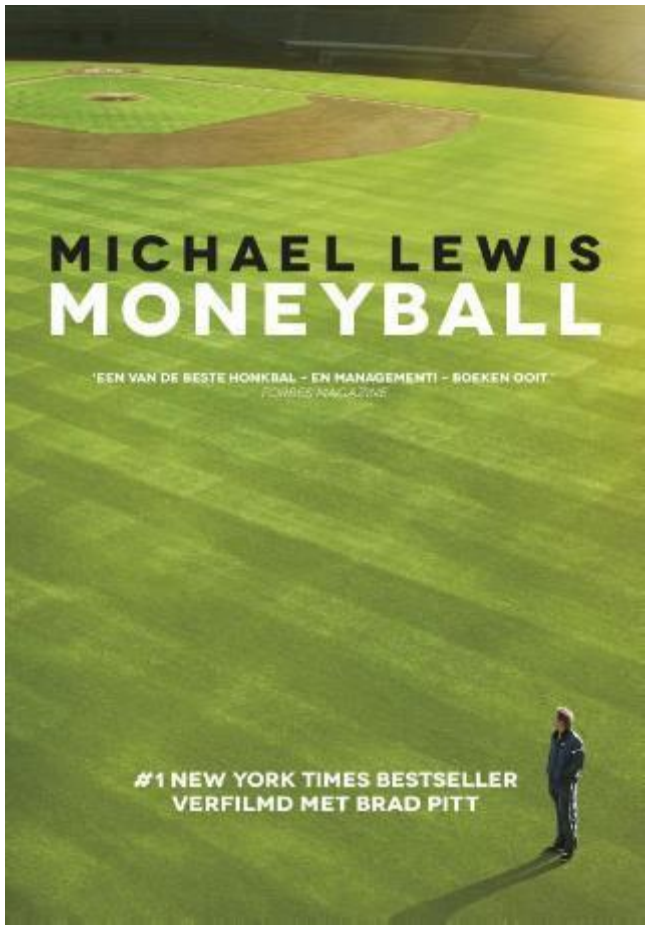
- Founder of Nursing
- Public Health Pioneer
- Statistician

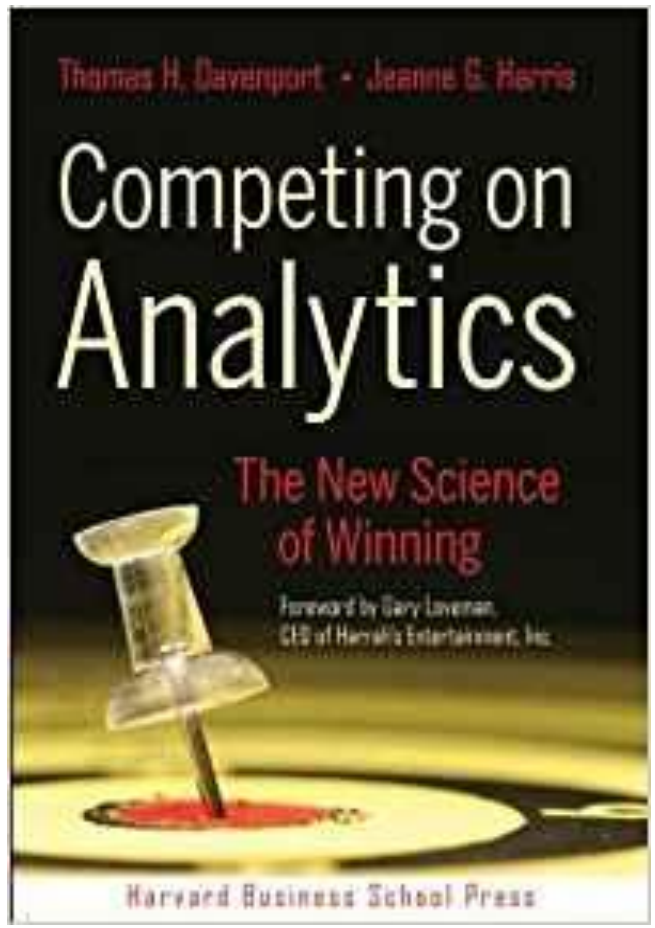
"Rose Chart" – shows the impact of poor sanitation and infectious disease on mortality rate among British soldiers during the Crimean War 1854-1856.



Our National Pastime

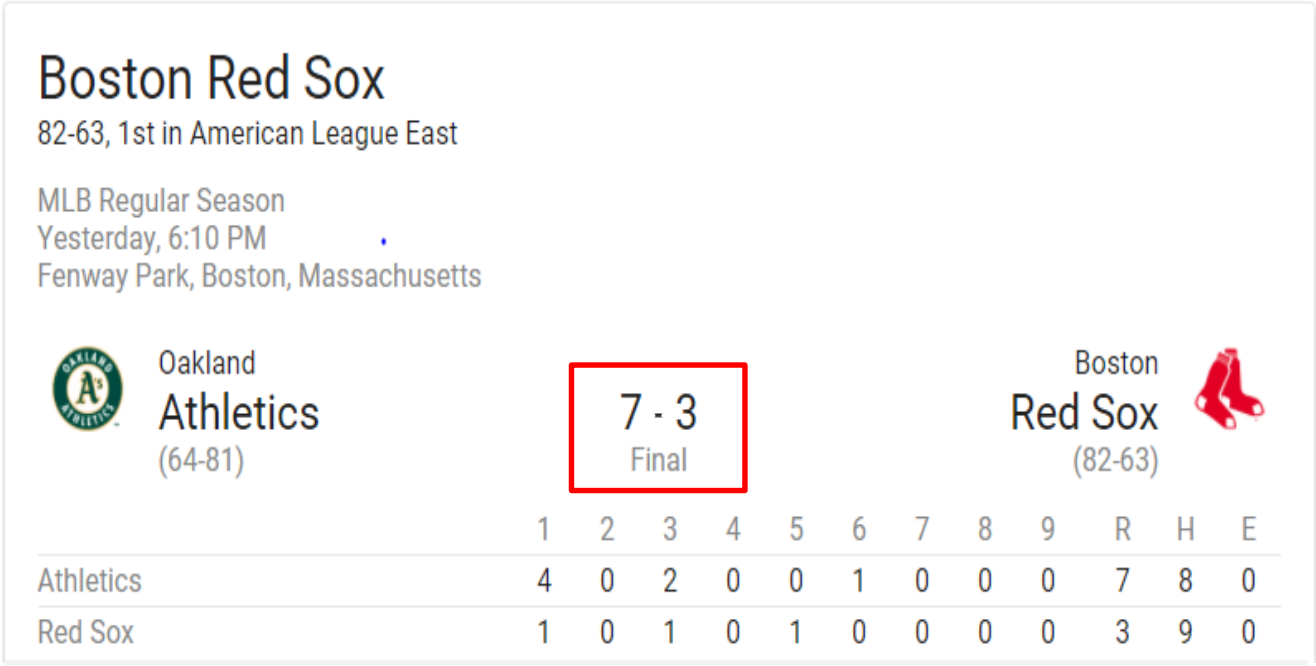
Baseball's New Frontier: Analytics





Focusing on Outcomes

Baseball



Boston Red Sox

82-63, 1st in American League East

MLB Regular Season
Yesterday, 6:10 PM
Fenway Park, Boston, Massachusetts

Oakland Athletics (64-81)

7 - 3
Final

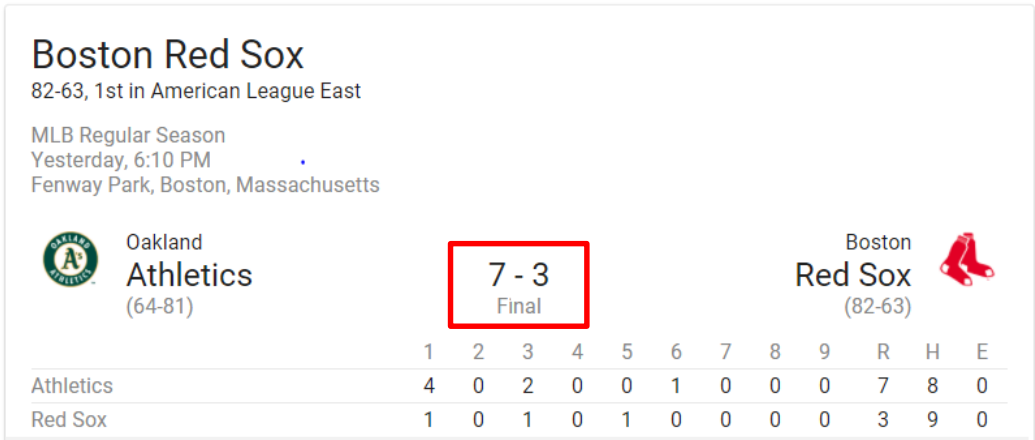
Boston Red Sox (82-63)

	1	2	3	4	5	6	7	8	9	R	H	E
Athletics	4	0	2	0	0	1	0	0	0	7	8	0
Red Sox	1	0	1	0	1	0	0	0	0	3	9	0

- Runs Scored / Prevented
- Games Won / Lost

Focusing on Outcomes

Baseball



Boston Red Sox
82-63, 1st in American League East

MLB Regular Season
Yesterday, 6:10 PM
Fenway Park, Boston, Massachusetts

Oakland Athletics (64-81) — 7 - 3 Final — Boston Red Sox (82-63)

	1	2	3	4	5	6	7	8	9	R	H	E
Athletics	4	0	2	0	0	1	0	0	0	7	8	0
Red Sox	1	0	1	0	1	0	0	0	0	3	9	0

- Runs Scored / Prevented
- Games Won / Lost

Healthcare

Population
- Years of Potential Life Lost
- Disease Incidence

Quality & Safety
- Medical Errors
- Preventable Admissions

Cost
- Per Person per Month
- Lifetime

Patient Experience
- Patient Centered Outcomes
- Self Perceived Wellness

What Impacts Outcomes?

Baseball

Year	AB	R	H	HR	RBI	SB	AVG	OBP
2017 Stats	512	80	137	8	54	15	.268	.333
MLB Career Stats	2359	346	665	49	275	41	.282	.336

- On Base Percentage (OBP)
- Slugging Percentage (SLG)
- Walks & Hits per Inning Pitched (WHIPS)

What Impacts Outcomes?

Baseball

Year	AB	R	H	HR	RBI	SB	AVG	OBP
2017 Stats	512	80	137	8	54	15	.268	.333
MLB Career Stats	2359	346	665	49	275	41	.282	.336

- On Base Percentage (OBP)
- Slugging Percentage (SLG)
- Walks & Hits per Inning Pitched (WHIPS)

Healthcare

Population

- Socio-economic Determinants
- Environment
- Lifestyle

Quality & Safety

- Evidence-based Care Compliance
- Coordination of Care

Cost of Care

- Efficiency
- Utilization

Patient Experience

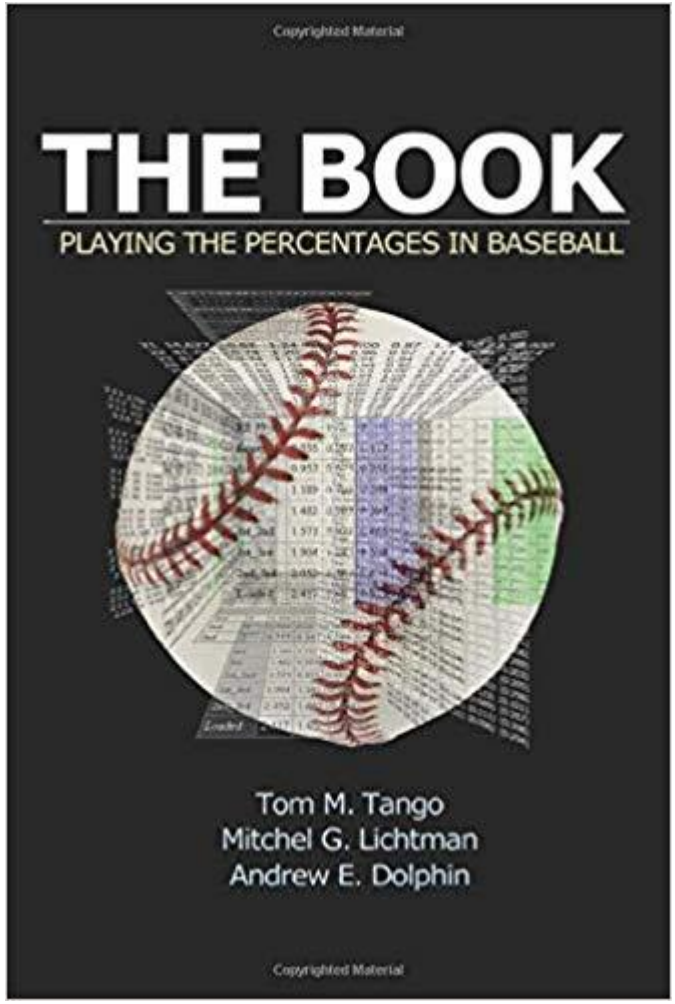
- Patient Engagement
- Compliance with Prescribed Care

Getting to Data Driven Decisions . . .

. . . Isn’t Easy.

Moneyball Clip

Data Driven Baseball



Old Wives Tales

- Hot & Cold Streaks: Batting
- Hot & Cold Streaks: Pitching

Evidence-based Practice

- Batting Order
- Pitcher's Days Rest
- Designated & Pinch Hitters

World Series Away Game Winners Analysis

The 2019 World Series was the first in history in which the *Away* team always won.

- How unusual is this -- really?
- How significant is the home field advantage?

Home / Away Winner Combinations in a 7 Game Series

	Game 1	Game 2	Game 3	Game 4	Game 5	Game 6	Game 7
Team A Wins	A - Home	A - Home	A - Away	B - Home	B - Home	B - Away	A - Home
	A - Home	A - Home	B - Home	A - Away	B - Home	B - Away	A - Home
	A - Home	A - Home	B - Home	B - Home	A - Away	B - Away	A - Home
	A - Home	A - Home	B - Home	B - Home	B - Home	A - Home	A - Home
	A - Home	B - Away	A - Away	A - Away	B - Home	B - Away	A - Home
	A - Home	B - Away	A - Away	B - Home	A - Away	B - Away	A - Home
	A - Home	B - Away	A - Away	B - Home	B - Home	A - Home	A - Home
	A - Home	B - Away	B - Home	A - Away	A - Away	B - Away	A - Home
	A - Home	B - Away	B - Home	A - Away	B - Home	A - Home	A - Home
	A - Home	B - Away	B - Home	B - Home	A - Away	A - Home	A - Home
	B - Away	A - Home	A - Away	A - Away	B - Home	B - Away	A - Home
	B - Away	A - Home	A - Away	B - Home	A - Away	B - Away	A - Home
	B - Away	A - Home	A - Away	B - Home	B - Home	A - Home	A - Home
	B - Away	B - Away	A - Away	A - Away	A - Away	B - Away	A - Home
	B - Away	B - Away	A - Away	A - Away	B - Home	A - Home	A - Home
	B - Away	B - Away	B - Home	A - Away	A - Away	A - Home	A - Home

Home / Away Winner Combinations in a 7 Game Series

	Game 1	Game 2	Game 3	Game 4	Game 5	Game 6	Game 7
Team B Wins	B - Away	B - Away	B - Home	A - Away	A - Away	A - Home	B - Away
	B - Away	B - Away	A - Away	B - Home	A - Away	A - Home	B - Away
	B - Away	B - Away	A - Away	A - Away	B - Home	A - Home	B - Away
	B - Away	B - Away	A - Away	A - Away	A - Away	B - Away	B - Away
	B - Away	A - Home	B - Home	B - Home	A - Away	A - Home	B - Away
	B - Away	A - Home	B - Home	A - Away	B - Home	A - Home	B - Away
	B - Away	A - Home	B - Home	A - Away	A - Away	B - Away	B - Away
	B - Away	A - Home	A - Away	B - Home	B - Home	A - Home	B - Away
	B - Away	A - Home	A - Away	B - Home	A - Away	B - Away	B - Away
	B - Away	A - Home	A - Away	A - Away	B - Home	B - Away	B - Away
	A - Home	B - Away	B - Home	B - Home	A - Away	A - Home	B - Away
	A - Home	B - Away	B - Home	A - Away	B - Home	A - Home	B - Away
	A - Home	B - Away	B - Home	A - Away	A - Away	B - Away	B - Away
	A - Home	A - Home	B - Home	B - Home	B - Home	A - Home	B - Away
	A - Home	A - Home	B - Home	B - Home	A - Away	B - Away	B - Away
	A - Home	A - Home	A - Away	B - Home	B - Home	B - Away	B - Away

World Series Away Game Winners Analysis

Number of combinations in a *Best of 7* series that goes to the full 7 games	32
Combinations where *Away* team always wins	1
Probability of *Away* team always winning	1 in 32 3.1%

World Series Away Game Winners Analysis

Number of World Series Games 1903 - 2019	115
Series that went to 7 Games	40
Times the *Away* team always won	1
Historical probability that *Away* team always wins	1 in 40 2.5%

Example from the Real World of Healthcare

CMS Comprehensive Joint Replacement Bundle Program (CJR)

- Objective: *Aims to support better and more efficient care for beneficiaries undergoing the most common inpatient surgeries for Medicare beneficiaries: hip and knee replacements.*
- *Outcome Measures*
  - 90 Day Episode Cost vs. Target
  - Patient Experience
    - HCAHPS Survey
    - Patient Centered Outcomes (optional)
  - Complications of Care

How can we impact these?

Monitoring Performance on CJR Bundle Program

Year-to-Date Performance from Jan 1st to August 30th

90 Day Episode Cost	Unplanned Readmissions	Complications of Care
Median $25,198	Rate 4.5%	Rate 1.2%

Patient Exp: HCAHPS	Pat. Centered Outcomes	Provider Experience
Percentile 85	Normalized Score 75	Composite Score 90

What can you do with this?

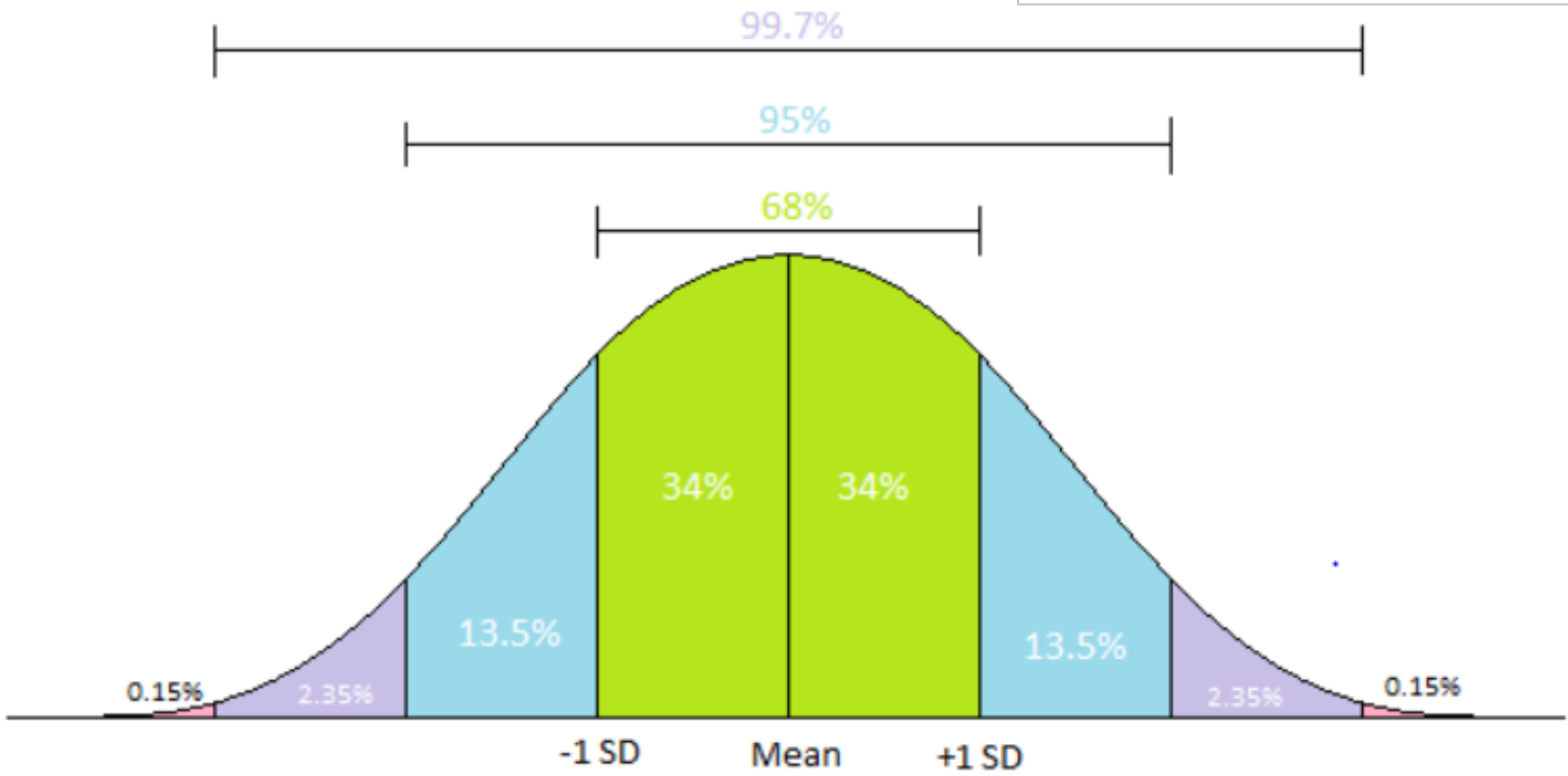
Numeracy: Basic Concepts & Tools

- Central Tendency: "The Middle"
  - Mean (Average)
  - Median
- Variation
  - Standard Deviation
  - Interquartile Range
  - Over Time
- Distribution
  - Normal
  - Something other than Normal

Normal Distribution & Standard Deviation

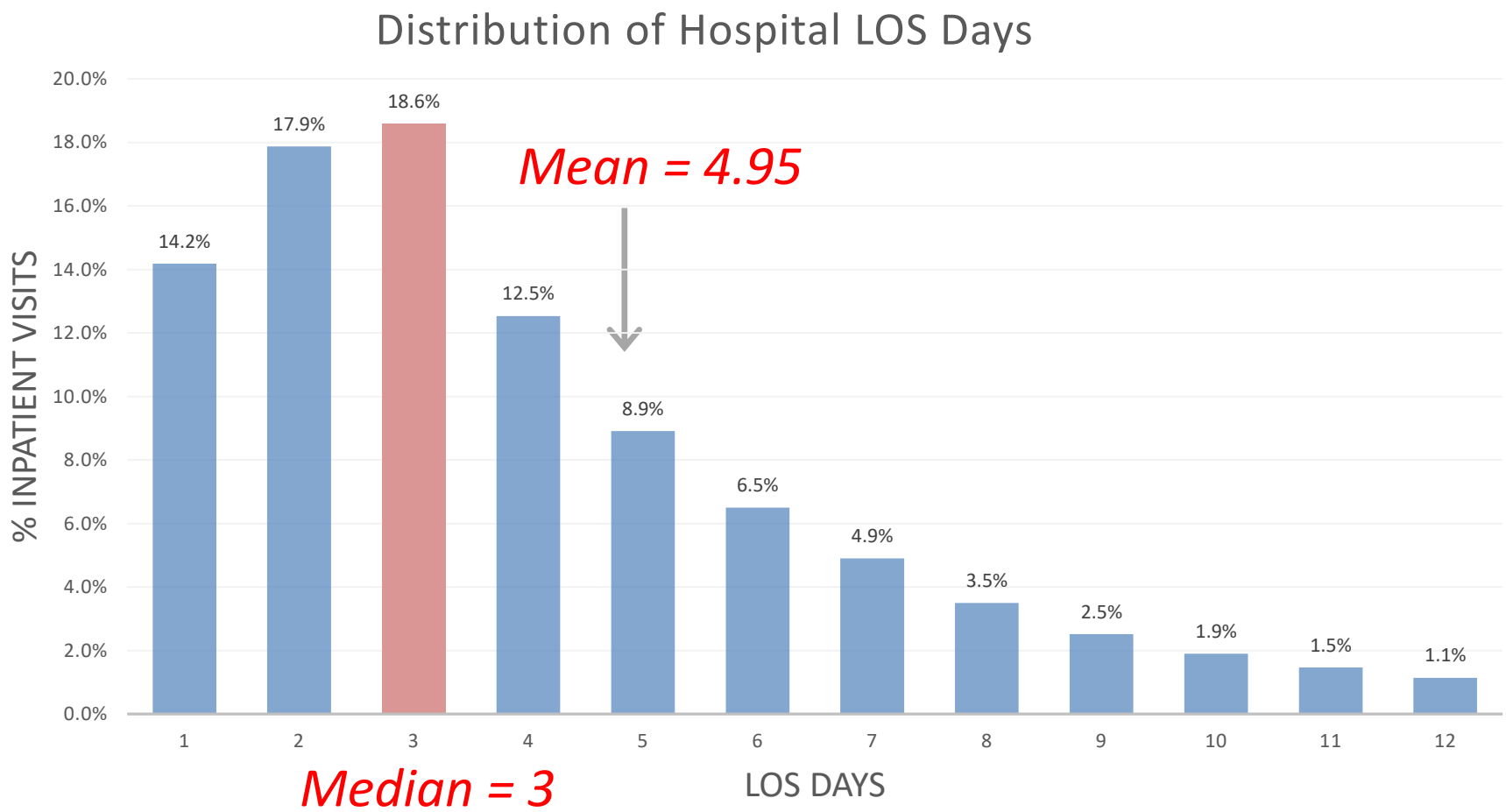
The underlying assumption for many statistical techniques

Beyond +/- 2 SD's: < 5% probability of value occurring in a the actual population.



Normal Distribution & Standard Deviation

But often the assumption doesn't hold up!



Box Plot & Interquartile Range

Working with skewed distributions using rank & relative standing

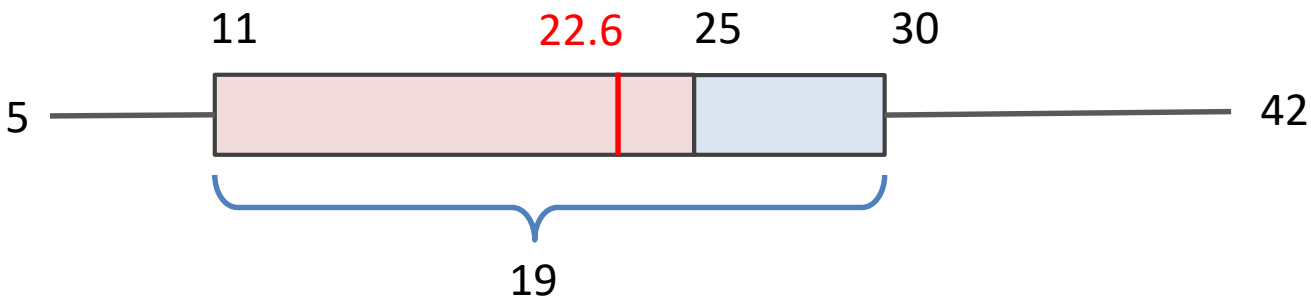
Five Number Summary

Min	1st Quartile	Median	3rd Quartile	Max
5	11	25	30	42

IQR
19

Mean = 22.6

Boxplot



John Tukey

The Future of Data Analysis, 1962
Exploratory Data Analysis, 1977

Applying Numeracy to the CJR Bundle Program

DRG 470 w/o Fracture 90 Day Episode Cost Summary

Our Median	CMS Target	+/-	% +/-
$25,149	$22,905	+2,290	+10.0%

Population: N = 589 procedures

7 Number Summary

Min	1st Quartile	Target	Median	Mean	3rd Quartile	Max
$18,430	$22,506	$22,905	$25,198	$25,698	$27,482	$50,196

IQR
$4,975

Box Plot

Looking for Patterns of Influence: Performing Surgeons

Principal Procedure Provider	Discharges	Mean Cost	Median Cost	IQR	Rank	Quartile
Aguirre, L (PHY_001367)	614	20,628	20,129	2,387	2	4
Gay, W (PHY_006173)	175	24,762	24,168	3,790	16	3
Bradford, S (PHY_015134)	99	26,103	25,948	2,523	27	3
Hawkins, A (PHY_001128)	95	24,194	23,964	3,313	14	4
Levy, C (PHY_000889)	93	24,142	23,444	2,632	13	4
Chase, J (PHY_001365)	93	25,307	24,770	4,695	20	3
Delacruz, Y (PHY_025250)	89	26,746	25,813	3,953	25	3
Miles, W (PHY_012926)	79	23,292	22,570	4,339	11	4
Mathews, K (PHY_026618)	79	23,203	23,157	3,061	12	4
Bernard, U (PHY_000647)	67	24,991	24,259	3,819	18	3
Marshall, V (PHY_015736)	64	26,703	26,439	3,111	32	2
Osborn, O (PHY_000819)	62	27,104	26,861	3,654	36	2
Tate, J (PHY_019982)	59	26,987	26,979	4,244	37	2
Gillespie, T (PHY_003778)	57	29,999	26,465	4,241	33	2
Ashley, U (PHY_003582)	52	26,849	26,780	4,427	35	2
Cotton, H (PHY_000495)	51	25,212	25,535	3,166	24	3
Hancock, K (PHY_019528)	46	27,847	27,309	3,187	40	2
Ochoa, J (PHY_000744)	36	25,601	25,090	2,477	22	3
Alvarez, Q (PHY_019909)	33	25,451	24,670	4,094	19	3
Mcdonald, H (PHY_002526)	32	24,655	24,188	2,834	17	3
Bryant, C (PHY_004340)	24	29,421	27,197	6,009	39	2
Martinez, A (PHY_004477)	23	28,179	28,064	2,259	41	1
Mclaughlin, N (PHY_000809)	16	23,117	21,968	2,590	8	4
Rocha, G (PHY_002484)	6	28,201	25,136	2,653	23	3
Maxwell, I (PHY_002240)	6	25,372	25,910	2,267	26	3

** Top 25 (of 53) surgeons by procedure volume account for 97% + of total volume.*

***Opportunities**:*

Highest Costs?

Lowest Costs?

Least Variation?

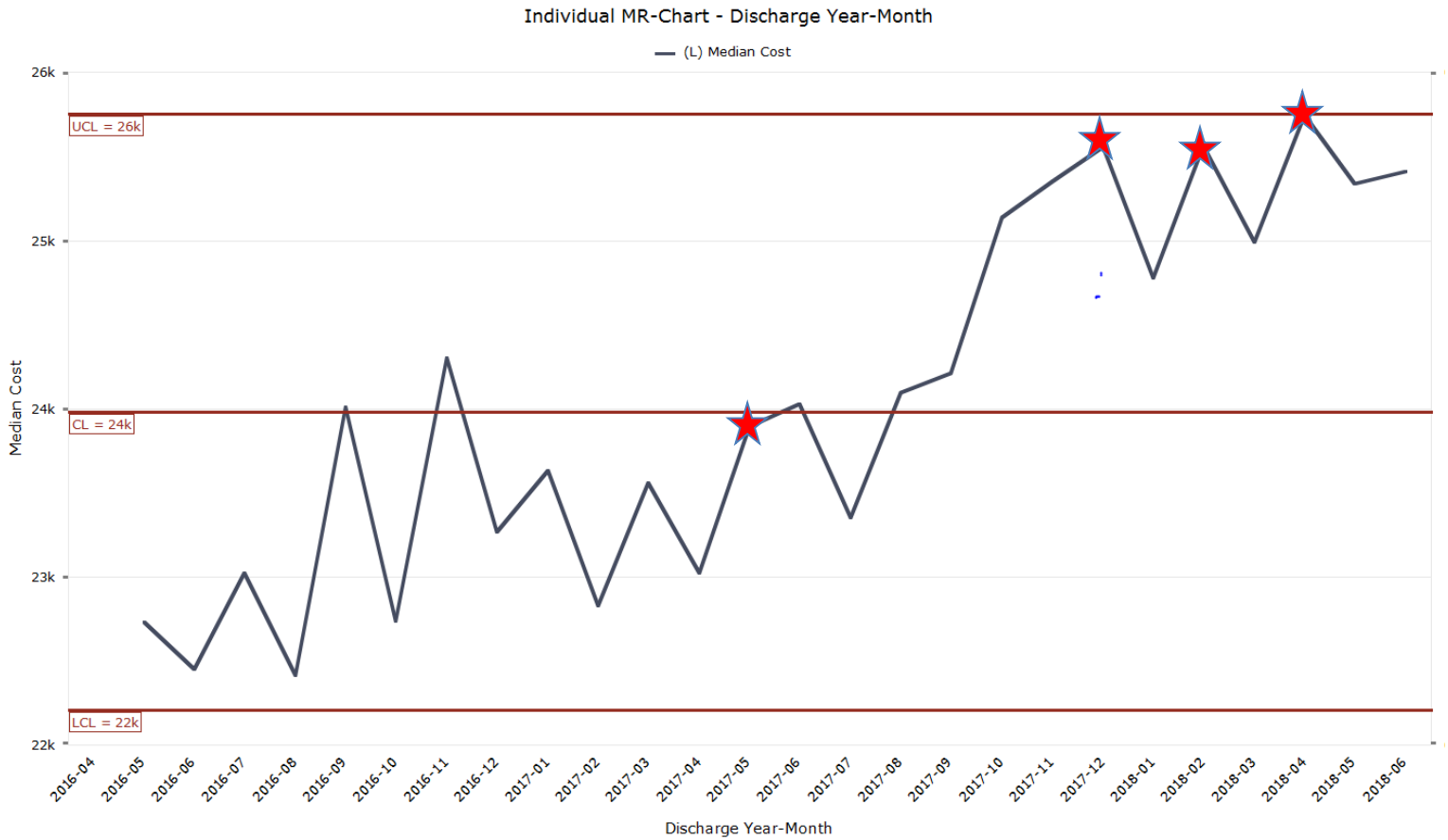
Most Variation?

Best Practices?

Inefficiencies?

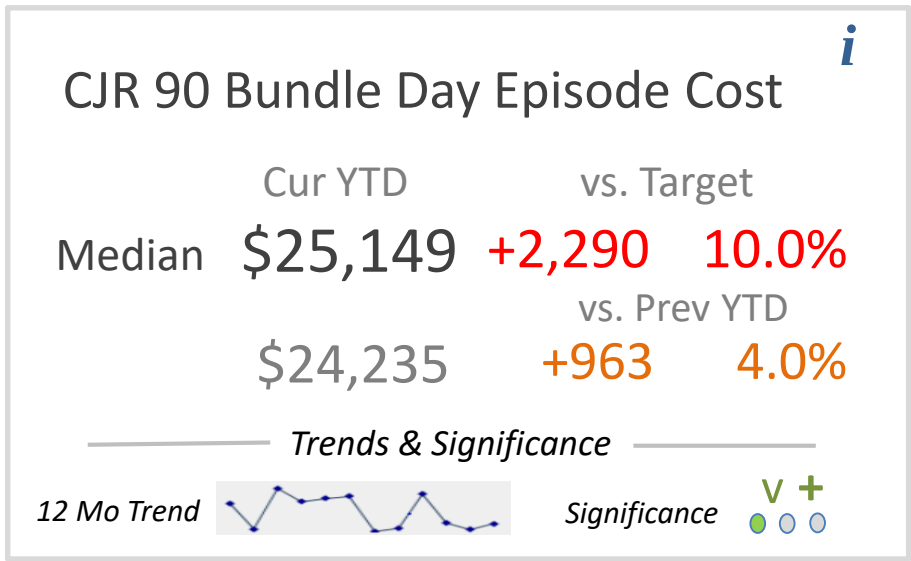
Detecting Significant Trends over Time

Statistical process control run charts can help detect whether change over time (such as an apparent improvement) is meaningful or just random, normal, expected variation.



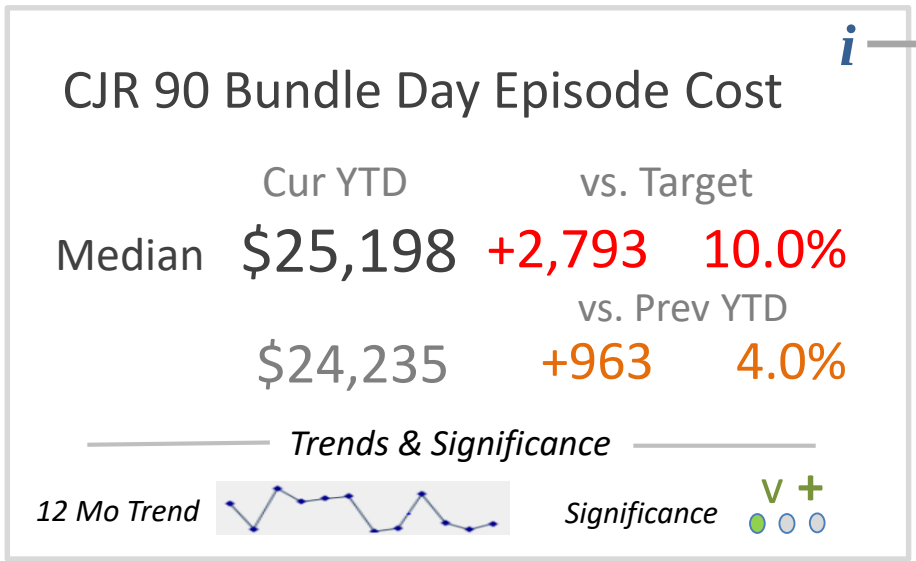
Monitoring Performance on CJR Bundle Program

Year-to-Date Performance from Jan 1st to August 30th



What can you do with this now?

Monitoring Performance on CJR Bundle Program



Definitions for Interpretation

90 Episode Cost

Definition:
Includes all payments made beginning 3 days before the procedure and ending 90 calendar days later. This includes all setting of care for which Medicare reimburses services.

Logic:
CJR Episode Costs / Episodes

Reference:
CMS Comprehensive Joint Replacement Bundle Program Specifications

Monitoring Performance on CJR Bundle Program

CJR 90 Bundle Day Episode Cost

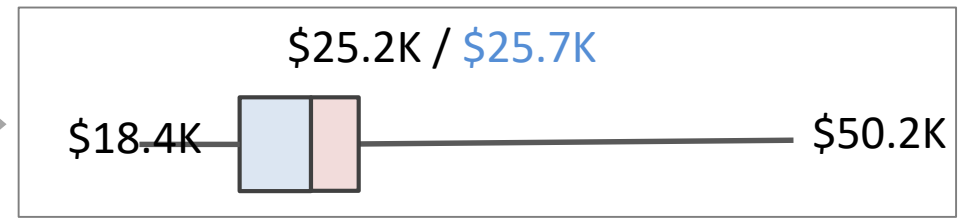
	Cur YTD	vs. Target	
Median	$25,198	+2,290	10.0%
		vs. Prev YTD	
	$24,235	+963	4.0%

Trends & Significance

12 Mo Trend 

Significance 

Distribution for Perspective



Monitoring Performance on CJR Bundle Program

CJR 90 Bundle Day Episode Cost

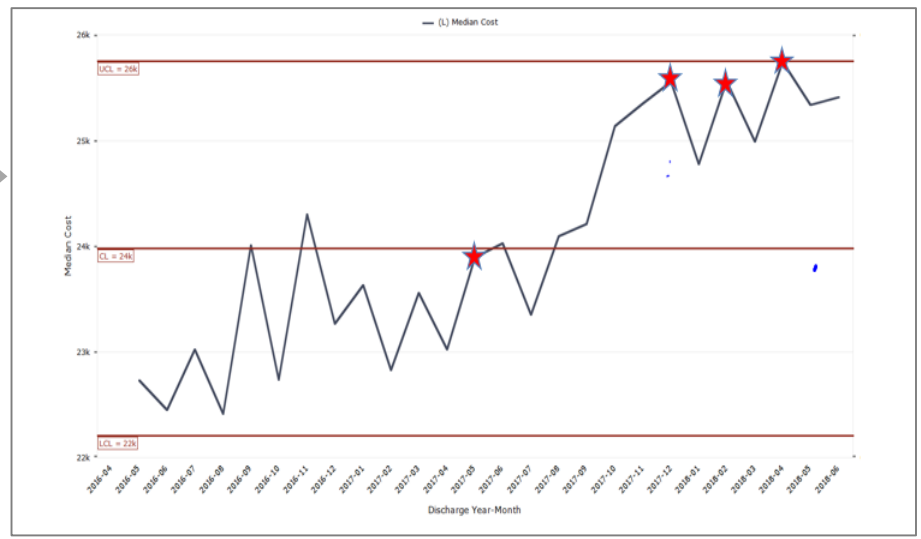
	Cur YTD	vs. Target	
Median	$25,198	+2,290	10.0%
		vs. Prev YTD	
	$24,235	+963	4.0%

Trends & Significance

12 Mo Trend 

Significance 

Trends for Context



Monitoring Performance on CJR Bundle Program

Details for Analysis

Principal Procedure Provider	Discharges	Mean Cost	Median Cost	IQR	Rank	Quartile
Aguirre, L (PHY_001367)	614	20,628	20,129	2,387	2	4
Gay, W (PHY_006173)	175	24,762	24,168	3,790	16	3
Bradford, S (PHY_015134)	99	26,103	25,948	2,523	27	3
Hawkins, A (PHY_001128)	95	24,194	23,964	3,313	14	4
Levy, C (PHY_000889)	93	24,142	23,444	2,632	13	4
Chase, J (PHY_001365)	93	25,307	24,770	4,695	20	3
Delacruz, Y (PHY_025250)	89	26,746	25,813	3,953	25	3
Miles, W (PHY_012926)	79	23,292	22,570	4,339	11	4
Mathews, K (PHY_026618)	79	23,203	23,157	3,061	12	4
Bernard, U (PHY_000647)	67	24,991	24,259	3,819	18	3
Marshall, V (PHY_015736)	64	26,703	26,439	3,111	32	2
Osborn, O (PHY_000819)	62	27,104	26,861	3,654	36	2
Tate, J (PHY_019982)	59	26,987	26,979	4,244	37	2
Gillespie, T (PHY_003778)	57	29,999	26,465	4,241	33	2
Ashley, U (PHY_003582)	52	26,849	26,780	4,427	35	2
Cotton, H (PHY_000495)	51	25,212	25,535	3,166	24	3
Hancock, K (PHY_019528)	46	27,847	27,309	3,187	40	2
Ochoa, J (PHY_000744)	36	25,601	25,090	2,477	22	3
Alvarez, Q (PHY_019909)	33	25,451	24,670	4,094	19	3
Mcdonald, H (PHY_002526)	32	24,655	24,188	2,834	17	3
Bryant, C (PHY_004340)	24	29,421	27,197	6,009	39	2
Martinez, A (PHY_004477)	23	28,179	28,064	2,259	41	1
Mclaughlin, N (PHY_000809)	16	23,117	21,968	2,590	8	4
Rocha, G (PHY_002484)	6	28,201	25,136	2,653	23	3
Maxwell, I (PHY_002240)	6	25,372	25,910	2,267	26	3

CJR 90 Bundle Day Episode Cost

	Cur YTD	vs. Target	
Median	$25,198	+2,290	10.0%
		vs. Prev YTD	
	$24,235	+963	4.0%

Trends & Significance

12 Mo Trend 

Significance 

A Few Words about Governance

More on Numeracy

- Analysis
- Interpretation
- Visualization
- Communication

See my *Practical Analysis* Blog!

- on LinkedIn (Connect with George Dealy)

Use Cases for Advanced Analytics

A glimpse into the future . . .

- Genetic Data & Personalized Medicine
- Artificial Intelligence (AI)-Supported Diagnostics
- Pattern Recognition in Comprehensive Population Datasets
- Better Opportunities to Understand & Impact Outcomes

Analysis in the 21st Century

- Big Data
- Machine Learning
- Pervasive Computing
- Artificial Intelligence

- Curiosity
- Passion
- Numeracy
- Communication

Don’t be afraid to

Think

Dimensional INSIGHT

HIMSS NEW ENGLAND Chapter

Advanced Analytics: *Back to Basics?*

Data Driven Decision Making in the 21st Century

George Dealy
Vice President, Healthcare Solutions
Dimensional Insight, Inc.

HIMSS Regional Conference, Portland, Maine
November 21, 2019