

BEYOND THE BOX: UTILIZATION OF UNDERLAYING JACS SIMULATION RESULTS FOR ADVANCED SRA AND JCL SERIES PAPER #A

JOHNSON SPACE CENTER

Exploration, Integration &
Science Directorate:

-Strategic Business Integration [XB]-

Steve Wilson & Mike Stelly

BEYOND THE BOX

AGENDA/SCHEDULE ESTIMATE

- 1. Opportunity Statement**
- 2. Solution Design**
- 3. Model Diagnostics - *(could run out of time here)***
- 4. Custom Visuals & Metrics - *(definitely run out of time here)***
- 5. Crazy Ideas & Forward Work - *(totally skippable)***

OPPORTUNITY STATEMENT

OPPORTUNITY STATEMENT

JCL = BIG, MAGICAL DATA?

- ...to what extent? **Big, fun answers, big fun policy** but generally **small, inflexible results sets**.
- **Legacy tools** (including *the ones that no longer exist*) have heavily relied on carbon-copy visuals & calculations within **Closed Boxes** 📞
 - ❑ *Mysterious calculations, often untraceable*
 - ❑ *Largely, same ole JCL metrics (duration cruciality much?)*
 - ❑ *Same standard JCL visuals (some unique)*
- Can this JCL thing this community spawned 10+ years ago **tell us more? Have we yet asked, honestly, What is its extensibility?**
- Lots of **good ideas** circulating... but **how many are implemented formally and widely?**



**JCL, SRA and the like
frameworks entail
unutilized potential
we can unlock
together.**

OPPORTUNITY STATEMENT

WE WANT TWO THINGS

➤ The Gateway programmatic analysis team, fast approaching its KDP-1, asked pressing questions...

1. *How can we use simulation results to assess the integrated model and underlying programmatic data?*
- B. *Is there a way to go beyond the typical results visuals and measures to give stakeholders what they could really want?*



BEST JCL
MEASURES
& VISUALS

MODEL
DIAGNOSTICS

Gateway pushed proactively for deep JCL data to drive model assessment & analysis.

SOLUTION DESIGN

SOLUTION DESIGN

CREATION: JCL TOOLBOX

- TRI's JACS offers a glimpse into the promise of **extensible** simulation analysis via its **cache file**
- This paper embraces the cache file and shows our nascent and evolving approach in transforming it into **new analytical measures and visuals** arsenal: Our **JCL Toolbox**
- This is the first in a perhaps unending series of papers we will offer on this topic; **stayed tuned for future NASA SCOPe and NCSS sessions.**



The JACS cache file in tandem with friendly applications are a bridge toward a better JCL analysis toolset.

SOLUTION DESIGN

WRANGLE THE MASSIVE CACHE FILE, MINE FOR USEFUL SIM DATA

- We've been sitting on this data horde.
- During each simulation, JACS writes to a **cache file** that captures all results
 - ❑ Large: for Gateway, between 600MG to +1GB
 - ❑ 307K rows...w/ analysis schedule clocking in @ ~12,000 rows @ 2000 iterations
- Do not let these numbers scare you: **Gateway went big** and tested (successfully) the limits of JACS, R, Excel, and the computational capability of our machines



*Gateway wanted more insight.
The cache file exists... so we
bent it, beat it, broke it to
visualize and analyze its data.*

Model Info

...

Calendars

...

Task Section

Task ID	Description	Context	General Info -->
1	Project Parent	Summary	...
2	Fab Unit #1	Task	...
3	Launch	Milestone	...
...

Task Start Dates (TSD)

Iterations Results -->

Task ID	Description	Point Estimate	Mean Date	1	2	3...	...	3000
1	Project Parent	Numeric Date A	Numeric Date X	ND[1,1]	ND[1,2]	ND[1,3]	...	ND [1,3000]
2	Fab Unit #1	Numeric Date B	Numeric Date Y	ND[2,1]	ND	ND	...	ND
3	Launch	Numeric Date C	Numeric Date Z	ND	ND	ND	...	ND
...

Task Finish Dates (TFD)

Iterations Results -->

Task ID	Description	Point Estimate	Mean Date	1	2	3...	...	3000
1	Project Parent	Numeric Date A	Numeric Date X	ND[1,1]	ND[1,2]	ND[1,3]	...	ND [1,3000]
2	Fab Unit #1	Numeric Date B	Numeric Date Y	ND[2,1]	ND	ND	...	ND
3	Launch	Numeric Date C	Numeric Date Z	ND	ND	ND	...	ND
...

Task Duration

Iterations Results -->

Task ID	Description	Point Estimate	Mean (Days)	1	2	3...	...	3000
1	Project Parent	Days A	Days X	Days [1,1]	Days [1,2]	Days	...	Days
...

CP (Critical Path)

Iterations Results -->

Task ID	Description	Point Estimate	Mean (aka Criticality)	1	2	3...	...	3000
1	Project Parent	1	0.75	1	1	0	...	1
2	Fab Unit #1	0	0.25	0	0	0	...	1
3	Launch	1	0.5	1	0	1	...	0
...

➤ Sections (by task) extend **vertically...**

➤ ...Iteration results from simulation extend **horizontally.**

➤ JACS's dynamic cache design allows for **same basic structure** regardless of schedule and simulation size

Cache file's robust build is perfect for simulation capture.

SOLUTION DESIGN CACHE FILE STRUCTURE

For Gateway saw its 307,000-row cache file and, behold, it was very good.



➤ Don't forget costs, which come in several forms.

➤ Like critical path section, risk section captures an activation indicator [0,1] per iteration.

➤ Data like this lends itself to conditional analysis...

Total (Cost)				Iterations Results -->				
Task ID	Description	Point Estimate	Mean \$	1	2	3...	...	3000
1	Project Parent	\$A	\$B	\$	\$	\$...	\$
...

TI Cost	Iterations Results -->
---------	------------------------

TD Cost	Iterations Results -->
---------	------------------------

Total Cost - Year 1	Iterations Results -->
---------------------	------------------------

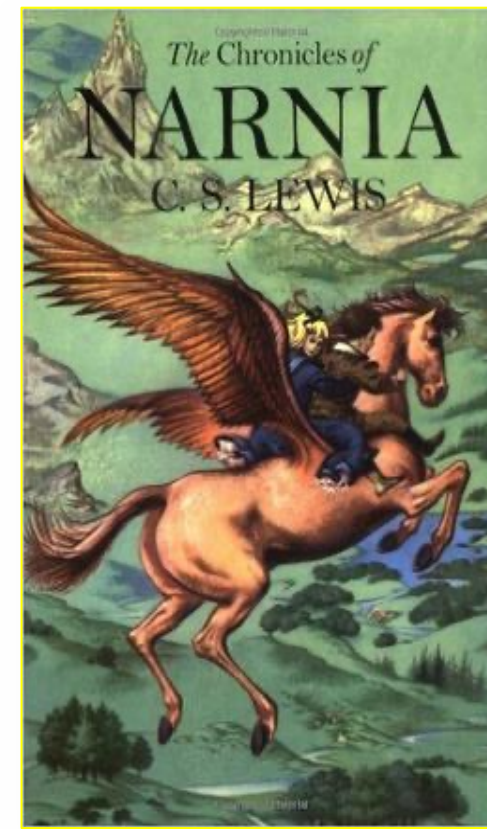
Total Cost - Year 2...	Iterations Results -->
------------------------	------------------------

IsActive (Risk)				Iterations Results -->				
Risk ID	Description	Point Estimate	Mean (aka Prob of Occurance)	1	2	3...	...	3000
R1	Test Failure	1	0.75	1	1	0	...	1
R2	Risk #2	0	0.25	0	0	0	...	1
R3	Risk #3	1	0.5	1	0	1	...	0
...

SOLUTION DESIGN

PROCESSING CACHE FILE

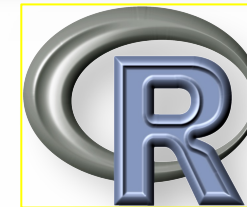
- **Started with Tableau Prep, but we oft make bad decisions...**
 - ❑ Prep is not to be trusted; yet another beta tool
 - ❑ Unstable, inconsistent behavior
 - ❑ Very limited customization
 - ❑ Limited statistical capability
- **Took the nerd plunge and moved to R**
 - ❑ See above, but invert
 - ❑ Highly extensible environment for data manipulation and statistical programming
 - ❑ Free (no lie)



Name of remote workstation we had
to use to process the cache file in R



+



=




1000 R code lines later, we can attest to the success of this combination.

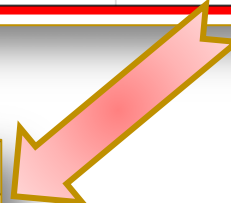
SOLUTION DESIGN PIVOT

- Native structure good for dynamic simulation results capture...
- ...but not for data crunching.
- Transform: Put all results per iteration into separate fields (rather than have each iteration occupy its own field.)

For visual and computational analysis, data is best-ish organized per each iteration.



Task Start Dates (TSD)				Iterations Results -->				
Task ID	Description	Point Estimate	Mean Date	1	2	3...	...	3000
1	Project Parent	Numeric Date A	Numeric Date X	ND[1,1]	ND[1,2]	ND[1,3]	...	ND [1,3000]
2	Fab Unit #1	Numeric Date B	Numeric Date Y	ND[2,1]	ND	ND	...	ND
3	Launch	Numeric Date C	Numeric Date Z	ND	ND	ND	...	ND
...



Task ID	Description	Context	Iteration	TSD
1	Project Parent	Summary	1	Date
1	Project Parent	Summary	2	Date
1	Project Parent	Summary	3	Date
1	Project Parent	Summary
1	Project Parent	Summary	3000	Date
2	Fab Unit #1	Task	1	Date
2	Fab Unit #1	Task	2	Date
...
...	3000	Date
3	Launch	Milestone	1	Date
...
R1	Test Failure	Risk Event	1	Date
R1	Test Failure	Risk Event	2	Date
R1	Test Failure	Risk Event

And then, wide iteration results becomes long.

SOLUTION DESIGN JOINS INTO CSV FILE

CSV that captures risk impacts also created.

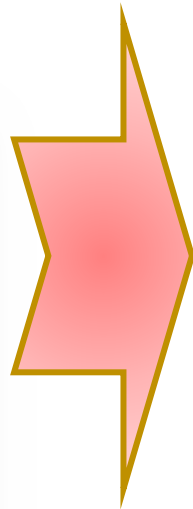
For Gateway, this pivot'n'join process turned ~307K rows into >24 Million.




➤ What you want: All your data in columns next to each other then declare **Big Data victory**

...and then, join all pivoted sections by iteration to arrive at a full results dataset: TASK ID X ITERATION

Task ID	Description	Context	Iteration	TSD
1	Project Parent	Summary	1	Date
1	Project Parent	Summary	2	Date
1	Project Parent	Summary	3	Date
1	Project Parent	Summary
1	Project Parent	Summary	3000	Date
2	Fab Unit #1	Task	1	Date
2	Fab Unit #1	Task	2	Date
...
...	3000	Date
3	Launch	Milestone	1	Date
...
R1	Test Failure	Risk Event	1	Date
R1	Test Failure	Risk Event	2	Date
R1	Test Failure	Risk Event



Primary - Output CSV								
Task ID	Description	Context	Iteration	TSD	TFD	CP	Total Cost	IsActive
1	Project Parent	Summary	1	Date	Date	1	\$	0
1	Project Parent	Summary	2	Date	Date	0	\$	0
1	Project Parent	Summary	3	Date	Date	1	\$	0
1	Project Parent	Summary
1	Project Parent	Summary	3000	Date	Date	0	\$	0
2	Fab Unit #1	Task	1	Date	Date	1	\$	0
2	Fab Unit #1	Task	2	Date	Date	0	\$	0
...
...	3000	Date	Date	0	\$	0
3	Launch	Milestone	1	Date	Date	1	\$	0
...
R1	Test Failure	Risk Event	1	Date	Date	1	\$	0
R1	Test Failure	Risk Event	2	Date	Date	0	\$	1
R1	Test Failure	Risk Event



MODEL DIAGNOSTICS

FIRST THING YOU WANT TO DO IS **NOT**
ANALYSIS

MODEL DIAGNOSTICS

WHAT'S WRONG (RIGHT, PERHAPS?) WITH YOUR BROKEN MODEL?

➤ And you may ask yourself...

- ☐ Do the stochastic critical paths make sense visually and mechanically? Are they wired properly?
- ☐ Has uncertainty been applied properly?
- ☐ Are your risks pinned correctly... and are they misbehaving?
- ☐ What is driving the shape of my s-curve? Why is it funky?

➤ None of these questions are easily answered within in the **SOUL CRUSHINGLY-DATED** MS Project environment or looking at raw sim results in tabular form...

➤ Data paired best with other **schedule assessment** tools (e.g. Fuse) and processes (see detailed guidance in **NASA Schedule Management Handbook**)

Overused
Quote

*All models are wrong
but some are useful*



George E.P. Box

Fred Kuo
Tribute

**Programmatic Assessment
via JCL Analysis via
Nerdy Tech**

MODEL DIAGNOSTICS CORE SCHEDULE BROWSER

Foundationally, cache file enables enhanced schedule surfing.

➤ **Cache file + R + Tableau
(or Power BI)
= Schedule browser**

- ❑ ...view where your risks fall on the critical paths per iteration
- ❑ ...view risks firing over iterations

➤ **Detect the weird iterations that are giving you worries...**

➤ **...or confirm schedule dynamics and flow.**



MODEL DIAGNOSTICS LAYERED SIMULATION BROWSER

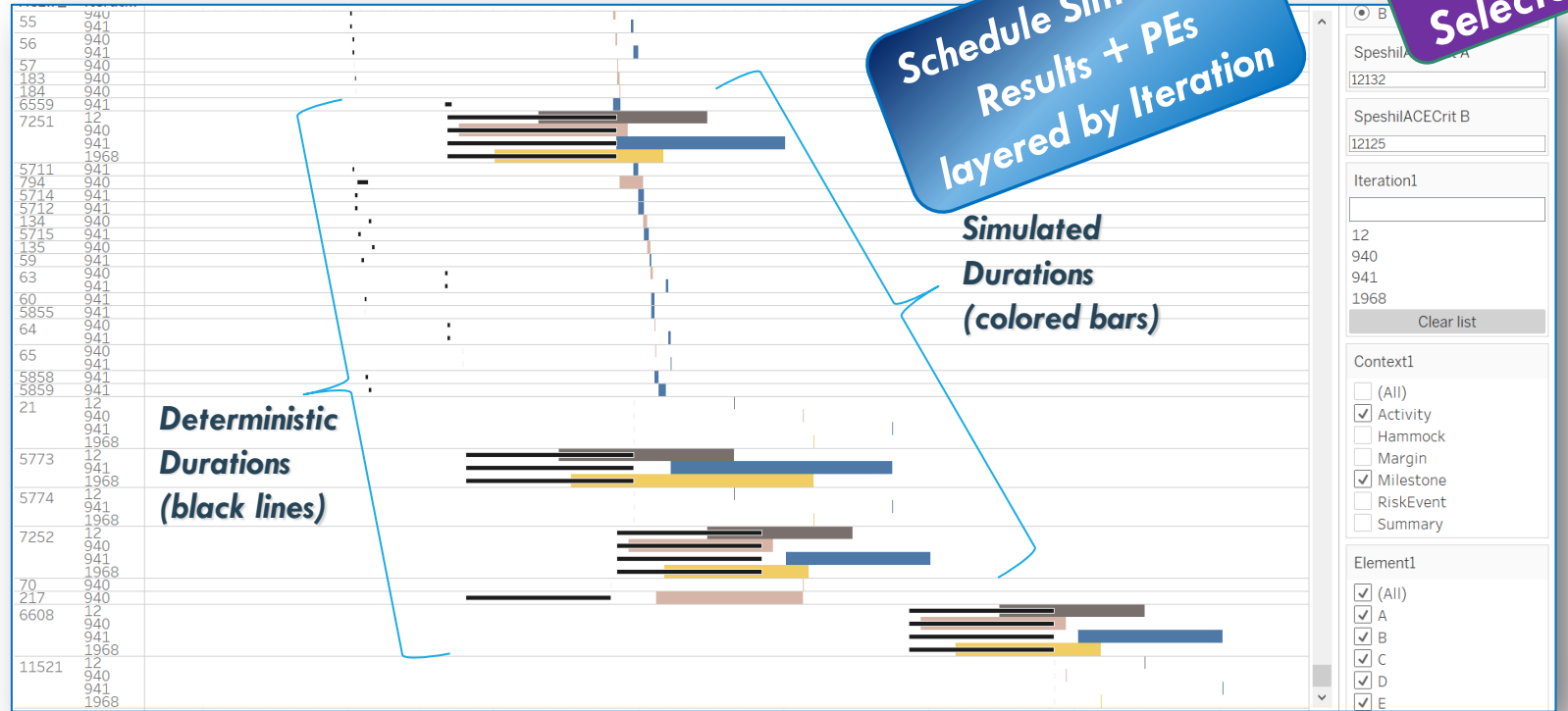
*Schedule Assessment isn't just for
deterministic schedules.*

➤ Simulation results per
iteration w/ PEs

➤ Extremely recently, we found
we just wanted to **isolate**
groups of iteration results to
compare what's going on...

➤ ...because any sufficiently
complex programmatic model
is going to behave strangely at
least occasionally. 🤪

➤ Code in this view can even
isolate which iterations
contain **multiple tasks**
collectively fall on same
critical path



*Simulation results help you understand
schedule artifacts that are forced to
dance (and dance) together.*

MODEL DIAGNOSTICS

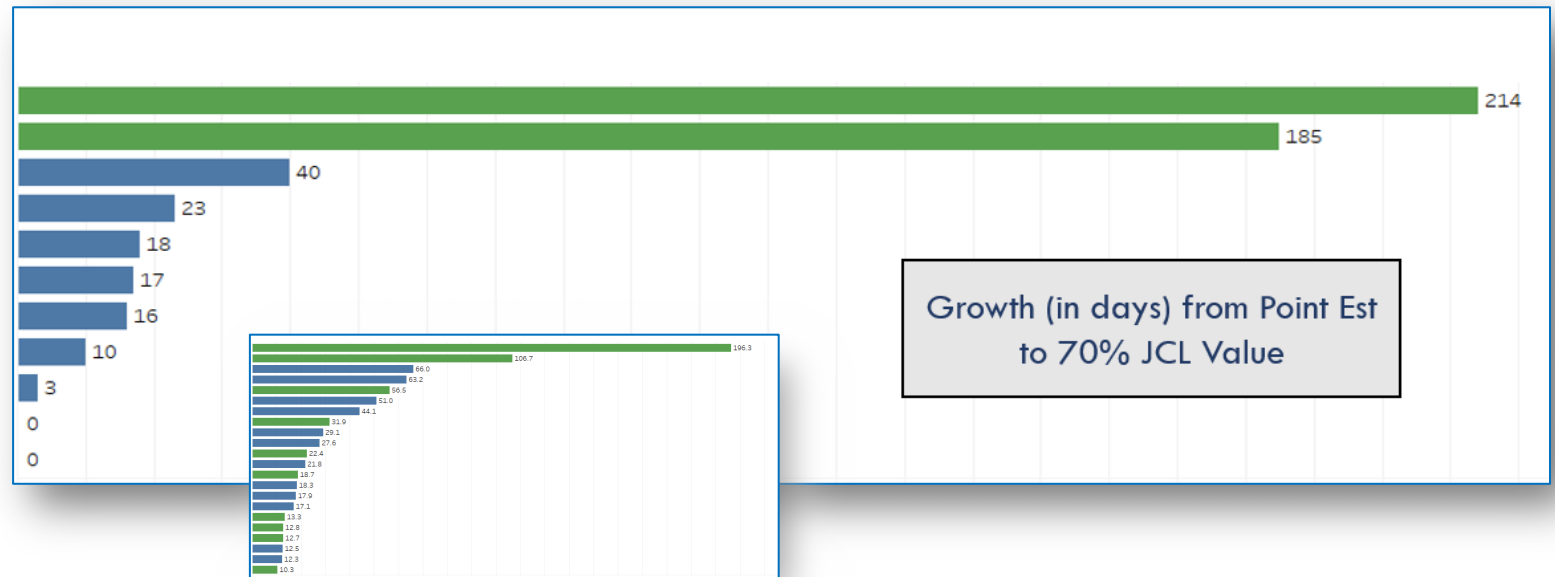
SCHEDULE AND COST GROWTH CHECKS

➤ Transformed data makes it **easy to check** how your cost and schedule elements are **growing**

➤ *Is this what you meant when you dialed in the risks and uncertainty?* (lognormal distribution*)

➤ Calibrated to any measure of growth you want: **Mean**, your favorite **CL**, maximum

➤ Doubles as bite-size stakeholder message




Big huge data enables customizable JCL assessment of input uncertainty and output growth.



AND SO ON

**MORE SCHEDULE ASSESSMENT TECHNIQUES VIA
THIS TECHNOLOGY WILL BE SHARED ON A
CONTINUOUS BASIS WITH THE COMMUNITY AS WE
MOVE FORWARD**



CUSTOM VISUALS & MEASURES

-PREVIEW-

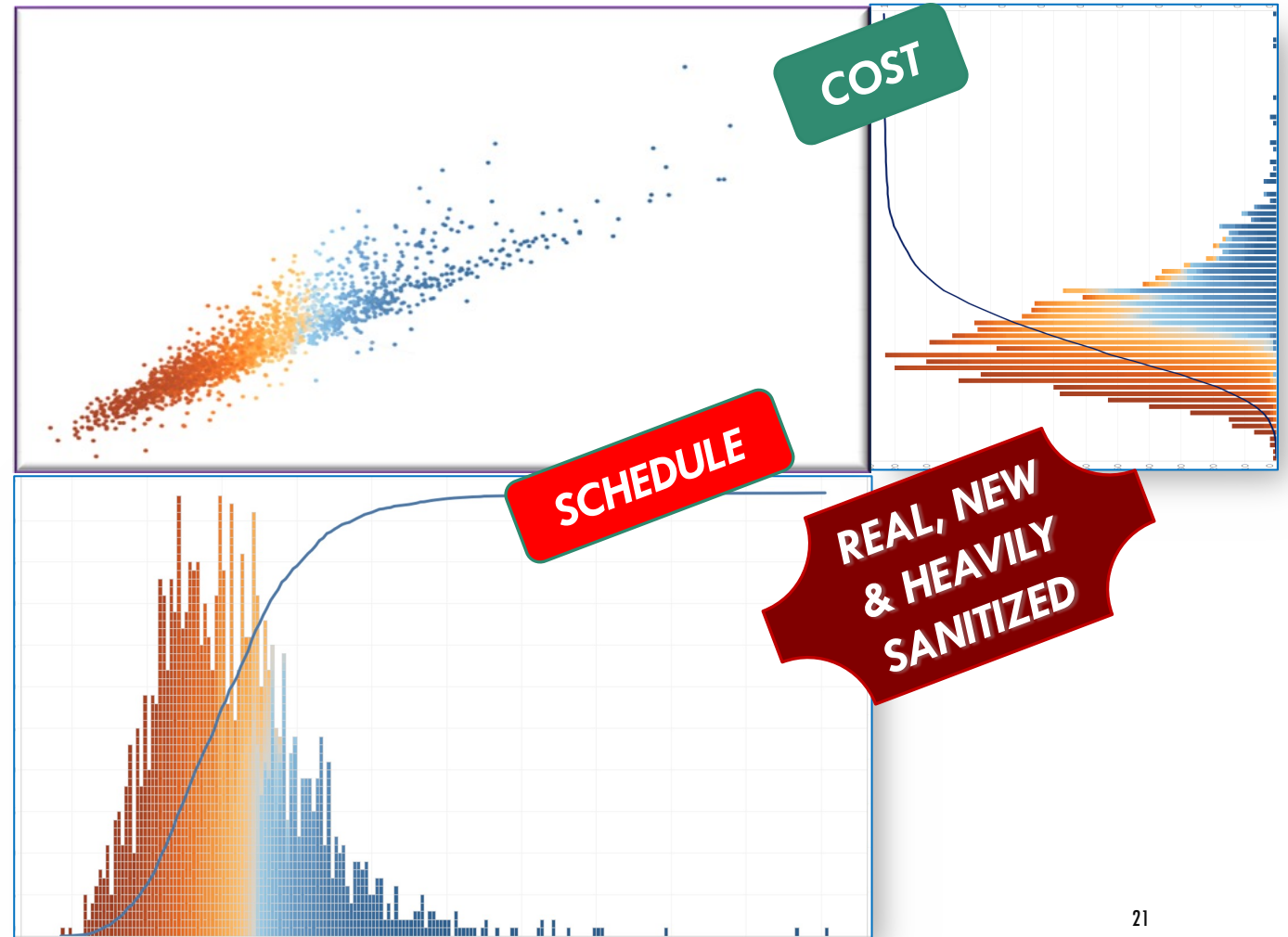
THESE ARE VERY NEW AND RAPIDLY
EVOLVING (BUT IF THEY'RE GOOD
ENOUGH FOR OUR STAKEHOLDERS...)

CUSTOM VISUALS & MEASURES

.....ALL THE TYPICAL VISUALS...

- CSV file provides enough data to recreate all the **traditional JCL charts and metrics**
- Note: you have seen these before a millions, so I will dispense with repetition

JCL Toolbox starts with legacy ideas.



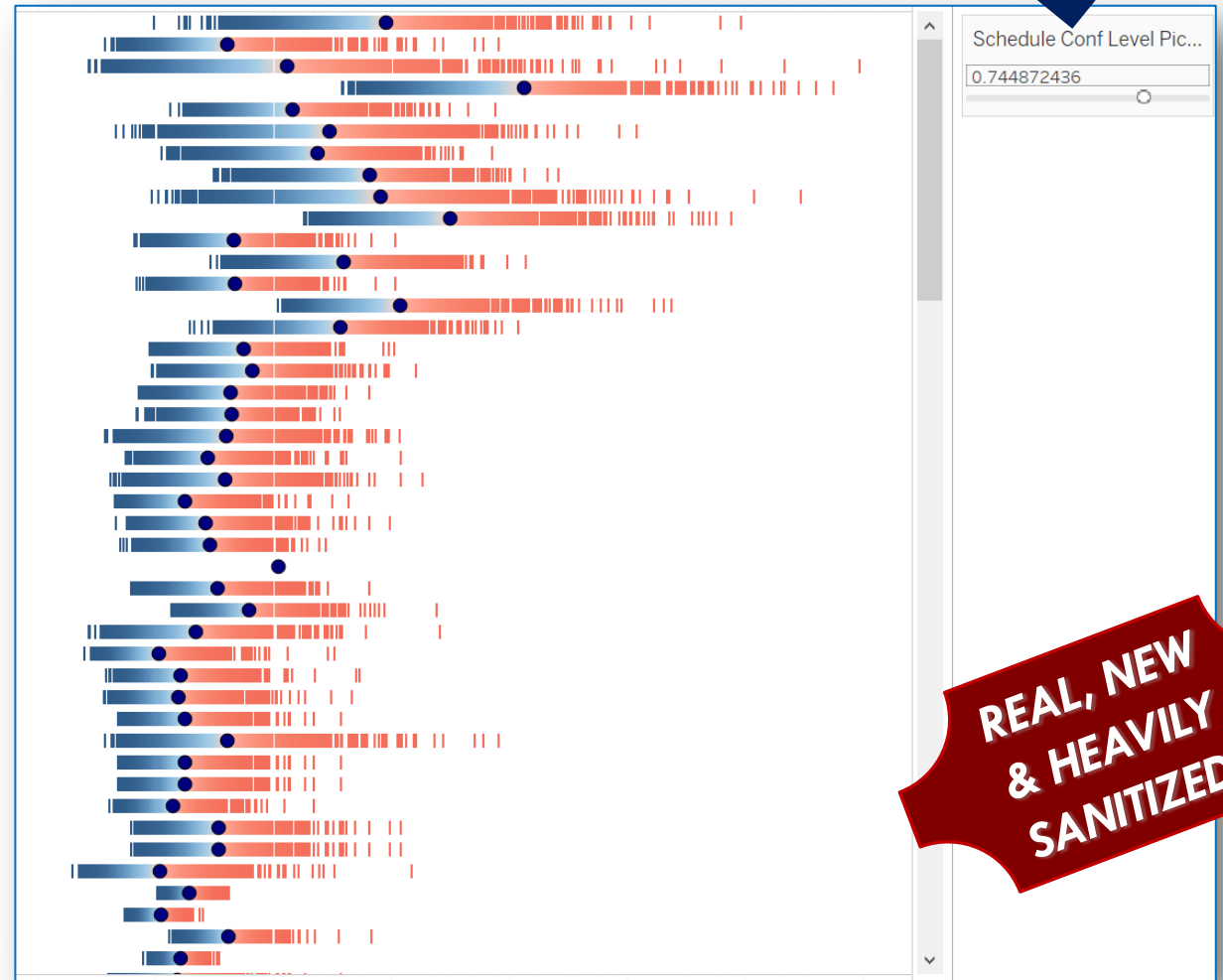
CUSTOM VISUALS & MEASURES

RISK-ADJUSTED SCHEDULE VIEWS

- **Finish date S-Curves** smashed flat and colored for confidence level
- ...+ a **confidence level slider** that moves your point estimate
- Communicates **ranges per schedule item** and potential temporal downsides (or upsides?)

We're looking for as many intuitive ways to map simulation results onto time as possible.

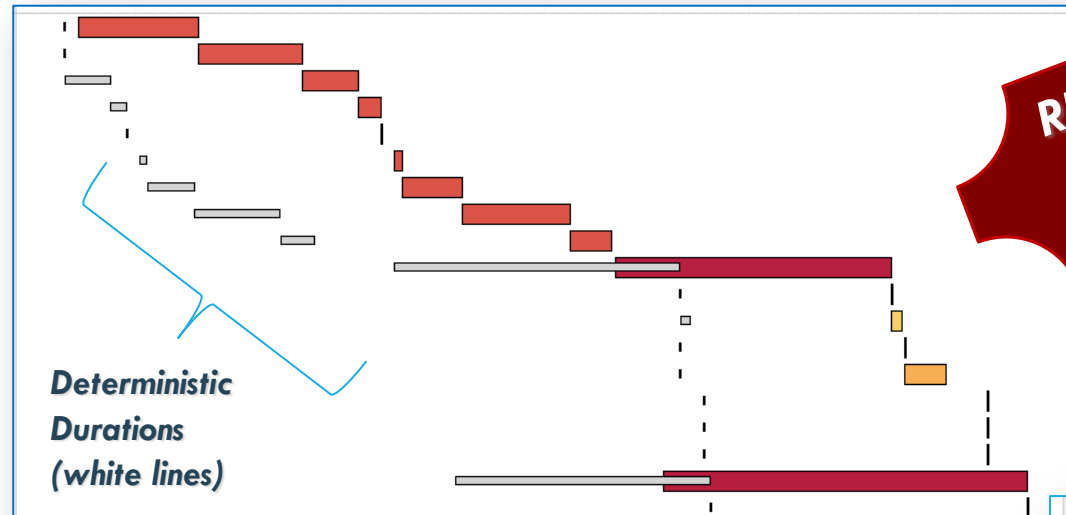
FUN CL Slider for the PE



CUSTOM VISUALS & MEASURES

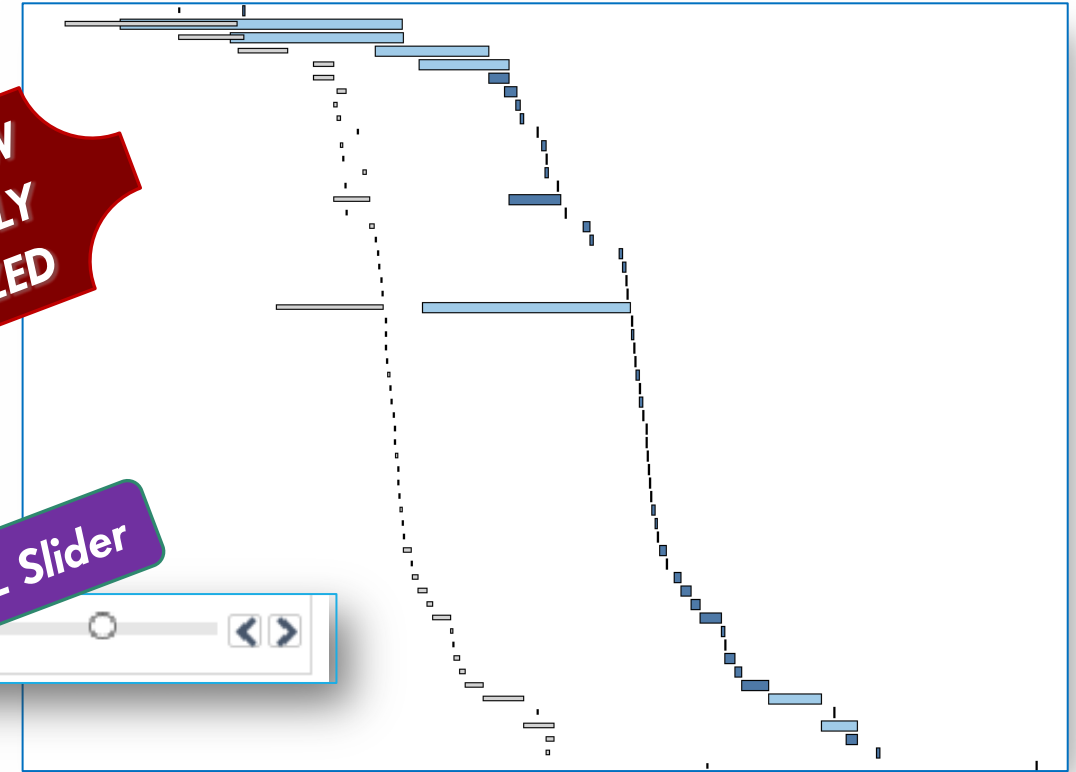
RISK-ADJUSTED SCHEDULE VIEWS

Toolbox allows for browsing of risk-adjusted scenarios of the schedule live.



REAL, NEW
& HEAVILY
SANITIZED

CL Slider



- **70% CL schedule solution w/ deterministic schedule superimposed.**
- **...also color-coded by criticality of the task.**

- **You can also dial in the CL. What if program manager wants to be 'success-oriented' and see the 25% CL schedule?**

CUSTOM VISUALS & MEASURES

RISK: CONDITIONAL 'IN/OUT' IMPACT



Risk CSV enables comparison of risk scenarios.

➤ Dataset allows for **extensive conditional analysis of risks**

➤ **Cost or schedule**

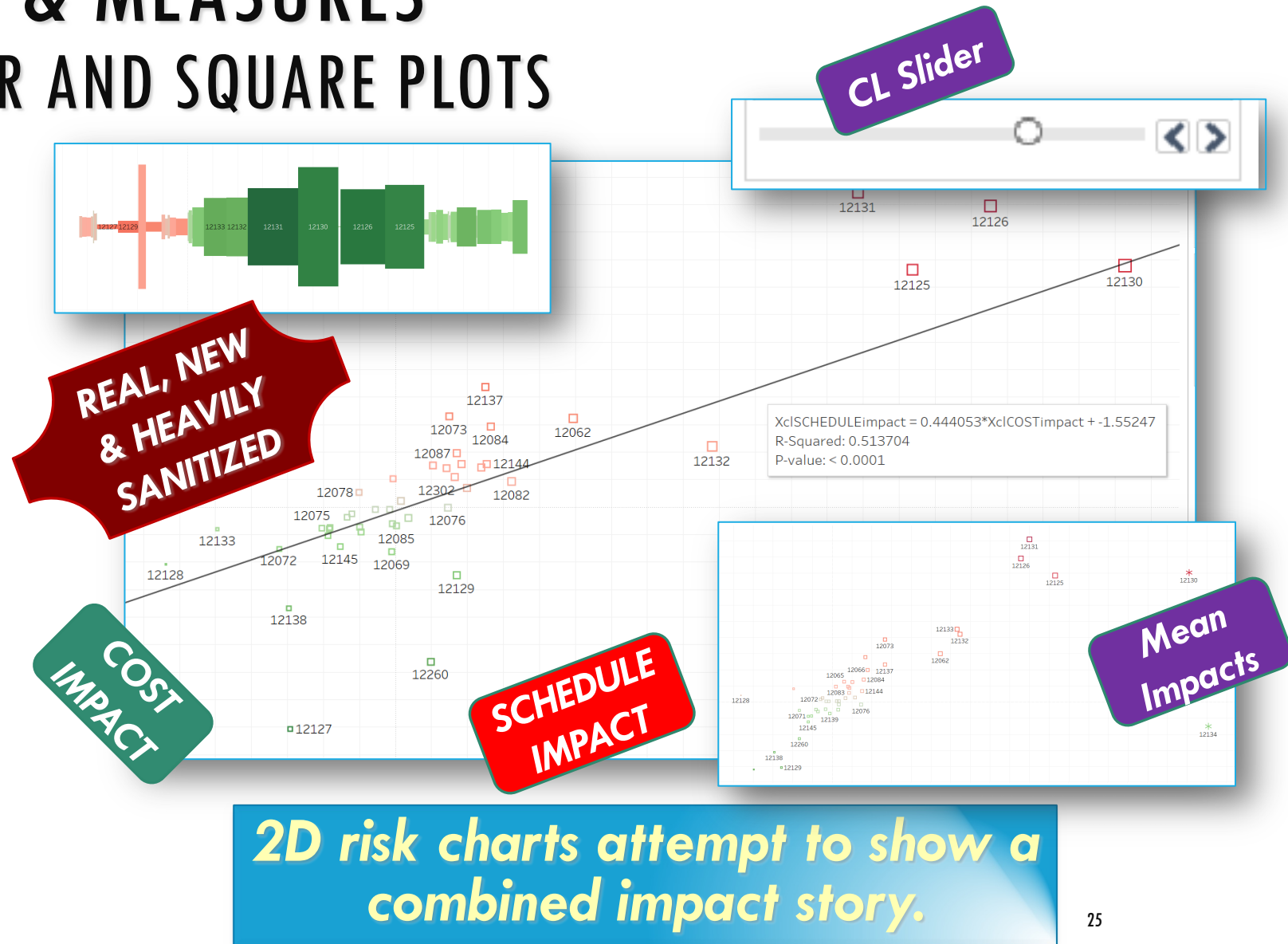
➤ This depiction shows effects of **risk omission vs activation** on the global finish date s-curve

➤ Mirrors **traditionally manual process** (take it from me)

CUSTOM VISUALS & MEASURES

RISK IMPACT: SCATTER AND SQUARE PLOTS

- **Cost and schedule** impacts can be scattered
- ...with their **relationship measured**
- Measured at mean OR **choose-your-own-CL** impact calibration
- Crazy cost-and-schedule, length-and-width '**square plot**' lends an **intuitive, geometric feel** to the size of risk impact.

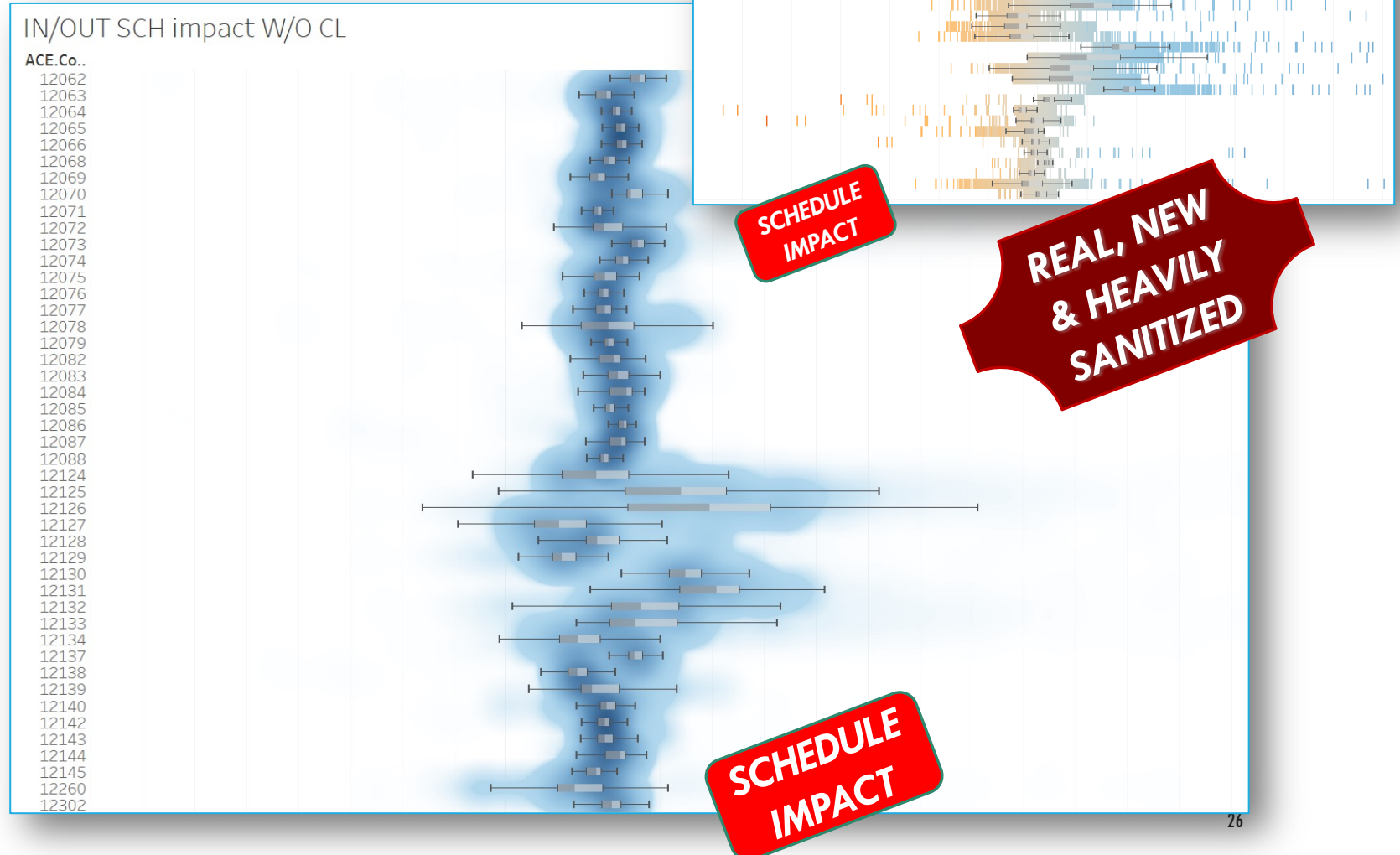


CUSTOM VISUALS & MEASURES

RISK RAINBOWS AND BOX PLOTS

- We had task rainbows, so....
- Smoothed density + box plots show **range of risks side-by-side**
- Can also be **colored vs confidence level** (not depicted)

Just starting to scratch surface on risk impact calcs and visuals.



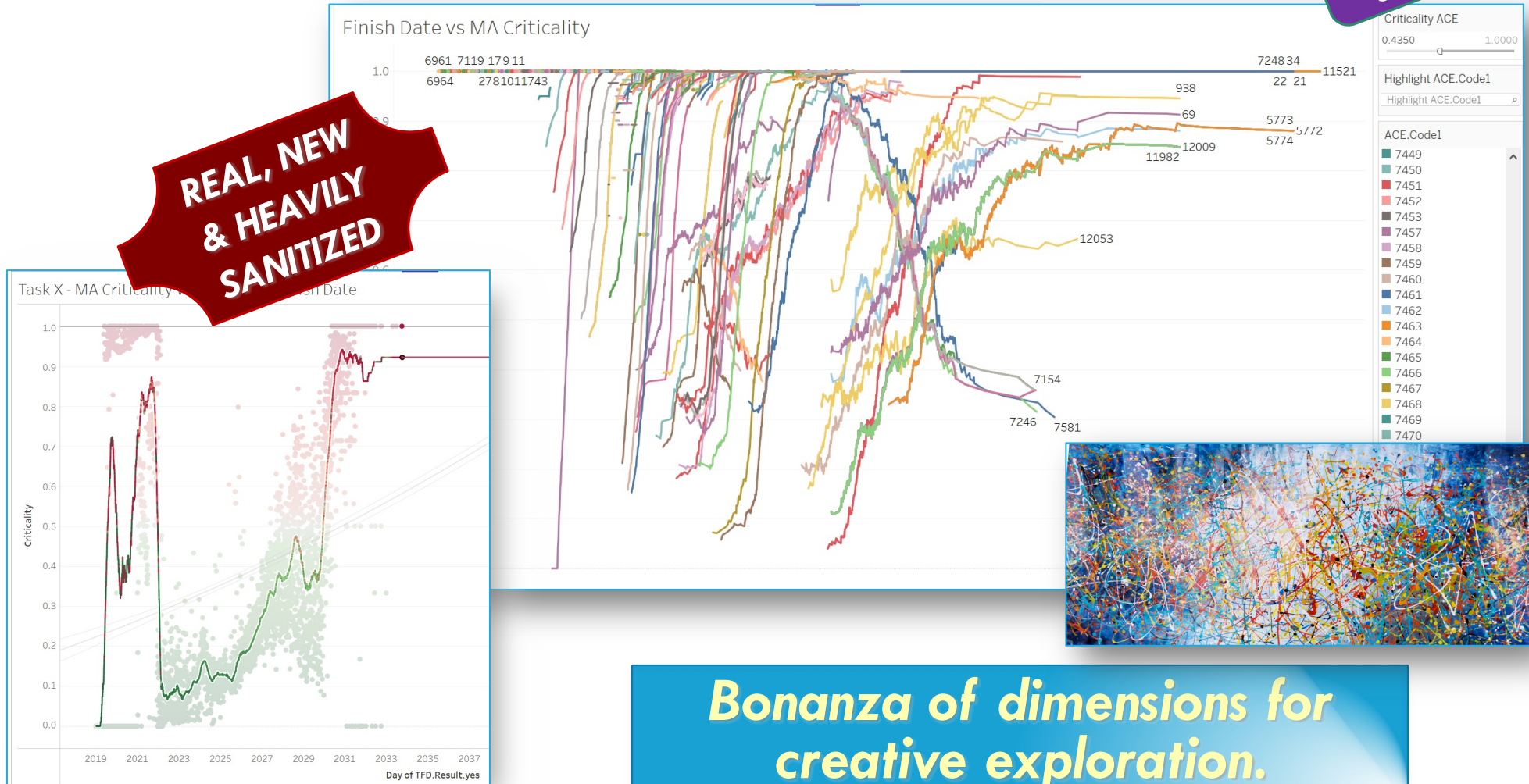
CUSTOM VISUALS & MEASURES AND MORE WEIRD, WILD STUFF WE'RE UNCOVERING...

Criticality
Filter
Slider

➤ Depiction of task **MA** criticality vs task finish date highlight the **nonlinearity of the schedule network** and the interplay amongst its elements.

➤ Early finding: A fixed criticality index **doesn't** tell the whole story

REAL, NEW
& HEAVILY
SANITIZED



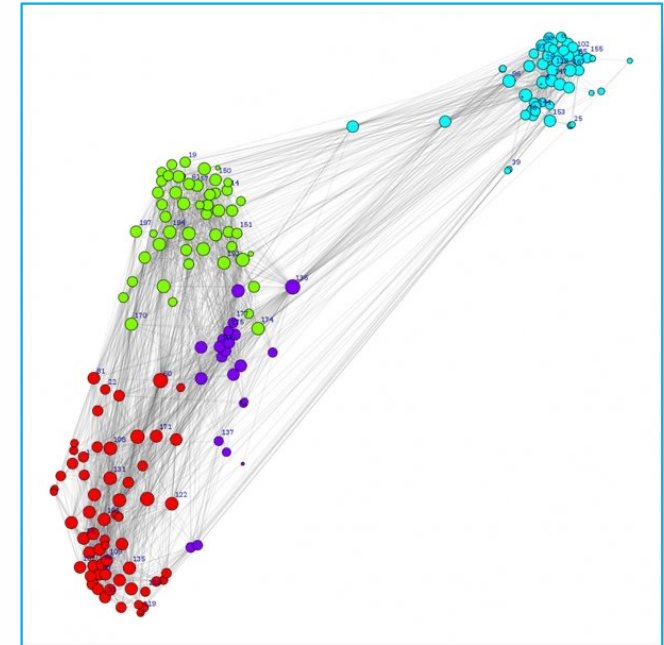
-CRAZY IDEAS-

IF YOU THOUGHT WE WERE SANE YOU'RE AT THE WRONG SYMPOSIUM

CRAZY IDEAS

FORWARD WORK: CRITICAL PATH CLUSTERING

- *How do we consider stochastic critical paths together?*
- Our data ninjitsu can produce a massive **critical path-indicator correlation matrix** amongst schedule tasks can be interpreted as a graph's **adjacency matrix**..
- Spectral graph theory (and other methods) could break down this graph into **critical path clusters**...
- ..which, as schedule chunks, could be used as a **basis for work, funding or risk-mitigation prioritization**



Schedules' stochastic critical paths could be conceptualized as agency nodes in a graph.

CRAZY IDEAS

FORWARD WORK: CONDITIONAL RISK IMPACT ANALYSIS

- Traditional ('Build-up' & 'Top-down') risk impact process, usually performed manually, **requires a TON of iterations to perform conditionally...**
- **Example:**
 - ❑ Given 50% prob risks, **vanishingly few iterations** will represent any active risk case combination as more risks are considered
 - ❑ $2^{30} = 1,073,741,824 > 1 \text{ Billion}$
Active/Inactive cases
 - ❑ A useful S-Curve needs way more than a single schedule impact iteration... like $>500X$. **Do the math. Your computer can't do this simulation.**



How can this explosive risk analysis be addressed by our conditional process address?

CRAZY IDEAS

FORWARD WORK: RISK PARALLELISM MATRIX

- Answer: let's start **measuring the overlap of risks** so that we can move towards avoiding manual processes.... as well as the massively iteration-intensive solutions
- Risks with **high parallelism could be grouped together** (via a **parallelism matrix**) and treated as a collective active/not condition
- Serial/parallel differentiation begin to fall away as progressive, **manual systems being to converge on simple, conditional active/inactive analysis....**
- ...meaning that the IN/OUT algorithm we've created would **render definitive schedule impacts for groups.**



Did you really plan mitigating your risks in isolation anyway? You did not.





-CRAZY IDEAS-
AND MORE TO COME AS
ENGAGEMENT CONTINUES



ONE MORE THING...

JCL TOOLBOX CODE V1.0 RELEASE — COMING SOON

VISUALS AND MEASURES

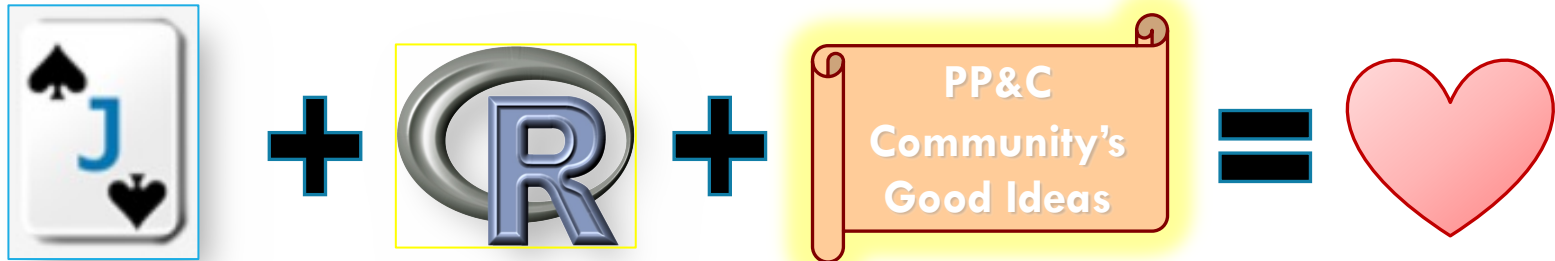
- We plan to release to this community for free use (release plan TBD)

- **Visualization focus**

- ☐ Current: Tableau
- ☐ Future: Adaptation for Power BI
- ☐ Future: R graphics and various pretty platforms

- **Measures focus**

- ☐ New, useful risk and driver measures
- ☐ Advanced metrics to supplement or replace current ones



*These are the continuing voyages
of our JCL **extensibility** journey.*

*Active community involvement will kickstart
exponential growth in sophistication.*

THANK YOU. SEE YOU @ SCOPE AND NCSS
LET'S CHAT. IF YOU'D LIKE TO LEARN MORE OR EXCHANGED IDEAS →

JOHNSON SPACE CENTER

Exploration, Integration & Science Directorate:

-Strategic Business Integration [XB]-

Steve Wilson – steve.m.wilson@nasa.gov

Mike Stelly – john.m.stelly@nasa.gov

BACKUP