



Enterprise Performance Strategies  
z/OS Performance Tuning Essentials

2023 Inperson Conference  
Where Technology and Talent  
Meet Tomorrow...

# Pontification of Data Visualization for z/OS Performance

Peter Enrico  
Email: [Peter.Enrico@EPStrategies.com](mailto:Peter.Enrico@EPStrategies.com)



Jamie Novotny  
Email: [Jamie.Novotny@EPStrategies.com](mailto:Jamie.Novotny@EPStrategies.com)



Enterprise Performance Strategies, Inc.

November 2023

Session QH



[www.epstrategies.com](http://www.epstrategies.com)



z/OS Performance  
Education, Software, and  
Managed Service Providers



Creators of Pivotor®



# Pontification of Data Visualization for z/OS Performance



Peter Enrico & Jamie Novotny  
Enterprise Performance Strategies, Inc.  
[Peter.Enrico@EPStrategies.com](mailto:Peter.Enrico@EPStrategies.com)  
[Jamie.Novotny@EPStrategies.com](mailto:Jamie.Novotny@EPStrategies.com)



# Contact, Copyright, and Trademarks



## Questions?

Send email to [performance.questions@EPStrategies.com](mailto:performance.questions@EPStrategies.com), or visit our website at <https://www.epstrategies.com> or <http://www.pivotor.com>.

## Copyright Notice:

© Enterprise Performance Strategies, Inc. All rights reserved. No part of this material may be reproduced, distributed, stored in a retrieval system, transmitted, displayed, published or broadcast in any form or by any means, electronic, mechanical, photocopy, recording, or otherwise, without the prior written permission of Enterprise Performance Strategies. To obtain written permission please contact Enterprise Performance Strategies, Inc. Contact information can be obtained by visiting <http://www.epstrategies.com>.

## Trademarks:

Enterprise Performance Strategies, Inc. presentation materials contain trademarks and registered trademarks of several companies.

The following are trademarks of Enterprise Performance Strategies, Inc.: **Health Check®**, **Reductions®**, **Pivotor®**

The following are trademarks of the International Business Machines Corporation in the United States and/or other countries: IBM®, z/OS®, zSeries®, WebSphere®, CICS®, DB2®, S390®, WebSphere Application Server®, and many others.

Other trademarks and registered trademarks may exist in this presentation

# Abstract



The authors of the presentation have years of experience helping people understand the performance of their z/OS systems, in part by visualizing the performance details that are stored in the SMF data. You probably won't be surprised to find that we're quite interested in data visualization as a discipline.

In this webinar, Peter Enrico and Jamie Novotny will explore the psychology behind data visualization, discuss some visualization best practices, ridicule examples of poor visualizations, and explain how performance reports work to avoid obfuscation and promote understanding.

# EPS: We do z/OS performance...



- We are z/OS performance!
- Pivotor
  - Performance reporting and analysis of your z/OS measurements
  - Example: SMF, DCOLLECT, other, etc.
  - Not just reporting, but cost-effective analysis-based reporting based on our expertise
- Performance Educational Workshops (while analyzing your own data)
  - Essential z/OS Performance Tuning
  - Parallel Sysplex and z/OS Performance Tuning
  - WLM Performance and Re-evaluating Goals
- Performance War Rooms
  - Concentrated, highly productive group discussions and analysis
- MSU reductions
  - Application and MSU reduction

# z/OS Performance workshops available



During these workshops you will be analyzing your own data!

- Essential z/OS Performance Tuning
- Parallel Sysplex and z/OS Performance Tuning
- WLM Performance and Re-evaluating Goals
- Also... please make sure you are signed up for our free monthly z/OS educational webinars! (email [contact@epstrategies.com](mailto:contact@epstrategies.com))

# Like what you see?



- Free z/OS Performance Educational webinars!
  - The titles for our Fall 2022-2023 webinars are as follows:
    - ✓ *LPAR Configurations to Avoid*
    - ✓ *How Different are High, Medium, and Low Pool Processors?*
    - *CPU Critical: A Modern Revisit of a Classic WLM Option*
    - *Mainframe Efficiency at High Utilizations (Bob Rogers)*
    - *I/O, I/O It's Home to Memory We (Should) Go*
    - *30th Anniversary of WLM : A Retrospective and Lessons Learned*
    - *Understanding and Measuring Warning Track on z/OS*
    - *30th Anniversary of Parallel Sysplex - A Retrospective and Lessons Learned*
    - *Batch Initiators – WLM Managed or JES Managed?*
    - *AI on Z: Exploring Common AI Terms on System Z*
    - *Analyzing 'Per CPU' Utilizations*
    - *AI on Z: Exploring new SMF Measurements*
  - Let me know if you want to be on our mailing list for these webinars
- If you want a free cursory review of your environment, let us know!
  - We're always happy to process a day's worth of data and show you the results
  - See also: <http://pivotor.com/cursoryReview.html>

# Who are we and why do we care?

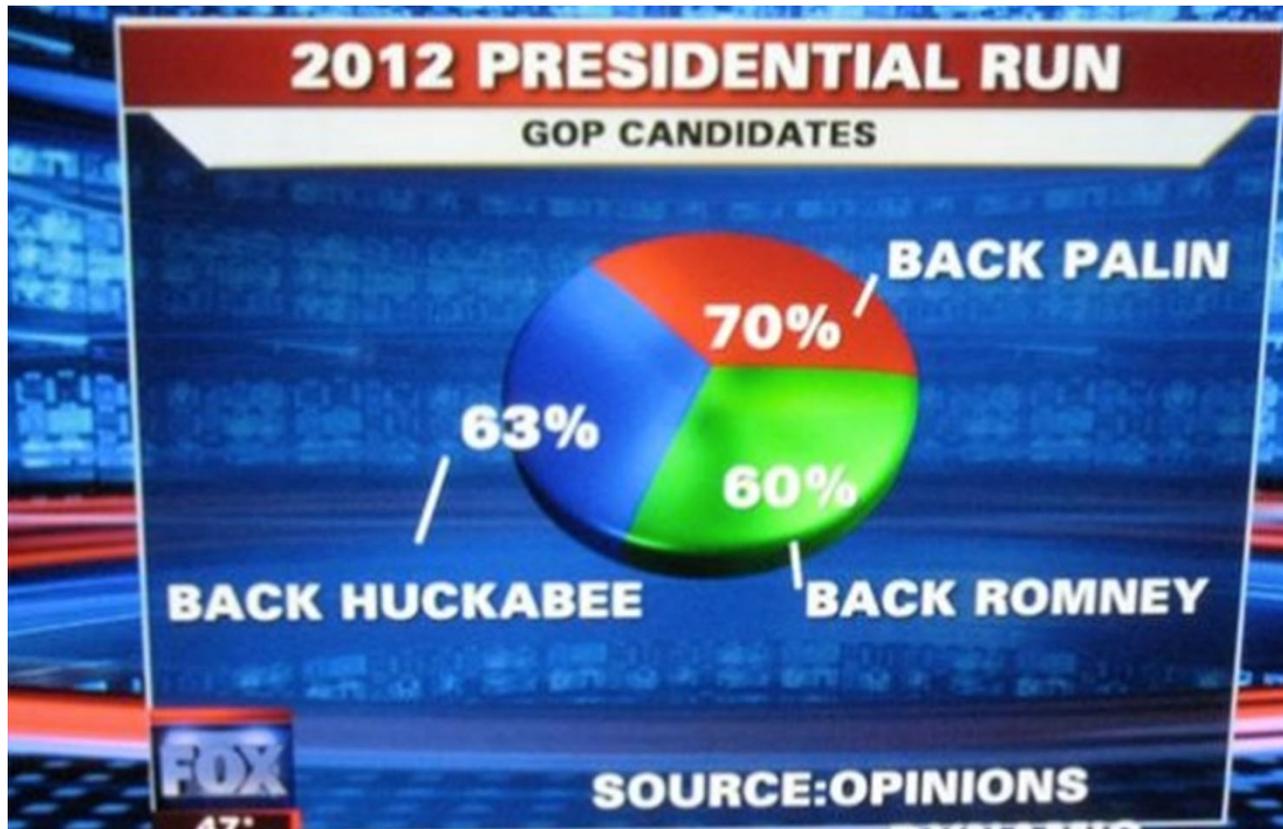


- Jamie Novotny
  - Interested in all things human behavior and brain development
- Peter Enrico
  - Mainframe performance expert



# Bad Visualizations Can Be Painful

# Good Visualizations Are Important

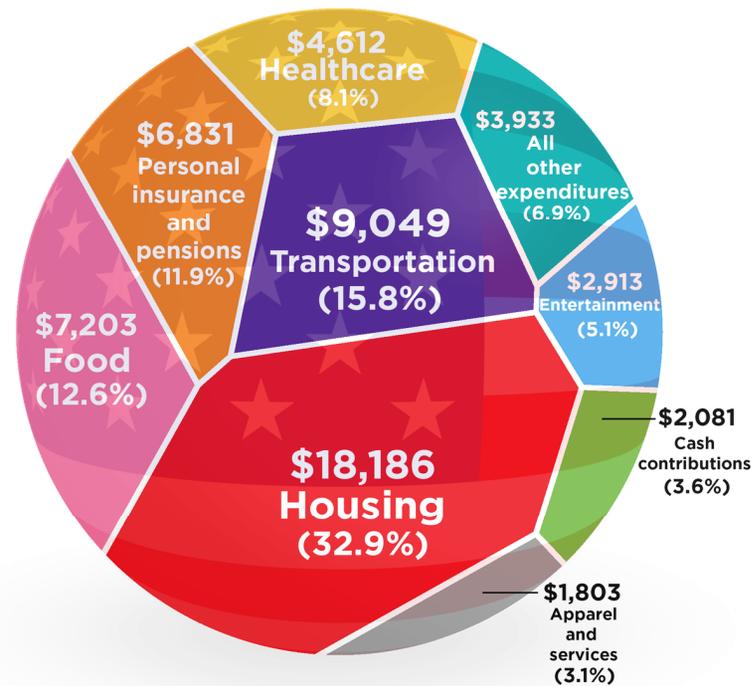


A pie chart adding up to 193%?!?

# How to make an even worse pie chart



## How Americans Spend Their Money

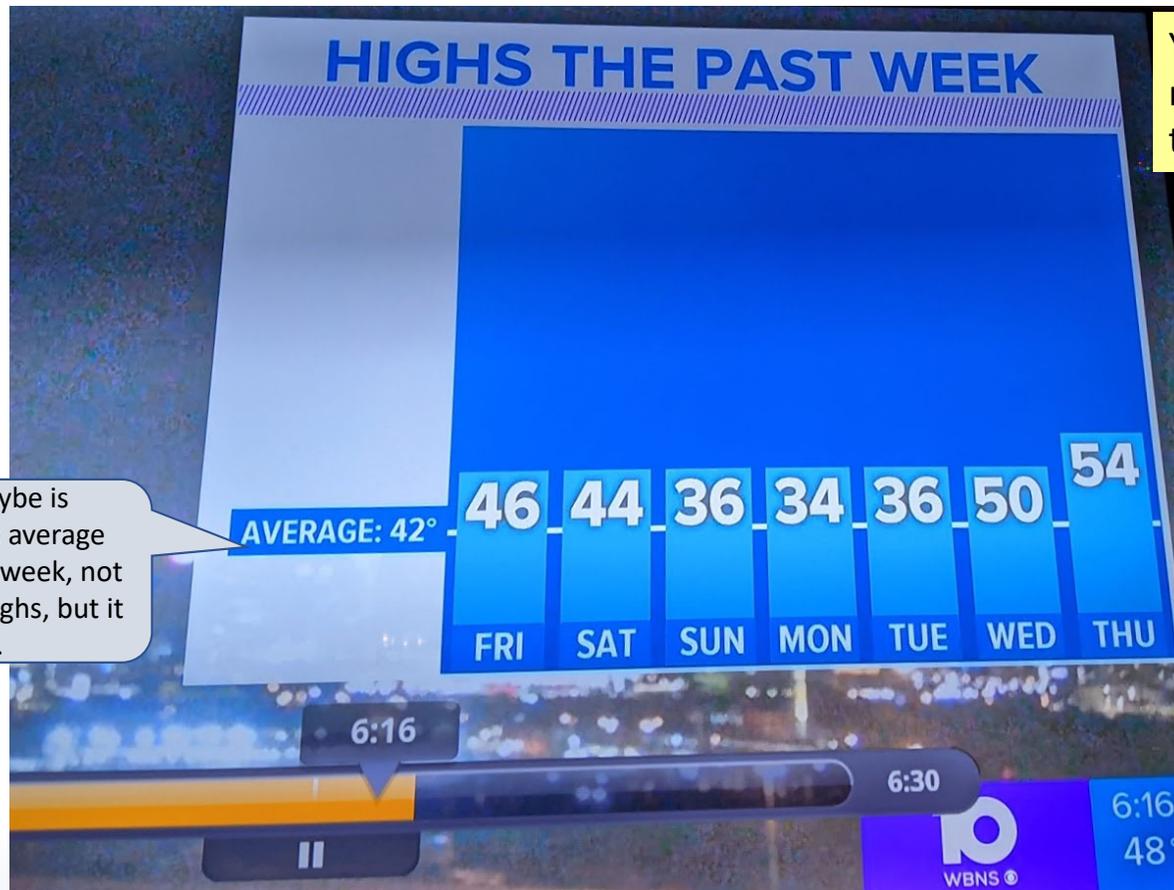


Average Income before taxes = \$74,664  
Article and sources:  
<https://howmuch.net/articles/how-americans-spend-their-money-2016>  
<https://www.bls.gov/news.release/pdf/cesan.pdf>

Average Annual Expenditures = \$57,311

**howmuch**.net

# Accuracy matters



Yes, I did pause the news just to capture this!

This average maybe is supposed to be the average temperature for the week, not the average of the highs, but it is not clear.

# TV stations must have terrible software



■ INFORMACIÓN CORONAVIRUS MÁS ALTA DE TODA LA PANDEMIA. LA OCUPACIÓN

■ DIRECTO

■ Redacción 'laSexta Noticias'

■ EVOLUCIÓN DEL CORONAVIRUS EN ESPAÑA

SANIDAD NOTIFICA 42.885 CASOS EN 24 HORAS, LA SEGUNDA CIFRA MÁS ALTA DE LA PANDEMIA

Este gráfico muestra el número de muertes notificadas por día en España durante la primera semana de la pandemia de COVID-19. El eje vertical representa el número de muertes, con marcas a los 100, 200, 300, 400, 500 y 600. El eje horizontal muestra los días: Lunes 18/01, Martes 19/01, Miércoles 20/01 (Acumulado fin de semana), Jueves 21/01 y Viernes 22/01. Los valores de las barras son: Lunes 18/01: 455; Martes 19/01: 404; Miércoles 20/01 (Acumulado fin de semana): 464; Jueves 21/01: 404; Viernes 22/01: 400. Fuente: Ministerio de Sanidad.

Día	Fecha	Muertes Notificadas
Lunes	18/01	455
Martes	19/01	404
Miércoles	20/01 (Acumulado fin de semana)	464
Jueves	21/01	404
Viernes	22/01	400

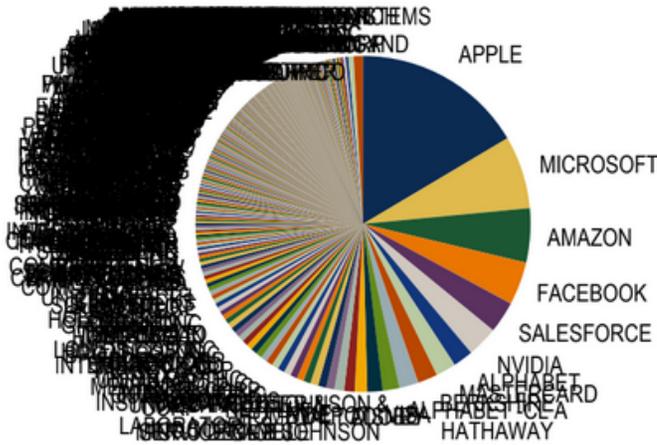
Fuente: Ministerio de Sanidad

This chart isn't helping!

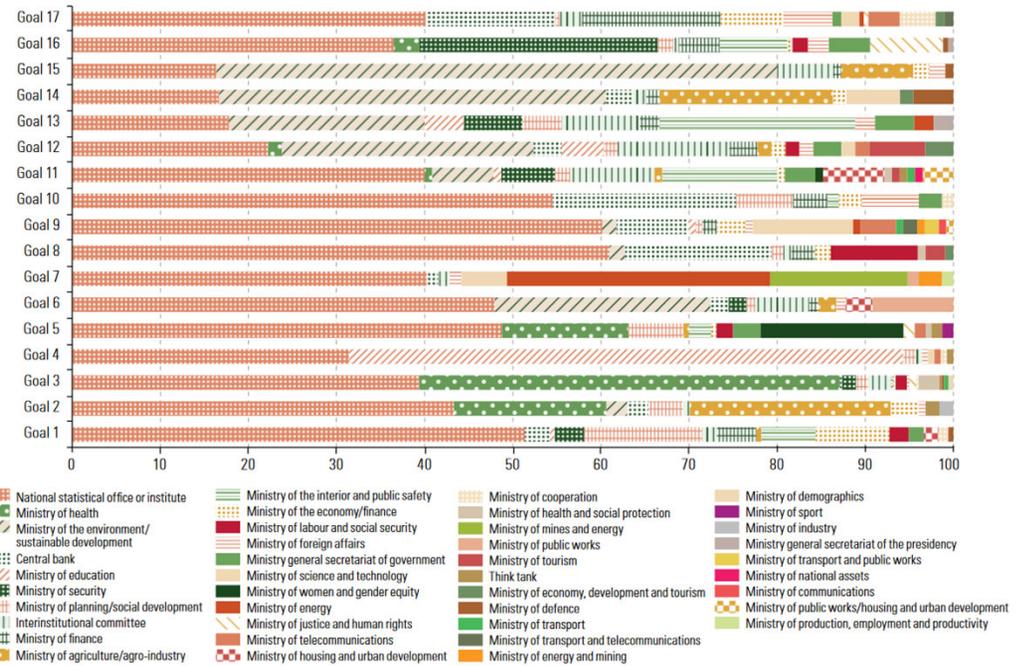
# More series is not more better!



Chart 3: 10 stocks in S&P500 accounted for >50% of August 7.2% return



Source: BofA Global Investment Strategy, Bloomberg

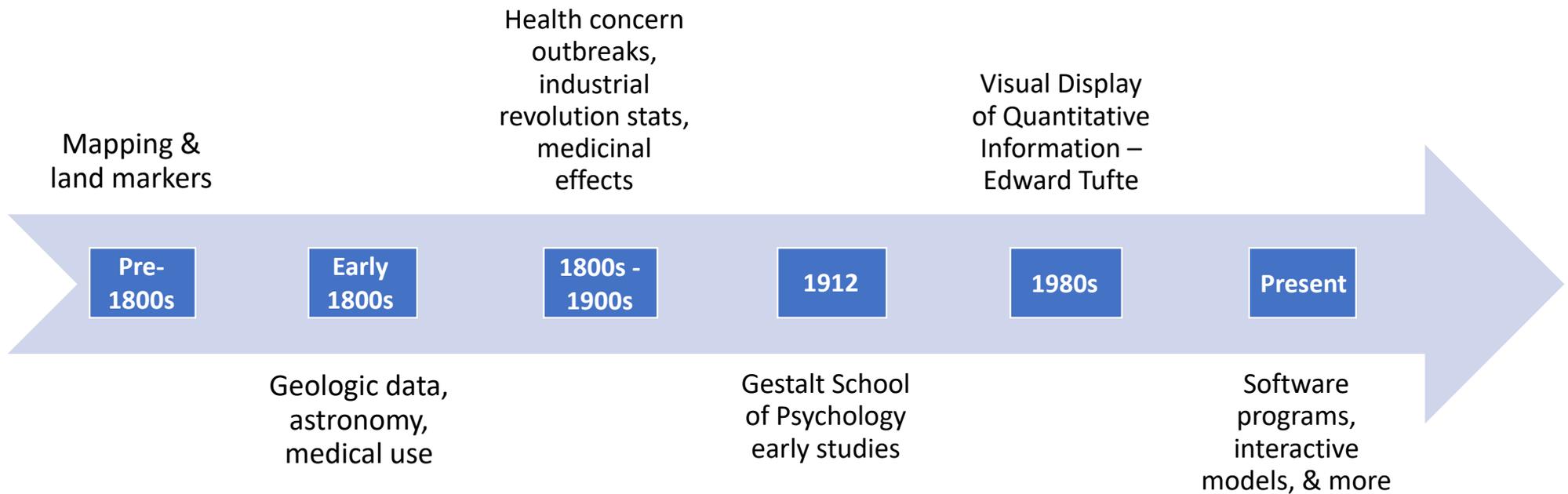


For more such fun, see: <https://www.reddit.com/r/dataisugly/>



# The History and Psychology of Visualizations

# History of Data Visualization & Perception



# The Importance of Representation



- Comprehensive
- Accurate
- Fitting
- Represents the situation
- No room for misinterpretation



- Difficult to understand/confusing
- Inaccurate
- Missing pieces
- Misrepresents the situation
- Misinterpretation → incorrect assumptions

# Gestalt Principles of Perception

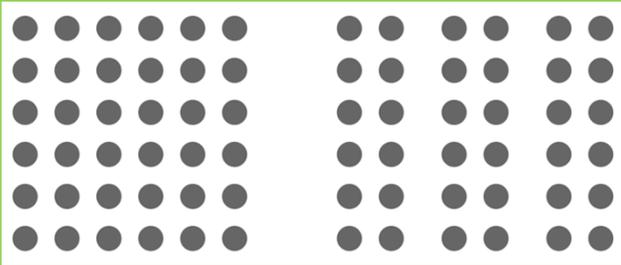


We know...

- The “human brain is hard-wired to identify patterns as a way to make sense of our world.” Psychologists Max Wertheimer, Wolfgang Köhler, and Kurt Koffka
- We create structure by default → “The whole is greater than the sum of its parts”
- Gestalt Principles ~6 laws/principles pinpointing the most telling patterns our brains find useful

### Proximity

Grouping is evident when items are close (or not) together, hinting a relationship



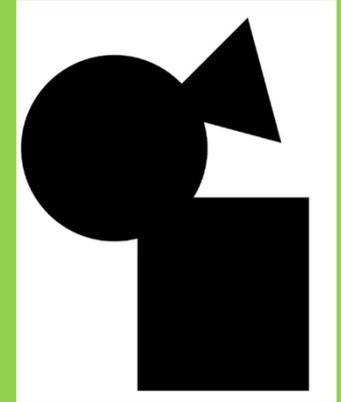
### Similarity

When similar characteristics or attributes are used (like color, shape, etc.), we group them together (or dissimilarity if we want something to stand out)



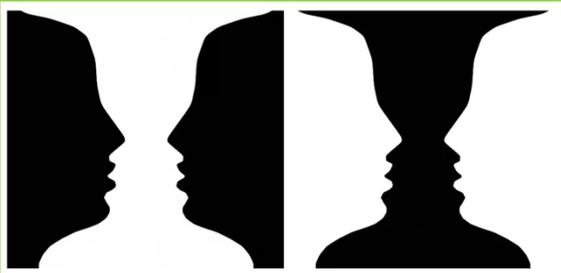
### Symmetry/Order

Ambiguous shapes will be perceived in a simpler way



### Figure/Ground

Use of contrast, colors, and size to create visual elements between background/foreground



### Continuity

When elements are in a line/curve, they are likely perceived as related, and our eyes will follow the smoothest path



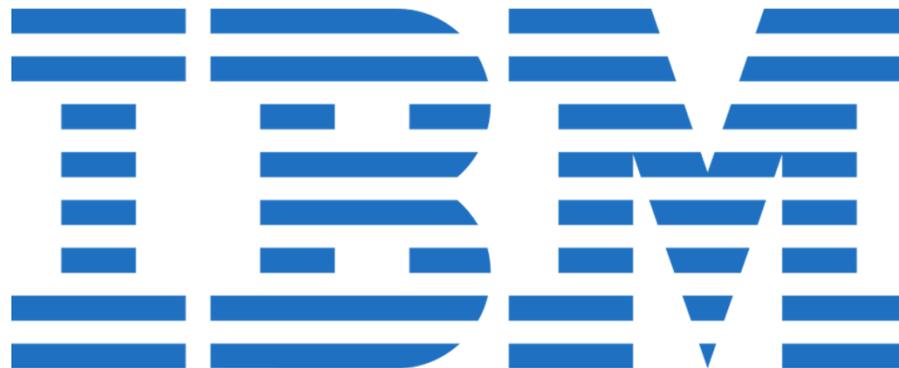
### Closure

A complex arrangement of elements are perceived by the brain as one recognizable object, we will fill in the missing pieces





FedEx



IBM





# Modern Visualization Recommendations

# First, Determine your message



- The purpose of data visualization is to tell a story or convey a message
  - Develop a clear vision of that story and/or message
- Then determine the best visualization for that story. For example:
  - Are you trying to show change over time?
    - Example: A before change / after change comparison
    - Example: A gradual change that requires showing a change over time
  - Are you trying to show a relationship between multiple variables or values
    - Example: The relationships between changing value
  - Are you trying to show comparisons?
    - Example: The ratio between two values
    - Example: The comparison to rules of thumb or exception thresholds
- What conclusion or answer do you want your audience to come to
  - If you find you have too much data, think about what really needs to be shown to support the intended conclusion
  - Consider highlighting data that supports the conclusion

# Determine needed dimensional data



- Dimensional Data - Data that gives metrics context
  - One should think of dimensional data as any data that could be used compare metrics
- Examples of Dimensional data
  - **Time** based dimensional data – When data is used in context of ‘**When?**’
    - Examples: Date, Time, Hour, Year, etc.
  - **Placement** based dimensional data – When data is used in context of ‘**Where?**’
    - Examples: Sysplex name, System name, Volser, Coupling Facility, etc.
  - **Subject** based dimensional data – When data is used in context of ‘**What?**’
    - Examples: Transaction name, Service class name, Volser, Job name, etc.
  - **Adjective** based dimensional data – When data is used in context of ‘**Which?**’
    - Examples: Type of processor (CP, zIIP, zAAP, IFL, ICF), type of WLM class (Report Class or Service Class), etc.
- How the dimensional data is used is dependent upon the question being asked

# Determine your measured metric



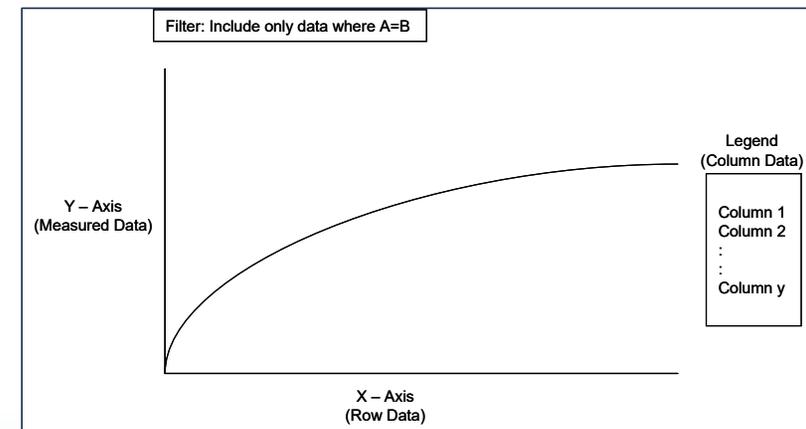
- Metrics Data

- What is typically thought of as the measurement to gauge the relative value of during an analysis or study
- One should think of metrics as any measured value that could be used to help gain an understanding needed to answer a question
- Many times, when reviewing measurements, metrics can be further defined
  - Cumulative / Interval
  - Snap-shot
  - Sampled
  - Etc.
- Examples of metrics (in a computer measurement environment)
  - Counts
  - Averages
  - Rates
  - Utilizations
  - Percentages
  - Ratios
  - Intensities
  - Etc.

# Determine the placement and grouping of data



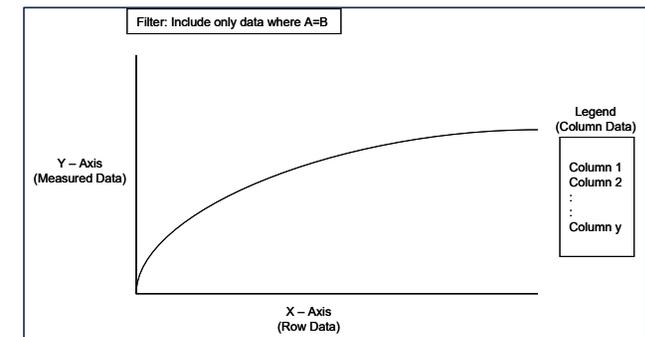
- Metric data
  - Data of study for the analyst of the report
  - Data values in this context could be the result of a collected value, complex formula, or could be nothing more than a Boolean indicator
- Row data
  - Group is data that provides context to measured data by separating summarized measured data into distinct groupings organized in row
- Column data
  - Grouping allows measured data to be logically grouped and summarized within a particular row
- Filter data
  - Grouping allows the data in any report to be narrowed down to a subset of data for reporting purposes



# Determine the report type



- Data can be reported in an infinite number of ways, but in general, there are 3 broad categories of reports
  - This is where the psychology of data visualization really starts
- Graphical Reports
  - Allow the visualization of data into charts
- Tabular Reports
  - Allow the reading of the data values in a grid format
- Customer Designed Reports
  - Allow the summarization of varying and different types of data in a custom designed report
  - Example: RMF or CMF Post Processor reports



	Column 1	Column 2	...	Column y
Row 1	Measurement 1,1	Measurement 1,2	...	Measurement 1,y
Row 2	Measurement 2,1	Measurement 2,2	...	Measurement 2,y
Row 3	Measurement 3,1	Measurement 3,2	...	Measurement 3,y
:	:	:	:	:
:	:	:	:	:
Row x	Measurement x,1	Measurement x,2	...	Measurement x,y

# Variables of Data Visualization



- The classic variables include:
  - Shape and type of chart
  - Position and size of chart and data
  - Color, hue, lightness of represented data
  - Orientation and texture of data
  - Creative data representation
  
- For z/OS performance, no need to get to fancy
  - Since most z/OS performance analyst are technical, it is mostly about quick analysis
  - Most common data visualization variables include:
    - Chart type
    - Color and hue
    - Position, size

# Picking a chart type: Line Chart



- Classic Line chart

- Use: Most typical when there is a need to examine values over time
  - How the value of something changes over time, and for value comparisons
  - Great for looking at counts, averages, percentage, utilizations, rates, intensities and ratios
- Pro: Widely used and easily understood
- Con: May be hard to find individual data values on the line
  - Consider adding data markers (carefully, can lead to cluttered chart)
  - Sometimes lines are cluttered and there is a need to remove lines for better messaging
  - When there is wide variability between data points can lead to difficult to read chart
    - Consider using data markers only—no line

# Line Chart Examples and Comparisons

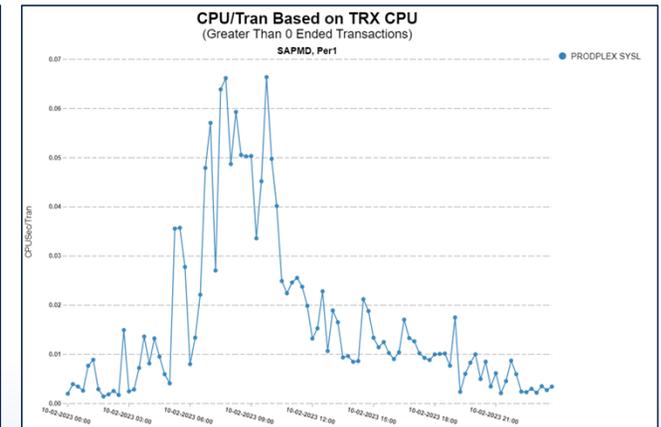
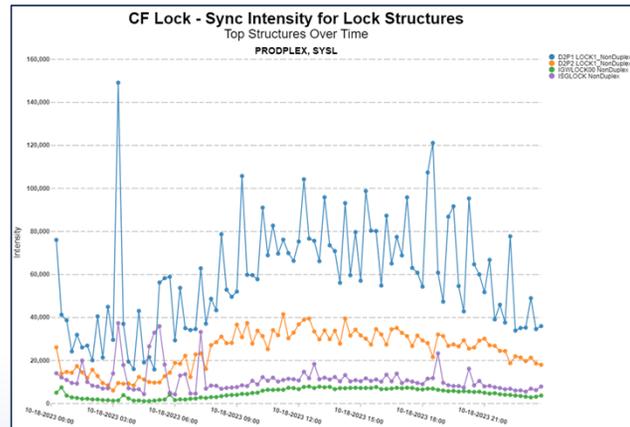
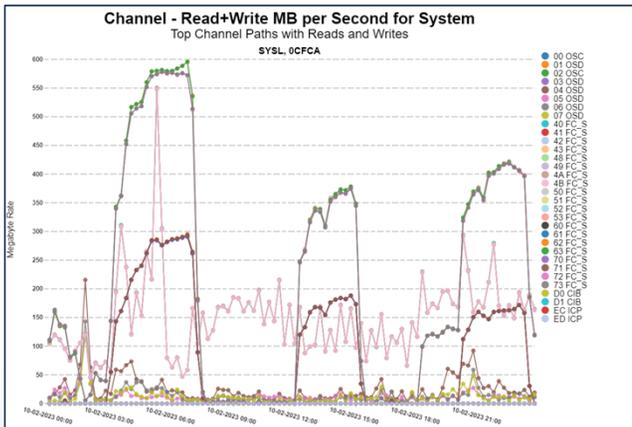
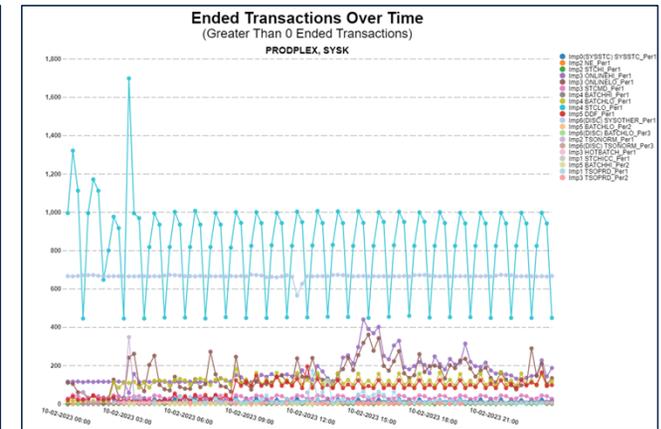
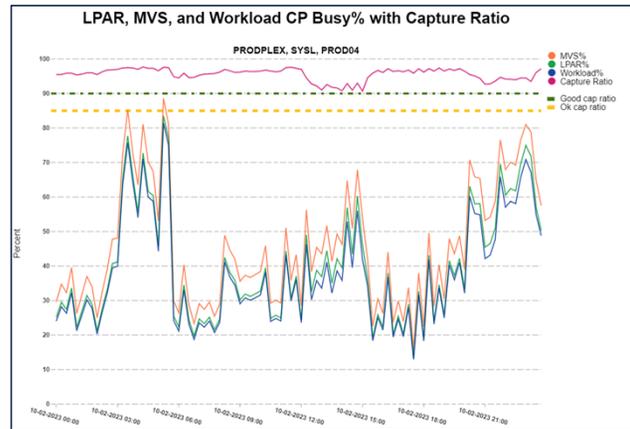


Line charts are great when comparing over time:

Examples where include:

- Compare various utilizations
- Compare number of ended transactions
- Compare rates of reads+writes per second
- Compare intensities
- Compare ratios

These same values would make little sense to visualize as area, column, pie, or most other chart types



# Picking a chart type: Area and Column



## ● Area chart

- Use: When looking at the collective 'volume' or 'usage' of something
- Pro: Visually great for looking at the total and accumulation of various values
  - Shape of chart and individual series shows the breakdown of the collective total
- Con: Area charts tend to 'smooth' the visual out over time.
  - Not a highly recommended chart type

## ● Column chart

- Use, Pros, and Cons are the same as area charts
- Additional Pro above area charts
  - Also allows you to more easily recognize the interval that is being reported

# Line Chart Examples and Comparisons

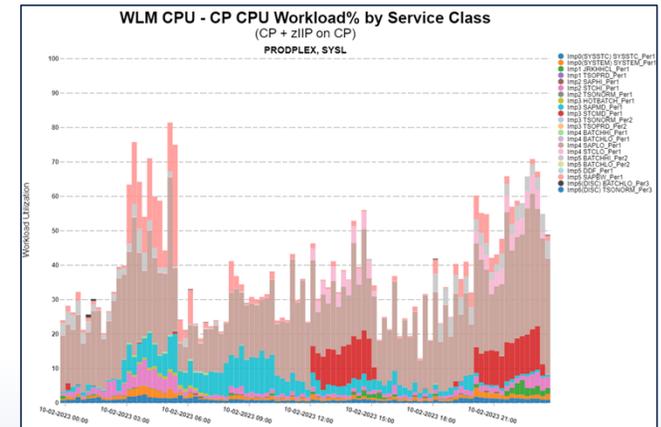
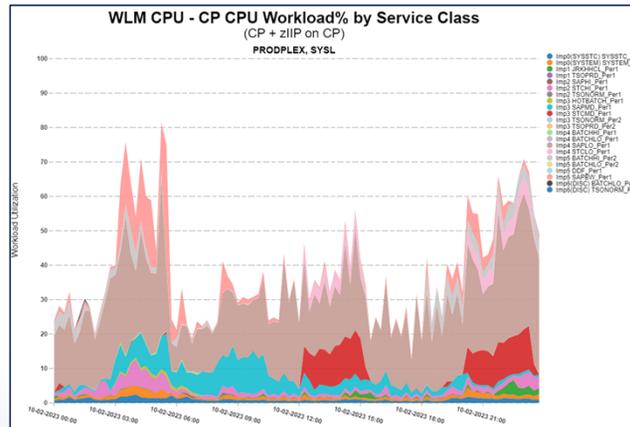
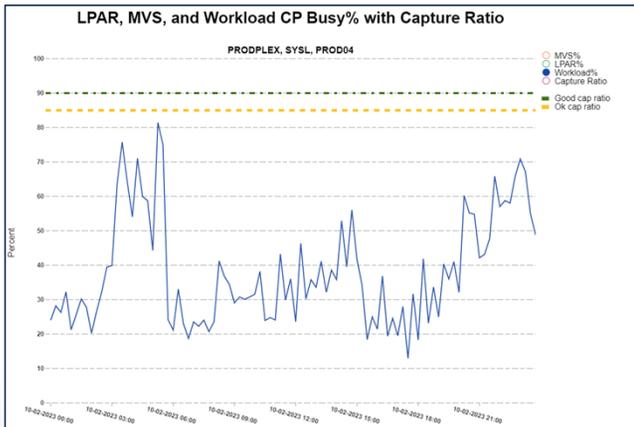
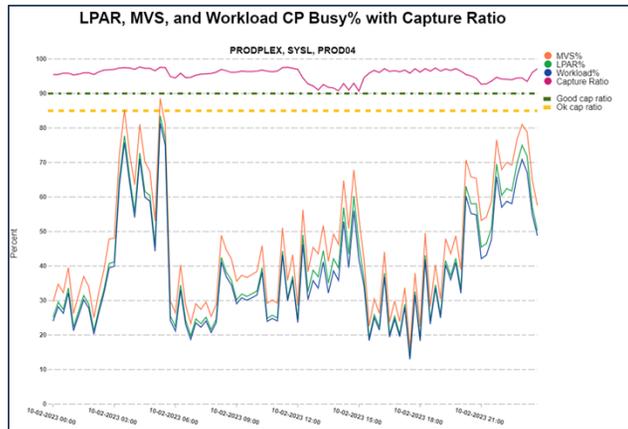


Line charts are great when comparing over time:

Examples where include:

- Compare various utilizations
- Compare number of ended transactions
- Compare rates of reads+writes per second
- Compare intensities
- Compare ratios

These same values would make little sense to visualize as area, column, pie, or most other chart types

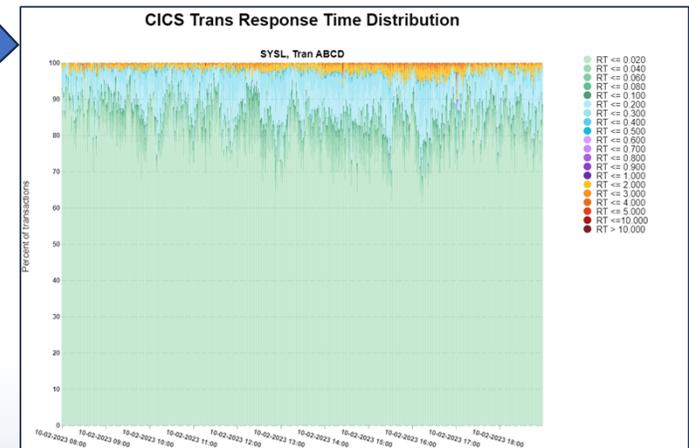
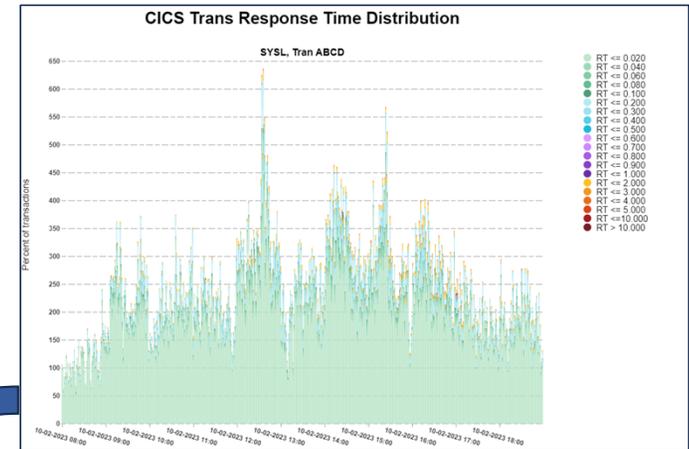
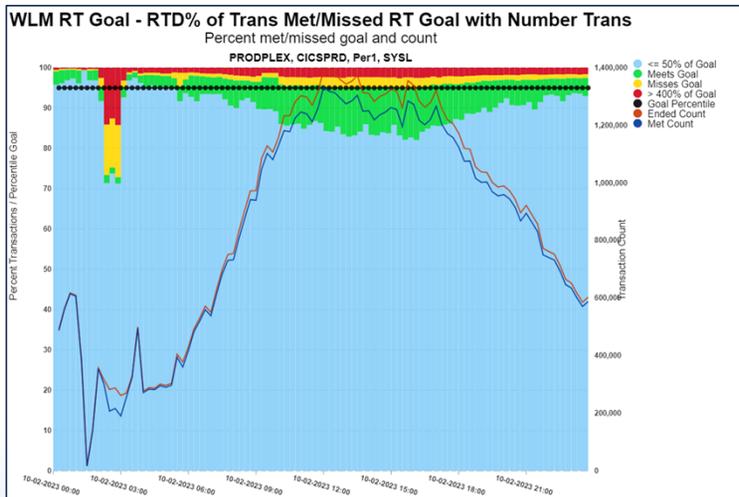


# Picking a chart type: Stacked 100%



## ● Stacked 100% chart

- Use: When need to visualize the percentage breakdown of a total
- Pro: Visually great for distributions
- Con: Total values can be drastically different, so always need to keep in context

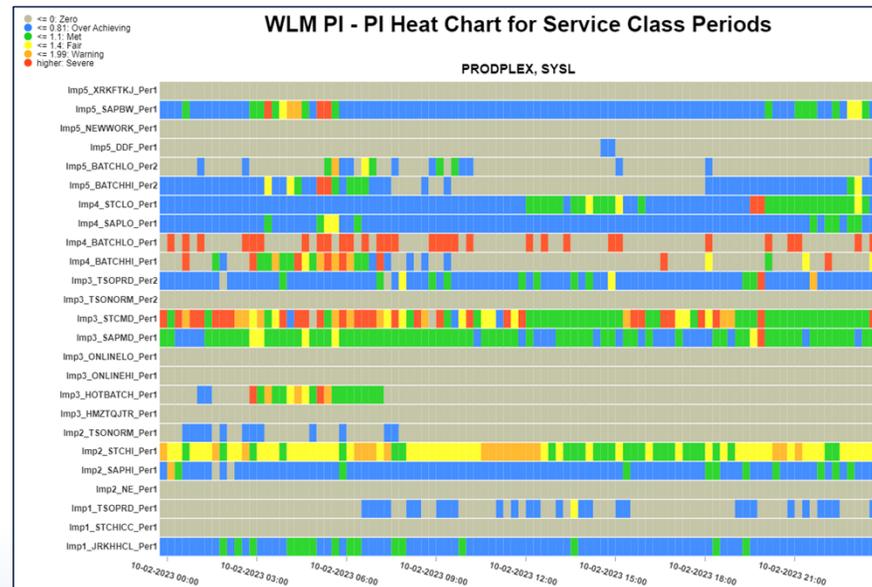


# Picking a chart type: Heat Charts



- Heat chart

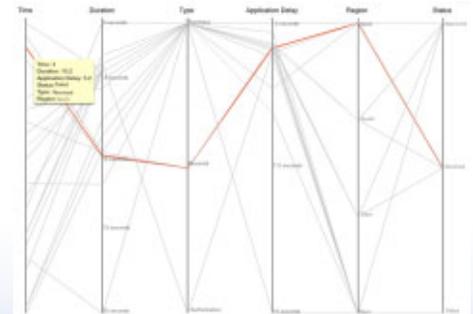
- Use: When needing to visualize time period and series to focus on
- Pro: Visually great seeing good and bad performance of series next to one another
- Con: Series heat indicators are the same for all series



# Picking a chart: Finding relationships



- Pie chart
  - More difficult to perceive differences between angles than length
  - If more than a few slices, labeling becomes difficult
- Horizontal bar charts
  - Bar length makes differences easier to perceive
  - Consider ordering the observations intelligently
  - Can effectively display many more values
- Scatter plots
  - Good for comparing two quantitatively values
  - Correlation generally stands out visually
- Bubble charts
  - Can be used similarly to scatter plots but variances in bubble size and color can encode two more variables
  - Can be difficult to discern small differences in size/color
  - Interactive bubble charts can be very compelling though
- Parallel Coordinates
  - Can be used when variables are both quantitative and qualitative
  - Can help you see correlations between multiple variables
  - Can be used with very large number of observations
  - Limited tooling available



# Colors



- Use white as your background for your chart
- Consider intensities of a single color for data ranges
- Use less saturated colors
  - Reserve vivid colors for highlighting particular data points
  - Consider gray scale for most data, reserving color for highlights
- Use different colors with similar intensities to denote categories of data
- Color blindness is common!
  - Red-green: 7-10%
  - Yellow-blue: 6%
  - Free check tool available at [vischeck.com](http://vischeck.com)
  - Phone apps available too, just hold your phone up to the screen to see what it might look like for a color blind person



# Chart Junk



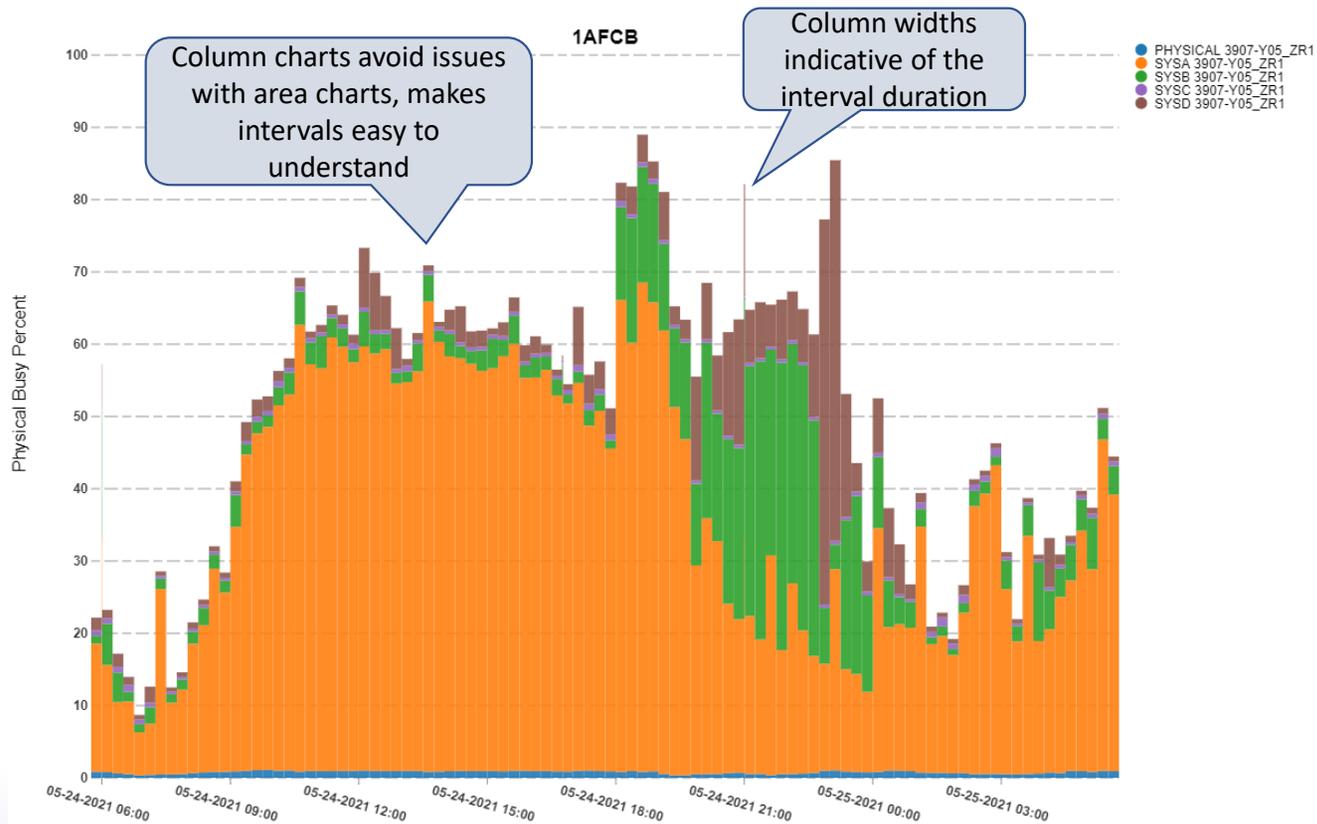
- Don't include what's not needed!
- Don't let visual effects distract the reader from the story of your data
  - Unless obfuscation *is* the goal
- 3-D effects are often overused and unnecessary
- Avoid unnecessary gradients, icons, and backgrounds
  - Sometimes a background indicating thresholds may be ok
- Grid lines don't need to be dark
- Y-axis should almost always start at zero



# Pivotor Prefers Column Charts



## CEC Physical Machine CP Busy% by CEC Serial Number



# Pivotor Avoids Chart Junk



User Interface elements limited to what is really needed

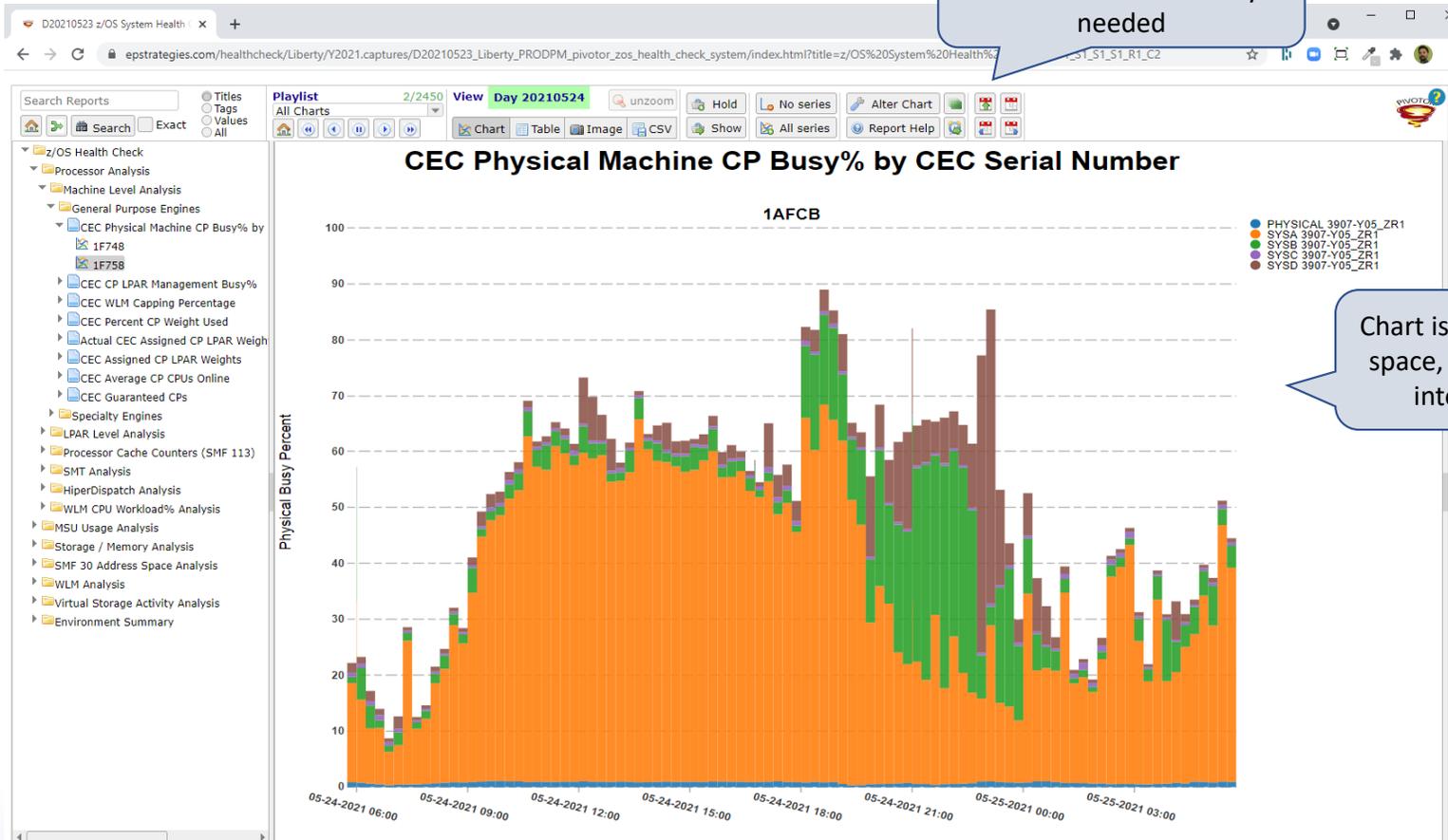
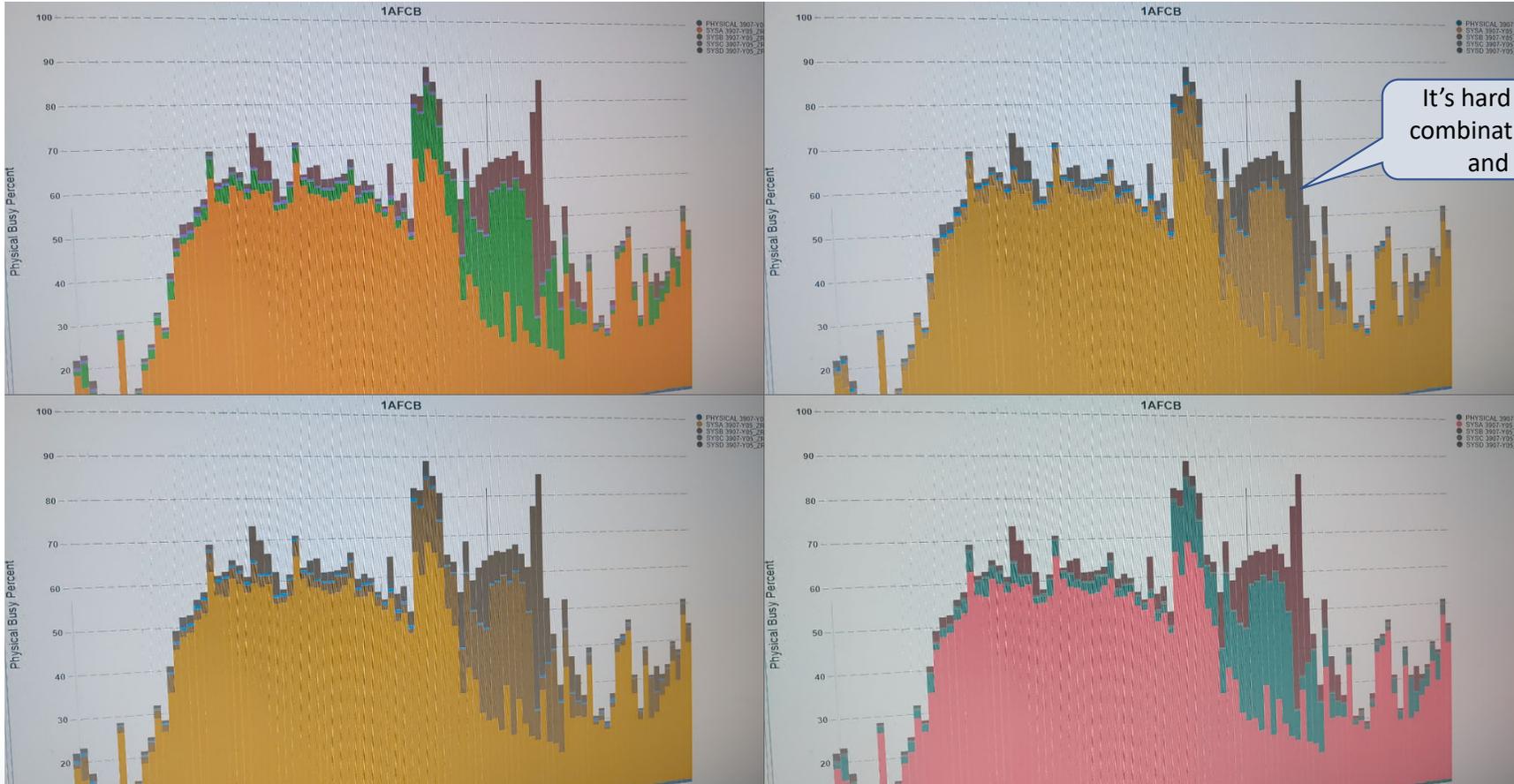


Chart is the majority of the space, making it easier to interpret the data

# Pivotor Tries to Address Color Blindness



It's hard to find a perfect combination for everybody and every chart!

# Feed your brain!

Data → Understanding



# Added Elements of Building a Visual



- Engaging – you’re telling a story!
  - Words + visual = 100
  - Pivotor helps + “fly-overs”
- Graph/report elements – key, title, labels
- The importance of “visual comparisons” and context
- Emotions play a role
  - This is more than just a table of numbers

# It takes “about 13 milliseconds for the brain to process an image”



- Visuals are perceived and processed quicker and easier than written text, tables, or numbers in a table (visual cortex)
- Reading text and thinking/cognition is slower and requires more conscious thinking (cerebral cortex)
- Variations in neural activity due to gender, age, various diagnoses can lead to differences into the speed and efficiency of the brain’s processing of colors, motion, fine detail, and more
- Emotions influence what certain statistics mean to us individually, how we are feeling when reviewing data, and subsequently how we respond to them
- Attention spans are changing, highlighting the importance for a quickly understood display of data and information

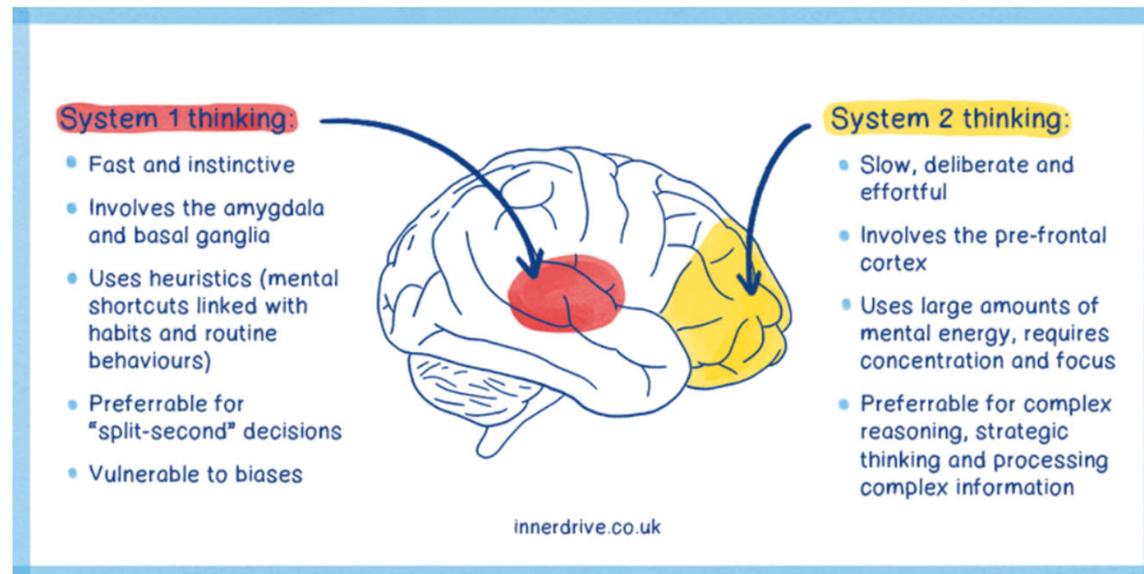


# Dual Processing Theory



Theory that thinking is categorized into 2 systems/types

- System 1 – often from visual processing, autonomous, influenced by emotions, “immediate grasp of complex info”
- System 2 – rational, requires working memory & cognitive stimulation, not autonomous



# Pre-attentive vs. Attentive Processing



**Preattentive attributes of visual perception**

**Form**

Length    Width    Orientation    Size    Shape

Curvature    Enclosure    Blur

**Color**    **Position**    **Motion**

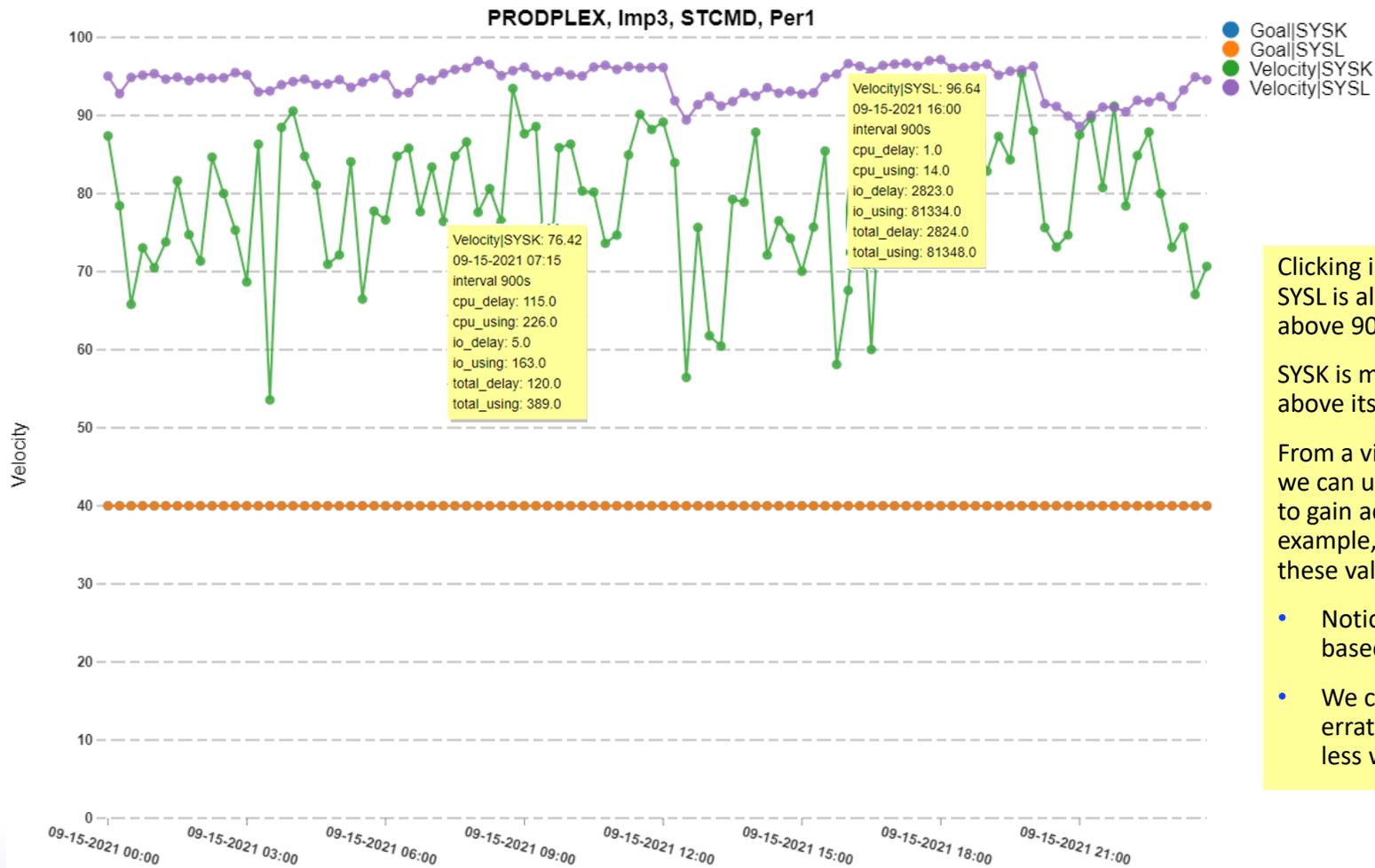
Hue    Intensity    2-D position    Spatial Grouping    Direction of Motion

- Pre-attentive processing includes basic attributes and identifiers then →
- Attentive processing takes place once pre-attentive processing has been done (about 10 milliseconds), for more complete analysis
- Visualizations are a smart option for quantitative data

# WLM Velocity Goal - Achieved Velocity Across Sysplex



# WLM Velocity Goal - Achieved Velocity Across Sysplex



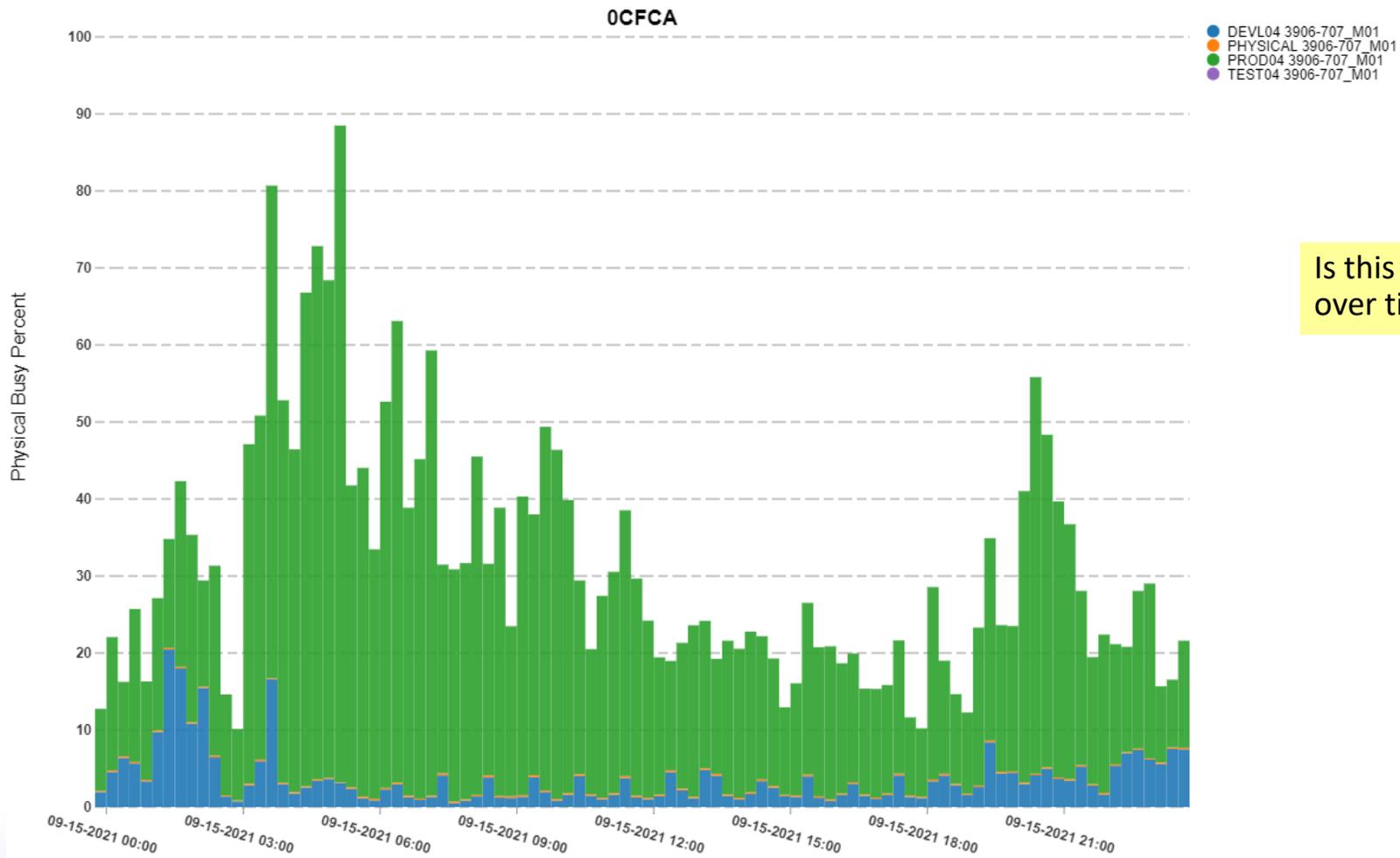
Clicking into this chart we see that SYSL is always achieving a velocity above 90.

SYSK is more variable, but always above its goal.

From a visualization point-of-view, we can use techniques like 'fly over' to gain additional insights. For example, the formula that calculates these values is based on samples:

- Notice the more stable velocity is based on far more samples
- We can also conclude that the erratic velocity of SYSK is due to less work

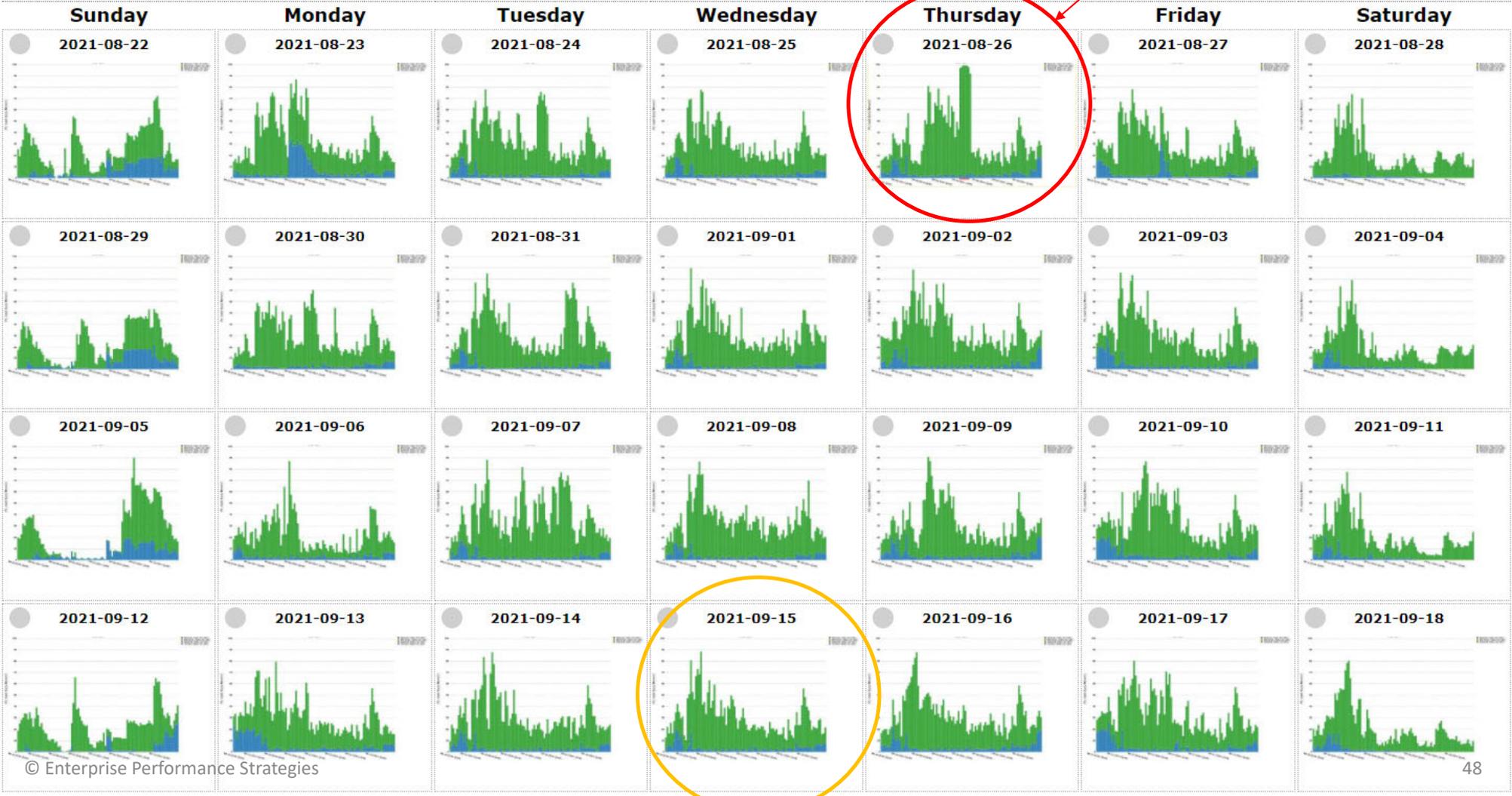
# CEC Physical Machine zIIP Busy%



Is this a common pattern over time?

### CEC Physical Machine zIIP Busy%

OCFCA

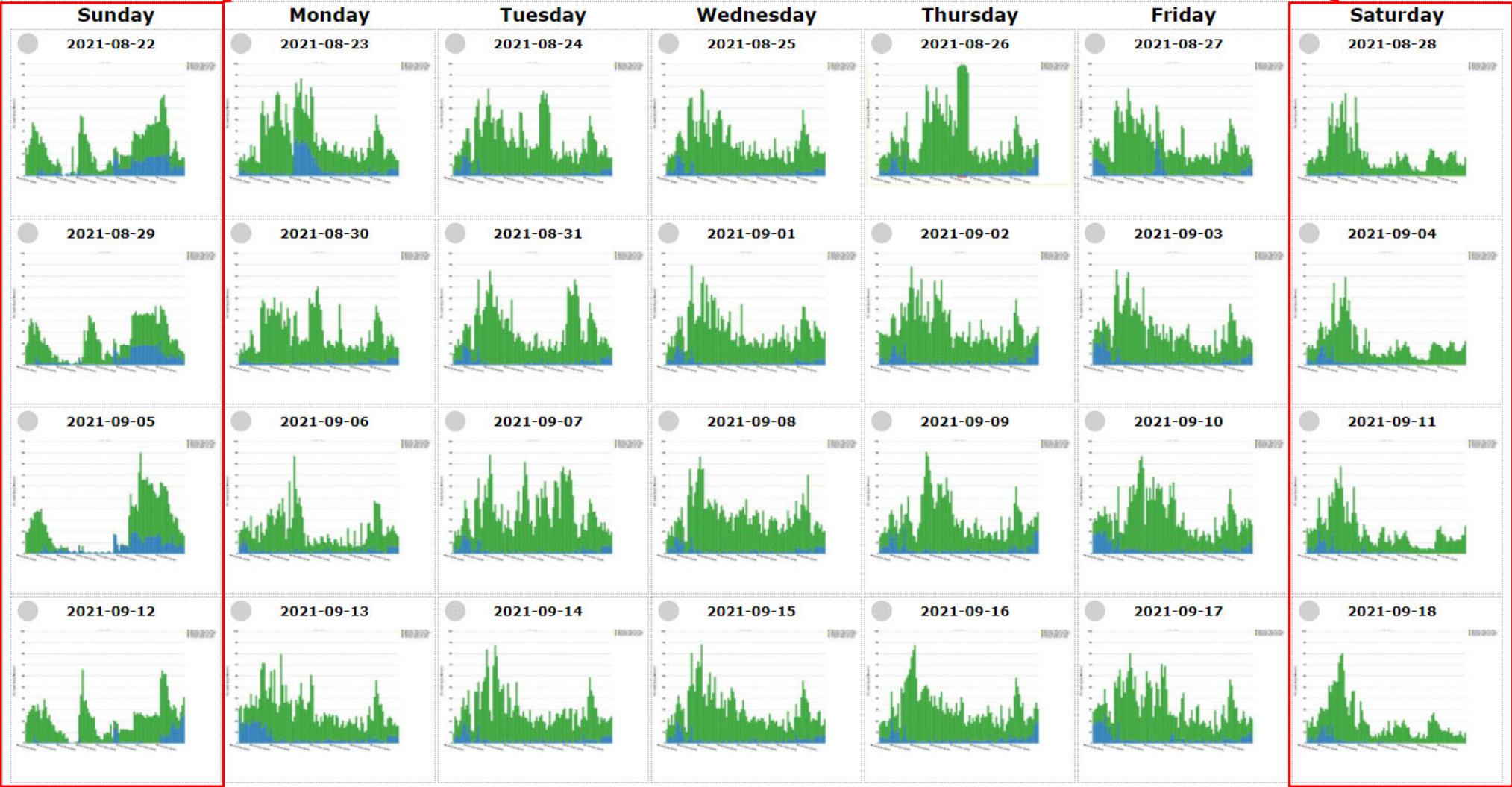


← Previous Week Load Selected Next Week →

Clear Close

### CEC Physical Machine zIIP Busy%

OCFCA





# Final Thoughts: SMF Visualization

# This has been a different webinar...



- Normally we're sharing something interesting about z/OS performance
- Hopefully today you found something interesting about data visualization
- Data visualization is an important performance topic
  - We think a good bit about this when creating new Pivotor reports
  - Poor visualization choices can hide problems
- If you're a Pivotor customer maybe you understand why some Pivotor things are the way they are now
- If you're making your own charts, maybe we've given you some new things to think about
- And if you just wish you had good visualizations of your z/OS performance: maybe we should talk! 😊

# Please submit your session feedback!



- All done via the Whova App
- QR Code to the right to download the Whova App
- This session is **QH**

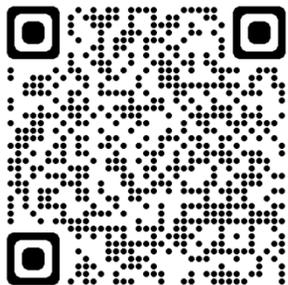


# GSE UK Conference 2023 Charities



- The GSE UK Region team hope that you find this presentation and others that follow useful and help to expand your knowledge of z Systems.
- Please consider showing your appreciation by kindly donating a small sum to our charities this year, Blood Bikes and LimbPower.

<https://www.justgiving.com/crowdfunding/mark-wilson-343>





Questions / Discussion?

# Sources



- 1 - <https://medium.com/vizzuality-blog/making-impactful-visualisations-using-psychology-797a6f0807ea>
- 2 - <https://treehousetechgroup.com/the-psychology-behind-data-visualization/>
- 3 - <https://www.empiricaldata.org/dataladyblog/6-lessons-from-human-psychology-for-effective-data-visualization>
- 4 - <https://www.washington.edu/news/2018/08/16/men-and-women-show-surprising-differences-in-seeing-motion/>
- 5- <https://www.nationalgeographic.com/culture/article/120907-men-women-see-differently-science-health-vision-sex>
- More: <https://www.interaction-design.org/literature/book/the-encyclopedia-of-human-computer-interaction-2nd-ed/data-visualization-for-human-perception>
- <https://medium.com/nightingale/how-to-apply-gestalt-psychology-principles-in-data-visualization-6242f4f1a3de>
- <https://www.dundas.com/resources/blogs/introduction-to-business-intelligence/brief-history-data-visualization>