An Approach To Data Visualization In Power BI

Meagan Longoria

January 12, 2017

About Me

Meagan Longoria

@mmarie

I live in Denver, Colorado. I'm proud to be a Microsoft Data Platform MVP. I'm a Solution Architect with BlueGranite who spends a lot of time thinking about how to use Power BI and data visualization techniques to make data useful for people. I enjoy speaking at conferences and user group meetings as well as blogging at DataSavvy.me.

BLUEGRANITE
BUSINESS INSIGHT. DELIVERED.

MVP

Programming Note

Data visualization can be used for

exploratory (sense-making) analysis
or
explanatory (communication) analysis

When we share Power BI reports, our goal is usually to communicate important information effectively by using visuals to:

- Clarify
- Provide memorable insights
- Help the audience make a decision or take action

What If I Told You…

Your explanatory data visualization success is largely determined before you ever place a chart on the canvas.

Do you know how to prepare?

Why Is Data Visualization So Important?

The greatest value of a picture is when it forces us to notice what we never expected to see. - John Tukey

Our (developers') outputs are decision-makers' inputs – and their outputs are what ultimately matter. – Rob Collie (Power Pivot Pro)

Noisy/clean data

Noisy/clean data viz

Noisy/clean decisions

Why Does a Data Visualization Fail?

Lack of appropriate data

Reports as intermediate steps

Poor presentation that makes it difficult to gain insight and take action

Poor presentation that discourages engagement

How To Get Started

Ask the right questions

Get The Scoop

Who is your audience? (Executives? Analysts? Website users?)

- Helps determine needs, priorities, and level of detail

What metrics are important? What is the dimensionality?

Is the report operational, analytical,
or a mix of both?

You'll Never Guess What Happens Next

What do they do with the data/information?

Sometimes reports are step 1 in a process. What comes next?

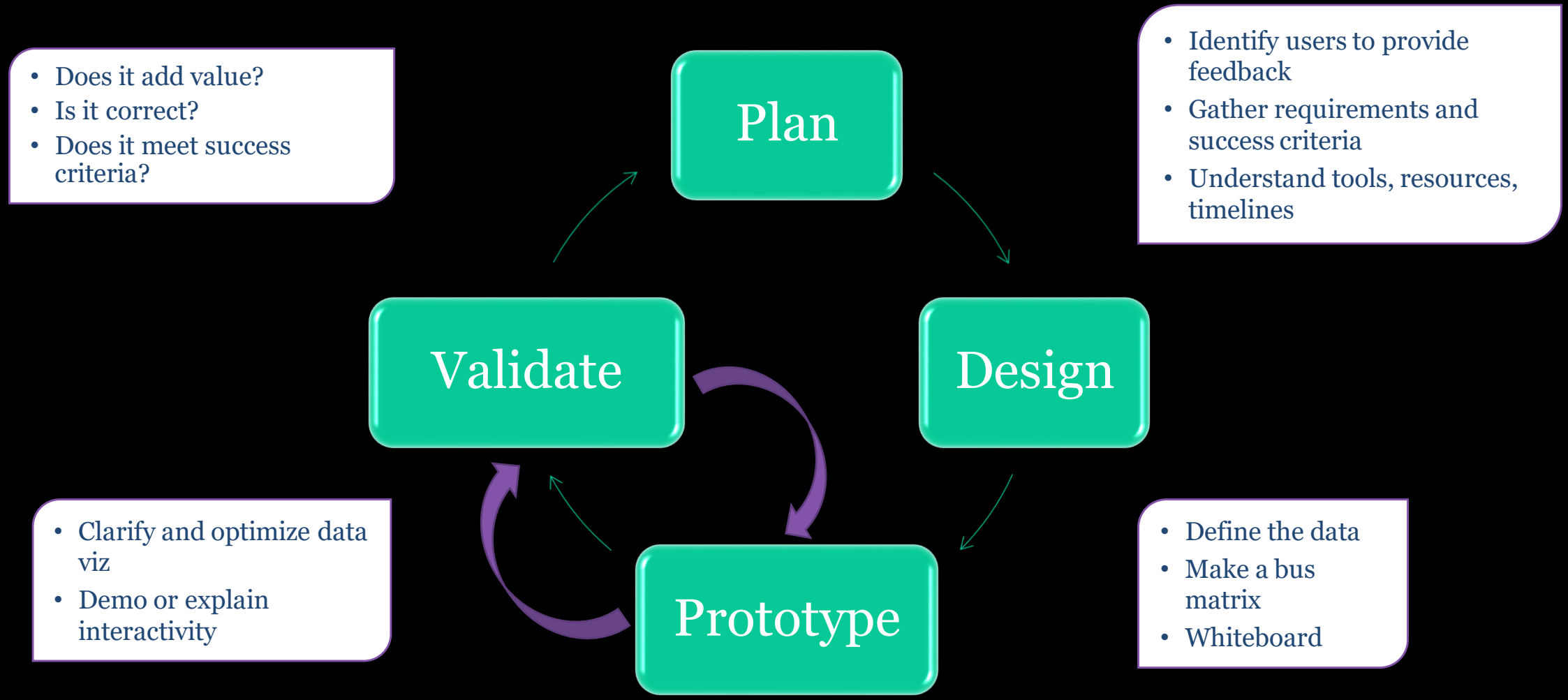
Borrow a page from the 5 Whys

We're developers. We automate repetitive and tedious things.

Can we add more value?

Add predictive or prescriptive capabilities? Push alerts?

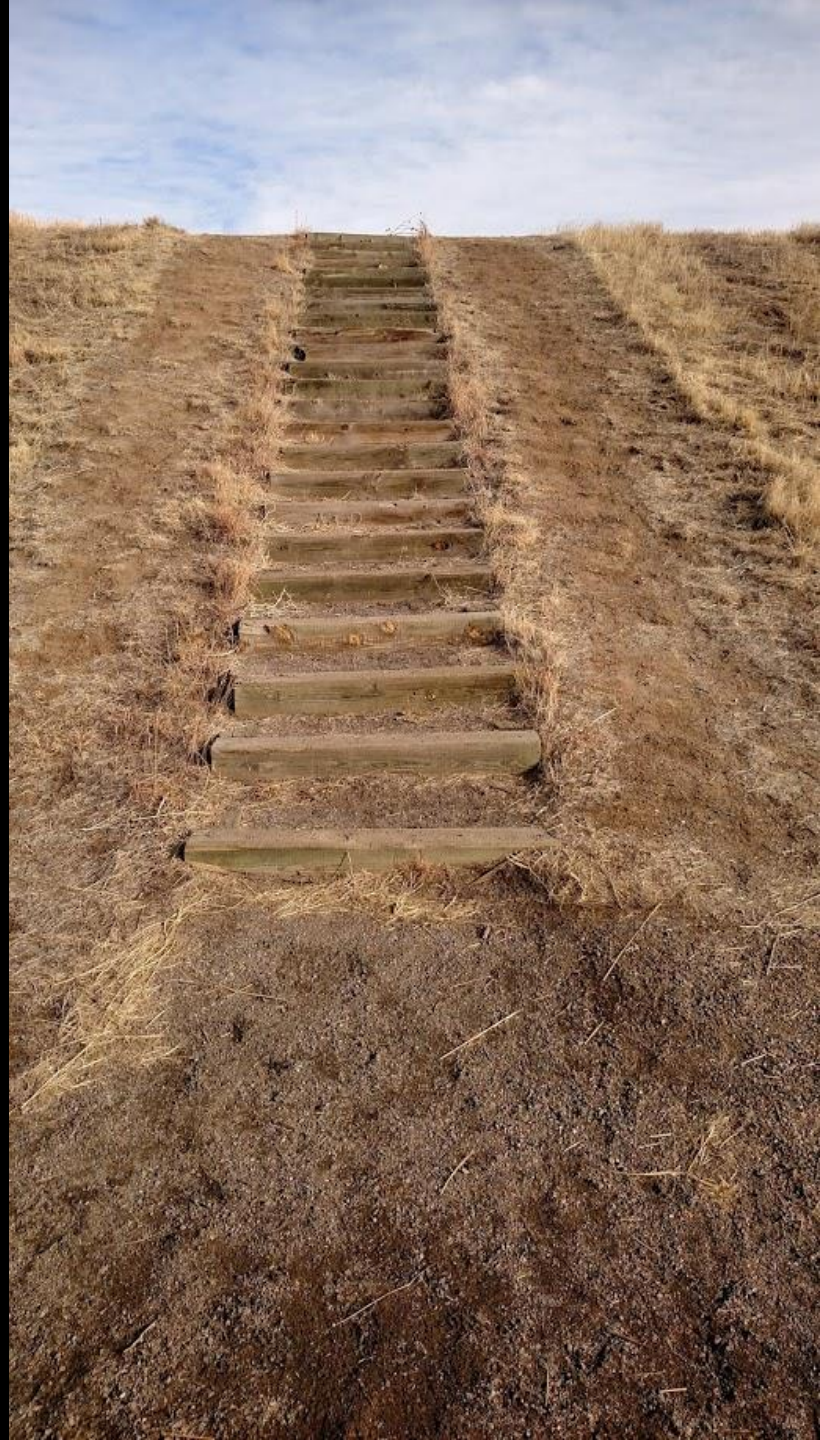
Begin The Development Process



The Data Viz Design Steps

1. Understand the context and craft your message
2. Choose an appropriate visual display
3. Eliminate clutter
4. Focus attention where you want it

(First 4 steps from Storytelling With Data book)



Context/Message

Who is your audience?

What do you want your audience to know or do?

How can you use data to help make your point?

“Know that even if you highlight or recommend the wrong thing, it prompts the right sort of conversation focused on action… If you simply present data, it's easy for your audience to say ‘Oh, that's interesting’ and move on to the next thing. But if you ask for action, your audience has to make a decision whether to comply or not.”
- Cole Nussbaumer Knaflic

My Report

Design: Bus Matrix

Beware of inferred relationships in flat data sets.

Think like the user in terms of relationships and business attributes.

Facts/Dimensions	Date	Item	Vendor	Distr Center	Shipper	Store	Customer	Promo
Purchase Orders	x	x	x	x				
Distribution Center Delivery	x	x	x	x	x			
Distribution Center Inventory	x	x		x				
Store Deliveries	x	x		x	x	x		
Store Inventory	x	x				x		
Store Sales	x	x				x	x	x
Returns	x	x				x	x	x

Design: Data Documentation

Dimension information	
Name	a simple concise name to help users identify it
Definition	a non-technical definition that explains what is being measured
Data Source(s)	the source system that contains the underlying data (not necessarily where you will query it)
Important/Commonly Used Attributes	the field the user slices by or pivots on to get useful information
Time reference	indicates if the dimension data is as originally entered, current state, or historical (type 0, 1, or 2)

KPI/Metric definition	
Name	a simple concise name to help users identify it
Definition	a non-technical definition that explains what is being measured
Related Business Objective	ties the KPI to an overall business objective so users understand the reason for the measurement
Unit Type	count, dollar amount, percent, etc.
Frequency of Measurement	indicates how often the underlying data is gathered and over what period of time
Data Source(s)	the source system that contains the underlying data (not necessarily where you will query it)
Target Goal	the goal against which performance is measured
Calculation	a technical definition of the metric/KPI (could be similar to an Excel formula)
Statuses	explanation of the possible states (red/yellow/green, 1 - 5, etc.)
Related Metrics or KPIs	any other KPIs that are related as a parent, child, or sibling of the metric/KPI

Design: Whiteboard

- Are you making a single chart, a report, or a dashboard?
- Come up with high-level ideas and place them on your canvas.
- What specific items of information should be displayed? What does each of these items tell you, and why is that important? At what level of summary or detail should the information be expressed?

- Use your Data Definitions and Bus Matrix as a catalog!

Design Questions

- Which items of information are most important for achieving your objectives?
- What are the logical groupings that could be used to organize items of information on the dashboard? In which of these groups does each item belong?
- What are the most useful comparisons that will allow you to see these items of information in meaningful context?

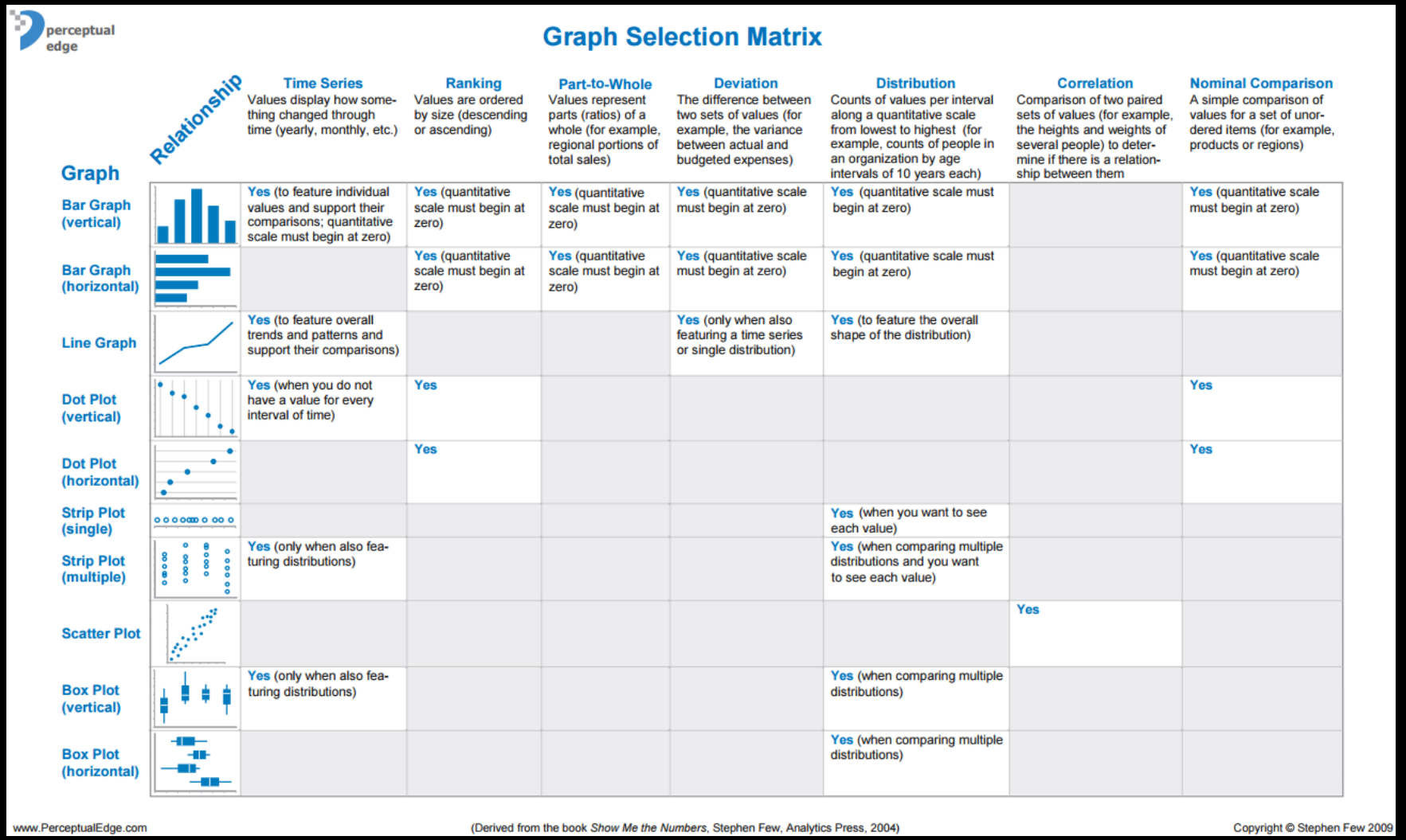
- (From Stephen Few's *Information Dashboard Design*)

Design: Chart Types

You can categorize charts into these types:

Categorical	Comparing categories and distributions of quantitative values
Hierarchical	Charting part-to-whole relationships and hierarchies
Relational	Graphing relationships to explore correlations and connections
Temporal	Showing trends and activities over time
Spatial	Mapping spatial patterns through overlays and distortions

From *Data Visualization: A Handbook for Data Driven Design* by Andy Kirk)



perceptual edge

Graph Selection Matrix

Graph (Relationship)	Time Series Values display how something changed through time (yearly, monthly, etc.)	Ranking Values are ordered by size (descending or ascending)	Part-to-Whole Values represent parts (ratios) of a whole (for example, regional portions of total sales)	Deviation The difference between two sets of values (for example, the variance between actual and budgeted expenses)	Distribution Counts of values per interval along a quantitative scale from lowest to highest (for example, counts of people in an organization by age intervals of 10 years each)	Correlation Comparison of two paired sets of values (for example, the heights and weights of several people) to determine if there is a relationship between them	Nominal Comparison A simple comparison of values for a set of unordered items (for example, products or regions)
Bar Graph (vertical)	Yes (to feature individual values and support their comparisons; quantitative scale must begin at zero)	Yes (quantitative scale must begin at zero)	Yes (quantitative scale must begin at zero)	Yes (quantitative scale must begin at zero)	Yes (quantitative scale must begin at zero)		Yes (quantitative scale must begin at zero)
Bar Graph (horizontal)		Yes (quantitative scale must begin at zero)	Yes (quantitative scale must begin at zero)	Yes (quantitative scale must begin at zero)	Yes (quantitative scale must begin at zero)		Yes (quantitative scale must begin at zero)
Line Graph	Yes (to feature overall trends and patterns and support their comparisons)			Yes (only when also featuring a time series or single distribution)	Yes (to feature the overall shape of the distribution)		
Dot Plot (vertical)	Yes (when you do not have a value for every interval of time)	Yes					Yes
Dot Plot (horizontal)		Yes					Yes
Strip Plot (single)					Yes (when you want to see each value)		
Strip Plot (multiple)	Yes (only when also featuring distributions)				Yes (when comparing multiple distributions and you want to see each value)		
Scatter Plot						Yes	
Box Plot (vertical)	Yes (only when also featuring distributions)				Yes (when comparing multiple distributions)		
Box Plot (horizontal)					Yes (when comparing multiple distributions)		

http://media.juiceanalytics.com/downloads/graphselectionmatrix_sfew.pdf

Design: Choose The Right Chart

What is the right graph for my situation? ...whatever is easiest for your audience to read.

- Cole Nussbaumer Knaflic

No chart is evil, they just have different roles & limitations.

- Andy Kirk

Check Yourself

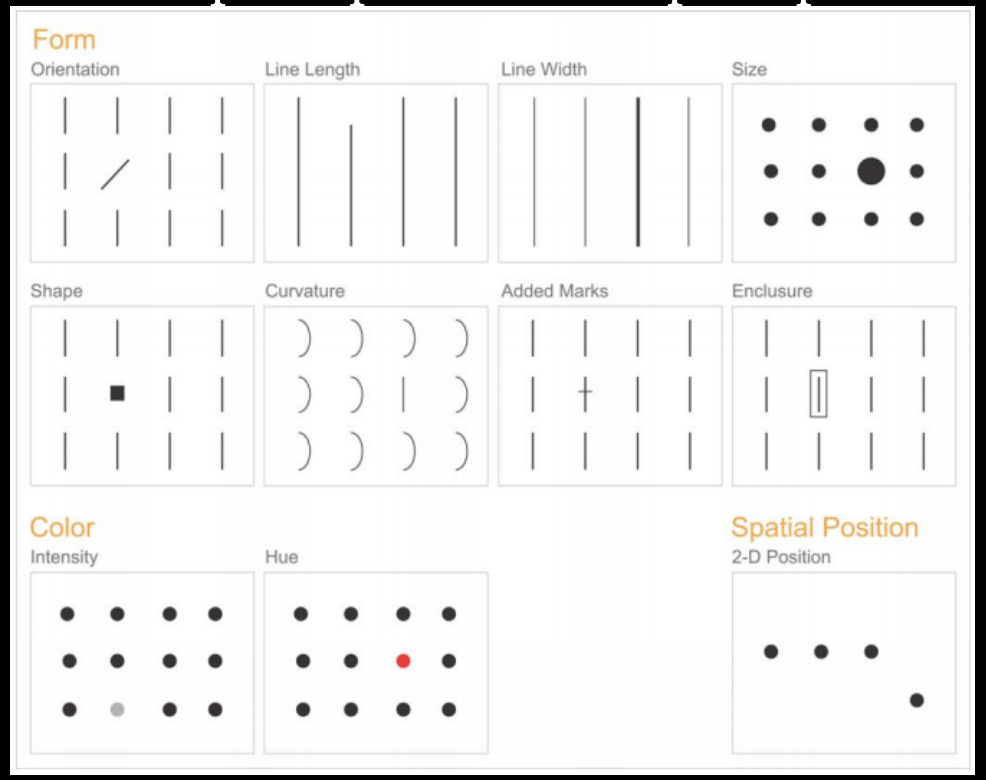
- Are the groupings of information obvious?
- Are the key metrics being featured adequately?
- Can you easily spot the items that need attention?
- Is enough information being displayed about the items that need attention to decide whether you must respond by taking action?

Prototype: Technical Items To Consider

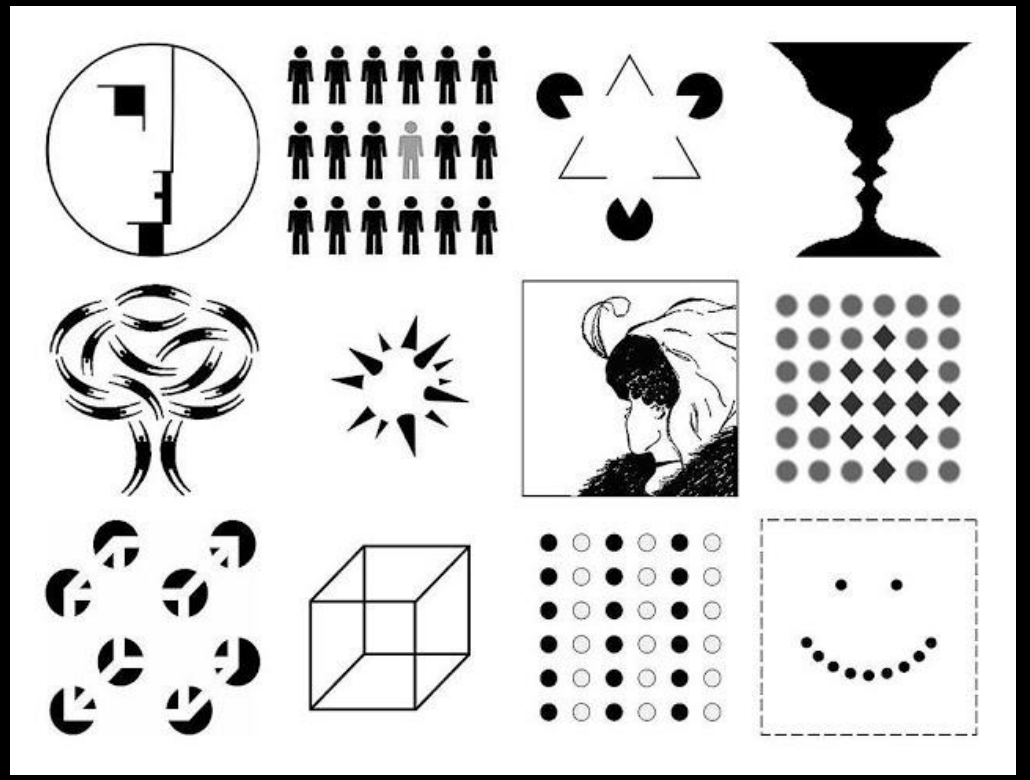
- Method of access (mobile?)
- Capabilities of the reporting tool
- Average or minimum screen resolution/size of users

Optimize Your Data Viz

Limits of working memory: 3 chunks at a time
Encoding data for rapid perception using preattentive attributes
Gestalt principles of visual perception

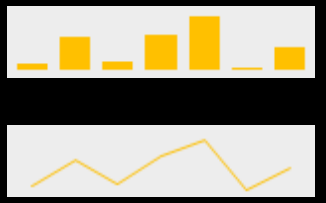


https://www.perceptualedge.com/articles/ie/visual_perception.pdf



More Tips

If you don’t want to include large graphs, consider bullet graphs and sparklines to provide visual context.



Use enough descriptive text to provide necessary context.

Put supplementary information within reach.

Reduce information to what's essential.

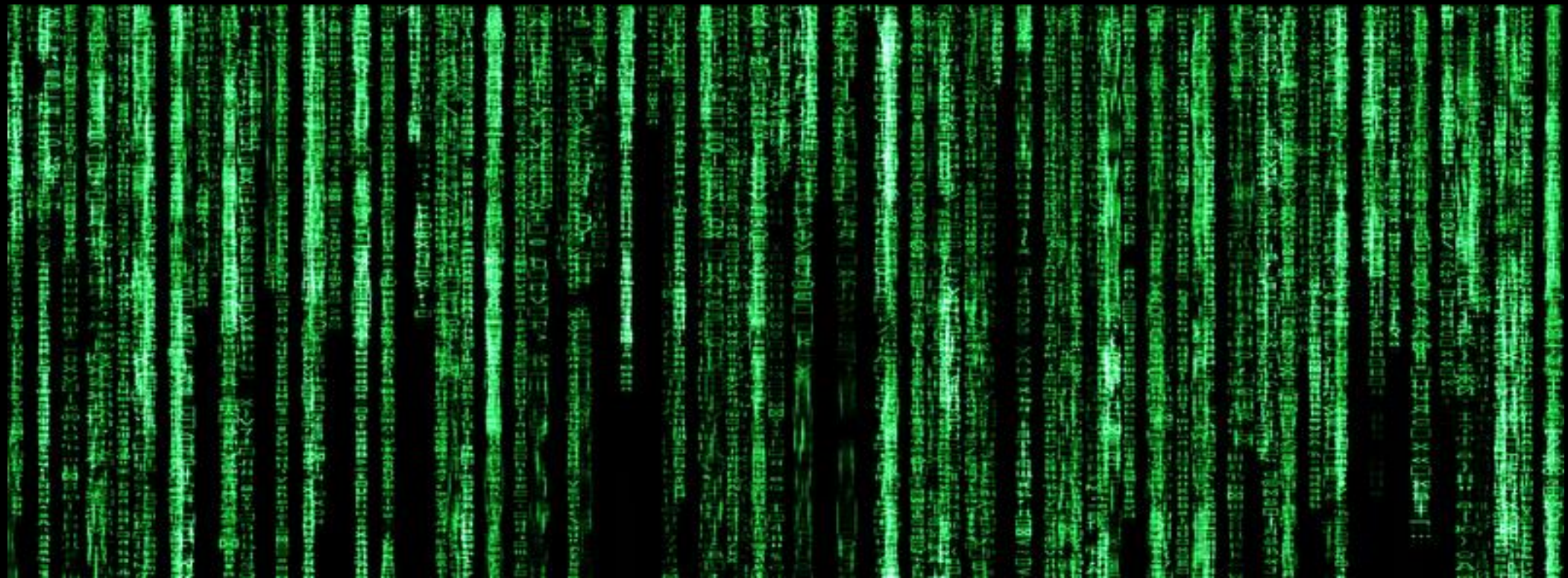
Make the experience aesthetically pleasing.

Review The Report

For Optimization And Essential Information

A Note On Big Data

Kirk: Visualizing big data isn't a data problem, it's a summarization problem. You've only got so many pixels on the screen. Summarize and then add interactivity to explore more detail.



Eliminate Clutter

Limits of working memory: 3 – 5 chunks at a time

Interpreting reports creates:

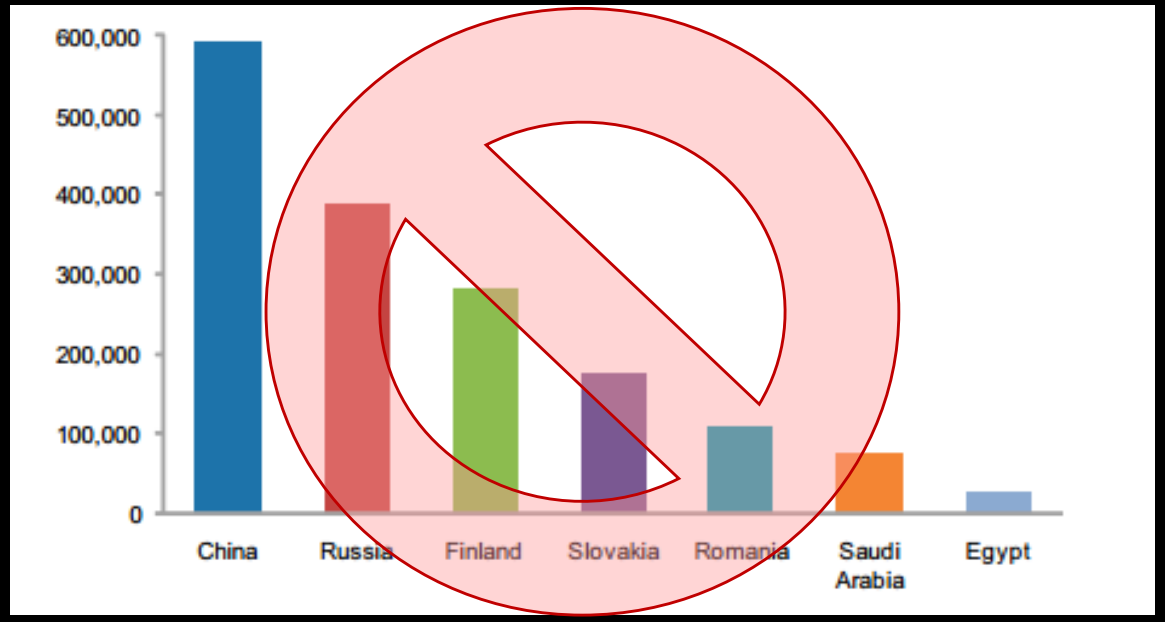
Cognitive Load

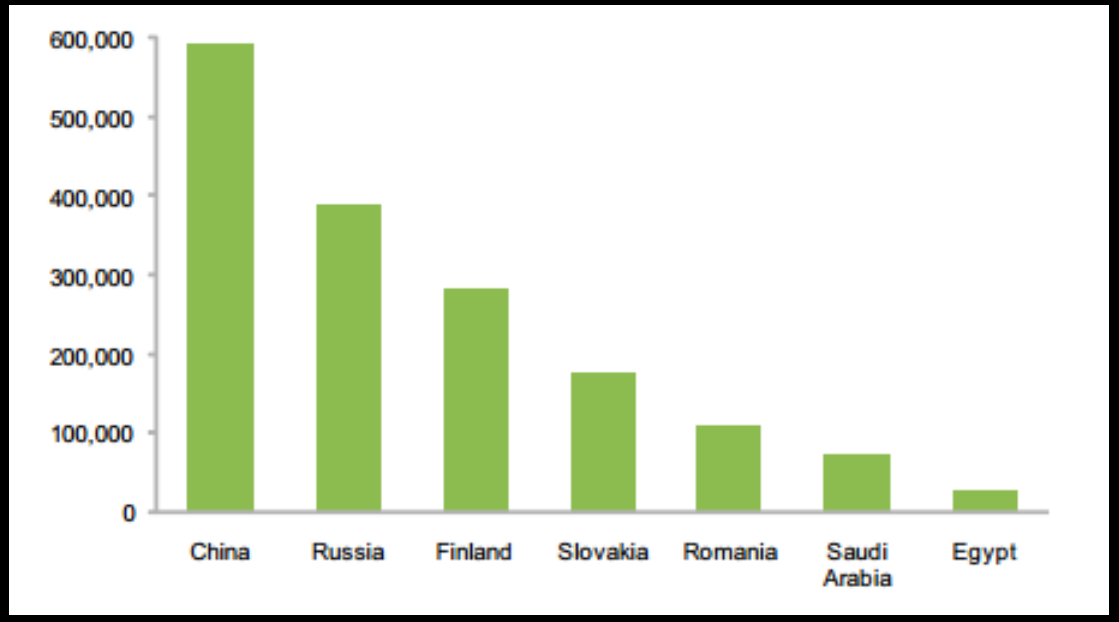
Few: Maximize data-ink ratio

Kirk: Balance data-ink ratio, maximize reward/effort

Choose Appropriate & Meaningful Colors

Changing colors indicates a difference



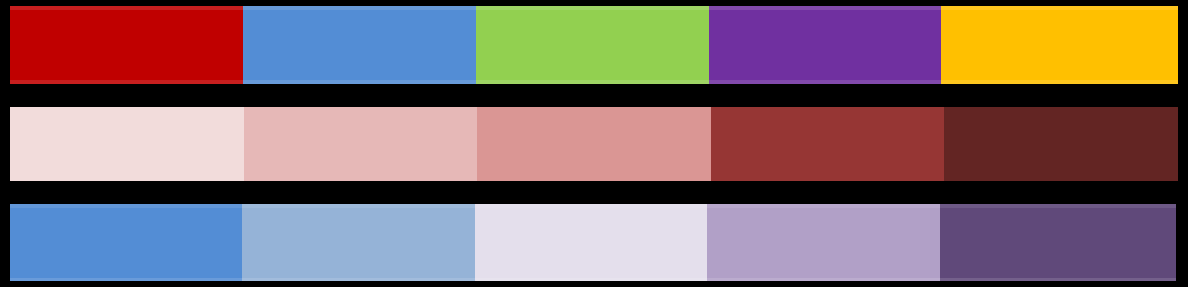


Color Palettes

Categorical

Sequential

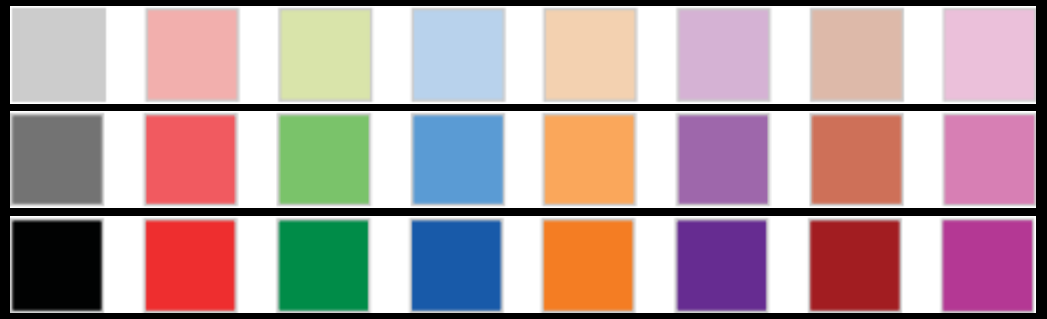
Diverging



Light

Medium

Dark



Good tool: https://txstate-etc.github.io/tints-and-shades/

Color Palettes

C

S

PLEASE STOP SCREAMING AT ME!

Good tool: [illegible]

The Squint Test

Shrink things down and/or half close your eyes to see what colored properties are most prominent and visible.

Are those the right ones?

Color Vision Deficiency

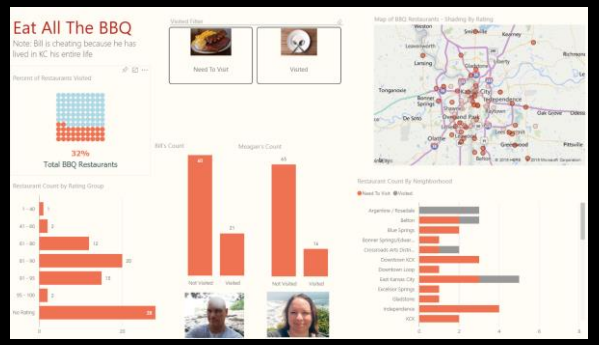
Color Vision Deficiency affects 1 in 12 men and 1 in 200 women.

Red-green color blindness is most common.

Blue and orange are good options for safe colors.

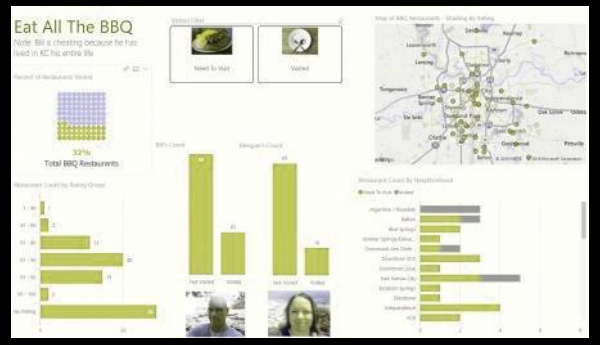
Use http://www.color-blindness.com/coblis-color-blindness-simulator/ to test your viz.

No CVD

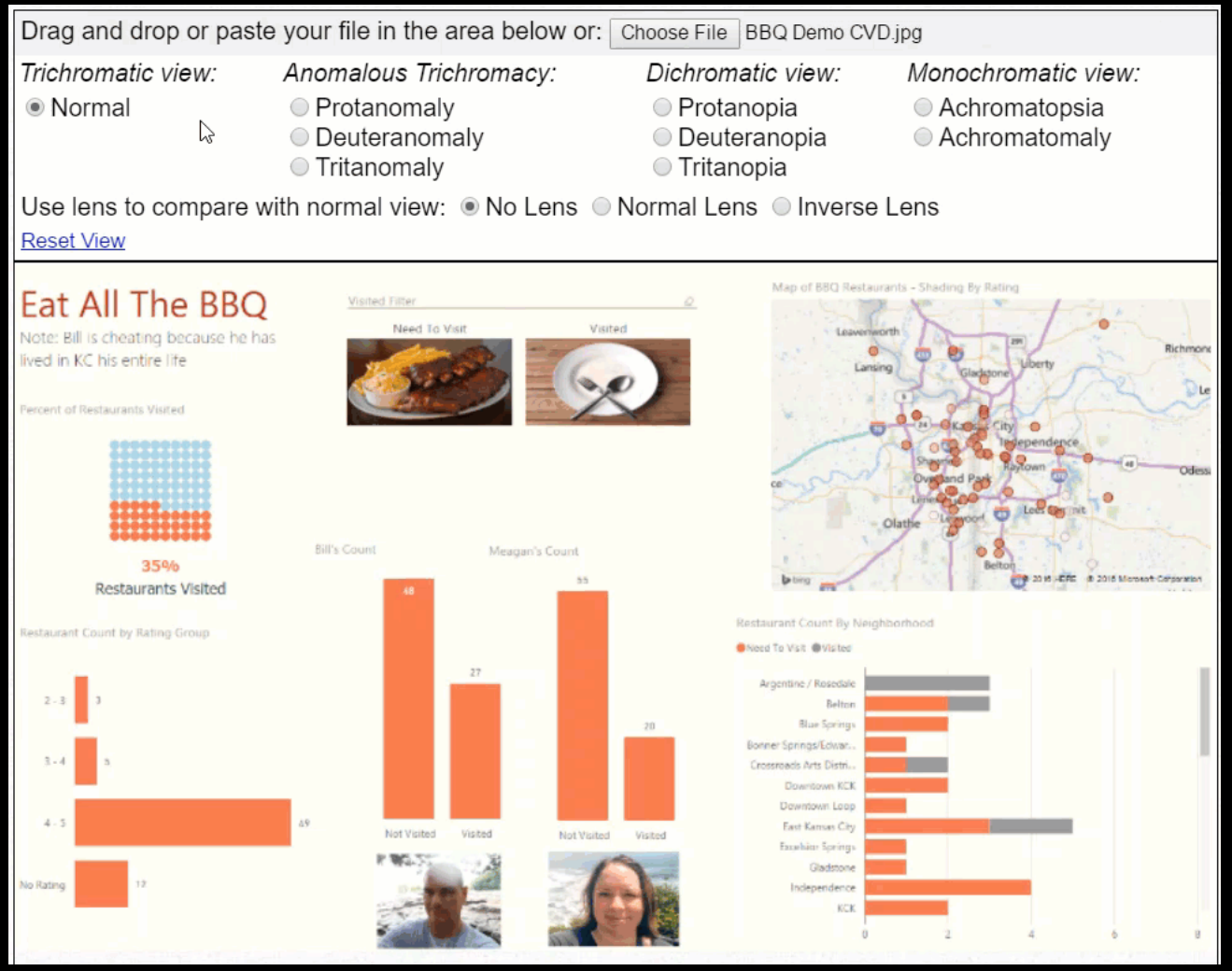


Deuteranomoly

Deuteranopia



Color Vision Deficiency Demo



Review Report

For Clutter and Use of Color

Remember This

Unless you are the main user, you are not building this data viz for you. Build your data viz to provide the most to your users.

Do not overpromise with your prototype and under deliver with your final product.

Don’t promise features that don’t yet exist, but don’t be afraid to get creative to meet user’s needs.

The one true measure of success is adoption/usage.

Getting people to engage is sometimes as important as building the cognitively most valid method. – Andy Cotgreave

Engagement

Visual Appeal

People perceive more aesthetic designs as easier to use and more readily accept and use them

Usability

Affordances – Make it obvious how the audience should interact with the visualization

Accessibility – Design that is usable by people of widely varying technical skills

- Don’t overcomplicate
- Text is your friend

The Mobile Report

Links for Further Learning

- Storytelling With Data: http://www.storytellingwithdata.com/
- Stephen Few/Perceptual Edge blog: http://www.perceptualedge.com/blog/
- Paul Turley - Transforming Reporting Requirements Into a Visual Masterpiece: https://www.youtube.com/watch?v=7c1hjdEzNfQ
- Preattentive Features and Tasks video: https://www.youtube.com/watch?v=wnvoZxe95bo
- Data Viz Done Right: http://www.datavizdoneright.com/
- Power Pivot Pro: A New Take on “Data Quality?”: http://www.powerpivotpro.com/2015/06/a-new-take-on-data-quality/ Improving Data Viz Effectiveness: http://www.blue-granite.com/blog/improving-data-visualization-effectiveness
- Andy Kirk: Separating Myth From Truth in Data Visualization: https://www.brighttalk.com/webcast/9059/193677
- Gestalt Principles Composition Image By Impronta (Own work) [CC BY-SA 3.0 (http://creativecommons.org/licenses/by-sa/3.0)], via Wikimedia Commons: http://commons.wikimedia.org/wiki/File:Gestalt_Principles_Composition.jpg

Questions & Final Comments

Keep In Touch

Slides are on my blog at https://datasavvy.me/presentations/

Tweet me at @mmarie

Come work with me: https://www.blue-granite.com/about