

Data Science Communication: Effective Visualizations

Shannon E. Ellis, Ph.D
UC San Diego

• • •

Department of Cognitive Science
sellis@ucsd.edu

Overview

01:

Your Audience



02:

Storytelling



03:

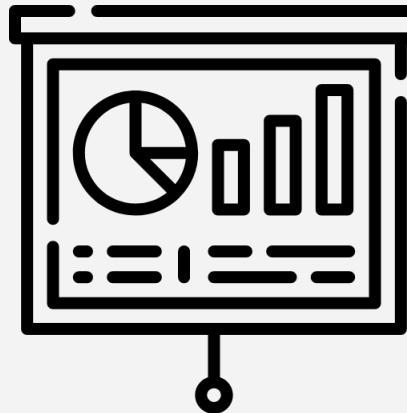
The Grammar of Graphics



04:

The Glamour of Graphics

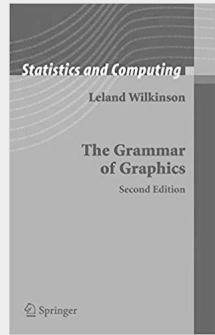




03:

The Grammar of Graphics

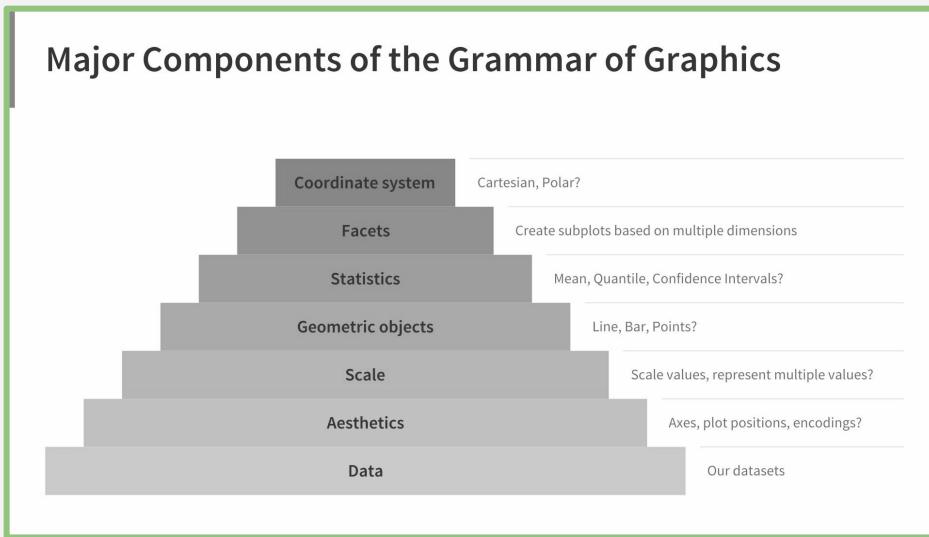
The Grammar of Graphics



In three parts

- Data
 - Aesthetics
 - Geometries
-

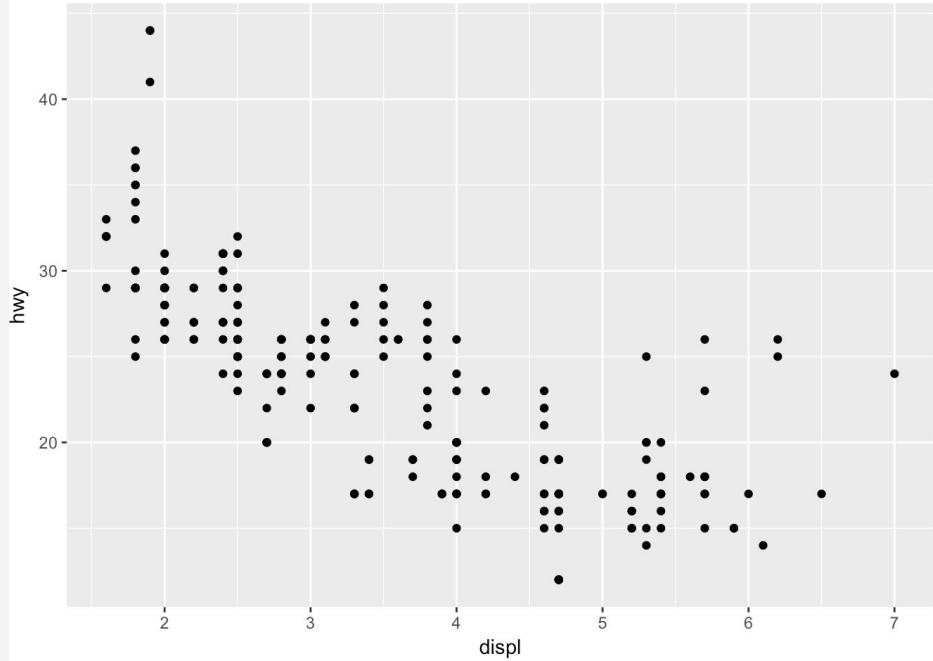
When Plotting Information...



1. Start with the data
2. Determine what positions the values should take, on what scale, and how encoded
3. Choose plot type
4. Stats, facets, coordinate system

Source: <https://towardsdatascience.com/a-comprehensive-guide-to-the-grammar-of-graphics-for-effective-visualization-of-multi-dimensional-1f92b4ed4149>

The Grammar of Graphics

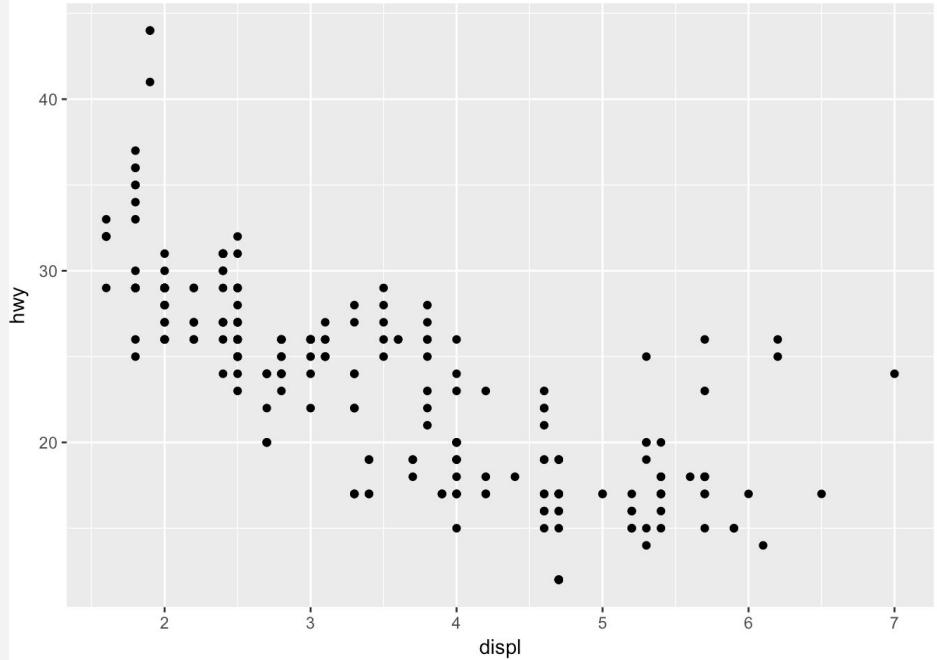


Data: what's being plotted

Aesthetic: maps data to geom

Geometry: Type of plot

The Grammar of Graphics



Data: what's being plotted
displ; hwy

Aesthetic: maps data to geom
x=displ; y=hwy

Geometry: Type of plot
scatterplot

The Grammar of Graphics

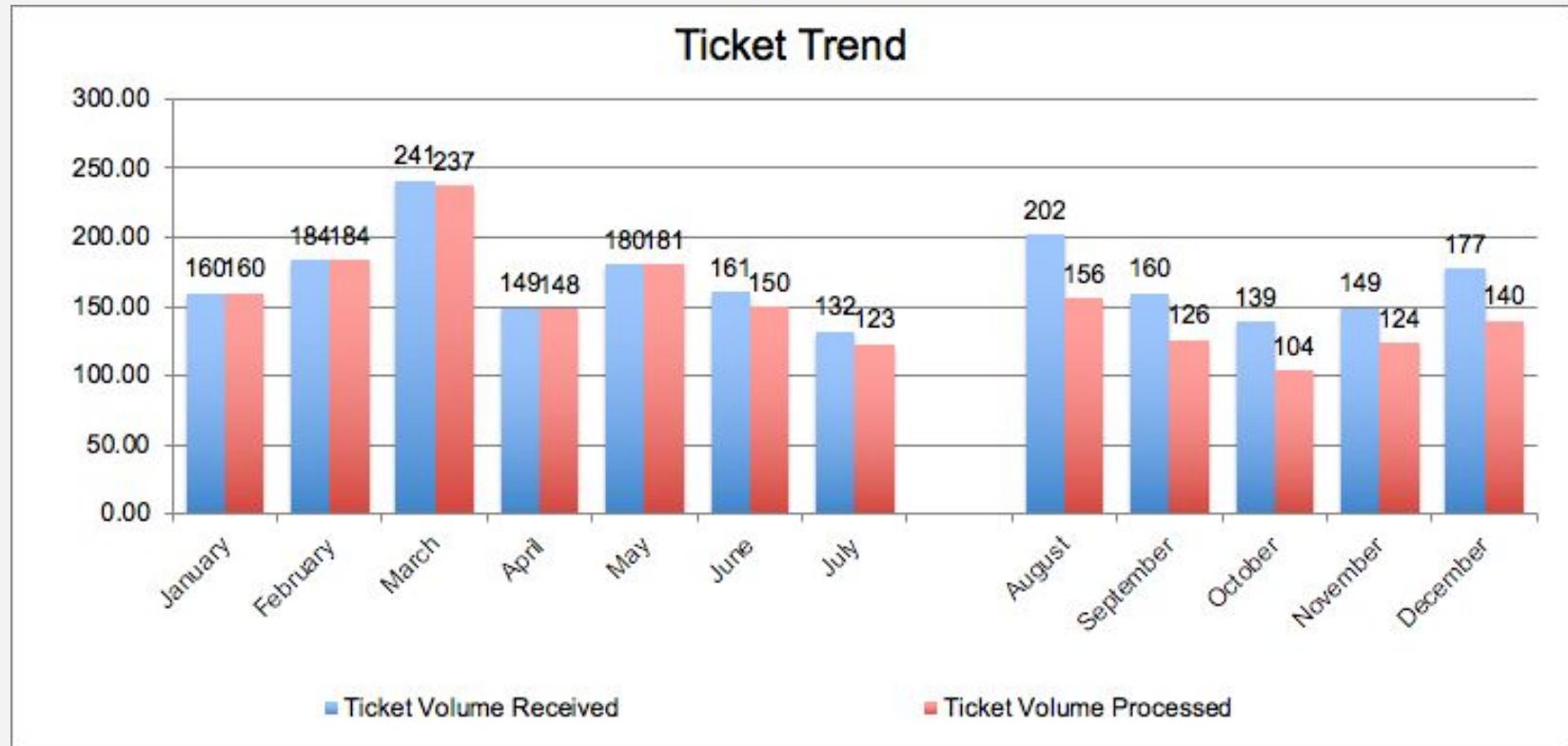
Data: what's being plotted
month, tickets received,
tickets processed

Aesthetic: maps data to geom

Geometry: Type of plot

Month	Tickets Received	Tickets Processed
Jan	160	160
Feb	181	184
Mar	241	237
...
Oct	139	104
Nov	149	124
Dec	177	140

The Grammar of Graphics

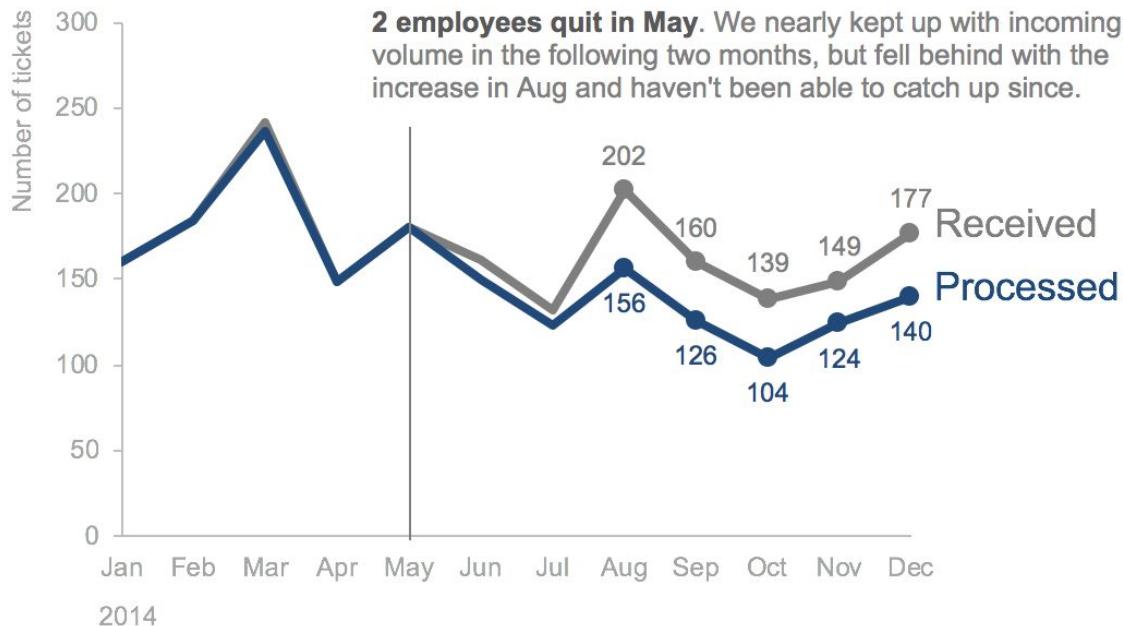


Storytelling with Data (Nussbaumer Knaflc)

Please approve the hire of 2 FTE

to backfill those who quit in the past year

Ticket volume over time



Data source: XYZ Dashboard, as of 12/31/2014 | A detailed analysis on tickets processed per person and time to resolve issues was undertaken to inform this request and can be provided if needed.

The Grammar of Graphics

from Data to Viz: If
you struggle to figure
out what graph is
most appropriate for
the data you have



from **Data** to **Viz**

<https://www.data-to-viz.com/>



04:

The Glamour of Graphics



Graphics: Ground Rules

1. Choose the right type of visualization
2. Include an informative title
3. Label your axes
4. Ensure text is readable

The Glamour of Graphics

1. What's your point?

Headline

2. How can you emphasize your point in your chart?

Chart type, data, color, highlights, annotations

3. What does the final chart show exactly?

Description, legend/keys, source

HYPOTHESIS

PROOF

EXPLAINING

determines

determines

The Glamour of Graphics



Why glamour?

- Engage more people
 - Aesthetic-usability effect
 - Respect your audience
-

Source: Will Chase's rstudio::conf 2020 talk

The Glamour of Graphics



In three parts

- Layout
 - Typography
 - Color
-

Source: Will Chase's rstudio::conf 2020 talk

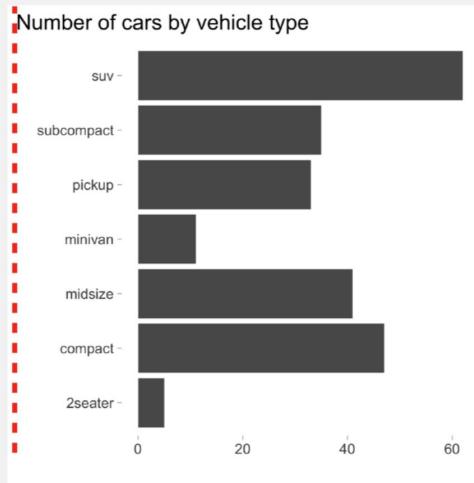
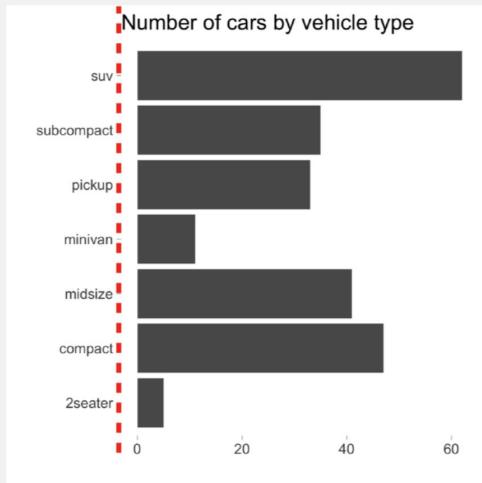
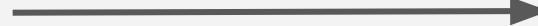


Good Figures: Ground Rules

1. Left-align titles at top-left

NOTE! These are directly taken from Will Chase's rstudio::conf2020 talk:
<https://www.williamrchase.com/slides/assets/player/KeynoteDHTMLPlayer.html>

The Glamour of Graphics



Source: Will Chase's rstudio::conf 2020 talk

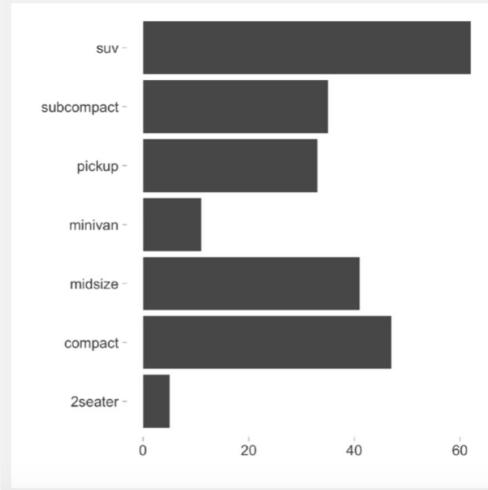
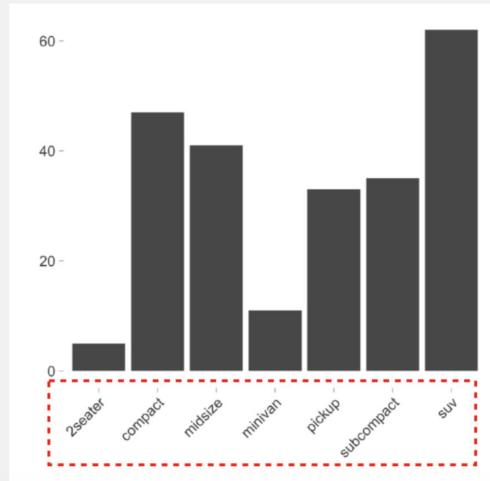
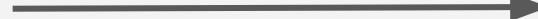
Left-Align Titles



Good Figures: Ground Rules

1. Left-align titles at top-left
2. Don't make people **tilt** their head

The Glamour of Graphics



Source: Will Chase's rstudio::conf 2020 talk

Avoid Head Tilting

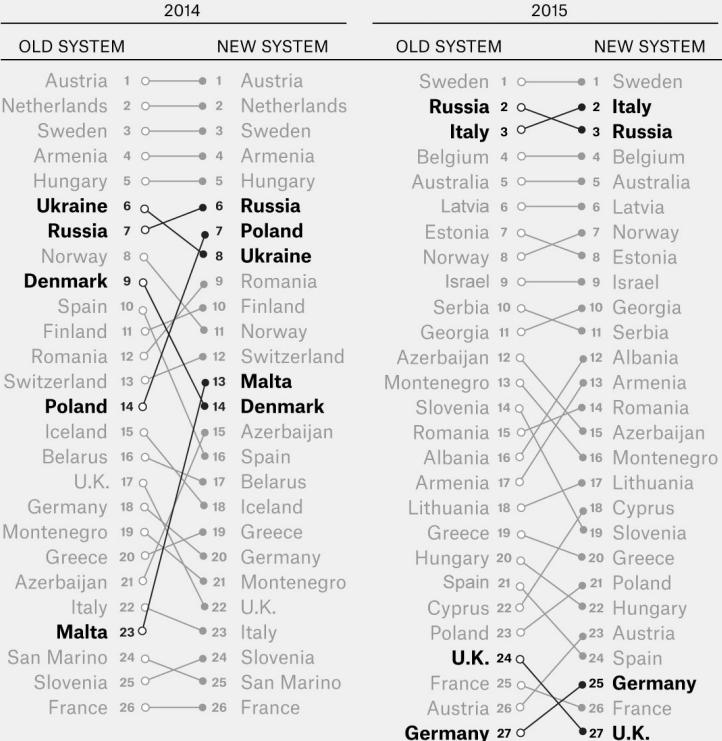


Good Figures: Ground Rules

1. Left-align titles at top-left
2. Don't make people tilt their head
3. Alignment should create **clean lines & symmetry**

The Glamour of Graphics

How the Eurovision rankings would change under the new scoring system



FIVETHIRTYEIGHT

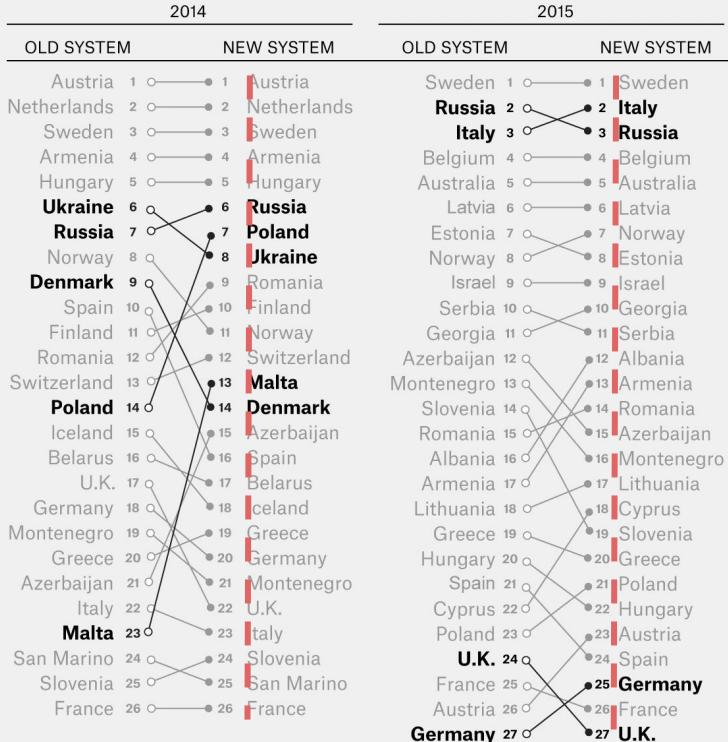
SOURCE: EUROVISION

Symmetric & Clean

The Glamour of Graphics

Left-align most text

How the Eurovision rankings would change under the new scoring system



FIVETHIRTYEIGHT

SOURCE: EUROVISION

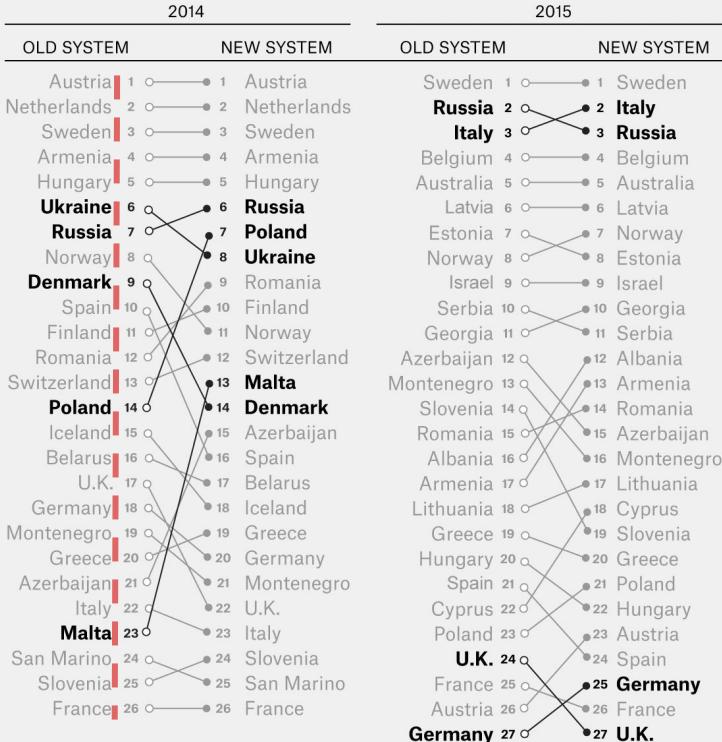
Symmetric & Clean

The Glamour of Graphics

Left-align most text

Some labels can be right-aligned

How the Eurovision rankings would change under the new scoring system



FIVETHIRTYEIGHT

SOURCE: EUROVISION

Symmetric & Clean

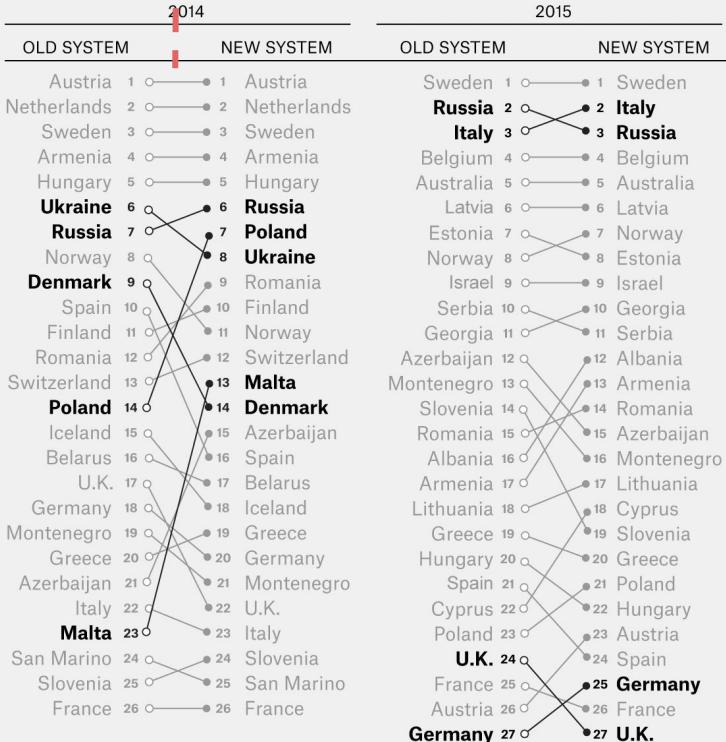
The Glamour of Graphics

Left-align most text

Some labels can be right-aligned

Some subtitles and labels can be center-aligned

How the Eurovision rankings would change under the new scoring system



FIVETHIRTYEIGHT

SOURCE: EUROVISION

Symmetric & Clean

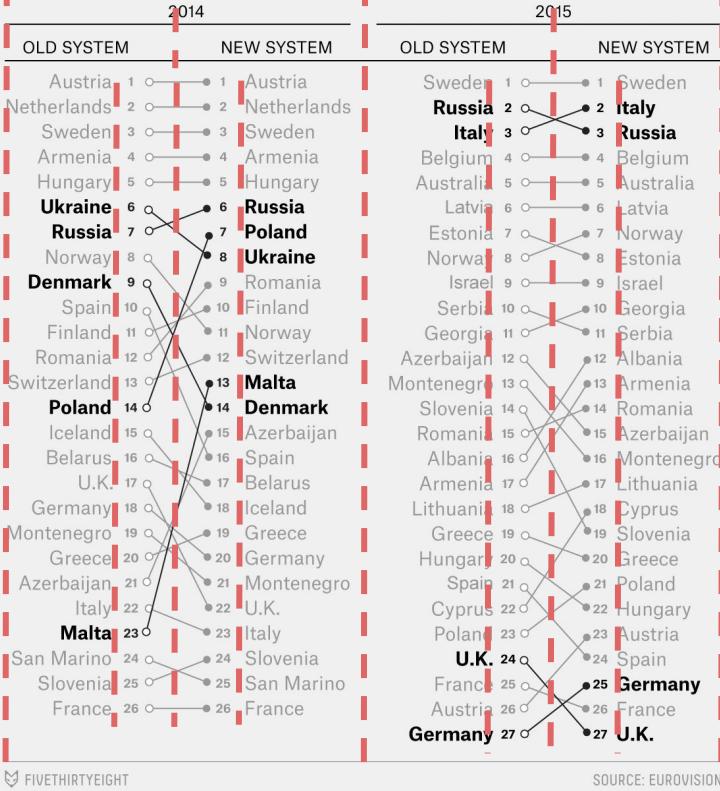
The Glamour of Graphics

Left-align most text

Some labels can be right-aligned

Some subtitles and labels can be center-aligned

How the Eurovision rankings would change under the new scoring system



Symmetric & Clean



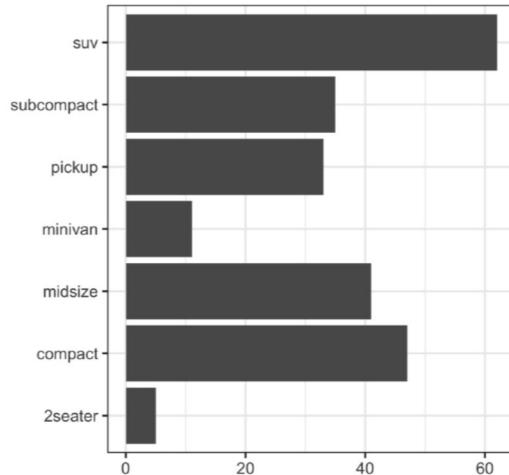
Good Figures: Ground Rules

1. Left-align titles at top-left
2. Don't make people tilt their head
3. Alignment should create clean lines & symmetry
4. ~~Borders~~ - don't use them

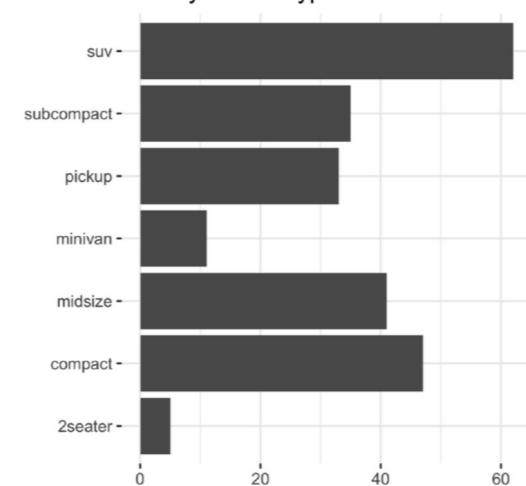
The Glamour of Graphics



Number of cars by vehicle type



Number of cars by vehicle type



Source: Will Chase's rstudio::conf 2020 talk

Remove Borders



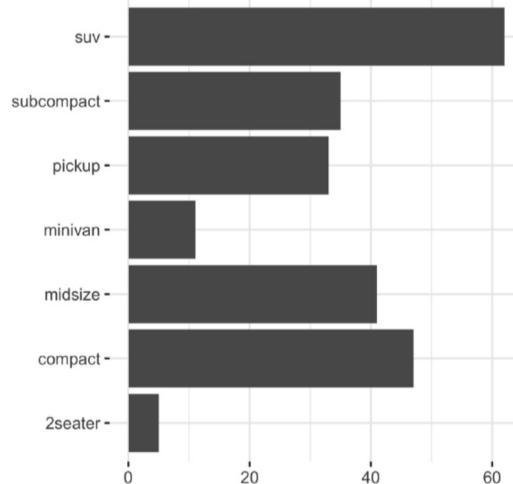
Good Figures: Ground Rules

1. Left-align titles at top-left
2. Don't make people tilt their head
3. Alignment should create clean lines & symmetry
4. ~~Borders~~ - don't use them
5. Remove and lighten as much as possible

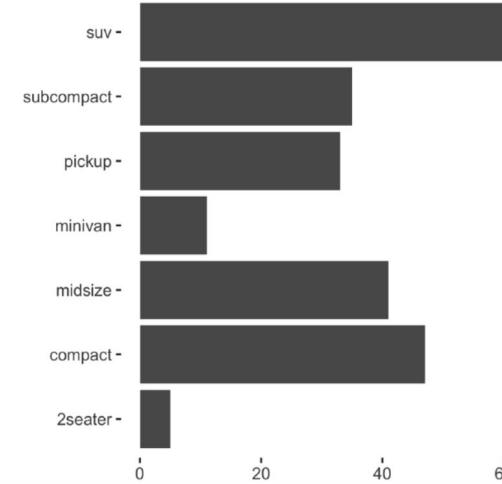
The Glamour of Graphics



Number of cars by vehicle type



Number of cars by vehicle type



Source: Will Chase's rstudio::conf 2020 talk

Remove & Lighten

less
is more

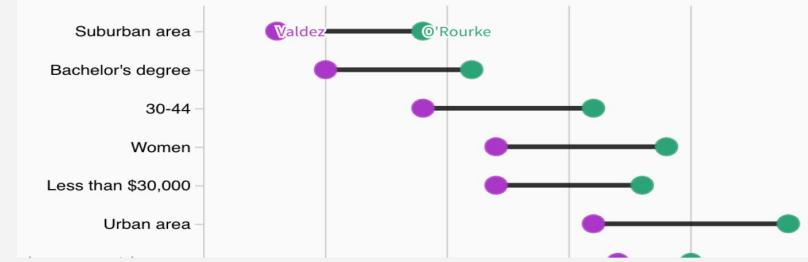
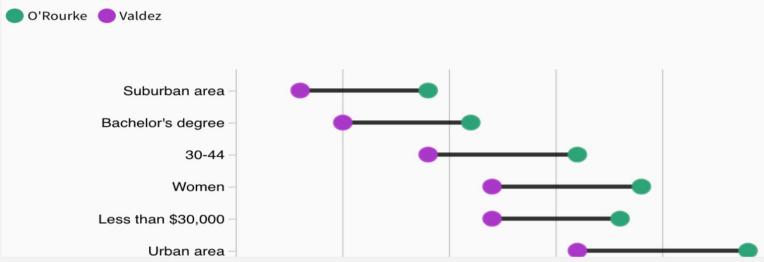
(effective)
(attractive)
(impactive)



Good Figures: Ground Rules

1. Left-align titles at top-left
2. Don't make people tilt their head
3. Alignment should create clean lines & symmetry
4. ~~Borders~~ - don't use them
5. Remove and lighten as much as possible
6. **Legends suck**

The Glamour of Graphics



Source: <https://flourish.studio/2018/11/15/legend-colors-in-header/>

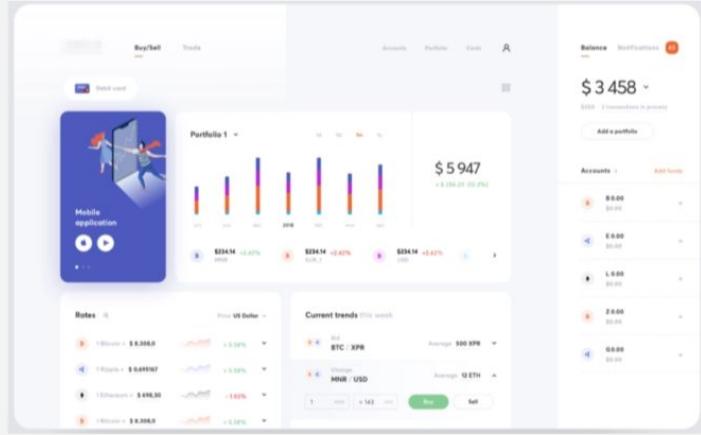
Legends Suck



Good Figures: Ground Rules

1. Left-align titles at top-left
2. Don't make people tilt their head
3. Alignment should create clean lines & symmetry
4. ~~Borders~~ - don't use them
5. Remove and lighten as much as possible
6. Legends suck
7. White space is like garlic - take the amount you need and triple it

The Glamour of Graphics



Source: Will Chase's rstudio::conf 2020 talk

Garlicky Whitespace



Good Figures: Ground Rules

1. Left-align titles at top-left
2. Don't make people tilt their head
3. Alignment should create clean lines & symmetry
4. ~~Borders~~ - don't use them
5. Remove and lighten as much as possible
6. Legends suck
7. White space is like garlic - take the amount you need and triple it

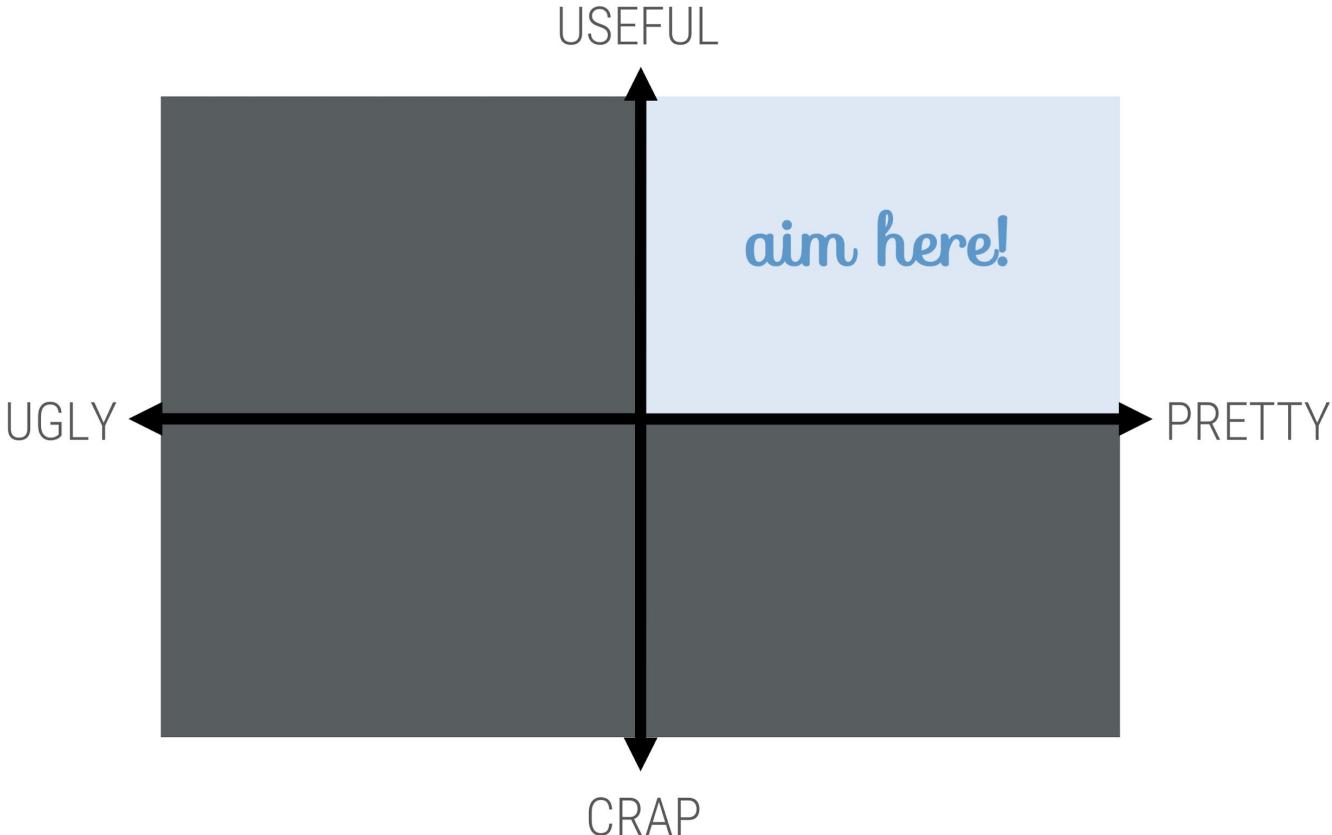
NOTE! These are directly taken from Will Chase's rstudio::conf2020 talk:
<https://www.williamrchase.com/slides/assets/player/KeynoteDHTMLPlayer.html>

Iteration

Take a sad plot &
make it better

Inspiration: Allison Hill & <https://policyviz.com/helpmeviz/>

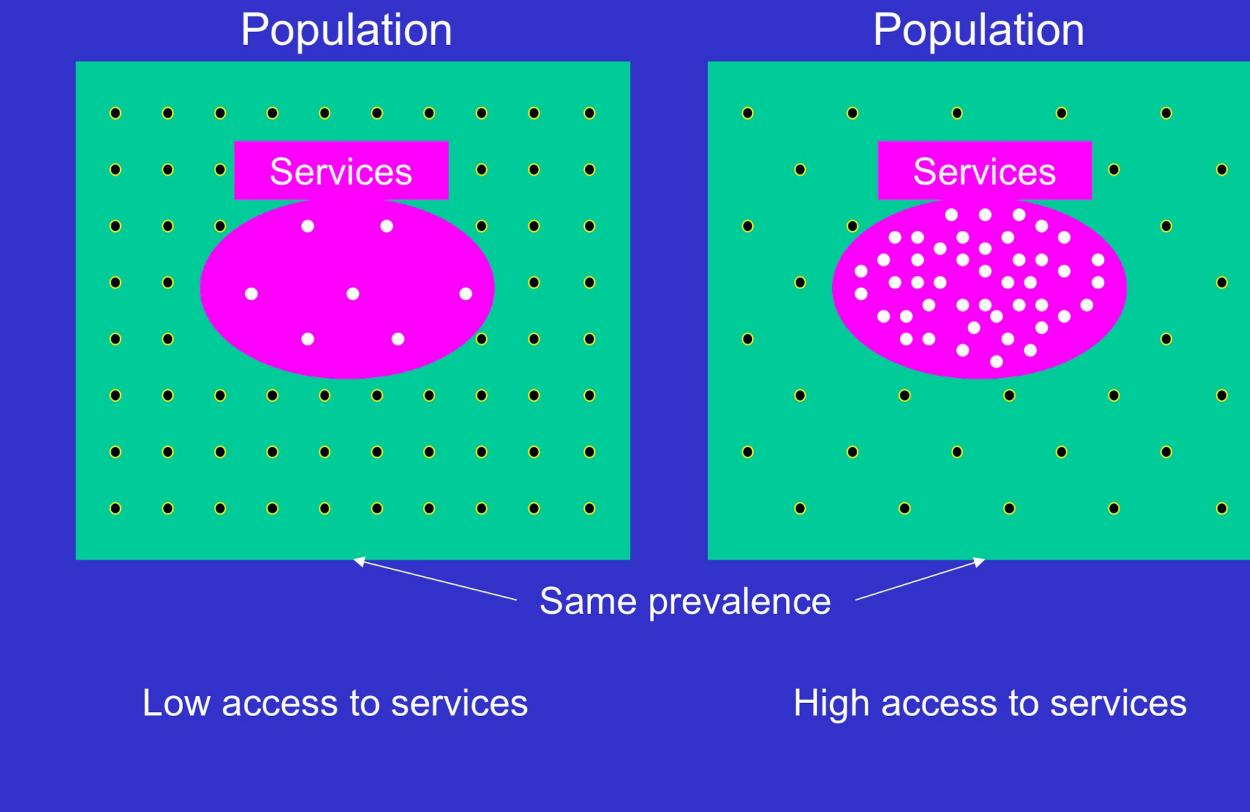
Iteration



Source: Jackie Wirz

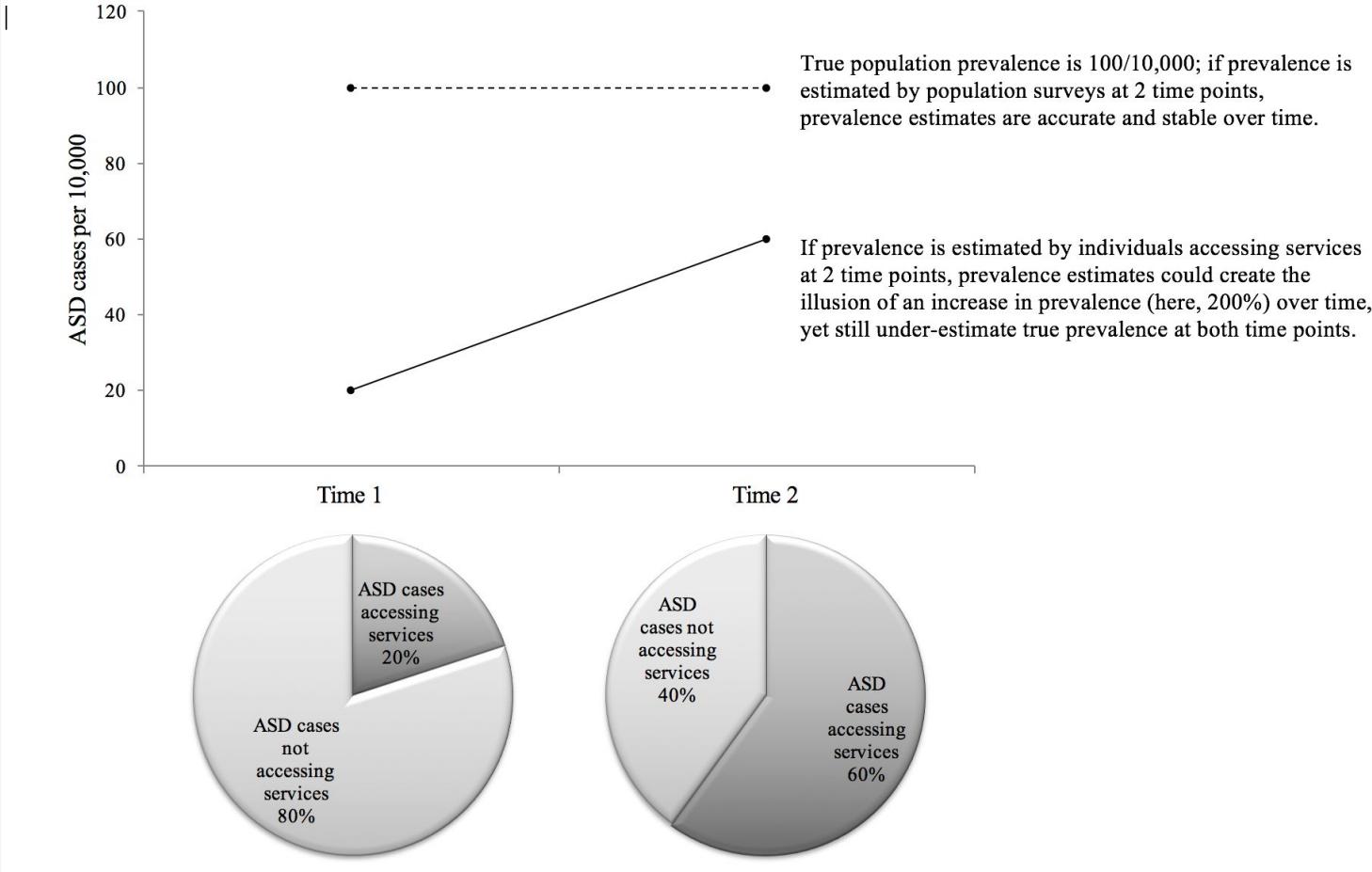
Iteration

Prevalence and access to services

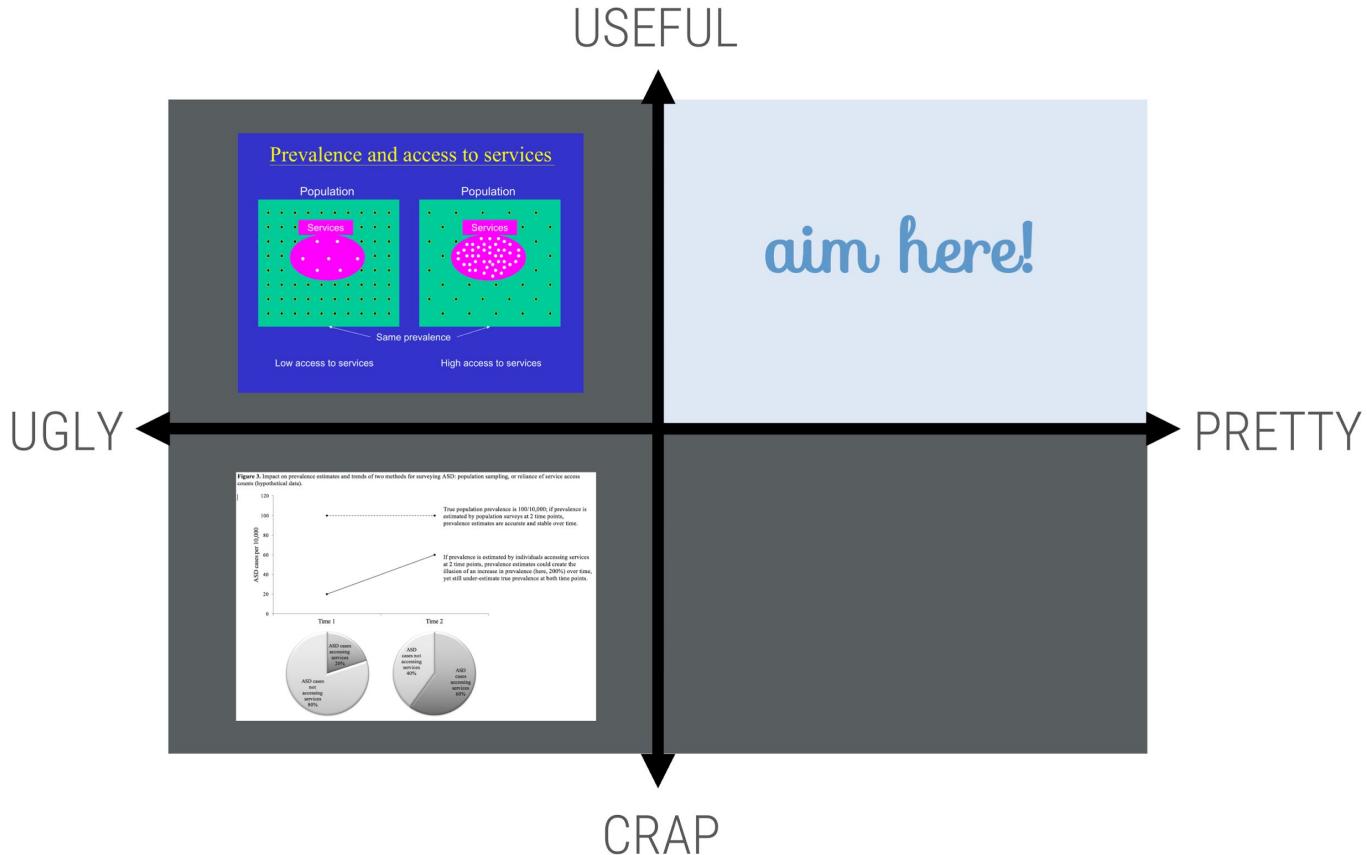


Iteration

Figure 3. Impact on prevalence estimates and trends of two methods for surveying ASD: population sampling, or reliance of service access counts (hypothetical data).

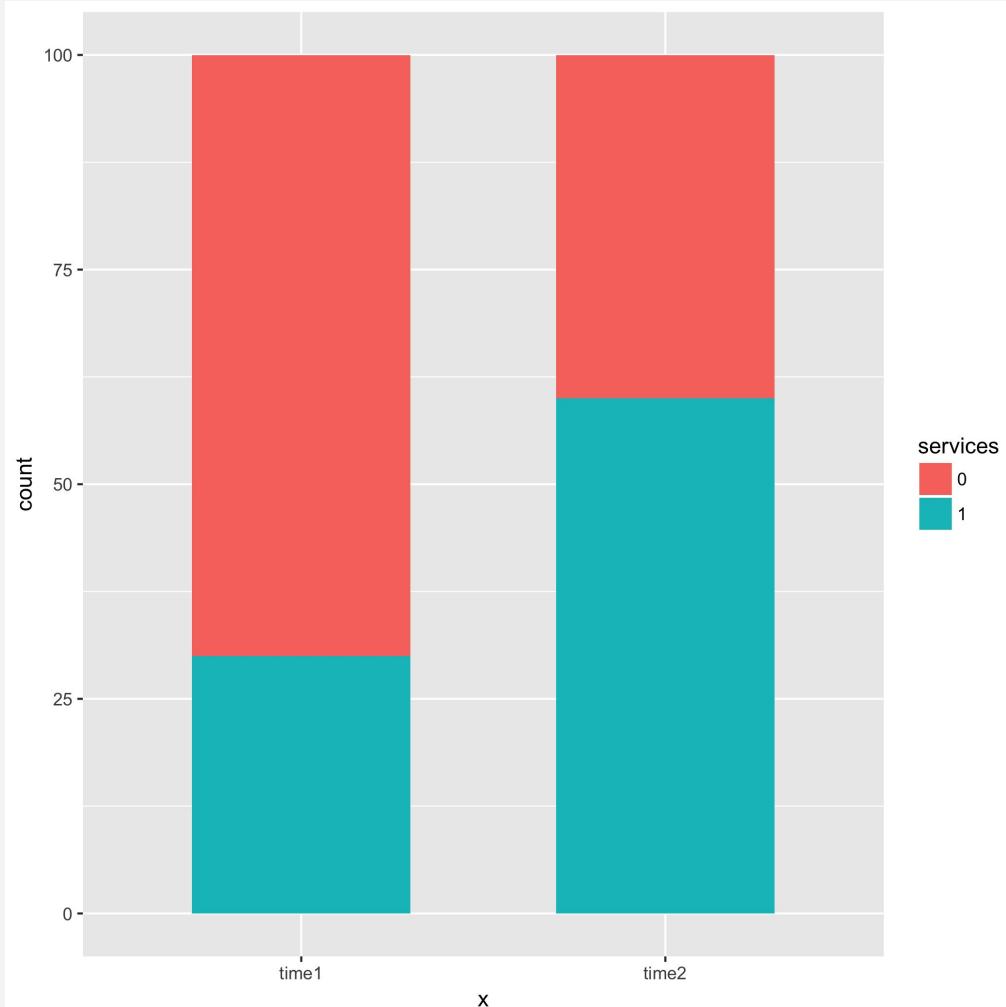


Iteration



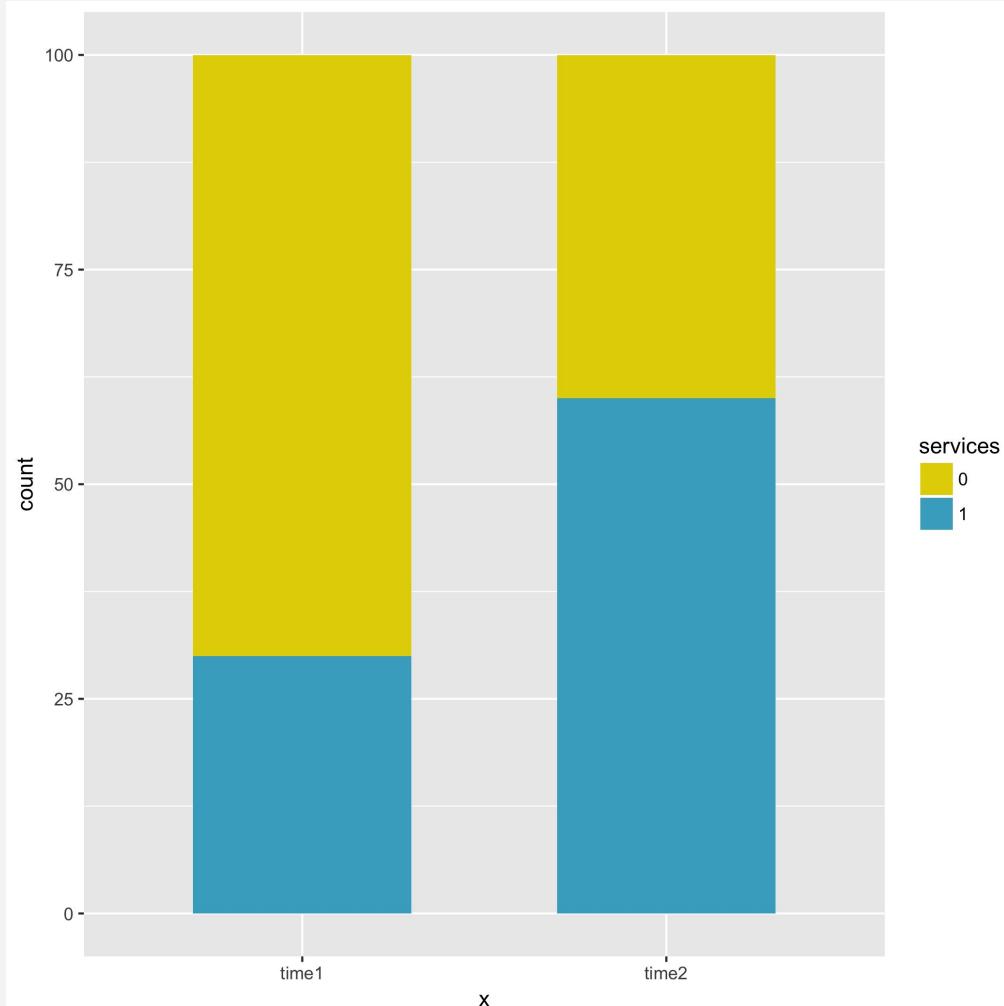
Source: Jackie Wirz

Iteration



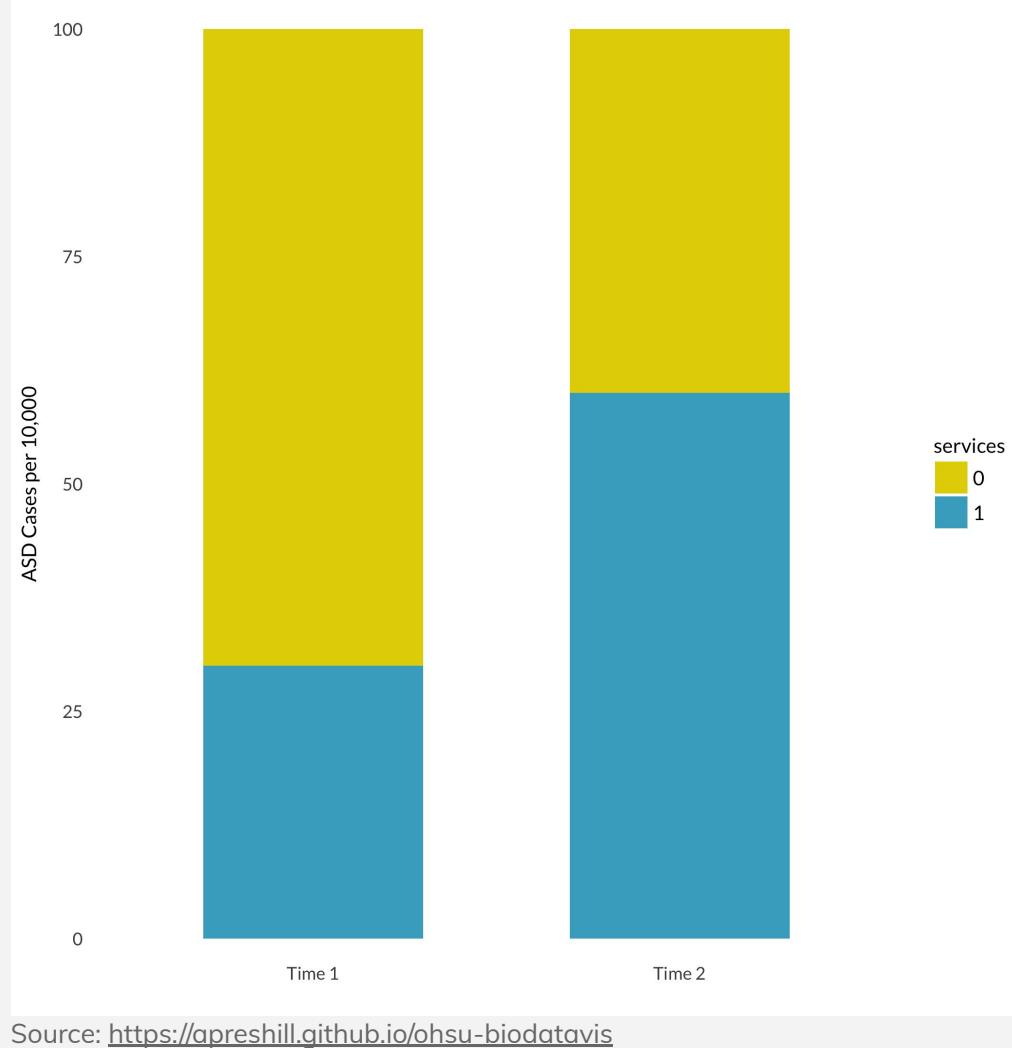
Source: <https://apreshill.github.io/ohsu-biodatavis>

Iteration

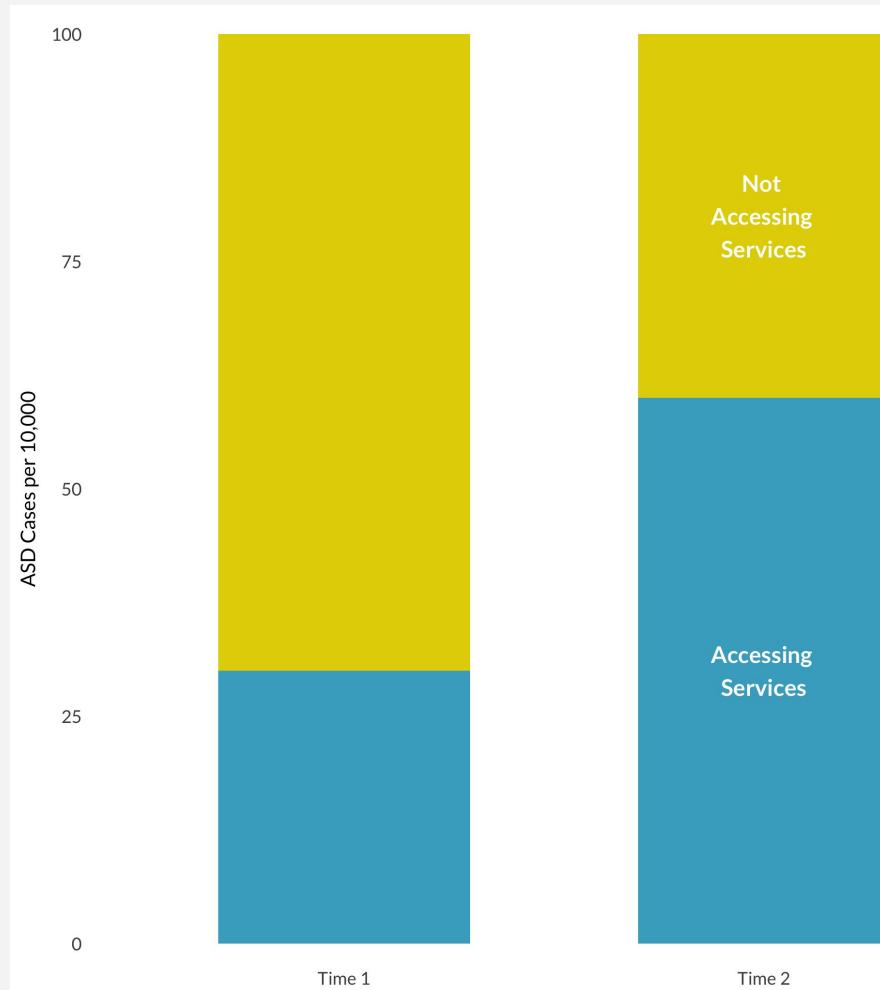


Source: <https://apreshill.github.io/ohsu-biodatavis>

Iteration

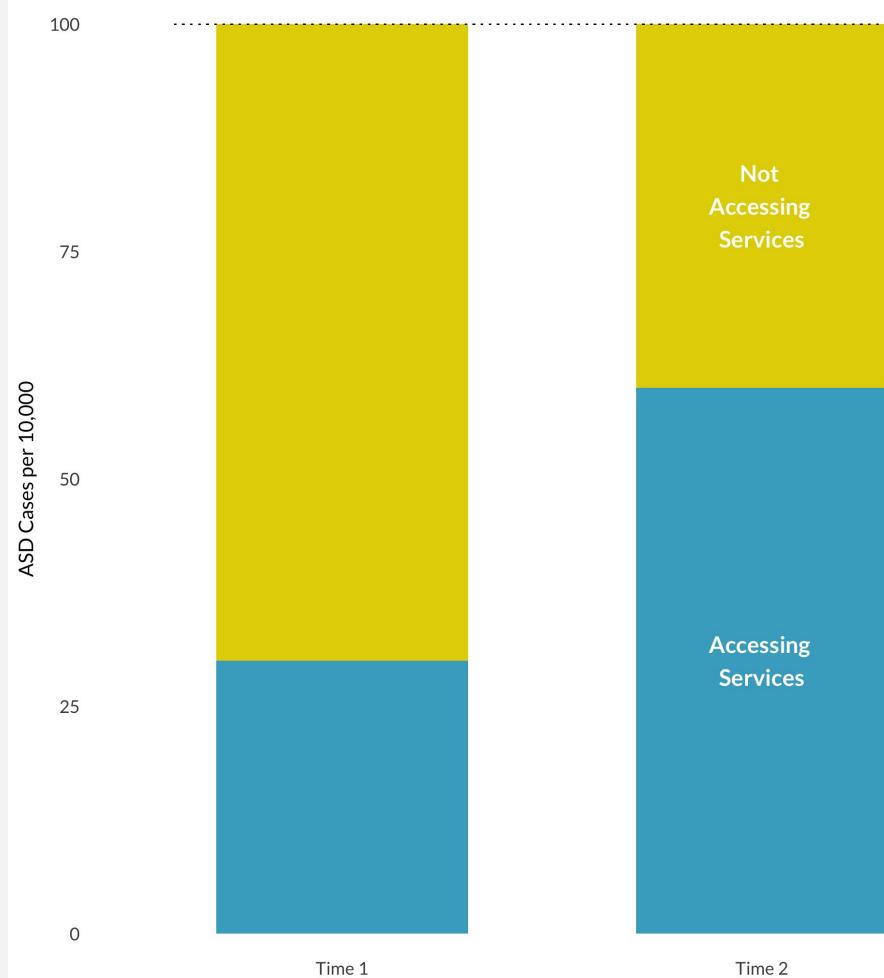


Iteration



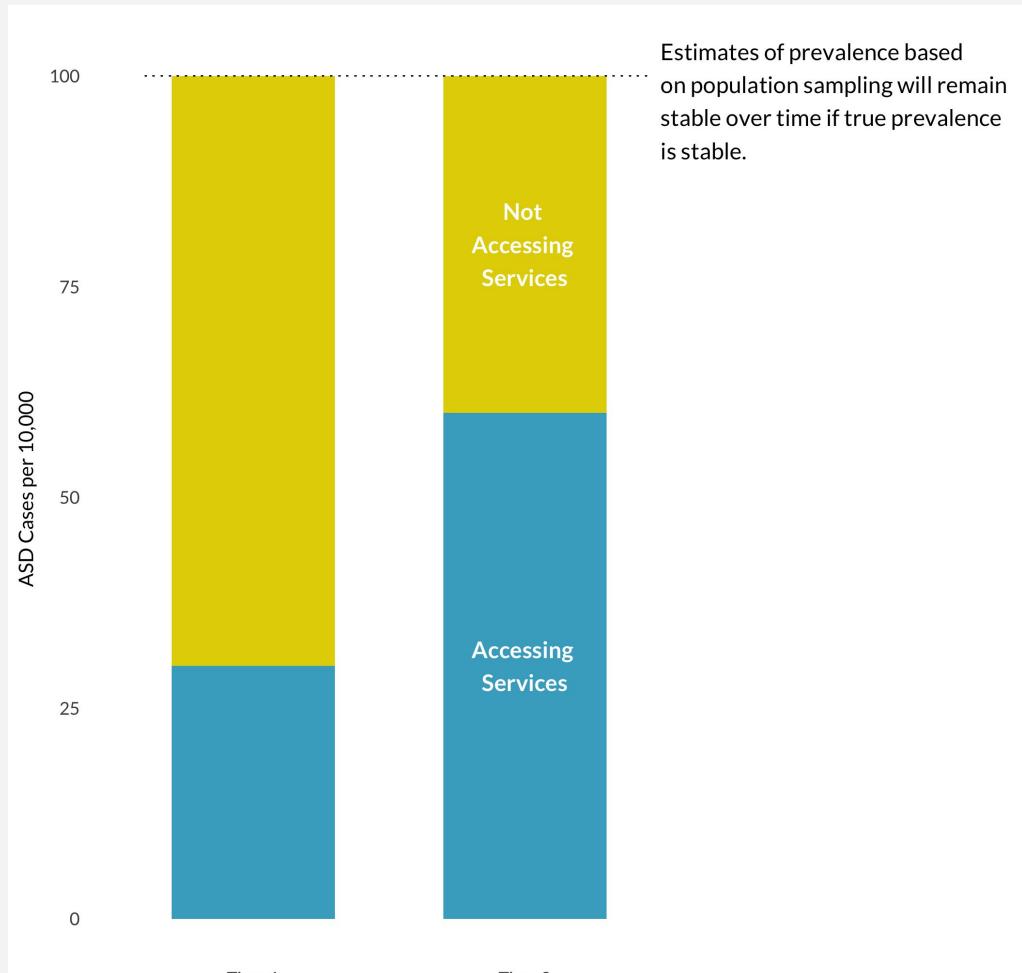
Source: <https://apreshill.github.io/ohsu-biodatavis>

Iteration



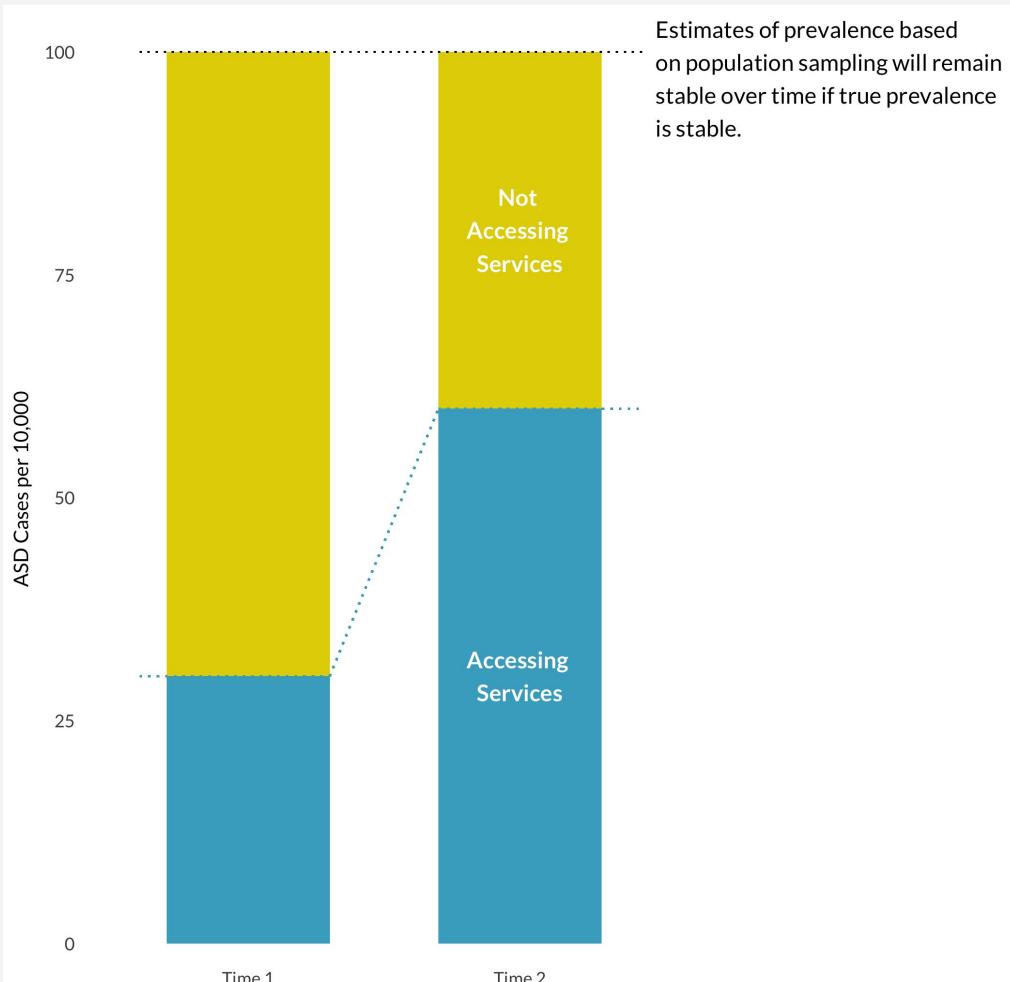
Source: <https://apreshill.github.io/ohsu-biodatavis>

Iteration



Source: <https://apreshill.github.io/ohsu-biodatavis>

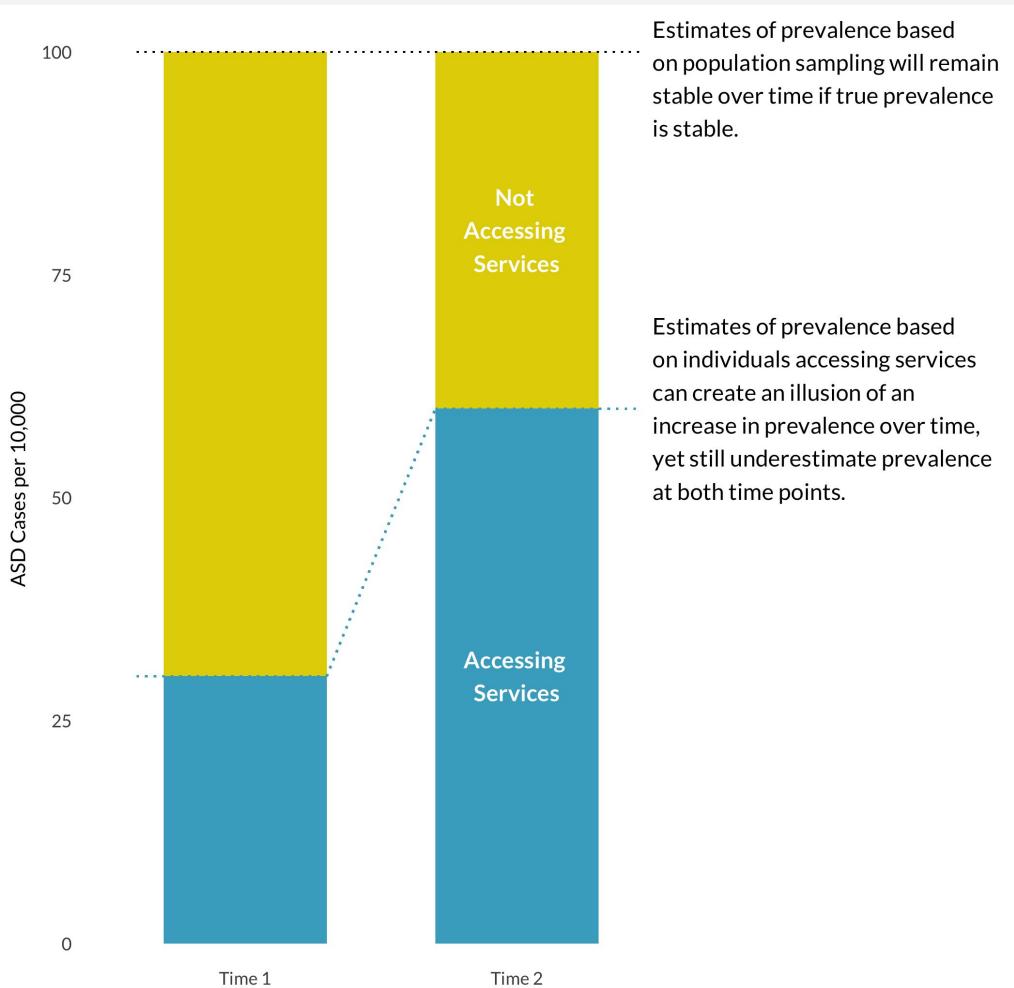
Iteration



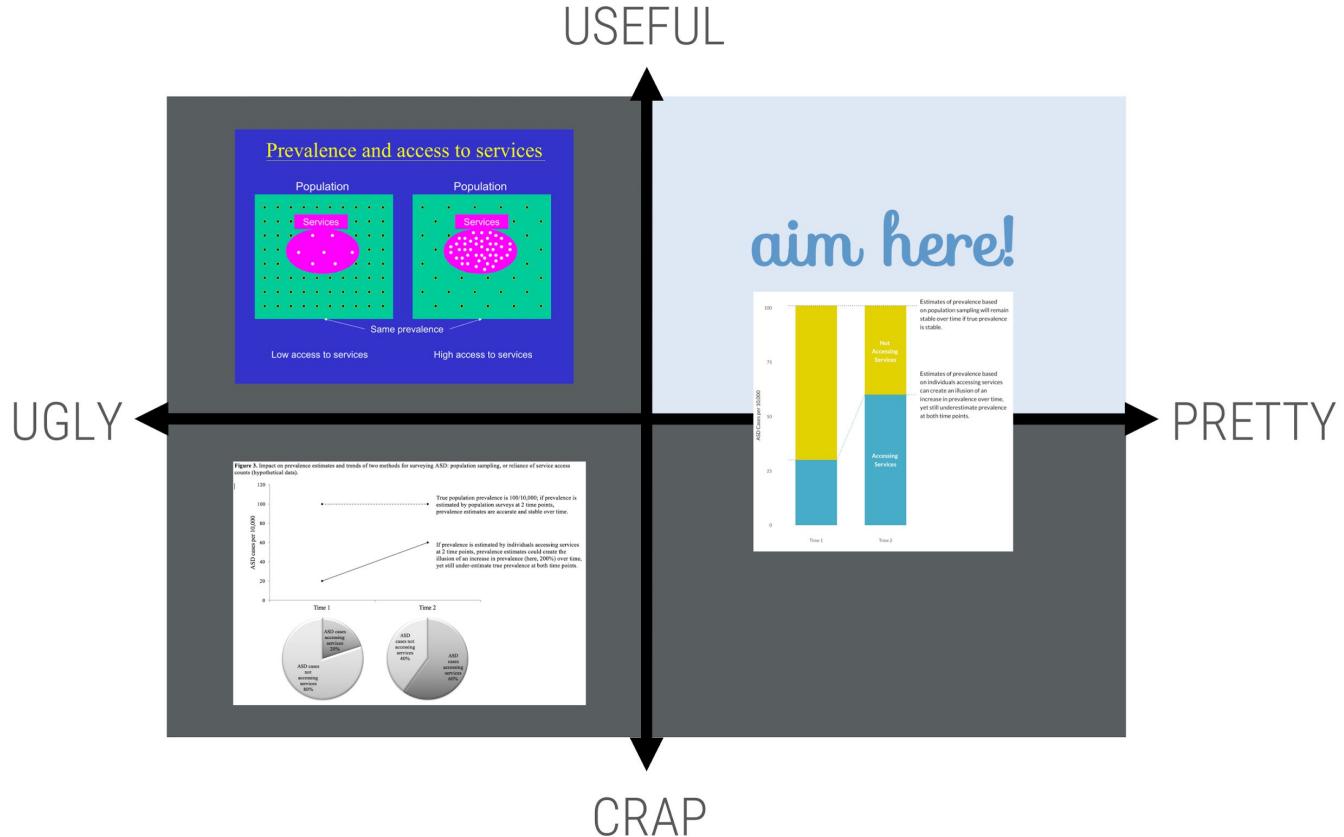
Source: <https://apreshill.github.io/ohsu-biodatavis>

Estimates of prevalence based on population sampling will remain stable over time if true prevalence is stable.

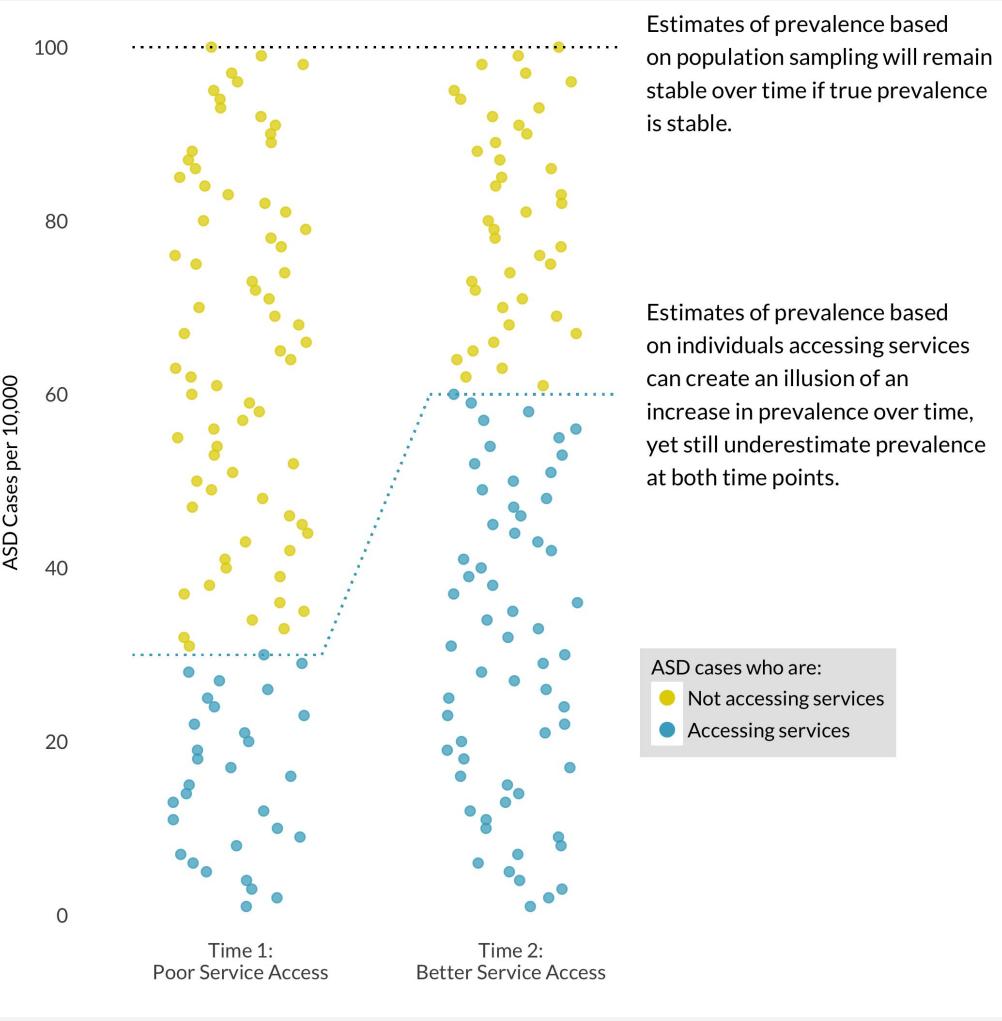
Iteration



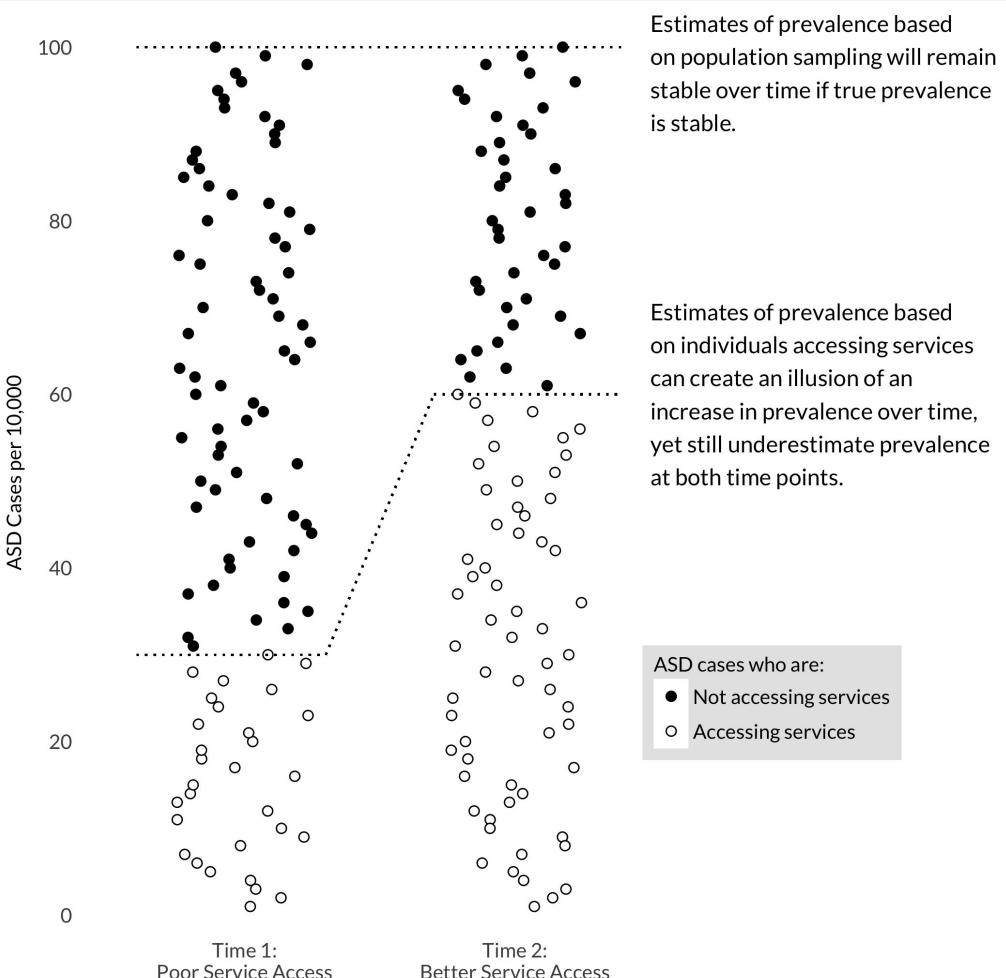
Activity



Iteration

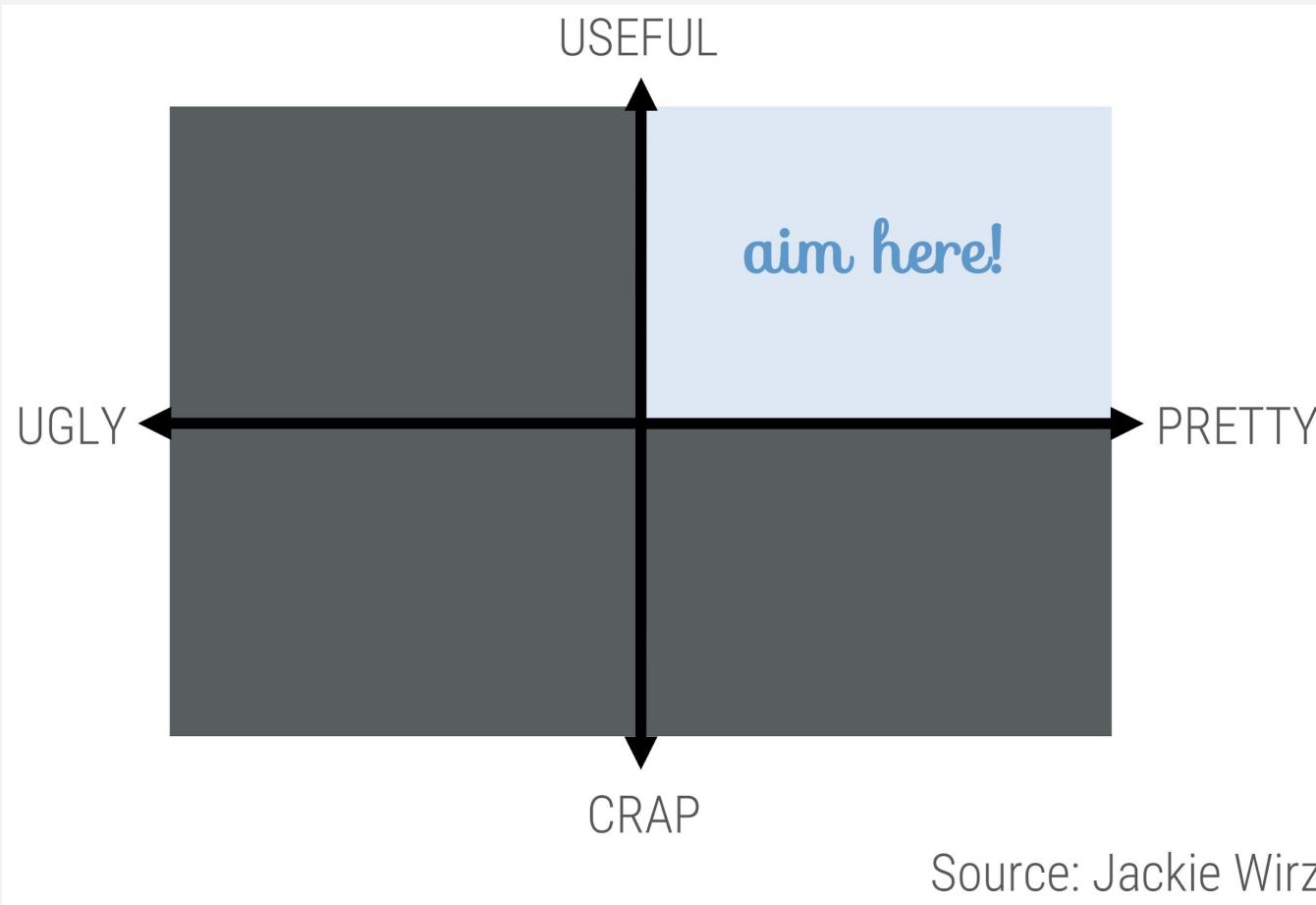


Iteration

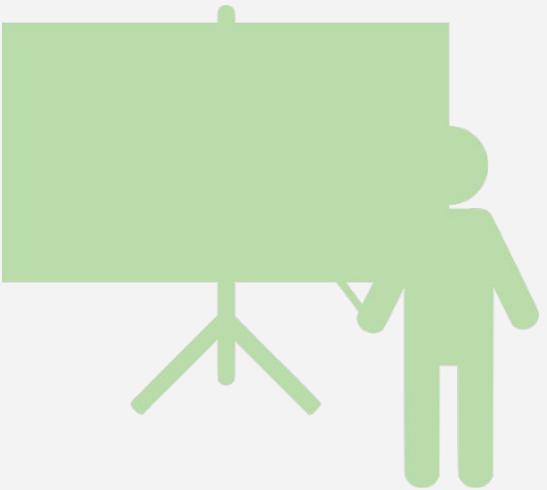


Iteration

Your turn: Take a sad plot & make it better



Source: Jackie Wirz

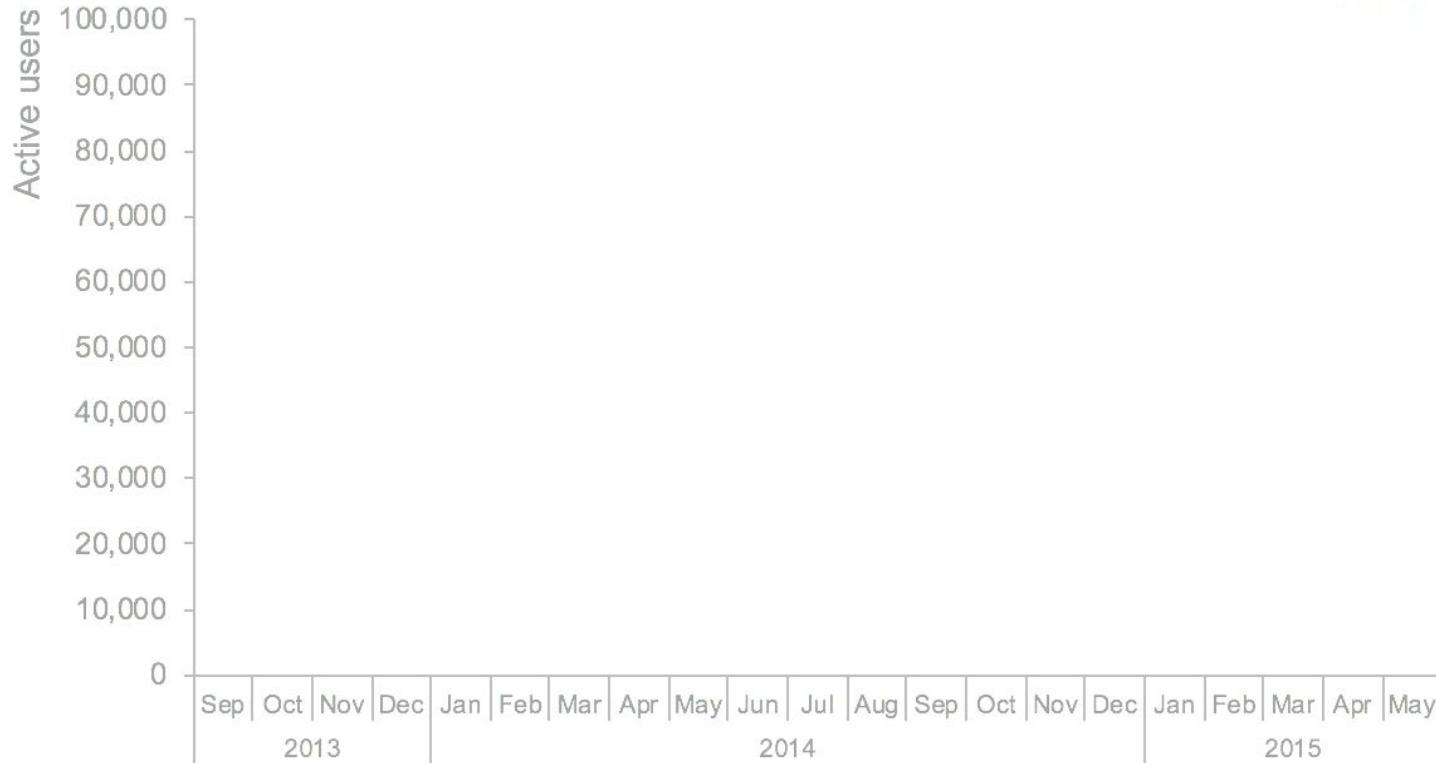


Presentations: for listening

- don't read directly off slides
- use animation to build your story (not to distract)
- introduce your axes
- benefit: words to explain out loud what you're showing

Your Audience

Moonville: active users over time



Data source: ABC Report. For purpose of analysis "active user" is defined as the number of unique users in the past 30 days.

Your Audience

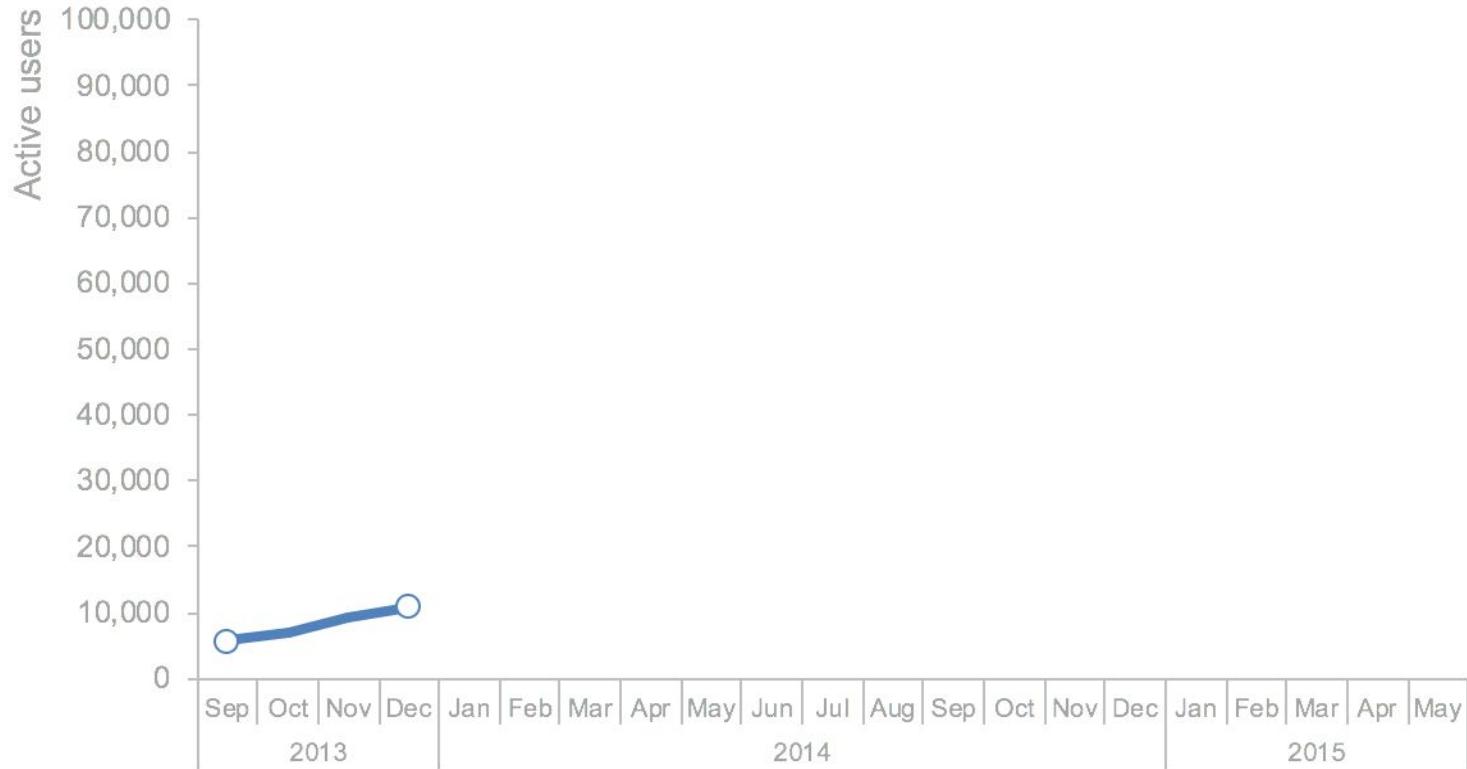
Moonville: active users over time



Data source: ABC Report. For purpose of analysis "active user" is defined as the number of unique users in the past 30 days.

Your Audience

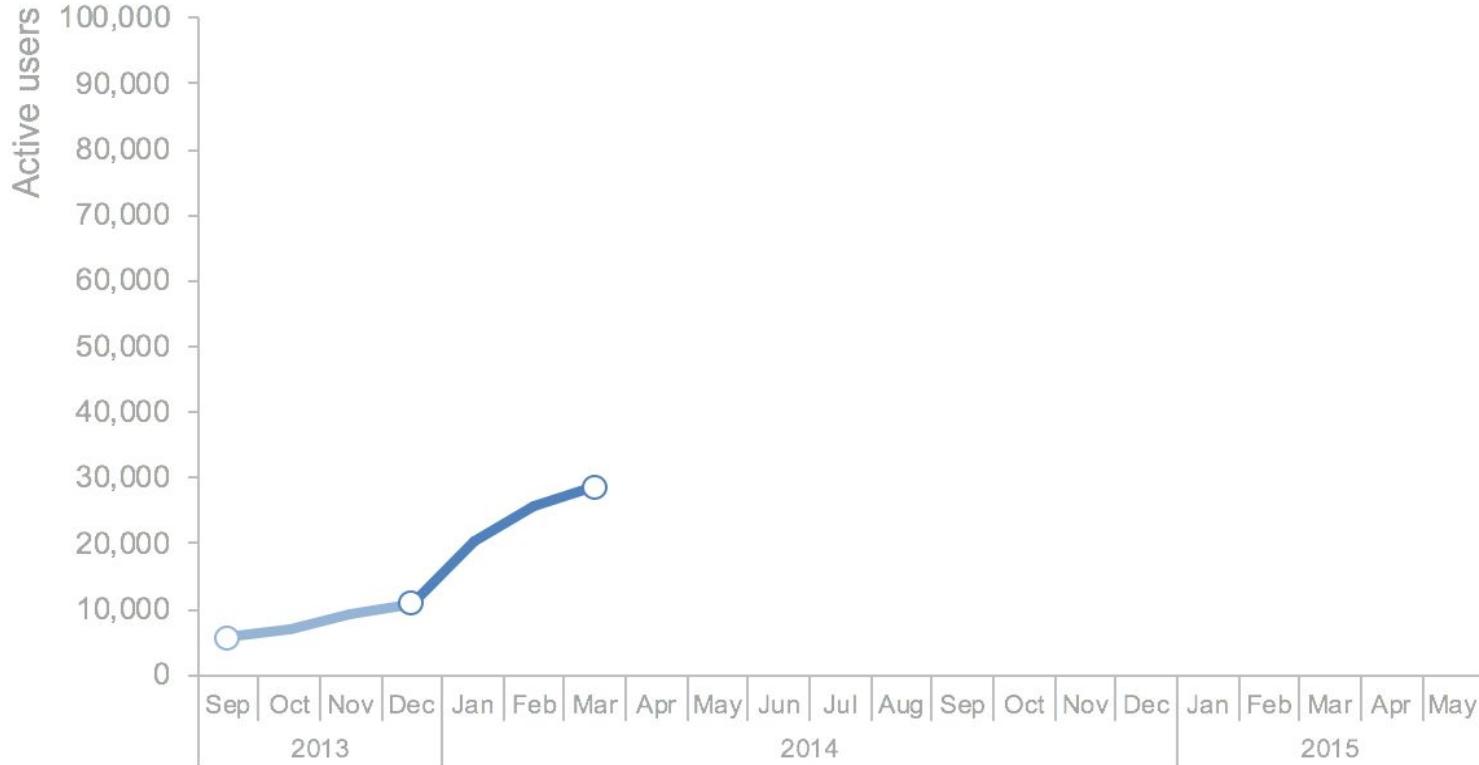
Moonville: active users over time



Data source: ABC Report. For purpose of analysis "active user" is defined as the number of unique users in the past 30 days.

Your Audience

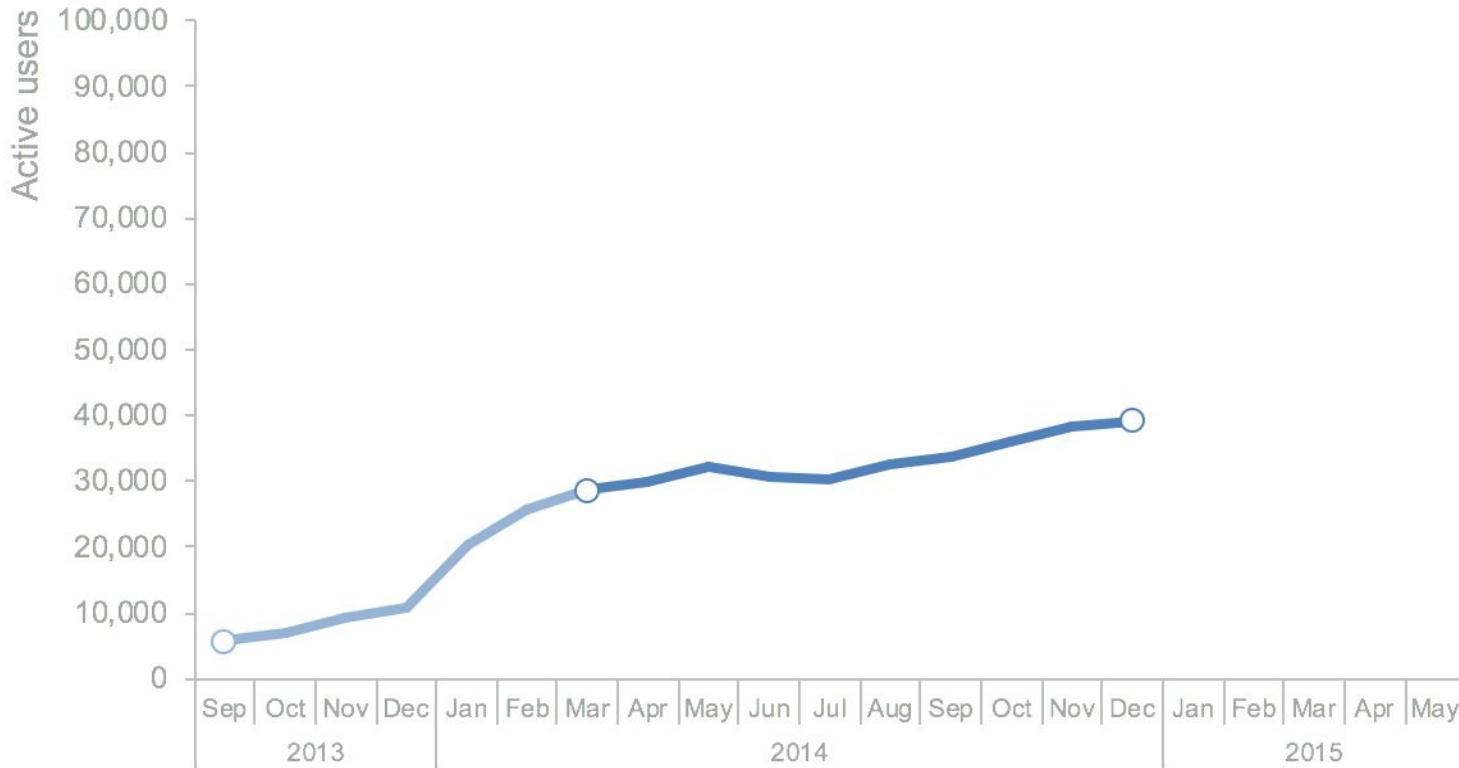
Moonville: active users over time



Data source: ABC Report. For purpose of analysis "active user" is defined as the number of unique users in the past 30 days.

Your Audience

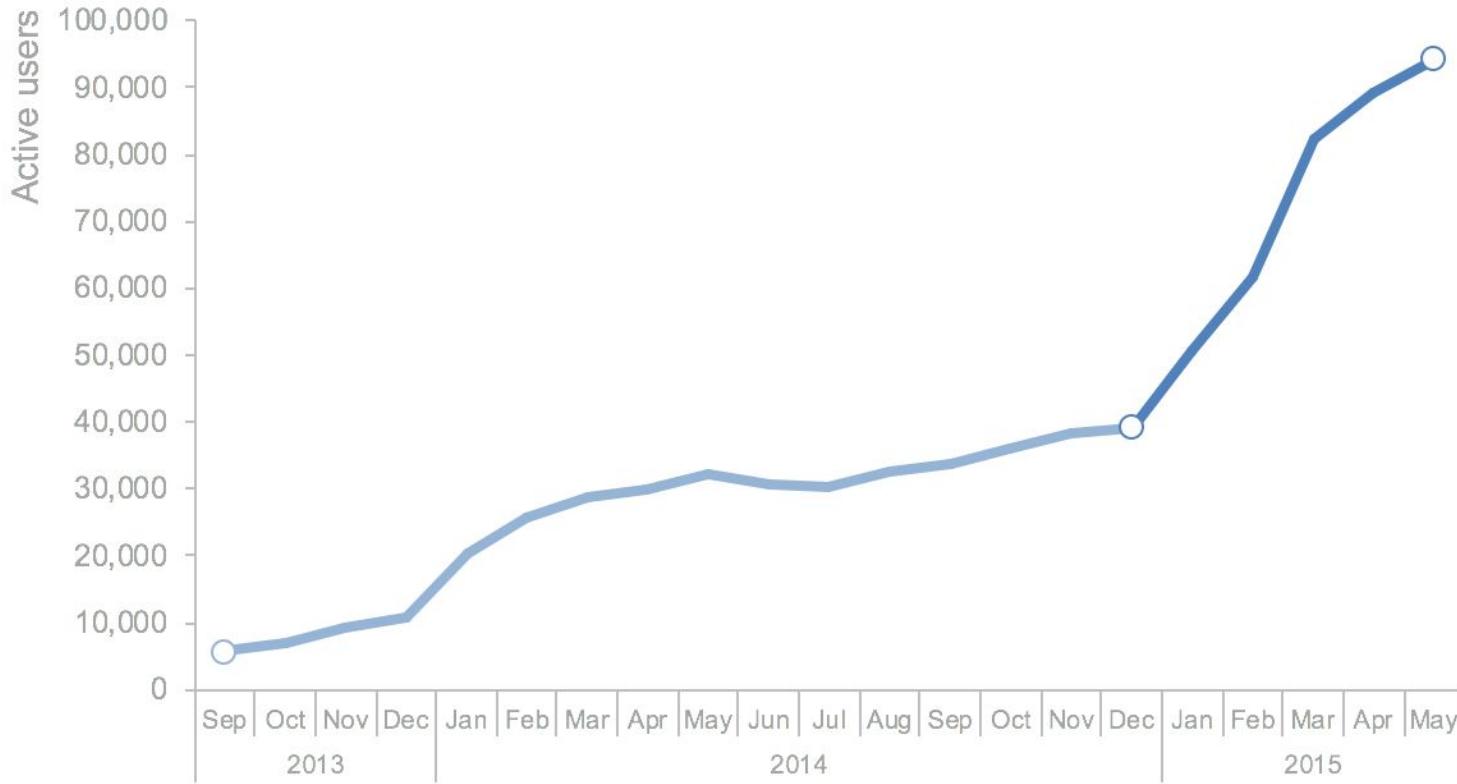
Moonville: active users over time



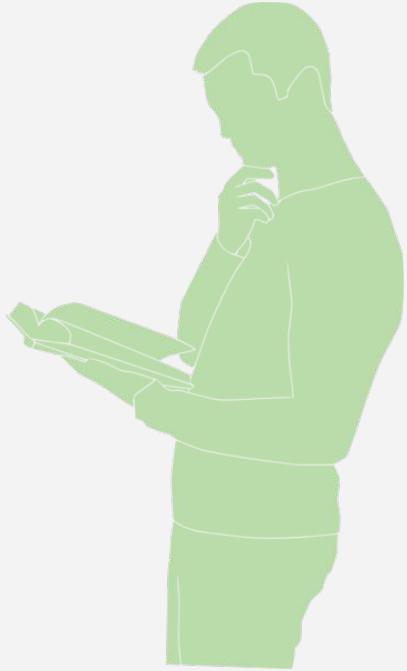
Data source: ABC Report. For purpose of analysis "active user" is defined as the number of unique users in the past 30 days.

Your Audience

Moonville: active users over time



Data source: ABC Report. For purpose of analysis "active user" is defined as the number of unique users in the past 30 days.

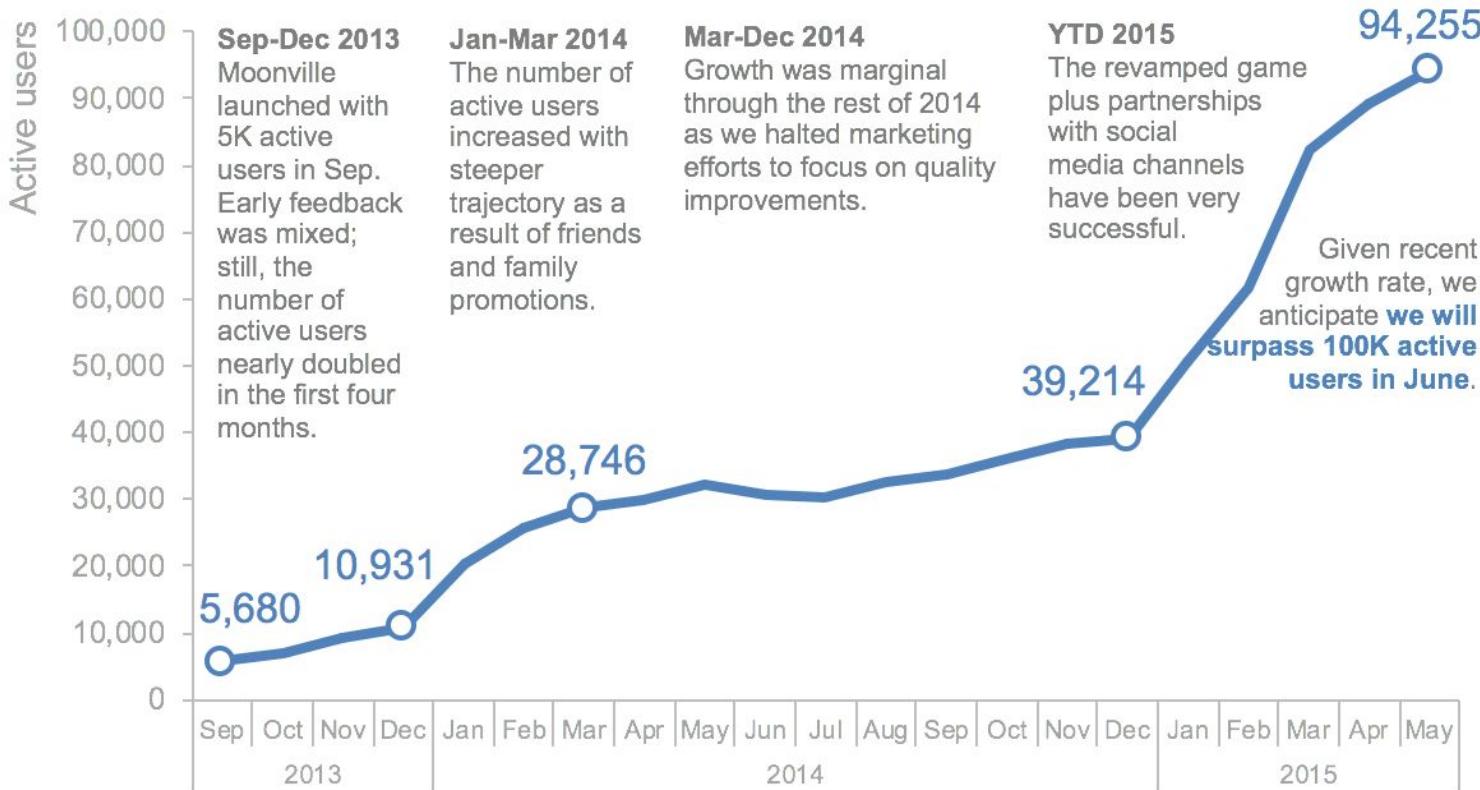


Reports: for reading

- more on a single visualization
- explanation must be there in text
- Benefit: people have time to look at what you've sent

Your Audience

Moonville: active users over time



Data source: ABC Report. For purpose of analysis "active user" is defined as the number of unique users in the past 30 days.

Checklist

Figures

- Have informative titles
- Appropriate figures used
- Adhering to 'less is more'
- Mappings clear (maybe a legend)
- Have consistent colors
- Take-home message clear
- Alignment all good
- Appropriate whitespace

Note: This **takes longer than you expect** it would. You can spend hours on a single figure. But, you worked really hard to understand the data and do the project. Now it's time to show it off! You get to do this through your presentation of your results - **spend the time to do this effectively.**