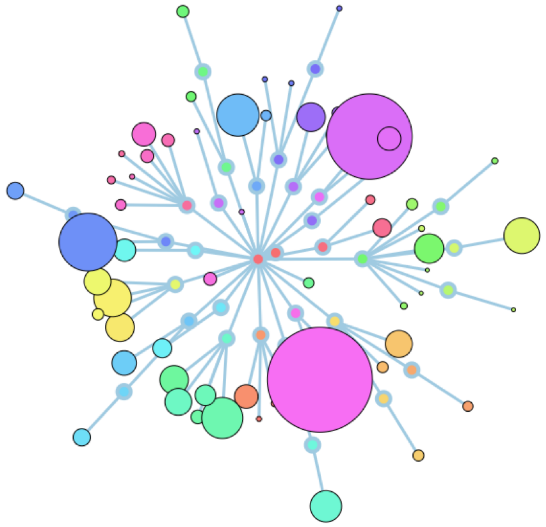
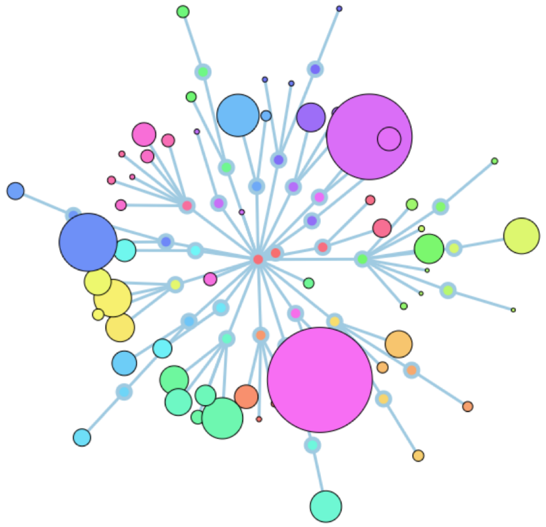
This is a talk about data viz

Jennifer Piscionere

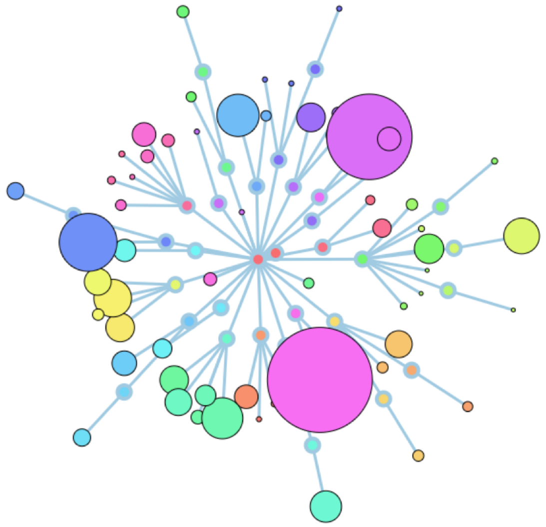
Data Scientist / Extronomer

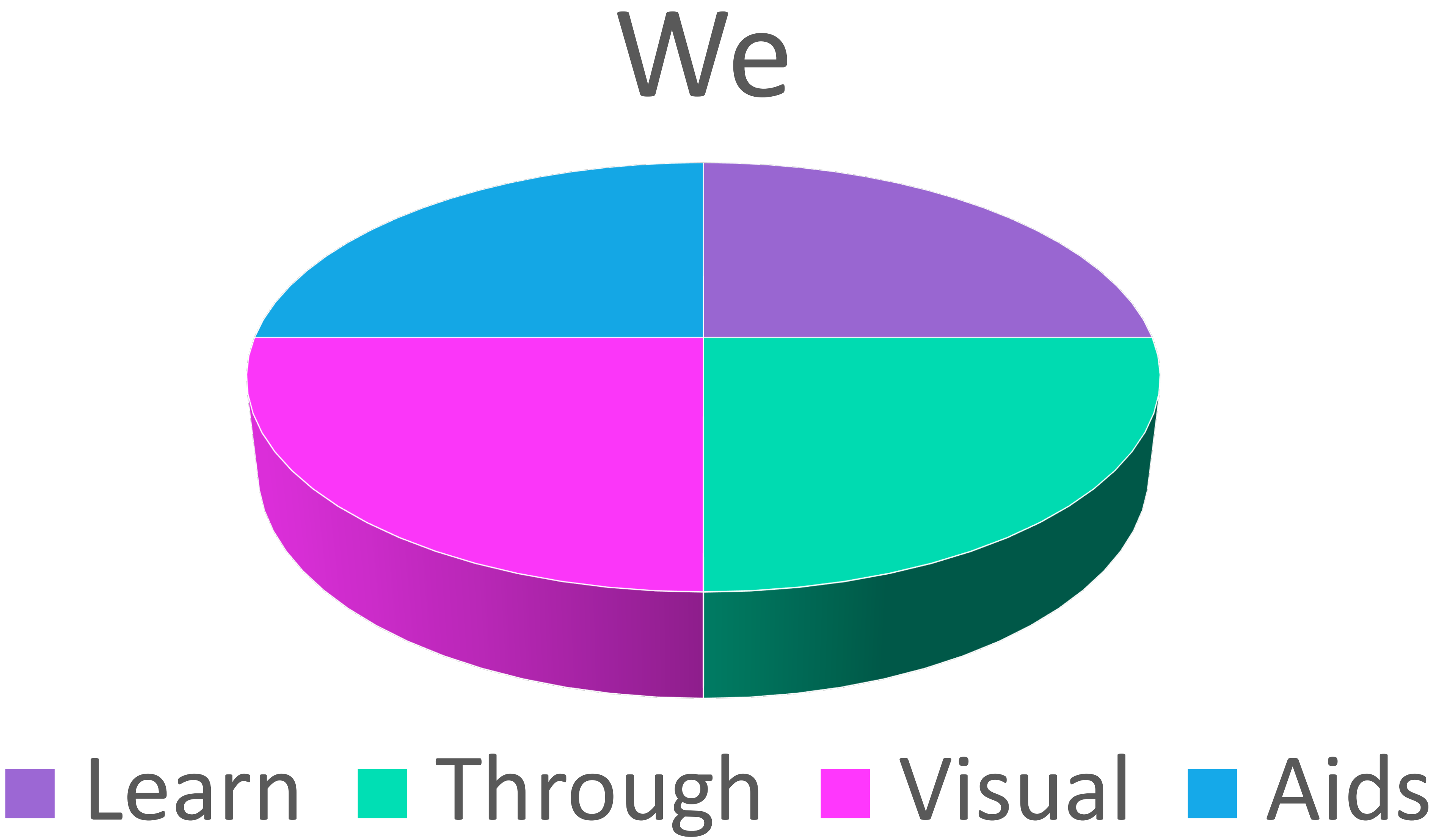


This is a talk about communication

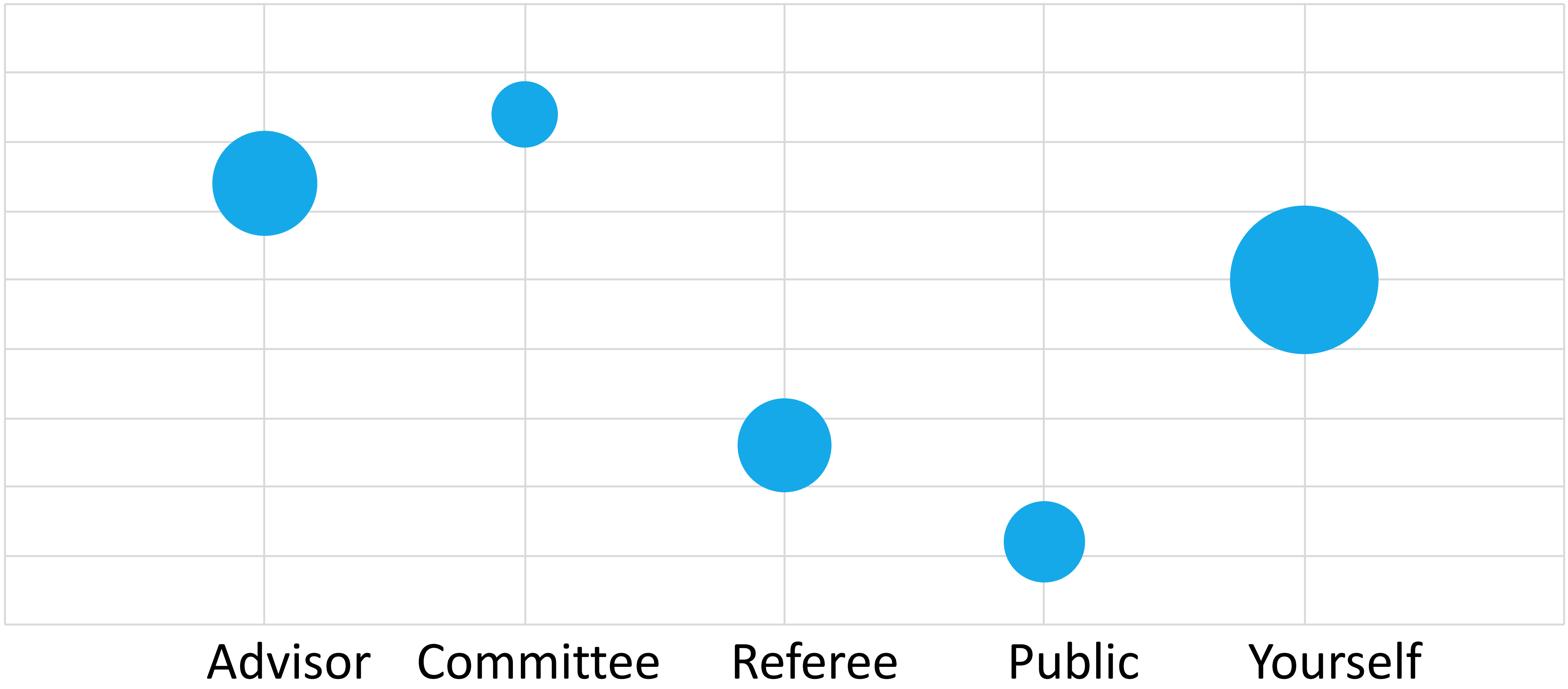


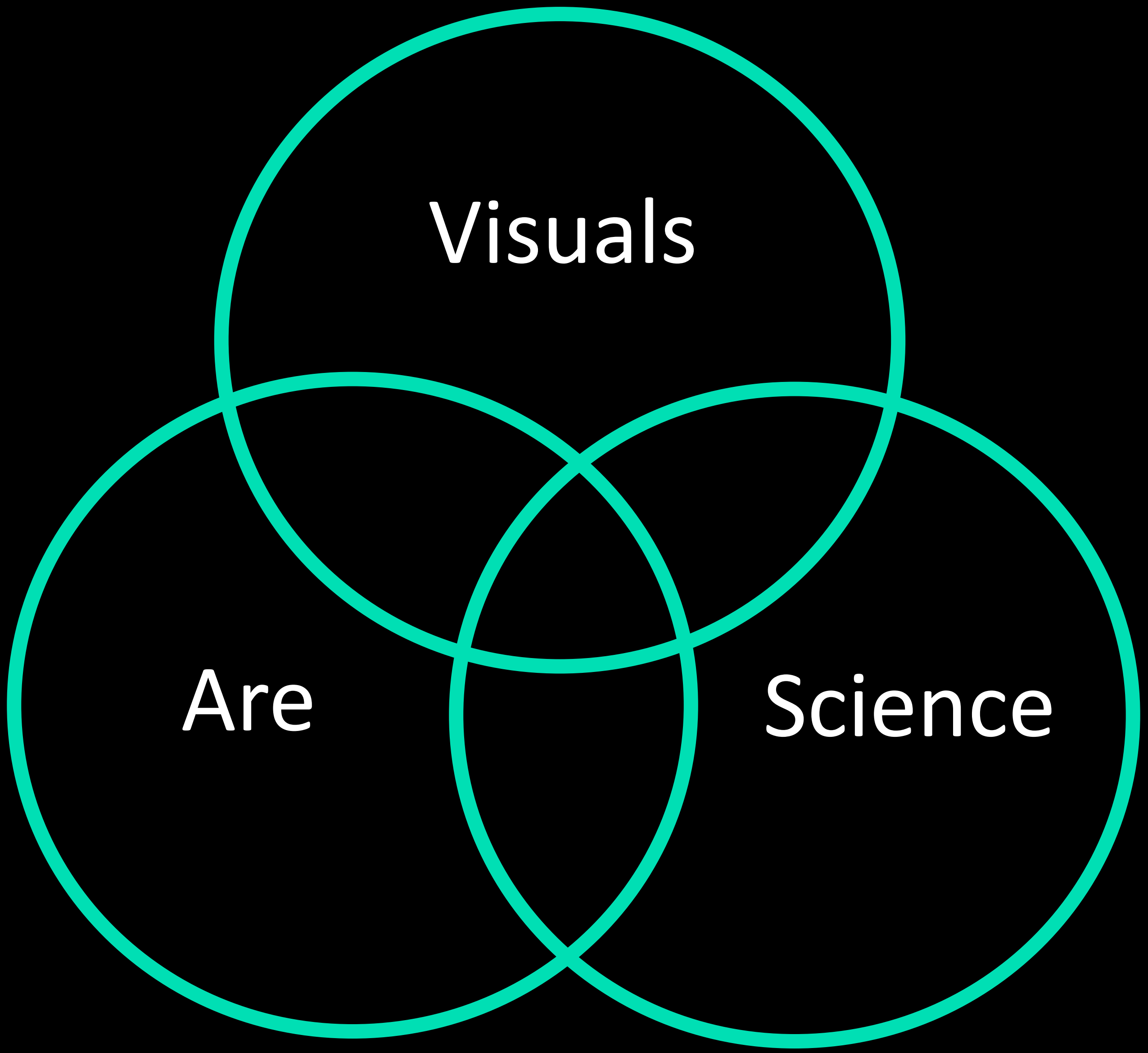
This is a talk about science

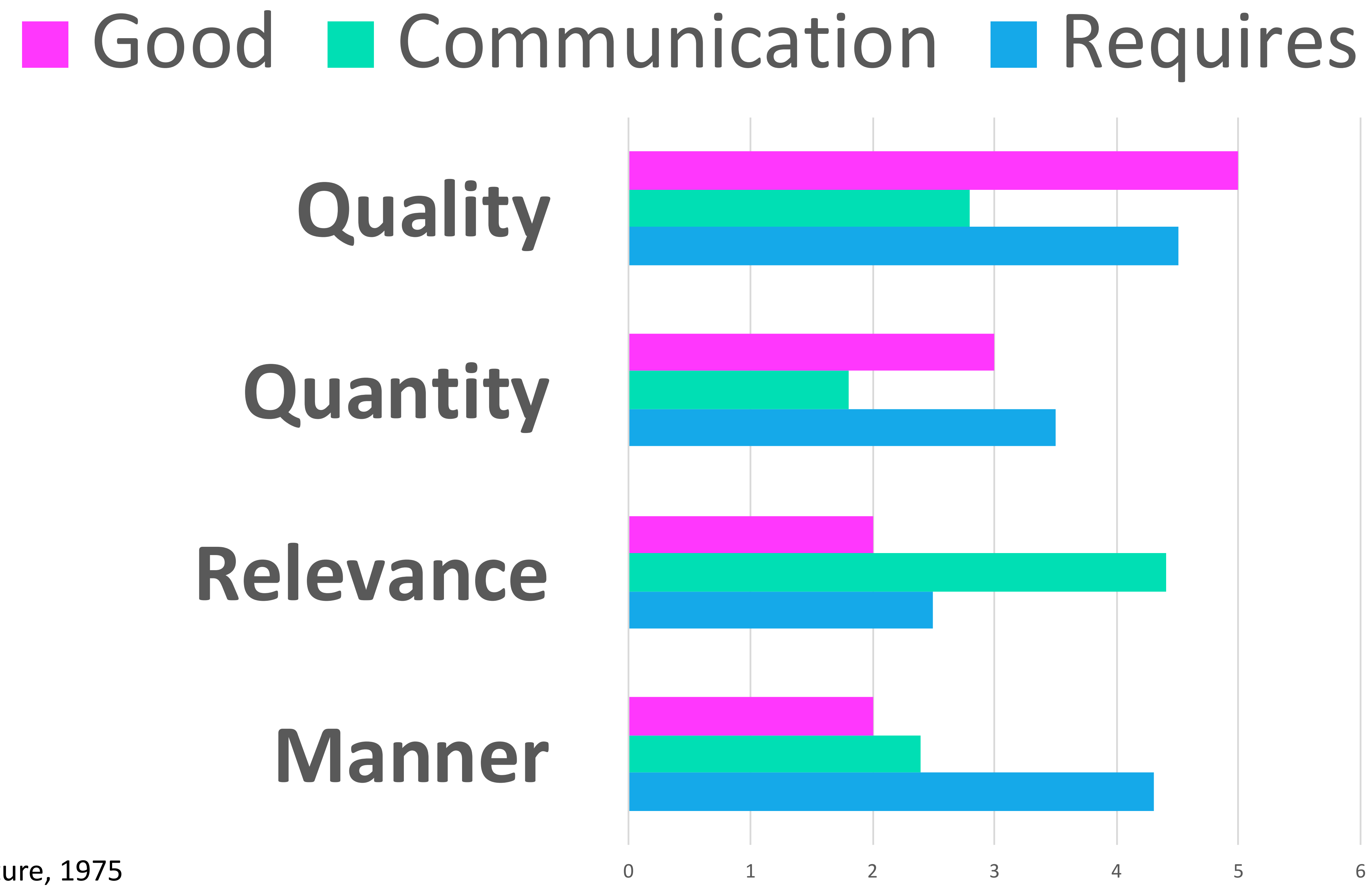



We
Learn
Through
Visual
Aids

Your job is to convince {\ref} of your science




Visuals
Are
Science



Grice, Implicature, 1975

Say everything you need to say, and nothing you don't.

Say everything you need to say, and nothing you don't.

Everything you say should be true and backed by evidence.

Say everything you need to say, and nothing you don't.

Everything you say should be true and backed by evidence.

Everything you say should be appropriate for the audience you are speaking to.

Say everything you need to say, and nothing you don't.

Everything you say should be true and backed by evidence.

Applicable to **plots** as well as **talks**

Everything you say should be appropriate for the audience you are speaking to.

Say everything you need to say, and nothing you don't.

Everything you say should be true and backed by evidence.

No viz is created in a **vacuum**

Everything you say should be appropriate for the audience you are speaking to.

Papers

Presentations

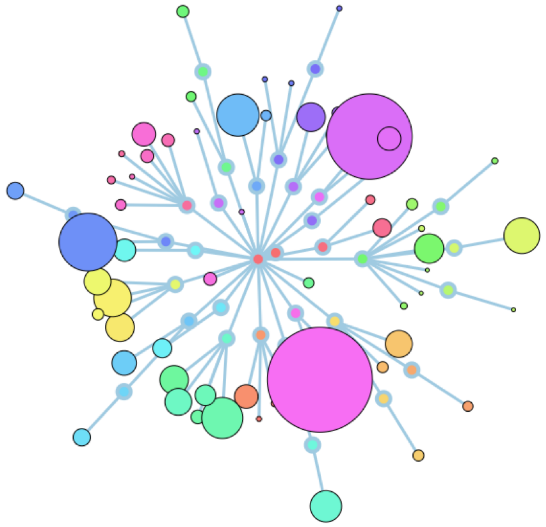
Special Projects

Papers

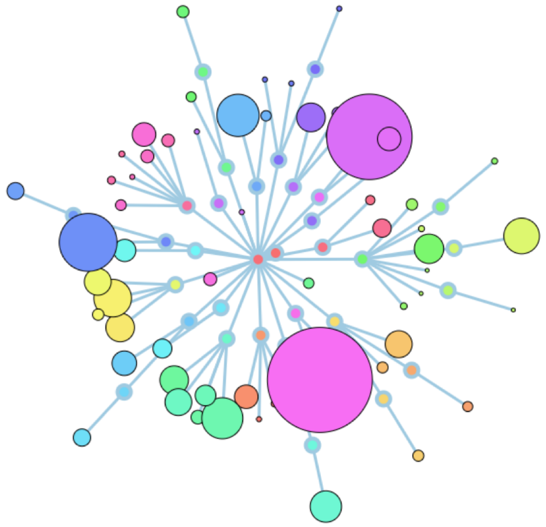
Presentations

Special Projects

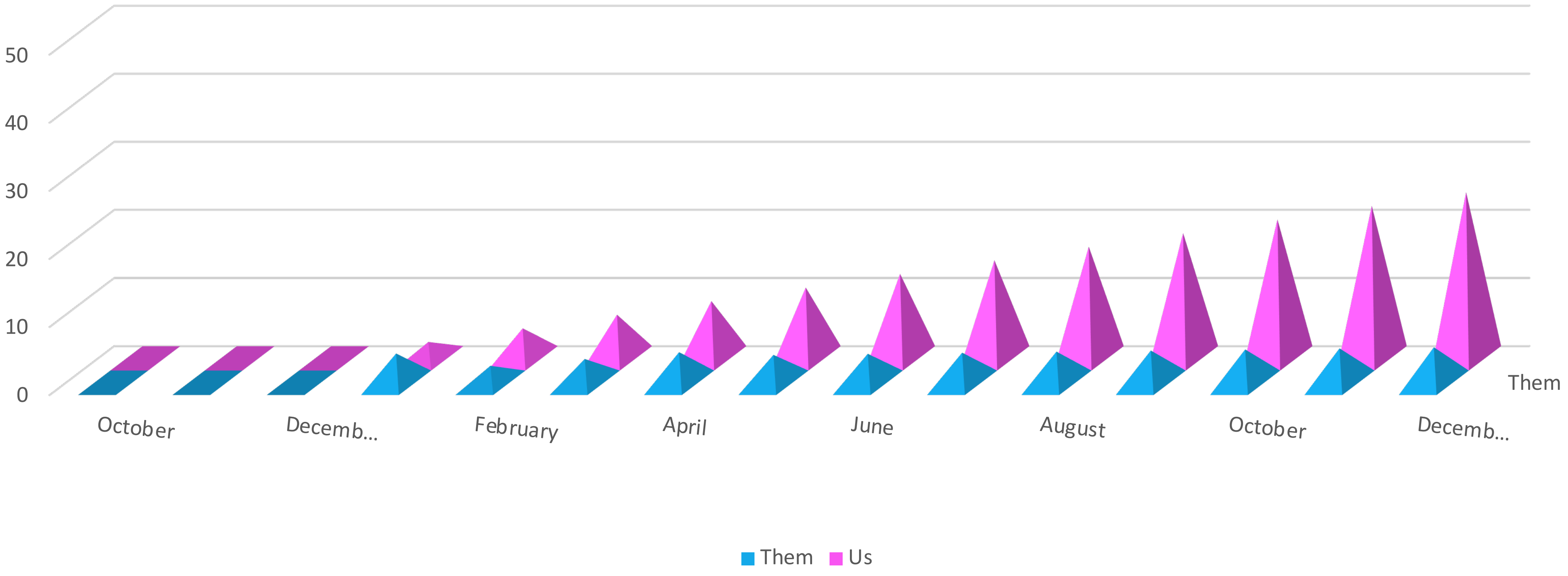
The key to a good plot



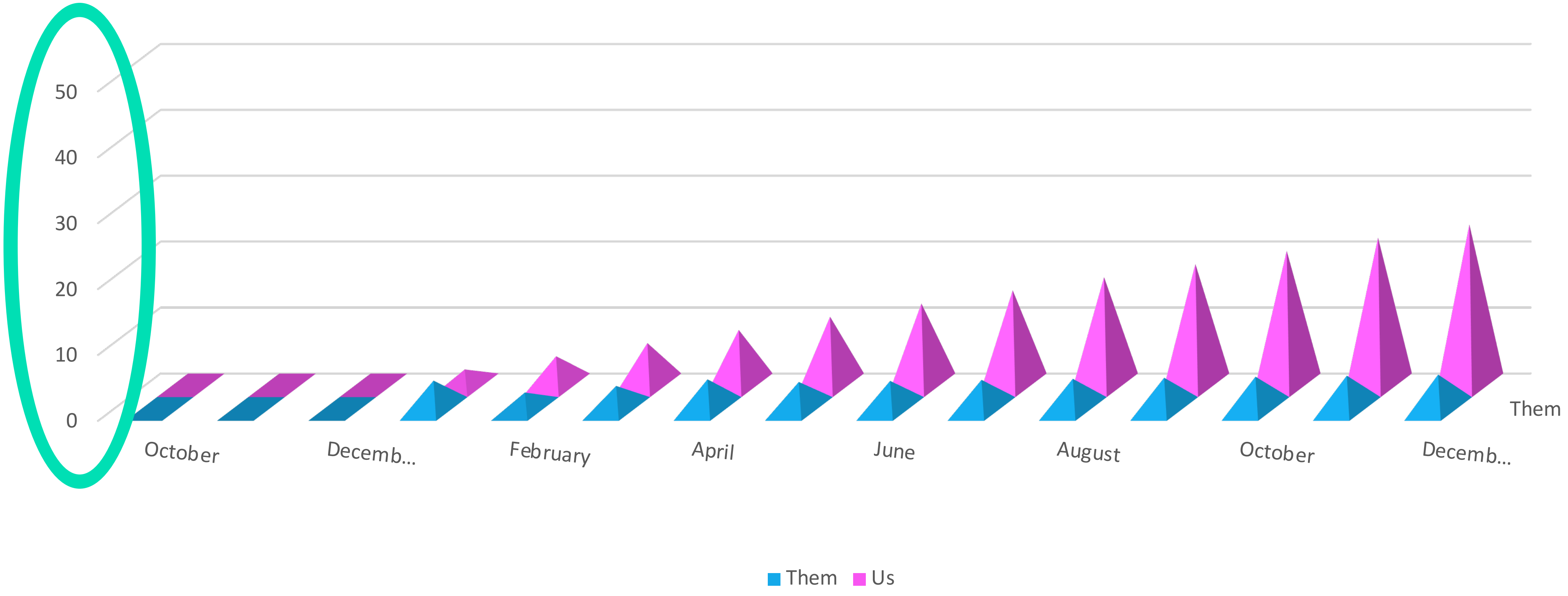
Remove unnecessary data ink and **emphasize** the most important data ink left.



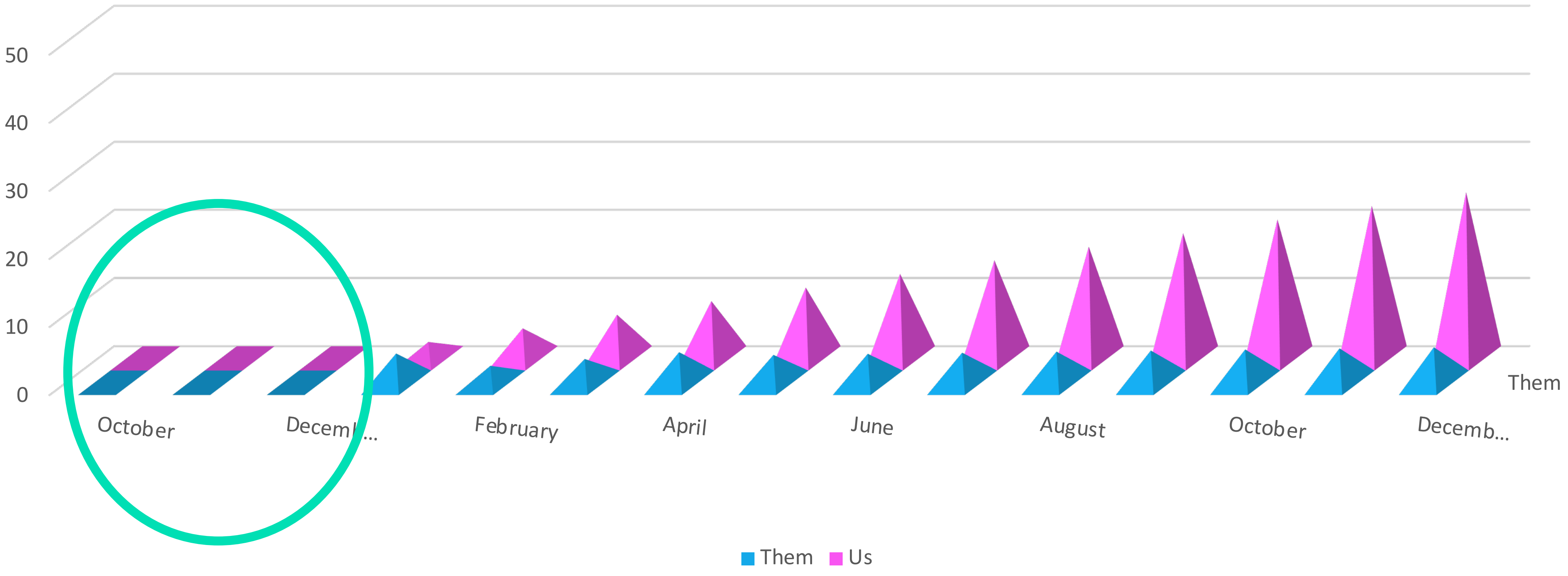
WHY is this a terrible plot?



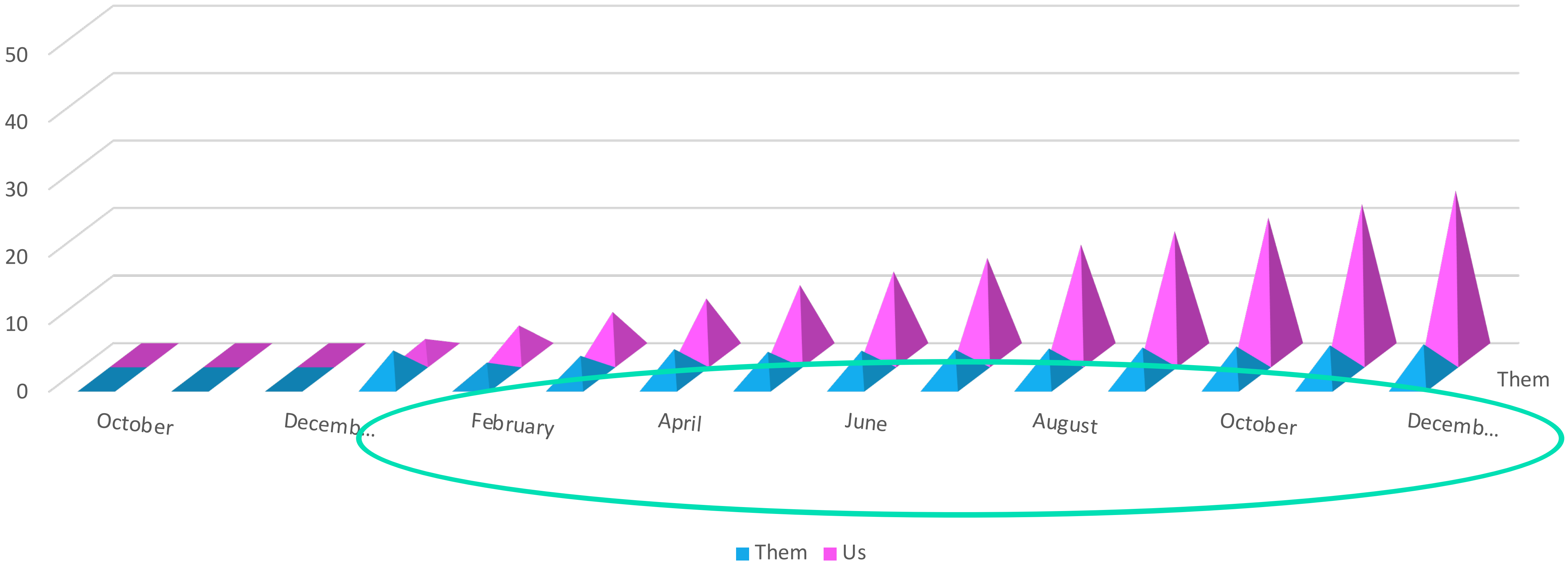
WHY is this a terrible plot?



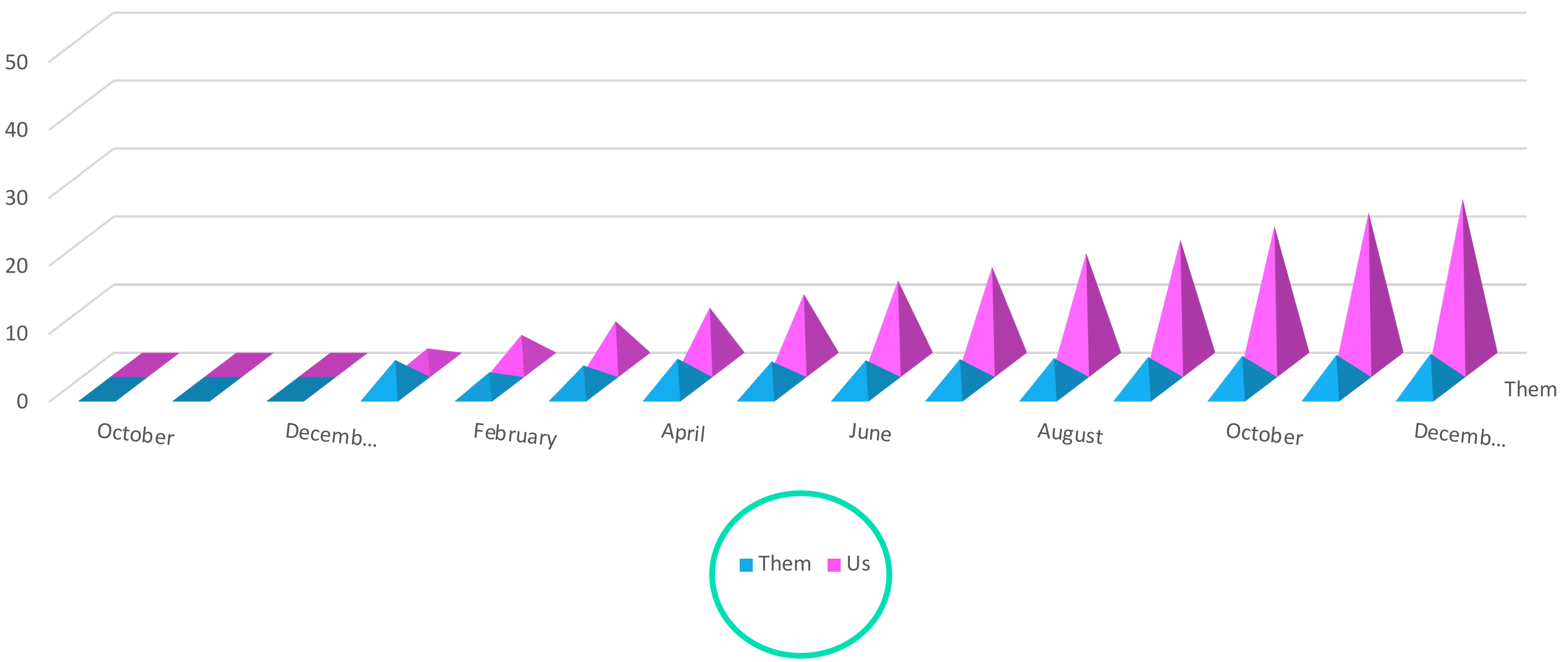
WHY is this a terrible plot?



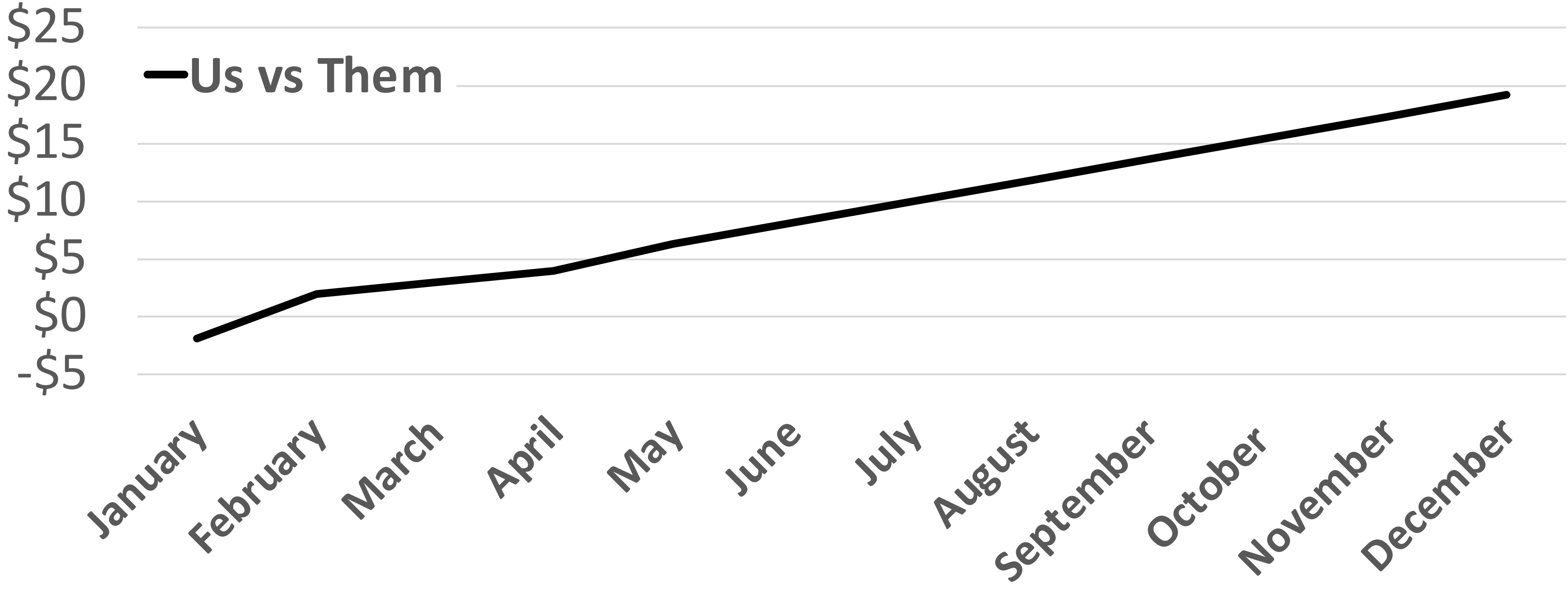
WHY is this a terrible plot?



WHY is this a terrible plot?

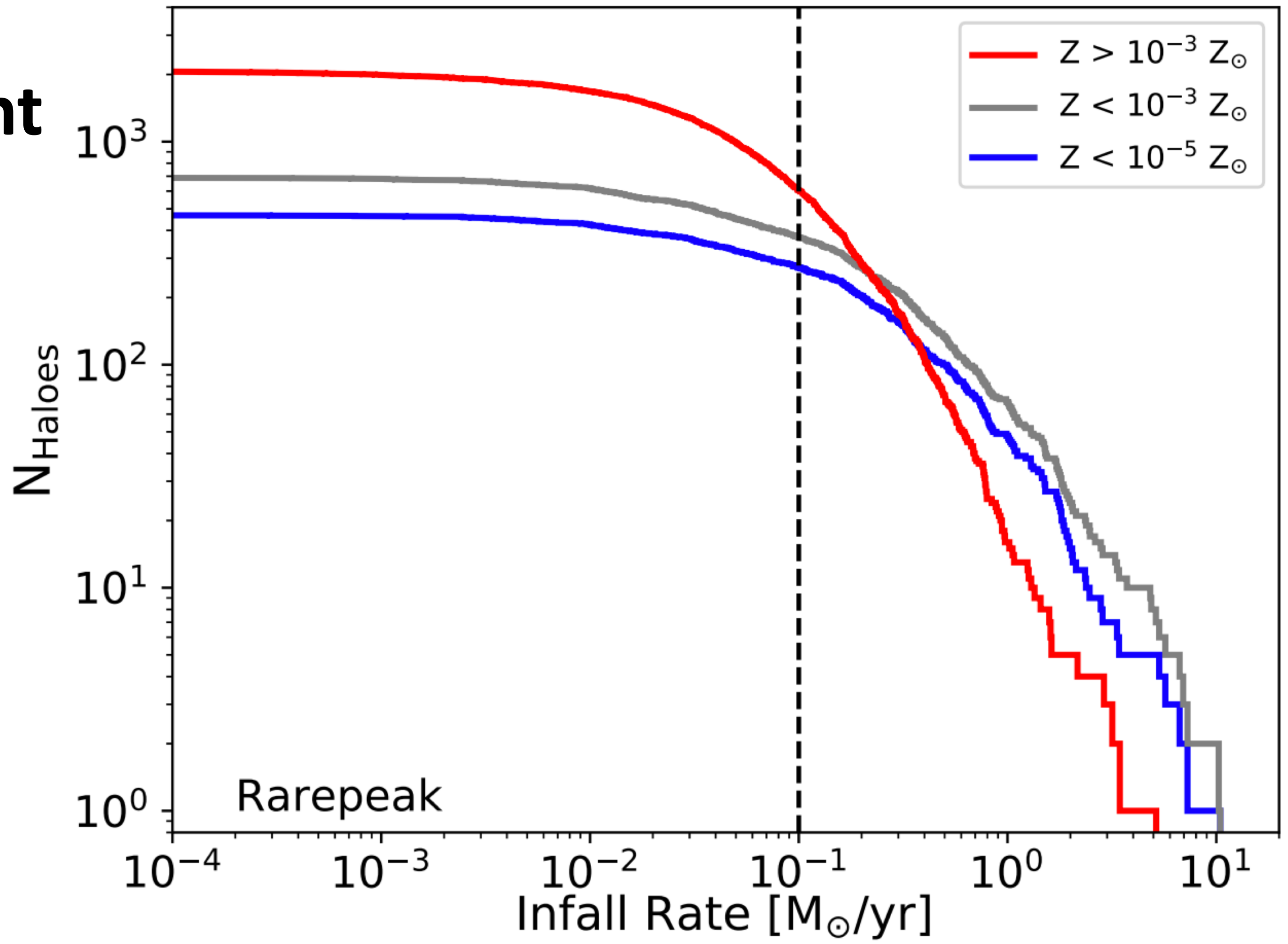


We Are Selling More Than Them



Anatomy of a Plot

1 **Graphical Element**



When we are examining data, what can we look for?

- Does this data describe a **geometric** object?
- Are the data points **connected** to each other?
- Can we describe data points with a fixed set of **categories**?
- Is there a **quantity** associated with the data?
- Are the datapoints **continuous** along one or more dimensions?

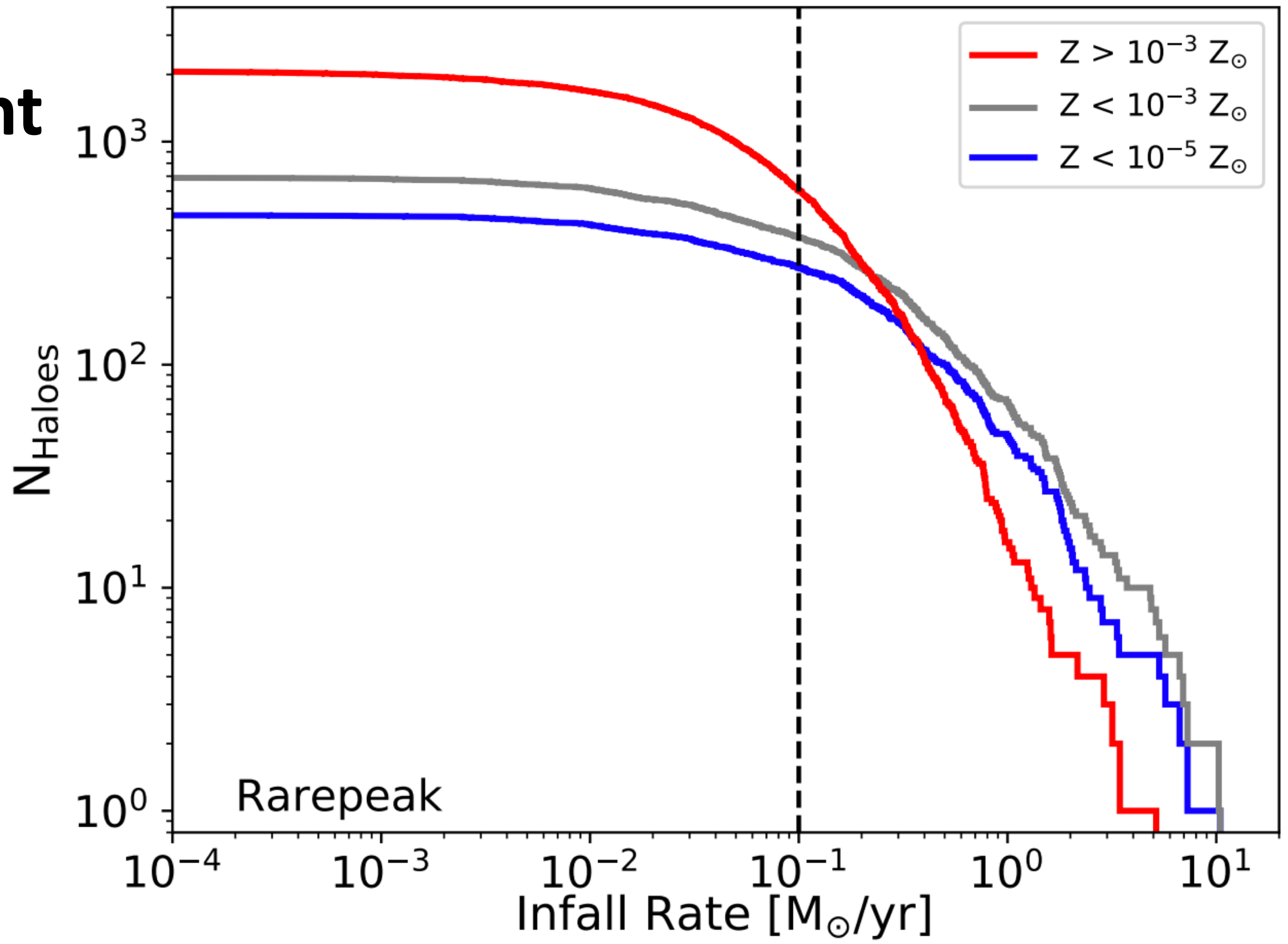
When we are examining data, what can we look for?

- Does this data describe a **geometric** object?
- Are the data points **connected** to each other?
- Can we describe data points with a fixed set of **categories**?
- Is there a **quantity** associated with the data?
- Are the datapoints **continuous** along one or more dimensions?

Is your data categorical or continuous?

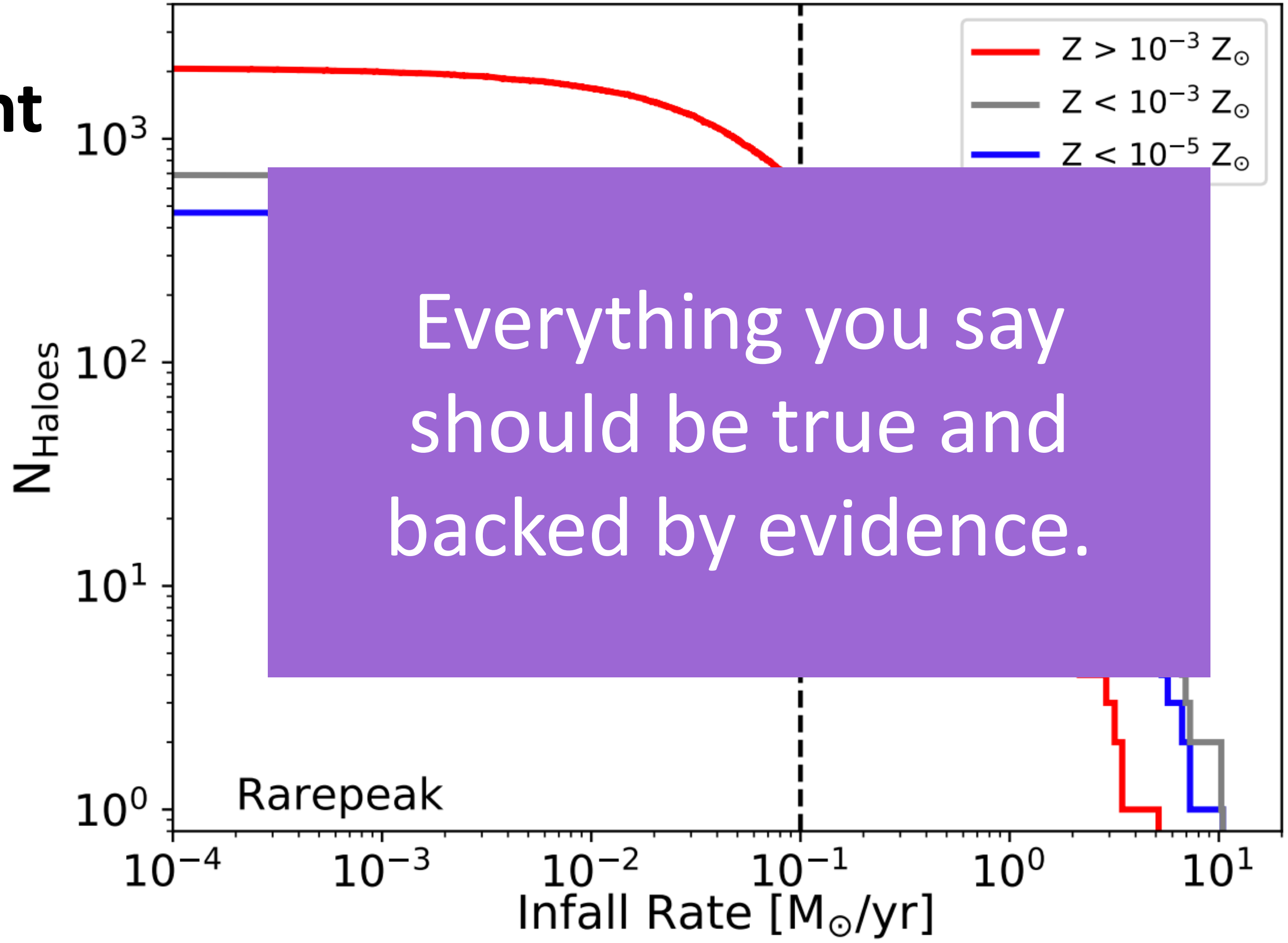
Anatomy of a Plot

1 **Graphical Element**



Anatomy of a Plot

1 **Graphical Element**



Everything you say should be true and backed by evidence.

A few tips

Be Mindful of Binning

Special Projects

A few tips

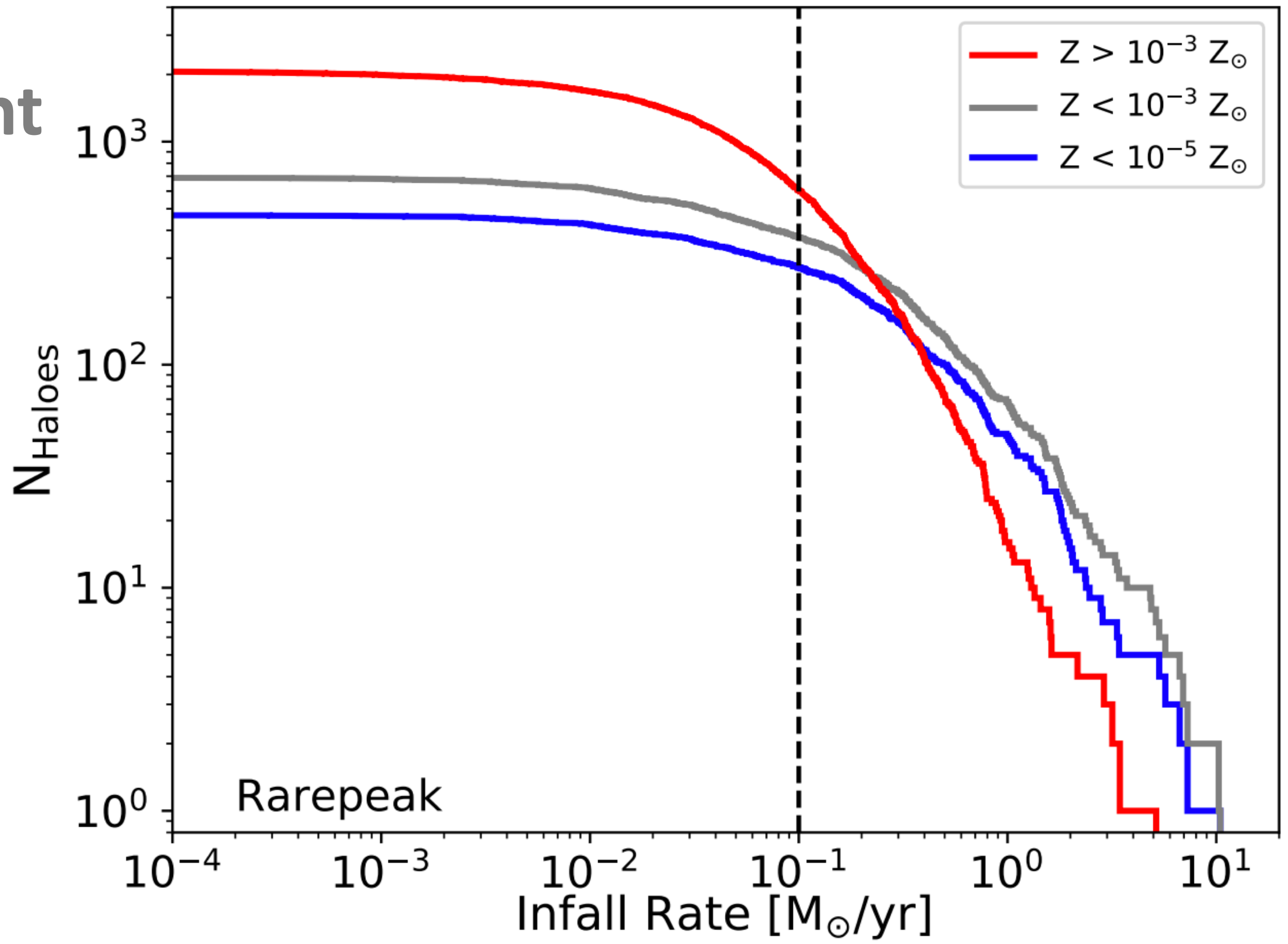
Avoid automatic line fitters and smoothers

Special Projects

Anatomy of a Plot

1 **Graphical Element**

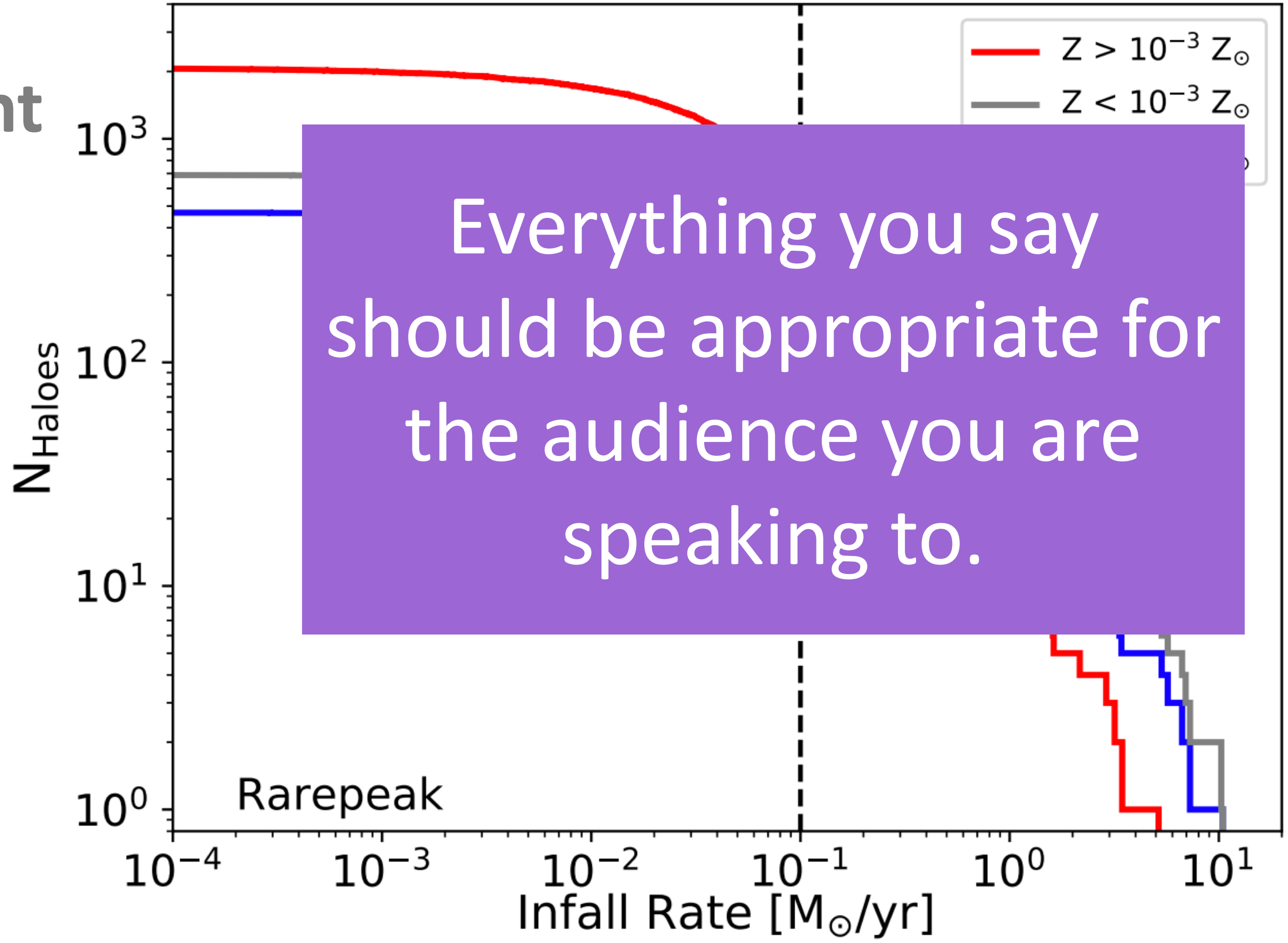
2 **Colors**



Anatomy of a Plot

1 **Graphical Element**

2 **Colors**



Everything you say should be appropriate for the audience you are speaking to.

A few tips

Make it black and white friendly

Special Projects

A few tips

Make it colorblind friendly

Special Projects

A few tips

Avoid Green

Special Projects

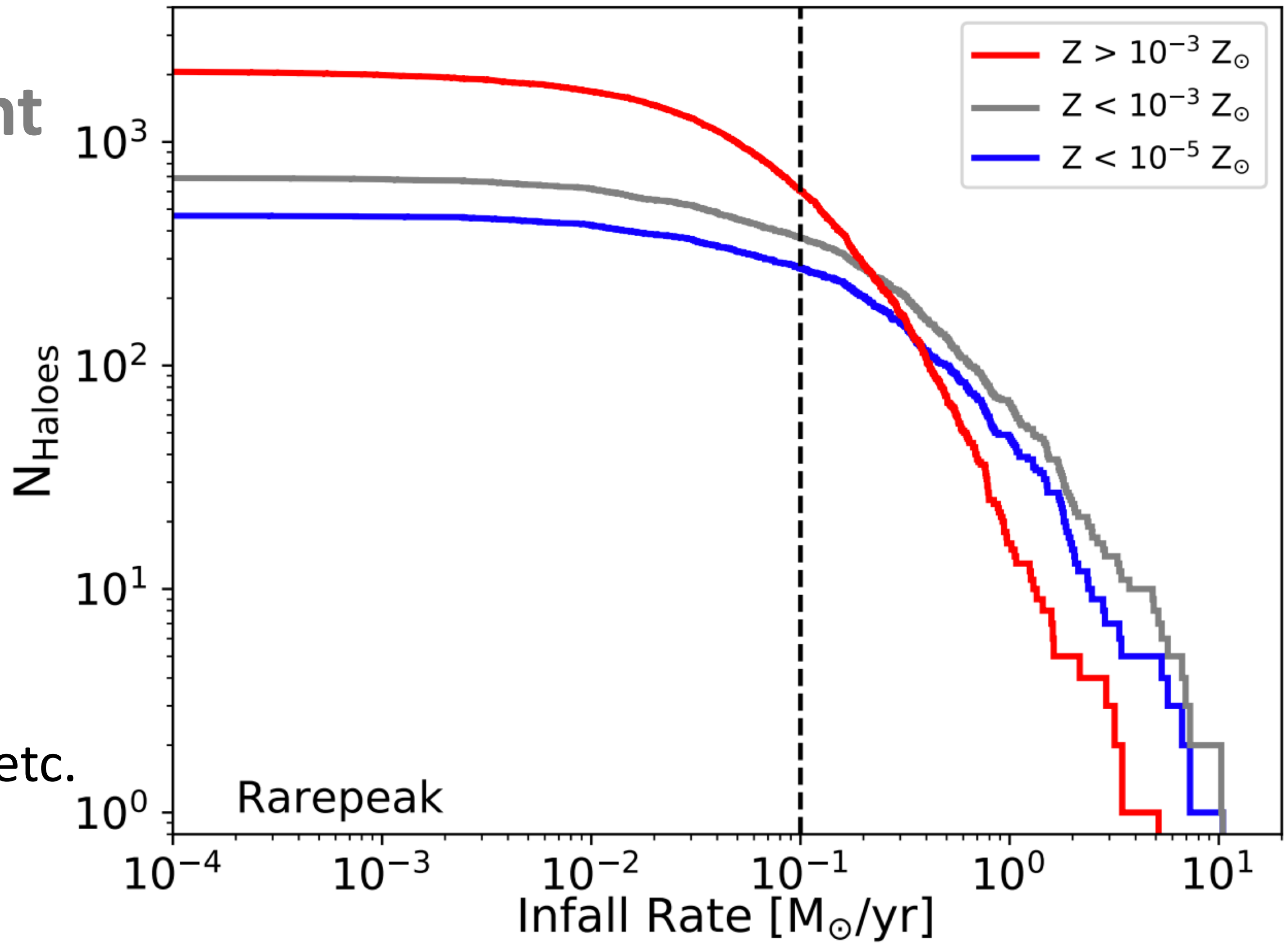
Anatomy of a Plot

1 **Graphical Element**

2 **Colors**

3 **Plot theme**

Axis labels, tick marks, etc.



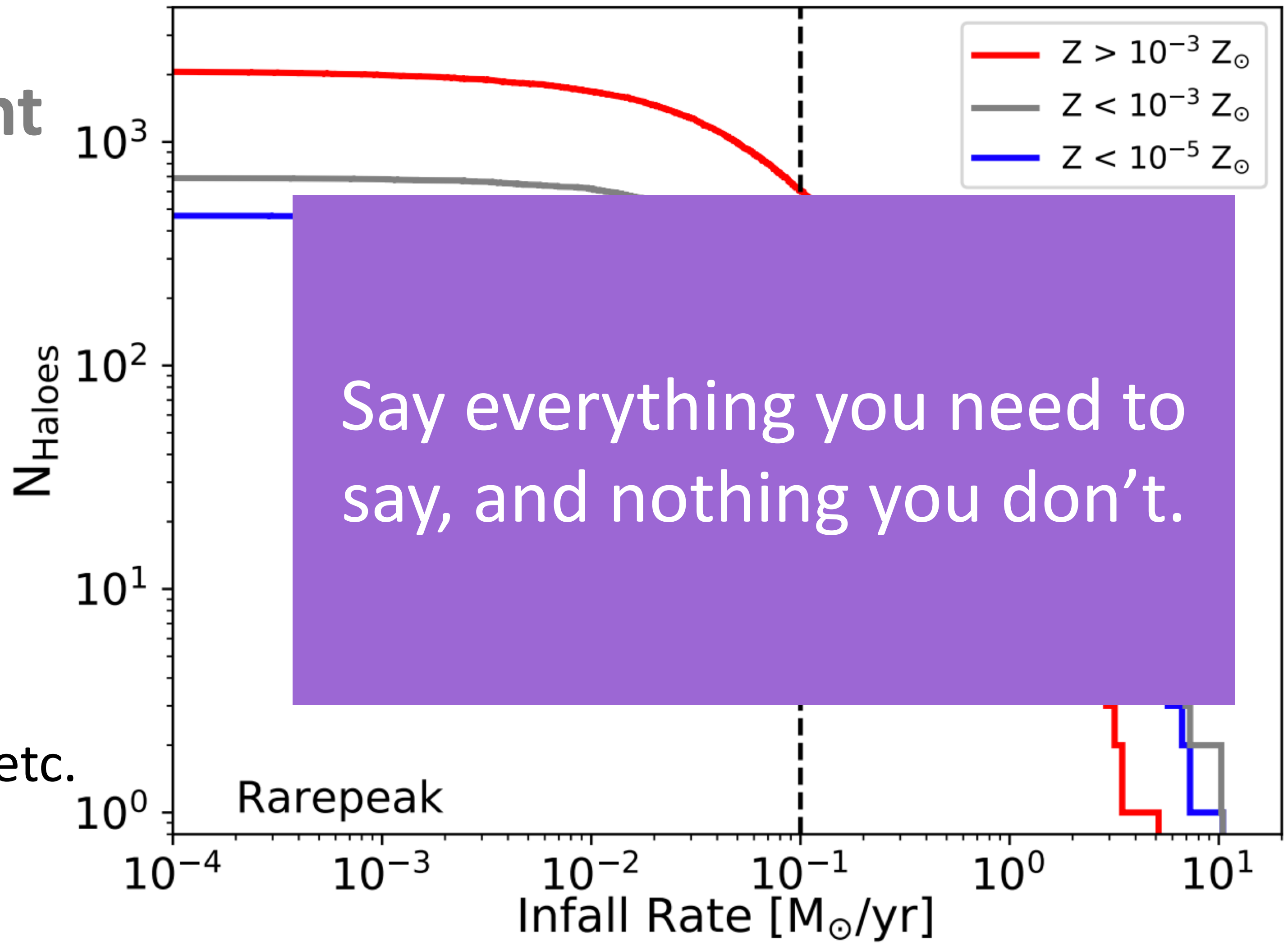
Anatomy of a Plot

1 **Graphical Element**

2 **Colors**

3 **Plot theme**

Axis labels, tick marks, etc.



A few tips

Set up your plot theme first

Special Projects

A few tips

This includes colors, tick mark size, axis placement, fonts, etc.

Special Projects

A few tips

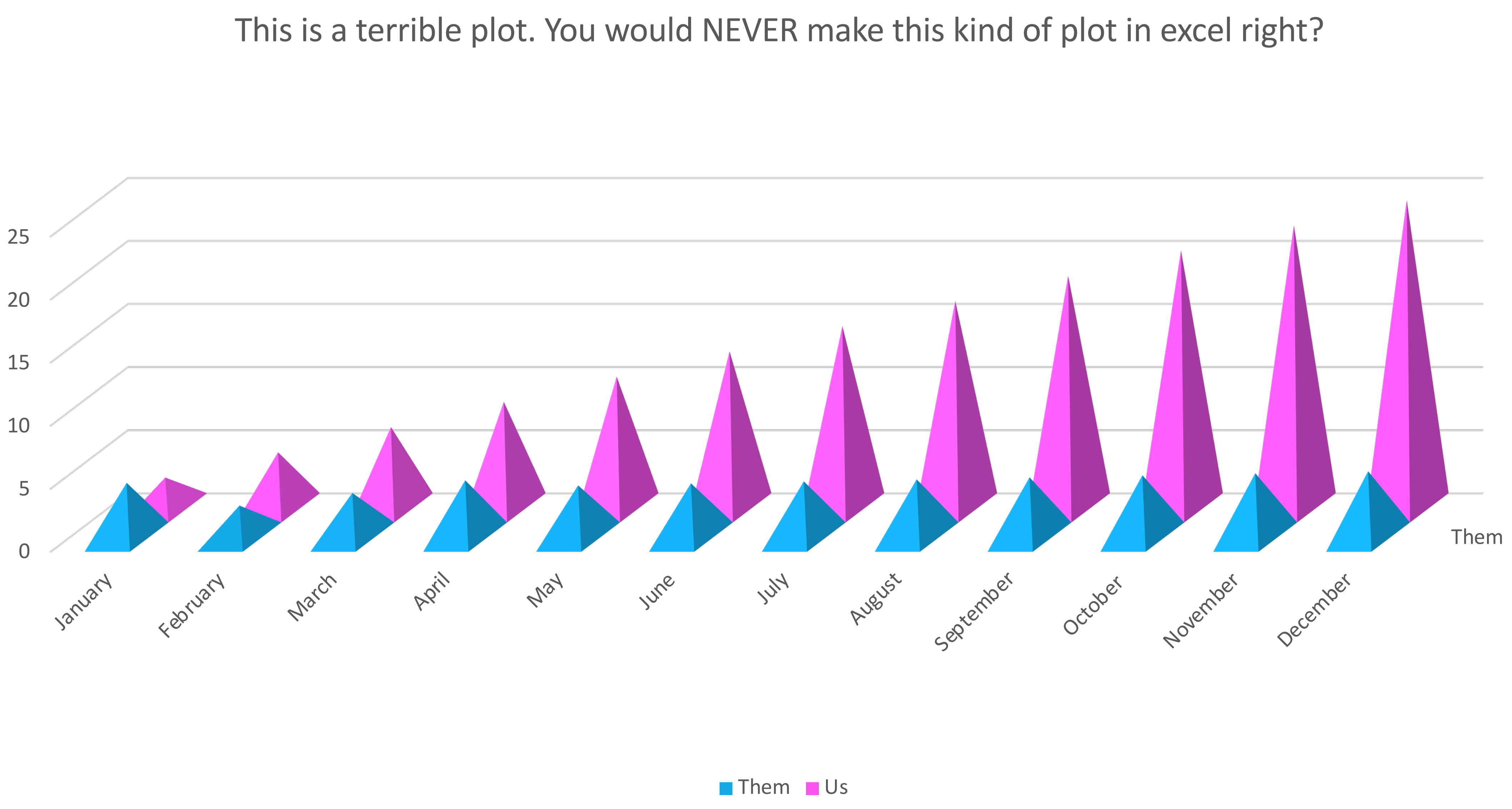
Be confident about your log-log axes

Special Projects

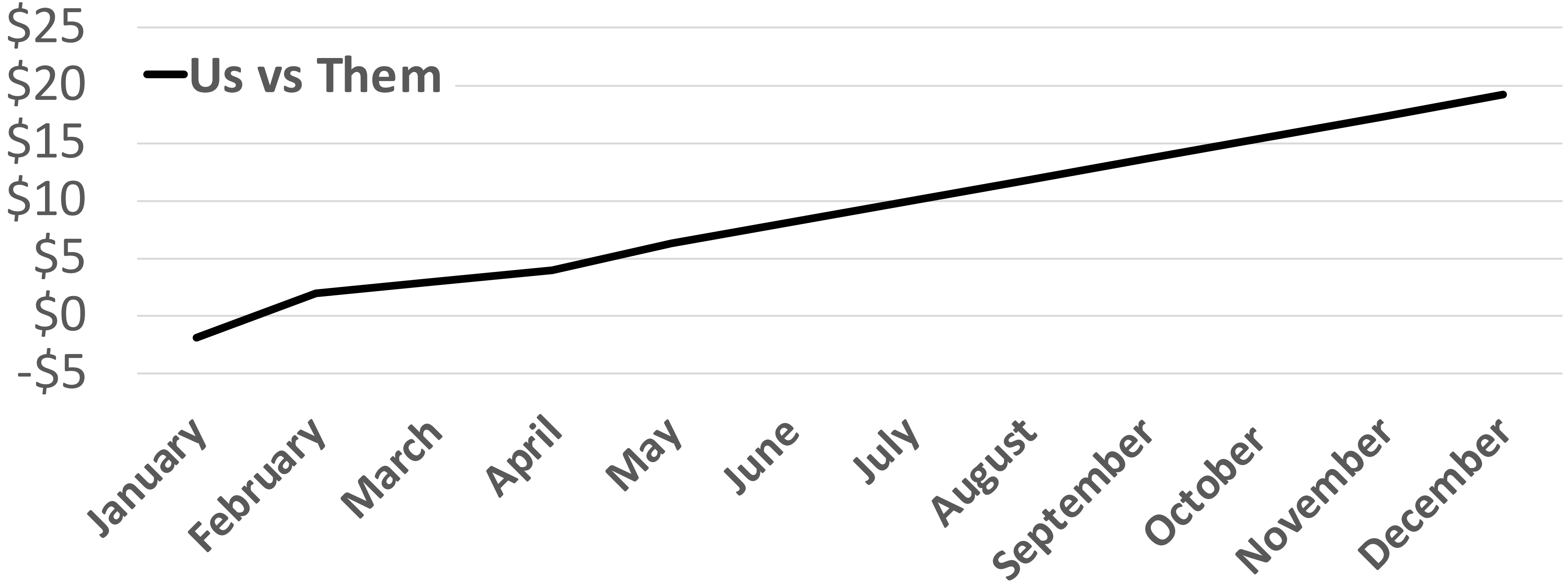
A few tips

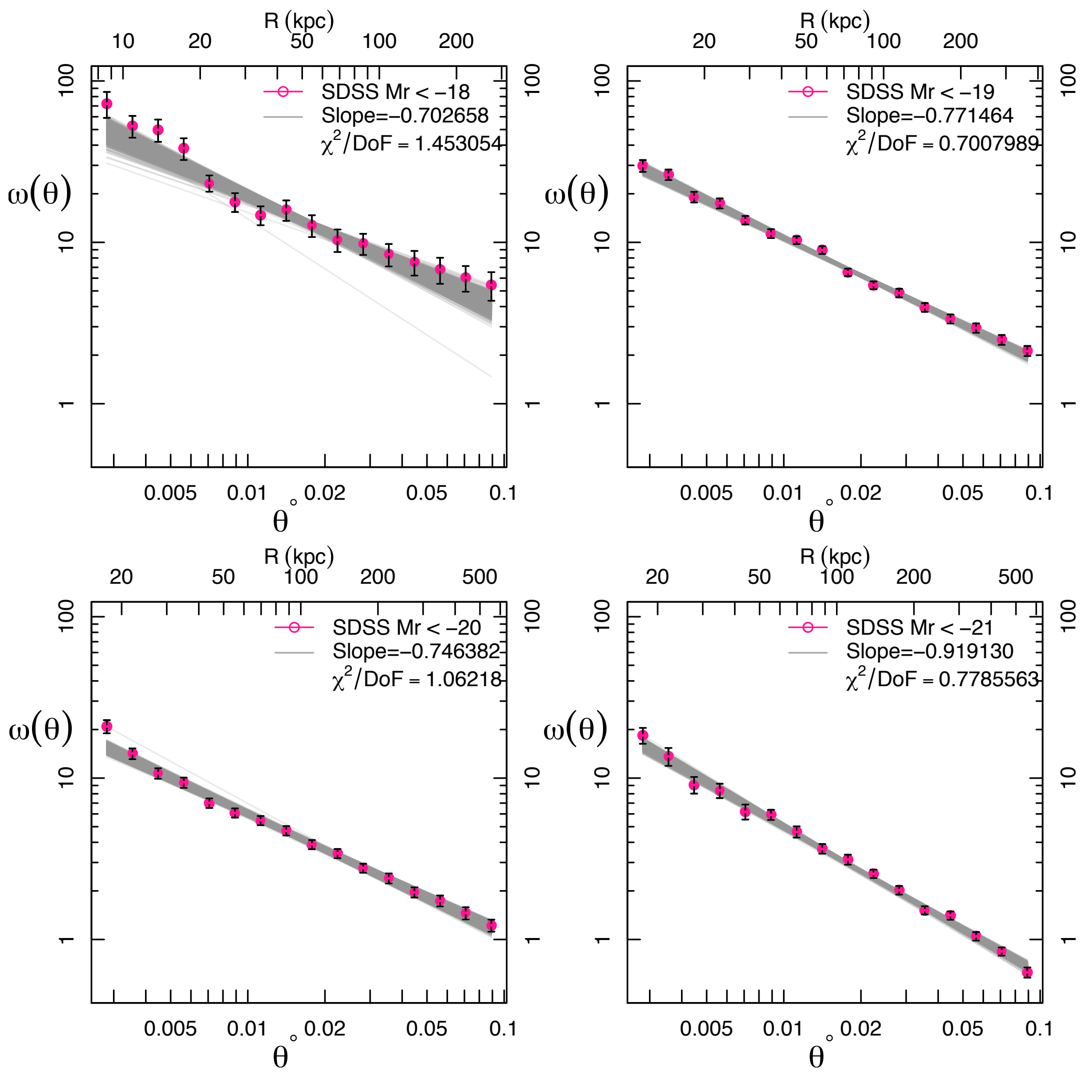
Don’t Make People Figure Out the Point

Special Projects


This is a terrible plot. You would NEVER make this kind of plot in excel right?
25
20
15
10
5
0
January
February
March
April
May
June
July
August
September
October
November
December
Them
Them
Us

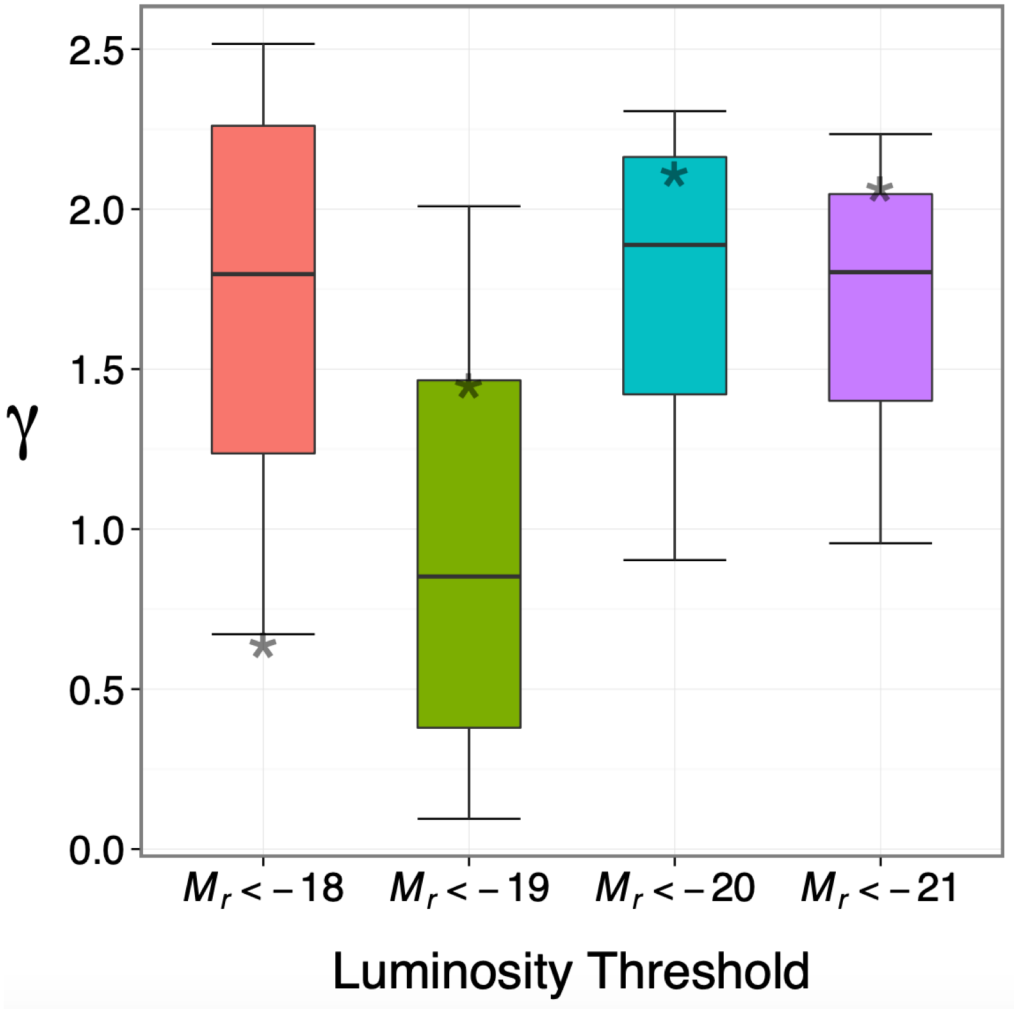
We Are Selling More Than Them





SDSS Mr < −18
Slope=−0.702658
χ²/DoF = 1.453054
SDSS Mr < −19
Slope=−0.771464
χ²/DoF = 0.7007989
SDSS Mr < −20
Slope=−0.746382
χ²/DoF = 1.06218
SDSS Mr < −21
Slope=−0.919130
χ²/DoF = 0.7785563
R (kpc)
ω(θ)
θ°

VS



γ
Mr < −18
Mr < −19
Mr < −20
Mr < −21
Luminosity Threshold

Papers

Presentations

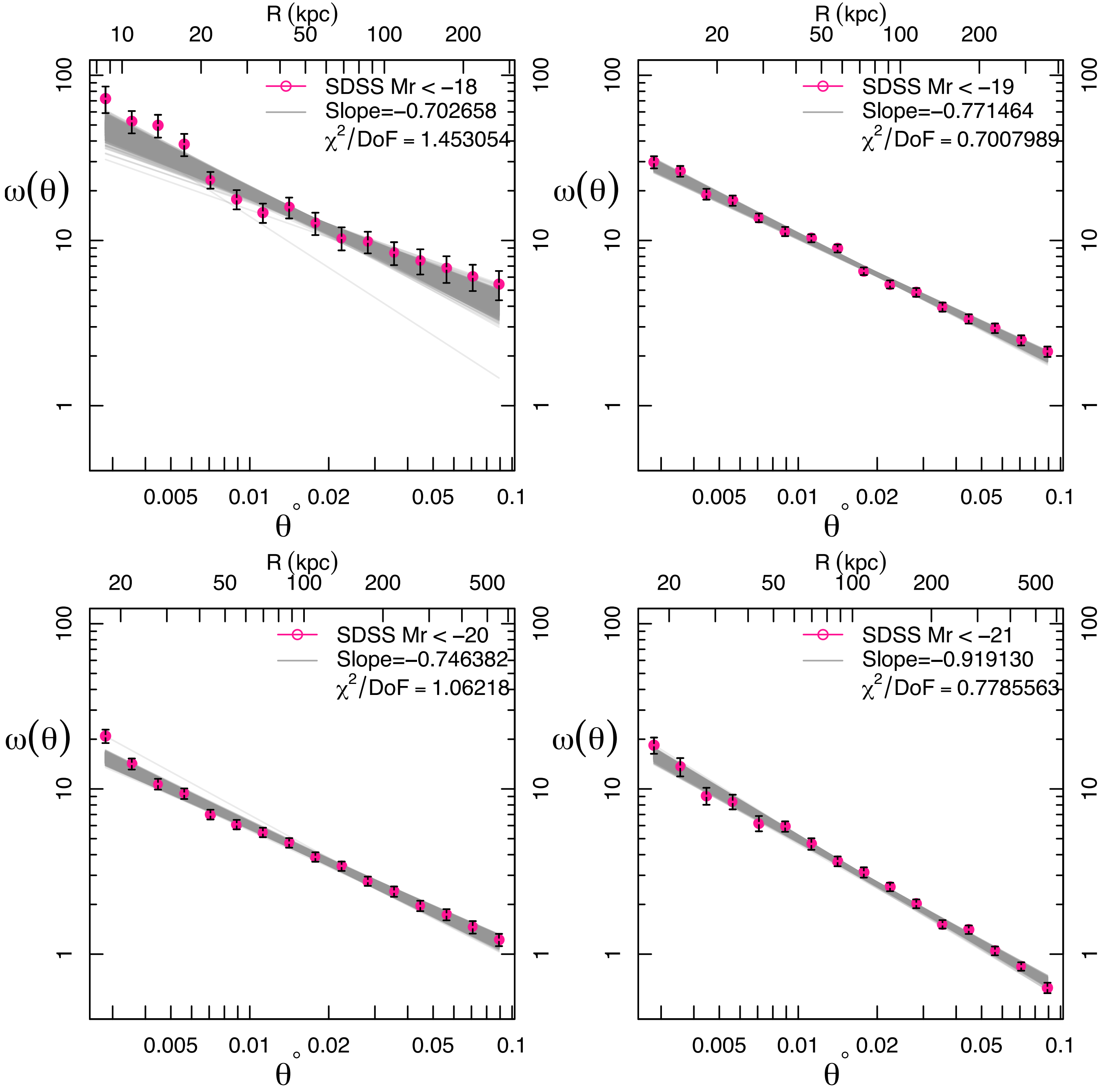
Special Projects

Papers

Turning your paper plots into presentation plots

Special Projects

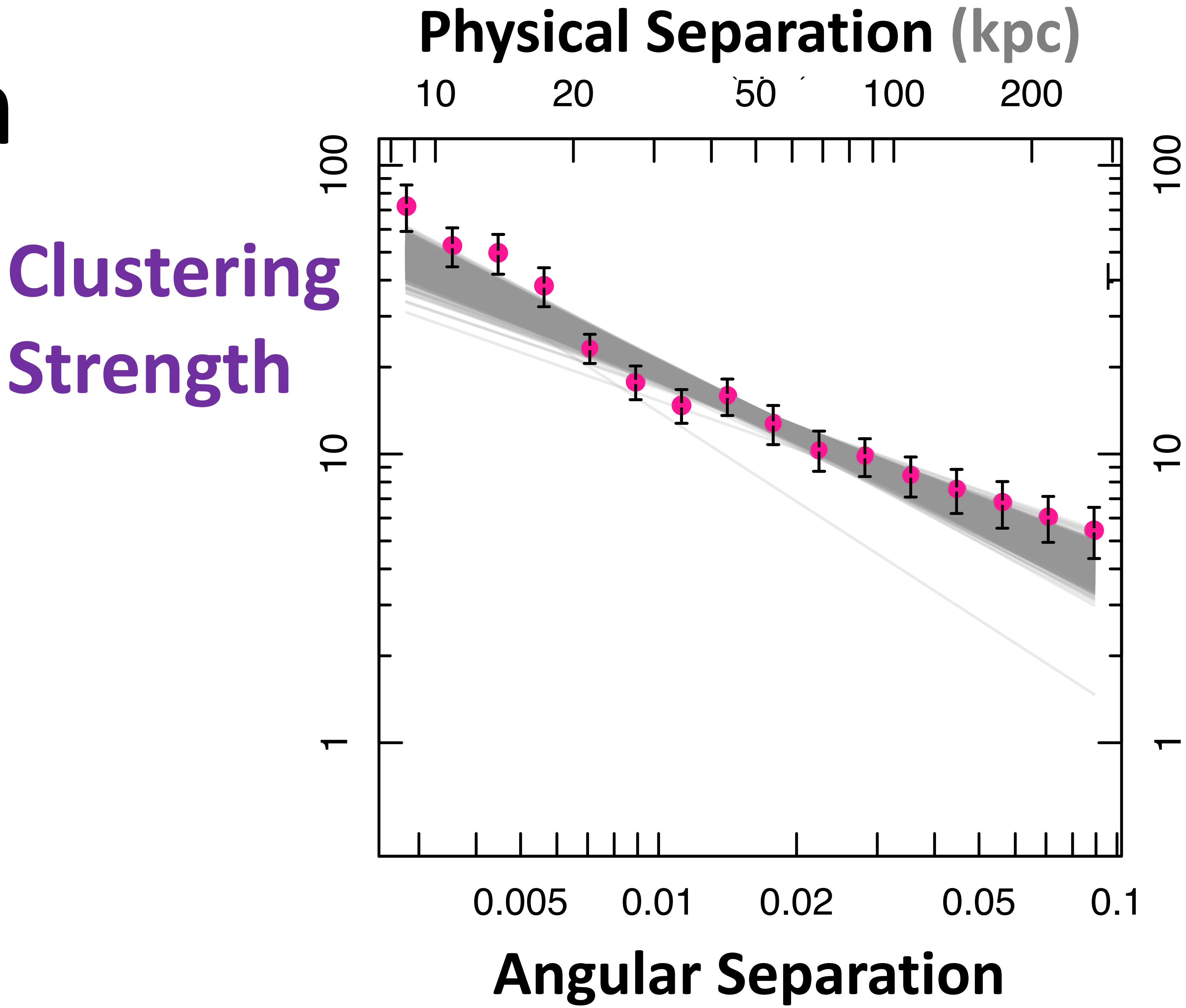
This is a paper plot



This is a

presentation

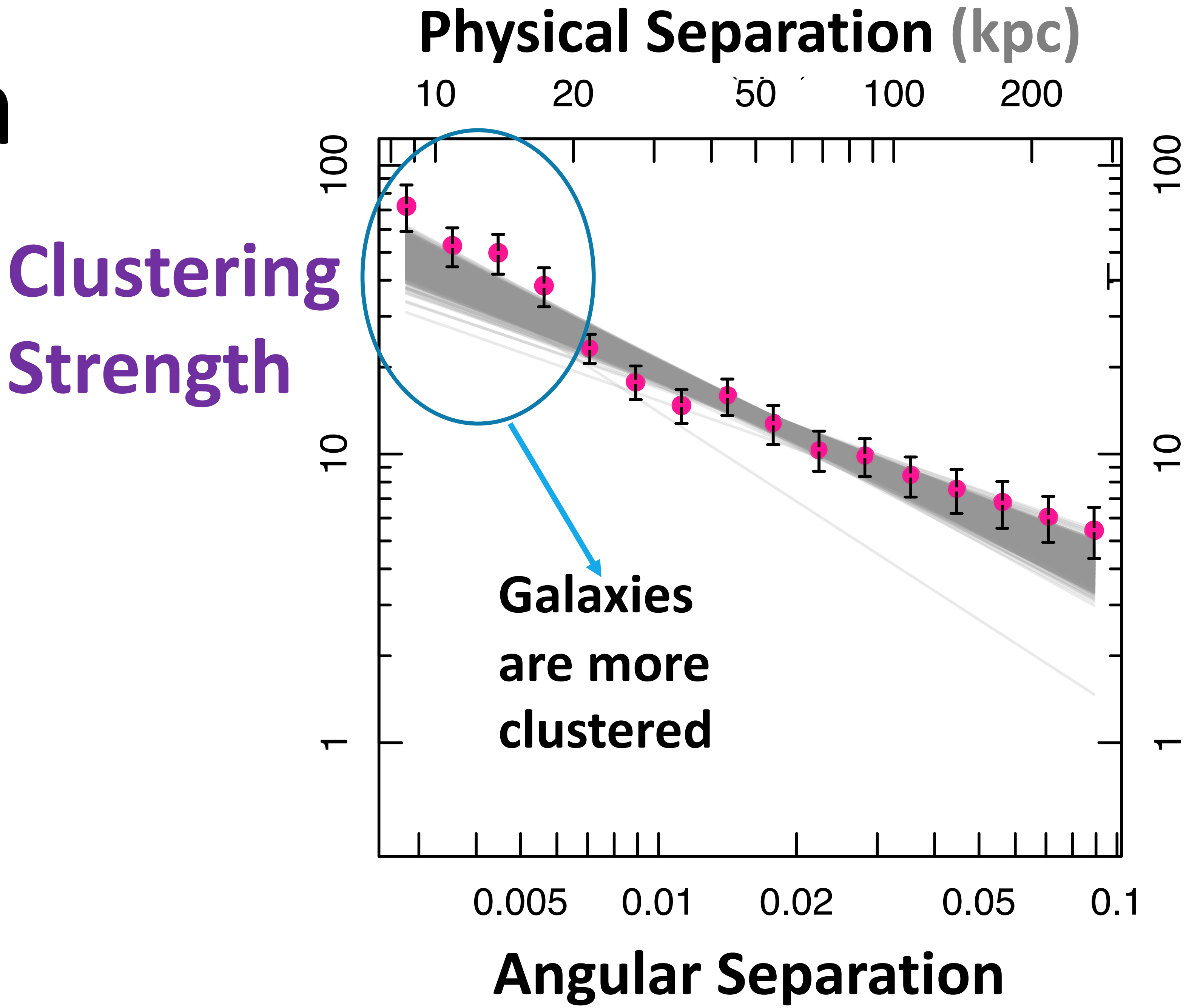
plot



This is a

presentation

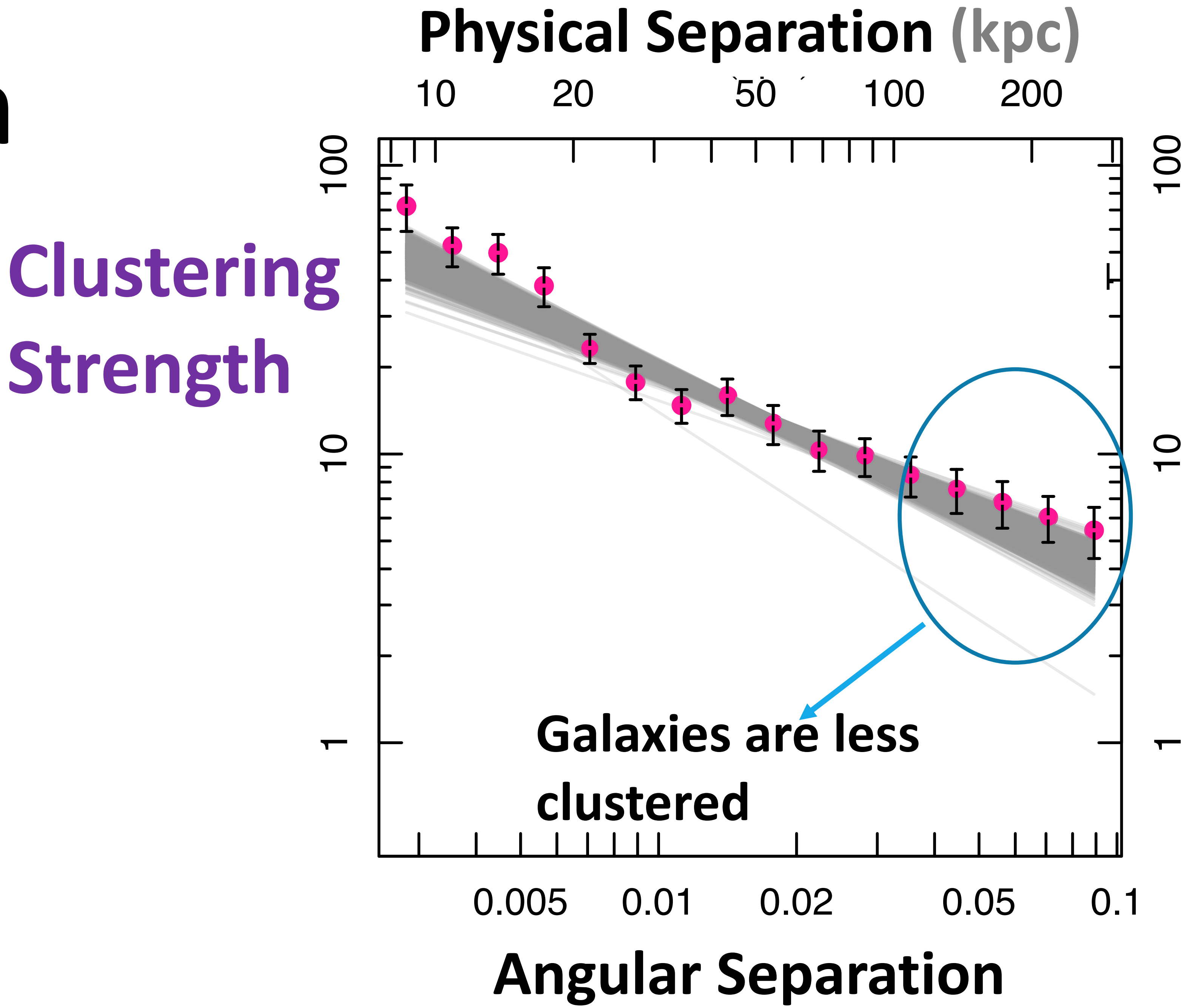
plot



This is a

presentation

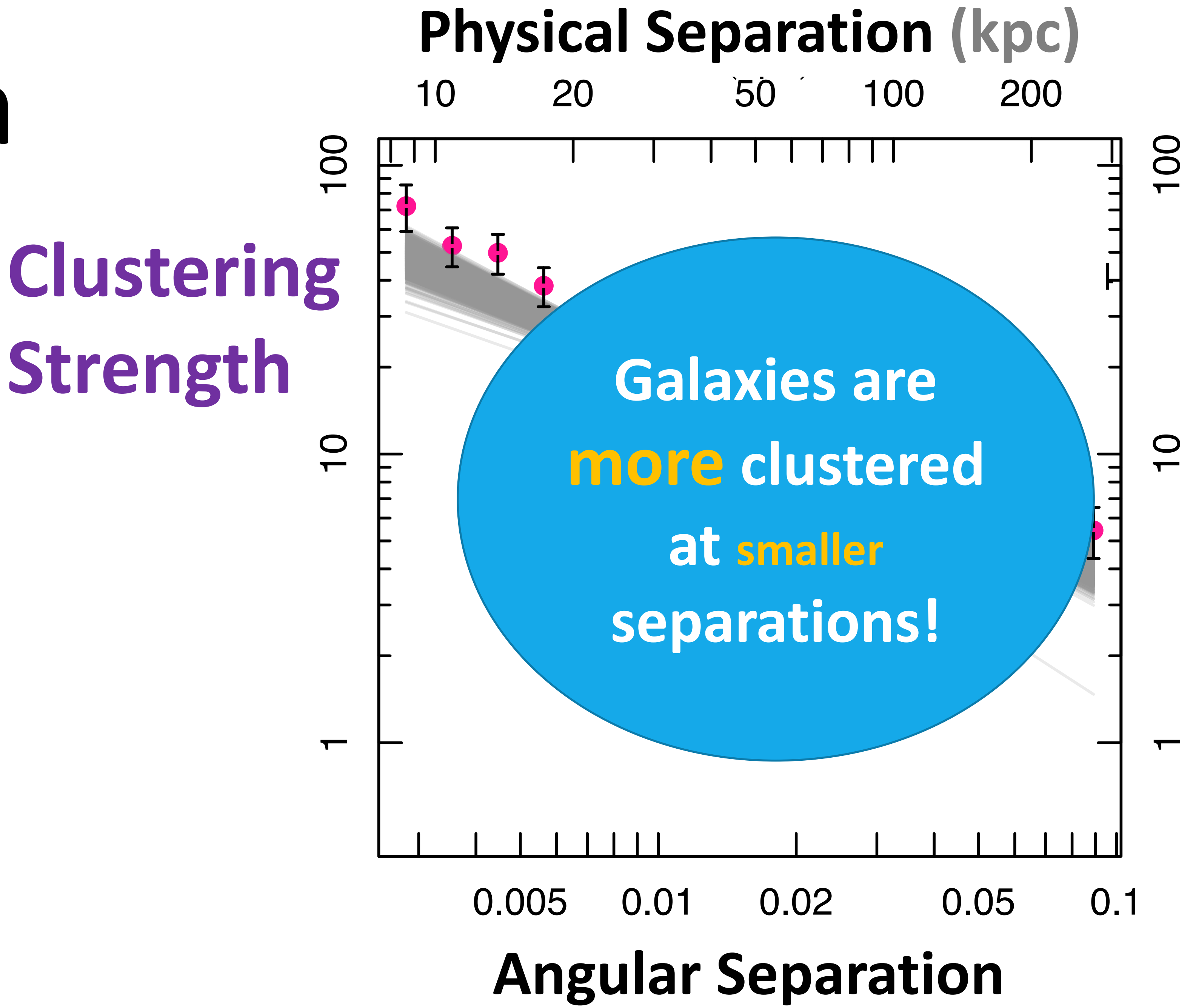
plot



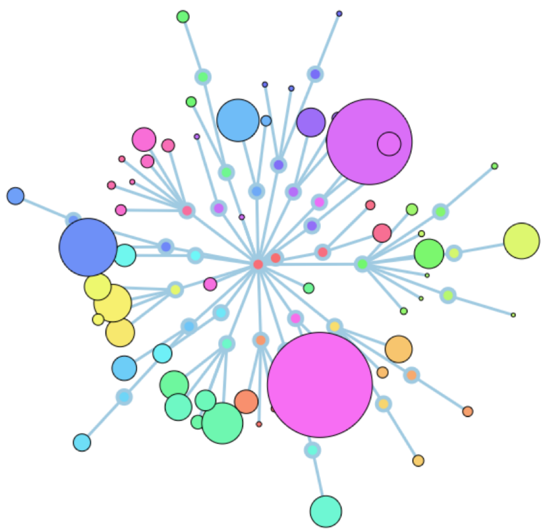
This is a

presentation

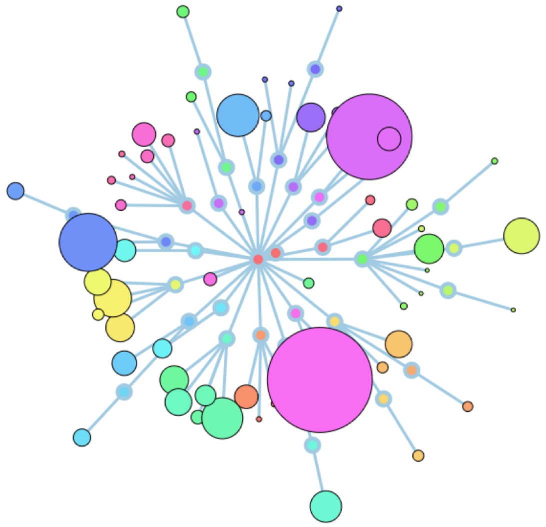
plot



Have a simple message you repeat often



An aside on powerpoint



A few tips

For general talks, pitch your presentation to a first year grad student

Special Projects

A few tips

Use Words Sparingly

Special Projects

If words are here

- They are not listening to you
  - This is a VERY important paragraph about how if you put everything on your slide, people won’t pay attention to you and instead read all the words in this very long run on sentence in too small font and maybe I should’ve used Helvetica?
- Definitely should’ve used Helvetica
- Is Helvetica even available on macs?
- I should watch the Helvetica documentary again
- That was wild.
- Update: Helvetica IS available on macs

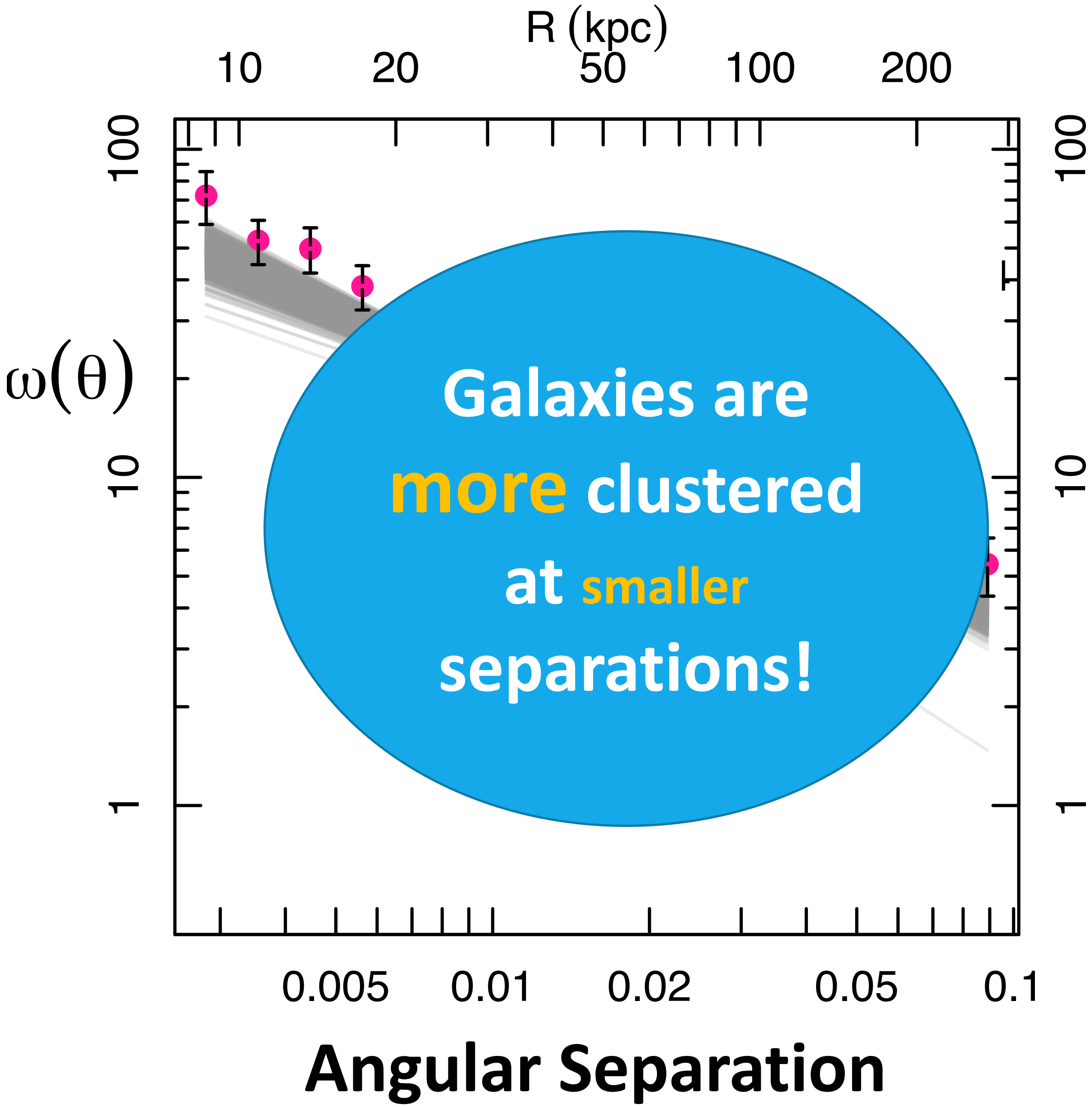
A few tips

Include reminder or catch-up slides

Special Projects

Did you fall asleep?

- Repeat your simple message often
  - Say everything you need and nothing you don't
  - Everything you say should be true and backed by evidence
  - Everything you say should be appropriate for your audience



A few tips

Do not put tables in your talk

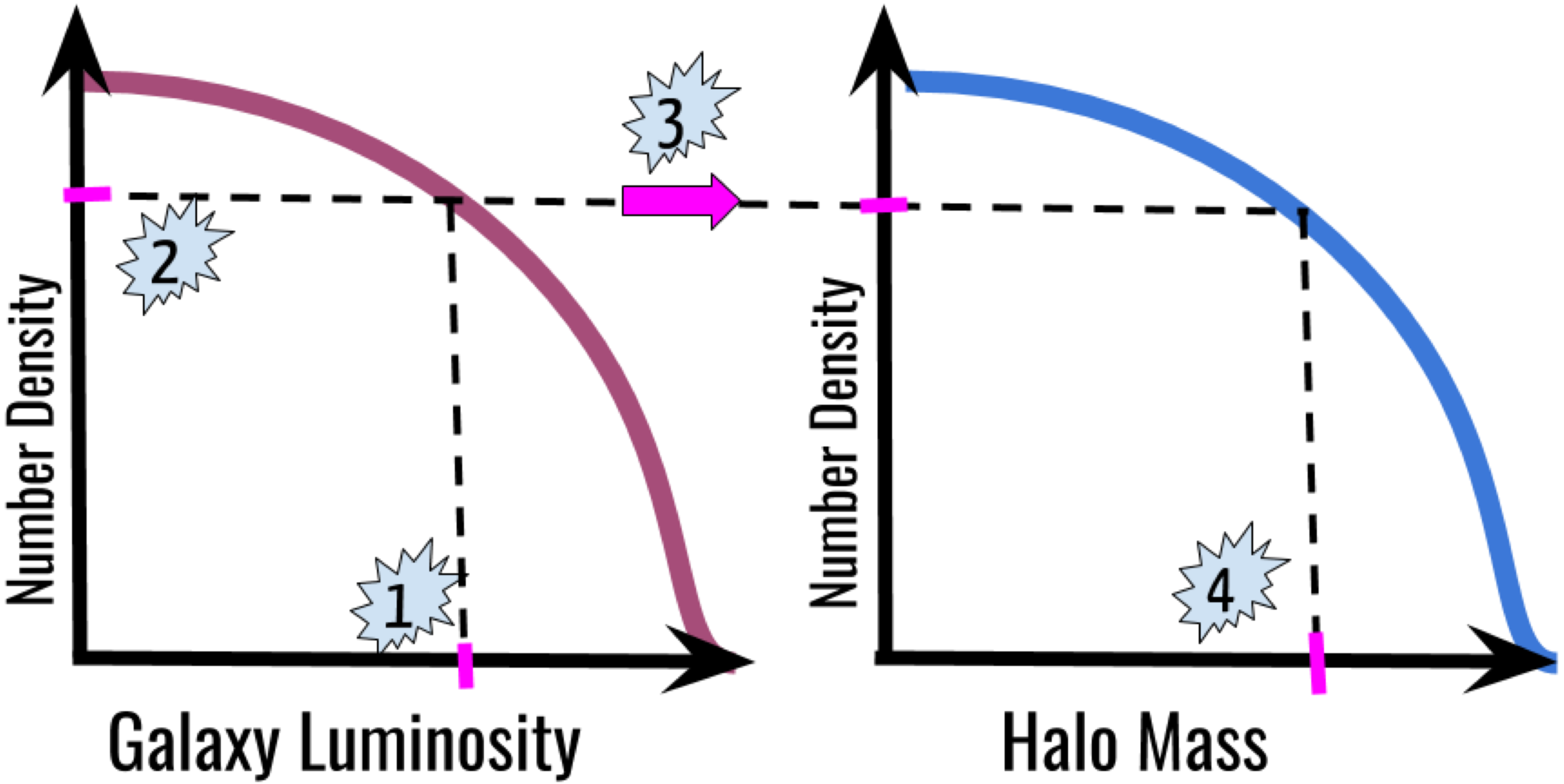
Special Projects

Please	Do not	Put	Tables in	Your	Presentations	
$0.2 < z < 0.5$	$(0.40^{+0.05}_{-0.05}, 0.22^{+0.06}_{-0.06})$	1.5	$(1.09^{+0.28}_{-0.10}, 10.33^{+0.86}_{-0.27}, -0.77^{+0.22}_{-0.31})$	0.7	7.2	0.2
$0.5 < z < 0.8$	$(0.46^{+0.07}_{-0.06}, 0.45^{+0.09}_{-0.08})$	8.0	$(1.42^{+0.13}_{-0.06}, 10.3^{+0.39}_{-0.17}, -0.83^{+0.18}_{-0.17})$	0.57	22.8	16.9
$0.8 < z < 1.1$	$(0.46^{+0.08}_{-0.3}, 0.68^{+0.45}_{-0.13})$	4.1	$(1.83^{+7.1}_{-0.20}, 10.7^{+33}_{-0.44}, -0.69^{+0.54}_{-0.35})$	25.0	14.9	21.9
$1.1 < z < 1.5$	$(0.63^{+0.05}_{-0.06}, 0.59^{+0.09}_{-0.06})$	5.8	$(1.93^{+0.18}_{-0.19}, 10.62^{+0.56}_{-0.3}, -0.93^{+0.23}_{-0.16})$	1.3	9.1	2.1
$1.5 < z < 2.0$	$(0.58^{+0.05}_{-0.06}, 0.86^{+0.11}_{-0.07})$	2.8	$(2.22^{+6.3}_{-0.27}, 10.95^{+13.7}_{-0.5}, -0.82^{+0.30}_{-0.35})$	10.77	-8.4	-36.3

A few tips

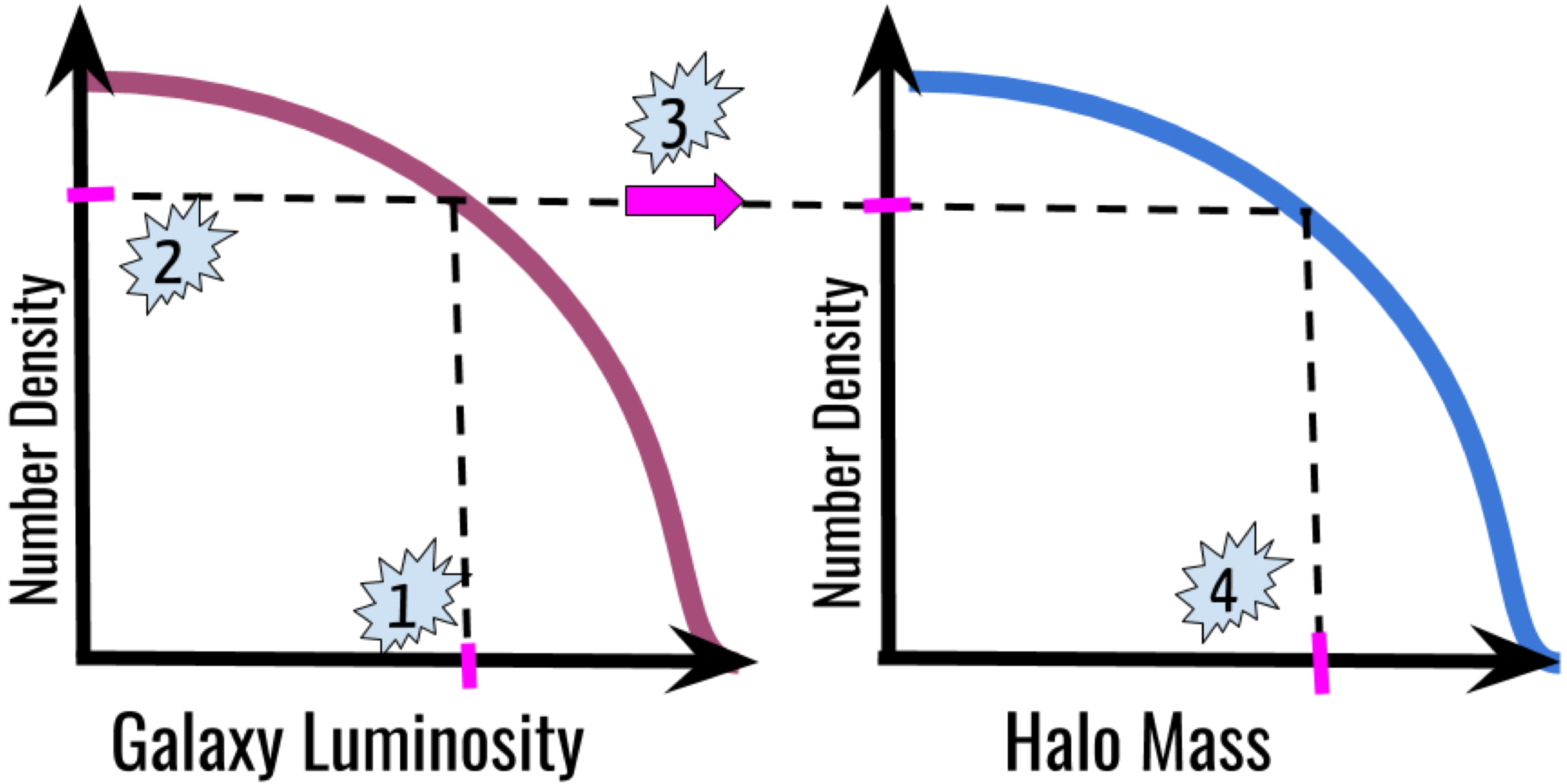
Use Cartoons

Special Projects



Number Density
Galaxy Luminosity
Number Density
Halo Mass
1
2
3
4

*Made With Google Drawings



A few tips

Avoid Green

Special Projects

A few tips

Don’t mess with fancy fonts

Special Projects

Final tip

Do not make a .ppt over 50mb

Special Projects

Papers

Presentations

Special Projects

Papers

Presentations

Posters

Papers

Presentations

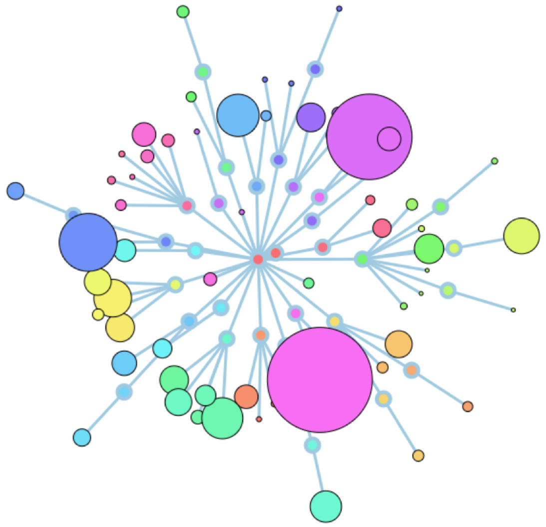
Websites

Papers

Presentations

Interactive objects

A ‘special project’ is how I made the **transition** from astronomy to data science



http://jpiscionere.github.io/Sequences_sunburst.html

Artist Tags

User Searches

Papers

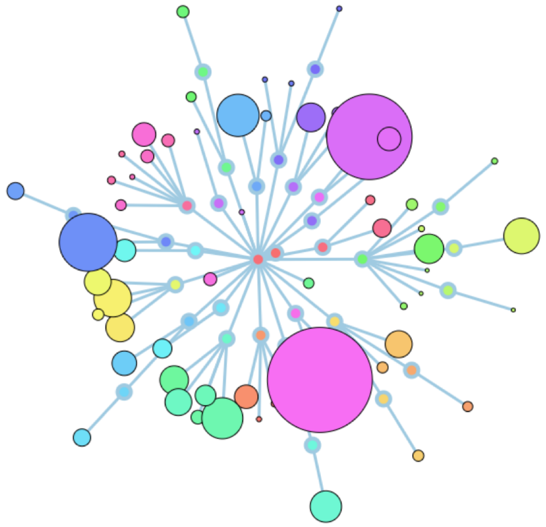
Presentations

Posters

General Tips

- First time- use a .ppt template
- Simple Background
- Edit on a simple printed page
- Don't print out full glossy – too heavy
  - Fabric is even better!
- Have printed handouts

Have fun and care about your audience



Resources

- ‘Show me the numbers’ by Stephen Few
  - https://nces.ed.gov/programs/slds/pdf/08_F_06.pdf
- Data Visualization by Jill P. Naiman
  - https://uiuc-ischool-dataviz.github.io/spring2019online/
- My slides on how to rip off a cool visualisation and put it on your website
  - http://jpiscionere.github.io/d3.pdf