



Everyone Counts: Interpreting Data

College of Letters & Science Student Academic Affairs,
Academic Information Management

Susan M. McMillan, April 2025



Welcome to “Everyone Counts: Interpreting Data”

Icebreaker: Would you rather find an answer to a data question using...

... a data table?

... or a graph ?

If you feel comfortable, please share with the group either verbally or in the chat.



What to expect today...

- What am I looking for in data?
 - Ideally, you're looking for the answer to a specific question
 - Structure of tables and graphs
- Is there a pattern?
 - Types of things to look for
- How can I be sure the pattern is "real"?
 - Types of interpretation difficulties and how to mitigate
- We'll work through a set of examples with "asides"



Data interpretation in general

- "The phenomena that rule our world are by definition complex, multifaceted and mostly difficult to grasp..."
 - Giorgia Lupi, "Data Humanism: My Manifesto for a New Data World"
- Data – no matter it's form – is only a representation of real-world people, activities, or things
- In tables and graphs, we are primarily looking at two dimensions of reality, sometimes three
- Interpreting data requires attention to the complexities, even when they are not immediately obvious.



Humans are good at patterns

- They help us organize complex information to...
 - Make predictions about what comes next
 - Form (simplified) decision rules
- When we “use data” for informed decision making, we’re often looking for a pattern that will suggest an answer to our analytic question.
- We sometimes (maybe often) see patterns that aren’t there...



Generic approach to data resources

- Start with the text
 - Title, notes, labels
- Examine the structure
 - Rows and columns for tables
 - Axis values
 - What encodes the data: Text? Numbers? Lines? Dots? Area?
- Content
 - Examine the values in the table or graph
 - Can you identify patterns?
- What is missing?



Identify patterns in data – always in context

- What is the range of values?
- Is there an obvious trend?
 - Is it regular? Up? Down? Stable?
 - Does it fluctuate regularly (as in seasonal data)?
 - Growing (or shrinking) rapidly?
- Are groups of values important?
- What might be missing?

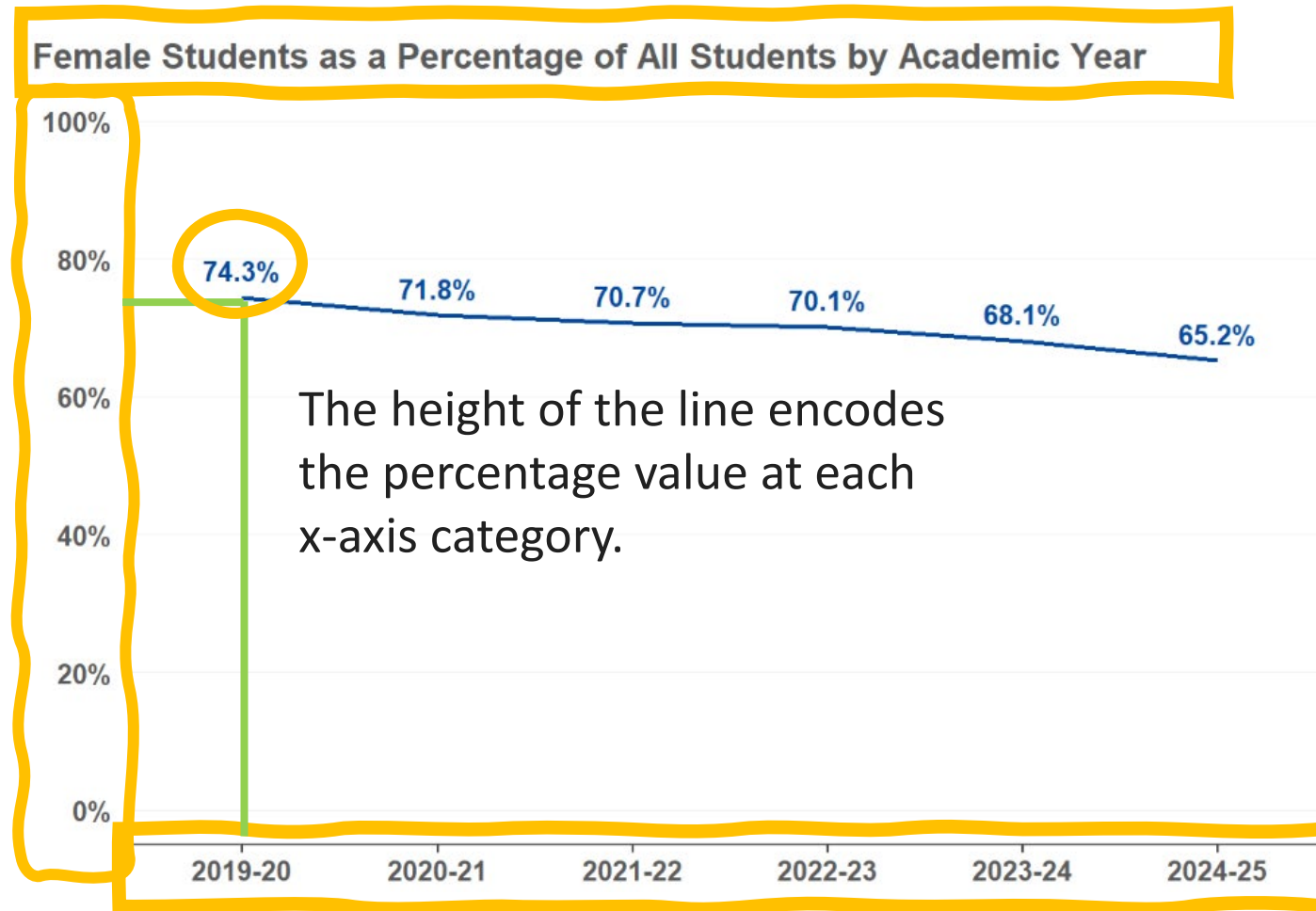


Example 1: Identify a trend

- Context:
 - We are in charge of a program that has been heavily dominated by female students.
 - We have been working a recruiting plan to diversify and get more male students, and we want to know...
- Is it working?
- Or do we need to change our tactics?



Identify a trend – in a graph





Aside: interpreting trends in percentages

- Every time you see a percentage, wonder about the parts in its calculation
 - Literally, every single time
- Behind every percentage is a fraction
 - Numerator ("top number")/denominator ("bottom number")
 - Multiply that by 100 to get the "per 100" interpretation we're used to
- Understanding our trend requires attention to both parts
 - Have the numbers of females changed?
 - Have the numbers of males changed?
 - Both?



Identify trends in a table

Our Students by Legal Sex*

Academic Year	Female Count	Female Percent	Male Count	Male Percent	Total Count
2019-20	156	74.3%	54	25.7%	210
2020-21	204	71.8%	80	28.2%	284
2021-22	241	70.7%	100	29.3%	341
2022-23	276	70.1%	118	29.9%	394
2023-24	312	68.1%	146	31.9%	458
2024-25	339	65.2%	181	34.8%	520

*All students selected either “Female” or “Male”.



Example 2: A different type of trend

- Context:
 - We have been appointed to a committee to see if we can simplify our college GPA policies
 - We want to know...
- What is the relationship between the cumulative and major GPA?
- Could we get away with fewer GPA requirements?



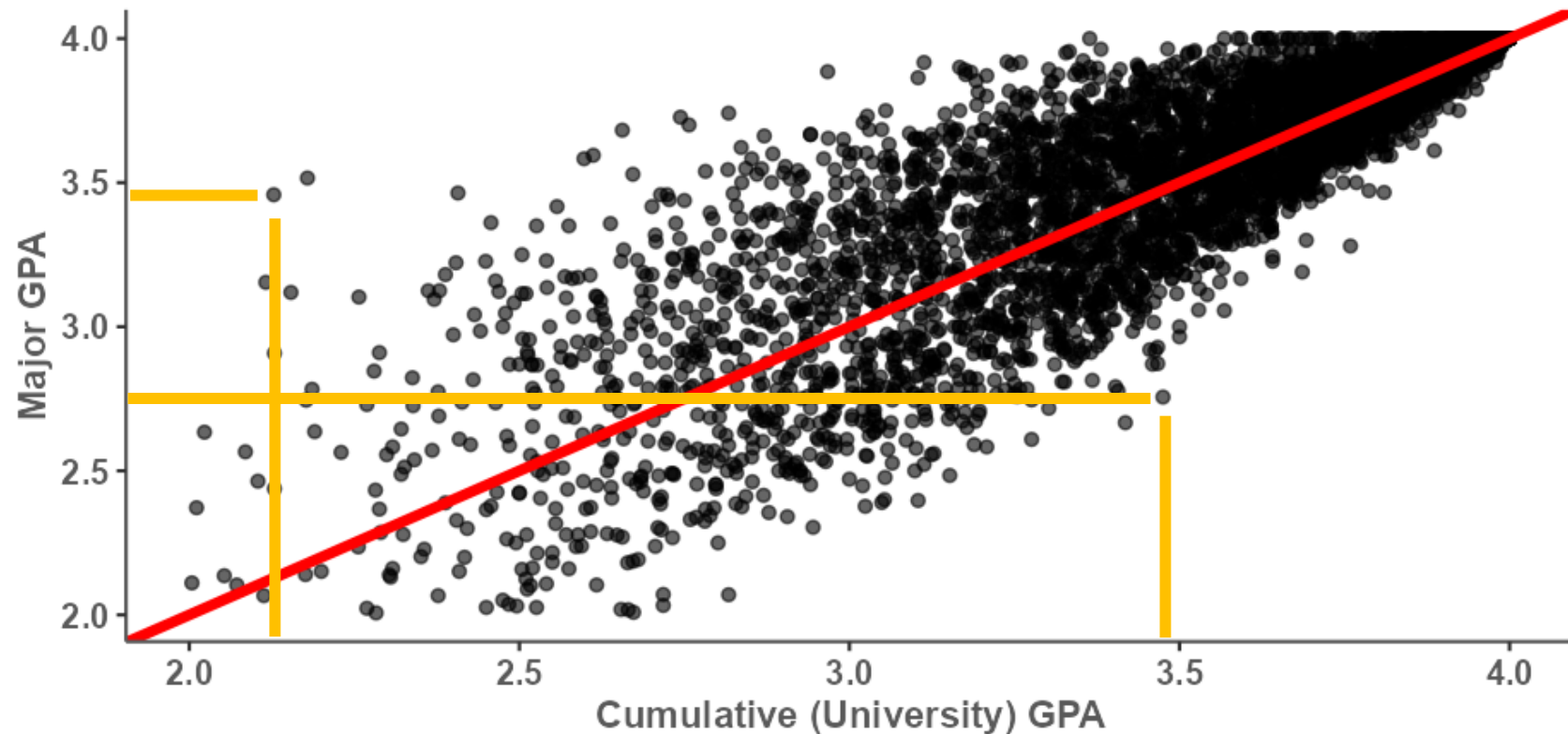
Another type of trend - association

Each dot encodes the cumulative and major GPA of an L&S undergraduate degree recipient.

(And this is really a subset of the data.)

Final GPA of L&S Undergrad. Degree Recipients

Award Academic Years 2010-11 through 2022-23



Solid diagonal line is the identity line ($x = y$)



A correlation table...

Pearson's Correlation Coefficients Between GPAs of L&S Undergrad. Degree Recipients,
2011-2023

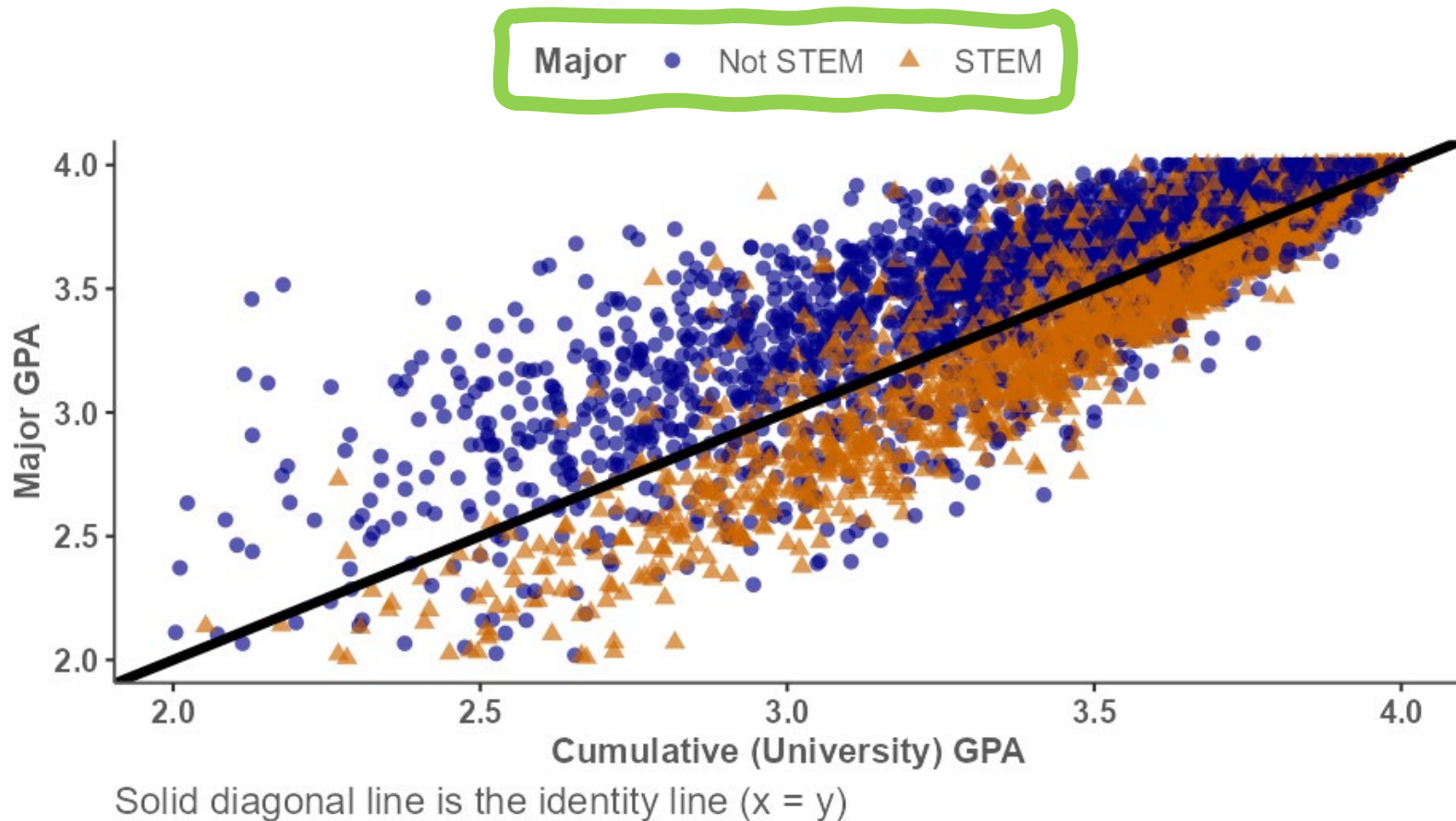
	Cumulative	LAS I/A	Major	Upper-level Major
Cumulative	1.000	0.957	0.879	0.841
LAS I/A	0.957	1.000	0.899	0.886
Major	0.879	0.899	1.000	0.948
Upper-level Major	0.841	0.886	0.948	1.000

Are groups important?

Now, in addition to marker location, the **color and shape** encode whether the graduate had a STEM major.
(still a subset of the real data...)

Final GPA of L&S Undergrad. Degree Recipients

Award Academic Years 2010-11 through 2022-23





Turning to common interpretation complications...

- Missing values
- Outliers (unusually high/low values)
- Small numbers of cases
- Confirmation bias (cognitive)



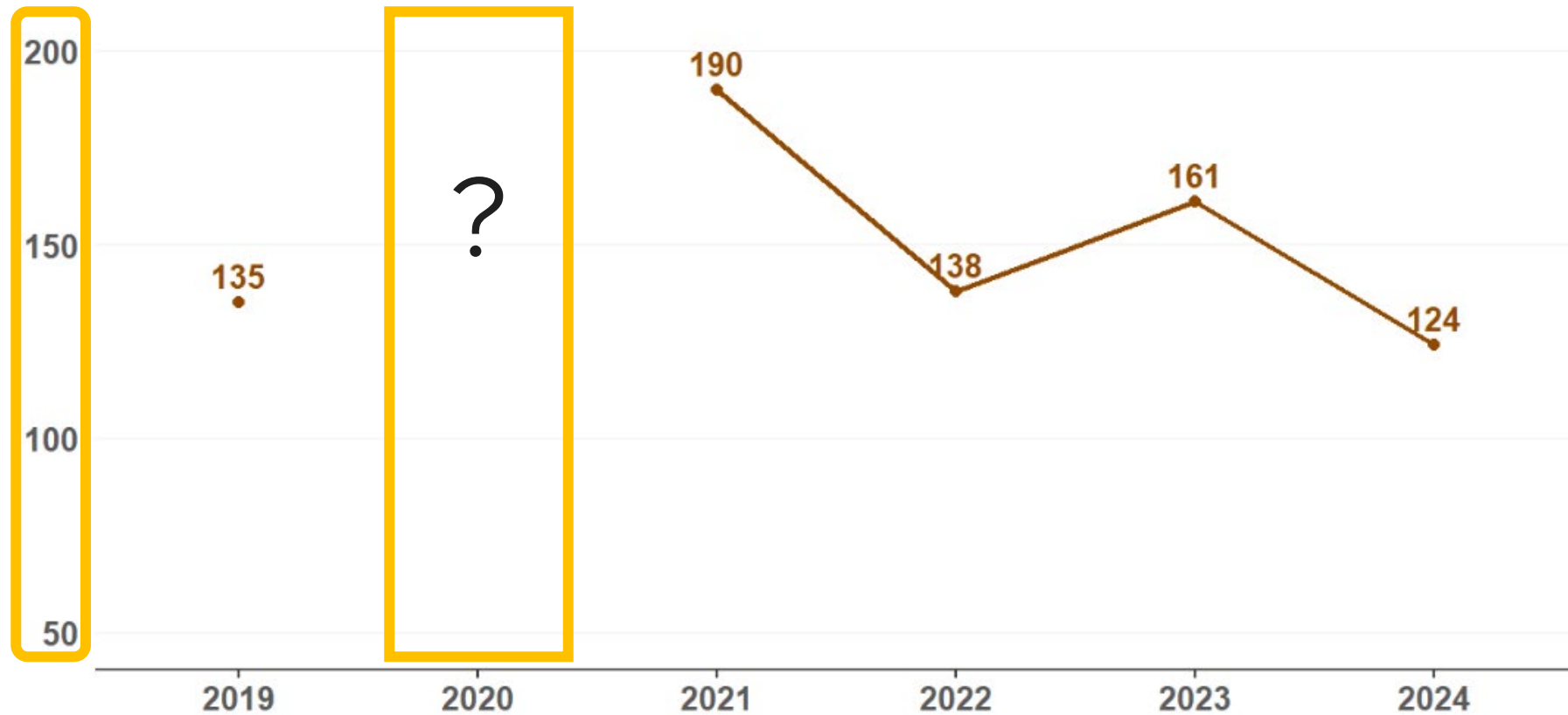
Example 3: Missing values (and outliers)

- Context:
 - We have a program for L&S students, and we want data to help us...
- Predict the number of students in our next cohort



Missing data (and possible outlier)

Number of Students by Cohort



Aside: outliers and averages, a well-known example



Current Cumulative GPA Credits [?]	Current Cumulative GPA [?]	Current Academic Standing
15	3.5	Good Academic Standing

For best accuracy, please make sure you enter above your current 4-digit CUM GPA (e.g., 2.979, 3.200, etc.) that appears on your DARS or unofficial transcript.

Current Term

Class:	Grade:	Credits:
Class 1	AB (3.5)	3
Class:	Grade:	Credits:
Class 2	AB (3.5)	3
Class:	Grade:	Credits:
Class 3	AB (3.5)	3
Class:	Grade:	Credits:
Class 4	AB (3.5)	3
Class:	Grade:	Credits:
Class 5	F (0.0)	5

Results:

Term GPA	New Cumulative GPA	New Academic Standing
2.470	3.027	Good Academic Standing



Example 3: Small numbers

- Context:
 - We have a potential donor who wants to support US Military veterans and wants to know...
- What is the retention rate for veterans on campus?



Aside: interpreting small numbers of cases

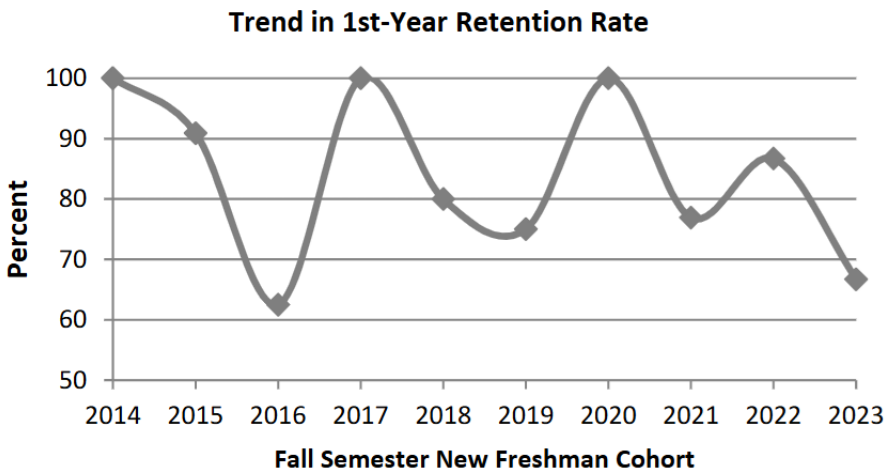
- Difficult to distinguish a potential pattern from randomness
- Problems with reliability (consistency) of any perceived pattern
- Observed pattern might be a misleading part of a larger pattern
 - "9 out of 10 dentists..." could be the odd ones out in "91 out of 100 dentists..."
- Description vs. inference
 - If we are only describing a situation, small numbers are not such a problem
 - Inference is more problematic, for the above reasons
 - Often, we are describing to infer



Small numbers...

Retention Rates of New Freshmen by Fall Semester Entrance Cohort

New Freshmen who Identify as U.S Military Veterans



Notes: Fall semester entrance cohorts include new freshmen who enrolled in the fall or in the immediately preceding summer term.

Data, Academic Planning & Institutional Research, Office of the Provost, clh

Source: <https://data.wisc.edu/retention-graduation/>

Fall	Cohort	Retention Rate After -		
	Headcount	1 Year	2 Years	3 Years
2014	12	100.0	91.7	75.0
2015	22	90.9	72.7	77.3
2016	8	62.5	62.5	50.0
2017	12	100.0	75.0	58.3
2018	15	80.0	80.0	73.3
2019	8	75.0	62.5	50.0
2020	14	100.0	64.3	42.9
2021	13	76.9	76.9	53.8
2022	15	86.7	80.0	---
2023	9	66.7	---	---

2014 $100 \times (12/12) = 100.0$

2015 $100 \times (20/22) = 90.9$

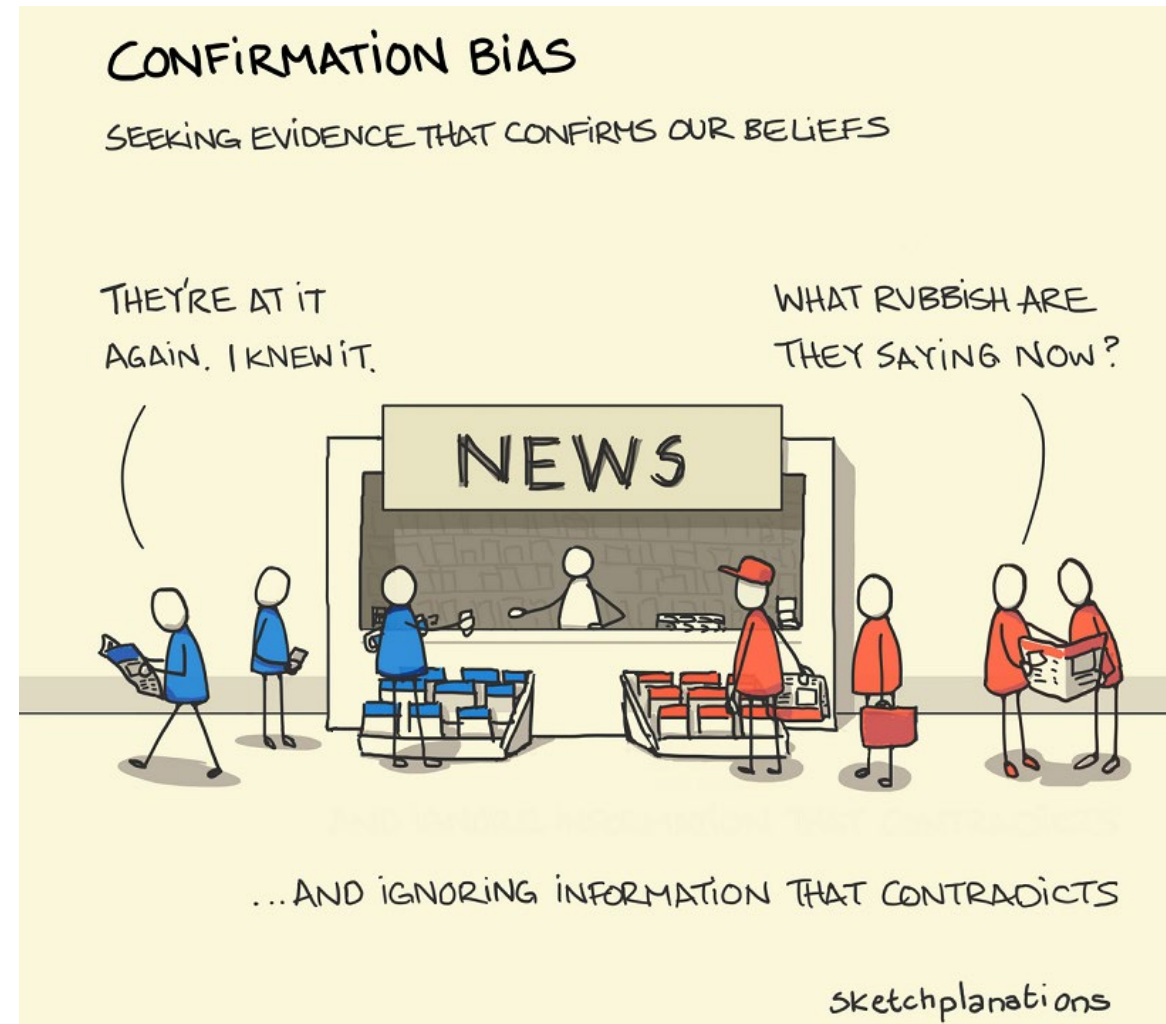
2016 $100 \times (5/ 8) = 62.5$



Example 4: Confirmation bias

- Context:
 - Someone in South Hall saw a Reddit post that “getting an A in an L&S course is super easy”.
 - That person happens to already think that grade inflation is rampant, so they request data that shows the trend in cumulative GPA at graduation for as many years as we have data...
 - Meanwhile a set of faculty members are confident that they have not in any way inflated grades in their subject, and that if grades have gone up, there are other explanations...
- We want to sort this out, knowing that...

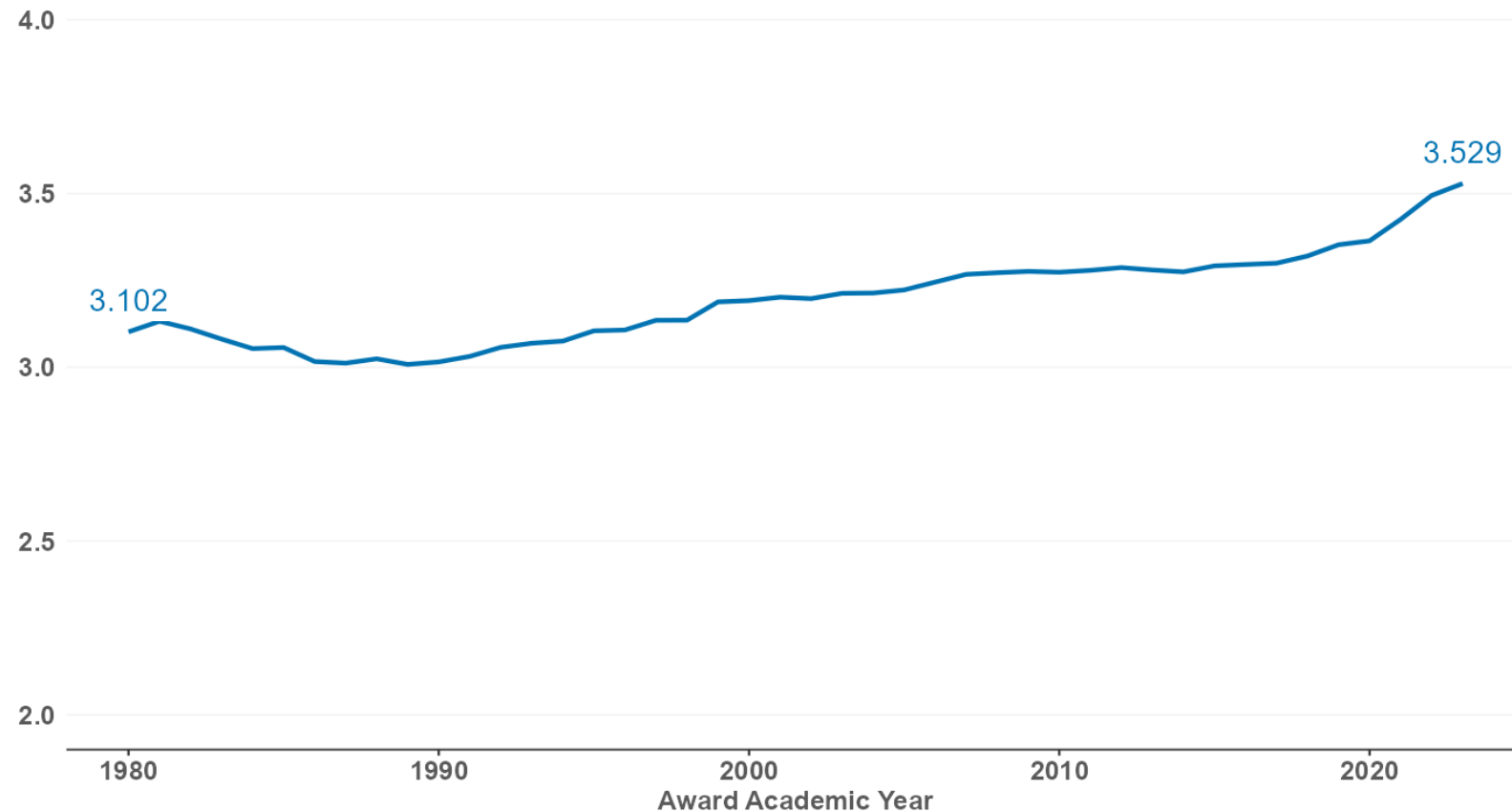
Cognitive shortcuts create interpretation problems





Do we have a problem with grade inflation?

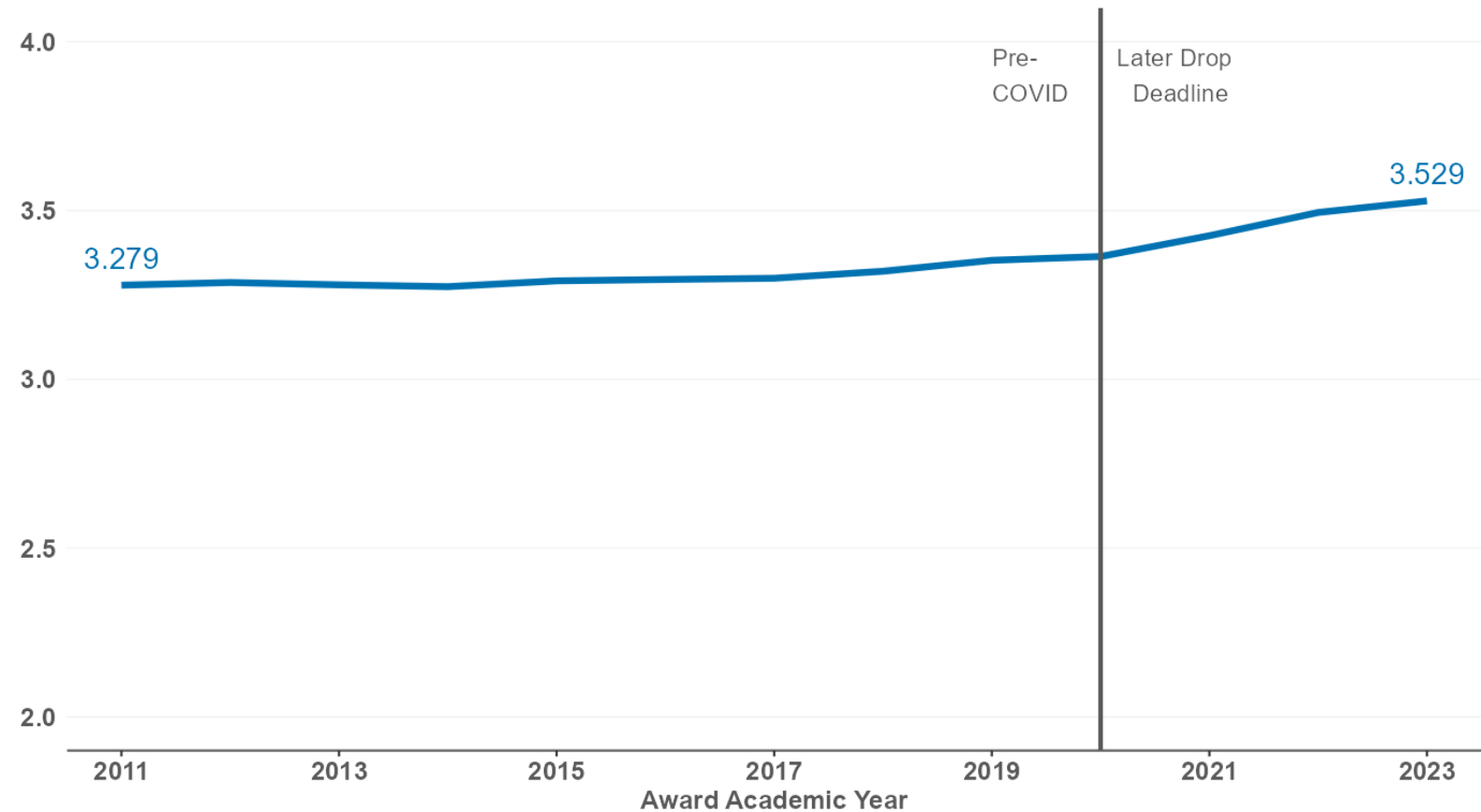
L&S Undergraduate Average Cumulative GPA at Graduation
1980-2023





Change the time frame, plus...

L&S Undergraduate Average Cumulative GPA at Graduation
2011-2023





How to counteract interpretation difficulties

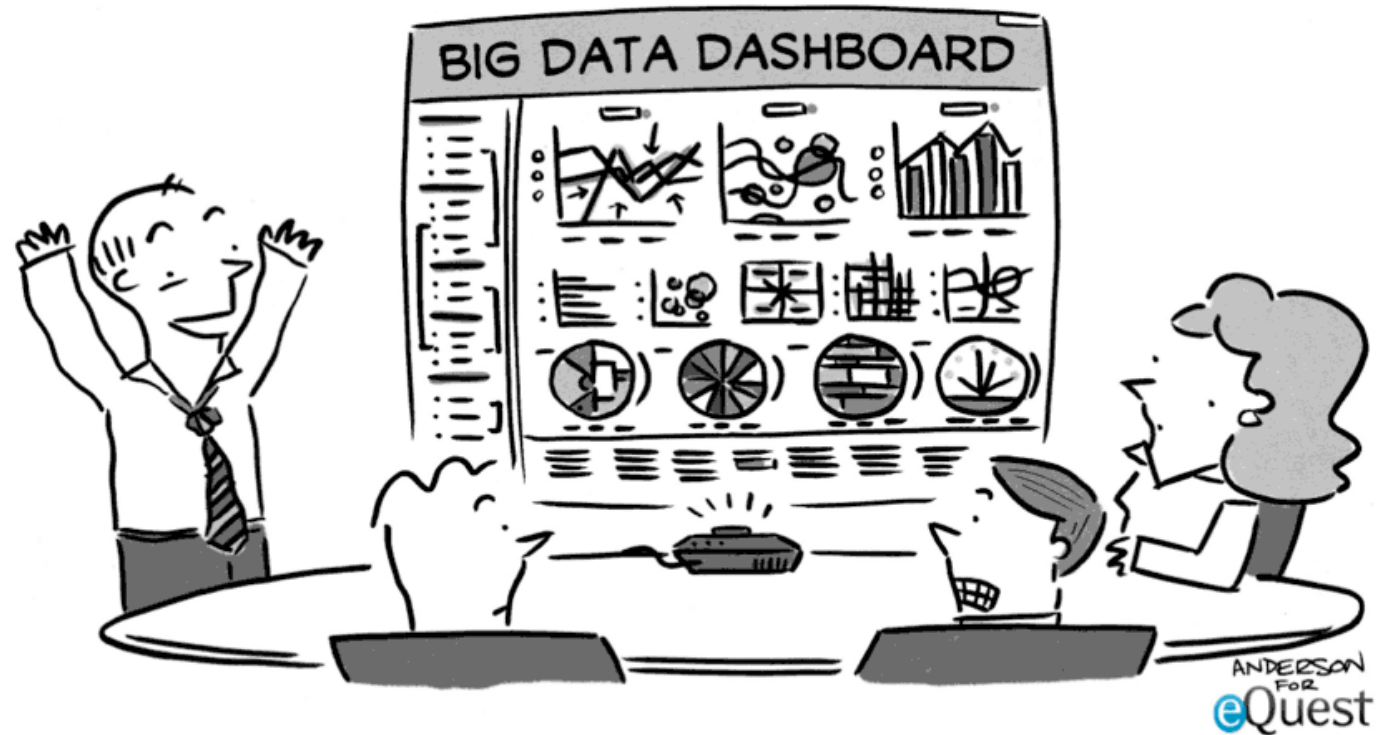
- Question your own assumptions about the...
 - Framing of the analytic question
 - The collection of the data
- Check for possibility that patterns are unduly influenced by outliers or small numbers of cases
- Consider whether tables and graphs exclude potentially relevant information
 - Is an important “confounding variable” (like group membership) missing?
- Pay attention to scale ranges
 - Does your conclusion change if the range is lengthened or shortened?



Your judgement matters

- All data analysis has imperfections
 - Remember the simplifications that have been made
 - Be very clear about who or what is included and why
- Work to develop good judgement in interpreting data
 - Pay attention to all the parts of a table or graph
 - Look for patterns
 - Look for common problems that might affect your interpretation
 - Attempt to mitigate those problems

Time to go...



"After careful consideration of all 437 charts, graphs, and metrics, I've decided to throw up my hands, hit the liquor store, and get snocked. Who's with me?!"

Source: <https://www.equest.com/cartoons/cartoons-2013/big-data-dashboard-dizziness-a-trendy-tool-with-little-utilization/>