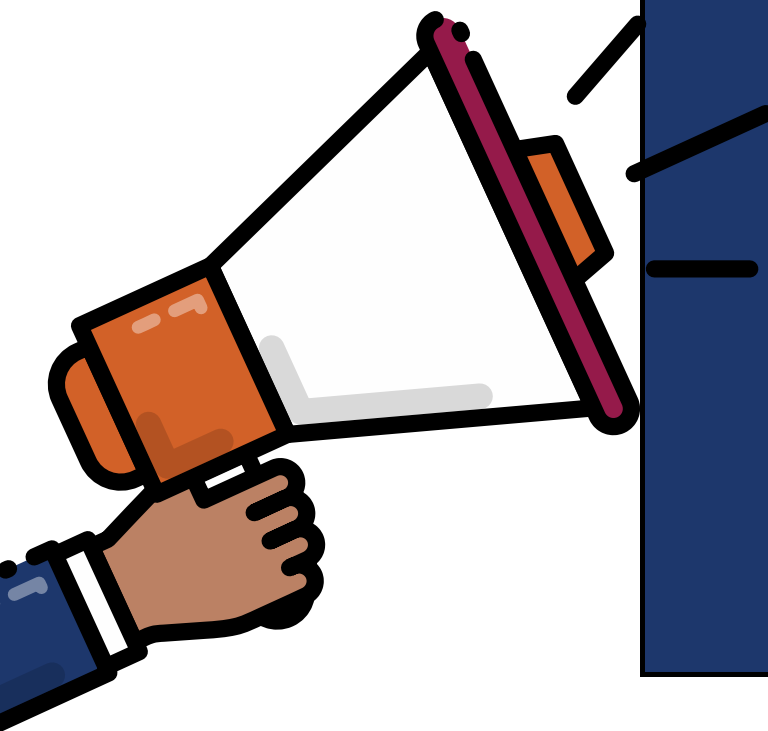


Welcome

Before we get started...



- Display Name: First & Last
- Program is recorded
 - Turn camera off to not be in recording
- Questions
 - Use chat function throughout session
 - Raise hand feature



OFFICE OF INDIAN EDUCATION

OIE Virtual Data Training

April 24, 2024



Today's Learning

We will discuss

- 01 Understanding Data and Visualizations
- 02 Selecting the Right Visualizations
- 03 Design Principles for Effective Visualization
- 04 Overview and Resources

Introductions



Russel Potter, Ph.D.

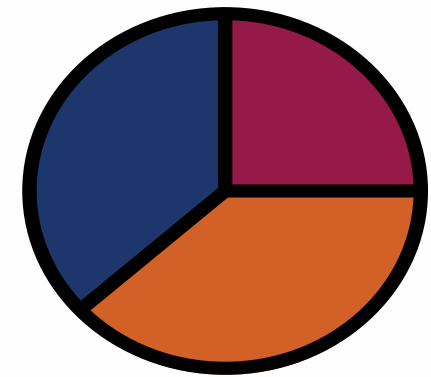
Data Director for the Office
of Indian Education

In the chat, please share

- Your name
- Role / District

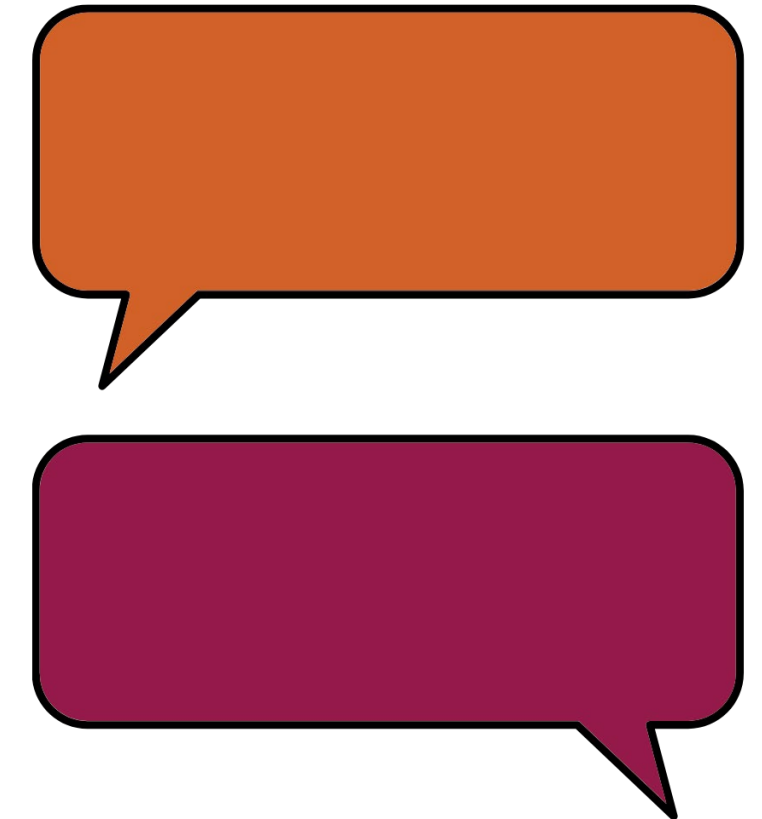


Understanding Data and Visualization



Understanding Your Data

Race	2022	2023
Asian	34295	37936
Black	63473	72072
Hispanic	526978	568407
Native American	47315	53218
Pacific	4216	4715
Multiple Races	44174	51874
Unlisted	98	6196
White	398322	431819



Share in the chat:

- What are these data telling us?
- Is that clear?
- Why do we care?

- *Are these good data?*
- *Does anything seem odd?*

Understanding Your Data

Race	2022	2023
Asian	34295	37936
Black	63473	72072
Hispanic	526978	568407
Native American	47315	53218
Pacific	4216	4715
Multiple Races	44174	51874
Unlisted	98	6196
White	398322	431819

Identify the Big Question

- What do these data address?
- In their most basic form, what are these data for?
- How do they address it?

How has enrollment changed in the state, for each racial group?

Understanding Your Data

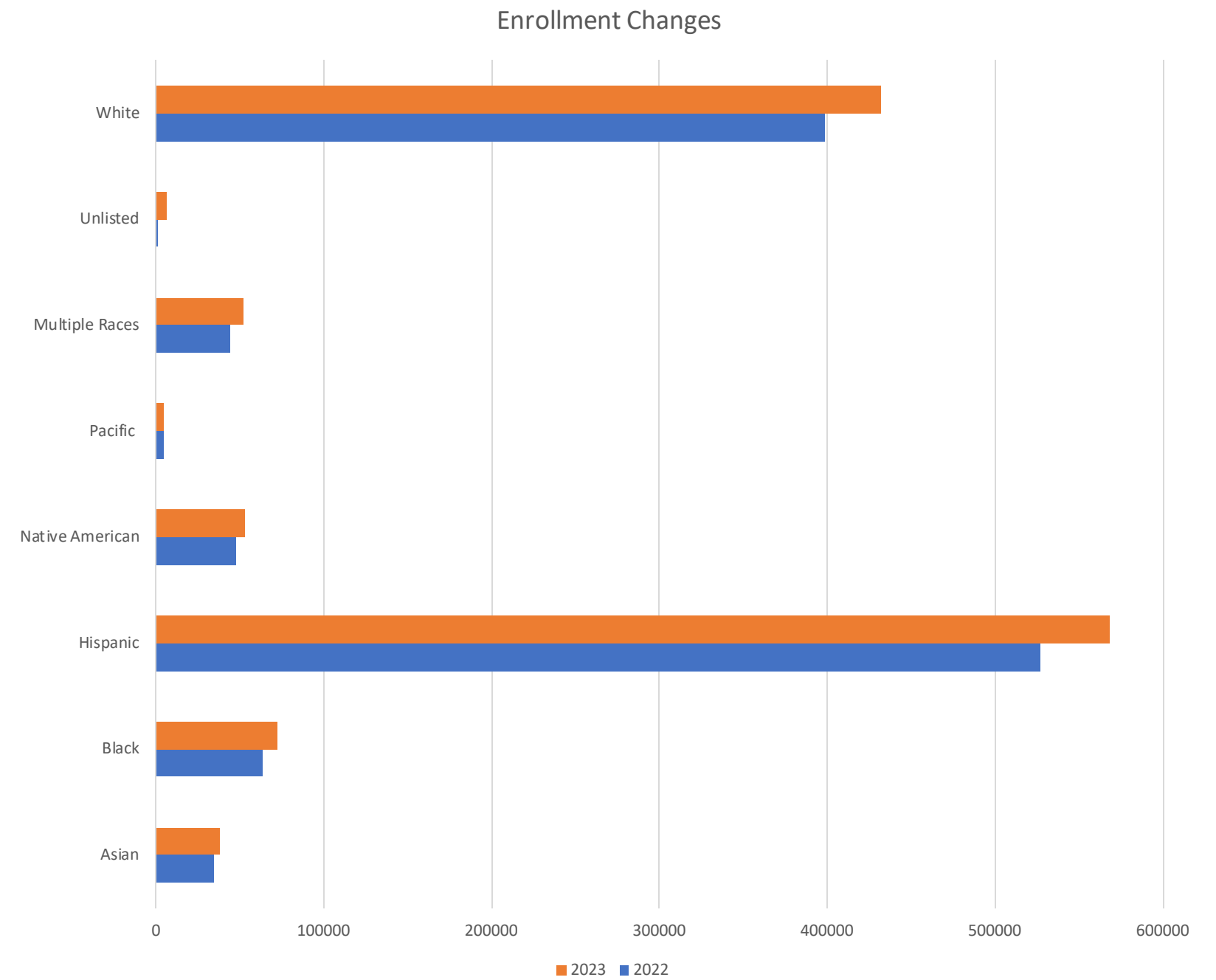
Race	2022	2023	Change	Percent Increase
Asian	34295	37936	3641	11%
Black	63473	72072	8599	14%
Hispanic	526978	568407	41429	8%
Native American	47315	53218	5903	12%
Pacific	4216	4715	499	12%
Multiple Races	44174	51874	7700	17%
Unlisted	98	6196	6098	6222%
White	398322	431819	33497	8%

How has enrollment changed in the state, for each racial group?

- Compare year to year
- Specify the difference
- Compare difference as proportion

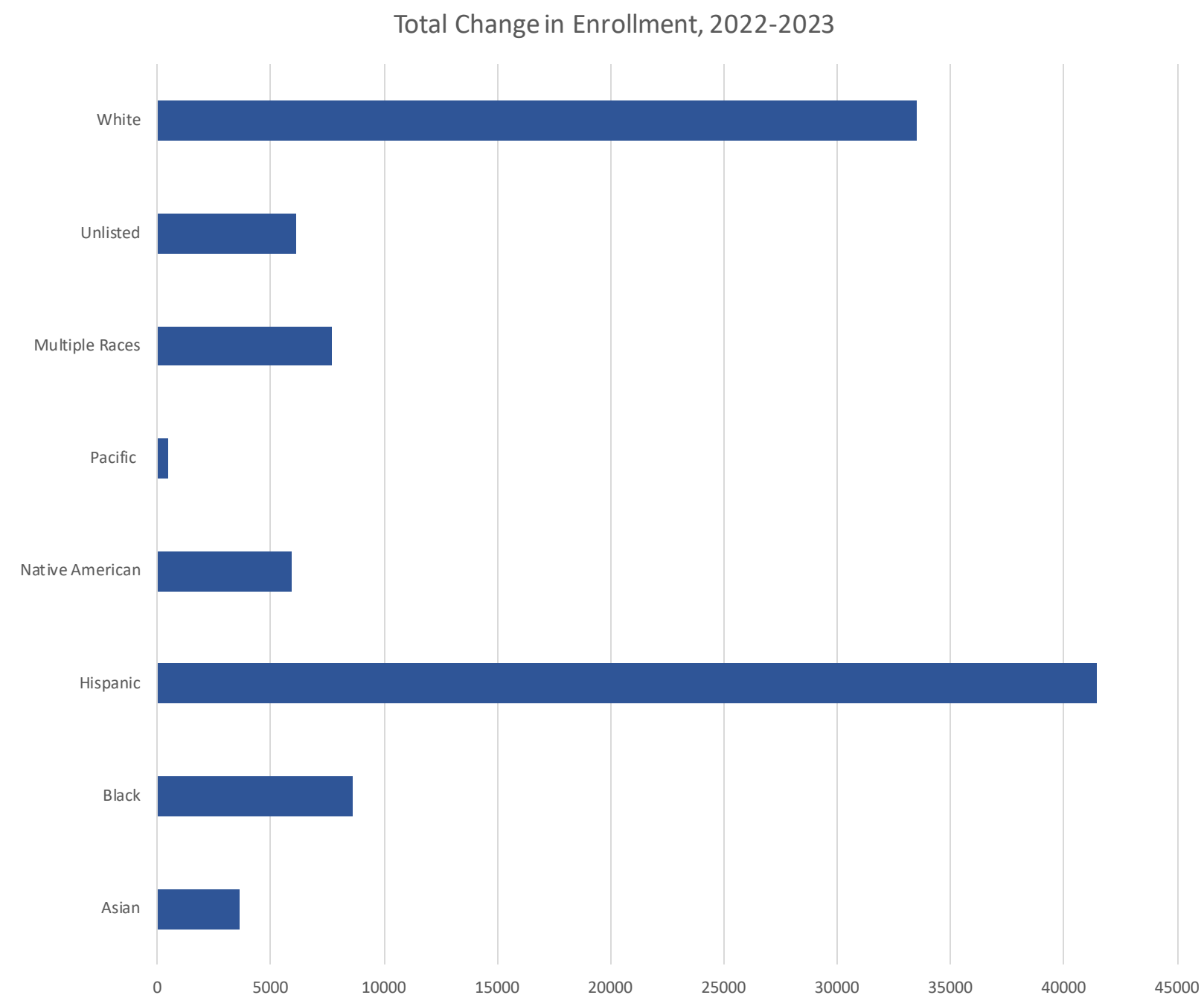
Year to Year Comparison

Race	2022	2023	Change	Percent Increase
Asian	34295	37936	3641	11%
Black	63473	72072	8599	14%
Hispanic	526978	568407	41429	8%
Native American	47315	53218	5903	12%
Pacific	4216	4715	499	12%
Multiple Races	44174	51874	7700	17%
Unlisted	98	6196	6098	6222%
White	398322	431819	33497	8%



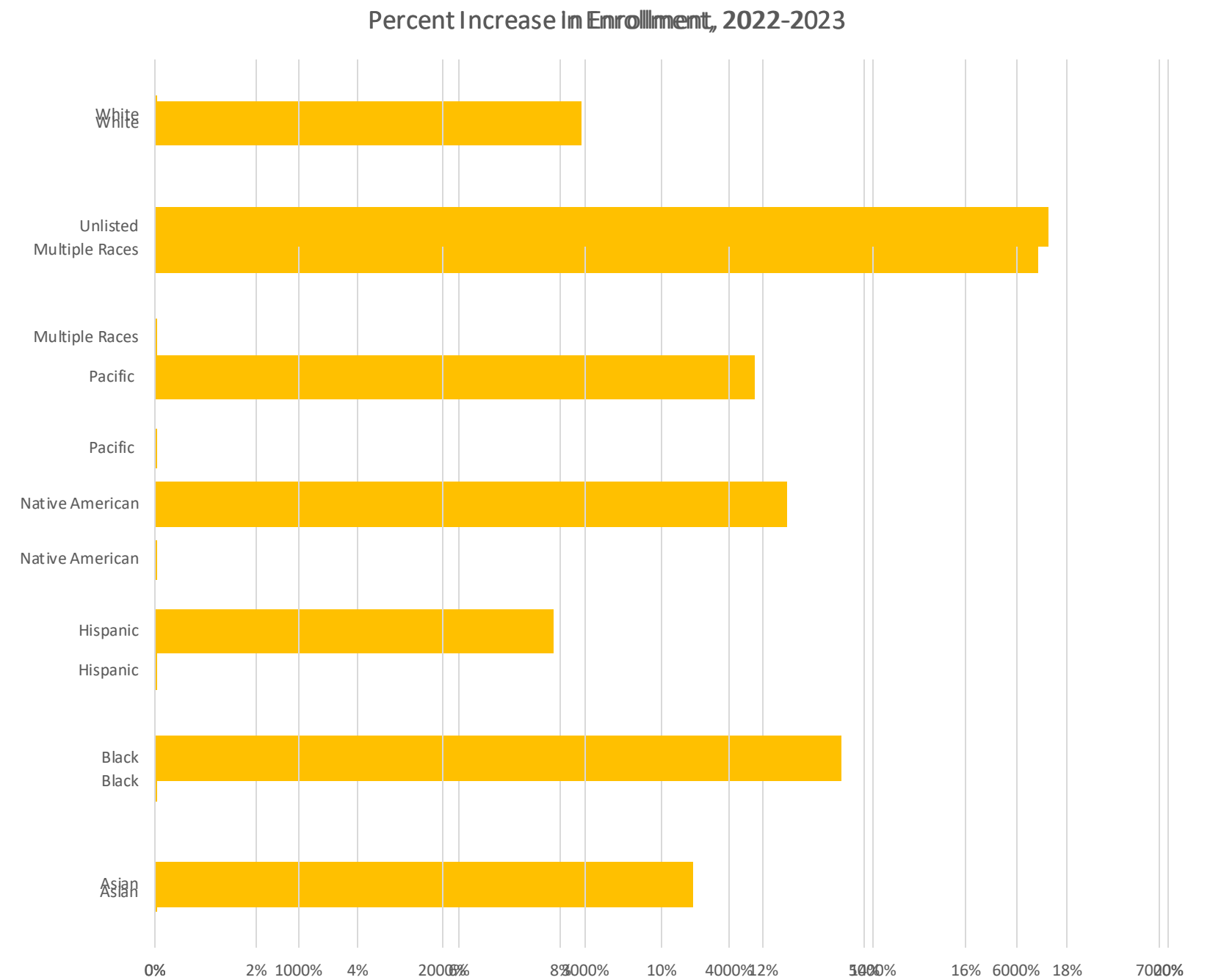
Raw Difference

Race	2022	2023	Change	Percent Increase
Asian	34295	37936	3641	11%
Black	63473	72072	8599	14%
Hispanic	526978	568407	41429	8%
Native American	47315	53218	5903	12%
Pacific	4216	4715	499	12%
Multiple Races	44174	51874	7700	17%
Unlisted	98	6196	6098	6222%
White	398322	431819	33497	8%



Proportional Difference

Race	2022	2023	Change	Percent Increase
Asian	34295	37936	3641	11%
Black	63473	72072	8599	14%
Hispanic	526978	568407	41429	8%
Native American	47315	53218	5903	12%
Pacific	4216	4715	499	12%
Multiple Races	44174	51874	7700	17%
Unlisted	98	6196	6098	6222%
White	398322	431819	33497	8%



Questions?

Raise Your Hand!

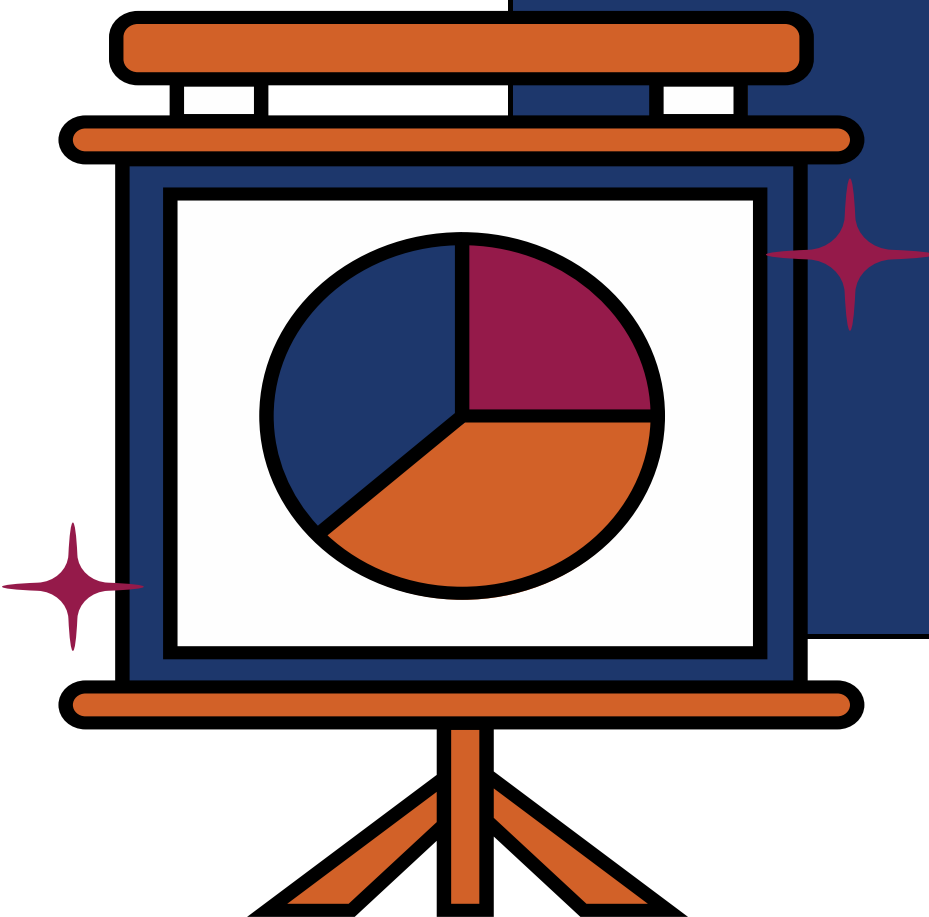


Drop in the Chat!





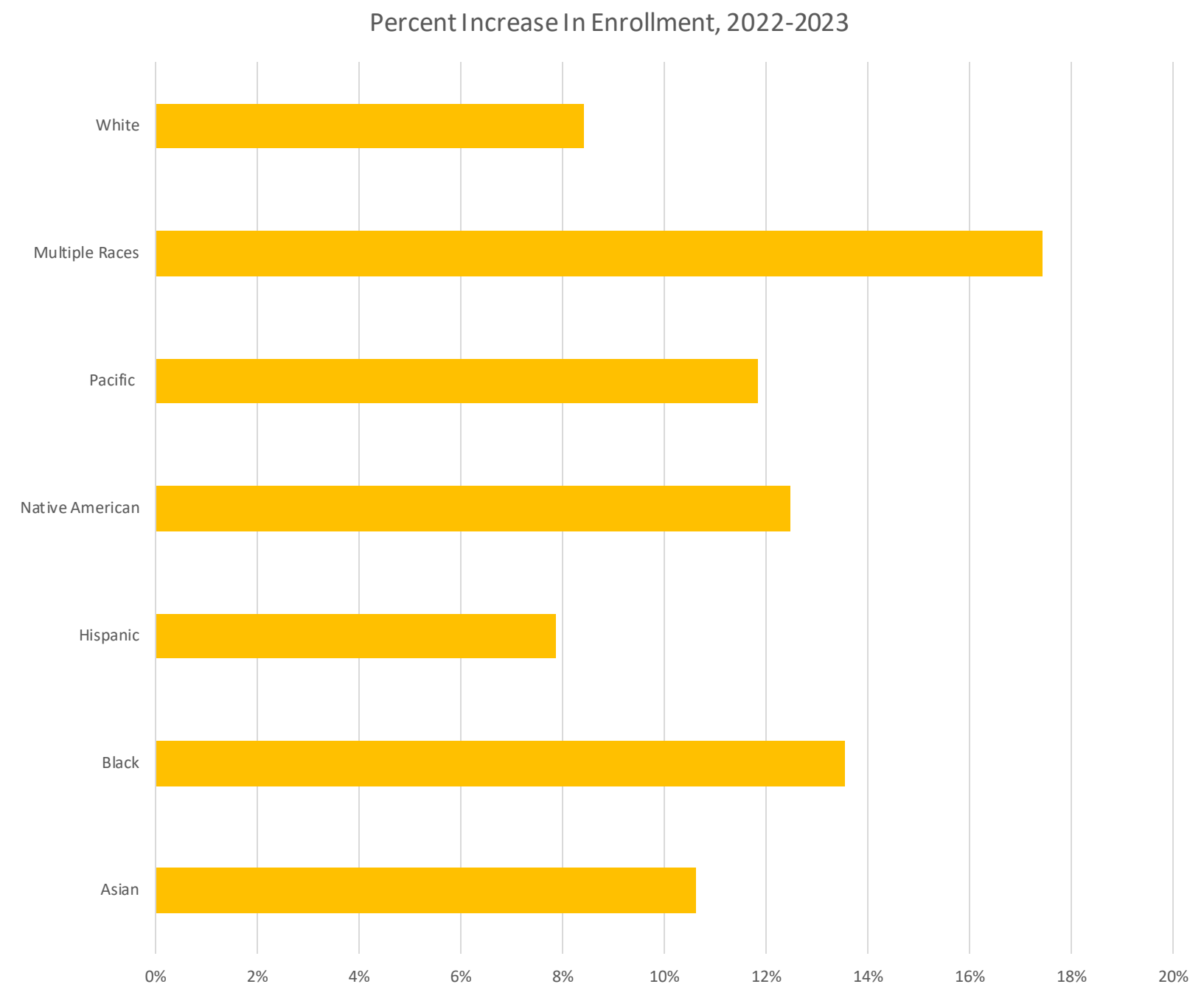
Selecting the Right Visualization



Using Proportions

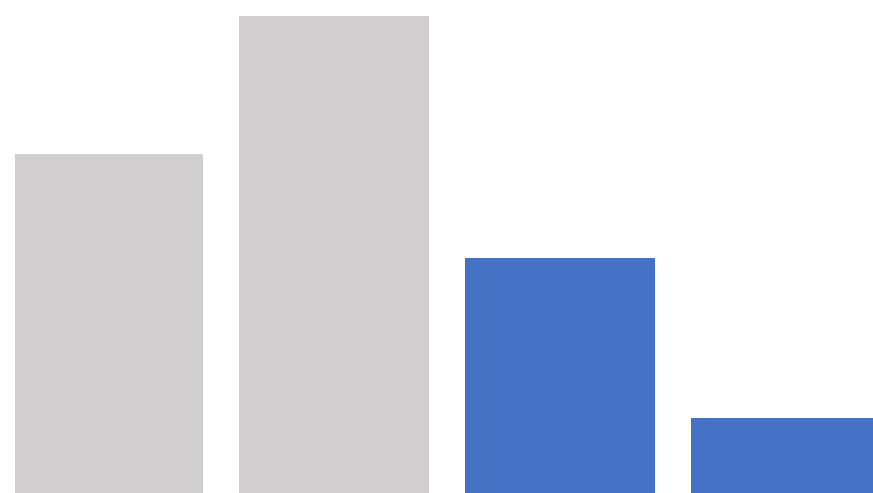
- In groups of different sizes or with different starting points...
 - What does a change of 10 mean in each group?
 - What does a change of 1000 mean in each group?
- If a change is not the same thing for *to the data story you're telling...*

PRINCIPLE:
Use proportions when dealing with populations of different sizes

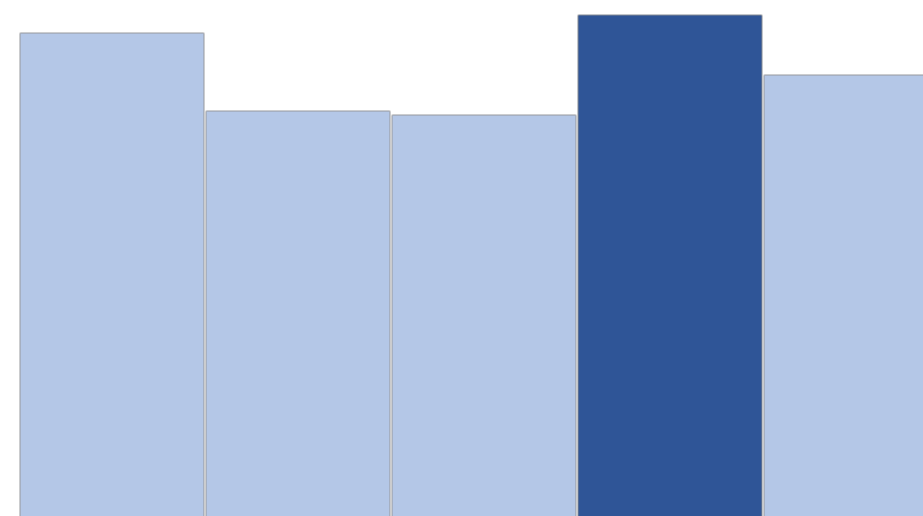


Two Numbers, Two Bars

- Discrete
 - **Only specific values in a range are possible**
 - Countable, but indivisible
 - Decimals usually don't mean anything
 - People, grade-levels, dates are usually *discrete*
- **Use bar graphs to display discrete data**
 - Bars don't touch
 - Data isn't continuous

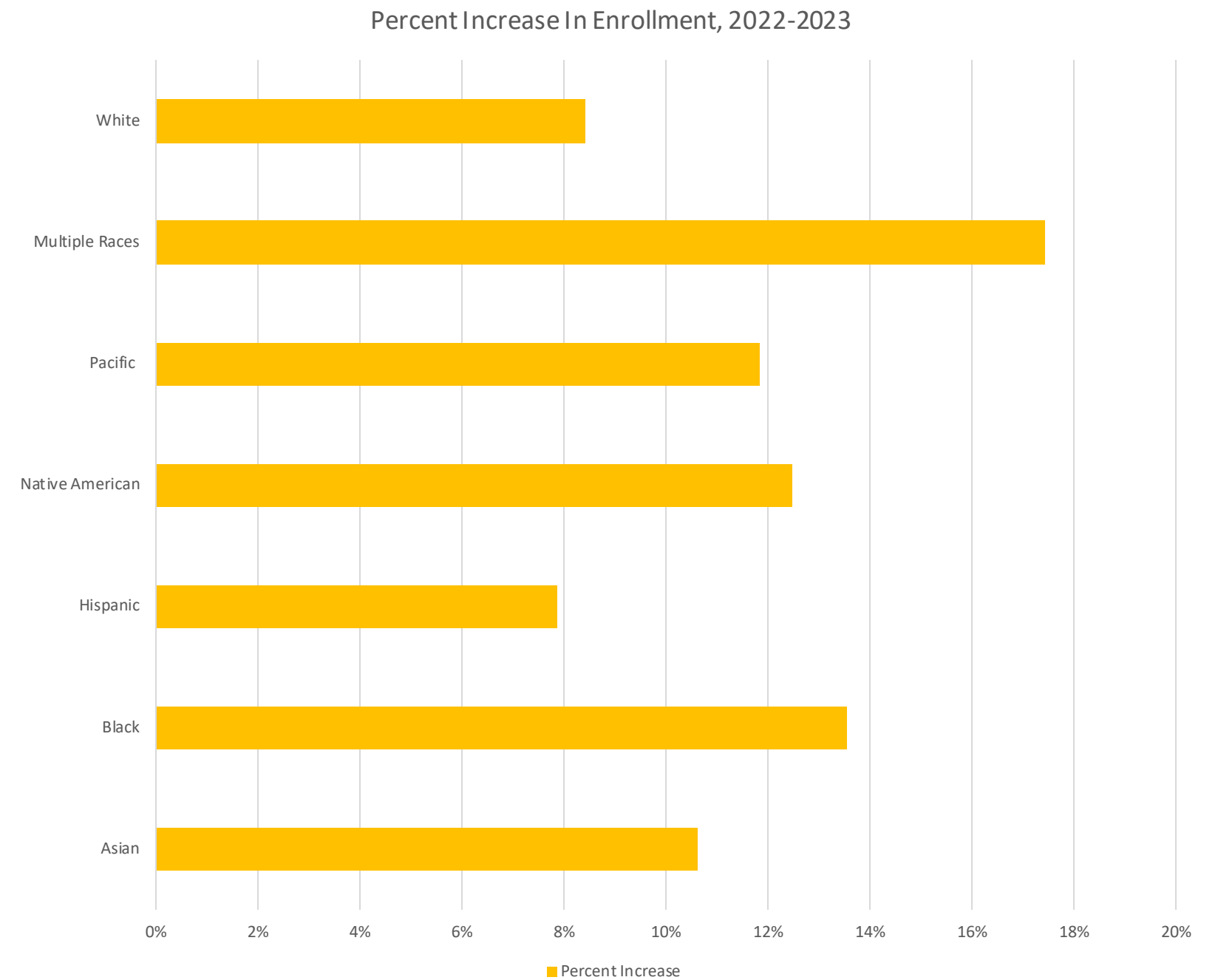


- Continuous
 - **Any value in a given range is possible**
 - Less countable, but infinitely divisible
 - Decimals have reasonable values
 - Time, money, age, height are usually *continuous*
- **Use Histograms for continuous data**
 - Bars touch
 - Data has no 'gaps'



Bar Graphs

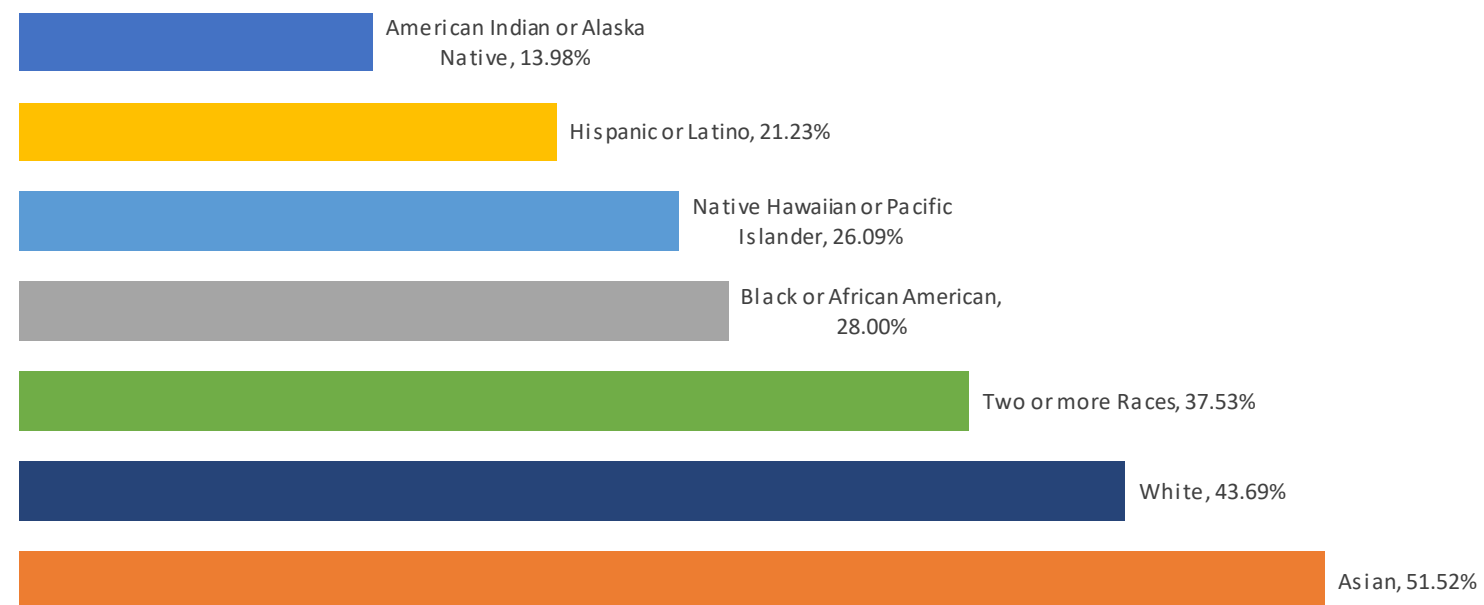
- Categorical against Discrete
 - Race and Counts
 - Age and Scores
 - Gender and Height
 - Describe with Numbers



Types of Bar Graph

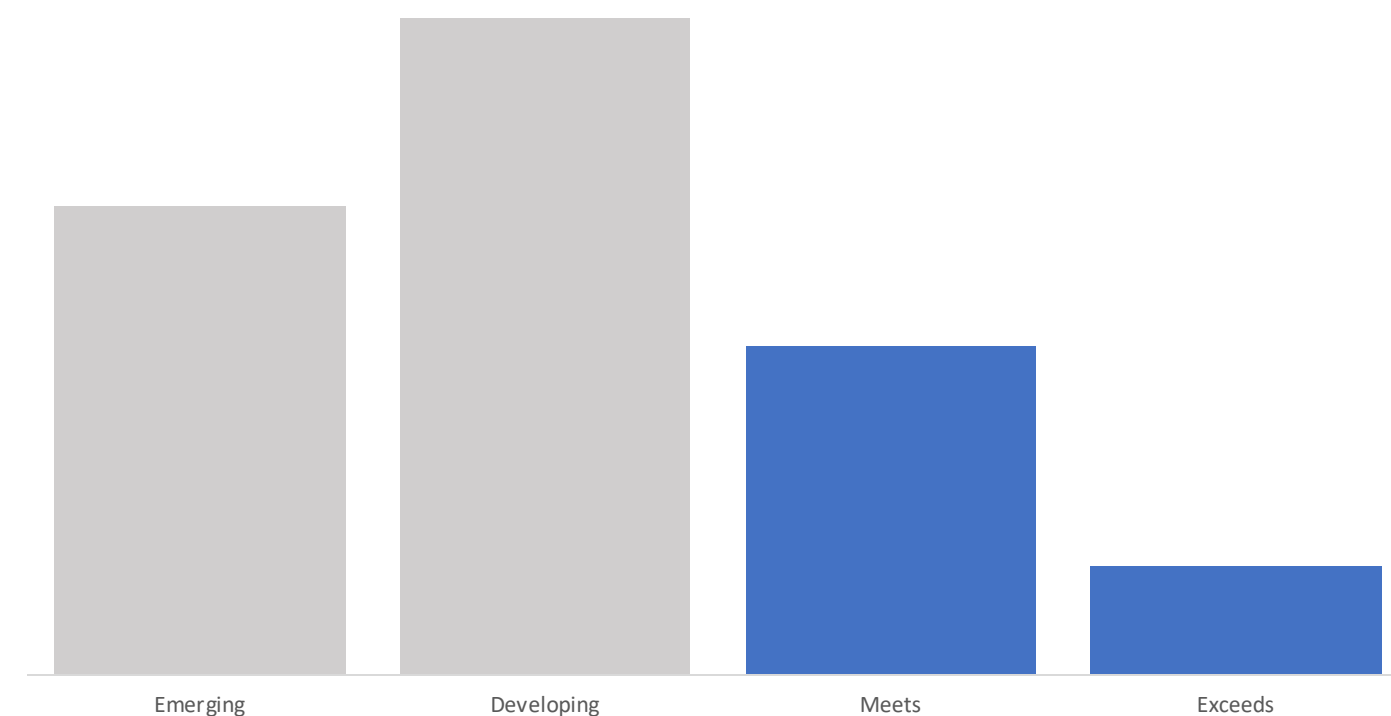
- Horizontal
 - Nominal and Categorical Data
 - Sort data by
 - greatest to least or least to greatest
 - target datum at the top

Perfect Proficient at High-Density Schools in SY23



- Vertical
 - Ordinal or Sequential data
 - Order time-data left-to-right
 - Groupings in logical order
 - If you have buckets, use a vertical bar chart

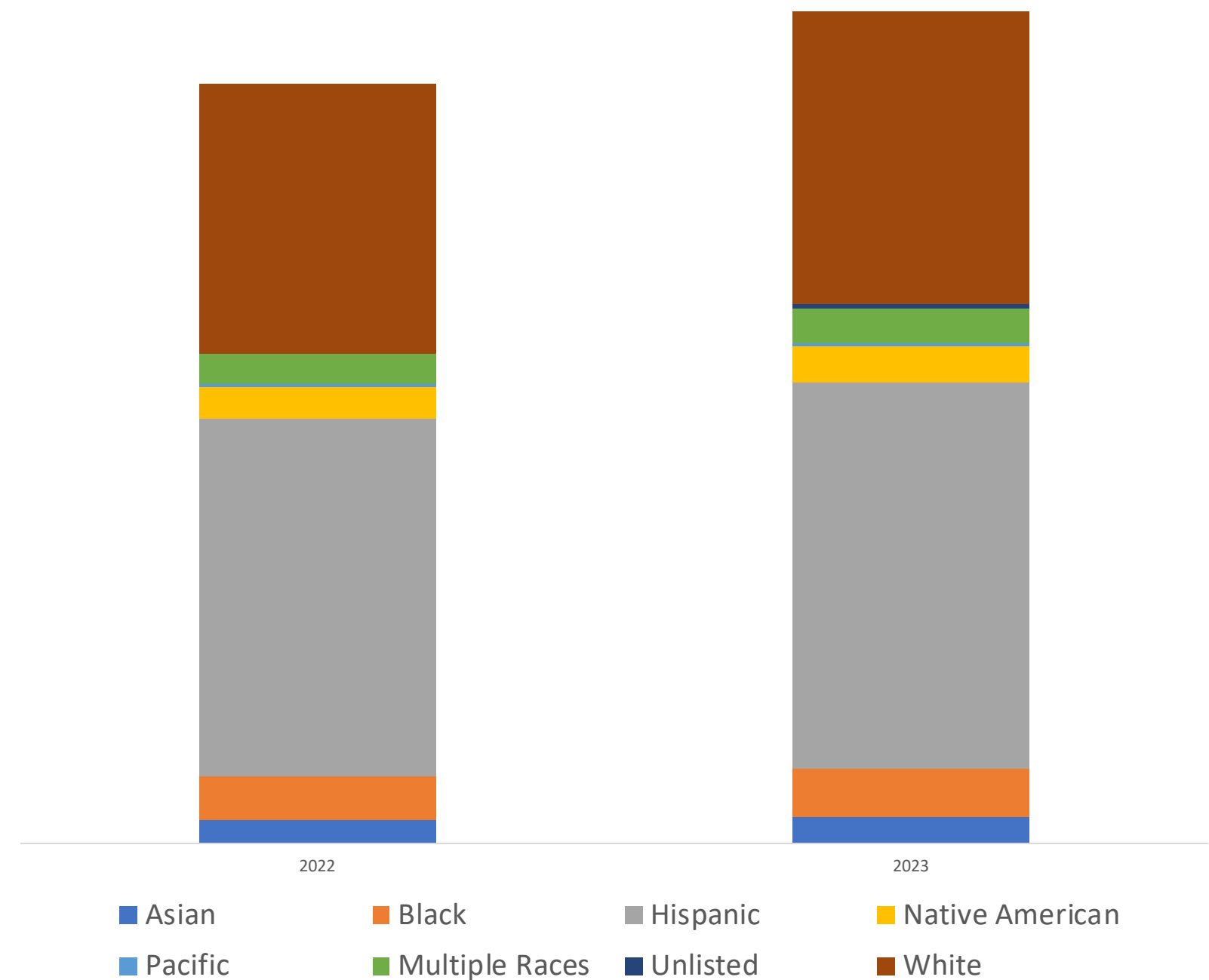
Ratio of Students Passing ELA



Stacked Bar Graphs

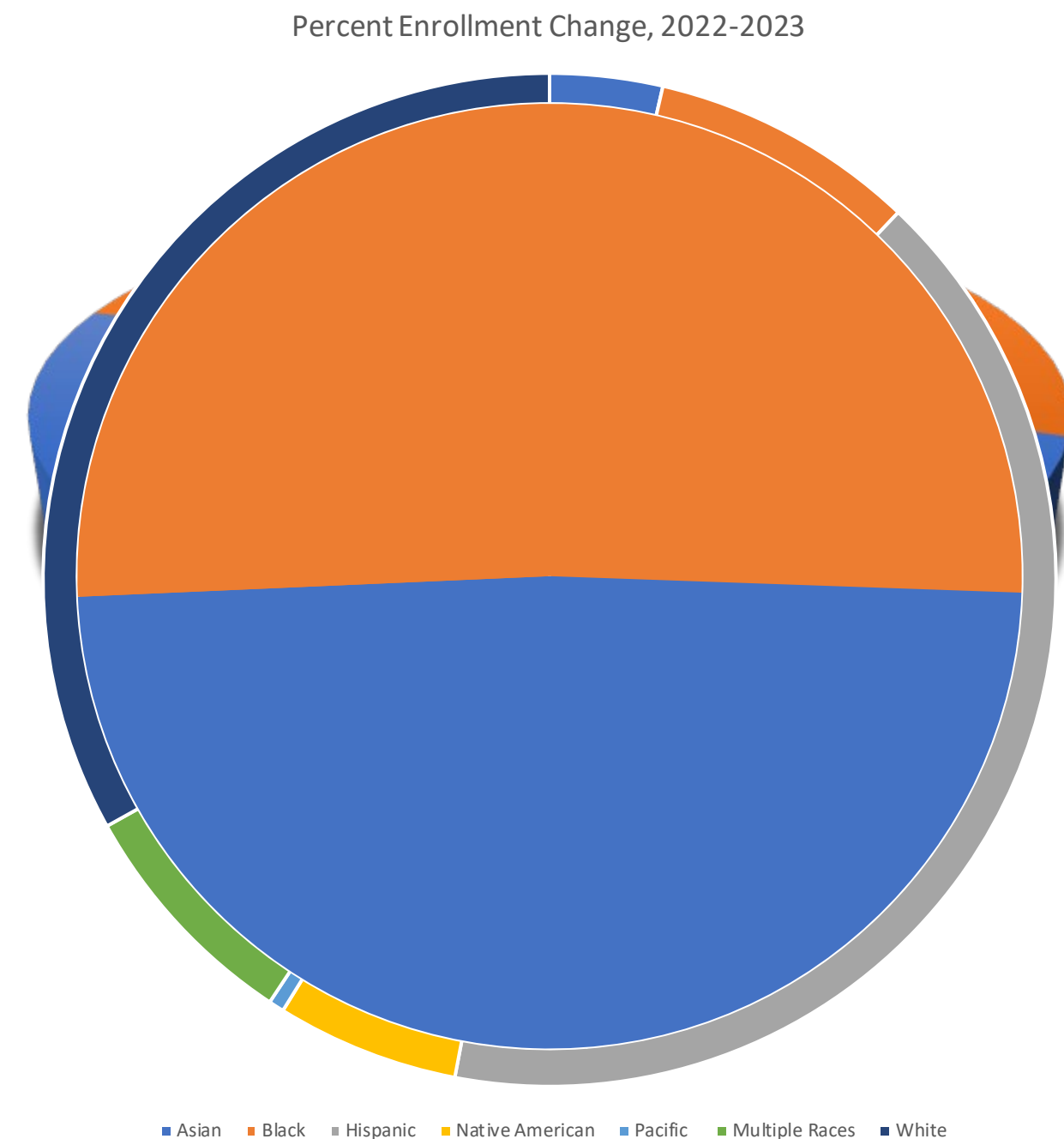
- When the total or aggregate is important but comprises smaller categorical groups.
- All the data for a value in one column.
 - Illustrates the composition of the total value
 - May occlude the details of parts that make up the population or sample.
- Comparing data year to year

Total Enrollment, 2022-2023



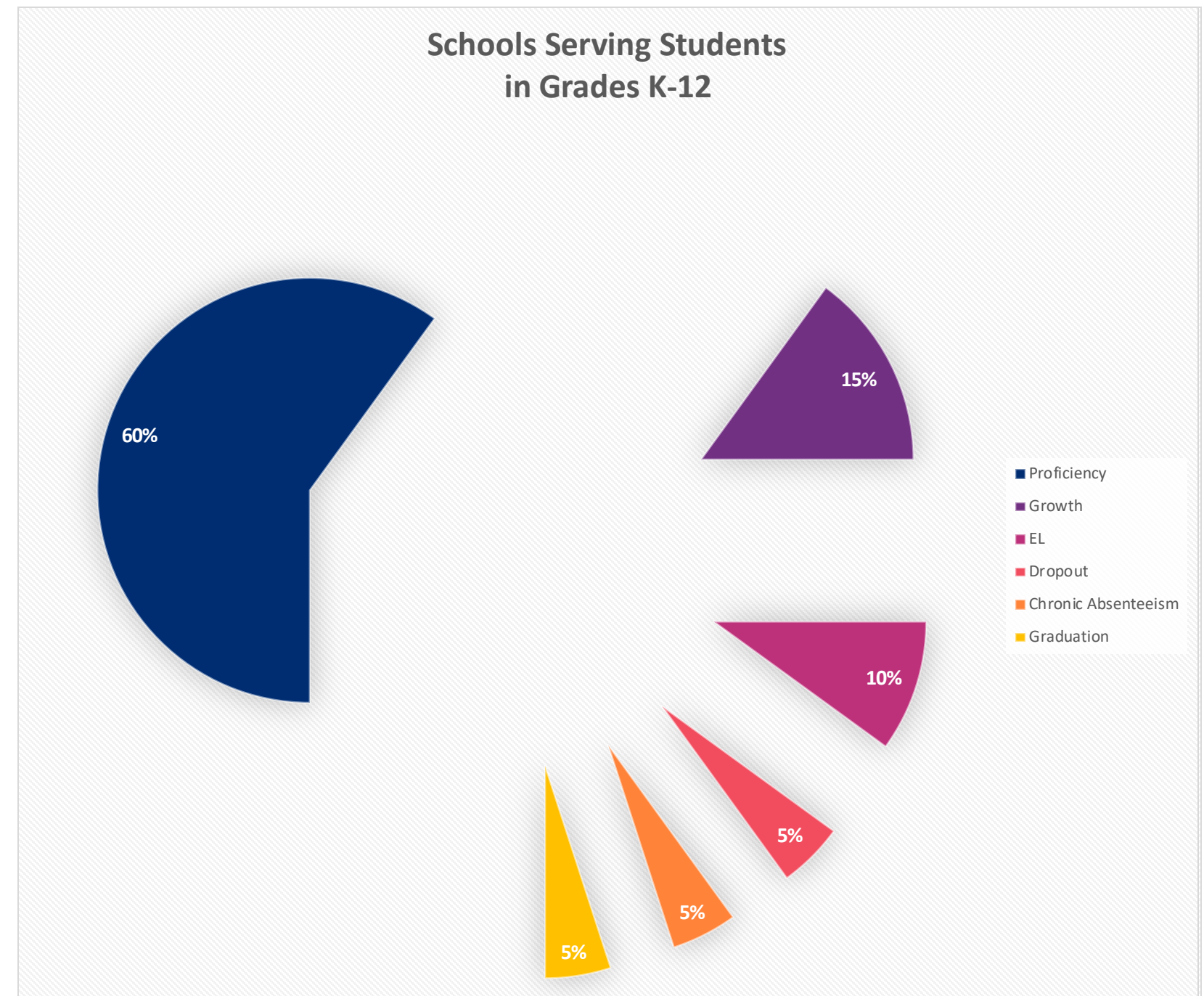
Pie Charts

- Categorical against Proportion
 - Only 2-3 Categories
 - Gender
 - Pass/Fail
 - Approve/Disapprove/No Opinion
- Area is hard to judge
 - Needs to be accurate
 - Difference must be clear, if any
 - Big divisions are vertical
 - Keep it flat

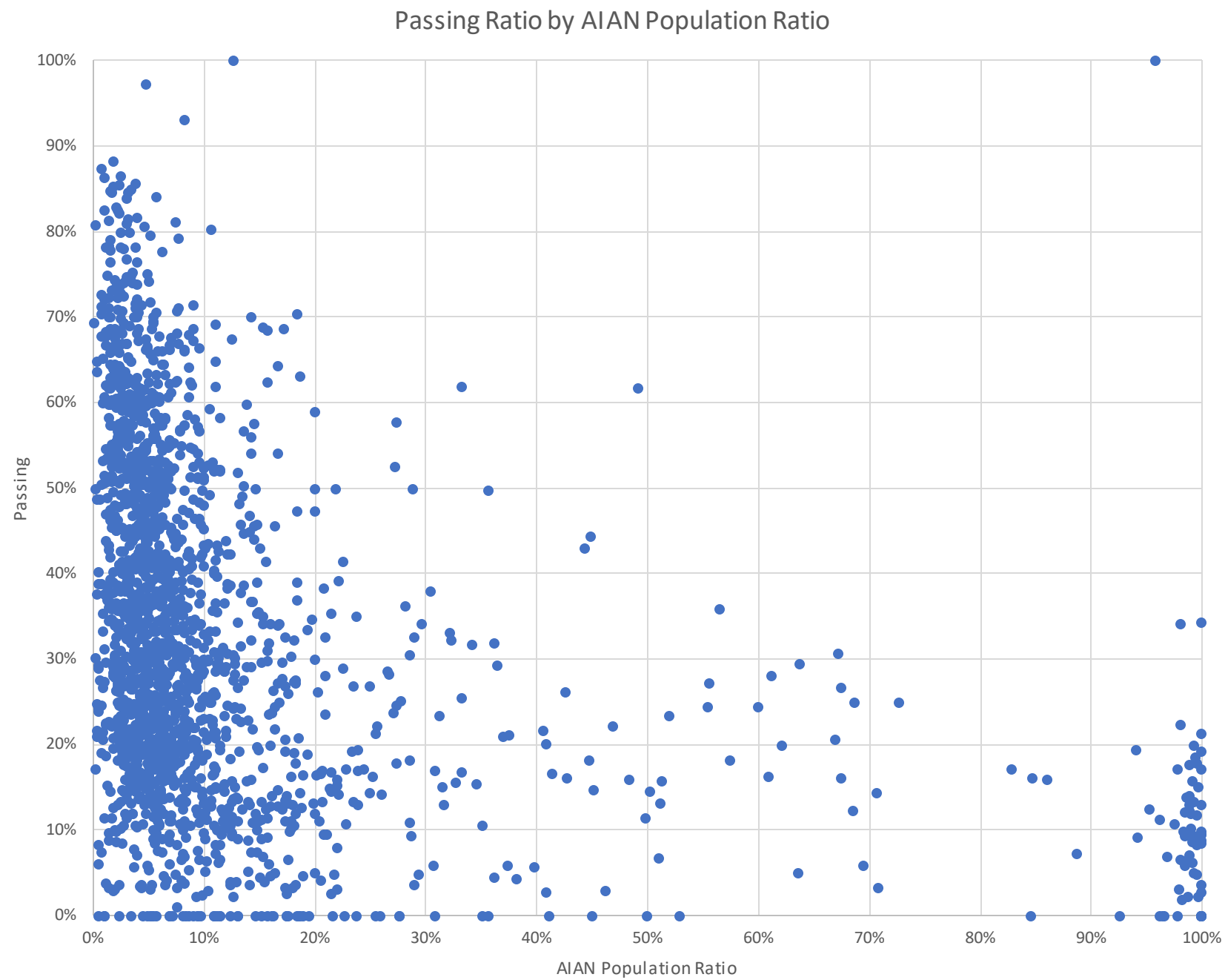


Exploding Pie Charts

- Exploding a pie chart with multiple categories *may* improve clarity.
- Beware:
 - Too many categories
 - Especially if differences aren't clear or useful
 - Too much explosion
 - A little goes a long way



Scatterplots



- Comparing two potentially-related numeric sets
 - FRPL and Proportion of Passing
 - Height and weight
 - Math score and English Score

Questions?

Raise Your Hand!



Drop in the Chat!

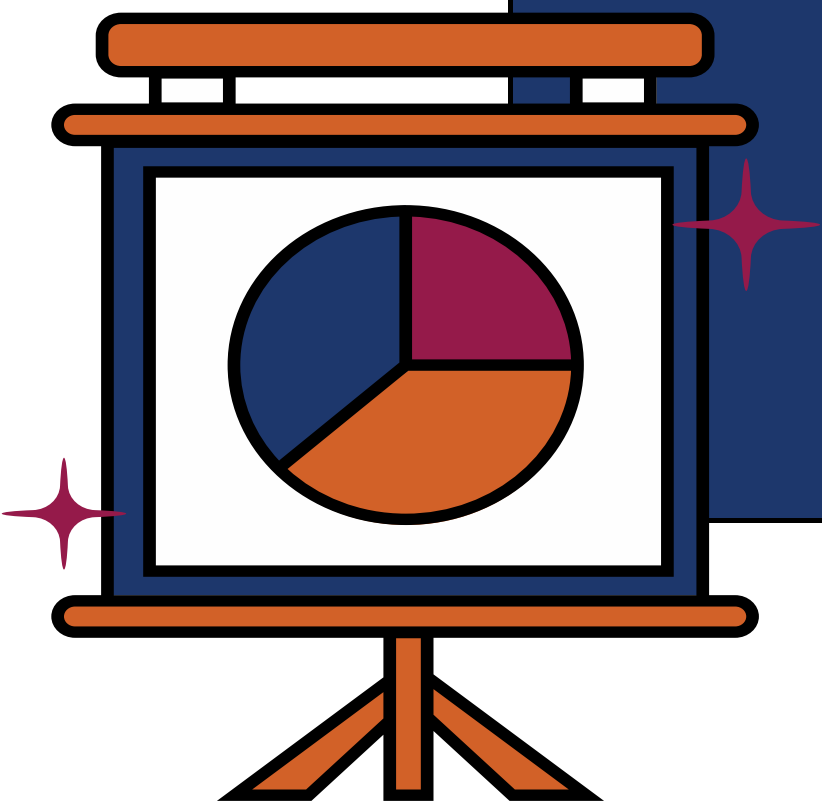


New stuff on the CNA

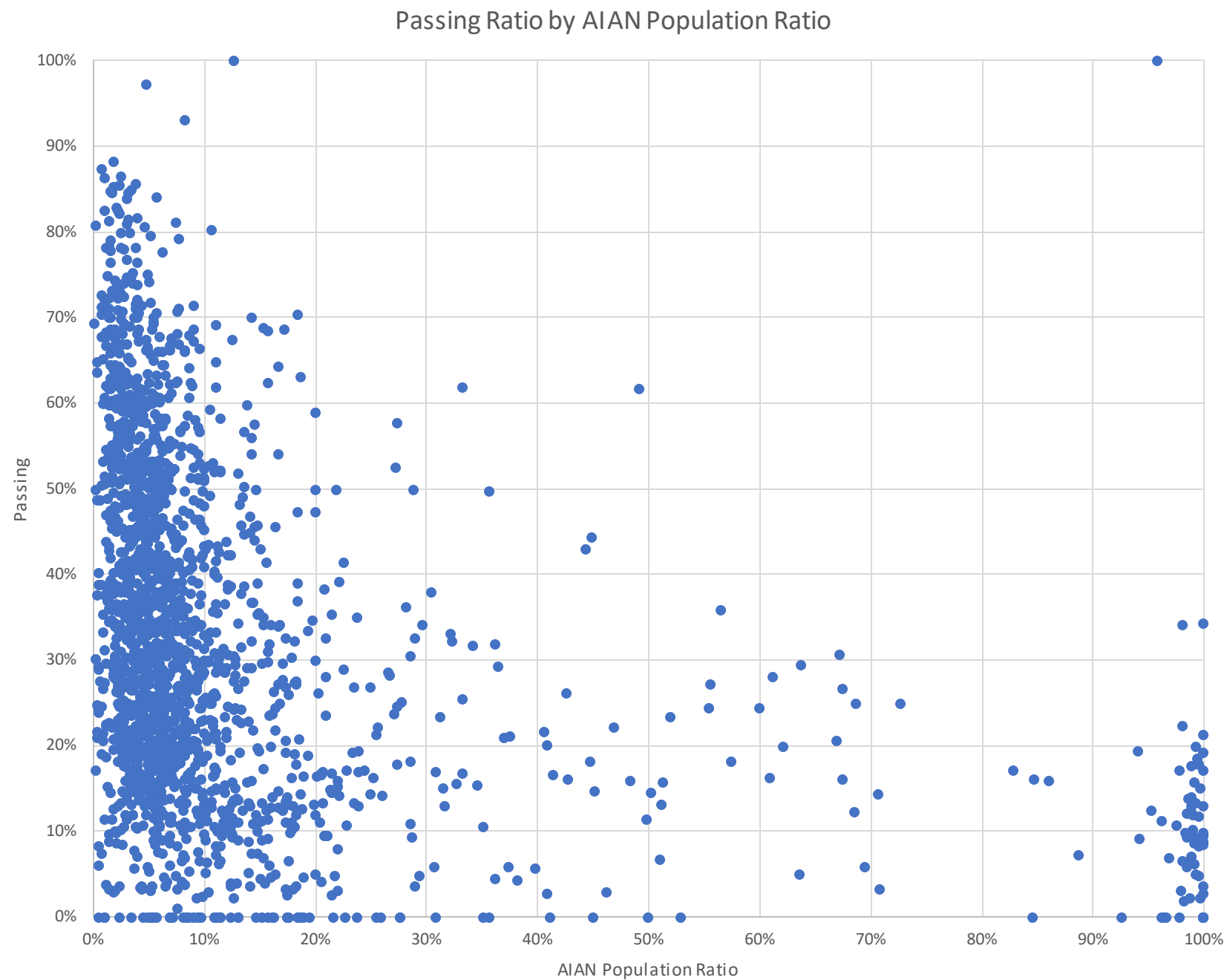
- Take 5-10 Minutes
 - Visualizations produced for the Comprehensive Needs Assessment (CNA)
- Think about these questions while you're looking at the CNA:
 - How useful is it?
 - How readable is it?
 - *Is it good?*



Principles for Effective Visualization



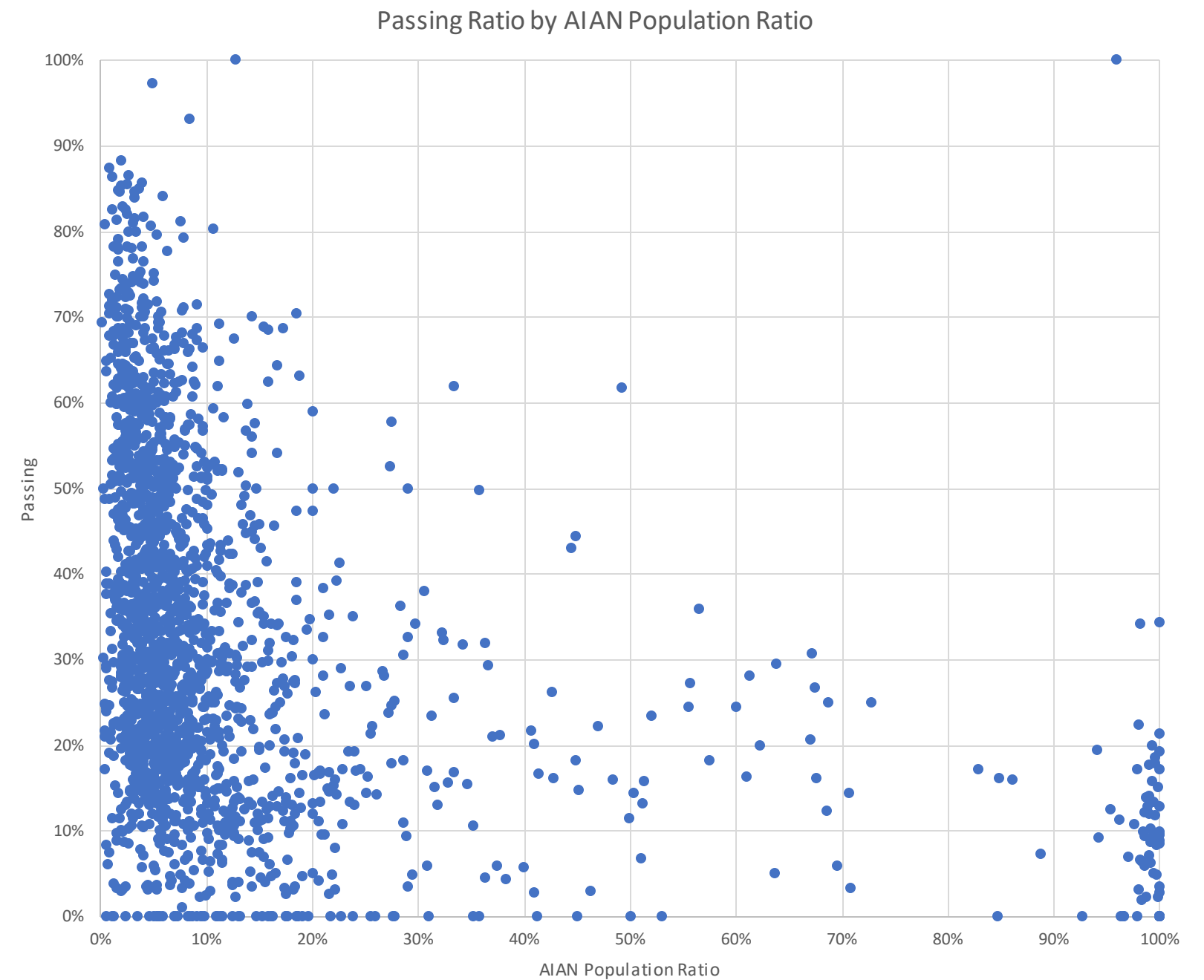
Positioning Data



- Where you place the data influences the viewer
- The shape of the data *and* the intended story should control placement.
 - Find the outliers
 - Find the trends
 - Which direction is “good”?

Repositioning Data

- How does changing the placement...
 - Help or hide the outliers?
 - Help or hide the trend?
 - Change the story?



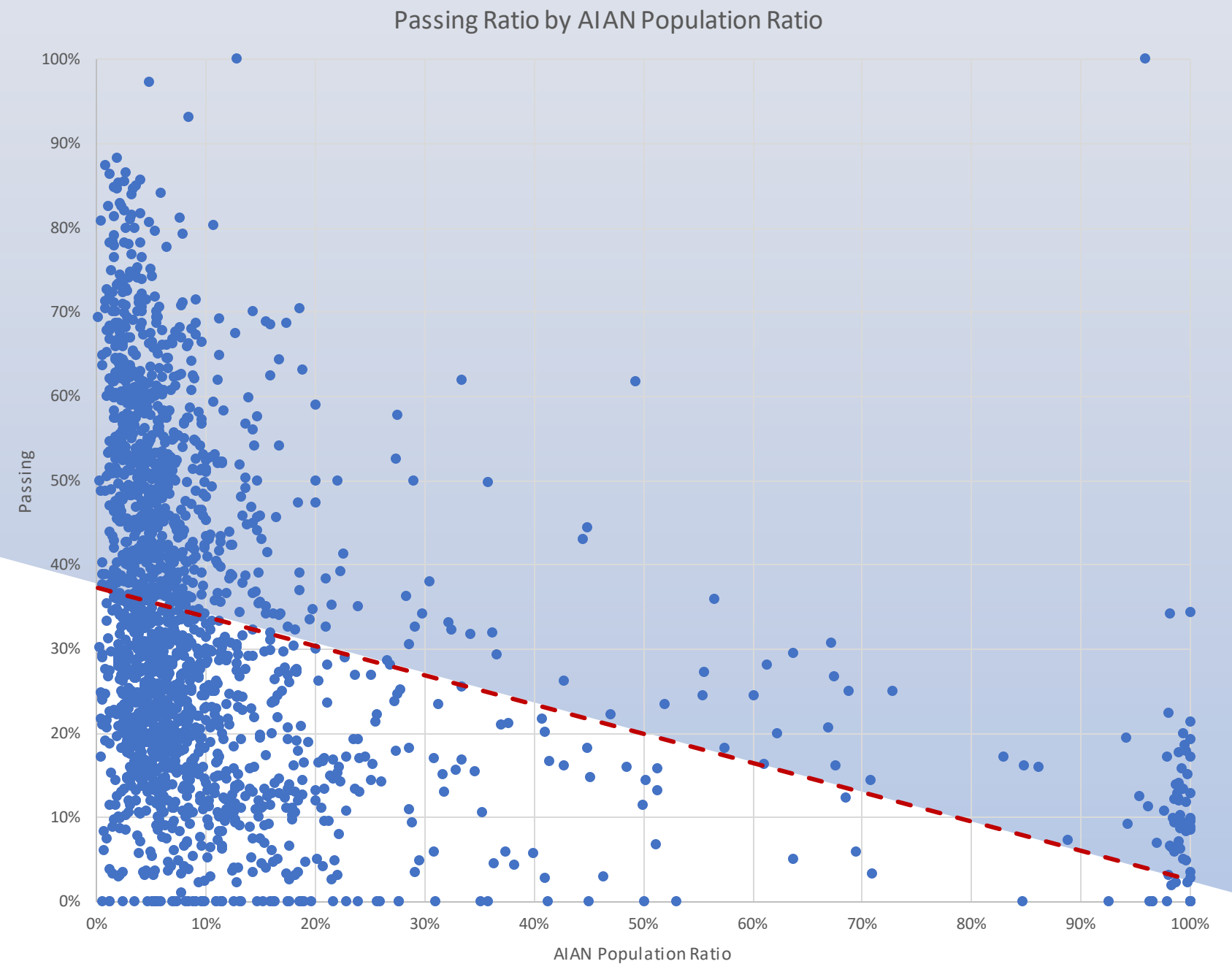
Positioning “Up”



- Trend lines influence the easy-to-read side of the chart
 - Negative trendlines should usually be on the left side of the page
 - Keep the ‘above the line’ area larger for betting impact

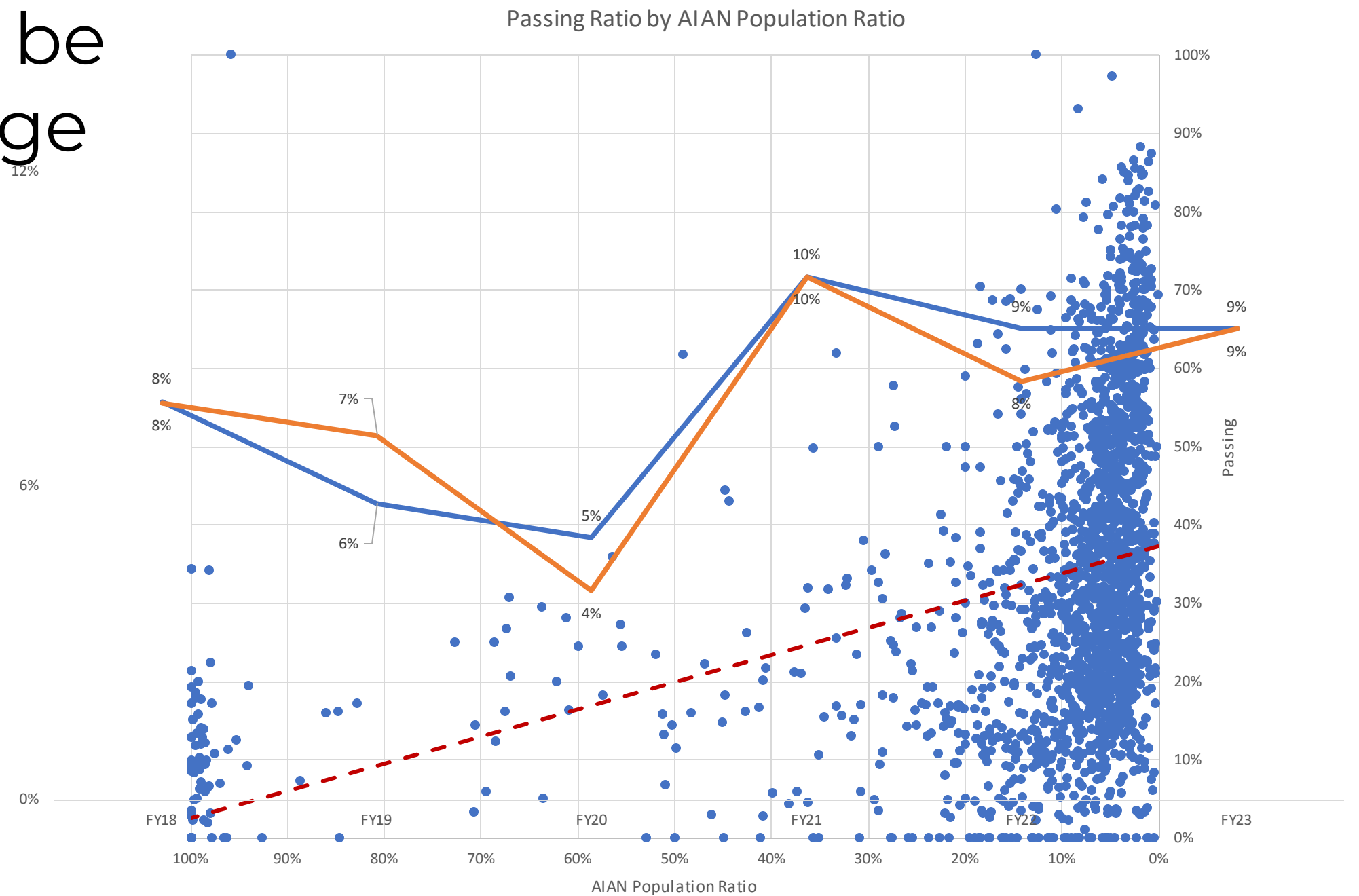
Positioning “Up”

- What the trendline bounds influences how we see the data.
- There’s a lot less “above” the line on this option.
 - Maybe “above” isn’t what you want
 - Sometime the story is under the line



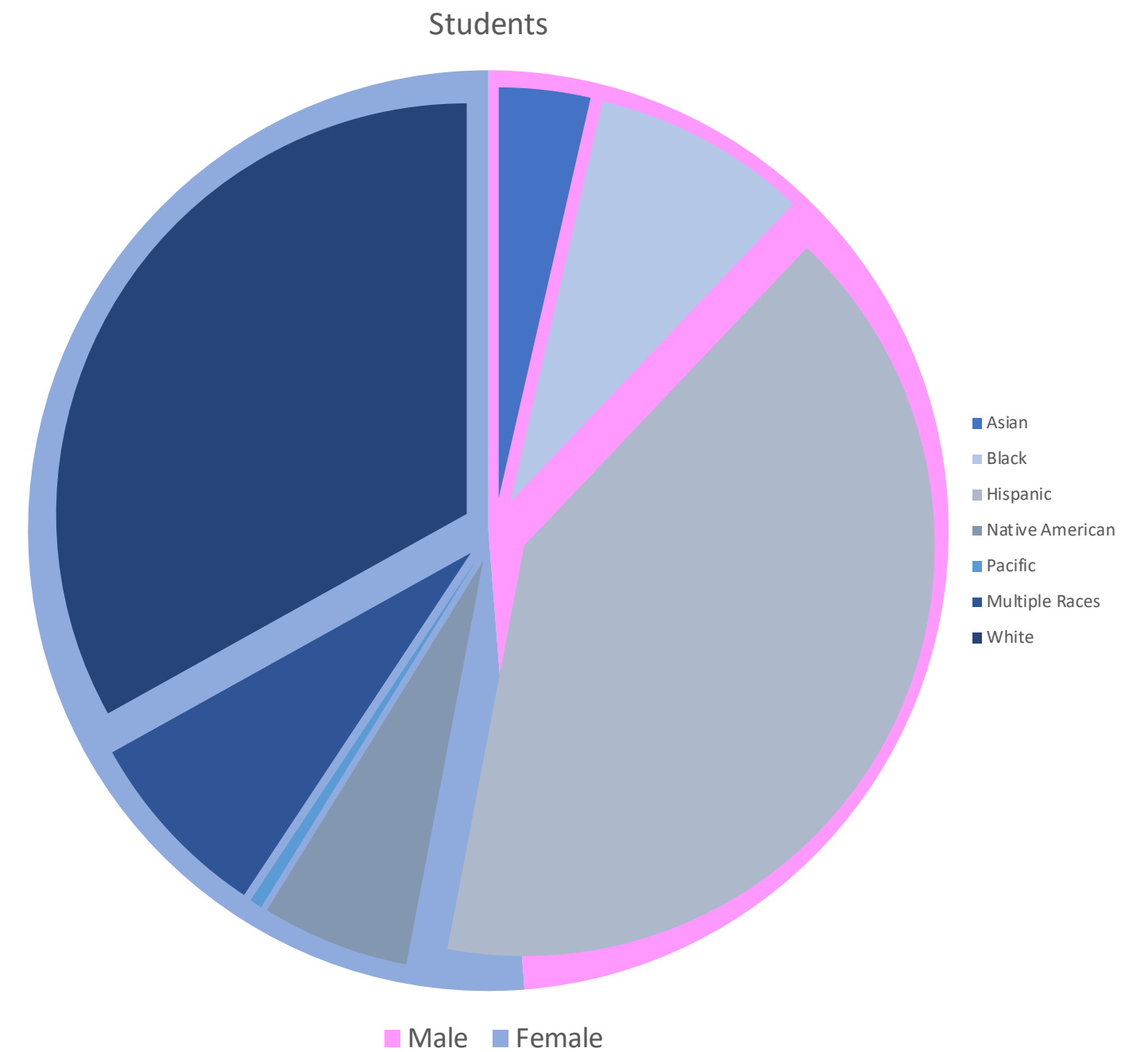
Positive Positioning

- Positive trendlines should be on the right side of the page
- Put the story in the 'good' section of the page



Principles of Color

- Make sense
- Distinct
 - [Data Viz Color Palette Generator](#)
- Consistent



Exceptions

- The story should be bounded by the visualization
 - If left to right, then 'positive' is also left to right
 - Read top down, then 'positive' is higher on the page.
- *There's always a good reason to do it another way.*



Simplicity: Data-Ink Ratio

Principles of good visualization

- Above all else show data.
 - Maximize the data-ink ratio.
 - Erase non-data-ink.
 - Erase redundant data-ink.
 - Revise and edit
- Data Ink: the non-erasable core of the visualization
 - Data Ink Ratio: the data ink divided by the total ink used to print the graphic.
 - *Get rid of EVERYTHING that doesn't add to the clarity.*

That's Not Data

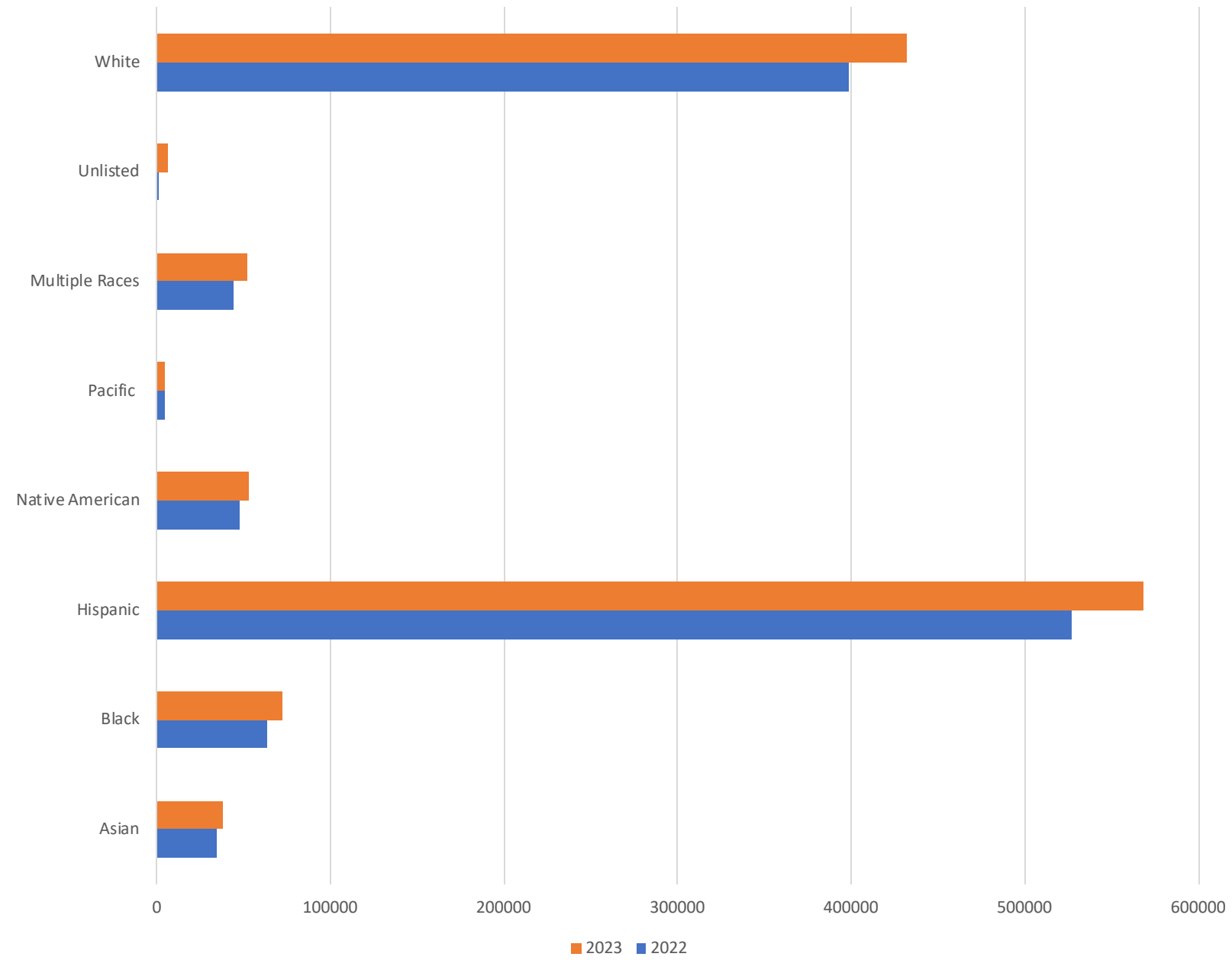
- Focus on the content
 - Lighten labels
 - Remove borders
 - Lighten or remove lines
 - Remove anything redundant
- That's not data:
 - Grid lines
 - Axis values/ticks
 - Color schemes

“Never underestimate your audience. It's the most common mistake made by presenters. It is not about you anymore. It's about your audience's relationship with your content.”

- -Edward Tufte

Chart Junk: Example 1

Enrollment Changes



Enrollment Changes, 2022 to 2023

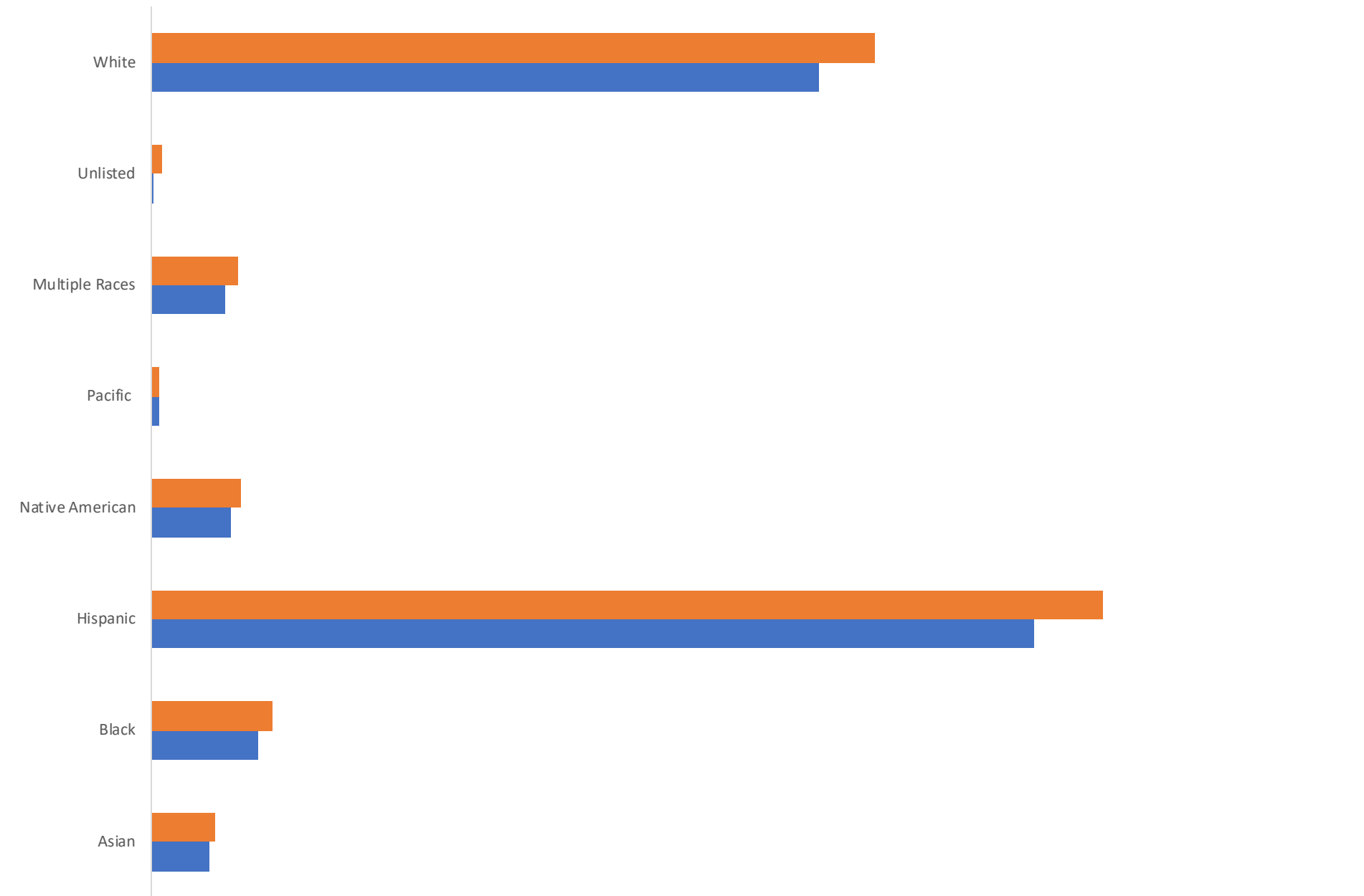
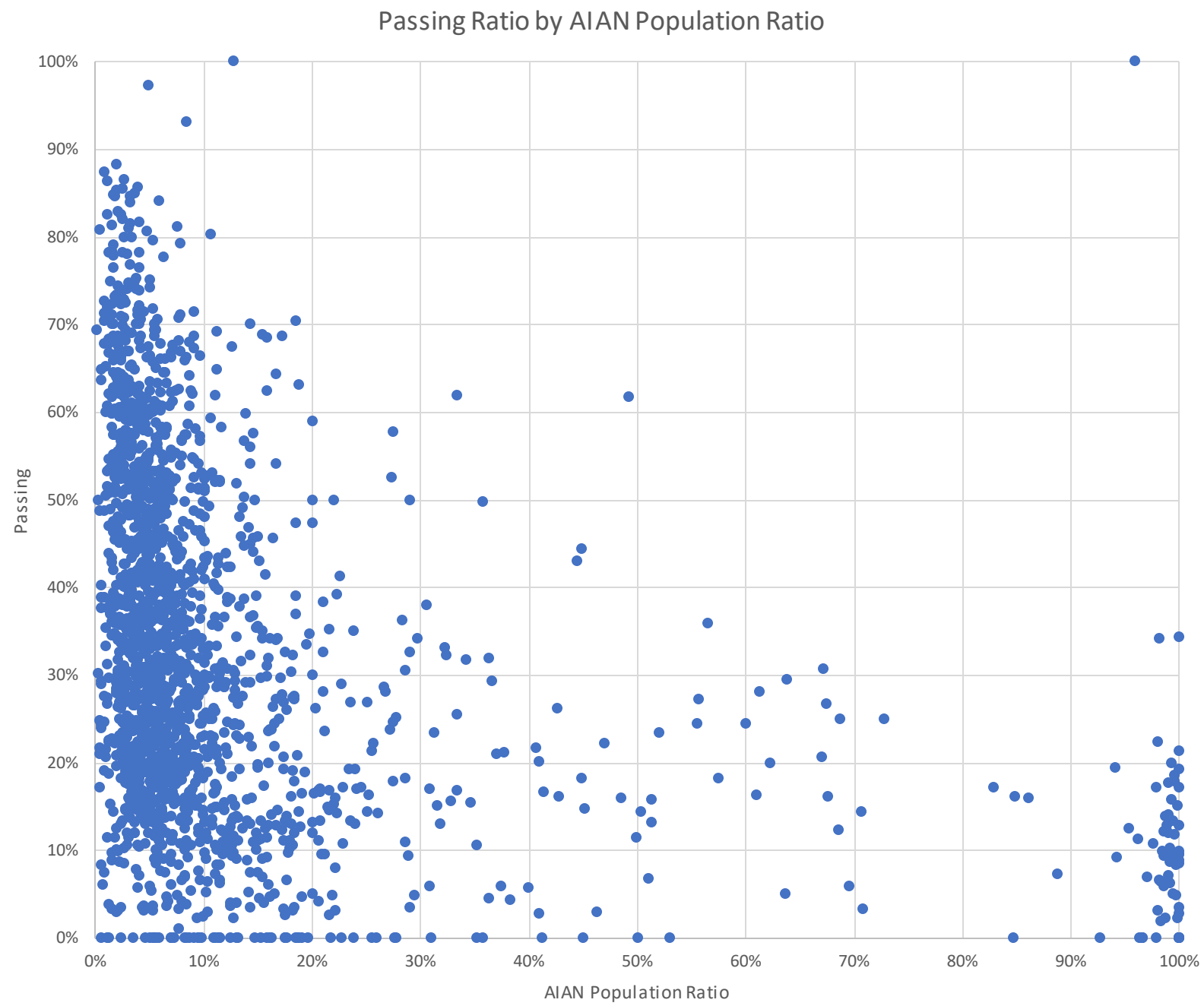


Chart Junk: Example 2



School Pass Ratio by AIAN Population

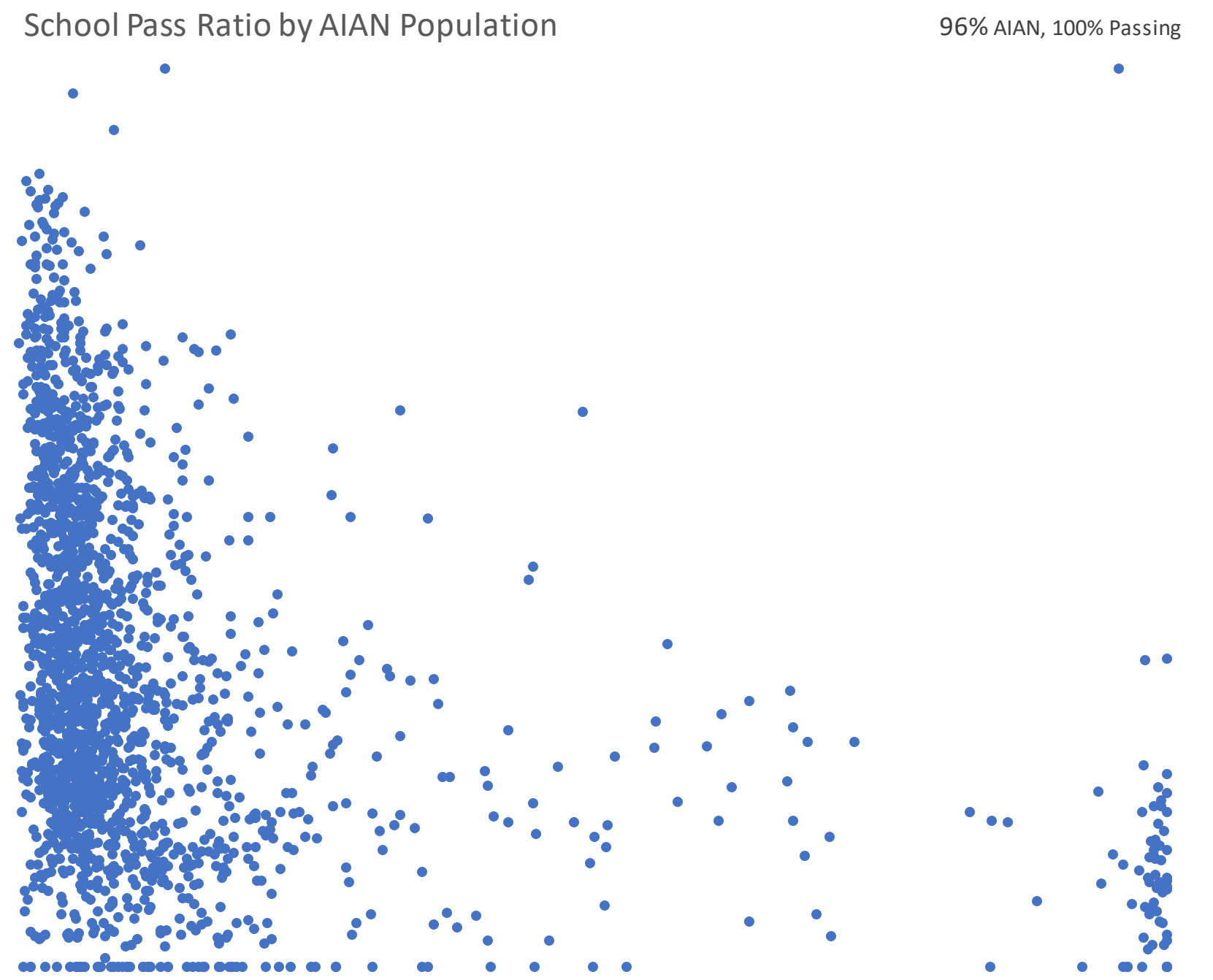
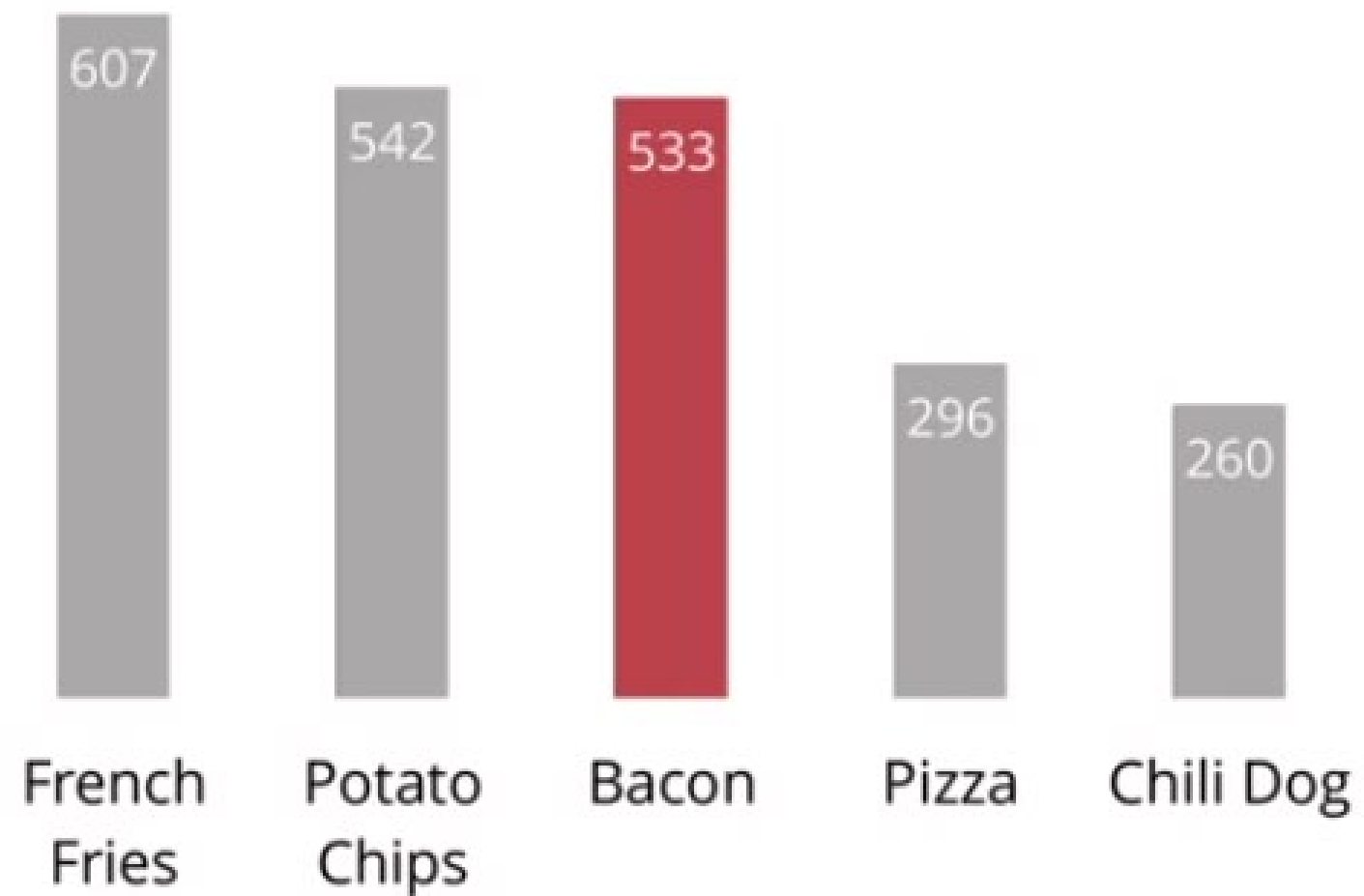


Chart Junk: Applied

Calories per 100g



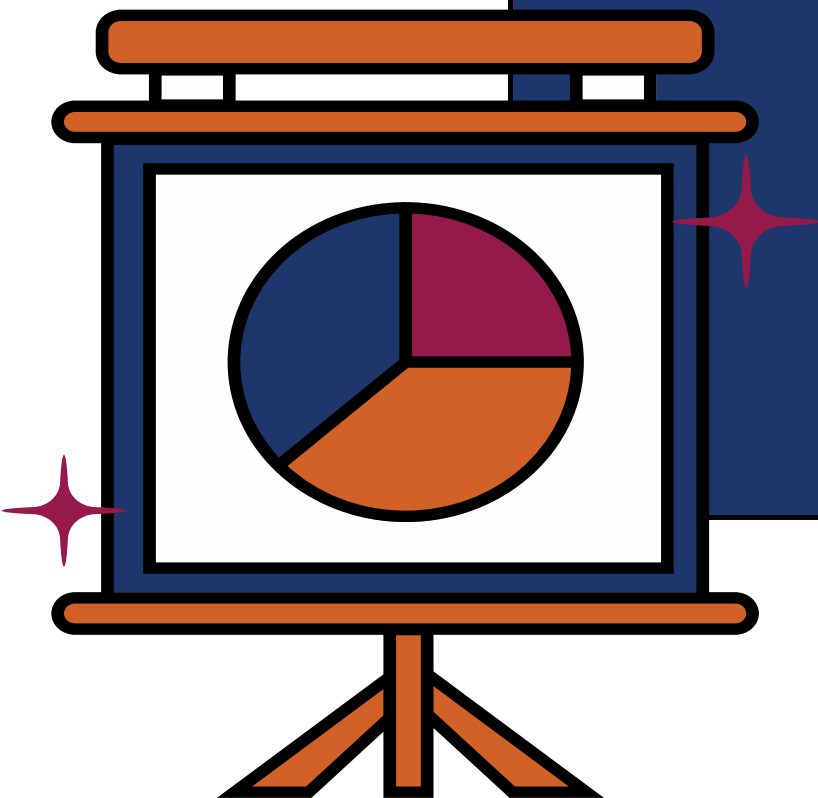
- Background color
- Grey background
- Category duplication
- Type of food
- ...for different foods
- Borders
- Shading/3D effects
- Colors aren't sensible (focus!)
- Unembedded data/Y-axis

Iterate and Commit

- Good visualization is a *process*
 - Many options for any datum
 - Try, edit, repeat
- Use your resources
 - Graphic Designers
 - Editors
- Don't let perfect be the enemy of good
 - You can spend *hours* on anything
 - Be conscious of your investment
 - If it works, move on



Recap and Resources

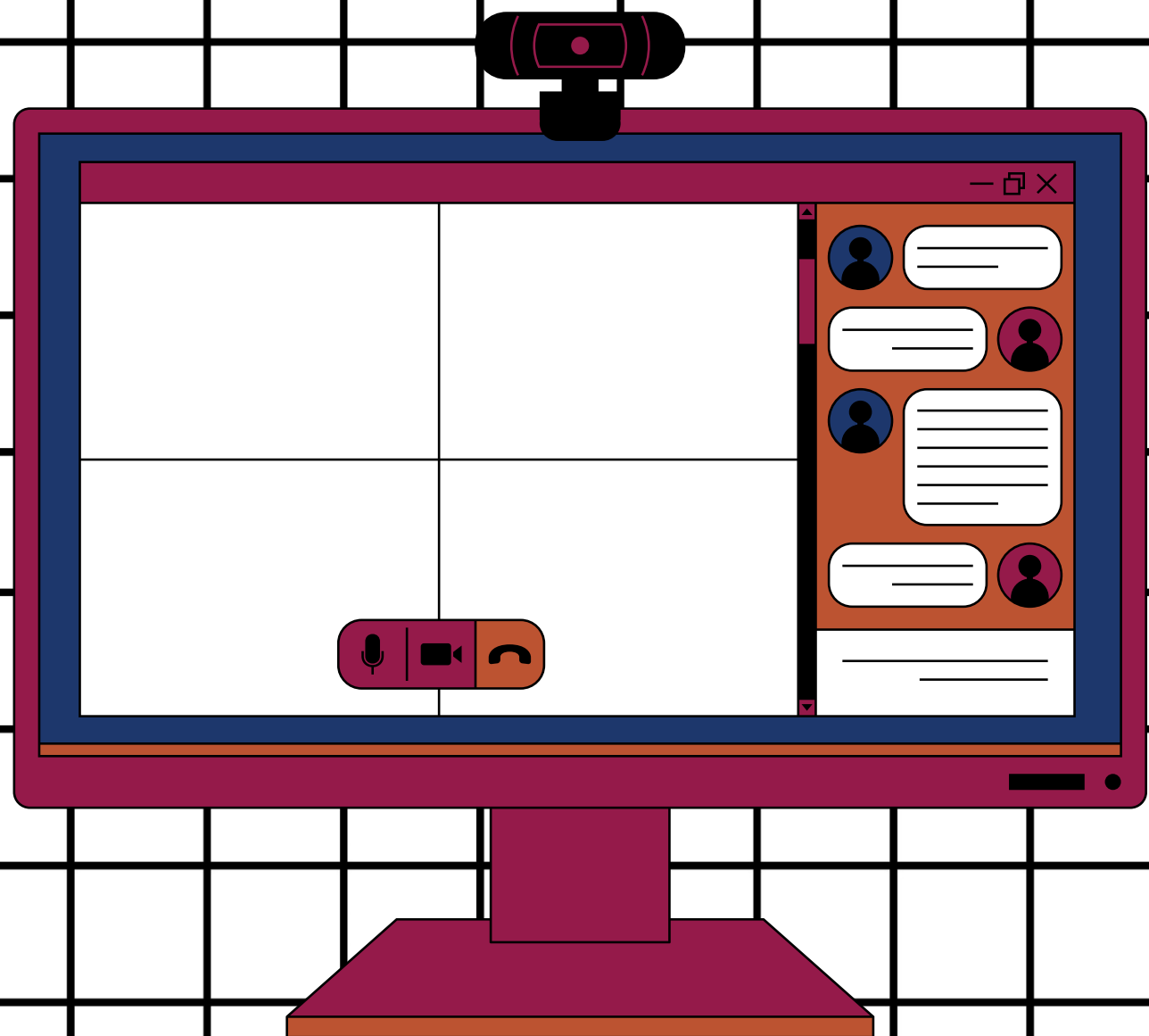


Review

- Identify the big question
- Seek the best visualization
 - What tells the story best?
 - Compare 2 values
 - Difference of 2 values
 - Proportional change
 - Types of visualizations
 - Bars and histograms
 - Pie Charts
 - Scatterplots
- Good visualization principles
 - Position and Color
 - Simplicity & Flatness
 - *Be reasonable with your resources: It's a process*
- **Maximize data-ink ratio**

The Visual Display of Quantitative Information, Edward R. Tufte

Breakout Discussions



- What are you taking away from this training?
- How will you apply what you've learned in your grant work?
- What else do you need to know?

More Resources



Don't forget, all data training recordings are on our website!



Scan the QR code or visit www.azed.gov/oie/professional-learning-resources

Thank you!



Feedback Survey

