

# Visualization

INT 93

---

Michael Topper

# Good Visualizations:

## Instant Gratification

- Interpretable with little effort from reader
- Self contained - nothing open to interpretation
- Interesting
- Simple
  - You will likely be trying to convey 1 idea
  - Complexity comes when showing same figure for multiple subgroups
  - Careful with complexity...could leave to misinterpretation from reader
- It will take many drafts to get the right figure!

# Planning a Figure:

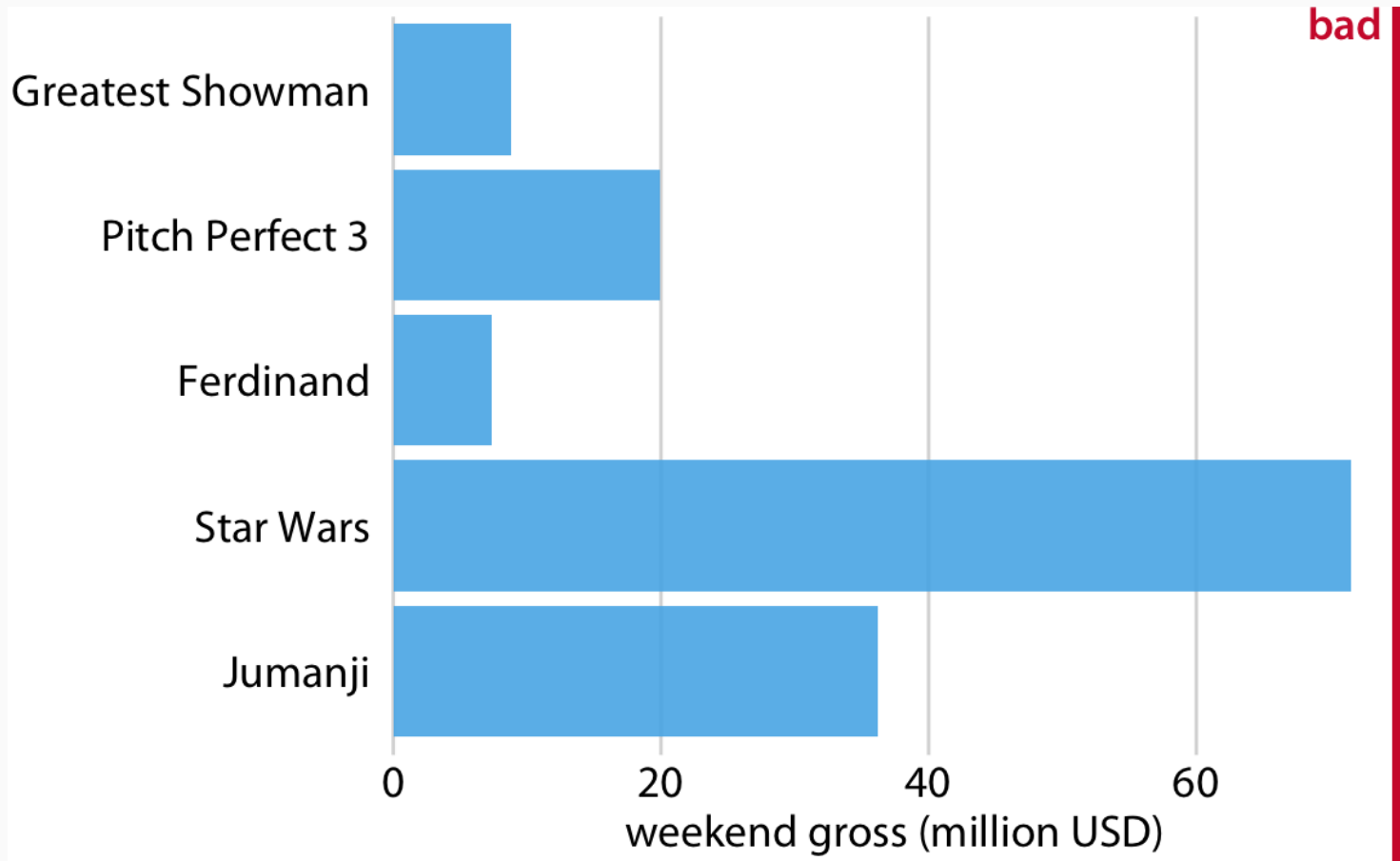
## Start with the result

- What result do you want to convey?
  - Think about what kind of figures best convey such result (histogram/barplot/time trend etc)
    - Would this figure be better as a table?
  - Draft out multiple ideas for the figure
  - Can you include interesting subsets in the figure with a slight modification?
- Use your group members!
  - Come up with a sketch of a graph and show to group members

# Exercise: Planning Figures Together

---

# Example: Bar Graph (Bad)



# Example: Bar Graph (Bad)

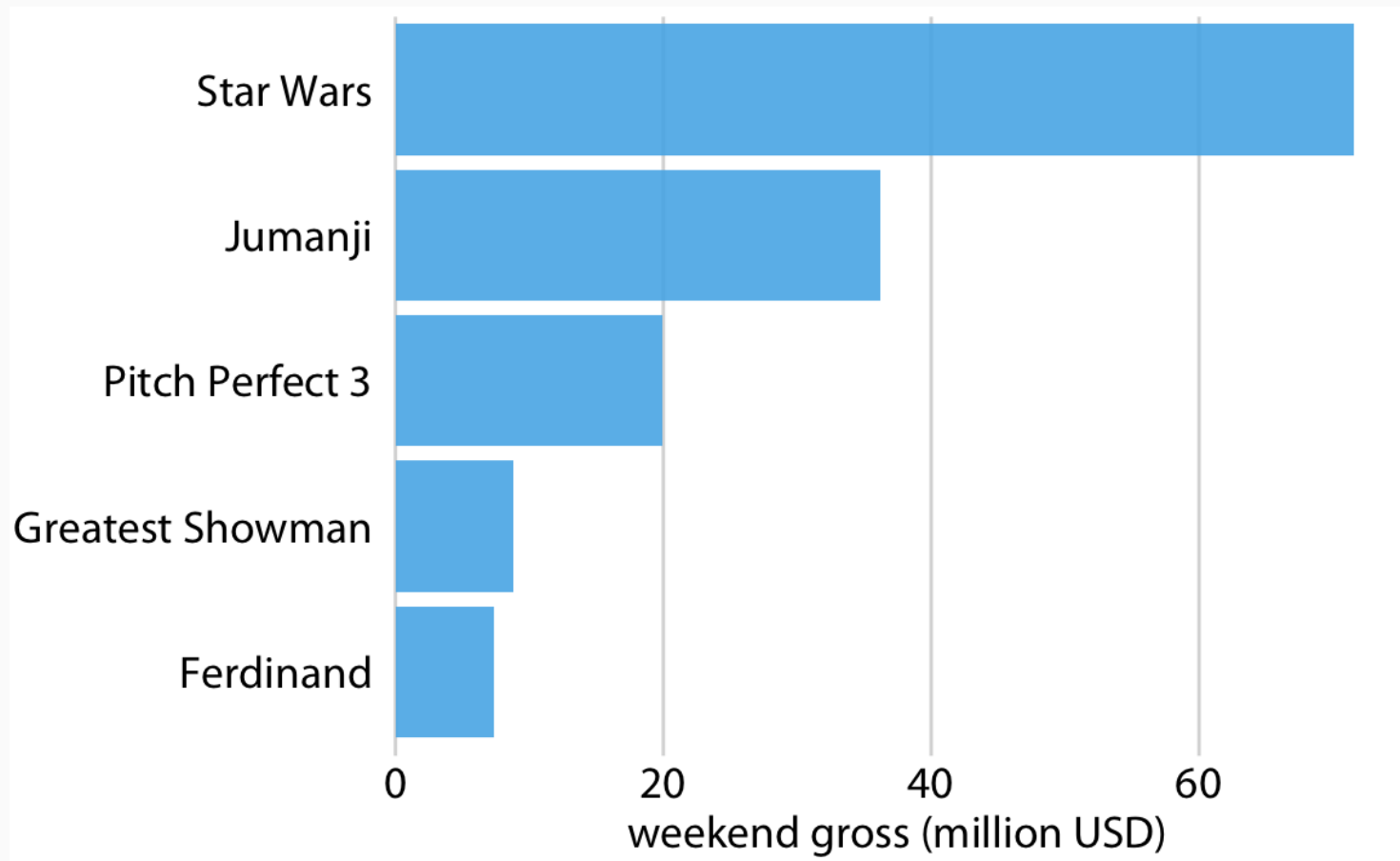
## Problems:

- Poor choice of ordering
- Hard to compare groups

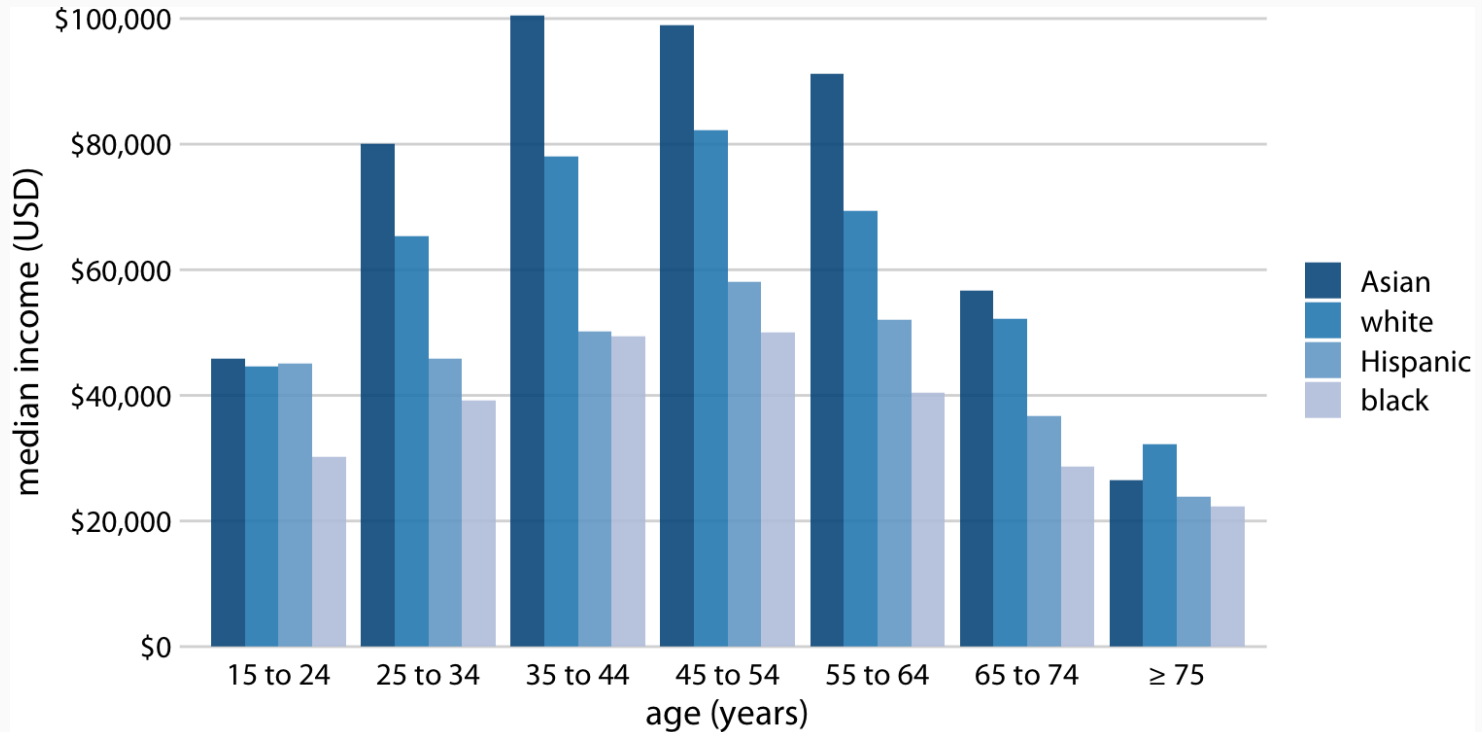
## Fix:

- Reorder from highest to lowest
- Easier comparison

# Example: Bar Graph (Fixed)



# Example: Bar Graph with Groups





# Example: Bar Graph with Groups

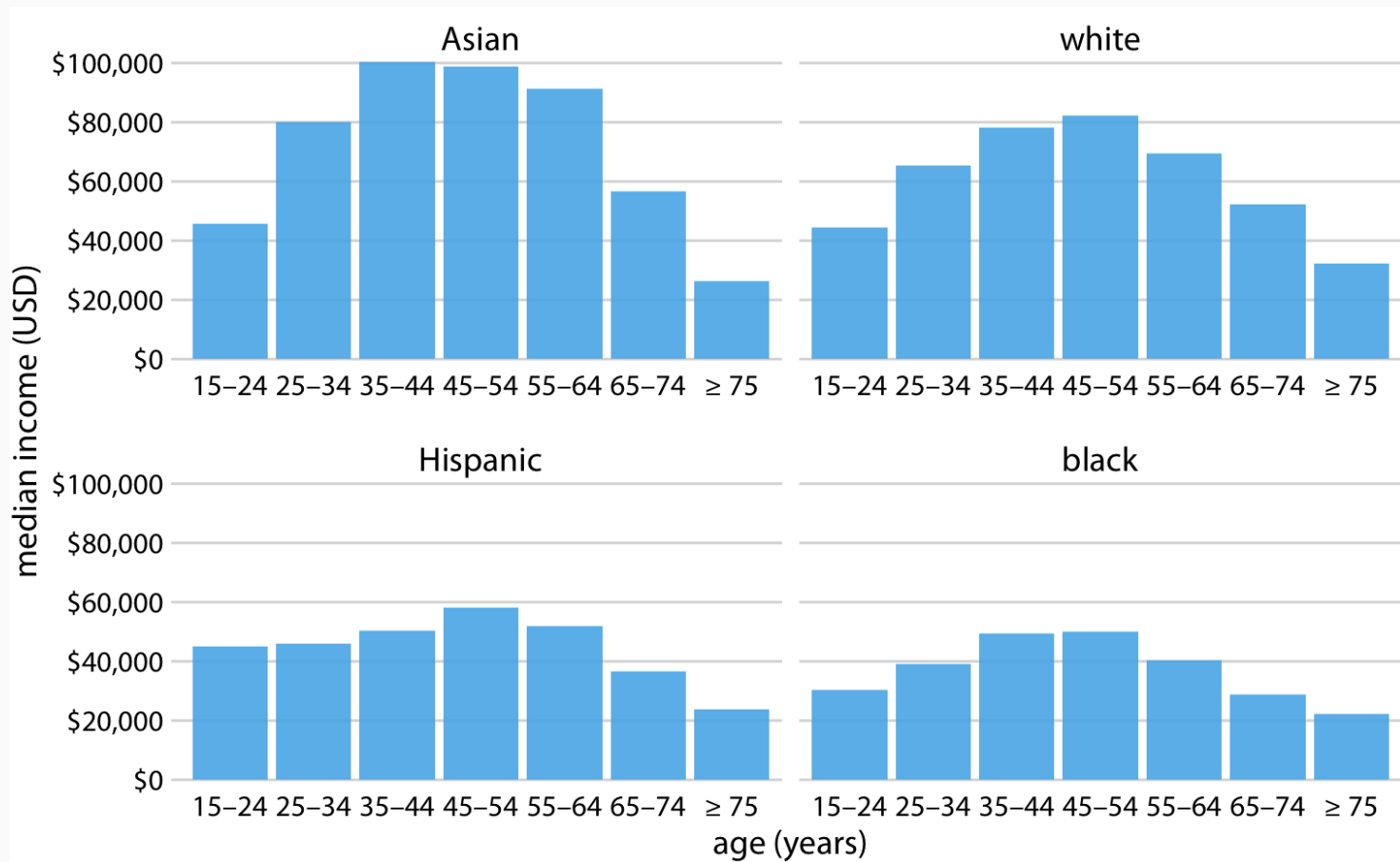
## Problems:

- Difficult to compare across races
- Wants to convey comparisons, but hard on the eye
- Inconsistent capitalization

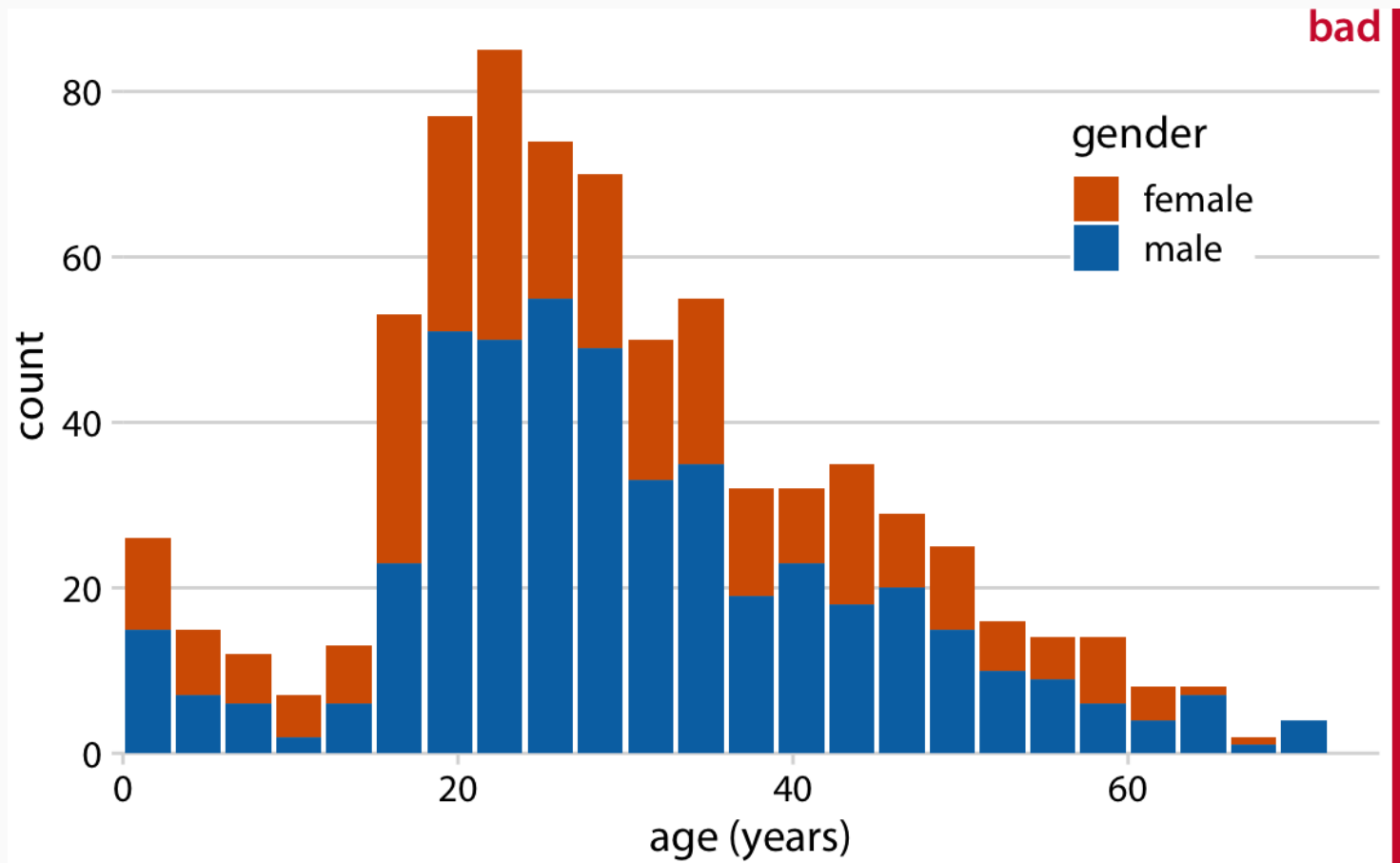
## Fix:

- "Facet" by race
  - A facet is an individual graph by a subset
  - 4 small figures by race
- This will help convey the message to the reader quicker

# Example: Bar Graph with Groups



# Example: Histogram (Bad)



# Example: Histogram (Bad)

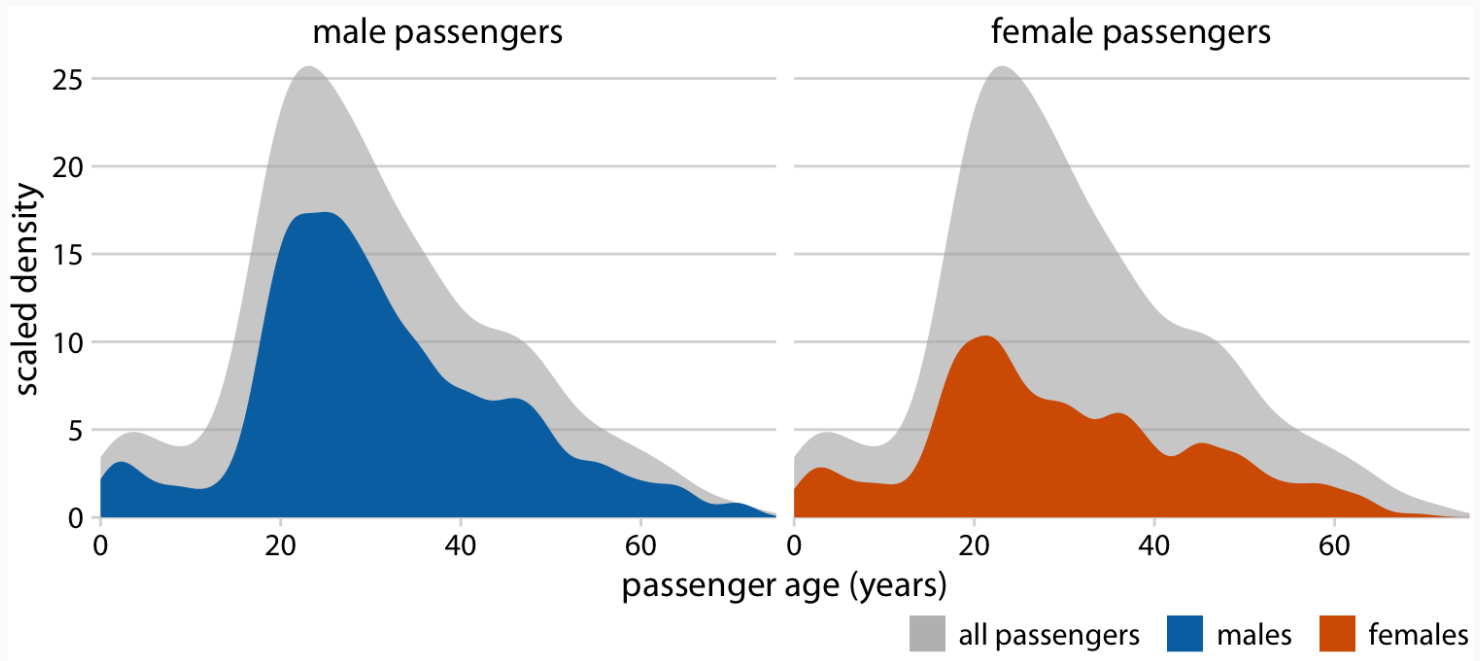
## Problems:

- Stacked histograms are easily confused with overlapping histograms!
- Is this a stacked histogram or overlapping
  - Can reach very different conclusions!

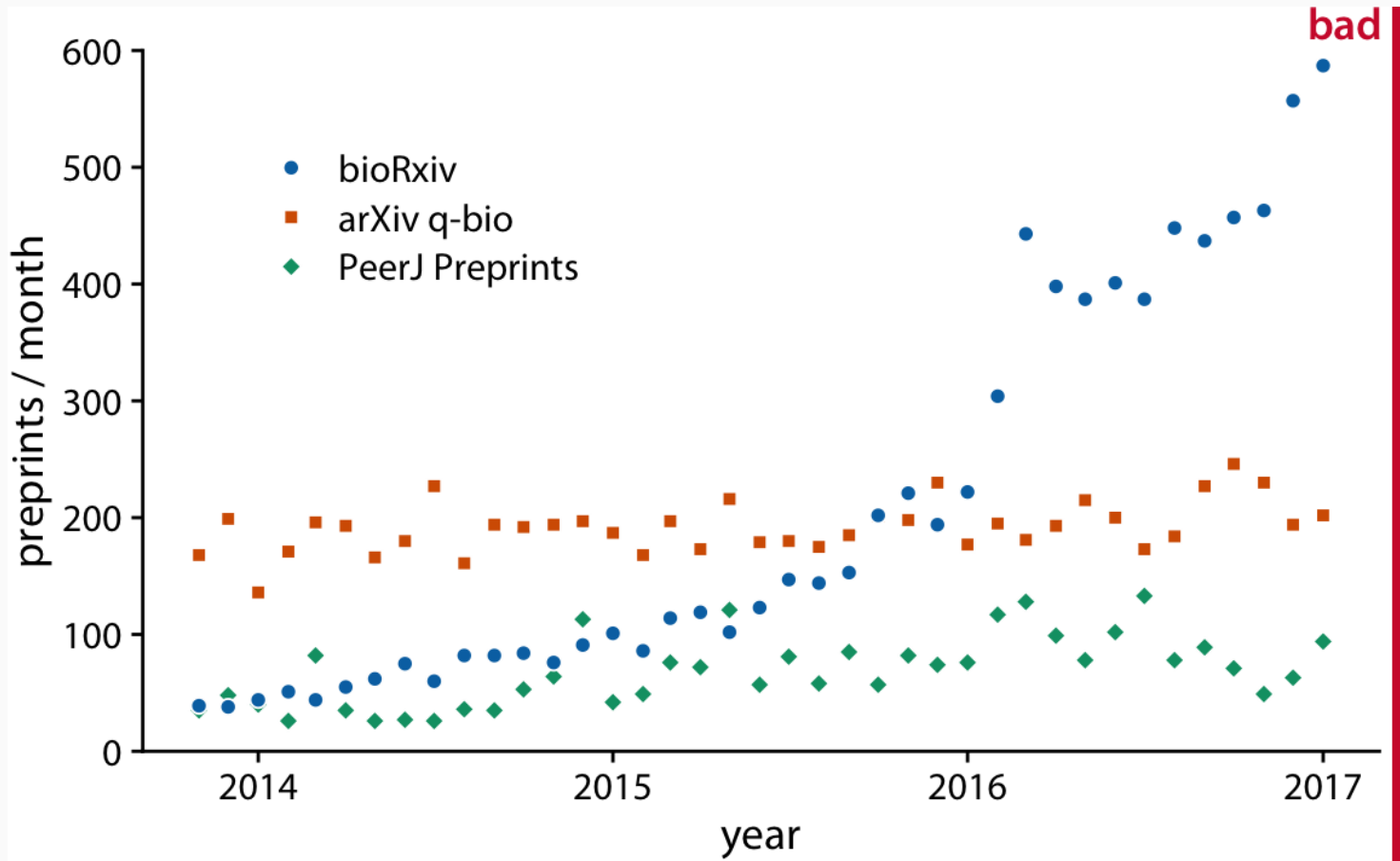
## Fix:

- Multiple solutions to fix here, but usually subgroups are best displayed with a facet.

# Example: Histogram (Fixed)



# Example: Time Trend (Bad)



# Example: Time Trend (Bad)

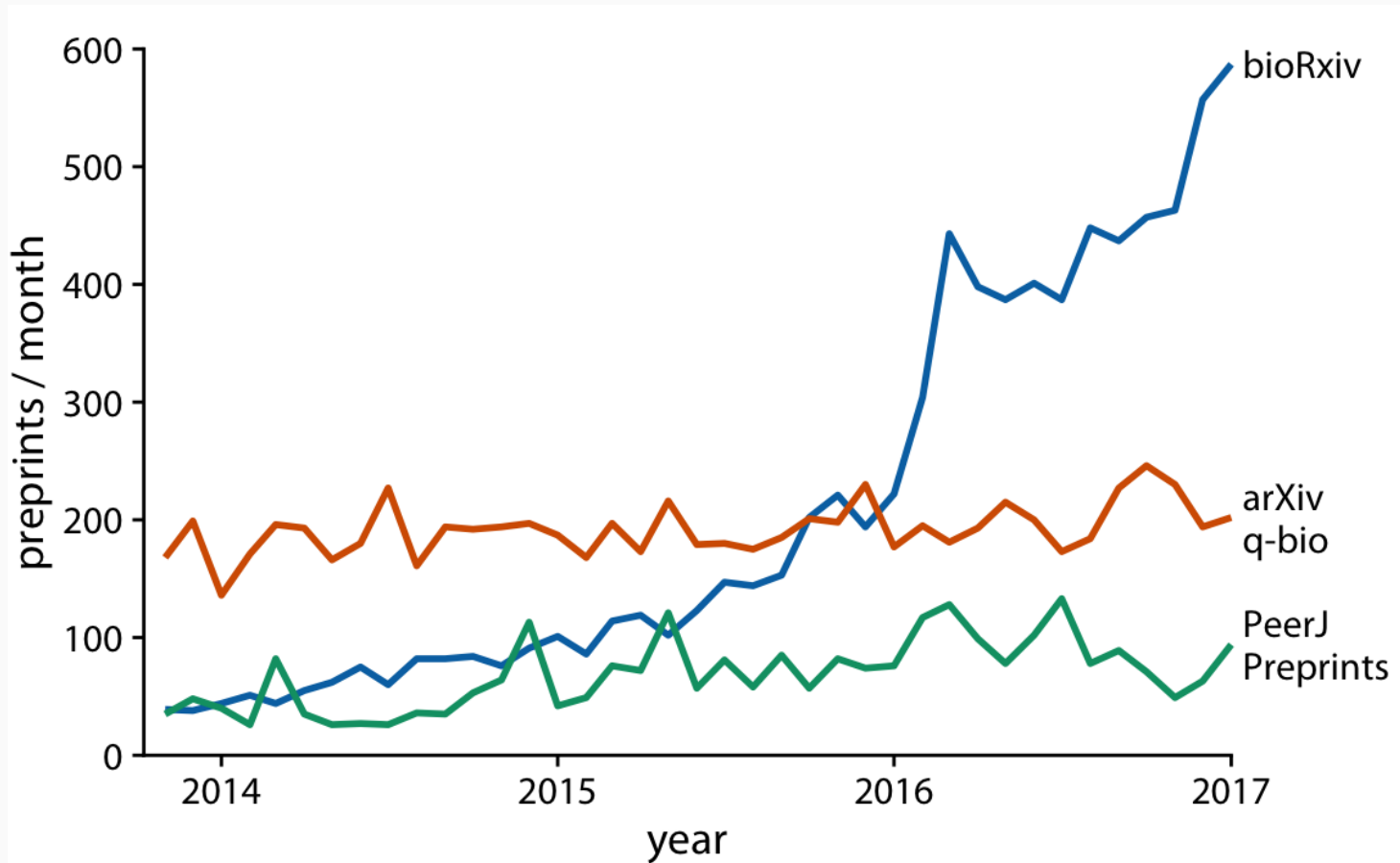
## Problems:

- Difficult for reader to see the trend
- Three different colors overlap with each other
- Reader needs to jump back and forth with the legend

## Fix:

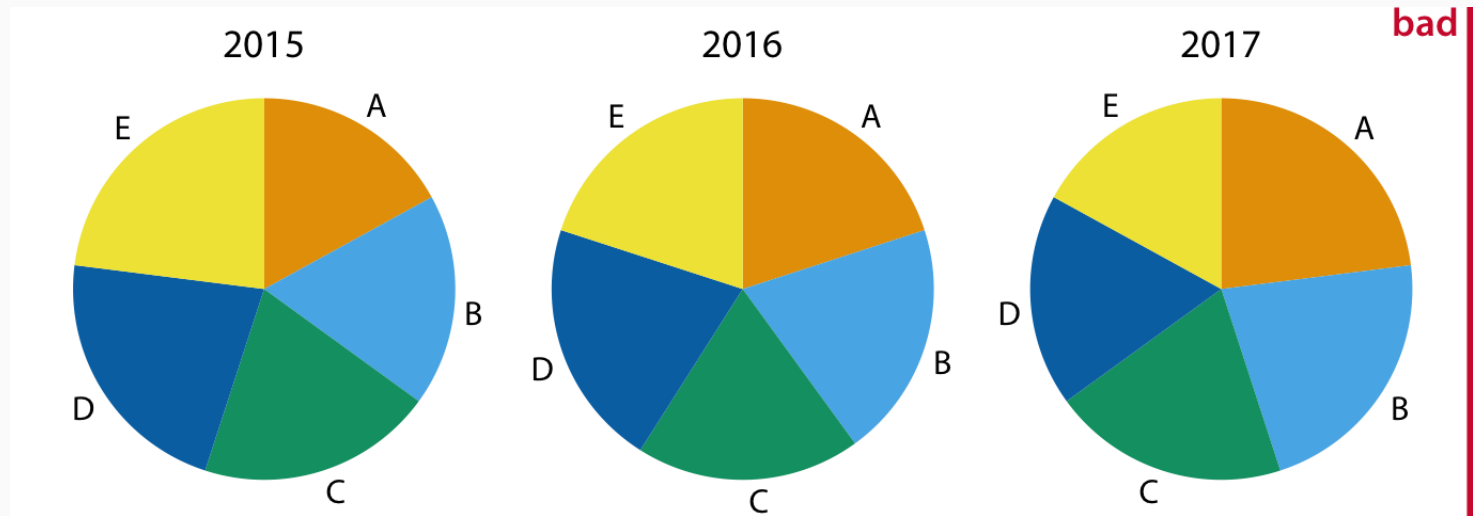
- Connect the dots
- Alternatively, you could facet by group...although this might be a little overkill
- Put the legend on the map

# Example: Time Trend (Fixed)





# Example: Proportions (Bad)

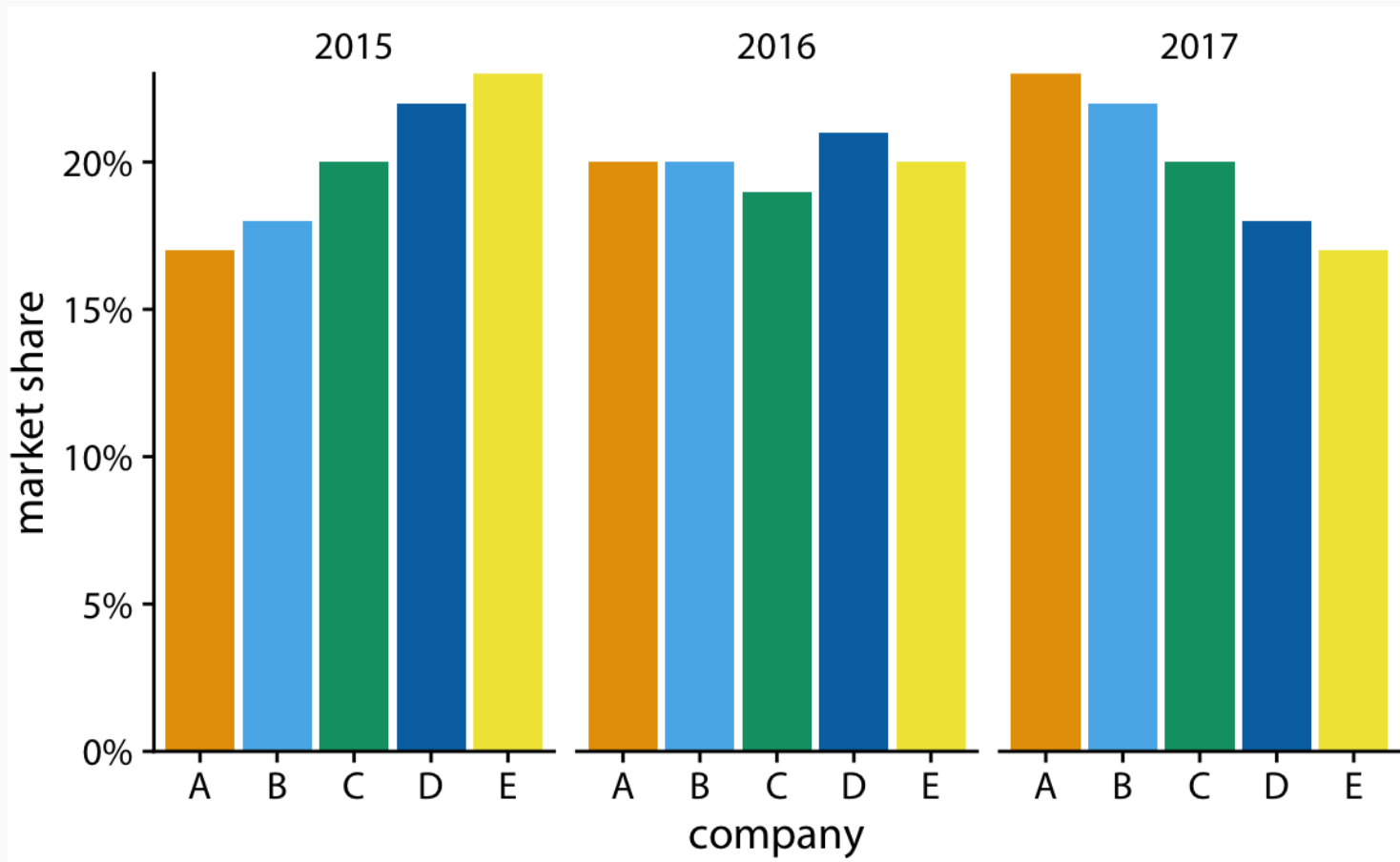


# Example: Proportions (Bad)

## Problems:

- Most people say pie charts are the worst form of visualization
  - In my opinion only good when looking at one group
- Poor method when comparing across groups
- Faceting will not work here - the pie chart is bad at these types of comparisons
- Try making a different figure:
  - Barplot side-by-side faceted by years

# Example: Proportions (Fixed)



# In Your Paper:

## Main Figures of the Paper

- Your paper should essentially be a few results summarized in a figure or two
- Most people will only look at this one figure
- Be sure this figure is clear to everyone—ask others!
- Make this figure *pop*
- Your parents should be able to understand your research question/paper/result in this one figure

# Examples:

## Police Mistrust Paper (Ang et al. 2022)

- Page 16 shows a graph that is the entire paper

## Fraternity Paper (Topper 2023)

- Main figure shows significant decline only in moratorium picture

# Further Readings:

## Great Books on Visualization Practice:

- [Fundamentals of Data Visualization](#)
- [Data Visualization](#)

## Great Sources of Visualization

- [New York Times](#)
- [FiveThirtyEight](#)

# Tables

---

# Summary Statistics Tables

## Good Summary Statistics:

- Self-contained
- For each outcome variable/control variable:
  - Shows the Mean/SD
  - Can possibly show Min/Max/Median if wanted
- Breaks down by interesting sub-groups (if there are any)

A stranger should be able to read the summary statistics table and know the people being studied, outcome being studied, and controls used!



# Example: Summary Stats

**This is an example of a good summary statistics table**

Sloan & Hoekstra 2020

- Does Race Matter for Police Use-of-Force?
  - Page 45

Things this does right:

- Self-contained notes
- Interesting sub-samples
- Means/SDs
  - Outcomes
  - Controls

# Results Tables

---

# Results Tables

## Good Results Tables:

- Self-contained
- Shows a main result, and possibly interesting sub-group result
- Not overcrowded with unnecessary information
  - You do not need to show all control variables if they are not important
  - Note these in the table note
- Contains notes
  - Notes should have all the necessary information to make the table self-contained
- Should include point estimates/standard errors
  - Generally helpful to include the mean of the dependent variable

# Paper Template

---

# Template

## Template for Paper

- Shared on Canvas: Week 2 Thursday
- DO NOT EDIT THE ORIGINAL
  - Make a copy
  - Edit the copy
- Follow style outlined

## Small Differences for Econ:

- Tables/Figures should be in their own sections/pages, after all text/references
- You may need to change the names of headings (abstract/intro/background/empirical strategy etc.)

# Progress Reports:

## At this point in time

- Your research question should be solidified
- You should know what your data is
- You should have opened your data

## Next steps for today

- Meet with me group-by-group
- Get some ideas of main figures/graphs
- Work on Progress Report 2
  - 1 Graph/Figure
  - 1 Regression Table
- Refine any broad ideas in your research