CSE 599K

Empirical Research Methods

Winter 2025

Data visualization and reporting

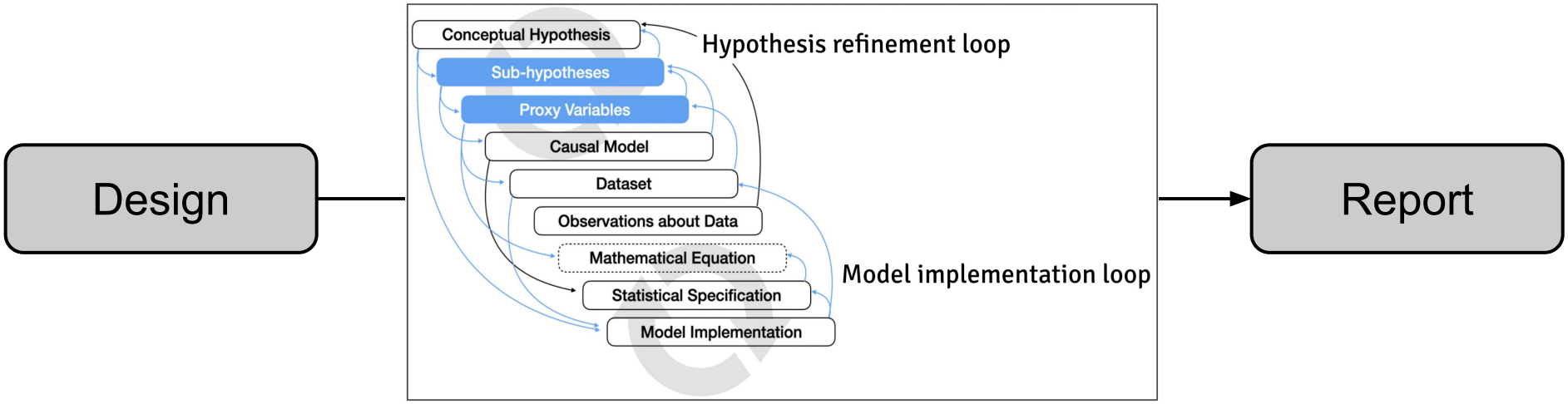
Today

- Effective tables and visualizations
  - Tables vs. graphs
  - Effective tables (xtable and knitr::kable)
  - Effective visualizations (ggplot2)

Tables vs. graphs

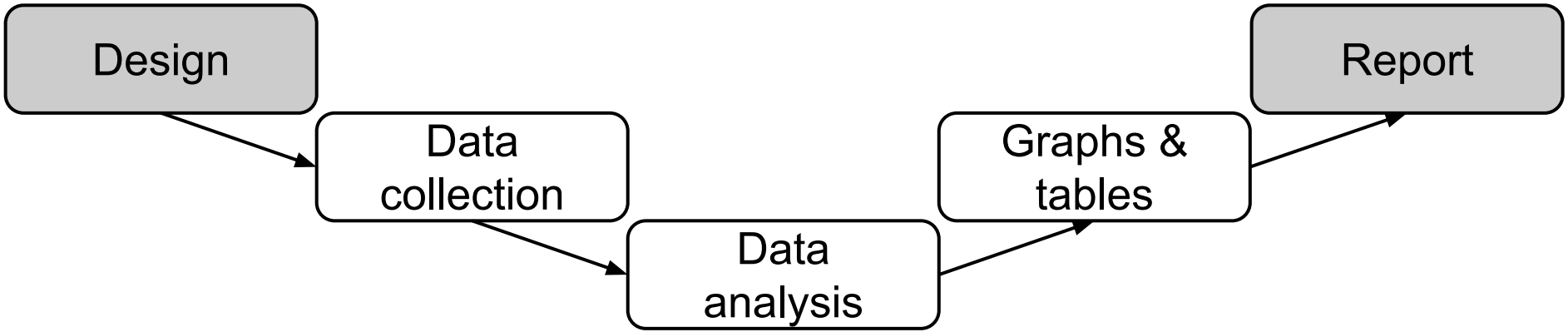
From analysis design to report

From analysis design to report



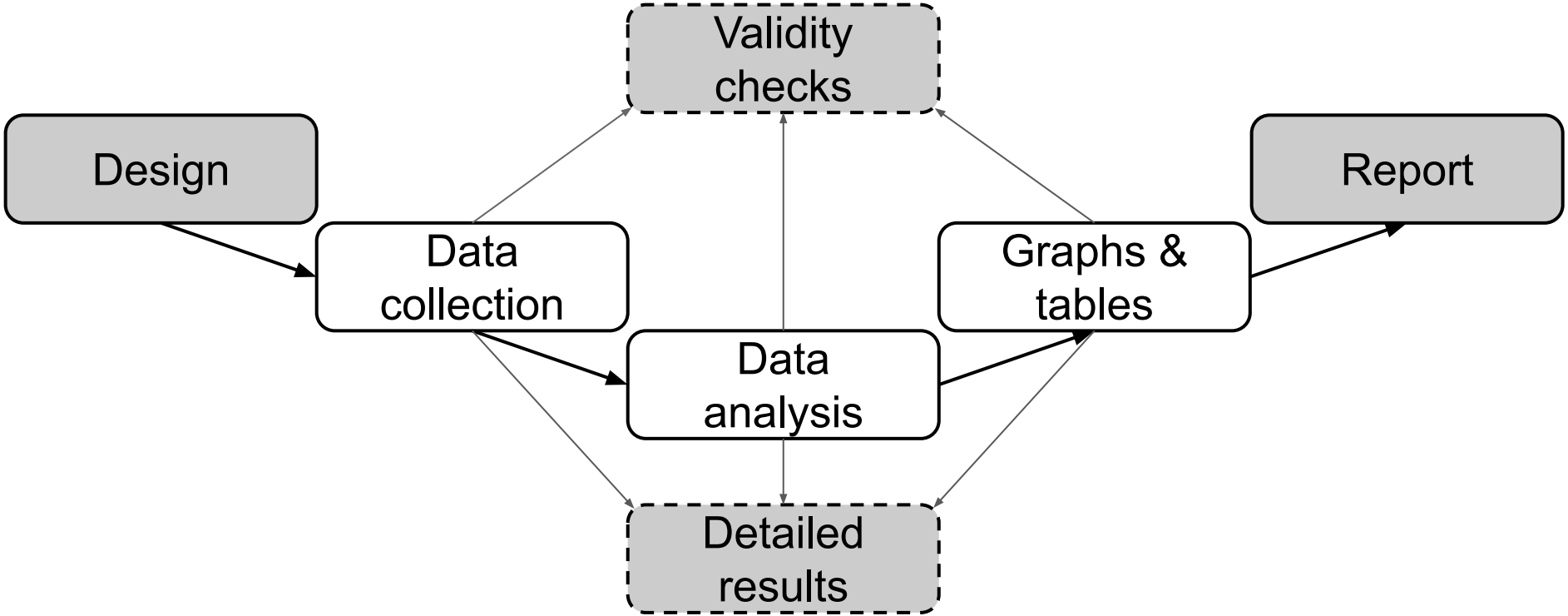
Hypothesis formalization: Empirical findings, software limitations, and design implications,
Jun et al., TOCHI 2022.

From analysis design to report



Do all analysis results go into the final report?

From analysis design to report

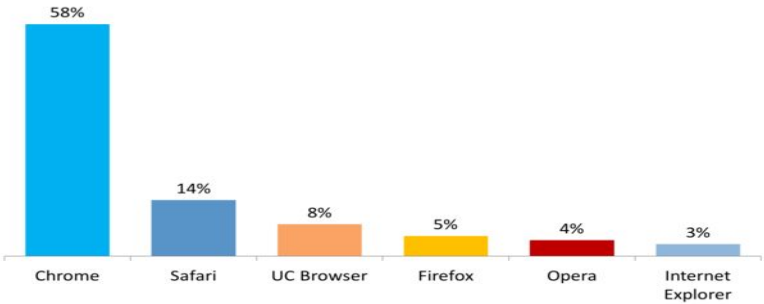


Tables vs. graphs

- When are tables useful?
  - Compare individual values
  - Values involve multiple units
  - Precise values are important

	Browser	Market share (%)	
		June 08	July 09
All users	Internet Explorer	75.4	67.7
	Firefox	18.9	22.5
	Safari	2.8	4.1
	Chrome	—	2.6
	Opera	2.1	2.0
	Netscape	0.5	0.7
	Other	0.2	0.5

- When are graphs useful?
  - Consider an entire set of values
  - Visualize trends and patterns
  - Relationships are more important than precise values



Market share 2018

Effective tables

Effective tables: the run-time data set

```
variant,naive,caching,forking,run,subject
11,     309.8,157.6,  144.8,  1,  "tax"
12,     379.5,237.4,  254.5,  1,  "tax"
13,     415.9,225.9,  225.9,  1,  "tax"
...
```

- Recall the run-time data set
  - 3 subjects (tax, tictactoe, triangle)
  - 3 strategies (naive, caching, forking)
  - 5 runs to account for the variation in run time

Goal: show run times and relative improvements in a table

Effective tables: layout

TABLE I
RUN TIMES AND IMPROVEMENTS.

Subject	RT-naive	RT-cache	RT-fork	I-cache	I-fork
tax	504.11	247.01	195.42	51.02%	61.31%
tictactoe	17.44	16.32	15.43	6.31%	11.49%
triangle	3.13	2.79	1.67	10.91%	46.62%

- Recall the run-time data set
  - 3 subjects (tax, tictactoe, triangle)
  - 3 strategies (naive, caching, forking)
  - 5 runs to account for the variation in run time

What are the pros/cons of Table I?

How would you improve it?

Effective tables: layout

Compare the two w.r.t. readability, clarity, and interpretability

TABLE I
RUN TIMES AND IMPROVEMENTS.

Subject	RT-naive	RT-cache	RT-fork	I-cache	I-fork
tax	504.11	247.01	195.42	51.02%	61.31%
tictactoe	17.44	16.32	15.43	6.31%	11.49%
triangle	3.13	2.79	1.67	10.91%	46.62%

TABLE II
RUN TIMES AND IMPROVEMENTS FOR THE **NAIVE**, CACHING (**CACHE**), AND FORKING (**FORK**) STRATEGIES. RUN TIMES ARE GIVEN IN SECONDS AND AVERAGED OVER FIVE RUNS.

Subject	**Run times**			**Improvements**	
	naive	**cache**	**fork**	**cache (vs. naive)**	**fork (vs. naive)**
Tax	504	247	195	51.0%	61.3%
TicTacToe	17.4	16.3	15.4	6.31%	11.5%
Triangle	3.13	2.79	1.67	10.9%	46.6%

Effective tables: content

Keep it simple

- Avoid mixing higher-is-better and lower-is-better numbers
- Allow for easy comparisons, primarily by row
- Be consistent about precision vs. significant digits
- Summarize the table (what is the bottom line?)

TABLE II

RUN TIMES AND IMPROVEMENTS FOR THE **NAIVE**, CACHING (**CACHE**), AND FORKING (**FORK**) STRATEGIES. RUN TIMES ARE GIVEN IN SECONDS AND AVERAGED OVER FIVE RUNS.

Subject	Run times			Improvements	
	naive	cache	fork	cache (vs. naive)	fork (vs. naive)
Tax	504	247	195	51.0%	61.3%
TicTacToe	17.4	16.3	15.4	6.31%	11.5%
Triangle	3.13	2.79	1.67	10.9%	46.6%

Effective tables: summaries

Subject	LOC	Speed up
Tax	8900	10.2%
TicTacToe	120	54.2%
Triangle	80	60.9%
Average	3393	41.8%

Total **vs.** **Average**

What are the downsides of these summaries?

Effective tables: best practices

Do

- **Use xtable (standalone scripts) and knitr::kable (Quarto)!**
- Automatically generate all table bodies for publications
- Make each table self-contained (content and caption)
- Use descriptive (hierarchical) headers
- Right align numbers
- Use meaningful totals or weighted averages
- Be consistent about precision vs. significant digits

Don’t

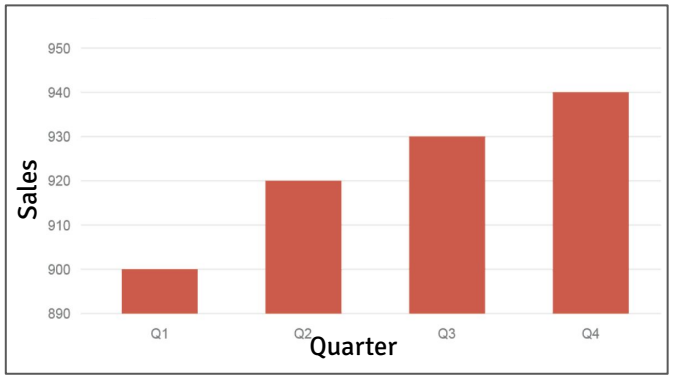
- Don’t use horizontal lines between related rows
- Don’t use vertical lines between related columns

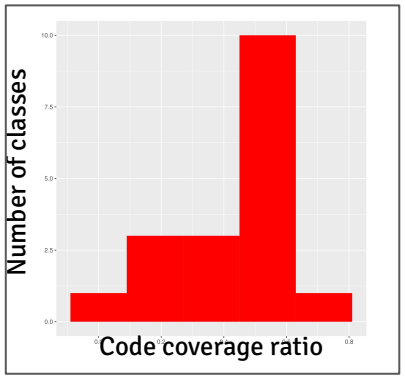
Effective graphs

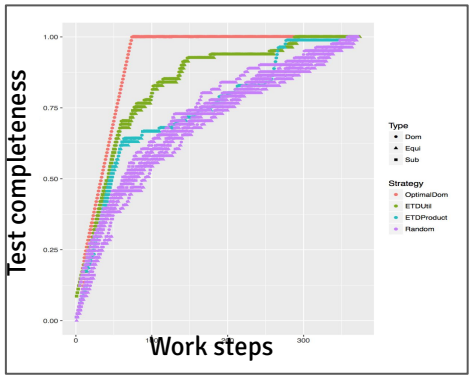
4 beautiful graphs

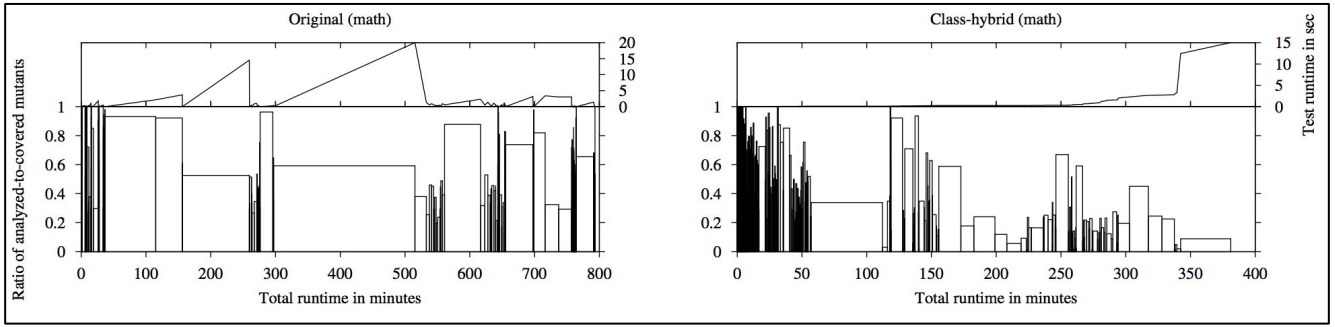
- Small groups of 2-3 students
- 4 example graphs
- For each graph
  - Discuss pros and cons
  - Propose improvements

4 beautiful graphs

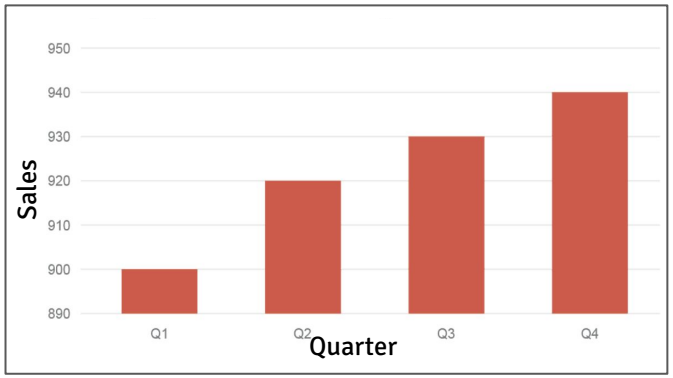






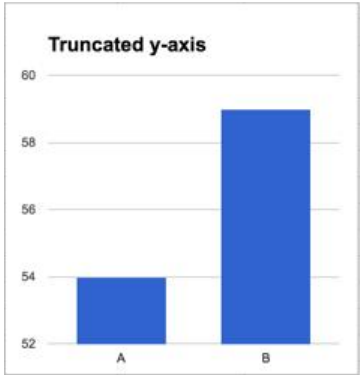


Example 1: bar charts



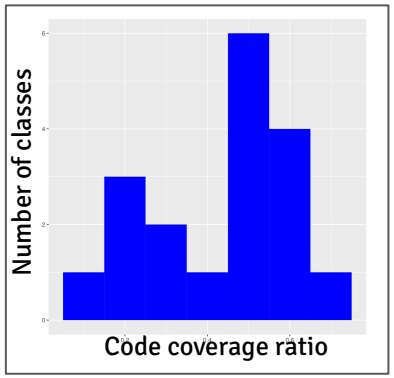
Truncated axes are misleading and
not a proper way to “demonstrate” effect size!

Example 1: bar charts

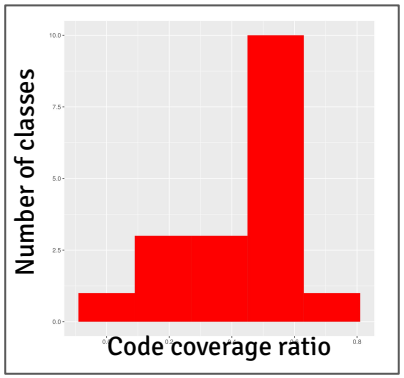


Truncated axes are misleading and
not a proper way to “demonstrate” effect size!

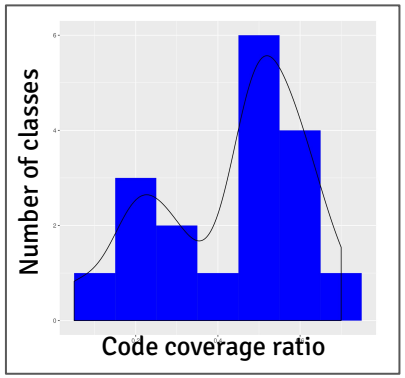
Example 2: histogram vs. density plot



Adequate binning



Changed binning

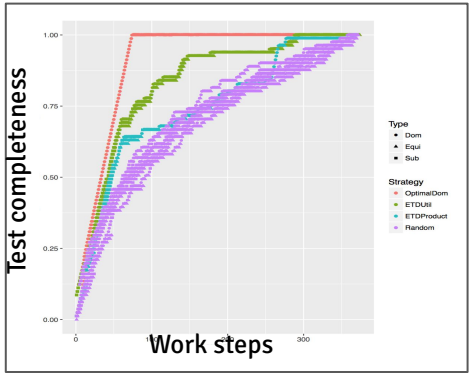


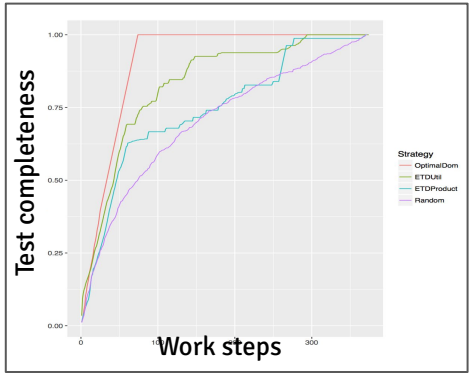
Kernel density overlay

Good visual summary of count data, but binning may be misleading.

Kernel density overlay can provide information about adequate binning.

Example 3: scatter plot vs. line plot



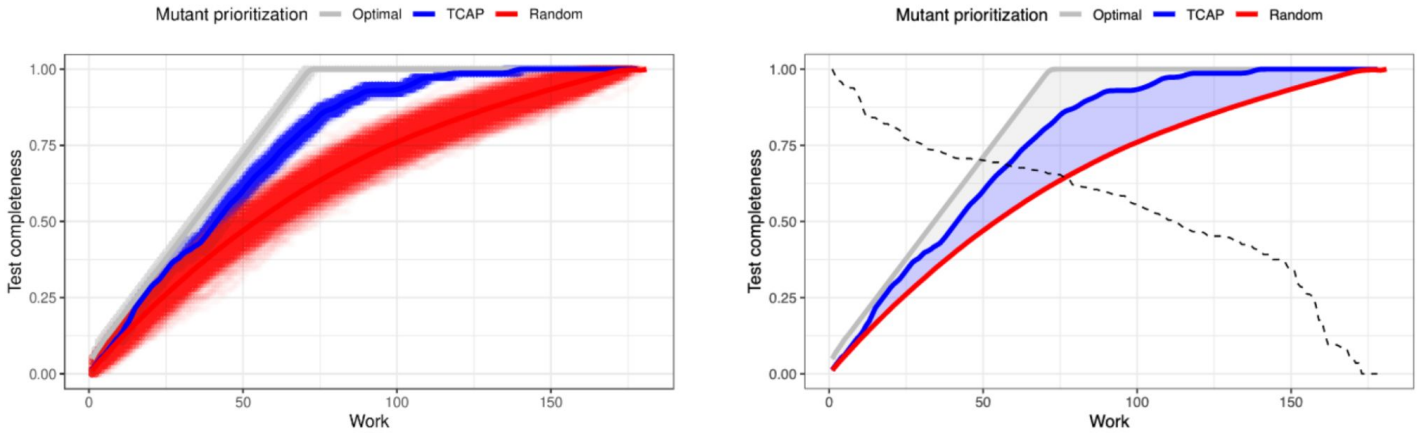


Good visual summary of point clouds, trends, and relationships.

May obscure relevant trends (overlapping points).

Hard to reason about density (without adding transparency).

Example 3: scatter plot vs. line plot

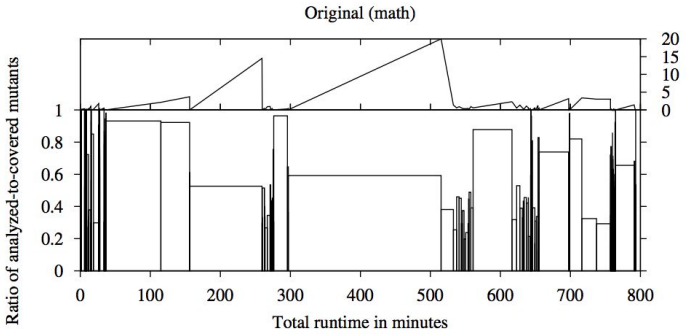


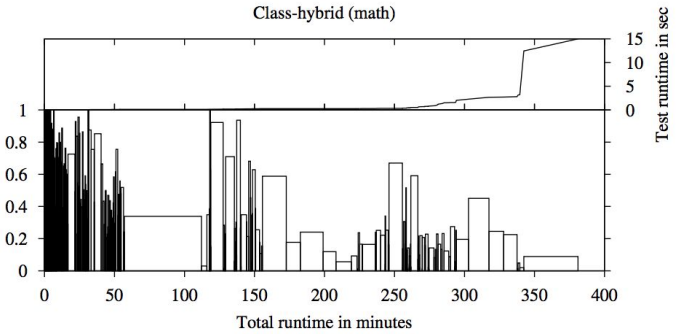
Good visual summary of point clouds, trends, and relationships.

May obscure relevant trends (overlapping points).

Hard to reason about density (without adding transparency).

Example 4: multi-plot visualization



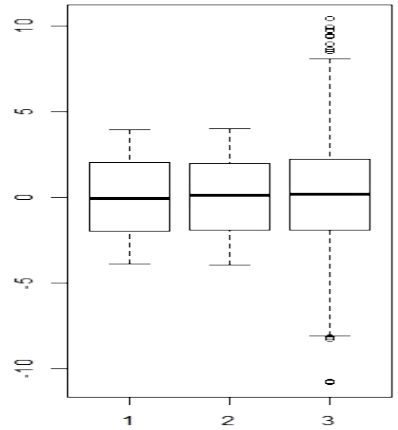


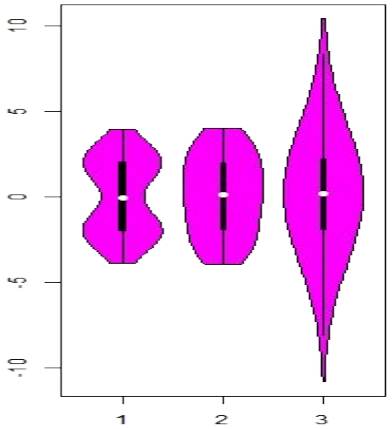
Way too many details!

The key trends and takeaways are obscured.

Good for detailed results but not a final report.

Effective graphs: box plots vs. violin plots

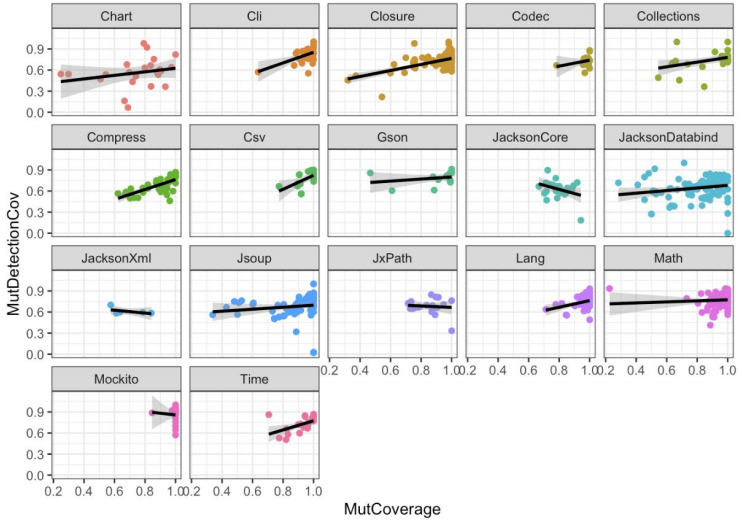




Box plots:

- Good visual data summary
- Nicely complements hypothesis tests
- May be misleading for multimodal data
- May be misleading for small samples

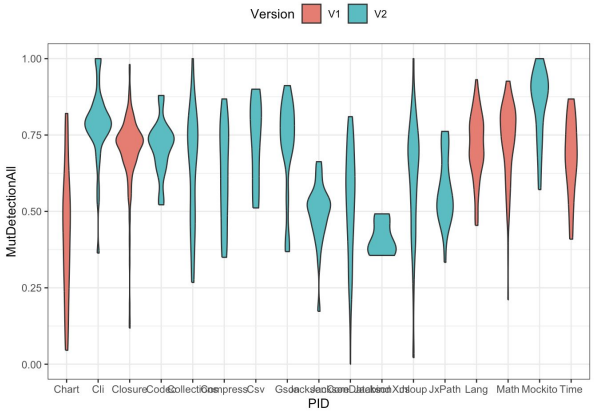
Effective graphs: facet plots



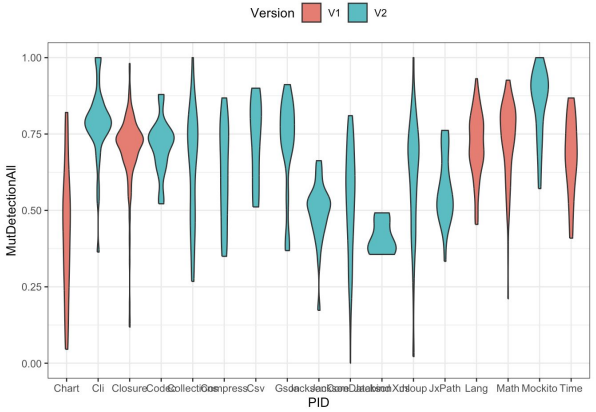
Facet plots:

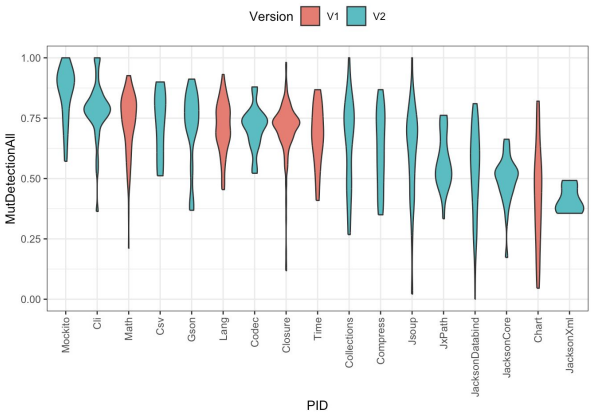
- Clean visualizations across multiple dimensions of interest
- Allows for comparisons within groups and across groups
- Complementary to other ggplot2 aesthetics (color, shape, etc.)
- Use ggplot's `facet_grid` for cross-product visualizations (formula syntax)

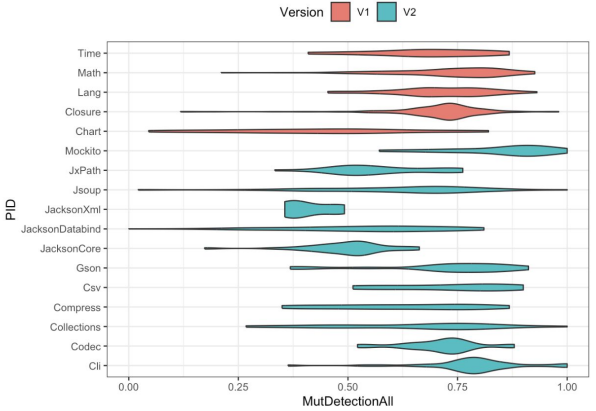
Effective graphs: reorder and/or flip axes



Effective graphs: reorder and/or flip axes







Reorder and/or flip axes:

- Reorder by mean/median or by groups of interest etc.
- Flip axes for readability if appropriate
- Favor short labels over rotated labels

Effective graphs: best practices

Do

- **Use ggplot2!**
- Make each plot self-contained (content and caption)
- Relate tables and graphs to tell a consistent story
- By default put the DV on the vertical axis
- Reduce complexity with facet plots

Don’t

- Don’t use multiple, unrelated axes
- Don’t connect unrelated data points
  (choose an appropriate graph instead)