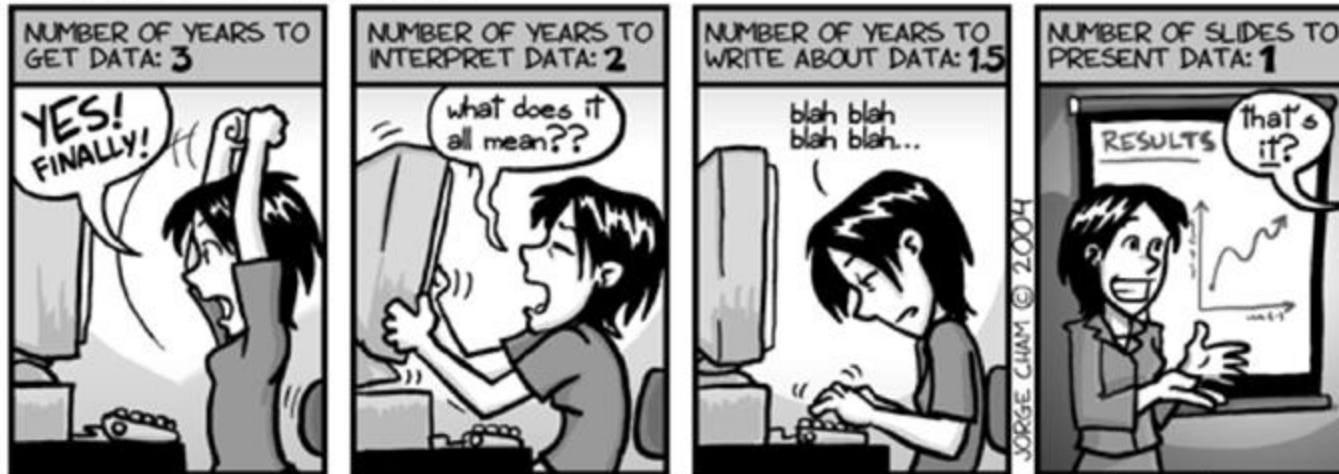


Data visualization

Ivano Malavolta

DATA: BY THE NUMBERS

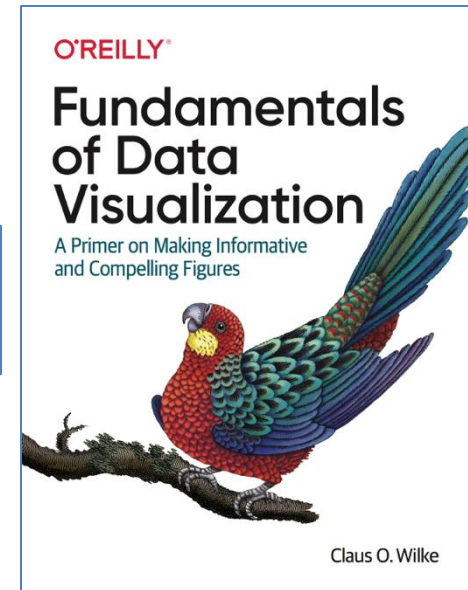


www.phdcomics.com

Resources

On-line book on data visualization

<https://clauswilke.com/dataviz/index.html>

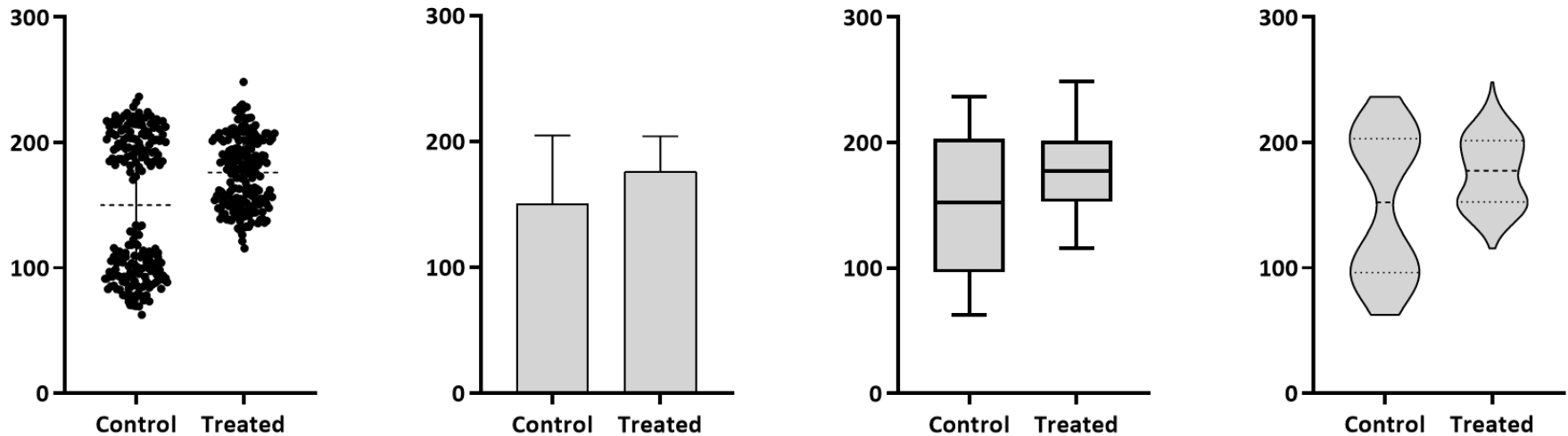


- Data to Viz decision tree: <https://www.data-to-viz.com>
- Data visualization caveats: <https://www.data-to-viz.com/caveats.html>
- Gallery (with R source code): <https://r-graph-gallery.com/>

Suggestions about data visualization

- Plotting is an art
- First think about the main message you want to give with your plot, then decide which type of plot better fits your message
e.g., comparison vs absolute values, demographics, distributions
- Get inspiration from the studies in the *papers_experiments* folder in Canvas

Is the boxplot my silver bullet for plotting?



Boxplots are generally good for comparing (1) the **central tendency** and (2) **spread** of your data

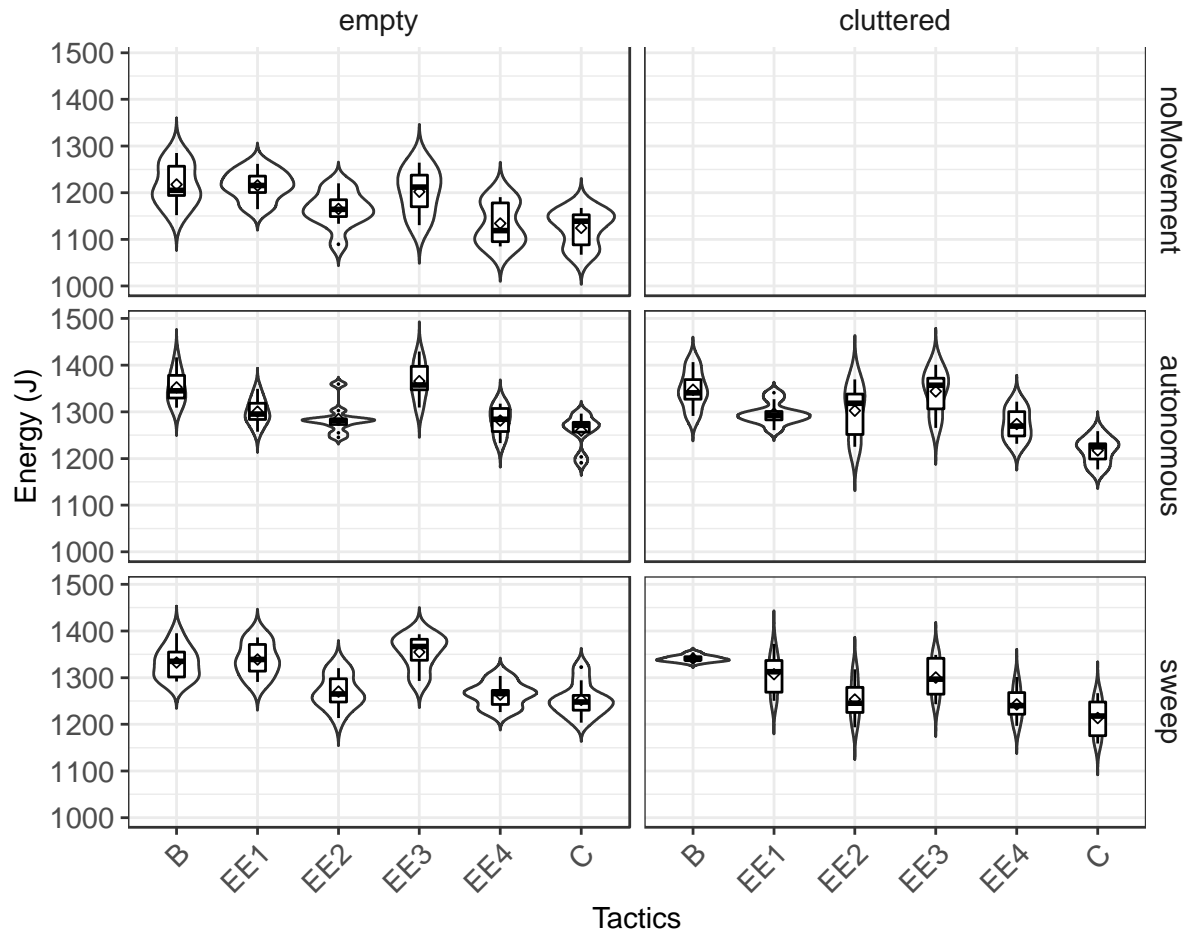
What you lose with boxplots is a precise visualization of the **distribution** of your data

For example, if your data is bimodal, you do not see it in a boxplot

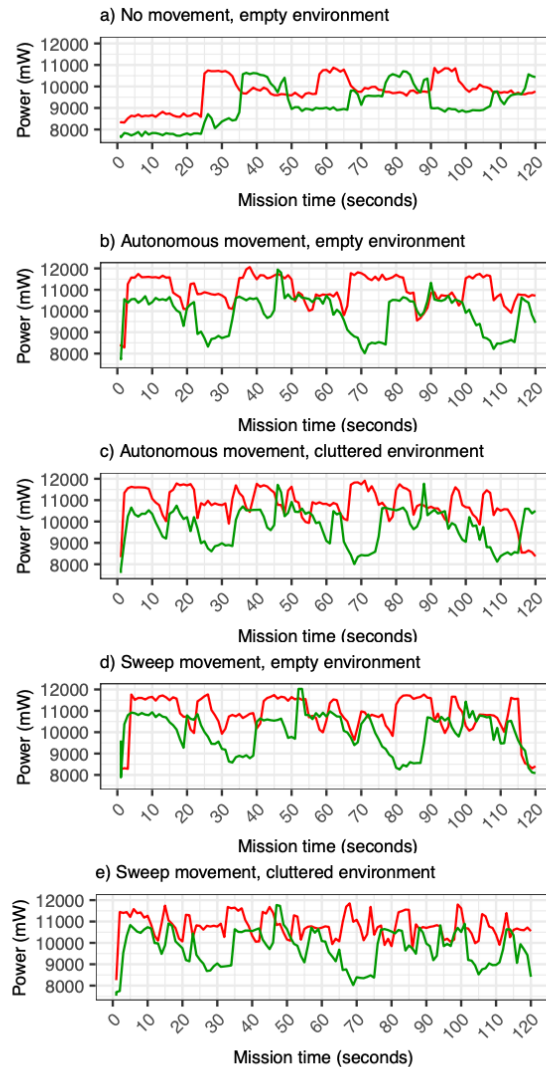
Examples of visualizations



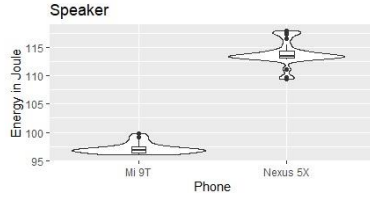
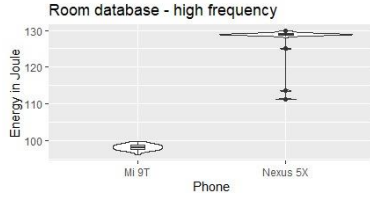
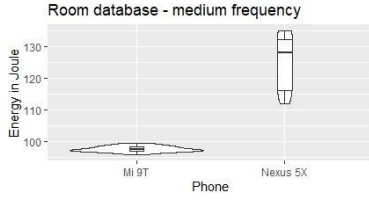
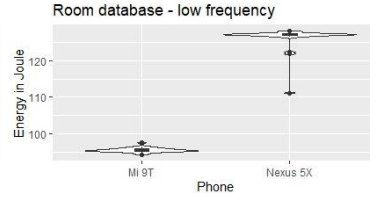
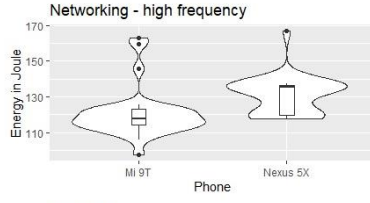
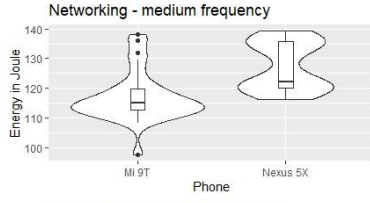
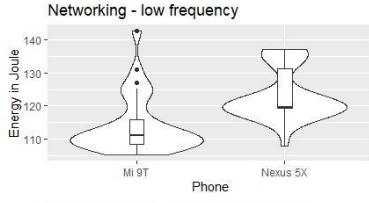
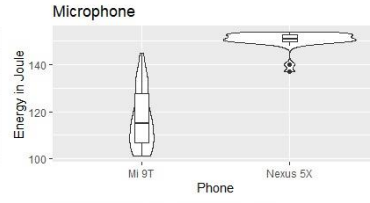
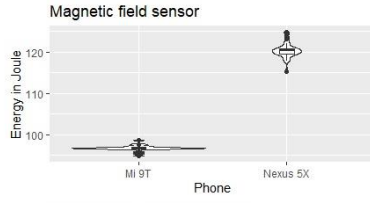
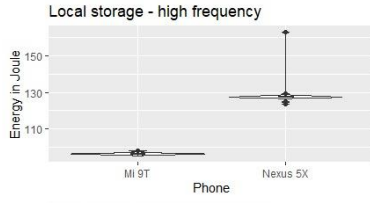
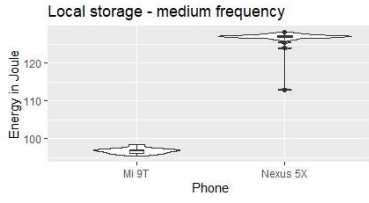
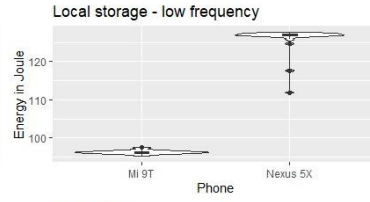
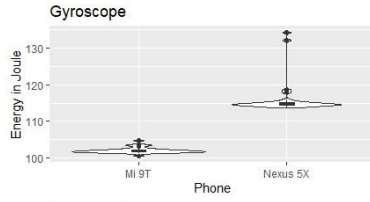
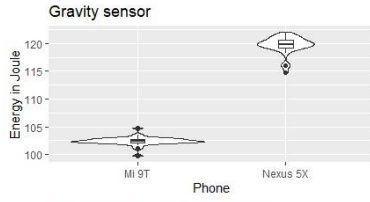
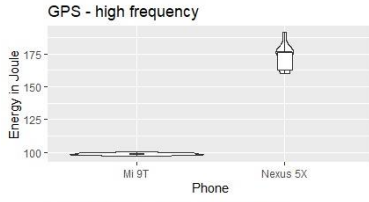
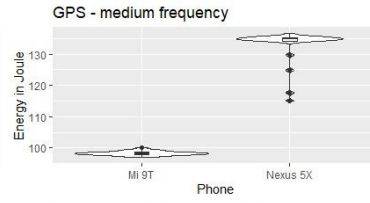
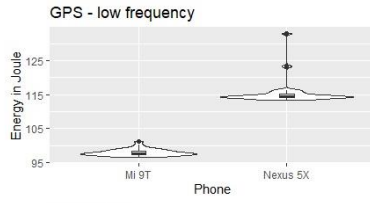
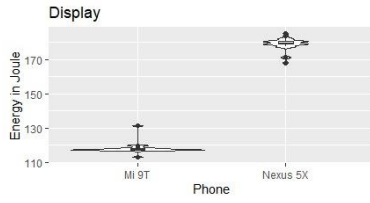
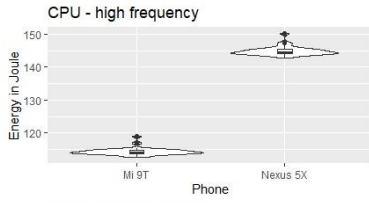
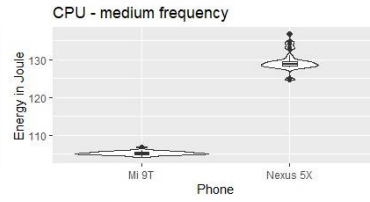
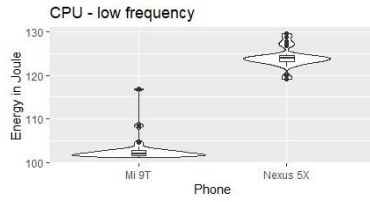
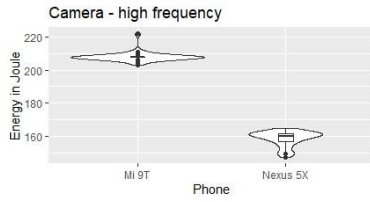
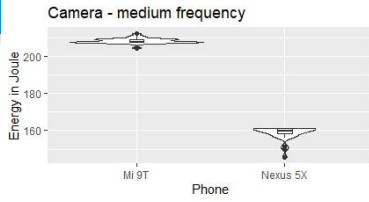
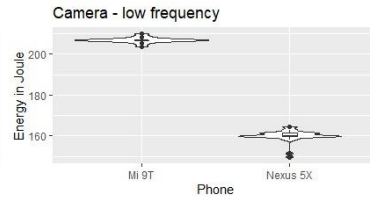
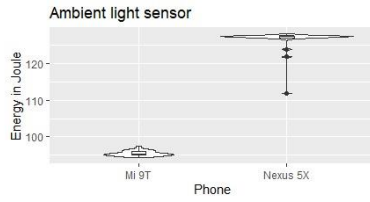
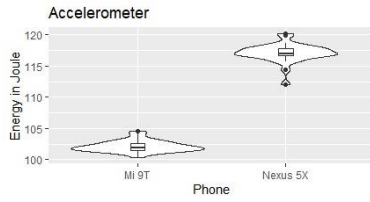
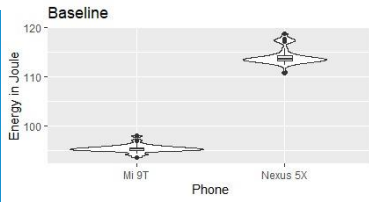
Examples of visualizations



Examples of visualizations

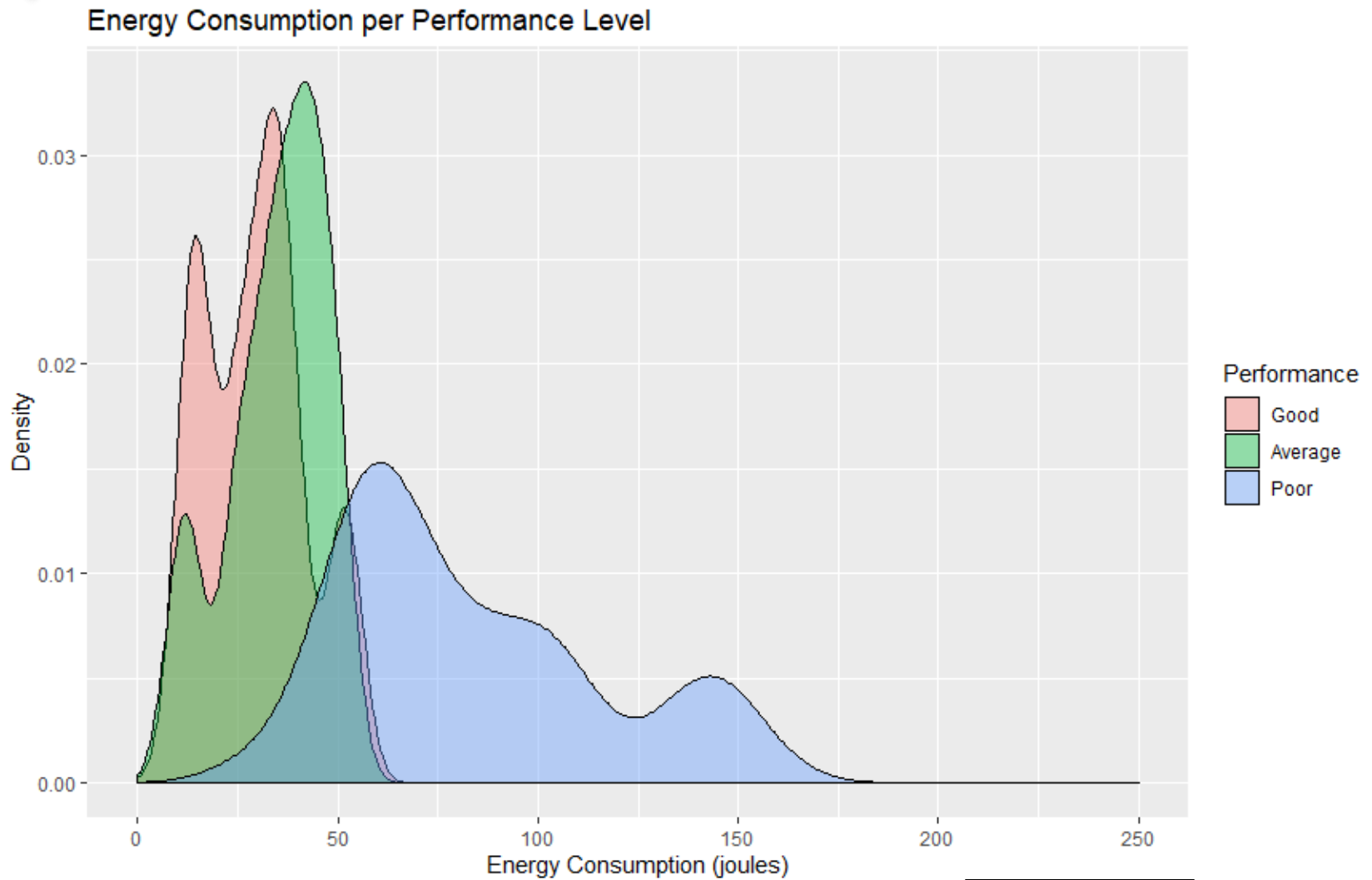


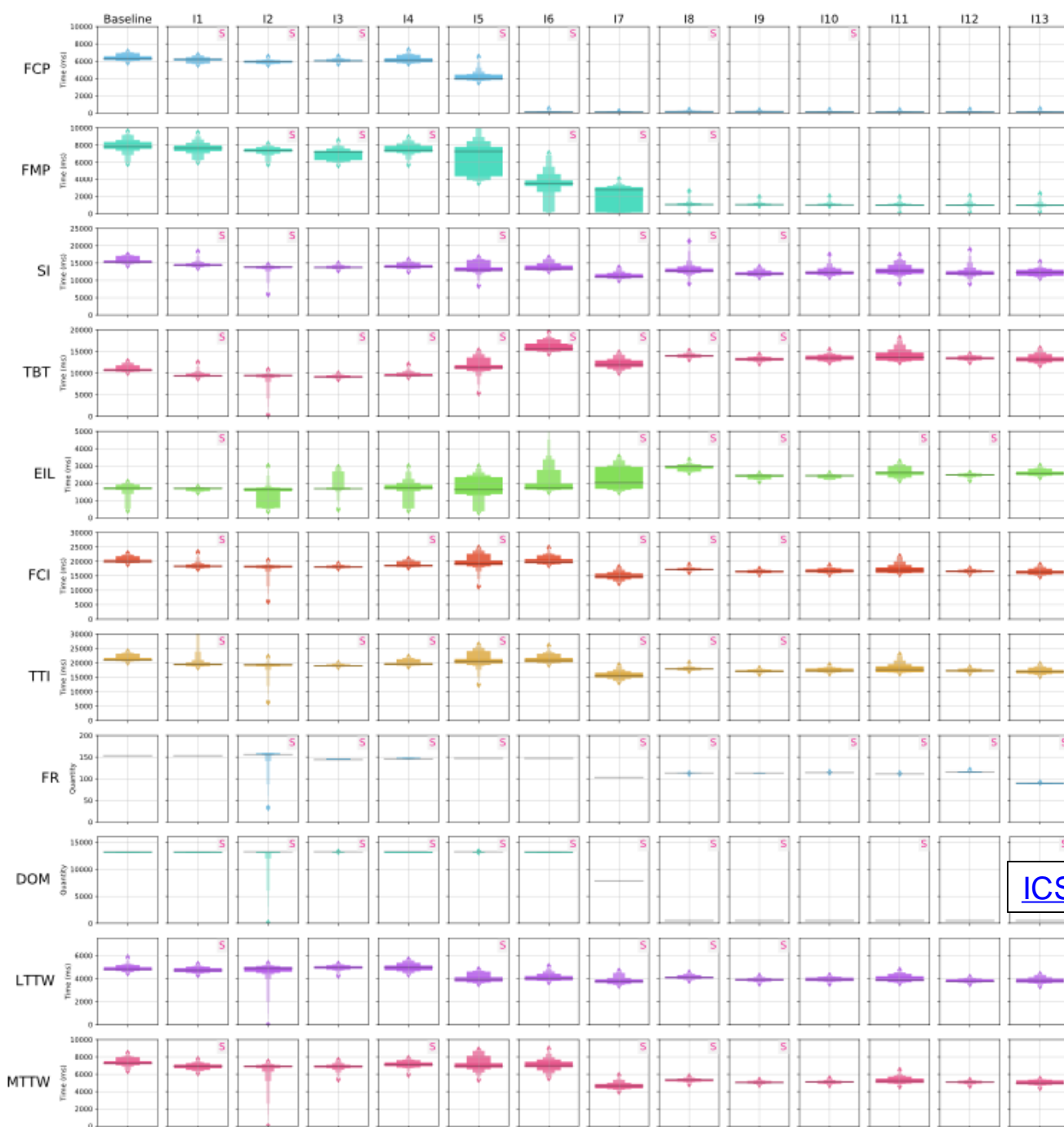
MSR 2021



A mobile 2020

Examples of visualizations





ICSME 2020

Fig. 2: Overview of the performance metrics across all interventions (S = statistically significant change after the intervention)

Tables shall be used to complement with fine-grained data

TABLE III: Final results of the case study

Metric	Baseline	After I_{13}	Δ %	p-value
FCP	6,290.52	153.33	-97.56%	2.59e-11
FMP	7,646.05	953.33	-75.01%	2.59e-11
SI	15,310.98	12,271.07	-19.85%	1.16e-10
TBT	10,708.86	13,180.75	23.08%	1.42e-10
EIL	1,701.13	2,549.60	49.88%	2.59e-11
FCI	19,928.40	16,344.58	-17.98%	2.59e-11
TTI	21,018.35	16,958.08	-19.32%	2.59e-11
NR	153	153	-41.18%	1.97e-12
DOM	13,219	507	-96.16%	1.97e-12
LTTW	4,796.3	3,798.72	-20.80%	2.59e-11
MTTW	7,280.71	5,006.26	-31.23%	2.59e-11

[ICSME 2020](#)

Top-down approach

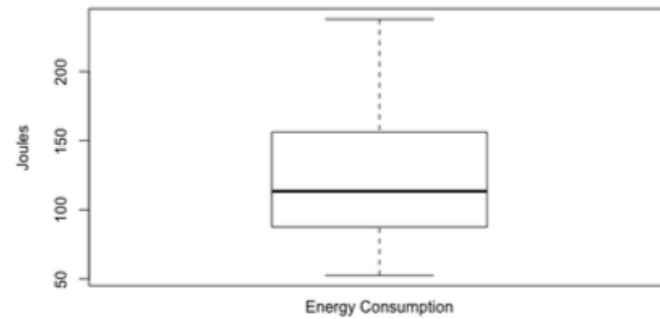


Fig. 2. Measured energy consumption values

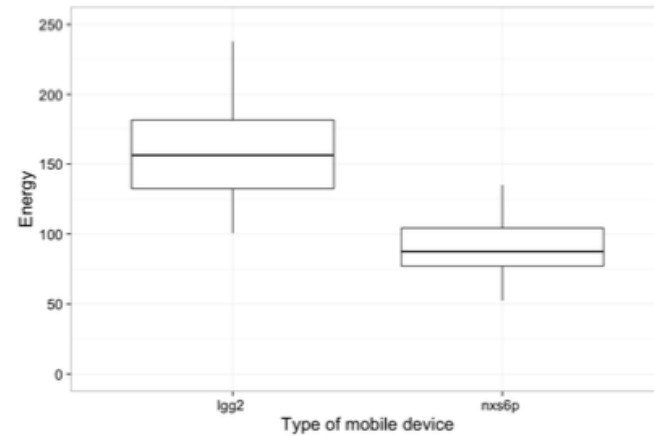


Fig. 3. Measured energy consumption values per mobile device (in Joules)

A. Descriptive Statistics

The energy consumption of all PWAs (two versions for each PWA, with service workers on and off) of our dataset is summarized in Table III.

[MOBILESoft 2017](#)

	Energy Consumption					
Phone	Min.	Max.	Median	Mean	SD	CV
<i>Both</i>	52.44	237.90	113.41	124.33	45.32	0.36
<i>LG G2</i>	100.61	237.90	156.21	157.38	37.78	0.24
<i>Nexus</i>	52.44	134.97	87.44	91.28	22.14	0.24

TABLE III
ICS FOR THE ENERGY CONSUMPTION VALUES (IN
STANDARD DEVIATION, CV = COEFFICIENT OF
VARIATION)

Show all subjects to spot interesting cases

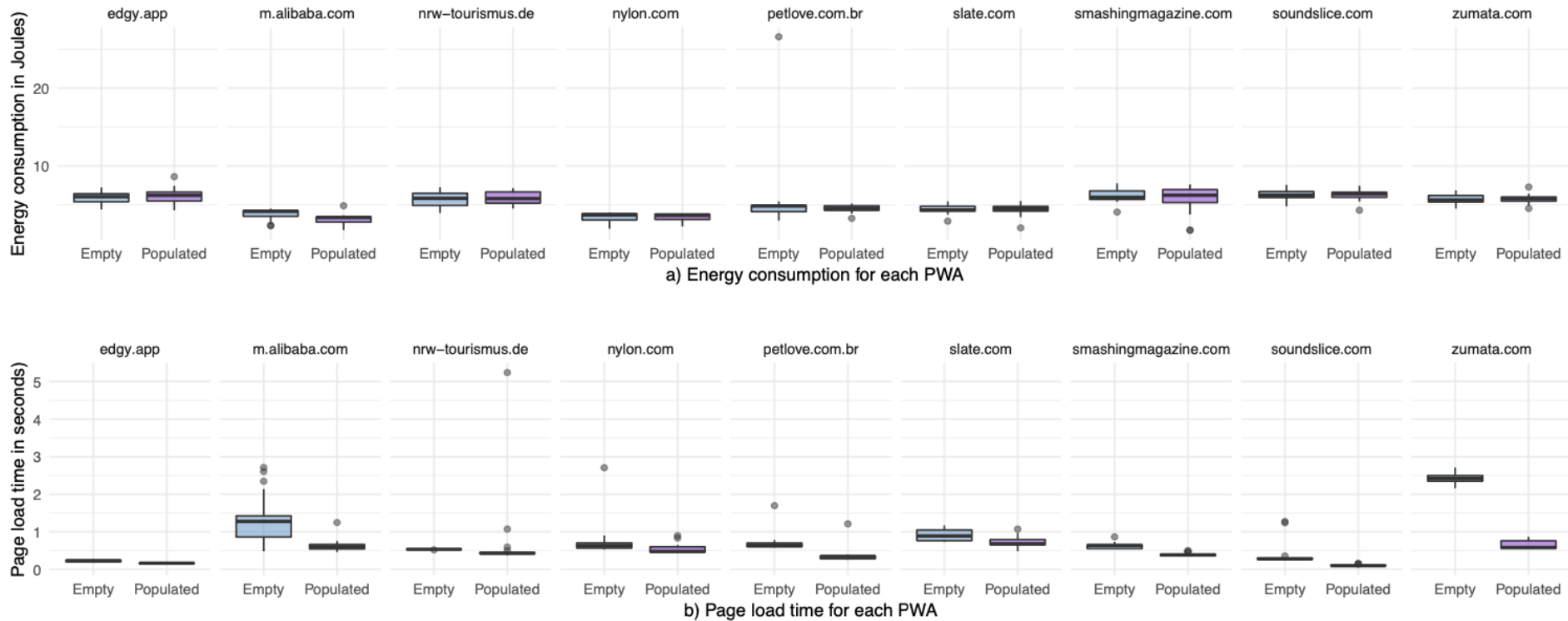


Figure 6: Energy consumption and page load time for each PWA

[MOBILESoft Caching PWA 2020](#)

What this lecture means to you?

- Think about the main message first, then choose the type of plot

Do not reinvent the wheel, follow what other researchers in the community are doing (and the resources in the first slide of the lecture)

- Complement plots with tables for the details
- Writing tip: first put plots and tables in the section, then describe them one by one in a top-down fashion
- Use colors parsimoniously

when you do, use them consistently across the whole report