John W. Tukey: Contributions to Statistics

John W. Tukey (1915–2000)

John Tukey was one of the most influential statisticians of the 20th century — blending deep theory, clever computation, and practical tools for understanding data.

Major Contributions & Inventions

Exploratory Data Analysis (EDA)

- Pioneered the **EDA mindset** — using graphics to *discover* rather than merely confirm.
- Famous quote: *"The greatest value of a picture is when it forces us to notice what we never expected to see."*

Boxplot & Five–Number Summary

- Invented the **box-and-whisker plot** (1970) and promoted the **five-number summary** (min, Q1, median, Q3, max).
- Still core to R (`boxplot()`, `fivenum()`).

Stem-and-Leaf Plot

- A quick, hand-calculable histogram alternative.
- Still available in R with `stem()`.

Tukey's HSD

- **Honestly Significant Difference** post-hoc test after ANOVA.
- R: `TukeyHSD()`.

Tukey's Ladder of Powers

- Transformations (log, √, reciprocal) to stabilize variance & symmetrize data.
- Basis for modern Box–Cox & `car::powerTransform()`.

Median Polish

- Robust decomposition for two-way tables (row + column + residual).
- R: `medpolish()`.

Resistant Lines & Robust Stats

- Advocated for **resistant** (robust) summaries & regressions decades before they were mainstream.

Fast Fourier Transform (FFT)

- Co-invented the **FFT** (Cooley–Tukey, 1965) — a cornerstone of signal processing.

Language & Ideas

- Coined **"bit"**, **"software"**, and popularized "ANOVA."
- Championed **graphics as thinking tools**.

Legacy in R

Tukey Method	R Function / Package
Boxplot, 5-number	`boxplot()`, `fivenum()`
Stem-and-leaf	`stem()`, `aplpack::stem.leaf()`
Tukey's HSD	`TukeyHSD()` (base)
Median Polish	`medpolish()` (base)
Ladder of Powers	`rcompanion`, `car`
Resistant Lines	`MASS::rlm()`
FFT	`fft()`

"An approximate answer to the right question is worth a great deal more than a precise answer to the wrong question."
— John W. Tukey