

## Article 2 : Principles and Procedures of Exploratory Data Analysis

The article discusses the increasing visibility of Exploratory Data Analysis (EDA) in the field of psychology, driven by the widespread availability of software for graphical data analysis and calls for its increased usage on epistemic grounds. Despite this visibility, the article highlights a significant gap in psychologists' explicit training in EDA beliefs and procedures.

The author references Huberty (1991) who noted that while statistical texts may briefly mention common EDA techniques, such as stem-and-leaf plots or box plots, they seldom integrate these techniques throughout. A survey of graduate training programs in psychology conducted by Aiken et al. (1990) corroborates this impression, revealing that only a minority of departments teach aspects of EDA in introductory graduate courses, and even fewer students can apply these learned techniques effectively.

The article aims to introduce researchers to the philosophical underpinnings and general heuristics of EDA through three sections: background and first principles, a primer on EDA heuristics, and implications for psychological methodology and training.

In the background and first principles section, the article clarifies the definition of EDA as a specific tradition of data analysis stemming from the work of John Tukey and associates since the early 1960s. This tradition emphasizes understanding data substantively, utilizing graphic representations, tentative model building, robust measures, and positions of skepticism and flexibility regarding methods.

The goal of EDA, akin to detective work, is to discover patterns in data, recognizing that the resulting description may not hold in subsequent samples. The article references classical works in this tradition and highlights the importance of distinguishing between exploratory and confirmatory modes of data analysis.

1. **Tukey's Perspective:** Tukey emphasized the value of graphics in EDA, stating that the greatest value of a picture is when it reveals unexpected insights in the data.
2. **Role of Graphics:** Graphics are essential in EDA because they can depict numerous data values simultaneously, unlike algebraic summaries which may overlook important patterns.
3. **Classical References:** The passage mentions several classical references in statistical graphics, highlighting the importance of this field in data analysis.
4. **Graphical Devices:** Various graphical devices, such as stem-and-leaf plots, dot plots, box plots, and kernel density smoothers, are discussed as effective tools for exploring and summarizing data.
5. **Interactive Graphics:** Interactive graphics, which allow users to manipulate and explore data dynamically, are highlighted as valuable for holding and assessing multiple working hypotheses.
6. **Example Analysis:** An example analysis using different graphical techniques is provided, demonstrating how each method can reveal different aspects of the data and contribute to exploratory analysis.

