Post hoc exploratory analysis

Post hoc exploratory analysis is a type of data analysis performed after the initial hypothesis testing or planned analyses have been completed. It involves examining the data for interesting patterns, trends, or relationships that were not originally anticipated or tested for. The term “post hoc” is Latin for “after this,” which emphasizes that the analysis occurs after the data has been collected and the initial analyses have been conducted.

Key characteristics of post hoc exploratory analysis include:

Flexibility: Unlike hypothesis-driven analysis, which follows a predetermined plan, post hoc exploratory analysis allows researchers to be flexible in exploring the data and generating new hypotheses.

Data-Driven: The analysis is driven by the data itself, rather than preconceived hypotheses. It involves visually inspecting data, calculating summary statistics, and using various data mining techniques to uncover interesting relationships or patterns.

Hypothesis Generation: Through post hoc exploration, researchers may generate new hypotheses that can be tested in future studies. These new hypotheses can lead to a deeper understanding of the data or phenomena under investigation.

Risk of Bias: Since the analysis is not based on predetermined hypotheses, there is an increased risk of false discoveries or spurious correlations. It is essential to be cautious and consider these findings as preliminary until they are confirmed through further research or experimental validation.

Exploratory Visualizations: Post hoc exploratory analysis often involves creating visualizations, such as scatter plots, histograms, heatmaps, or other graphical representations, to gain insights and identify potential relationships within the data.

No P-hacking: Researchers should be careful not to engage in “data fishing” or “p-hacking,” which means selectively analyzing the data until a significant result is found. This can lead to false-positive findings and undermine the credibility of the analysis.

Post hoc exploratory analysis can be a valuable step in the research process, as it helps researchers gain a more comprehensive understanding of the data and may lead to new directions for further investigation. However, it should be transparently reported as an exploratory analysis, distinct from pre-planned hypothesis testing, to avoid misinterpretations and overgeneralizations.