What is Exploratory Data Analysis

Edris Safari

Bellevue University, Nebraska U.S.A.

Abstract

Exploratory Data Analysis was coined by John W. Tukey, who wrote the book Exploratory Data Analysis in 1977. It has been widely known and accepted that it is an essential part of any data science project. Methodologies such as the CRISP-DM, SEMMA , JTA, and others highlight the necessity of EDA in any data science project. This paper offers a brief description of the process of data exploration and its role in a data science project.

Keywords: CRISP-DM,SEMMA,JTA,EDA

What is Exploratory Data Analysis

In the discussion of data science, the terms Machine Learning, and Data Mining are used interchangeably and often in the same sentence. The words probability & statistics, algorithms, and various algorithm techniques are also widely used in data science discussions. The bottom line for any data science project is that it should produce and answer to a question. The answer should add value to the questioner who posed the question in the first place. The questions are typically asked in a meeting room filled with executives, managers and directors. These people

References

https://www.kdnuggets.com/2017/04/value-exploratory-data-analysis.html

Kelleher, John D.. Data Science (MIT Press Essential Knowledge series) (p. 1). The MIT Press.