Stephanie Nelson

April 13th, 2023

Professor Bemley

**Title**: “Exploratory Data Analysis”

**Author**: Unknown (Link on Blackboard)

**I. Introduction**

In the article “Exploratory Data Analysis”, We get a better understanding of what data is, the different categories, types of EDA, Characteristics of data, types of data plots, and correlation. We understand the different thought processes of EDA. This gives us a step-by-step breakdown of the EDA process.

**II. Exploratory Data Analysis**

This article breaks down how to look at data and the process of EDA. The EDA process is to prevent mistakes in data, random assumptions, how to choose correct models and data relationships. An important part of collecting and cleaning data is how to properly represent the data. If you are unable to visualize the data to show your assumptions and findings. EDA is broken up into non-graphical which include summary statistics. Whereas graphical findings represent the statistics in graph forms. Categorical data sort data into categories. For example, the Colors of cars would be represented as red, blue, pink, orange, white, and black. Some types of graphs you can use in EDA are histograms, box plots, stem and leaf plots, Q plots, and Cross tabulation data. Plots and graphs allow your audience to understand your findings so it is crucial to choose plots that correctly display the information. It is also vital to look at your axis and change it to display your data to best represent your findings.

**III. Conclusion**

In all, the article played a key role in explaining how to begin turning your findings from exploratory data analysis into graphs. These allow the audience to understand your findings as well as tie them to the bigger picture of your research. EDA is an important part of the data science pipeline and is needed to display your findings.

**III. References**

Chapter 4: Exploratory Data Analysis. <https://bsuonline.blackboard.com/bbcswebdav/pid-17985980-dt-content-rid-28398452_1/xid-28398452_1>