

Introduction

This lesson will introduce exploratory data analysis.

WE'LL COVER THE FOLLOWING



- What is Exploratory Data Analysis?

What is Exploratory Data Analysis?

Exploratory Data Analysis (EDA) is the process of analyzing datasets with the aim of understanding them more deeply. As the term “exploratory” suggests, during EDA the focus is to *explore* or *understand* the data better. We try to make sense of the data in the context of our problem.

During EDA, we can do any kind of analysis that gives us some insight into our data. Some of the common practices in EDA are:

- Looking at the data types of the variables
- Identifying the most important variables
- Looking at the distributions of the variables
- Summarizing the data
- Finding biases in the data
- Looking at the different trends in data
- Studying relationships among quantities
- Spotting anomalies in the data
- Visualizing the data





We will use some of the techniques we learned in chapter 2 to perform EDA and learn some new techniques as well. Python is well suited for this kind of work. So, in the next lesson, we will look at analyzing individual columns.