Exploratory Data Analysis

DevFest 2020

Description

Exploratory Data Analysis (EDA) is an approach to extract the information enfolded in the data and summarize its main characteristics . It is considered to be a crucial step in any data science project since it allows us to uncover patterns and insights within our data. The expected output of this challenge is a well presented notebook that illustrates the Exploratory Data Analysis applied to the given dataset.

To pass the Challenge:

Create a notebook in which you do the following:

1. Identification of variables and data types
2. Analyzing the basic metrics
3. Non-Graphical Univariate Analysis
4. Graphical Univariate Analysis
5. Bivariate Analysis
6. Variable transformations
7. Missing value treatment
8. Outlier treatment
9. Correlation Analysis
10. Dimensionality Reduction

Data guide

These datasets are meteorological data provided by NASA from the HI-SEAS weather station from four months (September through December 2016) between Mission IV and Mission V. The dataset's size is 2.82 MB.

Technical Requirements:

1. Use the given dataset.
2. The result of each treatment should appear in the notebook.
3. The submitted file should have the **.html** extension.
4. the notebook **must be well commented and presented**.