

24th March, 2023

Pandas

import pandas as pd

Exploratory data analysis : This is a method of examining your dataset to identify the main characteristics, relationships between them, outliers or anomalies, etc.

`describe()` → gives info about the descriptive statistics of the data. Mean, standard deviation, max, min and percentile, etc. and count

'count' gives information about missing value.

`corr()` → gives information about the correlation between parameters. This is between

-1 to +1. (+1) means there is a high positive correlation between parameters and vice versa. 0 means no correlation exist between the parameters

To visually see the correlation among parameters one can use

`plt.matshow(df.corr())`