Python
Case study
Data Manipulation
WEEK 3

Data Science
Academy

Disclaimer: The dataset has minor modifications to protect its confidentiality.

You have been given Train Data.
Answer following questions:

1. Show first 7 observations in dataset.
2. Check if there are any missing values.
3. Find mean of each variable according to target.
4. Count number of values for each target.
5. Remove **ID_code** column.
6. Get statistical summary for each variable.
    a. Give interpretation for **std, min, max, mean**.
    b. Show mean of **var_0** by using indexing.
7. Filter **var_0** variable from minimum to mean.
8. Filter **var_6** greater than **5** and **var_196** less than **2**.
9. Show observations between 20000 and 30000 for 11th column.
10. Create features list including variables from **var_0** to **var_9**.
    a. Find correlations in features.
    b. Sort most correlated variables in features.
    c. Show top 10 most correlated variables in features.
11. Remove outliers from dataset.