

Case Study 04

[Python ilə Data Science]

Çətinlik Dərəcəsi: 3-dən 2

Qeyd: QSS Analytics Tədqiqat və İnkişaf Mərkəzi data və layihənin reallığı barədə iddia etmir. Tapşırıq tədris məqsədi ilə hazırlanmışdır.

1. Load data files
2. List of column names
3. Summary statistics
- DATA CLEANING AND PREPROCESSING
4. Find missing values
5. Impute missing values with mean (numerical variables)
6. Impute missing values with mode (categorical variables)
- PREDICTIVE MODELLING
7. Remove Loan_ID variable – Irrelevant
8. Create target variable
9. Build dummy variables for categorical variables
10. Split train data for cross validation
11. Apply algorithm (Logistic regression, Naive Bayes, SVM, Decision Tree)