

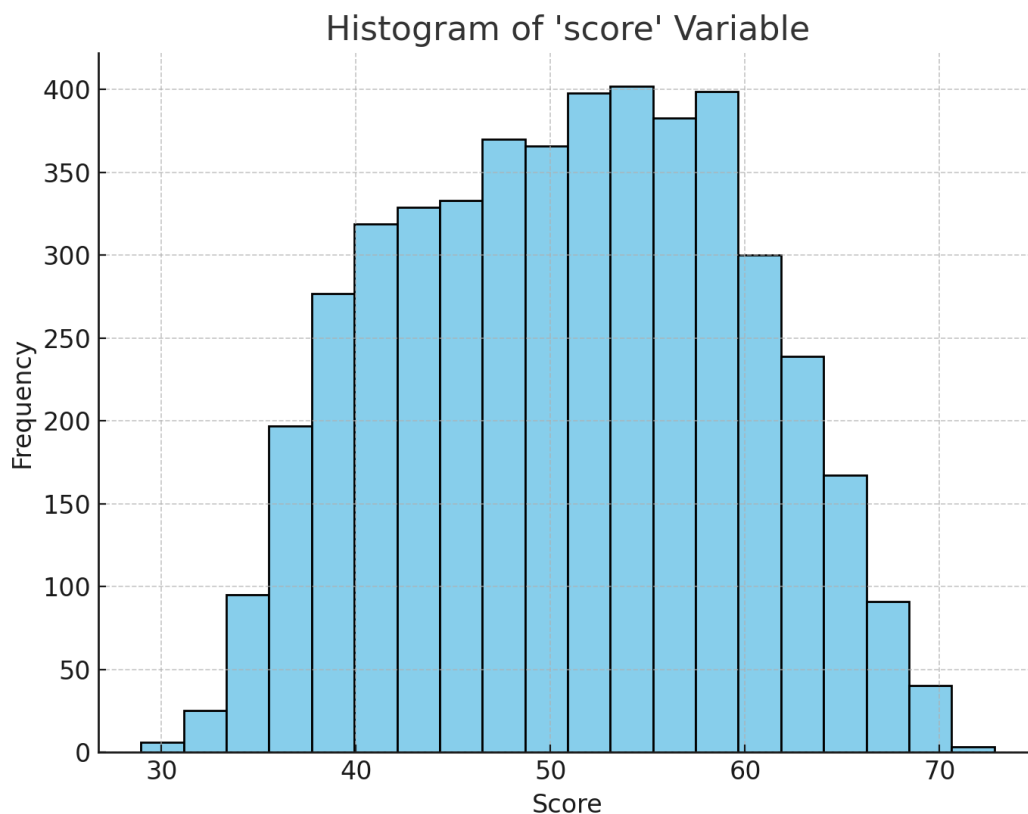
## Lab3 - Analizator Wynikow - Dokumentacja

Cel analizy: Na podstawie dostarczonego zbioru danych, zbudowano model predykcyjny przewidujący zmienną 'score'. Proces obejmował eksplorację danych, inżynierię cech, wybór modelu oraz jego ocenę.

### Eksploracja Danych

Dane zawierały 4739 rekordów i 15 kolumn, z których 8 było zmiennymi kategorycznymi, a pozostałe zmiennymi numerycznymi. Nie wykryto brakujących wartości. Podstawowe statystyki dla zmiennej 'score' to:

Srednia: 50.89, Odchylenie standardowe: 8.70.



### Inżynieria Cech

Przeprowadzono skalowanie zmiennych numerycznych oraz kodowanie zmiennych kategoriycznych.

Dane zostały podzielone na zbiór treningowy i testowy w stosunku 80:20.

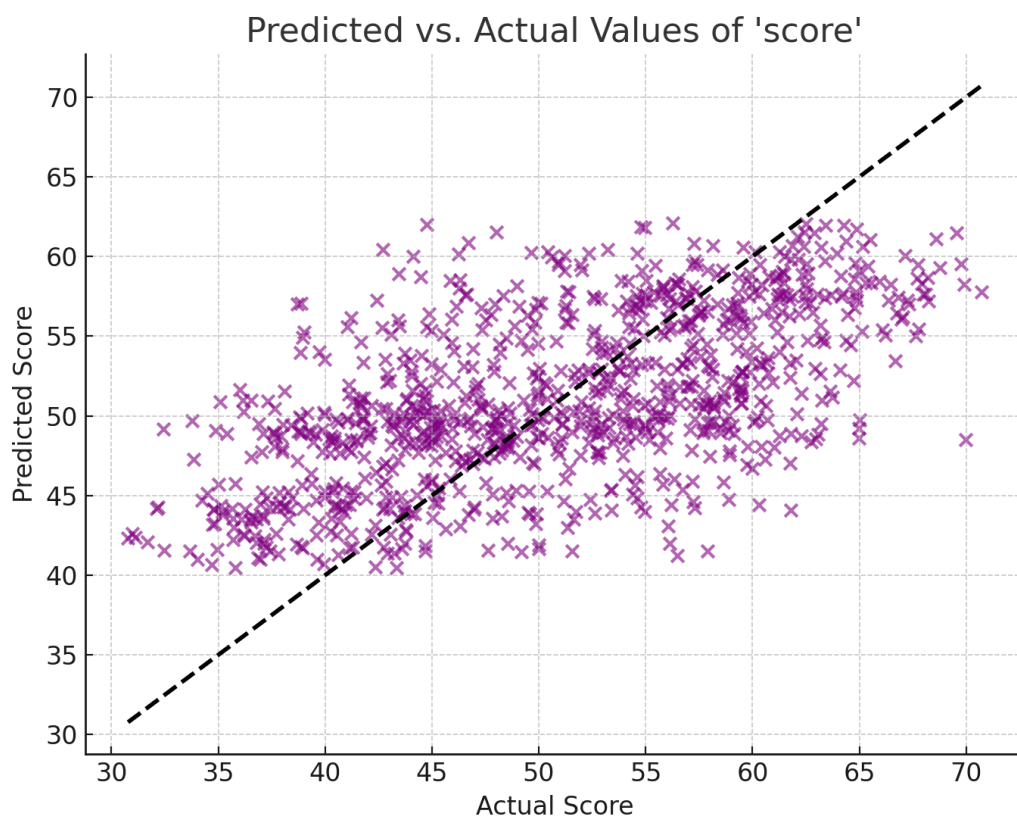
## Modelowanie i Ocena

Model regresji liniowej został wytrenowany na zbiorze treningowym. Wyniki na zbiorze testowym są następujące:

MAE: 5.75

MSE: 49.11

$R^2$ : 0.35.



## Analiza Reszt

Na poniższym wykresie przedstawiono histogram reszt (różnice między wartościami przewidywanymi a rzeczywistymi).

Histogram of Residuals

