*Aleksandra Stanisławka*

*Michał Żyrek*

*Agata Posadzy*

Zespół miał za zadanie przygotować dane ocen wina do analizy.

W celu oczyszczenia danych zostały podjęte następujące kroki:

* oczyszczenie danych, których typ był inny niż numeryczny ze znaków specjalnych (myślniki, znak podkreślenia)
* zmiana danych w kolumnach *citric acid*, *residual sugar*, *chloides* z typu *object* na typ danych numerycznych - *float64*
* zamiana pustych wartości na NaN we wszystkich kolumnach
* sprawdzenie, w których kolumnach występują wartości NaN, a następnie zamiana tych wartości na liczby odpowiadające specyfice danej cechy
* zamiana wartości ujemnych na wartości przeciwne ( absolute value)
* wykluczenie odstających wartości poszczególnych cech, które mogłyby zniekształcić końcowe wyniki
* standaryzacja danych za pomocą użycia StandardScaler