

# **Descriptive Statistics**

In the water\_potability.csv dataset we made a new variable called Drinkable which puts each sample into drinkable or non drinkable groups. this helped us compare the Total Dissolved Solids (TDS) levels between the two types. the dataset has 1998 non drinkable samples and 1278 drinkable ones.

Overall the TDS values range from 320.9 mg/L up to 61227.2 mg/L with a mean around 22014.1 mg/L.

For non-drinkable water TDS goes from min-320.9 to max-61227.2 mg/L with a mean of 21777.5 mg/L.

For drinkable water the range is min-728.8 to max-56488.7 mg/L and the mean is about 22384.0 mg/L.

These descriptive stats give a basic idea of how TDS varies between the drinkable and non drinkable samples.