Water Quality

- **Abstract:**

Safe drinking water is a basic human right and a component of a good health-protection strategy. On a national, regional, and local level, this is a major health and development concern. Investments in water supply and sanitation have been demonstrated to provide a net economic benefit in some locations since the reductions in negative health consequences and health expenses surpass the price of implementing the interventions.
In this project, AKN used historical samples from different resources and restricted standers for water quality based on the chemical, physical, and biological specifications of the sample that were specified as drinkable or not drinkable.

- **Question/Problem statement:**
  What are the chemical components that affect water?

- **Content:**
  Data size (no. rows, and no. of columns) 8000 Rows and 21 Columns.

- Description:

All attributes are numeric variables and they are listed bellow:

- aluminium - dangerous if greater than 2.8
- ammonia - dangerous if greater than 32.5
- arsenic - dangerous if greater than 0.01
- barium - dangerous if greater than 2
- cadmium - dangerous if greater than 0.005
- chloramine - dangerous if greater than 4
- chromium - dangerous if greater than 0.1
- copper - dangerous if greater than 1.3
- flouride - dangerous if greater than 1.5
- bacteria - dangerous if greater than 0
- viruses - dangerous if greater than 0
- lead - dangerous if greater than 0.015

- nitrates - dangerous if greater than 10
- nitrites - dangerous if greater than 1
- mercury - dangerous if greater than 0.002
- perchlorate - dangerous if greater than 56
- radium - dangerous if greater than 5
- selenium - dangerous if greater than 0.5
- silver - dangerous if greater than 0.1
- uranium - dangerous if greater than 0.3
- is_safe - class attribute {0 - not safe, 1 - safe}

- **Tools:**

  - Programs: Jupyter, Spyder, DB browser.
  - Libraries: Sklearn, NuPIC, Pandas, Matplotlib, Seaborn and NumPy.
  - Functions: Group by, Join, count, and others functions in python.
  - Plots: Scatter, Bar graph, forecast, bar chart.

- **MVP Goal:**

  The goal of the project is to classify the water usage (potable or not) based on chemical properties. To be more specific the goal is to classify if the water is potable without misclassifying not safe to potable therefore we used precision as our metric for the project.