

GYM DATASET

Source:

<https://www.kaggle.com/nsrose7224/crowdedness-at-the-campus-gym/downloads/crowdedness-at-the-campus-gym.zip/2>

Dataset description:

| Variable | Description |
|-----------------------------|---|
| number_people | The number of people who were at the gym counted within an hour |
| date | The date when the measurements were taken |
| is_weekend | If weekend = 1, if not = 0 |
| is_holiday | If holiday = 1, if not = 0 |
| is_start_of_semester | If yes = 1, if not = 0 |
| is_during_semester | If yes = 1, if not = 0 |
| month | The month of the year |
| hour | The hour when the calculations were taken |
| celcius | Temperature in Celsius |
| crowded | If number_people <= 29, crowded = 0, otherwise crowded = 1 |
| weekday | 1 = Sunday, 2 = Monday etc. |

The dataset has been modified – temperature variable has been transformed from Fahrenheit to Celsius and has been dropped after. Weekday has been created by taking a function `weekday(date)`, where 1 stands for Sunday. Crowded has been created by using an if-statement. The mean number_people equals around 29. A gym would be considered crowded when a number of visitors exceeds 29 people and crowded =1, otherwise it equals 0.