**Dataset Name:** Absenteeism at Work

**Abstract:**

The dataset contains records of hours of absence of 36 distinct employees at a courier company in Brazil, recorded from July 2007 to July 2010.

**Source:**

Creators, original owner, and donors: Andrea Martiniano (1), Ricardo Pinto Ferreira (2), and Renato Jose Sassi (3).

**Attribute information**

1. Individual Identification (ID): 🡺 Deleted from dataset
2. Reason (Reason for absence): 🡺 Categorical
3. Justified\_Absence 🡺 Categorical (Justified vs unjustified)
4. Month: 🡺 Categorical
5. Day: 🡺 Categorical
6. Season: 🡺 Categorical

Converted to binary variables:

* 1. Summer (0, 1): 🡺 0 = Not summer, 1 = Summer
  2. Fall (0, 1)
  3. Winter (0, 1)
  4. Spring (0, 1)

1. Trans expense (Transportation expense measured in BRL): 🡺 Numerical
2. DFRW (Distance from Residence to work measured in km): 🡺 Numerical
3. ST (Service time, or length of employment measured in years): 🡺 Numerical
4. Age: 🡺 Numerical
5. WLA (Workload Average or how much the employee brings to the company, measured in BRL): 🡺 Numerical
6. Education (Highest level of education attained): 🡺 Categorical

Converted to binary variables

* 1. High school (0, 1): 🡺 0 = False, 1 = True
  2. Grad (0, 1)
  3. Postgrad (0, 1)
  4. Master/PhD (0, 1)

1. Children (Number of Children): 🡺 Numerical
2. Social drinker: 🡺 Binary (0 = False, 1 = True)
3. Social smoker: 🡺 Binary (0 = False, 1 = True)
4. Pet (No. of pets): 🡺 Numerical
5. Weight (measured in kg): 🡺 Numerical
6. Height (measured in m): 🡺 Numerical
7. BMI (Body mass index): 🡺 Numerical
8. AT (Absenteeism time in hours): 🡺 Numerical (dependent variable)