

Applied Intelligent Data and Text Analytics

Absenteeism at Work Dataset

Dataset Information:

The dataset allows for several new combinations of attributes and attribute exclusions, or the modification of the attribute type (categorical, integer, or real) depending on the purpose of the research. The dataset (Absenteeism at Work - Part I) was used in academic research at the Universidade Nove de Julho - Postgraduate Program in Informatics and Knowledge Management.

Link:

<https://archive.ics.uci.edu/ml/datasets/Absenteeism+at+work>

Attribute Information:

1. Individual identification (ID)
2. Reason for absence (ICD).

Absences attested by the International Code of Diseases (ICD) stratified into 21 categories (I to XXI) as follows:

- I. Certain infectious and parasitic diseases
- II. Neoplasms
- III. Diseases of the blood and blood-forming organs and certain disorders involving the immune mechanism
- IV. Endocrine, nutritional, and metabolic diseases
- V. Mental and behavioural disorders
- VI. Diseases of the nervous system
- VII. Diseases of the eye and adnexa
- VIII. Diseases of the ear and mastoid process
- IX. Diseases of the circulatory system
- X. Diseases of the respiratory system
- XI. Diseases of the digestive system
- XII. Diseases of the skin and subcutaneous tissue
- XIII. Diseases of the musculoskeletal system and connective tissue
- XIV. Diseases of the genitourinary system
- XV. Pregnancy, childbirth, and the puerperium
- XVI. Certain conditions originating in the perinatal period
- XVII. Congenital malformations, deformations, and chromosomal abnormalities
- XVIII. Symptoms, signs, and abnormal clinical and laboratory findings, not elsewhere classified
- XIX. Injury, poisoning and certain other consequences of external causes

XX. External causes of morbidity and mortality

XXI. Factors influencing health status and contact with health services.

And 7 categories without (ICD) patient follow-up (22), medical consultation (23), blood donation (24), laboratory examination (25), unjustified absence (26), physiotherapy (27), dental consultation (28).

3. Month of absence

4. Day of the week (Monday (2), Tuesday (3), Wednesday (4), Thursday (5), Friday (6))

5. Seasons (summer (1), autumn (2), winter (3), spring (4))

6. Transportation expense

7. Distance from Residence to Work (kilometres)

8. Service time

9. Age

10. Workload Average/day

11. Hit target

12. Disciplinary failure (yes=1; no=0)

13. Education (high school (1), graduate (2), postgraduate (3), master and doctor (4))

14. Son (number of children)

15. Social drinker (yes=1; no=0)

16. Social smoker (yes=1; no=0)

17. Pet (number of pet)

18. Weight

19. Height

20. Body mass index

21. Absenteeism time in hours (target)