UCI Machine Learning Repository

Center for Machine Learning and Intelligent Systems

Search

● Repository   ○ Web

Google™

View ALL Data Sets

STUDENT ALCOHOL CONSUMPTION Data Set

Download: Data Folder, Data Set Description

Abstract: The result also provides the correlation between alcohol usage and the social, gender and study time attributes for each student.

Data Set Characteristics:	Multivariate	Number of Instances:	1044	Area:	Social
Attribute Characteristics:	Integer	**Number of Attributes:**	32	**Date Donated**	2016-03-03
Associated Tasks:	Classification	**Missing Values?**	N/A	**Number of Web Hits:**	11994

Source:

Fabio Pagnotta, Hossain Mohammad Amran.
Email:fabio.pagnotta@studenti.unicam.it, mohammadamra.hossain '@' studenti.unicam.it
University Of Camerino

Data Set Information:

P. Cortez and A. Silva. Using Data Mining to Predict Secondary School Student Performance. In A. Brito and J. Teixeira Eds., Proceedings of 5th FUture BUsiness TEChnology Conference (FUBUTEC 2008) pp. 5-12, Porto, Portugal, April, 2008, EUROSIS, ISBN 978-9077381-39-7.

Attribute Information:

Attributes for both student-mat.csv (Math course) and student-por.csv (Portuguese language course) datasets:
1 school - student's school (binary: 'GP' - Gabriel Pereira or 'MS' - Mousinho da Silveira)
2 sex - student's sex (binary: 'F' - female or 'M' - male)
3 age - student's age (numeric: from 15 to 22)
4 address - student's home address type (binary: 'U' - urban or 'R' - rural)
5 famsize - family size (binary: 'LE3' - less or equal to 3 or 'GT3' - greater than 3)
6 Pstatus - parent's cohabitation status (binary: 'T' - living together or 'A' - apart)
7 Medu - mother's education (numeric: 0 - none, 1 - primary education (4th grade), 2 Ã¢â¬â€œ 5th to 9th grade, 3

Ã¢â¬â€œ secondary education or 4 Ã¢â¬â€œ higher education)
8 Fedu - father's education (numeric: 0 - none, 1 - primary education (4th grade), 2 Ã¢â¬â€œ 5th to 9th grade, 3 Ã¢â¬â€œ secondary education or 4 Ã¢â¬â€œ higher education)
9 Mjob - mother's job (nominal: 'teacher', 'health' care related, civil 'services' (e.g. administrative or police), 'at_home' or 'other')
10 Fjob - father's job (nominal: 'teacher', 'health' care related, civil 'services' (e.g. administrative or police), 'at_home' or 'other')
11 reason - reason to choose this school (nominal: close to 'home', school 'reputation', 'course' preference or 'other')
12 guardian - student's guardian (nominal: 'mother', 'father' or 'other')
13 traveltime - home to school travel time (numeric: 1 - <15 min., 2 - 15 to 30 min., 3 - 30 min. to 1 hour, or 4 - >1 hour)
14 studytime - weekly study time (numeric: 1 - <2 hours, 2 - 2 to 5 hours, 3 - 5 to 10 hours, or 4 - >10 hours)
15 failures - number of past class failures (numeric: n if 1<=n<3, else 4)
16 schoolsup - extra educational support (binary: yes or no)
17 famsup - family educational support (binary: yes or no)
18 paid - extra paid classes within the course subject (Math or Portuguese) (binary: yes or no)
19 activities - extra-curricular activities (binary: yes or no)
20 nursery - attended nursery school (binary: yes or no)
21 higher - wants to take higher education (binary: yes or no)
22 internet - Internet access at home (binary: yes or no)
23 romantic - with a romantic relationship (binary: yes or no)
24 famrel - quality of family relationships (numeric: from 1 - very bad to 5 - excellent)
25 freetime - free time after school (numeric: from 1 - very low to 5 - very high)
26 goout - going out with friends (numeric: from 1 - very low to 5 - very high)
27 Dalc - workday alcohol consumption (numeric: from 1 - very low to 5 - very high)
28 Walc - weekend alcohol consumption (numeric: from 1 - very low to 5 - very high)
29 health - current health status (numeric: from 1 - very bad to 5 - very good)
30 absences - number of school absences (numeric: from 0 to 93)

these grades are related with the course subject, Math or Portuguese:
31 G1 - first period grade (numeric: from 0 to 20)
31 G2 - second period grade (numeric: from 0 to 20)
32 G3 - final grade (numeric: from 0 to 20, output target)

Relevant Papers:

Using Data Mining To Predict Secondary School Student Alcohol Consumption
Fabio Pagnotta, Hossain Mohammad Amran,
Department of Computer Science, University of Camerino
Available at: [Web Link]

Citation Request:

Please include this citation if you plan to use this database:
Using Data Mining To Predict Secondary School Student Alcohol Consumption.
Fabio Pagnotta, Hossain Mohammad Amran
Department of Computer Science,University of Camerino

Supported By: In Collaboration With:

About || Citation Policy || Donation Policy || Contact || CML