

# Data Sets Description

## ***bostonhousing.csv***

### ***Concerns housing values in suburbs of Boston***

crim: per capita crime rate by town  
zn: proportion of residential land zoned for lots over 25,000 sq.ft.  
indus: proportion of non-retail business acres per town  
chas: charles river dummy variable (= 1 if tract bounds river; 0 otherwise)  
nox: nitric oxides concentration (parts per 10 million)  
rm: average number of rooms per dwelling  
age: proportion of owner-occupied units built prior to 1940  
dis: distances to five boston employment centres  
rad: index of accessibility to radial highways  
tax: full-value property-tax rate per \$10,000  
ptratio: pupil-teacher ratio by town  
lstat: % lower status of the population  
medv: median value of owner-occupied homes in \$1000's

## ***cpuperform.csv***

### ***Relative CPU Performance Data***

vendor: vendor name  
myct: machine cycle time in nanoseconds (integer)  
mmin: minimum main memory in kilobytes (integer)  
mmax: maximum main memory in kilobytes (integer)  
cach: cache memory in kilobytes (integer)  
chmin: minimum channels in units (integer)  
chmax: maximum channels in units (integer)  
prp: published relative performance (integer)  
erp: estimated relative performance from the original article (integer)

## ***credit.csv***

### ***Concerns bank customers data***

age: customer's age in years

income: customers' income (1 – low, 2 – medium, 3 – high)

cards: number of credit cards

education: education level (1 – lower education, 2 – higher education)

carloans: whether the customer has car loans (1 – yes, 2 – no)

rating: credit score (0 – bad credit, 1 – good credit)

## ***credit\_card.csv***

### ***Information about credit card transactions***

gender: customers' gender (0 – male, 1 – female)

card: primary credit card (1 – American Express, 2 – Visa, 3 – Mastercard, 4 – Discover, 5 – other)

type: type of transaction (1 – grocery, 2 – retail, 3 – entertainment, 4 – travel, 5 – other)

items: number of items bought

spent: amount spent

## ***directmail.csv***

### ***Direct mail company database***

previous: did they respond to the previous messages?

age: customers' age

income: income category

education: education level

reside: for how long they reside at their current address

gender: customers' gender

children: number of children

region: the region where they live

## ***education.csv***

### ***Education expenditure in 50 US states***

state: state name

region: region (1 – Northeastern, 2 – North Central, 3 – Southern, 4 – Western)

urban: number of residents per thousand residing in urban areas

income: per capita personal income

teen: number of residents per thousand under 18 years of age

expend: per capita expenditure on public education in a state

### ***forestfires.csv***

#### ***Forest fires data***

temp: temperature in Celsius degrees

RH: relative humidity in %

wind: wind speed in km/h

rain: outside rain in mm/m2

fire: whether a fire took place (0 – no, 1 – yes)

### ***housedata.csv***

#### ***House data***

(Only the variables of interest are listed here)

price: house price in dollars

bedrooms: number of bedrooms

bathrooms: number of bathrooms

sqft\_living: living area surface in square feet

sqft\_lot: land surface in square feet

floors: number of floors

grade: slope around the house foundation

sqft\_basement: basement surface in square feet

old: house age in years

## ***ulcer\_recurrence.csv***

### ***Concerns data about ulcer patients***

age: patient's age

duration: duration of disease (1 – less than 5 years, 2 – equal to or greater than 5 years)

visit: time of last visit in months

result: whether the ulcer is present (0 – no ulcer, 1 – ulcer)

## ***vehicles.csv***

### ***Concerns technical data of vehicles***

engine: engine displacement

horse: horsepower

weight: vehicle weight

length: vehicle length

fuelcap: fuel capacity

type: vehicle type (0 – automobile, 1 – truck)

## ***winequality.csv***

### ***Concerns data about wines***

type: wine type (red or white)

fixed acidity

volatile acidity

citric acid

residual sugar

chlorides

free sulfur dioxide

total sulfur dioxide

density

pH

sulfates

alcohol: content of alcohol (%)

quality: wine quality (on a scale from 1 to 7)

## youngpeoplesurvey.csv

Healthyeating	I live a very healthy lifestyle.
Dailyevents	I take notice of what goes on around me.
Prioritisingworkload	I try to do tasks as soon as possible and not leave them until last minute.
Writingnotes	I always make a list so I don't forget anything.
Workaholism	I often study or work even in my spare time.
Thinkingahead	I look at things from all different angles before I go ahead.
Finaljudgement	I believe that bad people will suffer one day and good people will be rewarded.
Reliability	I am reliable at work and always complete all tasks given to me.
Keepingpromises	I always keep my promises.
Lossofinterest	I can fall for someone very quickly and then completely lose interest.
Friendsversusmoney	I would rather have lots of friends than lots of money.
Funniness	I always try to be the funniest one.
Fake	I can be two faced sometimes.
Criminaldamage	I damaged things in the past when angry.
Decisionmaking	I take my time to make decisions.
Elections	I always try to vote in elections.
Selfcriticism	I often think about and regret the decisions I make.
Judgmentcalls	I can tell if people listen to me or not when I talk to them.
Hypochondria	I am a hypochondriac.
Empathy	I am empathetic person.
Eatingtosurvive	I eat because I have to. I don't enjoy food and eat as fast as I can.
Giving	I try to give as much as I can to other people at Christmas.
Compassiontoanimals	I don't like seeing animals suffering.
Borrowedstuff	I look after things I have borrowed from others.
Loneliness	I feel lonely in life.
Cheatinginschool	I used to cheat at school.
Health	I worry about my health.
Changingthepast	I wish I could change the past because of the things I have done.
God	I believe in God.
Dreams	I always have good dreams.
Charity	I always give to charity.
Gender	Respondent's gender
Onlychild	I am the only child
Villagetown	I spent most of my childhood in a...