|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| |  |  | | --- | --- | | **Flags Data Set**  *Download*: [Data Folder](http://archive.ics.uci.edu/ml/machine-learning-databases/flags/), [Data Set Description](http://archive.ics.uci.edu/ml/machine-learning-databases/flags/flag.names)  **Abstract**: From Collins Gem Guide to Flags, 1986 | http://archive.ics.uci.edu/ml/assets/MLimages/Large40.jpg |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | | **Data Set Characteristics:** | Multivariate | **Number of Instances:** | 194 | **Area:** | N/A | | **Attribute Characteristics:** | Categorical, Integer | **Number of Attributes:** | 30 | **Date Donated** | 1990-05-15 | | **Associated Tasks:** | Classification | **Missing Values?** | No | **Number of Web Hits:** | 79868 | |

**Attribute Information:**

1. name: Name of the country concerned   
2. landmass: 1=N.America, 2=S.America, 3=Europe, 4=Africa, 4=Asia, 6=Oceania   
3. zone: Geographic quadrant, based on Greenwich and the Equator; 1=NE, 2=SE, 3=SW, 4=NW   
4. area: in thousands of square km   
5. population: in round millions   
6. language: 1=English, 2=Spanish, 3=French, 4=German, 5=Slavic, 6=Other Indo-European, 7=Chinese, 8=Arabic, 9=Japanese/Turkish/Finnish/Magyar, 10=Others   
7. religion: 0=Catholic, 1=Other Christian, 2=Muslim, 3=Buddhist, 4=Hindu, 5=Ethnic, 6=Marxist, 7=Others   
8. bars: Number of vertical bars in the flag   
9. stripes: Number of horizontal stripes in the flag   
10. colours: Number of different colours in the flag   
11. red: 0 if red absent, 1 if red present in the flag   
12. green: same for green   
13. blue: same for blue   
14. gold: same for gold (also yellow)   
15. white: same for white   
16. black: same for black   
17. orange: same for orange (also brown)   
18. mainhue: predominant colour in the flag (tie-breaks decided by taking the topmost hue, if that fails then the most central hue, and if that fails the leftmost hue)   
19. circles: Number of circles in the flag   
20. crosses: Number of (upright) crosses   
21. saltires: Number of diagonal crosses   
22. quarters: Number of quartered sections   
23. sunstars: Number of sun or star symbols   
24. crescent: 1 if a crescent moon symbol present, else 0   
25. triangle: 1 if any triangles present, 0 otherwise

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| |  |  | | --- | --- | | **Mushroom Data Set**  *Download*: [Data Folder](http://archive.ics.uci.edu/ml/machine-learning-databases/mushroom/), [Data Set Description](http://archive.ics.uci.edu/ml/machine-learning-databases/mushroom/agaricus-lepiota.names)  **Abstract**: From Audobon Society Field Guide; mushrooms described in terms of physical characteristics; classification: poisonous or edible | http://archive.ics.uci.edu/ml/assets/MLimages/Large73.jpg |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | | **Data Set Characteristics:** | Multivariate | **Number of Instances:** | 8124 | **Area:** | Life | | **Attribute Characteristics:** | Categorical | **Number of Attributes:** | 22 | **Date Donated** | 1987-04-27 | | **Associated Tasks:** | Classification | **Missing Values?** | Yes | **Number of Web Hits:** | 143732 | |

**Attribute Information:**

1. cap-shape: bell=b,conical=c,convex=x,flat=f, knobbed=k,sunken=s   
2. cap-surface: fibrous=f,grooves=g,scaly=y,smooth=s   
3. cap-color: brown=n,buff=b,cinnamon=c,gray=g,green=r, pink=p,purple=u,red=e,white=w,yellow=y   
4. bruises?: bruises=t,no=f   
5. odor: almond=a,anise=l,creosote=c,fishy=y,foul=f, musty=m,none=n,pungent=p,spicy=s   
6. gill-attachment: attached=a,descending=d,free=f,notched=n   
7. gill-spacing: close=c,crowded=w,distant=d   
8. gill-size: broad=b,narrow=n   
9. gill-color: black=k,brown=n,buff=b,chocolate=h,gray=g, green=r,orange=o,pink=p,purple=u,red=e, white=w,yellow=y   
10. stalk-shape: enlarging=e,tapering=t   
11. stalk-root: bulbous=b,club=c,cup=u,equal=e, rhizomorphs=z,rooted=r,missing=?   
12. stalk-surface-above-ring: fibrous=f,scaly=y,silky=k,smooth=s   
13. stalk-surface-below-ring: fibrous=f,scaly=y,silky=k,smooth=s   
14. stalk-color-above-ring: brown=n,buff=b,cinnamon=c,gray=g,orange=o, pink=p,red=e,white=w,yellow=y   
15. stalk-color-below-ring: brown=n,buff=b,cinnamon=c,gray=g,orange=o, pink=p,red=e,white=w,yellow=y   
16. veil-type: partial=p,universal=u   
17. veil-color: brown=n,orange=o,white=w,yellow=y   
18. ring-number: none=n,one=o,two=t   
19. ring-type: cobwebby=c,evanescent=e,flaring=f,large=l, none=n,pendant=p,sheathing=s,zone=z   
20. spore-print-color: black=k,brown=n,buff=b,chocolate=h,green=r, orange=o,purple=u,white=w,yellow=y   
21. population: abundant=a,clustered=c,numerous=n, scattered=s,several=v,solitary=y   
22. habitat: grasses=g,leaves=l,meadows=m,paths=p, urban=u,waste=w,woods=d

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| |  |  | | --- | --- | | **Census Income Data Set**  *Download*: [Data Folder](http://archive.ics.uci.edu/ml/machine-learning-databases/adult/), [Data Set Description](http://archive.ics.uci.edu/ml/machine-learning-databases/adult/adult.names)  **Abstract**: Predict whether income exceeds $50K/yr based on census data. Also known as "Adult" dataset. | http://archive.ics.uci.edu/ml/assets/MLimages/Large2.jpg |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | | **Data Set Characteristics:** | Multivariate | **Number of Instances:** | 48842 | **Area:** | Social | | **Attribute Characteristics:** | Categorical, Integer | **Number of Attributes:** | 14 | **Date Donated** | 1996-05-01 | | **Associated Tasks:** | Classification | **Missing Values?** | Yes | **Number of Web Hits:** | 96964 | |

age: continuous.   
workclass: Private, Self-emp-not-inc, Self-emp-inc, Federal-gov, Local-gov, State-gov, Without-pay, Never-worked.   
fnlwgt: continuous.   
education: Bachelors, Some-college, 11th, HS-grad, Prof-school, Assoc-acdm, Assoc-voc, 9th, 7th-8th, 12th, Masters, 1st-4th, 10th, Doctorate, 5th-6th, Preschool.   
education-num: continuous.   
marital-status: Married-civ-spouse, Divorced, Never-married, Separated, Widowed, Married-spouse-absent, Married-AF-spouse.   
occupation: Tech-support, Craft-repair, Other-service, Sales, Exec-managerial, Prof-specialty, Handlers-cleaners, Machine-op-inspct, Adm-clerical, Farming-fishing, Transport-moving, Priv-house-serv, Protective-serv, Armed-Forces.  
relationship: Wife, Own-child, Husband, Not-in-family, Other-relative, Unmarried.   
race: White, Asian-Pac-Islander, Amer-Indian-Eskimo, Other, Black.   
sex: Female, Male.   
capital-gain: continuous.   
capital-loss: continuous.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| |  |  | | --- | --- | | **Automobile Data Set**  *Download*: [Data Folder](http://archive.ics.uci.edu/ml/machine-learning-databases/autos/), [Data Set Description](http://archive.ics.uci.edu/ml/machine-learning-databases/autos/imports-85.names)  **Abstract**: From 1985 Ward's Automotive Yearbook | http://archive.ics.uci.edu/ml/assets/MLimages/Large10.jpg |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | | **Data Set Characteristics:** | Multivariate | **Number of Instances:** | 205 | **Area:** | N/A | | **Attribute Characteristics:** | Categorical, Integer, Real | **Number of Attributes:** | 26 | **Date Donated** | 1987-05-19 | | **Associated Tasks:** | Regression | **Missing Values?** | Yes | **Number of Web Hits:** | 108391 |   1. symboling: -3, -2, -1, 0, 1, 2, 3.  2. normalized-losses: continuous from 65 to 256.  3. make:  alfa-romero, audi, bmw, chevrolet, dodge, honda,  isuzu, jaguar, mazda, mercedes-benz, mercury,  mitsubishi, nissan, peugot, plymouth, porsche,  renault, saab, subaru, toyota, volkswagen, volvo   4. fuel-type: diesel, gas.  5. aspiration: std, turbo.  6. num-of-doors: four, two.  7. body-style: hardtop, wagon, sedan, hatchback, convertible.  8. drive-wheels: 4wd, fwd, rwd. |
|  |