

Dataset Information

This is a list of what all of the variables are and where the dataset is from.

| | |
|---------------------------------|---|
| Babies | 1 |
| Ball Wear | 2 |
| Cars | 2 |
| Diamonds | 2 |
| Empty Dataset for Editing | 3 |
| Kiwi | 3 |
| Marathon | 3 |
| Rugby | 4 |
| Sharks | 4 |
| Sports Science | 4 |
| TS Note | 5 |
| TS - Births and Deaths | 5 |
| TS - Forestry | 5 |
| TS - Imports | 5 |
| TS - Jobs | 5 |
| TS - Penguin | 6 |
| TS - Sea Ice | 7 |
| TS - Sunglasses | 7 |
| TS - Visitors | 7 |

Babies

The data on 189 births were collected at Baystate Medical Center, Springfield, Mass. during 1986. The goal of this study was to identify risk factors associated with giving birth to a low birth weight baby (weighing less than 2500 grams). Data was collected on 189 women, 59 of which had low birth weight babies and 130 of which had normal birth weight babies.

| Variable | Description |
|----------------|--|
| LowBirthWeight | No = Birth Weight \geq 2500g Yes = Birth Weight $<$ 2500g |
| MothersAge | Age of the Mother in Years |
| Race | Race of the mother |
| MotherSmoke | Smoking Status During Pregnancy |
| FTV | Number of Physician Visits During the First Trimester |
| BirthWeight | Birth Weight in Grams |

Ball Wear

Data was recorded of students going to the school ball in 2012 as to how much they spent on their clothing and accessories.

| Variable | Description |
|--------------|--|
| Gender | Boy = new student is male Girl = new student is female |
| Amount.spent | The amount spent on clothing and accessories in New Zealand Dollars. |

Cars

With rising costs of owning and running a car, and environmental awareness, buyers are becoming more conscious of the features when purchasing new cars. The data supplied is for new vehicles sold in America in 1993.

| Variable | Description |
|---------------------|--|
| Vehicle Name | |
| Origin | Country of manufacture · America · Foreign |
| Price | US \$1000 |
| Type | Small, midsize, large, compact, sporty, van |
| City | Fuel efficiency in kilometres per litre in cities and on motorways |
| OpenRoad | Fuel efficiency in kilometres per litre on country and open roads |
| Drive Train | Front Wheel Drive Rear Wheel Drive |
| Engine Size | Size in litres |
| Manual Transmission | Yes No |
| Weight | Weight of car in Kg |

Diamonds

Every diamond is unique, and there are a variety of factors which affect the price of a diamond. Insurance companies in particular are concerned that stones are valued correctly. Data on 308 round diamond stones was collected from a Singapore based retailer of diamond jewellery, who had the stones valued.

| Variable | Description |
|----------|--|
| Carat | Weight of diamond stones in carat units 1 carat = 0.2 grams |
| Colour | Numerical value given for quality of colour ranging from 1=colourless to 6=near colourless |
| Clarity | Average = score 1, 2 or 3 Above average = score 4, 5 or 6 |
| Lab | Laboratory that tested & valued the diamond 1 = laboratory 1 2 = laboratory 2 |
| Price | Price in US dollars |

Empty Dataset for Editing

This is a blank dataset designed for entering experimental data.

Kiwi

A sample of kiwi birds around New Zealand was collected in order to help with conservation efforts. The original data is from: <http://www.kiwisforkiwi.org/> and was sourced from the secondary school guides (<http://seniorsecondary.tki.org.nz/Mathematics-and-statistics/Achievement-objectives/AOs-by-level/AO-S7-1>)

| Variable | Description |
|------------|--|
| Species | GS-Great Spotted NIBr-NorthIsland Brown Tok-Southern Tokoeka |
| Gender | M-Male F-Female |
| Weight(kg) | The weight of the kiwi bird in kg |
| Height(cm) | The height of the kiwi bird in cm |
| Location | NWN-North West Nelson SF-South Fiordland CW-Central Westland N-Northland EC-Eastern Canterbury E-East North Island StI-Stewart Island W-West North Island NF-North Fiordland |

Teachers note: this is a synthesised dataset based on real data. At the time of creating the data set there were around 25,000 brown, 17,000 great spotted and 34,500 southern tokoeka. These numbers formed the basis of the data set, but instead of being out of around 76,000 the data set contains around 700 birds.

The data was generated using the population parameters, including gender, location, height, weight and species in Fathom. The size of the population was so that it was too big to use all the data (when doing by hand) but not too big that it couldn't be created for students to use as a "population" to sample from.

Marathon

The data is a sample taken from marathons in NZ.
It is a simple random sample of 200 athletes.

| Variable | Description |
|----------------|--|
| Minutes | How many minutes they completed the marathon in |
| Gender | Male (M) or Female (F) |
| AgeGroup | Younger (under 40) or older (over 40) |
| StridelengthCM | The persons average stride length over the marathon in cm. |

Rugby

The data is real data and comes from <http://www.rugby-sidestep-central.com/>

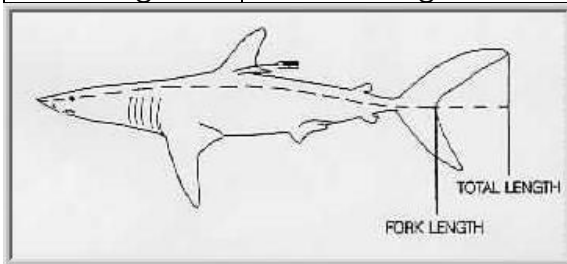
| Variable | Description |
|----------|--|
| Country | New Zealand or South Africa |
| Position | Forward or Back |
| Weight | The weight of the player in kilograms (kg) |
| Height | The height of the player in metres (m) |

Sharks

The data is real and comes from the MPI centralised observer database:

<http://www.fish.govt.nz/mi-nz/Research+Services/Research+Database+Documentation/Cod/default.htm>

| Variable | Description |
|---------------|-------------------------------------|
| Calendar Year | Year which the data was recorded in |
| Fish Sex | The gender of the shark |
| Total Length | The total length of the shark |
| Fork Length | The fork length of the shark |



Sports Science

The data is real data and comes from <http://www.statsci.org/data/oz/ais.html>

The data set provides information about 102 male athletes and 100 female athletes at the Australian Institute of Sport.

| Variable | Description |
|----------|---|
| Sex | male or female |
| Sport | sport played |
| Ht | height in cm |
| Wt | weight in kg |
| LBM | lean body mass in kg |
| %Bfat | % body fat |
| BMI | body mass index (weight/height ²) |
| RCC | red blood cell count |
| WCC | white blood cell count |
| Hc | haematocrit |
| Hg | haemoglobin |
| Ferr | plasma ferritin concentration |
| SSF | sum of skin folds |

Teachers note: this is the dataset used in the TKI Exemplar A

TS Note

All datasets prefixed with a TS that are preloaded in NZGrapher are time series datasets, and not particularly well suited to bivariate or multivariate analysis.

TS - Births and Deaths

Data on the number of births and deaths in New Zealand.
The data is sourced from Statistics New Zealand.

| Variable | Description |
|--------------------|--|
| Quarter | Quarterly |
| Male Live Births | Number of males born during the quarter |
| Female Live Births | Number of females born during the quarter |
| Male Deaths | Number of male deaths during the quarter |
| Female Deaths | Number of female deaths during the quarter |

TS - Forestry

The volume of wood removed from different types of forests in New Zealand.
The data is sourced from the Ministry for Primary Industries.

| Variable | Description |
|--------------------|--|
| Quarter | Quarterly |
| Natural Forests | The volume of wood removed from Natural Forests in millions of m ³ |
| Plantation Forests | The volume of wood removed from Plantation Forests in millions of m ³ |

TS - Imports

Information on imports to and from New Zealand.
The data is sourced from Statistics New Zealand.

| Variable | Description |
|-----------------------|--|
| Month | Monthly |
| TotalAirportsCIF | Cost, insurance and freight of imported goods in NZ\$(000) |
| TotalParcelPostCIF | Cost, insurance and freight of imported goods in NZ\$(000) |
| TotalSeaportsCIF | Cost, insurance and freight of imported goods in NZ\$(000) |
| TotalAirportsWeight | Weight of imported goods in tonnes |
| TotalParcelPostWeight | Weight of imported goods in tonnes |
| TotalSeaportsWeight | Weight of imported goods in tonnes |

TS - Jobs

The number of people in employment in New Zealand.
The data is sourced from Statistics New Zealand.

| Variable | Description |
|-------------------|--|
| Month | Monthly |
| Total Filled Jobs | The number of jobs that are filled in millions |
| NZ Population | The population of New Zealand in millions |

TS - Penguin

Data on the number of penguins at the Phillip Island Penguin Parade in Australia.

| Variable | Description |
|----------|--------------------------------------|
| Month | Monthly |
| Number | The number of penguins in the colony |

TS - Sea Ice

The data is the surface area of sea ice in millions of square kilometres.
The data is sourced from the National Snow and Ice Data Center.

| Variable | Description |
|------------|--|
| Time | Monthly |
| Arctic | Million Square Kilometres of Ice in the Arctic |
| Antarctica | Million Square Kilometres of Ice in Antarctica |

TS - Sunglasses

Data on the value of sunglasses sold.

| Variable | Description |
|----------|---|
| Quarter | Quarterly |
| Sales | Amount of sales in thousands of dollars |

TS - Visitors

The visitors' dataset is the number of people entering New Zealand on a Visitor Visa from Australia, China, Japan and the UK.
The data is sourced from Statistics New Zealand.

| Variable | Description |
|-----------------------------|--|
| Date | Quarterly |
| Australia | Number of visitors in the quarter from Australia |
| China, People's Republic of | Number of visitors in the quarter from China |
| Japan | Number of visitors in the quarter from Japan |
| United Kingdom | Number of visitors in the quarter from the UK |