

Chapter 0: Description of the cases

Table of Contents

1. The Decathlon Dataset	1
2. The temperature data set	2

1. The Decathlon Dataset

This dataset contains the results of decathlon events during two athletic meetings which took place one month apart in 2004: the Olympic Games in Athens which took place on August 23 and 24; and the Decastar 2004 which took place on 25 and 26 September.

For both competitions, the following information is available for each athlete: performance for each of the 10 events, total number of points (for each event, an athlete earns points based on performance; here the sum of points scored) and the final ranking.

The 10 events took place in the following order: 100 metres, long jump, shot put, high jump, 400 metres and 110 metre hurdles, discus, pole vault, javelin, 1500 metres.

The variable Competition indicates whether the results are from the Decastar competition or the Olympic Games.

	row.names	100m	Long.jump	Shot.put	High.jump	400m	110m.hurdle	Discus	Pole.vault	Javeline	1500m	Rank	Points	Competition
1	SEBRLE	11.04	7.58	14.83	2.07	49.81	14.69	43.75	5.02	63.19	291.70	1	8217	Decastar
2	CLAY	10.76	7.40	14.26	1.86	49.37	14.05	50.72	4.92	60.15	301.50	2	8122	Decastar
3	KARPOV	11.02	7.30	14.77	2.04	48.37	14.09	48.95	4.92	50.31	300.20	3	8099	Decastar
4	BERNARD	11.02	7.23	14.25	1.92	48.93	14.99	40.87	5.32	62.77	280.10	4	8067	Decastar
5	YURKOV	11.34	7.09	15.19	2.10	50.42	15.31	46.26	4.72	63.44	276.40	5	8036	Decastar
6	WARNERS	11.11	7.60	14.31	1.98	48.68	14.23	41.10	4.92	51.77	278.10	6	8030	Decastar
7	ZSIVOCZKY	11.13	7.30	13.48	2.01	48.62	14.17	45.67	4.42	55.37	268.00	7	8004	Decastar
8	McMULLEN	10.83	7.31	13.76	2.13	49.91	14.38	44.41	4.42	56.37	285.10	8	7995	Decastar
9	MARTINEAU	11.64	6.81	14.57	1.95	50.14	14.93	47.60	4.92	52.33	262.10	9	7802	Decastar
10	HERNU	11.37	7.56	14.41	1.86	51.10	15.06	44.99	4.82	57.19	285.10	10	7733	Decastar

The dataset is available in the package FactoMineR:

```
install.packages("FactoMineR")
library(FactoMineR)
data(decathlon)
```

Remark:

13 athletes performed in both competitions.

2. The temperature data set

In this example, we investigate the climates of different European countries. To do so, temperatures (in Celsius) were collected monthly for the main European capitals and other major cities. In addition to the monthly temperatures, the average annual temperature and the thermal amplitude (difference between the maximum monthly average and the minimum monthly average of a city) were recorded for each city.

We also included two quantitative positioning variables (latitude and longitude) as well as two categorical variables: Area (with values north, south, east and west of Europe) and warm (1: if city is considered as a warm city, ; 0 if the city is considered to be rather cold).

row.names	January	February	March	April	May	June	July	August	September	October	November	December	Annual	Amplitude	Latitude	Longitude	Area
Amsterdam	2.9	2.5	5.7	8.2	12.5	14.8	17.1	17.1	14.5	11.4	7.0	4.4	9.9	14.6	52.2	4.5	West
Athens	9.1	9.7	11.7	15.4	20.1	24.5	27.4	27.2	23.8	19.2	14.6	11.0	17.8	18.3	37.6	23.5	South
Berlin	-0.2	0.1	4.4	8.2	13.8	16.0	18.3	18.0	14.4	10.0	4.2	1.2	9.1	18.5	52.3	13.2	West
Brussels	3.3	3.3	6.7	8.9	12.8	15.6	17.8	17.8	15.0	11.1	6.7	4.4	10.3	14.4	50.5	4.2	West
Budapest	-1.1	0.8	5.5	11.6	17.0	20.2	22.0	21.3	16.9	11.3	5.1	0.7	10.9	23.1	47.3	19.0	East
Copenhagen	-0.4	-0.4	1.3	5.8	11.1	15.4	17.1	16.6	13.3	8.8	4.1	1.3	7.8	17.5	55.4	12.3	North
Dublin	4.8	5.0	5.9	7.8	10.4	13.3	15.0	14.6	12.7	9.7	6.7	5.4	9.3	10.2	53.2	6.1	North
Elsinki	-5.8	-6.2	-2.7	3.1	10.2	14.0	17.2	14.9	9.7	5.2	0.1	-2.3	4.8	23.4	60.1	25.0	North

The dataset is available from the temp_warm.txt file.

```
# import temp_warm.txt
temperature <- read.table(file=file.choose(), header=TRUE)
names(temperature)
```