**Data Analytics**

Assignment 2

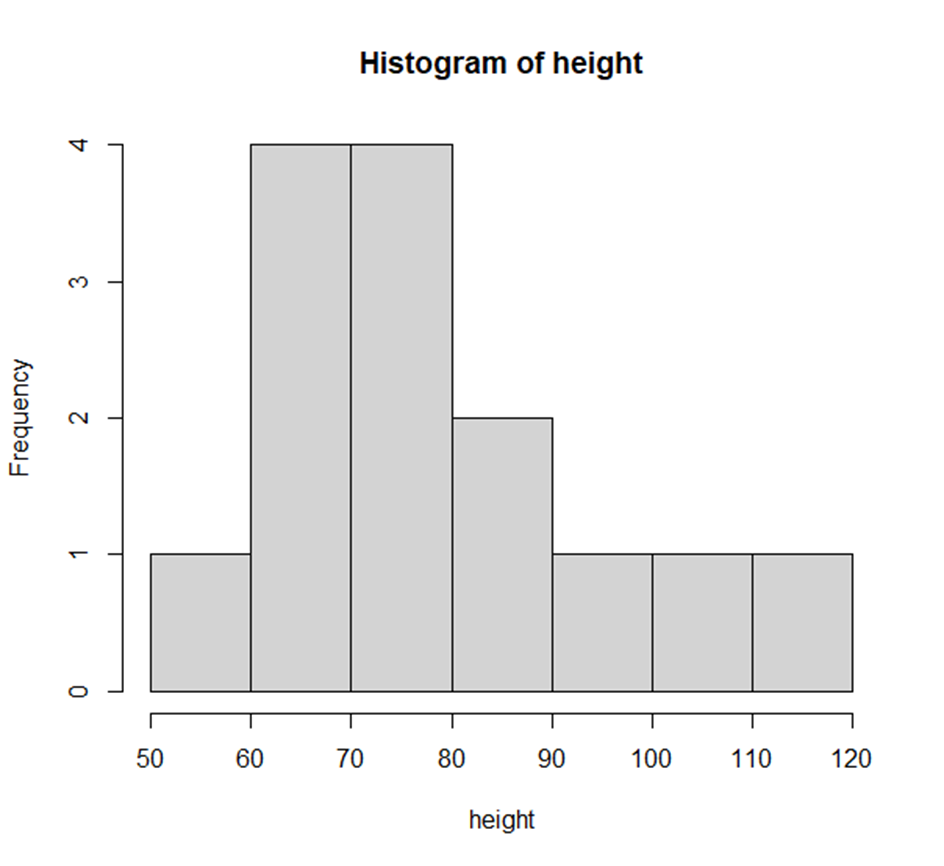
Table 1. Friends data set

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Friend** | **Max temp** | **Weight** | **Height** | **Years** | **Gender** | **Company** |
| Andrew | 25 | 77 | 175 | 10 | M | Good |
| Bernhard | 31 | 110 | 195 | 12 | M | Good |
| Carolina | 15 | 70 | 172 | 2 | F | Bad |
| Dennis | 20 | 85 | 180 | 16 | M | Good |
| Eve | 10 | 65 | 168 | 0 | F | Bad |
| Fred | 12 | 75 | 173 | 6 | M | Good |
| Gwyneth | 16 | 75 | 180 | 3 | F | Bad |
| Hayden | 26 | 63 | 165 | 2 | F | Bad |
| Irene | 15 | 55 | 158 | 5 | F | Bad |
| James | 21 | 66 | 163 | 14 | M | Good |
| Kevin | 30 | 95 | 190 | 1 | M | Bad |
| Lea | 13 | 72 | 172 | 11 | F | Good |
| Marcus | 8 | 83 | 185 | 3 | F | Bad |
| Nigel | 12 | 115 | 192 | 15 | M | Good |

1. The absolute and relative frequencies and respective cumulative frequencies for the attribute “Weight” in Friend data sets are presented in the table below.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Weight** | **Abs. freq.** | **Rel. freq.** | **Abs. cum. freq.** | **Rel. cum. freq.** |
| 55 | 1 | 1/14=7.14% | 1 | 1/14=7.14% |
| 63 | 1 | 1/14=7.14% | 2 | 2/14=14.29% |
| 65 | 1 | 1/14=7.14% | 3 | 3/14=21.43% |
| 66 | 1 | 1/14=7.14% | 4 | 4/14=28.57% |
| 70 | 1 | 1/14=7.14% | 5 | 5/14=35.71% |
| 72 | 1 | 1/14=7.14% | 6 | 6/14=42.86% |
| 75 | 2 | 2/14=14.29% | 8 | 8/14=57.14% |
| 77 | 1 | 1/14=7.14% | 9 | 9/14=64.29% |
| 83 | 1 | 1/14=7.14% | 10 | 10/14=71.43%% |
| 85 | 1 | 1/14=7.14% | 11 | 11/14=78.57% |
| 95 | 1 | 1/14=7.14% | 12 | 12/14=85.71% |
| 110 | 1 | 1/14=7.14% | 13 | 13/14=92.86% |
| 115 | 1 | 1/14=7.14% | 14 | 14/14=100% |

1. Histogram plot for the “height” attribute in Table 1



1. The type of relation between the two attributes in the scatter plot shown in Figure 1 is a **negatively correlated** type.

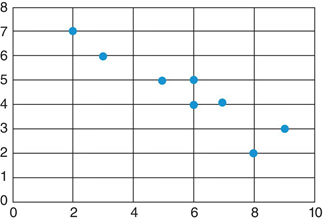


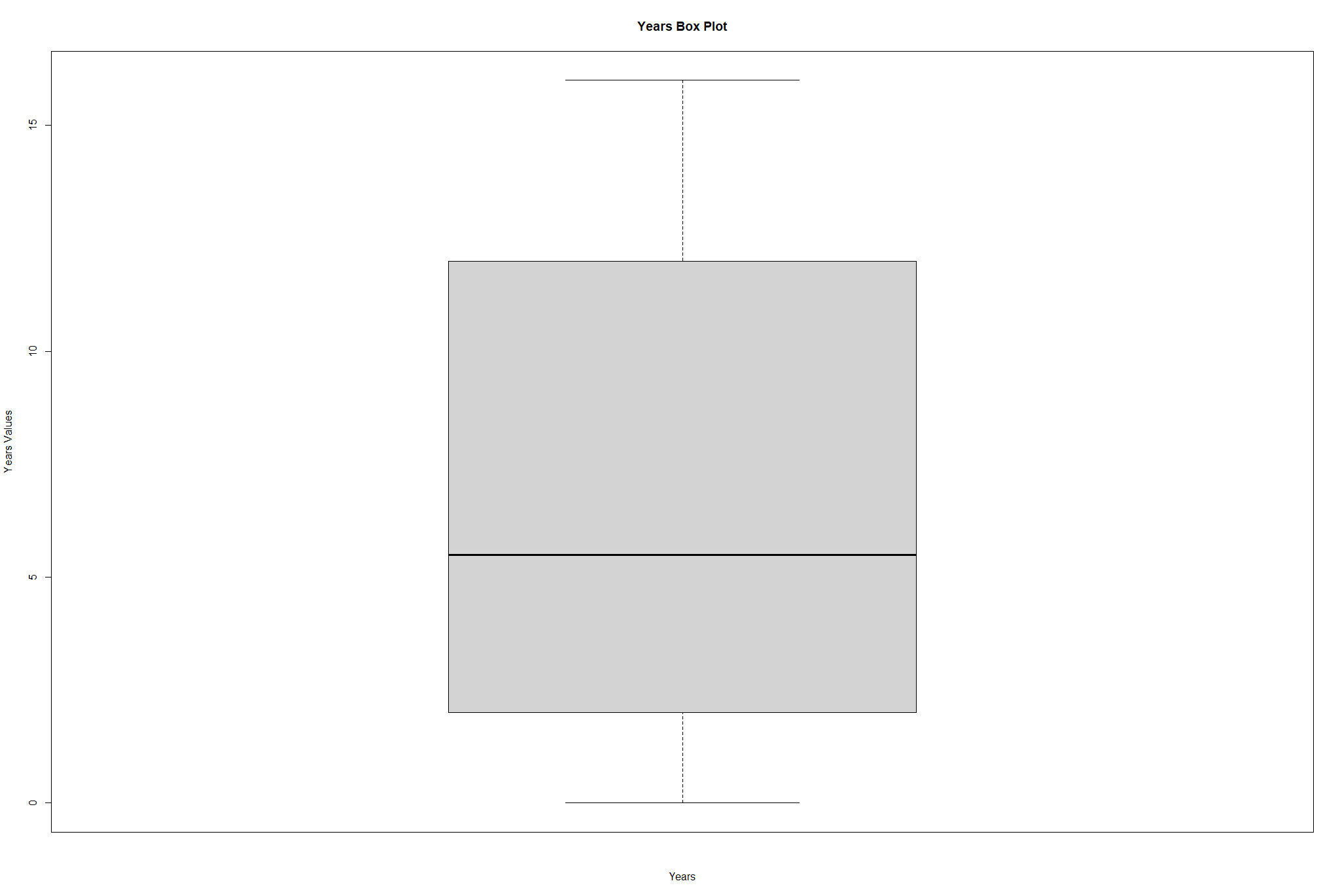
Fig. 1 Scatter plot

**Negative correlation** describes when two variables tend to move in opposite size and direction from one another, such that when one increases the other variable decreases, and vice-versa.

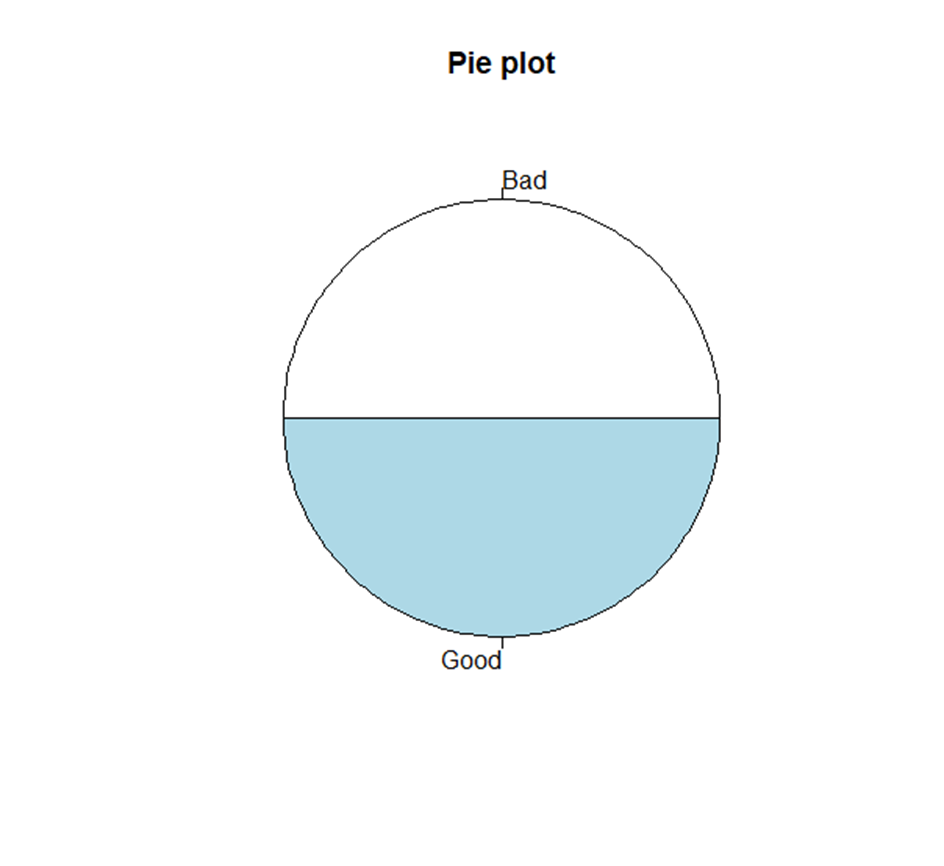
1. Contingency table for the attributes “gender” and “company” in Table 1.

|  |  |
| --- | --- |
| **Company** | |
| Good | Bad |
| **Gender** | Male | 6 | 1 | 7 |
| Female | 1 | 6 | 7 |
| 7 | 7 | 14 |

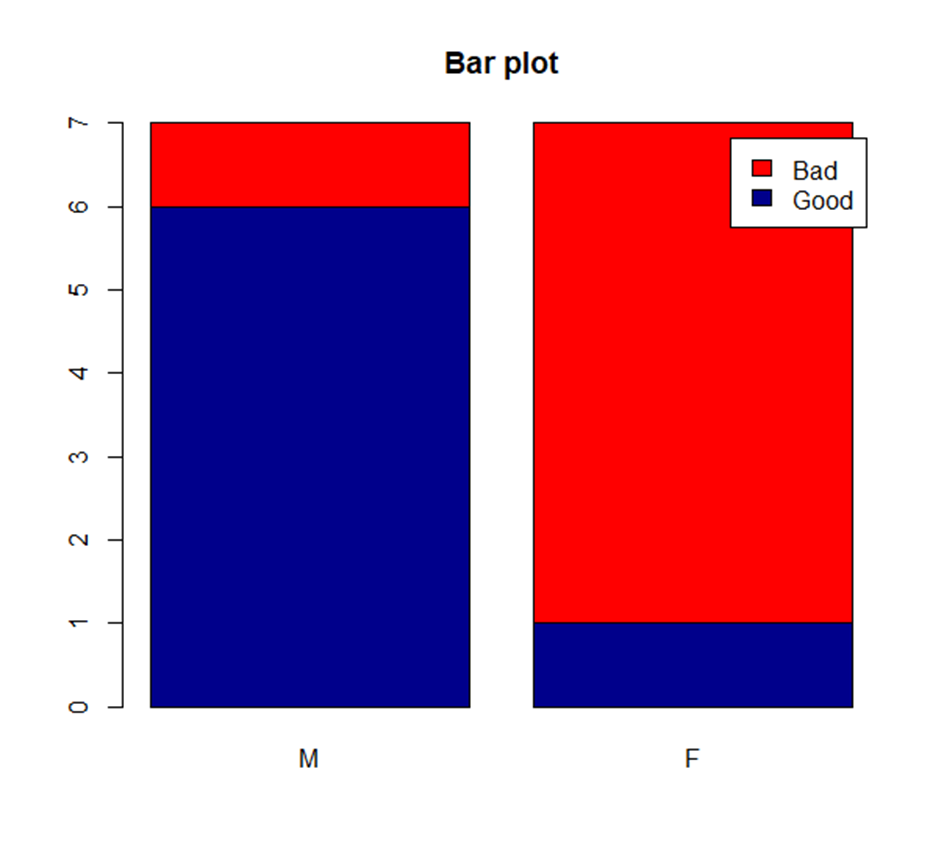
1. The box plot for the attribute “Years” in Table 1.



1. The pie plot for the attribute “Company” in Table 1.



1. The stacked bar plot for the attributes “Gender” and “Company” in Table 1.



1. The mosaic plot for the attributes “Gender” and “Company” in Table 1.



1. The scatter plot for the attributes “Max temp” and “weight” in Table 1.

