Overview of "divusa" Dataset from "faraway" Package

Julian Hatwell

January 3, 2016

This document provides a brief overview of the divusa dataset in the faraway R package.

```
##      year           divorce        unemployed          femlab
##  Min.   :1920   Min.   : 6.10   Min.   : 1.200   Min.   :22.70
##  1st Qu.:1939   1st Qu.: 8.70   1st Qu.: 4.200   1st Qu.:27.47
##  Median :1958   Median :10.60   Median : 5.600   Median :37.10
##  Mean   :1958   Mean   :13.27   Mean   : 7.173   Mean   :38.58
##  3rd Qu.:1977   3rd Qu.:20.30   3rd Qu.: 7.500   3rd Qu.:47.80
##  Max.   :1996   Max.   :22.80   Max.   :24.900   Max.   :59.30
##     marriage          birth          military
##  Min.   : 49.70   Min.   : 65.30   Min.   : 1.940
##  1st Qu.: 61.90   1st Qu.: 68.90   1st Qu.: 3.469
##  Median : 74.10   Median : 85.90   Median : 9.102
##  Mean   : 72.97   Mean   : 88.89   Mean   :12.365
##  3rd Qu.: 80.00   3rd Qu.:107.30   3rd Qu.:14.266
##  Max.   :118.10   Max.   :122.90   Max.   :86.641
```

From the summary, and the associated help (not shown), the following observations can be made:

The dataframe contains 77 rows and 7 columns.
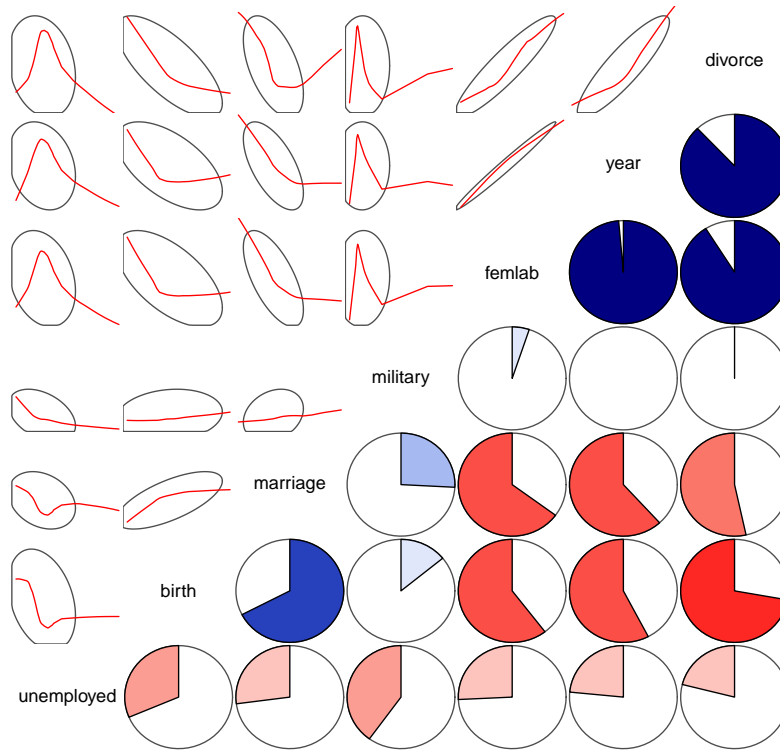
Figure 1: multi-variate comparisons

```
# Year is an integer value and should be
# used as a time series.

df <- df %>% mutate(year = ts(year, frequency = 1))
```
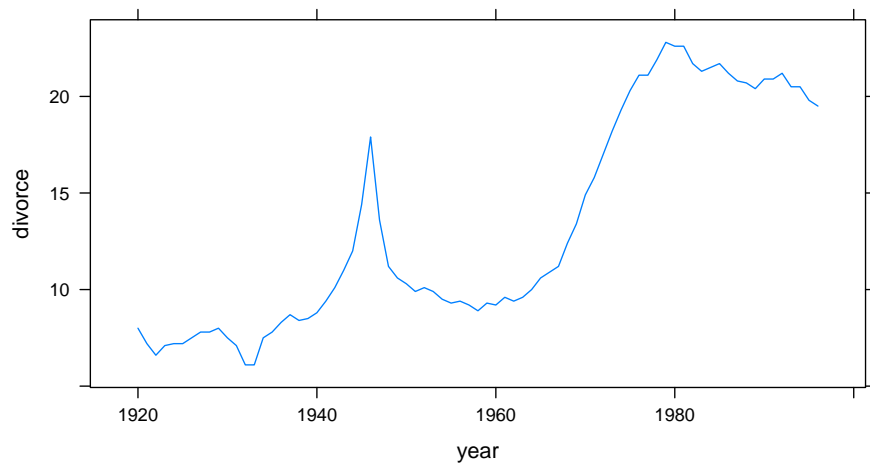
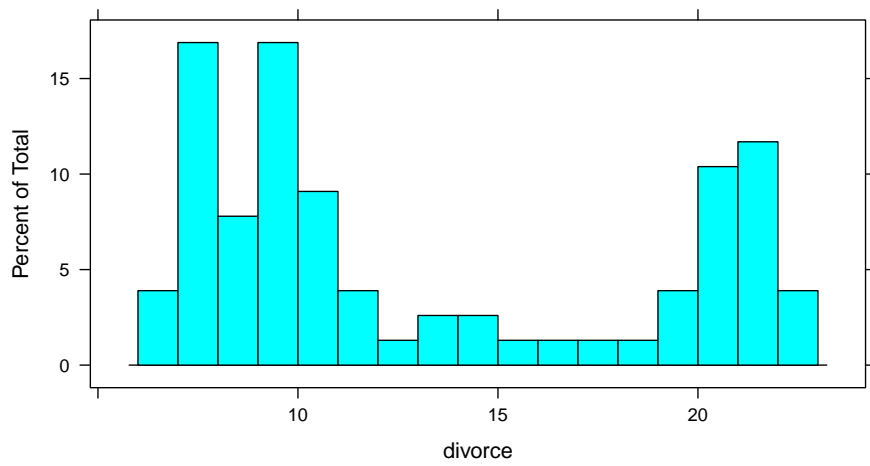Figure 2: Time Series plot of divorce against year

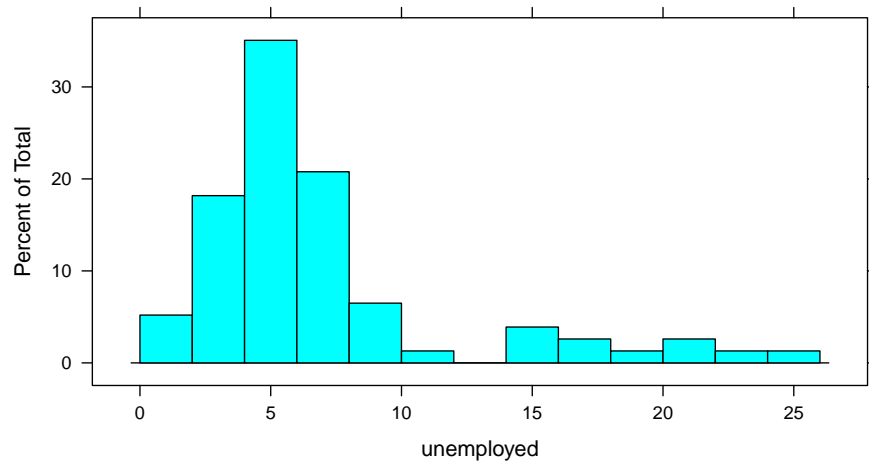
Figure 3: Histogram of the divorce variable
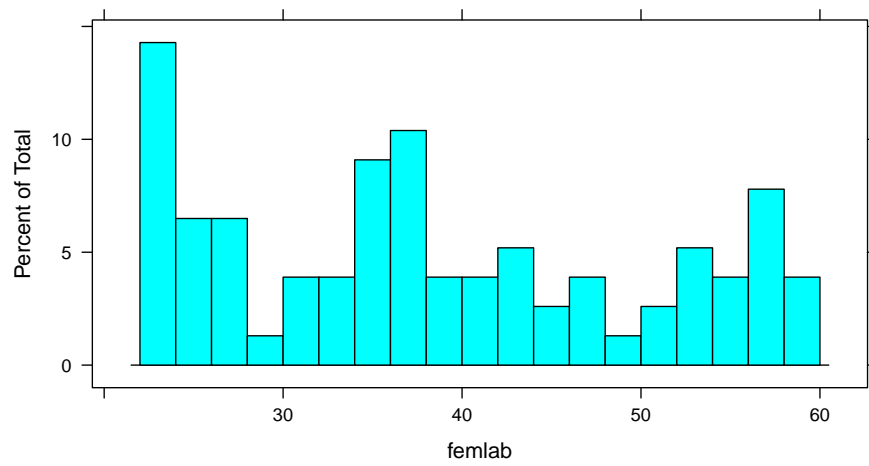
Figure 4: Histogram of the unemployed variable

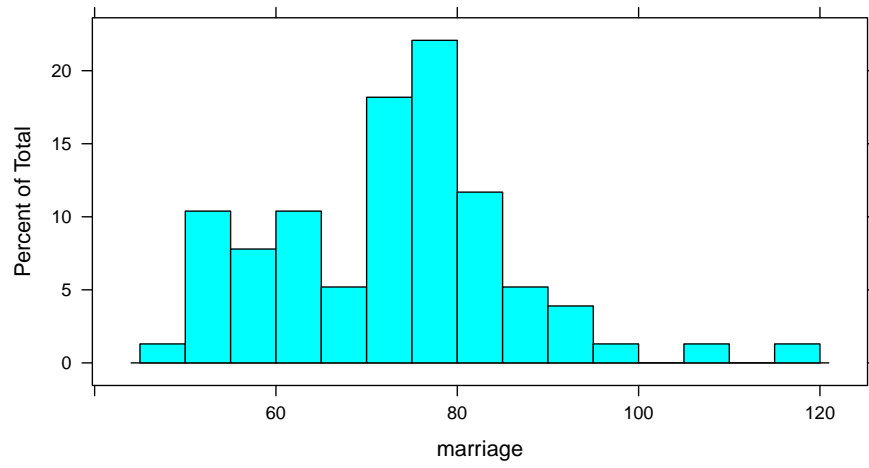Figure 5: Histogram of the femlab variable
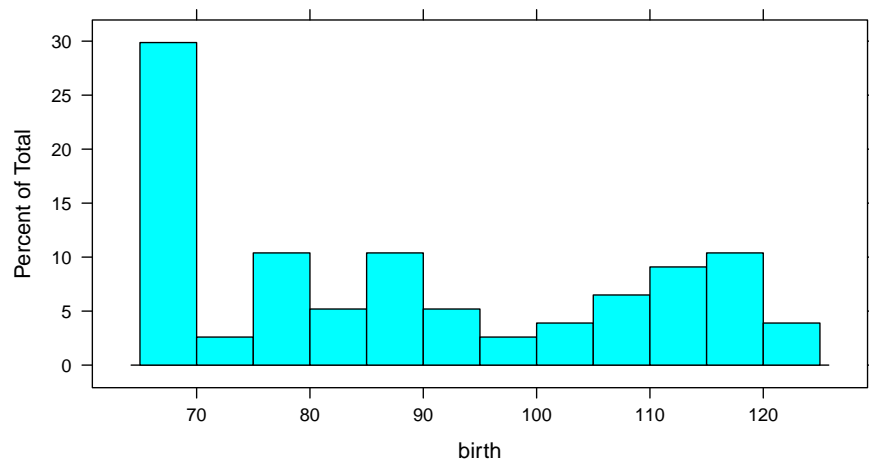
4

Figure 6: Histogram of the marriage variable

Figure 7: Histogram of the birth variable

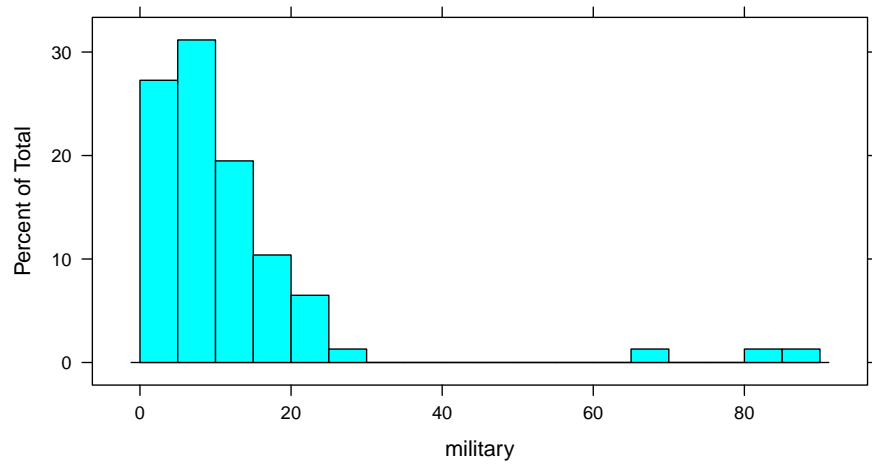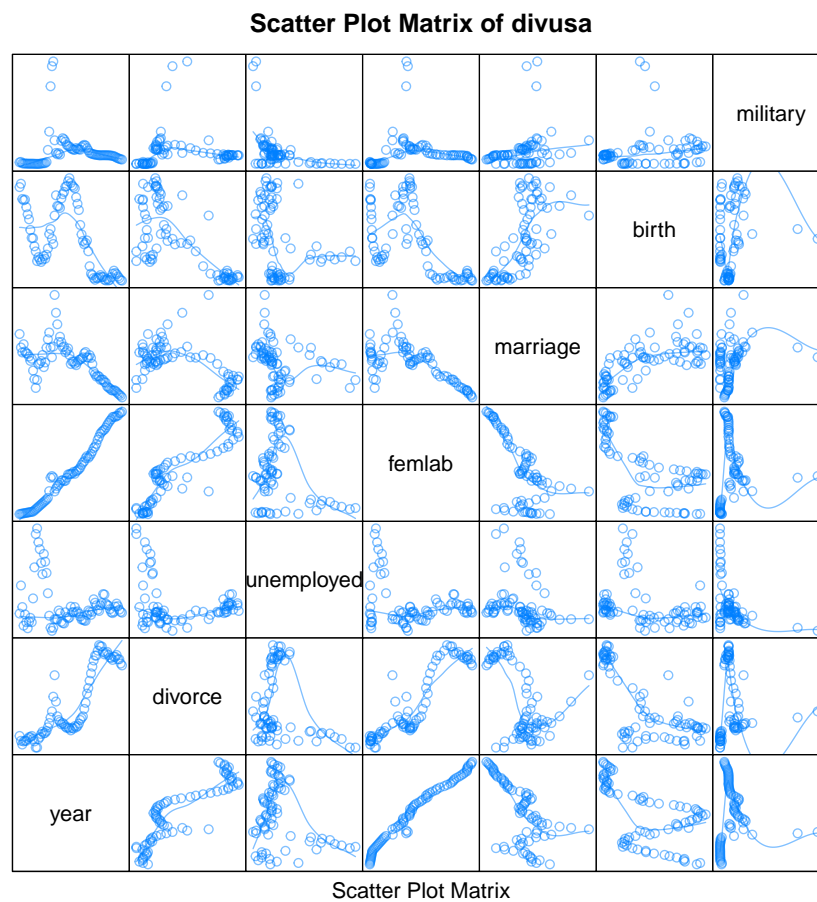Figure 8: Histogram of the military variable

Scatter Plot Matrix of divusa

Scatter Plot Matrix

Figure 9: multi-variate comparisons