```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
%matplotlib inline
%matplotlib inline
%load_ext rpy2.ipynb
```

```
%load_ext rpy2.ipynb
%%R
df = read.csv("household.csv")
```

```
%%R
head(df,5)
```

```
  GEO_REGION GEO_PROVINCE GEO_CENTER_STATUS  YEAR CODE_HH  HH_TYPE  LAT  LONG
1   BETWEEN       NAN!       NAN!       NA    NA     NA     NA    NA    NA
2   BETWEEN       NAN!       NAN!       NA    NA     NA     NA    NA    NA
3   BETWEEN       NAN!       NAN!       NA    NA     NA     NA    NA    NA
4   BETWEEN       NAN!       NAN!       NA    NA     NA     NA    NA    NA
5   BETWEEN       NAN!       NAN!       NA    NA     NA     NA    NA    NA
  Household_type Income_per_month Income_per_capita Income_per_consumption
1        Family             NaN             NaN                   NaN
2        Family             NaN             NaN                   NaN
3        Family             NaN             NaN                   NaN
4        Family             NaN             NaN                   NaN
5        Family             NaN             NaN                   NaN
  Income_per_expenditure
1                    NaN
2                    NaN
3                    NaN
4                    NaN
5                    NaN
```

5 rows of 20 columns

```
%%R
dim(df)
```

```
[1] 1000   20
```

```
%%R
colnames(df)
```

```
 [1] "GEO_REGION"          "GEO_PROVINCE"        "GEO_CENTER_STATUS"
 [4] "YEAR"                "CODE_HH"             "HH_TYPE"
 [7] "LAT"                 "LONG"                "Household_type"
[10] "Income_per_month"    "Income_per_capita"   "Income_per_consumption"
[13] "Income_per_expenditure"
```

```
%%R
summary(df)
```

```
   GEO_REGION         GEO_PROVINCE       GEO_CENTER_STATUS      YEAR
 Length:1000        Min.   :   NA    Min.   :   NA    Min.   :   NA
 Class :character   1st Qu.:   NA    1st Qu.:   NA    1st Qu.:   NA
 Mode  :character   Median :   NA    Median :   NA    Median :   NA
                    Mean   :  NaN    Mean   :  NaN    Mean   :  NaN
                    3rd Qu.:   NA    3rd Qu.:   NA    3rd Qu.:   NA
                    Max.   :   NA    Max.   :   NA    Max.   :   NA
                    NA's   :1000     NA's   :1000     NA's   :1000
     CODE_HH        HH_TYPE          LAT             LONG       Household_type
 Min.   :   NA    Min.   :   NA    Min.   :   NA    Min.   :   NA    Family:   0
 1st Qu.:   NA    1st Qu.:   NA    1st Qu.:   NA    1st Qu.:   NA
 Median :   NA    Median :   NA    Median :   NA    Median :   NA
 Mean   :  NaN    Mean   :  NaN    Mean   :  NaN    Mean   :  NaN
 3rd Qu.:   NA    3rd Qu.:   NA    3rd Qu.:   NA    3rd Qu.:   NA
 Max.   :   NA    Max.   :   NA    Max.   :   NA    Max.   :   NA
 NA's   :1000     NA's   :1000     NA's   :1000     NA's   :1000
 Income_per_month Income_per_capita Income_per_consumption
 Min.   :   NA    Min.   :   NA    Min.   :   NA
 1st Qu.:   NA    1st Qu.:   NA    1st Qu.:   NA
 Median :   NA    Median :   NA    Median :   NA
 Mean   :  NaN    Mean   :  NaN    Mean   :  NaN
 3rd Qu.:   NA    3rd Qu.:   NA    3rd Qu.:   NA
 Max.   :   NA    Max.   :   NA    Max.   :   NA
 NA's   :1000     NA's   :1000     NA's   :1000
 Income_per_expenditure
 Min.   :   NA
 1st Qu.:   NA
 Median :   NA
 Mean   :  NaN
 3rd Qu.:   NA
 Max.   :   NA
 NA's   :1000
```

```
%%R
str(df)
```

```
'data.frame':	1000 obs. of  20 variables:
 $ GEO_REGION          : chr  "BETWEEN" "BETWEEN" "BETWEEN" "BETWEEN" ...
 $ GEO_PROVINCE        : int  NA NA NA NA NA NA NA NA NA NA ...
 $ GEO_CENTER_STATUS   : int  NA NA NA NA NA NA NA NA NA NA ...
 $ YEAR                : int  NA NA NA NA NA NA NA NA NA NA ...
 $ CODE_HH             : int  NA NA NA NA NA NA NA NA NA NA ...
 $ HH_TYPE             : int  NA NA NA NA NA NA NA NA NA NA ...
 $ LAT                 : int  NA NA NA NA NA NA NA NA NA NA ...
 $ LONG                : int  NA NA NA NA NA NA NA NA NA NA ...
 $ Household_type      : chr  "Family" "Family" "Family" "Family" ...
 $ Income_per_month    : num  NA NA NA NA NA NA NA NA NA NA ...
 $ Income_per_capita   : num  NA NA NA NA NA NA NA NA NA NA ...
 $ Income_per_consumption: num  NA NA NA NA NA NA NA NA NA NA ...
 $ Income_per_expenditure: num  NA NA NA NA NA NA NA NA NA NA ...
 $ Income_per_income   : num  NA NA NA NA NA NA NA NA NA NA ...
 $ Income_per_income_per:
```