한국인의 삶을 파악하라

- 2006~2015년까지 전국에서 7000여 가구를 선정하여 매년 추적 조사한 자료
- 데이터 셋 : Koweps_hpc10_2015_beta1.sav
  - 2016년도 발간한 복지패널 데이터 6,914가구, 16,664명에 대한 정보

In [1]:

```
install.packages("foreign")
```

```
Warning message:
"unable to access index for repository http://www.stats.ox.ac.uk/pub/RWin/bin/window
s/contrib/3.5: (http://www.stats.ox.ac.uk/pub/RWin/bin/windows/contrib/3.5:)
  URL 'http://www.stats.ox.ac.uk/pub/RWin/bin/windows/contrib/3.5/PACKAGES'를 열 수
 없습니다"

package 'foreign' successfully unpacked and MD5 sums checked

The downloaded binary packages are in
        C:₩Users₩WITHJS₩AppData₩Local₩Temp₩RtmpCgEfb6₩downloaded_packages
```

In [2]:

```
library(foreign)
library(dplyr)
library(ggplot2)
library(readxl)
```

```
Attaching package: 'dplyr'

The following objects are masked from 'package:stats':

    filter, lag

The following objects are masked from 'package:base':

    intersect, setdiff, setequal, union
```

In [3]:

```
dat_welfare <- read.spss(file="D:₩₩dataset₩₩R_Doit₩₩Koweps_hpc10_2015_beta1.sav", to.data.frame=T)
welfare <- dat_welfare
```

```
Warning message in read.spss(file = "D:₩₩dataset₩₩R_Doit₩₩Koweps_hpc10_2015_beta1.sa
v", :
"D:₩dataset₩R_Doit₩Koweps_hpc10_2015_beta1.sav: Compression bias (0) is not the usua
l value of 100"
```

데이터 탐색해 보기

- head(welfare)
- tail(welfare)

- View(welfare)
- dim(welfare)
- str(welfare)
- summary(welfare)

In [4]:

```
welfare <- rename(welfare,
                  sex=h10_g3,
                  birth=h10_g4,
                  marriage=h10_g10,
                  religion=h10_g11,
                  income=p1002_8aq1,
                  code_job=h10_eco9,
                  code_region=h10_reg7)
names(welfare)
```

'h10_id' 'h10_ind' 'h10_sn' 'h10_merkey' 'h_new' 'h10_cobf' 'h10_reg5'
'code_region' 'h10_din' 'h10_cin' 'h10_flag' 'p10_wgl' 'p10_wsl' 'p10_wgc' 'p10_wsc'
'h10_hc' 'nh1001_1' 'nh1001_2' 'h1001_1' 'h10_pind' 'h10_pid' 'h10_g1' 'h10_g2'
'sex' 'birth' 'h10_g6' 'h10_g7' 'h10_g8' 'h10_g9' 'marriage' 'religion' 'h10_g12'
'h1001_110' 'h1001_5aq1' 'h1001_5aq2' 'h1001_5aq3' 'h1001_5aq4' 'h10_med1'
'h10_med2' 'h10_med3' 'h10_med4' 'h10_med5' 'h10_med6' 'h10_med7' 'h10_med8'
'h10_g9_1' 'h10_med9' 'h10_med10' 'h10_eco1' 'h10_eco2' 'h10_eco3' 'h10_eco4'
'h10_eco4_1' 'h10_eco5_1' 'h10_eco6' 'h10_eco_7_1' 'h10_eco_7_2' 'h10_eco_7_3'
'h10_eco8' 'code_job' 'h10_eco10' 'h10_eco11' 'h10_soc1' 'h10_soc_2' 'h10_soc_3'
'h10_soc_4' 'h10_soc_5' 'h10_soc_6' 'h10_soc_7' 'h10_soc_8' 'h10_soc_9'
'h10_soc_10' 'h10_soc_11' 'h10_soc8' 'h10_soc9' 'h10_soc11' 'h10_soc10'
'h10_soc_12' 'h10_soc_13' 'h1005_1' 'h1005_3aq1' 'h1005_2' 'h1005_3' 'h1005_4'
'h1005_5' 'h1005_6' 'h1005_7' 'nh1005_8' 'nh1005_9' 'h1005_3aq2' 'h1006_aq1'
'h1006_1' 'h1006_2' 'h1006_4' 'h1006_5' 'h1006_3' 'h1006_6' 'h1006_8' 'h1006_9'
'h1006_aq2' 'h1006_aq3' 'h1006_10' 'h1006_11' 'h1006_12' 'h1006_13' 'h1006_14'
'h1006_21' 'h1006_22' 'h1006_23' 'h1006_24' 'h1006_25' 'h1006_27' 'h1006_30'
'h1006_33' 'h1006_36' 'h1006_39' 'h1006_3aq1' 'h1007_3aq1' 'h1007_3aq2'
'h1007_5aq1' 'h1007_3aq3' 'h1007_3aq4' 'h1007_3aq5' 'h1007_6aq1' 'h1007_3aq6'
'h1007_5aq2' 'h1007_3aq7' 'h1007_3aq8' 'h1007_3aq9' 'h1007_3aq10' 'h1007_3aq11'
'h1007_5aq3' 'h1007_5aq4' 'h1007_3aq13' 'h1007_6aq4' 'h1007_6aq6' 'h1007_3aq14'
'h1007_3aq15' 'h1007_3aq16' 'h1007_3aq17' 'h1007_4' 'h1007_6aq7' 'h1007_6aq8'
'h1007_6aq9' 'h1007_6aq10' 'h1007_6aq11' 'h1007_5' 'h1007_6aq12' 'h1007_6aq13'
'h1007_6aq14' 'h1007_9' 'h1009_9' 'h1009_6aq4' 'h10_inc1' 'h10_inc2_1'
'h10_inc2_2' 'h10_inc3_1' 'h10_inc3_2' 'h10_inc4_1' 'h10_inc4_2' 'h10_inc5_1'
'h10_inc5_2' 'h10_inc6_1' 'h10_inc6_2' 'h10_inc7_1' 'h10_inc7_2' 'h1008_106'
'h1008_107' 'h1008_108' 'h1008_109' 'h1008_110' 'h1008_111' 'h10_inc2_3'
'h10_inc2' 'h10_inc3_6' 'h10_inc3' 'h10_inc4_7' 'h10_inc4' 'h10_inc4_8' 'h10_inc4_9'
'h1008_155' 'h1008_156' 'h1008_157' 'h1008_158' 'h1008_160' 'h1008_159'
'h1008_3aq3' 'h1008_161' 'h1008_162' 'h1008_163' 'h1008_164' 'h1008_166'
'h1008_165' 'h1008_3aq4' 'h1008_167' 'h1008_168' 'h1008_169' 'h1008_170'
'h10_inc7_3' 'h10_inc7' 'h1008_aq9' 'h1008_aq10' 'h1008_aq11' 'h1008_aq12'
'h1008_aq13' 'h1008_aq14' 'h1008_aq15' 'h1008_6aq1' 'h1008_aq16' 'h1008_aq17'
'h1008_10aq1' 'h1008_aq19' 'h1008_aq20' 'h1008_aq21' 'h1008_5aq3' 'h1008_7aq1'
'h1008_aq22' 'h1008_7aq2' 'h1008_aq23' 'h1008_aq24' 'h1008_4aq116'
'h1008_4aq117' 'h1008_5aq1' 'h1008_7aq4' 'h1008_7aq5' 'h1008_7aq6' 'h1008_7aq7'
'h1008_7aq8' 'h1008_7aq9' 'h1008_aq25' 'h1008_7aq10' 'h1008_aq26' 'h1008_aq27'
'h1008_aq28' 'h1008_aq29' 'h1008_3aq5' 'h1008_4aq118' 'h1008_aq30' 'h1008_6aq3'
'h1008_3aq6' 'h1008_3aq7' 'nh1008_3aq1' 'h1008_aq32' 'h1008_aq33' 'h1008_aq34'
'h1008_3aq8' 'h1008_195' 'h1008_7aq11' 'h1009_aq1' 'h1009_aq2' 'h1009_aq3'
'h1009_aq4' 'h1009_aq5' 'h1009_aq6' 'h1009_aq7' 'h1009_aq8' 'h1010_aq1'
'h1010_aq2' 'h1010_aq3' 'h1010_aq4' 'h1010_aq5' 'h1010_aq6' 'h1010_aq7'

'h1010_aq8' 'h1010_aq9' 'h1010_aq10' 'h1010_aq11' 'h1010_aq12' 'h1010_aq13'
'h1010_aq14' 'h1010_aq15' 'h1010_aq16' 'h1010_aq17' 'h1010_aq18' 'h1010_aq19'
'h1010_aq20' 'h1010_26' 'h1010_27' 'h1010_aq23' 'h1010_aq24' 'h1010_aq25'
'h1010_aq26' 'h1011_2' 'h1011_3' 'h1011_4' 'h1011_5' 'h1011_6' 'h1011_7'
'h1011_8' 'h1011_3aq1' 'h1011_3aq2' 'h1011_3aq3' 'h1011_3aq4' 'h1011_3aq5'
'h1011_3aq6' 'h1011_3aq7' 'h1012_1' 'nh1012_1' 'h1012_2' 'h1012_3' 'h1012_4'
'h1012_5' 'h1012_6' 'h1012_7' 'nh1012_7' 'nh1012_8' 'h1012_9' 'h1012_10'
'h1012_11' 'h1012_12' 'h1012_13' 'h1012_14' 'h1012_15' 'h1012_3aq1' 'h1012_16'
'h1012_17' 'h1012_18' 'h1012_19' 'h1012_1_4aq1' 'h1012_1_5aq1' 'h1012_1_5aq2'
'h1012_1_5aq3' 'h1012_1_5aq4' 'h1012_1_5aq5' 'h1012_1_4aq2' 'h1012_1_4aq3'
'h1013_2' 'h1013_6' 'h1013_10' 'h1013_14' 'h1013_18' 'h1013_22' 'h1013_26'
'h1013_8aq1' 'h1013_5aq1' 'h1013_8aq2' 'h1013_4aq1' 'h1013_4aq2' 'h1013_4aq4'
'h1013_4aq6' 'h1013_4aq8' 'h1013_4aq10' 'h1013_5aq4' 'h1013_5aq6' 'h1013_5aq8'
'h1013_6aq1' 'h1013_4aq14' 'h1013_4aq15' 'h1013_4aq16' 'h1013_4aq17'
'h1013_4aq18' 'h1013_4aq20' 'h1013_4aq22' 'h1013_4aq24' 'h1013_4aq26'
'h1013_4aq28' 'h1013_4aq30' 'h1013_4aq32' 'h1014_4' 'h1014_8' 'h1014_12'
'h1014_16' 'h1014_20' 'h1014_24' 'h1014_28' 'h1014_32' 'h1014_36' 'h1014_3aq1'
'h1014_4aq1' 'h1015_4' 'h1015_8' 'h1015_12' 'h1015_20' 'h1015_25' 'h1015_29'
'h1015_33' 'h1015_37' 'h1015_4aq1' 'h1015_7aq1' 'h1015_aq1' 'h1015_40'
'h1015_41' 'h1015_42' 'h1015_43' 'h1015_44' 'h1015_45' 'h1015_46' 'h1015_47'
'h1015_48' 'h1015_49' 'h1015_50' 'h1015_51' 'h1015_52' 'h1015_53' 'h1015_54'
'h1015_55' 'h1015_56' 'h1015_57' 'h1015_60' 'h1015_aq2' 'h1015_61' 'h1015_62'
'h1015_63' 'h1015_66' 'h1015_67' 'h1015_68' 'h1015_aq3' 'h1015_69' 'h1015_70'
'h1015_71' 'h1015_74' 'h1015_75' 'h1015_76' 'h1015_aq4' 'h1015_77' 'h1015_78'
'h1015_79' 'h1015_82' 'h1015_83' 'h1015_84' 'h1015_aq5' 'h1015_85' 'h1015_86'
'h1015_87' 'h1015_90' 'h1015_91' 'h1015_92' 'h1015_aq6' 'h1015_93' 'h1015_94'
'h1015_95' 'h1015_98' 'h1015_99' 'h1015_100' 'h1015_aq7' 'h1015_101' 'h1015_102'
'h1015_103' 'h1015_106' 'h1015_107' 'h1016_6aq1' 'h1016_6aq4' 'h1016_8'
'h1016_24' 'h1016_4aq1' 'h1016_4aq4' 'h1016_32' 'h1016_36' 'h1016_40'
'h1016_4aq7' 'h1016_56' 'h1016_60' 'h1016_64' 'h1017_1' 'h1017_2' 'h1017_3'
'h1017_4' 'h1017_5' 'h1017_6' 'h1017_7' 'p10_fnum' 'p10_tq' 'p10_cp' 'p1001_1'
'p1001_2' 'p1001_aq1' 'p1001_3' 'p1001_4' 'p1001_5' 'p1001_6' 'p1001_10'
'p1001_11' 'p1001_12' 'p1001_13' 'p1001_14' 'p1001_7' 'p1001_8' 'p1001_9'
'p1001_15' 'p1001_16' 'p1001_17' 'p1001_18' 'p1001_19' 'p1001_20' 'p1001_21'
'p1001_22' 'p1001_23' 'p1001_24' 'p1001_25' 'p1001_27' 'p1001_28' 'p1001_29'
'p1001_30' 'p1001_31' 'p1001_32' 'p1001_33' 'p1001_34' 'p1002_1' 'p1002_2'
'p1002_3' 'p1002_4' 'p1002_5' 'p1002_6' 'p1002_7' 'p1002_8' 'income' 'p1002_9'
'p1002_8aq2' 'p1002_4aq1' 'p1002_10' 'p1002_11' 'p1002_12' 'p1002_8aq3'
'p1002_8aq4' 'p1002_31' 'p1002_4aq2' 'p1002_4aq3' 'p1002_4aq4' 'p1002_8aq5'
'p1002_8aq6' 'p1002_8aq7' 'p1002_8aq8' 'p1002_8aq9' 'p1002_8aq10' 'p1002_8aq11'
'p1002_8aq12' 'p1002_8aq13' 'p1002_8aq14' 'p1002_8aq15' 'p1002_aq2' 'p1002_aq3'
'p1002_aq4' 'p1002_aq5' 'p1003_1' 'p1003_2' 'p1003_5' 'p1003_6' 'p1003_7'
'p1003_8' 'p1003_9' 'p1003_10' 'p1003_11' 'p1003_12' 'p1004_1' 'p1004_2'
'p1004_3' 'p1004_4' 'p1004_5' 'p1004_6' 'p1004_7' 'p1004_8' 'p1004_9' 'p1004_10'
'p1004_11' 'p1004_12' 'p1004_13' 'p1004_3aq1' 'p1004_3aq2' 'p1004_3aq3'
'p1004_3aq4' 'p1004_3aq5' 'p1004_3aq6' 'p1004_3aq7' 'p1004_3aq8' 'p1005_3aq1'
'p1005_3aq2' 'p1005_3aq3' 'p1005_3aq4' 'p1005_3aq5' 'p1005_3aq6' 'p1005_3aq7'
'p1005_3aq8' 'p1005_3aq9' 'p1005_3aq10' 'p1005_2' 'p1005_3' 'p1005_4aq1'
'p1005_4aq2' 'p1005_4aq3' 'p1005_4aq4' 'p1005_4aq5' 'p1005_4aq6' 'p1005_4aq7'
'p1005_4aq8' 'p1005_5' 'p1005_6' 'p1005_7' 'p1005_8' 'p1005_3aq11' 'p1005_9'
'p1005_10' 'p1005_11' 'p1005_12' 'p1005_13' 'p1005_14' 'p1005_15' 'p1005_16'

'p1005_17' 'p1005_18' 'p1005_19' 'p1005_20' 'p1005_21' 'p1005_22' 'p1005_23'
'p1005_24' 'p1005_25' 'p1005_26' 'p1005_27' 'p1005_28' 'p1005_29' 'p1005_4aq9'
'p1005_4aq10' 'p1005_4aq11' 'p1005_aq1' 'p1005_aq2' 'p1005_aq3' 'p1005_aq4'
'p1005_6aq1' 'p1005_6aq2' 'p1005_6aq3' 'p1005_6aq4' 'p1005_6aq5' 'p1005_6aq6'
'p1005_6aq7' 'p1005_6aq8' 'p1005_6aq9' 'p1005_7aq1' 'p1005_7aq2' 'p1005_7aq3'
'p1007_3aq1' 'p1007_3aq2' 'p1007_3aq3' 'p1007_3aq4' 'p1007_3aq5' 'p1007_3aq6'
'p1007_3aq7' 'np1006_1' 'np1006_2' 'np1006_3' 'np1006_4' 'np1006_5' 'np1006_6'
'np1006_7' 'np1006_8' 'np1006_9' 'np1006_10' 'np1006_11' 'np1006_12' 'np1006_13'
'np1006_14' 'np1006_15' 'np1006_16' 'np1006_17' 'np1006_18' 'np1006_19'
'np1006_20' 'np1006_21' 'np1006_22' 'np1006_23' 'np1006_24' 'np1006_25'
'np1006_26' 'np1006_27' 'np1006_28' 'np1006_29' 'np1006_30' 'np1006_31'
'np1006_32' 'np1006_33' 'np1006_34' 'np1006_35' 'np1006_36' 'np1006_37'
'np1006_38' 'np1006_39' 'np1006_40' 'np1006_41' 'np1006_42' 'np1006_43'
'np1006_44' 'p1006_3aq1' 'c10_fnum' 'c10_cp' 'c10_grade' 'c1001_1' 'c1001_2'
'c1001_3' 'c1001_4' 'c1001_5' 'c1001_6' 'c1001_7' 'c1001_8' 'c1001_9' 'c1001_10'
'c1001_11' 'c1001_12' 'c1001_13' 'c1001_7aq5' 'c1001_7aq6' 'c1001_7aq7'
'c1001_7aq8' 'c1001_4aq1' 'c1001_4aq2' 'c1001_4aq3' 'c1001_4aq4' 'c1001_4aq5'
'c1001_4aq6' 'c1002_1' 'c1002_2' 'c1002_3' 'c1002_4' 'c1002_5' 'c1002_6'
'c1002_7' 'c1002_8' 'c1002_9' 'c1002_10' 'c1002_11' 'c1002_12' 'c1002_13'
'c1002_14' 'c1002_15' 'c1002_16' 'c1002_17' 'c1002_18' 'c1002_19' 'c1002_20'
'c1002_21' 'c1002_22' 'c1002_23' 'c1002_24' 'c1002_25' 'c1002_26' 'c1002_4aq1'
'c1002_27' 'c1002_28' 'c1002_29' 'c1002_30' 'c1002_31' 'c1002_32' 'c1002_33'
'c1002_34' 'c1002_35' 'c1002_36' 'c1002_37' 'c1002_38' 'c1002_39' 'c1002_40'
'c1002_41' 'c1002_42' 'c1002_43' 'c1002_44' 'c1002_45' 'c1002_46' 'c1002_47'
'c1002_48' 'c1002_49' 'c1002_50' 'c1002_51' 'c1002_52' 'c1002_53' 'c1002_54'
'c1002_55' 'c1002_56' 'c1002_57' 'c1002_58' 'c1002_4aq2' 'c1002_59' 'c1002_60'
'c1002_61' 'c1002_62' 'c1002_63' 'c1002_64' 'c1002_65' 'c1002_66' 'c1002_67'
'c1002_68' 'c1002_69' 'c1002_70' 'c1002_71' 'c1002_72' 'c1002_73' 'c1002_74'
'c1002_75' 'c1002_76' 'c1002_77' 'c1002_7aq1' 'c1002_7aq2' 'c1002_7aq3'
'c1002_7aq4' 'c1002_7aq5' 'c1002_7aq6' 'c1002_7aq7' 'c1002_7aq8' 'c1002_7aq9'
'c1002_7aq10' 'c1002_7aq11' 'c1002_7aq12' 'c1002_7aq13' 'c1002_7aq14' 'c1002_78'
'c1002_79' 'c1002_80' 'c1002_81' 'c1002_82' 'c1002_83' 'c1002_84' 'c1002_85'
'c1002_86' 'c1002_87' 'c1002_88' 'c1002_89' 'c1002_90' 'c1002_91' 'c1002_92'
'c1002_93' 'c1002_94' 'c1002_95' 'c1002_96' 'c1002_97' 'c1002_98' 'c1002_4aq3'
'c1002_4aq4' 'c1002_4aq5' 'c1002_4aq6' 'c1002_4aq7' 'c1002_4aq8' 'c1002_7aq15'
'c1003_4aq1' 'c1003_4aq2' 'c1003_4aq3' 'c1003_4aq4' 'c1003_4aq5' 'c1003_4aq6'
'c1003_4aq7' 'c1003_4aq8' 'c1003_4aq9' 'c1003_6' 'c1003_7' 'c1003_8' 'c1003_9'
'c1003_13' 'c1003_14' 'c1003_15' 'c1004_4aq1' 'c1004_4aq2' 'c1004_4aq3'
'c1004_4aq4' 'c1004_4aq5' 'c1004_4aq6' 'c1004_1' 'c1004_2' 'c1004_3' 'c1004_4'
'c1004_5' 'c1004_6' 'c1004_7' 'c1004_8' 'c1004_9' 'c1004_10' 'c1004_4aq7'
'c1004_4aq8' 'c1005_12' 'c1005_13' 'c1005_14' 'c1005_15' 'c1005_16' 'c1005_17'
'c1005_4aq1' 'c1005_18' 'c1005_19' 'c1005_7aq1' 'c1005_7aq2' 'c1005_7aq3'
'c1005_7aq4' 'c1005_7aq5' 'c1005_7aq6' 'c1005_7aq7' 'c1005_7aq8' 'c1005_7aq9'
'c1005_7aq10' 'c1005_7aq11' 'c1005_7aq12' 'c1005_7aq13' 'c1005_7aq14'
'c1005_7aq15' 'c1005_7aq16' 'c1005_7aq17' 'c1005_7aq18' 'c1005_7aq19'
'c1005_7aq20' 'c1005_20' 'c1005_21' 'c1005_22' 'c1005_23' 'c1005_24' 'c1005_25'
'c1005_26' 'c1005_27' 'c1005_28' 'c1005_29' 'c1005_30' 'c1005_31' 'c1005_32'
'c1005_36' 'c1005_4aq2' 'c1005_38' 'c1005_4aq3' 'c1005_40' 'c1005_10aq1'
'c1005_42' 'c1005_4aq4' 'c1005_44' 'c1005_4aq5' 'c1005_46' 'c1005_4aq6'
'c1005_48' 'c1005_4aq7' 'c1005_4aq8' 'c1005_4aq9' 'c1005_4aq10' 'c1005_4aq11'
'c1005_4aq12' 'c1005_4aq13' 'c1005_4aq14' 'c1005_4aq15' 'c1007_4aq1'

'c1007_4aq2' 'c1007_7aq1' 'c1007_7aq2' 'c1007_7aq3' 'c1007_4aq3' 'c1007_4aq4'
'c1007_4aq5' 'c1007_4aq6' 'c1007_4aq7' 'c1007_4aq8' 'c1007_4aq9' 'c1007_4aq10'
'c1007_4aq11' 'c1007_4aq12' 'c1007_4aq13' 'c1007_4aq14' 'c1007_4aq15'
'c1007_4aq16' 'c1007_4aq17' 'c1007_4aq18' 'c1007_4aq19' 'c1007_4aq20'
'c1007_4aq21' 'c1007_4aq22' 'c1007_4aq23' 'h10_pers_income1' 'h10_pers_income2'
'h10_pers_income3' 'h10_pers_income4' 'h10_pers_income5'

In [7]:

```
welfare_m <- select(welfare, sex, birth, marriage, religion, income, code_job, code_region)
names(welfare_m)
```

'sex' 'birth' 'marriage' 'religion' 'income' 'code_job' 'code_region'

In [10]:

```
head(welfare_m,10)
```

sex	birth	marriage	religion	income	code_job	code_region
2	1936	2	2	NA	NA	1
2	1945	2	2	NA	NA	1
1	1948	2	2	120	942	1
1	1942	3	1	200	762	1
2	1923	2	1	NA	NA	1
1	1962	1	1	NA	530	1
2	1963	1	1	NA	NA	1
2	2003	0	1	NA	NA	1
1	1927	1	1	NA	NA	1
2	1934	1	1	NA	NA	1

미션 - 성별에 따른 월급 차이는 있을까?

- 변수 : 성별, 월급
- 성별, 월급 평균표 만들기
- 그래프 확인

In [12]:

```
str(welfare_m)
```

```
'data.frame':   16664 obs. of  7 variables:
 $ sex        : num  2 2 1 1 2 1 2 2 1 2 ...
 $ birth      : num  1936 1945 1948 1942 1923 ...
 $ marriage   : num  2 2 2 3 2 1 1 0 1 1 ...
 $ religion   : num  2 2 2 1 1 1 1 1 1 1 ...
 $ income     : num  NA NA 120 200 NA NA NA NA NA NA ...
 $ code_job   : num  NA NA 942 762 NA 530 NA NA NA NA ...
 $ code_region: num  1 1 1 1 1 1 1 1 1 1 ...
 - attr(*, "variable.labels")= Named chr  "가구 패널 ID(h10_id)" "가구생성차수(h10_i
nd)" "가구분리일련번호(h10_sn)" "패널간 가구머지 키변수(h10_merkey)" ...
  ..- attr(*, "names")= chr  "h10_id" "h10_ind" "h10_sn" "h10_merkey" ...
```

01 성별 검토

In [13]:

```
class(welfare_m$sex)
```

'numeric'

In [14]:

```
table(welfare_m$sex)
```

```
   1    2
7578 9086
```

- 1: 남자
- 2: 여자
- 9: 응답 없음.

만약 존재할 수 있으므로, 결측치 처리해야함.

In [15]:

```
welfare_m$sex <- ifelse(welfare_m$sex == 9, NA, welfare_m$sex)
table(is.na(welfare_m$sex))  # 결측치 확인
```

```
FALSE
16664
```

변수 1(남자), 2(여자)로 전처리

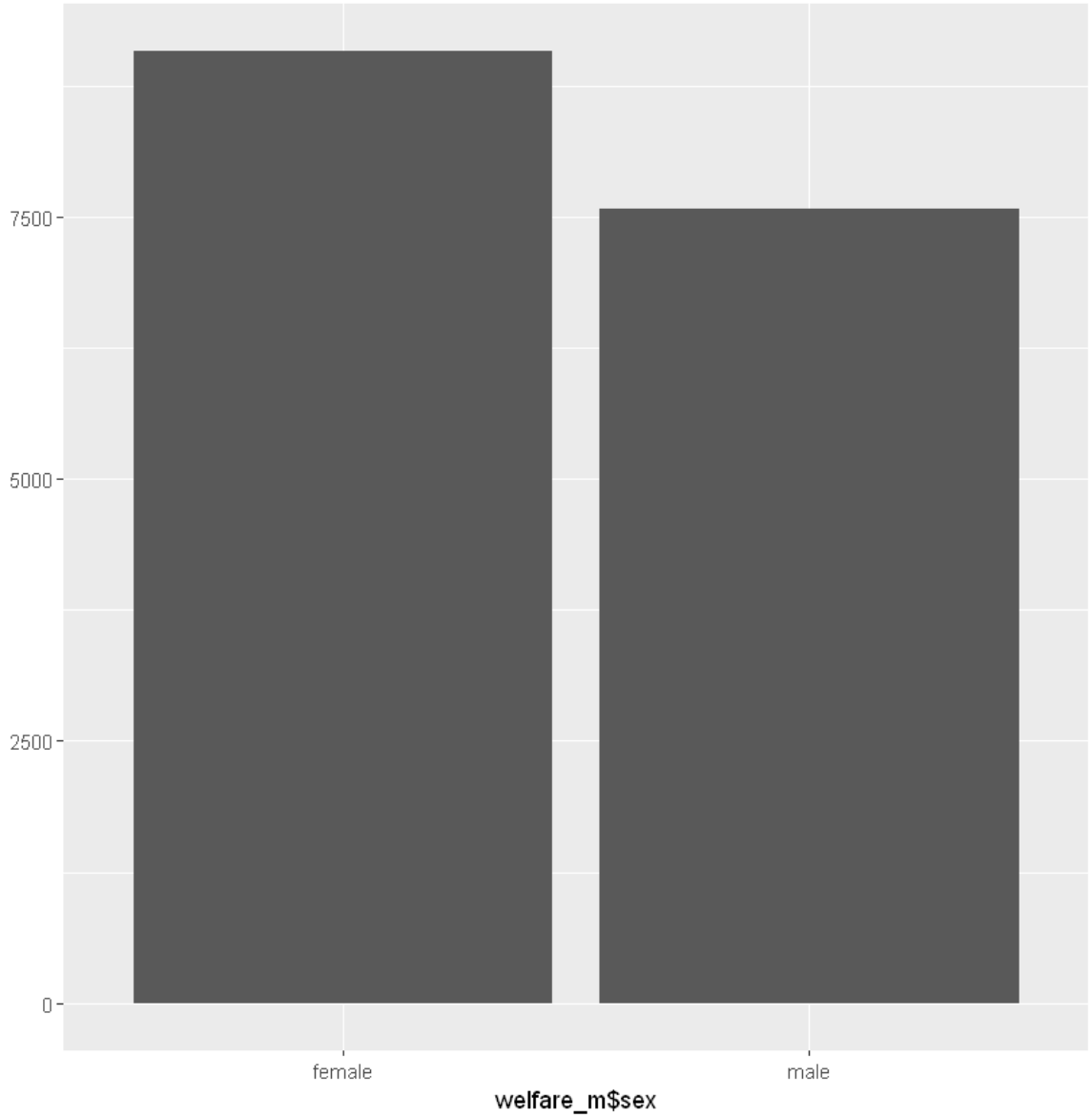
In [18]:

```
welfare_m$sex <- ifelse(welfare_m$sex == 1, "male", "female")
table(welfare_m$sex)
```

```
female   male
  9086   7578
```

In [19]:

```
qplot(welfare_m$sex)
```



02 월급 검토

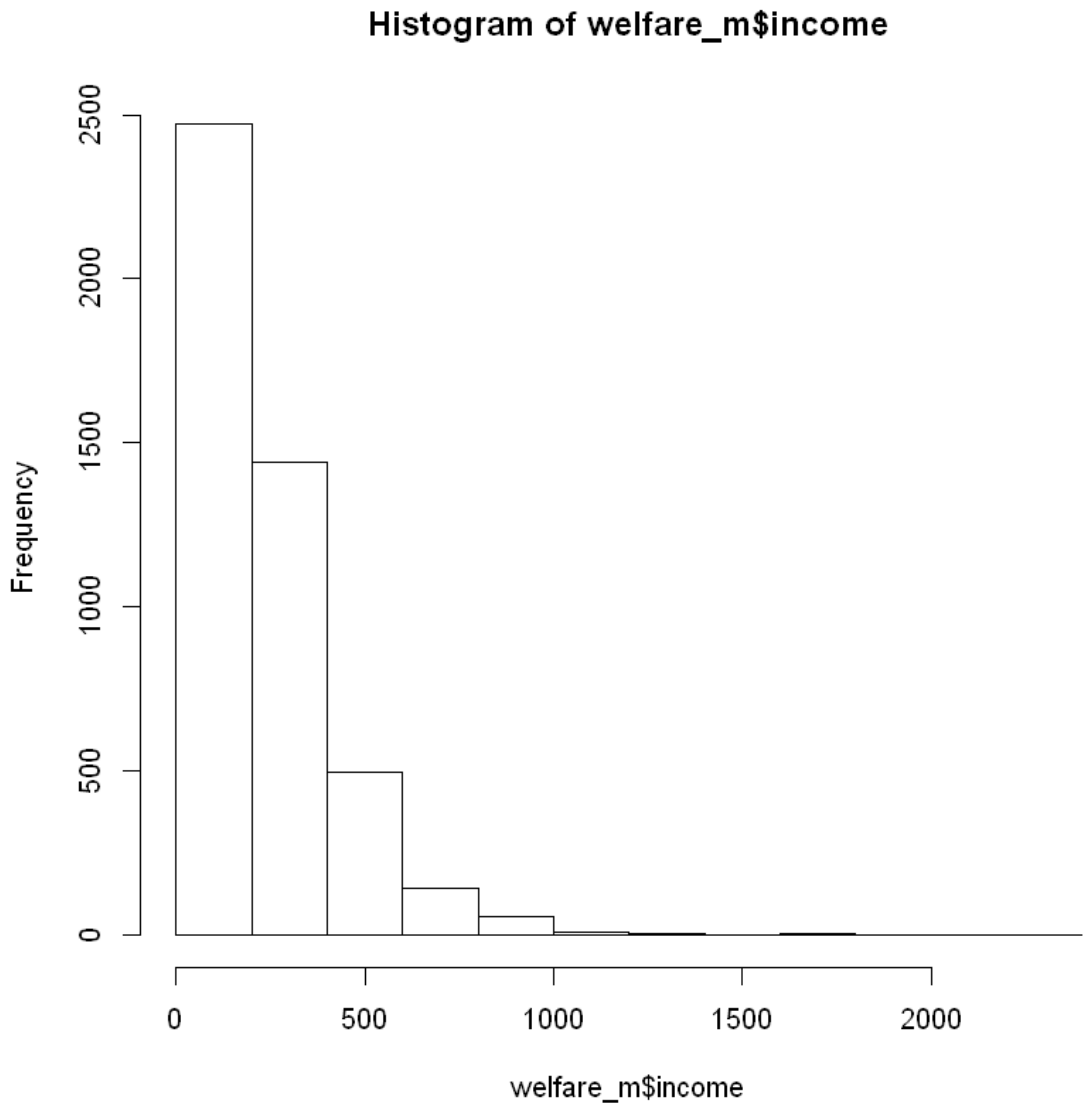
In [20]:

```
names(welfare_m)
```

'sex' 'birth' 'marriage' 'religion' 'income' 'code_job' 'code_region'

In [21]:

```
hist(welfare_m$income)
```



In [22]:

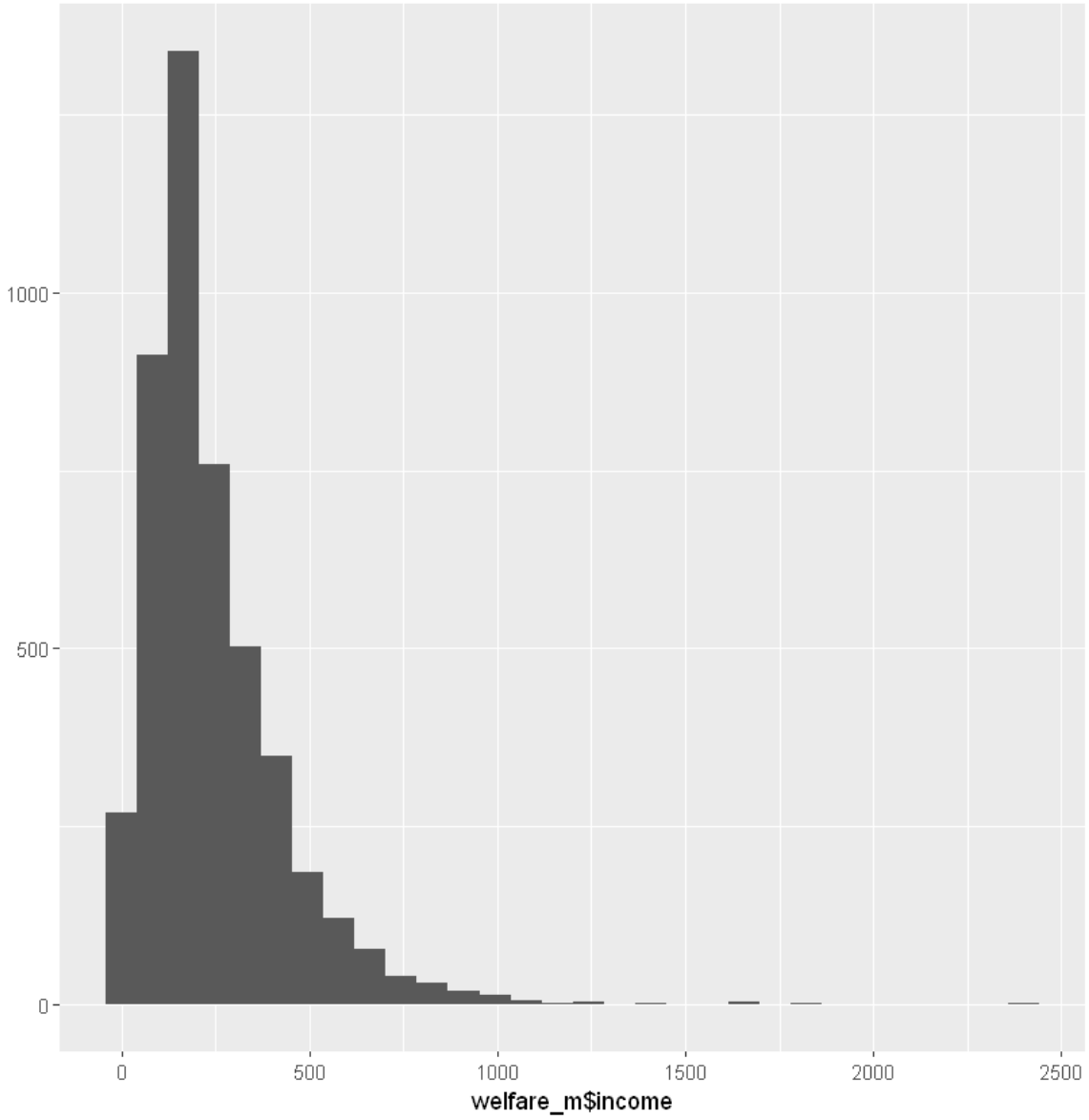
```
summary(welfare_m$income)
```

```
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's
    0.0   122.0   192.5   241.6   316.6  2400.0   12030
```

In [23]:

```
### 자세히 보자.
qplot(welfare_m$income)
```

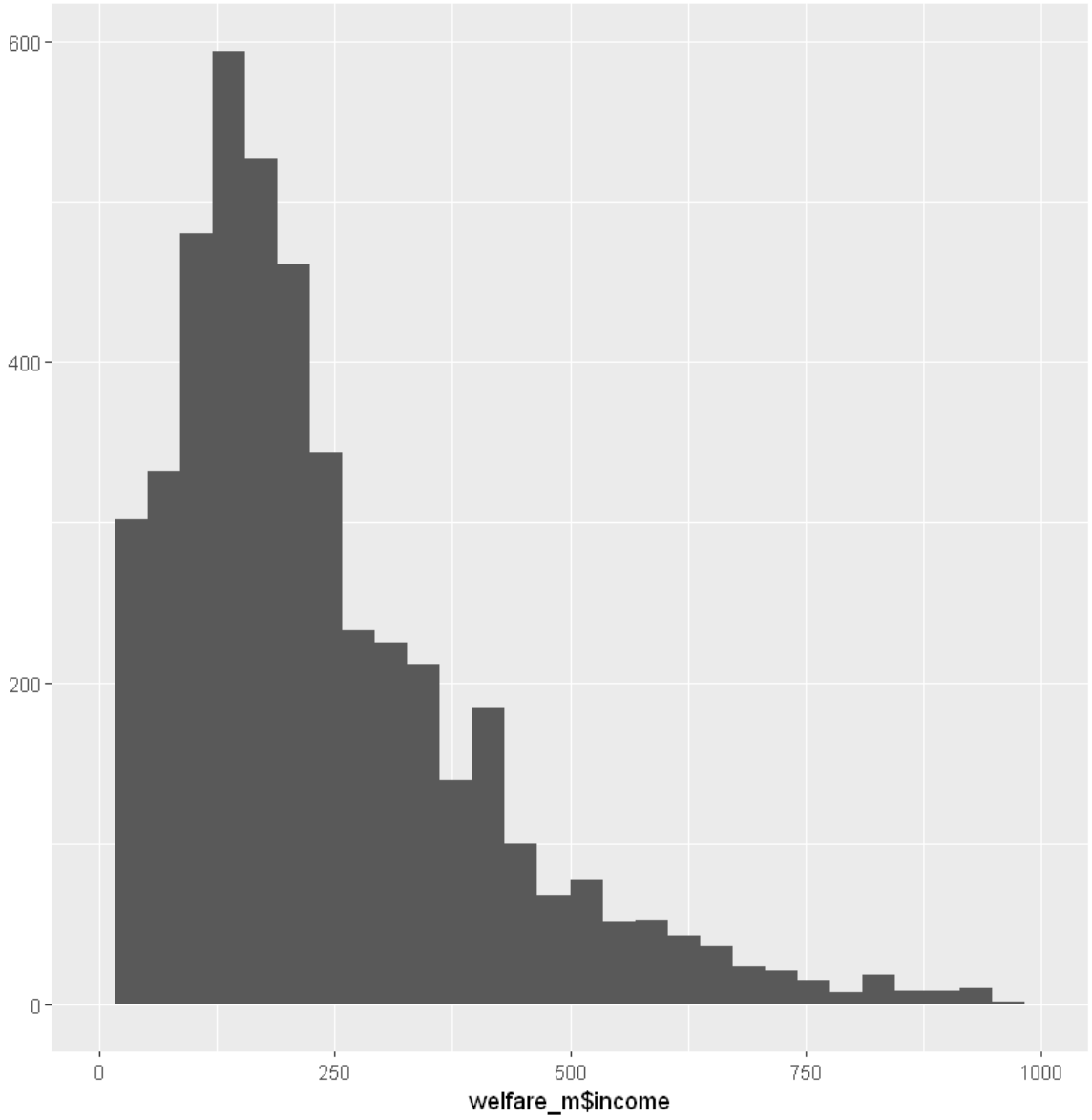
```
`stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
Warning message:
"Removed 12030 rows containing non-finite values (stat_bin)."
```



In [24]:

```
### 자세히 보자.
qplot(welfare_m$income) + xlim(0,1000)
```

```
`stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
Warning message:
"Removed 12051 rows containing non-finite values (stat_bin)."
```



NA를 전처리

In [25]:

```
summary(welfare_m$income)
```

```
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's
    0.0   122.0   192.5   241.6   316.6  2400.0   12030
```

In [26]:

```
### 모름/무응답 = 9999
### 범위 1~9998 이므로 0도 결측치 처리
welfare_m$income <- ifelse(welfare_m$income %in% c(0,9999), NA, welfare$income)
table(is.na(welfare_m$income))
```

```
FALSE  TRUE
 4620 12044
```

성별에 따른 월급 차이 분석

In [27]:

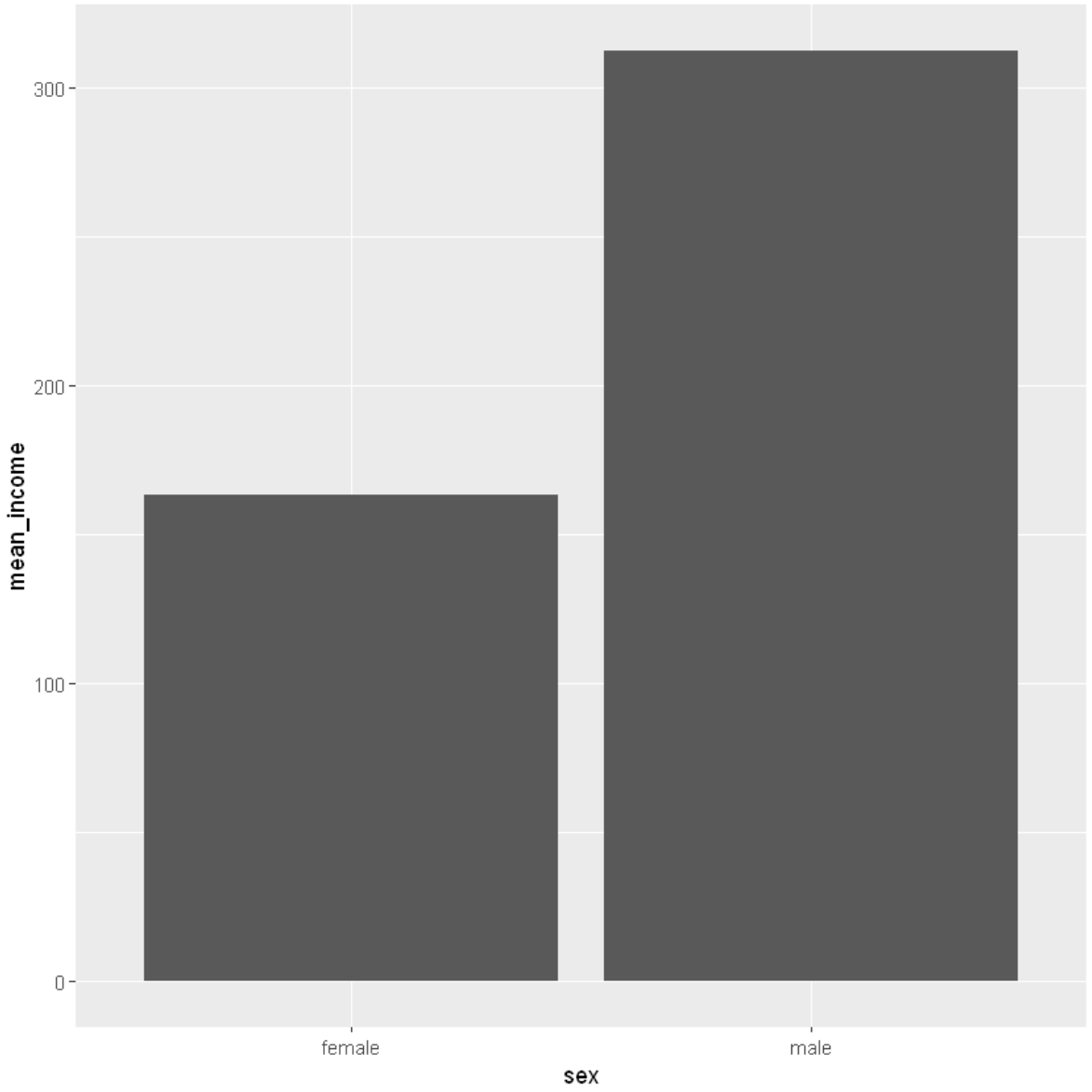
```
sex_income <- welfare_m %>% filter(!is.na(income)) %>%
                            group_by(sex) %>%
                            summarise(mean_income = mean(income))

sex_income
```

sex	mean_income
female	163.2471
male	312.2932

- 월급 평균은 남자가 312만원, 여자는 163만원으로 평균적으로 여성보다 남성의 월급이 약 150만원 많다.

In [28]:

```
ggplot(data = sex_income, aes(x=sex, y=mean_income)) + geom_col()
```



함수	설명
geom_point()	산점도
geom_col()	막대 그래프, X축, Y축을 모두 설정
geom_bar()	막대 그래프, X축만 설정, Y축은 해당 데이터의 수량
geom_line()	선 그래프
geom_boxplot()	박스 그래프

In [ ]:

```
```