Econ 613: Homework 1
Julian Sauvage

Exercise 1

Number of households surveyed in 2007: 10498

Number of households with marital status “Couple with kids” in 2005: 3374

Number of individuals surveyed in 2008: 25510

Number of individuals aged between 25 and 35 in 2016: 2765

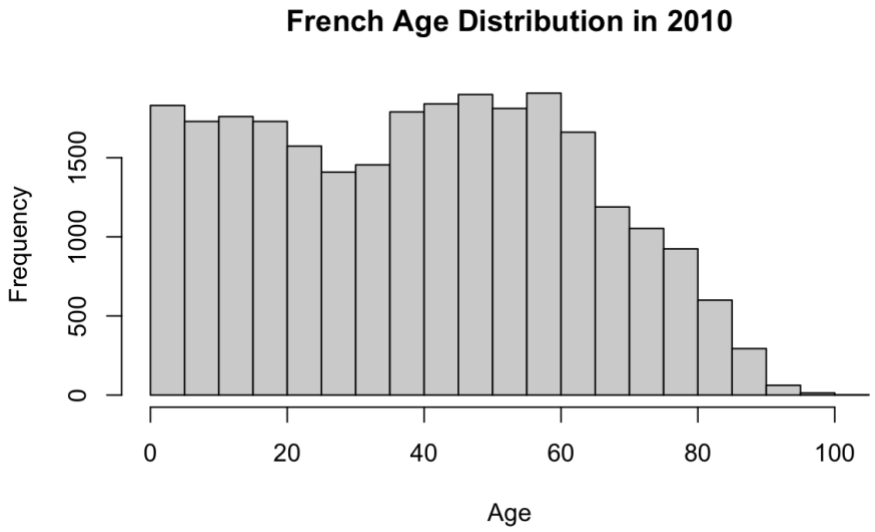
Cross-table gender/profession in 2009:

ind2009$profession	ind2009$gender Female	Male
0	11	19
11	30	57
12	8	19
13	29	78
21	63	213
22	65	114
23	8	48
31	68	98
33	85	107
34	184	142
35	50	59
37	179	260
38	78	368
42	258	110
43	437	117
44	1	2
45	153	95
46	410	340
47	82	429
48	22	215
52	782	169
53	27	182
54	584	98
55	353	101
56	696	74
62	64	443
63	35	520
64	29	246
65	19	159
67	147	237
68	120	177
69	40	82

Distribution of wages in 2005 and 2019. Report the mean, the standard deviation, the inter-decile ratio D9/D1 and the Gini coefficient:
2005 – Mean Wage: 22,443 Euros, SD: 18,076 Euros, D9/D1 ratio: 8.896, Gini: 0.667
2019 - Mean Wage: 27,578 Euros, SD: 25,107 Euros, D9/D1 ratio: 13.86, Gini: 0.665

While average wages have increased, and Gini has fallen marginally, the wage difference between low and high earners has increased significantly in the last 15 years.

Distribution of age in 2010. Plot a histogram. Is there any difference between men and women?



The age distributions by gender are largely similar, the only apparent different is that there is a steeper decline in men above the age of 60, relative to that of women. This is particularly true when we examine individuals above the age of 80; there are more women than men for every age above 80.

Number of individuals in Paris in 2011: 3514

Exercise 2

Variables that are simultaneously present in the individual and household datasets: “idmen” and “year”

Number of households in which there are more than four family members: 67,069 households

Number of households in which at least one member is unemployed: 139,312 households

Number of households in which at least two members are of the same profession:

Number of individuals in the panel that are from household-Couple with kids: 209,382 individuals

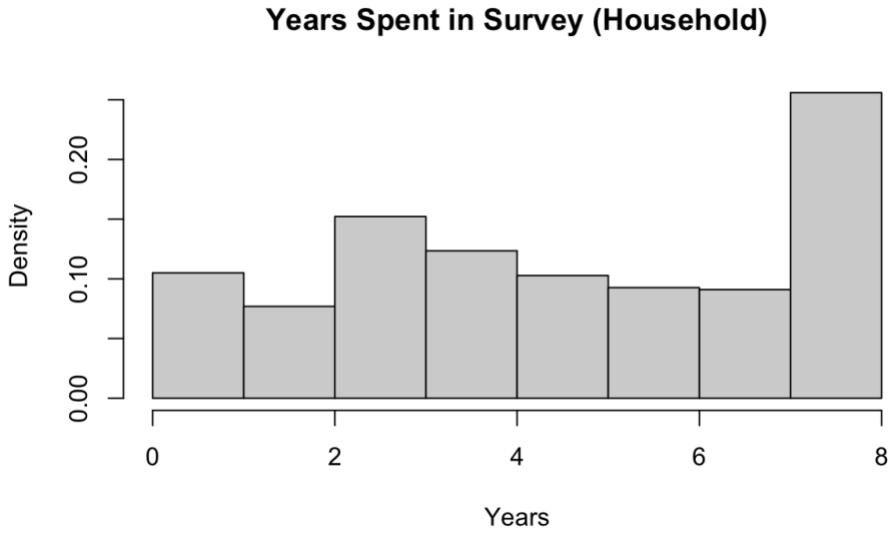
Number of individuals in the panel that are from Paris: 51,904 individuals

Find the household with the most number of family members. Report its idmen: Largest household has 14 members, idmen: 2207811124040100

Number of households present in 2010 and 2011: 17,698 households

Exercise 3

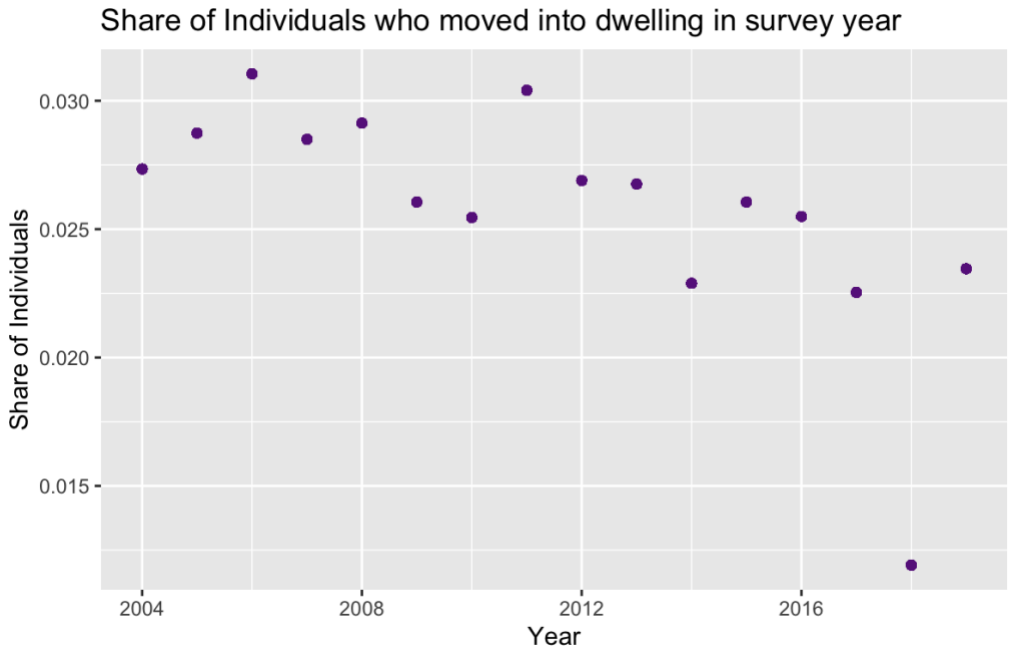
Distribution of time spent in survey by households:



Based on datent, Percent of households that moved into dwelling in year of survey
First 10 rows of result:

```
> head(tmpdat, 10)
# A tibble: 10 × 10
# Groups:   year [1]
       X   idmen  year datent myear mstatus          move location ident   perc
   <int>   <dbl> <int>  <int> <int> <chr>           <int> <chr>    <dbl>  <dbl>
 1     1 1.20e15  2004   2000  2000 Single             NA Paris        0 0.0318
 2     2 1.20e15  2004   2001  2001 Single Parent      NA Paris        0 0.0318
 3     3 1.20e15  2004   2000  2000 Couple, No kids    NA Paris        0 0.0318
 4     4 1.20e15  2004   1957  1957 Single             NA Paris        0 0.0318
 5     5 1.20e15  2004   2001  2001 Couple, No kids    NA Paris        0 0.0318
 6     6 1.20e15  2004   1990  1990 Single Parent      NA Paris        0 0.0318
 7     7 1.20e15  2004   2000  2000 Couple, No kids    NA Paris        0 0.0318
 8     8 1.20e15  2004   1948  1988 Other              NA Rural        0 0.0318
 9     9 1.20e15  2004   1979  1979 Single             NA Rural        0 0.0318
10    10 1.20e15  2004   1984  1981 Other              NA Rural        0 0.0318
```

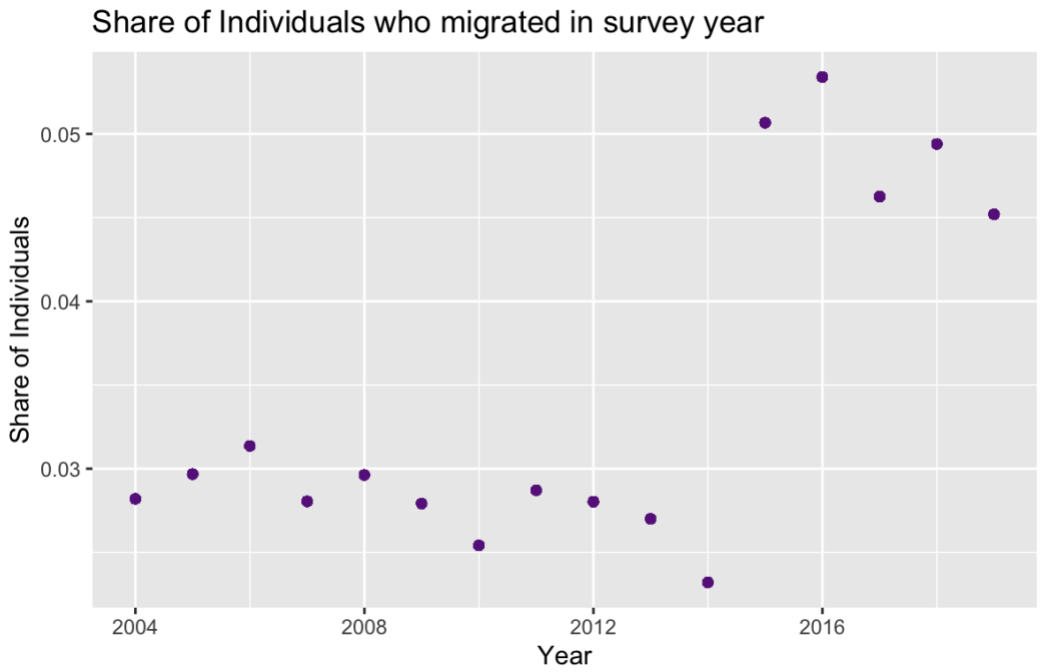
Plot the share of individuals in this situation:



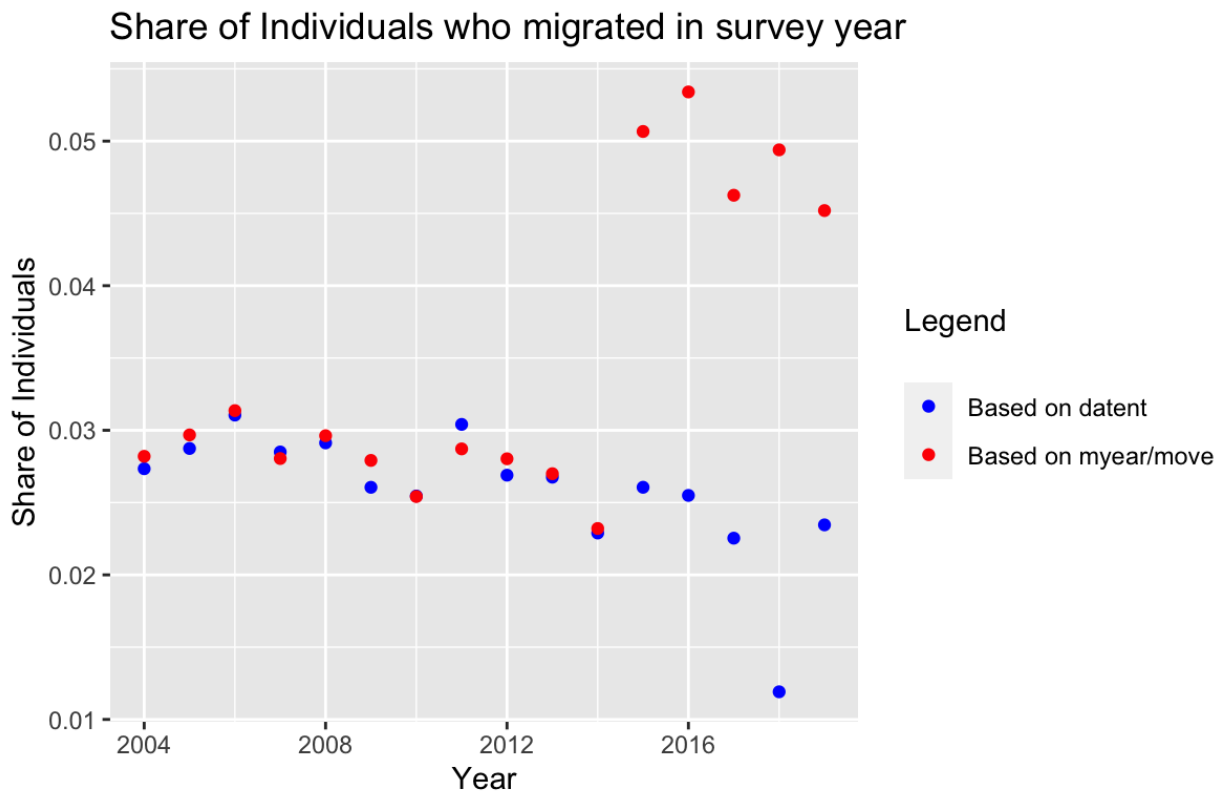
Based on myear and move, identify whether or not household migrated at the year of survey
First 10 rows:

```
> head(tmpdat, 10)
# A tibble: 10 × 10
# Groups:   year [1]
       X   idmen  year datent myear mstatus          move location ident   perc
   <int>   <dbl> <int>  <int> <int> <chr>           <int> <chr>    <dbl>  <dbl>
 1     1 1.20e15  2004   2000  2000 Single             NA Paris        0 0.0326
 2     2 1.20e15  2004   2001  2001 Single Parent      NA Paris        0 0.0326
 3     3 1.20e15  2004   2000  2000 Couple, No kids    NA Paris        0 0.0326
 4     4 1.20e15  2004   1957  1957 Single             NA Paris        0 0.0326
 5     5 1.20e15  2004   2001  2001 Couple, No kids    NA Paris        0 0.0326
 6     6 1.20e15  2004   1990  1990 Single Parent      NA Paris        0 0.0326
 7     7 1.20e15  2004   2000  2000 Couple, No kids    NA Paris        0 0.0326
 8     8 1.20e15  2004   1948  1988 Other              NA Rural        0 0.0326
 9     9 1.20e15  2004   1979  1979 Single             NA Rural        0 0.0326
10    10 1.20e15  2004   1984  1981 Other              NA Rural        0 0.0326
```

Plot the share of individuals in this situation:



Combined Graph



Exercise 4

year	attritionRate
2004	NA
2005	0.1276644
2006	0.1727239
2007	0.1548115
2008	0.2038831
2009	0.1803998
2010	0.1589942
2011	0.1727413
2012	0.1435115
2013	0.2351230
2014	0.1676090
2015	0.1867324
2016	0.1848821
2017	0.2206627
2018	0.2024250
2019	0.2077091