

Multiple Regression in R

We are interested in whether house prices in 250 regions of the UK can be predicted by:

- (a) Population size
- (b) Crime rate (per 10,000 people)
- (c) Average age of people in the region
- (d) Average household income in the region.

Including our predictor and a column identifying our Regions, our datasets consists of 6 variables.

```
> str(data)
Classes 'tbl_df', 'tbl' and 'data.frame': 250 obs. of 6 variables:
 $ Region      : num  1 2 3 4 5 6 7 8 9 10 ...
 $ House_price : num  193735 201836 191643 215952 203295 ...
 $ Population  : num  49004 48307 50379 53664 45481 ...
 $ Crime       : num  14 25 19 17 22 32 21 21 20 25 ...
 $ Average_age : num  72.7 78.1 71.4 72.1 76.1 ...
 $ Household_income: num  20843 19130 20411 16863 19964 ...
```

```
> head(data)
# A tibble: 6 x 6
   Region House_price Population Crime Average_age Household_income
   <dbl>      <dbl>      <dbl> <dbl>      <dbl>      <dbl>
1     1     193735     49004     14       72.7     20843.
2     2     201836     48307     25       78.1     19130.
3     3     191643     50379     19       71.4     20411.
4     4     215952     53664     17       72.1     16863.
5     5     203295     45481     22       76.1     19964.
6     6     191795     51582     32       81.2     20207.
```