- Extract raw datasets and do data cleaning
- Create table with AreaName+IndicatorID+Age+Sex+Value (ds1)
- Create wide column, combine IndicatorID+Age+Sex as column names (analysis)
- Remove variable with >7 NA values
- Define 'Suicide' = 'Suicide: age-standardised rate per 100,000 population (3 year average) 10+ yrs_Persons' column only/remove columns containing 'suicide'.
- Rename columns with shorter name (analysis1)
- Remove area name (analysis_final)
- Get the dataprocessing.R

The resulted tables:

The independent variables (31) and dependent variable (suicide_rate) used in dataprocessing.r are:

1. analysis1 table (with area name)
2. analysis_final_new table (without area name)

- Create ggplot for suicide rate based on gender (male and female) and year + area
- Create ggplot for suicide rate based on gender (male, female and persons) and year
- Create correlation matrices between variables
- Calculate fitted value → input the variables value to the equations
- Create ggplot for fitted value to actual value of suicide rate
- Calculate MAE, RMSE and Rsquared to see the best model
- Create prediction model and draw the plot
- Calculate and sort the most significance variables on suicide rate
- Produced importance variables on suicide rate

- Create table with column names: Area Name, Indicator ID 1, Indicator ID 2, etc
- Create table with column names: Area ID, Area Name, Indicator ID 1, Indicator ID 2, etc
- Create Indicator ID-Indicator table for reference

Produced m9 table contains of: AreaID, AreaName, IndicatorID (34)

Work on the draft of the final paper

Conduct factor analysis, the results are in the table shown below

Results:

Multiple Linear Regression (coefficients are sorted)

Indicator	Coefficient
`Estimated prevalence of common mental disorders: % of population aged 65 & over_65+ yrs_Persons`	12.08689416
`Mental Health: QOF prevalence (all ages)_All ages_Persons`	4.378776439
`Adults in treatment at specialist drug misuse services: rate per 1000 population_18+ yrs_Persons`	1.47072257
`Marital breakup: % of adults_18+ yrs_Persons`	1.428965251
`Adults in treatment at specialist alcohol misuse services: rate per 1000 population_18+ yrs_Persons`	1.184669242
`Older people living alone: % of households occupied by a single person aged 65 or over_65+ yrs_Persons`	0.840570556
`Emergency Hospital Admissions for Intentional Self-Harm_All ages_Persons`	0.386866829
`Unemployment (model-based)_16+ yrs_Persons`	0.262807421
`Long-term health problem or disability: % of population_All ages_Persons`	0.246600386

`Social Isolation: percentage of adult carers who have as much social contact as they would like_18+ yrs_Persons`	0.123091282
`Successful completion of alcohol treatment_18+ yrs_Persons`	0.093595094
`Successful completion of drug treatment - non-opiate users_18+ yrs_Persons`	0.023150345
`Admission episodes for alcohol-related conditions (Broad): Old Method_All ages_Persons`	0.021063819
`Children in care_<18 yrs_Persons`	0.010110135
`Long term claimants of Jobseeker's Allowance_16-64 yrs_Persons`	0.006855004
`Children leaving care: rate per 10,000 children aged under 18_<18 yrs_Persons`	0.006278076
`Social Isolation: percentage of adult carers who have as much social contact as they would like_65+ yrs_Persons`	0.005628362
`Admission episodes for alcohol-related conditions (Broad): Old Method_All ages_Male`	-0.012008835
`Admission episodes for alcohol-related conditions (Broad): Old Method_All ages_Female`	-0.013099296
`Social Isolation: percentage of adult carers who have as much social contact as they would like_18-64 yrs_Persons`	-0.020703138
`Children in the youth justice system (10-17 yrs)_10-17 yrs_Persons`	-0.028814498
`Social Isolation: percentage of adult carers who have as much social contact as they would like_All ages_Persons`	-0.04253679
`Self-reported wellbeing - people with a high anxiety score_16+ yrs_Persons`	-0.107666507
`Emergency Hospital Admissions for Intentional Self-Harm_All ages_Male`	-0.169290448
`Estimated prevalence of opiate and/or crack cocaine use_15-64 yrs_Persons`	-0.171873474
`Successful completion of drug treatment - opiate users_18+ yrs_Persons`	-0.18353718
`Emergency Hospital Admissions for Intentional Self-Harm_All ages_Female`	-0.187550606
`Social Isolation: percentage of adult social care users who have as much social contact as they would like_18+ yrs_Persons`	-0.323564467
`Depression: Recorded prevalence (aged 18+)_18+ yrs_Persons`	-0.336023524
`People living alone: % of all usual residents in households occupied by a single person_All ages_Persons`	-0.534240784
`Estimated prevalence of common mental disorders: % of population aged 16 & over_16+ yrs_Persons`	-7.385760068

Linear Regression analyse one-by-one in initial order

No	Indicator ID	Indicator name	Coefficient	p-Value	Sig. <0.05
1		`Admission episodes for alcohol-related conditions (Broad): Old Method_All ages_Female`	0.004578	0.03852	Yes
2		`Admission episodes for alcohol-related conditions (Broad): Old Method_All ages_Male`	8.319e-04	0.4518	No
3		`Admission episodes for alcohol-related conditions (Broad): Old Method_All ages_Persons`	0.002052	0.1795	No
4		`Adults in treatment at specialist alcohol misuse services: rate per 1000 population_18+ yrs_Persons`	6.263	7.421e-09	Yes
5		`Adults in treatment at specialist drug misuse services: rate per 1000 population_18+ yrs_Persons`	1.5611	2.083e-06	Yes
6		`Children in care_<18 yrs_Persons`	0.1252	3.055e-09	Yes
7		`Children in the youth justice system (10-17 yrs)_10-17 yrs_Persons`	-0.1415	0.6609	No
8		`Children leaving care: rate per 10,000 children aged under 18_<18 yrs_Persons`	0.04555	0.3887	No
9		`Depression: Recorded prevalence (aged 18+)_18+ yrs_Persons`	1.6814	9.262e-10	Yes
10		`Emergency Hospital Admissions for Intentional Self-Harm_All ages_Female`	0.041093	6.673e-11	Yes
11		`Emergency Hospital Admissions for Intentional Self-Harm_All ages_Male`	0.071827	7.845e-14	Yes
12		`Emergency Hospital Admissions for Intentional Self-Harm_All ages_Persons`	0.055727	1.14e-12	Yes
13		`Estimated prevalence of common mental disorders: % of population aged 16 & over_16+ yrs_Persons`	-0.1233	0.6682	No
14		`Estimated prevalence of common mental disorders: % of population aged 65 & over_65+ yrs_Persons`	0.664	0.6639	No
15		`Estimated prevalence of opiate and/or crack cocaine use_15-64 yrs_Persons`	0.6320	0.001504	Yes
16		`Long-term health problem or disability: % of population_All ages_Persons`	1.7518	< 2.2e-16	Yes
17		`Long term claimants of Jobseeker's Allowance_16-64 yrs_Persons`	0.7403	0.04014	Yes
18		`Marital breakup: % of adults_18+ yrs_Persons`	7.377	1.21e-11	Yes

19		`Mental Health: QOF prevalence (all ages)_All ages_Persons`	-0.2132	0.9572	No
20		`People living alone: % of all usual residents in households occupied by a single person_All ages_Persons`	8.266	8.593e-14	Yes
21		`Self-reported wellbeing - people with a high anxiety score_16+ yrs_Persons`	1.0327	0.002597	Yes
22		`Social Isolation: percentage of adult carers who have as much social contact as they would like_18-64 yrs_Persons`	-0.006432	0.9812	No
23		`Social Isolation: percentage of adult carers who have as much social contact as they would like_18+ yrs_Persons`	0.387	0.3868	No
24		`Social Isolation: percentage of adult carers who have as much social contact as they would like_18+ yrs_Persons`	1.899	0.05956	No
25		`Social Isolation: percentage of adult carers who have as much social contact as they would like_65+ yrs_Persons`	2.773	0.006297	Yes
26		`Social Isolation: percentage of adult carers who have as much social contact as they would like_All ages_Persons`	0.2580	0.01289	Yes
27		`Social Isolation: percentage of adult social care users who have as much social contact as they would like_18+ yrs_Persons`	0.5344	0.002105	Yes
28		`Successful completion of alcohol treatment_18+ yrs_Persons`	-0.1519	0.093	No
29		`Successful completion of drug treatment - non-opiate users_18+ yrs_Persons`	-0.13108	0.106	No
30		`Successful completion of drug treatment - opiate users_18+ yrs_Persons`	-1.0901	0.006836	Yes
31		`Unemployment (model-based)_16+ yrs_Persons`	0.2338	0.6304	No

Standardized Multiple Linear Regression (coefficients are sorted)

Variables	Coefficient
`Emergency Hospital Admissions for Intentional Self-Harm_All ages_Persons`	3.794797
`Estimated prevalence of common mental disorders: % of population aged 65 & over_65+ yrs_Persons`	2.0584
`Admission episodes for alcohol-related conditions (Broad): Old Method_All ages_Persons`	1.150939
`Adults in treatment at specialist drug misuse services: rate per 1000 population_18+ yrs_Persons`	0.360171
`Marital breakup: % of adults_18+ yrs_Persons`	0.194371
`Older people living alone: % of households occupied by a single person aged 65 or over_65+ yrs_Persons`	0.189303

`Mental Health: QOF prevalence (all ages)_All ages_Persons`	0.09231
`Social Isolation: percentage of adult carers who have as much social contact as they would like_18+ yrs_Persons`	0.09116
`Long-term health problem or disability: % of population_All ages_Persons`	0.088713
`Adults in treatment at specialist alcohol misuse services: rate per 1000 population_18+ yrs_Persons`	0.086481
`Successful completion of alcohol treatment_18+ yrs_Persons`	0.086268
`Unemployment (model-based)_16+ yrs_Persons`	0.04529
`Children in care_<18 yrs_Persons`	0.037748
`Successful completion of drug treatment - non-opiate users_18+ yrs_Persons`	0.023808
`Children leaving care: rate per 10,000 children aged under 18_<18 yrs_Persons`	0.009939
`Social Isolation: percentage of adult carers who have as much social contact as they would like_65+ yrs_Persons`	0.004413
`Long term claimants of Jobseeker's Allowance_16-64 yrs_Persons`	0.00158
`Children in the youth justice system (10-17 yrs)_10-17 yrs_Persons`	-0.00748
`Social Isolation: percentage of adult carers who have as much social contact as they would like_18-64 yrs_Persons`	-0.01703
`Self-reported wellbeing - people with a high anxiety score_16+ yrs_Persons`	-0.03312
`Social Isolation: percentage of adult carers who have as much social contact as they would like_All ages_Persons`	-0.03398
`Successful completion of drug treatment - opiate users_18+ yrs_Persons`	-0.03766
`Estimated prevalence of opiate and/or crack cocaine use_15-64 yrs_Persons`	-0.07104
`Depression: Recorded prevalence (aged 18+)_18+ yrs_Persons`	-0.0961
`People living alone: % of all usual residents in households occupied by a single person_All ages_Persons`	-0.12848
`Social Isolation: percentage of adult social care users who have as much social contact as they would like_18+ yrs_Persons`	-0.1534
`Admission episodes for alcohol-related conditions (Broad): Old Method_All ages_Female`	-0.49233
`Admission episodes for alcohol-related conditions (Broad): Old Method_All ages_Male`	-0.90897
`Emergency Hospital Admissions for Intentional Self-Harm_All ages_Male`	-1.34197
`Estimated prevalence of common mental disorders: % of population aged 16 & over_16+ yrs_Persons`	-2.15004
`Emergency Hospital Admissions for Intentional Self-Harm_All ages_Female`	-2.32023

Correlation matrix

On excel files

1. cor_new_marked (with Suicide_Rate variable)
2. cor_factor_new (without Suicide_Rate variable)

Conclusion on top 6 most influenced variables based on 4 statistics analyses (blank: `Estimated prevalence of common mental disorders: % of population aged 65 & over_65+ yrs_Persons`)

No.	Multiple Linear Regression	One-by-one Linear Regression	Standardized Multiple Linear Regression
1	`Estimated prevalence of common mental disorders: % of population aged 65 & over_65+ yrs_Persons`	`People living alone: % of all usual residents in households occupied by a single person_All ages_Persons`	`Emergency Hospital Admissions for Intentional Self-Harm_All ages_Persons`
2	`Mental Health: QOF prevalence (all ages)_All ages_Persons`	`Marital breakup: % of adults_18+ yrs_Persons`	
3	`Adults in treatment at specialist drug misuse services: rate per 1000 population_18+ yrs_Persons`	`Adults in treatment at specialist alcohol misuse services: rate per 1000 population_18+ yrs_Persons`	`Admission episodes for alcohol-related conditions (Broad): Old Method_All ages_Persons`
4	`Marital breakup: % of adults_18+ yrs_Persons`	`Social Isolation: percentage of adult carers who have as much social contact as they would like_65+ yrs_Persons`	`Adults in treatment at specialist drug misuse services: rate per 1000 population_18+ yrs_Persons`
5	`Adults in treatment at specialist alcohol misuse services: rate per 1000 population_18+ yrs_Persons`	`Long-term health problem or disability: % of population_All ages_Persons`	`Marital breakup: % of adults_18+ yrs_Persons`
6	`Older people living alone: % of households occupied by a single person aged 65 or over_65+ yrs_Persons`	`Depression: Recorded prevalence (aged 18+)_18+ yrs_Persons`	`Older people living alone: % of households occupied by a single person aged 65 or over_65+ yrs_Persons`

Based on Conclusion table above, there are 2 variables consistently stated in 3 analyses:

1. `Adults in treatment at specialist drug misuse services: rate per 1000 population_18+ yrs_Persons`
2. `Marital breakup: % of adults_18+ yrs_Persons`
3. `Adults in treatment at specialist alcohol misuse services: rate per 1000 population_18+ yrs_Persons`

And 2 variables consistently stated in 2 analyses:

1. `Estimated prevalence of common mental disorders: % of population aged 65 & over_65+ yrs_Persons
2. `Older people living alone: % of households occupied by a single person aged 65 or over_65+ yrs_Persons`

Week 10 (15 November – 21 November)

Factor Analysis result:

There are 7 factors consist of 31 variables:

```
$Factor1
 [1] "Children in care_<18 yrs_Persons"
 [2] "Children leaving care: rate per 10,000 children aged under 18_<18
yrs_Persons"
 [3] "Children in the youth justice system (10-17 yrs)_10-17 yrs_Persons"
 [4] "Unemployment (model-based)_16+ yrs_Persons"
 [5] "Long term claimants of Jobseeker's Allowance_16-64 yrs_Persons"
 [6] "Adults in treatment at specialist drug misuse services: rate per
1000 population_18+ yrs_Persons"
 [7] "Mental Health: QOF prevalence (all ages)_All ages_Persons"
 [8] "Estimated prevalence of opiate and/or crack cocaine use_15-64
yrs_Persons"
 [9] "Estimated prevalence of common mental disorders: % of population
aged 65 & over_65+ yrs_Persons"
[10] "Estimated prevalence of common mental disorders: % of population
aged 16 & over_16+ yrs_Persons"

$Factor2
[1] "Social Isolation: percentage of adult social care users who have as
much social contact as they would like_18+ yrs_Persons"
```

[2] "Depression: Recorded prevalence (aged 18+)_18+ yrs_Persons"
[3] "Marital breakup: % of adults_18+ yrs_Persons"
[4] "Long-term health problem or disability: % of population_All
ages_Persons"
[5] "Older people living alone: % of households occupied by a single
person aged 65 or over_65+ yrs_Persons"
[6] "Adults in treatment at specialist alcohol misuse services: rate per
1000 population_18+ yrs_Persons"

$Factor3
[1] "Emergency Hospital Admissions for Intentional Self-Harm_All
ages_Persons"
[2] "Emergency Hospital Admissions for Intentional Self-Harm_All
ages_Female"
[3] "Emergency Hospital Admissions for Intentional Self-Harm_All
ages_Male"

$Factor4
[1] "Social Isolation: percentage of adult carers who have as much social
contact as they would like_18+ yrs_Persons"
[2] "Social Isolation: percentage of adult carers who have as much social
contact as they would like_All ages_Persons"
[3] "Social Isolation: percentage of adult carers who have as much social
contact as they would like_18-64 yrs_Persons"
[4] "Social Isolation: percentage of adult carers who have as much social
contact as they would like_65+ yrs_Persons"

$Factor5
[1] "Admission episodes for alcohol-related conditions (Broad): Old
Method_All ages_Female"
[2] "Admission episodes for alcohol-related conditions (Broad): Old
Method_All ages_Persons"
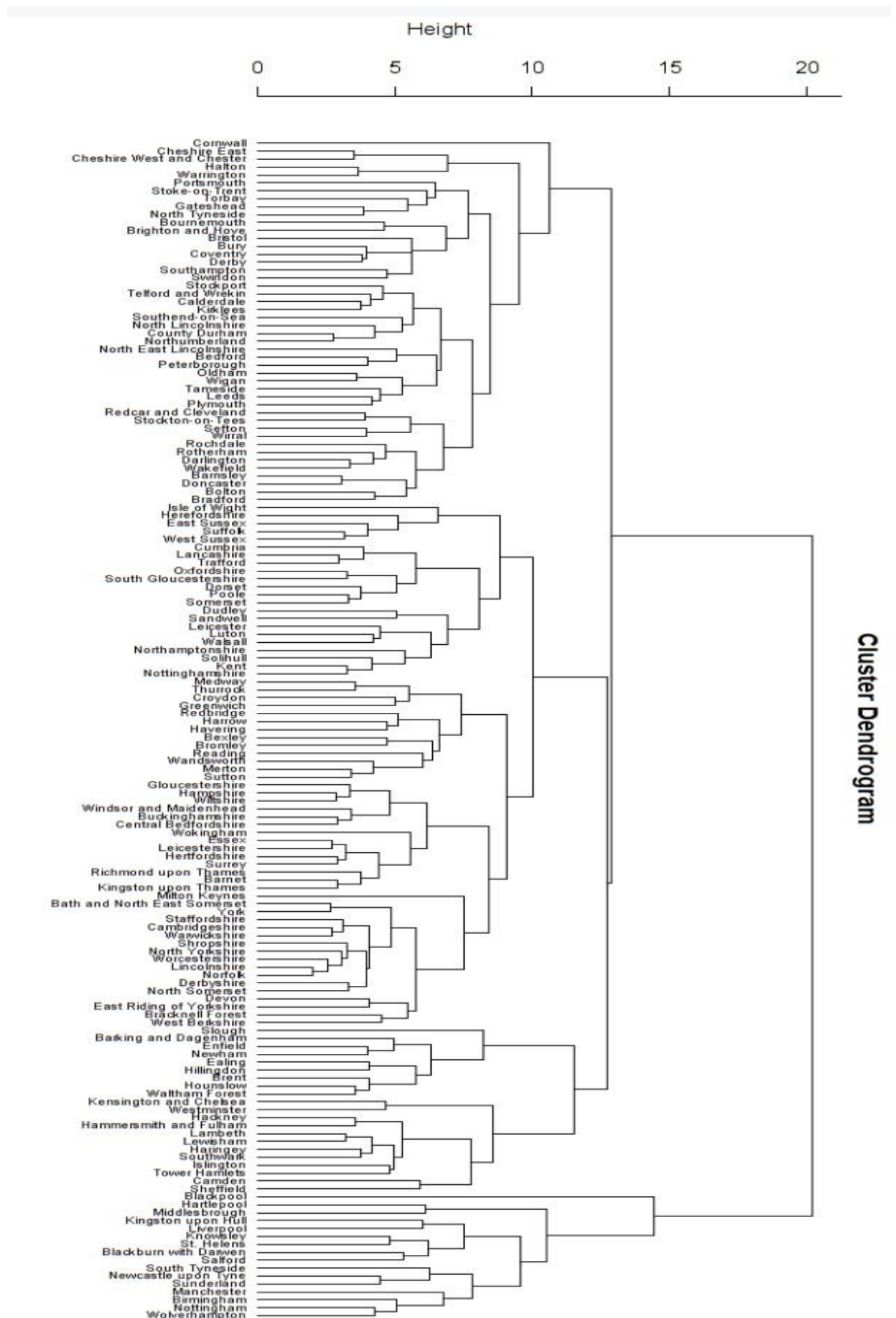[3] "Admission episodes for alcohol-related conditions (Broad): Old
Method_All ages_Male"

$Factor6
[1] "Successful completion of alcohol treatment_18+ yrs_Persons"
[2] "Successful completion of drug treatment - non-opiate users_18+
yrs_Persons"
[3] "Successful completion of drug treatment - opiate users_18+
yrs_Persons"

$Factor7
[1] "Self-reported wellbeing - people with a high anxiety score_16+
yrs_Persons"
[2] "People living alone: % of all usual residents in households occupied
by a single person_All ages_Persons"

Hierarchical clustering

Cluster Dendrogram

Hierarchical clustering result:

Areas is divided to 3-8 clusters in England. Determining the number of clusters is a hard task for humans to perform (unless the data are low-dimensional). Not only that, it's just as hard to explain what it is we're looking for.