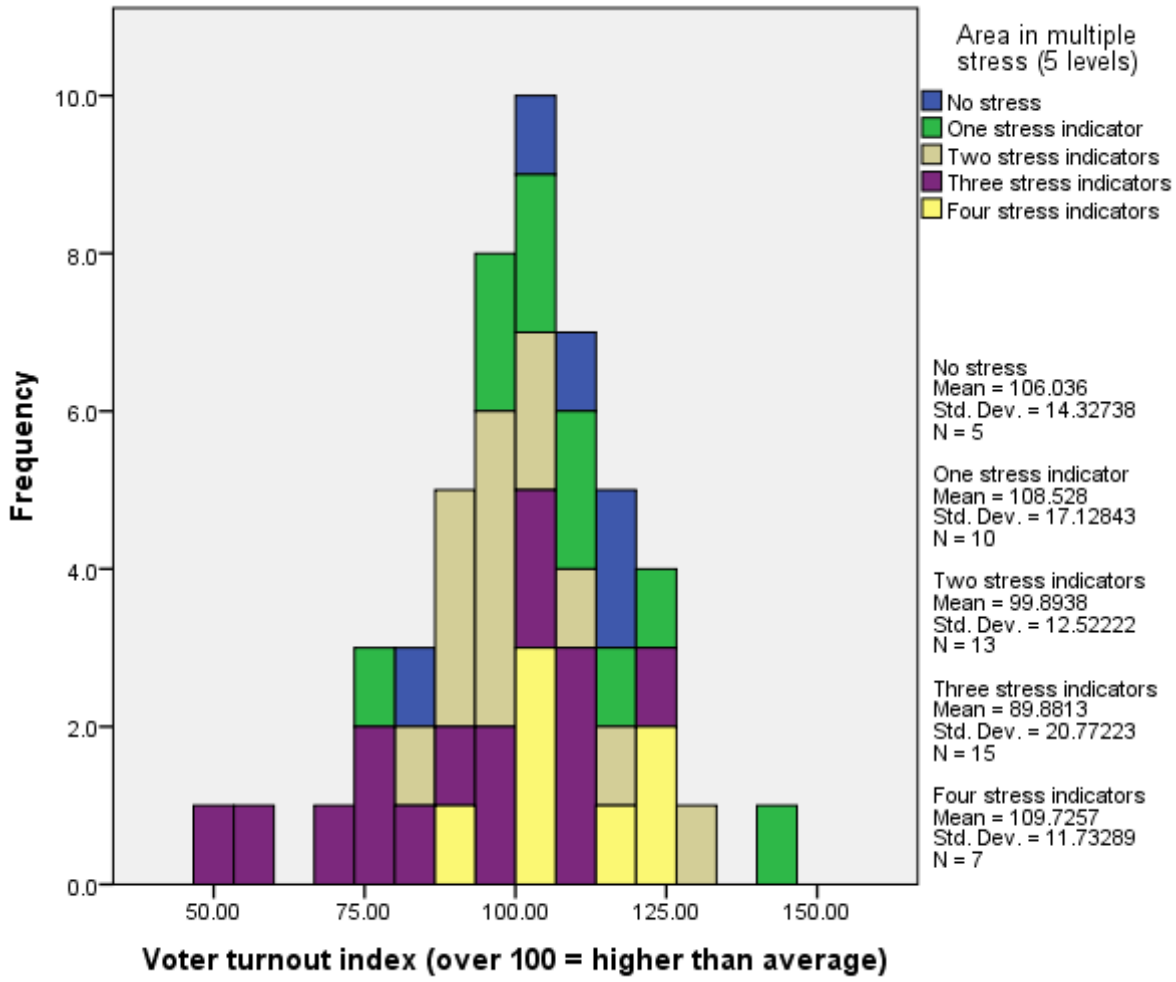
Exercises completed using SPSS, ArcGIS, GAM, Getis Ord Gi*

1. Descriptive statistics for the Voter Turnout Index and for one other variable

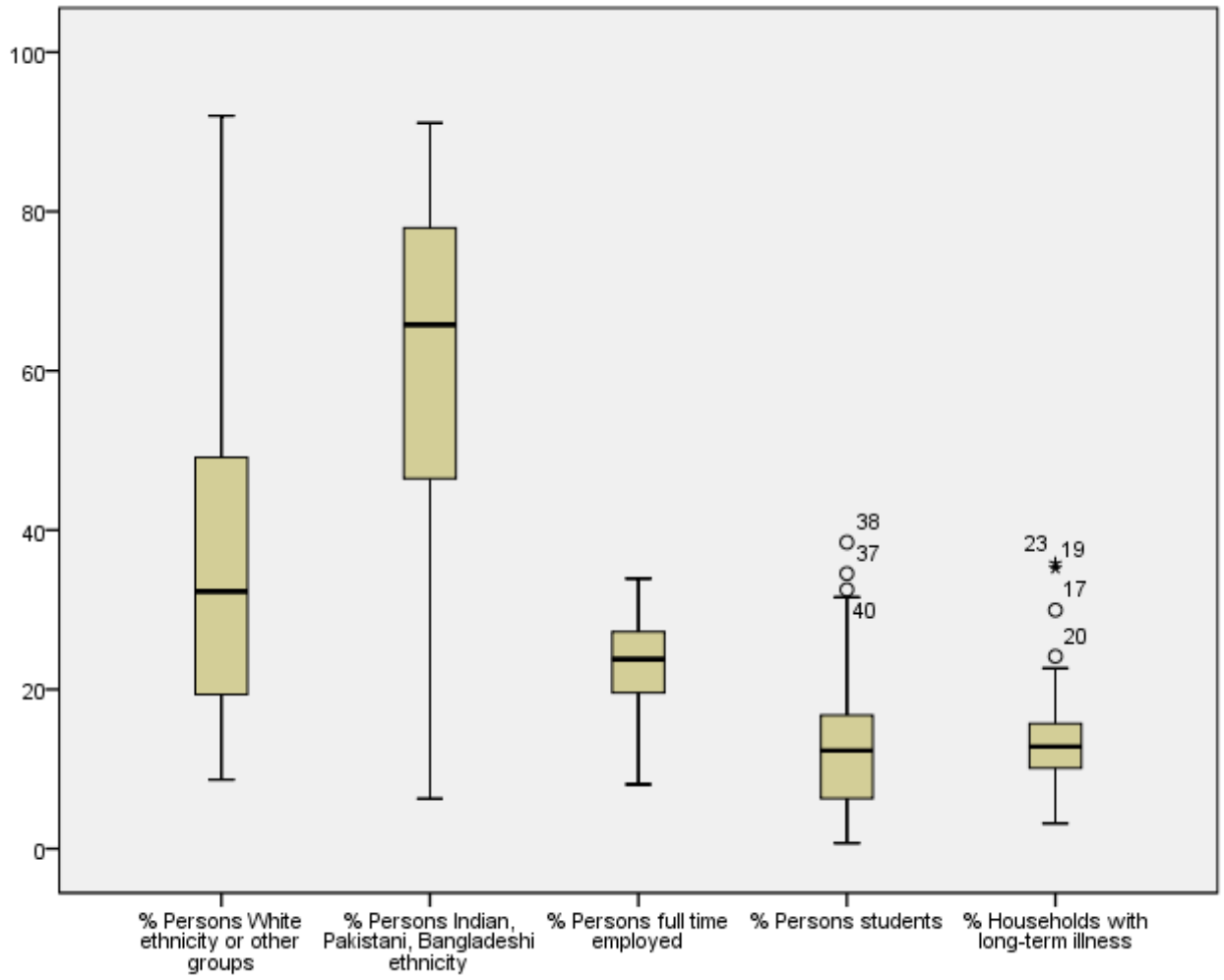
Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
Voter turnout index (over 100 = higher than average)	50	51.42	140.79	100.6076	17.66990
% Persons students	50	.70	38.46	13.5032	9.29519
Valid N (listwise)	50				

2. A histogram for the Voter Turnout Index and for one other variable



3. A boxplot of four or five variables which are percentages



4. The frequency count of different categories of 'Areas of Stress'

Area in multiple stress (5 levels)

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	No stress	5	10.0	10.0	10.0
	One stress indicator	10	20.0	20.0	30.0
	Two stress indicators	13	26.0	26.0	56.0
	Three stress indicators	15	30.0	30.0	86.0
	Four stress indicators	7	14.0	14.0	100.0
	Total	50	100.0	100.0	

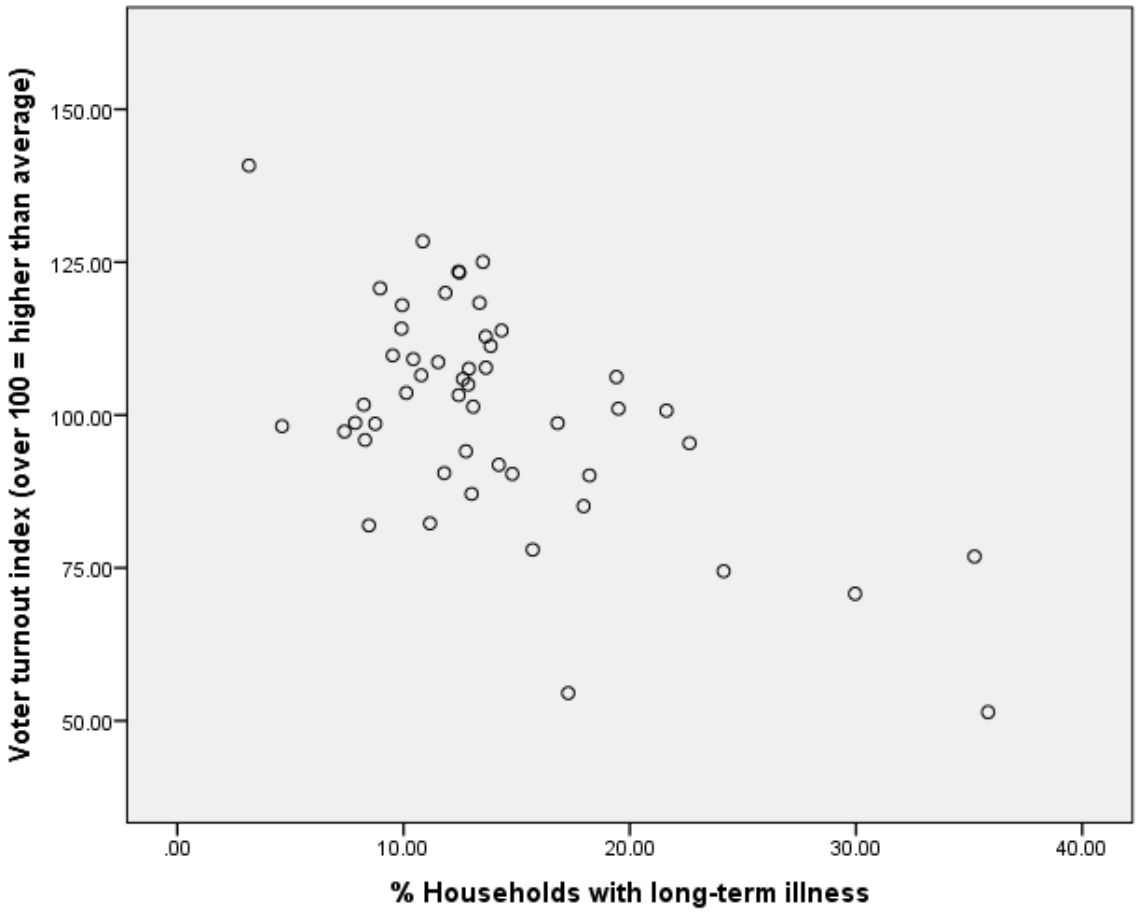
1. The correlation between the Voter Turnout Index and % Households with long-term illness

Correlations

		Voter turnout index (over 100 = higher than average)	% Households with long-term illness
Voter turnout index (over 100 = higher than average)	Pearson Correlation	1	-.597**
	Sig. (2-tailed)		.000
	N	50	50
% Households with long-term illness	Pearson Correlation	-.597**	1
	Sig. (2-tailed)	.000	
	N	50	50

**. Correlation is significant at the 0.01 level (2-tailed).

2. A scatterplot of Voter Turnout Index and % Households with long-term illness



3. Present the table of coefficients of an OLS regression between Voter Turnout Index (dependent) and % Households with long-term illness (independent). Write the R square model fit.

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients		
		B	Std. Error	Beta	t	Sig.
1	(Constant)	123.241	4.834		25.496	.000
	% Households with long-term illness	-1.608	.312	-.597	-5.157	.000

a. Dependent Variable: Voter turnout index (over 100 = higher than average)

The R Square Model fit is .357

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.597[a]	.357	.343	14.32136

a. Predictors: (Constant), % Households with long-term illness

1. Report the strongest two *positively* correlating variables with Voter Turnout Index

% Persons Indian, Pakistani, Bangladeshi ethnicity = .691
% Persons students = .224

2. Report the strongest two *negatively* correlating variables with Voter Turnout Index

% Persons White ethnicity or other groups = .-.666
% Households with long-term illness = -.597

3. After checking for collinearity between any possible independent variables, include here the table of coefficients for a multiple regression model with two or more variables you consider to be useful to include. Report the adjusted R square value

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients		
		B	Std. Error	Beta	t	Sig.
1	(Constant)	87.800	8.604		10.204	.000
	% Persons Indian, Pakistani, Bangladeshi ethnicity	.422	.090	.522	4.665	.000
	% Households with long-term illness	-.899	.302	-.334	-2.982	.005

a. Dependent Variable: Voter turnout index (over 100 = higher than average)

Model Summary

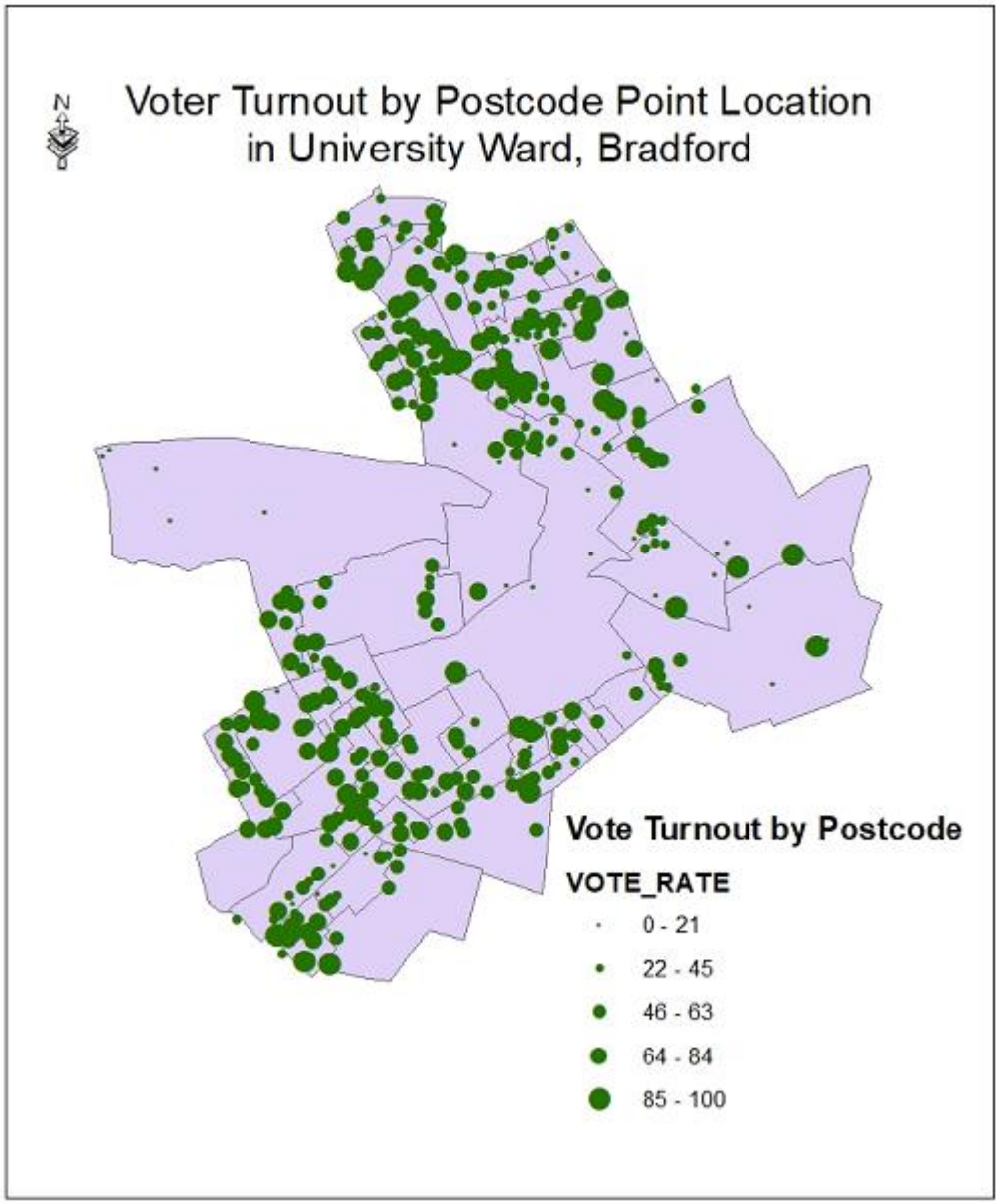
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.748[a]	.560	.541	11.96552

a. Predictors: (Constant), % Households with long-term illness, % Persons Indian, Pakistani, Bangladeshi ethnicity

Practical: Using the various resources on spatial clustering

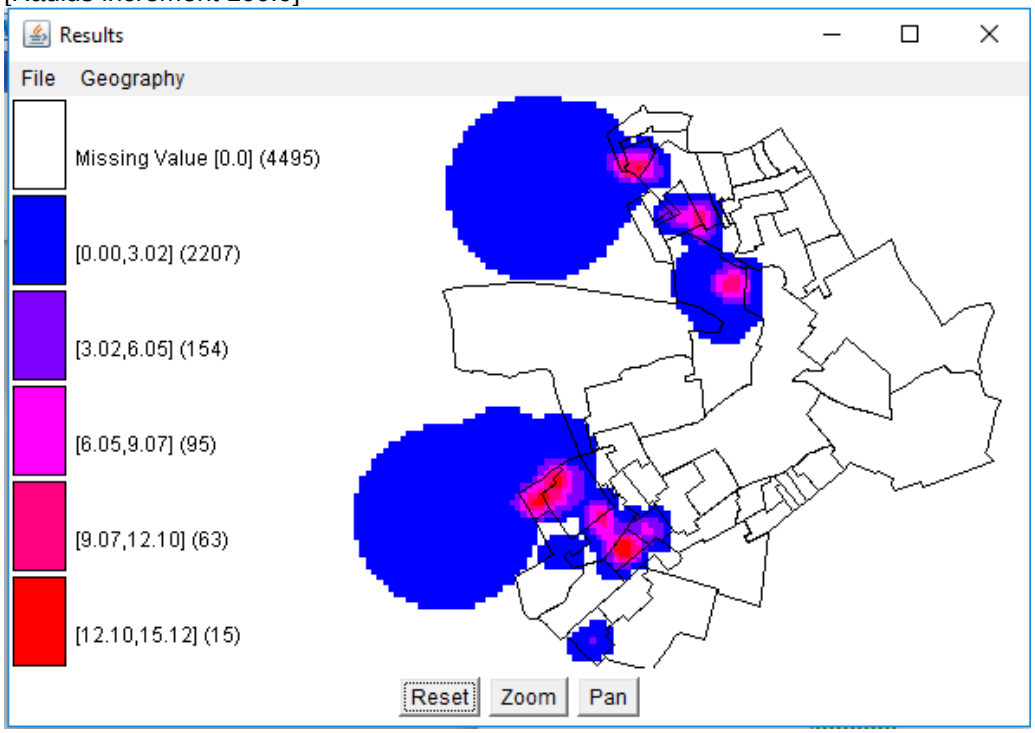
Copy and paste here:

1. A graduated symbol map of voter turnout by postcode point location

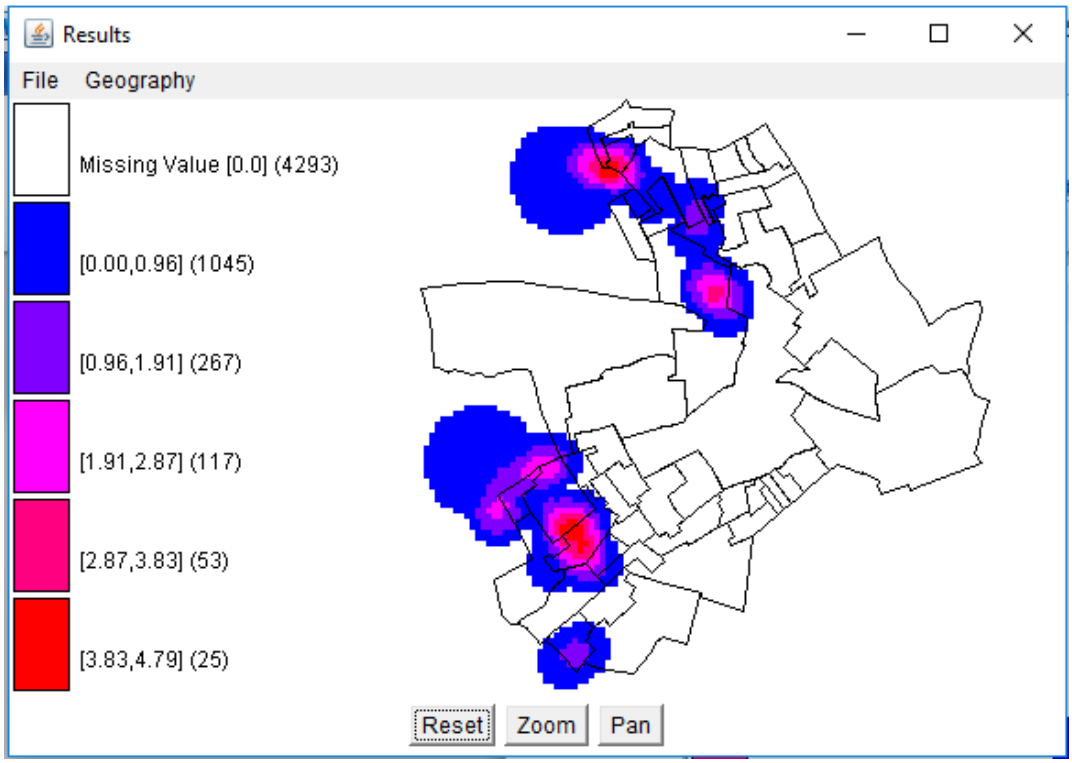


2. Two jpg outputs of spatial clustering from GAM which use different parameters (minimum / maximum radius, etc.)

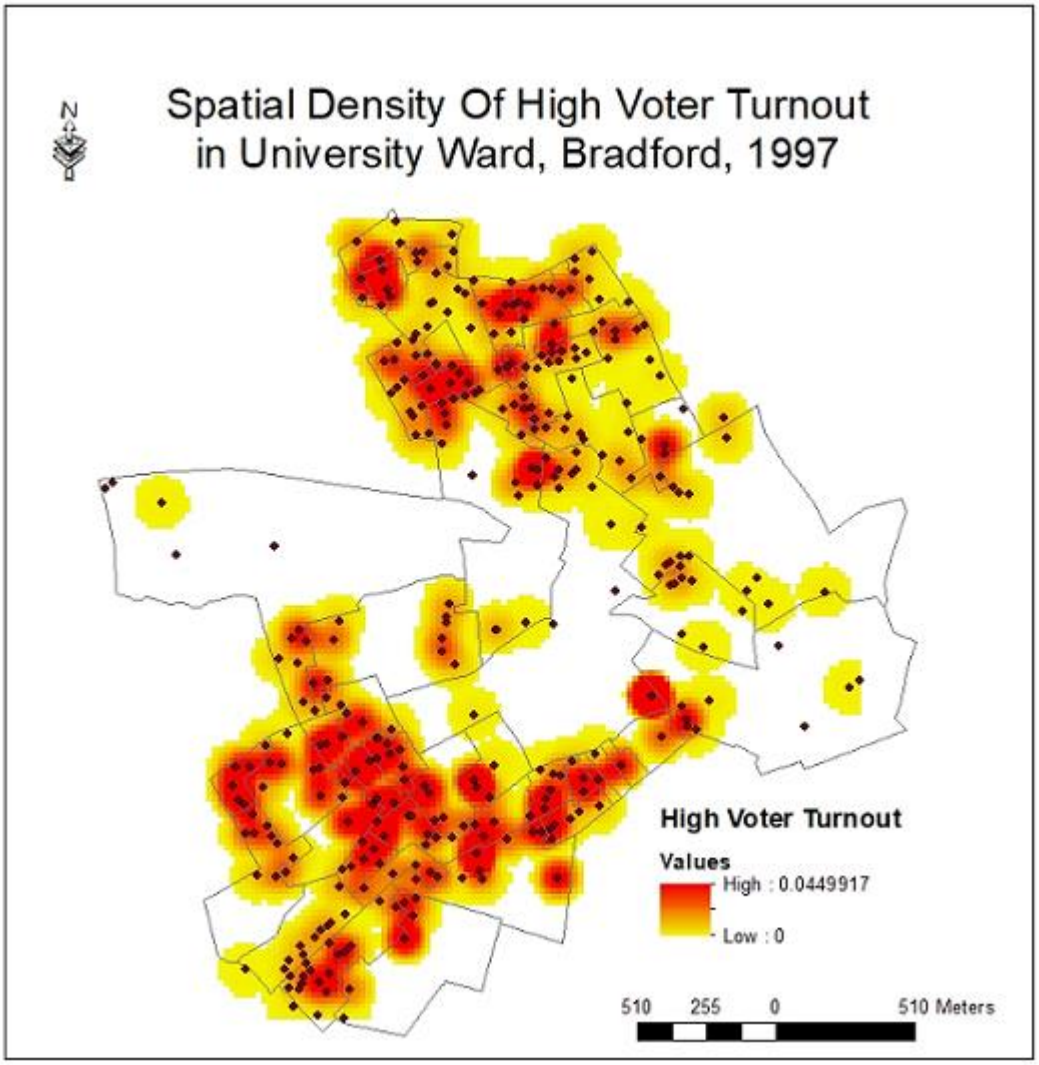
GAM output with parameters:
[Minimum radius 100.0]
[Maximum radius 500.0]
[Radius increment 100.0]

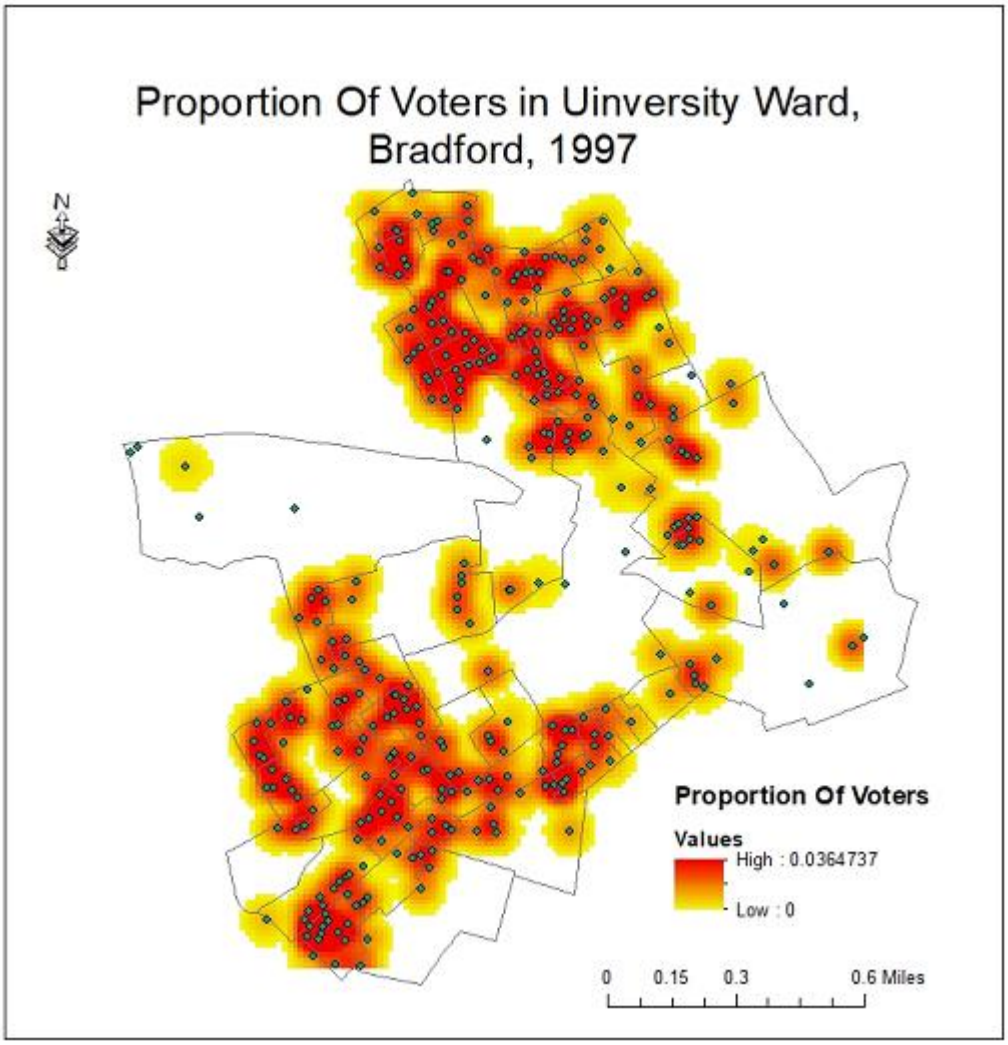


GAM with parameters:
[Minimum radius 150.0]
[Maximum radius 700.0]
[Radius increment 150.0]
[Circle overlap 0.5]

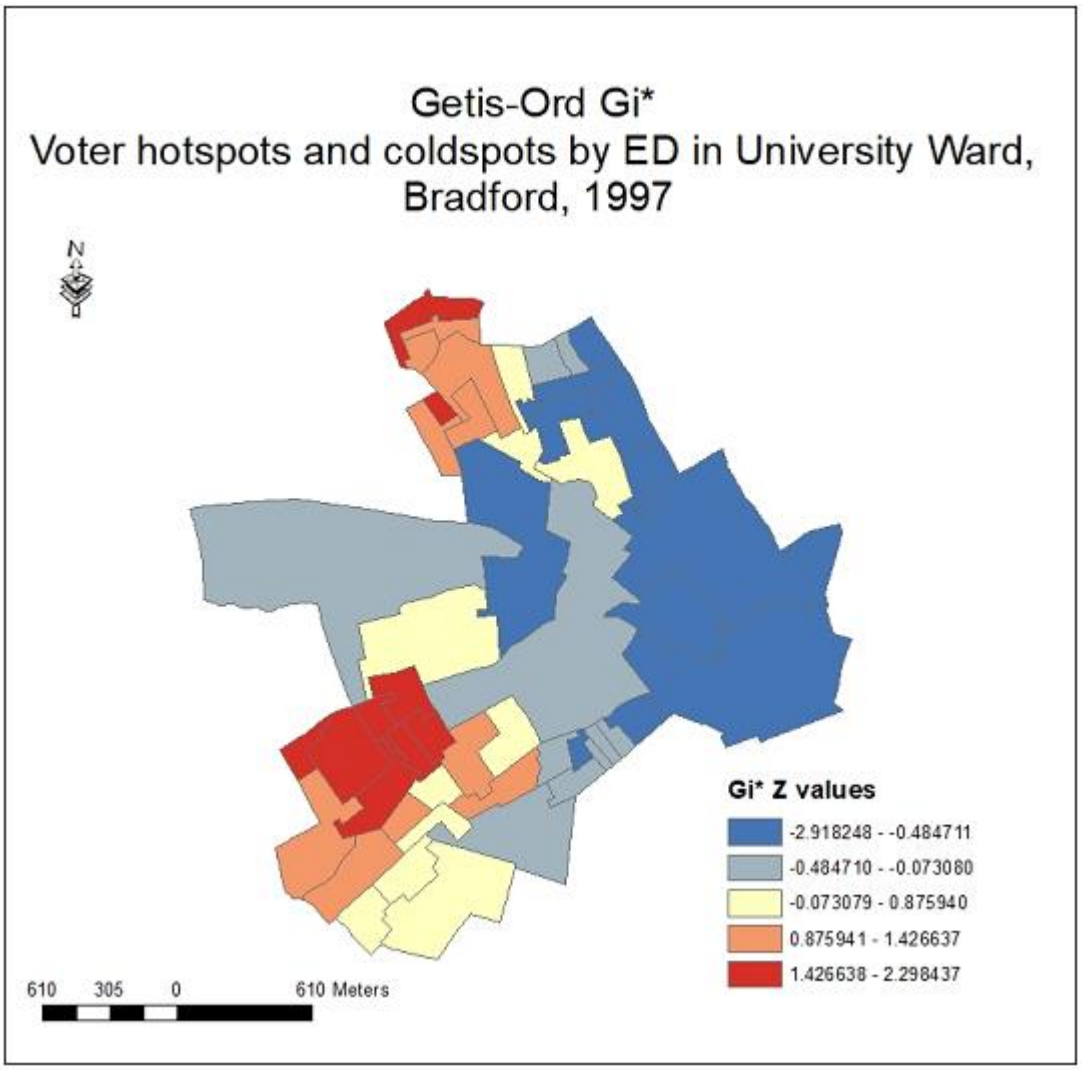


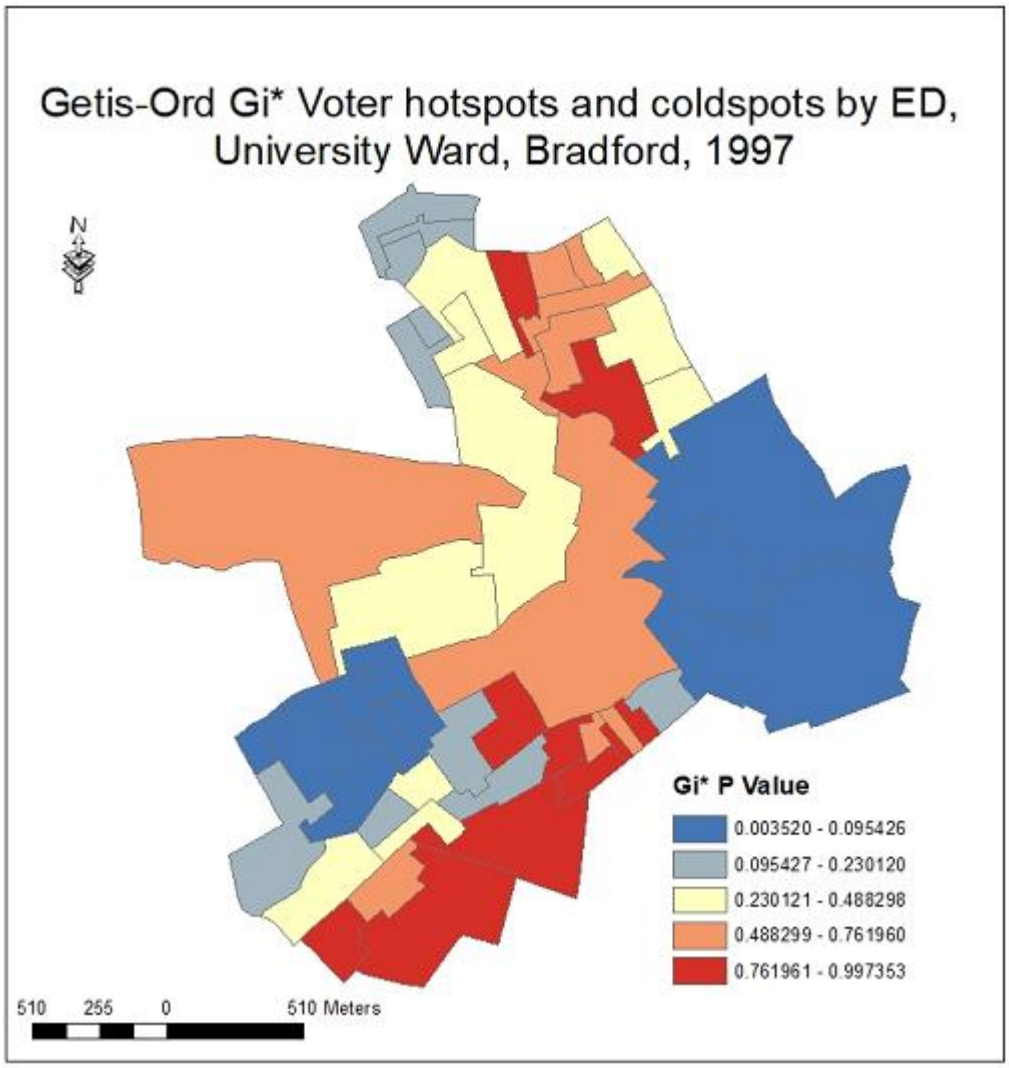
3. Two examples of Kernel Density and / or Getis Ord Gi* outputs


N
Spatial Density Of High Voter Turnout
in University Ward, Bradford, 1997
High Voter Turnout
Values
High : 0.0449917
Low : 0
510 255 0 510 Meters


Proportion Of Voters in Uinversity Ward, Bradford, 1997
N
Proportion Of Voters
Values
High : 0.0364737
Low : 0
0
0.15
0.3
0.6 Miles

Gi Ord Gi*


Getis-Ord Gi*
Voter hotspots and coldspots by ED in University Ward, Bradford, 1997
N
Gi* Z values
-2.918248 - -0.484711
-0.484710 - -0.073080
-0.073079 - 0.875940
0.875941 - 1.426637
1.426638 - 2.298437
610 305 0 610 Meters



When you have developed a classification using k-means, using information from your SPSS Output Window (or equivalent from other software), include below:

1. The frequency count of the Number of Cases (i.e. areas) in each Cluster

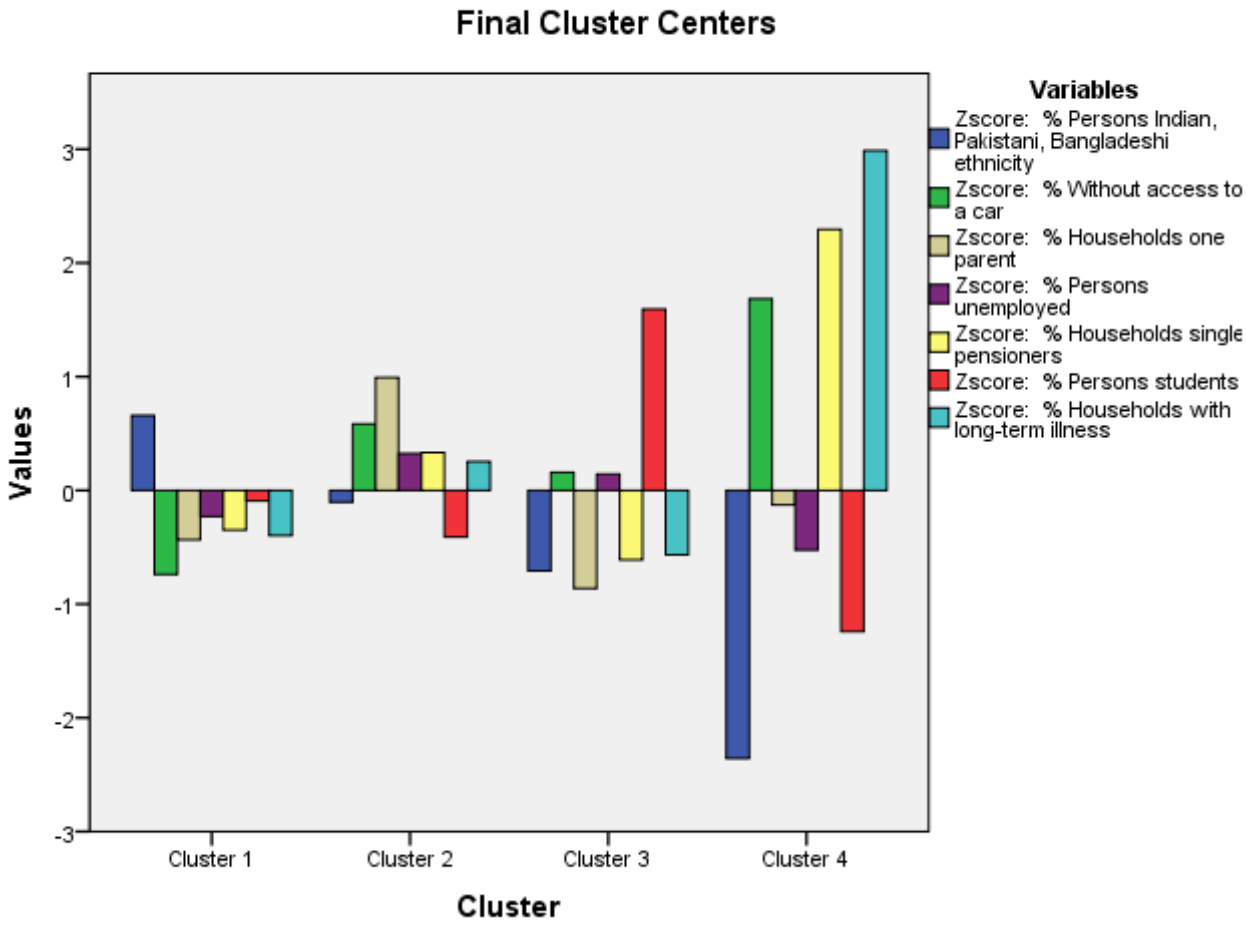
Number of Cases in each Cluster

Cluster	1	22.000
	2	17.000
	3	8.000
	4	3.000
Valid		50.000
Missing		.000

2. A table of the Final Cluster Centres or a graph to illustrate these (using SPSS or Excel)

Final Cluster Centers

	Cluster			
	1	2	3	4
Zscore: % Persons Indian, Pakistani, Bangladeshi ethnicity	.65955	-.10472	-.70701	-2.35794
Zscore: % Without access to a car	-.73888	.58458	.15823	1.68388
Zscore: % Households one parent	-.43447	.99105	-.86335	-.12754
Zscore: % Persons unemployed	-.22996	.32295	.14269	-.52417
Zscore: % Households single pensioners	-.34853	.33324	-.61018	2.29466
Zscore: % Persons students	-.09291	-.41074	1.59390	-1.24149
Zscore: % Households with long-term illness	-.39740	.25418	-.56738	2.98691



3. Stratify the Voter Turnout Index across your classification using the Aggregate procedure to show how the mean level of voting varies by area type. Paste the table below.

Cluster	VoteIndex_mean
South Asian Cluster (1)	109.10
Single Persons & Deprived Cluster (2)	95.90
Student Cluster (3)	100.09
Pensioners & Deprived Cluster (4)	66.36