APPENDIX G FACTOR LOADINGS AND COMMUNALITIES FOR YOUNG TRAFFIC OFFENDER QUESTIONNAIRE ITEMS

Table G.1

Factor loadings and communalities for personality measures

Personality measures	Factor loadings	Communalities (h^2)
Assertiveness (accounts for 17% of variance)		
I will hesitate to make phone calls to business establishments and institutions	.39	.16
I am very quick to express my opinion	.23	.05
There are times when I just can't say anything	.09	.01
I often don't know what to say to people I find attractive	.51	.26
I have hesitated to make or accept dates because of 'shyness'	.62	.38
Depression (accounts for 27% of variance)		
Living is a wonderful adventure for me	.50	.25
I feel blue and depressed	.52	.27
The future looks so gloomy that I wonder if I should go on	.54	.29
My future looks hopeful and promising	.54	.30
I often wish I was never born	.51	.27
I feel that there is more disappointment in life than satisfaction	.52	.27
I feel that life is drudgery and boredom	.50	.25
I am generally a happy person	.45	.21
Things have worked out well for me	.59	.35
Emotional adjustment (accounts for 34% of variance)		
I would call myself a tense or 'highly strung' person	.46	.21
I would say that I am fairly self-confident	.54	.30
I am often troubled with feelings of inferiority	.60	.36
My feelings are rather easily hurt	.66	.33
I would call myself a nervous person	.63	.39
I sometimes feel 'just miserable' for no good reason	.58	.34
Sensation seeking (two-factor solution)		
Thrill and adventure seeking (accounts for 18% of variance)		
I often wish I could be a mountain climber	.42	.22
I would like to take up water skiing	.46	.31
I would like to sail a long distance in a small but seaworthy sailing boat	.34	.23
I would like to learn to fly an aeroplane	.50	.41

Personality measures cont.	Factor loadings	Communalities (h^2)
Thrill and adventure seeking cont.		
A sensible person avoids activities that are dangerous	.32	.11
I would like to try surfing	.52	.35
I would never want to try jumping out of a plane with or without a parachute	.24	.07
I would like to go scuba diving	.48	.31
I think I would enjoy the sensations of skiing very fast down a large mountain slope	.54	.36
I like to dive off the high board	.57	.32
Disinhibition (accounts for 12% of variance)		
A person should have considerable sexual experience before marriage	.23	.17
I like to date members of the opposite sex who are physically exciting	.44	.32
Keeping the drinks full is the key to a good party	.60	.48
I enjoy the company of the 'in' crowd	.47	.24
I often like to get 'high' (drink alcohol or smoke marijuana)	.54	.43
I like wild 'uninhibited' parties	.49	.41
I like to have new and exciting experiences and sensations even if they are a little frightening, unconventional or illegal	.22	.32

Table G.2

Factor loadings and communalities for hostility and aggression measures

Hostility and aggression measures	Factor loadings	Communalities (h^2)
Assaultiveness (accounts for 21% of variance)		
If somebody hits me first, I let them have it	.45	.20
I have known people who have pushed me so far that we have come to blows	.45	.20
If I have to resort to physical violence to defend my rights, I will	.53	.28
Once in awhile I cannot control my urge to harm others	.39	.15
Whoever insults me or my family is asking for a fight	.55	.31
When I really lose my temper, I am capable of slapping someone	.50	.25
I can think of no good reason for ever hitting anyone	.33	.11
People who continually pester you are asking for a punch in the nose	.62	.39
I seldom strike back, even if someone hits me first	.05	.01
Indirect hostility (accounts for 33% of variance)		
I sometimes pout when I don't get my own way	.61	.37
Since the age of ten, I have never had a temper tantrum	.62	.38
I sometimes gossip about people I don't like	.61	.38
When I am angry, I sometimes sulk	.50	.25
When I am mad, I sometimes slam doors	.53	.28
Verbal hostility (accounts for 11% of variance)		
I often make threats I don't really mean to carry out	.07	.01
I could not put someone in their place even if they needed it	.11	.01
I would rather concede a point than get into an argument	.28	.08
I generally cover up my poor opinions of others	.17	.03
When I disapprove of my friend's behaviour, I let them know it	.34	.12
I can't help getting into arguments when people disagree with me	.38	.15
If somebody annoys me, I tell them what I think of them	.50	.25
I often find myself disagreeing with people	.47	.22
I demand that people respect my rights	.36	.13
Irritability (accounts for 17% of variance)		
I am always patient with others	.11	.01
I sometimes carry a chip on my shoulder	.42	.18
It makes me angry when some one make fun of me	.23	.05
Lately, I have been kind of grouchy	.56	.31
I am irritated a great deal more than people are aware	.59	.34

Hostility and aggression measures cont.	Factor loadings	Communalities (h^2)
Irritability cont.		
I often feel like a 'dynamite' ready to explode	.51	.26
I don't let a lot of unimportant things irritate me	.18	.03
Sometimes people bother me just by being around	.38	.14
Resentment (accounts for 38% of variance)		
At times, I feel I get a raw deal out of life	.64	.40
Other people always seem to get the breaks	.60	.36
If I let people see the way I feel, I'd be considered a hard person to get along with	.47	.22
When I look back on what's happened to me, I can't help feeling mildly resentful	.72	.51

Table G.3

Factor loadings and communalities for driving-related measures

Driving-related measures	Factor loadings	Communalities (h^2)
Aggression (accounts for 32% of variance)		
I often make rude signs at other motorists who annoy me	.68	.47
I lose my temper when another driver does something stupid	.59	.35
I am not easily provoked or angered when driving	.40	.16
I have given chase to a driver who has annoyed me	.64	.41
I find it difficult to control my temper when driving	.58	.33
I have been known to flash my car lights at others in anger	.41	.17
I swear out aloud at other drivers	.67	.45
I use my horn a great deal	.51	.26
If a driver follows too closely, I might hit the brakes to teach him or her a lesson	.58	.34
If the driver behind me has their lights shining in my mirror, I pay them back in some way	.51	.26
Competitive speed (accounts for 49% of variance)		
It's fun to manoeuvre and weave through traffic	.71	.50
It's fun to outwit other drivers	.73	.54
I like to pass other cars on the highway even if I'm not in a hurry	.58	.33
It's fun to beat other drivers when taking off from traffic lights	.78	.60
Driving at high speeds is exciting	.70	.49
Inhibition (accounts for 61% of variance)		
When I am feeling annoyed or angry I tend to drive more carefully because I am afraid of losing control of the car	.73	.53
When I am angry or stressed I make a conscious effort to make sure I drive safely	.81	.66
I generally become more cautious while driving when I am upset	.80	.64
Tension reduction (accounts for 79% of variance)		
I find driving a form of relaxation which I use when I feel tense	.89	.79
When I am upset, driving helps soothe my nerves	.89	.79

Table G.4

Factor loadings and communalities for selected measures

Additional measures	Factor loadings	Communalities (h^2)
Mild social deviance (accounts for 31% of variance)		
Ride on public transport without paying a fare	.59	.34
Park in a 'no standing' zone	.63	.39
Earn cash payments without paying income tax on them	.59	.34
Leave a shop with goods that you have not paid for	.57	.32
Drive the wrong way down a one-way street	.71	.51
Keep a $50 note which you have found in the street	.36	.13
Hit someone who has annoyed or upset you	.53	.28
Take time off work sick when you have something more interesting to do	.42	.17
Driving style (accounts for 60% of variance)		
I take risks	.66	.44
I race other cars	.83	.69
I cut in and out of the traffic	.82	.67
I pass other cars	.81	.65
I get angry with slow drivers	.67	.45
I like to drive fast	.85	.73
I exceed the speed limit	.76	.58
Attitudes (accounts for 14% of variance)		
I think it's OK to speed if the traffic conditions allow you to do so.	.19	.04
It is immoral to drink and drive.	.21	.04
The risk of dying young in a traffic crash is so low that you can ignore it.	.23	.05
Most of my friends drive safely.	.25	.06
I am not likely to be caught by police if committing a traffic offence.	.03	.01
Hurting someone else with my car would scar me for life.	.52	.27
I see most traffic hazards when driving.	.55	.30
I usually keep a sufficient following distance.	.59	.34

APPENDIX H METHODS FOR DETERMINING THE NUMBER OF CLUSTERS IN THE STUDENT AND OFFENDER DATA SETS

Procedures for Determining the Number of Clusters in a Data Set

Four procedures or methods and one cluster visualisation were used to determine the number of clusters in each data set: Pseudo F, the Cubic Clustering Criterion, the Ball and Hall method, and Pseudo t^2. The first three are adopted from Milligan and Cooper (1985).

The Pseudo *F* method is provided in the SAS program (Sarle, 1983) and was developed by Calanski and Harabasz (1974). Milligan and Cooper (1985) state that it "is computed as [trace $B/(k-1)$]/[trace $W/(n-k)$] where n and k are the total number of items and the number of clusters in the solution, respectively. The B and W terms are the between and pooled within cluster sum of squares and cross products matrices" (p. 163). The maximum value across the hierarchy levels is used to indicate the optimal number of clusters in the data.

Sarle (1983) conducted extensive simulations to develop the cubic clustering criterion, a test statistic provided in the SAS program. "The index is the product of two terms. The first term is the natural logarithm of $(1 - E(R^2))/(1 - R^2)$ where R^2 is the proportion if variance accounted for by the clusters and its expected value is determined under the assumption that the data have been sampled from a uniform distribution based on a hyperbox. The second term is $((np/2)^{.5})/((.001 + E(R2))^{1.2})$, where p is an estimate of the dimensionality of the between cluster variation. The constant terms were chosen on the basis of extensive simulation results (Sarle, 1983)" (Milligan & Cooper, 1985, p. 164). The maximum hierarchy level is used to indicate the correct number of clusters in the data.

The Ball and Hall (1965) method uses the average distance of the items to their respective cluster centroids as a measure of the number of clusters in the data. The largest difference between levels (success fits of k-means) is used to indicate the optimal solution.

A statistic referred to as *Je*(2)/*Je*(1) by Duda and Hart (1973) can be transformed into the Pseudo t^2 statistic. *Je*(2)/*Je*(1) is a ratio criterion where *Je*(2) is the sum of squared errors within clusters when the data are divided into two clusters, and *Je*(1) is the squared errors when only one cluster is present. The hypothesis of one cluster is rejected if the ratio is smaller than a specified critical value. (The critical value is computed from a formula given by the authors and is a function of several terms including a standard normal score, the sample size and the number of dimensions). Note that the Pseudo t^2 statistic can only be applied to hierarchical data. A small value of the Pseudo t^2 statistic and a larger Pseudo t^2 for the next cluster fusion (i.e. a large difference) suggest an optimal cluster solution.

An Isomap is a two-dimensional visualisation that reflects the inter point distances of data in space. Each data point in is connected to its nearest neighbours (Ripley, 1996) with respect to a given dissimilarity (i.e., Euclidean distance in the present data sets). One data point can only reach another data point by taking a route via data points that are connected to each other (i.e. no direct routes). The graph distance allocated to a pair of data points is the length of the shortest path via connected points. Multi Dimensional Scaling is performed on the graph distance matrix (Shepard, 1962). This is to ensure that the projection of points onto the eigenvectors with highest eigenvalues shows the configuration of points in a low dimensional Euclidean space that optimally preserves the graph distances (Purwins et al., 2004).

References

Ball, G. H., & Hall, D. J. (1965). *Isodata: a novel method of data analysis and pattern recognition*. Menlo Park: Stanford Research Institute.

Calinski, T., & Harabasz, J. (1974). A dendrite method for cluster analysis. *Commumications in Statistics, 3*(1), 1-27.

Duda, R. O., & Hart, P. E. (1973). *Pattern classification and scene analysis*. New York: Wiley.

Milligan, G. W., & Cooper, M. C. (1985). An examination of procedures for determining the number if clusters in a data set. *Psychometrika, 50*(2), 159-179.

Purwins, H., Blankertz, B., Dornhege, G., & Obermayer, K. (2004). Scale degree profiles from audio investigated with machine learning, *116th Convention of the Audio Engineering Society*. Berlin, Germany: Audio Engineering Society.

Ripley, B. D. (1996). *Pattern recognition and neural networks*. Cambridge: Cambridge University Press.

Sarle, W. S. (1983). *Cubic clustering criterion* (No. SAS Technical Report A-108). Cary, NC: SAS Institute Inc.

Shepard, R. N. (1962). The analysis of proximities: Multidimensional scaling with an unknown distance function. *Psychometrika*.

Results of Methods for Determining the Number of Clusters in the Student Data

Table H.1

SAS cluster history output from using Ward's Method for the student data

Cluster History										
NCL	Clusters Joined		FREQ	SPRSQ	RSQ	ERSQ	CCC[a]	PSF[b]	PST2[c]	Tie
10	CL25	CL45	18	0.0155	.345	.262	20.6	15.2	5.8	
9	CL19	CL17	48	0.0155	.329	.244	21.3	16.0	6.7	
8	CL18	CL13	65	0.0182	.311	.225	21.7	16.9	7.7	
7	CL15	CL26	38	0.0189	.292	.203	22.6	18.1	7.4	
6	CL16	CL10	41	0.0256	.267	.180	22.5	19.2	8.4	
5	CL9	CL8	113	0.0292	.237	.154	22.5	20.6	11.2	
4	CL7	CL6	79	0.0399	.197	.125	20.9	21.8	12.1	
3	CL12	CL5	147	0.0427	.155	.091	20.8	24.4	15.0	
2	CL3	CL11	191	0.0560	.099	.051	20.5	29.3	19.2	
1	CL4	CL2	270	0.0987	.000	.000	0.00	.	29.3	

[a] Cubic Clustering Criterion
[b] Pseudo *F* statistic
[c] Pseudo t^2 statistic

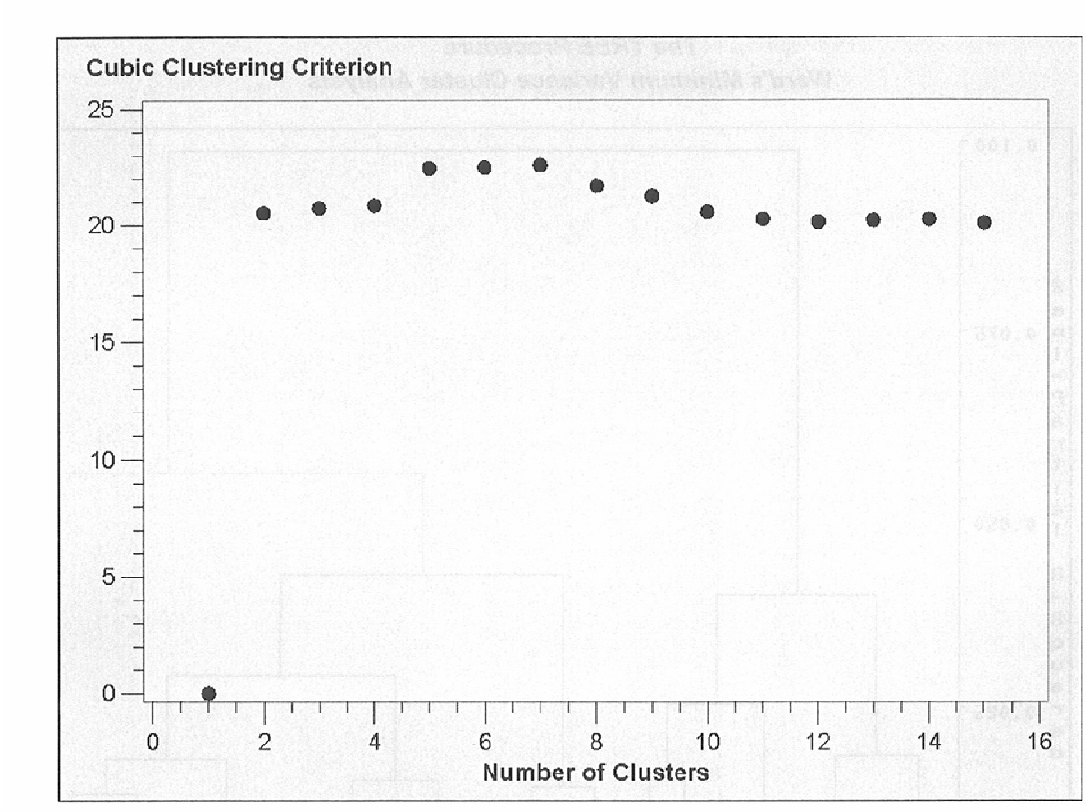

Figure H.1 SAS output of the Cubic Clustering Criterion for student data

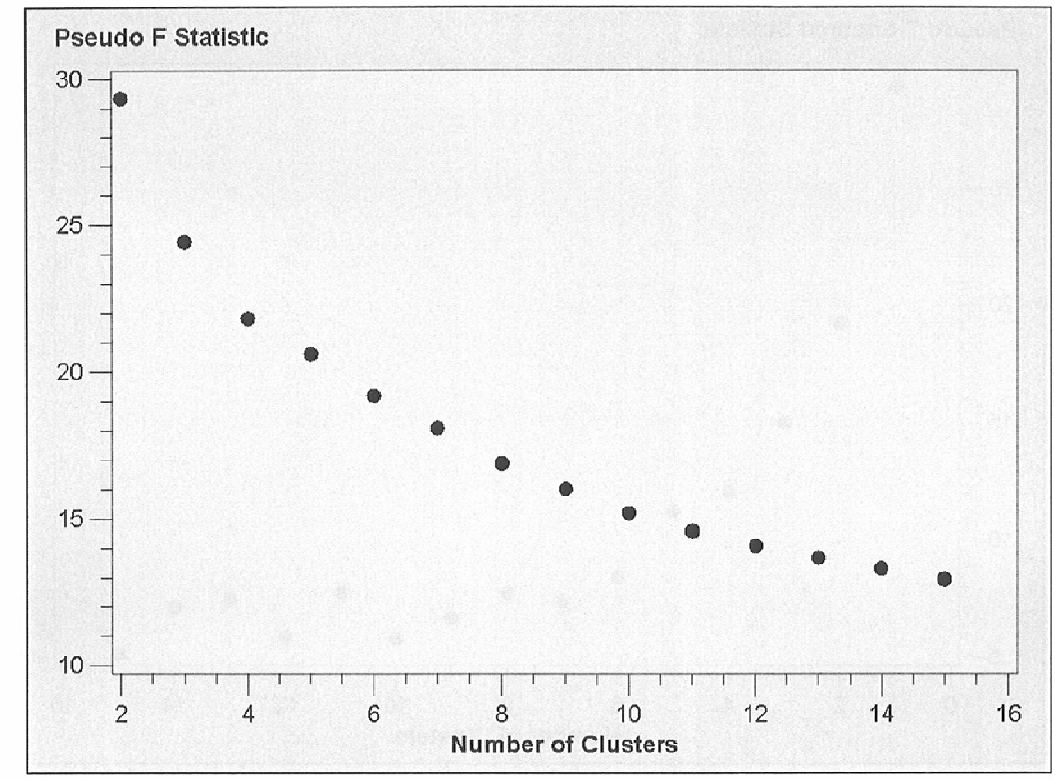

Figure H.2 SAS output of the Pseudo F for student data

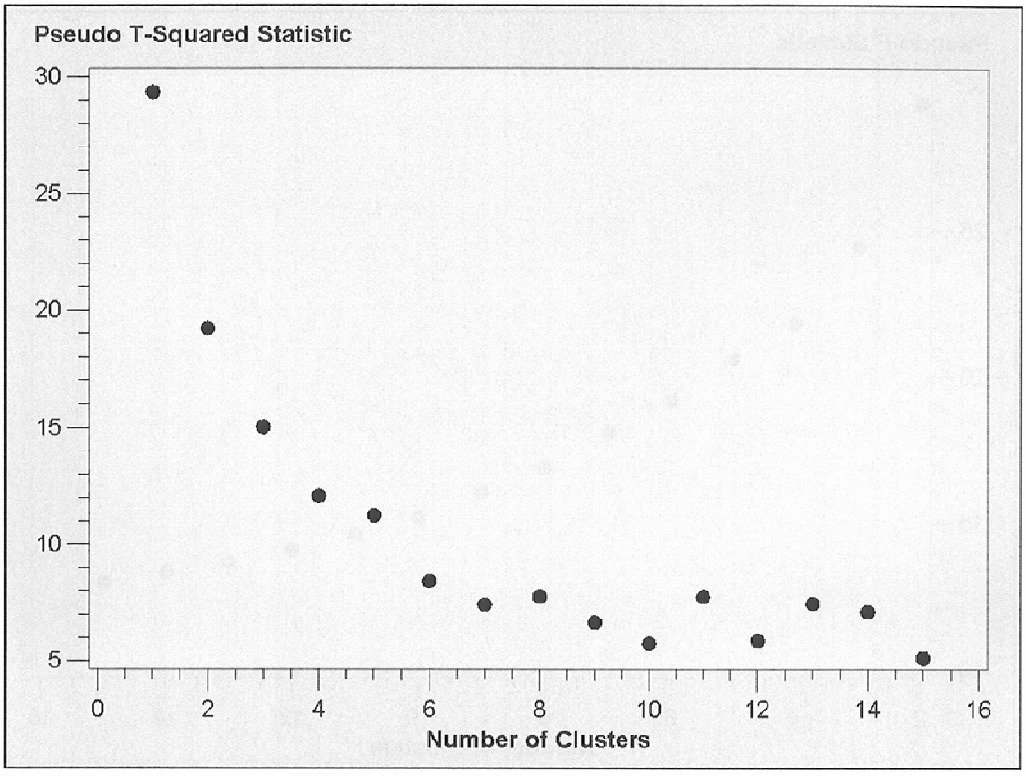

Figure H.3 SAS output of the Pseudo t^2 statistic for student data

Table H.2

Results for Ball and Hall procedure for student data

Clusters	Ball and Hall Method
10	-0.1719
9	-0.0894
8	-0.0569
7	-0.0464
6	-0.0304
5	-0.0277
4	-0.0203
3	-0.0274
2	-0.0163

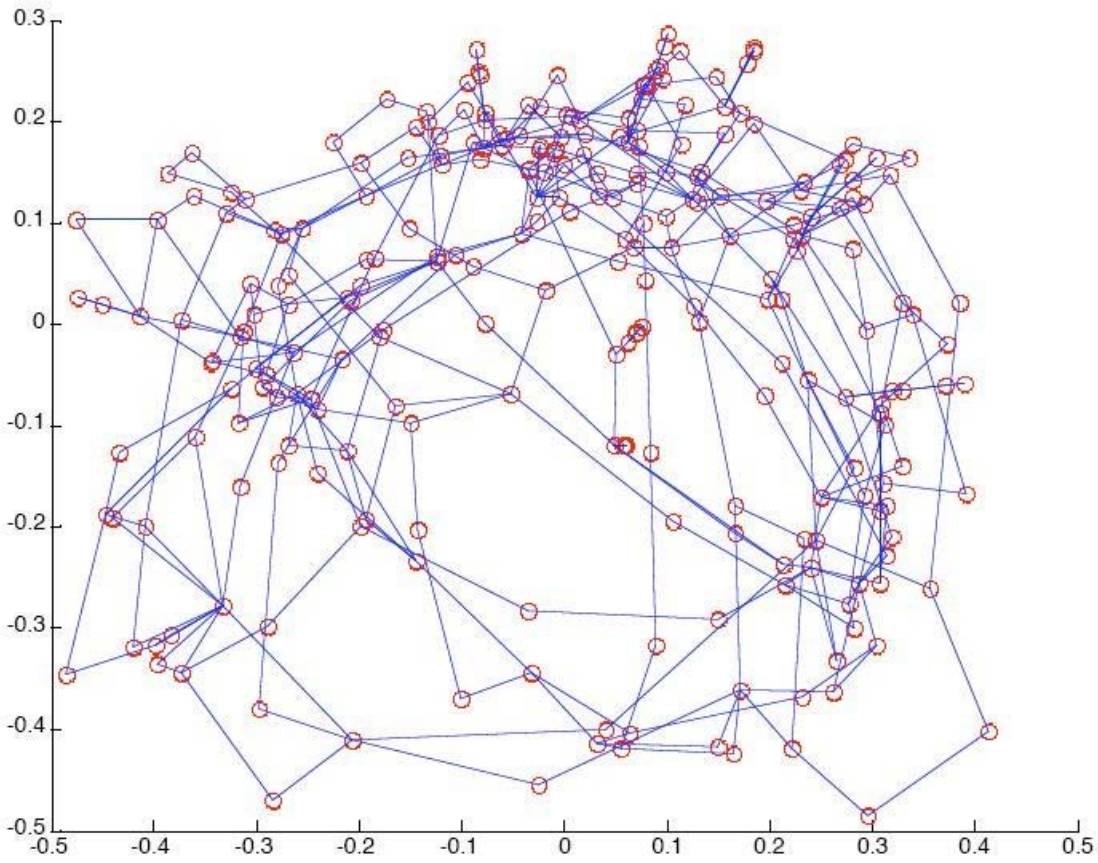

Figure H.4 Isomap visualisation for student data

Results of Methods for Determining the Number of Clusters in the Young Traffic Offender Data

Table H.3

SAS cluster history output from using Ward's Method for young traffic offender data

Cluster History										
NCL	Clusters Joined		FREQ	SPRSQ	RSQ	ERSQ	CCC[a]	PSF[b]	PST2[c]	Tie
10	CL19	CL31	47	0.0161	.379	.314	15.5	22.1	8.2	
9	CL12	CL17	50	0.0173	.362	.294	16.2	23.2	8.9	
8	CL14	CL21	83	0.0178	.344	.271	17.3	24.6	10.3	
7	CL23	CL13	49	0.0190	.325	.247	18.7	26.4	8.1	
6	CL18	CL11	73	0.0238	.301	.219	19.8	28.5	14.7	
5	CL15	CL6	107	0.0246	.277	.189	21.9	31.7	13.1	
4	CL7	CL9	99	0.0331	.244	.154	23.8	35.7	13.4	
3	CL4	CL10	146	0.0505	.193	.113	23.7	39.9	19.3	
2	CL8	CL5	190	0.0610	.132	.063	26.6	50.9	30.3	
1	CL2	CL3	336	0.1322	.000	.000	0.00	.	50.9	

[a] Cubic Clustering Criterion
[b] Pseudo *F* statistic
[c] Pseudo t^2 statistic

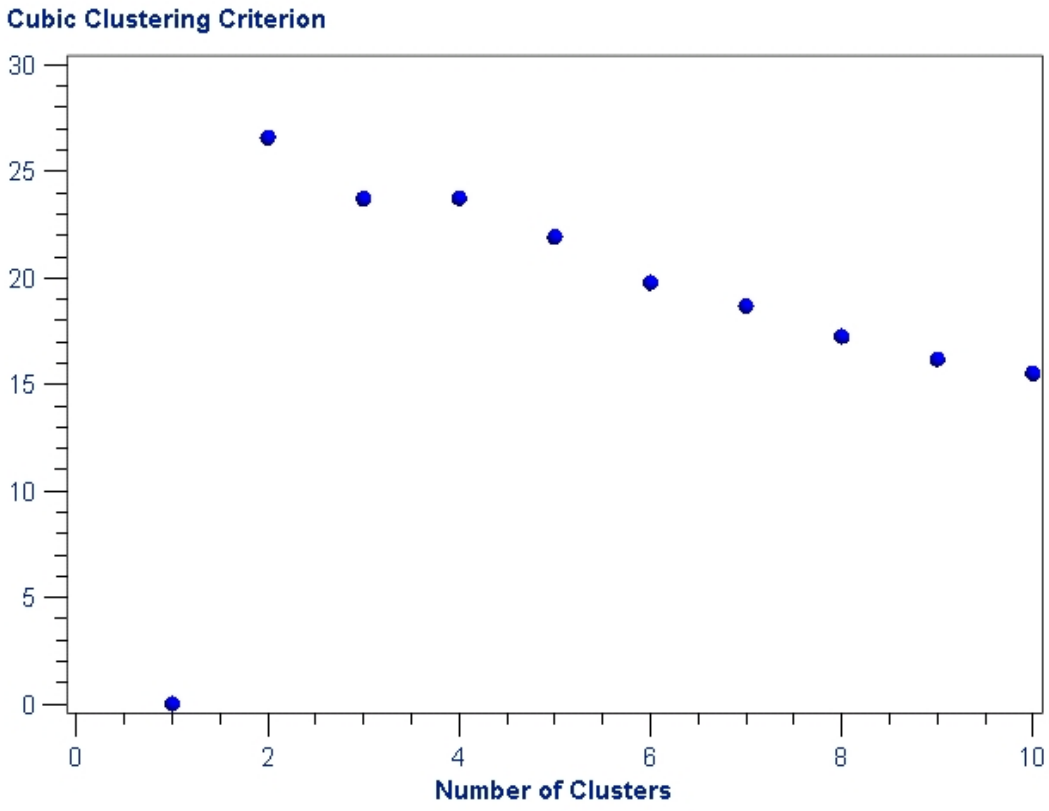

Figure H.5 SAS output of the Cubic Clustering Criterion for young traffic offender data

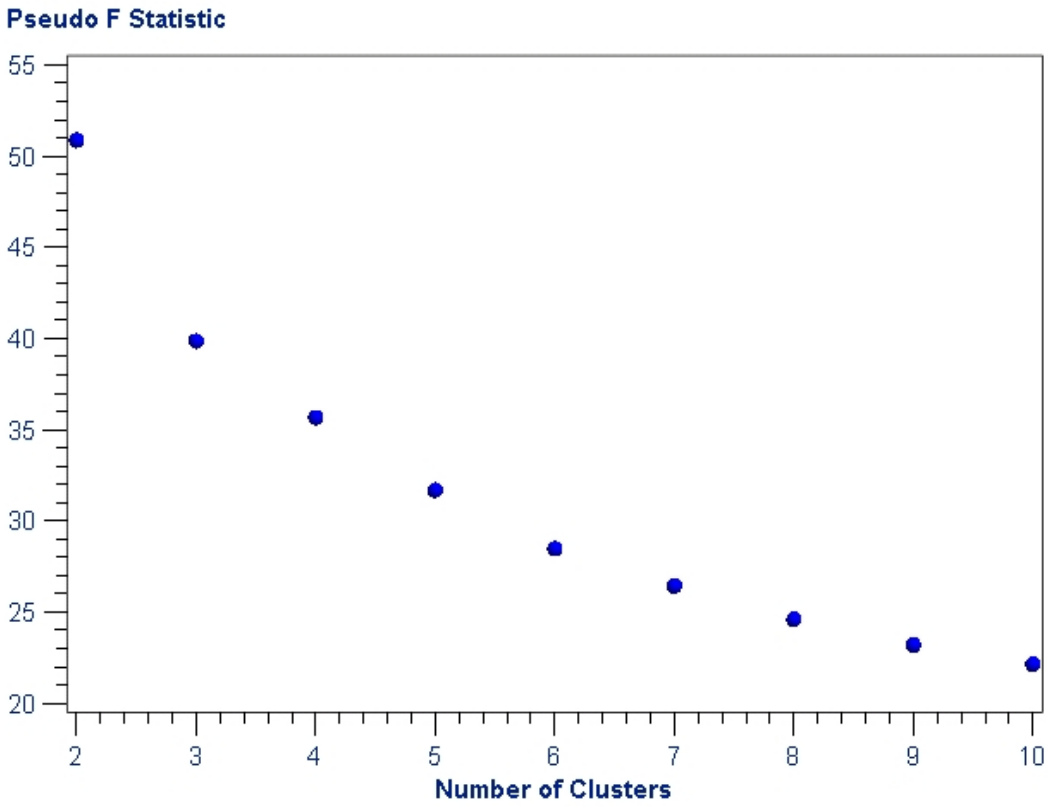

Figure H.6 SAS output of the Pseudo F statistic for young traffic offender data

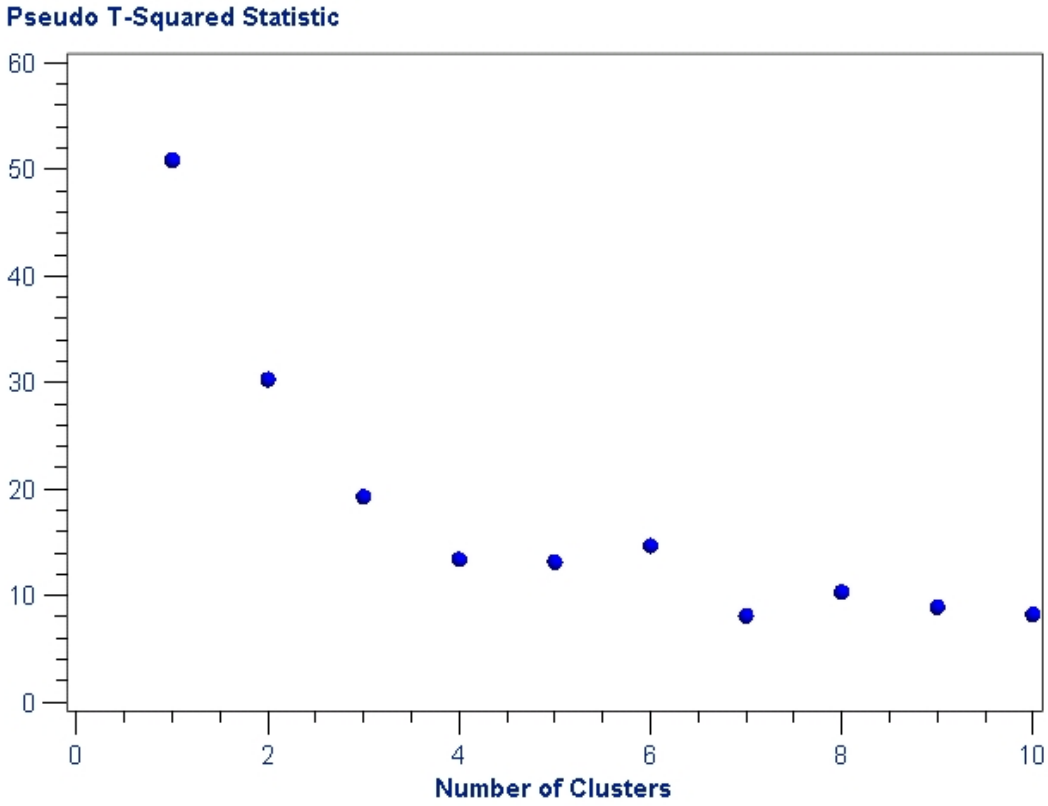

Figure H.7 SAS output of the Pseudo t^2 statistic for young traffic offender data

Table H.4

Results for Ball and Hall procedure for young traffic offender data

Clusters	Ball and Hall Method
10	-0.1728
9	-0.0956
8	-0.0617
7	-0.0410
6	-0.0351
5	-0.0269
4	-0.0289
3	-0.0182
2	-0.0154

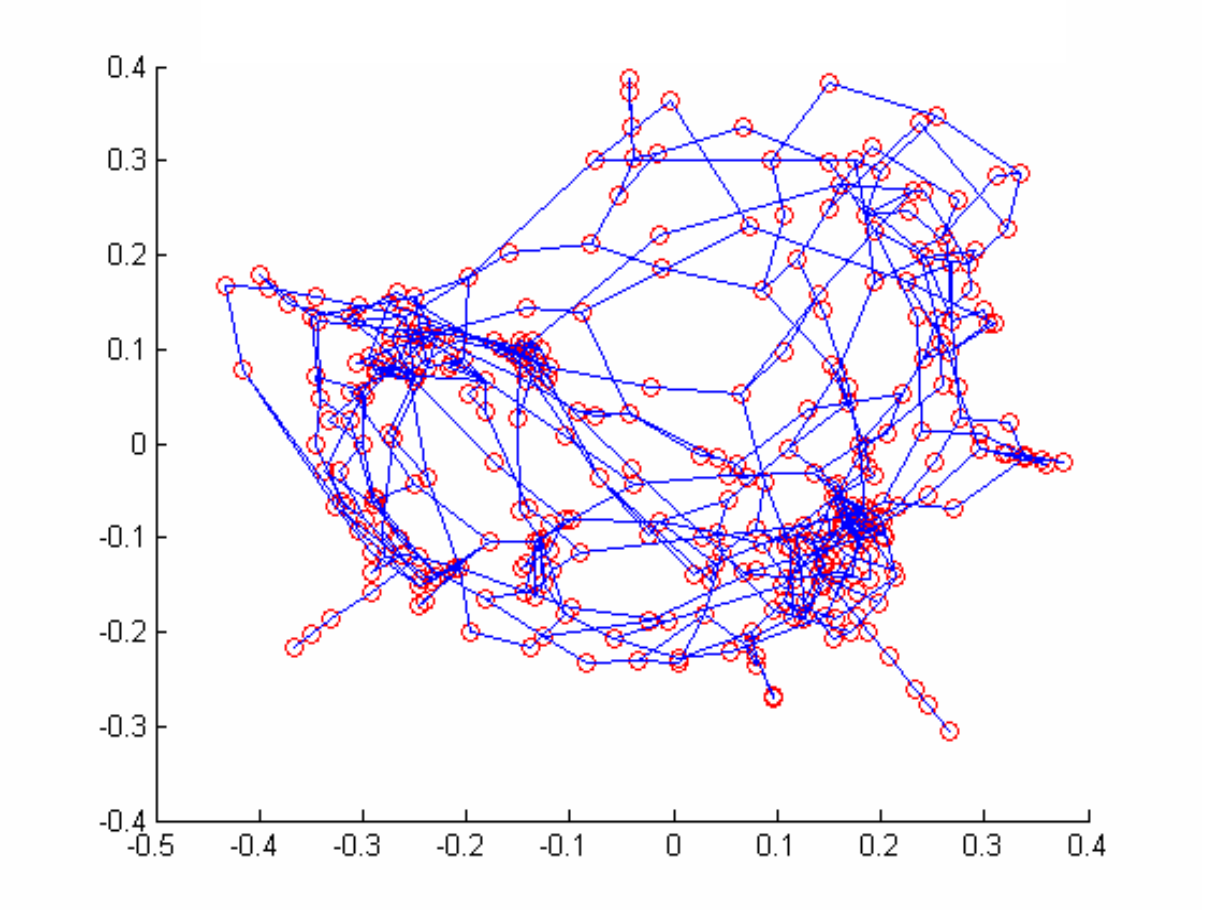

Figure H.2 Isomap visualisation for young traffic offender data