Cluster Analysis

COM 631, Spring 2019

Matt Erxleben

I. Model

Internal/Clustering Variables

- PolActivity2 – 10 item scale (alpha = .75)
    - Q83: attended meetings of your town or city council
    - Q84: attended a political meeting or rally
    - Q85: worn a button or put a sticker on your car
    - Q86: voted in the 2004 presidential election
    - Q87: participated in a march or rally
    - Q88: helped circulate or signed a petition
    - Q89: solicited political funds
    - Q90: contributed money to a party or candidate
    - Q91: contacted a public official about some issue
    - Q92: emailed, written a letter, or called media about some public issue

- Cosmopoliteness – 2 item scale (alpha = .54)
    - ZQ97: I enjoy learning about other peoples and cultures
    - ZQ98: I think of myself as a citizen of the world

- NeighTies2 – 4 item scale (alpha = .81)
    - ZQ34: Of the ten closest neighbors, how many of their home have you visited?
    - ZQ35: How many of your ten closest neighbors do you know by name or well enough to say hello when you see them on the street?
    - ZQ36: What percentage of your closest friends live in the same community?
    - ZQ2: How long have you lived in your neighborhood or community?

External Variables

- Q103 – Age (1-7 Scale)
    - 1: 18-20
    - 2: 21-30
    - 3: 31-40
    - 4: 41-50

- 5: 51-60
- 6: 61-70
- 7: 71 or older
- CivicEngagement – 2 item scale
  - ZPolActivity1
  - ZTotalOrgs
- Q57 – Comfort talking w/ strangers on street (1-4 Scale)
  - 1: "very uncomfortable"
  - 4: "very comfortable"
- Q21 – Comfort voicing complaints at public meetings (0-10 Scale)
  - 0: "completely disagree"
  - 10: "completely agree"
- Q60 – Freq read community newspaper (0-5 Scale)
  - 0: "no community newspaper"
  - 1: "almost never"
  - 5: "all the time"
- Q79 – Perceived political knowledge (0-10 Scale)
  - 0: "not at all knowledgeable"
  - 10: "very knowledgeable"
- Q82 – Number of people talk politics with regularly (0-5 Scale)
  - 0: "none"
  - 1: "one"
  - 2: "two or three"
  - 3: "five to ten"
  - 4: "more than ten"
- Q100 – Frequently watch TV news (0-6 Scale)
  - 0: "never"
  - 1: "less often than [choice 2]"
  - 2: "one or two days a week"
  - 3: "three or four days a week"
  - 4: "five or six days a week"
  - 5: "about once a day"
  - 6: "several times a day"
- Q107 – Household income (1-9 Scale)
  - 1: $10,000 or less
  - 2: $10,001 - $20,000
  - 3: $20,001 - $30,000
  - 4: $30,001 - $40,000
  - 5: $40,001 - $50,000
  - 6: $50,001 - $60,000
  - 7: $60,001 - $75,000
  - 8: $75,001 - $100,000
  - 9: $100,001 - $150,000

- Q109 – Gender
  - 1: Male
  - 2: Female
- Race – % white
- Q6 – value family (0-10 Scale)
  - 0: "totally unimportant"
  - 10: "extremely important"
- Q7 – value work (0-10 Scale)
  - 0: "totally unimportant"
  - 10: "extremely important"
- Q10 – value religion (0-10 Scale)
  - 0: "totally unimportant"
  - 10: "extremely important"
- Q12 – value being an American (0-10 Scale)
  - 0: "totally unimportant"
  - 10: "extremely important"

II. Running SPSS

1. Standardize internal variables if necessary
2. Analyze → Classify → Hierarchical Cluster

3. Select your internal variables
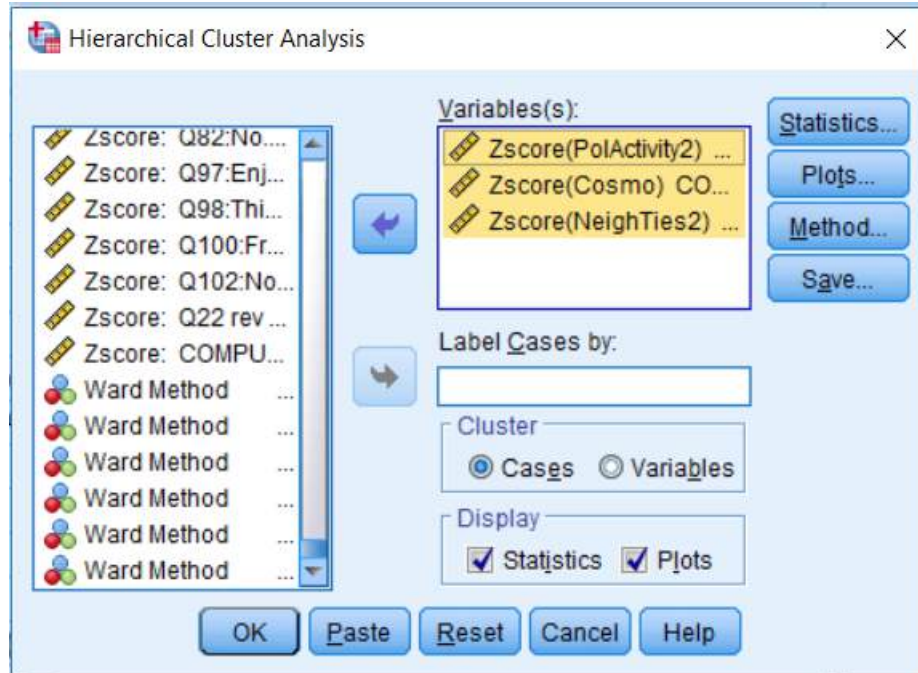
4. Click "Statistics"

→ check "Agglomeration schedule"

→ select "Range of solutions" → indicate minimum and maximum number of clusters
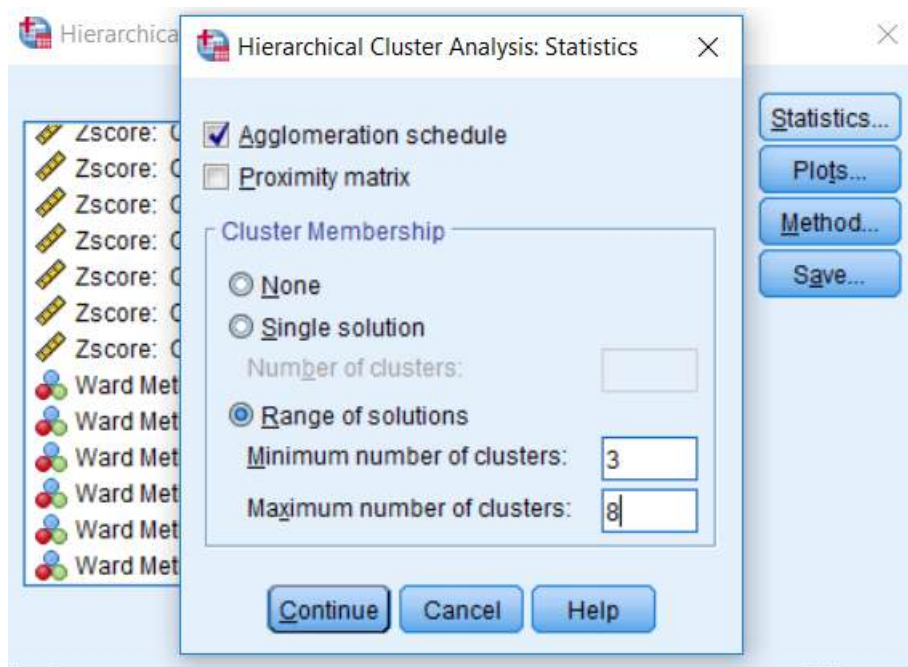
→ click "Continue"

5. Click "Plots"

→ select either "Dendrogram" or one of the "Icicle" options

→ click "Continue"
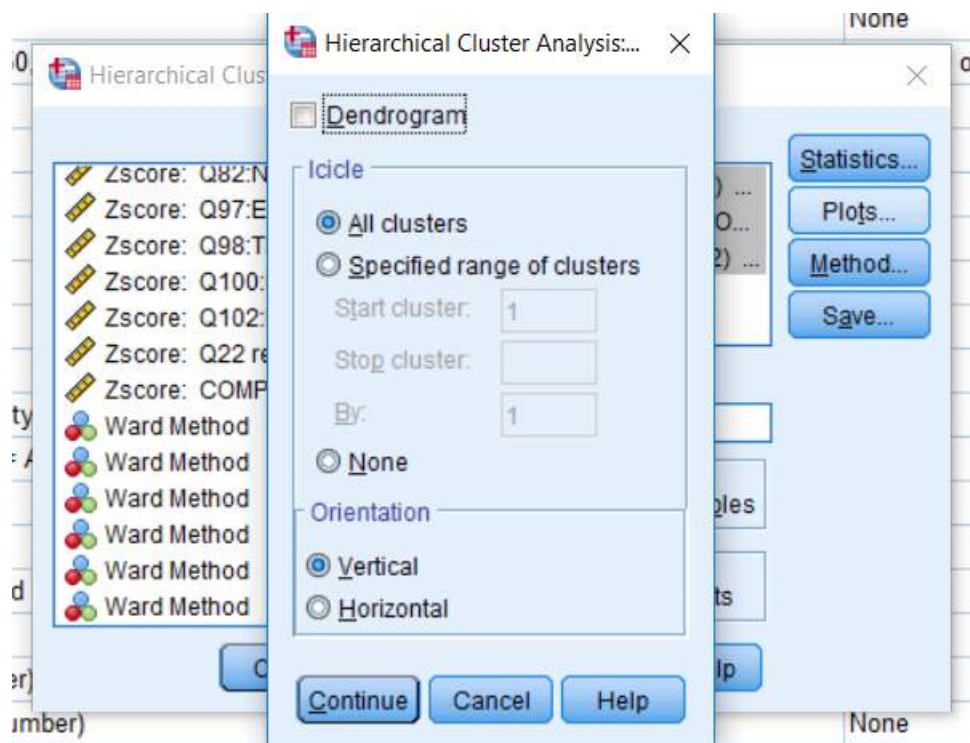
6. Click "Method"

    → click "Cluster Method" → select "Ward's Method"

    → click "Interval" → select "Squared Euclidean Distance"

    → click "Continue"

7. Click "Save"

 → select "Range of solutions" → indicate minimum and maximum number of clusters

 → click "Continue"
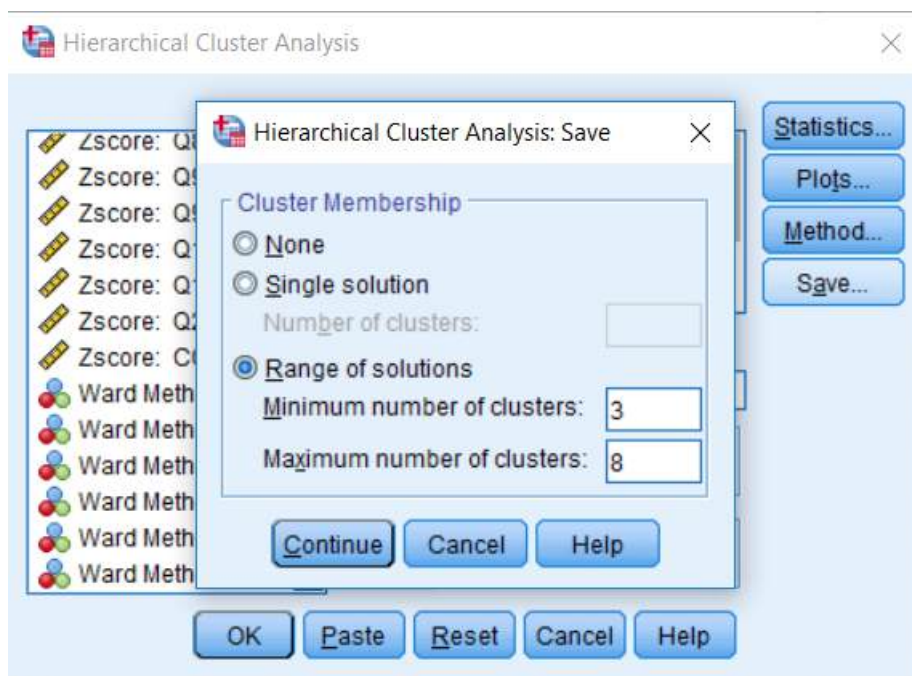
8. Click "OK" or "Paste

NOTE: This is the end of the actual Cluster procedure. This has produced the Agglomerative Schedule and Cluster Membership Table in the output. It has also saved the created clusters as new variables (8 cluster variable = CLU8_1, 7 cluster variables = CLU7_1, etc.)

****

NEXT: Use Frequencies and ANOVA procedures to decide which cluster solution to use.

1. Analyze → Descriptive Statistics → Frequencies

2. Select the cluster variables (will be at the bottom of the list)
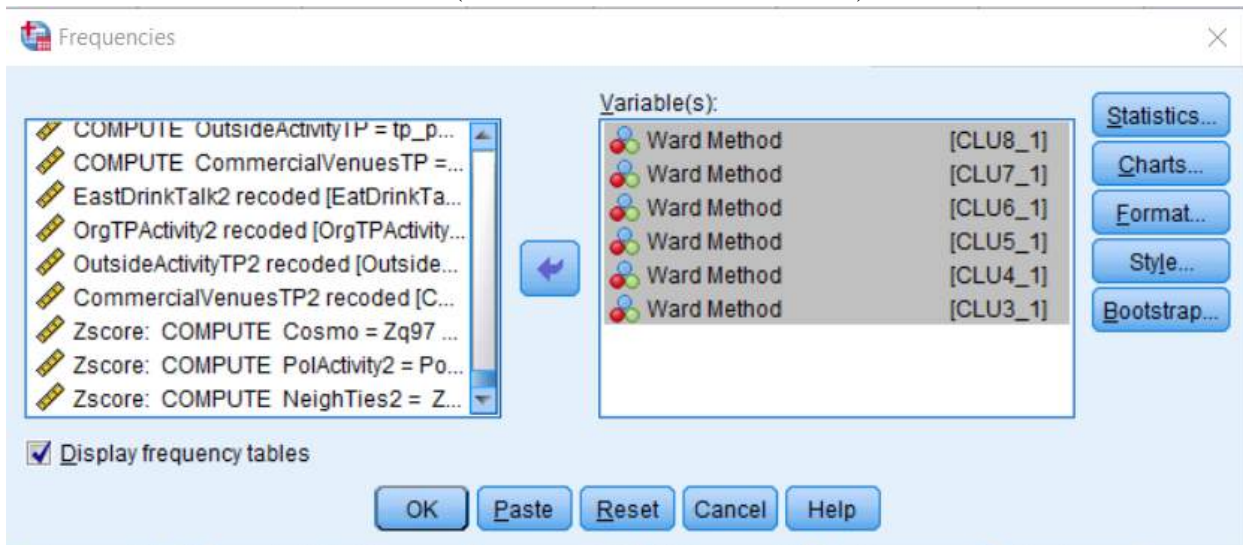
3. Click "OK" or "Paste"

NOTE: There are several ways to compare means using an ANOVA test. I used One-Way ANOVA.

1. Analyze → Compare Means → One-Way ANOVA
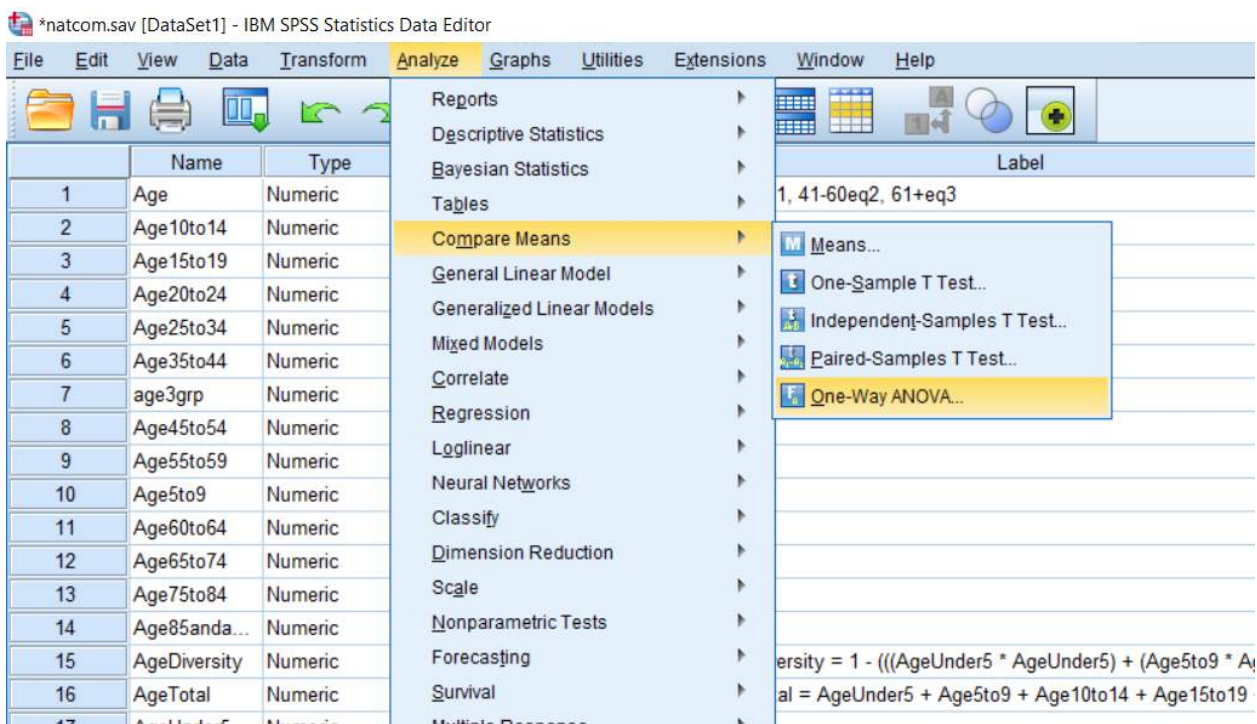
2. Enter your internal variables in the "Dependent List"
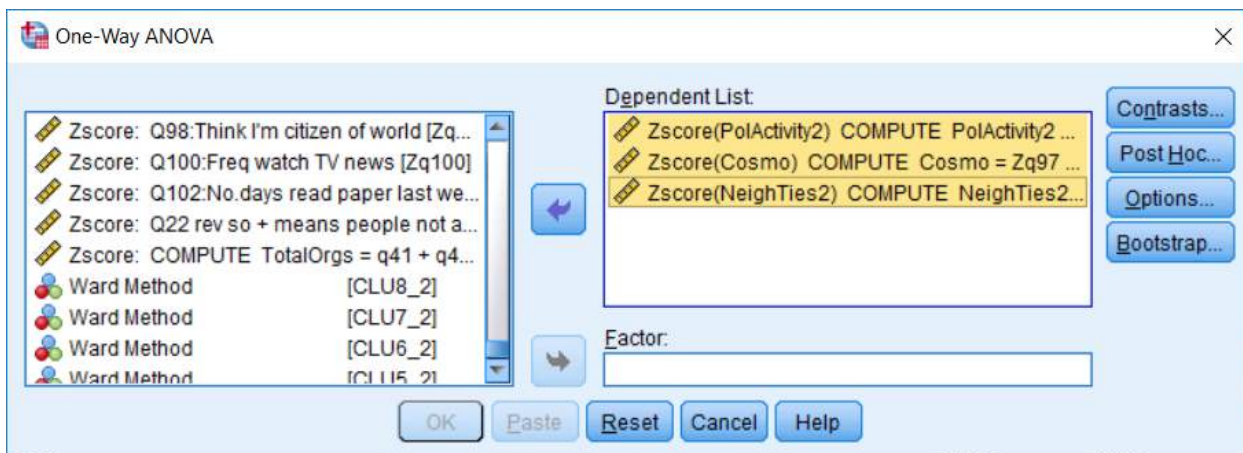
3. Enter your external variables in the "Dependent List"
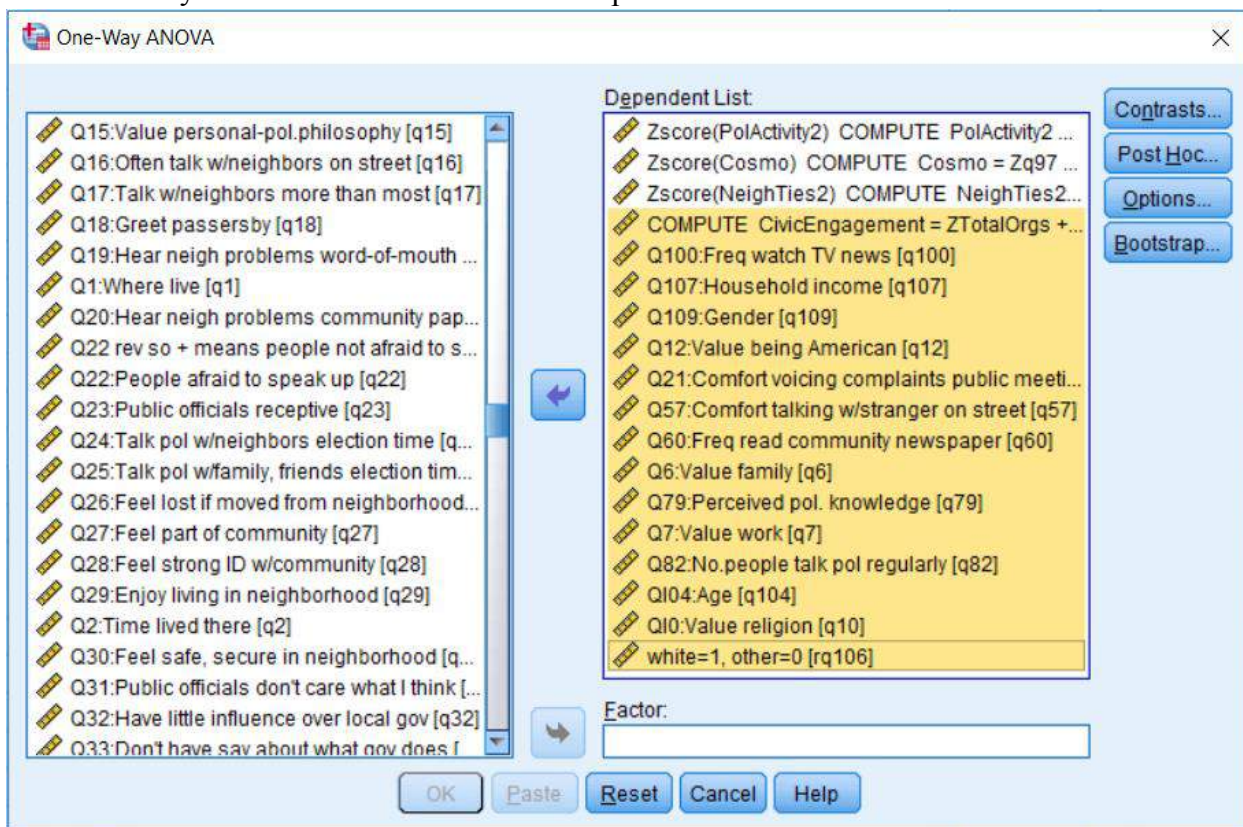
4. Enter the cluster solution of your choice under "Factor"
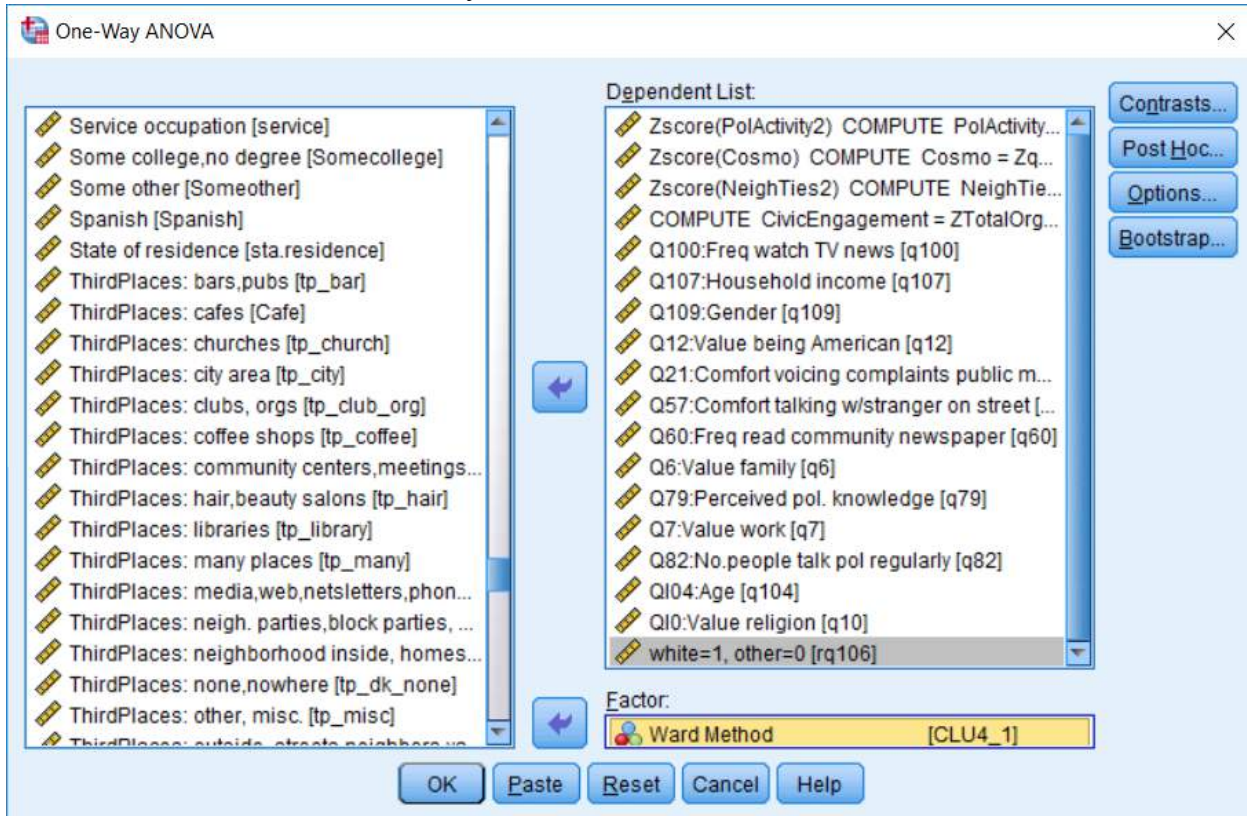
5. Click "Post Hoc"

→ under "Equal Variance Assumed" → select post-hoc tests of your choice

→ click "Continue"
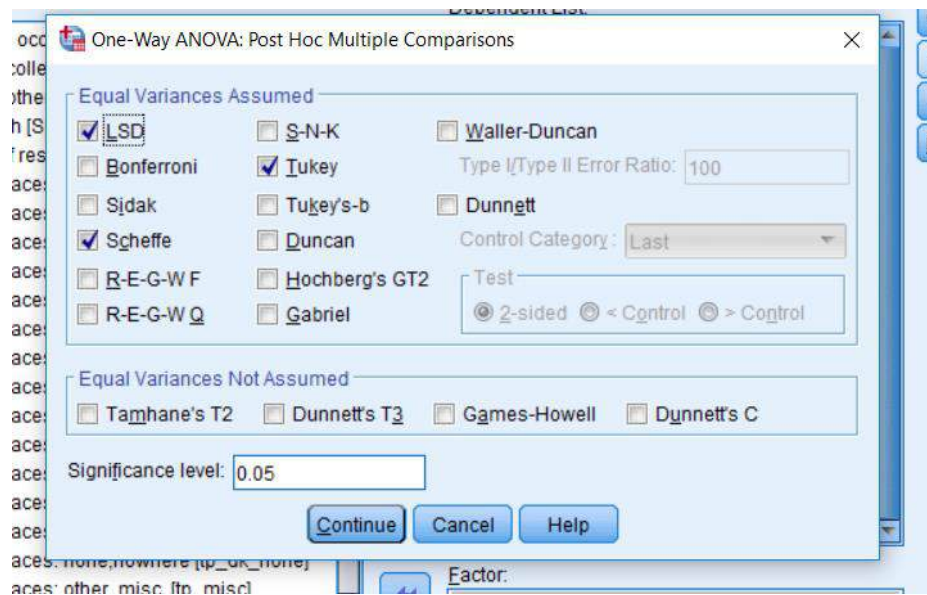
6. Click "Options"

→ check "Descriptive" (I did something different for this)

→ click "Continue"
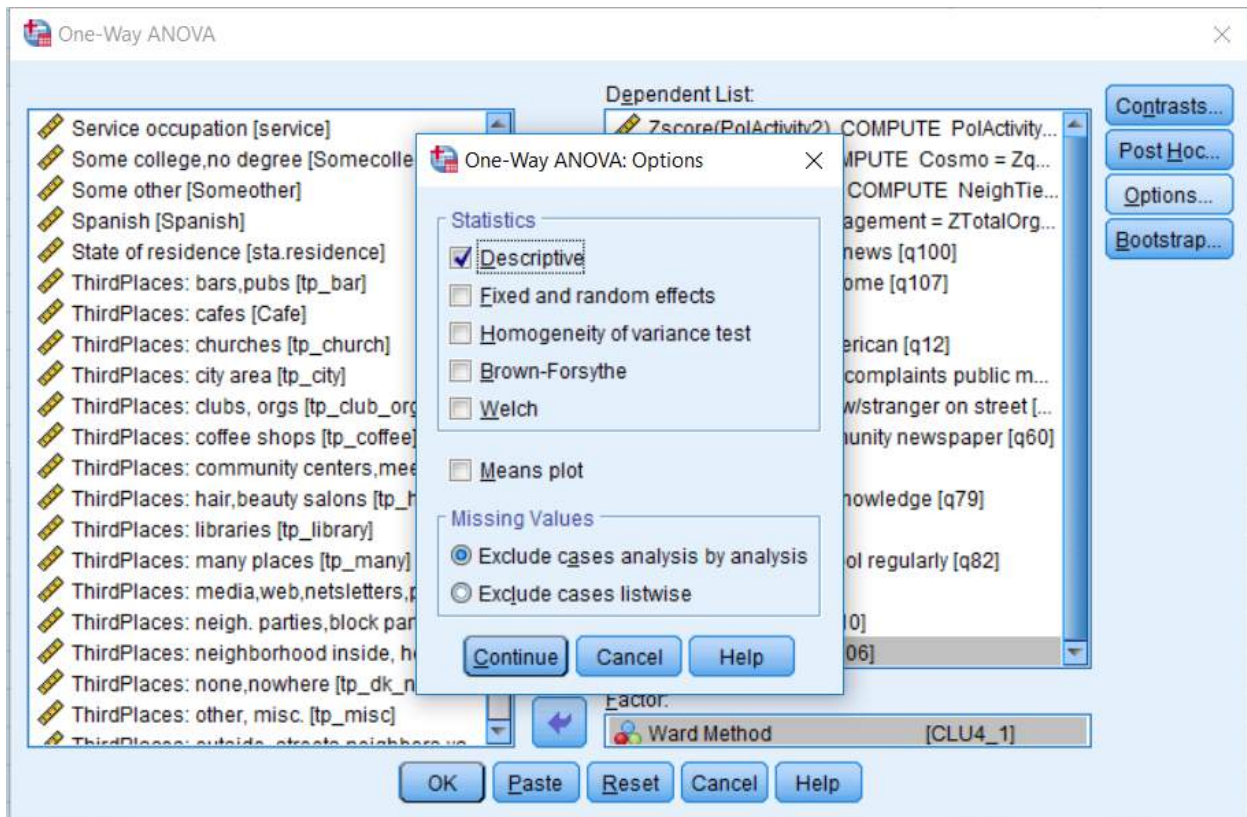
7. Click "OK" or "Paste"

III. SPSS Output

```
DESCRIPTIVES VARIABLES=Cosmo PolActivity2 NeighTies2
  /SAVE
  /STATISTICS=MEAN STDDEV MIN MAX.
```

Descriptives

Notes

Output Created		01-MAY-2019 14:32:21
Comments		
Input	Data	C:\Users\merxl\Downloads\natcom.sav
	Active Dataset	DataSet1
	File Label	CP05
	Filter	<none>
	Weight	<none>
	Split File	<none>
	N of Rows in Working Data File	477
Missing Value Handling	Definition of Missing	User defined missing values are treated as missing.
	Cases Used	All non-missing data are used.
Syntax		DESCRIPTIVES VARIABLES=Cosmo PolActivity2 NeighTies2 /SAVE /STATISTICS=MEAN STDDEV MIN MAX.
Resources	Processor Time	00:00:00.02
	Elapsed Time	00:00:00.02
Variables Created or Modified	ZSco01	Zscore(Cosmo) COMPUTE Cosmo = Zq97 + Zq98 (COMPUTE)
	ZSco02	Zscore(PolActivity2) COMPUTE PolActivity2 = PolActivity1 + q91 + q92 (COMPUTE)

	ZSco03	Zscore(NeighTies2) COMPUTE NeighTies2 = Zq34 + Zq35 + Zq36 + Zq2 (COMPUTE)

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
COMPUTE Cosmo = Zq97 + Zq98 (COMPUTE)	432	-5.00	1.78	-.0010	1.65882
COMPUTE PolActivity2 = PolActivity1 + q91 + q92 (COMPUTE)	434	.00	10.00	3.6774	2.61230
COMPUTE NeighTies2 = Zq34 + Zq35 + Zq36 + Zq2 (COMPUTE)	454	-6.52	5.81	.0257	2.76402
Valid N (listwise)	421				

```
CLUSTER   ZPolActivity2 ZCosmo ZNeighTies2
  /METHOD WARD
  /MEASURE=SEUCLID
  /PRINT SCHEDULE CLUSTER(3,8)
  /PLOT VICICLE
  /SAVE CLUSTER(3,8).
```

Cluster

Notes

Output Created		01-MAY-2019 14:32:21
Comments		
Input	Data	C:\Users\merxl\Downloads\natcom.sav
	Active Dataset	DataSet1
	File Label	CP05
	Filter	<none>
	Weight	<none>
	Split File	<none>
	N of Rows in Working Data File	477
Missing Value Handling	Definition of Missing	User-defined missing values are treated as missing.

Cases Used			Statistics are based on cases with no missing values for any variable used.
Syntax			CLUSTER ZPolActivity2 ZCosmo ZNeighTies2
			/METHOD WARD
			/MEASURE=SEUCLID
			/PRINT SCHEDULE CLUSTER(3,8)
			/PLOT VICICLE
			/SAVE CLUSTER(3,8).
Resources	Processor Time		00:00:00.55
	Elapsed Time		00:00:00.42
Variables Created or Modified	Cluster Membership	CLU8_2	Ward Method
		CLU7_2	Ward Method
		CLU6_2	Ward Method
		CLU5_2	Ward Method
		CLU4_2	Ward Method
		CLU3_2	Ward Method

Case Processing Summary[a,b]

Cases					
Valid		Missing		Total	
N	Percent	N	Percent	N	Percent
421	88.3	56	11.7	477	100.0

a. Squared Euclidean Distance used

b. Ward Linkage

Ward Linkage

Agglomeration Schedule

	Cluster Combined			Stage Cluster First Appears		
Stage	Cluster 1	Cluster 2	Coefficients	Cluster 1	Cluster 2	Next Stage
394	1	29	136.252	366	330	407
395	23	36	141.283	381	328	406

396	20	44	146.327	374	382	410
397	5	50	152.275	373	371	401
398	8	22	158.706	347	354	407
399	15	26	165.695	357	392	412
400	2	27	173.536	350	391	408
401	5	54	181.601	397	361	413
402	18	108	190.236	386	362	413
403	3	13	198.971	365	390	414
404	7	9	209.043	375	387	409
405	12	46	220.039	385	388	412
406	23	69	231.328	395	370	411
407	1	8	243.232	394	398	410
408	2	14	255.507	400	389	418
409	7	35	272.656	404	378	415
410	1	20	289.960	407	396	416
411	21	23	317.974	393	406	415
412	12	15	346.611	405	399	414
413	5	18	380.046	401	402	416
414	3	12	433.082	403	412	419
415	7	21	499.274	409	411	418
416	1	5	571.698	410	413	417
417	1	358	666.883	416	344	420
418	2	7	809.108	408	415	419
419	2	3	1011.517	418	414	420
420	1	2	1251.748	417	419	0

Cluster Membership

Case	8 Clusters	7 Clusters	6 Clusters	5 Clusters	4 Clusters	3 Clusters
438	3	3	3	3	3	3
441	1	1	1	1	1	1
443	3	3	3	3	3	3
444	3	3	3	3	3	3
445	7	6	5	4	4	2
446	3	3	3	3	3	3
447	6	3	3	3	3	3
449	6	3	3	3	3	3
450	3	3	3	3	3	3

451	6	3	3	3	3	3
453	7	6	5	4	4	2
454	8	7	6	5	1	1
455	8	7	6	5	1	1
456	5	5	5	4	4	2
458	2	2	2	2	2	2
459	1	1	1	1	1	1
460	7	6	5	4	4	2
461	5	5	5	4	4	2
463	7	6	5	4	4	2
464	6	3	3	3	3	3
465	2	2	2	2	2	2
466	7	6	5	4	4	2
467	1	1	1	1	1	1
468	2	2	2	2	2	2
469	3	3	3	3	3	3
470	7	6	5	4	4	2
471	4	4	4	1	1	1
472	1	1	1	1	1	1
473	6	3	3	3	3	3
474	7	6	5	4	4	2
475	6	3	3	3	3	3
476	2	2	2	2	2	2
477	7	6	5	4	4	2
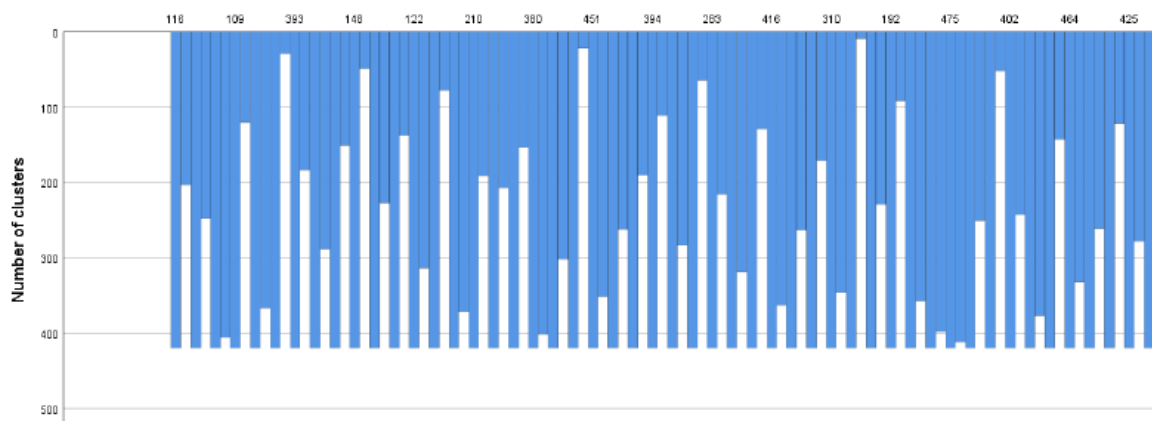
```
FREQUENCIES VARIABLES= CLU8_1 CLU7_1 CLU6_1 CLU5_1 CLU4_1 CLU3_1
  /ORDER=ANALYSIS.
```

Frequencies

Notes

Output Created		01-MAY-2019 14:32:22
Comments		
Input	Data	C:\Users\merxl\Downloads\natcom.sav
	Active Dataset	DataSet1
	File Label	CP05
	Filter	<none>
	Weight	<none>
	Split File	<none>
	N of Rows in Working Data File	477
Missing Value Handling	Definition of Missing	User-defined missing values are treated as missing.
	Cases Used	Statistics are based on all cases with valid data.
Syntax		FREQUENCIES VARIABLES= CLU8_1 CLU7_1 CLU6_1 CLU5_1 CLU4_1 CLU3_1 /ORDER=ANALYSIS.
Resources	Processor Time	00:00:00.02

Elapsed Time	00:00:00.02

Statistics

		Ward Method	Ward Method	Ward Method	Ward Method	Ward Method	
N	Valid	421	421	421	421	421	
	Missing	56	56	56	56	56	

Frequency Table

Ward Method

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1	62	13.0	14.7	14.7
	2	64	13.4	15.2	29.9
	3	54	11.3	12.8	42.8
	4	55	11.5	13.1	55.8
	5	36	7.5	8.6	64.4
	6	89	18.7	21.1	85.5
	7	53	11.1	12.6	98.1
	8	8	1.7	1.9	100.0
	Total	421	88.3	100.0	
Missing	System	56	11.7		
Total		477	100.0		

Ward Method

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1	62	13.0	14.7	14.7
	2	64	13.4	15.2	29.9
	3	143	30.0	34.0	63.9
	4	55	11.5	13.1	77.0
	5	36	7.5	8.6	85.5
	6	53	11.1	12.6	98.1

		Frequency	Percent	Valid Percent	Cumulative Percent
	7	8	1.7	1.9	100.0
	Total	421	88.3	100.0	
Missing	System	56	11.7		
Total		477	100.0		

Ward Method

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1	62	13.0	14.7	14.7
	2	64	13.4	15.2	29.9
	3	143	30.0	34.0	63.9
	4	55	11.5	13.1	77.0
	5	89	18.7	21.1	98.1
	6	8	1.7	1.9	100.0
	Total	421	88.3	100.0	
Missing	System	56	11.7		
Total		477	100.0		

Ward Method

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1	117	24.5	27.8	27.8
	2	64	13.4	15.2	43.0
	3	143	30.0	34.0	77.0
	4	89	18.7	21.1	98.1
	5	8	1.7	1.9	100.0
	Total	421	88.3	100.0	
Missing	System	56	11.7		
Total		477	100.0		

Ward Method

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1	125	26.2	29.7	29.7
	2	64	13.4	15.2	44.9

		Frequency	Percent	Valid Percent	Cumulative Percent
	3	143	30.0	34.0	78.9
	4	89	18.7	21.1	100.0
	Total	421	88.3	100.0	
Missing	System	56	11.7		
Total		477	100.0		

Ward Method

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1	125	26.2	29.7	29.7
	2	153	32.1	36.3	66.0
	3	143	30.0	34.0	100.0
	Total	421	88.3	100.0	
Missing	System	56	11.7		
Total		477	100.0		

```
DATASET ACTIVATE DataSet1.
RECODE q106 (2=1) (ELSE=0) INTO White.
VARIABLE LABELS  White 'percent white'.
EXECUTE.

DATASET ACTIVATE DataSet1.
ONEWAY ZPolActivity2 ZCosmo ZNeighTies2 q104 CivicEngagement Q57 Q21 Q60 Q79
Q82 Q100 Q107 Q109 White Q6 Q7 Q10 Q12 BY CLU4_1
  /MISSING ANALYSIS
  /POSTHOC=TUKEY SCHEFFE LSD ALPHA(0.05).
```

Oneway

Notes

Output Created		01-MAY-2019 14:32:22
Comments		
Input	Data	C:\Users\merxl\Downloads\natcom.sav
	Active Dataset	DataSet1
	File Label	CP05
	Filter	<none>
	Weight	<none>
	Split File	<none>

	N of Rows in Working Data File	477
Missing Value Handling	Definition of Missing	User-defined missing values are treated as missing.
	Cases Used	Statistics for each analysis are based on cases with no missing data for any variable in the analysis.

Syntax		ONEWAY ZPolActivity2 ZCosmo ZNeighTies2 q104 CivicEngagement Q57 Q21 Q60 Q79 Q82 Q100 Q107 Q109 White Q6 Q7 Q10 Q12 BY CLU4_1
		  /MISSING ANALYSIS
		  /POSTHOC=TUKEY SCHEFFE LSD ALPHA(0.05).

Resources	Processor Time	00:00:00.13
	Elapsed Time	00:00:00.11

ANOVA

		Sum of Squares	df	Mean Square	F	Sig.
Zscore: COMPUTE PolActivity2 = PolActivity1 + q91 + q92 (COMPUTE)	Between Groups	248.048	3	82.683	209.730	.000
	Within Groups	164.395	417	.394		
	Total	412.443	420			
Zscore: COMPUTE Cosmo = Zq97 + Zq98 (COMPUTE)	Between Groups	184.413	3	61.471	112.535	.000
	Within Groups	227.780	417	.546		
	Total	412.193	420			
Zscore: COMPUTE NeighTies2 = Zq34 + Zq35 + Zq36 + Zq2 (COMPUTE)	Between Groups	152.405	3	50.802	77.116	.000
	Within Groups	274.707	417	.659		
	Total	427.112	420			
Ql04:Age	Between Groups	47.024	3	15.675	5.750	.001
	Within Groups	1125.786	413	2.726		
	Total	1172.811	416			
COMPUTE	Between Groups	504.845	3	168.282	91.687	.000

CivicEngagement = ZTotalOrgs + ZPolActivity1 (COMPUTE)	Within Groups	750.677	409	1.835		
	Total	1255.522	412			
Q57:Comfort talking w/stranger on street	Between Groups	9.967	3	3.322	3.796	.010
	Within Groups	360.588	412	.875		
	Total	370.555	415			
Q21:Comfort voicing complaints public meeting	Between Groups	468.105	3	156.035	14.797	.000
	Within Groups	4270.672	405	10.545		
	Total	4738.778	408			
Q60:Freq read community newspaper	Between Groups	70.855	3	23.618	7.865	.000
	Within Groups	1222.186	407	3.003		
	Total	1293.041	410			
Q79:Perceived pol. knowledge	Between Groups	324.878	3	108.293	15.783	.000
	Within Groups	2813.228	410	6.862		
	Total	3138.106	413			
Q82:No.people talk pol regularly	Between Groups	70.402	3	23.467	18.604	.000
	Within Groups	523.483	415	1.261		
	Total	593.885	418			
Q100:Freq watch TV news	Between Groups	28.594	3	9.531	2.893	.035
	Within Groups	1370.654	416	3.295		
	Total	1399.248	419			
Q107:Household income	Between Groups	77.139	3	25.713	5.263	.001
	Within Groups	1670.791	342	4.885		
	Total	1747.931	345			
Q109:Gender	Between Groups	3.101	3	1.034	4.232	.006
	Within Groups	101.110	414	.244		
	Total	104.211	417			
percent white	Between Groups	.865	3	.288	1.523	.208
	Within Groups	78.940	417	.189		
	Total	79.805	420			
Q6:Value family	Between Groups	10.935	3	3.645	1.088	.354
	Within Groups	1377.393	411	3.351		
	Total	1388.328	414			
Q7:Value work	Between Groups	12.701	3	4.234	.509	.676
	Within Groups	3150.453	379	8.313		
	Total	3163.154	382			
Ql0:Value religion	Between Groups	102.618	3	34.206	3.268	.021
	Within Groups	4312.322	412	10.467		
	Total	4414.940	415			

Q12:Value being	Between Groups	75.182	3	25.061	3.537	.015
American	Within Groups	2932.897	414	7.084		
	Total	3008.079	417			

Post Hoc Tests

Multiple Comparisons

Dependent Variable		(I) Ward Method	(J) Ward Method	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
Zscore: COMPUTE PolActivity2 = PolActivity1 + q91 + q92 (COMPUTE)	Tukey HSD	1	2	1.89612649*	.09650787	.000	1.6471922	2.1450608
			3	1.47525902*	.07688139	.000	1.2769496	1.6735684
			4	1.74975215*	.08708307	.000	1.5251283	1.9743760
		2	1	-1.89612649*	.09650787	.000	-2.1450608	-1.6471922
			3	-.42086747*	.09442870	.000	-.6644387	-.1772962
			4	-.14637434	.10290526	.486	-.4118102	.1190615
		3	1	-1.47525902*	.07688139	.000	-1.6735684	-1.2769496
			2	.42086747*	.09442870	.000	.1772962	.6644387
			4	.27449313*	.08477305	.007	.0558278	.4931584
		4	1	-1.74975215*	.08708307	.000	-1.9743760	-1.5251283
			2	.14637434	.10290526	.486	-.1190615	.4118102

	(I)	(J)	Mean Difference (I-J)	Std. Error	Sig.	Lower Bound	Upper Bound
		3	-.27449313*	.08477305	.007	-.4931584	-.0558278
Scheffe	1	2	1.89612649*	.09650787	.000	1.6252349	2.1670181
		3	1.47525902*	.07688139	.000	1.2594577	1.6910603
		4	1.74975215*	.08708307	.000	1.5053154	1.9941889
	2	1	-1.89612649*	.09650787	.000	-2.1670181	-1.6252349
		3	-.42086747*	.09442870	.000	-.6859230	-.1558120
		4	-.14637434	.10290526	.568	-.4352230	.1424743
	3	1	-1.47525902*	.07688139	.000	-1.6910603	-1.2594577
		2	.42086747*	.09442870	.000	.1558120	.6859230
		4	.27449313*	.08477305	.016	.0365404	.5124458
	4	1	-1.74975215*	.08708307	.000	-1.9941889	-1.5053154
		2	.14637434	.10290526	.568	-.1424743	.4352230
		3	-.27449313*	.08477305	.016	-.5124458	-.0365404
LSD	1	2	1.89612649*	.09650787	.000	1.7064239	2.0858290
		3	1.47525902*	.07688139	.000	1.3241357	1.6263824
		4	1.74975215*	.08708307	.000	1.5785757	1.9209286

			Mean Difference	Std. Error	Sig.	Lower Bound	Upper Bound
	2	1	-1.89612649*	.09650787	.000	-2.0858290	-1.7064239
		3	-.42086747*	.09442870	.000	-.6064830	-.2352519
		4	-.14637434	.10290526	.156	-.3486520	.0559033
	3	1	-1.47525902*	.07688139	.000	-1.6263824	-1.3241357
		2	.42086747*	.09442870	.000	.2352519	.6064830
		4	.27449313*	.08477305	.001	.1078574	.4411289
	4	1	-1.74975215*	.08708307	.000	-1.9209286	-1.5785757
		2	.14637434	.10290526	.156	-.0559033	.3486520
		3	-.27449313*	.08477305	.001	-.4411289	-.1078574
Q100:Freq watch TV news	Tukey HSD	1	2	.055	.279	.997	-.67	.78
			3	-.572	.223	.051	-1.15	.00
			4	-.273	.252	.701	-.92	.38
		2	1	-.055	.279	.997	-.78	.67
			3	-.627	.273	.100	-1.33	.08
			4	-.328	.297	.689	-1.10	.44
		3	1	.572	.223	.051	.00	1.15
			2	.627	.273	.100	-.08	1.33
			4	.300	.245	.613	-.33	.93
		4	1	.273	.252	.701	-.38	.92
			2	.328	.297	.689	-.44	1.10
			3	-.300	.245	.613	-.93	.33
	Scheffe	1	2	.055	.279	.998	-.73	.84
			3	-.572	.223	.087	-1.20	.05
			4	-.273	.252	.760	-.98	.44
		2	1	-.055	.279	.998	-.84	.73

		(I)	(J)	Mean Difference	Std. Error	Sig.	Lower Bound	Upper Bound
			3	-.627	.273	.154	-1.39	.14
			4	-.328	.297	.750	-1.16	.51
		3	1	.572	.223	.087	-.05	1.20
			2	.627	.273	.154	-.14	1.39
			4	.300	.245	.684	-.39	.99
		4	1	.273	.252	.760	-.44	.98
			2	.328	.297	.750	-.51	1.16
			3	-.300	.245	.684	-.99	.39
	LSD	1	2	.055	.279	.844	-.49	.60
			3	-.572*	.223	.011	-1.01	-.13
			4	-.273	.252	.280	-.77	.22
		2	1	-.055	.279	.844	-.60	.49
			3	-.627*	.273	.022	-1.16	-.09
			4	-.328	.297	.271	-.91	.26
		3	1	.572*	.223	.011	.13	1.01
			2	.627*	.273	.022	.09	1.16
			4	.300	.245	.222	-.18	.78
		4	1	.273	.252	.280	-.22	.77
			2	.328	.297	.271	-.26	.91
			3	-.300	.245	.222	-.78	.18
QI0:Value religion	Tukey HSD	1	2	-1.177	.500	.088	-2.47	.11
			3	-1.018	.398	.053	-2.05	.01
			4	-1.099	.450	.071	-2.26	.06
		2	1	1.177	.500	.088	-.11	2.47
			3	.159	.491	.988	-1.11	1.42
			4	.078	.534	.999	-1.30	1.45
		3	1	1.018	.398	.053	-.01	2.05
			2	-.159	.491	.988	-1.42	1.11
			4	-.081	.440	.998	-1.22	1.05
		4	1	1.099	.450	.071	-.06	2.26
			2	-.078	.534	.999	-1.45	1.30
			3	.081	.440	.998	-1.05	1.22
	Scheffe	1	2	-1.177	.500	.138	-2.58	.23
			3	-1.018	.398	.090	-2.14	.10
			4	-1.099	.450	.115	-2.36	.16
		2	1	1.177	.500	.138	-.23	2.58
			3	.159	.491	.991	-1.22	1.54
			4	.078	.534	.999	-1.42	1.58

	3	1		1.018	.398	.090	-.10	2.14
		2		-.159	.491	.991	-1.54	1.22
		4		-.081	.440	.998	-1.32	1.15
	4	1		1.099	.450	.115	-.16	2.36
		2		-.078	.534	.999	-1.58	1.42
		3		.081	.440	.998	-1.15	1.32
LSD	1	2		-1.177*	.500	.019	-2.16	-.19
		3		-1.018*	.398	.011	-1.80	-.24
		4		-1.099*	.450	.015	-1.98	-.21
	2	1		1.177*	.500	.019	.19	2.16
		3		.159	.491	.747	-.81	1.12
		4		.078	.534	.885	-.97	1.13
	3	1		1.018*	.398	.011	.24	1.80
		2		-.159	.491	.747	-1.12	.81
		4		-.081	.440	.854	-.95	.78
	4	1		1.099*	.450	.015	.21	1.98
		2		-.078	.534	.885	-1.13	.97
		3		.081	.440	.854	-.78	.95

*. The mean difference is significant at the 0.05 level.

Homogeneous Subsets

Zscore: COMPUTE PolActivity2 = PolActivity1 + q91 + q92 (COMPUTE)

	Ward Method	N	Subset for alpha = 0.05		
			1	2	3
Tukey HSD[a,b]	2	64	-.7497875		
	4	89	-.6034131		
	3	143		-.3289200	
	1	125			1.1463390
	Sig.		.373	1.000	1.000
Scheffe[a,b]	2	64	-.7497875		
	4	89	-.6034131		
	3	143		-.3289200	
	1	125			1.1463390
	Sig.		.459	1.000	1.000

Means for groups in homogeneous subsets are displayed.

a. Uses Harmonic Mean Sample Size = 95.570.

b. The group sizes are unequal. The harmonic mean of the group sizes is used. Type I error levels are not guaranteed.

Zscore: COMPUTE   Cosmo = Zq97 + Zq98 (COMPUTE)

	Ward Method	N	Subset for alpha = 0.05		
			1	2	3
Tukey HSD[a,b]	4	89	-1.1896205		
	1	125		-.0181914	
	2	64			.4304713
	3	143			.5706180
	Sig.		1.000	1.000	.557
Scheffe[a,b]	4	89	-1.1896205		
	1	125		-.0181914	
	2	64			.4304713
	3	143			.5706180
	Sig.		1.000	1.000	.633

Means for groups in homogeneous subsets are displayed.

a. Uses Harmonic Mean Sample Size = 95.570.

b. The group sizes are unequal. The harmonic mean of the group sizes is used. Type I error levels are not guaranteed.

Zscore: COMPUTE   NeighTies2 =   Zq34 + Zq35 + Zq36 + Zq2 (COMPUTE)

	Ward Method	N	Subset for alpha = 0.05		
			1	2	3
Tukey HSD[a,b]	2	64	-1.1702535		
	4	89		-.0866558	
	1	125		-.0760447	
	3	143			.6657909
	Sig.		1.000	1.000	1.000
Scheffe[a,b]	2	64	-1.1702535		
	4	89		-.0866558	
	1	125		-.0760447	
	3	143			.6657909
	Sig.		1.000	1.000	1.000

Means for groups in homogeneous subsets are displayed.

a. Uses Harmonic Mean Sample Size = 95.570.

b. The group sizes are unequal. The harmonic mean of the group sizes is used. Type I error levels are not guaranteed.

QI04:Age

| | Ward Method | N | Subset for alpha = 0.05 | |
			1	2
Tukey HSD[a,b]	2	64	3.83	
	1	123	4.31	4.31
	4	88		4.45
	3	142		4.82
	Sig.		.187	.140
Scheffe[a,b]	2	64	3.83	
	1	123	4.31	4.31
	4	88	4.45	4.45
	3	142		4.82
	Sig.		.079	.204

Means for groups in homogeneous subsets are displayed.

a. Uses Harmonic Mean Sample Size = 94.874.

b. The group sizes are unequal. The harmonic mean of the group sizes is used. Type I error levels are not guaranteed.

COMPUTE   CivicEngagement = ZTotalOrgs + ZPolActivity1 (COMPUTE)

| | Ward Method | N | Subset for alpha = 0.05 | | |
			1	2	3
Tukey HSD[a,b]	2	63	-1.2493		
	4	88	-.9415		
	3	140		-.2997	
	1	122			1.6095
	Sig.		.405	1.000	1.000
Scheffe[a,b]	2	63	-1.2493		
	4	88	-.9415		
	3	140		-.2997	
	1	122			1.6095
	Sig.		.490	1.000	1.000

Means for groups in homogeneous subsets are displayed.

a. Uses Harmonic Mean Sample Size = 93.949.

b. The group sizes are unequal. The harmonic mean of the group sizes is used. Type I error levels are not guaranteed.

Q57:Comfort talking w/stranger on street

	Ward Method	N	Subset for alpha = 0.05	
			1	2
Tukey HSD[a,b]	4	88	2.88	
	2	64	2.95	2.95
	3	140	3.17	3.17
	1	124		3.27
	Sig.		.130	.099
Scheffe[a,b]	4	88	2.88	
	2	64	2.95	2.95
	3	140	3.17	3.17
	1	124		3.27
	Sig.		.192	.153

Means for groups in homogeneous subsets are displayed.

a. Uses Harmonic Mean Sample Size = 94.796.

b. The group sizes are unequal. The harmonic mean of the group sizes is used. Type I error levels are not guaranteed.

Q21:Comfort voicing complaints public meeting

	Ward Method	N	Subset for alpha = 0.05		
			1	2	3
Tukey HSD[a,b]	4	85	4.80		
	2	63	5.35		
	3	138		6.59	
	1	123		7.60	
	Sig.		.657	.150	
Scheffe[a,b]	4	85	4.80		
	2	63	5.35	5.35	
	3	138		6.59	6.59
	1	123			7.60
	Sig.		.722	.079	.216

Means for groups in homogeneous subsets are displayed.

a. Uses Harmonic Mean Sample Size = 92.993.

b. The group sizes are unequal. The harmonic mean of the group sizes is used. Type I error levels are not guaranteed.

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
Zscore: COMPUTE PolActivity2 = PolActivity1 + q91 + q92 (COMPUTE)	434	-1.40773	2.42031	.0000000	1.00000000
Zscore: COMPUTE Cosmo = Zq97 + Zq98 (COMPUTE)	432	-3.01281	1.07481	.0000000	1.00000000
Zscore: COMPUTE NeighTies2 = Zq34 + Zq35 + Zq36 + Zq2 (COMPUTE)	454	-2.36967	2.09446	.0000000	1.00000000
Ql04:Age	433	1	7	4.45	1.673
COMPUTE CivicEngagement = ZTotalOrgs + ZPolActivity1 (COMPUTE)	425	-2.60	5.62	.0026	1.77060
Q57:Comfort talking w/stranger on street	447	1	4	3.08	.961
Q21:Comfort voicing complaints public meeting	449	0	10	6.31	3.440
Q60:Freq read community newspaper	440	0	5	3.23	1.782
Q79:Perceived pol. knowledge	441	0	10	6.06	2.793
Q82:No.people talk pol regularly	446	0	4	1.72	1.195
Q100:Freq watch TV news	434	0	6	4.27	1.822
Q107:Household income	359	1	9	4.72	2.244
Q109:Gender	441	1	2	1.53	.500
percent white	477	.00	1.00	.6730	.46963
Q6:Value family	464	0	10	9.32	1.882
Q7:Value work	425	0	10	7.05	2.908
Ql0:Value religion	462	0	10	7.51	3.183
Q12:Value being American	467	0	10	8.32	2.679
Valid N (listwise)	293				

IV. Tabling

Table 1. Cluster Profiling

Cluster Name	HMM Pol Actives	LHL Cosmopolites (Only)	MHH Worldly Neighbors	LLM Non-Pol Localites	Total	F	Sig
Internal Variables	1(125)	2(64)	3(143)	4(89)	421		
ZPolActivity2 = (q83,q84,q85,q86,q87, q89,q90,q91,q92)	1.146[c]	-0.750[a]	-0.329[b]	-0.603[a]	0	209.7	<.001
Zcosmo = (Zq94, Zq98)	-0.018[b]	0.430[c]	0.571[c]	-1.190[a]	0	112.5	<.001
ZNeighTies2 = (Zq34, Zq35, Zq36, Zq2)	-0.076[b]	-1.170[a]	0.666[b]	-0.087[b]	0	77.1	<.001
External Variables							
Age (1-7 Scale)	4.31[a,b]	3.83[a]	4.82[b]	4.45[b]	4.45	5.75	0.001
CivicEngagement (Standardized)	1.61[c]	-1.25[a]	-0.30[b]	-0.94[a]	0	91.7	<.001
Comfort Talking w/ Strangers on Street (1-4 Scale)	3.27[a]	2.95[a,b]	3.17[a,b]	2.88[a]	3.08	3.80	0.010
Comfort Voicing Complaints at Public Meetings (0-10 Scale)	7.60[b]	5.35[a]	6.59[b]	4.80[a]	6.31	14.8	<.001
Freq Read Community Newspaper (0-5 Scale)	3.39[b,c]	2.62[a]	3.65[c]	2.75[a,b]	3.23	7.87	<.001
Percieved Political Knowledge (0-10 Scale)	7.24[c]	5.43[a,b]	6.15[b]	4.83[a]	6.06	15.8	<.001
No. People talk Politics with Regularly (0-4 Scale)	2.31[b]	1.27[a]	1.65[a]	1.34[a]	1.72	18.6	<.001
Freq Watch TV News* (0-6 Scale)	4.01[a]	3.95[b]	4.58[b]	4.28[a,b]	4.27	2.89	0.035
Household Income (1-9 Scale)	5.33[b]	4.48[a,b]	4.69[a,b]	3.97[a]	4.72	5.26	0.001
Gender (% Female)	39%[a]	58%[a,b]	59%[b]	56%[a,b]	53%	4.23	0.006
Race (% White)	78%	64%	76%	75%	67%	1.52	0.208
Value Family (0-10 Scale)	9.19	9.63	9.31	9.52	9.32	1.09	0.354
Value Work (0-10 Scale)	7.15	7.32	6.81	7.08	7.05	0.51	0.676
Value Religion* (0-10 Scale)	6.70[a]	7.87[b]	7.71[b]	7.80[b]	7.51	3.27	0.021
Value Being an American (0-10 Scale)	7.79[a]	8.08[a,b]	8.63[a,b]	8.83[b]	8.32	3.54	0.015

NOTE: Means that do not share a superscript are significantly different at $p < .05$, via the LSD, Scheffe, and Tukey post hoc tests.
* - Indicates that the post hoc difference of means is significant only for the LSC (not Scheffe nor Tukey)

V. Write-Up

The National Community Study (Jeffres, 2006) was chosen for cluster analysis. Three internal variables were used, each of which was an additive scale. Scale one, named Political Activity, includes ten items all measured with Yes/No responses: attended meetings of your town or city council (Q83), attended a political meeting or rally (Q84), worn a button or put a sticker on your car (Q85), voted in the 2004 presidential election (Q86), participated in a march or rally (Q87), helped circulate or signed a petition (Q88), solicited political funds (Q89), contributed money to a party or candidate (Q90), contacted a public official about some issue (Q91), and emailed, written a letter, or called media about some public issue (Q92) (alpha = .75). Scale two, named Cosmopoliteness, includes the standardized scores of two items measured on a 7-point Likert-type scale (0 = "Completely Disagree"; 10 = "Completely Agree"): I enjoy learning about other peoples and cultures (ZQ97) and I think of myself as a citizen of the world (ZQ98) (alpha = .54). Scale three, name Neighborhood Ties, includes the standardized scores of four items which are measured on a variety of Likert-type scales: Of the ten closest neighbors, how many of their home have you visited? (ZQ34), How many of your ten closest neighbors do you know by name or well enough to say hello when you see them on the street? (ZQ35), What percentage of your closest friends live in the same community? (ZQ36), and How long have you lived in your neighborhood or community? (ZQ2) (alpha = .81).

The fifteen external or "profiling" variables include: Age, Civic Engagement, comfort talking with strangers on the street (Q57), comfort voicing complaints at public meetings (Q21), frequently read community newspaper (Q60), perceived political knowledge (Q79), number of people talk politics with regularly (Q82), frequently watch TV news (Q60), household income, gender (% female), race (% white), value family (Q6), value work (Q7), value religion (Q10),

and value being an American (Q12).

A hierarchical agglomerative cluster analysis was performed to discover the natural grouping of participants. A four-cluster solution was chosen using Ward's Method (with Squared Euclidean Distance). This choice of four clusters was chosen due to size requirements and was supported by an examination of changes in the agglomeration coefficients from the agglomeration table. An icicle plot was run to give a visual representation of the data clusters. A Frequency procedure was run to examine cluster sizes to ensure all clusters had a reasonable *n*. A One-Way ANOVA was run to examine the differences among the four clusters with regard to all three internal variables. As expected, all internal variables were significantly different among the four clusters. The four clusters have been named: "Politically Actives", "Cosmopolites (Only)", "Worldly Neighbors", and "Non-Political Localites" (See Table 1). To further profile the four clusters, a complementary set of ANOVA analyses were conducted to test the significance of the differences among the four clusters against the fifteen external variables. All four of the internal variables showed highly significant differences across the four clusters ($p < .001$). Of the external variables, all showed significant differences ($p < .05$) across the four clusters except for race, value family, and value work.

Cluster 1 ($n = 125$) is labeled "Politically Actives" because this group ranked highest in political activity, but was slightly lower than average for cosmopoliteness and neighborhood ties. This cluster was significantly different from all others in many ways, ranking highest in civil engagement, perceived political knowledge, and number of people talk politics with regularly, as well as ranking lowest in value religion. While not significantly so, this group also tends to be the wealthiest, whitest, least female, and values being an American the least. Cluster 2 ($n = 64$) is labelled "Cosmopolites (Only)" because it is above average in cosmopoliteness but the lowest of the four clusters in political activity and neighborhood ties. While not significantly so, this

cluster is the youngest, least white, and least civically engaged. This group also values family, work, and religion more than any other, although, again, this is not significant. Cluster 3 (n = 143) is labelled "Worldly Neighbors" because it is above average in cosmopoliteness and neighborhood ties. While not significantly so, this cluster is the oldest, most female, and reads community newspapers and watches TV news the most frequently. Cluster 4 (n = 89) is labelled "Non-Political Localites" because they are average in neighborhood ties and below average in political activity and cosmopoliteness. While not significantly so, this cluster is lowest in comfort with talking to strangers on the street, comfort voicing complaints at public meetings, perceived political knowledge, and household income, while highest in value being an American.