Automated Data Analysis Report

Table of Contents

1. Clustering Results

Best Parameters: {'epsilon': 2.393369097964607, 'min_samples': 6, 'silhouette': 0.33287232534725236}, Best Silhouette Score: 0.333

2. ANOVA Results

Results for wife_religion: F-value = 20296.886, P-value = 0.000
Tukey-HSD Test Results: Multiple Comparison of Means - Tukey HSD, FWER=0.05
== group1 group2 meandiff p-adj lower upper reject ---------- -1 1 1.8704 0.0 1.7661 1.9746 True -1 3 1.8704 0.0 1.7712 1.9696 True -1 4 1.8704 0.0 1.7697 1.971 True -1 5 -0.9352 0.0 -1.0377 -0.8327 True -1 6 -0.9352 0.0 -1.0422 -0.8282 True 1 3 -0.0 1.0 -0.0359 0.0359 False 1 4 -0.0 1.0 -0.0398 0.0398 False 1 5 -2.8055 0.0 -2.8498 -2.7613 True 1 6 -2.8055 0.0 -2.8593 -2.7518 True 3 4 0.0 1.0 -0.0237 0.0237 False 3 5 -2.8055 0.0 -2.8361 -2.7749 True 3 6 -2.8055 0.0 -2.8488 -2.7623 True 4 5 -2.8055 0.0 -2.8406 -2.7704 True 4 6 -2.8055 0.0 -2.8521 -2.759 True 5 6 -0.0 1.0 -0.0504 0.0504 False ----------

Results for wife_working: F-value = 3630.540, P-value = 0.000
Tukey-HSD Test Results: Multiple Comparison of Means - Tukey HSD, FWER=0.05
== group1 group2 meandiff p-adj lower upper reject ---------- -1 1 0.8207 0.0 0.582 1.0594 True -1 3 1.3462 0.0 1.119 1.5735 True -1 4 -0.9616 0.0 -1.1922 -0.731 True -1 5 1.3462 0.0 1.1114 1.5811 True -1 6 -0.9616 0.0 -1.2067 -0.7165 True 1 3 0.5255 0.0 0.4434 0.6077 True 1 4 -1.7823 0.0 -1.8734 -1.6912 True 1 5 0.5255 0.0 0.4242 0.6269 True 1 6 -1.7823 0.0 -1.9055 -1.6591 True 3 4 -2.3078 0.0 -2.3621 -2.2536 True 3 5 -0.0 1.0 -0.0701 0.0701 False 3 6 -2.3078 0.0 -2.4069 -2.2087 True 4 5 2.3078 0.0 2.2274 2.3882 True 4 6 0.0 1.0 -0.1066 0.1066 False 5 6 -2.3078 0.0 -2.4233 -2.1923 True ----------

Results for media_exposure: F-value = 10811.880, P-value = 0.000
Tukey-HSD Test Results: Multiple Comparison of Means - Tukey HSD, FWER=0.05
== group1 group2 meandiff p-adj lower upper reject ---------- -1 1 1.2734 0.0 1.1315 1.4153 True -1 3 -2.5468 0.0 -2.6819 -2.4117 True -1 4 -2.5468 0.0 -2.6839 -2.4097 True -1 5 -2.5468 0.0 -2.6864 -2.4072 True -1 6 -2.5468 0.0 -2.6925 -2.4011 True 1 3 -3.8202 0.0 -3.869 -3.7713 True 1 4 -3.8202 0.0 -3.8743 -3.766 True 1 5 -3.8202 0.0 -3.8804 -3.7599 True 1 6 -3.8202 0.0 -3.8934 -3.7469 True 3 4 -0.0 1.0 -0.0323 0.0323 False 3 5 -0.0 1.0 -0.0417 0.0417 False 3 6 -0.0 1.0 -0.0589 0.0589 False 4 5 0.0 1.0 -0.0478 0.0478 False 4 6 0.0 1.0 -0.0634 0.0634 False 5 6 0.0 1.0 -0.0687 0.0687 False ----------

Results for age_children_interaction: F-value = 15.977, P-value = 0.000
Tukey-HSD Test Results: Multiple Comparison of Means - Tukey HSD, FWER=0.05
== group1 group2 meandiff p-adj lower upper reject ---------- -1 1 -1.2763 0.0003 -2.1265 -0.4261 True -1 3 -1.7529 0.0 -2.5622 -0.9436 True -1 4 -1.9621 0.0 -2.7835 -1.1407 True -1 5 -1.8269 0.0 -2.6633 -0.9905 True -1 6 -1.9877 0.0 -2.8605 -1.1148 True 1 3 -0.4766 0.0001 -0.7692 -0.184 True 1 4 -0.6858 0.0 -1.0103 -0.3614 True 1 5 -0.5506 0.0002 -0.9115 -0.1897 True 1 6 -0.7114 0.0001 -1.1501 -0.2727 True 3 4 -0.2092 0.025 -0.4025 -0.016 True 3 5 -0.074 0.9587 -0.3237 0.1756 False 3 6 -0.2348 0.4035 -0.5878 0.1181 False 4 5 0.1352 0.7583 -0.1511 0.4215 False 4 6 -0.0256 1.0 -0.4054 0.3542 False 5 6 -0.1608 0.8751 -0.5721 0.2505 False ----------

Results for edu_interaction: F-value = 55.231, P-value = 0.000

Tukey-HSD Test Results: Multiple Comparison of Means - Tukey HSD, FWER=0.05 == group1 group2 meandiff p-adj lower upper reject -- -1 1 -1.1287 0.0009 -1.9296 -0.3278 True -1 3 0.0767 0.9997 -0.6857 0.8391 False -1 4 0.1563 0.9926 -0.6174 0.9301 False -1 5 0.6846 0.1309 -0.1033 1.4725 False -1 6 0.7676 0.0832 -0.0546 1.5899 False 1 3 1.2054 0.0 0.9298 1.4811 True 1 4 1.285 0.0 0.9794 1.5907 True 1 5 1.8133 0.0 1.4733 2.1533 True 1 6 1.8964 0.0 1.4831 2.3097 True 3 4 0.0796 0.8133 -0.1025 0.2617 False 3 5 0.6079 0.0 0.3727 0.843 True 3 6 0.6909 0.0 0.3584 1.0234 True 4 5 0.5283 0.0 0.2586 0.798 True 4 6 0.6113 0.0 0.2536 0.9691 True 5 6 0.0831 0.9902 -0.3044 0.4705 False --

3. Cluster Variability

		antecedent support	consequent support	support	confidence	lift	\
	count	8.000000	8.000000	8.000000	8.000000	8.000000	
	mean	0.166524	0.387272	0.096747	0.607547	1.566041	
	std	0.135460	0.047679	0.072409	0.105193	0.162729	
	min	0.075342	0.328767	0.047945	0.456311	1.387945	
	25%	0.086758	0.333333	0.053653	0.559852	1.465085	
	50%	0.104452	0.409247	0.061644	0.594118	1.496000	
	75%	0.171518	0.428082	0.104737	0.662126	1.715358	
	max	0.428082	0.428082	0.213470	0.769231	1.796923	

		leverage	conviction	zhangs_metric	total_items	coverage
	count	8.000000	8.000000	8.000000	8.00000	8.000000
	mean	0.033805	1.653912	0.435987	2.75000	0.166524
	std	0.023634	0.423086	0.096721	0.46291	0.135460
	min	0.013738	1.234589	0.311414	2.00000	0.075342
	25%	0.016914	1.404793	0.349562	2.75000	0.086758
	50%	0.026111	1.521366	0.465875	3.00000	0.104452
	75%	0.040476	1.738544	0.492161	3.00000	0.171518
3	max	0.070776	2.478311	0.579718	3.00000	0.428082

		antecedent support	consequent support	support	confidence	lift	\
	count	8.000000	8.000000	8.000000	8.000000	8.000000	
	mean	0.125000	0.363511	0.077206	0.627152	1.750577	
	std	0.034488	0.072536	0.019847	0.107882	0.272198	
	min	0.084559	0.264706	0.051471	0.466667	1.565566	
	25%	0.095588	0.330882	0.055147	0.590909	1.575758	
	50%	0.123162	0.375000	0.084559	0.606265	1.646041	
	75%	0.161765	0.380515	0.095588	0.652174	1.789041	
	max	0.161765	0.496324	0.095588	0.851852	2.365217	

		leverage	conviction	zhangs_metric	total_items	coverage
	count	8.000000	8.000000	8.000000	8.000000	8.000000
	mean	0.031292	1.863555	0.476686	2.875000	0.125000
	std	0.004928	0.660548	0.068971	0.353553	0.034488
	min	0.022275	1.378676	0.418131	2.000000	0.084559
	25%	0.029314	1.527778	0.435897	3.000000	0.095588
	50%	0.033379	1.560633	0.449626	3.000000	0.123162
	75%	0.034926	1.923713	0.491658	3.000000	0.161765
4	max	0.035291	3.399816	0.630522	3.000000	0.161765

	antecedent support	consequent support	support	confidence	lift	\
count	0.0	0.0	0.0	0.0	0.0	
mean	NaN	NaN	NaN	NaN	NaN	
std	NaN	NaN	NaN	NaN	NaN	
min	NaN	NaN	NaN	NaN	NaN	
25%	NaN	NaN	NaN	NaN	NaN	
50%	NaN	NaN	NaN	NaN	NaN	
75%	NaN	NaN	NaN	NaN	NaN	
max	NaN	NaN	NaN	NaN	NaN	

	leverage	conviction	zhangs_metric	total_items
count	0.0	0.0	0.0	0.0
mean	NaN	NaN	NaN	NaN
std	NaN	NaN	NaN	NaN
min	NaN	NaN	NaN	NaN
25%	NaN	NaN	NaN	NaN
50%	NaN	NaN	NaN	NaN
75%	NaN	NaN	NaN	NaN
1 max	NaN	NaN	NaN	NaN

	antecedent support	consequent support	support	confidence	lift	\
count	0.0	0.0	0.0	0.0	0.0	
mean	NaN	NaN	NaN	NaN	NaN	
std	NaN	NaN	NaN	NaN	NaN	
min	NaN	NaN	NaN	NaN	NaN	
25%	NaN	NaN	NaN	NaN	NaN	
50%	NaN	NaN	NaN	NaN	NaN	
75%	NaN	NaN	NaN	NaN	NaN	
max	NaN	NaN	NaN	NaN	NaN	

	leverage	conviction	zhangs_metric	total_items
count	0.0	0.0	0.0	0.0
mean	NaN	NaN	NaN	NaN
std	NaN	NaN	NaN	NaN
min	NaN	NaN	NaN	NaN
25%	NaN	NaN	NaN	NaN
50%	NaN	NaN	NaN	NaN
75%	NaN	NaN	NaN	NaN
5 max	NaN	NaN	NaN	NaN

	antecedent support	consequent support	support	confidence	lift	\
	count	0.0	0.0	0.0	0.0	0.0
	mean	NaN	NaN	NaN	NaN	NaN
	std	NaN	NaN	NaN	NaN	NaN
	min	NaN	NaN	NaN	NaN	NaN
	25%	NaN	NaN	NaN	NaN	NaN
	50%	NaN	NaN	NaN	NaN	NaN
	75%	NaN	NaN	NaN	NaN	NaN
	max	NaN	NaN	NaN	NaN	NaN

		leverage	conviction	zhangs_metric	total_items
	count	0.0	0.0	0.0	0.0
	mean	NaN	NaN	NaN	NaN
	std	NaN	NaN	NaN	NaN
	min	NaN	NaN	NaN	NaN
	25%	NaN	NaN	NaN	NaN
	50%	NaN	NaN	NaN	NaN
	75%	NaN	NaN	NaN	NaN
6	max	NaN	NaN	NaN	NaN

	antecedent support	consequent support	support	confidence	\
count	29.000000	29.000000	29.000000	29.0	
mean	0.097701	0.442529	0.097701	1.0	
std	0.032035	0.099860	0.032035	0.0	
min	0.083333	0.250000	0.083333	1.0	
25%	0.083333	0.416667	0.083333	1.0	
50%	0.083333	0.500000	0.083333	1.0	
75%	0.083333	0.500000	0.083333	1.0	
max	0.166667	0.500000	0.166667	1.0	

		lift	leverage	conviction	zhangs_metric	total_items	coverage
	count	29.000000	29.000000	29.0	29.000000	29.0	29.000000
	mean	2.427586	0.054837	inf	0.619122	3.0	0.097701
	std	0.775931	0.022298	NaN	0.116632	0.0	0.032035
	min	2.000000	0.041667	inf	0.545455	3.0	0.083333
	25%	2.000000	0.041667	NaN	0.545455	3.0	0.083333
	50%	2.000000	0.041667	NaN	0.545455	3.0	0.083333
	75%	2.400000	0.062500	NaN	0.636364	3.0	0.083333
-1	max	4.000000	0.125000	inf	0.900000	3.0	0.166667

4. Rule Metrics Comparison

mean	std	min	25%	50%	75%
075467511729772	0.10519266426148435	0.4563106796116505	0.5598518518518518	0.5941176470588236	0.662126400996
271519048692962	0.10788154109625285	0.4666666666666667	0.5909090909090908	0.6062653562653563	0.652173913043
nan	nan	nan	nan	nan	nan
nan	nan	nan	nan	nan	nan
nan	nan	nan	nan	nan	nan

1.0	0.0	1.0	1.0	1.0	1.0
679802563448316	0.1380484715492421	0.3770491803278688	0.6289384502656165	0.7142857142857142	0.782462004162

5. Top Unique Rules per Cluster

Cluster 3:

Rule: frozenset({'age_children_interaction_(164.0, 768.0]', 'edu_interaction_(12.0, 16.0]'}) -> frozenset({'standard_of_living_index_4'}) (Support: 0.055, Confidence: 0.727, Lift: 1.699)
Rule: frozenset({'edu_interaction_(12.0, 16.0]'}) -> frozenset({'standard_of_living_index_4'}) (Support: 0.213, Confidence: 0.640, Lift: 1.496)
Rule: frozenset({'age_children_interaction_(42.0, 87.0]', 'edu_interaction_(12.0, 16.0]'}) -> frozenset({'standard_of_living_index_4'}) (Support: 0.048, Confidence: 0.600, Lift: 1.402)
Rule: frozenset({'standard_of_living_index_3', 'age_children_interaction_(42.0, 87.0]'}) -> frozenset({'edu_interaction_(6.0, 12.0]'}) (Support: 0.054, Confidence: 0.580, Lift: 1.486)
Rule: frozenset({'standard_of_living_index_4'}) -> frozenset({'edu_interaction_(12.0, 16.0]'}) (Support: 0.213, Confidence: 0.499, Lift: 1.496)

Cluster 4:

Rule: frozenset({'edu_interaction_(6.0, 12.0]', 'age_children_interaction_(42.0, 87.0]'}) -> frozenset({'husband_occupation_2'}) (Support: 0.055, Confidence: 0.652, Lift: 2.365)
Rule: frozenset({'husband_occupation_2', 'age_children_interaction_(42.0, 87.0]'}) -> frozenset({'edu_interaction_(6.0, 12.0]'}) (Support: 0.055, Confidence: 0.652, Lift: 1.867)
Rule: frozenset({'standard_of_living_index_2'}) -> frozenset({'husband_occupation_3'}) (Support: 0.096, Confidence: 0.591, Lift: 1.576)
Rule: frozenset({'standard_of_living_index_2', 'Cluster_(3.0, 4.0]'}) -> frozenset({'husband_occupation_3'}) (Support: 0.096, Confidence: 0.591, Lift: 1.576)
Rule: frozenset({'standard_of_living_index_2'}) -> frozenset({'husband_occupation_3', 'Cluster_(3.0, 4.0]'}) (Support: 0.096, Confidence: 0.591, Lift: 1.576)

Cluster 1:

Cluster 5:

Cluster 6:

Cluster -1:

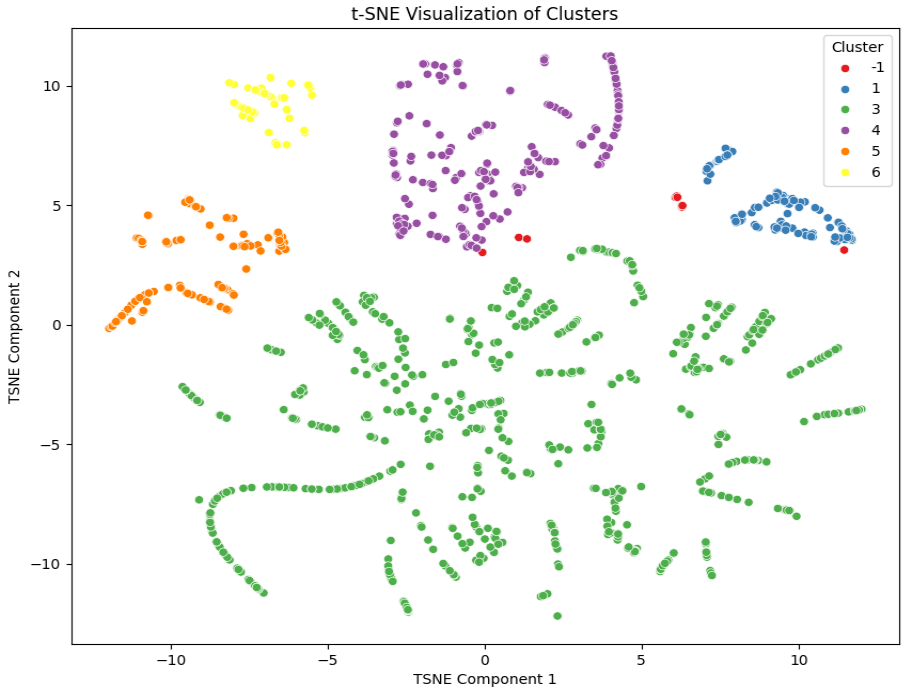
Rule: frozenset({'husband_occupation_3', 'age_children_interaction_(164.0, 768.0]'}) -> frozenset({'edu_interaction_(6.0, 12.0]'}) (Support: 0.083, Confidence: 1.000, Lift: 2.000)
Rule: frozenset({'standard_of_living_index_3', 'husband_occupation_2'}) -> frozenset({'edu_interaction_(6.0, 12.0]'}) (Support: 0.167, Confidence: 1.000, Lift: 2.000)

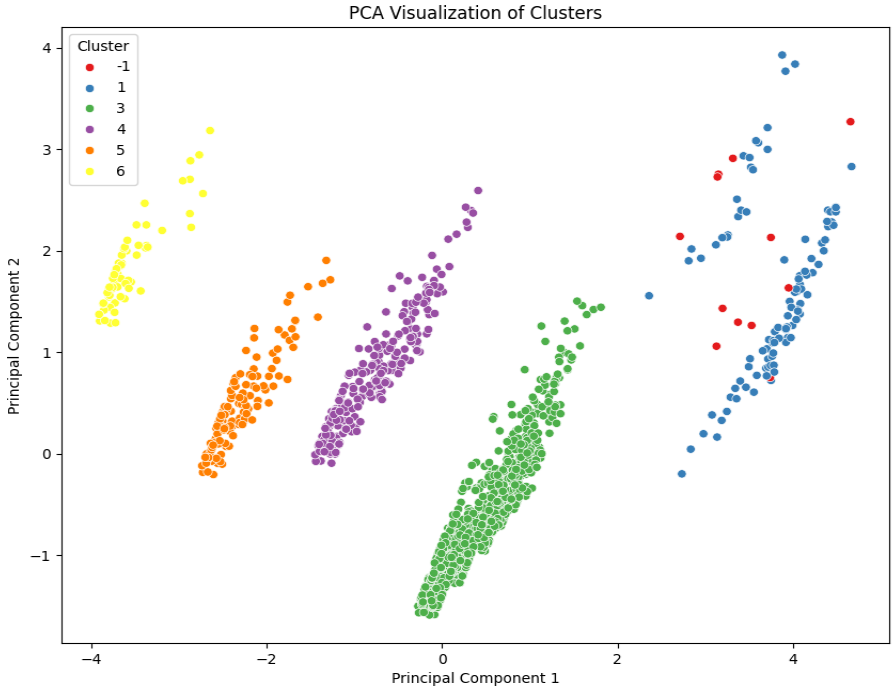
Rule: frozenset({'standard_of_living_index_3', 'edu_interaction_(12.0, 16.0]'}) ->
frozenset({'husband_occupation_3'}) (Support: 0.083, Confidence: 1.000, Lift: 2.000)
Rule: frozenset({'standard_of_living_index_3', 'husband_occupation_3'}) ->
frozenset({'edu_interaction_(12.0, 16.0]'}) (Support: 0.083, Confidence: 1.000, Lift: 4.000)
Rule: frozenset({'edu_interaction_(12.0, 16.0]', 'age_children_interaction_(87.0, 164.0]'}) ->
frozenset({'husband_occupation_3'}) (Support: 0.083, Confidence: 1.000, Lift: 2.000)

6. Top 10 Common Rules Sorted by Absolute Coverage Difference

Rule: frozenset({'standard_of_living_index_4', 'edu_interaction_(12.0, 16.0]',
'age_children_interaction_(87.0, 164.0]'}) (Abs Coverage Difference: 0.051)
Rule: frozenset({'standard_of_living_index_4', 'edu_interaction_(12.0, 16.0]',
'age_children_interaction_(87.0, 164.0]'}) (Abs Coverage Difference: 0.047)
Rule: frozenset({'standard_of_living_index_4', 'edu_interaction_(12.0, 16.0]',
'age_children_interaction_(87.0, 164.0]'}) (Abs Coverage Difference: 0.041)
Rule: frozenset({'standard_of_living_index_4', 'edu_interaction_(12.0, 16.0]',
'age_children_interaction_(87.0, 164.0]'}) (Abs Coverage Difference: 0.029)
Rule: frozenset({'standard_of_living_index_4', 'edu_interaction_(12.0, 16.0]',
'age_children_interaction_(87.0, 164.0]'}) (Abs Coverage Difference: 0.023)
Rule: frozenset({'standard_of_living_index_4', 'edu_interaction_(12.0, 16.0]',
'age_children_interaction_(87.0, 164.0]'}) (Abs Coverage Difference: 0.020)
Rule: frozenset({'standard_of_living_index_4', 'edu_interaction_(12.0, 16.0]',
'age_children_interaction_(87.0, 164.0]'}) (Abs Coverage Difference: 0.018)
Rule: frozenset({'standard_of_living_index_4', 'edu_interaction_(12.0, 16.0]',
'age_children_interaction_(87.0, 164.0]'}) (Abs Coverage Difference: 0.017)
Rule: frozenset({'standard_of_living_index_4', 'edu_interaction_(12.0, 16.0]',
'age_children_interaction_(87.0, 164.0]'}) (Abs Coverage Difference: 0.010)
Rule: frozenset({'standard_of_living_index_4', 'edu_interaction_(12.0, 16.0]',
'age_children_interaction_(87.0, 164.0]'}) (Abs Coverage Difference: 0.010)

7. Cluster Visualizations


t-SNE Visualization of Clusters
Cluster
-1
1
3
4
5
6
TSNE Component 2
TSNE Component 1
10
5
0
−5
−10
−10
−5
0
5
10


PCA Visualization of Clusters
Cluster
-1
1
3
4
5
6
Principal Component 2
Principal Component 1
4
3
2
1
0
−1
−4
−2
0
2
4