Automated Data Analysis Report

Table of Contents

1. Clustering Results

Best Parameters: {'epsilon': 3.6292714046523606, 'min_samples': 4, 'silhouette': 0.3172086854992034}, Best Silhouette Score: 0.317

Train Silhouette Score: 0.317, Test Silhouette Score: 0.335

2. ANOVA Results

Results for wife_age: F-value = 12.908, 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 ---------- 0 1 -0.3845 0.0003 -0.5945 -0.1745 True ----------

Results for num_children: F-value = 19.372, 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 ---------- 0 1 -0.4698 0.0 -0.6792 -0.2604 True ----------

Results for wife_religion: F-value = 5.549, P-value = 0.019
Tukey-HSD Test Results: Multiple Comparison of Means - Tukey HSD, FWER=0.05
== group1 group2 meandiff p-adj lower upper reject ---------- 0 1 -0.2529 0.0186 -0.4635 -0.0423 True ----------

Results for media_exposure: F-value = inf, 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 ---------- 0 1 -3.6903 0.0 -3.6903 -3.6903 True ----------

Results for age_children_interaction: F-value = 20.310, 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 ---------- 0 1 -0.4808 0.0 -0.6902 -0.2715 True ----------

Results for edu_interaction: F-value = 152.790, 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 ---------- 0 1 1.2514 0.0 1.0527 1.45 True ----------

3. Cluster Variability

	antecedent support	consequent support	support	confidence	\
count	36.000000	3.600000e+01	36.000000	36.0	
mean	0.021277	5.319149e-02	0.021277	1.0	
std	0.015676	7.037323e-18	0.015676	0.0	
min	0.010638	5.319149e-02	0.010638	1.0	
25%	0.010638	5.319149e-02	0.010638	1.0	
50%	0.010638	5.319149e-02	0.010638	1.0	
75%	0.021277	5.319149e-02	0.021277	1.0	
max	0.053191	5.319149e-02	0.053191	1.0	

	lift	leverage	conviction	zhangs_metric	total_items	\
count	3.600000e+01	36.000000	36.0	36.000000	36.000000	
mean	1.880000e+01	0.020145	inf	0.967638	2.944444	
std	3.603109e-15	0.014843	NaN	0.015828	0.232311	
min	1.880000e+01	0.010072	inf	0.956989	2.000000	
25%	1.880000e+01	0.010072	NaN	0.956989	3.000000	
50%	1.880000e+01	0.010072	NaN	0.956989	3.000000	
75%	1.880000e+01	0.020145	NaN	0.967391	3.000000	
max	1.880000e+01	0.050362	inf	1.000000	3.000000	

		coverage
	count	36.000000
	mean	0.021277
	std	0.015676
	min	0.010638
	25%	0.010638
	50%	0.010638
	75%	0.021277
0	max	0.053191

	antecedent support	consequent support	support	confidence	\
count	56.000000	56.000000	56.000000	56.000000	
mean	0.079599	0.206593	0.058942	0.788020	
std	0.039463	0.040478	0.023889	0.165349	
min	0.026753	0.122694	0.025830	0.439759	
25%	0.052352	0.153137	0.047740	0.694544	
50%	0.068727	0.233395	0.058118	0.760692	
75%	0.086716	0.233395	0.067343	0.985714	
max	0.233395	0.248155	0.152214	1.000000	

		lift	leverage	conviction	zhangs_metric	total_items	coverage
	count	56.000000	56.000000	56.000000	56.000000	56.000000	56.000000
	mean	3.860137	0.043264	inf	0.798565	2.964286	0.079599
	std	0.638008	0.018162	NaN	0.056736	0.187256	0.039463
	min	2.958051	0.018116	1.565945	0.701087	2.000000	0.026753
	25%	3.210243	0.035511	2.577905	0.749096	3.000000	0.052352
	50%	3.957374	0.039979	3.441767	0.802143	3.000000	0.068727
	75%	4.284585	0.049396	53.662362	0.839435	3.000000	0.086716
1	max	5.224096	0.116473	inf	0.998155	3.000000	0.233395

4. Rule Metrics Comparison

nt	mean	std	min	25%	50%	
0	1.0	0.0	1.0	1.0	1.0	
0	0.7880204008988905	0.16534902408176852	0.43975903614457834	0.6945436507936508	0.7606918238993711	0.98571
0	0.9120349159602523	0.10767059241948497	0.7	0.8314270152505447	0.9402985074626865	

5. Top Unique Rules per Cluster

Cluster 0:

Rule: frozenset({'standard_of_living_index_3', 'wife_edu_2'}) -> frozenset({'husband_occupation_4'}) (Support: 0.011, Confidence: 1.000, Lift: 18.800)
Rule: frozenset({'wife_age_(26.0, 32.0]', 'standard_of_living_index_3'}) -> frozenset({'husband_occupation_4'}) (Support: 0.011, Confidence: 1.000, Lift: 18.800)
Rule: frozenset({'age_children_interaction_(90.0, 168.0]', 'edu_interaction_(12.0, 16.0]'}) -> frozenset({'wife_edu_4'}) (Support: 0.021, Confidence: 1.000, Lift: 18.800)
Rule: frozenset({'wife_edu_4', 'wife_age_(39.0, 49.0]'}) -> frozenset({'edu_interaction_(12.0, 16.0]'}) (Support: 0.011, Confidence: 1.000, Lift: 18.800)
Rule: frozenset({'wife_age_(39.0, 49.0]', 'edu_interaction_(12.0, 16.0]'}) -> frozenset({'wife_edu_4'}) (Support: 0.011, Confidence: 1.000, Lift: 18.800)

Cluster 1:

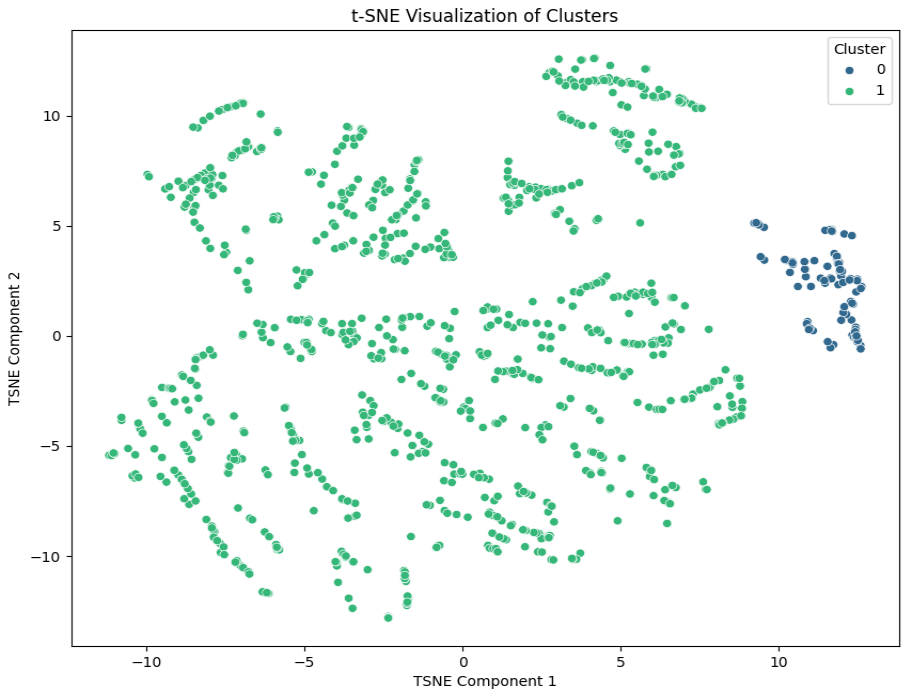
Rule: frozenset({'wife_age_(39.0, 49.0]', 'num_children_(2.0, 3.0]'}) -> frozenset({'age_children_interaction_(90.0, 168.0]'}) (Support: 0.032, Confidence: 1.000, Lift: 4.030)
Rule: frozenset({'wife_age_(26.0, 32.0]', 'num_children_(3.0, 5.0]'}) -> frozenset({'age_children_interaction_(90.0, 168.0]'}) (Support: 0.067, Confidence: 1.000, Lift: 4.030)
Rule: frozenset({'wife_age_(32.0, 39.0]', 'num_children_(2.0, 3.0]'}) -> frozenset({'age_children_interaction_(90.0, 168.0]'}) (Support: 0.054, Confidence: 1.000, Lift: 4.030)
Rule: frozenset({'num_children_(5.0, 16.0]', 'wife_age_(32.0, 39.0]'}) -> frozenset({'age_children_interaction_(168.0, 768.0]'}) (Support: 0.053, Confidence: 1.000, Lift: 4.285)
Rule: frozenset({'husband_occupation_2', 'num_children_(5.0, 16.0]'}) -> frozenset({'age_children_interaction_(168.0, 768.0]'}) (Support: 0.044, Confidence: 1.000, Lift: 4.285)

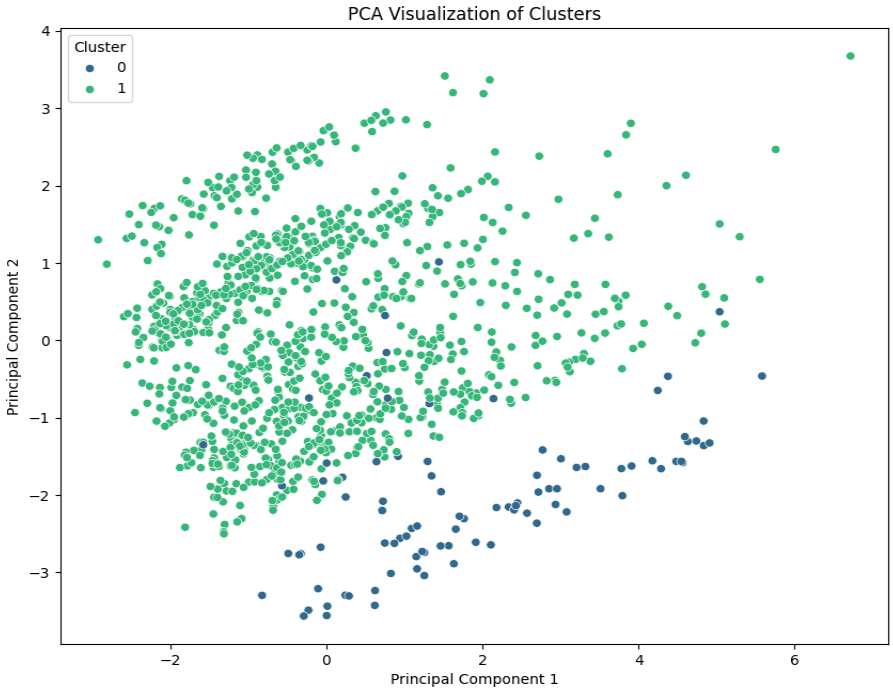
6. Top 10 Common Rules Sorted by Absolute Coverage Difference

Rule: frozenset({'wife_edu_4', 'edu_interaction_(12.0, 16.0]'}) (Abs Coverage Difference: 0.338)
Rule: frozenset({'wife_edu_4', 'edu_interaction_(12.0, 16.0]'}) (Abs Coverage Difference: 0.338)
Rule: frozenset({'husband_edu_4', 'wife_edu_4', 'edu_interaction_(12.0, 16.0]'}) (Abs Coverage Difference: 0.338)
Rule: frozenset({'husband_edu_4', 'wife_edu_4', 'edu_interaction_(12.0, 16.0]'}) (Abs Coverage Difference: 0.338)
Rule: frozenset({'husband_edu_4', 'wife_edu_4', 'edu_interaction_(12.0, 16.0]'}) (Abs Coverage Difference: 0.338)
Rule: frozenset({'husband_edu_4', 'wife_edu_4', 'edu_interaction_(12.0, 16.0]'}) (Abs Coverage Difference: 0.338)

Rule: frozenset({'husband_edu_4', 'wife_edu_4', 'edu_interaction_(12.0, 16.0]'}) (Abs Coverage Difference: 0.314)
Rule: frozenset({'husband_edu_4', 'wife_edu_4', 'edu_interaction_(12.0, 16.0]'}) (Abs Coverage Difference: 0.314)
Rule: frozenset({'husband_edu_4', 'wife_edu_4', 'edu_interaction_(12.0, 16.0]'}) (Abs Coverage Difference: 0.314)
Rule: frozenset({'husband_edu_4', 'wife_edu_4', 'edu_interaction_(12.0, 16.0]'}) (Abs Coverage Difference: 0.314)

7. Cluster Visualizations




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