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 = 25388.402, 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 1.7706 1.9701 True -1 2 -0.9352 0.0 -1.0321 -0.8383 True -1 3 1.8704 0.0 1.7754 1.9653 True -1 4 1.8704 0.0 1.774 1.9667 True 1 2 -2.8055 0.0 -2.845 -2.7661 True 1 3 -0.0 1.0 -0.0343 0.0343 False 1 4 -0.0 1.0 -0.0381 0.0381 False 2 3 2.8055 0.0 2.7805 2.8305 True 2 4 2.8055 0.0 2.7756 2.8355 True 3 4 0.0 1.0 -0.0227 0.0227 False ----------

Results for wife_working: F-value = 1159.187, 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 0.411 1.2303 True -1 2 0.6169 0.0002 0.2188 1.015 True -1 3 1.3462 0.0 0.9563 1.7362 True -1 4 -0.9616 0.0 -1.3573 -0.5659 True 1 2 -0.2038 0.0056 -0.366 -0.0416 True 1 3 0.5255 0.0 0.3846 0.6665 True 1 4 -1.7823 0.0 -1.9386 -1.626 True 2 3 0.7294 0.0 0.6267 0.832 True 2 4 -1.5785 0.0 -1.7014 -1.4556 True 3 4 -2.3078 0.0 -2.401 -2.2147 True ----------

Results for media_exposure: F-value = 13524.062, 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 1.1376 1.4092 True -1 2 -2.5468 0.0 -2.6787 -2.4148 True -1 3 -2.5468 0.0 -2.676 -2.4175 True -1 4 -2.5468 0.0 -2.6779 -2.4156 True 1 2 -3.8202 0.0 -3.8739 -3.7664 True 1 3 -3.8202 0.0 -3.8669 -3.7734 True 1 4 -3.8202 0.0 -3.872 -3.7684 True 2 3 0.0 1.0 -0.034 0.034 False 2 4 -0.0 1.0 -0.0407 0.0407 False 3 4 -0.0 1.0 -0.0309 0.0309 False ----------

Results for age_children_interaction: F-value = 19.657, 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 0.0002 -2.0901 -0.4624 True -1 2 -1.8777 0.0 -2.6686 -1.0868 True -1 3 -1.7529 0.0 -2.5275 -0.9782 True -1 4 -1.9621 0.0 -2.7483 -1.1759 True 1 2 -0.6014 0.0 -0.9237 -0.2792 True 1 3 -0.4766 0.0 -0.7567 -0.1965 True 1 4 -0.6858 0.0 -0.9964 -0.3752 True 2 3 0.1248 0.4521 -0.0792 0.3289 False 2 4 -0.0844 0.8797 -0.3286 0.1598 False 3 4 -0.2092 0.0175 -0.3942 -0.0242 True ----------

Results for edu_interaction: F-value = 68.975, 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 0.0006 -1.8951 -0.3623 True -1 2 0.7108 0.0697 -0.034 1.4557 False -1 3 0.0767 0.9985 -0.6528 0.8063 False -1 4 0.1563 0.9785 -0.5841 0.8967 False 1 2 1.8395 0.0 1.5361 2.143 True 1 3 1.2054 0.0 0.9417 1.4692 True 1 4 1.285 0.0 0.9925 1.5775 True 2 3 -0.6341 0.0 -0.8262 -0.442 True 2 4 -0.5545 0.0 -0.7845 -0.3246 True 3 4 0.0796 0.7232 -0.0946 0.2538 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	\
count	22.000000	22.000000	22.000000	22.000000	
mean	0.166517	0.250675	0.069307	0.448865	
std	0.103736	0.103635	0.050485	0.166932	
min	0.039604	0.148515	0.029703	0.235294	
25%	0.071782	0.168317	0.029703	0.301389	
50%	0.163366	0.207921	0.049505	0.420833	
75%	0.198020	0.336634	0.099010	0.597059	
max	0.356436	0.485149	0.198020	0.750000	

		lift	leverage	conviction	zhangs_metric	total_items	coverage
	count	22.000000	22.000000	22.000000	22.000000	22.000000	22.000000
	mean	1.838268	0.029797	1.449535	0.530245	2.636364	0.166517
	std	0.372597	0.021153	0.328799	0.124024	0.492366	0.103736
	min	1.485294	0.009705	1.108911	0.347368	2.000000	0.039604
	25%	1.555517	0.016175	1.155106	0.423868	2.000000	0.071782
	50%	1.666830	0.019116	1.344554	0.550575	3.000000	0.163366
	75%	2.117388	0.043280	1.597772	0.622761	3.000000	0.198020
1	max	2.754545	0.078032	2.122772	0.769335	3.000000	0.356436

	antecedent support	consequent support	support	confidence	lift	\
count	2.000000	2.000000	2.000000	2.000000	2.000000	
mean	0.160377	0.266509	0.075472	0.470996	1.767141	
std	0.006671	0.010006	0.000000	0.019591	0.007161	
min	0.155660	0.259434	0.075472	0.457143	1.762078	
25%	0.158019	0.262972	0.075472	0.464069	1.764610	
50%	0.160377	0.266509	0.075472	0.470996	1.767141	
75%	0.162736	0.270047	0.075472	0.477922	1.769673	
max	0.165094	0.273585	0.075472	0.484848	1.772205	

		leverage	conviction	zhangs_metric	total_items	coverage
	count	2.000000	2.000000	2.000000	2.0	2.000000
	mean	0.032763	1.387150	0.517035	3.0	0.160377
	std	0.000173	0.032456	0.001377	0.0	0.006671
	min	0.032641	1.364201	0.516061	3.0	0.155660
	25%	0.032702	1.375675	0.516548	3.0	0.158019
	50%	0.032763	1.387150	0.517035	3.0	0.160377
	75%	0.032824	1.398625	0.517522	3.0	0.162736
2	max	0.032885	1.410100	0.518008	3.0	0.165094

		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%
5467511729772	0.10519266426148435	0.4563106796116505	0.5598518518518518	0.5941176470588235	0.6621264009
1519048692962	0.10788154109625285	0.4666666666666667	0.5909090909090908	0.6062653562653563	0.6521739130
36513145202983	0.16693177684232688	0.23529411764705882	0.3013888888888889	0.4208333333333334	0.5970588235
99567099567097	0.019590837227679257	0.45714285714285713	0.4640692640692641	0.47099567099567097	0.4779220779
1.0	0.0	1.0	1.0	1.0	1.0
5911364808293	0.12253010097450433	0.4113475177304965	0.5471247563352827	0.6484178353110392	0.7190061475

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:

Rule: frozenset({'standard_of_living_index_3', 'age_children_interaction_(42.0, 87.0]'}) -> frozenset({'husband_occupation_3'}) (Support: 0.030, Confidence: 0.750, Lift: 1.546)
Rule: frozenset({'standard_of_living_index_3', 'husband_occupation_2'}) -> frozenset({'age_children_interaction_(164.0, 768.0]'}) (Support: 0.050, Confidence: 0.714, Lift: 1.640)
Rule: frozenset({'age_children_interaction_(164.0, 768.0]', 'standard_of_living_index_2'}) -> frozenset({'husband_occupation_2'}) (Support: 0.109, Confidence: 0.688, Lift: 2.042)
Rule: frozenset({'husband_occupation_2', 'age_children_interaction_(164.0, 768.0]'}) -> frozenset({'standard_of_living_index_2'}) (Support: 0.109, Confidence: 0.611, Lift: 1.715)
Rule: frozenset({'edu_interaction_(6.0, 12.0]', 'husband_occupation_3'}) -> frozenset({'age_children_interaction_(87.0, 164.0]'}) (Support: 0.030, Confidence: 0.600, Lift: 2.755)

Cluster 2:

Cluster -1:

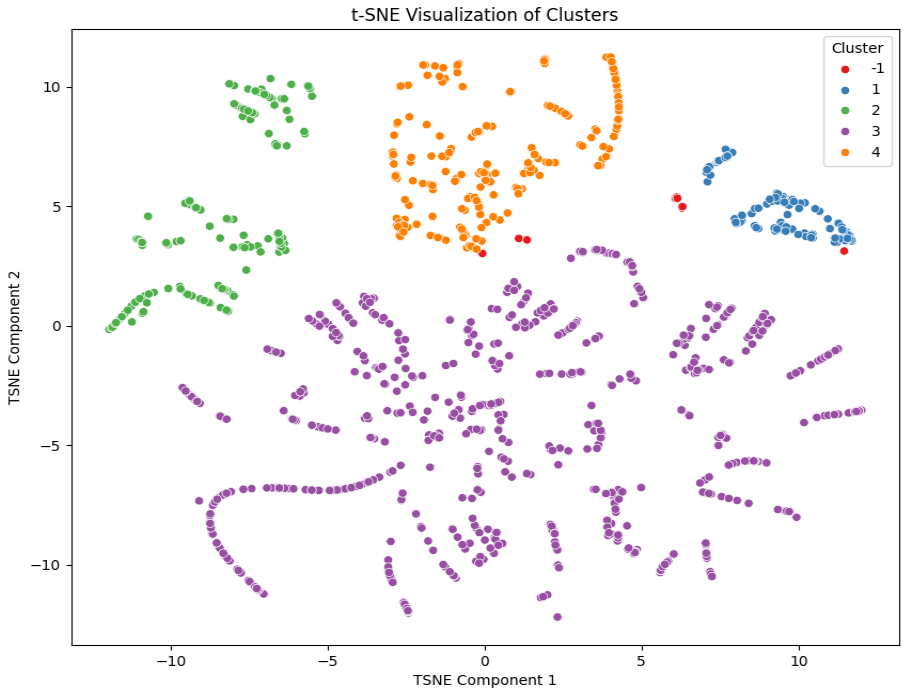
Rule: frozenset({'husband_occupation_3', 'age_children_interaction_(164.0, 768.0]'}) -> frozenset({'edu_interaction_(6.0, 12.0]'}) (Support: 0.083, Confidence: 1.000, Lift: 2.000)
Rule: frozenset({'edu_interaction_(6.0, 12.0]', 'husband_occupation_2'}) -> frozenset({'standard_of_living_index_3'}) (Support: 0.167, Confidence: 1.000, Lift: 4.000)
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)

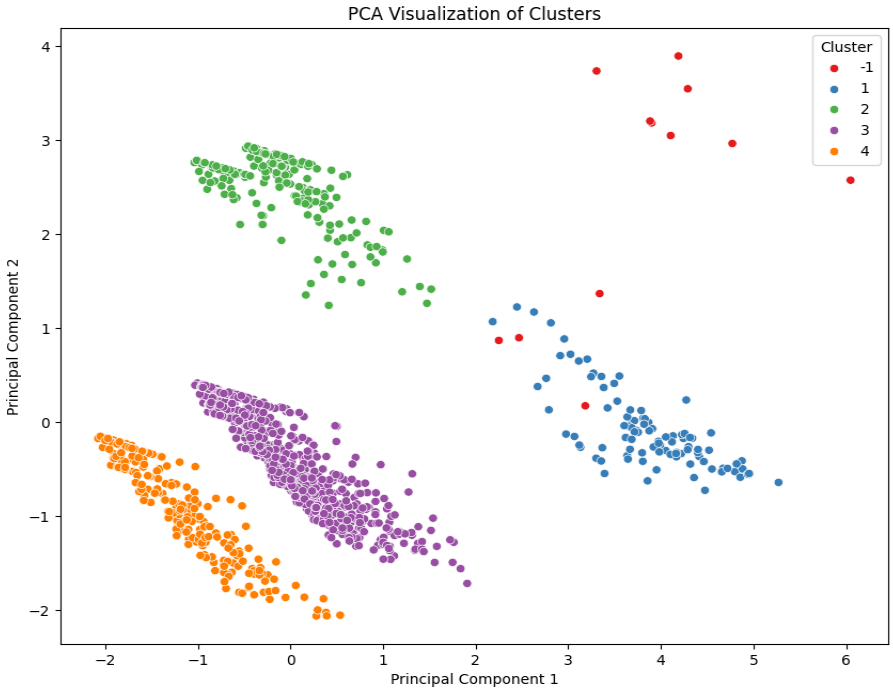
6. Top 10 Common Rules Sorted by Absolute Coverage Difference

Rule: frozenset({'edu_interaction_(12.0, 16.0]', 'standard_of_living_index_4'}) (Abs Coverage Difference: 0.131)
Rule: frozenset({'edu_interaction_(12.0, 16.0]', 'standard_of_living_index_4'}) (Abs Coverage Difference: 0.059)
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({'edu_interaction_(12.0, 16.0]', 'standard_of_living_index_4'}) (Abs Coverage Difference: 0.036)
Rule: frozenset({'edu_interaction_(12.0, 16.0]', 'standard_of_living_index_4'}) (Abs Coverage Difference: 0.036)
Rule: frozenset({'age_children_interaction_(42.0, 87.0]', 'edu_interaction_(12.0, 16.0]', 'standard_of_living_index_4'}) (Abs Coverage Difference: 0.030)
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)

7. Cluster Visualizations




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