Multseed macro for finding multiple cluster solutions with different starting seeds

The UNIVARIATE Procedure
Variable: OVER_ALL (Statistic Applying Over All Variables)

Moments			
N	100	Sum Weights	100
Mean	0.89241833	Sum Observations	89.2418335
Std Deviation	0.04622916	Variance	0.00213714
Skewness	0.60590473	Kurtosis	−1.3468191
Uncorrected SS	79.8526248	Corrected SS	0.2115764
Coeff Variation	5.18021196	Std Error Mean	0.00462292

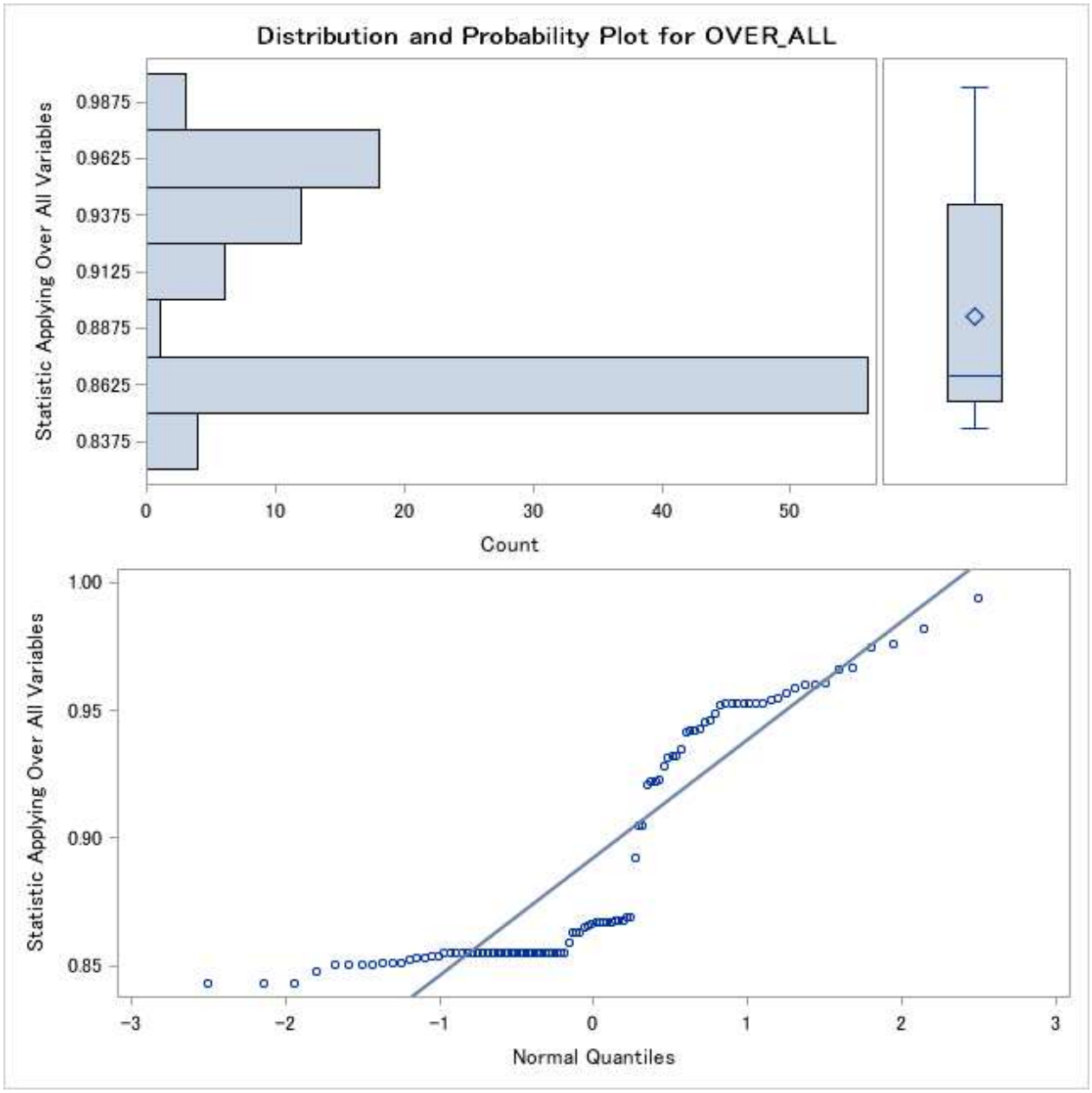
Basic Statistical Measures			
Location		Variability	
Mean	0.892418	Std Deviation	0.04623
Median	0.866572	Variance	0.00214
Mode	0.854992	Range	0.15112
		Interquartile Range	0.08727

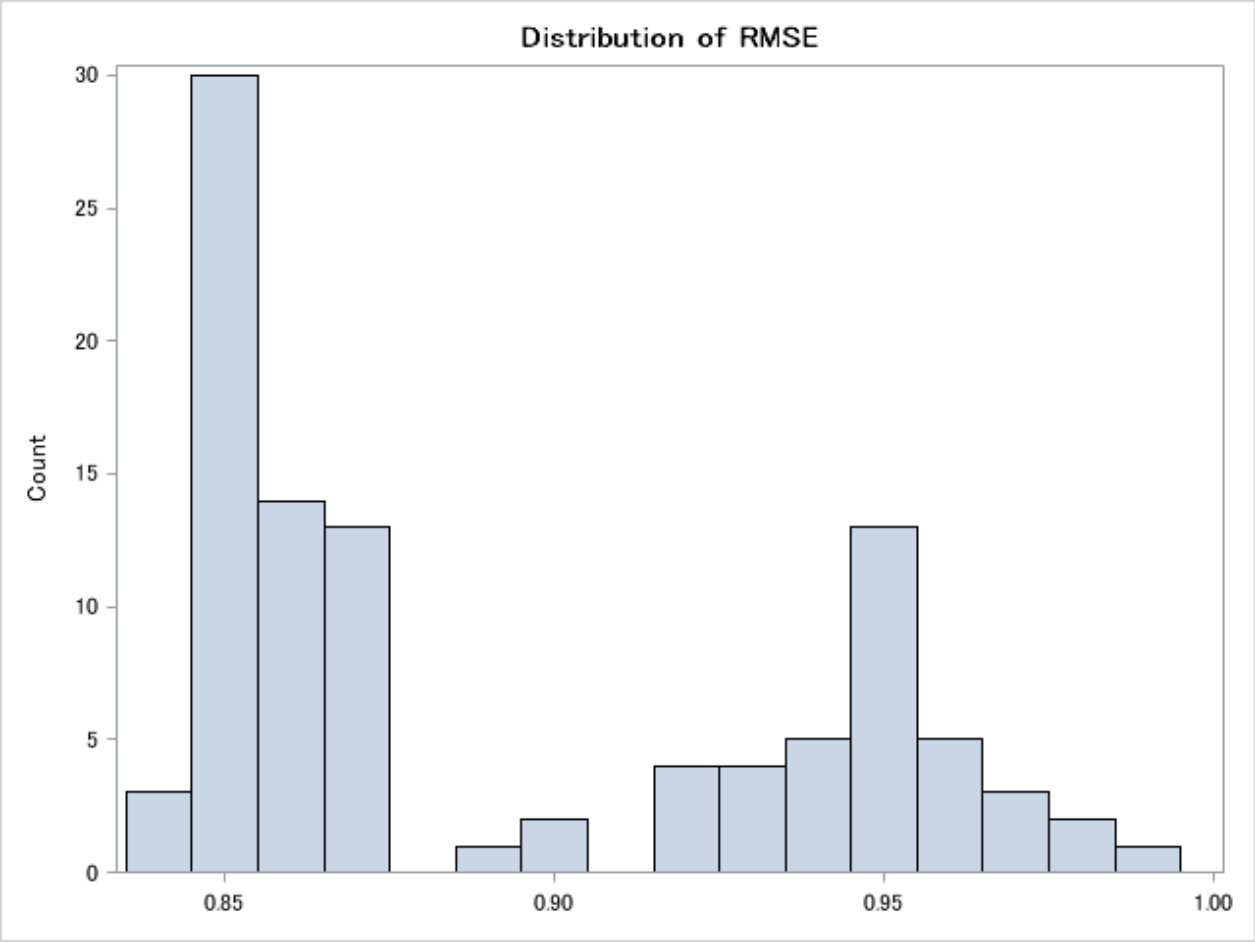
Tests for Location: Mu0=0				
Test	Statistic		p Value	
Student's t	t	193.0423	Pr > |t|	<.0001
Sign	M	50	Pr >= |M|	<.0001
Signed Rank	S	2525	Pr >= |S|	<.0001

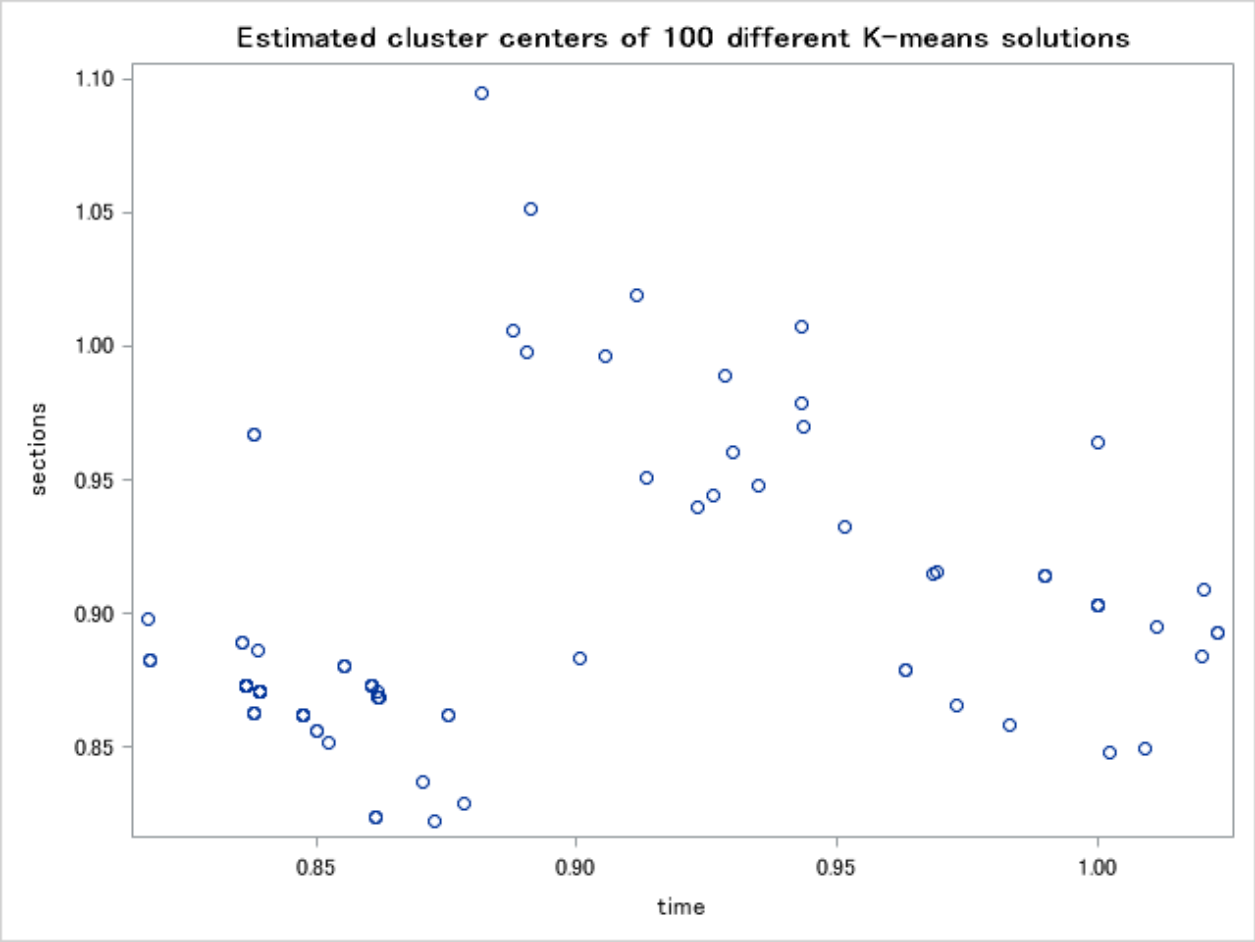
Quantiles (Definition 5)	
Level	Quantile
100% Max	0.994026
99%	0.988074
95%	0.966498
90%	0.957895
75% Q3	0.942261
50% Median	0.866572
25% Q1	0.854992
10%	0.851047
5%	0.850464
1%	0.842902
0% Min	0.842902

Extreme Observations	

Lowest			Highest		
Value	seed	Obs	Value	seed	Obs
0.842902	53	48	0.966878	41	60
0.842902	62	39	0.974660	81	20
0.842902	77	24	0.975765	51	50
0.847661	55	46	0.982123	95	6
0.850464	26	75	0.994026	46	55




Distribution of RMSE
Count
30
25
20
15
10
5
0
0.85
0.90
0.95
1.00


Estimated cluster centers of 100 different K-means solutions
sections
time
1.10
1.05
1.00
0.95
0.90
0.85
0.85
0.90
0.95
1.00

Best solution out of 100 seeds

The SURVEYSELECT Procedure

Selection Method	Simple Random Sampling

Input Data Set	NP
Random Number Seed	53
Sample Size	5
Selection Probability	0.09434
Sampling Weight	10.6
Output Data Set	SEEDS

Best solution out of 100 seeds

The FASTCLUS Procedure
Replace=FULL Drift Radius=0 Maxclusters=5 Maxiter=10 Converge=0 Least=2

Initial Seeds		
Cluster	time	sections
1	0.000000000	0.000000000
2	1.000000000	0.000000000
3	1.000000000	3.000000000
4	6.000000000	4.000000000
5	7.000000000	3.000000000

Minimum Distance Between Initial Seeds =	2.893742

Iteration History						
		Relative Change in Cluster Seeds				
Iteration	Criterion	1	2	3	4	5
1	0.9154	0.3297	0.3281	0.5919	0.1061	0.0666
2	0.8562	0.00950	0.1266	0.2789	0.1387	0
3	0.8450	0	0	0.0787	0.0904	0
4	0.8433	0	0	0.0808	0.0912	0
5	0.8422	0	0	0	0	0

Convergence criterion is satisfied.

Criterion Based on Final Seeds =	0.8422

Cluster Summary						
Cluster	Frequency	RMS Std Deviation	Maximum Distance from Seed to Observation	Radius Exceeded	Nearest Cluster	Distance Between Cluster Centroids
1	509	0.4478	1.7390		2	2.7843
2	725	0.8692	2.1947		1	2.7843
3	670	1.0265	2.2464		2	3.2549
4	607	0.8084	1.6397		5	3.2584
5	428	0.8855	1.9979		4	3.2584

Statistics for Variables				
Variable	Total STD	Within STD	R-Square	RSQ/(1-RSQ)
time	2.36950	0.86156	0.867971	6.574116
sections	2.25379	0.82382	0.866572	6.494669

OVER-ALL	2.31237	0.84290	0.867307	6.536171

Pseudo F Statistic =	4794.28

Approximate Expected Over-All R-Squared =	0.80081

Cubic Clustering Criterion =	28.706

WARNING: The two values above are invalid for correlated variables.

Cluster Means		
Cluster	time	sections
1	0.278978389	0.249508841
2	2.360000000	2.099310345
3	2.473134328	5.352238806
4	6.238879736	5.925864909
5	6.095794393	2.670560748

Cluster Standard Deviations		
Cluster	time	sections
1	0.461905222	0.433154464
2	1.018502855	0.688149146
3	1.039587432	1.013268949
4	0.762532733	0.851773457
5	0.754490740	0.999471834


Plot with the best means and cluster assignments
sections
time
0
2
4
6
Cluster 1 2 3 5 4 Center