

The CONTENTS Procedure

Data Set Name	WORK.IMPORT	Observations	1048562
Member Type	DATA	Variables	11
Engine	V9	Indexes	0
Created	06/13/2024 21:04:37	Observation Length	96
Last Modified	06/13/2024 21:04:37	Deleted Observations	0
Protection		Compressed	NO
Data Set Type		Sorted	NO
Label			
Data Representation	SOLARIS_X86_64, LINUX_X86_64, ALPHA_TRU64, LINUX_IA64		
Encoding	utf-8 Unicode (UTF-8)		

Engine/Host Dependent Information	
Data Set Page Size	131072
Number of Data Set Pages	770
First Data Page	1
Max Obs per Page	1363
Obs in First Data Page	1329
Number of Data Set Repairs	0
Filename	/saswork/SAS_workC21E00018D41_odaws01-apse1-2.oda.sas.com/SAS_workE6A600018D41_odaws01-apse1-2.oda.sas.com/import.sas7bdat
Release Created	9.0401M7
Host Created	Linux
Inode Number	201346800
Access Permission	rw-r--r--
Owner Name	u63751674
File Size	96MB
File Size (bytes)	101056512

Alphabetic List of Variables and Attributes					
#	Variable	Type	Len	Format	Informat
3	amount	Num	8	BEST12.	BEST32.
11	isFlaggedFraud	Num	8	BEST12.	BEST32.
10	isFraud	Num	8	BEST12.	BEST32.
7	nameDest	Char	11	\$11.	\$11.
4	nameOrig	Char	11	\$11.	\$11.
9	newbalanceDest	Num	8	BEST12.	BEST32.
6	newbalanceOrig	Num	8	BEST12.	BEST32.
8	oldbalanceDest	Num	8	BEST12.	BEST32.
5	oldbalanceOrg	Num	8	BEST12.	BEST32.
1	step	Num	8	BEST12.	BEST32.
2	type	Char	8	\$8.	\$8.

The CONTENTS Procedure

Data Set Name	WORK.IMPORT	Observations	1048562
Member Type	DATA	Variables	11
Engine	V9	Indexes	0
Created	06/13/2024 21:04:37	Observation Length	96
Last Modified	06/13/2024 21:04:37	Deleted Observations	0
Protection		Compressed	NO
Data Set Type		Sorted	NO
Label			
Data Representation	SOLARIS_X86_64, LINUX_X86_64, ALPHA_TRU64, LINUX_IA64		
Encoding	utf-8 Unicode (UTF-8)		

Engine/Host Dependent Information	
Data Set Page Size	131072
Number of Data Set Pages	770
First Data Page	1
Max Obs per Page	1363
Obs in First Data Page	1329
Number of Data Set Repairs	0
Filename	/saswork/SAS_workC21E00018D41_odaws01-apse1-2.oda.sas.com/SAS_workE6A600018D41_odaws01-apse1-2.oda.sas.com/import.sas7bdat
Release Created	9.0401M7
Host Created	Linux
Inode Number	201346800
Access Permission	rw-r--r--
Owner Name	u63751674
File Size	96MB
File Size (bytes)	101056512

Alphabetic List of Variables and Attributes					
#	Variable	Type	Len	Format	Informat
3	amount	Num	8	BEST12.	BEST32.
11	isFlaggedFraud	Num	8	BEST12.	BEST32.
10	isFraud	Num	8	BEST12.	BEST32.
7	nameDest	Char	11	\$11.	\$11.

Alphabetic List of Variables and Attributes					
#	Variable	Type	Len	Format	Informat
4	nameOrig	Char	11	\$11.	\$11.
9	newbalanceDest	Num	8	BEST12.	BEST32.
6	newbalanceOrig	Num	8	BEST12.	BEST32.
8	oldbalanceDest	Num	8	BEST12.	BEST32.
5	oldbalanceOrg	Num	8	BEST12.	BEST32.
1	step	Num	8	BEST12.	BEST32.
2	type	Char	8	\$8.	\$8.

Obs	step	type	amount	nameOrig	oldbalanceOrg	newbalanceOrig	nameDest	oldbalanceDest	newbalanceDest	isFraud	isFlaggedFraud
1	1	PAYMENT	9839.64	C1231006815	170136	160296.36	M1979787155	0	0	0	0
2	1	PAYMENT	1864.28	C1666544295	21249	19384.72	M2044282225	0	0	0	0
3	1	TRANSFER	181	C1305486145	181	0	C553264065	0	0	1	0
4	1	CASH_OUT	181	C840083671	181	0	C38997010	21182	0	1	0
5	1	PAYMENT	11668.14	C2048537720	41554	29885.86	M1230701703	0	0	0	0

The MEANS Procedure

Analysis Variable : amount	
Mean	Std Dev
160265.45	259276.50

The FASTCLUS Procedure
Replace=FULL Radius=0 Maxclusters=5 Maxiter=100 Converge=0.02

Initial Seeds	
Cluster	amount
1	28.56759076
2	18.79159625
3	-0.61812523
4	37.95073829
5	9.09196709

Minimum Distance Between Initial Seeds = 9.383148

Iteration History						
Iteration	Criterion	Relative Change in Cluster Seeds				
		1	2	3	4	5
1	0.9468	0.1236	0.2482	0.0583	0.0531	0.2960
2	0.6813	0.2293	0.3327	0.00289	0	0.1263
3	0.6376	0.1725	0.2316	0.00187	0	0.0809
4	0.6079	0.2270	0.1522	0.00141	0	0.0597
5	0.5859	0.2037	0.1027	0.00128	0.1324	0.0490
6	0.5679	0.2561	0.0783	0.00137	0.1164	0.0433
7	0.5501	0.2316	0.0672	0.00161	0.0992	0.0400
8	0.5304	0.1717	0.0634	0.00195	0.1085	0.0378
9	0.5076	0.1314	0.0614	0.00240	0.0506	0.0354
10	0.4814	0.1024	0.0562	0.00287	0.2794	0.0315
11	0.4539	0.0762	0.0504	0.00322	0.0955	0.0265
12	0.4283	0.0609	0.0429	0.00319	0.0889	0.0206
13	0.4075	0.0524	0.0360	0.00288	0.0808	0.0154
14	0.3919	0.0440	0.0297	0.00237	0.1549	0.0112
15	0.3811	0.0380	0.0247	0.00189	0.1659	0.00824
16	0.3735	0.0318	0.0212	0.00146	0.0448	0.00607
17	0.3684	0.0246	0.0173	0.00112	0.0429	0.00457
18	0.3650	0.0209	0.0148	0.000847	0.0272	0.00347
19	0.3626	0.0166	0.0119	0.000657	0.0631	0.00267
20	0.3609	0.0139	0.0101	0.000503	0.0853	0.00212
21	0.3596	0.0128	0.00828	0.000403	0.1062	0.00167
22	0.3585	0.0138	0.00758	0.000327	0.1598	0.00138
23	0.3569	0.0143	0.00689	0.000266	0.1625	0.00116
24	0.3549	0.0161	0.00668	0.000222	0.1691	0.00102
25	0.3522	0.0182	0.00683	0.000202	0.1430	0.000967
26	0.3492	0.0201	0.00696	0.000186	0.1135	0.000897
27	0.3462	0.0225	0.00755	0.000170	0.1061	0.000907
28	0.3427	0.0237	0.00837	0.000161	0.0780	0.000906
29	0.3396	0.0220	0.00902	0.000161	0.0596	0.00100
30	0.3368	0.0213	0.00947	0.000178	0.0482	0.00109
31	0.3343	0.0182	0.00942	0.000193	0.0342	0.00118
32	0.3322	0.0159	0.00922	0.000218	0.0274	0.00129
33	0.3303	0.0144	0.00915	0.000235	0.0221	0.00137
34	0.3286	0.0130	0.00889	0.000259	0.0204	0.00148
35	0.3269	0.0122	0.00839	0.000271	0.0189	0.00151

Convergence criterion is satisfied.

Criterion Based on Final Seeds = 0.3253

Cluster Summary						
Cluster	Frequency	RMS Std Deviation	Maximum Distance from Seed to Observation	Radius Exceeded	Nearest Cluster	Distance Between Cluster Centroids
1	11776	1.0783	2.4837		2	3.0576
2	41131	0.6504	1.5486		5	1.5969
3	666755	0.1867	0.4558		5	0.9000
4	2231	3.0406	27.7524		1	4.9189
5	277912	0.3303	0.8298		3	0.9000

Statistics for Variables				
Variable	Total STD	Within STD	R-Square	RSQ/(1-RSQ)
amount	1.00000	0.32448	0.894714	8.497912
OVER-ALL	1.00000	0.32448	0.894714	8.497912

Pseudo F Statistic = 2124053

Approximate Expected Over-All R-Squared = 0.96000

Cubic Clustering Criterion = -717.739

Cluster Means	
Cluster	amount
1	5.11281281
2	2.05523072
3	-0.44168248
4	10.03174281
5	0.45831415

Cluster Standard Deviations	
Cluster	amount
1	1.078339537
2	0.650432293
3	0.186711386
4	3.040595135
5	0.330260520

The CONTENTS Procedure

Data Set Name	WORK.CLUSTERED	Observations	999805
Member Type	DATA	Variables	13
Engine	V9	Indexes	0
Created	06/13/2024 21:04:39	Observation Length	112
Last Modified	06/13/2024 21:04:39	Deleted Observations	0
Protection		Compressed	NO
Data Set Type		Sorted	NO
Label			
Data Representation	SOLARIS_X86_64, LINUX_X86_64, ALPHA_TRU64, LINUX_IA64		
Encoding	utf-8 Unicode (UTF-8)		

Engine/Host Dependent Information	
Data Set Page Size	131072
Number of Data Set Pages	857
First Data Page	1
Max Obs per Page	1168
Obs in First Data Page	1137
Number of Data Set Repairs	0
Filename	/saswork/SAS_workC21E00018D41_odaws01-apse1-2.oda.sas.com/SAS_workE6A600018D41_odaws01-apse1-2.oda.sas.com/clustered.sas7bdat
Release Created	9.0401M7
Host Created	Linux
Inode Number	201344907
Access Permission	rw-r--r--
Owner Name	u63751674
File Size	107MB
File Size (bytes)	112459776

Alphabetic List of Variables and Attributes						
#	Variable	Type	Len	Format	Informat	Label
12	CLUSTER	Num	8			Cluster
13	DISTANCE	Num	8			Distance to Cluster Seed
3	amount	Num	8	BEST12.	BEST32.	
11	isFlaggedFraud	Num	8	BEST12.	BEST32.	
10	isFraud	Num	8	BEST12.	BEST32.	
7	nameDest	Char	11	\$11.	\$11.	
4	nameOrig	Char	11	\$11.	\$11.	

Alphabetic List of Variables and Attributes						
#	Variable	Type	Len	Format	Informat	Label
9	newbalanceDest	Num	8	BEST12.	BEST32.	
6	newbalanceOrig	Num	8	BEST12.	BEST32.	
8	oldbalanceDest	Num	8	BEST12.	BEST32.	
5	oldbalanceOrg	Num	8	BEST12.	BEST32.	
1	step	Num	8	BEST12.	BEST32.	
2	type	Char	8	\$8.	\$8.	

Obs	step	type	amount	nameOrig	oldbalanceOrg	newbalanceOrig	nameDest	oldbalanceDest	newbalanceDest	isFraud	isFlaggedFraud	CLUSTER	DISTANCE	_CLUSTER_
1	1	PAYMENT	-0.580175242	C1231006815	170136	160296.36	M1979787155	0	0	0	0	.	0.14107	.
2	1	PAYMENT	-0.610935299	C1666544295	21249	19384.72	M2044282225	0	0	0	0	.	0.17183	.
3	1	TRANSFER	-0.617427519	C1305486145	181	0	C553264065	0	0	1	0	.	0.17833	.
4	1	CASH_OUT	-0.617427519	C840083671	181	0	C38997010	21182	0	1	0	.	0.17833	.
5	1	PAYMENT	-0.573122925	C2048537720	41554	29885.86	M1230701703	0	0	0	0	.	0.13402	.

The CONTENTS Procedure

Data Set Name	WORK.TRANSACTIONS_CLEAN	Observations	999805
Member Type	DATA	Variables	13
Engine	V9	Indexes	0
Created	06/13/2024 21:04:42	Observation Length	120
Last Modified	06/13/2024 21:04:42	Deleted Observations	0
Protection		Compressed	NO
Data Set Type		Sorted	NO
Label			
Data Representation	SOLARIS_X86_64, LINUX_X86_64, ALPHA_TRU64, LINUX_IA64		
Encoding	utf-8 Unicode (UTF-8)		

Engine/Host Dependent Information

Data Set Page Size	131072
Number of Data Set Pages	918
First Data Page	1
Max Obs per Page	1090
Obs in First Data Page	1061
Number of Data Set Repairs	0
Filename	/saswork/SAS_workC21E00018D41_odaws01-apse1-2.oda.sas.com/SAS_workE6A600018D41_odaws01-apse1-2.oda.sas.com/transactions_clean.sas7bdat
Release Created	9.0401M7
Host Created	Linux
Inode Number	201344914
Access Permission	rw-r--r--
Owner Name	u63751674
File Size	115MB
File Size (bytes)	120455168

Alphabetic List of Variables and Attributes					
#	Variable	Type	Len	Format	Informat
3	amount	Num	8	BEST12.	BEST32.
12	customer_id	Char	11		
11	isFlaggedFraud	Num	8	BEST12.	BEST32.
10	isFraud	Num	8	BEST12.	BEST32.
7	nameDest	Char	11	\$11.	\$11.
4	nameOrig	Char	11	\$11.	\$11.
9	newbalanceDest	Num	8	BEST12.	BEST32.
6	newbalanceOrig	Num	8	BEST12.	BEST32.
8	oldbalanceDest	Num	8	BEST12.	BEST32.
5	oldbalanceOrg	Num	8	BEST12.	BEST32.
1	step	Num	8	BEST12.	BEST32.
13	transaction_time	Num	8		
2	type	Char	8	\$8.	\$8.

The MEANS Procedure

Analysis Variable : transaction_rate	
Mean	Std Dev
0.3909600	0.4401191



