

```

1  /*
2  Merging tables in SAS is a fundamental operation in statistical programming for combining data
3  from multiple sources. Here are some short notes on merging tables in SAS:
4
5  Purpose:
6  Merging tables allows you to combine data based on common variables, enabling comprehensive
7  analysis by consolidating information from different sources into a single dataset.
8
9  Types of merges:
10  SAS offers several types of merges, including the basic merge (one-to-one),
11  one-to-many merge, many-to-one merge, and many-to-many merge. The type of merge depends on the
12  relationship between the common variables in the datasets being merged.
13
14  Merge statement:
15  The MERGE statement is used in SAS to merge datasets. It specifies the datasets to be merged,
16  the common variable(s) to match on, and the type of merge to perform.
17
18  Key variables:
19  Common variables (also called key variables) are essential for merging tables.
20  They act as a link between the datasets, allowing SAS to match observations based on their values.
21
22  Resulting dataset:
23  The merge operation creates a new dataset that contains all the variables from both datasets involved in the merge.
24  The resulting dataset can be further manipulated and analyzed.
25
26  Data integrity:
27  It's crucial to ensure data integrity during the merge process.
28  Checking for duplicates, missing values, and sorting the datasets can help avoid issues and produce accurate results.
29
30  Merging options:
31  SAS provides additional options to customize the merge process,
32  such as IN= and OUT= options to control which datasets are used and where the result is stored, respectively.
33
34  Performance considerations:
35  When merging large datasets, optimizing performance is essential.
36  Properly indexing the datasets, using appropriate merge types, and reducing unnecessary variables can improve efficiency.
37
38  DATA Student;
39  input Id First_Name$ Last_Name$ DOB;
40  informat DOB ddmmyy10.;
41  format DOB ddmmyy10.;
42  datalines;
43  2109 Pradip Fulpagare 09/07/2000
44  2119 Kanchan Mahajan 30/07/2000
45  2131 Nikita Patil 12/07/2000
46  2144 Chetan Salunke 03/05/2000
47  ;
48  proc print data=Student;
49
50  data score;
51  input percentage CGPA;
52  datalines;
53  71 8.6
54  70 8.4
55  65 7.8
56  72 8.7
57  ;
58  /*-----*/
59  data combine;
60  set Student;
61  set score;
62  run;
63  proc print data=combine;
64
65  /*-----*/
66  data domain;
67  input Id First_Name$ Last_Name$ domain$30.;
68  datalines;
69  2109 Pradip Fulpagare Statistician
70  2119 Kanchan Mahajan Statistical_Programmer
71  2131 Nikita Patil Statistician
72  2144 Chetan Salunke Machine_Learning_Engineer
73  ;
74
75
76  DATA Student2;
77  input Id First_Name$ Last_Name$ DOB;
78  informat DOB ddmmyy10.;
79  format DOB ddmmyy10.;
80  datalines;

```

```

81 2111 Sachine Ghogare 09/07/2000
82 2125 Komal Borse 30/07/2000
83
84 ;
85 proc print data=Student2;
86
87 /*-----*/
88 data combine2;
89 set Student domain;
90 proc print data=combine2;
91 /*-----*/
92
93 data combine3;
94 set Student2;
95 set Domain;
96 proc print data=combine3;
97 /*-----*/
98 data score2;
99 input ID percentage CGPA;
100 datalines;
101 2109 71 8.6
102 2119 70 8.4
103 2131 65 7.8
104 2144 72 8.7
105 ;
106
107 /*-----*/
108
109 proc sort data=Student;by Id;
110 proc sort data=Score2;by Id;
111 data combine4;
112 merge Student Score2;
113 by Id;
114 proc print data=combine4;
115
116 data combine5;
117 set Student Score2;
118 proc print data=combine5;
119
120 data combine6;
121 set Student;
122 set Score2;
123 proc print data=combine6;
124
125
126 data combine7;
127 set Student Student2;
128 proc print data=combine7;
129
130
131 data cust1;
132 input ContactName$ Transaction_Amount;
133 datalines;
134 Chetan 500
135 Pradip 400
136 Nikita 300
137 Kanchan 400
138 ;
139
140 data cust2;
141 input ContactName$ City$ Country$;
142 datalines;
143 Sachin Nagar PAK
144 Komal Vadaji UK
145 Chetan Chopda India
146 Pradip Vele USA
147 ;
148
149 proc sort data=cust1;by ContactName ;
150 proc sort data=cust2;by ContactName;
151 data com_;
152 merge cust1 cust2;
153 by ContactName;
154
155 proc print data=com_;
156
157 /*-----*/
158
159 proc sort data=cust1;by ContactName ;
160 proc sort data=cust2;by ContactName;
161 data com_condition;

```

```

162 merge cust1(in=a)
163     cust2(in=b);
164 by ContactName;
165 if a;
166
167 proc print data=com_condition;
168 /*-----*/
169
170 proc sort data=cust1;by ContactName ;
171 proc sort data=cust2;by ContactName;
172 data com_condition2;
173 merge cust1(in=a)
174     cust2(in=b);
175 by ContactName;
176 if a and b;
177
178 proc print data=com_condition2;
179 /*-----*/
180 DATA Student;
181 input Id First_Name$ Last_Name$ DOB;
182 informat DOB ddmmyy10.;
183 format DOB ddmmyy10.;
184 datalines;
185 2109 Pradip Fulpagare 09/07/2000
186 2119 Kanchan Mahajan 30/07/2000
187 2131 Nikita Patil 12/07/2000
188 2144 Chetan Salunke 03/05/2000
189 ;
190 DATA Student2;
191 input Id First_Name$ Last_Name$ DOB;
192 informat DOB ddmmyy10.;
193 format DOB ddmmyy10.;
194 datalines;
195 2111 Sachine Ghogare 09/07/2000
196 2125 Komal Borse 30/07/2000
197
198 ;
199
200 proc append base=Student data=Student2;
201 proc print data=Student; /* Original Data get change*/
202
203
204
205
206
207
208
209

```