

SAS Data set for Q1 to Q6: WORK.creditcard

Q1. Observing the data set 'creditcard' in WORK library. Answer how many numeric and character variables respectively in this table?

Q2. Try to de-duplicate the data set based on the variables 'Cardtype' and 'Owner' using the 'PROC SORT' with 'NODUPKEY' option. You will get the final output table 'creditcard_1'. Please draw some conclusion after comparing the two data sets 'creditcard' and 'creditcard_1'.

Q3. Try to de-duplicate the data set based on the fields 'Cardtype' and 'Owner' using 'PROC SORT' with 'NODUPREC' option. You obtain the output data set 'creditcard_2'. Please summarize your finding after comparing the table 'creditcard' and 'creditcard_2'?

Q4. Using 'PROC PRINT' and the data set 'creditcard_2' created in Q3 to print (output) the following results in SAS output (listing) Window. Where the output only contains the records under the following filter condition Owner='Big Banks'. Please explain why the non-table-variable 'Obs' appears in the output?

The SAS System				21:05 Monday,
----- Cardtype=MasterCard -----				
Obs	Owner	Area	Balance	Limit
1	Big Banks	ALL	1814	8795
2	Big Banks	East	2037	8306
3	Big Banks	Middle	1647	8543
----- Cardtype=Visa -----				
Obs	Owner	Area	Balance	Limit
11	Big Banks	ALL	2334	9391
12	Big Banks	East	2640	9457
13	Big Banks	Middle	2123	8964
14	Big Banks	West	2582	10069

Q5. Using 'PROC PRINT' and the SAS table 'creditcard_2' created in Q3 to print the following results in SAS output (listing) Window. Note,

- The resulting output contains the whole data records.
- For each type of 'Cardtype', you should append a summed value of the variable 'Balance'

----- Cardtype=MasterCard -----

Obs	Area	Balance	Limit
1	ALL	1814	8795
2	East	2037	8306
3	Middle	1647	8543
4	ALL	1877	7285
5	ALL	1840	6926
6	East	1711	6207
7	Middle	1507	6777
8	West	1756	7412
9	ALL	1213	5376
10	Middle	1494	7370
-----		-----	
Cardtype		16896	

----- Cardtype=Visa -----

Obs	Area	Balance	Limit
11	ALL	2334	9391
12	East	2640	9457
13	Middle	2123	8964
14	West	2582	10069
15	Middle	1742	7074
16	ALL	928	4790
17	East	826	4516
18	ALL	1100	8061
19	East	2331	8783
20	Middle	1977	8720
21	West	2276	9468
22	ALL	1662	5648
-----		-----	
Cardtype		22521	
		=====	
		39417	

Q6. Observing the results in listing Window after you have run the following SAS codes:

```
PROC print data=creditcard_2;
VAR Area Balance limit;
SUM Balance;
BY Cardtype Owner;
SUMBY Cardtype;
RUN;
```

Please explain the output by matching the options in above SAS codes. Can you use only 'SUMBY Cardtype;' statement without 'BY Cardtype Owner;' statement in above SAS program? Please explain the reason.

SAS Data set for Q7 to Q8: WORK.Callcenter

Q7.Checking the 'Callcenter' SAS table in WORK library, display the first 5 observations in data set using the procedure 'PROC PRINT'.

Q8. Using 'PROC TRANSPOSE' to create a new set 'callcenter_tr'

	BUSINESS_HOME	NAME OF FORMER VARIABLE	MONDAY	TUESDAY	WEDNESDAY	THURSDAY	FRIDAY
1	BUSINESS	last_hours	53	80	67	100	60
2	HOME	last_hours	108	161	120	178	89

Where the values of column 'WEEK' (i.e. MONDAY, TUESDAY...) in the original table 'callcenter' become (is transposed) the variable names in the table above. As you can find that the values of the column 'last_hours' are transposed separately within "BUSINESS" and "HOME" group.

SAS Data set for Q9: WORK.product

Q9. Using 'PROC TRANSPOSE' to create a data set 'product_tr'

	NAME OF FORMER VARIABLE	Food	Clothes	Toy	Jewelry	Drink
1	tran_count	25	5	18	1	26
2	tran_amount	256	305	98	390	107

Note,

- The characters 'Food', 'Clothes', 'Toy', 'Jewelry' and 'Drink' are all the labels in the transposed table.
- The column names of these columns are actually 'P1', 'P2', 'P3', 'P4' and 'P5'.
- The transposed variables are 'tran_count' and 'tran_amount'.