

Solutions for Integration Assignment 0.5

Solution

Question 1

What is the balance for Acc_number 12346 for 03/01/2018?

Please also provide the code that produced the answer for the question.

Your answer

```
proc import datafile='/home/ermin0/Balance_Bank.xlsx' out=sheet1 dbms=xlsx;
sheet=sheet1;
run;
```

```
proc import datafile='/home/ermin0/Balance_Bank.xlsx' out=sheet2 dbms=xlsx;
sheet=sheet2;
run;
```

```
proc contents data=sheet1; run;
```

```
proc sort data=sheet1 out=sheet1_sorted;
by Acc_number;
run;
```

```
proc sort data=sheet2 out=sheet2_sorted;
by Acc_number;
run;
```

```
data my_balance;
merge sheet1_sorted sheet2_sorted;
by Acc_number;
retain Balance;
if first.Acc_number then Balance=Credit-Debit;
else Balance=Balance+Credit-Debit;
run;
```


Question 2

How would you explain the role of first. in your code?

Your answer

When we use the BY statement, we are telling SAS to group by a particular variable. In our case, we told it to group by Acc_number. We also sorted our Acc_number to be in order.

By using first.Acc_number, this indicator variable will identify if a particular row/observation should be included into a group. For example, most customers will have many transactions, therefore, their Acc_number will be in many rows/observations to reflect all their transactions. We need first. to tell us if a transaction on a particular row should be grouped with a particular Acc_number.

first. will identify if a particular row/observation will be included for a particular customer/Acc_number by assignning = 1 for an observation when an observation is the first observation in each group of values (first time it sees a particular Acc_number) and assign = 0 when a row/obs is not the first observation in each group of values. (the 2nd, 3rd, 4th, etc. times it sees that Acc_number).

So, when it moves on to a new Acc_number it once again assigns = 1, then 0's and so it goes on.

The solution for Integration Assignment 0.5 is attached.

