

MET MA 603:  
Assignment Project Exam Help  
SAS Programming and  
<https://powcoder.com>  
Applications  
Add WeChat powcoder

*Combining Datasets with Set*

# Combining Datasets with Set

The Set Statement can be used to combine two or more SAS datasets. Usually, the datasets will have exactly the same, or a similar, group of variables.

The number of observations in the output dataset is the sum of the observations in the input datasets.

The number of variables in the output dataset is the union of the variables in the input datasets. If a variable is not present in one of the datasets, missing values are written for those observations.

There are two ways that datasets can be combined with the Set statement: Stacking and Interleaving. The only difference between the two methods is the order of observations in the output dataset.

# Stacking

When datasets are combined with **Stacking**, the order of the observations in the output dataset is determined by the order in which the datasets are listed in the Set statement.

```
Data  integers_stacked ;  
Set   integers_odd integers_even;  
run ;
```

Assignment Project Exam Help

<https://powcoder.com>

In the example above, two datasets are being combined. In the output dataset, ~~integers\_stacked~~, the order of observations will be: first, all of the observations from integers\_odd, in the order found in integers\_odd, next, all of the observations from integers\_even, in the order found in integers\_even.

Add WeChat powcoder

# Interleaving

When datasets are combined with **Interleaving**, the order of the observations in the output dataset is determined by one or more of the variables in the input datasets.

The BY statement is used to indicate which variable or variables to interleave by.

The input datasets must already be sorted by the variables that is being used to interleave the datasets.

```
Data    integers_stacked ;  
Set      integers_odd integers_even;  
By       number ;  
run ;
```

# Practice

Use the Losses\_Weather.sas7bdat and Losses\_Nonweather.sas7bdat datasets.

First, add a variable called WeatherFlag to each dataset, such that 1 corresponds to a weather loss and 0 corresponds to a non-weather loss.

Then, combine the datasets such that the output dataset is ordered by the weather flag.

Finally, combine the datasets such that the output dataset is order by date of loss.

# Readings

- Textbook sections 6.2, 6.3

Assignment Project Exam Help

<https://powcoder.com>

Add WeChat powcoder