



Calling R from SAS

sp4r01d01.sas

Note: Calling R from SAS requires an installed version of R, and a configuration setting within the SAS environment. These settings will be covered in the video "Readying Your Machine to Call R" in Course 4: The Interactive Matrix Language and Open Source Integration. Because you can't modify the SAS University Edition configuration this PROC IML step will not execute successfully in SAS University Edition. For this demonstration, you can learn by just watching the "Calling R from SAS" video.

Use the randomForest package in R. Send the **birth** data set to R. Use the randomForest() function to create a predictive model, and return the results to SAS.

1. Invoke SAS/IML and send the **birth** data set in the **Work** library to R. Name the data frame **birth** as well.

```
proc iml;  
    call ExportDataSetToR("work.birth","birth");
```

2. Write your R code between the SUBMIT and ENDSUBMIT statements. Use the randomForest package in R and the randomForest() function to estimate a model with **BWT** as the dependent variable and **Smoke**, **HT**, **LWT**, and **PTL** as independent variables. Use the SUMMARY statement to print the details of the analysis to the console. Finally, create a data frame with the actual and predicted values, given the model, and name the variables **Actual** and **Predicted**.

```
submit / r;  
    library(randomForest)  
    rf = randomForest(BWT ~ SMOKE + HT + LWT + PTL,  
                      data=birth,ntree=200,importance=TRUE)  
    summary(rf)  
    actual = birth$BWT  
    pred = predict(rf,data=birth)  
    actual.pred = cbind(actual,pred)  
    colnames(actual.pred) <- c("Actual","Predicted")  
endsubmit;
```

	Length	Class	Mode
call	5	-none-	call
type	1	-none-	character
predicted	189	-none-	numeric
mse	200	-none-	numeric
rsq	200	-none-	numeric
oob.times	189	-none-	numeric
importance	8	-none-	numeric
importanceSD	4	-none-	numeric
localImportance	0	-none-	NULL
proximity	0	-none-	NULL
ntree	1	-none-	numeric
mtry	1	-none-	numeric
forest	11	-none-	list
coefs	0	-none-	NULL
y	189	-none-	numeric
test	0	-none-	NULL
inbag	0	-none-	NULL
terms	3	terms	call



The output generated in the R console was printed in the SAS Results page.

- Return the data frame to a SAS data set with the name **Rdata**.

```
call ImportDataSetFromR("Rdata", "actual.pred");
quit;
```



If you are running SAS Studio in client-server mode, you do *not* have access to the **Work** library on a point-and-click basis. You must use the PRINT procedure to view the results.