



```
/*This document is a code snippet, showcasing how to build a multiple linear  
regression model native on hana */
```

```
-----  
/* MULTIPLE LINEAR REGRESSION: PREDICT*/  
-----
```

```
-----  
/*DATA TABLE USED TO PREDICT: DEFFINITION*/  
-----
```

```
/*
```

The data definition phase, has the sole purpose of laying the foundation for the procedure, when generating the procedure, we must feed the parameter tables.

The parameter tables have no data in them so far and their structure must be defined.

All the tables will be in another table ("RM_PAL_FMLR_PDATA_TBL"), this will be specified in the procedure generation wrapper

```
*/
```



```
SET SCHEMA PAL;
```

```
DROP TYPE RM_PAL_FMLR_PREDICT_T;
```

```
CREATE TYPE RM_PAL_FMLR_PREDICT_T AS TABLE(
```

```
    "ID" INTEGER NOT NULL PRIMARY KEY GENERATED BY DEFAULT AS IDENTITY,
```

```
    "STORE" INTEGER,
```

```
    "STORE_TYPE" VARCHAR(10),
```

```
    "ASSORTMENT" varchar(10),
```

```
    "PROMO" INTEGER,
```

```
    "PROMO2" INTEGER,
```

```
    "PROMO2_SINCE_WEEK" INTEGER,
```

```
    "COMP_OP_Y" INTEGER,
```

```
    "SCHOOL_H" VARCHAR(10));
```

```
-----  
/*COEFFICIENT TABLE: DEFFINITION*/
```

```
-----  
/* this table will contain the coefficient values of the independent variables from  
the model's "RESULT" table */
```

```
DROP TYPE RM_PAL_FMLR_COEFFICIENT_T;
```

```
CREATE TYPE RM_PAL_FMLR_COEFFICIENT_T AS TABLE("Coefficient" varchar(50),  
"CoefficientValue" DOUBLE);
```



/*CONTROL TABLE: DEFFINITION*/

DROP TYPE RM_PAL_CONTROL_T;
CREATE TYPE RM_PAL_CONTROL_T AS TABLE("NAME" VARCHAR(100), "INTARGS" INT,
"DOUBLEARGS" DOUBLE,"STRINGARGS" VARCHAR(100));

/*FITTED (PREDICTED) VALUES TABLE: DEFFINITION*/

DROP TYPE RM_PAL_FMLR_FITTED_T;
CREATE TYPE RM_PAL_FMLR_FITTED_T AS TABLE("ID" INT,"Fitted" DOUBLE);

/* PARAMATER TABLE CREATION */

DROP table RM_PAL_FMLR_PDATA_TBL;
CREATE column table RM_PAL_FMLR_PDATA_TBL("POSITION" INT, "SCHEMA_NAME"
NVARCHAR(256), "TYPE_NAME" NVARCHAR(256), "PARAMETER_TYPE" VARCHAR(7));
INSERT INTO RM_PAL_FMLR_PDATA_TBL values (1,'PAL','RM_PAL_FMLR_PREDICT_T','IN');
INSERT INTO RM_PAL_FMLR_PDATA_TBL values
(2,'PAL','RM_PAL_FMLR_COEFFICIENT_T','IN');
INSERT INTO RM_PAL_FMLR_PDATA_TBL values (3,'PAL','RM_PAL_CONTROL_T','IN');
INSERT INTO RM_PAL_FMLR_PDATA_TBL values
(4,'PAL','RM_PAL_FMLR_FITTED_T','OUT');



```
/* PROCEDURE GENERATION */
```

```
-----
```

```
/*
```

Similarly to the model's procedure, this one will also contain the parameter table and all the table types defined.

From the "ALFPAL" area we will now use the "FORECASTWITHLR" function,

The procedure will be stored in the "PAL" Schema'

"RM_PAL_FORECAST_LR_PROC" is the name of our created procedure

"RM_PAL_FMLR_PDATA_TBL" is the parameter table

```
*/
```

```
CALL SYS.AFLLANG_WRAPPER_PROCEDURE_DROP('PAL','RM_PAL_FORECAST_LR_PROC');
```

```
CALL
```

```
SYS.AFLLANG_WRAPPER_PROCEDURE_CREATE('ALFPAL','FORECASTWITHLR','PAL','RM_PAL_FORECAST_LR_PROC',RM_PAL_FMLR_PDATA_TBL);
```

```
/*Now that the preparation part is complete it's time to import the coefficient's table ("Result table") from the model and feed the data table the independent variables*/
```

```
-----
```

```
/* CONTROL TABLE: INPUT TABLE */
```

```
-----
```

```
set schema pal;
```

```
DROP TABLE #RM_PAL_CONTROL_TBL;
```

```
CREATE LOCAL TEMPORARY COLUMN TABLE #RM_PAL_CONTROL_TBL ("NAME" VARCHAR(100), "INTARGS" INT, "DOUBLEARGS" DOUBLE,"STRINGARGS" VARCHAR(100));
```

```
INSERT INTO #RM_PAL_CONTROL_TBL VALUES ('THREAD_NUMBER',2,null,null);
```



/* DATA TABLE CREATION:INPUT TABLE */

```
DROP TABLE RM_PAL_FMLR_PREDICTDATA_TBL;
CREATE COLUMN TABLE RM_PAL_FMLR_PREDICTDATA_TBL LIKE
RM_PAL_FMLR_PREDICT_T;
INSERT INTO RM_PAL_FMLR_PREDICTDATA_TBL
("STORE","STORE_TYPE","ASSORTMENT","PROMO","PROMO2","PROMO2_SINCE_WEEK","C
OMP_OP_Y",
    "SCHOOL_H")
SELECT
    CAST("GPE_RMSTR" AS INTEGER),
    "GPE_RMSTP",
    "GPE_RMART",
    CAST("GPE_RMPRM" AS INTEGER),
    CAST("GPE_RMPR2" AS INTEGER),
    CAST("GPE_RMP2M" AS INTEGER),
    CAST("GPE_RMCOY" AS INTEGER),
    "GPE_RMIHD"
FROM "_SYS_BIC"."ZGPE.ROSSMAN/Z_PA_ROSSMAN_TRAIN_SET_CV"
WHERE CAST("GPE_RMSTR" AS INTEGER)<=10 AND CAST("GPE_RMOPN" AS
INTEGER)=1;
```



```
-----  
/* COEFFICIENT TABLE CREATION:INPUT TABLE */  
-----
```

```
/*we will be using the generated coefficient table created by the model*/
```

```
SET SCHEMA PAL;
```

```
DROP TABLE RM_PAL_FMLR_COEFFICIENT_TBL;
```

```
CREATE COLUMN TABLE RM_PAL_FMLR_COEFFICIENT_TBL ("Coefficient" varchar(50),  
"CoefficientValue" DOUBLE);
```

```
INSERT INTO RM_PAL_FMLR_COEFFICIENT_TBL("Coefficient","CoefficientValue")
```

```
    SELECT * FROM RM_PAL_MLR_RESULTS_TBL
```

```
    ;
```

```
-----  
/* FITTED VALUES TABLE CREATION:OUTPUT TABLE */  
-----
```

```
DROP TABLE RM_PAL_FMLR_FITTED_TBL;
```

```
CREATE COLUMN TABLE RM_PAL_FMLR_FITTED_TBL ("ID" INT,"Fitted" DOUBLE);
```

```
-----  
/* CALL PREVIOUSLY GENERATED PROCEDURE AND SPECIFY THE INPUT TABLES AS  
PARAMETERS */  
-----
```

```
/*Now that all the input parameter tables contain the needed data we just need to  
call the previously generated procedure so that it results in a prediction for the  
target column(dependent variable)*/
```

```
CALL PAL.RM_PAL_FORECAST_LR_PROC(RM_PAL_FMLR_PREDICTDATA_TBL,  
RM_PAL_FMLR_COEFFICIENT_TBL, "#RM_PAL_CONTROL_TBL",  
RM_PAL_FMLR_FITTED_TBL) with overview;
```



/* VIEW RESULTS */

```
SELECT * FROM RM_PAL_FMLR_FITTED_TBL;  
SELECT A."ID",A."SALES", B."Fitted" FROM RM_PAL_MLR_DATA_TBL A JOIN  
RM_PAL_FMLR_FITTED_TBL B ON A."ID"=B."ID"
```