

## Euredit utility function: xml\_read

### 1 Purpose

**xml\_read** reads data from an Euredit XML file.

### 2 Specification

```
#include <euredit_sys.h>
```

```
void xml_read (char *xmlfname, char *xdrfname, double miss_val, long n,
               long m, double *data, char **v_name, long *n_cat, long *cat_val,
               long maxcat, long mset, long *nmiss, long info[2])
```

### 3 Parameters

#### **xmlfname**

*Input:* the file name of the Euredit XML data file.

#### **xdrfname**

*Input:* the file name of the Euredit XDR associated with the XML file.

#### **miss\_val**

*Input:* the number that is used to represent a missing value.

#### **n**

*Input:* the number of observations in the data.

*Constraint:*  $n \geq 1$ .

#### **m**

*Input:* the number of variables in the data.

*Constraint:*  $m \geq 1$ .

#### **data[n\*m]**

*Output:* the data stored by row.

#### **v\_name[m][max\_name]**

*Output:* the names of the  $m$  variables. Each row in **v\_name** must be at least **max\_name** in length, where the value of **max\_name** is found by a call to **xml\_data\_size** prior to the call to **xml\_read**.

#### **n\_cat[m]**

*Output:* the number of categories present in a variable, if **n\_cat**[ $i$ ] = 0 then the variable is continuous

#### **cat\_val[m\*maxcat]**

*Output:* the categories observed for the categorical variables. The categories for the  $i$  variable are stored in **catval**[ $i*\text{maxcat}+j$ ] for  $j = 1, 2, \dots, n\_cat[i]$ .

#### **maxcat**

*Input:* the maximum number of categories in a variable.

#### **mset**

*Input:* indicated what level of data is to be set to missing.

If **mset** = 1 all non-valid values, including imputed values.

If **mset** = 2 all suspect, in error and missing values.

If **mset** = 3 all in error and missing values.

If **mset**  $\geq 14$  only missing data values.

*Constraint:* **mset**  $\geq 1$ .

#### **nmiss**

*Output:* the number of missing values in **data**.

**info**

*Output:* information on the success of the function call.

**info**[0] = 0: the function successfully completed its task.

**info**[0] =  $i$ : the specification of the  $i$ th formal parameter was incorrect,  $i = 1, 2, \dots, 13$ .

**info**[0] = 99: the function failed to allocate enough memory.

**info**[1] contains additional information for system debugging.