

# Utils

---

[Go Up](#)

## Table of Contents

<a href="#">Fat.ecl</a>
Will take a potentially sparse file d and fill in the missing
<a href="#">FatD.ecl</a>
Will take a potentially sparse file d and fill in the missing
<a href="#">Gini.ecl</a>
Creates a file of pivot/target pairs with a Gini impurity value
<a href="#">SequenceInField.ecl</a>
Given a file which is sorted by the work item identifier and INFIELD (and possibly other values), add sequence numbers within the range of each infield

[Go Up](#)

## **IMPORTS**

ML\_Core.Types |

## **DESCRIPTIONS**

### **FUNCTION** **Fat**

<code>DATASET(Types.NumericField)</code>	<b>Fat</b>
<code>(DATASET(Types.NumericField) d0, Types.t_FieldReal v=0)</code>	

Will take a potentially sparse file d and fill in the missing with value v for Numeric Field datasets

**PARAMETER** **d0** They myriad format Numeric Field dataset to be filled

**PARAMETER** **v** The value to assign missing records

**RETURN** A full Numeric Field dataset with every field populated

---

# ML\_Core/ Utils/ FatD

---

[Go Up](#)

## IMPORTS

ML\_Core.Types |

## DESCRIPTIONS

### **FUNCTION** FatD

<code>DATASET(Types.DiscreteField)</code>	<b>FatD</b>
<code>(DATASET(Types.DiscreteField) d0, Types.t_Discrete v=0)</code>	

Will take a potentially sparse file d and fill in the missing with value v for Discrete Field datasets

**PARAMETER** d0 They myriad format Discrete Field dataset to be filled

**PARAMETER** v The value to assign missing records

**RETURN** A full Discrete Field dataset with every field populated

---

# ML\_Core/ Utils/ Gini

---

[Go Up](#)

## DESCRIPTIONS

### MACRO Gini

	Gini
<code>(infile, pivot, target, wi_name='wi')</code>	

Creates a file of pivot/target pairs with a Gini impurity value.

**PARAMETER** infile the input file, any type with a work item field

**PARAMETER** pivot the name of the pivot field

**PARAMETER** target the name of the field used as the target

**PARAMETER** wi\_name the name of the work item field, default is "wi" return A table by Work Item and Pivot value giving count and Gini impurity value

---

# ML\_Core/ Utils/ SequenceInField

---

[Go Up](#)

## DESCRIPTIONS

### MACRO SequenceInField

SequenceInField
<code>(infile,infield,seq,wi_name='wi')</code>

Given a file which is sorted by the work item identifier and INFIELD (and possibly other values), add sequence numbers within the range of each infield. Slightly elaborate code is to avoid having to partition the data to one value of infield per node and to work with very large numbers of records where a global count project would be inappropriate. This is useful for assigning rank positions with the groupings.

**PARAMETER** infile the input file, any type

**PARAMETER** infield field name of grouping field

**PARAMETER** seq name of the field to receive the sequence number

**PARAMETER** wi\_name work item field name, default is wi

**RETURN** a file of the same type with sequence numbers applied

---