

我添加了新的 class, CosanRawData. 添加了读取 csv 后 stdout 有用的一些信息。如下图所示。改写了 load\_csv, 成为 CosanRawData 的 private function。现在可以兼顾 missing, +inf,-inf, categorical data types。原来在 io/utlis.h 的 load\_csv 改写成了 load\_csv1. (后面我可能不再用它)

example\_data/toy2 里面创立了一些有 categorical, missing data, inf 的 data。

test1.cpp is the playground for rewriting load\_csv

test2.cpp snippet to print out the new member variables.

test.cpp as a test program on CosanRawData.

documentation 在 ./docs/html/index.html

```
Siri-Pro:cosan xinyuzhang$ gcc -I/Users/xinyuzhang/Desktop/Spring2021/c++pattern/proje
ct/cosan -O2 test.cpp
Siri-Pro:cosan xinyuzhang$ ./a.out
Number of rows: 100
Number of columns: 11
Number of positive infinity values: 1. They are at
(1,2)
Number of negative infinity values: 1. They are at
(2,2)
Number of missing values: 6. They are at (0,2) (50,0) (51,0) (52,0) (67,1) (97,2)
Columns of categorical values: Column 10
```

Cosan::CosanRawData Class Reference

#include <CosanData.h>

Inheritance diagram for Cosan::CosanRawData:

```
graph BT
    CosanData[Cosan::CosanData] --> CosanRawData[Cosan::CosanRawData]
    CosanRawData --> CosanBO[Cosan::CosanBO]
```

Public Member Functions

	CosanRawData ()=delete
	CosanRawData (const std::string &srcX)
	CosanRawData (const std::string &srcX, const std::string &srcY)
void	SetInput (const std::string &srcX)
void	SetTarget (const std::string &srcY)
CosanMatrix &	GetInput ()
const CosanMatrix &	GetTarget () const
virtual const std::string	GetName () const
virtual const std::string	GetSummaryMessageX () const
virtual const std::string	GetSummaryMessageY () const

Public Member Functions Inherited from Cosan::CosanBO

Protected Attributes

CosanMatrix	X
CosanMatrix	Y
std::string	SummaryMessageX
std::string	SummaryMessageY
std::vector< std::vector< uint >>	IdxpInFX
std::vector< std::vector< uint >>	IdxminFX
std::vector< std::vector< uint >>	IdxmissingX
std::vector< std::vector< uint >>	IdxpInFY
std::vector< std::vector< uint >>	IdxminFY
std::vector< std::vector< uint >>	IdxmissingY
std::set< uint >	colCatX
std::set< uint >	colCatY
bool	catY = false
uint	rowsX = 0
uint	colsX = 0
uint	rowsY = 0
uint	colsY = 0
std::vector< std::string >	svaluesX
std::vector< std::string >	svaluesY

Private Member Functions

template<typename Matrix >	
std::tuple< uint, uint, std::string >	_load_csv (const std::string &path, CosanMatrix &X, std::vector< std::vector< uint >> &IdxpInF, std::vector< std::vector< uint >> &IdxminF, std::vector< std::vector< uint >> &Idxmissing, std::vector< std::string > &svalues, std::set< uint > &colCat)