

# Overview

Brought to you by [YData](#)

Overview

Alerts 4

Reproduction

## Dataset statistics

Number of variables	8
Number of observations	541909
Missing cells	136534
Missing cells (%)	3.1%
Duplicate rows	4879
Duplicate rows (%)	0.9%
Total size in memory	192.8 MiB
Average record size in memory	373.0 B

## Variable types

Text	3
Numeric	3
DateTime	1
Categorical	1

# Variables

Select Columns ▼

## InvoiceNo

Text

<b>Distinct</b>	25900
<b>Distinct (%)</b>	4.8%
<b>Missing</b>	0
<b>Missing (%)</b>	0.0%
<b>Memory size</b>	32.6 MiB



## StockCode

Text

<b>Distinct</b>	4070
<b>Distinct (%)</b>	0.8%
<b>Missing</b>	0
<b>Missing (%)</b>	0.0%
<b>Memory size</b>	32.1 MiB



## Description

Text

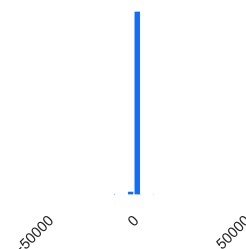
<b>Distinct</b>	4223
<b>Distinct (%)</b>	0.8%
<b>Missing</b>	1454
<b>Missing (%)</b>	0.3%
<b>Memory size</b>	43.2 MiB



## Quantity

Real number ( $\mathbb{R}$ )

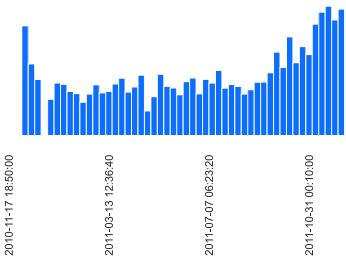
<b>Distinct</b>	722	<b>Minimum</b>	-80995
<b>Distinct (%)</b>	0.1%	<b>Maximum</b>	80995
<b>Missing</b>	0	<b>Zeros</b>	0
<b>Missing (%)</b>	0.0%	<b>Zeros (%)</b>	0.0%
<b>Infinite</b>	0	<b>Negative</b>	10624
<b>Infinite (%)</b>	0.0%	<b>Negative (%)</b>	2.0%
<b>Mean</b>	9.5522495	<b>Memory size</b>	4.1 MiB



# InvoiceDate

Date

Distinct	23260	Minimum	2010-12-01 08:00:00
Distinct (%)	4.3%	Maximum	2011-12-09 12:00:00
Missing	0	Invalid dates	0
Missing (%)	0.0%	Invalid dates (%)	0.0%
Memory size	4.1 MiB		

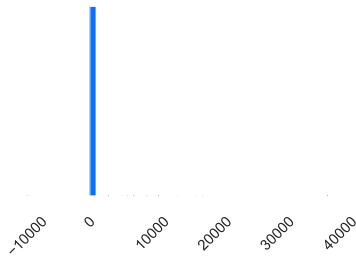


# UnitPrice

Real number ( $\mathbb{R}$ )

Skewed

Distinct	1630	Minimum	-11062.06
Distinct (%)	0.3%	Maximum	38970
Missing	0	Zeros	2515
Missing (%)	0.0%	Zeros (%)	0.5%
Infinite	0	Negative	2
Infinite (%)	0.0%	Negative (%)	< 0.1%
Mean	4.6111136	Memory size	4.1 MiB

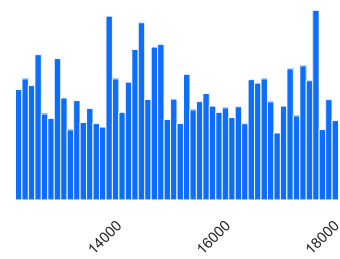


## CustomerID

Real number ( $\mathbb{R}$ )

Missing

Distinct	4372	Minimum	12346
Distinct (%)	1.1%	Maximum	18287
Missing	135080	Zeros	0
Missing (%)	24.9%	Zeros (%)	0.0%
Infinite	0	Negative	0
Infinite (%)	0.0%	Negative (%)	0.0%
Mean	15287.691	Memory size	4.1 MiB



## Country

Categorical

Imbalance

Distinct	38	United Ki...	495478
Distinct (%)	< 0.1%	Germany	9495
Missing	0	France	8557
Missing (%)	0.0%	EIRE	8196
Memory size	36.4 MiB	Spain	2533
		Other val...	17650

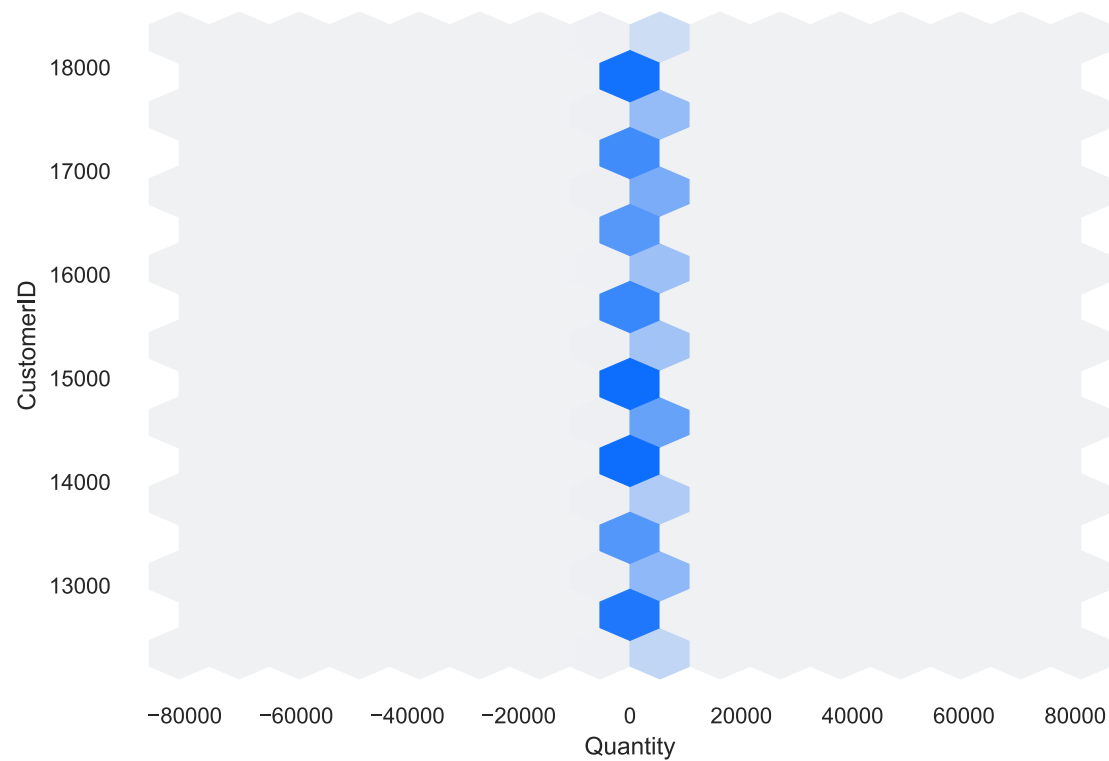
# Interactions

Quantity	UnitPrice	CustomerID
----------	-----------	------------

CustomerID

Quantity

UnitPrice

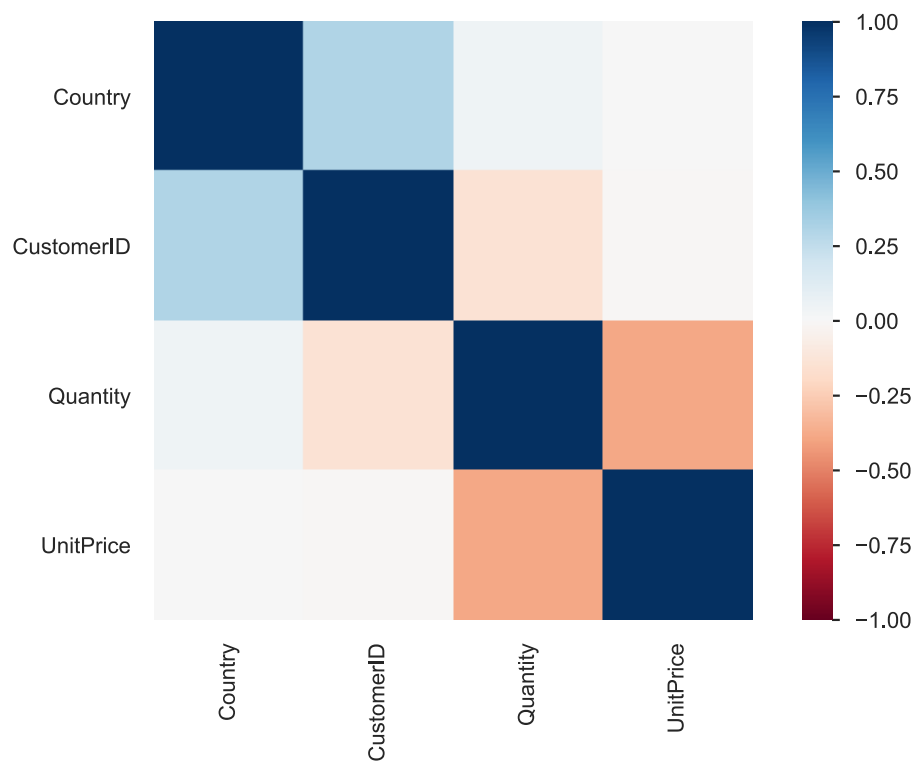


# Correlations

Auto

Heatmap

Table

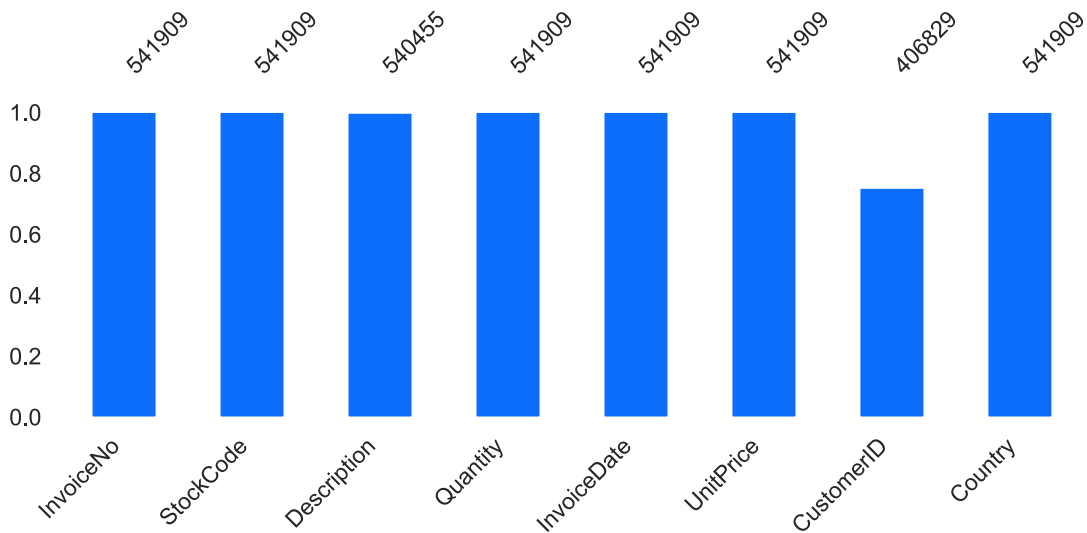


## Missing values

Count

Matrix

Heatmap



A simple visualization of nullity by column.

## Sample

First rows

[Last rows](#)

	InvoiceNo	StockCode	Description	Quantity
0	536365	85123A	WHITE HANGING HEART T-LIGHT HOLDER	6
1	536365	71053	WHITE METAL LANTERN	6
2	536365	84406B	CREAM CUPID HEARTS COAT HANGER	8
3	536365	84029G	KNITTED UNION FLAG HOT WATER BOTTLE	6
4	536365	84029E	RED WOOLLY HOTTIE WHITE HEART.	6
5	536365	22752	SET 7 BABUSHKA NESTING BOXES	2
6	536365	21730	GLASS STAR FROSTED T-LIGHT HOLDER	6



	InvoiceNo	StockCode	Description	Quantity
7	536366	22633	HAND WARMER UNION JACK	6
8	536366	22632	HAND WARMER RED POLKA DOT	6
9	536367	84879	ASSORTED COLOUR BIRD ORNAMENT	32

## Duplicate rows

### Most frequently occurring

	InvoiceNo	StockCode	Description	Quantity	InvoiceDate	UnitPrice
1613	555524	22698	PINK REGENCY TEACUP AND SAUCER	1	6/5/11 11:37	2.95
1612	555524	22697	GREEN REGENCY TEACUP AND SAUCER	1	6/5/11 11:37	2.95
3202	572861	22775	PURPLE DRAWERKNOB ACRYLIC EDWARDIAN	12	10/26/11 12:46	1.25
347	538514	21756	BATH BUILDING BLOCK WORD	1	12/12/10 14:27	5.95
474	540524	21756	BATH BUILDING BLOCK WORD	1	1/9/11 12:53	5.95
528	541266	21754	HOME BUILDING BLOCK WORD	1	1/16/11 16:25	5.95

	InvoiceNo	StockCode	Description	Quantity	InvoiceDate	UnitPrice
<b>529</b>	541266	21755	LOVE BUILDING BLOCK WORD	1	1/16/11 16:25	5.95
<b>3146</b>	572344	M	Manual	48	10/24/11 10:43	1.50
<b>4249</b>	578289	23395	BELLE JARDINIERE CUSHION COVER	1	11/23/11 14:07	3.75
<b>186</b>	537224	70007	HI TEC ALPINE HAND WARMER	1	12/5/10 16:24	1.65

Report generated by [YData](#).