

# Overview

## Dataset statistics

<b>Number of variables</b>	25
<b>Number of observations</b>	12684
<b>Missing cells</b>	794
<b>Missing cells (%)</b>	0.3%
<b>Duplicate rows</b>	74
<b>Duplicate rows (%)</b>	0.6%
<b>Total size in memory</b>	2.4 MiB
<b>Average record size in memory</b>	200.0 B

## Variable types

<b>Categorical</b>	25
--------------------	----

## Alerts

toCoupon_GEQ5min has constant value "1"	Constant
Dataset has 74 (0.6%) duplicate rows	Duplicates
destination is highly overall correlated with passanger and 3 other fields (passanger, time, direction_same, direction_opp)	High correlation
passanger is highly overall correlated with destination	High correlation
weather is highly overall correlated with temperature	High correlation
temperature is highly overall correlated with weather	High correlation
time is highly overall correlated with destination	High correlation
maritalStatus is highly overall correlated with has_children	High correlation
has_children is highly overall correlated with maritalStatus	High correlation
direction_same is highly overall correlated with destination and 1 other fields (destination, direction_opp)	High correlation
direction_opp is highly overall correlated with destination and 1 other fields (destination, direction_same)	High correlation
CoffeeHouse has 217 (1.7%) missing values	Missing
CarryAway has 151 (1.2%) missing values	Missing
RestaurantLessThan20 has 130 (1.0%) missing values	Missing
Restaurant20To50 has 189 (1.5%) missing values	Missing

## Reproduction

# Variables

Select Columns ▾

## destination

Categorical

<b>Distinct</b>	3
<b>Distinct (%)</b>	< 0.1%
<b>Missing</b>	0
<b>Missing (%)</b>	0.0%
<b>Memory size</b>	99.2 KiB

## Length

<b>Max length</b>	15
<b>Median length</b>	4
<b>Mean length</b>	9.4488332
<b>Min length</b>	4

## Characters and Unicode

<b>Total characters</b>	119849
<b>Distinct characters</b>	17
<b>Distinct categories</b>	3 ( <a href="https://en.wikipedia.org/wiki/Unicode_character_property#General_Category">https://en.wikipedia.org/wiki/Unicode_character_property#General_Category</a> )
<b>Distinct scripts</b>	2 ( <a href="https://en.wikipedia.org/wiki/Script_(Unicode)#List_of_scripts_in_Unicode">https://en.wikipedia.org/wiki/Script_(Unicode)#List_of_scripts_in_Unicode</a> )
<b>Distinct blocks</b>	1 ( <a href="https://en.wikipedia.org/wiki/Unicode_block">https://en.wikipedia.org/wiki/Unicode_block</a> )

The Unicode Standard assigns character properties to each code point, which can be used to analyse textual variables.

## Unique

<b>Unique</b>	0	?
<b>Unique (%)</b>	0.0%	?

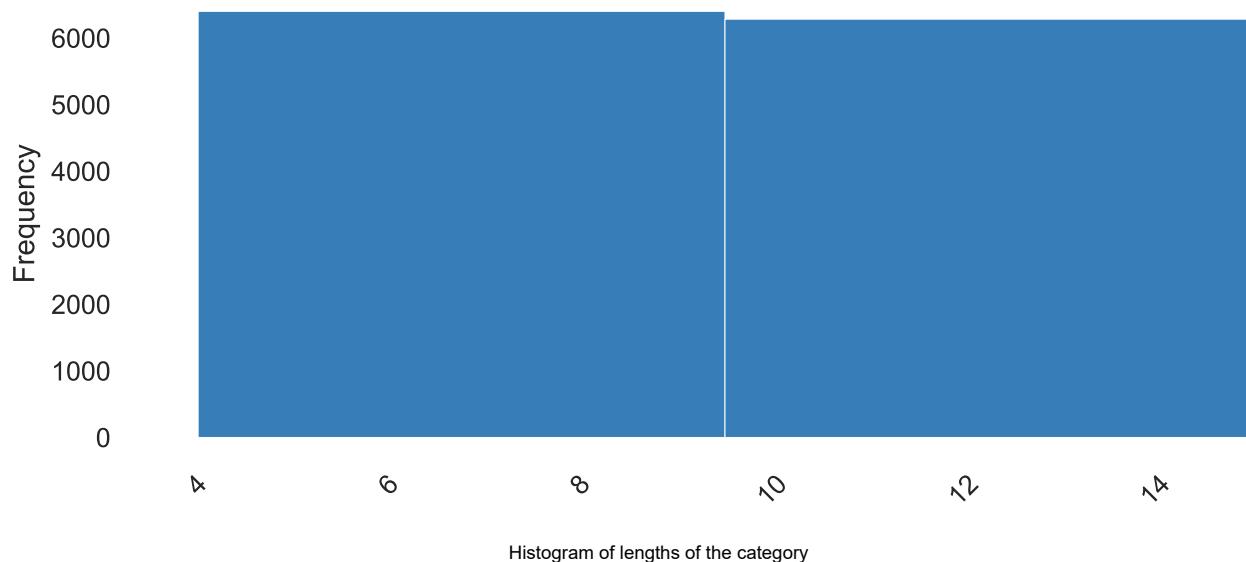
## Sample

<b>1st row</b>	No Urgent Place
<b>2nd row</b>	No Urgent Place
<b>3rd row</b>	No Urgent Place
<b>4th row</b>	No Urgent Place
<b>5th row</b>	No Urgent Place

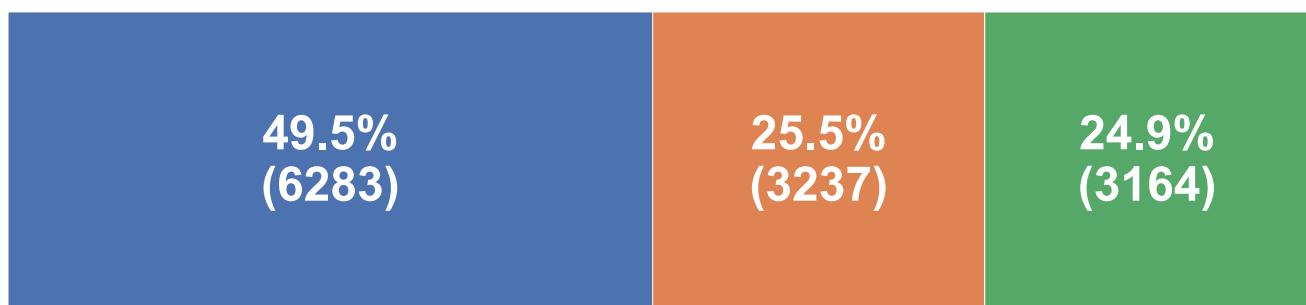
### Common Values

Value	Count	Frequency (%)
No Urgent Place	6283	49.5%
Home	3237	25.5%
Work	3164	24.9%

### Length



### Common Values (Plot)



No Urgent Place



Home



Work

Value	Count	Frequency (%)
no	6283	24.9%
urgent	6283	24.9%
place	6283	24.9%
home	3237	12.8%
work	3164	12.5%

## Most occurring characters

Value	Count	Frequency (%)
e	15803	13.2%
o	12684	10.6%
	12566	10.5%
r	9447	7.9%
N	6283	5.2%
P	6283	5.2%
c	6283	5.2%
a	6283	5.2%
l	6283	5.2%
t	6283	5.2%
Other values (7)	31651	26.4%

### Most occurring categories

Value	Count	Frequency (%)
Lowercase Letter	82033	68.4%
Uppercase Letter	25250	21.1%
Space Separator	12566	10.5%

### Most frequent character per category

#### *Lowercase Letter*

Value	Count	Frequency (%)
e	15803	19.3%
o	12684	15.5%
r	9447	11.5%
c	6283	7.7%
a	6283	7.7%
l	6283	7.7%
t	6283	7.7%
n	6283	7.7%
g	6283	7.7%
m	3237	3.9%

#### *Uppercase Letter*

Value	Count	Frequency (%)
N	6283	24.9%
P	6283	24.9%
U	6283	24.9%
H	3237	12.8%
W	3164	12.5%

#### *Space Separator*

Value	Count	Frequency (%)
	12566	100.0%

## Most occurring scripts

Value	Count	Frequency (%)
Latin	107283	89.5%
Common	12566	10.5%

## Most frequent character per script

*Latin*

Value	Count	Frequency (%)
e	15803	14.7%
o	12684	11.8%
r	9447	8.8%
N	6283	5.9%
P	6283	5.9%
c	6283	5.9%
a	6283	5.9%
l	6283	5.9%
t	6283	5.9%
n	6283	5.9%
Other values (6)	25368	23.6%

*Common*

Value	Count	Frequency (%)
	12566	100.0%

### Most occurring blocks

Value	Count	Frequency (%)
ASCII	119849	100.0%

### Most frequent character per block

ASCII

Value	Count	Frequency (%)
e	15803	13.2%
o	12684	10.6%
	12566	10.5%
r	9447	7.9%
N	6283	5.2%
P	6283	5.2%
c	6283	5.2%
a	6283	5.2%
l	6283	5.2%
t	6283	5.2%
Other values (7)	31651	26.4%





**passanger**  
Categorical

<b>Distinct</b>	4
<b>Distinct (%)</b>	< 0.1%
<b>Missing</b>	0
<b>Missing (%)</b>	0.0%
<b>Memory size</b>	99.2 KiB

**Length**

<b>Max length</b>	9
<b>Median length</b>	5
<b>Mean length</b>	6.2888679
<b>Min length</b>	5

**Characters and Unicode**

<b>Total characters</b>	79768
<b>Distinct characters</b>	16
<b>Distinct categories</b>	4 ( <a href="https://en.wikipedia.org/wiki/Unicode_character_property#General_Category">https://en.wikipedia.org/wiki/Unicode_character_property#General_Category</a> )
<b>Distinct scripts</b>	2 ( <a href="https://en.wikipedia.org/wiki/Script_(Unicode)#List_of_scripts_in_Unicode">https://en.wikipedia.org/wiki/Script_(Unicode)#List_of_scripts_in_Unicode</a> )
<b>Distinct blocks</b>	1 ( <a href="https://en.wikipedia.org/wiki/Unicode_block">https://en.wikipedia.org/wiki/Unicode_block</a> )

The Unicode Standard assigns character properties to each code point, which can be used to analyse textual variables.

**Unique**

<b>Unique</b>	0	?
<b>Unique (%)</b>	0.0%	?

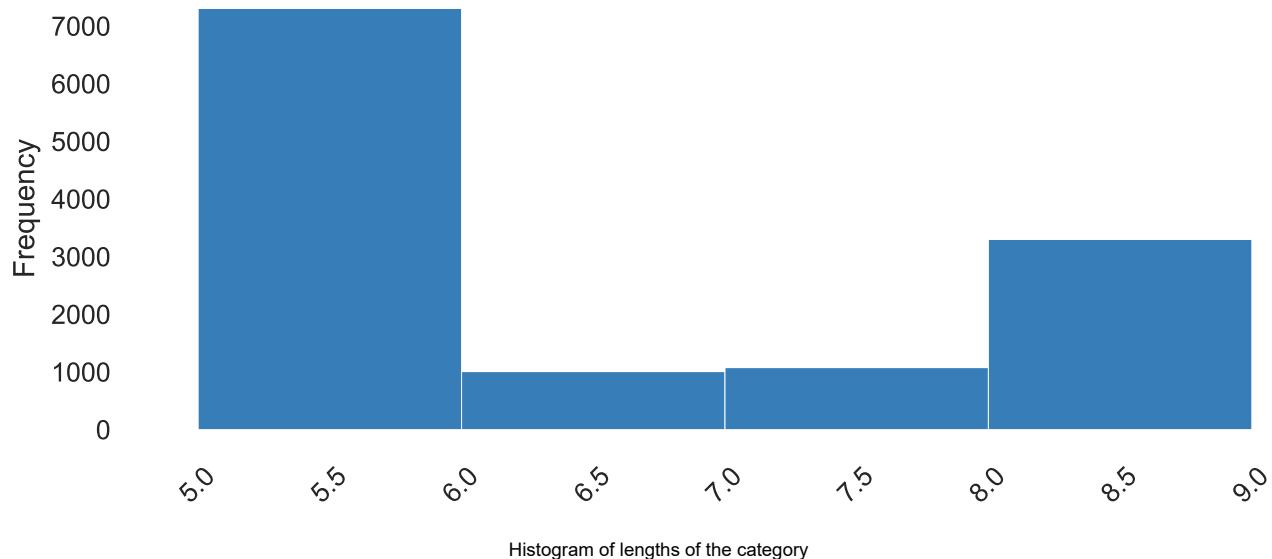
**Sample**

<b>1st row</b>	Alone
<b>2nd row</b>	Friend(s)
<b>3rd row</b>	Friend(s)
<b>4th row</b>	Friend(s)
<b>5th row</b>	Friend(s)

### Common Values

Value	Count	Frequency (%)
Alone	7305	57.6%
Friend(s)	3298	26.0%
Partner	1075	8.5%
Kid(s)	1006	7.9%

### Length



### Common Values (Plot)



Value	Count	Frequency (%)
alone	7305	57.6%
friend(s	3298	26.0%
partner	1075	8.5%
kid(s	1006	7.9%

## Most occurring characters

Value	Count	Frequency (%)
n	11678	14.6%
e	11678	14.6%
A	7305	9.2%
l	7305	9.2%
o	7305	9.2%
r	5448	6.8%
i	4304	5.4%
d	4304	5.4%
(	4304	5.4%
s	4304	5.4%
Other values (6)	11833	14.8%

### Most occurring categories

Value	Count	Frequency (%)
Lowercase Letter	58476	73.3%
Uppercase Letter	12684	15.9%
Open Punctuation	4304	5.4%
Close Punctuation	4304	5.4%

### Most frequent character per category

#### *Lowercase Letter*

Value	Count	Frequency (%)
n	11678	20.0%
e	11678	20.0%
l	7305	12.5%
o	7305	12.5%
r	5448	9.3%
i	4304	7.4%
d	4304	7.4%
s	4304	7.4%
a	1075	1.8%
t	1075	1.8%

#### *Uppercase Letter*

Value	Count	Frequency (%)
A	7305	57.6%
F	3298	26.0%
P	1075	8.5%
K	1006	7.9%

#### *Open Punctuation*

Value	Count	Frequency (%)
(	4304	100.0%

#### *Close Punctuation*

Value	Count	Frequency (%)
)	4304	100.0%

## Most occurring scripts

Value	Count	Frequency (%)
Latin	71160	89.2%
Common	8608	10.8%

## Most frequent character per script

*Latin*

Value	Count	Frequency (%)
n	11678	16.4%
e	11678	16.4%
A	7305	10.3%
I	7305	10.3%
o	7305	10.3%
r	5448	7.7%
i	4304	6.0%
d	4304	6.0%
s	4304	6.0%
F	3298	4.6%
Other values (4)	4231	5.9%

*Common*

Value	Count	Frequency (%)
(	4304	50.0%
)	4304	50.0%

### Most occurring blocks

Value	Count	Frequency (%)
ASCII	79768	100.0%

### Most frequent character per block

ASCII

Value	Count	Frequency (%)
n	11678	14.6%
e	11678	14.6%
A	7305	9.2%
I	7305	9.2%
o	7305	9.2%
r	5448	6.8%
i	4304	5.4%
d	4304	5.4%
(	4304	5.4%
s	4304	5.4%
Other values (6)	11833	14.8%





**weather**

Categorical

<b>Distinct</b>	3
<b>Distinct (%)</b>	< 0.1%
<b>Missing</b>	0
<b>Missing (%)</b>	0.0%
<b>Memory size</b>	99.2 KiB

**Length**

<b>Max length</b>	5
<b>Median length</b>	5
<b>Mean length</b>	5
<b>Min length</b>	5

**Characters and Unicode**

<b>Total characters</b>	63420
<b>Distinct characters</b>	9
<b>Distinct categories</b>	<a href="https://en.wikipedia.org/wiki/Unicode_character_property#General_Category">2 (https://en.wikipedia.org/wiki/Unicode_character_property#General_Category)</a>
<b>Distinct scripts</b>	<a href="https://en.wikipedia.org/wiki/Script_(Unicode)#List_of_scripts_in_Unicode">1 (https://en.wikipedia.org/wiki/Script_(Unicode)#List_of_scripts_in_Unicode)</a>
<b>Distinct blocks</b>	<a href="https://en.wikipedia.org/wiki/Unicode_block">1 (https://en.wikipedia.org/wiki/Unicode_block)</a>

The Unicode Standard assigns character properties to each code point, which can be used to analyse textual variables.

**Unique**

<b>Unique</b>	0	?
<b>Unique (%)</b>	0.0%	

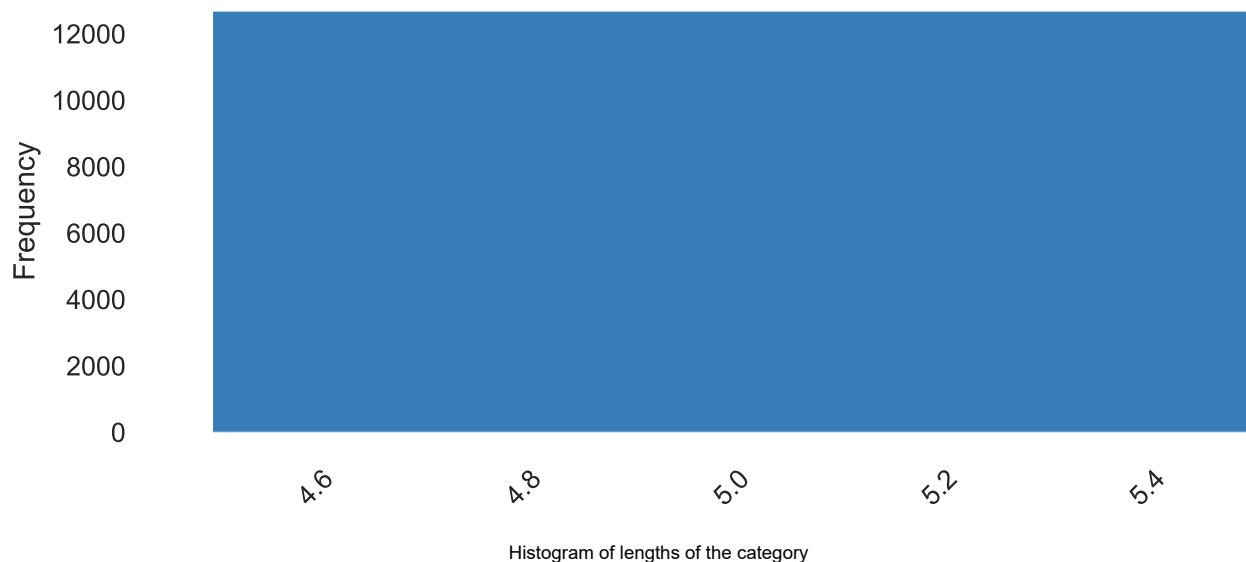
**Sample**

<b>1st row</b>	Sunny
<b>2nd row</b>	Sunny
<b>3rd row</b>	Sunny
<b>4th row</b>	Sunny
<b>5th row</b>	Sunny

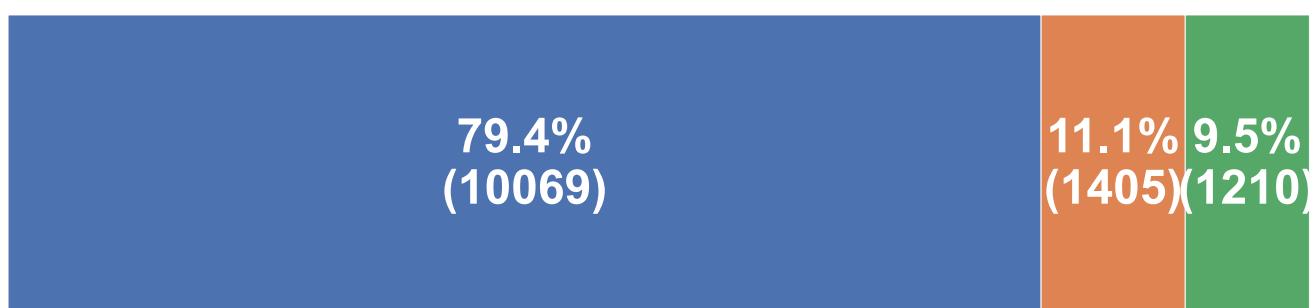
### Common Values

Value	Count	Frequency (%)
Sunny	10069	79.4%
Snowy	1405	11.1%
Rainy	1210	9.5%

### Length



### Common Values (Plot)



Value	Count	Frequency (%)
sunny	10069	79.4%
snowy	1405	11.1%
rainy	1210	9.5%

## Most occurring characters

Value	Count	Frequency (%)
n	22753	35.9%
y	12684	20.0%
S	11474	18.1%
u	10069	15.9%
o	1405	2.2%
w	1405	2.2%
R	1210	1.9%
a	1210	1.9%
i	1210	1.9%

### Most occurring categories

Value	Count	Frequency (%)
Lowercase Letter	50736	80.0%
Uppercase Letter	12684	20.0%

### Most frequent character per category

#### *Lowercase Letter*

Value	Count	Frequency (%)
n	22753	44.8%
y	12684	25.0%
u	10069	19.8%
o	1405	2.8%
w	1405	2.8%
a	1210	2.4%
i	1210	2.4%

#### *Uppercase Letter*

Value	Count	Frequency (%)
S	11474	90.5%
R	1210	9.5%

### Most occurring scripts

Value	Count	Frequency (%)
Latin	63420	100.0%

### Most frequent character per script

*Latin*

Value	Count	Frequency (%)
n	22753	35.9%
y	12684	20.0%
S	11474	18.1%
u	10069	15.9%
o	1405	2.2%
w	1405	2.2%
R	1210	1.9%
a	1210	1.9%
i	1210	1.9%

### Most occurring blocks

Value	Count	Frequency (%)
ASCII	63420	100.0%

### Most frequent character per block

ASCII

Value	Count	Frequency (%)
n	22753	35.9%
y	12684	20.0%
S	11474	18.1%
u	10069	15.9%
o	1405	2.2%
w	1405	2.2%
R	1210	1.9%
a	1210	1.9%
i	1210	1.9%





**temperature**

Categorical

<b>Distinct</b>	3
<b>Distinct (%)</b>	< 0.1%
<b>Missing</b>	0
<b>Missing (%)</b>	0.0%
<b>Memory size</b>	99.2 KiB

**Length**

<b>Max length</b>	2
<b>Median length</b>	2
<b>Mean length</b>	2
<b>Min length</b>	2

**Characters and Unicode**

<b>Total characters</b>	25368
<b>Distinct characters</b>	4
<b>Distinct categories</b>	<a href="https://en.wikipedia.org/wiki/Unicode_character_property#General_Category">1 (https://en.wikipedia.org/wiki/Unicode_character_property#General_Category)</a>
<b>Distinct scripts</b>	<a href="https://en.wikipedia.org/wiki/Script_(Unicode)#List_of_scripts_in_Unicode">1 (https://en.wikipedia.org/wiki/Script_(Unicode)#List_of_scripts_in_Unicode)</a>
<b>Distinct blocks</b>	<a href="https://en.wikipedia.org/wiki/Unicode_block">1 (https://en.wikipedia.org/wiki/Unicode_block)</a>

The Unicode Standard assigns character properties to each code point, which can be used to analyse textual variables.

**Unique**

<b>Unique</b>	0	?
<b>Unique (%)</b>	0.0%	

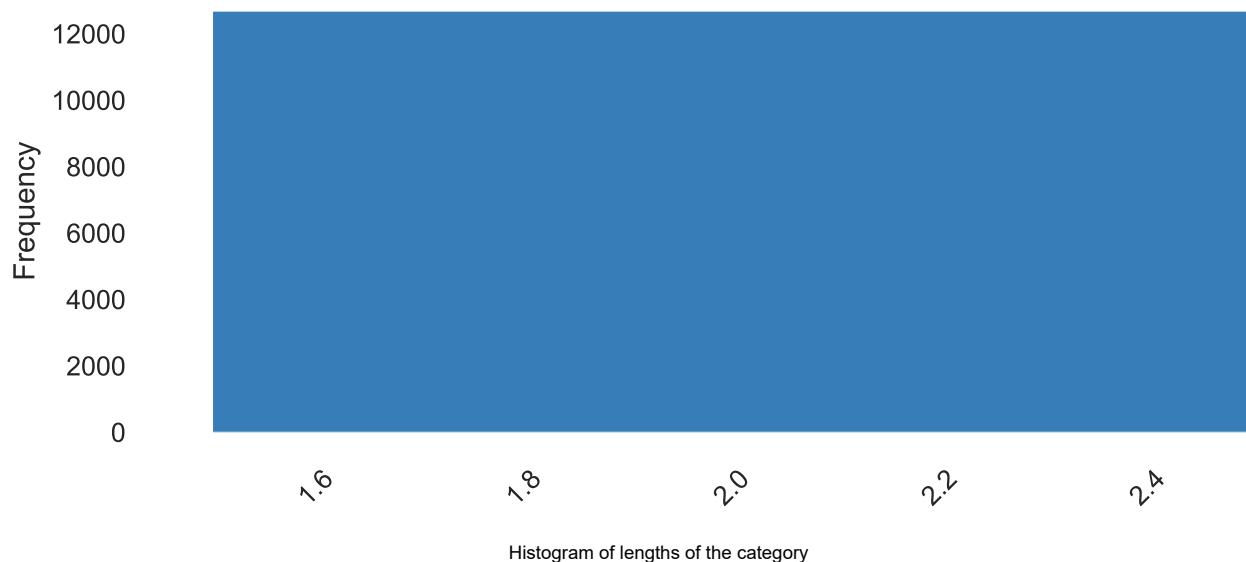
**Sample**

<b>1st row</b>	55
<b>2nd row</b>	80
<b>3rd row</b>	80
<b>4th row</b>	80
<b>5th row</b>	80

### Common Values

Value	Count	Frequency (%)
80	6528	51.5%
55	3840	30.3%
30	2316	18.3%

### Length



### Common Values (Plot)



Value	Count	Frequency (%)
80	6528	51.5%
55	3840	30.3%
30	2316	18.3%

## Most occurring characters

Value	Count	Frequency (%)
0	8844	34.9%
5	7680	30.3%
8	6528	25.7%
3	2316	9.1%

### Most occurring categories

Value	Count	Frequency (%)
Decimal Number	25368	100.0%

### Most frequent character per category

*Decimal Number*

Value	Count	Frequency (%)
0	8844	34.9%
5	7680	30.3%
8	6528	25.7%
3	2316	9.1%

### Most occurring scripts

Value	Count	Frequency (%)
Common	25368	100.0%

### Most frequent character per script

*Common*

Value	Count	Frequency (%)
0	8844	34.9%
5	7680	30.3%
8	6528	25.7%
3	2316	9.1%

### Most occurring blocks

Value	Count	Frequency (%)
ASCII	25368	100.0%

### Most frequent character per block

ASCII

Value	Count	Frequency (%)
0	8844	34.9%
5	7680	30.3%
8	6528	25.7%
3	2316	9.1%





**time**

Categorical

<b>Distinct</b>	5
<b>Distinct (%)</b>	< 0.1%
<b>Missing</b>	0
<b>Missing (%)</b>	0.0%
<b>Memory size</b>	99.2 KiB

**Length**

<b>Max length</b>	4
<b>Median length</b>	3
<b>Mean length</b>	3.3375118
<b>Min length</b>	3

**Characters and Unicode**

<b>Total characters</b>	42333
<b>Distinct characters</b>	8
<b>Distinct categories</b>	<a href="https://en.wikipedia.org/wiki/Unicode_character_property#General_Category">2 (https://en.wikipedia.org/wiki/Unicode_character_property#General_Category)</a>
<b>Distinct scripts</b>	<a href="https://en.wikipedia.org/wiki/Script_(Unicode)#List_of_scripts_in_Unicode">2 (https://en.wikipedia.org/wiki/Script_(Unicode)#List_of_scripts_in_Unicode)</a>
<b>Distinct blocks</b>	<a href="https://en.wikipedia.org/wiki/Unicode_block">1 (https://en.wikipedia.org/wiki/Unicode_block)</a>

The Unicode Standard assigns character properties to each code point, which can be used to analyse textual variables.

**Unique**

<b>Unique</b>	0	?
<b>Unique (%)</b>	0.0%	?

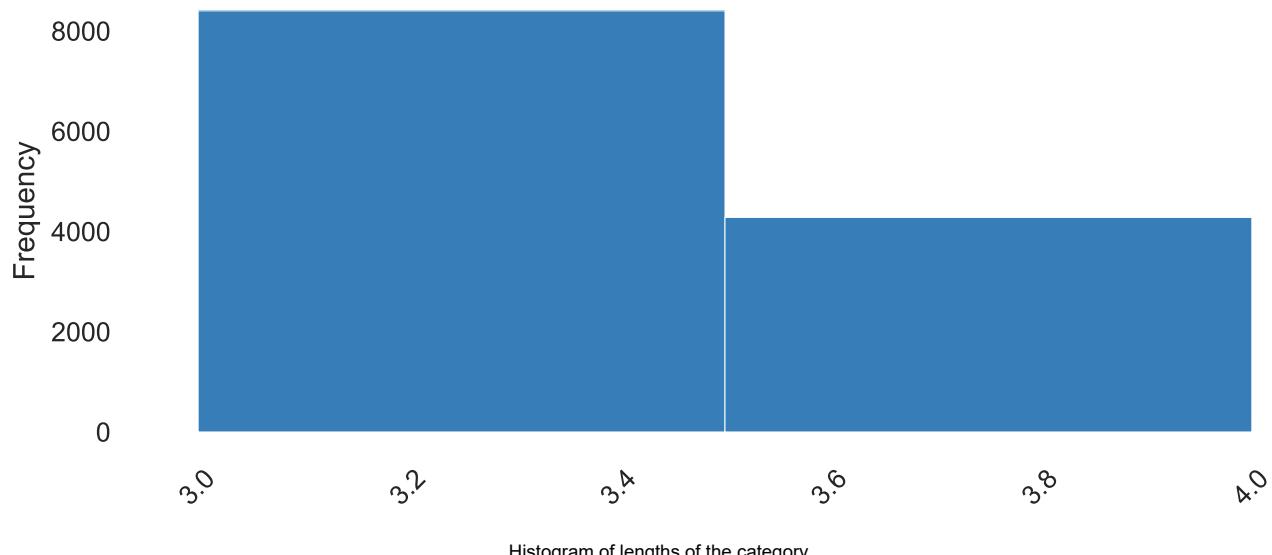
**Sample**

<b>1st row</b>	2PM
<b>2nd row</b>	10AM
<b>3rd row</b>	10AM
<b>4th row</b>	2PM
<b>5th row</b>	2PM

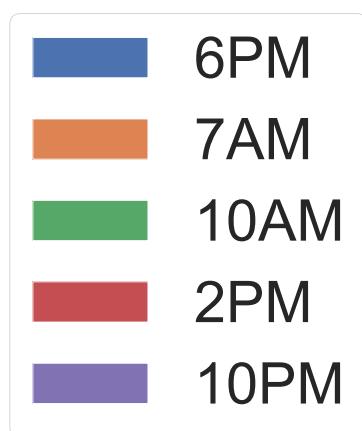
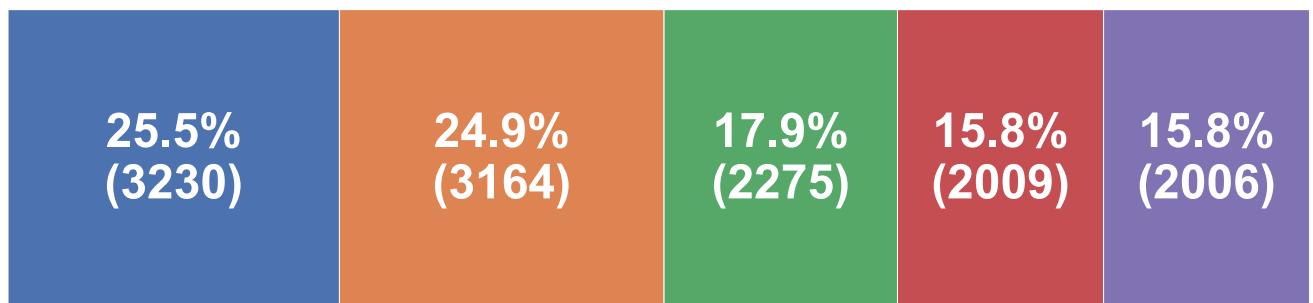
### Common Values

Value	Count	Frequency (%)
6PM	3230	25.5%
7AM	3164	24.9%
10AM	2275	17.9%
2PM	2009	15.8%
10PM	2006	15.8%

### Length



### Common Values (Plot)



Value	Count	Frequency (%)
6pm	3230	25.5%
7am	3164	24.9%
10am	2275	17.9%
2pm	2009	15.8%
10pm	2006	15.8%

## Most occurring characters

Value	Count	Frequency (%)
M	12684	30.0%
P	7245	17.1%
A	5439	12.8%
1	4281	10.1%
0	4281	10.1%
6	3230	7.6%
7	3164	7.5%
2	2009	4.7%

### Most occurring categories

Value	Count	Frequency (%)
Uppercase Letter	25368	59.9%
Decimal Number	16965	40.1%

### Most frequent character per category

*Decimal Number*

Value	Count	Frequency (%)
1	4281	25.2%
0	4281	25.2%
6	3230	19.0%
7	3164	18.7%
2	2009	11.8%

*Uppercase Letter*

Value	Count	Frequency (%)
M	12684	50.0%
P	7245	28.6%
A	5439	21.4%

### Most occurring scripts

Value	Count	Frequency (%)
Latin	25368	59.9%
Common	16965	40.1%

### Most frequent character per script

*Common*

Value	Count	Frequency (%)
1	4281	25.2%
0	4281	25.2%
6	3230	19.0%
7	3164	18.7%
2	2009	11.8%

*Latin*

Value	Count	Frequency (%)
M	12684	50.0%
P	7245	28.6%
A	5439	21.4%

### Most occurring blocks

Value	Count	Frequency (%)
ASCII	42333	100.0%

### Most frequent character per block

ASCII

Value	Count	Frequency (%)
M	12684	30.0%
P	7245	17.1%
A	5439	12.8%
1	4281	10.1%
0	4281	10.1%
6	3230	7.6%
7	3164	7.5%
2	2009	4.7%





**coupon**  
Categorical

<b>Distinct</b>	5
<b>Distinct (%)</b>	< 0.1%
<b>Missing</b>	0
<b>Missing (%)</b>	0.0%
<b>Memory size</b>	99.2 KiB

## Length

<b>Max length</b>	21
<b>Median length</b>	17
<b>Mean length</b>	13.513876
<b>Min length</b>	3

## Characters and Unicode

<b>Total characters</b>	171410
<b>Distinct characters</b>	26
<b>Distinct categories</b>	9 ( <a href="https://en.wikipedia.org/wiki/Unicode_character_property#General_Category">https://en.wikipedia.org/wiki/Unicode_character_property#General_Category</a> )
<b>Distinct scripts</b>	2 ( <a href="https://en.wikipedia.org/wiki/Script_(Unicode)#List_of_scripts_in_Unicode">https://en.wikipedia.org/wiki/Script_(Unicode)#List_of_scripts_in_Unicode</a> )
<b>Distinct blocks</b>	1 ( <a href="https://en.wikipedia.org/wiki/Unicode_block">https://en.wikipedia.org/wiki/Unicode_block</a> )

The Unicode Standard assigns character properties to each code point, which can be used to analyse textual variables.

## Unique

<b>Unique</b>	0	?
<b>Unique (%)</b>	0.0%	

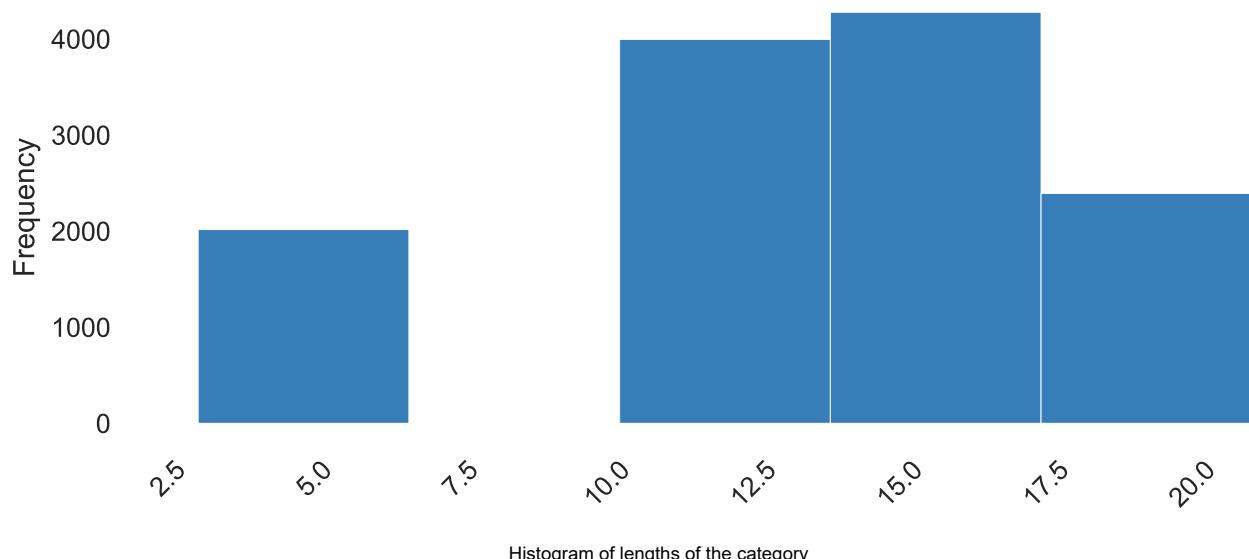
## Sample

<b>1st row</b>	Restaurant(<20)
<b>2nd row</b>	Coffee House
<b>3rd row</b>	Carry out & Take away
<b>4th row</b>	Coffee House
<b>5th row</b>	Coffee House

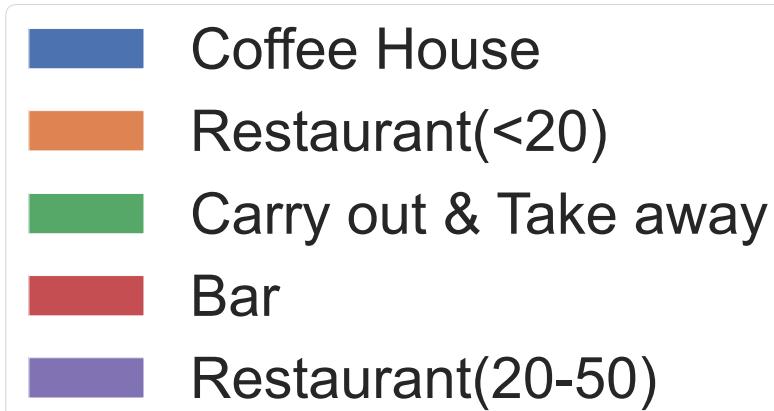
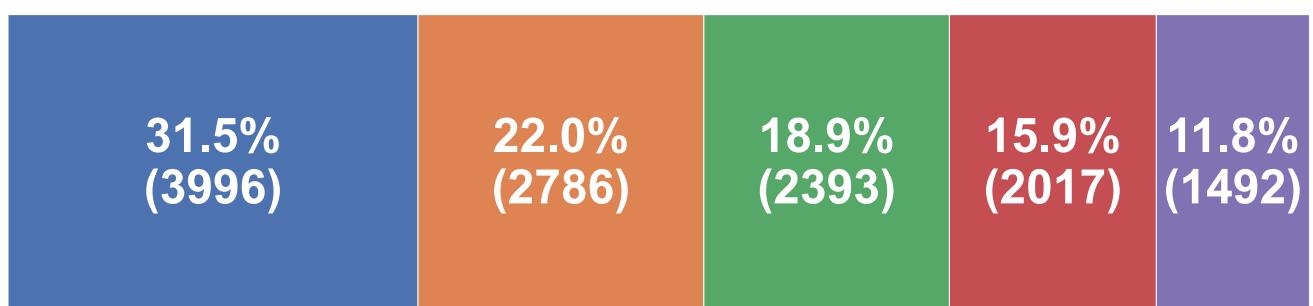
## Common Values

Value	Count	Frequency (%)
Coffee House	3996	31.5%
Restaurant(<20)	2786	22.0%
Carry out & Take away	2393	18.9%
Bar	2017	15.9%
Restaurant(20-50)	1492	11.8%

## Length



## Common Values (Plot)



Value	Count	Frequency (%)
coffee	3996	15.2%
house	3996	15.2%
restaurant(<20	2786	10.6%
carry	2393	9.1%
out	2393	9.1%
	2393	9.1%
take	2393	9.1%
away	2393	9.1%
bar	2017	7.7%
restaurant(20-50	1492	5.7%

## Most occurring characters

Value	Count	Frequency (%)
a	20145	11.8%
e	18659	10.9%
	13568	7.9%
r	11081	6.5%
t	10949	6.4%
u	10667	6.2%
o	10385	6.1%
s	8274	4.8%
f	7992	4.7%
C	6389	3.7%
Other values (16)	53301	31.1%

### Most occurring categories

Value	Count	Frequency (%)
Lowercase Letter	112002	65.3%
Uppercase Letter	19073	11.1%
Space Separator	13568	7.9%
Decimal Number	11540	6.7%
Close Punctuation	4278	2.5%
Open Punctuation	4278	2.5%
Math Symbol	2786	1.6%
Other Punctuation	2393	1.4%
Dash Punctuation	1492	0.9%

### Most frequent character per category

#### *Lowercase Letter*

Value	Count	Frequency (%)
a	20145	18.0%
e	18659	16.7%
r	11081	9.9%
t	10949	9.8%
u	10667	9.5%
o	10385	9.3%
s	8274	7.4%
f	7992	7.1%
y	4786	4.3%
n	4278	3.8%
Other values (2)	4786	4.3%

#### *Uppercase Letter*

Value	Count	Frequency (%)
C	6389	33.5%
R	4278	22.4%
H	3996	21.0%
T	2393	12.5%
B	2017	10.6%

#### *Decimal Number*

Value	Count	Frequency (%)
0	5770	50.0%
2	4278	37.1%
5	1492	12.9%

#### *Space Separator*

Value	Count	Frequency (%)
	13568	100.0%

*Close Punctuation*

Value	Count	Frequency (%)
)	4278	100.0%

*Open Punctuation*

Value	Count	Frequency (%)
(	4278	100.0%

*Math Symbol*

Value	Count	Frequency (%)
<	2786	100.0%

*Other Punctuation*

Value	Count	Frequency (%)
&	2393	100.0%

*Dash Punctuation*

Value	Count	Frequency (%)
-	1492	100.0%

## Most occurring scripts

Value	Count	Frequency (%)
Latin	131075	76.5%
Common	40335	23.5%

## Most frequent character per script

*Latin*

Value	Count	Frequency (%)
a	20145	15.4%
e	18659	14.2%
r	11081	8.5%
t	10949	8.4%
u	10667	8.1%
o	10385	7.9%
s	8274	6.3%
f	7992	6.1%
C	6389	4.9%
y	4786	3.7%
Other values (7)	21748	16.6%

*Common*

Value	Count	Frequency (%)
0	13568	33.6%
2	5770	14.3%
)	4278	10.6%
(	4278	10.6%
<	2786	6.9%
&	2393	5.9%
-	1492	3.7%
5	1492	3.7%

### Most occurring blocks

Value	Count	Frequency (%)
ASCII	171410	100.0%

### Most frequent character per block

ASCII

Value	Count	Frequency (%)
a	20145	11.8%
e	18659	10.9%
r	13568	7.9%
t	11081	6.5%
u	10949	6.4%
o	10667	6.2%
s	10385	6.1%
f	8274	4.8%
C	7992	4.7%
Other values (16)	6389	3.7%
	53301	31.1%





**expiration**

Categorical

<b>Distinct</b>	2
<b>Distinct (%)</b>	< 0.1%
<b>Missing</b>	0
<b>Missing (%)</b>	0.0%
<b>Memory size</b>	99.2 KiB

**Length**

<b>Max length</b>	2
<b>Median length</b>	2
<b>Mean length</b>	2
<b>Min length</b>	2

**Characters and Unicode**

<b>Total characters</b>	25368
<b>Distinct characters</b>	4
<b>Distinct categories</b>	2 ( <a href="https://en.wikipedia.org/wiki/Unicode_character_property#General_Category">https://en.wikipedia.org/wiki/Unicode_character_property#General_Category</a> )
<b>Distinct scripts</b>	2 ( <a href="https://en.wikipedia.org/wiki/Script_(Unicode)#List_of_scripts_in_Unicode">https://en.wikipedia.org/wiki/Script_(Unicode)#List_of_scripts_in_Unicode</a> )
<b>Distinct blocks</b>	1 ( <a href="https://en.wikipedia.org/wiki/Unicode_block">https://en.wikipedia.org/wiki/Unicode_block</a> )

The Unicode Standard assigns character properties to each code point, which can be used to analyse textual variables.

**Unique**

<b>Unique</b>	0	?
<b>Unique (%)</b>	0.0%	

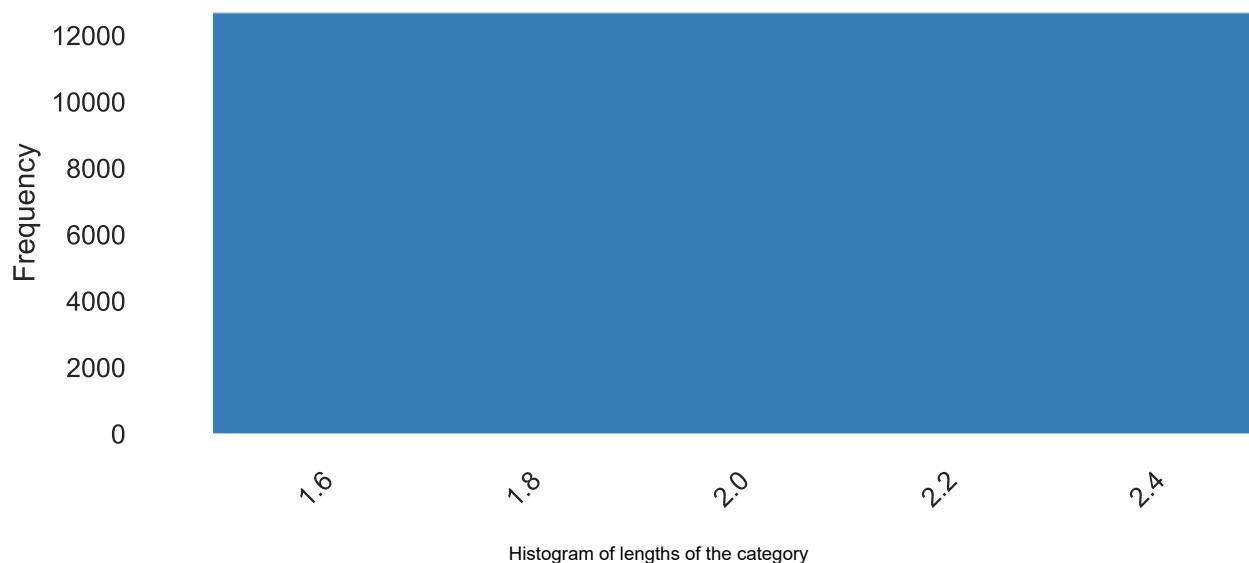
**Sample**

<b>1st row</b>	1d
<b>2nd row</b>	2h
<b>3rd row</b>	2h
<b>4th row</b>	2h
<b>5th row</b>	1d

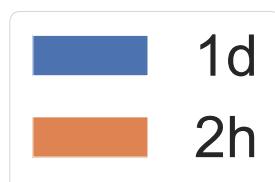
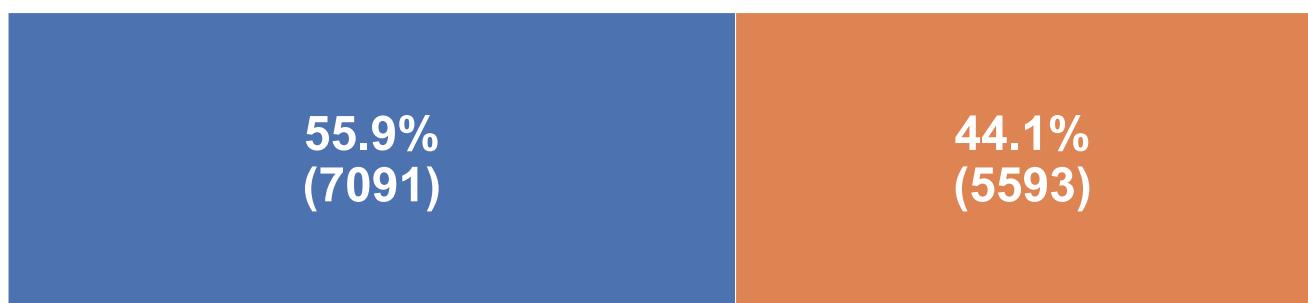
## Common Values

Value	Count	Frequency (%)
1d	7091	55.9%
2h	5593	44.1%

## Length



## Common Values (Plot)



Value	Count	Frequency (%)
1d	7091	55.9%
2h	5593	44.1%

## Most occurring characters

Value	Count	Frequency (%)
1	7091	28.0%
d	7091	28.0%
2	5593	22.0%
h	5593	22.0%

### Most occurring categories

Value	Count	Frequency (%)
Decimal Number	12684	50.0%
Lowercase Letter	12684	50.0%

### Most frequent character per category

#### *Decimal Number*

Value	Count	Frequency (%)
1	7091	55.9%
2	5593	44.1%

#### *Lowercase Letter*

Value	Count	Frequency (%)
d	7091	55.9%
h	5593	44.1%

### Most occurring scripts

Value	Count	Frequency (%)
Common	12684	50.0%
Latin	12684	50.0%

### Most frequent character per script

*Common*

Value	Count	Frequency (%)
1	7091	55.9%
2	5593	44.1%

*Latin*

Value	Count	Frequency (%)
d	7091	55.9%
h	5593	44.1%

### Most occurring blocks

Value	Count	Frequency (%)
ASCII	25368	100.0%

### Most frequent character per block

ASCII

Value	Count	Frequency (%)
1	7091	28.0%
d	7091	28.0%
2	5593	22.0%
h	5593	22.0%





## gender

Categorical

Distinct	2
Distinct (%)	< 0.1%
Missing	0
Missing (%)	0.0%
Memory size	99.2 KiB

## Length

Max length	6
Median length	6
Mean length	5.0266477
Min length	4

## Characters and Unicode

Total characters	63758
Distinct characters	6
Distinct categories	2 ( <a href="https://en.wikipedia.org/wiki/Unicode_character_property#General_Category">https://en.wikipedia.org/wiki/Unicode_character_property#General_Category</a> )
Distinct scripts	1 ( <a href="https://en.wikipedia.org/wiki/Script_(Unicode)#List_of_scripts_in_Unicode">https://en.wikipedia.org/wiki/Script_(Unicode)#List_of_scripts_in_Unicode</a> )
Distinct blocks	1 ( <a href="https://en.wikipedia.org/wiki/Unicode_block">https://en.wikipedia.org/wiki/Unicode_block</a> )

The Unicode Standard assigns character properties to each code point, which can be used to analyse textual variables.

## Unique

Unique	0	?
Unique (%)	0.0%	?

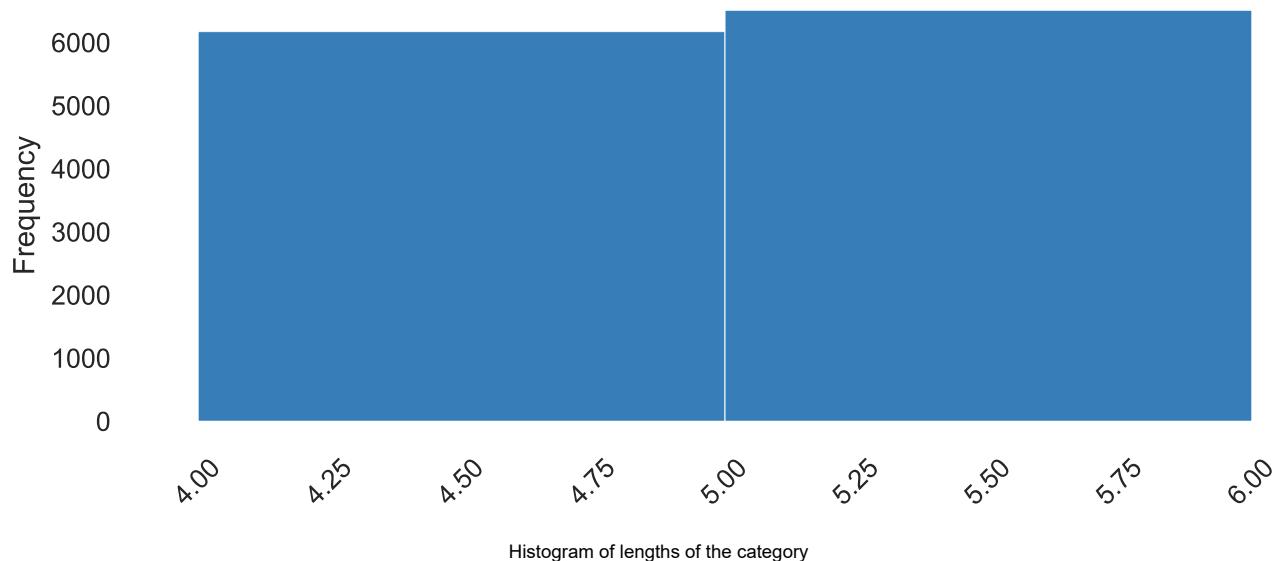
## Sample

1st row	Female
2nd row	Female
3rd row	Female
4th row	Female
5th row	Female

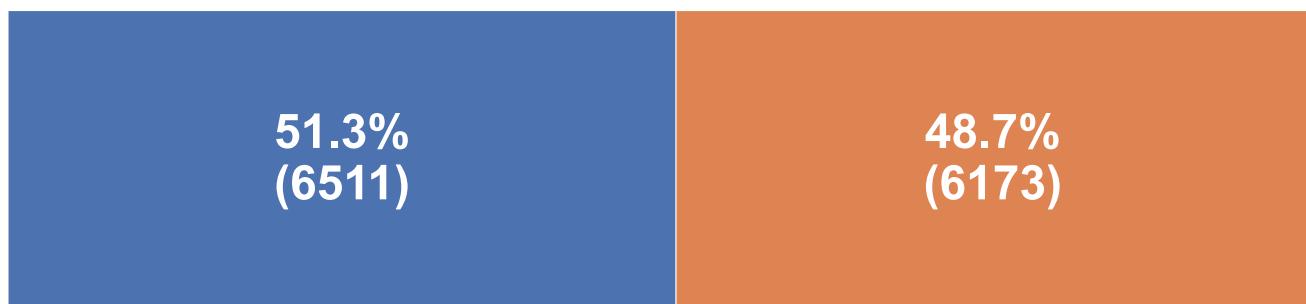
### Common Values

Value	Count	Frequency (%)
Female	6511	51.3%
Male	6173	48.7%

### Length



### Common Values (Plot)



Value	Count	Frequency (%)
female	6511	51.3%
male	6173	48.7%

## Most occurring characters

Value	Count	Frequency (%)
e	19195	30.1%
a	12684	19.9%
l	12684	19.9%
F	6511	10.2%
m	6511	10.2%
M	6173	9.7%

### Most occurring categories

Value	Count	Frequency (%)
Lowercase Letter	51074	80.1%
Uppercase Letter	12684	19.9%

### Most frequent character per category

#### *Lowercase Letter*

Value	Count	Frequency (%)
e	19195	37.6%
a	12684	24.8%
i	12684	24.8%
m	6511	12.7%

#### *Uppercase Letter*

Value	Count	Frequency (%)
F	6511	51.3%
M	6173	48.7%

### Most occurring scripts

Value	Count	Frequency (%)
Latin	63758	100.0%

### Most frequent character per script

*Latin*

Value	Count	Frequency (%)
e	19195	30.1%
a	12684	19.9%
i	12684	19.9%
F	6511	10.2%
m	6511	10.2%
M	6173	9.7%

### Most occurring blocks

Value	Count	Frequency (%)
ASCII	63758	100.0%

### Most frequent character per block

ASCII

Value	Count	Frequency (%)
e	19195	30.1%
a	12684	19.9%
i	12684	19.9%
F	6511	10.2%
m	6511	10.2%
M	6173	9.7%





**age**

Categorical

<b>Distinct</b>	8
<b>Distinct (%)</b>	0.1%
<b>Missing</b>	0
<b>Missing (%)</b>	0.0%
<b>Memory size</b>	99.2 KiB

**Length**

<b>Max length</b>	7
<b>Median length</b>	2
<b>Mean length</b>	2.779486
<b>Min length</b>	2

**Characters and Unicode**

<b>Total characters</b>	35255
<b>Distinct characters</b>	15
<b>Distinct categories</b>	<a href="https://en.wikipedia.org/wiki/Unicode_character_property#General_Category">2</a> ( <a href="https://en.wikipedia.org/wiki/Unicode_character_property#General_Category">https://en.wikipedia.org/wiki/Unicode_character_property#General_Category</a> )
<b>Distinct scripts</b>	<a href="https://en.wikipedia.org/wiki/Script_(Unicode)#List_of_scripts_in_Unicode">2</a> ( <a href="https://en.wikipedia.org/wiki/Script_(Unicode)#List_of_scripts_in_Unicode">https://en.wikipedia.org/wiki/Script_(Unicode)#List_of_scripts_in_Unicode</a> )
<b>Distinct blocks</b>	<a href="https://en.wikipedia.org/wiki/Unicode_block">1</a> ( <a href="https://en.wikipedia.org/wiki/Unicode_block">https://en.wikipedia.org/wiki/Unicode_block</a> )

The Unicode Standard assigns character properties to each code point, which can be used to analyse textual variables.

**Unique**

<b>Unique</b>	0	?
<b>Unique (%)</b>	0.0%	?

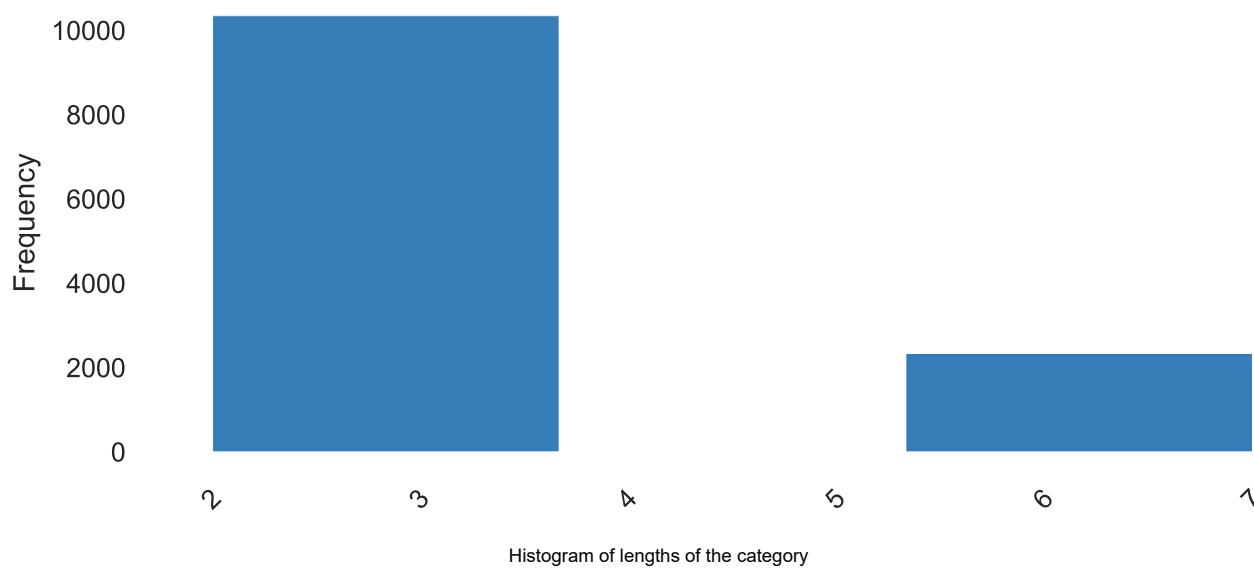
**Sample**

<b>1st row</b>	21
<b>2nd row</b>	21
<b>3rd row</b>	21
<b>4th row</b>	21
<b>5th row</b>	21

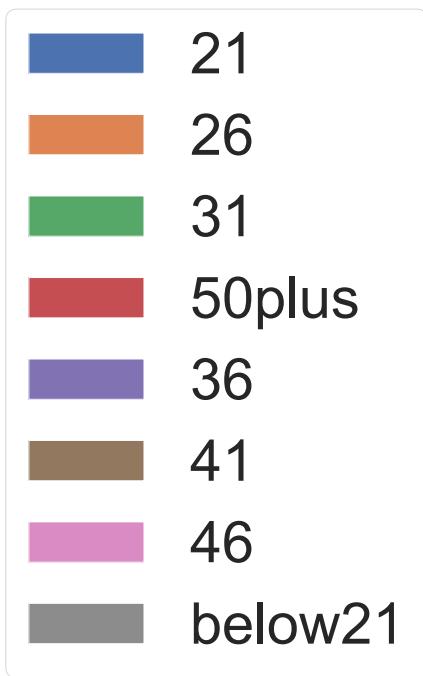
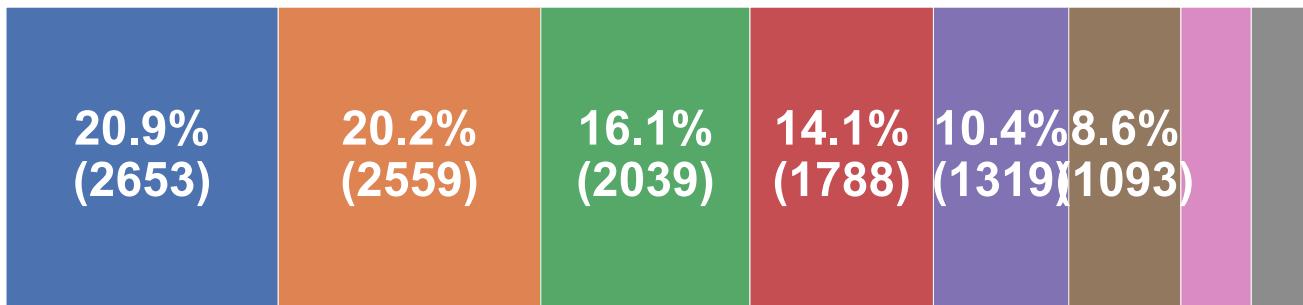
### Common Values

Value	Count	Frequency (%)
21	2653	20.9%
26	2559	20.2%
31	2039	16.1%
50plus	1788	14.1%
36	1319	10.4%
41	1093	8.6%
46	686	5.4%
below21	547	4.3%

### Length



Common Values (Plot)



Value	Count	Frequency (%)
21	2653	20.9%
26	2559	20.2%
31	2039	16.1%
50plus	1788	14.1%
36	1319	10.4%
41	1093	8.6%
46	686	5.4%
below21	547	4.3%

## Most occurring characters

Value	Count	Frequency (%)
1	6332	18.0%
2	5759	16.3%
6	4564	12.9%
3	3358	9.5%
I	2335	6.6%
5	1788	5.1%
0	1788	5.1%
p	1788	5.1%
u	1788	5.1%
s	1788	5.1%
Other values (5)	3967	11.3%

## Most occurring categories

Value	Count	Frequency (%)
Decimal Number	25368	72.0%
Lowercase Letter	9887	28.0%

## Most frequent character per category

*Lowercase Letter*

Value	Count	Frequency (%)
l	2335	23.6%
p	1788	18.1%
u	1788	18.1%
s	1788	18.1%
b	547	5.5%
e	547	5.5%
o	547	5.5%
w	547	5.5%

*Decimal Number*

Value	Count	Frequency (%)
1	6332	25.0%
2	5759	22.7%
6	4564	18.0%
3	3358	13.2%
5	1788	7.0%
0	1788	7.0%
4	1779	7.0%

## Most occurring scripts

Value	Count	Frequency (%)
Common	25368	72.0%
Latin	9887	28.0%

## Most frequent character per script

*Latin*

Value	Count	Frequency (%)
l	2335	23.6%
p	1788	18.1%
u	1788	18.1%
s	1788	18.1%
b	547	5.5%
e	547	5.5%
o	547	5.5%
w	547	5.5%

*Common*

Value	Count	Frequency (%)
1	6332	25.0%
2	5759	22.7%
6	4564	18.0%
3	3358	13.2%
5	1788	7.0%
0	1788	7.0%
4	1779	7.0%

### Most occurring blocks

Value	Count	Frequency (%)
ASCII	35255	100.0%

### Most frequent character per block

ASCII

Value	Count	Frequency (%)
1	6332	18.0%
2	5759	16.3%
6	4564	12.9%
3	3358	9.5%
I	2335	6.6%
5	1788	5.1%
0	1788	5.1%
p	1788	5.1%
u	1788	5.1%
s	1788	5.1%
Other values (5)	3967	11.3%





**maritalStatus**

Categorical

<b>Distinct</b>	5
<b>Distinct (%)</b>	< 0.1%
<b>Missing</b>	0
<b>Missing (%)</b>	0.0%
<b>Memory size</b>	99.2 KiB

**Length**

<b>Max length</b>	17
<b>Median length</b>	15
<b>Mean length</b>	11.606118
<b>Min length</b>	6

**Characters and Unicode**

<b>Total characters</b>	147212
<b>Distinct characters</b>	21
<b>Distinct categories</b>	<a href="https://en.wikipedia.org/wiki/Unicode_character_property#General_Category">3 (https://en.wikipedia.org/wiki/Unicode_character_property#General_Category)</a>
<b>Distinct scripts</b>	<a href="https://en.wikipedia.org/wiki/Script_(Unicode)#List_of_scripts_in_Unicode">2 (https://en.wikipedia.org/wiki/Script_(Unicode)#List_of_scripts_in_Unicode)</a>
<b>Distinct blocks</b>	<a href="https://en.wikipedia.org/wiki/Unicode_block">1 (https://en.wikipedia.org/wiki/Unicode_block)</a>

The Unicode Standard assigns character properties to each code point, which can be used to analyse textual variables.

**Unique**

<b>Unique</b>	0	?
<b>Unique (%)</b>	0.0%	?

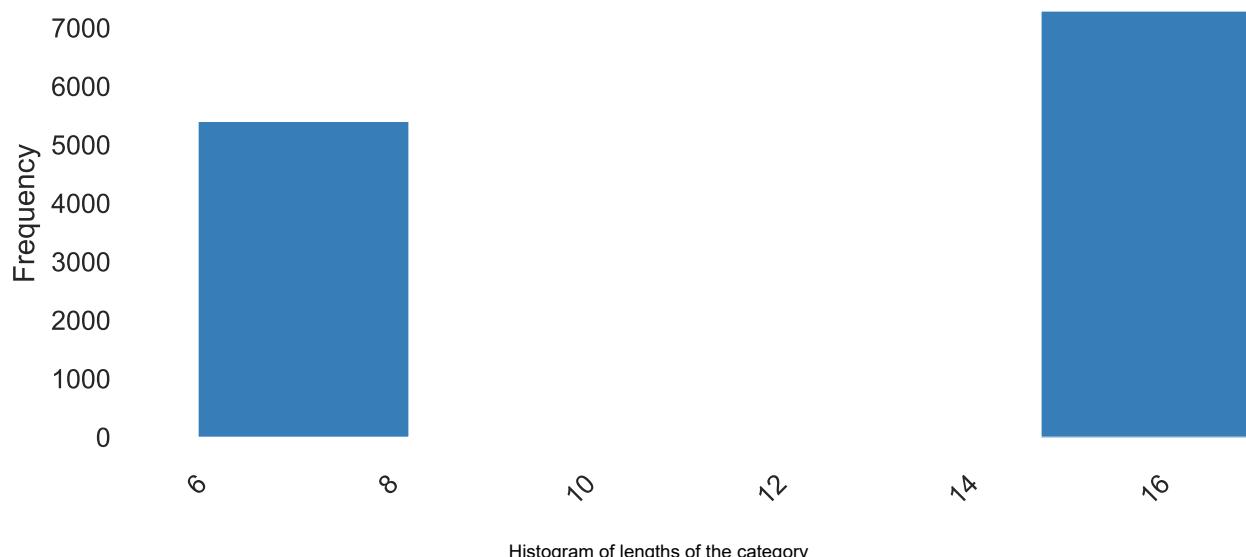
**Sample**

<b>1st row</b>	Unmarried partner
<b>2nd row</b>	Unmarried partner
<b>3rd row</b>	Unmarried partner
<b>4th row</b>	Unmarried partner
<b>5th row</b>	Unmarried partner

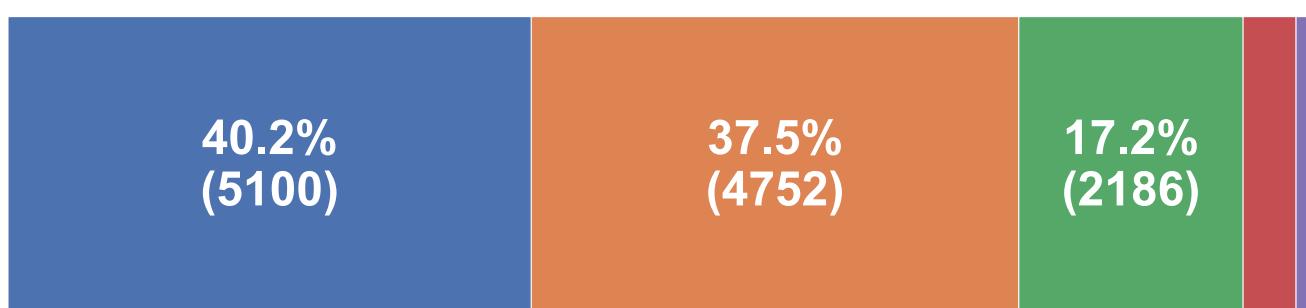
## Common Values

Value	Count	Frequency (%)
Married partner	5100	40.2%
Single	4752	37.5%
Unmarried partner	2186	17.2%
Divorced	516	4.1%
Widowed	130	1.0%

## Length



## Common Values (Plot)



Value	Count	Frequency (%)
partner	7286	36.5%
married	5100	25.5%
single	4752	23.8%
unmarried	2186	10.9%
divorced	516	2.6%
widowed	130	0.7%

## Most occurring characters

Value	Count	Frequency (%)
r	29660	20.1%
e	19970	13.6%
a	14572	9.9%
n	14224	9.7%
i	12684	8.6%
d	8062	5.5%
p	7286	4.9%
t	7286	4.9%
M	5100	3.5%
Other values (11)	21082	14.3%

## Most occurring categories

Value	Count	Frequency (%)
Lowercase Letter	127242	86.4%
Uppercase Letter	12684	8.6%
Space Separator	7286	4.9%

## Most frequent character per category

*Lowercase Letter*

Value	Count	Frequency (%)
r	29660	23.3%
e	19970	15.7%
a	14572	11.5%
n	14224	11.2%
i	12684	10.0%
d	8062	6.3%
p	7286	5.7%
t	7286	5.7%
l	4752	3.7%
g	4752	3.7%
Other values (5)	3994	3.1%

*Uppercase Letter*

Value	Count	Frequency (%)
M	5100	40.2%
S	4752	37.5%
U	2186	17.2%
D	516	4.1%
W	130	1.0%

*Space Separator*

Value	Count	Frequency (%)
	7286	100.0%

## Most occurring scripts

Value	Count	Frequency (%)
Latin	139926	95.1%
Common	7286	4.9%

## Most frequent character per script

*Latin*

Value	Count	Frequency (%)
r	29660	21.2%
e	19970	14.3%
a	14572	10.4%
n	14224	10.2%
i	12684	9.1%
d	8062	5.8%
p	7286	5.2%
t	7286	5.2%
M	5100	3.6%
l	4752	3.4%
Other values (10)	16330	11.7%

*Common*

Value	Count	Frequency (%)
	7286	100.0%

### Most occurring blocks

Value	Count	Frequency (%)
ASCII	147212	100.0%

### Most frequent character per block

ASCII

Value	Count	Frequency (%)
r	29660	20.1%
e	19970	13.6%
a	14572	9.9%
n	14224	9.7%
i	12684	8.6%
d	8062	5.5%
	7286	4.9%
p	7286	4.9%
t	7286	4.9%
M	5100	3.5%
Other values (11)	21082	14.3%





**has\_children**

Categorical

<b>Distinct</b>	2
<b>Distinct (%)</b>	< 0.1%
<b>Missing</b>	0
<b>Missing (%)</b>	0.0%
<b>Memory size</b>	99.2 KiB

**Length**

<b>Max length</b>	1
<b>Median length</b>	1
<b>Mean length</b>	1
<b>Min length</b>	1

**Characters and Unicode**

<b>Total characters</b>	12684
<b>Distinct characters</b>	2
<b>Distinct categories</b>	1 ( <a href="https://en.wikipedia.org/wiki/Unicode_character_property#General_Category">https://en.wikipedia.org/wiki/Unicode_character_property#General_Category</a> )
<b>Distinct scripts</b>	1 ( <a href="https://en.wikipedia.org/wiki/Script_(Unicode)#List_of_scripts_in_Unicode">https://en.wikipedia.org/wiki/Script_(Unicode)#List_of_scripts_in_Unicode</a> )
<b>Distinct blocks</b>	1 ( <a href="https://en.wikipedia.org/wiki/Unicode_block">https://en.wikipedia.org/wiki/Unicode_block</a> )

The Unicode Standard assigns character properties to each code point, which can be used to analyse textual variables.

**Unique**

<b>Unique</b>	0	?
<b>Unique (%)</b>	0.0%	

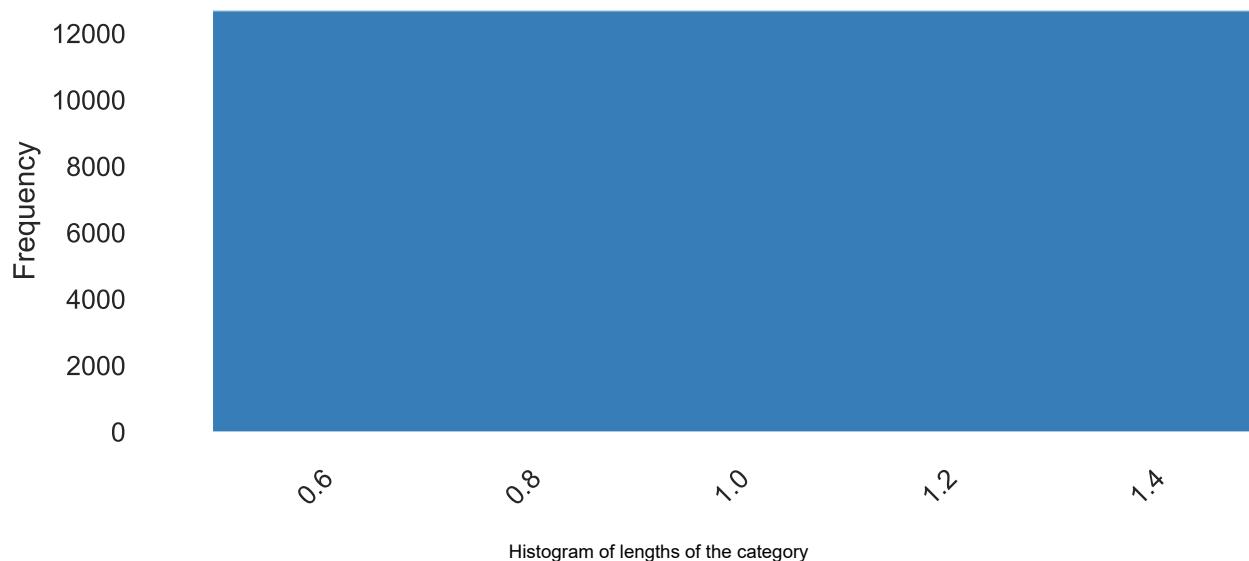
**Sample**

<b>1st row</b>	1
<b>2nd row</b>	1
<b>3rd row</b>	1
<b>4th row</b>	1
<b>5th row</b>	1

## Common Values

Value	Count	Frequency (%)
0	7431	58.6%
1	5253	41.4%

## Length



## Common Values (Plot)



Value	Count	Frequency (%)
0	7431	58.6%
1	5253	41.4%

## Most occurring characters

Value	Count	Frequency (%)
0	7431	58.6%
1	5253	41.4%

### Most occurring categories

Value	Count	Frequency (%)
Decimal Number	12684	100.0%

### Most frequent character per category

*Decimal Number*

Value	Count	Frequency (%)
0	7431	58.6%
1	5253	41.4%

### Most occurring scripts

Value	Count	Frequency (%)
Common	12684	100.0%

### Most frequent character per script

*Common*

Value	Count	Frequency (%)
0	7431	58.6%
1	5253	41.4%

### Most occurring blocks

Value	Count	Frequency (%)
ASCII	12684	100.0%

### Most frequent character per block

ASCII

Value	Count	Frequency (%)
0	7431	58.6%
1	5253	41.4%





**education**

Categorical

<b>Distinct</b>	6
<b>Distinct (%)</b>	< 0.1%
<b>Missing</b>	0
<b>Missing (%)</b>	0.0%
<b>Memory size</b>	99.2 KiB

**Length**

<b>Max length</b>	38
<b>Median length</b>	24
<b>Mean length</b>	22.332781
<b>Min length</b>	16

**Characters and Unicode**

<b>Total characters</b>	283269
<b>Distinct characters</b>	26
<b>Distinct categories</b>	6 ( <a href="https://en.wikipedia.org/wiki/Unicode_character_property#General_Category">https://en.wikipedia.org/wiki/Unicode_character_property#General_Category</a> )
<b>Distinct scripts</b>	2 ( <a href="https://en.wikipedia.org/wiki/Script_(Unicode)#List_of_scripts_in_Unicode">https://en.wikipedia.org/wiki/Script_(Unicode)#List_of_scripts_in_Unicode</a> )
<b>Distinct blocks</b>	1 ( <a href="https://en.wikipedia.org/wiki/Unicode_block">https://en.wikipedia.org/wiki/Unicode_block</a> )

The Unicode Standard assigns character properties to each code point, which can be used to analyse textual variables.

**Unique**

<b>Unique</b>	0	?
<b>Unique (%)</b>	0.0%	?

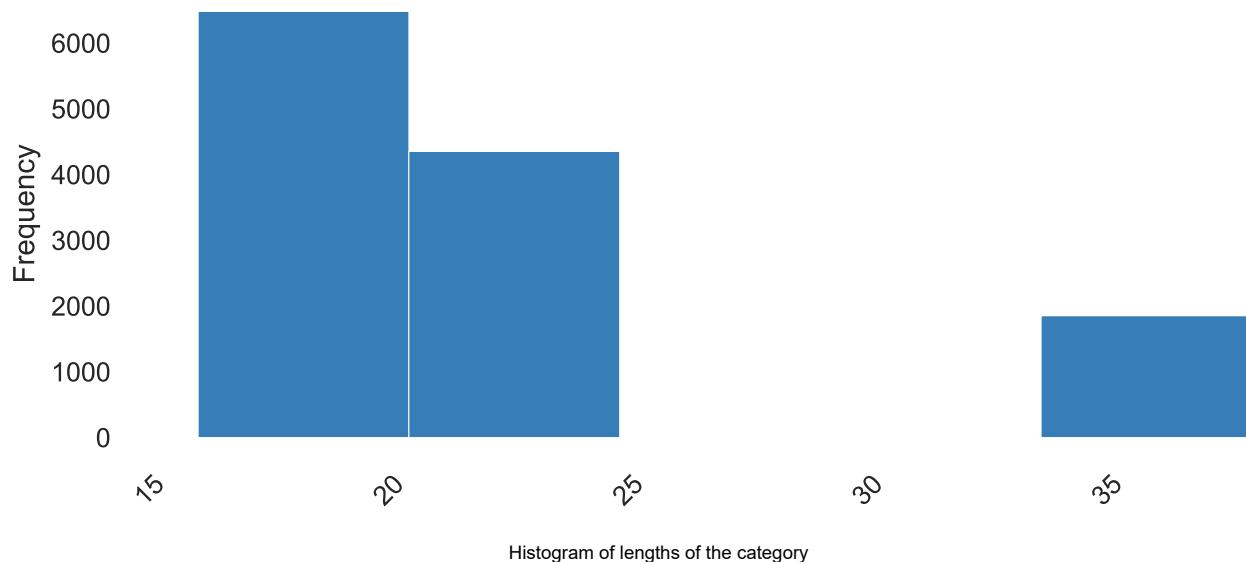
**Sample**

<b>1st row</b>	Some college - no degree
<b>2nd row</b>	Some college - no degree
<b>3rd row</b>	Some college - no degree
<b>4th row</b>	Some college - no degree
<b>5th row</b>	Some college - no degree

### Common Values

Value	Count	Frequency (%)
Some college - no degree	4351	34.3%
Bachelors degree	4335	34.2%
Graduate degree (Masters or Doctorate)	1852	14.6%
Associates degree	1153	9.1%
High School Graduate	905	7.1%
Some High School	88	0.7%

### Length



### Common Values (Plot)



- Some college - no degree
- Bachelors degree
- Graduate degree (Masters or Doctorate)
- Associates degree
- High School Graduate
- Some High School

Value	Count	Frequency (%)
degree	11691	26.0%
some	4439	9.9%
college	4351	9.7%
	4351	9.7%
no	4351	9.7%
bachelors	4335	9.6%
graduate	2757	6.1%
masters	1852	4.1%
or	1852	4.1%
doctorate	1852	4.1%
Other values (3)	3139	7.0%

## Most occurring characters

Value	Count	Frequency (%)
e	60163	21.2%
	32286	11.4%
o	26171	9.2%
r	24339	8.6%
g	17035	6.0%
a	14706	5.2%
d	14448	5.1%
l	14030	5.0%
c	12684	4.5%
s	11498	4.1%
Other values (16)	55909	19.7%

### Most occurring categories

Value	Count	Frequency (%)
Lowercase Letter	224554	79.3%
Space Separator	32286	11.4%
Uppercase Letter	18374	6.5%
Dash Punctuation	4351	1.5%
Open Punctuation	1852	0.7%
Close Punctuation	1852	0.7%

### Most frequent character per category

*Lowercase Letter*

Value	Count	Frequency (%)
e	60163	26.8%
o	26171	11.7%
r	24339	10.8%
g	17035	7.6%
a	14706	6.5%
d	14448	6.4%
l	14030	6.2%
c	12684	5.6%
s	11498	5.1%
t	9466	4.2%
Other values (5)	20014	8.9%

*Uppercase Letter*

Value	Count	Frequency (%)
S	5432	29.6%
B	4335	23.6%
G	2757	15.0%
M	1852	10.1%
D	1852	10.1%
A	1153	6.3%
H	993	5.4%

*Space Separator*

Value	Count	Frequency (%)
	32286	100.0%

*Dash Punctuation*

Value	Count	Frequency (%)
-	4351	100.0%

*Open Punctuation*

Value	Count	Frequency (%)
(	1852	100.0%
<i>Close Punctuation</i>		
Value	Count	Frequency (%)
)	1852	100.0%

## Most occurring scripts

Value	Count	Frequency (%)
Latin	242928	85.8%
Common	40341	14.2%

## Most frequent character per script

*Latin*

Value	Count	Frequency (%)
e	60163	24.8%
o	26171	10.8%
r	24339	10.0%
g	17035	7.0%
a	14706	6.1%
d	14448	5.9%
l	14030	5.8%
c	12684	5.2%
s	11498	4.7%
t	9466	3.9%
Other values (12)	38388	15.8%

*Common*

Value	Count	Frequency (%)
-	32286	80.0%
(	4351	10.8%
)	1852	4.6%
)	1852	4.6%

### Most occurring blocks

Value	Count	Frequency (%)
ASCII	283269	100.0%

### Most frequent character per block

ASCII

Value	Count	Frequency (%)
e	60163	21.2%
o	32286	11.4%
r	26171	9.2%
g	24339	8.6%
a	17035	6.0%
d	14706	5.2%
l	14448	5.1%
c	14030	5.0%
s	12684	4.5%
Other values (16)	11498	4.1%
	55909	19.7%





**occupation**

Categorical

<b>Distinct</b>	25
<b>Distinct (%)</b>	0.2%
<b>Missing</b>	0
<b>Missing (%)</b>	0.0%
<b>Memory size</b>	99.2 KiB

**Length**

<b>Max length</b>	41
<b>Median length</b>	36
<b>Mean length</b>	18.81859
<b>Min length</b>	5

**Characters and Unicode**

<b>Total characters</b>	238695
<b>Distinct characters</b>	40
<b>Distinct categories</b>	<a href="https://en.wikipedia.org/wiki/Unicode_character_property#General_Category">4 (https://en.wikipedia.org/wiki/Unicode_character_property#General_Category)</a>
<b>Distinct scripts</b>	<a href="https://en.wikipedia.org/wiki/Script_(Unicode)#List_of_scripts_in_Unicode">2 (https://en.wikipedia.org/wiki/Script_(Unicode)#List_of_scripts_in_Unicode)</a>
<b>Distinct blocks</b>	<a href="https://en.wikipedia.org/wiki/Unicode_block">1 (https://en.wikipedia.org/wiki/Unicode_block)</a>

The Unicode Standard assigns character properties to each code point, which can be used to analyse textual variables.

**Unique**

<b>Unique</b>	0	?
<b>Unique (%)</b>	0.0%	?

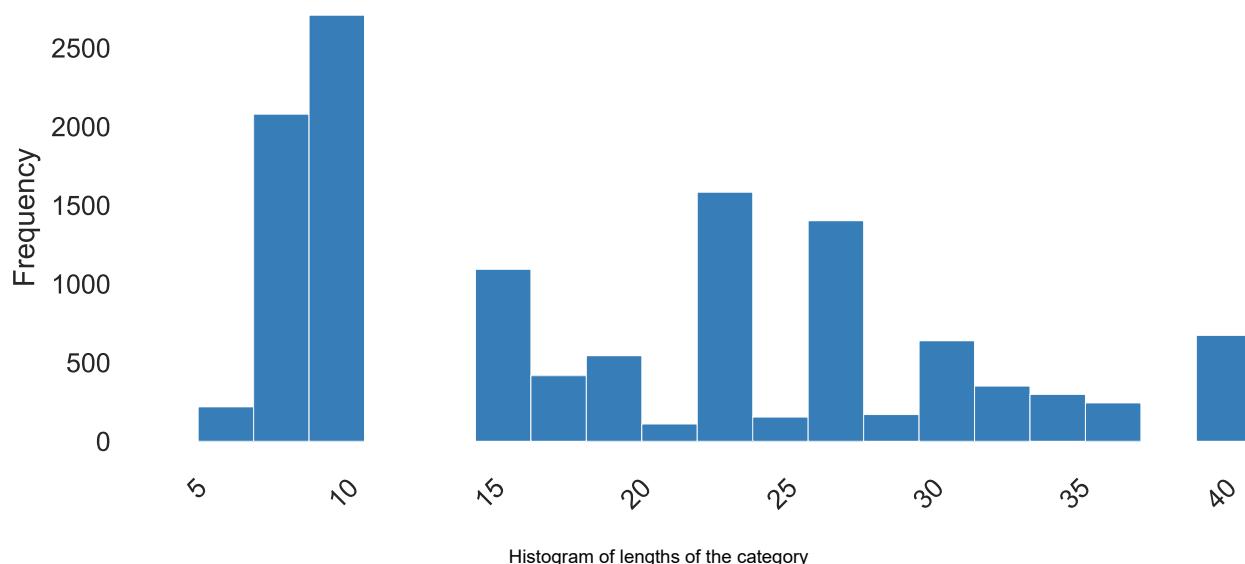
**Sample**

<b>1st row</b>	Unemployed
<b>2nd row</b>	Unemployed
<b>3rd row</b>	Unemployed
<b>4th row</b>	Unemployed
<b>5th row</b>	Unemployed

## Common Values

Value	Count	Frequency (%)
Unemployed	1870	14.7%
Student	1584	12.5%
Computer & Mathematical	1408	11.1%
Sales & Related	1093	8.6%
Education&Training&Library	943	7.4%
Management	838	6.6%
Office & Administrative Support	639	5.0%
Arts Design Entertainment Sports & Media	629	5.0%
Business & Financial	544	4.3%
Retired	495	3.9%
Other values (15)	2641	20.8%

## Length



Value	Count	Frequency (%)
	6082	20.2%
unemployed	1870	6.2%
student	1584	5.3%
computer	1408	4.7%
mathematical	1408	4.7%
related	1391	4.6%
sales	1093	3.6%
education&training&library	943	3.1%
support	881	2.9%
management	838	2.8%
Other values (45)	12607	41.9%

## Most occurring characters

Value	Count	Frequency (%)
e	25178	10.5%
t	20552	8.6%
a	19380	8.1%
	17421	7.3%
i	17417	7.3%
n	17336	7.3%
r	13026	5.5%
o	9537	4.0%
l	8583	3.6%
d	8047	3.4%
Other values (30)	82218	34.4%

### Most occurring categories

Value	Count	Frequency (%)
Lowercase Letter	187397	78.5%
Uppercase Letter	25909	10.9%
Space Separator	17421	7.3%
Other Punctuation	7968	3.3%

### Most frequent character per category

#### *Lowercase Letter*

Value	Count	Frequency (%)
e	25178	13.4%
t	20552	11.0%
a	19380	10.3%
i	17417	9.3%
n	17336	9.3%
r	13026	7.0%
o	9537	5.1%
l	8583	4.6%
d	8047	4.3%
c	7604	4.1%
Other values (11)	40737	21.7%

#### *Uppercase Letter*

Value	Count	Frequency (%)
S	5657	21.8%
M	3488	13.5%
C	2022	7.8%
R	2019	7.8%
E	1901	7.3%
U	1870	7.2%
A	1443	5.6%
T	1405	5.4%
L	1332	5.1%
P	1172	4.5%
Other values (7)	3600	13.9%

#### *Space Separator*

Value	Count	Frequency (%)
	17421	100.0%

#### *Other Punctuation*

Value	Count	Frequency (%)
&	7968	100.0%

## Most occurring scripts

Value	Count	Frequency (%)
Latin	213306	89.4%
Common	25389	10.6%

## Most frequent character per script

*Latin*

Value	Count	Frequency (%)
e	25178	11.8%
t	20552	9.6%
a	19380	9.1%
i	17417	8.2%
n	17336	8.1%
r	13026	6.1%
o	9537	4.5%
l	8583	4.0%
d	8047	3.8%
c	7604	3.6%
Other values (28)	66646	31.2%

*Common*

Value	Count	Frequency (%)
&	17421	68.6%
&	7968	31.4%

### Most occurring blocks

Value	Count	Frequency (%)
ASCII	238695	100.0%

### Most frequent character per block

ASCII

Value	Count	Frequency (%)
e	25178	10.5%
t	20552	8.6%
a	19380	8.1%
	17421	7.3%
i	17417	7.3%
n	17336	7.3%
r	13026	5.5%
o	9537	4.0%
l	8583	3.6%
d	8047	3.4%
Other values (30)	82218	34.4%





**income**

Categorical

<b>Distinct</b>	9
<b>Distinct (%)</b>	0.1%
<b>Missing</b>	0
<b>Missing (%)</b>	0.0%
<b>Memory size</b>	99.2 KiB

**Length**

<b>Max length</b>	16
<b>Median length</b>	15
<b>Mean length</b>	15.082151
<b>Min length</b>	15

**Characters and Unicode**

<b>Total characters</b>	191302
<b>Distinct characters</b>	23
<b>Distinct categories</b>	6 ( <a href="https://en.wikipedia.org/wiki/Unicode_character_property#General_Category">https://en.wikipedia.org/wiki/Unicode_character_property#General_Category</a> )
<b>Distinct scripts</b>	2 ( <a href="https://en.wikipedia.org/wiki/Script_(Unicode)#List_of_scripts_in_Unicode">https://en.wikipedia.org/wiki/Script_(Unicode)#List_of_scripts_in_Unicode</a> )
<b>Distinct blocks</b>	1 ( <a href="https://en.wikipedia.org/wiki/Unicode_block">https://en.wikipedia.org/wiki/Unicode_block</a> )

The Unicode Standard assigns character properties to each code point, which can be used to analyse textual variables.

**Unique**

<b>Unique</b>	0	?
<b>Unique (%)</b>	0.0%	?

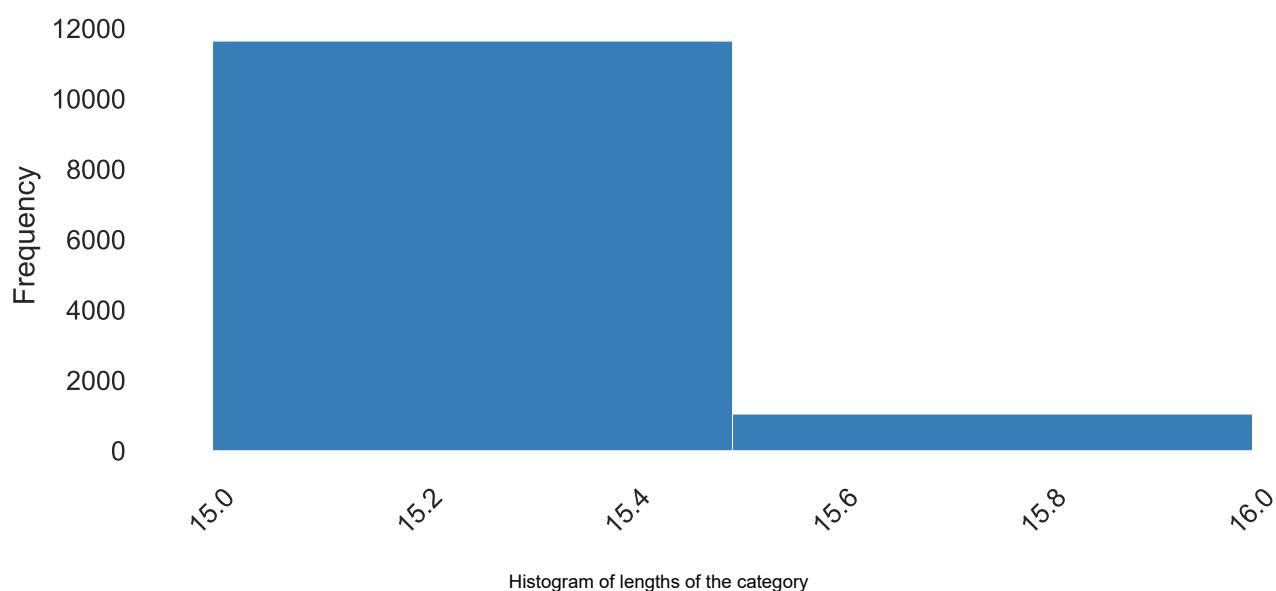
**Sample**

<b>1st row</b>	\$37500 - \$49999
<b>2nd row</b>	\$37500 - \$49999
<b>3rd row</b>	\$37500 - \$49999
<b>4th row</b>	\$37500 - \$49999
<b>5th row</b>	\$37500 - \$49999

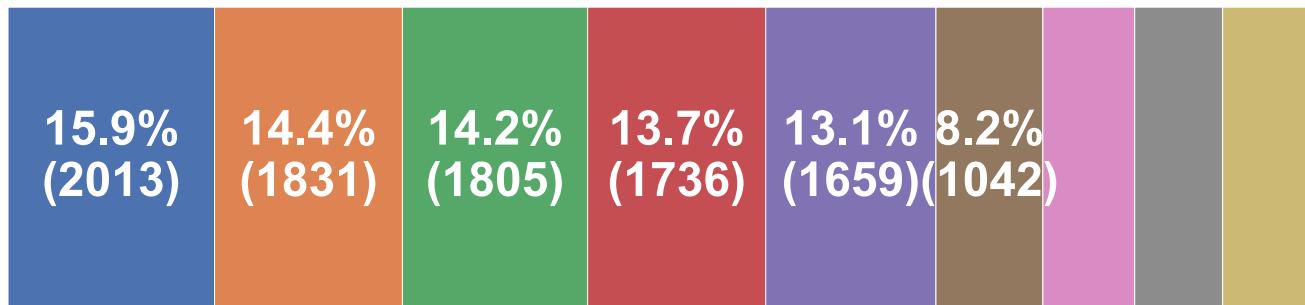
### Common Values

Value	Count	Frequency (%)
\$25000 - \$37499	2013	15.9%
\$12500 - \$24999	1831	14.4%
\$37500 - \$49999	1805	14.2%
\$100000 or More	1736	13.7%
\$50000 - \$62499	1659	13.1%
Less than \$12500	1042	8.2%
\$87500 - \$99999	895	7.1%
\$75000 - \$87499	857	6.8%
\$62500 - \$74999	846	6.7%

### Length



Common Values (Plot)



Value	Count	Frequency (%)
	9906	26.0%
12500	2873	7.6%
25000	2013	5.3%
37499	2013	5.3%
24999	1831	4.8%
37500	1805	4.7%
49999	1805	4.7%
100000	1736	4.6%
or	1736	4.6%
more	1736	4.6%
Other values (10)	10598	27.9%

## Most occurring characters

Value	Count	Frequency (%)
0	36764	19.2%
9	28784	15.0%
	25368	13.3%
\$	22590	11.8%
5	10948	5.7%
-	9906	5.2%
2	9222	4.8%
4	9011	4.7%
7	7273	3.8%
1	4609	2.4%
Other values (13)	26827	14.0%

### Most occurring categories

Value	Count	Frequency (%)
Decimal Number	114686	60.0%
Space Separator	25368	13.3%
Currency Symbol	22590	11.8%
Lowercase Letter	15974	8.4%
Dash Punctuation	9906	5.2%
Uppercase Letter	2778	1.5%

### Most frequent character per category

#### *Decimal Number*

Value	Count	Frequency (%)
0	36764	32.1%
9	28784	25.1%
5	10948	9.5%
2	9222	8.0%
4	9011	7.9%
7	7273	6.3%
1	4609	4.0%
3	3818	3.3%
6	2505	2.2%
8	1752	1.5%

#### *Lowercase Letter*

Value	Count	Frequency (%)
r	3472	21.7%
o	3472	21.7%
e	2778	17.4%
s	2084	13.0%
t	1042	6.5%
h	1042	6.5%
a	1042	6.5%
n	1042	6.5%

#### *Uppercase Letter*

Value	Count	Frequency (%)
M	1736	62.5%
L	1042	37.5%

#### *Space Separator*

Value	Count	Frequency (%)
	25368	100.0%

*Currency Symbol*

Value	Count	Frequency (%)
\$	22590	100.0%

*Dash Punctuation*

Value	Count	Frequency (%)
-	9906	100.0%

### Most occurring scripts

Value	Count	Frequency (%)
Common	172550	90.2%
Latin	18752	9.8%

### Most frequent character per script

#### *Common*

Value	Count	Frequency (%)
0	36764	21.3%
9	28784	16.7%
\$	25368	14.7%
5	22590	13.1%
-	10948	6.3%
2	9906	5.7%
4	9222	5.3%
7	9011	5.2%
1	7273	4.2%
Other values (3)	4609	2.7%
	8075	4.7%

#### *Latin*

Value	Count	Frequency (%)
r	3472	18.5%
o	3472	18.5%
e	2778	14.8%
s	2084	11.1%
M	1736	9.3%
L	1042	5.6%
t	1042	5.6%
h	1042	5.6%
a	1042	5.6%
n	1042	5.6%

### Most occurring blocks

Value	Count	Frequency (%)
ASCII	191302	100.0%

### Most frequent character per block

ASCII

Value	Count	Frequency (%)
0	36764	19.2%
9	28784	15.0%
	25368	13.3%
\$	22590	11.8%
5	10948	5.7%
-	9906	5.2%
2	9222	4.8%
4	9011	4.7%
7	7273	3.8%
1	4609	2.4%
Other values (13)	26827	14.0%





**Bar**

Categorical

<b>Distinct</b>	5
<b>Distinct (%)</b>	< 0.1%
<b>Missing</b>	107
<b>Missing (%)</b>	0.8%
<b>Memory size</b>	99.2 KiB

**Length**

<b>Max length</b>	5
<b>Median length</b>	5
<b>Mean length</b>	4.3801383
<b>Min length</b>	3

**Characters and Unicode**

<b>Total characters</b>	55089
<b>Distinct characters</b>	13
<b>Distinct categories</b>	<a href="https://en.wikipedia.org/wiki/Unicode_character_property#General_Category">3 (https://en.wikipedia.org/wiki/Unicode_character_property#General_Category)</a>
<b>Distinct scripts</b>	<a href="https://en.wikipedia.org/wiki/Script_(Unicode)#List_of_scripts_in_Unicode">2 (https://en.wikipedia.org/wiki/Script_(Unicode)#List_of_scripts_in_Unicode)</a>
<b>Distinct blocks</b>	<a href="https://en.wikipedia.org/wiki/Unicode_block">1 (https://en.wikipedia.org/wiki/Unicode_block)</a>

The Unicode Standard assigns character properties to each code point, which can be used to analyse textual variables.

**Unique**

<b>Unique</b>	0	?
<b>Unique (%)</b>	0.0%	

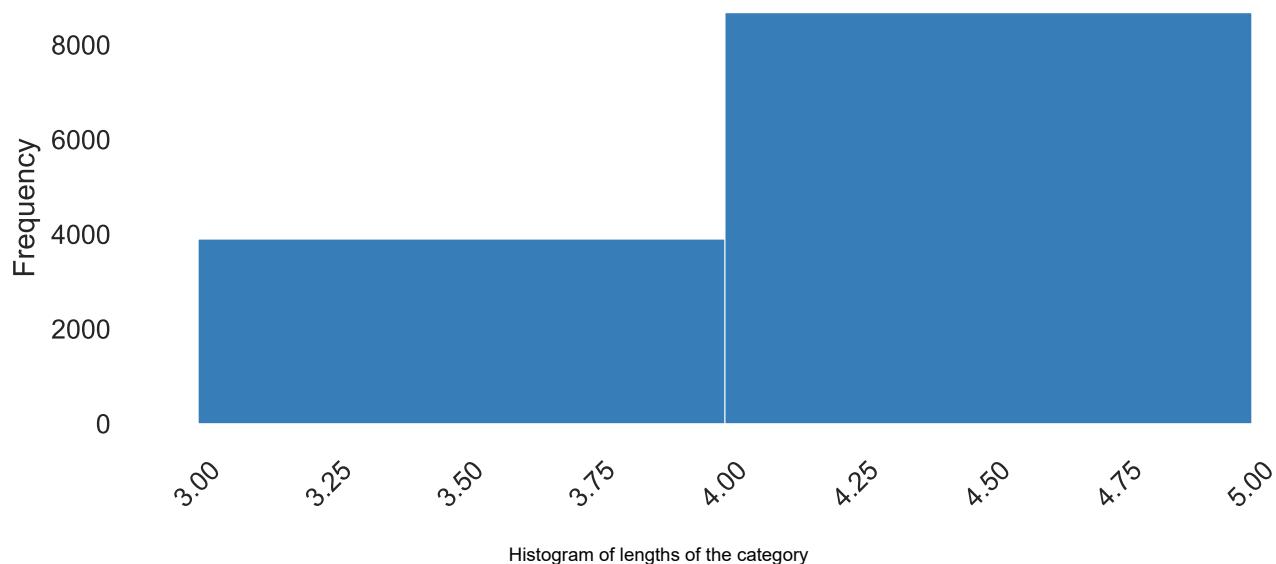
**Sample**

<b>1st row</b>	never
<b>2nd row</b>	never
<b>3rd row</b>	never
<b>4th row</b>	never
<b>5th row</b>	never

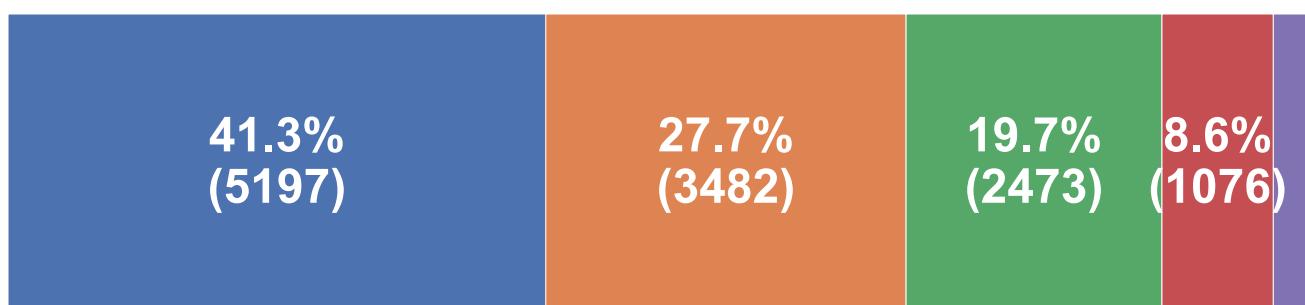
## Common Values

Value	Count	Frequency (%)
never	5197	41.0%
less1	3482	27.5%
1~3	2473	19.5%
4~8	1076	8.5%
gt8	349	2.8%
(Missing)	107	0.8%

## Length



## Common Values (Plot)



Value	Count	Frequency (%)
never	5197	41.3%
less1	3482	27.7%
1~3	2473	19.7%
4~8	1076	8.6%
gt8	349	2.8%

## Most occurring characters

Value	Count	Frequency (%)
e	13876	25.2%
s	6964	12.6%
l	5955	10.8%
n	5197	9.4%
v	5197	9.4%
r	5197	9.4%
~	3549	6.4%
l	3482	6.3%
3	2473	4.5%
8	1425	2.6%
Other values (3)	1774	3.2%

## Most occurring categories

Value	Count	Frequency (%)
Lowercase Letter	40611	73.7%
Decimal Number	10929	19.8%
Math Symbol	3549	6.4%

## Most frequent character per category

*Lowercase Letter*

Value	Count	Frequency (%)
e	13876	34.2%
s	6964	17.1%
n	5197	12.8%
v	5197	12.8%
r	5197	12.8%
l	3482	8.6%
g	349	0.9%
t	349	0.9%

*Decimal Number*

Value	Count	Frequency (%)
1	5955	54.5%
3	2473	22.6%
8	1425	13.0%
4	1076	9.8%

*Math Symbol*

Value	Count	Frequency (%)
~	3549	100.0%

## Most occurring scripts

Value	Count	Frequency (%)
Latin	40611	73.7%
Common	14478	26.3%

## Most frequent character per script

*Latin*

Value	Count	Frequency (%)
e	13876	34.2%
s	6964	17.1%
n	5197	12.8%
v	5197	12.8%
r	5197	12.8%
l	3482	8.6%
g	349	0.9%
t	349	0.9%

*Common*

Value	Count	Frequency (%)
1	5955	41.1%
~	3549	24.5%
3	2473	17.1%
8	1425	9.8%
4	1076	7.4%

### Most occurring blocks

Value	Count	Frequency (%)
ASCII	55089	100.0%

### Most frequent character per block

ASCII

Value	Count	Frequency (%)
e	13876	25.2%
s	6964	12.6%
l	5955	10.8%
n	5197	9.4%
v	5197	9.4%
r	5197	9.4%
~	3549	6.4%
l	3482	6.3%
3	2473	4.5%
8	1425	2.6%
Other values (3)	1774	3.2%





**CoffeeHouse**

Categorical

<b>Distinct</b>	5
<b>Distinct (%)</b>	< 0.1%
<b>Missing</b>	217
<b>Missing (%)</b>	1.7%
<b>Memory size</b>	99.2 KiB

**Length**

<b>Max length</b>	5
<b>Median length</b>	5
<b>Mean length</b>	4.0182081
<b>Min length</b>	3

**Characters and Unicode**

<b>Total characters</b>	50095
<b>Distinct characters</b>	13
<b>Distinct categories</b>	<a href="https://en.wikipedia.org/wiki/Unicode_character_property#General_Category">3 (https://en.wikipedia.org/wiki/Unicode_character_property#General_Category)</a>
<b>Distinct scripts</b>	<a href="https://en.wikipedia.org/wiki/Script_(Unicode)#List_of_scripts_in_Unicode">2 (https://en.wikipedia.org/wiki/Script_(Unicode)#List_of_scripts_in_Unicode)</a>
<b>Distinct blocks</b>	<a href="https://en.wikipedia.org/wiki/Unicode_block">1 (https://en.wikipedia.org/wiki/Unicode_block)</a>

The Unicode Standard assigns character properties to each code point, which can be used to analyse textual variables.

**Unique**

<b>Unique</b>	0	?
<b>Unique (%)</b>	0.0%	

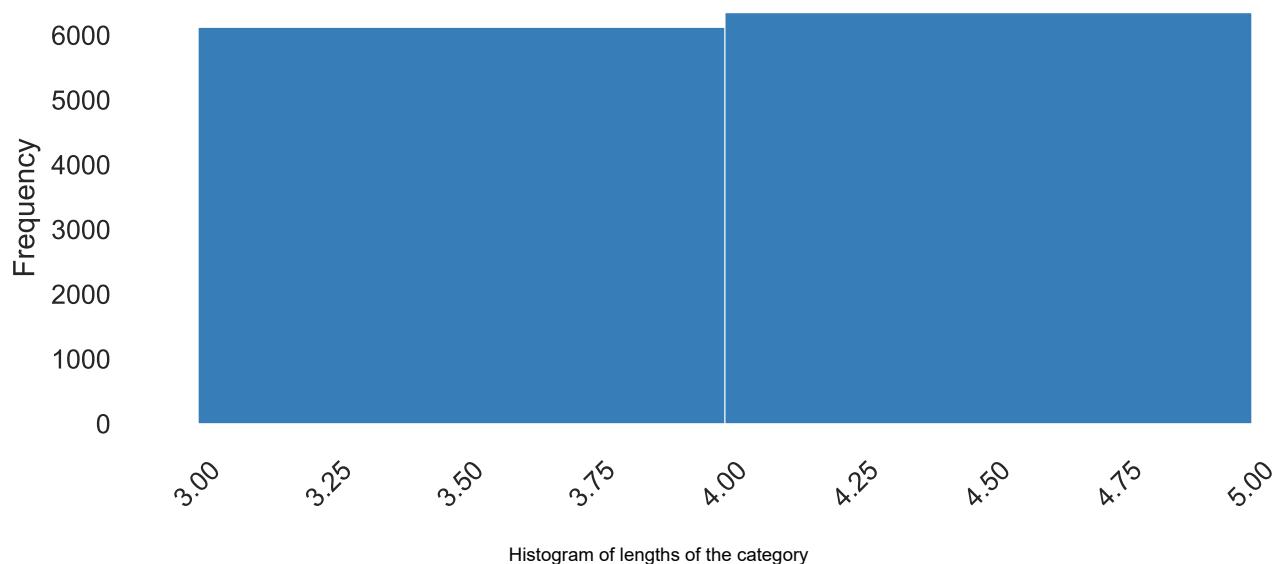
**Sample**

<b>1st row</b>	never
<b>2nd row</b>	never
<b>3rd row</b>	never
<b>4th row</b>	never
<b>5th row</b>	never

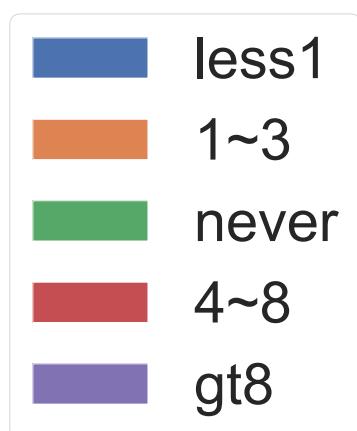
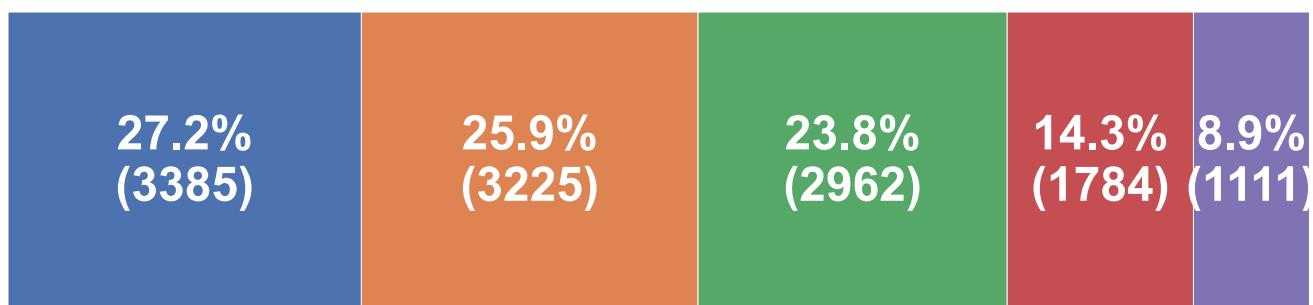
## Common Values

Value	Count	Frequency (%)
less1	3385	26.7%
1~3	3225	25.4%
never	2962	23.4%
4~8	1784	14.1%
gt8	1111	8.8%
(Missing)	217	1.7%

## Length



## Common Values (Plot)



Value	Count	Frequency (%)
less1	3385	27.2%
1~3	3225	25.9%
never	2962	23.8%
4~8	1784	14.3%
gt8	1111	8.9%

## Most occurring characters

Value	Count	Frequency (%)
e	9309	18.6%
s	6770	13.5%
1	6610	13.2%
~	5009	10.0%
l	3385	6.8%
3	3225	6.4%
n	2962	5.9%
v	2962	5.9%
r	2962	5.9%
8	2895	5.8%
Other values (3)	4006	8.0%

## Most occurring categories

Value	Count	Frequency (%)
Lowercase Letter	30572	61.0%
Decimal Number	14514	29.0%
Math Symbol	5009	10.0%

## Most frequent character per category

*Lowercase Letter*

Value	Count	Frequency (%)
e	9309	30.4%
s	6770	22.1%
l	3385	11.1%
n	2962	9.7%
v	2962	9.7%
r	2962	9.7%
g	1111	3.6%
t	1111	3.6%

*Decimal Number*

Value	Count	Frequency (%)
1	6610	45.5%
3	3225	22.2%
8	2895	19.9%
4	1784	12.3%

*Math Symbol*

Value	Count	Frequency (%)
~	5009	100.0%

## Most occurring scripts

Value	Count	Frequency (%)
Latin	30572	61.0%
Common	19523	39.0%

## Most frequent character per script

*Latin*

Value	Count	Frequency (%)
e	9309	30.4%
s	6770	22.1%
l	3385	11.1%
n	2962	9.7%
v	2962	9.7%
r	2962	9.7%
g	1111	3.6%
t	1111	3.6%

*Common*

Value	Count	Frequency (%)
1	6610	33.9%
~	5009	25.7%
3	3225	16.5%
8	2895	14.8%
4	1784	9.1%

### Most occurring blocks

Value	Count	Frequency (%)
ASCII	50095	100.0%

### Most frequent character per block

ASCII

Value	Count	Frequency (%)
e	9309	18.6%
s	6770	13.5%
1	6610	13.2%
~	5009	10.0%
l	3385	6.8%
3	3225	6.4%
n	2962	5.9%
v	2962	5.9%
r	2962	5.9%
8	2895	5.8%
Other values (3)	4006	8.0%





**CarryAway**

Categorical

<b>Distinct</b>	5
<b>Distinct (%)</b>	< 0.1%
<b>Missing</b>	151
<b>Missing (%)</b>	1.2%
<b>Memory size</b>	99.2 KiB

**Length**

<b>Max length</b>	5
<b>Median length</b>	3
<b>Mean length</b>	3.3205936
<b>Min length</b>	3

**Characters and Unicode**

<b>Total characters</b>	41617
<b>Distinct characters</b>	13
<b>Distinct categories</b>	<a href="https://en.wikipedia.org/wiki/Unicode_character_property#General_Category">3 (https://en.wikipedia.org/wiki/Unicode_character_property#General_Category)</a>
<b>Distinct scripts</b>	<a href="https://en.wikipedia.org/wiki/Script_(Unicode)#List_of_scripts_in_Unicode">2 (https://en.wikipedia.org/wiki/Script_(Unicode)#List_of_scripts_in_Unicode)</a>
<b>Distinct blocks</b>	<a href="https://en.wikipedia.org/wiki/Unicode_block">1 (https://en.wikipedia.org/wiki/Unicode_block)</a>

The Unicode Standard assigns character properties to each code point, which can be used to analyse textual variables.

**Unique**

<b>Unique</b>	0	?
<b>Unique (%)</b>	0.0%	

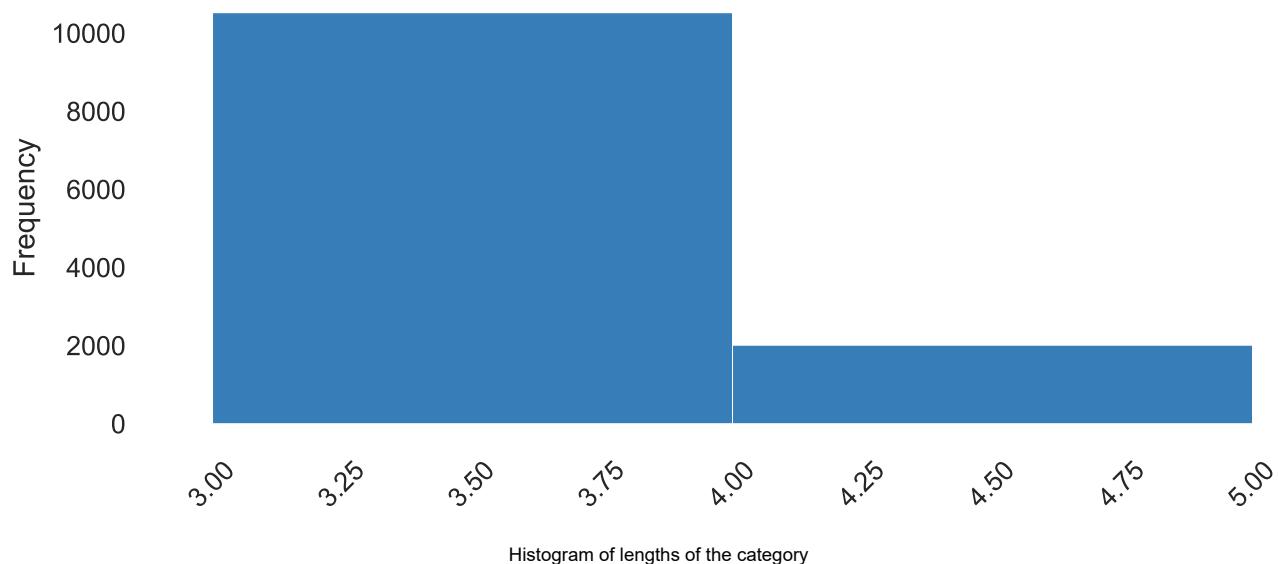
**Sample**

<b>1st row</b>	4~8
<b>2nd row</b>	4~8
<b>3rd row</b>	4~8
<b>4th row</b>	4~8
<b>5th row</b>	4~8

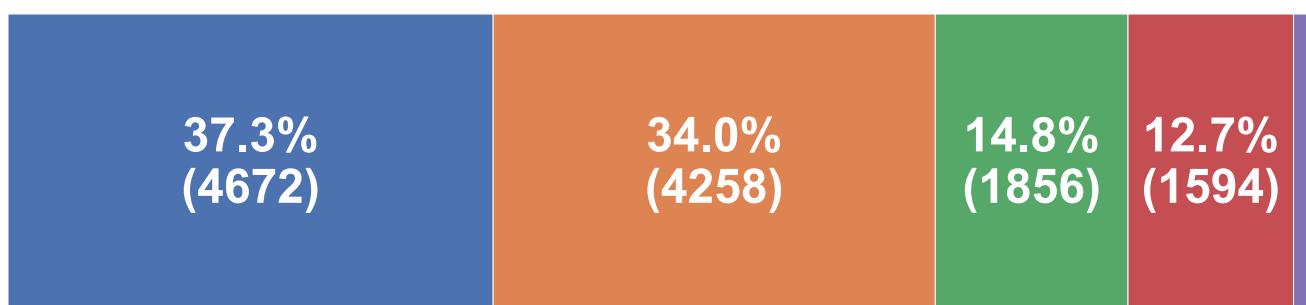
## Common Values

Value	Count	Frequency (%)
1~3	4672	36.8%
4~8	4258	33.6%
less1	1856	14.6%
gt8	1594	12.6%
never	153	1.2%
(Missing)	151	1.2%

## Length



## Common Values (Plot)



Value	Count	Frequency (%)
1~3	4672	37.3%
4~8	4258	34.0%
less1	1856	14.8%
gt8	1594	12.7%
never	153	1.2%

## Most occurring characters

Value	Count	Frequency (%)
~	8930	21.5%
1	6528	15.7%
8	5852	14.1%
3	4672	11.2%
4	4258	10.2%
s	3712	8.9%
e	2162	5.2%
l	1856	4.5%
g	1594	3.8%
t	1594	3.8%
Other values (3)	459	1.1%

## Most occurring categories

Value	Count	Frequency (%)
Decimal Number	21310	51.2%
Lowercase Letter	11377	27.3%
Math Symbol	8930	21.5%

## Most frequent character per category

*Lowercase Letter*

Value	Count	Frequency (%)
s	3712	32.6%
e	2162	19.0%
l	1856	16.3%
g	1594	14.0%
t	1594	14.0%
n	153	1.3%
v	153	1.3%
r	153	1.3%

*Decimal Number*

Value	Count	Frequency (%)
1	6528	30.6%
8	5852	27.5%
3	4672	21.9%
4	4258	20.0%

*Math Symbol*

Value	Count	Frequency (%)
~	8930	100.0%

## Most occurring scripts

Value	Count	Frequency (%)
Common	30240	72.7%
Latin	11377	27.3%

## Most frequent character per script

*Latin*

Value	Count	Frequency (%)
s	3712	32.6%
e	2162	19.0%
l	1856	16.3%
g	1594	14.0%
t	1594	14.0%
n	153	1.3%
v	153	1.3%
r	153	1.3%

*Common*

Value	Count	Frequency (%)
~	8930	29.5%
1	6528	21.6%
8	5852	19.4%
3	4672	15.4%
4	4258	14.1%

### Most occurring blocks

Value	Count	Frequency (%)
ASCII	41617	100.0%

### Most frequent character per block

ASCII

Value	Count	Frequency (%)
~	8930	21.5%
1	6528	15.7%
8	5852	14.1%
3	4672	11.2%
4	4258	10.2%
s	3712	8.9%
e	2162	5.2%
l	1856	4.5%
g	1594	3.8%
t	1594	3.8%
Other values (3)	459	1.1%





**RestaurantLessThan20**

Categorical

<b>Distinct</b>	5
<b>Distinct (%)</b>	< 0.1%
<b>Missing</b>	130
<b>Missing (%)</b>	1.0%
<b>Memory size</b>	99.2 KiB

**Length**

<b>Max length</b>	5
<b>Median length</b>	3
<b>Mean length</b>	3.3684881
<b>Min length</b>	3

**Characters and Unicode**

<b>Total characters</b>	42288
<b>Distinct characters</b>	13
<b>Distinct categories</b>	<a href="https://en.wikipedia.org/wiki/Unicode_character_property#General_Category">3 (https://en.wikipedia.org/wiki/Unicode_character_property#General_Category)</a>
<b>Distinct scripts</b>	<a href="https://en.wikipedia.org/wiki/Script_(Unicode)#List_of_scripts_in_Unicode">2 (https://en.wikipedia.org/wiki/Script_(Unicode)#List_of_scripts_in_Unicode)</a>
<b>Distinct blocks</b>	<a href="https://en.wikipedia.org/wiki/Unicode_block">1 (https://en.wikipedia.org/wiki/Unicode_block)</a>

The Unicode Standard assigns character properties to each code point, which can be used to analyse textual variables.

**Unique**

<b>Unique</b>	0	?
<b>Unique (%)</b>	0.0%	

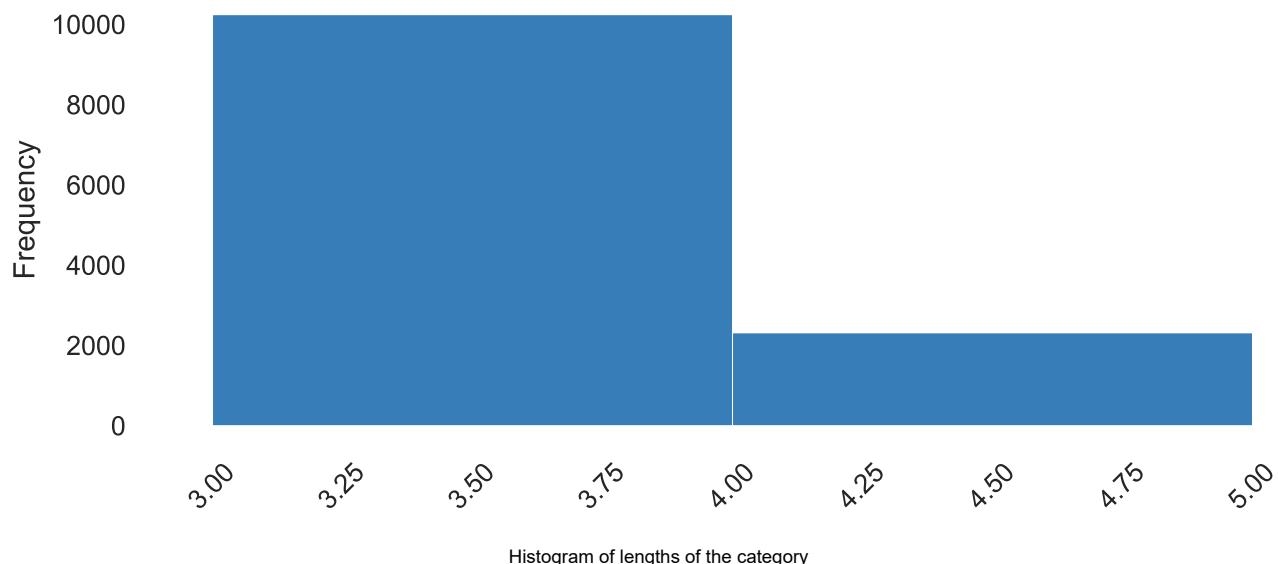
**Sample**

<b>1st row</b>	4~8
<b>2nd row</b>	4~8
<b>3rd row</b>	4~8
<b>4th row</b>	4~8
<b>5th row</b>	4~8

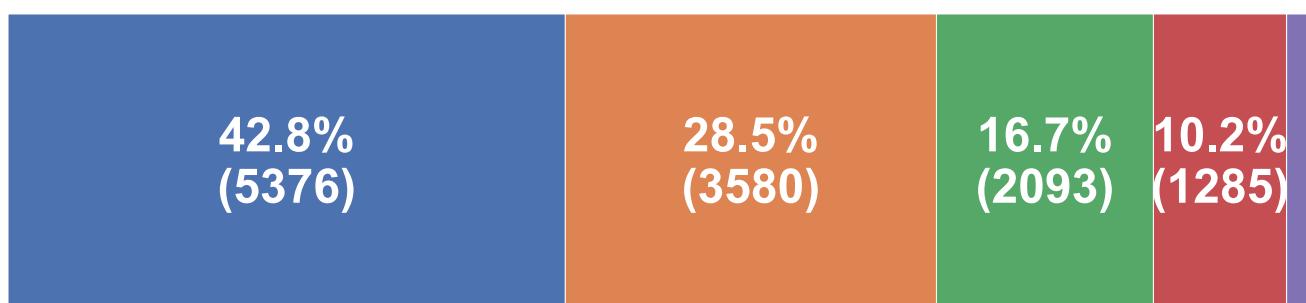
## Common Values

Value	Count	Frequency (%)
1~3	5376	42.4%
4~8	3580	28.2%
less1	2093	16.5%
gt8	1285	10.1%
never	220	1.7%
(Missing)	130	1.0%

## Length



## Common Values (Plot)



Value	Count	Frequency (%)
1~3	5376	42.8%
4~8	3580	28.5%
less1	2093	16.7%
gt8	1285	10.2%
never	220	1.8%

## Most occurring characters

Value	Count	Frequency (%)
~	8956	21.2%
1	7469	17.7%
3	5376	12.7%
8	4865	11.5%
s	4186	9.9%
4	3580	8.5%
e	2533	6.0%
l	2093	4.9%
g	1285	3.0%
t	1285	3.0%
Other values (3)	660	1.6%

## Most occurring categories

Value	Count	Frequency (%)
Decimal Number	21290	50.3%
Lowercase Letter	12042	28.5%
Math Symbol	8956	21.2%

## Most frequent character per category

*Lowercase Letter*

Value	Count	Frequency (%)
s	4186	34.8%
e	2533	21.0%
l	2093	17.4%
g	1285	10.7%
t	1285	10.7%
n	220	1.8%
v	220	1.8%
r	220	1.8%

*Decimal Number*

Value	Count	Frequency (%)
1	7469	35.1%
3	5376	25.3%
8	4865	22.9%
4	3580	16.8%

*Math Symbol*

Value	Count	Frequency (%)
~	8956	100.0%

## Most occurring scripts

Value	Count	Frequency (%)
Common	30246	71.5%
Latin	12042	28.5%

## Most frequent character per script

*Latin*

Value	Count	Frequency (%)
s	4186	34.8%
e	2533	21.0%
l	2093	17.4%
g	1285	10.7%
t	1285	10.7%
n	220	1.8%
v	220	1.8%
r	220	1.8%

*Common*

Value	Count	Frequency (%)
~	8956	29.6%
1	7469	24.7%
3	5376	17.8%
8	4865	16.1%
4	3580	11.8%

### Most occurring blocks

Value	Count	Frequency (%)
ASCII	42288	100.0%

### Most frequent character per block

ASCII

Value	Count	Frequency (%)
~	8956	21.2%
1	7469	17.7%
3	5376	12.7%
8	4865	11.5%
s	4186	9.9%
4	3580	8.5%
e	2533	6.0%
l	2093	4.9%
g	1285	3.0%
t	1285	3.0%
Other values (3)	660	1.6%





**Restaurant20To50**

Categorical

<b>Distinct</b>	5
<b>Distinct (%)</b>	< 0.1%
<b>Missing</b>	189
<b>Missing (%)</b>	1.5%
<b>Memory size</b>	99.2 KiB

**Length**

<b>Max length</b>	5
<b>Median length</b>	5
<b>Mean length</b>	4.3146058
<b>Min length</b>	3

**Characters and Unicode**

<b>Total characters</b>	53911
<b>Distinct characters</b>	13
<b>Distinct categories</b>	<a href="https://en.wikipedia.org/wiki/Unicode_character_property#General_Category">3 (https://en.wikipedia.org/wiki/Unicode_character_property#General_Category)</a>
<b>Distinct scripts</b>	<a href="https://en.wikipedia.org/wiki/Script_(Unicode)#List_of_scripts_in_Unicode">2 (https://en.wikipedia.org/wiki/Script_(Unicode)#List_of_scripts_in_Unicode)</a>
<b>Distinct blocks</b>	<a href="https://en.wikipedia.org/wiki/Unicode_block">1 (https://en.wikipedia.org/wiki/Unicode_block)</a>

The Unicode Standard assigns character properties to each code point, which can be used to analyse textual variables.

**Unique**

<b>Unique</b>	0	?
<b>Unique (%)</b>	0.0%	

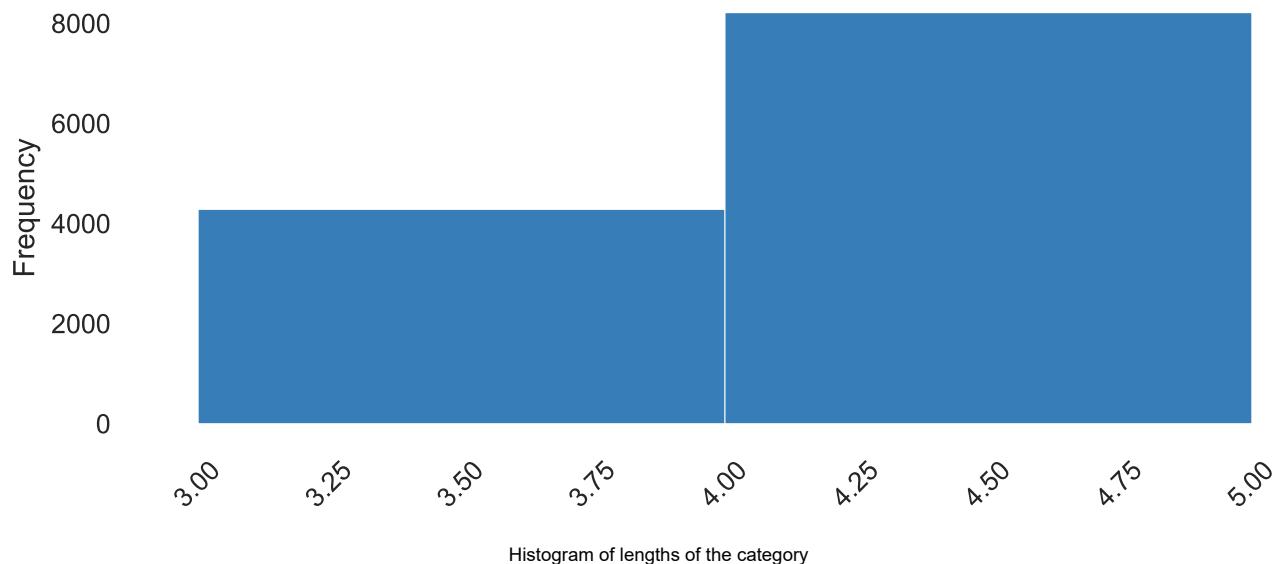
**Sample**

<b>1st row</b>	1~3
<b>2nd row</b>	1~3
<b>3rd row</b>	1~3
<b>4th row</b>	1~3
<b>5th row</b>	1~3

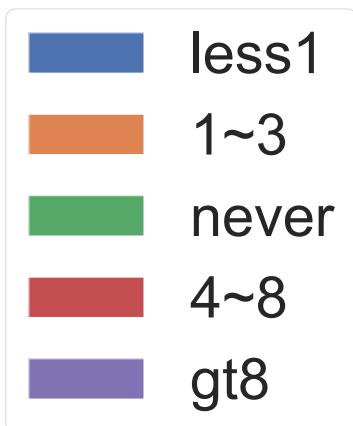
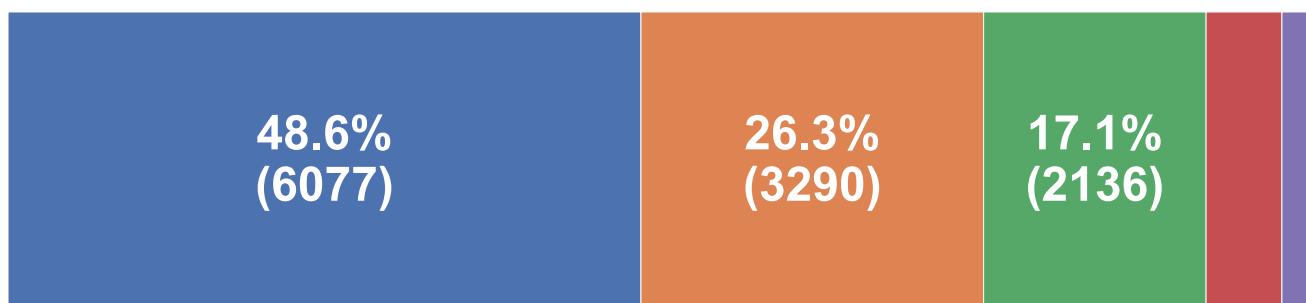
## Common Values

Value	Count	Frequency (%)
less1	6077	47.9%
1~3	3290	25.9%
never	2136	16.8%
4~8	728	5.7%
gt8	264	2.1%
(Missing)	189	1.5%

## Length



## Common Values (Plot)



Value	Count	Frequency (%)
less1	6077	48.6%
1~3	3290	26.3%
never	2136	17.1%
4~8	728	5.8%
gt8	264	2.1%

## Most occurring characters

Value	Count	Frequency (%)
s	12154	22.5%
e	10349	19.2%
1	9367	17.4%
l	6077	11.3%
~	4018	7.5%
3	3290	6.1%
n	2136	4.0%
v	2136	4.0%
r	2136	4.0%
8	992	1.8%
Other values (3)	1256	2.3%

## Most occurring categories

Value	Count	Frequency (%)
Lowercase Letter	35516	65.9%
Decimal Number	14377	26.7%
Math Symbol	4018	7.5%

## Most frequent character per category

*Lowercase Letter*

Value	Count	Frequency (%)
s	12154	34.2%
e	10349	29.1%
l	6077	17.1%
n	2136	6.0%
v	2136	6.0%
r	2136	6.0%
g	264	0.7%
t	264	0.7%

*Decimal Number*

Value	Count	Frequency (%)
1	9367	65.2%
3	3290	22.9%
8	992	6.9%
4	728	5.1%

*Math Symbol*

Value	Count	Frequency (%)
~	4018	100.0%

## Most occurring scripts

Value	Count	Frequency (%)
Latin	35516	65.9%
Common	18395	34.1%

## Most frequent character per script

*Latin*

Value	Count	Frequency (%)
s	12154	34.2%
e	10349	29.1%
l	6077	17.1%
n	2136	6.0%
v	2136	6.0%
r	2136	6.0%
g	264	0.7%
t	264	0.7%

*Common*

Value	Count	Frequency (%)
1	9367	50.9%
~	4018	21.8%
3	3290	17.9%
8	992	5.4%
4	728	4.0%

### Most occurring blocks

Value	Count	Frequency (%)
ASCII	53911	100.0%

### Most frequent character per block

ASCII

Value	Count	Frequency (%)
s	12154	22.5%
e	10349	19.2%
1	9367	17.4%
l	6077	11.3%
~	4018	7.5%
3	3290	6.1%
n	2136	4.0%
v	2136	4.0%
r	2136	4.0%
8	992	1.8%
Other values (3)	1256	2.3%





**toCoupon\_GEQ5min**

Categorical

<b>Distinct</b>	1
<b>Distinct (%)</b>	< 0.1%
<b>Missing</b>	0
<b>Missing (%)</b>	0.0%
<b>Memory size</b>	99.2 KiB

**Length**

<b>Max length</b>	1
<b>Median length</b>	1
<b>Mean length</b>	1
<b>Min length</b>	1

**Characters and Unicode**

<b>Total characters</b>	12684
<b>Distinct characters</b>	1
<b>Distinct categories</b>	1 ( <a href="https://en.wikipedia.org/wiki/Unicode_character_property#General_Category">https://en.wikipedia.org/wiki/Unicode_character_property#General_Category</a> )
<b>Distinct scripts</b>	1 ( <a href="https://en.wikipedia.org/wiki/Script_(Unicode)#List_of_scripts_in_Unicode">https://en.wikipedia.org/wiki/Script_(Unicode)#List_of_scripts_in_Unicode</a> )
<b>Distinct blocks</b>	1 ( <a href="https://en.wikipedia.org/wiki/Unicode_block">https://en.wikipedia.org/wiki/Unicode_block</a> )

The Unicode Standard assigns character properties to each code point, which can be used to analyse textual variables.

**Unique**

<b>Unique</b>	0	?
<b>Unique (%)</b>	0.0%	

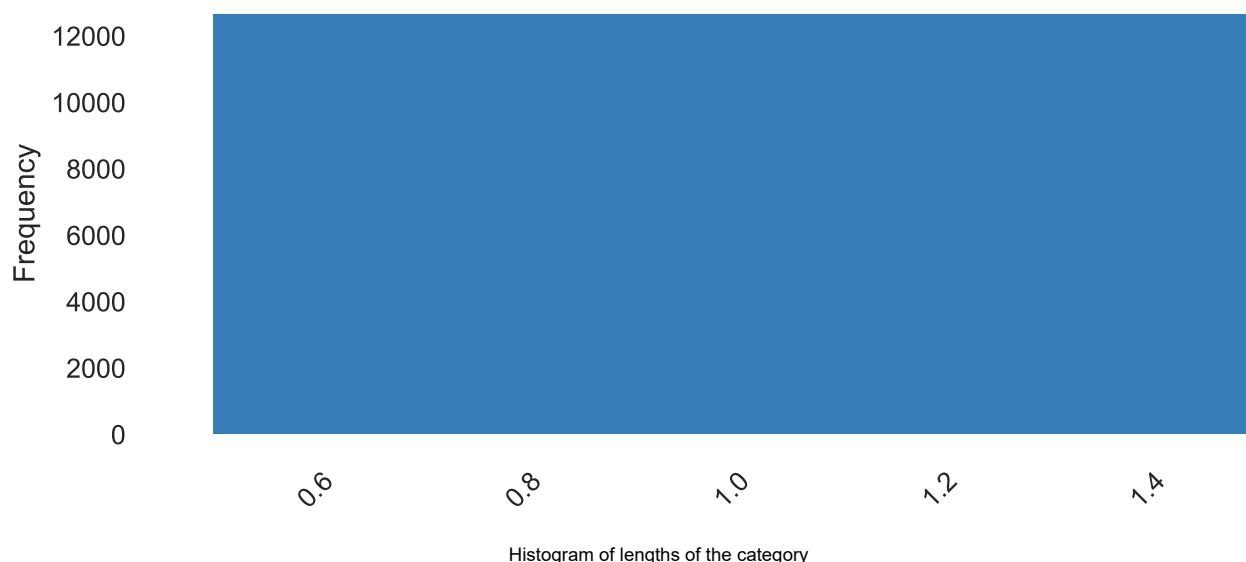
**Sample**

<b>1st row</b>	1
<b>2nd row</b>	1
<b>3rd row</b>	1
<b>4th row</b>	1
<b>5th row</b>	1

## Common Values

Value	Count	Frequency (%)
1	12684	100.0%

Length



Histogram of lengths of the category

Common Values (Plot)



Value	Count	Frequency (%)
1	12684	100.0%

## Most occurring characters

Value	Count	Frequency (%)
1	12684	100.0%

### Most occurring categories

Value	Count	Frequency (%)
Decimal Number	12684	100.0%

### Most frequent character per category

*Decimal Number*

Value	Count	Frequency (%)
1	12684	100.0%

### Most occurring scripts

Value	Count	Frequency (%)
Common	12684	100.0%

### Most frequent character per script

*Common*

Value	Count	Frequency (%)
1	12684	100.0%

### Most occurring blocks

Value	Count	Frequency (%)
ASCII	12684	100.0%

### Most frequent character per block

Value	Count	Frequency (%)
1	12684	100.0%





**toCoupon\_GEQ15min**

Categorical

<b>Distinct</b>	2
<b>Distinct (%)</b>	< 0.1%
<b>Missing</b>	0
<b>Missing (%)</b>	0.0%
<b>Memory size</b>	99.2 KiB

**Length**

<b>Max length</b>	1
<b>Median length</b>	1
<b>Mean length</b>	1
<b>Min length</b>	1

**Characters and Unicode**

<b>Total characters</b>	12684
<b>Distinct characters</b>	2
<b>Distinct categories</b>	<a href="https://en.wikipedia.org/wiki/Unicode_character_property#General_Category">1 (https://en.wikipedia.org/wiki/Unicode_character_property#General_Category)</a>
<b>Distinct scripts</b>	<a href="https://en.wikipedia.org/wiki/Script_(Unicode)#List_of_scripts_in_Unicode">1 (https://en.wikipedia.org/wiki/Script_(Unicode)#List_of_scripts_in_Unicode)</a>
<b>Distinct blocks</b>	<a href="https://en.wikipedia.org/wiki/Unicode_block">1 (https://en.wikipedia.org/wiki/Unicode_block)</a>

The Unicode Standard assigns character properties to each code point, which can be used to analyse textual variables.

**Unique**

<b>Unique</b>	0	?
<b>Unique (%)</b>	0.0%	

**Sample**

<b>1st row</b>	0
<b>2nd row</b>	0
<b>3rd row</b>	1
<b>4th row</b>	1
<b>5th row</b>	1

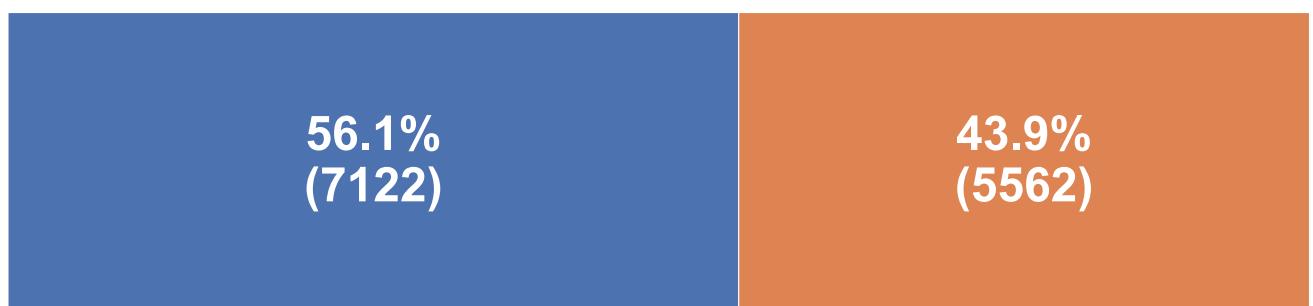
## Common Values

Value	Count	Frequency (%)
1	7122	56.1%
0	5562	43.9%

## Length



## Common Values (Plot)



Value	Count	Frequency (%)
1	7122	56.1%
0	5562	43.9%

## Most occurring characters

Value	Count	Frequency (%)
1	7122	56.1%
0	5562	43.9%

### Most occurring categories

Value	Count	Frequency (%)
Decimal Number	12684	100.0%

### Most frequent character per category

*Decimal Number*

Value	Count	Frequency (%)
1	7122	56.1%
0	5562	43.9%

### Most occurring scripts

Value	Count	Frequency (%)
Common	12684	100.0%

### Most frequent character per script

*Common*

Value	Count	Frequency (%)
1	7122	56.1%
0	5562	43.9%

### Most occurring blocks

Value	Count	Frequency (%)
ASCII	12684	100.0%

### Most frequent character per block

ASCII

Value	Count	Frequency (%)
1	7122	56.1%
0	5562	43.9%





**toCoupon\_GEQ25min**

Categorical

<b>Distinct</b>	2
<b>Distinct (%)</b>	< 0.1%
<b>Missing</b>	0
<b>Missing (%)</b>	0.0%
<b>Memory size</b>	99.2 KiB

**Length**

<b>Max length</b>	1
<b>Median length</b>	1
<b>Mean length</b>	1
<b>Min length</b>	1

**Characters and Unicode**

<b>Total characters</b>	12684
<b>Distinct characters</b>	2
<b>Distinct categories</b>	1 ( <a href="https://en.wikipedia.org/wiki/Unicode_character_property#General_Category">https://en.wikipedia.org/wiki/Unicode_character_property#General_Category</a> )
<b>Distinct scripts</b>	1 ( <a href="https://en.wikipedia.org/wiki/Script_(Unicode)#List_of_scripts_in_Unicode">https://en.wikipedia.org/wiki/Script_(Unicode)#List_of_scripts_in_Unicode</a> )
<b>Distinct blocks</b>	1 ( <a href="https://en.wikipedia.org/wiki/Unicode_block">https://en.wikipedia.org/wiki/Unicode_block</a> )

The Unicode Standard assigns character properties to each code point, which can be used to analyse textual variables.

**Unique**

<b>Unique</b>	0	?
<b>Unique (%)</b>	0.0%	

**Sample**

<b>1st row</b>	0
<b>2nd row</b>	0
<b>3rd row</b>	0
<b>4th row</b>	0
<b>5th row</b>	0

## Common Values

Value	Count	Frequency (%)
0	11173	88.1%
1	1511	11.9%

## Length



## Common Values (Plot)



Value	Count	Frequency (%)
0	11173	88.1%
1	1511	11.9%

## Most occurring characters

Value	Count	Frequency (%)
0	11173	88.1%
1	1511	11.9%

### Most occurring categories

Value	Count	Frequency (%)
Decimal Number	12684	100.0%

### Most frequent character per category

*Decimal Number*

Value	Count	Frequency (%)
0	11173	88.1%
1	1511	11.9%

### Most occurring scripts

Value	Count	Frequency (%)
Common	12684	100.0%

### Most frequent character per script

*Common*

Value	Count	Frequency (%)
0	11173	88.1%
1	1511	11.9%

### Most occurring blocks

Value	Count	Frequency (%)
ASCII	12684	100.0%

### Most frequent character per block

ASCII

Value	Count	Frequency (%)
0	11173	88.1%
1	1511	11.9%





**direction\_same**

Categorical

<b>Distinct</b>	2
<b>Distinct (%)</b>	< 0.1%
<b>Missing</b>	0
<b>Missing (%)</b>	0.0%
<b>Memory size</b>	99.2 KiB

**Length**

<b>Max length</b>	1
<b>Median length</b>	1
<b>Mean length</b>	1
<b>Min length</b>	1

**Characters and Unicode**

<b>Total characters</b>	12684
<b>Distinct characters</b>	2
<b>Distinct categories</b>	1 ( <a href="https://en.wikipedia.org/wiki/Unicode_character_property#General_Category">https://en.wikipedia.org/wiki/Unicode_character_property#General_Category</a> )
<b>Distinct scripts</b>	1 ( <a href="https://en.wikipedia.org/wiki/Script_(Unicode)#List_of_scripts_in_Unicode">https://en.wikipedia.org/wiki/Script_(Unicode)#List_of_scripts_in_Unicode</a> )
<b>Distinct blocks</b>	1 ( <a href="https://en.wikipedia.org/wiki/Unicode_block">https://en.wikipedia.org/wiki/Unicode_block</a> )

The Unicode Standard assigns character properties to each code point, which can be used to analyse textual variables.

**Unique**

<b>Unique</b>	0	?
<b>Unique (%)</b>	0.0%	

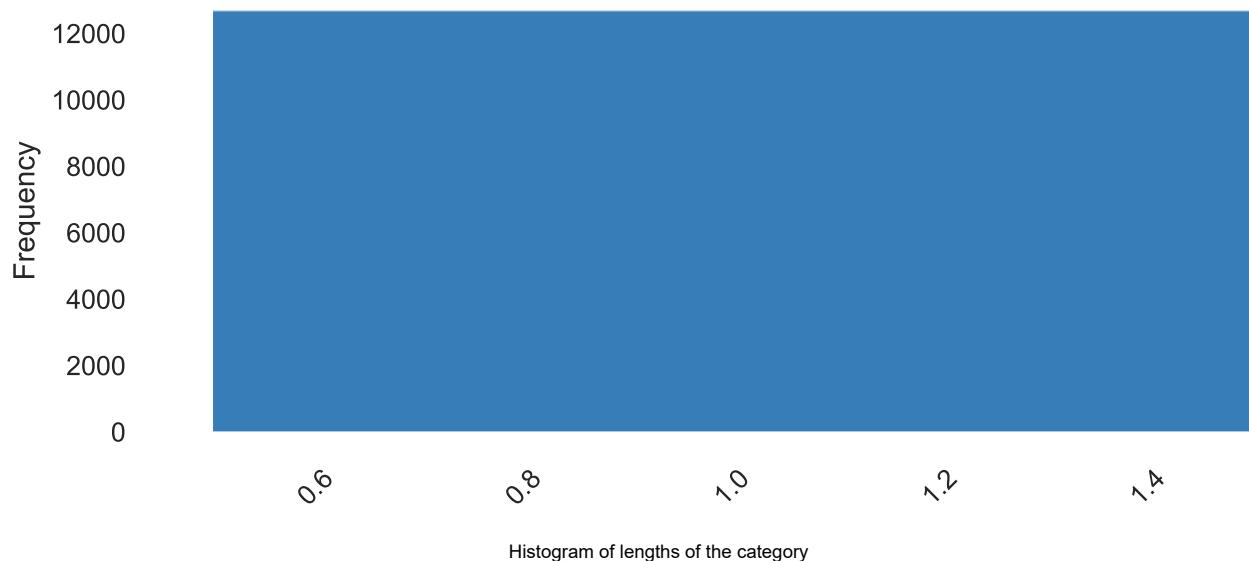
**Sample**

<b>1st row</b>	0
<b>2nd row</b>	0
<b>3rd row</b>	0
<b>4th row</b>	0
<b>5th row</b>	0

## Common Values

Value	Count	Frequency (%)
0	9960	78.5%
1	2724	21.5%

## Length



## Common Values (Plot)



Value	Count	Frequency (%)
0	9960	78.5%
1	2724	21.5%

## Most occurring characters

Value	Count	Frequency (%)
0	9960	78.5%
1	2724	21.5%

### Most occurring categories

Value	Count	Frequency (%)
Decimal Number	12684	100.0%

### Most frequent character per category

*Decimal Number*

Value	Count	Frequency (%)
0	9960	78.5%
1	2724	21.5%

### Most occurring scripts

Value	Count	Frequency (%)
Common	12684	100.0%

### Most frequent character per script

*Common*

Value	Count	Frequency (%)
0	9960	78.5%
1	2724	21.5%

### Most occurring blocks

Value	Count	Frequency (%)
ASCII	12684	100.0%

### Most frequent character per block

ASCII

Value	Count	Frequency (%)
0	9960	78.5%
1	2724	21.5%





**direction\_opp**

Categorical

<b>Distinct</b>	2
<b>Distinct (%)</b>	< 0.1%
<b>Missing</b>	0
<b>Missing (%)</b>	0.0%
<b>Memory size</b>	99.2 KiB

**Length**

<b>Max length</b>	1
<b>Median length</b>	1
<b>Mean length</b>	1
<b>Min length</b>	1

**Characters and Unicode**

<b>Total characters</b>	12684
<b>Distinct characters</b>	2
<b>Distinct categories</b>	1 ( <a href="https://en.wikipedia.org/wiki/Unicode_character_property#General_Category">https://en.wikipedia.org/wiki/Unicode_character_property#General_Category</a> )
<b>Distinct scripts</b>	1 ( <a href="https://en.wikipedia.org/wiki/Script_(Unicode)#List_of_scripts_in_Unicode">https://en.wikipedia.org/wiki/Script_(Unicode)#List_of_scripts_in_Unicode</a> )
<b>Distinct blocks</b>	1 ( <a href="https://en.wikipedia.org/wiki/Unicode_block">https://en.wikipedia.org/wiki/Unicode_block</a> )

The Unicode Standard assigns character properties to each code point, which can be used to analyse textual variables.

**Unique**

<b>Unique</b>	0	?
<b>Unique (%)</b>	0.0%	

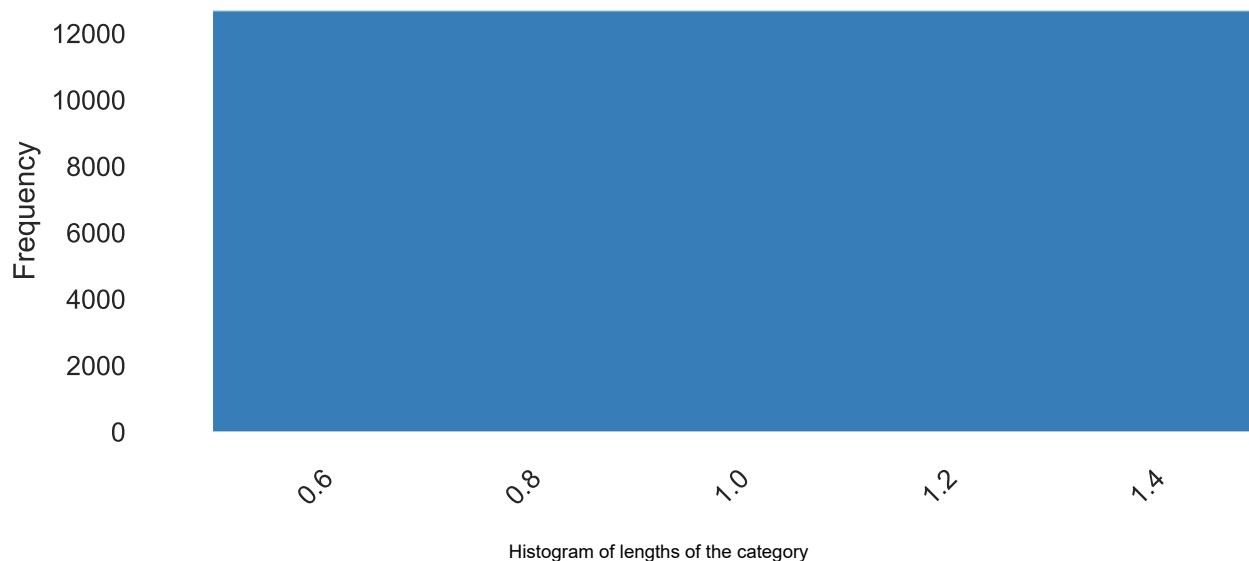
**Sample**

<b>1st row</b>	1
<b>2nd row</b>	1
<b>3rd row</b>	1
<b>4th row</b>	1
<b>5th row</b>	1

## Common Values

Value	Count	Frequency (%)
1	9960	78.5%
0	2724	21.5%

## Length



## Common Values (Plot)



Value	Count	Frequency (%)
1	9960	78.5%
0	2724	21.5%

## Most occurring characters

Value	Count	Frequency (%)
1	9960	78.5%
0	2724	21.5%

### Most occurring categories

Value	Count	Frequency (%)
Decimal Number	12684	100.0%

### Most frequent character per category

*Decimal Number*

Value	Count	Frequency (%)
1	9960	78.5%
0	2724	21.5%

### Most occurring scripts

Value	Count	Frequency (%)
Common	12684	100.0%

### Most frequent character per script

*Common*

Value	Count	Frequency (%)
1	9960	78.5%
0	2724	21.5%

### Most occurring blocks

Value	Count	Frequency (%)
ASCII	12684	100.0%

### Most frequent character per block

ASCII

Value	Count	Frequency (%)
1	9960	78.5%
0	2724	21.5%





Y  
Categorical

<b>Distinct</b>	2
<b>Distinct (%)</b>	< 0.1%
<b>Missing</b>	0
<b>Missing (%)</b>	0.0%
<b>Memory size</b>	99.2 KiB

Length

<b>Max length</b>	1
<b>Median length</b>	1
<b>Mean length</b>	1
<b>Min length</b>	1

Characters and Unicode

<b>Total characters</b>	12684
<b>Distinct characters</b>	2
<b>Distinct categories</b>	1 ( <a href="https://en.wikipedia.org/wiki/Unicode_character_property#General_Category">https://en.wikipedia.org/wiki/Unicode_character_property#General_Category</a> )
<b>Distinct scripts</b>	1 ( <a href="https://en.wikipedia.org/wiki/Script_(Unicode)#List_of_scripts_in_Unicode">https://en.wikipedia.org/wiki/Script_(Unicode)#List_of_scripts_in_Unicode</a> )
<b>Distinct blocks</b>	1 ( <a href="https://en.wikipedia.org/wiki/Unicode_block">https://en.wikipedia.org/wiki/Unicode_block</a> )

The Unicode Standard assigns character properties to each code point, which can be used to analyse textual variables.

Unique

<b>Unique</b>	0	?
<b>Unique (%)</b>	0.0%	

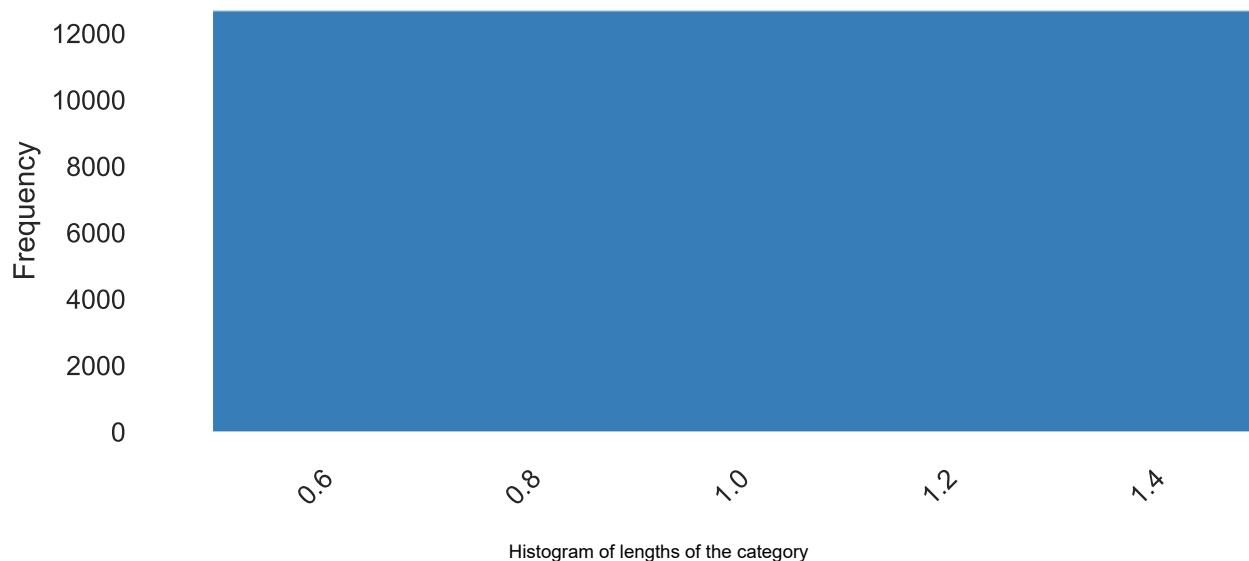
Sample

<b>1st row</b>	1
<b>2nd row</b>	0
<b>3rd row</b>	1
<b>4th row</b>	0
<b>5th row</b>	0

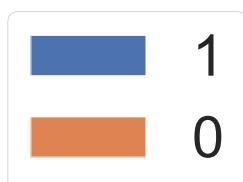
## Common Values

Value	Count	Frequency (%)
1	7210	56.8%
0	5474	43.2%

## Length



## Common Values (Plot)



Value	Count	Frequency (%)
1	7210	56.8%
0	5474	43.2%

## Most occurring characters

Value	Count	Frequency (%)
1	7210	56.8%
0	5474	43.2%

### Most occurring categories

Value	Count	Frequency (%)
Decimal Number	12684	100.0%

### Most frequent character per category

*Decimal Number*

Value	Count	Frequency (%)
1	7210	56.8%
0	5474	43.2%

### Most occurring scripts

Value	Count	Frequency (%)
Common	12684	100.0%

### Most frequent character per script

*Common*

Value	Count	Frequency (%)
1	7210	56.8%
0	5474	43.2%

### Most occurring blocks

Value	Count	Frequency (%)
ASCII	12684	100.0%

### Most frequent character per block

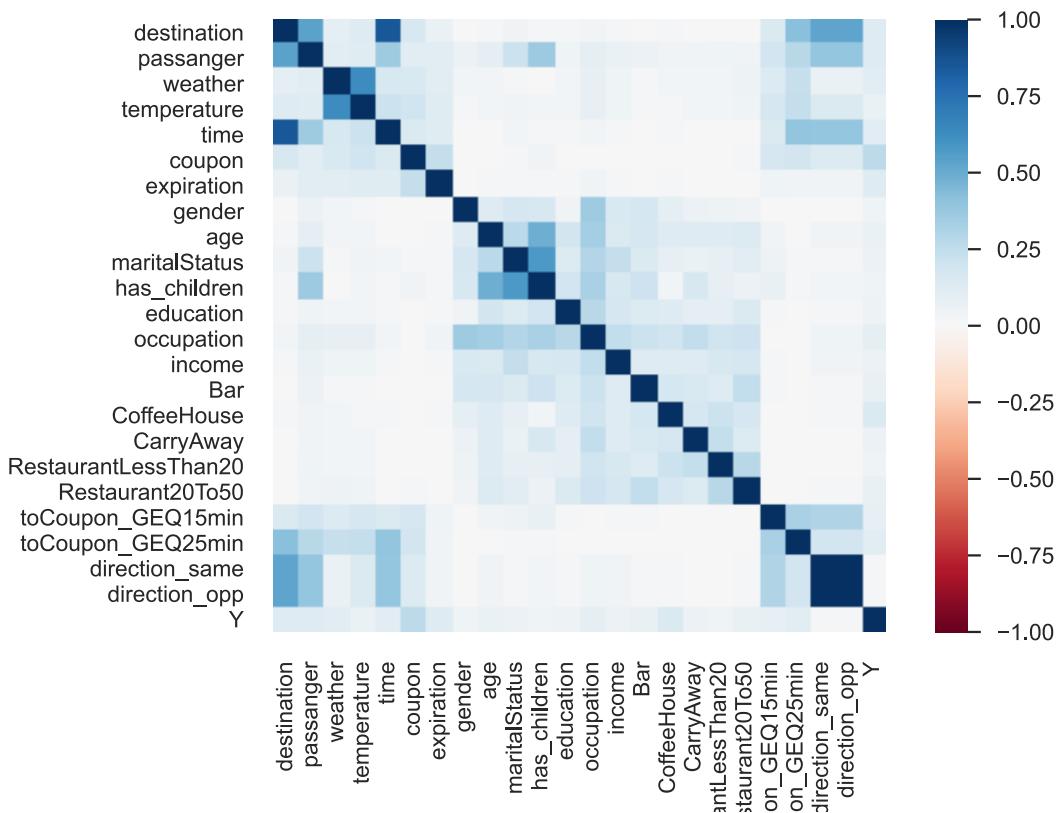
ASCII

Value	Count	Frequency (%)
1	7210	56.8%
0	5474	43.2%





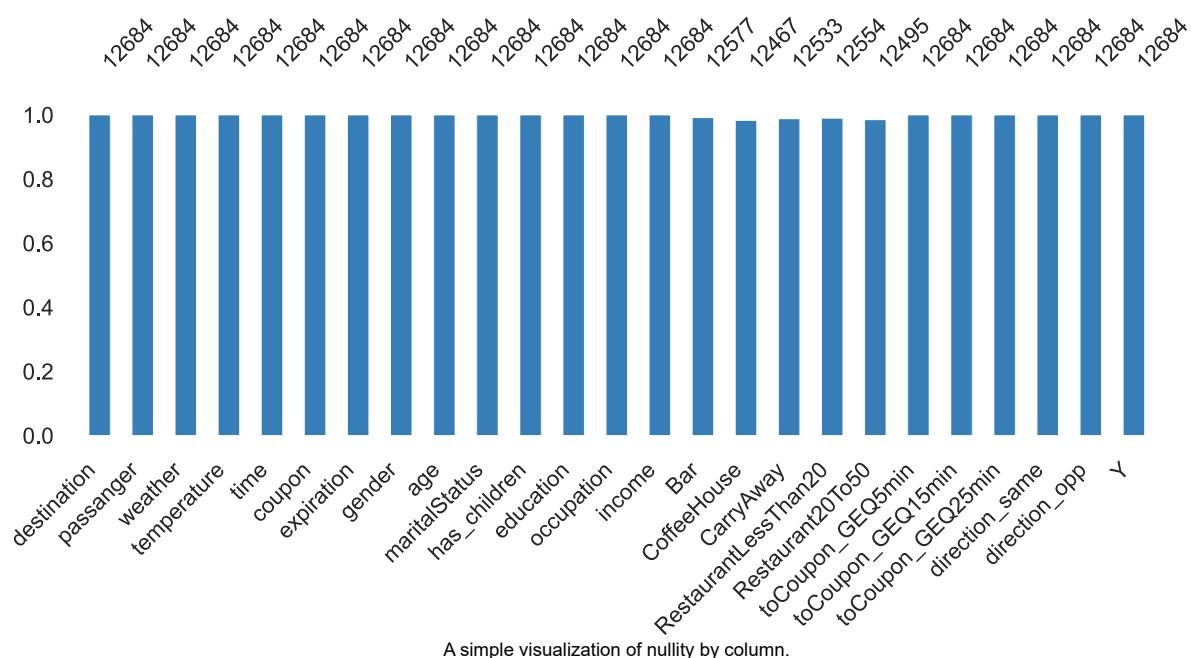
# Correlations



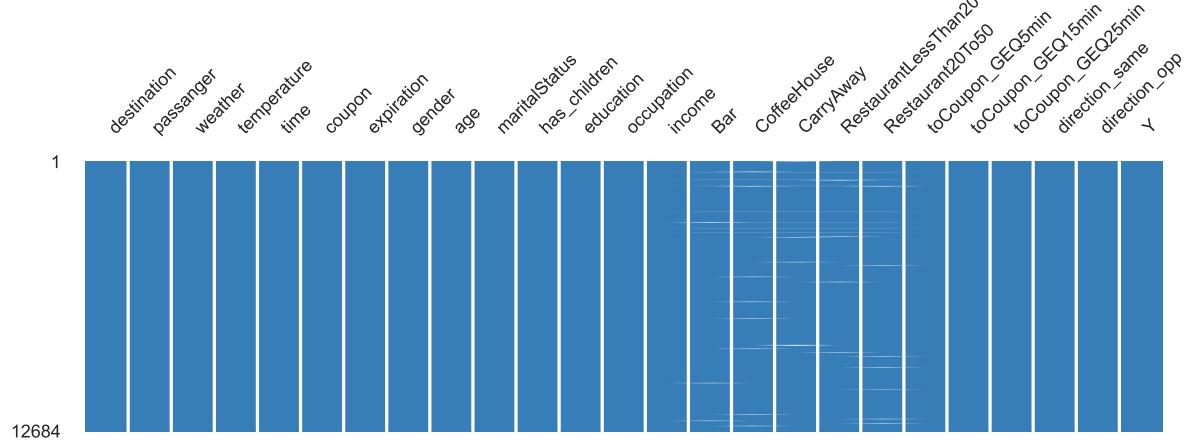
	destination	passanger	weather	temperature	time	coupon	expiration	gender	age	maritalStatus	I
destination	1.000	0.538	0.092	0.126	0.845	0.157	0.067	0.000	0.012	0.036	0
passanger	0.538	1.000	0.113	0.125	0.353	0.116	0.114	0.056	0.096	0.216	0
weather	0.092	0.113	1.000	0.626	0.158	0.155	0.114	0.027	0.029	0.017	0
temperature	0.126	0.125	0.626	1.000	0.213	0.199	0.125	0.023	0.028	0.037	0
time	0.845	0.353	0.158	0.213	1.000	0.146	0.124	0.000	0.006	0.029	0
coupon	0.157	0.116	0.155	0.199	0.146	1.000	0.236	0.000	0.000	0.020	0
expiration	0.067	0.114	0.114	0.125	0.124	0.236	1.000	0.000	0.020	0.023	0
gender	0.000	0.056	0.027	0.023	0.000	0.000	0.000	1.000	0.130	0.171	0
age	0.012	0.096	0.029	0.028	0.006	0.000	0.020	0.130	1.000	0.273	0



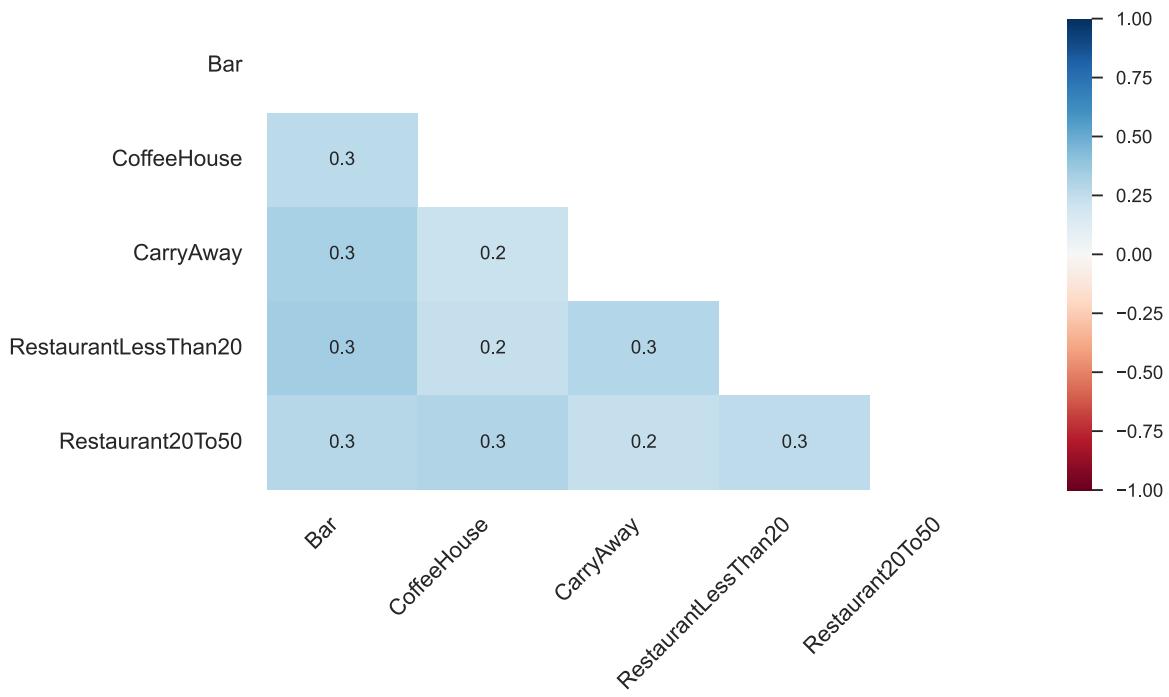
# Missing values



A simple visualization of nullity by column.



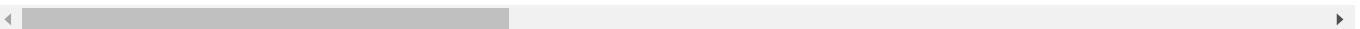
Nullity matrix is a data-dense display which lets you quickly visually pick out patterns in data completion.



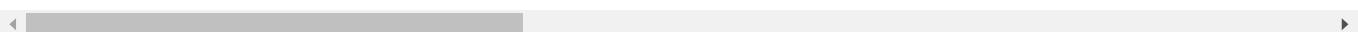
The correlation heatmap measures nullity correlation: how strongly the presence or absence of one variable affects the presence of another.

# Sample

	destination	passanger	weather	temperature	time	coupon	expiration	gender	age	maritalStatus	has_child
0	No Urgent Place	Alone	Sunny	55	2PM	Restaurant(<20)	1d	Female	21	Unmarried partner	1
1	No Urgent Place	Friend(s)	Sunny	80	10AM	Coffee House	2h	Female	21	Unmarried partner	1
2	No Urgent Place	Friend(s)	Sunny	80	10AM	Carry out & Take away	2h	Female	21	Unmarried partner	1
3	No Urgent Place	Friend(s)	Sunny	80	2PM	Coffee House	2h	Female	21	Unmarried partner	1
4	No Urgent Place	Friend(s)	Sunny	80	2PM	Coffee House	1d	Female	21	Unmarried partner	1
5	No Urgent Place	Friend(s)	Sunny	80	6PM	Restaurant(<20)	2h	Female	21	Unmarried partner	1
6	No Urgent Place	Friend(s)	Sunny	55	2PM	Carry out & Take away	1d	Female	21	Unmarried partner	1
7	No Urgent Place	Kid(s)	Sunny	80	10AM	Restaurant(<20)	2h	Female	21	Unmarried partner	1
8	No Urgent Place	Kid(s)	Sunny	80	10AM	Carry out & Take away	2h	Female	21	Unmarried partner	1
9	No Urgent Place	Kid(s)	Sunny	80	10AM	Bar	1d	Female	21	Unmarried partner	1



	destination	passanger	weather	temperature	time	coupon	expiration	gender	age	maritalStatus	has_children
12674	Home	Alone	Rainy	55	10PM	Coffee House	2h	Male	26	Single	0
12675	Home	Alone	Snowy	30	10PM	Coffee House	2h	Male	26	Single	0
12676	Home	Alone	Sunny	80	6PM	Restaurant(20-50)	1d	Male	26	Single	0
12677	Home	Partner	Sunny	30	6PM	Restaurant(<20)	1d	Male	26	Single	0
12678	Home	Partner	Sunny	30	10PM	Restaurant(<20)	2h	Male	26	Single	0
12679	Home	Partner	Rainy	55	6PM	Carry out & Take away	1d	Male	26	Single	0
12680	Work	Alone	Rainy	55	7AM	Carry out & Take away	1d	Male	26	Single	0
12681	Work	Alone	Snowy	30	7AM	Coffee House	1d	Male	26	Single	0
12682	Work	Alone	Snowy	30	7AM	Bar	1d	Male	26	Single	0
12683	Work	Alone	Sunny	80	7AM	Restaurant(20-50)	2h	Male	26	Single	0



# Duplicate rows

Most frequently occurring

	destination	passanger	weather	temperature	time	coupon	expiration	gender	age	maritalStatus	has_children	e
0	Home	Alone	Snowy	30	10PM	Restaurant(20-50)	2h	Female	31	Married partner	0	S
1	Home	Alone	Snowy	30	6PM	Coffee House	1d	Male	46	Married partner	1	C
2	Home	Alone	Sunny	30	6PM	Carry out & Take away	2h	Male	46	Married partner	1	G
3	Home	Alone	Sunny	80	6PM	Bar	2h	Female	31	Married partner	0	S
4	Home	Alone	Sunny	80	6PM	Bar	2h	Male	46	Married partner	1	G
5	Home	Partner	Sunny	30	10PM	Carry out & Take away	2h	Male	46	Married partner	1	C
6	Home	Partner	Sunny	80	6PM	Coffee House	1d	Female	31	Married partner	0	S
7	No Urgent Place	Alone	Sunny	55	6PM	Coffee House	2h	Male	46	Married partner	1	C
8	No Urgent Place	Alone	Sunny	80	10AM	Coffee House	1d	Female	31	Married partner	0	S
9	No Urgent Place	Friend(s)	Rainy	55	10PM	Bar	2h	Female	31	Married partner	0	S

Report generated by YData ([https://ydata.ai/?utm\\_source=opensource&utm\\_medium=pandasprofiling&utm\\_campaign=report](https://ydata.ai/?utm_source=opensource&utm_medium=pandasprofiling&utm_campaign=report)).