

Overview

Dataset statistics

Number of variables	14
Number of observations	89121
Missing cells	753
Missing cells (%)	0.1%
Duplicate rows	0
Duplicate rows (%)	0.0%
Total size in memory	12.2 MiB
Average record size in memory	143.7 B

Variable types

Categorical	11
DateTime	2
Numeric	1

Alerts

agent_login has a high cardinality: 1961 distinct values	High cardinality
finish_dttm.1 has a high cardinality: 48358 distinct values	High cardinality
queue_desc has a high cardinality: 155 distinct values	High cardinality
group_nm has a high cardinality: 52 distinct values	High cardinality
agent_login.1 has a high cardinality: 1961 distinct values	High cardinality
source_system_cd is highly overall correlated with group_nm and 1 other fields (group_nm, management_nm)	High correlation
using_flg is highly overall correlated with hit_status_result_desc	High correlation
hit_status_result_desc is highly overall correlated with using_flg	High correlation
group_nm is highly overall correlated with source_system_cd and 1 other fields (source_system_cd, management_nm)	High correlation
management_nm is highly overall correlated with source_system_cd and 1 other fields (source_system_cd, group_nm)	High correlation
using_flg is highly imbalanced (87.5%)	Imbalance
finish_dttm.1 is uniformly distributed	Uniform
duration_sec has 50630 (56.8%) zeros	Zeros

Reproduction

Analysis started	2024-04-07 15:44:21.431487
Analysis finished	2024-04-07 15:44:30.164615
Duration	8.73 seconds
Software version	pandas-profiling v3.6.6 (https://github.com/pandas-profiling/pandas-profiling)
Download configuration	config.json (data:text/plain;charset=utf-8,%7B%22title%22%3A%20%22df%22%2C%20%22dataset%22%3A%20%7B%22description%22%3A%20%22%22%2C%20%22c

Variables

Select Columns ▼

agent_login
Categorical

Distinct	1961
Distinct (%)	2.2%

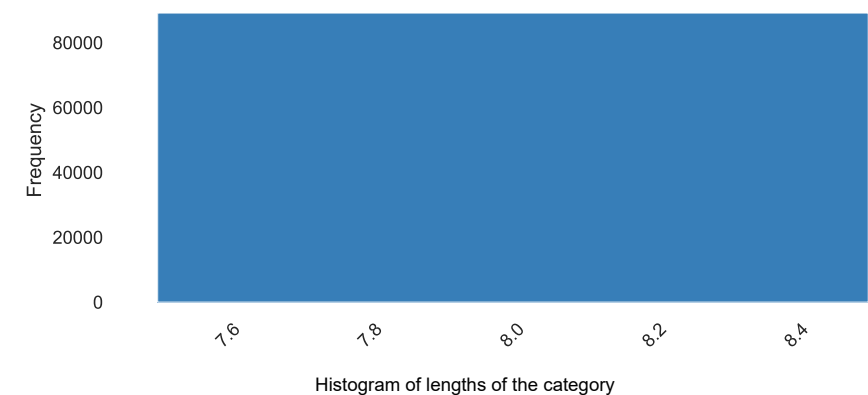
Missing	0
Missing (%)	0.0%
Memory size	1.4 MiB

Length		Characters and Unicode	Unique	Sample
Max length	8		Unique 40 ?	1st row alovmma
Median length	8	Total characters 712968	Unique (%) < 0.1%	2nd row 6446kmmi
Mean length	8	Distinct characters 35		3rd row zovaiaoz
Min length	8	Distinct categories 3 (https://en.wikipedia.org/wiki/Unicode_character_property#General_Category)		4th row 7787nyua
		Distinct scripts 2 (https://en.wikipedia.org/wiki/Script_(Unicode)#List_of_scripts_in_Unicode)		5th row kovadsku ?
		Distinct blocks 1 (https://en.wikipedia.org/wiki/Unicode_block)		?
		The Unicode Standard assigns character properties to each code point, which can be used to analyse textual variables.		

Common Values

Value	Count	Frequency (%)
covaadru	184	0.2%
enkovaab	159	0.2%
levaadko	156	0.2%
kinalapo	152	0.2%
8857iaro	151	0.2%
nayaivza	149	0.2%
rovaeagu	148	0.2%
0489vvvl	146	0.2%
novaadtr	143	0.2%
novannxo	143	0.2%
Other values (1951)	87590	98.3%

Length



Value	Count	Frequency (%)
covaadru	184	0.2%
enkovaab	159	0.2%
levaadko	156	0.2%
kinalapo	152	0.2%
8857iaro	151	0.2%

Value	Count	Frequency (%)
nayaivza	149	0.2%
rovaeagu	148	0.2%
0489vvvl	146	0.2%
novaadtr	143	0.2%
novannxo	143	0.2%
Other values (1951)	87590	98.3%

Most occurring characters

Value	Count	Frequency (%)
a	120983	17.0%
v	57275	8.0%
o	47582	6.7%
n	44696	6.3%
e	37203	5.2%
k	32807	4.6%
s	30318	4.3%
i	25419	3.6%
y	22256	3.1%
m	22167	3.1%
Other values (25)	272262	38.2%

Most occurring categories

Value	Count	Frequency (%)
Lowercase Letter	583407	81.8%
Decimal Number	129439	18.2%
Other Punctuation	122	< 0.1%

Most frequent character per category

Lowercase Letter

Value	Count	Frequency (%)
a	120983	20.7%
v	57275	9.8%
o	47582	8.2%
n	44696	7.7%
e	37203	6.4%
k	32807	5.6%
s	30318	5.2%
i	25419	4.4%
y	22256	3.8%
m	22167	3.8%
Other values (14)	142701	24.5%

Decimal Number

Value	Count	Frequency (%)
8	13752	10.6%
6	13601	10.5%
5	13541	10.5%
1	13407	10.4%
0	13258	10.2%
3	13106	10.1%
7	12499	9.7%
4	12492	9.7%
9	12237	9.5%
2	11546	8.9%

Other Punctuation

Value	Count	Frequency (%)
.	122	100.0%

Most occurring scripts

Value	Count	Frequency (%)
Latin	583407	81.8%
Common	129561	18.2%

Most frequent character per script

Latin

Value	Count	Frequency (%)
a	120983	20.7%
v	57275	9.8%
o	47582	8.2%
n	44696	7.7%
e	37203	6.4%
k	32807	5.6%
s	30318	5.2%
i	25419	4.4%
y	22256	3.8%
m	22167	3.8%
Other values (14)	142701	24.5%

Common

Value	Count	Frequency (%)
8	13752	10.6%
6	13601	10.5%
5	13541	10.5%
1	13407	10.3%
0	13258	10.2%
3	13106	10.1%
7	12499	9.6%

Value	Count	Frequency (%)
4	12492	9.6%
9	12237	9.4%
2	11546	8.9%

Most occurring blocks

Value	Count	Frequency (%)
ASCII	712968	100.0%

Most frequent character per block

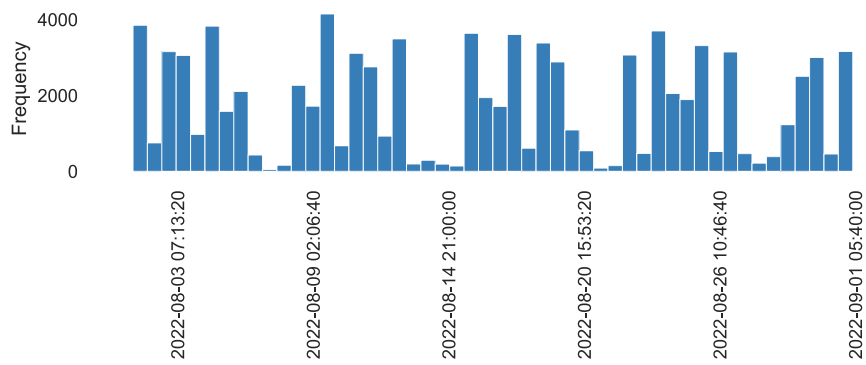
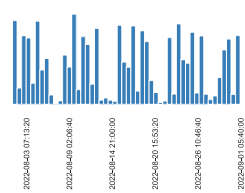
ASCII

Value	Count	Frequency (%)
a	120983	17.0%
v	57275	8.0%
o	47582	6.7%
n	44696	6.3%
e	37203	5.2%
k	32807	4.6%
s	30318	4.3%
i	25419	3.6%
y	22256	3.1%
m	22167	3.1%
Other values (25)	272262	38.2%

finish_dttm
Date

Distinct	84984
Distinct (%)	95.4%
Missing	0
Missing (%)	0.0%
Memory size	1.4 MiB

Minimum	2022-08-01 04:36:31
Maximum	2022-08-31 22:01:45

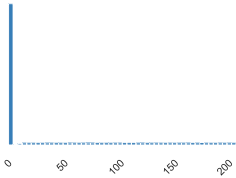


Histogram with fixed size bins (bins=50)

duration_sec
Real number (R)

Distinct	17177
Distinct (%)	19.3%
Missing	0
Missing (%)	0.0%
Infinite	0
Infinite (%)	0.0%
Mean	47.563788

Minimum	0
Maximum	210
Zeros	50630
Zeros (%)	56.8%
Negative	0
Negative (%)	0.0%
Memory size	1.4 MiB

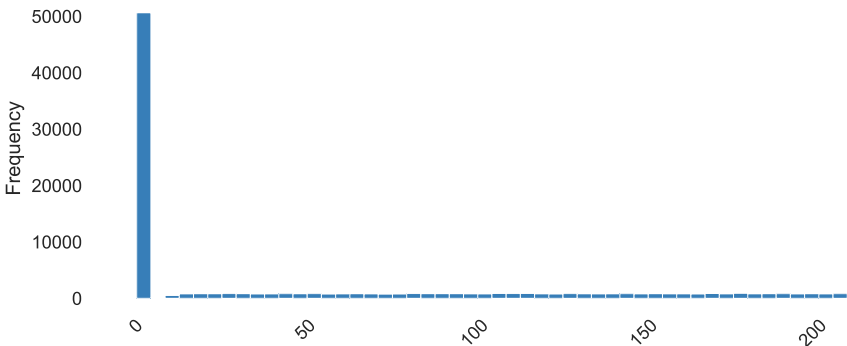


Quantile statistics

Minimum	0
5-th percentile	0
Q1	0
median	0
Q3	94.56
95-th percentile	186.81
Maximum	210
Range	210
Interquartile range (IQR)	94.56

Descriptive statistics

Standard deviation	66.463415
Coefficient of variation (CV)	1.3973533
Kurtosis	-0.33891705
Mean	47.563788
Median Absolute Deviation (MAD)	0
Skewness	1.071968
Sum	4238932.4
Variance	4417.3855
Monotonicity	Not monotonic



Histogram with fixed size bins (bins=50)

Value	Count	Frequency (%)
0	50630	56.8%
68.18	9	< 0.1%
208.28	9	< 0.1%
159.28	9	< 0.1%
95.33	8	< 0.1%
61.44	8	< 0.1%
153.91	8	< 0.1%
32.22	8	< 0.1%
28.96	8	< 0.1%
171.62	8	< 0.1%
Other values (17167)	38416	43.1%

Value	Count	Frequency (%)
0	50630	56.8%
3.57	1	< 0.1%

Value	Count	Frequency (%)
3.89	1	< 0.1%
4.4	1	< 0.1%
4.47	1	< 0.1%
4.99	1	< 0.1%
5.02	1	< 0.1%
6	1	< 0.1%
6.03	1	< 0.1%
6.19	1	< 0.1%

Value	Count	Frequency (%)
210	2	< 0.1%
209.99	3	< 0.1%
209.97	5	< 0.1%
209.96	2	< 0.1%
209.95	2	< 0.1%
209.94	3	< 0.1%
209.93	2	< 0.1%
209.91	1	< 0.1%
209.89	1	< 0.1%
209.88	3	< 0.1%

source_system_cd
Categorical

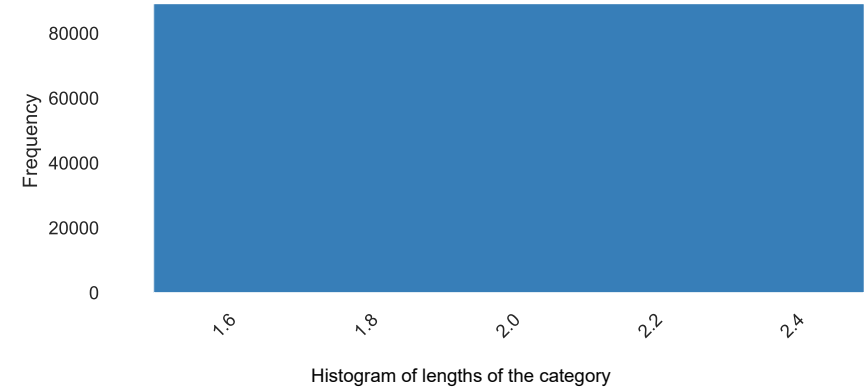
Distinct	3
Distinct (%)	< 0.1%
Missing	0
Missing (%)	0.0%
Memory size	1.4 MiB

Length		Characters and Unicode		Unique		Sample	
Max length	2	Total characters	178242	Unique	0 ?	1st row	GI
Median length	2			Unique (%)	0.0%	2nd row	GI
Mean length	2	Distinct characters	6			3rd row	GI
Min length	2					4th row	GI
		Distinct categories	1 (https://en.wikipedia.org/wiki/Unicode_character_property#General_Category)			5th row	GI ?
		Distinct scripts	1 (https://en.wikipedia.org/wiki/Script_(Unicode)#List_of_scripts_in_Unicode)				
		Distinct blocks	1 (https://en.wikipedia.org/wiki/Unicode_block)				
<p>The Unicode Standard assigns character properties to each code point, which can be used to analyse textual variables.</p>							

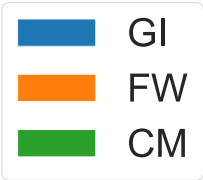
Common Values

Value	Count	Frequency (%)
GI	73333	82.3%
FW	11997	13.5%
CM	3791	4.3%

Length



Common Values (Plot)



Value	Count	Frequency (%)
gi	73333	82.3%
fw	11997	13.5%
cm	3791	4.3%

Most occurring characters

Value	Count	Frequency (%)
G	73333	41.1%
I	73333	41.1%
F	11997	6.7%
W	11997	6.7%
C	3791	2.1%
M	3791	2.1%

Most occurring categories

Value	Count	Frequency (%)
Uppercase Letter	178242	100.0%

Most frequent character per category

Uppercase Letter

Value	Count	Frequency (%)
G	73333	41.1%
I	73333	41.1%
F	11997	6.7%
W	11997	6.7%
C	3791	2.1%
M	3791	2.1%

Most occurring scripts

Value	Count	Frequency (%)
Latin	178242	100.0%

Most frequent character per script

Latin

Value	Count	Frequency (%)
G	73333	41.1%
I	73333	41.1%
F	11997	6.7%
W	11997	6.7%
C	3791	2.1%
M	3791	2.1%

Most occurring blocks

Value	Count	Frequency (%)
ASCII	178242	100.0%

Most frequent character per block

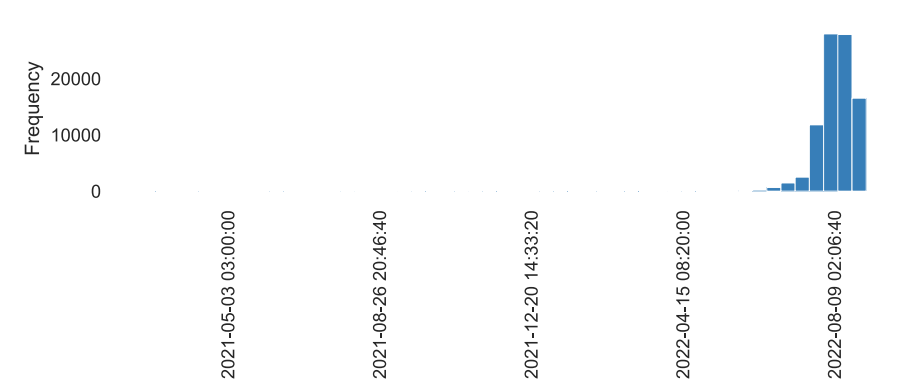
ASCII

Value	Count	Frequency (%)
G	73333	41.1%
I	73333	41.1%
F	11997	6.7%
W	11997	6.7%
C	3791	2.1%
M	3791	2.1%

create_dttm
Date

Distinct	47933
Distinct (%)	53.8%
Missing	0
Missing (%)	0.0%
Memory size	1.4 MiB

Minimum	2021-03-07 13:47:58
Maximum	2022-08-31 20:53:17



Histogram with fixed size bins (bins=50)

finish_dttm.1

Categorical

HIGH CARDINALITY UNIFORM

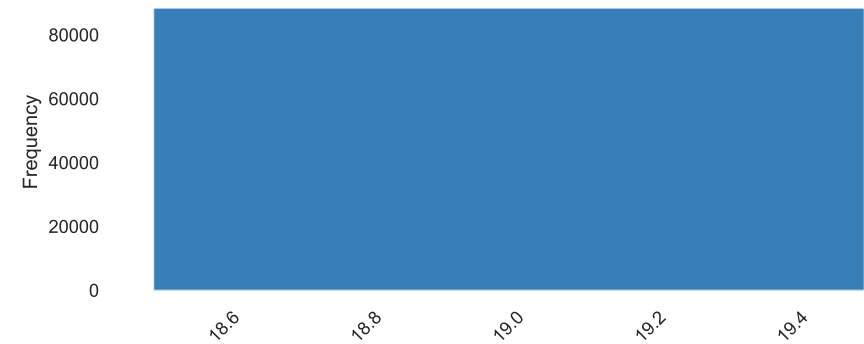
Distinct	48358
Distinct (%)	54.7%
Missing	751
Missing (%)	0.8%
Memory size	1.4 MiB

Length		Characters and Unicode		Unique		Sample	
Max length	19	Total	1679030	Unique	27614	1st row	2022-10-04 00:02:24
Median length	19	characters			?		
Mean length	19	Distinct	13	Unique (%)	31.2%	2nd row	2022-08-15 13:32:10
Min length	19	characters				3rd row	2022-08-15 16:19:49
		Distinct categories	4 (https://en.wikipedia.org/wiki/Unicode_characters)			4th row	2022-08-19 10:37:09
		Distinct scripts	1 (https://en.wikipedia.org/wiki/Script_(Unicode))			5th row	2022-08-15 16:43:08
		Distinct blocks	1 (https://en.wikipedia.org/wiki/Unicode_blocks)				
		The Unicode Standard assigns character properties to each code point, which can be used to analyse textual variables.					

Common Values

Value	Count	Frequency (%)
2022-08-22 10:46:38	18	< 0.1%
2022-08-19 10:03:50	12	< 0.1%
2022-08-19 13:12:03	12	< 0.1%
2022-08-26 15:24:00	11	< 0.1%
2022-08-23 09:11:35	11	< 0.1%
2022-09-26 11:30:30	11	< 0.1%
2022-09-16 13:33:11	11	< 0.1%
2022-08-23 14:28:46	10	< 0.1%
2022-08-18 14:47:17	10	< 0.1%
2022-08-10 15:59:03	10	< 0.1%
Other values (48348)	88254	99.0%
(Missing)	751	0.8%

Length



Value	Count	Frequency (%)
2022-08-15	4691	2.7%
2022-08-29	4676	2.6%
2022-08-22	4095	2.3%
2022-08-30	3733	2.1%
2022-08-08	3695	2.1%
2022-08-26	3569	2.0%
2022-08-23	3477	2.0%
2022-08-19	3261	1.8%
2022-08-18	3102	1.8%
2022-08-09	3100	1.8%
Other values (28996)	139341	78.8%

Most occurring characters

Value	Count	Frequency (%)
2	361124	21.5%
0	285896	17.0%
-	176740	10.5%
:	176740	10.5%
1	168092	10.0%
8	103040	6.1%
	88370	5.3%
3	70773	4.2%
5	67368	4.0%
4	63533	3.8%
Other values (3)	117354	7.0%

Most occurring categories

Value	Count	Frequency (%)
Decimal Number	1237180	73.7%
Dash Punctuation	176740	10.5%
Other Punctuation	176740	10.5%
Space Separator	88370	5.3%

Most frequent character per category

Decimal Number

Value	Count	Frequency (%)
2	361124	29.2%
0	285896	23.1%
1	168092	13.6%
8	103040	8.3%
3	70773	5.7%

Value	Count	Frequency (%)
5	67368	5.4%
4	63533	5.1%
9	54092	4.4%
6	35210	2.8%
7	28052	2.3%

Dash Punctuation

Value	Count	Frequency (%)
-	176740	100.0%

Other Punctuation

Value	Count	Frequency (%)
:	176740	100.0%

Space Separator

Value	Count	Frequency (%)
	88370	100.0%

Most occurring scripts

Value	Count	Frequency (%)
Common	1679030	100.0%

Most frequent character per script

Common

Value	Count	Frequency (%)
2	361124	21.5%
0	285896	17.0%
-	176740	10.5%
:	176740	10.5%
1	168092	10.0%
8	103040	6.1%
	88370	5.3%
3	70773	4.2%
5	67368	4.0%
4	63533	3.8%
Other values (3)	117354	7.0%

Most occurring blocks

Value	Count	Frequency (%)
ASCII	1679030	100.0%

Most frequent character per block

ASCII

Value	Count	Frequency (%)
2	361124	21.5%
0	285896	17.0%
-	176740	10.5%
:	176740	10.5%
1	168092	10.0%
8	103040	6.1%
	88370	5.3%
3	70773	4.2%
5	67368	4.0%
4	63533	3.8%
Other values (3)	117354	7.0%

hid
Categorical

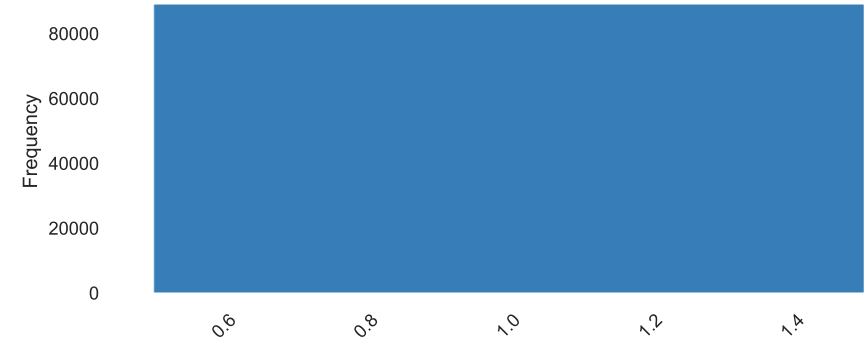
Distinct	4
Distinct (%)	< 0.1%
Missing	0
Missing (%)	0.0%
Memory size	1.4 MiB

Length		Characters and Unicode		Unique		Sample		
Max length	1	Total characters	89121	Unique	0	?	1st row	4
Median length	1			Unique (%)	0.0%	2nd row	1	
Mean length	1					3rd row	4	
Min length	1					4th row	4	
						5th row	4	?
		Distinct categories	1	(https://en.wikipedia.org/wiki/Unicode_character_property#General_Category)				?
		Distinct scripts	1	(https://en.wikipedia.org/wiki/Script_(Unicode)#List_of_scripts_in_Unicode)				
		Distinct blocks	1	(https://en.wikipedia.org/wiki/Unicode_block)				?
The Unicode Standard assigns character properties to each code point, which can be used to analyse textual variables.								

Common Values

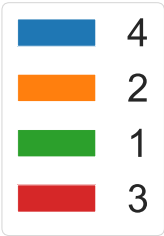
Value	Count	Frequency (%)
4	61388	68.9%
2	17779	19.9%
1	6434	7.2%
3	3520	3.9%

Length



Histogram of lengths of the category

Common Values (Plot)



Value	Count	Frequency (%)
4	61388	68.9%
2	17779	19.9%
1	6434	7.2%
3	3520	3.9%

Most occurring characters

Value	Count	Frequency (%)
4	61388	68.9%
2	17779	19.9%
1	6434	7.2%
3	3520	3.9%

Most occurring categories

Value	Count	Frequency (%)
Decimal Number	89121	100.0%

Most frequent character per category

Decimal Number

Value	Count	Frequency (%)
4	61388	68.9%
2	17779	19.9%
1	6434	7.2%
3	3520	3.9%

Most occurring scripts

Value	Count	Frequency (%)
Common	89121	100.0%

Most frequent character per script

Common

Value	Count	Frequency (%)
4	61388	68.9%
2	17779	19.9%
1	6434	7.2%
3	3520	3.9%

Most occurring blocks

Value	Count	Frequency (%)
ASCII	89121	100.0%

Most frequent character per block

ASCII

Value	Count	Frequency (%)
4	61388	68.9%
2	17779	19.9%
1	6434	7.2%
3	3520	3.9%

using_flg
Categorical

HIGH CORRELATION (This variable has a high overall correlation with 1 fields:
hit_status_result_desc) IMBALANCE

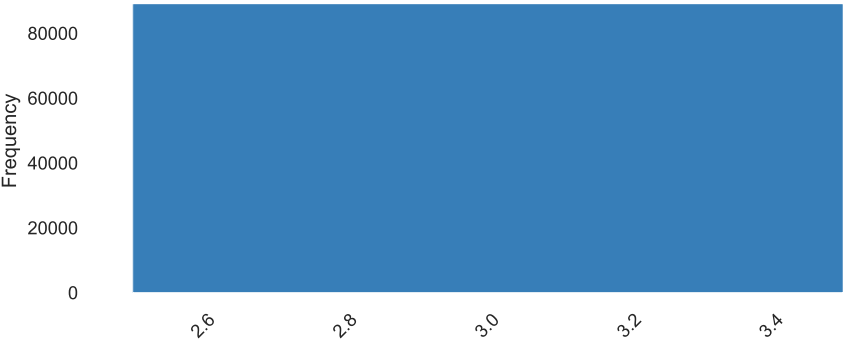
Distinct	2
Distinct (%)	< 0.1%
Missing	0
Missing (%)	0.0%
Memory size	1.4 MiB

Length		Characters and Unicode		Unique		Sample	
Max length	3	Total	267363	Unique	0	1st row	0.0
Median length	3	characters		Unique (%)	0.0%	2nd row	0.0
Mean length	3	Distinct	3			3rd row	0.0
Min length	3	characters				4th row	0.0
		Distinct	2			5th row	0.0
		categories	(https://en.wikipedia.org/wiki/Unicode_character_property#General_Category)				
		Distinct	1				?
		scripts	(https://en.wikipedia.org/wiki/Script_(Unicode)#List_of_scripts_in_Unicode)				
		Distinct	1	(https://en.wikipedia.org/wiki/Unicode_block)			
		blocks					?
The Unicode Standard assigns character properties to each code point, which can be used to analyse textual variables.							

Common Values

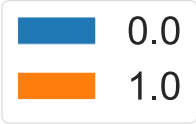
Value	Count	Frequency (%)
0.0	87589	98.3%
1.0	1532	1.7%

Length



Histogram of lengths of the category

Common Values (Plot)



Value	Count	Frequency (%)
0.0	87589	98.3%
1.0	1532	1.7%

Most occurring characters

Value	Count	Frequency (%)
0	176710	66.1%
.	89121	33.3%
1	1532	0.6%

Most occurring categories

Value	Count	Frequency (%)
Decimal Number	178242	66.7%
Other Punctuation	89121	33.3%

Most frequent character per category

Decimal Number

Value	Count	Frequency (%)
0	176710	99.1%
1	1532	0.9%

Other Punctuation

Value	Count	Frequency (%)
.	89121	100.0%

Most occurring scripts

Value	Count	Frequency (%)
Common	267363	100.0%

Most frequent character per script

Common

Value	Count	Frequency (%)
0	176710	66.1%
.	89121	33.3%

Value	Count	Frequency (%)
1	1532	0.6%

Most occurring blocks

Value	Count	Frequency (%)
ASCII	267363	100.0%

Most frequent character per block

ASCII

Value	Count	Frequency (%)
0	176710	66.1%
.	89121	33.3%
1	1532	0.6%

hit_status_result_desc
Categorical

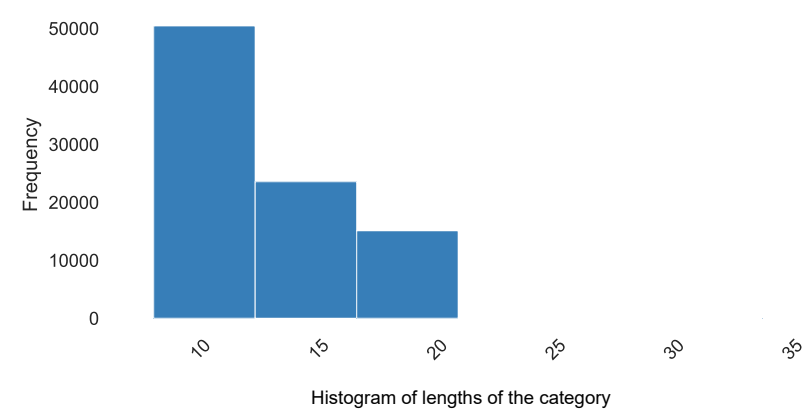
Distinct	7
Distinct (%)	< 0.1%
Missing	0
Missing (%)	0.0%
Memory size	1.4 MiB

Length		Characters and Unicode		Unique		Sample	
Max length	38	Total characters	1009379	Unique	0 ?	1st row	Дозвон, Перезвонить
Median length	8	Distinct characters	27	Unique (%)	0.0%	2nd row	Недозвон
Mean length	11.325939	Distinct categories	4 (https://en.wikipedia.org/wiki/Unicode_characters)			3rd row	Недозвон ?
Min length	8	Distinct scripts	2 (https://en.wikipedia.org/wiki/Script_(Unicode_block))			4th row	Недозвон ?
		Distinct blocks	2 (https://en.wikipedia.org/wiki/Unicode_block)			5th row	Недозвон ?
The Unicode Standard assigns character properties to each code point, which can be used to analyse textual variables.							

Common Values

Value	Count	Frequency (%)
Недозвон	50440	56.6%
Дозвон, Отказ	18445	20.7%
Дозвон, Перезвонить	15098	16.9%
Дозвон, Успешно	3308	3.7%
Дозвон, Отложить	1674	1.9%
Не было звонка	130	0.1%
Дозвон, Некорректное задание по звонку	26	< 0.1%

Length



Common Values (Plot)



Недозвон
Дозвон, Отказ
Дозвон, Перезвонить
Дозвон, Успешно
Дозвон, Отложить
Не было звонка
Дозвон, Некорректное задание по звонку

Value	Count	Frequency (%)
недозвон	50440	39.4%
дозвон	38551	30.1%
отказ	18445	14.4%
перезвонить	15098	11.8%
успешно	3308	2.6%
отложить	1674	1.3%
не	130	0.1%
было	130	0.1%
звонка	130	0.1%
некорректное	26	< 0.1%
Other values (3)	78	0.1%

Most occurring characters

Value	Count	Frequency (%)
о	198426	19.7%
з	122716	12.2%
н	107605	10.7%
в	104245	10.3%
е	84178	8.3%
Н	50596	5.0%
д	50466	5.0%
	38889	3.9%
Д	38551	3.8%
,	38551	3.8%
Other values (17)	175156	17.4%

Most occurring categories

Value	Count	Frequency (%)
Lowercase Letter	804267	79.7%

Value	Count	Frequency (%)
Uppercase Letter	127672	12.6%
Space Separator	38889	3.9%
Other Punctuation	38551	3.8%

Most frequent character per category

Lowercase Letter

Value	Count	Frequency (%)
о	198426	24.7%
з	122716	15.3%
н	107605	13.4%
в	104245	13.0%
е	84178	10.5%
д	50466	6.3%
т	36917	4.6%
к	18653	2.3%
а	18627	2.3%
и	16798	2.1%
Other values (10)	45636	5.7%

Uppercase Letter

Value	Count	Frequency (%)
Н	50596	39.6%
Д	38551	30.2%
О	20119	15.8%
П	15098	11.8%
У	3308	2.6%

Space Separator

Value	Count	Frequency (%)
	38889	100.0%

Other Punctuation

Value	Count	Frequency (%)
,	38551	100.0%

Most occurring scripts

Value	Count	Frequency (%)
Cyrillic	931939	92.3%
Common	77440	7.7%

Most frequent character per script

Cyrillic

Value	Count	Frequency (%)
о	198426	21.3%

Value	Count	Frequency (%)
з	122716	13.2%
н	107605	11.5%
в	104245	11.2%
е	84178	9.0%
Н	50596	5.4%
д	50466	5.4%
Д	38551	4.1%
т	36917	4.0%
О	20119	2.2%
Other values (15)	118120	12.7%

Common

Value	Count	Frequency (%)
	38889	50.2%
,	38551	49.8%

Most occurring blocks

Value	Count	Frequency (%)
Cyrillic	931939	92.3%
ASCII	77440	7.7%

Most frequent character per block

Cyrillic

Value	Count	Frequency (%)
о	198426	21.3%
з	122716	13.2%
н	107605	11.5%
в	104245	11.2%
е	84178	9.0%
Н	50596	5.4%
д	50466	5.4%
Д	38551	4.1%
т	36917	4.0%
О	20119	2.2%
Other values (15)	118120	12.7%

ASCII

Value	Count	Frequency (%)
	38889	50.2%
,	38551	49.8%

queue_desc
Categorical

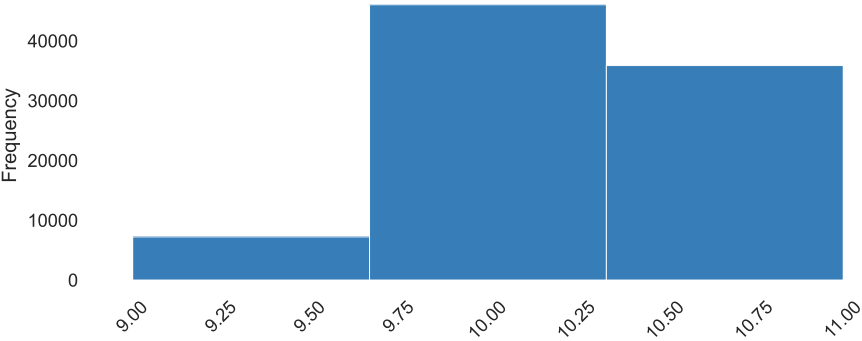
Distinct	155
Distinct (%)	0.2%
Missing	2
Missing (%)	< 0.1%
Memory size	1.4 MiB

Length		Characters and Unicode		Unique		Sample			
Max length	11	Total characters	919838	Unique	8	?	1st row	Очередь 50	
Median length	10			Unique (%)	< 0.1%	2nd row	Очередь 78		
Mean length	10.321458	Distinct characters	17	3rd row				Очередь 119	
Min length	9	Distinct categories	4	4th row				Очередь 8	?
		Distinct scripts	2	5th row				Очередь 119	?
		Distinct blocks	2 (https://en.wikipedia.org/wiki/Unicode_block)						?
The Unicode Standard assigns character properties to each code point, which can be used to analyse textual variables.									

Common Values

Value	Count	Frequency (%)
Очередь 119	8874	10.0%
Очередь 50	6572	7.4%
Очередь 41	5350	6.0%
Очередь 123	4076	4.6%
Очередь 7	3978	4.5%
Очередь 58	3942	4.4%
Очередь 143	3160	3.5%
Очередь 23	3153	3.5%
Очередь 148	2774	3.1%
Очередь 8	2263	2.5%
Other values (145)	44977	50.5%

Length



Histogram of lengths of the category

Value	Count	Frequency (%)
очередь	89119	50.0%
119	8874	5.0%
50	6572	3.7%
41	5350	3.0%
123	4076	2.3%
7	3978	2.2%
58	3942	2.2%
143	3160	1.8%
23	3153	1.8%
148	2774	1.6%
Other values (146)	47240	26.5%

Most occurring characters

Value	Count	Frequency (%)
e	178238	19.4%
O	89119	9.7%
p	89119	9.7%
д	89119	9.7%
ь	89119	9.7%
	89119	9.7%
ч	89119	9.7%
1	55922	6.1%
3	26314	2.9%
5	23446	2.5%
Other values (7)	101204	11.0%

Most occurring categories

Value	Count	Frequency (%)
Lowercase Letter	534714	58.1%
Decimal Number	206886	22.5%
Uppercase Letter	89119	9.7%
Space Separator	89119	9.7%

Most frequent character per category

Decimal Number

Value	Count	Frequency (%)
1	55922	27.0%
3	26314	12.7%
5	23446	11.3%
4	21751	10.5%
2	19761	9.6%
9	17210	8.3%

Value	Count	Frequency (%)
8	13345	6.5%
0	11354	5.5%
7	9514	4.6%
6	8269	4.0%

Lowercase Letter

Value	Count	Frequency (%)
e	178238	33.3%
p	89119	16.7%
д	89119	16.7%
ь	89119	16.7%
ч	89119	16.7%

Uppercase Letter

Value	Count	Frequency (%)
О	89119	100.0%

Space Separator

Value	Count	Frequency (%)
	89119	100.0%

Most occurring scripts

Value	Count	Frequency (%)
Cyrillic	623833	67.8%
Common	296005	32.2%

Most frequent character per script

Common

Value	Count	Frequency (%)
	89119	30.1%
1	55922	18.9%
3	26314	8.9%
5	23446	7.9%
4	21751	7.3%
2	19761	6.7%
9	17210	5.8%
8	13345	4.5%
0	11354	3.8%
7	9514	3.2%

Cyrillic

Value	Count	Frequency (%)
e	178238	28.6%
О	89119	14.3%

Value	Count	Frequency (%)
р	89119	14.3%
д	89119	14.3%
ь	89119	14.3%
ч	89119	14.3%

Most occurring blocks

Value	Count	Frequency (%)
Cyrillic	623833	67.8%
ASCII	296005	32.2%

Most frequent character per block

Cyrillic

Value	Count	Frequency (%)
е	178238	28.6%
О	89119	14.3%
р	89119	14.3%
д	89119	14.3%
ь	89119	14.3%
ч	89119	14.3%

ASCII

Value	Count	Frequency (%)
	89119	30.1%
1	55922	18.9%
3	26314	8.9%
5	23446	7.9%
4	21751	7.3%
2	19761	6.7%
9	17210	5.8%
8	13345	4.5%
0	11354	3.8%
7	9514	3.2%

group_nm
Categorical

HIGH CARDINALITY HIGH CORRELATION (This variable has a high overall correlation with 2 fields:
source_system_cd, management_nm)

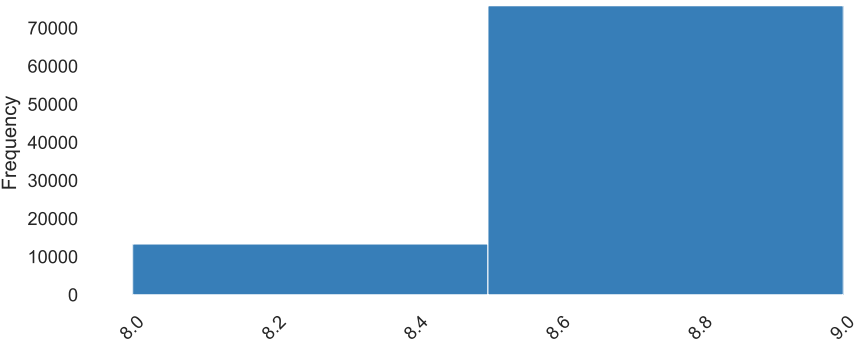
Distinct	52
Distinct (%)	0.1%
Missing	0
Missing (%)	0.0%
Memory size	1.4 MiB

Length		Characters and Unicode	Unique		Sample		
Max length	9		Unique	0 ?	1st row	Группа 41	
		Total characters	788785	Unique (%)	0.0%	2nd row	Группа 36
Median length	9	Distinct characters	16		3rd row	Группа 3	
Mean length	8.8507198				4th row	Группа 19	
Min length	8	Distinct categories	4 (https://en.wikipedia.org/wiki/Unicode_character_property#General_Category)		5th row	Группа 43 ?	
		Distinct scripts	2 (https://en.wikipedia.org/wiki/Script_(Unicode)#List_of_scripts_in_Unicode)			?	
		Distinct blocks	2 (https://en.wikipedia.org/wiki/Unicode_block)			?	
The Unicode Standard assigns character properties to each code point, which can be used to analyse textual variables.							

Common Values

Value	Count	Frequency (%)
Группа 36	16652	18.7%
Группа 29	6200	7.0%
Группа 5	5916	6.6%
Группа 19	5624	6.3%
Группа 30	4633	5.2%
Группа 23	4535	5.1%
Группа 41	4350	4.9%
Группа 52	3740	4.2%
Группа 48	2859	3.2%
Группа 45	2715	3.0%
Other values (42)	31897	35.8%

Length



Histogram of lengths of the category

Value	Count	Frequency (%)
гррпппа	89121	50.0%
36	16652	9.3%
29	6200	3.5%
5	5916	3.3%
19	5624	3.2%
30	4633	2.6%
23	4535	2.5%
41	4350	2.4%
52	3740	2.1%
48	2859	1.6%
Other values (43)	34612	19.4%

Most occurring characters

Value	Count	Frequency (%)
п	178242	22.6%
Г	89121	11.3%
р	89121	11.3%
у	89121	11.3%
а	89121	11.3%
	89121	11.3%
3	37506	4.8%
1	24016	3.0%
2	23642	3.0%
6	20883	2.6%
Other values (6)	58891	7.5%

Most occurring categories

Value	Count	Frequency (%)
Lowercase Letter	445605	56.5%
Decimal Number	164938	20.9%
Uppercase Letter	89121	11.3%
Space Separator	89121	11.3%

Most frequent character per category

Decimal Number

Value	Count	Frequency (%)
3	37506	22.7%
1	24016	14.6%
2	23642	14.3%
6	20883	12.7%
5	15513	9.4%
4	15284	9.3%

Value	Count	Frequency (%)
9	12453	7.6%
0	5995	3.6%
8	5482	3.3%
7	4164	2.5%

Lowercase Letter

Value	Count	Frequency (%)
n	178242	40.0%
p	89121	20.0%
y	89121	20.0%
a	89121	20.0%

Uppercase Letter

Value	Count	Frequency (%)
Г	89121	100.0%

Space Separator

Value	Count	Frequency (%)
	89121	100.0%

Most occurring scripts

Value	Count	Frequency (%)
Cyrillic	534726	67.8%
Common	254059	32.2%

Most frequent character per script

Common

Value	Count	Frequency (%)
	89121	35.1%
3	37506	14.8%
1	24016	9.5%
2	23642	9.3%
6	20883	8.2%
5	15513	6.1%
4	15284	6.0%
9	12453	4.9%
0	5995	2.4%
8	5482	2.2%

Cyrillic

Value	Count	Frequency (%)
п	178242	33.3%
Г	89121	16.7%
р	89121	16.7%
у	89121	16.7%

Value	Count	Frequency (%)
a	89121	16.7%

Most occurring blocks

Value	Count	Frequency (%)
Cyrillic	534726	67.8%
ASCII	254059	32.2%

Most frequent character per block

Cyrillic

Value	Count	Frequency (%)
п	178242	33.3%
Г	89121	16.7%
р	89121	16.7%
у	89121	16.7%
а	89121	16.7%

ASCII

Value	Count	Frequency (%)
	89121	35.1%
3	37506	14.8%
1	24016	9.5%
2	23642	9.3%
6	20883	8.2%
5	15513	6.1%
4	15284	6.0%
9	12453	4.9%
0	5995	2.4%
8	5482	2.2%

management_nm
Categorical

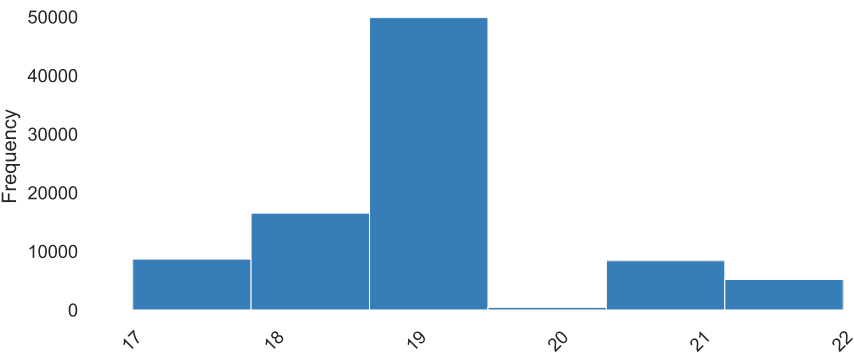
Distinct	45
Distinct (%)	0.1%
Missing	0
Missing (%)	0.0%
Memory size	1.4 MiB

Length		Characters and Unicode		Unique		Sample	
Max length	22	Total characters	1692291	Unique	0 ?	1st row	Управление КЦ Парис
Median length	19	Distinct characters	38	Unique (%)	0.0%	2nd row	Управление КЦ Чизан
Mean length	18.98869	Distinct categories	3 (https://en.wikipedia.org/wiki/Unicode_characters)			3rd row	Управление КЦ Кирис ?
Min length	17	Distinct scripts	2 (https://en.wikipedia.org/wiki/Script_(Unicode_block))			4th row	Управление КЦ Верс ?
		Distinct blocks	2 (https://en.wikipedia.org/wiki/Unicode_block)			5th row	Управление КЦ Мерс ?
The Unicode Standard assigns character properties to each code point, which can be used to analyse textual variables.							

Common Values

Value	Count	Frequency (%)
Управление КЦ Чизан	21328	23.9%
Управление КЦ Верс	6545	7.3%
Управление КЦ Дирим	6200	7.0%
Управление КЦ Ромашка	5916	6.6%
Управление КЦ Страница	5207	5.8%
Управление КЦ Парис	4350	4.9%
Управление КЦ Кас	4267	4.8%
Управление КЦ Неп	2859	3.2%
Управление КЦ Мазан	2715	3.0%
Управление КЦ Тазан	2585	2.9%
Other values (35)	27149	30.5%

Length



Histogram of lengths of the category

Value	Count	Frequency (%)
управление	89121	33.3%
кц	89121	33.3%
чизан	21328	8.0%
верс	6545	2.4%
дирим	6200	2.3%
ромашка	5916	2.2%
страница	5207	1.9%
парис	4350	1.6%
кас	4267	1.6%
неп	2859	1.1%
Other values (37)	32449	12.1%

Most occurring characters

Value	Count	Frequency (%)
е	196143	11.6%
	178242	10.5%
а	159716	9.4%
и	154498	9.1%
р	129981	7.7%
н	123067	7.3%
К	101255	6.0%
л	92639	5.5%
п	92431	5.5%
У	89121	5.3%
Other values (28)	375198	22.2%

Most occurring categories

Value	Count	Frequency (%)
Lowercase Letter	1157565	68.4%
Uppercase Letter	356484	21.1%
Space Separator	178242	10.5%

Most frequent character per category

Lowercase Letter

Value	Count	Frequency (%)
е	196143	16.9%
а	159716	13.8%
и	154498	13.3%
р	129981	11.2%
н	123067	10.6%
л	92639	8.0%
п	92431	8.0%

Value	Count	Frequency (%)
В	89121	7.7%
с	34273	3.0%
з	29142	2.5%
Other values (11)	56554	4.9%

Uppercase Letter

Value	Count	Frequency (%)
К	101255	28.4%
У	89121	25.0%
Ц	89121	25.0%
Ч	21700	6.1%
Д	8859	2.5%
В	8704	2.4%
С	7436	2.1%
П	6918	1.9%
М	6753	1.9%
Р	6718	1.9%
Other values (6)	9899	2.8%

Space Separator

Value	Count	Frequency (%)
	178242	100.0%

Most occurring scripts

Value	Count	Frequency (%)
Cyrillic	1514049	89.5%
Common	178242	10.5%

Most frequent character per script

Cyrillic

Value	Count	Frequency (%)
е	196143	13.0%
а	159716	10.5%
и	154498	10.2%
р	129981	8.6%
н	123067	8.1%
К	101255	6.7%
л	92639	6.1%
п	92431	6.1%
у	89121	5.9%
в	89121	5.9%
Other values (27)	286077	18.9%

Common

Value	Count	Frequency (%)
	178242	100.0%

Most occurring blocks

Value	Count	Frequency (%)
Cyrillic	1514049	89.5%
ASCII	178242	10.5%

Most frequent character per block

Cyrillic

Value	Count	Frequency (%)
е	196143	13.0%
а	159716	10.5%
и	154498	10.2%
р	129981	8.6%
н	123067	8.1%
к	101255	6.7%
л	92639	6.1%
п	92431	6.1%
у	89121	5.9%
в	89121	5.9%
Other values (27)	286077	18.9%

ASCII

Value	Count	Frequency (%)
	178242	100.0%

agent_login.1
Categorical

Distinct	1961
Distinct (%)	2.2%
Missing	0
Missing (%)	0.0%
Memory size	1.4 MiB

Length		Characters and Unicode		Unique		Sample			
Max length	8	Total characters	712968	Unique	40 ?	1st row	alovmma		
Median length	8			Unique (%)	< 0.1%	2nd row	6446kmmi		
Mean length	8								
Min length	8								
		Distinct characters	35						
		Distinct categories	3	5th row kovadsku ?					
		(https://en.wikipedia.org/wiki/Unicode_character_property#General_Category)							
		Distinct scripts	2	?					
		(https://en.wikipedia.org/wiki/Script_(Unicode)#List_of_scripts_in_Unicode)							
		Distinct blocks	1	(https://en.wikipedia.org/wiki/Unicode_block) ?					
</									

Value	Count	Frequency (%)
covaadru	184	0.2%
enkovaab	159	0.2%
levaadko	156	0.2%
kinalapo	152	0.2%
8857iario	151	0.2%
nayaivza	149	0.2%
rovaeagu	148	0.2%
0489vvvl	146	0.2%
novaadtr	143	0.2%
novannxo	143	0.2%
Other values (1951)	87590	98.3%

Most occurring characters

Value	Count	Frequency (%)
a	120983	17.0%
v	57275	8.0%
o	47582	6.7%
n	44696	6.3%
e	37203	5.2%
k	32807	4.6%
s	30318	4.3%
i	25419	3.6%
y	22256	3.1%
m	22167	3.1%
Other values (25)	272262	38.2%

Most occurring categories

Value	Count	Frequency (%)
Lowercase Letter	583407	81.8%
Decimal Number	129439	18.2%
Other Punctuation	122	< 0.1%

Most frequent character per category

Lowercase Letter

Value	Count	Frequency (%)
a	120983	20.7%
v	57275	9.8%
o	47582	8.2%
n	44696	7.7%
e	37203	6.4%
k	32807	5.6%
s	30318	5.2%

Value	Count	Frequency (%)
i	25419	4.4%
y	22256	3.8%
m	22167	3.8%
Other values (14)	142701	24.5%

Decimal Number

Value	Count	Frequency (%)
8	13752	10.6%
6	13601	10.5%
5	13541	10.5%
1	13407	10.4%
0	13258	10.2%
3	13106	10.1%
7	12499	9.7%
4	12492	9.7%
9	12237	9.5%
2	11546	8.9%

Other Punctuation

Value	Count	Frequency (%)
.	122	100.0%

Most occurring scripts

Value	Count	Frequency (%)
Latin	583407	81.8%
Common	129561	18.2%

Most frequent character per script

Latin

Value	Count	Frequency (%)
a	120983	20.7%
v	57275	9.8%
o	47582	8.2%
n	44696	7.7%
e	37203	6.4%
k	32807	5.6%
s	30318	5.2%
i	25419	4.4%
y	22256	3.8%
m	22167	3.8%
Other values (14)	142701	24.5%

Common

Value	Count	Frequency (%)
8	13752	10.6%
6	13601	10.5%
5	13541	10.5%
1	13407	10.3%
0	13258	10.2%
3	13106	10.1%
7	12499	9.6%
4	12492	9.6%
9	12237	9.4%
2	11546	8.9%

Most occurring blocks

Value	Count	Frequency (%)
ASCII	712968	100.0%

Most frequent character per block

ASCII

Value	Count	Frequency (%)
a	120983	17.0%
v	57275	8.0%
o	47582	6.7%
n	44696	6.3%
e	37203	5.2%
k	32807	4.6%
s	30318	4.3%
i	25419	3.6%
y	22256	3.1%
m	22167	3.1%
Other values (25)	272262	38.2%

horoscope
Categorical

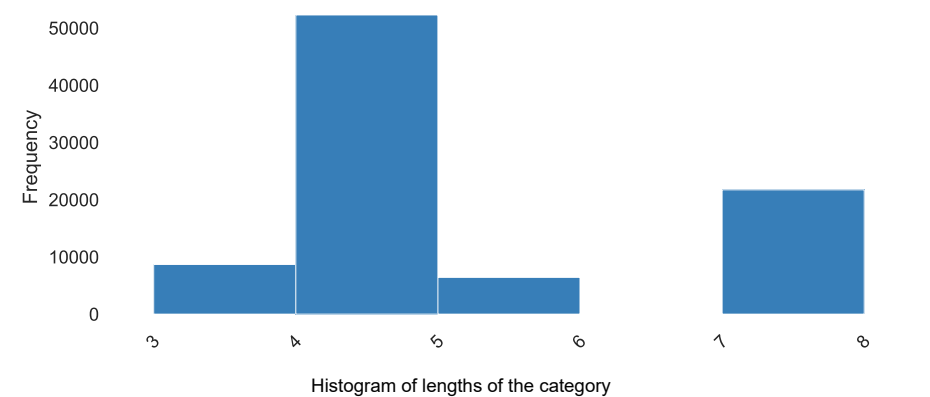
Distinct	12
Distinct (%)	< 0.1%
Missing	0
Missing (%)	0.0%
Memory size	1.4 MiB

Length		Characters and Unicode		Unique		Sample	
Max length	8	Total characters	428388	Unique	0 ?	1st row	Водолей
Median length	4	Distinct characters	28	Unique (%)	0.0%	2nd row	Овен
Mean length	4.8068132	Distinct categories	2 (https://en.wikipedia.org/wiki/Unicode_character_property#General_Category)			3rd row	Лев
Min length	3	Distinct scripts	1 (https://en.wikipedia.org/wiki/Script_(Unicode)#List_of_scripts_in_Unicode)			4th row	Овен
		Distinct blocks	1 (https://en.wikipedia.org/wiki/Unicode_block)			5th row	Водолей ?
							?
		The Unicode Standard assigns character properties to each code point, which can be used to analyse textual variables.					

Common Values

Value	Count	Frequency (%)
Дева	34967	39.2%
Рак	7366	8.3%
Близнецы	7294	8.2%
Весы	6970	7.8%
Телец	6424	7.2%
Водолей	6133	6.9%
Овен	5291	5.9%
Рыбы	5059	5.7%
Стрелец	3350	3.8%
Козерог	3257	3.7%
Other values (2)	3010	3.4%

Length



Value	Count	Frequency (%)
дева	34967	39.2%
рак	7366	8.3%
близнецы	7294	8.2%
весы	6970	7.8%
телец	6424	7.2%
водолей	6133	6.9%
овен	5291	5.9%
рыбы	5059	5.7%
стрелец	3350	3.8%
козерог	3257	3.7%
Other values (2)	3010	3.4%

Most occurring characters

Value	Count	Frequency (%)
е	84778	19.8%
а	42333	9.9%
в	41576	9.7%
Д	34967	8.2%
ы	24382	5.7%
л	23201	5.4%
о	22164	5.2%
ц	17068	4.0%
н	14277	3.3%
В	13103	3.1%
Other values (18)	110539	25.8%

Most occurring categories

Value	Count	Frequency (%)
Lowercase Letter	339267	79.2%
Uppercase Letter	89121	20.8%

Most frequent character per category

Lowercase Letter

Value	Count	Frequency (%)
е	84778	25.0%
а	42333	12.5%
в	41576	12.3%
ы	24382	7.2%
л	23201	6.8%
о	22164	6.5%
ц	17068	5.0%
н	14277	4.2%

Value	Count	Frequency (%)
з	10551	3.1%
к	9058	2.7%
Other values (9)	49879	14.7%

Uppercase Letter

Value	Count	Frequency (%)
Д	34967	39.2%
В	13103	14.7%
Р	12425	13.9%
Б	7294	8.2%
Т	6424	7.2%
О	5291	5.9%
С	5042	5.7%
К	3257	3.7%
Л	1318	1.5%

Most occurring scripts

Value	Count	Frequency (%)
Cyrillic	428388	100.0%

Most frequent character per script

Cyrillic

Value	Count	Frequency (%)
е	84778	19.8%
а	42333	9.9%
в	41576	9.7%
Д	34967	8.2%
ы	24382	5.7%
л	23201	5.4%
о	22164	5.2%
ц	17068	4.0%
н	14277	3.3%
В	13103	3.1%
Other values (18)	110539	25.8%

Most occurring blocks

Value	Count	Frequency (%)
Cyrillic	428388	100.0%

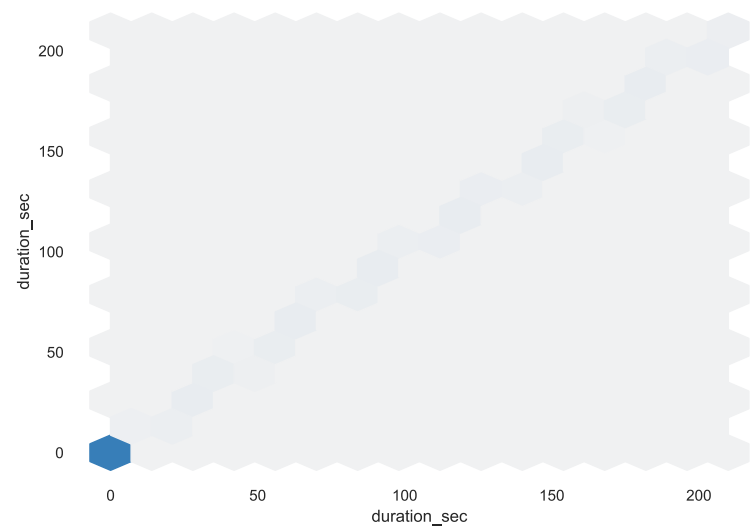
Most frequent character per block

Cyrillic

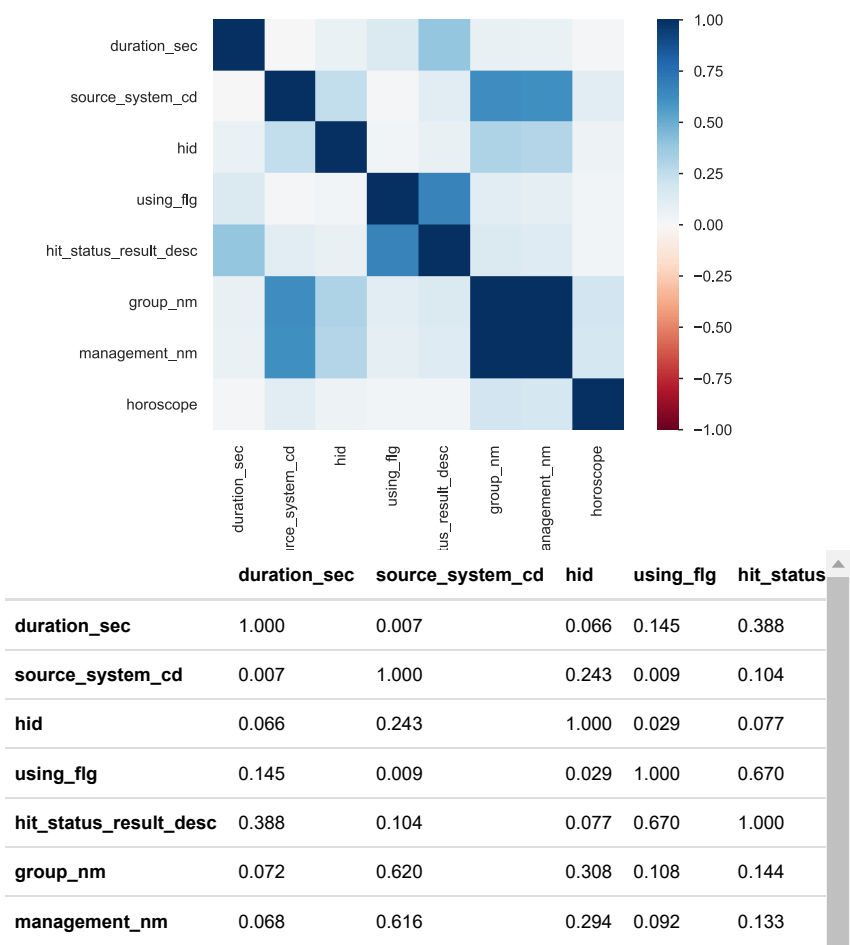
Value	Count	Frequency (%)
е	84778	19.8%

Value	Count	Frequency (%)
а	42333	9.9%
в	41576	9.7%
д	34967	8.2%
ы	24382	5.7%
л	23201	5.4%
о	22164	5.2%
ц	17068	4.0%
н	14277	3.3%
в	13103	3.1%
Other values (18)	110539	25.8%

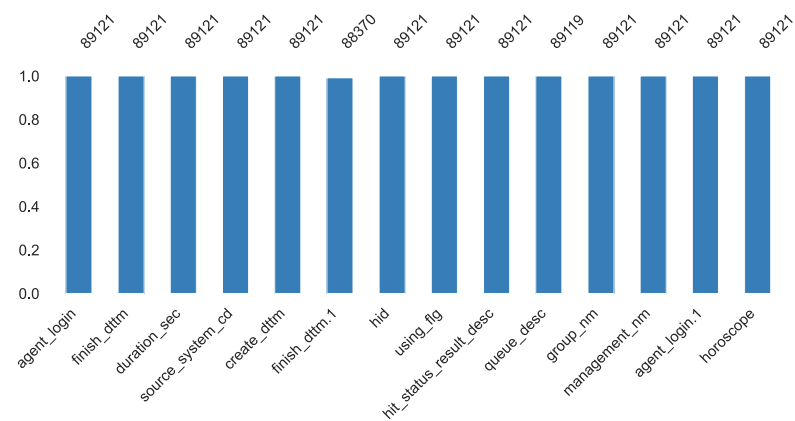
Interactions



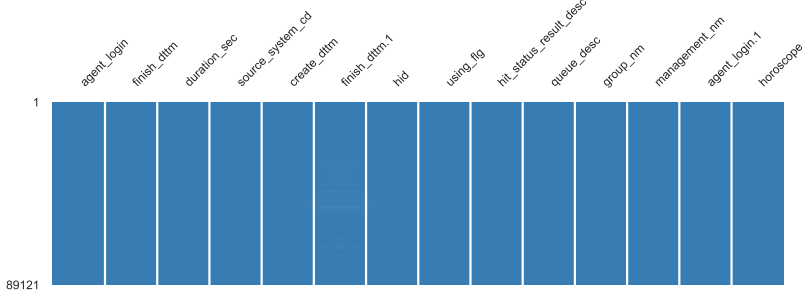
Correlations



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

	agent_login	finish_dttm	duration_sec	source_system_cd	create_dttm
0	alovmma	2022-08-15 15:52:33	160.63	GI	2022-08-03 03:44:01
1	6446kmmi	2022-08-15 10:29:33	0.00	GI	2022-08-14 14:00:51
2	zovaiaoz	2022-08-15 16:19:44	0.00	GI	2022-08-06 03:14:48
3	7787nyua	2022-08-15 13:32:01	0.00	GI	2022-08-15 12:31:49
4	kovadsku	2022-08-15 16:43:03	0.00	GI	2022-08-04 02:25:08
5	kinasgsa	2022-08-15 16:12:54	0.00	GI	2022-08-15 16:05:52
6	novaslp	2022-08-15 06:56:18	159.26	GI	2022-08-06 17:16:06
7	0218sase	2022-08-15 15:12:39	0.00	GI	2022-08-11 00:13:38
8	kinavase	2022-08-15 08:19:44	103.47	GI	2022-07-25 02:41:45
9	8269mnis	2022-08-15 10:40:06	0.00	GI	2022-08-14 13:36:54

	agent_login	finish_dttm	duration_sec	source_system_cd	create_dttm
89164	eninviks	2022-08-10 11:06:07	207.66	GI	2022-08-10 04:49:
89165	4453avsk	2022-08-15 14:36:28	0.00	GI	2022-08-10 05:14:
89166	kovaaavi	2022-08-15 15:56:38	79.90	GI	2022-08-05 02:44:
89167	5431asku	2022-08-10 17:43:37	172.54	GI	2022-08-04 20:52:
89168	kovaorer	2022-08-15 14:43:34	0.00	GI	2022-07-28 23:56:
89169	5037rbuz	2022-08-10 09:40:15	105.70	CM	2022-08-07 13:46:
89170	1116ovyu	2022-08-15 09:21:34	0.00	GI	2022-08-12 22:32:
89171	kovalaby	2022-08-10 16:16:17	0.00	GI	2022-08-08 15:46:
89172	ryanigxa	2022-08-15 16:41:16	0.00	GI	2022-07-16 05:03:
89173	7754kvmi	2022-08-10 13:50:01	44.14	GI	2022-08-08 17:44: