Overview

Dataset statistics

Number of variables	7
Number of observations	500
Missing cells	0
Missing cells (%)	0.0%
Duplicate rows	0
Duplicate rows (%)	0.0%
Total size in memory	251.9 KiB
Average record size in memory	516.0 B

Variable types

Text	7

Alerts

`_id` has unique values	Unique
`nconst` has unique values	Unique
`knownForTitles` has unique values	Unique

Reproduction

Analysis started	2024-02-03 05:47:14.180489
Analysis finished	2024-02-03 05:47:14.675686
Duration	0.5 seconds
Software version	ydata-profiling vv4.6.2 (https://github.com/ydataai/ydata-profiling)
Download configuration	config.json (data:text/plain;charset=utf-8,%7B%22title%22%3A%20%22Pandas%20Profiling%20Report%22%2C%20%22dataset%22%3A%20%7B%22description%22%3A%20%22%22%2C%20%22creator%22%3A%20%22%22%2C%20%22author%22%3A%2

Variables

Select Columns

_id

Text

UNIQUE

Distinct	500
Distinct (%)	100.0%
Missing	0
Missing (%)	0.0%
Memory size	39.7 KiB

Length

Max length	24
Median length	24
Mean length	24
Min length	24

Characters and Unicode

Total characters	12000	
Distinct characters	16	
Distinct categories	2 (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)	?

Unique

Unique	500	?
Unique (%)	100.0%	

Sample

1st row	617082fa2299cb9f2c2717ac
2nd row	617082fa2299cb9f2c2717ad
3rd row	617082fa2299cb9f2c2717ae
4th row	617082fa2299cb9f2c2717af
5th row	617082fa2299cb9f2c2717b0

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

Value	Count	Frequency (%)
617082fa2299cb9f2c2717ac	1	0.2%
617082fa2299cb9f2c2717bb	1	0.2%
617082fa2299cb9f2c2717ae	1	0.2%

Value	Count	Frequency (%)
617082fa2299cb9f2c2717af	1	0.2%
617082fa2299cb9f2c2717b0	1	0.2%
617082fa2299cb9f2c2717b1	1	0.2%
617082fa2299cb9f2c2717b2	1	0.2%
617082fa2299cb9f2c2717b3	1	0.2%
617082fa2299cb9f2c2717b4	1	0.2%
617082fa2299cb9f2c2717b5	1	0.2%
Other values (490)	490	98.0%

Most occurring characters

Value	Count	Frequency (%)
2	2563	21.4%
9	1723	14.4%
7	1147	9.6%
f	1064	8.9%
c	1064	8.9%
1	1063	8.9%
8	819	6.8%
6	563	4.7%
0	563	4.7%
b	563	4.7%
Other values (6)	868	7.2%

Most occurring categories

Value	Count	Frequency (%)
Decimal Number	8630	71.9%
Lowercase Letter	3370	28.1%

Most frequent character per category

Decimal Number

Value	Count	Frequency (%)
2	2563	29.7%
9	1723	20.0%
7	1147	13.3%
1	1063	12.3%
8	819	9.5%
6	563	6.5%
0	563	6.5%
5	63	0.7%
4	63	0.7%
3	63	0.7%

Lowercase Letter

Value	Count	Frequency (%)
f	1064	31.6%
c	1064	31.6%
b	563	16.7%
a	551	16.4%
e	64	1.9%
d	64	1.9%

Most occurring scripts

Value	Count	Frequency (%)
Common	8630	71.9%
Latin	3370	28.1%

Most frequent character per script

Common

Value	Count	Frequency (%)
2	2563	29.7%
9	1723	20.0%
7	1147	13.3%
1	1063	12.3%
8	819	9.5%
6	563	6.5%
0	563	6.5%
5	63	0.7%
4	63	0.7%
3	63	0.7%

Latin

Value	Count	Frequency (%)
f	1064	31.6%
c	1064	31.6%
b	563	16.7%
a	551	16.4%
e	64	1.9%
d	64	1.9%

Most occurring blocks

Value	Count	Frequency (%)
ASCII	12000	100.0%

Most frequent character per block

ASCII

Value	Count	Frequency (%)
2	2563	21.4%
9	1723	14.4%
7	1147	9.6%
f	1064	8.9%
c	1064	8.9%
1	1063	8.9%
8	819	6.8%
6	563	4.7%
0	563	4.7%
b	563	4.7%
Other values (6)	868	7.2%

nconst

Text

UNIQUE

Distinct	500
Distinct (%)	100.0%
Missing	0
Missing (%)	0.0%
Memory size	32.4 KiB

Length

Max length	9
Median length	9
Mean length	9
Min length	9

Characters and Unicode

Total characters	4500	
Distinct characters	12	
Distinct categories	2 (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)	?

Unique

Unique	500	?
Unique (%)	100.0%	

Sample

1st row	nm0000001
2nd row	nm0000005
3rd row	nm0000006
4th row	nm0000007
5th row	nm0000008

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

Value	Count	Frequency (%)
nm0000001	1	0.2%
nm0000018	1	0.2%
nm0000006	1	0.2%
nm0000007	1	0.2%
nm0000008	1	0.2%

Value	Count	Frequency (%)
nm0000009	1	0.2%
nm0000010	1	0.2%
nm0000011	1	0.2%
nm0000012	1	0.2%
nm0000004	1	0.2%
Other values (490)	490	98.0%

Most occurring characters

Value	Count	Frequency (%)
0	2200	48.9%
n	500	11.1%
m	500	11.1%
2	201	4.5%
1	200	4.4%
3	200	4.4%
4	199	4.4%
5	102	2.3%
7	100	2.2%
6	100	2.2%
Other values (2)	198	4.4%

Most occurring categories

Value	Count	Frequency (%)
Decimal Number	3500	77.8%
Lowercase Letter	1000	22.2%

Most frequent character per category

Decimal Number

Value	Count	Frequency (%)
0	2200	62.9%
2	201	5.7%
1	200	5.7%
3	200	5.7%
4	199	5.7%
5	102	2.9%
7	100	2.9%
6	100	2.9%
8	99	2.8%
9	99	2.8%

Lowercase Letter

Value	Count	Frequency (%)
n	500	50.0%
m	500	50.0%

Most occurring scripts

Value	Count	Frequency (%)
Common	3500	77.8%
Latin	1000	22.2%

Most frequent character per script

Common

Value	Count	Frequency (%)
0	2200	62.9%

Value	Count	Frequency (%)
2	201	5.7%
1	200	5.7%
3	200	5.7%
4	199	5.7%
5	102	2.9%
7	100	2.9%
6	100	2.9%
8	99	2.8%
9	99	2.8%

Latin

Value	Count	Frequency (%)
n	500	50.0%
m	500	50.0%

Most occurring blocks

Value	Count	Frequency (%)
ASCII	4500	100.0%

Most frequent character per block

ASCII

Value	Count	Frequency (%)
0	2200	48.9%
n	500	11.1%
m	500	11.1%
2	201	4.5%
1	200	4.4%
3	200	4.4%

Value	Count	Frequency (%)
4	199	4.4%
5	102	2.3%
7	100	2.2%
6	100	2.2%
Other values (2)	198	4.4%

primaryName

Text

Distinct	499
Distinct (%)	99.8%
Missing	0
Missing (%)	0.0%
Memory size	34.8 KiB

Length

Max length	23
Median length	20
Mean length	13.078
Min length	4

Characters and Unicode

Total characters	6539	
Distinct characters	60	
Distinct categories	5 (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	2 (https://en.wikipedia.org/wiki/Unicode_block)	?

Unique

Unique	498	?
Unique (%)	99.6%	

Sample

1st row	Fred Astaire
2nd row	Ingmar Bergman
3rd row	Ingrid Bergman
4th row	Humphrey Bogart
5th row	Marlon Brando

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

Value	Count	Frequency (%)
john	16	1.6%
robert	10	1.0%
james	10	1.0%
david	9	0.9%
jennifer	8	0.8%

Value	Count	Frequency (%)
michael	8	0.8%
richard	7	0.7%
lee	7	0.7%
de	6	0.6%
baldwin	4	0.4%
Other values (789)	939	91.7%

Most occurring characters

Value	Count	Frequency (%)
e	670	10.2%
a	556	8.5%
	524	8.0%
n	521	8.0%
r	473	7.2%
i	393	6.0%
o	346	5.3%
l	325	5.0%
t	226	3.5%
s	200	3.1%
Other values (50)	2305	35.3%

Most occurring categories

Value	Count	Frequency (%)
Lowercase Letter	4947	75.7%
Uppercase Letter	1042	15.9%

Value	Count	Frequency (%)
Space Separator	524	8.0%
Other Punctuation	16	0.2%
Dash Punctuation	10	0.2%

Most frequent character per category

Lowercase Letter

Value	Count	Frequency (%)
e	670	13.5%
a	556	11.2%
n	521	10.5%
r	473	9.6%
i	393	7.9%
o	346	7.0%
l	325	6.6%
t	226	4.6%
s	200	4.0%
h	167	3.4%
Other values (21)	1070	21.6%

Uppercase Letter

Value	Count	Frequency (%)
J	93	8.9%
B	88	8.4%
C	82	7.9%
D	81	7.8%
A	70	6.7%
H	66	6.3%
M	65	6.2%

Value	Count	Frequency (%)
S	63	6.0%
L	62	6.0%
G	56	5.4%
Other values (15)	316	30.3%

Other Punctuation

Value	Count	Frequency (%)
.	12	75.0%
'	4	25.0%

Space Separator

Value	Count	Frequency (%)
	524	100.0%

Dash Punctuation

Value	Count	Frequency (%)
-	10	100.0%

Most occurring scripts

Value	Count	Frequency (%)
Latin	5989	91.6%
Common	550	8.4%

Most frequent character per script

Latin

Value	Count	Frequency (%)
e	670	11.2%
a	556	9.3%

Value	Count	Frequency (%)
n	521	8.7%
r	473	7.9%
i	393	6.6%
o	346	5.8%
l	325	5.4%
t	226	3.8%
s	200	3.3%
h	167	2.8%
Other values (46)	2112	35.3%

Common

Value	Count	Frequency (%)
	524	95.3%
.	12	2.2%
-	10	1.8%
'	4	0.7%

Most occurring blocks

Value	Count	Frequency (%)
ASCII	6526	99.8%
None	13	0.2%

Most frequent character per block

ASCII

Value	Count	Frequency (%)
e	670	10.3%
a	556	8.5%
	524	8.0%

Value	Count	Frequency (%)
n	521	8.0%
r	473	7.2%
i	393	6.0%
o	346	5.3%
l	325	5.0%
t	226	3.5%
s	200	3.1%
Other values (45)	2292	35.1%

None

Value	Count	Frequency (%)
é	5	38.5%
è	3	23.1%
ó	3	23.1%
ñ	1	7.7%
ç	1	7.7%

birthYear

Text

Distinct	91
Distinct (%)	18.2%
Missing	0
Missing (%)	0.0%
Memory size	29.9 KiB

Length

Max length	4
Median length	4
Mean length	3.968
Min length	2

Characters and Unicode

Total characters	1984	
Distinct characters	12	
Distinct categories	3 (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)	?

Unique

Unique	14	?
Unique (%)	2.8%	

Sample

1st row	1899
2nd row	1918
3rd row	1915
4th row	1899
5th row	1924

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

Value	Count	Frequency (%)
1958	20	4.0%
1964	17	3.4%
1962	16	3.2%
1960	15	3.0%
1946	15	3.0%

Value	Count	Frequency (%)
1956	13	2.6%
1950	12	2.4%
1963	12	2.4%
1968	12	2.4%
1969	12	2.4%
Other values (81)	356	71.2%

Most occurring characters

Value	Count	Frequency (%)
1	553	27.9%
9	535	27.0%
6	167	8.4%
5	155	7.8%
4	141	7.1%
0	88	4.4%
7	88	4.4%
8	84	4.2%
2	81	4.1%
3	76	3.8%
Other values (2)	16	0.8%

Most occurring categories

Value	Count	Frequency (%)
Decimal Number	1968	99.2%
Other Punctuation	8	0.4%

Value	Count	Frequency (%)
Uppercase Letter	8	0.4%

Most frequent character per category

Decimal Number

Value	Count	Frequency (%)
1	553	28.1%
9	535	27.2%
6	167	8.5%
5	155	7.9%
4	141	7.2%
0	88	4.5%
7	88	4.5%
8	84	4.3%
2	81	4.1%
3	76	3.9%

Other Punctuation

Value	Count	Frequency (%)
\	8	100.0%

Uppercase Letter

Value	Count	Frequency (%)
N	8	100.0%

Most occurring scripts

Value	Count	Frequency (%)
Common	1976	99.6%
Latin	8	0.4%

Most frequent character per script

Common

Value	Count	Frequency (%)
1	553	28.0%
9	535	27.1%
6	167	8.5%
5	155	7.8%
4	141	7.1%
0	88	4.5%
7	88	4.5%
8	84	4.3%
2	81	4.1%
3	76	3.8%

Latin

Value	Count	Frequency (%)
N	8	100.0%

Most occurring blocks

Value	Count	Frequency (%)
ASCII	1984	100.0%

Most frequent character per block

ASCII

Value	Count	Frequency (%)
1	553	27.9%
9	535	27.0%
6	167	8.4%

Value	Count	Frequency (%)
5	155	7.8%
4	141	7.1%
0	88	4.4%
7	88	4.4%
8	84	4.2%
2	81	4.1%
3	76	3.8%
Other values (2)	16	0.8%

deathYear

Text

Distinct	65
Distinct (%)	13.0%
Missing	0
Missing (%)	0.0%
Memory size	29.2 KiB

Length

Max length	4
Median length	2
Mean length	2.552
Min length	2

Characters and Unicode

Total characters	1276	
Distinct characters	12	
Distinct categories	3 (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)	?

Unique

Unique	23	?
Unique (%)	4.6%	

Sample

1st row	1987
2nd row	2007
3rd row	1982
4th row	1957
5th row	2004

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

Value	Count	Frequency (%)
n	362	72.4%
2020	5	1.0%
2014	5	1.0%
2011	4	0.8%
1993	4	0.8%

Value	Count	Frequency (%)
1997	4	0.8%
2022	4	0.8%
1982	4	0.8%
1973	4	0.8%
2010	3	0.6%
Other values (55)	101	20.2%

Most occurring characters

Value	Count	Frequency (%)
\	362	28.4%
N	362	28.4%
1	120	9.4%
9	116	9.1%
0	90	7.1%
2	84	6.6%
7	33	2.6%
8	33	2.6%
6	26	2.0%
3	19	1.5%
Other values (2)	31	2.4%

Most occurring categories

Value	Count	Frequency (%)
Decimal Number	552	43.3%
Other Punctuation	362	28.4%

Value	Count	Frequency (%)
Uppercase Letter	362	28.4%

Most frequent character per category

Decimal Number

Value	Count	Frequency (%)
1	120	21.7%
9	116	21.0%
0	90	16.3%
2	84	15.2%
7	33	6.0%
8	33	6.0%
6	26	4.7%
3	19	3.4%
4	17	3.1%
5	14	2.5%

Other Punctuation

Value	Count	Frequency (%)
\	362	100.0%

Uppercase Letter

Value	Count	Frequency (%)
N	362	100.0%

Most occurring scripts

Value	Count	Frequency (%)
Common	914	71.6%
Latin	362	28.4%

Most frequent character per script

Common

Value	Count	Frequency (%)
\	362	39.6%
1	120	13.1%
9	116	12.7%
0	90	9.8%
2	84	9.2%
7	33	3.6%
8	33	3.6%
6	26	2.8%
3	19	2.1%
4	17	1.9%

Latin

Value	Count	Frequency (%)
N	362	100.0%

Most occurring blocks

Value	Count	Frequency (%)
ASCII	1276	100.0%

Most frequent character per block

ASCII

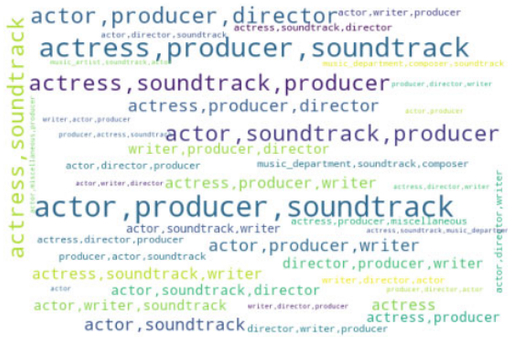
Value	Count	Frequency (%)
\	362	28.4%
N	362	28.4%
1	120	9.4%

Value	Count	Frequency (%)
9	116	9.1%
0	90	7.1%
2	84	6.6%
7	33	2.6%
8	33	2.6%
6	26	2.0%
3	19	1.5%
Other values (2)	31	2.4%

primaryProfession

Text

Distinct	106
Distinct (%)	21.2%
Missing	0
Missing (%)	0.0%
Memory size	39.8 KiB



Length

Max length	42
Median length	37
Mean length	24.21
Min length	5

Characters and Unicode

Total characters	12105	
Distinct characters	23	
Distinct categories	3 (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)	?

Unique

Unique	52	?
Unique (%)	10.4%	

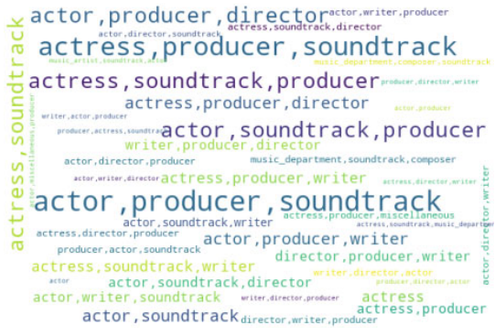
Sample

1st row	soundtrack,actor,miscellaneous
2nd row	writer,director,actor
3rd row	actress,soundtrack,producer
4th row	actor,soundtrack,producer
5th row	actor,soundtrack,director

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

Value	Count	Frequency (%)
actor,producer,soundtrack	55	11.0%
actress,producer,soundtrack	46	9.2%
actor,soundtrack,producer	30	6.0%
actress,soundtrack	29	5.8%
actor,producer,director	27	5.4%

Value	Count	Frequency (%)
actress,soundtrack,producer	26	5.2%
actress,producer,director	13	2.6%
actor,soundtrack	13	2.6%
actor,producer,writer	13	2.6%
actress,producer,writer	10	2.0%
Other values (96)	238	47.6%



Most occurring characters

Value	Count	Frequency (%)
r	1968	16.3%
c	1276	10.5%
t	1105	9.1%
o	1061	8.8%
e	935	7.7%
,	914	7.6%
a	832	6.9%
s	820	6.8%
d	785	6.5%
u	680	5.6%
Other values (13)	1729	14.3%

Most occurring categories

Value	Count	Frequency (%)
Lowercase Letter	11145	92.1%
Other Punctuation	914	7.6%

Value	Count	Frequency (%)
Connector Punctuation	46	0.4%

Most frequent character per category

Lowercase Letter

Value	Count	Frequency (%)
r	1968	17.7%
c	1276	11.4%
t	1105	9.9%
o	1061	9.5%
e	935	8.4%
a	832	7.5%
s	820	7.4%
d	785	7.0%
u	680	6.1%
p	374	3.4%
Other values (11)	1309	11.7%

Other Punctuation

Value	Count	Frequency (%)
,	914	100.0%

Connector Punctuation

Value	Count	Frequency (%)
_	46	100.0%

Most occurring scripts

Value	Count	Frequency (%)
Latin	11145	92.1%

Value	Count	Frequency (%)
Common	960	7.9%

Most frequent character per script

Latin

Value	Count	Frequency (%)
r	1968	17.7%
c	1276	11.4%
t	1105	9.9%
o	1061	9.5%
e	935	8.4%
a	832	7.5%
s	820	7.4%
d	785	7.0%
u	680	6.1%
p	374	3.4%
Other values (11)	1309	11.7%

Common

Value	Count	Frequency (%)
,	914	95.2%
–	46	4.8%

Most occurring blocks

Value	Count	Frequency (%)
ASCII	12105	100.0%

Most frequent character per block

ASCII

Value	Count	Frequency (%)
r	1968	16.3%
c	1276	10.5%
t	1105	9.1%
o	1061	8.8%
e	935	7.7%
,	914	7.6%
a	832	6.9%
s	820	6.8%
d	785	6.5%
u	680	5.6%
Other values (13)	1729	14.3%

knownForTitles

Text

UNIQUE

Distinct	500
Distinct (%)	100.0%
Missing	0
Missing (%)	0.0%
Memory size	47.0 KiB

Length

Max length	41
Median length	39
Mean length	39.012
Min length	39

Characters and Unicode

Total characters	19506
Distinct characters	12
Distinct categories	3 (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) ?

Unique

Unique	500	?
Unique (%)	100.0%	

Sample

1st row	tt0050419,tt0072308,tt0031983,tt0053137
2nd row	tt0083922,tt0069467,tt0050986,tt0050976
3rd row	tt0034583,tt0038787,tt0038109,tt0036855
4th row	tt0043265,tt0037382,tt0042593,tt0034583
5th row	tt0047296,tt0068646,tt0070849,tt0078788

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

Value	Count	Frequency (%)
tt0050419,tt0072308,tt0031983,tt0053137	1	0.2%
tt0050825,tt0054331,tt0080736,tt0049456	1	0.2%
tt0034583,tt0038787,tt0038109,tt0036855	1	0.2%
tt0043265,tt0037382,tt0042593,tt0034583	1	0.2%
tt0047296,tt0068646,tt0070849,tt0078788	1	0.2%

Value	Count	Frequency (%)
tt0059749,tt0061184,tt0087803,tt0057877	1	0.2%
tt0042041,tt0035575,tt0031867,tt0029870	1	0.2%
tt0035896,tt0044706,tt0027996,tt0034167	1	0.2%
tt0056687,tt0035140,tt0031210,tt0042192	1	0.2%
tt0078723,tt0080455,tt0072562,tt0077975	1	0.2%
Other values (490)	490	98.0%

Most occurring characters

Value	Count	Frequency (%)
t	4000	20.5%
0	3741	19.2%
1	1867	9.6%
,	1500	7.7%
8	1123	5.8%
2	1092	5.6%
9	1091	5.6%
7	1070	5.5%
4	1019	5.2%
3	1004	5.1%
Other values (2)	1999	10.2%

Most occurring categories

Value	Count	Frequency (%)
Decimal Number	14006	71.8%
Lowercase Letter	4000	20.5%

Value	Count	Frequency (%)
Other Punctuation	1500	7.7%

Most frequent character per category

Decimal Number

Value	Count	Frequency (%)
0	3741	26.7%
1	1867	13.3%
8	1123	8.0%
2	1092	7.8%
9	1091	7.8%
7	1070	7.6%
4	1019	7.3%
3	1004	7.2%
6	1004	7.2%
5	995	7.1%

Lowercase Letter

Value	Count	Frequency (%)
t	4000	100.0%

Other Punctuation

Value	Count	Frequency (%)
,	1500	100.0%

Most occurring scripts

Value	Count	Frequency (%)
Common	15506	79.5%
Latin	4000	20.5%

Most frequent character per script

Common

Value	Count	Frequency (%)
0	3741	24.1%
1	1867	12.0%
,	1500	9.7%
8	1123	7.2%
2	1092	7.0%
9	1091	7.0%
7	1070	6.9%
4	1019	6.6%
3	1004	6.5%
6	1004	6.5%

Latin

Value	Count	Frequency (%)
t	4000	100.0%

Most occurring blocks

Value	Count	Frequency (%)
ASCII	19506	100.0%

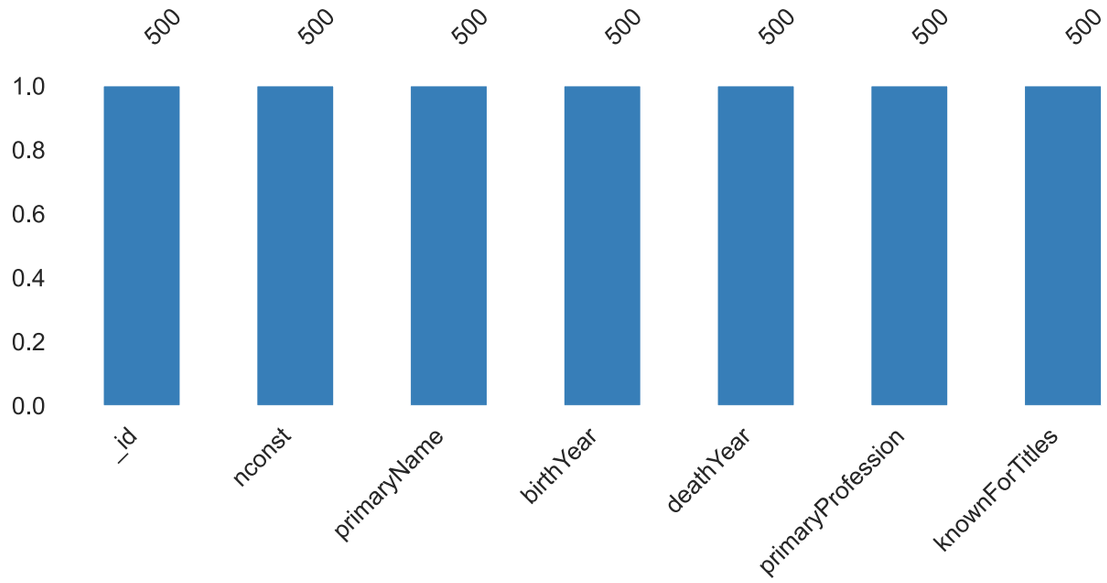
Most frequent character per block

ASCII

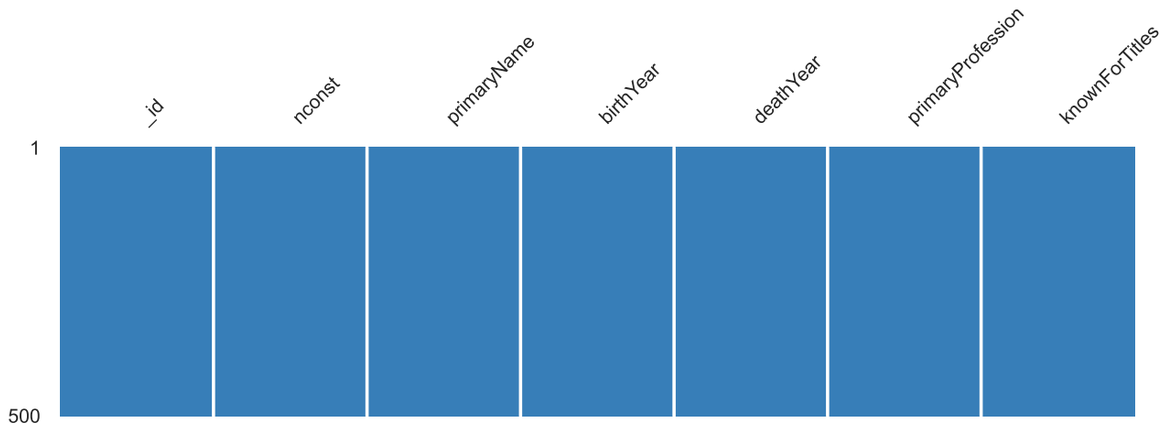
Value	Count	Frequency (%)
t	4000	20.5%
0	3741	19.2%
1	1867	9.6%

Value	Count	Frequency (%)
,	1500	7.7%
8	1123	5.8%
2	1092	5.6%
9	1091	5.6%
7	1070	5.5%
4	1019	5.2%
3	1004	5.1%
Other values (2)	1999	10.2%

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.

Sample

	_id	nconst	primaryName	birthYear	deathYear	primaryProfession	knownForTitles
0	617082fa2299cb9f2c2717ac	nm0000001	Fred Astaire	1899	1987	soundtrack,actor,miscellaneous	tt0050419,tt0072308,tt0031983,tt(
1	617082fa2299cb9f2c2717ad	nm0000005	Ingmar Bergman	1918	2007	writer,director,actor	tt0083922,tt0069467,tt0050986,tt(
2	617082fa2299cb9f2c2717ae	nm0000006	Ingrid Bergman	1915	1982	actress,soundtrack,producer	tt0034583,tt0038787,tt0038109,tt(
3	617082fa2299cb9f2c2717af	nm0000007	Humphrey Bogart	1899	1957	actor,soundtrack,producer	tt0043265,tt0037382,tt0042593,tt(
4	617082fa2299cb9f2c2717b0	nm0000008	Marlon Brando	1924	2004	actor,soundtrack,director	tt0047296,tt0068646,tt0070849,tt(
5	617082fa2299cb9f2c2717b1	nm0000009	Richard Burton	1925	1984	actor,soundtrack,producer	tt0059749,tt0061184,tt0087803,tt(
6	617082fa2299cb9f2c2717b2	nm0000010	James Cagney	1899	1986	actor,soundtrack,director	tt0042041,tt0035575,tt0031867,tt(
7	617082fa2299cb9f2c2717b3	nm0000011	Gary Cooper	1901	1961	actor,soundtrack,stunts	tt0035896,tt0044706,tt0027996,tt(
8	617082fa2299cb9f2c2717b4	nm0000012	Bette Davis	1908	1989	actress,soundtrack,make_up_department	tt0056687,tt0035140,tt0031210,tt(
9	617082fa2299cb9f2c2717b5	nm0000004	John Belushi	1949	1982	actor,soundtrack,writer	tt0078723,tt0080455,tt0072562,tt(

	_id	nconst	primaryName	birthYear	deathYear	primaryProfession	knownForTitles
490	617082fa2299cb9f2c271996	nm0000492	Jennifer Jason Leigh	1962	\N	actress,producer,soundtrack	tt0108122,tt0105414,tt3460252,tt
491	617082fa2299cb9f2c271997	nm0000493	Jack Lemmon	1925	2001	actor,soundtrack,producer	tt0053604,tt0053291,tt0104348,tt
492	617082fa2299cb9f2c271998	nm0000495	Téa Leoni	1966	\N	actress,producer,soundtrack	tt0218967,tt0163025,tt0112442,tt
493	617082fa2299cb9f2c271999	nm0000494	Robert Sean Leonard	1969	\N	actor,soundtrack	tt0412142,tt0107616,tt0097165,tt
494	617082fa2299cb9f2c27199a	nm0000496	Juliette Lewis	1973	\N	actress,soundtrack,director	tt0116367,tt0110632,tt0101540,tt1
495	617082fa2299cb9f2c27199b	nm0000497	Jennifer Lien	1974	\N	actress,producer,script_department	tt0112178,tt0133189,tt0120586,tt
496	617082fa2299cb9f2c27199c	nm0000498	Matthew Lillard	1970	\N	actor,producer,soundtrack	tt0331632,tt0267913,tt0133189,tt
497	617082fa2299cb9f2c27199d	nm0000499	Bai Ling	\N	\N	actress,producer,writer	tt1121931,tt0109506,tt0119994,tt
498	617082fa2299cb9f2c27199e	nm0000500	Richard Linklater	1960	\N	producer,director,writer	tt0243017,tt1065073,tt0405296,tt2
499	617082fa2299cb9f2c27199f	nm0000502	Christopher Lloyd	1938	\N	actor,soundtrack,producer	tt0106220,tt0088763,tt0101272,tt

Report generated by YData (https://ydata.ai/?utm_source=opensource&utm_medium=pandasprofiling&utm_campaign=report).