

# Overview

Dataset statistics	
Number of variables	14
Number of observations	10000
Missing cells	0
Missing cells (%)	0.0%
Duplicate rows	0
Duplicate rows (%)	0.0%
Total size in memory	1.1 MiB
Average record size in memory	112.0 B

Variable types	
Numeric	7
Categorical	7

Alerts

Surname has a high cardinality: 2932 distinct values	High cardinality
NumOfProducts is highly overall correlated with Exited	High correlation
Exited is highly overall correlated with NumOfProducts	High correlation
RowNumber is uniformly distributed	Uniform
RowNumber has unique values	Unique
CustomerId has unique values	Unique
Tenure has 413 (4.1%) zeros	Zeros
Balance has 3617 (36.2%) zeros	Zeros

Reproduction

Analysis started	2022-12-03 23:43:46.476309
Analysis finished	2022-12-03 23:43:58.393134
Duration	11.92 seconds
Software version	pandas-profiling vv3.5.0 ( <a href="https://github.com/pandas-profiling/pandas-profiling">https://github.com/pandas-profiling/pandas-profiling</a> )
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%

Variables

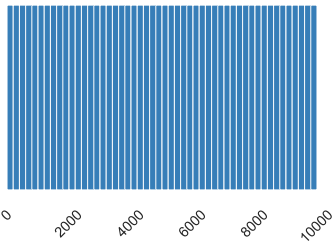
Select Columns ▾

RowNumber

Real number (ℝ)

UNIFORM

UNIQUE	
Distinct	10000
Distinct (%)	100.0%
Missing	0
Missing (%)	0.0%
Infinite	0
Infinite (%)	0.0%
Mean	5000.5
Minimum	1
Maximum	10000
Zeros	0
Zeros (%)	0.0%
Negative	0
Negative (%)	0.0%
Memory size	78.2 KiB



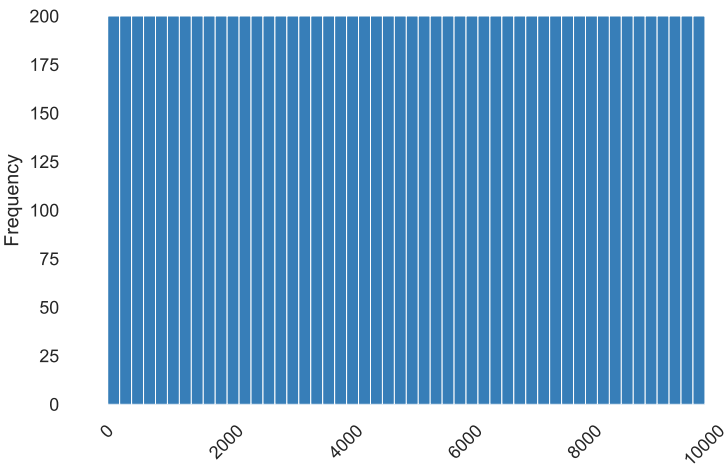
Quantile statistics

Minimum	1
5-th percentile	500.95
Q1	2500.75
median	5000.5
Q3	7500.25
95-th percentile	9500.05
Maximum	10000
Range	9999
Interquartile range (IQR)	4999.5

Descriptive statistics

Standard deviation	2886.8957
Coefficient of variation (CV)	0.5773214
Kurtosis	-1.2
Mean	5000.5
Median Absolute Deviation (MAD)	2500

Skewness	0
Sum	50005000
Variance	8334166.7
Monotonicity	Strictly increasing



Histogram with fixed size bins (bins=50)

Value	Count	Frequency (%)
1	1	< 0.1%
6671	1	< 0.1%
6664	1	< 0.1%
6665	1	< 0.1%
6666	1	< 0.1%
6667	1	< 0.1%
6668	1	< 0.1%
6669	1	< 0.1%
6670	1	< 0.1%
6672	1	< 0.1%
Other values (9990)	9990	99.9%

Value	Count	Frequency (%)
1	1	< 0.1%
2	1	< 0.1%
3	1	< 0.1%
4	1	< 0.1%
5	1	< 0.1%
6	1	< 0.1%
7	1	< 0.1%
8	1	< 0.1%
9	1	< 0.1%
10	1	< 0.1%



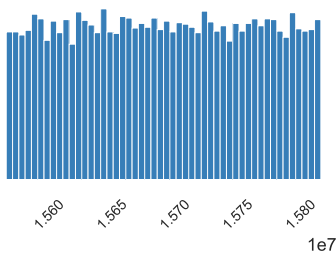
Value	Count	Frequency (%)
10000	1	< 0.1%
9999	1	< 0.1%
9998	1	< 0.1%
9997	1	< 0.1%
9996	1	< 0.1%
9995	1	< 0.1%
9994	1	< 0.1%
9993	1	< 0.1%
9992	1	< 0.1%
9991	1	< 0.1%





CustomerId  
Real number ( $\mathbb{R}$ )

Distinct	10000
Distinct (%)	100.0%
Missing	0
Missing (%)	0.0%
Infinite	0
Infinite (%)	0.0%
Mean	15690941
Minimum	15565701
Maximum	15815690
Zeros	0
Zeros (%)	0.0%
Negative	0
Negative (%)	0.0%
Memory size	78.2 KiB

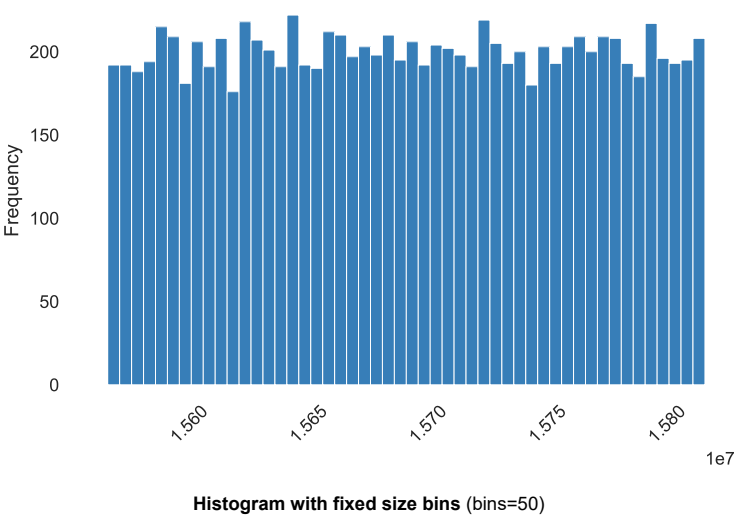


Quantile statistics

Minimum	15565701
5-th percentile	15578824
Q1	15628528
median	15690738
Q3	15753234
95-th percentile	15803034
Maximum	15815690
Range	249989
Interquartile range (IQR)	124705.5

Descriptive statistics

Standard deviation	71936.186
Coefficient of variation (CV)	0.0045845681
Kurtosis	-1.1961125
Mean	15690941
Median Absolute Deviation (MAD)	62432.5
Skewness	0.0011491459
Sum	$1.5690941 \times 10^{11}$
Variance	$5.1748149 \times 10^9$
Monotonicity	Not monotonic



Value	Count	Frequency (%)
15634602	1	< 0.1%
15667932	1	< 0.1%
15766185	1	< 0.1%
15667632	1	< 0.1%
15599024	1	< 0.1%
15798709	1	< 0.1%
15741921	1	< 0.1%
15793671	1	< 0.1%
15797900	1	< 0.1%
15795933	1	< 0.1%
Other values (9990)	9990	99.9%

Value	Count	Frequency (%)
15565701	1	< 0.1%
15565706	1	< 0.1%
15565714	1	< 0.1%
15565779	1	< 0.1%
15565796	1	< 0.1%
15565806	1	< 0.1%
15565878	1	< 0.1%
15565879	1	< 0.1%
15565891	1	< 0.1%
15565996	1	< 0.1%

Value	Count	Frequency (%)
15815690	1	< 0.1%
15815660	1	< 0.1%
15815656	1	< 0.1%
15815645	1	< 0.1%
15815628	1	< 0.1%
15815626	1	< 0.1%
15815615	1	< 0.1%
15815560	1	< 0.1%
15815552	1	< 0.1%
15815534	1	< 0.1%







Surname  
Categorical

Distinct	2932
Distinct (%)	29.3%
Missing	0
Missing (%)	0.0%
Memory size	78.2 KiB

Length

Max length	23
Median length	16
Mean length	6.4349
Min length	2

Characters and Unicode

Total characters	64349	
Distinct characters	55	
Distinct categories	5 ( <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	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> )	?
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	1558	?
Unique (%)	15.6%	

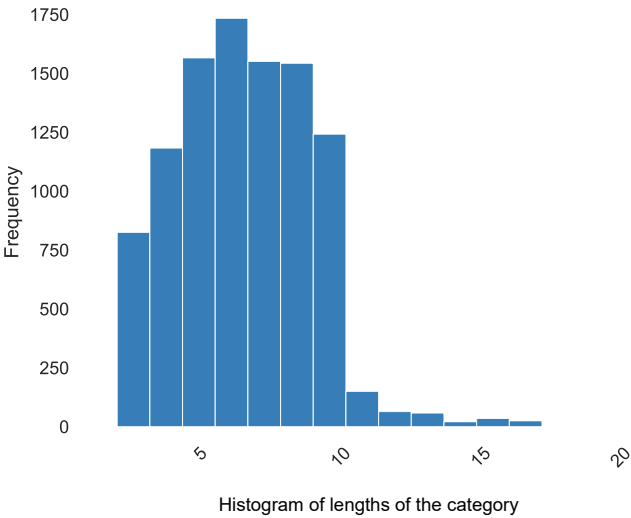
Sample

1st row	Hargrave
2nd row	Hill
3rd row	Onio
4th row	Boni
5th row	Mitchell

Common Values

Value	Count	Frequency (%)
Smith	32	0.3%
Scott	29	0.3%
Martin	29	0.3%
Walker	28	0.3%
Brown	26	0.3%
Yeh	25	0.2%
Shih	25	0.2%
Genovese	25	0.2%
Maclean	24	0.2%
Wright	24	0.2%
Other values (2922)	9733	97.3%

Length



Value	Count	Frequency (%)
lo	33	0.3%
smith	32	0.3%
martin	29	0.3%
scott	29	0.3%
walker	28	0.3%
brown	26	0.3%
yeh	25	0.2%
shih	25	0.2%
genovese	25	0.2%
maclean	24	0.2%
Other values (2931)	9779	97.3%

Most occurring characters

Value	Count	Frequency (%)
a	5799	9.0%
e	5764	9.0%
n	5235	8.1%
o	4905	7.6%
i	4491	7.0%
r	3547	5.5%
l	2921	4.5%
s	2592	4.0%
u	2552	4.0%
h	2150	3.3%
Other values (45)	24393	37.9%

Most occurring categories

Value	Count	Frequency (%)
Lowercase Letter	53647	83.4%
Uppercase Letter	10299	16.0%
Other Punctuation	329	0.5%
Space Separator	55	0.1%
Dash Punctuation	19	< 0.1%

Most frequent character per category

Lowercase Letter

Value	Count	Frequency (%)
a	5799	10.8%
e	5764	10.7%
n	5235	9.8%
o	4905	9.1%
i	4491	8.4%
r	3547	6.6%
l	2921	5.4%
s	2592	4.8%
u	2552	4.8%
h	2150	4.0%
Other values (16)	13691	25.5%

Uppercase Letter

Value	Count	Frequency (%)
C	1106	10.7%
M	1004	9.7%
B	707	6.9%
S	685	6.7%
H	661	6.4%
T	573	5.6%
L	545	5.3%
W	481	4.7%
P	466	4.5%
G	442	4.3%
Other values (15)	3629	35.2%

Other Punctuation

Value	Count	Frequency (%)
,	237	72.0%
?	92	28.0%



Space Separator

Value	Count	Frequency (%)
	55	100.0%

Dash Punctuation

Value	Count	Frequency (%)
-	19	100.0%

Most occurring scripts

Value	Count	Frequency (%)
Latin	63946	99.4%
Common	403	0.6%

Most frequent character per script

Latin

Value	Count	Frequency (%)
a	5799	9.1%
e	5764	9.0%
n	5235	8.2%
o	4905	7.7%
i	4491	7.0%
r	3547	5.5%
l	2921	4.6%
s	2592	4.1%
u	2552	4.0%
h	2150	3.4%
Other values (41)	23990	37.5%

Common

Value	Count	Frequency (%)
,	237	58.8%
?	92	22.8%
	55	13.6%
-	19	4.7%

Most occurring blocks

Value	Count	Frequency (%)
ASCII	64349	100.0%

Most frequent character per block

ASCII

Value	Count	Frequency (%)
a	5799	9.0%
e	5764	9.0%
n	5235	8.1%
o	4905	7.6%
i	4491	7.0%
r	3547	5.5%
l	2921	4.5%
s	2592	4.0%
u	2552	4.0%
h	2150	3.3%
Other values (45)	24393	37.9%

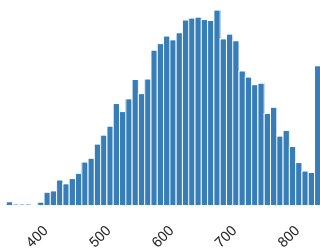




CreditScore

Real number ( $\mathbb{R}$ )

Distinct	460
Distinct (%)	4.6%
Missing	0
Missing (%)	0.0%
Infinite	0
Infinite (%)	0.0%
Mean	650.5288
Minimum	350
Maximum	850
Zeros	0
Zeros (%)	0.0%
Negative	0
Negative (%)	0.0%
Memory size	78.2 KiB

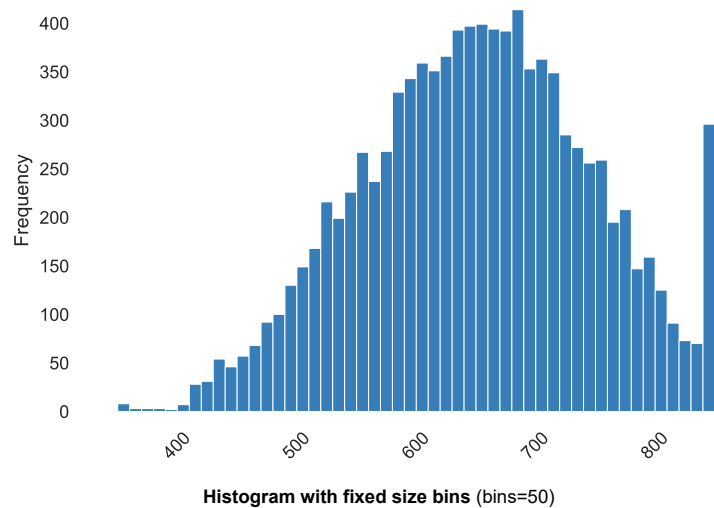


Quantile statistics

Minimum	350
5-th percentile	489
Q1	584
median	652
Q3	718
95-th percentile	812
Maximum	850
Range	500
Interquartile range (IQR)	134

Descriptive statistics

Standard deviation	96.653299
Coefficient of variation (CV)	0.14857651
Kurtosis	-0.42572568
Mean	650.5288
Median Absolute Deviation (MAD)	67
Skewness	-0.071606608
Sum	6505288
Variance	9341.8602
Monotonicity	Not monotonic



Value	Count	Frequency (%)
850	233	2.3%
678	63	0.6%
655	54	0.5%
705	53	0.5%
667	53	0.5%
684	52	0.5%
670	50	0.5%
651	50	0.5%
683	48	0.5%
652	48	0.5%
Other values (450)	9296	93.0%



Value	Count	Frequency (%)
350	5	0.1%
351	1	< 0.1%
358	1	< 0.1%
359	1	< 0.1%
363	1	< 0.1%
365	1	< 0.1%
367	1	< 0.1%
373	1	< 0.1%
376	2	< 0.1%
382	1	< 0.1%

Value	Count	Frequency (%)
850	233	2.3%
849	8	0.1%
848	5	0.1%
847	6	0.1%
846	5	0.1%
845	6	0.1%
844	7	0.1%
843	2	< 0.1%
842	7	0.1%
841	12	0.1%





Geography  
Categorical

Distinct	3
Distinct (%)	< 0.1%
Missing	0
Missing (%)	0.0%
Memory size	78.2 KiB

Length

Max length	7
Median length	6
Mean length	6.0032
Min length	5

Characters and Unicode

Total characters	60032	
Distinct characters	12	
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%	

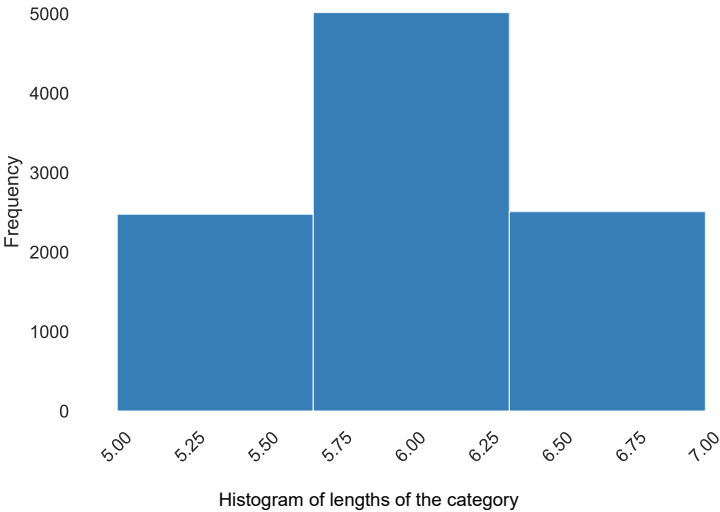
Sample

1st row	France
2nd row	Spain
3rd row	France
4th row	France
5th row	Spain

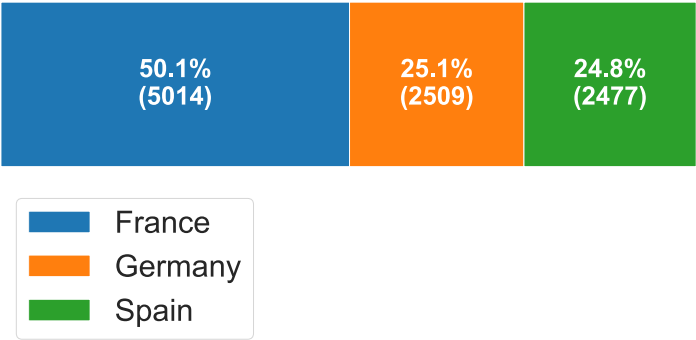
Common Values

Value	Count	Frequency (%)
France	5014	50.1%
Germany	2509	25.1%
Spain	2477	24.8%

Length



Common Values (Plot)



Value	Count	Frequency (%)
france	5014	50.1%
germany	2509	25.1%
spain	2477	24.8%



Most occurring characters

Value	Count	Frequency (%)
a	10000	16.7%
n	10000	16.7%
r	7523	12.5%
e	7523	12.5%
F	5014	8.4%
c	5014	8.4%
G	2509	4.2%
m	2509	4.2%
y	2509	4.2%
S	2477	4.1%
Other values (2)	4954	8.3%

Most occurring categories

Value	Count	Frequency (%)
Lowercase Letter	50032	83.3%
Uppercase Letter	10000	16.7%

Most frequent character per category

Lowercase Letter

Value	Count	Frequency (%)
a	10000	20.0%
n	10000	20.0%
r	7523	15.0%
e	7523	15.0%
c	5014	10.0%
m	2509	5.0%
y	2509	5.0%
p	2477	5.0%
i	2477	5.0%

Uppercase Letter

Value	Count	Frequency (%)
F	5014	50.1%
G	2509	25.1%
S	2477	24.8%

Most occurring scripts

Value	Count	Frequency (%)
Latin	60032	100.0%

Most frequent character per script

Latin

Value	Count	Frequency (%)
a	10000	16.7%
n	10000	16.7%
r	7523	12.5%
e	7523	12.5%
F	5014	8.4%
c	5014	8.4%
G	2509	4.2%
m	2509	4.2%
y	2509	4.2%
S	2477	4.1%
Other values (2)	4954	8.3%

Most occurring blocks

Value	Count	Frequency (%)
ASCII	60032	100.0%

Most frequent character per block

ASCII

Value	Count	Frequency (%)
a	10000	16.7%
n	10000	16.7%
r	7523	12.5%
e	7523	12.5%
F	5014	8.4%
c	5014	8.4%
G	2509	4.2%
m	2509	4.2%
y	2509	4.2%
S	2477	4.1%
Other values (2)	4954	8.3%





Gender

Categorical

Distinct	2
Distinct (%)	< 0.1%
Missing	0
Missing (%)	0.0%
Memory size	78.2 KiB

Length

Max length	6
Median length	4
Mean length	4.9086
Min length	4

Characters and Unicode

Total characters	49086	
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

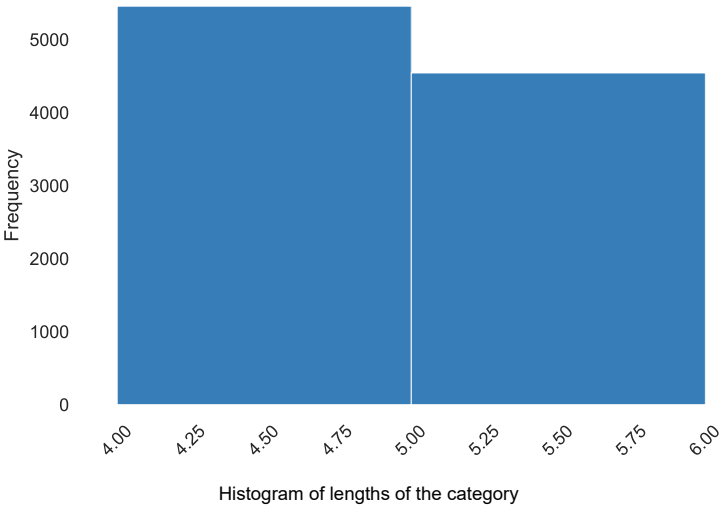
1st row	Female
2nd row	Female
3rd row	Female
4th row	Female
5th row	Female



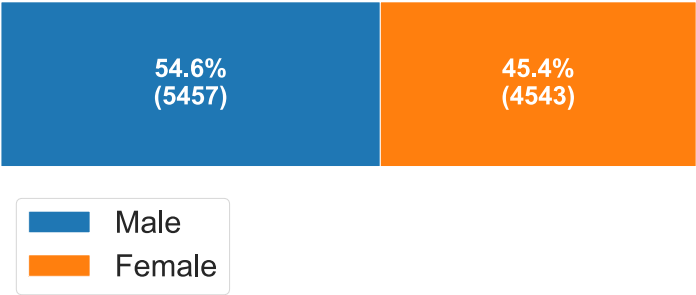
Common Values

Value	Count	Frequency (%)
Male	5457	54.6%
Female	4543	45.4%

Length



Common Values (Plot)



Value	Count	Frequency (%)
male	5457	54.6%
female	4543	45.4%

Most occurring characters

Value	Count	Frequency (%)
e	14543	29.6%
a	10000	20.4%
l	10000	20.4%
M	5457	11.1%
F	4543	9.3%
m	4543	9.3%

Most occurring categories

Value	Count	Frequency (%)
Lowercase Letter	39086	79.6%
Uppercase Letter	10000	20.4%

Most frequent character per category

Lowercase Letter

Value	Count	Frequency (%)
e	14543	37.2%
a	10000	25.6%
l	10000	25.6%
m	4543	11.6%

Uppercase Letter

Value	Count	Frequency (%)
M	5457	54.6%
F	4543	45.4%

Most occurring scripts

Value	Count	Frequency (%)
Latin	49086	100.0%

Most frequent character per script

Latin

Value	Count	Frequency (%)
e	14543	29.6%
a	10000	20.4%
l	10000	20.4%
M	5457	11.1%
F	4543	9.3%
m	4543	9.3%

Most occurring blocks

Value	Count	Frequency (%)
ASCII	49086	100.0%

Most frequent character per block

ASCII

Value	Count	Frequency (%)
e	14543	29.6%
a	10000	20.4%
l	10000	20.4%
M	5457	11.1%
F	4543	9.3%
m	4543	9.3%



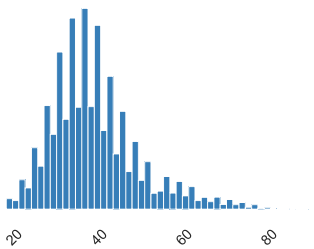




Age

Real number ( $\mathbb{R}$ )

Distinct	70
Distinct (%)	0.7%
Missing	0
Missing (%)	0.0%
Infinite	0
Infinite (%)	0.0%
Mean	38.9218
Minimum	18
Maximum	92
Zeros	0
Zeros (%)	0.0%
Negative	0
Negative (%)	0.0%
Memory size	78.2 KiB

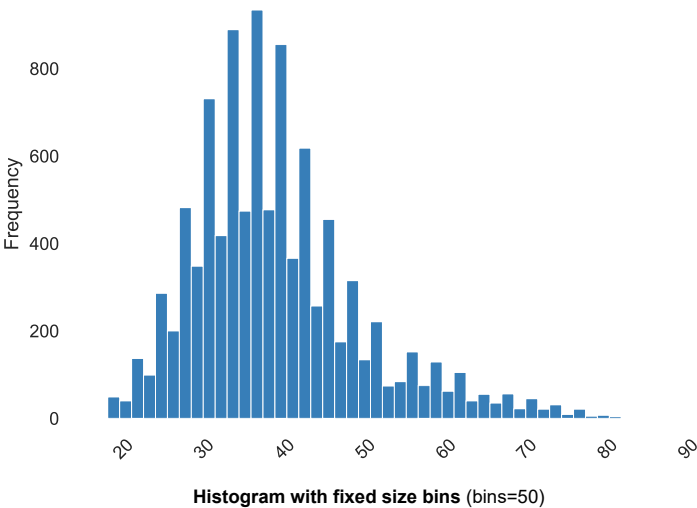


Quantile statistics

Minimum	18
5-th percentile	25
Q1	32
median	37
Q3	44
95-th percentile	60
Maximum	92
Range	74
Interquartile range (IQR)	12

Descriptive statistics

Standard deviation	10.487806
Coefficient of variation (CV)	0.26945841
Kurtosis	1.3953471
Mean	38.9218
Median Absolute Deviation (MAD)	6
Skewness	1.0113203
Sum	389218
Variance	109.99408
Monotonicity	Not monotonic



Value	Count	Frequency (%)
37	478	4.8%
38	477	4.8%
35	474	4.7%
36	456	4.6%
34	447	4.5%
33	442	4.4%
40	432	4.3%
39	423	4.2%
32	418	4.2%
31	404	4.0%
Other values (60)	5549	55.5%

Value	Count	Frequency (%)
18	22	0.2%
19	27	0.3%
20	40	0.4%
21	53	0.5%
22	84	0.8%
23	99	1.0%
24	132	1.3%
25	154	1.5%
26	200	2.0%
27	209	2.1%

Value	Count	Frequency (%)
92	2	< 0.1%
88	1	< 0.1%
85	1	< 0.1%
84	2	< 0.1%
83	1	< 0.1%
82	1	< 0.1%
81	4	< 0.1%
80	3	< 0.1%
79	4	< 0.1%
78	5	0.1%

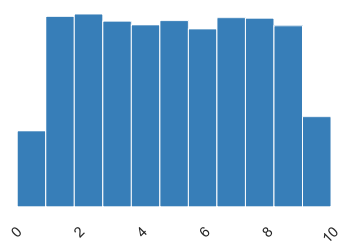




Tenure

Real number ( $\mathbb{R}$ )

Distinct	11
Distinct (%)	0.1%
Missing	0
Missing (%)	0.0%
Infinite	0
Infinite (%)	0.0%
Mean	5.0128
Minimum	0
Maximum	10
Zeros	413
Zeros (%)	4.1%
Negative	0
Negative (%)	0.0%
Memory size	78.2 KiB



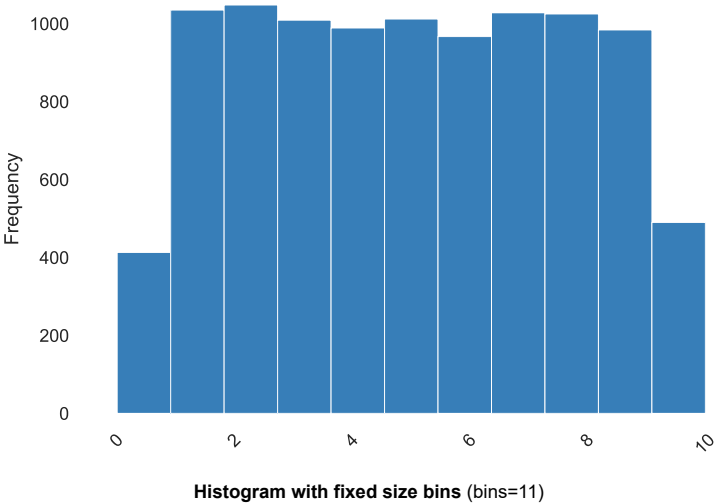
Quantile statistics

Minimum	0
5-th percentile	1
Q1	3
median	5
Q3	7
95-th percentile	9
Maximum	10
Range	10
Interquartile range (IQR)	4

Descriptive statistics

Standard deviation	2.8921744
Coefficient of variation (CV)	0.57695786
Kurtosis	-1.1652252
Mean	5.0128
Median Absolute Deviation (MAD)	2
Skewness	0.010991458
Sum	50128
Variance	8.3646726
Monotonicity	Not monotonic





Value	Count	Frequency (%)
2	1048	10.5%
1	1035	10.3%
7	1028	10.3%
8	1025	10.2%
5	1012	10.1%
3	1009	10.1%
4	989	9.9%
9	984	9.8%
6	967	9.7%
10	490	4.9%

Value	Count	Frequency (%)
0	413	4.1%
1	1035	10.3%
2	1048	10.5%
3	1009	10.1%
4	989	9.9%
5	1012	10.1%
6	967	9.7%
7	1028	10.3%
8	1025	10.2%
9	984	9.8%

Value	Count	Frequency (%)
10	490	4.9%
9	984	9.8%
8	1025	10.2%
7	1028	10.3%
6	967	9.7%
5	1012	10.1%
4	989	9.9%
3	1009	10.1%
2	1048	10.5%
1	1035	10.3%

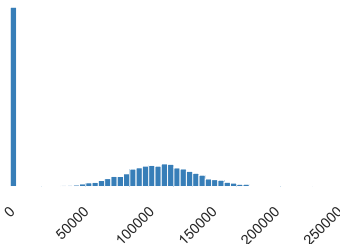




Balance

Real number (ℝ)

Distinct	6382
Distinct (%)	63.8%
Missing	0
Missing (%)	0.0%
Infinite	0
Infinite (%)	0.0%
Mean	76485.889
Minimum	0
Maximum	250898.09
Zeros	3617
Zeros (%)	36.2%
Negative	0
Negative (%)	0.0%
Memory size	78.2 KiB

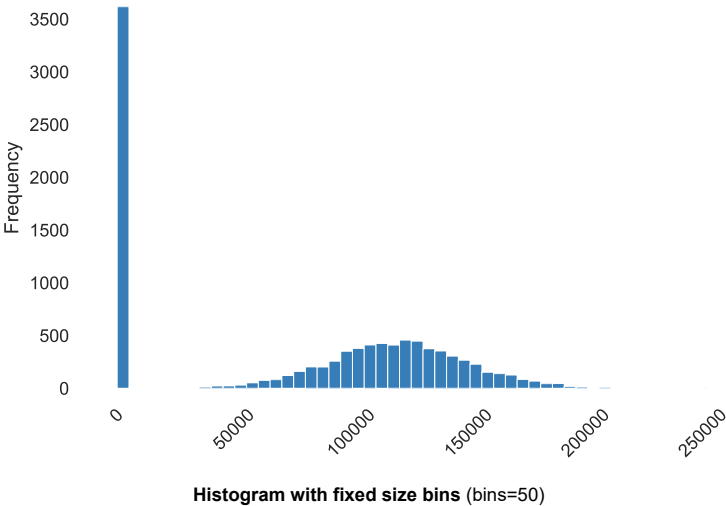


Quantile statistics

Minimum	0
5-th percentile	0
Q1	0
median	97198.54
Q3	127644.24
95-th percentile	162711.67
Maximum	250898.09
Range	250898.09
Interquartile range (IQR)	127644.24

Descriptive statistics

Standard deviation	62397.405
Coefficient of variation (CV)	0.81580283
Kurtosis	-1.4894118
Mean	76485.889
Median Absolute Deviation (MAD)	46766.79
Skewness	-0.14110871
Sum	$7.6485889 \times 10^8$
Variance	$3.8934362 \times 10^9$
Monotonicity	Not monotonic





Value	Count	Frequency (%)
0	3617	36.2%
130170.82	2	< 0.1%
105473.74	2	< 0.1%
85304.27	1	< 0.1%
159397.75	1	< 0.1%
144238.7	1	< 0.1%
112262.84	1	< 0.1%
109106.8	1	< 0.1%
142147.32	1	< 0.1%
109109.33	1	< 0.1%
Other values (6372)	6372	63.7%

Value	Count	Frequency (%)
0	3617	36.2%
3768.69	1	< 0.1%
12459.19	1	< 0.1%
14262.8	1	< 0.1%
16893.59	1	< 0.1%
23503.31	1	< 0.1%
24043.45	1	< 0.1%
27288.43	1	< 0.1%
27517.15	1	< 0.1%
27755.97	1	< 0.1%

Value	Count	Frequency (%)
250898.09	1	< 0.1%
238387.56	1	< 0.1%
222267.63	1	< 0.1%
221532.8	1	< 0.1%
216109.88	1	< 0.1%
214346.96	1	< 0.1%
213146.2	1	< 0.1%
212778.2	1	< 0.1%
212696.32	1	< 0.1%
212692.97	1	< 0.1%





NumOfProducts  
Categorical

Distinct	4
Distinct (%)	< 0.1%
Missing	0
Missing (%)	0.0%
Memory size	78.2 KiB

Length

Max length	1
Median length	1
Mean length	1
Min length	1

Characters and Unicode

Total characters	10000	
Distinct characters	4	
Distinct categories	1 ( <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%	

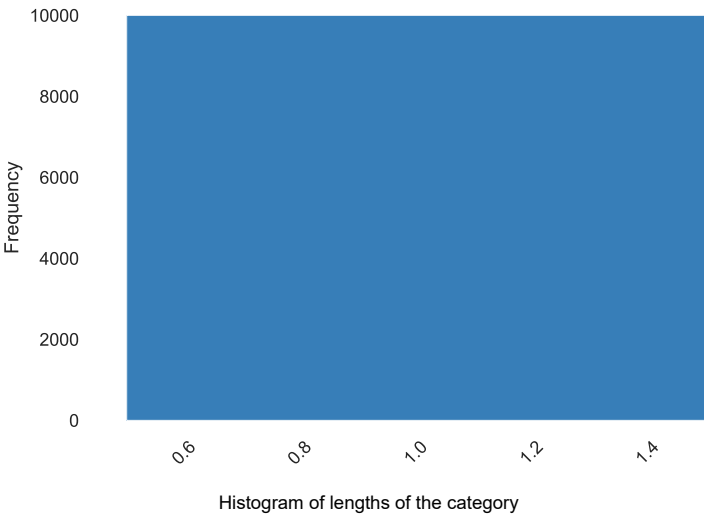
Sample

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

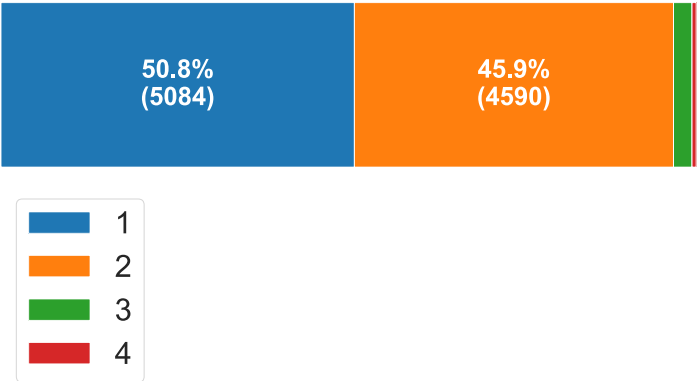
Common Values

Value	Count	Frequency (%)
1	5084	50.8%
2	4590	45.9%
3	266	2.7%
4	60	0.6%

Length



Common Values (Plot)





Value	Count	Frequency (%)
1	5084	50.8%
2	4590	45.9%
3	266	2.7%
4	60	0.6%

Most occurring characters

Value	Count	Frequency (%)
1	5084	50.8%
2	4590	45.9%
3	266	2.7%
4	60	0.6%

Most occurring categories

Value	Count	Frequency (%)
Decimal Number	10000	100.0%

Most frequent character per category

Decimal Number

Value	Count	Frequency (%)
1	5084	50.8%
2	4590	45.9%
3	266	2.7%
4	60	0.6%

Most occurring scripts

Value	Count	Frequency (%)
Common	10000	100.0%

Most frequent character per script

Common

Value	Count	Frequency (%)
1	5084	50.8%
2	4590	45.9%
3	266	2.7%
4	60	0.6%

Most occurring blocks

Value	Count	Frequency (%)
ASCII	10000	100.0%

Most frequent character per block

ASCII

Value	Count	Frequency (%)
1	5084	50.8%
2	4590	45.9%
3	266	2.7%
4	60	0.6%





HasCrCard  
Categorical

Distinct	2
Distinct (%)	< 0.1%
Missing	0
Missing (%)	0.0%
Memory size	78.2 KiB



Length

Max length	1
Median length	1
Mean length	1
Min length	1

Characters and Unicode

Total characters	10000	
Distinct characters	2	
Distinct categories	1 ( <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%	

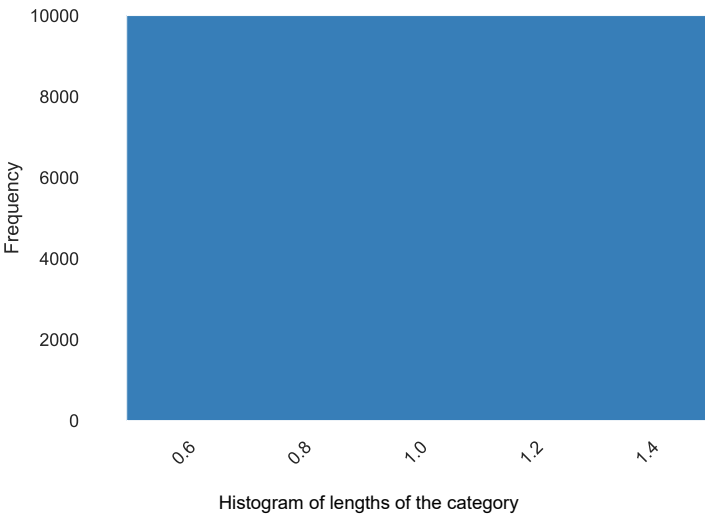
Sample

1st row	1
2nd row	0
3rd row	1
4th row	0
5th row	1

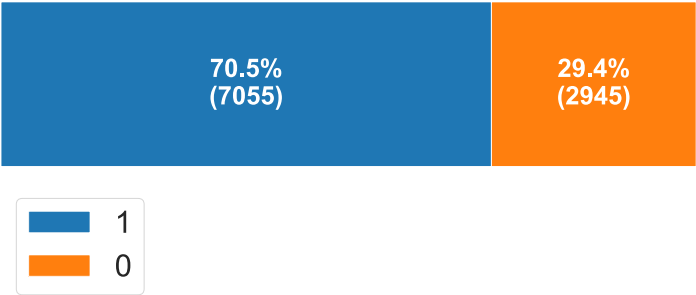
Common Values

Value	Count	Frequency (%)
1	7055	70.5%
0	2945	29.4%

Length



Common Values (Plot)



Value	Count	Frequency (%)
1	7055	70.5%
0	2945	29.4%

Most occurring characters

Value	Count	Frequency (%)
1	7055	70.5%
0	2945	29.4%

Most occurring categories

Value	Count	Frequency (%)
Decimal Number	10000	100.0%

Most frequent character per category

Decimal Number

Value	Count	Frequency (%)
1	7055	70.5%
0	2945	29.4%

Most occurring scripts

Value	Count	Frequency (%)
Common	10000	100.0%

Most frequent character per script

Common

Value	Count	Frequency (%)
1	7055	70.5%
0	2945	29.4%

Most occurring blocks

Value	Count	Frequency (%)
ASCII	10000	100.0%

Most frequent character per block

ASCII

Value	Count	Frequency (%)
1	7055	70.5%
0	2945	29.4%







IsActiveMember  
Categorical

Distinct	2
Distinct (%)	< 0.1%
Missing	0
Missing (%)	0.0%
Memory size	78.2 KiB

Length

Max length	1
Median length	1
Mean length	1
Min length	1

Characters and Unicode

Total characters	10000	
Distinct characters	2	
Distinct categories	1 ( <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%	

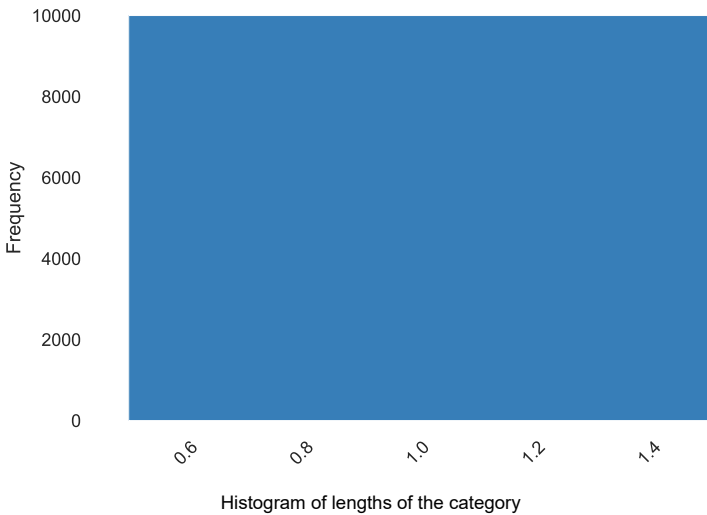
Sample

1st row	1
2nd row	1
3rd row	0
4th row	0
5th row	1

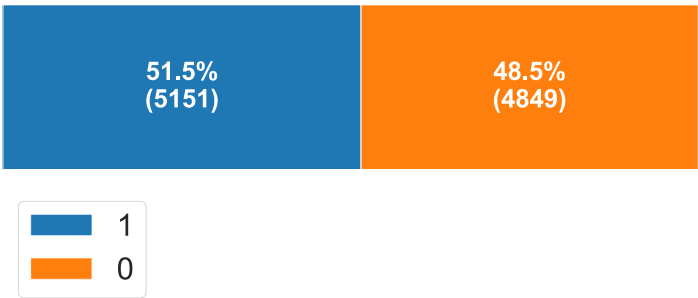
Common Values

Value	Count	Frequency (%)
1	5151	51.5%
0	4849	48.5%

Length



Common Values (Plot)



Value	Count	Frequency (%)
1	5151	51.5%
0	4849	48.5%

Most occurring characters

Value	Count	Frequency (%)
1	5151	51.5%
0	4849	48.5%

Most occurring categories

Value	Count	Frequency (%)
Decimal Number	10000	100.0%

Most frequent character per category

Decimal Number

Value	Count	Frequency (%)
1	5151	51.5%
0	4849	48.5%

Most occurring scripts

Value	Count	Frequency (%)
Common	10000	100.0%

Most frequent character per script

Common

Value	Count	Frequency (%)
1	5151	51.5%
0	4849	48.5%



Most occurring blocks

Value	Count	Frequency (%)
ASCII	10000	100.0%

Most frequent character per block

ASCII

Value	Count	Frequency (%)
1	5151	51.5%
0	4849	48.5%

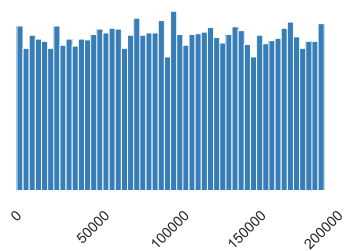




EstimatedSalary

Real number (ℝ)

Distinct	9999
Distinct (%)	> 99.9%
Missing	0
Missing (%)	0.0%
Infinite	0
Infinite (%)	0.0%
Mean	100090.24
Minimum	11.58
Maximum	199992.48
Zeros	0
Zeros (%)	0.0%
Negative	0
Negative (%)	0.0%
Memory size	78.2 KiB

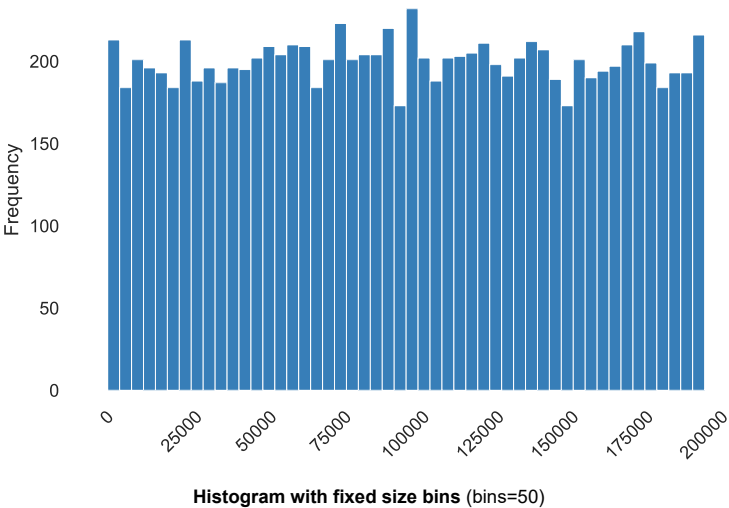


Quantile statistics

Minimum	11.58
5-th percentile	9851.8185
Q1	51002.11
median	100193.91
Q3	149388.25
95-th percentile	190155.38
Maximum	199992.48
Range	199980.9
Interquartile range (IQR)	98386.137

Descriptive statistics

Standard deviation	57510.493
Coefficient of variation (CV)	0.57458642
Kurtosis	-1.1815184
Mean	100090.24
Median Absolute Deviation (MAD)	49198.15
Skewness	0.0020853577
Sum	$1.0009024 \times 10^9$
Variance	$3.3074568 \times 10^9$
Monotonicity	Not monotonic



Value	Count	Frequency (%)
24924.92	2	< 0.1%
101348.88	1	< 0.1%
55313.44	1	< 0.1%
72500.68	1	< 0.1%
182692.8	1	< 0.1%
4993.94	1	< 0.1%
124964.82	1	< 0.1%
161971.42	1	< 0.1%
39488.04	1	< 0.1%
187811.71	1	< 0.1%
Other values (9989)	9989	99.9%

Value	Count	Frequency (%)
11.58	1	< 0.1%
90.07	1	< 0.1%
91.75	1	< 0.1%
96.27	1	< 0.1%
106.67	1	< 0.1%
123.07	1	< 0.1%
142.81	1	< 0.1%
143.34	1	< 0.1%
178.19	1	< 0.1%
216.27	1	< 0.1%

Value	Count	Frequency (%)
199992.48	1	< 0.1%
199970.74	1	< 0.1%
199953.33	1	< 0.1%
199929.17	1	< 0.1%
199909.32	1	< 0.1%
199862.75	1	< 0.1%
199857.47	1	< 0.1%
199841.32	1	< 0.1%
199808.1	1	< 0.1%
199805.63	1	< 0.1%







Exited  
Categorical

Distinct	2
Distinct (%)	< 0.1%
Missing	0
Missing (%)	0.0%
Memory size	78.2 KiB

Length

Max length	1
Median length	1
Mean length	1
Min length	1

Characters and Unicode

Total characters	10000	
Distinct characters	2	
Distinct categories	1 ( <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%	

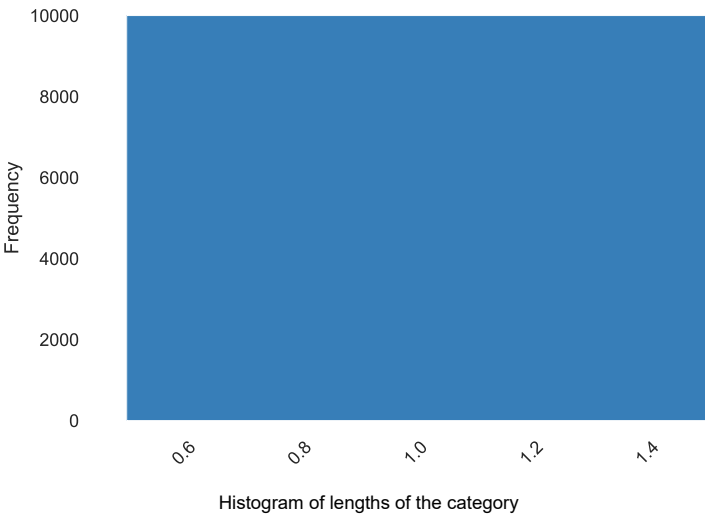
Sample

1st row	1
2nd row	0
3rd row	1
4th row	0
5th row	0

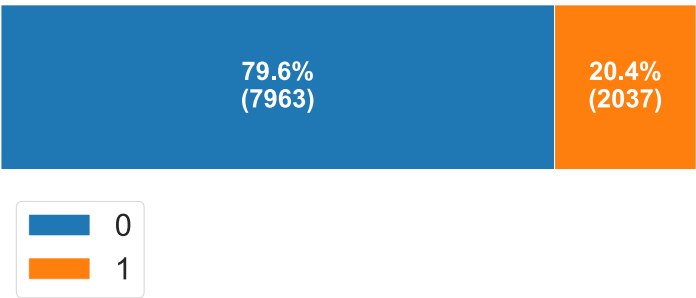
Common Values

Value	Count	Frequency (%)
0	7963	79.6%
1	2037	20.4%

Length



Common Values (Plot)



Value	Count	Frequency (%)
0	7963	79.6%
1	2037	20.4%

Most occurring characters

Value	Count	Frequency (%)
0	7963	79.6%
1	2037	20.4%

Most occurring categories

Value	Count	Frequency (%)
Decimal Number	10000	100.0%

Most frequent character per category

Decimal Number

Value	Count	Frequency (%)
0	7963	79.6%
1	2037	20.4%



Most occurring scripts

Value	Count	Frequency (%)
Common	10000	100.0%

Most frequent character per script

Common

Value	Count	Frequency (%)
0	7963	79.6%
1	2037	20.4%

Most occurring blocks

Value	Count	Frequency (%)
ASCII	10000	100.0%

Most frequent character per block

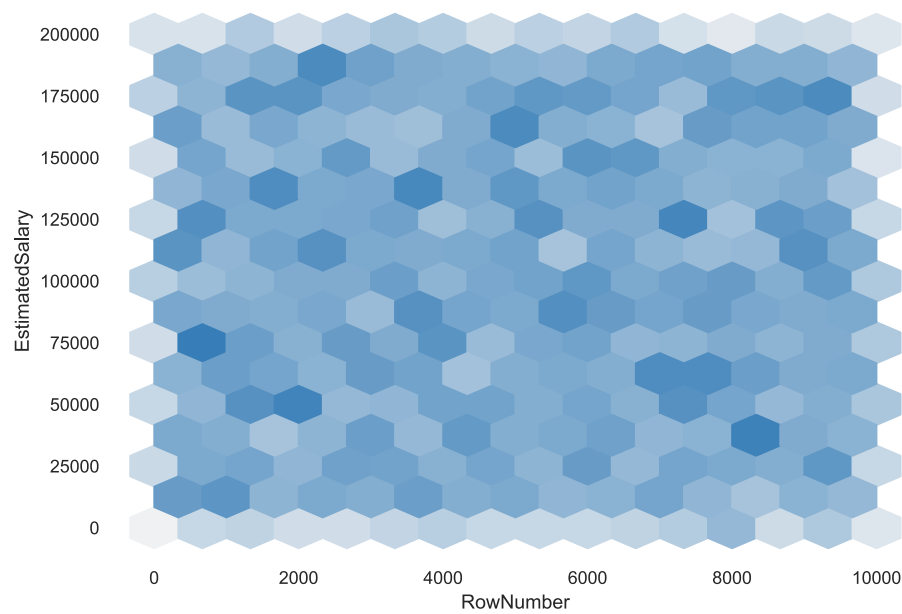
ASCII

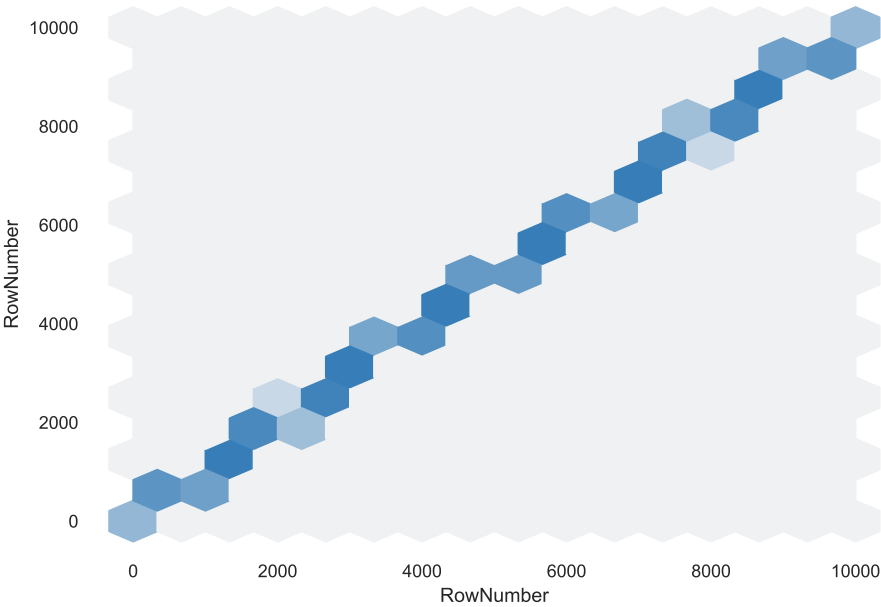
Value	Count	Frequency (%)
0	7963	79.6%
1	2037	20.4%

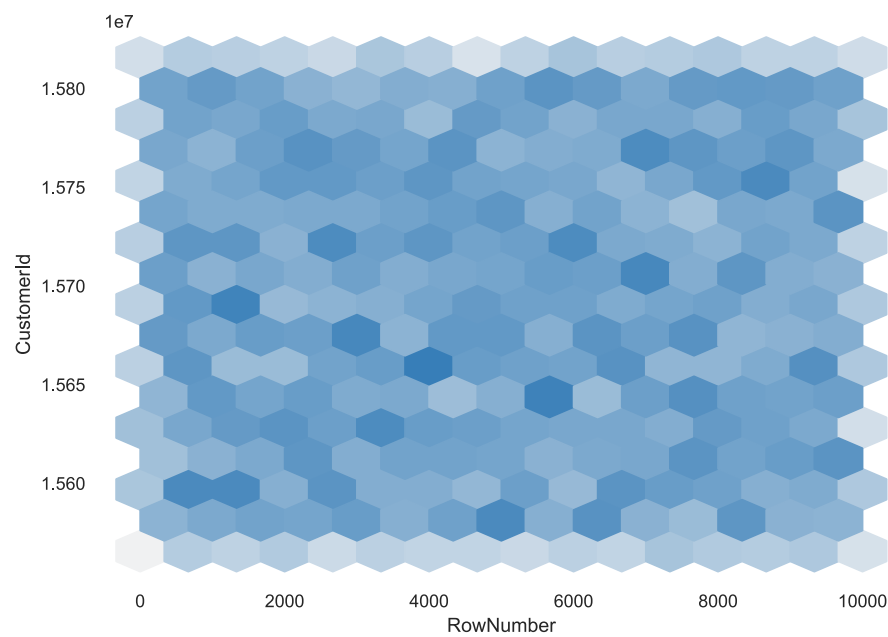


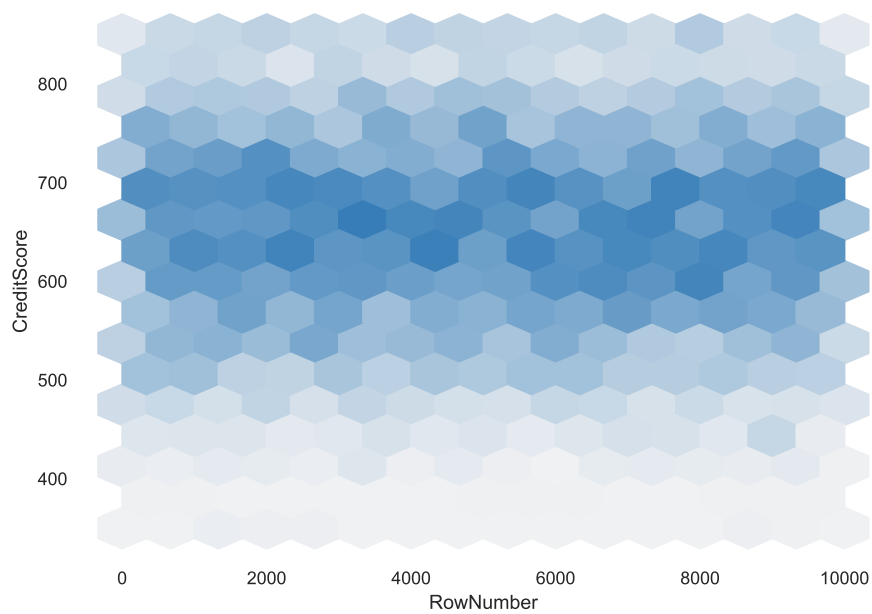


# Interactions

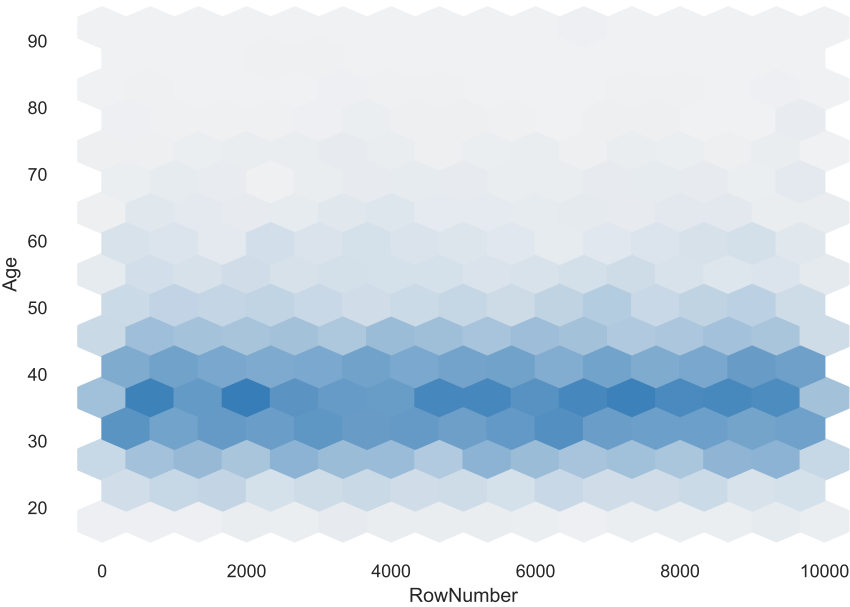


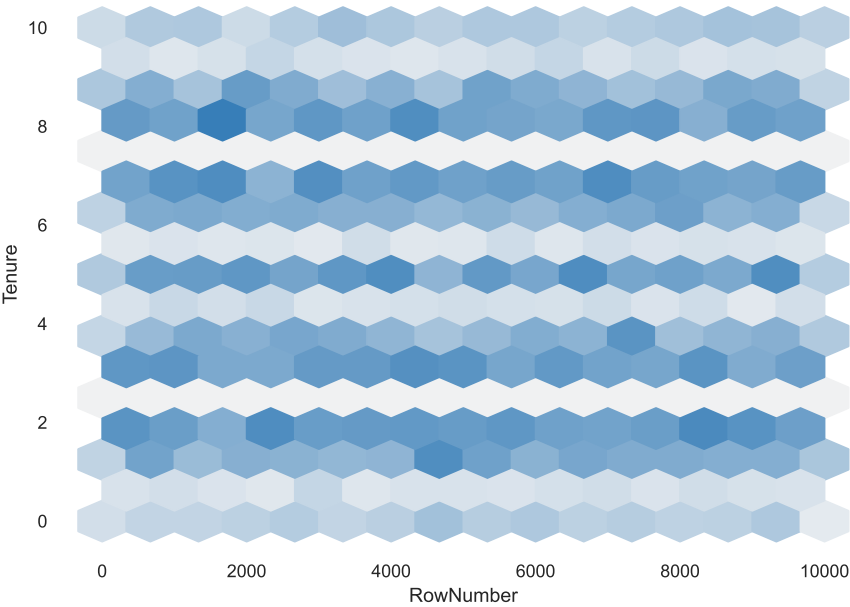


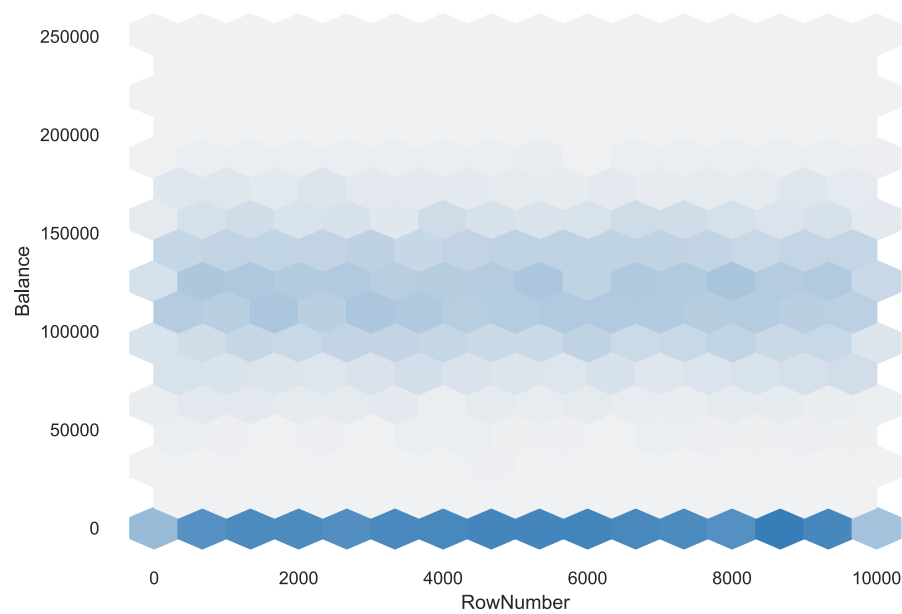




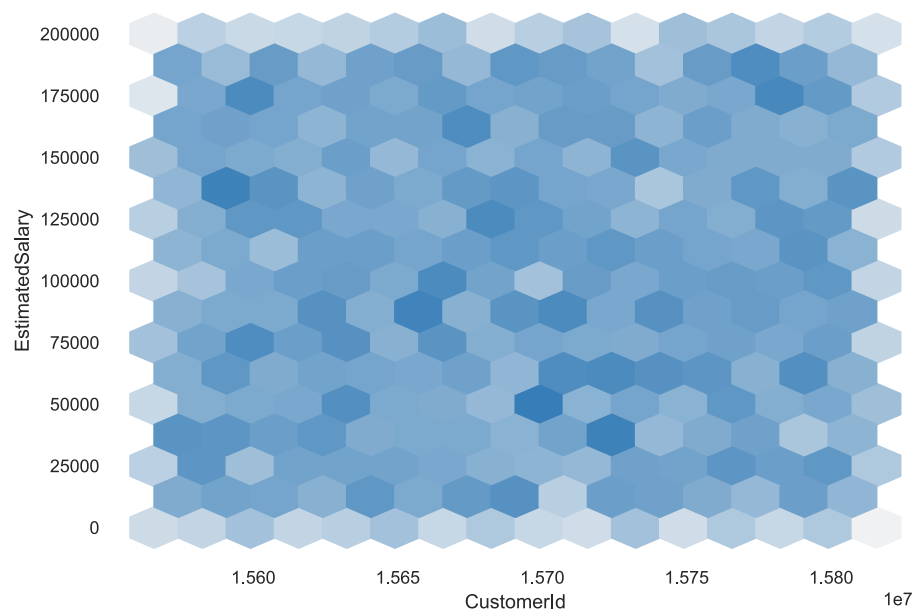


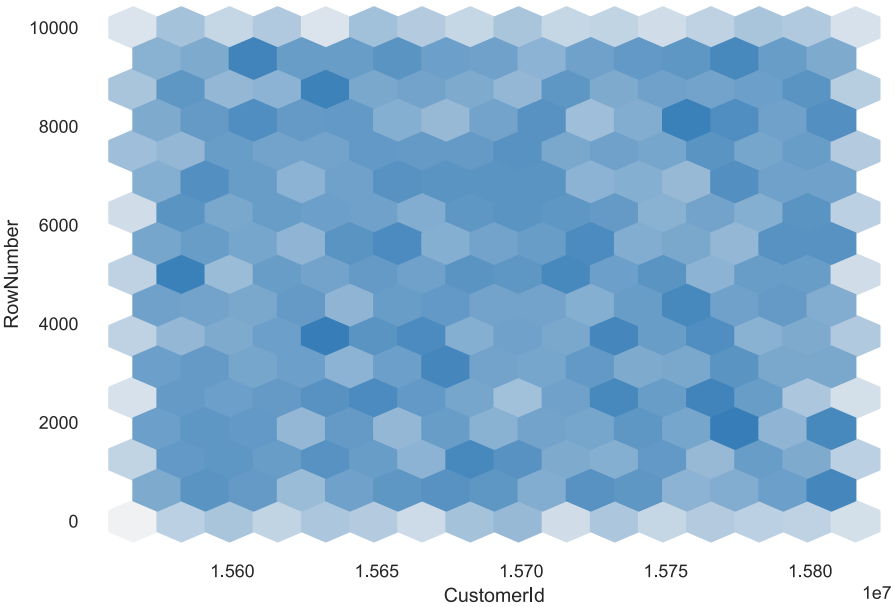


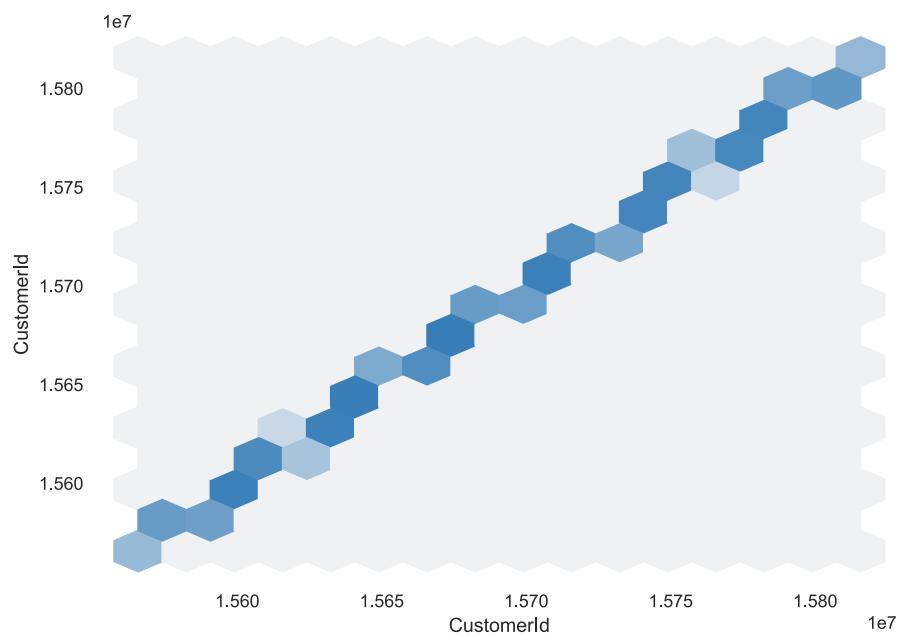


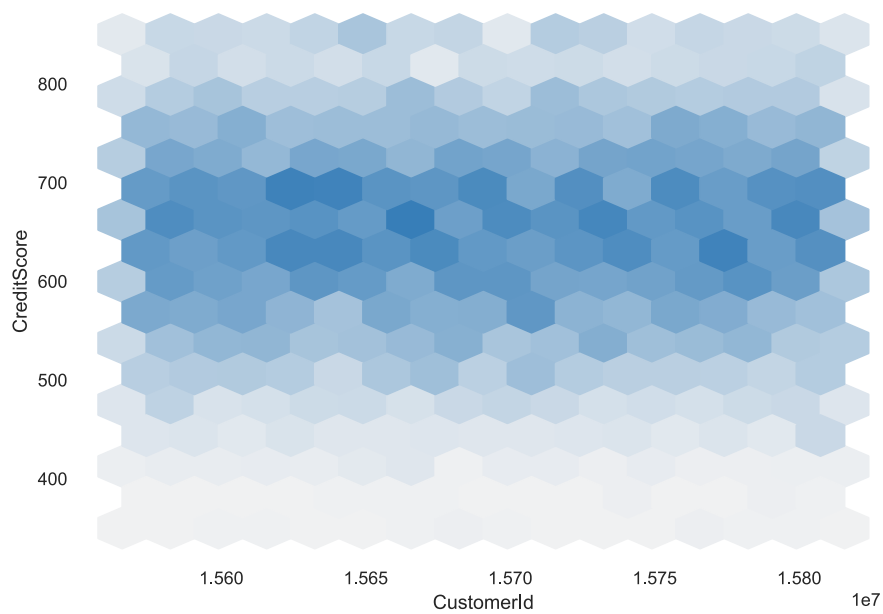




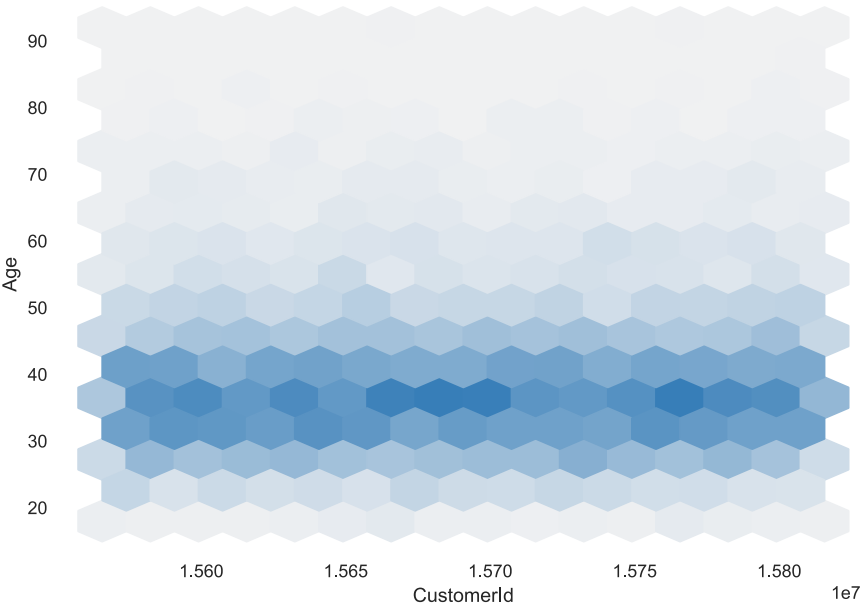


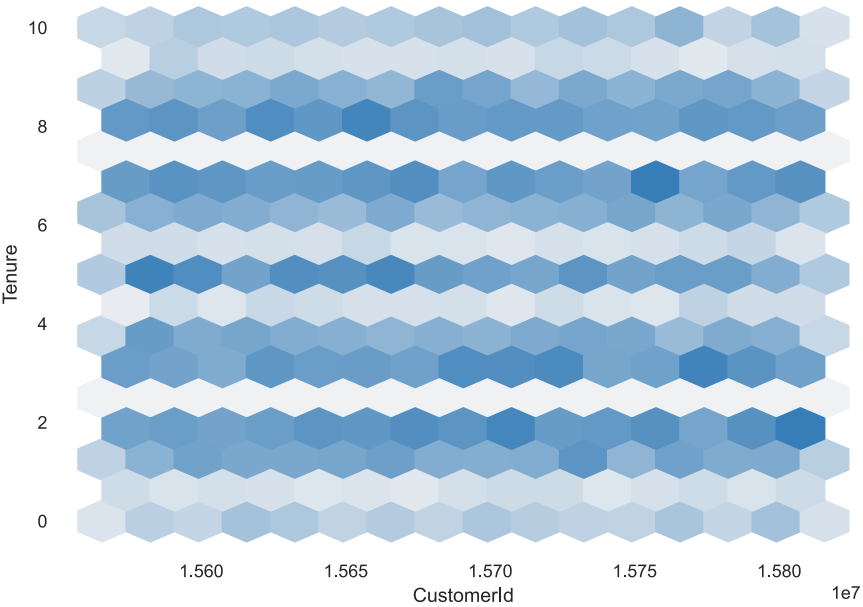


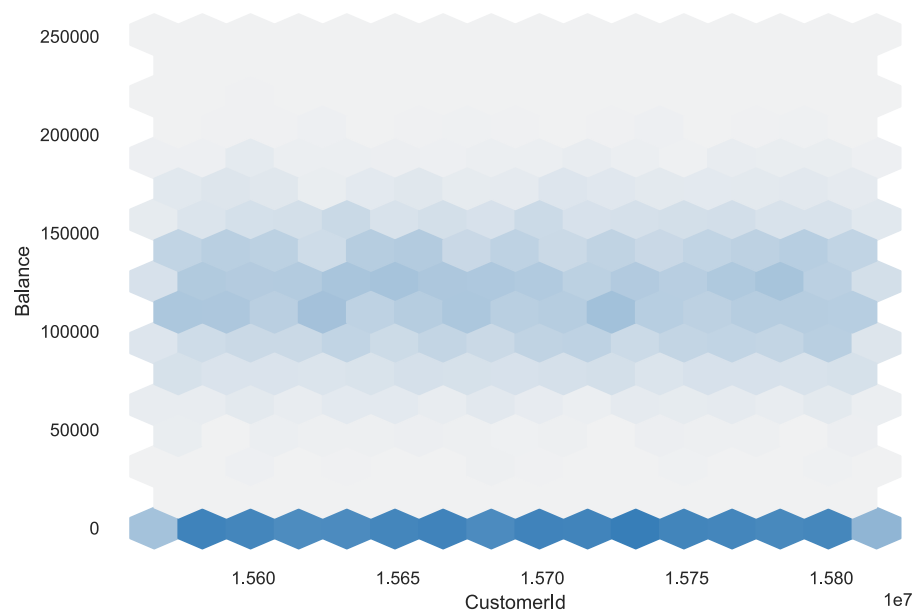




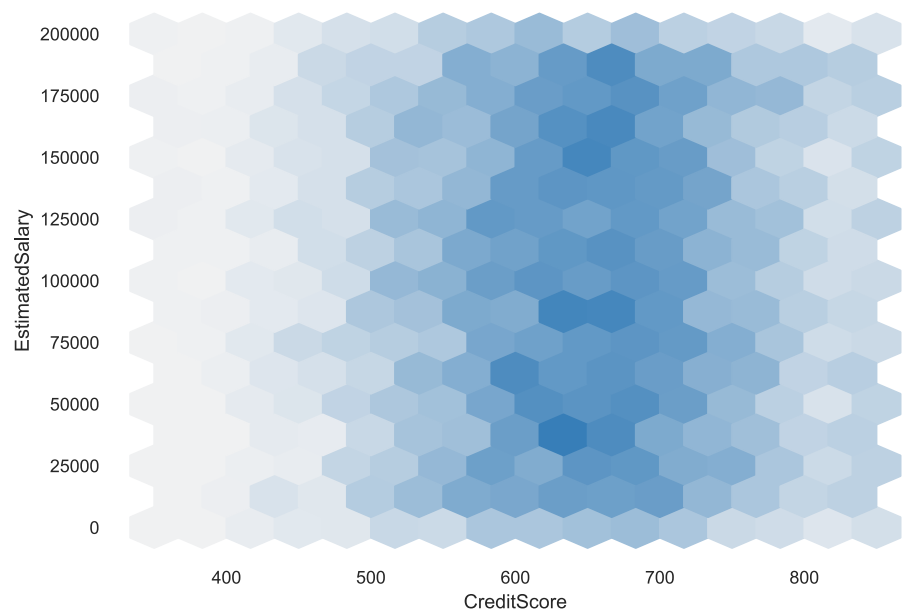


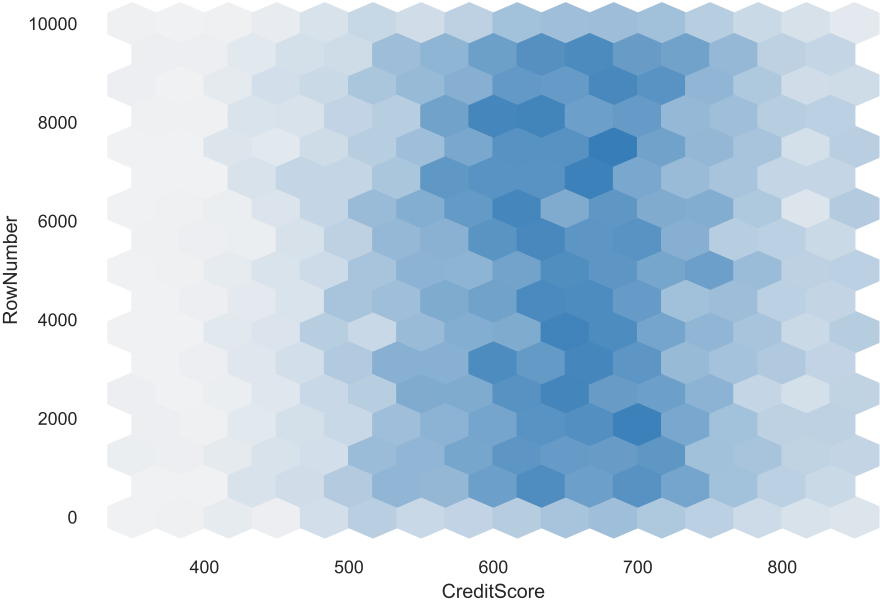


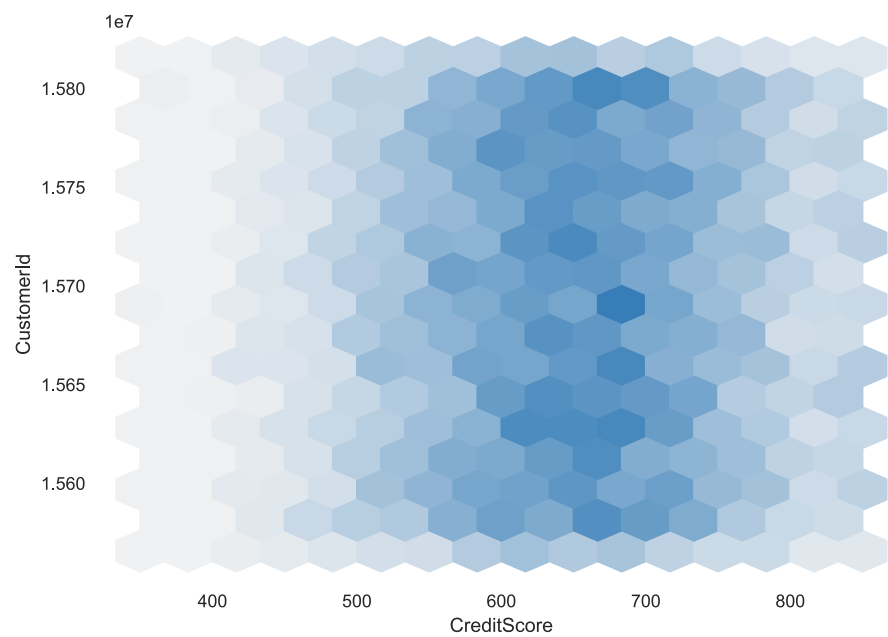


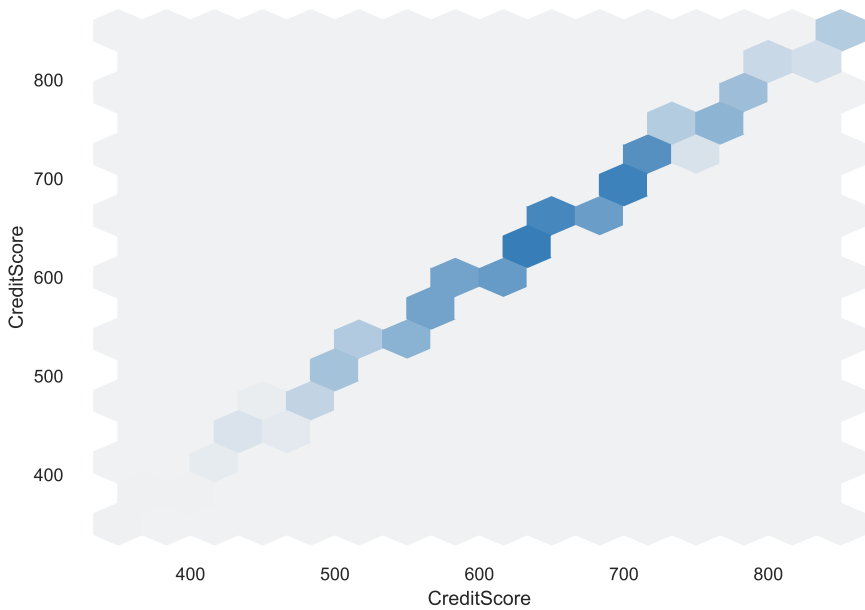




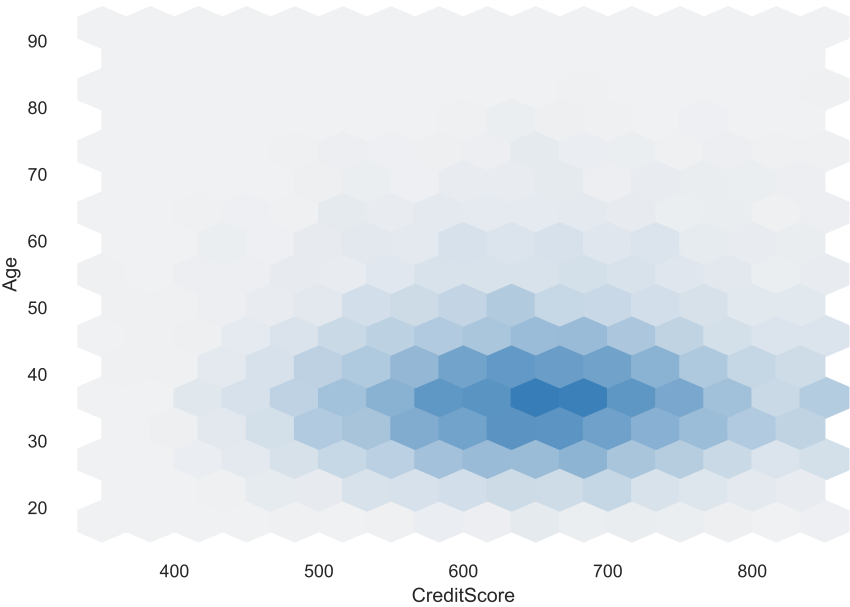


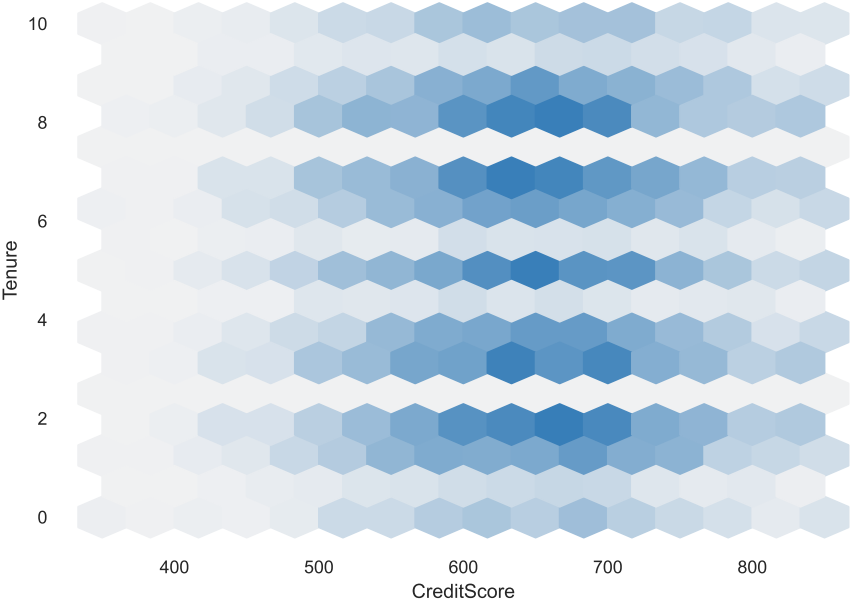


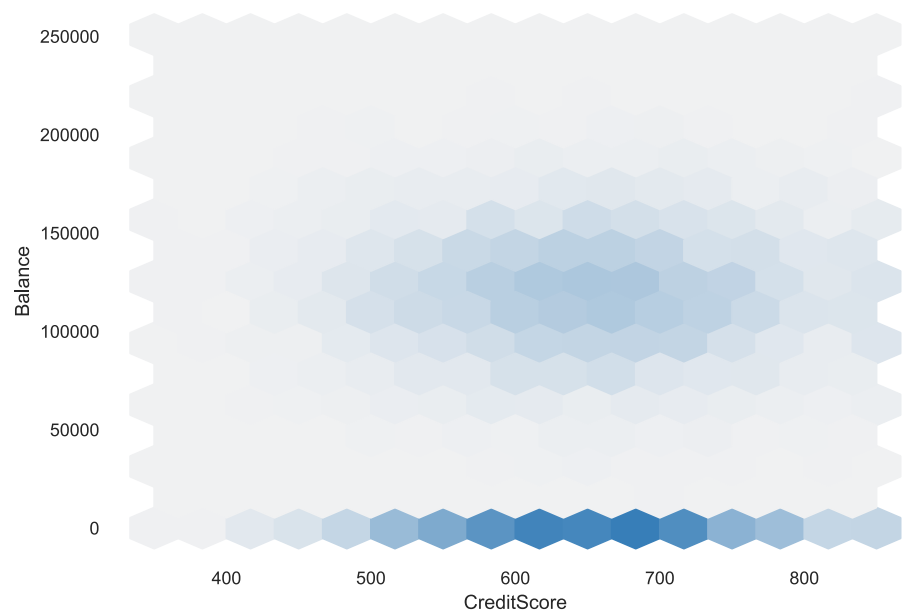




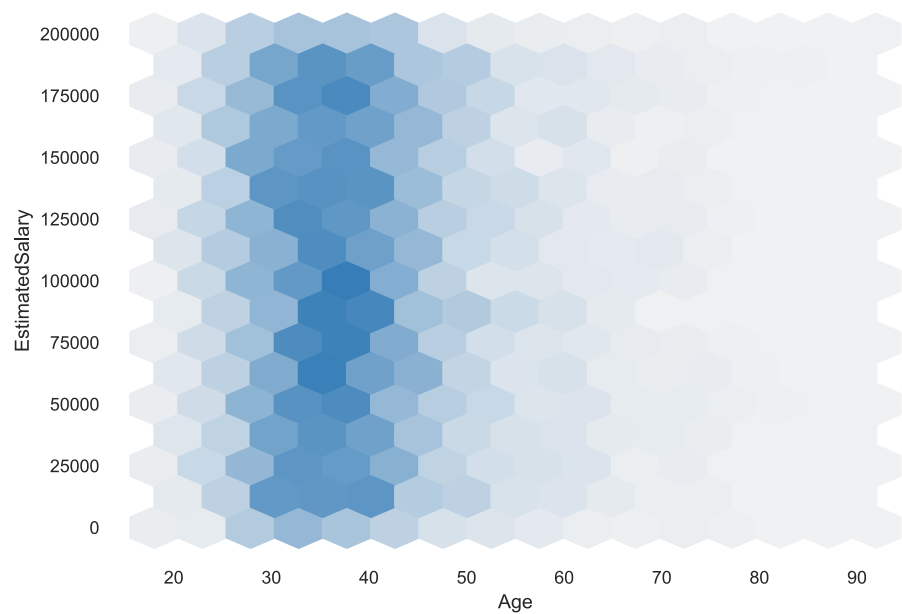


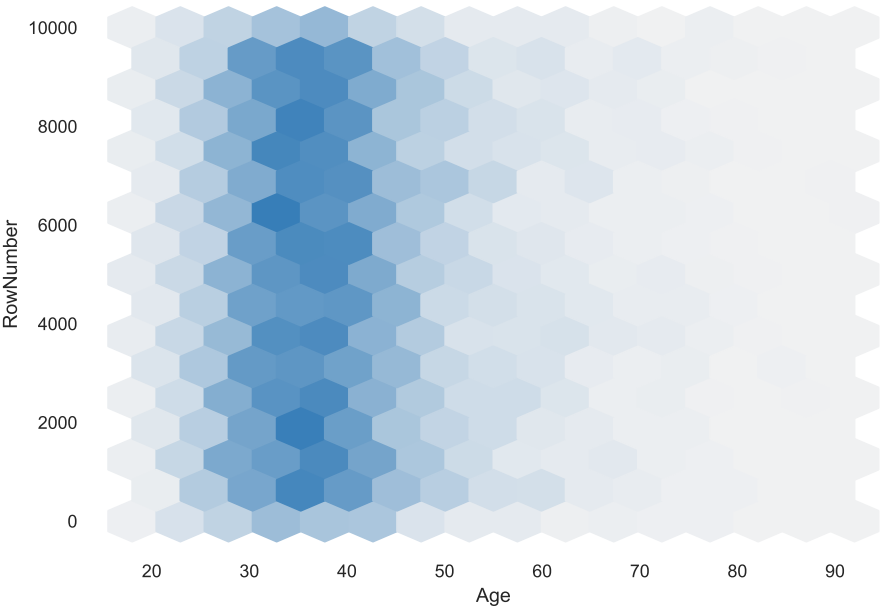


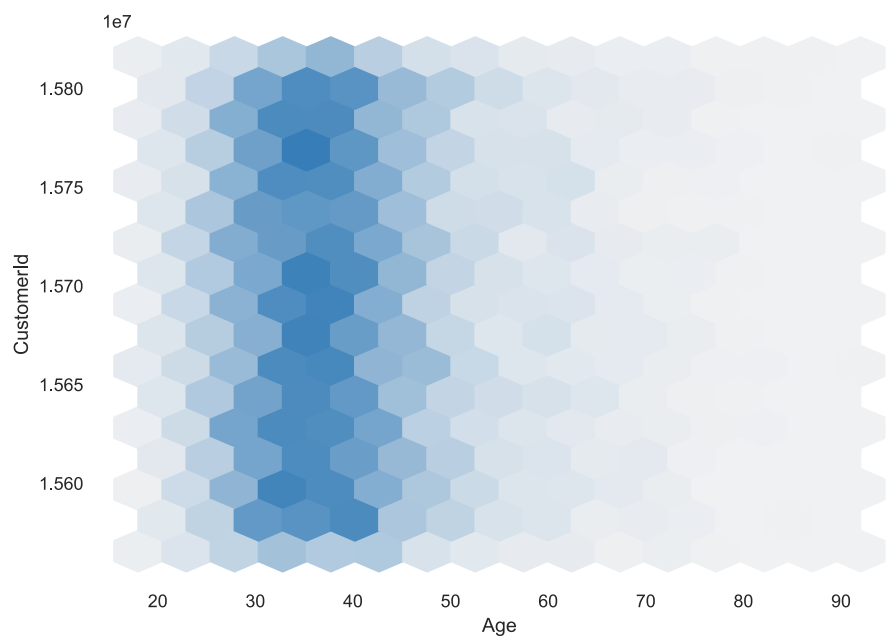


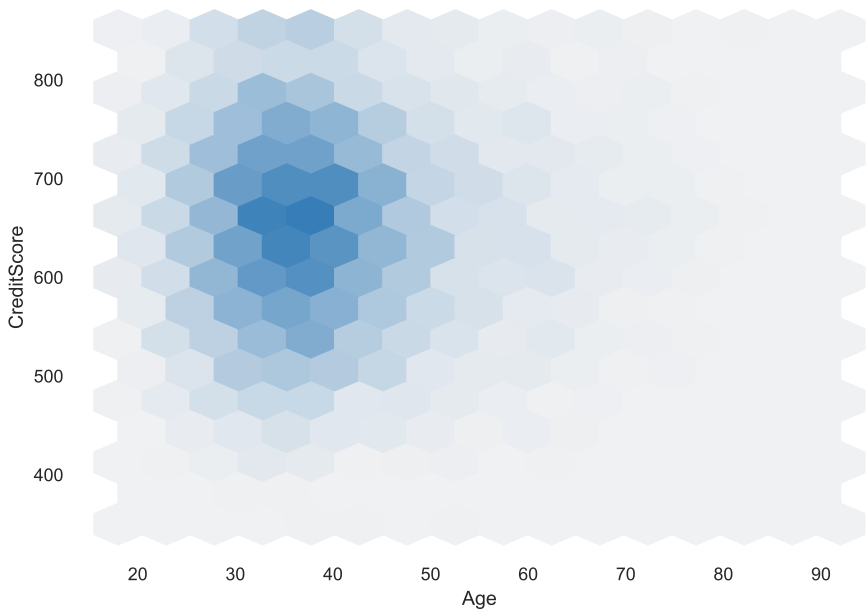




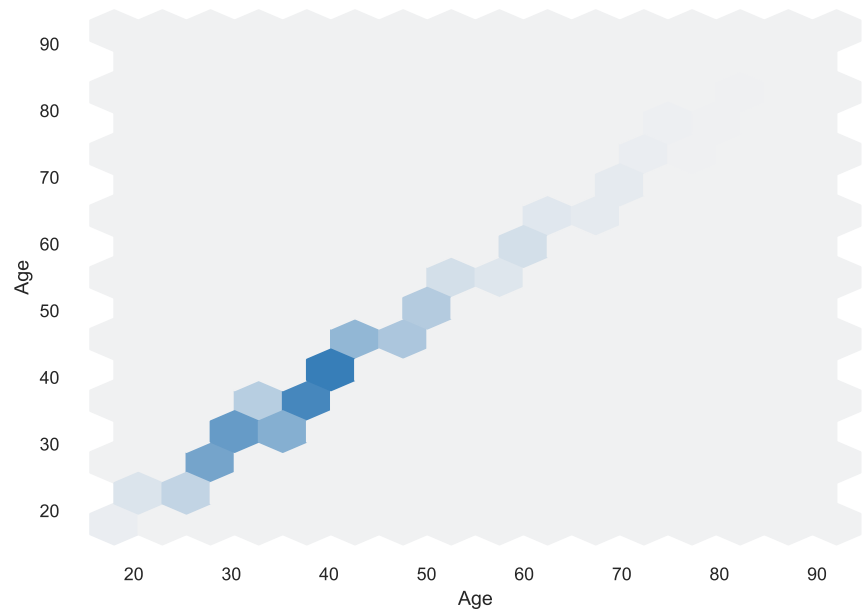


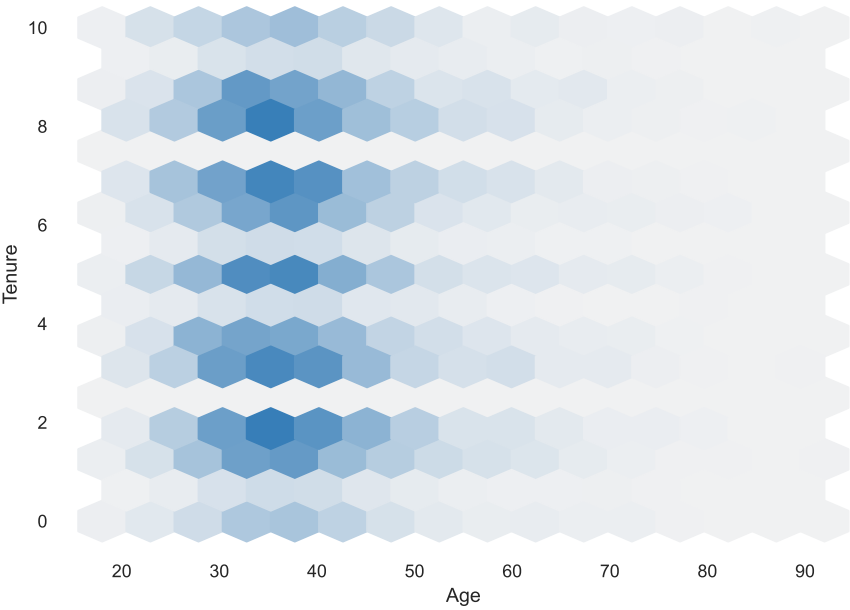


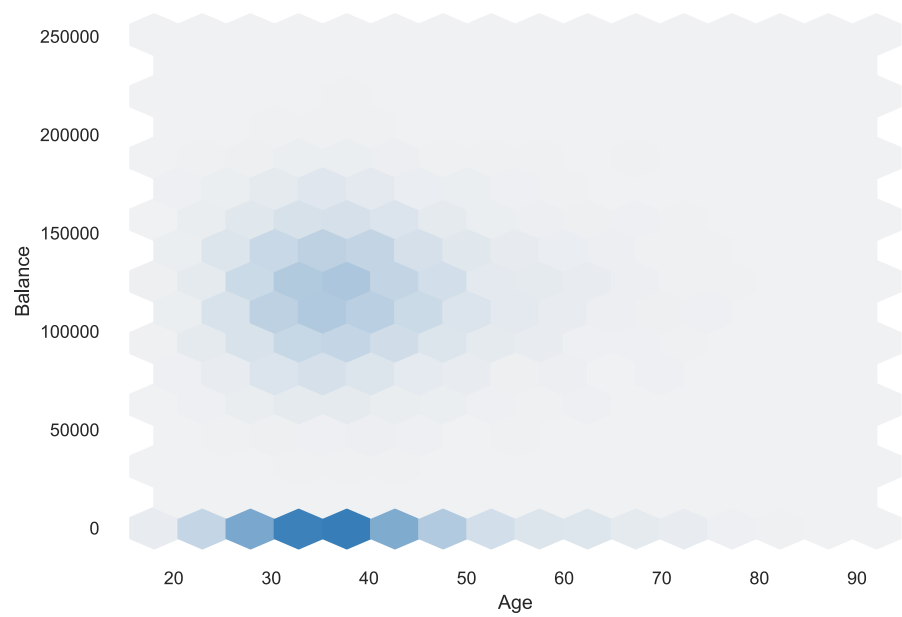




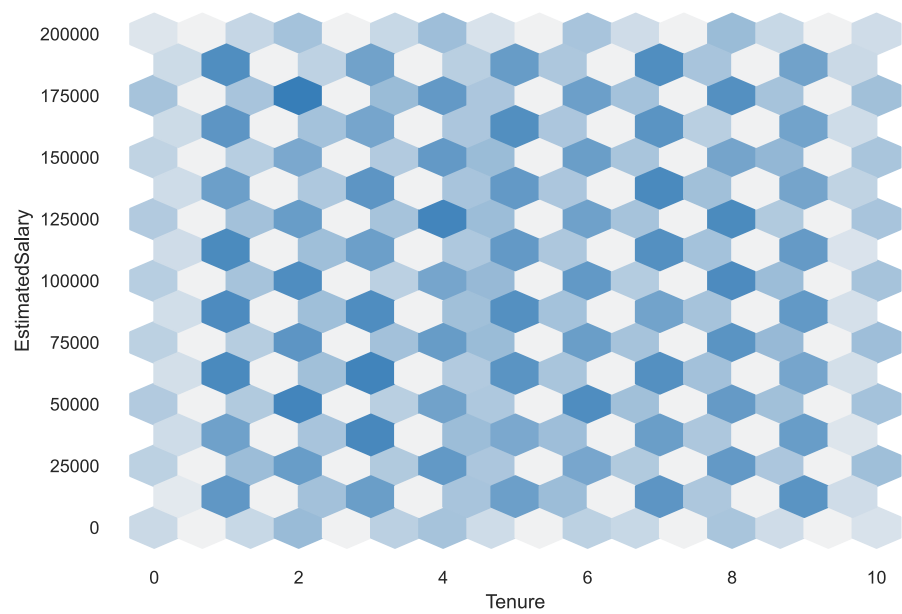


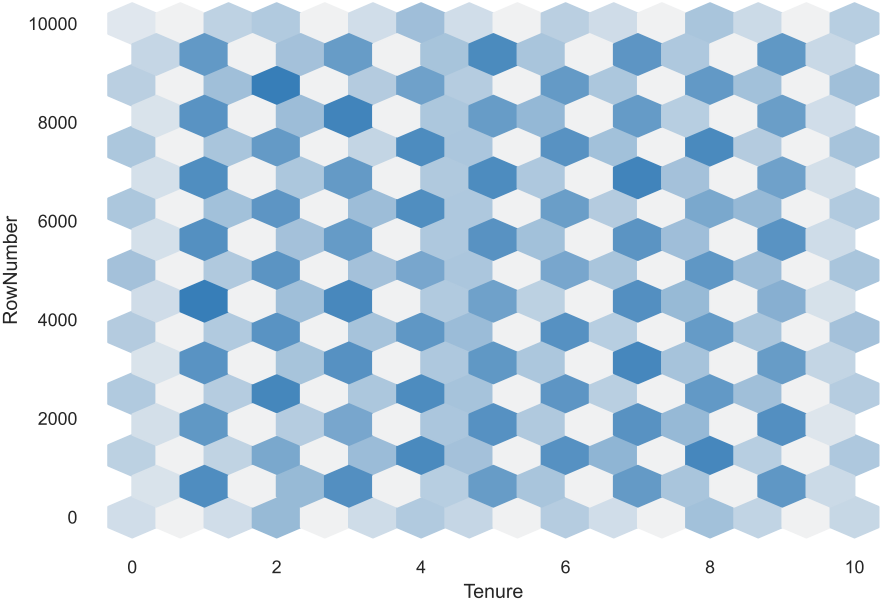


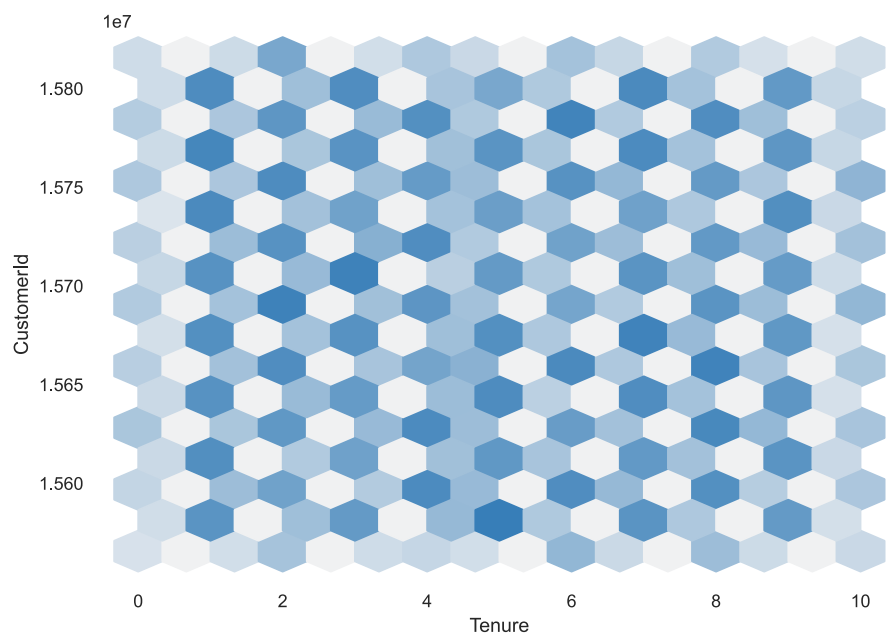


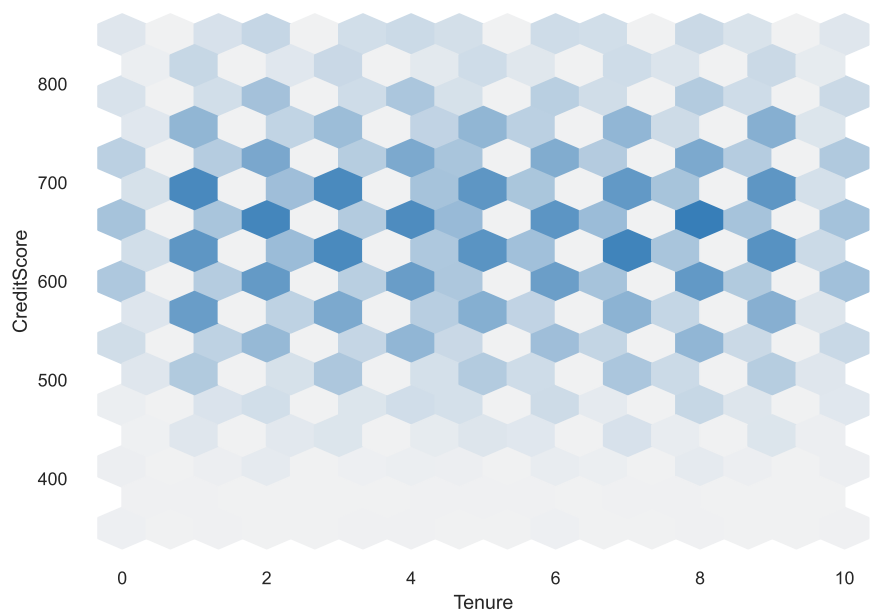




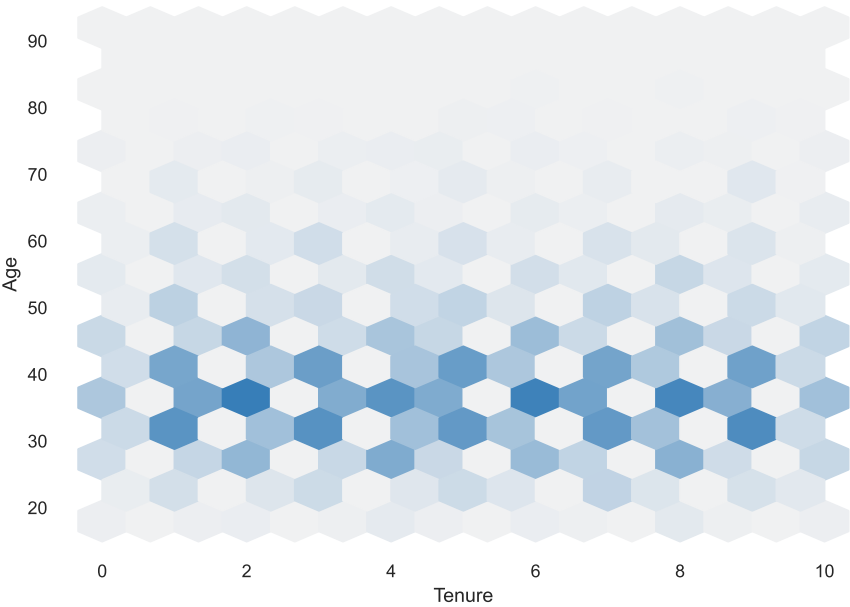


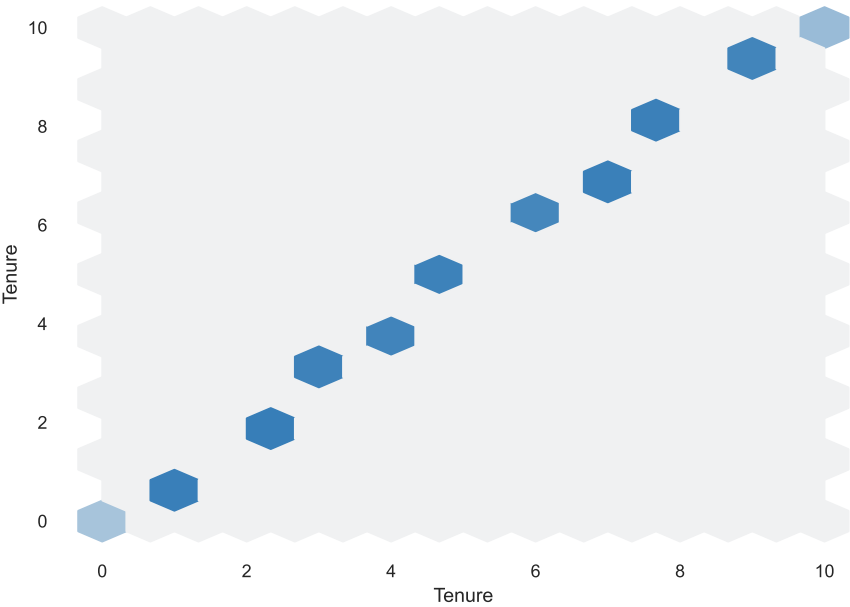


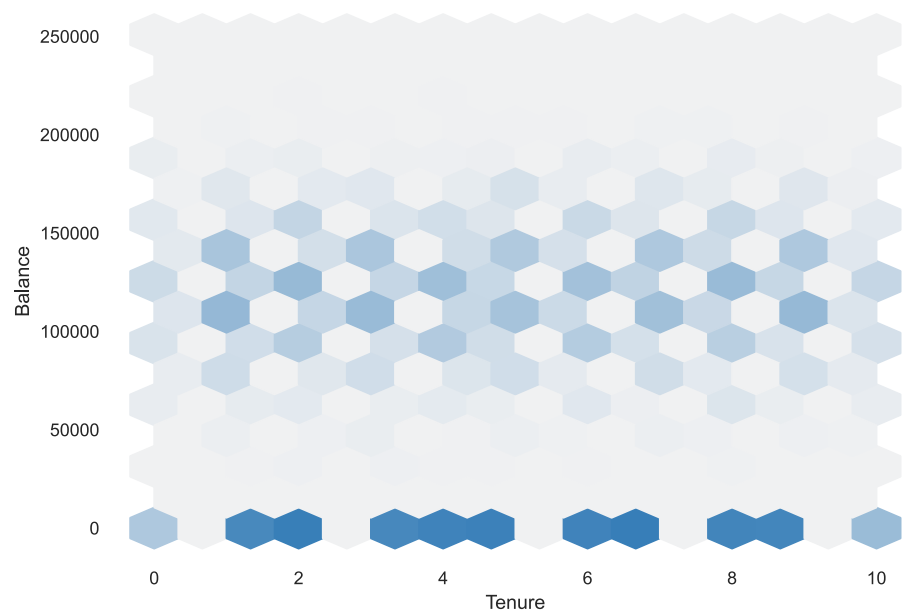




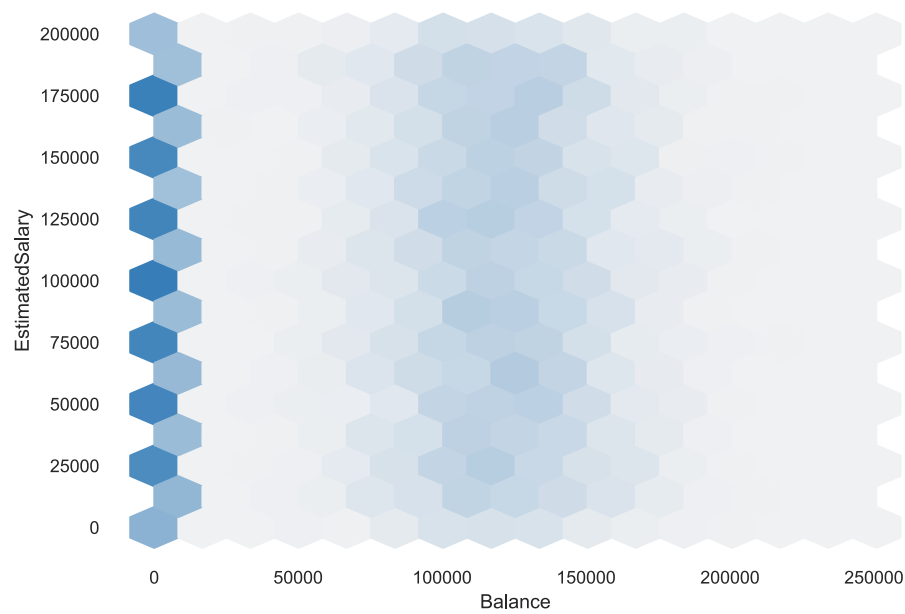


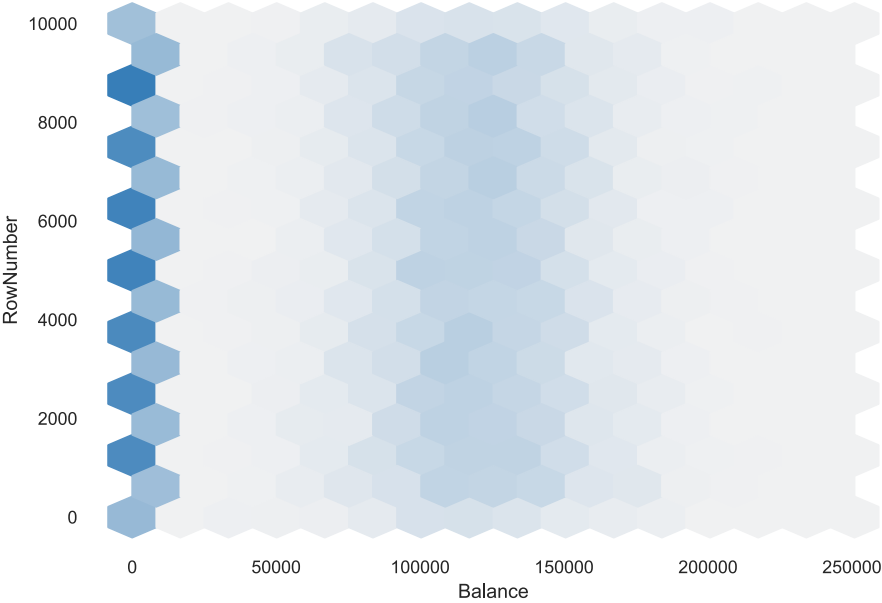


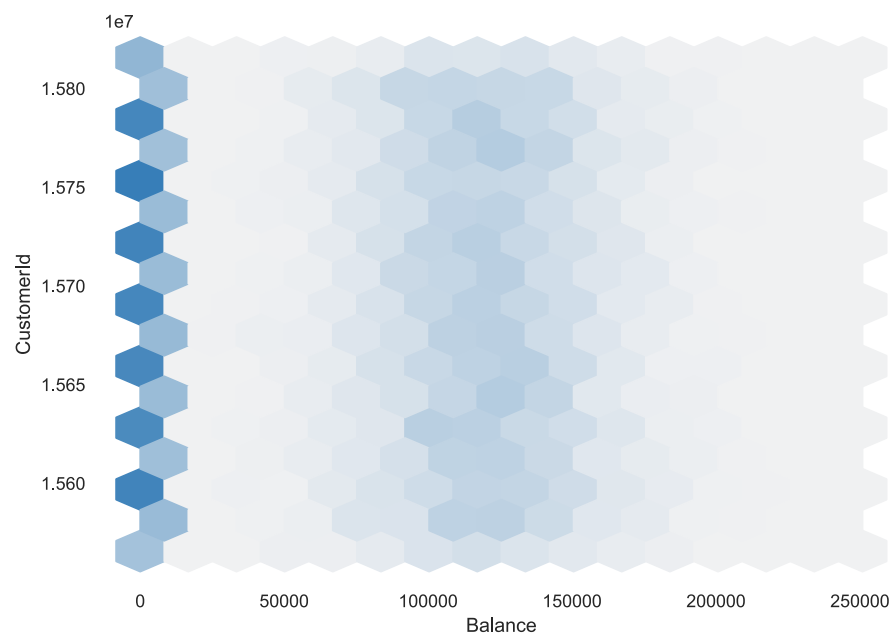


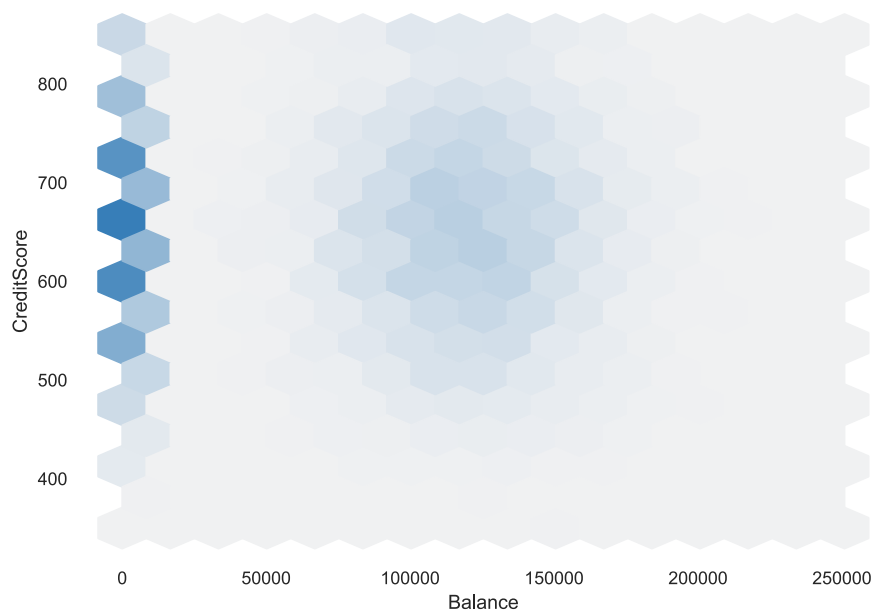




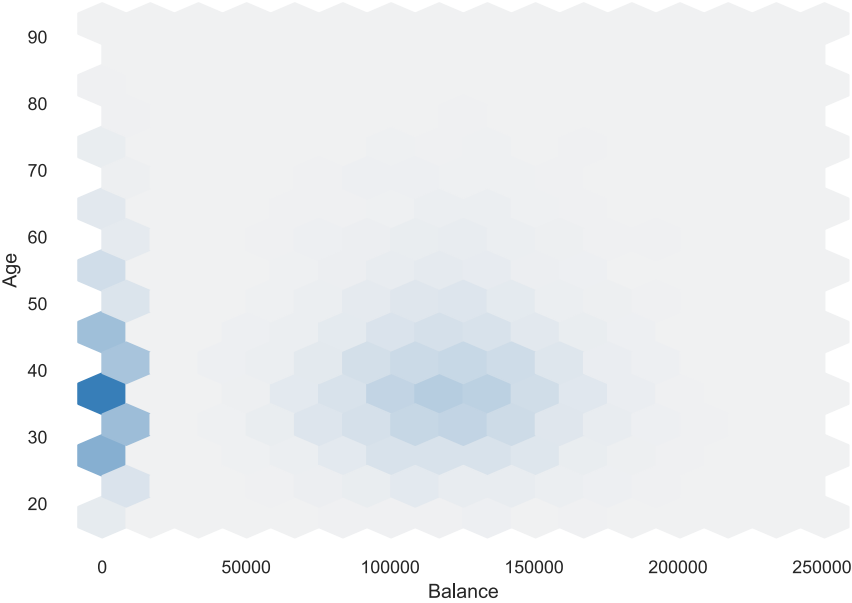


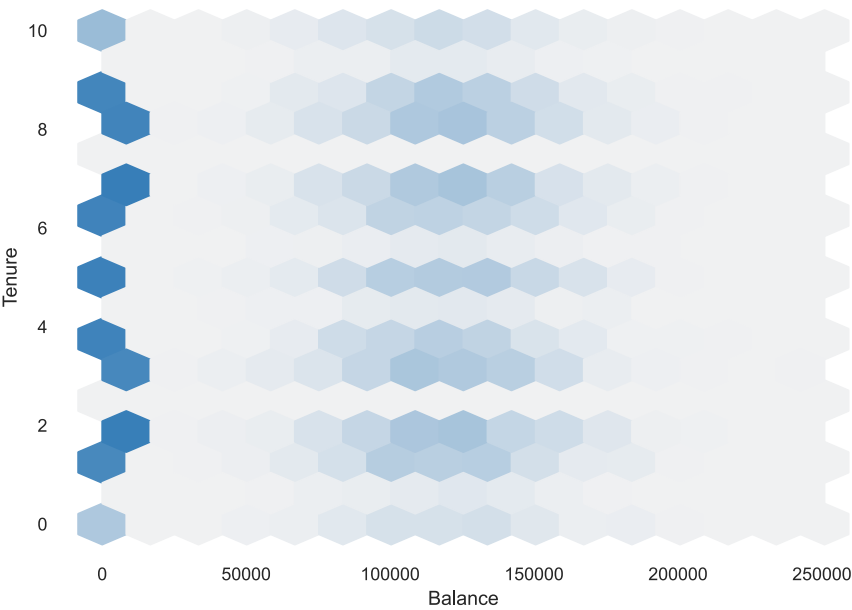


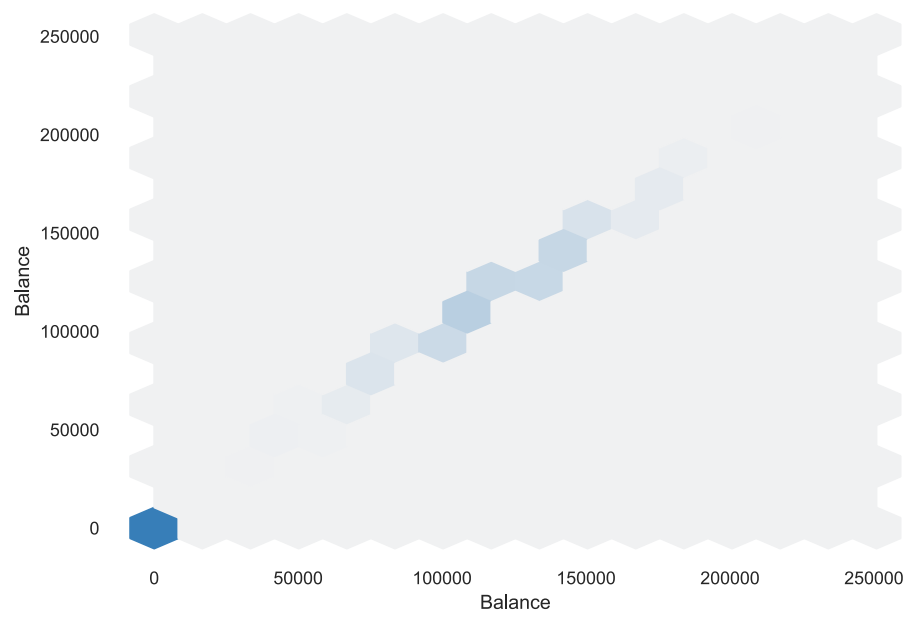




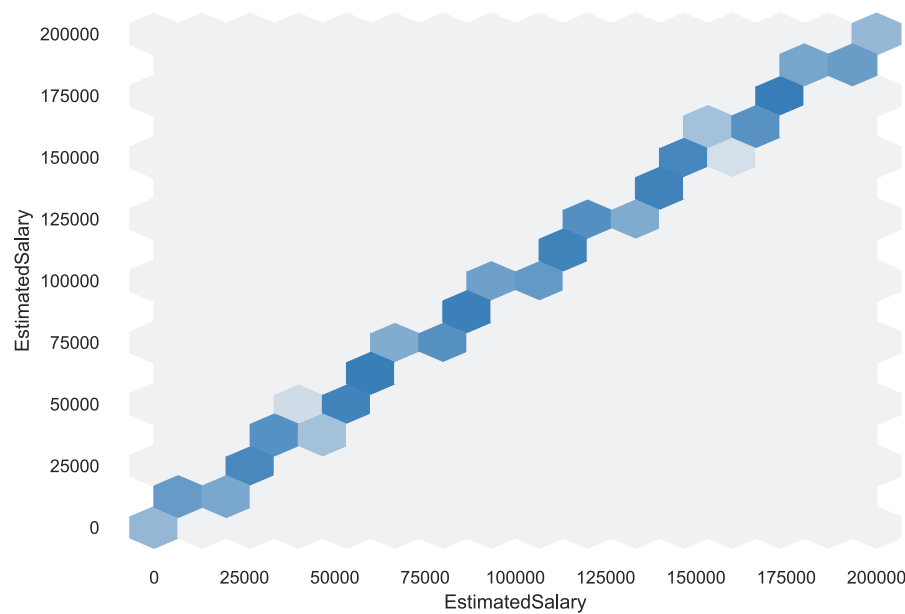


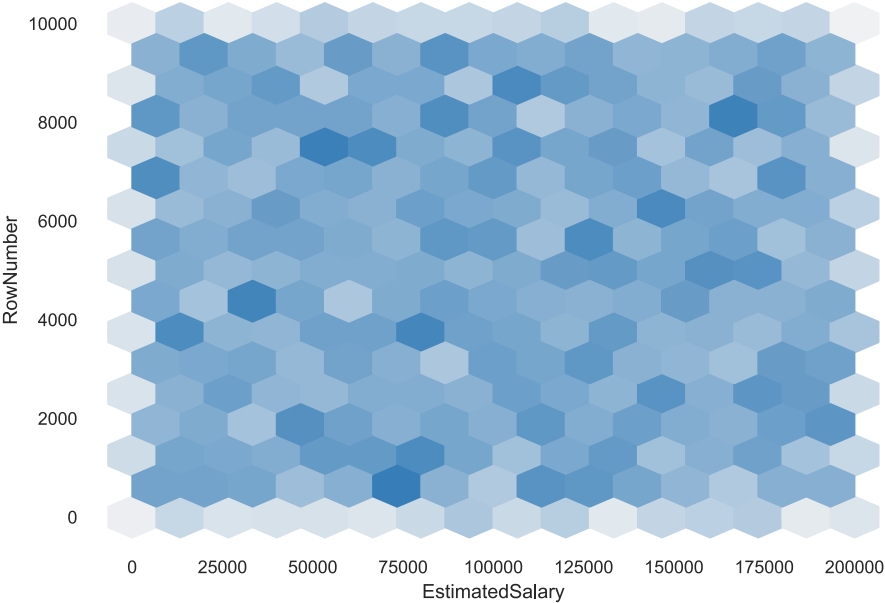


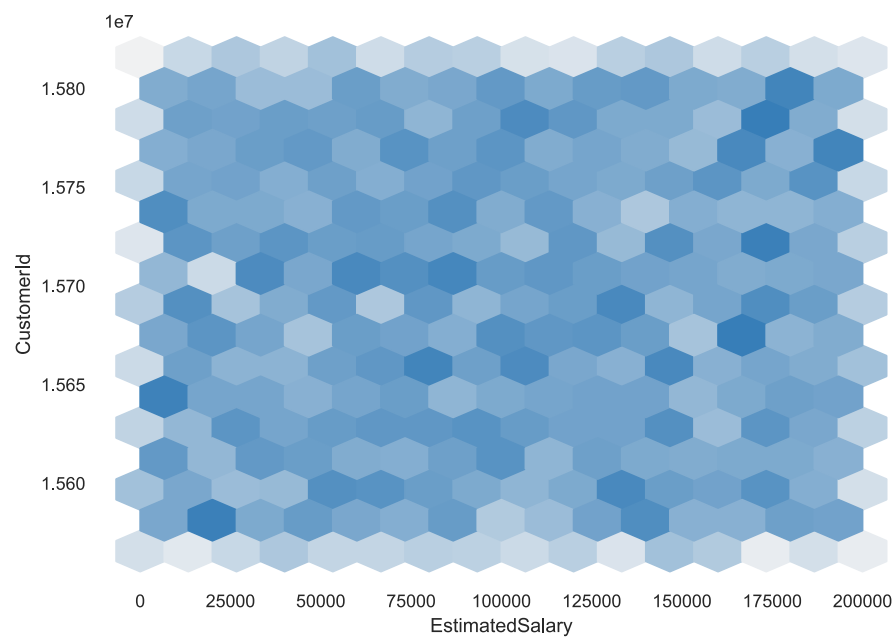


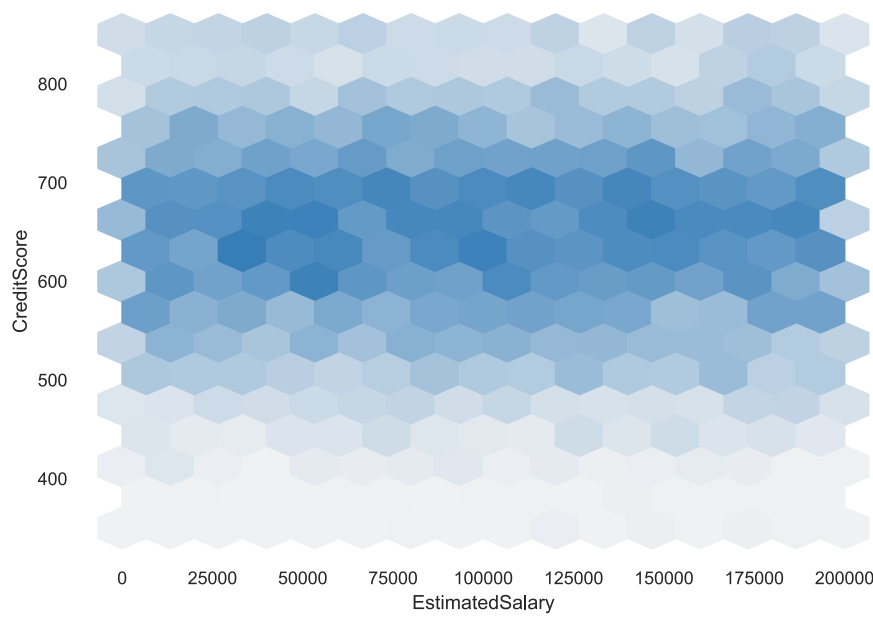




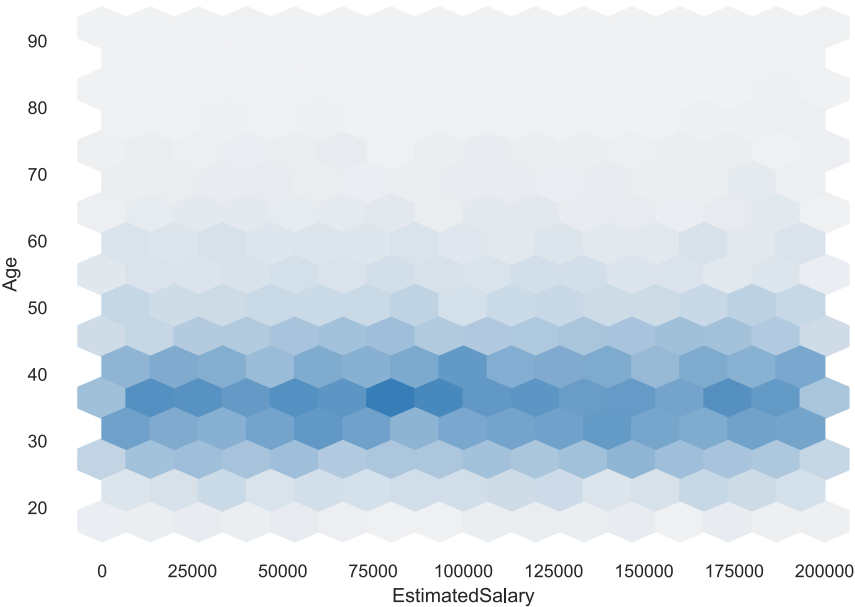


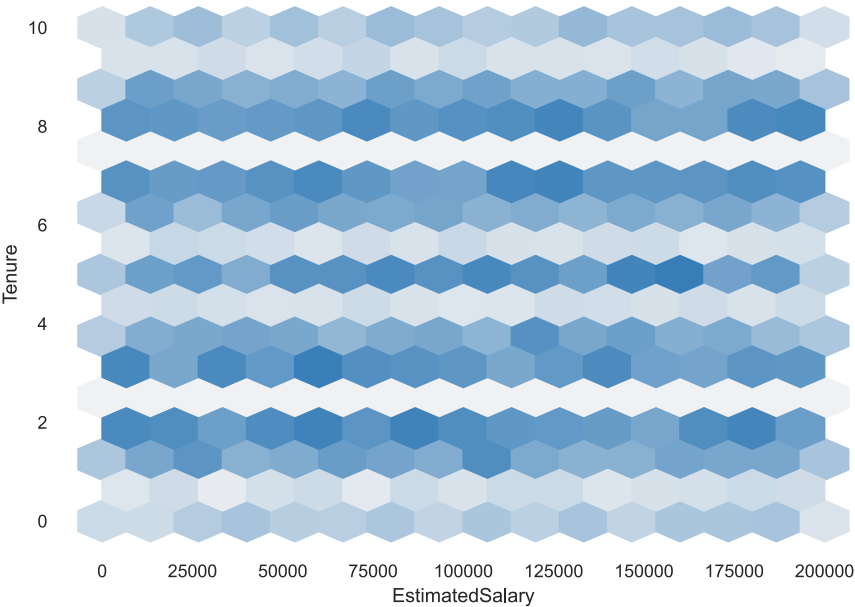


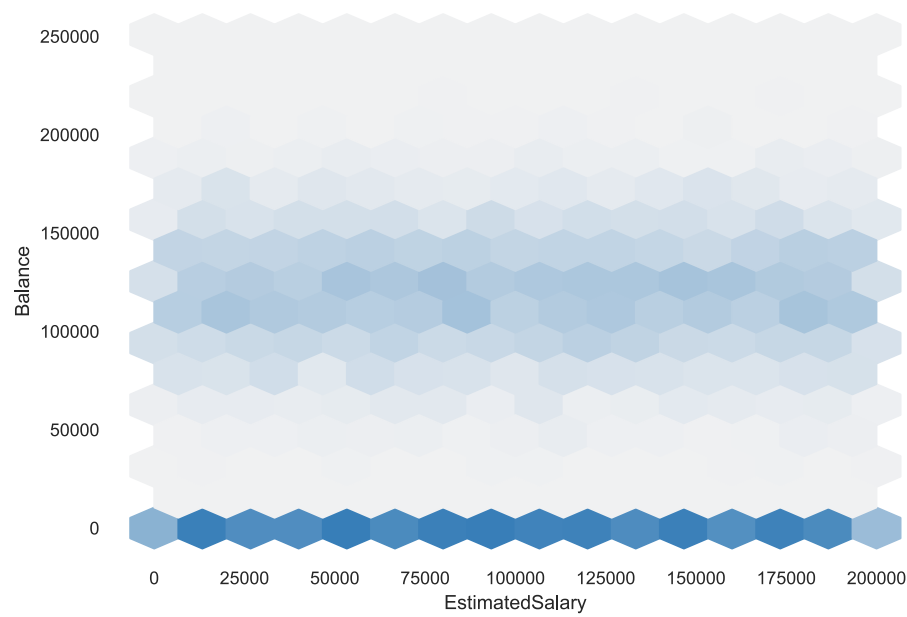






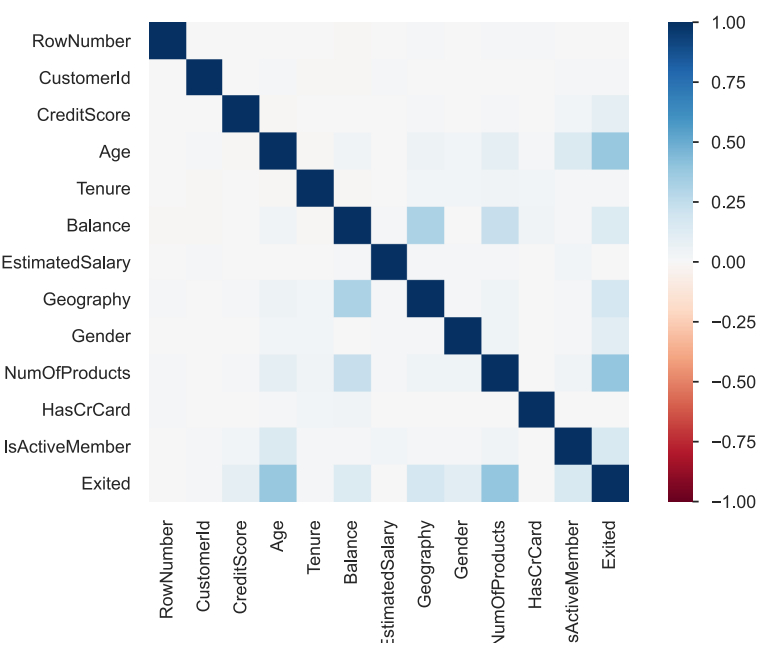


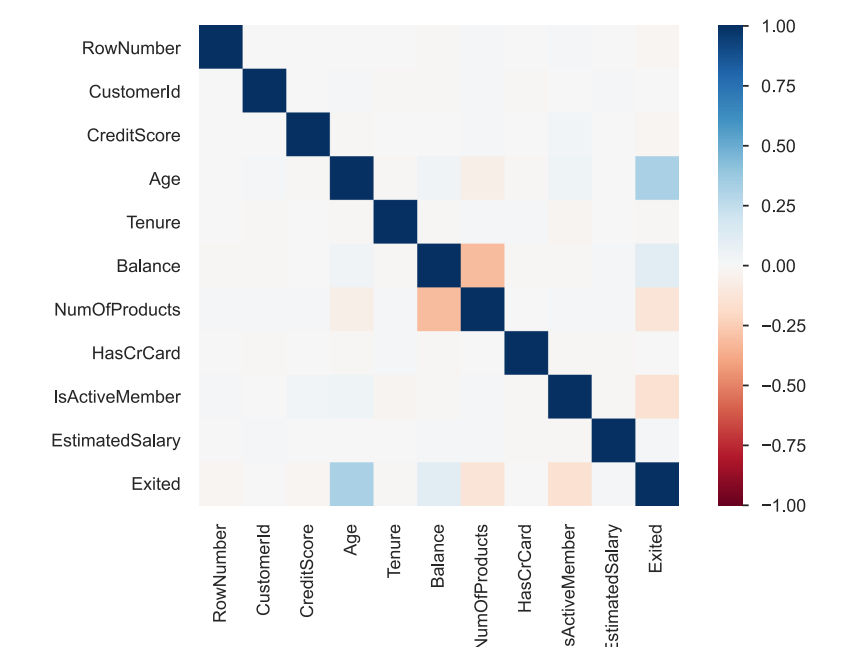


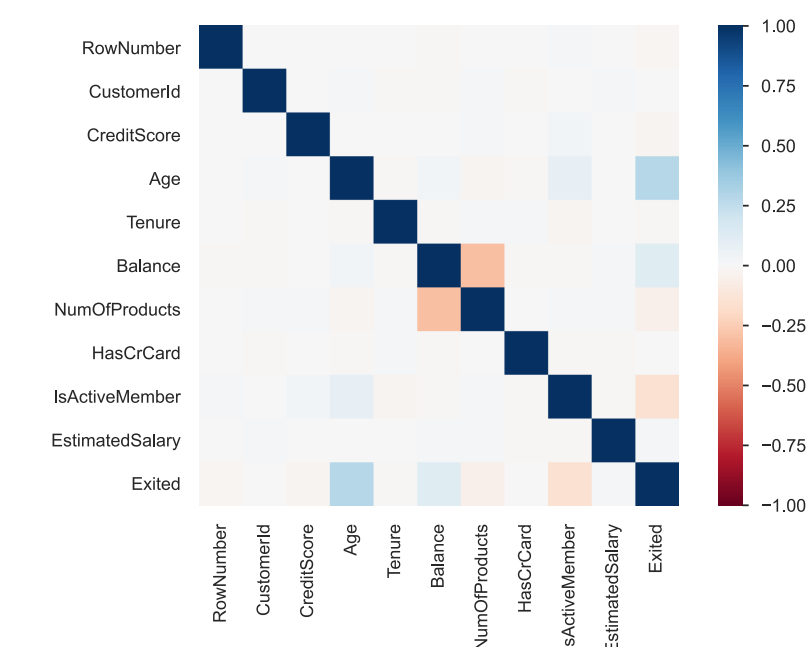


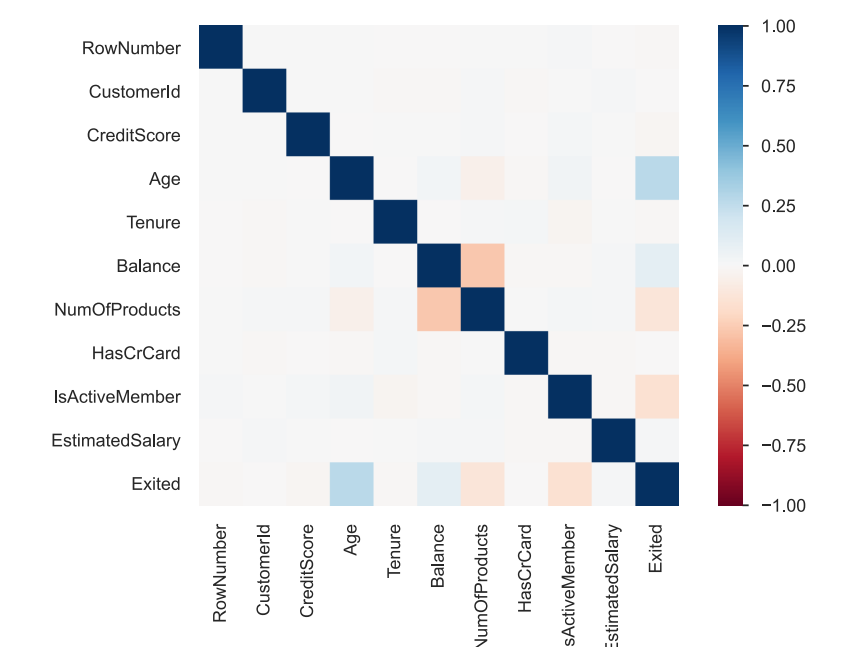


# Correlations

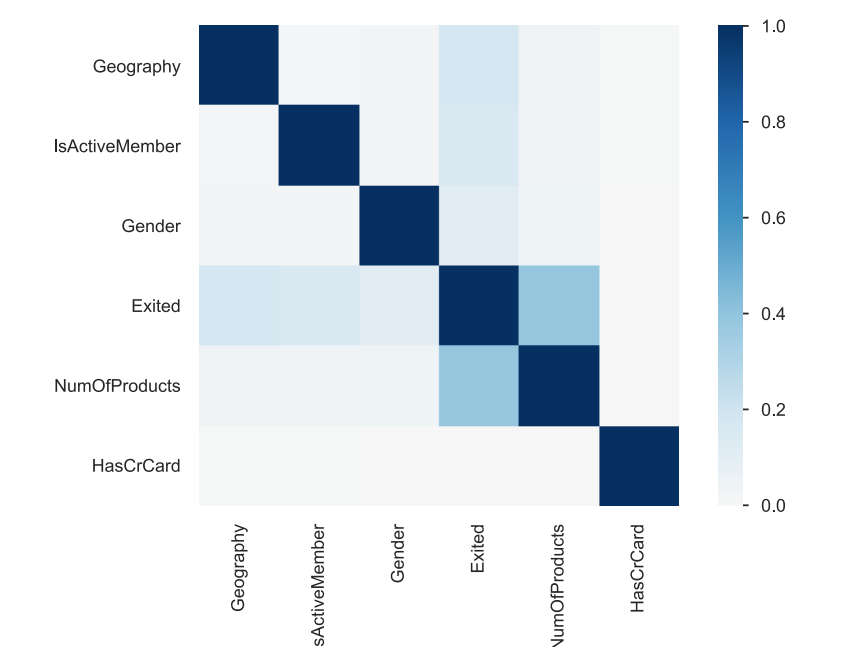


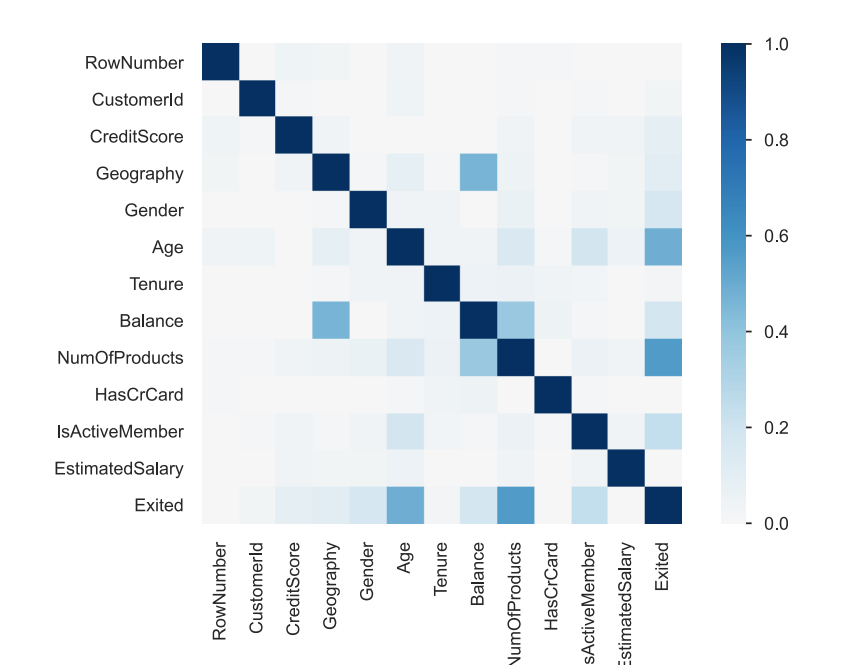




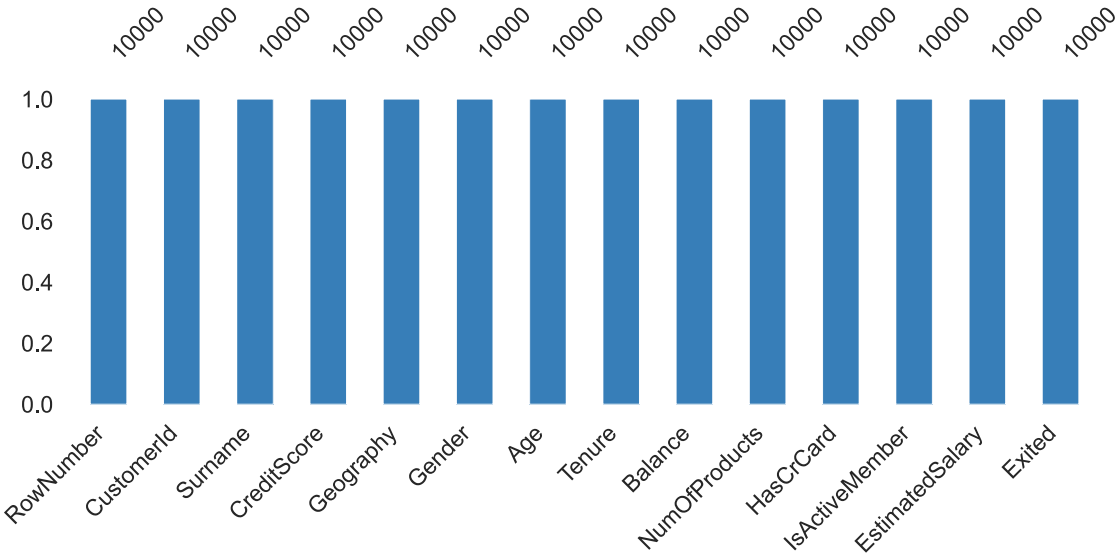




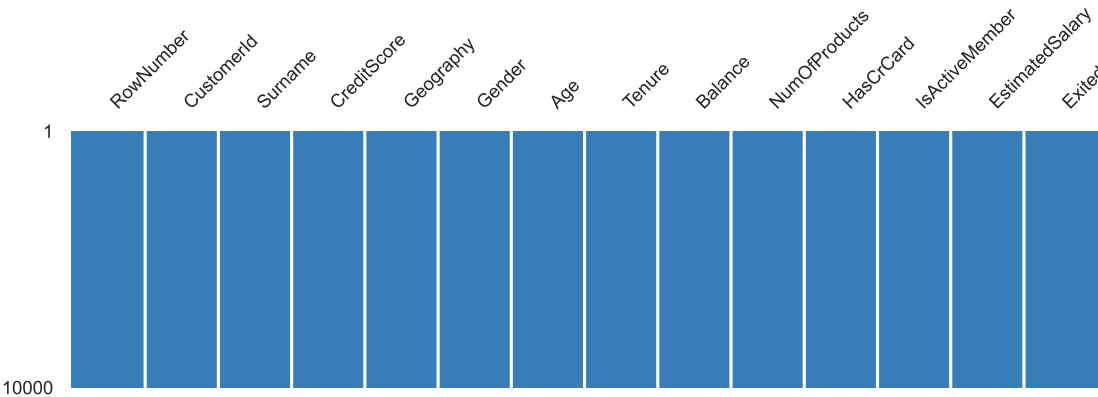




# Missing values



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.

# Sample

	RowNumber	CustomerId	Surname	CreditScore	Geography	Gender	Age	Tenure	Balance	NumOfProducts	HasCrCard	IsAct
0	1	15634602	Hargrave	619	France	Female	42	2	0.00	1	1	1
1	2	15647311	Hill	608	Spain	Female	41	1	83807.86	1	0	1
2	3	15619304	Onio	502	France	Female	42	8	159660.80	3	1	0
3	4	15701354	Boni	699	France	Female	39	1	0.00	2	0	0
4	5	15737888	Mitchell	850	Spain	Female	43	2	125510.82	1	1	1
5	6	15574012	Chu	645	Spain	Male	44	8	113755.78	2	1	0
6	7	15592531	Bartlett	822	France	Male	50	7	0.00	2	1	1
7	8	15656148	Obinna	376	Germany	Female	29	4	115046.74	4	1	0
8	9	15792365	He	501	France	Male	44	4	142051.07	2	0	1
9	10	15592389	H?	684	France	Male	27	2	134603.88	1	1	1

	RowNumber	CustomerId	Surname	CreditScore	Geography	Gender	Age	Tenure	Balance	NumOfProducts	HasCrCard
9990	9991	15798964	Nkemakonam	714	Germany	Male	33	3	35016.60	1	1
9991	9992	15769959	Ajuluchukwu	597	France	Female	53	4	88381.21	1	1
9992	9993	15657105	Chukwualuka	726	Spain	Male	36	2	0.00	1	1
9993	9994	15569266	Rahman	644	France	Male	28	7	155060.41	1	1
9994	9995	15719294	Wood	800	France	Female	29	2	0.00	2	0
9995	9996	15606229	Obijiaku	771	France	Male	39	5	0.00	2	1
9996	9997	15569892	Johnstone	516	France	Male	35	10	57369.61	1	1
9997	9998	15584532	Liu	709	France	Female	36	7	0.00	1	0
9998	9999	15682355	Sabbatini	772	Germany	Male	42	3	75075.31	2	1
9999	10000	15628319	Walker	792	France	Female	28	4	130142.79	1	1

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)).