

Overview

Dataset info

Number of variables	99
Number of observations	537081
Total Missing (%)	20.6%
Total size in memory	405.7 MiB
Average record size in memory	792.0 B

Variables types

Numeric	32
Categorical	26
Boolean	4
Date	0
Text (Unique)	0
Rejected	37
Unsupported	0

Warnings

activity_year	has constant value 2018	Rejected
lei	has a high cardinality: 836 distinct values	Warning
state_code	has constant value NY	Rejected
county_code	is highly skewed ($\gamma_1 = -109.41$)	Skewed
census_tract	is highly correlated with county_code ($\rho = 1$)	Rejected
conforming_loan_limit	has 6979 / 1.3% missing values	Missing
purchaser_type	has 353713 / 65.9% zeros	Zeros
open-end_line_of_credit	is highly correlated with reverse_mortgage ($\rho = 0.98618$)	Rejected
business_or_commercial_purpose	is highly correlated with open-end_line_of_credit ($\rho = 0.9861$)	Rejected
loan_amount	is highly skewed ($\gamma_1 = 104.26$)	Skewed
loan_to_value_ratio	has 144138 / 26.8% missing values	Missing
loan_to_value_ratio	has a high cardinality: 32994 distinct values	Warning
interest_rate	has 182122 / 33.9% missing values	Missing
interest_rate	has a high cardinality: 2341 distinct values	Warning
rate_spread	has 265405 / 49.4% missing values	Missing
rate_spread	has a high cardinality: 18393 distinct values	Warning
total_loan_costs	has 285632 / 53.2% missing values	Missing
total_loan_costs	has a high cardinality: 152314 distinct values	Warning
total_points_and_fees	has 520206 / 96.9% missing values	Missing
total_points_and_fees	has a high cardinality: 673 distinct values	Warning
origination_charges	has 285854 / 53.2% missing values	Missing
origination_charges	has a high cardinality: 59913 distinct values	Warning
discount_points	has 433433 / 80.7% missing values	Missing
discount_points	has a high cardinality: 40806 distinct values	Warning
lender_credits	has 431416 / 80.3% missing values	Missing
lender_credits	has a high cardinality: 32608 distinct values	Warning
loan_term	has 12214 / 2.3% missing values	Missing
loan_term	has a high cardinality: 568 distinct values	Warning
prepayment_penalty_term	has 484006 / 90.1% missing values	Missing
intro_rate_period	has 383860 / 71.5% missing values	Missing
intro_rate_period	has a high cardinality: 91 distinct values	Warning
negative_amortization	is highly correlated with business_or_commercial_purpose ($\rho = 0.97315$)	Rejected
interest_only_payment	is highly correlated with negative_amortization ($\rho = 1$)	Rejected
balloon_payment	is highly correlated with interest_only_payment ($\rho = 1$)	Rejected
other_nonamortizing_features	is highly correlated with balloon_payment ($\rho = 1$)	Rejected
property_value	has 77294 / 14.4% missing values	Missing
property_value	has a high cardinality: 1664 distinct values	Warning
manufactured_home_secured_property_type	is highly correlated with other_nonamortizing_features ($\rho = 0.98656$)	Rejected
manufactured_home_land_property_interest	is highly correlated with manufactured_home_secured_property_type ($\rho = 1$)	Rejected
multifamily_affordable_units	has 518232 / 96.5% missing values	Missing
multifamily_affordable_units	has a high cardinality: 91 distinct values	Warning
income	has 53414 / 9.9% missing values	Missing
income	is highly skewed ($\gamma_1 = 679.53$)	Skewed
debt_to_income_ratio	has 147171 / 27.4% missing values	Missing
applicant_credit_score_type	is highly correlated with manufactured_home_land_property_interest ($\rho = 0.99752$)	Rejected
co-applicant_credit_score_type	is highly correlated with applicant_credit_score_type ($\rho = 0.99976$)	Rejected
applicant_ethnicity-2	has 525024 / 97.8% missing values	Missing
applicant_ethnicity-3	has 536813 / 100.0% missing values	Missing
applicant_ethnicity-4	is highly correlated with applicant_ethnicity-3 ($\rho = 0.91826$)	Rejected
applicant_ethnicity-5	is highly correlated with applicant_ethnicity-4 ($\rho = 1$)	Rejected
co-applicant_ethnicity-2	has 532543 / 99.2% missing values	Missing
co-applicant_ethnicity-3	has 536991 / 100.0% missing values	Missing
co-applicant_ethnicity-4	is highly correlated with co-applicant_ethnicity-3 ($\rho = 1$)	Rejected
co-applicant_ethnicity-5	has 537080 / 100.0% missing values	Missing
applicant_race-2	is highly correlated with applicant_ethnicity-5 ($\rho = 0.99028$)	Rejected
applicant_race-3	has 535974 / 99.8% missing values	Missing
applicant_race-4	has 536965 / 100.0% missing values	Missing

applicant_race-5 has 537042 / 100.0% missing values **Missing**

co-applicant_race-1 is highly correlated with co-applicant_ethnicity-4 ($p = 1$) **Rejected**

co-applicant_race-2 has 528406 / 98.4% missing values **Missing**

co-applicant_race-3 is highly correlated with applicant_race-5 ($p = 1$) **Rejected**

co-applicant_race-4 is highly correlated with applicant_race-5 ($p = 1$) **Rejected**

co-applicant_race-5 is highly correlated with co-applicant_race-4 ($p = 0.9538$) **Rejected**

applicant_race_observed is highly correlated with co-applicant_race-5 ($p = 0.94141$) **Rejected**

co-applicant_race_observed is highly correlated with co-applicant_race-5 ($p = 0.94141$) **Rejected**

applicant_sex is highly correlated with co-applicant_ethnicity-4 ($p = 1$) **Rejected**

co-applicant_sex is highly correlated with co-applicant_race_observed ($p = 0.93907$) **Rejected**

applicant_sex_observed is highly correlated with applicant_race_observed ($p = 0.97696$) **Rejected**

co-applicant_sex_observed is highly correlated with co-applicant_sex ($p = 0.9419$) **Rejected**

applicant_age_above_62 has 49419 / 9.2% missing values **Missing**

co-applicant_age_above_62 has 326524 / 60.8% missing values **Missing**

submission_of_application is highly correlated with co-applicant_credit_score_type ($p = 0.99872$) **Rejected**

initially_payable_to_institution is highly correlated with submission_of_application ($p = 1$) **Rejected**

aus-1 is highly correlated with initially_payable_to_institution ($p = 0.99775$) **Rejected**

aus-2 has 511796 / 95.3% missing values **Missing**

aus-3 has 528144 / 98.3% missing values **Missing**

aus-4 is highly correlated with aus-3 ($p = 0.95423$) **Rejected**

aus-5 is highly correlated with aus-4 ($p = 0.96911$) **Rejected**

denial_reason-2 has 506868 / 94.4% missing values **Missing**

denial_reason-3 is highly correlated with applicant_race-5 ($p = 1$) **Rejected**

denial_reason-4 is highly correlated with applicant_race-3 ($p = 1$) **Rejected**

tract_minority_population_percent is highly correlated with co-applicant_ethnicity-4 ($p = 1$) **Rejected**

tract_to_msa_income_percentage is highly correlated with co-applicant_ethnicity-4 ($p = 1$) **Rejected**

tract_median_age_of_housing_units is highly correlated with co-applicant_ethnicity-4 ($p = 1$) **Rejected**

Dataset has 507 duplicate rows **Warning**

Variables

activity_year
Constant

This variable is constant and should be ignored for analysis

Constant value 2018

lei
Categorical


Distinct count	836
Unique (%)	0.2%
Missing (%)	0.0%
Missing (n)	0

7H6GLXDRUGQFU57RNE97	41377
KB1H1DSPRFMYMCUFXT09	38135
549300FGXN1K3HLB1R50	22467
Other values (833)	435102

[Toggle details](#)

derived_msa-md
Numeric

Distinct count	47	Mean	36588
Unique (%)	0.0%	Minimum	0
Missing (%)	0.0%	Maximum	99999
Missing (n)	0	Zeros (%)	0.7%
Infinite (%)	0.0%		
Infinite (n)	0		





[Toggle details](#)

state_code
Constant

This variable is constant and should be ignored for analysis

Constant value NY

<div>county_code</div> <div>Numeric</div> <div><div><div><div>Distinct count</div><div>111</div></div><div><div>Unique (%)</div><div>0.0%</div></div><div><div>Missing (%)</div><div>0.6%</div></div><div><div>Missing (n)</div><div>3181</div></div><div><div>Infinite (%)</div><div>0.0%</div></div><div><div>Infinite (n)</div><div>0</div></div></div><div><div><div></div></div></div></div> <div><div>Mean</div><div>36067</div></div> <div><div>Minimum</div><div>4013</div></div> <div><div>Maximum</div><div>55115</div></div> <div><div>Zeros (%)</div><div>0.0%</div></div>		<div><div></div></div> <div><div></div></div>		<div><div></div></div> <div><div></div></div>		<div>Toggle details</div>
<div>census_tract</div> <div>Highly correlated</div> <div><div>This variable is highly correlated with county_code and should be ignored for analysis</div></div> <div><div>Correlation</div><div>1</div></div>						
<div>conforming_loan_limit</div> <div>Categorical</div> <div><div><div><div>Distinct count</div><div>3</div></div><div><div>Unique (%)</div><div>0.0%</div></div><div><div>Missing (%)</div><div>1.3%</div></div><div><div>Missing (n)</div><div>6979</div></div></div><div><div><div></div></div></div></div> <div><div>C</div><div>498261</div></div> <div><div>NC</div><div>31841</div></div> <div><div>(Missing)</div><div>6979</div></div>						<div>Toggle details</div>
<div>derived_loan_product_type</div> <div>Categorical</div> <div><div><div><div>Distinct count</div><div>7</div></div><div><div>Unique (%)</div><div>0.0%</div></div><div><div>Missing (%)</div><div>0.0%</div></div><div><div>Missing (n)</div><div>0</div></div></div><div><div><div></div></div></div></div> <div><div>Conventional:First Lien</div><div>344446</div></div> <div><div>Conventional:Subordinate Lien</div><div>107546</div></div> <div><div>FHA:First Lien</div><div>66325</div></div> <div><div>Other values (4)</div><div>18764</div></div>						<div>Toggle details</div>
<div>derived_dwelling_category</div> <div>Categorical</div> <div><div><div><div>Distinct count</div><div>4</div></div><div><div>Unique (%)</div><div>0.0%</div></div><div><div>Missing (%)</div><div>0.0%</div></div><div><div>Missing (n)</div><div>0</div></div></div><div><div><div></div></div></div></div> <div><div>Single Family (1-4 Units):Site-Built</div><div>522134</div></div> <div><div>Single Family (1-4 Units):Manufactured</div><div>7968</div></div> <div><div>Multifamily:Site-Built</div><div>6917</div></div>						<div>Toggle details</div>
<div>derived_ethnicity</div> <div>Categorical</div> <div><div><div><div>Distinct count</div><div>5</div></div><div><div>Unique (%)</div><div>0.0%</div></div><div><div>Missing (%)</div><div>0.0%</div></div><div><div>Missing (n)</div><div>0</div></div></div><div><div><div></div></div></div></div> <div><div>Not Hispanic or Latino</div><div>383706</div></div> <div><div>Ethnicity Not Available</div><div>108904</div></div> <div><div>Hispanic or Latino</div><div>35370</div></div> <div><div>Other values (2)</div><div>9101</div></div>						<div>Toggle details</div>

<div><div>derived_race</div><div>Categorical</div><div><div><div>Distinct count</div><div>9</div></div><div><div>Unique (%)</div><div>0.0%</div></div><div><div>Missing (%)</div><div>0.0%</div></div><div><div>Missing (n)</div><div>0</div></div></div><div><div>White</div><div>338210</div></div><div><div>Race Not Available</div><div>111186</div></div><div><div>Asian</div><div>41763</div></div><div><div>Other values (6)</div><div>45922</div></div><div>Toggle details</div></div>			
<div><div>derived_sex</div><div>Categorical</div><div><div><div>Distinct count</div><div>4</div></div><div><div>Unique (%)</div><div>0.0%</div></div><div><div>Missing (%)</div><div>0.0%</div></div><div><div>Missing (n)</div><div>0</div></div></div><div><div>Joint</div><div>176626</div></div><div><div>Male</div><div>167586</div></div><div><div>Female</div><div>110503</div></div><div>Toggle details</div></div>			
<div><div>action_taken</div><div>Numeric</div><div><div><div>Distinct count</div><div>8</div></div><div><div>Unique (%)</div><div>0.0%</div></div><div><div>Missing (%)</div><div>0.0%</div></div><div><div>Missing (n)</div><div>0</div></div><div><div>Infinite (%)</div><div>0.0%</div></div><div><div>Infinite (n)</div><div>0</div></div></div><div><div>Mean</div><div>2.425</div></div><div><div>Minimum</div><div>1</div></div><div><div>Maximum</div><div>8</div></div><div><div>Zeros (%)</div><div>0.0%</div></div><div></div><div>Toggle details</div></div>			
<div><div>purchaser_type</div><div>Numeric</div><div><div><div>Distinct count</div><div>11</div></div><div><div>Unique (%)</div><div>0.0%</div></div><div><div>Missing (%)</div><div>0.0%</div></div><div><div>Missing (n)</div><div>0</div></div><div><div>Infinite (%)</div><div>0.0%</div></div><div><div>Infinite (n)</div><div>0</div></div></div><div><div>Mean</div><div>3.7427</div></div><div><div>Minimum</div><div>0</div></div><div><div>Maximum</div><div>72</div></div><div><div>Zeros (%)</div><div>65.9%</div></div><div></div><div>Toggle details</div></div>			
<div><div>preapproval</div><div>Boolean</div><div><div><div>Distinct count</div><div>2</div></div><div><div>Unique (%)</div><div>0.0%</div></div><div><div>Missing (%)</div><div>0.0%</div></div><div><div>Missing (n)</div><div>0</div></div></div><div><div>Mean</div><div>1.98</div></div><div><div>2</div><div>526324</div></div><div><div>1</div><div>10757</div></div><div>Toggle details</div></div>			
<div><div>loan_type</div><div>Numeric</div><div><div><div>Distinct count</div><div>4</div></div><div><div>Unique (%)</div><div>0.0%</div></div><div><div>Missing (%)</div><div>0.0%</div></div><div><div>Missing (n)</div><div>0</div></div><div><div>Infinite (%)</div><div>0.0%</div></div></div><div><div>Mean</div><div>1.1981</div></div><div><div>Minimum</div><div>1</div></div><div><div>Maximum</div><div>4</div></div><div><div>Zeros (%)</div><div>0.0%</div></div><div></div><div></div></div>			

Infinite (n)

0

Toggle details

loan_purpose

Numeric

Distinct count

6

Unique (%)

0.0%

Missing (%)

0.0%

Missing (n)

0

Infinite (%)

0.0%

Infinite (n)

0

Mean

11.038

Minimum

1

Maximum

32

Zeros (%)

0.0%

Toggle details

lien_status

Boolean

Distinct count

2

Unique (%)

0.0%

Missing (%)

0.0%

Missing (n)

0

Mean

1.2006

1

429322

2

107759

Toggle details

reverse_mortgage

Numeric

Distinct count

3

Unique (%)

0.0%

Missing (%)

0.0%

Missing (n)

0

Infinite (%)

0.0%

Infinite (n)

0

Mean

33.157

Minimum

1

Maximum

1111

Zeros (%)

0.0%

Toggle details

open-end_line_of_credit

Highly correlated

This variable is highly correlated with reverse_mortgage and should be ignored for analysis

Correlation

0.98618

business_or_commercial_purpose

Highly correlated

This variable is highly correlated with open-end_Line_of_credit and should be ignored for analysis

Correlation

0.9861

loan_amount

Numeric

Distinct count

1282

Unique (%)

0.2%

Missing (%)

0.0%

Missing (n)

0

Infinite (%)

0.0%

Infinite (n)

0

Mean

355650

Minimum

5000

Maximum

500000000

Zeros (%)

0.0%

1e8

Toggle details

<div>loan_to_value_ratio</div> <div>Categorical</div> <div><div><div>Distinct count</div><div>32994</div></div><div><div>Unique (%)</div><div>6.1%</div></div><div><div>Missing (%)</div><div>26.8%</div></div><div><div>Missing (n)</div><div>144138</div></div></div>				<div><div>80.0</div><div>43687</div></div> <div><div>Exempt</div><div>15533</div></div> <div><div>96.5</div><div>15054</div></div> <div><div>Other values (32990)</div><div>318669</div></div> <div><div>(Missing)</div><div>144138</div></div>				<div>Toggle details</div>	
<div>interest_rate</div> <div>Categorical</div> <div><div><div>Distinct count</div><div>2341</div></div><div><div>Unique (%)</div><div>0.4%</div></div><div><div>Missing (%)</div><div>33.9%</div></div><div><div>Missing (n)</div><div>182122</div></div></div>				<div><div>4.75</div><div>30837</div></div> <div><div>4.5</div><div>27803</div></div> <div><div>4.625</div><div>24299</div></div> <div><div>Other values (2337)</div><div>272020</div></div> <div><div>(Missing)</div><div>182122</div></div>				<div>Toggle details</div>	
<div>rate_spread</div> <div>Categorical</div> <div><div><div>Distinct count</div><div>18393</div></div><div><div>Unique (%)</div><div>3.4%</div></div><div><div>Missing (%)</div><div>49.4%</div></div><div><div>Missing (n)</div><div>265405</div></div></div>				<div><div>Exempt</div><div>15432</div></div> <div><div>0.0</div><div>1171</div></div> <div><div>4.75</div><div>543</div></div> <div><div>Other values (18389)</div><div>254530</div></div> <div><div>(Missing)</div><div>265405</div></div>				<div>Toggle details</div>	
<div>hoepa_status</div> <div>Numeric</div> <div><div><div>Distinct count</div><div>3</div></div><div><div>Unique (%)</div><div>0.0%</div></div><div><div>Missing (%)</div><div>0.0%</div></div><div><div>Missing (n)</div><div>0</div></div><div><div>Infinite (%)</div><div>0.0%</div></div><div><div>Infinite (n)</div><div>0</div></div></div>				<div><div>Mean</div><div>2.428</div></div> <div><div>Minimum</div><div>1</div></div> <div><div>Maximum</div><div>3</div></div> <div><div>Zeros (%)</div><div>0.0%</div></div>				<div>Toggle details</div>	
<div>total_loan_costs</div> <div>Categorical</div> <div><div><div>Distinct count</div><div>152314</div></div><div><div>Unique (%)</div><div>28.4%</div></div><div><div>Missing (%)</div><div>53.2%</div></div><div><div>Missing (n)</div><div>285632</div></div></div>				<div><div>Exempt</div><div>15528</div></div> <div><div>0.0</div><div>9147</div></div> <div><div>485.0</div><div>449</div></div> <div><div>Other values (152310)</div><div>226325</div></div> <div><div>(Missing)</div><div>285632</div></div>				<div>Toggle details</div>	

total_points_and_fees					
Categorical					
Distinct count	673				
Unique (%)	0.1%				
Missing (%)	96.9%				
Missing (n)	520206				
		Exempt	15528		
		0.0	387		
		661.5	45		
		Other values (669)	915		
		(Missing)		520206	
		Toggle details			

origination_charges					
Categorical					
Distinct count	59913				
Unique (%)	11.2%				
Missing (%)	53.2%				
Missing (n)	285854				
		0.0	34100		
		Exempt	15528		
		995.0	6175		
		Other values (59909)		195424	
		(Missing)			285854
		Toggle details			

discount_points					
Categorical					
Distinct count	40806				
Unique (%)	7.6%				
Missing (%)	80.7%				
Missing (n)	433433				
		Exempt	15527		
		750.0	257		
		1500.0	250		
		Other values (40802)	87614		
		(Missing)		433433	
		Toggle details			

lender_credits					
Categorical					
Distinct count	32608				
Unique (%)	6.1%				
Missing (%)	80.3%				
Missing (n)	431416				
		Exempt	15527		
		500.0	2779		
		250.0	2278		
		Other values (32604)		85081	
		(Missing)			431416
		Toggle details			

loan_term					
Categorical					
Distinct count	568				
Unique (%)	0.1%				
Missing (%)	2.3%				
Missing (n)	12214				
		360		370060	
		180	40220		
		240	34446		
		Other values (564)	80141		
		Toggle details			

<div>prepayment_penalty_term</div> <div>Categorical</div> <div><div><div>Distinct count</div><div>17</div></div><div><div>Unique (%)</div><div>0.0%</div></div><div><div>Missing (%)</div><div>90.1%</div></div><div><div>Missing (n)</div><div>484006</div></div></div> <div><div><div>36</div><div>31295</div></div><div><div>Exempt</div><div>15528</div></div><div><div>24</div><div>4848</div></div><div><div>Other values (13)</div><div>1404</div></div><div><div>(Missing)</div><div>484006</div></div></div> <div>Toggle details</div>			
<div>intro_rate_period</div> <div>Categorical</div> <div><div><div>Distinct count</div><div>91</div></div><div><div>Unique (%)</div><div>0.0%</div></div><div><div>Missing (%)</div><div>71.5%</div></div><div><div>Missing (n)</div><div>383860</div></div></div> <div><div><div>1</div><div>87592</div></div><div><div>Exempt</div><div>15526</div></div><div><div>60</div><div>10095</div></div><div><div>Other values (87)</div><div>40008</div></div><div><div>(Missing)</div><div>383860</div></div></div> <div>Toggle details</div>			
<div>negative_amortization</div> <div>Highly correlated</div> <div><div><div><div>This variable is highly correlated with <code>business_or_commercial_purpose</code> and should be ignored for analysis</div></div></div></div> <div><div>Correlation</div><div>0.97315</div></div>			
<div>interest_only_payment</div> <div>Highly correlated</div> <div><div><div><div>This variable is highly correlated with <code>negative_amortization</code> and should be ignored for analysis</div></div></div></div> <div><div>Correlation</div><div>1</div></div>			
<div>balloon_payment</div> <div>Highly correlated</div> <div><div><div><div>This variable is highly correlated with <code>interest_only_payment</code> and should be ignored for analysis</div></div></div></div> <div><div>Correlation</div><div>1</div></div>			
<div>other_nonamortizing_features</div> <div>Highly correlated</div> <div><div><div><div>This variable is highly correlated with <code>balloon_payment</code> and should be ignored for analysis</div></div></div></div> <div><div>Correlation</div><div>1</div></div>			
<div>property_value</div> <div>Categorical</div> <div><div><div>Distinct count</div><div>1664</div></div><div><div>Unique (%)</div><div>0.3%</div></div><div><div>Missing (%)</div><div>14.4%</div></div><div><div>Missing (n)</div><div>77294</div></div></div> <div><div><div>Exempt</div><div>15532</div></div><div><div>125000</div><div>10881</div></div><div><div>135000</div><div>10870</div></div><div><div>Other values (1660)</div><div>422504</div></div><div><div>(Missing)</div><div>77294</div></div></div> <div>Toggle details</div>			
<div>construction_method</div> <div>Boolean</div> <div><div><div>Distinct count</div><div>2</div></div><div><div>Mean</div><div>1.015</div></div></div>			

Unique (%)0.0%

Missing (%)0.0%

Missing (n)0

1529051

28030

Toggle details

occupancy_type

Numeric

Distinct count3

Unique (%)0.0%

Missing (%)0.0%

Missing (n)0

Infinite (%)0.0%


Infinite (n)0

Mean1.1746

Minimum1

Maximum3

Zeros (%)0.0%



Toggle details

manufactured_home_secured_property_type

Highly correlated

This variable is highly correlated with other_nonamortizing_features and should be ignored for analysis

Correlation0.98656

manufactured_home_land_property_interest

Highly correlated

This variable is highly correlated with manufactured_home_secured_property_type and should be ignored for analysis

Correlation1

total_units

Categorical

Distinct count9

Unique (%)0.0%

Missing (%)0.0%

Missing (n)0

1468247

248907

39244

Other values (6)10683

Toggle details

multifamily_affordable_units

Categorical

Distinct count91

Unique (%)0.0%

Missing (%)96.5%

Missing (n)518232

Exempt15527

02681

100252

Other values (87)389

(Missing)518232

Toggle details

income

Numeric

Distinct count3014

Unique (%)0.6%

Missing (%)9.9%

Missing (n)53414

Infinite (%)0.0%

Infinite (n)0

Mean178.52



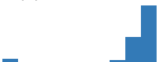
Minimum-17862

Maximum11971000

Zeros (%)0.8%

1e7

</

Toggle details			
<div><div>applicant_ethnicity-4</div><div>Highly correlated</div><div>This variable is highly correlated with applicant_ethnicity-3 and should be ignored for analysis</div><div>Correlation0.91826</div></div>			
<div><div>applicant_ethnicity-5</div><div>Highly correlated</div><div>This variable is highly correlated with applicant_ethnicity-4 and should be ignored for analysis</div><div>Correlation1</div></div>			
<div><div>co-applicant_ethnicity-1</div><div>Numeric</div><div><div><div>Distinct count10</div><div>Unique (%)0.0%</div><div>Missing (%)0.0%</div><div>Missing (n)69</div><div>Infinite (%)0.0%</div><div>Infinite (n)0</div></div><div><div>Mean3.8183</div><div>Minimum1</div><div>Maximum14</div><div>Zeros (%)0.0%</div></div></div><div></div><div>Toggle details</div></div>			
<div><div>co-applicant_ethnicity-2</div><div>Numeric</div><div><div><div>Distinct count7</div><div>Unique (%)0.0%</div><div>Missing (%)99.2%</div><div>Missing (n)532543</div><div>Infinite (%)0.0%</div><div>Infinite (n)0</div></div><div><div>Mean12.534</div><div>Minimum1</div><div>Maximum14</div><div>Zeros (%)0.0%</div></div></div><div></div><div>Toggle details</div></div>			
<div><div>co-applicant_ethnicity-3</div><div>Numeric</div><div><div><div>Distinct count7</div><div>Unique (%)0.0%</div><div>Missing (%)100.0%</div><div>Missing (n)536991</div><div>Infinite (%)0.0%</div><div>Infinite (n)0</div></div><div><div>Mean12.711</div><div>Minimum1</div><div>Maximum14</div><div>Zeros (%)0.0%</div></div></div><div></div><div>Toggle details</div></div>			
<div><div>co-applicant_ethnicity-4</div><div>Highly correlated</div><div>This variable is highly correlated with co-applicant_ethnicity-3 and should be ignored for analysis</div><div>Correlation1</div></div>			
<div><div>co-applicant_ethnicity-5</div><div>Boolean</div><div><div><div>Distinct count2</div><div>Unique (%)0.0%</div><div>Missing (%)100.0%</div><div>Missing (n)537080</div></div><div><div>Mean14</div><div>14.0 (Missing)1</div><div>537080</div></div></div><div>Toggle details</div></div>			

applicant_ethnicity_observed

Numeric

Distinct count	3	Mean	2.0737
Unique (%)	0.0%	Minimum	1
Missing (%)	0.0%	Maximum	3
Missing (n)	0	Zeros (%)	0.0%
Infinite (%)	0.0%		
Infinite (n)	0		



[Toggle details](#)

co-applicant_ethnicity_observed

Numeric

Distinct count	4	Mean	3.1427
Unique (%)	0.0%	Minimum	1
Missing (%)	0.0%	Maximum	4
Missing (n)	0	Zeros (%)	0.0%
Infinite (%)	0.0%		
Infinite (n)	0		



[Toggle details](#)

applicant_race-1

Numeric

Distinct count	19	Mean	5.2061
Unique (%)	0.0%	Minimum	1
Missing (%)	0.0%	Maximum	44
Missing (n)	107	Zeros (%)	0.0%
Infinite (%)	0.0%		
Infinite (n)	0		



[Toggle details](#)

applicant_race-2

Highly correlated

This variable is highly correlated with applicant_ethnicity-5 and should be ignored for analysis

Correlation 0.99028

applicant_race-3

Numeric

Distinct count	16	Mean	21.963
Unique (%)	0.0%	Minimum	1
Missing (%)	99.8%	Maximum	44
Missing (n)	535974	Zeros (%)	0.0%
Infinite (%)	0.0%		
Infinite (n)	0		



[Toggle details](#)

applicant_race-4

Numeric

Distinct count	13	Mean	21.345
Unique (%)	0.0%	Minimum	4
Missing (%)	100.0%	Maximum	44
Missing (n)	536965	Zeros (%)	0.0%
Infinite (%)	0.0%		
Infinite (n)	0		



applicant_race-5

Distinct count	9
Unique (%)	0.0%
Missing (%)	100.0%
Missing (n)	537042
Infinite (%)	0.0%
Infinite (n)	0

Mean	16.897
Minimum	5
Maximum	44
Zeros (%)	0.0%

~~co-applicant_race-1~~

This variable is highly correlated with co-applicant_ethnicity-4 and should be ignored for analysis

co-applicant_race-2

Distinct count	17
Unique (%)	0.0%
Missing (%)	98.4%
Missing (n)	528406
Infinite (%)	0.0%
Infinite (n)	0

Mean	20.761
Minimum	1
Maximum	44
Zeros (%)	0.0%

~~co-applicant_race-3~~

This variable is highly correlated with applicant_race-5 and should be ignored for analysis

~~co-applicant_race-4~~

This variable is highly correlated with applicant_race-5 and should be ignored for analysis

~~co-applicant_race-5~~

This variable is highly correlated with co-applicant_race-4 and should be ignored for analysis

~~applicant_race_observed~~



This variable is highly correlated with co-applicant_race-5 and should be ignored for analysis





~~co_applicant_race_observed~~

This variable is highly correlated with co-applicant race-5 and should be ignored for analysis

Correlation 0.94141

<div><div>applicant_sex</div><div>Highly correlated</div><div>This variable is highly correlated with co-applicant_ethnicity-4 and should be ignored for analysis</div><div>Correlation1</div></div>
<div><div>co-applicant_sex</div><div>Highly correlated</div><div>This variable is highly correlated with co-applicant_race_observed and should be ignored for analysis</div><div>Correlation0.93907</div></div>
<div><div>applicant_sex_observed</div><div>Highly correlated</div><div>This variable is highly correlated with applicant_race_observed and should be ignored for analysis</div><div>Correlation0.97696</div></div>
<div><div>co-applicant_sex_observed</div><div>Highly correlated</div><div>This variable is highly correlated with co-applicant_sex and should be ignored for analysis</div><div>Correlation0.9419</div></div>
<div><div>applicant_age</div><div>Categorical</div><div><div><div>Distinct count</div><div>Unique (%)</div><div>Missing (%)</div><div>Missing (n)</div></div><div><div>8</div><div>0.0%</div><div>0.0%</div><div>0</div></div></div><div><div>45-54</div><div>115764</div></div><div><div>35-44</div><div>114737</div></div><div><div>55-64</div><div>92696</div></div><div><div>Other values (5)</div><div>213884</div></div><div>Toggle details</div></div>
<div><div>co-applicant_age</div><div>Categorical</div><div><div><div>Distinct count</div><div>Unique (%)</div><div>Missing (%)</div><div>Missing (n)</div></div><div><div>9</div><div>0.0%</div><div>0.0%</div><div>0</div></div></div><div><div>9999</div><div>288493</div></div><div><div>35-44</div><div>47945</div></div><div><div>45-54</div><div>47067</div></div><div><div>Other values (6)</div><div>153576</div></div><div>Toggle details</div></div>
<div><div>applicant_age_above_62</div><div>Categorical</div><div><div><div>Distinct count</div><div>Unique (%)</div><div>Missing (%)</div><div>Missing (n)</div></div><div><div>3</div><div>0.0%</div><div>9.2%</div><div>49419</div></div></div><div><div>No</div><div>400671</div></div><div><div>Yes</div><div>86991</div></div><div><div>(Missing)</div><div>49419</div></div><div>Toggle details</div></div>
<div><div>co-applicant_age_above_62</div><div>Categorical</div><div><div><div>Distinct count</div><div>Unique (%)</div><div>Missing (%)</div><div>Missing (n)</div></div><div><div>3</div><div>0.0%</div><div>60.8%</div><div>326524</div></div></div><div><div>No</div><div>171424</div></div></div>

<div>Yes39133</div> <div>(Missing)326524</div> <div>Toggle details</div>			
<div><div>submission_of_application</div><div>Highly correlated</div><div>This variable is highly correlated with co-applicant_credit_score_type and should be ignored for analysis</div><div>Correlation0.99872</div></div>			
<div><div>initially_payable_to_institution</div><div>Highly correlated</div><div>This variable is highly correlated with submission_of_application and should be ignored for analysis</div><div>Correlation1</div></div>			
<div><div>aus-1</div><div>Highly correlated</div><div>This variable is highly correlated with initially_payable_to_institution and should be ignored for analysis</div><div>Correlation0.99775</div></div>			
<div><div>aus-2</div><div>Numeric</div><div><div><div>Distinct count6</div><div>Unique (%)0.0%</div><div>Missing (%)95.3%</div><div>Missing (n)511796</div><div>Infinite (%)0.0%</div><div>Infinite (n)0</div></div><div><div>Mean3.2647</div><div>Minimum1</div><div>Maximum5</div><div>Zeros (%)0.0%</div></div></div><div></div><div>Toggle details</div></div>			
<div><div>aus-3</div><div>Numeric</div><div><div><div>Distinct count6</div><div>Unique (%)0.0%</div><div>Missing (%)98.3%</div><div>Missing (n)528144</div><div>Infinite (%)0.0%</div><div>Infinite (n)0</div></div><div><div>Mean3.1947</div><div>Minimum1</div><div>Maximum5</div><div>Zeros (%)0.0%</div></div></div><div></div><div>Toggle details</div></div>			
<div><div>aus-4</div><div>Highly correlated</div><div>This variable is highly correlated with aus-3 and should be ignored for analysis</div><div>Correlation0.95423</div></div>			
<div><div>aus-5</div><div>Highly correlated</div><div>This variable is highly correlated with aus-4 and should be ignored for analysis</div><div>Correlation0.96911</div></div>			
<div><div>denial_reason-1</div><div>Numeric</div><div><div><div>Distinct count11</div><div>Unique (%)0.0%</div><div>Missing (%)0.0%</div><div>Missing (n)0</div><div>Infinite (%)0.0%</div></div><div><div>Mean25.996</div><div>Minimum1</div><div>Maximum1111</div><div>Zeros (%)0.0%</div></div></div></div>			

<div><div>Infinite (n)0</div><div></div></div>	<div>Toggle details</div>
<div><div>denial_reason-2</div><div>Numeric</div><div><div><div>Distinct count10</div><div>Unique (%)0.0%</div><div>Missing (%)94.4%</div><div>Missing (n)506868</div><div>Infinite (%)0.0%</div><div>Infinite (n)0</div></div><div><div>Mean3.7831</div><div>Minimum1</div><div>Maximum9</div><div>Zeros (%)0.0%</div></div></div><div></div></div>	<div>Toggle details</div>
<div><div>denial_reason-3</div><div>Highly correlated</div><div><div>This variable is highly correlated with applicant_race-5 and should be ignored for analysis</div><div>Correlation1</div></div></div>	
<div><div>denial_reason-4</div><div>Highly correlated</div><div><div>This variable is highly correlated with applicant_race-3 and should be ignored for analysis</div><div>Correlation1</div></div></div>	
<div><div>tract_population</div><div>Numeric</div><div><div><div>Distinct count3414</div><div>Unique (%)0.6%</div><div>Missing (%)0.0%</div><div>Missing (n)0</div><div>Infinite (%)0.0%</div><div>Infinite (n)0</div></div><div><div>Mean4756.1</div><div>Minimum0</div><div>Maximum28926</div><div>Zeros (%)0.7%</div></div></div><div></div></div>	<div>Toggle details</div>
<div><div>tract_minority_population_percent</div><div>Highly correlated</div><div><div>This variable is highly correlated with co-applicant_ethnicity-4 and should be ignored for analysis</div><div>Correlation1</div></div></div>	
<div><div>ffiec_msa_md_median_family_income</div><div>Numeric</div><div><div><div>Distinct count52</div><div>Unique (%)0.0%</div><div>Missing (%)0.0%</div><div>Missing (n)0</div><div>Infinite (%)0.0%</div><div>Infinite (n)0</div></div><div><div>Mean84760</div><div>Minimum0</div><div>Maximum123400</div><div>Zeros (%)0.7%</div></div></div><div></div></div>	<div>Toggle details</div>
<div><div>tract_to_msa_income_percentage</div><div>Highly correlated</div><div><div>This variable is highly correlated with co-applicant_ethnicity-4 and should be ignored for analysis</div><div>Correlation1</div></div></div>	

tract_owner_occupied_units

Numeric

Distinct count	1842	Mean	1189.8
Unique (%)	0.3%	Minimum	0
Missing (%)	0.0%	Maximum	4682
Missing (n)	0	Zeros (%)	0.8%
Infinite (%)	0.0%		
Infinite (n)	0		



Toggle details

tract_one_to_four_family_homes

Numeric

Distinct count	2231	Mean	1543.8
Unique (%)	0.4%	Minimum	0
Missing (%)	0.0%	Maximum	6541
Missing (n)	0	Zeros (%)	0.8%
Infinite (%)	0.0%		
Infinite (n)	0		



Toggle details

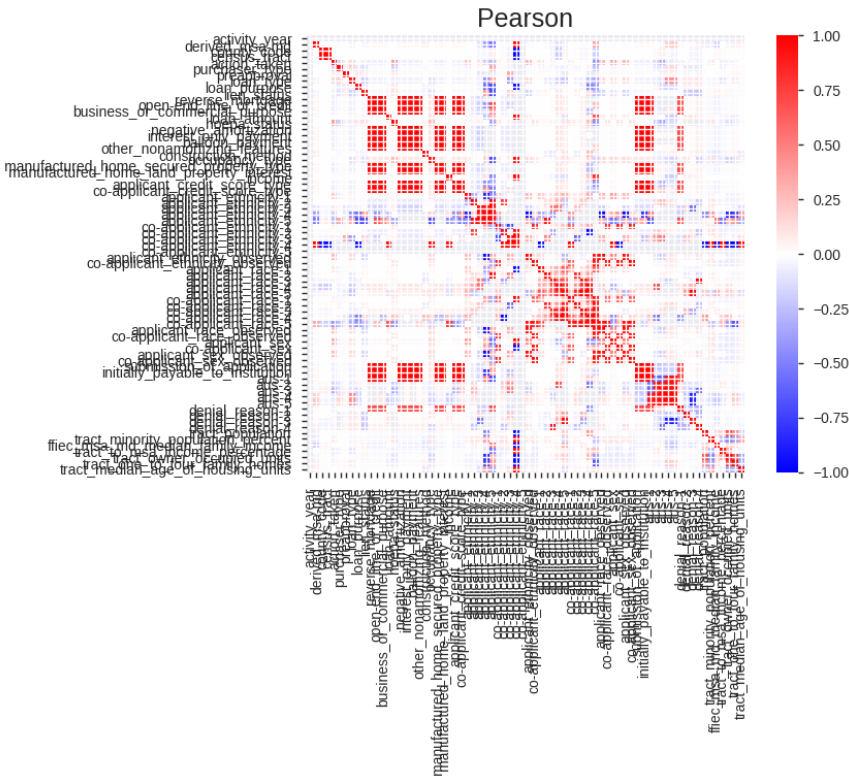
tract_median_age_of_housing_units

Highly correlated

This variable is highly correlated with co-applicant_ethnicity-4 and should be ignored for analysis

Correlation 1

Correlations





	activity_year	lei	derived_msa-md	state_code	county_code	census_year
0	2018	TR24TWEY5RVRQV65HD49	35614	NY	36085.0	3.6085
1	2018	TR24TWEY5RVRQV65HD49	35614	NY	36085.0	3.6085
2	2018	TR24TWEY5RVRQV65HD49	35614	NY	36061.0	3.6061
3	2018	TR24TWEY5RVRQV65HD49	35614	NY	36047.0	3.6047
4	2018	TR24TWEY5RVRQV65HD49	35614	NY	36081.0	3.6081