Credit Score and Default Risk

William Brickowski, Ben Miller, Spencer Hanson, Max Moghadam

Data Sets

Freddie Mac Single Family Loan-Level Dataset
http://www.freddiemac.com/research/datasets/sf_loanlevel_dataset.html
1. Origination Data File
2. Monthly Performance Data File

FBI Crime Data
https://ucr.fbi.gov/

Census Data
https://usa.ipums.org/usa/

Research Questions

How do societal and monetary factors affect default rates?

- Through the use of FBI crime statistics and Macroeconomic Census Data such as GDP and Inflation, we would like to quantify default rates in the Freddie Mac data set in order to formulate a better Credit Score

What attributes make it likely that a person will default on their loan?

- Correlation coefficient between default rate and varying attributes to determine a weighted credit score
- Probit & Logit regression models to predict the probability of defaulting on loan

Attributes of the Origination Data File

ORIGINATION DATA FILE				
COLUMN POSITION†	FORMAL NAME AND DEFINITION	VALID VALUES/ CALCULATIONS	TYPE	LENGTH
1	CREDIT SCORE - A number, prepared by third parties, summarizing the borrower's creditworthiness, which may be indicative of the likelihood that the borrower will timely repay future obligations. Generally, the credit score disclosed is the score known at the time of acquisition and is the score used to originate the mortgage.	• 301 - 850 • Space (3) = Unknown, if Credit Score is < 301 or > 850.	Numeric	3

#	Field	Values	Type	Length
2	**FIRST PAYMENT DATE** - The date of the first scheduled mortgage payment due under the terms of the mortgage note.	• YYYYMM	Date	6
3	**FIRST TIME HOMEBUYER FLAG** - Indicates whether the Borrower, or one of a group of Borrowers, is an individual who (1) is purchasing the mortgaged property, (2) will reside in the mortgaged property as a primary residence and (3) had no ownership interest (sole or joint) in a residential property during the three-year period preceding the date of the purchase of the mortgaged property. With certain limited exceptions, a displaced homemaker or single parent may also be considered a First-Time Homebuyer if the individual had no ownership interest in a residential property during the preceding three-year period other than an ownership interest in the marital residence with a spouse. Investment Properties, Second Homes and Refinance transactions are not eligible to be considered First-Time Homebuyer transactions. Therefore First Time Homebuyer does not apply and will be disclosed as "Not Applicable", which will be indicated by a blank space.	• Y = Yes • N = No • Space (1) = Unknown or Not Applicable	Alpha	1
4	**MATURITY DATE** - The month in which the final monthly payment on the mortgage is scheduled to be made as stated on the original mortgage note.	• YYYYMM	Date	6
5	**METROPOLITAN STATISTICAL AREA (MSA) OR METROPOLITAN DIVISION** - This disclosure will be based on the designation of the Metropolitan Statistical Area or Metropolitan Division based on 2010 census (for Mar 2013 and May 2013 releases) and 2013 census (for Aug 2013 and Dec 2013 releases) data. Metropolitan Statistical Areas (MSAs) are defined by the United States Office of Management and Budget (OMB) and have at least one urbanized area with a population of 50,000 or more inhabitants. OMB refers to an MSA containing a single core with a population of 2.5 million or more, which may be comprised of groupings of counties, as a Metropolitan Division. If an MSA applies to a mortgaged property, the applicable five-digit value is disclosed; however, if the mortgaged property also falls within a Metropolitan Division classification, the applicable five-digit value for the Metropolitan Division takes precedence and is disclosed instead. Changes and/or updates in designations of MSAs or Metropolitan Division will not be reflected in the Single Family Historical Dataset.	• Metropolitan Division or MSA Code. • Space (5) = Indicates that the area in which the mortgaged property is located is a) neither an MSA nor a Metropolitan Division, or b) unknown.	Numeric	5
6	**MORTGAGE INSURANCE PERCENTAGE (MI %)** - The percentage of loss coverage on the loan, at the time of Freddie Mac's purchase of the mortgage loan that a mortgage insurer is providing to cover losses incurred as a result of a default on the loan. Only primary mortgage insurance that is purchased by the Borrower, lender or Freddie Mac is disclosed. Mortgage insurance that constitutes "credit enhancement" that is not required by Freddie Mac's Charter is not disclosed. Amounts of mortgage insurance reported by Sellers that are less than 1% or greater than 55% will be disclosed as "Unknown," which will be indicated by three blank spaces. No MI will be indicated by three zeros.	• 1% - 55% • 000 = No MI • Space(3) = Unknown	Numeric	3
7	**NUMBER OF UNITS** - Denotes whether the mortgage is a one-, two-, three-, or four-unit property.	• 1 = one-unit • 2 = two-unit • 3 = three-unit • 4 = four-unit • Space(1) = Unknown	Numeric	1
8	**OCCUPANCY STATUS** - Denotes whether the mortgage type is owner occupied, second home, or investment property.	• O = Owner Occupied	Alpha	1

		• I = Investment Property • S = Second Home • Space (1) = Unknown		
9	**ORIGINAL COMBINED LOAN-TO-VALUE (CLTV)** – In the case of a purchase mortgage loan, the ratio is obtained by dividing the original mortgage loan amount on the note date plus any secondary mortgage loan amount disclosed by the Seller by the lesser of the mortgaged property's appraised value on the note date or its purchase price. In the case of a refinance mortgage loan, the ratio is obtained by dividing the original mortgage loan amount on the note date plus any secondary mortgage loan amount disclosed by the Seller by the mortgaged property's appraised value on the note date. If the secondary financing amount disclosed by the Seller includes a home equity line of credit, then the CLTV calculation reflects the disbursed amount at closing of the first lien mortgage loan, not the maximum loan amount available under the home equity line of credit. In the case of a seasoned mortgage loan, if the Seller cannot warrant that the value of the mortgaged property has not declined since the note date, Freddie Mac requires that the Seller must provide a new appraisal value, which is used in the CLTV calculation. In certain cases, where the Seller delivered a loan to Freddie Mac with a special code indicating additional secondary mortgage loan amounts, those amounts may have been included in the CLTV calculation. If the calculated CLTV is < 0 or > 200, and the LTV is > 80 and <= 200 set the CLTV to the LTV. If the LTV is < 80 or > 200 or unknown, set the CLTV to 'unknown.' If the CLTV is < LTV, set the CLTV to 'unknown.' Unknown is indicated by a null value. This disclosure is subject to the widely varying standards originators use to verify Borrowers' secondary mortgage loan amounts and will not be updated.	• 0% - 200% • Null = Unknown	Full Dataset: Numeric Literal Decimal; Sample Dataset: Numeric	Full Dataset: 7; Sample Dataset: 3
10	**ORIGINAL DEBT-TO-INCOME (DTI) RATIO** - Disclosure of the debt to income ratio is based on (1) the sum of the borrower's monthly debt payments, including monthly housing expenses that incorporate the mortgage payment the borrower is making at the time of the delivery of the mortgage loan to Freddie Mac, divided by (2) the total monthly income used to underwrite the loan as of the date of the origination of the such loan. Ratios greater than 65% are indicated by three (3) blank spaces and if unknown are indicated by a Null value. All loans in the HARP dataset will be disclosed as unknown. This disclosure is subject to the widely varying standards originators use to verify Borrowers' assets and liabilities and will not be updated.	• 0%<DTI<=65% • Space(3) >65% • Null = Unknown HARP ranges: • Null = Unknown	Numeric	3
11	**ORIGINAL UPB** - The UPB of the mortgage on the note date.	• Amount will be rounded to the nearest $1,000.	Numeric	12
12	**ORIGINAL LOAN-TO-VALUE (LTV)** - In the case of a purchase mortgage loan, the ratio obtained by dividing the original mortgage loan amount on the note date by the lesser of the mortgaged property's appraised value on the note date or its purchase price. In the case of a refinance mortgage loan, the ratio obtained by dividing the original mortgage loan amount on the note date and the mortgaged property's appraised value on the note date. In the case of a seasoned mortgage loan, if the Seller cannot warrant that the value of the mortgaged property has not declined since the	• 6% - 105% • Space(3) = Unknown HARP ranges: • 81% - 999% • Space(3) = Unknown	Numeric	3

	note date, Freddie Mac requires that the Seller must provide a new appraisal value, which is used in the LTV calculation. Ratios below 6% or greater than 105% will be disclosed as "Unknown," indicated by three (3) blank spaces. For loans in the HARP dataset, LTV ratios less than or equal to 80% and greater than 999% will be disclosed as unknown.			
13	**ORIGINAL INTEREST RATE** - The original note rate as indicated on the mortgage note.		Numeric Literal decimal	6
14	**CHANNEL** - Disclosure indicates whether a Broker or Correspondent, as those terms are defined below, originated or was involved in the origination of the mortgage loan. If a Third Party Origination is applicable, but the Seller does not specify Broker or Correspondent, the disclosure will indicate "TPO Not Specified". Similarly, if neither Third Party Origination nor Retail designations are available, the disclosure will indicate "TPO Not Specified." If a Broker, Correspondent or Third Party Origination disclosure is not applicable, the mortgage loan will be designated as Retail, as defined below. Broker is a person or entity that specializes in loan originations, receiving a commission (from a Correspondent or other lender) to match Borrowers and lenders. The Broker performs some or most of the loan processing functions, such as taking loan applications, or ordering credit reports, appraisals and title reports. Typically, the Broker does not underwrite or service the mortgage loan and generally does not use its own funds for closing; however, if the Broker funded a mortgage loan on a lender's behalf, such a mortgage loan is considered a "Broker" third party origination mortgage loan. The mortgage loan is generally closed in the name of the lender who commissioned the Broker's services. Correspondent is an entity that typically sells the Mortgages it originates to other lenders, which are not Affiliates of that entity, under a specific commitment or as part of an ongoing relationship. The Correspondent performs some, or all, of the loan processing functions, such as: taking the loan application; ordering credit reports, appraisals, and title reports; and verifying the Borrower's income and employment. The Correspondent may or may not have delegated underwriting and typically funds the mortgage loans at settlement. The mortgage loan is closed in the Correspondent's name and the Correspondent may or may not service the mortgage loan. The Correspondent may use a Broker to perform some of the processing functions or even to fund the loan on its behalf; under such circumstances, the mortgage loan is considered a "Broker" third party origination mortgage loan, rather than a "Correspondent" third party origination mortgage loan. Retail Mortgage is a mortgage loan that is originated, underwritten and funded by a lender or its Affiliates. The mortgage loan is closed in the name of the lender or its Affiliate and if it is sold to Freddie Mac, it is sold by the lender or its Affiliate that originated it. A mortgage loan that a Broker or Correspondent completely or partially originated, processed, underwrote, packaged, funded or closed is not considered a Retail mortgage loan. For purposes of the definitions of Correspondent and Retail, "Affiliate" means any entity that is related to another party as a consequence of the entity, directly or indirectly, controlling the other party, being controlled by the other party, or being under common control with the other party.	• R = Retail • B = Broker • C = Correspondent • T = TPO Not Specified • Space = Unknown	Alpha	1

15	**PREPAYMENT PENALTY MORTGAGE (PPM) FLAG** - Denotes whether the mortgage is a PPM. A PPM is a mortgage with respect to which the borrower is, or at any time has been, obligated to pay a penalty in the event of certain repayments of principal.	• Y = PPM • N = Not PPM • Space (1) = Unknown	Alpha	1
16	**PRODUCT TYPE** - Denotes that the product is a fixed-rate mortgage.	• FRM – Fixed Rate Mortgage	Alpha	5
17	**PROPERTY STATE** - A two-letter abbreviation indicating the state or territory within which the property securing the mortgage is located.	• AL, TX, VA, etc.	Alpha	2
18	**PROPERTY TYPE** - Denotes whether the property type secured by the mortgage is a condominium, leasehold, planned unit development (PUD), cooperative share, manufactured home, or Single Family home. If the Property Type is unknown, this will be indicated by two (2) blank spaces.	• CO = Condo • LH = Leasehold • PU = PUD • MH = Manufactured Housing • SF = 1-4 Fee Simple • CP = Co-op • Space (2) = Unknown	Alpha	2
19	**POSTAL CODE** – The postal code for the location of the mortgaged property	• ###00, where "###" represents the first three digits of the 5-digit postal code • Space(5)= Unknown	Numeric	5
20	**LOAN SEQUENCE NUMBER** - Unique identifier assigned to each loan.	F1YYQnXXXXXX • F1 = product (Fixed Rate Mortgage); • YYQn = origination year and quarter; and, • XXXXXX = randomly assigned digits	Alpha-numeric	12
21	**LOAN PURPOSE** - Indicates whether the mortgage loan is a Cash-out Refinance mortgage, No Cash-out Refinance mortgage, or a Purchase mortgage. Generally, a Cash-out Refinance mortgage loan is a mortgage loan in which the use of the loan amount is not limited to specific purposes. A mortgage loan placed on a property previously owned free and clear by the Borrower is always considered a Cash-out Refinance mortgage loan. Generally, a No Cash-out Refinance mortgage loan is a mortgage loan in which the loan amount is limited to the following uses: Pay off the first mortgage, regardless of its age Pay off any junior liens secured by the mortgaged property, that were used in their entirety to acquire the subject property Pay related closing costs, financing costs and prepaid items, and Disburse cash out to the Borrower (or any other payee) not to exceed 2% of the new refinance mortgage loan or $2,000, whichever is less. As an exception to the above, for construction conversion mortgage loans and renovation mortgage loans, the amount of the interim construction financing secured by the mortgaged property is considered an amount used to pay off the first mortgage. Paying off unsecured liens or construction costs paid by the Borrower outside of the secured interim construction financing is considered cash out to the Borrower, if greater than $2000 or 2% of loan amount. This disclosure is subject to various special exceptions used by Sellers to determine whether a mortgage loan is a No Cash-out Refinance mortgage loan.	• P = Purchase • C = Cash-out Refinance • N = No Cash-out Refinance • Space = Unknown	Alpha	1

22	**ORIGINAL LOAN TERM** - A calculation of the number of scheduled monthly payments of the mortgage based on the First Payment Date and Maturity Date. Loans with original term of 420 or more, or 300 or less, are excluded from the Dataset if originated prior to 1/1/2005. If loan was originated on/after 1/1/2005, this exclusion does not apply.	• Calculation: (Loan Maturity Date (MM/YY) – Loan First Payment Date (MM/YY) + 1)	Numeric	3
23	**NUMBER OF BORROWERS** - The number of Borrower(s) who are obligated to repay the mortgage note secured by the mortgaged property. Disclosure denotes only whether there is one borrower or more than one borrower associated with the mortgage note. This disclosure will not be updated to reflect any subsequent assumption of the mortgage note.	• 01 = 1 borrower • 02 = > 1 borrowers • Space (2) = Unknown	Numeric	2
24	**SELLER NAME** - The entity acting in its capacity as a seller of mortgages to Freddie Mac at the time of acquisition. Seller Name will be disclosed for sellers with a total Original UPB representing 1% or more of the total Original UPB of all loans in the Dataset for a given calendar quarter. Otherwise, the Seller Name will be set to "Other Sellers".	Name of the seller, or "Other Sellers"	Alpha-numeric	20
25	**SERVICER NAME** - The entity acting in its capacity as the servicer of mortgages to Freddie Mac as of the last period for which loan activity is reported in the Dataset. Servicer Name will be disclosed for servicers with a total Original UPB representing 1% or more of the total Original UPB of all loans in the Dataset for a given calendar quarter. Otherwise, the Servicer Name will be set to "Other Servicers".	Name of the servicer, or "Other Servicers"	Alpha-numeric	20
26	**Super Conforming Flag** – For mortgages that exceed conforming loan limits with origination dates on or after 10/1/2008 and settlements on or after 1/1/2009	• Y = Yes • Space (1) = Not Super Conforming	Alpha	1
27	**Pre-HARP LOAN SEQUENCE NUMBER** – The Loan Sequence Number link that associates this HARP loan to the Pre-HARP Loan Sequence Number in the Single Family loan level dataset already being published on a quarterly basis.	F1YYQnXXXXXX • F1 = product (Fixed Rate Mortgage); • YYQn = origination year and quarter; and, • XXXXXX = randomly assigned digits	Alpha-numeric	12

MONTHLY PERFORMANCE DATA FILE

COLUMN POSITION	FORMAL NAME AND DEFINITION	VALID VALUES/ CALCULATIONS	TYPE	LENGTH
1	**LOAN SEQUENCE NUMBER** - Unique identifier assigned to each loan.	F1YYQnXXXXXX • F1 = product (Fixed Rate Mortgage); • YYQn = origination year and quarter; and, • XXXXXX = randomly assigned digits	Alpha-numeric	12

#		Field Description	Values/Calculation	Type	Max Length
2		**MONTHLY REPORTING PERIOD** – The as-of month for loan information contained in the loan record.	YYYYMM	Date	6
3		**CURRENT ACTUAL UPB** - The Current Actual UPB reflects the mortgage ending balance as reported by the servicer for the corresponding monthly reporting period. For fixed rate mortgages, this UPB is derived from the mortgage balance as reported by the servicer and includes any scheduled and unscheduled principal reductions applied to the mortgage. For mortgages with loan modifications, as indicated by "Y" in the Modification Flag field, the current actual unpaid principal balance may or may not include partial principal forbearance. If applicable, for loans with partial principal forbearance, the current actual unpaid principal balance equals the sum of interest bearing UPB (the amortizing principal balance of the mortgage) and the deferred UPB (the principal forbearance balance). Current UPB will be rounded to the nearest $1,000 for the first 6 months after origination date. This was previously reported as zero for the first 6 months after the origination date.	**Calculation:** (interest bearing UPB) + (non-interest bearing UPB)	Numeric Literal decimal	12
4		**CURRENT LOAN DELINQUENCY STATUS** – A value corresponding to the number of days the borrower is delinquent, based on the due date of last paid installment ("DDLPI") reported by servicers to Freddie Mac, and is calculated under the Mortgage Bankers Association (MBA) method. If a loan has been acquired by REO, then the Current Loan Delinquency Status will reflect the value corresponding to that status (instead of the value corresponding to the number of days the borrower is delinquent).	• XX = Unknown • 0 = Current, or less than 30 days past due • 1 = 30-59 days delinquent • 2 = 60 – 89 days delinquent • 3 = 90 – 119 days delinquent • And so on… • R = REO Acquisition • Space (3) = Unavailable	Alpha-numeric	3
5		**LOAN AGE** - The number of months since the note origination month of the mortgage. To ensure the age measurement commences with the first full month after the note origination month, subtract 1.	**Calculation:** ((Monthly Reporting Period) – Loan Origination Date (MM/YY)) – 1 month	Numeric	3
6		**REMAINING MONTHS TO LEGAL MATURITY** - The remaining number of months to the mortgage maturity date. For mortgages with loan modifications, as indicated by "Y" in the Modification Flag field, the calculation uses the modified maturity date.	**Calculation:** (Maturity Date (MM/YY) – Monthly Reporting Period (MM/YY)	Numeric	3
7		**REPURCHASE FLAG** - Indicates loans that have been repurchased or made whole (not inclusive of pool-level repurchase settlements). This field is only populated only at loan termination month.	• N = Not Repurchased • Y = Repurchased • Space (1) = Not Applicable	Alpha	1
8		**MODIFICATION FLAG** – For mortgages with loan modifications, indicates that the loan has been modified.	• Y = Yes • Space (1) = Not Modified	Alpha	1

			Type	Size
9	**ZERO BALANCE CODE** - A code indicating the reason the loan's balance was reduced to zero.	• 01 = Prepaid or Matured (Voluntary Payoff) • 03 = Foreclosure Alternative Group (Short Sale, Third Party Sale, Charge Off or Note Sale) • 06 = Repurchase prior to Property Disposition • 09 = REO Disposition • Space(2) = Not Applicable	Numeric	2
10	**ZERO BALANCE EFFECTIVE DATE** - The date on which the event triggering the Zero Balance Code took place.	• YYYYMM • Space(6) = Not Applicable	Date	6
11	**CURRENT INTEREST RATE** - Reflects the current interest rate on the mortgage note, taking into account any loan modifications.		Numeric Literal Decimal	8
12	**CURRENT DEFERRED UPB**: The current non-interest bearing UPB of the modified mortgage.	$ Amount. Non-Interest Bearing UPB.	Numeric	12
13	**DUE DATE OF LAST PAID INSTALLMENT (DDLPI)**: The due date that the loan's scheduled principal and interest is paid through, regardless of when the installment payment was actually made.	YYYYMM	Date	6
14	**MI RECOVERIES** - Mortgage Insurance Recoveries are proceeds received by Freddie Mac in the event of credit losses. These proceeds are based on claims under a mortgage insurance policy.	$ Amount. MI Recoveries.	Numeric Literal Decimal	12
15	**NET SALES PROCEEDS** - The amount remitted to Freddie Mac resulting from a property disposition once allowable selling expenses have been deducted from the gross sales proceeds of the property. A value of "C" in Net Sales Proceeds stands for Covered, which means that as part of the property disposition process, Freddie Mac was "Covered" for its total indebtedness (defined as UPB at disposition plus delinquent accrued interest) and net sale proceeds covered default expenses incurred by Servicer during the disposal of the loan. A value of "U" indicates that the amount is unknown.	$ Amount. Gross Sale Proceeds – Allowable Selling Expenses. C = Covered U = Unknown	Alpha-numeric Literal Decimal	14
16	**NON MI RECOVERIES**: Non-MI Recoveries are proceeds received by Freddie Mac based on repurchase/make whole proceeds, non-sale income such as refunds (tax or insurance), hazard insurance proceeds, rental receipts, positive escrow and/or other miscellaneous credits.	$ Amount. Non MI Recoveries.	Numeric Literal Decimal	12
17	**EXPENSES** - Expenses will include allowable expenses that Freddie Mac bears in the process of acquiring, maintaining and/ or disposing a property (excluding selling expenses, which are subtracted from gross sales proceeds to derive net sales proceeds). This is an aggregation of Legal Costs, Maintenance and Preservation Costs, Taxes and Insurance, and Miscellaneous Expenses	$ Amount. Allowable Expenses.	Numeric Literal Decimal	12

18	**Legal Costs** - The amount of legal costs associated with the sale of a property (but not included in Net Sale Proceeds). Prior to population of a Zero Balance Code equal to 03 or 09, this field will be populated as "Not Applicable," Following population of a Zero Balance Code equal to 03 or 09, this field will be updated (as applicable) to reflect the cumulative total. Space(12) – Not applicable	$ Amount	Numeric Literal Decimal	12
19	**Maintenance and Preservation Costs –** The amount of maintenance, preservation, and repair costs, including but not limited to property inspection, homeowner's association, utilities, and REO management, that is associated with the sale of a property (but not included in Net Sale Proceeds). Prior to population of a Zero Balance Code equal to 03 or 09, this field will be populated as "Not Applicable," Following population of a Zero Balance Code equal to 03 or 09, this field will be updated (as applicable) to reflect the cumulative total. Space(12) – Not applicable	$ Amount	Numeric Literal Decimal	12
20	**Taxes and Insurance –** The amount of taxes and insurance owed that are associated with the sale of a property (but not included in Net Sale Proceeds). Prior to population of a Zero Balance Code equal to 03 or 09, this field will be populated as "Not Applicable,". Following population of a Zero Balance Code equal to 03 or 09, this field will be updated (as applicable) to reflect the cumulative total. Space(12) – Not applicable	$ Amount	Numeric Literal Decimal	12
21	**Miscellaneous Expenses** - Miscellaneous expenses associated with the sale of a property (but not included in Net Sale Proceeds). Prior to population of a Zero Balance Code equal to 03 or 09, this field will be populated as "Not Applicable,". Following population of a Zero Balance Code equal to 03 or 09, this field will be updated (as applicable) to reflect the cumulative total.. Space(12) – Not applicable	$ Amount	Numeric Literal Decimal	12
22	**Actual Loss Calculation** Actual Loss was calculated using the below approach: Actual Loss = (Default UPB – Net Sale_Proceeds) + Delinquent Accrued Interest - Expenses – MI Recoveries – Non MI Recoveries. Delinquent Accrued Interest = (Default_Upb – Non Interest bearing UPB)* (Current Interest rate – 0.35) * (Months between Last Principal & Interest paid to date and zero balance date) * 30/360/100. Please note that the following business rules are applied to this calculation:	$ Amount	Numeric Literal Decimal	

	a. For all loans, 35 bps is used as a proxy for servicing fee. b. The Actual Loss Calculation will be set to zero for loans with Repurchase Flag ='Y' c. The Actual Loss Calculation will be set to zero for loans with Net Sale Proceeds='C (i.e, Covered)'. d. The Actual Loss Calculation will be set to zero for loans with Net Sales Proceeds = 'U" (Net Sales Proceeds are missing or expenses are not available. e. The Actual Loss Calculation will be set to missing for loans disposed within three months prior to the performance cutoff date. f. Modification Costs are currently not included in the calculation of The Actual Loss Calculation field.			
23	**Modification Cost** – The cumulative modification cost amount calculated when Freddie Mac determines such mortgage loan has experienced a rate modification event. Modification Cost is applicable for loans with rate changes only. This amount will be calculated on a monthly basis beginning with the first reporting period a modification event is reported, and disclosed in the last performance record. For example: (Original Interest Rate/1200 * Current Actual UPB) – (Current Interest Rate/1200 * (sum(Current Actual UPB, -Current Deferred UPB)) and aggregate each month since modification through the Performance Cutoff Date into a cumulative amount	$ Amount	Numeric Literal Decimal	12

Prior Work

Freddie Mac computers credit scores based on the given attributes in the data set, however, we have not found any published work on computed an enhanced credit score by merging such attributes with macroeconomic data and crime statistics. This, we believe, will give us an edge in producing a credit score that competes with the one's published in the Freddie Mac data set.

Proposed Work

- ○ Data cleaning: Some of the dates in the dataset are different, so we will need to standardize it
- ○ Data preprocessing: Will need to filter out some of the useless data columns
- ○ Data integration: Mixing the loans with crime statistics and economic indicators
- ○ Data regression: We will need to formulate regressions that will predict the probability of defaulting on loans. This can be done with hazard rate modeling, reduced form credit models, weight of evidence models, linear or logistic regression

Tools

- SciKit - Learn
- Pandas
- Numpy
- Python 3

Evaluation

Were we able to calculate loan default probability to a similar or greater degree than Freddie-Mac credit score?

Summary of our Project

Title: Credit Score and Default Risk

Team members: Spencer Hanson, William Brickowski, Ben Miller, Max Moghadam

- **Description:** Study of loan data, macroeconomic data, and crime data to come up with a credit score by correlating and using regression on the varying attributes with respect to loan default rates in hopes of coming up with more accurate credit scores than the one's computed by Freddie Mac.
- **Research Questions:**
  - How do societal and monetary factors, like GDP and crime rate, affect default rates?
  - What attributes make it likely that a person will default on their loan?
- **Prior Work:**
- **Datasets:**
  - Mac Single Family Loan-Level Dataset
    - http://www.freddiemac.com/research/datasets/sf_loanlevel_dataset.html
    - Origination Data File
    - Monthly Performance Data File
  - FBI Crime Data
    - https://ucr.fbi.gov/
  - **Census Data**
    - https://usa.ipums.org/usa/
- **Proposed work**: what do you need to do?
  - Data cleaning:
  - Data preprocessing:
  - Data integration:
  - Data regression and correlation: We will need to formulate regressions that will predict the probability of defaulting on loans. This can be done with hazard rate modeling, reduced form credit models, weight of evidence models, linear or logistic regression.
    - Correlation coefficient between default rate and varying attributes to determine a weighted credit score
    - Probit & Logit regression models to predict the probability of defaulting on loan
- **List of tool(s):** Python, Pandas, Numpy, SciKit-Learn
- **Evaluation:** Were we able to calculate loan default probability to a similar or greater degree than Freddie-Mac credit score?