# Final Project Description:

Below is the definition of the Freddie Mac data for loans Single Family homes/condos. The first file called ‘ORIGINATION DATA FILE’ has attributes that are considered to make the loan. For example, the Credit score, occupancy status, loan-to-value ratio etc. The second file/table called ‘MONTHLY PERFORMANCE DATA FILE’ contains data regarding the performance of the loans. Data is available for years between 1999 to 2017. I am planning on using just data from 1999 for this project. That by itself will provide over 300,000 rows of loans data.

Question: Can we use the attributes in the ‘Origination Data file’ and the data provided in the ‘Monthly Performance Data file’ to ask one or both of the following questions,

* Which loan will default?
* What is the probability of a loan defaulting?

**ORIGINATION DATA FILE:**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Column Position | Name | Valid Values | Type | Length |
| 1 | Credit Score | • 301 - 850 • 9999 = Not Available, if Credit Score is < 301 or > 850. | Numeric | 4 |
| 2 | FIRST PAYMENT DATE | YYYYMM | Date | 6 |
| 3 | FIRST TIME HOMEBUYER FLAG | • Y = Yes  • N = No  • 9 = Not Available or Not Applicable | Alpha | 1 |
| 4 | MATURITY DATE | YYYYMM | Date | 6 |
| 5 | METROPOLITAN STATISTICAL AREA (MSA) OR METROPOLITAN DIVISION | • 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 %) | • 1% - 55%  • 000 = No MI  • 999 = Not Available | Numeric | 3 |
| 7 | NUMBER OF UNITS | • 1 = one-unit  • 2 = two-unit  • 3 = three-unit  • 4 = four-unit  • 99 = Not Available | Numeric | 2 |
| 8 | OCCUPANCY STATUS | • P = Primary Residence  • I = Investment Property  • S = Second Home • 9 = Not Available | Alpha | 1 |
| 9 | ORIGINAL COMBINED LOAN-TO-VALUE (CLTV) | • 0% - 200%  • 999 = Not Available | Numeric Literal Decimal | 7 |
| 10 | ORIGINAL DEBT-TO-INCOME (DTI) RATIO | • 0%<DTI<=65%  • 999 = Not Available HARP ranges:  • 999 = Not Available | Numeric | 3 |
| 11 | ORIGINAL UPB | Amount will be rounded to the nearest $1,000. | Numeric | 12 |
| 12 | ORIGINAL LOAN-TO-VALUE (LTV) | • 6% - 105%  • 999 = Not Available HARP ranges:  • 81% - 999%  • 999 = Not Available | Numeric | 3 |
| 13 | ORIGINAL INTEREST RATE |  | Numeric Literal decimal | 6 |
| 14 | CHANNEL | • R = Retail  • B = Broker  • C = Correspondent  • T = TPO Not Specified  • 9 = Not Available | Alpha | 1 |
| 15 | PREPAYMENT PENALTY MORTGAGE (PPM) FLAG | • Y = PPM  • N = Not PPM | Alpha | 1 |
| 16 | PRODUCT TYPE | FRM – Fixed Rate Mortgage | Alpha | 5 |
| 17 | PROPERTY STATE | • AL, TX, VA, etc. | Alpha | 2 |
| 18 | PROPERTY TYPE | • CO = Condo  • PU = PUD  • MH = Manufactured Housing  • SF = 1-4 Fee Simple  • CP = Co-op  • 99 = Not Available | Alpha | 2 |
| 19 | POSTAL CODE | • ###00, where “###” represents the first three digits of the 5- digit postal code  ORIGINAL LOAN TERM  • Space(5)= Unknown | Numeric | 5 |
| 20 | LOAN SEQUENCE NUMBER | F1YYQnXXXXXX • F1 = product (Fixed Rate Mortgage);  • YYQn = origination year and quarter; and,  • XXXXXX = randomly assigned digits | Alpha-numeric | 12 |
| 21 | LOAN PURPOSE | • P = Purchase  • C = Cash-out Refinance  • N = No Cash-out Refinance  • 9 =Not Available | Alpha | 1 |
| 22 | ORIGINAL LOAN TERM | Calculation: (Loan Maturity Date (MM/YY) – Loan First Payment Date (MM/YY) + 1) | Numeric | 3 |

**MONTHLY PERFORMANCE DATA FILE:**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Column Position | Name | Valid Values | Type | Length |
| 1 | LOAN SEQUENCE NUMBER | F1YYQnXXXXXX • F1 = product (Fixed Rate Mortgage);  • YYQn = origination year and quarter; and,  • XXXXXX = randomly assigned digits | Alpha-numeric | 12 |
| 2 | MONTHLY REPORTING PERIOD | YYYYMM | Date | 6 |
| 3 | CURRENT ACTUAL UPB | Calculation: (interest bearing UPB) + (noninterest bearing UPB) | Numeric Literal decimal | 12 |
| 4 | CURRENT LOAN DELINQUENCY STATUS | • 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 | Alphanumeric | 3 |
| 5 | LOAN AGE | Calculation: ((Monthly Reporting Period) – Loan Origination Date (MM/YY)) – 1 month | Numeric | 3 |
| 6 | REMAINING MONTHS TO LEGAL MATURITY | Calculation: (Maturity Date (MM/YY) – Monthly Reporting Period (MM/YY) | Numeric | 3 |
| 7 | REPURCHASE FLAG | • N = Not Repurchased  • Y = Repurchased • Space (1) = Not Applicable | Alpha | 1 |
| 8 | MODIFICATION FLAG | • Y = Yes  • Space (1) = Not Modified | Alpha | 1 |
| 9 | ZERO BALANCE CODE | • 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 | • YYYYMM  • Space(6) = Not Applicable | Date | 6 |
| 11 | CURRENT INTEREST RATE |  | Numeric Literal Decimal | 8 |
| 12 | CURRENT DEFERRED UPB | $ Amount. Non-Interest Bearing UPB. | Numeric | 12 |
| 13 | DUE DATE OF LAST PAID INSTALLMENT (DDLPI) | YYYYMM | Date | 6 |
| 14 | MI RECOVERIES | $ Amount. MI Recoveries. | Numeric Literal Decimal | 12 |
| 15 | NET SALES PROCEEDS | $ Amount. Gross Sale Proceeds – Allowable Selling Expenses. C = Covered U = Unknown | Alphanumeric Literal Decimal | 14 |
| 16 | NON MI RECOVERIES | $ Amount. Non MI Recoveries. | Numeric Literal Decimal | 12 |
| 17 | EXPENSES | $ Amount. Allowable Expenses. | Numeric Literal Decimal | 12 |
| 18 | LEGAL COSTS | $ Amount | Numeric Literal Decimal | 12 |
| 19 | MAINTENANCE AND PRESERVATION COSTS | $ Amount | Numeric Literal Decimal | 12 |
| 20 | TAXES AND INSURANCE | $ Amount | Numeric Literal Decimal | 12 |
| 21 | MISCELLANEOUS EXPENSES | $ Amount | Numeric Literal Decimal | 12 |
| 22 | ACTUAL LOSS CALCULATION | $ Amount | Numeric Literal Decimal | 12 |
| 23 | MODIFICATION COST | $ Amount | Numeric Literal Decimal | 12 |