

# Loss Data Analysis

## Tutorial 101:

Improve your understanding of Fannie Mae's credit risk performance data

**July 27, 2015**

# Disclaimer

The information provided in this presentation is intended to provide an introduction to the analysis and understanding of single family mortgage loan performance data published by Fannie Mae. The tools and methods presented in the tutorials are not intended to provide comprehensive instruction as to the use and analysis of the data, and may not reveal trends in performance or other information that may be meaningful to particular users of the data. Data users, together with their financial and other advisors, must determine for themselves the most appropriate methods to use to analyze the data and should ensure they are comfortable with the sufficiency of such analysis before using the data to assist in making investment decisions. Fannie Mae shall have no liability for any errors or misunderstandings resulting from misapplication of the information presented in the tutorials. Data users should also note that all loan performance data referenced in the tutorials is historical performance data and may not be predictive of future loan performance.

## Loss Data Webinar Series

- These tutorials are an accompaniment to our latest loan performance data release.
- They are designed to share best practices and methodology for using the data so that users can interpret it appropriately.

### **Tutorial 101**

Focuses on new loss data elements and how to prepare data for analysis and create statistical summaries for analysis.

### **Tutorial 102**

Builds off concepts featured in Tutorial 101 and introduces more advanced modeling and analysis methodology.

# Tutorial 101 Objectives

<b>1</b>	<b>Program Background</b>	Overview of loan performance data – background on program and the new data elements added to release
<b>2</b>	<b>Getting Started</b>	Quick walk-through and Q&A on process for finding, downloading data – and using an analysis application
<b>3</b>	<b>Creating a Dataset</b>	Steps involved in creating a unified dataset that combines variables from acquisition and performance data files into single dataset
<b>4</b>	<b>Creating Summary Tables</b>	Process for creating statistical summary tables that can be used to verify accurate capture of data

# Background



- In support of its credit risk sharing programs Fannie Mae released an extensive dataset beginning in 2013 that provides insight into the credit performance of a portion of Fannie Mae's single-family book of business.
- Now consisting of nearly 22 million records, the dataset provides monthly loan-level detail and is offered to help investors gain a better understanding of the credit performance of a portion of single-family loans owned or guaranteed by Fannie Mae.
- The public dataset includes a subset of Fannie Mae's 30-year, fixed-rate, fully documented, single-family amortizing loans that the company owned or guaranteed on or after January 1, 2000.

**In July of 2015, Fannie Mae enhanced the data offering to include additional credit performance data that now enables investors to model credit risk from loan acquisition through property disposition.**

*\*As with all historic datasets, past performance is no guarantee of future results.*

# Acquisition Statistical Summary Table



Origination Year	Loan Count	Total Orig. UPB (\$M)	Avg. Orig UPB (\$)	Acquisition Characteristics <sup>1</sup>						
				Borrower Credit Score	Co-Borrower Credit Score	LTV Ratio	CLTV Ratio <sup>2</sup>	DTI	Note Rate	
1999	127,179	\$ 15,948	\$ 125,402	716	724	79.5	79.5	34.9	7.80	
2000	1,070,195	\$ 140,963	\$ 131,717	718	726	79.1	79.3	35.7	8.13	
2001	2,346,511	\$ 349,702	\$ 149,031	719	726	75.4	75.8	34.0	6.99	
2002	2,390,308	\$ 374,469	\$ 156,661	723	730	72.9	73.4	34.0	6.50	
2003	3,009,007	\$ 497,095	\$ 165,202	725	732	70.8	71.5	33.5	5.75	
2004	1,192,756	\$ 200,864	\$ 168,403	721	728	72.1	73.7	36.6	5.84	
2005	1,131,259	\$ 208,483	\$ 184,293	725	732	71.2	73.3	38.2	5.84	
2006	894,631	\$ 172,510	\$ 192,828	724	732	71.8	73.9	39.3	6.42	
2007	1,063,517	\$ 218,063	\$ 205,039	724	732	73.4	75.5	39.4	6.36	
2008	1,181,458	\$ 262,804	\$ 222,441	744	752	73.4	75.0	38.5	6.04	
2009	1,756,148	\$ 417,064	\$ 237,488	764	770	67.9	69.4	34.2	4.97	
2010	1,198,294	\$ 295,123	\$ 246,286	768	774	69.6	71.0	32.8	4.72	
2011	1,002,875	\$ 235,291	\$ 234,617	767	774	71.7	73.0	33.1	4.55	
2012	1,711,293	\$ 417,948	\$ 244,229	770	775	71.6	72.9	31.9	3.84	
2013	1,522,367	\$ 356,118	\$ 233,924	762	768	75.2	76.3	33.4	4.05	
2014	347,254	\$ 76,291	\$ 219,698	752	759	78.8	79.6	34.9	4.62	
Total	21,945,052	\$ 4,238,736	\$ 193,152	742	751	72.4	73.6	34.7	5.52	

<sup>1</sup> Acquisition Characteristics are UPB-weighted averages, based on UPB at origination

<sup>2</sup> Missing CLTVs have been set to OLTV in this view

This table shows the acquisition profile of loans and can be used to identify shifts in key stats over time.

# Performance Statistical Summary Table



			Active Loans		Inactive Loans (Loan Count)				Total Mods to Date <sup>2</sup>				
Origination Year	Loan Count	Total Orig. UPB (\$M)	Loan Count (Active)	Active UPB (\$M)	Prepaid	Repurchased <sup>1</sup>	Alternative Disposition	REO Disposition	Loan Count	D180 UPB (\$M) <sup>3,4</sup>	D180 % of Orig. UPB <sup>3,4</sup>	Default UPB (\$M) <sup>5</sup>	Loss Rate (%) <sup>5</sup>
1999	127,179	\$ 15,948	2,719	\$ 155	122,135	613	319	1,393	846	\$ 298	1.9%	\$ 155	0.1%
2000	1,070,195	\$ 140,963	14,955	\$ 875	1,038,110	3,159	2,079	11,892	6,237	\$ 2,148	1.5%	\$ 1,262	0.1%
2001	2,346,511	\$ 349,702	65,824	\$ 4,796	2,246,209	3,921	4,537	26,020	16,127	\$ 5,062	1.4%	\$ 2,945	0.2%
2002	2,390,308	\$ 374,469	134,293	\$ 11,574	2,217,026	3,643	5,595	29,751	21,119	\$ 6,380	1.7%	\$ 3,388	0.3%
2003	3,009,007	\$ 497,095	381,116	\$ 39,448	2,565,274	4,620	12,134	45,863	44,869	\$ 13,882	2.8%	\$ 6,410	0.4%
2004	1,192,756	\$ 200,864	186,442	\$ 21,340	964,746	2,388	9,108	30,072	30,716	\$ 9,960	5.0%	\$ 4,883	0.9%
2005	1,131,259	\$ 208,483	207,233	\$ 28,230	854,703	2,860	19,873	46,590	50,970	\$ 18,810	9.0%	\$ 10,694	2.2%
2006	894,631	\$ 172,510	143,919	\$ 21,323	677,368	3,345	21,113	48,886	56,261	\$ 21,083	12.2%	\$ 12,264	3.2%
2007	1,063,517	\$ 218,063	190,308	\$ 30,673	777,176	8,246	26,149	61,638	81,875	\$ 30,672	14.1%	\$ 16,227	3.0%
2008	1,181,458	\$ 262,804	193,083	\$ 30,538	927,295	8,617	15,130	37,329	55,681	\$ 20,374	7.8%	\$ 9,580	1.1%
2009	1,756,148	\$ 417,064	558,385	\$ 94,986	1,184,564	2,298	3,233	7,668	10,596	\$ 4,646	1.1%	\$ 1,811	0.1%
2010	1,198,294	\$ 295,123	546,249	\$ 99,273	648,333	1,045	701	1,966	3,075	\$ 1,296	0.4%	\$ 347	0.0%
2011	1,002,875	\$ 235,291	562,564	\$ 104,244	438,784	469	244	814	1,846	\$ 665	0.3%	\$ 109	0.0%
2012	1,711,293	\$ 417,948	1,414,197	\$ 315,862	295,613	968	114	401	888	\$ 459	0.1%	\$ 38	0.0%
2013	1,522,367	\$ 356,118	1,361,094	\$ 300,648	158,495	2,591	43	144	267	\$ 259	0.1%	\$ 8	0.0%
2014	347,254	\$ 76,291	307,343	\$ 63,298	39,566	340	1	4	3	\$ 22	0.0%	\$ -	0.0%
Total	21,945,052	\$ 4,238,736	6,269,724	\$ 1,167,262	15,155,401	49,123	120,373	350,431	381,376	\$ 136,017	3.2%	\$ 70,121	0.6%

<sup>1</sup> Reflects loans repurchased up to and after 180 days of delinquency. Previous versions of the Statistical Summary reflected in this column only loans repurchased prior to the occurrence of a credit event.

<sup>2</sup> Only one modification is counted per loan.

<sup>3</sup> D180 Rates included here are calculated in the same methodology as prior statistical summaries, they are included for comparison purposes only.

<sup>4</sup> Reflects the outstanding available UPB at D180 as reflected in the dataset.

<sup>5</sup> Default rates and UPB in this view are for completed foreclosures only. These are defined as loans with a zero balance code of '09' or '03' and non-null disposition dates.

This example aggregates balance and performance information of loans and identifies loans that are active, in a delinquency state, prepaid, or repurchased.

# Loss/Severity Statistical Summary Table

By Origination Year



	Origination Year															Total
	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	
Default UPB (\$M)	\$155	\$1,262	\$2,945	\$3,388	\$6,410	\$4,883	\$10,694	\$12,264	\$16,227	\$9,580	\$1,811	\$347	\$109	\$38	\$8	\$70,121
Default Rate (%)	1.0%	0.9%	0.8%	0.9%	1.3%	2.4%	5.1%	7.1%	7.4%	3.7%	0.4%	0.1%	0.1%	0.0%	0.0%	1.7%
<b>EXPENSES:</b>																
Delinquent Interest	12%	12%	11%	10%	10%	10%	9%	10%	10%	9%	7%	6%	6%	4%	3%	10%
Total Liquidation Exp.	10%	11%	11%	11%	11%	10%	8%	7%	7%	7%	7%	7%	7%	6%	4%	8%
Foreclosure	4%	5%	4%	4%	4%	4%	3%	3%	2%	2%	2%	3%	2%	2%	2%	3%
Property Preservation	3%	3%	3%	3%	3%	2%	2%	2%	2%	1%	2%	2%	2%	2%	1%	2%
Asset Recovery	1%	1%	1%	1%	1%	1%	1%	1%	1%	1%	1%	1%	1%	1%	1%	1%
Misc. Holding Expenses/Credits	1%	1%	1%	1%	1%	1%	1%	1%	1%	1%	0%	0%	0%	0%	0%	1%
Associated Taxes	2%	2%	2%	3%	3%	3%	2%	2%	2%	2%	2%	2%	2%	1%	1%	2%
Total Costs	122%	123%	121%	122%	121%	120%	117%	118%	117%	117%	114%	114%	112%	110%	107%	118%
<b>PROCEEDS:</b>																
Net Sales Proceeds	81%	77%	74%	72%	76%	72%	66%	61%	60%	63%	76%	81%	83%	82%	91%	66%
Credit Enhancement	17%	17%	14%	12%	9%	8%	6%	6%	8%	10%	5%	6%	8%	9%	10%	8%
Repurchase/Make Whole	6%	8%	5%	4%	2%	2%	3%	5%	8%	12%	8%	6%	4%	1%	0%	6%
Other	4%	4%	5%	4%	3%	2%	1%	1%	1%	1%	1%	1%	1%	1%	1%	2%
Total Proceeds	109%	107%	98%	93%	89%	84%	75%	72%	78%	86%	89%	93%	96%	93%	101%	82%
Severity	12.9%	15.9%	23.3%	29.2%	31.4%	36.4%	42.0%	45.5%	39.9%	30.4%	24.6%	20.5%	16.6%	17.2%	5.3%	36.6%
Total Net Loss (\$M)	\$20	\$201	\$685	\$989	\$2,012	\$1,775	\$4,494	\$5,576	\$6,474	\$2,908	\$445	\$71	\$18	\$7	\$0	\$25,676

This example aggregates information related to defaulted loans in the dataset organized by origination vintage. It shows expense line items, proceeds line items, severity percentage, and total net loss for a given vintage.



# Loss/Severity Statistical Summary Table

By Disposition Year



	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	Total
Default UPB (\$M)	\$1	\$69	\$340	\$826	\$1,181	\$1,197	\$1,130	\$1,337	\$2,139	\$5,364	\$11,082	\$15,764	\$13,429	\$9,404	\$6,859	\$70,121
<b>EXPENSES:</b>																
Delinquent Interest	5%	8%	9%	9%	9%	9%	10%	9%	8%	7%	8%	9%	10%	12%	15%	10%
Total Liquidation Exp.	3%	6%	6%	7%	7%	9%	10%	9%	7%	6%	6%	7%	8%	11%	14%	8%
Foreclosure	1%	2%	3%	3%	3%	4%	4%	4%	3%	2%	2%	2%	3%	3%	4%	3%
Property Preservation	1%	2%	1%	1%	2%	2%	3%	2%	2%	1%	1%	2%	2%	2%	3%	2%
Asset Recovery	1%	0%	1%	0%	0%	1%	1%	0%	0%	0%	1%	1%	1%	1%	1%	1%
Misc. Holding Expenses/Credits	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	1%	1%	1%	2%	1%
Associated Taxes	0%	1%	1%	1%	1%	2%	2%	2%	2%	2%	2%	2%	2%	3%	4%	2%
Total Costs	108%	113%	115%	115%	116%	118%	119%	118%	115%	113%	114%	116%	118%	123%	129%	118%
<b>PROCEEDS:</b>																
Net Sales Proceeds	101%	76%	80%	77%	78%	79%	78%	76%	71%	63%	62%	59%	64%	71%	74%	66%
Credit Enhancement	7%	16%	16%	16%	15%	15%	16%	14%	11%	9%	8%	7%	7%	7%	7%	8%
Repurchase/Make Whole	0%	15%	12%	13%	10%	8%	5%	3%	4%	8%	9%	8%	5%	3%	1%	6%
Other	0%	4%	4%	4%	5%	5%	5%	5%	2%	1%	1%	1%	1%	2%	2%	2%
Total Proceeds	109%	111%	112%	110%	108%	108%	104%	99%	89%	81%	79%	75%	77%	82%	84%	82%
Severity	-1.1%	2.2%	3.2%	5.8%	7.2%	10.1%	15.1%	19.3%	26.5%	31.8%	34.7%	41.2%	40.9%	40.3%	45.2%	36.6%
Total Net Loss (\$M)	(\$0)	\$2	\$11	\$48	\$84	\$121	\$170	\$258	\$566	\$1,706	\$3,843	\$6,491	\$5,490	\$3,788	\$3,098	\$25,676

This example aggregates information related to defaulted loans in the dataset organized by disposition year.

# Where to begin?



## 1 Identify preferred analysis software

- For these webinars, Fannie Mae will use R
- Download R at: [www.r-project.org](http://www.r-project.org) and the supporting user interface Rstudio at: [www.rstudio.org](http://www.rstudio.org)

## 2 Download the latest analysis code or script

- For these webinars Fannie Mae created a script for R as well as SAS
- These are posted on our website here:  
[www.fanniemae.com/loanperformance](http://www.fanniemae.com/loanperformance)


## 3 Access and download the data

- Navigate from the Fannie Mae website at:  
[www.fanniemae.com/loanperformance](http://www.fanniemae.com/loanperformance)
- Or go directly to the Loan Performance Data Application at:  
[loanperformancedata.fanniemae.com/lppub/index.html](http://loanperformancedata.fanniemae.com/lppub/index.html)

# Download the data



The data is grouped according to Fannie Mae's acquisition year and performance quarter. For each available period, two files exist; one static file for loan acquisition characteristics and one dynamic file for performance activity.

Loan Performance Data


Home
Download

Loan Performance Data Files

A statistical summary of the loan and performance data and a representation of the risk profiles are available [here](#).  
 To facilitate the download of the data, the files do not include column headings. The [file layout](#) and sample files ([Acquisitions](#) and [Performance](#)) are provided here for your reference.

Year	Q1 Records		Q2 Records		Q3 Records		Q4 Records	
2000	<a href="#">Acquisitions</a>	<a href="#">Performance</a>	<a href="#">Acquisition</a>	<a href="#">Performance</a>	<a href="#">Acquisition</a>	<a href="#">Performance</a>	<a href="#">Acquisition</a>	<a href="#">Performance</a>
2001	<a href="#">Acquisitions</a>	<a href="#">Performance</a>	<a href="#">Acquisition</a>	<a href="#">Performance</a>	<a href="#">Acquisition</a>	<a href="#">Performance</a>	<a href="#">Acquisition</a>	<a href="#">Performance</a>
2002	<a href="#">Acquisitions</a>	<a href="#">Performance</a>	<a href="#">Acquisition</a>	<a href="#">Performance</a>	<a href="#">Acquisition</a>	<a href="#">Performance</a>	<a href="#">Acquisition</a>	<a href="#">Performance</a>
2003	<a href="#">Acquisitions</a>	<a href="#">Performance</a>	<a href="#">Acquisition</a>	<a href="#">Performance</a>	<a href="#">Acquisition</a>	<a href="#">Performance</a>	<a href="#">Acquisition</a>	<a href="#">Performance</a>
2004	<a href="#">Acquisitions</a>	<a href="#">Performance</a>	<a href="#">Acquisition</a>	<a href="#">Performance</a>	<a href="#">Acquisition</a>	<a href="#">Performance</a>	<a href="#">Acquisition</a>	<a href="#">Performance</a>
2005	<a href="#">Acquisitions</a>	<a href="#">Performance</a>	<a href="#">Acquisition</a>	<a href="#">Performance</a>	<a href="#">Acquisition</a>	<a href="#">Performance</a>	<a href="#">Acquisition</a>	<a href="#">Performance</a>
2006	<a href="#">Acquisitions</a>	<a href="#">Performance</a>	<a href="#">Acquisition</a>	<a href="#">Performance</a>	<a href="#">Acquisition</a>	<a href="#">Performance</a>	<a href="#">Acquisition</a>	<a href="#">Performance</a>
2007	<a href="#">Acquisitions</a>	<a href="#">Performance</a>	<a href="#">Acquisition</a>	<a href="#">Performance</a>	<a href="#">Acquisition</a>	<a href="#">Performance</a>	<a href="#">Acquisition</a>	<a href="#">Performance</a>
2008	<a href="#">Acquisitions</a>	<a href="#">Performance</a>	<a href="#">Acquisition</a>	<a href="#">Performance</a>	<a href="#">Acquisition</a>	<a href="#">Performance</a>	<a href="#">Acquisition</a>	<a href="#">Performance</a>
2009	<a href="#">Acquisitions</a>	<a href="#">Performance</a>	<a href="#">Acquisition</a>	<a href="#">Performance</a>	<a href="#">Acquisition</a>	<a href="#">Performance</a>	<a href="#">Acquisition</a>	<a href="#">Performance</a>
2010	<a href="#">Acquisitions</a>	<a href="#">Performance</a>	<a href="#">Acquisition</a>	<a href="#">Performance</a>	<a href="#">Acquisition</a>	<a href="#">Performance</a>	<a href="#">Acquisition</a>	<a href="#">Performance</a>
2011	<a href="#">Acquisitions</a>	<a href="#">Performance</a>	<a href="#">Acquisition</a>	<a href="#">Performance</a>	<a href="#">Acquisition</a>	<a href="#">Performance</a>	<a href="#">Acquisition</a>	<a href="#">Performance</a>
2012	<a href="#">Acquisitions</a>	<a href="#">Performance</a>	<a href="#">Acquisition</a>	<a href="#">Performance</a>	<a href="#">Acquisition</a>	<a href="#">Performance</a>	<a href="#">Acquisition</a>	<a href="#">Performance</a>
2013	<a href="#">Acquisitions</a>	<a href="#">Performance</a>	<a href="#">Acquisition</a>	<a href="#">Performance</a>	<a href="#">Acquisition</a>	<a href="#">Performance</a>	<a href="#">Acquisition</a>	<a href="#">Performance</a>
2014	<a href="#">Acquisitions</a>	<a href="#">Performance</a>	<a href="#">Acquisition</a>	<a href="#">Performance</a>	Not Available	Not Available	Not Available	Not Available

For questions regarding the data, contact the Fixed-Income Marketing at 1-800-237-8627, 9:00 a.m. to 5:30 p.m., ET, every business day.

To create a complete dataset for analysis, users should download all newly released public files.

- Acquisition files, for the past three years, are also updated quarterly – download these every quarter as well
- Performance files are updated quarterly – download all files every quarter

# General characteristics of the data

**Data are a subset of Fannie Mae's single-family, conventional mortgages and includes the following attributes *only*:**

- 30-year fixed-rate
- Fully amortizing
- Fully documented
- Acquired by Fannie Mae on or after January 1, 2000

## **Data does not include:**

- Home Affordable Refinance Program (HARP) mortgage loans
- Refi Plus™ mortgage loans
- Mortgage loans with LTVs greater than 97%
- Alt-A or other mortgage loans with reduced documentation and/or streamlined processing
- Loans sold to Fannie Mae with lender recourse or subject to certain other third-party, risk-sharing arrangements (other than primary mortgage insurance)
- Loans acquired under certain programs or negotiated variances that are no longer eligible for delivery to Fannie Mae
- Adjustable-Rate mortgage loans
- Balloon mortgage loans
- Interest-only mortgage loans
- Mortgage loans with prepayment penalties
- Government insured mortgage loans
- Loans sold to Fannie Mae on a negotiated bulk basis

# General characteristics of the data

- “Acquisition” file includes static mortgage loan data at the time of the mortgage loan’s origination and delivery to Fannie Mae.
- 1 row per loan
- “Performance” file provides monthly performance data for each loan, from acquisition up until its current status as of the previous quarter.
- 1 row per month of loan activity per loan

## Fields included in the **acquisition** file:

POSITION	FIELD NAME	TYPE
1	LOAN IDENTIFIER	ALPHA-NUMERIC
2	CHANNEL	ALPHA-NUMERIC
3	SELLER NAME	ALPHA-NUMERIC
4	ORIGINAL INTEREST RATE	NUMERIC
5	ORIGINAL UNPAID PRINCIPAL BALANCE (UPB)	NUMERIC
6	ORIGINAL LOAN TERM	NUMERIC
7	ORIGINATION DATE	DATE
8	FIRST PAYMENT DATE	DATE
9	ORIGINAL LOAN-TO-VALUE (LTV)	NUMERIC
10	ORIGINAL COMBINED LOAN-TO-VALUE (CLTV)	NUMERIC
11	NUMBER OF BORROWERS	NUMERIC
12	DEBT-TO-INCOME RATIO (DTI)	NUMERIC
13	BORROWER CREDIT SCORE	NUMERIC
14	FIRST-TIME HOME BUYER INDICATOR	ALPHA-NUMERIC
15	LOAN PURPOSE	ALPHA-NUMERIC
16	PROPERTY TYPE	ALPHA-NUMERIC
17	NUMBER OF UNITS	ALPHA-NUMERIC
18	OCCUPANCY STATUS	ALPHA-NUMERIC
19	PROPERTY STATE	ALPHA-NUMERIC
20	ZIP (3-DIGIT)	ALPHA-NUMERIC
21	MORTGAGE INSURANCE PERCENTAGE	NUMERIC
22	PRODUCT TYPE	ALPHA-NUMERIC
23	CO-BORROWER CREDIT SCORE	NUMERIC

## Fields included in the **performance** file:

POSITION	FIELD NAME	TYPE
1	LOAN IDENTIFIER	ALPHA-NUMERIC
2	MONTHLY REPORTING PERIOD	DATE
3	SERVICER NAME	ALPHA-NUMERIC
4	CURRENT INTEREST RATE	NUMERIC
5	CURRENT ACTUAL UNPAID PRINCIPAL BALANCE (UPB)	NUMERIC
6	LOAN AGE	NUMERIC
7	REMAINING MONTHS TO LEGAL MATURITY	NUMERIC
8	ADJUSTED REMAINING MONTHS TO MATURITY	NUMERIC
9	MATURITY DATE	DATE
10	METROPOLITAN STATISTICAL AREA (MSA)	ALPHA-NUMERIC
11	CURRENT LOAN DELINQUENCY STATUS	ALPHA-NUMERIC
12	MODIFICATION FLAG	ALPHA-NUMERIC
13	ZERO BALANCE CODE	ALPHA-NUMERIC
14	ZERO BALANCE EFFECTIVE DATE	DATE
15	LAST PAID INSTALLMENT DATE	DATE
16	FORECLOSURE DATE	DATE
17	DISPOSITION DATE	DATE
18	FORECLOSURE COSTS	NUMERIC
19	PROPERTY PRESERVATION AND REPAIR COSTS	NUMERIC
20	ASSET RECOVERY COSTS	NUMERIC
21	MISCELLANEOUS HOLDING EXPENSES AND CREDITS	NUMERIC
22	ASSOCIATED TAXES FOR HOLDING PROPERTY	NUMERIC
23	NET SALE PROCEEDS	NUMERIC
24	CREDIT ENHANCEMENT PROCEEDS	NUMERIC
25	REPURCHASE MAKE WHOLE PROCEEDS	NUMERIC
26	OTHER FORECLOSURE PROCEEDS	NUMERIC

**Note:** Loans are not reported on after liquidation

# Defining our combined dataset

- Create a dataset ready for analysis → **one loan per row**
  - Static variables from the acquisition files
    - Loan characteristics
    - Borrower characteristics
  - Monthly and “last status” variables from the performance files
    - Flags for intermediate performance milestones (e.g. date of first 180 delinquency)
    - Values from the last, most-recent row (e.g. zero balance code or current status)

A combined dataset compresses relevant information from many rows of loan activity into a single row to create efficient storage of data for calculating summary statistics and basic modeling of loan outcomes.

- The following examples use data from Q1 2006 acquisition and performance files.
- The entire process is included in the R and SAS scripts posted on the website.



*By applying simple logic and some basic manipulation, we can create a dataset that will be both smaller, and easier to work with.*

# From files to dataset

By keeping only the pertinent fields, we can create a population or “pop” dataset that reflects dynamic activity through a specific activity reporting period and removes the data we do not need.

LOAN IDENTIFIER	CHANNEL	SELLER NAME	ORIGINAL INTEREST RATE	ORIGINAL UNPAID PRINCIPAL BALANCE (UPB)	ORIGINAL LOAN TERM	ORIGINATION DATE	FIRST PAYMENT DATE	ORIGINAL LOAN-TO-VALUE (LTV)	ORIGINAL COMBINED LOAN-TO-VALUE (LTV)	NUMBER OF BORROWERS	DEBT-TO-INCOME RATIO	BORROWER CREDIT SCORE	FIRST-TIME HOME BUYER INDICATOR	LOAN PURPOSE	PROPERTY TYPE	NUMBER OF UNITS	OCCUPANCY STATUS	PROPERTY STATE	ZIP (3-DIGIT)	MORTGAGE INSURANCE PERCENTAGE	PRODUCT TYPE	CO BORROWER CREDIT SCORE	SERVICER NAME	CURRENT INTEREST RATE	CURRENT ACTUAL UNPAID PRINCIPAL BALANCE (UPB)	LOAN AGE	CURRENT MONTHS TO LEGAL MATURITY	ADJUSTED REMAINING MONTHS TO MATURITY	MATURITY DATE	METROPOLITAN STATISTICAL AREA	CURRENT LOAN DELINQUENCY STATUS	MODIFICATION FLAG	ZERO BALANCE CODE	ZERO BALANCE EFFECTIVE DATE	MINIMUM CREDIT SCORE	ORIGINAL HOME VALUE	DELINQUENT INTEREST EXPENSE	FIRST CREDIT EVENT DATE	FIRST CREDIT EVENT UPB	FIRST 180 DAY DELINQUENCY DATE	FIRST 180 DAY DELINQUENCY UPB	LAST STATUS	FOREGONE INTEREST COST	TOTAL NET LOSS	SEVERITY		
880684726544	C	BANK OF AMERICA, N	6.625	156000	360	Jan-06	Mar-06	80	80	1	53	742	N	P	SF	1	I	FL	325		FRM		BANK OF AMERICA, N	6.625	146996.64	57	303	0	Feb-36	18880	X	N	1	Nov-10	742	195000							P				

Note: Certain fields have been removed from this view to fit the data to the slide

From this example, we can see that this investor loan was originated in Florida in January 2006. The borrower made payments from March 2006 through October 2010 before prepaying the remaining UPB in November 2010.



# Dataset calculations

- Borrower minimum credit score
- Original Home Value

*Derived from acquisition variables*

- Last Status
- Last Date
- Credit Event Date
- Credit Event UPB
- First 180 Date
- First 180 UPB
- Foregone Interest Cost
- Net Loss
- Net Severity

*Derived from performance variables*

*Exact calculations for each of these variables are included in the associated R program posted with this webinar*

# Acquisition variables to calculate

1 2

## 1. Minimum credit score:

- If there is both a borrower credit score and co-borrower credit score on a mortgage loan, the minimum credit score is determined by taking the lowest of the two credit scores in order to create one column for analysis

## 2. Original home value:

- Original Loan-to-Value Ratio (OLTV) is the proportion of the original loan amount relative to the price of the property
- Purchase mortgages: Original home value = minimum of purchase price or appraisal
- Refinance mortgages: Original home value = appraisal
- Purchase price =  $\text{Original Loan Amount} / (\text{OLTV}/100)$

# Working with the performance files

Create *intermediate variables* to distill multiple activity records into a one-record-per-loan format.

- First, aggregate performance files
- Then, sort by loan ID and monthly reporting period
- Next, for loans with a “terminal” status, create flags to capture events. For example:
  - When Current Loan Delinquency Status first equals or exceeds 6, make F180\_DTE = <Monthly Reporting Period> and F180\_UPB = <Current Actual UPB>
  - When Current Loan Delinquency Status first equals or exceeds 6 or zero-balance code indicates an REO or REO alternative disposition, make FCE\_DTE <Monthly Reporting Period> and FCE\_UPB = <Current Actual UPB>
- Finally, for loans with an “active” status, capture the status as of the last monthly reporting period. For example:
  - Find the maximum of <Monthly Reporting Period> and set LAST\_STAT = <Current Loan Delinquency Status>

# Performance variables to calculate

1 2 3 4 5 6 7 8 9

## 1. Last Status:

- For loans in a “continuous” status (i.e., non-liquidated) the last date will be the latest monthly reporting period.

## 2. Last Date:

- For loans in a “terminal” status (i.e., liquidated) the last date will be the date on which the terminal status occurred.

## 3. Credit Event Date:

- The date that the credit event occurred.

## 4. Credit Event UPB:

- The outstanding unpaid principal balance corresponding to the credit event date.

# Performance variables to calculate

1 2 3 4 5 6 7 8 9

## 5. First 180 Date:

- The date that the first 180 day delinquency occurred.

## 6. First 180 UPB:

- The outstanding unpaid principal balance corresponding to the First 180 Date.

## 7. Delinquent (Foregone) Interest Expense:

- This represents the foregone interest Fannie Mae would otherwise earn on a performing loan.

# Performance variables to calculate

1 2 3 4 5 6 7 8 9

## 8. Net Loss:

- This represents the loss to Fannie Mae, including delinquent interest, net of any proceeds. It is represented in dollars.

## 9. Net Severity:

- This represents the net total loss to Fannie Mae, as a percentage of defaulted UPB. It is usually represented as a percentage, and is sometimes referred to as a “loss given default” statistic.

# Credit Event Definitions

Credit Event outcomes that would result in principal write-downs for CAS investors should be defined as:

## 1. Fixed Severity CAS Transactions

Credit Event defined as either a 180 day delinquency (Delq.Status GE 6) or a pre-180 day delinquency foreclosure or foreclosure alternative outcome (ZB Code 03 or 09), net of Post-D180 Repurchases (ZB Code 06)

## 2. Actual Loss CAS Transactions

Credit Event defined as REO Disposition or other disposition type. (ZB Code 03 or 09)

*Let's now look at an actual loan level example...*

This loan has both delinquency and loss information that we will want to capture in our final dataset.

*Again, by applying simple logic and some basic manipulation, we can create a dataset that will be smaller, easier to work with and still contain all the information required for our analysis.*



# From files to dataset

- Notice how pertinent dates and UPBs are kept, while other activity is dropped.
- Finally, **Total Net Loss** and **Severity** are calculated.

LOAN IDENTIFIER	MONTHLY REPORTING PERIOD	SERVICER NAME	CURRENT INTEREST RATE	CURRENT ACTUAL UNPAID PRINCIPAL BALANCE (UPB)	LOAN AGE	REMAINING MONTHS TO LEGAL MATURITY	ADJUSTED REMAINING MONTHS TO MATURITY	MATURITY DATE	METROPOLITAN STATISTICAL AREA	CURRENT LOAN DELINQUENCY STATUS	MODIFICATION FLAG	ZERO BALANCE CODE	ZERO BALANCE EFFECTIVE DATE	LAST PAID INSTALLMENT DATE	FORECLOSURE DATE	DISPOSITION DATE	FORECLOSURE COSTS	PROPERTY PRESERVATION AND REPAIR COSTS	ASSET RECOVERY COSTS	MISCELLANEOUS HOLDING EXPENSES AND CREDITS	ASSOCIATED TAXES FOR HOLDING PROPERTY	NET SALES PROCEEDS	CREDIT ENHANCEMENT PROCEEDS	REPURCHASE MAKE WHOLE PROCEEDS	OTHER FORECLOSURE PROCEEDS
637548085401	3/1/2006	CITIMORTGAGE	7	NA	1	359	359	Feb-36	26420	0	N														
637548085401	4/1/2006		7	NA	2	358	358	Feb-36	26420	0	N														

637548085401	5/1/2008		7	102813.4	27	333	328	Feb-36	26420	0	N														
637548085401	6/1/2008		7	102708.3	28	332	327	Feb-36	26420	0	N														
637548085401	7/1/2008		7	102602.7	29	331	326	Feb-36	26420	0	N														
637548085401	8/1/2008		7	102496.4	30	330	325	Feb-36	26420	0	N														
637548085401	9/1/2008		7	102389.5	31	329	324	Feb-36	26420	0	N														
637548085401	10/1/2008		7	102389.5	32	328	324	Feb-36	26420	0	N														
637548085401	11/1/2008		7	102389.5	33	327	324	Feb-36	26420	1	N														
637548085401	12/1/2008		7	102389.5	34	326	324	Feb-36	26420	2	N														
637548085401	1/1/2009		7	102389.5	35	325	324	Feb-36	26420	3	N														
637548085401	2/1/2009		7	102389.5	36	324	324	Feb-36	26420	4	N														
637548085401	3/1/2009		7	102389.5	37	323	324	Feb-36	26420	5	N														
637548085401	4/1/2009		7	102389.5	38	322	0	Feb-36	26420	X	N	9	Apr-09	Oct-08	Apr-09	Aug-09	2165.13	1140	760.76	217.97		71335.71	26891.22		

CURRENT LOAN DELINQUENCY STATUS	MODIFICATION FLAG	ZERO BALANCE CODE	ZERO BALANCE EFFECTIVE DATE	LAST PAID INSTALLMENT DATE	FORECLOSURE DATE	DISPOSITION DATE	FORECLOSURE COSTS	PROPERTY PRESERVATION AND REPAIR COSTS	ASSET RECOVERY COSTS	MISCELLANEOUS HOLDING EXPENSES AND CREDITS	ASSOCIATED TAXES FOR HOLDING PROPERTY	NET SALES PROCEEDS	CREDIT ENHANCEMENT PROCEEDS	REPURCHASE MAKE WHOLE PROCEEDS	OTHER FORECLOSURE PROCEEDS	MINIMUM CREDIT SCORE	ORIGINAL HOME VALUE	FIRST CREDIT EVENT DATE	FIRST CREDIT EVENT UPB	FIRST 180 DAY DELINQUENCY DATE	FIRST 180 DAY DELINQUENCY UPB	LAST STATUS	DELINQUENT INTEREST EXPENSE	TOTAL EXPENSE	TOTAL PROCEEDS	TOTAL NET LOSS	SEVERITY
X	N	9	Apr-09	Oct-08	Apr-09	Aug-09	2165	1140	761	218		71336	26891			811	11777	4/1/2009	2390			F	5674	4284	98227	14121	14%

- On the next slide, we'll explain exactly what we're looking at...

# From files to dataset



LOAN IDENTIFIER	CHANNEL	SELLER NAME	ORIGINAL INTEREST RATE	ORIGINAL UNPAID PRINCIPAL BALANCE (UPB)	ORIGINAL LOAN TERM	ORIGINATION DATE	FIRST PAYMENT DATE	ORIGINAL LOAN-TO-VALUE (LTV)	ORIGINAL COMBINED LOAN-TO-VALUE (LTV)	NUMBER OF BORROWERS	DEBT-TO-INCOME RATIO	BORROWER CREDIT SCORE	FIRST-TIME HOME BUYER INDICATOR	LOAN PURPOSE	PROPERTY TYPE	NUMBER OF UNITS	OCCUPANCY STATUS	PROPERTY STATE	ZIP (3-DIGIT)	MORTGAGE INSURANCE PERCENTAGE	PRODUCT TYPE	CO BORROWER CREDIT SCORE
637548085401	C	CITIMORTGAGE, INC.	7	106000	360	Jan-06	Mar-06	90	90	1	12	811	N	P	PU	1	I	TX	770	25	FRM	

*In January 2006, a purchase money investor loan was originated for \$106,000. The loan had a 30-year term, 90% LTV and 811 FICO score. There was only one borrower.*

# From files to dataset



	CURRENT LOAN DELINQUENCY STATUS																										
	MODIFICATION FLAG																										
	ZERO BALANCE CODE																										
	ZERO BALANCE EFFECTIVE DATE																										
	LAST PAID INSTALLMENT DATE																										
	FORECLOSURE DATE																										
	DISPOSITION DATE																										
	FORECLOSURE COSTS																										
	PROPERTY PRESERVATION AND REPAIR COSTS																										
	ASSET RECOVERY COSTS																										
	MISCELLANEOUS HOLDING EXPENSES AND CREDITS																										
	ASSOCIATED TAXES FOR HOLDING PROPERTY																										
	NET SALES PROCEEDS																										
	CREDIT ENHANCEMENT PROCEEDS																										
	REPURCHASE MAKE WHOLE PROCEEDS																										
	OTHER FORECLOSURE PROCEEDS																										
	MINIMUM CREDIT SCORE																										
	ORIGINAL HOME VALUE																										
	FIRST CREDIT EVENT DATE																										
	FIRST CREDIT EVENT UPB																										
	FIRST 180 DAY DELINQUENCY DATE																										
	FIRST 180 DAY DELINQUENCY UPB																										
	LAST STATUS																										
	DELINQUENT INTEREST EXPENSE																										
	TOTAL EXPENSE																										
	TOTAL PROCEEDS																										
	TOTAL NET LOSS																										
	SEVERITY																										
X	N	9	Apr-09	Oct-08	Apr-09	Aug-09	2165	1140	761	218		71336	26891			811	117778	4/1/2009	102390			F	5674	4284	98227	14121	14%

*The borrower reached 90 day delinquency in January 2009. The property was foreclosed on in April of 2009 and disposed of in August 2009. This resulted in a \$14,121 loss to Fannie Mae, which equates to a 14% severity rate.*

# Final Dataset – acquisition variables

Position	Field Name	Column ID	Type	Max Length	Allowable Values
1	LOAN IDENTIFIER	LOAN_ID	Alpha-Numeric	20	
2	CHANNEL	ORIG_CHN	Alpha-Numeric	1	R - "Retail" C - "Correspondent" B - "Broker"
3	SELLER NAME	Seller.Name	Alpha-Numeric	80	Name of Seller
4	ORIGINAL INTEREST RATE	ORIG_RT	Numeric	14,10	Blank = Unknown
5	ORIGINAL UNPAID PRINCIPAL BALANCE (UPB)	ORIG_AMT	Numeric	11,2	
6	ORIGINAL LOAN TERM	ORIG_TRM	Numeric	3,0	301 - 419
7	ORIGINATION DATE	ORIG_DTE	Date	MM/YYYY	MM/YYYY
8	FIRST PAYMENT DATE	FRST_DTE	Numeric	MM/YYYY	MM/YYYY
9	ORIGINAL LOAN-TO-VALUE (LTV)	OLTV	Numeric	14,10	0 - 97% / Blank (unknown)
10	ORIGINAL COMBINED LOAN-TO-VALUE (LTV)	OCLTV	Numeric	14,10	0 - 200% / Blank (if CLTV > 200 or unknown)
11	NUMBER OF BORROWERS	NUM_BO	Numeric	3,0	1 - 10
12	DEBT-TO-INCOME RATIO	DTI	Numeric	14,10	1 - 64% / Blank (if DTI is 0 or ≥ 65 or unknown)
13	BORROWER CREDIT SCORE	CSCORE_B	Alpha-Numeric	3,0	300 - 850 / Blank (if <300 or >850 or unknwn)
14	FIRST-TIME HOME BUYER INDICATOR	FTHB_FLG	Alpha-Numeric	1	Y - "First Time Home Buyer" N - "Not First Time Home Buyer" U - "Unknown"
15	LOAN PURPOSE	PURPOSE	Alpha-Numeric	1	P - "Purchase" R - "No Cash-out Refinance" C - "Cash-out Refinance" U - "Refinance - Not Specified"
16	PROPERTY TYPE	PROP_TYP	Alpha-Numeric	2	SF - "Single Family" CO - "Condo" CP - "Co-Op" MH - "Manufactured Housing" PU - "Planned Urban Development"
17	NUMBER OF UNITS	NUM_UNIT	Alpha-Numeric	10	1 - 4
18	OCCUPANCY STATUS	OCC_STAT	Alpha-Numeric	1	P - "Principal" S - "Second" I - "Investor" U - "Unknown"
19	PROPERTY STATE	STATE	Alpha-Numeric	20	
20	ZIP (3-DIGIT)	ZIP_3	Alpha-Numeric	10	XXX - First three digits of property's zip code
21	MORTGAGE INSURANCE PERCENTAGE	MI_PCT	Numeric	14,10	1 - 50% / Blank (if not applicable or < 1% or > 50%)
22	PRODUCT TYPE	Product.Type	Alpha-Numeric	20	FRM - "Fixed-rate mortgage loan"
23	CO BORROWER CREDIT SCORE	CSCORE_C	Numeric	3,0	300 - 850 / Blank (if <300 or >850, unknwn, or is not applicable)

# Final Dataset – performance variables

Position	Field Name	Column ID	Type	Max Length	Allowable Values
1	LOAN IDENTIFIER	LOAN_ID	Alpha-Numeric	20	
2	MONTHLY REPORTING PERIOD	Monthly.Rpt.Prđ	Date	MM/DD/YYYY	MM/DD/YYYY
3	SERVICER NAME	Servicer.Name	Alpha-Numeric	80	Name of Servicer / Other / Blank (unknown)
4	CURRENT INTEREST RATE	LAST_RT	Numeric	14,10	
5	CURRENT ACTUAL UNPAID PRINCIPAL BALANCE (UP	LAST_UPB	Numeric	11,2	
6	LOAN AGE	Loan.Age	Numeric	10,0	= [Monthly Reporting Period - First Payment Date] + 1
7	REMAINING MONTHS TO LEGAL MATURITY	Months.To.Legal.Ma	Numeric	3,0	=Maturity Date - Monthly Reporting Period
8	ADJUSTED REMAINING MONTHS TO MATURITY	Adj.Month.To.Mat	Numeric	3,0	
9	MATURITY DATE	Maturity.Date	Date	MM/YYYY	MM/YYYY
10	METROPOLITAN STATISTICAL AREA	MSA	Alpha-Numeric	5	XXXXX (five-digit MSA code)
11	CURRENT LOAN DELINQUENCY STATUS	Delq.Status	Alpha-Numeric	5	0 - "Current or less than 30 days past due" 1 - "30 - 59 days past due" 2 - "60 - 89 days past due" 3 - "90 - 119 days past due" 4 - "120 - 149 days past due" 5 - "150 - 179 days past due" 6 - "180 Day Delinquency" 7 - "210 Day Delinquency" 8 - "240 Day Delinquency" 9 - "270 Day Delinquency" / "270+ Day Delinquency" X - "Unknown"
12	MODIFICATION FLAG	MOD_FLAG	Alpha-Numeric	1	N - "No" Y - "Yes"
13	ZERO BALANCE CODE	Zero.Bal.Code	Alpha-Numeric	2	01 - "Prepaid or matured" 03 - "Short-sale, Third Party Sale, Note Sale" 06 - "Repurchased" 09 - "Deed-in-lieu or REO Disposition"
14	ZERO BALANCE EFFECTIVE DATE	LAST_DTE	Date	MM/YYYY	MM/YYYY
15	LAST PAID INSTALLMENT DATE	LPI_DTE	Date	MM/DD/YYYY	MM/DD/YYYY
16	FORECLOSURE DATE	FCC_DTE	Date	MM/DD/YYYY	MM/DD/YYYY
17	DISPOSITION DATE	DISP_DT	Date	MM/DD/YYYY	MM/DD/YYYY
18	FORECLOSURE COSTS	FCC_COST	Numeric	27,12	
19	PROPERTY PRESERVATION AND REPAIR COSTS	PP_COST	Numeric	27,12	
20	ASSET RECOVERY COSTS	AR_COST	Numeric	27,12	
21	MISCELLANEOUS HOLDING EXPENSES AND CREDITS	IE_COST	Numeric	27,12	
22	ASSOCIATED TAXES FOR HOLDING PROPERTY	TAX_COST	Numeric	27,12	
23	NET SALES PROCEEDS	NS_PROCS	Numeric	27,12	
24	CREDIT ENHANCEMENT PROCEEDS	CE_PROCS	Numeric	27,12	
25	REPURCHASE MAKE WHOLE PROCEEDS	RMW_PROCS	Numeric	27,12	
26	OTHER FORECLOSURE PROCEEDS	O_PROCS	Numeric	27,12	

# Final Dataset – calculated variables

Field Name	Column ID	Type	Max Length	Allowable Values
MINIMUM CREDIT SCORE	CSCORE_MN	Numeric	3,0	300 - 850 / Blank (if <300 or >850 or unknown)
ORIGINAL HOME VALUE	ORIG_VAL	Numeric	11,2	Various
FIRST CREDIT EVENT DATE	FCE_DTE	Date	MM/DD/YYYY	MM/DD/YYYY
FIRST CREDIT EVENT UPB	FCE_UPB	Numeric	11,2	Various
FIRST 180 DAY DELINQUENCY DATE	F180_DTE	Date	MM/DD/YYYY	MM/DD/YYYY
FIRST 180 DAY DELINQUENCY UPB	F180_UPB	Numeric	11,2	Various
LAST STATUS	LAST_STAT	Alpha-Numeric	1	C - "Current"
				1 - "30 Day Delinquency"
				2 - "60 Day Delinquency"
				3 - "90 Day Delinquency"
				4 - "120 Day Delinquency"
				5 - "150 Day Delinquency"
				6 - "180 Day Delinquency"
				7 - "210 Day Delinquency"
				8 - "240 Day Delinquency"
				9 - "270 Day Delinquency" / "270+ Day Delinquency"
				P - "Prepaid"
				S - "Foreclosure Alternative (Short Sale / Third Party Sale)"
				R - "Repurchase"
				F - "REO Disposition / Deed-in-lieu"
DELINQUENT INTEREST EXPENSE	INT_COST	Numeric	11,2	
TOTAL EXPENSE	TOTAL_COST	Numeric	11,2	
TOTAL PROCEEDS	TOTAL_PROCS	Numeric	11,2	
TOTAL NET LOSS	NET_LOSS	Numeric	11,2	
SEVERITY	NET_SEV	Numeric	3,0	

# Great! We have a dataset... now what!?

1. **Get familiar with the variables in dataset:**
  - Run frequency distributions on key fields
  - Identify outliers, nulls and missing values
2. **Begin the analysis**

# Example frequency distributions

Loan Count						Loan Count				
	Purchase	Rate/ Term	Cash out	Unknown Refinance	Total		Primary	Second Home	Investor	Grand Total
1999	90,989	18,793	17,397		127,179	1999	118,249	3,818	5,112	127,179
2000	803,533	128,235	138,402	25	1,070,195	2000	983,193	37,991	49,011	1,070,195
2001	915,552	774,688	655,213	1,058	2,346,511	2001	2,160,254	61,838	124,419	2,346,511
2002	828,049	829,833	730,566	1,860	2,390,308	2002	2,169,090	71,948	149,270	2,390,308
2003	802,471	1,255,930	945,203	5,403	3,009,007	2003	2,732,557	93,223	183,227	3,009,007
2004	523,406	299,705	368,304	1,341	1,192,756	2004	1,069,811	53,344	69,601	1,192,756
2005	486,502	199,136	445,566	55	1,131,259	2005	1,011,478	56,979	62,802	1,131,259
2006	405,225	135,512	353,892	2	894,631	2006	786,013	46,044	62,574	894,631
2007	437,875	209,845	415,795	2	1,063,517	2007	919,889	51,158	92,470	1,063,517
2008	532,638	286,805	362,015		1,181,458	2008	1,002,479	58,799	120,180	1,181,458
2009	488,142	732,295	535,711		1,756,148	2009	1,584,951	87,276	83,921	1,756,148
2010	473,193	431,105	293,983	13	1,198,294	2010	1,035,866	65,890	96,538	1,198,294
2011	468,116	331,744	203,015		1,002,875	2011	833,872	59,674	109,329	1,002,875
2012	672,346	735,930	303,016	1	1,711,293	2012	1,467,078	83,516	160,699	1,711,293
2013	806,983	456,442	258,942		1,522,367	2013	1,295,922	73,436	153,009	1,522,367
2014	241,741	49,981	55,532		347,254	2014	292,469	17,893	36,892	347,254
Total	8,976,761	6,875,979	6,082,552	9,760	21,945,052	Total	19,463,171	922,827	1,559,054	21,945,052

*You will notice two distinct refinance waves, one in 2003 and one in 2009.*



# Example frequency distributions

Last Status	Loan Count
C	6,073,210
1	76,388
2	21,647
3	9,748
4	7,589
5	6,867
6	5,324
7	4,598
8	4,121
9	60,232
F	350,431
P	15,155,401
R	49,123
S	120,373
Total	21,945,052



Frequency Distribution	Borrower Credit Score	Original Loan Amount	OLTV
Min.	300	\$ 4,000	1
1st Quartile	694	\$ 114,000	65
Median	742	\$ 170,000	77
Mean	732	\$ 193,152	73
3 <sup>rd</sup> Quartile	777	\$ 250,000	80
Max.	850	\$ 1,403,000	97
Missing	71343	\$ -	11

*Loan counts are great for verifying two populations, but quartile frequency distributions are far better for getting a sense of continuous variables.*

# Min. credit scores across origination vintages



Origination Year	760+	[720 - 760]	[680 - 720]	[620 - 680]	< 620	Missing	Total
1999	25,791	33,233	28,738	30,109	7,454	1,854	127,179
2000	237,459	277,631	238,198	241,055	61,238	14,614	1,070,195
2001	546,106	604,398	528,279	513,607	133,377	20,744	2,346,511
2002	647,969	603,736	507,181	489,265	131,331	10,826	2,390,308
2003	862,887	761,972	639,255	602,976	134,291	7,626	3,009,007
2004	314,523	274,752	259,923	273,561	65,649	4,348	1,192,756
2005	338,686	243,368	240,354	248,215	57,520	3,116	1,131,259
2006	266,131	183,762	184,924	202,861	55,575	1,378	894,631
2007	323,522	210,846	213,885	245,533	68,533	1,198	1,063,517
2008	484,604	285,769	231,806	148,424	29,947	908	1,181,458
2009	1,034,922	423,549	217,949	71,062	7,286	1,380	1,756,148
2010	738,105	269,659	141,535	47,523	417	1,055	1,198,294
2011	597,859	228,088	125,564	50,317	117	930	1,002,875
2012	1,065,473	382,297	193,349	69,312	185	677	1,711,293
2013	820,742	377,392	227,408	96,159	115	551	1,522,367
2014	158,882	86,668	64,825	36,735	6	138	347,254
Total	8,463,661	5,247,120	4,043,173	3,366,714	753,041	71,343	21,945,052

*Missing or unknown values are concentrated in the earlier origination vintages*

# Options for addressing missing/null values

- Start with count of missing values across key variables
  - Where are they concentrated?
  - When are they concentrated?
- Using the metric you are interested in, determine what category/cohort the null values most closely resemble
  - Measure metric you are interested in against the concentration of missing values and one other metric, e.g., origination year.
  - Remember to consider potential underlying causes.
- Set missing values to the next best possible value
  - Based on this metric, what group does the null value most closely resemble?

# Addressing null values – FICO

Default Rate (%) by Origination Year																
FICO	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	Total
760+	0.2%	0.2%	0.2%	0.2%	0.4%	0.8%	2.1%	3.0%	3.3%	1.5%	0.2%	0.1%	0.0%	0.0%	0.0%	0.5%
[720 - 760)	0.4%	0.3%	0.4%	0.5%	0.9%	1.6%	4.1%	6.0%	6.4%	3.5%	0.6%	0.1%	0.1%	0.0%	0.0%	1.3%
[680 - 720)	0.8%	0.8%	0.8%	1.0%	1.5%	2.8%	6.1%	8.3%	8.8%	5.5%	0.9%	0.3%	0.1%	0.0%	0.0%	2.4%
[620 - 680)	2.0%	1.9%	1.7%	1.8%	2.5%	4.1%	8.4%	11.1%	11.4%	7.7%	1.4%	0.5%	0.2%	0.0%	0.0%	4.2%
< 620	3.3%	2.9%	2.4%	2.5%	2.9%	4.3%	7.9%	10.8%	11.3%	7.6%	2.0%	0.9%	0.0%	0.0%	0.0%	5.1%
Missing	2.1%	2.7%	1.8%	4.3%	5.2%	5.0%	6.6%	7.1%	7.7%	3.7%	0.6%	0.2%	0.0%	0.0%	0.0%	3.2%
Total	1.0%	0.9%	0.8%	0.9%	1.3%	2.4%	5.0%	7.0%	7.3%	3.6%	0.4%	0.1%	0.0%	0.0%	0.0%	1.6%

Grouping the missing FICOs into a single other FICO bucket will inevitably cause bias in other vintages, therefore, we will keep these records separate for the purpose of our analysis in the 102 module.

# Addressing missing/null values

Options for addressing missing/null values:

1. **OCLTV** → set OCLTV to OLTV, or keep as null
  2. **Number of borrowers** → set to count of FICO scores, set to single borrower, or keep as null
  3. **Purpose** → set unknown refinances to cash-out, or keep as null
  4. **FICO** → set FICO to 620, set to average of vintage, or keep as null
  5. **DTI** → set to 45%, set to average of vintage, or keep as null
- > Users may also want to create a flag to identify the loans where data has been overwritten. This would enable them to also examine the impact of the corrections are having on their analysis.

# Next up... summary statistics!



Origination Year	Loan Count	Total Orig. UPB (\$M)	Avg. Orig UPB (\$)	Acquisition Characteristics <sup>1</sup>					
				Borrower Credit Score	Co-Borrower Credit Score	LTV Ratio	CLTV Ratio <sup>2</sup>	DTI	Note Rate
1999	127,179	\$ 15,948	\$ 125,402	716	724	79.5	79.5	34.9	7.80
2000	1,070,195	\$ 140,963	\$ 131,717	718	726	79.1	79.3	35.7	8.13
2001	2,346,511	\$ 349,702	\$ 149,031	719	726	75.4	75.8	34.0	6.99
2002	2,390,308	\$ 374,469	\$ 156,661	723	730	72.9	73.4	34.0	6.50
2003	3,009,007	\$ 497,095	\$ 165,202	725	732	70.8	71.5	33.5	5.75
2004	1,192,756	\$ 200,864	\$ 168,403	721	728	72.1	73.7	36.6	5.84
2005	1,131,259	\$ 208,483	\$ 184,293	725	732	71.2	73.3	38.2	5.84
2006	894,631	\$ 172,510	\$ 192,828	724	732	71.8	73.9	39.3	6.42
2007	1,063,517	\$ 218,063	\$ 205,039	724	732	73.4	75.5	39.4	6.36
2008	1,181,458	\$ 262,804	\$ 222,441	744	752	73.4	75.0	38.5	6.04
2009	1,756,148	\$ 417,064	\$ 237,488	764	770	67.9	69.4	34.2	4.97
2010	1,198,294	\$ 295,123	\$ 246,286	768	774	69.6	71.0	32.8	4.72
2011	1,002,875	\$ 235,291	\$ 234,617	767	774	71.7	73.0	33.1	4.55
2012	1,711,293	\$ 417,948	\$ 244,229	770	775	71.6	72.9	31.9	3.84
2013	1,522,367	\$ 356,118	\$ 233,924	762	768	75.2	76.3	33.4	4.05
2014	347,254	\$ 76,291	\$ 219,698	752	759	78.8	79.6	34.9	4.62
<b>Total</b>	<b>21,945,052</b>	<b>\$ 4,238,736</b>	<b>\$ 193,152</b>	<b>742</b>	<b>751</b>	<b>72.4</b>	<b>73.6</b>	<b>34.7</b>	<b>5.52</b>

<sup>1</sup> Acquisition Characteristics are UPB-weighted averages, based on UPB at origination

<sup>2</sup> Missing CLTVs have been set to OLTV in this view

# Computing weighted averages

- For the purpose of cohort analysis, general practice is to weight statistics by loan amount.
- This recognizes that larger loans make up a larger portion of a pool and should be represented as such.
- It is important to consider the time period for which you are trying to calculate a weighted-average statistic for:
  - For origination statistics, use origination UPB
  - For periods following origination, use the activity UPB for the period

$$\text{Formula: } \Sigma(\text{Statistic} * \text{UPB}) / \Sigma(\text{UPB})$$

*Records with missing values should be excluded from the numerator and denominator*

- Statistics
  - Original Interest Rate
  - Current Interest Rate
  - Loan Term
  - OLTV
  - OCLTV
  - Number of borrowers
  - DTI
  - Credit Score
  - Number of Units
  - Mortgage Insurance Percentage

# Computing population percentages

- Find percentage of loans in loan pool that match a specific criteria:
  - Create a binary (1/0) variable for the loans that match criteria, and either:
    - Take the average of the flagged loans, for a simple average, or,
    - Take a weighted-average of flagged loans, with appropriate UPB as weight
- Common criteria include:
  - Percentage of cohort with 2nd lien
  - Percentage of cohort with non-owner occupied loans
  - Percent of cohort with primary mortgage insurance



# Acquisition statistical summary table



Origination Year	Loan Count	Total Orig. UPB (\$M)	Avg. Orig UPB (\$)	Acquisition Characteristics <sup>1</sup>					
				Borrower Credit Score	Co-Borrower Credit Score	LTV Ratio	CLTV Ratio <sup>2</sup>	DTI	Note Rate
1999	127,179	\$ 15,948	\$ 125,402	716	724	79.5	79.5	34.9	7.80
2000	1,070,195	\$ 140,963	\$ 131,717	718	726	79.1	79.3	35.7	8.13
2001	2,346,511	\$ 349,702	\$ 149,031	719	726	75.4	75.8	34.0	6.99
2002	2,390,308	\$ 374,469	\$ 156,661	723	730	72.9	73.4	34.0	6.50
2003	3,009,007	\$ 497,095	\$ 165,202	725	732	70.8	71.5	33.5	5.75
2004	1,192,756	\$ 200,864	\$ 168,403	721	728	72.1	73.7	36.6	5.84
2005	1,131,259	\$ 208,483	\$ 184,293	725	732	71.2	73.3	38.2	5.84
2006	894,631	\$ 172,510	\$ 192,828	724	732	71.8	73.9	39.3	6.42
2007	1,063,517	\$ 218,063	\$ 205,039	724	732	73.4	75.5	39.4	6.36
2008	1,181,458	\$ 262,804	\$ 222,441	744	752	73.4	75.0	38.5	6.04
2009	1,756,148	\$ 417,064	\$ 237,488	764	770	67.9	69.4	34.2	4.97
2010	1,198,294	\$ 295,123	\$ 246,286	768	774	69.6	71.0	32.8	4.72
2011	1,002,875	\$ 235,291	\$ 234,617	767	774	71.7	73.0	33.1	4.55
2012	1,711,293	\$ 417,948	\$ 244,229	770	775	71.6	72.9	31.9	3.84
2013	1,522,367	\$ 356,118	\$ 233,924	762	768	75.2	76.3	33.4	4.05
2014	347,254	\$ 76,291	\$ 219,698	752	759	78.8	79.6	34.9	4.62
Total	21,945,052	\$ 4,238,736	\$ 193,152	742	751	72.4	73.6	34.7	5.52

<sup>1</sup> Acquisition Characteristics are UPB-weighted averages, based on UPB at origination

<sup>2</sup> Missing CLTVs have been set to OLTV in this view

*You should now have all of the tools you need to produce this table!*

# Performance statistical summary table

			Active Loans		Inactive Loans (Loan Count)				Total Mods to Date <sup>2</sup>					
Origination Year	Loan Count	Total Orig. UPB (\$M)	Loan Count (Active)	Active UPB (\$M)	Prepaid	Repurchased <sup>1</sup>	Alternative Disposition	REO Disposition	Loan Count	D180 UPB (\$M) <sup>3,4</sup>	D180 % of Orig. UPB <sup>3,4</sup>	Default UPB (\$M) <sup>5</sup>	Loss Rate (%) <sup>5</sup>	
1999	127,179	\$ 15,948	2,719	\$ 155	122,135	613	319	1,393	846	\$ 298	1.9%	\$ 155	0.1%	
2000	1,070,195	\$ 140,963	14,955	\$ 875	1,038,110	3,159	2,079	11,892	6,237	\$ 2,148	1.5%	\$ 1,262	0.1%	
2001	2,346,511	\$ 349,702	65,824	\$ 4,796	2,246,209	3,921	4,537	26,020	16,127	\$ 5,062	1.4%	\$ 2,945	0.2%	
2002	2,390,308	\$ 374,469	134,293	\$ 11,574	2,217,026	3,643	5,595	29,751	21,119	\$ 6,380	1.7%	\$ 3,388	0.3%	
2003	3,009,007	\$ 497,095	381,116	\$ 39,448	2,565,274	4,620	12,134	45,863	44,869	\$ 13,882	2.8%	\$ 6,410	0.4%	
2004	1,192,756	\$ 200,864	186,442	\$ 21,340	964,746	2,388	9,108	30,072	30,716	\$ 9,960	5.0%	\$ 4,883	0.9%	
2005	1,131,259	\$ 208,483	207,233	\$ 28,230	854,703	2,860	19,873	46,590	50,970	\$ 18,810	9.0%	\$ 10,694	2.2%	
2006	894,631	\$ 172,510	143,919	\$ 21,323	677,368	3,345	21,113	48,886	56,261	\$ 21,083	12.2%	\$ 12,264	3.2%	
2007	1,063,517	\$ 218,063	190,308	\$ 30,673	777,176	8,246	26,149	61,638	81,875	\$ 30,672	14.1%	\$ 16,227	3.0%	
2008	1,181,458	\$ 262,804	193,083	\$ 30,538	927,299	8,617	15,130	37,329	55,681	\$ 20,374	7.7%	\$ 9,580	1.1%	
2009	1,756,148	\$ 417,064	558,385	\$ 94,986	1,184,564	2,298	3,233	7,668	10,596	\$ 4,646	1.1%	\$ 1,811	0.1%	
2010	1,198,294	\$ 295,123	546,249	\$ 99,273	648,333	1,045	701	1,966	3,075	\$ 1,296	0.4%	\$ 347	0.0%	
2011	1,002,875	\$ 235,291	562,564	\$ 104,244	438,784	469	244	814	1,846	\$ 665	0.3%	\$ 109	0.0%	
2012	1,711,293	\$ 417,948	1,414,197	\$ 315,862	295,613	968	114	401	888	\$ 459	0.1%	\$ 38	0.0%	
2013	1,522,367	\$ 356,118	1,361,094	\$ 300,648	158,495	2,591	43	144	267	\$ 259	0.1%	\$ 8	0.0%	
2014	347,254	\$ 76,291	307,343	\$ 63,298	39,566	340	1	4	3	\$ 22	0.0%	\$ -	0.0%	
Total	21,945,052	\$ 4,238,736	6,269,724	\$ 1,167,262	15,155,401	49,123	120,373	350,431	381,376	\$ 136,017	3.2%	\$ 70,121	0.6%	

<sup>1</sup> Reflects loans repurchased up to and after 180 days of delinquency. Previous versions of the Statistical Summary reflected in this column only loans repurchased prior to the occurrence of a credit event.

<sup>2</sup> Only one modification is counted per loan.

<sup>3</sup> D180 Rates included here are calculated in the same methodology as prior statistical summaries, they are included for comparison purposes only.

<sup>4</sup> Reflects the outstanding available UPB at D180 as reflected in the dataset.

<sup>5</sup> Default rates and UPB in this view are for completed foreclosures only. These are defined as loans with a zero balance code of '09' or '03' and non-null disposition dates.

*This next section will prepare you to analyze performance outcomes*

# Performance outcome identification

- Zero Balance Codes identify outcomes resulting in loan liquidations

ZB Code	Terminal Outcomes	LAST_STAT
01	Prepaid	P
03	Foreclosure Alternatives (includes short sales & third party sales)	S
06	Repurchased	R
09	REO Disposition or Deed-in-lieu	F

- To consider the impact of modifications or repurchases, we often:
  - Separate current/prepaid loans into previously modified or non-modified cohorts for analysis, and/or
  - Filter repurchases out of severity analysis because those loans would be removed from the CAS reference pool

## Creating the “last status” variable

- The Last Status, or “LAST\_STAT”, variable represents the latest status of a loan as available in the dataset.
- In R, we program this variable as:

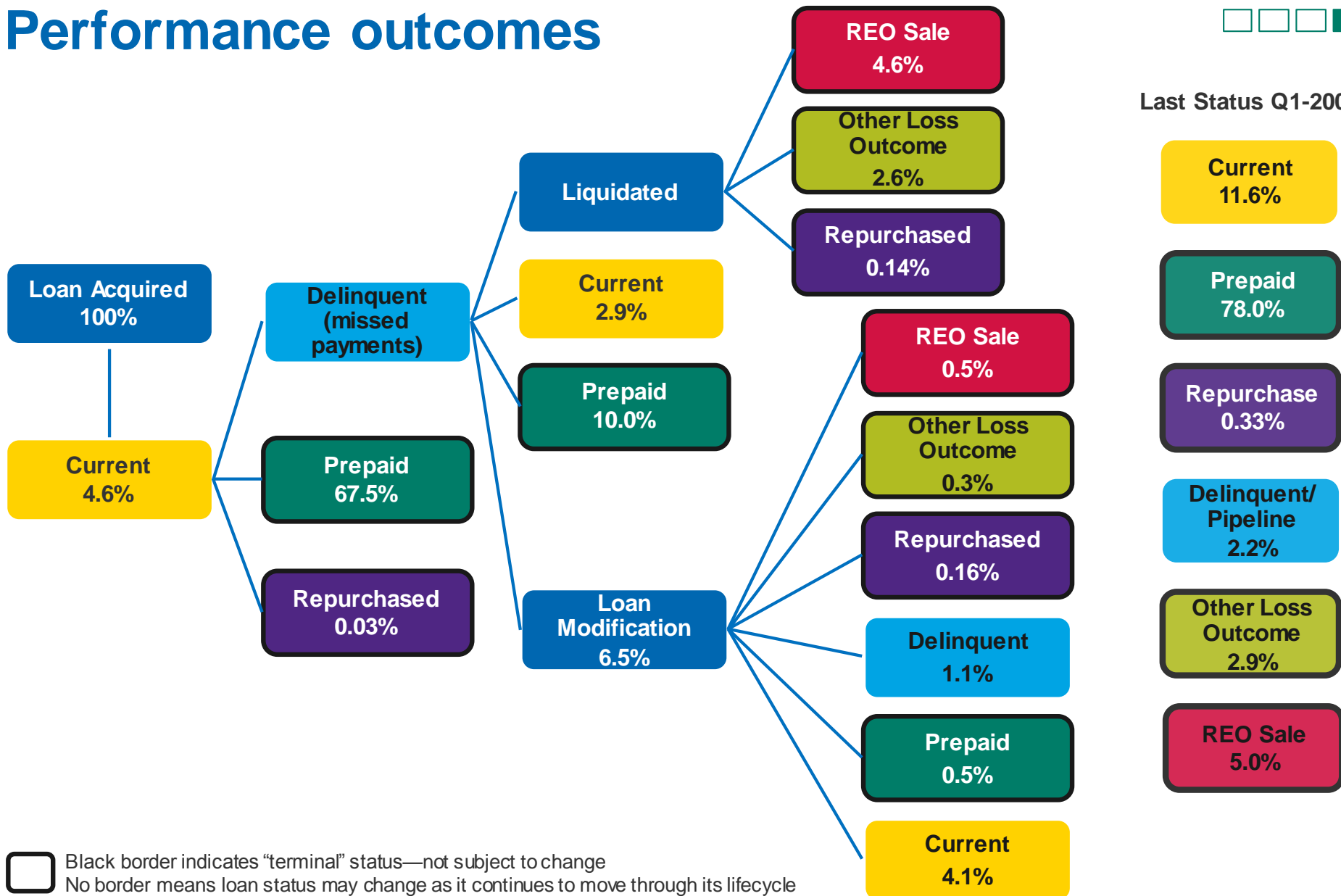
```
ifelse(Zero.Bal.Code=='01','P',ifelse(Zero.Bal.C  
ode=='03','S', ifelse(Zero.Bal.Code=='06', 'R',  
ifelse(Zero.Bal.Code=='09', 'F',  
ifelse(!(Delq.Status %chin% c('C', '1', '2',  
'3', '4', '5', '6', '7', '8', '9', 'X'))), '9',  
Delq.Status))))
```

- This script cycles through the Zero Balance Codes and then the Current Loan Delinquency Status codes to determine what activity state a particular loan is in as of the last information available for the loan in the dataset.

# Performance outcomes



Last Status Q1-2006

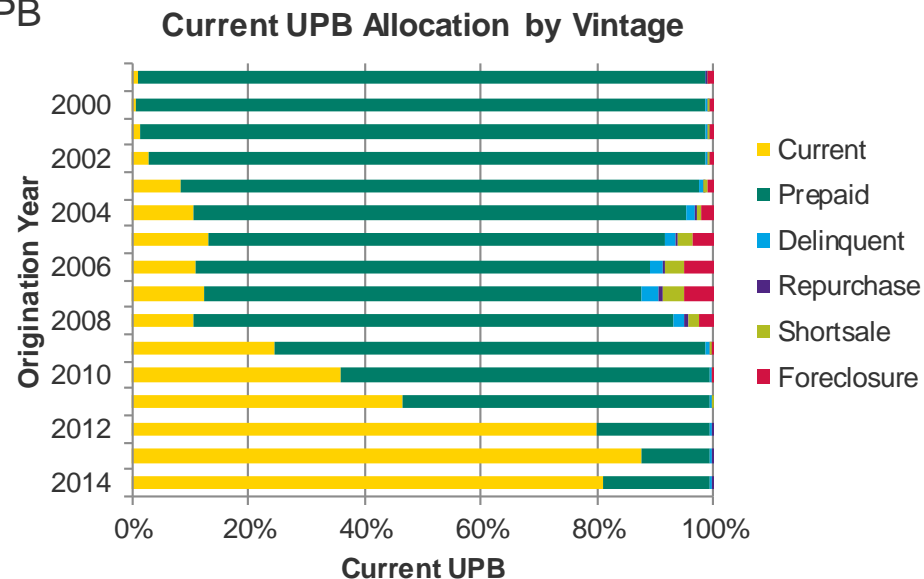


# Computing performance statistics



- Similar to calculating acquisitions population percentages:
  - Flag matching loans and sum total UPB → for example, flag loans with ZB “01” or last status “P”
  - Divide sum by total UPB for pool
- Examples:
  - **Prepayment Rate** = UPB of prepaid loans / Origination UPB
  - **Credit Event Rate** = (UPB of loans with a first 180 date + UPB of loans with a default date prior to their D180 date) / Origination UPB
  - **Default Rate** = Default UPB / Origination UPB
  - **Loss Rate** = Total Loss / Origination UPB

*Pre-crisis loans have nearly completely prepaid, while post-crisis loans are largely active, with low delinquencies as of the July 2015 data release.*



# Performance Statistical Summary Table



			Active Loans		Inactive Loans (Loan Count)				Total Mods to Date <sup>2</sup>					
Origination Year	Loan Count	Total Orig. UPB (\$M)	Loan Count (Active)	Active UPB (\$M)	Prepaid	Repurchased <sup>1</sup>	Alternative Disposition	REO Disposition	Loan Count	D180 UPB (\$M) <sup>3,4</sup>	D180 % of Orig. UPB <sup>3,4</sup>	Default UPB (\$M) <sup>5</sup>	Loss Rate (%) <sup>5</sup>	
1999	127,179	\$ 15,948	2,719	\$ 155	122,135	613	319	1,393	846	\$ 298	1.9%	\$ 155	0.1%	
2000	1,070,195	\$ 140,963	14,955	\$ 875	1,038,110	3,159	2,079	11,892	6,237	\$ 2,148	1.5%	\$ 1,262	0.1%	
2001	2,346,511	\$ 349,702	65,824	\$ 4,796	2,246,209	3,921	4,537	26,020	16,127	\$ 5,062	1.4%	\$ 2,945	0.2%	
2002	2,390,308	\$ 374,469	134,293	\$ 11,574	2,217,026	3,643	5,595	29,751	21,119	\$ 6,380	1.7%	\$ 3,388	0.3%	
2003	3,009,007	\$ 497,095	381,116	\$ 39,448	2,565,274	4,620	12,134	45,863	44,869	\$ 13,882	2.8%	\$ 6,410	0.4%	
2004	1,192,756	\$ 200,864	186,442	\$ 21,340	964,746	2,388	9,108	30,072	30,716	\$ 9,960	5.0%	\$ 4,883	0.9%	
2005	1,131,259	\$ 208,483	207,233	\$ 28,230	854,703	2,860	19,873	46,590	50,970	\$ 18,810	9.0%	\$ 10,694	2.2%	
2006	894,631	\$ 172,510	143,919	\$ 21,323	677,368	3,345	21,113	48,886	56,261	\$ 21,083	12.2%	\$ 12,264	3.2%	
2007	1,063,517	\$ 218,063	190,308	\$ 30,673	777,176	8,246	26,149	61,638	81,875	\$ 30,672	14.1%	\$ 16,227	3.0%	
2008	1,181,458	\$ 262,804	193,083	\$ 30,538	927,299	8,617	15,130	37,329	55,681	\$ 20,374	7.8%	\$ 9,580	1.1%	
2009	1,756,148	\$ 417,064	558,385	\$ 94,986	1,184,564	2,298	3,233	7,668	10,596	\$ 4,646	1.1%	\$ 1,811	0.1%	
2010	1,198,294	\$ 295,123	546,249	\$ 99,273	648,333	1,045	701	1,966	3,075	\$ 1,296	0.4%	\$ 347	0.0%	
2011	1,002,875	\$ 235,291	562,564	\$ 104,244	438,784	469	244	814	1,846	\$ 665	0.3%	\$ 109	0.0%	
2012	1,711,293	\$ 417,948	1,414,197	\$ 315,862	295,613	968	114	401	888	\$ 459	0.1%	\$ 38	0.0%	
2013	1,522,367	\$ 356,118	1,361,094	\$ 300,648	158,495	2,591	43	144	267	\$ 259	0.1%	\$ 8	0.0%	
2014	347,254	\$ 76,291	307,343	\$ 63,298	39,566	340	1	4	3	\$ 22	0.0%	\$ -	0.0%	
Total	21,945,052	\$ 4,238,736	6,269,724	\$ 1,167,262	15,155,401	49,123	120,373	350,431	381,376	\$ 136,017	3.2%	\$ 70,121	0.6%	

<sup>1</sup> Reflects loans repurchased up to and after 180 days of delinquency. Previous versions of the Statistical Summary reflected in this column only loans repurchased prior to the occurrence of a credit event.

<sup>2</sup> Only one modification is counted per loan.

<sup>3</sup> D180 Rates included here are calculated in the same methodology as prior statistical summaries, they are included for comparison purposes only.

<sup>4</sup> Reflects the outstanding available UPB at D180 as reflected in the dataset.

<sup>5</sup> Default rates and UPB in this view are for completed foreclosures only. These are defined as loans with a zero balance code of '09' or '03' and non-null disposition dates.

*Piece of cake! Right??*

# Loss/Severity Statistical Summary Table



	Origination Year															Total
	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	
Default UPB (\$M)	\$155	\$1,262	\$2,945	\$3,388	\$6,410	\$4,883	\$10,694	\$12,264	\$16,227	\$9,580	\$1,811	\$347	\$109	\$38	\$8	\$70,121
Default Rate (%)	1.0%	0.9%	0.8%	0.9%	1.3%	2.4%	5.1%	7.1%	7.4%	3.7%	0.4%	0.1%	0.1%	0.0%	0.0%	1.7%
<b>EXPENSES:</b>																
Delinquent Interest	12%	12%	11%	10%	10%	10%	9%	10%	10%	9%	7%	6%	6%	4%	3%	10%
Total Liquidation Exp.	10%	11%	11%	11%	11%	10%	8%	7%	7%	7%	7%	7%	7%	6%	4%	8%
Foreclosure	4%	5%	4%	4%	4%	4%	3%	3%	2%	2%	2%	3%	2%	2%	2%	3%
Property Preservation	3%	3%	3%	3%	3%	2%	2%	2%	2%	1%	2%	2%	2%	2%	1%	2%
Asset Recovery	1%	1%	1%	1%	1%	1%	1%	1%	1%	1%	1%	1%	1%	1%	1%	1%
Misc. Holding Expenses/Credits	1%	1%	1%	1%	1%	1%	1%	1%	1%	1%	0%	0%	0%	0%	0%	1%
Associated Taxes	2%	2%	2%	3%	3%	3%	2%	2%	2%	2%	2%	2%	2%	1%	1%	2%
Total Costs	122%	123%	121%	122%	121%	120%	117%	118%	117%	117%	114%	114%	112%	110%	107%	118%
<b>PROCEEDS:</b>																
Net Sales Proceeds	81%	77%	74%	72%	76%	72%	66%	61%	60%	63%	76%	81%	83%	82%	91%	66%
Credit Enhancement	17%	17%	14%	12%	9%	8%	6%	6%	8%	10%	5%	6%	8%	9%	10%	8%
Repurchase/Make Whole	6%	8%	5%	4%	2%	2%	3%	5%	8%	12%	8%	6%	4%	1%	0%	6%
Other	4%	4%	5%	4%	3%	2%	1%	1%	1%	1%	1%	1%	1%	1%	1%	2%
Total Proceeds	109%	107%	98%	93%	89%	84%	75%	72%	78%	86%	89%	93%	96%	93%	101%	82%
Severity	12.9%	15.9%	23.3%	29.2%	31.4%	36.4%	42.0%	45.5%	39.9%	30.4%	24.6%	20.5%	16.6%	17.2%	5.3%	36.6%
Total Net Loss (\$M)	\$20	\$201	\$685	\$989	\$2,012	\$1,775	\$4,494	\$5,576	\$6,474	\$2,908	\$445	\$71	\$18	\$7	\$0	\$25,676

*Establishing severity and loss expectations are key to analyzing credit risk*



# Foregone/Delinquent interest expense

- For performing loans on Fannie Mae's balance sheet, Fannie Mae receives full interest payment from borrowers, less servicing, for each month a loan sits on Fannie Mae's balance sheet.

- We calculate foregone interest as:

**Delinquent Interest** = Default UPB \* ((Current Interest Rate/100 – 0.0035)/12) \* months from LPI to disposition

- In R, this is programmed as:

```
ifelse(((LAST_STAT == "F" | LAST_STAT == "S") &
!is.na(DISP_DT)), LAST_UPB * (((LAST_RT/100) -
.0035)/12) * lpi2disp, 0)
```

- Where lpi2disp is defined as:

```
ifelse(Data_P$LPI_DTE != "" &
!(is.na(Data_P$DISP_DT)), as.numeric((year(DISP_DT) -
year(as.yearmon(LPI_DTE, "%m/%d/%Y")) * 12 + month(DISP_DT) -
month(as.yearmon(LPI_DTE, "%m/%d/%Y"))), 0))
```

# Computing loss statistics

- To compute loss statistics, carefully identify the numerator and denominator.
- For loss rates, the denominator will be the sum of the original UPB of the cohort.
- For a loss given default rate, the denominator will be the sum of the defaulted UPB of the cohort → this ensures that the formula for calculating a net loss rate from its components (default rate and severity, or loss given default) holds true.
- The new expense and proceed fields added to our data release are:
  1. Foreclosure Costs
  2. Property Preservation and Repair Costs
  3. Asset Recovery Costs
  4. Miscellaneous Holding Expenses and Credits
  5. Associated Taxes for Holding Property
  6. Net Sales Proceeds
  7. Credit Enhancement Proceeds
  8. Repurchase/Make-Whole Proceeds
  9. Other Foreclosure Proceeds

$$\text{Default Rate} * \text{Severity} = \text{Net Loss Rate}$$

# Computing loss statistics

1 2 3 4 5 6 7 8 9

## 1. Foreclosure Costs:

- Expenses associated with obtaining title to property from the mortgagor, valuing the property, and maintaining utility services to the property. Includes costs and fees associated with bankruptcy and foreclosure.

## 2. Property Preservation and Repair Costs:

- Expenses associated with securing and preserving the property. The expenses associated with securing and preserving the property including two major categories: maintenance and repairs. Maintenance costs are associated with preserving the property through normal upkeep, while repairs are associated with either avoiding deterioration of the asset or a marketing decision to help maximize sales proceeds upon final disposition.

# Computing loss statistics

1 2 3 4 5 6 7 8 9

## 3. Asset Recovery Costs:

- Expenses associated with removing occupants and personal property from an occupied property post foreclosure. Such expenses include relocation assistance, along with fees and costs associated with vacating a property.

## 4. Miscellaneous Holding Expenses and Credits:

- Expenses and credit associated with holding the property post foreclosure, including Homeowners Association and other dues; flood, hazard, and MI premiums and refunds; rental income; and title insurance costs.

# Computing loss statistics

1 2 3 4 5 6 7 8 9

## 5. Associated Taxes for Holding Property:

- Payment of taxes associated with holding the property.

## 6. Net Sales Proceeds:

- Total cash received from the sale of the property net of any applicable selling expenses, such as fees and commissions, allowable for inclusion under the terms of the property sale, as currently reported on the HUD-1 or other settlement statement.

## 7. Credit Enhancement Proceeds:

- Proceeds from primary mortgage insurance policy claims and recourse and indemnification payments from lenders under arrangements designed to limit credit exposure to Fannie Mae. Includes only amounts actually collected.

# Computing loss statistics

1 2 3 4 5 6 7 8 9

## 8. Repurchase/Make-Whole Proceeds:

- Amounts received by Fannie Mae under the terms of our representation and warranty arrangements for the repurchase of the mortgage loan or the subject property or loss reimbursement subsequent to property disposition.

## 9. Other Foreclosure Proceeds:

- Amounts, other than sale proceeds, received by Fannie Mae following foreclosure of a property, including redemption proceeds received from the mortgagor.

# Identifying “completed” property dispositions

- Due to the timing of various events throughout the foreclosure and disposition process, it is crucial to ensure the loan that is being analyzed is a “completed” disposition when looking at loss or severity statistics. To help ensure that we are reasonably close to the final loss/gain for a given property, expenses and proceeds are surprised in our data release until 90-days after the recorded property disposition date.
- In this July 2015 data release, “complete” dispositions are identified as a Last Status of F or S (as defined by the R code shared here) and a non-null Disposition Date.
- In our R program, you will see:

```
LAST_STAT %chin% c("F", "S") & DISP_DT!=""
```

# Computing loss statistics

<b>Delinquent Interest Expense</b>	Default UPB * ((Current Interest Rate/100 – 0.0035)/12) * months from LPI to disposition
▶ <b>Total Costs</b>	Foreclosure Costs + Property Preservation and Repair Costs + Asset Recovery Costs + Miscellaneous Holding Expenses and Credits + Associated Taxes for Holding Property
▶ <b>Total Proceeds</b>	Net Sales Proceeds + Credit Enhancement Proceeds + Repurchase Make Whole Proceeds + Other Foreclosure Proceeds
<b>Total Net Loss</b>	Defaulted UPB + Accrued Interest + Total Costs – Total Proceeds
▶ <b>Severity</b>	Total Net Loss / Defaulted UPB

*You should now have everything you need to produce the following table!*



# Loss/Severity statistical summary table



	Origination Year															Total
	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	
Default UPB (\$M)	\$155	\$1,262	\$2,945	\$3,388	\$6,410	\$4,883	\$10,694	\$12,264	\$16,227	\$9,580	\$1,811	\$347	\$109	\$38	\$8	\$70,121
Default Rate (%)	1.0%	0.9%	0.8%	0.9%	1.3%	2.4%	5.1%	7.1%	7.4%	3.7%	0.4%	0.1%	0.1%	0.0%	0.0%	1.7%
<b>EXPENSES:</b>																
Delinquent Interest	12%	12%	11%	10%	10%	10%	9%	10%	10%	9%	7%	6%	6%	4%	3%	10%
Total Liquidation Exp.	10%	11%	11%	11%	11%	10%	8%	7%	7%	7%	7%	7%	7%	6%	4%	8%
Foreclosure	4%	5%	4%	4%	4%	4%	3%	3%	2%	2%	2%	3%	2%	2%	2%	3%
Property Preservation	3%	3%	3%	3%	3%	2%	2%	2%	2%	1%	2%	2%	2%	2%	1%	2%
Asset Recovery	1%	1%	1%	1%	1%	1%	1%	1%	1%	1%	1%	1%	1%	1%	1%	1%
Misc. Holding Expenses/Credits	1%	1%	1%	1%	1%	1%	1%	1%	1%	1%	0%	0%	0%	0%	0%	1%
Associated Taxes	2%	2%	2%	3%	3%	3%	2%	2%	2%	2%	2%	2%	2%	1%	1%	2%
Total Costs	122%	123%	121%	122%	121%	120%	117%	118%	117%	117%	114%	114%	112%	110%	107%	118%
<b>PROCEEDS:</b>																
Net Sales Proceeds	81%	77%	74%	72%	76%	72%	66%	61%	60%	63%	76%	81%	83%	82%	91%	66%
Credit Enhancement	17%	17%	14%	12%	9%	8%	6%	6%	8%	10%	5%	6%	8%	9%	10%	8%
Repurchase/Make Whole	6%	8%	5%	4%	2%	2%	3%	5%	8%	12%	8%	6%	4%	1%	0%	6%
Other	4%	4%	5%	4%	3%	2%	1%	1%	1%	1%	1%	1%	1%	1%	1%	2%
Total Proceeds	109%	107%	98%	93%	89%	84%	75%	72%	78%	86%	89%	93%	96%	93%	101%	82%
Severity	12.9%	15.9%	23.3%	29.2%	31.4%	36.4%	42.0%	45.5%	39.9%	30.4%	24.6%	20.5%	16.6%	17.2%	5.3%	36.6%
Total Net Loss (\$M)	\$20	\$201	\$685	\$989	\$2,012	\$1,775	\$4,494	\$5,576	\$6,474	\$2,908	\$445	\$71	\$18	\$7	\$0	\$25,676

*Piece of cake! Right??*

# Introducing Tutorial 102

- In our next module, we will discuss:
  - How to use the research dataset to analyze credit risk from a default and loss perspective, across risk gradients and through time.
  - How to replicate the “comping” analysis included in our CAS deal marketing materials
  - How to use FHFA’s Home Price Index to mark a loan’s LTV to market

*We look forward to seeing you there. If there was anything in this webinar that you would like to discuss one-on-one, we would be happy to do so.*

## Contact Information

### **Sonja Beaubien**

*Director*

[sonja\\_beaubien@fanniemae.com](mailto:sonja_beaubien@fanniemae.com)

202.752.8290

### **Nick Sapirie**

*Director*

[nicholas\\_sapirie@fanniemae.com](mailto:nicholas_sapirie@fanniemae.com)

202.752.5151

### **Patty Koscinski**

*Director*

[patty\\_koscinski@fanniemae.com](mailto:patty_koscinski@fanniemae.com)

202.752.3661

### **Stephen Schwartz**

*Director*

[stephen\\_c\\_schwartz@fanniemae.com](mailto:stephen_c_schwartz@fanniemae.com)

202.752.2795

### **Nick Leonard**

Financial Economist

[nicholas\\_c\\_leonard@fanniemae.com](mailto:nicholas_c_leonard@fanniemae.com)

202.752.5579

### **Inquiries**

[credit\\_securities@fanniemae.com](mailto:credit_securities@fanniemae.com)

*If there was anything in this webinar that you would like to discuss one-on-one, please feel free to reach out to any of the above contacts.*