

# Credit Score Basics, Part 1: What's Behind Credit Scores?

October 2011

## OVERVIEW

Today, credit scores are often used synonymously as an absolute statement of consumer credit risk. Or, credit scores are so misunderstood that they are viewed as a de facto measurement of borrowing strength and mistakenly seen as including assets, income or net worth in the calculation. Consumers are regularly advised that the three digit credit score is important to their financial well-being. It is not uncommon to hear people boast of excellent credit scores or anxiously ask for help when their scores are low. However, a wide knowledge gap exists about credit scores. The gap is most significant among consumers, but also exists across the entire spectrum of credit score users.

To help clear-up the muddle, VantageScore Solutions, LLC has prepared three white papers containing some of the most basic information about credit scores. Our goal is to have this information become broadly known, and through that effort, help everyone become smarter about credit scores. This paper is the first in our three-part series and will illuminate the appropriate use of a credit score as an indicator of the likelihood of defaulting on a debt and how to interpret what a credit score value means.

Indeed, one common miscalculation is attempting to describe whether a particular credit score value is good or bad without the context of how to interpret the number. This can be confusing at best, and at worst, can create increased risk exposure.

To illustrate the point, it's not widely known that hundreds of credit score models are available to lenders today and that many of those models employ different scale ranges. Yet, understanding the scale range that a particular score falls within is required in order to properly interpret the risk level associated with the score. If a consumer has a score of 700, and that 700 falls on a scale of 500 to 1000, then the 700 might represent 'C' quality credit. But a 700 score on a range of 300 to 900 is more likely to represent 'B' quality credit. The mathematics of credit scoring always defines the score in the context of its range—the minimum and maximum possible values that can be achieved. When scores are quoted, it's critical to also quote the range in order to convey an accurate understanding.

The example above showcases just one area of broad misunderstanding about credit scores. The discussion that follows addresses the primary design elements within credit scores that relate to accurately interpreting a consumer credit score.

## SUMMARY HIGHLIGHTS

- The three digit credit score value is a numerical representation of the likelihood that a consumer within a specific population will become 90 days or more past due on a debt obligation in a two-year timeframe. This is a level of risk referred to as “propensity to default.”
- Credit scores assess the risk of each individual within the context of the total risk for the entire population being scored.
  - » Credit scores are not an absolute statement of risk for an individual consumer, rather they state a consumers’ risk in relation to other consumers.
- The level of risk associated with a specific score is based on the credit score model being used, the population being scored, specific lender products and timeframe at the point the scores are generated. Altering any of those factors will alter the risk level associated with a specific score value.

## THE RELATIONSHIP BETWEEN CREDIT SCORES AND CONSUMER CREDIT RISK

A credit score is a numeric interpretation of a consumer’s risk level relative to the risk of other consumers in the same population. In other words, it is the likelihood that the consumer will allow one debt to become 90 days or more delinquent, also known as “propensity to default” (PD). For example, if a score of 700 represents a PD of 5%, then the consumer has a 5% chance of becoming 90 days or more late on an account over a given time period, typically two years. The relationship between the score and PD is determined by apportioning the total risk for the entire population among each consumer in the population according to their credit management behavior. Consumers with good credit management behavior have a low propensity to default and a high credit score, and vice versa. It’s important to note that the actual PD value at a specific score can vary based on a number of factors:

- The population being evaluated
- The product and lender mix
- The type of decision being made (acquisition versus account management)
- The time period when the consumer’s score was calculated

### FOR EXAMPLE

Let’s assume that we have a population of just two consumers. One consumer has extremely strong credit management skills and is therefore high credit quality. The other consumer has very poor credit management skills and is very low credit quality. If the overall risk for the population is 50% then one half of the population (or one of the two consumers) is likely to default. Given their credit management profiles, the high credit quality consumer has a PD of less than one percent and consequently an extremely high credit score, say 990. (The full VantageScore® range is 501-990). The poor credit quality consumer has a PD of 99% which results in a credit score at the bottom of the range, say 510. Combining the individual PDs of both consumers results in the 50% default risk for the overall population. This same principle of “score to PD relationship” applies to most commercially available credit risk scores.

**THE  
RELATIONSHIP  
BETWEEN CREDIT  
SCORES AND  
CONSUMER  
CREDIT RISK**  
(Cont.)

In the real world, most lender portfolios are comprised of millions of people that exhibit a broad variety of behaviors, resulting in a spectrum of PDs ranging from very low to very high. The scoring industry represents this information using *performance charts* (Figure 1—shown below), also known as odds charts. The performance chart evaluates and summarizes a population of consumers based on their probability to default. Typically, performance charts are produced annually and reflect consumer default levels occurring in the prior two-year time period (e.g., 2008-2010 as of 2011).

For ease of interpretation and application, performance charts sort consumers in 20-point intervals from highest score to lowest score. The first interval includes the highest credit quality consumers. Therefore, based on the VantageScore scale of 501-990, consumers in the first interval are those whose scores fall between 971 and 990, inclusive. The second interval includes consumers with scores between 951 and 970, and so on.

**FIGURE 1  
PERFORMANCE CHART**

INTERVAL	SCORE RANGE	PROBABILITY OF DEFAULT	
		INTERVAL	CUMULATIVE
1	971-990	0.15%	0.15%
2	951-970	0.13%	0.14%
3	931-950	0.14%	0.14%
4	911-930	0.19%	0.16%
5	891-910	0.19%	0.17%
6	871-890	0.24%	0.19%
7	851-870	0.38%	0.24%
8	831-850	0.49%	0.28%
9	811-830	0.80%	0.35%
...	...	...	...
21	571-590	26.25%	4.47%
22	551-570	31.16%	4.79%
23	531-550	36.84%	5.02%
24	501-530	46.33%	5.37%

For interval 1, consumers with scores between 971 and 990, the probability of default is 0.15%. For interval 24, consumers with scores between 501 and 530, the probability of default is 46.33%.

The cumulative probability of default reflects the total risk level as you move deeper into the population. For example, the total risk for consumers with scores between 571 and 990 is 4.47%. Said another way, 4.47% of the population with credit scores between 571 and 990 are likely to default.

VantageScore performance charts are released in Q3 each year and reflect probability of default levels for a two-year window, June to June.

## HOW DO SCORE DESIGNERS PICK THE SCORE RANGE?

The underlying mathematics of credit scoring usually allows the designers to use the range they desire for the score. Typically, a range is selected to have enough span so that consumers are distributed across the full range in a manner allowing lenders to effectively slice the population into manageable sizes, reflecting the varying risk levels within the population. For example, a score range of 10 points (1 to 10), when applied to the U.S. population of 200 million consumers, results in tens of millions of consumers receiving the same score. Distributing the U.S. population across a range of just 10 points means that, on average, 20 million consumers receive the same score (200 million divided by 10), providing a lender with no meaningful risk differentiation and no real capability to segment the population. Conversely, a score range of one million points (1 to 1,000,000) results in just a few hundred consumers receiving a particular score, which can be equally cumbersome to manage as a segmentation tool. Generally, model designers have used a range that spans from 500 to 900 points in order to sufficiently disburse the consumer population. The VantageScore range spans 489 points since it is 501-990. Other models may use a similar breadth of range but could have a different starting point such as 100 or 300.

## DOES THE “PROPENSITY TO DEFAULT” FOR A GIVEN SCORE IN A PERFORMANCE CHART ALWAYS REMAIN THE SAME “PROPENSITY TO DEFAULT” FOR ALL FUTURE TIMEFRAMES?

The score and “propensity to default” (PD) values seen on performance charts are calculated on a specific population of consumers based on their behavior using a specific suite of products from a specific suite of lenders during a recent and specific two-year timeframe. Generally, performance charts need to be updated annually to account for changed consumer behavior or new product offerings from lenders. For the relationship between the score and the PD for a given performance chart to remain the same in future two-year timeframes, the same population of consumers must exhibit essentially the same behavior, and lenders must offer products that remain essentially the same as those observed during the two-year timeframe used to build the initial performance chart. If the population remains the same but their behavior changes, as seen during the 2007-2009 recession, then the PD is likely to be different. Similarly, if the future population substantively changes in demographic composition, then the PD is also very likely to be different at each score value.

The graph in Figure 2 below demonstrates this point. The graph depicts PDs for six timeframes for consumers who have either real estate or bankcard loans. All consumers fall into the same credit score tier (731-750). The PD for each population is calculated as the percentage of these consumers who are in default on their loan. Varying levels of default propensity are observed.

*Why do these populations have different PDs within the same timeframe?*

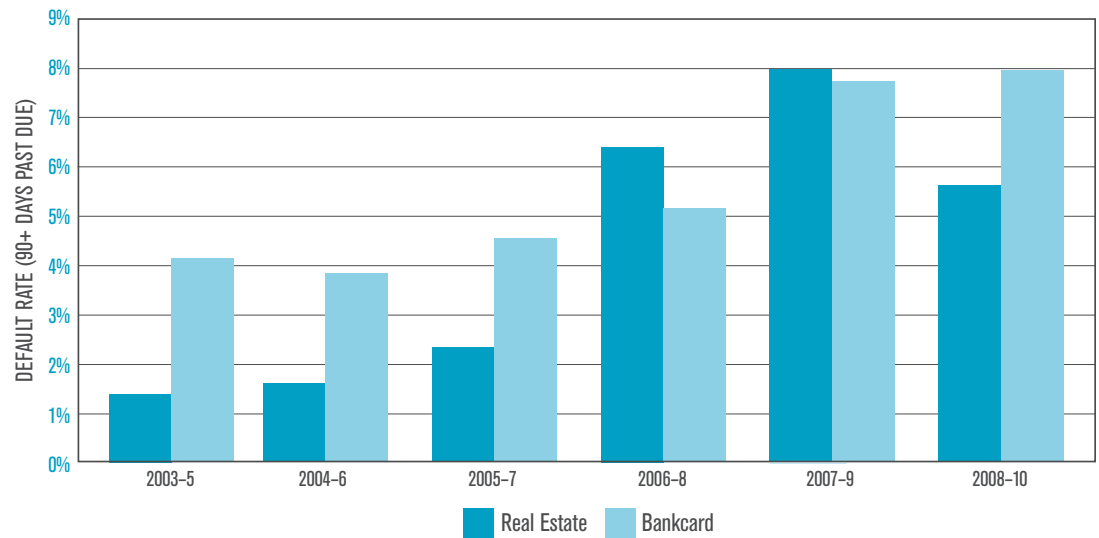
Focusing on the 2003-2005 timeframe, consumers viewed their real estate loans as the first priority for payment. Therefore the overall default level for real estate loans was extremely low at that time, just 1.25% for **all consumers with a mortgage**. As a result of distributing that total risk across consumers based on their credit management behavior, the portion of U.S. consumers scoring 731 to 750 have a default level of 1.46%. Bankcards were viewed as having lower priority in the payment hierarchy by consumers, perhaps due to the consequences of defaulting on a bankcard being far less severe than defaulting on a mortgage. As a result, the default level for bankcard products **across the entire population** in this same timeframe was much higher, 5.57%, and specifically was 4.14% for those consumers scoring 731 to 750.

DOES THE  
 “PROPENSITY TO  
 DEFAULT” FOR A  
 GIVEN SCORE IN  
 A PERFORMANCE  
 CHART ALWAYS  
 REMAIN THE SAME  
 “PROPENSITY  
 TO DEFAULT”  
 FOR ALL FUTURE

This result demonstrates that default levels can vary by loan type based on consumer hierarchy of payments. Propensity to default values can also vary for many other reasons, such as the structure of the loan, lender strategy, regional variances or macroeconomic influences.

These two industry populations were again scored in the 2007-2009 timeframe. Even though consumers have the same score range (731-750), the PDs have dramatically increased for both populations, indicative of a widespread behavioral change as a result of the recession. In fact, default levels for consumers in this score range with real estate accounts increased by a factor of five to 7.97%.

FIGURE 2  
 CONSUMERS WITH SCORE: 731-750



Key insights to remember:

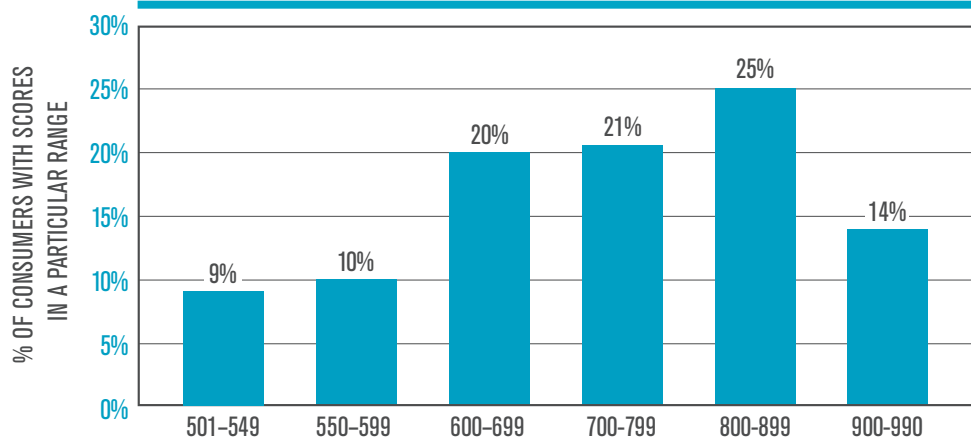
- Understand the range that the credit score is based upon.
- When considering whether the PD on a performance chart is a reasonable representation of the future consumer risk, determine whether the underlying population, product and lender mix is essentially the same with reasonably similar behaviors as the population used to generate the performance chart.

ALTERNATIVE REPRESENTATIONS OF CREDIT SCORES

Two additional representations of credit scores began appearing in the mainstream dialogue in 2011. New risk-based pricing disclosure rules and changes to adverse action requirements mean lenders will inform consumers of their credit score when credit files and/or scores are used in a pricing decision for credit. Lenders will use one of the two following formats to explain how a consumer’s credit score compares to the rest of the U.S. population from whichever credit reporting company supplied the credit score.

1. The bar graph below (Figure 3) represents the first option and presents the distribution of the U.S. population by credit score in six primary tiers, identified on the horizontal axis. A consumer with a score of 800, for example, can see that 25% of the scored U.S. population falls into the same tier, 14% of the U.S. population has a higher credit score than their tier (scoring 900-990) and 60% of the population scores in a lower tier (501 – 799).
2. The second option is seen in Figure 4, where the same information is represented in a more granular form. A consumer with a score of 900 ranks higher than 85% of U.S. consumers.

FIGURE 3



ALTERNATIVE REPRESENTATIONS OF CREDIT SCORES

FIGURE 4

LOW SCORE	HIGH SCORE	"Your credit score ranks higher than [X] percent of U.S. consumers." CUMULATIVE %
...	...	...
562	566	12%
567	569	13%
570	574	14%
575	579	15%
580	585	16%
586	590	17%
...	...	...
878	881	79%
882	884	80%
885	888	81%
889	891	82%
892	895	83%
896	898	84%
899	902	85%
903	906	86%
907	910	87%
911	914	88%
915	918	89%
...	...	...

Consumers are increasingly likely to see this information when they are approved or denied credit. Consumers in the high credit tiers or who rank higher than the majority of the rest of the U.S. population will likely be able to access more credit at more favorable pricing.

CONCLUSION

Credit scoring processes and their design have unnecessarily remained a black box for many years. Through the simple conceptual discussions in this paper, VantageScore Solutions LLC strives to provide greater transparency and comprehension to those using credit scores and to those engaging in risk and credit scoring dialogue.