Data and Bias, Part 1:

Question
[*the input your machine learning program treats as an*]
Authority

by Kevin Loney

Copyright Kevin Loney 2019. Posted to http://kevinloney.com
All rights reserved. May not be reprinted or reposted to any domain but kevinloney.com.
The views and opinions expressed in this article are my own and do not necessarily represent official policy, practices, or position of any company or organization with which I have ever been affiliated.

(Note: For consistency and length, the focus in this article is on the US economy.)

To start, a quick coding problem:

Assume that your company hardcodes its business rules into software code, and the code base is being enhanced to let the software learn and improve its calculations as it runs. In testing a new module that identifies optimal customers by recursively matching pairs of customers against each other, you are using the Transunion credit score as the primary input. For the first pair being evaluated, the input credit scores for the individuals are 649 and 613, respectively.

Which of those two should be chosen?

———

Does your machine learning algorithm generate unbiased outcomes, or do programs encode their creators' biases into business processes?

Is it even possible for code to be biased?

One common response to that last question: "*Code doesn't have 'bias' — that's a human trait. The algorithm is just doing math, and math is just numbers, and numbers aren't biased. It's just math —*"

But algorithms contain the rules and scores that are used to measure and direct society. They use inputs and internal mechanisms to generate values which influence the opportunities open to people and set the costs for them — and commonly used algorithms have already been publicly found to be biased, with corrective actions taken. It's no longer a question of whether a system is biased, it's a matter of how accurately we account for the bias when using its output.

This series of articles will focus on bias because the rapid movement to a data-based environment <u>requires</u> developers and architects and users to understand where that data comes from. As systems are built on systems, ingesting data from upstream sources that were built based on varied sets of assumptions and goals, it is easy to lose track of what the data actually means. As users, we cannot take these meanings for granted when they directly impact our lives, our jobs, our businesses, and our economy. That impact is documented in reports and working papers published separately over the last decade by the Federal Reserve, FINRA, and the credit reporting agencies, and it may be influencing your business processes. As in the coding problem this article started with, we'll start with something everybody already knows intimately: a credit score.

What's the score?

Questions to keep in mind for this next discussion:
What's your credit score?
What would your life be like if your credit score was 120 points lower?

For something as influential in our personal financial lives as the credit score, one would expect every person would be fully aware of how the score is calculated and how those calculations have changed over the years. Who tracks how those calculations are evaluated by the credit reporting agencies themselves? Which credit score impacted your most recent loan application process (since there are multiple scores in use)? And how do companies counter any detected issues with the scores? The credit score, after all, is not only used when evaluating applications for credit; potential renters, utility companies, phone companies, and insurance companies use credit scores to determine your eligibility for products they offer. Potential employers may check your score as a pre-employment validation step (if permitted by state and federal law, if it's for a managerial position, or if it's for a financial institution). Although its intended use may be valid, these additional uses of the credit score are beyond the scope of the original credit scoring effort, and taken together they magnify its effect.

The most commonly used credit score, FICO, was introduced in 1989. The intent of the FICO score was to quantify the likelihood a borrower would be 90 days delinquent on a loan payment at some point in the next two years. That's a very specific use case and metric. The effectiveness of the FICO scoring method for its use case led to other industries using that score for other purposes, looking to build on its foundation and embed it into other processes while adding attributes and algorithms as needed.

In generating the lending risk score for an individual, a lender would be interested in a number of criteria, and would give consideration to past behavior in the credit market, assuming prior behavior is indicative of future behavior. For the FICO score, there are five broad categories of input data. In order of importance, they are (roughly)

35% your payment history
30% your debt burden (including total debt and amounts across accounts)
15% length of credit history (the age of the accounts)
10% types of credit used (loans, cards, mortgage, etc)
10% recent requests for credit

(Note: *How those values are populated, weighed and evaluated differs by credit agency and score. See* https://en.wikipedia.org/wiki/Credit_score_in_the_United_States. *Negative information such as late payments are required by law to stay on the report for 7 years; positive information can stay on the report indefinitely. If you dispute a claim, both the original claim and the dispute resolution may be kept on the credit history depending on the dispute status. See* https://www.experian.com/blogs/ask-experian/how-disputing-information-on-your-credit-report-affects-your-credit/).

If a person has made a number of recent requests for credit (for example, by applying for multiple apartment rentals that trigger new credit checks) and has a low number of types of credit (no current mortgage, and has reduced the number of loans by consolidating them onto one relatively new card), then that person's credit score may reflect that negatively. A lower credit score makes a person appear riskier to lenders, who in turn may offer a higher interest rate on credit products or a higher rental rate than others may be offered. If a customer challenges their first quoted rate, lenders may point to the customer's credit score as part of the justification.

The use of the score as a potential driver in the economic system (determining the eligibility and rate going forward) should lead us to look at the scoring system as a whole. As it assigns outcomes to each individual based on past performance and current holdings, does the output of the credit scoring process become self-justifying — influencing, perpetuating and increasing the differences between the individuals it is measuring? Does it reward the behaviors it wants to encourage? Does the system of measurement contribute to the differences between those with the lowest and highest scores, the subprime and superprime borrowers?

(*In this article, 'subprime' borrowers have a credit score below 640. 'Prime' extends from 640 to 740, and above 740 is 'superprime.' There is not a completely consistent definition for these terms across the credit rating agencies*.)

Borrowers with lower credit scores may be offered credit with higher rates (since they are believed to be at a higher risk of defaulting on the payments) to cover the lender's expected risk. Those borrowers are therefore paying more than average for each dollar they borrow. On the other end of the financial spectrum, those with high credit scores will be more likely to be approved for new credit offerings at low rates, making it less burdensome for them to repay the loan. A high credit score supports future positive behavior, since it leads to lower costs per dollar borrowed. In terms of rewarding behavior, neither borrower has direct visibility to the financial terms for the other, limiting the system's ability to effectively inspire behavior changes in the participants. (For that matter, the data and ratings used by the agencies were not available to the citizens they were rating until the passage of the FACT Act in 2003.)

In comparing the repayment behavior of people with different credit scores, studies should account for the fact that those individuals are not operating in the same economic environment. The credit marketplace does not operate blindly on behalf of its shareholders — it offers different rates to different customers, changing the per-dollar borrowing costs for each person. This, in turn, may impact the ease and speed with which low- and high-scoring individuals can raise their scores in any of the categories listed above.

Is the credit score inherited?

To further understand the credit score data we are ingesting into our hypothetical code base, let's examine the role family background and environment plays in the credit score. When first pursuing this question, I was focused on quantifying the role played by private equity — either the ability to borrow money from relatives at low interest rates (but not generating a repayment history), or the drag of being the sole significant earner in a family with long-term financial liabilities. Recent working papers from the Federal Reserve went well beyond that question's scope (benefitting from access to a broad range of data sets) to include parents' college degree level, undergraduate college financing, school choices, and many other societal factors.

Note: In the next several pages I'll be referencing a working document, and I am including its disclaimer here. Any citation of this article would need to go back to the original citation: *Goodman, Sarena, Alice Henriques, and Alvaro Mezza (2017). "Where Credit is Due: The Relationship between Family Background and Credit Health," Finance and Economics Discussion Series 2017-032. Washington: Board of Governors of the Federal Reserve System, https://doi.org/10.17016/FEDS.2017.032.*

NOTE: Staff working papers in the Finance and Economics Discussion Series (FEDS) are preliminary materials circulated to stimulate discussion and critical comment. The analysis and conclusions set forth are those of the authors and do not indicate concurrence by other members of the research staff or the Board of Governors. References in publications to the Finance and Economics Discussion Series (other than acknowledgement) should be cleared with the author(s) to protect the tentative character of these papers.

People may move up or down from their original financial and social circumstances. A person receiving an inheritance may make poor investment choices; a person with only herself to rely on may build a successful business from the ground up. People may have supportive parents, or may have parents who damage their credit by opening accounts in their children's names and then defaulting on them. Many different life options were considered as part of the study documented in the Fed's working session papers. The papers, titled "*Where Credit is Due: The Relationship between Family Background and Credit Health*," were issued in 2017 as part of a Finance and Economics Discussion Series (FEDS). (*See* https://www.federalreserve.gov/econres/feds/files/2017032pap.pdf *and prior citation*).

In its abstract, the authors address the challenge for lenders: one of the strongest factors driving credit health of a 30-year-old is that person's family background and early financial support environment, but that data *cannot* be used as part of credit scoring:

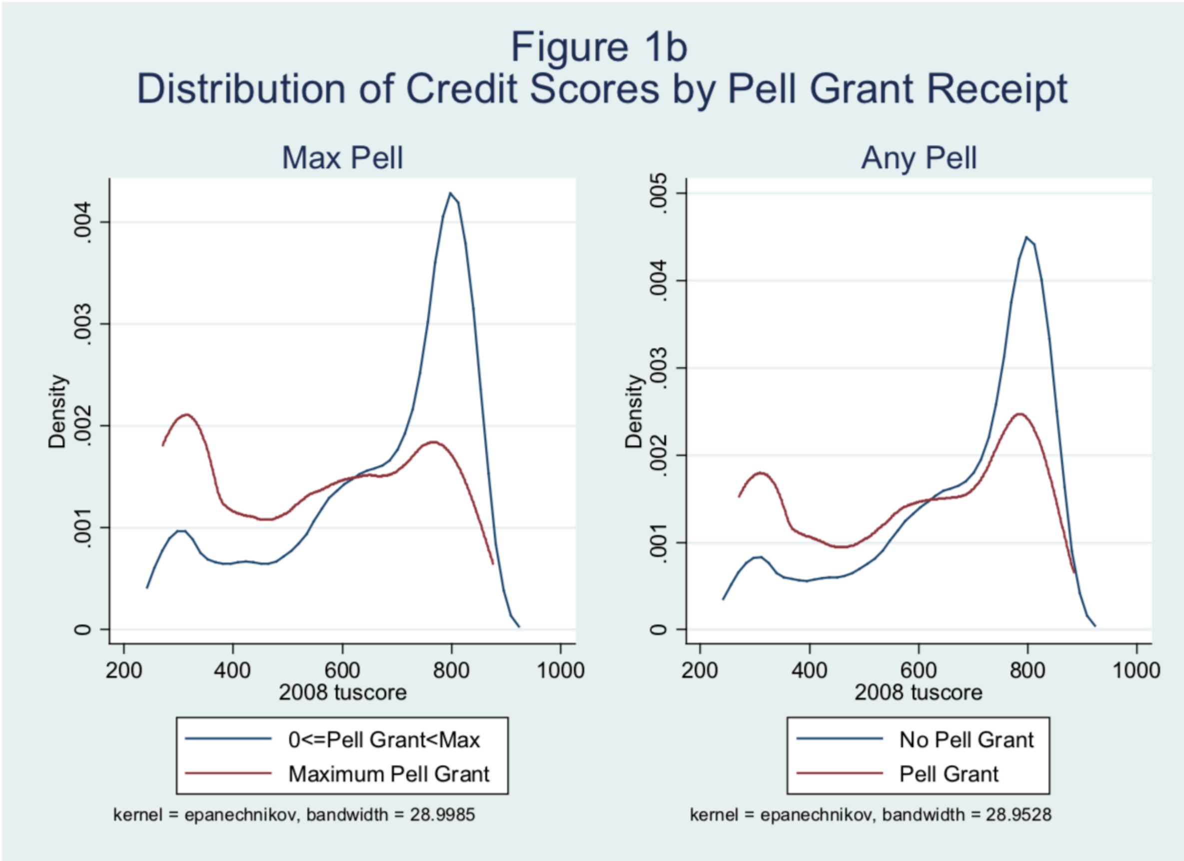
"*…we document that, even though it is not, and* cannot *be, used by credit agencies in assigning risk, family background is a strong predictor of early-career credit health.*"

What difference does that make? [Bold added for emphasis:]

"**[*At age 30:*] *In our simplest specification, we estimate that credit scores are about 100 points lower for individuals from disadvantaged backgrounds, and that such individuals are about 20 percentage points more likely to be subprime*.**"

A 100 point drop in credit score is significant. As people enter young adulthood, the papers conclude a financially disadvantaged background in youth will lower their credit score by 20%, on average, unless they manage to overcome the constraints inherent in their financial environment. The Fed discussion papers include the following graph in the appendix: the distribution of credit scores (in 2008, at age 30) based on whether the individuals in their data sets had taken any Pell Grant money during their college years. Pell Grants, which are available only for undergraduate work, provide up to $6,195 per academic year for a student whose family income level is so low that their expected family contribution for college is $0. Qualification for the maximum Pell grant amount is indicative of the family poverty level, the lack of student savings available, and the student's pursuit of a college degree.

In the graph on the left, the red line shows the distribution of credit scores in 2008 for 30-year-olds who qualified for the maximum Pell grants when they were college undergraduates. Almost a decade after starting college, more than half of that population — a group that was admitted to and attended college — had a credit score below 600, below the subprime cutoff point. In the graph on the right, a separate pool of students is shown: If *all* Pell students were evaluated (the red line), including those who qualified for less than the maximum grant, the credit score distribution shifts toward prime credit score range but still significantly trails the credit score distribution of those students in the study whose wealth or income disqualified them from receiving *any* Pell grants (the blue line in the graph on the right).



(*Note that all of these populations were tracked during an 8 year period ending in 2008 and may have been impacted by the economic downturns in 2001 and 2008; the ability to take positive action in that economic climate may have further influenced outcomes for some of the participants. The research papers state that the results are very similar for the same individuals 6 years later, in 2014, but do not provide those charts.)*

There are many variables behind these graphs, since Pell grant qualification reflects a complex environment that is not identical from one family to the next. But there are larger, consistent themes as described in the report. For example, a student who receives a Pell grant is more likely to take student loans, which are directly associated with lower credit health. The authors looked at a wide variety of variables, including SAT scores (which themselves are the subject of well-documented bias studies), school quality (students of low socioeconomic status are more likely to attend less selective schools), and graduate level coursework.

(*Note: The study did not specifically examine any bank or lender's credit products to determine how or if their practices or products impacted the economic gaps noted. It made no judgment about the fitness of any such products or practices. The study's focus was on the score as an outcome, and its change over time for specific populations and individuals, not on the downstream uses over which the rating agencies had no direct control*.)

While student loans help people attend school, raise their education levels and qualify for job opportunities, those gains are often offset by their repayment costs being incurred at a time when those people are financially fragile. Those students who rely on loans may need to discontinue attendance — and trigger repayment requirements — due to small but unaffordable increases in tuition, housing, book costs, and other fees. The report concludes:

"… *we find that taking federal student loans…is consistently negatively associated with credit health, even after accounting for all of our other controls*."

In other working papers, analysts determine it is not the size of the student loan balance that is the determining factor in delinquencies; it is the early credit histories of the young borrowers that are highly predictive.

For full details on the student loan repayment analysis in a separate working paper, see https://www.federalreserve.gov/econresdata/feds/2015/files/2015098pap.pdf
Mezza, Alvaro A., and Kamila Sommer (2015). "A Trillion Dollar Question: What Predicts Student Loan Delinquencies?," Finance and Economics Discussion Series 2015-098. Washington: Board of Governors of the Federal Reserve System, http://dx.doi.org/10.17016/FEDS.2015.098.
NOTE: Staff working papers in the Finance and Economics Discussion Series (FEDS) are preliminary materials circulated to stimulate discussion and critical comment. The analysis and conclusions set forth are those of the authors and do not indicate concurrence by other members of the research staff or the Board of Governors. References in publications to the Finance and Economics Discussion Series (other than acknowledgement) should be cleared with the author(s) to protect the tentative character of these papers.

Regarding student loan repayment: For a total of $26,000 in federal college loans, at a standard 6.8% rate and spread over 20 years, the monthly payment (due starting 6 months after graduation or non-enrollment) would be $198.47. Over 20 years, that makes the total repaid for the loan roughly $47,632. Over 45% of that total repayment is interest.

What about high achievers? Or grad students? The analysis concludes:

"*Further, a gap remains even after accounting for achievement, postsecondary schooling, and key elements of early credit histories…*"

Those who choose to repeat the errors of history —

The working papers observe: [bold added for emphasis]

***"The resilience of this relationship suggests that the credit market could be amplifying the transmission of economic well-being across generations*."**

After studying the credit score, the paper's focus shifts to the credit market — how the score is used by processes and algorithms. Does a lagging indicator based on a borrower's repayment history in some manner play a part in societal differentiation and economic disparity? This analysis of credit usage and its impact on populations over generations parallels the cross-generational analysis of income and opportunity in the US by Dr. Raj Chetty. *See* https://www.theatlantic.com/magazine/archive/2019/08/raj-chettys-american-dream/592804/ .

The Fed working papers end with consideration of the impact of outright discrimination and how the recursive use of prior scores in credit scoring mechanisms might perpetuate the outcomes of discriminatory practices: [*numbers added below for reference*]

"Finally, if discriminatory lending practices restrict certain groups' ability to access credit, these groups may have a more difficult time accumulating a strong credit history, which could then affect their scores.

Future research should explore which of the above mechanisms underlie the early gaps in credit health we detect and the effectiveness of policies in ameliorating them. In particular, a key question is whether the differences in credit scores that we document by socioeconomic group [1] *stem solely from the underlying default risk of different household types or* [2] *are partially an unintended artifact of how credit scores are constructed."*

There is a third option: not just the default risk of the household types or the method of constructing the credit score, but [3] how the credit score is used in the execution of business processes. Given the score as input, what is your machine learning code assuming about it? What is your code's downstream impact? As noted earlier, credit scores drive decisions in the insurance, utilities, rental, and phone industries in addition to the lending industry.

It should be noted that the FEDS study was conducted as a retrospective observation by a third party. Companies engaged in lending rely on factors beyond the credit agency credit scores to determine the customer pools to engage and the terms to offer them, and they actively monitor and manage those factors and review them with regulators. The FEDS working papers consider the most widely known factor and extrapolate from there for the subset of individuals they tracked across the study's time period.

Protecting populations

The 2017 report is not the first detailed study of the US credit scoring systems. The 2003 FACT Act, the same legislation that mandated free annual credit reports for consumers, included a provision that required the credit reporting agencies to check themselves for impact to protected classes. (*See* https://uscode.house.gov/view.xhtml?req=granuleid%3AUSC-prelim-title15-chapter41-subchapter3&edition=prelim)

Section 215-a(3) of the FACT Act calls out the need to investigate what parameters lead to an impacted credit score based on an array of personal characteristics associated with protected classes of people:
"(3) the extent to which, if any, the use of credit scoring models, credit scores, and credit-based insurance scores impact on the availability and affordability of credit and insurance to the extent information is currently available or is available through proxies, by geography, income, ethnicity, race, color, religion, national origin, age, sex, marital status, and creed, including the extent to which the consideration or lack of consideration of certain factors by credit scoring systems could result in negative or differential treatment of protected classes under the Equal Credit Opportunity Act, and the extent to which, if any, the use of underwriting systems relying on these models could achieve comparable results through the use of factors with less negative impact;"

The results of the study were provided four years later, in a report to Congress from the credit reporting agencies. (*For the full report, see* https://www.federalreserve.gov/boarddocs/rptcongress/creditscore/creditscore.pdf)

The 2007 report from the agencies, "*Report to the Congress on Credit Scoring and Its Effects on the Availability and Affordability of Credit*," details the analyses that checked for correlations between individual characteristics and the composite credit score. While most individual parameters were deteremined to not be directly correlated with the credit score, one parameter was correlated in a way that did not directly reflect the individual's credit worthiness: age. People under the age of 30 (with short credit histories) had a score that was lower than it should have been given their later repayment behavior, and those over the age of 62 had a score that overly benefitted from the length of their credit histories.

Although not required by the FACT Act-mandated analysis as an area to investigate, there was another group that was called out by the agencies as being negatively impacted the same way the under-30 population was: recent immigrants. The report states:

"*Recent immigrants appear to have somewhat lower scores in the FRB base model than would be appropriate given their performance … it is attributable to the tendency of recent immigrants to have credit profiles similar to those of young people in terms of the lengths of their credit histories, as reflected in their U.S. credit records*."

Lenders may request additional materials (utility bills, evidence from foreign banks, etc.) that justify a credit request as part of the application approval process. As of this writing, at least one 2020 presidential candidate is pushing to make expedited submissions of additional information part of *initial* credit applications, which may enable those populations with short credit histories to be more competitive during time-critical credit applications (such as bidding on houses) that may have dependencies on quick approvals.

Meanwhile, as the credit reporting industry learns more about its scores it enhances its algorithms. In August 2018, over 60% of the US population had its credit scores increased when core parts of the calculations were changed (S*ee* https://www.usatoday.com/story/money/personalfinance/2018/08/27/credit-score-may-have-gone-up-why-calculation-changed/1111647002/).

With so much riding on getting those numbers right, the credit reporting agencies will continue to evaluate the criteria they use and the weights they assign to each. They will continue to monitor these revisions as people acquire, use, and pay off credit. And the downstream systems that rely on that data will have to react to each new version of the different scores.

As this is happening, society is changing along with it. Every business process that considers the credit score of an individual may be using it properly, or it may be turning that score into an unacknowledged tool to change society. And that tool, in the words of the Fed working papers, may be "*amplifying the transmission of economic well-being across generations*", such as through altering access to opportunities. In so doing, that tool may broaden opportunity or may more deeply entrench a societal economic gap whose primary mechanism for upward movement, higher education, comes with a financial burden the FEDS papers found to be directly associated with broadening that gap.

To revisit the starting question:
In the coding example we started with, how old were the two individuals?

Rethinking the question

One of the assumptions of a functioning economic system is that all valid participants have access to the resources that can legally be purchased. If a participant's current status is based on past status, then that same assumption regarding access to resources should have been valid in the past as well; otherwise the system will perpetuate and memorialize past errors until they are actively counterbalanced. In the recent past, depending on where people lived or what type of mortgage company they were dealing with, companies may have chosen to withhold or limit their mortgage offerings, negatively impacting distinct populations. As cited by the National Consumer Law Center (*see* https://www.nclc.org/images/pdf/credit_discrimination/Past_Imperfect050616.pdf), for decades the Federal Housing Authority (FHA) "*refused to guarantee home loans made in African American communities, thus depriving them of the ability to accumulate wealth through homeownership*."

This practice was called *redlining*. From the time the FHA was put in place in 1934, certain minority zones were marked off by mortgage lenders with no support for housing ownership, blocking equity accrual by the residents (*see* https://en.wikipedia.org/wiki/Redlining). Transparency and remedies related to this process were not passed until the 1970s, several generations later. The effect (and underlying behavior) is still happening in pockets this decade, with a 2015 judgment passed against a lender for systematic exclusion of majority-minority communities in its lending. (*See, for example,* https://ag.ny.gov/press-release/ag-schneiderman-secures-agreement-evans-bank-ending-discriminatory-mortgage-redlining). There are many excellent articles and books providing research and guidance on the workings of the mortgage and loan guarantee systems for minorities in the US, which is beyond the scope here other than to note its impact to the unevenness of the historical economic playing field for individuals whose scores factor their histories into their algorithms.

(One quick note on the social impact of redlining: the FHA is funded by the government, which in turn is partly funded through taxation of its citizens. As the author Austin Channing Brown points out in her book I'm Still Here: Black Dignity in a World made for Whiteness, minorities were paying taxes that partly funded FHA programs that through Jim Crow laws were systematically denied them while those same programs benefitted everyone else. Each injury was incurred twice, with subsequent impacts to the wealth gap and credit health gaps.)

In the US, real estate ownership has been historically segregated in many areas (*see* https://en.wikipedia.org/wiki/Housing_segregation_in_the_United_States *for an overview*), and credit-impacting foreclosures are disproportionately prevalent in minority-intensive areas. As illustrated in Matthew Desmond's Pulitzer Prize-winning book Evicted, the rental units available in minority-intensive areas are priced nearly as high as those in better (and less accessible) areas. Minorities who were denied the right to purchase property in non-integrated communities outside their neighborhoods were then blocked (by redlining) from purchasing homes in their own neighborhoods, depriving them of the opportunity to invest in an appreciating asset while their rental rates rose decade after decade. Checking the credit score components again, the only short-term option to increase their credit scores would be to take on more debt, which leads to long-term credit problems.

Other currently protected classes have experienced similar obstacles first-hand. As recently as the early 1970s, single, widowed, or divorced women in the US could not get a credit card without bringing along a man to cosign the loan agreement. Passed in 1974, the Equal Credit Opportunity Act (see https://www.law.cornell.edu/uscode/text/15/1691) outlawed discrimination against any credit transaction applicant "*on the basis of race, color, religion, national origin, sex or marital status, or age (provided the applicant has the capacity to* *contract);* [*or*] *because all or part of the applicant's income derives from any public assistance program*." The act included the 'sex or marital status' clause due to the leadership of Lindy Boggs (*see* https://abcnews.go.com/Politics/congresswoman-ambassador-lindy-boggs-dies-97/story?id=19792180).

Is there a difference in credit availability based on sex? A 2012 paper published by the FINRA Investor Education Foundation (http://www.usfinancialcapability.org/downloads/In%20Our%20Best%20Interest%20Mottola%202013.pdf) found that for the people in their survey (respondents in a 2009 study), women paid more than men for credit: "*Even after controlling for a host of variables, women pay almost half a percentage point more in credit card interest rates than men*."

For the full paper from the FINRA Investor Education Foundation, see
Mottola, Gary R. (2013) "In Our Best Interest: Women, Financial Literacy, and Credit Card Behavior," Numeracy: Vol. 6: Iss. 2, Article 4.
DOI: http://dx.doi.org/10.5038/1936-4660.6.2.4
Available at: http://scholarcommons.usf.edu/numeracy/vol6/iss2/art4

The FINRA paper illustrates how issuers alter rates based on risk levels: "*That said, the overwhelming driver of credit card interest rates is clearly credit score. Respondents with a credit score of 620 or less paid credit card interest rates that were 3.96 percentage points higher than respondents with credit scores above 620*."

While these topics can be discussed from a detached perspective (viewing the credit system as a system with inputs, outputs, and controls that have changed over time), that is not how people experience it. If you as a developer are going to use a system's prior output as an input into its current processing, then that prior output must be understood and any control required must be applied as the system runs. For a significant portion of the US population, the system's prior operation has not been consistently positive, and has been determined in court to have been used in prejudicial ways against them — including in this decade — with legislative corrective controls attempted at all levels of government. If those impacted individuals were to approach this topic, the question would not be what I started with — "Is the credit score inherited?" — but rather, "Of course wealthy people start with high scores and poor people have low scores before they even start, and that impacts the opportunities available to each. How can we fix that access and the cost for opportunity?"

Now about your code (a recursive sequel) —

Given that -

Assume that your company hardcodes its business rules into software code, and the code base is being enhanced to let the software learn and improve its calculations as it runs. In testing a new module that identifies optimal customers by recursively matching pairs of customers against each other, you are using the Transunion credit score as the primary input. For the first pair being evaluated, the input credit scores for the individuals are 649 and 613, respectively.

Which of those two should be chosen?

- Kevin Loney, August 2019.

Thanks to Emily Loney and Arlene Harrison for their critiques on early drafts of this article.

Copyright Kevin Loney 2019. Posted to http://kevinloney.com
All rights reserved. May not be reprinted or reposted to any domain but kevinloney.com.
The views and opinions expressed in this article are my own and do not necessarily represent
official policy, practices or position of any company or organization
with which I have ever been affiliated.